2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

2nd edition

ENCYCLOPEDIA OF

PHILOSOPHY

ABBAGNANO – BYZANTINE PHILOSOPHY

1 volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT

Editor in Chief

MACMILLAN REFERENCE USA

An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE™

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON
GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

editorial board

editorial board

contents

editorial and production staff

executive vice president and publisher
Frank Menchaca

director, new product development
Hélène Potter

project editors
Jane A. Malonis
Carol A. Schwartz

contributing editors
Erin Bealmear, Deirdre S. Blanchfield, Steve Cusack, Angela Doolin, Susan Doty, Jason Everett, Alan Hedblad, Monica Hubbard, Lynn Koch, Melissa McDade, Bradley J. Morgan, Scot Peacock, Drew Silver, Ken Wachsberger

editorial technical support
Josh Kondek, Andrew Malonis, Mark Mikula, Mark Springer

manuscript editors
Robert A. Arlt, Dorothy Bauhoff, Sharon R. Gunton, William Kaufman, Eric Lagergren, Steven M. Long, Gina Misiroglu, Marie L. Thompson, Alan Thwaits, Amy Unterburger

proofreaders
Archie Hobson, John Krol, Amy Unterburger

bibliographic researcher
Michael Farmer

translators
Names of translators appear throughout the body of the *Encyclopedia*, at the end of each article that has been rendered into English.

indexer
Coughlin Indexing Services

product design
Kate Scheible

graphic art
Argosy Publishing

composition
Evi Seoud

manufacturing
Wendy Blurton

preface to the second edition

Nearly four decades ago, in 1967, Macmillan published its eight-volume *Encyclopedia of Philosophy*. With Paul Edwards as its exceptionally able editor in chief, the *Encyclopedia* became a highly respected, premier reference work consulted by countless professors and students as they pursued the examined life. Indeed, it would be safe to say that most if not all of the scholars who have contributed to the new Second Edition of the *Encyclopedia* leaned on the First Edition for philosophical insight during their formative years as young academicians. For them to be able to participate in reshaping a reference resource that figured importantly in their intellectual development has been a unique opportunity and a privilege.

When Macmillan invited me to serve as editor in chief for the new ten-volume Second Edition, the task appeared daunting because of its magnitude. But it also seemed manageable because backing me up was a valuable learning experience I had as the editor in chief for Macmillan's single-volume *Supplement*, published in 1996, that updated the *Encyclopedia*. Among the insights I gained from that experience three were especially important.

First, it seemed that the *Encyclopedia* had gained the respect of academicians because its articles provided substantive discussions by exceptionally competent scholars and its coverage embraced a wide range of topics in philosophy broadly construed. That was a winning formula: substantive articles by talented scholars exploring the full spectrum of philosophical topics. It would also guide the Second Edition.

Second, while that winning formula involved in-depth and broad coverage, nevertheless it did not and could not aspire to exhaustive coverage of all philosophical topics given the constraints imposed by the limited print space available. Whether the space available was the eight volumes of the First Edition or the one volume of the *Supplement* or the ten volumes of the Second Edition, a policy of selectivity had to be pursued with the unavoidable exclusion of some material that could have been, and perhaps should have been, included.

Third, to maintain the tradition of excellence established by the First Edition, an editor in chief needs to be surrounded by a group of distinguished philosophers who represent expertise in diverse subfields and who are willing to commit considerable time and effort to serve on an editorial board. I was fortunate indeed to have the support of an editorial team for the *Supplement* consisting of K. Danner Clouser, Paul Horwich, Jaegwon Kim, Joseph J. Kockelmans, Helen E. Longino, Vann McGee, Louis Pojman, Ernest Sosa, and Michael Tooley. Because of them, and the highly competent authors they helped to recruit, the *Supplement* continued Macmillan's tradition of publishing highly regarded reference works.

EDITORIAL BOARD FORMATION

Upon accepting the role of editor in chief for the Second Edition, I immediately turned to three of my former editorial colleagues—Jaegwon Kim, Michael Tooley, and Ernest Sosa—and invited them to become the core of a new Board of Associate Editors that would assist me in planning the new edition. The guidance provided by these three colleagues has been astute, seasoned, and truly indispensable from the early planning stages until the day of publication. With their assistance we were able to recruit Don Garrett, Barry Loewer, Doug MacLean, and Susan Wolf to join the Board of Associate Editors. Then we constituted a Board of Consulting Editors that would add expertise in specific subfields of philosophy not already covered by the specializations of the associate editors. The result was the impressive editorial team of distinguished philosophers listed below. Their areas of editorial oversight are noted after their names.

The Board of Associate Editors

Don Garrett—Modern Philosophy

Jaegwon Kim—Philosophy of Mind

Barry Loewer—Philosophy of Science

Doug MacLean—Ethics and Applied Ethics

Ernest Sosa—Epistemology

Michael Tooley—Metaphysics

Susan Wolf—Ethics and Applied Ethics

The Board of Consulting Editors

Louise Antony—Feminist Philosophy

John Burgess—Logic, Philosophy of Logic, Philosophy of Mathematics

Victor Caston—Ancient Philosophy, Medieval Philosophy

Richard P. Hayes—Buddhist Philosophy

Jeffrey King—Philosophy of Language

Oliver Leaman—Islamic Philosophy, Judaic Philosophy

Vladimir Marchenkov—Russian Philosophy

Thomas Nenon—Continental Philosophy

Karl H. Potter—Indian Philosophy

Philip Quinn—Philosophy of Religion

Jenefer Robinson—Aesthetics, Philosophy of Art

Kwong-loi Shun—Chinese Philosophy

James Sterba—Social and Political Philosophy

Charles Taliaferro—Philosophy of Religion

From the very beginning, our project's goal was not to replace the First Edition and the *Supplement* but to build the Second Edition on the foundation of their outstanding scholarly work. Accordingly, the task set before each editor was to analyze all the entries in the First Edition and the *Supplement* that were pertinent to his or her domain in order to determine which entries should be retained "as is" in the Second Edition with perhaps only a bibliographical update, which entries should be retained but needed an updating addendum, and which entries should be replaced by entirely new ones. In addition, all editors were given the opportunity to commission entirely new entries in their subfields. Each editor also had the responsibility to review and assess all new material appearing in his or her subfield. This generic description of the work of our subfield editors for the Second Edition masks all too easily the many hours of painstaking effort devoted to this project by these scholars.

In early autumn of 2004, regrettably, our editorial colleague Phil Quinn passed away after a brief struggle with esophageal cancer. Prior to his death, however, Phil had overseen his domain with an extraordinarily watchful and skilled eye. He had analyzed in detail every entry relating to the philosophy of religion in the First Edition and the *Supplement*, and sent me copious notes and recommendations for either improving, retaining, or replacing those entries. He also made specific recommendations for new entries to be commissioned and wrote detailed scope descriptions for those entries. When his illness forced him to withdraw from his teaching at the University of Notre Dame, he continued to work on the Second Edition, which provided concrete purpose for the day at hand. Phil worked carefully, deliberately, and had his eye on the prize of excellence. His fine work made it relatively easy for our colleague Charles Taliaferro to assume Phil's responsibilities on the editorial team.

If the Second Edition continues the tradition of excellence initiated by the First Edition, as I believe it will, that accomplishment will be due in no small measure to the exceptionally high quality work provided by our editors who, like Phil, have given of their time and talent to enhance the work of philosophy.

DEVELOPMENT OF THE SECOND EDITION'S CONTENT

Our strategy of building the Second Edition on the foundation of the First Edition and the *Supplement* requires a few additional comments.

Carefully and judiciously our editorial team selected those entries from the First Edition and the *Supplement*

that were so well done that they merited retention. To virtually all of these entries we added bibliographical updates and to many of them we added substantive addenda. We prized these entries because, appearing together with the new entries, they enabled the reader to view high quality philosophizing over the course of almost a half century thereby adding a measure of historical *gravitas* to our project.

Notwithstanding our respect for the First Edition and the *Supplement,* we added 450 entries on new topics, and nearly 300 completely fresh and newly authored treatments of important topics that were originally covered within the First Edition or Supplement. The presence of all of this new material is a clear indication of the vigorous and innovative philosophical activity that has occurred within the discipline since the *Encyclopedia* made its debut almost four decades ago. Entirely new subfields have appeared such as feminist philosophy, the philosophy of sex and love, and applied ethics. New important topics in virtually every subfield have been explored ranging from artificial intelligence to animal rights. New scholars, whose distinctive contributions to the discipline needed description in substantive personal entries, have appeared on the philosophical landscape. Among such individuals are Karl-Otto Apel, Mohammed Arkoun, Nancy Cartwright, Daniel Dennett, Fred Dretske, Ronald Dworkin, John Earman, Hassan Hanafi, Virginia Held, Julia Kristeva, Jacques Lacan, John McDowell, Ruth Millikan, Richard Montague, Thomas Nagel, Seyyed Hossein Nasr, Martha Nussbaum, Derek Parfit, Hilary Putnam, Peter Singer, Gregory Vlastos, Richard Wollheim, and many, many more.

We also added updates to 90 articles, with those updates provided by their original authors. Additionally, 150 scholarly updates to existing articles have been included by means of "addenda," with each addendum compiled by an author other than the original writer, thus allowing for a fresh perspective that augments discussion of the topic at hand. Approximately 430 of the almost 1,200 classic First Edition or *Supplement* articles that appear in the Second Edition have been strengthened further by the inclusion of new bibliographic citations. Classic articles from the First Edition and *Supplement* are clearly identifiable via specific dates in the author bylines that follow each article. Author bylines followed by "(1967)" indicate that the article originally appeared in the First Edition, while bylines followed by "(1996)" indicate first publication within the *Supplement.* The designation "(2005)" denotes first publication within the Second Edition.

We have modified and expanded the philosophical inclusiveness of the First Edition in several ways. Both the analytic and continental philosophical traditions are well represented in the new topics and new personal entries, as well as in the style of presentation offered by our authors. In addition, enhanced cultural diversity is evident in the major space we have provided for topics relating to Buddhist philosophy, Chinese philosophy, Islamic philosophy, and Indian philosophy. Because of space limitations a number of First Edition entries devoted to national philosophies (such as American, British, and German) were not retained. The major figures from those countries and their contributions to philosophy have, however, been included in the Second Edition via personal and topical entries. Importantly, we have retained and expanded the entries on Japanese philosophy, Latin American philosophy, and Russian philosophy, and have added entries on African philosophy and Korean philosophy.

To preserve and enhance the detailed record of philosophical bibliographies, dictionaries, encyclopedias, and journals contained in the First Edition entries devoted exclusively to these topics, we moved these articles to the last volume of the Second Edition and increased substantially the space that had been allocated to them in the First Edition. The very large number of new philosophical bibliographies, dictionaries, encyclopedias, and journals that have been published in a multitude of languages during the last half century testifies not only to the vitality of philosophy but also to the increasing cultural diversity on its landscape.

A FEW FINAL POINTS

Several additional features of our editorial practices are important to note. In retaining entries from the First Edition, we have studiously avoided changing the text of those entries in the interest of preserving the philosophical and authorial integrity of those entries. Some of the authors, however, of those First Edition entries were available and wished to revise their entries. We, of course, welcomed their modifications. On some occasions, without compromising the integrity of an entry, we made some minor changes in the retained First Edition entries, such as inserting the year of death in the biographical part of a personal entry.

The entries in the Second Edition vary in readability level. Many entries will be readily accessible to the general public. Others will require some familiarity with the specialized vocabulary of philosophers. Still other entries will presuppose some acquaintance with logic. All the

entries, it would be safe to say, require the kind of careful reading that is customary in the humanities and that helps to fashion liberally educated persons.

A good number of entries—such as those dealing with ancient, Buddhist, Chinese, Islamic, Judaic, and Russian philosophies—use non-English language words that required transliteration and the use of diacritical marks. In our transliterations and use of diacritical marks we have tried to follow the standard practice adopted by the contemporary leading scholars and the leading journals in the particular subfield to which the entry belongs.

The bibliographies that accompany the entries are selective rather than exhaustive. They provide the references to the works of the scholars cited in the text of an entry. The bibliographical entries in the tenth volume, however, which provide a record of philosophical bibliographies, dictionaries, encyclopedias, and journals, are much more extensive but are not exhaustive.

Volume 10 fulfills at least three important purposes. First, it houses the Appendix, which enabled us to include in the *Encyclopedia* a number of entries that, for a number of reasons, did not move through the editorial process in time to be included in the main alphabetical arrangement of the entries. For example, a few of our contributors encountered unexpected delays in completing their entries because of illness, and a few needed extra time because of other demanding professional commitments. Second, it provided a discrete location where the three lengthy comprehensive bibliographical entries on philosophical dictionaries and encyclopedias, journals, and bibliographies could be bundled together so that they would not distract from the topical and personal entries listed alphabetically in the main body of the set. Third, it contains the Index, a critical access tool for the book's readers.

SPECIAL ACKNOWLEDGMENTS

As editor in chief of this large project I owe a debt of gratitude to many people. I begin with my colleagues at Ohio University. The members of the Philosophy Department were a reservoir of philosophical expertise, good will, and seasoned professional advice. The Philosophy Department's Administrative Assistant, Penny Schall, helped to lighten my tasks, especially with her computer skills. Michael Farmer, the Head of Monographic Cataloging at Ohio University's Alden Library, devoted many painstaking hours to updating the bibliographies of scores of First Edition entries being retained in the Second Edition. The College of Arts and Sciences provided me a professional leave at a crucial juncture in the project so that I could work on the *Encyclopedia* without the standard professorial demands on my time.

Also, I wish to note with appreciation the role played by LinDa L. Grams, the Administrative Assistant in the Philosophy Department at the University of Notre Dame, who graciously served as a conduit of communication between Phil Quinn and me during his all too brief service as the editor overseeing the philosophy of religion.

In addition, there are four groups of people to whom all of us who use the Second Edition owe an expression of appreciation. The first group is the staff of Macmillan Reference and Thomson Gale. Frank Menchaca, Executive Vice President and Publisher, gave the support and encouragement of upper management to the Second Edition to ensure that it would go to press in 2005 and that it would continue the tradition of excellence that has been the hallmark of the reference works published by Macmillan through the years. Hélène Potter, Director of New Product Development, aided by her associates in the New York office, initiated the project and ever so adroitly assisted the editorial team to plan the structure and content of the new edition, and to operationalize those plans in each editor's domain of oversight. The five-person editorial team at Macmillan in Farmington Hills, Michigan, has exhibited seemingly untiring energy to bring the project to press at the targeted time. The core team consisted of Carol Schwartz, Senior Editor and Project Manager, who quarterbacked the team; Jane Malonis, Senior Editor and Project Manager; Brad Morgan, Senior Editor; Deirdre S. Blanchfield, Editor; and Lynn Koch, Associate Editor. This editorial team demonstrated the capacity to multi-task with incredible patience, resilience, diplomacy, and creativeness under many stressful conditions.

The second group to whom we owe words of gratitude consists of the hundreds of scholars who have contributed the multitude of articles that are the substance of the Second Edition. The extraordinarily fine entries that constitute the Second Edition were prepared by scholars with recognized expertise in the topics on which they have written. That fact should assure the reader that forays into the new edition of the *Encyclopedia* will prove to be always educationally valuable. We are deeply grateful for the intellectual heft that these distinguished authors have contributed to the Second Edition.

The third group that merits our appreciation is one that is almost invisible. I refer to the friends and families of our contributors who stood by patiently waiting for our contributors to complete their commitments to our project. Their patience is appreciated. The important contribution to learning that will be made by the new

Second Edition will ensure that the patience of these friends and family members will not have been in vain.

The fourth and final group that deserves appreciation is the team of associate and consulting editors who served on the *Encyclopedia*'s board. They are all very busy, very talented, and very distinguished philosophers. I am amazed and delighted that they were able to find the time to do the tasks that Macmillan and I laid on them. I daresay, however, that they had a special reward accruing from the many hours they devoted to the project. Each of them was asked to assess the new entries in their subfields as those entries were submitted by the authors to Macmillan. The editors were asked to indicate on a review sheet if, in their judgment, the entry at hand should be approved as is, if the entry needed revision, or if the entry should be rejected. As I reviewed the editors' assessments, I marveled at how often editors would characterize the entries as "superb" or "excellent" or "outstanding," and I could almost feel the editor's delight as those words were written on the review sheets. Occasionally, I even saw the words "the finest piece of this length on this topic that has yet been written." Those words exuded the joy and intellectual excitement which are truly the abiding rewards that the editors, and hopefully all readers, will receive from this project.

Donald M. Borchert, 2005

introduction to the first edition, 1967

The last and, in fact, the only previous major philosophical reference work in the English language, J. M. Baldwin's *Dictionary of Psychology and Philosophy*, appeared in 1901. While it was in many ways an admirable work (it numbered among its contributors men of such caliber as Charles Peirce and G. E. Moore), the scope of Baldwin's *Dictionary* was quite limited. The great majority of articles were exceedingly brief, providing concise definitions of technical terms sometimes accompanied by additional information of a historical nature. There were articles about individual philosophers, but these usually amounted to no more than a few lines. Baldwin himself insisted that his work was primarily a dictionary and not an encyclopedia, but he did feature several articles of "encyclopedic character" dealing with important movements in the history of philosophy and the general divisions of philosophy. Some of these "special" articles, as Baldwin called them, were of the highest quality and have become justly famous. Even they, however, were relatively brief—according to Baldwin's own estimate, they varied in length from 1,000 to 5,000 words—and many important questions were entirely neglected or treated in a very cursory fashion. In Baldwin's own day there was undoubtedly room for a philosophical reference work of more ambitious scope. Since then, especially in the light of the revolutionary developments in philosophy and related fields, the need for a truly encyclopedic presentation of philosophical theories and concepts has become increasingly acute.

The present encyclopedia is intended to fill this need. It has been our aim to cover the whole of philosophy as well as many of the points of contact between philosophy and other disciplines. The *Encyclopedia* treats Eastern and Western philosophy; it deals with ancient, medieval, and modern philosophy; and it discusses the theories of mathematicians, physicists, biologists, sociologists, psychologists, moral reformers, and religious thinkers where these have had an impact on philosophy. The *Encyclopedia* contains nearly 1,500 articles of ample length which can be of value to the specialist, while most of them are sufficiently explicit to be read with pleasure and profit by the intelligent nonspecialist. Some of the longer articles, such as those dealing with the history of the various fields of philosophical investigation or the work of the most influential philosophers, are in effect small books, and even the shorter articles are usually long enough to allow a reasonably comprehensive treatment of the subject under discussion. We believe that there is no philosophical concept or theory of any importance that is not identified and discussed in the *Encyclopedia*, although not every concept or theory has a separate article devoted to it. In apportioning the space at our disposal, we were guided by the thought that the majority of readers would derive more benefit from a smaller number of long and integrated articles than from a multitude of shorter entries.

Throughout we have aimed at presentations which are authoritative, clear, comprehensive, and interesting.

Reference works have a reputation, not altogether undeserved, for being deadly dull. There are notable exceptions to this rule, but by and large it is true that the articles in both general and specialized encyclopedias are written in the most colorless prose and shy away from controversial issues. The authors frequently adopt a pose of complete neutrality and Olympian superiority to the conflicts of warring schools of thought, but in practice this usually amounts to an endorsement of safe positions and to neglect or even misrepresentation of radical thinkers, especially if they are contemporaries. Whatever else may be said about it, we do not believe that the present work will be condemned as either dull or timid. Radical movements and thinkers are given their full due, and the most controversial contemporary issues are discussed at great length. Moreover, the authors of the relevant articles were free and welcome to express their own views and in some instances to propose new solutions. It should be added that our contributors were not required to be serious and solemn at all costs, and some of our articles are certain to offend those who believe that philosophy and laughter are incompatible. As a consequence of our approach, the present work may in some respects have a greater resemblance to Dr. Johnson's *Dictionary* and even to Diderot's *Encyclopedia* than to the uncontroversial reference works to which the public has become accustomed in more recent times.

I have no doubt that in years to come a number of the articles in the *Encyclopedia* will be regarded as original contributions to philosophy. This comment refers in particular to articles which deal with controversial philosophical issues, but many of our historical articles also embody original research and in some instances treat topics which have not previously been the subject of thorough scholarly investigations. We have also made it a special point to rescue from obscurity unjustly neglected figures, and in such cases, where the reader would find it almost impossible to obtain reliable information in standard histories or in general encyclopedias, we have been particularly generous in our space allotments. In addition, the reader will find a number of articles on unexpected subjects—such as "Greek Drama," "If," "Nothing," and "Popular Arguments for the Existence of God"—that we considered sufficiently intriguing to be given individual attention.

In the attempt to make the articles interesting, we did not, however, lose sight of the basic goal of any reference work—to supply information in a clear and authoritative fashion. We have been fortunate in obtaining the collaboration of a large number of the foremost philosophers in the world, representing all shades of opinion. It is notorious that philosophy differs from the natural sciences in having no body of generally accepted conclusions. There are, for example, no answers to the problem of causation or the mind-body problem which have the endorsement of all competent students of the subjects; and the same is true of all or nearly all other philosophical problems. However, it is possible to provide an authoritative account of the nature of philosophical problems and of the various attempts to answer them. As far as exposition is concerned, the articles in the *Encyclopedia* are meant to be authoritative: although our contributors were free to express their own opinions, this was never done at the expense of providing the necessary information. To the attentive reader it will always be clear where a writer's exposition ends and the statement of his personal position begins.

Something should perhaps be said at this stage about the question of editorial bias, a subject on which there exists a great deal of confusion. It is important to distinguish two very different varieties of bias. The first is what we may call "polemical" bias—the kind that is operative in political campaigns, in the lower forms of journalism, and wherever fanatics of any kind discuss the views of their opponents. The stock in trade of this kind of partisanship is familiar: where the writer does not resort to deliberate forgery, he nevertheless frequently distorts his opponent's position by quoting out of context and in general by making him look as foolish as possible. Regrettably, philosophers, including some very great ones, have not been above employing such weapons, but in this *Encyclopedia* the use of such techniques has not been allowed. There is, however, another kind of bias which cannot be totally eliminated. No matter how fair and equitable an editor may try to be, his personal views and commitments are bound to affect the organization of the work, the space allotted to different subjects, and the criteria employed in judging the quality of contributions. If this kind of bias cannot be eliminated, its influence can at least be restricted, and it also can and should be openly acknowledged. One method that was used to limit the influence of editorial opinions was to assign articles, wherever possible, to authors who were to some considerable extent sympathetic to the theory or the figure they were to discuss. This rule was adhered to in most, though not in all, cases. It was not applied when there was a serious conflict with other criteria which were also relevant to the selection of contributors. If, for example, an author was in our opinion far superior to all other available writers in such qualifications as intellectual incisiveness and capacity for clear statement, he was chosen even if his

sympathies for the subject of the article were limited. This happened in a few cases, but for the most part we succeeded in finding contributors who met all of our criteria.

It would, nevertheless, be idle to pretend that this *Encyclopedia* is free from bias and that my own ideological commitments have not significantly influenced its content. Like the majority of my closest advisers, I have been raised in the empirical and analytic tradition of Anglo-Saxon philosophy. There can be no doubt that if the *Encyclopedia* had been edited by a follower of Hegel or by a phenomenologist, assuming him to make every effort to be fair and equitable to other viewpoints, it would have looked very different. The topics chosen for separate articles would not have been the same, the space allotments would probably have been appreciably different, and there would undoubtedly have been a significantly different list of contributors. I doubt that an editor with such a background would have featured such articles as "Any and All," "Paradigm-case Argument," and "Proper Names and Descriptions," to give just a few illustrations, or that he would have devoted the same space to logic or to the philosophy of language. I am not here concerned with arguing that what we have done is right and that what other editors, with different commitments, would have done is wrong. I merely wish to remind the reader that in producing an encyclopedia one has to make a vast number of decisions and that one is not in the fortunate position of copying a pre-existing heavenly original. The decisions may be more or less justifiable, but in the last resort they always reflect the beliefs and sympathies of the editors.

We are presenting more than 900 articles on individual thinkers, and any responsible editor, no matter what his viewpoint, would have decided to include articles on the great majority of these. On the other hand, some figures have been omitted who, in the opinion of competent judges, have as good a claim to a separate article as some of those now included. We may as well here and now offer our apologies to all whose lists would have been different and who find that their favorites do not receive adequate attention. Some of these omissions can fairly be blamed on editorial judgment, but others are the result of accidental circumstances. For a number of relatively minor figures even the most diligent search failed to locate a contributor who could write an authoritative and readable article. In such cases it was decided that the space could be put to better use. Fortunately, these omissions are very few, and the ideas of most of the philosophers about whom we should have had separate articles are covered in various of our survey articles on the history of philosophy in different countries, in the articles on philosophical schools and movements, and sometimes also in those dealing with the history of the branches of philosophy. Nevertheless, there are some regrettable gaps, and we can only plead that if one works with over 500 contributors living in every corner of the globe, it is almost impossible that all one's plans should materialize.

One of the most difficult problems confronting the editor of any reference work is that of avoiding duplication without destroying the sense and continuity of individual articles. To be sure, not all duplication is undesirable, especially in a subject in which there is so much disagreement as in philosophy; and in the present work we have not tried to prevent discussions of the same topic in different contexts and from different viewpoints. To give one example, Zeno's paradoxes are discussed in the article bearing the philosopher's name and in the article "Infinity in Mathematics and Logic." The former article critically analyzes the paradoxes considered in the wider context of Greek thought, while in the latter the paradoxes are examined in order to cast light on problems concerning mathematical infinity. We have done our best, however, to avoid all duplication that would not serve a useful purpose. To achieve this end, it was necessary to be extremely flexible in the relative space provisions for various articles. It seemed unwise, for example, to have a lengthy review of the theories of Husserl once in the article bearing his name and then again in the article on phenomenology. In this particular instance we decided to feature a short article under "Husserl" but a very long one under "Phenomenology." This need for flexibility in order to use the available space to maximum advantage will account for many apparent disproportions in our space allotments. The articles on Marx and Engels, to give another illustration, are quite brief—much briefer than those on thinkers who have been far less influential; but this does not mean that Marxism has been neglected in the *Encyclopedia*. For, in addition to the biographical articles on Marx and Engels (and other Marxist thinkers), the *Encyclopedia* contains the very comprehensive articles "Dialectical Materialism," "Historical Materialism," and "Marxist Philosophy," as well as several shorter pieces, in all of which the theories of Marx and Engels are discussed. Our very elaborate index, prepared by a staff of specialists, and our system of cross references have made it possible to avoid a good deal of duplication.

The *Encyclopedia* is primarily the creation of the contributors, and I wish here to record our gratitude to the many fine scholars who have given so much of their time

and energy to this enterprise. A certain type of reader drawn to philosophy is not happy unless he finds a plentiful supply of obscure and high-flown phraseology. Such readers will be disappointed by the present work. Those, on the other hand, who prefer simple and unpretentious language will (we hope) find our *Encyclopedia* to their liking. Nothing can make philosophy into an easy subject, but by taking very great pains it is possible to offer a lucid presentation even of extremely difficult and abstruse philosophical theories. If the majority of our articles are entirely intelligible to most educated readers, this is due to the special care taken by our contributors.

It should also be mentioned that although we were, unfortunately, compelled to reject a number of articles, this in no way reflects on their quality. Many of them were excellent studies and were excluded only for reasons pertaining to problems of space, duplication of material, or other technical considerations. The understanding and patience of all contributors as well as of all whose articles could not be used is greatly appreciated.

We are also very much indebted to the members of the editorial board, whose advice was constantly sought and always readily given. They aided us in a great many ways at all stages—they helped in mapping out the table of contents, in locating suitable contributors, and in evaluating manuscripts. When in the spring and summer of 1965 some absolutely indispensable articles had not arrived, it was chiefly through the intervention of members of the editorial board that outstanding scholars agreed to write the missing articles within the space of a few months. We would like to thank the following contributors for coming to our rescue at the last moment: William P. Alston, Stephen Barker, Thomas G. Bergin, George Boas, Vernon J. Bourke, Wing-tsit Chan, Arthur C. Danto, Phillip H. De Lacy, Ronald Grimsley, Philip P. Hallie, Peter L. Heath, John Hick, Paul O. Kristeller, Hugh R. MacCallum, James E. McClellan, Alasdair Maclntyre, John Macquarrie, F. S. Northedge, Robert G. Olson, John Passmore, Bede Rundle, Colin Smith, W. H. Walsh, and Edward Wasiolek. We are particularly grateful to Professor G. B. Kerferd for writing the article on Aristotle at incredibly short notice. That our extremely detailed and exhaustive article on the history of logic was completed in time is in large measure due to the tireless efforts of Professor A. N. Prior, who was wonderfully helpful in a great many other ways as well.

It would be impossible to praise too highly the performance of the members of the editorial staff. The best testimony to their skill and devotion is the fact that a work of this scope could be completed in a relatively short time by such a small group of people. Ann Trabulsi had the very difficult task of coordinating the work of contributors, editors, copy editors, and the production staff. Her admirable calm and self-possession resolved many a potentially explosive situation, while her tact and firmness worked wonders with even the most reluctant contributors. Philip Cummings, Donald Levy, Sandra Litt, and Margaret Miner were the four full-time editors. Their high standards of scholarship and accuracy, their fine feeling for language, and their unfailing good sense again and again evoked admiring comments and expressions of gratitude from our contributors. Their enthusiasm and their delightful and contagious sense of humor made my own share of the work not only less burdensome but frequently a great deal of fun. Dr. Albert Blumberg joined the editorial staff on a part-time basis early in 1964. It is largely owing to his rich knowledge and painstaking labors that our articles on logic and foundations of mathematics are, as we believe, of an exceedingly high quality. Alix Shulman assisted us during the last year in dealing with various tricky editorial problems, and we are most grateful to her for the excellence of her work. Dr. Murray Greene and Sheila Meyer worked for extended periods in the very onerous position of managing editor, and to both of them I wish to express my appreciation of their valuable contributions. I should also like to thank Mr. Sidney Solomon, who designed the *Encyclopedia* and who was involved in the project from the beginning, for giving valuable advice and assistance on many occasions. Finally, we are all indebted to our editorial secretary, Eunice Dean, whose careful management of our vast and complicated records and correspondence has been an indispensable aid to the production of the *Encyclopedia*.

I have left to the last obligations of a more personal nature. Four of my own articles—"Atheism," "Life, Meaning and Value of," "'My Death,'" and "Why" were written during the academic year 1964/1985 while I held a John Simon Guggenheim Foundation Research Fellowship. The award of this fellowship made it possible for me to take a leave of absence from my teaching duties, and I wish to thank the Guggenheim Memorial Foundation for its generous aid. I should also like to thank the following friends and colleagues for reading one or more of my own articles and for offering criticism and suggestions: Reuben Abel, F. M. Barnard, Sandra Bartky, Milič Čapek, Gertrude Ezorsky, Antony Flew, Peter Heath, Martin Lean, Ruth Barcan Marcus, C. Douglas McGee, Sidney Morgenbesser, Mary Mothersill, Ernest Nagel, Andrew Oldenquist, Robert Olson, Richard Popkin, Bertrand Russell, J. B. Schneewind, Elmer Sprague, and Carl Wellman. In connection with the difficult article about Wil-

helm Reich I am especially grateful for advice and comments to Mr. A. S. Neill, Drs. Allan Cott and Ola Raknes (all of whom knew Reich well), and to Sir Karl Popper, Alasdair Maclntyre, Sidney Hook, and Michael Scriven. Needless to say, none of those who kindly helped me with my articles is responsible for any of the views expressed in them. To my dear friend and teacher, Ernest Nagel, I am deeply grateful for his unfailing encouragement and moral support ever since I began to edit the *Encyclopedia*. In spite of his many obligations he always found time to listen to our problems and to offer suggestions based on his immense erudition and his acquaintance with scholars in the most diverse fields.

Paul Edwards, Brooklyn College, March 1966

list of contributors

Contributors to the encyclopedia are listed below in alphabetic order followed by their academic affiliations and the article(s) they contributed. Articles reprinted from the first edition and supplement are indicated respectively by (1967) or (1996) following the article name. Affiliations provided for the authors of these articles were their 1967 or 1996 affiliations. New or updated articles are indicated by (2005) and include the current affiliation for the author.

Nicola Abbagnano
Professor, History of Philosophy, University of Turin
ALIOTTA, ANTONIO (1967)
ARDIGÒ, ROBERTO (1967)
FERRI, LUIGI (1967)
LEONARDO DA VINCI (1967)
POSITIVISM (1967)
PSYCHOLOGISM (1967)
RENSI, GIUSEPPE (1967)
RIGNANO, EUGENIO (1967)

Reuben Abel
Adjunct Associate Professor of Philosophy, Graduate Faculty, New School for Social Research, and Chairman of the Division of Humanities
SCHILLER, FERDINAND CANNING SCOTT (1967)

Francine F. Abeles
Mathematics, Kean University
LOGIC, HISTORY OF: MODERN LOGIC: THE BOOLEAN PERIOD: CARROLL (2005)

Raziel Abelson
Associate Professor and Chairman, Philosophy, University College, New York University
DEFINITION (1967)
ETHICS, HISTORY OF (1967)

Peter Achinstein
Professor of Philosophy, Johns Hopkins University
BRAITHWAITE, RICHARD BEVAN (1967, 2005)

Alparslan Açıkgenç
Professor, Philosophy, Fatih University, Istanbul
CAUSATION IN ISLAMIC PHILOSOPHY (2005)

H. B. Acton
Professor of Moral Philosophy, University of Edinburgh; Editor of Philosophy
ABSOLUTE, THE (1967)
BERKELEY, GEORGE (1967)
BOSANQUET, BERNARD (1967)
BRADLEY, FRANCIS HERBERT (1967)
DIALECTICAL MATERIALISM (1967)
HEGEL, GEORG WILHELM FRIEDRICH (1967)
HISTORICAL MATERIALISM (1967)
IDEALISM (1967)

David Adams
Professor of Philosophy, California State Polytechnic University, Pomona
PHILOSOPHY OF LAW, HISTORY OF [ADDENDUM] (2005)

E. M. Adams
Professor of Philosophy, University of North Carolina
LEWIS, CLARENCE IRVING (1967)

Robert M. Adams
Yale University
PHILOSOPHY OF RELIGION (1996)

Alfred Owen Aldridge
Head of Department of Comparative Literature, University of Maryland
PAINE, THOMAS (1967)

Peter Alexander
Reader in Philosophy, University of Bristol
DUHEM, PIERRE MAURICE MARIE (1967)
HERTZ, HEINRICH RUDOLF (1967)
PEARSON, KARL (1967)
POINCARÉ, JULES HENRI (1967)
SENSATIONALISM (1967)

W. M. Alexander
Associate Professor of Religion and Philosophy, St. Andrews College
HAMANN, JOHANN GEORG (1967)

Edwin Allaire
Professor of Philosophy, University of Texas at Austin
BERGMANN, GUSTAV (2005)

James Allen
Professor, Philosophy, University of Pittsburgh
ANCIENT SKEPTICISM (2005)
ANTIOCHUS OF ASCALON (2005)
ARCESILAUS (2005)
CARNEADES (2005)
PHILO OF LARISSA (2005)

Felix Alluntis O.F.M.
Ordinary Professor of Philosophy, Catholic University of America
VITORIA, FRANCISCO DE (1967)

Robert Almeder
McCullough Professor of Philosophy, Hamilton College, Clinton, NY
RESCHER, NICHOLAS (2005)

William P. Alston
Professor of Philosophy, University of Michigan
PHILOSOPHY OF RELIGION, PROBLEMS OF (1967)
PLEASURE (1967)
PSYCHOANALYTIC THEORIES, LOGICAL STATUS OF (1967)
RELIGION (1967)
RELIGION, NATURALISTIC RECONSTRUCTIONS OF (1967)
RELIGION, PSYCHOLOGICAL EXPLANATIONS OF (1967)
RELIGIOUS LANGUAGE (1967)
TELEOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD (1967)
TILLICH, PAUL (1967)

Andrew Altman
Professor of Philosophy; Director, Jean Beer Blumenfeld Center for Ethics, Georgia State University
DWORKIN, RONALD (2005)

Robert Anchor
Instructor of History, Yale University
RICKERT, HEINRICH (1967)
RITSCHL, ALBRECHT BENJAMIN (1967)

C. Anthony Anderson
Professor, Philosophy, University of California, Santa Barbara
CHURCH, ALONZO (2005)

Stephen C. Angle
Associate Professor of Philosophy, Wesleyan University
ZHU XI (CHU HSI) (2005)

Aldo Antonelli
Professor of Logic and Philosophy of Science, University of California, Irvine
FEMINIST PHILOSOPHY (2005)
INNATE IDEAS, NATIVISM (2005)
NON-MONOTONIC LOGIC (2005)
SEXISM (2005)

Roger Ariew
Professor and Chair, Philosophy, University of South Florida
CLAUBERG, JOHANNES (2005)
CORDEMOY, GÉRAUD DE (2005)

David M. Armstrong
University of Syndey
LAWS OF NATURE (1996)

Douglas Arner
Professor of Philosophy, Arizona State University
MCCOSH, JAMES (1967)

John Arthur
Professor of Philosophy; Director, Program in Philosophy, Politics and Law, Binghamton University
STATE [ADDENDUM] (2005)

Robert J. Arway, C.M.
Associate Professor of Philosophy, St. John's University
GODFREY OF FONTAINES (1967)

R. W. Ashby
Lecturer in Philosophy, King's College, University of London
BASIC STATEMENTS (1967)
VERIFIABILITY PRINCIPLE (1967)

Margaret Atherton
Professor of Philosophy, University of Wisconsin–Milwaukee
ASTELL, MARY (2005)
CAVENDISH, MARGARET (2005)
COCKBURN, CATHARINE TROTTER (2005)
SHEPHERD, MARY (2005)

Samuel Atlas
Professor, Philosophy, Hebrew Union College, New York
JACOBI, FRIEDRICH HEINRICH (1967)
SCHULZE, GOTTLOB ERNST (1967)

Robert Audi
Professor of Philosophy and David E. Gallo Chair in Ethics, University of Notre Dame
PHILOSOPHY (2005)

Bruce A. Aune
Emeritus Professor of Philosophy, University of Massachusetts, Amherst
CAN (1967)
POSSIBILITY (1967, 2005)
SELLARS, WILFRID (1996, 2005)
THINKING (1967)

Franz Austeda
Councilor of the School Board of Vienna
AVENARIUS, RICHARD (1967)
FISCHER, KUNO (1967)
JODL, FRIEDRICH (1967)
PETZOLDT, JOSEPH (1967)
RORETZ, KARL (1967)
STÖHR, ADOLF (1967)
WAHLE, RICHARD (1967)
ZIEHEN, THEODOR (1967)

Kent Bach
Professor of Philosophy, San Francisco State University
PERFORMATIVE UTTERANCES (2005)

Keith Michael Baker
Assistant Professor of History, University of Chicago
CONDORCET, MARQUIS DE (1967)

Edward G. Ballard
Professor of Philosophy, Tulane University
LACHELIER, JULES (1967)

Han Baltussen
Lecturer, Classics, University of Adelaide, South Australia
THEOPHRASTUS (2005)

Renford Bambrough
Fellow, Dean and Director, Studies in Moral Sciences, St. John's College
DEMIURGE (1967)
GREEK DRAMA (1967)

Julian Barbour
Director, The Leibniz Institute, Oxfordshire, U.K.
MACH, ERNST (2005)

Yehoshua Bar-Hillel
Professor of Logic and Philosophy of Science, Hebrew University of Jerusalem
BOLZANO, BERNARD (1967)
LOGIC, HISTORY OF: PRECURSORS OF MODERN LOGIC: BOLZANO (1967)

SYNTACTICAL AND SEMANTICAL CATEGORIES (1967)

Frederick M. Barnard
Associate Professor, Political Science, University of Saskatchewan
BACHOFEN, JOHANN JAKOB (1967)
MORGAN, LEWIS HENRY (1967)
REINHOLD, KARL LEONHARD (1967)
SPINOZISM (1967)
SUMNER, WILLIAM GRAHAM (1967)
WEBER, ALFRED (1967)

Rachel Barney
Canada Research Chair in Classical Philosophy; Associate Professor, Departments of Classics and Philosophy, University of Toronto
GORGIAS OF LEONTINI (2005)
NOMOS AND PHUSIS (2005)
SOPHISTS (2005)

Jeffrey A. Barrett
Professor, Logic and Philosophy of Science, University of California Irvine
MANY WORLDS/MANY MINDS INTERPRETATION OF QUANTUM MECHANICS (2005)

Jean-Pierre Barricelli
Associate Professor of Romance Languages and Comparative Literature; Chair, Department of French and Italian, University of California, Riverside
LEOPARDI, COUNT GIACOMO (1967)

Irving H. Bartlett
Professor and Chair, Philosophy, Duke University
CHANNING, WILLIAM ELLERY (1967)
PARKER, THEODORE (1967)

Heather D. Battaly
Assistant Professor, Philosophy, California State University, Fullerton
ALSTON, WILLIAM P. (2005)

Robert Batterman
Rotman Canada Research Chair in Philosophy of Science, University of Western Ontario
REDUCTION (2005)

Margaret Pabst Battin
Distinguished Professor, Philosophy, and Adjunct Professor, Internal Medicine, Division of Medical Ethics, University of Utah, Salt Lake City
SUICIDE (2005)

Charles A. Baylis
Professor and Chair, Philosophy, Duke University
CONSCIENCE (1967)

Ken Baynes
Professor, Philosophy, Syracuse University
APEL, KARL-OTTO (2005)

George Bealer
University of Colorado, Boulder
INTUITION [ADDENDUM 1] (1996)

Monroe C. Beardsley
Professor and Acting Chairman, Philosophy, Swarthmore College
AESTHETICS, HISTORY OF (1967)

Tom L. Beauchamp
Georgetown University
APPLIED ETHICS (1996)

Lewis White Beck
Burbank Professor of Intellectual and Moral Philolosphy, University of Rochester
NEO-KANTIANISM (1967)
STERN, LOUIS WILLIAM (1967)

Morton O. Beckner
Associate Professor of Philosophy, Pomona College
DARWINISM (1967)
ORGANISMIC BIOLOGY (1967)
TELEOLOGY (1967)
VITALISM (1967)

Jason BeDuhn
Associate Professor of Religious Studies, Humanities, Arts, and Religion, Northern Arizona University
MANI AND MANICHAEISM [BIBLIOGRAPHY] (2005)

Gordon Belot
Associate Professor of Philosophy, University of Pittsburgh
CONSERVATION PRINCIPLE (2005)

John W. Bender
Professor of Philosophy, Ohio University
ART, EXPRESSION IN (1996, 2005)
COHERENTISM (1996)

Stanley I. Benn
Senior Fellow in Philosophy, Australian National University
DEMOCRACY (1967)
EQUALITY, MORAL AND SOCIAL (1967)
NATIONALISM (1967)
POWER (1967)
PROPERTY (1967)
PUNISHMENT (1967)
SOCIETY (1967)
SOVEREIGNTY (1967)
STATE (1967)

Carlton W. Berenda
Professor of Philosophy, University of Oklahoma
WEYL, (CLAUS HUGO) HERMANN (1967)

Thomas Goddard Bergin
Sterling Professor of Romance Languages and Mast of Timothy Dwight College, Yale University
DANTE ALIGHIERI (1967)

Robert Bernasconi
Moss Professor of Philosophy, University of Memphis
ALTERITY (2005)
DERRIDA, JACQUES (1996, 2005)
PHILOSOPHY OF LANGUAGE IN CONTINENTAL PHILOSOPHY (2005)

Paul Bernays
Retired Professor, Swiss Federal Institute of Technology (Zurich)
HILBERT, DAVID (1967)

Arthur Berndston
Professor, Chair, Philosophy, University of Missouri
DEUSTUA, ALEJANDRO O. (1967)
CASO, ANTONIO (1967)
INGENIEROS, JOSÉ (1967)
KORN, ALEJANDRO (1967)
ROMERO, FRANCISCO (1967)
VASCONCELOS, JOSÉ (1967)
VAZ FERREIRA, CARLOS (1967)

Richard J. Bernstein
Chair, Philosophy, Haverford College
DEWEY, JOHN (1967)

Sylvia Berryman
Assistant Professor, Philosophy, University of British Columbia
PNEUMA (2005)
STRATO AND STRATONISM (2005)

Peter A. Bertocci
Bowne Professor of Philosophy, Brown University (Emeritus)
BOWNE, BORDEN PARKER (1967)
HOWISON, GEORGE HOLMES (1967)

Gábor Betegh
Associate Professor, Philosophy, Central European University, Budapest
DIOGENES OF APOLLONIA (2005)
MOIRA/TYCHĒ/ANANKĒ (2005)
ORPHISM [ADDENDUM] (2005)

Joël Biard
Professeur des Universités, Centre d'Études Supérieures de la Renaissance, Université de Tours
ALBERT OF SAXONY (2005)

Cristina Bicchieri
Carol and Michael Lowenstein Endowed Term Professor; Director, Philosophy, Politics and Economics Program; Professor of Philosophy, University of Pennsylvania
GAME THEORY (2005)

John Bickle
Professor, Philosophy and Neuroscience Graduate Program, University of Cincinnati
NEUROSCIENCE (2005)

John Bigelow
Professor, School of Philosophy and Bioethics, Monash University, Australia
NUMBER (2005)

Robert Bird
Assistant Professor, Slavic Languages and Literatures, University of Chicago
IVANOV, VIACHESLAV IVANOVICH (2005)
ROZANOV, VASILII VASIL'EVICH [ADDENDUM] (2005)

Robert Bishop
Lecturer, Faculty of Philosophy, University of Oxford
DETERMINISM AND INDETERMINISM (2005)

Michel Bitbol
Directeur de recherche au CNRS; Chargé de cours à l'Université Paris I, Centre de Recherches en Epistémologie Appliquée, (CREA/Ecole Polytechnique)
SCHRÖDINGER, ERWIN (2005)

Max Black
Professor, Philosophy, Cornell University
INDUCTION (1967)
RAMSEY, FRANK PLUMPTON (1967)

Simon Blackburn
University of North Carolina, Chapel Hill
ERROR THEORY OF ETHICS (1996)
RULE FOLLOWING (1996)

Robert Blanché
Professor of Philosophy, Faculty of Letters and Humane Sciences, University of Toulouse
COUTURAT, LOUIS (1967)
GOD, CONCEPTS OF (1967)
MEYERSON, ÉMILE (1967)
MILHAUD, GASTON (1967)
ROUGIER, LOUIS (1967)
WHEWELL, WILLIAM (1967)

Brand Blanshard
Sterling Professor Emeritus, Yale University
WISDOM (1967)

J. L. Blau
Professor, Religion, Columbia University
ALBO, JOSEPH (1967)
BAḤYA BEN JOSEPH IBN PAQUDA (1967)
CORDOVERO, MOSES BEN JACOB (1967)
HICKOK, LAURENS PERSEUS (1967)
IBN ZADDIK, JOSEPH BEN JACOB (1967)
ISRAELI, ISAAC BEN SOLOMON (1967)
JAMES, HENRY (1967)
KABBALAH (1967)
MATHER, COTTON (1967)
MUQAMMIṢ, DAVID BEN MERWAN AL- (1967)
PORTER, NOAH (1967)
SELLARS, ROY WOOD (1967)
WAYLAND, FRANCIS (1967)

Ned Block
Massachusetts Institute of Technology
FUNCTIONALISM (1996)

H. Gene Blocker
Professor, Philosophy, Ohio University
CHINESE PHILOSOPHY: SOCIAL AND POLITICAL THOUGHT (2005)
JAPANESE PHILOSOPHY (2005)

Mary K. Bloodsworth-Lugo
Associate Professor, Philosophy, Washington State University, Pullman
CIXOUS, HÉLÈNE (2005)

Philip Blosser
Professor of Philosophy, School of History, Philosophy, and Religion, Lenoir-Rhyne College
SCHELER, MAX (2005)

William T. Bluhm
Associate Professor of Political Science, University of Rochester
HARRINGTON, JAMES (1967)

Lawrence A. Blum
Professor of Philosophy, Distinguished Professor of Liberal Arts and Education, University of Massachusetts, Boston
ETHICS AND MORALITY (2005)
MURDOCH, IRIS (2005)

James Blumenthal
Associate Professor, Philosophy, Oregon State University, Corvallis
BUDDHISM—SCHOOLS: DGE-LUGS (2005)

George Boas
Professor Emeritus of the History of Philosophy, Johns Hopkins University
BONALD, LOUIS GABRIEL AMBROISE, VICOMTE DE (1967)
BURTHOGGE, RICHARD (1967)
CHATEAUBRIAND, FRANÇOIS RENÉ DE (1967)
COUSIN, VICTOR (1967)
DESTUTT DE TRACY, ANTOINE LOUIS CLAUDE, COMTE (1967)
JOUFFROY, THÉODORE SIMON (1967)
LAMENNAIS, HUGUES FÉLICITÉ ROBERT DE (1967)
LAROMIGUIÈRE, PIERRE (1967)

LOVE (1967)
LOVEJOY, ARTHUR ONCKEN (1967)
MAISTRE, COMTE JOSEPH DE (1967)
RAVAISSON-MOLLIEN, JEAN GASPARD FÉLIX (1967)
RENOUVIER, CHARLES BERNARD (1967)
ROYER-COLLARD, PIERRE PAUL (1967)
STAËL-HOLSTEIN, ANNE LOUISE GERMAINE NECKER, BARONNE DE (1967)
TRADITIONALISM (1967)

Susanne Bobzien
Professor, Philosophy, Yale University
LOGIC, HISTORY OF: ANCIENT LOGIC (2005)

Margaret Boden
Research Professor of Cognitive Science, Centre for Cognitive Science, University of Sussex
ARTIFICIAL INTELLIGENCE (1996, 2005)
COGNITIVE SCIENCE (1996, 2005)

István M. Bodnár
Philosophy, Central European University
ALEXANDER OF APHRODISIAS (2005)

Hermann Boeschenstein
Head, Department of German, University of Toronto
OKEN, LORENZ (1967)

Paul Artin Boghossian
New York University
ANALYTICITY (1996)

Bernadine M. Bonansea, O.F.M
Ordinary Professor of Philsophy, Catholic University of America
CAMPANELLA, TOMMASO (1967)
SCOTISM (1967)
TELESIO, BERNARDINO (1967)

David Boonin
Associate Professor, Philosophy, University of Colorado, Boulder
ABORTION [BIBLIOGRAPHY] (2005)

Vernon J. Bourke
Professor, Philosophy, St. Louis University; President of the World Union of Catholic Philosophical Societies
BÁÑEZ, DOMINIC (1967)
BELLARMINE, ST. ROBERT (1967)
BIEL, GABRIEL (1967)
CAPREOLUS, JOHN (1967)
FONSECA, PETER (1967)
JOHN OF ST. THOMAS (1967)
MARIANA, JUAN DE (1967)
SOTO, DOMINIC DE (1967)
SYLVESTER OF FERRARA, FRANCIS (1967)
THOMAS AQUINAS, ST. (1967)
TOLETUS, FRANCIS (1967)
VASQUEZ, GABRIEL (1967)

Bernard R. Boxill
Pardue Professor, Philosophy, University of North Carolina at Chapel Hill
AFFIRMATIVE ACTION (2005)
KING, MARTIN LUTHER (2005)

Michael Boylan
John J. McDonnell Jr. Chair in Ethics, Marymount University
JEFFERSON, THOMAS (2005)
MACHIAVELLI, NICCOLÒ [ADDENDUM] (2005)
MARSILIUS OF PADUA [ADDENDUM] (2005)
OAKESHOTT, MICHAEL (2005)

G. R. Boys-Stones
Senior Lecturer in Classics, Department of Classics and Ancient History, Durham University
CLEMENT OF ALEXANDRIA [BIBLIOGRAPHY] (2005)
NUMENIUS OF APAMEA [BIBLIOGRAPHY] (2005)

Germaine Brée
Vilas Professor, Institute for Research in the Humanities, Madison, WI
MALRAUX, GEORGES-ANDRÉ (1967)
WEIL, SIMONE (1967)

Ignatius Brady, O.F.M
Prefect, Theological Comission, Collegio di S. Bonaventura, Franciscan International College of Research, Quaracchi, Italy
ALEXANDER OF HALES (1967)
JOHN OF LA ROCHELLE (1967)
PETER LOMBARD (1967)

Craig Brandist
Reader in Cultural Theory and Intellectual History, Bakhtin Centre and Department of Russian and Slavonic Studies, University of Sheffield
BAKHTIN CIRCLE, THE (2005)

Frithhiof Brandt
Professor of Philosophy (Emeritus), University of Copenhagen
HØFFDING, HARALD (1967)

Richard B. Brandt
Professor, Philosophy, University of Michigan
EPISTEMOLOGY AND ETHICS, PARALLEL BETWEEN (1967)
ETHICAL RELATIVISM (1967)
HEDONISM (1967)

Johan Brännmark
Philosophy, Lund University, Sweden
GOOD, THE (2005)

David Braun
Professor, Philosophy, University of Rochester
DEMONSTRATIVES (2005)

Samantha Brennan
Associate Professor, Philosophy, The University of Western Ontario
CARD, CLAUDIA (2005)
HELD, VIRGINIA (2005)

Phillip Bricker
University of Massachusetts, Amherst
IDENTITY (1996)
PROPERTIES (1996)

Harvey H. Brimmer II
Assistant Professor, Philosophy, University of Maine
LEQUIER, (JOSEPH LOUIS) JULES (1967)

Crane Brinton
McLean Professor of Ancient and Modern History, Harvard University
ROMANTICISM (1967)

Susan J. Brison
Associate Professor, Philosophy, Dartmouth College
BEAUVOIR, SIMONE DE (1996, 2005)

Justin Broackes
Brown University
COLORS (1996)

EXTRINSIC AND INTRINSIC PROPERTIES (1996)

Dan W. Brock
Brown University
INFORMED CONSENT (1996)

Boruch A. Brody
Fulbright Fellow, Oxford University
LOGICAL TERMS, GLOSSARY OF (1967)

Sylvain Bromberger
Massachusetts Institute of Technology
PHONOLOGY (1996)

Andrew Brook
Professor of Philosophy, Director, Institute of Cognitive Science; Member, Canadian Psychoanalytic Society
DENNETT, DANIEL CLEMENT (2005)

Charlotte R. Brown
Associate Professor, Philosophy, Illinois Wesleyan University
ALTRUISM (2005)
SHAFTESBURY, THIRD EARL OF (ANTHONY ASHLEY COOPER) (2005)
WOLLASTON, WILLIAM (2005)

Eric Brown
Associate Professor, Philosophy, Washington University in St. Louis
EPICTETUS (2005)
MARCUS AURELIUS ANTONINUS (2005)
PANAETIUS OF RHODES (2005)

Robert Brown
Senior Fellow, Philosophy, Institute of Advanced Studies, Australian National University
BROAD, CHARLIE DUNBAR (1967)

Stephen F. Brown
Director of the Institute of Medieval Philosophy and Theology; Professor, Theology, Boston College
GODFREY OF FONTAINES [BIBLIOGRAPHY] (2005)
GREGORY OF RIMINI [BIBLIOGRAPHY] (2005)

Anthony Brueckner
Professor, Philosophy, University of California, Santa Barbara
SKEPTICISM, CONTEMPORARY (1996, 2005)

Leendert Brummel
Professor, Library Science, University of Amsterdam
HEMSTERHUIS, FRANS (1967)

Alexander Brungs
Research Fellow, Philosophy, Universität Zürich, Switzerland
INNER SENSES (2005)

Ronald Bruzina
Professor, Philosophy, University of Kentucky, Lexington
FINK, EUGEN (2005)

Jeffrey Bub
Chair, Committee for Philosophy and the Sciences, Philosophy Department, University of Maryland, College Park
COPENHAGEN INTERPRETATION (2005)
QUANTUM COMPUTING AND TELEPORTATION (2005)

Allen Buchanan
James B. Duke Professor of Philosophy and Public Policy Studies, Duke University
APPLIED ETHICS (2005)

Gerd Buchdahl
Lecturer in Philosophy of Science; Head, History and Philosophy of Science, Cambridge University
CAMPBELL, NORMAN ROBERT (1967)

Malcolm Budd
Emeritus Grote Professor of Philosophy of Mind and Logic, University College London
WOLLHEIM, RICHARD (2005)

Otávio Bueno
Associate Professor, Philosophy, University of South Carolina
VAN FRAASSEN, BAS (2005)

John P. Burgess
Professor, Philosophy, Princeton University
KRIPKE, SAUL (1996, 2005)
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL [OVERVIEW] (2005)
QUANTIFIERS IN FORMAL LOGIC (2005)

Keith Burgess-Jackson
Associate Professor, Philosophy, The University of Texas at Arlington
FEINBERG, JOEL (2005)

Charles Burnett
Professor of the History of Arabic/Islamic Influence in Europe, Warburg Institute, University of London
HERMETICISM [ADDENDUM] (2005)

David Burrell, C.S.C.
Hesburgh Professor in Philosophy and Theology, University of Notre Dame
AL-GHAZĀLĪ, MUHAMMAD [ADDENDUM] (2005)
ETERNITY [ADDENDUM 2] (2005)

Sam Butchart
Research Fellow, School of Philosophy and Bioethics, Monash University, Australia
NUMBER (2005)

Alex Byrne
Associate Professor, Philosophy, Massachusetts Institute of Technology
PRIVATE LANGUAGE PROBLEM [ADDENDUM] (2005)

Duane L. Cady
Hamline University
VIOLENCE (1996)

Steven M. Cahn
Doctoral Candidate, Columbia University
CHANCE (1967)

Craig Callender
Associate Professor of Philosophy, University of California, San Diego
TIME IN PHYSICS (2005)

Massimo Campanini
Reader in Arab Culture and Civilization, Faculty of Letters and Philosophy, University of Milan
HANAFĪ, HASSAN (2005)

A. H. Campbell
Regius Professor of Public Law, University of Edinburgh
DEL VECCHIO, GIORGIO (1967)

Keith Campbell
Emeritus Professor of Philosophy, University of Sydney
MACKIE, JOHN LESLIE (1996, 2005)
MATERIALISM (1967, 2005)
NATURALISM (2005)
ONTOLOGY (2005)
SMART, JOHN JAMIESON CARSWELL (1996, 2005)

Mariano Campo
Professor Emeritus of the History of Philosophy, University of Trieste
LIEBMANN, OTTO (1967)
RIEHL, ALOIS (1967)

Kenneth L. Caneva
Professor, History, University of North Carolina at Greensboro
KUHN, THOMAS (2005)

Walter F. Cannon
Curator of Astronomy and Physics; Curator in Charge of Division of Physical Sciences, Smithsonian Institution
HERSCHEL, JOHN (1967)

Milič Čapek
Professor of Philosophy, Boston University
AMPÉRE, ANDRÉ MARIE (1967)
ETERNAL RETURN (1967)
OSTWALD, WILHELM (1967)
RIBOT, THÉODULE ARMAND (1967)
TAINE, HIPPOLYTE-ADOLPHE (1967)

Ben Caplan
Associate Professor, Philosophy, University of Manitoba
KAPLAN, DAVID (2005)

A. Robert Caponigri
Professor, Philosophy, University of Notre Dame
GIOBERTI, VINCENZO (1967)
MARTINETTI, PIERO (1967)
ROSMINI-SERBATI, ANTONIO (1967)
SCIACCA, MICHELE FEDERICO (1967)
STEFANINI, LUIGI (1967)

George Cardona
Professor Emeritus of Linguistics, University of Pennsylvania, Philadelphia
PHILOSOPHY OF LANGUAGE IN INDIA (2005)

Stefano Caroti
Professore Ordinario di Storia della Filosofia Medievale, Dipartimento di Filosofia, Università degli Studi di Parma
ORESME, NICOLE (2005)

Brian Carr
Honorary University Fellow, University of Exeter
CAUSATION IN INDIAN PHILOSOPHY (2005)

David Carr
Charles Howard Candler Professor of Philosophy, Emory University
PHILOSOPHY OF HISTORY (2005)

Meyrick H. Carré
Reader in Philosophy at University of Bristol (retired)
PHYSICOTHEOLOGY (1967)

John Carroll
Professor of Philosophy, Department of Philosophy and Religion, North Carolina State University, Raleigh
LAWS OF NATURE [ADDENDUM] (2005)

Noel Carroll
Professor, Philosophy, University of Wisconsin, Madison
ART, DEFINITIONS OF (2005)
ART, INTERPRETATION OF (2005)
ART, TRUTH IN (2005)

Scott Carson
Ohio University
ALBERT THE GREAT (2005)
EUDAIMONIA (2005)
PHRONÊSIS (2005)
SOPHIA (2005)
SÔPHROSUNÊ (2005)
VALENTINUS AND VALENTINIANISM (2005)
XENOPHON [ADDENDUM] (2005)

Thomas Carson
Professor of Philosophy, Loyola University Chicago
METAETHICS (2005)

Robyn Carston
Professor of Linguistics, Department of Phonetics and Linguistics, University College
PRAGMATICS [ADDENDUM] (2005)

Benjamin Carter
Special Lecturer, Historical Studies, University of Bristol
PANNENBERG, WOLFHART (2005)

Andre Carus
Graduate Student, Philosophy, University of Chicago
CARNAP, RUDOLF (2005)
POSITIVISM [BIBLIOGRAPHY] (2005)

John Carvalho
Associate Professor, Philosophy, Villanova University
BARTHES, ROLAND (2005)

Héctor-Ner Castanñeda
Professor of Philosophy, Wayne State University
PRIVATE LANGUAGE PROBLEM (1967)

Albert Casullo
Professor, Philosophy, University of Nebraska-Lincoln
KNOWLEDGE, A PRIORI (2005)
KNOWLEDGE AND MODALITY (1996, 2005)

Walter Cerf
Visiting Professor, Philosophy, University of Wisconsin; Professor, Philosophy, City University of New York, Brooklyn College
HARTMANN, NICOLAI (1967)

Henry Chadwick
Regius Professor of Divinity at Oxford University
LESSING, GOTTHOLD EPHRAIM (1967)

Arindam Chakrabarti
Professor, Philosophy, University of Hawaii at Manoa
UNIVERSAL PROPERTIES IN INDIAN PHILOSOPHICAL TRADITIONS (2005)

Alan Chan
Professor, Philosophy, National University of Singapore
GUO XIANG (2005)
WANG BI (2005)

Wing-Tsit Chan
Professor of Chinese Culture and Philosophy, Dartmouth College; Adjunct Professor of Chinese Thought, Columbia University

CHENG HAO (1967)
CHENG YI (1967)
CHINESE PHILOSOPHY [OVERVIEW] (1967)
ZHUANGZI (1967)

Carsun Chang
President, Institute of Political Science, Shanghai (retired)
LU XIANGSHAN (1967)

Christopher K. Chapple
Professor, Theological Studies, Loyola Marymount University
MEDITATION IN INDIAN PHILOSOPHY (2005)

Louis Charland
Associate Professor, Philosophy, University of Western Ontario
EMOTION (2005)

Sebastien Charles
Associate Professor, Philosophy, Université de Sherbrooke
VOLTAIRE, FRANÇOIS-MARIE AROUET DE (2005)

Jo-shui Chen
Research Fellow, Institute of History and Philology, Academia Sinica, Taiwan
HAN YU (2005)
LI AO (2005)
WANG CHONG (2005)

Chung-ying Cheng
Professor of Philosophy, University of Hawaii at Manoa
DAI ZHEN (2005)
ZHANG ZAI (2005)

(John) Hsueh-li Cheng
Professor, Philosophy and Religious Studies, University of Hawaii at Hilo
CHINESE PHILOSOPHY: BUDDHISM (2005)

Roderick M. Chisholm
Romeo Elton Professor of Natural Theology and Professor of Philosophy, Brown University
BRENTANO, FRANZ (1967)
INTENTIONALITY (1967)
MARTY, ANTON (1967)
MEINONG, ALEXIUS (1967)

William Chittick
Professor, Asian and Asian-American Studies, Stony Brook University
IBN AL-ʿARABĪ (2005)

Peter Cholak
Professor, Mathematics, University of Notre Dame
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL: FRIEDMAN AND REVERSE (2005)
REVERSE MATHEMATICS (2005)

Edith Clowes
Professor, Slavic Languages and Literatures, University of Kansas, Lawrence
SHESTOV, LEV ISAAKOVICH [ADDENDUM] (2005)

Duane L. Cody
Hamline University
PACIFISM (1996)

Carl Cohen
Professor of Philosophy, The University of Michigan, Ann Arbor
DEMOCRACY [ADDENDUM] (2005)

L. J. Cohen
Queen's College, Oxford University, England
PRIOR, ARTHUR NORMAN (1996)

Robert S. Cohen
Professor and Chair, Physics, Boston University
NEURATH, OTTO (1967)

Ted Cohen
Professor, Philosophy, University of Chicago
AESTHETIC JUDGMENT (2005)
ART, FORMALISM IN (2005)

Margaret Cole
President, Fabian Society, Vice-Chairman of the Further and Higher Education Committee of the Inner London Education Authority
SOCIALISM (1967)

R. L. Cole
Professor, History and English, University of Iowa
ARMINIUS AND ARMINIANISM (1967)

Jules L. Coleman
Wesley Newcomb Hohfeld Professor of Jurisprudence, Yale Law School, Professor of Philosophy, Yale University
ANALYTIC JURISPRUDENCE (2005)
LEGAL POSITIVISM: ANGLO-AMERICAN LEGAL POSITIVISM SINCE H. L. A. HART (2005)

James Collins
Professor, Philosophy, St. Louis University
NEWMAN, JOHN HENRY (1967)

Juan Comesaña
Assistant Professor, Philosophy, University of Wisconsin-Madison
PYRRHONIAN PROBLEMATIC, THE (2005)

Earl Conee
Professor of Philosophy, University of Rochester, NY
EPISTEMOLOGY (1996, 2005)
EVIDENTIALISM (1996)

James I. Conway
Chair, Philosophy, Fordham University, New York, NY (retired)
MARÉCHAL, JOSEPH (1967)

Roy T. Cook
Visiting Professor, Philosophy; Villanova University, and Associate Research Fellow, Arché: The AHRC Centre for the Philosophy of Logic, Language, Mathematics, and Mind, University of St. Andrews
INFINITY IN MATHEMATICS AND LOGIC (2005)

John M. Cooper
Stuart Professor of Philosophy, Princeton University
OWEN, G. E. L. (2005)

David Copp
University of California, Davis
MORAL SKEPTICISM (1996)

Henry Corbin
Professor of Islamism at the École des Hautes Études, University of Paris at the Sorbonne; Director of the Department of Iranology, Institut franco-iranien (Tehran)
AL-GHAZĀLĪ, MUHAMMAD (1967)
IBN BĀJJA (1967)
IBN ṬUFAYL (1967)

SUHRAWARDĪ, SHIHĀB AL-DĪN YAḤYĀ (1967)

John Corcoran
Professor of Philosophy, University of Buffalo, State University of New York at Buffalo
BOOLE, GEORGE (2005)
LOGIC, HISTORY OF: MODERN LOGIC: THE BOOLEAN PERIOD: BOOLE (2005)

Michael Corrado
Allen Professor of Law and Professor of Philosophy, University of North Carolina at Chapel
POSNER, RICHARD (2005)

Gerald R. Cragg
Professor of Church History, Andover Newton Theological School
LAW, WILLIAM (1967)
MELANCHTHON, PHILIPP (1967)

William Lane Craig
Research Professor of Philosophy, Talbot School of Theology, Biola University
COSMOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD (2005)

Maurice Cranston
Reader in Political Science, University of London
BURKE, EDMUND (1967)
FASCISM (1967)
LIBERALISM (1967)
MONTESQUIEU, BARON DE (1967)
TOLERATION (1967)

Richard Creath
Arizona State University
VERIFIABILITY PRINCIPLE [ADDENDUM] (1996)

M. J. Cresswell
Professor of Philosophy, The University of Auckland and Texas A&M University
SEMANTICS, HISTORY OF [ADDENDUM] (2005)

Stephen D. Crites
Assistant Professor of Religion, Wesleyan University
BAUER, BRUNO (1967)
MEGARIANS (2005)
ROSENKRANZ, JOHANN KARL FRIEDRICH (1967)

L. G. Crocker
Dean, Graduate School, and W. G. Leutner Distinguised Professor of Romance Languages, Western Reserve University
BONNET, CHARLES (1967)
CABANIS, PIERRE-JEAN GEORGES (1967)
NAIGEON, JACQUES-ANDRÉ (1967)
ROBINET, JEAN-BAPTISTE-RENÉ (1967)
SAINT-HYACINTHE, THÉMISEUL DE (1967)
VAUVENARGUES, LUC DE CLAPIERS, MARQUIS DE (1967)
VOLNEY, CONSTANTIN-FRANÇOIS DE CHASSEBOEUF, COMTE DE (1967)

Richard Cross
Fellow and Tutor in Theology, Oriel College, University of Oxford
DUNS SCOTUS, JOHN [ADDENDUM] (2005)

Troy Cross
Assistant Professor of Philosophy, Yale University
DETERMINABLES AND DETERMINATES [ADDENDUM] (2005)

Antonio S. Cua
Professor Emeritus, Philosophy, Catholic University of America
WANG YANG-MING (2005)
XUNZI (2005)

Ann E. Cudd
Professor of Philosophy and Women's Studies, Director of Women's Studies, University of Kansas, Lawrence
ANALYTIC FEMINISM (1996, 2005)
FRYE, MARILYN (2005)

Philip W. Cummings
Lecturer in Philosophy at City University of New York, Hunter College
KÖHLER, WOLFGANG (1967)
POLITICAL PHILOSOPHY, HISTORY OF (1967)
RACISM (1967)

Phillip D. Cummins
Assistant Professor, Philosophy, University of Iowa
LE CLERC, JEAN (1967)

Patricia Curd
Professor of Philosophy, Purdue University
PARMENIDES OF ELEA [ADDENDUM] (2005)

Edwin Curley
James B. and Grace J. Nelson Professor of Philosophy, University of Michigan, Ann Arbor
DESCARTES, RENÉ (2005)
SPINOZA, BENEDICT (BARUCH) DE (2005)

Randall Curren
Professor of Philosophy and Professor of Education, University of Rochester
PHILOSOPHY OF EDUCATION, HISTORY OF: CONTEMPORARY ISSUES: ETHICAL AND POLITICAL (2005)

Haskell B. Curry
Evan Pugh Research Professor, University of Pennsylvania
COMBINATORY LOGIC (1967)

Fred D'Agostino
University of New England, Armidale, Australia
CHOMSKY, NOAM (1996)

Daniel O. Dahlstrom
Professor of Philosophy, Boston University
SCHILLER, FRIEDRICH (2005)
SCHLEGEL, FRIEDRICH VON (2005)

Hans Daiber
Universitätsprofessor und Direktor des Orientalischen Seminars, Universität Frankfurt, Germany
PYTHAGORAS AND PYTHAGOREANISM [ADDENDUM 2] (2005)

Richard C. Dales
Associate Professor, History, University of Southern California
GROSSETESTE, ROBERT (1967)
PSEUDO-GROSSETESTE (1967)
THOMAS OF YORK (1967)

Arleen B. Dallery
La Salle University
FEMINIST PHILOSOPHY OF SCIENCE (1996)

Jonathan Dancy
Professor of Philosophy, The University of Reading, U.K., and The University of Texas at Austin
INTUITIONISM, ETHICAL (2005)

Arthur C. Danto
Associate Professor of History, Columbia University
PHILOSOPHY OF SCIENCE, PROBLEMS OF (1967)

B. A. Dar
Director, Iqbal Academy (Karachi), Managing Editor of the Pakistan Philosophical Journal
IQBAL, MUHAMMAD (1967)

Stephen Darwall
University of Michigan, Ann Arbor
RATIONALISM IN ETHICS (PRACTICAL-REASON APPROACHES) (1996)

Joseph W. Dauben
Distinguished Professor of History and History of Science, Department of History, Herbert H. Lehman College, City University of New York (CUNY), and PhD Program in History, The Graduate Center, CUNY
CANTOR, GEORG (2005)

George E. Davie
Reader in Logic and Metapysics, Philosophy, University of Edinburgh
BAIN, ALEXANDER (1967)
FERRIER, JAMES FREDERICK (1967)
HODGSON, SHADWORTH HOLLOWAY (1967)
MANSEL, HENRY LONGUEVILLE (1967)

David Davies
Associate Professor of Philosophy, McGill University
ART, STYLE AND GENRE IN (2005)

Martin Davies
Professor of Philosophy, Research School of Social Sciences, Australian National University
EVANS, GARETH (2005)
LANGUAGE AND THOUGHT (1996)
LANGUAGE OF THOUGHT (1996)
MEANING (1996)

Stephen Davies
Associate Professor, Philosophy, University of Auckland
ART, PERFORMANCE IN (2005)

Martin Davis
New York University
INFINITESIMALS (1996)

Michael Davis
Professor of Philosophy, Humanties Department; Senior Fellow, Center for the Study of Ethics in the Professions, Illinois Institute of Technology
PUNISHMENT [ADDENDUM] (2005)

P. H. DeLacy
Professor, Classics, Cornell University
CICERO, MARCUS TULLIUS (1967)
EPICUREANISM AND THE EPICUREAN SCHOOL (1967)

Alain de Libera
Professoresseur ordinaire, Philosophie médiévale, Université de Genève; Directeur d'études, Histoire des théologies chrétiennes dans l'Occident médiéval, École pratique des Hautes Études, Section des sciences religieuses, Paris
AVERROISM [BIBLIOGRAPHY] (2005)

Julien Deonna
Assistant Professor, Philosophy, University of Lausanne
SYMPATHY AND EMPATHY (2005)

Michael DePaul
Professor, Philosophy, University of Notre Dame
MORAL EPISTEMOLOGY (1996, 2005)
REFLECTIVE EQUILIBRIUM (2005)

William H. Desmonde
Research Staff Member of the IBM Corp.; Lecturer, New School for Social Research
MEAD, GEORGE HERBERT (1967)

Penelope Deutscher
Associate Professor, Philosophy, Northwestern University
IRIGARAY, LUCE (2005)

Mary Devereaux
Ethicist, Research Ethics Program, University of California, San Diego
BEAUTY [BIBLIOGRAPHY] (2005)
UGLINESS [BIBLIOGRAPHY] (2005)

Garrett DeWeese
Associate Professor of Philosophy, Biola University
IMMORTALITY [ADDENDUM] (2005)

P. Diamandopoulos
Dean of Faculty and Associate Professor, Philosophy, Brandeis University
ANAXIMENES (1967)

Cora A. Diamond
University of Virginia, Charlottesville
ANSCOMBE, GERTRUDE ELIZABETH MARGARET (1996, 2005)
WITTGENSTEIN, LUDWIG JOSEF JOHANN [ADDENDUM 1] (1996)

George Dickie
Professor Emeritus, Philosophy, University of Illinois, Chicago
BEARDSLEY, MONROE C. (2005)

Alfred Di Lascia
Associate Professor, Philosopy, Manhattan College
STURZO, LUIGI (1967)

Frank B. Dilley
Emeritus Professor, Philosophy, University of Delaware, Newark
PARAPSYCHOLOGY (2005)

John Dillon
Regius Professor of Greek, School of Classics, Trinity College, Dublin
ALCINOUS (2005)
IAMBLICHUS (2005)

Martin Dillon
Professor of Philosophy, Binghamton University, State University of New York
MERLEAU-PONTY, MAURICE (2005)

Robin S. Dillon
Associate Professor, Philosophy, Lehigh University
RESPECT (2005)

John Dilworth
Professor, Philosophy, Western Michigan University
LANGER, SUSANNE K. (2005)

John Divers
Professor of Philosophy, University of Sheffield
FICTIONALISM (2005)

Zoltan Domotor
Professor of Philosophy, Biochemistry and Biophysics, University of Pennsylvania
MEASUREMENT AND MEASUREMENT THEORY (2005)

Alan Donagan
Professor, Philosophy, University of Illinois
COLLINGWOOD, ROBIN GEORGE (1967)

Thomas Donaldson
Mark O. Winkelman Professor, The Wharton School, and Professor of Philosophy, University of Pennsylvania
BUSINESS ETHICS (2005)

Willis Doney
Associate Professor, Philosophy, Dartmouth College
CARTESIANISM (1967)
GEULINCX, ARNOLD (1967)

Keith S. Donnelan
Associate Professor, Philosophy, Cornell University
PARADIGM-CASE ARGUMENT (1967)

Gerald Doppelt
Professor of Philosophy and Sciences Studies, University of California, San Diego
NATIONALISM [ADDENDUM] (2005)
SCIENTIFIC REVOLUTIONS (2005)

Douglas F. Dowd
Professor and Chair, Economics, Cornell University
VEBLEN, THORSTEIN BUNDE (1967)

Stillman Drake
Municipal Finance Consultant, San Francisco
GALILEO GALILEI (1967)

William H. Dray
Professor, Philosophy, University of Toronto
DETERMINISM IN HISTORY (1967)
HOLISM AND INDIVIDUALISM IN HISTORY AND SOCIAL SCIENCE (1967)
SPENGLER, OSWALD (1967)

James Dreier
Brown University
PROJECTIVISM (1996)

John Driscoll
Former Lecturer in Philosophy, San Francisco State University
OUSIA (2005)

Julia Driver
Professor, Philosophy, Dartmouth College
MORAL PSYCHOLOGY (2005)

John Drummond
Professor, Philosophy, Fordham University
HUSSERL, EDMUND (2005)

James Duerlinger
Professor of Philosophy, University of Iowa
VASUBANDHU (2005)

Timothy J. Duggan
Associate Professor and Chair, Philosophy, Dartmouth College
HAMILTON, WILLIAM (1967)

Daniel Dumouchel
Associate Professor, Philosophy, Université de Montréal
BATTEUX, ABBÉ CHARLES (2005)
BOILEAU, NICOLAS (2005)
DUBOS, ABBE JEAN BAPTISTE (2005)
GOTTSCHED, JOHANN CHRISTOPH (2005)

Harold B. Dunkel
Professor, Education, University of Chicago
HERBART, JOHANN FRIEDRICH (1967)

John D. Dunne
Assistant Professor, Languages and Cultures of Asia, University of Wisconsin, Madison
BUDDHIST EPISTEMOLOGY (2005)

John Dupré
Professor of Philosophy of Science and Director, ESRC Centre for Genomics in Society, University of Exeter
CARTWRIGHT, NANCY (2005)
NATURAL KINDS (2005)

Detlef Durr
Mathematisches Institut, Ludwig-Maximilians-Universität, München
BELL, JOHN, AND BELL'S THEOREM (2005)
BOHM, DAVID (2005)
BOHMIAN MECHANICS (2005)

Gerald Dworkin
Professor of Philosophy, University of California, Davis
LIBERTY (1996)
PATERNALISM [BIBLIOGRAPHY] (2005)

William James Earle
Instructor, Philosophy, Long Island University
JAMES, WILLIAM (1967)

Lloyd Easton
Professor and Chair, Philosophy, Ohio Wesleyan University
HARRIS, WILLIAM TORREY (1967)

A. W. Eaton
Assistant Professor, Philosophy, University of Illinois, Chicago
FEMINIST AESTHETICS AND CRITICISM (2005)

Marcia Muelder Eaton
Professor, Philosophy, University of Minnesota
AESTHETIC EXPERIENCE (2005)

Julius Ebbinghaus
Professor ordinarius Emeritus, University of Marburg
COHEN, HERMANN (1967)

Gary Ebbs
Professor, Philosophy, University of Illinois at Urbana-Champaign
RULE FOLLOWING [ADDENDUM] (2005)

Christopher J. Eberle
Associate Professor, Philosophy, United States Naval Academy
RELIGION AND POLITICS (2005)

Theodor Ebert
Professor, Philosophy, Universität Erlangen-Nürnberg
DIODORUS CRONUS (2005)
PHILO OF MEGARA (2005)

Ludwig Edelstein
Professor of the History of Science and Philsophy, The Rockefeller Institute
POSIDONIUS (1967)

Paul Edwards
Associate Professor, Philosophy, City University of New York, Brooklyn College; Lecturer, Philosophy, New School for Social Research
ATHEISM (1967)
ATHEISMUSSTREIT (1967)
COMMON CONSENT ARGUMENTS FOR THE EXISTENCE OF GOD (1967)
LIFE, MEANING AND VALUE OF (1967)
PANPSYCHISM (1967)
POPPER-LYNKEUS, JOSEF (1967)
REICH, WILHELM (1967)
WHY (1967)

Frances Egan
Associate Professor, Philosophy and Center for Cognitive Science, Rutgers University
COMPUTATIONALISM (2005)

Philip Ehrlich
Professor, Philosophy, Ohio University
CONTINUITY (2005)

Nader El-Bizri
Research Associate in Philosophy, Institute of Ismaili Studies, London; Affiliated Research Scholar, History and Philosophy of Science, University of Cambridge
IKHWĀN AL-ṢAFĀ' (2005)

Richard Eldridge
Charles and Harriett Cox McDowell Professor of Philosophy and Religion, Swarthmore College
HÖLDERLIN, JOHANN CHRISTIAN FRIEDRICH (2005)

Bernard Elevitch
Assistant Professor, Philosophy, University of Massachusetts, Boston
BRUNSCHVICG, LÉON (1967)

Catherine Elgin
Professor of the Philosophy of Education, Graduate School of Education, Harvard University
GOODMAN, NELSON [ADDENDUM] (1996, 2005)

Mircea Eliade
Sewell L. Avery Distinguished Service Professor of History of Religions and Professor of the Committee on Social Thought, University of Chicago
IONESCU, NAE (1967)
RĂDULESCU-MOTRU, CONSTANTIN (1967)

Charles Elkan
University of California, San Diego
FUZZY LOGIC (1996)

George F. R. Ellis
Professor Emeritus, Mathematics, University of Cape Town, Cape Town
COSMOLOGY [ADDENDUM] (2005)

Reinaldo Elugardo
Professor, Philosophy, University of Oklahoma, Norman
BAKER, LYNNE RUDDER (2005)

Lester Embree
William F. Dietrich Eminent Scholar in Philosophy, Florida Atlantic University; and President, the Center for Advanced Research in Phenomenology, Inc.
CAIRNS, DORION (2005)

Caryl Emerson
A. Watson Armour III University Professor of Slavic Languages and Literatures, Princeton University
BAKHTIN, MIKHAIL MIKHAILOVICH [ADDENDUM] (2005)

Steven M. Emmanuel
Virginia Wesleyan College
KIERKEGAARD, SØREN AABYE [ADDENDUM] (1996)

James Antony Emmen, O.F.M.
Member of the Theological Section, Collegio di S. Bonaventura, Franciscan International College of Research, Quaracchi, Italy
MATTHEW OF ACQUASPARTA (1967)
PETER AUREOL (1967)

Dorothy M. Emmet
Sir Samuel Hall Professor of Philosophy, University of Manchester
ALEXANDER, SAMUEL (1967)
FUNCTIONALISM IN SOCIOLOGY (1967)
WHITEHEAD, ALFRED NORTH (1967)

Herbert B. Enderton
Adjunct Professor, Mathematics, University of California, Los Angeles
COMPUTABILITY THEORY (2005)
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL: TURING AND COMPUTABILITY THEORY (2005)

Ronald Endicott
Department of Philosophy and Religion, Program Director for Cognitive Science, North Carolina State University, Raleigh
MULTIPLE REALIZABILITY (1996, 2005)

Edward Erwin
Professor of Philosophy, University of Miami
PSYCHOANALYSIS (2005)

Michael Esfeld
Full Professor of Epistemology and Philosophy of Science, University of Lausanne
HOLISM AND INDIVIDUALISM IN HISTORY AND SOCIAL SCIENCE [ADDENDUM] (2005)

Girard J. Etzkorn
Professor Emeritus, The Franciscan Institute, St. Bonaventure University
CHATTON, WALTER (2005)
MARSTON, ROGER (1967, 2005)
PECKHAM, JOHN (1967, 2005)
RICHARD OF MEDIAVILLA (1967)

Joseph W. Evans
Associate Professor, Philosophy, University of Notre Dame
MARITAIN, JACQUES (1967)

R. E. Ewin
Emeritus Professor of Philosophy, University of Western Australia
HOBBES, THOMAS [ADDENDUM] (1996, 2005)

Gertrude Ezorsky
Assistant Professor, Philosophy, City University of New York, Brooklyn College
PERFORMATIVE THEORY OF TRUTH (1967)

Rick Fairbanks
Philosophy, St. Olaf College
DEATH [ADDENDUM] (2005)

Eugene R. Fairweather
Keble Professor of Divinity, Trinity College, University of Toronto
CAROLINGIAN RENAISSANCE (1967)
DAVID OF DINANT (1967)
HENRY OF GHENT (1967)
ISAAC OF STELLA (1967)
PETER DAMIAN (1967)
WILLIAM OF MOERBEKE (1967)

Cynthia Farrar
Lecturer, Political Science, Yale University
THUCYDIDES (2005)

Sidney B. Fay
Professor Emeritus, Harvard University
MEINECKE, FRIEDRICH (1967)

Susan L. Feagin
Editor, The Journal of Aesthetics and Art Criticism; Research Professor, Department of Philosophy, Temple University
VISUAL ARTS, THEORY OF THE (2005)

Anita Feferman
Biographer, Independent scholar
TARSKI, ALFRED [ADDENDUM] (2005)

Solomon Feferman
Professor of Mathematics and Philosophy, Stanford University
PROOF THEORY (1996)
TARSKI, ALFRED [ADDENDUM] (2005)

Herbert Feigl
Director, Minnesota Center for Philosophy of Science, and Professor of Philosophy, University of Minnesota
MILLER, DICKINSON S. (1967)

Joel Feinberg
University of Arizona, Tucson
PATERNALISM (1996)

Richard Feldman
Professor of Philosophy, University of Rochester
EPISTEMOLOGY (1996, 2005)
EVIDENTIALISM (1996)

Otis Fellows
Professor of French Literature and Chair of Italian Department, Columbia University
BUFFON, GEORGES-LOUIS LECLERC, COMTE DE (1967)
CYRANO DE BERGERAC, SAVINIEN DE (1967)
FONTENELLE, BERNARD LE BOVIER DE (1967)
MAILLET, BENOÎT DE (1967)

M. Jamie Ferreira
Carolyn M. Barbour Chair of Religious Studies, University of Virginia, Charlottesville
NEWMAN, JOHN HENRY [ADDENDUM] (2005)

James H. Fetzer
Distinguished McKnight University Professor, Department of Philosophy, University of Minnesota, Duluth
FRAME PROBLEM (2005)
HEMPEL, CARL GUSTAV (2005)

Paul K. Feyerabend
Professor, Philosophy, University of California, Berkeley
HEISENBERG, WERNER (1967)
PLANCK, MAX (1967)
SCHULTZ, JULIUS (1967)

James Fieser
Professor, Philosophy, University of Tennessee at Martin
BEATTIE, JAMES (2005)

Vladimir Filipović
Professor and Head, Philosophy, Facuty of Arts, University of Zagreb
MARULIĆ, MARKO (1967)

Arthur Fine
Professor of Philosophy, University of Washington, Seattle
EINSTEIN, ALBERT (2005)

Stephen Finlay
Assistant Professor of Philosophy, University of Southern California
EMOTIVE THEORY OF ETHICS (2005)

Kai von Fintel
Associate Professor of Lingustics, Department of Linguistics and Philosophy, Massachusetts Institute of Technology
MODALITY AND LANGUAGE (2005)

John Fischer
Professor, Philosophy, University of California, Riverside
FRANKFURT, HARRY (2005)

Alden L. Fisher
Professor, Philosophy, St. Louis University
MERCIER, DÉSIRÉ JOSEPH (1967)

Sterling Fishman
Assistant Professor, Department of History and Educational Policy Studies, University of Wisconsin
LASSALLE, FERDINAND (1967)

Anthony Flew
Professor, Philosophy, University of Keele
IMMORTALITY (1967)
MALTHUS, THOMAS ROBERT (1967)
MIDDLETON, CONYERS (1967)
MIRACLES (1967)
PRECOGNITION (1967)

Sten G. Flygt
Professor of German, Vanderbilt University
BAHRDT, CARL FRIEDRICH (1967)
HEBBEL, CHRISTIAN FRIEDRICH (1967)

Thomas R. Flynn
Samuel Candler Dobbs Professor of Philosophy, Emory University
BAD FAITH (1996)
CONSCIOUSNESS IN PHENOMENOLOGY (2005)
EXISTENTIALISM [ADDENDUM] (1996)
EXISTENTIAL PSYCHOANALYSIS [ADDENDUM] (1996)
SARTRE, JEAN-PAUL (2005)

Robert J. Fogelin
Associate Professor, Philosophy, Pomona College
BLANSHARD, BRAND (1967)

Richard Foley
Professor of Philosophy; Ehrenkranz Dean of the Faculty of Arts and Science, New York University
CHISHOLM, RODERICK (1996, 2005)
SUBJECTIVIST EPISTEMOLOGY (1996)

Graeme Forbes
Celia Scott Weatherhead Distinguished Professor of Philosophy, Tulane University
INTENSIONAL TRANSITIVE VERBS (2005)

Peter Forrest
Professor of Philosophy, University of New England, Armidale, Australia
RELIGION, NATURALISTIC RECONSTRUCTIONS OF [ADDENDUM] (2005)

Michael Forster
Professor, Philosophy, University of Chicago
HERDER, JOHANN GOTTFRIED (2005)

Alan Fox
Associate Professor, Philosophy, University of Delaware
BUDDHISM—SCHOOLS: HUA YAN (2005)

Eli Franco
Professor, Institute for Indology and Central Asian Studies, University of Leipzig
KNOWLEDGE IN INDIAN PHILOSOPHY (2005)

Daniel H. Frank
Professor of Philosophy and Director of the Jewish Studies Program, Purdue University
MAIMONIDES [ADDENDUM] (2005)
PHILO JUDAEUS [ADDENDUM] (2005)
SAADYA (1967)
SAADYA [ADDENDUM] (2005)

Charles Frankel
Assistant U.S. Secretary of State for Educational and Cultural Affairs; Chair, Committee on Professional Ethics, American Association of University Professors
PROGRESS, THE IDEA OF (1967)

Lois Frankel
Ewing, New Jersey
CONWAY, ANNE (1996)

William K. Frankena
Professor, Philosophy, University of Michigan
VALUE AND VALUATION (1967)

Harry G. Frankfurt
Associate Professor, Rockefeller University
DOUBT (1967)

Gad Freudenthal
Permanent Senior Research Fellow, Centre National de la Recherche Scientifique (CNRS), Paris
GERSONIDES [ADDENDUM] (2005)

Gideon Freudenthal
Professor of Philosophy and History of Science, Cohn Institute for the History and Philosophy of Science and Ideas, Tel-Aviv University
MAIMON, SALOMON (2005)

R. G. Frey
Bowling Green University
SPECIESISM (1996)

Elizabeth Fricker
University Lecturer in Philosophy and Fellow, Magdalen College, Oxford University
TESTIMONY (1996, 2005)

Russell L. Friedman
Associate Professor, Philosophy, Catholic University of Leuven
DURANDUS OF SAINT-POURÇAIN [BIBLIOGRAPHY] (2005)
PETER AUREOL [ADDENDUM] (2005)

Wolfgang Friedman
Barrister at Law, Middle Temple, England; Professor of Law and Director of International Legal Research, Columbia University
GROTIUS, HUGO (1967)
RADBRUCH, GUSTAV (1967)
STAMMLER, RUDOLF (1967)

Horace L. Friess
Buttenwieser Professor of Human Relations; Member of the Departments of Philosophy and Religion at Columbia University
STEINER, RUDOLF (1967)

Robert Frodeman
Associate Professor and Chair, Philosophy and Religion Studies, University of North Texas
SCIENCE POLICY (2005)

Karin Fry
Assistant Professor, Philosophy, University of Wisconsin, Stevens Point
LYOTARD, JEAN FRANÇOIS (2005)

Northrop Frye
Principal of Victoria College, University of Toronto
BLAKE, WILLIAM (1967)

Alan Fuchs
Professor of Philosophy, College of William & Mary
RAWLS, JOHN (1996, 2005)

Richard Fumerton
F. Wendell Miller Professor of Philosophy, University of Iowa
CLASSICAL FOUNDATIONALISM (1996, 2005)
SOLIPSISM (2005)

David J. Furley
Reader in Greek and Latin, University College, University of London
HOMER (1967)
MELISSUS OF SAMOS (1967)
PARMENIDES OF ELEA (1967)

Michael Gagarin
James R. Dougherty, Jr. Centennial Professor of Classics, The University of Texas at Austin
ANTIPHON (2005)
DIKÉ (2005)

Piama Gaidenko
Russian Academy of Sciences, Institute of Philosophy
SOLOV'ËV (SOLOVYOV), VLADIMIR SERGEEVICH (2005)
TRUBETSKOI, SERGEI NIKOLAEVICH (2005)

Elizabeth Cameron Galbraith
Associate Professor of Philosophy and Religion, St. Olaf College
RAHNER, KARL (2005)

Laura L. Garcia
Adjunct Assistant Professor, Philosophy, Boston College
TELEOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD [ADDENDUM] (2005)

Patrick Gardiner
Fellow and Tutor in Philosophy, Magdalen College, Oxford University
BUCKLE, HENRY THOMAS (1967)
BURCKHARDT, JAKOB (1967)
IRRATIONALISM (1967)
SAINT-SIMON, CLAUDE-HENRI DE ROUVROY, COMTE DE (1967)
SCHOPENHAUER, ARTHUR (1967)
TOYNBEE, ARNOLD JOSEPH (1967)

Martin Gardner
Editor and Writer, Mathematical Games Department of Scientific American
LOGIC DIAGRAMS (1967)
LOGIC MACHINES (1967)

Sebastian Gardner
Professor of Philosophy, University College, London
ROMANTICISM [ADDENDUM] (2005)

Jay Garfield
Doris Silbert Professor in the Humanities and Professor of Philosophy, Smith College; Professor of Philosophy, University of Melbourne; Adjunct Professor of Philosophy, Central Institute of Higher Tibetan Studies
NĀGĀRJUNA (2005)

Eugenio Garin
Ordinary Professor of the History of Philosophy, Faculty of Letters and Philosophy, University of Florence
BANFI, ANTONIO (1967)
BONATELLI, FRANCESCO (1967)
GALLUPPI, PASQUALE (1967)
GENOVESI, ANTONIO (1967)
PASTORE, VALENTINO ANNIBALE (1967)
ROMAGNOSI, GIAN DOMENICO (1967)
VANINI, GIULIO CESARE (1967)

Richard T. Garner
Professor of Philosophy, Ohio State University
NONCOGNITIVISM (1996)
STEVENSON, CHARLES L. (1996, 2005)

Aaron Garrett
Associate Professor, Philosophy, Boston University
FERGUSON, ADAM (2005)

James W. Garson
Professor, Philosophy, University of Houston
MODALITY AND QUANTIFICATION (2005)

Newton Garver
State University of New York Distinguished Service Professor, University at Buffalo
BLACK, MAX (1967, 2005)
SUBJECT AND PREDICATE (1967)

Stephen Gaukroger
Professor of History of Philosophy and History of Science, University of Sydney
BACON, FRANCIS (2005)

Deno J. Geanakoplos
Professor of Medieval and Byzantine History, University of Illinois; Coeditor, Greek, Roman, Byzantine Studies
PLETHO, GIORGIUS GEMISTUS (1967)

R. Douglas Geivett
Professor of Philosophy, Talbot Department of Philosophy, Biola University
MIRACLES [ADDENDUM] (2005)

Hester Goodenough Gelber
Associate Professor, Religious Studies, Stanford University
HOLKOT, ROBERT (2005)

Tamar Szabo Gendler
Associate Professor of Philosophy, Cornell University
IMAGINATION [ADDENDUM] (2005)
THOUGHT EXPERIMENTS IN SCIENCE (2005)

Charles Genequand
Professor, Faculté des lettres, University of Geneva
ALEXANDER OF APHRODISIAS [ADDENDUM] (2005)

Jean-François Genest
Chercheur au Centre National de la Recherche Scientifique, Institut de Recherche et d'Histoire des Textes, Paris
BRADWARDINE, THOMAS (2005)

William Gerber
Economist, U.S. Department of Labor; Associate Professor, Philosophy, University of Maryland
TAGORE, RABINDRANATH (1967)

Hanna-Barbara Gerl-Falkovitz
Chair, Philosophy of Religion and Comparative Religious Studies, Technical University of Dresden
STEIN, EDITH (2005)

B. A. Gerrish
Associate Professor of Historical Theology, Divinity School, University of Chicago
LUTHER, MARTIN (1967)
REFORMATION (1967)

Stephen E. Gersh
Professor, Medieval Institute, and Concurrent Professor, Philosophy, University of Notre Dame
PLATONISM AND THE PLATONIC TRADITION (2005)

Bernard Gert
Stone Professor of Intellectual and Moral Philosophy, Dartmouth College; Adjunct Professor of Psychiatry, Dartmouth Medical School
BAIER, KURT (2005)
EUTHANASIA (1996, 2005)
IMPARTIALITY (1996, 2005)

Brie Gertler
Philosophy, University of Virginia
KNOWLEDGE ARGUMENT (2005)

Alan Gewirth
Professor, Philosophy, University of Chicago
MARSILIUS OF PADUA (1967)

John Gibbons
Assistant Professor of Philosophy, University of Nebraska–Lincoln
KNOWLEDGE, THE PRIORITY OF (2005)

Roger F. Gibson
Washington University, St. Louis
QUINE, WILLARD VAN ORMAN [ADDENDUM] (1996)

Ronald Giere
Professor of Philosophy, University of Minnesota, Twin Cities
NATURALIZED PHILOSOPHY OF SCIENCE (2005)

Felix Gilbert
Professor, School of Historical Studies, Institute for Advanced Study, Princeton, NJ
MACHIAVELLI, NICCOLÒ (1967)

Neal W. Gilbert
Professor, School of Historical Studies, Institute for Advanced Study, Princeton, NJ
JUNGIUS, JOACHIM (1967)
MAJOR, JOHN (1967)
RENAISSANCE (1967)
VALLA, LORENZO (1967)
VIVES, JUAN LUIS (1967)

Thomas Gilby, O.P.
S.T.M., PhD, Blackfriars, Cambridge University
THOMISM (1967)

Carl Gillett
Associate Professor, Philosophy, Illinois Wesleyan University
SPECIAL SCIENCES (2005)

Brendan S. Gillon
Associate Professor, Linguistics, McGill University
LOGIC, HISTORY OF: LOGIC AND INFERENCE IN INDIAN PHILOSOPHY (2005)
PLURALS AND PLURALITY (2005)

Richard C. Gilman
President, Occidental College, Los Angeles
HOCKING, WILLIAM ERNEST (1967)

Carl Ginet
Professor of Philosophy Emeritus, Cornell University
VOLITION (2005)

James Ginther
Associate Professor of Medieval Theology, St. Louis University
GROSSETESTE, ROBERT [ADDENDUM] (2005)

Michael Glanzberg
Associate Professor of Philosophy, University of California, Davis
QUANTIFIERS IN NATURAL LANGUAGE (2005)

Nahum Norbert Glatzer
Professor of Jewish History, Chair, Near Eastern and Judaic Studies, Brandeis University
ROSENZWEIG, FRANZ (1967)

Lydia Goehr
Professor of Philosophy, Columbia University
ADORNO, THEODOR WIESENGRUND (2005)
BENJAMIN, WALTER (2005)

Ludmila Gogotishvili
Senior Research Associate, Russian Academy of Sciences, Institute of Philosophy
BAKHTIN, MIKHAIL MIKHAILOVICH (2005)

Sanford Goldberg
Associate Professor of Philosophy and Director of Cognitive Science, University of Kentucky
PROPOSITIONAL ATTITUDES: ISSUES IN PHILOSOPHY OF MIND AND PSYCHOLOGY (2005)

Joshua L. Golding
Associate Professor of Philosophy, Bellarmine University
FAITH [ADDENDUM] (2005)

M. P. Golding
Associate Professor, Philosophy, Columbia University
PHILOSOPHY OF LAW, HISTORY OF (1967)

Alan H. Goldman
William R. Kenan, Jr. Professor of Humanities, Department of Philosophy, College of William & Mary
AESTHETIC QUALITIES (2005)
CAUSAL OR CONDITIONAL OR EXPLANATORY-RELATION ACCOUNTS (1996)

Alvin Goldman
Board of Governors Professor, Department of Philosophy, Rutgers, The State University of New Jersey
SIMULATION THEORY (2005)

Sheldon Goldstein
Professor of Mathematics, Rutgers University
BELL, JOHN, AND BELL'S THEOREM (2005)
BOHM, DAVID (2005)
BOHMIAN MECHANICS (2005)

Herman H. Goldstine
Director of Scientific Development, Data Processing Division, IBM
NEUMANN, JOHN VON (1967)

Michael F. Goodman
Professor, Philosophy, Humboldt State University
PERSONS (2005)

Russell B. Goodman
Professor of Philosophy, University of New Mexico, Albuquerque
NEW ENGLAND TRANSCENDENTALISM [ADDENDUM] (2005)

Robert M. Gordon
Research Professor in Philosophy of Mind and Cognitive Science, University of Missouri, St. Louis
EMOTION (2005)

Eva Gossman
Lecturer in Philosophy, Goucher College (Towson, MD)
FRANK, ERICH (1967)

Rubin Gotesky
Lecturer, Philosophy, Goucher College (Towson, MD)
CARUS, CARL GUSTAV (1967)
EUCKEN, RUDOLF CHRISTOPH (1967)
LOTZE, RUDOLF HERMANN (1967)

Roger S. Gottlieb
Professor of Philosophy, Department of Humanities and Arts, Worcester Polytechnic Institute
MARXIST PHILOSOPHY [ADDENDUM] (1996, 2005)

T. A. Ġoudge
Chair, Philosophy, University of Toronto; Fellow of the Royal Society of Canada
BERGSON, HENRI (1967)
BERTALANFFY, LUDWIG VON (1967)
BUTLER, SAMUEL (1967)
DARWIN, CHARLES ROBERT (1967)
DARWIN, ERASMUS (1967)

GRAY, ASA (1967)
HUXLEY, THOMAS HENRY (1967)
LAMARCK, CHEVALIER DE (1967)
MORGAN, C. LLOYD (1967)
SMUTS, JAN CHRISTIAAN (1967)
TEILHARD DE CHARDIN, PIERRE (1967)
WALLACE, ALFRED RUSSEL (1967)
WOODGER, JOSEPH HENRY (1967)

Josiah B. Gould Jr.
Assistant Professor and Chair, Philosophy, Claremont Graduate School (Claremont, CA)
CHRYSIPPUS (1967)

Jorge Gracia
Samuel P. Capen Chair and State University of New York Distinguished Professor, Philosophy, State University of New York at Buffalo
HISTORY AND HISTORIOGRAPHY OF PHILOSOPHY (2005)
LATIN AMERICAN PHILOSOPHY (2005)

A. C. Graham
Lecturer in Chinese, School of Oriental and African Studies, University of London
LOGIC, HISTORY OF: CHINESE LOGIC (1967)

Daniel W. Graham
Abraham Owen Smoot Professor of Philosophy, Brigham Young University
ANAXAGORAS OF CLAZOMENAE (2005)
ARCHĒ (2005)
LOGOS (2005)

Gordon Graham
Henry Luce III Professor of Philosophy and the Arts, Princeton Theological Seminary
ART, VALUE IN (2005)

Richard E. Grandy
McManis Professor, Philosophy and Cognitive Sciences, Rice University
GRICE, HERBERT PAUL (1996, 2005)

Herbert Granger
Professor, Philosophy, Wayne State University
HERACLITUS OF EPHESUS (2005)

Robert M. Grant
Professor of New Testament and Early Christianity, Divinity School, University of Chicago
APOLOGISTS (1967)
CELSUS (1967)
EUSEBIUS (1967)
NEMESIUS OF EMESA (1967)
ORIGEN (1967)
PATRISTIC PHILOSOPHY (1967)
TERTULLIAN, QUINTUS SEPTIMIUS FLORENS (1967)

S. A. Grave
Professor of Philosophy, University of Western Australia
BROWN, THOMAS (1967) (1967)
COMMON SENSE (1967)

Margaret Graver
Associate Professor, Classics, Dartmouth College
SENECA, LUCIUS ANNAEUS (2005)

Joseph Grcic
Associate Professor, Philosophy, Indiana State University
LIBERALISM [ADDENDUM] (2005)

John Greco
Associate Professor, Philosophy, Fordham University
INTERNALISM VERSUS EXTERNALISM (1996, 2005)
VIRTUE EPISTEMOLOGY (1996, 2005)

Karen Green
Head of School, School of Philosophy & Bioethics, Monash University
LLOYD, GENEVIEVE (2005)

Michael Griffin
Visiting Assistant Professor, Department of Philosophy, Central European University
MOLINA, LUIS DE (2005)

A. Phillips Griffiths
Professor of Philosophy, University of Warwick
MORAL PRINCIPLES: THEIR JUSTIFICATION (1967)

Ronald Grimsley
Professor, French, University of Bristol
ROUSSEAU, JEAN-JACQUES (1967)

Peter Groff
Assistant Professor of Philosophy, Bucknell University
DIALECTIC IN ISLAMIC AND JEWISH PHILOSOPHY (2005)

Adolf Grünbaum
Andrew Mellon Professor of Philosophy of Science, Research Professor of Psychiatry, Chairman, Center for Philosophy of Science, University of Pittsburgh
FREUD, SIGMUND (2005)

Anil Gupta
Indiana University, Bloomington
LIAR PARADOX, THE (1996)

Bina Gupta
Curators' Professor, Professor of Philosophy; Director, South Asian Studies Program, University of Missouri
BRAHMAN (2005)

W. K. C. Guthrie
Laurence Professor of Ancient Philosophy and Master of Downing College, Cambridge University
PRE-SOCRATIC PHILOSOPHY (1967)
PYTHAGORAS AND PYTHAGOREANISM (1967)

Paul Guyer
Florence R. C. Murray Professor in the Humanities, University of Pennsylvania
AESTHETICS, HISTORY OF [ADDENDUM] (2005)
BULLOUGH, EDWARD (2005)
CATEGORICAL IMPERATIVE (2005)
MORITZ, KARL PHILIPP (2005)

Susan Haack
University of Miami
PRAGMATISM [ADDENDUM] (1996)
PRAGMATIST EPISTEMOLOGY (1996)

Alexander Haardt
Professor and Doctor of Philosophy, Institute of Philosophy (Department of Philosophy of Modern Times), Ruhr Universität Bochum, Germany
SHPET, GUSTAV GUSTAVOVICH (2005)

Jeremiah Hackett
Professor and Chair, Philosophy, University of South Carolina, Columbia
BACON, ROGER [ADDENDUM] (2005)

Adrian Haddock
Lecturer, Philosophy, University of Stirling
NATURAL KINDS (2005)

Garry Hagberg
James H. Ottaway Jr. Professor of Philosophy and Aesthetics, Bard College
WITTGENSTEIN, LUDWIG JOSEF JOHANN [ADDENDUM 2] (2005)

Alan Hájek
Professor of Philosophy, Research School of the Social Sciences, Australian National University
CHANCE (2005)

Roland Hall
Lecturer, Philosophy, University of St. Andrews; Assistant Editor, Philosophical Quarterly
DIALECTIC (1967)
MONISM AND PLURALISM (1967)

Morris Halle
Massachusetts Institute of Technology
PHONOLOGY (1996)

Philip P. Hallie
Griffin Professor and Chair, Philosophy, Wesleyan University
MAINE DE BIRAN (1967)

Stephen Halliwell
Professor of Greek, School of Classics, University of St Andrews
KATHARSIS (2005)
MIMESIS (2005)

G. M. Hamburg
Otho M. Behr Professor of History, Claremont McKenna College
CHICHERIN, BORIS NIKOLAEVICH [ADDENDUM] (2005)

D. W. Hamlyn
Professor of Philosophy, Birkbeck College, University of London
ANALYTIC AND SYNTHETIC STATEMENTS (1967)
A PRIORI AND A POSTERIORI (1967)
EMPIRICISM (1967)
EPISTEMOLOGY, HISTORY OF (1967)

Hassan Hanafi
Professor of Philosophy, Cairo University
LAROUI, ABDULLAH (2005)

Roger Hancock
Assistant Professor, Philosophy, University of Missouri
METAPHYSICS, HISTORY OF (1967)

Michael Hand
Professor of Philosophy, Texas A&M University
DUMMETT, MICHAEL ANTHONY EARDLEY (1996, 2005)

Rollo Handy
Professor and Chair, Philosophy; Chair, Division of Philosophy and the Social Sciences, State University of New York, Buffalo
HAECKEL, ERNST HEINRICH (1967)
MOLESCHOTT, JACOB (1967)
VAIHINGER, HANS (1967)

R. J. Hankinson
Professor of Philosophy and Classics, University of Texas at Austin
AENESIDEMUS (2005)
AGRIPPA (2005)
AITIA (2005)
HIPPOCRATES AND THE HIPPOCRATIC CORPUS (2005)
IMPETUS (2005)
PYRRHO (2005)
SEXTUS EMPIRICUS (2005)
TIMON OF PHLIUS (2005)

Peter Hanks
Assistant Professor, Philosophy, University of Minnesota, Twin Cities
PROPOSITIONS [ADDENDUM] (2005)
QUESTIONS (2005)

Chad Hansen
Chair Professor of Chinese Philosophy, Department of Philosophy, University of Hong Kong
CHINESE PHILOSOPHY: DAOISM (2005)

Norwood Russell Hanson
Professor, Philosophy, Yale University
COPERNICUS, NICOLAS (1967)

Valerie Gray Hardcastle
Associate Dean, College of Liberal Arts and Human Sciences; Professor and Head, Department of Science and Technology in Society, Virginia Tech
ELIMINATIVE MATERIALISM, ELIMINATIVISM (1996)
PAIN (2005)

Stevan Harnad
Canada Research Chair in Cognitives Sciences, Université du Québec à Montréal
CHINESE ROOM ARGUMENT (2005)

Robert M. Harnish
Professor of Philosophy and Linguistics and Research; Professor of Cognitive Science, University of Arizona Tuscon
SEARLE, JOHN (2005)

Vicki Harper
Assistant Professor of Philosophy, St. Olaf College
IONESCU, NAE [BIBLIOGRAPHY] (2005)
STRAWSON, PETER FREDERICK [BIBLIOGRAPHY] (2005)

William L. Harper
Professor, Philosophy, University of Western Ontario
NEWTON, ISAAC (2005)
SCIENTIFIC METHOD (2005)

R. Harre
Fellow of Linacre College, Oxford University, and University Lecturer, Philosophy of Science
LAPLACE, PIERRE SIMON DE (1967)

Karsten Harries
Assistant Professor, Philosophy, Yale University
KEYSERLING, HERMANN ALEXANDER, GRAF VON (1967)
KLEIST, HEINRICH VON (1967)
NOVALIS (1967)
SOLGER, KARL WILHELM FERDINAND (1967)

H. S. Harris
Professor, Philosophy, Glendom College, York University, Toronto
CROCE, BENEDETTO (1967)
GENTILE, GIOVANNI (1967)
SPAVENTA, BERTRANDO (1967)
SPIRITO, UGO (1967)

Jonathan Harrison
Professor, Philosophy, Glendom College, York University
ETHICAL NATURALISM (1967)
ETHICAL SUBJECTIVISM (1967)

H. L. A. Hart
Professor of Jurisprudence, Oxford University
LEGAL POSITIVISM (1967)
PHILOSOPHY OF LAW, PROBLEMS OF (1967)

Klaus Hartmann
Docent, Bonn University
EHRENFELS, CHRISTIAN FREIHERR VON (1967)
SCHUPPE, ERNST JULIUS WILHELM (1967)

William Hasker
Professor Emeritus of Philosophy, Huntington College
EPISTEMOLOGY, RELIGIOUS (1996)
EPISTEMOLOGY, RELIGIOUS [ADDENDUM] (2005)

William H. Hay
Professor, Philosophy, University of Wisconsin
CARUS, PAUL (1967)
MURPHY, ARTHUR EDWARD (1967)

Richard P. Hayes
Assistant Professor, Philosophy, University of New Mexico
BUDDHISM (2005)
NIRVĀṆA (2005)

Allen P. Hazen
Lecturer, Philosophy, University of Melbourne
TYPE THEORY (2005)

P. L. Heath
Professor, Philosophy, University of Virginia
BALFOUR, ARTHUR JAMES (1967)
CARROLL, LEWIS (1967)
DE MORGAN, AUGUSTUS (1967)
EXPERIENCE (1967)
JEVONS, WILLIAM STANLEY (1967)
LOGIC, HISTORY OF: MODERN LOGIC: THE BOOLEAN PERIOD: VENN; DE MORGAN; HAMILTON; JEVONS (1967)
NOTHING (1967)
VENN, JOHN (1967)

Michael Heidelberger
Chair for Logic and Science Theory, Philosophisches Seminar, Universität Tübingen
EXPERIMENTATION AND INSTRUMENTATION (2005)

Steven Heine
Professor and Director of Asian Studies, Florida International University
DŌGEN (2005)

Susan Hekman
Professor of Political Science and Director of Graduate Humanities, University of Texas at Arlington
FEMINISM AND CONTINENTAL PHILOSOPHY (2005)

Lisa Heldke
Professor, Philosophy, Gustavus Adolphus College
FEMINISM AND PRAGMATISM (2005)

Geoffrey Hellman
Professor of Philosophy, University of Minnesota, Minneapolis
STRUCTURALISM, MATHEMATICAL (2005)

Robin F. Hendry
Senior Lecturer in Philosophy, Department of Philosophy, University of Durham
LAVOISIER, ANTOINE (2005)
PAULING, LINUS (2005)

Desmond Paul Henry
Senior Lecturer in Philosophy, University of Manchester
MEDIEVAL PHILOSOPHY (1967)

Grete Henry-Hermann
Professor, Pädagogischen Hocschule, Bremen, Germany
NELSON, LEONARD (1967)

Ronald W. Hepburn
Professor of Philosophy, University of Edinburgh
AGNOSTICISM (1967)
BULTMANN, RUDOLF (1967)
MORAL ARGUMENTS FOR THE EXISTENCE OF GOD (1967)
MYSTICISM, NATURE AND ASSESSMENT OF (1967)
NATURE, PHILOSOPHICAL IDEAS OF (1967)
RELIGIOUS EXPERIENCE, ARGUMENT FOR THE EXISTENCE OF GOD (1967)

Jennifer Herdt
Associate Professor of Theology, University of Notre Dame
CAMBRIDGE PLATONISTS (2005)
CUDWORTH, RALPH (2005)

Ulrike Heuer
Lecturer, School of Philosophy, University of Leeds
INTERNALISM AND EXTERNALISM IN ETHICS (2005)

Joh's Erich Heyde
Ordinary Professor of Philosophy, Technical University of Berlin
REHMKE, JOHANNES (1967)

John Hick
Lecturer in Philosophy of Religion, University of Cambridge
CHRISTIANITY (1967)
EVIL, THE PROBLEM OF (1967)
FAITH (1967)
OMAN, JOHN WOOD (1967)
ONTOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD (1967)
RELIGIOUS PLURALISM (1996)
REVELATION (1967)
TENNANT, FREDERICK ROBERT (1967)

Pamela Hieronymi
Assistant Professor, Philosophy, University of California, Los Angeles
FORGIVENESS (2005)

James Higginbotham
Somerville College, Oxford University
SYNTAX (1996)

Jocelyn Nigel Hillgarth
Lecturer in History, Harvard University
LULL, RAMÓN (1967)

David Hills
Acting Assistant Professor, Philosophy, Stanford University
ART, REPRESENTATION IN (2005)
METAPHOR [ADDENDUM] (2005)

Iwao Hirose
Donnelley Junior Research Fellow, University College, Oxford
SEN, AMARTYA K. (2005)

R. J. Hirst
Professor and Head, Logic, Glasgow University
ILLUSIONS (1967)
PERCEPTION (1967)
PHENOMENALISM (1967)
PRIMARY AND SECONDARY QUALITIES (1967)
REALISM (1967)
SENSA (1967)

Christopher R. Hitchcock
Professor of Philosophy, Division of Humanities and Social Sciences, California Institute of Technology
CAUSATION: PHILOSOPHY OF SCIENCE (2005)

Henry Hiż
Professor of Linguistics, University of Pennsylvania
CHWISTEK, LEON (1967)

Joshua P. Hochschild
Assistant Professor, Philosophy, Mount St. Mary's University
CAJETAN, CARDINAL (2005)

Andrew Hodges
Lecturer in Mathmatics, Wadham College, University of Oxford
TURING, ALAN M. (2005)

Wilfrid Hodges
Professor of Mathematics, Queen Mary, University of London
FIRST-ORDER LOGIC (2005)
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL: DECIDABLE AND UNDECIDABLE THEORIES; MODEL THEORY: ROBINSON; MODEL THEORY: TARSKI (2005)
MODEL THEORY (2005)

Carl Hoefer
Research Professor at ICREA and the Autonomous University of Barcelona
CHANCE (2005)
CONVENTIONALISM (2005)
HOLE ARGUMENT (2005)

Frank J. Hoffman
Associate Professor, Philosophy, West Chester University
MIND AND MENTAL STATES IN BUDDHIST PHILOSOPHY (2005)

Robert Holmes
Professor, Philosophy, University of Rochester
PEACE, WAR, AND PHILOSOPHY [ADDENDUM] (2005)

Nancy Holmstrom
Chair, Associate Professor, Philosophy, Rutgers University, Newark
FEMINIST SOCIAL AND POLITICAL PHILOSOPHY (2005)

Tze-ki Hon
Associate Professor, History, State University of New York–Geneseo
ZHOU DUNYI (2005)

Ted Honderich
Grote Professor Emeritus, University College London
DETERMINISM AND FREEDOM (1996, 2005)

Bradford W. Hooker
Professor of Moral Philosophy, University of Reading
MORAL RULES AND PRINCIPLES (2005)
UTILITARIANISM [ADDENDUM] (2005)

Vincent Hope
Former Fellow of the School of Philosophy, Psychology and Language Sciences, University of Edinburgh
STEWART, DUGALD (2005)

Burt C. Hopkins
Professor, Philosophy, Seattle University
LANDGREBE, LUDWIG (2005)

Patrick D. Hopkins
Associate Professor, Philosophy, Millsaps College
HETEROSEXISM (2005)
NATURAL LAW (2005)

Terence E. Horgan
University of Memphis
CONNECTIONISM (1996)

Irving Louis Horowitz
Professor of Sociology, Washington University
DE SANCTIS, FRANCESCO (1967)

Sergey Horujy
Director of the Institute of Synergetic Anthropology; Professor of the Institute of Philosophy of Russian Academy of Sciences; Honorary Professor of UNESCO (the Chair of Comparative Studies of Religious Traditions)
FLORENSKII, PAVEL ALEKSANDROVICH (2005)
FLOROVSKII, GEORGII VASIL'EVICH (2005)
KARSAVIN, LEV PLATONOVICH (2005)
TRUBETSKOI, EVGENII NIKOLAEVICH (2005)

Nathan Houser
Indiana University, Purdue University
PEIRCE, CHARLES SANDERS [ADDENDUM] (1996)

Daniel Howard-Snyder
Professor of Philosophy, Western Washington University
HIDDENNESS OF GOD (2005)

Bruce W. Hozeski
Chair, Department of English, Ball State University
HILDEGARD OF BINGEN (2005)

Pamela M. Huby
Reader in Philosophy (Retired), University of Liverpool
AGENT INTELLECT (2005)

Carl A. Huffman
Professor of Classics, DePauw University
ALCMAEON OF CROTON (2005)
ARCHYTAS OF TARENTUM (2005)
PHILOLAUS OF CROTON (2005)

Nicholas Huggett
Associate Professor of Philosophy, University of Illinois at Chicago
BLACK HOLES (2005)
FIELDS AND PARTICLES (2005)
SPACE IN PHYSICAL THEORIES (2005)

Namjin Huh
Professor, Philosophy, Seoul National University
KOREAN PHILOSOPHY (2005)

Paul Humphreys
Professor, Corcoran Department of Philosophy, University of Virginia

EMERGENCE (2005)
SALMON, WESLEY (2005)
SUPPES, PATRICK (2005)

David P. Hunt
Professor, Philosophy, Whittier College
FOREKNOWLEDGE AND FREEDOM, THEOLOGICAL PROBLEM OF (2005)

Bruce Hunter
Professor and Head, Logic, Glasgow University
CRITERIOLOGY (1996)

Thomas Hurka
Jackman Distinguished Chair in Philosophical Studies, Philosophy, University of Toronto
INTRINSIC VALUE (2005)
MOORE, GEORGE EDWARD [ADDENDUM] (2005)
TELEOLOGICAL ETHICS (2005)

Rosalind Hursthouse
Professor of Philosophy, University of Auckland
VIRTUE ETHICS [BIBLIOGRAPHY] (2005)

Katerina Ierodiakonou
Associate Professor, Ancient Philosophy, Department of the Philosophy and History of Science, University of Athens
BYZANTINE PHILOSOPHY (2005)
PLETHO, GIORGIUS GEMISTUS [BIBLIOGRAPHY] (2005)

Shams Inati
Professor, Islamic Studies, Villanova University
DETERMINISM, THEOLOGICAL (2005)

David B. Ingram
Professor of Philosophy, Loyola University, Chicago
ARENDT, HANNAH (1996, 2005)
POSTMODERNISM (2005)

Brad Inwood
Professor of Classics and Philosophy, University of Toronto
CLEANTHES (2005)
HELLENISTIC THOUGHT (2005)
STOICISM (2005)

Anna Maria Ioppolo
Full Professor, Ancient Philosophy, Dipartimento di Scienze Filosofiche ed Epistemologiche, Università di Roma "La Sapienza"
ARISTO OF CHIOS (2005)

Michela Ippolito
Assistant Professor of Linguistics, Department of Modern Foreign Languages and Literatures, Boston University
TENSE (2005)

Howard Isham
Associate Professor, Humanities, San Francisco State College
HUMBOLDT, WILHELM VON (1967)

Frank C. Jackson
Director, Research, School of Social Sciences, Australian National University
ARMSTRONG, DAVID M. (1996, 2005)

Pierre Jacob
Director of Institut Jean Nicod, CNRS/EHESS/ENS, Paris
INTENTIONALITY [ADDENDUM] (2005)

Theordore E. James
Associate Professor, Philosophy, Manhattan College
IBN GABIROL, SOLOMON BEN JUDAH (1967)

Dale Jamieson
Professor of Environmental Studies and Philosophy, New York University, Steinhardt School, HMSS
SINGER, PETER (2005)

M. Jammer
Head, Physics; Professor of Physicas and Philosophy of Science, Bar-Ilan University, Israel
ENERGY (1967)
FORCE (1967)
MASS (1967)
MOTION, A HISTORICAL SURVEY (1967)

Richard Janko
Professor and Chair, Classical Studies, Rackham Graduate School, University of Michigan
HOMER [BIBLIOGRAPHY] (2005)

Joyce L. Jenkins
Associate Professor, Philosophy Department, University of Manitoba
SELF-INTEREST (2005)

Robert Johnson
Associate Professor of Philosophy, University of Missouri
PRACTICAL REASON (2005)

Hans Jonas
Professor, Philosophy, Graduate Faculty of Political and Social Science, New School for Social Research
GNOSTICISM (1967)

Alexander Jones
Professor, Classics and the History and Philosophy of Science and Technology, University of Toronto
HELLENISTIC THOUGHT (2005)

Charles Jones
Associate Professor, Political Science, University of Western Ontario
COSMOPOLITANISM (2005)

Karen Jones
Lecturer, Philosophy, The University of Melbourne
BAIER, ANNETTE (2005)
FEMINIST EPISTEMOLOGY (2005)

Inge Jonsson
Docent, History of Literature, University of Stockholm
SWEDENBORG, EMANUEL (1967)

Z. A. Jordan
Lecturer, Philosophy of Science, University of Reading
KOTARBIŃSKI, TADEUSZ (1967)

Lawrence J. Jost
Professor of Philosophy, University of Cincinnati
VIRTUE AND VICE (2005)

James Joyce
Associate Professor of Philosophy, University of Michigan
DECISION THEORY (2005)
SAVAGE, LEONARD (2005)

Eric T. Juengst
Associate Professor of Bioethics, School of Medicine, Case Western Reserve University

GENETICS AND REPRODUCTIVE TECHNOLOGIES [ADDENDUM] (2005)

Béla Juhos
Professor of Theological Philosophy, University of Vienna
SCHLICK, MORITZ (1967)

Elzbieta Jung
Professor, Philosophy, University of Lodz
KILVINGTON, RICHARD (2005)

Guy Kahane
Research Associate, Uehiro Centre for Practical Ethics, Faculty of Philosophy, Oxford University
PAIN, ETHICAL SIGNIFICANCE OF (2005)

Russell Kahl
Associate Professor, Philosophy, San Francisco State College
HELMHOLTZ, HERMANN LUDWIG VON (1967)

Charles H. Kahn
Professor of Philosophy, University of Pennsylvania
ANAXIMANDER (1967)
EMPEDOCLES (1967)
PLATO (2005)

Irene Kajon
Ordinary Professor, Dipartimento di Ricerche Storico-filosofiche e Pedagogiche, Università di Roma "La Sapienza"
COHEN, HERMANN [ADDENDUM] (2005)

Ibrahim Kalin
Assistant Professor of Islamic Studies, Department of Religious Studies, College of the Holy Cross
CORBIN, HENRY (2005)
EPISTEMOLOGY, HISTORY OF [ADDENDUM] (2005)
MULLĀ ṢADRĀ [ADDENDUM] (2005)
NASR, SEYYED HOSSEIN (2005)

Paul Kalligas
Assistant Professor, Philosophy and History of Science, University of Athens
PLOTINUS [BIBLIOGRAPHY] (2005)

Akihiro Kanamori
Professor, Mathematics, Boston University
SET THEORY (2005)

David Kaplan
Assistant Professor of Philosophy, Department of Philosophy and Religion Studies, University of North Texas
RICOEUR, PAUL (2005)

Elizabeth Karger
Chargée de Recherche, CNRS, Paris
WODEHAM, ADAM (2005)

George Kateb
William Nelson Cromwell Professor of Politics, Emeritus, Princeton University
UTOPIAS AND UTOPIANISM (1967, 2005)

Arnold S. Kaufman
Professor, Philosophy, Princeton University
RESPONSIBILITY, MORAL AND LEGAL (1967)

Asaf Kedar
Doctoral Student, Political Science, University of California, Berkeley
HISTORICISM [ADDENDUM] (2005)

Samuel McMurray Keen
Associate Professor of Philosophy and Christian Faith, Louisville Presbyterian Seminary
MARCEL, GABRIEL (1967)

Morris Keeton
Professor of Philosophy and Religion; Dean of the Faculty, Antioch College
MONTGOMERY, EDMUND DUNCAN (1967)

John Kekes
Research Professor, University at Albany, State University of New York
CONSERVATISM (2005)

Birgit Kellner
Institute for South Asian, Tibetan and Buddhist Studies
NEGATION IN INDIAN PHILOSOPHY (2005)

Douglas Kellner
Professor and George F. Kneller Philosophy of Education Chair, Graduate School of Education, University of California, Los Angeles
BAUDRILLARD, JEAN (2005)
HORKHEIMER, MAX (2005)

W. E. Kennick
Professor, Philosophy, Amherst College
APPEARANCE AND REALITY (1967)

G. B. Kerferd
Professor, Classics, University College of Swansea, University of Wales
APEIRON/PERAS (1967)
CRATYLUS (1967)
HEN/POLLA (1967)
HIPPIAS OF ELIS (1967)
PERIPATETICS (1967)
PRODICUS OF CEOS (1967)
PROTAĔORAS OF ABDERA (1967)
PSYCHĒ (1967)

Ralph Ketchum
Professor of Political Science and American Studies, Syracuse University
FRANKLIN, BENJAMIN (1967)

Jeffrey Ketland
Lecturer, Philosophy, University of Edinburgh
CRAIG'S THEOREM (2005)
SECOND-ORDER LOGIC (2005)

I. G. Kidd
Senior Lecturer, Greek, University of St. Andrews
ANTISTHENES (1967)
CYNICS (1967)
DIOGENES OF SINOPE (1967)
GREEK ACADEMY (1967)

Kihyeon Kim
Associate Professor, Philosophy, Seoul National University
KOREAN PHILOSOPHY (2005)

Jeffrey C. King
Professor of Philosophy, University of Southern California
ANAPHORA [ADDENDUM] (2005)
SEMANTICS (2005)

Peter King
Professor of Philosophy and of Mediaeval Studies, University of Toronto
ANSELM, ST. (2005)
AUGUSTINE, ST. [ADDENDUM1] (1996)
WILLIAM OF CHAMPEAUX (2005)

John Kinnaird
Assistant Professor, English, University of Maryland
HAZLITT, WILLIAM (1967)

Eva F. Kittay
State University of New York at Stony Brook
METAPHOR (1996)

Peter Kivy
Board of Governors Professor of Philosophy, Rutgers University
HUTCHESON, FRANCIS [ADDENDUM] (2005)
MUSIC, PHILOSOPHY OF (2005)
SIBLEY, FRANK (2005)
SMITH, ADAM [ADDENDUM] (2005)

Pauline Kleingeld
Professor of Philosophy, Leiden University
PATRIOTISM (2005)

Gyula Klima
Professor, Philosophy, Fordham University
OCKHAMISM [BIBLIOGRAPHY] (2005)

George L. Kline
Professor, Philosophy, Bryn Mawr College
CHICHERIN, BORIS NIKOLAEVICH (1967)
FRANK, SEMËN LIUDVIGOVICH (1967)
HERZEN, ALEKSANDR IVANOVICH (1967)
KAREEV, NIKOLAI IVANOVICH (1967)
KAVELIN, KONSTANTIN DMITRIEVICH (1967)
LUNACHARSKII, ANATOLII VASIL'EVICH (1967)
PISAREV, DMITRI IVANOVICH (1967)
SHESTOV, LEV ISAAKOVICH (1967)
SKOVORODA, HRYHORII SAVYCH (GRIGORII SAVVICH) (1967)
VOLSKI, STANISLAV (1967)

Boris C. A. Kment
Princeton University
CONDITIONALS (2005)

William C. Kneale
White's Professor of Moral Philosophy, University of Oxford
ETERNITY (1967)

David Knowles
Honorary Fellow of Peterhouse and Christ's College, Cambridge University
BERNARD OF CLAIRVAUX, ST. (1967)
BOETHIUS, ANICIUS MANLIUS SEVERINUS (1967)
GERBERT OF AURILLAC (1967)
JOHN OF SALISBURY (1967)

Noretta Koertge
Professor Emeritus, History & Philosophy of Science, Indiana University
SCIENCE STUDIES (2005)

Peter Koestenbaum
Professor of Philosophy, San Jose State College
JASPERS, KARL (1967)
UNAMUNO Y JUGO, MIGUEL DE (1967)

Arthur Koestler
Novelist, Essayist, Man of Letters, Fellow, Royal Society of Literature
KEPLER, JOHANNES (1967)

Barry S. Kogan
Efroymson Professor of Philosophy and Jewish Religious Thought, Hebrew Union College–Jewish Institute of Religion, Cincinnati, Ohio
HALEVI, YEHUDA (2005)

Eckehart Köhler
Member of Phlilosophisches Seminar II, University of Munich, MA Candidate, New York University
SCHOLZ, HEINRICH (1967)

Niko Kolodny
Assistant Professor, Philosophy, University of California, Berkeley
LOVE [ADDENDUM] (2005)
OBJECTIVITY IN ETHICS (2005)

David Konstan
John Rowe Workman Distinguished Professor of Classics and Professor of Comparative Literature, Brown University
LUCIAN OF SAMOSATA (2005)
LUCRETIUS (2005)

Milton R. Konvitz
Professor of Law and Professor of Industrial and Labor Relations, Cornell University
HISTORICAL SCHOOL OF JURISPRUDENCE (1967)
SAVIGNY, FRIEDRICH KARL VON (1967)

Hilary Kornblith
Professor, Philosophy, University of Massachusetts
GOLDMAN, ALVIN (2005)

Stephan Körner
Head, Physics; Professor of Physicas and Philosophy of Science, Bar-Ilan University,
CASSIRER, ERNST (1967)
CONTINUITY (1967)
LAWS OF THOUGHT (1967)

Viacheslav Koshelev
Professor, Novgorod State University named after Yaroslav Mudryi, Member of the International Academy of Higher Education
CHAADAEV, PËTR IAKOVLEVICH (2005)
KHOMIAKOV, ALEKSEI STEPANOVICH (2005)

Kathrin Koslicki
Assistant Professor, Philosophy, Tufts University
NOUNS, MASS AND COUNT (2005)

Janet A. Kourany
Associate Professor, Philosophy, University of Notre Dame
FEMINIST PHILOSOPHY OF SCIENCE: CONTEMPORARY PERSPECTIVES (2005)

Julius Kovesi
Lecturer in Philosophy, University of Western Australia
PALÁGYI, MENYHERT (1967)
PAULER, AKOS (1967)

A. J. Krailsheimer
University Lecturer and College Tutor in French, Christ Church, Oxford University
BOSSUET, JACQUES BÉNIGNE (1967)
FÉNELON, FRANÇOIS DE SALIGNAC DE LA MOTHE (1967)
LA BRUYÈRE, JEAN DE (1967)
LA ROCHEFOUCAULD, DUC FRANÇOIS DE (1967)

Jill Kraye
Professor of the History of Renaissance, Philosophy, Warburg Institute, University of London
HUMANISM (2005)

Norman Kretzmann
Associate Professor, Philosophy, Cornell University
SEMANTICS, HISTORY OF (1967)
WILLIAM OF SHERWOOD (1967)

Yervant H. Krikorian
Professor Emeritus, Philosophy, City University of New York, City College
COHEN, MORRIS RAPHAEL (1967)

Paul Oskar Kristellar
Professor, Philosophy, Columbia University
FICINO, MARSILIO (1967)
FLORENTINE ACADEMY (1967)
PETRARCH (1967)
PICO DELLA MIRANDOLA, COUNT GIOVANNI (1967)
POMPONAZZI, PIETRO (1967)

George Krzywicki-Herburt
Associate Professor, Philosophy, City University of New York, Queens College
TWARDOWSKI, KAZIMIERZ (1967)

Taneli Kukkonen
Canada Research Chair in the Aristotelian Tradition, University of Victoria
ARISTOTELIANISM (2005)

Rahul Kumar
Associate Professor of Philosophy, Queen's University, Kingston, Canada
CONTRACTUALISM (2005)

Joel J. Kupperman
Professor of Philosophy, University of Connecticut
VALUE AND VALUATION [ADDENDUM] (2005)

Paul Kurtz
Professor, Philosophy, State University of New York at Buffalo
PALMER, ELIHU (1967)

Roxanne Marie Kurtz
Assistant Professor, Philosophy, University of Illinois, Springfield
PERSISTENCE (2005)

Douglas Kutach
Assistant Professor, Philosophy, Brown University
COUNTERFACTUALS IN SCIENCE (2005)

Jonathan Kvanvig
Professor and Chair, Philosophy, University of Missouri, Columbia
KNOWLEDGE AND TRUTH, THE VALUE OF (2005)

Kai Man Kwan
Associate Professor, Religion and Philosophy, Hong Kong Baptist University
MORAL ARGUMENTS FOR THE EXISTENCE OF GOD [ADDENDUM] (2005)
MYSTICISM, NATURE AND ASSESSMENT OF [ADDENDUM] (2005)

Will Kymlicka
Canada Research Chair in Political Philosophy, Queen's University
COMMUNITARIANISM (1996, 2005)

Hugh Lacey
Senior Research Scholar/Scheuer Family Professor Emeritus of Philosophy, Swarthmore College; Visiting Professor Universidade de São Paulo; Lecturer, University of Pennsylvania
SKINNER, B. F. (2005)

John Ladd
Professor, Philosophy, Brown University; Secretary-Treasurer of the American Society for Political and Legal Science
LOYALTY (1967)

James Ladyman
Reader in Philosophy, University of Bristol
THEORIES AND THEORETICAL TERMS (2005)

Henrik Lagerlund
Associate Professor in Philosophy, Uppsala University; Research Associate at CRASSH, University of Cambridge
KILWARDBY, ROBERT (2005)

Sterling P. Lamprecht
Professor Emeritus of Philosophy, Amherst College
WOODBRIDGE, FREDERICK JAMES EUGENE (1967)

Irene Lancaster
Honorary Research Fellow, Centre for Jewish Studies, University of Manchester
KABBALAH [ADDENDUM] (2005)

Marc Lange
Professor, Philosophy, University of North Carolina at Chapel Hill
CLASSICAL MECHANICS, PHILOSOPHY OF (2005)
ENERGY [ADDENDUM] (2005)
LAWS, SCIENTIFIC (2005)

Peter Laslett
Fellow of Trinity College, Cambridge, and Lecturer in History, University of Cambridge, Cofounder of the Cambridge Group for the History of Population and Social Structure
FILMER, ROBERT (1967)
POLITICAL PHILOSOPHY, HISTORY OF (1967)
SOCIAL CONTRACT (1967)

John H. Lavely
Professor and Chair, Philosophy, Boston University; Editor of the Philosophical Forum
BRIGHTMAN, EDGAR SHEFFIELD (1967)
PERSONALISM (1967)

James M. Lawler
Philosophy Department, State University of New York at Buffalo
COMMUNISM (2005)
DIALECTICAL MATERIALISM [ADDENDUM] (2005)

Krista Lawlor
Assistant Professor, Philosophy, Stanford University
MILLIKAN, RUTH (2005)

Leonard Lawlor
Faudree-Hardin University Professor of Philosophy; Graduate Admissions Coordinator, Philosophy Department; At Large Member of the Society for Phenomenology and Existential Philosophy, The University of Memphis
HYPPOLITE, JEAN (2005)
TIME IN CONTINENTAL PHILOSOPHY (2005)

Oliver Leaman
Professor of Philosophy, University of Kentucky
AL-FĀRĀBĪ [ADDENDUM] (2005)
AL-KINDĪ, ABŪ-YŪSUF YAʿQŪB IBN ISḤĀQ [ADDENDUM] (2005)
ARKOUN, MOHAMMED (2005)
AVERROES [ADDENDUM] (2005)
AVERROISM IN MODERN ISLAMIC PHILOSOPHY (2005)
AVICENNA [ADDENDUM] (2005)
BAḤYA BEN JOSEPH IBN PAQUDA [ADDENDUM] (2005)
CODOVERO, MOSES BEN JACOB [BIBLIOGRAPHY] (2005)
COSTA, URIEL DA [BIBLIOGRAPHY] (2005)
CRESCAS, HASDAI [ADDENDUM] (2005)
EMANATIONISM [ADDENDUM] (2005)
ENLIGHTENMENT, ISLAMIC (2005)
ENLIGHTENMENT, JEWISH (2005)
ESSENCE AND EXISTENCE [ADDENDUM] (2005)
HOLOCAUST (2005)
IBN BĀJJA [ADDENDUM] (2005)
IBN GABIROL, SOLOMON BEN JUDAH [BIBLIOGRAPHY] (2005)
IBN KHALDŪN [ADDENDUM] (2005)
IBN ṬUFAYL [ADDENDUM] (2005)
IBN ZADDIK, JOSEPH BEN JACOB [BIBLIOGRAPHY] (2005)
ISLAMIC PHILOSOPHY [ADDENDUM] (2005)
ISRAELI, ISAAC BEN SOLOMON [BIBLIOGRAPHY] (2005)
JEWISH AVERROISM (2005)
JEWISH PHILOSOPHY [ADDENDUM] (2005)
MENASSEH (MANASSEH) BEN ISRAEL [BIBLIOGRAPHY] (2005)
MENDELSSOHN, MOSES [BIBLIOGRAPHY] (2005)
MUQAMMIṢ, DAVID BEN MERWAN AL- [BIBLIOGRAPHY] (2005)
NEOPLATONISM [ADDENDUM] (2005)
SHARIATI, ALI (2005)

Mark LeBar
Associate Professor of Philosophy, Ohio University
KANTIAN ETHICS (2005)

Grace Ledbetter
Associate Professor of Classics and Philosophy, Swarthmore College
GREEK DRAMA [BIBLIOGRAPHY] (2005)

Callan Ledsham
Hoger Instituut voor Wijsbegeerte, Katholieke Universiteit Leuven
MARSILIUS OF INGHEN (2005)

Stephen Leeds
Professor of Philosophy, University of Wisconsin, Milwaukee
FIELD, HARTRY (2005)

Gordon Leff
Reader in Medieval History, University of York
AILLY, PIERRE D' (1967)
GILES OF ROME (1967)
GREGORY OF RIMINI (1967)

Brian Leftow
Nolloth Professor of the Philosophy of the Christian Religion, Oxford University
ETERNITY [ADDENDUM 1] (2005)
GOD, CONCEPTS OF [ADDENDUM] (2005)

Czeslaw Lejewski
Senior Lecturer in Philosophy, University of Manchester
LOGIC, HISTORY OF: MODERN LOGIC: FROM FREGE TO GÖDEL [OVERVIEW] (1967)
ŁUKASIEWICZ, JAN (1967)

Karl-Heinz Lembeck
Universitätsprofessor, Institut für Philosophie, Bayerische Julius-Maximilians-Universität Würzburg
NATORP, PAUL (2005)

Noah M. Lemos
Professor, The College of William and Mary
EPISTEMOLOGY, CIRCULARITY IN (2005)

James Lennox
Professor of History and Philosophy of Science, University of Pittsburgh
PHILOSOPHY OF BIOLOGY (2005)

Maria Lucrezia Leone
Postdoctoral Research Fellow, Philosophy, University of Bari (Italy) and Catholic University of Leuven (Belgium)
HENRY OF GHENT [BIBLIOGRAPHY AND ADDENDUM] (2005)

Ernest Lepore
Director, Center for Cognitive Science, Rutgers University
ANALYTIC AND SYNTHETIC STATEMENTS [ADDENDUM] (2005)
PHILOSOPHY OF LANGUAGE (2005)

Joseph Levine
Philosophy, Ohio State University
QUALIA (1996, 2005)
SUBJECTIVITY (1996, 2005)

Jerrold Levinson
University of Maryland at College Park
ART, AUTHENTICITY IN (1996)

Donald Levy
Faculty Member, New School for Social Research
MACROCOSM AND MICROCOSM (1967)

H. D. Lewis
Head, Department of History and Philosophy of Religion, King's College, University of London, and Fellow of King' College; Dean of the Faculty of Theology, University of London; President of the Society for the Study of Theology, Chairman of the Council of the Royal Institute of Philosophy
GUILT (1967)
PHILOSOPHY OF RELIGION, HISTORY OF (1967)

Neil T. Lewis
Associate Professor, Philosophy, Georgetown University
GROSSETESTE, ROBERT [BIBLIOGRAPHY] (2005)
WILLIAM OF AUVERGNE (2005)

Leonard Lewisohn
Iran Heritage Foundation Fellow in Classical Persian and Sufi Literature, The Institute of Arab and Islamic Studies, University of Exeter, England
AL-GHAZĀLĪ, AHMAD (2005)
SUFISM (2005)

Anatoly Liberman
Professor, German, Scandinavian and Dutch, University of Minnesota, Minneapolis
TRUBETSKOI, NIKOLAI SERGEEVICH (2005)

David Liggins
ANALYSIS Student, Faculty of Philosophy, University of Cambridge
FICTIONALISM (2005)

Leonard Linsky
Professor of Philosophy, University of Illinois
SYNONYMITY (1967)

Peter Lipton
Kings College, Cambridge University
INFERENCE TO THE BEST EXPLANATION (1996, 2005)

Iurii Lisitsa
Professor of Mathematical Analysis and Function Theory Department, Russian University of Peoples' Friendship; Head of Faculty of Religion, Russian Orthodox Saint Tikhon Humanistic University, Moscow
IL'IN, IVAN ALEKSANDROVICH (2005)

Jeeloo Liu
Assistant Professor, California State University, Fullerton
WANG FUZHI (2005)

Shu-hsien Liu
Adjunct Research Fellow, Institute of Chinese Literature and Philosophy, Academia Sinica, Taipei; Tuan-mu Kai Chair; Professor, Soochow University, Taipei; Emeritus Professor of Philosophy The Chinese University of Hong Kong
CHINESE PHILOSOPHY: CONTEMPORARY (2005)
HUANG ZONGXI (2005)

Wu-chi Liu
Professor of Chinese; Chair of Department of East Asian Languages and Literature, Indiana University
DONG ZHONGSHU (1967)

Paisley Livingston
Professor, Philosophy, Lingnan University, Hong Kong
CREATIVITY (2005)
VALÉRY, PAUL (2005)

A. C. Lloyd
Professor, Philosophy, University of Liverpool
ALEXANDER OF APHRODISIAS (1967)
PORPHYRY (1967)

G. E. R. Lloyd
University Assistant Lecturer in Classics and Fellow of King's College, Cambridge University
LEUCIPPUS AND DEMOCRITUS (1967)

L. E. Loemker
Charles Howard Candler Professor of Philosophy, Emory University
DEUSSEN, PAUL (1967)
HARTMANN, EDUARD VON (1967)
LIEBERT, ARTHUR (1967)
MONAD AND MONADOLOGY (1967)
PAULSEN, FRIEDRICH (1967)
PESSIMISM AND OPTIMISM (1967)
RINTELEN, FRITZ-JOACHIM VON (1967)
SPRANGER, (FRANZ ERNST) EDUARD (2005)

Barry Loewer
Professor II, Philosophy, Rutgers University
CONTENT, MENTAL (1996, 2005)
PHILOSOPHY OF PHYSICS (2005)

Charles Lohr
Professor Emeritus, History of Medieval Theology, Universität Freiburg
LULL, RAMÓN [BIBLIOGRAPHY] (2005)

Lawrence Brian Lombard
Wayne State University
EVENT THEORY (1996)

Franco Lombardi
Ordinary Professor of Moral Philosophy, University of Rome; Director of De Homine
BLOCH, ERNST (1967)

John L. Longeway
Associate Professor, Philosophy, University of Wisconsin at Parkside
HEYTESBURY, WILLIAM (2005)

Robert B. Louden
University of Southern Maine
VIRTUE ETHICS (1996)

Andrew Louth
Professor of Patristic and Byzantine Studies, University of Durham
JOHN OF DAMASCUS (2005)
PSEUDO-DIONYSIUS [BIBLIOGRAPHY] (2005)

Michael Loux
Schuster Professor of Philosophy, University of Notre Dame
METAPHYSICS, HISTORY OF [ADDENDUM] (2005)

E. J. Lowe
Professor, Philosophy, University of Durham
AGENT CAUSATION (2005)
BENNETT, JONATHAN (2005)

Thomas Luckmann
Professor, University of Frankfurt; Visiting Professor, Graduate Faculty, New School for Social Research
PLESSNER, HELMUT (1967)

Peter Ludlow
State University of New York at Stony Brook
PRESUPPOSITION (1996)

Kirk Ludwig
Professor, Philosophy, University of Florida, Gainesville
BELIEF (2005)

Rossella Lupacchini
Philosophy Department, University of Bologna
COMPUTING MACHINES (2005)

David Luscombe
Fellow and Director of Studies in History, Churchill College, Cambridge University
BERNARD, CLAUDE (1967)
BERNARD OF CHARTRES (1967)
BERNARD OF TOURS (1967)
CHARTRES, SCHOOL OF (1967)
GILBERT OF POITIERS (1967)
SAINT VICTOR, SCHOOL OF (1967)
THEODORIC OF CHARTRES (1967)
WILLIAM OF CONCHES (1967)

Dan Lusthaus
Visiting Professor, Boston University
BUDDHISM—SCHOOLS: YOGACĀRĀ (2005)

J. Rebecca Lyman
Samuel Garrett Professor of Church History Emerita, Church Divinity School of the Pacific
ARIUS AND ARIANISM (2005)

Michael Lynch
Associate Professor of Philosophy, University of Connecticut
RORTY, RICHARD (2005)

William Lyons
University of Dublin, Ireland
INTROSPECTION (1996)

Danielle Macbeth
Professor, Philosophy, Haverford College
MCDOWELL, JOHN (2005)

H. R. MacCallum
Associate Professor, English, University of Toronto
MILTON, JOHN (1967)

Stuart MacClintock
U.S. Government, Department of Defense
AVERROES (1967)
AVERROISM (1967)
JOHN OF JANDUN (1967)

Cynthia MacDonald
Professor of Philosophy, Queen's University Belfast
ANOMALOUS MONISM (1996)
PHYSICALISM (1996, 2005)
SHOEMAKER, SYDNEY (2005)

C. A. Mace
Emeritus Professor, University of London
PSYCHOLOGY (1967)
STOUT, GEORGE FREDERICK (1967)

Tibor Machan
R. C. Hoiles Professor of Business Ethics, Argyros School of Business & Economics, Chapman University
POLITICAL PHILOSOPHY, HISTORY OF [ADDENDUM] (2005)
PROPERTY [ADDENDUM] (2005)

Alasdair MacIntyre
Professor, Sociology, University of Essex
BEING (1967)
BRUNNER, EMIL (1967)
EGOISM AND ALTRUISM (1967)
ESSENCE AND EXISTENCE (1967)
EXISTENTIALISM (1967)
JUNG, CARL GUSTAV (1967)
KIERKEGAARD, SØREN AABYE (1967)
MYTH (1967)
ONTOLOGY, HISTORY OF (1967)
PANTHEISM (1967)

J. L. Mackie
Professor of Philosophy, University of York
FALLACIES (1967)
MILL'S METHODS OF INDUCTION (1967)
WESTERMARCK, EDWARD ALEXANDER (1967)

Ruth Macklin
Bronx, New York
GENETICS AND REPRODUCTIVE TECHNOLOGIES (1996)

John MacQuarrie
Professor of Systematic Theology, Union Theological Seminary
BLONDEL, MAURICE (1967)
GOGARTEN, FRIEDRICH (1967)
HARNACK, CARL GUSTAV ADOLF VON (1967)
HEIM, KARL (1967)
INGE, WILLIAM RALPH (1967)
LABERTHONNIÈRE, LUCIEN (1967)
PIETISM (1967)
TAYLOR, ALFRED EDWARD (1967)
VARISCO, BERNARDINO (1967)

Edward H. Madden
Professor of Philosophy, State University of New York at Buffalo; General Editor of Source Books in the History of Science (Harvard University Press)
WRIGHT, CHAUNCEY (1967)

Patrick Maher
Professor, Philosophy, University of Illinois at Urbana-Champaign
CONFIRMATION THEORY (2005)

James Edwin Mahon
Assistant Professor, Philosophy, Washington and Lee University
LYING (2005)

Rudolf Makkreel
Charles Howard Candler Professor of Philosophy, Emory University
DILTHEY, WILHELM (2005)

Norman Malcolm
Susan Linn Sage Professor of Philosophy and Chair of the Department of Philosophy, Cornell University; Managing Editor of the Philosophical Review
WITTGENSTEIN, LUDWIG JOSEF JOHANN (1967)

Paolo Mancosu
Associate Professor of Philosophy, University of California, Berkeley
HILBERT, DAVID [ADDENDUM] (2005)

Maurice Mandelbaum
Professor, Philosophy, Johns Hopkins University
HISTORICISM (1967)

Jon Mandle
Associate Professor of Philosophy, University at Albany (State University of New York)
GENERAL WILL, THE (2005)

William E. Mann
Marsh Professor of Intellectual and Moral Philosophy, University of Vermont
PATRISTIC PHILOSOPHY [BIBLIOGRAPHY] (2005)

A. R. Manser
Senior Lecturer, Philosophy, University of Southampton
DREAMS (1967)
IMAGES (1967)
IMAGINATION (1967)

Vladimir Marchenkov
Assistant Professor of Aesthetics, Ohio University
BAKUNIN, MIKHAIL ALEKSANDROVICH [BIBLIOGRAPHY] (2005)
BELINSKII, VISSARION GRIGOR'EVICH [BIBLIOGRAPHY] (2005)
BULGAKOV, SERGEI NIKOLAEVICH [BIBLIOGRAPHY] (2005)
CHERNYSHEVSKII, NIKOLAI GAVRILOVICH [BIBLIOGRAPHY] (2005)
FRANK, SEMËN LIUDVIGOVICH [BIBLIOGRAPHY] (2005)

HERZEN, ALEKSANDR IVANOVICH [BIBLIOGRAPHY] (2005)
KAREEV, NIKOLAI IVANOVICH [BIBLIOGRAPHY] (2005)
KAVELIN, KONSTANTIN DMITRIEVICH [BIBLIOGRAPHY] (2005)
KOZLOV, ALEKSEI ALEKSANDROVICH [BIBLIOGRAPHY] (2005)
KROPOTKIN, PËTR ALEKSEEVICH [BIBLIOGRAPHY] (2005)
LAPSHIN, IVAN IVANOVICH [BIBLIOGRAPHY] (2005)
LAVROV, PËTR LAVROVICH [BIBLIOGRAPHY] (2005)
LOPATIN, LEV MIKHAILOVICH [BIBLIOGRAPHY] (2005)
LOSEV, ALEKSEI FËDOROVICH (2005)
LUNACHARSKII, ANATOLII VASIL'EVICH [BIBLIOGRAPHY] (2005)
MIKHAILOVSKII, NIKOLAI KONSTANTINOVICH [BIBLIOGRAPHY] (2005)
PAVLOV, IVAN PETROVICH [BIBLIOGRAPHY] (2005)
PLEKHANOV, GEORGII VALENTINOVICH [BIBLIOGRAPHY] (2005)
RUSSIAN PHILOSOPHY (2005)
VYSHESLAVTSEV, BORIS PETROVICH [BIBLIOGRAPHY] (2005)

John Marenbon
Senior Research Fellow, Trinity College, Cambridge University
ABELARD, PETER (2005)

Adam Margoshes
Assistant Professor of Psychology, Shippensburg State College
BAADER, FRANZ XAVIER VON (1967)
SCHELLING, FRIEDRICH WILHELM JOSEPH VON (1967)

Jacqueline Mariña
Associate Professor of Philosophy; Chair, Religious Studies Program, Purdue University
SCHLEIERMACHER, FRIEDRICH DANIEL ERNST [ADDENDUM] (2005)

R. A. Markus
Senior Lecturer in Medieval History, University of Liverpool
AUGUSTINE, ST. (1967)
ILLUMINATION (1967)

Michael E. Marmura
Associate Professor in the Department of Islamic Studies, University of Toronto
AVICENNA (1967)

Donald Marquis
University of Kansas, Lawrence
ABORTION (1996)

Christopher J. Martin
Associate Professor, Philosophy, University of Auckland
BOETHIUS, ANICIUS MANLIUS SEVERINUS [ADDENDUM] (2005)
LOGIC, HISTORY OF: MEDIEVAL (EUROPEAN) LOGIC (2005)
ROSCELIN (2005)

Rex Martin
Professor of Philosophy, University of Kansas; Honorary Professor, School of European Studies, Cardiff University
PHILOSOPHY OF LAW, PROBLEMS OF [ADDENDUM] (2005)

Wayne M. Martin
Reader in Philosophy, Essex University
HEGELIANISM (2005)

Martin E. Marty
Fairfax M. Cone Distinguished Service Professor Emeritus, The University of Chicago
BONHOEFFER, DIETRICH (1967, 2005)

Elinor Mason
Lecturer, Philosophy, University of Edinburgh
CONSEQUENTIALISM [BIBLIOGRAPHY] (2005)

Kelby Mason
PhD candidate, Philosophy, Rutgers University
EVOLUTIONARY PSYCHOLOGY (2005)

Michelle Mason
Assistant Professor, Philosophy, University of Minnesota
MORAL SENTIMENTS (2005)

Heath Massey
Assistant Professor of Philosophy, Beloit College
TIME IN CONTINENTAL PHILOSOPHY (2005)

Wallace Matson
Professor of Philosophy, Emeritus, University of California, Berkeley
ARISTIPPUS OF CYRENE (2005)
CYRENAICS (2005)

Gareth B. Matthews
Professor of Philosophy, University of Massachusetts, Amherst
AUGUSTINIANISM [BIBLIOGRAPHY] (2005)

Robert Matthews
Rutgers University
MENTAL REPRESENTATION (1996)

Tim Maudlin
Professor, Philosophy, Rutgers University
GAUGE THEORY (2005)
NON-LOCALITY (2005)
QUANTUM LOGIC AND PROBABILITY (2005)
QUANTUM MECHANICS (2005)
RELATIVITY THEORY (2005)

Armand A. Maurer
Professor of Philosophy, Pontifical Institute of Medieval Studies and University of Toronto
BOETIUS OF DACIA (1967)
BROWNSON, ORESTES AUGUSTUS (1967)
EDWARDS, JONATHAN (1967)
HENRY OF HARCLAY (1967)
NICHOLAS OF CUSA (1967)

Wolfe Mays
Senior Lecturer in Philosophy, University of Manchester;Visting Professor of Philosophy, Northwestern University
PIAGET, JEAN (1967)

Bruce Mazlish
Professor, History, Massachusetts Institute of Technology
COMTE, AUGUSTE (1967)

William L. McBride
Arthur G. Hansen Distinguished Professor of Philosophy, Purdue University
IDEOLOGY (2005)
SOCIETY [ADDENDUM] (2005)

Edwin McCann
Professor of Philosophy, University of Southern California
LOCKE, JOHN (2005)

Hugh J. McCann
Professor of Philosophy, Texas A&M University
CREATION AND CONSERVATION, RELIGIOUS DOCTRINE OF (2005)

Thomas McCarthy
Northwestern University
CRITICAL THEORY (1996)
DISCOURSE ETHICS (1996)
HABERMAS, JÜRGEN (1996)

James E. McClellan Jr.
Director, Foundations of Education Department and the General Education Program for Teachers; Professor of Education, Temple University
PHILOSOPHY OF EDUCATION, INFLUENCE OF MODERN PSYCHOLOGY ON (1967)

John J. McDermott
Distinguished Professor of Philosophy and Humanities, Texas A&M University
DEWEY, JOHN [BIBLIOGRAPHY AND ADDENDUM] (1996)
JAMES, WILLIAM [ADDENDUM] (1996, 2005)

Vann McGee
Professor, Linguistics and Philosophy, Massachusetts Institute of Technology
GÖDEL'S INCOMPLETENESS THEOREMS (2005)
LOGICAL PARADOXES (2005)

Sarah McGrath
Assistant Professor, Philosophy, College of the Holy Cross
THOMSON, JUDITH JARVIS (2005)

Ralph McInerny
Michael P. Grace Professor of Medieval Studies, University of Notre Dame
MARITAIN, JACQUES [ADDENDUM] (2005)

Neil McInnes
Special writer in Paris for Barron's Financial Weekly and the Wall Street Journal
ENGELS, FRIEDRICH (1967)
GRACIÁN Y MORALES, BALTASAR (1967)
GRAMSCI, ANTONIO (1967)
LABRIOLA, ANTONIO (1967)
LUKÁCS, GEORG (1967)
MARÍAS, JULIÁN (1967)
MARXIST PHILOSOPHY (1967)
ORTEGA Y GASSET, JOSÉ (1967)
SOREL, GEORGES (1967)
ZUBIRI, XAVIER (1967)

Sean McKeever
Assistant Professor, Philosophy, Davidson College
ETHICS, HISTORY OF: OTHER DEVELOPMENTS IN TWENTIETH-CENTURY ETHICS (2005)

William McKenna
Chair, Department of Philosophy, Miami University of Ohio
GURWITSCH, ARON (2005)

Robert J. McKim
Professor in the Program for the Study of Religion and Professor of Philosophy, University of Illinois at Urbana-Champaign
BERKELEY, GEORGE [ADDENDUM] (1996, 2005)

Richard McKirahan
E.C. Norton Professor of Classics and Professor of Philosophy, Pomona College
MELISSUS OF SAMOS [BIBLIOGRAPHY] (2005)
XENOPHANES OF COLOPHON (2005)
ZENO OF ELEA (2005)

Brian McLaughlin
Professor of Philosophy, Rutgers University
MENTAL CAUSATION (1996, 2005)
MIND-BODY PROBLEM (2005)
SELF-KNOWLEDGE (2005)
SUPERVENIENCE (1996, 2005)

Y. P. Mei
Professor, Chair, Chinese and Oriental Studies; Director of the Center for Far Eastern Studies, University of Iowa
MOZI (1967)

Jørgen Mejer
Director, Danish Institute; Reader in Classical Philology, University of Copenhagen, Institute for Greek and Latin
DIOGENES LAERTIUS (2005)

Alfred R. Mele
William H. and Lucyle T. Werkmeister Professor of Philosophy, Florida State University
ACTION (1996, 2005)
INTENTION (2005)
WEAKNESS OF THE WILL (2005)

Henry Mendell
California State University, Los Angeles
COSMOS (2005)

Eduardo Mendieta
Associate Professor, Philosophy, Stony Brook University, State University of New York
POSTCOLONIALISM (2005)

Stephen Menn
Associate Professor, Philosophy, McGill University
ARISTOTLE (2005)

Christia Mercer
Gustave M. Berne Professor; Philosophy, Columbia University; North American Editor, Archiv für Geschichte der Philosophie
LEIBNIZ, GOTTFRIED WILHELM (2005)

Philip Merlan
Professor of German Philosophy and Literature, Scripps College and Claremont Graduate School, CA
EMANATIONISM (1967)
PLOTINUS (1967)

Pierre Mesnard
Professor of Philosophy; Faculty of Letters, University of Orléans-Tours; Director of Centre d'Études Supérieures de la Renaissance de Tours
BODIN, JEAN (1967)

Algis Mickunas
Philosophy, Ohio University
LANDGREBE, LUDWIG (2005)

Leland Miles
Dean of the College of Arts and Sciences, University of Bridgeport
COLET, JOHN (1967)

T. R. Miles
Professor of Psychology, University College of North Wales, Bangor
GESTALT THEORY (1967)
KOFFKA, KURT (1967)

Elizabeth Millán-Zaibert
Assistant Professor of Philosophy, DePaul University
LATIN AMERICAN PHILOSOPHY (2005)

Robert G. Miller, C.S.B.
Chairman of the Department of Philosophy, St. John Fisher College
GILSON, ÉTIENNE HENRY (1967)

Peter Milne
Reader, School of Philosophy, Psychology and Language Sciences of the University of Edinburgh
DE FINETTI, BRUNO (2005)

Edward H. Minar
Associate Professor of Philosophy, University of Arkansas, Fayetteville
CAVELL, STANLEY (2005)

Robert C. Miner
Associate Professor of Philosophy in the Honors College, Baylor University
VICO, GIAMBATTISTA (2005)

Carl Mitcham
Professor, Liberal Arts and International Studies, Colorado School of Mines
PHILOSOPHY OF TECHNOLOGY (1996, 2005)
SCIENCE POLICY (2005)
SOCIAL CONSTRUCTIONISM (2005)

Phillip Mitsis
A.S. Onassis Professor of Hellenic Studies, New York University
EPICURUS (2005)

Marc A. Moffett
Assistant Professor of Philosophy, University of Wyoming
SYNONYMITY [ADDENDUM] (2005)

Jitendra N. Mohanty
Temple University
PHENOMENOLOGY [ADDENDUM] (1996)

Michel Mohr
Lecturer, Institute for Language and Culture, Doshisha University
BUDDHISM—SCHOOLS: CHAN AND ZEN (2005)

D. H. Monro
Professor of Philosophy, Monash University
GODWIN, WILLIAM (1967)
HUMOR (1967)
SHELLEY, PERCY BYSSHE (1967)

Josep Puig Montada
Professor, Arabic and Islamic Studies, Universidad Complutense, Madrid
AL-JABIRI, ʿABD (2005)

Michelle Montague
Assistant Professor, Philosophy, University of California, Irvine
COUNTERFACTUALS (2005)

Ernest A. Moody
Professor of Philosophy, University of California, Los Angeles
LOGIC, HISTORY OF: MEDIEVAL (EUROPEAN) LOGIC (1967)
OCKHAMISM (1967)
WILLIAM OF OCKHAM (1967)

James H. Moor
Professor, Philosophy, Dartmouth College
COMPUTER ETHICS (1996, 2005)
MACHINE INTELLIGENCE (2005)

Kevin Moore
PhD candidate, Joint Program at University of Denver and Iliff School of Theology
HERMETICISM [BIBLIOGRAPHY] (2005)

Merritt Hadden Moore
Professor and Head of the Department of Philosophy, University of Tennessee
COURNOT, ANTOINE AUGUSTIN (1967)

Nancy J. Moore
Professor of Law and Nancy Barton Scholar, Boston University School of Law
INFORMED CONSENT IN THE PRACTICE OF LAW (2005)

Michael Moran
Lecturer in Philosophy, School of European Studies, University of Sussex
CARLYLE, THOMAS (1967)
COLERIDGE, SAMUEL TAYLOR (1967)
EMERSON, RALPH WALDO (1967)
NEW ENGLAND TRANSCENDENTALISM (1967)
THOREAU, HENRY DAVID (1967)

Parviz Morewedge
Director, Global Scholarly Publications, New York; Honorary Professor, National University, Bishkek, Kyrgyz Republic; Adjunct Professor, Philosophy, Fordham University; Instructor, Philosophy and Religion, Rutgers University
NAṢĪR AL-DĪN AL-ṬŪSĪ (2005)
SCHOOL OF QOM, THE (2005)

John Morreall
Professor, Religious Studies, College of William and Mary
HUMOR [ADDENDUM] (2005)

Herbert Morris
Professor of Philosophy and Law at University of California, Los Angeles
AUSTIN, JOHN (1967)

John Morrison
President of University College, Cambridge University
ORPHISM (1967)

Paul Moser
Professor and Chair of Philosophy, Loyola University of Chicago
PHILOSOPHY OF RELIGION [ADDENDUM] (2005)
PROPOSITIONAL KNOWLEDGE, DEFINITION OF (1996)
RATIONALITY (1996)

Albert G. Mosley
Professor of Philosophy, Smith College
AFRICAN PHILOSOPHY (2005)
PLURALISM (1996)
RACISM [ADDENDUM] (1996, 2005)

Ernest Campbell Mossner
Professor of English, University of Texas, Austin; Joint Editor of "Texas Studies in Literature and Language"
ANNET, PETER (1967)
BLOUNT, CHARLES (1967)
BOLINGBROKE, HENRY ST. JOHN (1967)
CHUBB, THOMAS (1967)
COLLINS, ANTHONY (1967)
DEISM (1967)
GIBBON, EDWARD (1967)

JOHNSON, SAMUEL (1967)
MORGAN, THOMAS (1967)
POPE, ALEXANDER (1967)
SWIFT, JONATHAN (1967)
TINDAL, MATTHEW (1967)
TOLAND, JOHN (1967)
WOOLSTON, THOMAS (1967)

Andrzej Mostowski
Professor of Mathematics, University of Warsaw; Corresponding Member of the Polish Academy of Sciences
TARSKI, ALFRED (1967)

Mary Mothersill
Professor and Chair, Philosophy, Barnard College
DUTY (1967)

Nelia Motroshilova
Head of Department of the History of Philosophy, Institute of Philosophy, Russian Academy of Sciences
MAMARDASHVILI, MERAB KONSTANTINOVICH (2005)

Bo Mou
CHINESE PHILOSOPHY: LANGUAGE AND LOGIC (2005)
GONGSUN LONG (2005)
HUI SHI (2005)

John A. Mourant
Professor of Philosophy, Pennsylvania State University
AUGUSTINIANISM (1967)
SCIENTIA MEDIA AND MOLINISM (1967)
SUÁREZ, FRANCISCO (1967)

Alexander P. D. Mourelatos
Assistant Professor of Philosophy, University of Texas at Austin
FRIES, JAKOB FRIEDRICH (2005)
PRE-SOCRATIC PHILOSOPHY [BIBLIOGRAPHY] (2005)

Charles Muller
Faculty of Humanities, Toyo Gakuen University
JINUL (2005)

C. W. K. Mundle
Head, Philosophy Department, University College of North Wales, University of Wales
TIME, CONSCIOUSNESS OF (1967)

Milton K. Muntiz
Professor of Philosophy, New York University
COSMOLOGY (1967)

Peter Munz
Associate Professor of Philosophy, University of Wellington
HOOKER, RICHARD (1967)

Murray G. Murphey
Associate Professor of American Civilization, University of Pennsylvania
PEIRCE, CHARLES SANDERS (1967)

Jeffrie G. Murphy
Regents' Professor of Law, Philosophy, and Religious Studies, Arizona State University
SHAME (2005)

Mark C. Murphy
Professor of Philosophy, Georgetown University
AUTHORITY (2005)
MACINTYRE, ALASDAIR (2005)

Mauro Murzi
Società Filosofica Italiana
REICHENBACH, HANS (2005)

Herbert Musurillo, S.J.
Professor of Classics, Fordham University
GREGORY OF NAZIANZUS (1967)
GREGORY OF NYSSA (1967)

Mechthild Nagel
Associate Professor of Philosophy, State University of New York, Cortland
FERGUSON, ANN (2005)

Thomas Nagel
University Professor, Professor of Philosophy, and Professor of Law, New York University
ETHICS (2005)
WILLIAMS, BERNARD (2005)

Jan Narveson
University of Waterloo, Canada
ETHICAL EGOISM (1996)

Seyyed Hossein Nasr
Professor of the History of Science and Philosophy, Tehran University
MULLĀ ṢADRĀ (1967)

Stephen Nathanson
Professor of Philosophy, Northeastern University
CIVIL DISOBEDIENCE (2005)

Stephen Neale
University of California, Berkeley
ANAPHORA (1996)

Karen Neander
Professor of Philosophy, University of California, Davis
EVOLUTIONARY THEORY (2005)
TELEOLOGY [ADDENDUM] (2005)

Jacob Needleman
Associate Professor of Philosophy, San Francisco State College
BINSWANGER, LUDWIG (1967)
EXISTENTIAL PSYCHOANALYSIS (1967)

Alex Neill
Senior Lecturer in Philosophy, University of Southampton
TRAGEDY (2005)

Susan Neiman
Director, Einstein Forum, Potsdam, Germany
EVIL (2005)

John O. Nelson
Professor of Philosophy, University of Colorado
INNATE IDEAS (1967)
MOORE, GEORGE EDWARD (1967)

Lynn Hankinson Nelson
Professor of Philosophy, University of Washington
CODE, LORRAINE (2005)
HARDING, SANDRA (2005)

Thomas Nenon
Professor of Philosophy, University of Memphis
ABBAGNANO, NICOLA [BIBLIOGRAPHY] (2005)
BINSWANGER, LUDWIG [BIBLIOGRAPHY] (2005)
BRUNSCHVICG, LÉON [BIBLIOGRAPHY] (2005)
CAMUS, ALBERT [BIBLIOGRAPHY] (2005)
CASSIRER, ERNST [BIBLIOGRAPHY] (2005)
CONSCIOUSNESS IN PHENOMENOLOGY (2005)
CONTINENTAL PHILOSOPHY (2005)

EXISTENTIAL PSYCHOANALYSIS [BIBLIOGRAPHY] (2005)
GEISTESWISSENSCHAFTEN (2005)
HARTMANN, NICOLAI [BIBLIOGRAPHY] (2005)
INGARDEN, ROMAN [BIBLIOGRAPHY] (2005)
JASPERS, KARL [BIBLIOGRAPHY] (2005)
KIERKEGAARD, SØREN AABYE [ADDENDUM] [BIBLIOGRAPHY] (2005)
LAVELLE, LOUIS [BIBLIOGRAPHY] (2005)
LUKÁCS, GEORG [BIBLIOGRAPHY] (2005)
MARCEL, GABRIEL [BIBLIOGRAPHY] (2005)
MOUNIER, EMMANUEL [BIBLIOGRAPHY] (2005)
SCHELLING, FRIEDRICH WILHELM JOSEPH VON [BIBLIOGRAPHY] (2005)
UNAMUNO Y JUGO, MIGUEL DE [BIBLIOGRAPHY] (2005)
WEIL, SIMONE [BIBLIOGRAPHY] (2005)
ZUBIRI, XAVIER [BIBLIOGRAPHY] (2005)

G. C. Nerlich
Senior Lecturer in Philosophy, University of Sydney
EDDINGTON, ARTHUR STANLEY (1967)
JEANS, JAMES HOPWOOD (1967)
POPULAR ARGUMENTS FOR THE EXISTENCE OF GOD (1967)
STEBBING, LIZZIE SUSAN (1967)

Ram Neta
Assistant Professor of Philosophy, University of North Carolina, Chapel Hill
CONTEXTUALISM (2005)

Richard R. Niebuhr
Florence Corliss Lamont Professor of Divinity, Harvard Divinity School, Harvard University
SCHLEIERMACHER, FRIEDRICH DANIEL ERNST (1967)

Kai Nielsen
Unversity of Calgary (Emeritus); Concordia University, Montreal (Canada)
SOCIALISM [ADDENDUM] (2005)

Robert Niklaus
Professor of French and Italian and Deputy Vice-Chancellor, University of Exeter; Member of the Executive Committees of the Modern Humanities Research Association and the Society for French Studies
BOULAINVILLIERS, HENRI, COMTE DE (1967)
CLANDESTINE PHILOSOPHICAL LITERATURE IN FRANCE (1967)

Daniel Nolan
Professor of Theoretical Philosophy, Departments of Philosophy, University of St. Andrews
LEWIS, DAVID (2005)

Harold Noonan
Professor, Philosophy, University of Nottingham
PERSONAL IDENTITY [ADDENDUM] (1996, 2005)
WIGGINS, DAVID (2005)

Calvin G. Normore
Professor of Philosophy, University of California, Los Angeles
SCOTISM (2005)

F. S. Northedge
Reader in International Relations, London School of Economics and Political Science, University of London
PEACE, WAR, AND PHILOSOPHY (1967)

Vivian Nutton
Professor, History of Medicine, University College London
GALEN (2005)

Michael Nylan
Professor, History, University of California, Berkeley
YANG XIONG (2005)

Thomas C. O'Brien, O.P.
Dominican House of Studies, Washington, DC
GARRIGOU-LAGRANGE, RÉGINALD MARIE (1967)

D. J. O'Connor
Professor of Philosophy, University of Exeter
AYER, ALFRED JULES (1967)
SUBSTANCE AND ATTRIBUTE (1967)

Graham Oddie
Professor, Philosophy, University of Colorado at Boulder
METAPHYSICS (2005)
TRUTHLIKENESS (1996)

Frederick A. Olafson
Professor of Education and Philosophy, Harvard Graduate School of Education, Harvard University
CAMUS, ALBERT (1967)
SANTAYANA, GEORGE (1967)

Andrew Oldenquist
Associate Professor of Philosophy, Ohio State University
FISKE, JOHN (1967)
SELF-PREDICTION (1967)

Kelly Oliver
W. Alton Jones Professor, Philosophy, Vanderbilt University
KRISTEVA, JULIA (2005)

Robert G. Olson
Associate Professor and Chairman of the Department of Philosophy, University College, Rutgers University
DEATH (1967)
NIHILISM (1967)

Eileen O'Neill
University of Massachusetts, Amherst
GOURNAY, MARIE LE JARS DE (1996)
WOMEN IN THE HISTORY OF PHILOSOPHY (1996)

William G. O'Neill
Associate Professor of Philosophy, Iona College, New Rochelle, NY
CONSCIENCE [BIBLIOGRAPHY] (2005)

Walter J. Ong, S.J.
Professor of English, St. Louis University
RAMUS, PETER (1967)

Jan Opsomer
Associate Professor, Philosophy, University of Cologne
PLUTARCH OF CHAERONEA (2005)

Leslie E. Orgel
Professor, Salk Institute for Biological Studies
LIFE, ORIGIN OF (2005)

E. F. Osborn
Professor of Biblical Studies, Queen's College, University of Melbourne; Editor of the Australian Biblical Review
CLEMENT OF ALEXANDRIA (1967)
PSEUDO-DIONYSIUS (1967)

Jennifer Ottman
Visiting Assistant Professor, History, Wake Forest University
RUFUS, RICHARD (2005)

H. P. Owen
Reader in the Philosophy of Religion, University of London
DOGMA (1967)
ESCHATOLOGY (1967)
INFINITY IN THEOLOGY AND METAPHYSICS (1967)
PERFECTION (1967)
PROVIDENCE (1967)

G. Michael Pace
Assistant Professor, Philosophy, Chapman University
PERCEPTION, CONTEMPORARY VIEWS (2005)

Alan G. Padgett
Professor of Systematic Theology, Luther Seminary, St. Paul, MN; Editor, Journal for Christian Theological Research
BARTH, KARL [ADDENDUM] (2005)
LUTHER, MARTIN [ADDENDUM] (2005)
RELIGION AND THE PHYSICAL SCIENCES (2005)

David A. Pailin
Emeritus, University of Manchester, School of Middle Eastern Studies & Theology
HERBERT OF CHERBURY (2005)

Richard F. Palmer
MacMurray College
GADAMER, HANS-GEORG (1996)

Claude Panaccio
Canada Research Chair in the Theory of Knowledge, Philosophy, University of Quebec at Montreal
WILLIAM OF OCKHAM [BIBLIOGRAPHY] (2005)

David Papineau
Professor of Philosophy, King's College London
CAUSAL CLOSURE OF THE PHYSICAL DOMAIN (2005)

John N. Pappas
Professor of French, Department of Romance Languages, University of Pennsylvania
ALEMBERT, JEAN LE ROND D' (1967)

H. O. Pappé
Member if the Faculty of the School of Social Sciences, University of Sussex
GEHLEN, ARNOLD (1967)
JÜNGER, ERNST (1967)
PHILOSOPHICAL ANTHROPOLOGY (1967)
SOMBART, WERNER (1967)

Charles Parsons
Edgar Pierce Professor of Philosophy, Emeritus, Harvard University
BROUWER, LUITZEN EGBERTUS JAN (1967)
MATHEMATICS, FOUNDATIONS OF (1967, 2005)

P. H. Partridge
Director of the Research School of Social Sciences and Professor of Social Philosophy, Australian National University
FREEDOM (1967)
MOSCA, GAETANO (1967)

John Passmore
Professor of Philosophy, Institute of Advanced Studies, Australian National University
ANDERSON, JOHN (1967)
BOYLE, ROBERT (1967)
COLLIER, ARTHUR (1967)
CULVERWEL, NATHANAEL (1967)
CUMBERLAND, RICHARD (1967)
FLUDD, ROBERT (1967)
HARVEY, WILLIAM (1967)
LOGICAL POSITIVISM (1967)
MORE, HENRY (1967)
NORRIS, JOHN (1967)
PRIESTLEY, JOSEPH (1967)
SMITH, JOHN (1967)
WHICHCOTE, BENJAMIN (1967)

Stanley L. Paulson
William Gardiner Hammond Professor of Law, and Professor of Philosophy, Washington University School of Law, St. Louis, MO
KELSEN, HANS (2005)

Vuko Pavićević
Professor of Ethics and Sociology, University of Belgrade
PETROVIĆ-NJEGOŠ, PETAR (1967)

Francis Jeffry Pelletier
Canada Research Chair in Cognitive Science, Professor of Philosophy, Professor of Linguistics, Simon Fraser University
GENERICS (2005)

Terence Penelhum
Professor Emeritus of Religious Studies, University of Calgary
BUTLER, JOSEPH (2005)
PERSONAL IDENTITY (1967)

Adriaan Peperzak
Loyola University of Chicago
LEVINAS, EMMANUEL (1996)

Derk Pereboom
Professor, Philosophy, University of Vermont
NONREDUCTIVE PHYSICALISM (2005)

Alan Perreiah
University of Kentucky
PAUL OF VENICE (2005)

Roy W. Perrett
Professor of Philosophy, University of Hawaii at Manoa
ATOMIC THEORY IN INDIAN PHILOSOPHY (2005)

John Perry
Stanford University
INDEXICALS (1996)
SELF (1996)

R. S. Peters
Professor of the Philosophy of Education, University of London Institute of Education
HOBBES, THOMAS (1967)
PSYCHOLOGY (1967)

Gajo Petrović
Associate Professor of Philosophy, University of Zagreb; President of the Yugoslav Philosophical Association
ALIENATION (1967)
PLEKHANOV, GEORGII VALENTINOVICH (1967)

Philip Pettit
L. S. Rockefeller University Professor of Politics and Human Values, Princeton University
PHILOSOPHY OF SOCIAL SCIENCES (1996)
RESPONSE-DEPENDENCE THEORIES (2005)

Stephen H. Phillips
Professor, Philosophy, University of Texas at Austin
TRUTH AND FALSITY IN INDIAN PHILOSOPHY (2005)

Charles R. Pidgen
University of Otago, New Zealand
HARE, RICHARD M. (1996)

Paul M. Pietroski
Professor of Philosophy, Professor of Linguistics, University of Maryland, College Park
EVENTS IN SEMANTIC THEORY (2005)
LOGICAL FORM (2005)

Kirk Pillow
Associate Dean of the Faculty, Associate Professor of Philosophy, Hamilton College, Clinton, NY
SUBLIME, THE (2005)

Shlomo Pines
Professor of General and Jewish Philosophy, Hebrew University of Jerusalem
JEWISH PHILOSOPHY (1967)
MAIMONIDES (1967)

Gino K. Piovesana, S.J.
Director of the Board of Regents and Professor of Philosophy, Sophia University, Japan
ANDŌ SHŌEKI (1967)
HATANO SEIICHI (1967)
HAYASHI RAZAN (1967)
ITŌ JINSAI (1967)
KAIBARA EKKEN (1967)
KUMAZAWA BANZAN (1967)
MIKI KIYOSHI (1967)
MINAGAWA KIEN (1967)
MIURA BAIEN (1967)
MURO KYŪSŌ (1967)
NAKAE TŌJU (1967)
NISHI AMANE (1967)
OGYŪ SORAI (1967)
WATSUJI TETSURŌ (1967)
YAMAGA SOKŌ (1967)
YAMAZAKI ANSAI (1967)

Robert B. Pippin
University of Chicago
HEGEL, GEORG WILHELM FRIEDRICH [ADDENDUM] (1996)

Fabienne Pironet
Professeure agrégée, Université de Montréal, Faculté des arts et des science, Département de philosophie
SIGER OF BRABANT (2005)

Alvin Plantinga
Professor of Philosophy, Calvin College
MALCOLM, NORMAN (1967)

Thomas Pogge
Professorial Research, Fellow, Centre for Applied Philosophy and Public Ethics, Australian National University
JUSTICE [BIBLIOGRAPHY] (2005)

Richard H. Popkin
Professor of Philosophy, University of California, Los Angeles
AGRIPPA VON NETTESHEIM, HENRICUS CORNELIUS (1967)
BAYLE, PIERRE (1967, 2005)
CHARRON, PIERRE (1967, 2005)
COSTA, URIEL DA (1967)
ERASMUS, DESIDERIUS (1967, 2005)
FIDEISM (1967)
GASSENDI, PIERRE (1967, 2005)
GLANVILL, JOSEPH (1967, 2005)
HUET, PIERRE-DANIEL (1967, 2005)
LA MOTHE LE VAYER, FRANÇOIS DE (1967, 2005)
LA PEYRÈRE, ISAAC (1967, 2005)
MENASSEH (MANASSEH) BEN ISRAEL (1967)
MERSENNE, MARIN (1967, 2005)
MONTAIGNE, MICHEL EYQUEM DE (1967, 2005)
OROBIO DE CASTRO, ISAAC (1967, 2005)
PASCAL, BLAISE (1967, 2005)
PICO DELLA MIRANDOLA, GIANFRANCESCO (2005)
SANCHES, FRANCISCO (1967, 2005)
SIMON, RICHARD (1967, 2005)
SKEPTICISM, HISTORY OF (1967, 2005)

Peter E. Pormann
Frances A. Yates Long-Term Research Fellow, Warburg Institute, School of Advanced Studies, University of London
GALEN [ADDENDUM] (2005)

Amanda Porter
PhD candidate, Philosophy, University of Western Ontario
CARD, CLAUDIA (2005)
HELD, VIRGINIA (2005)

Karl H. Potter
Professor Emeritus, University of Washington, Seattle
INDIAN PHILOSOPHY (2005)

C. F. Presley
Head of the Department of Philosophy, University of the Queensland (Australia)
QUINE, WILLARD VAN ORMAN (1967)

Kingsley Price
Professor of Philosophy, Johns Hopkins University
PHILOSOPHY OF EDUCATION, HISTORY OF (1967)

Graham Priest
Boyce Gibson Professor of Philosophy, University of Melbourne; Arche Professorial Fellow, Department of Logic and Metaphysics, University of St. Andrews
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL: THE PROLIFERATION OF NONCLASSICAL (2005)
LOGIC, NON-CLASSICAL (2005)
MANY-VALUED LOGICS (2005)
MOTION (2005)
PARACONSISTENT LOGICS (2005)
RELEVANCE (RELEVANT) LOGICS (2005)

Jesse Prinz
Associate Professor, Philosophy, University of North Carolina at Chapel Hill
CONCEPTS (2005)

A. N. Prior
Professor of Philosophy, University of Manchester; Coeditor of the Journal of Symbolic Logic; Fellow of the British Academy
CORRESPONCENCE THEORY OF TRUTH (1967)
EXISTENCE (1967)

LOGIC, HISTORY OF: MODERN LOGIC: THE BOOLEAN PERIOD [OVERVIEW]; JOHNSON; KEYNES; PEIRCE; THE HERITAGE OF KANT AND MILL (1967)
LOGIC, HISTORY OF: PRECURSORS OF MODERN LOGIC [OVERVIEW] (1967)
LOGIC, TRADITIONAL (1967)
NEGATION (1967)
RUSSELL, BERTRAND ARTHUR WILLIAM (1967)

Mary Prior
Co-author (with A. N. Prior), "Erotetic Logic," Philosophical Review (Vol. 64)
WHATELY, RICHARD (1967)

Duncan Pritchard
Reader in Philosophy, University of Stirling
RELEVANT ALTERNATIVES (2005)

Benjamin S. Pryor
Assistant Professor, Philosophy; Co-Director of the Program in Law and Social Thought, University of Toledo;
FOUCAULT, MICHEL (2005)

Stathis Psillos
Associate Professor, Philosophy and History of Science, University of Athens
PHILOSOPHY OF SCIENCE, HISTORY OF (2005)
SCIENTIFIC REALISM (2005)
UNDERDETERMINATION THESIS, DUHEM-QUINE THESIS (2005)

Joseph Pucci
Associate Professor of Classics and in the Program in Medieval Studies; Associate Professor of Comparative Literature, Brown University
CAROLINGIAN RENAISSANCE [BIBLIOGRAPHY] (2005)

Richard Purtill
Western Washington University
DIVINE COMMAND THEORIES OF ETHICS (2005)
LEWIS, C. S. (CLIVE STAPLES) (2005)
PHILOSOPHY OF RELIGION, HISTORY OF [ADDENDUM] (2005)
RELIGION AND MORALITY (2005)
THEISM, ARGUMENTS FOR AND AGAINST (1996)

Anthony Quinton
University Lecturer in Philosophy and Fellow of New College, Oxford University
KNOWLEDGE AND BELIEF (1967)
POPPER, KARL RAIMUND (1967)

Michael R. Rackett
Cary, NC
PELAGIUS AND PELAGIANISM (2005)

Diana Raffman
Professor, Philosophy, University of Toronto
MARCUS, RUTH BARCAN (1996)

Fazl-Ur- Rahman
Director of the Central Institute of Islamic Research (Karachi)
ISLAMIC PHILOSOPHY (1967)

Bjørn T. Ramberg
Universitetet i Oslo
DAVIDSON, DONALD (1996, 2005)

Albert G. Ramsperger
Professor of Philosophy, University of Wisconsin
CRITICAL REALISM (1967)

David M. Rasmussen
Professor, Philosophy, Boston College; Editor in Chief, Philosophy and Social Criticism
HABERMAS, JÜRGEN (2005)

Michael Rea
Associate Professor, Philosophy, University of Notre Dame
PLANTINGA, ALVIN (2005)

Miklós Rédei
Eötvös Loránd Tudományegyetem
COMMON CAUSE PRINCIPLE (2005)

Joan Wynn Reeves
Reader in Psychology, Bedford College, University of London
BINET, ALFRED (1967)

Marjorie E. Reeves
Vice-Principal and Fellow of St. Anne's College and University Lecturer, Oxford University
JOACHIM OF FIORE (1967)

Thomas Regan
Emeritus Professor of Philosophy, North Carolina State University
ANIMAL RIGHTS AND WELFARE (1996, 2005)

David A. Reidy
Assistant Professor, Philosophy, University of Tennessee, Knoxville
PHILOSOPHY OF LAW, PROBLEMS OF [ADDENDUM] (2005)

Nicholas Rescher
Professor of Philosophy and Associate Director of the Center for Philosophy of Science, University of Pittsburgh
LOGIC, HISTORY OF: LOGIC IN THE ISLAMIC WORLD (1967)

David Resnik
Bioethicist, National Institute of Environmental Health Sciences, National Institutes of Health
SCIENCE, RESEARCH ETHICS OF (2005)

Georges Rey
Professor, Philosophy, University of Maryland, College Park
BEHAVIORISM (2005)
FODOR, JERRY A. (2005)

Gretchen A. Reydams-Schils
Associate Professor, Program of Liberal Studies and Department of Philosophy, University of Notre Dame
MUSONIUS RUFUS (2005)

Nicholas V. Riasanovsky
Professor of History, University of California, Berkeley
FOURIER, FRANÇOIS MARIE CHARLES (1967)

Mark Richard
Professor and Chair, Philosophy, Tufts University
BELIEF ATTRIBUTIONS (1996)
NON-TRUTH-CONDITIONAL MEANING (2005)
PROPOSITIONS (1996)

Henry Richardson
Professor, Philosophy, Georgetown University
DEONTOLOGICAL ETHICS (2005)

Aaron Ridley
Professor of Philosophy, University of Southampton
WILDE, OSCAR FINGAL O'FLAHERTIE WILLS (2005)

Miles Rind
Independent scholar
ADDISON, JOSEPH (2005)
LONGINUS (PSEUDO)

Fritz-Joachim von Rintelen
Professor of Philosophy, University of Mainz
GEYSER, JOSEPH (1967)

Carolyn Ristau
Adjunct Associate Professor, Psychology, Barnard College
ANIMAL MIND (2005)

David B. Robinson
Lecturer in Greek, University of Edinburgh
XENOPHON (1967)

Jenefer Robinson
Professor of Philosophy, University of Cincinnati
AESTHETICS, PROBLEMS OF (2005)

Thomas Robischon
Associate Professor and Chair, Philosophy, Tuskegee Institute
HOLT, EDWIN BISSELL (1967)
MCGILVARY, EVANDER BRADLEY (1967)
MONTAGUE, WILLIAM PEPPERELL (1967)
NEW REALISM (1967)
PERRY, RALPH BARTON (1967)

Heiner Roetz
Professor for Chinese History and Philosophy, Faculty of East Asian Studies, Ruhr-University, Bochum, Germany
CONFUCIUS (2005)

Yosal Rogat
Associate Professor, Political Science, University of Chicago
LEGAL REALISM (1967)

Robin Rollinger
Research Editor, Katholieke Universiteit Leuven
LIPPS, THEODOR (2005)
PFÄNDER, ALEXANDER (2005)

Mark Rollins
Associate Professor and Chair, Philosophy, Washington University in St. Louis
IMAGERY, MENTAL (2005)

Patrick Romanell
H. Y. Benedict Professor of Philosophy, University of Texas, El Paso
ABBAGNANO, NICOLA (1967)

Grace G. Roosevelt
New York University
ROUSSEAU, JEAN-JACQUES [ADDENDUM] (1996)

Richard M. Rorty
Associate Professor of Philosophy, Princeton University
INTUITION (1967)
RELATIONS, INTERNAL AND EXTERNAL (1967)

Connie Rosati
Associate Professor, Philosophy, University of Arizona
BRANDT, R. B. (2005)

Philipp W. Rosemann
Associate Professor of Philosophy, University of Dallas
PETER LOMBARD [ADDENDUM] (2005)

Gideon Rosen
Professor of Philosophy, Princeton University
NOMINALISM, MODERN (2005)
REALISM [ADDENDUM] (1996)

Roger D. Rosenkrantz
Independent scholar
FISHER, R. A. (2005)
INFORMATION THEORY (2005)
STATISTICS, FOUNDATIONS OF (2005)

David M. Rosenthal
Professor of Philosophy and Coordinator of Cognitive Science, Graduate Center, The City University of New York
CONSCIOUSNESS (2005)

Adina Roskies
Assistant Professor, Philosophy, Dartmouth College
KITCHER, PATRICIA (2005)

James F. Ross
Professor of Philosophy and Law, Philosophy Department, University of Pennsylvania
ANALOGY IN THEOLOGY (2005)

Stephanie Ross
Professor, Philosophy, University of Missouri, St. Louis
ENVIRONMENTAL AESTHETICS (2005)

Ferruccio Rossi-Landi
Research Member of the Staff of the State University of Milan
CALDERONI, MARIO (1967)
CATTANEO, CARLO (1967)
DINGLER, HUGO (1967)
PEANO, GIUSEPPE (1967)
VAILATI, GIOVANNI (1967)

Christopher J. Rowe
Professor of Greek, Durham University
KALON (2005)

William L. Rowe
Professor of Philosophy, Purdue University
EVIL, THE PROBLEM OF [ADDENDUM] (1996, 2005)

Anthony Rudd
Visiting Assistant Professor of Philosophy, St. Olaf College
CALVIN, JOHN [ADDENDUM] (2005)

Richard S. Rudner
Professor and Chairman of the Department of Philosophy, Washington University (St. Louis); Editor in Chief of Philosophy of Science
GOODMAN, NELSON (1967)

Laura Ruetsche
Associate Professor, Philosophy, University of Pittsburgh
EARMAN, JOHN (2005)
STRING THEORY (2005)

T. S. Rukmani
Professor and Chair in Hindu Studies, Concordia University, Montreal, Canada
GOD IN INDIAN PHILOSOPHY (2005)

Bede Rundle
Fellow and Lecturer in Philosophy, Trinity College, Oxford University
LOGIC, HISTORY OF: MODERN LOGIC: FROM FREGE TO GÖDEL: BROUWER AND INTUITIONISM; FREGE; GÖDEL; HERBRAND (1967); HILBERT AND

FORMALISM (2005);
LÖWENHEIM; PEANO; POST;
RAMSEY; SKOLEM (1967);
WHITEHEAD AND RUSSELL
(2005)
LOGIC, HISTORY OF: MODERN
LOGIC: SINCE GÖDEL: CHURCH;
GENTZEN (1967)

Joseph Runzo
Professor, Philosophy and Religious Studies, Chapman University; Life Member, Clare Hall, University of Cambridge
LIFE, MEANING AND VALUE OF [ADDENDUM] (2005)

Michael Ruse
Lucyle T. Werkmeister Professor of Philosophy; Director of the Program in the History and Philosophy of Science, Department of Philosophy, Florida State University
EVOLUTIONARY ETHICS (1996, 2005)
HUMAN GENOME PROJECT (2005)
RELIGION AND THE BIOLOGICAL SCIENCES (2005)
WILSON, EDWARD O. (2005)

Bruce Russell
Professor and Chair of Philosophy, Wayne State University
INTUITION [ADDENDUM 2] (2005)

Cheyney Ryan
Professor, Philosophy Department, University of Oregon
BERLIN, ISAIAH (2005)

Todd Ryan
Assistant Professor, Trinity College, Hartford, CT
LE CLERC, JEAN [ADDENDUM] (2005)

Martin Ryder
Adjunct Professor, Information and Learning Technologies, University of Colorado, Denver
SOCIAL CONSTRUCTIONISM (2005)

Daniel Rynhold
Lecturer in Judaism, Department of Theology and Religious Studies, King's College, London
ALBO, JOSEPH [ADDENDUM] (2005)

David Rynin
Professor of Philosophy, University of California, Berkeley
JOHNSON, ALEXANDER BRYAN (1967)

Hassan Saab
Professor, Lebanese Unviersity and St. Joseph University of Beirut
IBN KHALDŪN (1967)

Marcelo Sabatés
Associate Professor and Head, Philosophy Department, Kansas State University, Manhattan
KIM, JAEGWON (2005)
REDUCTIONISM IN THE PHILOSOPHY OF MIND (2005)

Nathan Salmon
Professor, Philosophy, University of California, Santa Barbara
PROPER NAMES AND DESCRIPTIONS (2005)

Norbert Samuelson
Grossman Chair in Jewish Studies, Philosophy Department, Arizona State University
ROSENZWEIG, FRANZ [ADDENDUM] (2005)

David Sanford
Instructor in Philosophy, Dartmouth College
DEGREES OF PERFECTION, ARGUMENT FOR THE EXISTENCE OF GOD [BIBLIOGRAPHY] (1967)

Jonathan J. Sanford
Department of Philosophy, Franciscan University of Steubenville
PETER DAMIAN (2005)

Antonio Santucci
University Professor of the History of Modern and Contemporary Philosophy; Extraordinary Professor of the History of Philosophy, Faculty of Education, University of Bologna
PAPINI, GIOVANNI (1967)

Virginia Sapiro
University of Wisconsin, Madison
WOLLSTONECRAFT, MARY (1996)

Jennifer M. Saul
Professor of Philosophy, University of Sheffield
CONVERSATIONAL IMPLICATURE (2005)

Jason L. Saunders
Professor of Philosophy, University of California, San Diego
LIPSIUS, JUSTUS (1967)
PATRIZI, FRANCESCO (1967)

James P. Scanlan
Emeritus, Philosophy, Ohio State University
BELINSKII, VISSARION GRIGOR'EVICH (1967)
BULGAKOV, SERGEI NIKOLAEVICH (1967)
CHERNYSHEVSKII, NIKOLAI GAVRILOVICH (1967)
DOSTOEVSKY, FYODOR MIKHAILOVICH (2005)
KOZLOV, ALEKSEI ALEKSANDROVICH (1967)
LAPSHIN, IVAN IVANOVICH (1967)
LAVROV, PËTR LAVROVICH (1967)
LENIN, VLADIMIR IL'ICH (1967, 2005)
LOPATIN, LEV MIKHAILOVICH (1967)
MIKHAILOVSKII, NIKOLAI KONSTANTINOVICH (1967)
RADISHCHEV, ALEKSANDR NIKOLAEVICH (1967)
ROZANOV, VASILII VASIL'EVICH (1967)
TOLSTOY, LEV (LEO) NIKOLAEVICH (2005)
VYSHESLAVTSEV, BORIS PETROVICH (1967)

Eva Schaper
Lecturer in Logic and Aesthetic Philosophy, University of Glasgow
KAUFMANN, WALTER ARNOLD (2005)
PATER, WALTER HORATIO (1967)
TROELTSCH, ERNST (1967)

Kevin Schilbrack
Associate Professor of Religious Studies, Wesleyan College, Macon, GA
MYTH [ADDENDUM] (2005)

G. Schlesinger
Professor of Philosophy, University of North Carolina
BRIDGMAN, PERCY WILLIAM (1967)
OPERATIONALISM (1967)

Antonia Ruth Schlette
University of Munich
CHAMBERLAIN, HOUSTON STEWART (1967)

Tad M. Schmaltz
Professor of Philosophy, Duke University
ARNAULD, ANTOINE (2005)
CARTESIANISM [ADDENDUM] (2005)
CONDILLAC, ÉTIENNE BONNOT DE (2005)
DESGABETS, ROBERT (2005)
JANSENISM (2005)
MALEBRANCHE, NICOLAS (2005)
NICOLE, PIERRE (2005)
REGIUS, HENRICUS (HENRY DE ROY)

Dennis Schmidt
Professor of Philosophy, Comparative Literature, and German, Pennsylvania State University
GADAMER, HANS-GEORG (2005)
HERMENEUTICS (2005)

James Schmidt
Professor of History and Political Science, Boston University
ENLIGHTENMENT (2005)

David Schmidtz
Professor of Philosophy, Joint Professor of Economics; Director, Program in Philosophy of Freedom, University of Arizona
ETHICAL EGOISM [BIBLIOGRAPHY](2005)
NOZICK, ROBERT (2005)
PHILOSOPHY OF ECONOMICS (2005)

Frederick F. Schmitt
Professor of Philosophy, Indiana University
NATURALIZED EPISTEMOLOGY (2005)

Richard Schmitt
Associate Professor of Philosophy, Brown University
PHENOMENOLOGY (1967)

J. B. Schneewind
Associate Professor of Philosophy, University of Pittsburgh
ELIOT, GEORGE (1967)
GROTE, JOHN (1967)
MARTINEAU, JAMES (1967)
MCTAGGART, JOHN MCTAGGART ELLIS (1967)
MILL, JOHN STUART (1967)
SIDGWICK, HENRY (1967)
STEPHEN, LESLIE (1967)

Malcolm Schofield
Professor of Ancient Philosophy, Faculty of Classics, University of Cambridge
ZENO OF CITIUM (2005)

Philip Schofield
Professor of the History of Legal and Political Thought, Faculty of Law, University College London
BENTHAM, JEREMY (2005)

Martin Schönfeld
Associate Professor, Philosophy, Core Faculty, Environmental Science and Policy, University of South Florida
WOLFF, CHRISTIAN (2005)

Alan D. Schrift
Professor of Philosophy; Director, Center for the Humanities, Grinnell College
DECONSTRUCTION (2005)
NIETZSCHE, FRIEDRICH (2005)
STRUCTURALISM AND POST-STRUCTURALISM (1996, 2005)

Mark Schroeder
Assistant Professor of Philosophy, University of Maryland, College Park
ETHICAL NATURALISM [ADDENDUM] (2005)

Oliver Schulte
Associate Professor, Philosophy, Simon Fraser University
SCIENTIFIC METHOD (2005)

R. Barton Schultz
Fellow and Lecturer, Division of the Humanities, University of Chicago
NUSSBAUM, MARTHA (2005)
SIDGWICK, HENRY [ADDENDUM] (1996, 2005)

George Schumm
Associate Professor, Philosophy, Ohio State University
MARCUS, RUTH BARCAN (1996)

Joachim Schummer
Editor, HYLE: International Journal for Philosophy of Chemistry; Heisenberg-Fellow, Philosophy, University of Darmstadt; Adjunct Professor, Philosophy, University of South Carolina
CHEMISTRY, PHILOSOPHY OF (2005)

Charles E. Scott
Distinguished Professor of Philosophy and Director of the Vanderbilt Center for Ethics, Vanderbilt University
HEIDEGGER, MARTIN (2005)

Dion Scott-Kakures
Professor of Philosophy, Scripps College
FOLK PSYCHOLOGY (1996, 2005)
SELF-DECEPTION (2005)

William Seager
Professor of Philosophy, University of Toronto at Scarborough
DRETSKE, FRED (2005)

John Searle
Slusser Professor of Philosophy, University of California, Berkeley
DETERMINABLES AND DETERMINATES (1967)
STRAWSON, PETER FREDERICK (1967)

Krister Segerberg
Visiting Professor of Philosophy, Stanford University
MODAL LOGIC (2005)
WRIGHT, GEORG HENRIK VON (1996, 2005)

Svetlana Semënova
Professor, Institute of World Literature, Moscow
FËDOROV, NIKOLAI FËDOROVICH (2005)

Mikhail Yu. Sergeev
Adjunct Associate Professor, Religion, University of the Arts, Philadelphia
LOSSKII, NIKOLAI ONUFRIEVICH (2005)
ZEN'KOVSKII, VASILII VASIL'EVICH (2005)

Bogdan Šešić
Ordinary Professor of Logic, University of Belgrade
PETRONIEVIĆ, BRANISLAV (1967)

Russ Shafer-Landau
Professor of Philosophy, University of Wisconsin, Madison
RATIONALISM IN ETHICS [ADDENDUM] (2005)

Scott A. Shalkowski
Lecturer, Philosophy, University of Leeds
MODALITY, PHILOSOPHY AND METAPHYSICS OF (2005)

Brian Shanley
Catholic University of America
THOMISM [ADDENDUM] (2005)

Lisa Shapiro
Associate Professor of Philosophy, Simon Fraser University
ELISABETH, PRINCESS OF BOHEMIA (2005)

Stewart Shapiro
O'Donnell Professor of Philosophy, The Ohio State University; Arché Professorial Fellow, University of St. Andrews
REALISM AND NATURALISM, MATHEMATICAL (2005)

Arvind Sharma
Birks Professor of Comparative Religion, McGill University
SELF IN INDIAN PHILOSOPHY (2005)

Vincent Shen
Lee Chair in Chinese Thought and Culture, Department of Philosophy and Department of East Asian Studies, University of Toronto
CHINESE PHILOSOPHY: METAPHYSICS AND EPISTEMOLOGY (2005)
YANG ZHU (2005)

Anne D. R. Sheppard
Senior Lecturer, Classics, Royal Holloway, University of London
ANCIENT AESTHETICS (2005)
PHANTASIA (2005)

Gila Sher
University of California, San Diego
LOGICAL TERMS (1996)

Nancy Sherman
University Professor in Philosophy, Adjunct Professor in Law, Georgetown University
FRIENDSHIP (2005)

Michael Shermer
Founding Publisher of Skeptic magazine; Director of the Skeptics Society; Columnist for Scientific American; Host of the Skeptics Distinguished Science Lecture Series, California Institute of Technology
SCIENCE AND PSEUDOSCIENCE (2005)

Sanford Shieh
Wesleyan University
LOGICAL KNOWLEDGE (1996)

Vincent Y. C. Shih
Professor of Chinese Philosophy and Literature, University of Washington
HU SHI (1967)

J. M Shorter
Professor of Philosophy, University of Canterbury (New Zealand)
OTHER MINDS (1967)

Kwong-loi Shun
Professor, Philosophy and East Asian Studies, University of Toronto
CHINESE PHILOSOPHY: CONFUCIANISM (2005)
MENCIUS (2005)

Alan Sidelle
University of Wisconsin, Madison
CONSTRUCTIVISM AND CONVENTIONALISM (1996)

David Sider
Professor of Classics, New York University
SIMPLICIUS (2005)

Mark Siderits
Professor, Philosophy, Illinois State University, Normal
BUDDHISM—SCHOOLS: MADHYAMAKA (2005)

Wilfried Sieg
Professor, Philosophy, Carnegie Mellon University
COMPUTING MACHINES (2005)

Harvey Siegel
Professor, Philosophy, University of Miami
PHILOSOPHY OF EDUCATION, HISTORY OF: CONTEMPORARY ISSUES: EPISTEMOLOGICAL (2005)

Hugh J. Silverman
State University of New York at Stony Brook
MODERNISM AND POSTMODERNISM (1996)

Anita Silvers
Professor, Philosophy, San Francisco State University
DANTO, ARTHUR (2005)

Keith Simmons
Professor of Philosophy, University of North Carolina at Chapel Hill
TRUTH (2005)

Lawrence H. Simon
Associate Professor, Philosophy and Environmental Studies, Bowdoin College
MARX, KARL (2005)

W. M. Simon
Professor of History, University of Keele
FOUILLÉE, ALFRED (1967)
LAAS, ERNST (1967)
LITTRÉ, ÉMILE (1967)
RENAN, JOSEPH ERNEST (1967)

Peter Simons
Professor of Philosophy, School of Philosophy, University of Leeds
LEŚNIEWSKI, STANISŁAW (2005)
MEREOLOGY (1996)

Marcus G. Singer
Professor and Chair, Department of Philosophy, University of Wisconsin; Chair of the Department of Philosophy of the University of Wisconsin Center System
GOLDEN RULE (1967)

Georgette Sinkler
University of Illinois at Chicago
HEYTESBURY, WILLIAM (2005)

Walter Sinnott-Armstrong
Professor of Philosophy, Hardy Professor of Legal Studies, Dartmouth College
MORAL DILEMMAS (2005)

John Sisko
Assistant Professor, Philosophy, The College of New Jersey

NOUS (2005)

Lawrence Sklar
Carl G. Hempel and William K. Frankena Distinguished University Professor of Philosophy, University of Michigan, Ann Arbor
BOLTZMANN, LUDWIG (2005)
GIBBS, JOSIAH (2005)
PHILOSOPHY OF STATISTICAL MECHANICS (2005)
PHYSICS AND THE DIRECTION OF TIME (1996, 2005)

Henryk Skolimowski
Associate Professor of Philosophy, University of Southern California
INGARDEN, ROMAN (1967)

J. J. C. Smart
Emeritus Professor, Australian National University
SPACE (1967)
TIME (1967, 2005)
UTILITARIANISM (1967)

Ninian Smart
H. G. Wood Professor of Theology, University of Birmingham
BARTH, KARL (1967)
BOEHME, JAKOB (1967)
ECKHART, MEISTER (1967)
HÜGEL, BARON FRIEDRICH VON (1967)
JOHN OF THE CROSS, ST. (1967)
KARMA (1967)
MYSTICISM, HISTORY OF (1967)
REINCARNATION (1967)
RUYSBROECK, JAN VAN (1967)
SUSO, HEINRICH (1967)
TAULER, JOHANNES (1967)
TERESA OF ÁVILA, ST. (1967)
THOMAS À KEMPIS (1967)
ZABARELLA, JACOPO (1967)
ZOROASTRIANISM (1967)

Andrew Smith
Professor of Classics, University College Dublin
PORPHYRY [BIBLIOGRAPHY] (2005)

Barry C. Smith
University of of London, England
LANGUAGE (1996)

Brent Smith
Claremont Graduate University
ASCETICISM [BIBLIOGRAPHY] (2005)

Colin Smith
Reader in French, University of London
BACHELARD, GASTON (1967)
HAMELIN, OCTAVE (1967)
JANKÉLÉVITCH, VLADIMIR (1967)
LALANDE, ANDRÉ (1967)
LE ROY, ÉDOUARD (1967)
LE SENNE, RENÉ (1967)
LOISY, ALFRED (1967)
MODERNISM (1967)
MOUNIER, EMMANUEL (1967)

Daniel W. Smith
Associate Professor, Philosophy, Purdue University
DELEUZE, GILLES (2005)

James Ward Smith
Professor of Philosophy, Princeton University
STACE, WALTER TERENCE (1967)

John E. Smith
Professor of Philosophy, Yale University; General Editor of the Yale Edition of Works of Jonathan Edwards
ROYCE, JOSIAH (1967)

Murray Smith
Professor of Film Studies, University of Kent
PHILOSOPHY OF FILM (2005)

Nicholas D. Smith
James F. Miller Professor of Humanities, Lewis and Clark College
LEHRER, KEITH (2005)

Quentin Smith
Western Michigan University
TIME, BEING, AND BECOMING (1996, 2005)

Howard E. Smokler
Visiting Associate Professor of Philosophy, Stanford University
CLIFFORD, WILLIAM KINGDON (1967)

Paul F. Snowdon
Grote Professor of Mind and Logic, University College London
RYLE, GILBERT [ADDENDUM] (2005)

Scott Soames
Professor, School of Philosophy, University of Southern California
ANALYSIS, PHILOSOPHICAL (2005)
ENTAILMENT, PRESUPPOSITION, AND IMPLICATURE (2005)
PROPOSITIONAL ATTITUDES: ISSUES IN SEMANTICS (2005)

Alan Soble
Professor of Philosophy and University Research Professor, University of New Orleans
PHILOSOPHY OF SEX (2005)

Miriam Solomon
Professor of Philosophy; Director of Graduate Studies, Temple University
LOGIC, HISTORY OF: MODERN LOGIC: SINCE GÖDEL: FRIEDMAN AND REVERSE (2005)
REVERSE MATHEMATICS (2005)
SOCIAL EPISTEMOLOGY (1996, 2005)

Mary Sommers
Director, Center for Thomistic Studies, University of St. Thomas, Houston, TX
BURLEY, WALTER (2005)

Roy Sorensen
Professor, Philosophy, Dartmouth College
KNOWLEDGE AND VAGUENESS (2005)
VAGUENESS (1996, 2005)

David Sosa
Associate Professor of Philosophy, University of Texas at Austin
HARMAN, GILBERT (2005)

Elmer Sprague
Associate Professor of Philosophy, City University of New York, Brooklyn College
BALGUY, JOHN (1967)
GAY, JOHN (1967)
HARTLEY, DAVID (1967)
HOME, HENRY (1967)
HUTCHESON, FRANCIS (1967)
MANDEVILLE, BERNARD (1967)
MORAL SENSE (1967)
PALEY, WILLIAM (1967)
PRICE, RICHARD (1967)
SMITH, ADAM (1967)

Joke Spruyt
Senior University Lecturer, Faculty of Arts and Culture, Maastricht University
PETER OF SPAIN (2005)

Ilja Srubar
Full Professor, Sociology, University of Erlangen
SCHUTZ, ALFRED (2005)

J. F. Staal
Professor of General and Comparative Philosophy and Director of the Instituut voor Filosofie, University of Amsterdam
LOGIC, HISTORY OF [OVERVIEW] (1967)

Werner Stark
Professor of Sociology, Fordham University
MANNHEIM, KARL (1967)
SOCIOLOGY OF KNOWLEDGE (1967)

Robert Stecker
Professor of Philosophy, Central Michigan University
ART, ONTOLOGY OF (2005)
LITERATURE, PHILOSOPHY OF (2005)

Carlos Steel
Professor, Philosophy, Catholic University of Leuven
DAMASCIUS (2005)
ERIGENA, JOHN SCOTUS (2005)
NEOPLATONISM (2005)
PROCLUS (2005)

Warren E. Steinkraus
Professor of Philosophy, State University of New York at Oswego
CREIGHTON, JAMES EDWIN (1967)

James P. Sterba
Professor of Philosophy, University of Notre Dame
DISTANT PEOPLES AND FUTURE GENERATIONS (1996)
GEWIRTH, ALAN (2005)
JUST WAR THEORY (2005)
LIBERTARIANISM (2005)
SOCIAL AND POLITICAL PHILOSOPHY (1996, 2005)
SOCIAL CONTRACT [ADDENDUM] (2005)
TERRORISM (2005)

J. P. Stern
Fellow and Tutor of St. John's College and University Lecturer in German, Cambridge University
BENN, GOTTFRIED (1967)
KAFKA, FRANZ (1967)
RILKE, RAINER MARIA (RENÉ) (1967)

Matthias Steup
Professor of Philosophy, St. Cloud State University
RELIABILISM (2005)
SOSA, ERNEST (2005)

Leslie Stevenson
Honorary Reader in Philosophy, University of St. Andrews
HUMAN NATURE (2005)

M. A. Stewart
Senior Research Fellow, Harris Manchester College, Oxford,U.K
STILLINGFLEET, EDWARD (2005)

Anfinn Stigen
Docent, Institute of Philosophy and History of Ideas, University of Oslo
TRESCHOW, NIELS (1967)

Allan Stoekl
Professor of French and Comparative Literature, Pennsylvania State University
BATAILLE, GEORGES (2005)
BLANCHOT, MAURICE (2005)
EXISTENTIALISM [ADDENDUM] (2005)

Andrija Stojković
Professor of Philosophy, University of Belgrade
MARKOVIĆ, SVETOZAR (1967)

Daniel Stoljar
Senior Fellow, Philosophy Program, Research School of Social Sciences (RSSS), Australian National University
MENTAL-PHYSICAL DISTINCTION (2005)

Jerome Stolnitz
Professor of Philosophy, University of Rochester
BEAUTY (1967)
UGLINESS (1967)

Leonid Stolovich
Professor Emeritus of Tartu University (Estonia)
LOTMAN, IURII MIKHAILOVICH (2005)

A. K. Stout
Professor Emeritus and Fellow of the Senate, University of Sydney
CAIRD, EDWARD (1967)
HOBHOUSE, LEONARD TRELAWNEY (1967)
PRINGLE-PATTISON, ANDREW SETH (1967)
RASHDALL, HASTINGS (1967)
ROSS, WILLIAM DAVID (1967)

D. Stove
Senior Lecturer, University of Sydney
KEYNES, JOHN MAYNARD (1967)

Michael W. Strasser
Associate Professor of Philosophy, Duquesne University
LIBER DE CAUSIS (1967)

Tony Street
Assistant Director of Research in Islamic Studies, Faculty of Divinity, University of Cambridge
LOGIC, HISTORY OF: LOGIC IN THE ISLAMIC WORLD [ADDENDUM] (2005)

Michael Strevens
Associate Professor, Philosophy, New York University
BAYES, BAYES' THEOREM, BAYESIAN APPROACH TO PHILOSOPHY OF SCIENCE (2005)
CHAOS THEORY (2005)
EXPLANATION (2005)
PROBABILITY AND CHANCE (2005)

Avrum Stroll
Professor of Philosophy, University of California, San Diego
PRESUPPOSING (1967)

Fred Gillette Sturm
Professor of Philosophy, Western College for Women (Ohio); Visiting Professor of Philosophy, Eastern Indiana Center of Indiana University
FARIAS BRITO, RAIMUNDO DE (1967)
MOLINA GARMENDIA, ENRIQUE (1967)
REALE, MIGUEL (1967)
VARONA Y PERA, ENRIQUE JOSÉ (1967)

Kathleen M. Sullivan
Stanford University
CENSORSHIP (1996)

L. W. Sumner
University Professor, Philosophy, University of Toronto
HAPPINESS (2005)

Edward Surtz, S.J.
Professor of English, Loyola University
MORE, THOMAS (1967)

John Sutton
Associate Professor, Philosophy, Macquarie University
MEMORY (2005)

Edith D. Sylla
Professor of History, North Carolina State University, Raleigh
SWINESHEAD, RICHARD (2005)

Zoltan G. Szabo
Associate Professor, Philosophy, Cornell University
COMPOSITIONALITY (2005)
SYNTACTICAL AND SEMANTICAL CATEGORIES [ADDENDUM] (2005)

Charles Taliaferro
Professor of Philosophy, St. Olaf College
DEATH [ADDENDUM] (2005)
IDEAL OBSERVER THEORIES OF ETHICS (2005)
NAGEL, THOMAS (2005)

Frank Talmage
Assistant Professor of Hebrew Studies, University of Wisconsin
CRESCAS, HASDAI (1967)
GERSONIDES (1967)

Scott Tanona
Assistant Professor, Philosophy, Kansas State University
BOHR, NIELS (2005)

C. C. W. Taylor
Emeritus Professor of Philosophy, Oxford University, and Emeritus Fellow of Corpus Christi College
LEUCIPPUS AND DEMOCRITUS [BIBLIOGRAPHY] (2005)
SOCRATES (2005)

Richard Taylor
Professor of Philosophy, University of Rochester
DETERMINISM, A HISTORICAL SURVEY (1967)
VOLUNTARISM (1967)

Paul Teller
Professor, University of California, Davis
UNITY AND DISUNITY OF SCIENCE (2005)

Larry S. Temkin
Professor of Philosophy, Rutgers, the State University of New Jersey
EQUALITY, MORAL AND SOCIAL [ADDENDUM] (2005)
PARFIT, DEREK (2005)

Roland J. Teske
Professor of Philosophy, Marquette University
HERVAEUS NATALIS (2005)

Paul Thagard
Professor, University Research Chair, Philosophy, University of Waterloo
PSYCHOLOGY [ADDENDUM] (2005)

Irving Thalberg
Assistant Professor of Philosophy, University of Illinois
ERROR (1967)

Peter Thielke
Assistant Professor, Philosophy, Pomona College
MAIMON, SALOMON (2005)
NAGEL, ERNEST (1967)
PRAGMATISM (1967)

Johannes M. M. H. Thijssen
Professor of Ancient and Medieval Philosophy, Radboud University Nijmegen
NICOLAS OF AUTRECOURT [ADDENDUM] (2005)

Paul Thom
Professor; Executive Dean, Faculty of Arts, Southern Cross University
AUGUSTINE, ST. [ADDENDUM 2] (2005)
PLOTINUS [ADDENDUM] (2005)

Ivo Thomas
Visiting Professor, Notre Dame University and Ohio State University
LOGIC, HISTORY OF: MODERN LOGIC: FROM FREGE TO GÖDEL: NINETEENTH-CENTURY MATHEMATICS (2005)
LOGIC, HISTORY OF: PRECURSORS OF MODERN LOGIC: EULER (1967)
LOGIC, HISTORY OF: PRECURSORS OF MODERN LOGIC: LAMBERT AND PLOUCQUET (1967)
LOGIC, HISTORY OF: PRECURSORS OF MODERN LOGIC: LEIBNIZ (1967)
LOGIC, HISTORY OF: THE INTERREGNUM (BETWEEN MEDIEVAL AND MODERN) (1967)

Richmond Thomason
University of Michigan
ARTIFICIAL AND NATURAL LANGUAGES (2005)
MONTAGUE, RICHARD (2005)

Manley Thompson
Professor and Chair, Philosophy, University of Chicago
CATEGORIES (1967)

Judith Jarvis Thomson
Associate Professor of Philosophy, Massachusetts Institute of Technology
WISDOM, (ARTHUR) JOHN TERENCE DIBBEN (1967)

S. Harrison Thomson
Professor Emeritus of History, University of Colorado
HUS, JOHN (1967)
SCOT, MICHAEL (1967)
WYCLYF, JOHN (1967)

Mark Timmons
University of Memphis
CONSTRUCTIVISM, MORAL (1996)

Robert B. Todd
Professor of Classics, University of British Columbia
THEMISTIUS (2005)

Vincent Tomas
Professor of Philosophy, Brown University
DUCASSE, CURT JOHN (1967)

Giorgio Tonelli
Professor of History of German Literature and History of Modern Philosophy, University of Pisa
BASEDOW, JOHANN BERNHARD (1967)
BAUMGARTEN, ALEXANDER GOTTLIEB (1967)
BILFINGER, GEORG BERNHARD (1967)
BUDDE, JOHANN FRANZ (1967)
CRUSIUS, CHRISTIAN AUGUST (1967)
EBERHARD, JOHANN AUGUST (1967)

GARVE, CHRISTIAN (1967)
KNUTZEN, MARTIN (1967)
LAMBERT, JOHANN HEINRICH (1967)
LAVATER, JOHANN KASPAR (1967)
MEIER, GEORG FRIEDRICH (1967)
MENDELSSOHN, MOSES (1967)
NICOLAI, CHRISTIAN FRIEDRICH (1967)
PLOUCQUET, GOTTFRIED (1967)
REIMARUS, HERMANN SAMUEL (1967)
RÜDIGER, ANDREAS (1967)
SULZER, JOHANN GEORG (1967)
TETENS, JOHANN NICOLAUS (1967)
THOMASIUS, CHRISTIAN (1967)
THÜMMIG, LUDWIG PHILIPP (1967)
TSCHIRNHAUS, EHRENFRIED WALTER VON (1967)
WINCKELMANN, JOHANN JOACHIM (1967)

Michael Tooley
Professor of Philosophy, University of Colorado, Boulder
CAUSAL APPROACHES TO THE DIRECTION OF TIME (2005)
CAUSATION: METAPHYSICAL ISSUES (1996, 2005)
METAPHYSICS, NATURE OF [ADDENDUM] (2005)

Roberto Torretti
Professor Emeritus, Philosophy, Universidad de Puerto Rico en Río Piedras
GEOMETRY (2005)

Norman L. Torrey
Emeritus Professor of French, Columbia University
DIDEROT, DENIS (1967)

Stephen E. Toulmin
Professor of History of Ideas and Philosophy, Brandeis University
MATTER (1967)

Dabney Townsend
Professor of Philosophy, Armstrong Atlantic State University
ALISON, ARCHIBALD (2005)
GERARD, ALEXANDER (2005)

Knut Erik Tranöy
Professor of Philosophy, University of Bergen (Norway)
STEFFENS, HENRICH (1967)

Charles Travis
Sterling University, Scotland
PUTNAM, HILARY (1996)

Simon Trépanier
Lecturer, Classics, School of History and Classics, University of Edinburgh
EMPEDOCLES [ADDENDUM] (2005)

Voula Tsouna
Associate Professor of Philosophy, University of California, Santa Barbara
PHILODEMUS (2005)

Nancy Tuana
Director, Rock Ethics Institute, Pennsylvania State University
FEMINISM AND THE HISTORY OF PHILOSOPHY (2005)

Roderich Tumulka
Wissenschaftlicher Assistent, Mathematics, Eberhard-Karls-Universität, Tübingen, Germany
BELL, JOHN, AND BELL'S THEOREM (2005)
BOHM, DAVID (2005)
BOHMIAN MECHANICS (2005)

Paul Turner
Lecturer in English and Fellow of Linacre College, Oxford University of London; Professor of English, Ankara University
LUCIAN OF SAMOSATA (1967)

Anna-Teresa Tymieniecka
Diplôme d'Études Supérieures (University of Paris, Sorbonne)
LAVELLE, LOUIS (1967)

Shizuteru Ueda
Professor Emeritus, Kyoto University
NISHIDA, KITARŌ (2005)

Robert Ulich
Professor of Education, Emeritus, Harvard University
APPERCEPTION (1967)
COMENIUS, JOHN AMOS (1967)
FROEBEL, FRIEDRICH (1967)
PESTALOZZI, JOHANN HEINRICH (1967)

Mark T. Unno
Associate Professor of East Asian Religions, Department of Religious Studies, University of Oregon, Eugene
SHINRAN (2005)

James O. Urmson
Fellow of Corpus Christi College, Oxford University
AUSTIN, JOHN LANGSHAW (1967)
IDEAS (1967)
RYLE, GILBERT (1967)

John Douglas Uytman
Senior Lecturer, Department of Psychiatry and Lecturer in Medical Psychology, St. Andrews University
ADLER, ALFRED (1967)
MCDOUGALL, WILLIAM (1967)
PAVLOV, IVAN PETROVICH (1967)

Jouko Väänänen
Professor, Mathematics, University of Helsinki
HINTIKKA, JAAKKO (2005)

Ezio Vailati
Professor, Philosophy, Southern Illinois University, Edwardsville
CLARKE, SAMUEL (2005)

Mark van Atten
Institut d'Histoire et de Philosophie des Sciences et des Techniques (CNRS/Paris 1/ENS)
BROUWER, LUITZEN EGBERTUS JAN [ADDENDUM] (2005)

James Van Cleve
Brown University
REID, THOMAS (2005)

Dirk van Dalen
Professor, Philosophy, Utrecht University
INTUITIONISM AND INTUITIONISTIC LOGIC (2005)

Philippe van Haute
Full Professor in Philosphical Anthropology, Radboud University (The Netherlands); Psychoanalyst (private practice), Leuven, Belgium
LACAN, JACQUES (2005)

Peter van Inwagen
John Cardinal O'Hara Professor of Philosophy, The University of Notre Dame
ONTOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD [ADDENDUM] (2005)

Andrew G. M. Van Melsen
Professor of Philosophy, University of Nijmegen
ATOMISM (1967)

Linda Van Norden
Professor of English, University of California, Davis
PARACELSUS (1967)

Mark van Roojen
University of Nebraska, Lincoln
ETHICAL SUBJECTIVISM [BIBLIOGRAPHY] (2005)

Aram Vartanian
Professor of French, New York University
HELVÉTIUS, CLAUDE-ADRIEN (1967)
HOLBACH, PAUL-HENRI THIRY, BARON D' (1967)
LA METTRIE, JULIEN OFFRAY DE (1967)
MAUPERTUIS, PIERRE-LOUIS MOREAU DE (1967)
MESLIER, JEAN (1967)
STAHL, GEORG ERNST (1967)

Pekka Vayrynen
Assistant Professor, Philosophy, University of California, Davis
MORAL REALISM (2005)

Theo Verbeek
Professor, History of Modern Philosophy, Utrecht University
GEULINCX, ARNOLD [ADDENDUM] (2005)

Rineke Verbrugge
University of Göteborg, Sweden
PROVABILITY LOGIC (1996)

Pieter Vermaas
Researcher, Philosophy, Delft University of Technology
MODAL INTERPRETATION OF QUANTUM MECHANICS (2005)

G. N. A. Vesey
Reader in Philosophy, University of London; Honorary Director of The Royal Institute of Philosophy
SOUND (1967)
TOUCH (1967)

Olga Volkogonova
Professor, Moscow State University
BERDYAEV, NIKOLAI ALEKSANDROVICH (2005)
KIREEVSKII, IVAN VASIL'EVICH (2005)

Georg Henrik von Wright
Professor-at-Large, Cornell University
LICHTENBERG, GEORG CHRISTOPH (1967)

Henry Vyverberg
Associate Professor of History, University of Akron
TURGOT, ANNE ROBERT JACQUES, BARON DE L'AULNE (1967)

William J. Wainwright
University of Wisconsin, Milwaukee
OTTO, RUDOLF (1967)
RELIGIOUS EXPERIENCE (1996)

Mary Ellen Waithe
Professor, Philosophy, Cleveland State University
HYPATIA (1996, 2005)

Rebecca L. Walker
Assistant Professor, Social Medicine; Adjunct Assistant Professor, Philosophy, University of North Carolina, Chapel Hill
BIOETHICS (2005)

R. Jay Wallace
Professor of Philosophy, University of California, Berkeley
PROMISES (2005)

Jerry L. Walls
Professor of Philosophy of Religion, Asbury Seminary, Wilmore, KY
HEAVEN AND HELL, DOCTRINES OF (2005)

W. H. Walsh
Professor of Logic and Metaphysics, University of Edinburgh
GREEN, THOMAS HILL (1967)
KANT, IMMANUEL (1967)
METAPHYSICS, NATURE OF (1967)

Kendall Walton
Charles L. Stevenson Collegiate Professor of Philosophy, University of Michigan
NONEXISTENT OBJECT, NONBEING (1996, 2005)

W. J. Waluchow
Professor, Philosophy, McMaster University
HART, HERBERT LIONEL ADOLPHUS (2005)

Richard Walzer
Honorary Professor and Reader, Oxford University
AL-KINDĪ, ABŪ-YŪSUF YAʿQŪB IBN ISḤĀQ (1967)

Richard Warner
Professor of Law, Chicago-Kent College of Law, Illinois Institute of Technology; Professor and Chair of American and Comparative Law, Catholic University of Lublin, Poland
GRICE, HERBERT PAUL (1996, 2005)

G. J. Warnock
Fellow and Tutor of Philosophy, Magdalen College, Oxford University
REASON (1967)

James Warren
University Lecturer in Classics, Corpus Christi College, Cambridge
EPICUREANISM AND THE EPICUREAN SCHOOL [BIBLIOGRAPHY] (2005)

Gary Watson
Professor of Philosophy, University of California, Riverside
FOOT, PHILIPPA (2005)

Richard A. Watson
Philosophy Professor Emeritus, Washington University; Philosophy Faculty Affiliate, University of Montana
FOUCHER, SIMON (1967, 2005)
RÉGIS, PIERRE-SYLVAIN (1967, 2005)
ROHAULT, JACQUES (1967, 2005)

W. Montgomery Watt
Professor and Head, Department of Islamic Studies, University of Edinburgh
AL-FĀRĀBĪ (1967)

Wayne Waxman
Independent scholar
HUME, DAVID (2005)

A. Wedberg
Professor of Philosophy, University of Stockholm

BOSTRÖM, CHRISTOPHER JACOB (1967)
HÄGERSTRÖM, AXEL (1967)

Laura Weed
Associate Professor of Philosophy, The College of St. Rose
RELIGION, PSYCHOLOGICAL EXPLANATIONS OF [ADDENDUM] (2005)

Kai F. Wehmeier
Associate Professor, Logic & Philosophy of Science, University of California, Irvine
FREGE, GOTTLOB (2005)

Vivian Weil
Professor of Ethics; Director, Center for the Study of Ethics in the Professions, Illinois Institute of Technology
ENGINEERING ETHICS (2005)

Gerson Weiler
Research Fellow, Philosophy, Institute of Advanced Studies, Australian National University
MAXWELL, JAMES CLERK (1967)

Julius R. Weinberg
Vilas Professor of Philosophy; Member of the Institute of Research in the Humanities, University of Wisconsin
NICOLAS OF AUTRECOURT (1967)

Kurt Weinberg
Professor of French, and Comparative Literature, University of Rochester
PANTHEISMUSSTREIT (1967)

Rudolph H. Weingartner
Associate Professor and Chair, Philosophy, San Francisco State College
SIMMEL, GEORG (1967)

James A. Weisheipl, O.P.
Professor of Medieval Thought, School of Philosophy, Aquinas Institute (River Forest, IL)
DURANDUS OF SAINT-POURÇAIN (1967)
JOHN OF PARIS (1967)
ULRICH (ENGELBERT) OF STRASBOURG (1967)

John Weiss
Associate Professor of History, Wayne State University
HESS, MOSES (1967)
KAUTSKY, KARL (1967)
SABATIER, AUGUSTE (1967)

Paul Weithman
Professor, Philosophy, University of Notre Dame
REPUBLICANISM (2005)
SOVEREIGNTY [ADDENDUM] (2005)

Morris Weitz
Professor of Philosophy, Ohio State University
PROUST, MARCEL (1967)

Albert Wellek
Professor and Director, Department of Psychology, University of Mainz
KRUEGER, FELIX (1967)
WUNDT, WILHELM (1967)

René Wellek
Sterling Professor and Chair, Department of Comparative Literature, Yale
MASARYK, TOMÁŠ GARRIGUE (1967)

Carl Wellman
Professor Emeritus, Washington University, St. Louis
ASCETICISM (1967)
RIGHTS (2005)

Peter Wenz
Emeritus Professor of Philosophy, University of Illinios at Springfield
ENVIRONMENTAL ETHICS (1996, 2005)

William Werkmeister
Director of the School of Philosophy, University of Southern California
DRIESCH, HANS ADOLF EDUARD (1967)

Karel Werner
Honorary Professorial Associate, Department of the Study of Religions, School of Oriental and African Studies, University of London
LIBERATION IN INDIAN PHILOSOPHY (2005)

Charles C. West
Professor Emeritus of Christian Ethics, Princeton Theological Seminary
CONSEQUENTIALISM (1996)
LIBERATION THEOLOGY (1996, 2005)

Henry R. West
Professor of Philosophy, Macalester College
MILL, JAMES [BIBLIOGRAPHY] (2005)
MILL, JOHN STUART [ADDENDUM] (1996, 2005)

Robin L. West
Professor of Law, Georgetown University Law Center
FEMINIST LEGAL THEORY (1996, 2005)

Willem. G. Weststeijn
Professor of Slavic Literatures, University of Amsterdam
EURASIANISM (2005)
LEONT'EV, KONSTANTIN NIKOLAEVICH (2005)

Rebecca Whisnant
Assistant Professor, Philosophy, University of Dayton
FEMINIST ETHICS (2005)

Alan R. White
Professor of Philosophy, University of Hull; Honorary Secretary of the Mind Association
COHERENCE THEORY OF TRUTH (1967)

Hayden V. White
Professor of History, University of Rochester
FEUERBACH, LUDWIG ANDREAS (1967)
GOBINEAU, COMTE JOSEPH ARTHUR DE (1967)
STRAUSS, DAVID FRIEDRICH (1967)
WINDELBAND, WILHELM (1967)

Roger White
Assistant Professor, Philosophy, New York University
ANTHROPIC PRINCIPLE, THE (2005)

Stephen A. White
Professor, Classics and Philosophy, University of Texas at Austin

CICERO, MARCUS TULLIUS [BIBLIOGRAPHY] (2005)
POSIDONIUS [ADDENDUM] (2005)
THALES OF MILETUS (2005)

Jacques J. Whitfield
Doctor of Law, Doctor of Economics, University of Würzburg, and other degrees; Author of articles on Michael Servetus and on the economic history of the 18th century
FRANCK, SEBASTIAN (1967)

Lancelot Law Whyte
Author, Next Development in Man; Unconscious Before Freud; Internal Factors in Evolution; Unitary Principle in Physics and Biology; Focus and Diversions; Accent on Form
BOSCOVICH, ROGER JOSEPH (1967)
UNCONSCIOUS (1967)

Henry Nelson Wieman
Professor Emeritus, University of Chicago; Distinguished Visiting Professor, Southern Illinois University
NIEBUHR, REINHOLD (1967)

Christian Wildberg
Professor of Classics, Princeton University
ANAXIMANDER [BIBLIOGRAPHY] (2005)
PHILOPONUS, JOHN (2005)

Bernard Williams
Professor of Philosophy, University of London
HAMPSHIRE, STUART NEWTON (1967)
RATIONALISM (1967)

George Hunston Williams
Professor of Divinity, Divinity School, Harvard University; Director of the Foundation for Reformation Research; Past President of the American Society of Church History
SERVETUS, MICHAEL (1967)
SOCINIANISM (1967)

Raymond Williams
Lecturer in English, Cambridge University, and Director of English Studies, Jesus College, Cambridge University
ARNOLD, MATTHEW (1967)
ELIOT, THOMAS STEARNS (1967)
RUSKIN, JOHN (1967)

Thomas Williams
Associate Professor of Philosophy, The University of Iowa
GAUNILO (2005)

Timothy Williamson
The University of Edinburgh, Scotland
REFERENCE (1996)
SENSE (1996)

Arthur M. Wilson
Daniel Webster Professor, Professor of Government and of Biology, Dartmouth College; Fellow of the Royal Historical Society (London)
ENCYCLOPÉDIE (1967)

Deirdre Wilson
University College London
PRAGMATICS (1996)

Jessica Wilson
Assistant Professor, Philosophy, University of Toronto
FORCE [ADDENDUM] (2005)

R. McL. Wilson
Senior Lecturer in New Testament Language and Literature, St. Mary's College, St. Andrews University
MANI AND MANICHAEISM (1967)
MARCION (1967)
NUMENIUS OF APAMEA (1967)
SIMON MAGUS (1967)

Peter Winch
Reader in Philosophy, University of London
DURKHEIM, ÉMILE (1967)
LÉVY-BRUHL, LUCIEN (1967)
PARETO, VILFREDO (1967)
SPANN, OTHMAR (1967)
WEBER, MAX (1967)

John Wippel
Theodore Basselin Professor of Philosophy, Catholic University of America
GILES OF ROME (2005)

Gene Witmer
Associate Professor, Philosophy, University of Florida
PHILOSOPHY OF MIND (2005)

Charlotte Witt
Professor of Philosophy and Humanities, University of New Hampshire
FEMINIST METAPHYSICS (2005)

Gerd Wolandt
Docent in Philosophy, University of Bonn
HÖNIGSWALD, RICHARD (1967)

Erik Wolf
Professor of Law, Director of Philosophy of Law Seminar, University of Freiburg im Briesgau
ALTHUSIUS, JOHANNES (1967)
PUFENDORF, SAMUEL VON (1967)

Marvin E. Wolfgang
Professor and Graduate Chair, Sociology, University of Pennsylvania; Codirector of the Center of Criminal Research, University of Pennsylvania; President of the Pennsylvania Prison Society
BECCARIA, CESARE BONESANA (1967)

Harry A. Wolfson
Nathan Littauer Professor Emeritus of Hebrew Literature and Philosophy, Harvard University
PHILO JUDAEUS (1967)

Allan B. Wolter, O.F.M.
Ordinary Professor of Philosophy, Catholic University of America; Editor of Quincy College Publications
BACON, ROGER (1967)
BONAVENTURE, ST. (1967)
DUNS SCOTUS, JOHN (1967)

Nicholas Wolterstorff
Professor of Philosophy, Calvin College
CALVIN, JOHN (1967)

David Wong
Professor of Philosophy, Duke University
CHINESE PHILOSOPHY: ETHICS (2005)
ETHICAL RELATIVISM [ADDENDUM] (2005)

Rega Wood
Research Professor, Philosophy, Stanford University
RUFUS, RICHARD (2005)

George Woodcock
Assistant Professor of English, University of Washington; Associate Professor of English, University of British Columbia
ANARCHISM (1967)
BAKUNIN, MIKHAIL ALEKSANDROVICH (1967)
KROPOTKIN, PËTR ALEKSEEVICH (1967)
PROUDHON, PIERRE-JOSEPH (1967)
STIRNER, MAX (1967)

Arthur E. Woodruff
Assistant Professor, Mathematics and Science Education, Belfer Graduate School of Science, Yeshiva University
FARADAY, MICHAEL (1967)

Paul Woodruff
Darrell K. Royal Professor in Ethics and American Society, Department of Philosophy, The University of Texas at Austin
ARETĒ/AGATHON/KAKON (2005)
HIPPIAS OF ELIS [BIBLIOGRAPHY] (2005)
PRODICUS OF CEOS [BIBLIOGRAPHY] (2005)
PROTAGORAS OF ABDERA [BIBLIOGRAPHY] (2005)
VLASTOS, GREGORY (2005)

A. D. Woozley
Professor of Moral Philosophy, St. Andrews University
UNIVERSALS, A HISTORICAL SURVEY (1967)

John Worrall
Professor of Philosophy of Science; Co-Director, Centre for Philosophy of Natural and Social Sciences, London School of Economics and Political Science, University of London
LAKATOS, IMRE (2005)

Don J. Wyatt
Professor, History, Middlebury College
SHAO YONG (2005)

Michael Wyschogrod
Assistant Professor of Philosophy, City University of New York, City College
BUBER, MARTIN (1967)
HALEVI, YEHUDA (1967)

Liu Xiaogan
Professor, Philosophy, The Chinese University of Hong Kong
LAOZI (2005)

Keith E. Yandell
Julius R. Weinberg Professor of Philosophy, University of Wisconsin, Madison
RELIGIOUS LANGUAGE [ADDENDUM] (2005)

Frances A. Yates
Reader in the History of Renaissance, University of London; Fellow of the Royal Society of Literature
BRUNO, GIORDANO (1967)
HERMETICISM (1967)

Lee Yearley
Walter Y. Evans-Wentz Professor, Religious Studies, Stanford University
CHINESE PHILOSOPHY: RELIGION (2005)

Palle Yourgrau
Associate Professor of Philosophy, Brandeis University
GÖDEL, KURT (1996, 2005)

Mikko Yrjönsuuri
Academy Researcher, Philosophy, University of Jyväskylä
GERSON, JEAN DE (2005)
JOHN OF MIRECOURT (2005)
OLIVI, PETER JOHN (2005)

Naomi Zack
Professor of Philosophy, University of Oregon, Eugene
MULTICULTURALISM (2005)

Taras Zakydalsky
Editor, Russian Studies in Philosophy
SKOVORODA, HRYHORII SAVYCH (GRIGORII SAVVICH) [ADDENDUM] (2005)

Nino Zanghi
Associate Professor, Theoretical Physics, University of Genova, Italy
BELL, JOHN, AND BELL'S THEOREM (2005)
BOHM, DAVID (2005)
BOHMIAN MECHANICS (2005)

Mary-Barbara Zeldin
Associate Professor of Philosophy, Hollins College
SPIR, AFRIKAN ALEXANDROVICH (1967)

Nai Z. Zia
Editor of Christian Classics Series; Former Professor at Lingnan University and the University of Nanking
HAN FEI (1967)

Hossein Ziai
Professor of Iranian and Islamic Studies, Director of Iranian Studies, Department of Near Eastern Languages and Cultures, University of California, Los Angeles
ILLUMINATIONISM (2005)
SUHRAWARDĪ, SHIHĀB AL-DĪN YAḤYĀ [ADDENDUM] (2005)

Dean Zimmerman
Associate Professor, Philosophy, Rutgers University
DUALISM IN THE PHILOSOPHY OF MIND (2005)

Günter Zöller
Professor of Philosophy, University of Munich
FICHTE, JOHANN GOTTLIEB (2005)
KANT, IMMANUEL [ADDENDUM] (1996, 2005)
SCHOPENHAUER, ARTHUR [BIBLIOGRAPHY] (2005)

Arthur Zucker
Associate Professor and Chair; Department of Philosophy; Director, Institute for Applied and Professional Ethics, Ohio University
MEDICAL ETHICS (2005)
PHILOSOPHY OF MEDICINE (1996, 2005)

Jack Zupko
Associate Professor of Philosophy, Emory University
BURIDAN, JOHN (2005)

Arnulf Zweig
Assistant Professor of Philosophy, University of Oregon
BECK, JAKOB SIGISMUND (1967)
BENEKE, FRIEDRICH EDUARD (1967)
DÜHRING, EUGEN KARL (1967)
FECHNER, GUSTAV THEODOR (1967)

GOETHE, JOHANN WOLFGANG VON (1967)
KRAUSE, KARL CHRISTIAN FRIEDRICH (1967)
KÜLPE, OSWALD (1967)
LANGE, FRIEDRICH ALBERT (1967)
SIGWART, CHRISTOPH (1967)
STUMPF, KARL (1967)

list of articles

A

B

C

D

E

F

I

J

K

L

M

Q

R

S

T

U

V

ABBAGNANO, NICOLA

(1901–1990)

Nicola Abbagnano, born in Salerno, was the chief exponent of Italian existentialism, which he defined as a militant and rational "philosophy of the possible." Originally a pupil of Antonio Aliotta at the University of Naples, Abbagnano began teaching at the University of Turin in 1936, where he also for years had been coediting the influential *Rivista di filosofia*. Practically since his first book, *Le sorgenti irrazionali del pensiero* (Naples, 1923), Abbagnano had been advocating a change of philosophical horizon suitable to the problematic nature of human life. This advocacy is reflected in a notable series of historical studies, culminating in the monumental three-volume work *Storia della filosofia* (Turin, 1946–1950; 2nd ed., 1963).

Reacting against the prevailing neo-Hegelianism of Benedetto Croce and Giovanni Gentile in Italy, Abbagnano was influenced, in turn, by Edmund Husserl's phenomenology and, later, by Søren Kierkegaard, Martin Heidegger, and Karl Jaspers; but he revealed in his first attempt at existentialism, *La struttura dell'esistenza* (Turin, 1939), that he was no mere expositor or disciple of German existentialism. In that work he took a stand against Heidegger and Jaspers; and in subsequent writings his polemic was sharpened and extended to French existentialism, including Jean-Paul Sartre on the one hand and Gabriel Marcel, Louis Lavelle, René Le Senne on the other. He groups Sartre with Kierkegaard under German existentialism, and the others under "theological or ontological existentialism."

According to Abbagnano, all forms of existentialism in vogue since Kierkegaard have been self-defeating, since they lead, on examination, to the negation of what is basic to their whole interpretation of human existence: "the primacy of possibility." He discerns two principal directions within the contemporary existentialist movement. One (the left wing) is associated with the early Heidegger, Jaspers, and Sartre; the other (the right wing), with Marcel, Lavelle, and Le Senne. The first group of existentialists negates existence as possibility by reducing human possibilities to *impossibilities*, with everything projected by finite man inevitably foredoomed to fail; the second group negates existence by "surreptitiously" transforming human possibilities into *potentialities*, necessarily destined to succeed in the end.

Even though for Abbagnano the left and the right wings of the existentialist movement are founded, technically, on opposite principles—"the impossibility of the

possible" and "the necessity of the possible," respectively—they at least share a common negative ground because each of them, in one way or another, ultimately makes possibility itself impossible. The only valid alternative to "negative existentialism," which for polemical reasons Abbagnano calls "positive existentialism," takes as its guiding principle "the possibility of the possible" or, in Kantian terminology, "transcendental possibility." In this view, an authentic possibility in human life is one that, once it has been chosen or realized, remains open to further choice or realization; that is, continues to be possible. In short, Abbagnano's alternative constitutes an *open possibilism*.

This alternative calls for a clarification and coherent use of the fundamental category of all existentialism: the modal category of possibility. It is perhaps here that Abbagnano made his greatest contribution to the entire existentialist movement, especially since in contemporary logic, as he himself observes, the concept of modality has not been given sufficient "analytic elaboration."

Ever since Aristotle, Abbagnano maintains, there has been confusion concerning the modal categories, particularly with respect to the meaning of the term *possible*. The possible in the empirical sense of what may be has been distinguished from the possible in the purely logical sense of the noncontradictory. But, unfortunately, it has been confused with the "potential" in Aristotle's sense and with the "contingent" in Avicenna's. Since potentiality signifies "*pre*-determination" of the actual, the potential excludes the possible, ex hypothesi. Aristotle did concede that not all potentialities are actualized, but this concession on his part was only introduced "surreptitiously." For, if the potential means what is destined to occur anyway, there is no room for possibility as such. As for Avicenna's concept of the contingent, there is no doubt about its necessitarian character. For he makes the contingent into a species of the necessary—the contingent being, by his own definition, whatever is necessary through another. Hence, it follows that the modal status of the potential and the contingent is not that of possibility, of what may be; but that of necessity, of what must be. Abbagnano concludes that those who think in such terms, including existentialists, are necessitarians in disguise.

Historically, Abbagnano sees his own version of existentialism as an attempt to relate Immanuel Kant and Kierkegaard in a complementary way. In Kant's Table of Categories three pairs of categories are listed under modality: possibility-impossibility, existence-nonexistence, and necessity-contingency. Abbagnano virtually reduces Kant's three pairs of modality categories to one primary pair: the necessary and the nonnecessary. The reason he gives for doing so is that necessity and contingency are not really opposites. Neither are possibility and impossibility. For impossibility is the negative of necessity, not the negative of possibility; what *can't* be at all being the opposite of what *must* be of necessity.

As an existential possibilist, Abbagnano defines existence as possibility, and nonexistence as "non-possibility," not as impossibility. While the nonnecessary excludes the necessary and the impossible, it includes the possible and the nonpossible. This means that man can neither be sure of realizing his conflicting possibilities, nor be sure of the impossibility of their realization. It also means that every concrete possibility open to man has two aspects, a promising (positive) prospect and an inauspicious (negative) aspect. To illustrate, the possibility of knowledge implies the possibility of error. Errors are not "impossible," since we do in fact make them, but they are "non-possible" in the sense that they are unverifiable when put to test. Thus, a double-aspect theory of possibility lies at the heart of Abbagnano's "positive existentialism."

Another distinctive feature of Italian existentialism in general and of Abbagnano's philosophy in particular is the deliberate focus on a problem that was originally foreign to German existentialism; to wit, the problem of value.

Starting with the assumption that the problem of value is the problem of what man *ought to be*, Abbagnano argues in effect that, since the *ought-to-be* is the possible in the normative sense, it is therefore the moral equivalent of the *may-be*, which is the possible in the empirical sense. As a consequence, the logic of possibility coincides with the ethics of possibility, and these two phases of the same problem come together in Abbagnano's possibilistic interpretation of human conduct. This interpretation stresses the "normativity" of human existence, which involves the problem of freedom in all its dimensions. Thus, Abbagnano's existentialism logically unites the complementary categories of possibility and freedom, as is clear from his important volume *Possibilità e libertà* (Turin, 1956).

In the mid-twentieth century, Abbagnano came to characterize the "New Enlightenment," of contemporary philosophy and openly declared his affinities with the neopositivistic and neonaturalistic movements in the Anglo American world. As a result, he developed the empirical and naturalistic strains in his existentialism, emphasizing the methodological connections between possibility as a generic criterion of existence and verifia-

bility as a specific criterion in scientific inquiry. This "transfiguration" of existentialism into scientific methodology is clearly evident in the article on existentialism in *Dizionario di filosofia* (Turin, 1961). However, Abbagnano thought that the romantic "myth of security" in Auguste Comte's positivism, typical of the nineteenth-century mentality, still survives in the scientific utopianism of the Vienna Circle; and although he sympathizes with the later Ludwig Wittgenstein's thesis that the meaning of words depends on their use, he contends that the leader of the analytic movement failed to give a philosophical analysis of the notion of "use" itself. Abbagnano's sympathies with North American naturalism are reflected in his writings on John Dewey and in his review of P. Romanell's volume *Toward a Critical Naturalism* (*Rivista di filosofia* 50 [1959]: 108–109).

See also Aristotle; Avicenna; Comte, Auguste; Croce, Benedetto; Dewey, John; Existentialism; Gentile, Giovanni; Heidegger, Martin; Jaspers, Karl; Kant, Immanuel; Kierkegaard, Søren Aabye; Lavelle, Louis; Le Senne, René; Logical Positivism; Marcel, Gabriel; Naturalism; Possibility; Sartre, Jean-Paul; Scientific Method; Value and Valuation; Wittgenstein, Ludwig Josef Johann.

Bibliography

ADDITIONAL WORKS BY ABBAGNANO

Historical

Il nuovo idealismo inglese e americano. Naples, 1927.

La filosofia di E. Meyerson e la logica dell'identità. Naples: Perrella, 1929.

Guglielmo di Ockham. Lanciano: Carabba, 1931.

La nozione del tempo secondo Aristotele. Lanciano, 1933.

Bernardino Telesio. Milan: Fratelli Bocca, 1941.

Systematic

Introduzione all'esistenzialismo. Milan, 1942.

Filosofia, religione, scienza. Turin: Taylor, 1947.

Esistenzialismo positivo. Turin: Taylor, 1948.

"Contemporary Science and Freedom." *Review of Metaphysics* 5 (3) (1952): 361–378.

Problemi di sociologia. Turin: Taylor, 1959.

"Existentialism in Italy." *Cesare Barbieri Courier* 3 (2) (1961): 12–18.

Critical Existentialism Translated with an introduction by Nino Langiulli. New York: Doubleday, 1969.

Fra il tutto e il nulla. Milan: Rizzoli, 1973.

La saggezza della vita. Milan: Rusconi, 1985.

Scritti esistenzialisti. Turin: UTET, 1988.

Scritti neoillumnistici (1948–1965). Turin: UTET, 2001.

The Human Project—The Year 2000. Translated by Bruno Martini e Nino Langiulli. Vol. 119 of the "Value Inquiry Book Series." Amsterdam/New York: Rodopi, 2002.

WORKS ON ABBAGNANO

Bruno Maiorca. *Bibliografia degli scritti di Nicola Abbagnano 1922–1992*. Bari: Laterza, 1993.

Giannini, G. *L'esistenzialismo positivo di N. Abbagnano*. Brescia, 1956.

Langiulli, Nino. *Possibility, Necessity and Existence: Abbagnano and His Predecessors*. Philadelphia: Temple University Press, 1992.

Santucci, Antonio. *Esistenzialismo e filosofia italiana*. Bologna: Mulino, 1959.

Simona, Maria Angela. *La notion de liberté dans l'existentialisme positif de Nicola Abbagnano*. Fribourg, Switzerland: Editions Universitaires, 1962.

Patrick Romanell (1967)

Bibliography updated by Thomas Nenon (2005)

ABELARD, PETER

(1079–1142)

Peter Abelard has been famous since the fourteenth century for his exchange of love letters with Héloïse, his former wife, written when he was a monk and she a nun. Nineteenth-century historians saw him as a rationalist critic of traditional Christian doctrine and a forerunner of modernity. More recently, Abelard's originality and power as a philosopher have come to be appreciated.

Abelard's working life splits into two main, slightly overlapping periods. From about 1100 until about 1125, his activity as a thinker and teacher revolved around the ancient logical texts available in Latin at that time—the so-called *logica vetus* ("Old Logic"). But from about 1120, Abelard started to become strongly interested in questions about Christian doctrine, to which he gradually came to give an increasingly ethical emphasis. The important works of the first phase of his career were thus the *Dialectica* (c. 1113–1116), a logical textbook, and the *Logica Ingredientibus* (c. 1119), commentaries on ancient logical texts (along with a shorter logical commentary, the *Logica Nostrorum Petitioni Sociorum*, from the mid-1120s). To the second phase belong his *Theologia*, mainly a philosophical investigation of the Trinity, which exists in three different, much altered versions: *Theologia Summi Boni* (1121), *Theologia Christiana* (c. 1125), *Theologia Scholarium* (c. 1133–1134); biblical commentaries, and a set of *Sentences* (c. 1134), which record his lectures on a wide range of theological topics; the *Collationes* (Comparisons), an imaginary dialogue between a Philosopher, a Jew, and a Christian (probably c. 1130); and the *Scito teipsum* (Know yourself!) or, as it is sometimes called, Abelard's *Ethics* (1138).

Although the division of his career into two phases was partly occasioned by his castration in 1117 (at the hands of ruffians hired by Héloïse's uncle, the canon of Notre-Dame), which put a violent end to his marriage, and his subsequent decision to become a monk of Saint-Denis, Abelard remained a teacher for most of his life. After studying with two of the most celebrated logicians of the time, Roscelin of Compiègne and William of Champeaux, both of whom later considered him an enemy, Abelard set up his own school and finally became the schoolmaster in Paris. He continued to teach as a monk of Saint-Denis and later, when he left that monastery to set up his own hermetic-monastic community. After a period as an unsuccessful reforming abbot of a remote Breton monastery, Abelard returned to the now numerous and flourishing Paris schools in the 1130s. He spent his final years at Cluny and its dependency, after his activity as a teacher was ended by his condemnation at the Council of Sens (1140).

LOGIC

The *logica vetus* included just two texts by Aristotle himself, the *Categories and On Interpretation*, along with the *Isagoge* (Introduction) to the *Categories* by Porphyry (c. 232–305 CE), and texts by Boethius (c. 475–c. 524 CE) on categorical and hypothetical syllogism, division, and topical inference. From this unpromising set of authorities, Abelard was able not merely to explore areas of formal logic untouched by Aristotle, but also to elaborate a whole metaphysics and semantics.

Ancient and medieval logicians worked in natural language, rather than devising a special logical symbolism. One of the hallmarks of Abelard's approach to logic was his awareness of the ambiguities in many ordinary sentences and the need to distinguish them carefully when constructing a logical argument. Abelard was not the first medieval logician to notice this point (Anselm of Canterbury, for instance, was an eleventh-century forerunner), but he placed an emphasis on it that would be taken up by many of his medieval successors. Consider, for instance, a sentence such as "Possibly the standing man sits." Abelard is quick to observe that it can be read in a composite sense (*This is possible: that the man is standing-and-sitting*) or in a divided sense (*The man is standing, and it is possible that he is sitting*). Although this distinction is made by Aristotle in his *Sophistical Refutations*, Abelard had already used it very widely in his *Dialectica* before he read it in the Aristotelian text.

Moreover, Abelard used this approach as the basis for devising—as Christopher Martin has shown—a genuinely propositional logic, to complement the term logic of Aristotelian syllogistic. In antiquity, the Stoics developed a propositional logic, and traces of their theory are found in Boethius's writings on topical argument and hypothetical syllogisms. Boethius, however, clearly neither developed a propositional logic nor understood it. His hypothetical syllogisms (for instance, "If it is day, it is light. It is day. So it is light") look like arguments in propositional logic, but Boethius takes them as being based on the relation between the terms *day* and *light*; and he cannot grasp the negation of a conditional such as, "If it is day, it is light," except as the negation of one of the terms ("If it is day, it is not light"). By contrast, Abelard has a clear notion of propositional negation (It is not the case that: If it is day, it is light), and it governs his reconstruction of the theory of topical argument. For Boethius the theory of topics is a sort of logic for constructing real arguments on the basis of commonly accepted maxims, which range from basic logical principles to (fairly dubious) rules of thumb, such as "What the experts think about something is true." Abelard retains only those maxims which underwrite conditionals that are not just logically necessary, but where the sense of the consequent is contained in that of the antecedent (for example, Abelard accepts "Whatever is predicated of the species is predicated of the genus," on which is based, for instance, "If it is a man, it is an animal"). The resulting system of propositional logic turns out to be more like some modern connexive logics than classical modern propositional calculus.

METAPHYSICS AND SEMANTICS

Aristotle's *Categories* provided Abelard and his contemporaries with a basic metaphysics. It proposes that the items that make up the world are either substances, which exist independently, or non-substances, which exist only in dependence on substances; and that they are either particular or universal. For example, John Marenbon is a particular substance and man (in general) a universal one; the whiteness of John's skin and his rationality are individual non-substances, and whiteness and rationality (in general) are universal non-substances. Abelard, however, is a nominalist. Following, but exploring in more depth, a lead given by others, including Roscelin, he contended that everything which exists is a particular. There are no universal things, he argued, because to be universal a thing would have to be both one and shared between many in a way that is impossible. Abelard had, then, to show how the basic structure of the universe can be

explained solely in terms of particular substance and non-substances.

Unlike many more recent nominalists, Abelard accepted that the best scientific description (Aristotle's, he thought) cuts nature at the joints: It is a fundamental truth, he believed, that some things are human beings and others dogs, and that human beings are human because they are mortal, rational animals. To be a mortal, rational animal, indeed, is to have the "status" of man, Abelard said. But, he quickly added, a status is not a thing. Every human, then, is alike in having his or her own particular rationality, mortality, and animality. But what about these particular non-substance things? They are, in Abelard's view, real items on an ontological checklist because, he says, it might have been the case that the particularity rationality R1 by which John is rational was the rationality by which William—who is in fact rational by rationality R^2—is rational, and vice versa; and so R^1 cannot be explained away as just being John insofar as he is rational. The non-substance particulars are dependent, however, because they cannot exist except in some substance or other, and they cannot exist in one substance and then afterward in another. Just as Abelard has to explain what it is that makes John and William both human beings, he must explain too what it is that makes R^1 and R^2 both rationalities. But he does not, as might be expected, try to speak of a status of being rational—analyzing rationality into certain patterns of behavior, for instance. Rather, he seems to admit, in all but name, that there is a universal rationality.

Abelard's nominalism also poses a semantic problem with regard to universal words. It is important to grasp that this problem is *not* one about reference. Once a kind-word is first imposed, it automatically refers to every particular which is really of that kind, even if the impositor himself has merely a vague or inaccurate idea of the internal structure which characterizes the species in question. (This feature, as Peter King [1982] has pointed out, brings Abelard's semantics uncannily close to the thought of contemporary philosophers such as Kripke.) By contrast, a word's signification is, for medieval authors in general, a causal, psychological notion: a word w signifies x by causing a thought of x in the listener's mind. The signification of "human being" in "John is a human being" is clearly universal: the x of which it causes a thought is a universal human being, not a particular one. But how can there be such an x, if every thing is particular? Abelard's answer is to say that universal words cause a mental image, a confused conception of, for instance, what humans have in common, which is not the image of any particular man. Such confused conceptions are not things, and it is these conceptions which universal words signify. The conceptions are not things, because they are not thoughts themselves (which Abelard would class as particular non-substance things), but the contents of thoughts—objects in the world envisaged, to use an anachronistic expression, under a certain mode of presentation.

Abelard also had a theory about the semantics of sentences. A sentence signifies neither the things to which its component words refer, nor the thought they produce, but rather its *dictum* (meaning "what it says"). At first sight, Abelard seems to mean by *dictum* what modern philosophers call a proposition, and he does indeed characterize those logical connections that he understands propositionally—as, for example, between the antecedent and consequent of a conditional—as holding between *dicta*. But it is not quite clear whether *dicta* are truth-bearers or rather, like facts, truth-makers. Moreover, Abelard insists that *dicta*—along with statuses and common conceptions—are not things. But whether he can coherently deny the reality of *dicta*, while at the same time using them to underpin his account of the workings of the universe, remains doubtful. Nonetheless, Abelard's metaphysics is bold and original, and it ranges into many areas other than those discussed here, such as parts and wholes, relations, the physical constitution of objects and their sensible properties, and the laws of nature.

ETHICS

Like any Christian thinker, Abelard held that every detail of world history is providentially ordained. Unlike the great theologians of the thirteenth century, such as Thomas Aquinas and John Duns Scotus, he did not accept that God has any freedom in choosing what the course of providence should be: God, he argues, must choose whatever is best to happen, and that, he believes, leaves no space for alternatives. Yet there is room, Abelard thought (contradicting the Platonizing tradition of Augustine and Anselm) for the existence of genuinely evil things, because—as he explains, citing the distinction between things and *dicta*—it is good that there is evil.

If God ordains the universe so that every human action, good or evil, contributes to the best providence, it is clear that ethical judgment cannot be based on consequences. Abelard is very often seen as a moral theorist who, rather, concentrates entirely on intentions, and subscribes to a subjective view of morality. Both aspects of this characterization need qualification. Following Augustine's lead, almost all medieval thinkers based

moral judgment on intentions. For instance, Abelard's immediate predecessors and contemporaries saw sinning as a stage-by-stage process of intending—a person begins to sin once he entertains a temptation to perform a forbidden act; as he thinks about it with pleasure and plans how to put it into effect, the sin becomes graver, and it is more serious still when he actually performs the act. By contrast, for Abelard someone is guilty of sinning when, and only when, he consents to the sin—when he is ready to perform it and will do so unless thwarted. Up until that moment, he is not guilty, and, once that moment is reached, his guilt is complete: performing the act will not increase it.

Abelard's account of what determines whether an action is sinful or not seems at first sight to be subjective. A person sins, he says, by showing contempt for God. It sounds, from this definition, as if it is the mere subjective state of someone's mind, and not what he does or plans to do, that makes him a sinner. But, for Abelard, one shows contempt for God precisely by consenting to an action one knows is divinely forbidden. Sinners do not usually want to perform a forbidden action because it is forbidden; rather, they perform it in spite of the fact that God forbids it, and very often with the fervent wish that it were licit. Moreover, he does not think that it is a matter of guesswork to decide which acts God forbids. Christians and Jews have scriptural revelation to guide them; but, in any case, Abelard believed, all people in all places and in all times, apart from children and the mentally incapable, are able to grasp natural law, which teaches them the fundamental rules for behavior ordained by God. Abelard would not hesitate, therefore, to say that, for example, it is and was always wrong for a mentally normal adult to commit adultery (unless, in some way, he is unaware that it is in this case adultery) because he could not fail to know that adultery is divinely forbidden and that, therefore, it shows contempt to God to perform it.

Abelard's account of acting well is less fully developed than his treatment of sinning. He takes over a list of four virtues (ultimately from Plato's *Republic*) from Cicero: prudence, justice, courage, and temperance. He does not, however, use these virtues to provide a view of the good life for human beings. Rather, he sees justice as the central virtue, by which a person acts in accord with God's commands as known through revelation or natural law. Prudence is a precondition for being just, but not a virtue itself. Courage and temperance are props of justice. A person may be deflected from just action by fear or by desire for pleasure; courage makes him stand firm, despite what threatens him; temperance makes him resist the blandishments of pleasure.

As this description suggests, Abelard tends to think of morally good action as a hard-won victory over sinning, which is usually the easier or the more pleasant choice. Yet he also wants to insist that there is something deficient in goodness about actions which, although carried out from excellent motives, fail to achieve their intended good effect; as, for example, if a person works hard in order to provide for the poor or the sick, but his plans are never realized. Abelard's ethical theory is further complicated by a somewhat unexpected twist. He believes that judgments made by human judges should be based on a utilitarian evaluation of the punishments given. A woman who entirely unintentionally smothers her baby (whom she was trying to keep warm) should be punished severely, although she has committed no sin, so as to discourage others from making the same mistake.

PHILOSOPHY OF RELIGION

Modern interpreters of Abelard tend to play down any tension between his rationalism and Christian belief: He used the tools of his logic, they say, to analyse Christian doctrines and criticize heretical distortions of them, but he was fully willing to accept the ultimate mysteriousness of doctrines such as the Trinity. Yet there is good reason to see Abelard's main project in the works of his last decade as being the presentation of a rationalized Christianity, which in important ways did not accord with the accepted beliefs of his time.

Abelard's conception of a universal natural law was not merely a foundation for his ethical theory. People at all times and in all places, he believed, have been able to grasp the fact the God exists, and that God is triune. Supposedly pagan sources, such as Plato, the Sibylls, and the writings attributed to Hermes Trismegistus, provide better testimony, he believes, to the Trinity than anything in the Old or even the New Testament. Although Abelard—under pressure to conform to an orthodoxy which, as it turned out, he was in any case accused of infringing—might accept a certain element of inexplicable mystery in the doctrine of divine triunity, he elaborated in the different versions of his *Theologia* a complex theory of sameness and difference, which seems to have been designed to explain in terms of logic how something can be three and yet one. And he considered that God's triune nature emerged just from thinking about the attributes an omniperfect being must have: "For God to be three persons—Father, Son and Holy Spirit—is," he explains at the beginning of the *Theologia Summi Boni*, "as if we were

to say that the divine substance is powerful, wise and benign. …" This attitude was part of Abelard's general, though nuanced, rejection of there being anything praiseworthy in the acceptance by faith of truths that are not understood, and of the limited function he gives to revelation. For most of his contemporaries, the Jews, to whom the Old Law had been revealed, were far closer to a grasp of the truth than the ancient pagans. For Abelard, the pagan philosophers, without revelation but using natural law, were able to live highly virtuous lives and to reach a better understanding of God than most of the Jews.

Abelard did not, however, think that every important theological truth could be grasped by reason, without revelation. In particular, only by revelation can people know of Christ's life and his death, and without this knowledge, he thought, no one can be saved. But Abelard went on to argue that God would reveal what was necessary for salvation to anyone who lived well, and also to give a rationalistic explanation of why it was necessary to know about Christ's crucifixion—because it set an example of love, indispensable for being able to overcome temptations. Similarly, while Abelard broadly accepted the biblical accounts of heaven and hell, he was one of the few medieval thinkers to insist that they should not be interpreted literally.

AFTER ABELARD

One of the schools of later twelfth-century philosophy, the *nominales*, probably consisted of Abelard's followers. But, apart from his letters to Héloïse, Abelard was not one of the authors who was much read after 1200. Elements of his approach to logic were absorbed into the developing medieval curriculum, although many of his subtlest ideas seem never to have been used. The type of doctrinal problems raised by him influenced the *Sentences*, written by Peter Lombard in the 1150s, and through this work, which became the standard textbook, the whole tradition of later medieval theology. Abelard's effect on the positions and arguments they developed was very limited, however, because the university theologians had their outlook formed by a reading of the whole range of Aristotle's philosophy and the Arabic commentary tradition. In many ways, however, Abelard's approach to metaphysics and the philosophy of religion, with its basis in logical and linguistic analysis, is closer to today's philosophical tastes than the grand systems of the thirteenth and early fourteenth-century philosophers.

See also Aristotelianism; Logic, History of: Ancient Logic; Logic, History of: Medieval (European) Logic; William of Champeaux.

Bibliography

WORKS BY ABELARD

Selected Latin Texts

Petri Abaelardi Dialectica, edited by Lambertus M. de Rijk. 2nd ed. Assen: Van Gorcum, 1970.

Petri Abaelardi opera theologica I–III, edited by Eligius Buytaert (and Constant J. Mews, vol. III). Turnhout: Brepols, 1967. Corpus christianorum, continuatio mediaeualis 13 (including the *Theologia* in its different versions).

Peter Abaelards philosophische Schriften, I. 3, edited by Bernhard Geyer. Münster: Aschendorff, 1927 (Beiträge zur Geschichte der Philosophie und Theologie des Mittlalters 21). Logical commentaries.

Sententiae magistri Petri Abaelardi (Sententie Hermanni), edited by Sandro Buzzetti. Florence: La nuova Italia, 1983.

English Translations

Five Texts on the Mediaeval Problem of Universals: Porphyry, Boethius, Abelard, Duns Scotus, Ockham, edited by Paul V. Spade. Indianapolis and Cambridge: Hackett, 1994. Extract from *Logica Ingredientibus*.

Peter Abelard's "Collationes," edited by J. Marenbon and G. Orlandi. Oxford: Oxford University Press, 2001. Latin and English.

Peter Abelard's "Ethics," edited by David E. Luscombe. Oxford: Oxford University Press, 1971. Latin and English.

WORKS ABOUT ABELARD

Brower, Jeffrey E., and Kevin Guilfoy, eds. *The Cambridge Companion to Abelard*. Cambridge, U.K., and New York: Cambridge University Press, 2004.

Clanchy, Michael. *Abelard: A Medieval Life*. Oxford: Blackwell, 1997.

De Rijk, Lambertus M. "Peter Abelard's Semantics and his Doctrine of Being." *Vivarium* 24 (1986): 85–128.

King, Peter O. "Peter Abailard and the Problem of Universals." PhD diss. Princeton University, 1982.

Marenbon, John. *The Philosophy of Peter Abelard*. Cambridge, U.K.: Cambridge University Press, 1997.

Martin, Christopher J. "Embarrassing Arguments and Surprising Conclusions in the Development of Theories of the Conditional in the Twelfth Century." In *Gilbert de Poitiers et ses contemporains*, edited by Jean Jolivet and Alain De Libera, 377–400. Naples: Bibliopolis, 1987.

Martin, Christopher J. "Logic." In *The Cambridge Companion to Abelard*, edited by Jeffrey E. Brower and Kevin Guilfoy, 158–199. Cambridge, U.K., and New York: Cambridge University Press, 2004.

Mews, Constant J. *Abelard and Heloise*. Oxford: Oxford University Press, 2005.

Mews, Constant J. *Peter Abelard*. Aldershot: Variorum, 1995.

Tweedale, Martin. *Abailard on Universals*. Amsterdam and New York: North Holland, 1976.

John Marenbon (2005)

ABORTION

The claims to which partisans on both sides of the "abortion" issue appeal seem, if one is not thinking of the abortion issue, close to self-evident, or they appear to be easily defensible. The case against abortion (Beckwith 1993) rests on the proposition that there is a very strong presumption that ending another human life is seriously wrong. Almost everyone who is not thinking about the abortion issue would agree. There are good arguments for the view that fetuses are both living and human. ("Fetus" is generally used in the philosophical literature on abortion to refer to a human organism from the time of conception to the time of birth.) Thus, it is easy for those opposed to abortion to think that only the morally depraved or the seriously confused could disagree with them.

Standard pro-choice views appeal either to the proposition that women have the right to make decisions concerning their own bodies or to the proposition that fetuses are not yet persons. Both of these propositions seem either to be platitudes or to be straightforwardly defensible. Thus, it is easy for pro-choicers to believe that only religious fanatics or dogmatic conservatives could disagree. This explains, at least in part, why the abortion issue has created so much controversy. The philosophical debate regarding abortion has been concerned largely with subjecting these apparently obvious claims to the analytical scrutiny philosophers ought to give to them.

Consider first the standard argument against abortion. One frequent objection to the claim that fetuses are both human and alive is that we do not know when life begins. The reply to this objection is that fetuses both grow and metabolize and whatever grows and metabolizes is alive. Some argue that the beginning of life should be defined in terms of the appearance of brain function, because death is now defined in terms of the absence of brain function (Brody 1975). This would permit abortion within at least eight weeks after conception. However, because death is, strictly speaking, defined in terms of the irreversible loss of brain function, the mere absence of brain function is not a sufficient condition for the absence of life. Accordingly, the claim that the presence of brain function is a necessary condition for the presence of life is left unsupported. Also, the standard antiabortion argument is criticized on the ground that we do not know when the soul enters the body. However, such a criticism is plainly irrelevant to the standard, apparently secular, antiabortion argument we are considering.

The Thomistic premise that it is always wrong intentionally to end an innocent human life is used by the Vatican to generate the prohibition of abortion. This premise is often attacked for presupposing "absolutism." This Vatican principle seems to render immoral active euthanasia, even when a patient is in excruciating, unrelievable pain or in persistent coma; it even seems to render immoral ending the life of a human cancer-cell culture. In none of these cases is the individual whose life is ended victimized. Thus, the Vatican principle seems most implausible.

Opponents of abortion are better off appealing to the weaker proposition that there is a very strong presumption against ending a human life (Beckwith 1993). Because this presumption can be overridden when the victim has no interest in continued life, use of this premise provides a way of dealing with the above counterexamples. However, this tactic provides room for another objection to the antiabortion argument. Some pro-choicers have argued that insentient fetuses have no interest in continued life. Because what is insentient does not care about what is done to it and because what does not care about what is done to it cannot have interests, insentient fetuses cannot have an interest in living. Therefore, abortion of insentient fetuses is not wrong (Steinbock 1992, Sumner 1981, and Warren 1987).

If this argument were sound, then it would also show that patients who are in temporary coma, and therefore insentient, do not have an interest in living. M. A. Warren (1987) attempts to avoid this counterexample by making the neurological capacity for sentience a necessary condition for having any interests at all and, therefore, for having an interest in living. This move does not solve the problem, however. Because the argument in favor of permitting the abortion of insentient fetuses generated an untenable conclusion, that argument must be rejected. Because the argument rests on an equivocation between what one takes an interest in and what is in one's interest, there are even better reasons for rejecting it. Accordingly, this objection to the standard antiabortion argument is unsupported.

The classic antiabortion argument is subject to a major theoretical difficulty. Antiabortionists have tried vigorously to avoid the charge that they are trying to force their religious views upon persons who do not share them. However, the moral rule to which the standard antiabortion argument appeals obtains its particular force in the abortion dispute because it singles out members of the species *Homo sapiens* (rather than persons or sentient beings or beings with a future like ours, for example). It is difficult to imagine how the *Homo sapiens*

rule could be defended against its competitors without relying upon the standard theological exegesis of the Sixth Commandment and upon the divine-command theory on which its moral standing rests. This leads to two problems. First, arguments against divine-command ethical theory seem compelling. Second, when arguments based on divine-command theory are transported into the Constitutional realm, First Amendment problems arise.

The philosophical literature contains two major kinds of pro-choice strategies. The personhood strategy appeals to the proposition that no fetuses are persons. If this is so, then, because a woman plainly has the right to control her own body if she does not directly harm another person, abortion is morally permissible. However, Judith Thomson (1971) has argued that a woman's right to control her own body can justify the right to an abortion in some situations even if fetuses are persons. This second strategy rests on the claim that no one's right to life entails the right to a life-support system provided by another's body even if use of that life-support system is the only way to save one's life. Thus, even if opponents of abortion are successful in establishing that fetuses have the right to life, they have not thereby established that any fetus has the right to anyone else's uterus.

It is widely believed that Thomson's strategy can justify abortion in cases of rape and in cases where the life of a pregnant woman is threatened by pregnancy (Warren 1973). There is much less unanimity concerning other cases, because it is generally believed that, if we create a predicament for others, we have special obligations to help them in their predicament. Furthermore, let us grant that A's right to life does not entail A's right to B's body even when A needs B's body to sustain life. Presumably, by parity of reasoning, B's right to B's body does not entail B's right to take A's life even if A's continuing to live severely restricts B's choices. Thus, we have a standoff, and the winner from the moral point of view will be that individual with the strongest right. Although Thomson's strategy has been widely discussed and raises interesting questions about the duty of beneficence, questions both about its philosophical underpinnings and about its scope suggest that philosophically inclined pro-choicers would be better off with a personhood strategy.

No doubt, this is why personhood strategies have dominated the pro-choice philosophical literature. Such strategies come in many varieties (Engelhardt 1986; Feinberg 1986; Tooley 1972, 1983, and 1994; and Warren 1973, 1987). Warren's 1973 version is most famous. She argued that reflection on our concept of person suggests that in order to be a person one must possess at least more than one of the following five characteristics: consciousness, rationality, self-motivated activity, the capacity to communicate, and the presence of a concept of self. Since no fetus possesses any of these characteristics, no fetus is a person. If only persons have full moral rights, then fetuses lack the full right to life. Therefore, abortion may never be forbidden for the sake of a fetus.

One might object to such a strategy on the ground that, since fetuses are potential persons, the moral importance of personhood guarantees them a full place in the moral community. The best reply to such an objection is that the claim that *X*'s have a right to *Y* does not entail that potential *X*'s have a right to *Y* (think of potential voters and potential presidents; Feinberg 1986).

Although personhood theorists (like antiabortionists) tend to say little about the moral theories on which their views rest (Engelhardt 1986 is an interesting exception), presumably most personhood theorists will turn out to be, when driven to the wall, social-contract theorists. Such theories, according to which morality is a self-interested agreement concerning rules of conduct among rational agents, tend to have problems accounting for the moral standing of those who are not rational agents—beings such as animals, young children, the retarded, the psychotic, and the senile. Thus, the personhood defense of the pro-choice position tends to have problems that are the inverse of those of the classic antiabortion argument.

Both standard antiabortion and personhood accounts appeal, in the final analysis, to the characteristics fetuses manifest at the time they are fetuses as a basis for their arguments concerning the ethics of abortion. This appeal may be a mistake both defenses share. My premature death would be a great misfortune to me because it would deprive me of a future of value. This is both generalizable and arguably the basis for the presumptive wrongness of ending human life. Such a view seems to imply that abortion is seriously immoral, seems to have a defensible intuitive basis, and seems to avoid the counterexamples that threaten alternative views (Marquis 1989). However, this view is subject to two major objections. One could argue that the difference between the relation of fetuses to their futures and the relation of adults to their futures would explain why adults are wronged by losing their futures but fetuses are not (McInerney 1990). One might also argue that because human sperm and ova have valuable futures like ours, the valuable future criterion for the wrongness of killing is

too broad (Norcross 1990). Not everyone believes these objections are conclusive.

See also Animal Rights and Welfare; Bioethics; Rights.

Bibliography

Beckwith, Francis J. *Politically Correct Death: Answering Arguments for Abortion Rights*. Grand Rapids, MI: Baker, 1993.

Boonin, David. *A Defense of Abortion*. Cambridge, U.K.: Cambridge University Press, 2003.

Brody, Baruch. *Abortion and the Sanctity of Human Life: A Philosophical View*. Cambridge, MA: MIT Press, 1975.

Burgess, J. A., and S. A. Tawia. "When Did You First Begin to Feel It? Locating the Beginning of Human Consciousness." *Bioethics* 10 (1) (1996): 1–26. An extremely useful discussion for sentience-based views.

Davis, Nancy. "Abortion and Self-Defense." In *Abortion: Moral and Legal Perspectives*, edited by Jay L. Garfield and Patricia Hennessey, pp. 186–210. Amherst: The University of Massachusetts Press, 1984.

Dworkin, Ronald. *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom*. New York: Knopf, 1993.

Engelhardt, H. Tristam, Jr. *The Foundations of Bioethics*. New York: Oxford University Press, 1986. An attempt to place the personhood strategy in the context of an ethical theory and other issues in bioethics.

English, Jane. "Abortion and the Concept of a Person" (1975). In *Arguing about Abortion*, edited by Lewis M. Schwartz, pp. 159–198. Belmont, CA: Wadsworth, 1993.

Feinberg, J. "Abortion." In *Matters of Life and Death: New Introductory Essays in Moral Philosophy*. 2nd ed., edited by Tom Regan. New York: Random House, 1986. The best account of the personhood strategy in a single essay.

Ford, Norman M. *When Did I begin? Conception of the Human Individual in History, Philosophy and Science*. Cambridge, U.K.: Cambridge University Press, 1988. An excellent discussion of issues relevant to the moral status of the early embryo.

Hare, R. M. "Abortion and the Golden Rule" (1975). In his *Essays on Bioethics*, pp. 147–167. Oxford: Clarendon, 1993.

Hare, R. M. "A Kantian Approach to Abortion" (revised version, 1989). In his *Essays on Bioethics*, pp. 168–184. Oxford: Clarendon, 1993.

Hursthouse, Rosalind. "Virtue Theory and Abortion." *Philosophy and Public Affairs* 20 (3) (Summer 1991): 223–246.

Kamm, F. M. *Creation and Abortion: A Study in Moral and Legal Philosophy*. New York: Oxford University Press, 1992. A defense of Thomson's strategy.

Lee, Patrick. *Abortion and Unborn Human Life.* Washington, DC: The Catholic University of America Press, 1996. One of the best book-length cases against abortion.

Marquis, Donald. "Why Abortion Is Immoral." *Journal of Philosophy* 86 (1989): 183–202.

McInerney, Peter. "Does a Fetus Already Have a Future-Like-Ours?" *Journal of Philosophy* 87 (1990): 265–268.

McDonagh, Eileen L. *Breaking the Abortion Deadlock: From Choice to Consent*. Oxford: Oxford University Press, 1996. Defends a self-defense variant of Thomson's strategy.

Norcross, A. "Killing, Abortion, and Contraception: A Reply to Marquis." *Journal of Philosophy* 87 (1990): 268–277.

Pojman, Louis J., and Francis J. Beckwith, eds. *The Abortion Controversy: 25 Years after* Roe v. Wade*: A Reader*. 2nd ed. Belmont, CA: Wadsworth, 1998. An excellent anthology.

Quinn, Warren. "Abortion: Identity and Loss." In his *Morality and Action*, pp. 20–51. Cambridge, U.K.: Cambridge University Press, 1993.

Schwarz, Stephen. *The Moral Question of Abortion*. Chicago: Loyola University Press, 1990. A useful book-length argument against abortion.

Steinbock, Bonnie. *Life before Birth: The Moral and Legal Status of Embryos and Fetuses*. New York: Oxford University Press, 1992. A sentience strategy.

Stone, Jim. "Why Potentiality Matters." *Canadian Journal of Philosophy* 17 (4) (December 1987): 815–830.

Sumner, L. W. *Abortion and Moral Theory*. Princeton, NJ: Princeton University Press, 1981. A sentience strategy.

Thomson, Judith Jarvis. "A Defense of Abortion." *Philosophy and Public Affairs* 1 (1971): 47–66. A classic paper.

Tooley, M. "Abortion and Infanticide." *Philosophy and Public Affairs* 2 (1972): 37–65. A classic paper.

Tooley, M. *Abortion and Infanticide*. New York: Oxford University Press, 1983.

Tooley, M. "In Defense of Abortion and Infanticide." In *The Abortion Controversy: A Reader*, edited by L. P. Poman and F. J. Beckwith. Boston: Jones and Bartlett, 1994.

Warren, Mary Anne. "The Abortion Issue." In *Health Care Ethics: An Introduction*, edited by D. VanDeVeer and T. Regan. Philadelphia: Temple University Press, 1987. A sentience strategy combined with a personhood strategy.

Warren, Mary Anne. "On the Moral and Legal Status of Abortion." *Monist* 57 (1973): 43–61. A classic paper.

Don Marquis (1996)

Bibliography updated by David Boonin (2005)

ABSOLUTE, THE

"The Absolute" is a term used by philosophers to signify the ultimate reality regarded as one and yet as the source of variety; as complete, or perfect, and yet as not divorced from the finite, imperfect world. The term was introduced into the philosophical vocabulary at the very end of the eighteenth century by Friedrich Wilhelm Joseph von Schelling and Georg Wilhelm Friedrich Hegel and was naturalized into English by Samuel Taylor Coleridge as early as 1809–1810 in *The Friend.* Later in the century it was an important term in the writings of such Idealist philosophers as James Frederick Ferrier, Francis Herbert Bradley, Bernard Bosanquet, and Josiah Royce.

INTRODUCTION OF THE TERM

One of the sources of the philosophy of the Absolute is the literature about Benedict (Baruch) de Spinoza commencing with Moses Mendelssohn's *Morgenstunden* (1785) and F. H. Jacobi's *Ueber die Lehre des Spinoza in Briefen an den Herrn Moses Mendelssohn* (1785). The expression "the Absolute" does not appear in these books, but there is a discussion of Spinoza's view that God does not transcend the world but is the sole infinite substance in which everything has its being. In the second edition of his book (1789), Jacobi printed as an appendix passages from Giordano Bruno's *De la causa, principio et uno* (1584) in order to call attention to a defense of pantheism that had, in Jacobi's view, influenced both Spinoza and Gottfried Wilhelm Leibniz.

Another source of the philosophy of the Absolute is Immanuel Kant's doctrine of the Reason as the faculty that aims at unified knowledge of the Unconditioned—"to find for the conditioned knowledge of the Understanding the Unconditioned that completes its unity" (*Critique of Pure Reason,* A307). In the Fourth Antinomy (A453) Kant writes of "an absolutely necessary being" (*ein Absolutnotwendiges*), and in the *Critique of Judgment,* in his account of the sublime, Kant distinguishes between what is great merely by comparison with something smaller (*comparative magnum*) and what is absolutely, not merely comparatively, great (*absolute magnum*). The former is a sensible concept, the latter is a concept of the Reason that "conducts the notion of nature to a supersensible substratum (underlying both nature and our faculty of thought) which is great beyond every standard of the senses" §26). Kant, of course, warned against supposing that these concepts of absolute unity and the absolutely unconditioned were more than Ideas that direct and regulate the search for empirical knowledge. But he himself, in the *Critique of Practical Reason* (1788), claimed to show that the reality of an unconditioned cause, and hence of freedom, could be *proved* "by means of an apodeictic law of the practical reason, and becomes the keystone of the whole edifice of a system of pure, even of speculative reason" (Preface). Thus Kant himself went some way toward repairing the destruction he had wrought upon "the edifice of speculative reason," and during his last years Johann Gottlieb Fichte and Schelling carried this work further in ways he by no means approved.

We have seen that Kant said that the Practical Reason provided proof of something Unconditioned, namely, of free, uncaused activity. Fichte, in his *Grundlage der gesammten Wissenschaftslehre* (1794), developed this aspect of Kant's teaching, arguing that a nonempirical, free, and active self must be regarded not merely as a condition of human knowledge, but also as the source and essence of all that is. (It is "All my I," as Coleridge derisively parodies it in the *Biographia Literaria.*) Thus the Transcendental Ego, which in Kant's philosophy was a logical or epistemological conception, was transformed by Fichte into the "absolute ego," a being that he later described as "the creator of all phenomena, including phenomenal individuals." Schelling's earliest writings were reinforcements of Fichte's views and shared his philosophical vocabulary.

By 1800, however, Schelling was moving toward a position of his own, and in his *System des transzendentalen Idealismus* of that year he writes of "an Absolute," and even, once or twice, of "the Absolute." In his *Darstellung meines Systems der Philosophie* (1801) he writes that "there is no philosophy except from the standpoint of the Absolute," and "Reason is the Absolute." In Hegel's *Differenz des Fichtischen und Schellingschen Systems der Philosophie* (1801) the Absolute is constantly referred to. Hegel writes, for example: "Division and conflict [*Entzweiung*] is the source of the need for philosophy, and in the form of the culture of the age, is its unfree, merely given aspect. What is merely an appearance of the Absolute has isolated itself from the Absolute and set itself up as independent." It will be noticed that in this passage the Absolute is contrasted with appearances and with what is "unfree," and that there is a further contrast between appearances that are falsely regarded as independent and appearances viewed in relation to the Absolute.

In 1803, there appeared the second edition of the essay by Schelling titled *Ideen zu einer Philosophie der Natur,* which had first appeared in 1797. In an appendix written for this new edition, Schelling argues that philosophy, as concerned with first principles, must be "an absolute science," that it is therefore concerned with what is absolute, and that, since all things (*Dinge*) are conditioned (*bedingt*), philosophy must be concerned with the activity of knowing rather than with things or objects. "Philosophy," he writes, "is the science of the Absolute," and the Absolute is the identity of the act of knowledge and of what is known. Schelling gives the name "Absolute Idealism" to the philosophy in which this identity is recognized. The exponent of Absolute Idealism, he argues, seeks out the intelligence that is necessarily embodied in nature, and he achieves by means of "intellectual intuition" a grasp of the identity between knower and known.

"The Absolute" was now well established in the vocabulary of Idealist philosophy.

SOME VIEWS ABOUT THE NATURE OF THE ABSOLUTE

We have seen that Schelling regarded the Absolute as that which intellectual intuition revealed as the identity of the knower and the known. He argued, furthermore, that knowledge is inseparable from will, so that the ultimate whole is active and free. The Absolute is manifested not only in nature but also in human history, which is a progress toward self-consciousness. An important thesis of Schelling's philosophy of the Absolute is that whereas in nature the Absolute is embodied in an unconscious way, in works of art it is consciously embodied, so that through his productions the artistic genius reveals the Absolute to humankind. In *Philosophie und Religion* (1804) Schelling tried to show how the finite, phenomenal world is related to the Absolute. He here had recourse to the notion of a fall that is a consequence of freedom and is yet, like the Absolute itself, outside time. He recognized that his view might be regarded as pantheistic (it was so regarded by Coleridge), and he attempted to show that human selves are, although finite, divine by nature. Thus the philosophy of the Absolute is developed as a sort of theology with some kinship to the speculations of Nicholas of Cusa.

It is well known that in his *Phenomenology of Mind* (1807) Hegel, by his characterization, "a night in which all cows are black," insinuated that Schelling's Absolute had no positive ascertainable features. Schelling, for his part, regarded Hegel's Absolute as "panlogistical"; that is, as nothing but an array of abstract categories. In his *Encyclopedia* Hegel presents various "definitions" of the Absolute in ascending order of complexity and adequacy. It is Being, he says, as Parmenides had held, but this is the least that can be said about it. It is also the self-identical, and, at a higher level, it is inference (*Schluss*—Wallace translates it "syllogism"). These definitions, from the *Logic,* appear to confirm Schelling's criticisms; but when Hegel comes to the *Philosophy of Mind,* the third part of the *Encyclopedia,* he writes that "the Absolute is mind: this is the highest definition of the Absolute." In his account of mind, Hegel shows how it develops as society moves toward higher levels of freedom in the course of human history, and how it reaches its fullest expression in the self-consciousness of the philosopher. Hegel's intention was to describe the Absolute in such a way that it would be seen to be infinite and yet comprise the finite within itself, and to be real and yet contain the apparent. But this intention was so ambitious that the result is ambiguity, and the Hegelian Absolute has been regarded by some, including Andrew Seth (later Pringle-Pattison), as "a single self" in which finite selves are lost, and by others, such as J. McT. E. McTaggart, as a society of individual, nontemporal selves. The ambiguity is also reflected in divergent interpretations of the religious significance of Hegel's Absolute, the majority of interpreters regarding it as equivalent to God, with others, for example, Bruno Bauer and Kojève, taking the view that "the Absolute" is Hegel's designation for man as a progressing historical individual.

In the nineteenth century and the early twentieth century, Absolutism became an important influence in the philosophy of Great Britain and the United States. J. S. Ferrier, who had written a life of Schelling and who had studied Coleridge and was aware of Schelling's influence on him, expounded, in his *Institutes of Metaphysics* (1854), a pluralistic Absolutism according to which there is a plurality of contingent "Absolute Existences" that are "minds-together-with-that-which-they-apprehend," and one "Absolute Existence which is strictly *necessary* … a supreme, infinite and everlasting Mind in synthesis with all things." But the most influential version of Absolute Idealism to be published in English was Bradley's *Appearance and Reality* (1893). In this book Bradley argued that mere appearances are conflicting and self-contradictory and that reality or the Absolute must therefore be harmonious and consistent. The self-contradictory character of appearances is due to their relatedness, and therefore the Absolute must not contain relations. Bradley maintained that the nature and possibility of a harmonious nonrelational whole is adumbrated in "immediate experience," the prereflective experience from which the world of distinct and related things emerges as we learn to talk and to judge. In this prereflective experience, subject and object are not yet differentiated, and there is diversity without numerical plurality. "From such an experience of unity below relations," Bradley writes, "we can rise to the idea of a superior unity above them." In this view, the Absolute is a suprarelational, differentiated harmony of experience. It is not a self, and it is not God, for "short of the Absolute, God cannot rest, and having reached that goal, he is lost and religion with him." Some have thought that this view of the Absolute is less open to the charge of panlogism than is that of Hegel. Before the publication of *Appearance and Reality,* Andrew Seth had, from within the Idealist school, criticized the line of thought that submerged individual selves in an impersonal or suprapersonal Absolute. McTaggart, we have seen, did not interpret Hegel in this way, and endeavored on his own

account to show that the unreality of the phenomenal world is consistent with the absolute existence of individual selves. Josiah Royce's solidly and persuasively argued *The World and the Individual* (1904) is another attempt to rescue individual minds from absorption in the Absolute.

CRITICAL COMMENTS

It is remarkable that a line of philosophical argument that set out to defend the reality of mind and of freedom should end up with minds that are self-contradictory appearances and an Absolute that alone is free. The Absolute was to have been the seat of freedom, reality, truth, and harmony; yet if Bradley was right, harmony and reality shut out the possibility of truth and freedom. Like Spinoza he tried to meet the difficulty with a doctrine of degrees of truth and freedom; and the comparison is revealing, for Spinoza is often regarded as a determinist. What went wrong? Coleridge, although greatly impressed by Schelling, argued in *The Friend* that Schelling's view, like that of Spinoza, was pantheistic. We may agree that Schelling sought for truth and freedom in the universe at large instead of in the limited beings to which they really belong. Schelling continued Kant's error of locating freedom outside the only world in which it is of importance, the world in which individual men decide and act. The view of Absolute Idealists is, however, that this world is merely phenomenal and must be contrasted with an infinite reality that contains it. The critic will ask whether this infinite reality must exist or whether it is only a projection from the finite. In adopting the former view, Absolutists have used arguments analogous to the Ontological Argument and to the Argument from the Contingency of the World. It would be self-contradictory, that is, to suppose that the Perfect could fail to exist; and in any case contingent being could not *be* unless there were a Necessary Being. Pierre Gassendi, Kant, and others have brought forward arguments against these so-called proofs, but it will not do merely to move forward these "disproofs" in opposition to Absolute Idealism. For the defenders of the Absolute do not allow that the distinctions made in these objections, between thought and reality or between concepts and things, are tenable just as they stand. Absolute Idealists cannot be refuted by arguments in which commonsense distinctions or the terms of an opposed philosophical tradition are uncritically presupposed. It is true that the conceptual adventurousness of Absolute Idealism was the occasion for the extreme conceptual conservatism of G. E. Moore and of those philosophers who insist on the essential rightness of ordinary language. But in the course of philosophical argument it has emerged that facts and concepts, the world and the ways in which it is thought about, cannot be isolated from one another as dogmatic common sense says they can be. On this matter the Absolutists' prejudice in favor of unity seems to have caused them to look in the right direction and to see how closely associated with one another are our conceptual framework and the world it is used to describe and classify.

See also Bosanquet, Bernard; Bradley, Francis Herbert; Bruno, Giordano; Coleridge, Samuel Taylor; Ferrier, James Frederick; Fichte, Johann Gottlieb; Gassendi, Pierre; Hegel, Georg Wilhelm Friedrich; Jacobi, Friedrich Heinrich; Kant, Immanuel; Leibniz, Gottfried Wilhelm; McTaggart, John McTaggart Ellis; Mendelssohn, Moses; Moore, George Edward; Ontological Argument for the Existence of God; Practical Reason; Pringle-Pattison, Andrew Seth; Reason; Royce, Josiah; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de.

Bibliography

Historians of philosophy do not seem to say much about the introduction of the term "the Absolute." Information can be obtained from Richard Kroner, *Von Kant bis Hegel,* 2nd ed. (Tübingen: Mohr, 1961) and from Frederick Copleston, S.J., *A History of Philosophy,* Vol. VII, *Fichte to Nietzsche* (London: Search Press, 1963).

On various views about the nature of the Absolute, see, in addition to the books mentioned in the text: Bruno Bauer, *Die Posaune des jüngsten Gerichts wider Hegel, den Atheisten und Antichristen* (Leipzig, 1841); Andrew Seth (later Pringle-Pattison), *Hegelianism and Personality* (London and Edinburgh: Blackwood, 1887); J. McT. E. McTaggart, *Studies in Hegelian Cosmology* (Cambridge: University Press, 1901); A. Kojève, *Introduction à la lecture de Hegel* (Paris: Editions Gallimard, 1947).

For criticisms of Absolutism, see: William James, *A Pluralistic Universe* (New York and London: Longmans, Green, 1909), Chs. II and III; G. E. Moore, *Some Main Problems of Philosophy* (London: Allen and Unwin, 1953), Chs. VIII–XII; A. C. Ewing, *Idealism, a Critical Survey* (London: Methuen, 1934). In Ewing, Ch. VIII, §3 is headed "The Absolute" and contains a brief discussion of the views of Bradley and Bosanquet.

H. B. Acton (1967)

ABUBACER

See *Ibn Ṭufayl*

ACADEMY

See *Florentine Academy; Greek Academy*

ACOSTA, GABRIEL

See *Costa, Uriel da*

ACTION

People speak not only of the actions of human beings and other intelligent animals but also of the actions of inanimate objects such as acids and waves. The philosophy of action, however, is not directly concerned with the actions of inanimate objects. Its primary subject matter is intentional action. Two questions are central in the philosophy of action: What are intentional actions? And how are intentional actions to be explained? An adequate answer to the first question would enable one to see how intentional actions differ from everything else—including the actions of acids and waves, nonactions, and unintentional actions. A successful answer to the second question would provide one with the theoretical machinery to use in explaining why you are reading this entry and why the author wrote it.

INTENTIONAL ACTION AND INDIVIDUATION

According to an attractive causal theory, intentional actions are, in one important respect, like money. The piece of paper with which Ann just purchased her drink is a genuine U.S. dollar bill partly in virtue of its having been produced (in the right way) by the U.S. Treasury Department. A duplicate bill produced with plates and paper stolen from the Treasury Department is a counterfeit bill, not a genuine one. Similarly, according to one kind of causal theory of intentional action, a certain event is Ann's buying a drink—an intentional action—partly in virtue of its having been produced in the right way by certain mental items. An event someone else covertly produces by remote control—one including visually indistinguishable bodily motions not appropriately produced by Ann's intentions or decisions (nor by physical states or events that realize the mental items)—is not Ann's intentional action, even if she feels as though she is in charge. (This view does not identify intentional actions with nonactional events—or nonintentional actions—caused in the right way. That would be analogous to identifying genuine U.S. dollar bills with pieces of printed paper that are not genuine U.S. dollar bills and are produced in the right way by the U.S. Treasury Department, which is absurd.)

The question "What are intentional actions?" directly raises two other questions. "How do intentional actions differ from everything else?" and, "How do intentional actions differ from one another?" A crude sketch of one answer to the first question about differences has just been provided. Intentional actions differ from other events in their causal history. Events that are intentional actions are produced in a certain way by mental items (or physical states and events that realize these items); events that are not intentional actions lack such a causal history (a topic picked up again in section 2.) Alternative conceptions of intentional action include (1) an internalist view, according to which intentional actions differ experientially from other events in a way that is essentially independent of how, or whether, they are caused; (2) a conception of intentional actions as composites of nonactional mental events or states (e.g., intentions) and pertinent nonactional effects (e.g., an arm's rising); and (3) views identifying an intentional action with the causing of a suitable nonactional product by appropriate nonactional mental events or states—or, instead, by an agent.

A debate over the second question about differences—the question of action individuation—has produced a collection of relatively precise alternatives: a coarse-grained view, a fine-grained view, and componential views. Donald Davidson writes, "I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home" (1980, p. 4). How many actions does the agent, Don, perform? Davidson's coarse-grained answer is one action "of which four descriptions have been given" (p. 4). The action is intentional under certain descriptions (e.g., "I flip the switch"), and unintentional under others (e.g., "I alert the prowler"). A fine-grained alternative view treats *A* and *B* as different actions if, in performing them, the agent exemplifies different action properties. In this view, Don has performed at least four actions (only some of which are intentional), because the action properties at issue are distinct. An agent may exemplify any of these action properties without exemplifying any of the others. One may even turn on a light in a room without illuminating the room (the light may be painted black). Componential views represent Don's illuminating the room as an intentional action having various components, including—but not necessarily limited to—his moving his arm, his flipping the switch, and the light's going on. Where proponents of the coarse-grained and fine-grained theo-

ries find, respectively, a single action under different descriptions and a collection of intimately related actions, advocates of the various componential views locate a larger action having smaller actions among its parts.

Davidson and Jennifer Hornsby hold that every action is intentional under some description. Proponents of alternative theories of action individuation may make an analogous claim: in every case of action something is done intentionally; when nothing is done intentionally, no action is performed. Where Davidson and Hornsby seek to distinguish descriptions under which an action is intentional from descriptions under which it is not, other philosophers may seek to distinguish intentional from unintentional actions in the same case of action. Either way, intentional actions are of primary importance.

This entry proceeds in a neutral way regarding the leading contending theories of individuation. Readers may treat the action variable *A* as a variable either for actions themselves (construed componentially or in a more fine-grained way) or for actions under *A*-descriptions, depending on their preferred mode of action individuation. The same goes for the term *action*.

CAUSALISM: BACKGROUND AND A CHALLENGE

One approach to understanding both the nature of intentional action and the explanation of intentional actions emphasizes causation. The conjunction of the following two theses may be termed *standard causalism*: (1) An event's being an intentional action depends on how it was caused; and (2) Proper explanations of intentional actions are causal explanations. Familiar causal theories feature as causes such psychological or mental items as beliefs, desires, intentions, and such related events as acquiring an intention to *A*.

Causalism typically is embraced as part of a naturalistic stand on agency, according to which mental items that play causal/explanatory roles in intentional action are in some way dependent on or realized in physical states and events. A range of options is open. Indeed, any viable solution to the mind-body problem that supports the idea that the mental has a significant causal/explanatory role in intentional action would, in principle, be welcomed by causalists.

Aristotle endorses the idea that intentional actions are to be explained, causally, in terms of mental states or events in his assertion that "the origin of action—its efficient, not its final cause—is choice, and that of choice is desire and reasoning with a view to an end" (Aristotle 1984, 1139a31–32). Davidson, in an influential article, "Actions, Reasons, and Causes," rebuts arguments against causalism, develops a positive causalist view, and presents noncausalists with what has proved to be a difficult challenge. Addressed to philosophers who hold that when people act intentionally they act for reasons, the challenge is to provide an account of the reasons for which people act that does not treat (people's having) those reasons as figuring in the causation of the relevant behavior (or, one might add, as realized in physical causes of the behavior). The challenge is acute when an agent has more than one reason for *A*-ing but *A*-s only for only one of them. Imagine that Al has a pair of reasons for mowing his lawn this morning. First, he wants to mow it this week and he believes that this morning is the most convenient time. Second, he has an urge to repay his neighbor for the rude awakening Al suffered recently when the neighbor turned on her mower at the crack of dawn; he believes that his mowing his lawn this morning would repay her. As it happens, Al mows his lawn this morning only for one of these reasons. In virtue of what is it true that he mowed his lawn for this reason, and not the other, if not that this reason—or his having this reason or what realizes either this reason or his having it—and not the other, played a suitable causal role in his mowing his lawn? Alfred Melerebuts detailed noncausalist attempts to answer this challenge in chapter two of *Motivation and Agency*. Space constraints preclude pursuing the issue here.

TWO ALLEGED PROBLEMS FOR CAUSALISM

Two alleged problems for causalism that continue to be lively topics of debate are causal deviance and vanishing agents.

CAUSAL DEVIANCE. Deviant causal chains raise difficulties for causal analyses of action itself and of doing something intentionally. The alleged problem is that whatever psychological causes are claimed to be both necessary and sufficient for a resultant event's being an action, or for an action's being intentional, cases can be described in which, owing to a deviant causal connection between the favored psychological antecedents—for example, events of intention acquisition—and a resultant event, that event is not an action, or a pertinent resultant action is not done intentionally.

The most common examples of deviance divide into two types: (1) Examples of primary deviance, which raise a problem about a relatively direct connection between

mental antecedents and resultant bodily motion; and (2) examples of secondary deviance, which highlight behavioral consequences of intentional actions and the connection between these actions and their consequences. In Davidson's well-known example of primary deviance, "A climber … want[s] to rid himself of the weight and danger of holding another man on a rope, and he … know[s] that by loosening his hold on the rope he [can] rid himself of the weight and danger. This belief and want … so unnerve him as to cause him to loosen his hold" unintentionally (1980, p. 79). In his equally well-known example of secondary deviance, "A man [tries] to kill someone by shooting at him. [He] misses his victim by a mile, but the shot stampedes a herd of wild pigs that trample the intended victim to death" (p. 78).

Instructive attempts to resolve the problems examples such as these pose highlight four points:

((1)) An event is an intentional action only if it is an action, and in many cases of deviance the pertinent event seems not to be an action. For example, the climber's "loosening his hold" is more aptly described as the rope's slipping from his trembling fingers.

((2)) An analysis of intentional action may preclude there being a gap between an action's psychological causal initiator and the beginning of the action. If, for example, every intentional action has the acquisition of a proximal intention—that is, an intention to *A* now or an intention to *A*, beginning now—as a proximate cause, there is no room between cause and the beginning of action for primary deviance. ("*Proximate* cause" may be defined as follows: *x* is a proximate cause of *y* if and only if *x* is a cause of *y* and there is nothing *z* such that *x* is a cause of *z* and *z* is a cause of *y*.)

((3)) Intention (or one's preferred psychological item) has a continuous guiding function in the development of intentional action.

((4)) An action's being intentional depends on its fitting the agent's conception or representation of the manner in which it will be performed—a condition violated in Davidson's shooting scenario.

George Wilson challenges point 2. Sometimes, Wilson observes, "intentions cause states of nervous agitation that positively *enable* the agent to perform the type of action intended" (1989, p. 252). He offers the example of a weightlifter whose "intention to lift the weight then caused a rush of nervous excitement that was, in fact, necessary for him to budge the great weight even slightly from off the floor" (1989, p. 252). However, this observation and example arguably leave the requirement of proximate causation unscathed. What is required is not that intention-inspired nervousness, agitation, and the like, play no role in the production of intentional actions, but rather that they not fill a gap between the acquisition of a pertinent proximal intention and action in such a way that intention acquisition figures only indirectly in the production of the corresponding action. In Wilson's example, one may contend, there is no gap between intention acquisition and the beginning of the lifting that is filled by nervousness. Rather, one may argue that intention acquisition proximately initiates the lifting—which action, according to some causalists, begins with a relevant brain event prior to the weight's rising—while also producing nervousness that is required for the agent's even budging the weight.

Proximal intentions typically are not momentary states, and the intention to lift the weight in the present case is at work as long as the lifting continues. Even if nervousness were somehow required for the occurrence of the agent's muscular movements themselves, a nervousness producing proximal intention to lift the weight whose acquisition plays a causal role in the production of a corresponding intentional lift would, in conjunction with the resultant nervousness, figure in the proximate initiation of those movements. If, alternatively, the causal role of an intention to lift the weight were exhausted by the intention's issuing in nervousness, and the nervousness were somehow to result in the upward movement of limbs and weight independently of any pertinent intention present at the time, the weightlifting would not be intentional. The case—aside from its failure to provide an intuitively appealing mechanistic explanation of the focal occurrence—would then be on par with familiar examples of nonintentional occurrences caused by intention-inspired nervousness (e.g., the climber's case).

The point about the continued functioning of proximal intentions blunts an objection John Bishop (1989) raises to Myles Brand's position on primary deviance. Bishop observes that deviance can break in after intention acquisition has (properly) initiated a causal chain—but before bodily movement occurs—and strip agents of control over their motions. In such cases, although agents' motions may accord with their intentions, they do not act intentionally. On Brand's view, however, the proximal intentions that initiate intentional actions also sustain and guide them: "Given that intention is in part guidance … of activity, the intention continues as long as

guidance … continues" (1984, p. 175). In a case of the kind Bishop imagines, guidance is absent.

Some causal theorists who have assessed cases of primary deviance as attempted counterexamples to a causal account of what it is for an action to be intentional have dismissed them on the grounds that they are not cases of action at all. If this diagnosis is correct, primary deviance poses an apparent problem for the project of constructing a causal analysis of action. Can causalists identify something of a causal nature in virtue of which it is false that the climber performed the action of loosening his grip on the rope?

In a discussion of primary deviance, Alvin Goldman remarks: "A complete explanation of how wants and beliefs lead to intentional acts would require extensive neurophysiological information, and I do not think it is fair to demand of a philosophical analysis that it provide this information.… A detailed delineation of the causal process that is characteristic of intentional action is a problem mainly for the special sciences" (1970, p. 62). This remark may strike some readers as evasive, but Goldman has a point. A deviant causal connection between an *X* and a *Y* is deviant relative to normal causal connections between *X*-s and *Y*-s. Moreover, what counts as normal in this context is perspective-relative. From the point of view of physics, for example, there is nothing abnormal about Davidson's examples of deviance. And, for beings of a particular kind, the normal route from intention to action may be best articulated partly in neurophysiological terms.

One way around the problem posed by incomplete neuroscientific knowledge is to design (in imagination, of course) an agent's motor control system. Knowing the biological being's design in that sphere, there is then a partial basis for distinguishing causal chains associated with overt action—that is, action essentially involving peripheral bodily motion—from deviant motion-producing chains. If one can distinguish deviant from non-deviant causal chains in designed agents—that is, chains not appropriate to action from action-producing chains—then the same may also be done for normal human beings, if much more than is currently known about the human body is discovered. (This line of thought is pursued in Mele 2003, ch. 2).

VANISHING AGENTS. Some philosophers claim that causalism precludes there being any actions at all and therefore makes agents vanish. According to Thomas Nagel, "The essential source of the problem is a view of persons and their actions as part of the order of nature.… That conception, if pressed, leads to the feeling that we are not agents at all.… *My doing* of an act—or the doing of an act by someone else—seems to disappear when we think of the world objectively. There seems no room for agency in [such] a world.… There is only what happens" (1986, pp. 110–111).

Nagel's worry is not worrisome. Cats and dogs are part of the natural order. If radical skeptical hypotheses are set aside—for example, the hypotheses that everything is a dream and that all biological entities are brains in vats—it is plain that cats and dogs act. They fight, eat, and play. When they do these things they are acting. The same is true of humans, even if people are part of the natural order. Supernatural beings (e.g., gods and ghosts) are not part of the natural order. That a being needs to be supernatural in order to act is an interesting proposition, but it is difficult to take that proposition seriously in the absence of a powerful argument for it.

J. David Velleman voices a variant of Nagel's worry. He contends that standard causal accounts of intentional action do not capture what "distinguishes human action from other animal behavior" and do not accommodate "human action par excellence" (2000, p. 124). He also reports that his objection to what he calls "the standard story of human action" (p. 123), a causal story, "is not that it mentions mental occurrences in the agent instead of the agent himself [but] that the occurrences it mentions in the agent are no more than occurrences in him, because their involvement in an action does not add up to the agent's being involved" (p. 125). Velleman says that this problem would remain even if the mind-body problem were solved, and, like Nagel, he regards the problem as "distinct from the problem of free-will" (p. 127).

Here, Velleman runs together two separate issues. Human agents may be involved in some of their actions in ways that cats and dogs are involved in many of their actions. Human agents do not vanish in such actions. Scenarios in which human agents vanish are one thing; scenarios in which actions of human agents do not come up to the level of human action par excellence, whatever that may be, are another.

Causalists are entitled to complain that Velleman has been unfair to them. His description of the standard story of human action is apparently a description of the sort of thing found in the work of causalists looking for what is common to all (overt) intentional actions, or all (overt) actions done for reasons, and for what distinguishes actions of these broad kinds from everything else. If some nonhuman animals act intentionally and for reasons, a story with that topic definitely should apply to them.

Also, human action par excellence may be intentional action and action done for a reason in virtue of its having the properties identified in standard causal analyses of these things. That the analyses do not provide sufficient conditions for—or a story about—human action par excellence is not a flaw in the analyses, given their targets. If Velleman were to believe that causalism lacks the resources for accommodating human action par excellence, he may attack the standard story on that front, arguing that it cannot be extended to handle such action. But Velleman himself is a causalist. Moreover, causalists have offered accounts of kinds of action—for example, free or autonomous action and action exhibiting self-control (the contrary of weakness of will)—that exceed minimal requirements for intentional action or action done for a reason. Their story about minimally sufficient conditions for action of the latter kinds is not their entire story about human actions.

REASONS, DESIRES, AND INTENTIONS

Reasons, desires, and intentions are featured in many theories about how intentional actions are to be explained. According to Davidson's influential view, reasons for action are complexes of beliefs and desires. Some philosophers claim that Davidsonian reasons for action really are not reasons at all. T. M. Scanlon, for example, argues that "desires almost never provide reasons for action in the way described by the standard desire model" (1998, p. 43).

Philosophical work on reasons for action tends to be guided primarily either by a concern with the explanation of intentional actions or by a concern with the evaluation of intentional actions or their agents. In work dominated by the former concern, reasons for action tend to be understood as states of mind, along broadly Davidsonian lines. Philosophers with the latter concern may be sympathetic or unsympathetic to this construal, depending on their views about standards for evaluating actions or agents. For example, a theorist whose evaluative concern is with rational action and who holds that the pertinent notion of rationality is subjective—in the sense that a proper verdict about the rationality or irrationality of an agent's intentional action is to be made from the perspective of the agent's own desires, beliefs, principles, and the like, rather than from some external, or partly external, perspective—may be happy to understand reasons for action as states of mind. A theorist with a more objective conception of rational action or rational agency also is likely to have a more objective conception of reasons for action. Such a theorist may find it natural to insist that many or all reasons for action are facts about the agent-external world. Consider Bob's starting a new diet after his doctor informs him that his cholesterol is dangerously high. Theorists with a subjective conception of rationality tend to regard Bob's reasons for starting the new diet as constituted by desires and beliefs (e.g., his desire to improve his health and his belief that the new diet will help him do that), whereas theorists with an objective conception of rationality tend to regard his reasons as objective facts (e.g., the diet will improve his health, or it is likely to do so). Alleged reasons of these two types may be termed, respectively, *agent-internal* and *agent-external* justificatory reasons.

COMBINING AGENT-INTERNAL AND AGENT-EXTERNAL REASONS. If there are agent-external justificatory reasons for action, it may be that intentional actions are to be relatively directly explained at least partially in terms of Davidsonian reasons, and that when agent-external justificatory reasons—for example, the new diet is likely to improve Bob's health—contribute to explanations of intentional actions, they do so less directly, by way of a causal contribution made by an agent's *apprehending* such a reason. For example, Bob's apprehension of the likelihood that the new diet will improve his health might, along with his desire for improved health, enter into a true causal explanation of Bob's starting the new diet. An exploration of the possibility of agent-external justificatory reasons and of their compatibility with the existence of Davidsonian reasons quickly takes one well beyond the philosophy of action into moral philosophy and value theory. Further discussion of this topic is beyond the scope of the present entry, but is discussed in chapters three through six of Mele's *Motivation and Agency* (2003).

DESIRES. There is a related controversy about the nature of desires. Scanlon's critique of what he calls "the standard desire model" (1998, p. 43) is framed partly in terms of his own account of "what is usually called desire" (p. 65). He contends that something's seeming to an agent to be a reason for *A*-ing is "the central element in what is usually called [a] desire" to *A* (p. 65). Seemings of this kind do important motivational work, according to Scanlon. He claims that in a thirsty man with a desire to drink, "the motivational work seems to be done by" the agent's taking "the pleasure to be obtained by drinking … to count in favor of drinking" (p. 38).

Scanlon's account of what is usually called a desire is overly intellectualized. Toddlers and pretoddlers are commonly thought to desire to do things—for example, to

drink some juice or to hug a teddy bear. This common thought is not that although these little agents have desires to act, they lack what is usually called a desire. The thought is that they have desires in a usual sense of the term. But because it is unlikely that toddlers have the concept of a reason for action (or of something's counting in favor of a course of action), it is unlikely that things seem to them to be reasons for action (or to count in favor of actions). There is good evidence that younger three-year-olds tend not to have the concept—or a proper concept—of belief and that the concept of desire normally does not emerge until around the age of two. Presumably, even if the concept of a reason for action were to have no conceptual ties to the concepts of belief and desire, it would be sufficiently sophisticated to be out of reach of children too young to have proper concepts of belief and desire. Even so, it is commonly and plausibly thought that such children act intentionally and for reasons. (They also have desires and beliefs, on the assumption that having such attitudes does not require possessing proper concepts of these attitudes.) In thirsty toddlers or pretoddlers, desires to drink—rather than any taking of the pleasure to be obtained by drinking to be a reason for drinking—seem to do the work of motivating drinking.

Thirsty toddlers are attracted by cups of juice, and not in the way moths are attracted by light. Toddlers are flexible in their approach to getting drinks: they try alternative means. Moths behave tropistically. Even though it is unlikely that thirsty toddlers have the conceptual wherewithal to take features—including anticipated consequences—of drinking to be reasons for (or count in favor of) drinking, they are attracted by cups of juice in a way characteristic of desiring agents. Being attracted to cups of juice owing to a sensitivity to certain of their features is distinguishable from being attracted to cups of juice owing to the agent's taking those features to be reasons. An agent's behavior may be sensitive to attractive features of things without the agent's taking those features to be reasons. If this were not so, a radically new theory of animal behavior would be required, one entailing either that only members of the most conceptually sophisticated species ever act intentionally (perhaps just human beings) or that many nonhuman species are much more conceptually sophisticated than anyone has thought.

When ordinary thirsty adults drink (intentionally, and in ordinary scenarios), they presumably are motivated at least partly by a desire to drink. The strength of the desire may sometimes be explained partly by their believing that drinking would be pleasant, or, more fully, by that belief together with a desire for pleasure. A toddler's desire to drink water and an adult's desire to drink water may admit of the same analysis. Just as something's seeming to be a reason for drinking is not a constituent of the toddler's desire, it may not be a constituent of the adult's desire either. If a seeming of this kind sometimes is at work in thirsty adults, it may function as a partial cause of the desire's strength or of the desire itself.

INTENTIONS. Next on the agenda are intentions, states of mind commonly regarded as being closely linked to desires and beliefs. Intention has a motivational dimension, and the word *desire* (like the word *want*) is often used in the literature as a generic term for motivation. Intention also is widely regarded as involving a belief condition of some sort. Few people are inclined to say that gamblers who believe that their chances of winning today's lottery are about one in a million intend to win the lottery. However, philosophers disagree about the tightness of the connection between intentions, on the one hand, and desires and beliefs, on the other. Some—attracted, perhaps, by the idea that desire and belief are the most fundamental representational states of mind—argue that intentions are reducible to combinations of desires and beliefs, whereas others argue that attempts at such reduction are doomed to failure.

The central issue is whether the settledness that intention encompasses can be articulated in terms of beliefs and desires. Ann wants to go to a 7:00 movie and she wants to attend a 7:00 lecture. She knows that she can do either but not both. Although Ann wants to see the movie more than she wants to attend the lecture and believes that, given what she usually does in such situations, she will probably go to the movie, she is unsettled about what to do. After further deliberation, Ann settles matters for herself by deciding to attend the lecture. In so deciding, she forms an intention to attend it. To intend to *A* is, at least in part, to be settled (but not necessarily irrevocably) on *A*-ing. Wanting or desiring to *A*—even when the desire is stronger than its competitors, and even when it is accompanied by a belief that one probably will *A*—is compatible with being unsettled about whether to *A*.

Functions plausibly attributed to intentions include initiating and sustaining intentional actions, guiding intentional actions, helping to coordinate agents' behavior over time and their interaction with others, and prompting and appropriately terminating practical reasoning. Some philosophers have advanced nonreductive accounts of intention designed to accommodate many or all of these functions. According to a representative

account of this kind, intentions are executive attitudes toward plans. Plans—which range from simple representations of simple actions to complex strategies for achieving remote goals—constitute the representational contents of intentions. What distinguishes intentions from other practical attitudes (e.g., desires to act), in this account, is their executive nature. The settledness on *A*-ing that is encompassed in an intention to *A* is a psychological commitment to executing the intention-embedded plan of action, a commitment of a kind arguably constituted exclusively by intentions.

ANALYZING INTENTIONAL ACTION: DIFFICULTIES

Attention to a trio of problems for the following pair of protoanalyses of intentional action sheds light on what the difficult project of analyzing intentional action encompasses:

> *A1*. *S* intentionally *A*-ed if and only if *S A*-ed in the way that *S* intended to *A*.
>
> *A2*. *S* intentionally *A*-ed if and only if *S A*-ed for a reason.

SIDE EFFECTS. Gilbert Harman discusses a scenario in which "In firing his gun," a sniper who is trying to kill a soldier, "knowingly alerts the enemy to his presence" (1997, p. 151). Harman claims that although the sniper "does not intend to alert the enemy," he intentionally alerts the enemy, "thinking that the gain is worth the possible cost." If Harman is right, both *A1* and *A2* are false. The sniper does not intend to alert the enemy, and he does not alert them for a reason either (even if his alerting them is part of some larger action that he does for a reason).

Because Harman's sniper does not unknowingly or accidentally alert the enemy, many people will deny that the sniper unintentionally alerted them. But the truth of that denial is consistent with the action's not being intentional, if there is a middle ground between intentional and unintentional action. Arguably, actions that an agent in no way aims at performing but that are not performed unknowingly or accidentally are properly located on that middle ground. They may be nonintentional, as opposed to unintentional. Of course, it also is arguable that Harman correctly assesses the sniper's case and that *A1* and *A2* are far too simple to be true.

BELIEF CONSTRAINTS. Some putative belief constraints on intentions or on rational intentions also pose problems for *A1*. Michael Bratman argues that intention has a normative side that requires that an agent's intentions be internally consistent (individually and collectively), consistent with the agent's beliefs, and means-end coherent. Rational intentions, he maintains, satisfy those requirements, and he contends that agents rationally intend to *A* only if, "other things being equal," they do "not have beliefs inconsistent with the belief that [they] will *A*" (1987, p. 116).

The normative demands figure prominently in an argument Bratman advances against what he calls "the Simple View"—the thesis that intentionally *A*-ing entails intending to *A*. The argument revolves around an example involving a pair of video games and an ambidextrous player who shall be called Vic. Vic's task is to hit targets with missiles. In the main case, he simultaneously plays two games, each with its own target and firing mechanism, and he knows that the machines are "so linked that it is impossible to hit both targets" (Bratman 1987, p. 114). (He knows that hitting a target ends both games, and that "if both targets are about to be hit simultaneously," both machines shut down before the targets can be hit.) Vic tries to hit the target on machine 1 while also trying to hit the target on machine 2. He succeeds in hitting the former—"in just the way that [he] was trying to hit it, and in a way which depends heavily on [his] considerable skill"—but, of course, he misses the latter.

If Vic hit target 1 intentionally, fans of the Simple View must say that he intended to hit it. Because Vic's attitude toward hitting that target is not relevantly different from his attitude toward hitting target 2, Simple View fans apparently must also say that he intended to hit target 2. Bratman contends that having both intentions, given what Vic knows—namely, that he cannot hit both targets—would be irrational. Yet, it seems perfectly rational of Vic to have proceeded as he did. So given the point about the symmetry of Vic's attitudes toward the targets, Bratman concludes that he did not have either intention. And if Vic hit target 1 intentionally in the absence of an intention to hit it, the Simple View and *A1* are false.

Some critics of the Simple View, including Bratman and Harman, also reject the idea that intentions are reducible to complexes of beliefs and desires; Hugh McCann argues that they are in danger of having to settle for an unwanted reductive analysis of intention (1998). Bratman, who suggests that a "guiding desire" (e.g., to hit target 1) can play the role of an intention (Bratman 1987, p. 137), is McCann's main target. McCann notes that once it is conceded that desires can stand in for intentions, reductionists will justifiably ask what need there is for a

notion of intention that is irreducible to desire and belief. However, philosophers who reject the Simple View need not follow Bratman in appealing to guiding desires. For example, it may be argued that intentions to try to *A* can stand in for intentions to *A*, and, of course, intentions to try to *A* are intentions. Presumably, Vic intends to try to hit target 1 while also intending to try to hit target 2.

LUCK. Instances of lucky success pose problems for *A1* and *A2*. Beth, who has never fired a gun, mistakenly thinks that modern technology makes target shooting fool proof, and she intends to hit the bull's-eye on a distant target by aiming and firing at it. She luckily hits it in just the way she intended, but was her hitting it an intentional action? Suppose that Beth has no natural talent with firearms: she fires hundreds of additional rounds at the target and does not even come close. Here philosophers' intuitions differ. According to Christopher Peacocke (1985), an agent who makes a successful attempt "to hit a croquet ball through a distant hoop" intentionally hits the ball through the hoop (p. 69). But Brian O'Shaughnessy (1980) maintains that a novice who similarly succeeds in hitting the bull's-eye on a dart board does not intentionally hit the bull's-eye. Readers inclined to regard Beth's hitting the bull's-eye as an intentional action should consider her brother Bob. He wants to save his town by disarming a bomb, and he believes that his punching in any ten-digit sequence of numbers will disarm it. In fact, only one ten-digit code will work. Bob intends to disarm the bomb by entering ten digits. If he luckily punches in the right code, thereby disarming the bomb, is his disarming it an intentional action? Or was his chance of success too low for that action to count as intentional? If the correct answer to the latter question is yes, *A1* is false.

Protoanalysis *A2* also is threatened by stories such as these. Probably, many people would happily (but perhaps mistakenly) say that Bob's disarming the bomb—that action—was done for a reason. After all, he wanted to save the town and knew that he must disarm the bomb to do so, and this helps to explain why he entered ten digits. But, again, was Bob's chance of success too low for the disarming to count as an intentional action?

Recall the two central questions identified in the introduction to this entry: What are intentional actions? And how are intentional actions to be explained? Depending on how nuanced a satisfactory answer to the first question is, philosophers of action working on the second question may do well to focus their efforts on core instances of intentional action. If the sniper's alerting the enemy is an intentional action, it is intentional in a different way than his firing his gun is. He fires his gun as a means to an end, but this is not true of his alerting the enemy. He also intends to fire his gun and fires it for a reason, but he does not intend to alert the enemy and does not alert them for a reason. One approach in looking for core instances of intentional action is to look for interesting properties that all cases of intentional action have in common, even if not all intentional actions have them. It may be discovered that there are no cases of intentional action in which the agent does not perform any intended intentional actions. (Even if Vic lacks an intention to hit target 1 in the video games example, he intends to fire at it and he intentionally fires at it.) If so, it may be fruitful for philosophers of action to focus primarily on intended intentional actions in developing their theories about how intentional actions are to be explained—theories in light of which it can explained why the author wrote this entry and why you are reading it, and explain how those actions are produced. Possibly, theories of this kind can then be augmented to cover all intentional actions.

See also Agent Causation; Weakness of the Will.

Bibliography

Aristotle. *Nicomachean Ethics*. In *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Audi, Robert. *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press, 1993.

Bishop, John. *Natural Agency*. Cambridge, U.K.: Cambridge University Press, 1989.

Brand, Myles. *Intending and Acting*. Cambridge, MA: MIT Press, 1984.

Bratman, Michael. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.

Bratman, Michael. *Faces of Intention*. Cambridge, U.K.: Cambridge University Press, 1999.

Davidson, Donald. "Actions, Reasons, and Causes." *Journal of Philosophy* 60 (1963): 685–700. Reprinted in Davidson 1980.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Frankfurt, Harry. *The Importance of What We Care About*. Cambridge, U.K.: Cambridge University Press, 1988.

Ginet, Carl. *On Action*. Cambridge, U.K.: Cambridge University Press, 1990.

Goldman, Alvin. *A Theory of Human Action*. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Harman, Gilbert. "Practical Reasoning." In *The Philosophy of Action*, edited by Alfred Mele. Oxford: Oxford University Press, 1997.

Hornsby, Jennifer. *Actions*. London: Routledge, 1980.

McCann, Hugh. *The Works of Agency*. Ithaca, NY: Cornell University Press, 1998.

Mele, Alfred. *Motivation and Agency*. New York: Oxford University Press, 2003.
Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.
O'Shaughnessy, Brian. *The Will*. Cambridge, U.K.: Cambridge University Press, 1980.
Peacocke, Christopher. "Intention and Akrasia." In *Essays on Davidson*, edited by Bruce Vermazen and Merrill Hintikka. Oxford: Clarendon Press, 1985.
Ruben, David-Hillel. *Action and Its Explanation*. Oxford: Clarendon Press, 2003.
Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
Schueler, G. F. *Reasons and Purposes*. Oxford: Clarendon Press, 2003.
Searle, J. *Intentionality*. Cambridge, U.K.: Cambridge University Press, 1983.
Velleman, J. David. *The Possibility of Practical Reason*. Oxford: Clarendon Press, 2000.
Wilson, George. *The Intentionality of Human Action*. Stanford, CA: Stanford University Press, 1989.

Alfred R. Mele (1996, 2005)

ADDISON, JOSEPH

(1672–1719)

Joseph Addison—Oxford scholar, poet, playwright, essayist, and politician—figures in the history of philosophy chiefly on the strength of his *Essay on the Pleasures of the Imagination*, published in 1712 as numbers 411 through 421 of his and Richard Steele's journal *The Spectator*.

Addison defines "pleasures of the imagination" as "such [pleasures] as arise from visible objects" (no. 411). He calls "primary" those derived from things present to vision, "secondary" those derived from things merely called to mind. There are three qualities of objects from which the primary pleasures may arise: greatness, novelty, and beauty. Greatness is an extensiveness that throws the viewer into "a pleasing astonishment," as in, for example, the sight of a mountain range. Novelty includes what is new or unfamiliar to the viewer, as a fresh meadow in spring may be, as well as what continually changes its appearance, for example, a waterfall. Beauty includes, on the one hand, whatever appearances effect sexual attraction, and on the other, "the gaiety or variety of colors," "the symmetry and proportion of parts," and "the arrangement and disposition of bodies" (no. 412).

Addison's account of the secondary pleasures is more complex. Such pleasures may be produced by mere spontaneous imaginings, or by representational artifacts, such as sculptures, paintings, some pieces of music, and descriptions. In these cases, we derive pleasure not merely from the object imagined, but also from the comparison of that object with that which represents it (no. 416). Addison also invokes comparison to explain the pleasure that we take in fictional descriptions of terrible things and events: our pleasure derives from our awareness that we ourselves are not actually threatened by the evils about which we read (no. 418).

Addison's *Essay* has been taken to mark the beginning of modern aesthetics. There are several grounds for such a claim. Addison, in contrast to previous writers on his various topics, investigates pleasures that can be derived from art and nature equally, treats the beautiful as merely one among several pleasing visual qualities, and centers his account on the mental activity of the onlooker rather than on the character of the object viewed. In all these respects, his *Essay* sets the direction for subsequent work in aesthetics.

At the same time, there are considerable differences of purview between Addison's investigation and later aesthetic thought. The sources of the pleasures of the imagination include works of art only so far as these either please the eye or awaken visual images; they do not include nonprogrammatic music, or even the nonimagistic aspects of literature. Further, for Addison, works of history, natural philosophy, travel narrative, and even criticism, morals, and speculative philosophy (so far as these use visual figures of speech) may be sources of the pleasures of the imagination just as much as works of fiction (nos. 420–421). Thus, for all the concerns and assumptions that Addison shares with subsequent writers on taste and the fine arts, the scope of his inquiry is distinctively his own.

See also Aesthetics, History of.

Bibliography

Addison, Joseph, and Richard Steele. *The Spectator.* 5 vols, edited by Donald F. Bond. Oxford: Clarendon Press, 1965. Nos. 411–421 are in vol. 3.
Hipple, Walter John, Jr. *The Beautiful, the Sublime, and the Picturesque in Eighteenth-Century British Aesthetic Theory*. Carbondale: Southern Illinois University Press, 1957. Chapter 1 is an exposition of Addison's *Essay*.
Rind, Miles. "The Concept of Disinterestedness in Eighteenth-Century British Aesthetics." *Journal of the History of Philosophy* 40 (2002): 67–87. A critique of Jerome Stolnitz's historiography of aesthetics.
Stolnitz, Jerome. "On the Origins of 'Aesthetic Disinterestedness.'" *Journal of Aesthetics and Art Criticism* 20 (1961): 131–144. A highly influential study that credits Addison with originating the concept of "aesthetic perception."

Walker, William. “Ideology and Addison’s Essays on the Pleasures of the Imagination.” *Eighteenth-Century Life* 24 (2000): 65–84. A critique of the widely held view that Addison’s *Essay* is a work of “bourgeois ideology.”

Miles Rind (2005)

ADLER, ALFRED

(1870–1937)

Alfred Adler, the medical psychologist and founder of Individual Psychology, was born in Vienna of Hungarian-Jewish parents. He received his MD from the University of Vienna in 1895 and practiced general medicine before turning to psychiatry. His soundest scientific works were written before World War I and largely prepared during his ambivalent association with the early Freudian group. After serving in the Austrian army he became concerned with child guidance as a method of preventive medical psychology, and gaining favor with the new Austrian government, opened child-guidance centers in Vienna, Berlin, and Munich schools. Family-guidance interviews in public, with general discussion periods, disseminated his methods and theories, particularly among educators. He became an international lecturer in Europe and the United States and was America’s first professor of medical psychology, at Long Island Medical School. In the 1930s his efforts to spread his doctrine of “social interest” in the face of Europe’s totalitarian nationalisms marked him as preacher rather than scientist, and his later published work served to promulgate a faith rather than to report scientific work. He died in Aberdeen, Scotland, during a lecture tour.

Adler’s first psychologically important work, the *Study in Organ Inferiority and Its Psychical Compensation* (1907), was “a contribution to clinical medicine” in constitutional pathology. In it Adler explored constitutional defects of structure and function and their physiopathological compensation and also described “psychical” compensatory changes in disposition and way of life; overcompensation could produce not only “genius,” like the deaf Ludwig van Beethoven, but also neurotic or psychotic responses, like hysteria or paranoia. Adler gave a causal-deterministic exposition of development as dependent upon constitutional endowments, innate biological drives, and environmental pressures. His papers of 1908 described as innate an “aggression drive” (to subdue the environment) and a “need for affection.” Both concepts were then rejected by Sigmund Freud’s group but reappeared in later psychoanalytic theories.

Adler himself modified both concepts and reformulated his whole psychology in *The Neurotic Constitution* (1912). He repudiated drive psychology and causal determinism. He viewed inferiority (vis-à-vis adults) and consequent “inferiority feeling” as experiences common to every child. The child responds as a whole individual with a “striving for superiority” (the former “aggression drive”) directed toward a “fictive goal” of manly strength and dominance, which is pursued through a “guiding fiction,” or life plan, modified by the “antifiction” of social demands. Goal and fiction are subjective creations of the individual’s making, but unrealistic, rigid, neurotic patterns may be favored by organ inferiority, pampering, or neglect in childhood, or the child’s age-ranking in the family. To Adler the Nietzschean “will to power” was this kind of neurotic pattern, not a universal human trait. He also described an opposite but equally effective response to increased insecurity:

> It is one of the triumphs of human wit to put through the guiding fiction by adapting it to the anti-fiction, … to conquer by humility and submissiveness … to cause pain to others by one’s own suffering, to strive to attain the goal of manly force by effeminate means, to make oneself small in order to appear great. Of such sort … are often the expedients of neurotics.

In contrast to the neurotic, the psychotic character attempts to shape reality to the fiction, while the normal character adapts itself to the environment.

Adler’s later works reiterated, renamed, elaborated, and finally, simplified and broadened the concepts on which he had founded Individual Psychology in 1912 after breaking with Freud. The basis of character was the response of the whole individual to a universal infantile inferiority feeling. Accentuated inferiority feeling became the celebrated “inferiority complex,” and a pathological striving for superiority was a “superiority complex.” The guiding fiction was renamed the “life style,” usually unconscious or “not understood,” which Adlerian analysis endeavored to illuminate with insight. The antifiction and the early “need for affection” fused in the important concept of social interest. Adler first diverged from psychoanalysis over Freud’s emphasis on sexual instincts. Ultimately, where Freud saw animal instincts humanized through repression, Adler described inborn trends—social interest and striving for superiority—whose full development perfected the personality. In summary, “Heredity only endows [the individual] with certain abilities. Environment only gives him certain impressions … it is his individual way of using these bricks, … his atti-

tude toward life, which determines [his] relationship to the outside world."

Despite their differences, Adler always acknowledged his debt to Freud's psychogenetic theory of neurosis. He acknowledged Pierre Janet's *sentiment d'incomplétitude*, a predecessor of the inferiority feeling. Adler's formulation of personality somewhat resembled the "psychic structure" and "attitudes" of Wilhelm Dilthey's psychology, but direct influence is unlikely: Adler never mentioned Dilthey, although he did cite a work of Dilthey's contemporary Hans Vaihinger, the *Philosophy of "As If"* (New York, 1924), for the theory of fictions. Individual psychology had a brushfire success in continental Europe and the United States, rather less in Britain; everywhere it found more acceptance among educators, psychologists, even writers than among physicians and psychiatrists.

Adler's work has been largely absorbed into practice and thought without retaining a separate identity despite the familiar phrases—"overcompensation," "inferiority complex," "organ jargon"—which enrich a conversational rather than a psychological vocabulary. Individual Psychology still has its own centers, schools, and work groups, but Adler's influence has permeated other psychologies. His "aggression drive" reappeared in the ego psychology of orthodox psychoanalysis; other Adlerian echoes are found in Karen Horney, Harry Stack Sullivan, and Franz Alexander, and in Ian Suttie's mother-relationship theories, which surely influenced the contemporary mother-need ethological school. Child-guidance practice is non-Adlerian, and his name is not now invoked in progressive pedagogy, but those who try to see the backward child, the delinquent, the psychopath, or the psychiatric patient as a whole person are sharing Adler's viewpoint.

Adler's approach to psychology, normal and abnormal, was speculative rather than scientific. From 1912 on, he sought the elegantly economical theory rather than the proven fact. At first he recognized his theory as a fiction in Vaihinger's nonpejorative sense; a person behaves "as if" compensating for inferiority feeling. Later this step was omitted—these things *were* so. Adler often illustrated his theory with case material, but this was invariably anecdotal and in excerpts, never statistically organized. He openly despised statistics. It is uncertain how many patients Adler treated in continuity, apart from single consultations to advise physicians or teachers. The same case histories appear as examples through many books over many years, with no systematic follow-up. He made no use of normal "controls," an omission he justified by his insistence upon the uniqueness of the individual, but this left unsolved the problem of why one creative self chose neurosis, another not. Adler never experimented, never firmly predicted, never attempted systematically to verify a hypothesis. He had great intuitive insight, the greater, perhaps, for having grown up as a second son and a sickly rachitic child of a Hungarian-Jewish family in the Austrian imperial capital. His intuitions and their formulations, if not so close to reality as he believed, remain as valuable guiding fictions.

See also Dilthey, Wilhelm; Freud, Sigmund; Psychoanalysis; Psychology; Unconscious; Vaihinger, Hans.

Bibliography

WORKS BY ADLER

Studie über die Minderwertigkeit von Organen. Vienna: Urban and Schwarzenberg, 1907. Translated by Smith Ely Jelliffe as *Inferiority and Its Psychical Compensation*. New York: Nervous and Mental Diseases Publishing, 1917.

Über den nervösen Charakter. Wiesbaden: Bergmann, 1912. Translated by Bernard Glueck and John E. Lind as *The Neurotic Constitution*. New York, 1926.

Praxis und Theorie der Individualpsychologie. Munich: Bergmann, 1920. Translated by P. Radin as *Practice and Theory of Individual Psychology*. New York: Harcourt Brace, 1924.

Menschenkenntnis. Leipzig: Hirzel, 1927. Translated by W. Béran Wolfe as *Understanding Human Nature*. London: Allen and Unwin, 1927.

Understanding Life: An Introduction to the Psychology of Alfred Adler (1927).Edited by Colin Brett.Oxford; Rockport, MA: Oneworld, 1997.

The Education of Children (1930). Chicago: Regnery, 1970.

Der Sinn des Lebens. Vienna: Rolf Passer, 1933. Translated by John Winton and Richard Vaughan as *Social Interest; a Challenge to Mankind*. London: Faber, 1938.

Superiority and Social Interest: A Collection of Letter Writings. Routledge & K. Paul, 1965.

With Ernst Jahn. *Religion und Individualpsychologie: Eine prinzipielle Auseinandersetzung über Menschenführung*. Frankfurt am Main: Fischer-Taschenbuch-Verlag, 1975,

Kindererziehung. Frankfurt am Main: Fischer Taschenbuch, 1976.

Das Problem der homosexualitat und sexueller Pervessionen: Erotisches Training und erotischer Rückzug. Edited by Wolfgang Metzger. Frankfurt am Main: Fischer Taschenbuch Verlag, 1977.

Co-operation between the Sexes: Writings on Women, Love and Marriage, Sexuality, and Its Disorders. Edited and translated by Heinz Ludwig Ansbacher and Rowena R. Ansbacher. Garden City, NY: Anchor Books, 1978.

Journal Articles:, 1910–1913: Elaborating on the Basic Principles of Individual Psychology. Edited by Gerald L. Liebenau, and Henry T. Stein. San Francisco: Classical Adlerian Translation Project, Alfred Adler Institute of San Francisco, 2003.

WORKS ON ADLER

Ansbacher, Heinz, and Rowena Ansbacher, eds. *The Individual Psychology of Alfred Adler*. New York: Basic, 1956. Extracts, full bibliography, and critical annotations, including evaluation of Adler's proclaimed finalistic subjectivism as approximating William James's "soft determinism."

Orgler, Hertha. *Alfred Adler, the Man and His Work*. London: Daniel, 1939; 3rd ed., London: Sidgwick and Jackson, 1963.

J. D. Uytman (1967)
Bibliography updated by Michael J. Farmer (2005)

ADORNO, THEODOR WIESENGRUND

(1903–1969)

Theodor Wiesengrund Adorno, philosopher, composer, sociologist, and aesthetic theorist, was born September 11, 1903, in Frankfurt am Main and died August 6, 1969. His last days were beset by the "emergencies in democracy" prompted by the student movement of the 1960s; the students simultaneously treated him as friend and foe.

LIFE AND WORK

Studying in Frankfurt in the 1920s, but increasingly unable to secure employment in the first years of Nazi Germany, Adorno moved to England in 1934. Four years later, with his new wife, Margarethe ("Gretel") Karplus (1902–1993), he moved to the United States, first to New York and then to Los Angeles. In 1949 they returned to Frankfurt where Adorno worked both as professor at the university and as public intellectual, participating in radio and television programs on philosophy, society, education, and the arts.

Born into a comfortable bourgeois home, he was the only son of a Protestant wine merchant of Jewish descent, Oscar Wiesengrund, and of a Catholic singer, Maria Calvelli-Adorno. Before his move to the United States he was known by his father's name and after by his mother's. However, though "Wiesengrund" was abbreviated to a middle initial, the name was honored in Thomas Mann's *Doctor Faustus* (1947), the exemplary novel on the fate of musical modernism to which Adorno significantly contributed. The Beethovenian tones of the *Wiesengrund*—meadow-ground—expressed an early promise of happiness for the bourgeois age that would eventually be shattered, leaving the ill-fated dodecaphonic composer Adrian Leverkühn no choice but to complete his life with a melancholic requiem composed to the former greatness of German art.

Adorno wrote broadly on metaphysics, epistemology, political philosophy, ethics, the history of philosophy, and the philosophy of history. He is most widely known for his attempt to reveal the intricate historical and dialectical relationships between philosophy, society and the arts, or between philosophy, sociology, and aesthetic theory.

PHILOSOPHY AND MUSIC

In the 1920s, Adorno worked as a music critic reflecting upon contemporary developments in both the high and popular forms of the arts. Following his graduation in 1924 with a critical dissertation on Husserl's phenomenology he moved to Vienna to study composition with Alban Berg, a member alongside Arnold Schoenberg and Anton Webern of the Second Viennese School. Torn initially between philosophy and music he finally chose both, in this way furthering a tradition that had its beginnings with Plato. Following Schopenhauer, Kierkegaard, and Nietzsche (and knowledgeable of his contemporary Ernst Bloch), Adorno gave pride of place to music in his philosophical thinking and to philosophy in his musical thinking. However, he never aimed to reduce one to the other. He aimed neither to produce a philosophy *of* music nor, indeed, a philosophy *of* anything else, as if, by this use of "of," philosophy was assumed to be the master method to which all other disciplines were subject(ed). Philosophy, rather, was one of many nonreducible modes of thinking, and music was another, through which truth might be approached. Like music, philosophy was to be treated critically and self-reflectively; neither offered a guarantee regarding the good, the true, or the beautiful. Both were conditioned by what was going on in history and society. Yet both at best challenged the terms of that conditioning: philosophy by means of reason and music by means of expression.

Philosophy and music stood in an antagonistic but intimate relation. Because music was the exemplary language of pure expression but of no concept, and philosophy that of pure concept but no expression, each yearned, as if seeking a (Goethean) affinity, for what the other had—rational articulation for the one, and expression for the other. In their productive but troubled yearning they jointly tracked the historical course of modernity. Adorno focused predominantly on German philosophy and German music as both consummate and cautionary of enlightenment.

COLLABORATIVE PROJECTS

Temperamentally allied to the solitary thinkers and lonely composers of modernity, Adorno's thinking was shaped by notions of exile, otherness, and alienation. However, this did not render him merely an isolated or esoteric thinker; much of his work was produced collaboratively and often under the auspices of publicly sponsored research projects.

A leading member of the Frankfurt-based Institute of Social Research, he worked most closely with its founder Max Horkheimer, but so too with other members like Herbert Marcuse and Leo Löwenthal. In his early years he was in close contact with Walter Benjamin and Siegfried Kracauer. In New York he worked, albeit with difficulty, under the leadership of the Austrian exiled sociologist Paul Lazarsfeld on the Princeton Radio Project. He worked specifically on the empirical testing (a method of which he was highly critical) of listening habits, opinions, and tastes shaped by the then new means of technological production. A significant proportion of his writing on the arts was devoted to the mass media, to the radio, record player, television, and film, and particularly to the changes in modes of reception each instigated. Generally Adorno showed more interest in developing a critical, sociological aesthetic of the ear than of the eye. He did, however, think about the prohibition of the image and then about the adaptation of that prohibition to word and tone within an increasingly censorious society.

In Los Angeles he collaborated with Horkheimer in research on authoritarianism, fascism, anti-Semitism, and prejudice. To their results they linked descriptions of what came to be called the culture or mass entertainment industry, an industry of cultural production and propaganda devoted to "administering" public opinion and taste. In relation to philosophy, society, and the arts they traced the tendencies they took to be equally prevalent in Germany and America, although in different degrees and modes of advancement. They traced the tendencies toward mass consumerism and standardization, toward conformism and adaptation (as part of their critique of *identity thinking*), and toward domestication and normalization, as if, they argued, that which was being sold to the public as "the good, the true and the beautiful" was nothing but obviously "authentic," "natural," or "self-evident." They picked out these latter terms just because they were the ones most often used in public discourse, where the understanding was that to declare something self-evident, for example, rendered any further justification or reasoning unnecessary. In general, their work aimed to disassemble the philosophical illusions and aesthetic appearances that sustained a modern society of self-evidence. The work culminated in their jointly authored *Dialectic of Enlightenment. Philosophical Fragments* (1944), Adorno's *Philosophy of New Music* (1948), and Horkheimer's *The Eclipse of Reason* (1947).

In tandem with the work he did with Horkheimer, Adorno argued against the false rationalizations offered on behalf of mainstream social and aesthetic forms: the *pseudo- individualization* associated with the mainstream production of jazz and popular music, the *pseudo-ritualization* of some of Igor Stravinsky's music, and the pseudo-naturalism of some of John Cage's. He objected to contemporary appeals made on behalf of particular arts to return to ritual, nature, or the individual, as if these things had not suffered what society in general had suffered. All had suffered the consequences of an ideology of progress or of enlightenment ideals gone wrong. Adorno wanted the contemporary forms of art to take account of what had historically occurred and not assume that good-sounding ideas and ideals remained guiltlessly in place.

While working with Horkheimer and Mann, Adorno also collaborated with the composer Hanns Eisler, a student of Schoenberg and collaborator also with Bertolt Brecht, all of whom were contemporaneously resident in Los Angeles. With Eisler, Adorno furthered his sociological aesthetic of listening. Together they wrote a primer (1947) for the composition of a progressive or new music for the film. They framed their recommendations by a sustained critique of the increasingly dominant Hollywood film industry.

CRITICAL THEORY

Adorno contributed significantly to the development of critical theory, a dialectical, historical approach to both thinking and writing that unrelentingly aimed to expose the errors of the dominant scientistic, empiricist, and positivist methods of the day. In 1961, in Tübingen, he engaged in the so-called positivist dispute with, among others, Karl Popper and Jürgen Habermas. What he argued was just a continuation of his life-long double-pronged critique of a reductionist or eliminativist method, on the one hand, and an overly grounded or too securely founded totalizing metaphysics, on the other. (With the latter he usually associated the work of Heidegger and the postwar Heideggerians.) His work in aesthetic theory mirrored the same double-pronged critical aim.

Influenced by Goethe, Kant, Beethoven, and Hegel at the one end of modernity, and by the post- Marxists and Freudians, Lukacs, Kracauer, and Benjamin at the other, Adorno traced the convergences between philosophy, society, and the arts, or the dialectical movement of reason and irrationality that reached its inconceivable extreme in the Nazi concentration camps. Reversing Hegel's dictum that "the true is the whole"—where the whole is the positive and absolute completion of the dialectical movement of *Geist*—Adorno described the complex tendencies that had historically led toward untruth in its varying regressive and progressive concrete arrangements. He encapsulated his entire philosophical, sociological, and aesthetic reflections in the thought that there is no life—and thus no thought, no art, and no action—that is lived rightly when the whole is false.

Adorno focused on the major thinkers and artists of his times, for example: on Husserl and Heidegger in philosophy, on Schoenberg, Berg, Stravinsky, and Cage in music, and on Brecht, Kafka and Beckett in literature and drama. He did so partially to assess their historical relation to their great predecessors: Goethe, Schiller, Kant, Hegel, Beethoven, Kierkegaard, Wagner, Balzac, Valéry, George, and Proust, to name only a very few of the many writers who absorbed Adorno's indefatigable attention. He explored the tense relation between ideas of tradition, establishment, the accepted, and the expected, on the one hand, and ideas of the new, the unfamiliar, the unexpected, the explosive, and the shocking, on the other. (He particularly liked to work with an analogy between the artwork and the firework.) When he spoke of the old and the new, he most often thought, with Goethe, about how the new comes to suffer from its own aging. In other terms, his aesthetic reflections were also reflections constitutive of a *Geschichtsphilosophie*: a philosophy of history that would attempt to resist either falling into the safety of conservative, nostalgic, or utopian pastures, on the one hand, or reaching absolute or positive end points on a road that had no end, on the other. Most of his thinking aimed to invert the movement of Hegelian spirit in the light of the concrete social changes that had occurred between Hegel's time and his own.

TENDENCIES AND CATEGORIES

Adorno approached history by describing how the general social tendencies toward regression and progression were always mediated by concrete or particular instances. Though he had a rhetorical tendency to make it seem as if all the many thinkers, artists, and composers about whom he wrote would duly be lined up on the side of "the good" or of "the bad," his more subtle aim was to show how particular thoughts, works, or genres were constellations of contradictory tendencies. Indeed, to show them as such was to counter the very tendency to which his rhetorical tendency pointed, namely, the extreme polarization into which modern, administered society had placed its products and its persons.

Adorno focused on categorization, on the social dynamics of organization that included the stereotying and pigeonholing of persons, the social classification and marketing of the arts, as well as the construction and use of philosophical concepts. In his work on listening, he produced a taxonomy of listeners, to show less the type of which he approved (although his own tastes and preferences were always explicit in his critique), and more the types of listening that had developed in relation to the production of modern, "high" and "low" forms of music. Labels designating one sort of music as "serious," "elite," "esoteric," "difficult" or "incomprehensible" maintained a dialectical relation to those that designated another sort of music as "popular" and "authentic." On either side, the labels deflected the listener's attention from the music itself and refocused it in terms of what best suited the listener as consumer. Concepts of the high and low were not "givens" of aesthetic practice; they were sociological categories used to encourage musicians to produce musics of perfect fit, equally "hit tunes" or "difficult works."

AESTHETIC THEORY AND NEGATIVE DIALECTICS

Adorno may be read through his many essays and books amounting to more than 20 volumes. Or he may be read through his two masterworks, his *Negative Dialectics* of 1966 and his unfinished and posthumously published *Aesthetic Theory* of 1970. More specifically, whether one reads his early *Kierkegaard: Construction of the Aesthetic* or his exemplary essay on the "Social Situation of Music," or one of his monographs on Richard Wagner, Gustav Mahler, or Alban Berg, or whether, rather, one reads only his last works, one sees immediately that his primary interest in music never confined him to this particular art. Music was the model through which to access the entire domain of the aesthetic if not also society. He pursued most of the traditional problems of classical, romantic and modernist aesthetic theory: judgment and experience; the sublime and the beautiful; form, content, and material; genre, movement, and style (naturalism, realism, expressionism, and surrealism); the fateful, tragic, and the comic; art's relation to nature, to time, temporality, history and movement, and to society, poli-

tics, and propaganda. He drew upon many concepts unfamiliar to us today as well as upon concepts that at the time had become overly standardized through long term (mis)use, notably: mimesis, autonomy, expression, remembrance, comportment, commitment, and convergence.

Central to his aesthetic theory were two dialectical relationships, first, between the concept of art and that of the work of art; second, between the articulated and the hidden, concealed, or unexpressed dimensions of meaning. To regard a work of art as a constellation of contradictory impulses was to regard it as suspended between historical, social, and aesthetic demands: for example, following Kant, between the demand that the work be a product of labor and construction and the demand that it be a product of genius and thus appear as if natural, spontaneous, and free; or, following Schiller, that the work embody the mutually antagonistic drives toward form and sensuousness; or, following Hegel, that a work tremble between freedom and necessity, or between form and content, or between the demands of the traditional and the new, or between the repetition of the same and the shock of the different, or, finally, between acceptance and exemplarity.

To the extent that a work maintained the tension between conflicting demands, the work, so Adorno argued, was truthful. To resolve the tension in any given direction tended to result in an ideologically, theory-laden, or aesthetically compromised product. Thus, the more autonomous, or the more philosophically and socially truthful a work, the more it failed to conceal its inherent tensions or contradictions behind the illusion of perfect order, the more it refused not to show the untruth of its times. The failure and refusal prompted Adorno to speak of a negative autonomy or of a negative dialectics. Following an old Platonic anxiety, art had the ability to expose the lie of appearance or the untruth of society at the same time that it was able to serve as the primary means (of appearance) by which to encourage and sustain the lie. Its double-sided character and dependence on appearance rendered it exemplary both as a means and as an object of critique.

For Adorno, artworks were social formations set at an aesthetic remove; as such they exhibited a drive toward order, harmony, and internal coherence. This drive was dominant in the very concept of a work, a concept coincident with the dialectical course of enlightenment. And precisely what this drive aimed to do was suppress its opposing drive, the drive that would itself attempt to flout the conditions or possibility of order in a work by mimetically conveying as residue the non-expressed expression implicit to the concept of art. Just as the one drive toward order couldn't do without the drive toward free expression, so, under the condition of modernity, the concept of a work couldn't do without the concept of art, despite the antagonism they displayed toward one another. Yet in this antagonism resided all that was most productive and exemplary in the world of art. Hence, the more autonomous a work, the more the work exhibited the mimetic tension between silence and expression, between what it brought to expression under the concept of the work and what was concealed or excluded of the concept of art thereby. That Adorno often pursued an analogy between the artwork and the person was not without relevance for the truth art could indirectly reveal about society as a whole. The greater society's untruth, the more reified or fetishized the work's or the person's relation to society. The greater society's untruth the more the work was inclined to show the achievement of workhood as consumer product. The work, like a person, could show the achievement in two ways, either by adapting to or by resisting the social situation.

AFTER CATASTROPHE

When Adorno returned to Germany in 1949 he was confronted with the fact of having survived the catastrophe. He asked what it meant for (West) Germany to become a democracy given what he understood to be a continuation of social injustice and prejudice. He used his experiences in America partially as a model of both the promise and the curse of democracy. While convinced that neither the philosopher nor the artist could assume an ahistorical vantage point from which to view society, Adorno was nonetheless convinced that by describing the dominant tendencies toward philosophical, social, and aesthetic untruth, one would thereby show by dialectical negation what remained as the residue or remainder of truth. With Walter Benjamin, he did not think that truth could be found or established in a sustained method of philosophical argument; he rather looked in the cracks of such arguments, in what was not said, in what had historically come to be concealed by dominant patterns, be they philosophical theories, social formations, or artistic movements.

After the war, Adorno wrote that "to still write a poem after Auschwitz is barbaric," a claim he later somewhat modified (1992, vol. 2, p. 87). However, in the claim he asked a question of despair, whether and how continuation in art or thought was possible in a society that now lived "metaphysically"—as he used that term in conclud-

ing his *Negative Dialectic*,—under the condition of death. His *Aesthetic Theory* had, however, opened with the same claim, that it "is self-evident that nothing concerning art is self-evident any more, not its inner life, not its relation to the world, not even its right to exist" (1997, p. 1). Here, the point was to use the concept of self-evidence to begin a critique of its social, philosophical and aesthetic forms, where self-evidence found its subjective side in the formation of public opinion and its objective side in the production of ordered-appearances (say, in works of art). His preoccupation with how art and philosophy could continue in modern times had begun around 1930 when he asked after their "*actuality*." Later, he posed the question again but now even more concretely against the background of the compromise the university and the concert hall had made under national socialism.

Adorno experimented with the essay form, as is shown in his exemplary essay in his *Notes to Literature* on the essay as form. He wrote his aesthetic theory conscious of aesthetic figuration, sometimes in aphorisms or fragments, sometimes in figures of montage, even if this text often reads as a single paragraph without end. He wrote in such a way as to show his interest both in the techniques of high modernism and in the use and mutilation of language (his own use included), be that language one of communication, speech, gesture, or expression. He often expressed his thoughts as catch-phrases articulated as statements of a negative dialectic: for example, only for the sake of happiness and beauty are happiness and beauty renounced; only in memory and longing is pleasure now possible in art; the old only has refuge in the new; dissonance is the truth about harmony. Adorno was an aesthetic thinker of exemplary modernist form; he mediated that thinking within a dialectical and materialist history of society.

See also Aesthetics, History of; Aesthetics, Problems of; Beauty; Benjamin, Walter; Bloch, Ernst; Critical Theory; Dialectical Materialism; Enlightenment; Goethe, Johann Wolfgang von; Habermas, Jürgen; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Horkheimer, Max; Husserl, Edmund; Kant, Immanuel; Kierkegaard, Søren Aabye; Lukács, Georg; Nietzsche, Friedrich; Popper, Karl Raimund; Proust, Marcel; Schiller, Friedrich; Schopenhauer, Arthur.

Bibliography

WORKS BY THEODOR W. ADORNO

Composing for the Films. With Hanns Eisler. New York: Oxford University Press, 1947.

Gesammelte Schriften, edited by Rolf Tiedemann. 20 vols. Frankfurt am Main: Suhrkamp, 1970–1986.

Negative Dialectics. Translated by E. B. Ashton. New York: Seabury Press, 1973.

Minima Moralia: Reflections from Damaged Life. Translated by E. F. N. Jephcott. London: Verso, 1974.

Introduction to the Sociology of Music. Translated by E. B. Ashton. New York: Seabury Press, 1976.

"The Actuality of Philosophy." *Telos* 31 (Spring 1977): 120–133.

In Search of Wagner. Translated by Rodney Livingstone. London: New Left Books, 1981.

Positivist Dispute in German Sociology. Aldershot, Hampshire: Ashgate Press, 1981.

Prisms. Translated by Samuel and Shierry Weber. Cambridge, MA: MIT Press, 1981.

Kierkegaard: Construction of the Aesthetic. Translated by Robert Hullot-Kentor. Minneapolis: University of Minnesota Press, 1989.

Alban Berg: Master of the Smallest Link. Translated by Juliane Brand and Christopher Hailey. Cambridge, U.K.: Cambridge University Press, 1991.

The Culture Industry: Selected Essays on Mass Culture, edited by J. M. Bernstein. London: Routledge, 1991.

Notes to Literature. Translated by Shierry Weber Nicholsen. 2 vols. New York: Columbia University Press, 1991–1992.

Mahler: A Musical Physiognomy. Translated by Edmund Jephcott. Chicago: University of Chicago Press, 1992.

Nachgelassene Schriften, edited by Rolf Tiedemann et al. Frankfurt am Main: Suhrkamp, 1993–.

Aesthetic Theory. Edited by Gretel Adorno and Rolf Tiedemann. Translated by Robert Hullot-Kentor. Minneapolis: University of Minnesota Press, 1997.

Beethoven: The Philosophy of Music, edited by Rolf Tiedemann. Translated by Edmund Jephcott. Stanford, CA: Stanford University Press, 1998.

Sound Figures. Translated by Rodney Livingstone. Stanford, CA: Stanford University Press, 1999.

The Dialectic of Enlightenment. With Max Horkheimer, edited by Gunzelin Schmid Noerr. Translated by Edmund Jephcott. Stanford, CA: Stanford University Press, 2002.

Critical Models: Interventions and Catchwords. Translated and with a preface by Henry W. Pickford, Introduction by Lydia Goehr, New York: Columbia University Press, 2005.

Philosophy of New Music. Translated and with an introduction by Robert Hullot-Kentor. Minneapolis: University or Minnesota Press, 2006.

WORKS ABOUT THEODOR W. ADORNO

Arato, A., and E. Gebhardt. *The Essential Frankfurt School Reader*. New York: Continuum, 1982.

Benjamin, Andrew, ed. *The Problems of Modernity: Adorno and Benjamin*. London and New York: Routledge, 1989.

Bernstein, J. M., ed. *The Culture Industry*. London: Routledge, 2001.

Bernstein, J. M. *The Fate of Art: Aesthetic Alienation from Kant to Derrida and Adorno*. University Park: Pennsylvania State University Press, 1992.

Buck-Morss, Susan. *The Origin of Negative Dialectics: Theodor W. Adorno, Walter Benjamin, and the Frankfurt Institute*. New York: Free Press, 1977.

de Nora, Tia. *After Adorno. Rethinking Music Sociology*. Cambridge, U.K.: Cambridge University Press, 2003.
Hohendahl, Peter Uwe. *Prismatic Thought: Theodor W. Adorno*. Lincoln and London: University of Nebraska Press, 1995.
Huhn, T., ed. *The Cambridge Companion to Adorno*, Cambridge, U.K.: Cambridge University Press, 2004.
Huhn, T., and L. Zuidervaart, eds. *The Semblance of Subjectivity: Essays in Adorno's Aesthetic Theory*. Cambridge, MA: MIT Press, 1997.
Jameson, Fredric. *Late Marxism: Adorno, or the Persistence of the Dialectic*. London and New York: Verso, 1990.
Jarvis, Simon. *Adorno: A Critical Introduction*. Cambridge, U.K.: Polity Press, 1998.
Jay, Martin. *Adorno*. London: Fontana, 1984.
Leppert, Richard, ed. *Essays on Music: Theodor W. Adorno*. Selected, with introduction, commentary, and notes by Richard Leppert. Translated by Susan H. Gillespie. Berkeley: University of California Press, 2002.
O'Connor, B. ed. *The Adorno Reader*. Oxford: Blackwell, 2000.
Paddison, Max. *Adorno, Modernism and Mass Culture. Essays in Critical Theory and Music*. London: Kahn and Averill, 1996.
Paddison, Max. *Adorno's Aesthetics of Music*. Cambridge, U.K.: Cambridge University Press, 1997.
Pensky, Max, ed. *The Actuality of Adorno: Critical Essays on Adorno and the Postmodern* New York: SUNY Press, 1997.
Reijen, Willem van. *Adorno: An Introduction*. Philadelphia: Pennbridge Books, 1992.
Rose, Gillian. *The Melancholy Science: An Introduction to the Thought of Theodor W. Adorno*. London: Macmillan, 1978.
Weber Nicholsen, Shierry. *Exact Imagination, Late Idyll: Adorno's Aesthetics*. Cambridge, MA: MIT Press, 1997.
Zuidervaart, Lambert. *Adorno's Aesthetic Theory: The Redemption of Illusion*. Cambridge, MA: MIT Press, 1991.

Lydia Goehr (2005)

AEGIDIUS COLONNA ROMANUS

See *Giles of Rome*

AENESIDEMUS

(1st century BCE)

Very little is known about Aenesidemus's life. He was associated with the Athenian Academy around the time of its collapse in 87 BCE; and he was party to the dispute between Philo of Larissa, who advocated a mild form of skepticism in the form of an externalist, coherentist epistemology, and Antiochus of Ascalon, whose epistemology was basically that of Stoic foundationalism. The Academy had been for two centuries the home of epistemological skepticism, directed largely against the optimistic epistemology of the Stoics, who posited "apprehensive impressions" (*phantasiai katalêptikai*), which carried their own guarantee of truth. Aenesidemus saw Philo and Antiochus as betraying that heritage, as "Stoics fighting with Stoics" (Photius, *Library Catalogue* 212), and resolved to "philosophize after the fashion of Pyrrho."

Aenesidemus wrote eight books of *Pyrrhonian Discourses*, which Photius summarized: "the whole aim of the book is to ground the view that there is no ground for apprehension, whether through perception or thought." The main burden of the *Discourses*, Photius says, is to establish that nobody really grasps anything. However, only Pyrrhonian skeptics are aware of this ignorance, while everyone else falsely considers themselves to be in possession of secure knowledge. This false conviction, and the inevitable disputes that follow from the evident fact that different people hold different and incompatible beliefs, leads the Dogmatists ("belief-holders," as skeptics styled their opponents) into "ceaseless torments." Skeptics, having no beliefs, avoid these torments; indeed they "are happy … in the wisdom of knowing that they have firm apprehension of nothing." "Apprehension" (*katalêpsis*) is the Stoic technical term for sure and unshakable knowledge based on apprehensive impressions. When Aenesidemus claims that Pyrrhonists have no apprehension of anything, he is careful not to say that they have *apprehension* of that second-order fact. Yet they may still be aware of it, since it is evident to them introspectively that they are not certain of anything (thus skeptics seek to avoid the charge that their position is self-refuting).

Moreover, "even in regard to what he knows [this is Photius's language; and he may well be less careful than Aenesidemus in avoiding apparent self-refutation], he takes care to assent no more to its affirmation than to its denial." "Assent" (*sunkatathesis*) is another Stoic term, denoting unwavering commitment to the truth of some proposition (positive or negative); and no skeptic will claim that sort of cognitive security, even in regard to his own claims: a skeptic's "positions" (insofar as he really has any) are invariably provisional. In the same vein, "no more" (*ou mallon*) is a skeptical slogan: things may appear to be thus and so, but in themselves they are no more one way rather than the other. Diogenes Laertius (DL 9.106) reports Aenesidemus as saying that appearances are the criterion for action; thus he seeks to evade the common charge brought against skeptics (most famously by Hume) that their refusal to hold beliefs renders life impossible (it is a further, difficult question how far this notion of appearance can really be divorced from some concept of belief).

In the first *Pyrrhonian Discourse*, according to Photius, Aenesidemus distanced himself from the Academics, since they "posit some things with confidence and deny others unambiguously, while Pyrrhonists are aporetic and devoid of dogma; they say neither that all things are inapprehensible, nor that they are apprehensible, but that they are no more so than not so, or sometimes so and sometimes not so, or so for one person but not for another." The Academics are negative dogmatists, positively affirming that nothing can be apprehended according to the Stoic criterion; Pyrrhonists, by contrast, will say that they do not seem to apprehend anything, but will not reject the possibility of there being apprehension. Crucially, "the Pyrrhonist determines absolutely nothing, not even this very proposition, that nothing is determined." That this is the authentic skeptical attitude is confirmed by Sextus Empiricus, *Outlines of Pyrrhonism* (PH) 1.187–209; and Sextus probably relies heavily upon Aenesidemus in that work.

The second *Pyrrhonian Discourse* casts doubt upon "truth, causes, effects, motion, generation and destruction," while the third "was also about motion and sense perception … working carefully through a similar set of contradictions, he puts them too beyond our grasp." These arguments about perception no doubt included the material of the so-called "Ten Modes of Aenesidemus," arguments designed to undermine the Dogmatists' truth-claims, and hence to induce *epochê*, or suspension of judgment, "which the skeptics say is the goal (*telos*), upon which tranquility follows like a shadow, according to Aenesidemus and Timon" (DL 9.107; cf. PH 1.25–30). Thus "Pyrrhonian discourse is a kind of recollection of appearances … , on the basis of which they are all brought into confrontation with one another, and when compared are found to cause much disparity and confusion; so says Aenesidemus in the summary of his *Pyrrhonics*" (DL 9.78).

The Ten Modes are attributed to Aenesidemus by Sextus (*Against the Professors* [M] 7.345); Aristocles ascribes nine Modes to him, and we know the number of the Modes to have been fluid (our earliest source, Philo of Alexandria, records only eight). Neither Sextus in his extant treatment of the Modes (PH 1.31–163), nor Diogenes in his shorter summary (DL 9.79–88) father them on Aenesidemus; but it is still likely that he was responsible for this organization of earlier skeptical material. The Modes share a common form, involving conflicting appearances: *x* appears *F* in conditions *C*, or to observer *O*, not-*F* in conditions *C**, or to observer *O**; there is no non-question-begging way of privileging either of *C* or *C**, *O* or *O**; so we should suspend judgment as to whether *x* is *F*. The Modes are differentiated by different fillers for *C* or *O*; thus the first (in Sextus's ordering) compares the different sensory representations of different animals, the second collects cases of dissonant judgment between different humans, the third conflicts in the deliverances of different sense-modalities, and the fourth includes discrepant reports from the same sense at different times. Other Modes collect cases of ethical or social discrepancy (the tenth), and point to the ways in which differing conditions of the perceiver may affect what they seem to perceive.

The upshot is that we cannot in any case say how things really are, but only how they seem in particular circumstances. Things are judged relatively to the perceiver and their circumstances. Sextus is careful not to draw relativistic conclusions (although the facts of relativity figure both as a particular Mode, the eighth, and in general in the articulation of all the Modes): He does not positively *assert* that things are for the observer as they appear. By contrast, Aenesidemus, judging from Photius's summary, is quite happy to accept the relative judgments as such, since they do not (cannot) count as Dogmatic.

In the fourth *Discourse*, Aenesidemus discussed signs. Sign-theory and its associated epistemology was of overwhelming importance in post-Aristotelian philosophy. The Stoics (along with various Dogmatic medical schools) held that it was possible to infer directly from the phenomena to the underlying structural conditions responsible for them. Skeptics (and Empiricist doctors) denied the validity of such inferences, allowing only that memories of past conjunctions of phenomena might allow us to expect (although fallibly) similar conjunctions in the future. Aenesidemus advanced the following paradigmatically skeptical argument: If apparent things appear alike to all in a similar condition, then signs should appear alike to all in a similar condition; but they do not; hence signs are not apparent (M 8.215). That is, it is not unequivocal what they are signs of—different doctors, for example, draw radically different conclusions from the same symptoms (M 8.219–220).

In the fifth *Discourse* Aenesidemus turned to causes; again Sextus retails some of his arguments (M 8.218–226)yes; crucial to them is the idea that a cause should operate from its own resources; but if it does, then, since it requires nothing else in order to exercise its causal power, it should do so invariably and continuously. More impressive are the Eight Modes against the Aetiologists, mentioned in Photius and ascribed to Aenesidemus by Sextus at PH 1.180–185. These are eight general argu-

ments against the possibility of inferring from evident phenomena to the hidden structures of things that are supposedly causally responsible for those phenomena, in the manner of Dogmatist philosophers and scientists (notably Epicureans, but also Peripatetics and Stoics). Aenesidemus's basic claim foreshadows the modern maxim that theories are invariably underdetermined by the available data. No amount of evidence can ever entail that any particular theory must be true: There are always many ways in principle of accounting for the same set of phenomena (1.181–182). Moreover (and here Aenesidemus turns from general methodological issues to castigating particular recurrent theoretical foibles), theorists sometimes offer piecemeal, unrelated explanations for what are evidently related sets of phenomena; and they tend to suppose, without justification, that the structure of the hidden, subperceptual realm will mirror in all important respects that of the phenomenal world (1.182; this point is particularly well-taken against Epicurean physics).

Furthermore, Aenesidemus notes (and this too is a staple of contemporary philosophy of science) that researchers are inclined to favor explanations that concur with their own prejudices (1.183), and indeed on occasion to prefer explanations that not only conflict with the facts, but also with their own theories (1.184). Finally, he notes that Dogmatists "frequently … seek to explain doubtful things on the basis of things equally doubtful" (1.184). Taken together, the eight Modes are an impressive attack on the possibility of arriving at any soundly based understanding of the hidden natures of things. As such, they are obviously of a piece with, and complement, the rest of Aenesidemus's skeptical argumentation. The last three *Pyrrhonian Discourses* dealt with ethical issues, with Aenesidemus arguing that the lack of philosophical agreement regarding good and bad, choice and avoidance, virtues, and finally the end, preclude the possibility of arriving at any secure judgments about them.

All of the evidence so far reviewed makes Aenesidemus a consistent and powerful skeptic. However, a number of passages in Sextus portray him in a much more Dogmatic light, as holding various views about the intellect (M 7.350), and endorsing the view that there are two types of change (M 10.38). Elsewhere he is said to be in agreement with Heraclitus, whom Sextus explicitly describes as a Dogmatist. These discrepancies are too widespread simply to be brushed aside. But there is as yet no scholarly agreement as to what to do about them.

See also Ancient Skepticism; Antiochus of Ascalon; Philo of Larissa; Pyrrho; Sextus Empiricus.

Bibliography

Annas, J., and J. Barnes. *The Modes of Scepticism*. Cambridge, U.K.: Cambridge University Press, 1985.

Hankinson, R. J. *The Skeptics*. London: Routledge, 1995.

Long, A. A, and D. N. Sedley. *The Hellenistic Philosophers*. Cambridge, U.K.: Cambridge University Press, 1987.

Rist, J. "The Heracliteanism of Aenesidemus." *Phoenix* 24 1970: 309–319.

Tarrant, H. *Skepticism or Platonism?* Cambridge, U.K.: Cambridge University Press, 1985.

R. J. Hankinson (2005)

AESTHETIC EXPERIENCE

An aesthetic experience arises in response to works of art or other aesthetic objects. Although the term *aesthetic* itself was not introduced until the eighteenth century, it is clear that what are identified in contemporary discussions as "aesthetic experiences" were "felt" by individuals long before this: for example, when Plato worried about excessively emotional reactions to recitations of poetry or when Aristotle described the positive effects of attending the theater. Nevertheless, the exact nature of aesthetic experience—even the idea that there is such a unique form of experience—remains a matter of controversy.

WHAT AESTHETIC EXPERIENCES FEEL LIKE

One area of contention concerns what it feels like to have an aesthetic experience—that is, whether there is some special emotion or attitude or other internal sign that enables one to recognize that what one is having is an aesthetic experience and not some other kind. Immanuel Kant, one of the first philosophers to have addressed these kinds of questions, characterizes aesthetic experiences as those pleasures associated with occasions when one judges something to be beautiful. He asserts that one recognizes that this pleasure does not result from a realization that an object is useful or agreeable to one because of special things about oneself. Instead the pleasure arises simply because the form of the object is delightful and could and should be enjoyed by anyone. Kant makes a sharp distinction between responding positively in this manner and responding positively for moral or scientific reasons. Although several theorists have disagreed with Kant's argument, most theorists agree that aesthetic experiences are identified as such at least partly because of an emotional involvement of the experiencer. One feels good (or bad) when one responds aesthetically to a beautiful

sunset or elegant poem (or to a messy waste dump or plodding verse).

But it is more than just a feeling of pleasure (or pain) that characterizes aesthetic experiences, according to many theorists. John Dewey (1958), for example, argues that aesthetic experiences are the most complete, the richest, and the highest experiences possible. One is actively engaged and conscious of the world's effect on one but at the same time appreciative of one's possibilities for acting on the world. One senses an organization, coherence, and satisfaction as well as an integration of the past, present, and future that ordinary nonaesthetic experiences lack.

More recently, Nelson Goodman (1976) has warned that too much emphasis on the pleasurable aspects of aesthetic experiences deprives them of much of their importance. What he derisively calls "tingle-immersion" theories overlook the crucial role of intellect, he cautions. In aesthetic experiences, the emotions function cognitively, he says; one "feels" a heightened operation of both cognition and emotion operating together.

WHAT AESTHETIC EXPERIENCES FOCUS ON

Another area of debate is the object of aesthetic experience. Many philosophers have insisted that the pleasurable (or painful) responses associated with an aesthetic experience must be connected with something special about some objects and events—properties that nonaesthetic or nonartistic objects and events lack—for clearly we do not have aesthetic experiences with regard to just any old thing.

Aristotle believed that the pleasure unique to dramatic tragedies consisted in a catharsis of the painful emotions of pity and fear and that this could occur only if a play had certain properties—the right sort of plot and characters. Kant, we saw above, thought that aesthetic experiences were pleasant when objects were such that mere apprehension of their form alone evoked delight. In general, theorists and critics described as "formalists" insist that in an aesthetic experience attention is directed solely to immediately perceivable properties of objects and events—shape, colors, tones, sounds, and patterns. Monroe Beardsley (1958), for instance, characterizes the focus of aesthetic experiences as formal unity and the intensity of regional quality. Clive Bell (1914) claims that emotional responses to objects exhibiting "significant form" can be so intense that one does not care at all about the content of some artworks; what matters is always form and not content. Jerome Stolnitz (1960) argues that one takes up a special attitude, distinterestedness, when one has an aesthetic experience. Ordinary everyday concerns or purposes are put aside, and one focuses on the form of an object for its sake alone, he believed.

An increasing number of theorists disagree with the formalist position that when one has an aesthetic experience one focuses solely on an object's formal properties and that one's scientific, moral, religious and other beliefs or concerns are put aside. For one thing, some insist, the expression of certain ideas plays a key role in some works of art, and surely thinking about these ideas (content) is an appropriate and important aspect of the aesthetic experiences of them. Even if focus on form is necessary to aesthetic experiences, it may be that content and context are also legitimate matters for aesthetic attention.

WHAT HAVING AN AESTHETIC EXPERIENCE REQUIRES

Even if one grants that aesthetic experiences arise only in the presence of objects that exhibit a form that pleases, many theorists have insisted that more than a formally pleasing object and passive viewer are required. Just as not every object gives rise to an aesthetic experience, so not all individuals have aesthetic experiences in reaction to the same objects. David Hume (1987) in the eighteenth century and, more recently, Frank Sibley (1959) in the twentieth, have insisted that only persons who have taste or special sensitivities are capable of responding aesthetically. Not all people are equally competent judges, Hume claims. Only people who are sensitive, attentive, open-minded, perceptive, clear-headed, trained, and experienced can tell a good poem from a bad poem. In the absence of sensitivity, one will be left completely cold by objects that enthrall a more acute and receptive observer.

Formalists, we saw above, insist that aesthetic experience requires an appropriate amount of distance—one must put aside beliefs or purposes and give oneself up entirely to the object. But others argue that precisely the opposite is the case. Contextualists insist that, before one can have an aesthetic response (or at least an appropriate or full one), one's intellect and moral beliefs must be engaged. Noel Carroll (2000), for example, argues that moral concerns may block or enhance aesthetic experiences. Kendall Walton (1970) asserts that one cannot interpret and otherwise respond to a work of art unless one is versed in the genre it represents. One cannot judge whether a sonnet is good or bad unless one knows that it is in fact a sonnet and not a haiku, for example. Allen Carlson (2000) points out that an aesthetic appreciation of nature requires an awareness that what one is appreci-

ating is nature (not a painted landscape, for instance). This in turn demands an understanding of how nature works. The person who brings a fair degree of scientific knowledge to a particular environmental system will have a much fuller, richer aesthetic experience of that environment. What is required by or, at the very least, relevant to aesthetic experience may be whatever directs one's attention as fully as possible to the potentially pleasurable formal properties of an object or event.

WHERE OR WHEN AESTHETIC EXPERIENCES OCCUR

The nature of aesthetic experience may not be fully accounted for even if one knows everything important about objects that occasion them—the context or circumstances attending an individual's response may prove critical. Some philosophers call attention to the viewing conditions: for example, whether a concert is live or recorded or whether a poem is read to oneself or recited aloud. Others focus on the political, economic, or social conditions of an experience. To what extent are aesthetic experiences socially constructed? Is responding pleasurably to the color of a flower, for instance, "natural" (in the way that hunger or sexual arousal is), is it taught (in the way that acquired tastes are), or is there some mix of innate and learned response? Herein lies another set of issues that philosophers and others (for example psychologists, sociologists, and economists) debate.

AESTHETIC VERSUS ARTISTIC EXPERIENCE

Art objects are examples of aesthetic objects. But not all aesthetic objects are artworks—for example, sunsets or mountain vistas. Whether there is a difference between aesthetic experience and artistic experience is still another question that theorists address. Kant notes that in appreciating art objects one is aware of the fact that a human created it (and, in the case of great Art, that someone of genius was responsible for it). Thus artistic experiences lack the "purity" associated with those disinterested pleasures that arise from form alone.

Arthur Danto (1986) has argued that developments in the history of Art (such as the appearance of rather odd artifacts in museums) mean that one cannot tell if something is a work of art or not in the absence of a theory of art. This is not the case for aesthetic objects, it would seem. One does not need a theory of the aesthetic in order to have an aesthetic response, for one can have such a response to anything at all. It may be that some experiences of art are not aesthetic at all. If one is primarily concerned with the history of an object or its economic or religious value, then one may not care about or may even completely ignore the formal properties of that object.

THE NEED FOR THE CONCEPT OF AESTHETIC EXPERIENCE

Finally it must be pointed out that not everyone believes that it is possible or necessary to distinguish aesthetic from other kinds of experiences. The whole notion is too vague and abstract, some philosophers argue. Reporting that one has had an aesthetic experience is no more informative than claiming that one has had an "economic experience" or an "automotive experience," according to some. One describes one's experience far better by saying things like "I bought some junk bonds yesterday" or "I had an exciting ride in a Porsche this morning" than by saying "I had an economic experience" or "I had an automotive experience." Similarly, one might do away completely with talk about aesthetic experiences and rely instead on discussions of reading particular poems or listening to pieces of music or birdsongs or looking at specific paintings or landscapes or drinking particular wines.

Nevertheless, people do talk about aesthetic experiences, and there might be good reason to try to articulate what they involve. If one goal of education is to improve the quality of life through aesthetic experiences, then it will be important to determine what such experiences feel like, focus on, and require. Moreover, if one fears that significant properties of objects or events will be overlooked if one confuses moral or scientific perspectives with aesthetic ones, then it may be necessary to distinguish the last from the former two.

See also Aesthetic Judgment; Aesthetic Qualities; Art, Interpretation of.

Bibliography

Beardsley, Monroe. *Aesthetics*. New York: Harcourt Brace, 1958.

Bell, Clive. *Art*. London: Chatto & Windus, 1914.

Carlson, Allen. *Aesthetics and the Environment*. London and New York: Routledge, 2000.

Carroll, Noel. *Theories of Art Today*. Madison: University of Wisconsin Press, 2000.

Danto, Arthur. *The Transformation of the Commonplace*. New York: Columbia University Press, 1986.

Dewey, John. *Art as Experience*. New York: Putnam's Sons, 1958.

Eaton, Marcia Muelder. *Merit, Aesthetic and Ethical*. New York: Oxford University Press, 2000.

Goodman, Nelson. *Languages of Art*. Indianapolis, IN: Hackett, 1976.

Hume, David. "Of a Standard of Taste" (1759). In *Essays Moral, Political and Literary*, edited by Eugene F. Miller. Indianapolis, IN: Liberty Fund, 1987.

Kant, Immanuel, *The Critique of Judgment* (1790). Translated by Werner S. Pluhar. Indianapolis, IN: Hackett, 1987.

Sibley, Frank. "Aesthetic Concepts." *Philosophical Review* 68 (1959): 421–450.

Stolnitz, Jerome. *Aesthetics and the Philosophy of Art Criticism*. New York: Houghton Mifflin, 1960.

Walton, Kendall. "Categories of Art." *Philosophical Review* 79 (1970): 334–367.

Marcia Muelder Eaton (2005)

AESTHETIC JUDGMENT

In recent analytic aesthetics, there have been two prominent questions about aesthetic judgments. One is how to distinguish aesthetic judgments from other judgments. Answering this question seems particularly urgent when an aesthetic judgment and a nonaesthetic judgment about the same object are incongruent. In such a case it seems that an object might be judged to have aesthetic value but also to be negatively judged, say ethically or in terms of its practical use. A corollary question is whether the negative value of a nonaesthetic judgment should affect the allegedly purely aesthetic judgment.

The other prominent question, a question present at least since the eighteenth century, is actually two questions: first, whether aesthetic judgments are objective or subjective, and second, whether aesthetic judgments can be verified or otherwise substantiated. Somewhat curiously, perhaps, some philosophers have thought that even though such judgments are subjective, they are still capable of being supported. David Hume is an example. In contrast, other philosophers have thought that even though such judgments are genuinely objective, they are nonetheless incapable of being verified by customary procedures. Frank Sibley has been the leading exponent of this opinion. A more obvious thesis is Immanuel Kant's, namely that aesthetic judgments are both subjective and impossible to support by any interpersonal means.

Hume (1987) believed that it is possible to identify certain judges as having especially reliable taste and then to take their subjective responses to objects as a standard in evaluating the objects. When such judges deliver what Hume called "a joint verdict," meaning, presumably, that they concur in taking pleasure in an object, taking pleasure in the object is then established as correct, in a sense, with at least customary probability, and any judge who fails to realize this pleasure is defective in his taste.

Kant, in contrast, thought that no corroboration of one's judgment is possible because a concurrence with or difference from the responses of other judges is logically irrelevant.

The idea of something explicitly called an aesthetic judgment seems first to have appeared in the eighteenth century and was formulated in detail by Kant (2000). By "aesthetic judgment" Kant meant a judgment based on a feeling. He was especially concerned to describe those feeling-based judgments in which an object is found beautiful, and then to show that we are entitled to make such judgments despite being unable to verify them. In his conviction that these judgments are essentially subjective (that is, derived from or based on the subject's feeling), Kant is in line with an earlier tradition. The most notable exponent of this tradition was Hume, though it remains unsettled just how much, if any, of Hume's writings on this topic were known to Kant. Yet Kant probably did know the earlier work of Francis Hutcheson, work in the spirit of Hume even if less compelling philosophically. In later developments of the idea of an aesthetic judgment, however, this feeling-based subjectivity has been less important than Kant's description of how an aesthetic judge attends to the object of his judgment.

The subjective character of judgments of beauty seemed obvious in the eighteenth century, especially to Hume and Kant, so obvious that neither of them argued for this notion but simply assumed it. Indeed, the etymology of the word "aesthetic" indicates that an aesthetic judgment must be essentially related to a feeling. The Greek term refers to sense perception, usually, but it has now come to refer to feelings in general, and in particular to feelings of pleasure. Hume does not use the term "aesthetic," and he speaks only of the exercise of taste in the discernment of beauty, but like Kant he takes it for granted that all judgments of beauty arise from feelings of pleasure experienced by the judge.

According to Hume, the term "beauty" does not correspond to any objective property of things, and so judgments of beauty cannot be correct or incorrect in any straightforward manner. Yet such judgments can be vindicated, he thought, by agreement with the judgments of especially well suited judges of the object. These exemplars of taste (whose responses, he said, constitute a "standard of taste") are identified by their stellar discernment, without prejudice, of all the properties of the objects being judged. There is no way to inspect an object for its beauty, Hume thought, because "beauty" does not

mark any property of an object, but it is possible, as a matter of empirical investigation, to determine whether any particular judge is an exemplary judge.

Kant, in describing what he calls "a pure judgment of taste," had a different idea. He thought that the judge must pay no attention to any use to which the object might be put, to any concept that applies to the object, or to any interest that the judge might have in the object. The judgment must thus be entirely disinterested and free of any thought that relates the object to anything else. It is a judgment about the object purely and simply in itself.

Kant first described aesthetic judgments made about natural objects (his leading example being a beautiful rose), and then extended such judgments to works of art. He thus effectively regarded successful works of art (which for him meant artificial beautiful objects) as loci for such judgments.

The idea that aesthetic judgment requires a detached state of mind has sometimes been developed as the idea that aesthetic judgments require an aesthetic attitude, a distinct mode of addressing objects. An early exponent of this idea was Arthur Schopenhauer, although he does not use the term "aesthetic attitude." Pursuing a line different from Kant's, Schopenhauer thought that contemplation of works of art was an activity in which one could escape the usual constraints on one's will.

In the early twentieth century, the idea of an aesthetic attitude was developed further, given this particular name, and given more detailed treatment, though it eventually became a problematic notion. An early formulation is Edward Bullough's (1957), although his interests were somewhat more psychological than philosophical. A later, more sophisticated treatment is to be found in the works of Jerome Stolnitz (1978). A useful canvass of the idea is in George Dickie's "The Myth of the Aesthetic Attitude" (1964), where Dickie seeks to do away with the idea.

Although continuing conceptions of aesthetic judgment in many respects derive from the early work of Hume and Kant, these conceptions have taken at least two noteworthy turns. In philosophy at the beginning of the twenty-first century, the term "aesthetics" has become a virtual synonym for "philosophy of art." This assimilation sometimes draws attention to a question, but at other times tends to cover it up—the question of which is basic, the idea of art or the idea of the aesthetic. In Kant and many of his followers, the idea of the aesthetic is basic, and the idea of art is, so to speak, constructed out of the idea of the aesthetic. Kant thus first characterizes aesthetic judgments and then essentially describes works of fine art as objects about which such judgments can be made. Richard Wollheim (1980), in contrast, reverses this dependence, declaring that to make an aesthetic judgment is to regard something as a work of art.

A radically different thesis is that of Frank Sibley (1959, 1965). Sibley takes aesthetic judgments to be judgments that apply aesthetic concepts to objects through the use of aesthetic terms. Rather than understand taste as Hume and Kant did, as the ability to take pleasure in the judgment of objects, Sibley takes taste to be the ability to use aesthetic terms and concepts. Furthermore, in view of his conviction that aesthetic judgments are objective, Sibley treats the term "beautiful" quite differently from his eighteenth-century predecessors. For Hume and Kant, the term "beauty" has very little semantic content, it indicating only that the object produces a particular feeling of pleasure in the judge. Sibley, in contrast, insists that the term refers to a property of the object being judged. Thus, for Sibley, "beautiful," "elegant," "graceful," and other terms indicated mainly by example are all aesthetic terms, and as such they all refer to objective properties, although only judges exercising what Sibley calls "taste" can detect these properties and hence correctly apply the terms. Thus, quite apart from the tradition of Hume and Kant, Sibley's thesis is that aesthetic judgments are perfectly objective, meaning that their terms refer to properties objectively present in the objects being judged. Yet Sibley's thesis, at least in one respect, is more like Hume's and Kant's than it is like Wollheim's. For Wollheim, to regard an object aesthetically is to regard it as a work of art. For Hume, Kant, and Sibley, aesthetic judgments are freely made of works of art but also of other objects, and in the latter case there is no need to treat these objects as works of art.

Even among those who regard the concept of art as more basic than the concept of the aesthetic, many such thinkers continue to insist, with Kant, that an aesthetic judgment must be disinterested and must not attend to anything besides the object itself. Those who believe aesthetic judgments to be a unique kind of judgment have been eager to distinguish aesthetic judgments from ethical judgments, in particular, and also from practical concerns. Others have wondered whether it is possible to make such a clear logical separation. When the question of design is raised, it becomes increasingly difficult to suppose that an aesthetic judgment about an object is entirely divorced from other considerations—an issue that is perhaps most acute in the case of architecture. If a building is beautiful to behold but ill suited to whatever activities it is meant to house, can one keep the building's

evident disutility from contaminating one's sense of the aesthetic value of the building? The same question arises, obviously, in many other cases of artistic design, ranging from automobiles to writing instruments to time-keeping devices. It seems clear that a genuinely ugly object might be a perfectly serviceable automobile or watch. It is less clear that that a poorly performing object can still be beautiful. On this matter, Kant's opinion is clear. He thought that it is one thing to judge a watch, say, to be a good watch because of its perspicuous time display and reliable time keeping, this being to judge the watch in terms relying on the concept of a watch; it is another thing to offer a pure judgment of taste. To other authors, this is not obvious, because for them, questions of utility are difficult to separate from questions of the aesthetic value of an object.

Recently much attention has been given to the separation of ethical concerns from aesthetic concerns (Levinson 2001), and in 2005 it is a much debated question whether the dubious moral character of an art work can be kept separate from its artistic or aesthetic value. There has thus been a renewal of interest in the question of the relations of ethics and aesthetics to one another.

See also Aesthetic Experience; Aesthetic Qualities; Aesthetics, History of; Art, Interpretation of; Beauty; Sublime, The; Ugliness.

Bibliography

Brady, Emily, and Jerrold Levinson, eds. *Aesthetic Concepts: Essays after Sibley*. New York: Oxford University Press, 2001.

Bullough, Edward. *Æsthetics: Lectures and Essays* Stanford, CA: Stanford University Press, 1957.

Dickie, George. "The Myth of the Aesthetic Attitude." *American Philosophical Quarterly* 1 (1) (1964): 56–65.

Guyer, Paul. "History of Modern Aesthetics." In *The Oxford Handbook of Aesthetics*, edited by Jerrold Levinson. New York: Oxford University Press, 2003.

Guyer, Paul. "The Origins of Modern Aesthetics: 1711–35." In *The Blackwell Guide to Aesthetics*, edited by Peter Kivy. Oxford: Blackwell, 2004.

Hume, David. *Essays, Moral, Political, and Literary*. Rev. ed. Indianapolis, IN: Liberty Fund, 1987.

Kant, Immanuel. *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. New York: Cambridge University Press, 2000.

Levinson, Jerrold. *Aesthetics and Ethics: Essays at the Intersection*. New York: Cambridge University Press, 2001.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne. Indian Hills, CO: Falcon's Wing Press, 1958.

Sibley, Frank. "Aesthetic and Non-aesthetic." *Philosophical Review* 74 (1965): 135–159.

Sibley, Frank. "Aesthetic Concepts." *Philosophical Review* 68 (1959): 421–450.

Stolnitz, Jerome. " "The 'Aesthetic Attitude'" in the Rise of Modern Aesthetics." *Journal of Aesthetics and Art Criticism* 36 (1978): 409–422.

Townsend, Dabney, ed. *Eighteenth-Century British Aesthetics*. Amityville, NY: Baywood, 1998.

Wollheim, Richard. *Art and Its Objects*. 2nd ed. New York: Cambridge University Press, 1980.

Ted Cohen (2005)

AESTHETIC QUALITIES

It is generally, although not universally, agreed among philosophers that there is an important distinction to be drawn between the aesthetic qualities of objects, especially art objects, and their nonaesthetic qualities: between being serene, stunning, or grating, and being square, in the key of A-minor, or weighing seven pounds. The concept of an aesthetic quality is a philosophical one, not in general use, but aestheticians appeal to it in clarifying the practice of art criticism, justifying aesthetic judgments, and evaluating artworks.

HISTORICAL BACKGROUND

Both David Hume (1963) and Immanuel Kant (1966) set the stage for this modern distinction in their discussions of aesthetic judgments, judgments regarding the beauty of objects. Both argued that such judgments differ in kind from judgments regarding ordinary perceptual properties. Both held that aesthetic judgments depend on subjective feelings of pleasure and affective responses, but both also sought a universal ground for such judgments. Unlike Francis Hutcheson (1971) before them, they did not find this ground in an objective property (for Hutcheson, unity in variety) that always gives rise to this pleasurable response in qualified observers. Instead, recognizing the normative force of ascriptions of beauty, the demand for agreement in one's ascriptions of this property, they sought a standard in universal subjective grounds of the judgments of qualified critics.

Hume emphasized that only the judgments of fully competent or ideal critics indicate the presence of beauty or aesthetic merit. The property of beauty is similar in this respect to secondary qualities like colors, as analyzed by John Locke. For Locke, the color red is a power in objects, based on objective properties of their surfaces, to cause red sensations in normal observers in normal conditions. For Hume, beauty is similarly a relation between various objective properties and subjective responses, dif-

ferences being that, as noted, there is no single objective property to be found here, and that qualified observers are rarer and more difficult to define. Such observers must have developed tastes, be knowledgeable of the type of work they are judging and of the historical tradition with which to compare the work, and be sensitive to the sorts of subtle relations on which the beauty of the work might depend. In the end, even such qualified critics might disagree in their comparative aesthetic judgments, Hume recognized.

Kant was both more emphatic than Hume that there are no universal objective grounds for ascriptions of beauty, and was more confident that such judgments should nevertheless be universally shared. For him, there are no principles that connect objective properties with correct ascriptions of beauty. Nevertheless, the pleasure derived from the disinterested perception of form should be universally felt, since common human faculties are involved in such perception. The perception of formal properties elicits a value-laden (pleasurable) response that is common to all disinterested observers and expressed in ascriptions of beauty. Since there is no objective property common to all beautiful objects (no objective concept of beauty), one cannot tell from a description of an object whether it is beautiful. One must experience the pleasure from perception of the object. But in judging an object to be beautiful, one demands the agreement of other observers, unlike in judging mere agreeableness.

THE NATURE OF AESTHETIC QUALITIES: REALISM

The contemporary discussion of aesthetic qualities began with Frank Sibley (1959). He first expanded the list of aesthetic qualities from beauty and sublimity to include emotion qualities like being sad or serene, evocative qualities like being powerful or dull, behavioral qualities like being jaunty or sluggish, formal-evaluative qualities like being graceful or tightly knit, and second-order perceptual qualities like being vivid or steely. A major philosophical question resulted from this expansion. What do these qualities have in common that distinguishes them from nonaesthetic qualities? Other questions remain from the discussions of Hume and Kant. What is the nature of these qualities, and how are they related to the nonaesthetic qualities of their objects?

In regard to the first question, some of the properties listed may be ascribed to artworks only metaphorically, but others are ascribed literally. If "sad" here can mean expressive of sadness, and "powerful" can refer to the power to evoke a strong response, then these two properties fall into the latter category.

According to Sibley, perceiving aesthetic properties requires taste. If taste is a special quasi-perceptual faculty different from the ordinary five senses, as his usage sometimes suggests, then its existence and operation becomes mysterious, as do the aesthetic qualities it alone can grasp. If taste refers simply to sensitivity to aesthetic properties, then there is a tight circularity in the definitions that needs to be removed. But appeal to taste here can have two other more plausible functions. First, it can indicate that the perception of all the relevant nonaesthetic properties of an object is not sufficient for the perception of its aesthetic properties. One must perceive nonaesthetic properties to perceive aesthetic qualities, but not vice versa.

Second, since "taste" in one of its senses refers to dispositions to evaluate in certain ways, appeal to taste here can indicate that ascribing aesthetic properties to artworks is always relevant to their evaluation. We justify aesthetic evaluations by pointing to the aesthetic properties of objects. Some of these properties, like being graceful or tightly knit, are typically value-laden in themselves. Others, like being sad, seem not to be. But if artworks not only have such properties, but, as Nelson Goodman (1969) claims, exemplify them, that is, refer to them and tell us something of their nature, then this is of some value. And experiencing such qualities can also be of value by being part of an overall response to an artwork that engages not only the emotions, but the perceptual, imaginative, and cognitive faculties as well.

Thus, we can define aesthetic qualities as those that contribute directly to an object's aesthetic value, positive or negative. Again, there is a circularity here, but it can be removed by defining aesthetic value without appealing to aesthetic qualities, perhaps in terms of the overall engagement of our mental faculties just alluded to. What has aesthetic value, according to this concept, simultaneously challenges and exercises all our mental capacities—perceptual, imaginative, affective, and cognitive. If the concept of art itself is in turn evaluative, if having aesthetic value in the sense indicated is both necessary and sufficient for being a (fine) artwork, then aesthetic qualities are also definitive of (fine) artworks. Taken in this sense, however, the concept of aesthetic properties has not only been broadened from the initial reference to beauty; it has also been narrowed to the domain of artworks, at least in its primary use.

In regard to the second question on the nature of aesthetic qualities, it is clear that they are relational prop-

erties, as Hume and Kant held, involving appreciative responses to the objective or base qualities of objects. These base qualities include structural properties of tones, shapes, and colors; syntactic and semantic properties of literary texts; and relations between these and similar properties in other works. Appeal to these base properties justifies ascriptions of aesthetic qualities, and appeal to these aesthetic qualities in turn justifies overall aesthetic evaluations.

That aesthetic qualities involve subjective responses does not imply that these qualities are not real. Real properties are those that are instantiated independently of observers' beliefs about them and of how they appear to particular observers. Secondary qualities like colors are real in this sense because, even though particular observers can disagree and even though colors can appear other than they are, normal observers in normal conditions can achieve consensus on colors. Such consensus among qualified observers is essential to the reality of such relational properties. A crucial question is whether we would find agreement in the ascription of aesthetic qualities among fully qualified art critics.

THE RELATION TO BASE PROPERTIES: RELATIVISM

Kant held that there are no principles linking objective properties to beauty, and Sibley held that nonaesthetic properties are never sufficient conditions for aesthetic properties. The lack of such principles is due to the fact that aesthetic qualities are not only relational, but relative in several different senses. First, they are relative to the contexts of the particular objects that instantiate them. A graceful passage in a Mozart piece would not be graceful at all in a piece by Charles Ives. Second, they are relative to differing interpretations of the same work. Iago's "Credo" aria in Giuseppe Verdi's *Otello* can be interpreted as boisterous and defiant or as sinister and brooding. Third, they are relative to historical context and change with changing historical contexts. The works of Antonio Salieri were heard as graceful before Mozart but as somewhat stilted and awkward after Mozart. Finally, as Hume in the end affirmed but Kant denied, they are relative to differing tastes of different critics. What is poignant to one is maudlin to another; what is striking and powerful to one is garish and grating to another.

That the latter disagreements occur at all levels of actual competence and sophistication indicates that even ideal critics would fail to reach consensus in ascribing aesthetic properties. For every such property, there would be some disagreements among fully qualified critics as to whether some objects had the property in question. And this would occur not only in borderline cases, indicating only vagueness in the concepts of such properties. A paradigm of poignancy for some critics, for example, a Tchaikovsky symphony or Puccini aria, is a paradigm of maudlin sentimentality for others.

It seems, therefore, that we must relativize ascriptions of aesthetic properties to both tastes and contexts (including work, historical, and interpretive contexts). The main problem with doing so is that it then becomes problematic to see opposed ascriptions as really in disagreement and difficult to explain why opposing critics argue for their interpretations and evaluations. Genuine disagreement and argument about the presence of an aesthetic property seem to assume a right answer to the question of whether or not the property is present. But if an artwork is powerful to one critic and not to another, then what are they disagreeing about? In short, the problem for the relativist is to account for the normative force of judgments regarding aesthetic qualities. Even if Kant was too strong in his claim that we demand universal agreement in our aesthetic judgments, surely the practice of critical argument reflects some demand for agreement.

To maintain a realist account of aesthetic qualities in the face of disagreement among fully qualified critics, one might say that an object really has an aesthetic quality only if the quality is experienced by all qualified critics, or, alternatively, that it really has the quality even if it is experienced only by some qualified critics. But the first response leaves artworks with too few aesthetic qualities and makes almost all aesthetic judgments false, while the second response ascribes too many aesthetic qualities, even incompatible ones, to the same objects. Another possibility for the realist is to hold that when critics disagree about the evaluative aesthetic properties they ascribe, there are nevertheless real nonevaluative aesthetic properties that they agree on in perceiving. When, for example, one critic sees a painting as elegant and another as insipid, they nevertheless see the same aesthetic quality underlying these opposed evaluative qualities. But the problem with this response is, first, that it splits the account of aesthetic qualities in two and, second, that it fails to specify what the underlying aesthetic quality might be. The critics seem to react to the base, nonaesthetic formal properties of the painting with different responses.

The relativist account therefore seems preferable. In addition, it explains why we cannot know from an objective description of an object whether it has a given aesthetic quality. We can infer that it does from testimony

only if we are certain that the testifier shares our taste. But the relativist must still account for the normativity of aesthetic judgments and how they are justified.

THE JUSTIFICATION OF ASCRIPTIONS OF AESTHETIC QUALITIES

Objective base properties justify ascriptions of aesthetic qualities, and these justify overall evaluations. But there are no principles at either level. On the second level, elegance, for example, usually contributes to a positive evaluation. But prose or painting styles can be too elegant for their subject matters, lessening the overall impact of their works. In view of the lack of principles and the relativity of aesthetic qualities to different tastes, how do these justifications work?

Ascriptions of aesthetic qualities are unjustified when based on inattention, bias, lack of knowledge of the formal properties of a work or its historical context, or an unacceptable interpretation. In asserting that an object has an aesthetic quality, one makes an implicit claim that one's judgment is not based on any of these disqualifying factors. This is equivalent to the claim that a fully competent or ideal critic who shares one's taste would respond to the object in the same way, would ascribe the same property to it. Thus, the relation between objective nonaesthetic properties and aesthetic qualities is simply that the former cause fully competent critics with certain tastes to respond in ways expressed by ascriptions of the aesthetic qualities.

Arguments over the presence of aesthetic qualities proceed until it is clear that both parties are fully competent in the circumstances to make the aesthetic judgments they make. Typically, critics proceed by pointing to the objective properties in the given historical context that elicit the responses expressed in their judgments, under the assumption that the other party has for one reason or another missed the relevance of the underlying base properties. But once the relevant base properties have been noted and interpretations agreed on, argument will cease, and the parties will have to accept ultimate differences in taste.

If aesthetic qualities are instantiated relative not only to contexts but, more significantly, to tastes of qualified critics, then two main questions remain. First, when do fully qualified critics share tastes? Can those who do share tastes nevertheless disagree about particular ascriptions of aesthetic qualities? Second, why should the judgments of such critics have normative force for others? If fully qualified or ideal critics who share tastes can disagree in their ascriptions of aesthetic qualities, and if objects have the relational properties that these critics ascribe, then the same problem that relativizing was intended to solve, the ascription of incompatible qualities to the same objects, reappears. When such critics disagree, they therefore have slightly different tastes. But if an ordinary observer who shares tastes with an ideal critic in all other aesthetic judgments disagrees in a particular case, this is a strong (but not infallible) indication that the observer is not making a sound aesthetic judgment, that he is mistaken in ascribing the aesthetic quality to the object. Clarifying argument is then in order. Only when all relevant base properties have been noted and acceptable interpretations agreed on can disagreements be explained away as reflecting different tastes. The object will then be asserted to have the disputed aesthetic qualities only relative to these different tastes.

To turn to the second question, when an ordinary observer disagrees with a fully competent critic who shares his taste, why should he accept the judgment of the critic as correct or normative for him? The answer can only be that such critics experience works more deeply—on cognitive, emotional, imaginative, and perceptual levels simultaneously. The works and their aesthetic qualities, when so appreciated, offer lasting satisfaction.

See also Aesthetic Experience; Aesthetic Judgment; Art, Interpretation of.

Bibliography

Beardsley, Monroe. "The Descriptivist Account of Aesthetic Attributions." *Revue internationale de philosophie* 28 (1974): 336–352.

Beardsley, Monroe. "What Is an Aesthetic Quality?" *Theoria* 39 (1973): 50–70.

Bender, John. "Realism, Supervenience, and Irresoluble Aesthetic Disputes." *Journal of Aesthetics and Art Criticism* 54 (1996): 371–381.

Budd, Malcolm. "Aesthetic Judgments, Aesthetic Principles, and Aesthetic Properties." *European Journal of Philosophy* 7 (1999): 295–311.

Cohen, Ted. "Aesthetic/Non-aesthetic and the Concept of Taste." *Theoria* 39 (1973): 113–152.

Currie, Gregory. "Supervenience, Essentialism, and Aesthetic Properties." *Philosophical Studies* 58 (1990): 243–257.

De Clercq, Rafael. "The Concept of an Aesthetic Property." *Journal of Aesthetics and Art Criticism* 60 (2002): 167–176.

Eaton, Marcia. "The Intrinsic, Non-supervenient Nature of Aesthetic Properties." *Journal of Aesthetics and Art Criticism* 52 (1994): 383–397.

Goldman, Alan. *Aesthetic Value*. Boulder, CO: Westview Press, 1995.

Goodman, Nelson. *Languages of Art*. Oxford, U.K.: Oxford University Press, 1969.

Hermeren, Goran. *The Nature of Aesthetic Qualities*. Lund, Sweden: Lund University Press, 1988.
Hume, David. "Of the Standard of Taste." In his *Essays, Moral, Political, and Literary*. Oxford, U.K.: Oxford University Press, 1963.
Hutcheson, Francis. *An Inquiry into the Original of Our Ideas of Beauty and Virtue*. New York: Garland, 1971.
Isenberg, Arnold. "Critical Communication." *Philosophical Review* 58 (1949): 330–344.
Kant, Immanuel. *Critique of Judgment*. Translated by J. H. Bernard. New York: Hafner, 1966.
Kivy, Peter. "Aesthetic Appraisals and Aesthetic Qualities." *Journal of Philosophy* 65 (1968): 85–93.
Levinson, Jerrold. "Aesthetic Properties, Evaluative Force, and Differences in Sensibility." In *Aesthetic Concepts: Essays after Sibley*, edited by Emily Brady and Jerrold Levinson. Oxford, U.K.: Oxford University Press, 2001.
Mitias, Michael, ed. *Aesthetic Quality and Aesthetic Experience*. Amsterdam: Rodolpi, 1988.
Sibley, Frank. "A Contemporary Theory of Aesthetic Qualities: Aesthetic Concepts." *Philosophical Review* 68 (1959): 421–450.
Walton, Kendall. "Categories of Art." *Philosophical Review* 79 (1970): 334–367.
Zangwill, Nick. *The Metaphysics of Beauty*. Ithaca, NY: Cornell University Press, 2001.
Zemach, Eddy. *Real Beauty*. University Park: Pennsylvania State University Press, 1997.

Alan H. Goldman (2005)

AESTHETICS

The Encyclopedia features two very detailed survey entries, *Aesthetics, History of*, and *Aesthetics, Problems of*, as well as the following entries: *Beauty*; *Humor*; *Metaphor*; *Tragedy*; and *Ugliness*.

AESTHETICS, HISTORY OF

In the West, the history of systematic philosophizing about the arts begins with Plato. But his great achievement was preceded, and prepared for, by certain developments in the preceding two hundred years, of which we know or can guess only a little. Thus, the famous aesthetic judgment—if such it was—of the picture on Achilles' shield, "That was a marvellous piece of work" (*Iliad* XVIII 548), hints at the beginning of wonder about imitation, i.e., the relation between representation and object, or appearance and reality. Plato shows the aesthetic consequences of the thinking on this problem by Democritus and Parmenides. Further, the elevation of Homer and Hesiod to the status of wise men and seers, and moral and religious teachers, led to a dispute over the truthfulness of poetry when they were attacked by Xenophanes and Heraclitus for their philosophical ignorance and misrepresentation of the gods. Homer and Hesiod themselves raised the question of the source of the artist's inspiration, which they attributed to divine power (*Odyssey* VIII; *Theogony* 22 ff.). Pindar traced this gift to the gods but allowed that the poet's skill can be developed by his own effort. Pythagoras and his Order discovered the dependence of musical intervals on the ratios of the lengths of stretched strings, generalized this discovery into a theory about the elements of the material world (that they either are, or depend upon, numbers), and developed an elaborate ethical and therapeutic theory of music, which, according to them, is capable of strengthening or restoring the harmony of the individual soul—*harmonia* being the term for the primary interval, the octave.

PLATO

Nearly all of the fundamental aesthetic problems were broached, and some were deeply considered, by Plato. The questions he raised and the arguments he framed are astonishingly varied and deep. They are scattered throughout his dialogues, but the principal discussions are in (*a*) the *Ion, Symposium*, and *Republic*, belonging to Plato's early, pre-Academy period (roughly 399–387 BCE); (*b*) the *Sophist* and *Laws*, written at the end of his life (roughly 367–348/347 BCE); and (*c*) the *Phaedrus*, which lies between these periods. Though perhaps not Plato's, the *Greater Hippias* is very Platonic and may be drawn upon. (In this entry, no distinction will be attempted between Plato's views and those of Socrates.)

ART AND CRAFT. When today we speak of Plato's aesthetics, we mean his philosophical views about those fine arts that he discusses: visual arts (painting, sculpture, architecture), literary arts (epic, lyric, and dramatic poetry), and mixed musical arts (dance and song). Plato does not himself assign them a special name; for him they belong in the more general class of "craft" (*technē*), which includes all skills in making or doing, from woodcraft to statecraft. In the *Sophist* (265–266), crafts are divided into "acquisitive" and "productive," the latter being subdivided into (1) production of actual objects, which may be either human or divine (plants and elements by god, houses and knives by men), and (2) production of "images" (*idola*), which may also be human or divine (reflections and dreams by god; pictures by men). Images, which imitate their originals but cannot fulfill their function, are further subdivided; the imitator may produce (1) a genuine likeness (*eikon*), with the same properties as his model, or (2) an apparent likeness, or semblance (*phantasma*),

which merely *looks* like the original (as when the architect makes his columns swell at the top so that they will not appear to diminish). There is thus false imitation, the making of deceptive semblances. Yet Plato finds this distinction troublesome to maintain, for it is essential to any imitation that in some way it falls short of its original; if it were perfect, it would not be an image (*eidolon*), but another example of the same thing, another bed or knife (*Cratylus* 432). So all imitation is in a sense both true and untrue, has both being and nonbeing (*Sophist* 240C).

IMITATION. The term "imitation" (*mimesis*) is one of the most troublesome in Plato's aesthetics, for its denotation constantly expands and contracts with the movement of the dialectic, along with that of its substitutes and near synonyms, *methexis* (participation), *homoiosis* (likeness), and *paraplesia* (resemblance).

If, in one sense, all created things are imitations of their eternal archetypes, or "forms," Plato seems also to regard paintings, dramatic poems, and songs as imitations in a narrower sense: They are images. It is this that places the arts at the second remove from the reality of the forms, on the lowest of the four levels of cognition, *eikasia* (imagining) (*Republic* 509–511). Some works of art, however—and Plato sometimes speaks as though he meant all of them—are imitative in the more pejorative sense, as deceptive semblances. In Book X of the *Republic*, the painter is said to represent the bed, not as it is but as it appears. It is this that puts him in the "tribe of imitators" (*Timaeus* 19D) and allies him with those pseudo craftsmen of the *Gorgias* (463–465) who do not possess a genuine craft, like medicine, but a pseudo craft, or knack (*tribē*), like cosmetics, which gives us the bloom of health rather than health itself.

BEAUTY. By this route, Plato approaches the question that is of great importance to him as a metaphysician: Do the arts contain, or convey, knowledge? Before coming to this question, there is another to be considered. If the architect, as a maker of semblances, changes reality to make it look better, why does he do this? He seeks those images that will appear beautiful (*Sophist* 236A). This is another basic fact about the arts, in Plato's view; they can embody in various degrees the quality of beauty (*to kalon*—a term that can branch out into more general senses of appropriateness or fitness to function but that often appears in a more strictly aesthetic sense). The beauty of concrete things may change or disappear, may appear to some but not to others (*Republic* 479A); but behind these temporal embodiments there is an eternal and absolute form of beauty. Its existence can be demonstrated dialectically, like that of the other forms; but direct acquaintance with it is to be sought, Plato says, via the partial and dimmer beauties open to the senses—and it is easier of access than the other forms (*Phaedrus* 249B–C).

The path to beauty is described most fully in the *Symposium* A man possessed by love (*eros*) of beauty is to progress from bodily beauty to beauty of mind, to beauty of institutions and laws and the sciences themselves, and finally to beauty in itself. It is noteworthy that Diotima of Mantineia, who presents this picture, does not assign to the arts any role in assisting this progress; that step was taken by Plato's successors.

It is also important to ask what beauty is, or, if that cannot be stated abstractly, what the conditions are under which beauty will be embodied in an object. The argument in the *Greater Hippias* takes up several possibilities, especially the possibility that the beautiful either is, or depends upon, what is beneficial or what pleases through the senses of hearing and sight. But in the *Philebus*, a careful discussion leads to the conclusion that beautiful things are made with care in the due proportion of part to part, by mathematical measurement (cf. *Timaeus* 87C–D; *Statesman* 284A). "The qualities of measure (*metron*) and proportion (*symmetron*) invariably … constitute beauty and excellence" (*Philebus* 64E, Hackforth translation). And because it is, or depends upon, measure, beauty is assigned a high place in the final list of goods (*Philebus* 66A–B; cf. *Sophist* 228B).

ART AND KNOWLEDGE. Knowledge (*episteme*), as distinct from mere opinion (*doxa*), is a grasp of the eternal forms; and Plato clearly denies it to the arts, as imitations of imitations (*Republic* 598–601). So the poet is placed on the sixth level of knowledge in the *Phaedrus* (248D), and Ion is said to interpret Homer not by "art or knowledge" (532C) but in an irrational way (cf. *Apology* 22), for he does not know what he is saying or why he might be right or wrong. On the other hand, a work of art that embodies beauty has some direct relation to one form. And if the artist inspired by the Muses is like a diviner in not knowing what he is doing (*Meno* 99C; *Timaeus* 71E–72A), he may have a kind of insight that goes beyond ordinary knowledge (cf. *Laws* 682A). His madness (*mania*) may be possession by a divinity that inspires him to truth (*Phaedrus* 245A; *Ion* 533E, 536B). Moreover, since the arts can give us genuine likenesses, not only of appearances but of actualities, and even imitate the ethical character of the human soul (*Republic* 400–401B; cf. Xenophon, *Memorabilia* III viii), it is possible, and indeed obligatory, to judge

them by their truth, or their resemblance to actuality. The competent judge, especially of dance and song, must have "first, a knowledge of the nature of the original; next, a knowledge of the correctness of the copy; and thirdly, a knowledge of the excellence with which the copy is executed" (*Laws* 669A–B, Bury translation).

ART AND MORALITY. The supreme craft, for Plato, is the art of the legislator and educator, who must have the final say about the arts, for his task is to insure that they play their proper role in the life of the entire social order. The first problem is to discover what effects the arts have on people, and this problem has two aspects. First, there is the enjoyability of art. On the one hand, just insofar as it has beauty, the pleasures art gives are pure, unalloyed, and harmless (*Phaedrus* 51B–C), unlike the pleasure of scratching an itch, which is preceded and followed by discomfort. But, on the other hand, dramatic poetry involves the representation of unworthy characters behaving in undesirable ways (ranting and wailing) and tempts the audience into immoderate laughter or weeping. Therefore its pleasures are to be condemned for their unworthy effect on character. Second, when we consider this tendency of the arts to influence character and conduct, there are again two sides to the matter. In his *Republic* and *Laws*, Plato makes it quite clear that he thinks the literary imitation of evil conduct is an implicit invitation to imitate the conduct in one's life (*Laws* 665B). Thus the stories of gods and heroes who behave immorally have to be excluded from the education of the young guardians in the *Republic*, and stories in which the gods and heroes behave as they should must either be found or written (*Republic* 376E–411; cf. *Laws* 800–802, 664A). Music composed in enervating modes must also be replaced by a suitable kind (*Republic* 398E, 411A).

But this does not mean that the arts have no role to play in the cultural life and education of the citizens. Indeed, the fear of their power that underlies Plato's severe censorship and regulation is accompanied by an equally great respect. The measure that is so closely allied to beauty is, after all, closely allied to goodness and virtue too (*Laws* 655A; *Protagoras* 326A–B; *Republic* 432). Music and poetry and dancing are, at their best, indispensable means of character education, able to make men better and more virtuous (*Laws* 653–654, 664). The problem, as Plato in his role of legislator sees it, is to ensure the social responsibility of the creative artist by insisting that his own good, like that of every citizen, be subordinated and made conducive to the good of all.

ARISTOTLE

Our knowledge of Aristotle's aesthetic theory comes chiefly from the little collection of lecture notes that has come down to us as the *Poetics*, composed probably about 347–342 BCE and later added to. The text is corrupt, the argument condensed and puzzling. No work in the history of aesthetics has given rise to such vexatious problems of interpretation; no work has had so great an influence on the theory and practice of literary criticism.

THE ART OF POETRY. Aristotle's first task is to define the art of poetry (*poietike*), which is his subject. He assumes a distinction between three kinds of "thought," knowing (*theoria*), doing (*praxis*), and making (*poiesis*) (see *Metaphysics* E 1; *Topics* VI 6); but in the *Poetics*, "poiesis" is taken in a narrower sense. One kind of making is imitation, which Aristotle seems to take fairly straightforwardly as representation of objects or events. The imitative art divides into (1) the art of imitating visual appearances by means of color and drawing and (2) the art of poetry, the imitation of a human action (*praxis*) through verse, song, and dance (*Poetics*, Ch. 1). Thus the art of poetry is distinguished from painting by its medium (words, melody, rhythm) and from versified history or philosophy (the poem of Empedocles) by virtue of the object it imitates. Two of the species of the poetic art are of primary concern to Aristotle: drama (either tragic or comic) and epic poetry, distinguished from comedy by the gravity of the actions imitated (Chs. 2, 6).

What is of the first importance in Aristotle's treatise is his method of inquiry, for he aims to present a systematic theory of a particular literary genre. He asks: What is the nature of the tragic art? And this leads him to inquire not only into its material, formal, and efficient causes (many of his observations under these headings are of permanent value to literary theory) but also into its final cause or end (*telos*). What is a good tragedy, and what makes it good; what are "the causes of artistic excellence and the opposite" (Ch. 26, G. F. Else translation)? This function of tragedy, he thinks, must be to provide a certain kind of enjoyable experience—the "proper pleasure" (*oikeia he-done*) of tragedy (Chs. 14, 23, 26)—and if the nature of this pleasure can be determined it will then be possible to justify the criteria by means of which one can say that one tragedy is better than another.

THE PLEASURE OF IMITATION. Aristotle suggests briefly (Ch. 4) two motives that give rise to tragedy. The first is that imitation is natural; and the recognizing of imitation is naturally pleasurable to man because man

finds learning pleasant, and recognizing, say, a picture of a dog, is a form of learning (cf. *Rhetoric* I xi). Since tragedy is an imitation of a special sort of object, namely fearful and pitiable events, its proper pleasure "is the pleasure that comes from pity and fear by means of imitation" (Ch. 4, Else translation). The problem that evidently arises is how we can derive pleasure from feeling emotions that are painful (cf. the definitions of "fear" and "pity" in *Rhetoric* II v, viii). Aristotle's nearest answer seems to be that though the object imitated may be in itself unpleasant to contemplate, the pleasure of seeing the imitation may overcome our distaste—as with skilled drawings of cadavers (see *De Partibus Animalium* I v; *Rhetoric* I xi). Here Aristotle is offering a partial answer to one of Plato's grounds for skepticism about art; he takes the basic aesthetic pleasure as a cognitive one, of the same genus as the philosopher's (though no doubt of a lower level).

THE PLEASURE OF BEAUTY. Tragedy also grows, Aristotle says (Ch. 4), out of our natural disposition to "melody and rhythm." He does not develop this point and may be postulating a kind of decorative impulse. But if we may think here of Plato's *Philebus*, our pleasure in melody and rhythm may be taken as pleasure in beauty in general. "A beautiful (*kalliste*) thing, either a living creature or any structure made of parts, must have not only an orderly arrangement of those parts, but a size which is not accidental" (Ch. 7). Thus a tragedy, or its plot, may be "beautiful," i.e., artistically excellent (Chs. 1, 13). And the "proper pleasure" of the epic, for example, depends on its unity, on being "like a single whole creature" (*zoon*) with a beginning, middle, and end (Ch. 23). This analogy echoes Plato's *Phaedrus* 264c. For the fineness of the object sensed or contemplated produces the highest degree of that pleasure that is proper to the organ sensing or mind contemplating (*Nicomachean Ethics* X iv).

THE UNIVERSAL. If the function of tragic poetry is to provide a certain species of enjoyment, we can then inquire into the features of a particular work that will promote or inhibit this enjoyment. Its concentration and coherence depend in large part upon the plot and the sense of inevitability in its development (Ch. 10). This is evidently achieved most fully when the characters act in accordance with their natures, when they do the "kinds of thing a certain kind of person will say or do in accordance with probability or necessity, which is what poetic composition aims at" (Ch. 9, Else translation). These sorts of behavior, i.e., behavior that is motivated in accordance with psychological laws, Aristotle calls "universal," contrasting them with the events in a historical chronicle, which he thinks of as a causally unconnected string of particular incidents ("what Alcibiades did or had done to him").

This famous passage has inspired many later theories about art imitating universals or essences, but the gist of it (for Aristotle) is that the poet must make his plot plausible by relying on general psychological truths. This important point adds another level to Aristotle's defense (against Plato) of the cognitive status of poetry, for the poet must at least understand human nature or he cannot even produce a good plot.

THE CATHARSIS. In Aristotle's definition of tragedy (Ch. 6) there is one phrase that has given rise to an enormous amount of interpretation: *di eleou kai phobou perainousa ten ton toiouton pathematon katharsin* (translated in the traditional way by Butcher: "through pity and fear effecting the proper purgation of these emotions"). Thus Aristotle is interpreted as having a further theory, not about the immediate pleasure of tragedy but about its deeper psychological effects. This phrase is the only basis for such an interpretation in the *Poetics*; but in the *Politics* (VIII 7), Aristotle clearly does propose a cathartic theory of music and even says he will explain catharsis further "when hereafter we speak of poetry"—a remark that possibly refers to the presumed lost parts of the *Poetics*. If tragedy produces a catharsis of the emotions, there are still other problems in deciding what Aristotle had in mind—whether, for example, he meant it in a medical sense (a purgation of the emotions, their elimination by mental physic) or in a religious and lustratory sense (a purification of emotions, their transformation into a less harmful form). Both senses had precedents. There is also the question whether Aristotle believed in a catharsis of pity and fear alone, or, through them, of all destructive emotions.

In any case, on this interpretation, Aristotle would be answering Plato's second objection to poetry in Book X of the *Republic*, by saying that poetry helps men to be rational. The traditional interpretation has been interestingly challenged in recent years by Professor Gerald F. Else, who argues that the catharsis is not an effect on the audience or reader but something accomplished in the play itself, a purification of the hero, a release from the "blood pollution" of his crime, through his recognition of it, his horror at it, and the discovery that it was due to a "serious mistake" (*hamartia*) on his part. This reading does not seem to fit some of the tragedies. If it is correct, Aristotle has no therapeutic theory of tragedy at all, but

he may still be replying to Plato that the immoral effects of tragedy are not to be feared, since the finest ones, at least, will have to show a kind of moral progress if they are to be structurally capable of moving the spectator tragically.

THE LATER CLASSICAL PHILOSOPHERS

Aristotle's *Poetics* does not seem to have been available to his successors. His ideas had some influence via the works (now largely lost) of his favorite pupil, Theophrastus; and the *Tractatus Coislinianus* (Greek, probably first century BCE) shows an acquaintance with his work, for its definition of comedy parallels remarkably Aristotle's definition of tragedy. During the later classical period, Stoicism, Epicureanism, skepticism, and Neoplatonism flourished competitively, and each of these schools of thought had some contribution to make to the history of aesthetics.

STOICISM. The Stoics were much interested in poetry and in problems of semantics and logic. Zeno, Cleanthes, and Chrysippus wrote treatises on poetry, no longer extant. From Philodemus we know of a work on music by the Stoic Diogenes of Babylon, and from Cicero's *De Officiis* of a work on beauty by Panaetius. Both seem to have held that beauty depends on the arrangement of parts (*convenientia partium*, in Cicero's phrase). The delight in beauty was connected with the virtue that expresses itself in an ordered life, with decorum (*to prepon*). Thus not only irrational pleasure (*hedone*), but a rational elevation of the soul (*chara*), in keeping with the Stoic goal of tranquillity, was thought to be obtainable from poetry of the right sort. The Stoics emphasized the moral benefit of poetry as its chief justification and held that it might allegorize true philosophy (see Strabo, *Geography* I, i, 10; I, ii, 3).

EPICUREANISM. The Epicureans are said (by Sextus Empiricus, *Against the Professors* VI, 27) to have disapproved of music and its pleasure, but it appears that this is partly based on a misunderstanding of Epicurus's aversion to music *criticism* (see Plutarch, *That It Is Not Possible to Live Pleasurably According to the Doctrine of Epicurus* 13). Two important works by Philodemus of Gadara (first century BCE), parts of which have been unearthed at Herculaneum, give further evidence of Epicurean thinking about the arts. In his work *On Music* (*Peri Mousikes*), Philodemus strikes the earliest known blow for what later was called "formalism," by arguing (against the Pythagoreans, Plato, and Aristotle) that music by itself—apart from the words, whose effects are often confused with the music itself—is incapable either of arousing emotions or of effecting ethical transformations of the soul. And in his work *On Poems* (*Peri Poematon*) he argued that specifically poetic goodness (*to poietikon agathon*) is not determined either by the moral-didactic aim (*didaskalia*), by the pleasure of technique and form (*psychagogia*), or by a mere addition of the two, but by a unity of form and content—his conception of which we do not now know.

The main lines of reflection about literature during the Roman period seem to have been practical and pedagogical. Two works were outstandingly influential (the second, however, not until its rediscovery in the modern period): the *Ars Poetica*, or *Epistle to the Pisos*, of Horace, which discusses many questions of style and form, and the work *On Elevation in Poetry* (*Peri Hypsous*, or *On the Sublime*), probably written during the first century CE, perhaps by a Greek named "Longinus." This lively and brilliant work defines the quality of great writing in affective terms, as that which transports the soul; and it investigates the stylistic and formal conditions of this effect.

PLOTINUS. The philosophical reflection that continued in the Platonic schools until the Academy at Athens was closed by Justinian I in CE, 529 culminated in the Neoplatonic system of Plotinus. Three of his fifty-four tractates, which make up the six *Enneads*, deal especially with aesthetic matters: "On Beauty" (I, vi); "On the Intellectual Beauty" (V, viii); and "How the Multiplicity of the Ideal-Forms Came into Being; and on the Good" (VI, vii).

Behind the visible world, in this view, stands "the one" (*to hen*), or "the first," which is ultimate reality in its first "hypostasis," or role, beyond all conception and knowledge. In its second hypostasis, reality is "intellect," or "mind" (*nous*), but also the Platonic forms that are known by mind. In its third hypostasis it is the "all-soul" (*psyche*), or principle of creativity and life. Within his scheme—infinite gradations of being "emanating" from the central "light"—Plotinus develops a theory of beauty that is highly original, though inspired by the *Symposium* and other Platonic dialogues. The tractate "On Beauty" (MacKenna and Page translation) begins by noting that Beauty lies in things seen and heard, and also in good character and conduct (I, vi, 1); and the question is, "What … is it that gives comeliness to all these things?"

The first answer considered, and rejected, is that of the Stoics. Beauty is, or depends on, symmetry. Plotinus argues that simple sense qualities (colors and tones), and also moral qualities, can have beauty though they cannot be symmetrical; moreover, an object can lose some of its beauty (as when a person dies) without losing any sym-

metry (VI, vii, 22). Therefore, symmetry is neither a necessary nor sufficient condition of beauty. It is not beauty but participation in ideal-form, that is, embodiment of Platonic ideas, that marks the difference in a stone before and after the sculptor carves it; for he gives it form. Where ideal-form enters, he says, confusion has been "rallied … into co-operation" (I, vi, 2): when an object becomes unified, "Beauty enthrones itself." A homogeneous thing, like a patch of color, is already unified by similarity throughout; a heterogeneous thing, like a house or ship, is unified by the dominance of the form, which is a divine thought (I, vi, 2). In the experience of beauty, the soul finds joy in recognizing in the object an "affinity" to itself; for in this affinity it becomes aware of its own participation in ideal-form and its divinity. Here is the historical source of mysticism and romanticism in aesthetics.

Love, in Plotinus's system, is always the love of beauty (III, v, 1) and of absolute and ultimate beauty through its lesser and dimmer manifestations in nature or in the work of the artist-craftsman (I, vi, 7; VI, ii, 18; V, viii, 8–10). Something of Plato's ambivalence toward art reappears in Plotinus's account at this point, though muted and closer to being overcome in the basic monism of the system. We ascend from the contemplation of sensuous beauty to delight in beautiful deeds, to moral beauty and the beauty of institutions, and thence to absolute beauty (I, vi, 8–9; II, ix, 16). Plotinus distinguishes three ways to truth, that of the musician, the lover, and the metaphysician (I, iii, 1–2); and he speaks of nature as offering a loveliness that cannot help but lead the admiring contemplator to thought of the higher beauties that are reflected there (II, ix, 7; V, viii, 2–3). Nor are the arts to be neglected, on the ground that they are mere imitations (here he comes closest to correcting the *Republic*, Book X), for both the painting and the object it copies are, after all, both imitations of the ideal-form; moreover, the painter may be able to imitate form all the more truly, to "add where nature is lacking" (V, viii, 1; cf. V, ix, 11). Yet, in his more religious mood, Plotinus reminds us that earthly and visible beauty may distract us from the infinite (V, v, 12), that "authentic beauty," or "beyond-beauty," is invisible (VI, vii, 33); and he who has *become* beautiful, and hence divine, no longer sees or needs it (V, vii, 11). The ladder, to use once more a too-familiar similitude, is kicked away by the philosophic mystic once he reaches home.

THE MIDDLE AGES

The early church Fathers were somewhat doubtful of beauty and the arts: They feared that a keen interest in earthly things might endanger the soul, whose true vocation lies elsewhere—especially since the literature, drama, and visual art they were acquainted with was closely associated with the pagan cultures of Greece and Rome. But despite the danger of idolatry, sculpture and painting became accepted as legitimate aids to piety, and literature became accepted as part of education in the liberal arts. Concern with aesthetic problems was not a prominent part of medieval philosophy, but some important lines of thought can be observed in the works of the two greatest thinkers.

ST. AUGUSTINE. In his *Confessions* (IV, xiii), Augustine tells a little of his lost early work, *De Pulchro et Apto* ("On the Beautiful and Fitting"), in which he distinguished a beauty that belongs to things in virtue of their forming a whole and a beauty that belongs to things in virtue of their fitting in with something else or being part of a whole. It is not possible to be sure, from his brief description, of the exact nature of this distinction. His later thoughts on beauty are scattered throughout his works, and especially in *De Ordine* ("Concerning Order," CE 386), *De Vera Religione* ("Concerning True Religion," CE 390), and *De Musica* (CE 388–391), a treatise on meter.

The key concepts in Augustine's theory are unity, number, equality, proportion, and order; and unity is the basic notion, not only in art (*De Ordine* II, xv, 42) but in reality. The existence of individual things as units, and the possibility of comparing them with respect to equality or likeness, gives rise to proportion, measure, and number (*De Musica* VI, xiv, 44; xvii, 56; *De Libero Arbitrio* II, viii, 22). Number, he emphasizes in various places, is fundamental both to being and to beauty—"Examine the beauty of bodily form, and you will find that everything is in its place by number" (*De Libero Arbitrio* II, xvi, 42, Burleigh translation). Number gives rise to order, the arrangement of equal and unequal parts into an integrated complex in accordance with an end. And from order comes a second-level kind of unity, the emergent unity of heterogeneous wholes, harmonized or made symmetrical through internal relations of likeness between the parts (*De Vera Religione* xxx, 55; xxxii, 59; *De Musica* VI, xvii, 58).

An important feature of Augustine's theory is that the perception of beauty involves a normative judgment. We perceive the ordered object as being what it ought to be, the disordered object as falling short; hence the painter can correct as he goes along and the critic can judge (*De Vera Religione* xxxii, 60). But this rightness or wrongness cannot be merely sensed (*De Musica* VI, xii,

34); the spectator must bring with him a concept of ideal order, given to him by a "divine illumination." It follows that judgment of beauty is objectively valid; there can be no relativity in it (*De Trinitate* IX, vi, 10; *De Libero Arbitrio* II, xvi, 41).

Augustine also wrestled with the problem of literary truth, and in his *Soliloquies* (CE 387) he proposed a rather sophisticated distinction between different sorts of lying or deception. In the perceptual illusion, the straight oar pretends to be bent, and could be bent, but the statue could not be a man and therefore is not "mendacious." So, too, the fictional character could not be real and does not pretend to be real by his own will, but only follows the will of the poet (II, ix, 16; x, 18; cf. *Confessions* III, vi).

ST. THOMAS AQUINAS. Thomas's account of beauty is given tersely, almost casually, in a few key passages that have become justly famous for their rich implications. Goodness is one of the "transcendentals" in his metaphysics, being predicable of every being and cutting across the Aristotelian categories; it is Being considered in relation to desire (*Summa Theologica* I, q. 5, art. 1). The pleasant, or delightful, is one of the divisions of goodness—"that which terminates the movement of appetite in the form of rest in the thing desired, is called the *pleasant*" (*S.T.* I, q. 5, art. 6, Dominican Fathers translation). And beauty is what pleases on being seen (*Pulchra enim dicuntur quae visa placent, S.T.* I, q. 5, art. 4).

Here, of course, "seeing" extends to all cognitive grasp; the perception of beauty is a kind of knowing (this explains why it does not occur in the lower senses of smell and taste, *S.T.* I–II, q. 27, art. 1). Since cognition consists in abstracting the form that makes an object what it is, beauty depends on the form. Thomas's best-known statement about beauty occurs in the course of a discussion of Augustine's attempt to identify the persons of the Trinity with some of his key concepts, the Father with unity, etc. Beauty, he says, "includes three conditions" (*S.T.* I, q. 39, art. 8). First, there is "integrity or perfection" (*integritas sive perfectio*)—broken or injured objects, incomplete objects, are ugly. Second, there is "due proportion or harmony" (*debita proportio sive consonantia*), which may refer partly to the relations between parts of the object itself but mainly refers to a relation between the object and the perceiver: that the eminently visible object, for example, is proportioned to the sight. Third, there is "brightness or clarity" (*claritas*), or brilliance (see also *S.T.* II–II, q. 145, art 2; q. 180, art. 2). The third condition has been variously explicated; it is connected with the medieval Neoplatonic tradition in which light is a symbol of divine beauty and truth (see the pseudo-Dionysius on the *Divine Names*, Ch. 4; Robert Grosseteste, *De Luce*, and his commentary on the *Hexaëmeron*). Clarity is that "splendor of form [*resplendentia formae*] shining on the proportioned parts of matter" in the opusculum *De Pulchro et Bono* (I, vi, 2), written either by the young Thomas or his teacher Albertus Magnus. The conditions of beauty can be stated univocally, but beauty, being a part of goodness, is an analogical term (that is, has different senses when applied to different sorts of things). It signifies a whole family of qualities, for each thing is beautiful in its own way (Aquinas, *Commentary on the Psalms*, Psalm xliv, 2; cf. *Commentary on the Divine Names* iv, 5).

THE THEORY OF INTERPRETATION. The consuming tasks of the early Fathers, clarifying, reconciling, and systematizing Biblical texts in order to defend Christianity against external enemies and heretical deviations, required a method of exegetical interpretation. The Greek tradition of allegorizing Homer and Hesiod and the Rabbinical tradition of allegorical exposition of Jewish scriptures had been brought together and elaborately refined by Philo of Alexandria. His methods were adopted by Origen, who distinguished three levels of meaning in scripture: the literal, the moral, and the spiritual or mystical (see *De Principiis* IV, i, 16, 18, 20). This method was taken into the West by Hilary of Poitiers and Ambrose, bishop of Milan, and further developed by John Cassian, whose formulation and examples became standard throughout the medieval period up to the time of Dante (see Dante's letter to Can Grande, 1319, the Preface to the *Paradiso*).

In Cassian's example (*Collationes* xiv, 8), Jerusalem, in the Old Testament, is, "literally" or "historically," the city of the Jews; on the "allegorical," or what came to be called the "typical," level, it refers prophetically to the later church of Christ; on the "tropological," or moral, level, to the individual soul; on the "anagogical" level, to the heavenly City of God. The last three levels together are sometimes called the "allegorical," or (as by St. Thomas) the "spiritual," meaning. As Thomas also indicates (*Summa Theologica* I, q. 1, art 10), the "literal" meaning also includes metaphorical statements.

Origen insisted that all Biblical texts must have the highest level of meaning, the "spiritual," though they may lack a moral sense and may even fail to make sense on the literal level, if too great an absurdity would be entailed by taking them that way. In this he was followed by St. Augustine (*De Doctrina Christiana* III, x, 14; xv, 23) but not by Hugh of St. Victor (*De Scripturis*, v; *Eruditiones*

Didascalicon VI, iv, viii–xi), who held that the second-level meanings are a function of the first level, and a first-level meaning can always be found if metaphor is included in it.

Because Christianity taught that the world was created *ex nihilo* by God, rather than generated or molded out of something else, Christian thinkers tended, in the Middle Ages, to hold that nature itself must carry the marks or signs of its origin and be a symbolic embodiment of the Word; in this respect, like Holy Scripture, God's other creation, it can be subjected to interpretation. Thus, nature becomes an allegory, and every natural object a symbol of something beyond. This view reaches its fullest development in John Scotus Erigena (*De Divisione Naturae* I, iii) and St. Bonaventure (*Collationes in Hexaëmeron* II, 27).

Though these reflections were primarily theological, rather than aesthetic, they were of great significance to the later history of aesthetics: They raised important questions about the nature of metaphor and symbol, in literature as well as in theology; they initiated reflection on the general problem of interpreting works of art; and they showed the possibility of a broad philosophy of symbolic forms, in which all art might be understood as a kind of symbolism.

THE RENAISSANCE

The most interesting philosophical development in the fifteenth and sixteenth centuries was the revival, by a number of thinkers, of Platonism and the creation of a vigorous Neoplatonism. Of these thinkers, Marsilio Ficino, translator of Plato and Plotinus and founder of the new Academy (1462), was the greatest. In *De Amore* (his commentary on the *Symposium*, written 1474–1475) and in his principal work, the *Theologia Platonica*, Ficino took over a number of the leading aesthetic notions of the Greeks and of St. Augustine, and to them he added one of his most original ideas, a theory of contemplation based on Plato's *Phaedo*. In contemplation, he held, the soul withdraws to some extent from the body into a purely rational consciousness of the Platonic forms. This inward concentration is required for artistic creation, which involves detachment from the real, to anticipate what does not yet exist, and also is required for the experience of beauty (this explains why beauty can be grasped only by the intellectual faculties—sight, hearing, and thinking—and not by the lower senses).

More significant for the future, however, were the changes taking place in basic assumptions about the arts and in attitudes toward them. The most significant works on the fine arts were the three books on painting, sculpture, and architecture by Leon Battista Alberti, the large collection of notes toward a systematic treatise on painting by Leonardo da Vinci, and surviving memoranda and the two books, on geometry and perspective and on human proportions by Albrecht Dürer.

One of the most serious endeavors of these artists and others was to establish a status for painting within the liberal arts, separating it from the other manual crafts among which it had been classified throughout the medieval period. The painter, Alberti argued (in his *Della pittura*, 1436), requires a special talent and skill; he needs a liberal education and a knowledge of human affairs and human nature; he must be a scientist, in order to follow the laws of nature and produce accurate representations of natural events and human actions. His scientific knowledge, indeed, must be basically mathematical, for the theory of proportions and the theory of linear perspective (which preoccupied Renaissance theorists, and especially Dürer) are mathematical studies; and they provide the principles in terms of which paintings can be unified and made beautiful, but at the same time made to depict correctly. Leonardo's argument for the superiority of painting to poetry and music (and also, in some degree, to sculpture) followed similar lines (see the first part of the *Treatise on Painting*).

The concern for faithfulness of representation that is fundamental to Renaissance fine arts theory is also found in the developing theory of music. The music theorists, aiming to secure the place of music as a humanistic discipline, sought for a vocal music that would attain the powerful emotional and ethical effects attributed to Greek music. They stressed the importance of making the music follow the text, to intensify the meanings of the words. These ideas were defended, for example, by Gioseffe Zarlino, in his *Istitutioni Armoniche* (1558) and by Vincenzo Galilei, in his *Dialogo della musica antica e della moderna* (1581).

Renaissance poetics was dominated by Aristotle (especially the concept of poetry as imitation of human action) and Horace (the thesis that poetry aims to delight and instruct—though this dualism was rejected by one of the major theorists, Lodovico Castelvetro, in his commentary on Aristotle's *Poetics*, 1570). The concept of imitation was variously interpreted and criticized by the Italian theorists. Among the chief points of disagreement and contention was the question whether poetry must belong to fixed genres and obey rigid rules, such as the dramatic "unities" adopted so adamantly by Julius Caesar Scaliger in his *Poetics* (1561), and the question (as dis-

cussed, for example, in Sidney's *Defense of Poesie*, 1595) whether the poet is guilty of telling lies and of leading his readers into immorality. In these discussions, the Aristotelian *katharsis* and Plato's condemnation of the poets were central and recurrent topics.

THE ENLIGHTENMENT: CARTESIAN RATIONALISM

Though Descartes had no aesthetic theory, and indeed wrote nothing about the arts apart from his early *Compendium Musicae* (1618), his epistemological method and conclusions were decisive in the development of neoclassical aesthetics. As in other areas, the search for clarity of concept, rigor of deduction, and intuitive certainty of basic principles penetrated the realm of critical theory, and its effects can be traced in numerous works, for example, in Nicolas Boileau-Despréaux's *L'art poétique* (1674); in Alexander Pope's *Essay on Criticism* (1711); in Charles Du Fresnoy's *De Arte Graphica* (translated into French by Roger de Piles, 1668, into English by Dryden, 1695); and in Jean Philippe Rameau's *Traité de l'harmonie réduite à ses principes naturels* (1722). Cartesian and Aristotelian elements combined in the richly polysemous concepts of reason and nature, which became central to all theories of the arts. To follow nature and to follow rules of reason were identified in counsel to the creative artist as well as in critical judgment.

In the sixteenth century, the rules for making and for judging works of art were generally (but not always) supported by authority, either the supposed authority of Aristotle or the models provided by classical writers. The new rationalism in aesthetics was the hope that these rules could be given a more solid, a priori, foundation by deduction from a basic self-evident axiom, such as the principle that art is imitation of nature—where nature comprised the universal, the normal, the essential, the characteristic, the ideal. So, in Samuel Johnson (*Preface to Shakespeare*, 1765), "just representations of general Nature" become the end of art; the painter "is to examine, not the individual, but the species" (*Rasselas*, 1759, Ch. 10). And in the *Discourses* (1778) of Sir Joshua Reynolds, the painter is advised to "consider nature in the abstract, and represent in every one of his figures the character of its species" (III).

THE PROBLEM OF THE RULES. The controversy over the authority and infallibility of the rules reflected a conflict between reason and experience, between less and more empirical approaches to art. For example, Corneille, in his three *Discourses* (1660), admitted the necessity of observing unity of space, time, and action in dramatic construction but confessed also that he was by no means their "slave" and sometimes had to break or modify them for the sake of dramatic effect or the audience's enjoyment. Molière, in his *Critique de l'école des femmes* (1663), was even more outspoken in making experiment the test. However, other theorists held the line in France, for example, George de Scudéry and Charles de Saint-Évremond. Dryden, in his *Defense of an Essay of Dramatic Poesy* (1668), suggested that if drama has a function or end, there must be rules, but the rules themselves are only probable and rest in part upon experience. In this spirit, Johnson criticized the pseudo-Aristotelian rules of time and place.

In music, the conflict between reason and experience appeared in controversies over harmony and consonance, as well as over the absoluteness of rules, such as the avoidance of parallel fifths. The followers of Zarlino insisted on a mathematical basis for acceptable chords; the followers of Vincenzo Galilei were more willing to let the ear be the judge. A kind of reconciliation of these views appears in Leibniz's theory (*Principles of Nature and of Grace*, 1714, § 17) that, like all sensations, musical tones are confused mélanges of infinite sets of *petites perceptions* that at every moment are in pre-established harmony with the perceptions of all other monads; in hearing a chord, the soul unconsciously counts the beats and compares the mathematical ratio which, when simple, produces concord.

TOWARD A UNIFIED AESTHETICS. The Cartesian theory of knowledge led to a more systematic attempt at a metaphysics of art in the *Meditationes Philosophicae de Nonnullis ad Poema Pertinentibus* (1735) of Alexander Gottlieb Baumgarten. Baumgarten, who coined the term "aesthetics," aimed to provide an account of poetry (and indirectly of all art) as involving a particular form, or level, of cognition—"sensory cognition." He began with Descartes's distinctions (*Principles of Philosophy* I, xlv–xlvi), elaborated by Leibniz (*Discourse on Metaphysics*, xxiv), between clear and obscure ideas, and between distinct and confused ideas. Sense data are clear but confused, and poetry is "sensate discourse," that is, discourse in which such clear–confused ideas are linked together into a structure. The "extensive clarity" of a poem consists in the number of clear ideas combined in it, and the rules for making or judging poetry have to do with ways in which the extensive clarity of a poem may be increased or diminished.

Baumgarten's book is remarkably concise, and its formalized deductive manner, with definitions and deri-

vations, goes out of its way to declare the possibility of dealing in an acceptably rigorous Cartesian way with matters apparently so little suited for rigorous treatment. Though he did not finish his *Aesthetics*, which would have generalized his study of poetry, the makings of a general theory are present in the *Meditations*. Its basic principle is still the imitation of nature—the principle that is also fundamental to the influential work of the Abbé Charles Batteux, *Les beaux arts réduits à un même principe* (1746), and to the important classification of the fine arts in d'Alembert's *Discours préliminaire* to the *Encyclopédie* (1751).

The importance of Lessing's *Laokoon oder über die Grenzen der Malerei und Poesie* (1766) is that, though he did not reject the possibility of a system that will relate all the arts, he attacked superficial and deadening analogies (many of them based on the Horatian formula, *ut pictura poesis*, torn from its context). He looked for the specific individual potentialities and values of painting and poetry in their own distinctive mediums. The medium of an art is, he says, the "signs" (*Zeichen*) it uses for imitation; and painting and poetry, when carefully examined for their capacities to imitate, turn out to be radically different. Consisting of shapes and colors, side by side, painting is best at picturing objects and visible properties, and can only indirectly suggest actions; poetry is just the opposite. When a secondary power of an art is made primary, it cannot do its best work. By the clarity and vigor of his argument and his sharp criticism of prevailing assumptions, Lessing gave a new turn to aesthetics.

THE ENLIGHTENMENT: EMPIRICISM

Contemporaneous with the development of neoclassical critical theory was the divergent line of aesthetic inquiry pursued principally, though not exclusively, by British theorists in the Baconian tradition of empiricism. They were greatly interested in the psychology of art (though they were not merely psychologists), especially the creative process and the effects of art upon the beholder.

THE IMAGINATION. That the imagination (or "fancy") plays a central, if mysterious, role in artistic creation had long been acknowledged. Its mode of operation—the secret of inventiveness and originality—was not systematically investigated before the empiricists of the seventeenth century. Among the rationalists, the imagination, considered as an image-registering faculty or as an image-combining faculty, played little or no role in knowledge. (See Descartes's Rule III of the *Regulae* ["the blundering constructions of imagination"]; *Principles* I, lxxi–lxxiii; and *Meditation* VI.) But Bacon's *Advancement of Learning* (1605) placed the imagination as a faculty alongside memory and reason and assigned poetry to it, as history and philosophy (including, of course, both moral and natural philosophy) were assigned to the other faculties.

Thomas Hobbes, in the first chapters of his *Leviathan* (1651), undertook to give the first analysis of imagination, which he defined as "decaying sense" (I, ii), the phantasms, or images, that remain when the physiological motions of sensation cease. But besides this "simple imagination," which is passive, there is also "compound imagination," which creates novel images by rearranging old ones. Hobbes stated that the mind's "trains" of thought are guided by a general principle of association (I, iii), but he did not work it out very fully. Nor did Locke develop this idea very far in the famous chapter "Of the Association of Ideas" (II, xxxiii) that he added to the fourth edition (1700) of his *Essay concerning Human Understanding* (1690). The tendency of ideas that have accompanied each other to stick together and pull each other into the mind was noted by Locke as a pathological feature of the understanding: It explains various sorts of error and the difficulty of eradicating them (cf. *Conduct of the Understanding*, §41). The work of fancy is best seen, according to Locke, in the tendency of poetic language to become figurative. As long as we are interested in pleasure, we cannot be troubled by such ornaments of style; but metaphors and similes are "perfect cheats" when we are interested in truth (III, x, 34; cf. *Conduct of Understanding*, §§32–42). Locke here reflects a widespread distrust of imagination in the later seventeenth century. It is shown in a famous passage from Sprat's *History of the Royal Society* (1702), in which Sprat describes the "close, naked, natural way of speaking," in clearly defined words, required for scientific discourse, and contrasts it with the "specious tropes and figures" of poetry.

The theory of the association of ideas was developed into a systematic psychology by Hume, in his *Treatise of Human Nature* (1739–1740), and Hartley, in his *Observations on Man* (1749). In Hume, the tendency of ideas to consort with one another because of similarity, propinquity, or causal connection became a powerful principle for explaining many mental operations; and Hartley carried the method further. Despite attacks upon it, associationism played a crucial role in several eighteenth-century attempts to explain the pleasures of art.

THE PROBLEM OF TASTE. The investigation of the psychological effects of art and of the aesthetic experience

(in modern terms) developed along two distinct, but occasionally intersecting, paths: (1) the search for an adequate analysis and explanation of certain basic aesthetic qualities (the beautiful, the sublime) or (2) an inquiry into the nature and justification of critical judgment, the problem of "taste." Without trying to keep these completely separate, let us first consider those philosophers in the early part of the eighteenth century in whose thinking the second problem was uppermost.

One phase of aesthetic thinking was launched by the very influential writings of the third earl of Shaftesbury (see especially his *Moralists*, 1709, III; *Inquiry concerning Virtue or Merit*, 1699, I; and *Characteristics*, 1711). Shaftesbury's philosophy was basically Neoplatonic, but to emphasize the immediacy of our impression of beauty, and also to underline his view that the harmony perceived as beauty is also perceived as virtue, Shaftesbury gave the name "moral sense" to that "inward eye" that grasps harmony in both its aesthetic and ethical forms. The concept of a special faculty of aesthetic apprehension was one form of the theory of taste. Shaftesbury's other contributions to the development of aesthetics are his description of disinterestedness as a characteristic of the aesthetic attitude (*Moralists* III) and his appreciation (along with his contemporaries John Dennis and Thomas Burnet) of wild, fearful, and irregular forms of nature—a taste that helped bring into prominence, in the eighteenth century, the concept of the sublime as an aesthetic quality distinct from beauty.

Joseph Addison's *Spectator* papers on aesthetic enjoyment (1712, Nos. 409, 411–421) conceived taste as simply the capacity to discern those three qualities that give rise to "the pleasures of the imagination," greatness (that is, sublimity), uncommonness (novelty), and beauty. Addison made some attempt to explain why it is that the perception of these qualities is attended by so much pleasure of so special a sort, but he did not go far; his service (earning the appreciation he received from succeeding thinkers) was the lively and provocative way in which he raised many of the basic questions.

The first real treatise on aesthetics in the modern world was Francis Hutcheson's *Inquiry concerning Beauty, Order, Harmony, and Design*, the first part of *An Inquiry into the Original of our Ideas of Beauty and Virtue* (1725). From Shaftesbury, Hutcheson took the idea of an inner sense; the "sense of beauty" is the power to frame the idea of beauty when confronted with those qualities of objects suited to raise it. The sense of beauty does not depend on judgment or reflection; it does not respond to intellectual or utilitarian features of the world, nor does it depend on association of ideas. His analysis showed that we sense beauty in an object when it presents "a compound ratio of uniformity and variety" (2d ed., p. 17), so that beauty varies with either of these, if the other is held constant. A basis is thus laid for a nonrelativistic standard of judgment, and variations in actual preference are explained away as due to different expectations with which the beautiful object, in art or nature, is approached.

The question of a standard of taste was the chief concern of David Hume's thinking on aesthetic matters. In his *Treatise* (II, i, 8), he suggested that "beauty is such an order and construction of parts, as either by the *primary constitution* of our nature, by *custom*, or by *caprice*, is fitted to give a pleasure and satisfaction to the soul," thus allowing, like Hutcheson, who influenced him considerably, an immediate delight in beauty, but allowing also for a transfer of this delight by association. For example, the appearance (not necessarily the actuality) of convenience or utility explains why many objects are esteemed beautiful (III, iii, 1). Some types of beauty, then, are simply seen or missed; judgments of them cannot be corrected. But in other cases, especially in art, argument and reflection can correct judgment (see *Enquiry concerning the Principles of Morals*, 1751, Sec. 1). This problem is discussed most carefully in the essay "Of the Standard of Taste" (in *Four Dissertations*, 1757). Hume argued that it is natural to seek for a standard of taste, by which aesthetic preferences can be called correct or incorrect, especially as there are clear cases of error ("Bunyan is a better writer than Addison"). The rules, or criteria, of judgment are to be established by inductive inquiry into those features of works of art that enable them to please most highly a qualified perceiver, that is, one who is experienced, calm, unprejudiced. But there will always be areas within which preference is due to temperament, age, culture, and similar factors unchangeable by argument; there is no objective standard by which such differences can be rationally resolved.

THE AESTHETIC QUALITIES. The search for necessary and sufficient conditions of beauty and other aesthetic qualities (the concept of the "picturesque" was added late in the century) was continued enthusiastically in the latter half of the eighteenth century. In this debate, an important part was played by Edmund Burke's youthful work, *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and Beautiful* (1757). Its argument develops on two levels, phenomenological and physiological. The first task is to explain by what qualities objects excite in us the feelings of beauty ("love" without desire) and sublimity ("astonishment" without actual danger). The

feeling of the sublime, to begin with, involves a degree of horror—controlled horror—the mind being held and filled by what it contemplates (II, 1). Thus, any object that can excite the ideas of pain and danger, or is associated with such objects, or has qualities that can operate in a similar way, can be sublime (I, 7).

Burke then goes on to argue that obscurity, power, privation and emptiness, vastness approaching infinity, etc. contribute to sublimity (II, 3–8). Beauty is analogously treated: The paradigm emotion is response to female beauty, minus lust; and objects that are small, smooth, gently varying, delicate, etc. can give the feeling of beauty (III, 1–16). The same scene can be both beautiful and sublime, but because of the opposition in several of their conditions it cannot be very intensely either if it is both.

Burke then moves to his second level of explanation (IV, 1, 5). He asks what enables the perceptual qualities to evoke the feelings of beauty and sublimity, and he answers that they do so by producing physiological effects like those of actual love and terror. "Beauty acts by relaxing the solids of the whole system" (IV, 19)—this is one of Burke's celebrated hypotheses, a pioneering attempt at physiological aesthetics.

In this very fertile period of aesthetic investigation, many other writers, of various degrees of sophistication, contributed to the theory of beauty and sublimity and to the foundations of taste. Among the most important works, still worth reading for some of their suggestions, are Alexander Gerard's *Essay on Taste* (written by 1756, published 1759; see also his *Essay on Genius*, 1774), which made much use of association in explaining our pleasure in beauty, novelty, sublimity, imitation, harmony, ridicule, and virtue; Henry Home's (Lord Kames) *Elements of Criticism* (1762); Hugh Blair's *Lectures on Rhetoric and Belles Lettres* (given from 1759 on, published 1783); Thomas Reid's essay on Taste in his *Essays on the Intellectual Powers of Man* (1785). On the Continent, the question whether there is a special aesthetic sense was discussed, along with many other problems, by Jean-Pierre de Crousaz, *Traité du beau* (1714), and the Abbé Dubos, *Réflexions critiques sur la poésie et sur la peinture* (1719). Noteworthy also are Voltaire's *Temple du goût* (1733), Yves-Marie André's *Essai sur le beau* (1741), and especially the article on beauty that Diderot wrote for the *Encyclopédie* (1751), in which the experience of beauty is analyzed as the perception of "relationships" (*rapports*).

In general, the later development of empiricist aesthetics involved increasingly ambitious attempts to explain aesthetic phenomena by means of association; a further broadening of the acknowledged aesthetic qualities, away from a limited concept of beauty; further reflection on the nature of "genius," the capacity to "snatch a grace beyond the reach of art"; and a growing conviction that critical principles have to be justified, if they can be justified at all, in terms of empirical knowledge of the characteristic effects of art. The achievements and the high level of discussion reached by the empiricist movement can be seen very well in a later treatise by Archibald Alison, his *Essays on the Nature and Principles of Taste* (1790; rev. ed., which became highly influential, 1811). Alison abandoned the hope for simple formulas of beauty and resolved the pleasure of taste into the enjoyment of following a train of imaginations, in which some of the ideas produce emotions and in which the entire train is connected by a dominant emotion. No special sense is required; the principles of association explain everything. And the arguments by which Alison supported his main theses, the careful inductions at all points, are models of one kind of aesthetics. For example, he showed, by experimental comparisons, that particular qualities of objects, or of Hogarth's "line of beauty" (II, iv, 1, Part II), do not produce aesthetic pleasure unless they become "expressive," or take on the character of signs, by being able to initiate a train of associations; and it is the same, he said, with colors: "Purple, for instance, has acquired a character of Dignity, from its accidental connection with the Dress of Kings" (II, iii, 1).

GERMAN IDEALISM

By assigning to the problems of aesthetic judgment the major part of his third *Critique* (*The Critique of Judgment*, 1790), Kant became the first modern philosopher to make his aesthetic theory an integral part of a philosophic system. For in this volume he aimed to link the worlds of nature and freedom, which the first two *Critiques* had distinguished and separated.

KANT'S ANALYSIS OF JUDGMENTS OF TASTE. Kant recast the problems of eighteenth-century aesthetic thought, with which he was thoroughly familiar, in the characteristic form of the critical philosophy: How are judgments of the beautiful and the sublime possible? That is, in view of their evident subjectivity, how is their implicit claim to general validity to be vindicated? That such judgments claim general validity and yet are also subjective is argued by Kant, in careful detail, in the "Analytic of the Beautiful" and the "Analytic of the Sublime."

Judgments of beauty (also called "judgments of taste") are analyzed in terms of the four "moments" of the

table of categories: relation, quantity, quality, and modality. First, the judgment of taste does not (like ordinary judgments) subsume a representation under a concept, but states a relation between the representation and a special disinterested satisfaction, that is, a satisfaction independent of desire and interest (§5). Second, the judgment of taste, though singular in logical form ("This rose is beautiful"), lays title to universal acceptance, unlike a report of mere sensuous pleasure, which imposes no obligation to agree. Yet, paradoxically, it does not claim to be supportable by reasons, for no arguments can constrain anyone to agree with a judgment of taste (§9; cf. §33). Third, aesthetic satisfaction is evoked by an object that is purposive in its form, though in fact it has no purpose or function: because of a certain wholeness, it looks as though it were somehow made to be understood (§10; cf. §65 and Introduction): it has "purposiveness without purpose" (*Zweckmässigkeit ohne Zweck*). Fourth, the beautiful is claimed by the judgment of taste to have a necessary reference to aesthetic satisfaction (§18): not that when we find ourselves moved in this way by an object we can guarantee that all others will be similarly moved, but that they *ought* to take the same satisfaction we do in it.

THE PROBLEM OF VALIDATION. It is the above four aspects of the judgment of beauty that give rise to the philosophical problem of validation, which Kant formulates as he had the parallel problems in the earlier *Critiques*: How can their claim to necessity (and subjective universality) be legitimized? This can only be done, he argues, if it can be shown that the conditions presupposed in such a judgment are not confined to the individual who makes it, but may reasonably be ascribed to all rational beings. A minor clue is offered by the disinterestedness of aesthetic sansfaction; for if our satisfaction is in no way dependent on individual interests, it takes on a kind of intersubjectivity (§6). But the validation of the synthetic a priori judgment of taste requires something more searching, namely, a transcendental deduction.

The gist of this argument is as follows: Empirical knowledge is possible because the faculty of judgment can bring together general concepts and particular sense-intuitions prepared for it in the imagination. These cases of *determinate* judgment presuppose, however, a general harmony between the imagination, in its freedom as synthesizer of representations, and the understanding, in its a priori lawfulness. The formal purposiveness of an object as experienced can induce what Kant calls "a free play of the imagination," an intense disinterested pleasure that depends not on any particular knowledge but just on consciousness of the harmony of the two cognitive powers, imagination and understanding (§9). This is the pleasure we affirm in the judgment of taste. Since the general possibility of sharing knowledge with each other, which may be taken for granted, presupposes that in each of us there *is* a cooperation of imagination and understanding, it follows that every rational being has the *capacity* to feel, under appropriate perceptual conditions, this harmony of the cognitive powers. Therefore a true judgment of taste can legitimately claim to be true for all (§9; cf. §§35–39).

Kant's system requires that there be a dialectic of taste with an antinomy to be dissolved on the principles of critical philosophy. This is a paradox about the role of concepts in the judgment of taste: If the judgment involves concepts, it must be rationally disputable, and provable by reasons (which it is not); if it does not involve concepts, it cannot even be the subject of disagreement (which it is). The solution is that no determinate concept is involved in such judgments, but only the indeterminate concept of the supersensible, or thing-in-itself that underlies the object as well as the judging subject (§§56–57).

KANT ON THE SUBLIME. Kant's analysis of the sublime proceeds on quite different grounds. Essentially, he explains this species of satisfaction as a feeling of the grandeur of reason itself and of humankind's moral destiny, which arises in two ways: (1) When we are confronted in nature with the extremely vast (the mathematical sublime), our imagination falters in the task of comprehending it and we become aware of the supremacy of reason, whose ideas reach toward infinite totality. (2) When we are confronted with the overwhelmingly powerful (the dynamical sublime), the weakness of our empirical selves makes us aware (again by contrast) of our worth as moral beings (see the "Analytic of the Sublime"). In this analysis, and again in his final remarks on beauty in nature, Kant goes some way toward re-establishing on one level a connection between realms whose autonomy he has fought for on a different level. As he had done earlier with the a priori concepts of the understanding and the sphere of morality, he has here tried to show that the aesthetic stands on its own feet, independent of desire and interest, of knowledge or morality. Yet because the experience of beauty depends upon seeing natural objects as though they were somehow the artifacts of a cosmic reason bent on being intelligible to us, and because the experience of the sublime makes use of natural formlessness and fearfulness to cel-

ebrate reason itself, these aesthetic values in the last analysis serve a moral purpose and a moral need, exalting and ennobling the human spirit.

SCHILLER. Kant's aesthetic theories were first made use of by the dramatic poet Friedrich Schiller, who found in them the key to a number of profound problems about culture and freedom that he had been meditating. In several essays and poems, and principally in the remarkable *Briefe über die ästhetische Erzieung des Menschen* ("Letters on the Aesthetic Education of Man," 1793–1795), he developed a neo-Kantian view of art and beauty as the medium through which humanity (and the human individual) advances from a sensuous to a rational, and therefore fully human, stage of existence. Schiller distinguishes (Letters 12–13) two basic drives in man, the sensuous impulse (*Stofftrieb*) and the formal impulse (*Formtrieb*), and argues that they are synthesized and lifted to a higher plane in what he calls the play impulse (*Spieltrieb*), which responds to the living shape (*Lebensform*) or beauty of the world (Letter 15). Play, in his sense, is a more concrete version of Kant's harmony of imagination and understanding; it involves that special combination of freedom and necessity that comes in voluntary submission to rules for the sake of the game. By appealing to the play impulse, and freeing man's higher self from dominance by his sensuous nature, art renders man human and gives him a social character (Letters 26–27); it is therefore the necessary condition of any social order that is based not upon totalitarian compulsion but upon rational freedom.

SCHELLING. Friedrich Wilhelm von Schelling was the first philosopher to claim to have discovered an "absolute standpoint" from which the dualisms and dichotomies of Kant's epistemology could be overcome, or overridden; and he was the first since Plotinus to make art and beauty the capstone of a system. In his *System of Transcendental Idealism* (1800), he attempted a reconciliation of all oppositions between the self and nature through the idea of art. In the artistic intuition, he says, the self is both conscious and unconscious at once; there is both deliberation, *Kunst*, and inspiration, *Poesie*. This harmony of freedom and necessity crystallizes and makes manifest the underlying harmony that exists between the self and nature. There is at work an unseen creative drive that is, on the unconsciousness level, the same as conscious artistic activity. In Schelling's lectures on the *Philosophy of Art* (given 1802–1803, but not published until 1859), transcendental idealism becomes "absolute idealism" and art becomes the medium through which the infinite "ideas," which are the expressions of the various "potencies" involved in the ultimate absolute self-identity, become embodied in finite form, and therefore the medium through which the absolute is most fully revealed. This same general position underlies the famous work *Über das Verhältniss der bildenden Künste zu der Natur (On the Relation Between the Plastic Arts and Nature*, 1807).

HEGEL. The most fully articulated idealistic system of aesthetics was that of George Friedrich Wilhelm Hegel, in his lectures between 1820 and 1829, the notes for which were published (1835) as his *Philosophy of Fine Art*. In art, he says, the "idea" (the notion at its highest stage of dialectical development) becomes embodied in sensuous form. This is beauty. Man thereby renders explicit to himself what he is and can be (see *Philosophy of Fine Art*, Osmaston translation, I, 41). When the sensuous is spiritualized in art (I, 53), there is both a cognitive revelation of truth, and also a reinvigoration of the beholder. Natural beauty is capable of embodying the idea to some degree, but in human art the highest embodiment takes place (see I, 39, 10–11, 208–214).

Hegel also worked out, in great detail, a theory of the dialectical development of art in the history of human culture, from Oriental "symbolic" art, in which the idea is overwhelmed by the medium; through its antithesis, classical art, in which the idea and the medium are in perfect equilibrium; to the synthesis, romantic art, in which the idea dominates the medium and spiritualization is complete (see Vols. III, IV). These categories were to prove very influential in nineteenth-century German aesthetic thought, in which the Hegelian tradition was dominant, despite attacks by the "formalists" (such as J. F. Herbart), who rejected the analysis of beauty in terms of ideas as an overintellectualization of the aesthetic and a slighting of the formal conditions of beauty.

ROMANTICISM

Without attempting to trace its roots and early stages, we may say that the romantic revolution in feeling and taste was fully under way in Schelling's philosophy of nature and in the new forms of literary creation explored by the German and English poets from about 1890 to 1910. From the start, these developments were accompanied by reflection on the nature of the arts themselves, and they led in time to fundamental changes in prevailing views about the arts.

EMOTIONAL EXPRESSION. The romantics generally conceived of art as essentially the expression of the artist's personal emotions. This view is central to such basic doc-

uments as Wordsworth's 1800 Preface to *Lyrical Ballads*, Shelley's *Defense of Poetry* (written 1819) Mill's "What is Poetry?" (1833), and the writings of the German and French romantics. The poet himself, his personality as seen through the "window" of the poem (Carlyle's term in "The Hero as Poet," 1841), becomes the center of interest, and sincerity (in Wordsworth, Carlyle, Arnold) becomes one of the leading criteria of criticism.

IMAGINATION. A new version of the cognitive view of art becomes dominant in the concept of the imagination as a faculty of immediate insight into truth, distinct from, and perhaps superior to, reason and understanding—the artist's special gift. The imagination is both creator and revealer of nature and what lies behind it—a romanticized version of Kant's transcendental idealism, ascribing the form of experience to the shaping power of the mind, and of Fichte's Ego "positing" the non-Ego. A. W. Schlegel, Blake, Shelley, Hazlitt, Baudelaire, and many others spoke of the imagination in these terms. Coleridge, with his famous distinction between imagination and fancy, provided one of the fullest formulations: The fancy is a "mode of memory," operating associatively to recombine the elementary data of sense; the imagination is the "coadunating faculty" that dissolves and transforms the data and creates novelty and emergent quality. The distinction (based on Schelling) between the "primary" and "secondary" imagination is between the unconscious creativity involved both in natural processes and in all perception and the conscious and deliberate expression of this in the artist's creating (see Chs. 13 and 14 of Coleridge's *Biographia Literaria*, 1817). Through most of Coleridge's work there runs his unfinished task of supplying a new theory of mind and of artistic creation that would replace the current associationism, which he had at first enthusiastically adopted and then, under the influence of Plotinus and the German idealists, came to reject.

ORGANISM. Another important, and related, aspect of Coleridge's critical theory was his distinction (derived essentially from A. W. Schlegel's Vienna *Lectures on Dramatic Art*, 1809–1811) between mechanical and organic form and his conception of a work of art as an organic whole, bound together by deeper and more subtle unity than that explicated in the neoclassic rules and having a vitality that grows from within (see his Shakespearean criticism for examples). The concept of nature as organic, and of art as growing out of nature like a living being, had already been developed by Johann Gottfried Herder (see, for example, his *Vom Erkennen und Empfinden der Menschlichen Seele*, 1778), and by Goethe, in some of his essays (e.g., "Vom Deutscher Baukunst," 1772; "Über Wahrheit und Wahrscheinlichkeit der Kunstwerke," 1797).

SYMBOLISM. The idea of the work of art as being, in some sense (in some one of many possible senses), a symbol, a sensuous embodiment of a spiritual meaning, though old in essence, as we have seen, came into a new prominence in the romantic period. Goethe distinguished allegory, a mechanical combination of universal and particular, and symbol, as a concrete unity (see "Über die Gegenstände der bildenden Kunst," 1797); and Friedrich and August Wilhelm Schlegel followed with a new interest in myth and metaphor in poetry. The English Romantic poets (notably Wordsworth) evolved a new lyric poetry in which the visible landscape took on the attributes of human experience. And in France, later in the century, the symbolist movement, launched by Jean Moréas in 1885, and the practice of such poets as Baudelaire, Rimbaud, and Mallarmé emphasized concrete symbolic objects as the heart of poetry.

SCHOPENHAUER. Though first written in the climate of post-Kantian idealism, and, in that context, largely ignored, Arthur Schopenhauer's *Die Welt als Wille und Vorstellung* ("World as Will and Idea," 1819; 2d ed. enlarged, 1844) came into its deserved fame in the second half of the century. Its romantic pessimism and intuitionism and, more particularly, the central position it assigned to the arts (especially music) made it one of the most important aesthetic documents of the century. Schopenhauer's solution of the basic Kantian dualism was to interpret the thing in itself, or noumenal world, as the "Will to Live" and the phenomenal world as the objectification, or expression, of that primal will. The objects of the phenomenal world fall into a hierarchy of types, or grades, that embody, according to Schopenhauer, certain universals or Platonic ideas, and it is these ideas that are presented to us for contemplation by works of art. Since the idea is timeless, the contemplation of it (as, for example, some general character of human nature in a poem or painting) frees us from subjection to the "principle of sufficient reason," which dominates our ordinary practical and cognitive consciousness, and hence from the constant pressure of the will. In this "pure will-less state," we lose individuality and pain.

Schopenhauer has much to say about the various arts and the forms of ideas suited to them; the uniqueness of music in this scheme is that it embodies not ideas but the will itself in its striving and urging and enables us to contemplate its awfulness directly, without involvement.

Schopenhauer's theory of music was one of his most important contributions to aesthetic theory and influenced not only those theorists, such as Richard Wagner (see his essay on Beethoven, 1870), who emphasized the representative character of music, but also those critical of this view, such as Eduard Hanslick in *Vom Musikalisch-Schönen* ("The Beautiful in Music," 1854).

NIETZSCHE. Friedrich Nietzsche repudiated romantic art as escapist, but his own aesthetic views, briefly sketched in the notes published posthumously as *The Will to Power* (1901), are best understood in relation to those of Schopenhauer. Nietzsche's early work, *The Birth of Tragedy from the Spirit of Music* (1872), presented a theory of tragedy as arising from the conjunction of two fundamental impulses, which Nietzsche called the Dionysian and Apollonian spirits: the one a joyful acceptance of experience, the other a need for order and proportion. In Nietzsche's later thinking about art, it is the former that becomes dominant; he insists, for example, as opposed to Schopenhauer, that tragedy exists not to inculcate resignation and a Buddhist negation of life, by showing the inevitability of suffering, but to affirm life in all its pain, to express the artist's overabundance of will to power. Art, he says, is a "tonic," a great "yea-sayer" to life.

THE ARTIST AND SOCIETY

Political, economic, and social changes in the nineteenth century, in the wake of the French Revolution and the rise of modern industry, raised in a new form the Platonic problem of the artists' relation to their society, their possibly conflicting obligations to their craft and to their fellow human beings. In the nineteenth century, an important part of aesthetic thinking was concerned with this problem.

ART FOR ART'S SAKE. One solution to the problem was to think of the artist as a person with a calling of his own, whose whole, or at least primary, obligation is to perfect his work, especially its formal beauty, whatever society may expect. Perhaps the artist, because of his superiority, or higher sensitivity, or the demands of his art, must be alienated from society, and, though perhaps doomed to be destroyed by it, can carry his curse as a pride. This notion stems from the German romantics, from Wilhelm Wacken-Roder, Johann Ludwig Tieck, and others. From 1820–1830 it became the doctrine of "art for art's sake," the center of continuing controversy in France and, later, in England. In its extreme forms, as reflected, for example, in Oscar Wilde (*Intentions*, 1891) and J. A. M. Whistler ("Ten O'Clock" lecture, 1885), it was sometimes a claim that art is more important than anything else and sometimes a flaunting of the artist's freedom from responsibility. More thoughtfully and fundamentally, as in Théophile Gautier (Preface to *Mademoiselle de Maupin*, 1835) and throughout Flaubert's correspondence with Louise Colet and others, *l'art pour l'art* was a declaration of artistic independence and a kind of professional code of dedication. In that respect, it owed much to the work of Kant in carving out an autonomous domain for art.

REALISM. The theory of realism (or, in Zola's sense, naturalism) arose as a broadened conviction of the cognitive duty of literature, a desire to give it an empirical, and even experimental status (in Zola's essay on "The Experimental Novel," 1880), as exhibitor of human nature and social conditions. In Flaubert and Zola, realism called for the cool, analytical eye of the novelist, treating virtue and vice, in Hippolyte Taine's words, as "products like vitriol and sugar"; see the Introduction to his *History of English Literature* (1863), in which Taine set forth his program for explaining art deterministically in terms of race, context, and epoch (*race, milieu, moment*). Among the Russian literary theorists, Vissarion G. Belinsky, Nikolai G. Chernyshevski ("The Aesthetic Relation of Art to Reality," 1855), and Dmitri I. Pisarev ("The Destruction of Aesthetics," 1865), all art was given a similar treatment—as a reproduction of factual reality (sometimes an aid in explaining it, which may have value as a substitute, like a photograph, says Chernyshevski) or as the bearer of social ideas (Pisarev).

SOCIAL RESPONSIBILITY. The theory that art is primarily a social force and that the artist has a social responsibility was first fully worked out by the French socialist sociologists. Claude Saint-Simon (*Du système industriel*, 1821), Auguste Comte (*Discours sur l'ensemble du positivisme*, 1848, Ch. 5), Charles Fourier (*Cités ouvrières*, 1849), and Pierre Joseph Proudhon (*Du principe de l'art et de sa destination sociale*, 1865) attacked the idea that art can be an end in itself and projected visions of future social orders free of violence and exploitation, in which beauty and use would be fruitfully combined and for which art will help prepare. In England, John Ruskin and William Morris were the great critics of Victorian society from an aesthetic point of view. They pointed to the degradation of the worker into a machine, unfree to express himself, the loss of good taste, the destruction of natural beauty, and the trivialization of art. Ruskin's essay on "The Nature of Gothic" (*Stones of Venice*, 1851) and many other lectures (for example those in *The Two Paths*,

1859; *Lectures on Art*, 1870) insisted on the social conditions and effects of art. Morris, in his lectures and pamphlets (see, for example, "Art under Plutocracy," 1883; "The Aims of Art," 1887; "Art and Socialism," 1884), argued that radical changes were needed in the social and economic order to make art what it should be: "… the expression of man's happiness in his labor … made by the people, and for the people, as a happiness to the maker and the user" ("The Art of the People," 1879).

The functionalist tendencies of Ruskin and Morris also turned up, even earlier, in the United States, in the trenchant views of Horatio Greenough ("American Architecture," 1843) and in some essays of Ralph Waldo Emerson ("Thoughts on Art," 1841; "Beauty," *Conduct of Life*, 1860; "Art," *Essays, First Series*, 1841).

TOLSTOY. It was, however, Leo Tolstoy who drove the social view of art to its farthest point in the nineteenth century and issued the most fundamental challenge to art's right to exist. In *What Is Art?* (first uncensored edition, 1898, in English), he asked whether all the social costs of art could be rationally justified. If, as he argued, art is essentially a form of communication—the transmission of emotion—then certain consequences can be deduced. Unless the emotion is one that can actually be shared by men in general—is simple and human—there is either bad art or pseudo art: this criterion rules out most of the supposedly great works of music and literature, including Tolstoy's own major novels. A work must be judged, in the end, by the highest religious criteria of the age; and in Tolstoy's age that meant, he said, its contribution to the sense of human brotherhood. Great art is that which transmits either simple feelings, drawing men together, or the feeling of brotherhood itself (*Uncle Tom's Cabin*). In no other way can it claim genuine social value (apart from the adventitious value of jewelry, etc.); and where it falls short of this high task (as it usually does), it can only be a social evil, dividing people into cliques by catering to sensuality, pride, and patriotism.

CONTEMPORARY DEVELOPMENTS

Aesthetics has never been so actively and diversely cultivated as in the twentieth century. Certain major figures and certain lines of work stand out.

METAPHYSICAL THEORIES. Though he later proposed two important changes in his central doctrine of intuition, the early aesthetic theory of Benedetto Croce has remained the most pervasively influential aesthetics of the twentieth century. The fullest exposition was given in the *Estetica come scienza dell'espressione e linguistica generale* ("Aesthetic as Science of Expression and General Linguistic," 1902), which is part of his *Filosofia dello spirito*. Aesthetics, in this context, is the "science" of images, or intuitive knowledge, as logic is knowledge of concepts—both being distinguished from "practical knowledge." At the lower limit of consciousness, says Croce, are raw sense data, or "impressions," which, when they clarify themselves, are intuïtions, are also said to be "expressed." To express, in this subjective sense, apart from any external physical activity, is to create art. Hence, his celebrated formula, "intuition = expression," on which many principles of his aesthetics are based. For example, he argued that in artistic failure, or "unsuccessful expression," the trouble is not that a fully formed intuition has not been fully expressed but that an impression has not been fully intuited. R. G. Collingwood, in his *Principles of Art* (1938), has extended and clarified Croce's basic point of view.

The theory of intuition presented by Henri Bergson is quite different but has also been eagerly accepted by many aestheticians. In his view, it is intuition (or instinct become self-conscious) that enables us to penetrate to the *durée*, or *élan vital*—the ultimate reality which our "spatializing" intellects inevitably distort. The general view is explained in his "Introduction à la métaphysique" (1903) and in *L'évolution créatrice* (1907) and applied with great ingenuity and subtlety to the problem of the comic in *Le rire* (1900).

NATURALISM. Philosophers working within the tradition of American naturalism, or contextualism, have emphasized the continuity of the aesthetic with the rest of life and culture. George Santayana, for example, in his *Reason in Art* (1903; Vol. IV of *The Life of Reason*), argues against a sharp separation of "fine" from "useful" arts and gives a strong justification of fine art as both a model and an essential constituent of the life of reason. His earlier book, *The Sense of Beauty* (1896), was an essay in introspective psychology that did much to restimulate an empirical approach to art through its famous doctrine that beauty is "objectified pleasure."

The fullest and most vigorous expression of naturalistic aesthetics is *Art as Experience* (1934), by John Dewey. In *Experience and Nature* (1925), Dewey had already begun to reflect upon the "consummatory" aspect of experience (as well as the instrumental aspects, which had previously occupied most of his attention) and had treated art as the "culmination of nature," to which scientific discovery is a handmaiden (see Ch. 9). *Art as Experi-*

ence, a book that has had incalculable influence on contemporary aesthetic thinking, develops this basic point of view. When experience rounds itself off into more or less complete and coherent strands of doing and undergoing, we have, he says, "*an* experience"; and such an experience is aesthetic to the degree in which attention is fixed on pervasive quality. Art is expression, in the sense that in expressive objects there is a "fusion" of "meaning" in the present quality; ends and means, separated for practical purposes, are reunited, to produce not only experience enjoyable in itself but, at its best, a celebration and commemoration of qualities ideal to the culture or society in which the art plays its part.

A number of other writers have worked with valuable results along similar lines, for example, D. W. Prall, *Aesthetic Judgment* (1929) and *Aesthetic Analysis* (1936); C. I. Lewis, *An Analysis of Knowledge and Valuation* (1946, Chs. 14, 15); and Stephen C. Pepper, *Aesthetic Quality* (1937), *The Basis of Criticism in the Arts* (1945), *The Work of Art* (1955).

SEMIOTIC APPROACHES. Since semiotics in a broad sense has undoubtedly been one of the central preoccupations of contemporary philosophy, as well as many other fields of thought, it is to be expected that philosophers working along this line would consider applying their results to the problems of aesthetics. The pioneering work of C. K. Ogden and I. A. Richards, *The Meaning of Meaning* (1923), stressed the authors' distinction between the "referential" and the "emotive" function of language. And they suggested two aesthetic implications that were widely followed: first, that the long-sought distinction between poetic and scientific discourse was to be found here, poetry being considered essentially emotive language; second, that judgments of beauty and other judgments of aesthetic value could be construed as purely emotive. This work, and later books of Richards, have been joined by a number of aesthetic studies in the general theory of (artistic) interpretation, for example, John Hospers, *Meaning and Truth in the Arts* (1946); Charles L. Stevenson, "Interpretation and Evaluation in Aesthetics" (1950); Morris Weitz, *Philosophy of the Arts* (1950); and Isabel C. Hungerland, *Poetic Discourse* (1958).

Meanwhile, anthropological interest in classical and primitive mythology, which became scientific in the nineteenth century, led to another semiotical way of looking at art, particularly literature. Under the influence of Sir James G. Frazer's *The Golden Bough* (1890–1915), a group of British classical scholars developed new theories about the relations between Greek tragedy, Greek mythology, and religious rite. Jane Ellen Harrison's *Themis: A Study of the Social Origins of Greek Religion* (1912) argued that Greek myth and drama grew out of ritual. This field of inquiry was further opened up, or out, by C. G. Jung, in his paper "On the Relation of Analytical Psychology to Poetic Art" (1922; see *Contributions to Analytical Psychology*, 1928) and in other works. Jung suggested that the basic symbolic elements of all literature are "primordial images" or "archetypes" that emerge from the "collective unconscious" of man. In recent years the search for "archetypal patterns" in all literature, to help explain its power, has been carried on by many critics and has become an accepted part of literary criticism.

The most ambitious attempt to bring together these and other lines of inquiry to make a general theory of human culture ("philosophical anthropology") is that of Ernst Cassirer. In his *Philosophie der Symbolischen Formen* (3 vols., 1923, 1925, 1929), the central doctrines of which are also explained in *Sprache und Mythos* (1925) and in *An Essay on Man* (1944), he put forward a neo-Kantian theory of the great "symbolic forms" of culture—language, myth, art, religion, and science. In this view, man's world is determined, in fundamental ways, by the very symbolic forms in which he represents it to himself; so, for example, the primitive world of myth is necessarily different from that of science or art. Cassirer's philosophy exerted a strong influence upon two American philosophers especially: Wilbur Marshall Urban (*Language and Reality*, 1939) argued that "aesthetic symbols" are "insight symbols" of a specially revelatory sort; and Susanne K. Langer has developed in detail a theory of art as a "presentational symbol," or "semblance." In *Philosophy in a New Key* (1942), she argued that music is not self-expression or evocation but symbolizes the morphology of human sentience and hence articulates the emotional life of man. In *Feeling and Form* (1953) and in various essays (*Problems of Art*, 1957), she applied the theory to various basic arts.

Charles W. Morris presented a closely parallel view in 1939, in two articles that (like Mrs. Langer's books) have been much discussed: "Esthetics and the Theory of Signs" (*Journal of Unified Science* [*Erkenntnis*], VIII, 1939–1940) and "Science, Art and Technology" (*Kenyon Review*, I, 1939; see also *Signs, Language and Behavior*, 1946). Taking a term from Charles Peirce, he treats works of art as "iconic signs" (i.e., signs that signify a property in virtue of exhibiting it) of "value properties" (e.g., regional properties like the menacing, the sublime, the gay).

MARXISM–LENINISM. The philosophy of dialectical materialism formulated by Karl Marx and Friedrich Engels contained, at the start, only the basic principle of an aesthetics, whose implications have been drawn out and developed by Marxist theoreticians over more than half a century. This principle is that art, like all higher activities, belongs to the cultural "superstructure" and is determined by sociohistorical conditions, especially economic conditions. From this it is argued that a connection can always be traced—and must be traced, for full understanding—between a work of art and its sociohistorical matrix. In some sense, art is a "reflection of social reality," but the exact nature and limits of this sense has remained one of the fundamental and persistent problems of Marxist aesthetics. Marx himself, in his *Contribution to the Critique of Political Economy* (1859), pointed out that there is no simple one-to-one correspondence between the character of a society and its art.

In the period before the October Revolution of 1917, Georgi V. Plekhanov (*Art and Social Life*, 1912) developed dialectical materialist aesthetics through attacks on the doctrine of art for art and the separation of artist from society, either in theory or in practice. After the Revolution, there ensued a period of vigorous and free debate in Russia among various groups of Marxists and others (e.g., the formalists, see below). It was questioned whether art can be understood entirely in sociohistorical terms or has its own "peculiar laws" (as Trotsky remarked in *Literature and Revolution*, 1924) and whether art is primarily a weapon in the class struggle or a resultant whose reformation awaits the full realization of a socialist society. The debate was closed in Russia by official fiat, when the party established control over the arts at the First All-Union Congress of Soviet Writers (1934). Socialist realism, as a theory of what art ought to be and as a guide to practice, was given a stricter definition by Andrei Zhdanov, who along with Gorki became the official theoretician of art. But the central idea had already been stated by Engels (letter to Margaret Harkness, April 1888): the artist is to reveal the moving social forces and portray his characters as expressions of these forces (this is what the Marxist means by a "typical" character), and in so doing he is to forward the revolutionary developments themselves. (See also Ralph Fox, *The Novel and the People*, 1937; Christopher Caudwell, *Illusion and Reality*, 1937, and other works.)

Indications of recent growth in dialectical materialist aesthetics, and of a resumption of the dialogue with other systems, can be seen in the important work of the Hungarian Marxist Georg Lukács (see, for example, *The Meaning of Contemporary Realism*, translated, 1962, from *Wider den missverstandenen Realismus*, 1958) and in the writings of the Polish Marxist, Stefan Morawski (see "Vicissitudes in the Theory of Socialist Realism," *Diogenes*, 1962).

PHENOMENOLOGY AND EXISTENTIALISM. Among many critics and critical theorists, there has been, in the twentieth century, a strong emphasis on the autonomy of the work of art, its objective qualities as an object in itself, independent of both its creator and its perceivers. This attitude was forcefully stated by Eduard Hanslick in *The Beautiful in Music* (1854); it was reflected in the work of Clive Bell (*Art*, 1914) and Roger Fry (*Vision and Design*, 1920); and it appeared especially in two literary movements. The first, Russian "formalism" (also present in Poland and Czechoslovakia), flourished from 1915 until suppressed about 1930. Its leaders were Roman Jakobson, Victor Shklovsky, Boris Eichenbaum, and Boris Tomashevsky (*Theory of Literature*, 1925). The second, American and British "New Criticism," was inaugurated by I. A. Richards (*Practical Criticism*, 1929), William Empson (*Seven Types of Ambiguity*, 1930), and others (see René Wellek and Austin Warren, *Theory of Literature*, 1949).

This emphasis on the autonomy of the work of art has been supported by Gestalt psychology, with its emphasis on the phenomenal objectivity of Gestalt qualities, and also phenomenology, the philosophical movement first developed by Edmund Husserl. Two outstanding works in phenomenological aesthetics have appeared. Working on Husserl's foundations, Roman Ingarden (*Das Literarische Kunstwerk*, 1930) has studied the mode of existence of the literary work as an intentional object and has distinguished four "strata" in literature: sound, meaning, the "world of the work," and its "schematized aspects," or implicit perspectives. Mikel Dufrenne (*Phenomenologie de l'expérience esthétique*, 2 vols., 1953), closer to the phenomenology of Maurice Merleau-Ponty and Jean-Paul Sartre, has analyzed the differences between aesthetic objects and other things in the world. He finds that the basic difference lies in the "expressed world" of each aesthetic object, its own personality, which combines the "being in itself" (*en-soi*) of a presentation with the "being for itself" (*pour-soi*) of consciousness and contains measureless depths that speak to the depths of ourselves as persons.

The "existential phenomenalism" of Heidegger and Sartre suggests possibilities for an existentialist philosophy of art, in the central concept of "authentic existence," which art might be said to further. These possibilities

have only begun to be worked out, for example, in Heidegger's paper "Der Ursprung des Kunstwerkes" (in *Holzwege*, 1950) and in a recent book by Arturo B. Fallico, *Art and Existentialism* (1962).

EMPIRICISM. The contemporary empiricist makes a cardinal point of attacking the traditional problems of philosophy by resolving them into two distinct types of questions: questions about matters of fact, to be answered by empirical science (and, in the case of aesthetics, psychology in particular), and questions about concepts and methods, to be answered by philosophical analysis.

Some empiricists emphasize the first type of question and have called for a "scientific aesthetics" to state aesthetic problems in such a way that the results of psychological inquiry can be brought to bear upon them. Max Dessoir, Charles Lalo, Étienne Souriau, and (in America) Thomas Munro have formulated this program (see, especially, Munro's *Scientific Method in Philosophy*, 1928, and later essays). The actual results of work in psychology, over the period since Fechner inaugurated experimental aesthetics (*Vorschule der Ästhetik*, 1876) to replace "aesthetics from above" by an "aesthetics from below," are too varied to summarize easily (see Bibliography). But two lines of inquiry have had an important effect on the way in which twentieth-century philosophers think about art. The first is Gestalt psychology, whose studies of perceptual phenomena and the laws of Gestalt perception have illuminated the nature and value of form in art (see, for example, Kurt Koffka's "Problems in the Psychology of Art," in *Art: A Bryn Mawr Symposium*, 1940; Rudolf Arnheim, *Art and Visual Perception*, 1954; Leonard Meyer, *Emotion and Meaning in Music*, 1956). The second is Freudian psychology, beginning with Freud's interpretation of Hamlet (*Interpretation of Dreams*, 1900) and his studies of Leonardo (1910) and Dostoyevsky (1928), which have illuminated the nature of art creation and appreciation. Description of aesthetic experience, in terms of concepts like "empathy" (Theodor Lipps), "psychical distance" (Edward Bullough), and "synaesthesis" (I. A. Richards), has also been investigated by introspective methods.

Analytical aesthetics, in both its "reconstructionist" and "ordinary language" forms, is more recent. This school considers the task of philosophical aesthetics to consist in the analysis of the language and reasoning of critics (including all talk about art), to clarify language, to resolve puzzles due to misapprehensions about language, and to understand its special functions, methods, and justifications (see M. C. Beardsley, *Aesthetics: Problems in the Philosophy of Criticism*, 1958; Jerome Stolnitz, *Aesthetics and Philosophy of Art Criticism*, 1960; William Elton, ed., *Aesthetics and Language*, 1954; Joseph Margolis, ed., *Philosophy Looks at the Arts*, 1962).

See also Addison, Joseph; Aesthetic Qualities; Albert the Great; Alembert, Jean Le Rond d'; Analysis, Philosophical; Aristotle; Arnold, Matthew; Art, Value in; Augustine, St.; Baumgarten, Alexander Gottlieb; Beauty; Belinskii, Vissarion Grigor'evich; Bergson, Henri; Blake, William; Burke, Edmund; Carlyle, Thomas; Cartesianism; Cassirer, Ernst; Chernyshevskii, Nikolai Gavrilovich; Chrysippus; Cicero, Marcus Tullius; Cleanthes; Coleridge, Samuel Taylor; Collingwood, Robin George; Comte, Auguste; Croce, Benedetto; Descartes, René; Dewey, John; Dialectical Materialism; Diderot, Denis; Dostoevsky, Fyodor Mikhailovich; Emerson, Ralph Waldo; Empiricism; Engels, Friedrich; Enlightenment; Epicureanism and the Epicurean School; Epicurus; Erigena, John Scotus; Existentialism; Fechner, Gustav Theodor; Fichte, Johann Gottlieb; Ficino, Marsilio; Fourier, François Marie Charles; Freud, Sigmund; Gestalt Theory; Goethe, Johann Wolfgang von; Greek Academy; Grosseteste, Robert; Hazlitt, William; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Herbart, Johann Friedrich; Herder, Johann Gottfried; Hobbes, Thomas; Home, Henry; Homer; Hume, David; Husserl, Edmund; Hutcheson, Francis; Idealism; Imagination; Johnson, Samuel; Jung, Carl Gustav; Kant, Immanuel; Koffka, Kurt; Langer, Susanne K.; Leibniz, Gottfried Wilhelm; Lessing, Gotthold Ephraim; Leucippus and Democritus; Lewis, Clarence Irving; Locke, John; Lukács, Georg; Marx, Karl; Marxist Philosophy; Merleau-Ponty, Maurice; Mill, John Stuart; Naturalism; Neo-Kantianism; Neoplatonism; Nietzsche, Friedrich; Origen; Parmenides of Elea; Peirce, Charles Sanders; Phenomenology; Philodemus; Philo Judaeus; Pisarev, Dmitri Ivanovich; Plato; Platonism and the Platonic Tradition; Plekhanov, Georgii Valentinovich; Plotinus; Pope, Alexander; Proudhon, Pierre-Joseph; Pythagoras and Pythagoreanism; Rationalism; Realism; Reid, Thomas; Renaissance; Romanticism; Ruskin, John; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Santayana, George; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von; Schiller, Friedrich; Schlegel, Friedrich von; Schopenhauer, Arthur; Sextus Empiricus; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Shelley, Percy Bysshe; Skepticism; Socrates; Stevenson, Charles L.; Stoicism; Taine, Hippolyte-Adolphe; Theophrastus; Thomas Aquinas, St.; Tolstoy, Lev (Leo) Nikolaevich; Wilde, Oscar Fingal O'Flahertie Wills; Xenophanes of Colophon; Zeno of Citium.

Bibliography

CLASSICAL GREEK PHILOSOPHERS

See J. W. H. Atkins, *Literary Criticism in Antiquity*, Vol. 1 (Cambridge, 1934), Chs. 1, 2; G. F. Else, "'Imitation' in the Fifth Century," in *Classical Philology*, Vol. 53, No. 2 (1958), 73–90; T. B. L. Webster, "Greek Theories of Art and Literature down to 400 BC" in *Classical Quarterly*, Vol. 33, Nos. 3, 4 (1939), 166–179; Alice Sperduti, "The Divine Nature of Poetry in Antiquity," in *Transactions and Proceedings, American Philological Association*, Vol. 81 (1950), 209–240.

Plato

See Raphael Demos, *The Philosophy of Plato* (New York, 1939), Chs. 11–13; Rupert C. Lodge, *Plato's Theory of Art* (London, 1953), G. M. A. Grube, "Plato's Theory of Beauty," in *Monist*, Vol. 37, No. 2 (1927), 269–288; J. Tate, "Plato and 'Imitation,'" in *Classical Quarterly*, Vol. 26, Nos. 3, 4 (1932), 161–169.

Aristotle

See G. F. Else, *Aristotle's Poetics: The Argument* (Cambridge, 1957); S. H. Butcher, *Aristotle's Theory of Poetry and Fine Art*, 4th ed. (London, 1923); Ingram Bywater, *Aristotle on the Art of Poetry* (Oxford, 1909); G. F. Else, "Aristotle on the Beauty of Tragedy," in *Harvard Studies in Classical Philology*, Vol. 49 (1938); Roman Ingarden, "A Marginal Commentary on Aristotle's Poetics," in *Journal of Aesthetics and Art Criticism*, Vol. 20, No. 2 (Winter, 1961), 163–173; *Ibid.*, No. 3 (Spring, 1962), 273–285; Richard McKeon, "Literary Criticism and the Concept of Imitation in Antiquity," in R. S. Crane, ed., *Critics and Criticism* (Chicago, 1952).

LATER CLASSICAL PHILOSOPHERS

See Phillip De Lacy, "Stoic Views of Poetry," in *American Journal of Philology*, Vol. 49, No. 275 (1948), 241–271; L. P. Wilkinson, "Philodemus and Poetry," in *Greece and Rome*, Vol. 2, No. 6 (1932–1933), 144–151; L. P. Wilkinson, "Philodemus on *Ethos* in Music," in *Classical Quarterly*, Vol. 32, Nos. 3, 4 (1938), 174–181; Craig La Drière, "Horace and the Theory of Imitation," in *American Journal of Philology*, Vol. 60, No. 239 (1939), 288–300; Philippus V. Pistorius, *Plotinus and Neoplatonism* (Cambridge, U.K., 1952), Ch. 7.

THE MIDDLE AGES

See Edgar de Bruyne, *Etudes d'esthétique médiévale*, 3 vols. (Brugge, 1946); Edgar de Bruyne, "Esthétique païenne, esthétique chrétienne," in *Revue Internationale de Philosophie*, No. 31 (1955), 130–144; Emmanuel Chapman, *Saint Augustine's Philosophy of Beauty* (New York, 1939); K. Svoboda, *L'esthétique de Saint Augustin et ses sources* (Brno, 1933); Jacques Maritain, *Art and Scholasticism*, translated by J. F. Scanlan (London, 1930), esp. Ch. 5; Maurice de Wulf, *Études historiques sur l'esthétique de S. Thomas d'Aquin* (Louvain, 1896); Leonard Callahan, *A Theory of Esthetic According to the Principles of St. Thomas Aquinas* (Washington, DC, 1927); Bernard F. Huppé, *Doctrine and Poetry* (Albany, NY, 1959), Chs. 1, 2; H. Flanders Dunbar, *Symbolism in Medieval Thought* (New Haven, CT, 1929); Murray Wright Bundy, *The Theory of Imagination in Classical and Medieval Thought*, in University of Illinois Studies in Language and Literature (Urbana, 1927), Chs. 8–12.

THE RENAISSANCE

See Nesca A. Robb, *Neoplatonism of the Italian Renaissance* (London, 1935), Ch. 7; Erwin Panofsky, *Idea*, 2nd ed. (Berlin, 1960); Erwin Panofsky, *The Codex Huygens and Leonardo da Vinci's Art Theory* (London, 1940); Erwin Panofsky, *The Life and Art of Albrecht Dürer*, 4th ed. (Princeton, NJ, 1955), Ch. 8; Rensselaer W. Lee, "*Ut Pictura Poesis*: The Humanistic Theory of Painting," in *Art Bulletin*, Vol. 22, No. 4 (1940), 197–269; Anthony Blunt, *Artistic Theory in Italy, 1450–1600* (Oxford, 1940); Edward E. Lowinsky, "Music in the Culture of the Renaissance," in *Journal of the History of Ideas*, Vol. 15, No. 4 (1954), 509–553; D. P. Walker, "Musical Humanism in the 16th and Early 17th Centuries," in *Music Review*, Vol. 2, No. 1 (1941), 1–13; *Ibid.*, No. 2, 111–121; *Ibid.*, No. 3, 220–227; *Ibid.*, No. 4, 288–308; Vol. 3, No. 1 (1942), 55–71; Bernard Weinberg, *A History of Literary Criticism in the Italian Renaissance*, 2 vols. (Chicago, 1961); Baxter Hathaway, *The Age of Criticism: The Late Renaissance in Italy* (Ithaca, NY, 1962).

THE ENLIGHTENMENT

Cartesian Rationalism

See Émile Krantz, *Essai sur l'esthétique de Descartes* (Paris, 1882); Brewster Rogerson, "The Art of Painting the Passions," in *Journal of the History of Ideas*, Vol. 14, No. 1 (1953), 68–94; Scott Elledge, "The Background and Development in English Criticism of the Theories of Generality and Particularity," in *PMLA*, Vol. 62, No. 1 (1947), 147–182; A. O. Lovejoy, "'Nature' as Aesthetic Norm," in *Modern Language Notes*, Vol. 42, No. 7 (1927), 444–450; Hoyt Trowbridge, "The Place of Rules in Dryden's Criticism," in *Modern Philology*, Vol. 44, No. 2 (1946–1947), 84–96; Meyer H. Abrams, *The Mirror and the Lamp* (Oxford, 1953), Chs. 1, 2; Samuel H. Monk, *The Sublime*, rev. ed. (Ann Arbor, MI, 1960), Ch. 9; Claude V. Palisca, "Scientific Empiricism in Musical Thought," in H. H. Rhys, ed., *Seventeenth Century Science and the Arts* (Princeton, NJ, 1961); Louis I. Bredvold, "The Tendency toward Platonism in Neo-classical Esthetics," in *ELH: A Journal of English Literary History*, Vol. 1, No. 2 (1934), 91–119; Cicely Davis, "Ut pictura poesis," *Modern Language Review*, Vol. 30, No. 2 (1935), 159–169; P. O. Kristeller, "The Modern System of the Arts," in *Journal of the History of Ideas*, Vol. 12, No. 4 (1951), 496–527; *Ibid.*, Vol. 13, No. 1 (1952), 17–46; Ernst Cassirer, *The Philosophy of the Enlightenment*, translated by Koelln and Pettegrove (Princeton, NJ, 1951), Ch. 7.

Empiricism

See Clarence DeWitt Thorpe, *The Aesthetic Theory of Thomas Hobbes* (Ann Arbor, MI, 1940); Donald F. Bond, "The Neo-classical Psychology of the Imagination," in *ELH: A Journal of English Literary History*, Vol. 4, No. 4 (1937), 245–264; Martin Kallich, "The Association of Ideas and Critical Theory: Hobbes, Locke, and Addison," in *ELH: A Journal of English Literary History*, Vol. 12, No. 4 (1945), 290–315; R. L. Brett, *The Third Earl of Shaftesbury* (London, 1951); Jerome Stolnitz, "On the Significance of Lord Shaftesbury in Modern Aesthetic Theory," in *Philosophical Quarterly*, Vol. 11, No. 43 (1961), 97–113; Clarence D. Thorpe, "Addison and Hutcheson on the Imagination," in *ELH: A Journal of English Literary History*, Vol. 2, No. 3 (1935), 215–234; Martin Kallich, "The Argument against the Association of Ideas in Eighteenth-Century Aesthetics," in *Modern*

Language Quarterly, Vol. 15, No. 2 (1954), 125–136; Margaret Gilman, *The Idea of Poetry in France from Houdar de la Motte to Baudelaire* (Cambridge, MA, 1958), Ch. 2; Wladyslaw Folkierski, *Entre le classicisme et le romantisme* (Krakow and Paris, 1925); Walter Jackson Bate, *From Classic to Romantic* (Cambridge, Mass., 1946); Ernest Lee Tuveson, *The Imagination as a Means to Grace* (Berkeley, CA, 1960); Jerome Stolnitz, "'Beauty': Some Stages in the History of an Idea," in *Journal of the History of Ideas*, Vol. 22, No. 2 (1961) 185–204; Walter John Hippie, Jr., *The Beautiful, The Sublime, and the Picturesque in Eighteenth-Century British Aesthetic Theory* (Carbondale, IL, 1957); Samuel H. Monk, *The Sublime*, rev. ed. (Ann Arbor, 1960).

GERMAN IDEALISM

See H. W. Cassirer, *A Commentary on Kant's Critique of Judgment* (London, 1938); James C. Meredith, *Kant's Critique of Aesthetic Judgement* (Oxford, 1911); Hermann Cohen, *Kant's Begründung der Ästhetik* (Berlin, 1889); Victor Basch, *Essai critique sur l'esthétique de Kant*, 2d ed. (Paris, 1927); Humayun Kabir, *Immanuel Kant on Philosophy in General*, essays on the first Introduction to the *Critique of Judgment* (Calcutta, 1935); G. T. Whitney and D. F. Bowers, eds., *The Heritage of Kant* (Princeton, NJ, 1939); Robert L. Zimmerman, "Kant: The Aesthetic Judgment," in *Journal of Aesthetics and Art Criticism*, Vol. 21, No. 3 (Spring, 1963), 333–344; Frederic Will, *Intelligible Beauty in Aesthetic Thought from Winckelmann to Victor Cousin* (Tubingen, 1958); S. S. Kerry, "The Artist's Intuition in Schiller's Aesthetic Philosophy," in *Publications of the English Goethe Society*, New Series 28 (Leeds, 1959); Elizabeth E. Bohning, "Goethe's and Schiller's Interpretation of Beauty," in *German Quarterly*, Vol. 22, No. 4 1949), 185–194; Jean Gibelin, *L'esthétique de Schelling d'après la philosophie de l'art* (Paris, 1934); E. L. Fackenheim, "Schelling's Philosophy of the Literary Arts," in *Philosophical Quarterly*, Vol. 4, No. 17 (1954), 310–326; H. M. Schueller, "Schelling's Theory of the Metaphysics of Music," *Journal of Aesthetics and Art Criticism*, Vol. 15, No. 4 (June, 1957), 461–476; W. T. Stace, *The Philosophy of Hegel* (London, 1924), Part IV, Third Div., Ch. 1; Israel Knox, *The Aesthetic Theories of Kant, Hegel, and Schopenhauer* (New York, 1936).

ROMANTICISM

See René Wellek, *A History of Modern Criticism: 1750–1950*, Vol. 2 (New Haven, 1955); M. H. Abrams, *The Mirror and the Lamp: Romantic Theory and the Critical Tradition* (Oxford 1953); W. J. Bate, *From Classic to Romantic* (Cambridge, MA, 1949), Chs. 5, 6; Paul Reiff, *Die Ästhetik der deutschen Frühromantik* in University of Illinois Studies in Language and Literature, Vol. 31 (Urbana, 1946); M. Z. Shroder, *Icarus: The Image of the Artist in French Romanticism* (Cambridge, MA, 1961); A. G. Lehmann, *The Symbolist Movement in France, 1885–1895* (Oxford, 1950); Joseph Ciari, *Symbolism from Poe to Mallarmé: The Growth of a Myth* (London, 1956); John M. Bullitt, "Hazlitt and the Romantic Conception of the Imagination," in *Philological Quarterly*, Vol. 24, No. 4 (1945), 343–361; J. B. Baker, *The Sacred River: Coleridge's Theory of the Imagination* (Baton Rouge, LA, 1957); James Benziger, "Organic Unity: Leibniz to Coleridge," in *PMLA*, Vol. 66, No. 2 (1951), 24–48; John Stokes Adams, *The Aesthetics of Pessimism* (Philadelphia, 1940); J. M. Stein, *Richard Wagner and the Synthesis of the Arts* (Detroit, 1960); E. A. Lippman, "The Esthetic Theories of Richard Wagner," in *Musical Quarterly*, Vol. 44 (1958), 209–220.

THE ARTIST AND SOCIETY

See R. F. Egan, "The Genesis of the Theory of 'Art for Art's Sake' in Germany and in England," in *Smith College Studies in Modern Languages*, Vols. 2, 5; Albert Cassagne, *La théorie de l'art pour l'art en France* (Paris, 1906); Irving Singer, "The Aesthetics of 'Art for Art's Sake,'" *Journal of Aesthetics and Art Criticism*, Vol. 12, No. 3 (March, 1954), 343–359; H. A. Needham, *La développement de l'esthétique sociologique en France et en Angleterre au XIXe siècle* (Paris, 1926); Bernard Weinberg, *French Realism: The Critical Reaction 1830–1870* (New York, 1937); H. M. Kallen, *Art and Freedom*, 2 vols. (New York, 1942); René Wellek, "Social and Aesthetic Values in Russian Nineteenth-Century Literary Criticism," in E. J. Simmons, ed., *Continuity and Change in Russian and Soviet Thought* (Cambridge, MA, 1955); F. D. Curtin, "Aesthetics in English Social Reform: Ruskin and his Followers," in Herbert Davis et al., eds., *Nineteenth-Century Studies* (Ithaca, NY, 1940).

CONTEMPORARY DEVELOPMENTS

Metaphysical Theories

See G. N. G. Orsini, *Benedetto Croce: Philosopher of Art and Literary Critic* (Carbondale, IL, 1961); John Hospers, "The Croce–Collingwood Theory of Art," in *Philosophy*, Vol. 31, No. 119 (1956), 291–308; Alan Donegan, "The Croce–Collingwood Theory of Art," in *Philosophy*, Vol. 33, No. 125 (1958), 162–167; T. E. Hulme, "Bergson's Theory of Art," in Herbert Read, ed., *Speculations* (New York and London, 1924); Arthur Szathmary, *The Aesthetic Theory of Bergson* (Cambridge, MA, 1937).

Naturalism

See W. E. Arnett, *Santayana and the Sense of Beauty* (Bloomington, IN, 1955); Jack Kaminsky, "Dewey's Concept of *An* Experience," in *Philosophy and Phenomenological Research*, Vol. 17, No. 3 (March 1957), 316–330; E. A. Shearer, "Dewey's Aesthetic Theory," *Journal of Philosophy*, Vol. 32, Nos. 23, 24 (1935), 617–627, 650–664; Sidney Zink, "The Concept of Continuity in Dewey's Theory of Esthetics," *Philosophical Review*, Vol. 52, No. 4 (1943), 392–400; S. C. Pepper, "The Concept of Fusion in Dewey's Aesthetic Theory," in *Journal of Aesthetics and Art Criticism*, Vol. 12, No. 2 (December, 1953), 169–176.

Semiotics

See Richard Rudner, "On Semiotic Aesthetics," in *Journal of Aesthetics and Art Criticism*, Vol. 10, No. 1 (September, 1951), 67–77; E. G. Ballard, "In Defense of Semiotic Aesthetics," in *Journal of Aesthetics and Art Criticism*, Vol. 12, No. 1 (September, 1953), 38–43; Max Rieser, "The Semantic Theory of Art in America," *Journal of Aesthetics and Art Criticism*, Vol. 15, No. 1 (September, 1956), 12–26.

Marxism-Leninism

See Karl Marx and Friedrich Engels, *Literature and Art: Selections from Their Writings* (New York, 1947); Victor Erlich, "Social and Aesthetic Criteria in Soviet Russian Criticism," R. M. Hankin, "Main Premises of the Communist Party in the Theory of Soviet Literary Controls," and E. J. Simmons, "Review," in Simmons, ed.,

Continuity and Change in Russian and Soviet Thought (Cambridge, MA, 1955); Max Rieser, "The Aesthetic Theory of Socialist Realism," in *Journal of Aesthetics and Art Criticism*, Vol. 16, No. 2 (December, 1957), 237–248; Ernst Fischer, *Von der Notwendigkeit der Kunst* (Dresden, 1959). Translated by Anna Bostock as *The Necessity of Art, A Marxist Approach* (Baltimore, 1963).

Phenomenology and Existentialism

See Herbert Spiegelberg, *The Phenomenological Movement: A Historical Introduction*, 2 vols. (The Hague, 1960); Fritz Kaufmann, "Art and Phenomenology," in Marvin Farber, ed., *Philosophical Essays in Memory of Edmund Husserl* (Cambridge, MA, 1940); Anna-Teresa Tymieniecka, *Phenomenology and Science in Contemporary European Thought* (New York, 1961); J.-Claude Piguet, "Esthétique et Phénomenologie" (discussion review of Dufrenne), in *Kantstudien* (1955–1956), 192–208; E. F. Kaelin, *An Existentialist Aesthetic: The Theories of Sartre and Merleau-Ponty* (Madison, WI, 1962).

Empiricism

See Douglas Morgan, "Psychology and Art: a Summary and Critique," in *Journal of Aesthetics and Art Criticism*, Vol. 9, No. 2 (December, 1950), 81–96; Douglas Morgan, "Creativity Today," in *Journal of Aesthetics and Art Criticism*, Vol. 12, No. 1 (September, 1953), 1–24; A. R. Chandler, *Beauty and Human Nature* (New York, 1934); C. W. Valentine, *The Experimental Psychology of Beauty* (London, 1962); Edward Bullough, in Elizabeth Wilkinson, ed., *Aesthetics: Lectures and Essays* (London, 1957); William Phillips, ed., *Art and Psychoanalysis* (New York, 1957); D. E. Schneider, *The Psychoanalyst and the Artist* (New York, 1950).

Bibliographies

For further bibliographies on contemporary aesthetics, see M. C. Beardsley, *Aesthetics* (New York, 1958); Guido Morpurgo-Tagliabue, *L'esthétique contemporaine* (Milan, 1960).

GENERAL HISTORIES

See the following general histories of aesthetics: Katherine Gilbert and Helmut Kuhn, *A History of Aesthetics* (New York, 1939; Bloomington, IN, 1954); M. C. Beardsley, *Aesthetics from Ancient Greece to the Present* (New York, 1965); Bernard Bosanquet, *A History of Aesthetic* (London, 1892; New York, 1957).

Monroe C. Beardsley (1967)

AESTHETICS, HISTORY OF [ADDENDUM]

TWENTIETH-CENTURY AESTHETICS

Aesthetics continued to be intensively cultivated in all the main schools of twentieth-century philosophy. The following survey emphasizes work that continues to be of interest at the beginning of the twenty-first century. It will focus first on the Anglo-American tradition, including continental work that has fed into it, and then will consider other work in the continental tradition.

ANGLO-AMERICAN AESTHETICS.

Naturalism, organicism, pragmatism. One main line of twentieth-century aesthetics begins with George Santayana's *The Sense of Beauty* of 1896. Santayana's book was a renewal of the empiricism and naturalism of the eighteenth century undertaken in opposition to the incorporation of aesthetics into speculative metaphysics by philosophers such as Schelling, Schopenhauer, and Hegel. Santayana held that beauty is "value positive, intrinsic, and objectified": a pleasurable emotion that is "pure gain" and that we regard as if it were a property of its object even though it depends upon our own response. The idea that beauty is objectified pleasure is found in writers from Hutcheson to Kant, but Santayana departed from the reductionism characteristic of many eighteenth-century authors by refusing to restrict the sources of such pleasure to a single category. He instead showed how such pleasure can arise from the materials of works of art, from their forms, and from their expression, which he defined broadly to include our emotional associations with objects. Santayana also rejected the attempt to justify the human interest in beauty, especially the often costly interest in artistic beauty, by claiming that it contributes to morality; for Santayana, morality is concerned with the removal of the evils of life, and thus exists only to facilitate the wider enjoyment of the positive pleasures of life, epitomized by beauty. In his second main work on aesthetics, *Reason in Art*, the fourth volume of his 1905–1906 *Life of Reason*, Santayana added that by the ability to adopt an aesthetic attitude and thus find beauty almost anywhere in nature, on the one hand, and by the ability to create art, on the other, we can augment our positive pleasure in life. In this work he also emphasized that the various arts have all arisen from the ordinary and natural activities of human beings, thus adding a pragmatist element to his naturalism and preparing the way for the later work of John Dewey.

Santayana's thesis that morality exists to remove the evils that stand in the way of the enjoyment of the positive pleasure of beauty anticipates the famous statement of G. E. Moore's *Principia Ethica* (1903) that "the most valuable things, which we can know or imagine, are … the pleasures of human intercourse and the enjoyment of beautiful objects" (Moore 1903, p.237), which would become the creed of the Bloomsbury group of artists and intellectuals. Moore treated "aesthetic appreciation" as an "organic whole" consisting of consciousness of both the beautiful qualities of an object and the feeling of its beauty, an idea that is related to Santayana's notion of beauty as objectified pleasure; but Moore also held that

beautiful objects are themselves organic unities, in the sense that the contemplation of the individual parts may have no value, but the contemplation of the whole loses value without the contemplation of those parts. Moore thus adopted a more restrictive analysis of the objects of aesthetic pleasure than had Santayana.

Moore influenced the critic Clive Bell, who in his 1914 book *Art* postulated a special aesthetic emotion in response to "significant form" in works of art. Edward Bullough, a professor of literature who in 1907 gave the first course on aesthetics at Cambridge, has also been considered a follower of Moore, but his theory is different from Bell's; according to Bullough's famous 1912 paper "'Psychical Distance' as a Factor in Art and an Aesthetic Principle," distancing oneself from the most obvious emotions that might be aroused by some object, such as the emotion of fear in response to a fog at sea, does not allow one to enjoy some special aesthetic emotion, but rather opens oneself up to a whole range of other feelings and emotions that can be aroused by the very same object, thereby increasing the richness and intensity of one's emotional experience of life as a whole. Instead of being closely associated with Moore and Bell, Bullough might thus be better placed on a line of thought leading from Santayana to Dewey.

Dewey's *Art as Experience* (1934) came late in his lengthy career, but remains his most widely read book as well as one of the still most widely read books of twentieth-century aesthetics. He anticipated its central idea of "consummatory experience" in his 1925 *Experience and Nature*. A consummatory experience is a moment felt as one of repose and equilibrium in the constant flow of energy, in stimulus and response, that constitutes human life, and it is paradigmatically produced by the experience of art. As Dewey put it in 1925, "art is the solvent union of the generic, recurrent, ordered, established phase of nature with its phase that is incomplete, going on, and hence still uncertain, contingent, novel, particular" (Dewey 1925, p.301), or as he said in 1934, "Art is the living and concrete proof that man is capable of restoring consciously, and thus on the plane of meaning, the union of sense, need, impulse and action characteristic of the live creature" (Dewey 1934, p. 25).

But in the later work Dewey also argued that art has a special role in the expression of emotion, not merely projecting our emotions onto objects but clarifying them by presenting the contexts in which they arise. Here Dewey's thought comes into contact with the next stream of aesthetic thought to be considered here, which makes the expression of emotion the core of aesthetic experience. But Dewey's pragmatism reveals itself in his insistence that the aesthetic "is the clarified and intensified development of traits that belong to every normally complete experience," and even more so with his argument that while the term *aesthetic* connotes the "consumer's rather than the producer's standpoint" and the term *art* "denotes a process of doing and making," there is a strong element of each in the other: The audience for art must take an active and imaginative role in appreciating it, while the artist must also adopt the standpoint of his audience to gauge the effect of his work—hence Dewey's title *Art as Experience*, blurring the line between the production and the reception of art (Dewey 1934, p.47). This is a theme that would also be stressed by the British philosopher R.G. Collingwood a few years later, who though not considered a pragmatist came out of a Hegelian background with affinities to that of Dewey.

Before we turn to the tradition with which Collingwood is associated, we may note that Monroe C. Beardsley, the author of the first part of this article, was himself the most important heir to Dewey's aesthetics in the period after World War II. Although there are certainly other influences at work, the central claim of Beardsley's 1958 *Aesthetics* was clearly Deweyan. Beardsley wrote that an experience has a marked "aesthetic character" when it includes "attention firmly fixed on a perceptual or intentional object; a feeling of freedom from concerns about matters outside that object; notable affect that is detached from practical ends; the sense of exercising powers of discovery; and the integration of the self and of its experiences" (Beardsley 1981, p. lxii). The most recent heir to Dewey and Beardsley, Richard Shusterman, has particularly stressed the experience of one's own body as part of the complete aesthetic experience (*Pragmatist Aesthetics*, 1992).

Expression. A second main line of twentieth-century aesthetics identifies the chief goal of art as the expression of emotion, a feature that was only one facet of Dewey's notion of aesthetic experience. This theory is often thought of as an alternative to the idea that beauty is the essence of art, but at least in its early stages the successful expression of emotion in art was intended as an explanation of its beauty. This is evident in the 1892 *History of Æsthetic* by Bernard Bosanquet and in the 1902 work by Benedetto Croce, *Estetica come scienze dell'espressione e linguistica generale* (The aesthetic as the science of expression and of the linguistic in general). Bosanquet argued that art operates "through that expansion of self which comes in utterance," that is, that content acquires beauty by passing through the crucible of an individual sensibil-

ity and style—even though this means that it may take others time to appreciate the beauty of a distinctive style of expression (Bosanquet 1904, p. 453). Croce wrote that the beautiful is "successful expression, or better, … expression *simpliciter*, since expression, when it is not successful, is not expression (Croce 1992, p. 87). Ten years after Croce, the neo-Kantian Hermann Cohen based an elaborate *Ästhetick des reinen Gefühls* (Aesthetics of pure feeling) on the premise that human feelings have their own distinctive forms, which are most clearly revealed by art.

The fullest development of the expression theory, however, is found in the 1938 *Principles of Art* by the Oxford philosopher (and archaeologist) R. G. Collingwood. Collingwood is often thought of as a follower of Croce, but his theory is more fully developed than Croce's, and it also overcomes the supposition that successful expression must be perceived as beautiful in some traditional sense. Collingwood begins by distinguishing art from craft, arguing that in the latter there is always a clear distinction between means and end, but that there is never such a distinction in the case of art proper. This leads to two important claims: that art is never intended merely to arouse emotions for the sake of magic or propaganda or to discharge them for the sake of amusement; and that the element of craft that is typically part of art, namely the production of a physical object, is not essential to the true work of art at all, which thus appears to exist complete in the mind of the artist without any physical expression.

The latter claim, however, is clearly modified over the rest of Collingwood's book. The second part of the book argues that there is an affective or emotional aspect of all perception and thought, and that the special function of art is to clarify that dimension of our experience so that we can understand and gain control over it. In the third part of his work, Collingwood then argues that the clarification of emotion takes place through the artist's interaction with a physical medium and an audience. So Collingwood's initial claim that the work of art exists complete in the mind of the artist turns out to be an overstatement of the claim that the effort in art is aimed at the clarification of emotion rather than at the production of a physical object for its own sake. Writing at a tense moment in the 1930s, Collingwood concludes by stressing that art proper is necessary for the survival of civilization precisely because it allows us to gain control over our own emotions rather than having our emotions controlled by the propaganda of others.

Art and language. Expression theorists such as Croce and Collingwood suggested that all art, whether in verbal media or not, can be regarded as using or creating languages for the expression of emotion. Beginning in the 1930s, many other varieties of aesthetic theory focused on linguistic aspects of the arts and of critical discourse about art. One important movement was logical positivism, represented above all by A. J. Ayer's 1936 *Language, Truth, and Logic*, which argued that aesthetic discourse does not consist of verifiable, descriptive propositions about its objects at all, but only expresses the response of the speaker to such objects, to which a prescriptive rather than descriptive recommendation of the object to others might also be added. This doctrine, which applied to ethical as well as aesthetic discourse, became known as "emotivism" and enjoyed considerable currency after its further development in C. L. Stevenson's *Ethics and Language* (1944). It would become one of the sources for hostility to traditional aesthetic theory during the heyday of "analytical" philosophy in the 1950s and 1960s.

A different strand of thought can be traced back to Ernst Cassirer's *Philosophie der symbolischen Formen*, published in German from 1923 to 1929 and translated into English (*Philosophy of Symbolic Forms*) only in 1953, but preceded by Cassirer's English-language summary of his position, *An Essay on Man* of 1944. Cassirer, a student of Hermann Cohen, held that human beings represent and deal with their environment through a variety of symbolic systems, including natural language, mathematical and scientific language, mythology, and the arts, each of which has its distinctive uses and none of which can simply be subordinated to the others.

Cassirer was a major influence on the American philosopher Susanne K. Langer, who interpreted human thought as using a variety of symbol-systems in her 1942 *Philosophy in a New Key* and dedicated her major work in aesthetics, the 1953 *Feeling and Form*, to the memory of Cassirer. She held that the arts do not employ "discursive" symbol-systems to analyze experience but instead use non-discursive symbols to capture the felt quality of experience itself. Using music as an example, she argued that the symbol-systems of the arts do not use "syntactical terms with fixed connotations, and syntactical rules for deriving complex connections," like ordinary and scientific language, but instead "present emotive experience through global forms that are as indivisible as the elements of chiaroscuro" (Langer 1942, p. 232). Her position thus looks back to Alexander Baumgarten's original distinction between logic and aesthetics, but also looks for-

ward to the 1968 *Languages of Art* of Nelson Goodman, who acknowledged affinities between his own approach and that of Langer as well as of Cassirer, Charles Sanders Pierce, and the semiotician C. W. Morris.

Goodman abjured any interest in the traditional topics of beauty and pleasure in the arts, and instead offered analyses of fictional and metaphorical depiction and of expression within the framework of an austerely nominalistic theory of language. But his affinity with Langer and indeed with Baumgarten became clear when he argued that symbols or uses of language are symptomatic of the aesthetic if they are syntactically and semantically dense rather than discrete, if they are replete, with many features of the symbol contributing to its meaning, and if they exemplify qualities metaphorically as well as literally. And while maintaining his emphasis on the cognitive rather than emotional or affective dimension of aesthetic experience, he also wrote about its dynamic rather than static character, its "restless, searching, testing" attitude, its creation and re-creation, in a way that ultimately makes clear the pleasurable character of the aesthetic form of cognition. At its deepest level, Goodman's aesthetics thus falls within the Kantian tradition.

A third major influence on modern thought about aesthetics and language was of course the philosophy of Ludwig Wittgenstein. Through the influence of his 1921 *Tractatus Logico-Philosophicus* on the so-called Vienna Circle, he was in the background of Ayer's *Language, Truth, and Logic*. In 1938 (thus the same year as the publication of Collingwood's *Principles of Art* in Oxford), he lectured on aesthetics in Cambridge. One central theme of these lectures, presumably directed against such nineteenth-century German psychologists as Hermann von Helmholtz and Gustav Theodor Fechner, was that aesthetics cannot be made into a science causally connecting measurable responses to measurable qualities of objects. Here Wittgenstein was in fact only reminding his auditors of an argument made long before by Hume and Kant. More influential themes of his lectures were, first, that aesthetic discourse does not typically work by using a general predicate like "beautiful" but instead uses more particular words and gestures to focus attention on particular aspects of objects that in their particular context look right or satisfying, and, second, that aesthetic response often involves imaginatively seeing an aspect or interpretation in an object.

Although these lectures were not published until 1967, the first of these themes was influential before that date. Thus Frank Sibley (himself a student of Gilbert Ryle) argued in 1959 that aesthetic concepts are not "condition-governed" but are instead highly context-sensitive; this theme was further developed in Peter Kivy's 1973 *Speaking of Art*. The second theme, which Wittgenstein would develop further in the major work of his late philosophy, the *Philosophical Investigations*, posthumously translated and published in 1953, was carried on in Roger Scruton's *Art and Imagination* (1974) and in Richard Wollheim's theory of "seeing-in" in his A. W. Mellon Lectures on the Fine Arts, *Painting as an Art* (1987).

The greatest influence of the *Philosophical Investigations*, however, came from its view that many concepts, including the concept of language itself, are not defined by a determinate set of necessary and sufficient conditions, but by a looser network of "family resemblances." Wittgenstein argued that a concept like that of games could only be understood in this way, and that the abstraction of "language" likewise consists of a loosely interconnected network of "language-games." In a famous paper of 1956, Morris Weitz argued that this model applied to the arts as well, thus that the concept of art is an "open concept" for which there could be no determinate definition of art of the kind to which traditional aesthetics had aspired. In an equally important paper of 1965, Maurice Mandelbaum replied that a determinate definition of an abstract concept like art is compatible with diversity and constant change at the level of the particular objects of art. This interchange as well as the history of developments in twentieth-century art, from the "readymades" of Marcel Duchamp through Dada to the Pop Art of Andy Warhol and Robert Rauschenberg, launched a debate about the possibility of a definition of art that was a central topic of analytical aesthetics from the 1960s into the 1980s.

In a 1964 paper on "The Artworld," Arthur C. Danto used the cases of artworks that are perceptually indiscernible either from other artworks or from ordinary objects that are not artworks at all to argue that an artwork is never identical to a physical object, but is rather a physical object embedded in a world of artistic theory. In his 1974 *Art and the Aesthetic*, George Dickie was inspired by Danto's concept of the "artworld" to offer a definition of a work of art as an artifact offered as a candidate for appreciation by an agent of the artworld, where he understood the latter in sociological terms as the social system of artists, dealers, curators, critics, and so on.

Danto's 1981 *Transfiguration of the Commonplace* made it clear that this was not what Danto had meant by an artworld, but that by this concept he instead meant the complex of meaning, metaphor, and style within which an artist intended his work to be received, a view that he

has refined in subsequent work, including his 2003 book *The Abuse of Beauty*, into the definition of art as "embodied meaning." Dickie acknowledged this basic difference in the understanding of the concept of an artworld in his 1984 book *The Art Circle: A Theory of Art*, and redefined an artworld as a set of artistic conventions rather than a sociological formation.

Jerrold Levinson and Noël Carroll subsequently developed historicized versions of Dickie's approach, arguing that a work of art is an object made within a historical tradition of art making. But from Danto's point of view, all such appeals to artistic conventions, histories, or traditions are circular without some definition of what makes the latter conventions, histories, or traditions of *art* in the first place. However, in his 1997 *Philosophies of Arts*, Peter Kivy argued against the assumption that all art has semantic meaning, which underlies Danto's definition of art, by appeal to "absolute" music and the decorative arts, which are not "about" anything.

The return of beauty. Danto's earlier work was very much under the influence of Marcel Duchamp's attack upon beauty as a mere "retinal flutter" inessential to the real character of art, and Goodman likewise dismissed beauty from the cognitive core of art. However, not all philosophers have been convinced of the inessentiality of beauty, and two important works of the 1980s offered detailed analyses of beauty while defending its centrality in the experience of art. In *The Test of Time* (1982), Anthony Savile argued that we find an object beautiful when we see it as a successful solution to its underlying problem or problems within its own style, that we are able to recognize a successful solution to a problem even when the problem is not our own, and that being beautiful in this sense, along with being deep—that is, revealing fundamental and general principles— and suggestive about the possibilities for successful forms of human life, is one of the things that enables a work of art to withstand the test of time.

Two years later, Mary Mothersill's *Beauty Restored* reached back to Hume and Kant and beyond them to Thomas Aquinas to argue that beauty is a disposition actualized when a person is pleased by the apprehension of the aesthetic qualities of objects, where the latter are precisely what distinguish an object from all others, and that beauty so understood is central to the ambitions of art. More recently, Alexander Nehamas has interpreted the traditional conception of beauty as a "promise of happiness" (a phrase that comes from Baudelaire) to mean that we find an object beautiful when it draws us into an ongoing engagement with itself and an open-ended network of related objects, and that this is essential to our experience of art, although he emphasizes that these networks are personal and that there is no reason to expect "universal validity" in responses to beauty. Art critics and literary theorists such as Dave Hickey, Elaine Scarry, and Wendy Steiner have also recently defended the importance of beauty in art.

Aesthetics and morality. One of the most significant developments in recent aesthetics is renewed interest in the relations between aesthetic experience and morality, one of the two issues initially raised by Plato's attack upon popular arts in the education of his guardians but one that had been largely neglected during the heyday of "analytical" aesthetics, when indeed traditional modes of theorizing in both aesthetics and ethics were under attack. Both Plato's original attack upon popular arts and contemporary versions thereof have themselves been subjects of recent investigations. Alexander Nehamas has examined parallels between the ancient and modern attacks in papers collected in his *Virtues of Authenticity* (1999), while in *A Philosophy of Mass Art* (1998), Noël Carroll has shown in detail how many forms of "mass" art engage their audiences in ways both cognitive and emotional that are no different from the ways in which "high" arts engage their audiences. This work may be considered as a rejoinder to the critique of the "culture industry" as necessarily a form of mass manipulation that was offered by Max Horkheimer and Theodor W. Adorno in their famous *Dialectic of Enlightenment*, first published in 1947 as *Dialektic der Aufklärung* (see below).

Most of the recent debates about aesthetics and morality, however, have focused on two distinguishable issues. The first concerns the value of the experience of art, especially literature, in moral education. One view here holds that the moral truths expressed in works of art are so obvious and general that there is no need to turn to art to learn them, thus that their role in moral education can hardly be central to the value we place on art. The opposing view concedes that it may be unnecessary to turn to art to learn general moral *principles*, but that we can learn a great deal from narrative art, particularly literature and cinema, about the emotions of both agents and patients in morally significant situations, and indeed that narrative art may well be the primary means by which we learn to be attentive to the details of the kinds of situations in which we will ultimately have to apply our general moral principles. This view has been defended in numerous works by Martha C. Nussbaum and Noël Carroll.

The current debate could be enriched by a return to its roots in the eighteenth century, where Kant recognized that the artistic presentation of examples of virtuous conduct are essential in teaching children not so much the content as the importance of aesthetic principles, while Schiller later argued that aesthetic experience sharpens our sensitivity to both general principles and particular situations in his letters *On the Aesthetic Education of Man* (1967). Another voice that needs to be incorporated into this debate is that of Stanley Cavell, who has argued in both his philosophical work such as *The Claim of Reason* (1979) and his critical work such as his 1969 essay "The Avoidance of Love" on Shakespeare's *King Lear* that a central lesson we can learn from art concerns the epistemology of conduct itself, that is, our need to act upon trust in both ourselves and others in the face of our always imperfect knowledge of self and others rather than being destroyed by fantasies about the perfection of knowledge and love that are beyond human powers.

The other recent debate has been about what has come to be called "ethical criticism" of the arts. Here the issue is whether what may be perceived as ethical defects of works of art, that is, defects in the moral views that may be expressed by works of art, are necessarily also aesthetic defects in those works, or whether our appreciation of the aesthetic merits of a work can be independent of any such ethical defects. The latter position, called "autonomism," has been defended by Daniel Jacobson and others; "moderate moralism," the position that ethical defects are at least *pro tanto* aesthetic defects in a work of art, although they may be outweighed by other aesthetic merits of the work, has been defended by Noël Carroll and Berys Gaut. Carroll has argued that some moral defects may prevent imaginative "uptake" of a work while others may not, that is, that some ethical defects may be sufficient to prevent an audience from identifying with the characters and standpoints of a work in the way necessary for it to accomplish its aesthetic goals, while others may not. The conditions under which "uptake" of a work may be facilitated or blocked would seem to be a subject for psychological investigation, and thus one of the points at which aesthetics can intersect with contemporary cognitive science.

Fictionality. Another area of contemporary debate where aesthetic theory can intersect with cognitive science is the recent discussion of the emotional impact of fictions. This debate too has roots in antiquity, namely the paradox of tragedy. One side of this paradox is related to the issue just discussed, namely, how we can take pleasure in the depiction of events that, were they real, we should surely abhor. But there is also an epistemological and psychological question here, namely, how we can have emotional responses to fictions that are anything like the emotional responses we would have to the depicted events if they were real, when we know that they are not?

In his 1990 book *Mimesis as Make-Believe*, Kendall Walton has argued that we use works of art as props in games of make-believe, that it is fictional rather than actual that we respond to the work with the emotions that the objects they depict would induce in ordinary life, for example, that we respond with fear to events depicted in a horror movie, and therefore that there is no paradox in either how we can like or how we can fear fictions, because we do not in fact have the same emotional responses to fictions that we do to reality. This leads to an interpretation of the experience of fiction as "simulation" that is also investigated in contemporary cognitive science. An alternative position holds that to experience a fiction is like entertaining but not asserting a thought, and that we can have the same emotional response to an unasserted as to an asserted thought. This position has been developed by Noël Carroll and Peter Lamarque, among others. It too seems suitable for investigation by cognitive scientists.

CONTINENTAL AESTHETICS. Just as the division between "analytical" and "continental" aesthetics is less than clear-cut, so any rigid division of the continental tradition into separate lines of development will also be misleading. Nevertheless, the present discussion will be organized around a division between Marxist, phenomenological, and post-structuralist aesthetics.

Marxist aesthetics. Both Marx and Engels included the arts among the cultural superstructure of societies, which is determined by their economic substructure, but neither provided an extended treatment of aesthetics. That awaited twentieth-century Marxism. In the early days of Bolshevism and Russian communism, both Lenin and Trotsky addressed the role of the arts at length. Lenin treated art as a category of "intellectual work" that, like any other form of labor, could be used for or against the revolution. He expected art to serve the political education of the proletariat and therefore to remain accessible through the use of conventional forms.

This line of thought led to the official adoption of the style of "Socialist realism," defined by Andrei Zhdanov, at the First All-Union Congress of Soviet Writers in 1934. By that time, Leon Trotsky, a less conventional thinker, had already been exiled from the Soviet Union.

Trotsky, who had published *Literatura i revoliutsiia* (Literature and revolution) in 1924, also argued that art should serve as a "hammer" for building the new society, but recognized that art also needed to be a "mirror" of existing society in order to reveal what needed to be corrected or rejected in it. Trotsky also kept in mind that the ultimate point of the revolution was supposed to be the extension of the enjoyment of freedom from an elite to the masses, and therefore held that art was not merely instrumental in value, but should enjoy some freedom of its own. In this regard Trotsky actually remained closer to the mainstream of modern Western aesthetics.

Trotsky's recognition that traditional forms of art could be used as a mirror for the flaws of existing society was developed by the Hungarian György Lukács. His first books, *Die Seele und die Formen* (The soul and the forms; 1910) and *The Theory of the Novel* (1916), (Die Theorie des Romans" [1916]), were written in neo-Kantian and Hegelian veins respectively, but after World War I, Lukács became a major communist theorist with *Geschichte und Klassenbewusstsein* (History and class consciousness) in 1923. He then devoted the rest of his career to aesthetics, culminating in his massive and untranslated *Die Eigenart des Ästhetischen* (The uniqueness of the aesthetic) in 1963.

Lukács held that every society is a complex whole in which all aspects of life reflect its underlying economics and politics; that individual psychologies form types that reflect the roles that are possible within their society; and that art, especially the novel, should represent the types of psychologies possible within the society that it depicts. Lukács became hostile to modernists such as Joyce and Kafka, whom he saw as expressing their own, individual psychologies without regard for the larger society of which they were a part. He recognized that all art involves some abstraction, but rejected abstraction as an end in art. This led him into debates with Ernst Bloch and Bertolt Brecht, who held that abstract and unconventional means of presentation might work more effectively than traditional forms of mimesis to expose the contradictions within society and to agitate for change.

A figure who was much less influential when he was alive but who gained prominence in later decades is the literary critic Walter Benjamin. Benjamin failed to make an academic career in the 1920s with his work on the German baroque and romanticism, but had more of an impact with his work on modernist literature and life: He spent much of the last part of his life working on a Marxist-inspired study of modern sensibility through the lens of the twentieth-century shopping mall and its mass-produced goods, his so-called "Arcades Project." Among aestheticians, however, his most influential work was his 1936 essay on "The Work of Art in the Age of its Mechanical Reproducibility," in which he argued that the "aura" of traditional art derived from its original cultic role and then from the restriction of its accessibility to elites, conditions that could not be maintained with contemporary mass arts such as cinema. But Benjamin's essay left it open whether the mass rather than cultic accessibility of modern media makes them instruments for even greater domination, now by commercial rather than religious elites, or creates increased room for individual autonomy in the exercise of taste and choice of pursuits.

The most influential neo-Marxist aesthetician working after World War II was Theodor W. Adorno. Adorno was a student of composition under Alban Berg in Vienna as well as a student of philosophy in Frankfurt, where he became an associate of the "Frankfurt school" of critical theory before the war and eventually, after his return from his wartime exile in Oxford and Los Angeles, its postwar leader. Adorno wrote in many areas, from sociology (he coauthored *The Authoritarian Personality* in 1950), literary criticism (*Noten zur Literatur* [Notes to literature]; 1958–1965), and music theory (*Philosophie der neuen Musik* [Philosophy of modern music]; 1949). With Max Horkheimer, the original director of the Frankfurt school, he coauthored the *Dialektic der Aufklärung* (Dialectic of enlightenment; 1947), which argued that, contrary to its intention, the European Enlightenment was actually an extension of the traditional drive to dominate the individual by mythology, and then that the contemporary "culture industry" continues the mass manipulation of the individual. Horkheimer and Adorno thus disambiguated Benjamin's ambivalent attitude toward modern media in favor of the more pessimistic interpretation.

Adorno's largest works were his *Negativ Dialektic* (Negative dialectics) of 1966 and the posthumous *Ästhetische Theorie* (Aesthetic theory) of 1970. In the latter, a more optimistic work than the *Dialektic der Aufklärung*, Adorno emphasized that even though art is always located in a historical context and therefore "refuses definition," it has always "turned against the status quo and what merely exists just as much as it has come to its aid by giving form to its elements" (Adorno 1997, p.2). On his account, art both reveals the contradictions of existing society and yet can make us aware of the possibility of something better. Art shows both the fissures in current society and the possibility of a non-coercive integration beyond those fissures. In spite of the length of the book,

much of Adorno's view remains programmatic. One distinctive feature often missing from other modern and especially Marxist aesthetics, however, is Adorno's reflection on the relation between artistic and natural beauty: He argues that natural beauty offers a model of integration or reconciliation that is often missing from the man-made, and thus that art often seeks to bring nature within its scope, but at the same time that we may easily be seduced by nature into thinking that reconciliation of social fissures will come automatically instead of by our own, intentional efforts.

The Frankfurt theorist who remained in America, Herbert Marcuse, drew on Freud to criticize orthodox Marxism in much of his late work, first in *Eros and Civilization* (1955), which as its title suggests argued for the necessity of Eros as well as social justice, and then in his last work, *The Aesthetic Dimension* (1978), which argued that art unequivocally reflects the human wish for life rather than death, that "Aesthetic form, autonomy, and truth… each *transcends* the socio-historical arena," and that art "challenges the monopoly of the established reality to determine what is 'real,' and it does so by creating a fictitious world which is nevertheless 'more real than reality itself'" (Marcuse 1978, p. 22). Marcuse's conviction that the resistance to the forces of Eros come primarily from politics rather than from the natural conditions of human life are regarded as naive by contemporary psychoanalysis.

The British literary theorist Terry Eagleton returned to more traditional Marxist-inspired critique of ideology in his 1990 *The Ideology of the Aesthetic*, arguing that the classical modern theory of the autonomy and universal validity of taste, which was developed simultaneously with the bourgeois domination of the economics and politics of European society beginning in the eighteenth century, was in fact a mask for that increasing domination.

The phenomenological tradition. The other main German-influenced line in twentieth-century aesthetics is the phenomenological tradition. This has its sources in both Wilhelm Dilthey and Edmund Husserl. Dilthey was a historian, biographer, and literary critic as well as a theorist of history, the arts, and the human sciences generally. He adopted the idea of hermeneutics from Friedrich Schleiermacher, the subject of one of his major studies, and introduced it into twentieth-century thought. He held that every society and period has a distinctive "worldview" (*Weltanschauung*), but that the modern worldview (since the Renaissance) has grown so complex that it can only be represented artistically, in virtue of art's powers of isolation or abstraction, concentration, and integration. In his view, hermeneutics is the method for interpreting the larger worldview expressed by a work of art.

Husserl, by contrast, started off as a technical philosopher of logic and mathematics, and then argued for a distinctive power of apprehending the essential structures of logical, mathematical, and scientific concepts, of ordinary objects, and of the social world that is independent of ordinary empirical investigation—what he called *Wesensschau*, or the intuition of essence. From such a premise, it would be natural to see art as a form of *Wesensschau*, especially in the pioneering period of abstract art. Husserl himself did not apply his phenomenology to the case of art, but the Pole Roman Ingarden did in *Literarische Kunstwerk* (The literary work of art; 1931). Ingarden employed Husserl's approach in seeing a work of literature (and by extension other works of art) as containing complex layers of intentionality, including meaningful words, meaningful combinations of words and elements, represented objects, and "schematized aspects" or implicit perspectives that need to be developed in the thought of the reader rather than the writer. In this regard Ingarden's works can be seen as a forerunner of the "reception aesthetics" of Wolfgang Iser (*The Act of Reading*, 1978; originally published as *Der Act des Lesens*, 1976) and Hans Robert Jauss (*Aesthetic Experience and Literary Hermeneutics*, 1982; originally published as *Äesthetische Erfahrung und literarische Hermeneutik*, 1977).

Martin Heidegger, however, was influenced by Wilhelm Dilthey as well as by Husserl, and it could be argued that in his case the influence of the former gradually overtook that of the latter: for Heidegger, art reveals more *Weltanschauung* than *Wesensschau*. Heidegger's magnum opus, *Being and Time* (1927) (*Sein und Zeit*) argues for the priority of the human experience (*Dasein*) of the world as an arena for agency with tools and instruments over the objectivist standpoint of science and traditional philosophy, which treats humans as more passive knowers of independent realities. Heidegger did not discuss art in this work, but it has proven tremendously influential on writers from Jean-Paul Sartre and Maurice Merleau-Ponty to the present, all of whom treat art as the special vehicle for the expression of the point of view of *Dasein* rather than objectifying science.

During the 1930s, as the style of his philosophy became more mythic (some would argue that this was a reflection of his allegiance to National Socialism), Heidegger lectured explicitly on aesthetics, culminating in

the essay on *The Origin of the Work of Art* (Der Ursprung des Kunstwerks [1950]) that was written in 1935–1936 but not published until 1950. Here Heidegger describes art as revealing both "world" and "earth," the former the complex of beliefs, practices, and feelings that characterizes a human way of life and the latter the chthonic domain and forces from which the human world emerges. One striking feature of this essay is that it begins by stressing that art is a form of work, thus a product of human activity, but ends with a theory of truth in which truth is *revealed* to the artist who knows chiefly how not to get in its way (an approach that goes back to Schopenhauer's interpretation of genius). The essay thus ends up with a peculiarly passive view of artistic creation and, by implication, reception; it is thus very much opposed to the model of artistic creation and reception to be found in such writers as Collingwood, Dewey, Ingarden, and the "reception" theorists.

Heidegger's most influential student in the arena of aesthetics was Hans-Georg Gadamer, whose major work was *Wahrheit und Methode* (Truth and method; 1960). Gadamer was also influenced directly by Dilthey, and was the major proponent of hermeneutics in the second half of the twentieth century. *Wahrheit und Methode* is a general theory of hermeneutics as the means, though not a formal method, for understanding oneself and others. But it begins with an attack upon traditional aesthetics, especially Kant's, for "subjectivizing" aesthetic experience, or for seeing art as a means to producing an experience in the subject (the audience), which might be shared with other subjects because a *sensus communis* is presupposed, but not as a means for building a *sensus communis* or intersubjective understanding in the first place. Gadamer calls the first, narrowly subjective kind of experience *Erlebnis*, but the fuller experience that is essentially intersubjective *Erfahrung*. On his account, while we always already understand ourselves and others from within some conceptual framework, art is a fundamental means for us to revise and expand our understanding of self and others, and thus to make the transition from *Erlebnis* to *Erfahrung*.

Jean-Paul Sartre never wrote a treatise on aesthetics, although a large part of his enormous oeuvre consists of books and essays on particular artists such as Flaubert, Baudelaire, Jean Genet, and Mallarmé. His early work *Imaginaire; psychologie* (The psychology of imagination; 1940), under the influence of Husserl, stressed the role of forming images in imagining, although his own creative as well as critical output was in the field of literature rather than the visual arts. Like Adorno and Marcuse later, he stressed the potential of art for non-alienating communication, in which the artist's expression of freedom invites the audience to experience their own freedom of imagination as well. In this regard, his view falls into the Kantian rather than Heideggerian tradition. Merleau-Ponty's emphasis on the primacy of perception over scientific understanding in his main work, *Phénoménologie de la perception* (The phenomenology of perception; 1945), is certainly influenced by the Heidegger of *Being and Time* as well as by Husserl's late emphasis on the *Lebenswelt* ("lifeworld"), but his three seminal essays on aesthetics, especially "Cézanne's Doubt," also stress the freedom and individuality of artistic vision.

Post-structuralism. Main voices in the French "post-structuralist" or "post-modernist" movement, which has had its primary influence on literary theory rather than philosophical aesthetics, include among others Roland Barthes, Michel Foucault, Jacques Derrida, Paul de Man, Jean-François Lyotard, and the sociologist Pierre Bourdieu. (An independent form of "post-modernism," and the term itself, originated within the precincts of architecture and architectural theory, beginning with the works of Robert Venturi and his widely read books *Complexity and Contradiction in Architecture* [1966] and *Learning from Las Vegas* [1972].)

With the exception of de Man, a literary critic who published only a few volumes of papers, all of these authors published a flood of works on a wide range of topics, and aesthetics in the traditional sense concerns only a small number of their works. Barthes's works in aesthetics touch on topics including criticism, fashion, and photography (*Camera Lucida*, 1980). Foucault's works, beginning with *Mots et les choses* (1966), translated as *The Order of Things* (1970), focused on the "archaeology of knowledge," an historicist analysis of cognitive systems and of the power relations underlying such systems. But both Barthes and Foucault were known particularly for the thesis of the "death of the author," which held that it is primarily the reader (or auditor or viewer) rather than the author who constitutes the meaning of a work of art, and thus of course that works of art do not have determinate meanings, since they may have many different audiences. This is an extreme version of what had been one aspect of aesthetic theory since Dewey and Collingwood, and a hyperbolic statement of the reception aesthetics developed in Germany by Iser and Jauss.

Derrida and de Man were the leaders of the movement called "deconstructionism," which pervaded literary studies in the last decades of the twentieth century. The central idea of this approach is that a text never has a

"transcendental" meaning outside of itself, but only an endless deferral of meaning from one sign to another both within itself and intertextually, that is, in other texts but not in some reality beyond texts altogether, and further that texts, especially philosophical texts, are typically built upon unsustainable distinctions that inevitably collapse. A classic example of deconstructionist analysis is Derrida's argument in *La vérité en peinture* (1978), translated as *The Truth in Painting* (1987), that Kant's distinction between a painting and its frame (its *parerga*) collapses because sometimes one cannot tell the difference between the painting and the frame, and therefore the distinction between art and non-art collapses as well. This is a misreading of Kant, who introduced the concept of the *parerga* only to show that even in the frame or drapery around a work of art we respond primarily to formal properties, as with the work itself, and not to define the difference between art and non-art; and it ignores the fact that in the vast majority of cases we can perfectly well tell the difference between the painting and its frame, even if in a small number of cases we cannot. Most of our empirical concepts have a penumbra of borderline cases, and yet we successfully use them in all sorts of contexts.

Lyotard extended Derrida's attack on the determinacy of language by arguing that figural or visual imagery often brings us closer to our real desires (*Discours, figure* [Discourse, figure]; 1971), and among other works also published lectures on Kant's concept of the sublime (1991) that manifest deconstructionism's fascination with the sublime as purported evidence of the ultimate ineffability of meaning. Finally, in his widely influential *Distinction: Critique sociale du jugement* (Distinction: A social critique of the judgment of taste; 1979), Bourdieu argued against the existence of any universal validity in matters of taste, thereby rejecting the ambitions of traditional aesthetics.

As the creation and reception of art in many different forms remain fundamental features of human life in many different cultures throughout the world, it can be expected that aesthetics will remain a central branch of philosophy in the twenty-first century as in centuries past.

See also Adorno, Theodor Wiesengrund; Aesthetic Experience; Ayer, Alfred Jules; Barthes, Roland; Baumgarten, Alexander Gottlieb; Beardsley, Monroe C.; Benjamin, Walter; Bloch, Ernst; Bosanquet, Bernard; Cassirer, Ernst; Cavell, Stanley; Cohen, Hermann; Collingwood, Robin George; Croce, Benedetto; Danto, Arthur; Derrida, Jacques; Dewey, John; Dilthey, Wilhelm; Engels, Friedrich; Enlightenment; Fechner, Gustav Theodor; Foucault, Michel; Gadamer, Hans-Georg; Goodman, Nelson; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heidegger, Martin; Helmholtz, Hermann von; Hermeneutics; Horkheimer, Max; Hume, David; Husserl, Edmund; Hutcheson, Francis; Ingarden, Roman; Kafka, Franz; Kant, Immanuel; Langer, Susanne K.; Lenin, Vladimir Il'ich; Lukács, Georg; Lyotard, Jean-François; Marx, Karl; Merleau-Ponty, Maurice; Moore, George Edward; Neo-Kantianism; Nussbaum, Martha; Peirce, Charles Sanders; Plato; Ryle, Gilbert; Santayana, George; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von; Schiller, Friedrich; Schleiermacher, Friedrich Daniel Ernst; Schopenhauer, Arthur; Stevenson, Charles L. ; Structuralism and Post-structuralism; Thomas Aquinas, St.; Wittgenstein, Ludwig Josef Johann.

Bibliography

Adorno, Theodore W. *Aesthetic Theory* (1970), edited by Gretel Adorno and Rolf Tiedemann. Translated by Robert Hullot-Kenter. Minneapolis: University of Minnesota Press, 1997.

Beardsley, Monroe C. *Aesthetics: Problems in the Philosophy of Criticism*. 2nd ed. Indianapolis: Hackett, 1981. 1st ed., New York: Harcourt, Brace & World, 1958.

Bosanquet, Bernard. *A History of Æsthetic*. 2nd ed. London: George Allen & Unwin, 1904.

Croce, Benedetto. *The Aesthetic as the Science of Expression and of the Linguistic in General*. Translated by Colin Lyas. Cambridge, U.K.: Cambridge University Press, 1992.

Dewey, John. *Art as Experience*. New York: G. P. Putnam's Sons, 1934.

Dewey, John. *Experience and Nature* (1925). 2nd ed., Chicago and LaSalle Open Court, 1929.

Langer, Suzanne K. *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*. Cambridge, MA: Harvard University Press, 1942.

Marcuse, Herbert. *The Aesthetic Dimension: Toward a Critique of Marxist Aesthetics*. Boston: Beacon Press, 1978.

Moore G[eorge] E[dward]. *Principia Ethica* (1903). 2nd ed., with other papers, edited by Thomas Baldwin. Cambridge, U.K.: Cambridge University Press, 1993.

Paul Guyer (2005)

AESTHETICS, PROBLEMS OF

The philosophical discipline of aesthetics deals with conceptual problems arising out of the critical examination of art and the aesthetic. Monroe Beardsley subtitled his 1958 book on general aesthetics *Problems in the Philosophy of Criticism*, implying that aesthetics is about philo-

sophical concepts that are used—often unthinkingly—by critics of the arts, when they say that a work of art such as a painting is *beautiful* or has *aesthetic* value, that it *represents* some subject matter, has a well-composed *form*, is in a particular *style*, and *expresses* some emotion. But aesthetics also deals more broadly with the aesthetics of nature (Budd 1996, Carlson 2000) and gardens (Ross 1998), and with the aesthetic appreciation of objects and activities in everyday life (Dewey 1934). And even when focused on the arts, philosophical aesthetics is concerned with the philosophical problems that arise from the artist's point of view as well as the critic's. Thus creativity, expression, representation, form, and style are problems that can be addressed from the artist's point of view as well as the spectator's. Moreover, "the philosophy of criticism" does not do justice to the breadth of concerns addressed by philosophical aesthetics today. Some of the thorniest issues in aesthetics relate directly to problems in general philosophy: What is aesthetic value? Do the arts provide knowledge? Is there a special kind of aesthetic experience or aesthetic perception?

Most of the questions that come up in theorizing about particular art forms—the philosophy of literature, the theory of the visual arts, the philosophy of music, the philosophy of film, environmental arts and so on—are general questions having implications for other art forms. Some theorists, however, think that the individual arts come with their own unique sets of philosophical problems (Kivy 1997). The problem of the experience and value of absolute music, for example, does not have a clear parallel in any of the other arts, including the other abstract arts (Kivy 1990). Authenticity is a particular problem in the performing arts such as dance and music. But for the most part, questions in the philosophy of art have general application across the arts. Thus the problem of the nature of fictional characters has usually been taken to be a problem about literature, but representational works of visual art also contain fictional people, objects and events (Walton 1990). Similarly, the question as to why people get emotionally involved with fictional characters may seem to be unique to films and novels (Carroll 1990, Currie 1990, Feagin 1996, Lamarque 1996), but it applies equally to fictions in works of visual art. Again, the question why people enjoy tragedies is not peculiar to tragedies: It is the same kind of question as the question why do people listen to sad music if it makes them feel sad (Davies 1994, Levinson 1990)?

This brief overview first discusses the aesthetic in general and then turns to problems peculiar to the arts. It ends with some general remarks about how aesthetics connects to more general questions about knowledge, emotion, and value. Some effort has been made to point out how the most important concepts of aesthetics came to be considered important. The tendency of late-twentieth-century philosophy—especially analytic philosophy—has been to treat the problems of aesthetics as timeless problems having correct answers that will be true of all art works and aesthetic experiences no matter where or when they occur. But if one approaches aesthetics with an eye to the historical background from which its characteristic problems emerged, one will have a better sense not only of what those problems are but also of the different ways they have been conceptualized and why.

THE AESTHETIC

What is the realm of the aesthetic? Should it be thought of as a special kind of pleasure, or, more broadly, as a special kind of experience, as a special type of judgment, as a special type of attitude toward the world, or as a special type of quality? All these options have been pursued. The term "aesthetics" derives from the Greek word *aesthesis*, meaning "perception." The German rationalist philosopher Alexander Baumgarten coined the term in 1735 to mean the science of "sensory perception," which was designed to contrast with logic, the science of "intellect" (Baumgarten 1954), and ever since, the term "aesthetic" has kept its connotation as having an essential connection to the perceptually discriminable.

Although German rationalism gave the field of aesthetics its name and a rationale, it was the British empiricists who established aesthetics as a philosophical discipline and who set the agenda for its subsequent development. The problem that chiefly exercised the eighteenth century thinkers in aesthetics was the nature of aesthetic pleasure and of aesthetic judgment, the judgment of "taste." If aesthetics were to be a serious philosophical discipline, then presumably there must be principles that would justify aesthetic judgments, and distinguish them from mere assertions of liking or disliking. At the same time it was taken for granted by the empiricists that aesthetic judgments depend on subjective feelings of pleasure. For Hutcheson (1973), Hume, and their successors, the aesthetic judgment was primarily a judgment that something is beautiful. So the challenge was to figure out if there was a special kind of pleasure that was the proper response to beauty or a special kind of judgment that was being made when one judged an object beautiful.

BEAUTY

The concept of beauty was an heirloom of ancient and Medieval philosophy. For Plato (1953), only the Idea of Beauty is really beautiful, since everything else is only beautiful in one respect or at one time rather than another or by comparison with one thing and not another. Beautiful people and things can only approach the Form of Beauty. The Medievals, under the influence of the Neoplatonist Plotinus, thought of beauty, the good, and other perfections, as true in the strictest sense only of the highest level of reality. Christianity echoed this idea in the doctrine that beauty is one of God's perfections. In this framework the beauty of the world is derivative from "an image and reflection of Ideal Beauty" (Eco 1986, p. 17). Augustine, for example, believed that a person possesses beauty of body or soul only to the degree that he or she approximates God's perfect beauty. Such a conception of beauty is a far cry from the way it has come to be thought about in modern aesthetics.

Since the Enlightenment, beauty has by and large no longer been regarded as having or being an ethical or religious value. Instead, the eighteenth-century empiricists thought of it simply as the capacity of an object to produce a particular kind of pleasurable experience. The judgment that something is beautiful was the paradigm of what they called the aesthetic judgment or judgment of taste. If, however, the judgment that something is beautiful is not to be a mere statement of liking or preference, then there must be a standard of taste, a principle of justification for claims that something is beautiful which nevertheless preserves the insight that judgments of the beautiful are based on subjective feelings of pleasure. It is this formulation of the problem of beauty and the aesthetic that has come down to us and which continues to exercise theorists.

THE AESTHETIC JUDGMENT

The empiricists rejected the idea that there are universal standards of beauty: The great variety of beautiful things suggests that there are no general canons or rules of beauty as assumed by some classical writers in the Renaissance. Hutcheson thought that the classical idea of "unity in variety" is the one property that reliably evokes aesthetic pleasure (Hutcheson 1973), but whether something has the right degree of unity or variety is itself problematic. Hume famously solved the dilemma by arguing that we are all so constituted as to be pleased by the same sorts of objects in nature and works of art but that we do not all have the same background of experience, delicacy of taste, good sense, ability to make comparisons and lack of prejudice that we ideally could and should have (Hume 1985). Those who have these abilities in the highest degree are the "ideal critics" to whom the rest of us should defer about what is beautiful, and in theory these ideal critics will all agree with one another. Even Hume himself, however, suspected that this would not do entirely, pointing out that younger people have different tastes from older, and that people from one culture might take no pleasure in the art of another if the values it assumes and promotes are sufficiently alien. Today, Marxist critics, reader-response theorists and feminist critics have all emphasized the difficulty of generalizing about the responses of perceptive critics with different background assumptions and points of view.

KANT AND FORMALISM

After Hume, Kant (2000) gave an equally famous *a priori* argument that judgments of taste, though based on subjective feelings of pleasure, lay claim to universality because the pleasure in question is neither pleasure in the sensuously pleasing nor pleasure in the useful, but a disinterested pleasure that arises from the harmonious free play of imagination and understanding, which are cognitive faculties common to all rational human beings. Since it derives from these shared abilities, this pleasure is itself shareable and communicable. Kant thought that an aesthetic judgment is disinterested because it is not addressed to anything in which we have an interest or personal stake but instead is a judgment about the form of an object. The object of aesthetic judgment is "purposiveness without purpose," the appearance something has of having being harmoniously put together for some end even though it lacks any specific end. Kant's examples of aesthetic judgment are drawn primarily from the beauties of nature such as the shape and sweetness of the rose, but his ideas were influential in fixing attention on the formal aspects of art works as well. Kant himself emphasized the role of art works in producing "aesthetic ideas," but critics who focus exclusively on the early part of the *Critique of Judgment* have found there a justification for the view that with respect to both nature and art, the aesthetic judgment or judgment of taste is directed exclusively to formal qualities. This idea no doubt ultimately derives from the classical notion that measure and symmetry are important or even definitive of beauty.

At any rate, Kant has, perhaps unjustly, been seen as the main source of formalism, the idea that the most or only important features of a work of art are its formal qualities. To twentieth-century critics of painting such as Clive Bell and Clement Greenberg, this means that only

color, line, and shape, and their inter-relations are of aesthetic importance and that content is aesthetically irrelevant. In music it is the doctrine that only structure is important. In literature, formalists have emphasized the structures of plots in narratives and the use of imagery and other rhetorical devices in poetry. There is something to be said for formalism—it draws people's attention to what is truly artistic in a work of art, the "art" with which it is put together—but it assumes a distinction between form and content that is very difficult—perhaps impossible—to make out.

Bell (1914) thought art could be defined as "significant form," suggesting that two paintings can imitate or represent the very same thing—the Virgin, say, or a field full of cows—yet one can be art and the other not, because of the way the artist has rendered the form of the work. Bell was part of the Art for Art's Sake movement that swept England in the late nineteenth and early twentieth century. The emphasis on form is congenial to critics of the abstract arts such as architecture and instrumental music, but it is far less plausible for such arts as literature and photography. Moreover, as has often been pointed out, Bell seems to be defining good art rather than art *simpliciter*, and in defining good art, he is attributing to it his own favored criterion of value.

AESTHETIC QUALITIES, AESTHETIC EXPERIENCE, AESTHETIC ATTITUDE

In the early eighteenth century the paradigm of an aesthetic judgment was taken to be the judgment that something is beautiful; and the beautiful was explained in terms of pleasure. In the later part of the century, however, the notion of aesthetic judgment was expanded to include judgments of the picturesque and the sublime, but the judgment of the sublime is no longer wholly pleasurable. Burke described the source of the feeling of the sublime as "whatever is fitted in any sort to excite the ideas of pain and danger" such as vastness, power and obscurity (Burke 1909, p. 36).

Once aesthetic judgments were no longer directed solely at the beautiful, the way was clear for thinking of the aesthetic not as one particular kind of pleasure or as one particular kind of judgment, but rather as a certain kind of quality of an object. Beauty and sublimity might then be merely two among a much broader class of aesthetic qualities, such as "dainty," "garish," "delicate," "insipid," and so on. One question raised by expanding the range of aesthetic qualities is whether all aesthetic qualities are correctly describable as formal qualities. Frank Sibley, who initiated the modern discussion of aesthetic qualities, includes on his list of examples not only clear-cut examples of formal qualities, such as "graceful" and "garish," but also qualities such as "melancholy," which are usually thought of as *expressive* properties, a special subset of aesthetic qualities (Sibley 1959).

Interestingly, very similar questions arise in connection with aesthetic qualities as formerly arose about beauty: Are they intrinsic or mind-dependent qualities? And if they are mind-dependent, do they behave like colors which are perceived similarly by everybody with properly functioning eyes, or are they more like the taste of curry or cilantro, which is perceived as delicious and piquant by some and disgusting by others? Is there a set of ideal critics, as Hume proposed, whose faculties are keener than those of the rest of us and who should be the true judges of aesthetic qualities? These are questions that are still being hotly debated.

The notion of a special aesthetic pleasure or aesthetic perception has also broadened since the eighteenth century into the more general concept of *aesthetic experience*. John Dewey is partly responsible for this change in emphasis. He wanted to stress the importance of having "experiences" in daily life that have the same wholeness, richness, and sense of integration that are characteristic of our encounters with works of art. Other theorists (for example, Schopenhauer [1958] and Stolnitz [1960]) have insisted that what marks out the aesthetic is a special kind of attitude, that should be taken to works of art but that can in theory be taken to anything whatsoever. It turns out that the aesthetic attitude has many of the features of an aesthetic judgment: It is a special kind of disinterested contemplation, often taking the form of an object or art work as the focus of attention.

THE THEORY OF THE ARTS: IMITATION AND REPRESENTATION

The idea that poetry and painting are arts of imitation derives from Plato, who likened imitations to shadows and reflections, and as such, he thought, led away from rather than toward the truth. Aristotle, too, thought that the arts of poetry and painting were imitations of reality, but, unlike Plato, he thought that we learn from imitations and that we take pleasure in doing so. Plato and Aristotle were the first in the western tradition to theorize about poetry and painting as arts of imitation, but they did not think of them as a special category of "fine arts" or Art with a capital "A." The Ancient Greeks had no concept of "the aesthetic" (Sparshott 1982). The arts of painting and sculpture were varieties of technē or craft. The word "art" derives from the Latinized form of the Greek

technē, meaning a "corpus of knowledge and skills organized for the production of changes of a specific kind in matter of a specific kind," like the arts of cobbling or leatherwork (Sparshott 1982, p. 26). The art of poetry had a more important educational role as a source of moral education but it too is an art of imitation. In the Renaissance and the Enlightenment under the influence of Aristotle and his descendants in the classical period, it became a commonplace that poems and paintings imitated or represented the world.

The first attempt to systematize the fine arts came in 1746 when the abbé Batteux grouped together poetry, painting, sculpture, dance and music under the rubric of the imitation of beautiful nature. This was a revolutionary idea in that it categorized together craftsmen such as sculptors and painters with the more highly educated poets, and implied that all the practitioners of the fine arts provided representations of the world that were potential sources of knowledge (Kristeller 1951–1952). Once the idea of the fine arts was established, it was possible to search for traits that they all have in common, and the search for a definition of the fine arts and eventually of "Art" was born.

From the beginning, the search for a definition has been challenged by the multiplicity of the arts. Thus the idea that the arts imitate or represent beautiful nature may have seemed plausible in the age of Pheidias and Praxiteles who made realistic but highly idealized sculptures of the human body, and similarly in the High Renaissance when the beautiful paintings of Raphael and Leonardo imitated the beautiful female form in their paintings of the Virgin, but the arts of "pure" music and dance are not obviously imitating anything. Architecture, too, is only exceptionally an art of imitation. In the eighteenth-century synthesis of the fine arts as arts of the imitation of beautiful nature, we see an attempt to fit together two different conceptual traditions, on the one hand the new empiricist concern with aesthetic judgment, the judgment of beauty, and on the other hand the classical idea—derived from Plato and Aristotle—that the fine arts are arts of imitation. Although buildings, dances and music do not fit very well under the description of arts of imitation, they can certainly be beautiful by satisfying the formal demand for "unity within variety." We see here the beginnings of a clash which lasts to our own day, roughly speaking, the clash between thinking of the arts as aspiring to beauty of form or as seeking to show us the way things are in the world.

The idea that the arts are all arts of imitation has seemed more and more far-fetched in the contemporary world, where a tendency toward abstraction is the rule in the visual arts, and even literature has drawn attention to its formal aspects rather than the story it tells. Perhaps in some very broad sense the arts are "about" the world, but even this has been denied by some defenders of "absolute music" who see it rather as a means of escape from the world (Kivy 1990).

At the same time the notion of "imitation" has come under attack as an account of representation. Many works of art, such as representational paintings, photographs, films, and sculptures represent the world, but it does not seem right to say that they *imitate* it. The role of convention and style is too important in all these genres to make a comparison with a mirror image plausible. Widely discussed theories of pictorial representation include Ernst Gombrich's view that the history of realistic painting is a history of "making and matching" (Gombrich 1960), and Richard Wollheim's theory that pictorial representation rests upon a prior capacity people have for "seeing in" (Wollheim 1987). In literature, a distinction has been made between literary narratives that talk about the world in some sense but arguably do not *represent* it and literary dramas that do represent the world, but perhaps not in quite the same sense that pictures do. Kendall Walton thinks that representations in general should be analyzed in terms of the concept of what a work prescribes us to imagine (Walton 1990). When, for example, we encounter a pictorial representation of a water mill, we imagine of our act of seeing that it is a seeing of a water-mill. His controversial theory of photography holds that, in contrast to paintings, we do not merely imagine but *really see* the object photographed that appears in the picture (Walton 1984).

EXPRESSION

In the Romantic period, artists and writers began to describe themselves not as merely imitating an inert reality but as expressing their own emotional perspectives on the world. Poetry, wrote Wordsworth in a famous phrase, is the "spontaneous overflow of powerful feelings" that are "recollected in tranquility" (Wordsworth 1963, p. 260). After the imitation theory, the next great attempt to define Art was the theory of art as expression. Kant had stressed the role of imagination in art, and the role of the genius that "gives the rule to art" (Kant 2000, p. 187), i.e. who makes up his own rules rather than obeys conventional canons. The Platonic notion of the craftsman who knew how to craft sculptures or poems and who was creative only insofar as he was inspired by the gods, gave way to the idea of the artist who used his creative imagination

to come up with novel expressions of novel ideas and emotions.

Kant's notion that the mark of genius is to come up with "aesthetic ideas" was taken up by Hegel, who argued that art is one of the modes of consciousness whereby man reaches knowledge of Absolute Spirit; specifically it is that mode of consciousness whereby ideas are embodied in some sensuous form. For Hegel, then, art was an important means to knowledge, but it was a special kind of knowledge that could not be detached from the medium in which it is conveyed. The theorists of expression, including the idealist R. G. Collingwood and the pragmatist John Dewey, echoed some of these ideas, insisting that artistic expression is a *cognitive* activity, a matter of elucidating and articulating emotions (Collingwood 1938, Dewey 1934). Like Hegel, they seemed to think that the emotional attitude embodied in a poem or painting was unique to that poem or painting: Any change in color or line in a painting, any change in imagery or rhythm in a poem would change the emotion expressed. Some theorists stressed not so much personal expression as the communication of emotion from one person to another (Tolstoy 1960).

Just as the definition of art as the imitation of reality fits well with eighteenth-century poems and paintings, so the theory of art as expression fits best with Romantic and Expressionist poetry, music, sculpture and painting. Once again architecture is a problem: Most buildings do not seem to express the personal emotions and attitudes of their makers.

The concept of expression has proved malleable, however. More recent theories include Goodman's view that expression is metaphorical exemplification (Goodman 1976). In this sense a work of architecture can express some of its aesthetic properties, its gracefulness, its minatory look, its wit, and it can literally exemplify its mass, its solidity, and perhaps its style. Likewise, a piece of music can metaphorically exemplify its melancholy or jovial character. Other theorists have argued that expression is nothing but the *possession* of a certain sort of aesthetic property (Hospers 1954–1955), namely expressive properties such as "melancholy," "jovial," "witty," and "lively," and have disputed about whether these properties are possessed metaphorically or literally (Davies 1994). In this discussion, too, we see a clash of different conceptual traditions. The idea that art is expression is far removed from the notion that art has a special set of aesthetic properties called "expressive" properties.

The idea that art has expressive properties is not a very surprising revelation, but it does have the advantage of being true across a wide range of art works. By contrast the Romantic, idealist theory of art as expression fits poorly with most of the works made before the end of the eighteenth century. And although twentieth century modernist artists thought of themselves as "embodying" ideas and emotions in a medium just as Collingwood recommended, in the postmodern world artists seem to want to convey their ideas by any means possible rather than "embodying" them in a carefully constructed work of Collingwoodian expression. At the same time, however, many artists continue to talk about *expressing* themselves in their work.

THE INSTITUTIONAL THEORY OF ART

The imitation theory, the theory of art as form, and the expression theory all seem incapable of providing a definition of art that covers all those things that people in Western societies generally want to count as art. Consequently, some have despaired of the possibility of defining art at all, and have retreated to the position that "art" is a "family resemblance" concept in Wittgenstein's sense (Weitz 1956). The more popular move, however, has been to look for a definition which does not appeal to "exhibited" properties such as the form of a work, its representational content or its expressive qualities, but rather to historical or contextual features of the work. Arthur Danto has proposed that we count something as art if there is an artistic theory behind it that links it to the history of art (Danto 1964, 1981). Just as the theory of art as imitation had its origins in the classical world and the theory of art as expression in the Romantic period, so Danto's theory is a response to the conceptual art of the late twentieth century, art that does not necessarily *embody* or exemplify its meaning but which needs to be decoded by those who have an understanding of "the artworld"—an "atmosphere of artistic theory, a knowledge of the history of art"—in virtue of which the work counts as art (Danto 1964, p. 580). Again, the theory is most appropriate to works of "high" art that are made within and in recognition of the contemporary institutions of art. Work of folk art—such as the tattoos and walrus tusk carvings of the ancient Inuit—do not fit very easily into this definition, because folk cultures often do not have a concept of "Art" as was developed in the West in the eighteenth century.

George Dickie has taken the concept of the artworld to refer not to a body of theory but to a particular group of people—artists, curators, art critics, museum-goers—and has argued that, roughly speaking, something is art if it is the sort of thing that is designed to be presented to

members of the artworld (Dickie 1984). But if we understand the artworld is this way, then once again the theory will not happily apply in cultures where there are no curators, critics or museums, and nothing approaching an "artworld." Modern attempts to surmount this problem (Levinson 1990, 1996; Carroll 2001) have emphasized the historical dimension of art and art appreciation: Perhaps we can define art in terms of the kinds of intention with which art works have traditionally been made or by the kinds of responses they have traditionally invited.

MEANING AND INTERPRETATION

In insisting that art works require an artistic theory to justify them, Danto is emphasizing that all art works have artistic meaning and require *interpretation*: One cannot just contemplate the beauty of an artwork; one needs to grasp the ideas that lie behind it, ideas that may not even be manifest in the aesthetic surface, at least until the artist or her surrogate has pointed them out. In Goodman's *Languages of Art*, art works are conceived of, by analogy with language, as symbols in different kinds of symbol system. As in Danto's theory, art is meant to be interpreted and understood, rather than merely contemplated and appreciated. The idea that works require interpretation fits well with the ethos of modernism. Modernist works are often difficult—one thinks of *The Wasteland* or the works of Schoenberg—and they *need* to be interpreted. Postmodern works may sometimes be more playful but often they too are mystifying unless you know the theory behind them, for example the stories of Italo Calvino or the late works of architecture by Peter Eisenman.

But what is it to interpret a work of art? In the late twentieth century there developed a sharp divide between the approach taken by analytic philosophers of literature who tend to stress the importance of understanding the author's probable intentions in constructing a work (Levinson 1996, Stecker 2003) and the various approaches taken by continental thinkers. German reception theory saw interpretation as primarily determined by readers' responses rather than the artist's intentions (Iser 1978). Thinkers in the structuralist and poststructuralist tradition emphasize the importance of how readers or viewers decode or deconstruct art works, thereby uncovering an abundance of possible meanings permitted by the interweaving structures of a text as well as by their interactions with further texts (Barthes 1974, Derrida 1974). Marxist, Freudian, and feminist theorists have reinterpreted works from the past from the perspective of the contemporary reader's assumptions, that might well not have been shared by the author of the work. In both analytic and continental traditions, however, the importance of taking account of the cultural context of artist and reader has been stressed.

The rage for interpretation has even reached the aesthetics of nature. Instead of just contemplating the beauties of a waterfall, a flower or a mountain, it has been argued that we should base our appreciation on scientific knowledge about what we are looking at (Carlson 2000) and that the more we know about it the more aesthetically pleased we will become. To others this seems doubtful about much of our experience of nature (Budd 1996). They could argue that the Romantics who first fostered interest in the wilder aspects of nature were no experts in the sciences of botany or geology, but were deeply moved by nature all the same.

ONTOLOGY

The question of interpretation is closely bound up with the ontological status of art works. What is it that we are interpreting when we interpret a work of art? On the face of it, paintings and sculptures and works of architecture are individual physical objects, whereas novels, symphonies, etchings and digital art works are types or abstract objects of some kind (Wollheim 1980). In addition, some arts are performing arts, requiring a performance in order to be experienced (Davies 2001). Performance arts such as dance and music raise additional questions about the authenticity of modern performances of older works. If performance practice has changed radically from when a piece was composed, are we really experiencing the work itself, a modified version of the work, or a wholly new work bearing some resemblance to the old?

Goodman distinguished allographic from autographic art forms, the former being identifiable as a structure or sequence of symbols, such as a novel, and the latter being identifiable only by means of the history of production of the artwork (Goodman 1976). One problem with this distinction is that even allographic art works may need to be distinguished by their history of production (Levinson 1990): if Smith in 2005 composes what we identify as Beethoven's Fifth in total ignorance of the "original" Beethoven's Fifth, he would on Goodman's view have composed the very same symphony. But if we take seriously the idea that a work of art is partly identified by when, where, and by whom it was made, then it would seem that Smith's "Fifth" is a different work. Confirming this conclusion is the fact that Smith's Fifth has

different artistic and aesthetic qualities from Beethoven's, being conventional and derivative, predictable and old-fashioned.

Works of art are cultural objects, objects with cultural significance, so they cannot be treated simply as individuals like tables and chairs on the one hand or like abstract types such as the standard meter on the other. Whether a work of art is an individual or a type, it has to be identified partly by means of the cultural context that spawned it, hence the importance of the artist's intentions and the historical, geographical, and intellectual context in which the artist operated (Margolis 1999). From this point of view, interpretation is necessarily bound up with ontology. Not everyone agrees, of course. But those who think that ontological questions should be kept separate from questions about interpretation have some difficulty in explaining how this is to be accomplished.

ART AND KNOWLEDGE

If art works are symbols that need to be pored over in order to release their meanings, then it is reasonable to expect them to advance our cognitive skills and to reveal truths about the world. This claim, however, has been controversial ever since Plato, who famously rejected the claims of poetry to knowledge, arguing that shadows and reflections lead away from rather than toward the truth. Aristotle, on the other hand, argued that poetry is more philosophical than history, because it is about universals rather than particulars, the probable rather than the actual (Janko 1987).

In the classical period, when the arts were thought of as arts of imitation, art works could be a means to knowledge in a very straightforward way: If a painting of Napoleon's Coronation is an imitation or representation of the coronation, then it can inform the world at large that Napoleon has been crowned emperor, what the event looked like, and how important it was. The absolute Idealists, however, made far weightier claims for art: For them it was a mode of knowledge of absolute Spirit. Shorn of its idealist underpinnings this idea can be seen to be a variety of a very old idea: that the artist is a special person with special insight into reality. In the Romantic period, when the arts were thought of as expressions of the artist's attitudes and emotions, the knowledge art works could be expected to provide was knowledge of the emotions, both the artist's and our own. The artist worked out his emotions for us in such a way that we can recreate them in imagination and thereby arrive at self-knowledge.

Current theories about the cognitive value of art are less ambitious. The tendency is to emphasize that works of art are not the best conduits for propositional scientific knowledge, but that they can teach us in other ways. Goodman stressed how paintings, sculptures, films and the other visual arts can teach us to become more adept at making perceptual discriminations of various kinds (Goodman 1976). Literary works in particular have often been thought to provide us with moral knowledge, knowledge of moral truths that can be expressed in propositional terms, as well as knowledge of how to live, how to balance different goods, how to treat one's friends and how to make moral decisions. Novels, films, plays and short stories are thought to be tailor-made to educate our emotions and teach us moral values (Nussbaum 1990, Robinson 2005). On the other hand, if we try to abstract what moral truths are taught by a great work of literature, the best we can often come up with is some banality that may not even be true: *King Lear* teaches us that love is exhibited in deeds, not words, *Anna Karenina* that misery ensues if you abandon your husband and children.

ART AND EMOTION

Goodman has suggested that in our appreciation of art works, the emotions function cognitively. This is an idea first found in Aristotle, who argues that the goal of tragedy is to evoke a catharsis of pity and fear. Although the meaning of "catharsis" has been much debated, nowadays it is usually thought to imply that the evocation of pity and fear is an aid to understanding, not just a fortuitous accompaniment of the tragedy. Aristotle is replying to Plato's denunciation of the art of tragedy as evoking emotions that weaken the moral fiber.

Goodman's idea is more general than Aristotle's. It suggests that understanding any kind of art work may be accomplished in part by having our emotions aroused. For example, feeling surprised, bewildered, and finally relieved by the way the themes and harmonies behave in a piece of music may alert us to the form or structure of the piece (Meyer 1956). Having our emotions aroused by the gradual unfolding of the plot of a novel may draw our attention to important structural high points. But in the literary case our emotions may also help us to understand not just the works of art themselves but also something of life itself. In responding sympathetically to how the characters are feeling and responding and what the significance of their various situations is, we learn what it is like to be in various unfamiliar situations. Responding sympathetically to characters in a novel can give us practice in

understanding other people in real life (Feagin 1996, Carroll 2001). More generally, imaginative engagement with works of literature, film, painting and so on can broaden our imaginative horizons.

The Expression Theory insists that art works do not merely arouse emotions in audiences but also express emotions themselves. This means that an art work can contain a point of view or attitude that gets articulated in the work (Robinson 2005), as, for example, Wordsworth's famous poem articulates the emotions of a stranger, a wanderer, who feels "lonely as a cloud," but becomes happy when he comes across a joyous crowd of daffodils. Paintings too can contain such emotional points of view, for example Monet's *The Seine in Thaw*, painted after the death of his wife Camille, which Wollheim sees as an expression of mourning (Wollheim 1987).

ART AND VALUE

Views about the value of art vary depending on what the essential features of art are taken to be (Budd 1995). For formalists, the value of art is likely to be purely aesthetic, the provision of aesthetic pleasure or aesthetic emotion (Bell 1914). Expression theorists value the arts for their ability to articulate the artist's emotions (Collingwood 1938, Dewey 1934) or to communicate emotions from one person to another (Tolstoy 1960). Cognitive theories of art stressing the meaning and interpretation of art works stress the cognitive values of art, its ability to improve our perceptual and emotional awareness of the world (Goodman 1976, Langer 1953). Of these kinds of value, aesthetic value seems to be a genuinely intrinsic value and a value intrinsic to art. Increased understanding and improved communication among people are no doubt intrinsic goods also, but they are not unique to the arts. By contrast, theories of art that define art in terms of its cultural context or the institutions that surround it do not seem to explain why art has value.

One problem that has been much discussed returns us to the origins of aesthetic theory in the eighteenth century. The question is whether the aesthetic value of the arts includes other sorts of value. Most thinkers on the subject have rejected the idea that monetary value has any bearing on aesthetic value, and most have also distinguished between the aesthetic value of an artwork and its value as a historical or archeological document. But there is no clear consensus on whether the value of art includes moral value, or whether we should keep a sharp divide between the realms of the moral and the aesthetic (Lamarque and Olsen 1994, Gaut 1998). Those who think that art works are primarily designed to provide aesthetic experiences (Beardsley 1958, Iseminger 2004), are more likely to think that moral value is irrelevant to aesthetic value. But to those who think that the arts are rich repositories of values of all sorts, including cognitive and emotional values (Goldman 1995), moral value will be just one more source of artistic value in a work.

See also Aesthetic Experience; Aesthetic Judgment; Aesthetic Qualities; Aesthetics, History of; Aristotle; Art, Value in; Batteux, Abbe Charles; Baumgarten, Alexander Gottlieb; Beardsley, Monroe C.; Beauty; Collingwood, Robin George; Continental Philosophy; Danto, Arthur; Dewey, John; Empiricism; Enlightenment; Feminist Aesthetics and Criticism; Goodman, Nelson; Hegel, Georg Wilhelm Friedrich; Hume, David; Hutcheson, Francis; Kant, Immanuel; Neoplatonism; Plato; Plotinus; Rationalism; Renaissance; Romanticism; Schopenhauer, Arthur; Sibley, Frank; Wittgenstein, Ludwig Josef Johann; Wollheim, Richard.

Bibliography

Barthes, Roland. *S/Z*. Translated by Richard Miller. New York: Farrar, Straus and Giroux, 1974.

Batteux, Abbé Charles. *Les Beaux-Arts réduits à un même principe* (1746). Geneva: Slatkine Reprints, 1969.

Baumgarten, Alexander. *Meditationes philosophicae de nonnulis ad poema pertinentilous*. Halle, 1735. Translated by K. Aschenbrenner and W. Holther as *Reflections on Poetry*. Berkeley: University of California Press, 1954.

Beardsley, Monroe C. *Problems in the Philosophy of Criticism*. New York: Harcourt, Brace, 1958.

Bell, Clive. *Art*. London: Chatto and Windus, 1914.

Budd, Malcolm. "The Aesthetics of Nature." *Proceedings of the Aristotelian Society* 100 (1996): 137–157.

Budd, Malcolm. *Values of Art: Painting, Poetry, and Music*. London: Penguin, 1995.

Burke, Edmund. *On the Sublime and Beautiful*, edited by Charles W. Eliot. New York: P.F. Collier, 1909.

Carlson, Allen. *The Aesthetics of the Environment: The Appreciation of Nature, Art and Architecture*. London: Routledge, 2000.

Carroll, Noël. *Beyond Aesthetics: Philosophical Essays*. Cambridge, U.K.: Cambridge University Press, 2001.

Carroll, Noël. *The Philosophy of Horror, or Paradoxes of the Heart*. New York: Routledge, 1990.

Collingwood, R. G. *The Principles of Art*. Oxford: Oxford University Press, 1938.

Currie, Gregory. *The Nature of Fiction*. Cambridge, U.K.: Cambridge University Press, 1990.

Danto, Arthur. "The Artworld." *Journal of Philosophy* 61 (1964): 571–584.

Danto, Arthur. *The Transfiguration of the Commonplace*. Cambridge, MA: Harvard University Press, 1981.

Davies, Stephen. *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press, 1994.

Davies, Stephen. *Musical Works and Performances: A Philosophical Exploration*. Oxford: Oxford University Press, 2001.

Derrida, Jacques. *Of Grammatology*. Translated by Gayatri Chavrakorty Spivak. Baltimore: Johns Hopkins University Press, 1974.

Dewey, John. *Art as Experience*. New York: Putnam, 1934.

Dickie, George. *The Art Circle*. New York: Haven, 1984.

Eco, Umberto. *Art and Beauty in the Middle Ages*. Translated by Hugh Bredin. New Haven, CT: Yale University Press, 1986.

Feagin, Susan. *Reading with Feeling*. Ithaca, NY: Cornell University Press, 1996.

Gaut, Berys. "The Ethical Criticism of Art." In *Aesthetics and Ethics: Essays at the Intersection*, edited by Jerrold Levinson, 182–203. Cambridge, U.K.: Cambridge University Press, 1998.

Goldman, Alan H. *Aesthetic Value*. Boulder, CO: Westview, 1995.

Gombrich, Ernst. *Art and Illusion*. London: Phaidon, 1960.

Goodman, Nelson. *Languages of Art*. Indianapolis, IN: Hackett, 1976.

Greenberg, Clement. *The Collected Essays and Criticism*. 2 vols. Chicago: University of Chicago Press, 1986.

Hegel, G.W. F. *Vorlesungen über die Ästhetik*. Berlin: 1835. Translated by T. M. Knox as *Lectures on Fine Art*. 2 vols. Oxford: Oxford University Press, 1975.

Hospers, John. "The Concept of Artistic Expression." *Proceedings of the Aristotelian Society* 55 (1954–1955): 313–344.

Hume, David. *Essays, Moral, Political, and Literary*, edited by Eugene F. Miller. Indianapolis: Liberty Classics, 1985.

Hutcheson, Francis. *An Inquiry concerning Beauty, Order, Harmony, Design*, edited by Peter Kivy. The Hague: Martinus Nijhoff, 1973.

Iseminger, Gary. *The Aesthetic Function of Art*. Ithaca, NY: Cornell University Press, 2004.

Iser, Wolfgang. *The Act of Reading: A Theory of Aesthetic Response*. London: Routledge and Kegan Paul, 1978.

Janko, R. *Aristotle: Poetics* Indianapolis, IN: Hackett, 1987.

Kant, Immanuel. *Critique of the Power of Judgment*, edited by P. Guyer. Translated by P. Guyer and E. Matthews. Cambridge, U.K.: Cambridge University Press, 2000.

Kivy, Peter. *Music Alone*. Ithaca, NY: Cornell University Press, 1990.

Kivy, Peter. *Philosophies of Arts: An Essay in Differences.* Cambridge, U.K.: Cambridge University Press, 1997.

Kristeller, Paul. "The Modern System of the Arts." *Journal of the History of Ideas* 12 (1951–1952): 496–527; *Journal of the History of Ideas* 13 (1951–1952): 17–46.

Lamarque, Peter. *Fictional Points of View*. Ithaca, NY: Cornell University Press, 1996.

Lamarque, Peter, and Stein Haugom Olsen. *Truth, Fiction, and Literature*. Oxford: Clarendon, 1994.

Levinson, Jerrold. *Music, Art, and Metaphysics*. Ithaca, NY: Cornell University Press, 1990.

Levinson, Jerrold. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996.

Margolis, Joseph. *What, After All, Is a Work of Art?* University Park: Pennsylvania State University Press, 1999.

Meyer, Leonard B. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.

Nussbaum, Martha. *Love's Knowledge*. Oxford: Oxford University Press, 1990.

Plato. *Dialogues*. Translated by B. Jowett. Oxford: Clarendon, 1953.

Robinson, Jenefer. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press, 2005.

Ross, Stephanie. *What Gardens Mean*. Chicago: University of Chicago Press, 1998.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne. Indian Hills, CO: Falcon's Wing Press, 1958.

Sibley, Frank. "Aesthetic Concepts." *Philosophical Review* 68 (1959): 421–450.

Sparshott, Francis. *The Theory of the Arts*. Princeton, NJ: Princeton University Press, 1982.

Stecker, Robert. *Interpretation and Construction: Art, Speech, and the Law*. Oxford: Blackwell, 2003.

Stolnitz, Jerome. *Aesthetics and the Philosophy of Criticism*. Boston: Houghton Mifflin, 1960.

Tolstoy, Leo. *What Is Art?* Translated by Almyer Maude. Indianapolis: Bobbs-Merrill, 1960.

Walton, Kendall. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press, 1990.

Walton, Kendall. "Transparent Pictures: On the Nature of Photographic Realism." *Critical Inquiry* 11 (1984): 246–277.

Weitz, Morris. "The Role of Theory in Aesthetics." *Journal of Aesthetics and Art Criticism* 15 (1956): 27–35.

Wollheim, Richard. *Art and Its Objects*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1980.

Wollheim, Richard. *Painting as an Art*. Princeton, NJ: Princeton University Press, 1987.

Wordsworth, William. Preface to *Lyrical Ballads*, edited by R. L. Brett and A. R. Jones. London: Methuen, 1963.

Jenefer Robinson (2005)

AFFIRMATIVE ACTION

Affirmative action is a policy applied in the United States and other countries that aims to enhance educational and career opportunities for minorities and women by granting them preferences in college and graduate school admissions, promotions, and contract awards. Its detractors argue that a policy of favoring some races and ethnic groups over others not only fosters resentments and unrest but also compromises educational and professional standards by considering race or ethnicity ahead of objective criteria of achievement and qualifications. But supporters of affirmative action maintain that it necessary to redress past injustices—in their view created in part by traditional forms of de facto affirmative action (such as university "legacy" admissions) that have benefited only privileged elites.

In the United States, the term "affirmative action" originally referred to a court order requiring companies that had engaged in illegal racial or sexual discrimination to compensate those they had wronged and to show that they planned to avoid future illegal discrimination. Although this ruling suggested that affirmative action was compensation for unjust discrimination, it could not explain why the main beneficiaries of affirmative action were young women and young African Americans who had not been discriminated against by the companies required to hire and promote them. Some defenders of affirmative action responded that women and African Americans were the victims of a generalized prejudice compounded by a legacy of slavery. But critics pointed out that it hardly followed that companies that had refrained from participating in the pervasive discrimination and prejudice were required to compensate its victims; these critics contended that although the slaves deserved compensation from their masters, it did not follow that the descendants of the slaves deserved compensation from the descendants of the masters.

The debate on these issues was lively, but it was never completely settled because affirmative action began to refer to policies that took race and sex into account in order to increase the number of women and racial minorities in universities and businesses, with no implication that the policies were justified because the universities and businesses had practiced illegal discrimination. The defense of affirmative action therefore came to emphasize future results as much as past injustices. One early argument was that affirmative action would reduce inequality; critics countered that although it might increase the number of blacks in the middle and upper classes, it might do little to reduce overall inequality.

The most popular current defense of affirmative action in higher education centers on the educational and cultural advantages of a racially and ethnically diverse student body; this rationale was introduced by Justice Lewis Powell in the 1978 *Regents of the University of California v. Bakke* decision. Writing for the majority, Justice Sandra Day O'Connor appealed to this argument in the *Grutter v. Bollinger* decision of 2003, but critics objected that the principle of strict scrutiny forbids the state from giving racial preference unless it demonstrates that they serve a compelling state interest, which, in their view, had not been demonstrated in these cases. Defenders of these decisions countered that states do indeed have an obvious and compelling interest in eliminating the racial subordination that would likely persist without some form of affirmative action. If these observers are right, the diversity argument for affirmative action may require supplementation with evidence that affirmative action is necessary to reduce racial subordination.

See also Racism; Social and Political Philosophy.

Bibliography

Anderson, Terry H. *The Pursuit of Fairness: A History of Affirmative Action*. Oxford: Oxford University Press, 2004.

Boxill, B. "The Morality of Preferential Hiring." *Philosophy and Public Affairs* 7 (1978): 248–263.

Boxill, B. "Affirmative Action in Higher Education." In *A Companion to the Philosophy of Education*, edited by Randall Curren, 593–604. Oxford, Blackwell Publishers, 2003.

Bowen, W., and D. Bok. *The Shape of The River*. Princeton: Princeton University Press, 1998.

Cohen, C., and James Sterba. *Affirmative Action and Racial Preference*. Oxford: Oxford University Press, 2003.

Dworkin, R. *Sovereign Virtue*. Cambridge, MA: Harvard University Press, 2000.

Ezorsky, G. *Racism and Justice: The Case for Affirmative Action*. Ithaca, NY: Cornell University Press, 1991.

Fullinwider R. *The Reverse Discrimination Controversy*. Lanham, MD: Rowman and Littlefield, 1980.

Hill, T. E., Jr. "The Message of Affirmative Action." *Social Philosophy and Policy* 8 (1991): 108–129.

Kershnar, S. "Are the Descendants of Slaves Owed Compensation For Slavery?" *Journal of Applied Philosophy* 16 (1999): 95–101.

McGary, H. *Race and Social Justice*. Malden, MA: Blackwell, 1999.

Mosley, A., and N. Capaldi. *Affirmative Action: Social Justice or Unfair Preference*. Lanham, MD: Rowman and Littlefield, 1996.

Nickel, J. 1975. "Preferential Policies in Hiring and Admissions: A Jurisprudential Approach." *Columbia Law Review* 75, 534–558

Roberts, R. "Why Have the Injustices Perpetrated against Blacks in America Not Been Rectified?" *Journal of Social Philosophy* 32 (2001): 357–373.

Sowell, Thomas. *Affirmative Action around the World: An Empirical Study*. New Haven: Yale University Press, 2004.

Thomson, J. "Preferential Hiring." *Philosophy and Public Affairs* 2 (1973): 364–384.

Thomas, L. "What Good Am I?" In *Affirmative Action and the University*, edited by S. Kahn, 125–131. Philadelphia: Temple University Press, 1993.

Wasserstrom, R. "The University and the Case for Preferential Treatment." *American Philosophical Quarterly* 13 (1976): 165–170.

Bernard R. Boxill (2005)

AFRICAN PHILOSOPHY

Many of the greatest thinkers of the modern era, including David Hume, Immanuel Kant, and Thomas Jefferson,

considered Africans and their descendants to be so intellectually handicapped as to make them philosophical invalids, incapable of moral and scientific reasoning. Thus, prior to the twentieth century, the idea of African Philosophy was, for most educated Europeans and Americans, an oxymoron (Eze 1997, pp. 4–5).

Moreover, the notion of African philosophy was provocative (in a way that the notion of British or French or German or Chinese philosophy was not) because the cultures of sub-Sahara Africa had no indigenous written languages in which issues were traditionally discussed and examined. Other than the Egyptians and Ethiopians, most African cultures developed a written script only in response to Islamic and European influences. Following the model of European and North American philosophy, one group of contemporary African philosophers has contended that philosophy requires a tradition of written communication, and that African cultures must evolve beyond traditional conceptions expressed in oral forms if they are to develop the levels of critical exchange required for sophisticated scientific and philosophical activities (Wiredu in Mosley 1995, pp. 160–169; Hountoundji 1983, p. 106). But others have argued that African philosophy should be sought in the values, categories, and assumptions that are implicit in the language, rituals, and beliefs of traditional African cultures. In this view, African philosophy is a form of ethnophilosophy—such as ethnobiology and ethnopharmacology—one of the many subject areas of ethnology.

AFRICAN PHILOSOPHY AS ETHNOPHILOSOPHY

One of the principal sources of African ethnophilosophy was the French philosopher Lucien Levy-Bruhl (1857–1939). Levy-Bruhl taught at the Sorbonne from 1896 to 1927 and was one of the leading ethnologists of his era. He argued that the primary concepts, causal relationships, and modes of reasoning used by non-European people were not the result of scripts developed through academic exercises to conform to the laws of Aristotelian logic. Rather, they were "collective representations" inculcated during rites and rituals as a result of intense affective and psychomotor experiences. The concepts of non-European people were felt rather than understood, mystical rather intellectual, and mediated relationships between both physical and nonphysical modes of being. Every event had not only a physical but a "mystical" significance, and the connections between physical and mystical realities were governed by "laws of participation" that transcended the laws of logic that structured thought in European cultures. In contrast to the law of the excluded middle and the law of noncontradiction, these "laws of participation" allowed things to be both themselves and something else, to be "here" and not here, and to exist both in the present and in the future. Medicine, magic, witchcraft, divination, and communication with the dead were made possible through mystical forces apprehended through "laws of participation" that could not be reduced to "rational explanations" structured by the laws of logic.

In *Bantu Philosophy* (1945), Father Placide Tempels proposed to articulate the structure of reality implicit in traditional African culture. For Tempels, the basic difference between European and African views of reality was ontological. Whereas the basic constituents of reality in European civilization tended to be things with fixed natures (atoms, minds, bodies), the basic constituents of reality in traditional African cultures were dynamic forces. These forces were organized hierarchically into divine, celestial, terrestrial, animal, plant, mineral (including fire, water, and air), and human forces. Good and evil were made manifest in the use of these forces to amplify or diminish the vitality of human beings. Through medicine, witchcraft, sorcery, and divination, certain individuals were able to manipulate these forces to the benefit or detriment of their communities.

Temple's analysis reflected in many respects the Sapir-Whorf thesis that the structure of a culture's language shapes the way that culture structures reality. In his book Whorf argued that the structure of Native-American languages such as Hopi gave rise to an ontology of fields and forces, whereas the structure of Indo-European languages gave rise to an ontology of discrete things. From this point of view, philosophical principles were implicit in the structure of the language, beliefs, and practices of a culture, whether or not they were stated explicitly by any member of that culture. Tempel's analysis was extended and refined by Father Alexis Kagame of Rwanda and by the Belgian ethnographer Jahnhein Janz.

In his influential book, *African Religions and Philosophy* (1969), Professor John Mbiti elaborated the view that implicit in African cultures were different concepts of causality, time, and personhood. Every event had both a physical and a spiritual cause, traceable to the influence of a continuum of spiritual beings (consisting of the living, the ancestral dead, deities, and God). Key to understanding this African metaphysic was a concept of time that consisted of an endless past (the *Zamani*), a living present (the *Sasa*), and a truncated future that returned to the past. Those who had recently died continue to interact

with the living for as long as they were remembered, and then they too returned to the *Zamani*.

One of the major expressions of philosophy as ethnology was *negritude*, a principal exponent of which was Leopold Senghor. Senghor argued that Africans have a distinctive approach to reality in which knowledge is based on emotion rather than logic, where the arts are privileged over the sciences, and where sensual participation is encouraged over cerebral analysis. For Senghor, the European analyzes reality from an objective distance whereas the African embraces reality by participating in it aesthetically and spiritually. This difference between African and European cultures was, for Senghor, physiologically based and inherited (Senghor 1962). However, for Aime Cesaire, the other principal exponent of negritude, though the differences between African and European cultures were real, they resulted primarily from historical circumstances rather than biological differences (Arnold 1981, p. 37).

Whether biologically, culturally, or historically determined, many have claimed that the African contribution to civilization was invaluable because it was unique and peculiar to Africans. Nationalists in Africa and in the diaspora—Edward Blyden, Martin Delany, Alexander Crummell, Ndabaningi Sithole, Kwame Nkrumah, Alex Quaison-Sackey, and Leopold Senghor—denied that the African was a degenerate form of the European, and instead held that Africans as a race embodied capacities and potentialities that were different from but equal to those of Europeans. Pan-African nationalists typically held that abolition of the slave trade, slavery, colonialism, and the return of Africans in the diaspora to Africa would reverse the paralyzing effect of European imperialism and make it possible for Africans to develop their peculiar contributions to the evolution of civilization. Africans who chose to remain in the diaspora nonetheless had an obligation to focus inward to develop their peculiar talents so as to address their peculiar problems, rather than looking to Europe for ideas and solutions. From a nationalist perspective, African philosophy should be concerned with articulating those factors that distinguish the African worldview. This orientation rejects the European Enlightenment focus on universal standards of reason, religion, and political development, relative to which every other culture was to be measured. Among European philosophers, it drew its support from Johann Herder, who championed a kind of cultural pluralism that encouraged each race or ethnic group to develop a national character that reflected its peculiar linguistic, historical, and cultural heritage.

CRITICISMS OF AFRICAN ETHNOPHILOSOPHY

Many critics of ethnophilosophy deny that the basis of African philosophy should be sought in the structure of traditional African culture, and tend to favor the more universalist outlook of the European Enlightenment. For Kwasi Wiredu, the development of philosophy in Africa should parallel the development of philosophy in Europe, and traditional African thought should not be considered the principal source of contemporary African philosophy any more than traditional European thought (of the Celtic and Nordic variety) is considered the primary source of contemporary European philosophy. Wiredu is critical of the tendency to preserve traditional beliefs and practices even when they have little rational justification or practical utility. He stresses the need to develop written modes of communication, arguing that literacy is a necessary condition of the transition from a prescientific to a scientific world view. In his view, it is likely that literacy will have as great an impact on the oral cultures of Africa as it had on the oral cultures of premodern Europe.

The fight against colonialism in Africa gave rise to many activists—such as Julius Nyerere, Kenneth Kaunda, Sekou Toure, and Leopold Senghor—who used philosophy for political purposes. But for the critics of ethnophilosophy, postcolonial philosophy in Africa is the era of the professional philosopher, whose interests have been formatively shaped by training in the European philosophical tradition. For the professional philosopher, just because something may have developed by Europeans is no argument against its proving useful for Africans. African philosophers have a pivotal responsibility to domesticate the products of European thought into materials usable by Africans both on the continent and in the diaspora.

But defenders of the professionalization of contemporary African philosophy are also critical of the tendency to automatically reject traditional African institutions and beliefs in favor of modern European ones. A central function of postcolonial African philosophy should be "conceptual decolonization," which means avoiding or reversing the unexamined assimilation of European ideas by African people. The necessity of a decolonization of the African mind derives from the imposition on Africa of foreign conceptual schemes through the mediums of language, religion, and politics. Wiredu, along with Kwame Gyekye (1995, 1997), Marcien Towa, and others, stress that the professional African philosopher must be prepared to utilize indigenous

sources of wisdom when they offer viable insights and options. Only by the critical assessment of both modern and traditional sources will Africa develop cultural variants that are not the result of the indiscriminate acceptance of either.

Thus, Wiredu defends professional African philosophers from the charge of inauthenticity, and challenges them with two important responsibilities: domesticating European ideas and adapting them to African needs; and reconstructing traditional African ideas so they are relevant to contemporary problems. With his colleague, Kwame Gyekye, the procedure he suggests for domesticating European ideas is that of translating European ideas into an indigenous African language. If an issue addressed in European languages (e.g., the mind-body problem) makes no sense when translated into one's indigenous African language, then it is likely to be an issue that is peculiar to its European origins, and may produce more problems than it solves when applied within the African context. But one must recognize that this test of relevancy is problematic. For given the multiplicity of languages in Africa, even within a single modern nation state, it is questionable whether what does not make sense in one African language (e.g., Akan, Ga) will also not make sense in other African languages (e.g., Xhosa, Zulu). And what of Africans in the diaspora, whose indigenous language is English or French or Portuguese?

UNAMISM

One of the chief criticisms of the ethnophilosophical approach to African Philosophy is its tendency to treat African cultures as if they all must have some essential feature in common. Paulin Hountoundji (1983, 2002) rejects the contention that there is some unarticulated collective philosophy imbedded within folk beliefs that all Africans adhere to, a view he calls "unamism." Too often, he argues, ethnophilosophers intentionally or unintentionally reconstruct traditional beliefs according to categories provided by Europeans to advance European interests. Thus, Hountoundji claims, Tempels' analysis was made in order to help European colonialists devise better ways to rule the Bantu people. The intent was to benefit not Africans, but Europeans. Likewise, it was European racists who characterized Africans as being ruled by their emotions, incapable of logical thought or the ability to effectively plan for the future. Valorizing these traits as definitive of traditional African cultures simply plays into the hands of the racists. In contrast, Hountoundji argues that African philosophy must be a critical literature produced by Africans for Africans. And philosophy, like science, must be a process of continual self-examination and critical reflection that requires a tradition of literacy. Only if ideas are recorded can energy be focused on assessing them rather than merely recalling them (Hountoundji 1983).

APPROACHES TO AFRICAN PHILOSOPHY

Whereas Wiredu and Hountoundji construe literacy as essential to the practice of African philosophy, others such as Odera Oruka (1990), Kwame Gyekye, and J. O. Sodipo insist that active engagement in critical reflection on the beliefs and practices of one's culture is a requirement sufficient for that culture to have a tradition of philosophy. From their perspective, African sages that critically reflect on the assumptions of their culture are just as much philosophers as was Socrates. Thus, one may legitimately consider proverbs to be the result of critical reflection in traditional African thought, their purpose being to provide, not a scripted system of abstract rules, but a situational model to guide concrete action. If one follows the orientation of traditional thought, Godwin Sogolo argues, the point of African philosophy would be more to guide people in how they should interact with the world rather than to provide them with a true understanding of it. Odera Oruka's conversations with Luo sages, Hallen and Sodipo's (1986) conversations with Yoruba Babalawo, and Marcel Griaule's conversations with Ogotemmeli show them to be individuals with levels of critical wisdom comparable to that of Socrates.

THE NATIONALIST-IDEOLOGICAL APPROACH. Another approach to African philosophy may be characterized as nationalist-ideological, hermeneutical, or liberationist. Its exponents would include Tsenay Serequeberhan, Franz Fanon, Kwame Nkrumah, Julius Nyerere, Amical Cabral, W. E. B. Dubois, Chubba Okadigbo, and Wamba Dia Wamba. In this approach, philosophy takes the lived experience of African people as its starting point, and the lived experience of most Africans revolves around a struggle to cope with the omnipresent effects of European colonialism and neocolonialism. As such, the principle objective of African philosophy must be how to achieve liberation from the injuries imposed by European imperialism. Traditional beliefs are not valuable in themselves, but have merit in modern Africa only to the extent that they contribute to this end. A focus on the past as the source of authenticity diverts attention from the regressive nature of many

beliefs and practices, and detracts from a critical posture that evaluates all practices, both traditional and modern, of both African and European origin, relative to their contribution to the liberation of Africa. African philosophy must address the fact that many traditional leaders were installed by European imperialists as mere mouthpieces of colonial rule, and many contemporary African leaders have remained neocolonial puppets, even as they have appropriated the symbols of traditional Africa with the power of the modern state.

In addressing the question of liberation, a central question for many African philosophers is the relative importance of race versus class. Many see race to be as or more important than class in the struggle for African liberation, and they doubt whether the white proletariat will abandon the privileges of white supremacy in order to form a united front with people of color. A case in point is the apartheid regime of South Africans, where poor whites who considered themselves Africans nonetheless insisted on privileges over black Africans. Even when race is secondary, the effects of colonial rule continue to divide Africans along tribal lines. Thus where Africans have replaced Europeans in neocolonial states, it is often tribal differences among Africans that is a source of current problems. As Kwame Gyekye (1997) points out, loyalty to family and tribal affiliations tends to breed nepotism, graft, and corruption when fostered by neocolonial ties. For Franz Fanon, racism was simply a way of justifying oppression by insisting on the inferiority of the oppressed. Africans would gain a sense of agency, he argued, only when, through struggle, they overcame the false separations of race and tribe introduced by colonialism. Africans must devise, through their own initiative, the means to liberate themselves (Fanon 1963). Cabral argued that this would require urban intellectuals to "return to the source" and form alliances with the agricultural peasantry in the fight for freedom from colonialism and neocolonialism. (Cabral 1979)

AFROCENTRISM. Afrocentricism is built around the claim that Black Africa's contributions to world culture have been denied in order to further a racist agenda. Afrocentrists take as their patron Cheik Anta Diop, who argued that Egypt was an African culture, and its achievements in science, mathematics, architecture, and philosophy were the basis for the flowering of classical Greek civilization. That the ancient Egyptians were black Africans was freely acknowledged in the ancient world but was denied and misrepresented by modern Europeans in order to justify racism, slavery, and colonialism. Diop uses language, rituals, and practices to trace the origins of the major sub-Saharan African cultures to ancient Egyptian civilization. As such, he denies that Africans are "naturally" more oriented towards the arts than to science and technology. Rather, he claims that European imperialism in the modern era impoverished Africa's resources and stifled it's scientific, technological, and political development. The imposition by Europe of a patriarchal ethical and social structure on an African orientation that was traditionally matriarchal further distorted Africa's social and political development.

THE PROBLEM WITH RACE. Kwame Appiah has mounted a sustained attack on the view that African philosophy should express the peculiar orientation of the African race. He argues in *In My Father's House* (1992) that, before their contact with Europeans beginning in the fifteenth century, people on the African continent did not view themselves as members of the same race. The notion of the African race was invented by Europeans to justify a generic form of continental oppression. Moreover, Appiah has argued that people should reject the notion of race because there is no biological or cultural basis for dividing humankind into races: there is more variation, he claims, both biologically and culturally, among those characterized as Africans than there is between the average African and European. Thus, the Pan-African ideal of uniting all members of the African race, both on the continent and in the diaspora, is flawed and is itself a form of "intrinsic racism." (Appiah 1992, p. 17) Attempts to identify some set of traits as the essence of the African race are misguided, whether the intent is to denigrate or valorize.

Appiah's views reflect a trend, since the end of WWII, of rejecting racism by rejecting the existence of races. However, within biology and anthropology this orientation is highly contentious. Many, including Diop, reject racial essentialism and racism but insist nonetheless that there are legitimate grounds for recognizing the existence of races. That Africa is the source of all humankind is one explanation for the huge range of variation among its people, who are moreover united by a history of super exploitation and denigration.

THE FEMINIST PERSPECTIVE European philosophy has typically assumed that the interest of males represents the interest of the species, just as it has assumed that European philosophy is the standard for judging all other attempts to do philosophy. Thus, given similar histories of struggling against domination, many feminist philosophers have shared with Africans and African Americans an interest in deconstructing traditional philosophical

methods and assumptions so as to expose implicit agendas of domination. Ifa Amadiume (1997) has elaborated Diop's contention that precolonial Africa was primarily matriarchal, but moves beyond Diop to stress the advantages of small political units such as the family and village over large political units such as nations and empires. Other African feminists not only deny that traditional African societies followed the European paradigm of privileging men over women but also consider patriarchy and matriarchy to be European categories imposed to configure Africa on a European standard.

Africa has had its biggest cultural impact on the direction of contemporary European culture, not in the sciences, but in the arts. African sculpture, painting, music, and dance have radically influenced the development of modern European art forms and aesthetic values. But traditional African art forms have differed from modern European art forms in several important respects. Modern art is often displayed in museums as objects to be viewed, not touched. But traditional African art played functional roles in addressing practical realities, and Beauty resided as much in what something did as in how it looked. Music and dance were activities to be participated in, not simply perceived from a distance, and they provided individuals with a model of how to situate themselves in a world in which they played an active role in creating.

The American feminist Sandra Harding has stressed the similarity between the struggle of Africans and the struggle of women against European male hegemony. Other American feminists have argued that values implicit in Africa's practice of the arts may help to develop a better appreciation of the ingredients of the ethical life and reinforce orientations that enhance people's ability to live together. In much of the European philosophical tradition, ethics involves the attempt to articulate principles that should guide and justify the choices one makes. But Cynthia Willett (1995) and Kathleen Higgins (1991) have attempted to ground ethical relationships in the music and dance traditions of the African aesthetic rather than in principles deriving from rational choice or compassionate care. In a similar vein stressing the importance of the aesthetic orientation in African philosophy, Richard Bell (2002) proposes that African philosophy should be conceived as embodied in narrative icons rather than verbal texts. These developments show how African philosophy should not be considered the exclusive domain of men, that it need not take science as its principal exemplar, and that one need not be African in order to address issues of central importance in African philosophy.

The domination of African states by repressive regimes of colonial and neocolonial tyrants has institutionalized violence throughout Africa and its diaspora. The Truth and Reconciliation tribunals of South Africa have provided a novel process for achieving justice. This approach recognizes that the purpose of seeking the truth concerning violence against the people is to seek atonement and reconciliation; and that this is something that is as much needed in dealing with crimes of Africans against Africans as in crimes of Europeans against Africans.

See also Aristotelianism; Enlightenment; Feminist Philosophy; Harding, Sandra; Herder, Johann Gottfried; Hermeneutics; Hume, David; Jefferson, Thomas; Kant, Immanuel; Lévy-Bruhl, Lucien; Mind-Body Problem; Multiculturalism; Racism; Socrates.

Bibliography

Amadiume, Ifa. *Reinventing Africa: Matriarchy, Religion, & Culture*. London: Zed Books, 1997.

Appiah, Kwame Anthony. *In My Father's House: Africa in the Philosophy of Culture*. New York: Oxford University Press, 1992.

Arnold, A. James. *Modernism and Negritude – The Poetry and Poetics of Aime Cesaire.* Cambridge, MA.: Harvard University Press, 1981.

Bell, Richard. *Understanding African Philosophy*. New York: Routledge, 2002.

Cabral, Amilcar. *Unity and Struggle.* New York: Monthly Review Press, 1979.

Diop, Cheikh Anta. *Civilization or Barbarism: An Authentic Anthropology*. Translated by Yaa-Lengi Meema Ngemi. Brooklyn, NY: Lawrence Hill Books, 1991.

English, Parker, and Kibujjo Kalumba, eds. *African Philosophy: A Classical Approach*.Englewood Cliffs, NJ: Prentice Hall, 1996.

Eze, Emmanuel Chukwudi, ed. *Postcolonial African Philosophy: A Critical Reader*. Cambridge, MA: Blackwell, 1997.

Eze, Emmanuel Chukwudi, ed. *Race and the Enlightment: A Reader*. Cambridge, MA: Blackwell Publishers, 1997.

Fanon, Franz. *The Wretched Of The Earth.* New York: Grove Press, 1963.

Gyekye, Kwame. *An Essay on African Philosophical Thought*. Philadelphia: Temple University Press, 1995.

Gyekye, Kwame. *Tradition and Modernity* Oxford: Oxford University Press, 1997.

Hallen, Barry, and J. O. Sodipo. *Knowledge, Belief, and Witchcraft*. London: Ethnographica Press, 1986.

Higgins, Kathleen. *The Music of Our Lives*. Philadelphia: Temple University Press, 1991.

Hountoundji, Paulin. *African Philosophy: Myth and Reality*. Bloomington: Indiana University Press, 1983.

Hountoundji, Paulin. *The Struggle for Meaning: Reflections on Philosophy, Culture, and Democracy in Africa*. Athens: Ohio University Press, 2002.

Masolo, D. A. *African Philosophy in Search of Identity*. Bloomington: Indiana University Press, 1994.

Mbiti, John. *African Religions and Philosophy.* London: Praeger Publishers, 1969.

Mosley, Albert, ed. *African Philosophy: Selected Readings* Englewood Cliffs, NJ: Prentice Hall, 1995.

Mutiso, Gideon-Cyrus, and S. W. Rohio. *Readings in African Political Thought*. London: Heinemann Educational Books, 1975.

Nkrumah, Kwame. *Neo-Colonialism: The Last Stage of Imperialism*. London: Nelson, 1965, 1965.

Nkrumah, Kwame. *Consciencism*. New York: Monthly Review Press, 1970.

Onyewumi, Oyeronke. *The Invention of Women*. Minneapolis: University of Minnesota Press, 1997.

Oruka, Henry Odera. *Sage Philosophy*. Leiden, Netherlands: Brill, 1990.

Senghor, Leopold, "On Negrohood: Psychology of the African Negro," *Diogenes,* Spring (1962), pp. 1–15.

Serequeberhan, Tsenay, ed. *African Philosophy: The Essential Readings*. New York: Paragon House, 1991.

Serequeberhan, Tsenay. *The Hermeneutics of African Philosophy*. New York: Routledge, 1994.

Tempels, Placide. *La philosophie bantoue* (Bantu philosophy). Elisabethville: Lovania, 1945. Paris: Presence Africaine, 1959.

Willett, Cynthia. *Maternal Ethics and Other Slave Moralities.* New York: Routledge, 1995.

Wiredu, Kwasi. *Conceptual Decolonization in African Philosophy: Four Essays.* Selected and Introduced by Olusegun Oladipo. Ibadan, Nigeria: Hope Publications, 1995.

Wiredu, Kwasi. "How Not to Compare African Thought with Western Thought." In *African Philosophy: Selected Readings*, edited by Albert Mosley. Englewood Cliffs, NJ: Prentice Hall, 1995.

Albert Mosley (2005)

AGAPE

See *Love*

AGENT CAUSATION

The concept of an agent's causing some event seems distinct from that of an event's causing another event, and this apparent distinctness has been exploited by some philosophers of action—agent causationists—to defend an incompatibilist and libertarian account of free will. Agent causationism is associated historically with, among others, the philosophers Francisco Suárez and Thomas Reid, and in more recent times has been defended by Richard Taylor and Roderick Chisholm.

AGENT CAUSATION AND EVENT CAUSATION

What is indisputable is that causal statements come in at least two forms, one in which a term denoting a person or persisting object is the subject of the verb *cause* and one in which a term denoting a particular event occupies this role. Compare, for example, "The bomb caused the collapse of the bridge" and "The explosion of the bomb caused the collapse of the bridge." Here it seems plausible to contend that the first of these statements is elliptical, meaning something such as "Some event involving the bomb caused the collapse of the bridge," and more generally that the causation of events by inanimate objects is always reducible to the causation of those events by other events involving those objects. However, it is less evident that this sort of analysis applies in cases in which a person or other intelligent agent is said to cause some event. Sentences containing transitive verbs of action generate many such cases, because an action sentence such as "John raised his arm" clearly entails a corresponding agent-causal sentence, "John caused a rising of his arm." What seems less clear is that the latter sentence entails an event-causal sentence, "Some event involving John caused a rising of his arm," at least on the assumption that John raised his arm as a so-called basic action.

A basic action is standardly taken to be one that is not done by doing anything else. An action such as closing a door is nonbasic, because one can only close a door by doing something to it, such as pushing it. It is possible to raise one's arm as a nonbasic action—for example, by pulling on a rope attached to the arm, using one's other arm. But, it seems, there is nothing one needs to do in order to raise one's arm when one raises it in the normal way. This appears to generate a difference between the case of the bomb's causing the collapse of the bridge and that of John's causing the rising of his arm: the bomb caused the collapse by exploding, but John, it seems, did not cause the rising by doing anything else. Consequently, it is not evident that there was any event involving John that could be said to have caused the rising in the way that the explosion of the bomb can be said to have caused the collapse of the bridge. In this case it appears that a statement of agent causation is not reducible to one of event causation.

Philosophers who favor a volitionist theory of action may dispute this suggestion. They may urge that there is in fact something that John did, and by doing this he

raised his arm—on the assumption, at least, that he did so voluntarily. Namely, John willed to raise his arm. It was by willing to raise his arm that he did raise it, and so it might be said that the agent-causal statement, "John caused the rising of his arm," is true only in virtue of the truth of the event-causal statement, "John's willing to raise his arm caused the rising of his arm." However, volitionism is now a minority position in the philosophy of action—in contrast with its heyday in the seventeenth and eighteenth centuries—because many philosophers are skeptical about the existence of volitions as a supposedly distinctive class of mental events. Proponents of the irreducibility of agent causation to event causation may take comfort in this fact, although they still have to face another and more prevalent kind of critic: the proponents of mainstream causal theories of action. These critics contend that intentional actions have mental causes of another sort—the onsets of states of belief and desire. While these philosophers may concede that there is no action by doing which John caused his arm to rise, they still contend that there was an event involving John that caused the rising of his arm—to wit, the onset of his desire to raise it and, perhaps, his belief that by raising it he could achieve some further desired end. This event was not an action of John's, to be sure, but it was nonetheless an event involving him that, like the exploding of the bomb, seems to explain how the effect he caused was brought about.

AGENT CAUSATIONISM AND FREE WILL

Agent causationists—that is, philosophers who maintain the irreducibility of agent causation to event causation—are opposed to mainstream causal theories of action, not least because the latter seem inhospitable to libertarianism (the doctrine that free actions lack determining causes in the form of antecedent events which causally necessitate their occurrence). Proponents of such mainstream theories are typically compatibilists concerning the relationship between free will and determinism. Agent causationists, in contrast, standardly hold that certain events caused by agents are not caused by any antecedent events, or at least that these certain events lack sufficient causes in the form of antecedent events. Some agent causationists maintain that the events in question are bodily movements—such as the rising of an arm—when these are the products of basic actions. Others maintain, perhaps more plausibly, that the events in question are certain neural events that are the causal precursors of bodily movements. Yet others seek to combine agent causationism with a form of volitionism by contending that what agents cause directly are their own volitions, choices, or endeavors. Thus, agent causationism is not necessarily opposed to volitionism, only to certain versions of it.

Common to all standard forms of agent causationism, however, is the doctrine that at least some cases of an agent *A*'s causing an event *e* do not consist in *e*'s being caused by any antecedent event involving *A*. This doctrine seems to help the case for libertarianism in the following way. The libertarian wants to say that in a case of free action, an event *e* occurs that lacks a sufficient cause in the form of antecedent events. But this prompts the objection that *e* would then be a mere chance event—such as the spontaneous decay of a radium atom—and as such would not exhibit the kind of freedom associated with an action for which an agent may be held morally responsible. The agent causationist may respond by urging that there is a significant difference between the decay of a radium atom and a case of free action because in the latter an event *e* occurs that, while lacking a sufficient cause in the form of antecedent events, still has a cause in the form of the agent whose action it is. A radium atom does not cause itself to decay, but a free human agent may cause him or herself to act in a certain way, according to agent causationism. Free agents, according to this conception, are unmoved movers or ultimate initiators of certain trains of events. And it is in having this capacity for initiation that their freedom allegedly lies, for it supposedly enables free agents to intervene in and affect the ongoing stream of events in which natural physical processes consist. Free agents' capacity for initiation is conceived to be a "two-way power"—a power either to cause or to refrain from causing an initial event of an appropriate kind.

OBJECTIONS TO AGENT CAUSATIONISM

Not surprisingly, agent causationism is subject to many criticisms—in particular from philosophers who advertise their own position as being "naturalistic"—and is charged with being mysterious and incompatible with the modern scientific worldview as revealed by physics and biology. More specifically, one popular objection is that agent causationism is committed to some form of substance dualism in the philosophy of mind, which in the eyes of most naturalistic philosophers would be enough to condemn it. However, whereas many agent causationists may in fact be substance dualists, it is not clear that their agent causationism requires them to be.

A more cogent objection to agent causationism, forcibly expressed by C. D. Broad, is that the agent causationist cannot explain why an event supposedly caused irreducibly by an agent should occur when it does (since agent-causes, unlike event-causes, are not datable items). The collapse of the bridge occurred when it did because the explosion, which was its cause, occurred when it did. But why should the many different events supposedly caused by a single agent during his or her lifetime have occurred when they did? One possible answer is that these events also have contributory causes in the form of antecedent events occurring at different times, even though each of them additionally requires causation by the agent for its occurrence. Another possible response, consistent with the first, is to appeal to temporal factors included in the agent's reasons for causing the various different events in question. (Agent causationists typically repudiate the doctrine that reasons are causes, and distinguish between reasons-explanations and causal explanations—or at least they deny that an agent's reasons, in the form of certain beliefs and desires of the agent, are part of a sufficient event-cause of the agent's action.)

Others may object that agent causationism does not really assist the case for libertarianism in the way it is alleged to for the following reason. Suppose, for the sake of argument, that instances of irreducible agent causation really do occur, and that sometimes it is the case that an agent *A* causes an event *e* in this irreducible fashion, while *e* lacks a sufficient cause in the form of antecedent events. It was suggested earlier that this still allows us to say that *e* was not just a chance event, because it was caused by *A*. However, what about the event of *A*'s causing *e*? It would seem that this event must either possess or lack a sufficient cause in the form of antecedent events. If the former, then it is hard to see how libertarianism is saved. If the latter, then it would seem that *A*'s causing *e* is itself just a chance event and so once again provides the wrong sort of freedom for moral responsibility.

Once more, various replies are available. One is simply to deny that *A*'s causing *e* qualifies as an event and as such is something eligible to possess a cause. After all, it seems odd to think of one event's causing another as itself being an event, just as it may be deemed equally odd to think of an agent's causing an event as itself being an event. Another possible reply is that when there is an instance of agent causation—*A*'s causing *e*—the agent *A* is not only the agent-cause of *e* but is also, by virtue of that fact, the agent-cause of this instance of agent causation. If this is the case, then the instance of agent causation is excluded from being a mere chance event for the same reason that the event *e* is thus excluded.

It does not appear, on close inspection, that there is anything incoherent in the notion of irreducible agent causation, but whether it really helps to solve the problem of free will and whether it is consistent with current scientific theories in physics and biology are questions that still remain open to further debate.

See also Action; Causation: Metaphysical Issues; Causation: Philosophy of Science; Chisholm, Roderick; Determinism and Freedom; Freedom; Libertarianism; Reid, Thomas; Suárez, Francisco.

Bibliography

Broad, C. D. "Determinism, Indeterminism, and Libertarianism." In *Ethics and the History of Philosophy: Selected Essays*. London: Routledge, 1952.

Chisholm, Roderick M. *Person and Object*. La Salle, IL: Open Court, 1976.

Clarke, Randolph. "Toward a Credible Agent-Causal Account of Free Will." *Noûs* 27 (1993): 191–203.

Danto, Arthur C. *Analytical Philosophy of Action*. Cambridge, U.K.: Cambridge University Press, 1973.

Kane, Robert, ed. *The Oxford Handbook of Free Will*. New York: Oxford University Press, 2002.

Lowe, E. J. "Personal Agency." In *Minds and Persons*, edited by Anthony O'Hear. Cambridge, U.K.: Cambridge University Press, 2003.

O'Connor, Timothy, ed. *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. New York: Oxford University Press, 1995.

O'Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press, 2000.

Reid, Thomas. *Essays on the Active Powers of the Human Mind*, edited by Baruch A. Brody. Cambridge, MA: MIT Press, 1969.

Suárez, Francisco. *On Efficient Causality: Metaphysical Disputations 17, 18, and 19*. Translated by Alfred J. Freddoso. New Haven, CT: Yale University Press, 1994.

Taylor, Richard. *Action and Purpose*. Englewood Cliffs, NJ: Prentice Hall, 1966.

Watson, Gary, ed. *Free Will*, 2nd ed. New York: Oxford University Press, 2003.

E. J. Lowe

AGENT INTELLECT, THE

In his *On the Soul*, iii 4–5, Aristotle wrote that there is one intellect that becomes all things and another that makes all things, just as light makes colors visible. It is separate, impassible, unmixed, and in essence activity; it alone is

immortal and eternal. Those few statements are the basis of the theory of the agent intellect.

Aristotle was studied with intense and sometimes imaginative care by ancient and medieval scholars, and his ideas were developed to the extent of dominating thought about human thinking. Our concept arose in Greek but was developed in Arabic and flowered in medieval Latin; "agent intellect" is the English rendering of the Latin *intellectus agens*, but behind that lie a number of other terms. Furthermore, English writers have sometimes used active instead of agent.

The field falls into three parts: the Greek commentators on Aristotle, the Arabic philosophers who developed his views, and the medieval Europeans who built on the rest. Aristotle himself was sparing with technical terms, and the text of *On the Soul*, iii 4 and 5. is in a poor state that raises several questions. Later thinkers brought in material from earlier parts of *On the Soul* (i 4, ii 2); part of *On the Generation of Animals* (ii 3) in which Aristotle says that in humans the intellect (unqualified) comes into the fetus from outside (*thurathen*); passages from his ethics and his metaphysics, in which the intellect is regarded as in some sense divine; and the end of his *Posterior Analytics* (ii 19). The result is far from anything Aristotle can have held.

Aristotle's student Theophrastus raised pertinent questions about the agent intellect, reported, perhaps unreliably, by Themistius (c. 317–88), who himself studied Aristotle with care and ingenuity. He reports one early view—that the agent intellect was the body of premises and deductions that form knowledge—but dismisses it, as he does the view of Alexander of Aphrodisias (c. 200 CE), who held that the productive or active or agent *nous* (now identified with the *nous thurathen* that for Aristotle was a biological concept) was identical with the First God or the unmoved mover of *Metaphysics* XII 8. It is clear that there had been much discussion about this already, and already we see a tendency to the hypostatization of various intellects.

Themistius and Alexander together influenced Arabic thinkers. Most important are Avicenna (980–1047) and Averroes (1126–98). Avicenna had a theory of celestial intellects, derived from Neoplatonist views as well as Aristotle's metaphysics and psychology; for him the agent intellect was the tenth and lowest of a chain descending from the First Intellect, far removed from the human soul. More accessible was Averroes's view, which started from Aristotle's distinction between intellect as potential and as active or agent but went on to argue that the agent intellect was one and the same in all men, leading on to the question whether the potential intellect was also one and the same in all men. The Arabic philosophers were also interested in this intellect as the source of prophecy, and in the possible conjunction with it of the human reason.

Arabic works were translated into Latin by Western medieval scholars, so that Europe became aware of much of Aristotle and of his Arabic interpreters at almost the same time; in the thirteenth century it was taken for granted that the words *intellectus agens* stood for something definite, but there remained many questions about it. Albert the Great (c.1200–1280) introduced the Latin expression *intellectus agens*. He got to know the Arabic evidence and dealt with the fourfold distinction of agent, possible, acquired, and speculative intellects, which became the basis for later discussion.

In his time there was an Averroist school of thought, particularly in Padua, which troubled more orthodox thinkers; even Thomas Aquinas (1225–1274) wrote against them. In his extensive writings he worked out a theory aimed at satisfying both Aristotelians and Christian theologians. He quoted Aristotle to disprove his opponents' views, using the *Physics*, *On the Soul*, and Themistius, Avicenna, and others. He was primarily concerned with whether there was but a single intellect for all men and the subsidiary question about the agent and the receptive intellects. In his *Summa* he concentrates on the internal features of the intellect, and the agent is that which by its light abstracts species from images.

A single agent intellect would not secure individual immortality as required by Christianity, and when many Averroist doctrines were condemned by the Church in 1277, a number were about the Agent Intellect. An anonymous work from the early fourteenth century covers sixteen supposed views about the agent intellect from Plato (who is said to have denied its existence) through the Arabs to a number of others; the writer favors Thomas Aquinas. An array of arguments, partly from Aristotle but partly independent, is deployed. There are questions about the existence of the agent intellect, and again about whether there is one in each person or only one for all, as there is one light source illuminating all illuminated objects.

Even in the Renaissance the concept is found in the Averroism of Pomponazzi (1462–1525): in his *On the Immortality of the Soul* he doubted immortality, but, opposed by the Church, argued that philosophy could not prove anything in this area. Zabarella (1533–1589), a logician, also still spoke of the agent intellect as playing a part in induction. Finally, Aristotle's dominance came to an

end, and his account of the intellect has been described recently as a museum piece. Instead of his metaphysical approach, a scientific psychology slowly developed, which was not interested in analyzing his actual words.

See also Aristotle; Averroes; Pomponazzi, Pietro; Thomas Aquinas, St.

Bibliography

Davidson, H Alfarabi. *Avicenna and Averroes on Intellect*. Oxford: Oxford University Press 1992.

Huby, P. M. "Stages in the Development of Language about Aristotle's Nous." In Blumenthal and Robinson, *Aristotle and the Later Tradition*. OSAP suppl. vol.. Oxford: Oxford University Press 1991, 129–143.

Kuksewicz, Z. "The Potential and the Agent Intellect." In *Cambridge History of Later Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1982.

Kurfess, H. *Zur Geschichte der Erklärung der Aristotelischen Lehre vom sog. NOUS POIHTIKOS und PAQHTIKOS*. Tübingen, 1911. Reprinted in *Aristotle and His Influence*, edited by L. Taran. New York and London: Garland, 1987.

Pamela M. Huby (2005)

AGNOSTICISM

In the most general use of the term, "agnosticism" is the view that we do not know whether there is a God or not. Although the history of agnosticism, in this general sense, is continuous with that of skepticism (thus reaching back to the ancients), the term itself was coined by T. H. Huxley and its distinctive philosophical bearings emerged in the course of the nineteenth-century debate on religious belief. Participants in that debate often used the word in a strong and specific sense: To be an agnostic was to hold that knowledge of God is impossible because of the inherent, insuperable limitations of the human mind. To assert confidently either the existence or the nonexistence of a deity with definite and intelligible attributes was to transgress these limits.

This consciousness of limitation is classically expressed in the "Transcendental Dialectic" of Immanuel Kant's *Critique of Pure Reason* (1781). There is a continual temptation, Kant stated, to raise questions about the totality of things; but these questions, he argued, are demonstrably unanswerable. Contradictions are encountered, for instance, whether it is assumed that the world is finite in space and time or infinite in space and time. Or, in another instance, one event may properly be called the cause of another event, but such a concept cannot be used to assert that something (a First Cause) is the cause of the universe as a whole. Of this "whole" one has, and can have, no experience. The main line of agnostic argument in the nineteenth century followed Kant closely in his criticism of cosmological reasoning, although many agnostic writers were not thoroughgoing Kantians. Nor did they have to be Humeans to have their metaphysical assurance called in question by David Hume's famous (or notorious) criticism of speculation in *An Enquiry concerning Human Understanding* (1748): "If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask, *Does it contain any abstract reasoning concerning quantity or number?* No. *Does it contain any experimental reasoning concerning matter of fact and existence?* No. Commit it then to the flames: for it can contain nothing but sophistry and illusion."

A person who calls himself agnostic commonly judges that he cannot have both agnosticism and, say, Christian belief. Yet the main positions of nineteenth-century agnosticism were in fact worked out and held by "religious agnostics," writers who argued that a very high degree of ignorance concerning the deity was nonetheless compatible with a religious commitment of some kind. In fact, if not in name, this view was also found in the twentieth century; it is essentially the view of those who disclaim metaphysical knowledge of God, but yet stake all upon "faith," "authority," or Christianity as a practical way of life. Kant may again provide the archetypal model: Having denied that theoretical reasoning could furnish arguments for the existence of God, he nevertheless claimed that God had to be "postulated" in order to make sense of moral experience.

In his most influential article, "Philosophy of the Unconditioned" (*Edinburgh Review,* 1829), Sir William Hamilton tersely introduced themes that were to be developed, refined, and repudiated by writer after writer to the end of the century and well beyond. "The mind," he wrote, "can … know only the *limited, and the conditionally limited.*" To attempt to think the unconditioned or absolute is to think away "those very conditions under which thought itself is realized." "Loath to admit that our science is at best the reflection of a reality we cannot know, we strive to penetrate to existence in itself; … But, like Ixion, we embrace a cloud for a divinity."

H. L. Mansel, in his Bampton Lectures, *The Limits of Religious Thought* (1858), tried to show in detail that alleged knowledge of the Absolute is self-contradictory at many points. One attributes personal qualities to God, for instance, and yet one cannot think through the notion of personality without the idea of limitation; thought must be distinguished from thinker, and so on. But limitation

is incompatible with infinite and absolute deity. The conclusion, however, is not a total religious skepticism. For although speculation about the divine nature is a vain attempt to escape the inescapable conditions of human thought, yet through the "feeling of dependence" and in moral conviction faith may still operate where speculative reason cannot.

Herbert Spencer in his *First Principles* (1862) accepted this picture of a limited human reason, aware of its limits and yet (in his view) aware also that those limits are decidedly not the limits of the real. Science and religion could, in fact, be reconciled by realizing that each of them testifies to a mystery, to an inscrutable Absolute, quite beyond the frontiers of knowledge or conception but yet not mere negation or nothingness.

The sources of nineteenth-century agnosticism—particularly the agnosticism of those who abandoned organized religion—were, however, more numerous and complex than has been indicated so far. It is rare indeed that a single line of philosophical argument produces by itself either religious conviction or disillusionment. At least three additional sources should be mentioned.

First, a growing mass of data and theory supplied by the physical sciences was prima facie at variance with biblical history and cosmology. There was the new time scale of geology, the impersonal and amoral Darwinian evolutionary theory, and the radical textual, historical criticism of the Bible itself.

Second, once the strong initial resistance to systematic and searching criticism of Christian teaching had been overcome, it was possible to express openly a good many moral misgivings about the Christian conception of God and his governance of the world. J. S. Mill declared it was impossible for a thoughtful person to ascribe "absolute perfection to the author and ruler of so clumsily made and capriciously governed a creation as this planet" (*Three Essays on Religion,* 1874). He found "moral difficulties" also in "the recognition … of the object of highest worship, in a being who could make a Hell" and create creatures whom he foreknew to be destined to suffer in it eternally. No less morally repugnant to many writers was the insistence of the orthodox that their dogmas required sheer unswerving acceptance, and that breakdowns in argument or intelligibility were simply occasions for the exercise of an intensified faith. T. H. Huxley was forthright. In "Agnosticism and Christianity" (1889) he wrote, "I, and many other Agnostics, believe that faith, in this sense, is an abomination." In "Agnosticism" (1889) he said, "I verily believe that the great good which has been effected … by Christianity has been largely counteracted by the pestilent doctrine … that honest disbelief in their more or less astonishing creeds is a moral offence, indeed a sin of the deepest dye."

Third, the same authors were vehemently critical of the standards of evidence and reasoning normal in theology, and contrasted them with the severe, rigorous, and dispassionate criteria of the sciences. To Mill, "The whole of the prevalent metaphysics of the present century is one tissue of suborned evidence in favour of religion." If one considers the nature of the world as one actually observes it, the very most one could dare to hazard is the existence of a good but finite deity; and Mill put forward even this possibility with a characteristically agnostic tentativeness. For Huxley agnosticism was "not a creed but a method, the essence of which lies in the rigorous application of a single principle": Reason should be followed "as far as it can take you," but undemonstrable conclusions should not be treated as if they were certain. "One may suspect," he said, "that a little more critical discrimination would have enlarged the Apocrypha not inconsiderably." In a similar vein, Leslie Stephen protested against theologians who ventured to define "the nature of God Almighty with an accuracy from which modest naturalists would shrink in describing the genesis of a black beetle" (*An Agnostic's Apology,* 1893).

It is not the purpose here to estimate how far theologians remedied, or could ever remedy, the deficiencies in their arguments that offended their agnostic critics. Some permanently valuable lessons can be learned, however, from the course of the controversies. An obvious one is the odd instability or ambiguity of certain agnostic positions. Let us suppose—as did many of the writers just quoted—that one ceases to find convincing the arguments for the existence of a deity. Experience, one now judges, is limited to the observable world; and reason, although it may lay bare the conditions and presuppositions of that experience, cannot extend our experience of what is. A religiously minded person, in this situation, is tempted to divide reality into the knowable and the unknowable and to attribute to the latter many of the lineaments of deity. Thus, "negative theology" and a religiously toned agnosticism can be the closest of relatives. No sweeping philosophical criticism can demonstrate that all such positions are untenable or involve a cryptotheism; each case must be scrutinized individually. Certain religious attitudes toward the unknown or unknowable—attitudes, for example, of wonderment and awe—can be perfectly appropriate and invulnerable to criticism, whereas others—such as the expectation of personal encounter with the unknown—are obviously

most vulnerable. One can turn to history for some examples.

In 1896 James Ward delivered his Gifford Lectures, *Naturalism and Agnosticism* (published in 1899), at Aberdeen University. These contained a vigorous attack on the basic presuppositions of the Hamilton-Mansel-Spencer approach. The sciences, Ward said, do not form a whole that floats in a surrounding "nescience." The world we know does not consist of "appearance" concealing an "ultimate reality" that lies behind or beyond it. In any case, nescience is nescience. "Where nescience is absolute, nothing can be said; neither that there is more to know nor that there is not." Spencer and like-minded writers had, however, said a good many mysterious things about their Absolute, things that, by their own account, were strictly unsayable.

R. Flint (*Agnosticism,* Croall lectures, 1887–1888, published in 1903) also denounced the equivocations (as he saw them) of a religious agnosticism. "All that the mind can do on the side of the Unknowable is to play at make-believe, to feign faith, to worship nothingness." "Call your doubts mysteries," said Stephen, satirizing the complacent, "and they won't disturb you any longer."

Is it possible for a reflective person to be an agnostic in the present time? Logical positivists have answered "No." In *Language, Truth and Logic* (1936), A. J. Ayer claimed that since "all utterances about the nature of God are nonsensical," the agnostic's statements about God are no less nonsensical than the theist's. Both assume, wrongly, that "the question whether a transcendent God exists is a genuine question." According to positivism and postpositivist logical analysis, the theological problem is not a problem of evidence and argument, but a problem of meaningfulness. If "God" is a meaningless word, the sentence "Perhaps God does not exist" is also meaningless.

In stating the situation thus, positivism was dramatically drawing attention to what it believed to be distinctive in its approach, but it simultaneously obscured some important lines of continuity with the earlier debate on agnosticism. Before the nineteenth century had ended, Flint had written, in criticism of Hamilton, "*Credo quia absurdum* can be the only appropriate motto of a philosopher who holds that we may believe in a God the very idea of whom we can perceive to be self-contradictory." The possibility of internal illogicality in the very notion of deity, the risk of the absurd and nonsensical, were well enough recognized. Spencer, wrestling with the problems of the world's origin and beginning, said that the questions here are not questions of credibility but of *conceivability.* Notions such as self-existence and creation by an external agency "involve symbolic conceptions of the illegitimate and illusive kind." The logical positivist tethered his theory of meaning to the demands for observational verification and falsification of our claims about existents. Compare Spencer once more, writing in 1899: "Intellect being framed simply by and for converse with phenomena, involves us in nonsense when we try to use it for anything beyond phenomena." It must, of course, be added that the positivists and later analysts carried out their austere program with far greater thoroughness and consistency than did their predecessors. But the lines of continuity are there; and they are—once more—those same lines that reach back to Kant's "Transcendental Dialectic" and to David Hume. They justify the use of "atheist" to describe one who rejects the performances and attitudes of religion on the grounds that talk about God is unverifiable talk, or that the concept *God* contains inner illogicalities.

But is there still room for agnosticism as undogmatic dubiety or ignorance about the existence of God? A case for saying that there is still room can be made on the following lines. Where one gives an account of an expression in our language, and where that expression is one that refers to an existent of some kind, one needs to provide not only a set of rules for the use of the expression, but also an indication of how the referring is to be done—through direct pointing, perhaps, or through giving instructions for an indirect method of identifying the entity. Can this be done in the case of God? Pointing, clearly, is inappropriate, God being no finite object in the world. The theologian may suggest a number of options at this point. He may say: God can be identified as that being upon whom the world can be felt as utterly dependent, who is the completion of its incompletenesses, whose presence is faintly adumbrated in experience of the awesome and the numinous. Clear direction-giving has here broken down; the theologian may well admit that his language is less descriptive or argumentative than obliquely evocative. Does this language succeed in establishing that statements about God have a reference? To persons susceptible to religious experience but at the same time logically and critically alert, it may seem just barely to succeed, or it may seem just barely to fail. Some may even oscillate uneasily between these alternatives without finding a definite procedure of decision to help them discriminate once for all. A person in this last category is surely an agnostic. His agnosticism takes full account of current linguistic criticisms of religion; it is in the course of his reflections upon meaning that he sees the necessity of relating the linguistic to the

extralinguistic, and his answers to this problem, the problem of reference, plunge him into the deepest uncertainty.

The temper of mind just outlined, with all its inner turbulence and anxiety, is probably the most creatively fruitful of the many varieties of agnosticism. Where there is no temptation to believe, there can be little philosophical interest in not believing. Where there has been little or no religious experience, no sense of the haunting strangeness that makes the believer wittingly violate language and logic to express it, there can be little incentive to explore minutely the possible interpretations—theistic, pantheistic, naturalistic—of that experience. As a matter of history, agnostics of this temper are to be found far more rarely today than at the height of the agnosticism controversy a century ago. For the great writers of that controversy were in most cases brought up within the Christian faith, had identified themselves with it, and subsequently suffered a bewildering disorientation. Yet, if one is to take seriously today the problems of philosophical theology, there must be some suspension of disbelief, at least an imaginative venture, in order to see why the believer feels compelled to use the extraordinary language he does use. He knows well enough that it is extraordinary; but he deems that it is ordinary language that is found wanting, and not his experiences and the interpretations he puts upon them. The agnostic knows that sometimes ordinary language needs to be violated, as a poet often violates it. He knows also that to disturb our linguistic apparatus in so radical a way can obscure some movements of thought of a very questionable (or downright invalid) logic. Has this happened in the particular case of theism? Searching in this obscurity, the agnostic reports that he cannot tell. For the health of philosophy and theology, it is well that he should continue to search.

See also Ayer, Alfred Jules; Empiricism; Hamilton, William; Hume, David; Huxley, Thomas Henry; Kant, Immanuel; Mansel, Henry Longueville; Mill, John Stuart; Skepticism, Contemporary; Skepticism, History of; Stephen, Leslie.

Bibliography

HUME AND KANT

Hume, David. *Dialogues concerning Natural Religion.* London: Robinson, 1779.

Hume, David. *Enquiry concerning Human Understanding.* London: A. Millar, 1748. Especially sections X and XI.

Kant, Immanuel. *Kritik der reinen Vernunft* (1781). In Berlin Academy, *Complete Works.* Berlin, 1902–1955.

See also the entries Empiricism; Skepticism, Contemporary; and Skepticism, History of.

NINETEENTH-CENTURY DOCUMENTS

Flint, R. *Agnosticism.* Edinburgh and London, 1903.

Hamilton, Sir William. "Philosophy of the Unconditioned." *Edinburgh Review* (October 1829).

Huxley, T. H. "Agnosticism and Christianity" (1889). In his *Collected Essays,* Vol. V. London: Macmillan, 1894. Entire volume is relevant.

Mansel, H. L. *The Limits of Religious Thought.* London, 1858.

Mill, J. S. *Three Essays on Religion.* London: Longmans, Green, Reader, and Dyer, 1874.

Spencer, Herbert. *First Principles.* London: Williams and Norgate, 1862.

Stephen, Leslie. *An Agnostic's Apology.* London: Smith Elder, 1893. First published as an essay in 1876.

Ward, James. *Naturalism and Agnosticism.* London: Macmillan, 1899; 2nd ed., with alterations, 1903.

SECONDARY SOURCES

Armstrong, R. A. *Agnosticism and Theism in the Nineteenth Century.* London: Green, 1905.

Britton, K. *John Stuart Mill.* London: Penguin, 1953.

Burtt, E. A. *Types of Religious Philosophy.* New York: Harper, 1951.

Collins, J. *God in Modern Philosophy.* Chicago: Regnery, 1959.

Garrigou-Lagrange, R. *Dieu, son existence et sa nature; solution thomiste des antinomies agnostiques.* Paris, 1915.

Hájek, Alan. "Agnosticism Meets Bayesianism." *Analysis* 58 (3) (1998): 199–206.

Kenny, Anthony. *Faith and Reason.* New York: Columbia University Press, 1983.

Lightman, Bernard. *The Origins of Agnosticism: Victorian Unbelief and the Limits of Knowledge.* Baltimore: Johns Hopkins University Press, 1987.

Marty, Martin E. *Varieties of Unbelief.* New York: Holt Rinehart Winston, 1964.

McGrath, P. J. "Atheism or Agnosticism." *Analysis* 47 (1987): 54–57.

Morris, Thomas. "Agnosticism." *Analysis* 45 (1985): 219–224.

Oppy, Graham. "Weak Agnosticism Defended." *International Journal for Philosophy of Religion* 36 (1994): 147–167.

Packe, M. St. John. *John Stuart Mill.* London: Secker & Warburg, 1954. Biography.

Passmore, J. *A Hundred Years of Philosophy.* London: Duckworth, 1957. Especially Ch. 2.

Pyle, Andrew, ed. *Agnosticism: Contemporary Responses to Spencer and Huxley.* Bristol, U.K.: Thoemmes, 1995.

Stephen, Leslie. *English Thought in the Eighteenth Century.* London: Smith, Elder & Company, 1876.

van Fraassen, Bas C. "The Agnostic Subtly Probabilified." *Analysis* 58 (3) (1998): 212–220.

Willey, Basil. *Nineteenth Century Studies.* New York: Columbia University Press, 1949.

Ronald W. Hepburn (1967)

Bibliography updated by Christian B. Miller (2005)

AGRIPPA

(c. 50 BCE–c. 150 CE)

Agrippa is known by way of one citation in Diogenes Laertius's *Lives of the Philosophers* (*DL* 9.88). Nothing is known of his life, and little of his dates (he lived between the mid-first century BCE and the second century CE). Yet Agrippa is indisputably a figure of the highest importance in the history of skepticism, indeed of epistemology in general. The citation attributes to him the invention (or at least the codification) of five "Modes," or argument patterns, which represent a new methodological rigor and self-consciousness in the development of Pyrrhonian skepticism. Earlier skeptics such as Aenesidemus had presented certain aspects of skeptical procedure in a more or less organized fashion; but the Ten Modes attributed to him are arranged according to the subject matter of the considerations appealed to. By contrast, the Modes of Agrippa seek to categorize skeptical practice according to the type and function of the argument patterns involved.

The Five Modes are summarized in two sources; in addition to Diogenes's brief notice (*DL* 9.88–89), a somewhat longer treatment survives (although without mentioning Agrippa by name) in Sextus Empiricus's *Outlines of Pyrrhonism* (*PH* 1.164–177). Taken together, they offer a general strategy for inducing doubt (and suspension of judgement, *epochē*) on every contentious issue. Two of the Modes, the First and the Third, may be described as material. The First Mode notes, in standard Pyrrhonian fashion, that most important issues are matters of dispute (*diaphōnia*) and if they are not, the skeptic will make them so. Sextus Empiricus describes skepticism as "a capacity for opposing appearances to appearances and judgments to judgments in whatever manner, so that we are brought … first to *epochē* and then to tranquility" (*PH* 1.8). Thus "we find an irresoluble conflict both among lay people and philosophers," which leads to these conditions "since we are unable either to assent or deny" (*PH* 1.165, cf. *DL* 9.88). The disputes are said to be "irresoluble" (*PH* 1.98, 212), because (skeptics allege) no independent criterion of judgment is available for them. Unpacking this claim involves invoking the other three, formal, Modes. This is because, as the Third Mode from Relativity holds, things are never apprehended in themselves and unalloyed, but only "along with something else" (*DL* 9.89): "the underlying object appears thus and so in relation to the one judging and concomitant circumstances, so we suspend judgment as to its real nature" (*PH* 1.167). Such considerations form the material for the Ten Modes of Aenesidemus, and via René Descartes and others came to dominate the landscape of epistemological scepticism (e.g., lights seem bright in the dark but dim in sunlight; oars seem straight in air, but bent in water: *PH* 1.119). Thus people can say how things appear to them but they have no grounds for any pronouncements as to how things really are.

But it is in the exposition and deployment of the three formal Modes that the power and originality of Agrippan skepticism becomes manifest. The Second is that from Regress: "what is adduced as confirmation for what is posited itself requires further confirmation, and that another, and so on *ad infinitum*" (*PH* 1.166). The Fourth is the Mode of Hypothesis, which the *Dogmatists* (Sextus's generic term for his nonskeptical opponents) resort to "when being forced to regress *ad infinitum*, take as an axiom something which they have not established, but see fit to assume as agreed without demonstration" (*PH* 1.168). This is hopeless, as Diogenes points out, because there is as a matter of fact no such agreement (*DL* 9.89). Finally the Fifth Mode, of Circularity, claims that "what ought to support the matter under investigation itself requires confirmation from that very matter" (*PH* 1.169). Diogenes adds an example: "as for instance someone seeking to confirm the existence of pores [in the skin] on the grounds of the emanations should establish the latter on the basis of the former" (*DL* 9.89).

The Modes lend themselves to use in combination. Take any dogmatic proposition *p*: one may ask what it is supposed to rest on. If the answer is "nothing," then it is a mere hypothesis, unworthy of credence by the Fourth Mode. If it is alleged to rest on *q*, one may ask the same question of *q*. If one gets the same answer, the same response applies. If *q* is said to rest on *p*, then the Mode of Circularity comes in; or else the process goes on, potentially *ad infinitum* in line with the Second Mode (*PH* 169–174). Credit for seeing the force of such objections is not due to Agrippa. Aristotle was aware of them (*Posterior Analytics* 1.3 [Barnes, ed. 1984]), and realized that any foundationalist epistemology requires its basic propositions to be more than mere assumptions. But how that is to be done—if it is to be done at all—is still a matter of dispute, apparently undecidable. Agrippa fashioned a powerful and elegant arsenal of skepticism, and all modern nonskeptical epistemologies sooner or later must confront them, and the challenge they pose, in one form or another.

See also Aenesidemus; Ancient Skepticism; Sextus Empiricus.

Bibliography

Annas, Julia, and Jonathan Barnes, eds. *Outline of Scepticism*. Translation of Sextus Empiricus's *Outlines of Pyrrhonism*. Cambridge, U.K.: Cambridge University Press, 1994.

Barnes, Jonathan. *The Toils of Scepticism*. Cambridge, U.K.: Cambridge University Press, 1989.

Barnes, Jonathan, ed. *The Complete Works of Aristotle*. 2 vols. Princeton, NJ: Princeton University Press, 1984.

Hankinson, R. J. *The Skeptics*. London: Routledge, 1995.

Hicks, R. D., ed. and trans. *Diogenes Laertius: Lives of the Eminent Philosophers.* 2 vols. Cambridge MA: Harvard University Press, 1925.

R. J. Hankinson (2005)

AGRIPPA VON NETTESHEIM, HENRICUS CORNELIUS

(1486–1535)

Henricus Cornelius Agrippa von Nettesheim, a colorful Renaissance figure—a diplomat, a military adventurer, a kabbalist, an expert on occult science, a medical doctor, a lawyer, a theologian, an early Reformer, as well as a troublesome and troubled intellectual—was born of minor nobility in or near Cologne. His first official position was that of a court secretary of the Holy Roman emperor. He was sent to Paris in 1506 and there joined a secret group of theosophists. He next became involved in a revolutionary plot in Catalonia. In 1509 he gave lectures at the University of Dôle, on Johannes Reuchlin's kabbalistic *De Verbo Mirifico*. He learned Hebrew and immersed himself in kabbalistic, Gnostic, and hermeneutic writings. This research culminated in three volumes on occult science, *De Occulta Philosophia*, written in 1509–1510 but not published until 1531–1533 in Cologne (trans. by J. F., London, 1651). At Dôle he also wrote on the superiority and nobility of women and entered into his first marriage. These early unpublished writings touched off a fight between Agrippa and certain conservative monks, who accused him, along with Reuchlin, Desiderius Erasmus, and the French humanist–Reformer Jacques Lefèvre d'Etaples, of being Judaizers and heretics.

In 1510 Agrippa was sent to London, where he lived with Erasmus's friend John Colet, who interested Agrippa in St. Paul's epistles. Next, Agrippa lectured on theology in Cologne. From 1511 to 1513 he fought in various Italian campaigns and engaged in theological battles, even with the pope. In 1515 he taught occult science at the University of Pavia. Three years later Agrippa became public advocate and orator of Metz and was soon embroiled again in theological battles and in defending a peasant woman accused of sorcery. The opposition of the inquisitor of Metz forced him to leave. Agrippa's wife died soon after, and he retired to Geneva. In 1522 he remarried and became a medical practitioner. He was appointed physician to the queen mother of France and became involved in a demoralizing struggle to collect his salary and to fulfill his duties. At the queen mother's orders he was stranded in Lyons from 1524 to 1526 without funds and without permission to leave. Agrippa wrote many bellicose letters to the court, antagonizing numerous people but settling nothing. His only official duty was the drawing up of horoscopes (which he knew were useless and fraudulent). In this period Agrippa wrote his major work, *De Incertitudine et Vanitate de Scientiarum et Artium* (Antwerp, 1530; trans, by James Sandford as *Of the Vanitie and uncertaintie of artes and sciences*, London, 1569), attacking every type of intellectual endeavor and art, as well as courtiers, princes, and monks. Even kabbalistic and occult researches were disowned as superstitious rhapsodies. Only pious Bible study remained worthwhile.

Agrippa abandoned hope of regaining court favor or receiving his salary and in 1528 went to Antwerp, where he had a brief flurry of success. He was appointed historiographer to Charles V, achieved success as a medical doctor, and finally published his works. This happy phase was soon followed by catastrophes. His second wife died of the plague. The publication of his *Vanity of the Sciences* outraged Charles V. Agrippa was jailed and branded a heretic. A disastrous marriage left him financially ruined and miserable. He returned to Germany, battled with the inquisitor of Cologne, and was banished in 1535. Having fled to France, he was arrested for having criticized the queen mother, was released, and died in Grenoble.

Agrippa was notorious as a magician and as a stormy opponent of the monks and the "establishment." He made his main intellectual contributions as an expositor of kabbalism and occult science, as a critic of all intellectual activities, and as a Reformer within Catholicism. His *De Occulta Philosophia* tried to explain the universe in terms of kabbalistic analyses of Hebrew letters and their relations to natural phenomena and divine understanding; in terms of the Pythagorean numerological symbols; and of the Christian interpretation of kabbalism and Pythagoreanism. *De Occulta Philosophia* played a major role in Renaissance magical and kabbalistic studies.

Agrippa's *Vanity of the Sciences* was one of the first contributions to the Renaissance revival of skepticism, but its weapons were denunciation and ridicule, not philosophical analysis. It is more a bitter version of Eras-

mus's *In Praise of Folly* than a serious epistemological examination of whether knowledge can be gained by human means. Its final appeal is to a type of fundamentalistic anti-intellectualism. The work represents a stage in Agrippa's journey from occult studies to a simple biblical faith opposed to late medieval Scholasticism. Agrippa, although he did not revolt against Catholicism, lacked Erasmus's patience and calm and became almost a Catholic Martin Luther, violently denouncing monks, Scholastic theologians, and others. In the end he rejected occult studies—and all others—as a way of penetrating the divine mysteries, and he proclaimed: "It is better therefore and more profitable to be idiots and know nothing, to believe by Faith and Charity, and to become next unto God, than being lofty and proud through the subtilties of sciences to fall into the possession of the Serpent."

See also Colet, John; Erasmus, Desiderius; Gnosticism; Hermeneutics; Kabbalah; Luther, Martin; Medieval Philosophy; Pythagoras and Pythagoreanism; Renaissance; Skepticism, History of.

Bibliography

WORKS BY AGRIPPA

Die Eitelkeit und Unsicherheit der Wissenschaften und die Verteidigungschrift, edited by Fritz Mauthner. Munich: Müller, 1913.

De occulta philosophia, edited by Karl Anton Nowotny. (Faksimile des ältesten Kölner Druckes, 1533.) Graz, Akademische Druck u. Verlagsanstalt, 1967.

Opera. Hildesheim, NY: G. Olms, 1970.

Three Books of Occult Philosophy or Magic. Book One: Natural Magic which Includes the Early Life of Agrippa, His Seventy-four Chapters on Natural Magic, New Notes, Illustrations, Index, and Other Original and Selected Matter. New York: S. Weiser, 1971.

Of the Vanitie and Uncertaintie of Artes and Sciences. Northridge: California State University, 1974.

Declamatio de nobilitate et praecellentia foeminei sexus edited by Elisabeth Gössmann. München: Iudicium. 2nd ed., überarbeitete und erw. Aufl., 1996.

WORKS ON AGRIPPA

Bayle, Pierre. "Agrippa." In *Dictionnaire historique et critique*. 2 vols. Rotterdam, 1695–1697.

Blau, Joseph. *The Christian Interpretation of the Cabala in the Renaissance*. New York: Columbia University Press, 1944.

Morley, Henry. *The Life of Henry Cornelius von Nettesheim*. 2 vols. London, 1856.

Nauert, Charles G. *Agrippa and the Crisis of Renaissance Thought*. Urbana: University of Illinois Press, 1965.

Nauert, Charles G. "Magic and Skepticism in Agrippa's Thought." *Journal of the History of Ideas* 18 (1957): 161–182.

Popkin, R. H. *History of Scepticism from Erasmus to Descartes*. Assen, 1960; New York: Humanities Press, 1964.

Prost, Auguste. *Les Sciences et les arts occultes au XVIe siècle: Corneille Agrippa, sa vie et ses oeuvres*. 2 vols. Paris: Champion, 1881–1882.

Walker, D. P. *Spiritual and Demonic Magic from Ficino to Campanella*. London: Warburg Institute, University of London, 1958.

Yates, Frances A. *Giordano Bruno and the Hermetic Tradition*. London, 1964.

Zarathustra, Frater. *Index to Agrippa's Occult Philosophy or Magic. Book 1, Natural magic*. Fremont, CA: Technology Group, 1993.

Richard H. Popkin (1967)
Bibliography updated by Michael J. Farmer (2005)

AILLY, PIERRE D'
(1350–1421)

Pierre d'Ailly, the Ockhamist philosopher, was born at Compiègne in France. He studied at the Navarre College in Paris in 1372, receiving his doctoral degree in 1380 and becoming chancellor of the university in 1389. He was made bishop of Puy in 1395 and bishop of Cambrai in 1396 and cardinal in 1411. He took a leading part in the Council of Constance (1414–1418), where he asserted the superiority of a general council of the church over the pope. He died as papal legate at Avignon.

D'Ailly's literary output was vast and wide-ranging. It comprehended philosophy, theology, scientific theory, political theory, canon law, and ecclesiastical politics and touched on mysticism. Among his more important writings were the treatise *De Anima*, commentaries on Boethius's *Consolation of Philosophy* and the four books of the *Sentences*, two studies of mysticism and asceticism, three works on different aspects of church government, and a series of works on logic, astronomy, and geography.

In his philosophical outlook d'Ailly seems to have been sympathetic to Ockhamism. Like so many fourteenth-century thinkers he postulated different degrees of certainty. The main distinction d'Ailly made was between what he called "natural light" and reason. Natural light corresponded to knowledge that was indubitable—namely, that which could be reduced to the principle of contradiction or immediate intuition of the existence of the self, in the manner of John of Mirecourt. Reason, on the other hand, was only relative in its certainty and was confined to the natural order. Included within it were the traditional arguments for God's existence, which d'Ailly treated as merely probable. The influence of William of Ockham is also apparent in d'Ailly's

treatment of God's omnipotence; since it was independent of the natural order, God was in no way bound to follow nature's laws. Accordingly, God could create the illusion that something existed when in fact it did not; this was one of the most insistent Ockhamist arguments against the infallibility of experiential knowledge. At the same time d'Ailly was careful to distinguish the realm of God's absolute power (*potentia absoluta*) from the realm subject to his ordained power (*potentia ordinata*). Whereas the first realm referred to God's omnipotence as such, the latter constituted the specific application of his omnipotence to this world; it provided the laws by which creation was regulated, and among them d'Ailly included the laws of physics. They therefore operated constantly and with certainty.

D'Ailly's debt to Ockham and John of Mirecourt is also to be seen in his views on essences. There was no inherent reason why hot was hot or cold cold other than God's willing it. The same applied to the moral order, where good and bad were such because of God's voluntary decree: "Nothing is good or bad of itself such that God must love or hate it." Similarly, a man was just not from possessing the intrinsic property of justice but because God accepted him as just. Here was the same absence of a constant scale of values that had proved so destructive of the traditional teachings in the time of Ockham and the first generation of his followers, who included Robert Holkot, Adam of Woodham, and John of Mirecourt. D'Ailly further emphasized the uncertain nature of natural experience by his acceptance of the so-called *complexe significabile*, by which an expression such as "sin" did not denote a specific object but was a description or statement that referred to an action. As employed by Nicolas of Autrecourt, it had denied the reality of a wide range of expressions. Thus the word *God* stood not for a specific being but for a verbal expression: supreme or highest being. As such it lacked correspondence to anything but a grouping of words. At the same time, in keeping at the natural level, d'Ailly granted a correspondingly wider area of jurisdiction to faith. Thus evidence for God's existence could be held only as a matter of belief.

See also Holkot, Robert; John of Mirecourt; Nicolas of Autrecourt; Ockhamism; William of Ockham; Wodeham, Adam.

Bibliography

WORKS BY D'AILLY

Tractatus Exponibilium Magistri Petri de Allyaco. Paris, 1494.

Quaestiones Super Primum, Tertium et Quartum Sententiarum. Venice, 1500.

De Anima. Paris, 1501.

Tractatus et Sermones. Douai, France, 1639. Includes *De Anima*.

Destructiones Modorum Significandi. Conceptus et Insolubilia Secondum Viam Nominalium Magistri Petri de Allyaco. (No date.)

WORKS ON D'AILLY

Dictionnaire de théologie catholique. Vol. I, 642–654. Paris, 1903–1951.

Gandillac, M. P. de. "Usage et valeur des arguments probables chez Pierre d'Ailly." *Archives d'histoire doctrinale et littéraire* 8 (1933): 43–91.

Gordon Leff (1967)

AITIA

The Greek word *aitia* (or *aition*) derives from the adjective *aitios*, meaning "responsible," and functions as such as early as the Homeric poems. It was originally applied to agents, and only later does it come to qualify nonsentient items—although owing to the fragmentary nature of earlier sources, it is by no means clear when this transition takes place. But certainly by the latter part of the fifth century BCE, Hippocratic doctors were using the term, as were the historians Herodotus and Thucydides. It is in the latter, as well as in some of the Hippocratic texts, that the beginnings of the distinction of causal terminology can be found. Similar fine distinctions are also beginning to appear in the forensic and rhetorical traditions. In his discussion of the plague at Athens (*Peloponnesian War* 2.47–54), Thucydides disavows any knowledge of its origins or "what causes (*aitiai*) may be adduced adequate to explain its powerful natural effects" (2.48), and notes that "in some cases there seemed to be no *prophasis*" (2.49). A *prophasis* is an external cause, or occasion, or antecedent event correlated with an outcome. This word, as well, has Homeric roots, but it also has a legal (and more general) sense of pretext. Hippocratic texts also contrast *prophaseis* with *aitiai*, and in the same general way: *Prophaseis* are the observable antecedent signs, *aitiai* the inferred inner, structural facts causally responsible for the outcome. *Aitiai* are now closely linked with the notion of *phusis* or nature, the primary matter of pre-Socratic inquiry. If things have natures—internal structures—then those natures will explain how and why things behave the way they do.

Plato was the first philosopher to subject the concept of an *aitia* to detailed examination. Whereas generally an *aition* is "that because of which something comes to be"

(*Cratylus* 413a), and "the cause and the productive may rightly be said to be identical" (*Philebus* 26e), Plato treats these characterizations generally—they do not restrict causation to efficient causation. Indeed, at *Phaedo* 95e–103b, he takes the pre-Socratics to task for concentrating on mechanical causation at the expense of teleology: It is only if you know why things are for the best that you understand them. Moreover, Plato elaborates a thesis of necessity and sufficiency with regard to cause and effect (or *explanans* and *explanandum*): If *F*'s cause *G*'s, then there is no *F* without a *G*, and vice versa.

Aristotle followed Plato in espousing teleological explanations, and referring to them in the language of *aitiai*. Final causes are one of his four causal (or explanatory) types, along with material, efficient, and formal (*Physics* 2.3). But unlike Plato, Aristotle's final causes in nature presuppose no agency. Where Plato spoke of the Craftsman who designed everything for the best (*Timaeus*), Aristotle makes finality an irreducible component of nature itself. Nature is goal-directed, and no adequate account of natural processes can ignore that fact (as those of the atomists and other mechanists do). As Plato had before him, Aristotle thinks explanatory resources available to pure mechanism are inadequate to give a satisfying explanation of the order and regularity of the cosmos. The four causes are designed primarily to account for substances, and only derivatively for events and processes. Thus one might ask what makes an oak tree what it is. Firstly, its efficient cause—namely its parent tree, which supplies the formal model from which it derives. Secondly, its material cause: There could be no oak tree without a suitable supply of matter for the form to mold. Thirdly, there is the form itself, deriving from the efficient cause—yet from the moment the seed is created (or at least begins to germinate) it is an independent structural principle. And finally there is the end—or completely expressed form—toward which the process of maturation is directed and in which it will culminate if all other (material) factors equal.

Aristotle seeks to apply this model, with varying success, to all cases of coming to be (although he allows that coincidences lack final causes: *Physics* 2.4–6); and that all of the factors involved may equally be called *aitiai*. Moreover, he believes that even abstract objects have formal causes (the formal cause of the octave is the ratio 2:1). Plato's Stoic successors, however, reserved the term *aition* for a physical productive cause—a body that brought about in another body an incorporeal effect, a predicate's coming to be true of it. They allowed matter a role in overall explanation, yet being passive by definition it could not be a cause; neither could it be disembodied goals or ends. These Stoic successors, or their contemporaries in the medical schools, turned to making further fine distinctions within the notion even more restricted, distinguishing between "perfect" or "sustaining" causes on the one hand and "antecedent" causes on the other. "Perfect" or "sustaining" causes were sufficient, necessary, and coterminous with their effects—and functionally correlated with them—in that any increase or decrease in intensity in the one is matched by a similar change in the other, "antecedent" causes answered roughly to the earlier *prophaseis*: prior events that set a causal process in train but are not sufficient for it (since they require suitably constituted bodies upon which to act).

Skeptics were to argue that causes could not both precede and be coterminous with their effects; and that because cause and effect are relative terms, they cannot be conceived independently, as they must be if one is to explain the other. This and other such attacks in turn prompted doctors and philosophers such as Galen to even further conceptual refinements that continued at least until the third century CE, whereas Neoplatonists like Proclus would later insist that, properly speaking, all causes were immaterial (being the action of soul).

See also Aristotle; Causation; Plato; Stoicism.

Bibliography

Barnes, J., ed. *The Complete Works of Aristotle*. 2 vols. Princeton, NJ: Princeton UIniversity Press, 1984.

Cooper, J. M., ed. *Plato: Complete Works*. Indianapolis: Hackett, 1997.

Hankinson, R. J. *Cause and Explanation in the Ancient Greek World*. Oxford: Oxford University Press, 1998.

Warner, R. *Thucydides: History of the Peloponnesian War*. Harmondsworth, U.K.: Penguin Books, 1953.

R. J. Hankinson (2005)

ALBERT OF SAXONY

(c. 1316–1390)

Sometimes nicknamed Albertucius to distinguish him from Albert the Great, Albert of Saxony was born in Rickensdorf in the region of Helmstedt (Lower Saxony), in present-day Germany. He did his early studies in his native region, then most likely took a trip to Elfurt. He later went to Prague and Paris, where he earned his master of arts degree in 1351. He was rector of the university in 1353. He taught the arts there for a decade, while studying theology at the Sorbonne, apparently without

earning a degree. After a few years as a diplomat mediating between Pope Urban VI and the Duke of Austria, he was called on to found the University of Vienna, becoming its first rector in 1365. He was appointed canon of Hildesheim in 1366 and became bishop of Halberstadt the same year. He served in that capacity until his death on July 8, 1390.

Albert of Saxony left behind no theological writings and is known primarily for his works in logic and natural philosophy. He also composed commentaries on Aristotle's *Nicomachean Ethics* and *Economics*, as well as a short mathematical treatise on the squaring of the circle.

In logic, his masterwork is a summa titled *Perutilis logica* (Very useful logic). He also composed a voluminous collection, the *Sophismata* (Sophisms), in which he examined many statements that raise difficulties of interpretation because they contain syncategoremes. In addition, he wrote *Commentarius in Posteriora Aristotelis* (Commentary on Aristotle's *Posterior Analytics*) and a collection of twenty-five questions (*Quaestiones circa logicam*) relating to semantic problems or the status of logic. He also commented on Aristotle's writings in logic. During Albert of Saxony's career, Jean Buridan enjoyed great renown at the faculty of arts in Paris. Albert's writings, however, attest to the influence exerted in Paris by English ideas. His *Very Useful Logic*, while developing treatises on obligations, insolubles, and consequences—topics that were becoming increasingly important during the period—was modeled after William of Ockham's *Summa logicae* (Summa of logic). Albert of Saxony adopted the Ockhamist conception of the sign and based signification on a referential relationship to a unique thing. He also subordinated the oral sign to the conceptual sign. He was an Ockhamist in his conception of the universal and, for the most part, in his theory of the supposition. In particular, he retained the notion of the simple supposition—that is, the reference of a term to a concept to which it is subordinated, even though it signifies an extra-mental thing. Finally, he was Ockhamist in his theory of categories. Unlike Jean Buridan, he refused to consider quantity an absolute reality and relegated it to a disposition of the substance and the quality.

On a few points, however, Albert departed from William of Ockham. Hence he rejected the idea that an ambiguous proposition ought to be assigned multiple meanings. Such a proposition can only be conceded, rejected, or called into doubt. In the *Sophisms*, William Heytesbury often served as his guide (for example, in the analysis of epistemic verbs and the study of the infinite). He grants the proposition a literal meaning, which is not that of its terms. Like the syncategoreme, the proposition signifies a "mode of being." Nevertheless, Albert of Saxony avoided accounting for these "modes of being" and, in the last analysis, transferred them to relationships between the things to which the terms referred. But he used the idea of a proposition's meaning to define truth and to deal with "insolubles," that is, semantic paradoxes. By virtue of its form, every proposition signifies that it is true; for that reason, the insoluble is false, since it signifies both that it is true and that it is false.

This analysis of language was combined with a gnoseological realism that stemmed in part from an analysis of the void. It is possible to imagine that the void exists by divine omnipotence, but no science of nature can integrate the existence of the void as a hypothesis. Albert refused to extend the referent of physics terms to supernatural possibilities. For him, physics cannot develop into a study of imaginary cases, despite what was being done at Oxford at the time. It must account for the natural course of things.

In addition to commentaries on Aristotle's *Physica* (Physics), Albert composed a commentary on Johannes de Sacrobosco's *De sphaera* (On the sphere) and a treatise on relationships inspired by Thomas Bradwardine. In pursuing the work of the Oxford Calculators and of Nicole d'Oresme in Paris, he created a compendium setting out the elements of the theory of relationships and their application to different motions, adopting the rule elaborated by Bradwardine on the relationship between powers of propulsion and resistance. In his physics texts, he also displayed a curiosity for many natural phenomena, taking an interest in motions of the earth, tides, and geology.

It was undoubtedly in the field of dynamics, however, that Albert's role was most important. To account for the motion of projectiles and the acceleration of falling objects, he adopted the Buridanian theory of impetus, a quality acquired by the body. He drew clearly the consequences of extending that notion to celestial movements, rejected the notion of a propulsive intelligence, and followed the same principles in studying celestial bodies and earthly bodies. His commentary on Aristotle's *De caelo* (On the sky) exerted a great influence in northern Italy. Albert of Saxony thus played a role in developing a vision of the cosmos that departed from conceptions inherited from Greco-Arab Peripateticism.

See also Bradwardine, Thomas; Buridan, Jean; Heytesbury, William; Impetus; Oresme, Nicholas; William of Ockham.

Bibliography

WORKS BY ALBERT OF SAXONY

Perutilis logica. Maracaibo: Universidad del Zulia, 1988. Based on the incunabulum Venice edition of 1522. Critical edition with Spanish translation by Angel Muñoz García.

Perutilis logica, Tractatus II (De proprietatibus terminorum). Leiden: E. J. Brill, 1993.

Quaestiones in artem veterem. Maracaibo: Universidad del Zulia, 1988. Critical edition with Spanish translation by Angel Muñoz García.

"Quaestiones circa logicam." In *Albert of Saxony's Twenty-Five Disputed Questions on Logic, a Critical Edition of His "Quaestiones circa logicam,"* edited by Michael Fitzgerald. Leyden: E. J. Brill, 2002.

Expositio et Quaestiones in Aristotelis Physicam ad Albertum de Saxonia attributae, edited by Benoît Patar. 3 vols. "Philosophes médiévaux" 39–41. Louvain-la-Neuve: Peeters, 1999.

WORKS ABOUT ALBERT OF SAXONY

Biard, Joël. "Albert de Saxe et les sophismes de l'infini." In *Sophisms in Medieval Logic and Grammar*, edited by Stephen Read, 288–303. Dordrecht: Kluwer, 1993.

Biard, Joël. "Les sophismes du savoir: Albert de Saxe entre Jean Buridan et Guillaume Heytesbury." *Vivarium* 27 (1989): 36–50.

Gonzales, A. "The Theory of Assertoric Consequences in Albert of Saxony." *Franciscan Studies* 18 (1958): 290–354; 19: 13–114.

Heidingsfelder, Georg. *Albert von Sachsen. Sein Lebensgang und sein Kommentar zur Nikomachischen Ethik des Aristoteles*. "Beiträge zur Geschichte der Philosophie des Mittelalters" 22, 3–4. Münster: Aschendorff, 1927. The biographical part is now dated, but the study on ethics remains valid.

Kann, Christoph. *Die Eigenschaften der Termini. Eine Untersuchung zur "Perutilis Logica" des Alberts von Sachsen*. Leiden: E. J. Brill, 1993.

Sarnowsky, Jürgen. "Place and Space in Albert of Saxony's Commentaries on the *Physics*." *Arabic Sciences and Philosophy* 9 (1999): 25–45.

Joël Biard (2005)

ALBERT THE GREAT

See Appendix, Vol. 10

ALBINUS

See *Alcinous*

ALBO, JOSEPH

(c. 1380–c. 1444)

The Spanish-Jewish preacher and philosopher Joseph Albo was the last major figure of the philosophical surge in medieval Jewry. Little is known about his early life; he was probably born at Monreal, in the kingdom of Aragon, and he asserted that Hasdai Crescas was his teacher. Albo was one of the principal apologists for the Jews at the Colloquium of Tortosa (February 7, 1413–November 3, 1414); his activities as apologist and preacher are reflected in the style of his philosophic classic, *Sefer ha-'Ikkarim* (*The Book of Roots*).

Albo's acknowledged and unacknowledged borrowings from other writers are so extensive that he was accused of plagiarism in his own age, as well as in more recent and more sensitive times. We must recognize, however, that Albo's purpose was to systematize and thus to defend the dogmas of Judaism rather than to produce an original philosophic work. Clarity and lucidity, systematic and easily remembered organization of materials, and simple and uninvolved style of presentation have made Albo's *The Book of Roots* one of the most popular works of medieval Hebrew literature. Indeed, it was one of the earliest printed Hebrew books, the first edition having been issued at Soncino, Italy, in 1485. Albo's occasional use of medical materials to illustrate his thought has suggested to critics that he may have been trained as a physician. He was well trained in Jewish philosophy, and in addition he knew, probably at second hand, the works of the Arabic Aristotelians.

Albo asserted that there are three essential dogmas ("roots") of Judaism: the existence of God, revelation, and reward and punishment. Seven secondary principles were derived from these three. The existence of God yields four: his unity, his incorporeality, his timelessness, and his perfection. From the dogma of revelation Albo derived two secondary principles: the prophets were the medium of revelation, and the Mosaic law will have binding force until another law is proclaimed with equal publicity; that is, before 600,000 men. God's providential knowledge in the matter of retribution was, for Albo, the sole secondary derivative from the doctrine of reward and punishment. Beyond these primary and secondary roots are other logically derived "branches" that every professing Jew must believe or be guilty of heresy, among them the doctrine of the Messiah.

It may be presumed that Albo removed the doctrine of the Messiah from the center of the Jewish faith as an important part of his polemic against Christianity, a

recurrent feature of *The Book of Roots.* As an aspect of this polemic, Book III, Chapter 25 contains an actual summary of a disputation between a Jew and a non-Jew (omitted in some editions).

See also Crescas, Hasdai; Jewish Philosophy.

Bibliography

BILINGUAL EDITION

Husik, Isaac. *Sefer ha-'Ikkarim,* 5 vols. Philadelphia: Jewish Publication Society, 1929–1930. Critical edition of the Hebrew text, with facing English translation.

DISCUSSIONS

Agus, Jacob B. *The Evolution of Jewish Thought.* London and New York: Abelard-Schuman, 1959.

Blau, Joseph L. *The Story of Jewish Philosophy.* New York: Random House, 1962.

Guttmann, Julius. *Philosophies of Judaism.* Translated by D. W. Silverman. New York: Holt, Rinehart and Winston, 1964.

Husik, Isaac. *History of Mediaeval Jewish Philosophy.* New York: Macmillan, 1916.

J. L. Blau (1967)

ALBO, JOSEPH [ADDENDUM]

Albo remains one of the lesser-studied philosophers of the medieval period, in part because his main work more apologetic than philosophical in nature. No full-length monograph has been written on him; rather, he is the subject of scattered articles on diverse topics. Not surprisingly, the most systematic work has been done on his dogmatics, with the place of dogma in Judaism generally arousing a measure of philosophical interest in the late twentieth and early twenty-first centuries. Albo follows Simeon ben Tzemach Duran (1361–1444) in reducing Maimonides's thirteen principles of faith to three—with eight, not seven, derivative principles: revelation yielding (1) God's knowledge; (2) prophecy; and (3) the authenticity of the divine messenger. Menachem M. Kellner, however, has argued that in his portrayal of Torah as having the axiomatic structure of a deductive science, Albo is the first to present the commandments rather than metaphysics as embodying this scientific structure. Support for this view may be found in Albo's account of the six nonessential beliefs, or branches, particular to Jewish dogmatics, which are not strictly entailed by the earlier fundamental or derivative principles, even though their denial constitutes heresy.

A further significant strand in Albo's thought in which this emphasis on practice can also be found is the shift from the intellectualism of the Aristotelians—for whom intellectual apprehension was the path to perfection—to an act-based theology in which acts are even at one point referred to as knowledge.

The particular scientific topic of time has been subject to detailed analysis by Warren Zev Harvey. Whereas Albo follows his teacher Crescas in asserting that time is independent of motion and therefore of the physical world, Harvey argues that Albo is the first to state that time is an imagined duration rather than one that is intellectually cognized. This is significant for Albo as a foundation for one of his derivative principles—that God is independent of time—though Harvey argues that the links here are not demonstrated. As with all of the above, Albo's remarks here are suggestive, but left underdeveloped.

See also Aristotelianism; Crescas, Hasdai; Dogma; Jewish Philosophy; Maimonides.

Bibliography

Harvey, Warren Zev. "Albo's Discussion of Time." *The Jewish Quarterly Review* 70 (4) (1980): 210–238.

Kellner, Menachem M. "The Conception of the Torah as a Deductive Science in Medieval Jewish Thought." *Revue des etudes juives* 146 (1987): 265–279.

Manekin, Charles H. "Hebrew Philosophy in the Fourteenth and Fifteenth Centuries: An Overview." In *History of Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman. New York: Routledge, 1997.

Shatz, David. "Freedom, Responsibility and Hardening of the Hearts: Albo *vs* Maimonides." *Faith and Philosophy* 14 (4) (1997): 478–509.

Sirat, Colette. *A History of Jewish Philosophy in the Middle Ages*, 374–381. Cambridge, U.K.: Cambridge University Press, 1985.

Daniel Rynhold (2005)

ALCINOUS
(fl. c. 150 CE)

Alcinous is the name that has come down to us in the manuscript tradition as the author of a handbook of Platonic doctrine, *Didaskalikos tôn Platonos dogmatôn*, probably from the second century CE. Following an 1879 suggestion by the German scholar Jacob Freudenthal, this figure was identified for more than a hundred years with the Middle Platonist philosopher known as Albinus, but this is now recognized to have been based on unsound

assumptions, and the work has been returned to its shadowy author. The *Didaskalikos* has much in common with another second-century handbook of Platonism, the *De Platone et eius dogmate* of Apuleius, but the similarities are not close enough to indicate that they emanate from the same school.

The *Didaskalikos* is an introduction to Platonism as it was understood in the first and second centuries CE, which means that it exhibits an amalgam of Platonic, Aristotelian, and Stoic formulations and doctrines, presented as clarifications and amplifications of Plato's views. Aristotelian influence is particularly to be seen in the sphere of logic, where the whole system of syllogistic is claimed for Plato; Stoic influence may be seen chiefly in ethics, in relation to the doctrine that virtue is sufficient for happiness. In either case the assumption is that Aristotle and the Stoics are only expounding Platonic doctrine.

The work is divided as follows. After three introductory chapters, concerned with the definition of philosophy and the distinction of its "parts"—physics, ethics, and logic—Alcinous proceeds to take these three topics in order, beginning with logic (chaps. 4–6), then turning to "physics" (chaps. 7–26), comprising both an account of first principles, Matter, Form, and God (chaps. 7–11), and then of the physical world, largely based on Plato's *Timaeus*, but also including discussions of the immortality of the soul, and of fate and free will (chaps. 12–26); and finally ethics, covering such topics as the virtues, happiness, the purpose of life (*telos*)—which he characterizes as "likeness to God"—the emotions, and political theory (chaps. 27–34). The work ends with a disquisition on the difference between the philosopher and the Sophist (chap. 35), and a brief conclusion (chap. 36).

While the *Didaskalikos* is not securely datable, there is nothing in it that cannot be seen as "Middle Platonist" in doctrine. Even the discussion of God in chapter 10, which has many intriguing aspects, including a distinction between a supreme God, a cosmic Intellect and a World Soul, can be accommodated within Middle Platonist parameters.

See also Aristotle; Plato; Stoicism.

Bibliography

Dillon, John M. *The Middle Platonists: A Study of Platonism, 80 B.C. to A.D. 220*. London: Duckworth, 1977. Rev. ed. Ithaca, NY: Cornell University Press, 1996.

Dillon, John M., trans. *Alcinous: The Handbook of Platonism*. Oxford, U.K.: Clarendon Press, 1993.

Giusta, M. *I dossografi di etica*. 2 vols. Turin: Giappichelli, 1967.

Whittaker, John. *Alcinoos: Enseignement des doctrines de Platon*. Budé ed. Paris: Belles lettres, 1990.

John Dillon (2005)

ALCMAEON OF CROTON

(c. 540–500 BCE)

Alcmaeon of Croton (a Greek city-state in southern Italy) was a pioneer in the study of human psychology and physiology. He published one book in the late sixth or first half of the fifth century BCE. Only two or three fragments of the book survive, but substantial reports of his views are preserved in authors such as Theophrastus. It is controversial whether Alcmaeon was primarily a physician and medical writer or whether he dealt with physiological issues as part of a typical pre-Socratic account of the cosmos. Beginning in the second century CE, some authors call him a Pythagorean, but the earliest sources do not. Aristotle appears to distinguish him from the Pythagoreans (*Metaph*. 986a22).

Alcmaeon is the earliest author to state the common ancient view that health depends on a balance of opposed powers in the body. Just as Anaximander used a political analogy to explain the workings of the cosmos, Alcmaeon said that "equality (*isonomia*) of powers (wet, dry, cold, hot, bitter, sweet, etc.) maintains health, but monarchy among them produces disease" (Fr. 4 Diels-Kranz). Alcmaeon may have excised an eyeball and observed passages (*poroi*—i.e., the optic nerve) leading from the eye toward the brain. Perhaps as a result of this observation, he was the first person in the Greek tradition to argue that the brain was the seat of thought. There is no evidence, however, that he used dissection to any further extent or that he practiced it systematically. He was the first to address a series of issues that would become standard in later writings on physiology, such as the causes of sleep, waking, and death. He argued that human seed came from the brain, that the brain was the first part of the embryo to develop and that both parents contributed seed in the production of children.

In contrast to the wealth of evidence for Alcmaeon's views on human physiology, the evidence for his cosmological views is sketchy. He may have believed that the cosmos, like the human body, arose from a balance of opposing powers. He also maintained that the sun was flat.

Alcmaeon argued that there was no human knowledge of what is not perceptible and that judgments about what is not perceptible can only be made on the basis of what is perceived. He was the first to make a clear distinction between animals, which only have sense perception, and human beings, who also have understanding. Alcmaeon may have originated the three-step empiricist epistemology found in both Plato (*Phaedo* 96a–b) and Aristotle (*Posterior Analytics* 100a3) that begins with sensations, which when collected become memories and opinions, which in turn become knowledge when they gain fixity. Finally, Alcmaeon gave the first argument for the immortality of the soul. The exact nature of Alcmaeon's argument is hard to reconstruct, because it was later developed by Plato in the *Phaedrus* (245c), but he appears to have argued that the soul was immortal because it was in constant motion.

See also Anaximander; Aristotle; Plato; Psychology; Pythagoras and Pythagoreanism; Theophrastus.

Bibliography

TEXTS AND COMMENTARIES

Diels, H., and W. Kranz. *Die Fragmente der Vorsokratiker.* 6th ed., vol. 1, 210–216. Dublin, Ireland: Weidmann, 1952. Contains the Greek texts with translations in German.

Wachtler, J. *De Alcmaeone Crotoniata*. Leipzig, Germany: Teubner, 1896. Contains Greek texts with commentary in Latin.

DISCUSSIONS

Lloyd, G. E. R. "Alcmeon and the Early History of Dissection." *Sudhoffs Archiv* 59 (1975): 113–147.

Mansfeld, J. "Alcmaeon: 'Physikos' or Physician." In *Kephalaion: Studies in Greek Philosophy and its continuation offered to Professor C. J. de Vogel*, edited by J. Mansfeld and L. M. de Rijk, 26–38. Assen, Netherlands: Van Gorcum, 1975.

Vlastos, Gregory. "Isonomia." *American Journal of Philology* 74 (4) (1953): 337–366.

Carl Huffman (2005)

ALEMBERT, JEAN LE ROND D'

(1717–1783)

The French mathematician and encyclopedist Jean Le Rond d'Alembert was the illegitimate son of Madame de Tencin and the artillery general Destouches-Canon. He was abandoned by his mother on the steps of the baptistry of Saint-Jean-Le-Rond in Paris, from which he received his name. Shortly afterward his father returned from the provinces, claimed the child, and placed him with Madame Rousseau, a glazier's wife, with whom d'Alembert remained until a severe illness in 1765 forced him to seek new quarters. Through the Destouches family, Jean Le Rond was placed in the exclusive Jansenist Collège de Mazarin and given the name of d'Aremberg, which he later changed to d'Alembert, no doubt for phonetic reasons. At the college an effort was made to win him over to the Jansenist cause, and he went so far as to write a commentary on St. Paul. The intense Jesuit-Jansenist controversy served only to disgust him with both sides, however, and he left the college with the degree of bachelor of arts and a profound distrust of, and aversion to, metaphysical disputes.

After attending law school for two years he changed to the study of medicine, which he soon abandoned for mathematics. His talent and fascination for mathematics were such that at an early age he had independently discovered many mathematical principles, only to find later that they were already known. In 1739 he submitted a *mémoire* on integral calculus to the Académie des Sciences, but it was his *Traité de dynamique* in 1743 that won him acclaim and paved the way for his entry into the academy that same year. The introduction to his treatise is significant as the first enunciation of d'Alembert's philosophy of science. He accepted the reality of truths rationally deduced from instinctive principles insofar as they are verifiable experimentally and therefore are not simply aprioristic deductions. Although admitting unproved axioms at the base of his principles of mechanics, thus revealing his debt to René Descartes, d'Alembert rejected metaphysical affirmations and the search for universals and expressed admiration for Bacon's experimental and inductive method.

The decade of the 1740s may be considered d'Alembert's mathematical period during which he made his most outstanding and fruitful contributions to that discipline. In addition to the *Traité de dynamique* he wrote *Mémoire sur la réfraction des corps solides* (1741); *Théorie de l'équilibre du mouvement des fluides* (1744 and 1751); *Réflexions sur la cause générale des vents,* which won him the prize of the Berlin Academy in 1746 as well as membership in that body; a *mémoire* on vibrating strings (*Recherches sur les cordes vibrantes*), written in 1747 for the Berlin Academy; *Recherches sur la précession des équinoxes et sur la nutation* (1749); *Réflexions sur la théorie de la résistance des fluides* (1752); *Recherches sur differents points importants du systéme du monde* (1754–1756), plus eight volumes of *Opuscules mathématiques* (1761–1780).

D'Alembert's first philosophical work, the *Discours préliminaire* to the *Encyclopédie,* appeared in 1751. As early as 1746 he, with Denis Diderot, had been on the publisher's payroll as translator, in connection with the projected French version of Chambers's *Cyclopaedia.* We may suppose that, like Diderot, he had already worked for the publishers as a translator of English works for French consumption, thus exposing himself to the writings of the English empiricists and supplementing the meager pension left him by his father. In any event, d'Alembert had read Bacon as early as 1741; and his *Discours préliminaire* revealed not only his debt to the Descartes of the *Regulae,* shorn of metaphysics, but his admiration for, and indebtedness to, Bacon for his experimental method; Isaac Newton, whom he admired for proving gravitational force without trying to explain its first cause; and John Locke, whose metaphysical method he adopted. While paying lip service to the traditional religious concepts of his time, d'Alembert used Lockian sensationalist theory to arrive at a naturalistic interpretation of nature. It is not through vague and arbitrary hypotheses that nature can be known, he asserted, but through a careful study of physical phenomena. He discounted metaphysical truths as inaccessible through reason. In the *Discours,* d'Alembert began by affirming his faith in the reliability of the evidence for an external world derived from the senses and dismissed the Berkeleian objections as metaphysical subtleties that are contrary to good sense. Asserting that all knowledge is derived from the senses, he traced the development of knowledge from the sense impressions of primitive man to their elaboration into more complex forms of expression. Language, music, and the arts communicate emotions and concepts derived from the senses and, as such, are imitations of nature. For example, d'Alembert believed that music that is not descriptive is simply noise. Since all knowledge can be reduced to its origin in sensations, and since these are approximately the same in all men, it follows that even the most limited mind can be taught any art or science. This was the basis for d'Alembert's great faith in the power of education to spread the principles of the Enlightenment.

In his desire to examine all domains of the human intellect, d'Alembert was representative of the encyclopedic eighteenth-century mind. He believed not only that humanity's physical needs are the basis of scientific and aesthetic pursuits, but also that morality too is pragmatically evolved from social necessity. This would seem to anticipate the thought of Auguste Comte, who also placed morality on a sociological basis, but it would be a mistake to regard d'Alembert as a Positivist in the manner of Comte. If d'Alembert was a Positivist, he was so through temporary necessity, based on his conviction that since ultimate principles cannot be readily attained, one must reluctantly be limited to fragmentary truths attained through observation and experimentation. He was a rationalist, however, in that he did not doubt that these ultimate principles exist. In the *Discours préliminaire* he expressed the belief that everything could be reduced to one first principle, the universe being "one great truth" if we could only see it in a broader perspective. Similarly, in the realm of morality and aesthetics, he sought to reduce moral and aesthetic norms to dogmatic absolutes, and this would seem to be in conflict with the pragmatic approach of pure sensationalist theories. He was forced, in such cases, to appeal to a sort of intuition or good sense that was more Cartesian than Lockian, but he did not attempt to reconcile his inconsistencies and rather sought to remain within the basic premises of sensationalism. D'Alembert's tendency to go beyond the tenets of his own theories, as he did, for example, in admitting that mathematical realities are a creation of the human intellect and do not correspond to physical reality, has led Ernst Cassirer to conclude that d'Alembert, despite his commitment to sensationalist theory, had an insight into its limitations.

During the early 1750s d'Alembert engaged actively in the polemics of the time, particularly in the defense of the *Encyclopédie* and the party it represented. Many of the articles he wrote for that publication, as well as his preface to Volume III (1753), were aimed at the enemies of the *Encyclopédie,* notably the Jesuits, who were among the first to attack it for its antireligious and republican orientation. In addition, he took part in the controversy over French versus Italian music, which was inflamed by Jean-Jacques Rousseau's attack on French music in "Lettre sur la musique française" (1753). D'Alembert had already published his *Éléments de musique* (1752), based on Jean-Philippe Rameau's theories on harmonics, and in 1754 he published anonymously his *Réflexions sur la musique en général et la musique française en particulier.*

D'Alembert's chief preoccupation at this period, however, was with philosophy and literature. His *Mélanges de littérature et de philosophie* appeared in 1753 in two volumes (expanded to four volumes in 1759, with a fifth volume added in 1767), and it is here that his skepticism concerning metaphysical problems is delineated. Proceeding on the premise that certainty in this field cannot be reached through reason alone, he considered the arguments for and against the existence of God and cautiously concluded in the affirmative, on the grounds that

intelligence cannot be the product of brute matter. Like Newton, d'Alembert viewed the universe as a clock, which necessarily implies a clockmaker, but his final attitude is that expressed by Montaigne's "*Que sais-je?*" Humankind's uncertainty before this enigmatic universe is the basis of d'Alembert's plea for religious tolerance. He maintained his skeptical deism as an official, public position throughout his life, but there is evidence for believing that in the late 1760s, under the influence of Diderot (whose *Rêve de d'Alembert* appeared in 1769), d'Alembert was converted to Diderot's materialism. In private correspondence with intimate friends, d'Alembert revealed his commitment to an atheistic interpretation of the universe. He accepted intelligence as simply the result of a complex development of matter and not as evidence for a divine intelligence.

Aside from the publication of a polemical brochure, *Histoire de la destruction des Jésuites,* in 1765 (with two additional *Lettres* on the subject in 1767), d'Alembert spent the last two decades of his life in furthering the cause of the *philosophes* in the Académie Française—by writing his *Éloges,* which were read in the Académie (and published in 1779), and by fostering the election of candidates of his own choice. Mademoiselle de Lespinasse's salon, where d'Alembert presided, became, in the words of Frédéric Masson, the "obligatory antechamber of the Académie." In this period he became influential with young aspiring men of letters, whom he recruited for his party and whose careers he fostered. The most notable of his disciples was the Marquis de Condorcet. After years of ill health, d'Alembert died of a bladder ailment and was buried as an unbeliever in a common, unmarked grave.

See also Bacon, Roger; Cassirer, Ernst; Comte, Auguste; Condorcet, Marquis de; Descartes, René Diderot, Denis; Locke, John; Music, Philosophy of; Newton, Isaac; Rousseau, Jean-Jacques.

Bibliography

WORKS BY D'ALEMBERT

Oeuvres philosophiques et littéraires. Edited by J. F. Bastien. 18 vols. Paris, 1805. Not so complete as the Belin edition but contains letters to d'Alembert not included elsewhere. *Nouvelle édition augmentée.* Edited by A. Belin. 5 vols. Paris, 1821. The most complete edition to date.

Oeuvres et correspondance inédites de d'Alembert. Edited by Charles Henry. Paris, 1887. Contains important supplements to above editions in the fields of philosophy, literature, and music, as well as additional correspondence.

Discours préliminaire de l'Encyclopédie. Edited by P. Picavet. Paris, 1912. Standard critical edition.

For a fuller listing, see *A Critical Bibliography of French Literature.* Edited by D. C. Cabeen. Syracuse, NY: Syracuse University Press, 1947–1951. Vol. IV, pp. 136–138.

WORKS ON D'ALEMBERT

Bertrand, Joseph. *D'Alembert.* Paris: Librarie Hochette, 1889. Despite shortcomings and reliance on Condorcet's *Eloge de d'Alembert,* the most complete biography to date.

Grimsley, Ronald. *Jean d'Alembert.* Oxford: Clarendon Press, 1963. A good, comprehensive treatment of d'Alembert's philosophy and ideas. Less concerned with biography.

Kunz, Ludwig. "Die Erkenntnistheorie d'Alemberts," in *Archiv für Geschichte der Philosophie,* Vol. 20 (1907), 96–126. Considers relation between d'Alembert's metaphysics and English empiricists. Presents him as a link between empiricists and Comte.

Misch, Georg. *Zur Entstehung des französischen Positivismus.* Berlin, 1900. Influence of d'Alembert's empiricism and materialistic viewpoint on Comte's Positivism.

Muller, Maurice. *Essai sur la philosophie de Jean d'Alembert.* Paris, 1926. Most important and complete study of d'Alembert's general philosophy.

Pappas, John N. *Voltaire and d'Alembert.* Bloomington: Indiana University Press, 1962. Considers d'Alembert's position and method in spreading the ideals of the Enlightenment and his influence on Voltaire.

John N. Pappas (1967)

ALEXANDER, SAMUEL

(1859–1938)

The British realist metaphysician Samuel Alexander was born in Sydney, New South Wales, and was educated at Wesley College, Melbourne. He came to England in 1877 on a scholarship to Balliol College, Oxford, where he read mathematics, classics, and philosophy (*literae humaniores*). In 1882 he was elected to a fellowship at Lincoln College, Oxford, becoming the first Jew to be a fellow of an Oxford or Cambridge college. His earliest work, the Green Prize essay in moral philosophy, subsequently published as *Moral Order and Progress* (1889), shows the influence of the idealist ethics dominant in Oxford at the time. But he soon began moving toward an approach to philosophy that could be more closely related to the development of the empirical sciences, particularly biology and psychology. He gave up his fellowship and spent a year in Hugo Münsterberg's psychological laboratory at Freiburg, Germany, continuing in private study until his election to the chair of philosophy at Owens College (later the Victoria University of Manchester) in 1893. He held the chair until his retirement in 1924 and lived in Manchester until his death in 1938, a beloved, influential, and, indeed, legendary figure in both city and university.

EMPIRICAL METAPHYSICS

Alexander wrote occasional papers and a small book on John Locke, but it was not until 1920 that he published his major work, *Space, Time and Deity* (delivered as the Gifford Lectures in Glasgow in 1915). This was a comprehensive and constructive system, which he claimed was metaphysics following an "empirical method." By this he meant that he understood metaphysics to be a very inclusive kind of science, differing from the special sciences "not in its spirit, but only in its boundaries, dealing with certain comprehensive features of experience which lie outside the purview of the special sciences." Alexander called these features "categorial" and "a priori" but said that this must not be taken to mean that they are imposed or constructed by thought; they are discerned by reflective inspection as pervasive features of the world. As such he called them "nonempirical," reserving the term *empirical* for the variable features of the world. But the *study* of both, as a study of what is found in experience, he called "empirical." This could be considered an empirical way of thinking only in a much broader and more speculative sense than subsequent forms of empiricism, with their stricter notions of what constitutes tests in observation and experiment. Nevertheless, Alexander insisted that his system not only was a speculative world view but also took account of certain ways of thinking he believed were suggested by work in contemporary experimental science. Here his starting point was probably his interest in physiological psychology (he had introduced this study into the University of Manchester at a time when British universities were still slow to recognize it).

MIND. In contrast to idealistic or dualistic views, Alexander regarded mind as, in one sense, identical with an organized structure of physiological and neural processes, there being no animistic or purely "mental" factor over and above these. But in another sense, mind could be looked on as a new "emergent"—when neural processes are organized in a certain way, they manifest a new quality, consciousness, or awareness.

EMERGENTS. By *emergents* (a term generally ascribed to C. Lloyd Morgan, though its first use can be found in G. H. Lewes) Alexander designated certain organized patterns which, he held, produce new qualitative syntheses that could not have been predicted from knowledge of the constituent elements of the pattern before they were so organized. Emergents are thus what others have called gestalt properties of organized systems; Alexander thought of them particularly as characteristics of those syntheses where some strikingly new quality can be discerned. He generalized the idea that new qualities emerge from patterns of subvening elements of certain degrees of complexity, so as to look on the world as a hierarchy of qualities, a hierarchy in which those higher in the scale depend on the lower but manifest something genuinely new.

SPACE-TIME. At the basis of nature Alexander set *space-time* as a continuum of interrelated complexes of motion. These can be analyzed into relations between "point-instants," a point-instant being the limiting case of a motion. Sometimes he spoke of point-instants as if they were real elements, the smallest instances of spatiotemporal motions, sometimes as if they were ideal concepts, the bare notion of time at a point or space at an instant, while any actual motion has a spatiotemporal spread.

Space-time was also distinguished into "perspectives." A perspective defines how space-time can be ordered with reference to particular point-instants. It is a line of advance, or phase of a spatiotemporal process, seen in relation to some point-instant as its center of reference. Alexander used the illustration of a tree sawn across. For the carpenter the concentric rings are simultaneous, but this is to look on it as an artificial section. For the botanist they are of different dates, carrying with them the history of the tree. Thus, a perspective is a historical phase of the process of nature, ordered with reference to some event, *e,* as center and integrating other events related to the event from which the perspective is developed. These may be integrated as observably contemporaneous or as earlier and later stages in motions of which *e* is a stage.

The definition of a perspective thus depends on the notion of motions and their interrelation, and even on their causal relations. It is difficult to see how these notions can be derived purely from that of structures within space-time. Indeed, the notion of space-time itself as the fundamental stuff or matrix out of which things arise is certainly not one that it is natural to see as an "empirical" description of the most general features of the world as it discloses itself to an observing mind.

CATEGORIES. It might be more plausible if Alexander could be taken to have meant that the basic universal feature of all experience is its spatiotemporal character. He did indeed claim to follow Kant in holding that the world is apprehended first and foremost as a spatiotemporal manifold, under categories. Apart from the union of space and time in a four-dimensional continuum, his categories follow closely the Kantian ones of substance,

cause, number, and relation. But Alexander insisted that these categories are discovered or discerned in the world and are not a conceptual framework imposed by the mind. Indeed, according to his realist theory of knowledge, thought does not construct or impose conceptual schemes. Knowledge is "contemplation" of an object where there is a relation of "compresence" between a mind and an object (except in the special case of a mind's knowledge of itself, for a mind cannot be compresent as an object to itself but is aware of itself as knowing and perceiving; Alexander calls this kind of knowing "enjoyment"). But it is surely difficult to understand why any mind compresent with the world of nature would see in it just these particular all-pervasive categorial features.

EMPIRICAL FEATURES OF REGIONS OF SPACE-TIME. Beside the categorial features, which Alexander called "nonempirical," meaning by this that they are invariable and all-pervasive, we discover "empirical" features, defined as variable qualities characterizing particular regions of space-time. "Universals" are discerned *in rebus,* as plans of configurations of motions in space-time showing persistent identities; Alexander called them "habits" of space-time. Within space-time arises the hierarchy of emergent qualities. The patterns of motions that differentiate it are in the first place bearers of the properties of extension and inertia that characterize "matter." These organized patterns of matter are bearers of the qualities found in physical structures and chemical syntheses. Some of these syntheses, in turn, are bearers of the quality of "life," and some living structures are bearers of mind or consciousness, which is the highest empirical quality known to us. There is no reason, however, to assume that this is the highest possible emergent quality. Alexander held that the structures that are bearers of "mind" may in their turn become productive of a new emergent quality, which he called "deity."

DEITY. The term *deity* does not here stand for a God who precedes the universe as its cause or creator. Alexander did not try to find in such terms an "explanation" of why the universe should exist. Existence, he held, should be accepted with "natural piety" (borrowing a phrase of William Wordsworth's), and its general character should then be described. This general character is first and foremost spatiotemporal. In addition, Alexander held that it exhibits a *nisus,* or creative tendency, toward the production of new qualitative syntheses. So in one sense God can be thought of as *Deus sive Natura,* the universe of space-time "pregnant" with emergent qualities. In another sense deity is "the next highest emergent quality which the universe is engaged in bringing to birth." This quality, Alexander suggested, may emerge in beings—we do not yet know what they would be like—who would be bearers of deity as we are of mind, and these in their turn might prepare the basis for a yet further emergent quality. Alexander held that the existence of religious sentiments and aspirations witnesses to an experience of the *nisus* toward the higher quality of deity in some of those who are already bearers of mind. Such religious feelings, he thought, are incipient aspirations toward a new level of development. It is toward this further stage of development, not toward an already existent object, that the religious sentiment is directed. Alexander claimed that he started from the empirical fact of this sentiment, rather than from a theory of its object, and asked what it suggests; the religious sentiment can be interpreted as the feelings of beings caught up in the *nisus* of a universe "pregnant" with the quality of deity.

TIME AS MIND. Is there any reason in the nature of space-time itself why there should be this *nisus*? Alexander sometimes spoke as though the mere fact of conjoining time with space in itself produces the possibility not only of a dynamic but even of a creative process. He summed this up in the saying "Time is the Mind of Space"—surely one of the most astonishing remarks ever made by a metaphysician. But it was not intended merely to shock. It should be read in connection with Alexander's interest in physiological psychology and the view of the body-mind relation that he derived from this and that he here extended in a daring analogy. Alexander reported that he reached his notion of perspectives in space-time by considering the unity of the self. There is no such thing as awareness of the self at an instant. The least moment of conscious experience is a "specious present" with a durational aspect and, as embodied, a spatialized aspect. Our consciousness of what we are thinking at any moment is linked with the memory of what we were thinking, for example, a fraction of a second ago, and it is directed in anticipation toward what we are going to think a fraction of a second from now. What we are, at any given stage, is partly constituted by memories of the past and anticipation of the future.

Hence, the unity of the self depends on events of different dates being brought into a perspective with reference to the self of "present" experience. Similarly, a physical perspective consists of all events that can be shown to be earlier or later stages in lines of development in which a given event, taken as center of reference, is a phase. A perspective thus describes a historical line of advance. The temporal aspect of this is said to be the ana-

logue of its "mind" and the spatial aspect the analogue of its "body." This is because mental experience is partly constituted by memory of the past and anticipation of the future and, more specifically, because the "mind" aspect of anything is looked on as the *new* quality it may exhibit at its latest point of development, whereas the organized structure that is the bearer of this property and could be described beforehand as accomplished fact is looked on as its body. Time is not mind in the sense of consciousness or thought, which is the distinctive quality characteristic of the level we call mind proper. It is "mind" in an analogical sense, as whatever is the new property characteristic of a new qualitative synthesis. Thus, for example, to Alexander the defining qualities of matter are the primary qualities, such as extension and inertia. Secondary qualities, such as color, are emergents from organized complexes of matter and may, as such, be called their "mind." This is not to give them some rudimentary degree of consciousness; it is to say that on each level there is an element that can be called the analogue of mind, as introducing something new. But what is new appears sometimes to be not describable as an *element,* but rather as a new way of *functioning* released in some particular kind of ordered structure. When this happens, the new way of functioning dominates the lower levels that support it but does not transform them into something different. Physicochemical processes continue to be physicochemical processes, and neural processes to be a form of physicochemical processes. But where there is conscious thinking, although no separate animistic or mental factor may be present, the whole ordered structure becomes a vehicle for this new activity, and we say we are confronted by an "embodied mind."

TIME AS AN ATTRIBUTE OF REALITY. Alexander's view of a hierarchy of syntheses with new emergent qualities may be significant, but can time, as the pure notion of irreversible succession, be sufficient to account for their possibility? To say that there is a general tendency for complexes of one order to combine and form complexes of what will become a new order must surely presuppose some fundamental property or properties in the world besides those of space and time; Alexander, in fact, admits this when he speaks of a *nisus,* or creative tendency, in space-time. But is this a necessary property of an infinite four-dimensional continuum, unless one can assume that the mere fact of succession entails creative advance? Alexander may have been near enough to nineteenth-century ideas of inevitable evolutionary progress to be able implicitly to assume this. In agreement with these ideas, he insisted that philosophers must "take Time seriously"; that is to say, they must incorporate a conception of time as an essential attribute of reality and not only as describing a way of experiencing or measuring a reality that is ultimately nontemporal. Alexander said that if Benedict (Baruch) de Spinoza could be rewritten with time as well as extension as an attribute of substance, this would represent the type of past philosophy most congenial to him; indeed, if someone were to write on his funerary urn "*Erravit cum Spinoza,*" he would be content.

REALITY AS PROCESS. Alexander's view of space-time as the final reality seems, however, open to two interpretations. On the one (perhaps the more Spinozistic) interpretation, space and time are the two necessary attributes of an infinite substance, distinguishable, it is true, into perspectives defined by reference to point-instants, but where "motions" (analogous to Spinoza's "modes") are simply the redistribution of spatiotemporal coefficients within the whole already existent space-time. In this view space-time is looked on as that out of which things come, and we can ask whether, as with the materialist's conception of matter, this is not to treat an abstraction as though it were a reality. In another sense Alexander was giving a view of reality as essentially a process, and as historical. There is an irreversible direction in it, defined by "time's arrow" (to use Arthur Stanley Eddington's expression). In this, nature is focused in lines of development whose "history" describes the successive levels of ordered structures they exhibit. At each stage in time, where there is a new emergent quality, this quality is the spearhead of a genuine creative advance. Yet if this new emergent quality at each stage is said to be analogous to mind, is it satisfactory to equate this with saying that it is analogous to time? It might be more plausible to say that it was Alexander's notion of the *nisus* in space-time that corresponds to the "mind" factor in those complexes whose extended patterns can be regarded as the analogue of the body. Or one might say that the "body" of anything is the external view of nature as unified in that particular perspective, while its "mind" is the "idea" of the distinctive internal quality of that particular perspective; this indeed suggests comparison with Spinoza's view of the body-mind relation.

VALUES

Alexander wrote no large work besides *Space, Time and Deity.* The volume *Beauty and the Other Forms of Value* (1933) is a collection of occasional papers and lectures on themes relating to aesthetics and ethics. The general notion underlying these is that of values as related to the satisfaction of impulses. Values are "tertiary qualities"

(supervening on the primary and secondary qualities), characterizing complexes where one component is a mind capable of interest or appreciation. The higher values—beauty, truth, and goodness—are qualities that arise in the satisfaction of certain impulses where these have become contemplative and disengaged from their immediate practical ends. Thus aesthetic creation and enjoyment grow out of the impulse to construct things, which Alexander traced down to the animals ("impulse," he thought, was a less question-begging term than "instinct"). The impulse to construct something out of physical materials, including sounds, becomes a contemplative delight in the form so imposed on the material. Truth is a value analogous to beauty, as that which satisfies the impulse of curiosity when this too becomes contemplative rather than practical. Moral value is a quality created out of natural impulses by the introduction of another natural impulse that can bring form and harmony into the impulses that are its materials. This impulse Alexander called "gregariousness." His interpretation of this was close to Adam Smith's view of "sympathy" as fellow feeling with the feelings of others. Gregariousness, like Smith's sympathy, becomes disinterested and so is able to act as a harmonizing agent both among a person's other impulses and in producing "sociality." The impulse of "sociality" was also invoked in support of Alexander's view that we are directly aware of other minds in such experiences as friendly conversation or quarrels, which are completed as experiences through reciprocated responses. These are not, in Alexander's view, adequately described as merely responses to behavior; they are responses to behavior as expressing the mind of the other person.

A collection of occasional papers and addresses, *Philosophical and Literary Pieces* (1939), was published posthumously by John Laird, prefaced by a memoir that gives a sympathetic account of Alexander the man, including a number of the stories, true or apocryphal, that were told about him. Some of the pieces on nontechnical themes—on Dr. Johnson, for instance, or Jane Austen, or Blaise Pascal—show Alexander in his happiest vein.

Alexander was awarded the Order of Merit in 1930. His appearance was impressive; a bust by Jacob Epstein in the entrance hall of the Arts Building of the University of Manchester gives a good impression of his massive head and beard but misses his kindliness. The library of the University of Manchester contains a large collection of letters written to him by his contemporaries, including the philosophers F. H. Bradley, G. F. Stout, and T. Percy Nunn, the physiologists C. Lloyd Morgan and Sir Charles Scott Sherrington, and the Jewish leaders Chaim Weizmann and Claude Montefiore.

See also Bradley, Francis Herbert; Emergence; Empiricism; Kant, Immanuel; Locke, John; Morgan, C. Lloyd; Pascal, Blaise; Smith, Adam; Spinoza, Benedict (Baruch) de; Stout, George Frederick.

Bibliography

WORKS BY ALEXANDER

Books

Moral Order and Progress. London: Trübner, 1889.

Locke. London: Constable, 1908.

Space, Time and Deity. 2 vols. London: Macmillan, 1920.

Beauty and Other Forms of Value. London: Macmillan, 1933.

Philosophical and Literary Pieces. Edited by J. Laird. London: Macmillan, 1939.

The papers of particular philosophical interest appearing in this volume are "Art and Instinct," Herbert Spencer Lecture (1927); "Artistic Creation and Cosmic Creation," Annual Philosophical Lecture of the British Academy (1927); and "Spinoza and Time," Arthur Davis Memorial Lecture (1921). The book also contains a list of Alexander's published works, including reviews and contributions to journals.

Periodicals

"The Method of Metaphysics; and the Categories." *Mind,* n.s., 21 (1912): 1–20.

"On Relations" and, in particular, "The Cognitive Relation." *Mind,* n.s., 21 (1912): 305–328.

"Collective Willing and Truth." *Mind,* n.s., 22 (1913): 14–47 and 161–189.

"The Basis of Realism." In *Proceedings of the British Academy,* 1914. Republished in *Realism and the Background of Phenomenology,* edited by Roderick M. Chisholm. Glencoe, IL: Free Press, 1960. This lecture relates Alexander's views on the theory of knowledge to those of other philosophers of the time and presents his theory that mind and object are both empirically within nature, their relation being a particular form of a more general relation of "compresence."

"Sense-Perception: A Reply to Mr. Stout." *Mind,* n.s., 32 (1923): 1–11. Written in answer to criticisms by G. F. Stout, in *Mind,* n.s., 31 (1922): 385–412.

WORKS ON ALEXANDER

Broad, C. D. *The Mind and Its Place in Nature,* pp. 646–650. New York: Harcourt, Brace, 1925. Discusses Alexander's "emergence" theory of mind. See also two reviews by Broad of *Space, Time and Deity,* in *Mind,* n.s., 30 (1921): 25–39 and 129–150.

Devaux, P. *Le système d'Alexander.* Paris: J. Vrin, 1929.

McCarthy, J. W. *The Naturalism of Samuel Alexander.* New York: King's Crown, 1948.

Dorothy M. Emmet (1967)

ALEXANDER OF APHRODISIAS

Alexander of Aphrodisias, who was teaching at Athens in 200 CE, was recognized for centuries as the most authoritative exponent of Aristotle. His influence has probably been most far-reaching in the development of the theory of universals because he emphasized certain elements in Aristotle's not always unambiguous account. These were the unqualified priority of the particular substance and the existence of universals only as concepts, or "acts of intellect." The form was what made "this" matter (that is, an identifiable piece) what it was, but it was contingent whether the form was universal in the sense of generic. (Alexander does not notice that a class with only one member, like his case of the sun, is still a class.) What the form is as a subject remains unclear.

More famous is his doctrine about soul and intellect. A human being's intellectual faculty can exist in three conditions, described as three intellects: (1) the "material" intellect (*intellectus possibilis*), which is nothing actual but the bare potentiality (so Aristotelian matter) of the body to develop reason—the condition of babies; (2) the intellect (*intellectus in habitu*) that is the possession of, in fact, is identical with, concepts, or universals gained from sense experience—the condition of adults; (3) the "active" intellect (*intellectus agens*), which is exercising the thoughts that form the *intellectus in habitu* and is thus equivalent to the intellect as aware of itself.

What is distinctively Alexandrist is the identification he made, or seemingly made, of the "active" intellect both with the intellect that Aristotle said entered the body "from outside" and with the intellect eternally thinking of itself that Aristotle said was God. Intellect was, of course, the highest part or function of the soul, but since only the "active" intellect, as a "separate form," could exist without matter, it followed that there was no individual immortality for human beings. The exact relation of the "active" intellect to the individual soul or intellect is obscure in Alexander. He does not describe an active intellect acting directly like an efficient or even formal cause on a passive intellect but suggests rather the quasi-logical relationship which was fundamental to Neoplatonism and which made the less perfect instance of a kind entail the existence of the perfect. Thus, it is not at all certain that he meant thinking itself to go the way of immortality. In the fifteenth century Italian philosophers known as Alexandristi defended this interpretation of Aristotle's psychology against both Averroes's version and the theologically orthodox version of Themistius and Thomas Aquinas.

In other subjects we see Alexander less original but often attacking Stoic doctrine, notably in his tracts *On Fate* and *On Mixture*. But the exact understanding of him is colored always by the difficulty of knowing how far we can trust the writings attributed to him. The commentary on Books E (VI) to N (XIV) of Aristotle's *Metaphysics* and parts of Book II of his own *De Anima* are probably not his. The latter includes the section *On Intellect* which greatly influenced later Greek, Arab, and medieval philosophers. But both may well depend on and be closer to his thought than is allowed by a modern tradition that underestimates Neoplatonizing features of Aristotle as well as of Alexander.

See also Aristotle; Averroes; Neoplatonism; Themistius; Thomas Aquinas, St.; Universals, A Historical Survey.

Bibliography

Alexander's works, including dubious ones, are in *Commentaria in Aristotelem Graeca*, Vols. I–III, and *Supplementum Aristotelicum*, Vol. II (Berlin, 1883–1901). P. Moraux, *Alexandre a'Aphrodise, exégète de la noétique d'Aristote* (Paris, 1942), includes a French translation of *On Intellect*. See also F. E. Cranz, in *Catalogus Translationum et Commentariorum: Mediaeval and Renaissance Latin Translations and Commentaries*, Vol. I (Washington, DC: Catholic University of America Press, 1960), pp. 77–135; Ernst Cassirer, *Das Erkenntnisproblem*, 3rd ed., Part I, "Die Reform der aristotelischen Psychologie" (Berlin: Cassirer, 1922); and J. H. Randall Jr., *The School of Padua and the Emergence of Modern Science* (Padua, 1961).

A. C. Lloyd (1967)

ALEXANDER OF APHRODISIAS [ADDENDUM]

Alexander of Aphrodisias's influence on Islamic philosophy was far reaching. In fact, it could appear to be somewhat out of proportion with his real importance as a thinker. The reason for this is partly fortuitous in that a large number of his works were preserved long enough for them to reach Baghdad in the ninth century CE and be translated into Arabic. Among the most significant of these are the following:

(1) The fragments of the Commentary on Aristotle's Metaphysics, book lambda (*lam* in Arabic) preserved by Ibn Rushd in his own Great Commentary on the same work. The original text is lost in Greek.

(2) The short treatise *On the Principles of the Universe* describing the mechanics of the heavenly motions and the mode of their influence on the sublunary world. It could be defined as a free synthesis of the main themes of Aristotle's *Physics* and *Metaphysics*, with some borrowings from the *De Anima* and the *Nicomachean Ethics*.

(3) A treatise, *On Providence*, preserved in two fairly different translations. This last work was of particular importance to the Muslims in that it provides an Aristotelian answer to a question that is crucial in the context of a monotheistic religion, but was never treated as such by Aristotle himself.

The main features of the philosophical system set forth in these works can be summarized as follows. The heavenly motions are caused by the souls of the spheres (which carry the stars) in their desire to imitate the First Mover of the universe. The counterpart of this upward motion is the influence that the contrasting motions of the stars exert on the world of nature. This influence is as a matter of fact identified by Alexander with nature and providence. But this providence, although emanating from the heavens, is not willed by them, because Alexander postulates that the superior cannot care for the inferior without debasing itself.

Another Alexandrian tenet that exerted a profound influence on the Arab philosophers is his identification of the Active Intellect of Aristotle's *De Anima* with the Unmoved Mover of the *Metaphysics*. The intellectual processes of the human mind were thus directly connected with the divine.

Bibliography

Genequand, Charles. *Alexander of Aphrodisias on the Cosmos*. Leiden, Netherlands: Brill, 2001.

Thillet, Pierre. *Alexandre d'Aphrodise. Traité de la Providence*. Paris: Verdier, 2003.

Charles Genequand (2005)

ALEXANDER OF HALES

(c. 1185–1245)

Alexander of Hales, "Doctor Irrefragabilis," friar minor, was an English Scholastic at the University of Paris. He was born in Hales Owen, Shropshire, and died in Paris.

Alexander was a student at Paris about 1200 and received his M.A. before 1210. He joined the faculty of theology, becoming a master regent about 1220. After 1222 Alexander made an innovation in the university by using the *Book of Sentences* of Peter Lombard as the basic text for theological courses. His newly published *Glossa* (identified only in 1945) was the result of this work. At the height of his career, about 1236, he became a Franciscan, "edifying the world and giving new status to the Order" (in the words of Roger Bacon). After he was put in charge of the school at the Paris friary, he continued his teaching, especially through his *Disputed Questions,* and had some part to play in the "great Summa weighing more than a horse, which the friars out of reverence ascribed to him and called 'the Summa of Friar Alexander'" (R. Bacon). At the same time, he participated in the affairs of the order, attending the chapter that deposed Brother Elias in 1239, and was a coauthor of an *Exposition of the Rule of St. Francis*; he was also active in the affairs of the church, both in the university and in the First Council of Lyon (1244–1245). His sudden death after his return from Lyon apparently resulted from an epidemic current in Paris. An epitaph in the convent church saluted him as *Gloria doctorum, decus et flos philosophorum* (Glory of learned men, the honor and pride of philosophers).

TEACHINGS

Alexander's own doctrines are found in his *Glossa* and *Disputed Questions* (which are divided in the British Museum manuscript *Royal 9. E. 14.* into two series: those written before and those written after he became a friar); the *Summa* ascribed to him does not necessarily represent his opinions. Both the *Gloss* and the *Questions* labor under the disadvantage of being students' reportations (although some copies seem to have had a kind of official approval); both, however, justify the encomium of Bernard of Bessa: *maximus in theologia et philosophia magister* (greatest master in theology and philosophy). Alexander is both theologian and philosopher, masterfully handling a wide range of questions. Undoubtedly a traditionalist whose prime sources are Augustine, John of Damascus, and Pseudo-Dionysius, and whose thought is close to the scholastic traditions of his predecessors, Alexander nonetheless surpasses his contemporaries in the breadth and profundity of his questions and in the new problems and tracts he introduced into theology. To this extent he was an innovator who helped open the way for the scholastic renaissance of the mid-thirteenth century. In particular, as head of the friars' studium at Paris, he initiated a certain approach that came to characterize such representatives of the Franciscan school as Odo Rigaldus, Bonaventure, and Matthew of Aquasparta.

The problems of the distinction between philosophy and theology, and the nature of theology as a science, much discussed after 1240, are not treated explicitly (though it is possible that Alexander wrote a question on the subject; see below). These problems are implicitly considered in scattered remarks on the kinds of human knowledge and the validity of arguments, in the general organization of material into specific questions and problems, and in the principles used in the solution of the problems. For example, our knowledge of God arises both from authority and from reason; that is, either from faith, which "depends on hearing" (Romans 10.17), or from knowledge drawn from the things God has made. Proofs of God's existence are suggested rather than developed at length: one is derived from the transcendental attributes of truth, goodness, and unity found in things; others are argued from the changing to the Unchanged, from dependent being to the Highest Being, from participated and partial good to the *summum bonum* (*Glossa* I, pp. 39–41). In the tradition of Augustine, Alexander finds analogies of the triune God in all creatures, thus setting the pattern for the Franciscan school, which, with St. Francis, delights to make of creation a "ladder" to the Creator. At the same time, Alexander shows the simplicity of the divine being to be in marked contrast to the composite character of all created being (*Glossa* I, p. 254; *Quaestiones,* pp. 14, 19). The doctrine here, that of *quo est* (the substance) and *quo est* (essence), is derived ultimately from Boethius, not from Avicenna, who seems to have been unknown to Alexander. In contrast to the *Summa Fratris Alexandri* and to Bonaventure, Alexander vehemently rejects any composition of matter and form either in angels or in the human soul (*Glossa* II, p. 28; other texts are in V. Doucet, *Prolegomena,* pp. 237, 268, n. 2). Apart from a lengthy question on immortality (*Quaestiones,* pp. 556–565), only passing remarks embody his notion of the soul. His attention is drawn more to the problem of free will (Ibid., pp. 566–608, plus an unedited question). Here, Alexander teaches that man by his nature is free and that freedom of choice resides both in the intellect and in the will. The primary purpose for which man has been given this freedom is to choose that which is morally good. Alexander considers the moral life of man in such *Disputed Questions* as "On Ignorance," "On Scandal," "Love of Neighbor," "Fraternal Correction," "On Impediments to Reason," "On Lying," and "Conscience" (the last two as yet unpublished). To the last question must be joined his study of synderesis (*Glossa* II, pp. 380–385), which seems to make Alexander, not Philip the Chancellor, the creator of such a tract in Scholasticism.

LITERARY PROBLEMS OF THE "SUMMA FRATRIS ALEXANDRI"

Since the *Summa* attributed to Alexander was unfinished at his death, William of Militona, who became master regent in 1248, seems to have undertaken its completion, for in 1255 Pope Alexander IV charged the provincial of Paris to supply Militona with capable assistants who without delay would bring the work to a finish. The text as it now stands consists of four parts. Book I deals with the nature of theology, the existence and nature of God, the divine names, and the Trinity. Book II is divided into two sections: II–1, creation in general, the angels, the six days of creation, the soul, the body, and the human composite, and II–2, a lengthy study of moral theology—the nature of evil, definition and classification of sins, and original and actual sins. Book III considers the Incarnation and mysteries of Christ's life, law (eternal, natural, positive, the commandments), grace, and faith (tome IV). Book IV treats of man's reparation through the sacraments, the mass, prayer, fasting, and almsgiving; quite evidently a section on "Last Things" was to be included as the climax of the work.

Except in a few manuscripts and in the protest of Roger Bacon, however, the compilatory nature of the *Summa* was forgotten. All four books came to be attributed to Alexander, despite the manifest contradictions and conflicting opinions in the various parts. Only since the end of the nineteenth century, with the renewal of interest in medieval Scholasticism, has the question of authorship attracted attention. A few writers, it is true, have gone to an extreme in claiming that the whole *Summa* was a compilation of the last half of the thirteenth century, in basic dependence on Thomas Aquinas, Albert the Great, and Bonaventure. But more mature and solid scholarship has established that, if by and large the *Summa* is a compilation, it existed as a whole by 1257. The first three books were in existence before the death of Alexander, with three notable exceptions: The last tract of Book I was added between 1250 and 1253, while in Book II–1 the two sections "On the Human Body" and "The Human Composite" were composed after Bonaventure, almost certainly in 1255–1257, as was the last book. On the other hand, modern research is forced to agree with Roger Bacon that Alexander was not the author, in the strict sense, of the pre-1245 *Summa.* At most, it appears that he planned and organized the work, while the details were left to others. Internal criticism of style, language, and doctrine would show essentially two authors at work, neither of whom, by reason of doctrinal positions, can be Alexander. Books I and III were almost certainly the work

of John of La Rochelle, although the presence of other collaborators may be detected. Both parts of Book II, on the other hand, were written or compiled by some unknown friar who possessed a keen philosophical mind and a greater spirit of independence.

DOCTRINES OF THE PRE-1245 "SUMMA." The work of the "Summists" was largely one of compilation, yet not without a certain new and fresh viewpoint. If they drew on earlier material, they did not hesitate to insert their own views or add fresh tracts written specifically for the *Summa.* Relatively new was the opening inquisition on the nature of theology, based on the tract in manuscript *Vatican Latin 782,* folio 184d–186d (which may be by Alexander himself); it bears witness to the growing influence of Aristotle's ideal of a science. This inquisition is followed by an original tract on natural theology, remarkable for its metaphysical doctrine of God and creatures. This doctrine holds that the very conditions of finite being demand the existence of a First Being, even as the positive perfections of finite things reflect and lead to the infinite. The unknown author of Book II does not hesitate to repeat some of this material in an interesting and well-balanced dissertation on Creator and creature; he examines in detail the meaning of the act of creation, the properties of created being that reflect the divine cause, and those properties peculiar to creatures: composition, changeableness, time and space, and the beauty and order of the universe. Several questions seem to have bearing on problems that arose in the early thirteenth century under the influence of the newly known Arabian philosophers.

The importance of the *Summa* lies chiefly, perhaps, in its presentation and defense of the so-called Augustinian traditions in theology and philosophy without neglecting whatever was solid in the new philosophical literature. It may rightly be called the *Summa Minorum,* embodying the fundamental doctrines of the Franciscan school of the early thirteenth century.

See also Albert the Great; Augustine, St.; Bacon, Roger; Bonaventure, St.; John of Damascus; John of La Rochelle; Matthew of Acquasparta; Peter Lombard; Pseudo-Dionysius; Thomas Aquinas, St.

Bibliography

For texts of Alexander, see *Glossa in Quattuor Libros Sententiarum Petri,* 4 vols. (Quaracchi, 1951–1957); *Quaestiones Disputatae "Antequam Esset Frater,"* 3 vols. (Quaracchi, 1960)—the series "after he became a friar" awaits publication at Quaracchi; *Expositio Quattuor Magistrorum Super Regulam Fratrum Minorum,* L. Oliger, ed. (Rome, 1950). The text of the *Summa* is *Doctoris Irrefragabilis Alexandri de Hales Summa Theologica,* 4 vols. (Quaracchi, 1924–1948), books I–III; Book IV is found in several early editions, such as Nuremberg, 1482, and Cologne, 1622.

With regard to studies of Alexander's works, writings previous to 1948 that are concerned with doctrinal problems rather than literary ones must be interpreted in the light of new discoveries. For literary and historical aspects, see V. Doucet, *Expositio Quattuor Magistrorum III Necnon in Libros I et II Summa Fratris Alexandri* (Quaracchi, 1948), partially translated in *Franciscan Studies,* Vol. 7 (1947): 26–41, 274–312; Ibid., Vol. 6 (1946): 403–417; and Doucet's introductions to the authentic *Glossa* and *Quaestiones.*

Doctrinal studies on the authentic Alexander are just beginning to appear: A. Fuerst, *A Historical Study of the Doctrine of the Omnipresence of God* (Washington, DC: Catholic University of America Press, 1951); E. Lio, *Determinatio "Superflui" in Doctrina Alexandri Halensis* (Rome: Apud Pontificium Athenaeum Antonianum, 1953); A. Hufnagel, "Die Wesensbestimmung der Person bei Alexander von Hales," in *Freiburger ZPT,* Vol. 4 (1957): 148–174; Bettoni, *Il problema della cognoscibilità di Dio nella scuola francescana* (Padua, 1950); see also *Franciscan Studies,* Vol. 10 (1950): 164–185, 286–312; Vol. 19 (1959): 334–383; Vol. 20 (1960): 96–148; Vol. 22 (1962): 32–149.

For earlier and somewhat outdated studies on Alexander as author of the *Summa,* see I. Herscher, "A Bibliography of Alexander of Hales," in *Franciscan Studies,* Vol. 5 (1945), 434–454; and P. Boehner, "The System of Metaphysics of Alexander of Hales," in Ibid., 366–414.

Ignatius Brady, O.F.M. (1967)

AL-FĀRĀBĪ

(c. 873–950)

Al-Fārābī, more fully Abū-Nasr Muhammad al-Fārābī, known in Latin as Alfarabius or Avennasar, was one of the greatest Muslim philosophers. He was widely known as "the second master," Aristotle being the first, and Ab-Ar-Rahman ibn-Khaldūn rates him above Avicenna and Averroes. He was of Turkish origin, and his name indicates that he came from the district of Fārāb, on the middle Jaxartes River (now Syr Darya).

One of al-Fārābī's teachers was the Nestorian Christian Yuhannā ibn-Haylān, who was noted as a logician; it is uncertain whether al-Fārābī studied with him in Merv (Persia) or Harran (Syria) or Baghdad. His principal teacher was Abū-Bishr Mattā ibn-Yūnus, the most prominent member of the school of Christian Aristotelians in Baghdad. Here al-Fārābī studied not merely the various branches of philosophy, but also physics, mathematics, astronomy, and music, even becoming a skilled musical performer. He spent the last few years of his life at the court of the ruler Sayf-ad-Dawla at Aleppo. He did not

seem to have had any regular occupation by which to earn a livelihood and lived frugally, even ascetically, often in solitude.

Al-Fārābī's philosophy is based on the teachings of Plato and Aristotle as they were interpreted in the school of Baghdad in the tenth century. Like all writers in Arabic he assumed there were no essential differences between the two, but he preferred the metaphysics of Aristotle, as interpreted by Neoplatonists. Plato, however, he regarded as superior in practical matters, and he wrote commentaries on the *Republic* and the *Laws.* What is often regarded as his major work is reminiscent of these books; it has the clumsy title "On the Principles of the Views of the Inhabitants of the Excellent State," often shortened in practice to "Der Musterstaat," or "The Ideal City" (*al-madīna al-fadila*). The first third of this work sets out al-Fārābī's metaphysical system, the second third his psychology (largely Aristotelian), and the concluding third his views on the ideal state and various imperfect states.

To those familiar with the intellectual environment in which al-Fārābī lived, it is immediately apparent that he wrote in such a way as to commend his views to as many different groups of people as possible. It has been alleged that he supported the Shiʿite sect of Islam, and certainly his last patron Sayf-ad-Dawla was a Shiʿite; features of his "ideal city," such as the dependence of all on the head, resemble Shiʿite conceptions. Yet it is also clear that he wrote in such a way as not to offend the Sunnite majority; for example, by avoiding such a technical Shiʿite term as *imam.* Indeed, his view of the relation of philosophy and religion led him to attach positive value to the religions, although he regarded them as inferior to philosophy. Philosophy was the supreme exercise of human reason and therefore the primary requirement of an ideal city. By it, humanity came to know the one ultimate truth about the universe. To this ultimate philosophical truth the symbolic representations of it found in the several religions stand in varying degrees of proximity and remoteness. Al-Fārābī paid particular attention, of course, to the forms of the main Islamic states of his time and developed his conception of the ideal city in such a way that the actual states he knew were within measurable distance of the ideal.

His metaphysics, similarly, resembles that implicit in the Qurʾan (Koran) and Islamic theology. God is the One or the First from whom all existence proceeds; and in this sense he accepts the Islamic doctrine that God is the creator of the world, although he also holds the heretical view that the world is eternal. In the relation of existent things to God there is a hierarchical order. Similarly in the ideal city there is a head (*raʾīs*) who is the source of all authority and who assigns men to their appropriate grades. This head is also described as commanding but not obeying; all the intermediate grades obey those above and command those below, and the lowest grade only obeys.

Interest has been shown, especially in recent times, in al-Fārābī's theory of prophecy; that is, in particular, how it was possible for Muhammad to receive the Qurʾan from God. Philosophic knowledge, the highest of all, he regarded as coming to the passive intellect of the philosopher from the Active Intellect, an existent below God in rank. Prophetic revelations also come from the Active Intellect but are received by the imagination of the prophet. In this al-Fārābī was able to accept the Qurʾan as coming from God and yet to place philosophy above it.

See also Aristotle; Averroes; Avicenna; Islamic Philosophy; Kant, Immanuel; Plato.

Bibliography

WORKS BY AL-FĀRĀBĪ

The "Ideal City" has been translated as *Der Musterstaat von al-Fārābī* by Friedrich Dieterici (Leiden: Brill, 1900) and as *Idées des habitants de la cité vertueuse* by R. P. Jaussen, Youssef Karam, and J. Chlala (Cairo, 1949). A critical edition of the Arabic text and an English translation by Richard Walzer is in preparation.

See also *Al-Fārābī's Short Commentary on Aristotle's Prior Analytics,* translated by Nicholas Rescher (Pittsburgh: University of Pittsburgh Press, 1963); *Al-Fārābī's philosophische Abhandlungen,* short essays translated by Friedrich Dieterici (Leiden, 1892); and *The Fusul al-Madani: Aphorisms of the Statesman of al-Fārābī,* with an English translation, introduction, and notes by D. M. Dunlop, ed. (Cambridge: Cambridge University Press, 1961).

WORKS ON AL-FĀRĀBĪ

De Boer, Tjitze J. *The History of Philosophy in Islam.* London: Luzac & Co., 1903. See especially pp. 106–128.

Rescher, Nicholas. *Al-Fārābī: An Annotated Bibliography.* Pittsburgh: University of Pittsburgh Press, 1962.

Rescher, Nicholas. *Studies in the History of Arabic Logic.* Pittsburgh: University of Pittsburgh Press, 1964.

Quadri, G. *La philosophie arabe dans l'Europe médiévale.* Paris: Payot, 1947. See especially pp. 71–94.

Walzer, Richard. *Greek into Arabic.* Oxford: B. Cassirer, 1962. See especially pp. 18–23, 206–219.

W. Montgomery Watt (1967)

AL-FĀRĀBĪ [ADDENDUM]

Al-Fārābī was a key figure in establishing much of the problematic of Islamic philosophy in the peripatetic tradition. He built on the earlier attempt by Abū-Yūsuf Yaʿqūb ibn Isḥāq al-Kindī to establish a technical language of philosophy in Arabic and presented a vocabulary and a curriculum that came to dominate for many centuries after his death. Al-Fārābī's epistemology and political philosophy were particularly influential. Firmly neoplatonic in tone, he differentiated between diverse kinds of intellect to describe human thought and gave an interesting and influential account of how knowledge can be analyzed in terms of a range of degrees of abstraction. The active intellect became a controversial topic in Islamic philosophy; it represented the highest one could go in one's thoughts and was responsible for emanating form to the world in which one lived. The nature of this concept came to dominate much of Islamic philosophy for a long time after al-Fārābī's death. There was a great deal of debate on the precise role and nature of the active intellect and whether the hidden agenda of its use by the philosophers was to limit human knowledge to a relatively low level of impersonal thought.

Similarly, the distinctions he made in his political thought won attention as a result of their conceptual clarity. Following Plato he distinguished between different kinds of state, and he used the concept of happiness as the ultimate aim of government. Different kinds of government can be distinguished from each other by their varying links with happiness, with corrupt states being very poor at reaching happiness while the virtuous states achieve that end to a high degree. Not surprisingly, the idea that philosophers make the best rulers was rather attractive to philosophers, and in al-Fārābī's case the skills of the philosopher need to be blended with those of a religious leader if the state is to be well organized and led. As with his work on epistemology and metaphysics, his writings on political philosophy produced a lively debate in Islamic philosophy on the role of philosophy and philosophers in the state and on the nature of the state itself.

Bibliography

Al-Fārābī. "Al-Madīna al-Fāḍila" (The virtuous city). In *Al-Farabi on the Perfect State: Abū Nasr al-Fārābī's Mabādī' Arā' Ahl al-Madīna al-Fāḍila*. Translated by Richard Walzer. Oxford, U.K.: Clarendon Press, 1985.

Alon, Ilai. "Farabi's Funny Flora: Al-Nawabit as Opposition." *Arabica* 37 (1990): 56–90.

Black, Deborah. "Al-Fārābī." In *History of Islamic Philosophy*. 2 vols., edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996.

Fakhry, Majid. *Al-Farabi, Father of Islamic Neoplatonism: His Life, Works, and Influence*. Oxford, U.K.: Oneworld, 2002.

Galston, Miriam. *Politics and Excellence: The Political Philosophy of Alfarabi*. Princeton, NJ: Princeton University Press, 1990.

Netton, Ian Richard. *Al-Fārābī and His School*. London: Routledge, 1992.

Netton, Ian Richard. *Allāh Transcendent: Studies in the Structure and Semiotics of Islamic Philosophy, Theology, and Cosmology*. London: Routledge, 1989.

Oliver Leaman (2005)

AL-GHAZĀLĪ, AḤMAD
(c. 1062–1126)

Aḥmad al-Ghazālī's reputation as an Islamic thinker has unfortunately been overshadowed by that of his more celebrated elder brother, Muḥammad al-Ghazālī, author of the famous *Revivification of the Sciences of Religion*. The former was in fact the foremost metaphysician of love in the Sufi tradition and the chief founder of the philosophy of love in mystical Islam, and his impact on the later Persian Sufi tradition was more profound than his brother the theologian.

He spent most of his life in his *khānaqāh* (Sufi cloister) in Qazvīn, where he was famed for his eloquence as a preacher, and died there in 1126. Al-Ghazālī was the teacher of Abū' l-Najīb al-Suhrawardī, who was in turn the master of his nephew Shihāb al-Dīn Yaḥyā Suhrawardī, founder of the Suhrawardī order, famed as the "mother of Sufi orders." He was also the master of the enigmatic mystical theologian 'Ayn al-Quḍāt al-Hamadhānī, who was executed in 1132 by fanatical Muslim clerics for his uncompromising Sufi beliefs. He features as a central figure in the initiatic chains of most of the great Islamic Sufi orders.

His fame derives mainly from his authorship of the first treatise on mystical love in Persian, the *Sawāniḥ al-ʿushshāq* (The lovers' experiences), a short work on the spiritual psychology of divine love couched in the terminology of human erotic relationships. The main subject of his philosophy is passionate love (*ʿishq*), which is not formally speaking "philosophy"—*Falsafa*—but rather comprises a sort of erotic theosophy apprehended by intuitional means (*dhawq*), based on contemplative experience rather than on rational meditations and deliberations. Expressing little of the same animosity to peripatetic philosophy manifested by his famous brother,

almost all his teachings are set in the context of commentary on Qurʾānic verses and prophetic traditions. Al-Ghazālī deliberately abstained from using any overt philosophical vocabulary in the text, employing instead terminology from a number of other fields, ethics, erotic poetry, and psychology, and so on. He follows Manṣūr al-Ḥallāj in identifying love with the divine essence as well as with the divine spirit. He maintained that knowledge (*ʿilm*) alone is unable to grasp love (*ʿishq*), comparing knowledge to the shore of the sea and love to a pearl in an oyster buried in its lowest depths. Forever shore-bound in immanence, neither dry reason (*ʿaql*) nor barren knowledge (*ʿilm*) can ever access or apprehend the transcendent truths of love's apophatic teachings. The summit of knowledge lies in a kind of drunken inapprehension that is nonetheless a kind of apprehension without any of the limitations of subjective consciousness. Al-Ghazālī paradoxically describes this understanding of love that is "beyond knowledge" as being a kind of surmise or conjecture. This conjectural wisdom is higher than certainty for it is only that surmise or conjecture that can swim love's ocean to dive under in pursuit of its pearl. Due to *Sawāniḥ* and the many works of imitations it spawned, al-Ghazālī has come to be generally regarded as the foremost metaphysician of love in the Sufi tradition and the founder of the literary topos and mystical persuasion known as the "religion of love" (*madhhab-i ʿishq*) in Islam.

See also al-Gḥazālī, Muḥammad; Sufism.

Bibliography

al-Gḥazālī, Ahmad. *Sawāniḥ*, edited by Helmut Ritter. Tehran, Iran: Tehran University Press, 1989.

Pūrjavādī, Nasr Allāh, trans. *Sawānih: Inspirations from the World of Pure Spirits: The Oldest Persian Sufi Treatise on Love*. London: KPI, 1986.

Leonard Lewisohn (2005)

AL-GHAZĀLĪ, MUḤAMMAD

(450 or 451 AH [1058 or 1059 CE]–505 AH [1111 CE])

Muḥammad al-Ghazālī (in Persian, "Ghazālī"), the Islamic theologian known to medieval Scholastics as Algazel, was born in Ghazāleh, a village on the outskirts of Tūs, in Khorāsān, northeastern Iran. His name is the same as that of his birthplace, which should be transcribed as Ghazālī, not as Ghazzālī. He died at Tūs. He was undoubtedly one of the strongest spiritual personalities of Islam, one of those who strove most effectively for the establishment of an "orthodox" Sufism that would transcend the legalistic and superficial religion of the doctors of the Law. Al-Ghazālī was well known to the medieval Scholastics through a Latin translation of an unfortunately truncated work, *Maqasīd al-Falāsifa* ("The Intentions of Philosophers"). As a result the true meaning of his work was completely misunderstood, and he was thought to be a philosopher, whereas in fact he was the most ardent critic of philosophy.

At the age of thirty-six, al-Ghazālī experienced a profound crisis, provoked by the problem of intellectual certitude. He abandoned his professorship and his position as rector of Niẓāmīya University of Baghdad. During a period of ten years, clothed in the characteristic wool garment of the Sufis and completely absorbed in spiritual practices, he made solitary pilgrimages throughout the Muslim world, to Syria, Egypt, Mecca, and Medina. What he conveyed in his doctrines cannot be separated from this pathetic experience. He solved the problem of knowledge and certitude by affirming a degree of comprehension that left the heart no room for doubt, a comprehension that is the essential apprehension of things. The thinking soul becomes the focus of the universal Soul's irradiations, the mirror of intelligible forms received from the universal Soul. This theme dominates certain characteristic short treatises (the *Monqidh,* or "Preservative from Error," and the *Risālat al-Ladonīya*) as well as the great synthesis titled *Ihyāʾ ʿUlūm ad-Dīn* ("Revival of the Religious Sciences"). But this theme had already been treated, undoubtedly without his knowledge, by the Imāms of Shiʿism, and it does not differ essentially from the *Ishrāq* of Sohrawardī. This very theme led Sohrawardī to advance philosophy on a new basis rather than destroy the efforts of philosophers as such.

It is principally this aspect of al-Ghazālī's work, developed in his *Tahāfut al-Falāsifa* ("Autodestruction of the Philosophers") that Westerners have been inclined to emphasize. An attempt has even been made to read into it a more incisive and decisive critique or metaphysics than that of Immanuel Kant. In fact, al-Ghazālī strove vehemently to destroy the demonstrative range that philosophers, Avicennians as well as others, accorded to their arguments regarding the eternity of the world, the procession of the Intelligences, the existence of purely spiritual substances, and the idea of spiritual resurrection. In general al-Ghazālī strove to refute the idea of any causality, of any necessary connection. According to him all that can be experimentally affirmed is, for example, that combustion of cotton occurs *at the moment of* contact with

fire; it cannot be shown that combustion takes place *because of* the contact between cotton and fire. Nor can it be shown that there is any cause whatsoever. From this bursts forth the paradox of a thinker who professes the inability of reason to attain certitude while maintaining the certitude of destroying, with massive doses of rational dialectic, the certitudes of the philosophers. Averroes clearly discerned this self-contradiction and replied to it with his celebrated *Tahāfut al-Tahāfut* ("Autodestruction of the Autodestruction").

The same paradox is apparent in al-Ghazālī's other polemical works; in the "Courteous Refutation of the Divinity of Jesus Christ according to the Gospel"; in his treatise in Persian against all sorts of "freethinkers," or heretical thinkers (*Ibāhīya*); and, finally, in the treatise against the Ismāʿīlites (the Bātinites, or "esoterics"). The last treatise was overly influenced by the fact that it had been commissioned for political reasons by the ʿAbbāsid caliph al-Mostażhir, and the savage *dialectic,* deployed against an essentially *hermeneutic* Shiʿite thought, rings false. The Ismāʿīlites met this attack in the twelfth century with a monumental response (a work of the fifth Yemenite Dāʿī, in 1,500 pages), which unfortunately, is still unedited.

In any case, these polemical works had but a limited echo; al-Ghazālī's influence made itself felt principally through the *Iḥya.* Without doubt this influence was, and remains, considerable in Sunnite Islam. In Shiʿite Islam, notably in Iran, it was another matter. First of all, his effort did not respond to the same necessity, since the teaching of the Imāms of Shiʿism had already opened the way to spiritual Islam. But his effort was not ignored in Shiʿism, especially in the Ispahan School. Moḥsen Fayż (d. 1091 AH/1680 CE), one of the most celebrated pupils of Mullā Ṣadrā Shīrāzi (d. 1050 AH/1640 CE), even went so far as to rewrite the whole *Iḥya* with a Shiʿite interpretation. (Certain authors believe with him, assuming the authenticity of the book titled *Sirr al-ʿĀlamayn,* "Secret of the Two Universes," that al-Ghazālī would finally have rallied to Shiʿism.) In any case, in Iran no one ever thought or heard it said, as in the West, that the Ghazalian critique might have rendered impossible the continuation of philosophy in Islam and that Islamic philosophy was perhaps obliged to transport itself to Andalusia, where its last flames glowed with Ibn Bājja, Ibn Ṭufayl, and Averroes. Avicennianism, for example, enriched and modified by diverse contributions, continued to develop in Shiʿite Iran, not only during the Safavid epoch but also afterward, even to this day.

See also al-Ghazālī, Ahmad; Averroes; Ibn Bājja; Ibn Ṭufayl; Suhrawardī, Shihāb al-Dīn Yaḥyā.

Bibliography

For works by al-Ghazālī, see *Al-Ghazali's Tahafut al-Falasifah* (*Incoherence of the Philosophers*), translated by S. A. Kamali (Lahore, Pakistan: Philosophical Congress, 1963); *Al-Ghazali's Iḥyāʾ ʿUlūm al-Dīn, Book XX,* translated and annotated by L. Zolondek (Leiden: Brill, 1963); W. Montgomery Watt, *The Faith and Practice of al-Ghazāli* (London: Allen and Unwin, 1951), which contains translations of the *Monqidh* under the title "Deliverance from Error" and of "The Beginning of Guidance"; Ignaz Goldziher, *Streitschrift des Ghazālī gegen die Bāṭinijja-Sekte* (Leiden: Brill, 1916; Neudruck, 1956), an abbreviated edition with introduction and summary of the work against the Ismāʿīlites; and *Réfutation excellente de la divinité de Jésus-Christ d'après les Évangiles,* a translation and commentary on the "Courteous Refutation," by R. Chidiac (Paris: Leroux, 1939). See also *Averroes' Tahafut al-Tahafut (The Incoherence of the Incoherence),* translated by Simon van den Bergh, 2 vols. (London: Luzac, 1954), which incorporates most of al-Ghazālī's *Tahāfut al-Falāsifah.*

For works on al-Ghazālī, see Henry Corbin, *Histoire de la philosophie islamique,* Vol. I (Paris: Gallimard, 1964), pp. 251–261, with a detailed bibliography on pp. 358–359. See pp. 278–283 for Ahmad al-Ghazālī. See also W. Montgomery Watt, *Islamic Philosophy and Theology,* pp. 114–124, which includes bibliography (Edinburgh: Edinburgh University Press, 1962); this is Vol. I of *Islamic Surveys.*

Henry Corbin (1967)

AL-GHAZĀLĪ, MUḤAMMAD [ADDENDUM]

For comparisons with the western Christian tradition, Augustine comes more readily to mind than Aquinas, yet al-Ghazālī fulfills something of the role of each. He realized that understanding can be perfected in a faithful response to divine revelation, and that human reason can elucidate that response by showing the way through many pitfalls. Al-Ghazālī is aware of the deleterious effect of a simple reading of the scriptures, and so helps his readers to a sophisticated yet respectful grasp of the Word of God in the Qurʾan, all the while insisting that variant readings need to be discerned by careful intellectual examination. He is acutely aware of the way in which ordinary philosophical categories need to be stretched to accommodate the "creator of heaven and earth," and so of the necessary negative moments in using the names which the Qurʾan itself gives to God. Al-Ghazālī's recommended way to engage in that negative moment is via Sufi meditation, which can alert both mind and heart to their

inadequacy as well as bolster both to continue the journey toward proximity with the divine. In this respect he can also be favorably compared with Moses Maimonides, who was probably cognizant of at least some of al-Ghazālī's writings.

See also al-Ghazālī, Ahmad.

Bibliography

RECENT TRANSLATIONS OF AL-GHAZĀLĪ

Aims of the Philosophers. Translated by David Burrell and Tony Street. Provo, UT: Brigham Young University Press, 2004.

Incoherence of the Philosophers. Translated by Michael Marmura. Provo, UT: Brigham Young University Press, 1997.

On the Boundaries of Theological Tolerance in Islam. Translated by Sherman Jackson. Oxford: Oxford University Press, 2002.

TRANSLATIONS OF KEY BOOKS FROM THE *IHYA'*

Deliverance from Error. Translated by R. J. McCarthy. Louisville, KY: Fons Vitae, 1999.

Disciplining the Soul and *Breaking the Two Desires*. Translated by T. J. Winter. Cambridge, U.K.: Islamic Texts Society, 1995.

Faith in Divine Unity and Trust in Divine Providence. Translated by David Burrell. Louisville, KY: Fons Vitae, 2001.

Invocations and Supplications. Translated by K. Nakamura. Cambridge, U.K.: Islamic Texts Society, 1990.

Niche of Lights. Translated by David Buchman. Provo, UT: Brigham Young University Press, 1998.

Ninety-nine Beautiful Names of God. Translated by David Burrell and Nazih Daher. Cambridge, U.K.: Islamic Texts Society, 1992.

Remembrance of Death and the Afterlife. Translated by T. J. Winter. Cambridge, U.K.: Islamic Texts Society, 1989.

COMMENTARIES

Frank, Richard. *Al-Ghazali and the Ash'arite School*. Durham, NC: Duke University Press, 1994.

Gianotti, Timothy. *Al-Ghazali's Unspeakable Doctrine of the Soul*. Leiden, Netherlands: Brill, 2001.

Shehadi, Fadlou. *Ghazali's Unique Unknowable God*. Leiden, Netherlands: Brill, 1964.

David Burrell (2005)

ALIENATION

The term *alienation* (estrangement) has many different meanings in everyday life, in science, and in philosophy; most of them can be regarded as modifications of one broad meaning which is suggested by the etymology and the morphology of the word—the meaning in which alienation (or estrangement) is the act, or result of the act, through which something, or somebody, becomes (or has become) alien (or strange) to something, or somebody, else.

In everyday usage alienation often means turning away or keeping away from former friends or associates. In law it usually refers to the transfer of property from one person to another, either by sale or as a gift. In psychiatry alienation usually means deviation from normality; that is, insanity. In contemporary psychology and sociology it is often used to name an individual's feeling of alienness toward society, nature, other people, or himself. For many sociologists and philosophers, alienation is the same as reification: the act (or result of the act) of transforming human properties, relations, and actions into properties and actions of things that are independent of man and that govern his life. For other philosophers, "alienation" means "self-alienation" (self-estrangement): the process, or result of the process, by which a "self" (God or man) through itself (through its own action) becomes alien (strange) to itself (to its own nature).

HISTORY OF THE CONCEPT

The concept of alienation was first philosophically elaborated by Georg Wilhelm Friedrich Hegel. Some writers have maintained that the Christian doctrine of original sin and redemption can be regarded as a first version of Hegel's doctrine of alienation and dealienation. According to others, the concept of alienation found its first expression in Western thought in the Old Testament concept of idolatry. Still others have maintained that the source for Hegel's view of nature as a self-alienated form of Absolute Mind can be found in Plato's view of the natural world as an imperfect picture of the sublime world of Ideas. As investigation continues, probably more forerunners of Hegel will be discovered. But it seems established that Hegel, Ludwig Feuerbach, and Karl Marx were the three thinkers who first gave an explicit elaboration of alienation and whose interpretation is the starting point for all discussions of alienation in present-day philosophy, sociology, and psychology.

HEGEL. It is a basic idea of Hegel's philosophy that whatever is, is, in the last analysis, Absolute Idea (Absolute Mind, Absolute Spirit, or, in popular language, God) and that Absolute Idea is neither a set of fixed things nor a sum of static properties but a dynamic Self, engaged in a circular process of alienation and dealienation. Nature is only a self-alienated (self-estranged) form of Absolute Mind, and man is the Absolute in the process of dealienation. The whole of human history is the constant growth of man's knowledge of the Absolute and, at the same

time, the development of self-knowledge of the Absolute, who through finite mind becomes self-aware and "returns" to himself from his self-alienation in nature. Finite mind, however, also becomes alienated. It is an essential characteristic of finite mind (man) to produce things, to express itself in objects, to objectify itself in physical things, social institutions, and cultural products; and every objectification is, of necessity, an instance of alienation: the produced objects become alien to the producer. Alienation in this sense can be overcome only in the sense of being adequately known. Again, it is the vocation of man as man to serve as the organon of the self-knowledge of the Absolute. To the extent that he does not perform this function, he does not fulfill his human essence and is merely a self-alienated man.

FEUERBACH. Feuerbach accepted Hegel's view that man can be alienated from himself, but he rejected both the view that nature is a self-alienated form of Absolute Mind and the view that man is Absolute Mind in the process of dealienation. Man is not self-alienated God. On the contrary, God is self-alienated man; he is man's essence absolutized and estranged from man. And man is not alienated from himself when he refuses to recognize nature as a self-alienated form of God; man is alienated from himself when he creates and puts above himself an imagined alien higher being and bows before that being as a slave. The dealienation of man consists in the abolition of that estranged picture of man which is God.

MARX. Marx praised Hegel for having grasped that the self-creation of man is a process of alienation and dealienation. But he criticized Hegel for, among other things, having identified objectification with alienation and the suppression of alienation with the abolition of objectivity, for having regarded man as self-consciousness and the alienation of man as the alienation of his self-consciousness, and for having assumed that the suppression of objectification and alienation is possible only and merely in the medium of pure thought. Marx agreed with Feuerbach's criticism of religious alienation, but he stressed that the religious alienation of man is only one among many forms of man's self-alienation. Man not only alienates a part of himself in the form of God; he also alienates other products of his spiritual activity in the form of philosophy, common sense, art, morals, and so on. He alienates products of his economic activity in the form of commodities, money, capital, etc.; he alienates products of his social activity in the form of the state, law, and social institutions. Thus, there are many forms in which man alienates from himself the products of his own activity and makes of them a separate, independent, and powerful world of objects toward which he is related as a slave, powerless and dependent.

Nevertheless, man not only alienates his own products from himself; he also alienates himself from the very activity through which these products are produced, from the natural world in which he lives, and from other men. All these kinds of alienation are, in the last analysis, one; they are only different aspects of man's self-alienation, different forms of the alienation of man from his human "essence" or "nature," from his humanity. The self-alienated man is a man who is really not a man, a man who does not realize his historically created human possibilities. A nonalienated man would be a man who really is a man, a man who fulfills himself as a free, creative being of praxis.

The concepts of alienation and dealienation were elaborated by Marx in his early writings, especially in his *Economic and Philosophical Manuscripts*, written in 1844 and first published in 1932. In his later works the two concepts were basic, but they were used implicitly rather than explicitly. Their importance was therefore overlooked. In no exposition or interpretation of Marx's views written in the nineteenth century or in the first three decades of the twentieth did the concepts of alienation and dealienation play any important role. But since the publication of the *Manuscripts* and especially since World War II, they have become the object of passionate discussions, not only among Marxists but also among non-Marxists (especially existentialists and personalists), and not only among philosophers but also among psychologists (especially psychoanalysts), sociologists, literary critics, and writers.

CONTEMPORARY INTERPRETATIONS AND DEFINITIONS

Present-day writers who use the term *alienation* differ very much in the ways in which they understand and define it. Some authors think that the concept can be applied both to man and to nonhuman entities (to God, world, and nature, for instance); but most writers insist that it is applicable only to humans. Some of those who apply it only to humans insist that it can refer only to individuals and not to society as a whole. According to a number of such authors, the nonadjustment of the individual to the society in which he lives is a sign of his alienation. Others maintain that a society also can be alienated, or "sick," so that an individual who cannot adapt to the existing society is not, of necessity, alienated.

Many of those who regard alienation as applicable merely to individuals conceive it as a purely psychological concept referring to a feeling, or a state of mind. Others insist that alienation is not only a feeling but that it is also an objective fact, a way of being. Some of the writers who characterize alienation as a state of mind regard it as a fact or concept of psychopathology; others insist that although alienation is not good or desirable, it is not strictly pathological. They often add that one should distinguish alienation (a psychological state of the individual characterized by feelings of estrangement) both from anomie (relative normlessness in a social system) and from personal disorganization (disordered behavior arising from conflict within the individual).

Those who oppose characterizing alienation as a psychological concept usually say that it is also (or primarily) an economic, or political, or sociological, or ethical concept. Some insist that it is basically a concept of general philosophy, or a concept of ontology and philosophical anthropology.

According to Gwynn Nettler, alienation is a certain psychological state of a normal person, and an alienated person is "one who has been estranged from, made unfriendly toward, his society and the culture it carries" ("A Measure of Alienation," p. 672). For Murray Levin, "the essential characteristic of the alienated man is his belief that he is not able to fulfill what he believes is his rightful role in society" (*Man Alone*, p. 227). According to Eric and Mary Josephson, alienation is "an individual feeling or state of dissociation from self, from others, and from the world at large" (Introduction to *Man Alone*, p. 13). For Stanley Moore, the terms *alienation* and *estrangement* "refer to the characteristics of individual consciousness and social structure typical in societies whose members are controlled by, instead of controlling, the consequences of their collective activity" (*The Critique of Capitalist Democracy*, p. 125). According to Jean-Yves Calvez, alienation is "a general type of the situations of the absolutized subject who has given a world to himself, a formal world, refusing in this way the true concrete and its requirements" (*La pensée de Karl Marx*, p. 51); and according to Erich Fromm, "Alienation (or 'estrangement') means, for Marx, that man does *not* experience himself as the acting agent in his grasp of the world, but that the world (nature, others and he himself) remain alien to him. They stand above and against him as objects, even though they may be objects of his own creation. Alienation is essentially experiencing the world and oneself passively, receptively, as the subject separated from the object" (*Marx's Concept of Man*, p. 44).

With such a variety of definitions, it is difficult to say which is the best one. One may reserve the term for a specific phenomenon in which one is interested and, consequently, define it in such a narrow way as to make the majority of existing uses of "alienation" entirely inadmissible; or one may define it so broadly as to make as many as possible of the existing uses at least partly admissible and then distinguish between different forms of alienation in order to account for the variety of phenomena and to prevent possible confusions. The latter course seems more promising.

FORMS OF ALIENATION

All authors who have used the concept of alienation have distinguished between different forms of alienation; but not all of them have done so explicitly. Hegel attempted no explicit classification of the forms of alienation; but since, for him, the essence of all development was a process of alienation and dealienation, different stages in the development of the Absolute could be regarded as so many forms of alienation. It would be much more difficult to develop a similar classification for Feuerbach's works because the essence of his philosophy was negation of systematic philosophy. "Alienated Labor," a well-known fragment in Marx's *Economic and Philosophic Manuscripts*, seems to suggest that we should distinguish between four forms of man's alienation: the alienation of man from the products of his own activity, the alienation of man from his productive activity itself, the alienation of man from his human essence, and the alienation of man from other men. But in other places Marx talked about other forms and subforms of alienation not mentioned in this fragment. The enumeration seems to be defective also in that it puts on the same level forms of alienation that should not be at the same level.

Twentieth-century writers differed greatly in their enumeration of the basic forms of alienation. Frederick A. Weiss distinguished three basic forms (self-anesthesia, self-elimination, and self-idealization); Ernest Schachtel distinguished four (the alienation of men from nature, from their fellow men, from the work of their hands and minds, and from themselves); Melvin Seeman, five (powerlessness, meaninglessness, social isolation, normlessness, and self-estrangement); and Lewis Feuer, six (the alienation of class society, of competitive society, of industrial society, of mass society, of race, and of generations).

In listing five different forms of alienation, Seeman tried to define them strictly. According to him, powerlessness is "the expectancy or probability held by the individ-

ual that his own behavior cannot determine the occurrence of the outcomes, or reinforcements, he seeks"; meaninglessness results "when the individual is unclear as to what he ought to believe—when the individual's minimal standards for clarity in decision-making are not met"; normlessness is the characteristic of a situation "in which there is a high expectancy that socially unapproved behaviors are required to achieve given goals"; isolation is characteristic of those who "assign low reward value to goals or beliefs that are typically highly valued in the given society"; and self-estrangement is "the degree of dependence of the given behavior upon anticipated future rewards, that is upon rewards that lie outside the activity itself" ("On the Meaning of Alienation," pp. 786, 788, 789, 790).

Instead of trying to enumerate all classifications of the forms of alienation that have been made so far, we shall only mention a few of the basic criteria according to which such classifications could be made and actually have been made.

(1) According to the nature of that which is alienated, we may distinguish between alienation of things and alienation of selves. And if we distinguish different types of things or selves, we may add further subdivisions. To those for whom the only self is man, alienation of self is only another name for the alienation of man. But they may distinguish between individual alienation and social alienation. We may classify as types of social alienation the alienation of societies as a whole (such as feudal societies and capitalist societies), the alienation of social groups (capitalists, workers, intellectuals, bureaucrats, producers, consumers, etc.), and the alienation of social institutions (such as the state, the church, and cultural institutions).

(2) According to the question, we can distinguish between alienation from something else or somebody else and alienation from oneself. The distinction is applicable only to alienation of selves; a thing cannot be alienated from itself. A self can be alienated either from something or somebody or from itself. According to the different kinds of "others" and according to the different aspects or sides of the self, further subdivisions can be added (for example, alienation from nature, alienation from fellow men, or alienation of the self from its body, its feelings, its needs, or its creative possibilities).

(3) According to whether that which is alienated is alienated through its own activity or through the activity of another, we could distinguish between alienation through others and alienation through oneself. Alienation of a thing can obviously be only an alienation through others. There can be different kinds of alienation of things (stealing, giving, and buying and selling). Alienation of self can be either alienation through others or an alienation through oneself.

SELF-ALIENATION

The concept of self-alienation, found in Hegel and Marx and of the greatest interest for philosophy, is a result of applying a combination of the above three basic criteria. What Hegel and Marx called self-alienation is alienation *of* a self *from* itself *through* itself. They differ in that Marx recognized only one self-alienated self (man), while Hegel recognized two (man and God, or Absolute). Some writers hold that one could also speak about self-alienation of nature or of the world. In religious myths we find self-alienated angels (for example, Lucifer), and in children's stories and fables we find self-alienated animals (the cowardly lion, the naive fox) and even plants (a humpy fir tree, a stinking rose). But the concept of a self-alienated man is basic.

In what sense is it possible for a self (either an individual man or a society) to be alienated from itself? It seems plausible to say that to be self-alienated means to be internally divided, split into at least two parts that have become alien to each other. But in that case, why talk of self-alienation; why not, instead, simply refer to an internal division or split? The term *self-alienation* seems to suggest some or all of the following points. (1) The division of the self into two conflicting parts was not carried out from the outside but is the result of an action of the self. (2) The division into conflicting parts does not annihilate the unity of the self; despite the split, the self-alienated self is nevertheless a self. (3) Self-alienation is not simply a split into two parts that are equally related to the self as a whole; the implication is that one part of the self has more right to represent the self as a whole, so that by becoming alien to it, the other part becomes alien to the self as a whole.

One way to specify and clarify the inequality of the two parts into which a self-alienated self is split is to describe the self-alienation as a split between man's real "nature," or "essence," and his factual "properties," or "existence." The self-alienated man in such a case is a man who is not in fact what he is in essence: a man whose actual existence does not correspond to his human essence. Similarly a self-alienated society would be a society whose factual existence does not correspond to the real essence of human society.

How can the actual existence of man deviate from his real essence or nature? If one were to conceive man's

essence as something shared by all men, then somebody alienated from man's essence could not be a man in fact. Accordingly, if alienation of man from his essence is possible, his essence must not be conceived as something that all men have in common.

One possible interpretation would be the conception of man's essence as an eternal or nontemporal idea of man toward which the real man ought to strive. This interpretation is full of difficulties and leads to unanswerable questions, such as Where and in what way does such an idea of man exist? What is the way or method to achieve an adequate knowledge of it? Why should a real man strive toward it?

Another interpretation would consist in conceiving man's essence as something actually belonging to men—not to all, but only to some men; for example, to the majority of all so-far-existing men or to the majority of future men. Whichever interpretation one chooses, new difficulties arise. Why should a majority be more representative of the nature of man than a minority? If we already allow the split into essence and existence, why should we not also allow the possibility of the split being present in the majority? And why should a future actuality have any advantage over the past and the present one?

The third, and perhaps the most promising, interpretation consists in saying that man's essence is neither an eternal idea nor a part of actuality, but the sum of historically created human possibilities. To say that a man alienates himself from his human essence would then mean that a man alienates himself from the realization of his historically created human possibilities. To say that a man is not alienated from himself would mean that a man stands on the level of his possibilities and that in realizing his possibilities he permanently creates new and higher ones. The third interpretation seems more plausible than the first two, but it too leads to difficulties. In what way do the possibilities exist, and how do we discover them? On what basis do we divide man's real possibilities into human and inhuman possibilities?

SELF-ALIENATION AND HISTORY

Another much-discussed question asks whether self-alienation is an essential, imperishable property of man as man or whether it is characteristic only of one historical stage in man's development. Some philosophers, especially existentialists, have maintained that alienation is a permanent structural moment of man's existence. Man as man is necessarily self-alienated; in addition to his authentic existence he leads a nonauthentic one, and it is an error to expect that he will one day live only authentically.

Opposed to this view is the view that the originally nonself-alienated man, in the course of development, alienated himself from himself, but that he will return to himself in the future. This view was held by Friedrich Engels and is accepted by many contemporary Marxists; Marx himself seems to have been inclined to think that man had always been self-alienated, but that in spite of this, he can and ought to overcome his self-alienation in the future. In this sense, Marx, in *Economic and Philosophical Manuscripts*, wrote about communism as the positive supersession of all alienation and the return of man from religion, family, state, etc., to his human (that is, social) existence. Such a conception of communism as a dealienation of human community formed the basis of all of Marx's other works.

ALIENATION IN PAST AND PRESENT. If we assume that the whole of history up to now has been a history of humanity's self-alienation, then it may be asked whether history has been characterized by the gradual elimination of alienation or by its permanent deepening. Those who believe in constant progress have maintained that alienation has always been diminishing. But many contemporary philosophers and sociologists have found that alienation has constantly increased, so that it is much deeper and more pervasive than ever before in contemporary capitalism and bureaucratic socialism. A third group of authors have maintained that alienation has diminished in some respects and increased in others. Some have insisted that the question cannot be answered simply in terms of more or less, that we should investigate different types of self-alienated individuals typical of different periods in human history. An interesting attempt in this direction was made by Erich Fromm, who distinguished four basic types of "nonproductive" (self-alienated) character orientations (the receptive, hoarding, exploitative, and marketing orientations), each typical of a successive stage of historical development. According to Fromm, all four are found in contemporary self-alienated society, but whereas the first three were inherited from earlier periods, the marketing orientation is "definitely a modern product," typical of twentieth-century capitalism (*Man for Himself*, pp. 62–81).

ALIENATION IN THE FUTURE. For those who regard alienation as a historical phenomenon, the question about a possible end of alienation (dealienation or disalienation) naturally arises. Two main answers have been given.

According to one group of thinkers, absolute dealienation is possible; all alienation, both social and individual, can be once and for all abolished. The most radical among this group have even maintained that all alienation has already in principle been eliminated in socialist countries, that it exists there only as a case of individual insanity or as an insignificant remnant of capitalism. More realistic representatives of this view have not denied facts showing that in countries considering themselves socialist, many old forms and even some new forms of alienation exist. But they have insisted that in more mature stages of socialism all these forms of alienation are destined to disappear.

According to a second group, only a relative dealienation is possible. It is impossible to eliminate alienation completely and finally because human nature is not something given and unchangeable that can be fulfilled once and for all. It is possible, however, to create a basically nonalienated society that would stimulate the development of nonalienated, really human individuals.

OVERCOMING ALIENATION. The means recommended for overcoming self-alienation differ according to one's view of the essence of self-alienation.

Those who regard self-alienation as a psychological fact, as a fact of the life of the individual human self, dispute the importance or even the relevance of any external changes in circumstances and suggest the individual's own moral effort, a revolution within the self, as the only cure. Those who regard self-alienation as a result of the neurotic process are quite consistent in offering a psychoanalytical medical treatment; they regard the new creative experience of acceptance and meeting in a warm, truly mutual and trusting doctor–patient relationship as the main therapeutic factor.

Diametrically opposed to this view are those philosophers and sociologists who, basing their ideas on a degenerate variant of Marxism called economic determinism, hold that individuals are the passive products of the social organization, that the whole of social organization is determined by the organization of economic life, and that all economic life is dependent on the question of whether the means of production are or are not private property. For economic determinists, the problem of dealienation is reduced to the problem of social transformation, and the problem of social transformation is reduced to the abolition of private property.

In criticizing "the materialist doctrine that men are products of circumstances and upbringing," Marx stressed that "it is men that change circumstances," so that "the coincidence of the changing of circumstances and of human activity can be conceived and rationally understood only as *revolutionizing practice* (Praxis)" (*Basic Writings on Politics and Philosophy*, with Engels, New York, 1959, p. 244).

Those who have tried to elaborate such a conception have insisted that dealienation of the society and dealienation of individuals are closely connected: One cannot be carried out without the other or reduced to the other. It is possible to create a social system that would enable and even stimulate the development of dealienated individuals, but it is impossible to organize a society that would automatically produce such individuals. A nonalienated individual is an individual who fulfills himself as a free and creative being of praxis, and free creativity is not something that can be given as a gift or forced upon anyone from outside. An individual can become free only through his own activity.

It is not simply that dealienation of individuals cannot be reduced to dealienation of society; the dealienation of society, in turn, cannot be conceived as a change in economic organization that will automatically be followed by change in all other fields and aspects of social life. Far from being an eternal fact of social life, the split of society into mutually independent and conflicting spheres and the predominance of the economic sphere is, according to Marx, a characteristic of a self-alienated society. Therefore, the dealienation of society is impossible without abolishing the alienation of the different human activities from each other.

Finally, the problem of dealienation of economic life cannot be solved by the abolition of private property. The transformation of private property into state property does not introduce an essential change in the situation of the working man, the producer. The dealienation of economic life also requires the abolition of state property, that is, its transformation into real social property; and this can be achieved only by organizing the whole of social life on the basis of the self-management of immediate producers.

See also Absolute, The; Engels, Friedrich; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Marx, Karl; Ontology; Philosophical Anthropology; Plato.

Bibliography

ANTHOLOGIES

Josephson, Eric, and Mary Josephson, eds. *Man Alone: Alienation in Modern Society*. New York: Dell, 1962.

Sykes, Gerald, ed. *Alienation: The Cultural Climate of Our Time*. 2 vols. New York: G. Braziller, 1964.

CLASSICAL WORKS

Feuerbach, Ludwig. *Das Wesen des Christentums*. Leipzig, 1841. Translated by Marian Evans as *The Essence of Christianity*, 2nd ed. London, 1882.

Hegel, G. W. F. *Encyclopädie der philosophischen Wissenschaften im Grundrisse*. Heidelberg, 1817.

Hegel, G. W. F. *Die Phänomenologie des Geistes*. Bamberg and Würzburg, 1807. Translated by J. B. Baillie as *Phenomenology of Mind*. London: S. Sonnenschein, and New York: Macmillan, 1910; 2nd ed., 1931.

Hegel, G. W. F. *Theologische Jugendschriften*, edited by H. Nohl. Tübingen: J.C.B. Mohr, 1907.

Marx, Karl. *Grundrisse der Kritik der politischen Ökonomie (Rohentwurf) 1857–1858*. 2 vols. Moscow: Verlag für Fremdsprachige Literatur, 1939–1941; 2nd ed., Berlin, 1953.

Marx, Karl. *Das Kapital*. 3 vols. Hamburg, 1867–1894. Translated by S. Moore, E. Aveling, and E. Untermann as *Capital*. 4 vols. London, 1887–1909. Especially Vol. I, Ch. 1, Sec. 4.

Marx, Karl. *Die oekonomisch-philosophischen Manuskripte aus dem Jahre 1844*. In *Historisch-kritische Gesamtausgabe*, edited by Karl Marx and Friedrich Engels. Berlin, 1932. Div. I, Vol. III. Translated by Martin Milligan as *Economic and Philosophical Manuscripts of 1844*. London, 1959.

Marx, Karl, and Friedrich Engels. *Die deutsche Ideologie* (1844–1845). In their *Historisch-kritische Gesamtausgabe*. Berlin, 1932. Div. I, Vol. III. Translated as *The German Ideology*, edited by R. Pascal. London: Lawrence and Wishart, 1938.

Works on Hegel, Marx, and Engels

Calvez, Jean-Yves. *La pensée de Karl Marx*. Paris: Éditions du Seuil, 1956.

Fromm, Erich. *Marx's Concept of Man*. New York: F. Ungar Publishing, 1961.

Hyppolite, Jean. *Études sur Marx et Hegel*. Paris: M. Rivière, 1955.

Kangrga, Milan. *Eticki Problem u Djelu Karla Marxa*. Zagreb, 1963.

Lukács, Georg. *Der junge Hegel und die Probleme der kapitalistischen Gesellschaft*. Zürich and Vienna: Aufbau-Verlag, 1948.

Marcuse, Herbert. *Reason and Revolution*. New York: Oxford University Press, 1941.

Oizerman, T. I. *Formirovanie Filosofii Marksizma*. Moscow: Izd-vo sotsial'no ekon lit-ry, 1962.

Popitz, Heinrich. *Der entfremdete Mensch. Zeitkritik und Geschichtsphilosophie des jungen Marx*. Basel: Verlag für Recht und Gesellschaft, 1953.

Tucker, Robert. *Philosophy and Myth in Karl Marx*. Cambridge, U.K.: Cambridge University Press, 1961.

RECENT STUDIES

Arendt, Hannah. *The Human Condition*. Chicago: University of Chicago Press, 1958.

Fromm, Erich. *Escape from Freedom*. New York: Farrar and Rinehart, 1941.

Fromm, Erich. *Man for Himself*. New York: Rinehart, 1947.

Fromm, Erich. *Sane Society*. New York: Rinehart, 1955.

Goldmann, Lucien. *Recherches dialectiques*, 3rd ed. Paris, 1959.

Lefebvre, Henri. *Critique de la vie quotidienne*. Paris: B. Grasset, 1947; 2nd ed., 1958.

Lefebvre, Henri. *Le matérialisme dialectique*. Paris: F. Alcan, 1939.

Levin, Murray B. *The Alienated Voter*. New York: Holt, Rinehart and Winston, 1960.

Lukács, Georg. *Geschichte und Klassenbewusstsein*. Berlin, 1923.

Mills, C. Wright. *White Collar*. New York: Oxford University Press, 1951.

Moore, Stanley. *The Critique of Capitalist Democracy*. New York: Paine-Whitman, 1957.

Naville, Pierre. *De l'aliénation à la jouissance*. Paris: M. Rivière, 1957.

Pappenheim, Fritz. *The Alienation of Modern Man: An Interpretation Based on Marx and Tönnies*. New York: Monthly Review Press, 1959.

Riesman, David, with Nathan Glazer and Reuel Denney. *The Lonely Crowd*. New Haven, CT: Yale University Press, 1950.

Whyte, William H., Jr. *The Organization Man*. New York: Simon and Schuster, 1956.

ARTICLES

Bell, Daniel. "The 'Rediscovery' of Alienation: Some Notes along the Quest for the Historical Marx." *Journal of Philosophy* 56 (1959): 933–957.

Braybrooke, David. "Diagnosis and Remedy in Marx's Doctrine of Alienation." *Social Research* 25 (1958): 325–345.

Cornu, Auguste. "L'idée d'aliénation chez Hegel, Feuerbach et K. Marx." *La pensée* no. 2 (1948): 65–75.

Dean, Dwight. "Alienation and Political Apathy." *Social Forces* 38 (1960).

Dean, Dwight. "Meaning and Measurement of Alienation." *American Sociological Review* 26 (1961): 753–758.

Duhrsen, Alfred. "Philosophic Alienation and the Problem of Other Minds." *Philosophic Review* 69 (1960): 211–220.

Easton, Loyd D. "Alienation and History in the Early Marx." *Philosophy and Phenomenological Research* 22 (1961): 193–205.

Feuer, Lewis. "What Is Alienation? The Career of a Concept." *New Politics* 1 (3) (1962): 116–134.

Garaudy, Roger. "O Ponjatii Otčuždenie." *Voprosi Filosofii* no. 8 (1959): 68–81.

Glazer, Nathan. "The Alienation of Modern Man." *Commentary* 3 (April 1947).

Kraft, Julius. "Die Selbstentfremdung des Menschen." *Geist und Tat* (March 1956).

Löwith, Karl. "Man's Self-Alienation in the Early Writings of Marx." *Social Research* 21 (1954): 204–230.

Nettler, Gwynn. "A Measure of Alienation." *American Sociological Review* 22 (1957): 670–677.

Petrović, Gajo. "Man as Economic Animal and Man as Praxis." *Inquiry* 6 (1963): 35–56.

Petrović, Gajo. "Marx's Theory of Alienation." *Philosophy and Phenomenological Research* 23 (1963): 419–426.

Rose, Arnold M. "Alienation and Participation: A Comparison of Group Leaders and the 'Mass.'" *American Sociological Review* 27 (6) (1962): 834–838.

Seeman, Melvin. "On the Meaning of Alienation." *American Sociological Review* 24 (1959): 783–791.

Sommer, Robert, and Hall, Robert. "Alienation and Mental Illness." *American Sociological Review* 23 (1958): 418–420.

"Symposium on Alienation and the Search for Identity." *American Journal of Psychoanalysis* 21 (2) (1961).

Vignaud, P. "L'aliénation selon Karl Marx." *La vie intellectuelle* (February 1937).

G. Petrović (1967)

ALIOTTA, ANTONIO

(1881–1964)

Antonio Aliotta, the Italian philosopher, was born in Palermo and taught at the universities of Padua and Naples. Moving from studies in experimental psychology, *La misura in psicologia sperimentale* (1905), Aliotta published in 1912 a vast critical analysis of contemporary philosophy titled *La reazione idealistica contro la scienza* (English translation, London, 1914) in which he defended a monadological spiritualism with a theistic tendency. When the shadow of the neo-Hegelianism of Benedetto Croce and Giovanni Gentile began to loom over Italy, Aliotta took sides with the opponents of this idealism and in his teaching and writings spread the news of other philosophical movements going on outside Italy, especially the philosophy of science, realism, and pragmatism.

From 1917 to 1936, in the mature phase of his thought, Aliotta's sympathies were above all with pragmatism, and his experimentalism suggests many points of similarity with the philosophies of William James and George Herbert Mead. Experimentation is the only means of establishing the truth of any knowledge whatever, even metaphysical and religious. By "experimentation," Aliotta does not mean simply the techniques of the laboratory but any kind of trial-and-error procedure in any field of human activity. History is a kind of grand laboratory in which people seek, through conflict, to attain more harmonious forms of life.

The success of the experiment, according to Aliotta, consists in the elimination of conflict and in the realization of a certain degree of harmony. "The quest for truth," he says in *Relativismo e idealismo*, "is the quest for a superior harmony of active human and non-human forces, operating in the universe of our experience." Obviously, the presupposition is that experience is not a single and continuous process, but is composed of a plurality of individual centers that meet and limit each other by stages and, through conflicts, try to realize a growing coordination. Common sense, science, and philosophy are the steps, or phases, of this coordination. The "thing" of common sense makes possible a certain degree of coordination between individual intuitions. The syntheses of science represent a superior degree of coordination, since they eliminate the disparity between the perspectives of common sense; and philosophical inquiry seeks to collect the remaining dissident elements, to correct the restricted vision of the particular sciences, and to achieve a more comprehensive view. The concept limit toward which this process tends is the coordination of all activities and their convergence to a single end, which is, in other terms, the Leibnizian monad of monads, or God.

Aliotta insists, however, on the social character, in Mead's sense, of all degrees of knowledge. He denies the absoluteness of truth and defends philosophical relativism, of which he sees implicit proof in the physics of Albert Einstein; and he holds that the measure of truth is in every case determined by the degree of coordination that is experimentally realized between the intuitions, the perspectives, and the individual points of view that constitute the rough fabric of experience.

In later writings, for example, *Il sacrificio come significato del mondo* (1943), Aliotta sought to extend this point of view to ethics with an inquiry into what he calls "the fundamental postulates of action." The indeterminacy of the world and its relative uniformity, the value of the human person and the transcendence of reality, and the plurality of persons and their tendency toward unity are among these postulates, but *the* fundamental postulate is that of the "perennial character of human-values" and of the existence of God, which guarantees this character. The spiritualistic and fideistic aspect prevails over the pragmatic and methodological aspect in this final phase of Aliotta's thought.

See also Croce, Benedetto; Gentile, Giovanni; Hegelianism; Idealism; James, William; Mead, George Herbert; Philosophy of Science, History of; Pragmatism; Realism.

Bibliography

WORKS BY ALIOTTA

La misura in psicologia sperimentale. Florence, 1905.

La reazione idealistica contro la scienza. Palermo, 1912. Translated by Agnes McCaskill as *The Idealistic Reaction against Science*. London: Macmillan, 1914.

La guerra eterna e il dramma dell'esistenza. Naples, 1917.

Relativismo e idealismo. Naples: Perrella, 1922.

La teoria di Einstein. Palermo, 1922.

L'esperimento nella scienza, nella filosofia, e nella religione. Naples, 1936.

Opere complete (Complete works). 7 vols. Rome, 1942–1954.
Il sacrificio come significato del mondo. Rome, 1947.
Evoluzionismo e spiritualismo. Naples: Libreria Scientifica, 1948.
Le origini dell'irrazionalismo contemporaneo. Naples, 1950.
Pensatori tedeschi della fine dell'800. Naples: Libreria Scientifica, 1950.

WORKS ON ALIOTTA

Carbonara, Cleto et al. *Lo sperimentalismo di Antonio Aliotta*. Naples, 1951. Essays on the occasion of Aliotta's 80th birthday.
Sciacca, M. F. *Il secolo XX*, 2nd ed., Vol. 1, 470–490. Milan, 1947.

Nicola Abbagnano (1967)
Translated by Nino Langiulli

ALISON, ARCHIBALD
(1757–1839)

Archibald Alison was born in Edinburgh, Scotland, and educated at Glasgow and Balliol College, Oxford. He was ordained in the Church of England and held positions in both England and Scotland. He married a daughter of John Gregory (1724–1773), who was a professor of philosophy and medicine at Aberdeen and an associate of Thomas Reid in the Aberdeen Philosophical Society. Alison preached at the Cowgate Chapel in Edinburgh from 1800 until his death. He published a volume of sermons, but is known primarily for his "Essays on the Nature and Principles of Taste," published in 1790 and reissued in 1810.

Alison's theory of taste breaks with earlier eighteenth-century theories in several respects while retaining other characteristic features. Like his predecessors, Alison regards beauty and sublimity as essentially emotional, hedonic experiences. Beauty is a form of pleasure, and as such it is found not in objects but in the mind. He accepts a faculty psychology that is essentially associative, and he regards what he is doing as a scientific investigation of the principles of human nature. In addition, Alison is the first theorist to clearly separate what he calls the emotions of taste—beauty, sublimity, and so on—from other kinds of pleasure. Although earlier theories speak of the pleasures of the imagination as special pleasures and sometimes suggest distinctions from other pleasures, it is Alison who first clearly appeals to a separate aesthetic pleasure that in his words is distinct from "every other emotion of pleasure" (1790/1999, p. 407).

Alison also argues that the ideas required to produce the emotions of taste must be complex. A simple idea, such as that of a color, which may be pleasant in itself, is only felt as beautiful when it enters into an associative complex. Thus, he rejects both the view that taste is an effect of an internal sense and the view that some single principle, such as relation, utility, or order and design, produces the emotions of taste. Alison believes that the emotion he seeks to describe is very much a product of an active mind. So he distinguishes two elements in complex emotions such as beauty. One is a simple idea and its accompanying emotion. Almost any simple emotion will do, including painful as well as pleasurable emotions. But the complex emotion of taste only appears when the simple emotion is acted on by the faculty of the imagination to produce "a consequent excitement. … The *peculiar* pleasure of the beautiful or the sublime is only felt when these two effects are conjoined, and the complex emotion produced" (1790/1999, p. 408).

Alison's theory of the imagination moves away from the earlier eighteenth-century theories of imagination according to which imagination is essentially a faculty that recombines preexisting ideas into new, artificial images—for example, a centaur is a combination of the ideas of horse and man. Alison still thinks of imagination as producing new ideas, but his emphasis is on its ability to detect resemblances, "trains of imagery" (1790/1999, p. 412), and expressive signs. So the faculty of imagination is essentially an active, associative faculty and the peculiar pleasure that it produces arises from the activity of the mind itself.

Alison draws a conclusion, which parallels Immanuel Kant's theories in many respects, that for the imagination to do its work it must be "free and unembarrassed" (1790/1999, p. 412)—that is, disinterested—"so little occupied by any private or particular object of thought, as to leave us open to all the impressions which the objects that are before us can produce" (p. 412). Whereas the earlier theories that suggest the need for disinterestedness understand it as a negative condition—a condition of good taste (Third Earl of Shaftesbury [Anthony Ashley Cooper]) or an avoidance of prejudice (David Hume) and thus a part of a theory of criticism, Alison treats it as a condition of experience. It is what allows the imagination to form the associations that are a necessary condition for the production of the complex emotion of beauty or sublimity. Alison goes so far as to describe a kind of free play of the imagination, which is opposed to attention. For Alison, however, these are competing mental habits and not Kantian epistemological principles.

Alison does draw the conclusion, common to some twentieth-century aesthetic attitude theories, that criti-

cism is incompatible with the emotion of taste. Thus, taste ceases to be a form of critical judgment. He acknowledges that an active imagination does not necessarily produce good taste—the young are undiscriminating, for example—but he does not seem to recognize that on his theory taste has ceased to be what it had been since the Renaissance formation of the idea—a form of judgment with social implications.

Instead, Alison develops two essentially romantic theses: “matter is not beautiful in itself, but derives its beauty from the expression of mind” (1790/1999, p. 417) and the qualities of matter that are productive of beauty or sublimity are either themselves immediately expressive of mental qualities or powers—for example, the activity of creation in the arts or of the divine creator in nature; or they are signs of mental qualities—for example, the tone of voice. So Alison’s theory combines three elements: imagination, association, and expression. He concludes, “[T]he beauty and sublimity which is felt in the various appearances of matter, are finally to be ascribed to their expression of mind; or to their being, either directly or indirectly, the signs of those qualities of mind which are fitted, by the constitution of our nature, to affect us with pleasing or interesting emotion” (p. 419).

Alison anticipates Kant and many of the features of romantic and twentieth-century aesthetics, therefore, without completely abandoning the tradition of theories of taste with which he is most closely associated—particularly those of Alexander Gerard and Reid. Although there are extensive references to the fine arts, Alison’s theory of the arts remains a theory of imitation, not a theory of artistic creation or genius. Natural beauty provides the paradigm for beauty in the arts. The only creative mind is the divine mind; artists can only discover beauty, not create it. At the same time, however, imagination and expression are given a new scope. They are the necessary faculties for an artist. Artistic imitation is an active, not a passive mental operation.

Alison does not go far in formulating the epistemological requirements of his theory. He is not prepared to go as far as Samuel Taylor Coleridge and declare that the artist is a second creator. He takes for granted a theory of natural signs, found also in Reid and drawn from earlier theories, and he depends on a theory of association that is rapidly losing its grounding in the theory of ideas developed by John Locke and Hume. This produces some obscurity about what aesthetic qualities in objects are, a good deal of rhetorical excess, and an avoidance of the problems that exist for a theory of taste in which taste is no longer a form of judgment. But the new scope given to the imagination makes Alison one of the first to formulate a full theory of aesthetics as expression.

See also Aesthetics, History of.

Bibliography

WORKS BY ALISON

“Essays on the Nature and Principles of Taste” (1790). In *Eighteenth-Century British Aesthetics*, edited by Dabney Townsend. Amityville, NY: Baywood, 1999.

WORKS ABOUT ALISON

Dickie, George. *The Century of Taste: The Philosophical Odyssey of Taste in the Eighteenth Century*. New York: Oxford University Press, 1996.

Hipple, Walter J. *The Beautiful, the Sublime, and the Picturesque in Eighteenth-Century British Aesthetic Theory*. Carbondale: Southern Illinois University Press, 1957.

Kivy, Peter. *The Seventh Sense: Francis Hutcheson and Eighteenth-Century British Aesthetics*. 2nd ed. New York: Oxford University Press, 2003.

McCosh, James. “Archibald Alison” (1875). In *The Scottish Philosophy, Biographical, Expository, Critical, from Hutcheson to Hamilton*. Hildesheim, Germany: Georg Olms, 1966.

Stolnitz, Jerome. “On the Origins of Aesthetic Disinterestedness.” *Journal of Aesthetics and Art Criticism* 20 (2) (1961): 131–143.

Townsend, Dabney. “Archibald Alison: Aesthetic Experience and Emotion.” *Journal of Aesthetics* 28 (2) (1988): 132–144.

Dabney Townsend (2005)

AL-JABIRI, ʿABD
(1935–)

Muḥammad ʿAbd al-Jabiri studied philosophy at Muḥammad V University in Rabat, Morocco, where he got his PhD in 1970. He had been a school teacher since 1957 and after successive promotions he became professor of philosophy at that university in 1971. Al-Jabiri has been involved in politics and journalism, and he is the main editor of the journal *Fikr wa-Naqd* (Thought and criticism) published in Rabat. His philosophy has to be understood in the context of the effort to modernize his country while at the same time preserving its cultural identity.

Al-Jabiri is a prolific writer; his large project, *The Critique of the Arab Mind*, is in three volumes: *Formation of the Arab Mind* (1984), *Structure of the Arab Mind* (1986), and *The Arab Political Mind* (1990). Al-Jabiri emphasizes the concept of cultural legacy (*turāth*) and analyses different readings of it. The fundamentalists (*al-salafīya*) search for the pristine Islam and they commit a

grave mistake because they ignore the historical factor. The original "authentic" form of Islam was valid in its time, but the fundamentalists do not see it as subject to the course of history, they consider its initial form perpetually valid. The liberals and the Orientalists read cultural legacy from the Western standpoint. Arab liberals suffer under such cultural alienation that they cannot perceive their own identity. As for the Marxists they expect tradition to develop into revolution and the revolution to develop into tradition, and they cannot escape this vicious circle.

Al-Jabiri's reading is based on his criticism of Arab rationality, or mind (*ʿaql*). To this purpose he follows a methodology to liberate the reader-subject from being a hostage as the read-object, that is, Arabic language and Arabic tradition. After gaining objectivity, the reader rejoins the object, apprehends it by means of intuition (*h?ads*), and recognizes the historicity of reason. According to him Arab reason started as a political instrument. Two trends existed within the Umayyad regime: the one rationalist and reformist—Muʿtazilite—and the other traditional and conservative—Sunnite; the Sunnites were in power, and the Muʿtazilites in opposition. When the Abbasids overthrew the Umayyads, the Muʿtazilites moved to the governing side, and the Sunnites to the opposition. Nevertheless, since the Muʿtazilites were not strong enough to face the challenge of esoteric movements, the caliph [Abū] al-Ma'mūn (786–833) turned to the philosophy of Aristotle for help.

For al-Jabiri philosophy in the Islamic East is radically different from that in the West. Avicenna in the East wanted to create the "Oriental" philosophy by combining Platonic philosophy with the Aristotelian and integrating esoteric Gnostic doctrines and Muʿtazilite theology; it survived only in Iranian Gnosticism. By contrast, Averroes in the West succeeded in standing by Aristotle and abandoning the other doctrines and solved the long-lasting issue of the relationship between revealed religion and philosophy by proving their coherence and continuity. Thus, al-Jabiri asserts that the future of Arab philosophy lies in Averroes's philosophical method and his rationalism (*ʿaqlānīya*).

See also Aristotle; Averroes; Averroism in Modern Islamic Philosophy; Gnosticism; Islamic Philosophy; Marxist Philosophy; Rationalism; Rationality.

Bibliography

WORKS BY AL-JABIRI

Arabic-Islamic Philosophy: A Contemporary Critique. Translated by Aziz Abbassi. Austin: University of Texas Press, 1999.

WORKS ABOUT AL-JABIRI

Ábed Yabri, Mohamed. *El legado filosófico árabe: Alfarabi, Avicena, Avempace, Averroes, Abenjaldún: Lecturas contemporáneas*. Madrid, Spain: Editorial Trotta, 2001.

Josep Puig Montada (2005)

AL-KINDĪ, ABŪ-YŪSUF YAʿQŪB IBN ISHĀQ

(ninth century)

Abū-Yūsuf Yaʿqūb ibn Ishāq al-Kindī was the first outstanding Arabic-writing philosopher. He was born in the Mesopotamian city of Basra and later held a distinguished position at the caliph's court in Baghdad, where he died shortly after 870. For about a century he enjoyed a reputation as a great philosopher in the Aristotelian-Neoplatonic tradition. He appears to have been the first to introduce the late Greek syllabus of philosophical learning into the Muslim world. It was mainly, though not exclusively, based on the *Corpus Aristotelicum* and its Peripatetic and Neoplatonic commentators. Numerous competent Arabic versions of Greek philosophical texts were available then, and al-Kindī himself commissioned translations of Aristotle's *Metaphysics* and of the so-called *Theology of Aristotle* (in fact a paraphrase of Plotinus) which are extant and available in print.

Al-Kindī's fame, however, was eclipsed by such later philosophers as al-Fārābī and Ibn-Sīnā (Avicenna). Only a few of his numerous treatises reached the Latin Schoolmen, but one recently discovered Arabic manuscript contains twenty-four of his otherwise unknown philosophical writings.

Two basic tenets of al-Kindī's, concerning prophecy and the creation of the world, were not accepted by his more famous Muslim successors. First, knowledge acquired through revelation in the Scriptures and from divinely inspired prophets is unambiguously superior to any knowledge acquired through philosophical training. In many cases, religious tradition and speculative, dialectical theology (repudiated emphatically by al-Fārābī) lead one to the same conclusions as philosophy and natural theology, which al-Kindī very consciously and proudly

introduced for the first time into the Muslim discussion. He maintained, however, that there are certain fundamental tenets of faith that are guaranteed by revelation alone and cannot be demonstrated by human reason.

Second, unlike the later Muslim philosophers, al-Kindī did not proclaim the eternity of the world and an eternal, emanating creation. Rather, he attempted to prove in philosophical terms that the world had been created from nothing, in time, through a divine creator, and that at some future date, according to divine dispensation, it would dissolve again into nothing. In doing this, he appears to use essentially the same arguments that were developed with more sophistication and subtlety by John Philoponus, the Christian Neoplatonic-Aristotelian philosopher, in sixth-century Alexandria. Al-Kindī also disagreed with the leading later thinkers by considering astrology to be a genuine branch of rational and methodical knowledge.

See also al-Fārābī; Aristotle; Avicenna; Islamic Philosophy; Philoponus, John; Plotinus.

Bibliography

WORKS BY AL-KINDĪ

An Arabic text is *Rasā'il al-Kindī al-falasafiyyah,* edited with an introduction by ʿAbd al-Hādī Abū Rīdah, 2 vols. (Cairo: n.p., 1950–1953), in which 24 scientific and philosophical texts are printed for the first time. An Arabic text with Italian translation is *Studi su Al-Kindī*: Vol. I was translated by M. Guidi and R. Walzer (Rome, 1940), and Vol. II was translated by H. Ritter and R. Walzer (Rome, 1938). An Arabic text with German translation is "Al-Kindi als Astrolog," translated by O. Loth, in *Morgenländische Forschungen fuer H. L. Fleischer* (Leipzig: Brockhaus, 1875), pp. 261ff. A Latin text with French translation is *Antécédents gréco-arabes de la psychologie,* a translation of *De Rerum Gradibus* by L. Gauthier (Beirut, 1939). A Latin text is found in *Die philosophischen Abhandlungen des Jaʿqūb ben Ishāg Al-Kindī,* edited by A. Nagy, which is Vol. II of C. Baeumker, ed., *Beiträge zur Geschichte der Philosophie des Mittelalters* (Münster: Aschendorff, 1897).

WORKS ON AL-KINDĪ

Works on al-Kindī are A. Altmann and S. M. Stern, *Isaac Israeli* (London: Oxford University Press, 1958), passim; F. Rosenthal, "Al-Kindī and Ptolemy," in *Studi orientalistici in onore di G. Levi della vida,* Vol. II (Rome: Istituto per l'Oriente, 1956), pp. 436ff.; and Richard Walzer, *Greek into Arabic* (Oxford: B. Cassirer, 1962), passim.

Richard Walzer (1967)

AL-KINDĪ, ABŪ-YŪSUF YAʿQŪB IBN ISḤĀQ [ADDENDUM]

Al-Kindī is important as the individual who established the earliest vocabulary for philosophy in the Islamic world. He was unusual in tending to avoid religious issues. In particular, in his ethics he tended to steer clear of specifically religious issues altogether. In this respect he followed a broadly Stoic line by advocating the life of the mind and the futility of relying on physical things to bring happiness. Virtue is attained by adhering to the middle ground and avoiding extremes. Toward the end of his life al-Kindī came under sustained attack by the local ruler. All in all, he did place philosophy in the Islamic world on a firm footing, and his influential disciples continued to debate and write along the lines their teacher had demonstrated.

See also Happiness; Islamic Philosophy; Stoicism; Virtue and Vice.

Bibliography

WORKS BY AL-KINDĪ

Al-Kindī's Metaphysics: A Translation of Yaʿqub ibn Isḥāq al-Kindī's Treatise "On First Philosophy" (fī al-Falsafah al-ūlā), edited and translated by Alfred L. Ivry. Albany: State University of New York Press, 1974.

"Uno scritto morale inedito di al-Kindi," edited and translated by H. Ritter and R. Walzer. *Memorie della Reale Accademia nazionale dei Lincei* (Rome), series VI, 8 (1) (1938): 47–62. Originally published as *Risalah fi al-hilah li-daf' al-ahzan* (On the art of averting sorrows).

Fī ḥudūd al-ashyā' wa-rusūmihā (On the definitions of things and their descriptions), edited by M. A. Abu Ridah. In *Rasā'il al-Kindī al-falsafiya*. Cairo, 1953. Translated by D. Gimaret in *Cinq épîtres*. Paris: Centre national de la recherche scientifique, 1976.

WORKS ON AL-KINDĪ

Klein-Franke, Felix. "Al-Kindi." In *History of Islamic Philosophy*, edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996. ch. 11, 165–177. I: 182–191

Oliver Leaman (2005)

AL-MUQAMMIṢ, DAVID BEN MERWAN

See *Muqammiṣ, David ben Merwan Al-*

ALSTON, WILLIAM P.

(1921–)

William P. Alston, an American philosopher, was born in Shreveport, Louisiana. He earned his Ph.D. from the University of Chicago (1951), and has taught at the University of Michigan (1949–1971), Rutgers University (1971–1976), the University of Illinois (1976–1980), and Syracuse University (1980–2000). Alston is a past president of the Central Division of the American Philosophical Association, the Society for Philosophy and Psychology, and the Society of Christian Philosophers as well as the founding editor of both *The Journal of Philosophical Research* and *Faith* and *Philosophy*. He is best known for his work in epistemology, the philosophy of religion, metaphysics, and the philosophy of language.

Alston made his early reputation in *Philosophy of Language* (1964), where he rejects the verifiability criterion of meaning and referential theories, and argues that the meaning of a sentence consists in its illocutionary act potential. He defends this view in his recent *Illocutionary Acts and Sentence Meaning* (2000), emphasizing the normative character of illocutionary acts. To illustrate, in uttering "Eat all of your vegetables," Trudy performs the illocutionary act of ordering the hearer to eat all of his vegetables only if she *takes responsibility* for the satisfaction of certain conditions, including: the hearer has some vegetables, it is possible for him to eat them, and Trudy has authority over him. So, Trudy performs the aforementioned illocutionary act only if she renders herself liable to censure in case these conditions are not satisfied—only if, Alston argues, she subjects her utterance to an illocutionary rule. Alston endorses a "Use Theory of Meaning," according to which a sentence's having a particular meaning consists in its being usable to play a particular role in communication. Because it is a sentence's illocutionary act potential that enables it to play this role, the meaning of a sentence consists in its usability to perform illocutionary acts of a particular type (in its being subject to a particular illocutionary rule).

Alston is also one of the leading proponents of realism about truth. In *A Realist Conception of Truth* (1996), he argues for alethic realism, the view that (1) truth is important and (2) a proposition is true if and only if what it claims to be the case *is* the case. Accordingly, the proposition that snow is white is true if and only if snow is white. Nothing else is necessary for the truth of that proposition. In opposition to epistemic conceptions of truth, a person need not be justified (rational, warranted) in believing that snow is white, nor must it be the case that she or he would be justified in believing it in ideal epistemic circumstances. Snow must simply be white. This is a minimalist—but not a deflationist—account of the *concept* of truth because the *property* of truth may have features that go beyond this concept. Consequently, Alston's realist conception of truth is consistent with the correspondence theory, but does not entail it. His conception of truth is also consistent with different types of metaphysical antirealism, including idealism and Hilary Putnam's conceptual relativism. In *A Sensible Metaphysical Realism* (2001), Alston defends his own version of *metaphysical* realism, according to which large and important stretches of reality do not depend on conceptual schemes for their existence.

Alston's early work in the philosophy of religion, much of which is collected in *Divine Nature and Human Language* (1989), focuses on the nature and properties of God, the literal application of predicates (e.g. "knowing") to God, and divine action. While Alston's views on philosophical theology are crucial contributions to the field, his most pioneering work is thought to be *Perceiving God* (1991), in which he develops a "doxastic practice" approach to the epistemology of religious experience. He argues that putative experiences of God can provide prima facie justification for beliefs about God. This is because mystical perception (MP), in which beliefs about a religiously construed ultimate reality are based directly on putative experiences of it, is a basic doxastic practice—a family of socially established belief-forming dispositions or mechanisms. MP (which includes Christian mystical perception [CMP], Hindu mystical perception, etc.) is analogous to sense perception—the basic practice of forming perceptual beliefs about the physical environment on the basis of sensory experience. Alston argues here, and in *The Reliability of Sense Perception* (1993), that any attempt to *show* that basic doxastic practices are reliable will be infected with epistemic circularity. Still, it is practically rational to suppose that CMP is reliable, and hence that the beliefs it generates are prima facie justified. It is also rational for practitioners of CMP, and practitioners of other forms of MP, to continue to engage in their respective practices.

Alston has had a striking impact on epistemology. His early work is devoted to defending fallibilist foundationalism, delineating and evaluating different concepts of epistemic justification, and advocating an account of justification that combines a core externalism with minimal accessibility to grounds. Rejecting perspectival internalism and higher-level requirements, Alston distinguishes between the activity of showing that a belief

is justified and a belief's being justified. In *Epistemic Justification* (1989), he argues that a belief's being prima facie justified consists in its being based on an adequate ground that is fairly readily accessible. The ground must *be* adequate—it must actually be sufficiently indicative of the truth of the belief. Because the subject need not have access to, or beliefs about, its adequacy, this is primarily an externalist, reliable-indicator account of justification. It anticipates the externalism of Alston's doxastic practice approach, according to which, for example, the socially established practice of sense perception must simply be reliable in order for a person's perceptual beliefs to be prima facie justified. In his recent work, he defends the Theory of Appearing as a superior alternative to sense-data and adverbial theories of the nature of perception. And, radically, in *Beyond "Justification": Dimensions of Epistemic Evaluation* (2005), he argues that there is no objective, epistemically crucial property of beliefs picked out by "justified." Consequently, epistemologists should dispense with the debate over justification, and instead investigate a plurality of epistemic desiderata, some of which are salvageable from it.

See also Epistemology; Metaphysics; Philosophy of Language; Philosophy of Religion.

Bibliography

PRIMARY WORKS

"Ontological Commitments." *Philosophical Studies* 9 (1958): 8–17.

Philosophy of Language. Englewood Cliffs, NJ: Prentice-Hall, 1964.

"The Deontological Conception of Epistemic Justification." In *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press, 1989.

Divine Nature and Human Language: Essays in Philosophical Theology. Ithaca, NY: Cornell University Press, 1989.

Perceiving God: The Epistemology of Religious Experience. Ithaca, NY: Cornell University Press, 1991.

The Reliability of Sense Perception. Ithaca, NY: Cornell University Press, 1993.

A Realist Conception of Truth. Ithaca, NY: Cornell University Press, 1996.

"Belief, Acceptance, and Religious Faith." In *Faith, Freedom, and Rationality: Philosophy of Religion Today*, edited by Jeff Jordan and Daniel Howard-Snyder. Lanham, MD: Rowman & Littlefield, 1996.

"Back to the Theory of Appearing." In *Epistemology: Philosophical Perspectives, 13*, edited by James Tomberlin. Oxford: Blackwell, 1999.

Illocutionary Acts and Sentence Meaning. Ithaca, NY: Cornell University Press, 2000.

A Sensible Metaphysical Realism. Milwaukee: Marquette University Press, 2001.

"Doing Epistemology without Justification." *Philosophical Topics* 29 (2001): 1–18.

Beyond "Justification": Dimensions of Epistemic Evaluation. Ithaca, NY: Cornell University Press, 2005.

SECONDARY WORKS

Battaly, Heather D., and Michael P. Lynch, eds. *Perspectives on the Philosophy of William P. Alston*. Lanham, MD: Rowman & Littlefield, 2005.

Senor, Thomas D., ed. *The Rationality of Belief and the Plurality of Faith: Essays in Honor of William P. Alston*. Ithaca, NY: Cornell University Press, 1995.

Heather D. Battaly (2005)

ALTERITY

The term *alterity* derives from the Latin word *alter*, which means "other." In contemporary philosophy the question of the other is primarily that of the other human being, the Other (*Autrui*, in French), although some thinkers have raised the question of whether the human other should be privileged in this way. However, the central question governing philosophical discussions of alterity is not that of who the other is, but that of our access to alterity. So-called continental philosophy highlights the ontological dimension of this question rather than its epistemological dimension, which was the focus in English-speaking philosophy of what, since the nineteenth century, has been called the problem of other minds.

In his *Cartesian Meditations* (1960 [1931]) Edmund Husserl offers an account of how, by an analogy with my own body, I recognize another body as organic and, by a kind of alienation in which I make myself other that we call empathy, constitute an other as an alter ego. Martin Heidegger in *Being and Time* (1996 [1927]) dismisses this approach as based on René Descartes's inadequate understanding of the human being as an isolated subject. Heidegger displaces the epistemological problem of alterity by issuing the ontological claim that the other possesses the kind of being that he calls *Mitsein* (literally "with-being"). Nevertheless, the problem of the other reappears in Jean-Paul Sartre's *Being and Nothingness* (1956 [1943]), where, in part under the impact of Georg Wilhelm Friedrich Hegel's account of the master-slave dialectic, the relation with the Other is presented as conflictual.

LEVINAS, DERRIDA, AND THE ABSOLUTE OTHER

In *Totality and Infinity* (1969 [1961]) Emmanuel Levinas radicalizes the problem of alterity by thinking of the other not as another subject like me, but as radically Other, the one who puts me in question and calls me to my responsibility. This ethical relation is asymmetrical in the sense the Other is accessible only starting from an I. However, the Other is no longer defined by his or her differences from me, but by the way he or she exceeds this relation in absolute separation from me. Thus, Levinas's conception of the absolute Other self-consciously breaks with the way that the other has been thought in the West since Plato's *Sophist*. According to Plato the other is always relative to some other (*Sophist* 255d), a formulation usually understood to mean that the other is "other than the same."

When Jacques Derrida challenges Levinas's account of the absolute Other in "Violence and Metaphysics" (1978 [1964]), he explicitly evokes Plato's critique that renders such a conception unthinkable, impossible, and unsayable (*Sophist* 238e). Without underwriting the legitimacy of Husserl's account of intersubjectivity Derrida asks whether Husserl's notion of an alter ego does not better secure the ethical character of the radical alterity of the other than does Levinas's notion of the absolutely other. Derrida's point is that the Other cannot be the Other of the Same except by being itself the same, that is, an ego, but he himself subsequently embraces Levinas's language of alterity with the phrase *tout autre est tout autre* (every other is wholly other) (1995 [1990]), p. 82).

Meanwhile, and in part in response to Derrida's essay, Levinas developed the fundamental idea of his later thought: the substitution of the one for the other. To the question of how it is possible for the Other to call me into question, Levinas, in *Otherwise Than Being* (1981 [1974]), gives the answer that it is possible because I am already for-the-other, that is to say, because the other is in me in the midst of my self-identification. A parallel gesture by which alterity is relocated within the same can be found in psychoanalytic literature, for example, in Julia Kristeva's *Strangers to Ourselves* (1991 [1988]). However, it can be argued that the new kind of cosmopolitanism she promotes retains the division between "them" and "us" and that it seeks to overcome, insofar as the world is now divided between those who recognize that there are no foreigners and those who do not.

To address the difficulty of thinking substitution, Levinas has recourse to Arthur Rimbaud's impossible phrase *je est un autre* (I is an other). Levinas uses the very difficulty of thinking and saying alterity not only to challenge the priority of ontology and proclaim the primacy of ethics but also to mark an exit from Western philosophy as he inherits it. This shows how far the question of alterity has departed from the Husserlian problem of intersubjectivity, as a regional problem, to become the philosophical site for explorations of the limits of thought and language.

See also Deleuze, Gilles; Derrida, Jacques; Levinas, Emmanuel.

Bibliography

Derrida, Jacques. *The Gift of Death* (1990). Translated by David Wills. Chicago: University of Chicago Press, 1995.

Derrida, Jacques. "Violence and Metaphysics." In *Writing and Difference* (1967). Translated by Alan Bass. Chicago: University of Chicago Press, 1978.

Heidegger, Martin. *Being and Time* (1927). Translated by Joan Stambaugh. Albany: SUNY Press, 1996.

Husserl, Edmund. *Cartesian Meditations: An Introduction to Phenomenology* (1931). Translated by Dorian Cairns. The Hague: Nijhoff, 1960.

Kristeva, Julia. *Strangers to Ourselves* (1988). Translated by Leon S. Roudiez. New York: Columbia University Press, 1991.

Levinas, Emmanuel. *Otherwise Than Being, or, Beyond Essence* (1974). Translated by Alphonso Lingis. The Hague: Nijhoff, 1981.

Levinas, Emmanuel. *Totality and Infinity* (1961). Translated by Alphonso Lingis. Pittsburgh, PA: Duquesne University Press, 1969.

Sartre, Jean-Paul. *Being and Nothingness: An Essay on Phenomenological Ontology* (1943). Translated by Hazel E. Barnes. New York: Philosophical Library, 1956.

Theunissen, Michel. *The Other: Studies in the Social Ontology of Husserl, Heidegger, Sartre, and Buber* (1977). Translated by Christopher Macann. Cambridge: MIT Press, 1984.

Robert Bernasconi (2005)

ALTHUSIUS, JOHANNES

(1557–1638)

Johannes Althusius, the German legal and political philosopher, was born at Diedenshausen, a village of the county of Wittgenstein-Berleburg in the Westphalian Circle. He is thought to have been the son of a farmer, although all data of his early youth are quite unknown. By 1581 he was studying Aristotle in Cologne, and he later studied Roman law at Basel. His experience of the Swiss way of life gave him a predilection for municipal freedom and self-government and for republican constitutionalism. Although deeply influenced by Calvinist piety, he

was eager to become a learned classical scholar. The forces of Christian faith, humanistic learning, and democratic feeling formed his character. He was both a man of strong will with a tendency to stubbornness and an austere moralist. It is, therefore, not surprising that he was a rigorous logical thinker and a systematic teacher as well as a realistic positivist with a desire to describe the empirical realities of social life.

Althusius passed his examination for the doctorate of civil and ecclesiastical law at Basel in 1586 with theses on the right of succession. In the same year he published a booklet, *Iurisprudentia Romana, vel Potius Iuris Romani Ars, 2 Libri, Comprehensa, et ad Leges Methodi Rameae Conformata* (Basel, 1586), that discussed fundamental questions of Roman law and that is also of philosophical interest. Through this work Althusius introduced into political science the systematic method of the French philosopher Petrus Ramus that contrasted with the prevailing humanistic method based on philological concerns. But although Ramus opposed the traditional Scholastic method of instruction, he had nevertheless retained the formalism of his predecessors insofar as he used the "method of dichotomy." This specific "ramistic" method divided every logical concept into two others, and each of them into two new concepts. This method of an endless, progressing, systematic presentation was applied by Althusius to all his later writings.

Soon after receiving his doctorate, Althusius became a lecturer in Roman law and in philosophy at Herborn, a newly established Calvinist college attended by students from many countries. In 1594 he became professor of law, and he was appointed rector of the college in 1597 and again in 1602. He also served as an advocate in the chancellery at Dillenburg. In this capacity he defended the rights of the college against the ambitions of the noblemen of the county. He was also involved in controversies with his colleague, the law professor Anton Matthäus (1564–1637), and with some of the Herborn theologians. In spite of these activities, he found time to write his most famous work, *Politica Methodice Digesta et Exemplis Sacris et Profanis Illustrata* (Politics methodically arranged and illustrated by holy and profane examples [Herborn, 1603; 2nd enlarged ed., Groningen, 1610; 3rd enlarged ed., Herborn, 1614]). This work was, as C. J. Friedrich wrote, "the culminating point of his life." The book clearly showed Althusius's systematic strength. He undertook to coordinate the diverse views of the Bible, Roman law, and the advocacy of the right to resist an unjust monarch of George Buchanan and the monarchomachs, and, on this basis, to write a compendium of political science.

The book was a natural and rational system of sociology, involving all the contemporary discussions of the problematical questions of theology, ethics, and jurisprudence. Althusius's fundamental view was that "politics is the science of linking human beings to each other for a social life." The whole of humankind, living in natural cooperative groups, builds up a universal community of civil and private corporations. The members join each corporation by the force of their sympathetic emotions. In this respect Althusius resembled both Hugo Grotius and Jean-Jacques Rousseau. However, he was a strong opponent of Jean Bodin's doctrine of royal absolutism, believing that the constituent power belongs to the community and that sovereignty is an attribute of the organized people, not of the king. The people decide all fundamental political questions through the representative assembly, and the chief of state is only a commissioner of the people and may be deposed if he acts contrary to the contract between him and the community. The representative assembly must obey the commandments of God and observe the natural laws. The necessities of human nature are as much a source of social order as is God's will.

Thus, Althusius held a threefold conception of social order: as a biopsychological social phenomenon, as a historically conditioned reality, and as a divinely limited work of man.

The principal sources of Althusius's thought were faith, reason, and experience. A major work composed somewhat later, *Dicaiologicae Libri Tres Totum et Universum Ius, Quo Utimur, Methodice Complectentes* (Digest of jurisprudence [Herborn, 1617]), is based on these three elements. In this work Althusius discussed the fundamental principles, institutions, and concepts of public and private law as they were found in the Roman jurisprudence of his day. By presenting the law as the realization of the concept of law and of its component legal categories, Althusius became one of the most important forerunners of modern Continental "legal conceptualism."

Meanwhile, in 1604 Althusius had been called as a syndic to Emden, a Calvinist city in eastern Frisia. He was soon appointed to the council, and he played an important part in the struggles of the city with the count of Frisia. He also became a dominant figure in the consistory of the Reformed Church in Emden.

See also Aristotle; Bodin, Jean; Grotius, Hugo; Political Philosophy, History of; Ramus, Peter; Rousseau, Jean-Jacques.

Bibliography

PRIMARY WORKS

"Politica Methodice Digesta" of Johannes Althusius (Althaus). Edited by Carl J. Friedrich. Cambridge, MA: Harvard University Press, 1932.

The Politics of Johannes Althusius. Abridged and translated by Frederick S. Carney. Boston: Beacon Press, 1964.

Politica. Edited and translated by Frederick S. Carney. Indianapolis, IN: Liberty Fund, 1995.

SECONDARY WORKS

Friedrich, Carl S. *Johannes Althusius und sein Werk im Rahmen der Entwicklung der Theorie von der Politik.* Berlin: Duncker und Humblot, 1975.

Gierke, Otto von. *The Development of Political Theory.* Translated by Bernard Freyd. New York: Norton, 1939.

Hueglin, Thomas O. *Early Modern Concepts for a Late Modern World: Althusius on Community and Federalism.* Waterloo, ON: Wilfrid Laurier University Press, 1999.

Wolf, Erik. *Grosse Rechtsdenker der deutsche Geistesgeschichte,* 4th ed. Tübingen: Mohr, 1963.

Erik Wolf (1967)
Bibliography updated by Philip Reed (2005)

ALTRUISM

While benevolence, compassion, and humanity were not major virtues for the ancient philosophers, modern moral philosophers generally agree that altruism is important to morality, although they disagree about what it is, how to explain it, and what its scope should be. The nineteenth-century French theorist Auguste Comte, who first coined the term *altruism*, claimed that the way to end social conflict is by training people to "live for others," rather than themselves. In a popular sense, altruism means something like noble self-sacrifice. A more minimal understanding, one that many philosophers favor, is an acknowledgment that the interests of others make claims on us and limit what we may do.

Altruism made its way into moral theory when Christian philosophers added the theological virtues of faith, hope, and charity to the cardinal virtues of the Greeks. Charity, the greatest of the theological virtues, was thought to be an inner spiritual orientation toward others. Charity is characterized as disinterested, universal, and unconditional. It should be directed to everyone, saint and sinner alike, regardless of merit.

The eighteenth-century Scottish philosopher Francis Hutcheson followed the Christian philosophers, claiming that everyone is capable of Christian love—calm universal benevolence—that aims at the good of all sentient creatures. He also distinguished two other types of benevolence: love directed toward smaller groups or particular persons, such as parental affection and friendship, and particular feelings of pity, sympathy, and gratitude. Christian love is the best form of benevolence; the other two are good so long as they do not counteract it.

Hutcheson's view about how altruistic we should be is even more radical than the Christian view. Reducing virtue to benevolence, he argues that none of the four cardinal virtues of the Greeks—temperance, courage, prudence, justice—are virtues unless their practice is motivated by love. Temperance is not a virtue, unless motivated by a concern to make ourselves fit to serve others. Courage is mere craziness, unless we face dangers in order to defend the innocent or to right wrongs. Prudence is not a virtue if it aims only at promoting our own interests. Justice is not a virtue unless it has a regard for the good of humankind. Hutcheson derives the utility principle—maximizing happiness for the greatest number—from the idea that the morally best motive is calm, universal benevolence.

Later utilitarians made the utility principle central to their account of moral rightness, but detached it from Hutcheson's basis in Christian love. Many utilitarians have argued that our duties of benevolence are extreme, so their view about the scope of benevolence is radical in another way. As long as I have the power to benefit others without hurting myself so much that total utility is reduced, I am obligated to help them. On this view, giving aid to famine relief, for example, is not a matter of charity but a duty.

There are two other ways of understanding altruism. One way, adopted by David Hume in the eighteenth century and by Bernard Williams as well as some feminist thinkers in the twentieth, characterizes altruism in terms of particular benevolent dispositions, desires, or affections. According to this view, you help others because you love them. Hume denied that we have the universal love of humankind to which Hutcheson and the Christian philosophers appealed, but thought that such benevolent dispositions as parental love and friendship were morally important character traits essential for virtue. Hume also thought that we possess the capacity to act from sympathy. When you see someone in distress, sympathy leads you to feel distress, which in turn motivates you to alleviate your distress by alleviating theirs. Sympathy enables

us to extend our love for particular individuals and smaller groups to larger groups of people.

Williams's view is similar to Hume's. Some of our particular benevolent desires are directed toward people we care about, for example, a daughter or friend, and are motivated by thoughts like "Mary needs help." Other benevolent desires are more general and impersonal concerns, motivated by thoughts like "someone needs help." Williams claims that the structure of the motivating thought in both cases is the same. Although altruism is not a rational requirement on action, Williams thinks that sympathetic reflection may move us from benevolent desires motivated by our love of particular individuals to more general altruistic dispositions.

Some feminist philosophers have argued that altruistic dispositions such as caring, compassion, and maternal love should be made the focus of morality. These philosophers claim that relationships should be at the heart of morality and that most of our relationships are not only intimate, but also involuntary. They argue that an ethics of care rather than an ethics of justice is appropriate for these types of relationships.

By contrast, philosophers in the Kantian tradition conceive of altruism as a rational requirement on action. They claim there is no need to postulate a benevolent desire to explain altruism. Kant's initial argument appeals to his requirement that we may only act on principles that we can will as universal laws. Willing a world in which everyone has a policy of not helping others, while knowing that you will need help, would be inconsistent, so we must will to help those who are in need. Kant also argues for a duty of beneficence on the basis of the requirement of treating humanity as an end in itself. He argues that you must treat the ends of others as you treat your own ends. You take your own ends to be good and worth pursuing, so consistency requires that you treat the ends of others as good and worth pursuing. This suggests that we have reason to help not only those in need, but anyone we are in a position to help.

Thomas Nagel follows Kant in thinking that the reasons of others directly provide us with reasons. Suppose someone wants you to stop tormenting him. How does that person's desire not to be treated that way give you a reason to stop? At an intuitive level, Nagel's argument appeals to the question: How would you like it if someone did that to you? You realize that if someone were tormenting you, you would not merely dislike what he was doing, you would resent it. Resentment is a response to the idea that someone has ignored a reason he has to not treat you badly. The reason in this case is your own desire not to be tormented. You think your desire not to be tormented is a reason for your tormentor to stop. Since you think that your reasons provide direct reasons for others, you must also think that the reasons of others provide you with reasons. The argument turns on the idea that your reasons and the reasons of your victim are the same: they are the reasons of a person. According to Nagel, the argument works only because you have the capacity to view yourself as just one person among others. Although Humeans and Kantians disagree about whether to explain altruism in terms of particular desires or to view it as a rational requirement on action, they agree that the force of altruism springs from our common humanity.

See also Egoism and Altruism; Ethical Egoism; Friendship; Human Nature; Love; Sympathy and Empathy; Virtue and Vice.

Bibliography

Hume, David. *Enquiry concerning the Principles of Morals* (1751), edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press, 1975.

Hume, David. *A Treatise of Human Nature* (1739–1740). 3 vols., edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press, 1985.

Hutcheson, Francis. *An Essay on the Nature and Conduct of Passions and Affections with Illustrations upon the Moral Sense* (1728), edited by Aaron Garrett. Indianapolis, IN: Liberty Fund, 2002.

Kant, Immanuel. *Grounding for the Metaphysics of Morals*. Translated by James W. Ellington. Indianapolis, IN: Hackett, 1993. Originally published as *Grundlegung zur Metaphysik der Sitten* (Riga: J. F. Hartknoch, 1875).

Nagel, Thomas. *The Possibility of Altruism*. Oxford: Clarendon Press, 1970.

Williams, Bernard. *Problems of the Self*. Cambridge, U.K.: Cambridge University Press, 1973.

Charlotte R. Brown (2005)

AMPÈRE, ANDRÉ MARIE

(1775–1836)

André Marie Ampère was a French physicist and philosopher; his main achievement in physics was the foundation of electrodynamics. He correctly recognized that Hans Christian Ørsted's discovery, in 1819, of the effect of electric current on a magnetic needle was merely a special case of the general correlation of electricity in motion with the rise of a magnetic field. His explanation of magnetism in terms of molecular electric currents was a bold anticipation of one feature of the later electron theory.

Shortly after Ampère's death his *Essai sur la philosophie des sciences* appeared with a biographical note by Charles-Augustin Sainte-Beuve and a warm appraisal by Émile Littré. Its subtitle, *Exposition analytique de toutes les connaissances humaines,* indicated that the main topic was the classification of sciences, in which Ampère was as much interested as his contemporary Auguste Comte. Ampère's main division of sciences into "cosmological" and "noological" was inspired by Cartesian dualism. The details of the classification, which also included "applied sciences"—medicine, agriculture, etc.—are now of only historical interest.

Far more interesting is *La philosophie des deux Ampères,* edited by J. Barthelémy Saint-Hilaire. The title is misleading because the only contribution of Ampère's son Jean Jacques is an introduction to the philosophy of his father. Besides this, the book contains some unfinished philosophical manuscripts as well as Ampère's letters to Maine de Biran, with whom he remained in personal contact and in correspondence until Maine de Biran's death in 1824. Ampère accepted the central idea of Maine de Biran's voluntaristic idealism that the true nature of the self is revealed in the introspective experience of effort. But unlike Maine de Biran, Ampère more cautiously differentiated what he called *emesthèse* (that is, consciousness of personal activity) from the sensation of muscular effort that can be induced by some external agency.

This was not the only instance of Ampère's remarkable gift for introspective analysis. In dealing with the association of ideas he distinguished two cases. The first is *commémoration,* or ordinary recall, when two associated ideas remain unaffected by their contiguity. The second is *concrétion,* when two ideas merge, for example, when the present perception of an object seen before blends with the recollection of its previous perception. But the main difference between Ampère and Maine de Biran concerned the problem of knowledge of the external world. Maine de Biran, under the influence of Immanuel Kant, denied any possibility of knowing things-in-themselves; Ampère, under the influence of Isaac Newton, John Locke, and his own scientific interests, believed in the possibility of knowing inferentially the relations between things-in-themselves. These "noumenal relations" are similar to Locke's primary qualities; they can be known when the general spatial, temporal, and numerical relations are divorced from the qualitative content (Locke's secondary qualities) of sensory experience. But unlike Locke, Ampère interpreted the impenetrability of matter dynamically, as being a result of inextensive resistances (*résistances inétendues*) of which there is an indefinite number in each body. This view of matter as being a product of inextensive dynamic centers is thus closer to the dynamism of Gottfried Wilhelm Leibniz, Roger Joseph Boscovich, and Michael Faraday than to the traditional atomism of Newton. On the other hand, Ampère remained a Newtonian in his insistence on the reality of absolute space and time, which he interpreted theologically, again like Newton, as attributes of God. Equally Newtonian was his rejection of the Cartesian plenum.

See also Boscovich, Roger Joseph; Cartesianism; Comte, Auguste; Faraday, Michael; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Littré, Émile; Locke, John; Maine de Biran; Newton, Isaac.

Bibliography

WORKS BY AMPÈRE

Théorie mathématique des phénomènes électrodynamiques, uniquement déduite de l'expérience. Paris, 1827. Ampère's account of electrodynamics.

Essai sur la philosophie des sciences, 2 vols. Paris: Chez Bachelier, 1834–1843.

La philosophie des deux Ampères. Edited by J. Barthelémy Saint-Hilaire. Paris: Didier, 1866; 2nd ed., 1870.

WORKS ON AMPÈRE

Broglie, Louis de. *Continu et discontinu en physique moderne.* Paris: A. Michel, 1941. Pp. 241–266.

Cantor, Georg. "Über verschiedene Theoreme aus der Theorie der Punktmengen." *Acta Mathematica* 7 (1885): 105–124.

"Lettres de Maine de Biran à A.-M. Ampère." *Revue de metaphysique et de morale* 1 (1893): especially 553.

Lorentz, Borislav. *Die Philosophie André-Marie Ampère.* Berlin, 1908. Inaugural dissertation.

Milič Čapek (1967)

ANALOGY IN THEOLOGY

Religious discourse has been under scrutiny since ancient Greece when Anaxagoras said if oxen and dogs could paint, they would depict the gods in their own likenesses. The Jewish, Christian, and Muslim scriptures depict the divine being in vivid humanlike traits while conveying the divine otherness, mystery, immateriality, and eternity. Thus there are religious currents of anthropomorphism, of transcendentalism, of metaphor and symbolism, and of literalism about the being and nature of God. The Greek philosophical ancestry of Western culture presents the divine as immaterial, immutable, everlasting, perfect, and incomprehensible. Both the Platonic and Aristotelian

metaphysicians developed theories of analogical predication that were later extended to theology, the study of the revealed divinity.

Theologians used a theory of analogy that had three parts: analogy of *being* (of reality between God and world, and among created realities, too); analogy of *meaning* (of words and concepts); and analogical thinking (of conception by proportionalities). The aim was to explain how words that apply to sensible things also adapt in meaning to apply literally, not only metaphysically, to the transcendent deity known only by inference, revelation, or mystical experience. Words applied to God—"wise" and "good," for example—are neither entirely equivocal (such as *bank*/savings; *bank*/river), nor merely metaphorical (*drop*/an argument), but rather, they are analogous; that is, they adjust in ways explained below to the context, just as words generally adjust to contrasting contexts, say, as "knows"/the way differs from "knows"/arithmetic, and as "exist" does in "*there exist* /trees/species/numbers/shapes." Metaphysics articulates *theoretical* truth-conditions for such statements and for ordinary religious beliefs—conditions not accessible without such metaphysics—the way science states the molecular structure for water.

1. SECULAR ORIGIN IN PLATO AND ARISTOTLE

The thesis that words fit in literal meaning to diverse verbal contexts that reflect differences of reality—the analogy theory—has its origin in secular philosophy. For Plato, things that share in the Forms are not said to exist in the same sense as the Forms (compare *Sophist*; *Parmenides*), and the Form "Human" is *what*-it-is-to-be-human, and thus is human, but not in the sense in which Callicles is human by participating in the Form. Further, Plato used the same names, such as the *courageous* man and the *courageous* act, *just*/state; *just*/man, for things related as cause to effect and sign to signified.

Aristotle used those distinctions, added more, and regarded *real*, entitative analogy, reflected in word-meaning, as central to his explanatory principles. (*Metaphysics* 1070a.31). Such predication is *literal*, as opposed to metaphorical (*Poet* 1457b)—for example, "the fields *smile* with the sunlight" (Aquinas called that *improper* proportionality [*Summa Theologiae* 1.13.3.ad 3]). Aristotle acknowledged analogy by *attribution* (relational naming: *healthy*/animal; *healthy*/diet), and by *proper proportionality* (e.g., genus is to species as body is to soul, namely, as *potency* is to *act*). The explanatory terms—for example, "act/potency"—apply to diverse things analogically (*Met* 1048b, 5–8). Aristotle further reasoned that qualities, such as color and shape, and other accidentals, are said *to be* derivatively (*pros hen*) to substances; and "analogically the same things are principles, i.e. actuality and potentiality; but they are not only different for different things, but also apply in different ways to them" (*Met* 1071a.5). Aristotle says the causes and principles of different things are analogous and are spoken of analogously (*Met* 1070a.31). Moreover, the contrast-dependent notions, "act/potency," "matter/form," "substance/accident," "cause/effect," are all *analogical* in meaning because the phenomena to which they apply are really, *de re*, analogous; for instance, body is matter for soul, and clay is matter for a statue.

2. TRANSITION TO THEOLOGY

The Arabic philosophers adopted Aristotle's views on analogy in their metaphysics and physics and in their discussion of the simplicity of God in the Qurʾan. That made the first connections of Aristotelian analogy-theory to scriptural theology. Islamic religious believers differed on how to interpret the physical descriptions of God's face, eyes, hands, speaking, sitting, and so on, in the Qurʾan, as well as the description of God's feelings—for example, wrath, satisfaction, and God's traits, such as cunning and patience—whether anthropomorphically, metaphorically, symbolically, and so on (compare Van Ess 1954). Al-Kindī (c. 850) thought a literal reading of the Qurʾan on creation is coherent with Aristotelian concepts. In his treatise "On the One True Agent" he holds God is literally the only agent bringing being from (absolute) nonbeing, whereas humans are only metaphorically (analogically?) agents, bringing being from potentiality. Al-Fārābī (c. 900) in chapter 1 of "On the Perfect State" says "existing," "having intellect," "knowing," "being wise," "real," "true," "living," and the like, are said of God in senses different from what we say of creatures because the divine being is simple, without composition or distinct traits. And Avicenna (980–1037) used Aristotle's views about analogy of meaning and of reality directly in his metaphysics and in his physics, where "motion," for instance, is said (as Aristotle also said) to apply analogously, to augmentation, alteration, and locomotion, and the analogy of "being" within the ten categories is acknowledged.

Avicenna reasoned that being and essence are really the same in God, and indicated that a creature's being is not explained by "*what*-it-is" as is the divine; Aquinas would adopt this. Avicenna also formulated the principle that God's knowledge is the cause of things (later used by

Aquinas as *cognito dei causa rerum*, *ST* 1.14.8), whereas our knowledge is posterior to things known. The Arabic writers, including the Jewish Moses Maimonides, all hold that God is simple; he is not a body, without any plurality of attributes except by attribution from the divine effects, infinite and incomprehensible. It is from those Arabic, chiefly Islamic, sources, along with the corpus of Aristotle, that the analogy theory came into Latin theology, Avicenna being the most influential in metaphysics.

Maimonides (1135–1204) argued that the eternity of the world is not demonstrated, and that it is both created and has a temporal beginning. Like Avicenna, he affirmed the divine simplicity in a strong sense, so that: "either every attribute we predicate of Him is an attribute of action [and so named from the received effect], or, if the attribute is intended for the apprehension of His essence and not of his action, it signifies the negation of the privation of the attribute in question" (*Guide for the Perplexed*, 1, p. 58). Thus, saying, "God is all knowing" means "God is not unknowing of anything," and saying "God is simple" means "God is not composite," and saying "God is eternal" means "God is without beginning or end." That came to be known as "negative theology," with no positive ascriptions to God, except existence and creation and the metaphors provided by scripture.

Christians, from the earliest fathers of the church, developed explanatory analogies—that is, proportional comparisons, say, of the Trinity to the unity amidst distinction of the essence, power, and operation of the human soul, and an analogy of the relation of the Father to the Son as "light from light" (in *Nicene Creed*, and Augustine's *De Trinitate*). Such explanatory analogies, not part of the theory described here, were devised throughout the predominantly neoplatonic first millennium of Christian thinking, for instance in Augustine's *De Trinitiate* (c. 410) and Boethius's *De Trinitate* (c. 510), the School of Chartres (twelfth century), and the School of St. Victor (twelfth century), and continued throughout the later history of theology (compare Chollet 1923–1967).

A neoplatonic writer historians call Pseudo-Dionysius (c. 500) was widely believed, but not by Aquinas, to have authority as a disciple of St. Paul. He proposed, in his *Divine Names*, that one first knows God by negation (*via negationis*), "not a body," "not with parts," and so on, then by inadequate affirmation as "wise," "good" "loving," qualified by "but not in the way of creatures," and then in a third stage by superlatives, such as "infinitely knowing" and "good beyond excellence" (*via eminentiae*). But in his *Mystical Theology* Pseudo-Dionysius is more restrictive, saying one starts *via remotionis* by denying of God the things most remote from him, such as "drunkenness and fury," then progressing by denial even through all the higher attributes of creatures until one reaches "the super-essential darkness," entering "the cloud of unknowing," mystically united to what is "wholly unknowable" (because of the limitations of the human mind). This work had a profound influence on the development of transcendentalism in medieval theology and even into the twenty-first century.

3. AQUINAS (1225–1274)

Aquinas combined the influences of Avicenna, Maimonides, and Pseudo-Dionysius, along with mastery of Aristotle and Plato. He held that God infinitely transcends every true description achieved by human philosophical efforts, but that, nevertheless, a great deal can be known and positively established about God; in fact, Aquinas believed, there can be both a philosophical science of God from unaided reason, and a *divine science* whose first principles are given by revelation (*ST*, 1.q.1.a.2). Furthermore, he absorbed Aristotle's notion of analogy of "being" (*pros hen*) for the ten Categories into his own wider theory of analogy between creatures and God by participation. Aquinas said "being can be essentially predicated of God alone, because to be divine is to be subsistent and absolute, whereas being is predicated of any creature by participation; for no creature *is* its own being, but is something having being," as the actuality (*esse*) of its potentiality (its essence), because creatures do not exist on account of *what* they are, but on account of God (*Quod*.2, q.2, 1.1). Further, *what* God is, essentially, is not naturally knowable to humans, though it is disclosed to the blessed by divine gift (*ST* 1.12.1).

Thus, Aquinas reasoned that our knowledge is not limited to what we can attribute negatively or only by metaphor, or merely by the *extrinsic* attribution that would make "God is good" mean merely "God is the *cause* of creaturely goodness" (*ST* 13. a.6) in the way that a person is called "captain" because of what he does. Many writers, influenced by Philo Judaeus (c. 20 BCE–40 CE) whose work came to the West through Clement of Alexandria and Origen, held that God is named only with names of his effects. Aquinas, however, says we can know that pure perfections (unmixed with limits, such as "educated") apply intrinsically to God by *explanatory priority* because the divine perfections are the cause and exemplar of all perfections in creatures, such as being, life, knowledge, freedom, and love. This position is variously developed in *Summa Theologiae* (q.13, a.4–5), and *Summa*

Contra Gentiles (1, chap. 34), and *Q. D. De Veritate* (q 2.a.1). Nevertheless, the names and concepts of pure perfections are acquired only through our perceptual experience with creatures (*ST* 1.q.13, a.6), even though their primary reality is in God. Thus the words "loves," "knows," "chooses," and so forth, used of God and of humans, have similar definitions but differing presuppositions that reflect the diverse manner of being of God and creatures, the perfections being prior and all the same as God's being, and finite, received, and really separable from one another in creatures.

So whatever is predicated positively of God is either by *attribution*, as God is called "creator" on account of what is made and "happy" because of his perfect enjoyment, or predicated by *proportionality* and priority, as God is said to be "knowing, loving, wise, excellent and beautiful," and so on, but in a manner explanatorily prior to the creature's imperfect and derived being and knowledge. Aquinas also acknowledged metaphorical predicates of God, too (*ST* 1.19.11), many sanctioned by scripture ("angry," "Prince of Peace"), and many useful negative ones ("not a body," "not in space," "not with parts or complexity," "not with a beginning or end").

The religiously and philosophically central attributes are predicated literally and intrinsically, with their presuppositions adjusted to religious discourse (e.g., "God chooses" but does not deliberate), and elaborated theoretically (e.g., God's attributes are all "really the same as the divine being, *esse*, differing from one another only in concepts"). They include "knowing," "loving," "good," "righteous," "just," "omnipotent," "omniscient," "immutable," and "present everywhere"—and every other unmixed perfection, too. They apply to God but are adjusted to the priority and perfection of divine being. Thus, God knows but does not find out; God loves but does not need. All creation participates in God's being, not as being divine in any way, but as being continuously *from and on account of* God, and thus, being said "to be" analogously. Created being is God's proper and continuous effect; the way setting-alight—igniting—is the *proper* effect of fire; and the illumination of the air is the *continuous* effect of the sun (*ST* 1, 8.1).

Aquinas thought the real analogy between divine subsistent being (*ipsum esse subsistens*) and creaturely, participated being is an adequate basis for demonstrative knowledge of the existence and of the many attributes of God by reasoning that he displayed in *Summa Contra Gentiles*.

Nevertheless, Aquinas emphasizes that because what is received is received in the manner of the recipient (*quidquid recipitur recipitur modo recipientis*, *ST* 1.75.5), God is disclosed through nature only as far as nature is capable, with all creatures falling infinitely short of the divine reality. And he holds that the divine biblical revelation, though vastly exceeding anything humans could discover or even conceive on their own, is proportioned to what is fitting for humankind, thus leaving the infinite divine mystery "wrapped in a mist" (*caligine abvoluta*, *Const. Dei Filius* ch. 4, Vat. 1), with the essence of God beyond all natural understanding.

By the Reformation in the sixteenth century, a religious role for scholastic philosophy was largely rejected, and the reformers held the faith to be in no need of fragile and contested support from philosophy. Biblical authority was said to stand on its own, to be understood by the "analogy of Faith" (*analogia fidei*, based in Rom. 12.6, according to both Luther and Calvin). Thus the analogy discussions dried up, except among Catholic philosophers such as Cardinal Cajetan (1458–1564), Sylvester of Ferrara (1474–1528), and Francisco Suárez (1548–1617), and mostly stayed that way, apart from the historical scholarship that continues to the present.

David Hume (1711–1776) inaugurated modern noncognitivism, consigning metaphysics to the flames (*Enquiry*, 1748), asserting that all truths are grounded in sense impressions or relations of ideas, thus setting the framework for twentieth-century verificationism and the attack on the cognitive content of religious discourse.

4. CONTEMPORARY CONTEXT

In the twentieth century, positivist philosophers, seeking to be like scientists, questioned whether talk about God had any cognitive content at all. Alfred Jules Ayer argued that talk about God is without content because it is unverifiable. Some believers, such as Richard Bevan Braithwaite and Frank Plumpton, proposed empirical understandings of its content; others, such as John Hick, even propose eschatological verifiability. Philosophers such as D. Z. Phillips argued that religious discourse belongs to its own "language game"—a notion adopted from Ludwig Wittgenstein—with its own conditions for meaningfulness, and its own conditions of rational belief, analogous to mathematics and aesthetics. Mostly, however, the discussion of meaningfulness was unconnected to the historical positions on analogy in metaphysics and theology.

One twentieth-century adaptation of the classical accounts (Ross 1981) reasoned that analogy, as "fit of word-meaning to contrasting contexts," is a universal feature of natural languages within which the Aristotelian

cases of relational and presuppositional adaptation are particular species, and that the cognitive content of utterances is a function of the family of statements and practices in which they are employed (and often craft-bound to specialized skills and tasks, such as medicine or sailing). Thus, analogy of meaning in religious contexts is a special case of the analogy phenomena found in all the neighborhoods of discourse, whether specialized or not. And Aquinas's metaphysical theory, say of participation and *esse subsistens*, was interpreted, on that account, as his articulating *theoretical* truth-conditions for the ordinary and analogous talk of divine existence, perfection, and action, the way a chemist might explain the atomic constitution of a commonly known metal such as lead.

Thus there are at least two additions to Aristotle's and Aquinas's work on analogy: first, that the linguistic phenomena involve differences of *discourse* commitment (e.g., "God decides," but does not deliberate), as well as the differing *theoretical* presuppositions articulated by metaphysicians, such as "all divine perfections are really *de re* the same"; and second, that analogous fit of meanings to diverse context is lawlike, universal, and dynamic in natural languages. But *lexical* meanings of words are not to be regarded as direct pairings of words to concepts (considered to be their meanings), but are relations of contrast-dependence within the language itself (compare Saussure 1915)—that is, relations of contrastive expressive *capacity*, so that *meanings* and the world are correlated in clouds or clusters of discourse, not simply item by item.

As Wittgenstein, Wilfrid Sellars, and others observed, the cognitive content of verbalized beliefs is a function of the community of social behavior in which they have a place in the giving of explanations, reasons, motivations for actions, and interpretations. Thus, although a lot of nonsense is easily formulated in religious talk—as in any other talk—expressed convictions that modify action and attitudes either reflect reality or fail to, and either do so poorly or well. They are thus suitable for epistemic attitudes such as belief and denial. Nevertheless, the truth or falsity of what is said by the religious may not be accessible from outside the practicing community, just as the truth of medical, musical, manufacturing, or scientific expert discourse is largely inaccessible from outside the community of expertise.

The late twentieth- and early twenty-first-century cognitivity issue for religion involved three challenges: (i) whether characteristic expressions (say "Jesus is my personal savior"; "There is one God in three Persons") have stable conditions of appropriate utterance, qualification, reasons, rejoinder, and so on, within a practicing (relatively narrow) religious community; (ii) whether the community practice is one of coherent stable conditions, positive or negative, for acceptable use and endorsement and reason-giving for such assertions; and, (iii) whether basic claims, say, about the existence and nature of God, or some of them (*praeambula fidei*), can be rationally accepted or rejected, as well, from outside the confessing community. The common core of Judeo-Christian-Islamic monotheism meets the challenges affirmatively, and many find it externally well supported, even demonstrated in part, though other competent assessors disagree.

Some participants, such as the Reformers, thought external assessment carries no weight or utility for religious faith, though it may have some value in defense of the faith (apologetics). Note also that, in general, the false may sometimes be rationally accepted and the true rationally rejected, as the history of medicine and physics illustrates. Nothing requires a body of convictions to be decidable entirely, or even at its heart, from outside the practice in which conviction is arrived at and sustained. Otherwise the fabric of science would be subject to nonscientific rejection, rather than just parts of it. The same holds for religion. Still, Augustine and Aquinas held that the scripture cannot mean literally what science has demonstrated to be false (*ST* 1.68.1).

Some recent writers talk as if words, including temporal ones, apply to God not only literally but univocally; for instance, Richard Swinburne, in *The Coherence of Theism*(1977) said he applies "good" to God in the "perfectly ordinary" sense in which he would say his grandmother was good, though the conditions differ (p. 71). That contrasts with those philosophers such as D. Cupitt or Bishop Robinson who regard talk about God as merely metaphorical. Perhaps, like Duns Scotus, Swinburne and others consider the meaning of the words to be unaffected by differences in the mode of a thing's being. Charles Hartshorne, a Whiteheadean "process metaphysician," came closer to anthropomorphic literalism when he said that God, in process of self-surpassing, can suffer, change, and have other temporal predicates. Analogy theory is often mistaken for a theory of *non*literal predication, when it is just the opposite: an account of the literal but not anthropomorphic.

Some theologians such as Karl Barth say the meanings of "God loves," "forgives," "redeems," and "commands" are determined by the *scriptural* context as understood by the church (the community of believers): "Language about God has the proper content, when it

conforms to the essence of the Church, i.e., to Jesus Christ.… according to the analogy of faith, (Rom 12.6)" (*Church Dogmatics*). In accord with Luther and Calvin, he probably meant that nothing *more* than the *analogia fidei*, as understood by the Church, determines what a faithful Christian is to believe and mean. But to say there can be *no further* truth-conditions at all, say, for "Jesus is the Son of God," would conflict with simple logic. So, sciences might investigate such conditions. And whether extrascriptural theoretical content is sometimes required for faithful belief (say, Eucharistic consubstantiation vs. transubstantiation, vs. mystical presence) is a matter not settled by *sola scriptura* and *analogia fidei*, unless theological inquiry is included.

Thus the analogy theorists, historically and in the twenty-first century—like the Reformers—and Barth, the Evangelicals, and philosophers such as Swinburne and Alvin Plantinga, hold that talk about God is neither empty of intelligible content (noncognitivism), nor only metaphorical, poetic, or symbolic (Paul Tillich); nor only negative, except for God's existence (Maimonides); nor positive only in superlatives (Pseudo-Denis—*via eminentiae*). And they reject the principle that what is not observationally verifiable or falsifiable is meaningless. They agree that the scripture is the norm for what is to be said about God as Revealed. But analogy theorists additionally maintain (i) that analogous predication is literal and perfectly common in discourse generally, and characteristic of discourse about God, and (ii) that the metaphysical exploration of the divine, even of what is revealed, discloses theoretical truth-conditions, not otherwise accessible, for claims that God exists and has the divine perfections, just as science discloses microconditions for water that are not contained on the surface of the ordinary vocabulary.

So it seems that analogy theory both as linguistic theory and as metaphysical account of being has more innings to play in the history of theology.

See also al-Fārābī; Anaxagoras of Clazomenae; Aristotelianism; Aristotle; Augustine, St.; Avicenna; Ayer, Alfred Jules; Barth, Karl; Boethius, Anicius Manlius Severinus; Braithwaite, Richard Bevan; Cajetan, Cardinal; Calvin, John; Clement of Alexandria; Creation and Conservation, Religious Doctrine of; Duns Scotus, John; Hume, David; Infinity in Theology and Metaphysics; Luther, Martin; Maimonides; Origen; Philo Judaeus; Philosophy of Religion, History of; Plantinga, Alvin; Plato; Platonism and the Platonic Tradition; Pseudo-Dionysius; Ramsey, Frank Plumpton; Reformation; Saint Victor, School of; Sellars, Wilfrid; Suárez, Francisco; Sylvester of Ferrara, Francis; Thomas Aquinas, St.; Tillich, Paul; Wittgenstein, Ludwig Josef Johann.

Bibliography

Aquinas, Thomas. *Summa Theologiae*. Vol. 1, edited by Thomas Gilby. Garden City, NY: Image Books, 1969.

Aquinas, Thomas. *Summa Contra Gentiles*. Translated by A. Pegis. Notre Dame, IN: University of Notre Dame Press, 1975. Both contain easily accessible and brief accounts of ideas developed in many other places of his work. See Wippel, below, for a comprehensive exposition of Aquinas.

Barth, Karl. *Church Dogmatics*. Vol. 2, edited by G. W. Bromiley and T. F. Torrance. Translated by G. W. Bromiley. Edinburgh: T. & T. Clark, 1957.

Cajetan, Thomas De-Vio. *The Analogy of Names, and The Concept of Being*. 1498. Translated by E. Bushinski and H. Koren. Pittsburgh: Duquesne University Press, 1953. An influential, brief, systematization and interpretation of Aquinas's positions, much disputed by later scholars, but still useful and the source of some classifications such as "analogy of proper proportionality."

Chollet, A. "Analogie." In *Dictionnaire de Theologie Catholique*, edited by A. Vacant et al. Paris: Librairie Letouzey, 1923–1967. A scholarly survey of analogy in Roman Catholic natural and dogmatic theology.

Klubertanz, George. *St. Thomas Aquinas on Analogy: A Textual Analysis and Systematic Synthesis*. Chicago: Loyola University Press, 1960. With an appendix of the passages (in Latin) in which Aquinas discussed analogy.

Lyttkens, H. *The Analogy Between God and the World: An Investigation of its Background and Interpretation of its Use by Thomas of Aquino*. Uppsala, Sweden: Lundequistska bokhandeln, 1953. A published doctoral dissertation with historical sweep and a comprehensive study of the primary sources.

McInerny, Ralph. *Aquinas and Analogy*. Washington, DC: Catholic University of America Press, 1996.

Ross, James F. *Portraying Analogy*. Cambridge, U.K.: Cambridge University Press, 1981. Expands Aristotle and Aquinas into a new account of analogy of *meaning*, its role in philosophy, and in the debate about the cognitive content of religious discourse. It does not address the *analogia entis*, the metaphysics (see Wippel).

Saussure, Ferdinand de. *Cour de linguistique generale*. Paris, 1915. Translated by W. Baskin in *Course in General Linguistics*. New York: Philosophical Library, 1959. By the "father" of modern linguistics, in which the "linguistic meaning is contrast of meaning," idea is developed along with the notion of paradigmatic contrasts that map meaning relationships (employed in Ross, 1981).

Van Ess, J. "Tashbih wa-Tanzih." In *Encyclopaedia of Islam*, edited by H. A. R. Gibb et al., vol. 10, 341–344. Leiden, Netherlands: Brill, 1960. Reference courtesy of Dr. Jon McGinnis.

Wippel, John, F. *The Metaphysical Thought of Thomas Aquinas: From Finite Being to Uncreated Being*. Washington, DC: Catholic University of America Press, 2000. The most comprehensive and up-to-date exposition of the whole of

Aquinas's theory, its development, and its rationales, especially in chapters 3 and 13.

James F. Ross (2005)

ANALYSIS, PHILOSOPHICAL

Philosophical analysis is a term of art. At different times in the twentieth century, different authors have used it to mean different things. What is to be analyzed (e.g., words and sentences versus concepts and propositions), what counts as a successful analysis, and what philosophical fruits come from analysis are questions that have been vigorously debated since the dawn of analysis as a self-conscious philosophical approach. Often, different views of analysis have been linked to different views of the nature of philosophy, the sources of philosophical knowledge, the role of language in thought, the relationship between language and the world, and the nature of meaning—as well as to more focused questions about necessary and apriori truth. Indeed the variety of positions is so great as to make any attempt to extract a common denominator from the multiplicity of views sterile and not illuminating.

Nevertheless analytic philosophy—with its emphasis on what is called "philosophical analysis"—is a clear and recognizable tradition. Although the common core of doctrine uniting its practitioners scarcely exceeds the platitudinous, a pattern of historical influence is not hard to discern. The tradition begins with G. E. Moore, Bertrand Russell, and Ludwig Wittgenstein (as well as Gottlob Frege, whose initial influence was largely filtered through Russell and Wittgenstein). These philosophers set the agenda, first, for logical positivists such as Rudolf Carnap, Carl Hempel, and A. J. Ayer and then later for Wittgenstein, who in turn ushered in the ordinary language school led by Gilbert Ryle and J. L. Austin. More recently the second half of the twentieth century has seen a revival of Russellian and Carnapian themes in the work of W. V. Quine, Donald Davidson, and Saul Kripke. Analytic philosophy, with its changing views of philosophical analysis, is a trail of influence, the broad outlines of which we will trace here.

G. E. MOORE

We begin with George Edward Moore, whose influence, along with that of his Cambridge classmate Bertrand Russell, was felt from their student days in the last decade of the nineteenth century throughout the whole of the twentieth. As a student Moore, who was to become the great defender of the Common Sense view of the world, was fascinated and perplexed by what he took to be the dismissive attitude toward common sense adopted by some of his philosophical mentors. He was particularly puzzled about the doctrines of absolute idealism that time is unreal (and so our ordinary belief that some things happen before other things must, in some way, be mistaken), that only the absolute truly exists (and so our ordinary conception of a variety of independently existing objects is incorrect), and that the essence of all existence is spiritual (and so our ordinary, non-mentalistic view of material objects is erroneous). Moore was curious how proponents of such doctrines could think themselves capable of so thoroughly overturning our ordinary ways of looking at things. How could anyone by mere reflection arrive at doctrines the certainty of which was sufficient to refute our most fundamental pre-philosophical convictions?

Before long he came to believe one couldn't. On the contrary, one's justification for a general principle of philosophy could never outweigh one's justification for the most basic tenets of the Common Sense view of the world. In essence he held that philosophers have no special knowledge that is prior to, and more secure than, the best examples of what we all pre-theoretically take ourselves to know. The effect of this position was to turn the kind of philosophy done by some of his teachers on its head. According to Moore the job of philosophy is not to prove or refute the most basic propositions, those we have no choice but to accept. It is however a central task of philosophy to explain *how* we know them. The key to doing so, he thought, was to analyze precisely what these propositions state, and hence what we know, when we know them.

Moore turned his method of analysis on two major subjects—perceptual knowledge and ethics. Although he achieved important results in both, they didn't fulfill his hopes for analysis. For example despite making a persuasive case in "A Defense of Common Sense" (1925) and "Proof of an External World" (1939) that we do know such elementary truths as *I am perceiving this and this is a human hand*, he never succeeded in explaining how, precisely, perception guarantees their truth. Moreover his speculative explorations of different analyses of their contents—briefly canvassed in "A Defense of Common Sense"—didn't advance the case very far. The paucity of these results—in which analysis aims at theoretical reconstructions of the contents of ordinary propositions—

contrasts with the modest but much more successful conception of analysis that emerges from his painstaking philosophical practice in papers such as "The Refutation of Idealism" (1903). The burden of that piece is to show that idealists who hold that all of reality is spiritual have no good reason for their view. A crucial step is the isolation and analysis of a premise—roughly *For anything to exist, or be real, is for it to be experienced*—that Moore takes to be crucial to their argument. His point is that in order to play the role required by the argument, it must be a necessary truth. But, he thinks, the only plausible ground for believing it to be necessary lies in wrongly taking the concept of being experienced to be (analytically) included in the concept of an object existing, or being real—a mistake, he thinks, that is akin to wrongly identifying the sensation of yellow with that of which it is a sensation. Putting aside the accuracy of Moore's depiction of his opponents, or of his contentious views of the distinction between analytic and synthetic propositions, the paper is a beautiful example of the theoretically modest but philosophically illuminating practice of analysis at which Moore excelled—conceptual clarification, the drawing of clear distinctions, avoidance of equivocation, logical rigor, and attention to detail.

Much the same can be said about his use of philosophical analysis in ethics. On the one hand his enormously influential view that *good* is unanalyzable may be criticized for falling prey to a crippling dilemma. On any understanding of analyzability on which the unanalyzability of good would justify Moore's claim that conclusions about what is good are not derivable from, or supported by, premises that don't contain it, his "open question" argument does not show that good is unanalyzable; whereas on any understanding of analyzability on which his argument does establish that good is unanalyzable, this result does not justify the claim that conclusions about what is good can't be derived from or supported by premises that don't talk about goodness. In this sense his most famous ethical analysis was unsuccessful. Moreover this failure was connected to his official view of analysis, which conferred a privileged status on those necessary, apriori truths that reflect part-whole relations between concepts—roughly those propositions expressed by sentences that can be reduced to logical truths by putting synonyms for synonyms (where pairs of synonyms are thought to be easily recognizable by anyone who understands them)—as opposed to those necessary, apriori truths that do not fall into this category. Far from a source of strength, this theoretically-loaded conception of analysis was, arguably, Moore's Achilles heel.

On the other hand the decidedly more modest, theoretically uncontentious, conception of analysis that emerged from his exemplary analytic practice of unrelenting, conceptual clarification undeniably advanced the subject and served as a model for generations of analytic philosophers to come. It also produced, in the first paragraph of *Principia Ethica* (1903), what may be the best expression of the guiding spirit of analytic philosophy, and philosophical analysis, ever written.

> It appears to me that in Ethics, as in all other philosophical studies, the difficulties and disagreements, of which its history is full, are mainly due to a very simple cause: namely to the attempt to answer questions, without first discovering precisely *what* question it is which you desire to answer. I do not know how far this source of error would be done away, if philosophers would *try* to discover what question they were asking, before they set about to answer it; for the work of analysis and distinction is often very difficult: we may often fail to make the necessary discovery, even though we make a definite attempt to do so. But I am inclined to think that in many cases a resolute attempt would be sufficient to ensure success; so that, if only this attempt were made, many of the most glaring difficulties and disagreements in philosophy would disappear. At all events, philosophers seem, in general, not to make the attempt, and, whether in consequence of this omission or not, they are constantly endeavoring to prove that that 'Yes' or 'No' will answer questions, to which *neither* answer is correct, owing to the fact that what they have before their minds is not one question, but several, to some of which the true answer is 'No', to others 'Yes.' (p. vii)

BERTRAND RUSSELL

Bertrand Russell's views on philosophical analysis are unique in two respects. They are more explicit, highly articulated, and theoretically fruitful than those of other leading figures; and their historical influence remains unsurpassed. The most well-known of his doctrines about philosophical analysis is his theory of descriptions presented in "On Denoting" (1905). The initial problem to be solved was an ontological one, posed by negative existentials—sentences of the form ⌜α doesn't exist⌝ in which α is a name or description. The puzzle posed by such a sentence is that if it is true then there would seem to be nothing named or described; but if α doesn't stand

for anything then it is hard to see how the sentence can be meaningful at all, let alone true. According to Russell the problem arises from false ideas about meaning—(i) the idea that the meaning of α is the entity it names or describes, and (ii) the idea that the meaning of ⌜α doesn't exist⌝ is a proposition that predicates non-existence of that entity. At first blush these ideas seem doubly problematic since, on the one hand, if α doesn't stand for anything then there is nothing for non-existence to be predicated of, and on the other if there is an object with the property of non-existence, it would seem that there must exist an object that doesn't exist, which is a contradiction. Since Russell thought that (i) and (ii) led to these paradoxical results, he rejected both. His theory of descriptions is a proposal for replacing them with a conception of meaning that avoids such paradox.

Russell begins by distinguishing grammatically proper names (like the ordinary names of people and places) from logically proper names (*this* and *that*). Whereas the meaning of a logically proper name is its referent, the meaning of a grammatically proper name n for a speaker s is given by some singular definite description, ⌜the F⌝, that s associates with n. When it comes to singular definite descriptions, Russell's view is that they are incomplete symbols, which have no meaning in isolation. By this he means three things: (i) that the objects (if any) they denote are not their meanings, (ii) that the propositions expressed by sentences containing them do not contain constituents corresponding to them, and (iii) that their meanings can be given by rules that explain the systematic contributions they make to the meanings of sentences containing them.

Consider, for example, the negative existential ⌜The F doesn't exist⌝. To understand this sentence is to grasp the proposition it expresses. However since for Russell its grammatical form is not the same as the logical form of the proposition p it expresses, he found it useful to translate it into a formula of his logical system the syntactic structure of which did match the logical structure of p. (Russell later came to think that he could dispense with propositions themselves as real entities, and get by with his logico-linguistic structures alone, but that may be regarded as a never-fully-worked-out afterthought.) The logical form of ⌜The F doesn't exist⌝ was identified with that of of $\ulcorner {\sim}\exists x \, \forall y \, (Fy \leftrightarrow y = x) \urcorner$—where the proposition expressed by this formula was seen as having three constituents: negation, the property expressed by '∃x', of being "sometimes true," and the propositional function f expressed by the sub formula $\ulcorner \forall y \, (Fy \leftrightarrow y = x) \urcorner$. This function assigns to any object o the proposition that says of o that it is identical with any object y if and only if y has the property expressed by F. Since o is identical with itself and nothing else, this means that the proposition f assigns to o is one that is true if and only if o, and only o, has the property expressed by F. Finally to say of a propositional function that it "is sometimes true" is to say that in at least one case it assigns a true proposition to an object. Putting all this together we get the result that the negative existential ⌜The F doesn't exist⌝ expresses a proposition which is true if and only if there is no object which is such that it, and only it, has the property expressed by F. Since this proposition simply denies that a certain propositional function has a certain property, neither the truth nor the meaningfulness of the negative existential that expresses it requires there to be any object with the property of non-existence.

Negative existentials were, in Russell's view, special in that they contain the grammatical predicate *exist*, which, on his analysis, does not function logically as a predicate of individuals. However his theory was intended to cover all sentences containing descriptions. Whenever ⌜is G⌝ does function as a predicate, the analysis of ⌜The F is G⌝ is $\ulcorner \exists x \, \forall y \, (Fy \leftrightarrow y = x) \,\&\, Gx \urcorner$, which may be paraphrased *there is something such that it, and only it, is F, and it is also G*. In "On Denoting," Russell showed how this analysis could be used to solve several logico-linguistic puzzles, and many other applications have been found since then. With the exception of Gottlob Frege's invention of the logical quantifiers in his *Begriffsschrift* (1879), one would be hard pressed to identify any comparably fruitful idea in the history of philosophical analysis.

Russell's revival of Frege's logicist program of reducing arithmetic to logic—in *Principia Mathematic* with Whitehead (1910, 1912) and *Introduction to Mathematical Philosophy* (1919)—represented a different, more philosophically ambitious kind of analysis. The task of deriving the axioms of Peano arithmetic from what Russell took to be axioms of pure logic required defining the arithmetical primitives *zero*, *successor*, and *natural number* in purely logical terms. Russell's approach (which he shared with Frege) was both elegant and natural. Let zero be the set whose only member is the empty set; let the successor of a set x (of sets) be a set y (of sets) with the following property: For each member of y the result of removing a member leaves one with a member of x. It follows that the successor of zero (i.e., the number one) is the set of all single-membered sets, the successor of one (i.e. the number two) is the set of all pairs, and so on. Note how natural this is. What is the number two? It is

that which all pairs have in common; more precisely, it is the set of which they, and only they, are members. Finally the set of natural numbers is defined as the smallest set containing zero and closed under the operation of successor.

With these definitions, together with Russell's proposed logical axioms (formulated within his theory of logical types, so as to avoid paradox), the axioms of Peano arithmetic can be derived as theorems. As a result, arithmetical sentences can be viewed as convenient abbreviations of the complex formulas associated with them by the Russellian definitions. Since the sentences of higher mathematics can themselves be viewed as abbreviations of complex arithmetical sentences, it seemed to many that Russell's reduction had succeeded in showing that all of mathematics can be regarded as an elaboration of pure logic and that all problems in the philosophy of mathematics could, in principle, be solved by a correct philosophical account of logic. Thus the reduction, in addition to being recognized as a substantial technical achievement, was viewed by many as a stunning demonstration of the extraordinary philosophical power of Russell's version of logico-linguistic analysis. No matter that his system of logic and theory of types was, in point of fact, epistemologically less secure than arithmetic itself; the program of attacking philosophical problems by associating the sentences that express them with hidden logical forms was considered to have taken a huge step forward.

Russell pushed the program further in *Our Knowledge of the External World* (1914), in which he applied his method of analysis to Moore's problem of the external world. The problem that perplexed Moore was that, although we know that there are material objects and although our evidence is perceptual, there seems to be a gap between this evidence and that which we know on the basis of it. Whereas material objects are public and independent of us, Moore had come to think of the data provided to us by our sensory impressions as logically private and dependent for their existence on the perceiver.

Russell set out to bridge this gap. His solution was to analyze material-object talk as talk about a system of interrelated private perspectives—a forerunner of the idea that material objects are logical constructions out of sense data. According to this view sentences that appear to be about material objects are really about the sense data of perceivers, and each material-object sentence is analyzable into a conjunction of categorical and hypothetical sentences about sense data. Apart from the obvious, Berkeleyan problems inherent in this view, its portents of the future of philosophical analysis were ominous. Prior to this Russell's main examples of analysis—his theory of descriptions and logicist reduction—were precisely formulated and well worked out. By contrast the supposed analysis of material object statements was highly programmatic—neither Russell nor anyone else ever attempted to provide a fully explicit and complete analysis of any material-object statement. It was supposed to be enough to sketch the outlines that presumed analyses were supposed to take.

This programmatic approach also characterized Russell's position in his 1918 lectures "The Philosophy of Logical Atomism," in which he sketched the outlines of an ambitious philosophical system that posited a thoroughgoing parallelism between language and the world. The idea was to use the techniques of logical and linguistic analysis to reveal the ultimate structure of reality. Before, Russell had offered analyses piecemeal—to provide solutions to different philosophical problems as they came up. Now he sought to develop a systematic framework in which philosophy would, for all intents and purposes, be identified with logico-linguistic analysis. However it was his former student, Ludwig Wittgenstein, who pushed this idea the furthest.

EARLY WITTGENSTEIN

The *Tractatus* (1922) is an intricate, ingenious, and highly idiosyncratic philosophical system of the general sort Russell had imagined. In it Wittgenstein presents his conception of a logically perfect language, which, he believes, underlies all language and, presumably, all thought. Crucial to the construction of a theory of meaning for this language is the account of its relation to the world, which we are told in the opening two sentences is the totality of facts rather than things. The simplest—atomic—sentences of language correspond (when true) to simple—atomic—facts. The constituents of these facts are metaphysically simple objects and universals named by linguistically simple expressions—logically proper names and predicates. All meaningful sentences are said to be truth functions of atomic sentences, each of which is logically independent of all other atomic sentences. Since atomic facts are similarly independent, all and only the possible assignments of truth values to atomic sentences determine possible worlds, which are possible constellations of atomic facts. The actual world is the combination all existing atomic facts.

For Wittgenstein what a sentence says is identified with the information it provides about the location of the actual world within the logical space of possible worlds. If S is atomic then S represents the actual world as being

one that contains the possible atomic fact the existence of which would make S true. If S is both meaningful and logically complex, then S is a truth function of a certain set A_s of atomic sentences, and S represents the actual world as containing a constellation of facts that corresponds to an assignment of truth values to A_s that would make S true. However, in the system of the *Tractatus*, P is a member of A_s only if there are situations in which the truth value of S is affected by which truth value is assigned to P—only if there are complete assignments of truth values to A_s which differ solely in what they assign to P that determine different truth values for S. Since, when S is a tautology, its truth does not depend on the truth values of any atomic sentence in this way, it follows that S isn't a truth function of any non-empty set of such sentences.

For Wittgenstein this means that tautologies don't provide any information about the world, and so, strictly speaking, don't say anything. In this sense tautologies are not fully meaningful, though we may regard them as meaningful in the degenerate sense of arising from meaningful atomic sentences by permitted applications of truth-functional operators. Thinking of them in this way we may take tautologies to be true, so long as we understand that they don't state or correspond to any facts. For Wittgenstein there are no necessary facts for necessary truths to correspond to. Rather their truth is an artifact of our linguistic system of representation. Because of this, he thought, they should be knowable apriori, simply by understanding them and recognizing their form.

Many philosophers found the strikingly simple Tractarian conception of necessity, apriority, and logical truth to be compelling. According to the *Tractatus* (i) all necessity is linguistic necessity, in the sense of being the result of our system of representing the world, rather than the world itself; (ii) all linguistic necessity is logical necessity, in that all necessary truths are tautologies; (iii) all tautologies are knowable apriori; and (iv) only necessary truths are apriori. In short the necessary, the apriori, and the logically true are one and the same. These truths make no claims about the world but instead constitute the domain of logic, broadly construed. All other truths are contingent and knowable only by empirical investigation. These truths do make claims about the world and constitute the domain of science.

There are no other meaningful sentences, save for the logically or contingently false. According to the *Tractatus*, all meaningful sentences are either tautologies, contradictions, or contingent, aposteriori statements which are truth functions of atomic sentences that describe possible combinations of the basic metaphysical simples that make up the world. Since virtually all of the traditional statements of ethics, philosophy, and religion seem to fall outside these categories, Wittgenstein concluded that these statements are nonsense. No aspect of his system was more fascinating to readers of the *Tractatus* than this consequence of his global criterion of intelligibility. Moreover his conclusion was not limited to language. If one assumes, as Wittgenstein clearly did, that all genuine thoughts are in principle expressible by meaningful sentences then his criterion not only fixes the limits of meaning but it also fixes the limits of thought. Since ethical, philosophical, and religious sentences are meaningless, they don't express propositions; since there are no such propositions for us to believe, we have no ethical, philosophical, or religious beliefs.

Where does this leave philosophy and philosophical analysis? The lesson of the *Tractatus* is that here are no meaningful philosophical claims and hence no genuine philosophical questions for philosophers to answer. What then is responsible for the persistence of the discipline and for the illusion that it is concerned with real problems for which solutions might be found? Linguistic confusion. As Wittgenstein saw it all the endless disputes in philosophy are due to this one source. If we could ever fully reveal the workings of language, our philosophical perplexities would vanish, and we would see the world correctly. Fortunately, philosophy can help. Although there are no new true propositions for philosophers to discover, they can clarify the propositions we already have. Like Russell, Wittgenstein believed that everyday language disguises thought by concealing true logical form. The proper aim of philosophy is to strip away the disguise and illuminate the form. In short, philosophy is a kind of linguistic analysis that doesn't solve problems but dissolves them. As he put it in his first post-*Tractatus* paper, "Some Remarks on Logical Form" (1929),

> The idea is to express in an appropriate symbolism what in ordinary language leads to endless misunderstandings. That is to say, where ordinary language disguises logical structure, where it allows the formation of pseudo-propositions, where it uses one term in an infinity of different meanings, we must replace it by a symbolism which gives a clear picture of the logical structure, excludes pseudo-propositions, and uses its terms unambiguously.
>
> (P. 163)

Though the Wittgenstein of the *Tractatus* did not himself practice this form of analysis, the vision of analysis he

articulated was one that later philosophers found attractive in its own right, quite apart from the doctrines that led him to it.

LOGICAL POSITIVISM

We now turn to something new—a self-conscious school of philosophy that arose through the collaborative efforts of several like-minded thinkers, including, most prominently, Rudolf Carnap, Moritz Schlick, Hans Reichenbach, A. J. Ayer, and Carl Hempel. The evolving creation of many minds, logical positivism was not monolithic; there was always plenty of disagreement on matters of detail, and even its central doctrines were never formulated in a way that commanded universal assent. The positivists did, however, share a common commitment to the development of certain themes inherited largely from Russell and Wittgenstein. From Russell they took the theory of descriptions as the paradigm of philosophical analysis (so characterized by F. P. Ramsey), the reduction of arithmetic to logic as the key to the nature of all mathematical truth (set out in Hempel's "On the Nature of Mathematical Truth," 1945), and the systematic, empiricist reconstruction of our knowledge of the external world—undertaken in Carnap's *The Logical Structure of the World* (1928). From Wittgenstein they took the idea of a test of intelligibility, the identification of necessary, apriori, and analytic truth, the bifurcation of all meaningful statements into the analytic versus empirical, the dismissal of whole domains of traditional philosophy as meaningless nonsense, and the goal of philosophy as the elimination of linguistic confusion by philosophical analysis.

The centerpiece of logical positivism was, of course, the empiricist criterion of meaning, which stated roughly that a non-analytic, non-contradictory sentence S is meaningful if and only if S is in principle verifiable or falsifiable—where verifiability and falsifiability are thought of as logical relations R_V and R_F between observation statements and S. Although the idea initially seemed simple, the devil proved to be in the details. One source of contention was the nature of observation statements. Initially Carnap, Schlick, and others construed them as reports of private sense data of observers. However the dangers of solipsism and phenomenalism soon forced a retreat to reports of (unaided) observation of everyday physical objects. Even then the theoretical / observational distinction proved elusive, with obvious strain on the clarity and plausibility of the criterion of meaning.

Defining the relations R_V and R_F that were to hold between meaningful (empirical) sentences and observation statements proved even more problematic. Initially it was hoped that the needed relations could be something quite strong—like the notion of being either conclusively verifiable (i.e., logically entailed by some finite, consistent set of observation statements) or conclusively falsifiable (i.e., something the negation of which is conclusively verifiable). However it soon became clear that when R_V and R_F are defined in this way, many obviously meaningful statements of science and everyday life are wrongly characterized as meaningless. This led to the attempt, illustrated by Ayer's proposal in the Introduction to the second edition of *Language, Truth and Logic* (1946), to define empirical meaningfulness in terms of a weak notion of verifiability—roughly that of being a statement which, when combined with an independently meaningful theory T, logically entailed one or more observation statements not entailed by T alone. However, as Alonzo Church demonstrated in his 1949 review of Ayer, this criterion was far too promiscuous, classifying no end of nonsense as meaningful.

There were of course other attempts to secure a workable empiricist theory of meaning, such as Carnap's criterion of translatability into an empiricist language, sketched in his 1936 essay "Testability and Meaning." But as Hempel showed in "Problems and Changes in the Empiricist Criterion of Meaning" (1950) this formulation runs into serious problems over theoretical terms in science. In Hempel's view the source of these problems is that sentences about theoretical entities are meaningful by virtue of being embedded in a network of hypotheses and observational statements, which as a whole makes testable predictions. As W. V. Quine emphasized even more forcefully in "Two Dogmas of Empiricism" (1951) these predictions are the product of all the different aspects of the system working together—in the sense that, given a set of observational predictions made by a theoretical system, one cannot in general match each prediction with a single discreet hypothesis, or small set of hypotheses.

Quine suggests that this is the crucial fact that makes it impossible to devise an adequate criterion of empirical meaningfulness for individual sentences. If for each sentence S we could isolate a set P of predictions made by S alone, and if P exhausted the contribution made by S to the predictions made by the theory as a whole then one could define S in terms of P. However the interdependence of S with other sentences in the system makes this impossible. Thus, Quine maintained, what we have to look for is not the empirical content of each statement taken in isolation, but rather its role in an articulated sys-

tem that, as a whole, has empirical content. This point effectively marked the end of the positivists' version of verificationism.

QUINE

From the *Tractatus* through logical positivism and beyond, many analytic philosophers identified the apriori with the necessary and attempted to explain both by appealing to the analytic. As they saw it there simply is no explaining what necessity is, how we can know any truth to be necessary, or how we can know anything apriori, without appealing to statements that are, and are known to be, true by virtue of meaning. From this point of view necessary and apriori truths had better be analytic, since if they aren't one can give no intelligible account of them at all. Ironically this theoretical weight placed on analyticity left the doctrines about necessity, apriority, and analyticity advocated by positivists and others vulnerable to a potentially devastating criticism. If it could be shown that analyticity cannot play the explanatory role assigned to it, then their commitment to necessity, apriority, and perhaps even analyticity itself might be threatened. This was precisely Quine's strategy.

He launched his attack in "Truth by Convention" (1936), the target of which is the linguistic conception of the apriori. According to this view all apriori knowledge is knowledge of analytic truths, which in turn is explained as arising from knowledge of the linguistic conventions governing our words. This view was attractive because it provided a seemingly innocuous answer to the question of how any statement could be known without empirical confirmation: A statement can be known in this way only if it is devoid of factual content—that is, only if its truth is entirely due to its meaning. Surely, it was thought, there is no mystery in our knowing what we have decided our words are to mean. But then, it was concluded, there must be no mystery in the idea that the truth of a sentence may follow, and be known to follow, entirely from such decisions. Putting these two ideas together, proponents of the linguistic conception of the apriori thought that they had found a philosophical explanation of something that otherwise would have been problematic.

Quine argued that this is not so. As noted, the proposed explanation rests on two bits of knowledge taken to be unproblematic—(i) knowledge of what our words mean, and (ii) knowledge that the truth of certain sentences follows from our decisions about meaning. However there is a problem here, located in the words *follows from*. Clearly we don't stipulate the meanings of all the necessary / apriori / analytic truths individually. Rather, it must be thought, we make some relatively small number of meaning stipulations and then draw out the consequences of these stipulations for the truth of an indefinitely large class of sentences. What is meant here by *consequences*? Not wild guesses or arbitrary inferences with no necessary connection to their premises. No, by *consequences* proponents of the linguistic apriori meant something like *logical consequences, knowable apriori to be true if their premises are true*. But now we have gone in a circle. According to these philosophers, all apriori knowledge of necessary truths—including apriori knowledge of logical truths—arises from our knowledge of the linguistic conventions we have adopted to give meanings to our words. However, in order to derive this apriori knowledge from our linguistic knowledge, one has to appeal to antecedent knowledge of logic itself. Either this logical knowledge is apriori or it isn't. If it is then some apriori knowledge is not explained linguistically; if it isn't then it is hard to see how any knowledge could qualify as apriori. Since neither alternative was acceptable to proponents of the linguistic apriori, Quine's attack was a telling one.

Fifteen years later, in "Two Dogmas of Empiricism" (1951), he renewed it. He agreed with the positivists' premise that there is no explaining necessity and apriority without appealing to analyticity. However he challenged the idea that any genuine distinction can be drawn between the analytic and the synthetic without presupposing the very notions they are supposed to explain—a point he sought to drive home by demonstrating the circularity of the most obvious attempts to define analyticity. Hence, he concluded, there is no way of explaining and legitimating necessity and apriority—or analyticity either. For him this meant that there is no genuine distinction to be drawn between the analytic and the synthetic, the necessary and the contingent, or the apriori and the aposteriori. The idea that any such distinctions exist was one of the two dogmas targeted in his article.

In assessing this argument it is important to remember that it was directed at a specific conception of analyticity, which was taken to be the source of necessity and apriority. Although this conception was widely held at the time Quine wrote, it is radically at variance with the post-Kripkean perspective according to which necessity and apriority are, respectively, metaphysical and epistemological notions that are non-coextensive and capable of standing on their own. From this perspective the attempt to explain necessity and apriority in terms of analyticity appears to be badly mistaken. Since Quine's circularity argument shares the problematic presupposition that all these notions are acceptable only if such an explanation

can be given, it doesn't come off much better. For this reason Quine should not be seen as giving a general argument against analyticity. At most his argument succeeds in undermining one particular conception that enjoyed a long run among analytic philosophers in the middle fifty years of the twentieth century.

The second dogma attacked by Quine is *radical reductionism*, the view that every meaningful sentence is translatable into sentences about sense experience. Quine points out that the two dogmas—(i) that there is a genuine analytic / synthetic distinction, and (ii) radical reductionism—are linked in empiricist thinking by verificationism. Roughly speaking, verificationism holds that two sentences have the same meaning if and only if they would be confirmed or disconfirmed by the same experiences. Given this notion of synonymy, one could define analyticity as synonymy with a logical truth. Thus if verificationism were correct then the analytic / synthetic distinction would be safe. Similarly if verificationism, or at any rate a particularly simple version of verificationism, were correct then any empirical sentence would be translatable into the set of observation sentences that would confirm it, and radical reductionism would be saved. For these reasons, Quine concludes, if simple verificationism were correct then the two dogmas of empiricism would be corollaries of it.

By the time Quine wrote "Two Dogmas," verificationism, as a theory of meaning for individual sentences, was already dead, as was radical reductionism. Nevertheless he noted that some philosophers still maintained a modified version of the latter according to which each (synthetic) statement is, by virtue of its meaning, associated with a unique set of possible observations that would confirm it and another that would disconfirm it. Against this Quine argued that verification is holistic, by which he meant that most sentences don't have predictive content in isolation but are empirically significant only insofar as they contribute to the predictive power of larger empirical theories. Since he continued to assume, with the positivists, that meaning is verification, his position was one of holistic verificationism. According to this view the meaning of a theory is, roughly, the class of possible observations that would support it, and two theories have the same meaning if and only if they would be supported by the same possible observations. Since individual sentences don't have meanings on their own, any sentence can be held true in the face of any experience (by making necessary adjustments elsewhere in one's overall theory), and no sentence is immune from revision—since given a theory T incorporating S, Quine thought that one could construct a different, but predictively equivalent, and hence synonymous, theory T incorporating the negation of S.

The resulting picture of philosophy and philosophical analysis that emerges from Quine's work is radically at variance with any we have seen. He rejects the doctrine that philosophical problems arise from confusion about the meanings of words or sentences, and with it the conception of philosophy as providing analyses of their meaning. He rejects these views because he rejects their presuppositions—that words and sentences have meanings in isolation and that we can separate out facts about meanings or linguistic conventions from the totality of all empirical facts. For Quine philosophy is continuous with science. It has no special subject matter of its own, and it is not concerned with the meanings of words in any special sense. Philosophical problems are simply problems of a more abstract and foundational sort than the ordinary problems of everyday science.

In later years Quine put less emphasis on holistic verificationism (which is itself beset with problems akin to earlier versions of verificationism), but he did not back away from his skepticism about our ordinary, pre-theoretic conception of meaning. Instead he deepened and extended his attack with his doctrine of the indeterminacy of translation in *Word and Object* (1960) and its corollary, the inscrutability of reference, in "Ontological Relativity" (1969). Since Quine, the naturalist, could find no place in nature for meaning and reference as ordinarily conceived, he repudiated both in favor of radically deflated, behaviorist substitutes. Thus it should not be surprising that there is no place in his brave new world for philosophical analysis as a distinctive intellectual activity. Nevertheless his actual philosophical practice is hard to discern from that of his analytic predecessors. Like them he does little, when arguing for his central doctrines, to delineate their alleged contributions to the observational predictions made by our overall theory of the world.

LATER WITTGENSTEIN

In *The Philosophical Investigations* (1953) Wittgenstein outlines a new, essentially social conception of meaning that contrasts sharply with the one presented in the *Tractatus*. In the earlier work language was viewed on the model of a logical calculus in which conceptual structure is identical with logical structure, and all meaningful sentences are truth-functions of atomic sentences that represent metaphysically simple objects standing in relations isomorphic to those in which logically proper names stand

in the sentences themselves. In the *Investigations* the picture is quite different. Language is no longer seen as a calculus, derivability by formal logical techniques is accorded no special role in explaining conceptual connections among sentences, and naming is not taken to be the basis of meaning. Instead, meaning arises from socially conditioned agreement about the use of expressions to coordinate the activities and further the purposes of their users. For the later Wittgenstein, to know the meaning of an expression is not to know what it names or how to define it but to know how to use it in interacting with others.

According to this conception of meaning, understanding a word is not a psychological state but rather a disposition to apply it in the correct way over a wide range of cases; where by *the correct way* we do not mean the way determined by a rule the speaker has internalized. The problem, as Wittgenstein sees it, with appealing to such rules to explain our understanding of words is that rules are themselves made up of symbols that must be understood if they are to be of any use. Obviously this sort of explanation can't go on forever. In the end we are left with a large class of words or symbols that we understand and are able to apply correctly, despite the fact that what guides us and makes our applications correct are not further rules of any sort. When we reach rock bottom we are not guided by rules at all; we simply apply expressions unthinkingly to new cases.

What determines whether these new applications are correct? The mere fact that I am inclined to call something F can't guarantee that I am right. If my use of F is to be meaningful, there must be some independent standard that my application is required to live up to in order to be correct. Wittgenstein thinks this standard can't come from me alone. The reason it can't is that the same argument that shows that the standard of correctness cannot be determined by an internalized rule can be repeated to establish that it can't be determined by any belief, intention, or other contended mental state of mine. The problem, Wittgenstein thinks, is that in order to perform such a role, any such mental state must itself have gotten its content from somewhere. A regress argument can then be used to conclude that the contents of all my words and all my mental states must, in the end, rest on something other than my mental states. Thus, he suggests, the standard of correctness governing my use of F cannot rest on anything internal to me, but must somehow come from the outside. What more natural place to look for this than in the linguistic community of which I am a part? Hence, he suggests, for me to use F correctly is for me to apply it in conformity with the way it is applied by others. For Wittgenstein this, in turn, implies that F must be associated with public criteria by which someone else can, in principle, judge whether my use of it is correct. Language is essentially public; there can be no logically private language.

This conception of language leads Wittgenstein to a new conception of philosophy and philosophical analysis. He continues to believe that philosophical problems are linguistic, and that philosophical analysis is the analysis of language—but this analysis is no longer seen as a species of logical analysis. According to the new conception there is no such thing as the logical form of a sentence, and one should not imagine that sentences have unique analyses. According to Wittgenstein we do not give an analysis of a sentence because there is anything wrong with it that demands clarification. We give an analysis when something about it leads us into philosophical confusion. The same sentence might even receive different analyses, if people become confused about it in different ways. In such a case each analysis may clear up a particular confusion, even if no analysis clears them all up.

Accompanying this deflationary view of analysis is a highly deflationary conception of philosophy. According to the *Investigations* the philosophical analysis of language does not aim at, and cannot issue in, theories of any kind. Philosophy, as Wittgenstein says in section 109, "is a battle against the bewitchment of our intelligence by means of language." According to this view the task of philosophy is essentially therapeutic. It is the untangling of linguistic confusions, achieved by examining our words as they are ordinarily used, and contrasting that use with how they are misused in philosophical theories and explanations.

This deflationary conception arises naturally from Wittgenstein's new ideas about meaning, plus certain unquestioned philosophical presuppositions that he brings to the enterprise. These include his long-held convictions (i) that philosophical theses are not empirical, and hence must be necessary and apriori, and (ii) that the necessary, the apriori, and the analytic are one and the same. Because he takes (i) and (ii) for granted, he takes it for granted that if there are any philosophical truths, they must be analytic. To this he adds his new conception of meaning—with its rejection of abstract logical forms, its deflationary view of rule-following and algorithmic calculation, and its emphasis on social conditioning as generating agreement in our instinctive applications of words. Having jettisoned his old conception of meaning as something hidden and replaced it with a conception of

meaning that sees it as arising from an unquestioning, socially-conditioned agreement, he has little room in his conceptual universe for surprising philosophical truths. Genuinely philosophical truths, if there are any, can only be necessary and apriori, which in turn are taken to be true in virtue of meaning.

But how are the analytic truths of interest to a philosopher to be established if they are not to be translated into the formulas of a logical calculus and demonstrated by being given rigorous but sometimes also innovative and insightful logical proofs? For the Wittgenstein of the *Investigations*, the answer is that they don't need to be established, since they are already implicitly recognized by competent language users. To be sure they may sometimes need to be brought into focus by assembling examples of ordinary use that illustrate the constitutive role they play in our language; but there is little room here for surprising philosophical discoveries. Such is the official view of the *Investigations*.

As with the *Tractatus*, there is an evident problem here. Wittgenstein's official view of philosophy is at variance with his own philosophical practice. His general theses about language and philosophy (to say nothing of his surprising and, arguably, revisionist views about sensation and other psychological language arising from the private language argument) are by no means obvious or already agreed upon; nor are they the sorts of things that one can just see to be true, once they are pointed out. On the contrary they require substantial explanation and argument, if they are to be accepted at all. As was so often the case throughout the twentieth century, the practice of philosophical analysis—understood as whatever it is that analytic philosophers do—eluded the official doctrines about analysis propounded by its leading practitioners.

THE ORDINARY LANGUAGE SCHOOL

This school, which received great impetus from the *Investigations*, was shaped by two leading ideas. The first was that since philosophical problems are due solely to the misuse of language, the job of the philosopher is not to construct elaborate theories to solve philosophical problems but to expose linguistic confusions that fooled us into thinking there were genuine problems to be solved in the first place. The second idea was that meaning itself—the key to progress in philosophy—is not to be studied from an abstract scientific or theoretical perspective. Rather philosophers were supposed to assemble observations about the ordinary use of words, and to show how misuse of certain words leads to philosophical perplexity. In retrospect this combination of views seems quite remarkable: All of philosophy depends on a proper understanding of something that there is no systematic way of studying. Fortunately this anti-theoretical approach changed over time with much of the progress in the period being marked by significant retreats from it—including Austin's theory of performatives in *How to Do Things with Words* and Paul Grice's theory of conversational implicature in "Logic and Conversation" (both originally delivered as the William James Lectures at Harvard, in 1955 and 1967, respectively).

A good example of the standard, anti-theoretical approach is Ryle's *Dilemmas* (1953), in which he identifies the main aim of philosophy as that of resolving dilemmas. For Ryle a dilemma arises when obvious theories or platitudes appear to conflict with one another. In such cases a view that is unobjectionable in its own domain comes to seem incompatible with another view that is correct when confined to a different domain. When this happens we find ourselves in the uncomfortable position of seeming to be unable jointly to maintain a pair of views, each of which appears correct on its own. Ryle believes that in most cases the apparent conflict is an illusion to be dispelled by philosophical analysis. However, the needed analysis is not a matter of defining key concepts or uncovering hidden logical forms. Although analysis is conceptual what is wanted is never a sequence of definitions that could in principle be presented one by one. Instead Ryle compares the required analysis to the description of the position of wicket keeper in cricket. Just as one can't describe that position without describing how it fits in with all the other positions in cricket, so, Ryle thinks, one cannot usefully analyze a concept without tracing its intricate connections with all the members of the family of concepts of which it is a part.

His most important application of this method is to psychological language, in *The Concept of Mind* (1949). There he rejects what he calls the myth of "the Ghost in the Machine," according to which belief and desire are causally efficacious, mental states of which agents are non-inferentially aware. Ryle takes this view to be "entirely false" and to be the result of what he calls "a category-mistake," by which he means that it represents mental facts as belonging to one conceptual type, when they really belong to another. He illustrates this with the analogy of someone who visits different buildings and departments of a university and then asks "But where is the university?" Here the category mistake is that of taking the university to be a separate building or department alongside the others the visitor has seen, rather than

being the way in which all the different buildings and departments are coordinated.

Similarly, Ryle maintains, someone who believes that the mind is something over and above the body fails to realize that the mind is not a separate thing, and that talk of the mental is really just talk about how an agent's actions are coordinated. According to this view, to attribute beliefs and desires to an agent is not to describe the internal causes of the agent's action but simply to describe the agent as one who would act in certain ways if certain conditions were fulfilled. This is rather surprising. According to Ryle's ordinary-language ideology, philosophy is not supposed to give us new theories but to untangle linguistic confusions—leaving us, presumably, with a less muddled version of what we pretheoretically thought. Here, however, his aim was to undermine a certain widely-held view of the mind and to provide what, arguably, amounts to a sweeping revision of our ordinary conception of the mental.

J. L. Austin was similarly ambitious. In his elegant classic *Sense and Sensibilia*, published in 1962 but delivered as lectures several times between 1947 and 1959, he attempted to dissolve, as linguistically confused, phenomenalism, skepticism about knowledge of the external world and the traditional sense-data analysis of perception. His goal was to show these positions to be incoherent by undermining the presupposition that our knowledge of the world always rests on conceptually prior evidence of how things perceptually appear. For this he employed two main strategies. One was to try to show that certain statements—such as, "there is a pig in front of me" in normal circumstances, with the animal in plain sight—are statements about which the claim that knowledge of them requires evidence of how things appear cannot be true. Austin drew this conclusion from the observation that it would be an abuse of language for the speaker in such a situation to say, "It appears that there is a pig in front of me," or "I have evidence that there is a pig in front of me." His other strategy was to argue that appearance statements themselves are parasitic on ordinary non-appearance statements and so cannot be regarded as conceptually prior to the latter.

Neither strategy was successful. The first was rebutted by Ayer in "Has Austin Refuted the Sense Datum Theory" (1967), in which he pointed out that the abuse that Austin spotted was, in effect, a matter of Gricean conversational implicature (*Don't make your conversational contribution too weak!*) from which no conclusion about the possibility of knowledge without evidence can be drawn. The general lesson here is that not all matters of language use (or misuse) are matters of meaning (or truth). Austin's second strategy, though not similarly rebutted, was not developed in enough detail to be compelling. In addition it faced the general difficulty (common to many ordinary-language attempts to undermine skepticism) of appealing to non-skeptical claims about meaning to refute the skeptic. Even if the view of meaning is correct, it may have little argumentative force against a determined skeptic.

By contrast the theory of performative utterances given in *How to Do Things with Words* (1962) has become an enduring fixture of the study of language. The idea, in its simplest form, is that utterances of sentences like "I promise to come" or "I name this ship The Ferdinand" are, in proper circumstances, not reports of actions but performances of them. Although there have been many disputes about how to develop this idea, there is no question that there is something to it. Austin himself was inclined to think that performative utterances of this sort were attempts, not to state facts, but to perform certain conventionally recognized speech acts.

For a time this idea generated considerable optimism about performative analyses of important philosophical concepts of the sort illustrated by Peter Strawson's 1949 paper, "Truth"—according to which ⌜It is true that S⌝ is analyzed as ⌜I concede / confirm / endorse that S⌝—and R. M. Hare's *The Language of Morals* (1952)—according to which ⌜ That is a good N⌝ is assimilated to ⌜ I commend that as an N⌝. However, these views, along with other ambitious attempts to use performative analyses to sweep away age-old philosophical problems, ran into serious difficulties. Chief among them was the point—made by Peter Geach in "Ascriptivism" (1960) and John Searle in "Meaning and Speech Acts" (1962)—that any analysis of the meaning of S must explain the contribution S makes to complex sentences of which it is a constituent. Since analyses that focus exclusively on the speech acts performed by utterances of S on its own don't—and often can't—do this, they cannot be taken to be correct accounts of meaning. This reinforced a message noted earlier; not all aspects of language use are aspects of meaning. As this point sunk in, the need for systematic theories to sort things out became clear, and the ordinary language era drew to a close.

LATER DEVELOPMENTS

Many philosophers found what they were looking for in Donald Davidson's attempt to construct, in the 1960s and 1970s, a theory of meaning for natural language modeled on Alfred Tarski's formal definition of truth for logic and

mathematics. According to Davidson it is possible to construct finitely axiomatizable theories of truth for natural languages L that allow one to derive—from axioms specifying the referential properties of its words and phrases—a T-sentence, ⌜'S' is a true sentence of L if and only if p⌝, for each sentence S of L, which gives its truth conditions. Since such a theory gives the truth conditions of every sentence on the basis of its semantically significant structure, it is taken to count as a theory of meaning for L. The theory is empirically tested by comparing the situations in which speakers hold particular sentences to be true with the truth conditions it assigns to those sentences. According to Davidson's view the correct theory of meaning is, roughly, the theory T_M according to which the conditions in which speakers actually hold sentences to be true most closely matches the conditions in which T_M, plus our theory of the world, predicts the sentences to be true. Roughly put Davidson takes the correct theory to be the one according to which speakers of L turn out to be truth tellers more frequently than on any other interpretation of L.

This bold idea generated a large volume of critical comment, both pro and con, over the next two decades. One important cluster of problems centers around the fact that the T-sentences generated by Davidsonian theories are material biconditionals and so provide truth conditions of object-language sentences only in the very weak sense of pairing each such sentence with some metalanguage sentence or other that has the same truth value.

One popular way of countering this difficulty is to strengthen the theory of meaning by putting it in the form of a theory of truth relative to a context of utterance and a possible world-state. This approach, widely known as *possible worlds semantics*, was pioneered from the 1940s through the 1970s by Carnap, Saul Kripke, Richard Montague, David Lewis, and David Kaplan, among others. As commonly pursued it involves enriching the formal languages amenable to Tarski's techniques, so that they incorporate more and more of the concepts found in natural language—including modal concepts expressed by words like *actual*, *necessary*, *possible*, *could*, and *would*, temporal concepts expressed by natural-language tenses, and indexical expressions like *I*, *we*, *you*, *he*, *now*, and *today*. By the end of the century it had become possible to imagine the day in which natural languages would be treatable in something close to their entirety by the descendants of the logical techniques initiated by Tarski. Analyses of central philosophical concepts, formulated in terms of possible world-states, had also become commonplace, as illustrated by the highly influential treatment of counterfactual conditionals given by Robert Stalnaker and David Lewis as well as Lewis's related analysis of causation.

However the most important philosophical development in the last half of the century occurred in Princeton in January of 1970 when Saul Kripke, then twenty-nine years old, delivered the three lectures that became *Naming and Necessity*. Their impact was profound, immediate, and lasting. In the philosophy of language Kripke's work ranks with that of Frege in the late nineteenth century, and of Russell and Tarski in the first half of the twentieth. Beyond the philosophy of language, it fundamentally changed the way in which much philosophy is done. The most important aspects of the work are (i) a set of theses about the meaning and reference of proper names according to which neither their meanings nor reference-determining conditions are determined by descriptions associated with them by speakers; (ii) a corresponding set of theses about the meaning and reference of natural kind terms such as *heat*, *light*, *water*, and *tiger*; (iii) a compelling defense of the metaphysical concepts of necessity and possibility; (iv) a sharp distinction between necessity and apriority; (v) forceful arguments that some necessary truths are knowable only aposteriori and some apriori truths are contingent; and (vi) a persuasive defense of the view that objects have some of their properties essentially and others accidentally. In addition to these explicit aspects of the work, Kripke's discussion had far-reaching implications for what has come to be known as externalism about meaning and belief—roughly the view that the meanings of one's words, as well as the contents of one's beliefs, are partly constituted by facts outside oneself. Finally *Naming and Necessity* played a large role in the implicit but widespread rejection of the view—so popular among earlier analytic philosophers—that philosophy is nothing more than the analysis of language.

See also Analytic and Synthetic Statements; Analytic Feminism; Austin, John Langshaw; Ayer, Alfred Jules; Carnap, Rudolf; Common Sense; Davidson, Donald; Frege, Gottlob; Grice, Herbert Paul; Hare, Richard M.; Hempel, Carl Gustav; Idealism; Kaplan, David; Kripke, Saul; Lewis, David; Logical Positivism; Materialism; Montague, Richard; Moore, George Edward; Philosophy of Language; Presupposition; Quine, Willard Van Orman; Ramsey, Frank Plumpton; Reichenbach, Hans; Russell, Bertrand Arthur William; Ryle, Gilbert; Schlick, Moritz; Searle, John; Strawson, Peter Frederick; Tarski, Alfred; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

Austin, John Langshaw. *How to Do Things with Words*. New York: Oxford University Press, 1962.

Austin, John Langshaw. *Sense and Sensibilia*. London: Oxford University Press, 1962.

Ayer, Alfred Jules. "Has Austin Refuted the Sense-Datum Theory?" *Synthese* 17(1967): 117–140.

Ayer, Alfred Jules. *Language, Truth and Logic*. New York: Dover, 1952.

Carnap, Rudolf. *The Logical Structure of the World* (1928). Berkeley: University of California Press, 1969.

Carnap, Rudolf. *Meaning and Necessity*. Chicago: University of Chicago, 1947.

Carnap, Rudolf. "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–71, and 4 (1937): 1–40.

Church, Alonzo. *Review of Language, Truth and Logic*. 2nd ed. *Journal of Symbolic Logic* 14 (1949): 52–53.

Davidson, Donald. "Truth and Meaning." *Synthese* 17 (1967): 304–323.

Davidson, Donald. "Radical Interpretation." *Dialectica* 27 (1973): 313–328. Reprinted in *Inquiries into Truth and Interpretation*. Oxford: Clarendon, 2001.

Davies, Martin. *Meaning, Quantification, and Necessity*. London: Routledge & Kegan Paul, 1981.

Dummett, Michael. "What Is a Theory of Meaning?" In *Mind and Language*, edited by Samuel Guttenplan. Oxford: Clarendon Press, 1975.

Dummett, Michael. "What Is a Theory of Meaning? (II)." In *Truth and Meaning*, edited by Gareth Evans and John McDowell. Oxford: Clarendon Press, 1976.

Foster, John A. "Meaning and Truth Theory." In *Truth and Meaning*, edited by Gareth Evans and John McDowell. Oxford: Clarendon Press, 1976.

Frege, Gottlob. *Begriffsschrift*. Halle, AS: Louis Nebert, 1879.

Geach, Peter. "Ascriptivism." *Philosophical Review* 69 (1960): 221–225.

Grice, Paul. "Logic and Conversation." In *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

Hare, Richard. *The Language of Morals*. Oxford: Clarendon Press, 1952.

Hempel, Carl. "On the Nature of Mathematical Truth." *American Mathematical Monthly* 52 (1945): 543–556. Reprinted in *The Philosophy of Mathematics*. 2nd ed., edited by Paul Benacerraf and Hilary Putnam. Cambridge, U.K.: Cambridge University Press, 1983.

Hempel, Carl. "Problems and Changes in the Empiricist Criterion of Meaning." *Revue Internationale de Philosophie* 4 (1950): 41–63. Reprinted in *Logical Positivism*, edited by Alfred Jules Ayer. New York: The Free Press, 1959.

Kaplan, David. "Demonstratives." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press, 1989.

Kripke, Saul. "A Completeness Theorem in Modal Logic." *Journal of Symbolic Logic* 24 (1959): 1–14.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980. Originally published in *Semantics of Natural Languages*, edited by Donald Davidson and Gilbert Harman. Dordrecht: Reidel, 1972.

Kripke, Saul. "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94. Reprinted in *Reference and Modality*, edited by Leonard Linsky. London: Oxford University Press, 1971.

Lewis, David, "Causation." *Journal of Philosophy* 70 (1973): 556–567.

Lewis, David. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

Montague, Richard. *Formal Philosophy*. New Haven: Yale University Press, 1974.

Moore, George Edward. *Principia Ethica*. London: Cambridge University Press, 1903.

Moore, George Edward. "The Refutation of Idealism." *Mind* 12 (1903): 433–453.

Moore, George Edward. "A Defense of Common Sense." In *Contemporary British Philosophy* (Second Series), edited by John Henry Muirhead. New York: Macmillan, 1925. Reprinted in *Philosophical Papers*. New York: Collier 1962.

Moore, George Edward. "Proof of an External World." *Proceedings of the British Academy* XXV (1939): 273–300.

Quine, Willard Van Orman. "Ontological Relativity." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Quine, Willard Van Orman. "Truth by Convention." *Philosophical Essays for A. N. Whitehead*, edited by O. H. Lee. New York: Longmans, 1936. Reprinted in *Ways of Paradox*. New York: Random House, 1966.

Quine, Willard Van Orman. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43.

Quine, Willard Van Orman. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Ramsey, Frank. *The Foundations of Mathematics*. London: Routledge and Kegan Paul, 1931.

Russell, Bertrand. *Introduction to Mathematical Philosophy*. New York: Macmillan, 1919. Reprinted New York: Dover, 1993.

Russell, Bertrand. "Knowledge by Acquaintance and Knowledge by Description." *Proceedings of the Aristotelian Society* 11 (1910–1911): 108–128.

Russell, Bertrand. "On Denoting." *Mind* 14 (1905): 479–493.

Russell, Bertrand. *Our Knowledge of the External World*. Chicago and London: Open Court, 1914. Reprinted London: Routledge, 1993.

Russell, Bertrand. "The Philosophy of Logical Atomism." *The Monist* 28 (1918): 495–527; and 29 (1919): 33–63, 190–222, 344–80; reprinted La Salle, IL: Open Court, 1985.

Russell, Bertrand, and Alfred North Whitehead. *Principia Mathematica*. Vol. 1. London: Cambridge University Press, 1910.

Russell, Bertrand, and Alfred North Whitehead. *Principia Mathematica*. Vol. 2. London: Cambridge University Press, 1912.

Ryle, Gilbert. *The Concept of Mind*. New York: Barnes and Noble, 1949.

Ryle, Gilbert. *Dilemmas*. Cambridge, U.K.: Cambridge University Press, 1953.

Searle, John. "Meaning and Speech Acts." *Philosophical Review* 71 (1962): 423–432.

Soames, Scott. *Philosophical Analysis in the Twentieth Century*, Vol. 1: *The Dawn of Analysis*. Princeton, NJ: Princeton University Press, 2003.

Soames, Scott. *Philosophical Analysis in the Twentieth Century*, Vol. 2: *The Age of Meaning*. Princeton, NJ: Princeton University Press, 2003.

Soames, Scott. "Semantics and Semantic Competence." *Philosophical Perspectives* 3 (1989): 575–596.

Soames, Scott. "Truth, Meaning, and Understanding." *Philosophical Studies* 65 (1992): 17–35.

Stalnaker, Robert. "A Theory of Conditionals." In *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series 2. Oxford: Blackwell, 1968. Reprinted in *Ifs*, edited by William Harper, Robert Stalnaker, and Glenn Pearce. Dordrecht: Reidel, 1980.

Tarski, Alfred. "The Concept of Truth in Formalized Languages." In *Logic, Semantics, and Metamathematics*, edited by John Corcoran. Indianapolis: Hackett, 1983.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Blackwell, 1953.

Wittgenstein, Ludwig. "Some Remarks on Logical Form." Proceedings of the *Aristotelian Society* supp. Vol. 9 (1929): 162–171.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. Translated by C. K. Ogden. London: Routledge & Kegan Paul, 1922.

Scott Soames (2005)

ANALYTIC FEMINISM

Analytic feminism applies analytic concepts and methods to feminist issues and applies feminist concepts and insights to issues that traditionally have been of interest to analytic philosophers. Analytic feminists, like analytic philosophers more generally, value clarity and precision in argument and use logical and linguistic analysis to help them achieve that clarity and precision. Unlike nonfeminists, they write against a background of recognition of sexism (practices that take women and feminine things to be inferior to men and masculine things) and androcentrism (practices that take males or men or men's life experiences to be the norm or the ideal for human life), and work with the aim of contesting both.

Analytic feminism holds that the best way for scholars to counter sexism and androcentrism in their work is through forming a clear conception of and pursuing truth, logical consistency, objectivity, rationality, justice, and the good, while recognizing that these notions have often been perverted by androcentrism throughout the history of philosophy. Analytic feminists engage the literature traditionally thought of as analytic philosophy, but also draw on other traditions in philosophy, as well as work by feminists working in other disciplines, especially the social and biological sciences.

Analytic feminists assert the sex/gender distinction, a distinction between the biological concept of sex and the socially constructed concept of gender (non-isomorphic to sex), though they may disagree widely on how this distinction is to be drawn and what moral or political implications it has. Although they share the conviction that the social constructions of gender create a fundamentally unjust imbalance in contemporary social and political arrangements, there is no other political thesis generally held by them. Analytic feminists who are political philosophers defend political views that reflect progressive positions found in contemporary nonfeminist political philosophy, from liberalism (Okin 1989, Nussbaum 1999) to republicanism (Phillips 2000) to socialism (MacKinnon 1989, Ferguson 1991). They also draw on views of previous generations of feminist political philosophers from John Stuart Mill and Mary Wollstonecraft to Friederich Engels, Emma Goldman, Charlotte Perkins Gilman, and Simone de Beauvoir. Analytic feminists, like nonanalytic feminists, have written much about social and political issues like abortion, pornography, prostitution, rape, sexual harassment, surrogacy, and violence against women. What characterizes analytic feminism here is the use of logical and conceptual analysis and, sometimes, decision theoretic analysis (see article by Cudd in Antony and Witt 2001).

Analytic feminists often defend traditional analytic methods and concepts against criticism from nonanalytic feminists. Many nonanalytic feminists charge (in various ways) that the notions of reason, truth, objectivity, or the methods of logical and linguistic analysis are hopelessly masculinist, and cannot be reclaimed for feminist purposes. They criticize canonical male philosophers, including Aristotle, Descartes, Kant, Rousseau, Frege, Quine, and Rawls, as sexist or at least androcentric, and at times suggest that these philosophers have nothing useful to say to women. These charges challenge feminist philosophers who have been trained in the analytic tradition and who find that tradition valuable. To reject philosophers on those grounds, they argue, would indict similarly almost the entire history of philosophy. The question analytic feminists ask is whether those androcentric or sexist writings can be corrected and rescued by an enlightened critical reader. Annette Baier's work on Hume in "Hume, the Women's Moral Theorist?" and "Hume, the Reflective Women's Epistemologist?" (Baier 1994), Marcia Homiak's work on Aristotle in "Feminism and Aristotle's Rational Ideal" (Antony and Witt 2001), Barbara Herman's work on Kant in "Could It Be Worth Thinking about Kant on Sex and Marriage?" (Antony and Witt 2001), and Peg O'Connor's work on Wittgenstein in *Oppression and*

Responsibility: A Wittgensteinian Approach to Social Practices and Moral Theory (2002) exemplify such attempts.

An important insight of feminism has been to expose the androcentric bias toward seeing human individuals as essentially isolated, epistemically, socially, and morally, from others. One early result of this insight was the ethics of care (Held 1995), which challenges the dominant tradition of ethical theory with the idea that caring for others is a central ethical activity. Eva Kittay developed the "dependency critique" (Kittay 1999) of Rawls's theory of justice, arguing that the capacity for caring for dependent others is one of the central moral powers, just as basic as the capacities to form a sense of the good and a sense of justice. Analytic feminists have joined other feminist theorists in focusing much of their recent attentions to questions of the self. In the anthology *Relational Autonomy: Feminist Perspectives on Automony, Agency, and the Social Self* (2000), several articles examine the notion of relational autonomy, which takes seriously the idea that humans must define their identities in relation to others in ways which challenge their ability to be completely autonomous in the traditional sense. These articles attempt to define a new notion of autonomy that incorporates that insight. Another important book on the self (Brison 2002) connects traditional theories of personal identity with recent research on trauma, arguing that the trauma arising from sexual violence, for example, challenges those theories.

Analytic feminism holds that many traditional philosophical notions are not only normatively compelling, but also in some ways empowering and liberating for women. While postmodern feminism rejects the universality of truth, justice, and objectivity and the univocality of "women," analytic feminism defends these notions. They recognize that to reject a view because it is false or oppressive to women, one needs some rational, objective ground from which we can argue that it is in fact false or oppressive. An important task for analytic feminism involves investigating the objectivity of science. Helen Longino's *Science as Social Knowledge* (1990) was the first such analytic feminist work. Elizabeth Anderson's "Feminist Epistemology: An Interpretation and a Defense" (Anderson 1995) shows how a carefully aimed feminist critique can improve the objectivity of science by distinguishing and illustrating four ways that feminist critiques have corrected the distorted lenses of masculinist science: through the critique of gendered structures in the social organization of science, through the analysis of gendered symbols in scientific models, through exposing sexism in scientific practices and focuses, and through revealing androcentrism in its concepts and theories.

Louise Antony, in "Quine as Feminist: The Radical Import of Naturalized Epistemology" (Antony and Witt 2001) presented what she called the *bias paradox*: Feminists (and others) want to criticize certain claims as false because they are biased, and yet feminism is also clearly a bias; in effect, a particular slant on the world. She locates a solution in naturalized epistemology. First we must see that what we can know necessarily comes through our particular human cognitive apparatus, which *biases* the content of our claims. Thus, bias *per se* is not the problem, but some biases lead us away from the truth. Her more recent work has emphasized the importance of embodiment generally in epistemology (Antony 2002), and she credits feminism in large part for this insight. Other analytic feminists (Grasswick and Webb 2002) have extended the naturalized epistemology analysis to argue for a social feminist epistemology, which asserts that socially induced sexist and androcentric biases can affect the content and justification of knowledge. In its analysis of traditional philosophical topics like objectivity and personal identity and new topics such as sexism in language (Vetterling-Braggin 1981), analytic feminism reveals the blurriness of the distinction between metaphysics, epistemology, and social/political philosophy.

See also Feminist Epistemology; Feminist Ethics; Feminist Metaphysics; Feminist Philosophy; Feminist Philosophy of Science.

Bibliography

Anderson, Elizabeth. "Feminist Epistemology: An Interpretation and a Defense." *Hypatia* 10 (1995): 50–84.

Antony, Louise M. "Embodiment and Epistemology." In *The Oxford Handbook of Epistemology*, edited by Paul K. Moser. New York: Oxford University Press, 2002.

Antony, Louise M., and Charlotte Witt, eds. *A Mind of One's Own: Feminist Essays on Reason and Objectivity*. 2nd ed. Boulder, CO: Westview, 2001.

Baier, Annette C. *Moral Prejudices*. Cambridge, MA: Harvard University Press, 1994.

Brison, Susan. *Aftermath: Violence and the Remaking of a Self*. Princeton, NJ: Princeton University Press, 2002.

Cudd, Ann E., and Virginia Klenk, eds. *Hypatia* 10 (1995). This is a special issue of the feminist philosophy journal devoted to analytic feminism.

Ferguson, Ann. *Sexual Democracy: Women, Oppression, and Revolution*. Boulder, CO: Westview, 1991.

Grasswick, Heidi E., and Mark Owen Webb. "Feminist Epistemology as Social Epistemology." *Social Epistemology* 16 (2002): 185–196.

Haslanger, Sally. "Gender and Race: (What) Are They? (What) Do We Want Them to Be?" *Nous* 34 (2000): 31–55.

Held, Virginia, ed. *Justice and Care: Essential Readings in Feminist Ethics*. Boulder, CO: Westview, 1995.

Kittay, Eva Feder. *Love's Labor: Essays on Women, Equality, and Dependency*. New York: Routledge, 1999.

Longino, Helen. *Science as Social Knowledge*. Princeton, NJ: Princeton University Press, 1990.

Mackenzie, Catriona, and Natalie Stoljar, eds. *Relational Autonomy: Feminist Perspectives on Automony, Agency, and the Social Self*. New York: Oxford University Press, 2000.

MacKinnon, Catharine A. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press, 1989.

Nelson, Lynn Hankinson. *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia, PA: Temple University Press, 1990.

Nussbaum, Martha. *Sex and Social Justice*. New York: Oxford University Press, 1999.

O'Connor, Peg. *Oppression and Responsibility: A Wittgensteinian Approach to Social Practices and Moral Theory*. University Park: Pennsylvania State University Press, 2002.

Okin, Susan Moller. *Justice, Gender, and the Family*. Boston: Basic Books, 1989.

Phillips, Anne. "Feminism and Republicanism: Is This a Plausible Alliance?" *Journal of Political Philosophy* 8 (2000): 279–293.

Vetterling-Braggin, Mary, ed. *Sexist Language: A Modern Philosophical Analysis*. Lanham, MD: Rowman and Littlefield, 1981.

Ann E. Cudd (1996, 2005)

ANALYTIC AND SYNTHETIC STATEMENTS

The distinction between analytic and synthetic judgments was first made by Immanuel Kant in the introduction to his *Critique of Pure Reason.* According to him, all judgments could be exhaustively divided into these two kinds. The subject of both kinds of judgment was taken to be some thing or things, not concepts. Synthetic judgments are informative; they tell something about the subject by connecting or synthesizing two different concepts under which the subject is subsumed. Analytic judgments are uninformative; they serve merely to elucidate or analyze the concept under which the subject falls. There is a prima facie difficulty as to how a judgment can be simultaneously about an object, uninformative in relation to it, and explicative of the concepts involved, but this question will be examined later.

Kant associated this distinction with the distinction between a priori and a posteriori judgments. The one distinction was taken to cut across the other, except that there are no analytic a posteriori judgments. The remaining three classifications were, in Kant's opinion, filled; there are analytic a priori judgments, synthetic a posteriori judgments, and synthetic a priori judgments. Since Kant there has been little argument concerning the first two of these, but considerable argument and opposition, chiefly from empiricists, about the last. Analytic a priori and synthetic a posteriori judgments correspond roughly to logically and empirically true or false judgments. In distinguishing them, Kant was following in the steps of Gottfried Wilhelm Leibniz and David Hume, both of whom had made a similar distinction, although in different terms. Leibniz had distinguished between truths of fact, guaranteed by the principle of sufficient reason, and truths of reason, guaranteed by the principle of contradiction. The latter were such that their denial involved a contradiction; they could indeed be reduced to identical propositions via chains of definitions of their terms. Hume had likewise distinguished between matters of fact and relations of ideas. The former were merely contingent, while the latter were necessary and such that their denial involved a contradiction. Kant's innovation was to connect this distinction with the two further distinctions between the analytic and the synthetic and the a priori and the a posteriori.

It should be noted that Kant's distinction between the analytic and the synthetic was made in terms of judgments and concepts. This gave it a psychological flavor for which it has been criticized by many modern philosophers. The notion of judgment is ambiguous between the act of judging and what is judged. One problem is how to extend what Kant said so that it applies only to what is judged or to propositions. Furthermore, an implication of Kant's formal account of the distinction was that it is limited in its application to subject-predicate judgments (although it was also one of Kant's doctrines that existential judgments are always synthetic).

KANT'S CRITERIA AND USE OF THE ANALYTIC/SYNTHETIC DISTINCTION

CRITERIA. Apart from the general distinction, Kant offered two criteria for it. According to the first criterion, an analytic judgment is one in which the concept of the predicate is contained (although covertly) in the concept of the subject, while in a synthetic judgment the concept of the predicate stands outside the concept of the subject. According to the second criterion, analytic judgments are such that their denial involves a contradiction, while this is not true of synthetic judgments of any kind. Kant was here following his predecessors, although, with Leibniz, he did not suggest that analytic truths can be reduced to simple identities. This criterion can scarcely be said to

suffice as a definition of an analytic statement, although it may provide grounds for saying whether a judgment is analytic or not. It will do the latter if it can be assumed that all analytic judgments are logically necessary, since reference to the principle of contradiction may provide the basis of logical necessity.

The first criterion seems on firmer ground in this respect, since it offers what seems to be a formal characteristic of all analytic judgments. It specifies what we must be doing in making an analytic judgment, in terms of the relations between the concepts involved. It has been objected that the idea of one concept being contained in another is also a psychological one, but this was certainly not Kant's intention. The point may perhaps be expressed in terms of meaning. When we make an analytic judgment, what we mean when we invoke the predicate concept is already included in what we mean by the subject concept. Just as the notion of a judgment is ambiguous, so a concept can mean either the act of conceiving or what is conceived, and it is the latter which is relevant here. By this criterion, therefore, a judgment is analytic when, in judging about something, what we judge about it is already included in what is meant by the term under which we subsume the subject. Kant assumed that all judgments of this kind are a priori, presumably on the grounds that their truth can be ascertained merely by considering the concepts involved, without further reference to the facts of experience.

CHARACTERISTICS OF ANALYTIC STATEMENTS.

Kant's criterion could be applied only to statements of subject-predicate form, and could not, therefore, be used to make an exhaustive distinction between all statements. If Kant's distinction is to be of use, however, it must be extended to cover propositions or statements and, moreover, statements of any form, not just those of subject-predicate form. If an analytic judgment is of an object, an analytic statement must similarly be about the object or objects referred to by the subject expression. Analytic statements cannot, therefore, be equated with definitions, for the latter are surely about words, not things. It has sometimes been said (for instance, by A. J. Ayer in his *Language, Truth and Logic*) that analytic statements make clear our determination to use words in a certain way. Apart from the fact that the use of words cannot be a simple matter of choice, what Ayer says cannot be the main function of analytic statements, since this would be to identify them with (possibly prescriptive) definitions. If we learn something about the use of words from analytic statements, this must at most be indirect.

Analyticity, a property of statements. We have seen that Kant's point of view might be represented as saying that only the meaning of the terms involved, the nature of the corresponding concepts, makes the judgment true. It might, therefore, seem feasible that an analytic statement could be characterized as a statement about something which says nothing about the thing but is such that the meanings of the words involved make it true. To be more exact, it would be the meanings of the words involved in a sentence—*any* sentence that expresses the statement—that make that statement true. It is important to stress the words "any sentence," for analytic truth can be a feature only of *statements.* It cannot be a feature of sentences per se, nor can it be limited to sentences in a given language (as Rudolf Carnap in effect supposes). Truth is a property of statements, not sentences, and the same must be the case with analytic truth. No account of analyticity which explains it in terms of what is the case with regard to sentences in any one language will do. If someone who says "All bodies are extended" makes an analytic statement, so will anyone who says the same thing in any other language.

Analyticity as a function of the meanings of words. What is meant by saying that the meanings of the terms involved *make a statement true?* Are analytic truths those which follow from the meanings of the words involved; that is, from their definitions? This cannot be so, since all that can follow from a definition is another definition, and how, in any case, can a statement about things follow directly from one about words? If analyticity is connected with meaning, it must be more indirectly. Friedrich Waismann has suggested that an analytic truth is one which is so *in virtue of* the meanings of the words involved. But the words "in virtue of" are themselves vague. It has been held by certain empiricists that "All bodies are extended" is analytic if and only if we use "body" in exactly the same way we use "extended thing"; that is, if we attach the same meaning to each expression. Nevertheless, the truth of "All bodies are extended" does not follow simply from the fact that the expressions "body" and "extended thing" have the same meaning, for the substitution of expressions equivalent in meaning leaves one with a statement corresponding in form to the law of identity. Hence, the original statement will be true only if the law of identity holds. In other words, an analytic statement will be one whose truth depends not only on the meanings of the words involved but also on the laws of logic. This raises the question of the status of these laws themselves. It is sometimes claimed that they, too, are analytic; but this cannot be so if a definition of analyticity involves reference to the laws of logic.

Analyticity as a function of the laws of logic. The necessity of referring to the laws of logic in any account of analyticity has been noted in modern times by many philosophers. Waismann, for example, eventually defines an analytic statement as one which reduces to a logical truism when substitution of definitional equivalents is carried out. Gottlob Frege had much earlier defined an analytic truth as one in whose *proof* one finds only "general logical laws and definitions," and he had sought to show that arithmetical propositions are analytic in this sense. Both of these accounts make reference to logical truisms or logical laws. Whatever the status of these, it certainly seems that analytic statements must depend for their validity not only on the meanings of the terms involved but also on the validity of the laws of logic; and these laws cannot themselves be analytic.

OBJECTIONS TO THE DISTINCTION

THE PROBLEM OF SYNONYMY. Nevertheless, objections to the notion of analyticity have been made, particularly by Willard Quine, on the basis of supposed difficulties about meaning itself, and not merely on those about the status of the truths of logic—although here, too, Quine has found difficulties. He distinguishes between two classes of analytic statements. There are, first, those which are logically true, such as "No unmarried man is married"; these are statements which are true and which remain true under all reinterpretations of their components other than the logical particles. Second, there are those, such as "No bachelor is married," which can be turned into logical truths by substituting synonyms for synonyms. It is the second kind of analytic statement that raises problems here, and these problems arise from the notion of synonymy or, to be precise, cognitive synonymy; that is, synonymy that depends on words having the same meaning for thought, as opposed to merely applying to the same things. The notion of definition which other philosophers have invoked in this connection rests, Quine maintains, on that of synonymy. How is this to be explained?

Quine's difficulties here are associated with general difficulties about synonymy raised by himself and Nelson Goodman in the effort to embrace a nominalism that does not involve the postulation of so-called meanings, and to push as far as possible the thesis that language is extensional; that is, such that it can be built up from variables and an indefinite set of one and many-place predicates, so that complex sentences are related to atomic sentences by truth-functional relationships and by quantification. In such a language, sameness of meaning might be equivalent to extensional equivalence, such that any two extensionally equivalent expressions are interchangeable *salva veritate*; that is, leaving unchanged the truth value of the statements in which they occur, wherever the expressions occur. The outcome of Goodman's argument in this connection is that since there may always be some occurrence in which the two expressions are not interchangeable *salva veritate,* no two expressions are identical in meaning. Quine himself recognizes something of this and has explored the restrictions which must be put upon the general thesis.

In the present connection, Quine explores the possibility that synonymity might be explained by interchangeability *salva veritate* except within words. But the interchangeability of, say, "bachelor" and "unmarried man" in this way may be due to accidental factors, as is the case with "creature with a heart" and "creature with kidneys." If it is the case that all—and only—creatures with a heart are creatures with kidneys, this is due simply to the fact that, as it happens, the two expressions always apply to the same things and not to any sameness of meaning. How do we know that the situation is not the same with "bachelor" and "unmarried man"? It is impossible to reply that it is because of the truth of "Necessarily, all—and only—bachelors are unmarried men," for the use of "necessarily" presupposes a nonextensional language. Furthermore, a sense has already been given to the kind of necessity involved here: analyticity. Hence, while cognitive synonymy might be explained in terms of analyticity, to try to explain analyticity in terms of cognitive synonymy would involve something like circularity.

Quine argues that similar considerations apply to attempts, such as Carnap's, to deal with the matter in terms of a semantic rule. Quine then considers the further possibility that, given that the truth of statements in general rests upon a linguistic component and a factual component, an analytic statement might be one in which the factual component is null. This, while apparently reasonable, has not, he objects, been explained; and the attempt by positivists to do so by reference to the verification theory of meaning (with its assumption that there are basic propositions in which the factual component is all that matters and, on the other hand, that there are analytic propositions in which the linguistic component is all that matters) involves reductionism, an unjustified dogma.

Synonymy and meaning. A possible objection to Quine—one in effect made by H. P. Grice and P. F. Strawson—is that his difficulty over synonymy involves a refusal to understand. There is a family of terms that

includes analyticity, necessity, and cognitive synonymy, and Quine will not accept, as explanations of any one of them, accounts that involve reference to other members of the family. On the other hand, to go outside the family in one's explanations, as is involved in having recourse to extensional equivalence, is necessarily an inadequate explanation. This is a situation that frequently occurs in philosophy, wherever one is confronted with families of terms between which and any other family there is a radical or categorical distinction. This is perhaps an oversimplification of the situation, true though it is. It must be remembered that Quine's basic urge is to do without meanings, so as not to introduce unnecessary entities into our ontology. The failure of this particular enterprise of defining synonymy is, however, in fact, a demonstration of its futility. Meaning is a notion which must be presupposed rather than explained away in this connection.

THE BOUNDARY BETWEEN ANALYTIC AND SYNTHETIC STATEMENTS. Quine also has a second thesis in connection with analyticity, a thesis that has been echoed in different forms by other philosophers. It is a quite general thesis, in the sense that it does not depend on considerations about synonymy and is not, therefore, restricted to statements whose truth turns on synonymy. This thesis states that even if a distinction could be drawn between analytic and synthetic statements or between logical and factual truth, it is impossible to draw a sharp boundary between them. The contrary supposal rests on the dogma of reductionism already referred to. On that thesis, there is clearly an absolute distinction to be made. The denial of the dogma entails that there can be, at the most, a relative distinction. Within any particular system it is possible to distinguish those statements, those of logic and mathematics, which we should be extremely reluctant to give up and those, on the other hand, which we should be ready to give up if required to do so. The former are entrenched because of their close connections with other elements of the system. It has often been pointed out that the giving up of some high-grade scientific statements would involve the giving up with them of whole scientific systems. On Quine's view, the situation is worse, but not intrinsically different, with logical statements. There are no statements that depend for their truth on a direct confrontation with experience. The best that can be produced in the way of a distinction between different kinds of statements is a relative distinction between those which are more or less entrenched. No absolute and sharp distinction between analytic and synthetic statements can be drawn. Quine's conventionalism here reflects pragmatist tendencies.

One possible reply to this thesis is that the rejection of the dogma of reductionism does not by itself dispose of an absolute distinction of this kind. Even if it is accepted that there are no statements in which the factual component is everything, it does not follow that there are no statements in which the linguistic component is everything. Despite what Quine says, the thesis that there is a distinction between analytic and synthetic statements is independent of that of reductionism. Grice and Strawson have also attempted to deal with the issue by making a distinction in terms of the responses to attempts to falsify a statement. Analytic statements are those that, in a falsifying situation, demand a revision in our concepts; synthetic statements are those that demand a revision in our view of the facts. It has frequently been pointed out that it is possible to preserve a scientific statement against falsifying circumstances by making it logically true and thus immune to falsification. In doing this, we revise our concepts but not our view of the facts. It is clear that Quine could not accept this suggestion as such, since it presupposes that an answer has been given to the first of his problems—the definition of analyticity—in terms of notions like those of a concept or meaning. But, given that Quine's thesis is untenable in this first respect, there is no reason for denying its untenability in the second.

STATEMENTS THAT ARE NEITHER ANALYTIC NOR SYNTHETIC. Other reasons for dissatisfaction with a sharp distinction between analytic and synthetic statements have been offered by other philosophers. Waismann, for example, has maintained that there are some statements which do not admit of a clear classification; for instance, "I see with my eyes." In this case there are reasons for saying that it is analytic, since whatever I see with might be called "eyes"; on the other hand, it might be said that it is a matter of fact that it is with my eyes that I see. Hence, Waismann maintains, such statements are neither analytic nor synthetic, properly speaking. The objection to this, as has been pointed out by W. H. Walsh, is that Waismann has failed to consider the contexts in which such statements are made. The sentence "I see with my eyes" may be used in one context to express an analytic statement and in another to express a synthetic one. The fact that the same sentence may have different uses and that the analyticity or syntheticity of a statement is a function of those uses (a statement is just the use of a sentence) shows nothing about the necessity of abandoning the analytic-synthetic distinction.

ARE THERE ANY ANALYTIC STATEMENTS? Emphasis of the point that analyticity is a function of use prompts the question of whether sentences which purport to express analytic statements have a use at all and whether, in consequence, there *are* any analytic statements. It has been emphasized from Kant onward that analytic statements are trivial, and similar things were said even before Kant—by John Locke, for instance. The truth of an analytic statement makes no difference to the world. It is, therefore, difficult to see why anyone should ever make an analytic statement. A possible reply is that such a statement might be made in order to clarify something about the concepts involved. If the statements in question are about concepts, however, rather than about the thing or things referred to by the subject expression, why are they not simply definitions? Definitions are not in themselves analytic statements, whatever their exact status. It could thus be argued that any statement which has a use either provides information about things or about the meanings of words, and in either case the statement would be synthetic, or at least not analytic. The only viable function remaining for the term *analytic* would be as a term of logical appraisal, not as a classificatory expression. That is to say, the use of the words "That is analytic" would not be to classify the statement in question, but to say, in effect, "You have not said anything."

Whether or not this is plausible in itself, the crucial question remains: How is it possible for a statement both to be about something and to elucidate the concepts involved? (The question is probably more crucial for judgments than for statements, since it might seem obvious what a judgment must be about, while the criteria of "aboutness" are less obvious in the case of statements.) The issues are simple. A statement is one use of a sentence, and an analytic statement is such a use that conforms to certain conditions—two of which are that it says nothing about its subject and that its truth depends at least in part on the meanings of the words involved. If this is so, it cannot be used to make clear those meanings. If an analytic statement does serve to make clear those meanings to someone, this must be an incidental and unintended consequence of its use, not an essential part of that use. On the other hand, if the triviality of analytic statements is accepted, there can be no argument to show that their use is impossible, for there is no reason why a statement, if it is to be about something, should also *say* something about that thing. The use of such statements would simply lack point.

A POSSIBLE WAY OF MAKING THE DISTINCTION

Ludwig Josef Johann Wittgenstein pointed out in the *Tractatus Logico-Philosophicus* (4.4611) that tautologies are senseless but not nonsense. By "senseless" he meant that they do not pick out any determinate state of affairs that makes a difference to our view of the world. They are, in effect, trivial. They are not nonsense, however, because they are part of our symbolism, just as "0" is part of the symbolism of arithmetic, although it is useless for counting. Given a system of symbolism, or a language, it must always be possible to construct sentences that could be used to assert analytic truths or falsehoods (contradictions), whether or not there would be any point in doing so. This possibility is a necessary consequence of the nature of language. A language, however, is not just a system of symbols; it is something whose function is, among other things, to state and communicate facts. Hence, it is possible to say that, given that these sentences have a use, the truth of their uses (or, in the case of contradictions, their falsity)—that is, the truth of the relevant statements—is a necessary condition of the employment of the language from which the corresponding sentences are drawn, or of any language in which there are sentences with the same meaning. More briefly, analytic statements will be those whose truth is necessary to the employment, as expressed in language, of the system of concepts on which they depend. Any statement of which this is not true will be synthetic. Of these other statements, many will be such that their truth is not necessary in any way, but there may be others whose truth is necessary in some way other than that of analytic statements—as Kant maintained about the synthetic a priori.

See also A Priori and A Posteriori; Ayer, Alfred Jules; Grice, Herbert Paul; Hume, David; Kant, Immanuel; Locke, John; Quine, Willard Van Orman; Strawson, Peter Frederick; Wittgenstein, Ludwig Josef Johann.

Bibliography

BOOKS

Ayer, A. J. *Language, Truth and Logic,* 2nd ed. London: V. Gollancz, 1946. Ch. 4.

Carnap, Rudolf. *Introduction to Semantics.* Cambridge, MA: Harvard University Press, 1942.

Frege, Gottlob. *Foundations of Arithmetic.* Translated by J. L. Austin. Oxford: Blackwell, 1950.

Hume, David. *Treatise of Human Nature.* Oxford: Clarendon Press, 1896. I.iii.1.

Kant, Immanuel. *Critique of Pure Reason.* Translated by Norman Kemp-Smith. London, 1953. Especially the introduction.

Leibniz, G. W., throughout his works. Especially *Monadology,* 31ff.
Pap, Arthur. *Semantics and Necessary Truth.* New Haven, CT: Yale University Press, 1958. Useful discussions of general issues.
Quine, W. V. *From a Logical Point of View.* Cambridge, MA, 1953. Ch. 2.

ARTICLES
Goodman, Nelson. "On Likeness of Meaning." *Analysis* 10 (1) (October 1949): 1–7.
Grice, H. P., and P. F. Strawson. "On Defence of a Dogma." *Philosophical Review* 65 (March 1956): 141–158.
Hamlyn, D. W. "Analytic Truths." *Mind,* n.s., 65 (259) (1956): 359–367.
Hamlyn, D. W. "On Necessary Truth." *Mind,* n.s., 70 (280) (1961): 514–525.
Putnam, Hilary. "The Analytic and the Synthetic" In *Minnesota Studies in the Philosophy of Science.* Minneapolis, 1962. Vol. III, pp. 358–397.
Waismann, Friedrich. "Analytic and Synthetic." *Analysis* 10, 11, and 13 (1949–1952). A series of articles.
Walsh, W. H. "Analytic and Synthetic." *PAS* 54 (1953/1954): 77–96.
White, M. "Analytic-Synthetic: An Untenable Dualism." In *John Dewey: Philosopher of Science and Freedom,* edited by Sydney Hook, pp. 316–330. New York, 1950.

D. W. Hamlyn (1967)

ANALYTIC AND SYNTHETIC STATEMENTS [ADDENDUM]

There are several major philosophical projects that having a viable analytic/synthetic distinction would advance. For example: Analytic (true) sentences are supposed to have their truth values solely in virtue of the meanings (together with the syntactic arrangement) of their constituents; in effect, their truth values are supposed to supervene on their linguistic properties alone (Quine 1953). So they are true in every possible world where they mean what they mean here. So they are necessarily true. So if there were a viable analytic/synthetic distinction, we would understand the necessity of at least some necessary truths. If, in particular, it were to turn out that the logical and/or the mathematical truths are analytic, we would understand why *they* are necessary (Gibson 1988, Quine 1998).

An account of necessity according to which necessary truths are analytic has special virtues. Necessity is not, of course, an epistemic notion. Still, suppose that the necessity of a sentence arises from the meanings of its parts. It is natural to assume that one of the things one knows in virtue of knowing one's language is what the expressions of the language mean (Boghossian 1994). A treatment of modality in terms of analyticity therefore connects the concept of necessity with the concept of knowledge; and knowledge *is*, of course, an epistemic property. So if there is a viable analytic/synthetic distinction, we could explain why the necessary truths, or at least some of the necessary truths, are knowable a priori by anybody who knows a language that can express them (Quine 1991). It bears emphasis that not every theory of necessity yields a corresponding treatment of apriority; doing so is a special virtue of connecting modality with meaning.

Many philosophers interested in the metaphysics of semantical properties find attractive the idea that the meaning of an expression supervenes on its conceptual/inferential role (Sellars 1954, Harman 1987, Block 1994 and references therein). It is, however, a plausible objection to conceptual role semantics that it courts a ruinous holism unless there is some way to distinguish meaning-constitutive inferences from the rest (Fodor and Lepore 1991, 1992). A tenable analytic/synthetic distinction might resolve this tension; the meaning constitutive-inferences could be identified with the analytic ones. In practice, it is pretty widely agreed that saving the analytic/synthetic distinction is a condition for saving conceptual role semantics (Block 1994, Peacocke 1992).

For many linguists, it is a main goal of "lexical semantics" to predict which sentences express them; typically, by "decomposing" the meanings of some words into their definitions. On this sort of view, intuitions of analyticity play much the role vis-à-vis theories of meaning that intuitions of grammaticality do vis-à-vis theories of syntax (Katz 1972).

For all of the aforementioned points, many philosophers have been persuaded (largely by considerations Quine raised) that there is no unquestion-begging way to formulate a serious analytic/synthetic distinction (Gibson 1988, Harman 1999, Lepore 1995). The moral might be that philosophy will have to do without it, even if, in consequence, notions like necessity, apriority and definition seem deeply mysterious.

Harman (following Quine) has famously offered an across-the-board argument that the notion of analyticity is untenable; namely, that the truth of analytic sentences is supposed somehow to be independent of "how the world is," but it is puzzling how the truth of anything could be independent of how the world is. How, for example, could a stipulation, or a linguistic convention (implicit or otherwise) make a proposition *true*? How

could our undertaking to respect the inference from "bachelor" to "unmarried" make it true that bachelors are unmarried?

There is an obvious problem in understanding how the truth of a statement can be independent of the way the world is and depend entirely on the meaning of the statement. Why is it not a fact about the world that copper is a metal such that, if this were not a fact, the statement "copper is a metal" would not express a truth? And why doesn't the truth expressed by "copper is copper" depend in part on the general fact that everything is self-identical (Harman 1999)?

Boghossian (1996) holds that a sentence can be made true by stipulation; and that that stipulation determines which proposition the sentence expresses. Call the sentence S and the proposition P. Surely, if S is true, P is true, since it is a truism (assuming sentences have truth values at all) that each sentence has the same truth-value as the proposition it expresses. It is thus unclear why making a sentence true by stipulation is not *thereby* making the corresponding proposition true by stipulation.

Still, whatever the truth maker for a proposition is, the proposition is true just in case its truth maker is "in place." Now consider the proposition expressed by a sentence that is true by stipulation. Presumably, the truth maker for that proposition *must be* in place since the sentence that expresses it is true. If so, then, Harman can object as follows: "It's not obvious how a stipulation could make the world such that a certain sentence is true of it. But it's also, and equally, not obvious how a stipulation could guarantee that the truth maker of the proposition that a sentence expresses is 'in place.'" In fact, the second question is plausibly just the first one all over again.

See also Analysis, Philosophical; Analyticity; Meaning; Quine, Willard Van Orman; Synonymity.

Bibliography

Block, N. "Advertisement for a Semantics for Psychology." In *Mental Representation: A Reader*, edited by Stephen Stich and Ted A. Warfield. Cambridge, MA: Blackwell, 1994.

Boghossian, P. "Analyticity Reconsidered." *Nous* 30 (3) (1996): 360–391.

Boghossian, P. "The Transparency of Mental Content." *Philosophical Perspectives* 8 (1994): 33–50.

Fodor, J., and E. Lepore. *Holism: A Shopper's Guide*. Oxford: Basil Blackwell, 1992.

Fodor, J., and E. Lepore. "Why Meaning Probably Isn't Conceptual Role." *Mind and Language* 6 (4) (1991): 329–343. Reprinted in *The Compositionality Papers*, edited by J. Fodor and E. Lepore. Oxford and New York: Clarendon Press, 2002.

Gibson, Roger. *Enlightened Empiricism: An Examination of W. V. Quine's Theory of Knowledge*. Tampa: University of South Florida Press, 1988.

Harman, G. "The Death of Meaning." In *Reasoning, Meaning, and Mind*, edited by G. Harman, 119–137. Oxford: Clarendon Press, 1999.

Harman, G. "(Non-Solipsistic) Conceptual Role Semantics." In *New Directions in Semantics*, edited by E. Lepore. London: Academic Press, 1987.

Katz, Jerrold. *Semantic Theory*. New York: Harper and Row, 1972.

Lepore, Ernest. "Two Dogmas of Empiricism and the Generality Requirement." *Nous* 24 (1995): 468–480.

Peacocke, C. *A Study of Concepts*. Cambridge, MA: MIT Press, 1992.

Quine, W. V. O. "Reply to Roger F. Gibson, Jr." In *The Philosophy of W. V. Quine*, edited by Lewis Edwin Hahn and Paul A. Schilpp, 684–685. Chicago: Open Court, 1998.

Quine, W. V. O. "Two Dogmas in Retrospect." In *Dear Carnap, Dear Van: The Quine Carnap Correspondence and Related Work*, edited by Richard Creath. Berkeley: University of California Press, 1991.

Quine, W. V. O. "Two Dogmas of Empiricism." In *From a Logical Point of View*, 20–46. Cambridge, MA: Harvard University Press, 1953.

Sellars, W. "Some Reflections on Language Games." *Philosophy of Science* 21 (1954): 204–228.

Ernest Lepore (2005)

ANALYTICITY

The idea of "analyticity"—or truth by virtue of meaning—can be understood in two different ways. On the one hand, it might stand for an epistemic notion, for the idea that mere grasp of the meaning of a sentence suffices for knowledge that it is true. On the other hand, it might stand for a metaphysical notion, for the idea that a statement owes its truth value completely to its meaning, and not at all to "the facts." We may call the first notion "epistemic analyticity" and the second "metaphysical analyticity." On the face of it, these are distinct notions that subserve distinct philosophical programs. Willard Van Orman Quine, whose writings are largely responsible for the contemporary rejection of analyticity, failed to distinguish between them; as a result, many philosophers came to assume that the two notions stand or fall together. However, it is the moral of recent work in this area that this assumption is mistaken: epistemic analyticity can be defended even while its metaphysical cousin is rejected.

The metaphysical concept of analyticity is presupposed by the logical positivist program of reducing all necessity to linguistic necessity. Guided by both the fear

that objective, language-independent necessary connections would be metaphysically odd, and that no empiricist epistemology could explain our knowledge of them, philosophers like Rudolf Carnap (1947) and A. J. Ayer (1946) attempted to show that all necessary truths are simply disguised decisions concerning the meanings of words. According to this view, there is no more to the truth of, say, "Either snow is white or it is not" than a decision concerning the meaning of the word "or." On this view, linguistic meaning by itself is supposed to generate necessary truth; a fortiori, linguistic meaning by itself is supposed to generate truth. Hence the play with the metaphysical notion of analyticity.

However, it is doubtful that this makes a lot of sense. What could it possibly mean to say that the truth of a statement is fixed exclusively by its meaning and not by the facts? Is it not in general true that for any statement **S,**

S is true if and only if (iff) for some **p, S** means that **p** and **p**?

How could the mere fact that **S** means that **p** make it the case that **S** is true? Doesn't it also have to be the case that **p** (see Harman, 1960)?

The proponent of the metaphysical notion does have a comeback, one that has perhaps not been sufficiently addressed. What he will say instead is that, in some appropriate sense, our meaning **p** by **S** makes it the case that **p.**

But this line is itself fraught with difficulty. For how are we to understand how our meaning something by a sentence can make something or other the case? It is easy to understand how the fact that we mean what we do by a sentence determines whether that sentence expresses something true or false. But as Quine (1951) points out, that is just the normal dependence of truth on meaning. What is not clear is how the truth of what the sentence expresses could depend on the fact that it is expressed by that sentence, so that we would be entitled to say that what is expressed would not have been true at all had it not been for the fact that it is expressed by that sentence. But are we really to suppose that, prior to our stipulating a meaning for the sentence

"Either snow is white or it is not"

it was not the case that either snow was white or it was not? Is it not overwhelmingly obvious that this claim was true *before* such an act of meaning, and that it would have been true even if no one had thought about it, or chosen it to be expressed by one of our sentences?

There is, then, very little to recommend the linguistic theory of necessity and, with it, the metaphysical notion of analyticity that is supposed to subserve it. Epistemic analyticity, by contrast, is not involved in that futile reductive enterprise. Its role, rather, is to provide a theory of a priori knowledge.

Intuitively speaking, it does seem that we can know certain statements—the truths of logic, mathematics, and conceptual analysis, most centrally—without recourse to empirical experience. The problem has always been to explain how.

The history of philosophy has known a number of answers to this question, among which the following has been very influential: We are equipped with a special evidence-gathering faculty of intuition, distinct from the standard five senses, that allows us to arrive at justified beliefs about the necessary properties of the world. By exercising this faculty, we are able to know a priori such truths as those of mathematics and logic.

The central impetus behind the analytic explanation of the a priori is to explain the possibility of a priori knowledge without having to postulate any such special faculty of "intuition," an idea that has never been adequately elaborated.

This is where the concept of epistemic analyticity comes in. If mere grasp of **S**'s meaning by **O** were to suffice for **O**'s being justified (with a strength sufficient for knowledge—henceforth, we will take this qualification to be understood) in holding **S** true, then **S**'s apriority would be explainable without appeal to a special faculty of intuition: the very fact that it means what it does for **O** would by itself explain why **O** is justified in holding it to be true.

How could mere grasp of a sentence's meaning justify someone in holding it true? Clearly, the answer to this question has to be semantical: something about the sentence's meaning, or about the way that meaning is fixed, must explain how its truth is knowable in this special way. What could this explanation be?

In the history of the subject, two different sorts of explanation have been especially important. Although these, too, have often been conflated, it is crucial to distinguish between them.

One idea was first formulated in full generality by Gottlob Frege (1884). According to this view, a statement's epistemic analyticity is to be explained by the fact that it is transformable into a logical truth by the substitution of synonyms for synonyms. We may call

statements that satisfy this semantical condition "Frege-analytic."

Quine's enormously influential "Two Dogmas of Empiricism," (1951) complained that there could not be any Frege-analytic statements because there could not be any synonymies. But, as Herbert P. Grice and Peter F. Strawson showed (1956), the arguments for this claim are highly disputable. And Paul Boghossian (1995) has added to this by arguing that Quine's negative arguments cannot plausibly stop short of his radical thesis of the indeterminacy of meaning, a thesis that most philosophers continue to reject.

The real problem with Frege-analyticity is not that there are not any instances of it, but that it is limited in its ability to explain the full range of a priori statements. Two classes remain problematic: a priori statements that are not transformable into logical truths by the substitution of synonyms for synonyms, and a priori statements that are trivially so transformable.

An example of the first class is the sentence "Whatever is red all over is not blue." Because the ingredient descriptive terms do not decompose in the appropriate way, this sentence is not transformable into a logical truth by substitution of synonyms.

The second class of recalcitrant statements consists precisely of the truths of logic. These truths satisfy, of course, the conditions on Frege-analyticity. But they satisfy them trivially. And it seems obvious that we cannot hope to explain our entitlement to belief in the truths of logic by appealing to their analyticity in this sense: Knowledge of Frege-analyticity presupposes knowledge of logical truth and so cannot explain it.

How, then, is the epistemic analyticity of these recalcitrant truths to be explained? The solution proposed by Carnap (1947) and the middle Ludwig Wittgenstein (1974) turned on the suggestion that such statements are to be thought of as "implicit definitions" of their ingredient terms. Applied to the case of logic (a similar treatment is possible in the case of the other class of recalcitrant truths), this suggestion generates the semantical thesis we may call:

> Implicit definition: It is by arbitrarily stipulating that certain sentences of logic are to be true, or that certain inferences are to be valid, that we attach a meaning to the logical constants. A particular constant means that logical object, if any, which makes valid a specified set of sentences and/or inferences involving it.

The transition from this sort of implicit definition account of grasp to an account of the apriority of logic can then seem immediate, and the following sort of argument would appear to be in place:

1. If logical constant **C** is to mean what it does, then argument-form **A** has to be valid, for **C** means whatever logical object in fact makes **A** valid.
2. **C** means what it does.

Therefore,

3. **A** is valid.

Quine's "Truth by Convention" (1936) and "Carnap and Logical Truth" (1976) raised several important objections against the thesis of implicit definition: first, that it leads to an implausible conventionalism about logical truth; second, that it results in a vicious regress; and third, that it is committed to a notion—that of a meaning-constituting sentence or inference—that cannot be made out.

Even the proponents of implicit definition seem to have agreed that some sort of conventionalism about logical truth follows from implicit definition. However, Nathan Salmon (1994) and Boghossian (1997) have argued that this is a mistake: No version of conventionalism follows from the semantical thesis of implicit definition, provided that a distinction is observed between a sentence and the claim that it expresses.

Quine's second objection is also problematic in relying on a defective conception of what it is for a person to adopt a certain rule with respect to an expression, according to which the adoption of a rule always involves explicitly stating in linguistic terms the rule that is being adopted. On the contrary, it seems far more plausible to construe **x**'s following rule **R** with respect to **e** as consisting in some sort of fact about **x**'s behavior with **e.**

In what would such a fact consist? Here there are at least two options of which the most influential is this: **O**'s following rule **R** with respect to **e** consists in **O**'s being disposed, under appropriate circumstances, to conform to rule **R** in his employment of **e.**

According to this view, then, the logical constants mean what they do by virtue of figuring in certain inferences and/or sentences involving them and not in others. If some expressions mean what they do by virtue of figuring in certain inferences and sentences, then some inferences and sentences are constitutive of an expression's meaning what it does, and others are not.

Quine's final objection to implicit definition is that there will be no way to specify systematically the meaning-constituting inferences, because there will be no

way to distinguish systematically between a meaning constituting inference and one that is not meaning-constituting but simply obvious. However, although this is a serious challenge, and although it remains unmet, there is every reason for optimism (see, for example, Peacocke 1994 and Boghossian 1995).

Quine helped us see the vacuity of the metaphysical concept of analyticity and, with it, the futility of the project it was supposed to underwrite—the linguistic theory of necessity. But those arguments do not affect the epistemic notion of analyticity, the notion that is needed for the purposes of the theory of a priori knowledge. Indeed, the analytic theory of apriority seems to be a promising research program, given reasonable optimism about the prospects both for a conceptual role semantics and for the idea of Frege-analyticity.

See also Ayer, Alfred Jules; Carnap, Rudolf; Conventionalism; Frege, Gottlob; Grice, Herbert Paul; Knowledge, A Priori; Moral Epistemology; Quine, Willard Van Orman; Strawson, Peter Frederick; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ayer, A. J. *Language, Truth and Logic.* London: Gollancz, 1946.

Boghossian, P. A. "Analyticity." In *A Companion to the Philosophy of Language,* edited by C. Wright and B. Hale. Cambridge, U.K.: Blackwell, 1997.

Carnap, R. *Meaning and Necessity.* Chicago: University of Chicago Press, 1947.

Dummett, M. *Frege: The Philosophy of Mathematics.* Cambridge, MA: Harvard University Press, 1991.

Frege, G. *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung uber den Begriff der Zahl* (1884). Translated by J. L. Austin as *The Foundations of Arithmetic.* Oxford: Blackwell, 1950.

Grice, H. P., and P. Strawson. "In Defense of a Dogma." *Philosophical Review* 65 (1956): 141–158.

Harman, G. "Quine on Meaning and Existence I." *Review of Metaphysics* 21 (1960): 124–151.

Pap, A. *Semantics and Necessary Truth.* New Haven, CT: Yale University Press, 1958.

Peacocke, C. "How Are A Priori Truths Possible?" *European Journal of Philosophy* 1 (1993): 175–199.

Peacocke, C. *A Study of Concepts.* Cambridge, MA: MIT Press, 1992.

Putnam, H. *Mind, Language and Reality—Philosophical Papers 2.* Cambridge: Cambridge University Press, 1975.

Quine, W. V. O. "Carnap and Logical Truth" (1954). Reprinted in *The Ways of Paradox,* Cambridge, MA: Harvard University Press, 1976.

Quine, W. V. O. "Truth by Convention." In *Philosophical Essays for A. N. Whitehead,* edited by O. H. Lee. New York: Longmans, Green, 1936. Reprinted in *The Ways of Paradox,* Cambridge, MA: Harvard University Press, 1976.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* (1951). Reprinted in *From a Logical Point of View,* Cambridge, MA: Harvard University Press, 1953.

Quine, W. V. O. *Word and Object.* Cambridge, MA: MIT Press, 1960.

Salmon, N. "Analyticity and Apriority." *Philosophical Perspectives* (1994).

Wittgenstein, L. J. J. *Philosophical Grammar* (1932–1934). Berkeley: University of California Press, 1974.

Paul Artin Boghossian (1996)

ANALYTIC JURISPRUDENCE

Analytic jurisprudence divides into two related areas: substantive and methodological. Until the late 1980s most analytic jurisprudence had been substantive. It focused on producing theories of the nature of law, the relationship between laws (particular legal standards) and law (a system of governance by laws), and the relationship of law to morality and other institutions for regulating human affairs and actions.

Whereas these debates in substantive jurisprudence remain as lively and urgent as ever, analytic jurisprudence has taken a decidedly methodological turn. Jurisprudence is a philosophical theory of the nature of law, not a historical, economic, or sociological one. But how can philosophy shed light on law? The conventional answer is that philosophy aims to uncover the nature of law.

But how can philosophy help uncover the nature of law? Since H. L. A. Hart, at least, the most prominent answer has been that philosophical theories of law are theories of the concept of law, of concepts related to it (such as obligation and authority), and of the relationships among these concepts. The philosophical method of jurisprudence is conceptual analysis. Thus, analytic jurisprudence is on the same footing as analytic epistemology, metaphysics, and metaethics. Analytic jurisprudence is conceptual analysis of the concept of law, just as epistemology is conceptual analysis of the concepts of epistemic warrant and knowledge.

The standard view is that competent speakers share the concept of law, though each has an incomplete grasp of it. While we take ourselves to be employing the same concept—a concept that regulates our usage and enables us to communicate meaningfully—in fact few competent speakers have theories of the concept in all its particulars. Constructing such a theory is the task of the jurisprudent. Such a theory refines and regulates our use and aims to

deepen our understanding. Constructing such a theory begins with ordinary use, which reflects a partial understanding, but does not end there. Nor is jurisprudence merely a descriptive activity reporting on common or shared understandings.

Since Hart, at least, this has been the dominant method for approaching the study of the nature of law. Of contemporary legal philosophers, Joseph Raz is perhaps the preeminent proponent of this way of understanding conceptual analysis as a distinctive philosophical approach to law.

The place of conceptual analysis within jurisprudence has recently come under sustained attack from several quarters. These attacks have been responsible for much of the current interest in the methods of analytic jurisprudence. The first line of attack raises doubts about conceptual analysis, not just in jurisprudence, but in philosophy more generally. This is the naturalist challenge. In its stronger forms, naturalized jurisprudence argues that conceptual analysis is a form of inquiry that proceeds by culling usage and then testing various refinements and revisions against intuitions about proper use, and that it turns philosophical inquiry into an irreducible battle among competing intuitions and is ultimately hopeless.

Naturalists invite us to understand law by taking our cues from the social-scientific theories that explore the role of law in our social lives. We revise or amend those accounts only insofar as the theories fail to deliver the requisite goods: to enable us to make our way through the social world.

Within law, naturalists, Brian Leiter in particular, have focused more narrowly on the theory of adjudication. If we take authoritative legal texts as inputs and judicial decisions as outputs, then a theory of adjudication is a set of norms that takes the set of relevant authoritative texts, together with pertinent factual premises, and generates correct judicial opinions from them. A theory of adjudication is an account of warranted or justified legal inferences or decisions. The naturalist rejects the view that the norms governing proper legal reasoning can be determined by a priori reflection on our practice. Instead of trying to determine the norms by which judges ought to decide cases, they urge that we study how judges in fact decide cases. In this way, the legal naturalist echoes the claim, often attributed to W. V. O. Quine, that properly understood, epistemology would be no more than a chapter in a psychology text book.

The second kind of objection does not reject the idea that jurisprudence aims to provide a theory of the concept of law. Rather, it focuses on the form of conceptual analysis that Hart and others have been committed to, according to which the goal of jurisprudence is to identify the rule or criteria for the proper use of the concept of law. There are several objections to this project. One worry is that the concept of law may not be governed by a rule for its proper use, at least not one that is fixed by the shared understandings and behavior of competent speakers. As some have put it, the concept of law may be an essentially contested concept, the criteria of its proper application being fundamentally and inevitably in dispute.

Ronald Dworkin, for one, views the concept of law as essentially contested. Because the criteria for its application are necessarily in dispute, the proper application of the concept cannot be determined by a rule, and certainly not one whose content is shared by competent speakers. The essentially contestable nature of the concept of law implies that a theory of it cannot be constructed from reports of common use and understandings, even suitably revised and refined. Instead, the method of conceptual analysis appropriate to law is "constructive interpretation." Such an interpretation requires first attributing a value or purpose to law. The purpose of law is introduced to explain why it would be rational for agents to participate in it, or in some other sense to legitimate the practice. The theory of the concept is constructed by imposing this value on the practice of law as a way of organizing it and determining which features of it are most important to explain. Most important, a constructive interpretation of the concept of law is a normative theory of law. The interpretation begins with a contestable claim about the value of law that can only be defended by appealing to substantive moral principles.

Interestingly, Dworkin shares more with the natural jurisprudent than one might think. Both feel that a descriptive account of our legal practice is best left to social scientists, not philosophers. The naturalist takes this to be reason enough to deny that philosophical jurisprudence is a distinctive endeavor. For Dworkin, it is reason to think that philosophical jurisprudence must be normative.

Interpretivism is one form of normative jurisprudence. Like the so-called descriptivists and Dworkin, and unlike the naturalists, Stephen Perry accepts that the project of jurisprudence is to analyze the concept of law. Like Dworkin, he thinks the descriptivists have gone awry by thinking that an analysis of the concept of law can be achieved by reflecting on ordinary use, that is, by culling data about use, then revising and refining accordingly.

Perry's argument for normative jurisprudence is very different than Dworkin's, however. His point of departure is the claim that every theory of law has embedded within it a range of normative premises—about the nature of human agency, the value of governance by law, and most important, the proper function of law. According to Perry, the best way to interpret Hart is as claiming that the function of law is to guide conduct by reasons. By the same token, the best way to understand Dworkin is as claiming that the function of law is to justify the application of coercive force in terms of the past political decisions of legal actors. For Perry, defending a theory of the concept of law requires defending one or another view about the proper function of law. Any such defense calls for arguments of political morality, not for reports of common use or understanding.

Finally, other philosophers of law, notably Jules Coleman and Ori Simchen, take issue with conceptual analysis as the method of jurisprudence in somewhat different terms. As they see it, there are at least two problems with conceptual analysis. The first is that it relies on an unsustainable formulation of the analytic/synthetic distinction, according to which theorizing about a subject has distinct conceptual and empirical dimensions. The role of philosophy is identified with the former; the rest is a matter of empirical science. This division of labor, they claim, relies on a way of distinguishing the analytic from the synthetic that is untenable. Those who identify philosophical inquiry with conceptual analysis believe that the role of philosophy is to explore the fundamental concepts of an area of study. Philosophy uncovers the nature of the things studied by uncovering the conditions for the proper application of the relevant concepts. But this again artificially constrains the role of philosophy. It may well be that we cannot study a subject without having a concept of it, but that does not mean that philosophical inquiry must be identified with determining the conditions for the proper application of the concept. There is no reason why philosophical inquiry into law cannot be a direct account of the significant features of legal practice itself, and not merely of the concepts used to refer to the practice of law.

To sum up, the methods of analytic jurisprudence are hotly contested. The partisans fall into two camps: those who identify the distinctive role of philosophical inquiry with traditional conceptual analysis and those who, in one way or another, reject this approach. Arguably, Raz falls into the first category, whereas Leiter, Dworkin, Perry, Coleman, and Simchen, among others, fall into the second. Those who reject traditional conceptual analysis do so for a variety of reasons. Some naturalists, such as Leiter, take the rejection of the analytic/synthetic distinction to mean in effect that there is no distinctive role for philosophy in jurisprudence. That is, they implicitly accept the view that what is distinctive of philosophy is conceptual analysis, but since conceptual analysis requires the analytic/synthetic distinction, rejecting the distinction implies abandoning a distinctive role for philosophy.

Defenders of normative jurisprudence, especially Dworkin, believe that if there is a role for philosophy in the wake of the rejection of the analytic/synthetic distinction, it must take the form of a normative theory of law. After all, ifjurisprudence cannot be conceptual, and empirical inquiries are best left to the social scientists, all that remains for philosophy is to advance a speculative normative philosophy of law. Still others who reject the analytic/synthetic distinction, like Coleman and Simchen, are inclined to the view that abandoning the distinction means that philosophical inquiry into law can be an amalgam of the empirical and the conceptual.

See also Feminist Legal Theory; Legal Realism; Natural Law; Positivism.

Bibliography

Coleman, Jules. "Methodology." In *The Oxford Handbook of Jurisprudence and Philosophy of Law*, edited by Jules Coleman and Scott Shapiro, 311–351. Oxford, U.K.: Oxford University Press, 2002.

Coleman, Jules. *The Practice of Principle*. Oxford, U.K.: Oxford University Press, 2001.

Coleman, Jules and Simchen, Ori. "Law." *Legal Theory* 9 (2003): 1–41.

Dickson, Julie. *Evaluation and Legal Theory*. Oxford, U.K.: Hart Publishing, 2001.

Dickson, Julie. "Methodology in Jurisprudence." *Legal Theory* 10 (2004): 117–156.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

Finnis, John. "Law and What I Truly Should Decide." *American Journal of Jurisprudence* 48 (2003): 107–129.

Finnis, John. *Natural Law and Natural Rights*. Oxford, U.K.: Oxford University Press, 1980.

Greenberg, Mark. "How Facts Make Law." *Legal Theory* 10 (2004): 157–198.

Hart, H. L. A. *The Concept of Law*. Oxford, U.K.: Clarendon Press, 1994.

Leiter, Brian. "Beyond the Hart/Dworkin Debate: The Methodology Problem in Jurisprudence." *American Journal of Jurisprudence* 48 (2003): 17–51.

Leiter, Brian. "Legal Realism, Hard Positivism, and the Limits of Conceptual Analysis." In *Hart's Postscript*, edited by Jules Coleman, 355–370. Oxford, U.K.: Oxford University Press, 2001.

Leiter, Brian. "Naturalism and Naturalized Jurisprudence." In *Analyzing Law: New Essays in Legal Theory*, edited by Brian Bix, 79–104. Oxford, U.K.: Clarendon Press, 1998.

Marmor, Andrei, ed. *Law and Interpretation*. Oxford, U.K.: Oxford University Press, 1995.

Perry, Stephen. "Hart's Methodological Positivism." In *Hart's Postscript*, edited by Jules Coleman, 311–354. Oxford, U.K.: Oxford University Press, 2001.

Perry, Stephen. "Holmes versus Hart: The Bad Man in Legal Theory." In *"The Path of Law" and Its Influence*, edited by Steven Burton, 158–169. Cambridge, U.K.: Cambridge University Press, 2000.

Perry, Stephen. "Interpretation and Methodology in Legal Theory." In *Law and Interpretation*, edited by Andrei Marmor, 97–136. Oxford, U.K.: Clarendon Press, 1995.

Postema, Gerald. "Jurisprudence as Practical Philosophy." *Legal Theory* 4 (1998): 329–357.

Raz, Joseph. "The Problem about the Nature of Law." In *Ethics in the Public Domain*, 179–193. Oxford, U.K.: Oxford University Press, 1994.

Raz, Joseph. "Two Views of the Nature of Law: A Partial Comparison." In *Hart's Postscript*, edited by Jules Coleman, 1–38. Oxford, U.K.: Oxford University Press, 2001.

Shapiro, Scott. "The Bad Man and the Internal Point of View." In *"The Path of Law" and Its Influence*, edited by Steven Burton, 197–210. Cambridge, U.K.: Cambridge University Press, 2000.

Stavropoulos, Nicos. "Hart's Semantics." In *Hart's Postscript*, edited by Jules Coleman. Oxford, U.K.: Oxford University Press, 2001.

Stavropoulos, Nicos. "Interpretivist Theories of Law." In *The Stanford Encyclopedia of Philosophy*, winter 2003 ed., edited by Edward N. Zalta. Available at http://plato.stanford.edu/archives/win2003/entries/law-interpretivist/.

Waldron, Jeremy. *Law and Disagreement*. Oxford, U.K.: Oxford University Press, 2001.

Jules L. Coleman (2005)

ANANKE

See *Moira/Tyche/Ananke*

ANAPHORA

The study of *anaphora* (from Greek, "carry back") is the study of the ways in which occurrences of certain expressions, particularly pronouns, depend for their interpretations upon the interpretations of occurrences of other expressions. Problems of anaphora are of interest to philosophy and logic because of their intersection with problems of ontology, quantification, and logical form.

REFERENTIAL ANAPHORA

Pronouns understood as anaphoric on referential noun phrases are plausibly viewed as referring to the same things as their antecedents. Sentences (1)–(3) permit such readings (coindexing will be used to indicate an intentional anaphoric connection):

(1) Jim_1 respects students who argue with him_1.

(2) Jim_1 loves his_1 mother.

(3) Jim_1 is here. He_1 arrived yesterday. I think he_1's asleep right now.

We might call these pronouns "referential anaphors."

It is sometimes suggested (see, e.g., Soames 1994) that anaphoric pronouns in such constructions can be understood in a second way. For example, although (2) might be understood as equivalent to "Jim loves Jim's mother," it might seem to admit of another interpretation that makes it equivalent to "Jim is a self's-mother-lover," the logical form of which is given by (2'):

(2') $\lambda x(x$ loves x's mother)Jim.

The contrast between the two readings emerges when (2) is embedded, as in

(4) Mary believes that Jim_1 loves his_1 mother.

Certainly, many of the traditional problems involved in interpreting proper names recur for pronouns anaphoric on names.

BOUND-VARIABLE ANAPHORA

Pronouns anaphoric on quantified noun phrases cannot be treated as straightforwardly referential. Consider the following:

(5) Every man_1 thinks he would be a good president_1.

(6) No man_1 respects his_1 brothers' friends.

There is no point inquiring into the referents of the pronouns in examples like these. Following W. V. Quine (1960) and P. Geach (1962), philosophers have tended to treat such pronouns as the natural-language analogs of the variables of quantification theory. Certainly, the logical forms of quantified sentences of the form "every F is G" and "some Fs and Gs" can be captured using the standard first-order quantifiers "$\forall$" and "$\exists$." But a comprehensive semantic theory must treat sentences containing noun phrases formed using "no," "the," "exactly one," "most," "few," and so on. This fact highlights two problems. Using the identity sign "=" and the negation sign "$\neg$," it is possible to use "$\forall$" and "$\exists$" to represent sentences containing "no," "the," "exactly one," "exactly two,"

and so forth, but the resulting formulae obscure the relationship between the surface syntax of a sentence and its logical form. For example, if Bertrand Russell is right that "the F is G" is true if and only if every F is G and there is exactly one F, then the logical form of this sentence is as follows:

(7) $(\exists x)((\forall y)(Fy \equiv y = x) \& Gx)$.

A more serious problem is that there are sentences that cannot be dealt with in first-order logic—for instance, sentences of the form "most Fs are Gs."

Both of these problems are solved if quantification in natural language is viewed as *restricted.* The basic idea here is that determiners combine with their complements (noun complexes) to form restricted quantifiers. So, for example, "every," "some," "most," "the," and so on combine with simple nouns such as "pig" (or "pigs"), "man" (or "men"), and so forth (or complex nouns such as "man who owns a pig," etc.) to form restricted quantifiers such as "some man," "most men," "every man who owns a pig," and so forth. We can represent a restricted quantifier "every man" as "[every x: man x]." This quantifier may combine with a predicate phrase such as "is mortal" (which we can represent as "x is mortal") to form the sentence "every man is mortal," which we can represent as

(8) [every x: man x]x is mortal.

Now consider sentences (5) and (6) again. If we treat the anaphoric pronouns in these examples as bound variables, their logical forms will be (abstracting somewhat):

(5') [every x: man x](x thinks x would be a good president).

(6') [no x: man x](x respects x's brothers' friends).

VARIABLE BINDING AND SCOPE

G. Evans (1977) has argued that not all pronouns anaphoric on quantified noun phrases are bound variables. Consider the following examples.

(9) Jim bought some pigs and Harry vaccinated them.

(10) Just one man ate haggis and he was ill afterward.

A bound-variable treatment of the occurrence of "them" in (9) yields the wrong result. On such an account, the logical form of the sentence will be

(9') [some x: pigs x](Jim bought x & Harry vaccinated x).

But (9') can be true even if Harry did not vaccinate *all* of the pigs Jim bought, whereas (9) cannot. (If Jim bought ten pigs and Harry vaccinated only two of them, (9') would be true whereas (9) would not.) And if the pronoun "he" in (10) is treated as a bound variable, the logical form of the sentence will be

(10') [just one x: man x](x ate haggis and x was ill afterward).

This is also incorrect; if two men ate haggis and only one was ill afterward, (10') will be true whereas (10) will be false.

There is a plausible syntactic explanation of these facts. In both (9) and (10), the pronoun is located outside the smallest sentence containing the quantifier upon which it is anaphoric and hence lies outside its scope, according to the most promising syntactic characterization of this notion. The scope of an expression α in a sentence of a natural language appears to correspond to the first branching node dominating α at the syntactic level relevant to semantic interpretation. If this is correct, and contemporary syntactic theory suggests it is, then syntactic theory explains why the pronouns in (9) and (10) are not understood as bound variables. There seem to be, therefore, anaphoric pronouns that are neither bound nor straightforwardly referential.

UNBOUND ANAPHORA

A plausible paraphrase of (9) is (9"):

(9") Jim bought some pigs and Harry vaccinated the pigs Jim bought.

In view of this, Evans (1977) suggests that the pronoun "them" in (9) is understood in terms of the plural description "the pigs Jim bought," as what he calls an "E-type" pronoun. An E-type pronoun has its reference fixed by description (in Saul Kripke's sense) and is therefore a rigid designator. On this account, in (9) the pronoun "them" is taken to refer to those objects satisfying "pigs Jim bought."

Similarly where the antecedent is singular. A plausible paraphrase of (11) is (11'):

(11) Jim bought a pig and Harry vaccinated it.

(11') Jim bought a pig and Harry vaccinated the pig Jim bought.

According to Evans, the pronoun "it" in (11) refers to the unique object satisfying "pig Jim bought."

This idea forms the basis of Evans's general account of the semantic content of unbound anaphors. The pronoun "he" in (10) has its reference fixed by "the man who

ate haggis"; and in (12) "they" has its reference fixed by "the philosophers who came":

(12) A few philosophers came. They drank far too much.

Evans's proposal can be summarized thus: if *P* is an unbound pronoun anaphoric on a quantified noun phrase "[DET x: ϕ]" occurring in a sentence "[DET x: ϕ]ψ," then the referent of *P* is fixed by the description "[the x: ϕ & ψ]."

Examination of more complex cases reveals weaknesses in Evans's theory (see below). The problems uncovered have tended to steer semanticists in one of two directions. First, there have been attempts to modify or refine Evans's framework (Davies 1981, Neale 1990). Second, there have been attempts to replace the entire framework with a uniform, discourse-based approach (Kamp 1981, Heim 1982). Both approaches will now be examined.

DESCRIPTIVE ANAPHORA

Evans rejected the view that unbound anaphors go proxy for descriptions (in favor of the view that they have their referents fixed by description) on the grounds that such pronouns, unlike overt descriptions, do not give rise to ambiguities of scope. But consider the following:

(14) A man murdered Smith, but Jim doesn't think he did it.

(15) A man murdered Smith. The police have reason to think he injured himself in the process.

If "he" goes proxy for "the man who murdered Smith," there will be two readings for each of the anaphor clauses in these examples—the so-called *de re* and *de dicto* readings—according as the description for which the pronoun goes proxy is given large or small scope:

(14a) [the x: man x & x murdered Smith] (Jim doesn't believe that x murdered Smith)

(14b) Jim doesn't believe that [the x: man x & x murdered Smith](x murdered Smith)

It is natural to interpret (14) as attributing to Jim a noncontradictory belief concerning the murderer to the effect that he is not the murderer. On the proxy view this is captured by the *de re* reading of the second conjunct. The *de dicto* reading is technically available to the proxy theorist but is obviously not the preferred interpretation. But with (15) the *de dicto* reading of the second sentence is actually the more natural; yet Evans's theory explicitly precludes its existence.

Further support for the proxy rather than reference-fixing approach comes from examples containing modal expressions:

(16) Mary wants to marry a rich man. He must be a banker.

The first sentence in (16) may be read either *de re* or *de dicto.* Moreover, the pronoun "he" can be anaphoric on "a rich man" on either reading. But as L. Karttunen (1976) points out, the modal expression has to be there for the anaphora to work if the antecedent sentence is to be interpreted *de dicto.* That is, in

(17) Mary wants to marry a rich man. He is a banker.

it is not possible to get the *de dicto* reading for the antecedent clause if "he" is anaphoric on "a rich man." This contrast between (16) and (17) is explicable on the assumption that the anaphoric pronoun in (16) goes proxy for the description "the man Mary marries" and may therefore take large or small scope with respect to the modal expression. On the *de dicto* reading of the antecedent clause, the *de re* reading of the anaphor clause is infelicitous because an implication of existence results from giving the description large scope. But the *de dicto* reading of the anaphor clause is fine because on such a reading the description is within the scope of the modal expression. In (17), on the other hand, since there is no modal operator with respect to which the pronoun can be understood with small scope, the sentence has no felicitous reading when the antecedent clause is read *de dicto.*

DONKEY ANAPHORA

H. Kamp (1981) and I. Heim (1982) have explored alternative approaches that aim to treat all anaphoric pronouns in a unitary fashion. One motivation is the problem of so called donkey anaphora, typified by sentences such as (18) and (19), originally discussed by Geach (1962):

(18) If a man buys a donkey he vaccinates it.

(19) Every man who buys a donkey vaccinates it.

Both Evans's theory and the simple proxy theory seem to fail here. For example, if the pronoun "it" in (19) is analyzed in terms of the singular description "the donkey he buys" (with "he" bound by "every man who buys a donkey") the sentence will be true just in case every man who buys a donkey vaccinates the unique donkey he buys. Consequently, it will be false if any man buys more than one donkey. But this is incorrect; the truth of (19) is quite compatible with some men owning more than one donkey, as long as every man who buys a donkey vaccinates

every donkey he buys. It would appear, then, that the indefinite description "a donkey"—which can normally be treated as an existentially quantified phrase—has the force of a *universally* quantified phrase in (19). And in (18) both "man" and "a donkey" appear to have universal force.

A common explanation of the "universalization" of the indefinite descriptions in such examples has been proposed by Kamp. The idea (roughly) is that noun phrases introduce variables to which common nouns and predicates supply "conditions" within a "discourse representation" (DR). Typically, the variable is bound by an existential quantifier taking scope over the entire discourse. On this account, an indefinite description is not inherently quantificational; rather, it introduces a variable with conditions on it imposed by, among other things, the predicative material it contains. The DR for (18) might be represented as:

(18') [man(x) & donkey(y) & buys(x,y)] IFTHEN [vaccinates(x,y)].

Kamp proposes that (18') is true if and only if every assignment of values to x and y that makes the antecedent true also makes the consequent true. The apparent universalization of the indefinite descriptions "a man" and "a donkey" is thus explained as a consequence of a general analysis of conditionals.

In the light of the equivalence of (18) and (19), Kamp suggests that, although (18) is not actually a conditional, because the subject quantifier is universal we get a DR in which the indefinite "a donkey" has universal force. That is, the DR for (19) is given by

(19') [man(x) & donkey(y) & buys(x,y)] EVERY [vaccinates(x,y)].

Like (18'), (19') is true if and only if every assignment of values to x and y that makes "[man(x) & donkey(y) & buys(x,y)]" true, also makes "[vaccinates(x,y)]" true.

One problem with this proposal is that it does not predict that indefinite descriptions "universalize" when they are embedded in *other* quantifiers and thus leads to the so-called proportion problem. Consider

(20) Most men who buy a donkey vaccinate it.

By analogy with (18') and (19'), the DR for (20) will be

(20') [man(x) & donkey(y) & buys(x,y)] MOST [vaccinates(x,y)]

which is true just in case *most* assignments of values to x and y that make "[man(x) & donkey(y) & buys(x,y)]" true also make "[vaccinates(x,y)]" true. But on its most natural reading, the truth of (20) requires that most men who buy a donkey vaccinate *every* donkey they buy, whereas (20') can be true as long as most of the donkeys that are bought by men are vaccinated by their respective buyers. Suppose Alan buys five donkeys, Bill buys one donkey, Clive buys one donkey, and no other man buys any donkeys. Sentence (20') will come out true if Alan vaccinates at least four of his donkeys, even if Bill and Clive do not vaccinate their respective donkeys; but in such a situation (20) would be false. (It has been suggested that there is another reading of (20), which requires that most men who buy at least one donkey vaccinate most of the donkeys they buy; but (20') does not capture this reading either.)

From this brief overview it should be clear that both the simple descriptive theory and the simple DR theory need to be refined if they are to do justice to the full range of antecedent/anaphor relations in natural language. For example, the descriptive approach needs to be modified if it is to handle donkey anaphora, perhaps allowing for the possibility of interpreting some donkey pronouns in terms of "all of the" rather than "the" (Davies 1981, Neale 1990). And the DR approach needs to be modified to avoid the proportion problem and also permit pronouns to be understood with various scopes. At the time of writing, more sophisticated versions of these theories are being developed, as are alternatives to both.

See also Kripke, Saul; Logical Form; Ontology; Philosophy of Language; Quine, Willard Van Orman; Russell, Bertrand Arthur William.

Bibliography

Davies, M. *Meaning, Quantification, Necessity.* London: Routledge and Kegan Paul, 1981.

Evans, G. *The Collected Papers.* Oxford: Clarendon Press, 1985.

Evans, G. "Pronouns." *Linguistic Inquiry* 11 (1980): 337–362. (Reprinted in Evans [1985], 214–248.)

Evans, G. "Pronouns, Quantifiers, and Relative Clauses (I)." *Canadian Journal of Philosophy* 7 (1977): 467–536. (Reprinted in Evans [1985], 76–152.)

Geach, P. *Reference and Generality.* Ithaca, NY: Cornell University Press, 1962.

Heim, I. *The Semantics of Definite and Indefinite Noun Phrases.* Ph.D diss. University of Massachusetts, Amherst, 1982.

Kamp, H. "A Theory of Truth and Semantic Interpretation." In *Formal Methods in the Study of Natural Language,* edited by J. Groenendijk et al. Amsterdam Centre, 1981.

Karttunen, L. "Discourse Referents." In *Syntax and Semantics* 7, *Notes from the Linguistic Underground,* edited by J. McCawley. New York: Academic Press, 1976.

Kripke, S. "Naming and Necessity." In *Semantics of Natural Language,* edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.

Neale, S. *Descriptions.* Cambridge, MA: MIT Press, 1990.
Quine, W. V. O. *Word and Object.* Cambridge, MA: MIT Press, 1960.
Soames, S. "Attitudes and Anaphora." *Philosophical Perspectives* 8 (1994): 251–272.

Stephen Neale (1996)

ANAPHORA [ADDENDUM]

Most recent work on anaphora has tended to focus on cases in which a pronoun is anaphoric on what appears to be a quantifier phrase, where it cannot be understood as a variable bound by that quantifier phrase (as it is in 5 and 6 above). Two central cases of this sort, to which attention will be confined here, are as follows. First, there is *discourse anaphora* in which the anaphoric expression is in a different sentence from its antecedent (see also 12, and 15–17 above):

(21) Few cars are gasoline and electric hybrids. They are expensive;

(22) A woman is at the door. She is from Santa Monica.

There are at least two reasons for thinking that the pronouns in 21 and 22 are not variables bound by its quantifier antecedents. Garreth Evans (1977) appears to be the first to discuss both reasons. Focusing on 21, the first reason is that such a treatment gets the wrong truth conditions for an example like 21. If *they* is a bound variable in 21, the two sentences of 21 together should be equivalent to: 21a) Few cars: x (x is a gasoline and electric hybrid and x is expensive). But this is obviously incorrect because 21 entails that few cars are hybrids, whereas 21a does not.

Second, it is generally thought that the scope of a quantifier cannot extend beyond the sentence in which it occurs. If that is correct, then the pronoun in 21 falls outside the scope of its quantifier antecedent and so cannot be bound by it. So, though the pronoun in 21 has a quantified antecedent, it cannot be understood as a variable bound by it.

Another sort of case in which this occurs is that of *donkey anaphora*, which comes in two varieties as illustrated above by 18 and 19. Let us call the former *conditional donkey anaphora* and the latter *relative clause donkey anaphora*. In the case of 18, all independent evidence suggests that (what appears to be) the quantifier "a donkey" cannot take wide scope over the conditional and bind the pronoun in the consequent. When one attempts to do this with other quantifiers, one cannot bind the pronoun, as is the case here:

(23) If John buys every $donkey_1$, he beats it_1.

And even if "a donkey" could somehow scope over the conditional and bind the pronoun in the consequent, assuming it is as usual an existential quantifier, we still wouldn't get the right truth conditions for 18, which intuitively require donkey-owning men to vaccinate every donkey they own. In the case of 19, again the independent evidence suggests that a quantifier occurring in a relative clause cannot bind a pronoun outside the relative clause:

(24) Every teacher who flunks every male $student_1$ hates him_1.

So both donkey anaphora and discourse anaphora require treating the anaphoric pronoun as something other than a bound variable. Thus both of these phenomena shall be grouped under the heading *unbound anaphora*. In addition to the descriptive and discourse representation approaches discussed above, there are at least two other attempts to treat unbound anaphora.

First, there are Dynamic Logic Accounts (DL), originally formulated by Jeroen Groenendijk and Martin Stokhof (1991). Other DL accounts have been suggested by Gennaro Chierchia (1995) and Makoto Kanazawa (1994a and 1994b). DL accounts, which are descended from DR accounts, characteristically claim that quantifiers can semantically bind pronouns even if those pronouns do not occur in their syntactic scopes. The pronouns in 18, 19, 21, and 22, then, are semantically (though not syntactically) bound by their quantifier antecedents even though they are beyond the syntactic scope of their antecedents. DL also provides new accounts of the semantics of conditionals and universal quantification in assigning to 18 and 19 the intuitively correct truth conditions.

The second other sort of approach to unbound anaphora (in addition to descriptive and DR approaches) is the Context Dependent Quantifier Approach (CDQ), which was first suggested by George Wilson (1984) and further articulated and defended by Jeffrey King (1987, 1991, 1994, 2005). The CDQ approach holds that in cases of unbound anaphora, the pronouns in question are quantifiers. The forces (universal or existential), restrictions (what the quantifiers range over—e.g., "every student" ranges over students), and relative scopes of these pronouns and quantifiers are determined by features of their linguistic contexts. Thus, according to CDQ, the

anaphoric pronouns in cases of unbound anaphora are contextually sensitive devices of quantification.

The precise natures of the quantifications they express, the forces, restrictions, and relative scope of the quantifications, are determined by features of their linguistic contexts. In 22, CDQ holds that the pronoun expresses the existential quantifier normally expressed outside any context by the indefinite "a woman who is at the door." Similar remarks apply to 18 and 19, except that the semantics of the conditional interacts with the pronoun qua quantifier to get the proper reading of 18 (King 2005); and in 19 the pronoun qua quantifier takes narrow scope with respect to the universal quantifier "every man" (because its antecedent does as well).

See also Quantifiers in Natural Language; Reference; Syntax.

Bibliography

Chierchia, Gennaro. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar*. Chicago: University of Chicago Press, 1995.

Evans, Gareth. "Pronouns, Quantifiers, and Relative Clauses." *Canadian Journal of Philosophy* 8 (3) (1977): 467–536.

Groenendijk, J., and M. Stokhof. "Dynamic Predicate Logic." *Linguistics and Philosophy* 14 (1991): 39–100.

Kanazawa, Makoto, ed. "Dynamic Generalized Quantifiers and Montonicity." In *Dynamics, Polarity, and Quantification*. Stanford, CA: Center for the Study of Language and Information, 1994a.

Kanazawa, Makoto. "Weak vs. Strong Readings of Donkey Sentences in a Dynamic Setting." *Linguistics and Philosophy* 17 (2) (1994b): 109–158.

King, Jeffrey C. "Anaphora and Operators."*Philosophical Perspectives, Logic, and Language* 8 (1994).

King, Jeffrey C. "Context Dependent Quantifiers and Donkey Anaphora." *Canadian Journal of Philosophy* (2005).

King, Jeffrey C. "Instantial Terms, Anaphora, and Arbitrary Objects." *Philosophical Studies* 61 (1991): 239–265.

King, Jeffrey C. "Pronouns, Descriptions, and the Semantics of Discourse." *Philosophical Studies* 51 (1987): 341–363.

Wilson, George. "Pronouns and Pronominal Descriptions: A New Semantical 'Category.'" *Philosophical Studies* 45 (1984): 1–30.

Jeffrey C. King (2005)

ANARCHISM

"Anarchism" is a social philosophy that rejects authoritarian government and maintains that voluntary institutions are best suited to express man's natural social tendencies. Historically the word *anarchist,* which derives from the Greek *an archos,* meaning "no government," appears first to have been used pejoratively to indicate one who denies all law and wishes to promote chaos. It was used in this sense against the Levelers during the English Civil War and during the French Revolution by most parties in criticizing those who stood to the left of them along the political spectrum. The first use of the word as an approbatory description of a positive philosophy appears to have been by Pierre-Joseph Proudhon when, in his *Qu'est-ce-que la propriété?* (*What is property*?, Paris, 1840), he described himself as an anarchist because he believed that political organization based on authority should be replaced by social and economic organization based on voluntary contractual agreement.

Nevertheless, the two uses of the word have survived together and have caused confusion in discussing anarchism, which to some has appeared a doctrine of destruction and to others a benevolent doctrine based on a faith in the innate goodness of man. There has been further confusion through the association of anarchism with nihilism and terrorism. In fact, anarchism, which is based on faith in natural law and justice, stands at the opposite pole to nihilism, which denies all moral laws. Similarly, there is no necessary connection between anarchism, which is a social philosophy, and terrorism, which is a political means occasionally used by individual anarchists but also by actionists belonging to a wide variety of movements that have nothing in common with anarchism.

Anarchism aims at the utmost possible freedom compatible with social life, in the belief that voluntary cooperation by responsible individuals is not merely more just and equitable but is also, in the long run, more harmonious and ordered in its effects than authoritarian government. Anarchist philosophy has taken many forms, none of which can be defined as an orthodoxy, and its exponents have deliberately cultivated the idea that it is an open and mutable doctrine. However, all its variants combine a criticism of existing governmental societies, a vision of a future libertarian society that might replace them, and a projected way of attaining this society by means outside normal political practice. Anarchism in general rejects the state. It denies the value of democratic procedures because they are based on majority rule and on the delegation of the responsibility that the individual should retain. It criticizes utopian philosophies because they aim at a static "ideal" society. It inclines toward internationalism and federalism, and, while the views of anarchists on questions of economic organization vary greatly, it may be said that all of them reject what William Godwin called accumulated property.

Attempts have been made by anarchist apologists to trace the origins of their point of view in primitive non-governmental societies. There has also been a tendency to detect anarchist pioneers among a wide variety of teachers and writers who, for various religious or philosophical reasons, have criticized the institution of government, have rejected political activity, or have placed a great value on individual freedom. In this way such varied ancestors have been found as Laozi, Zeno, Spartacus, Étienne de La Boétié, Thomas Münzer, François Rabelais, François Fénelon, Denis Diderot, and Jonathan Swift; anarchist trends have also been detected in many religious groups aiming at a communalistic order, such as the Essenes, the early Christian apostles, the Anabaptists, and the Doukhobors. However, while it is true that some of the central libertarian ideas are to be found in varying degrees among these men and movements, the first forms of anarchism as a developed social philosophy appeared at the beginning of the modern era, when the medieval order had disintegrated, the Reformation had reached its radical, sectarian phase, and the rudimentary forms of modern political and economic organization had begun to appear. In other words, the emergence of the modern state and of capitalism is paralleled by the emergence of the philosophy that, in various forms, has opposed them most fundamentally.

WINSTANLEY

Although Proudhon was the first writer to call himself an anarchist, at least two predecessors outlined systems that contain all the basic elements of anarchism. The first was Gerrard Winstanley (1609–c. 1660), a linen draper who led the small movement of the Diggers during the Commonwealth. Winstanley and his followers protested in the name of a radical Christianity against the economic distress that followed the Civil War and against the inequality that the grandees of the New Model Army seemed intent on preserving. In 1649–1650 the Diggers squatted on stretches of common land in southern England and attempted to set up communities based on work on the land and the sharing of goods. The communities failed, but a series of pamphlets by Winstanley survived, of which *The New Law of Righteousness* (1649) was the most important. Advocating a rational Christianity, Winstanley equated Christ with "the universal liberty" and declared the universally corrupting nature of authority. He saw "an equal privilege to share in the blessing of liberty" and detected an intimate link between the institution of property and the lack of freedom. In the society he sketched, work would be done in common and the products shared equally through a system of open storehouses, without commerce.

Like later libertarian philosophers, Winstanley saw crime as a product of economic inequality and maintained that the people should not put trust in rulers. Rather, they should act for themselves in order to end social injustice, so that the land should become a "common treasury" where free men could live in plenty. Winstanley died in obscurity and, outside the small and ephemeral group of Diggers, he appears to have wielded no influence, except possibly over the early Quakers.

GODWIN

A more elaborate sketch of anarchism, although still without the name, was provided by William Godwin in his *Enquiry concerning Political Justice* (1793). Godwin differed from most later anarchists in preferring to revolutionary action the gradual and, as it seemed to him, more natural process of discussion among men of good will, by which he hoped truth would eventually triumph through its own power. Godwin, who was influenced by the English tradition of Dissent and the French philosophy of the Enlightenment, put forward in a developed form the basic anarchist criticisms of the state, of accumulated property, and of the delegation of authority through democratic procedure. He believed in a "fixed and immutable morality," manifesting itself through "universal benevolence"; man, he thought, had no right "to act anything but virtue and to utter anything but truth," and his duty, therefore, was to act toward his fellow men in accordance with natural justice. Justice itself was based on immutable truths; human laws were fallible, and men should use their understandings to determine what is just and should act according to their own reasons rather than in obedience to the authority of "positive institutions," which always form barriers to enlightened progress. Godwin rejected all established institutions and all social relations that suggested inequality or the power of one man over another, including marriage and even the role of an orchestra conductor. For the present he put his faith in small groups of men seeking truth and justice; for the future, in a society of free individuals organized locally in parishes and linked loosely in a society without frontiers and with the minimum of organization. Every man should take part in the production of necessities and should share his produce with all in need, on the basis of free distribution. Godwin distrusted an excess of political or economic cooperation; on the other hand, he looked forward to a freer intercourse of individuals through the progressive breaking down of social and economic barri-

ers. Here, conceived in the primitive form of a society of free landworkers and artisans, was the first sketch of an anarchist world. The logical completeness of *Political Justice,* and its astonishing anticipation of later libertarian arguments, make it, as Sir Alexander Gray said, "the sum and substance of anarchism."

NINETEENTH-CENTURY EUROPEAN ANARCHISM

Despite their similarities to later libertarian philosophies, however, the systems of Winstanley and Godwin had no perceptible influence on nineteenth-century European anarchism, which was an independent development and which derived mainly from the peculiar fusion of early French socialist thought and German neo-Hegelianism in the mind of Proudhon, the Besancon printer who has been called the father of anarchism. This tradition centered largely on a developing social revolutionary movement that attained mass dimensions in France, Italy, and Spain (where anarchism remained strong until the triumph of Franco in 1939), and to a lesser extent in French-speaking Switzerland, the Ukraine and Latin America. Apart from Proudhon, its main advocates were Michael Bakunin, Prince Peter Kropotkin, Errico Malatesta, Sebastien Faure, Gustav Landauer, Elisée Reclus, and Rudolf Rocker, with Max Stirner and Lev Tolstoy on the individualist and pacifist fringes, respectively. Also, there arose among nineteenth-century anarchists a mystique that action and even theory should emerge from the people. Libertarian attitudes, particularly in connection with the anarchosyndicalism of France and Spain, were influenced by the rationalization and even romanticization of the experience of social struggle; the writings of Fernand Pelloutier and Georges Sorel in particular emanate from this aspect of the anarchist movement. Nineteenth-century anarchism assumed a number of forms, and the points of variation between them lie in three main areas: the use of violence, the degree of cooperation compatible with individual liberty, and the form of economic organization appropriate to a libertarian society.

INDIVIDUALIST ANARCHISM. Individualist anarchism lies on the extreme and sometimes dubious fringe of the libertarian philosophies since, in seeking to assure the absolute independence of the person, it often seems to negate the social basis of true anarchism. This is particularly the case with Max Stirner, who specifically rejected society as well as the state and reduced organization to a union of egoists based on the mutual respect of "unique" individuals, each standing upon his "might." French anarchism during the 1890s was particularly inclined toward individualism, which expressed itself partly in a distrust of organization and partly in the actions of terrorists like "Ravachol" and Émile Henry, who alone or in tiny groups carried out assassinations of people over whom they had appointed themselves both judges and executioners. A milder form of individualist anarchism was that advocated by the American libertarian writer Benjamin Tucker (1854–1939), who rejected violence in favor of refusal to obey and who, like all individualists, opposed any form of economic communism. What he asked was that property should be distributed and equalized so that every man should have control over the product of his labor.

MUTUALISM. Mutualism, developed by Proudhon, differed from individualist anarchism in its stress on the social element in human behavior. It rejected both political action and revolutionary violence—some of Proudhon's disciples even objected to strikes as a form of coercion—in favor of the reform of society by the peaceful spread of workers' associations, devoted particularly to mutual credit between producers. A recurrent mutualist plan, never fulfilled, was that of the people's bank, which would arrange the exchange of goods on the basis of labor notes. The mutualists recognized that workers' syndicates might be necessary for the functioning of industry and public utilities, but they rejected large-scale collectivization as a danger to liberty and based their economic approach as far as possible on individual possession of the means of production by peasants and small craftsmen united in a framework of exchange and credit arrangements. The mutualists laid great stress on federalist organization from the local commune upward as a substitute for the national state. Mutualism had a wide following among French artisans during the 1860s. Its exponents were fervently internationalist and played a great part in the formation of the International Workingmen's Association in 1864; their influence diminished, however, with the rise of collectivism as an alternative libertarian philosophy.

COLLECTIVISM. Collectivism is the form of anarchism associated with Michael Bakunin. The collectivist philosophy was developed by Bakunin from 1864 onward, when he was forming the first international organizations of anarchists, the International Brotherhood and the International Alliance of Social Democracy. It was collectivist anarchism that formed the principal opposition to Marxism in the International Workingmen's Association and thus began the historic rivalry between libertarian and

authoritarian views of socialism. Bakunin and the other collectivists agreed with the mutualists in their rejection of the state and of political methods, in their stress on federalism, and in their view that the worker should be rewarded according to his labor. On the other hand, they differed in stressing the need for revolutionary means to bring about the downfall of the state and the establishment of a libertarian society. Most important, they advocated the public ownership and the exploitation through workers' associations of the land and all services and means of production. While in mutualism the individual worker had been the basic unit, in collectivism it was the group of workers; Bakunin specifically rejected individualism of any kind and maintained that anarchism was a social doctrine and must be based on the acceptance of collective responsibilities.

ANARCHIST COMMUNISM. Collectivism survived as the dominant anarchist philosophy in Spain until the 1930s; elsewhere it was replaced during the 1870s by the anarchist communism that was associated particularly with Kropotkin, although it seems likely that Kropotkin was merely the most articulate exponent of a trend that grew out of discussions among anarchist intellectuals in Geneva during the years immediately after the Paris Commune of 1871. Through Kropotkin's literary efforts anarchist communism was much more elaborately worked out than either mutualism or collectivism; in such books as *La conquête du pain* (*The conquest of bread,* 1892) and *Fields, Factories and Workshops* (1899) Kropotkin elaborated the scheme of a semiutopian decentralized society based on an integration of agriculture and industry, of town life and country life, of education and apprenticeship. Kropotkin also linked his theories closely with current evolutionary theories in the fields of anthropology and biology; anarchism, he suggested in *Mutual Aid* (1902), was the final stage in the development of cooperation as a factor in evolution.

Anarchist communism differed from collectivism on only one fundamental point—the way in which the product of labor should be shared. In place of the collectivist and mutualist idea of remuneration according to hours of labor, the anarchist communists proclaimed the slogan "From each according to his means, to each according to his needs" and envisaged open warehouses from which any man could have what he wanted. They reasoned, first, that work was a natural need that men could be expected to fulfill without the threat of want and, second, that where no restriction was placed on available goods, there would be no temptation for any man to take more than he could use. The anarchist communists laid great stress on local communal organization and even on local economic self-sufficiency as a guarantee of independence.

ANARCHOSYNDICALISM. Anarchosyndicalism began to develop in the late 1880s, when many anarchists entered the French trade unions, or syndicates, which were just beginning to reemerge after the period of suppression that followed the Paris Commune. Later, anarchist militants moved into key positions in the Confédération Générale du Travail, founded in 1895, and worked out the theories of anarchosyndicalism. They shifted the basis of anarchism to the syndicates, which they saw as organizations that united the producers in common struggle as well as in common work. The common struggle should take the form of "direct action," primarily in industry, since there the workers could strike most sharply at their closest enemies, the capitalists; the highest form of direct action, the general strike, could end by paralyzing not merely capitalism but also the state.

When the state was paralyzed, the syndicates, which had been the organs of revolt, could be transformed into the basic units of the free society; the workers would take over the factories where they had been employees and would federate by industries. Anarchosyndicalism created a mystique of the working masses that ran counter to individualist trends; and the stress on the producers, as distinct from the consumers, disturbed the anarchist communists, who were haunted by the vision of massive trade unions ossifying into monolithic institutions. In France, Italy, and Spain, however, it was the syndicalist variant that brought anarchism its first and only mass following. The men who elaborated the philosophy of anarchosyndicalism included militants, such as Fernand Pelloutier, Georges Yvetot, and Émile Pouget, who among them created the vision of a movement arising from the genius of the working people. There were also intellectuals outside the movement who drew theoretical conclusions from anarchosyndicalist practice; the most important was Sorel, the author of *Réflexions sur la violence* (*Reflections on violence,* 1908), who saw the general strike as a saving "social myth" that would maintain society in a state of struggle and, therefore, of health.

PACIFIST ANARCHISM. Pacifist anarchism has taken two forms. That of Tolstoy attempted to give rational and concrete form to Christian ethics. Tolstoy rejected all violence; he advocated a moral revolution, its great tactic the refusal to obey. There was much, however, in Tolstoy's criticisms of contemporary society and his suggestions for the future that paralleled other forms of anarchism. He denounced the state, law, and property; he foresaw

cooperative production and distribution according to need.

Later a pacifist trend appeared in the anarchist movement in western Europe; its chief exponent was the Dutch ex-socialist, Domela Nieuwenhuis. It differed from strict Tolstoyism by accepting syndicalist forms of struggle that stopped short of violence, particularly the millenarian general strike for the abolition of war.

Despite their differences, all these forms of anarchism were united not merely in their rejection of the state, of politics, and of accumulated property, but also in certain more elusive attitudes. In its avoidance of partisan organization and political practices, anarchism retained more of the moral element than did other movements of protest. This aspect was shown with particular sharpness in the desire of its exponents for the simplification of life, not merely in the sense of removing the complications of authority, but also in eschewing the perils of wealth and establishing a frugal sufficiency as the basis for life. Progress, in the sense of bringing to all men a steadily rising supply of material goods, has never appealed to the anarchists; indeed, it is doubtful if their philosophy is at all progressive in the ordinary sense. They reject the present, but they reject it in the name of a future of austere liberty that will resurrect the lost virtues of a more natural past, a future in which struggle will not be ended, but merely transformed within the dynamic equilibrium of a society that rejects utopia and knows neither absolutes nor perfections.

The main difference between the anarchists and the socialists, including the Marxists, lies in the fact that while the socialists maintain that the state must be taken over as the first step toward its dissolution, the anarchists argue that, since power corrupts, any seizure of the existing structure of authority can only lead to its perpetuation. Anarchosyndicalists, however, regard their unions as the skeleton of a new society growing up within the old.

The problem of reconciling social harmony with complete individual freedom is a recurrent one in anarchist thought. It has been argued that an authoritarian society produces antisocial reactions, which would vanish in freedom. It has also been suggested, by Godwin and Kropotkin particularly, that public opinion will suffice to deter those who abuse their liberty. George Orwell, however, has pointed out that the reliance on public opinion as a force replacing overt coercion might lead to a moral tyranny which, having no codified bounds, could in the end prove more oppressive than any system of laws.

See also Bakunin, Mikhail Aleksandrovich; Communism; Diderot, Denis; Fénelon, François de Salignac de la Mothe; Godwin, William; Kropotkin, Pëtr Alekseevich; Laozi; Pacifism; Proudhon, Pierre-Joseph; Sorel, Georges; Stirner, Max; Swift, Jonathan; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

Avrich, Paul. *Anarchist Portraits.* Princeton: Princeton University Press, 1988.

Bakunin, Michael. *Michael Bakunin: Statism and Anarach.* Translated and edited by Marshall S. Shatz. Cambridge U.K.: Cambridge University Press, 1990.

Crowder, George. *Classical Anarchism: The Political Thought of Godwin, Proudhon, Bakunin, and Kropotkin.* Oxford: Oxford University Press, 1991.

Eltzbacher, Paul. *Anarchism.* Translated by Steven T. Byington. New York: Tucker, 1908.

Gans, Chaim. *Philosophical Anarchism and Political Disobedience.* Cambridge, U.K.: Cambridge University Press, 1992.

Godwin, William. *Political and Philosophical Writings of William Godwin,* 7 vols. Edited by Mark Philp. London: Pickering, 1993.

Gray, Alexander. *The Socialist Tradition, Moses to Lenin.* London: Longmans, Green, 1946.

Guérin, Daniel. *Anarchism: From Theory to Practice.* Translated by Mary Klopper. New York: Monthly Review Press, 1970.

Joll, James. *The Anarchists.* London: Eyre and Spottiswoode, 1964.

May, Todd. *The Political Philosophy of Poststructuralist Anarchism.* University Park: Pennsylvania State University Press, 1994.

Miller, David. *Anarchism.* London: Dent, 1984.

Nettlau, Max. *Der Anarchismus von Proudhon zu Kropotkin.* Berlin: Syndikalist, 1927.

Nettlau, Max. *Anarchisten und Social-Revolutionäre.* Berlin: Syndikalist, 1931.

Nettlau, Max. *Der Vorfrühling der Anarchie.* Berlin: Syndikalist, 1925.

Ritter, Alan. *Anarchism: A Theoretical Analysis.* Cambridge, U.K.: Cambridge University Press, 1980.

Rocker, Rudolf. *Anarcho-Syndicalism.* London: Secker and Warburg, 1938.

Russell, Bertrand. "Bakunin and Anarchism." In *Proposed Roads to Freedom.* New York: Holt, 1919.

Wolff, Robert Paul. *In Defense of Anarchism.* New York: Harper and Row, 1970.

Woodcock, George. *Anarchism: A History of Libertarian Ideas and Movements.* Cleveland: World Publishing, 1962.

Zenker, E. V. *Anarchism.* London: Methuen, 1898.

George Woodcock (1967)

Bibliography updated by Philip Reed (2005)

ANAXAGORAS OF CLAZOMENAE

(c. 500–428 BCE)

One of the leading philosophers of the fifth century BCE, Anaxagoras continued the cosmological style of philosophy begun in Miletus in the preceding century. Born in Clazomenae in Asia Minor around 500, he came to Athens and spent thirty years there, enjoying access to intellectual circles through his friendship with Pericles. Two alternate accounts of his dates in Athens are available: either he came around 480 and stayed until 450, or he came around 460 and stayed until 430. Because his name is associated with a meteor that fell near Aegospotami around 467, and his theory of the Nile floods was known to Aeschylus (d. 456), it appears that his work was well known already in the 460s, supporting an early date at least for his philosophical activity. Anaxagoras is said to have fled Athens to avoid prosecution on a charge of impiety, and he finished his days in Lampsacus in northern Greece, where he died in 428. He wrote a book that was well-known in Athens in the late fifth century BCE and was available until the sixth century CE. About twenty fragments of the book survive, describing some key points of his theory. Although Anaxagoras wrote in simple Ionic prose, many details of his theory remain obscure and controversial.

Like most natural philosophers of his time, Anaxagoras tells how the world arose from a primeval chaos. Initially, all things (kinds of matter, presumably) were mixed together to such an extent that nothing was differentiated. But Mind (*Nous*) caused a whirling motion to start, which caused different materials to separate out, as in a centrifuge, leading to the articulation of the cosmos. At the center of the cosmos is a flat earth, surrounded by stony bodies in the heavens carried around by the cosmic vortex motion. The sun is a hot stone and the moon an earthy body. This cosmogony broadly follows the pattern set by Anaximander, and it shows the influence of Anaximenes in some details. In making the heavenly bodies spherical bodies, Anaxagoras may be following the pattern of Parmenides's cosmology.

ANAXAGORAS'S PRINCIPLES

What is innovative about Anaxagoras's theory is not the sequence of his cosmogony, but the principles on which he bases it. In the first place, he adheres to a principle of No Becoming—previously articulated by Parmenides—according to which nothing can come to be out of what is not. But whereas Parmenides seems to have meant this principle as a grounds for ruling out cosmological theories, Anaxagoras uses it as a restriction on what kind of explanation is allowed. Second, Anaxagoras follows a principle of Universal Mixture, which he states repeatedly, to the effect that everything is mixed with everything. There has been much controversy among interpreters about what the domain of "everything," in its two occurrences, is. Whatever the precise interpretation, the principle clearly applies to the primeval chaos insofar as all stuffs seem to be thoroughly mixed; but Anaxagoras maintains that even when the separation process takes over, some quantity of every stuff remains mixed with any given stuff. Third, Anaxagoras holds to a principle of Infinite Divisibility, according to which there are no minimal particles of matter—no atoms. Finally, Anaxagoras accepts a principle of Predominance, such that any stretch of matter manifests the properties of whatever stuff it has the largest quantity. Thus, if there is more water than salt in a mixture, people perceive it to be water; if more salt than water, they then perceive it to be salt.

It is generally agreed that the point of Anaxagoras's principles is to account for change with the least allowance for novelty: When one thing seems to change into another thing, the second does not arise out of nothing, but was already present (if in a lower concentration). Thus there is change, but no radical coming to be out of what is not—a possibility forbidden by Parmenides. What is less clear is whether Anaxagoras succeeds in formulating a coherent account of change. Whether he succeeds depends in large measure on how one interprets the details of his theories of matter and change, which will be discussed briefly below.

CONTROVERSIES

A fifth principle that is often attributed to Anaxagoras is Homoiomereity, using a Greek term of Aristotle's that designates a stuff in which the parts are like the whole. Thus if one divides a quantity of water in half, one gets two (smaller) quantities of water; but if one divides a chair in half one does not get two chairs. The former sort of being is called *homoiomerous*. Aristotle calls Anaxagoras's basic stuffs homoiomeries, but it is not clear whether he intends to say of them that they have the property of homoiomereity; or whether he simply wishes to denote things that in Aristotle's own system are homoiomeries (e.g., flesh and bone), whatever their properties for Anaxagoras. In any case, neither the term nor the property is found in the fragments of Anaxagoras—except as

the property is applied to Mind, which does not behave like a physical element.

Another controversy concerns the relationship between stuffs and contraries, or qualities in general. Anaxagoras mentions qualities such as hot and cold alongside stuffs such as earth (fr. 4) and maintains that contrary qualities cannot be cut off from each other (fr. 8). Does Anaxagoras recognize a strong categorial difference between stuffs and qualities, and if so, what is their relationship? According to one interpretation, the stuffs are composed of qualities, such that different amounts of hot, cold, wet, dry, and so on, combine to constitute different stuffs. Thus Universal Mixture signifies that every stuff is potentially in every stuff because by changing the ratios of qualities one can produce other results. On this view Anaxagoras is a reductionist who reduces stuffs to qualities. Defenders of this view have sometimes claimed that only on this interpretation can Anaxagoras's principles be rendered consistent. Yet, other interpreters have shown that his principles can be made consistent without reducing stuffs to qualities. Critics of the reductionist view see Anaxagoras's stuffs as elemental bodies. Qualities could be either stuffs in their own right, or simply properties that happen to describe certain stuffs.

Another controversial question is the meaning of the seeds Anaxagoras refers to in fr. 4 as being part of the original mixture. Are these biological seeds, as some interpreters hold, from which the first plants and animals grew? Or are they structural principles generally, to account for the presence of shapes and structures which emerge from the formless mixture (perhaps including, but not confined to, shapes of plants and animals)? Or are they small particles of a given stuff that are present as starting points for the growth of that given stuff? (A number of other hypotheses have also been advanced.) On many of these hypotheses, at least, no stretch of matter could be homoiomerous, for it would contain seeds having a different composition from the whole. Anaxagoras does not say enough about seeds to allow scholars to make a clear determination in favor of one of these hypotheses.

In another difficult saying in fr. 4, Anaxagoras talks about an alternative to "our"world. But is his statement merely counterfactual, or does he hold that there are other worlds, like the atomists; or worlds within worlds, as among Leibniz's monads; or repeating worlds, as does Empedocles? There is no more consensus on this question than on the other controversies mentioned.

MIND AND KNOWLEDGE

One of Anaxagoras's most interesting and innovative theories is his philosophy of mind. As has been shown, Anaxagoras makes mind responsible for the beginning of the cosmic vortex. He says that mind is "boundless, autonomous, and mixed with no object" (fr. 12). If it were not "by itself" it would be mixed with everything, by Universal Mixture, which would hinder it from ruling things. As it is, mind is "the finest of all objects and the purest, and it exercises complete oversight over everything and prevails above all" (fr. 12). Mind is present in some things, namely those things that have soul, but it does not mix with them. Thus mind is not immaterial, but it is not material in the same way as the stuffs are. It exercises control over the stuffs of the world and comprehends all things. Anaxagoras's theory suggests a dualism of mind and matter, though it is not nearly as radical as Descartes's dualism. In any case, Anaxagoras is the first philosopher to recognize mind as a distinct reality alongside physical entities. In cosmology, Anaxagoras is the first philosopher to support creationism—involving not a creation ex nihilo, to be sure, but an organization of preexisting elements by an intelligent agency distinct from those elements.

Anaxagoras accounts for sense perception as the effect of opposite qualities on opposites; thus one perceives hot by cold and wet by dry. He observes that because of the weakness of the senses people are not able to perceive the truth (fr. 21). But, on the positive side, "Appearances are a vision of the invisible" (fr. 21a). A serious philosophical problem for Anaxagoras is how humans can have knowledge at all. Because everything is mixed in everything, if I perceive something as water, I may infer that it is composed of water, and salt, and every other kind of stuff. But how can I say that I know the body before me as water if I have to analyze it as, among other things, water, and the water that I analyze it into is a theoretical entity I do not perceive? I seem to be involved in a regress that keeps me from knowing anything, except in a purely hypothetical way, in which everything has exactly the same components (all the stuffs there are), all of which are perceptually inaccessible to me.

Here, fr. 21a (just cited) provides a possible way out of the problem. People know by an inference to the best (or only possible) explanation that there are countless basic stuffs in the world. Further, they are acquainted with those stuffs by their manifestations to sense experience. Because when some stuff predominates, it gives its character to the whole body it predominates in, they can infer the character of, for instance, elemental water from

the character of phenomenal water in bodies of water they encounter. Similarly, people can infer the character of all other basic stuffs from their appearances, because, by hypothesis, the basic stuffs are like their phenomenal counterparts. People cannot give an adequate verbal definition of water, but they can give an ostensive definition of it. Thus they know the structure of reality by theory, but the content of reality by experience. This same sort of strategy appears to have appealed to Democritus, who approved of Anaxagoras's statement in B21a and applied it to his atoms (for a limited range of properties).

PHYSICAL THEORY

Though in some ways conventional, Anaxagoras's physical theories made some important advances. Like most of his predecessors, Anaxagoras envisioned a flat earth at the center of the circling heavens; the earth is held in place by air pressure, and the solstices of the sun are caused by winds in the heavens. He explained the annual floods of the Nile as the result of melting snows in southern Africa (the Nile is in fact fed by melting snows, but the floods are caused by monsoon rains), a view cited by Aeschylus, Sophocles, and Euripides, and criticized by Herodotus. He gave an essentially correct explanation of hail. His view that the heavenly bodies are earthy or stony was probably novel, and he believed that invisible stones were also carried about aloft with the vortex—in effect, he posited asteroids. When a large meteor fell at Aegospotami, Anaxagoras was given credit for predicting it, and henceforth meteors were included among data to be explained by cosmological theories. He gave the first correct explanation of solar and lunar eclipses (perhaps inspired by Parmenides's recognition that the moon gets its light from the sun), a feat that Aristotle regarded as a paradigm of scientific discovery. He also correctly hypothesized that the moon had mountains and valleys on its surface. In his physical theory he was followed by Archelaus of Athens, and in his teleological tendencies by Diogenes of Apollonia.

After Anaxagoras, natural philosophers mostly accepted his theory of eclipses and his view of heavenly bodies as spherical solid bodies. Though his astronomy was influential among intellectuals, it clashed with popular religious views according to which the sun and moon were gods, and led to an indictment of impiety in Athens. It was the sort of theory that Plato criticized in the *Laws* as leading to atheism. Anaxagoras presumably would reply that his views offered grounds for a more enlightened religion than those based on worshiping forces of nature.

Plato saw one of Anaxagoras's views as offering a new approach to cosmology. If Mind ordered everything with a view to the best, then philosophers should be able to explain the structure of the cosmos on the basis not of how it arose from a primeval chaos, but how it manifests order and value. Plato reports that Anaxagoras's book was disappointing because it failed to exploit this insight, and Aristotle agrees. In fact, Anaxagoras used the same style of explanation as other pre-Socratics stressing the natural capacities of different kinds of matter. But Plato later used Anaxagoras's insight to construct the cosmos on teleological principles in his *Timaeus*. In a sense, then, Anaxagoras provided the impulse for the rational cosmologies of Plato and Aristotle that dominated ancient and medieval thought. He was the first philosopher to make his home in Athens, and also the first to offend the Athenian people. Through the Athenian philosophical tradition he had a lasting influence.

See also Anaximander; Aristotle; Atomism; Cosmology; Descartes, René; Diogenes of Apollonia; Dualism in the Philosophy of Mind; Empedocles; Leucippus and Democritus; Nous; Parmenides of Elea; Philosophy of Mind; Plato; Pre-Socratic Philosophy; Sensa.

Bibliography

Barnes, Jonathan. *The Presocratic Philosophers*. Revised edition. Chaps. 16, 19. London: Routledge, 1982.

Graham, Daniel W. "The Postulates of Anaxagoras." *Apeiron* 27 (1994): 77–121.

Graham, Daniel W. "Was Anaxagoras a Reductionist?" *Ancient Philosophy* 24 (2004): 1–18.

Kerferd, G. B. "Anaxagoras and the Concept of Matter before Aristotle." In *The Pre-Socratics: A Collection of Critical Essays*, edited by Alexander P. D. Mourelatos, 489–503. Garden City, NY: Doubleday, 1974.

Kirk, G. S., J. E. Raven, and M. Schofield. *The Presocratic Philosophers*. 2nd edition, chap. 12. Cambridge, U.K.: Cambridge University Press, 1983.

Lesher, James H. "Mind's Knowledge and Powers of Control in Anaxagoras DK B12." *Phronesis* 40 (1995): 125–142.

Schofield, Malcolm. *An Essay on Anaxagoras*. Cambridge, U.K.: Cambridge University Press, 1980.

Sider, David. *The Fragments of Anaxagoras*. Meisenheim am Glan, Germany: Verlag Anton Hain, 1981.

Taylor, C. C. W. "Anaxagoras and the Atomists." In *Routledge History of Philosophy*, vol. 1: *From the Beginning to Plato*, edited by C. C. W. Taylor, 208–243. London: Routledge, 1997.

Vlastos, Gregory. "The Physical Theory of Anaxagoras." *Philosophical Review* 59 (1950): 31–57. Reprint in Gregory Vlastos's *Studies in Greek Philosophy*, vol. 1, 303–327. Princeton, NJ: Princeton University Press, 1995.

Daniel W. Graham (2005)

ANAXIMANDER

(c. 610 BCE–after 546 BCE)

Anaximander is the first Greek scientist and philosopher whose thought is known to us in any detail. He was born in Miletus c. 610 BCE and died shortly after 546 BCE. He was thus in his twenties in 585 BCE, the year of the famous solar eclipse that Thales is said to have predicted. According to the ancient tradition, Anaximander was the "pupil and successor of Thales"; but in view of our ignorance of Thales' real achievements, it is perhaps Anaximander who should be considered the founder of Greek astronomy and natural philosophy. Nothing is known of his life except an unverifiable report that he led a Milesian colony to Apollonia, on the Black Sea. His lifetime corresponds with the great age of Miletus, when it was the richest and most powerful Greek city in Asia Minor.

His scientific achievements are said to include the first Greek world map, the first Greek star map or celestial globe, and the invention, or rather adaptation, of the gnomon (the vertical pointer of a sundial) for use in measuring the hours of the day and annual variations in the course of the sun. According to Pliny, he also traced the sun's annual path in the ecliptic and noted its inclination with regard to the celestial axis. This last discovery may really belong to a later age, but there is no doubt that Anaximander conceived (and almost certainly constructed) a spherical model for the heavens, in the center of which was placed Earth, as a disk or cylinder whose height was one-third its diameter. The ratio 1:3 seems also to have been used in the spacing of the celestial circles or rings assigned to stars, moon, and sun: The conjectural sizes for these rings are 9, 18, and 27 Earth diameters, respectively. (His strange error in assigning the lowest circle to the stars is unexplained. There is, unfortunately, no evidence to support J. Burnet's attractive suggestion that this circle corresponds not to the fixed stars but to bright planets such as Venus. If we could accept this, the fixed stars might then be assigned to their natural place at the periphery of the celestial sphere.)

Anaximander is thus the author of the first geometrical model of the universe, a model characterized not by vagueness and mystery but by visual clarity and rational proportion, and hence radically different in kind from all known "cosmologies" of earlier literature and myth. The highly rational character of the scheme (despite its factual errors) is best indicated by Anaximander's explanation of Earth's stable position in the center: It remains at rest because of its equal distance from all points of the celestial circumference, having no reason to move in one direction rather than in another. This argument from symmetry contrasts not only with all mythic views but also with the doctrine ascribed to Thales: that Earth floats on water. Here Anaximander is clearly the precursor of the mathematical approach to astronomy developed later by the Pythagoreans, Eudoxus, and Aristarchus.

The book of Anaximander, quoted later under the standard title *On the Nature of Things (peri physeôs)*, seems to have contained a description of his map and celestial model, as well as an account of how the natural world functions and how it reached its present form. Beginning from a first principle called the Boundless or Infinite (*to apeiron*: see below), he describes how "something capable of generating Hot and Cold was separated off … and a sphere of fire from this source grew around the air in the region of earth like bark around a tree. When this sphere was torn off and enclosed in certain rings, the sun and moon and stars came into existence" (Diels-Kranz, 12 A 10). These heavenly bodies are "wheel-like, compressed masses of air filled with fire, which exhale flames from an orifice at one point" (Diels-Kranz, 12 A 17a).

Eclipses and lunar phases are explained by obstruction of the orifices. The sea is what remains of the primeval moisture, the rest having been evaporated as air or dried up by the celestial fire to form Earth. Land, sea, air, and heavens are thus all explained by a continual process of separating off from the primeval pair of Hot (dry) and Cold (wet). Wind, rain, lightning, thunder, and related phenomena are explained by the interaction of these elemental principles (water, air, fire) and opposite powers (hot, cold; dry, moist; thick, thin; light, dark). The origin of living things is explained as part of the same process: They arose as aquatic beings in moisture and later transferred to dry land. The first examples of each species developed to maturity within a protective membrane. In an interesting anticipation of modern ideas, Anaximander remarked that the first human beings could never have survived as helpless infants, but must have been born "from living things of another kind, since the other animals are quickly able to look for their own food, while only man requires prolonged nursing" (Diels-Kranz, 12 A 10).

The one quotation from Anaximander's book that seems to have been preserved in very nearly the original wording is his famous statement on cosmic justice: "Out of those things whence is the generation of existing things, into them also does their destruction take place, as is right and due; for they make retribution and pay the penalty to one another for their offense [or "injustice,"

adikia], according to the ordering of time." The interpretation of this oldest surviving philosophic text has been a subject of much controversy. The earlier commentators (including Friedrich Nietzsche) interpreted the "injustice" as the separation of individual things from their infinite source and saw the eventual reabsorption of all things back into the *apeiron* as their only fitting atonement. This fails to explain how the things that perish can pay the penalty *to one another*, or why the source of generation is referred to in the plural. It is now generally agreed that offense and compensation must both refer to the strife of opposing principles (such as the hot and cold), and that the "ordering of time" stands primarily for periodic regularity in the daily and seasonal variation of heat, moisture, daylight, and the like. Whether there is also a reference here to a larger cycle in which the cosmos itself would perish into its source is more doubtful.

Anaximander's fame rests chiefly on his doctrine of the Boundless as the *arche*, the starting point and origin of the cosmic process. For him, the term *apeiron* meant "untraversable" or "limitless" rather than "infinite" in any precise mathematical sense. He described this principle with the Homeric epithets for divinity, calling it "ageless and immortal," and probably even "the divine" (*to theion*). This *apeiron* surrounds and embraces all things and apparently "steers" or governs them as well. It seems to have been conceived as ungenerated as well as imperishable, and thus contrasts in every respect with the limited, perishable world it engenders. Our sources refer to "worlds" (*kosmoi*) in the plural; a succession rather than a simultaneous plurality of worlds seems to be meant. The Boundless transcends this process of world creation, circumscribing each individual world in space, outlasting all of them in time, and providing the inexhaustible material source, the eternal motive power, the vital energy, and (presumably) the geometrical form and cyclical regularity for the cosmic process as a whole. In its archaic complexity, the *apeiron* is thus both a physical and a metaphysical or theological concept, and points the way not only to the infinite void of the atomists but also to the cosmic deity of Xenophanes, Aristotle, and the Stoics.

See also Thales of Miletus.

Bibliography

ANCIENT EVIDENCE

Conche, Marcel. *Anaximandre, fragments et témoignages: Texte grec, traduction, introduction & commentaire*. Paris: Presses Universitaire de France, 1991.

Diels, Hermann, and Walther Kranz. *Die Fragmente der Vorsokratiker*. 6th ed. Berlin: Weidmann, 1951. Vol. I, Ch. 12.

Kirk, Geoffrey S., John E. Raven, and Malcolm Schofield, Malcolm. *The Presocratic Philosophers*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983, Ch. 3.

IMPORTANT SPECIAL STUDIES

Asmis, Elizabeth. "What Is Anaximander's *Apeiron*?" *Journal of the History of Philosophy* 19 (1981): 279–297.

Barnes, Jonathan. *The Presocratic Philosophers*. 2 vols. London: A & C. Black, 1979.

Burnet, J. *Early Greek Philosophy*. 3rd ed. London: A. and C. Black, 1920.

Classen, C. Joachim. "Anaximander and Anaximenes: The Earliest Greek Theories of Change." *Phronesis* 22 (1977): 89–102.

Couprie, Dirk L., Robert Hahn, and Gerard Robert. *Anaximander in Context: New Studies in the Origins of Greek Philosophy*. Albany: State University of New York Press, 2003.

Engmann, Joyce. "Cosmic Justice in Anaximander." *Phronesis* 36 (1991): 1–26.

Freudenthal, Gad. "The Theory of the Opposites and an Ordered Universe: Physics and Metaphysics in Anaximander." *Phronesis* 31 (1986): 197–228.

Hahn, Robert. *Anaximander and the Architects: The Contributions of Egyptian and Greek Architectural Technologies to the Origins of Greek Philosophy*. Albany: State University of New York Press, 2001.

Kahn, Charles H. *Anaximander and the Origins of Greek Cosmology*. New York: Columbia University Press, 1960. Reprint, Indianapolis, IN: Hackett, 1994.

McKirahan, Richard. "Anaximander's Infinite Worlds." In *Essays in Ancient Greek Philosophy VI: Before Plato*, edited by Anthony Preus, 49–65. Albany: State University of New York Press, 1989.

Robinson, J., "Anaximander and the Problem of the Earth's Immobility." In *Essays in Ancient Greek philosophy I*, edited by John P. Anton and George L. Kustas, 111–118. Albany: State University of New York Press, 1971.

Shelley, Cameron. "The Influence of Folk Meteorology in the Anaximander Fragment." *Journal for the History of Ideas* 61 (2000): 1–17.

Vlastos, Gregory. "Equality and Justice in the Early Greek Cosmogonies." *Classical Philology* 42 (1947): 156–178.

Charles H. Kahn (1967)
Bibliography updated by Christian Wildberg (2005)

ANAXIMENES

(6th century BCE)

Anaximenes was the third and last member (the others were Thales and Anaximander) of what is traditionally called the Milesian school of natural philosophers (*physiologoi*). The date of his death is estimated 528/526 BCE; it is probable that he "flourished" about 545 BCE. Although little is known about his life and work, fragments of ancient testimony credit him with studies under his older contemporary, Anaximander; with the writing

of a book in "simple Ionic"; and with the doctrine that air is the underlying principle of the universe, changes in physical state being the result of its condensation and rarefaction. It is likely that Aristotle read Anaximenes' book and that Theophrastus had access to it. Several of the doxographers (Aëtius, Hippolytus, Diogenes Laertius) may have read later Hellenistic versions of the work.

On the strength of ancient testimony, historians of philosophy after Aristotle regarded Anaximenes' doctrine as a contribution to the Milesian debates on Nature. They assumed that from Thales to Anaximenes there was a continuous development in physical thought, and they insisted that this development was intelligible only in terms of the supposedly unique problem of the period: the birth and structure of the physical world. On this interpretation, Anaximenes' air was taken to be an *arche,* and his condensation-rarefaction doctrine was construed as a theory about physical transformations. The physical system reconstructed along these lines was then usually shown to be, in comparison with that of Anaximander, not as cogent; and whatever could not be accommodated within such a reconstruction was relegated to Anaximenes' "retrogressive astronomy."

Recent studies in mythical and early cosmogonic discourse (Hesiod) perhaps call for some revision of the traditional estimate. At a time when mechanical change and biological growth had not yet been distinguished from each other, when physical permanence was regarded as incomprehensible apart from "justice" between the warring Opposites, when inanimate continuity was mistaken for animal kinship, when experience was permitted only to illustrate but never to refute supposed insight, when meteorology served as the foundation for astrophysics—several of Anaximenes' ideas were pioneering. A schematic reconstruction of some of these ideas follows.

The fundamental and most pervasive thing in the world is air (*aer*), according to Anaximenes. Air is infinitely vast in extent but perfectly determinate in character: It is ordinary atmospheric air, invisible where most even in consistency, visible through the Hot and Cold and Damp and motion. It is from air that all the things that exist, have existed, or will exist come into being. This applies to gods and divine things and also to the rest of the world, inasmuch as the world is compounded out of the offspring of air. On this account, Anaximenes suggests, the primordial air is continually in motion, and this motion is the cause of alternating physical states. Condensation and rarefaction are the key manifestations of changing air: rarefied air generates fire; condensed air creates winds; condensed winds, clouds; condensed clouds, water; condensed water, earth; earth, stones and the rest of the world.

Throughout the process of cosmic change, the Hot and the Cold are dominant states of physical activity, but in no way are they forces distinct from air. They never come out of air by "separating off" (*ekkrisis*); rather, they are "attributes" of air when it condenses through "felting" or is rarefied through "loosening up."

From the genesis of the universe at large, Anaximenes moves to the description of the shape of Earth and of the visible sky. Earth, according to him, is broad, flat, and shallow—tablelike. All the heavenly bodies are fires in the sky, caused by the moist exhalations of Earth. The heavenly bodies are nailed on a hemispherical diaphanous membrane and move around Earth like a cap that can be turned around one's head, and not under Earth. The stars do not produce any sensible heat because of their distance. When the sun, moon, and stars disappear, they are hidden by the distant elevations of Earth. The stars may also be likened to fiery leaves floating on the air.

Clouds, rain, hail, and snow—all these phenomena, too, are caused by condensed air. And the same is true of the violent breaks of the clouds that produce lightning and thunder.

With the elements of his cosmology worked out, Anaximenes seems to need a general natural law guaranteeing the regularity of the world. He observes that as our souls, being air (according to an ancient tradition), hold us together, so does the cosmic Air hold the world together by enclosing it. Presumably what Anaximenes meant by this was that the regularity of an animated world is reliable and intelligible, as is the regularity of an animated body, a body that is organically self-regulative and autonomous—a microcosm. For Anaximenes, law-like regularities were inconceivable without access to the idea of cause. The notion of physical constraint was accordingly effected through containment. The divine Air, by encasing the world, successfully regulates it.

See also Anaximander; Aristotle; Diogenes Laertius; Pre-Socratic Philosophy; Thales of Miletus.

Bibliography

"Anaximenes." In Pauly-Wissowa, *Realenzyklopädie der altertums Wissenschaft,* Vol. I, p. 2086. Stuttgart, 1893.

Burnet, J. *Early Greek Philosophy,* 4th ed. London: Black, 1930. Pp. 72–79.

Coxon, A. H. "Anaximenes." In *Oxford Classical Dictionary.* London: Clarendon Press, 1949. P. 51.

Diels, Hermann, and Walther Kranz, eds. *Fragmente der Vorsokratiker,* 5th ed. Berlin: Weidmannsche Verlagbuchhandlung, 1954. Vol. I, pp. 90–96.

Guthrie, W. K. C. "Anaximenes and *τὸκρυσταλλοειδές.*" *Classical Quarterly* (1956): 40–44.

Guthrie, W. K. C. *A History of Greek Philosophy.* Cambridge, U.K.: Cambridge University Press, 1962. Vol. I, pp. 115–150.

Heidel, W. A. "The *δίνη* in Anaximenes and Anaximander." *Classical Philology* (1906): 279–282.

Kerferd, G. B. "The Date of Anaximenes." *Museum Helveticum* (1954): 117–121.

Kirk, G. S., and J. E. Raven. *The Presocratic Philosophers.* Cambridge, U.K.: Cambridge University Press, 1957. Pp. 143–162.

Tannery, P. "Un fragment d'Anaximene." *Memoires scientifiques* 8 (1925).

P. Diamandopoulos (1967)

ANCIENT AESTHETICS

In antiquity, aesthetics did not form a distinct branch of philosophy. Ancient philosophers discussed literature, music, and the visual arts and reflected on the nature of beauty in a variety of contexts. Since the Greek word for "beautiful" or "fine," *kalos*, is a very general value term that can also be used to describe what is morally admirable, ancient discussion of beauty is often embedded in wider-ranging discussion of values. Literature, music, and the visual arts are frequently considered in an educational and political context; at the same time, most ancient philosophers' views about the arts are strongly influenced by other aspects of their philosophy, in particular their metaphysics.

The earliest Greek philosophy does include some suggestive remarks on aesthetic topics, notably some comments by Gorgias in his *Encomium of Helen*, written in the fifth century BCE, about the power of speech. However, in aesthetics as elsewhere, it was Plato and Aristotle who set the philosophical agenda for all subsequent discussion. We shall therefore begin with Plato and Aristotle and shall trace the development of what we now call aesthetics through the Hellenistic and Roman periods into late antiquity.

PLATO

Plato raises questions about the arts, and about beauty, in a number of different dialogues. Poetry is the art to which he devotes the most discussion, but this entry will also discuss his attitude to rhetoric, his use of the visual arts to illustrate his arguments about both poetry and rhetoric, his comments on music, and finally his view of beauty.

Plato alludes to "an old quarrel" between philosophy and poetry (*Republic* 10.607B). He saw dangers in the widespread use of Homer in classical Greek education and in the role played by tragic drama in classical Athens, a role comparable to that of the mass media in modern society. He therefore argues in the *Ion* and in *Republic* 10 that poets, unlike philosophers, do not have knowledge, and in *Republic* 2 and 3 he places strict limits on the amount of Homer that the future guardians of his ideal state may read and on the type of dramatic performance in which they may take part. In *Republic* 3 he describes drama as imitation (*mimesis* in Greek) and regards both acting and viewing drama as dangerous, both because playing a variety of different roles can destabilize the personality and because imitation of evil characters may likewise make us evil. Since poets lack knowledge, their poetry, according to Plato, appeals not to reason but to the emotions. This point recurs in the *Ion*, in *Republic* 2 and 3, and in *Republic* 10, where it is made using the theory of three parts of the soul first set out in *Republic* 4.

Traditionally Greek poets claimed to be inspired by the Muses. Plato too regards poets as inspired, in the *Ion* and elsewhere, but since such inspiration is contrasted with knowledge, it may not be worth much. However, he does suggest at *Phaedrus* 245A that inspired poetry is more valuable than poetry produced by technical skill alone.

In *Republic* 10, Plato puts forward perhaps his most famous and influential argument to distinguish poetry from philosophy, using the metaphysics developed in the central books of the *Republic*. According to that metaphysics, the physical world is only an imitation (*mimesis*, again) of a world of transcendent Forms. In *Republic* 10, Plato suggests that painters simply copy objects in the physical world and are thus at two removes from the true reality of the world of Forms. The point is then immediately applied to poets, who are regarded as low-grade copyists of the same kind. The scope of *mimesis* is now much wider than in *Republic* 3, where it applied only to drama; here Plato treats virtually all poetry as mimetic and so banishes it from his ideal state.

Plato is as harshly critical of rhetoric as he is of poetry, and for similar reasons. In classical Athens, teachers of rhetoric were popular and rhetorical skill was widely seen as the passport to a successful political career. Many of the Sophists, such as Gorgias and Thrasymachus, taught rhetoric. Plato regularly sets up an opposition between the Sophists, as false teachers, and his own mentor, Socrates; in dialogues such as the *Gorgias*, he contrasts the persuasive power of rhetoric, aimed only at

pleasing the audience, with philosophy, which aims for knowledge of the truth. Similarly, in the *Sophist*, at 235Bff., Plato defines the Sophist as a maker of images, comparing his techniques to those used by sculptors and painters. In the *Phaedrus*, however, although Socrates criticizes severely a speech said to be by the orator Lysias, he also raises the possibility that there could be an ideal kind of rhetoric, based on knowledge.

Plato makes occasional remarks about music and in *Republic* 3.398Cff. proposes to regulate the music to which the future guardians of the ideal state may listen, just as he regulates the poetry that they may study. He assumes that music, like poetry, affects the emotions, and he distinguishes between musical modes such as the Dorian and the Lydian on ethical grounds: the future guardians should listen to music that will make them brave and warlike, not to music that will encourage excessive indulgence in unmanly emotions such as grief.

When Plato discusses poetry, rhetoric, and music, sometimes using the visual arts to illustrate and support his argument, he says little or nothing about beauty. He considers beauty in a quite different context in *Symposium* 210ff. where Socrates, speaking in praise of Love (*Erōs* in Greek), reports what he says he was told by a wise woman, Diotima. This passage describes, in lyrical, poetic language, a progression from the love of physical beauty to the love of moral and intellectual beauty and finally to the Form of Beauty itself. Plato here makes no direct reference to the arts, but it is worth noting that in the *Phaedrus* too he recognizes love as a powerful but nonrational motive force in the human soul. The *Phaedrus* also contains a mythical account of how the human soul, before it entered the body, was able to see the Forms, including the Form of Beauty. As we have seen, the *Phaedrus* includes some favorable comments on inspired poetry and the suggestion that an ideal rhetoric, based on knowledge, could be devised. It is therefore tempting to suggest that the right kind of poetry and rhetoric could find a place among the moral and intellectual beauties mentioned in the *Symposium*. Yet we should note that even if this is correct, such beauties will still be left behind by the soul that ascends to the Forms, the ultimate object of philosophical inquiry.

ARISTOTLE

Whereas Plato always discusses poetry and the other arts within a broader context, Aristotle devotes the *Poetics* solely to an examination of poetry. In fact the scope of the *Poetics* as we have it is narrower still: after some introductory remarks about the nature of poetry in general, Aristotle concentrates on tragedy and epic; a lost second book was devoted to comedy. Although the *Poetics* is the main source for Aristotle's aesthetic thought, there is a brief but important discussion of music in the *Politics* that supplements the single allusion in the *Poetics* to *katharsis*, and his views on rhetoric, expounded in the *Rhetoric*, are also of interest.

Like Plato, Aristotle regards poetry as a form of *mimesis*, or "imitation," but since Aristotle's metaphysics differs radically from Plato's, his understanding of *mimesis* is also radically different. For Aristotle forms are immanent in matter, not transcendent. Poetry imitates the world around us, and Aristotle is happy to accept both that we enjoy recognizing such imitation and that we can learn from it. Tragedy, for Aristotle, is an imitation of an action, and Aristotle focuses not on the characters represented but on the plot. Although he does discuss what kind of tragic hero is best, for example, his concern is primarily with what makes a good play. For that reason he has often, with some justice, been regarded as the first formalist in literary theory. He stresses the importance of a unified plot, arguing, for instance, in 1459a that Homer's *Iliad* and *Odyssey* are superior to other epics such as the *Cypria* or the *Little Iliad* in being less episodic. He illustrates his argument with many examples from classical Greek plays, particularly Sophocles' *Oedipus Tyrannus*, which he admires as a supreme example of a well-constructed tragedy.

Yet Aristotle's approach to poetry is not purely formal. He regards the action and the characters of a tragedy as morally significant and believes that poetry can convey universal truths, claiming, at 1451b, that it is closer to philosophy than to history in that respect. Like Plato he recognizes that poetry has a powerful effect on the emotions and like Plato he holds that tragedy arouses both pity and fear. However, whereas Plato, in *Republic* 10 and elsewhere, argues that tragedy and other forms of poetry overstimulate these emotions, Aristotle has a more complex view. When he gives a definition of tragedy in *Poetics* 1449b, he describes it as bringing about a *katharsis* of pity and fear and in *Politics* 8. 1341bff., in a discussion of music, he mentions a similar *katharsis* effected by the healing use of music in certain religious rites. There has been much scholarly discussion of just what Aristotle means by *katharsis*. Arguably it is best understood in the light of Aristotle's ethics: Aristotle holds that in order to act virtuously we need to feel the right emotions at the right time, in the right way and toward the right objects; in some way that is not fully explained, our feeling pity and fear as we watch a good tragedy brings about the

result that, when we leave the theater, we feel not too much pity and fear, as Plato supposed, but just the right amount that we need for ethical action.

The rest of Aristotle's discussion of music in *Politics* 8 assumes, as Plato did in the *Republic*, that music has a powerful effect on the emotions. He criticizes some details of Plato's argument in *Republic* 3.398Cff., and by introducing the notion of *katharsis*, Aristotle opens up the possibility that music can be used for therapeutic purposes.

A similar interest in the effect of art on the emotions can be seen in Aristotle's *Rhetoric*. Aristotle devotes much of *Rhetoric* 2 to a discussion of the emotions because the orator will need to understand his audience's emotional responses in order to persuade them effectively. The *Rhetoric* also contains important discussions of rhetorical reasoning and of prose style. Just as the *Poetics* is not a handbook for poets but a philosophical treatise on poetry based on close study of tragedy and epic, so the *Rhetoric* is not a handbook for orators but a philosophical treatise based on close study of rhetorical practice.

THE HELLENISTIC AND ROMAN PERIODS

After the death of Aristotle, Greek philosophy became increasingly diverse. While the Platonist and Aristotelian schools continued, the Epicureans and the Stoics developed new approaches to many issues. Aristotle's pupil, Theophrastus, was interested in the therapeutic powers of music and claimed that music could even cure bodily afflictions such as sciatica. Another pupil of Aristotle, Aristoxenus, studied music from an empirical point of view, opposing the mathematical approach that had been taken by the Pythagoreans.

The Stoics regarded both the order of the universe and moral virtue as beautiful, and their interest in the philosophy of language led them to discuss poetry and rhetoric. They thought poetry could express truth, as we can see from Cleanthes' choice of verse to convey his philosophy in the *Hymn to Zeus* and from the way in which critics such as Heraclitus and Cornutus used allegorical interpretation of poetry and mythology. By contrast, Epicurus appears to have rejected the idea that poetry could have any value as a means of instruction, although he was prepared to accept that it could be a source of pleasure.

In the first century BCE, Philodemus, an Epicurean, wrote his important works *On Poems* and *On Music*, which survive only in fragmentary form in papyrus scrolls found at Herculaneum. Much of Philodemus's work took the form of attacks on other critics and theorists. His own view was that poetry, and music, do not give pleasure by their sound alone. Music at this time was normally an accompaniment to poetry, and Philodemus holds that the value of music comes from the poetry that is performed with it, and the value of that poetry comes from the thought that it expresses; he also holds that form and content go closely together and that a poem cannot be good in thought but bad in composition.

Although Philodemus influenced the Roman poets Virgil and Horace, Epicurean views remained outside the mainstream of thinking about the arts in the first century BCE and the first century CE. Many educated writers of this period combine together ideas from more than one philosophical school. In both Cicero (*Orator* 8) and Seneca (*Letters* 58 and 65) we find an important new idea about the metaphysical status of works of art. Both these writers suggest that rather than merely imitating objects in the physical world, which are themselves copies of transcendent Forms, the artist looks to ideas in his own mind, which are themselves reflections of the Platonic Forms, understood by the Platonists of this period as the thoughts of God. The Greek sculptor Phidias, famous for his statue of the supreme god, Zeus, is used as an example of an artist who worked in this way.

Throughout the Hellenistic and Roman periods there continued to be great interest in the moral effect of the arts and the role of the arts in education. Plutarch (c. 45–c. 120 CE) discusses poetry from a moral point of view in his *De audiendis poetis*, in a way that reflects the continuing influence of Plato's views. He is familiar with the idea that music can be used as psychological therapy and associates this with Pythagoreanism in *De Iside* 384A.

The Pythagoreans, as we saw earlier, were also credited with a mathematical approach to music. Ptolemy's *Harmonics*, written in the second century CE, contrasts Pythagorean and Aristoxenian views of music. Ptolemy agrees with the Pythagoreans that musical structures must be analyzed in mathematical terms but criticizes them for neglecting empirical, perceptual evidence.

Literary criticism in the Hellenistic and Roman periods was closely intertwined with the theory and practice of rhetoric. Writers such as Cicero and Quintilian discuss literary and aesthetic matters in the context of rhetorical education. The work *On the Sublime* attributed to Longinus, which probably dates from the first century CE, blends ideas drawn from the rhetorical tradition of literary criticism with ideas drawn from philosophy, particularly from Platonism. The work is unusual among surviving ancient works of literary criticism both because

the author develops the view that the best works of literature have an essential quality of sublimity that explains their enduring appeal and because he illustrates his view with detailed discussion of examples in a way that combines technical analysis with judgment of literary value.

LATE ANTIQUITY

Plotinus, writing in the third century CE, regarded himself as a Platonist but is now labeled rather a "Neoplatonist" because he elaborated a more complex metaphysics than previous Platonists, postulating a transcendent One beyond the realm of the Platonic Forms. In aesthetics, Plotinus combined the suggestion that the artist uses ideas in his own mind that directly reflect the Forms, already found in Cicero and Seneca, with the account of the ascent to the Form of Beauty offered in Plato's *Symposium*. *Ennead* 1.6 begins by rejecting a Stoic account of beauty as symmetry of parts, arguing that incomposite things can also be beautiful and that they derive their beauty from a higher source. Plotinus then draws on Plato's *Symposium* and *Phaedrus* to describe an ascent from physical beauty through moral and intellectual beauty to the Form itself, and ultimately to the One beyond the Form of Beauty. *Ennead* 1.6 has often been regarded as presenting an aesthetic theory, but we must recognize, first, that Plotinus is not talking just about "beauty" in the modern sense and, second, that his theory implies that beauty in the physical world is to be valued only insofar as it leads us to a higher realm. As noted at the beginning of this entry, the Greek word *kalos*, standardly translated as "beautiful," is a very general value term. It would be a mistake to say that Plotinus is aestheticizing morality when, like the Stoics and Plato before him, he describes moral virtue as *kalos*; it would be more correct to say that, like most ancient thinkers, he makes no distinction between aesthetic and moral value. It is also important to recognize that for Plotinus our ultimate goal is union with the One; intellectual contemplation is the next best thing, and appreciation of beauty is only a means to achieving these goals, not something valued for its own sake.

Plotinus says little or nothing about art in *Ennead* 1.6, but in *Ennead* 5.8.1 he combines the view of beauty found in 1.6 with the suggestion that the artist can imitate the Forms directly, using principles in his own mind that derive from the Forms. He uses the standard example of Phidias's statue of Zeus and suggests, very politely, that Plato's argument in *Republic* 10 is mistaken in representing works of art as imitating only objects in nature. According to Plotinus's argument, art itself is superior to its products, and he moves on in the rest of 5.8 to discuss the intellectual beauty of the world of Forms.

Although Plotinus himself shows only limited interest in the arts, his view of beauty led to important developments in poetic and musical theory. His views were applied to poetry by the later Neoplatonist Proclus, in the fifth century CE. In his *Commentary on the Republic*, Proclus argues that much of Homer's poetry is not after all vulnerable to Plato's criticisms, since it is not mimetic but inspired. Just as Phidias's Zeus, for both Plotinus and Proclus, portrays the god, capturing something of divine beauty in the statue we see, so Homeric poetry conveys truths about the divine world of Neoplatonic metaphysics. In order to maintain this view of Homer, Proclus resorts to allegorical interpretation of episodes criticized by Plato, drawing on a long tradition of such interpretation by Stoics and others.

Proclus and other later Neoplatonists also devoted attention to music. On the one hand, they integrated traditional views about the effect of music on the emotions into their philosophical system. On the other, they regarded music as one of the mathematical sciences, following a Pythagorean rather than an Aristoxenian approach. They perceived the same mathematical patterns in music as in the physical universe, believing that the beauty of such perceptible order derived from the ordered structure of the intelligible world. The *Institutio musica* of Boethius (c. 480–c. 524 CE), written in Latin, draws heavily on these ideas.

See also Aesthetics, History of; Aristotle; Boethius; Anicius Manlius Severinus; Gorgias of Leontini; Kalon; Katharsis; Mimesis; Neoplatonism; Philodemus; Plato; Plotinus; Proclus.

Bibliography

TEXTS AND TRANSLATIONS

Aristotle. *Ars rhetorica*. Edited by Rudolf Kassel. Berlin and New York: de Gruyter, 1976.

Aristotle. *De arte poetica* (Oxford Classical Text). Edited by Rudolf Kassel. Oxford: Oxford University Press, 1965.

Aristotle. *On Rhetoric: A Theory of Civic Discourse*. Translated by George A. Kennedy. New York and Oxford: Oxford University Press, 1991.

Aristotle. *Poetics*. Translated by Malcolm Heath. London: Penguin, 1996.

Aristotle. *Politica* (Oxford Classical Text). Edited by W. D. Ross. Oxford: Oxford University Press, 1957.

Barker, Andrew D. *Greek Musical Writings*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1984, 1989.

Boethius. *De institutione arithmetica; De institutione musica*. Edited by Gottfried Friedlein. Leipzig: Teubner, 1867.

Boethius. *Fundamentals of Music*. Translated by Calvin M. Bower. New Haven and London: Yale University Press, 1989.
Longinus. *On the Sublime*. Edited by Donald A. Russell. Oxford: Oxford University Press, 1968.
Philodemus. *On Poems Book One*. Edited by Richard Janko. Oxford: Oxford University Press, 2000.
Plato. *Complete Works* (in English translation). Edited by John M. Cooper. Indianapolis and Cambridge: Hackett, 1997.
Plato. *Opera* (Oxford Classical Text). Edited by John Burnet. Oxford: Oxford University Press, 1902–1906.
Plato. *Opera*, vol. 1 (Oxford Classical Text). Edited by E. A. Duke and others. Oxford: Oxford University Press, 1995.
Plato. *Republic* (Oxford Classical Text). Edited by Simon R. Slings. Oxford: Oxford University Press, 2003.
Plotinus. *Enneads* (Loeb Classical Library). Edited by A. Hilary Armstrong. Cambridge, MA, and London: Harvard University Press and Heinemann, 1966–1988.
Russell, Donald A., and Michael Winterbottom. *Ancient Literary Criticism: The Principal Texts in New Translations*. Oxford: Clarendon Press, 1972.

STUDIES
Chadwick, Henry. *Boethius. The Consolations of Music, Logic, Theology, and Philosophy*. Oxford: Clarendon Press, 1981.
Halliwell, Stephen. *The Aesthetics of Mimesis*. Princeton, NJ, and Oxford: Princeton University Press, 2002.
Halliwell, Stephen. *Aristotle's Poetics*. London: Duckworth, 1986.
Janaway, Christopher. *Images of Excellence: Plato's Critique of the Arts*. Oxford: Clarendon Press, 1998.
Janko, Richard. *Aristotle on Comedy: Towards a Reconstruction of Poetics II*. London: Duckworth, 1984.
Kennedy, George A., ed. *The Cambridge History of Literary Criticism*, Vol. 1. Cambridge, U.K.: Cambridge University Press, 1989.
Lamberton, Robert. *Homer the Theologian*. Berkeley and Los Angeles: University of California Press, 1986.
Moravcsik, Julius, and Philip Temko, eds. *Plato on Beauty, Wisdom and the Arts*. Totowa, NJ: Rowman and Littlefield, 1982.
Obbink, Dirk, ed. *Philodemus and Poetry: Poetic Theory and Practice in Lucretius, Philodemus, and Horace*. Oxford: Oxford University Press, 1995.
Russell, Donald A. *Criticism in Antiquity*. London: Duckworth, 1981.

Anne Sheppard (2005)

ANCIENT SKEPTICISM

Tradition recognizes two schools of ancient skepticism: the Academics and the Pyrrhonists. The ancient Greek term "skeptic" was used by the Pyrrhonists to describe themselves. They denied that it described the Academics, but this point could be and was disputed, and later in antiquity the word may have been used as a common designation for both schools. Our use of the term in this way goes back to the seventeenth century.

The term itself is derived from a verb in common use meaning "to inquire" or "to investigate"—hence the skeptic as inquirer. This is a surprise. We take skepticism, roughly speaking, to imply a denial of the possibility of knowledge. Yet Sextus Empiricus, the second-century CE Pyrrhonist—and the only member of the school whose works have survived intact and in bulk—is quite firm on this point. In the opening chapter of his *Outlines of Pyrrhonism*, he distinguishes three types of philosophers: those who take themselves to have discovered the truth, those who hold that it cannot be apprehended, and those who persist in inquiring. Philosophers of the first type he calls "dogmatists," members of the last group "skeptics," and those of the middle tendency "Academics."

This is unfair. Even Academics like Philo of Larissa, who did hold that nothing can be apprehended, did not conclude from this that inquiry was pointless. Though they held that certain knowledge was unobtainable, they believed that it was possible to identify views that enjoyed a high degree of probability or verisimilitude—among them, the view that nothing can be known for certain—and they regarded inquiry for the sake of such discoveries as eminently worthwhile. What is more, Academics like Carneades and Clitomachus were no more convinced that nothing can be known than the Pyrrhonists, and they and deserved to be described as inquirers at least as much.

These facts only add to the puzzle, however. If not only the Pyrrhonists but also many Academics were skeptics in Sextus's sense, why the persistent tendency, beginning with the ancient skeptics' own contemporaries, to equate skepticism with one of the positions that Sextus expressly opposes to it? And why should a dedication to inquiry set the skeptics apart from members of other schools? Philo of Alexandria, who was active in the first century CE, was able to use the term "skeptikos" (in the sense of "inquirer") of philosophers quite generally.

Sextus's idea seems to be this: Inquiry into a particular question comes to a natural end either when the question that set the inquiry in train is resolved or when it becomes plain that it cannot be resolved. Absent either outcome, further inquiry is indicated. Dogmatists take themselves to have brought many inquiries to a successful conclusion in the first way. Negative dogmatists, or dogmatic skeptics as we may also call them, have satisfied themselves that the questions are beyond resolution. By contrast, skeptics, properly so called, find that question after question remains open and hence calls for further inquiry. On their view, dogmatists of both the positive and negative variety were guilty of calling off their inquiries prematurely. And the fault about which ancient

skeptics complain most frequently is rashness or precipitate judgment.

The condition in which skeptics find themselves regarding the questions they investigate resembles that of negative dogmatists or dogmatic skeptics in being one of not knowing. But those who saw Academics and Pyrrhonists as skeptics in the modern sense were not simply confusing the condition of the inquirer with the dogmatic rejection of the possibility of knowledge. They were reacting to the fact that skeptics of both schools devoted far more time and energy to the case that nothing can be known than to arguments bearing on any other question.

The reason for this seems to be the following: It is possible to pursue unresolved inquiries into all sorts of questions without ever doubting that knowledge is, at least sometimes, achievable. But it is also possible to make the nature and possibility of knowledge an object of inquiry. If questions about knowledge remain stubbornly open, one of the things that one will not know, and that will require further study, is whether one can know at all; and from this central epistemological question the skeptical condition will spread to other inquiries, which can be brought to a conclusion only by justified claims to knowledge that the skeptic cannot make with confidence about anything. The inquiry into the possibility of knowledge remains open because of the persistent lack of satisfactory answers to the powerful arguments that knowledge is impossible. And ancient skeptics pursued the inquiry into the nature and possibility of knowledge chiefly by confronting the best theories of knowledge with these arguments.

Because the ancient skeptics consistently declined to make knowledge claims and constantly argued that nothing can be known, it is hardly surprising that outsiders took them to hold the position that nothing can be known and to hold it because they were convinced by the arguments they advanced in support of it. But if, in deference to tradition, we call this position the skeptical position and arguments supporting it skeptical arguments, for most ancient skeptics, being a skeptic was not a matter of holding the skeptical position in this traditional sense, and their reason for arguing skeptically was not to establish or defend the skeptical position. Rather, their skepticism was a matter of being unable to terminate the inquiries in which they were engaged—chiefly about the possibility of knowledge, but about the other matters as well.

PRECURSORS

The history of ancient Greek philosophy before the emergence of the main skeptical schools contains many figures who expressed doubts about the possibility of knowledge. Some of these were collected by skeptical Academics in order to provide themselves with a distinguished lineage.

Already in the sixth century BCE, the pre-Socratic philosopher Xenophanes composed some verses about the impossibility of human beings ever knowing for sure whether they had hit upon the truth or not. Perhaps the most important pre-Socratic precursor of skepticism was Democritus, who observed that his theory of atomism, which he took to be based ultimately on the evidence of the senses, had the consequence that the senses were unreliable, since the colors and flavors with which they appear to put us in contact would have no real existence if he were correct. It was characteristic of Academic argument especially, but also of many Pyrrhonian arguments, to proceed in the same way by deducing consequences imperiling the possibility of knowledge from dogmatic theories about knowledge. Though we are not well informed about the details, it is clear that a tradition calling the possibility of knowledge into question arose among philosophers influenced by Democritus. They include Metrodorus of Chios (fourth century BCE), whose work on nature begins, “None of us knows anything, not even this, whether we know or do not know,” and Pyrrho of Elis (circa 365–275 BCE), who is traditionally, though probably wrongly, viewed as the founder of the school which bears his name.

Unsurprisingly, skeptical Academics in search of illustrious antecedents appealed to the example of Socrates, who was the teacher of the Academy’s founder, Plato, and well known for claiming that he knew nothing except perhaps that he knew nothing. This was Socrates’ explanation for the pronouncement of the oracle in Delphi that he was the wisest man in Greece. The wisdom that set him apart from others, he conjectured, could only be his recognition that he lacked knowledge, whereas others, who were no more knowledgeable, deluded themselves and others into believing that they had knowledge.

Academic skeptics were inspired by at least two other characteristics of Socrates. First, though he set the highest possible value on knowledge and devoted his life to the pursuit of wisdom, Socrates lived an exemplary life without having attained it, thus providing the Academic skeptics with a model of the life they took themselves to be leading. Second, Socrates was a master of dialectic. A dialectical argument involves two parties: a questioner and an answerer. The answerer commits himself to a the-

sis, which it is his task to defend. The questioner aims to construct an argument to the contradictory of the answerer's thesis from grounds acceptable to the answerer, and he poses his questions with this end in view. When the questioner succeeds, it is through an argument all of whose premises have been conceded by the answerer. The answerer is thereby shown to lack the kind of understanding of the subject under discussion that Socrates' interlocutors typically claimed. The dialectical inquiry thus exposes problems inherent in the answerer's position or his defense of it or both. Since this kind of refutation can be accomplished by a questioner with no independent knowledge of the matters in contention, dialectical argument recommended itself to committed inquirers like Socrates, and it became the principal method of the Academic skeptics, who drew their inspiration from him.

Attempts were also made in the Academy to interpret Plato as a skeptic. The argument is based on his many expressions of caution and his manifest willingness in the dialogues to raise difficulties without resolving them. Whatever the merits of this claim, questions about the possibility of knowledge were not as prominent among Socrates' and Plato's concerns as they were among those of the Academics and Pyrrhonians.

Although book 9 of Diogenes Laertius's *Lives and Opinions of Eminent Philosophers* (the fullest treatment Pyrrhonism apart from Sextus Empiricus) includes a list of Pyrrhonists extending from Pyrrho to Sextus and beyond, the first part of it is almost certainly a construction. Pyrrho should probably not be viewed as the founder of a skeptical school. The modern scholarly consensus is that the Pyrrhonian school was founded in the first century BCE by Aenesidemus, who appears to have been an Academic dissatisfied with what he saw as the drift to dogmatism in the Academy of his time. He and his followers seem to have turned to Pyrrho in an effort to create an alternative history of skepticism that would make his school the legitimate heir of an older skeptical tradition.

Pyrrho wrote nothing but made a strong impression on his contemporaries, at least as much through his character as through his teachings. Figures with no sympathy for the positions he is thought to have held praised his imperturbability, lack of conceit, and tranquility. His views are elusive, however. Cicero seems to have known of him only as a moralist. He grouped Pyrrho together with figures like the heterodox Stoic Aristo of Chios (third century BCE). Such thinkers, he maintains, by making virtue the sole human good, fail to supply it with an object outside itself and so produce ethical theories incapable of furnishing practical guidance. The poet Timon of Phlius (c. 325–c. 238 BCE) became a follower of Pyrrho, whom he celebrated in a number of works that were probably the later Pyrrhonists' principal source of information about Pyrrho.

There was enough of an affinity between Pyrrho and Arcesilaus, the school leader of the Academy responsible for its skeptical turn, for their relationship to be the subject of a satirical verse by Aristo. Later on the characteristics of the Pyrrhonian school were imputed to Pyrrho. But whether and in what way Pyrrho was himself a skeptic remains subject to controversy. The most complete surviving account of his views is a late antique quotation of a first century CE citation of Timon. According to it, Pyrrho maintained that "things are equally indifferent, unmeasurable, and undecidable," and he went on to say that "neither our perceptions nor our opinions are true or false." According to one school of interpretation, the first claim is best viewed an epistemological thesis that Pyrrho deduced from the second, which, on this view, is an assertion about the apparent impossibility of distinguishing true from false beliefs. This interpretation would make him a skeptic, albeit probably a dogmatic one. But others have argued that the claim that "things are equally indifferent, unmeasurable, and undecidable" is a metaphysical thesis about the nature of reality from which Pyrrho inferred that perceptions and opinions cannot be true or false. In any case, he maintained that the proper response was to be without opinion, and he claimed that the result for those who attain this condition is tranquility.

ACADEMIC SKEPTICISM

Arcesilaus (316/15–241/40 BCE) became the fifth head of the Academy after Plato and was responsible for the school's turn to skepticism. To mark this change in outlook, later ancient writers speak of Arcesilaus as the founder of the New Academy as opposed to the Old Academy of Plato and his earliest successors; sometimes the Academy of Arcesilaus and his successors is called the Middle or Second Academy to distinguish it from the New or Third Academy of Carneades and his followers. (None of these distinctions corresponds to changes in the Academy as an institution.)

Like Socrates, Arcesilaus wrote nothing but was distinguished by his mastery of dialectic in face-to-face conversation. Rather than expound or defend views of his own, he would let his interlocutors put forward a view that he would then subject to dialectical examination. His decision to make Stoic epistemology the principal object

of his inquiries exerted a decisive influence on the subsequent history of ancient skepticism.

The Stoics took wisdom to mean a firm grasp of the truth, entirely free from error. They maintained that, though exceedingly rare and difficult of attainment, wisdom was nevertheless within the power of human beings. The key concept in the Stoics' account of wisdom, what they called cognitive impressions—their criterion of truth—which they define as impressions "from what is, stamped and impressed in exact accordance with what is, and such as could not be from what is not." In the paradigm case of perceptual impressions, this means that cognitive impressions arise in a way that ensures that they capture their objects with perfect accuracy, thus guaranteeing their truth, and at the same time impart to them a character that human beings can discern.

Assent to a cognitive impression is a cognition or apprehension, and, if further conditions are satisfied, it will qualify as knowledge. Assent to anything but a cognitive impression is opinion, and, according to the Stoics, the wise avoid error by remaining entirely free of opinion. Arcesilaus began the long Academic tradition of arguing that there are no cognitive impressions, which in the context of Stoic epistemology amounts to arguing that knowledge is impossible. He did this by arguing for indiscernibility— that is, he held that the character purportedly peculiar to cognitive impressions could also belong to impressions that did not arise in the required truth-guaranteeing way and were in fact false. His arguments were based as much as possible on considerations that the Stoics would have to acknowledge, either because they were drawn from Stoic theory or could be rejected only at a high cost in plausibility.

The idea that there are no cognitive impressions ("inapprehensibility" for short) was the first skeptical proposition with which the Academy came to be associated. The second, that it is incumbent on the wise to suspend judgment on all matters, Arcesilaus deduced from the first, along with the Stoic doctrine that wisdom is incompatible with opinion. Together they make up what we might call a skeptical position.

On a strictly dialectical interpretation of Arcesilaus's arguments, the conclusions he drew need tell us nothing about what views, if any, he held. The propositions that make up the skeptical position follow in the context of arguments dominated by Stoic assumptions about what is to count as knowledge and about the incompatibility of wisdom with opinion; these issues raise problems for the Stoics to solve. To be sure, Arcesilaus responded to Stoic arguments that action was impossible without assent, and assent senseless in the absence of cognitive impressions, by defending the possibility of a life in which all judgment is suspended. But this argument may only have shown that the Stoics were not in a position to easily escape the difficulties raised by his first set of arguments. And the fact that his response to the Stoics was based so closely on their theory of action as to have no force outside this debate lends support to this suggestion.

It is clear, however, that Academics after Arcesilaus interpreted him as endorsing the skeptical propositions in a certain way. This was their own view, and they may have been right about Arcesilaus. Thus a skeptical stance or outlook arose in the Academy as a result of a dialectical dispute with the Stoa that was expressed by means of the skeptical propositions. But the Academic followers of Arcesilaus seem not to have subscribed to the skeptical propositions in the ordinary way. Instead, their situation is akin to that of the skeptics described by Sextus: They were not in a position to conclude the inquiry into the nature and possibility of knowledge or other inquiries dependent on its resolution. And it is this condition that they described in terms borrowed from their debate with the Stoa—inapprehensibility and suspension of judgment—not the condition of being convinced by the arguments on their side of the debate.

We know little about Arcesilaus's successors before Carneades. Carneades was another exceptionally gifted dialectician and nonwriter. It is likely that he supplemented and refined the arguments against cognitive impressions that he inherited from his predecessors, but his most distinctive contribution was his response to the Stoics' argument that without cognitive impressions and assent, action and life are impossible. Whereas Arcesilaus's response stayed very close to Stoic theory, Carneades's did not. Instead he seems to have worked out a full-blown theory of so-called probable impressions (*probabilis* was Cicero's Latin translation of the Greek *pithanos*, meaning *persuasive*). And he appealed to them to explain how life, even a life of wisdom, was possible without the perfectly secure foundation provided by cognitive impressions.

As Arcesilaus had done before him, Carneades defended the possibility of acting without assent. There is, he argued, a way of using or following probable impressions that does not amount to assent but is adequate for action and inquiry. But he also sometimes conceded that assent was essential in order to argue that even this consession did not vindicate Stoic claims about the cognitive impression. For, he suggested, it was permissible for the wise to form opinions by assenting to noncogni-

tive impressions, opinions held in the full consciousness that they were only opinions and might be wrong.

This line of argument is behind the view that Carneades relaxed or weakened the more militantly skeptical stance of Arcesilaus. But perhaps the new features of Carneades's arguments are part of a broadly dialectical form of argument. The Stoics believe their views should win acceptance not because they are theirs but because they do justice to common assumptions about human nature—its needs and the resources available to it—as no others can. The challenge that Carneades accepted, then, was to show that the ready availability of equally sound or even better alternatives ought to discourage a premature embrace of the Stoic position.

This posture makes it hard to know whether Carneades actually subscribed to any of the views he defended. And his students and successors interpreted him in different ways. Clitomachus, his student and eventual successor, held that one should suspend judgment and that this had been Carneades's view. Philo of Larissa, who succeeded Clitomachus, contended that Carneades believed that the wise were permitted to form opinions in the absence of cognitive impressions and that one of the probable views deserving assent was inapprehensibility. Philo was, then, a dogmatic skeptic, who championed one of the skeptical propositions simply because he was convinced by the arguments for it. There is an air of paradox about this position, but it must be remembered that he did not claim to know for certain that nothing can be known for certain, but rather that it was highly probable, which, if nothing can be known for certain, is the most that can be said for it.

PYRRHONISM

It seems to have been Philo of Larissa's dogmatic skepticism that moved Aenesidemus to found or revive a competing school of Pyrrhonian skepticism in the first century BCE. The Pyrrhonian school he founded existed past the time of Sextus Empiricus, who is usually thought to have been active in the latter part of the second century CE. Although none of Aenesidemus's works have survived, a summary of eight books of his Pyrhhonian Arguments by Photius (ninth century CE) has. From it we learn that Aenesidemus, who had been an Academcie himself, charged the Academics of his time with being little more than Stoics fighting Stoics, disagreeing only about cognitive impressions while agreeing about many other issues. Though the decision by Aenesidemus and his followers to call themselves "Pyrrhonists" does not imply a direct line of descent from Pyrrho, it is probable that they were influenced by traditions about Pyrrho. Another important influence came from the Empirical school of medicine, with which Pyrrhonism maintained close ties and shared many members including Sextus Empiricus (whose name means "the Empiricist").

In view of the school's origins, it is not surprising to find many points of contact between it and Academic skepticism. The Pyrrhonists describe the skeptical condition with the aid of terms like "inapprehensibility" and "suspension of judgment," which have their origins in the epistemological debate between the Academy and the Stoa. They view this condition as the result of a standoff or impasse between their arguments and those of their dogmatic opponents, not as the result of being convinced by their own skeptical arguments. And they explain that they are able to act and to live despite suspending judgment on all questions. This argument hinges on a distinction between two senses of "belief" (Greek: dogma) that is indebted to Carneades's and Clitomachus's contrast between assenting to an impression and using or following it. In the former sense, the Pyrrhonists had no beliefs, but in the latter sense they did have beliefs, which were able to serve as a basis for action. The two works of Sextus that have come down to us, the *Outlines of Pyrrhonism* in three books and *Against the Mathematicians* in nine, are packed with arguments against dogmatic positions, many of which are of Academic origin.

There are, however, equally notable differences between the two schools, some of which may reflect other influences on Aenesidemus and his followers. The most striking and important of these is the positive value the Pyrrhonists seem to attach to the skeptical suspension of judgment about all matters. According to Sextus, Pyrrhonism has a telos, a supreme aim or goal in life: tranquility (and, where that is unattainable, moderation in one's emotions). Suspension of judgment is recommended because it gives rise to tranquility. This recommendation is not based on a theory of human nature that would explain why it finds fulfillment in tranquility. Rather the argument seems to presupposed that tranquility is humans' goal. This assumption commands greater credibility if viewed not as a claim about the essential nature of the best life for human beings, which would elicit vehement disagreement from some ancient philosophical schools, but as a weaker claim that such a life will somehow involve tranquility. And the Pyrrhonists do not pretend to be able to explain why suspension of judgment should give rise to tranquility; they claim to have made this discovery only by accident. Tranquility is supposed to arise in a manner exemplified by the famous story of

Apelles the painter, who, despairing of being able paint the foam on the neck of a racing horse, gave up and threw his sponge at the painting, thereby producing by chance what he had been unable to achieve deliberately.

The idea of a correlation between freedom from opinion and tranquility may have been the Pyrrhonian school's truest debt to Pyrrho. This idea sets it clearly apart from the Academy. The Academy attached the highest value to knowledge and regarded the skeptical condition as a stop-gap, albeit a surprisingly congenial one. Asa we have seen, the Pyrrhonists were officially committed to the quest for knowledge. But the accounts of Pyrrhonism in Sextus and Diogenes Laertius give evidence of a positive attachment to suspending judgment as a means to tranquility. Arguments and argumentative strategies are recommended for their efficacy in bringing about equipollence, the condition in which arguments on either side of a question are of apparently equal force; and equipollence is cultivated not as a means to cognitive certainty but to the suspension of judgment that leads to tranquility. Thus there is a sense in which Academics like Arcesilaus and Carneades exemplified true "skepticism," in the sense of open-minded inquiry, more than the Pyrrhonists did.

There is also a difference in the kinds of arguments the two schools used. Sextus and our other sources give pride of place to the so-called modes or tropes of argument that bring about suspension of judgment. There is a set of ten such tropes, which seem to go back to Aenesidemus, and a later set of five ascribed to Agrippa, who may, however, be a fictional character in a Pyrrhonian work. (There is also a set of two tropes, and a further set of eight tropes concerning causal explanation, which is likewise credited to Aenesidemus). The ten tropes appear to be the oldest, and they draw on arguments and examples with a long history. Book Gamma of Aristotle's *Metaphysics* is already familiar, with arguments resembling those in the ten tropes. Most of the ten aim to demonstrate that there are undecidable conflicts between the appearances perceived by different species or different human beings or the different senses or by the same human being in different conditions or between the appearances presented by objects in different circumstances. The existence of such conflicts is illustrated by a wealth of examples, some of them fanciful. Left unclear are the exact arguments envisaged and how they relate to the official program, which calls for the production of equipollence by the balancing of arguments. The tropes seemingly aim to elicit from these conflicting appearances a thesis of undecidability thatrequires suspension of judgment. That is, it appears as though undecidability arises from an argument whose premises would command the assent of the skeptic. But perhaps the arguments for the undecidability of conflicts are meant to oppose arguments that they are decidable, and it is the equipollence between these arguments that is supposed to lead to suspension of judgment.

Even so, by comparison with Academic arguments, and with the arguments found elsewhere in Sextus, the trope-based arguments appear somewhat naive. Substantial assumptions about species, perceptual faculties, and the requisite conditions for the acceptance of an impression as true enter the argument without being marked as dialectical concessions or without comment of any kind about their status. Perhaps the material collected in the ten tropes arose from traditions of dogmatic skeptical thinking outside the Academy and maybe even from Pyrrho himself. There is a problem with the trope of relativity, which may suggest a similar conclusion about origins. According to this trope, since all things are relative, we must suspend judgment about their real natures. Though Sextus makes an attempt to correct for this, the conclusion of this argument is not, properly speaking, skeptical.

The five Agrippan modes are (i) disagreement, (ii) regress to infinity, (iii) relativity, (iv) hypothesis, and (v) circularity. Except for relativity, they form a system by means of which dogmatic attempts to justify a disputed claim can be counteracted. Any claim put forward invites disagreement. Further claims enlisted in support of it will lead to an infinite regress, by requiring justification themselves, unless the process is brought to an arbitrary halt with a hypothesis or the justification depends on the originally disputed claim. To judge from the enormous mass of arguments preserved by Sextus, neither set of tropes consistently guided the Pyrrhonists as they collected and composed arguments to further their skeptical purposes.

See also Aenesidemus; Agrippa; Aristo of Chios; Aristotle; Carneades; Cicero, Marcus Tullius; Diogenes Laertius; Dogma; Greek Academy; Leucippus and Democritus; Philo Judaeus; Philo of Larissa; Plato; Pyrrho; Sextus Empiricus; Skepticism, History of; Socrates; Stoicism; Timon of Phlius; Xenophanes of Colophon.

Bibliography

Algra, Keimpe, Jonathan Barnes, Jaap Mansfeld, and Malcolm Schofield, eds. *The Cambridge History of Hellenistic Philosophy*. Cambridge, U.K.: Cambridge University Press, 1999.

Allen, James. "Academic Probabilism and Stoic Epistemology." *Classical Quarterly* 44 (1994): 85–113.

Allen, James. "Pyrrhonism and Medical Empiricism: Sextus Empiricus on Evidence and Inference." *Aufstieg und Niedergang der Römischen* Welt II. 37.1, edited by Wolfgang Haase. Berlin: Walter de Gruyter. 646–90.

Annas, Julia. "Plato the Sceptic." *Oxford Studies in Ancient Philosophy*, suppl. vol. (1992): 43–72.

Bailey, Alan. *Sextus Empiricus and Pyrrhonean Skepticism*. Oxford: Oxford University Press, 2002.

Bett, Richard. *Pyrrho, His Antecedents and his Legacy*. Oxford: Oxford University Press, 2000.

Brittain, Charles. *Philo of Larissa: The Last of the Academic Sceptics*. Oxford: Oxford University Press, 2001.

Brochard, Victor. *Les sceptiques grecs.* 2nd ed. Paris: Vrin, 1887.

Brunschwig, Jacques. *Papers in Hellenistic Philosophy*. Cambridge, U.K.: Cambridge University Press, 1994.

Burnyeat, Myles, ed. *The Skeptical Tradition*. Berkeley: University of California Press, 1983.

Burnyeat, Myles, and Michael Frede, eds. *The Original Sceptics: A Controversy*. Indianapolis: Hackett, 1997.

Cicero. *De natura deorum, Academica*. Translated by H. Rackham. Cambridge, MA: Harvard University Press, 1933.

Diogenes Laertius. *Diogenes Laertius: Lives of Eminent Philosophers.* 2 vols. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1929.

Frede, Michael. *Essays in Ancient Philosophy*. Oxford: Oxford University Press, 1987.

Glucker, John. *Antiochus and the Late Academy Hypomnemata 56*. Göttingen: Vandenhoeck und Ruprecht, 1978.

Groarke, Leo. *Greek Scepticism: Anti-realist Trends in Ancient Thought*. Montreal: McGill–Queen's University Press, 1990.

Hankinson, R. James. *The Sceptics*. London: Routledge, 1995.

Inwood, Brad, and Jaap Mansfeld, eds. *Assent and Argument: Studies in Cicero's Academic Books*. Leiden: Brill, 1997.

Robin, Léon. *Pyrrhon et le scepticisme grec*. Paris: Presses Universitaires de France, 1944.

Schofield, Malcolm, Burnyeat, Myles, and Barnes, Jonathan,eds. *Doubt and Dogmatism: Studies in Hellenistic Epistemology*. Oxford: Oxford University Press, 1980.

Sextus Empiricus. *Outlines of Pyrrhonism, Against the Professors.* 4 vols. Translated by R. G. Bury. Cambridge, MA: Harvard University Press, 1955.

Sextus Empiricus. *Outlines of Scepticism*. Translated by Julia Annas and Jonathan Barnes. Cambridge, U.K.: Cambridge University Press, 1994.

Sextus Empiricus. *The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism*. Translated by Benson Mates. New York: Oxford University Press, 1996.

Stough, Charlotte. *Greek Skepticism*. Berkeley: University of California Press, 1969.

Striker, Gisela. *Essays on Hellenistic Epistemology and Ethics*. Cambridge, U.K.: Cambridge University Press, 1996.

Tarrant, Harold. *Scepticism or Platonism? The Philosophy of the Fourth Academy*. Cambridge, U.K.: Cambridge University Press, 1985.

James Allen (2005)

ANDERSON, JOHN

(1893–1962)

John Anderson, the Scottish-born Australian philosopher, was the son of a politically radical headmaster. Born at Stonehouse, Lanarkshire, and educated at Hamilton Academy and at the University of Glasgow, which he entered in 1911, he was at first principally interested in mathematics and physics; he turned to philosophy partly under the influence of his brother William, then a lecturer at Glasgow and later professor of philosophy at Auckland University College, New Zealand. Anderson graduated with an M.A. in 1917, with first-class honors both in the school of philosophy and in the school of mathematics and natural philosophy (physics). He lectured at Cardiff (1918–1919), Glasgow (1919–1920), and Edinburgh (1920–1927) before accepting an appointment in 1927 as professor of philosophy at the University of Sydney, Australia. He remained there, except for a visit to Scotland and the United States in 1938, until his retirement in 1958. He had almost no personal contact with philosophers in England, a country he regarded with the suspicion characteristic of a Scottish radical.

Anderson's career as a professor was an unusually stormy one. He attacked whatever he took to encourage an attitude of servility—and this included such diversified enemies as Christianity, social welfare work, professional patriotism, censorship, educational reform of a utilitarian sort, and communism. For a time he was closely associated with the Communist Party, seeing in it the party of independence and enterprise, but he broke with it in the early 1930s. His passionate concern for independence and his rejection of any theory of "natural subordination" were characteristic of his whole outlook—political, logical, metaphysical, ethical, and scientific. Attempts were made to silence him and even to remove him from his professorship; he was subjected to legislative censure and clerical condemnation. In the debates that these attacks provoked, he spoke out forcibly and fearlessly in defense of free speech and university autonomy.

METAPHYSICS AND EPISTEMOLOGY

Anderson was trained at Glasgow as an Absolute Idealist. However, he soon abandoned Idealism, influenced by William James, whom he studied very closely, G. E. Moore, Bertrand Russell, the American "new realists," and, most significantly, Samuel Alexander, whose Gifford Lectures on *Space, Time and Deity* he attended in Glasgow in 1916–1918. James and Alexander taught him that

it was possible to reject absolute idealism without, like Russell, reverting to a modified version of traditional British empiricism. Anderson set out to show that continuity, stressed by absolute idealists, and distinction, stressed by empiricists, are equally real and equally involved in every experience. In experience, he argued, we encounter neither an undifferentiated continuum nor isolated sense data; our experience is of complex states of affairs, or "propositions," understood not as sentences, but as what true utterances assert to be the case. These propositions do not mediate between ourselves and reality; to take that view, Anderson argued, is to leave us in a state of invincible ignorance about this supposed "reality." To be real simply is to be "propositional," that is, to be a thing of a certain description, or, in Anderson's view, a complex of activities in a spatiotemporal region.

Unlike many of his British contemporaries, Anderson was by no means opposed to the use of philosophical labels; he was prepared to describe himself as an empiricist, a realist, a pluralist, a determinist, a materialist, or a positivist—but always in a somewhat individual sense. For example, although he insisted that he was an empiricist, he rejected what is usually taken to be the most characteristic doctrine of empiricism—that our experience is of "impressions" or "sense data." For Anderson, empiricism consisted in the rejection of the view that there is anything "higher" or "lower" than complex states of affairs as we encounter them in everyday experience; he rejected ultimates of every sort, whether in the form of ultimate wholes, like Francis Herbert Bradley's Absolute, or ultimate units, such as "sense data" or "atomic propositions."

Similarly, although he agreed with positivists in their opposition to metaphysics, when it is understood as the revelation of realities "beyond facts," he shared neither the positivist hostility to traditional philosophy as such, nor its conception of experience as consisting in "having sensations," nor its interpretation of logic and mathematics as calculi. He was a realist, insofar as he argued that what we perceive exists independently of our perceiving it; but he forcibly criticized the phenomenalism characteristic of so many twentieth-century realists. He described himself as a pluralist, but whereas classical pluralism had defended the thesis that there is a plurality of ultimate simples, everything, for Anderson, is complex. No state of affairs is analyzable into just so many ingredients—whether in the form of sense data or of abstract qualities. Pluralities, in his view, consist of pluralities, not of simples. For the same reason he was not a determinist in the classical sense, because for him no description of a situation was ever complete; his determinism consisted only in his holding that there are sufficient and necessary conditions for the occurrence of any state of affairs. Finally, his materialism did not incorporate the classical conception of matter; what is essential to his view is the idea that every state of affairs is describable in terms of physical laws—which does not exclude its also being describable in terms of biological, psychological, or sociological laws.

The arguments by which Anderson attempted to establish his philosophical conclusions were manifold and diverse. What was perhaps his fundamental argument can be put thus: As soon as we try to describe "ultimate" entities or offer any account of their relation to those "contingent" entities whose existence and behavior they are supposed to explain, we find ourselves obliged, by the very nature of the case, to treat the alleged "ultimates" as possessing such-and-such properties as a "mere matter-of-fact." The metaphysician either sees his ultimate entities vanish into emptiness—like John Locke's "substance"—or else he is forced to admit that they exhibit precisely the logical characteristics which were supposed to indicate that a thing is not ultimate.

The emptiness of ultimates, Anderson thought, is often disguised by the fact that they are defined in wholly relational terms—as when, for example, substance is defined as "that which underlies qualities," or a sense datum as "that which is an object of immediate perception." Anderson attacked this procedure as "relativism," that is, as the attempt to think of an entity or a quality as being wholly constituted by its relation to something else. To be related, Anderson argued, an entity must be qualitatively describable; relational definitions, it follows, cannot be used to avoid the conclusion that the "ultimate," if it exists at all, must itself be a thing of a certain description. According to Anderson, every state of affairs is "ultimate," in the sense that it is something we have to take account of; but it is contingent, too, in the sense that there are circumstances in which it might not have come about. There is nothing whose nature is such that it must exist, but there is nothing, either, whose nature is exhausted by its relation to other states of affairs.

Particularly in Anderson's lectures, through which his influence has been mainly exerted, such general considerations were supported by detailed analyses of specific philosophical theories. Although he was not, in a professional sense, a scholar, it was his habit both to develop his own views by way of a criticism of his predecessors and also to ascribe to those predecessors—espe-

cially perhaps to Heraclitus and to the Plato of the later dialogues—the views that he took to be correct.

LOGIC AND MATHEMATICS

Anderson's approach to philosophy was in some respects formal. He agreed with the Russell of *Our Knowledge of the External World* that logic is the essence of philosophy—if by this is meant that philosophical problems are to be settled by an analysis of propositions. But despite strong mathematical interests, he was only to a very limited degree influenced by Russell's mathematical logic. He worked out, and defended against Russell's criticisms, a reformulated version of the traditional formal logic, which he tried to show had a much greater range and power than its critics would allow to it. He related logic very closely to discussion: the conception of an "issue," of what is before a group for consideration, bulks very large in his logic. The issue, he thought, is always whether some kind of thing is of a certain description, and discussion consists in drawing attention to connections between such descriptions. Unless these connections actually hold, discussion falsifies unless it is actually the case, for example, that what one person brings forward as an objection is logically inconsistent with what another person has said. To point to logical relations, Anderson concluded, is to assert that something is the case, just as much as to draw attention to any other sort of relation.

He took a similar stand concerning mathematics, which, he argued, can be applied to the world only in virtue of the fact that it describes that world. "Application," in Anderson's view, consists in drawing conclusions from what is being applied. If mathematics offered no description of the world, no application of it could describe the world.

He did not, however, agree with John Stuart Mill that mathematical propositions are "inductions from experience." He was a vigorous critic of induction. If, as traditional empiricists had assumed, all our experience is of "pure particulars," then, according to Anderson we would not have the slightest ground for believing in—we could not even conceive the possibility of—general connections. But, in fact, the least we can be acquainted with is not a bare particular but a particular state of affairs; from the very beginning, generality is an ingredient of our experience. We can recognize directly that, say, fire burns, although we can be mistaken in this as in any other of our beliefs; for to "recognize" is nothing more or less than to hold a belief.

AESTHETICS, ETHICS, AND POLITICAL PHILOSOPHY

Although even in his aesthetic, ethical, and political writings, Anderson was constantly concerned to make formal points—as, for example, that the definition of good as "that whose nature it is to be an end" exhibits the vice of "relativism"—yet he was also a good deal influenced by, and deeply concerned with, the issues raised by economists like Alfred Marshall, social theorists like Karl Marx and Georges Sorel, critics like Matthew Arnold, psychologists like Sigmund Freud, and novelists like James Joyce and Fëdor Mikhailovich Dostoevsky. His aesthetic, ethical, and political writings conjoin the logical and the concrete; in virtue of this fact he has influenced many Australian intellectuals who would not accept his formal analyses.

In his aesthetics, Anderson argued that the beauty of works of art is independent of the observer; and similarly in ethics, that acts are good or bad in themselves. He was influenced by Moore's *Principia Ethica* but critical of Moore's attempt to treat "good" as being a simple and indefinable quality and at the same time to define it as "that which ought to be," and thus a quality. Anderson took "good" to be a predicate of certain forms of mental activity—the spirit of inquiry, love, courage, artistic creation, and appreciation—and tried to work out a theory of the connection and distinction between these different forms of activity.

In his political theory, Anderson attacked, on the one hand, the view that human society has a single "good" to which all activity ought to be subordinated, and, on the other hand, the doctrine that it is a set of contractual relations between individuals. Society, as he saw it, is a complex of complex institutions, of which the state is only one. A community flourishes when this fact is fully realized, when no attempt is made to enforce uniformity upon these diverse competing and cooperating types of institutions. The attempt to achieve absolute security by social planning, Anderson held, is doomed to failure and is stultifying in its effects in society.

INFLUENCE. Anderson's ideas were presented in a series of articles, mainly in the *Australasian Journal of Philosophy,* and in his influential lectures at the University of Sydney, where he founded what has been described as "the only indigenous school of philosophy in Australia." Among those philosophers who have, in varying degrees, felt his influence, the best known are D. M. Armstrong, A. J. Baker, Eugene Kamenka, J. L. Mackie, P. H. Partridge, and J. A. Passmore.

See also Aesthetics, History of; Alexander, Samuel; Armstrong, David M.; Bradley, Francis Herbert; Dostoevsky, Fyodor Mikhailovich; Freud, Sigmund; Heraclitus of Ephesus; Idealism; James, William; Locke, John; Mackie, John Leslie; Marx, Karl; Mill, John Stuart; Moore, George Edward; Plato; Positivism; Russell, Bertrand Arthur William; Sorel, Georges.

Bibliography

Anderson's principal contributions to periodicals, together with two previously unpublished papers, have been brought together as *Studies in Empirical Philosophy* (Sydney: Angus and Robertson, 1962), with an introduction by J. A. Passmore, which contains a full bibliography. For further information, see Gilbert Ryle, "Logic and Professor Anderson," in *Australian Journal of Philosophy* 28 (3) (1950): 137–153; J. L. Mackie, "The Philosophy of John Anderson," in *Australian Journal of Philosophy* 40 (3) (1962): 264–281; J. A. Passmore, "Philosophy in Australia," in *Australian Culture* (Ithaca, NY: Cornell University Press, 1963).

John Passmore (1967)

ANDŌ SHŌEKI

Andō Shōeki was a critical thinker in the Tokugawa period of Japan. All that is known of his life is that he was born in Akita toward the end of the seventeenth century and died in the second half of the eighteenth century, that his profession was medicine, and that he went to Nagasaki, the first Japanese port to receive Western trade, where he learned about conditions in foreign countries. He is described as a man of stern character who in his teaching never quoted, except to criticize, the Chinese classical books, meaning that he followed only his own ideas, a very unorthodox way of teaching for Tokugawa Confucianists. Very fond of the peasant class, he insisted that his pupils, and he had very few, should do manual work to be in contact with nature, the greatest master of all. Until recently he was virtually unknown, because of his nonconformist ideas, although nowadays he is overpraised. His manuscripts were found only in 1889, and only in part. They were published with difficulty. The better-known are *Shizen shin-eidō* (The way and activity of nature, written in 1755) and *Tōdō shinden* (A true account of the ruling of the way). They are the most devastating critique ever made of Tokugawa society and of every kind of Japanese ideology.

Andō's iconoclasm was directed first of all against Shintoism and Buddhism. He sharply attacked Shinto mythology and Prince Shōtoku (574–622) for his role in spreading Buddhism. Other rulers, too, and priests of all sects came under his critical scrutiny, which is too negative. Nor had he a better appreciation of the different schools of Confucianism, for he accused them of perverting the teaching of the old sages in their interpretation of nature.

Nature for Andō is an eternal *ki,* or material energy, in perpetual motion. Nature is not to be conquered but to be known; and in following nature man attains the ideal. More positive were his ideas about society; he was the only genuine equalitarian of Tokugawa Japan, arguing against the evils of a system which oppressed the peasant. He cannot be considered completely iconoclastic, since he was not against authority as such, nor was he an atheist, and even his alleged materialism has to be qualified.

See also Japanese Philosophy.

Bibliography

For a guide to primary sources, see bibliography in the Japanese Philosophy entry. See also E. H. Norman, "Andō Shōeki and the Anatomy of Japanese Feudalism," *Transactions of the Asiatic Society of Japan,* 3rd series, 2 (1949): 1–340; and Y. B. Radul-Zatulovskij, *Andō Shōeki, Filosof Materialist XVIII Veka* (Moscow, 1961).

Gino K. Piovesana, S.J. (1967)

ANIMAL MIND

Mind is considered in terms of contents or processes or both. The term usually includes both conscious and unconscious events. In the case of the term *animal mind*, there is intense scientific and philosophical disagreement as to whether animal minds are unconscious or can include conscious events as well. In particular, even among scientists who may accede to the possibility of animal consciousness, there is great reluctance to consider the issue as amenable to scientific study. Donald R. Griffin is a particularly notable exception, who has made the issue a focus of his scientific attention.

OVERVIEW OF PHILOSOPHICAL AND SCIENTIFIC HISTORY

Concerns that still strongly engage philosophers and psychologists to this day were raised by the opposing ideas of John Locke (1632–1704) and Rene Descartes (1596–1650). In Locke's accounting, the elements of mind are ideas. Ideas are written by experience onto the blank slate

of the mind, the *tabula rasa* first proposed by Aristotle. Descartes claimed that ideas are innate.

Locke considered that human's ideas are created through sensations; furthermore human minds can reflect upon their ideas. According to Locke, an automatic process of association is an essential mechanism in the linkage between ideas. Descartes, too, proposed automatic, mechanistic connections to explain the mind and behavior of animals and much of humans'. Descartes emphasized automatic reflexes, which are connections between stimulating sensations of the external world and the organism's response to those sensations. For humans, Descartes proposed a mediating influence that could be exerted on reflexes by the soul operating through the brain. These views of Locke and Descartes strongly determined the field of experimental psychology; the reflex and the process of association formed the basis of the phenomenon of classical conditioning.

Both Locke's and Descartes's ideas impacted directly on the study of animal mind. Descartes had claimed that man has a soul, while animals do not; they are mere automata. Humans too have automatic processes, but humans are conscious, feel pain, and experience pleasure, while animals do not. Locke considered animals to have memory and to be capable of simple cognitive processes, including simple reasoning. They lack, however, the capacity to manipulate their ideas, to reflect upon them, as humans can. Essentially, "Brutes abstract not" (Locke 1690, website, p. 31).

With the advent of Charles Darwin's theory of evolution, the proposed continuity between humans and animals promoted a search for animal abilities that were precursors of human abilities. Darwin's *The Expression of Emotions in Man and Animals* (1872) not only proposed that animals experience emotions, but that, indeed, the expression of human emotion is in many ways similar to and derivative from that of certain other species, particularly nonhuman primates. At about that time, George Romanes (1882) compiled numerous examples of animal intelligence; the range of presumed capacities startled the public and scientists were criticized. The criticism, especially in later decades, decried the anecdotal nature of many observations and stressed the need for experimental verification. These issues continue to trouble the adequate documentation of observed instances of intelligent behavior, which would most plausibly be revealed in single, unique instances as an organism attempts to deal with a novel situation.

In the 1920s, Ivan Pavlov's study of digestion in dogs led him to discover that the dogs learned to anticipate the arrival of food via signals in the environment, such as his entry into the room. Evidence was the dogs salivating well before food was in their mouths. Pavlov's many subsequent detailed studies revealed underlying laws of classical conditioning.

The behavioristic approach was further espoused by James Watson and then by B. F. Skinner. They argued that private mental states cannot be the subject matter of science, only public events can be. Concentration was on learned behavior, reducible to stimulus-response units, which were subject to psychological laws. The laws of behavior were derivative of Locke's postulated process of association and, with the Pavlovian laws of classical conditioning, dominate experimental psychology even into the twenty-first century.

In a more cognitive approach, Edward Tolman's studies (1948) of rats learning their way through complicated mazes led him to propose that rats create a tentative, cognitive map indicating routes and environmental relationships, which determine the rats' responses. He struggled with the issue of behavioral indices of mental states. Of particular interest to him was specifying descriptive properties of a behavior to indicate that it is purposive. Tolman's views met with skepticism and interest in them faded until the concept of animals' cognitive maps was revived in an important book by John O'Keefe and Lynn Nadel (1978).

Griffin's influential book, *The Question of Animal Awareness* (1976), and his several subsequent books, reawakened both interest and controversy about animal awareness and thinking. His emphasis that animal awareness is an issue amenable to scientific study spurred investigations into animal cognitive capacities, both in the lab and the field.

Yet just what cognitive processes animal minds possess is controversial. Most contemporary experimental psychologists prefer to examine such processes without relevance to issues of consciousness. In an effort to create highly replicable experimental paradigms in controlled laboratory settings, the scientists can justifiably be strongly criticized for setting for their subjects very simplistic tasks, many of little or no relevance to the organism in its natural life, situating them in impoverished environments for rearing and testing (e.g. T- mazes or Skinner boxes/operant chambers) and for ignoring the contextual effects that are always part of the experimental conditions (e.g., as Pavlov had noted, the dogs in his study began salivating before his original digestion experiment had officially begun). Furthermore the subject of choice is most often the white rat, a genetically inbred

docile animal, which may well have lost some of the cognitive and other traits essential to survival in the complex, treacherous world of the wild rat.

In addition to psychologists, ethologists, in particular cognitive ethologists, also study animal minds. Cognitive ethology is a field established by Griffin, being the study of animals' mental experiences, particularly in the course of their daily lives in their natural environment. Data are gathered either from observations or experiments in the field, initial observations often forming the basis for creating the experimental investigations. Likewise some laboratory work has become more naturalistic, employing larger spaces and other means to simulate the organism's niche. Philosophers of science and of mind have shown interest in the field of cognitive ethology, and some, such as Daniel Dennett and Colin Allen, have collaborated in varying degrees with ethologists. Other very relevant contributions have been made, such as Ruth Millikan's (1984, 1995) analysis of natural functions and both Jonathan Bennett's (1989)and John Searle's (2000) considerations of intentionality, belief-desire systems, and consciousness.

CAPACITIES OF ANIMAL MIND

The aspects of animal mind include cognitive, emotional, moral, and communicative capacities and consciousness.

COGNITIVE CAPACITIES. By defining cognition very broadly, one can include the simplest processes, for example, habituation, found in fairly simple creatures such as the sea slug, *Aplysia*, to processes of abstraction, inference, and deception, credited to some primates and selected other species. (Habituation is a process whereby an organism decreases responding to a repeated stimulus.) In most psychological analyses, investigators assume that processes found at the lowest evolutionary levels are similarly to be found in any and all higher organisms (insofar as a hierarchical notion of evolution is appropriate). This is the model of experimental psychologists who conduct laboratory studies of white rats and pigeons.

However neurophysiological studies of simple organisms such as *Aplysia* and the mollusk *Hermissenda* do note different biochemical and neural mechanisms that may underlie similar psychological processes (e.g., classical conditioning at the cellular level). Ethologists, too, are quick to note species specific and niche specific behavioral traits, which often depend upon specialized sensory receptors. Without the capacity to detect certain information, there is no opportunity to develop advanced cognitive capacities dependent upon such information. Thus bats can echo-locate and, thereby, in the dark navigate through obstacles and catch minute insects; dogs can follow faint odor trails of individuals; humans can do neither.

Psychologists would argue that the same basic psychological laws can be applied to different sensory systems, but there is mounting evidence against this interpretation. Rather than the laws of classical conditioning applying equally to all stimuli, evidence shows, for example, a bias for associating stimuli that are involved with ingestive systems. Thus, in laboratory experiments, rats tend to associate taste with apparent nausea (induced by X-rays) while visual and auditory stimuli are readily associated with painful exteroceptive stimuli. The latter biases are usually interpreted as associations most relevant in predatory-prey interactions and in other dangerous environmental events, as seen in work by John Garcia and R. A. Koelling (1966). Pigeons are biased to associate visual cues with X-ray induced illness; for the pigeon, vision is most essential in detecting their appropriate foods, such as grains. Simply put, organisms have evolved to readily learn which food associated stimuli make them ill, and thus are better able to avoid such. And further they can associate the stimuli with an illness occurring several hours later, contrary to assumed need for temporal contiguity.

In brief, all animal species, including some insects that have been studied, and probably even some single-celled animals, have been shown to be capable of at least the following: habituation, classical conditioning, and operant conditioning. But since the 1960s, important constraints on those basic processes have been recognized. Classical conditioning most simply refers to the learning process whereby a previously neutral stimulus (the conditioned stimulus or CS), when paired with a noxious or positive stimulus (the unconditioned stimulus or US), comes to elicit a response preparatory for or similar to that elicited by the US.

Since the 1960s important constraints on the basic learning processes have been recognized. Lab experiments showed that necessary conditions for classical conditioning were not merely those of temporal association as indicated by Locke and Pavlov. In addition, the CS had to have predictive value; thus if the US occurred too frequently not preceded by the CS, the CS was no longer predictive and much reduced conditioning occurred, if any (Rescorla 1966, 1988). These matters become of special significance when interpreting the overall cognitive abilities of animals: Are many processes most properly interpreted as simple, automatic, stamping in of associa-

tions, or should they be considered as expectancies and predictions that organisms hold about their world?

The same issues arise in considering operant conditioning, the strengthening of responses which are followed by reinforcement, or colloquially, rewards. This is basically the question: How do organisms learn how to behave in the world? Are the laws governing response learning automatic, generally applicable processes? Can animals learn behaviors without responding at all? An example might be the ability to form a cognitive map simply from observation. An early experiment had cats towed about in carts through a maze, so they never made responses to be reinforced; nevertheless the cats later could walk correctly thorough the maze. This may not seem surprising to many readers, but to psychologists intent on establishing simple, noncognitive, stimulus-response laws; this was anathema.

Animals are capable of many advanced abilities; certainly Locke was wrong in proclaiming, "Brutes abstract not." Even lab pigeons can learn natural, humanmade or even arbitrary categories. Pigeons were trained to peck for food reward at various slides including: tree/non-trees, water/non-water, people/non-people, scenes with a particular person/scenes with other people or no people, the letter *A* in various fonts/other letters, fish/non-fish (a natural category but not one within a pigeon's usual experience) and a random selection of fish versus non-fish versus another random collection of the same types. The pigeons succeeded at all these discriminations as indicated by differential pecking rates and were able to generalize appropriately to novel instances. Interestingly the birds took far longer to learn the arbitrary sets. And they were capable of correctly categorizing together such examples of water as a droplet or a pond.

Precisely what the pigeons were learning is open to question and beyond the scope of this limited survey. It has not been definitively demonstrated that the birds had formed concepts of tree and non-tree; they could have pecked upon detecting leafiness or trunkness; they could have refrained from pecking at various sub-groupings rather than non-tree. Numerous other studies do not resolve the issue to the satisfaction of all, though at least some species, particularly ravens, parrots, and great apes, can form concepts to criteria acceptable by very many researchers.

COGNITIVE MAPS. The study of cognitive maps in animals has produced evidence of impressive abilities. After training, pigeons shown a photograph with objects and food in it can go correctly to that location in a lab room. Pigeons that have flown around a campus can, from an aerial photograph, learn to go to designated locations, including untrained sites (Honig 1991). In bird species that cache food for the winter, numerous experiments indicate that birds not only recall the placement of hidden seeds, but they recall better those seeds which they have hidden themselves. Experiments involving displaced landmarks indicate that rats and avian species studied use geometric information from their stored representations. Chimpanzees hide stones for later use as tools, and retrieve them using near optimal paths to do so. Succinctly put, pigeons, rats, and other species have been shown, with experimental evidence, to form concepts and cognitive maps, though the precise definitions of those terms is debated.

Animal knowledge of time presents a challenge to investigating scientists. There are many reports of animals returning at appropriate times to access regularly occurring food arrivals; the most notable may be the return of bees just before tea time each afternoon to the garden tea table of the famous bee scientist Karl von Frisch. Laboratory studies indicate that rats and pigeons can learn complicated schedules of responding for food, and can estimate time durations on the order of seconds very accurately.

But there are other aspects to the knowledge of time. As the philosopher Ludwig Wittgenstein noted when discussing Locke's ideas, "We say a dog is afraid his master will beat him, but not, he is afraid his master will beat him tomorrow." (Wittgenstein 1963, vol.1, p. 650). Relevant to this concern is research with scrub jays, a species that caches food for use at a later time; the work indicates use of elapsed time information in a fairly subtle way. According to the work of Nicola Clayton and colleagues (2003), these birds can discriminate and preferentially retrieve, depending on time elapsed since storage, either a preferred food (larvae) with a shorter time until decay or a less preferred food (peanuts), which lasts longer. Some of the ape cognition and language studies do include reports by apes of past occurrences, but those data do not appropriately tackle the issue of animal knowledge of time past, present, and future. In summary, by the current two-system hypothesis, both simple, automatic learning processes and more sophisticated cognitive skills are characteristic of both animals and humans.

MOTIVATIONAL, EMOTIONAL, AND MORAL CAPACITIES OF ANIMALS. These capacities have received far less investigation than have the cognitive. Motives and emotions have been studied in the laboratory and occa-

sionally in the field (Robert Sapolsky), particularly in reference to possible practical applications to humans. Thus theoretical and neurophysiogical/hormonal models have been proposed with regard to stress, addiction, learned helplessness, and depression. Experimental psychologists are in a dubious position, whereby some deny the possibility of animal consciousness or its scientific study, while others use animals as models for human emotions and motivations.

A possible evolutionary basis for morality reawakened research interest, beginning in the 1990s with neuroanatomical investigations and field studies. Apparent animal altruism has long intrigued scientists, resulting in theoretical models drawn, for example, from economic theorizing. Some suggest that the basis for human morality can be found in human's capacity for empathy, for understanding another's thoughts and feelings. Neurological studies confirm that merely viewing pictures of people injuring themselves, even stubbing their toes, activates some of the same brain regions that are engaged when people stub their own toes. Relevant animal research could be undertaken with potentially important results.

COMMUNICATION. Griffin suggests that animal communication may well serve as a window on animal minds, and thus provide evidence relevant to animal consciousness. Comparisons are frequently made to human linguistic communication, provoking agreement and controversy. To be discussed here are both natural and artificial communication systems.

Natural animal communication systems. Late-twentieth-century research has developed beyond the rigid stimulus-response model of classical ethology and the notion that at least some animal communication is merely a by-product of an internal state, what Griffin (1992) has termed the Groan of Pain (GOP) interpretation. Central issues now concern *what* is being communicated.

An important approach to communication by W. John Smith (1977) stresses an interactional, informational framework, which, however, has not received adequate attention. He notes that animals' signals by themselves do not provide sufficient information to enable recipients to choose appropriate responses. The context of the signal, including the roles of the specific interactants, their past history, and the environmental characteristics, all help determine meaning. This evaluation of information implies complex cognitive processes. In Smith's view, communication importantly allows interactants to predict the other's behavior; he avoids use of intentional terms, but his analyses are indeed amenable to such.

Beginning with mere insects, one finds surprising complexity and versatility in the genetically based dance communication system of honeybees. It has been known since the time of von Frisch, from studies begun in the 1920s, that the figure eight shaped waggle dances that honeybees perform inside their darkened hive convey information about the distance, direction, and desirability of a food source, though many academic battles were fought until that information was accepted. Later research indicated the dances could convey the same information about a potential new hive location, even including site height. The dance itself seems able to persuade other dancers to change their steps, and sometimes a recipient will begin to dance about a new location, sight unseen.

Of particular interest in the continuing controversy about the distinctions between human and animal communication, is the fact that several investigations indicate that some species' signals appear to be referential, that is, the calls specify the type of predator that has been detected. The species include vervet monkeys that appear able to indicate their three major predator types: the martial eagle, the leopard or other large carnivore, and the python. Diana and Campbells monkeys likewise have two different alarm calls, one for each of their major aerial and ground predators. Even some lemurs, primitive primate-like creatures, have calls specific to raptors, as does a mongoose species.

Sometimes level of arousal is included in the information of these various species' calls. Prairie dog calls reputedly identify predator types, even conveying information about the intruder's color and size, but the research needs further verification. Note, however, that the term *alarm call* is controversial, for some scientists, such as Smith, emphasize the broader use of some such calls. Peter Marler and his colleagues (1986) have also investigated reference in alarm and other calls, emphasizing the role of the audience, both that present and that to which a call is directed, in determining if a signal is given and which signal is made. It is also the case that many scientists are very reluctant to accept referential use of a signal by a nonhuman animal.

Artificial communication systems. Scientists have also undertaken studies of communication in apes and other species using modified forms of human sign language, plastic chips, computerized geometric figures (lexigrams), and spoken words. It is beyond the scope of this

entry to discuss these investigations fully, but it should be noted that some of the chimpanzees can respond to and produce strings of words similarly to the behaviors of a two-and-one-half-year-old human. That is not to say that the understandings of the humans and other species are the same. Whether the units should properly be termed words and whether the behavior should be termed language use is hotly debated (Terrace 1979); linguists are the strongest dissenters.

However both apes and African Grey parrots can use the communication units to indicate the color, number, and shape of items, and can accomplish cognitive tasks such as indicating same-different. Some of the apes understand and use artificial units, while also appropriately responding to some spoken English words. Apes have been reported to use the lexigrams to express simple thoughts and emotional feelings (Ristau and Robbins 1982, Ristau 1991, Savage-Rumbaugh 1998, and others).

ANIMAL CONSCIOUSNESS. To study consciousness, it is first necessary to delineate possible levels or kinds of consciousness, a task for both psychologists and philosophers. Since the topic is beyond the scope of this entry, note at least that consciousness can refer most simply to perceptual consciousness or awareness of sensations and pain and in more complex states to consciousness of self through past, present and anticipated future and to metacognition, or thoughts about one's thoughts and knowing *that* one knows. Yet even at a primitive level, it is difficult to imagine that a sensing creature, infant or animal, does not in some way distinguish between an external world and that which belongs to itself—such as its own paw.

Griffin has suggested the following as kinds of evidence for consciousness:

(1) An argument from evolution: Given that many other aspects of human structure and function are derived from those of other animal species, why should not consciousness likewise be part of the continuum?

(2) An argument from neurology: No Consciousness producing neurological structure or process can be found in humans, but absent from nonhuman animals. On the contrary, similar electrical brain waves are correlated to apparently similarly psychological functions in both humans and animals.

(3) As Griffin notes, "Appropriate responses to novel challenges for which the animal has not been prepared by genetic programming or previous experience provide suggestive evidence of animal consciousness because such versatility is most effectively organized by conscious thinking" (Griffin and Speck 2003, p. 5).

(4) Animal communication may well serve as a window to the minds of animals, revealing their subjective experiences, including intentions.

In his books and papers, Griffin (1976, 2001, 2003) reviews many experiments that provide evidence for consciousness. A few examples are noted. Beginning in the late 1970s, experiments examined the ability of chimpanzees to recognize themselves in a mirror (Gallup 1970). Children can do this after about eighteen months of age, but up to that time, they react socially to the mirror, interacting with their reflection as though it were another child. Chimpanzees also react socially, unless they have had extensive experience with mirrors. Results are mixed for other great apes, with controversial evidence from monkey species and no positive results from chickens and a myriad of other animals. Yet some monkey species, unable to recognize themselves by the mirror test, can nevertheless use a mirror to help them in a task with their otherwise unseen hand. Whatever the final evidence and interpretations, the mirror test can imply only some sense of the self as a body and not necessarily of the self as a mind, or as a self persisting from the past into the future.

A more limited claim, that rats can discriminate their own behaviors, derives from a task in which rats learned to push one of four different levers when a buzzer sounded, depending upon their own activity at the moment, for example, face washing, rearing, walking, or immobility (Beninger et al. 1974). Again interpretations of the results vary; for example, whether a rat is associating a particular lever with kinesthetic feedback from its behavior or whether a rat is indicating, " Now I am walking."

There is evidence that monkeys sometimes know what they know. As Griffin notes, "Consciously considering the contents of memory, in contrast to automatically using stored information, is a kind of metacognition, which many are still hesitant to infer in animals" (Griffin and Speck 2003, p. 13). Yet the ability to consciously consider uncertainties faced in nature is indeed an asset for an animal in a critical situation. In experiments investigating metacognition, monkeys had a choice of pressing one lever, thereby producing a less preferred food, or another lever requiring correct stimulus selection in order to receive a more preferred reward. Correct selection was difficult if monkeys had to delay their respond-

ing after seeing the stimulus they were to match. On such trials, the monkeys most often chose the less desirable reward, rather than take the test and quite likely get no reward. The author concludes that the monkeys can report the presence or absence of memory (Hampton 2001).

Creative tool making by crows, and indeed by other species, is another ability that strongly suggests conscious deliberation. There has long been considerable evidence of tool making by chimpanzees and orangutans, but less so for other species. New Caledonian crows studied in lab aviaries spontaneously used sticks to reach food in a clear cylinder, most often selecting sticks of the proper length. In other experiments, the crows selected a hooked wire, rather than a straight one to reach food most readily gained with a hook. When only a straight wire was available, the female crow, never having seen the process of wire bending, bent the wire herself to make a hook and thereby obtain the food.

Experiments have also been conducted in the field, suggesting purposeful, strategic behavior by the organisms involved. For example, Carolyn Ristau (1991) conducted experiments with piping plovers, birds that perform broken-wing or distraction displays at an intruder's approach to their nest/young. She suggested criteria for purposive behavior and found that the birds met such criteria. The plovers used the display correctly, so as to draw a human intruder away from the nest/young, positioning themselves in the intruder's front visual field. When plovers flew to reposition themselves before displaying, they went nearer the intruder and the center of the intruder's visual field. Plovers, even mid-display, monitored the intruders. Should the intruder not follow the birds' displays, the plovers modified their behaviors, re-approaching the intruder or increasing display intensity. Other experiments indicated the plovers' awareness of the direction of an intruder's attention, by becoming more aroused when a passing intruder looked toward their nest area in the dunes in contrast to looking towards the sea. Alexandra Horowitz (2002) has further developed Ristau's criteria for intentional behavior and has applied the ideas to dogs' interactive behavior.

In research by David Premack (1978, 1992), Daniel Povinelli (2000), Michael Tomasello (1997), and Frans de Waal (2003) and their colleagues, chimpanzees have been shown to be capable of complex problem solving and social understanding, sometimes interpreted as the ability to attribute and to understand other minds. Such abilities include determining the intentions of others, detecting, understanding and engaging in deception, and distinguishing between knowledge held by another in contrast to another's visual perception. Many aspects of these capacities seem reasonable evidence for consciousness.

In summary, though unresolved in the view of some, many behavioral scientists appear to be coming to agree that animals are conscious. The matter of proof of the content of mental states of any creature, human or otherwise, remains a philosophical problem. There simply are no incontrovertible means by which external behaviors, linguistic or otherwise, provide absolute proof of specific mental states. One can be certain only of one's own consciousness; this is the extreme version of the solipsistic position.

PHILOSOPHICAL IMPLICATIONS

The essential problem confronting the study of animal minds as conscious entities is that of solipsism. However, in order to survive in daily life, one cannot accept the solipsistic position. In science, one can at least recognize that to declare that animals are not conscious is not a neutral stance, but one that demands proof. As Griffin notes, the probability of awareness (pA) must be assumed to be 0.5, not 0.0. So the scientific task becomes one of accumulating evidence that shifts pA in either direction, noting that level of awareness for a particular task does not necessarily imply the organism's consciousness during another task.

PHILOSOPHICAL PROBLEMS. Several traditional philosophical lines of inquiry are to be considered in the study of animal mind, certainly the philosophy of science and of mind including the nature of scientific evidence, solipsism, nature of experience (e.g., *qualia*), intentionality and gradations of belief-desire systems, linguistic concerns, nature of a referent and of representation, nature of specific cognitive capacities, and defining levels of awareness/consciousness and at least suggestive evidence for each.

POTENTIAL ROLES FOR PHILOSOPHERS

In the past, philosophers were usually dismissive of the need for scientific data in pursuing philosophical problems. Fortunately, that attitude has changed. Philosophers cognizant of the data in their area of interest can play much needed roles in elucidating unidentified assumptions in scientists' work. They can suggest the kinds of data and experimental designs required to provide insight into mental states. Philosophical examinations of Kantian

and other concepts of space and time as relevant to animal minds would likewise be helpful.

But philosophers also need to accept real-world constraints on their thinking, prime among them being temporal: Organisms act in a time-limited world and often the most dangerous situations they face require making very rapid decisions. Thus organisms often operate using default mechanisms as well as more time-consuming, deliberative, or trial-and-error methods. Organisms, both animal and human, are often overloaded with information; thus simple heuristics must often suffice. Aware of constraints such as these, as well as the need to communicate effectively to those in other fields, philosophers' contributions to the understanding of animal mind can be outstanding.

See also Animal Rights and Welfare; Aristotle; Bennett, Jonathan; Darwin, Charles Robert; Dennett, Daniel Clement; Descartes, René; Locke, John; Millikan, Ruth; Pavlov, Ivan Petrovich; Qualia; Searle, John; Skinner, B. F.; Speciesism; Wittgenstein, Ludwig Josef Johann.

Bibliography

Beninger, R. J., S. R. Kendall, and C. H. Vanderwolf, "The Ability of Rats to Discriminate Their Own Behavior." *Canadian Journal of Psychology* 28 (1974): 79–91.

Bennett, Jonathan Francis. *Rationality: An Essay towards an Analysis*. Indianapolis, IN: Hackett, 1989.

Clayton Nicola S, Kara Shirley Yu, and Anthony Dickinson. "Interacting Cache Memories: Evidence for Flexible Memory Use by Western Scrub-Jays (*Aphelocoma californica*)." *Journal of Experimental Psychology: Animal Behavior Processes* 29 (2003): 14–22.

Darwin, Charles R. *The Expression of the Emotions in Man and Animals*. London: Appleton, 1872.

Dennett, Daniel C. *Kinds of Minds*. New York: Basic Books, 1996.

Frisch, Karl von. *The Dance Language and Orientation of Bees*. Cambridge, MA: Harvard University Press, 1967.

Gallup, Gordon G. "Chimpanzees: Self-Recognition." *Science*. Vol. 167 (3914) (January 1970): 86–87.

Garcia, John, and R. A. Koelling. "Relation of Cue to Consequence in Avoidance Learning." *Science* 4 (1966): 123–124.

Griffin, Donald R. *Animal Minds, Beyond Cognition to Consciousness*. Chicago: University of Chicago Press, 2001.

Griffin, Donald R. *The Question of Animal Awareness*. New York: Rockefeller University Press, 1976.

Griffin, Donald R., and Gayle B. Speck. "New Evidence of Animal Consciousness." *Animal Cognition* 7 (1) (2003): 5–18.

Hampton, Robert Russell. "Rhesus Monkeys Know When They Remember." *Proceedings of the National Academy of Sciences USA* 98 (2001): 5359–5362.

Honig, Werner K. "Structure and Function in the Spatial Memory of Animals," edited by Wickliffe C. Abraham, Michael Corballis, et al. In *Memory Mechanisms: A Tribute to G. V. Goddard*, pp. 293–313. Hillsdale, NJ: Lawrence Erlbaum Associates, 1991.

Horowitz, Alexandra C. "The Behaviors of Theories of Mind and a Case Study of Dogs at Play." PhD diss., University of California, San Diego, 2002. Available from http://www.crl.ucsd.edu/~ahorowit/frontmatters.pdf.

Locke, John. *An Essay concerning Human Understanding* (1690). Printed by Elizabeth Holt for Thomas Basset. Reprinted by Pelagus Free Books. www.pelagus.org/books/AN_ESSAY_CONCERNING_HUMAN_ UNDERSTANDING,_by_John_Locke_31.html

Marler, Peter, Alan Dufty, and Roberta Pickert. "Vocal Communication in the Domestic Chicken: II. Is a Sender Sensitive to the Presence and Nature of a Receiver?" *Animal Behaviour* 34 (1986): 194–198.

Millikan, Ruth Garrett. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press, 1984.

Millikan, Ruth Garrett. *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press, 1995.

O'Keefe, John, and Lynn Nadel. *Hippocampus as a Cognitive Map*. Oxford: Clarendon Press, 1978.

Povinelli, Daniel J. *Folk Physics for Apes: The Chimpanzees' Theory of How the World Works*. New York: Oxford University Press, 2000.

Pepperberg, Irene M. "Evolution of Avian Intelligence, with an Emphasis on Grey Parrots (*Psittacus erithacus*)." In *The Evolution of Intelligence*, edited by R. S. Sternberg and J. C. Kaufman, 315–337. Mahwah, NJ: Lawrence Erlbaum, 2002.

Premack, David. "On the Origins of Domain-Specific Primitives." *Cognition: Conceptual and Methodological Issues*, edited by Herbert L Pick, Jr. and Paulus Willem van den Broek, 189–212. Washington, DC: American Psychological Association, 1992.

Premack, David, and G. Woodruff. "Does the Chimpanzee Have a Theory of Mind?" *Behavioral and Brain Sciences* 4 (1978): 515–526 and commentaries.

Rescorla, Robert A. "Pavlovian Conditioning: It's Not What You Think It Is." *American Psychologist* 43 (1988): 151–160.

Rescorla, Robert A. "Predictability and Number of Pairings in Pavlovian Fear Conditioning." *Psychonomic Society* 4 (1966): 383–384.

Ristau, Carolyn A. "Aspects of the Cognitive Ethology of an Injury-Feigning Bird, the Piping Plover." In *Cognitive Ethology: The Minds of Other Animals: Essays in Honor of Donald R. Griffin*, edited by Carolyn A. Ristau, 91–126. Hillsdale, NJ: Lawrence Erlbaum, 1991.

Ristau, Carolyn A., and Donald Robbins. "Language in the Great Apes: A Critical Review." In *Advances in the Study of Behavior*. Vol. 12, edited by J. S. Rosenblatt, et al., 142–255. New York: Academic Press, 1982.

Romanes, George J. *Animal Intelligence*. London: Kegan Paul, 1882.

Savage-Rumbaugh Sue, S. G. Shanker, and T. Taylor. *Apes, Language, and the Human Mind*. New York: Oxford University Press, 1998.

Searle, John R. "Consciousness." *Annual Review of Neuroscience* 23 (2000): 557–578.

Smith, W. John. *The Behavior of Communicating*. Cambridge, MA: Harvard University Press, 1977.
Terrace, H. S. *Nim*. New York: Knopf, 1979.
Tolman, Edwin C. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (1948): 189–208.
Tomasello, Michael, and Josep Call. *Primate Cognition*. New York: Oxford University Press, 1997.
Waal, Frans B. M. de, and Peter L. Tyack. *Animal Social Complexity: Intelligence, Culture, and Individual Societies*. Cambridge, MA: Harvard University Press, 2003.
Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Basil Blackwell, 1963.

Carolyn A. Ristau (2005)

ANIMAL RIGHTS AND WELFARE

Although all the major moral philosophers in the Western tradition have had something to say about the moral status of animals, they have commented infrequently and for the most part only in brief. This tradition of neglect changed dramatically during the last quarter of the twentieth century, when dozens of works in ethical theory, hundreds of professional essays, and more than a score of academic conferences were devoted to the moral foundations of human treatment of nonhuman animals.

Two main alternatives—animal welfare and animal rights—have come to be recognized. Animal welfarists accept the permissibility of human use of nonhuman animals as a food source and in biomedical research, for example, provided such use is carried out humanely. Animal rightists, by contrast, deny the permissibility of such use, however humanely it is done.

Differ though they do, both positions have much in common. For example, both reject Descartes's view that nonhuman animals are *automata*. Those animals raised for food and hunted in the wild have a subjective presence in the world; in addition to sharing sensory capacities with human beings, they experience pleasure and pain, satisfaction and frustration, and a variety of other mental states. There is a growing consensus that many nonhuman animals have a mind that, in Charles Darwin's words, differs from the human "in degree and not in kind."

Proponents of animal welfare and animal rights have different views about the moral significance of human psychological kinship with other animals. Animal welfarists have two options. First, they can argue that we ought to treat animals humanely because this will lead us to treat one another with greater kindness and less cruelty. On this view we have no duties to animals, only duties involving them; and all those duties involving them turn out to be, as Kant wrote, "indirect duties to Mankind" (Immanuel Kant, "Duties to Animals," in Regan and Singer, 1991, p. 23). Theorists as diverse as Kant, St. Thomas Aquinas, and John Rawls favor an indirect-duty account of the moral status of nonhuman animals.

Second, animal welfarists can maintain that some of our duties are owed directly to animals. This is the alternative favored by utilitarians, beginning with Jeremy Bentham and John Stuart Mill and culminating in the work of Peter Singer (1990). Animal pain and pleasure count morally in their own right, not only indirectly through the filter of the human interest in having humans treated better. The duty not to cause animals to suffer unnecessarily is a duty owed directly to animals.

Of the two options the latter seems the more reasonable. It is difficult to understand why the suffering of animals should count morally only if it leads to human suffering in the future. Imagine that a man sadistically tortures a dog and dies of a heart attack as a result of his physical exertion; what he does seems clearly wrong even though he does not live long enough to mistreat a human being. If this is true, then we have at least some direct duties to animals.

Animal welfarists who are utilitarians (Singer is the most notable example) use utilitarian theory to criticize how animals are treated in contemporary industries (animal agriculture and biomedical research, for example). For in these industries animals are made to suffer and, Singer alleges, to suffer unnecessarily.

Other animal welfarists who are utilitarians disagree. Government and industry leaders agree that some animals sometimes suffer in the course of being raised for food or used in biomedical research; but they deny that they are made to suffer unnecessarily.

Consider organ transplant research. Research on animals in this quarter involves transplanting some internal organ from one healthy animal to another; the "donor" animal, who is under anesthetic, is killed, but the "receiver" animal is permitted to recover and doubtless experiences no small amount of postoperative pain before being humanely killed.

Is the pain unnecessary? In one sense it clearly is. For since the organ was not transplanted for the good of the recipient animal, all the pain that animal experienced was unnecessary. However, this is not the real question, given the utilitarian perspective. The pain caused to this partic-

ular animal is only one part of the overall calculation that needs to be carried out. One also needs to ask about the possible benefits for humans who are in need of organ transplants, the value of the skills surgeons acquire carrying out animal organ transplants, the value of knowledge for its own sake, and so on. After these questions have been answered and the overall benefits impartially calculated, then an informed judgment can be made about whether organ transplant research involving nonhuman animals does or does not cause unnecessary suffering.

As this example illustrates, animal welfarists who are utilitarians can disagree about when animals suffer unnecessarily. As such, these animal welfarists can differ in judging whether animals are being treated humanely and, if not, how much reform is called for.

Advocates of animal rights advance a position that avoids the always daunting, frequently divisive challenge of carrying out uncertain utilitarian calculations. Central to their view is the Kantian idea that animals are never to be treated merely as a means to human ends, however good these ends might be. The acquisition of knowledge, including biological knowledge, is surely a good end, as is the promotion of human health. But the goodness of these ends does not justify the utilization of nonhuman animals as means. Thus, even if animal-model organ transplant research can be justified on utilitarian grounds, animal rights advocates would judge it immoral.

Of the two main options—animal welfare and animal rights—it is the latter that attempts to offer a basis for a radical reassessment of how animals are treated. Animal welfare, provided the calculations work out a certain way, enables one to call for reforms in human institutions that routinely utilize nonhuman animals. But animal rights, independent of such calculations, enables one to call for the abolition of all forms of institutional exploitation.

However these matters are resolved, one should note the major contribution philosophers have made in placing the "animal question" before a wider audience. Despite their philosophical differences, none of the philosophers participating in the debate is satisfied with how animals are treated by the major animal user industries. This consensus has meant that those who manage these industries have had to respond to new forms of moral criticism. Collectively, these philosophers have been and will continue to be a powerful voice calling for better treatment of animals.

In addition, the interest philosophers have shown in the "animal question" has spilled over into other disciplines, including sociology, history, anthropology, and law. The latter is of particular interest. Whereas thirty years ago not a single law school in America offered courses on animals and the law, upwards of thirty do so today. The evidence suggests that a new field of inquiry, Human and Animal Studies, is in the offing.

See also Darwin, Charles Robert; Descartes, René; Speciesism; Utilitarianism.

Bibliography

Carruthers, Peter. *The Animals Issue*. Cambridge, U.K.: Cambridge University Press, 1992.

Clark, S. S. L. *The Moral Status of Animals*. Oxford: Oxford University Press, 1977.

Cohen, C. and T. Regan. *The Animal Rights Debate*. Lanham, MD: Rowman and Littlefield, 2001.

Favre, David. *Animals: Welfare, Interests, and Rights*. East Lansing, MI: Animal Legal & Historical Center, 2003.

Francione, G. *Animals, Property, and the Law*. Philadelphia: Temple University Press, 1995.

Frey, R. G. *Interests and Rights: The Case Against Animals*. Oxford: Oxford University Press, 1980.

Frey, R. G. *Rights, Killing, and Suffering*. Oxford: Basil Blackwell, 1983.

Leahy, M. *Against Liberation: Putting Animals in Perspective*. Lanham, MD: Rowman and Littlefield, 1991.

Machan, T. *Putting People First: Why Humans are Favored by Nature*. Lanham, MD: Rowman and Littlefield, 2004.

Midgley, M. *Animals and Why They Matter*. New York: Thompson-Shore, 1983.

Pluhar, Evelyn. *Beyond Prejudice: The Moral Significance of Human and Nonhuman Animals*. Durham: Duke University Press, 1995.

Rachels, J. *Created from Animals: The Moral Implications of Darwinism*. Oxford: Oxford University Press, 1990.

Regan, Tom. *Animal Rights, Human Wrongs: An Introduction to Moral Philosophy*. Lanham, MD: Rowman and Littlefield, 2004.

Regan, Tom. *The Case for Animal Rights*. 2nd ed. Berkeley: University of California Press, 2004.

Regan, Tom, and Peter Singer, eds. *Animal Rights and Human Obligations*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

Rollin, B. *Animal Rights and Human Morality*. Buffalo, NY: Prometheus Press, 1981.

Rowland, Mark. *Animal Rights: A Philosophical Defense*. New York: MacMillan, 1998.

Sapontzis, S. *Morals, Reason, and Animals*. Philadelphia: Temple University Press, 1981.

Scruton, R. *Animal Rights and Wrongs*. London: Metro Press, 2000.

Singer, Peter. *Animal Liberation*. 2nd ed. New York: New York Review of Books: Distributed by Randon House, 1990.

Singer, Peter, ed. *In Defense of Animals*. New York: Blackwell, 1985.

Wise, Steven M. *Rattling the Cage: Toward Legal Rights for Animals*. Cambridge, U.K.: Cambridge University Press, 2000.

Tom Regan (1996, 2005)

ANIMAL SOUL

See *Animal Mind*

ANIMISM

See *Macrocosm and Microcosm; Panpsychism*

ANNET, PETER

(1693–1769)

Peter Annet, an English freethinker and deist, was by profession a schoolmaster. He lost his employment in 1744 because of his outspoken attacks on certain Christian apologists. A debater at the Robin Hood Society (named after a public house where the meetings were held), he soon became a popular lecturer. The first published result was a pamphlet of 1739, titled *Judging for Ourselves: Or Free-Thinking, the Great Duty of Religion. Display'd in Two Lectures, deliver'd at Plaisterers-Hall,* "By P. A. Minister of the Gospel. With A Serious Poem address'd to the Reverend Mr. Whitefield." The tone of the work is indicated by the statement: "If the Scriptures are Truth, they will bear Examination; if they are not, let 'em go." This was followed by several tracts directly attacking Thomas Sherlock, bishop of London: *The Resurrection of Jesus Considered: In Answer To the Tryal of the Witnesses* "By a Moral Philosopher," which ran through three editions in 1744; *The Resurrection Reconsidered* (1744); *The Sequel of the Resurrection of Jesus Considered* (1745); and *The Resurrection Defenders stript of all Defence* (1745).

In *Social Bliss Considered* (1749) Annet, like John Milton before him, advocated the liberty of divorce. He answered Gilbert West's *Observations on the Resurrection of Jesus Christ* (1747) in *Supernaturals Examined* (1747) and George Lyttleton's *Observations on the Conversion and Apostleship of St. Paul in a Letter to Gilbert West* (1747) in *The History and Character of St. Paul Examined* (1748). Arguing that all miracles are incredible, Annet proceeded to attack Old Testament history in his journal, *The Free Enquirer* (9 numbers, October 17, 1761–December 12, 1761). For this work he was accused of blasphemous libel before Lord Mansfield in the Court of King's Bench in the Michaelmas term of 1762. There is some evidence that Lord Mansfield, urged on by Bishop Warburton and others, used Annet as a scapegoat after a fruitless attempt had been made to suppress the publication of David Hume's *Four Dissertations* of 1757.

Annet pleaded guilty to the charge. "In consideration of which, and of his poverty, of his having confessed his errors in an affidavit, and of his being seventy years old, and some symptoms of wildness that appeared on his inspection in Court; the Court declared they had mitigated their intended sentence to the following, viz., to be imprisoned in Newgate for a month; to stand twice in the pillory [Charing Cross and the Royal Exchange] with a paper on his forehead, inscribed Blasphemy; to be sent to the house of correction [Bridewell] to hard labour for a year; to pay a fine of 6s.8d.; and to find security, himself to 100 £ and two sureties in 50 £. each, for his good behaviour during life." Having survived this "mitigated," charitable, and humane punishment based on the iniquitous Blasphemy Act of 1698, Annet returned to schoolmastering. Archbishop Secker is said to have so far relented as to afford aid to the culprit until his death in 1769. In 1766 Annet issued *A Collection of Tracts of a Certain Free Enquirer noted by his sufferings for his opinions,* a work containing all of the tracts mentioned above.

Annet was long thought to have been the author of *The History of the Man after God's Own Heart* (1761), in which the writer took exception to a parallel drawn by a divine between George II and King David. The anonymous writer argued that such a comparison was an insult to the late king. Recent scholarship has proved that the real author was John Noorthouck, a respected member of the Stationers' Company.

Among his accomplishments, Annet was the inventor of a system of shorthand. Unlike most of the leading English deists, Annet had relatively little formal education and spoke and wrote plainly and forcefully directly to the masses. He was the last to suffer physical punishment for his heterodox religious opinions.

See also Deism; Hume, David; Milton, John.

Bibliography

There is no collected edition of Annet's works, and the texts mentioned in the article are extremely rare. A useful article, however, regarding the authorship of *The History of the Man after God's Own Heart,* is the anonymous "John Noorthouck, 'The Man after God's Own Heart,'" in *Times Literary Supplement* (August 25, 1945): 408.

See also *The English Reports,* Vol. 96, King's Bench Division, XXV (Edinburgh and London), 1909; and E. C. Mossner, "Hume's *Four Dissertations*: An Essay in Biography and Bibliography," in *Modern Philology* 48 (1950): 37–57.

Ernest Campbell Mossner (1967)

ANOMALOUS MONISM

Originated by Donald Davidson, "anomalous monism" is a nonreductive, token physicalist position on the relation between the mental and the physical. According to it, each mental event is a physical event, although mental descriptions are neither reducible to nor nomologically correlated with physical ones. In terms that are ontologically more robust than those used by Davidson, the position asserts identities between individual mental and physical events while denying that mental types or properties are either identical with, or nomologically connected with, physical ones. The position specifically concerns intentional mental phenomena such as beliefs and desires, although it is arguable that it can be extended to cover other mental phenomena such as sensations.

Davidson's argument for this position results from an attempt to reconcile three apparently inconsistent principles, two of which he finds independently plausible and the third of which he defends at length. The first is the principle of causal interaction (PCI), which states that mental events cause physical events and vice versa, causality being understood as relating events in extension. The second is the principle of the nomological character of causality (PNCC), which states that events that are causally related have descriptions under which they instantiate strict causal laws. The third is the principle of the anomalism of the mental (PAM), which states that there are no strict laws in which mental terms figure. The principles appear to conflict in that the first two imply what the third seems to deny—namely that there are strict laws governing causal interactions between mental and physical events.

Davidson argues that the principles can be reconciled by adopting the thesis that each mental event has a physical description and so is a physical event. He further suggests that a sound argument can be constructed from these principles to this thesis. Suppose a mental event, *m,* causes a physical event, *p.* Then, by the PNCC, *m* and *p* have descriptions under which they instantiate a strict causal law. By PAM this cannot be mental in that it cannot contain mental terminology. Therefore *m* must have a physical description under which it instantiates a strict causal law, which is to say that it is a physical event. Although the argument is formulated in terms of events and their descriptions, it can be formulated equally effectively in the terminology of events and their properties.

Davidson does not take PAM to be obvious. His defense of it involves the idea that laws bring together terms from the same or similar conceptual domains. Using this idea he argues that the constraints that govern the application of mental terms and their associated concepts to things are normative in nature, involving "constitutive" principles of rational coherence, deductive and inductive consistency, and the like. These principles constitute the distinctive rationalistic normativity that is the earmark of the intentional domain; and Davidson argues that they have no place in physical theory.

The argument for anomalous monism appears to work because of the extensionality of the causal relation and the intensionality of nomologicality. Events are causally related no matter how described; but they are governed by laws only as they are described one way rather than another. This opens up a conceptual space between causality and nomologicality that makes it possible to hold both that mental events that interact causally with physical ones are governed by laws and that there are no strict psychological or psychophysical laws.

Davidson's argument has had a profound effect on discussions of mental causation and token physicalism. Many have found either the PNCC or the PAM questionable and have taken issue with it. However, the main objection to the argument is that, on a certain conception of the relation between causality and laws, it leads either to inconsistency or to epiphenomenalism. According to this conception, laws link events causally by linking certain, but not all, of their descriptions or properties, the causally relevant ones. The question now arises, In virtue of which of their properties do mental events interact causally with physical ones? If the answer is the mental ones, then anomalous monism is threatened with inconsistency since this implies that there are laws in which mental descriptions/properties figure. If the answer is the physical ones, then anomalous monism is threatened with epiphenomenalism since it is in virtue of their physical properties that mental events are causally efficacious. Since PAM is a crucial premise in the argument for anomalous monism, it is the epiphenomenalism charge that poses the real threat to the position.

There is a general question of whether nonreductive token physicalist theories count as proper forms of physicalism since they recognize the existence of irreducibly mental properties. Davidson himself favors supplement-

ing his position with some sort of supervenience thesis, according to which, necessarily, if things (events) are the same with regard to their physical descriptions/properties, then they are the same with regard to their mental descriptions/properties. The principal difficulty in formulating such a thesis is in specifying a dependency relation strong enough to ensure that physical properties determine mental ones without leading to reducibility and hence to type physicalism.

See also Davidson, Donald; Mental Causation; Philosophy of Mind; Physicalism; Supervenience.

Bibliography

Davidson, D. "Mental Events." In *Experience and Theory,* edited by L. Foster and J. W. Swanson. Amherst: University of Massachusetts Press, Reprinted in D. Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980). The classic statement of the argument for anomalous monism.

Davidson, D. "Psychology as Philosophy." In *Philosophy of Psychology,* edited by S. C. Brown. London: Macmillan, 1974. Reprinted in D. Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980). Discusses anomalous monism and the argument against psychophysical laws.

Davidson, D. "Thinking Causes." In *Mental Causation,* edited by J. Heil and A. Mele. Oxford: Clarendon Press, 1993. Responds to the charge that anomalous monism leads to the causal inefficacy of the mental.

Honderich, T. "The Argument for Anomalous Monism." *Analysis* 42 (1982). Classic statement of the inconsistency-or-epiphenomenalism objection to anomalous monism.

Kim, J. "The Myth of Nonreductive Materialism." *Proceedings of the American Philosophical Association* 63 (1989): 31–47. Argues that nonreductive materialism leads to epiphenomenalism.

Kim, J. "Psychophysical Laws." In *Actions and Events: Perspectives on the Philosophy of Donald Davidson,* edited by E. LePore and B. McLaughlin. Oxford: Blackwell, 1985. Discusses and defends an interpretation of Davidson's argument against psychophysical laws.

LePore, E., and B. Loewer. "Mind Matters." *Journal of Philosophy* 84 (1987): 630–641. Discusses the causal efficacy of the mental within the context of physicalism.

LePore, E., and B. McLaughlin, eds. *Actions and Events: Perspectives on the Philosophy of Donald Davidson.* Oxford: Blackwell, 1985. Articles on Davidson's argument for anomalous monism.

Macdonald, C. *Mind-Body Identity Theories.* London: Routledge, 1989. Surveys various type-type and token identity theories, and defends a version of nonreductive monism.

Macdonald, C., and G. Macdonald. "Mental Causes and Explanation of Action." *Philosophical Quarterly* 36 (1986): 145–158. Reprinted in *Mind, Causation, and Action,* edited by L. Stevenson, R. Squires, and J. Haldane (Oxford: Blackwell, 1986). Defends anomalous monism against the charge of epiphenomenalism.

McLaughlin, B. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989). Discusses the problem of mental causation for anomalous monism.

Cynthia Macdonald (1996)

ANSCOMBE, GERTRUDE ELIZABETH MARGARET

(1919–2001)

G. E. M. Anscombe, English philosopher, was educated at Sydenham High School and St. Hugh's College, Oxford, where she read *Literae Humaniores* (Greats). She went as a research student to the University of Cambridge, where she became a pupil of Ludwig Wittgenstein (1889–1951). He and Aristotle were the most important influences on her philosophical thought. Anscombe became a Roman Catholic while in her teens, and her Catholicism was also a shaping influence. She was a Fellow for many years of Somerville College, Oxford, and held the Chair of Philosophy at the University of Cambridge from 1970 until 1986. A philosopher of great range, she made important contributions to ethics, philosophy of mind and action, metaphysics, epistemology, philosophical logic, and philosophy of language. Much of her most interesting work was in the history of philosophy; her discussions of ancient, medieval, and modern philosophers combine illuminating accounts of challenging texts with penetrating treatment of the philosophical problems themselves. As one of Wittgenstein's literary executors, as an editor and translator of his writings, and as a writer and lecturer about Wittgenstein, she has done more than anyone else to make his work accessible. Her *Introduction to Wittgenstein's "Tractatus"* (1959) is a superb introduction to the central themes of that work, making clear the character of the problems (like that of negation) treated in it.

Long before it became fashionable in the 1970s for moral philosophers to concern themselves with practical problems, Anscombe was writing about them. Her first published essay, in 1939, shortly after the beginning of World War II, concerned the justice of that war. She discussed closely related topics in her protest against the honorary degree that Oxford University awarded Harry Truman in 1957 and in connection with the policy of nuclear deterrence. She wrote also on contraception, murder, and euthanasia. All her writings on such questions reflect her belief that the concepts of action and intention are important for ethics, especially in connection with questions about our responsibility for the con-

sequences of our actions. She explained and defended the doctrine of double effect (it is sometimes permissible to cause, as a side effect, a merely foreseen harm that is forbidden if sought intentionally). She argued that denial of this doctrine "has been the corruption of non-Catholic thought and its abuse the corruption of Catholic thought" (1981, 3:54).

Anscombe's interest in war and in the concept of murder led her also to more general philosophical questions about political authority. "Modern Moral Philosophy" (1981, 2005) has been the most influential of her papers on ethics and was an important impetus for the development of virtue ethics (which emphasizes the character traits a human being needs in order to flourish). She defended three theses in the paper: that moral philosophy cannot be done until we have an adequate philosophical psychology; that the concepts of moral obligation, moral duty, and moral "ought" are survivals from a now largely abandoned conception of ethics, are incoherent outside that framework, and should therefore be abandoned if possible; and that English moral philosophers from Henry Sidgwick (1838–1900) on differ only in superficial ways. In explaining the third thesis, Anscombe introduced the term "consequentialism" for the common view that right and wrong are determined by consequences (including among consequences the promotion of intrinsic values), and she argued that consequentialism is a corrupt and shallow philosophy.

In her ground-breaking monograph *Intention* (1957), Anscombe raised and discussed questions about intention, action, and practical thought (practical reasoning and practical knowledge). Widely prevalent philosophical ideas about intention had treated it as some special kind of mental state or event. Departing radically from that tradition, Anscombe gave an account of intentional action in terms of the applicability to it of a kind of question asking for the agent's reason. This account enabled her to show how conceptions of good are important for practical thought. The questions with which Anscombe was concerned frequently straddled metaphysics, philosophy of logic, and philosophy of mind. For example, in "The First Person" (1981), she explained how we are led into confusion by misunderstandings of "I" on the model of a proper name. In "The Intentionality of Sensation: A Grammatical Feature" (1981) she drew on philosophy of language in explaining grammatical analogies between intention and sensation, and was able to give a very interesting and original account of what is right in sense-impression philosophy and of what is misleading in it.

Anscombe explored the topic of causation in several papers, questioning in them widely held assumptions. "Causality and Determination" (1981) begins by formulating two such assumptions: that causality is a necessary connection of some kind, and that it involves a universal generalization connecting events of two kinds. One or the other or both of the assumptions are accepted by virtually all writers on causation, but Anscombe questioned both, together with the related idea that if two courses of events appear similar but have different outcomes, there must be some further relevant difference. She argued that the root idea in all our causal notions is that of one thing deriving from another, and that this need not involve necessitation. In "Times, Beginnings, and Causes" (1981) she challenged two widely accepted views of Hume's: that causal relations are never logically necessary, and that logically something can begin to exist without being caused to do so. Questions about time figure centrally in other papers as well. For example, in "The Reality of the Past" (1981) she treats a problem raised by Parmenides (b. c. 515 BCE) and shows how attempts to explain the concept of the past by reference to memory must fail. This paper also contains one of the best short discussions of Wittgenstein's later approach to philosophy.

While Anscombe worked within the tradition of twentieth-century analytic philosophy, she challenged many of the assumptions of her contemporaries. Although her work, especially on intention and action, has exercised wide influence, much of her thought has not yet been assimilated, owing partly to the fact that she maintained a critical distance from the ideas of her contemporaries and partly to the fact that many of her later papers are not readily accessible.

See also Aristotle; Consequentialism; Euthanasia; Feminist Philosophy; Intention; Metaphysics, History of; Philosophy of Language; Philosophy of Mind; Sidgwick, Henry; Virtue Ethics; Wittgenstein, Ludwig Josef Johann; Women in the History of Philosophy.

Bibliography

WORKS BY ANSCOMBE

Intention. Oxford: Blackwell, 1957. 2nd ed., 1963.

An Introduction to Wittgenstein's "Tractatus." 3rd ed. London: Hutchinson, 1959, 1971.

With P. T. Geach. *Three Philosophers*. Oxford: Blackwell, 1963.

Collected Philosophical Papers. Vol. 1: *From Parmenides to Wittgenstein*. Vol. 2: *Metaphysics and the Philosophy of Mind*. Vol. 3: *Ethics, Religion, and Politics*. Oxford: Blackwell, 1981.

Human Life, Action and Ethics: Essays by G. E. M. Anscombe, edited by Mary Geach and Luke Gormally. Exeter: Imprint Academic, 2005.

WORKS ON ANSCOMBE

Diamond, Cora, and Jenny Teichman, eds. *Intention and Intentionality: Essays in Honour of G. E. M. Anscombe*. Brighton, U.K.: Harvester Press, 1979.

Gormally, Luke, ed. *Moral Truth and Moral Tradition: Essays in Honour of Peter Geach and Elizabeth Anscombe*. Dublin: Four Courts Press, 1994.

Haber, Joram Graf, ed. *Absolutism and Its Consequentialist Critics*. Lanham, MD: Rowman & Littlefield, 1994.

Müller, Anselm. "G. E. M. Anscombe (1919–2001)." In *A Companion to Analytic Philosophy*, edited by A. P. Martinich and David Sosa, 315–325. Oxford: Blackwell, 2001.

Teichman, Jenny. "Gertrude Elizabeth Margaret Anscombe, 1919–2001." *Proceedings of the British Academy* 115 (2001): 31–50.

Teichmann, Roger, ed. *Logic, Cause, and Action: Essays in Honour of Elizabeth Anscombe*. Cambridge, U.K.: Cambridge University Press, 2000.

Cora Diamond (1996, 2005)

ANSELM, ST.

(1033–1109)

The greatest philosopher of the eleventh century, Anselm of Canterbury was the author of some dozen works whose originality and subtlety earned him the title of "Father of Scholasticism." Best known in the modern era for his "Ontological Argument," designed to prove God's existence, Anselm made significant contributions to metaphysics, ethics, and philosophy of language.

Anselm was born in Aosta, in the Piedmont region of the kingdom of Burgundy, near the border with Lombardy. His family was noble but of declining fortunes. Anselm remained at home until he was twenty-three; after the death of his mother he quarreled irrevocably with his father and left home, wandering for some years before arriving at the Benedictine Abbey at Bec in Normandy. Impressed by the abbey's prior Lanfranc, who had a reputation as a scholar and teacher of dialectic, Anselm joined the monastery as a novice in 1060. Such was his ability that in 1063 he was elected prior and in 1078 abbot, a position he held until his elevation as archbishop of Canterbury in 1093. While at Bec, Anselm wrote his *Monologion*, *Proslogion*, and the four dialogues *De grammatico*, *De veritate*, *De libertate arbitrii*, and *De casu Diaboli*. While archbishop, Anselm wrote his *De incarnatione Verbi*, *Cur Deus homo*, *De conceptu virginali*, *De processione Spiritus Sancti*, and *De concordia*. Perhaps from this time also date his fragmentary notes on power, ability, and possibility. Anselm's archepiscopate was marked by controversy with the English kings William Rufus and Henry I over royal privileges and jurisdiction; Anselm spent the years from 1097 to 1100 and from 1103 to 1107 in exile. After a brief illness, Anselm died on April 21, 1109, in Canterbury, where he is interred in the Cathedral.

METHOD

Most of Anselm's work systematically reflects on the content of Christian doctrine: Trinity, Incarnation, the procession of the Holy Spirit, original sin, the fall of Lucifer, redemption and atonement, virgin conception, grace and foreknowledge, the divine attributes, and the nature of sin. He called this reflective activity "meditation" and also, in a famous phrase, "faith in search of understanding" (*fides quaerens intellectum*). His search for understanding is of interest to philosophers for three reasons. First, he often addresses arguments to those who do not share his dogmatic commitments—that is, he offers proofs based only on natural reason. He begins the *Monologion*, for example, with the claim that a person who does not (initially) believe that there is a God with the traditional divine attributes "can at least persuade himself of most of these things by reason alone if he has even moderate ability." Likewise, the "Ontological Argument" of the *Proslogion*, and indeed the treatise as a whole, is addressed to the Biblical Fool, who denies the existence of God. This approach, later known as "natural theology," may be given in support of but does not depend upon particular points of doctrine.

Second, even when Anselm assumes certain dogmatic theses, his analysis is often directed to specifically philosophical issues in the case at hand, and thereby has broader implications. While discussing Lucifer's sin and subsequent fall in his *De casu Diaboli*, for instance, Anselm formulates a series of general theses about responsibility and motivation that hold not only of Lucifer's primal sin (or Adam's original sin), but which apply to ordinary cases of choice. Elsewhere he offers a defense of metaphysical realism (*De incarnatione Verbi*), a reconciliation of foreknowledge with the freedom of the will (*De concordia*), an account of sentential truth-conditions (*De veritate*), and so on.

Third, even when pursuing his doctrinal agenda, Anselm is always a philosopher's philosopher: Distinctions are drawn and defended, theories proposed, examples given to support theses, and tightly constructed arguments are the means by which he meditates on Christian themes. He uses the selfsame method when no

doctrinal commitment is at stake, as in the semantic analysis of the *De grammatico*, the account of power and ability in his fragmentary notes, or the analysis of freedom of choice in *De libertate arbitrii*. For Anselm, understanding—the very understanding for which faith is searching—is a philosophical enterprise, and his treatment of even the knottiest doctrinal difficulties is clearly philosophical in character. Intellectual integrity, he held, demands it. (He further held that although a philosophical approach to matters of faith is necessary, it is not sufficient; hence, in addition to systematic treatises, Anselm also composed prayers and devotional works.)

METAPHYSICS

Following Augustine, Anselm is, broadly speaking, a platonist in metaphysics. A thing has a feature in virtue of its relation to something paradigmatically exhibiting that feature. Anselm begins the *Monologion*, for example, by noting the diversity of good things in the world, and argues that we should hold that "there is some one thing through which all goods whatsoever are good" and that that one thing "is itself a great good … and indeed supremely good" (chap. 1). He reasons that we can judge that some things are better or worse than others only if there is something, namely goodness, which is the same in each, though in different degrees—a claim sometimes dubbed "the Platonic Principle" for Plato's use of it in the case of equal sticks and stones in his *Phaedo*. To establish the uniqueness of this one thing, Anselm applies the Platonic Principle again and rules out an infinite regress. Furthermore, since the goodness of good things is derivative, and things might be good in any degree imaginable, it follows that the one thing through which all good things are good must be supremely good; it can be neither equaled nor excelled by the goodness of any good thing that is good through it. Note that the Supreme Good does not strictly speaking "have" goodness but rather *is* goodness itself, a quasi-substantial entity whose nature is goodness.

Much of Anselm's metaphysics is a sustained study of such relations of dependence and independence: things may be the way they are "through themselves" (*per se*) or "through another" (*per aliud*), Anselm holds, and roughly the same reasoning can be applied to features other than goodness. The later medieval tradition called such features "pure perfections," and their defining characteristic is that it is unqualifiedly better to have them than not. Just as the presence of goodness in things leads to the conclusion that there is some one thing that is paradigmatically good, through which all good things have their goodness, Anselm argues that so too the bare fact of their existence leads to the conclusion that there is some one thing through which everything else exists. Moreover, this one thing "paradigmatically" exists, namely, it exists through itself and of necessity: it is existence itself, something whose nature is existence (chaps. 3–4).

Anselm drops from the Platonic Principle the requirement that things having a certain feature exhibit it in varying degrees; rather, the possession of the same feature by itself licenses the inference that there is something each thing has, something exemplifying the feature itself. Likewise, the key move in his argument that there is only one such thing that exists through itself, rather than a plurality of independent things each equally existing through itself, is to apply the Platonic Principle to the feature of *self-existence* itself; this entails that there is a unique self-existent nature. Furthermore, since it is better to exist through oneself than through another (independence is better than dependence), the Supreme Good must exist through itself, and hence is identical with the self-existent nature, the source of the existence and goodness of all else there is. Anselm concludes that "there is accordingly a certain nature (or substance or essence) that through itself is good and great, and through itself is what it is, and through which anything that exists is genuinely either good or great or anything at all" (chap. 4). In short order Anselm shows that this being is appropriately called "God," and the remainder of the *Monologion* is devoted to establishing that God has the full range of divine attributes: simplicity, unchangeableness, eternality, triune nature of persons, and the like.

The existence of God is therefore the most fundamental metaphysical truth. Anselm tells us that he sought to replace the chain of arguments outlined above with "a single argument that needed nothing but itself alone to prove its conclusion, and would be strong enough to establish that God truly exists and is the Supreme Good, depending on nothing else, but on whom all other things depend for their existence and well-being." In doing so, he devised one of the most-discussed arguments in the history of philosophy, presented in *Proslogion* 2 as follows:

> Therefore, Lord, You Who give understanding to faith, give me understanding to the extent You know to be appropriate: that You are as we believe, and You are that which we believe. And, indeed, we believe You to be something than which nothing greater can be thought. Or is there is not some such nature, then, since "The Fool hath said in his heart: There is no God" [Psalms 13:1]? But certainly that same Fool,

> when he hears this very thing I say, 'something than which nothing greater can be thought', understands what he hears; and what he understands is in his understanding, even if he were not to understand that to be. It is one matter that a thing is in the understanding, another to understand a thing to be. For when the painter thinks beforehand what is going to be done, he has it in the understanding but does not yet understand to be what he does not yet make. Yet once he has painted, he both has it in the understanding and also understands to be what he now makes. Therefore, even the Fool is convinced that there is in the understanding even something than which nothing greater can be thought, since when he hears this he understands, and whatever is understood is in the understanding. And certainly that than which a greater cannot be thought cannot be in the understanding alone. If indeed it is even in the understanding only, it can be thought to be in reality, which is greater. Thus if that than which a greater cannot be thought is in the understanding alone, the very thing than which a greater cannot be thought is that than which a greater can be thought. But certainly this cannot be. Therefore, without a doubt something than which a greater is not able to be thought exists (*exsistit*), both in the understanding and in reality.

The logical analysis, validity, and soundness of this argument have been a matter of debate since Anselm came up with it. Yet its general drift is clear. God, Anselm tells us, is something than which nothing greater can be thought. (Note that he does not present this formula as a definition or part of the meaning of "God" but rather only as a claim that is true of God; the indirect negative formulation is important since we cannot adequately think of or conceive God as such.) So understood, the denial of God's existence leads to a contradiction, as follows. That than which a greater cannot be thought cannot itself be thought not to exist, since if it were, we could think of something greater than it, namely that than which nothing greater can be thought existing in reality. But it is logically impossible to think of something greater than that than which nothing greater can be thought. Thus the denial of God's existence must be rejected, and so God's existence affirmed. Hence Anselm's argument as a whole is *ad hominem*, directed against someone who accepts the claim that God is something than which nothing greater can be thought; once accepted, Anselm offers a *reductio ad absurdum* of the denial of God's existence.

Anselm's argument (as it was known in the Middle Ages) attracted attention from the very first. When the *Proslogion* was initially circulated, Gaunilon, a monk of Marmoutiers, wrote a brief in defense of the Fool; Anselm wrote a gracious reply and directed that thereafter the treatise should be copied with their exchange.

In the *Monologion* and *Proslogion*, Anselm says that he is trying to establish the existence of a "nature" (or equally an essence or a substance). The divine nature is identical with the very qualities of which it is the paradigm, and furthermore is also a concrete particular: God is an individual, albeit a three-in-one individual. In addition to such an extraordinary nature, there are also common natures, such as human nature, which is present in each human being as his or her individual nature. Anselm holds that such common natures "become singular" when combined with a collection of distinctive properties (*proprietates*) that distinguish an individual from all others (*De incarnatione Verbi* 11). In the same work he inveighs against the extreme nominalism of Roscelin of Compiègne that anyone taking universals to be no more than vocal utterances deserves no hearing on theological matters; Roscelin cannot understand how a plurality of humans are one human in species, and cannot understand how anything is a human being if not an individual (chap. 1).

While the extent of Anselm's metaphysical realism is a matter of debate, remarks such as these make it clear that he countenanced some form of realism about universals. Whereas some form of platonic exemplarism works for features that are identical with the divine essence, a more traditional realism applies to nondivine natures in the mundane world of creatures. From Boethius, Anselm adopts the standard metaphysical framework of substances and accidents, sorted into the ten Aristotelian categories. In the case of substances, Anselm holds that common names designate common natures, while proper names designate individuals metaphysically composed of a nature combined with distinctive properties with further accidental qualities. In addition, there are nonsubstantial qualities such as whiteness, instances of which may be found in individuals. Anselm speaks occasionally of form and of matter, but does not have a developed hylomorphic theory.

ETHICS

Anselm's positive ethical theory is grounded on his theory of the will and free choice, one of his most striking and

original contributions. The traditional account of free will holds that an agent is free when there are genuine alternatives open to her, so that she can do one or another of them as she pleases. This traditional account is sometimes called "bilateral" since the agent must have at least two possible courses of action in order to act freely. In his *De libertate arbitrii*, by contrast, Anselm defends a unilateral normative conception of freedom, according to which an agent is free when two conditions are jointly satisfied: (a) she has the ability to perform a given action; and (b) that action is the one she ought to perform, that is, it is objectively the right action and hence the one she ought to want to perform—roughly, that an agent is free when she can act as she ought, regardless of alternatives. (Anselm, like all medieval philosophers, holds that what an agent ought to do is an objective matter.)

Note that Anselm is careful to say that an agent is free when she *can* act as she ought, not that she *does* so act; we commit wrongdoing freely when the right course of action is open to us but we fail to pursue it. The crucial issue, of course, is when an agent has the ability to perform a given action. Anselm devotes most of his fragmentary notes on ability and power to investigating this issue. His analysis tracks connections among ascriptions of ability, responsibility, and the cause of an action, much in the spirit of contemporary philosophical reflections on tort law. Very roughly, Anselm thinks there are a variety of freedom-canceling conditions; some of these, such as compulsion, are extremely sensitive to the kind of ability at stake.

One case in particular attracts Anselm's attention in his *De libertate arbitrii*. Some abilities can be exercised by an agent more or less at will: lifting a book, thinking about Rome, deciding not to eat pork, playing the piano. Other abilities depend on external factors, which may include the actions and abilities of other agents. It takes two to tango, a multitude of musicians to play a symphony, other runners to have a race. These are all necessarily dependent abilities: They require other agents acting appropriately for their exercise. But consider a case in which an ability that could be exercised at will can no longer be so exercised, though the agent retains the ability. A ballerina tied to a chair cannot dance but still has the ability to do so. More exactly, Anselm holds, she does not have the opportunity to exercise the ability, though she retains the ability; were the constraint removed, she could exercise her ability at will. Anselm argues that the ballerina's ability to dance is what matters to her free choice, according to (a), not whether she currently has the opportunity to exercise her ability.

Now suppose that the ballerina, no longer tied to a chair, has through excessive dancing injured her legs so badly that she can dance only if a doctor operates on her legs. Here too, Anselm maintains, she has not lost the ability to dance but only the opportunity to exercise her ability, and can regain the opportunity only if a doctor helps her to do so. This is the situation in which Anselm finds the human race. Through the (wrongful) exercise of our free choice in original sin, we have lost the opportunity to freely do what is right, and can only recover it through the actions of another (namely through God's grace). We can legitimately be faulted for not doing what is right even now, despite the fact that we cannot do what is right at will, by our unaided efforts; we have the ability, and we lost the opportunity to exercise it through its improper use, but these facts do not stand in the way of our being free to act rightly; hence our culpability for failing to do so. Whether we agree with Anselm or not, his analysis is subtle and provocative, and represents a new level of sophistication in the analysis of free choice.

Following Augustine, Anselm argues that we abandon rectitude of will only by our own choice. Many things can happen against one's will, but it is impossible to will against one's will, since that would require both willing something and willing not to will it—but that can be done by simply not willing it in the first place. Not even God can take away our rectitude of will, Anselm maintains, since rectitude of will is doing what God wants; if God wanted to deprive our wills of rectitude, He would want us to not do what he wants, and whether we try to obey or to disobey, we wind up doing as He wants. Thus abandoning rectitude must be through our own choice, since it cannot happen against our will or by external (even divine) compulsion. The responsibility for wrongdoing rests squarely on our shoulders.

Anselm returns to these topics in his *De casu Diaboli*, perhaps returning to the traditional bilateral conception of freedom in the process. In Chapter 12 he puts forward a famous thought-experiment in which God creates an angel with free will, but without any motive for action whatsoever—a free being with no ends at all. Anselm argues that such a being would never act, since any action is motivated by pursuit of an end, and by hypothesis the angel has no ends. (Nor is an angel ever prompted by biological needs, and this is the point of using an angel rather than a human being in the example.) From this case Anselm and later philosophers drew the moral that at least some ultimate end has to be given to agents in order for there to be action at all, and hence the possibility of

moral action. An agent must therefore have at least one ultimate end, an end she does not choose.

Yet one end is not enough for moral agency. Anselm argues that there must be two ultimate and incommensurable ends to make sense of moral choices, and specifically of moral dilemmas. He reasons as follows. If an agent had only a single end, she would always act in pursuit of that end, unless deceived or misled through ignorance. There would be no moral conflict; her motives and reasons for action would be transparently in the service of her single ultimate end. This is quite similar to the life of nonrational animals. A dog pursues only its apparent good, as defined by its nature (which establishes its ultimate end). Dogs naturally aim at their own "perfection," as Anselm puts it. But human beings are more complicated. We face choices in which each alternative serves a distinct end, the ends being ultimate and incommensurable. Anselm holds that this fact explains moral agency and the possibility of moral wrongdoing, for rational agents have two distinct ultimate ends: they seek their own happiness (through advantage or benefit) on the one hand, and they seek justice (rectitude of will) on the other hand.

This is the core of Anselm's so-called "two-will theory" of motivation. Moral conflicts and dilemmas arise when we are faced with the choice between happiness and justice, between individual self-interest and impersonal fairness. Each end is a genuine good to the individual agent, and the conflict between them is real. Morality demands that we favor justice over happiness in such conflicts; wrongdoing is explained as the choice of happiness over justice. A thief prefers his own advantage to following the laws. While we might not side with the thief, his choice is not inexplicable; indeed, we may even sympathize with him while deploring his actions. The possibility of an irreducible clash between ultimate ends that we cannot forgo gives us the ability to explain moral agency. To say that justice and happiness can conflict is of course not to say that they do; if we are lucky, we might avoid moral dilemmas. Nevertheless, our actions are free because of the pull between these ends, even if we consistently take one side or the other.

Human fulfillment for Anselm thus turns out to be surprisingly paradoxical. We do not deserve to be happy unless we are prepared on principle to forgo happiness for justice. Indeed, only by pursuing justice for its own sake can we attain the self-interested happiness we have scorned. The price of moral agency is that happiness is the reward for those who do not pursue it.

PHILOSOPHY OF LANGUAGE

Anselm adopts Augustine's view of language as a system of signs. This general category covers linguistic items, such as utterances, inscriptions, gestures, and at least some acts of thought; it also covers nonlinguistic items, such as icons, statues, smoke (a sign of fire), and even human actions, which Anselm says are signs that the agent thinks the action should be done. Roughly, a sign signifies something by bringing it to mind; this single semantic relation, founded on psychology, is the foundation of Anselm's semantics.

As noted above, common names—at least natural-kind terms—signify common natures, and proper names signify the common nature in combination with distinctive properties. Nondenoting terms are problematic; "nothing" seems to be significant only by signifying nothing, a paradox that perplexes Anselm in several treatises. Troublesome as they are, Anselm directs his most sustained inquiry into semantics not at empty names but at "denominative" terms, roughly what we call adjectives.

The difficulty he addresses in his *De grammatico* can be stated simply: "white" cannot signify whiteness ("whiteness" does that); nor can it signify what is white ("snow" does that); what then does it signify? Anselm's answer depends on several distinctions, the most important of which is between direct and indirect signification (*per se* and *per aliud* signification). A term signifies directly if it brings the proper and customary signification to mind; it signifies other things indirectly, perhaps things linked somehow to what the term directly signifies. As a first approximation, then, Anselm holds that 'whiteness' directly signifies whiteness, whereas 'white' directly signifies whiteness and indirectly signifies things that have whiteness (and is used to pick out the latter).

Verbs, for Anselm, signify actions or "doings" of some sort, broadly speaking, including even passive processes; that is their distinguishing feature. Names and subjects, respectively, signify subjects and their doings; when combined in a sentence, the truth of the sentence reflects the underlying metaphysical dependence of doings on doers, of actions on subjects. Now just as Anselm's theory of meaning applies to more than words, so too his theory of truth applies to more than statements. In the *De veritate*, Anselm puts forward an account that recognizes a wide variety of things to be capable of truth—statements, thoughts, volitions, actions, the senses, even the very being of things. Truth, for Anselm, is a normative notion: Something is true when it is as it ought to be. Thus truth is in the end a matter of correctness (*rectitudo*), the correctness appropriate

in each instance (*De veritate* 11). For statements there are actually two forms of correctness: A given statement ought to signify what it was designed to express, and, if assertoric, it ought to signify the world the way it is. The first is a matter of the propositional content of an utterance, the second whether that propositional content is asserted (or denied). The statement "snow is white" does what it should do when it succeeds in signifying that snow is white; it also does what it should do when it succeeds in signifying that snow is white in the circumstances that snow really is white. The latter is the closest to our contemporary notion of truth for statements, but Anselm insists that the former is a kind of truth too (he calls it the "truth of signification"), and indeed can hold even if the world changes such that snow is no longer white.

See also Aristotle; Augustine, St.; Ontological Argument for the Existence of God; Plato; Roscelin.

Bibliography

Anselm's works have been critically edited by F. S. Schmitt, *S. Anselmi Cantuariensis Archiepiscopi opera omnia*, vols. 1–6 (Edinburgh: T. Nelson, 1946–1961), supplemented by fragments and miscellaneous materials in R. W. Southern and F. S. Schmitt, *Memorials of St. Anselm* (Auctores Britannici Medii Aevi 1), published for The British Academy by Oxford University Press, 1969. Complete translations of Anselm's major works may be found in Jasper Hopkins and Herbert Richardson, *Anselm of Canterbury*, vols. 1–3 (New York: Edwin Mellen Press, 1976), as well as in Brian Davies and Gillian R. Evans, eds., *The Major Works: Anselm of Canterbury* (Oxford and New York: Oxford University Press, 1998). The most thorough modern biography of Anselm is by R. W. Southern, *Saint Anselm: A Portrait in a Landscape* (Cambridge and New York: Cambridge University Press, 1990).

Overviews of Anselm's thought are given in Jasper Hopkins, *A Companion to the Study of St. Anselm* (Minneapolis: University of Minnesota Press, 1972); Gillian R. Evans, *Anselm* (London: Chapman, 1989); Brian Davies and Brian Leftow, eds., *The Cambridge Companion to Anselm* (Cambridge and New York: Cambridge University Press, 2004). On the Ontological Argument, see Alvin Plantinga, ed., *The Ontological Argument* (Garden City, NY: Anchor Books, 1965); John Hick and Arthur McGill, eds., *The Many-Faced Argument* (New York: Macmillan, 1967); David Lewis, "Anselm and Actuality," in *Nous* 4 (1970): 175–188; Jonathan Barnes, *The Ontological Argument* (London: Macmillan, 1972); Paul Oppenheimer and Edward Zalta, "On the Logic of the Ontological Argument," in *Philosophical Perspectives*, vol. 5, edited by James Tomberlin (Atascadero, CA: Ridgeview, 1991). Issues having to do with free choice are discussed in John O'Neill, "Anselm on Conflicting Oughts," in *Heythrop Journal* 35 (1994): 312–314; Calvin Normore, "Picking and Choosing: Anselm and Ockham on Choice" in *Vivarium* 36 (1998): 23–39; Thomas Williams and Sandra Visser, "Anselm's Account of Freedom" in *Canadian Journal of Philosophy* 31 (2001): 221–244. Anselm's theory of action is analyzed in Eileen Serene, "Anselmian Agency in the Lambeth Fragments: A Medieval Perspective on the Theory of Action" in *Anselm Studies* 1 (1983): 143–156; Douglas Walton, "Anselm and the Logical Syntax of Agency" in *Franciscan Studies* 14 (1976): 298–312. On Anselm's philosophy of language see Desmond P. Henry, *The Logic of Saint Anselm* (Oxford: Clarendon Press, 1967); and Marilyn Adams, "Saint Anselm's Theory of Truth" in *Documenti e studi sulla tradizione filosofica medievale* 1 (1990): 353–372.

Peter King (2005)

ANTHROPIC PRINCIPLE, THE

The term *Anthropic Principle* (AP) was introduced by the physicist Brandon Carter, who stated that "what we can expect to observe must be restricted by the conditions necessary for our presence as observers" (Carter 1974, p. 292). The central idea of AP could be put as follows: We can observe only those states of affairs that are compatible with the existence of observers.

The term has subsequently been applied to all manner of claims, variously obscure and bizarre. This entry restricts its attention to the central and philosophically interesting idea. Carter distinguished what he called the weak version of the principle, according to which "our location in the universe is necessarily privileged to the extent of being compatible with our existence as observers" (Carter 1974, p. 293), and the strong version, which states that "the universe (and hence the fundamental parameters on which it depends) must be such as to admit the creation of observers within it at some stage" (p. 295). The distinction was meant merely to apply the simple insight of AP on the one hand to local conditions at places and times within the universe, and on the other to features of the universe as a whole. The unfortunate wording of Carter's strong principle has led many to misunderstand it as attributing necessity to the universe's fundamental parameters. Whatever appeal this idea has seems to derive from a simple scope confusion. AP tells that necessarily, if humans observe a universe, then it has the parameters that allow for the development of observers. It does not follow that if humans observe a universe then the conditions required for observers take hold necessarily.

APPLICATIONS WITHIN THE UNIVERSE

AP is obviously true, and may appear too obvious to be of any interest. It is said to play a crucial role in explanation and theorizing about the universe. But how could a seemingly trivial, necessary truth enter into scientific explanations and inferences at all?

One can begin to answer this question by recalling that failure to consider the limits on what can be observed often leads to errors in scientific reasoning. This is well illustrated by one of John Leslie's cases of selection bias in *Universes* (1989). If a person finds all of the fish he or she has caught to be more than five inches long, this person may be tempted to inductively infer that all fish in the lake are longer than five inches. But the strength of this inference is undermined by noting that the net used cannot hold smaller fish. One can understand this epistemic situation in terms of competing explanations. The hypothesis that all fish in the lake are more than five inches long may, in principle, explain the failure to observe any shorter fish: One has not seen short fish because there are none around to see (perhaps chemical waste has rendered the adult fish population infertile). But such an explanation becomes redundant when it is noted that the method of observation prevents one from seeing smaller fish, whether there are any in the lake or not. If this person had been fishing with a regular reel and bait, it would be remarkable that he or she would have failed to catch small fish, and the hypothesis that all the fish in the lake are longer should be taken more seriously. The inference to all fish being more than five inches long is undermined by eliminating its use as an explanatory hypothesis.

It pays to be clear on what is explained here and what is not. The observational limitation—using a net with large holes—does nothing to explain, for any particular fish, why that fish is longer than five inches. What is explained is the failure to observe anything but long fish.

In a similar way, AP can serve as a check on overly zealous use of what is known as the Copernican Principle in cosmology. Copernicus famously taught that the Earth is not central to the solar system, let alone the universe. Taking this lesson to heart, cosmologists have been wary of theories that attribute special characteristics to the Earth's spatio-temporal position. The Copernican Principle instructs, roughly, to proceed on the assumption that the conditions that take hold within one's observable neighborhood are more or less the same throughout space-time. As a guard against gratuitous biases about the human place in the universe, the Copernican Principle is appropriate. But it would be equally arbitrary to rule out the possibility that in the vastness of space-time, there are isolated pockets with strikingly unique features. And it is not out of the question that humans may happen to occupy one of these special regions. Indeed, if these rare conditions are necessary for the development of intelligent life, this is just where humans should expect to find themselves. It would be a mistake akin to that in the fishing story to extrapolate too eagerly from observations of local conditions to the wider universe if these locally observed features are a necessary condition of one's being here to observe anything. For in this case one can adequately explain the failure to observe any other features, even if most of the universe is different. (One of the earliest influential appeals to AP by the physicist R. H. Dicke, in "Dirac's Cosmology and Mach's Principle" [1961], uses this kind of strategy.)

One must be careful to distinguish this lesson from some more grandiose claims made on behalf of AP. Some incautious statements by physicists have been taken to suggest that human's existence and ability to observe the universe helps to explain why those observed features took hold. Clearly this explanation goes in the wrong direction. From human existence together with certain laws of nature, it may be possible to deduce that certain conditions took hold; this is not, however, sufficient for explanation. It is the required conditions that (partly) explain why humans are here, and not the other way round. Human observational limits no more explain why any observed conditions took hold than the use of a fishing net with large holes explains the length of any fish. In each case it is only one's failure to make contrary observations that is explained.

APPLICATIONS TO THE UNIVERSE'S FUNDAMENTAL PARAMETERS

According to contemporary cosmology, if the values of various fundamental parameters of the universe—such as force strengths and particle masses—differed ever so slightly from their actual values, life could not possibly have developed anywhere in the universe. And it appears that these parameters could easily have been different. It is as if the universe were the product of a machine with dozens of dials that determined its features. The vast majority of dial combinations result in a universe that collapses within seconds, or that contains nothing but hydrogen, or nothing but black holes. Only the most delicate adjustment of the dials will produce a stable universe, capable of supporting life at some time and place. Without the aid of deliberate adjustment, the odds of the

big bang producing a life-permitting universe appear extremely low.

In the light of these data, that the universe meets the conditions for life has struck many scientists and philosophers as a striking fact that requires explanation (whether or not this attitude is appropriate may be questioned, but it is only on this initial assumption that uses of AP arise). Some have taken these facts as the basis of a new version of the argument from design. The remarkable coincidence of physical parameters required for life may be explained by the actions of a rational agent. Others have suggested that the solution may lie in a more fundamental theory, with laws constraining the range of values that crucial parameters can take. The application of AP is supposed to undermine the need for such hypotheses. The simplest anthropic-style response takes the form of a popular glib reply: "If the physical parameters hadn't been just so, then we wouldn't be here puzzling about the matter!" The inadequacy of this response is well illustrated by the following analogy from Leslie (1989). Standing before a firing squad, a dozen guns are fired your way, but not a single bullet hits you. Clearly you have grounds to be astonished and wonder why you have been so lucky. Did they all deliberately miss? Did they fill their guns with blanks? It is possible that their missing you is just a fluke, but this seems incredible. It becomes no more credible when one considers that if the gunmen had not all missed, you would not be here to puzzle about it. Given that people do observe a universe, it is no surprise that they see one that meets the conditions for observers to exist. But they may well still wonder how they, or anyone, are here to see anything at all.

MULTIPLE UNIVERSES

More serious uses of AP couple it with the suggestion that this universe is just one of an enormous variety of actually existing universes. (Here "universe" does not refer to the totality of what there is, but rather to a large, more or less isolated aggregate of matter in space-time.) Of course this strategy is viable only to the extent that reason exists to suppose that there are a great number of universes. This is highly controversial. One of the proposed universe-generating models is the inflationary theory in cosmology. The multiple-universe hypothesis is distinct from the many-worlds interpretation of quantum mechanics, but some have appealed to the latter as a way in which the required variety of universes might be generated (see Leslie 1989).

How could the existence of other universes help solve a puzzle about this universe? For any improbable outcome of an event, if you repeat the type of event enough times you can expect to get an outcome of that type eventually. To take the popular example, if a monkey types for long enough, or a large enough army of monkeys is put to work, it is all but certain that somewhere, at some time, a monkey will type a sonnet. Similarly, whereas it may be extremely unlikely that any particular universe meets the conditions for life, if there are a large enough number of them, it is to be expected that at least one of them will by chance be life-permitting. The vast majority of universes will be rather bland, containing no stars or planets, let alone life. There should be no room to wonder why humans have been lucky enough to see only one of the nice universes. They may note by AP that they could not possibly have found themselves in any other kind of universe, as those universes fail to meet the conditions for human existence.

The same explanatory strategy has been employed in areas of science as diverse as statistical mechanics and evolutionary biology. Ludwig Boltzman (1895) suggested a similar idea to account for the extremely low level of entropy (i.e., roughly, the high degree of order) in the observable neighborhood. Boltzman's speculation was that the universe is extremely large in space and time, with disorder on the large scale, but large, finite regions of order within. One can picture this view as like an infinite number of coins tossed on an infinite expanse. The big picture will almost certainly be a random, disordered mess. But with maximum probability there will also be enormous finite stretches of nothing but heads, and vast regions of beautiful and orderly patterns. Boltzman noted that it is only in regions of low entropy that living organisms such as humans can be found. So on this hypothesis, people should not be surprised to find that theirs is a low-entropy environment. Similar principles are applied in Darwinian explanations of the evolution of organisms. The tree of life consists of an enormous variety of branches produced by random mutations. Most of these are hidden from human view. It is only those that have the remarkable ability to sustain themselves and reproduce that people are able to observe.

As before, care needs to be taken in stating what has been explained and what has not. The plenitude of universes does not explain why this particular universe humans inhabit is life-permitting. The answer to the question "Why is this universe suitable for life?" is not "Because there are many other universes." The existence of many universes does not raise the probability that any particular one such as this can support living creatures. At most, what is to be expected is some universe will do this.

What the hypothesis of many universes may do, however, is remove the urgency of explanation regarding the particular universe in which humans find themselves. That this universe is life-permitting seems remarkable only insofar as it seemed remarkable that there was life at all. But if there are many universes, then it is not surprising that somewhere in some universe life can develop. The more specific question of why it is this universe and not another one appears less urgent, such as the question of why Jones won the lottery, or why the golf ball landed on this blade of grass. An adequate answer may be along the lines of "That's just how it turned out."

See also Cosmology; Many Worlds/Many Minds Interpretation of Quantum Mechanics; Philosophy of Statistical Mechanics.

Bibliography

Barrow, John, and Frank Tipler. *The Anthropic Cosmological Principle*. Oxford: Clarendon Press, 1986.

Boltzman, Ludwig. "On Certain Questions of the Theory of Gases." *Nature* 51 (1895): 413–415.

Carter, Brandon. "Large Number Coincidences and the Anthropic Principle in Cosmology." In *Confrontation of Cosmological Theories with Observational Data*, edited by M. Longair, 291–298. Dordrecht, Netherlands: D. Reidel, 1974. Reprinted in *Modern Cosmology and Philosophy*, 2nd edition, edited by John Leslie. New York: Prometheus, 1998.

Dicke, R. H. "Dirac's Cosmology and Mach's Principle." *Nature* 192 (1961): 440–441. Reprinted in *Modern Cosmology and Philosophy.* 2nd edition, edited by John Leslie, 127–130. New York: Prometheus, 1998.

Earman, John. "The SAP Also Rises: A Critical Examination of the Anthropic Principle." *American Philosophical Quarterly* 24 (4) (1987): 307–317.

Leslie, John, ed. *Modern Cosmology and Philosophy.* 2nd edition. New York: Prometheus, 1998.

Leslie, John. *Universes*. London: Routledge, 1989.

Roger White (2005)

ANTHROPOLOGY

See *Philosophical Anthropology*

ANTIOCHUS OF ASCALON

(130/120?–68/7 BCE)

Antiochus joined the Academy, the school founded by Plato, late in the second century BCE, when Philo of Larissa was its head. Philo was (at this time) a moderate Academic skeptic who had been convinced by the Academy's anti-Stoic arguments that nothing can be known for certain, but he did not embrace the other doctrine for which Academic skeptics argued—suspension of judgment. According to Philo, although certain knowledge is unobtainable, it is possible to identify highly probable impressions, and there is no reason not to accept them, provided that one realizes that one might be wrong. Prominent among them is the impression that nothing can be known.

After defending this view for many years, Antiochus became a dogmatist by accepting that knowledge is possible. His epistemological position was now essentially that of the Stoa. He responded to accusations that he had left the Academy for the Stoa by claiming that Zeno of Citium (335–263 BCE), the founder of Stoicism, had introduced a new vocabulary but was otherwise in essential agreement with the schools of Plato and Aristotle. Far from abandoning the Academy, Antiochus maintained, he had returned it to its true self. For this reason, he and his followers styled themselves the Old Academy. It is unclear what institutional status this group enjoyed or whether Antiochus ever officially succeeded Philo.

Antiochus regarded the criterion of truth and the goal of life as the most important concerns of philosophy, and his ethical theory is the other area about which we are well informed. In opposition to the Stoics, who maintained that virtue is the sole good and therefore sufficient for happiness, Antiochus held that there were also bodily and external goods. He rightly took this to be the view of Aristotle and the Old Academy, but the form in which he presented his theory owes a good deal to the Hellenistic schools. Thus Antiochus relied heavily on the so-called cradle argument, which takes the uncorrupted behavior of infants as its starting point. Antiochus combined evidence from this source about the objects of our first natural concern with the general principle that what accords with a creature's natural impulses is its good, to derive his account of the goal.

He was in broad agreement with the Stoics that our constitution and things that preserve and develop it are the first objects of our natural concern, and not pleasure as Epicurus supposed. But the Stoics take it that this natural concern is replaced by a unique attachment to virtue. Antiochus held that, as the perfection of reason, which is the most important part of our constitution, virtue is the chief good. But he also regarded the other objects of natural concern as goods, albeit lesser goods, and therefore a part of the goal.

Antiochus wanted to claim that, even so, virtue is sufficient for happiness. To this end, he distinguished

between the happy life (*vita beata*), for which virtue was enough, and the entirely happy life (*vita beatissima*), which requires other goods as well.

None of Antiochus's books have survived, but he is known to have written a work about epistemology, the *Canonica*, and another epistemological work, the *Sosus* against Philo of Larissa's late views. A book in which he stressed the close relation between the Peripatos, Aristotle's school, and the Stoa is attested, and Cicero tells us that he wrote in many places about his views concerning happiness and virtue.

See also Ancient Skepticism; Cicero, Marcus Tullius; Philo of Larissa; Stoicism; Zeno of Citium.

Bibliography

Barnes, Jonathan. "Antiochus of Ascalon." In *Philosophia Togata: Essays on Philosophy*, edited by Jonathan Barnes and Miriam Griffin. Oxford, U.K.: Oxford University Press, 1989.

Glucker, John. *Antiochus and the Late Academy*. Hypomnemata 56. Göttingen: Vandenhoeck und Ruprecht, 1978.

Mette, H. J. "Philon von Larissa und Antiochos von Askalon." *Lustrum* 28–29 (1986–1987): 9–63.

Striker, Gisela. "Academics Fighting Academics." In *Assent and Argument: Studies in Cicero's Academic Books*, edited by Brad Inwood and Jaap Mansfeld. Leiden, Netherlands: Brill, 1997.

James Allen (2005)

ANTIPHON

(c. 480–411 BCE)

Antiphon was an Athenian sophist, author of *Truth*, *Concord*, and—if identical with the same person as Antiphon of Rhamnus—three *Tetralogies* and many court speeches. The identity of the sophist and the speechwriter remains uncertain but is increasingly accepted (see Gagarin 2002; for contra, Pendrick 2002). If the two are the same, Antiphon was an aristocratic Athenian, admired by Thucydides (*History of the Peloponnesian War* 8.68), who wrote sophistic works, taught, gave legal and political advice, and wrote speeches for litigants in court. He was a leader of an oligarchic coup in 411 BCE and was tried and executed after the coup quickly failed.

Antiphon's *Tetralogies*, probably his earliest works (450–430 BCE), were intended for intellectual stimulation and pleasure and perhaps for public performance. Each group of four speeches (two on each side) treats a hypothetical case of homicide. In the *First Tetralogy*, the identity of the killer is uncertain, and arguments are based on the likelihood (*eikos*) that the defendant is the killer. The *Second* disputes whether a young man who threw a javelin that killed a boy is responsible for the death. The *Third* questions who is responsible for a man's death during a drunken fight. None of the *Tetralogies* has a conclusion or verdict. Their aim is to explore issues and forms of argument (likelihood vs. truth, fault and responsibility, cause and effect) with subtlety and cleverness. They also raise questions about the relationship of *logos* (speech, argument) to reality, and the relationship between opposed arguments when each claims, with some justification, to speak the truth.

Perhaps in the 420s BCE, Antiphon composed the sophistic works *Truth* and *Concord*—only fragments of which now remain—and the even more fragmentary *Politicus* and *Dream-Interpretations*. *Truth* explored a wide range of issues, including mathematics (squaring the circle), meteorology, and natural philosophy. The largest surviving fragments show Antiphon exploring the relationship between *nomos* (law, convention) and *physis* (nature), particularly with respect to law and justice. He may be saying that law is purely a matter of convention, and that a person may violate the law as long as no one else will know of it.

See also Sophists.

Bibliography

Antiphon. "Antiphon." Translated by J. S. Morrison. In *The Older Sophists*, edited by Rosamond K. Sprague, 106–240. Columbia: University of South Carolina Press, 1972. English translation of all surviving works, including fragments, and of the ancient testimony for Antiphon.

Gagarin, Michael. *Antiphon the Athenian: Oratory, Law and Justice in the Age of the Sophists*. Austin: University of Texas Press, 2002. Assumes a single Antiphon and analyzes all the works, including the three surviving court speeches.

Pendrick, Gerard J. *Antiphon the Sophist: The Fragments*. Cambridge, U.K.: Cambridge University Press, 2002. Introduction, Greek text with translation, and full scholarly commentary on *Truth*, *Concord*, and the other sophistic fragments.

Thucydides. *History of the Peloponnesian War*. Edited by Robert B. Strassler. New York: Free Press, 1996.

Michael Gagarin (2005)

ANTIREALISM

See *Realism*

ANTISTHENES

(ante 443 BCE–post 366 BCE)

Antisthenes, son of an Athenian father and Thracian mother, was a pupil of the rhetorician Gorgias and an intimate and admirer of Socrates. He taught professionally at Athens, maintaining his own interpretation of Socrates against other Socratics such as Plato and Aristippus. There is, however, only one reference in classical literature to Antistheneans (Aristotle, *Metaphysics* 1043b24); later antiquity saw him as a founder of Cynicism, a view that may have gained support through later historical systematization or from Stoics attempting to trace their philosophical pedigree to Socrates. Nevertheless, while the historical relationship between Antisthenes and Diogenes remains obscure, there were elements in Antisthenes' thought that heralded and may have given some impulse to Diogenes. His numerous works have not survived (a list of titles is found in Diogenes Laertius's *Lives,* 6.15–18); but he is characterized in Xenophon, and Diogenes preserves a doxographical and anecdotal account. Antisthenes had rhetorical and sophistic interests and was famed for his style and his myths as well as for his Socratic dialogues.

The influence of Socrates shaped Antisthenes' overriding interest in practical ethics. He held happiness to be dependent solely on moral virtue, which involved practical intelligence and so could be taught, partly from a study of the names of things and definitions. But the good man also required strength of mind and character; for by contrasting external goods with the inviolability of the "wealth of the soul," Antisthenes came to stress the importance of self-control by a hostility to luxury and sensual pleasure that went some way toward Cynic asceticism. Thus, the achievement of virtue necessitated a mental and physical effort to toil through opposing difficulties, suffering, and pain. Antisthenes glorified this struggle in the myths of Heracles; and for Cynics "toil" (*ponos*) became a technical good and Heracles a saint.

Antisthenes combined a moral interest in politics with a wariness of the dangers of participation, and attacked the rules of convention when they were in opposition to the laws of virtue. He denounced famous statesmen of previous generations and outlined his own ideal king, whose preeminence was due to his own moral self-mastery.

Most tantalizing is the brief glimpse Aristotle affords of Antisthenes' interest in the logic of predication and definition. He denied the possibility of contradiction (*Topics* 104b21), apparently because he believed (*Metaphysics* 1024b27 ff.) that each object could be spoken of only by its own peculiar verbal designation that said what it was; that is, words corresponded directly with reality, and since predication was confined to assigning names to things, or limited to formulas determining their real structure, any other predicative account must then refer to something different or to nothing at all, and contradiction did not arise. There was a similar difficulty with falsity. Elsewhere (*Metaphysics* 1043b23ff.) the Antistheneans are said to have denied the possibility of defining what a thing (like silver) was; one could only explain what sort of thing it was (for instance, "like tin"). Aristotle's context referred to simple substances that could not be analyzed but only named or described. Similar problems to these occur in Plato (as in *Sophist* 251A f.; *Theaetetus* 201C ff.; *Euthydemus* 283E ff.; *Cratylus* 429B ff.). It has been argued that in one or more of these passages Plato had Antisthenes in mind, but this is not at all certain. The problems were common to the period. Interesting similarities have been pointed out between Antisthenes' logic and the nominalism of Thomas Hobbes.

See also Aristotle; Cynics; Diogenes Laertius; Hobbes, Thomas; Plato; Socrates.

Bibliography

CLASSICAL

Antisthenis Fragmenta. Edited by A. W. Winkelmann. Zürich, 1842.

Diogenes Laertius. *Lives.* 6.1–19.

Xenophon. *Symposium* and *Memorabilia.*

MODERN

Dudley, D. R. *A History of Cynicism from Diogenes to the 6th Century.* London: Methuen, 1937.

Dümmler, F. *Antisthenica.* Halle, 1882.

Field, G. C. *Plato and His Contemporaries.* London: Methuen, 1930; 2nd ed., 1948.

Fritz, K. v. "Zur Antisthenischen Erkenntnistheorie und Logik." *Hermes* 62 (1927): 453–484.

Fritz, K. v. "Antisthenes und Sokrates in Xenophons Symposium." *Rheinisches Museum* n. F. 84 (1935): 19–45.

Gillespie, C. M. "The Logic of Antisthenes." *Archiv für Geschichte der Philosophie* 26 (1913): 479–500; 27 (1914): 17–38.

Höistad, R. *Cynic Hero and Cynic King.* Uppsala, Sweden: n.p., 1949.

Sayre, F. *The Greek Cynics.* Baltimore: Furst, 1948.

G. Kidd (1967)

APEIRON/PERAS

The Greek term *Apeiron,* meaning originally "boundless" rather than "infinite," was used by Anaximander for the ultimate source of his universe. He probably meant by it something spatially unbounded, but since out of it arose the primary opposite substances (such as the hot and the cold, the dry and the wet) it may have been regarded also as qualitatively indeterminate. Aristotle, summarizing the views of certain early Pythagoreans (*Metaphysics* A, 5), puts the pair *Peras* ("Limit") and *Apeiron* ("Unlimited") at the head of a list of ten opposites. *Peras* is equated with (numerical) oddness, unity, rest, goodness, and so on; *Apeiron* is equated with evenness, plurality, motion, badness. The two principles *Peras* and *Apeiron* constituted an ultimate dualism, being not merely attributes but also themselves the substance of the things of which they are predicated. From the Pythagoreans on, the opposition of *Peras* and *Apeiron* was a standard theme in Greek philosophy.

Parmenides (fr. 8, 42ff.) seems to have accepted Limit and rejected the Unlimited for his One Being. The later Pythagoreans removed unity from the list of identities with *Peras* and argued that unity was the product of the imposition of the *Peras* upon the *Apeiron,* or else it was the source of both of them. Plato in the *Philebus* regards *Peras* and *Apeiron* as contained in all things, and supposes that it is through limit that intelligibility and beauty are manifested in the realm of Becoming. Exactly how the Ideas fit into this scheme is controversial, but in the doctrine of ideal numbers which Aristotle attributes to him Plato seems finally to have identified a material principle with the *Apeiron* and a formal principle with the *Peras.* Both principles apply to the ideal as well as to the sensible world. This leads in due course to the doctrine in Proclus (*Elementa* 89–90) that true being is composed of *Peras* and *Apeiron,* and beyond being there is a first *Peras* and a first *Apeiron.* The Christian writer known as Dionysius the Areopagite identified this doubled First Principle with God.

INFINITY

The concept of infinity, for long wrongly regarded as contrary to the whole tenor of Greek classicism, was in fact a Greek discovery, and by the fifth century BCE the normal meaning of *Apeiron* was "infinite." Infinite spatial extension was implied in the doctrines of Anaximander, Anaximenes, and Xenophanes and was made explicit by the Pythagoreans (see Aristotle, *Physics* IV, 6). Denied by Parmenides, it was reasserted for the Eleatics by Melissus (frs. 3–4) and adopted by the Atomists. Plato, however (in the *Timaeus*), and Aristotle (*Physics* III) insisted upon a finite universe, and in this they were followed by the Stoics and most subsequent thinkers until the Renaissance. Aristotle had, however, admitted that infinity could occur in counting and he stated the concept clearly for the first time. He also accepted infinite divisibility (*Physics* VI), which had been "discovered" by Zeno and adopted wholeheartedly by Anaxagoras. It was rejected by the Atomists. Plato rejected it in the *Timaeus,* although he seems to have admitted it at the precosmic stage in *Parmenides* 158B–D, 164C–165C. Aristotle accepted infinite divisibility for movements, for magnitudes in space, and for time. The concept of a continuum so reached has been a basic concept in physical theory ever since. The mathematical concept of infinitesimal numbers associated with infinite divisibility and also with the doctrine of incommensurables remained important until the development of calculus in modern times.

See also Anaximander; Aristotle; Parmenides of Elea; Plato.

Bibliography

Mondolfo, R. *L'infinito nel pensiero dell'antichita,* 2nd ed. Florence, 1956.

Solmsen, Friedrich. *Aristotle's System of the Physical World.* Ithaca, NY: Cornell University Press, 1960. Ch. 8.

G. B. Kerferd (1967)

APEL, KARL-OTTO

(1922–)

Karl-Otto Apel (born in Düsseldorf) is an influential post-World War II German philosopher responsible for creatively introducing analytic linguistic philosophy to the German philosophical tradition. He fought in the German army on the eastern front and, in fact, began his university studies while a prisoner-of-war in France. He completed his doctoral dissertation on Martin Heidegger in Bonn in 1950, wrote his *Habilitation* ("The Idea of Language in the Tradition from Dante to Vico") in Mainz in 1960, and, after several years teaching at the Universities of Kiel and Saarbrücken, spent the rest of his academic career at the Goethe University in Frankfurt am Main (where Jürgen Habermas, whom he had known since his student years in Bonn, was his colleague). He is best known for his development of transcendental semiotics that, as a first philosophy distinct from both tradi-

tional metaphysics and a modern (e.g., Cartesian, Kantian, or Husserlian) philosophy of the subject, provides an ultimate foundation (*Letzbegründung*) for knowledge (1998, chapter 2).

His so-called transformation of philosophy represents an ambitious attempt to bring together in a systematic form analytic philosophy of language, American pragmatism (especially Charles Sanders Peirce), and philosophical hermeneutics (Heidegger and Hans-Georg Gadamer). According to Apel, in light of these innovative traditions, the transcendental philosophy of Immanuel Kant must be fundamentally reconceived. In particular, the conditions for intersubjectively valid knowledge cannot be explicated in terms of the structure of consciousness or the cognitive capacities of the individual knowing subject but only through a systematic investigation of language as the medium of symbolically mediated knowledge. The pragmatic turn, initiated by Peirce and Charles W. Morris (1901–1979) and continued in the early twenty-first century in speech act theory, further implies that an adequate explanation of how meaningful communication is possible cannot be achieved by a semantic theory alone. Rather, it must be supplemented by a pragmatic study of the relation between linguistic signs and the conditions of their use by speakers. Apel's strong thesis is that his transcendental semiotics yields a set of normative conditions and validity claims presupposed in any critical discussion or rational argumentation. Central among these is the presupposition that a participant in a genuine argument is at the same time a member of a counterfactual, ideal communication community that is in principle equally open to all speakers and that excludes all force except the force of the better argument. Any claim to intersubjectively valid knowledge (scientific or moral-practical) implicitly acknowledges this ideal communication community as a metainstitution of rational argumentation, to be its ultimate source of justification (1980).

Drawing on the Continental tradition, Apel argues that the most important contribution of philosophical hermeneutics, Gadamer's in particular, has been to show that interpretation is not another method of investigation in addition to the methods used within the hard sciences, but an unavoidable dimension of all understanding. Every empirical investigation of a domain of objects implies at the same time a relation to other subjects, to a community of interpreters. Thus, the attempt to study language from an exclusively objectivistic or naturalistic perspective involves an abstraction from the inquirer's own membership in a linguistic community. The inquirer's verbal behavior must also be interpreted by the community of investigators and this interpretive moment can never itself be displaced by objectivistic investigation. In fact, such investigation itself presupposes a communication community. But Apel's transcendental hermeneutics departs from Gadamer's historicism in that successive interpretations not only purport to understand differently but also raise an implicit claim to truth or correctness that can be clarified, once again, with reference to the ideal communication community. Furthermore, like Habermas, Apel does not exclude the possibility of introducing causal or functional explanations to clarify systematic distortions to communication, so long as they are "considered to be capable of conversion into a reflexively heightened self-understanding of the communicating parties" (1980, p. 125). In a response to externalist approaches (such as the strong program in the sociology of knowledge) Apel proposes a principle of self-appropriation that further develops this internalist (or rationalist) theme (see Kettner 1996).

In an important critique of the critical rationalism of Karl Raimund Popper and his followers, Apel further clarifies the status of transcendental pragmatics. He suggests that their skepticism with regard to the possibility of ultimate philosophical grounding is based on an abstractive fallacy in which sentences are viewed in isolation from the pragmatic contexts of argumentation. The so-called Münchhausen trilemma—that is, that all attempts to discover ultimate foundations result in either logical circularity, infinite regress, or an arbitrary end to the process of justification—can be overcome by moving from the level of semantic analysis to the level of pragmatics and recognizing that some presuppositions are necessary for the very possibility of intersubjectively valid criticism and argumentation. Similarly, he argues, even the "principle of fallibilism" (which holds that any claim can, in principle, be doubted) is only meaningful within an "institution of argumentation," where some pragmatic rules and norms are not open to question. Thus, contrary to the claim of critical rationalism, the principle of fallibilism does not exclude the notion of philosophical foundations and, Apel argues, certainly could not replace it as the basic principle of rational discourse (1998, chapter 4).

In a series of essays and in *Diskurs und Verantwortung* (1988) Apel argues that transcendental pragmatics can be used to develop an ethics of communication or *Diskursethik* that closely parallels the moral theory of Habermas. Like other cognitivist approaches, this ethics rejects the claim that moral judgments are ultimately the expressions of subjective preferences or an arbitrary will

and hence beyond the reach of rational justification. By elucidating its basic principle in relation to the pragmatic presuppositions of argumentation in general, Apel seeks a more secure foundation than Kant's appeal to a fact of reason or John Rawls's reflective equilibrium. According to the basic principle of his ethics of communication, only those norms are justified that could meet with the agreement of all concerned as participants in a practical discourse. However, in contrast to Habermas, to avoid an abstract utopianism, Apel (1988) maintains that this basic principle must be supplemented by a further principle of responsibility. Taken together, however, these two basic principles offer a secular foundation for a new global ethics.

See also Critical Theory.

Bibliography

WORKS BY APEL

Towards a Transformation of Philosophy. Translated Glyn Adey and David Frisby. London: Routledge and Kegan Paul, 1980.

Diskurs und Verantwortung: Das Problem des Übergangs zur postkonventionellen Moral. Frankfurt am Main, Germany: Suhrkamp, 1988.

From a Transcendental-Semiotic Point of View, edited by Marianna Papastephanou. New York: Manchester University Press, 1998.

WORKS ABOUT APEL

Kettner, Matthias. "Karl-Otto Apel's Contribution to Critical Theory." In *The Handbook of Critical Theory*, edited by David M. Rasmussen. Cambridge, MA: Blackwell, 1996.

Mendieta, Eduardo. *The Adventures of Transcendental Philosophy: Karl-Otto Apel's Semiotics and Discourse Ethics*. Lanham, MD: Rowman & Littlefield, 2002.

Kenneth Baynes (2005)

APOLOGISTS

"Apologists" is the term used historically in reference to Christian teachers from the second century to the fourth who wrote treatises defending their religion against charges of godlessness and immorality and usually ascribing these traits to their opponents. The way had been prepared for such writings in Hellenistic Judaism when Philo of Alexandria wrote an apologetic *Hypothetica* (now lost). All his extant writings can be regarded as attempts to set forth the nature of Judaism in a way comprehensible to a Greek audience. Josephus had explained away the revolt against Rome (*History of the Jewish War*), had rewritten the history of Israel (*Antiquities of the Jews*), and had provided an explicitly apologetic defense of Judaism (*Against Apion*). In addition, fragments of apologetic sermons are preserved in the New Testament book of Acts (14.15–17; 17.22–31), and perhaps may be reflected in I Thessalonians 1.9, I Corinthians 12.2, and Romans 1.18–32. The earliest known Christian apologists, however, wrote early in the second century.

Quadratus apparently wrote at Athens in the reign of Hadrian (117–138), and the one extant fragment of his work contrasts "our Savior" with some other savior. He argues that Jesus' healings and revivifications were authentic because some of the beneficiaries survived until Quadratus' own time. The *Apology* of Aristides (second century) begins with a semi-Stoic definition of God and goes on to show that all the gods of popular cult and legend cannot be gods because their deeds or sufferings are not in harmony with the definition. Finally, Aristides provides rather faint praise of Jews and high commendation for Christians. These writings cannot have won much, if any, favor with the pagans who read them.

The principal Christian apologist of the second century was Justin (c. 100–c. 165), born in Samaria of Greek parents and converted to Christianity (c. 130) after a fruitless quest for truth that had led him to religious-minded Middle Platonism. His education, he says, had not included many of the liberal arts; and from his account of his conversion, it is evident that he knew little about philosophy. A Christian whom he met by chance used Peripatetic arguments to indicate inconsistencies in Platonism. Justin, seeking new authority, was given the Old Testament prophecies. He had already admired the constancy of Christian martyrs; he soon became a Christian himself and instructed others, first in Asia Minor, later at Rome. He was martyred there between 163 and 167. Three of his works have survived: his *Apology,* written about 150 to show that Christians are not immoral and that Christ's life was foretold in the Old Testament; the *Dialogue with Trypho,* written about 160, developing this argument from the Old Testament; and an appendix to the *Apology,* also written about 160. His writings reflect a combination of Middle Platonism with Stoic terminology; he speaks of the divine Logos ("Word" for earlier Christians, "Reason" for Philo and the apologists), which was seminally present in some Greek philosophers but was incarnate in Christ. By working out some of the implications of this identification, Justin produced the first semiphilosophical Christian theology. It is possible that he knew something about Philo, but he cannot have understood his writings.

Justin's disciple Tatian (born c. 120), who later left the church, knew little about philosophy except for odd

details from philosophers' biographies, although like Justin he discussed the Logos as God's agent in creation and criticized the Stoic doctrine of fate. From Alexandria, perhaps, came the *Plea for the Christians* by Athenagoras (second century). He is the first Christian writer to reflect knowledge of the compendium of philosophical opinions apparently used in school teaching, especially by Skeptics. On the basis of earlier arguments in the schools, Athenagoras constructed a defense of the unity of God; and his later work *On Resurrection* contains a similar rearrangement of arguments from the schools to prove that God is able and willing to raise corpses, and will do so because man is a unity of soul and body. The last Greek apologist of the second century was Theophilus, bishop of Antioch, whose work in three books, *To Antolycus,* is concerned with the works of the invisible God (Philonic-Platonic arguments), God's revelation to the prophets and his six-day work of creation, and Christian ethics and the antiquity of the Jewish-Christian revelation. Theophilus used handbooks for much of his information about philosophy, but he may have read some works by Plato.

Generally speaking, the second-century apologists knew something about Platonism (that is, Middle Platonism) and Stoicism (largely the older Stoics) and made use of philosophy at points where it supported—or could be made to support—their own ideas of revelation, creation, providence, free will, divine punishment, and resurrection. They reinterpreted the Johannine "Word" as the divine Reason, instrumental in creation and revelation alike; Justin, unlike the others, used this Reason to explain how it was that some Greeks possessed inklings of the truth. The apologists also stressed the disputes among various schools in order to show how wrong the Greek philosophers usually were and how subjective their knowledge was.

At the very end of the second century an ex-lawyer named Tertullian produced two apologies in Latin. The first, *Ad Nationes,* is not very original, since much of it is derived from Varro's critique of Roman religion; the second, the *Apologeticum,* is a completely rewritten, and much more brilliant, revision of the first. Either before or after these works were published, another Latin apology, the *Octavius* of Minucius Felix, appropriated much of Cicero's treatise *De Natura Deorum* to Christian use. Both Tertullian and Minucius also made use of their Greek predecessors' writings.

Greek apologetic continued to be produced in the third century; examples include the anonymous booklet *To Diognetus,* the *Protrepticus* by Clement of Alexandria, and the highly important work *Against Celsus* by Origen, in which the author often makes use of philosophical topoi (commonplaces) in his argument (for instance, Platonic discussions of the divine nature; Stoic arguments in favor of providence) and reveals that he shares many presuppositions with Celsus himself. Apparently some of the writings later ascribed to Justin, such as the *Cohortatio* and the *Oration,* also come from the third century. In them we find extensive use of handbooks and a little first-hand knowledge of philosophical writings.

Stimulus for the production of further apologies was provided about 260, when the Neoplatonist Porphyry produced a work in fifteen books, *Against the Christians.* Now lost because it was later proscribed, this work criticized the Old and New Testament, the apostles, and the life and thought of the church. The *Praeparatio Evangelica* of Eusebius is primarily a reply to it and to the similar work by Hierocles. In the fourth century the emperor Julian composed a work in three books, *Against the Galileans*; this was answered by Theodoret and Cyril of Alexandria, among others. Among the later Latin apologists we should mention Arnobius (d. c. 330, vaguely acquainted with Neoplatonism), Lactantius (c. 240–c. 320, who relied extensively on Cicero), and—above all—Augustine, whose *City of God* contains much from Varro and sets forth a Christian philosophy of history in response to Porphyry and other critics.

The significance of the apologists lies not so much in what they actually wrote (their works seem to have been read chiefly within the church) but in the influence their effort had on one another's thought and on the thought of later theologians. Their criticisms of Greco-Roman philosophy compelled them not only to learn something about it but also to employ its modes of discourse and some of its axioms in expounding the nature of Christianity. It was through the apologists that philosophical theology entered, and to some measure shaped, Christian thought. To be sure, later theologians could not accept the apologists' rather naive theologies (Irenaeus, for example, learned from the apologists but also corrected some of their statements); but impetus for philosophical study was given in the apologists' works and by the school of Alexandria, whose members were more at home in philosophy, especially Platonism.

All the early apologists, and most of the later ones, admired Plato and were influenced by Middle Platonism; the work they valued most highly was the *Timaeus,* in which they found intimations of Christianity (sometimes explained as derived from the Old Testament). They usually employed traditional Stoic arguments in defense of

providence and anti-Stoic arguments in opposition to fate. When they dealt with pagan mythology, they often employed the arguments of Skeptics. Their approach, then, was eclectic; and the famous statement of Justin, "Whatever has been well spoken by anyone belongs to us," had been made by eclectic philosophers. At the same time, the apologists were aware of the difference between all philosophies and their own cardinal doctrines of God (*Creator ex ouk ontōn,* "wrathful against sin"), the Incarnation, and the future corporeal resurrection. Even those apologists who were most eager to express their doctrines in philosophical modes of discourse were usually aware that the basic beliefs could not be so expressed. Theophilus, for example, defines *pistis* (faith) in a manner strongly reminiscent of the probabilism of Carneades and then provides analogies to the resurrection of the body that are based on Stoic arguments for the cosmic cycle. He admits, however, that only faith is ultimately convincing.

See also Augustine, St.; Cicero, Marcus Tullius; Clement of Alexandria; Eusebius; Origen; Platonism and the Platonic Tradition; Porphyry; Tertullian, Quintus Septimius Florens.

Bibliography

Amand, David. *Fatalisme et liberté.* Louvain: Bibliothèque de l'Université, 1945.

Becker, Carl. *Tertullians Apologeticum.* Munich: Kösel-Verlag, 1954.

Canivet, Pierre. *Histoire d'une enterprise apologétique au V^{e} siècle.* Paris: Bloud and Gay, 1957.

Daniélou, Jean. *Message évangélique et culture hellénistique.* Tournai: Desclée, 1961.

Geffcken, Johannes. *Zwei griechische Apologeten.* Leipzig: B.G. Teubner, 1907.

Pellegrino, Michele. *Gli apologeti.* Rome: Anonima Veritas, 1947.

Puech, Aimé. *Les apologistes grecs.* Paris: Hachette, 1912.

Robert M. Grant (1967)

APPEARANCE AND REALITY

In *The Problems of Philosophy* Bertrand Russell referred to the distinction between appearance and reality as "one of the distinctions that cause most trouble in philosophy." Why it should cause trouble in philosophy, however, when it causes little or no trouble outside of philosophy, Russell did not say. The distinction has played an important part in the thinking of many philosophers, and some of them, including Russell, have employed it in curious ways to support odd and seemingly paradoxical claims. It may be this last fact that Russell had in mind when he spoke of trouble.

Before turning to some of its troublesome uses in philosophy, let us consider some of its relatively untroublesome uses in everyday discourse.

LOOKS AND APPEARANCES

There is a potentially troublemaking ambiguity in the term *to appear* and its cognates. (This ambiguity is not peculiar to English but is also to be found, for example, in the Greek verb *phainesthai* and its cognates.) Contrary to Russell's suggestion, the distinction between appearance and reality is not simply the distinction "between what things seem to be and what they are," more precisely, the distinction between what things seem to be and what they are is not a simple distinction. There are at least two groups of appearance idioms—what might be called "seeming idioms" and "looking idioms." The first group typically includes such expressions as "appears to be," "seems to be," "gives the appearance of being"; the second, such expressions as "appears," "looks," "feels," "tastes," "sounds."

The two groups are not always as obviously distinct as these examples make them appear to be. The same expression, particularly one from the second group (notoriously, "appears," but also such expressions as "looks as if"), may be used either as a seeming expression or as a looking expression. For example, "The oar appears bent" may mean either "The oar looks bent" or "The oar appears to be bent." These are by no means the same. I may say that the oar appears to be bent *because* it looks bent, and this is not to say that the oar appears to be bent because it appears to be bent or that it looks bent because it looks bent. Nor is there any necessary connection between the two statements—or, generally, between statements employing seeming idioms and those employing looking idioms. "The oar looks bent" does not imply or entail "The oar appears to be bent"; for the oar may look bent—immersed in water, it naturally does—without appearing to be bent. As St. Augustine put it in a striking passage in *Contra Academicos* (III, xi, 26): "'Is that true, then, which the eyes see in the case of the oar in water?' 'Quite true. For since there is a special reason for the oar's looking (*videretur*) that way, I should rather accuse my eyes of playing me false if the oar looked straight (*rectus appareret*) when dipped in water; for in that case my eyes would not be seeing what, under the circumstances, ought to be seen.'" (Compare J. L. Austin, *Sense and Sensibilia,* p. 26.) The oar's looking bent in water is not an

illusion, something that appears to be the case but is not; but this does not mean that the oar does not look bent. Conversely, "The oar appears to be bent" does not imply "The oar looks bent"; for the oar may appear to be bent without its looking bent; there may be reasons for saying that it appears to be bent (evidence that suggests that it is bent) other than its looking bent. (On this distinction, compare C. D. Broad, *Scientific Thought,* pp. 236–237.)

An example of the troublemaking neglect—or at least apparent neglect—of this distinction is to be found in Russell (op. cit.): "Although I believe that the table is 'really' of the same colour all over, the parts that reflect the light look much brighter than the other parts, and some parts look white because of reflected light. I know that, if I move, the parts that reflect the light will be different, so that the apparent distribution of colours on the table will change." But further on he wrote: "To return to the table. It is evident from what we have found, that there is no colour which pre-eminently appears to be *the* colour of the table, or even of any one particular part of the table—it appears to be of different colours from different points of view, and there is no reason for regarding some of these as more really its colour than others." But if all we have found is that the parts of the table that reflect the light *look* brighter than the others, it is by no means "evident" that there is no color which *appears to be* the color of the table.

SEEMING IDIOMS. Seeming idioms have nothing strictly to do with the senses; looking idioms characteristically do. From the evidence at hand, it may *appear,* or *look as if,* there will be an economic recession within the year. The characteristic uses of seeming idioms are to express what one believes is probably the case, to refrain from committing oneself, or to express hesitancy about what *is* the case. (Compare G. J. Warnock, *Berkeley,* p. 186: "The essential function of the language of 'seeming' is that it is noncommittal as to the actual facts.") Hence, "I know that *X* is *Y,* but it appears (to me) that it is not *Y*" is odd or paradoxical in much the same way as is "I know that *X* is *Y,* but it may not be the case that it is." From "*X* appears to be *Y*" (though *not* "merely appears to be *Y*"), I cannot validly infer either "*X* is *Y*" or "*X* is not *Y.*" But "*X* appears to be *Y*" entails that it is possible that *X* is *Y* and possible that *X* is not *Y.*

The same is not true of looking idioms, except in so far as they double as seeming idioms. No oddity or paradox is involved in saying such things as "I know that the two lines in Müller-Lyer's drawing *are* the same length, but one of them still *looks* longer than the other."

LOOKING IDIOMS. Looking idioms have a number of uses or senses that must be kept distinct.

Noticing resemblances. To notice that an inkblot has the appearance of (looks like) a face or that Alfredo's voice sounds like Caruso's is to note a visible resemblance between the inkblot and a face or an audible resemblance between Alfredo's voice and Caruso's. Here appearance does not normally contrast with what is possibly reality; rather it is a reality. "Alfredo's voice sounds like Caruso's" does not mean either "Alfredo's voice appears to be Caruso's" or "Alfredo's voice (merely) sounds like Caruso's, but it isn't Caruso's voice." To be sure, in certain circumstances one might be misled by appearances. For instance, by the audible resemblance between Alfredo's voice and Caruso's one might suppose that he was hearing Caruso's voice. Compare, however, "At a distance (in this light, at a quick glance) that looks like blood (a dollar bill), but it's really just red paint (a soap coupon)."

Describing. To describe something's appearance may merely be to describe its perceptible (visible, audible, tactile) features, and as such it is to describe how something *is,* not how it looks or appears as possibly *opposed* to how it is. Here the *apparent* qualities of something are the *real* perceptible qualities of it. To describe a man's appearance, as opposed, say, to his character, is to describe those features of him (his "looks") that he can be seen to possess. Appearances in this sense are what are most often referred to as phenomena in the nonphilosophical use of the latter term, in such phrases as "biological phenomena."

"Looks" and "merely looks." The phrase "mere appearance" ("merely looks, sounds") shows that there is a sense of "appears" as a looking idiom which is neutral with respect to how things are. "*X* merely looks red (to me, or under such-and-such conditions)" implies that *X* is not (really) red. But simply from "*X* looks red (to me, or under such-and-such conditions)" I cannot validly infer either that *X* (really) is red or that *X* (really) is not red. If it is possible, however, for *X* to look (sound, feel, taste) *Y,* it must at least be possible for *X* (really) to be *Y.* This logical feature of looking idioms, which—in this sense—they share with seeming idioms, may be the source of some confusion between them.

PROTAGOREAN RELATIVISM

According to Plato (*Theaetetus,* 152; Cornford trans.), Protagoras held that "man is the measure of all things—alike of the being of things that are and of the non-being of things that are not." And by this he meant that "any given thing is to me such as it appears to me, and is to you

such as it appears to you." This statement can be read in two different ways, depending on whether "appears" is construed as a seeming idiom or a looking idiom. In either interpretation, however, it is a paradox or else a tautology.

Expressions such as "is for me" and "is for you" are distinctly odd, and one is puzzled to know what to make of them. If they are construed to mean the same as "is," Protagoras' statement then becomes manifestly paradoxical. For if "*X* appears to me to be *Y* (or looks *Y* to me)" and "*X* appears to you to be *Z* (or looks *Z* to you)" are equivalent respectively to "*X* is *Y*" and "*X* is *Z*," where *Y* and *Z* represent logically incompatible predicates, then the joint affirmation of two (possibly) true propositions, "*X* looks *Y* to me" and "*X* looks *Z* to you," would be equivalent to the necessarily false proposition that *X* is both *Y* and *Z*.

On the other hand, if we interpret "is for me" to mean the same as "appears to me" and "is for you" as "appears to you," Protagoras' dictum reduces to a tautology. For if "*X* appears to me to be *Y*" and "*X* appears to you to be *Z*" are equivalent respectively to "*X* is *Y* for me" and "*X* is *Z* for you," then, even if *Y* and *Z* represent logically incompatible predicates, the equivalent statements can be substituted for one another. In that case, Protagoras' dictum, generalized, reduces to either "Everything is for any given person such as it is for that person" or "Everything appears to any given person such as it appears to that person." But since the two statements are themselves equivalent, the effect of Protagoras' dictum is to obliterate any possible distinction between appearance and reality, or to claim what is clearly false, that there is no such distinction.

Protagoras' statement can be read in yet another way, but read in that way it is also a truism. The Greek verb *phainesthai,* especially with the participle, was used to state, not that something (merely) appears to be so, but that something manifestly is so. Read in this way, Protagoras' claim that appearance is reality is simply the claim that what is manifestly the case is the case. This innocent truism may have been intended to remind those of Protagoras' contemporaries who contemned the common run of men for living by appearances, which they equated with error, that what is reliably observed to be the case is justifiably said to be the case.

THE ARGUMENT FROM ILLUSION

What has been called the "argument from illusion" has been used by many philosophers (for example, George Berkeley in *Three Dialogues, I,* and A. J. Ayer in *Foundations of Empirical Knowledge,* pp. 3–5) to justify some form of phenomenalism or subjective idealism. The argument rests on the fact that things sometimes appear (for example, look) different to different observers or to the same observer in different circumstances. This fact is supposed to show that sensible qualities, such as colors or odors, are not really "in" things. For if things can, say, look one color when they are (supposedly) really another, then we can never say what color they really are, what color really "inheres" in them. For all sensible qualities, as Berkeley put it, "are equally apparent"; he seems to have meant that for every putatively veridical perception there is a possible corresponding illusory one (or wherever it is possible that "*X* is *Y*" is true, it is equally possible that "*X* merely looks *Y*" is true). Hence, given any perception, *P,* it is possible that *P* is veridical and possible that *P* is illusory. But since there is no apparent or observable difference between a veridical *P* and an illusory *P,* we cannot in principle tell which it is. We cannot, for example, say what colors things *are*; we can only say what colors they *look.*

The consequence of this argument is the same as that of Protagoras' dictum, namely, to obliterate in principle any distinction between "is" and "(merely) looks or sounds." But this is a distinction on which the argument itself rests: if the distinction cannot, in principle, be made, then the argument cannot get off the ground; but if the distinction can, in principle, be made, the conclusion of the argument cannot be true.

"IS Y" AS A FUNCTION OF "APPEARS Y." Many philosophers who have used the argument from illusion have attempted to resist the consequence that there is then no distinction between "is" and "(merely) looks." Berkeley, for example, protested that "the distinction between realities and chimeras retains its full force" (*Principles of Human Knowledge,* §34). He was able to suppose that it does because he supposed that "*X* is *Y*" is a logical function of "*X* appears (appears to be *or,* for example, looks) *Y*": when the appearances of *X* are not only "lively" but "steady," "orderly," and "coherent," we say that *X is* (really) *Y* and not that it merely *appears Y.* Being is orderly and coherent appearing (*Principles,* §29).

But if this is so, the distinction between realities and chimeras does not retain its full force. "*X* appears *Y* consistently (steadily, in an orderly and coherent way)" neither is equivalent to, nor does it entail, "*X* is *Y*"; for it is possible that the former is true while the latter is false. The truth of the former may be *evidence for* the truth of the latter, but the latter is not a logical function of the former. (Compare Warnock, op. cit., pp. 180–182.) The same

holds for such claims as that of G. E. Moore (*Commonplace Book,* p. 145) that "'This book is blue' = This book looks (or *would* look) blue to normal people … who look at it by *good* daylight at *normal distances,* i.e. not too far off or too near."

PHENOMENA AND THINGS-IN-THEMSELVES

One of the foundation stones of Immanuel Kant's philosophy is the claim that "we can know objects only as they *appear* to us (to our senses), not as they may be in themselves" (*Prolegomena,* §10.) Read in one way, Kant's claim is tautologous. If by "an appearance" we mean a possible object of knowledge and by "a thing-in-itself" something that can be "thought" but cannot be known, the claim reduces to "What we can know, we can know; and what we cannot know, we cannot know." As such, this tells us nothing about the limits of knowledge, about what we can know, any more than "God can do everything that it is possible for God to do" tells us anything about the extent of God's powers.

Kant may, however, have meant the following: I can know that *X* is *Y* only if *X* can appear (to be) *Y*; if, in principle, *X* cannot appear (to be) *Y,* then I cannot know that *X* is *Y.* This, too, is a truism. But it does not follow from this that "the things we intuit are not in themselves what we intuit them as being. … As appearances, they cannot exist in themselves, but only in us" (*Critique of Pure Reason,* A42; Kemp Smith trans.). That is, it does not follow that *X* as it appears is not what it is apart from how it appears; nor does it follow that what *X* is apart from how it appears is different from how it appears. To allow Kant's inference is implicitly to endorse a paradox or to adopt a new use of "appears" to which no sense has been given. For if something appears (to be) so, it must be *possible* for it to *be* so "in itself"; and this is precisely the possibility which Kant does not allow.

APPEARANCES OF THE IMPOSSIBLE. Closely related to Kant's distinction between appearances and things-in-themselves is the notion of appearances of the impossible. According to Parmenides and Zeno, multiplicity and motion, empty space and time, are impossible; yet things appear to be many, some of them appear to move, and so on. Similarly, for Gottfried Wilhelm Leibniz bodies with their qualities, such as colors, are well-founded appearances (*phaenomena bene fundata*), mere appearances "grounded" in monads and their perceptions; in reality there can be no such things as colored bodies. And according to F. H. Bradley in *Appearance and Reality,* space, time, motion and change, causation, things, and the self are "unreal as such" because they "contradict themselves"; hence, they are "mere appearances" or "contradictory appearances."

Taken at face value, this view is blatantly paradoxical: If for something to appear (to be) the case it must be *possible* for it "really" to be the case, then if it is *impossible* for it to be the case, it is impossible for it to appear (to be) the case. (Compare Morris Lazerowitz, *The Structure of Metaphysics,* pp. 208–209.) The metaphysician of "contradictory appearances," however, may mean that for certain kinds of things, *t,* it is *never* permissible to say "There are *t*'s," but only "There appear to be *t*'s." But this, as Lazerowitz has pointed out (op. cit., esp. p. 225), has the consequence of obliterating the distinction between "is" and "appears" and hence of depriving "appears" of its meaning. For if "There are *t*'s" is in principle disallowed, "There appear to be *t*'s" loses its sense.

See also Augustine, St.; Austin, John Langshaw; Ayer, Alfred Jules; Berkeley, George; Bradley, Francis Herbert; Illusions; Kant, Immanuel; Moore, George Edward; Plato; Russell, Bertrand Arthur William.

Bibliography

Augustine, St. *Contra Academicos.* Translated by John J. O'Meara as *Against the Academics.* Westminster, MD: Newman Press, 1951.

Austin, J. L. *Sense and Sensibilia.* Oxford: Clarendon Press, 1962.

Ayer, A. J. *The Foundations of Empirical Knowledge.* London: Macmillan, 1940.

Berkeley, George. *A Treatise concerning the Principles of Human Knowledge.* Dublin: A. Rhames for J. Pepyat, 1710.

Berkeley, George. *Three Dialogues between Hylas and Philonous.* London: Henry Clements, 1713.

Bradley, F. H. *Appearance and Reality,* 2nd ed. Oxford: Clarendon Press, 1897.

Bradley, F. H. *Essays on Truth and Reality.* Oxford: Clarendon Press, 1914. Ch. 9.

Broad, C. D. *Perception, Physics and Reality.* Cambridge, U.K.: Cambridge University Press, 1914. Ch. 2.

Broad, C. D. *Scientific Thought.* London: K. Paul, Trench, Trubner, 1923. Part 2.

Chisholm, R. M. *Perceiving: A Philosophical Study.* Ithaca, NY: Cornell University Press, 1957.

Chisholm, R. M. "The Theory of Appearing." In *Philosophical Analysis,* edited by Max Black, pp. 102–118. Ithaca, NY: Cornell University Press, 1950.

Chisholm, R. M. "Theory of Knowledge." In his *Philosophy,* pp. 233–344. Humanistic Scholarship in America, The Princeton Studies. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Kant, Immanuel. *Critique of Pure Reason.* Translated by Norman Kemp Smith. London, 1919.

Lazerowitz, Morris. "Appearance and Reality." In his *The Structure of Metaphysics.* London: Routledge and Paul, 1955. Ch. 10.
Lean, Martin. *Sense-Perception and Matter.* London: Routledge and Paul, 1953.
Moore, G. E. *The Commonplace Book 1919–1953.* London: Allen and Unwin, 1962. Passim.
Moore, G. E. "The Conception of Reality." In his *Philosophical Studies.* London: K. Paul, Trench, Trubner, 1922. Ch. 6.
Plato. *Theaetetus.* Translated by F. M. Cornford as *Plato's Theory of Knowledge.* London: K. Paul, Trench, Trubner, 1935.
Price, H. H. "Appearing and Appearances." *American Philosophical Quarterly* 1 (1964): 3–19.
Price, H. H. *Perception.* London: Methuen, 1932.
Prichard, H. A. "Appearances and Reality." *Mind* 15 (1906): 223–229.
Russell, Bertrand. *The Problems of Philosophy.* New York: Henry Holt, 1912. Ch. 1.
Ryle, Gilbert. *The Concept of Mind.* London: Hutchinson's University Library, 1949. Ch. 7.
Sibley, Frank. "Aesthetics and the Looks of Things." *Journal of Philosophy* 16 (1959): 905–915.
Taylor, A. E. *Elements of Metaphysics.* London: Methuen, 1903. Bk. 2, Ch. 3.
Warnock, G. J. *Berkeley.* London: Penguin, 1953. Ch. 9.
Wollheim, Richard. *F. H. Bradley.* Harmondsworth, U.K.: Penguin, 1959. Ch. 5.

W. E. Kennick (1967)

APPERCEPTION

Apperception is usually defined as the mental process that raises subconscious or indistinct impressions to the level of attention and at the same time arranges them into a coherent intellectual order. The term *apperception*, however, has been used ambiguously, sometimes to mean merely consciousness or awareness, at other times to mean the acts of concentration and assimilation. Inevitably, a process of such significance has implicitly and explicitly been dealt with by philosophers ever since they first concerned themselves with the cognitive process. Aristotle, the Church Fathers, and the Scholastics all distinguished between vague notions and feelings on the one hand, and conceptions brought about by an act of intellectual willing on the other.

DESCARTES

The concept of apperception (in the form of the verb *apercevoir*) appears in René Descartes's *Traité des passions*.

Later writers generally use the term *perception* for denoting a state of dim awareness. So John Locke believes that perception is "the first step and degree towards knowledge, and the inlet of all materials of it." It "is in some degree in all sorts of animals" (*Essay concerning Human Understanding*, Book II, Ch. 9). On the other hand, apperception denotes a state of conscious or reflecting awareness.

In contrast, Descartes makes no distinction between the two. But he stresses the volitional element (which he calls passion) in the cognitive process: "For it is certain that we would not even know how to will something, unless we had apperceived it by the same medium by which we will. And just as one can say with regard to our soul that willing is a form of action, so one can also say that there is in the soul an element ["passion"] by which it apperceives that which it wills" (*Traité des passions*).

LEIBNIZ

It was Gottfried Wilhelm Leibniz who introduced the concept of apperception into the more technical philosophical tradition. In his *Principes de la nature fondés en raison et de la grâce* he says: "One should distinguish between *perception*, which is an inner state of the monad reflecting the outer world, and *apperception*, which is our conscious reflection of the inner state of the monad."

For the understanding of Leibniz's ideas about perception and apperception, one should also refer to his *Nouveaux essais sur l'entendement humain*, which contain a discussion of Locke's *Essay concerning Human Understanding*. There Leibniz objects to Locke's *tabula rasa* theory, according to which "there are no innate principles in the mind" (Book I, Ch. 2). Leibniz's insistence on innate mental powers had a decisive influence on the idealism of Immanuel Kant and Johann Friedrich Herbart.

KANT

The concept of apperception was taken up by Kant in his *Critique of Pure Reason*. There he distinguished between empirical apperception, the person's awareness of himself which depends on the changing conditions of his consciousness, and transcendental apperception, or "pure reason," the inner, unchangeable fundamental, and therefore "transcendental" unity of consciousness. This transcendental unity of consciousness precedes all data of perception and makes possible their inner order and meaning ("Transcendental Logic," Para. 12). It consists of the ideas of space and time, which are not objects of perception but modes of perceiving, and a number of categories which Kant orders under the headings of quantity, quality, relation, and modality. Kant's attempt to organize these categories and their subcategories according to a symmetrical scheme has been generally rejected as artifi-

cial. Kant's rejection of the opinion, however, that our conscious reasoning about the world reflects the world as it really is remains as one of the great epistemological problems in his concept of apperception.

IDEALISTS

The self-critical quality in Kant's philosophy was not heeded by romantic idealists impatient to achieve a complete insight into the essence of all existence. Thus Johann Gottlieb Fichte turned Kant's self-critical concept of apperception into the absolute self; Hegel developed logical idealism; and Friedrich Wilhelm Joseph von Schelling maintained in his philosophy of identity that the evolution of mind or consciousness is nothing but the evolution of ultimate reality from its prerational and groping state of willing toward self-consciousness and self-direction, toward the discovery of its inherent and universal laws. Whatever we think about Schelling's lofty speculation, it led its author to the understanding of myth. For in myth, so Schelling concluded, the human mind in its prerational state creates its first perceptions of reality in the form of artistic intuition and imagery. Myth, so we could say with Schelling, is not untruth but pretruth. About half a century later, following Schelling's lead, Wilhelm Wundt became one of the foremost interpreters of prerational or mythical thinking.

HERBART

In contrast with the romanticists, Kant's successor, Johann Friedrich Herbart, insisted on a less romantic and more empirical interpretation of the transcendentalist position. In the second part of his *Psychologie als Wissenschaft*, however, Herbart characterizes the gift of apperception as one—though not the only one—of the qualities that distinguish man from animal because it gives him the power of reflection. In the human soul, so Herbart says, there are operating series of presentations, combinations, and whole masses of perceptions that are sometimes completely and sometimes incompletely interwoven, in part conforming and in part opposed to each other. It is the function of apperception to assimilate the various and often divergent ideas. In this process the older apperceptive mass, consisting of concepts, judgments, and maxims, will tend to assimilate more recent and less settled impressions. No one, however, can measure how strong the older apperceptive mass must be in order to fulfill effectively the function of assimilation.

Obviously, the power of apperception as conceived by Herbart is closely related to a person's inner stability, self-consciousness, and self-identity. Apperception requires will and attention in order to function adequately. A mentally sick person will be unable to perform it.

Inevitably, the concept of apperception plays a decisive role in Herbart's pedagogical theory. In his *Allgemeine Pädagogik aus dem Zweck der Erziehung Abgeleitet*, Herbart emphasizes the obligation of the teacher to arrange the course of instruction in such a way that the new material can be properly integrated with the already available store of knowledge. If the two fall apart, the learner cannot assimilate the new experience and will feel frustrated.

WUNDT

The qualities of will and attention, which from Descartes to Herbart were emphasized as inherent in the apperceptive process, are still more accentuated by Wilhelm Wundt. In his *Grundriss der Psychologie*, Wundt distinguishes between passive apperception, in which the consciousness simply accepts impressions, and active apperception, in which the new impression is met by an emotional state of tension followed by a sense of satisfaction. Furthermore, in all apperception a personifying element is at work in that the apperceived objects are colored by the mode of the apperceiving subject. This is the reason why we tend to identify apperceived objects with our own form of existence. The most obvious historical example of this tendency is myth, in which, for example, animals, the forces of nature, and the gods appear in anthropomorphic transfiguration.

Entirely in the spirit of Wundt is the following (freely translated) passage from the well-known *Grundriss der Geschichte der Philosophie seit Beginn des neunzehnten Jahrhunderts*:

> There is nothing inside and outside of man which he could call totally his own but his will. … Hence, looking for the terminus of individual psychological regression, we discover the *inner will* or the *pure apperception*, which is not in a state of quiet, but in a state of never resting activity. The apperceptive will is not an a-posteriori conception, but an a-priori, postulated by reason, a transcendental quality of the soul, postulated by empirical psychology as the ultimate source of all mental processes, yet at the same time beyond the competence of the empirical psychologist.

THE DEEPER UNITY

In quoting the foregoing passage (omitted in later editions of Ueberweg-Heinze) we have already indicated the deeper unity that in spite of all differences underlies the apperception theories of Leibniz, Kant, Herbart, and Wundt. They predicate a transcendental element, or an inherent logos, in the human process of cognition because they are convinced that there is no other explanation for its uniting and ordering capacity. They belong, in the wide sense of the term, to the "idealistic" tradition of the *philosophia perennis*, although they are in no way opposed to painstaking empirical and statistical inquiry, as the examples of Herbart and Wundt prove.

In postulating a transempirical factor as the condition of experience, however, they expose themselves to the reproach of mysticism by the empiricist. And there can be no doubt that the modern experimental, associationist, and behaviorist schools have made us more critical of psychological concept. Nevertheless, it still seems to many contemporary philosophers and psychologists that a purely empirical account of knowledge is inadequate and that in order to achieve a defensible position it is necessary to have recourse to nonempirical factors such as apperception.

See also Aristotle; Descartes, René; Fichte, Johann Gottlieb; Herbart, Johann Friedrich; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Patristic Philosophy; Schelling, Friedrich Wilhelm Joseph von; Wundt, Wilhelm.

Bibliography

In addition to the works cited in the text, the following may be consulted: Benno Erdmann, "Zur Theorie der Apperzeption," in *Vierteljahrsschrift für wissenschaftliche Philosophie* 10 (1886): 307ff.; Karl Lange, *Ueber Apperception*, 6th rev. ed. (Leipzig, 1899), translated by E. E. Brown (Boston, 1893); L. H. Lüdtke, "Kritische Geschichte der Apperzeptionsbegriffs," in *Zeitschrift für Philosophie* (1911); Hugo Münsterberg, *Grundzüge der Psychologie* (Leipzig: J.A. Barth, 1900), pp. 436–457; G. F Stout, "Apperception and the Movement of Attention," in *Mind* 16 (1891): 23–53, and *Analytic Psychology* (London, 1896); and Friedrich Ueberweg, *Grundriss der Geschichte der Philosophie seit Beginn des neunzehnten Jahrhunderts*, 10th ed., edited by Max Heinze. (Berlin: Mittler, 1902).

OTHER RECOMMENDED TITLES

Allison, Henry E. "Apperception and Analyticity in the B-Deduction." *Grazer Philosophische Studien* 44 (1993): 233–252.

BonJour, Laurence, and Ernest Sosa. *Epistemic Justification*. Oxford: Blackwell, 2003.

Brandom, Robert, B. "Leibniz and Degrees of Perception." *Journal of the History of Philosophy* 19 (1981): 447–479.

Castañeda, Hector-Neri. "The Role of Apperception in Kant's Transcendental Deduction of the Categories." *Nous* 24 (1) (1990): 147–157.

Howell, Robert. "Apperception and the 1787 Transcendental Deduction." *Synthese* 47 (1981): 385–448.

Kitcher, Patricia. *Apperception and Epistemic Responsibility in Central Themes in Early Modern Philosophy*. Indianapolis, IN: Hackett, 1990.

Kulstad, Mark. *Leibniz on Apperception, Consciousness and Reflection*. Germany: Philosophia, 1990.

Robert Ulich (1967)
Bibliography updated by Benjamin Fiedor (2005)

APPLIED ETHICS

Moral philosophers have traditionally aspired to normative theories of what is right or wrong that are set out in the most general terms. But a practical price is paid for generality in ethical theory: It is often unclear whether and, if so, how theory is to be applied in specific cases and contexts. The terms applied ethics and practical ethics came in vogue in the 1970s, when philosophical ethics began to address issues in professional ethics as well as social problems such as capital punishment, abortion, environmental responsibility, and affirmative action. Philosophers interested in applying their training to such problems share with persons from numerous other fields the conviction that decision making in these areas is fundamentally moral and of the highest social importance.

Philosophers working in applied ethics sometimes do more than teach and publish articles about applications of ethical theory. Their work involves actual applications. They serve as consultants to government agencies, hospitals, law firms, physician groups, business corporations, and engineering firms. Branching out further, they serve as advisers on ethics to radio and educational television, serve on national and state commissions on ethics and policy, and give testimony to legislative bodies. Occasionally, they draft public policy documents, some with the force of law.

Controversies have arisen about whether philosophers have an ethical expertise suited to such work and also about whether the work is philosophical in any interesting sense. Enthusiasm about applied ethics is mixed in academic philosophy. It has been criticized as lacking in serious scholarship, and many philosophers regard it as reducing ethics to engineering—a mere device of problem solving. Some philosophers are not convinced that philosophical theories have a significant role to play in

the analysis of cases or in policy and professional contexts, and others are skeptical that philosophical theories have direct practical implications.

DEFINITIONAL PROBLEMS

"Applied ethics" has proved difficult to define, but the following is a widely accepted account: Applied ethics is the application of general ethical theories to moral problems with the objective of solving the problems. However, this definition is so narrow that many will not recognize is as reflecting their understanding of either the appropriate method or content. "Applied ethics" is also used more broadly to refer to any use of philosophical methods critically to examine practical moral decisions and to treat moral problems, practices, and policies in the professions, technology, government, and the like. This broader usage permits a range of philosophical methods (including conceptual analysis, reflective equilibrium, phenomenology, etc.) and does not insist on problem solving as the objective.

Biomedical ethics, political ethics, journalistic ethics, legal ethics, environmental ethics, and business ethics are fertile areas for such philosophical investigation. However, "applied ethics" is not synonymous with "professional ethics" (a category from which business ethics is often excluded). Problems such as the allocation of scarce social resources, just wars, abortion, conflicts of interest in surrogate decision making, whistleblowing, the entrapment of public officials, research on animals, and the confidentiality of tax information extend beyond professional conduct, but all are in the domain of applied ethics. Likewise, professional ethics should not be viewed as a part of the wider domain of applied ethics. The latter is usually understood as the province of philosophy, the former as reaching well beyond philosophy and into the professions themselves.

HISTORY

Philosophers from Socrates to the present have been attracted to topics in applied ethics such as civil disobedience, suicide, and free speech; and philosophers have written in detail about practical reasoning. Nonetheless, it is arguably the case that there never has been a genuine practical program of applied philosophy in the history of philosophy (the casuists possibly qualifying as an exception). Philosophers have traditionally tried to account for and justify morality, to clarify concepts, to examine how moral judgments and arguments are made, and to array basic principles—not to use either morality or theories to solve practical problems.

This traditional set of commitments began to undergo modification about the time the *Encyclopedia of Philosophy* was first published in 1967. Many hypotheses can be invoked to explain why. The most plausible explanation is that law, ethics, and many of the professions—including medicine, business, engineering, and scientific research—were profoundly affected by issues and concerns in the wider society regarding individual liberties, social equality, and various forms of abuse and injustice. The issues raised by civil rights, women's rights, the consumer movement, the environmental movement, and the rights of prisoners and the mentally ill often included ethical issues that stimulated the imagination of philosophers and came to be regarded by many as essentially philosophical problems. Teaching in the philosophy classroom was influenced by these and other social concerns, most noticeably about unjust wars, dramatic ethical lapses in institutions, domestic violence, and international terrorism. Increases in the number of working women, affirmative action programs, escalation in international business competition, and a host of other factors heightened awareness. Classroom successes propelled the new applied ethics in philosophy throughout the 1970s, when few philosophers were working in the area but public interest was increasing.

It is difficult to identify landmark events that stimulated philosophers prior to *Roe v. Wade* (the U.S. Supreme Court decision on abortion in 1973), which deeply affected applied philosophical thinking. But at least one other landmark deserves mention. Research ethics had been poorly developed and almost universally ignored in all disciplines prior to the Nuremberg Trials. This apathy was shaken when the Nuremberg Military Tribunals unambiguously condemned the sinister political motivation and moral failures of Nazi physicians. The ten principles constituting the "Nuremberg Code" served as a model for many professional and governmental codes formulated in the 1950s and 1960s and eventually influenced philosophers as well.

In the late 1960s and early 1970s there emerged a rich and complex interplay of scholarly publications, journalism, public outrage, legislation, and case law. The 1970s and 1980s saw the publication of several books devoted to philosophical treatments of various subjects in applied ethics, concentrating first on biomedical ethics and second on business ethics. Virtually every book published in these applied fields prior to 1979 was organized topically; none was developed explicitly in terms of moral principles or ethical theory. Philosophers had by this time been working in areas of applied ethics for several years with

an interest in the connection between theory, principles, practical decision making, and policy. However, in retrospect, it appears that these connections and their problems were not well understood prior to the mid-1980s.

MODELS OF APPLICATION, REASONING, AND JUSTIFICATION

When applied ethics began to receive acceptance in philosophy, it was widely presumed that the "applied" part involves the application of basic moral principles or theories to particular moral problems or cases. This vision suggests that ethical theory develops general principles, rules, and the like, whereas applied ethics treats particular contexts through less general, derived principles, rules, judgments, and the like. From this perspective applied ethics is old morality or old ethical theory applied to new areas. New, derived precepts emerge, but they receive their moral content from the old precepts. Applied work need not, then, generate novel ethical content. Applied ethics requires only a detailed knowledge of the areas to which the ethical theory is being applied (medicine, engineering, journalism, business, public policy, court cases, etc.).

Many philosophers reject this account because it reduces applied ethics to a form of deductivism in which justified moral judgments must be deduced from a preexisting theoretical structure of normative precepts that cover the judgment. This model is inspired by justification in disciplines such as mathematics, in which a claim is shown to follow logically (deductively) from credible premises. In ethics the parallel idea is that justification occurs if and only if general principles or rules, together with the relevant facts of a situation (in the fields to which the theory is being applied) support an inference to the correct or justified judgment(s). In short, the method of reasoning at work is the application of a norm to a clear case falling under the norm.

This deductive model is sometimes said to be a top-down "application" of precepts. The deductive form in the application of a rule is the following:

1. Every act of description A is obligatory. (rule)

2. Act b is of description A. (fact)

Therefore,

3. Act b is obligatory. (applied moral conclusion)

This structure directs attention from particular judgments to a covering level of generality (rules and principles that cover and justify particular judgments) and then to the level of ethical theory (which covers and warrants rules and principles).

This model functions smoothly whenever a fact circumstance can be subsumed directly under a general precept, but it does not adequately capture how moral reasoning and justification proceed in complicated cases. The failure to explain complex moral decision making and innovative moral judgment has led to a widespread rejection of deductivism as an appropriate model for applied ethics. Among the replacements for deductivism as a model of application, two have been widely discussed in the literature: case-based reasoning and reflective equilibrium.

CASE-BASED REASONING (A FORM OF CASUISTRY). This approach focuses on practical decision making about particular cases, where judgments cannot simply be brought under general norms. Proponents are skeptical of principles, rules, rights, and theory divorced from history, circumstances, and experience: One can make successful moral judgments of agents and actions, they say, only when one has an intimate understanding of particular situations and an appreciation of the record of similar situations. They cite the use of narratives, paradigm cases, analogies, models, classification schemes, and even immediate intuition and discerning insight.

An analogy to the authority operative in case law is sometimes noted: When the decision of a majority of judges becomes authoritative in a case, their judgments are positioned to become authoritative for other courts hearing cases with similar facts. This is the doctrine of precedent. Defenders of case-based reasoning see moral authority similarly: Social ethics develops from a social consensus formed around cases, which can then be extended to new cases without loss of the accumulated moral wisdom. As a history of similar cases and similar judgments mounts, a society becomes more confident in its moral judgments, and the stable elements crystallize in the form of tentative principles; but these principles are derivative, not foundational.

In addition to having a history dating from medieval casuistry, the case method, as it is often called, has long been used in law schools and business schools. Training in the case method is widely believed to sharpen skills of legal and business reasoning as well as moral reasoning. One can tear a case apart and then construct a better way of treating similar situations. In the thrust-and-parry classroom setting, teacher and student alike reach conclusions about rights, wrongs, and best outcomes in cases. The objective is to develop a capacity to grasp problems

and to find novel solutions that work in the context: Knowing how to reason and act is more prized then knowing that something is the case on the basis of a foundational rule.

The case method in law has come to be understood as a way of learning to assemble facts and judge the weight of evidence—enabling the transfer of that weight to new cases. This task is accomplished by generalizing and mastering the principles that control the transfer, usually principles at work in the reasoning of judges. Use of the case method in business schools springs from an ideal of education that puts the student in the decision-making role after an initial immersion in the facts of a complex situation. Here the essence of the case method is to present a situation replete with the facts, opinions, and prejudices that one might encounter and to find a way of making appropriate decisions in such an environment.

REFLECTIVE EQUILIBRIUM (A FORM OF COHERENCE THEORY). Many now insist that the relationship between general norms and the particulars of experience is bilateral (not unilateral). Moral beliefs arise both by generalization from the particulars of experience (cases) and by making judgments in particular circumstances by appeal to general precepts. John Rawls's celebrated account of "reflective equilibrium" has been the most influential model of this sort. In developing and maintaining a system of ethics, he argues, it is appropriate to start with the broadest possible set of considered judgments about a subject and to erect a provisional set of principles that reflects them. Reflective equilibrium views investigation in ethics (and theory construction) as a reflective testing of moral principles, theoretical postulates, and other relevant moral beliefs to make them as coherent as possible. Starting with paradigms of what is morally proper or morally improper, one then searches for principles that are consistent with these paradigms as well as one another. Widely accepted principles of right action and considered judgments are taken, as Rawls puts it, "provisionally as fixed points" but also as "liable to revision."

"Considered judgments" is a technical term referring to judgments in which moral beliefs and capacities are most likely to be presented without a distorting bias. Examples are judgments about the wrongness of racial discrimination, religious intolerance, and political conflict of interest. By contrast, judgments in which one's confidence level is low or in which one is influenced by the possibility of personal gain are excluded from consideration. The goal is to match, prune, and adjust considered judgments so that they coincide and are rendered coherent with the premises of theory. That is, one starts with paradigm judgments of moral rightness and wrongness and then constructs a more general theory that is consistent with these paradigm judgments (rendering them as coherent as possible); any loopholes are closed, as are all forms of incoherence that are detected. The resultant action guides are tested to see if they too yield incoherent results. If so, they are readjusted or given up, and the process is renewed, because one can never assume a completely stable equilibrium. The pruning and adjusting occur by reflection and dialectical adjustment, in view of the perpetual goal of achieving reflective equilibrium.

This model demands the best approximation to full coherence under the assumption of a never-ending search for defects of coherence, for counterexamples to beliefs, and for unanticipated situations. From this perspective moral thinking is analogous to hypotheses in science that are tested, modified, or rejected through experience and experimental thinking. Justification is neither purely deductivist (giving general action guides preeminent status), nor purely inductivist (giving experience and analogy preeminent status). Many different considerations provide reciprocal support in the attempt to fit moral beliefs into a coherent unit. This is how we test, revise, and further specify moral beliefs. This outlook is very different from deductivism, because it holds that ethical theories are never complete, always stand to be informed by practical contexts, and must be tested for adequacy by their practical implications.

METHOD AND CONTENT: DEPARTURES FROM TRADITIONAL ETHICAL THEORY

In light of the differences in the models just explored and the enormously diverse literature in applied philosophy it is questionable whether applied ethics has a special philosophical method. Applied philosophers appear to do what philosophers have always done: They analyze concepts, examine the hidden presuppositions of moral opinions and theories, offer criticism and constructive accounts of the moral phenomena in question, and criticize strategies that are used to justify beliefs, policies, and actions. They seek a reasoned defense of a moral viewpoint, and they use proposed moral frameworks to distinguish justified moral claims from unjustified ones. They try to stimulate the moral imagination, promote analytical skills, and weed out prejudice, emotion, misappropriated data, false authority, and the like.

Differences between ethical theory and applied ethics are as apparent over content as over method. Instead of analyzing general terms such as "good", "rationality", "ideals", and "virtues", philosophers interested in applied ethics attend to the analysis of concepts such as confidentiality, trade secrets, environmental responsibility, euthanasia, authority, undue influence, free press, privacy, and entrapment. If normative guidelines are proposed, they are usually specific and directive. Principles in ethical theory are typically general guides that leave considerable room for judgment in specific cases, but in applied ethics proponents tend either to reject principles and rules altogether or to advance precise action guides that instruct persons how to act in ways that allow for less interpretation and discretion. Examples are found in literature that proposes rules of informed consent, confidentiality, conflict of interest, access to information, and employee drug testing.

However, in philosophy journals that publish both applied and theoretical work no sharp line of demarcation is apparent between the concepts and norms of ethical theory and applied ethics. There is not even a discernible continuum from theoretical to applied concepts or principles. The applied/theoretical distinction therefore needs to be used with great caution.

COMPETING THEORIES AND PROBLEMS OF SPECIFICITY

One reason theory and application are merged in the literature is that several different types of ethical theories have been employed in attempts to address practical problems. At least the following types of theories have been explicitly invoked: (1) utilitarianism, (2) Kantianism, (3) rights theory, (4) contract theory, (5) virtue theory, (6) communitarianism, (7) casuistry, and (8) pragmatism. Many proponents of these theories would agree that specific policy and practical guidelines cannot be squeezed from appeals to these philosophical ethical theories and that some additional content is always necessary.

Ethical theories have rarely been able to raise or answer the social and policy questions commonplace in applied ethics. General theories are ill suited for this work, because they address philosophical problems and are not by their nature practical or policy oriented. The content of a philosophical theory, as traditionally understood, is not of the right sort. Philosophical theories are about morality, but they are primarily attempts to explain, unify, or justify morality, not attempts to specify the practical commitments of moral principles in public policy or in particular cases. In applied ethics, ethical theory is often far less important than moral insight and the defense and development of appropriate guidelines suited to a complex circumstance.

Every general ethical norm contains an indeterminacy requiring further development and enrichment to make it applicable in a complex circumstance. To have sufficient content, general theories and principles must be made specific for contexts; otherwise, they will be empty and ineffectual. Factors such as efficiency, institutional rules, law, and clientele acceptance must be taken into account to make them more specific. An ethics useful for public and institutional policies needs to prove a practical strategy that incorporates political procedures, legal constraints, uncertainty about risk, and the like. Progressive specification of norms will be required to handle the variety of problems that arise, gradually reducing dilemmas, policy options, and contingent conflicts that abstract theory and principle are unable to handle.

Some philosophers view this strategy of specification as heavily dependent upon preexistent practices. They maintain that major contributions in philosophical ethics have run from "applied" contexts to "general" theory rather than the reverse. In examining case law and institutional practices, they say, philosophers have learned about morality in ways that require rethinking and modifying general norms of truth telling, consenting, confidentiality, justice, and so forth. To the extent that sophisticated philosophical treatments of such notions are now emerging, they move, not from theory application (including specification), but from practice to theory. Traditional ethical theory, from this perspective, has no privileged position and has more to learn from "applied contexts" than the other way around.

Nonetheless, there are problems with attempts to base applied ethics entirely in practice standards. A practice standard often does not exist within the relevant field, group, or profession. If current standards are low, they could not legitimately determine what the appropriate standards should be. Most moral problems present issues that have to be thought through, not issues to which good answers have already been provided, which explains why many in the professions have turned to philosophers for help in developing professional ethics. Applied philosophers are often most useful to those with whom they collaborate in other fields when practice standards are defective or deficient and a vacuum needs filling by reflection on, criticism of, and reformulation of moral viewpoints or standards.

See also Abortion; Affirmative Action; Business Ethics; Communitarianism; Deontological Ethics; Environmental Ethics; Justice; Metaethics; Pragmatism; Rawls, John; Rights; Utilitarianism; Virtue Ethics.

Bibliography

Beauchamp, T. L. "On Eliminating the Distinction between Applied Ethics and Ethical Theory." *The Monist* 67 (1984): 514–31.

Brock, D. W. "Truth or Consequences: The Role of Philosophers in Policy-Making." *Ethics* 97 (1987): 786–791.

Caplan, A. L. "Ethical Engineers Need Not Apply: The State of Applied Ethics Today." *Science, Technology, and Human Values* 6 (Fall 1980): 24–32.

DeGrazia, D. "Moving Forward in Bioethical Theory: Theories, Cases, and Specified Principlism." *Journal of Medicine and Philosophy* 17 (1992): 511–539.

Feinberg, J. *The Moral Limits of the Criminal Law*. 4 vols. New York: Oxford University Press, 1984–1987.

Fullinwider, R. K. "Against Theory, or: Applied Philosophy—A Cautionary Tale." *Metaphilosophy* 20 (1989): 222–234.

Gert, B. "Licensing Professions." *Business and Professional Ethics Journal* 1 (1982): 51–60.

Gert, B. "Moral Theory and Applied Ethics." *The Monist* 67 (1984): 532–548.

Jonsen, A., and S. Toulmin. *The Abuse of Casuistry: A History of Moral Reasoning*. Berkeley: University of California Press, 1988.

MacIntyre, A. "Does Applied Ethics Rest on a Mistake?" *The Monist* 67 (1984): 498–513.

MacIntyre, A. "What Has Ethics to Learn from Medical Ethics?" *Philosophic Exchange* 2 (1978): 37–47.

Noble, C. "Ethics and Experts." *Hastings Center Report* 12 (June 1982): 7–10, with responses by four critics.

Professional Ethics 1, 1–2 (Spring–Summer 1992). Special issue on applied ethics.

Rawls, J. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Regan, T., ed. *Matters of Life and Death*. 3rd ed. New York: McGraw-Hill, 1993.

Reich, Warren, ed. *Encyclopedia of Bioethics*. 2nd ed. New York: Macmillan, 1995.

Richardson, H. "Specifying Norms as a Way to Resolve Concrete Ethical Problems." *Philosophy and Public Affairs* 19 (1990): 279–310.

Singer, P. *Practical Ethics*. 2nd ed. New York: Cambridge University Press, 1993.

Winkler, E. R., and J. R. Coombs, eds. *Applied Ethics: A Reader*. Oxford: Blackwell, 1993.

Tom L. Beauchamp (1996)

A PRIORI AND A POSTERIORI

The distinction between the a priori and the a posteriori has always been an epistemological one; that is to say, it has always had something to do with knowledge. The terms *a priori* and *a posteriori* are Scholastic terms that have their origin in certain ideas of Aristotle; but their use has been considerably extended in the course of history, and their present use stems from the meaning given to them by Immanuel Kant. The terms literally mean "from what is prior" and "from what is posterior." According to Aristotle, *A* is prior to *B* in nature if and only if *B* could not exist without *A; A* is prior to *B* in knowledge if and only if we cannot know *B* without knowing *A*. It is possible for these two senses of "prior" to have an application in common; substance, for example, is prior to other things in both of these senses and in others. It follows that to know something from what is prior is to know what is, in some sense, its cause. Aristotle believed that it is possible to demonstrate a causal relationship by means of a syllogism in which the term for the cause is the middle term. Hence, to know something in terms of what is prior is to know it in terms of a demonstrable causal relationship. To know something from what is posterior, on the other hand, can involve no such demonstration, since the knowledge will be inductive in form.

The transition to Kant's conception of the matter is evident in Gottfried Wilhelm Leibniz. According to the latter, to know reality a posteriori is to know it from what is actually found in the world, that is, by the senses, by the effects of reality in experience; to know reality a priori is to know it "by exposing the cause or the possible generation of the definite thing" (*Nouveaux Essais*, Bk. III, Ch. 3). It is also possible to speak of a priori proofs. As a general consequence of this, Leibniz could distinguish between "truths a posteriori, or of fact," and "truths a priori, or of reason" (ibid., Bk. IV, Ch. 9); for a priori truths can be demonstrated in terms of their being based on identical propositions, while a posteriori truths can be seen to be true only from experience. Thus the distinction between the a posteriori and the a priori comes to be a distinction between what is derived from experience and what is not, whether or not the notion of the a priori also has the notion of demonstration in terms of cause or reason associated with it. Such is the distinction in Kant, and it has remained roughly the same ever since. Since in Kant there is no simple opposition between sense experience and reason (there being also the understanding), it is not possible to express the distinction he laid down as one

between what is derived from experience and what is derived from reason.

The distinction, then, is roughly equivalent to that between the empirical and the nonempirical. Kant also connected it with the distinction between the necessary and the contingent, a priori truths being necessary and a posteriori truths contingent. But to assume without further argument that the two distinctions coincide in their application is to assume too much. The same is true of the distinction between the analytic and the synthetic; this too cannot be assimilated without argument to that between the a priori and a posteriori. Whether or not these distinctions coincide in their applications, they certainly cannot have the same meaning. The distinction between the a priori and a posteriori is an epistemological one; it is certainly not evident that the others are.

THE DISTINCTION APPLIED TO CONCEPTS

The distinction between the a priori and the a posteriori has been drawn not only in connection with truths or propositions but also in connection with concepts. Indeed, some truths are doubly a priori; not only is their truth knowable independently of experience but the concepts that they involve are similarly independent of experience. The distinction between a posteriori and a priori concepts may seem a perspicuous one, for it may be thought to be a distinction between concepts that we derive from experience by building them up therefrom and concepts that we have independently of experience. It has sometimes been said also that the latter concepts are innate ideas, with which we are born, so that we have no need to acquire them. But the question whether ideas are innate or acquired seems to be one of psychology, as is the question how we acquire ideas if we do. The distinction under consideration, being an epistemological one, has no direct connection with psychology. A concept that is independent of experience may or may not be innate; and although it cannot be acquired directly from experience, it may still be that experience is in some way a necessary condition of our having the concept. What then does it mean to say that a concept is independent of experience? The answer must be in terms of the validation of the concept.

It may be assumed for present purposes that a concept is what is meant by the corresponding term (although this may not be a fully adequate view and bypasses the question whether concepts are independent of words). To have a concept will thus at least be to understand the corresponding term. Perhaps, then, an a posteriori concept is one expressed by a term understandable purely in terms of experience, and an a priori concept one that does not satisfy this condition. The point has sometimes been made by saying that an a posteriori, or empirical, concept or term is one that is cashable in terms of sense experience. This is of course a metaphor, and what it means is that the meaning of empirical terms can be given by definitions that must ultimately depend on ostensive definitions only. Ostensive definitions are those which provide the definition of a term by a direct confrontation with experience. To define a term ostensively it is necessary only to repeat the expression together with some form of pointing to the object or phenomenon in question. It is highly questionable, however, whether any performance of this kind could ever constitute definition as such. For the meaning of a word to be taught in this way there would have to be (as Ludwig Wittgenstein in effect pointed out at the beginning of his *Philosophical Investigations*) a previous understanding that the noise made was a word in a language and in a language of a definite sort. Furthermore, it would have to be understood what sort of term was being defined—whether it was descriptive and, if so, what range of phenomena it was being used to describe. If all this must be understood, it can scarcely be said that the term in question is defined purely by reference to sense experience.

Nevertheless, there is some distinction to be made here. Even if such terms as "red" cannot be defined purely by reference to experience, they could not be understood fully without experience, for example, by someone who does not possess and never has possessed sight. There is a sense in which the blind *can*, up to a point, understand terms such as "red," in that they can know that red is a color and even a color of a certain sort related to other colors in certain ways. But since they cannot know when to apply the term in fact, there is an obvious sense in which they do not have a full understanding of it—and the same applies to the notion of color itself. A posteriori terms and concepts may thus be defined as those that directly require our having experience in order for us to apply them or those that can only be fully understood by reference to terms that directly require our having experience to apply them. Whether or not a creature without experience could ever come to have a concept such as, for example, validity, it is clear that being able to apply the concept does not directly require experience. This may afford the basis of a distinction between a posteriori and a priori concepts. There may be various views about a priori concepts, concerning, for example, whether they are to be restricted to concepts of, or concepts involved in, mental operations on a posteriori concepts. Empiricists

have in general held that the only a priori concepts are those that express relations of ideas. The field is thus restricted to the concepts of logic and mathematics.

THE DISTINCTION APPLIED TO PROPOSITIONS

In a sense, the distinction between concepts presupposes the distinction between propositions, since concepts can be applied only in propositions. According to the rough criteria already mentioned, an a priori proposition will be one whose truth is knowable independently of experience. It may be questioned, however, whether there are any truths that can be known if the subject has no experiences whatever. Hence, the matter is better put in terms of the validation of the proposition in question, in terms of its verification or falsification. It has sometimes been suggested that a proposition is a priori if its truth is ascertainable by examination of it alone or if it is deducible from such propositions. An a priori proposition would thus be one that provides its own verification; it is true in itself. This account is too restrictive, since there may be propositions whose truth is ascertainable by argument that makes no reference to empirical matters of fact, but that may not be deducible from any propositions of the kind previously mentioned. That is to say, there may be circumstances in which it is possible to validate propositions by argument that makes no reference to matters of fact discoverable by experience. Empiricists have generally denied this, but the possibility of what Kant called "transcendental arguments" cannot be so lightly dismissed. Aristotle's argument for the truth of the principle of contradiction would be a case in point, namely, that a denial of it already presupposes it.

On the other hand, to say simply that a priori propositions are those whose truth can be discovered without reference to experience is too wide a definition. For it may be argued that the terms in which many such propositions are expressed could only be fully understood by reference to experience. A proposition may be a priori without its involving terms that are without exception a priori. It was for this reason that Kant distinguished between a priori and *pure* a priori judgments; only in the latter are all the terms a priori. In view of this, an a priori proposition may be defined as one whose truth, given an understanding of the terms involved, is ascertainable by a procedure that makes no reference to experience. The validation of a posteriori truths, on the other hand, necessitates a procedure that does make reference to experience.

CAN ANALYTIC PROPOSITIONS BE A POSTERIORI? It has already been mentioned that Kant superimposed upon the a priori–a posteriori distinction the distinction between the analytic and the synthetic. There are difficulties involved in defining this latter distinction, but for present purposes it is necessary to note that Kant assumed it impossible for analytic judgments to be a posteriori. He does this presumably on the grounds that the truth of an analytic judgment depends upon the relations between the concepts involved and is ascertainable by determining whether the denial of the judgment gives rise to a contradiction. This latter procedure is surely one that makes no reference to experience. Kant is clearly right in this. As already seen, it is not relevant to object that since analytic judgments, propositions, or statements need not involve purely a priori terms, evaluation of the truth of some analytic propositions will involve reference to experience; for in determining whether a proposition is a priori, it is necessary to take as already determined the status of the terms involved. It is similarly irrelevant to maintain that it is sometimes possible to come to see the truth of an analytic proposition through empirical means. It may be possible, for example, for a man to realize the truth of "All bachelors are unmarried men" as an analytic proposition as a consequence of direct experience with bachelors. But this consequence will be an extrinsic one. That is to say that while the man may attain this insight in this way, it will be quite accidental; the validity of the insight does not depend upon the method by which it is acquired. That is why the definition of an a priori proposition or statement involves the idea that its truth must be ascertainable without reference to experience. As long as a nonempirical procedure of validation exists, the proposition in question will be a priori, whether or not its truth is always ascertained by this procedure. It is quite impossible, on the other hand, for an a posteriori proposition to be validated by pure argument alone.

MUST A POSTERIORI PROPOSITIONS BE CONTINGENT? Given that all analytic propositions are a priori, it is a further question whether all synthetic propositions must be a posteriori. This is a hotly debated question, with empiricists maintaining that they must be. But first it is necessary to consider the relation between the a priori–a posteriori dichotomy and the necessary–contingent one.

Kant certainly associated the a priori with the necessary, and there is a prima facie case for the view that if a proposition is known a posteriori, its truth must be contingent. For how can experience alone tell us that some-

thing must be so? On the other hand, it might be maintained that we can learn inductively that a connection between characteristics of things holds as a matter of necessity. Some philosophers maintain that natural laws represent necessary truths, and they do not all think this incompatible with the view that natural laws can be arrived at through experience. What is sometimes called intuitive induction—a notion originating in Aristotle—is also something of this kind; we see by experience that something is essentially so and so. An even greater number of philosophers would be willing to assert that, in *some* sense of the word "must," experience can show us that something must be the case. Certainly the "must" in question is not a logical "must," and empiricists have tended to maintain that all necessity is logical necessity. This, however, is just a dogma. It seems plausible to assert that an unsupported body must in normal circumstances fall to the ground.

Yet it must be admitted that the normal philosophical conception of necessity is more refined than this, and to say that an unsupported body must in normal circumstances fall to the ground need not be taken as incompatible with saying that this is a contingent matter. Similarly, there is an important sense in which natural laws are contingent; they are about matters of fact. If we also think of them as necessary, the necessity in question stems from the conceptual framework into which we fit them. It is possible to conceive of empirical connections in such a way that, within the framework of concepts in which we place them, they are treated as holding necessarily. It is still a contingent matter whether the whole conceptual framework has an application. If propositions expressing such connections are a priori, it is only in a relative sense.

MUST A PRIORI PROPOSITIONS BE NECESSARY? It seems at first sight that there is no necessity for nonempirical propositions to be necessary, or rather that it is possible to construct propositions which, if true, must be true a priori, while they apparently remain contingent. These are propositions that are doubly general. They may be formalized in such a way as to contain both a universal and an existential quantifier, for example, $(x) \cdot \exists y \cdot \phi xy$. Such propositions have been called by J. W. N. Watkins (following Karl Popper) "all and some propositions." Because they have this kind of double generality, they are both unverifiable and unfalsifiable. The element corresponding to the universal quantifier makes them unverifiable; that corresponding to the existential quantifier makes them unfalsifiable. Under the circumstances they can hardly be said to be empirical. An example of this kind of proposition is the principle of universal causality, "Every event has a cause," which is equivalent to "For every event there is some other event with which it is causally connected." It has been claimed by some philosophers, for instance, G. J. Warnock, that this proposition is vacuous, since no state of affairs will falsify it. But the most that can be claimed in this respect is that no *particular* state of affairs *which can be observed* will falsify it. It is clearly not compatible with any state of affairs whatever, since it is incompatible with the state of affairs in which there is an event with no cause. It remains true that it is impossible to verify that an event has no cause.

Watkins does not claim that the proposition is necessary, although the principle of causality has been held by many, for instance, Kant, to be an example of a necessary truth, and it could no doubt be viewed as such. But it is also possible to treat it as a contingent truth, one that holds only in the contingency of every event being causally determined. How we could know that such a contingency held is a further question. It is clear that nothing that we could observe would provide such knowledge. Such propositions certainly could not be *known* a posteriori; if true, they must be known a priori if they are to be known at all. The difficulty is just this—how *are* they to be known at all? Thus, it may be better to distinguish between a priori propositions and nonempirical propositions of this kind. A priori propositions are those which can be known to be true and whose truth is ascertainable by a procedure that makes no reference to experience; nonempirical propositions of the kind in question are not like this, for their truth is, strictly speaking, not ascertainable at all. If we accept them, it must be as mere postulates or as principles whose force is regulative in some sense.

This does not exclude the possibility that there are other propositions whose truth can be ascertained by a nonempirical procedure but that are less than necessary. It has been argued by J. N. Findlay that there are certain propositions asserting connections between concepts that are only probable, as opposed to the commonly held view that all connections existing among concepts are necessary. He maintains that our conceptual systems may be such that there are connections between their members that are by no means analytic; the connections do not amount to entailments. Perhaps something like the Hegelian dialectic is the prototype of this. Findlay argues, for example, that if one has likings, there is the presumption that one will like likings of this sort; on this sort of basis one could move toward the notion of a community of ends. It is difficult to speak more than tentatively here. Given, however, that the propositions stating these con-

ceptual connections are, if true, then true a priori (as they surely must be), it is not clear that it is necessary to claim only that what one knows in relation to them is probable. Certainly the connections do not constitute entailments; but this of itself does not mean that what one knows is only probable. The fact that the argument for a certain position is not a strictly deductive one does not mean that the position cannot be expressed by truths that are necessary and can be known to be so. For the argument may justify the claim to such knowledge in spite of the fact that the argument is not deductively valid in the strict sense. If such a necessary proposition does not seem to have universal application, this may be due to the fact that it holds under certain conditions and that its necessity is relative to these conditions. This was Kant's position over the principle of universal causality. He held that the principle that every event has a cause is necessary only in relation to experience. If propositions of this sort lack absolute necessity, they need not lack necessity altogether. The tentative conclusion of this section is that while some propositions may in a certain sense be both nonempirical and contingent, it nevertheless remains true that if a proposition is known a priori, it must be necessarily true in some sense or other.

MUST A PRIORI PROPOSITIONS BE ANALYTIC? It has been suggested in the previous section that there may be a priori propositions that are not analytic. They depend for their validation on a priori argument but cannot be given a deductive proof from logical truths. The question of the synthetic a priori is one of the most hotly debated topics in philosophy and has, indeed, been so ever since Kant first stated the issues explicitly. Empiricists have always vehemently denied the possibility of such truths and have even tried to show that a proposition that is a priori must be analytic by definition. Most attempts of this sort rest on misconceptions of what is meant by these terms.

Kant's synthetic a priori. Kant claimed that synthetic a priori truths were to be found in two fields—mathematics and the presuppositions of experience or science—although he denied that there was a place for them in dogmatic metaphysics. He maintained that although mathematics did contain some analytic truths (since there were propositions which summed up purely deductive steps), the main bulk of mathematical truths were synthetic a priori; they were informative, nonempirical, and necessary, but not such that their denial gave rise to a contradiction. These characteristics were in large part due to the fact that mathematical knowledge involved intuitions of time (in the case of arithmetic) and space (in the case of geometry). Kant's conception of arithmetic has not found much support, and his view of geometry has often been considered to have been undermined by the discovery of non-Euclidean geometries. It is doubtful, however, whether the situation is quite so simple as this, for what Kant maintained was that an intuition of space corresponding to Euclidean geometry was necessary at any rate *for creatures with sensibility like ours*. That is to say, what we perceive of the world must conform to Euclidean geometry, whether or not it can be conceived differently in abstraction from the conditions of perception. Whether or not this is true, it is not obviously false.

The main attack on the Kantian view of arithmetic, and thereafter on that view of other branches of mathematics, came from Gottlob Frege and from Bertrand Russell and Alfred North Whitehead. Frege defined an analytic proposition as one in the proof of which one comes to general logical laws and definitions only; and he attempted to show that arithmetical propositions are analytic in this sense. The crucial step in this program is Frege's definition of "number" roughly in terms of what Russell called one-to-one relations. (Russell himself gave a parallel definition in terms of similarity of classes.) Given Frege's definition of number, arithmetical operations had to be expressed in terms of the original definition. It is at least an open question whether this attempt was successful. The definition has been accused of being circular and/or insufficient. This being so, the most that can be claimed is that arithmetic, while not reducible to logic, has a similar structure. Nevertheless, Gödel's proof that it is impossible to produce a system of the whole of formal arithmetic that is both consistent and complete may be taken to cast doubt even on this claim. At all events, the exact status of arithmetical truths remains arguable.

Other synthetic a priori truths claimed by Kant were the presuppositions of objective experience. He tried to demonstrate that the truth of such propositions as "Every event has a cause" is necessary to objective experience. These propositions indeed express the necessary conditions of possible experience and of empirical science. As such, their validity is limited to experience, and they can have no application to anything outside experience, to what Kant called "things-in-themselves." According to Kant, these principles—which are of two kinds, constitutive or regulative in relation to experience—are ultimately derived from a list of a priori concepts or categories, which he claims to derive in turn from the traditional logical classification of judgments. These principles, in a form directly applicable to empirical

phenomena, are also established by transcendental arguments. In the "Second Analogy" of the *Critique of Pure Reason*, for example, Kant sought to show that unless objective phenomena were irreversible in time, and therefore subject to rule, and therefore due to causes, it would be impossible to distinguish them from merely subjective phenomena. Causality is therefore a condition of distinguishing phenomena as objective at all. The cogency of this position depends upon the acceptability of the arguments, and it is impossible to examine them here. It is to be noted, however, that what the arguments seek to show is that certain necessary connections between concepts must be accepted if we are to give those concepts any application. The connection between the concepts of "objective event" and "cause" is not an analytic one, but it is a connection that must be taken as obtaining if the concepts are to have any application to empirical phenomena.

Another instance of this kind of situation, perhaps more trivial, can be seen in such a proposition as "Nothing can be red and green all over at the same time in the same respect." This proposition has sometimes been classified as empirical, sometimes as analytic; but it has been thought by empiricists a more plausible candidate for synthetic a priori truth than any of Kant's examples. There is clearly some kind of necessity about this proposition. It may be possible for something to *appear* red and green all over, but to suggest that something might *be* red and green all over or that one might produce examples of such a thing has little plausibility; for in some sense red excludes green. The question is, In what sense? Since "red" does not mean "not-green" and cannot be reduced to this (for terms such as "red" and "green" do not seem capable of analysis), the proposition under consideration cannot, strictly speaking, be analytic. How can red and green exclude each other without this being a logical or analytic exclusion? It is not merely a contingent exclusion, since it is clearly impossible to produce something that is red and green all over (shot silk, for example, although it appears so, does not conform to the conditions of being two-colored all over), and we cannot imagine what such a thing would be like.

It may be suggested that red and green are different determinates of the same determinable—color. We distinguish colors and use different terms in order to do so. To allow, then, that something might be described by two such terms at the same time would be to frustrate the purposes for which our system of color classification was devised. However, this may sound too arbitrary. After all, given two colors that do in fact shade into each other, we might be less reluctant to allow that something might be both of them at once. It is no accident that we distinguish colors as we do. For creatures of our kind of sensibility, as Kant would put it, colors have a definite structure; it is natural to see them in certain ways and to conceive of them accordingly. We then fit them to a conceptual scheme that reflects those distinctions. If we will not allow that something may be red and green all over, it is because the mutual exclusion of red and green is a necessary feature of our scheme of color concepts. Yet the whole scheme has application to the world only because we see colors as we do.

THE RELATIVE AND ABSOLUTE A PRIORI. Because of the empirical preconditions for our scheme of color concepts, if we maintain that we can know a priori that something cannot be red and green all over, it cannot be absolutely a priori. For the truth that something that is red cannot also be green at the same time and in the same respect can scarcely be said to be ascertainable without any reference whatever to experience. The same is true of the principle of universal causality discussed earlier. It might be maintained that the truth that every event has a cause is necessary because "cause" and "event" are so definable that there is an analytic connection between them (implausible as this may be in fact). In that case the proposition in question would be true in all possible worlds (to use a Leibnizian phrase), since its truth would not depend on what is. In a world in which no events occurred, it would be true, in this view, that every event (if there were any) would have a cause. We can know the truth of this proposition absolutely a priori. However, if the principle is not analytic (and it is clearly not, in its ordinary interpretation) but is still thought to be necessary, this can be so only because the connection between cause and event is necessary to our conception of the world as we see it. Similarly, the mutual exclusion of red and green is necessary to our conception of colors as we see them. These propositions are not true in all possible worlds, and while their truth can be known a priori, it is not known absolutely a priori.

On the other hand, the so-called laws of thought, such as the principle of contradiction, while not analytic, must again be known absolutely a priori, whatever the kind of necessity they possess. The truth of the principle of contradiction is necessary to the possibility of thought in general, including the thought of the principle itself. It is not possible even to deny the principle without presupposing it. It cannot be maintained that its truth is in any way ascertainable by a procedure that makes reference to experience. Its truth is a necessity of thought, not of expe-

rience, and is not relative to experience. Hence it may be said to be known absolutely a priori.

Of those propositions that are absolutely a priori there are two kinds—analytic truths and the principles of logic themselves. (It is perhaps not surprising that these have sometimes been classified together, even if wrongly so.) On the other hand, there are some truths that are necessary but known only relatively a priori—truths such as the principle of causality and the principle of the incompatibility of colors. Finally, of course, there is that large class of truths which can only be known a posteriori. But for philosophers these are naturally much less interesting than truths of the first two kinds—those which are a priori in some sense or other. And over these there is still much argument.

See also Analytic and Synthetic Statements; Aristotle; Empiricism; Frege, Gottlob; Gödel, Kurt; Gödel's Theorem; Hegelianism; Innate Ideas; Kant, Immanuel; Knowledge, A Priori; Laws of Thought; Leibniz, Gottfried Wilhelm; Logic, History of; Mathematics, Foundations of; Popper, Karl Raimund; Propositions; Russell, Bertrand Arthur William; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

For Kant's distinction between a priori and a posteriori, see the Norman Kemp Smith translation of the *Critique of Pure Reason* (London, 1953), especially the introduction but also the chapters on the aesthetic and the analytic of principles. For the precedents to Kant's distinction, see Aristotle, *Posterior Analytics*, especially Bk. I.2; Gottfried Leibniz, *Nouveaux Essais*, translated by A. G. Langley as *New Essays concerning Human Understanding* (Chicago: Macmillan, 1916), especially III.3 and IV.9. See also Arthur Pap, *Semantics and Necessary Truth* (New Haven, CT: Yale University Press, 1958), Pt. 1.

For the application of the distinction to concepts or terms, see H. H. Price, *Thinking and Experience* (London: Hutchinson, 1953); for criticisms of the notion of ostensive definition, see the opening sections of Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1953); and Peter Geach, *Mental Acts* (London: Routledge and Paul, 1957), especially sections 6–11; and for discussion of the a priori in relation to analyticity, see G. H. Bird, "Analytic and Synthetic," *Philosophical Quarterly* 11 (1961): 227–237; L. J. Cohen, *The Diversity of Meaning* (London: Methuen, 1962), Chs. 6 and 10; and Ch. 5 of the work by Arthur Pap cited above.

For discussion of nonempirical propositions that are not necessary, see J. N. Findlay, *Values and Intentions* (London: Allen and Unwin, 1961); two articles by J. W. N. Watkins: "Between Analytic and Empirical," *Philosophy* 32 (1957): 112–131, and "Confirmable and Influential Metaphysics," *Mind* 67 (1958): 344–365; and compare G. J. Warnock, "Every Event Has a Cause," in *Logic and Language* (Oxford: Blackwell, 1953), Vol. II.

For Kant's views on the possibility of the synthetic a priori, see the Kemp Smith translation cited above. For discussions of examples, see: D. W. Hamlyn, "On Necessary Truth," *Mind* 70 (1961): 514–525; S. N. Hampshire, "Identification and Existence," in *Contemporary British Philosophers* (3d Series) (London, 1956); D. J. O'Connor, "Incompatible Properties," *Analysis* 15 (1955): 109–117; D. F. Pears, "Incompatibility of Colours," in *Logic and Language* (Oxford: Blackwell, 1953), Vol. II; and compare Aristotle's discussion of the principle of contradiction in *Metaphysics* IV.4.

OTHER RECOMMENDED TITLES

Boghossian, P., and C. Peacocke, eds. *New Essays on the A Priori*. Oxford: Clarendon Press, 2000.

BonJour, L. *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge, U.K.: Cambridge University Press, 1998.

Casullo, Albert. *A Priori Justification*. New York: Oxford University Press, 2003.

Casullo, Albert, ed. *A Priori Knowledge*. Aldershot, U.K.: Ashgate, 1999.

Coffa, J. A. *The Semantic Tradition from Kant to Carnap*. Cambridge, U.K.: Cambridge University Press, 1991.

DePaul, M. R., and W. Ramsey, eds. *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield, 1998.

Foley, Richard. *A Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Foley, Richard. *Working without a Net*. New York: Oxford University Press, 1993.

Fumerton, Richard. "A Priori Philosophy after an A Posteriori Turn." *Midwest Studies in Philosophy* 23 (1999): 21–33.

Gendler, T. S., and J. Hawthorne, eds. *Conceivability and Possibility*. Oxford: Oxford University Press, 2002.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Hanson, P., and B. Hunter, eds. *Return of the A Priori*. Calgary, AB: University of Calgary Press, 1992.

Hawthorne, John P. "Deeply Contingent A Priori Knowledge." *Philosophy and Phenomenological Research* (September 2002): 247–269.

Kitcher, Philip. *The Nature of Mathematical Knowledge*. New York: Oxford University Press, 1983.

Kripke, Saul A. "Identity and Necessity." In *Naming, Necessity, and Natural Kinds*, edited by S. P. Schwartz, 66–101. Ithaca, NY: Cornell University Press, 1977.

Kripke, Saul A. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Kripke, Saul A. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press, 1982.

Moser, Paul, ed. *A Priori Knowledge*. Oxford: Oxford University Press, 1987.

Moser, Paul. *Knowledge and Evidence*. Cambridge, U.K.: Cambridge University Press, 1989.

D. W. Hamlyn (1967)
Bibliography updated by Benjamin Fiedor (2005)

A PRIORI KNOWLEDGE

See *A Priori and A Posteriori; Innate Ideas; Knowledge, A Priori; Rationalism*

AQUINAS, ST. THOMAS

See *Thomas Aquinas, St.*

ARABIC PHILOSOPHY

See *Islamic Philosophy*

ARCESILAUS

(316/315–241/240 BCE)

Arcesilaus was born in Pitane, a Greek city on the coast of Asia Minor. In Athens, after a period of study with Theophrastus—Aristotle's successor as head of the Peripatos—he joined the Academy, Plato's school, which was then dominated by Crantor, Polemon, and Crates. Arcesilaus succeeded Crates, Polemon's successor, as head of the Academy and was responsible for the school's turn to skepticism. From this point, the skeptical examination of other schools' theories replaced the elaboration of its own positive doctrines as the Academy's principal occupation. This change in the Academy's direction is recognized in the ancient tradition that credits Arcesilaus with founding the second or Middle Academy, which replaced the first or Old Academy and gave way in turn to the third or New Academy of Carneades.

Like Socrates before him and Carneades after him, Arcesilaus wrote nothing, but made his mark in face to face philosophical argument. His practice was not to present views of his own, but instead to invite his interlocutors to put forward their views, which he then subjected to rigorous scrutiny. His method was dialectical: He put questions to his interlocutors from the answers to which he aimed to deduce conclusions at odds with their positions. The effect was to uncover difficulties internal to the interlocutors' positions without committing him to a position of his own. These arguments were conceived by Arcesilaus and his Academic followers as their contribution to argument on both sides of the question, which they regarded as the best way to discover the truth—their ultimate aim. The resemblance to Socrates is unmistakable and was much emphasized by the Academics.

Their principal target was Stoic epistemology. According to Zeno of Citium, the founder of Stoicism and an older contemporary of Arcesilaus's, it is possible for human beings to free themselves entirely from opinion—that is, false or insecure belief—and to attain the kind of knowledge that qualifies as wisdom. In the Socratic tradition, Zeno held that wisdom was identical with virtue and as such the one necessary and sufficient condition for happiness. A necessary condition for knowledge on the Stoic view was the existence of cognitive impressions (*kataleptikai phantasiai*). Each of these is a perceptual impression that arises in conditions which both ensure that, by capturing its objects with perfect accuracy, it is true while at the same time imparting to it a character that belongs only to impressions that arise in this way and which human beings can discriminate. According to Stoic epistemology, all knowledge depends in one way or another on cognitive impressions, which is why the cognitive impression is the school's criterion of truth. By restricting one's assent (in the sphere of perception) to impressions with this character, one can avoid ever assenting to a false perceptual impression. If further conditions are satisfied, one can avoid error altogether.

Arcesilaus and the Academics argued that, on any plausible account of it, the character allegedly proper to cognitive impressions was not in fact confined to impressions produced in the specified truth-guaranteeing way, but also belonged to false impressions. As a result, the former, though true, are indistinguishable from the latter and therefore unable to serve as a criterion. It follows on Stoic assumptions about knowledge that nothing can be known. This is the first of the two propositions most closely associated with ancient skepticism. The second—that one ought to suspend judgment—Arcesilaus deduced from the first, together with the Stoic insistence that wisdom is incompatible with mere opinion. Assent to a noncognitive impression (or an impression that does not stand in the proper relation to cognitive impressions), automatically results in opinion, so that, in the absence of cognitive impressions, a wise person can avoid opinion only by suspending judgment entirely.

On a strictly dialectical interpretation of his arguments, Arcesilaus did nothing more than present his Stoic opponents with a set of difficulties. On this view, it was their task either to resolve the difficulties or to abandon or modify the position that had given rise to them. Some ancient authorities held that Arcesilaus's arguments against the possibility of knowledge and in favor of suspension of judgment had implications only for the Stoa and did not prevent him from espousing a form of dog-

matic Platonism within the Academy. But according to another, better-founded tradition in the Academy itself, Arcesilaus agreed with Zeno that opinion is utterly alien to wisdom and that it is a grave failing—indeed a sin—to allow assent to run ahead of knowledge. But, according to this tradition, the lesson he drew from the difficulties that he had uncovered in the Stoic position among others, was that he and his opponents were not, or not yet, in a position to assent, secure in the conviction that they were in possession of the truth. In these conditions, suspension of judgment and continued open-minded inquiry were indicated. The skepticism characterized by this attitude was a matter of intellectual honesty and prudence, rather than convinced adherence to the skeptical proposition that nothing can be known. And Arcesilaus, it is told, was careful to maintain that one could not even know that one could not know anything.

The Stoics responded to Arcesilaus by arguing that, if nothing can be known and people are therefore obliged to withhold assent, life becomes impossible, as there can no basis for judgment without a criterion nor any possibility of action without assent. Arcesilaus's answer appears to have been an extension of his first dialectical arguments, for it aimed to show that Stoic epistemology and moral psychology had the resources to explain how a human being may proceed in the absence of cognitive impressions and act without assent. In these conditions, one will be guided by what is reasonable, the notion that the Stoics had used to explain how the wise will act when certainty is not available. It will, for instance, be reasonable to expect a successful voyage if the weather is fair, the crew skilled and so on. Action, on the other hand, requires only that an impression elicit an impulse, which the Stoics used to explain the behavior of nonrational animals and human children, who lack the power of assent.

Arcesilaus's explanation of how action is possible without assent, at least in the form in which it has survived, is sketchy, and it is not clear that it can do justice to the concerns that moved the Stoics and other philosophers to develop their theories in the first place. It plainly does not have the independent appeal of Carneades's theory of probability or his suggestion that assent, as the Stoics conceived it, could be replaced either by qualified assent or a way of following or using impressions that did not entail assent. Nonetheless Arcesilaus's proposals marked the beginning of a long tradition of defending the skepticism as a way of life, of which Carneades's probabilism and Pyrrhonism were later examples.

The example of Arcesilaus continued to inspire members of the Academy until the end of the school's history and thereafter Pyrrhonists, who regarded New Academics such as Carneades as apostates from the true skeptical way but acknowledged a kinship with Arcesilaus.

See also Ancient Skepticism; Carneades; Stoicism; Zeno of Citium.

Bibliography

WORKS GIVING TEXTS OF ARCESILAUS

Long, A. A., and D. N. Sedley, eds. *The Hellenistic Philosophers.* 2 vols. Chapters 68–70. Cambridge, U.K.: Cambridge University Press, 1987.

Mette, H. J. "Zwei Akademiker heute: Krantor von Soloi und Arkesilaos von Pitane." *Lustrum* 26 (1984): 7–94.

STUDIES OF ARCESILAUS

Cooper, John. "Arcesilaus: Socratic and Skeptic." in *Knowledge, Nature and the Good*, by John Cooper. Princeton, NJ: Princeton University Press, 2004.

Couissin, P. "L'Origine et l'evolution de l'epoche." *Revue des études grecques* 42 (1929): 373–397.

Couissin, P. "The Stoicism of the New Academy." In *The Skeptical Tradition*, edited by Myles Burnyeat, 31–63. Berkeley: University of California Press, 1983.

Long, A. A. "Diogenes Laertius, Life of Arcesilaus." *Elenchos* 7 (1986): 429–449.

Schofield, Malcolm. "Academic Epistemology." In *The Cambridge History of Hellenistic Philosophy*, edited by Keimpe Algra, Jonathan Barnes, Jaap Mansfeld, and Malcolm Schofield. Cambridge, U.K.: Cambridge University Press, 1999.

Striker, Gisela. "Sceptical Strategies." In *Doubt and Dogmatism: Studies in Hellenistic Epistemology*, edited by Malcolm Schofield, Myles Burnyeat, and Jonathan Barnes. Oxford: Oxford University Press, 1980.

James Allen (2005)

ARCHĒ

The Greek term *archē* refers to the original stuff from which the world came to be, according to pre-Socratic philosophers. In his *Metaphysics* Aristotle explains:

> Of the first philosophers, the majority thought the sources [*archai*, plural] of all things were found only in the class of matter. For that of which all existing things consist, and that from which they come to be first and into which they perish last—the substance continuing but changing in its attributes—*this*, they say, is the element and this the source [*archē*] of existing

things. Accordingly they do not think anything either comes to be or perishes, inasmuch as that nature is always preserved. … For a certain nature always exists, either one or more than one, from which everything else comes to be while this is preserved. All, however, do not agree on the number and character of this source, but Thales, the originator of this kind of theory, says it is water.…

(*METAPHYSICS* 983B 6–21)

Aristotle seems to use the term *archē* to refer to several different notions that he holds are all part of the pre-Socratics' conception: (1) a primeval chaos in which only one element (or one set of elements) exists; (2) the primeval element that constituted the primitive state, from which all the bodies of the present world were formed; (3) that same fundamental element insofar as it even now constitutes the world; (4) the principle of explanation, or explanatory source (identified with the primeval element), that logically and causally accounts for the phenomena of the world.

According to Aristotle, the pre-Socratic philosophers with cosmological theories agreed in explaining all phenomena as deriving from a single stuff or set of stuffs (sense 4). They disagreed about whether there was only one stuff or several. Those who held that there was only one stuff (monists) disagreed as to what it was: Thales said water; Anaximander said the Boundless; Anaximenes said air; and Heraclitus said fire. Those who held there were several stuffs or elements (pluralists) disagreed among themselves as to what those were: Empedocles said earth, water, air, and fire; Anaxagoras said an unlimited number of homogeneous stuffs including flesh, gold, wood; the atomists said an infinite number of atomic particles of differing shapes.

Aristotle's account, partly through the writings of his colleague Theophrastus on the history of philosophical opinions, dominated ancient and then modern interpretations. Unfortunately, there are a number of problems with his account. First, it seems to conflate two different types of theory, that of the alleged monists, and that of the pluralists, which may operate on different principles. Second, it ignores theories that posit a stable cosmology (in which the world does not arise out of a primeval chaos), such as those of Xenophanes and Heraclitus. Third, it seems to project back onto cosmologists of the sixth century BCE the theory of changeless being that Parmenides invented in the early fifth century BCE. Fourth, it assumes a sophisticated theory of matter in which a subject is distinguished from attributes or properties, which seems to arise only in the fourth century BCE. Fifth, it embodies a tendentious interpretation of how the pre-Socratics understood causal explanation.

The term *archē* itself in the sense of "beginning, starting point" might have been used by early pre-Socratics such as Anaximander, but there are no extant quotations to verify this. In the late fifth century Diogenes of Apollonia used the term to mean something like "starting point," with a possible implication of being an explanatory principle. (fr. 1). But the term only seems to become a philosophically important one when one considers that Plato described an archē as a principle to which nothing is prior (Republic 511b, Phaedrus 245c-d), in effect as supplying a metaphysical ground and a logical axiom. Aristotle himself distinguished six senses of the term, only the last of which is a technical philosophical one, reflecting Plato's use (Metaphysics V.1). Aristotle's account of the archē as a principle of explanation among the pre-Socratics is highly suggestive but should not be accepted uncritically.

Most of the pre-Socratics were interested in explaining how the present world arose out of a primeval chaos, and also in identifying the basic realities from which the world arose. In those two senses, they sought through their studies and writings to elucidate the sources, the archai, of the world. Whether, or in what sense, their basic realities were material, and whether they were changeless, are controversial questions scholars still wrestle with.

See also Aristotle; Pre-Socratic Philosophy.

Bibliography

Algra, Keimpe. "The Beginnings of Cosmology." In *The Cambridge Companion to Early Greek Philosophy*, edited by A. A. Long. Cambridge, U.K.: Cambridge University Press, 1999.

Alt, Karin. "Zum Satz des Anaximenes über die Seele: Untersuchung von Aetios *Peri archōn*." *Hermes* 101 (1973): 129–164.

Barnes, Jonathan. *The Presocratic Philosophers*. Rev. ed. London: Routledge, 1982.

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 1: *The Earlier Presocratics and the Pythagoreans*. Cambridge, U.K.: Cambridge University Press, 1962.

Stokes, Michael C. *One and Many in Presocratic Philosophy*. Washington: Center for Hellenic Studies, 1971.

Daniel W. Graham (2005)

ARCHYTAS OF TARENTUM

(c. 425 BCE—c. 350 BCE)

Archytas of Tarentum was active in the first half of the fourth century BCE as a mathematician and a philosopher in the Pythagorean tradition. He is famous for having sent a ship in 361 BCE to rescue Plato from Dionysius II, tyrant of Syracuse. Archytas is unique among ancient philosophers for his success in the political sphere—he was elected general seven consecutive times in a democratically governed Tarentum (at the time one of the most important Greek city-states in southern Italy).

More texts have been preserved in Archytas's name than in that of any other Pythagorean, but the majority of these texts are spurious. The pseudo-Pythagorean treatises of the first century BCE and later were often written in his name, considering him the latest of the three great early Pythagoreans (following Pythagoras himself and Philolaus). The spurious works on categories in Archytas's name were regarded as genuine by the commentators on Aristotle's *Categories* and were frequently cited. Four fragments survive from Archytas's genuine works, of which *Harmonics* was the most important, and there is a relatively rich set of testimonia.

Archytas provided the first solution to one of the most celebrated problems in ancient mathematics, the duplication of the cube. One romantic version of the problem reports that the inhabitants of the island of Delos were commanded by the god to build an altar double the size of the current altar, which had the shape of a cube. The problem was thus to determine the length of the side on which to build a cube of double the volume. Archytas's solution is a masterpiece of mathematical imagination, requiring one to envision the intersection of two lines drawn on the surface of a semicylinder—one by a rotating semicircle and one by a rotating triangle. In later antiquity, a story arose that Plato was critical of Archytas for using mechanical instruments to find the solution and thus perverting the true function of mathematics—that is, to direct the soul to the intelligible realm. This story was probably invented to explain the separation of the science of mechanics from philosophy. No physical instruments are employed in Archytas's solution, and it was criticized by some ancient authors as too abstract and of little practical application. Although Plato's complaints about the state of solid geometry in his day (*Rep*. 528b–d) may be directed at Archytas, they focus not on the use of instruments but rather on the failure of its practitioners to develop a coherent science of solid geometry.

Fr. 1, the beginning of Archytas's *Harmonics*, is the earliest text to identify a quadrivium of four sciences (the science of number, geometry, astronomy, and music). Archytas praises the sciences for beginning by distinguishing the universal concepts relevant to the specific science, but he regards their ultimate goal as an account of individual things in the world in terms of number, thus building on Philolaus's insight that all things are known through number. Archytas's own *Harmonics* begins by distinguishing important general conceptions in acoustics. His mistaken view that pitch depends on the speed with which a sound travels—it depends, in fact, on the frequency of impacts in a given period—was adopted with modifications by both Plato and Aristotle and was the most common view in antiquity. Archytas provided an important proof that the basic musical intervals such as the octave, which correspond to ratios of the form (n+1)/1, cannot be divided in half.

The goal of Archytas's harmonics, however, was the description of a particular set of phenomena—in this case the musical scales in use in his day—in terms of specific numerical ratios. Plato complained that such a science of harmonics sought numbers in the sensible world rather than ascending to more abstract problems, which were independent of the phenomena (*Rep*. 531c). For Archytas, however, there was no split between the intelligible and sensible world. Logistic, the science of number and proportion, was the master science for Archytas, because all other sciences ultimately rely on number to provide knowledge of individual things (Fr. 4). Just as his science aimed at mathematical description of concrete phenomena, so Archytas also developed a theory of definition that earned Aristotle's praise (*Metaph*. 1043a14–26) for taking into account not just the limiting (formal) aspect of the definiendum but also the unlimited (material) aspect.

Archytas argued that number was crucial in the political and ethical sphere as well. The stability of the state is based on the widely held human ability to calculate, which convinces the rich and the poor that they have their fair share (Fr. 3). Archytas regarded bodily pleasure as inimical to the rational calculation that should guide one's life, because, he believed, someone in the throes of the most intense pleasure (e.g. sexual orgasm) is manifestly unable to calculate.

There is little evidence for Archytas's cosmology, but he developed the most powerful ancient argument for the infinity of the universe. Archytas assumes that, if the universe is limited, it has an edge (modern science would question this assumption) and asks whether or not some-

one standing at the edge would be able to extend his or her hand beyond the edge. Normal assumptions about space suggest that it would be paradoxical if the person could not extend a hand beyond the edge. Archytas can ask the same question about any supposed limit, and hence the universe will not have a limit and will extend indefinitely. Versions of this argument were adopted by the Epicureans, Stoics, Locke, and Newton—although both Plato and Aristotle rejected it. Aristotle wrote three books—now lost—on Archytas and presents him favorably. Plato was impressed with Archytas's work in mathematics, but the two philosophers disagreed sharply on important philosophical issues.

See also Philolaus of Croton; Pythagoras and Pythagoreanism.

Bibliography

TEXTS AND COMMENTARIES

Huffman, Carl A. *Archytas of Tarentum: Pythagorean, Philosopher and Mathematician King*. Cambridge, U.K.: Cambridge University Press, 2005.

DISCUSSIONS

Barker, Andrew D. *Greek Musical Writings, Vol. II: Harmonic and Acoustic Theory*. New York: Cambridge University Press, 1989.

Huffman, Carl A. "The Authenticity of Archytas Fr. 1." *Classical Quarterly* 35 (2) (1985): 344–348.

Knorr, Wilbur Richard. *The Ancient Tradition of Geometric Problems*. New York: Dover, 1993.

Lloyd, G. E. R. "Plato and Archytas in the Seventh Letter." *Phronesis* 35 (2) (1990): 159–174.

Winnington-Ingram, R. P. "Aristoxenus and the Intervals of Greek Music." *Classical Quarterly* 26 (1932): 195–208.

Carl A. Huffman (2005)

ARDIGÒ, ROBERTO
(1828–1920)

Roberto Ardigò, the principal figure in Italian positivism, was born in Casteldidone in Cremona. He became a Catholic priest, but left the priesthood when, at the age of forty-three, he found it no longer compatible with his beliefs, particularly his conviction that human knowledge originates in sensation—a conviction that came to him suddenly, as he recounted it, while staring at the red color of a rose (*Opere,* Vol. III, p. 368). From 1881 to 1909 he taught history of philosophy at the University of Padua. He spent the last years of his life defending and illustrating his fundamental ideas and debating with the prevailing idealism, which had supplanted positivism as the dominant viewpoint within and without the Italian universities during the last three decades of the nineteenth century. He died in Padua after two attempts at suicide.

The basic interests of Ardigò's positivism were not historical and social, as were Auguste Comte's, but scientific and naturalistic, like Herbert Spencer's. From Comte, Ardigò accepted the principle that facts are the only reality and that the only knowledge possible is the knowledge of facts, which consists in placing one fact in relation to others either immediately or by means of those mental formations that constitute ideas, categories, and principles. When these relations are established, the fact is "explained." Science, therefore, is the only kind of knowledge possible; and philosophy itself is a science that, like all other sciences, uses induction and does not have at its disposal privileged principles or procedures. Metaphysics, which claims to start from principles independent of facts and to use deduction, is a fictitious science. Yet philosophy is not just a "synthetic" discipline in Spencer's sense of the unifier of the general results of the individual sciences. On the one hand, it is a complex of special disciplines that is left after the natural sciences have gone their way. As such, it encompasses the disciplines that are concerned with the "phenomena of thought" and finds articulation in two spheres: psychology, which includes logic, "gnosis" (epistemology), and aesthetics; and sociology, which includes ethics, *dikeika* (doctrine of justice or of law), and economics. On the other hand, to philosophy belongs the field of the *indistinct,* which lies outside the realm of the distinct, which constitutes the object of the individual sciences (matter, for physics; life, for biology; society, for sociology; mind, for psychology, etc.). This realm of the indistinct constitutes the unique and common origin of all the realms of the distinct, and it is the object of philosophy as *peratology* (*Opere,* Vol. X, p. 10).

The indistinct in the philosophy of Ardigò had the same function as the unknowable in Spencer. Ardigò distinguished it from the unknowable in that the indistinct is not that which is not known but that which is not yet known distinctly. It is a relative concept, because the distinct that emerges from some knowledge is in its turn indistinct with respect to further knowledge insofar as it is that which produces, solicits, and explains that knowledge (*Opere,* Vol. II, p. 350). The indistinct-distinct relationship was, moreover, used by Ardigò—in a manner analogous to the way Spencer used the homogeneous-heterogeneous relation—to explain "the natural formation" of every known reality. Every natural formation, in the solar system as well as in the human spirit, is a passage

from the indistinct to the distinct. This passage occurs necessarily and incessantly, regulated by a constant rhythm, that is, by an immutable order. But the distinct never exhausts the indistinct, which both underlies and transcends it; and since the distinct is the finite, then we must admit that, beyond the finite, lies the infinite as indistinct. Ardigò conceived the infinite as a progressive development without beginning or end (the analogue to Spencer's evolution), denying that such a development leads to a transcendent cause or principle (*Opere,* Vol. II, p. 129; Vol. III, p. 293; Vol. X, p. 519). All natural formations, including thought, which is a kind of "meteor" in the life of the universe, emerge from and return to this infinite (*Opere,* Vol. II, p. 189).

In the domain of psychology, Ardigò held that the I (self) and natural things are constituted by neutral elements, that is, sensations. The self and things differ, therefore, only by the nature of the synthesis, that is, by the connections that are established among the sensations. Those sensations that refer to an internal organ and have the character of continuity are associated in the "autosynthesis," or the self. Those sensations that refer to an external organ and are discontinuous are associated in the "heterosynthesis" that gives rise to things (*Opere,* Vol. IV, p. 529 ff.). This doctrine, propounded by Ardigò in his very first work, *La psicologia come scienza positiva* (Mantua, 1870), is similar to that later propounded by Ernst Mach in *Die Analyse der Empfindungen* (Jena, 1886).

In the moral domain Ardigò carried on a polemic against every kind of religious and rationalistic ethic. It is a fact, according to Ardigò, that humans are capable of disinterested or altruistic actions, but such actions can be explained by recourse to natural and social factors. The ideals and the prescriptive maxims that determine them derive from the reactions of society to acts that either preserve or damage it—reactions that impress the individual and become fixed in his conscience as norms or moral imperatives. That which is called "conscience," therefore, is the progressive interiorization accomplished by the repeated and constant experience of the external sanctions that the antisocial act encounters in society (*Opere,* Vol. III, p. 425; Vol. X, p. 279).

Finally, Ardigò tried to mitigate the rigorous determinism found in all forms of positivism by giving some emphasis to the notion of chance. Chance consists in the intersecting of various causal series that, taken together, constitute the order of the universe. These intersections are unpredictable, though the events that constitute every individual series are not unpredictable. So-called human "freedom" is an effect of the plurality of the psychical series, that is, of the multiplicity of the possible combinations of various causal orderings that constitute man's psychical life (*Opere,* Vol. III, p. 122).

See also Comte, Auguste; Determinism and Freedom; Idealism; Mach, Ernst; Positivism.

Bibliography

WORKS BY ARDIGÒ

Opere, 12 vols. Padua, 1882–1912.

La scienza dell'educazione. Padua, 1893; 2nd ed., 1903. Not included in the *Opere.*

WORKS ON ARDIGÒ

Amerio, F. *Ardigò.* Milan, 1957. With bibliography.

Bluwstein, J. *Die Weltanschauung Roberto Ardigòs.* Leipzig: Eckardt, 1911.

Marchesiani, G., and A. Groppali, eds. *Nel 70o anniversario di Roberto Ardigò.* Turin, 1898.

Marchesiani, G. *La vita e il pensiero di Roberto Ardigò.* Milan, 1907.

Marchesiani, G. *Lo spirito evangelico di Roberto Ardigò.* Bologna, 1919.

Marchesiani, G. *Roberto Ardigò, l'uomo e l'umanista.* Florence, 1922.

Nicola Abbagnano (1967)
Translated by Nino F. Langiulli

ARENDT, HANNAH
(1906–1975)

Hannah Arendt, American philosopher and political scientist, was born in Hanover, Germany. In 1928 she completed her PhD under Karl Jaspers at the University of Heidelberg, having previously studied with Martin Heidegger at the University of Marburg. Upon immigrating to the United States in 1941, she became director of several Jewish organizations and served as chief editor of Schocken Books before being appointed to the Committee on Social Thought at the University of Chicago in 1963. She taught at the New School for Social Research in New York from 1967 until her death.

THE INFLUENCE OF HEIDEGGER AND PHENOMENOLOGY

Despite sharing Jaspers's views about the existential importance of communication, Arendt's philosophy mainly bears the imprint of Heidegger's phenomenology. Following Edmund Husserl, Heidegger argued that the scientific worldview conceals the genuine appearances of things as they are directly presented within lived experi-

ence. By abstracting from, and thereby concealing, the primal experience of meaning and value, this worldview provokes a crisis of nihilism, the practical upshot of which—foreshadowed in Friedrich Nietzsche's thought—is a technological "will to power" that reduces all of being to the status of a predictable and useful object. Because modern science is but the culmination of a metaphysical tradition dating back to Plato, Heidegger turned to the pre-Socratics and the archaic language and life of the ancient Greek *polis* for disclosing a more original experience of things. Arendt followed him in this respect, but with different results. Heidegger's supreme estimation of the revelatory power of the lone thinker/poet/artist to open up a new experience of community and world—coupled with his contempt for the indecisiveness of public opinion and democratic political debate—led him to embrace the resolve of a Nazi führer who embodied the will of the German people. In Arendt's judgment, Heidegger's politics betrayed his own critique of European metaphysics as an elitist form of idealism that conceals the common roots of meaning and value in democratic action.

FREEDOM AND POLITICAL ACTION

Action is part of a triad of comportments that together make up the active life definitive of the human condition. As the quintessential appearance of human freedom, political action, Arendt argued, must be distinguished from both work and cultural fabrication. Laboring to procure life's necessities is unfree; and the freedom of artistic creation is at best hidden and derivative. As distinct from the solitary application of means in pursuit of ends, true freedom must be communicated publicly, in political deeds and words. For this there must be a public space—exemplified by the Greek *polis* and such modern-day equivalents as the worker council and town hall meeting—wherein equals representing diverse opinions meet and deliberate together.

Arendt often invoked Augustine's comment on the miracle of birth, or what she called *natality*, in capturing the distinctive capacity of political action to initiate new beginnings. The concept owes much to Arendt's lifelong obsession with modernity and political revolutions, although she traces it back to ancient Greek and early Christian notions of freedom. In discussing the archaic Greek notion of freedom (*archein* = to begin or initiate), Arendt stresses the utter unpredictability of actions that draw their meaning and identity from the distinctiveness of the individual actors whose personality they express. Early Christian thinkers such as Paul and Augustine develop this idea further in discussing religious conversion as spiritual rebirth. The existential pathos of continually breaking with the past and remaking oneself also informs modern revolutionary thought, which appeals to free consent rather than traditional authority as the principal underlying political life.

THE TENSION BETWEEN FREEDOM AND SOCIAL EQUALITY

Although modern revolutions exemplify political freedom, Arendt thought that their failure to distinguish this end from the social struggle for equality conflated the imperatives of political action with those of economic production and consumption. The subsequent substitution of efficient administration for political action is especially apparent in the revolutionary movement inaugurated by the French Revolution and brought to completion in twentieth-century communist and fascist revolutions. Here, freedom is reduced to the sovereign legislation of a unified will that seeks to administer the general welfare of all citizens with the ultimate aim of remaking them into a single, harmonious body. In Arendt's opinion, the American Revolution evolved differently, partly because it was not faced with the same social problems, and partly because it was nourished on Protestant individualism rather than Catholic paternalism. It was not driven by economic need and class struggle, and the remnants of feudalism—mainly concentrated in the slave economy of the South—had already been eclipsed by the modern commercial economies of the North. Yet according to Arendt, the individualistic spirit of commercial life that compelled the Founding Fathers to adopt limited and divided forms of governance would also prove to be the undoing of their revolution. As Americans became more preoccupied with their private economic pursuits and problems of class developed within industrial capitalism, political life receded in importance and a paternalistic welfare state eventually emerged.

POWER, VIOLENCE, AND LEGITIMACY

Arendt's distinction between political power and political violence builds upon her critique of the welfare state. Contrary to the dominant view held by the Weberian school, political power is not equivalent to wielding a monopoly of instruments that can be brought to bear in top-down fashion by governmental elites in coercively defending and administering a state. On the contrary, political power consists in popular consent and public opinion nourished in open discussion. As such, its vital-

ity depends on multiplying resistances rather than by concentrating forces, a condition that is best promoted by encouraging the flourishing of open debate. Following Baron de Montesquieu, Arendt held that policies that preserve this discursive plurality by separating or dividing governmental powers and instituting a system of checks and balances are more powerful and enduring than ones that do not. Totalitarian regimes that dispense with the rule of law and concentrate all power in the hands of a single leader are notoriously unstable and weak because they deprive their own citizens of the public space necessary for taking independent initiative and uniting politically.

According to Arendt, the violence exercised by totalitarian regimes against their own citizens is but the reverse side of their impotence. Arendt equates violence with any coercive, instrumental action that lacks prior popular consent. Although it can never be legitimate, or politically justified, violence may sometimes be morally justified as a necessary means for avoiding great evil. Emergencies of state sometimes call for violent measures, but as Arendt notes, liberal democracies often use this pretext to suppress political action unjustly, and indeed any unilateral governmental intervention, however bureaucratically routine, bears traces of violence.

THE DECLINE OF AUTHORITY AND THE CRISIS IN CULTURE

Many of Arendt's studies—on totalitarianism, evil, revolution, and the Jewish question—document the political impact wrought by the decline of traditional authority and the crisis of culture. Although she did not blame secular Enlightenment and its revolutionary offspring for this decline, she nonetheless believed that the destruction of the old Roman trinity of religion, tradition, and authority contributed to a crisis of culture that undermined essential differences—between public and private, political and economic, action and work—on which the survival of a public political space depended. Transcendent authority anchored the autonomy of the public realm as a space for manifesting immortal deeds in beautiful words; the waning of authority diminishes that autonomy, thereby enabling the assimilation of both culture and politics to economic life.

Arendt's diagnosis of the crisis in culture bears directly on her political concerns. She appealed to the Greek ideal of culture as a religious memorialization of political community. In the absence of traditional religious authority, culture can provide those standards of judgment so essential for maintaining a common space for action. Political life is thus jeopardized whenever culture loses its normative authority—that is to say, whenever it is monopolized by elites, manipulated by government for purposes of propaganda, or is degraded to the mundane level of mass consumption and entertainment.

TOTALITARIANISM AND RADICAL EVIL

According to Arendt, the crisis of culture is symptomatic of all mass societies, or societies wherein individuals—isolated from one another in the lonely pursuit of familial and vocational aims—cease to engage in political action; and it is therefore one of the main conditions paving the way for modern totalitarianism. Under these conditions, it is the state, not politically engaged individuals, that assumes responsibility for integrating the masses, even when doing so renders individuality and life itself superfluous.

By engendering a system in which life is made superfluous, totalitarianism represents the epitome of evil. Contrary to popular opinion, such evil is seldom if ever motivated by diabolical intentions. Adolph Eichmann's evil, Arendt observed, simply consisted in his banal "thoughtlessness." Like most persons living in mass society, he confused moral duty with the duty to obey authority. However, Arendt also believed that the "absolute goodness" and violence born of idealism (as personified in Melville's Billy Budd) are as pernicious as the radical evil and destructiveness born of any workmanlike devotion to order. In both instances, the critical check provided by consulting the opinion of others who comprise an enlarged public sphere is totally absent.

JUDGMENT AND POLITICAL ACTION

Arendt's appeal to an enlarged public sphere touches upon the importance of judgment in sustaining political action. In the classical tradition of Aristotle and Aquinas, the judgment that guides action is intimately connected to a practical wisdom (*phronesis*), or prudential art, cultivated by experience and habituation in customary modes of behavior. In modern times, beginning with Immanuel Kant, judgment acquires an altogether different sense, one based on an impartial consideration of possible points of view. These two senses of judgment—the former typically associated with the standpoint of the political or moral actor faced with practical decision, the latter with the historical or aesthetic spectator who understands, interprets, and narrates action retrospectively and disinterestedly—intersect in Arendt's thought.

Prior to *The Life of the Mind* (1978), Arendt still affirmed the intimate connection between "a judgment of the intellect" and knowledge of the rightness and wrongness of practical aims (1968, p. 152). Indeed, she insisted that moral and political agents living in modern conditions are especially obligated to judge the laws, opinions, and actions of their society from the common—if not universal—standpoint of "all those who happen to be present" (p. 221).

Arendt's late lectures on Kant's political philosophy revise this connection between action and judgment. With the deterioration of public spaces requisite for exercising practical judgment, judgment ceases to be linked with the two faculties of practical reasoning—knowing and willing—and instead takes on the function of retrospective interpretation. As a vicarious form of action, historical spectatorship preserves the memory of those all-too-rare and tragically ill-fated moments of political action—such as the Paris Commune of 1871, the resistance of the Warsaw Ghetto, and the Hungarian revolt of 1956—by judging their universal, exemplary validity. Rescuing these unprecedented displays of spontaneous self-determination from the oblivion of history, judgment dignifies what otherwise appears to be an unbearable, arbitrary, compulsive—in short, utterly contingent and irresponsible—act of freedom.

Jürgen Habermas and others have rightly criticized Arendt for dissociating the common sense guiding judgment from any relationship to truth or justice. Her earlier work, for example, links the cultivation of common sense to the agonal exchange of opinions. Because this communication is constrained by the real effects of social domination, it remains prejudiced by ideological distortions. By contrast, her later work (following Kant) links historical judgment to an ideal *sensus communis*, or hypothetical community of taste (feeling). Here judgment achieves impartiality by imaginatively representing the standpoints of other persons as they may have been communicated had these persons been free from domination and constraint. No doubt, an accurate account of responsible judging lies somewhere between these extremes of realism and idealism, as even Arendt herself suggests; for judging, it seems, bears witness to rationality only when tempered by the real—mutual and impartial—criticism that obtains between actors who aspire to ideal freedom and equality.

See also Aristotle; Augustine, St.; Evil, The Problem of; Freedom; Habermas, Jürgen; Heidegger, Martin; Jaspers, Karl; Kant, Immanuel; Rousseau, Jean-Jacques; Thomas Aquinas, St.

Bibliography

WORKS BY ARENDT

The Human Condition. Chicago: University of Chicago Press, 1958.

The Origins of Totalitarianism. Enlarged edition. New York: Harcourt Brace & Co., 1958.

On Revolution. New York: Viking, 1963.

Eichmann in Jerusalem: A Report on the Banality of Evil. New York: Viking, 1965.

Men in Dark Times. New York: Harvest Books, 1968.

On Violence. New York: Harcourt Brace & World, 1970.

Crisis of the Republic. New York: Harcourt Brace & Jovanovich, 1972.

Rahel Varnhagen: The Life of a Jewish Woman. New York: Harcourt, 1974.

Between Past and Future: Eight Exercises in Political Thought. New York: Viking Press, 1978.

The Jew as Pariah: Jewish Identity and Politics in the Modern Age, edited by Ron Feldman. New York: Grove Press, 1978.

The Life of the Mind. 2 vols., edited by Mary McCarthy. New York: Harcourt Brace & Jovanovich, 1978.

Lectures on Kant's Political Philosophy, edited by R. Beiner. Chicago: University of Chicago Press, 1982.

Correspondence: 1926–1969: Hannah Arendt/Karl Jaspers, edited by Lotte Kohler and H. Saner. New York: Harcourt, Brace & Co., 1992.

Essays in Understanding: 1930–1954, edited by J. Kohn. New York: Harcourt, Brace & Co., 1994.

Love and Saint Augustine, edited by Joanna Scott & Judith Stark. Chicago: University of Chicago Press, 1996.

Responsibility and Judgment, edited by J. Kohn. New York: Schocken, 2003.

Letters: 1925–1975: Hannah Arendt/Martin Heidegger, edited by Ursula Ludz. New York: Harcourt Inc., 2004.

WORKS ON ARENDT

Benhabib, S. *The Reluctant Modernism of Hannah Arendt*. Beverly Hills, CA: Sage, 1996.

Bernauer, J., ed. *Amor Mundi: Explorations in the Faith and Thought of Hannah Arendt*. Dordrecht, Netherlands: Martinus Nijhoff, 1987.

Bernstein, R. *Hannah Arendt and the Jewish Question*. London: Polity, 1996.

Bowen-Moore, P. *Hannah Arendt's Philosophy of Natality*. New York: St. Martin's Press, 1989.

Bradshaw, L. *Acting and Thinking: The Political Thought of Hannah Arendt*. Toronto: University of Toronto Press, 1989.

Carnovan, M. *Hannah Arendt: A Reinterpretation of Her Thought*. New York: Cambridge University Press, 1992.

Disch, L. *Hannah Arendt and the Limits of Philosophy*. Ithaca, NY: Cornell University Press, 1994.

Garner, R., ed. *In the Realm of Humanitas: Responses to the Writings of Hannah Arendt*. New York: Peter Lang, 1990.

Gottsegen, M. G. *The Political Thought of Hannah Arendt*. Albany, NY: State University of New York, 1994.

Hill, M. A., ed. *Hannah Arendt: Recovery of the Public World*. New York: St. Martin's Press, 1979.

Hinchman, L. P., and K. Sandra, eds. *Hannah Arendt: Critical Essays*. Albany: State University of New York, 1994. Contains essays by Arendt.

Honig, B., ed. *Feminist Interpretations of Hannah Arendt*. Philadelphia: Pennsylvania State University, 1995.

Kateb, G. *Hannah Arendt: Politics, Conscience, Evil*. Totowa, NJ: Rowman & Allanheld, 1984.

Kohn, J., and L. May. *Hannah Arendt: Twenty Years Later*. Cambridge, MA: MIT Press, 1996.

Parekh, B. C. *Hannah Arendt and the Search for a New Political Philosophy*. Atlantic Highlands, NJ: Humanities Press, 1981.

Pitkin, H. F. *The Attack of the Blob: Hannah Arendt's Concept of the Social*. Chicago: University of Chicago, 1998.

Villa, D. *Arendt, Heidegger, and the Fate of the Political*. Princeton, NJ: Princeton University Press, 1996.

Villa, D., ed. *The Cambridge Companion to Hannah Arendt*. Cambridge, U.K.: Cambridge University Press, 2000.

Watson, D. *Hannah Arendt*. London: Fotaten Press, 1992.

Young-Bruehl, E. *Hannah Arendt: For Love of the World*. New Haven, CT: Yale University Press, 1982. Contains a bibliography.

David Ingram (1996, 2005)

ARETĒ/AGATHON/KAKON

Aretē, meaning "excellence" or "virtue," is central to ancient Greek ethics, from the early poets through Plato and Aristotle to the Stoics. It is a quality necessary for success, and the *aretai* for moral success are moral virtues. *Agathon,* meaning "good," implies virtue when used to describe human beings, as does *kalon* (meaning "noble" or "beautiful"), the adjective most closely associated with aretē and nearly synonymous with agathon. *Kakon* implies the lack of virtue. In Hesiod and Solon the moral use of these terms is well established, and it is clearly prefigured in Homer. Virtue, to such poets, no less than to Plato, is long lasting and independent of wealth and power. The principal virtues under discussion before Socrates were shame (*aidos*), reverence (*hosion*), and justice (*dike*). Protagoras evidently considered shame and justice to be essential to a stable society.

Socrates and Plato taught that virtue is to the soul as health is to the body. In addition to reverence and justice, they treated wisdom, courage, and sound-mindedness (or temperance; in Greek, *sôphrosunê*) as virtues. Plato represents Socrates in the early dialogues as unsuccessfully seeking definitions for the virtues, while hypothesizing that they are in some way identical with each other. Socrates is often thought to have held an intellectualist account of arête.

In the *Republic* Plato works out a theory of virtue from his account of health in the soul: Justice is the quality that allows the parts of the soul to work together in harmony, and the other virtues depend on that harmony. In a related context Plato somewhat mysteriously compares the form of the good to the sun; what the sun does to illuminate and nourish the world humans can merely see with their senses, and what the good does for the world humans can investigate with their intellect.

Aristotle's ethics begins from the hypothesis that all things aim at the good (agathon). The good for human beings, he says, is flourishing or happiness (*eudaimonia*), and the qualities that enable people to reach these goals he calls virtues (aretai). His account of virtues has been fundamental to all subsequent discussion of the subject in the European tradition. Moral and intellectual virtues are both necessary for human flourishing, and for each other. Moral virtues temper the soul to enjoy what is good, rather than what is bad, and consist in a disposition to experience emotions that lie on a mean between excess and defect. Courage, for example, belongs to a soul that is neither too rash nor too timid. In Stoic theory, nothing is entirely good but virtue, and this consists mainly in the ability to resist powerful emotion.

Some early Greek authors distinguish aristocrats as *agathoi* from common people as *kakoi.* The scholar A. W. H. Adkins identified the virtues that marked this class difference as competitive (as opposed to moral) virtues; he argued that in the time of Socrates and Plato, Greek thought about virtue underwent a major shift, and the philosophers brought the first usage of these terms that is moral in human sense. Hugh Lloyd-Jones and Bernard Williams contested Adkins's arguments, and the emerging consensus among scholars favors a more unified account of these terms.

See also Aristotle; Eudaimonia; Plato; Socrates.

Bibliography

Adkins, A. W. H. *Merit and Responsibility: A Study in Greek Values.* Chicago: University of Chicago Press, 1975.

Annas, Julia. *The Morality of Happiness.* New York: Oxford University Press, 1993.

Cooper, John. *Reason and Human Good in Aristotle.* Cambridge, MA: Harvard University Press, 1975.

Irwin, Terence. *Plato's Ethics.* New York: Oxford University Press, 1995.

Lloyd-Jones, Hugh. *Sather Classical Lectures.* Vol. 41: *The Justice of Zeus.* Berkeley, CA: University of California Press, 1983.

Williams, Bernard. *Sather Classical Lectures.* Vol. 57: *Shame and Necessity.* 2nd ed. Berkeley: University of California Press, 1993.

Paul Woodruff (2005)

ARISTIPPUS OF CYRENE

(c. 435–c. 356 BCE)

Aristippus of Cyrene, founder of the Cyrenaic school of philosophy, was born in the Greek North African port city of Cyrene (now Shahhat, Libya). Attracted to Athens by the fame of Socrates, he became a member of the Socratic circle and probably associated with Protagoras and Gorgias as well.

Like Socrates, Aristippus concentrated on ethics, conceived as endeavoring to determine the good life for the individual, and rejected the study of nature as both uncertain and useless for furthering the good. He gave a simple answer to the question of the goal of life: It is pleasure and nothing else. The wise man will arrange his life so that, as far as possible, one pleasure follows another and pains are kept to a minimum. He will not forgo a present pleasure for the sake of one to come, for the future is uncertain. But he will be master of his pleasures, as Socrates was, unperturbed when they must be done without.

Pleasures are individual episodes of internal feeling, not mere absence of pain but positive bodily sensations as experienced in eating, drinking, and sex. All pleasures, considered as pleasures, are equal, he declared, though they may differ in intensity, which is why those of the body take precedence over those of the mind. They are still pleasures even if produced by activities conventionally regarded as shameful. Virtues and friendships are goods only insofar as they are productive of pleasures. He found proof that pleasure is the goal of life in the (alleged) fact that all animals, as well as uninstructed human beings, pursue it by nature.

Aristippus taught his philosophy in the marketplace (unlike Plato, who taught in his gated Academy) and charged substantial fees. Like a modern psychiatrist, he regarded his services as therapy: liberation from superstitions and irrational conventions; and the fees, illustrating (so he claimed) the proper use of money, were part of the treatment. He also showed his pupils how to get along with anybody in any situation.

Many stories illustrate how Aristippus lived by his own principles, such as they were. Notorious for his involvement with the famous and expensive prostitute Lais, he insisted, "I have her, she doesn't have me." (As Cicero remarked, this sounds better in Greek.) And, he averred, having sex with one who has sex with many is no different from voyaging on a ship that carries other passengers. He perfumed himself. Sojourning in Syracuse at the court of Dionysius, he dressed in women's clothing for a party at the tyrant's behest. (Plato, there at the same time, refused.) When a client protested the high price he asked for educating his son, saying that for the same amount of money he could buy a slave, Aristippus told him to go ahead and buy the slave: Then he would have *two* slaves, the one he bought and his own son.

Traveling widely, Aristippus was pleased to be "everywhere a stranger," freeloading the advantages of city life without incurring the burdens of citizenship. Freedom, he held, consists not only in not being ruled but also in not ruling, for the ruler is the slave of those he rules.

See also Cyrenaics; Socrates.

Bibliography

Döring, Klaus. *Der Sokratesschüler Aristipp und die Kyrenaiker*. Mainz, Germany: Akademie der Wissenschaften und der Literatur; Stuttgart, Germany: F. Steiner Verlag Wiesbaden, 1988.

Mannebach, Erich, ed. *Aristippi et Cyrenaicorum fragmenta*. Leiden, Netherlands: E. J. Brill, 1961.

Wallace Matson (2005)

ARISTO OF CHIOS

(third century BCE)

Aristo of Chios was a disciple of Zeno of Citium, the founder of Stoicism. The scant biographical information that exists, from Diogenes Laertius (VII 160–64), describes him as an unorthodox Stoic, who later abandoned the school to found one of his own. There is some question in Diogenes' sources as to whether works ascribed to him are genuine or belong to the peripatetic Aristo of Ceos. But there are difficulties about his views as well. Like Zeno, he accepted the Socratic and cynic principle that virtue was sufficient for happiness. But whereas Zeno identified this with "living consistently," Aristo understood it as an internal consistency, where one behaved indifferently toward anything that was not virtue or vice (*adiaphoria*). At the core of his philosophy is the view that moral values are absolute: Only virtue is good and only vice bad; everything else intermediate between these is absolutely indifferent and equal. The third head of the school, Chrysippus, who polemicized against Aristo, was successful in establishing his own interpretation of Zeno's thought as the orthodox Stoic position, thus leading to Aristo's marginalization. But Aristo was held in high esteem by his contemporaries: Eratosthenes of Cyrene (c. 276–c. 194 BCE) maintained that Aristo's

philosophy, along with that of the skeptic Arcesilaus, was the most important of his time.

The confusion with Aristo of Ceos makes it difficult to attribute fragments that do not specify the author's origin. The most important is the summary given by Philodemus (*PHerc.* 1008, columns 10–23 Jensen), which has been attributed by Wehrli (1952), to the peripatetic Aristo of Ceos, but a study of the language and philosophical terminology reveals similarities with the surviving fragments of the Stoic Aristo. (Although this is included by Wehrli in the fragments of the peripatetic Aristo, a study of the language and philosophical terminology reveals similarities with the surviving fragments of the Stoic Aristo.)

See also Chrysippus; Stoicism; Zeno of Citium.

Bibliography

Arnim, Hans Friedrich August von. *Stoicorum Veterum Fragment: Collegit Ioannes ab Arnim*. Vol. 1 (1905). Stuttgart: B. G. Teubneri, 1968.

Ioppolo, Anna Maria. *Aristone di Chio e lo Stoicismo antico*. Naples, Italy: Bibliopolis, 1980.

Ioppolo, Anna Maria. "Una polemica antiscettica in Filodemo?" In *Epicureismo greco e romano*. Vol. II, edited by Gabriele Giannantoni and Marcello Gigante, 715–734. Naples, Italy: Bibliopolis, 1996.

Porter, James. "The Philosophy of Aristo of Chios." In *The Cynics: The Cynic Movement in Antiquity and Its Legacy*, edited by R. Bracht Branham and Marie-Odile Goulet-Cazé, 156–189. Berkeley: University of California Press, 1996.

Wehrli, Fritz. *Die Schule des Aristoteles, Texte und Kommentar, herausgegeben von Fritz Wehrli*. Vol. VI. Basel: Benno Schwabe, 1952.

Anna Maria Ioppolo (2005)

ARISTOTELIANISM

The question of what it means to be an Aristotelian—whether this requires adherence to a specific set of doctrines, a certain methodological approach, or the fulfilment of some other set of conditions—is a vexed one and has exercised the minds of self-professed Aristotelians and anti-Aristotelians alike over the course of twenty centuries. Like many problems of definition, it is best approached indirectly (as indeed Aristotle would likely have approached it). This historical overview starts from the broad assumption that one may consider Aristotelian all those thinkers who have either (a) considered Aristotle's texts a suitable point of departure for an enquiry into a given subject, or (b) thought themselves to be extending a peripatetic approach to a subject not covered by Aristotle himself. This assumption will have the consequence of making Aristotelians out of many whom modern reckoning would not readily count as philosophers. The result is not untoward because Aristotle's own enterprise extended far beyond philosophy thus narrowly defined.

THE FIRST *PERIPATOS*

Upon returning to Athens in 335 BCE, Aristotle founded a school in a grove consecrated to Apollo Lyceus. Hence the school was termed the Lyceum, yet it became forever known as the Peripatos for its covered colonnade. Indeed, in the annals references to "Peripatetics" greatly outnumber those made to "Aristotelians."

Aristotle's school was both a teaching and a research institution, with scholars pursuing interests ranging from musicology and the cataloguing of Greek forms of government to public lectures on popular subjects. The school survived Aristotle's departure from Athens and subsequent death in 322 BCE: Indeed, it flourished under Aristotle's successor and close collaborator, Theophrastus (372–287 BCE), who is reported to have presided over some 2,000 students.

Theophrastus expanded upon Aristotle's philosophical and scientific program. Theophrastus's botanical studies are pioneering works; the ancients especially valued his contributions to the categorical and hypothetical syllogistic. Theophrastus adheres to an aporetic methodology in the philosophical treatises while amassing observations in the scientific; this commitment to a peripatetic approach even leads Theophrastus to criticize Aristotle's *Metaphysics* Lambda in his own work on first philosophy. Theophrastus questions the extent to which teleological language, central to Aristotle's explanation of living nature, is applicable in a cosmic context: In effect, Theophrastus questions whether Aristotle is Aristotelian enough.

The Lyceum's independent spirit is further manifested in how its third head, Strato of Lampsacus (d. 269), departed from Aristotle on several important points, notably in natural philosophy. The diffuse activities and conflicting viewpoints countenanced within the Peripatos may have worked to its detriment in an age of intensifying competition between the philosophical schools. Strato's stewardship coincided with a decline in the school's fortunes, and within two generations it had all but disappeared from view.

THE IMPERIAL AGE

The nascent Hellenistic schools found elements to their liking in Aristotle's now-lost dialogues, praised for their style by Cicero and plundered for their edifying materials. Through criticism and creative appropriation, the Stoa in particular remained indebted to the peripatetics, who in the second century enjoyed a measure of resurgence under Critolaus. Still, self-professed peripatetics are hard to come by before Andronicus of Rhodes presented the ancient world with his authoritative collection of Aristotle's school works c. 50 BCE. Thereafter appear figures such as the Augustan intellectual Nicholas of Damascus, whose self-portrait is a model of Aristotelian virtue and who is credited with writing a compendium of Aristotelian philosophy, and Alexander of Aigai, teacher to Nero.

Andronicus's epoch-making edition is as important for the organization of its materials as for its contents, which quickly became canon. Immediately the impression is one of a full-fledged curriculum: The acquisition of methodological tools—the *Organon* of reasoned argument—is followed by an account of natural principles and natural bodies. After this comes living nature, then first philosophy (now dubbed "metaphysics"), and then the practical and productive sciences. Aristotle's widely varied investigations take on the appearance of a system here and retained it thereafter.

In Andronicus's wake there are two signal developments. First, propounding Aristotelian doctrine comes to be viewed as involving the writing of commentaries, starting with the *Categories* and *On Interpretation*. Second, in the first century BCE the Academician Antiochus thinks to present Aristotle as belonging essentially to the Platonic tradition. This classification set the tone for much of the imperial period. The fundamental continuity of Plato's and Aristotle's projects was correctly ascertained by late ancient thinkers and seized upon with momentous consequences.

The most important late ancient philosopher of purely peripatetic persuasion was Alexander of Aphrodisias. Around 200 CE Alexander was appointed to Athens's imperial chair in Aristotelian philosophy: He expounded his master's teaching in a series of magisterial commentaries *ad litteram*. Alexander's commentaries remain unsurpassed for erudition and insight, taking on all comers in a spirited defence of the Aristotelian worldview. Alexander's sharp, down-to-earth observations—for instance his unflinching admission that Aristotelian psychology makes no provisions for an immortal soul—provided a sobering reminder to later commentators who approached Aristotle's texts with loftier aspirations and syncretistic leanings. Though Alexander's Aristotle is undeniably a system-builder—it is with Alexander that the Aristotelian program of "saving the appearances" becomes a desire to explain each Aristotelian sentence by reference to another—he occasionally advances different interpretations without feeling the need to come down on one side. Alexander also wrote new treatises where he felt a lacuna existed in the extant corpus; and from his circle derives the peripatetic genre of disputed questions.

A different approach to Aristotle's texts is offered by Themistius, a late-fourth-century senator and proconsul of Constantinople. Themistius wrote paraphrases rather than commentaries; aporias and scholarly disputes take a back seat to a clear exposition of the main points. Yet Themistius positions himself as a peripatetic: his works and Alexander's provided a touchstone for later scholars who sought a genuine understanding of Aristotle's meaning.H

THE LATE ANCIENT SYNTHESIS

Plotinus (d. 270) is credited with an impressive dismantling of Aristotle's criticisms of Plato, and with the subsequent triumph of (neo-)Platonism in antiquity. But in the process, Plotinus also consolidated the assimilation of central Aristotelian concepts into a Platonic framework: for example, the potentiality-actuality distinction and the notion of pure contemplation as self-reflective. Plotinus's pupil Porphyry (d. 309) went further, attempting to show how nothing in Aristotle's virulently anti-Platonic categorical scheme in fact speaks against the primacy of separate Forms. The *Categories*, in its own words, purports to detail how things are spoken of: its universals are those abstracted from sense-particulars. The suggestion, embedded in Porphyry's enormously influential introduction (*Eisagôgê*) to Aristotle's *Organon*, is that Aristotelian science deals with substances prior to us, not with those prior by nature. This move made Aristotelian logic, and by extension natural philosophy, innocuous to ancient Platonists. It also set up the protracted Latin debate concerning the universals.

The Platonist appropriation of Aristotle was made complete in the fifth-century revival of the Athenian and Alexandrian schools. Aristotle was considered a largely reliable guide to the workings of the sensible cosmos: His works became positioned between those Platonic dialogues that were considered propaedeutic in character and those that disclosed the higher realities that Aristotle either failed to mention or knew nothing about. Though committed to the supremacy of the "divine" Plato over

the "daemonic" Aristotle, late ancient Platonists were thus Aristotelians, too, in their fashion. The voluminous commentaries on Aristotle's logic and natural philosophy testify to the care and attention devoted to subtle points of argument and doctrine. In negotiating tensions between Aristotle's treatises and Plato's dialogues, notably the *Timaeus* (a prime target of Aristotle's but a treatise that the Platonists ranked high), both reconciliation and taking Plato's side could produce philosophically interesting work, as the examples of Simplicius (*fl.* in the 530s) and Proclus (d. 485) show. So could an unorthodox mindset coupled with a healthy self-image and a nascent Christian agenda, as witnessed by the groundbreaking work of John Philoponus (d. 574).

Opinion varied about how far harmony extended in the direction of Plato's supernal principles. Iamblichus (d. 325) came under fire for suggesting that Aristotle would have subscribed to Plato's Forms, while Ammonius's (d. 517/526) equally hyperbolic claim that Aristotle's Prime Mover was intended as a divine creative force was broadly accepted. Ammonius's project of harmonizing Aristotle with Plato thus made Aristotle more acceptable to monotheists both in the Arabic-speaking East and, eventually, the Latin Christian West.

As for the Eastern Roman Empire, after the decline of Alexandria, the next high point for Aristotelian studies came with the Aristotelian circle assembled by Princess Anna Comnena in early-twelfth-century Constantinople. This activity resulted in commentaries by Eustratius and Michael of Ephesus and helped secure the transmission of Aristotelian materials to the Latin world.

ARISTOTELES ARABUS

Legend depicts Greek wisdom as passing from Alexandria to Baghdad: Although the chain of transmission is not as ironclad as Arabic-speaking Hellenophiles liked to pretend, the story contains a kernel of truth. The philosophy the Islamic world inherited, in particular, was Alexandrian and hence broadly Aristotelian. Aristotle's works were translated mostly through Syriac, by Christians. Many went through several recensions because the audience's growing scholarly acuity demanded progressively more exacting translations. By 950, all of Aristotle except for the *Politics* was available in Arabic (Plato's *Republic* replacing the latter), along with a host of commentaries. Creative reflection was underway among Muslims, Christians, and Jews alike, all of who wrote in Arabic.

A reliance on Alexandrian learning, which for the most part accepted the "lower" calling of explaining Aristotle, had the effect of making of Aristotle the preeminent sage of old. In the Arabic understanding, Aristotle had perfected, but also corrected, the views of other ancient thinkers, including his teacher Plato: The well-known adage of Aristotle considering "truth a truer friend" is traceable to al-Ghazālī (1058–1111), who can thus mockingly position himself as a peripatetic in spirit even when questioning the cogency of the Muslim *falâsifah*. But the Arabic Aristotle also manifested Platonic traits. This was due partly to the pseudonymous *Theology of Aristotle* and *Epistle concerning the Pure Good* (really Plotinus and Proclus in disguise), and mostly to a comfortable familiarity with the synthesis effected in late antiquity. The Peripateticism taught in the wake of al-Kindī (d. ca. 873) and al-Fārābī (d. 950) was both theist and emanationist.

The most powerful synthetic mind in Arabic philosophy, and the man responsible for tying the disparate threads of Aristotelianisms past into the service of a singular vision, was also the philosopher who eclipsed Aristotle in the East. Ibn Sīnā (the Latin Avicenna, 980–1037) progressed from traditional commentary to comprehensive philosophical encyclopaedias "presented in the manner of the peripatetics" to free-form expositions of his own views. Too Platonizing to be considered purely peripatetic, altogether too Aristotelian to be mistaken for a Platonist, lifting materials from the Muslim dialectical theologians as needed, Ibn Sīnā's philosophy constitutes an original achievement, one whose success is measured by the fact that in the East his works supplanted Aristotle's as the basis for study and philosophical reflection. It is thanks to Ibn Sīnā that mainstream Islamic philosophy to this day retains a broadly peripatetic vocabulary and orientation. Yet his substantial revisions to Aristotelian metaphysics, psychology, and logic, among other areas, were presented in such an attractive package that later philosophers rarely paused to consider whether Ibn Sīnā's philosophy faithfully reflected that of Aristotle. More important was that it conveyed truth. The subsequent period is consequently more rightfully called Avicennian than Aristotelian.

From this perspective, Ibn Rushd (Averroes, 1126–1198) appears a man out of place. Following upon al-Ghazālī's criticisms of Ibn Sīnā, Ibn Rushd advocated a return to an undiluted Aristotle, undertaking a massive commentary project worthy of Alexander or Themistius, both of whom he used extensively. This Cordovan commentator regarded Aristotle as a model of human perfection (*In De anima* III, comm. 14). For him, this faith in Aristotle's exemplary rationality and consistency held the key to settling any outstanding scholarly dispute. Sidelined in Islamic philosophy, Ibn Rushd became fabulously

influential among Jews and Christians, who viewed him as *the* Commentator (in antiquity, Alexander was similarly honoured).

ARISTOTELES LATINUS

The story of the Latin Aristotle begins with Boethius's (d. 525) stated intention of translating all of Aristotle's works. The project only got so far as the logical treatises; until the mid-twelfth century, of these only the *Categories* and *On Interpretation* circulated, making of Aristotle primarily a logician, and a curious one at that. A slow dissemination of the "new logic" (the full *Organon*) occurred in the twelfth century: acquaintance with Arabic philosophy—above all, Avicenna's *De anima* and *Metaphysics*—helped raise interest in Aristotle's natural philosophy and metaphysics, which were then translated in short order, often concurrently from the Greek and the Arabic in a race to get to the heart of the matter.

The theologically suspect aspects of Aristotelian teaching, which the Arabic tradition helpfully pointed out, promptly resulted in the 1210 and 1215 bans in Paris, then the most prestigious of the rising universities. This did little to stem the tide. By mid-century, studying the entire range of Aristotle's works—often coupled with Averroes's commentaries—was commonplace in the arts faculties. Aristotle himself was so ubiquitous that writers could refer to him simply as "the Philosopher."

Thereafter, Aristotle dominated philosophical teaching in the Christian West for three centuries. Hundreds upon hundreds of commentaries were produced at the height of scholasticism; the list of the major commentators was a roll call of the best and brightest of the schoolmen: Albertus Magnus, Thomas Aquinas, John Duns Scotus, William Ockham, Jean Buridan, and so on—for every major figure, there was a score more. As in ancient scholasticism, considerable philosophical ingenuity and innovation went on under a nominal exposition of the text (the *quaestiones* providing an even more congenial setting).

Especially going into the nominalist phase, the question arises: To what extent are some of these thinkers to be considered Aristotelian at all? Clearly, greater liberties were being taken; but this freedom would be expected following a period of assimilation. Moreover, adherence to a tradition need not stifle creative thought. The fallout from the famous condemnations of 1270–1277 spotlights the complex dynamic. For the most part, the condemnations were directed against the allegedly heterodox teachings of the so-called "Latin Averroists." But just because their radicalism was so resolutely Aristotelian—upholding the world's eternity and the unity of the intellect, and so on—the condemnations could be interpreted as an invitation to read Aristotle more creatively. And could the resultant bold conceptual and scientific inquiry not be considered more authentically Aristotelian than a single-minded adherence to the master's letter? "Radical" Aristotelians and radical "Aristotelians" were similarly drawn to the spirit of Aristotle's texts, in equal parts confident and intellectually curious. They merely took their admiration of the master in different directions.

THE MODERN AGE

The Renaissance humanists' newfound appreciation for the breadth of ancient culture put an end to Aristotle's supremacy. With the intellectual scene splintering into multiple incommensurable paradigms, Aristotle was effectively demoted to the headmastership of one school once more after long representing Greek wisdom in its entirety. As the quality of texts and translations came under scrutiny, the very state of Aristotle's preserved writings was found wanting. What to make of this was less evident. The Ciceronian Mario Nizilio could claim that wrinkles in expression signaled confusion in thought, whereas others blamed Andronicus's editorializing. Yet others took refuge in the ancient tradition, so that by the sixteenth century any configuration of Alexandrine, Themistian, Averroist, and even neo-Platonist tenets could be combined in an attempted rehabilitation of Aristotle, as exemplified by the works of philosophers such as Nicoletto Vernia (1442–1499) Agostino Nifo (d. 1538), and Pietro Pomponazzi (1462–1525). The textual drive had other unforeseen consequences. Elegant new translations of works such as the zoological and elemental treatises excited new scholarship, and Aristotle's *Poetics* at last found an appreciative audience among the literati.

With the Reformation, new complications emerged. Martin Luther's attitude towards Aristotle was ambivalent, but Melanchton enthusiastically endorsed the teaching of solid scholarly materials (excepting the *Metaphysics*). The Counter-Reformation likewise gravitated towards neoscholasticism. The late sixteenth and early seventeenth centuries thus saw a resurgence in the fortunes of Aristotle's works, which for a time were studied with equal intensity in the Protestant north and the Catholic south. Of particular note are the efforts of Francisco Suarez (1548–1617) and the Coimbra commentators.

By comparison, the seventeenth to nineteenth centuries represent a true dry spell for Aristotelian philoso-

phy. One may ask why; a tentative answer, if necessarily incomplete and hesitant, may yet tell us something about Aristotelianism as a historical force. Part of Aristotle's attraction had been the promise of a comprehensive, largely unified worldview, with pressure points doubling as the main locus for scientific advancement (discrepancies calling for new solutions). With the new sciences wresting fields of inquiry from the philosophers' hands, discipline by discipline, what appeared to be left of Aristotle was the barest husk of a system—in effect its extremities, logic and moral education. And of these, the nineteenth century threatened to supplant Aristotle's logic, long regarded as his lasting achievement. Antiquarian interest, it seems, could not of itself make Aristotelianism thrive. It could, however, help keep it alive, at least for a time.

THE NEW ARISTOTLE

The post-Enlightenment rise in Classical scholarship eventually brought about a renewed interest in ancient philosophy. But the philological and historical orientation of the new generation meant that Aristotle (along with Plato) returned with a difference. Instead of unity, the new scholarship sought signs of discrepancy, editorial interference, and intellectual development. Werner Jaeger's (1888–1961) studies mark a watershed, representing the culmination of a century's worth of textual work but also providing the launching point for countless philosophical studies sharing the same problem-oriented, if not aporetic, approach to Aristotle's works. An alternative to the genetic method would be to treat individual treatises as essentially closed units, examined closely but in splendid isolation. Twentieth-century analytic philosophy produced many such Aristotelian essays, while thinkers such as Brentano, Husserl, and Heidegger took more general inspiration from the Stagirite's writings.

Within the Catholic Church's sphere of influence, the nineteenth century witnessed the ascendancy of neoscholasticism, culminating in Leo III's 1879 encyclical officially endorsing Aquinas. Pius X further singled out twenty-four Thomist tenets to be taught in all Catholic institutions. This development injected a more systematic impulse into Aristotelian studies because Aquinas's Aristotle was the undisputed "master of those who know." Still, questions about Aristotle's perennial wisdom—as opposed to his historically conditioned contributions—persisted. The Thomist revival undoubtedly perpetuated a medieval understanding of Aristotle. But it also represented an important moment in the recovery of the medieval Aristotelian tradition as a whole.

Late-twentieth-century philosophers discovered in Aristotle new things again. As virtue ethics flourished, some proponents declared themselves neo-Aristotelians (Alasdair Macintyre, Martha Nussbaum), while others were so labeled. Philosophers of mind and biology found intriguing formulations in Aristotle's studies on living nature; even Aristotle's notoriously problematic modal syllogistic has garnered newfound respect as a philosophically sophisticated formalization of an essentialist metaphysics. In each case many have determined that Aristotle is best approached through an analytic engagement with his commentators—itself an ancient strategy.

This interplay of historical and systematic concerns prompts one final observation. Aristotle's works have been said to present a system *in potentia*. One possible history of Aristotelianism would accordingly unfold as a series of attempts by different thinkers in different ages to map out and explore the conceptual possibilities and limitations embedded in the texts received as Aristotle's. Such a story would span the history of Western thought, because no other philosopher has enjoyed such sustained attention (admittedly, Plato comes close). A welcome corollary is that the contemporary student has at her disposal a kaleidoscope of "Aristotelianisms" to aid in further understanding and exploration.

See also Alexander of Aphrodisias; al-Fārābī; Aristotle; Averroes; Avicenna; Brentano, Franz; Buridan, Jean; Duns Scotus, John; Heidegger, Martin; Husserl, Edmund; Luther, Martin; Metaphysics; Philoponus, John; Plato; Plotinus; Pomponazzi, Pietro; Simplicius; Suàrez, Francisco; Substance and Attribute; Themistius; Theophrastus; Thomas Acquinas, St.; Universals, A Historical Survey; Virtue Ethics; William of Ockham.

Bibliography

Barnes, Jonathan, and Miriam Griffin, eds. *Philosophia Togata II. Plato and Aristotle at Rome*. Oxford: Clarendon Press, 1997.

Blackwell, C., and S. Kusukawa, eds. *Philosophy in the Sixteenth and Seventeenth Centuries: Conversations with Aristotle*. Aldershot: Ashgate, 1999.

Commentaria in Aristotelem Graeca. 23 vols. and 2 supplementary vols. Berlin: G. Reimer: 1882–1907.

Des Chene, Dennis. *Physiologia. Natural Philosophy in Late Aristotelian and Cartesian Thought*. Ithaca: Cornell University Press, 1996.

Düring, Ingemar. *Aristotle in the Ancient Biographical Tradition*. Gothenburg: Institute of Classical Studies, 1957.

Gutas, Dimitri. *Avicenna and the Aristotelian Tradition*. Leiden: E. J. Brill, 1988.

Kessler, Eckhard, Charles H. Lohr, and Walter Sparn, eds. *Aristotelismus und Renaissance*. Wiesbaden: Otto Harrassowitz, 1988.

Kraye, Jill, W. F. Ryan, and Charles B. Schmitt, eds. *Pseudo-Aristotle in the Middle Ages: The Theology and other Texts*. London: The Warburg Institute, 1986.

Kretzmann, Norman, Anthony Kenny, and Jan Pinborg, eds. *The Cambridge History of Later Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1982.

Lohr, Charles. *Latin Aristotle Commentaries*. 2 vols. Florence: L. S. Olschki, 1988.

Moraux, Paul. *Der Aristotelismus bei den Griechen*. 3 vols. Berlin: Walter De Gruyter, 1973–2001.

Olivieri, Luigi, ed. *Aristotelismo veneto e scienza moderna*. 2 vols. Padua: Editrice Antenore, 1983.

Pade, Marianne, ed. *Renaissance Readings of the* Corpus Aristotelicum. Copenhagen: Museum Tusculanum Press, 2001.

Peters, Francis E. *Aristoteles Arabus*. Leiden: E. J. Brill, 1968.

Sandbach, F. H. *Aristotle and the Stoics*. Cambridge, U.K.: Proceedings of the Cambridge Philological Society, Supplement 10, 1985.

Schmitt, Charles B. *Aristotle and the Renaissance*. Cambridge, MA: Harvard University Press, 1983.

Schmitt, Charles B. *The Aristotelian Tradition and Renaissance Universities*. London: Variorum Reprints, 1984.

Sharples, Robert W. "The Peripatetic School." *Routledge History of Philosophy. Vol. II: From Aristotle to Augustine*, edited by David Furley. London and New York: Routledge, 1999.

Sharples, Robert W., ed. *Whose Aristotle? Which Aristotelianism?* Aldershot: Ashgate, 2001.

Sorabji, Richard, ed. *Aristotle Transformed: The Ancient Commentators and their Influence*. London: Duckworth, 1990.

Wardy, Robert. *Aristotle in China: Language, Categories, and Translation*. Cambridge, U.K.: Cambridge University Press, 2000.

Wisnovsky, Robert. *Avicenna's Metaphysics in Context*. Ithaca: Cornell University Press, 2003.

Taneli Kukkonen (2005)

ARISTOTLE

(384 BCE–322 BCE)

Aristotle was born in Stagira, a Greek colony in Macedonia. His father was physician to the Macedonian king, and the family had both a tradition of learning and connections to the Macedonian elite. At the age of seventeen Aristotle came to Athens to study in Plato's Academy (he may also have briefly studied rhetoric under Isocrates). The community of the Academy included some people who would stay for a few years to learn some philosophy before pursuing political careers in their native cities, and others for whom philosophy was an end in itself, and who might spend their entire lives in the Academy. Aristotle was one of the latter, and stayed in the Academy for twenty years, until Plato's death in 348, when Plato's nephew Speusippus succeeded him as head of the Academy, while the other most prominent Academics, Aristotle and Xenocrates, went to Assos in Asia Minor. There they seem to have formed a kind of local branch of the Academic community under the patronage of the tyrant Hermias of Atarneus, whose niece (and adopted daughter) Aristotle married.

Aristotle spent thirteen years around the north and east Aegean: in Assos; on Lesbos, where he did biological research; in Macedonia, as tutor to the future Alexander the Great; and in Stagira, where he is said to have given laws when it was rebuilt after the Macedonians burned it. He returned to Athens only in 335 (after the Macedonians had attained supremacy over Greece in 338, and after Alexander had succeeded his father in 336), not to the Academy, where Xenocrates had succeeded Speusippus, but to found his own school in the Lyceum, later called the Peripatetic school. He taught there until, after Alexander's death in 323, the Athenians revolted against Macedonia, and Aristotle was charged with impiety for a poem he had written that was held to have given divine honors to Hermias. He left Athens for family property in Stagira's mother-city, Chalcis on Euboia, where he died the following year.

With Aristotle, much more than with Plato, most of the preserved writings are closely connected with his teaching activity. Many of Aristotle's writings bear titles which remain the names of disciplines today (*Physics*, *Politics*, etc.), and much of Aristotle's work was either to introduce these disciplines into the Academy and its daughter communities, or to turn them from less systematic practices into systematically teachable disciplines. "Philosophy" in fifth–fourth century Athens meant simply "higher education," that is, whatever disciplines, beyond elementary education in gymnastics and "grammar" and "music" (including poetry), might be needed for someone who wishes to live well and to rule his city (or even his own household) well. For different teachers, this would cover different disciplines. For Isocrates, "philosophy" meant rhetoric. For Plato, to judge from the ideal curriculum of *Republic* VII, it meant mathematics (arithmetic, plane and solid geometry, astronomy, and "harmonics" or music theory) and dialectic (an art of regimented discussion, in which a respondent defends some thesis, typically a definition, and a questioner tries to refute it by yes-no questions leading to a contradiction); these are the means that will lead to knowledge of

what really and eternally *is*, and ultimately of divine things (the Forms and the Good).

Plato conspicuously leaves out rhetoric, which deals with mere opinions rather than with how things really are. He also leaves out pre-Socratic–style "physics" or "natural history," which he thinks is approximate and probable rather than precise and certain, and which explains things by placing them in a grand cosmogonic narrative of how things come to be, rather than (like mathematics and dialectic) by defining and demonstrating what things eternally are. Aristotle teaches all of these disciplines, without claiming that they are all equally scientific; he introduces a hierarchy of disciplines, from those accessible even to an aspiring politician with no great patience for philosophy, up through more strictly scientific disciplines, to the most demanding but most intrinsically rewarding philosophical wisdom.

Aristotle's introduction of rhetoric (probably already in the Academy) should be seen in the context of the conflict between the Academy and the school of Isocrates. Plato draws a sharp contrast between dialectic and rhetoric: that is, between using question and answer to refute a single respondent on a universal question and using long speeches to persuade a group about such particular questions as are discussed in meetings of a citizen assembly (deliberative rhetoric) or a jury (forensic rhetoric). Plato thinks that only dialectic is worthy of the philosopher. But rhetoric is the path to political success, and so students flock to Isocrates instead. Aristotle thinks that, however narrowly practical many students are, "we Academics," with our philosophical knowledge, ought to be able to educate them better than Isocrates can. (Aristotle is said to have justified his teaching of rhetoric by saying "it were shameful to keep silence and let Isocrates speak," varying a line of tragedy, "it were shameful to keep silence and let barbarians speak.")

This might be merely a practical compromise. More shocking to a Platonist is Aristotle's claim that "rhetoric is the counterpart of dialectic"—they are both, not sciences of any one subject matter, but sub-scientific abilities to discover and arrange and express arguments, applicable equally to any subject. Rhetoric also requires rudimentary knowledge of ethics and politics, because these are the subjects about which we must persuade, and because we must know how to project a given character or emotional state, and how an audience of given character and political background will react; the focus remains on argument.

Plato thinks that dialectic, by allowing us to arrive at definitions, gives us a scientific knowledge of eternal Forms existing apart from the sensible world. Aristotle, who has participated in the same dialectical practice as Plato, thinks this claim about its status is spurious. Dialectic is not scientific knowledge of eternal separate Forms, since there are no such Forms. Aristotle is willing to speak of forms present within sensible things (a form is whatever is the object of scientific definition), but dialectic is not scientific knowledge of these forms either, since scientific definition of (say) lunar eclipse depends on specific knowledge of the cause of lunar eclipse, which the dialectician, as a generalist, does not have; it is not the dialectician but the physicist who grasps forms of physical things. Dialectic remains a valuable preliminary training because, by showing what can be refuted, it rules out wrong definitions and helps us find the right ones, and because, by allowing us to find arguments on both sides, it sets out puzzles that science must solve, but it is not itself science or philosophy. And while Plato speaks not of *teaching* in dialectic, but only of a communal *practice* of questioning and answering, Aristotle demystifies the practice, and claims in his *Topics* to teach rules for discovering dialectical arguments, just as his *Rhetoric* teaches rules for discovering rhetorical arguments.

The average practically minded student will probably study only rhetoric and not dialectic, but Aristotle hopes to lure the better students on further to more scientific disciplines. In the first place, this means ethics and politics, which are philosophical, that is, scientific or causal discussions of what is good for individuals and cities, based on an understanding of what human beings and cities are. But Aristotle distinguishes these "practical sciences" from the "theoretical sciences," that is, kinds of knowledge valued purely for the sake of knowing them, which are capable of greater precision and are more intrinsically worth knowing, though less useful. Against Plato, physics or natural science (in the broadest sense, including biology and psychology) is a theoretical science: when done correctly, it grasps forms of natural things, and proceeds by definition and demonstration, but the forms it grasps are inseparable from matter and motion, and many of its results hold only "for the most part," or *ceteris paribus*, rather than universally.

Aristotle agrees with Plato that the highest wisdom, the knowledge most intrinsically worth knowing, must be a science of things existing eternally apart from matter, and ultimately of the Good. But neither dialectic nor physics is such a wisdom (nor is mathematics, which is not about separately existing objects, but about ordinary objects hypothetically idealized), and so Aristotle announces, beyond dialectic and physics, a new discipline

of “first philosophy” (what commentators of Aristotle since antiquity have called “metaphysics”), which will provide the theoretical wisdom that he thinks both Plato and the pre-Socratics have failed to deliver.

WRITINGS

We can broadly divide Aristotle’s writings into three classes:

> “Exoteric,” or “published” writings, were intended for circulation outside the circle of philosophers, elegantly written and sometimes in dialogue form (also the poem for Hermias and a similar poem for Plato). All such writings are lost, but there are substantial fragments; we have perhaps as much as half of Aristotle’s *Protrepticus*, or *Exhortation to Philosophy*, addressed to a royal patron, which remains an excellent introduction to Aristotelian philosophy. (Aristotle’s will is also preserved, in Diogenes Laertius.)
>
> Collections of data, classified but not written up with any literary pretensions, were intended as raw material to be further used in philosophical research and writing and teaching. These texts may have been “loose-leaf,” with new material constantly added, some of it perhaps by members of the school other than Aristotle. Extant writings of this type are the *History of Animals*, the *Constitution of Athens* (discovered in a papyrus in 1890 and not quite complete, a fragment of a vast series of 158 *Constitutions* of different cities), and the *Physical Problems*.
>
> “Acroamatic” writings, that is, writings related to Aristotle’s lectures, form the bulk of the surviving corpus. This does not mean that the texts are verbatim identical with the lectures; while Aristotle sometimes speaks as if addressing a live audience, that is compatible with the texts being notes written beforehand as a basis for lectures, or a later revision retaining the lecture style (as in published Gifford or Sather lectures), and the treatises contain many passages which no student then or now could endure if read verbatim as a lecture. The problem is not special to Aristotle; most Greek literature was intended for oral performance, and in every case it is difficult to determine how close the transmitted text is to any given performance. Performances would vary, and the written text is not a transcript of any one occasion but a model for varying expanded or abridged oral performances. In Aristotle’s case, while usually only one written version survives for each lecture series, occasionally (as in the *Ethics*) we can compare two and gain a sense of the range of variation.

The transmitted texts of the acroamatic writings vary greatly in style. Some passages are highly literary (often marked by avoidance of “hiatus”—the juxtaposition of a vowel at the end of a word with a vowel at the beginning of the next—as in Isocrates and in Plato’s late dialogues), whether because they have been more thoroughly revised toward eventual publication, or because Aristotle delivered some pieces (especially the beginnings of works) in more elaborate form, or because they are excerpted from Aristotle’s exoteric works. Other passages are long strings of brutally truncated arguments for the same conclusion, connected merely by “also”; in performance Aristotle would have selected only some arguments, and filled them out and connected them better.

The transmitted texts contain many references to “what we have said previously/elsewhere” or “what we will say,” sometimes with a title “in the [writings or lectures] on *x*.” (It is possible, but should not be the default assumption, that some of these cross-references were added by later editors.) While we can often supply a plausible page reference, we should beware of assuming that Aristotle’s references are to texts now extant, or else to lost parallel texts: They are not necessarily to fixed texts at all, but to earlier and later parts of an idealized curriculum, each part of which would be repeatedly given (with variations) as a lecture, and also written down and occasionally updated, even if no actual student ever heard the whole series in order. “We have said” and “we will say” refer not to order of composition but to order in the curriculum; however, while Aristotle is *mostly* consistent about the ideal order, there are contradictions that may indicate that he changed his mind on the appropriate sequence of the psychological-zoological writings. There is no real contradiction in the fact that Aristotle (and his followers) cite the same work under different titles; the curriculum may be subdivided more or less finely, and the same title may be used generically for a large section or specifically for a smaller subsection: “physics” or “on nature” may refer to the entire physical-biological corpus or to something as narrow as *Physics* I–IV (with *Physics* V–VIII cited contrastively as “on motion”).

Some ancient catalogs list Aristotle’s works by shorter units and some by longer units (the catalogs may also contain duplications, and some catalogs refer to works not available to other catalogs, or to us). Andronicus of Rhodes in the first century BCE attempted to introduce order by determining the correct titles and sequence, generally opting for longer “works,” and it is

probably roughly his decisions that won out. (But the story that Andronicus, drawing on a rediscovery of Aristotle's library, made the acroamatic works available for the first time and so touched off a renaissance of Aristotelianism, is mostly or wholly fiction.) Following a Stoic division of philosophy, Andronicus organized the corpus into first "logical" writings, then "physical" (or more broadly "theoretical" writings, to include the *Metaphysics*, concerned with nonphysical things), then "ethical" writings (or more broadly "practical," to include the *Politics*).

Many of the texts are now lost. As with the rest of Greek literature, what survived was generally only what was used and copied for educational purposes, which explains why the "exoteric" works are lost and why usually only one version of each "acroamatic" text survives. The surviving texts have been edited many times since the invention of printing, often in complete editions that generally try to follow Aristotle's and other ancient indications of the correct sequence of the corpus (although these are not fully consistent and, for example, give no hint how to order the three surviving ethical works, which all fill the same place in the curriculum).

Immanuel Bekker's nineteenth-century edition has become standard. Modern editions and translations give "Bekker pages" in the margins (e.g., "1042b5," where "a" or "b" is a column of a double-columned page), and Aristotle is always cited in this form where possible; editions that aspire to completeness print the texts in Bekker's sequence. The editions divide Aristotle's treatises into books and chapters; the book divisions have (not always undisputed) ancient authority and may in some cases go back to Aristotle himself, but the chapter divisions are modern artifacts and deserve no deference (medieval authors use a different division into "lectiones"). Ancient writers cite the books of multibook treatises by Greek letter-names; modern writers generally use numbers, but prefer letters in the *Metaphysics*, where the presence of two books alpha (conventionally designated A and α) disrupts the usual letter-number conversion.

The following list presents the texts in Bekker's sequence, leaving out texts currently agreed to be spurious, and marking with an asterisk texts whose authenticity is currently controversial. The traditional Latin titles are added where these sound significantly different from the English.

Logical writings (*Organon*): *Categories* (the title is controversial), *On Interpretation*, *Prior Analytics*, *Posterior Analytics*, *Topics*, *On Sophistical Refutations* (*De sophisticis elenchis*).

Theoretical writings: *Physics*, *On the Heaven* (*De caelo*), *On Generation and Corruption*, *Meteorology*, *On the Soul* (*De anima*), *Parva naturalia* (including *On Sense and Sensibilia*, *On Memory*, *On Sleep* [*De somno*], *On Dreams* [*De insomniis*], *On Divination in Sleep*, *On Length and Shortness of Life*, and *On Youth, Old Age, Life, Death and Respiration* [*De juventute* for short]), *History of Animals*, *Parts of Animals*, *Movement of Animals*, *Progression of Animals* (*De incessu animalium*), *Generation of Animals*, **Physical Problems*, *Metaphysics*.

Practical writings: *Nicomachean Ethics* (abbreviated "NE" or "EN"), **Magna Moralia*, *Eudemian Ethics* ("EE"), *Politics*. (In a peculiar situation, three central books are identical: NE V–VII = EE IV–VI. These books are usually printed with the NE, but most modern scholars agree that they were originally written with the EE instead.)

Bekker puts at the end the *Rhetoric* and *Poetics*, under the head of "productive philosophy" (i.e., philosophy to guide production, in this case of speeches or poems); their place is controversial, and they had sometimes been put at the end of the *Organon*. The **Constitution of Athens* and other texts not printed by Bekker (fragments discovered on papyrus or in later ancient citations or translations) are often placed at the end.

With Aristotle, as with Plato, there have been attempts to determine the order of composition of the works, distinguishing "early," "middle," and "late." Sometimes stylometric tests are applied. Some scholars, like Jaeger, assume that Aristotle moved from an early Platonism, to a critical revision of Platonism, to an independent mature philosophy. Such "developmental" studies have had the merit of bringing out tensions in Aristotle's work, and calling attention to works (often fragmentary, like the *Protrepticus*) that had been ignored or deemed spurious because they seemed embarrassingly close to Plato. Some chronological results have won widespread assent, notably that the *Protrepticus* is early, and the EE earlier than the NE. But dating has not generally been successful, and for a good reason, namely that Aristotle regularly revised his work, so that a single text may show both "early" and "late" features and thus resist easy classification. Aristotle was trying to present his treatises as parts of a synchronic system, ordered by pedagogical role; tensions remain, and while sometimes these tensions are best explained diachronically, this is not always the case. In what follows, the most important texts will be discussed, not in Bekker's order or in a presumed chronological order, but in roughly the order of increasing difficulty:

probably many of Aristotle's students dropped out early in this sequence, and only a few remained until the end.

ETHICS AND POLITICS

Aristotle conceives ethics as a part of political philosophy: We cannot understand and evaluate different political structures unless we understand individual character, and conversely, we cannot fully describe the best life for an individual without reference to the city in which he lives and is educated. Many comments in the ethical works assume that the reader or hearer is (or wants to be) a *politikos*, or statesman, and Aristotle assumes that the best life for an individual and the best *politeia* or constitution for a city, whatever they turn out to be, will be analogous. The ethical works, then, emerge from popular lectures to aspiring *politikoi*, who have come to hear lectures by a philosopher in the hope that it will make them happier and more successful *politikoi*, but who do not intend to spend their lives on philosophy.

Aristotle can be seen as trying to repair the damage that Plato did in his lecture on the Good, where an audience who had come expecting to hear about "health or wealth or some marvelous happiness" were surprised to find that the lecture was about numbers and that its conclusion was that the Good was the One, with the result that some of the audience gave up on philosophy altogether, while others presumably turned to the more practical philosophy of teachers like Isocrates (see Aristoxenus, *Elements of Harmonics* II,1, and cf. EE I,8). Aristotle is in part rejecting Plato's conclusions (he thinks mathematics has nothing to do with goodness), in part simply rejecting his method of presentation: we must start with what the audience antecedently believes and values, get them to see the difficulties, and so introduce philosophical doctrines (including any doctrine of a higher good) as solutions to those difficulties. But in ethics, as in rhetoric, he thinks that the Academics should be able to educate them better than Isocrates can.

Anyone who can choose how to live, and who wants to approach the question rationally, must first clarify what he is aiming at—what is the chief good of human life. Everyone agrees that the aim is *eudaimonia*—usually translated as "happiness," but perhaps best neutrally as "success"; it need not be introspectible, must be evaluated over a lifetime rather than at one moment, and can be said of cities as well as of individuals—but they disagree about what *eudaimonia* consists in. The three plausible contenders for the best way of life—the pleasure-seeking life, the active or political life, and the contemplative or philosophical life (Aristotle thinks the money-making life is chosen only from necessity)—go with different contenders for the human good. The pleasure seekers think it is pleasure; the *politikoi* may think it is fame or honor or, more appropriately, that it is *aretê*, virtue, or excellence (what *deserves* honor).

Among the philosophers, Socrates thinks that virtue (consisting in some kind of knowledge) is necessary and sufficient for happiness, and Plato talks about the Form of the Good or about the One. Aristotle creates an aporia by using these views against each other and raising objections against each, in order to motivate his own account of happiness, and the conceptual distinction on which it is based, as a solution to the aporia. Happiness or success in life is not virtue, which is a stable *hexis* ("habit" or acquired state) persisting even when it is not exercised, but rather the *energeia* (exercise or activity) of virtue throughout a complete lifespan. We can thus avoid the paradox of saying that the good person is happy even when poor, sick, and unjustly despised by his fellow citizens; in such a condition he remains virtuous but cannot exercise his virtue, or is greatly hampered in exercising it. The happy life will *involve* virtue, and it will also involve wealth if the virtues (say, generosity) need wealth to be exercised, and these facts explain the temptation to identify happiness with virtue or even wealth. Likewise, the happy life is pleasant, since Aristotle analyzes pleasure as being (or following upon) the exercise of a natural state, but its pleasantness is not what *makes* it happy or worthy of choice. (This is against the view of some Academics that pleasure is always a process, the *restoration* of a natural state, and that the happiest life is a steady natural state without deficiency or restoration. Aristotle avoids the paradox that the happiest life is without activity or pleasure by arguing that there are *energeiai* that are not processes.)

Aristotle applies the same method of setting out competing beliefs and arguments, resolving the aporia through a distinction, and showing how justice can be done to all sides, to resolve Socrates' paradoxical argument that incontinence is impossible: I can do something wrong if I have *hexis*-knowledge that this type of action is wrong, but not if I am applying the *hexis* and have *energeia*-knowledge that this particular action is wrong. It must be stressed that this is a *teaching* method, designed to motivate Aristotle's doctrines and conceptual distinctions for his audience and to make softened versions of Socratic or Platonic paradoxes more palatable. We do not know that this is how Aristotle himself arrived at his conclusions.

Aristotle also tries to show what is right in the Socratic and Platonic conclusions that virtue and happiness consist in knowledge, perhaps knowledge of a transcendent Good. The work or task or function (*ergon*) of a human being is rational activity, and a virtue is a condition that disposes to such activity. But there are two kinds of virtues: "intellectual virtues," or virtues of the rational soul, and "moral virtues," conditions of an irrational part of the soul, according to which it is disposed to act as reason would require.

Genuine moral virtue is not simply habituation to desire the right amount, but involves choice, which involves deliberation or means-end reasoning: so moral virtue is not possible without the intellectual virtue of *phronêsis* ("prudence," "practical wisdom") or deliberative ability. (Nor, conversely, is *phronêsis* possible without moral virtue, since uncontrolled passions will warp our deliberations.) But *phronêsis* is not identical with the highest intellectual virtue, *sophia* ("wisdom"), knowledge of the divine things that are intrinsically most worth knowing. *Sophia* is exercised only in contemplation (*theôria*) and not in action: we cannot deduce, from these necessary eternal things, knowledge of the contingent temporal objects of practical choice. But *sophia* gives a starting point for deliberative reasoning in another way, because contemplation is itself the exercise of the highest virtue, and is therefore the highest happiness we can try to achieve. (It is the only exercise of virtue we can attribute to the gods, who can hardly be courageous or temperate.) So while happiness is possible with only moral virtue and practical intelligence, the highest happiness needs theoretical intelligence as well.

When Aristotle says that maximizing contemplation is the highest goal of human planning, he means not only planning an individual life, but also a statesman's planning for the city. (The statesman may have only *phronêsis*, but needs proper respect for *sophia*.) Happiness, for cities as for individuals, is an exercise of virtue, and while this may require material conditions (prosperity and external peace), the statesman's main concern should be making the city virtuous. And, for cities as for individuals, some virtues are more worth exercising than others: courage and military solidarity are virtues we would rather not have occasion to exercise. While a city must be able to defend itself, its highest goal is the exercise of the virtues of peaceful leisure. This is *theôria*, not only in its metaphorical sense (the philosopher's contemplation of nature or of incorporeal divine things), but also in its ordinary sense: attendance at civic religious festivals, including the musical-poetic contests (of tragedies, comedies, etc.), which may be occasions for private or communal moral and political reflection. (The *Poetics* defends the value of such musical-poetic performances, and inquires how it comes about; it thus elaborates an important point too briefly treated in the *Politics*.)

Aristotle's main goal in the *Politics* is the construction of an ideal *politeia* (constitution or collective mode of life and governance), a critical revision of Plato's *Republic* and *Laws*. But he also discusses less ideal *politeiai*, how they are preserved by proper legislation, and how they are corrupted, leading to revolution; the trained *politikos* will be useful even to a non-ideal *politeia*, helping to preserve it by moderating and improving it. (And Aristotle's 158 collected *Politeiai* will help give an empirical base.) The central thesis of the *Politics* is the distinction between genuinely political rule (rule over free fellow citizens, in the interest of the ruled) and despotic rule (rule as of a master over slaves, in the interest of the ruler). While *Politics* I is notorious for defending slavery, Aristotle's main interest is to make clear the differences between despotic rule (legitimate only within the household) and political rule. (He thus also defuses the Socratic paradox that there is only one art of ruling, depending on philosophical knowledge of the good: to the Athenian bourgeois, this suggests that the Academics are claiming the right to rule over their fellow citizens, while not allowing ordinary citizens even to give orders to their servants.)

Within the city, not only tyranny but also oligarchy and democracy are despotic: even when they are ruled by law, their laws express the economic and political interests of a ruling individual or group (the rich few in an oligarchy, the poor majority in a democracy). But rather than conclude, with Thrasymachus, that all rule is despotic, Aristotle argues, with Plato, that genuinely political rule is possible. Officially (like Plato's *Statesman*) Aristotle has a two-by-three grid of constitutions: corresponding to tyranny, oligarchy, and democracy are three good constitutions, kingship, aristocracy, and (what Aristotle calls in a narrower sense) *politeia*, which are the rule of the one, the few, or the many in the interest of the whole city. But actually Aristotle treats kingship and aristocracy as an ideal constitution run by morally and practically virtuous people and aiming at the development and exercise of virtue; *politeia* is a more attainable ideal, a "mixed constitution" between democracy and oligarchy, as the moral virtues lie between vices of excess and deficiency. *Politeia*, though a "virtuous" constitution, does not aim at virtue in the citizens and does not choose officials for their virtue, but at least its laws, balancing the

interests of different groups and designed to preserve peace between them, do not impose a partisan "justice" that would conflict with genuine moral virtue in the individual.

DIALECTIC AND ANALYTICS

Aristotle's logical treatises are usually grouped as the *Organon*, or "instrument"; against the Stoics, who make logic a part of philosophy alongside physics and ethics, the Peripatetics say that logic is a mere instrument of philosophy, valuable neither intrinsically nor as guiding action, but only as guiding reasoning in other fields. Also ancient is the arrangement of the *Organon*: first the *Categories*, dealing with single terms and the simple objects (substances, quantities, qualities, relations, actions, passions, "where," "when," positions [e.g., standing], states [e.g., armed]) that they signify; then the *De interpretatione*, dealing with propositions composed of two terms linked by a copula (affirmative or negative, universal or particular, assertoric or modal); then the *Prior Analytics*, dealing with syllogisms, valid arguments composed of three propositions sharing three terms (e.g., "A belongs to no B, C belongs to all B, therefore A does not belong to all C," valid since Aristotle rejects empty terms).

Then come treatises dealing with different types of syllogism: the *Posterior Analytics*, with scientific or demonstrative syllogism, where the premises must be true and causally explanatory of the conclusion; the *Topics*, with dialectical syllogism, where the premises need only be plausible; the *Sophistical Refutations*, with sophistical or pseudo-dialectical syllogisms, which are only apparently valid or have only apparently plausible premises; some ancient writers add "rhetorical" and even "poetic" syllogisms. At the end of the *Sophistical Refutations*, Aristotle says that while he has perfected earlier teaching of rhetoric, in the case of the syllogism there had been no such teaching before him; Aristotle has been taken as here summing up the *Organon* and reflecting on his crucial discovery, the syllogism.

However, Aristotle has no conception of "logic," but of two different disciplines, analytics (*Prior* and *Posterior*), and dialectic (the *Topics*, taken as including the *Sophistical Refutations*); the *Categories* and *De Interpretatione* seem designed to support the *Topics* rather than the *Analytics*. We have spoken above of dialectic, the practice of regimented discussion in which a questioner seeks to refute a respondent's thesis by a series of yes-no questions. The end of the *Sophistical Refutations* is summing up not the entire *Organon* but only the *Topics*, which has for the first time made dialectic a teachable art and has shown how to discover syllogisms to deduce the contradictory of the respondent's thesis.

These arguments, unlike rhetorical arguments, can proceed only from premises the respondent will grant, and by steps he must accept as valid. Dialectic must proceed from plausible (*endoxa*) premises, since these are just those premises that a respondent will concede (if he does not see that they favor or hurt his thesis). It is a mistake to turn dialectic into "argument from prereflective intuitions," detached from the context of refutation, and to give it a foundational role in philosophy. Aristotle does say that dialectic gives a path to the principles of the sciences, but these principles are, especially, definitions, and, as in Socratic dialogue, dialectic is chiefly devoted to testing and refuting proposed definitions. The structure of the *Topics* brings this out: successive books give rules for testing claims that *P* belongs to *S*, that *P* is or contains the genus of *S*, and that *P* is proper (*idion*) to *S* (i.e., belongs to every *S* and no non-*S*), which are necessary but insufficient conditions for *P* to be the definition of *S*, and then give special rules for testing claims of definition. Aristotle also gives advice on how to order your questions, how to proceed as respondent, and background knowledge the dialectician should have.

The *Categories* and *De interpretatione*, as well as *Topics* I, seem to give such background knowledge; the most recent edition of the *Categories* prefers the alternative ancient title *The Before the Topics*, in part because the text is not just about categories. First, Aristotle distinguishes simple from complex expressions; then, what is signified by a simple expression is signified either synonymously (univocally) or homonymously (equivocally) or paronymously (denominatively). Two things are synonymous if they are signified by the same name and according to the same definition; homonymous if signified by the same name according to two different definitions (bank and bank, but also *mousikê*, the art of music, and *mousikê*, a female musician); paronymous if one name is derived from the other ("just" is paronymous or derived from "justice," not because the word "justice" is older, but because something is called "just" *because* there is justice in it).

Only synonymous things, not homonymous or paronymous, can be given genus-differentia definitions ("just" is neither a species of animal nor a species of virtue). Synonymous things that are in a subject (like justice) fall under one of the nine categories of accidents; synonymous things that are not in a subject are substances. (Substances can be "said of" something, but cannot be "in" something: horse is said of Bucephalus, since

Bucephalus is said to be a horse, but there is not a horse *in* Bucephalus. "Primary substances," like Bucephalus, are neither in anything nor said of anything.) Aristotle gives tests for when a thing falls under each category, which are needed to apply the rules of the *Topics* (thus if *P* is the genus of *S*, *S* and *P* must belong to the same category, but we need tests to determine to which category they belong). Likewise, after the categories proper, Aristotle gives accounts of the different kinds of opposition, priority and simultaneity, motion and having, which serve similar functions in dialectic.

Sometimes a dialectical questioner poses a series of questions that appear to necessitate the contradictory of the respondent's thesis, but which contain some hidden fallacy; the respondent must avoid assenting to what does not follow, and must be able to explain *why* it does not follow, in order to avoid appearing, to the spectators and perhaps even to himself, to have been refuted. The *Sophistical Refutations*, which may be considered as a final book of the *Topics*, is devoted to classifying such "sophisms," or "sophistical refutations," explaining how each type arises, and advising the respondent on how to recognize and to "solve" or "resolve" each such sophism as it comes at him in questioning.

Sophisms are not intrinsically dishonest: They are puzzles demanding solution. We should imagine, not an arms race between sophists devising offensive weapons and philosophers improving defenses, but a single intellectual community exploring sophisms and discussing the merits of different possible solutions. Often the most philosophically interesting sophisms are "sophisms of figure of speech," arising when the grammatical form of a term misrepresents its logical form: these include the family of sophisms concluding that "there is a third man" beyond mortal individuals and the Platonic Form, which turn on treating "man" as "signifying some this." Aristotle himself, in the fragmentary *On Ideas*, constructs a series of such philosophically serious sophisms, giving for each Platonic argument for the Forms a parallel argument to an unacceptable conclusion, such as the third man. Each sophism challenges the Platonists: "dismantle my sophistical argument without at the same time dismantling your own allegedly probative arguments for the Forms." The *Categories* helps solve sophisms of figure of speech by testing what category each term signifies, and its distinction between primary and secondary substances can solve many third man sophisms; but if Platonists accept these solutions, they risk undermining their own favorite arguments and conclusions.

A syllogism or deduction is "a discourse in which, some things being supposed, something different results of necessity through their being so." Syllogisms are as old as thought and language, and Aristotle does not claim to have invented them. What the *Analytics* invents is a method for *analyzing* them: that is, for classifying them and then, by giving a few primitive argument forms and derivation rules for generating more complicated forms, explaining why syllogism comes about. In every case, syllogism depends on two premises sharing a common term (the syllogism will be in different "figures," depending on whether the shared term is subject of one premise and predicate of the other, predicate of both, or subject of both; some "moods" will be valid and others not, depending on whether the premises are affirmative or negative, particular or universal, assertoric or modal). Aristotle's analysis depends on the realization that the necessity or validity of an argument, once all premises are made explicit, depends only on its form, so that the same analysis applies whether the premises are true or false; this realization presumably arose from the deliberate exploration, in dialectic, of the consequences of false hypotheses.

But Aristotle sharply distinguishes dialectical from scientific or causal reasoning, and he devotes the *Posterior Analytics* to analyzing "demonstrations" or *scientific* syllogisms, arguments that give their possessors knowledge or science (*epistêmê*) of some object; here *epistêmê* is a cognitive state that not only grasps an object as it is, without the possibility of falsehood, but also understands *why* the object is as it is. It seems surprising that mere arguments, without direct contact with the object, can give such knowledge, and Aristotle tries to analyze the conditions under which this can happen. The premises must be true, necessary, and better known than the conclusion; they must also express the causes that explain why the conclusion is true. We can of course come to know an object by reasoning from effects to causes, but properly scientific and explanatory knowledge must reason from causes to effects; the logical structure of the argument will mirror the causal structure of the world.

On pain of circularity or regress, the first principles of demonstrations must be known by some means other than demonstration (Aristotle calls the non-demonstrative grasp of first principles *nous* rather than *epistêmê*). Apart from some topic-neutral principles of reasoning ("axioms"), these will be either "hypotheses" that the objects of each science exist, or "definitions" of those objects; we accept without demonstration both the existence and definitions of the simple objects of the science (e.g., for geometry, point and straight line) and pre-

liminary definitions of the complex objects (e.g., regular pentagon), but we demonstrate the existence of complex objects satisfying those definitions. Dialectic can reach these preliminary definitions, but we can give properly scientific definitions of complex objects only once we demonstrate their existence from simple causes (thus not "thunder is noise in the clouds" but "thunder is noise of extinction of fire in the clouds"). We cannot give justificatory explanations of how we know the first principles of the sciences, but only causal explanations of how the human mind, primed by experience, comes to grasp them by *nous*; Aristotle's account is compressed enough that it has been read both as an empiricist account of induction and as a friendly revision of Plato's theory of recollection.

Aristotle's account of science is clearly modeled on geometry. But he tries to show that physics too can be a science, beginning from a grasp of the forms of natural things.

PHYSICS AND COSMOLOGY

Aristotle's project in physics is a response to Platonic challenges both to the narrative method and to the content of pre-Socratic physics. Anaxagoras's physics—to take a typical pre-Socratic example—narrates the origin of everything from a cosmogonic vortex, whose rotation and centrifugal force explain the separation of heaven from earth, the rotation of the heavens, the motion of heavy bodies down and light bodies up and the sorting of like bodies to like, and then the formation of the first plants and animals and humans out of seeds present in the precosmic mixture. Plato thinks such narrative can never be scientific; science must be concerned not with how things come to be but with what they are, beginning from their forms as grasped by definitions, and proceeding to demonstration.

Plato also complains that pre-Socratics explain the emergence of the cosmos by reference not to a rational plan or to some good to be accomplished, but through violence; if things are where they are because of a vortex (i.e., through being shoved by other bodies that are shoved into them) rather than because it is best for them to be there, then there will be no explanation of the goodness and orderliness of the universe, as manifested in the mathematically precise motions of the planets. In the *Timaeus*, Plato addresses the second objection by sketching an alternative teleological physics; but this too follows a narrative method, and even a reformed physics cannot be science but only a likely story.

Aristotle tries to address both objections and to produce a genuinely scientific physics, explaining the physicists' traditional explananda (rotating heavens, fall of heavy bodies, lightning, earthquakes, animals …) not in a narrative sequence but in a causal or explanatory sequence, beginning from the form or nature of each body, which is the object of a properly physical definition. Aristotle broadly accepts the *Timaeus*'s picture of the cosmos: a spherical earth is at rest at the center of a single spherical cosmos. The cosmos is made of earth, water, air, and fire intertransformed and combined, teleologically organized to support living things, and surrounded by heavenly bodies that are themselves living and divine; these move in several uniform circular motions, which combine to produce complex astronomical phenomena, and they are ultimately governed by an incorporeal god or gods. But Aristotle's method contrasts with the *Timaeus*, and leads him to challenge particular claims of the *Timaeus* as well as of the pre-Socratics.

Aristotle's particular physical treatises—the *De caelo*, *On Generation and Corruption*, *Meteorology*, and psychological-zoological writings—follow roughly what had been the traditional narrative sequence of the explananda. Thus the *De caelo* treats the rotation of the heavens and the motions of heavy and light bodies, traditionally explained through a cosmogonic vortex. But Aristotle rejects explanations through vortices or any other violent cause. What happens to a thing violently, contrary to its own nature, cannot happen always or for the most part, but only as a temporary interruption of a thing's natural behavior (e.g., a stone being thrown upward). Physics, as a science, seeks to explain what happens always or for the most part, and must therefore start by grasping the nature of each thing, where "nature" means "principle of natural motion"; so the nature of heavy bodies is to move toward the center of the cosmos, and thus teleology is built into each nature. (Thus physical definitions necessarily involve motion, and the forms they describe cannot exist separately from matter, as the Form described by a Platonic dialectical definition is supposed to. And fire and so on must be defined physically by their motions, rather than mathematically by their shapes as in Democritus and the *Timaeus*.)

Aristotle draws the conclusion that the heavens cannot be made of the four standard elements; since these naturally move in straight lines toward or away from the center, the heavens would have to be constrained to circular motion by violence (whether by a vortex or by a providential soul as in the *Timaeus*), and such motion could not be regular or permanent. Consequently, the heavens are made out of a fifth element (sometimes called "aether") whose natural motion is around the cen-

ter. The aether is free of the accidents that obstruct natural motion in the sublunar world, so it rotates eternally without interruption or irregularity. Because this motion arises eternally from the nature of the thing, Aristotle rejects the claim of the pre-Socratics and the *Timaeus* that the rotation of the heavens and the separation of the elements into an ordered cosmos arose (by what could only be a violent process) from a precosmic chaos; the ordered world and its more-or-less regular phenomena have always existed, and a narrative explanation is excluded, since there is no precosmic situation from which a narrative could begin. Rather, the phenomena must be explained by the influence of the naturally rotating heavens on naturally moving sublunar elements.

The *On Generation and Corruption* and *Meteorology* continue this program. If it were not for the rotation of the heavens, the four sublunar elements would separate out into concentric spheres of heavier and lighter elements, with no intertransformation or combination, and therefore no living things. But the regular daily rotation of the heavens, combined with the regular rotation of the sun through the inclined circle of the ecliptic, bring it about that the sun is above the horizon more of the time in the summer than in the winter, causing regular cycles of heating and cooling, and thus of evaporation and condensation.

Aristotle sees evaporation as a genuine transformation of water into air, and likewise of earth into fire; when a heavy element is transformed into a light element, it begins to rise (and when a light element is condensed, it falls), and this cycle keeps the elements from separating and gives rise to combinations. But properly the light elements are not "air" and "fire" but "moist exhalation" and "dry exhalation"; air is a mixture of both, and the portion of the dry exhalation that gathers above the air and beneath the sphere of the moon is not actually fire, but is a fuel that easily becomes inflamed, as it does in comets and shooting stars.

Since Tycho Brahe proved that comets and novae are supralunar, Aristotle's account has been regarded as a desperate attempt to save his theory of immutable heavens by moving all changes in the heavens to a fictional sublunar fire-sphere governed by a fantastic exhalation process. Historically this is the wrong attitude. Aristotle's explanation of comets is among the most traditional parts of his physics: Heraclitus explains even the sun through a continuous process of exhalations rising from the sea and becoming inflamed. Aristotle's innovation is to separate out from meteorological phenomena genuinely astronomical things like the sun, which are not dependent on the sublunar world but are governed only by themselves and by unchangeable incorporeal things, and therefore have eternally constant motions and can be objects of precise mathematical science; it is only because these things are perfectly regular that they can impose even an approximate regularity on the sublunar world.

The *Physics* in the narrower sense is a deliberately non-cosmological prolegomenon to the physical works, describing the principles from which all natural things arise and the necessary conditions (above all, motion) for anything to arise from these principles, and using a definition of "nature" to delimit the physicist's domain and methods and the causes or explanations that he must invoke in tracing natural things back to their principles. Aristotle begins, traditionally enough, with the *archai*, the principles or starting points of natural things—whatever must exist before each natural thing comes to be, and can be used in explaining it. (For narrative physics these would be whatever existed before the cosmos, e.g., for Empedocles the four elements and love and strife, for the *Timaeus* the Forms and receptacle and demiurge; but Aristotle's *archai* do not exist before the cosmos, since his cosmos never came to be.) We will infer to the *archai* by analyzing the characteristic effect that arises from them, which is, most generally, motion or change—not only change of place (locomotion) but also change of quality (alteration), change of quantity (growth and diminution), and the coming to be and passing away of substances (generation and corruption).

Aristotle argues that whenever some new *F* comes to be, in any category, there must be some persisting substratum that was not *F* and comes to be *F*; this analysis shifts *F* to predicate position. The subject that persists through even substantial change is one *archê*, the matter. This echoes the *Timaeus*'s argument that the apparent substantial change of (say) water into air shows that the real *archê* is not water or air, but the receptacle, the persisting substratum that appears now watery, now airy. But the *Timaeus* seems to infer that the change is not really substantial, that all sensible things are just accidental modifications of this single persisting substance. Aristotle argues that there is real substantial change, that the substance of a natural thing is not the matter that persists through the thing's generation and corruption, but the form that comes to be in the matter. Both form and matter are *archai* of natural things, and while the matter is *potentially* this or that substance, the form, as what makes each substance *actually* that substance, is substance in a stronger sense. (Plato would reply that while form as well as matter is an *archê* and a substance, the real Form is

eternal and separate, and what comes to be in the matter is a nonsubstantial image of the Form.)

How do we tell when a form acquired through change is a new substance, and when it is merely a new accident of a persisting substance? The shape of an artifact is merely an accident, but the *nature* of a natural thing, that is, the distinctive "principle of motion and rest" within it that is responsible for its carrying out its characteristic activities, is a substance. *Physics* II argues that the nature of a natural thing is more properly its form than its matter, and therefore that the physicist must study form as well as matter; thus, as we have seen, physics must define and not merely narrate, giving definitions that, unlike Platonic dialectical definitions, are inseparable from motion and thus from matter—natures are like "snubness," which is neither the matter "nose" nor the form "concave," but a form that cannot be defined without reference to its appropriate matter, the nose. A natural thing acts *for the sake of* actualizing the characteristic potentialities of its nature, and so the physicist will give explanations not only through the material and formal causes and through the mover or efficient cause, but also through the final cause. Aristotle thus, like Plato in *Laws* X, argues against many pre-Socratic physicists that purposive activity is prior to chance and violence, but he does this while preserving what is specific to nature, and without reducing natural things to artifacts of a designing soul.

Nothing will arise from matter and form without motion; motion depends on time and place and (some people think) on void; also a motion, to be a single motion, must be continuous, and continuity implies infinite divisibility. All these concepts are problematic, and Aristotle tries to define, and to resolve *aporiai* about, motion, place, and time, and to show that the infinite and the void do not exist (except in specially qualified senses). He then turns to the "*On Motion*," *Physics* V–VIII (*Physics* VII seems to interrupt the argument, and may be a survivor of an earlier stage of Aristotle's work). *Physics* V–VI give non-causal considerations that would apply equally to natural and violent motion, notably about when a motion is a single motion, about when two motions or a motion and a rest are contrary, and about the continuity of motion, place, and time; they seem to be there chiefly to supply premises for the causal argument of *Physics* VIII. *Physics* VIII, relying only on the abstract concepts of the *Physics* and not on empirical cosmology, gives an elaborate argument from the natural motions of corruptible things, first to the eternity of motion as such, then to self-moved movers (empirically, animals) and unmoved movers (their souls), then to an eternally continuous motion (the motion of the heavens), and finally to an eternally unmoved cause outside the cosmos. This bravura display reaches beyond physics to metaphysics or theology, and Aristotle relies crucially on it in *Metaphysics* Λ, discussed below.

PSYCHOLOGY AND ZOOLOGY

Narrative physics typically ends with the production of plants and of animals, including humans, before turning to human societies and conventions, which Aristotle treats under practical philosophy. Aristotle devotes a large part of his writing to animals, complemented by Theophrastus's studies of plants. But his program of denarrativizing physics, and of physical teleology and physical definition, entail major differences from earlier accounts of animals; Aristotle also integrates an account of soul into his study of animals, though not as fully as we might expect. The crucial methodological texts are *Parts of Animals* Book I, which serves as an introduction to the zoological works generally, and *De anima* I,1.

A narrative physicist believes he has accounted for the elephant once he has taken the cosmogonic narrative far enough to generate the first elephant. This means that he puts his "Generation of Animals" before his "Parts of Animals." (The parts of an elephant are simply whatever results from the prior generative process: Thus the Hippocratic *On Fleshes* gives a cosmogonic account of the generation of each tissue, with no regard to how the tissues are arranged in the animal, what animal they are parts of, or what functions they have.) Such a physicist will also be more concerned with the hard problem of the "spontaneous" (nonsexual) generation of the first elephant than with the easier problem of how to get more elephants out of the elephants there already are.

For Aristotle, however, the whole cosmos with all its species has existed from eternity, so there is no reason to believe elephants were ever generated spontaneously. We never see elephants generated spontaneously anymore, and while nature might have had greater generative force at some past time when it was undergoing more violent motions (see *Physical Problems* X,13), when we understand the extremely complex arrangement of parts required for a functioning, self-sustaining elephant, it becomes incredible that the crude natural powers of the pre-elephantine era could have combined to produce it. (Plato might say that God intervened to produce the first elephant, but Aristotle thinks that God acted no more or less then than now, and that his activity simply sustains the regular activities of natures. While Aristotle is now

notorious for defending spontaneous generation, he actually allows less scope to spontaneous generation than any other Greek philosopher, restricting it to lower life-forms.)

Thus when Aristotle studies the generation of living things, he is chiefly studying their generation out of already existing members of the same species. And we can understand this process not in narrative sequence but only backward, starting from the arrangement of parts that the generative process is for the sake of producing; so methodologically the *Parts of Animals* must precede the *Generation of Animals*. And the parts themselves must be explained teleologically, not through the generative process but through their function in the animal. Different species of animals will have different strategies for survival and reproduction, thus different characteristic activities, requiring different characteristic parts; the scientist will define each animal species by describing its characteristic parts, defining each part as an "organ" or instrument of some activity and deducing its shape and matter from its function.

Aristotle describes the parts, and the whole animal, as organs of the *soul*, that is, instruments through which the soul's powers are exercised. Because they cannot be defined without reference to the soul, it belongs to the natural scientist to study soul, or at least those powers of soul that are exercised through bodily instruments—all powers except, possibly, intellect (*nous*). Aristotle is trying here both to reform physics by making it include the soul, and also to make the study of soul scientific by bringing it under physics. However, he also makes the study of soul *further* from physics as usually conceived, by denying that the soul is moved, either in moving the body or in sensing and thinking. In *De anima* I he says that earlier philosophers have approached the soul either from its capacity to originate motion in the body, concluding that it is a self-moving source of motion; or from its ability to represent all things, concluding that it is composed of the elementary constituents of all knowable objects; or from its "bodilessness," identifying it either with fire or air or with something entirely incorporeal.

The *Timaeus* combines all of these approaches but, Aristotle thinks, in a mistaken way, representing the soul as a magical quasi-body interwoven with visible bodies, moved in the same way that bodies are, and moving bodies and being moved by them in the same way that bodies move each other. In *De anima* II, Aristotle instead defines the soul by its relation to its *energeiai*, the activities it carries out through the body. Soul is the *dunamis* (power, potentiality, capacity) for these *energeiai*, or it is that which, added to a potentially living thing (a seed or embryo), makes it an actually living thing, where to be an actually living thing is to have the potentiality to carry out an appropriate range of the vital activities (nourishment, growth, reproduction, sensation, memory, imagination, desire, locomotion, intellection). In Aristotle's formula, soul is "the first actuality [*entelecheia*] of a potentially living body," the second actuality being the vital activities; soul stands to these activities as a *hexis* of science stands to the exercise of that science in contemplation, or as a productive art stands to its exercise in production.

Aristotle spells out his definition by saying that soul is "the first actuality of a natural organic body." Modern connotations of "organic" are misleading here: an organic body is an *instrumental* body, as is, for example, a hammer; the living body is the instrument of the soul as the hammer is an instrument of the art of carpentry. But the hammer is an *artificial* organic body, while the living body is a *natural* one, meaning (by the definition of the *Physics*) that it has an internal principle of motion and rest. So while the art of carpentry moves the hammer from outside (by inhabiting the body of the carpenter), the soul is a *nature* moving the body in a quasi-artistic way from inside, in producing and maintaining its natural instrument (nutrition, growth, reproduction) and in further using that instrument (sensation and the higher activities). The arts give us a model for how the soul can move its body without itself being moved (unlike a body pushing or pulling another body): though the carpenter's hand is moved when he moves the hammer, his art of carpentry is not. The arts also give a model for the cognitive powers, since an art contains the "formula," the definition or perhaps recipe, of its objects, without containing their matter; and arts can recognize individual objects through cognitive instruments (the art of measuring might use scales), as well as moving them through instruments of action.

The vegetative powers (powers shared even by plants) and the sensitive powers (powers shared by irrational animals) are "not without" their appropriate bodily instruments, as snubness is "not without" nose. So souls of plants and irrational animals cannot exist when separated from their bodies. The question whether any soul can so exist, and thus whether any soul is immortal (besides the souls of the heavens, which have immortal bodies), depends on whether all psychic powers are similarly dependent on bodily instruments. Sensation is not without its instruments, and imagination is not without sensation, so these are inseparable.

Some passages in *De anima* III suggest that intellection is not without imagination, so that it too is inseparable; other passages suggest that a special kind of intellection, of special matterless objects, is separable. (Fragments of Aristotle's "exoteric" works also argue that soul is immortal; perhaps Aristotle changed his mind from these early texts to the *De anima*, perhaps the texts can be reconciled, or perhaps the "exoteric" texts should be regarded as a popular approximation to a more precise truth.) *De anima* III,5 says that "the passive *nous* is corruptible," and that only the active or productive *nous* is immortal. But what is this productive *nous* and what does it do? Since it is eternally and essentially intellectually cognizing, it seems that it must not be a part of the human soul, but rather a separate immaterial divine thing that acts on the "passive *nous*" in the soul. This recalls Platonic texts on *nous* (here best translated as "reason" or "rationality") as a separately existing virtue in which souls participate, the *nous* apparently personified as the divine craftsman of the *Timaeus*. Aristotle rejects all other separately existing virtues, because they are "not without" the irrational soul and the conditions of the body, but he has no reason to reject this one; and he too in *Metaphysics* Λ will identify such a *nous* with a world-governing divine *archê*.

For Aristotle, we can fully understand soul only by understanding its specific powers, their activities, and the objects and instruments of those activities; the *De anima* gives a general abstract account, which is filled in by the *Parva naturalia*, which treats of the actions and passions "common to soul and body"—and almost all the soul's actions and passions are in common with the body—and by the accounts of the instruments and activities of different animal species in the zoological works. But the neat sequence of "psychological works" (*De anima* and *Parva naturalia*) followed by "zoological" or "biological" works (the *History*, *Parts*, *Movement*, *Progression*, and *Generation of Animals*), as presented in Bekker and other modern editions, is probably an illusion. The texts themselves frequently refer to what has preceded or what will follow, but they seem to indicate two different sequences. Some texts, especially the *Parts* and *Generation of Animals*, imply a sequence in which the *Parts* would lead immediately into the *Generation* (both presupposing the *History*, as giving the facts for which they will supply the causes); the *De anima* and *Parva naturalia* would be a separate sequence, if anything more likely to come after than before (Aristotle refers to a lost part of the *Parva naturalia*, on the principles of health and disease, as the end point of natural philosophy).

But other texts imply a different order. Call "*Parva naturalia* Group I" the treatises connected with sensation, the *On Sensation*, *On Memory*, *On Sleep*, *On Dreams*, and *On Divination in Sleep*; "*Parva naturalia* Group II" would be the *On Length and Shortness of Life*, *On Youth, Old Age, Life, Death and Respiration*, and the lost treatise on health. There are many indications for a sequence *Parts of Animals*, *Progression of Animals*, *De anima*, *Parva naturalia* Group I, *Motion of Animals*, *Generation of Animals*, *Parva naturalia* Group II, and perhaps a treatise on plants. It seems most likely that Aristotle began with the *Parts-Generation* sequence, and later inserted the other texts between the *Parts* and *Generation*, treating reproduction, like sensation and breathing, as an activity involving soul as well as body. No evidence supports putting the *De anima* before the *Parts of Animals*; one option is to regard biology as beginning with the body, turning to the soul, and then exploring how they act together.

METAPHYSICS

Sophia as an intellectual virtue—"*epistêmê* and *nous* of what is most noble by nature"—had been discussed in the *Ethics*. In the *Metaphysics*, Aristotle tries to provide a new discipline to bring us to this virtue, because he thinks that the existing disciplines with a claim to yield theoretical wisdom—physics, mathematics, and dialectic—are insufficient. The awkward title, literally "The [books or things] after the physical [books or things]," first attested in Nicolaus of Damascus (1st century CE), reflects the difficulty of fitting the treatise into the standard scheme of disciplines: it belongs to theoretical philosophy, and draws on physics, but does not belong to physics, because the divine things it considers (unlike the heavens, also divine) exist separately from matter and motion.

The unity of the treatise is problematic. It is clear that Aristotle intended to write a long treatise on *sophia*, and that most of the books of the *Metaphysics* were intended as materials for such a treatise. But it is also clear (from almost verbatim duplication between A9 and M4–5, verbatim duplication between the latter part of K and parts of the *Physics*, looser duplication between the former part of K and BΓE, and the coexistence of two books called alpha [now distinguished as A and α]) that Aristotle never finished the treatise to his satisfaction. Perhaps he would have discarded some parts of the *Metaphysics*, and perhaps some were never intended for the treatise; and there are grounds for suspecting that K is a student's reworking of Aristotle's lectures. In what follows, it will be assumed that all the books except α and K were intended to belong to the treatise, but many scholars

have doubted this for Δ and (less plausibly) Λ. (There is nothing to support the view, popular among nonspecialists, that the *Metaphysics* consists of fourteen independent books assembled by later Peripatetic editors; there are many forward and backward references between the books, including Δ and Λ.)

Aristotle says different things in different parts of the *Metaphysics* about the object of wisdom (or "first philosophy," as he says when distinguishing it from physics; once he calls it "theology"). Jaeger took these as evidence of different chronological strata: Aristotle would first have conceived the project of wisdom as searching for divine substances to replace the Platonic Forms, then reconceived it as a general study of being. But usually, and rightly, Aristotle's descriptions of wisdom are thought to be compatible. They are best taken as part of a developing strategy in the *Metaphysics* to narrow down and finally to acquire *sophia*.

It is perhaps most often thought that Aristotle aims at a universal science; that this project faces a difficulty, because "being" is said in different ways of things in different categories; that Aristotle proposes to solve this by discovering things that *are* in the primary way (these things, whatever they are, will be called substances, and once we understand their mode of being, we can understand the derivative modes of being of other things); that there are different and sometimes conflicting criteria for something to *be* in the primary way; that forms meet these different criteria better than matter or matter-form composites, but that the forms of corruptible things do so only imperfectly (because they are not separable except by reason); and that Aristotle therefore turns to divine forms (forms existing separately from matter), which will allow us to understand the derivative modes of being of other forms, other substances, and non-substances. This would explain why Aristotle can say that wisdom is about being, that it is about substance, and that it is about divine things.

However, the *Metaphysics* does not actually follow this program, and Aristotle nowhere calls divine things "forms," and nowhere says that they are beings or substances in any stronger sense than ordinary form-matter composites are (still less does he use them to understand the inferior modes of being of other things). Instead, Aristotle begins with an ethical characterization of wisdom, infers that wisdom will be a science of the *archai* (the "principles," or first of all things) and of first causes, then specifies these as causes of *being*, then reaches an account of divine things as *archai* and first causes of being, not as instances of a special sense of being. Theology is not a means to ontology; rather, ontology is a means to theology, or more precisely to "archeology" (knowledge of the *archai* might still count as wisdom even if there were nothing divine to know).

Metaphysics A begins by characterizing wisdom as the kind of knowledge intrinsically most worth having, setting aside practical consequences; Aristotle then argues that this is knowledge of the *archai*, and that these *archai* will be first causes of all things. Indeed, all philosophers who believe in theoretical wisdom claim knowledge of some *archai*; for pre-Socratic physicists, these are whatever existed from eternity before the ordered world arose out of them; for Platonic dialecticians, the Forms, especially maximally universal forms like being and unity; for Pythagorizing mathematicians, the one and the two or the infinite. We cannot directly observe any of these claimed *archai*, but must infer them as causes of more manifest things. Aristotle asks how each philosopher uses his *archai* as causes—that is, how the things he posits at the beginning of his account function in explaining the things he describes as arising later.

The best earlier philosophers, Anaxagoras and Empedocles and Plato, agree that among the *archai* is a Good and cause of goodness to the world. But, Aristotle claims, Anaxagoras and Empedocles cite only material and efficient causes (using *nous* or Love, their good *archai*, as efficient causes), and Plato cites only material and formal causes (using the one, his good *archê*, as a formal cause). Aristotle's main point is not that earlier philosophers have been discovering the four causes of his *Physics*, but that no one has yet used the Good as a *final* cause, thus no one has made it a cause whose *goodness* is explanatory. Aristotle thus motivates a new search for *archai* which will lead, in Λ, to a good *archê* as a final cause. He thus hopes to vindicate a key aspiration of Platonism, which Plato had undermined in his lecture on the Good by reinterpreting the Good as mathematical unity. Aristotle's rival Speusippus had concluded that the *archê* is One but not good; Aristotle makes the opposite decision, to discard the mathematics and save goodness. (Aristotle gives detailed objections against Academic accounts of Forms and numbers and their *archai* in *Metaphysics* MN.)

Metaphysics B raises a series of *aporiai*, some about how the science of *archai* should proceed, some about the *archai* themselves, some about what things exist "by themselves" or as substances. If some *X* (a genus or a number, or being or unity) is not a substance, but is merely an attribute of some other underlying nature, then *X* is posterior to that nature and cannot be among

the *archai* that wisdom seeks. Now while we know that the *archai* will be first causes, this does not tell us how to find them, since there are different kinds of causes, and different effects we might seek to explain.

Metaphysics Γ proposes to find the highest causes as causes of the most universal effects: "there is a science of being, inasmuch as it is being, and its *per se* attributes"—a science that knows the causes to all things of the facts that they *are*, that they are each *one*, are severally *many*, and so on. It is sufficient to study the causes of being to *substances*, since the being of accidents is dependent on that of substances. (Γ argues that this science will also give explanatory understanding of the principles of noncontradiction and excluded middle.)

Metaphysics Δ distinguishes different senses of "*archê*," "cause," "one," "being," and other terms necessary for the investigation. Δ7 argues that "being" is said in several ways: being in different senses will have different kinds of cause, and confusion will result if we look for causes of being without drawing the necessary distinctions.

Metaphysics EZHΘ investigate causes of being in these different senses. E1 sets out the program of looking for the *archai* as causes of being, and specifically for *archai* which will be eternally unmoved and exist separately (not as attributes of something else); physics fails to reach unmoved *archai*, and mathematics fails to reach separately existing *archai*, and a new discipline of first philosophy or "theology" is needed. This might be dialectic, if Plato were right that the formal causes of things were eternal and separate, but he is not; E1 argues that physics, not first philosophy, understands the formal causes of natural things. E2–3 investigate the causes of "being *per accidens*," and E4 the causes of "being as truth," both concluding that no science (and certainly not wisdom) deals with these causes; the serious possibilities are "being as said of the categories," primarily of substance and derivatively of accidents, treated in ZH, and "being as actuality and potentiality," treated in Θ. *Metaphysics* I ("Iota") deals with causes of *per se* attributes of being such as unity, difference, and contrariety, arguing that these do not lead to a separately existing one-itself or first pair of contraries, but only to a unit or a contrariety within each genus.

Metaphysics Z examines the causes of being as said of substances and accidents, but quickly restricts itself to the primary case, causes of substance. Aristotle speaks interchangeably of "the cause of substance to *X*" and "the substance of *X*." The conventional translation "substance" for *ousia* (the nominalization of the verb "to be") obscures the point that the *ousia* of *X* is whatever answers the question "what is *X*?".

There are several ways we might answer this question, notably by giving the *subject* of *X* (i.e., a *Y* such that *Y* is *X*: "what is Socrates?" "this flesh and these bones"), or by giving the *essence* of *X* (i.e., a *Y* such that *X is Y*, or such that for *X* to be is for it to be *Y*: "what is man?" "wingless biped animal"), or by giving some part of the essence of *X*, such as a universal or genus under which *X* falls. The *ousia* of *X* taken the first way is its material cause; the *ousia* of *X* taken the second way is its formal cause. A philosopher might hope to reach *archai*, eternal and prior to sensible things, by starting with some sensible substance and asking "what is it?" repeatedly, in one of these ways, until some ultimate answer is reached: this might be, as a material cause, atoms and the void, or earth, water, air, and fire, or the "receptacle" of the *Timaeus*; or, as a formal cause, Platonic Forms, especially the genera and being and unity.

Z devotes much ingenuity to showing that these projects do not succeed; what a sensible substance *is* is most properly its form, not a separate eternal form but one that does not exist prior to the form-matter composite. Plato might argue that, if the composite *X* came to be, there must already have been a form or essence of *X* for the process of coming-to-be to aim at; Aristotle agrees, but argues that this is not a separate eternal form, but a form existing in a generator of the same species (e.g., for an animal, the father) or in the soul of the artisan who produced *X*. Aristotle also argues that if the parts of the essence mentioned in the definition of *X* (like three lines in the definition of triangle, or like the four elements in Empedocles' definition of blood as "earth, water, air, and fire in equal proportions," or like animal and biped in the definition of man) were *archai* existing in actuality prior to *X*, *X* would not be one thing but many things (thus, as a *reductio ad absurdum* of Plato, there would not be one Form, Man, but two Forms, Animal and Biped). This argument might seem to make definition impossible, since the *definientia* are supposed to be prior to the *definiendum*; but Aristotle argues that they can be definitionally prior without being capable of separate existence. There is no Animal that is just animal, prior to the differentiae of animals: an actual animal is always a biped animal or a quadruped animal or the like, and the genus "animal" is merely a potentiality for these differentiae. Likewise, actual matter is always hot or cold, wet or dry, and the common matter that underlies all sensible changes is only a potentiality, not something actually existing prior to all sensible things. *Metaphysics* ZH thus

give rules for definition, with implications for understanding the relations between genus and differentia, universal and particular, form and matter; but the *archai* of these definitions are not the eternal, separately existing first things sought by wisdom (whether pre-Socratic physics or Platonic dialectic or Aristotelian first philosophy); prior in definition but not in separate existence, they are objects of Aristotelian physics.

Metaphysics Θ examines causes of being as actuality and as potentiality. A power or potentiality (*dunamis*), whether an active power to produce *X* or a passive power to undergo or become *X*, is a cause of *X*'s existing potentially (*dunamei*). Most of the *archai* of the physicists would be potentialities or potential causes. Thus the "seeds" in Anaxagoras's precosmic mixture *can become* plants and animals and their functional parts; Anaxagoras's *nous*, or the demiurge of the *Timaeus*, prior to the cosmos, *can act* to produce order, but are not yet doing so. But such causes explain only the potential existence of the cosmos, and give no sufficient reason why the active *archê* should begin to act on the passive *archê*. That the effect exists *actually* (*energeiâ[i]*) requires an *activity* (*energeia*) or an actual cause ("housebuilder" is a potential efficient cause, "housebuilder housebuilding" an actual efficient cause). Aristotle tries, both to extract general concepts of *dunamis* and *energeia*, and to argue that *energeia* is prior to *dunamis*: seeds are not prior to mature living things (since a seed exists dependently on a previous mature member of the species), and the *archai* in the strict sense, the first of all things, are not *dunameis* or potential causes, but *energeiai* or actual causes. Thus against (say) Anaxagoras's conception of the *archai* as temporally and narratively prior to the cosmos, the *archai* must from all eternity have been acting to produce the cosmos, so the cosmos too must have existed from eternity.

Metaphysics Λ pulls together the threads of ZHΘ and draws conclusions for what causal chains lead up from changing sensible things to separate eternal *archai*. There is no single separately existing matter of all changeable things, nor a single form even for all things in the same species. While the form of a natural composite substance does not exist before the composite, its generator, a previous mature member of the same species, does exist before; but this chain of efficient causes goes back *ad infinitum*, without leading to a separate eternal *archê*. But Aristotle argues (drawing on *Physics* VIII) that the eternal continuance and approximate periodicity of sublunar generation require a further cause: not simply the sublunar generators, but something eternal and perfectly regular—namely, the rotations of the heavens—that sets the precise time lengths that sublunar cycles aim to approximate. Especially the daily and yearly motions of the sun, yielding the cycle of the seasons, serve to regulate cycles of generation.

Furthermore, these eternally unchanging motions require eternally unchanging substances as their efficient causes. Aristotle accepts Anaxagoras's and Plato's description of the mover of at least the first motion, the daily rotation of the whole heaven, as *nous*. But, using the premise that the *archê* must be pure *energeia*, he critically examines earlier philosophers' descriptions of *nous*'s causality, rejecting anything that would imply *dunamis* or changeability. Notably, *nous* must always move the heavens in the same way, and it must not move them in such a way as to be reciprocally affected by them. "Purifying" the Anaxagorean and Platonic accounts in this way, Aristotle concludes that *nous* moves the heaven only by causing the heaven to know and desire it: *Nous* is an efficient cause only by being a final cause. (When the heaven desires its mover, what does it desire to do, and how does this explain its motion? It should, like humans, order its actions toward contemplating God; and presumably its eternally unchanging motion is the best available imitation of God's eternally unchanging *energeia*.)

The premise that this *nous* is pure *energeia* also allows Aristotle (drawing on *De anima* III on "active *nous*") to "purify" earlier accounts of how it thinks and what it thinks. It is not a cognitive ability that could be applied to many objects, but a single eternal act of cognition of a single eternal object—the best object, or "good-itself." If this object were outside *nous*, *nous* would depend on something external to complete its act of cognition, and would of its own essence be merely potential *nous*; Aristotle concludes that *nous* is identical with the good-itself that it contemplates. This result allows Aristotle to fulfill various promises about wisdom from *Metaphysics* A, showing how the good is a cause, *qua* good (as a final cause), and not just as an efficient or formal cause. He vindicates Plato's promise of a single first good *archê* against Speusippus's criticism, but only by giving up on talk of the One, and finding a causal route up to the *archê* from physics rather than from mathematics.

INFLUENCE

Aristotle's immediate influence came through the Peripatetic school, led after Aristotle's death by his student Theophrastus; other important students were Eudemus, Aristoxenus, and Dicaearchus. The Peripatetic Demetrius of Phalerum governed Athens, backed by Macedonian power, from 317 to 307 BCE. But the Peripatetic school

declined after this (perhaps in part because of the reaction against Demetrius at Athens); for most of the Hellenistic period (323–30 BCE), the dominant schools were the Academics, Stoics, and Epicureans. Peripatetics turn up more at Alexandria than at Athens, and more in biography and literary scholarship than in scientific philosophy.

However, there was a revival of Aristotle, as well as of Plato, in the first centuries BCE–CE, and attention turned from the "exoteric" texts to the "acroamatic" texts as offering a systematic teaching in all philosophical disciplines. Teaching would take place, by oral exposition of the texts of Aristotle, in whatever was thought to be the correct sequence, accompanied by refutations of more recent schools and solutions to new *aporiai*. This oral teaching is reflected in written commentaries, of which the most important are those of Alexander of Aphrodisias (c. 200 CE); we also have paraphrases of several Aristotelian treatises by Themistius (fourth century CE).

Besides the Peripatetics, late ancient Platonists make use of Aristotelian concepts in trying to extract a systematically teachable technical philosophy out of Plato's dialogues, and often wind up incorporating Aristotelian doctrines. In particular, they share Aristotle's concern to avoid inappropriately assimilating soul or *nous* or other divine realities to lower things, notably by attributing to them extension or change or *dunamis*. (Aristotle is here seen as an ally especially against Stoic corporealism.) Where Plato describes intelligible forms as conspecific with sensible things, the demiurge as acting after a period of inactivity, or thinking as a circular motion of soul, the Platonists use Aristotle's arguments, together with a principle of charity, to argue that Plato must have intended these comparisons to sensible things to be understood allegorically; they say either that Aristotle's criticisms of Plato are misunderstandings, or that Aristotle intended to criticize not Plato, but only disciples who took Plato's metaphors literally. At the same time, they reinterpret Platonic forms as a plurality of sciences in God, weakening Aristotle's insistence on the singleness of God's knowledge. Thus fifth- and sixth-century commentaries both on Plato and on Aristotle harmonize the two authors to some extent. This is taken furthest by Simplicius; by contrast, John Philoponus, for Christian reasons, defends some specifically Platonic doctrines, including creation in time, against Aristotle.

After the mid-sixth century, the teaching of philosophy collapses, beyond introductions to philosophy and lectures on Porphyry's *Introduction* (*Isagoge*) to the *Organon* and Aristotle's *Categories*, *De interpretatione*, and the elementary part of the *Prior Analytics*. The recovery of the rest of Aristotle's work as a basis for systematic philosophical instruction occurred first in the Islamic world; key figures are al-Fārābī and Avicenna (Ibn Sīnā). These thinkers accept the guidance of late ancient commentators, and thus share to some extent in the harmonizing of Aristotle and Plato; by contrast, Averroes (Ibn Rushd) champions Alexander of Aphrodisias against harmonistic commentators, and tries to defend "scientific" Aristotelian philosophy against what he sees as unscientific Platonist contamination. Versions of Avicennian philosophy are taught in Iran to the present day.

In Greece, Michael Psellus revived late ancient Platonic-Aristotelian philosophy, which remained vital until the fall of the Byzantine Empire. In the Latin West, knowledge of Aristotelian philosophy survived in the translations and commentaries on the *Organon* by Boethius; Abelard and his twelfth century contemporaries began a renaissance of Aristotelian philosophy based almost wholly on the logic. Around 1200, translations of the whole Aristotelian corpus became available (first from Arabic, along with Arabic commentaries and treatises, then directly from Greek), and a systematic teaching of Aristotelian philosophy ("scholasticism") became the basis for university instruction, and a prerequisite for the study of Christian theology; different solutions were proposed to the conflicts between Aristotle and biblical revelation (key figures are Thomas Aquinas, John Duns Scotus, and William of Ockham).

While scholastic Aristotelianism flourished in the Renaissance (key figures are Pietro Pomponazzi in Italy, Francisco Suárez in Spain and Portugal), there is also much Renaissance polemic against Aristotle. The charge may be that he is irreligious (he makes the causal connection of God with the world too thin, and seems to deny providence over the sublunar world and the immortality of human souls; he certainly denies miracles such as creation in time or resurrection); that his claims of scientific knowledge cannot overcome skeptical challenges; or that his explanations are tautologous, multiplying words without practical consequences either technical or moral. These criticisms are taken up by the mechanical philosophers of the seventeenth century (notably Descartes, Gassendi, Hobbes), who aim to give a systematic replacement for Aristotelian physics, doing without forms or qualities superadded to matter (except possibly the human rational soul), abolishing the distinction between heavenly and earthly matter, and deriving phenomena from a natural tendency of bodies to persist in rectilinear motion, and from the results of collisions between bod-

ies. Since the successes of this new physics (culminating in Newton), no systematic revival of Aristotelian philosophy has been possible; likewise, modern mathematical logic has permanently eclipsed Aristotelian syllogistic. Kantians often accuse Aristotle of uncritical realism in epistemology.

But Aristotle continues to be central in philosophical education, and to be a source of inspiration, chiefly in practical philosophy, metaphysics, and the philosophy of mind. Often philosophers have turned back to Aristotle for a description of phenomena of ordinary experience and language, careful attention to which (it is claimed) would undermine the appeal of oversimple modern reductionist theories (utilitarianism, associationism, materialism), without positing anything radically beyond ordinary experience (categorical imperatives of pure reason, intellectual intuitions, incorporeal substances). Neo-Aristotelians prefer intensional to extensional distinctions: a soul is not a substance other than the living body, but is the body itself *qua* living and not merely *qua* body. And these intensional differences are discerned, not by intellectual intuition, or by Kantian *a priori* synthesis, but by ordinary perception disciplined to recognize things *as* what they are. (Thus the practical rationality required for virtue is neither means-end reasoning nor a Kantian faculty of rules, but a sensitivity to morally salient features of situations.) Aristotle is seen as seeking a "middle way," for example, between pre-Socratic materialism and Platonic metaphysics, that could be a model for modern philosophers. Such interpretations tend to understate the commitments that Aristotle shares with Plato, and his internal criticisms and refinements of the Platonic philosophical (and theological) project; the use of Aristotle for inspiration in contemporary philosophy should be balanced by an awareness of the risks of removing Aristotle from his context and reducing him to what seems usable for current philosophical problems.

See also Anaxagoras of Clazomenae; Ancient Aesthetics; Aristotelianism; Logic, History of: Ancient Logic; Neoplatonism; Plato; Theophrastus.

Bibliography

There is an excellent annotated bibliography in Jonathan Barnes, ed., *The Cambridge Companion to Aristotle* (Cambridge, U.K.: Cambridge University Press, 1995); there are also comprehensive bibliographies in Hellmut Flashar, ed., *Grundriss der Geschichte der Philosophie* ("Ueberweg"), *Die Philosophie der Antike*, Vol. 3, pp. 407–492 (Basel: Schwabe, 2004), and, for the *Metaphysics*, in Robero Radice and Richard Davies, *Aristotle's Metaphysics: Annotated Bibliography of the Twentieth-Century Literature* (Leiden: Brill, 1997). Only primary sources, indispensable scholarly tools, and starting points for entering into scholarly discussions, including some particularly important publications that have appeared since Barnes, will be listed here. All the transmitted works and many fragments are translated in Jonathan Barnes, ed., *The Complete Works of Aristotle*, 2 vols. (Princeton, NJ: Princeton University Press, 1984); the translations are not uniform, and some are better than others. Many works have been well translated, with useful notes, in the Clarendon Aristotle series from Oxford University Press (thus far: *Categories and De interpretatione*, translated by J. L. Ackrill, 1962; *Posterior Analytics*, translated by Jonathan Barnes, 2nd ed., 1994; *Topics* I and VIII, translated by Robin Smith, 1994; *Physics* I–II, translated by William Charlton, 1980; *Physics* III–IV, translated by Edward Hussey, 1983; *Physics* VIII, translated by Daniel Graham, 1999; *De generatione et corruptione*, translated by C. J. F. Williams, 1982; *De anima* II–III, translated by D. W. Hamlyn, 2nd ed., 1993; *De partibus animalium* I and *De generatione animalium* I, translated by David Balme, 2nd ed., 1993; *De partibus animalium*, translated by James Lennox, 2001; *Metaphysics* Β, translated by Arthur Madigan, 1999; *Metaphysics* ΓΔΕ, translated by Christopher Kirwan, 2nd ed., 1993; *Metaphysics* ΖΗ, translated by David Bostock, 1994; *Metaphysics* ΜΝ, translated by Julia Annas, 1977; *Eudemian Ethics* I, II, and VIII, translated by Michael Woods, 2nd ed., 1992; *Nicomachean Ethics* VIII–IX, translated by Michael Pakaluk, 1998; *Politics* I–II, translated by Trevor Saunders, 1996; *Politics* III–IV, translated by Richard Robinson, 1962; *Politics* V–VI, translated by David Keyt, 1999; *Politics* VII–VIII, translated by Richard Kraut, 1997). The editions and translations in the Loeb Classical Library are usually less good. Hackett has published good annotated translations of the *Prior Analytics* (translated by Robin Smith, 1989), *Nicomachean Ethics* (translated by Terence Irwin, 1985), *Politics* (translated by C. D. C. Reeve, 1998), and *Poetics* (translated by Richard Janko, 1987), all Indianapolis, IN: Hackett. Also useful is George Kennedy's annotated translation of the *Rhetoric*, *On Rhetoric: A Theory of Civic Discourse*, New York: Oxford University Press, 1991. Many works are also available in the Budé series, from Editions Les Belles Lettres, Paris: Presses Universitaires de France, Greek text with facing French translation, and in the Oxford Classical Texts (Greek only, many of them edited by W. D. Ross) and the Teubner series (Greek text only; previously Leipzig: Teubner or Stuttgart: Teubner, now Munich: Saur). W. D. Ross's *editiones maiores* (Greek text with English introduction and commentary) of the *Metaphysics* (2 vols., 1924), *Physics* (1946), *Analytics* (1955), *De anima* (1961) and *Parva naturalia* (1955), and Harold Joachim's *editio maior* of the *On Generation and Corruption* (1922), all Oxford: Oxford University Press, are extremely useful. However, on the *Metaphysics* and *De anima*, it is advisable to compare Werner Jaeger's Oxford Classical Text (1957) and Antonio Jannone and Edmond Barbotin's Budé (1966), respectively; with Joachim compare Marwan Rashed's Budé (2005). The Budé texts are generally more prudent and reliable than Ross's; some have very useful introduction and notes, others are only sparsely annotated. For the *Categories*, compare the Oxford Classical Text of Lorenzo Minio-Paluello (also including the *De*

interpretatione, 1949) with Richard Bodéüs' Budé (2001), which puts "[*Catégories*]" in brackets on the title page and then switches to *Les Avant les Lieux* ("The Before the Topics"). Jacques Brunschwig's Budé *Topiques* (Vol. 1, 1967; Vol. 2, forthcoming) particularly stands out. For the *Rhetoric*, the standard text is Rudolfus Kassel, *Aristotelis Ars rhetorica*, Berlin: De Gruyter, 1976. For the *De motu animalium*, see Martha Nussbaum, Aristotle's *De motu animalium*, Princeton, NJ: Princeton University Press, 1978; for the *Historia animalium*, see David Balme, *Aristotle: Historia animalium*, Cambridge, U.K.: Cambridge University Press (Vol. 1, 2002; Vol. 2, forthcoming). For the fragmentary *Protrepticus*, compare Ingemar Düring, *Aristotle's Protrepticus: An Attempt at Reconstruction*, Göteborg: Acta Universitatis Gothoburgensis, 1961 (including an edition of the fragments with English translation) with the forthcoming translation by Douglas Hutchinson and Monte Johnson (expected Indianapolis: Hackett). For the fragmentary *On Ideas*, see Gail Fine, *On Ideas*, Oxford: Oxford University Press, 1993, and Walter Leszl, *Il "De ideis" di Aristotele e la teoria platonica delle idee*, Florence: Olschki, 1975 (including a critical edition by Dieter Harlfinger). An extremely useful tool for studying concepts across the Aristotelian corpus (unfortunately presupposing Latin as well as Greek, and using a reference system that takes some getting used to) is Hermann Bonitz, *Index Aristotelicus*, Berlin: Reimer, 1870.

The most important secondary literature on Aristotle remains the late ancient commentaries collected in *Commentaria in Aristotelem Graeca*, 23 vols., Berlin: Reimer, 1882–1909, especially those of Alexander of Aphrodisias and Simplicius. Many of these volumes have been translated into English in the series *Ancient Commentators on Aristotle*, edited by Richard Sorabji, London: Duckworth, and Ithaca, NY: Cornell University Press. On the Greek commentators see Richard Sorabji, ed., *Aristotle Transformed*, London: Duckworth, 1990, and Richard Sorabji, ed., *Philoponus and the Rejection of Aristotelian Science*, London: Duckworth, 1987. For the earlier Greek reception of Aristotle, see Paul Moraux, *Der Aristotelismus bei den Griechen*, 2 vols., Berlin: De Gruyter, 1973–1984. For the availability of medieval Arabic and Latin commentaries on Aristotle, see the articles on, especially, al-Fārābī, Averroes (Ibn Rushd), and Thomas Aquinas.

Articles on Aristotle, edited by Jonathan Barnes, Malcolm Schofield, and Richard Sorabji, 4 vols., London: Duckworth, 1975–1979, is a good cross section of modern scholarship up to its time, with non-English articles translated. There are also excellent collections of articles, usually on a particular work of Aristotle, often arranged so as to form a collective commentary, in the proceedings of the triennial Symposia Aristotelica, including thus far: Ingemar Düring and G. E. L. Owen, eds., *Aristotle and Plato in Mid-Fourth Century*, Göteborg: Acta Universitatis Gothoburgensis, 1960; Suzanne Mansion, ed., *Aristote et les problèmes de méthode*, Louvain: Publications universitaires, 1961; G. E. L. Owen, ed., *Aristotle on Dialectic: The Topics*, Oxford: Oxford University Press, 1968; Ingemar Düring, ed., *Naturphilosophie bei Aristoteles und Theophrast*, Heidelberg: Stiehm, 1969; Paul Moraux, ed., *Untersuchungen zur Eudemischen Ethik*, Berlin: De Gruyter, 1971; Pierre Aubenque, ed., *Etudes sur la Métaphysique d'Aristote*, Paris: Vrin, 1979; G. E. R. Lloyd and G. E. L. Owen, eds., *Aristotle on Mind and the Senses*, Cambridge, U.K.: Cambridge University Press, 1978; Enrico Berti, ed., *Aristotle on Science: the Posterior Analytics*, Padova: Editrice Antenore, 1981; Paul Moraux and Jürgen Wiesner, eds., *Zweifelhaftes im Corpus Aristotelicum*, Berlin: De Gruyter, 1983; Andreas Graeser, ed., *Mathematics and Metaphysics in Aristotle* (= *Metaphysics* MN), Bern: Haupt, 1987; Günther Patzig, ed., *Aristoteles' Politik*, Göttingen, 1990; David Furley and Alexander Nehamas, eds., *Philosophical Aspects of Aristotle's Rhetoric*, Princeton, NJ: Princeton University Press, 1994; Michael Frede and David Charles, eds., *Aristotle's Metaphysics Lambda*, Oxford: Oxford University Press, 2000; and Frans de Haas and Jaap Mansfeld, eds., *Aristotle's On Generation and Corruption I*, Oxford: Oxford University Press, 2004.

The classic attempt to write Aristotle's intellectual biography is Werner Jaeger, *Aristotle: Fundamentals of the History of his Development*, translated by Richard Robinson, Oxford: Oxford University Press, second edition, 1948. For a survey of the *status quaestionis* on Aristotle's development, see William Wians, ed., *Aristotle's Philosophical Development*, Lanham, MD: Rowman and Littlefield, 1996. The ancient sources on Aristotle's life are collected by Ingemar Düring, *Aristotle in the Ancient Biographical Tradition*, Göteborg: Acta Universitatis Gothoburgensis, 1957. On the titles and histories of Aristotle's writings, see Paul Moraux, *Les listes anciennes des ouvrages d'Aristote*, Louvain: Editions universitaires de Louvain, 1951; on stories about the rediscovery of the texts and on the possible role of Andronicus of Rhodes in editing and ordering Aristotle's texts, compare the treatment of Paul Moraux, *Der Aristotelismus bei den Griechen* (cited above) with that of Jonathan Barnes, "Roman Aristotle," in *Philosophia Togata* II, edited by Jonathan Barnes and Miriam Griffin, pp. 1–69, Oxford: Oxford University Press, 1997; and see Stephen Menn, "The Editors of the *Metaphysics*," in *Phronesis* 40, no. 2 (July 1995): 202–208. On Aristotle's treatment of his predecessors, the classic (but highly controversial) works are Harold Cherniss, *Aristotle's Criticism of Presocratic Philosophy*, Baltimore, MD: Johns Hopkins Press, 1935, and *Aristotle's Criticism of Plato and the Academy*, Vol. 1 (Vol. 2 was never published), Baltimore, MD: Johns Hopkins Press, 1944.

On the *Nicomachean Ethics*, starting points are R. A. Gauthier and J. Y. Jolif, *L'Ethique à Nicomaque*, introduction, traduction et commentaire, 2 vols. in 3, Louvain: Publications universitaires de Louvain, 1958–1959; W. F. R. Hardie, *Aristotle's Ethical Theory*, 2nd ed., Oxford: Oxford University Press, 1980; Sarah Broadie, *Ethics with Aristotle*, Oxford: Oxford University Press, 1991; and Amélie Rorty, ed., *Essays on Aristotle's Ethics*, Berkeley: University of California Press, 1980. On the relation between the different ethical works, starting points are (besides Michael Woods's Clarendon *Eudemian Ethics* I, II, and VIII) Anthony Kenny, *The Aristotelian Ethics*, Oxford: Oxford University Press, 1978 (although Kenny's claim that the *Nicomachean Ethics* is earlier than the *Eudemian* is idiosyncratic); and John Cooper, "The *Magna Moralia* and Aristotle's Moral Philosophy," in his *Reason and Emotion*, pp. 195–211, Princeton, NJ: Princeton University Press, 1999. On the *Politics*, besides the introductory volume and the essays and notes of the classic W. L. Newman, *The Politics of Aristotle*, 4

vols., Oxford: Oxford University Press, 1887–1902, see now Richard Kraut, *Aristotle: Political Philosophy*, Oxford: Oxford University Press, 2002; and, for the context of political debate, Josiah Ober, *Political Dissent in Democratic Athens*, Princeton, NJ: Princeton University Press, 1998. On the *Constitution of the Athenians*, see P. J. Rhodes, *A Commentary on Aristotle's Athênaiôn Politeia*, Oxford: Oxford University Press, 1981; and John Keaney, *The Composition of Aristotle's Athênaiôn Politeia*, New York: Oxford University Press, 1992.

On dialectic, besides Robin Smith's Clarendon *Topics* I and VIII, the introduction to Jacques Brunschwig's Budé *Topiques*, and G. E. L. Owen's Symposium Aristotelicum volume, all cited above, see the introduction to Eleonore Stump, *Boethius De differentiis topicis*, Ithaca, NY: Cornell University Press, 1978; and Robin Smith, "Aristotle on the Uses of Dialectic," in *Synthese* 96 (1993): 335–358. The majority of English-language literature on dialectic is really about the use of pre-reflective intuitions or ordinary language in argument, and has little to do with what Aristotle calls dialectic. On the *Categories*, see Michael Frede, "Title, Unity and Authenticity of the Aristotelian *Categories*," in his *Essays on Ancient Philosophy*, pp. 11–28, Oxford: Oxford University Press, 1987; and Stephen Menn, "Metaphysics, Dialectic, and the Categories," in *Revue de Métaphysique et de Morale*100, no. 3 (July–September 1995): 311–337; on the *On Interpretation*, see C. W. A. Whittaker, *Aristotle's De interpretatione: Contradiction and Dialectic*, Oxford: Oxford University Press, 1996; on the *Sophistical Refutations*, see Louis-André Dorion, Aristote, *Les refutations sophistiques, introduction, traduction et commentaire*, Paris: Vrin, 1995. On Aristotle's syllogistic, besides Robin Smith's introduction to his Hackett *Prior Analytics*, see Jan Łukasiewicz, *Aristotle's Syllogistic*, 2nd ed., Oxford: Oxford University Press, 1957; Günther Patzig, *Aristotle's Theory of the Syllogism*, Dordrecht: Reidel, 1968; and Paul Thom, *The Syllogism*, Munich: Philosophia, 1981; on demonstration, see Enrico Berti's Symposium Aristotelicum volume (*Aristotle on Science*) and Jonathan Barnes' Clarendon *Posterior Analytics*. On the *Rhetoric*, the Symposium Aristotelicum volume (D. Furley and A. Nehamas, *Aristotle's Rhetoric*) gives one approach, and George Kennedy's translation (mentioned above), and his *The Art of Persuasion in Greece*, Princeton, NJ: Princeton University Press, 1963, give another. On the *Poetics*, good starting points are Richard Janko's Hackett and Amélie Rorty, ed., *Essays on Aristotle's Poetics*, Princeton, NJ: Princeton University Press, 1982.

On physics and cosmology, there is unfortunately not much to recommend, beyond the *editiones maiores* and Symposium Aristotelicum volumes noted above; still useful is Friedrich Solmsen, *Aristotle's System of the Physical World*, Ithaca, NY: Cornell University Press, 1960; on Aristotle on causality see Richard Sorabji, *Necessity, Cause and Blame*, London: Duckworth, 1980; and for the issues of Greek cosmology, see David Furley, *The Greek Cosmologists*, Vol. 1, Cambridge, U.K.: Cambridge University Press, 1987 (the promised Vol. 2 has not yet appeared), and his collection of articles *Cosmic Problems*, Cambridge, U.K.: Cambridge University Press, 1989. By contrast, there is a large literature on Aristotle's psychology and zoology. For orientation to the literature on the psychology, see Martha Nussbaum and Amélie Rorty, eds., *Essays on Aristotle's De anima*, 2nd ed., Oxford: Oxford University Press, 1995; and see Stephen Menn, "Aristotle's Definition of Soul and the Programme of the *De anima*," in *Oxford Studies in Ancient Philosophy*, 22 (Summer 2002): 83–139. On Aristotle's ordering and reordering of the psychological-physiological-zoological corpus, see Marwan Rashed, "Agrégat de parties ou *vinculum substantiale*? Sur une hésitation conceptuelle et textuelle du *corpus* aristotélicien," in André Laks and Marwan Rashed, eds., *Aristote et le mouvement des animaux: dix études sur le De motu animalium*, pp. 185–202, Lille: Presses universitaires du Septentrion, 2004. On zoology, besides the translations and commentaries of Lennox, Balme, and Nussbaum, cited above, the best starting point is Alan Gotthelf and James Lennox, eds., *Philosophical Perspectives on Aristotle's Biology*, Cambridge, U.K.: Cambridge University Press, 1985.

For the *Metaphysics*, it is always helpful to compare Theophrastus's *Metaphysics*, for which see the Budé by André Laks and Glenn Most, 1993; for an English translation, see W. D. Ross and F. H. Fobes, eds., Theophrastus, *Metaphysics*, Oxford: Oxford University Press, 1929. Modern landmark studies on the *Metaphysics* are Werner Jaeger, *Studien zur Entstehungsgeschichte der aristotelischen Metaphysik*, Berlin: Weidmann, 1912; Pierre Aubenque, *Le problème de l'être chez Aristote*, Paris: Presses universitaires de France, 1966; Joseph Owens, *The Doctrine of Being in the Aristotelian Metaphysics*, Toronto: Pontifical Institute of Medieval Studies, 3rd ed., 1978; Günther Patzig, "Theology and Ontology in Aristotle's *Metaphysics*" in *Articles on Aristotle* (cited above), Vol. 3, pp. 33–49; Vianney Décarie, *L'objet de la métaphysique selon Aristote*, Paris: Vrin, 1961; Giovanni Reale, *The Concept of First Philosophy and the Unity of the Metaphysics of Aristotle*, translated by John Catan, Albany: SUNY Press, 1980; Michael Frede, "The Unity of General and Special Metaphysics: Aristotle's Conception of Metaphysics," in his *Essays on Ancient Philosophy* (cited above), pp. 81–95; Michael Frede and Günther Patzig, *Aristoteles: Metaphysik Z: Text, Übersetzung, Kommentar*, Munich: Beck, 1988; and Myles Burnyeat, *A Map of Metaphysics Zeta*, Pittsburgh: Mathesis Publications, 2001.

Stephen Menn (2005)

ARITHMETIC

See *Mathematics, Foundations of; Number*

ARIUS AND ARIANISM

Arius was a controversial fourth-century Christian thinker in Alexandria, Egypt, who was condemned by the first ecumenical council at Nicaea in 325. Because most of his writings were destroyed as heretical and "Arianism" as a movement developed only after his death, historians continue to debate both the content and the purpose of his teaching. Theological debate continued for a century

within Christianity, prompting a number of councils and creeds as well as a voluminous literature exploring the definition of God as Trinity, the origin of the divine Son, and the nature of salvation. From these events "Arianism" has been traditionally defined in theological polemic as a denial of the essential divinity of the Son and therefore of both the orthodox doctrines of Incarnation and Trinity.

ARIUS AND THE COUNCIL OF NICAEA

As a Christian presbyter in Alexandria, Arius claimed a connection to a famous martyr and theologian, Lucian of Antioch. Philosophically educated as well as an exegete of scripture, he criticized Bishop Alexander of Alexandria for using language of eternity and nature with regard to the generation of the Son from the Father; this defended a common shared divinity, but muddled their separate personal identities. Arius argued that the Father, defined as the creator of all existence, could not share his uncaused nature or being with the Son. To speak of a shared divine nature would compromise biblical monotheism as well as contradict the definition of the creator as unbegun by nature, therefore opposing the first principle of all existence. The Son had to be of a separate nature because he was created or generated at some point as the offspring of the Father.

Contrary to earlier theologians, Arius argued that the Son could not be eternally begotten or he would be a coexistent principle. Instead, the Son possessed divinity from his direct creation by the Father and preexistence before all creation, but this was a separate and secondary divinity. Early authors had also interpreted the title of Word from the Gospel of John or Wisdom from Proverbs to show the Son's eternal presence with the Father as a mental attribute. By contrast, Arius accepted the traditional titles, but denied the eternal presence; he also denied that the Son knew the Father apart from what knowledge the Father had bestowed upon him. The secondarily divine Son remained the revealer of the Father, the agent of creation, and the mediator of the divine will and salvation through Incarnation.

The origins and motivations of Arius's views remain controversial, and no single interpretation has yet to persuade all scholars. Only three of his documents remain, and his opponent, Athanasius, preserved fragments of his theological poem, *Thalia*. All historians emphasize his indebtedness to earlier theologians, such as Origen, who described the Son as Word, and ascribed a lesser and derivative nature to the Son. This hierarchical model echoed both the philosophy of Numenius and Philo, in which the Logos was the mediator between transcendent reality and the material world, as well as biblical accounts of the Son's obedience to the Father. However, by contrast, Arius denied any communication or participation of essence between the Father and the Son; apophatic theology became central to his thought. This highly significant shift may well reflect changes in contemporary Platonism, such as the increased transcendence of Plotinus's thought, but the parallels are not entirely conclusive.

Arius may also be defending the theology of Lucian of Antioch, which emphasized the will of God and the agency of the Son. The emphasis on the distinct nature of the Son may have been to portray him as a moral exemplar and mediator, in line with the New Testament model of the obedient Christ in Luke; the evidence for this interpretation, however, remains contested. Finally, Arius's rejection of coexistent principles could also be linked to the growing presence of Manichees in Egypt, who taught two eternal principles of good and evil. Clearly, Arius was a creative and powerful thinker who was revising traditional categories to clarify the singularity of the Father and the mediation of the Son.

After local councils did not succeed in reconciling or suppressing the controversy, Emperor Constantine convened a council of bishops from the East and West at Nicaea in 325. The accounts of the council show the difficulty of using scriptural language that, insofar as it was metaphorical, did not solve analytical difficulties concerning causality or nature. *Homoousios*, or "of the same nature," was adopted as a definition of the relation of the Father and Son, less as a positive definition than as a term rejected by Arius and others. However, the creed was not readily adopted by the larger church, and other councils were held over the next five decades to find more acceptable language. Constantine accepted a later statement by Arius that avoided discussing the nature of the Father and Son, if affirming the priority of the Father. Arius died in 336 before being accepted back into communion with the church, perhaps by poisoning.

"ARIANISM" AFTER ARIUS

The issues of divine causality and saving knowledge raised by the Arian controversy and the Nicene definition were strenuously debated for several decades. We may best speak of these shifting alliances as "non-Nicenes" rather than use the older categories of "Arians" or "Semi-Arians" to describe all those who for various reasons rejected the authority of the creed of Nicaea. Many were content to avoid substance language or affirm a "like substance" (*homoiousios*) between the Father and the Son,

maintaining a traditional hierarchy of being and action. Aetius and Eunomius, often called "Neo-Arians," were the most strenuous opponents of Nicene theology; they argued that the Father and Son must be dissimilar in nature since the divine essence was "unbegun," and insisted this description was fully revelatory of God. These varied opponents of Nicaea thinkers did not describe themselves as followers of Arius; rather, they were tagged with his condemnation by opponents, such as Athanasius and Marcellus of Ancyra, in order to discredit their theological positions.

The separation of divine nature between the Father and the Son also had implications for salvation and incarnation. The author of the Latin *Opus imperfectum* insisted that the created Son was able to suffer authentically on the cross; he criticized the Christology of the orthodox of Docetism, since they claimed only the body suffered and not the eternal Word. A series of legislative acts curtailed the activities of the non-Nicenes after the council of Constantinople in 381.

LATER ARIANISM

The Christianization of the Goths occurred during this theological turmoil, and they were baptized as "Arians." The destruction of the Visigothic library in medieval Spain erased documents that might have provided significant clues to Gothic theology. In the seventeenth century "Arianism" was embraced by a number of English theologians, including Isaac Newton and William Whiston, who questioned the logic of the Trinity and the biblical authority of creeds.

See also Christianity; Newton, Isaac; Platonism and the Platonic Tradition; Plotinus.

Bibliography

Barnes, Michel R., and Daniel H. Williams. *Arianism after Arius: Essays on the Development of the Fourth Century Trinitarian Conflicts*. Edinburgh: T. & T. Clark, 1993.

Gregg, Robert, and Dennis Groh. *Early Arianism: A View of Salvation*. New York: Fortress Press, 1981.

Hanson, Richard P. C. *The Search for the Christian Doctrine of God*. Edinburgh: T. & T. Clark, 1988.

Kopecek, Thomas. *A History of Neo-Arianism*. 2 vols. Cambridge, MA: Philadelphia Patristic Foundation, 1979.

Vaggione, Richard P. "'Arius, Heresy and Tradition' by Rowan Williams: A Review Article." *Toronto Journal of Theology* 5 (1989): 63–87.

Vaggione, Richard P. *Eunomius of Cyzicus and the Nicene Revolution*. Oxford: Oxford University Press, 2000.

Wiles, Maurice. *Archetypal Heresy: Arianism through the Centuries*. Oxford: Oxford University Press, 1996.

Williams, Daniel. *Ambrose of Milan and the End of the Nicene-Arian Conflicts*. Oxford: Oxford University Press, 1995.

Williams, Rowan. *Arius: Heresy and Tradition*. 2nd ed. Grand Rapids. MI: Wm. Erdman's, 2001.

J. Rebecca Lyman (2005)

ARKOUN, MOHAMMED

(1928–)

Mohammed Arkoun was born in Kabylia, Algeria, and spent much of his career at the Sorbonne in Paris. His early work in philosophy was in the history of Islamic philosophy, and in particular the thought of the Persian philosopher Miskawayh. Like so many modern Arab philosophers, Arkoun is part both of the Islamic world and of the secular European world, and how to reconcile those two worlds has been a continuing issue for those philosophers. It has been a continuing issue of interest to them how to reconcile these two worlds. Arkoun, on the one hand, has in general been supportive of *laïcité*, the determined secularism of France that he argues preserves the freedom of all to follow their religions. On the other hand, he has roundly criticized the ways in which the Islamic and the non-Islamic worlds have cast each other in the role of the Other. He outlines in his work how a tradition creates a world of discourse, but at the same time also cuts people off from other forms of discourse. Thus traditions, and in particular religious traditions, can be seen to have both positive and negative features. Arkoun suggests that it is not acceptable for a tradition to rule out some ways of thinking, because in order to understand the whole range of alternatives that are available, people first need to contemplate a wide range of options.

But does this not contravene the idea that a tradition establishes rules about what can and cannot be thought? Here Arkoun broadens his analysis to suggest that traditions are not pure, and so do not have fixed boundaries. Traditions need to be applied to the world of experience; in turn, experiences will affect traditions on a piecemeal basis, and followers of a tradition will have to inevitably consider their responses to those experiences and the affect they had on the tradition. This brings out the problem with traditions that see the different approaches as representing the Other, because the distinctions between the tradition and the Other are often slight and difficult to determine. It follows that a program of secularism is not in opposition to religion, but should be seen as providing space for religions, and their opposites, to flourish

and think through their foundations. He also argues that Islam's renaissance in the nineteenth century is incomplete, that Muslims should radically examine the roots of their faith and establish it in line with contemporary forms of reality. If there is a theme in Arkoun's work it is the significance of history. History shows that doctrines such as Islam are never finished and complete, but continue to develop. History also shows that the antagonisms and conflicts between different ways of looking at the world are variable. An investigation of history allows people to ground their understanding of significant ideas within a particular context and thus acquire a critical understanding of them. There is a tension in this thesis—which owes much to the thought of Foucault—and the transcendent role that any religion seeks to establish for itself. Much of Arkoun's work tries to reconcile the clash between these two intellectual positions.

See also Enlightenment, Islamic; Foucault, Michel; Islamic Philosophy; Thinking; Traditionalism.

Bibliography

Arkoun, Muhammad. *Rethinking Islam: Common Questions, Uncommon Answers*. Boulder, CO: Westview Press, 1994.

Arkoun, Muhammad. *The Unthought in Contemporary Islamic Thought*. London: Saqi, 2002.

Oliver Leaman (2005)

ARMINIUS AND ARMINIANISM

Jacobus Arminius (Jacob Harmanszoon, 1560–1609), who gave his name to a variant of Reformed belief, was born in Oudewater, Holland. After his father's early death, the boy was protected in turn by a minister, who converted him to Protestantism; by Rudolphus Snel van Rooijen the mathematician; and by Pieter Bertius of Rotterdam. With Pieter Bertius Jr., later important in the great Arminian disputes, Arminius studied at Leiden under the French Protestant Lambertus Danaeus. Later Arminius studied under Theodorus Beza in Geneva, where he met Johannes Uytenbogaert (Wtenbogaert), the chief proponent of Arminian doctrines after the death of Arminius.

Soon after his ordination (1588), Arminius was called upon by the ecclesiastical court of Amsterdam to refute the arguments of the Dutch "libertine" theologian Dirck Volckertszoon Coornhert, an exercise that undermined Arminius' orthodox Calvinism. He came to doubt the deterministic doctrine of damnation, and believed that election, dependent in part on man's free will, was not arbitrary but arose from God's pity for fallen men. Arminius was consistently attacked by orthodox clergymen (notably Petrus Plancius and Franciscus Gomarus) for his alleged Pelagianism; in spite of all opposition, however, he was made professor of theology at Leiden in 1603 and thereafter exercised great influence upon the next generation of divines. He died just prior to the national schism brought about by his beliefs.

ARMINIANISM

In 1610 the Arminian clergy published their Great Remonstrance, a codification of Arminius' creed. This work dealt with five doctrinal points: It rejected the doctrine of election and predestination, both supralapsarian and sublapsarian. It rejected the idea that Christ died for the elect alone and belief in irresistible grace. It asserted belief in the sufficient power of saints, rejecting the idea that saints could fall from grace.

To the orthodox, these were Romish heresies; for eight years the battle of the pulpits raged, with Uytenbogaert, Bertius, and Hugo Grotius the great defenders of the Remonstrance. A theological question of this magnitude necessarily involved political theory and practice: the Remonstrants developed several versions of a theory by which, to protect consciences, the magistrate, rather than the Dutch Reformed Church, was given final say in matters of religion. Naturally, since such a theory favored republican administration, Arminianism gained support in the town governments and in the States-General, particularly in the figure of the pensionary of Holland, Jan van Olden Barneveldt.

In 1618 a synod was called to rule on Remonstrant doctrine, with the open support of the *stadholder,* Prince Maurice of Orange, who realized that the theological controversy might be used to curb the power of the States-General. For the hearing at Dordrecht (Dort), Arminian tenets were slightly modified by Uytenbogaert. Election was interpreted as God's grace to true believers; but this grace was not irresistible, and salvation still depended on the cooperation of the human will, which was sufficiently strong to overcome the temptations of evil. By the time the sessions began, the leading Arminian laymen had been arrested for treason: Olden Barneveldt was sentenced to be beheaded in The Hague; Grotius and Rombout Hogerbeets were imprisoned in Loevestein Castle.

The Synod was international: Representatives from Germany, Geneva, and England took part in the hearings, but the Remonstrants were barely allowed to be heard.

Their five tenets were declared inadmissible, or heretical, and orthodox Calvinism was upheld. Remonstrants were given the choice of recantation or exile.

Most chose exile—in France, Geneva, or England. Until the death of Prince Maurice in 1625, Arminianism was persecuted in Holland; but with the accession to the *stadholderate* of the tolerant Frederick Henry, Arminians began to return, particularly to the great cities of Amsterdam and Rotterdam. In 1630 a church was organized in Amsterdam, to which in 1632 an academy was attached, to train Remonstrant clergymen and the sons of Remonstrants barred from studying at the universities.

Dutch Arminianism was closely allied with advanced secular learning, both philosophical and scientific. The Remonstrant "Illustre School" (later the nucleus of the University of Amsterdam) was distinguished for its mathematical and medical, as well as its theological and philosophical, faculties. Whatever the philosophical implications of Arminius' humanistic doctrine, in the seventeenth century it was coupled with broad learning: An Arminian professor translated René Descartes's *Discourse upon Method* into Latin for the general use of the learned world; Arminian professors contributed to the periodicals of the republic of letters; and John Locke found a home among the Arminians during his exile from England.

See also Determinism and Freedom; Grotius, Hugo; Locke, John; Pelagius and Pelagianism.

Bibliography

PRIMARY SOURCES

Arminius, Jacobus. *Disputationes XXIV.* Leiden, 1609.

Arminius, Jacobus. *The Works.* Translated and edited by James Nichols, 3 vols. London, 1825–1875.

Brandt, Gerard. *The History of the Reformation and Other Ecclesiastical Transactions in the Low Countries,* 4 vols. London: Timothy Childe, 1720–1723.

Limborch, Philippus van, and Christiaan Hartsoeker. *Praestantium ac Eruditorum Virorum Epistolae Ecclesiasticae et Theologicae.* Amsterdam, 1684. Letters of Uytenbogaert, Arminius, Grotius, et al.

Triglandius, Jacobus. *Kerckelycke Geschiedenissen, Begrypende de Swaere Ende Bekommerlijke Geschillen, in de Vereenigde Nederlanden Voor-gevallen, met Derselver Beslissinge.* Leiden: Ghedruckt by A. Wyngaerden, 1650.

Uytenbogaert, Johannes van. *Kerckelycke Historie.* Rotterdam, 1647.

SECONDARY SOURCES

Colie, R. L. *Light and Enlightenment: A Study of the Cambridge Platonists and the Dutch Arminians.* Cambridge, U.K.: Cambridge University Press, 1957.

Davies, Godfrey. "Arminian versus Puritan in England, c. 1620–1650" *Huntington Library Bulletin* 5 (1934).

Harrison, A. W. *Arminianism.* London: Duckworth, 1937.

Harrison, A. W. *The Beginnings of Arminianism.* London: University of London Press, 1926.

Itterzoon, G. P. van. "Koning Jacobus I en de Synode van Dordrecht." *Nederlandsch Archief voor Kerkgeschiedenis* 24 (1932).

Nobbs, Douglas. *Theocracy and Toleration. A Study in the Disputes in Dutch Calvinism from 1600 to 1650.* Cambridge: Cambridge University Press, 1938.

Rogge, H. C. *Johannes Wtenbogaert en Zijn Tijd,* 2 vols. Amsterdam, 1874–1876.

Tideman, Johannes. *De Remonstrantie en het Remonstrantisme.* Amsterdam, 1851.

R. L. Colie (1967)

ARMSTRONG, DAVID M.
(1926–)

David Malet Armstrong was born in Melbourne, Australia. He received his first degree at the University of Sydney where John Anderson held the Challis chair of philosophy. He then completed the bachelor of philosophy at Oxford (in 1954), being one of the first of the many Australian philosophers in the 1950s and 1960s to take that degree. After a short spell at Birkbeck College, London, he accepted a position at the University of Melbourne. In 1964 he took up the Challis chair in Sydney where he stayed until his retirement in 1991.

Armstrong has made influential contributions to a remarkable range of major topics in epistemology and metaphysics, including perception, materialism, bodily sensations, belief and knowledge, laws, universals, and the metaphysics of possibility. Recurrent themes have been the need to reconcile what the philosopher says with the teachings of science, a preference for realist over instrumentalist theories, and an interest in the fundamental elements of being. A feature of his work is his ability to write about difficult issues with directness and clarity without sacrificing rigor.

Armstrong's *A Materialist Theory of the Mind* (1968) is a seminal and comprehensive presentation of the mind-brain identity theory, the view that mental states are states of the brain. Armstrong argues that for each mental state, there is a distinctive functional role. For each mental state, we can specify what it does by way of mediating between inputs, outputs, and other mental states. For example, pain is typically caused by bodily injury and typically causes behavior that tends to alleviate it; thirst is typically caused by lack of water and typi-

cally gives rise to behavior that leads to drinking water, provided there is water knowingly available to the subject. This means, Armstrong argues, that the question of the identity of a given mental state is nothing more than the question of the identity of that which plays the functional role distinctive of that state: Thirst is that which plays the role just described. It is then a question for science what state in fact plays that role, and that it will in fact be some state of the brain. Thus, Armstrong derives the mind-brain identity theory from a view about the distinctive roles played by mental states, combined with a view about what kinds of states—namely brain states—play those roles.

In the philosophy of perception he was one of the first to argue that we must move away from the tradition that thinks of perception as acquaintance with a special, mental item sometimes called a "sense datum." Instead, we should adopt an account that analyses perception as the acquisition of putative belief about our world—an account that has the signal advantage of making sense of the role of perception in our traffic with the world. Armstrong saw bodily sensations as being a special kind of perception—in the case of pain, a perception of putative damage in a part of one's body, accompanied by a desire that it cease. His work on sensations and perception may be seen as a precursor to currently much discussed representationalist accounts that analyze an experience in terms of how the experience represents things as being.

Armstrong revived interest in F. P. Ramsey's view that belief is like a map by which we steer, in opposition to approaches that think of belief as a kind of "saying to oneself." His account of knowledge is a version of reliabilism: S's true belief that P is knowledge if it is an empirically reliable sign that P.

Armstrong is a realist about universals: they exist, they are not reducible to sets of particulars (squareness is not the set of square things), and although they serve as the truth makers for predication, there is no simple one-to-one correspondence between predicates and universals. But there are no uninstantiated universals, so Armstrong is not a realist in Plato's sense. Armstrong deploys his realism about universals to deliver an account of laws of nature and of possibility. Laws are to be understood in terms of relations of nomic necessitation between universals: Roughly "Every F is G" is a fundamental law if being F necessitates being G. Armstrong's account of possibility is a combinatorial one. The various possibilities are the various combinations and recombinations of particulars (individuals) and universals that obey the right combinatorial rules (for example, combining being square with not being square does *not* deliver a possibility).

Armstrong's overall position in analytic ontology—that part of metaphysics that seeks to inventory at the most fundamental level what there is—is given in *A World of States of Affairs* (1997), where states of affairs are the basis on which accounts of properties, relations, numbers, necessity, dispositions, classes, causes, and laws are constructed.

See also Anderson, John; Being; Functionalism; Laws of Nature; Pain; Perception; Plato; Ramsey, Frank Plumpton; Realism; Reliabilism; Universals, A Historical Survey.

Bibliography

ARMSTRONG'S MAJOR WORKS

Berkeley's Theory of Vision. Melbourne: Melbourne University Press, 1960.

Perception and the Physical World. London: Routledge, 1961.

Bodily Sensations. London: Routledge, 1962.

A Materialist Theory of the Mind. London: Routledge, 1968.

Belief, Truth and Knowledge. Cambridge, U.K.: Cambridge University Press, 1973.

Universals and Scientific Realism. 2 vols. Cambridge, U.K.: Cambridge University Press, 1973.

What Is a Law of Nature? Cambridge, U.K.: Cambridge University Press, 1983.

With Norman Malcolm. *Consciousness and Causality: A Debate on the Nature of Mind.* Oxford: Basil Blackwell, 1984.

A Combinatorial Theory of Possibility. Cambridge, U.K.: Cambridge University Press, 1989.

Universals: An Opinionated Introduction. Boulder, CO: Westview Press, 1989.

A World of States of Affairs. Cambridge, U.K.: Cambridge University Press, 1997.

The Mind-Body Problem: An Opinionated Introduction. Boulder, CO: Westview Press, 1999.

Frank Jackson (1996, 2005)

ARNAULD, ANTOINE

(1612–1694)

Antoine Arnauld, a Jansenist theologian and Cartesian philosopher, was one of the most skilled philosophical and theological controversialists of the seventeenth century. His reputation was such that he was known in the early modern period as *le grand Arnauld*. Arnauld was born in Paris on February 8, 1612, the last of twenty children of Catherine Marion de Druy and the elder Antoine Arnauld. Arnauld's father served as an attorney for Queen

Catherine de Médicis, and at the beginning of the seventeenth century he successfully argued the case in the Parlement de Paris for the expulsion of the Jesuits from France. Arnauld's sister, Mère Angélique Arnauld, was installed as abbess of Port-Royal des Champs at the age of thirteen and became a prominent member of the convent. Though Arnauld initially intended to follow in his father's footsteps by becoming a lawyer, he later changed his mind and began to study theology in 1633. He received his baccalaureate in theology in 1635, and soon thereafter came under the influence of Jean Duvergier de Hauranne, the abbé de Saint-Cyran, who was then closely linked to Port-Royal. Because Saint-Cyran was also a political opponent of Cardinal Richelieu, Arnauld was prevented from receiving a doctorate from the Sorbonne during Richelieu's life. Soon after Richelieu's death, however, Arnauld received his doctorate in 1641 and became a member of the Sorbonne.

In 1641 Arnauld also wrote a critically sharp yet sympathetic set of objections to René Descartes's *Meditations*, an event that marks the start of his lifelong association with Cartesianism. In 1643 he published *De la fréquente communion*, an attack on the penitential theology of the Jesuits that earned him the enmity of members of that order. At the urging of Saint-Cyran, Arnauld also responded that same year to the criticisms of the theological account of grace and sin in the *Augustinus* (1640) by Cornelius Jansen, the bishop of Ypres, against the criticisms of Isaac Habert, a Paris theologian. In particular Arnauld insisted that Jansenius's views were in line with the criticisms in St. Augustine of the heretical Pelagian view that salvation depends on one's free will rather than on the workings of grace.

After 1648 Arnauld lived near Port-Royal as one of the solitaires associated with the convent. He was forced to go into hiding during this time because his opposition to the campaign against the *Augustinus* brought him into conflict with Cardinal Mazarin, the French first minister. This opposition also set Arnauld against the decision in Rome to condemn five propositions purportedly drawn from the *Augustinus* in 1653 and to attribute those propositions to Jansenius's text in 1656. Arnauld criticized those who refused absolution to the Duc de Liancourt because of his failure to assert that Jansenius affirmed these propositions. For his efforts, Arnauld was excluded from the Sorbonne in 1656, after a celebrated trial. In defense of Arnauld, the Port-Royal solitaire Blaise Pascal wrote a series of *Lettres provinciales* (1656–1657) attacking the moral theology of the Jesuits. In further response to the 1656 papal bull attributing the condemned propositions to the *Augustinus*, Arnauld argued that, though the pope's word is definitive with respect to the *question de droit* regarding the unacceptability of the propositions, it is not authoritative with respect to the *question de fait* concerning the presence of the propositions in Jansenius. He advocated a "respectful silence" in response to the pope's opinion on the latter question.

LATER CAREER. After 1661, when Louis XIV took sole control of the government following the death of Mazarin, considerable pressure was placed on those associated with Port-Royal to bring them into conformity with the official church rejection of Jansenism. This pressure involved the closing of the *petite écoles* at Port-Royal, but the instruction there informed two books that Arnauld coauthored with fellow Port-Royalists, the *Grammaire generale et raisonée*, which he authored with Claude Lancelot in 1660, and the *Logique ou l'art de penser*, which he authored with Pierre Nicole in 1662. Noam Chomsky (1966) emphasizes the importance of the view in the former work that there is an innate "universal grammar" responsible for language (compare Arnauld and Lancelot 1975). The latter work served as a popular Cartesian alternative to scholastic texts on logic, and indeed the University of Paris formally adopted it in 1720 for use with Descartes's *Meditations*.

In 1669 the campaign against Jansenism was brought to a temporary end by the Peace of the Church that Pope Clement IX established in concert with Louis XIV. During this temporary truce, which allowed for the respectful silence concerning the heretical nature of the *Augustinus*, Arnauld turned his attention to his work with Nicole on the three-volume *Perpétuité de la foi* (1669–1674), which attacked the Eucharistic theology of the Calvinists. In 1672 Arnauld met the German intellectual Gottfried Wilhelm Leibniz during the latter's visit to Paris, and in 1679 he met his fellow Cartesian, Nicolas Malebranche. Both of these meetings set the stage for important later exchanges on philosophical and theological matters.

In 1679 Louis XIV forced Arnauld to leave France, bringing to an end the Clementine Peace. Arnauld took up residence in the Spanish Netherlands, where he lived the rest of his life. In 1683 he composed a critique of Malebranche's *Recherche de la vérité*, which triggered a long and increasingly bitter dispute with Malebranche over issues concerning the nature of ideas and of grace and divine providence. In 1686 Arnauld began a brief but important correspondence with Leibniz on a summary of Leibniz's *Discourse on the Metaphysics and the Monadology*. This correspondence addressed issues concerning the

nature of divine freedom and creation as well as the tenability of a Cartesian conception of the material world.

In 1690 Arnauld succeeded in his campaign to have certain works of Malebranche placed on the Roman *Index of Prohibited Books*. During this same time he engaged in disputes with Nicole over general grace and one's knowledge of moral truth. Arnauld was also involved in several disputes with the Jesuits in the Spanish Netherlands. He died in Brussels on August 8, 1694, and was buried in the Church of St. Catherine in that city. His heart was buried in Port-Royal, and after the destruction of the latter in 1710 it was moved to the Church de Palaiseau.

FAITH AND FREEDOM

Arnauld was fond of the Augustinian slogan that "what we know, we owe to reason; what we believe, to authority" (1964–1967, p. 38:94). This slogan reflects Arnauld's own view that philosophy and theology are distinct disciplines with their own standards. Philosophical questions are to be resolved through the use of reason, and he took issue with scholastics who attempted to settle such questions by means of an appeal to the authority of Aristotle. In contrast, Arnauld insisted that questions pertaining to religious belief, and in particular to the content of the Catholic faith, are to be decided by an appeal to the authority of Scripture, interpreted in light of the church tradition. Here, he took issue with Jesuit critics who attempted to use their Aristotelian philosophy to explicate the mysteries of the faith.

Arnauld did recognize a distinction between "sacred theology" concerning Catholic doctrine and "natural theology" concerning theological truths accessible to reason. Indeed, one of the reasons he defended Cartesian philosophy so vigorously, even in the face of opposition from his fellow Port-Royalists, was that he took it to provide compelling arguments for the existence of a transcendent God and for the real distinction of the human soul from body. For Arnauld, Descartes's theistic and dualistic system complemented perfectly a theology based on the authority of Augustine.

Arnauld began by defending the particular version of Augustinian theology in Jansen's work. In particular, he was concerned to argue with Jansenius for the view that meritorious action is the result of grace that is "efficacious in itself," that is, that brings about the relevant action. This position conflicted with the Jesuit insistence on one's ability to freely reject the divine grace that is offered. In the last decade of his life, however, Arnauld rejected Jansen's account of grace in terms of a prevenient state of delight that causes the meritorious action. He claimed to find in St. Thomas Aquinas the alternative view that efficacious grace is simply the meritorious act of will that God produces in each person. His final position did not bring him closer to the Jesuits, however, and is in fact similar to the view of the Dominican Domingo Bañez, which the Jesuits had opposed, that God causally determines free human action.

EUCHARISTIC THEOLOGY

The other theological issue of most importance to Arnauld concerned the Catholic doctrine of the Eucharist. According to the Council of Trent there is in this sacrament a "marvelous and unique change of the whole substance of the bread into the body [of Christ], and of the whole substance of the wine into the blood, while only the appearances [*species*] of bread and wine remain." Arnauld and Nicole composed *Perpétuité de la foi*, in which they defended the Tridentine doctrine against the Calvinist position that Christ has a merely "spiritual presence" in this sacrament. In 1680 Arnauld wrote in defense of the compatibility of Descartes's view with Catholic teachings on the Eucharist to silence critics, including some Port-Royalists, who charged that Cartesianism has heretical implications. His "Examen" considers a text in which it is argued that since Christ's body must be present in the sacrament without its extension, it cannot be the case, as the Cartesian doctrine, that extension constitutes the essence of body. Arnauld countered that Catholic teaching requires only that Christ's body is present without the impenetrability by means of which it is enclosed in a place.

Though Arnauld thought of himself primarily as a theologian, his writings on both human freedom and the Eucharist reflect his ability to grapple with the subtle philosophical issues pertaining to theological topics. This facility with philosophical discourse is revealed as well in his interaction with three of his great philosophical contemporaries: Descartes, Malebranche, and Leibniz.

ARNAULD AND RENÉ DESCARTES

***MEDITATIONS* AND CORRESPONDENCE.** Arnauld's set of objections to the *Meditations* prompted Descartes to comment that he could not have asked for a more perceptive critic. Arnauld was particularly sympathetic to those aspects of the *Meditations* that he took to be in line with Augustinian views of the soul and God. The first two sections of Arnauld's Fourth Objections are in fact devoted to these two topics. He offered penetrating objections in these sections to Descartes's arguments for mind-body distinctness and for the existence of God, as well as

mentioning difficulties concerning Descartes's denial of the souls of nonhuman animals, his discussion of the "material falsity" of sensations, and the circularity of his defense of the truth of clear and distinct perceptions. Still, Arnauld also emphasized the Augustinian nature of Descartes's insistence that the intellect is distinct from and epistemically superior to the senses, and he showed himself to be sympathetic throughout to the central conclusions of the *Meditations*.

In a final section, "Points Which May Cause Difficulty to Theologians," Arnauld insisted that Descartes's principle that proper assent is governed by clear and distinct perception be restricted to intellectual matters to allow for the Augustinian conclusion that one's religious beliefs are grounded in one's acceptance of religious authority. He further noted that what is "likely to give the greatest offense to theologians" is the appearance that Descartes's view that bodily modes are inseparable from the substance they modify conflicts with the Catholic doctrine that in the Eucharist the sensible species of the bread and wine remain without inhering in any substance.

In 1648 Arnauld renewed contact with Descartes while in hiding because of the political controversies in France involving Jansenism. Arnauld asked for clarification on several issues pertaining to the nature of memory, the relation of particular thoughts to the attribute of thought, the duration of mind as a thinking thing, and Descartes's argument for the impossibility of a vacuum. In responding to questions concerning this argument, Descartes cited his view that all truths depend on God's omnipotence in warning against the claim that God cannot create a vacuum. Neither in this correspondence nor in his later exchanges with Malebranche and Leibniz, where this view was broached, did Arnauld take a firm position on Descartes's doctrine that all truths depend on God's will. However, Arnauld did profess himself satisfied with Descartes's responses to his questions concerning the nature of mind and its relation to body, concluding that "what you wrote concerning the distinction between the mind and the body seems to me very clear, evident, and divinely inspired" (1964–1967, vol. 5, p. 186).

QUALIFICATIONS OF CARTESIANISM. In his correspondence with Descartes Arnauld professed satisfaction with Descartes's solution to the problem concerning the Eucharist raised in the Fourth Objections. However, he mentioned as a further difficulty that Descartes's identification of the extension of a body with its quantity seems to conflict with the Catholic teaching that Christ's body is present in this sacrament without local extension. Descartes did not respond to this difficulty, even though in earlier correspondence with the Jesuit Denis Mesland he had proposed that the physical presence of Christ in the Eucharist is explained by the union of His soul with the matter of the elements. This proposal provided the basis for a Cartesian account of the Eucharist in the *Considérations sur l'état present* (1671) by the French Benedictine Robert Desgabets. Louis XIV's confessor declared the *Considérations* to be heretical, and Louis had his archbishop of Paris condemn it. When called before the archbishop, Arnauld and Nicole denounced the work, in part to disassociate it from their own account of the Eucharist in their writings against the Calvinists. In his later 1680 "Examen" Arnauld did offer his own version of a Cartesian account of the Eucharist. However, his version deviates from Descartes's own views insofar as it requires the possibility of the existence of the extension of Christ's body apart from the quantity by means of which it occupies a place.

Arnauld also departed from Descartes's views on human freedom. Although he approved of the account that Descartes provided in the Fourth Meditation, he was less happy with later correspondence in which Descartes attempted to accommodate the Jesuit position that free action involves an indifference that explains the power of the agent to act otherwise. Indeed, in response to Desgabets's claim that Descartes is "exceedingly enlightened in matters of religion," Arnauld responded that Descartes's "letters are full of Pelagianism and, outside of the points which he was convinced by his philosophy—like the existence of God and the immortality of the soul—all that can be said of him to his greatest advantage is that he always seemed to submit himself to the Church" (1964–1967, vol. 1, p. 671). Therefore, Arnauld's theological commitments placed clear constraints on what he could accept from Descartes's own writings.

ARNAULD AND NICOLAS MALEBRANCHE

THE *SEARCH* AND IDEAS. During the early 1670s Arnauld was on friendly terms with his younger Cartesian colleague, Malebranche. He also had an initially positive view of Malebranche's masterwork, the *Search after Truth* (first published 1674–1675). After seeing an initial draft of Malebranche's *Treatise on Nature and Grace* (1680), however, Arnauld had a more negative view. In a meeting with Malebranche in 1679, just before he left France for good, Arnauld took exception to the claim in that work that though God wills that all be saved, his wis-

dom requires that he distribute grace by means of a "general will" that allows for the salvation of only a few. Arnauld objected that this emphasis on the role of the general will is a novelty that is out of line with the view, deriving from the work of the church fathers, that God exhibits a "particular providence" in distributing grace to those whom he has predestined for salvation.

After Malebranche decided to publish his *Treatise* in 1680, Arnauld decided to go public with his criticisms of Malebranche. The public debate began with the publication in 1683 of Arnauld's *On True and False Ideas*, and it lasted until Arnauld's death in 1694. During his lifetime Arnauld published eight critiques of Malebranche and Malebranche published seven responses. A further text from Arnauld was published after his death, and Malebranche published three further responses to Arnauld, with the last appearing in 1709. The debate ranged over several topics, the most well known being the nature of ideas, but included as well the relation of "intelligible extension" to God, the relation of pleasure to happiness, the nature of causation, miracles, the efficacy of grace, divine providence, and divine freedom.

The issue of ideas is prominent at the start of the debate, for Arnauld's *On True and False Ideas* focuses on the doctrine in the *Search after Truth* that "we see all things in God," and more specifically, that one perceives bodies by means of ideas that exist in the divine intellect. For Arnauld, such a doctrine has the "bizarre" consequence that "we see God when we see bodies, the sun, a horse, or a tree" (1964–1967, p. 38:236). Still, Arnauld objected not only to the placement of ideas of material objects in God but also, and more basically, to the reification of the ideas. As an alternative to Malebranche's claim that the ideas one perceives are "representative beings" distinct from one's perceptions, he offered the position, which he claimed to find in Descartes's Third Meditation, that such ideas are merely the "objective reality" of perceptions, that internal feature of the perceptions in virtue of which they represent particular objects. Malebranche sometimes offered a different reading of this text, on which the objective reality of an idea is something distinct from the perception as a modification of mind. However, he typically appealed not to Descartes but to the view, which he claimed to find in Augustine, that "archetypes" in the divine intellect serve as the principle of one's knowledge of objects. In response, Arnauld insisted that it was never Augustine's intention to hold that one apprehends features of God's essence in perceiving objects.

GOD AND GENERAL WILL. The debate over the nature of ideas held the attention of the early modern intellectual community, with philosophers as diverse as John Locke, Leibniz, Pierre Bayle, and Pierre-Sylvain Regis offering commentaries on it. Indeed, Arnauld's friend, Nicole, claimed that the preoccupation with the topic of ideas served to divert attention from more important theological issues. Even so, most of the exchanges between Arnauld and Malebranche concerned just such issues. As discussed earlier, Arnauld's initial concerns with Malebranche's system derived from the claim in Malebranche that God distributes grace by means of His "general will." But Arnauld also objected that the stress on the generality of God's action undermined the belief in miraculous exceptions to the natural order. Most fundamentally, Arnauld was worried that the introduction of Malebranche's impersonal "God of the philosophers" would displace the "God of Abraham, Isaac, and Jacob," a God who takes a personal interest in the welfare of His creatures. Arnauld held that the latter sort of God governs by means of particular volitions, even in the case where He acts in accord with general laws.

Arnauld further protested against the suggestion in Malebranche that God cannot act to correct certain deficiencies in creation since He is constrained to act by means of His general will. For Arnauld, such a suggestion involves an unacceptable limitation of God's freedom. On this point Arnauld showed some sympathy for considerations that led Descartes to affirm that God is not constrained by the eternal truths since they derive from his free will. Even so, he never did explicitly affirm this doctrine in his exchange with Malebranche. One can speculate that Arnauld was reluctant to endorse this philosophical position due to his uncertainty about its implications for theology. This would at least be in keeping with his concern, evident in his long debate with Malebranche, to purify theology of various novelties deriving from philosophy.

FREEDOM AND CAUSATION. Arnauld's lifelong preoccupation with theological issues involving Jansenism is reflected in his objections to the view in Malebranche that meritorious action involves one's free and undetermined "consent" to the promptings of divine grace. Arnauld commented that he did not think that "Pelagius ever said anything more pelagian" (1964–1967, p. 37:648f). Though Arnauld later retracted his original endorsement of the view in Jansenius that this consent is determined by a psychological state of delight deriving from grace, he consistently held that such consent must be determined by God's action. It is interesting, however, that Arnauld at

the same time took exception to the occasionalist position in Malebranche that God is the only real cause and that creatures serve merely as inefficacious "occasional causes" for the exercise of divine power. Since Arnauld held that mind is "more noble than" body, and since he accepted the Augustinian dictum that the less noble cannot act on the more noble, he allowed that bodily events can be only occasional causes of changes in mental states. He apparently saw no difficulty in allowing for the action of bodies on each other or the action of mind on body.

ARNAULD AND GOTTFRIED WILHELM LEIBNIZ

FATALISM AND ACTUALISM. On February 11, 1686, in the midst of Arnauld's polemical exchanges with Malebranche, Leibniz sent a request to the Landgrave Ernst von Hessen-Rheinfels to pass along to Arnauld "a short discourse" on "questions of grace, the concourse of God and creatures, the nature of miracles, the cause of sin, the origin of evil, the immortality of the soul, ideas, etc." (Mason 1967, p. 3). This discourse was simply a list of the titles of the thirty-seven articles of what became Leibniz's *Discourse on Metaphysics*. Arnauld engaged somewhat reluctantly with Leibniz on the content of some of these articles. In the end the two exchanged through Hessen-Rheinfels some dozen letters before Arnauld, preoccupied with other matters, failed to respond to Leibniz's letter to him of October 9, 1687. Leibniz attempted to reengage the correspondence in 1688 and 1690 letters to Arnauld, but without success.

In his initial response to Leibniz, Arnauld took exception to the claim in the title to article thirteen of Leibniz's discourse that "since the individual concept of each person contains once for all everything that will ever happen to him, one sees in it the a priori proofs or reasons for the truth of each event, or why the event has occurred rather than another," even though such truths "are nevertheless contingent, being based on the freewill of God and creatures" (Mason 1967, p. 5). He held that this claim is "shocking" since it seems to imply that everything that happens is obliged to do so "through a more than fatal necessity" (p. 9). In particular, God would have no choice, having decided to create Adam, to create all the features of the world that Adam actually inhabits.

After Leibniz bitterly rejected the charge of fatalism and some further letters were exchanged, Arnauld withdrew his charge in a letter of September 28, 1686. His willingness to do so was prompted by Leibniz's insistence that certain truths that are present in the individual concept of a person are present there only contingently. Even so, Arnauld mentioned in this letter that he still had qualms about Leibniz's conception of God as "having chosen the universe amongst an infinite number of other possible universes which he saw but did not wish to create." A hint concerning the source of these qualms is provided by Arnauld's insistence in an earlier letter that God's omnipotence, being a "pure act," does "not permit the existence in it of any possibility" (Mason 1967, p. 31f). On Arnauld's view here, possibilities pertain only to the substances that God has freely created. On the basis of such a view, one commentator claims to find in Arnauld an "actualism" that contrasts with a "possibilism" in Leibniz that allows for possibilities founded in nothing external to the divine intellect (Nelson 1993). A further development of this sort of actualism may have led Arnauld to endorse some version of Descartes's doctrine of the creation of the eternal truths. As in the case of his debate with Malebranche, however, Arnauld failed in his correspondence with Leibniz to take any explicit stand on this doctrine.

CONCOMITANCE AND CARTESIANISM. In contrast to his treatment of Leibniz's critique of the eternal truths doctrine, Arnauld did engage both Leibniz's "hypothesis of concomitance or agreement between substances" and his claim that the reality of material objects depends on their possession of a "substantial form." Arnauld urged that the hypothesis of concomitance is not distinct in the end from the occasionalist position that God brings about the harmony among various substances by means of an eternal act of will. Moreover, he objected to Leibniz's claim that the soul expresses everything in its body on the Cartesian grounds that the soul must have some thought or knowledge to express anything. Since the soul has no more thought or knowledge "of the movements of lymph in the lymphatic vessels than of the movements of Saturn's satellites" (Mason 1967, p. 132), it cannot intelligibly be said to express this aspect of its body. Arnauld's Cartesianism is also evident in his response to Leibniz's position that to be substantial, material objects must have a unity conferred on them by an immaterial and indivisible substantial form. Assuming the Cartesian identification of matter with extension, Arnauld held that all material objects are mere composites and that their unity derives not from Leibniz's substantial form but from the functional interrelation of their parts.

In the note to Hessen-Rheinfels accompanying his final letter to Leibniz, Arnauld expressed the opinion that it would be "preferable" if Leibniz, a lifelong Protestant, "gave up, at least for a time, this sort of speculation, and applied himself to the greatest business he can have, the

choice of the true religion" (Mason 1967, p. 138). This opinion indicates Arnauld's own preference for theology over philosophical speculation. Even so, his philosophically rich exchanges with Descartes, Malebranche, and Leibniz provide reason for philosophers to be grateful that he did not give up philosophical speculation entirely in the interests of furthering acceptance of the Catholic faith.

See also Aristotle; Augustine, St.; Augustinianism; Bayle, Pierre; Cartesianism; Descartes, René; Desgabets, Robert; Jansenism; Leibniz, Gottfried Wilhelm; Locke, John; Logic, History of; Malebranche, Nicolas; Nicole, Pierre; Pascal, Blaise; Pelagius and Pelagianism; Thomas Aquinas, St.

Bibliography

Arnauld, Antoine, and Claude Lancelot. *The Port-Royal Grammar: General and Rational Grammar*. Translated and edited by Jacques Rieux and Bernard E. Rollin. The Hague: Mouton, 1975.

Arnauld, Antoine, and Pierre Nicole. *Logic, or, The Art of Thinking*. 5th ed. Translated by Jill Vance Buroker. New York: Cambridge University Press, 1996.

WORKS BY ARNAULD

Oeuvres de Messire Arnauld. 42 vols. Brussels: Culture et Civilisation, 1964–1967.

On True and False Ideas, New Objections of Descartes' Meditations and Descartes' Replies. Translated by Elmar J. Kremer. Lewiston, NY: Mellen Press, 1990.

Textes philosophiques. Translated by Denis Moreau. Paris: Presses Universitaires de France, 2001.

WORKS ON ARNAULD

"Antoine Arnauld (1612–1694)." *Philosophe, Écrivain, Théologien, Chroniques de Port-Royal* (Paris) 44 (1995).

Chomsky, Noam. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper and Row, 1966.

Kilcullen, John. *Sincerity and Truth: Essays on Arnauld, Bayle, and Toleration*. New York: Oxford University Press, 1988.

Kremer, Elmar J., ed. *The Great Arnauld and Some of His Philosophical Correspondents*. Toronto: University of Toronto Press, 1994.

Kremer, Elmar J., ed. *Interpreting Arnauld*. Toronto: University of Toronto Press, 1996.

Laporte, Jean. *La doctrine de Port-Royal*. 2 vols. Paris: Presses Universitaires de France, 1923.

Mason, H. T., ed. and trans. *The Leibniz-Arnauld Correspondence*. Manchester, U.K.: Manchester University Press, 1967.

Moreau, Denis. *Deux Cartésiens, La Polemique entre Antoine Arnauld et Nicolas Malebranche*. Paris: J. Vrin, 1999.

Nadler, Steven M. *Arnauld and the Cartesian Philosophy of Ideas*. Princeton, NJ: Princeton University Press, 1989.

Ndiaye, Aloyse Raymond. *La Philosophie d'Antoine Arnauld*. Paris: J. Vrin, 1991.

Nelson, Alan. "Cartesian Actualism in the Leibniz-Arnauld Correspondence." *Canadian Journal of Philosophy* 23 (4) (1993): 675–694.

Pariente, Jean-Claude. *L'analyse du langue à Port-Royal: Six études logico-grammaticales*. Paris: Les Éditions de Minuit, 1985.

Sleigh, R. C., Jr. *Leibniz and Arnauld: A Commentary on Their Correspondence*. New Haven, CT: Yale University Press, 1990.

Yolton, John W. *Perceptual Acquaintance: From Descartes to Reid*. Minneapolis: University of Minnesota Press, 1984.

Tad M. Schmaltz (2005)

ARNOLD, MATTHEW

(1822–1888)

Matthew Arnold, the English poet and social and literary critic, was the son of Dr. Thomas Arnold, headmaster of Rugby. Matthew Arnold was educated at Winchester and Rugby and entered Balliol College, Oxford, in 1841. In 1847 he became private secretary to Lord Lansdowne, who in 1851 appointed Arnold inspector of schools, a position he held until 1886. In 1857 he was elected professor of poetry at Oxford.

As a critic, Arnold ranged over a broad spectrum from literary criticism through educational theory to politics, social thought, and religion.

Arnold's most important contribution to nineteenth-century thought was his discussion of the significance of culture as a social ideal. His related discussion of the function of criticism has been widely influential. He also contributed to the dispute over the relation between the Christian Scriptures and belief.

In *Culture and Anarchy* (London, 1869), Arnold defined "culture" as "a pursuit of our total perfection by means of getting to know, on all the matters which most concern us, the best which has been thought and said in the world; and, through this knowledge, turning a stream of fresh and free thought upon our stock notions and habits." Culture is thus a process of learning, which can refine individuals and reform societies. Arnold often attacked the kind of reforming or progressive spirit that is not governed by this humane reference. At the same time, he made it clear that the object of the learning and refining process was indeed reform. He laid great stress on the development of the individual through the right use of literature and knowledge, but the pursuit of total perfection was still the ultimate objective. He argued that culture taught men "to conceive of true human perfection as a *harmonious* process, developing all sides of our humanity; and as a *general* perfection, developing all parts of our

society." Perfection, although an "*internal* condition," is nevertheless "not possible while the individual remains isolated. The individual is required, under pain of being stunted and enfeebled in his own development if he disobeys, to carry others along with him in his march towards perfection, to be continually doing all he can to enlarge and increase the volume of the human stream sweeping thitherwards."

This position illuminates some of the apparent paradoxes of Arnold's thinking. In one sense, he was clearly a liberal thinker, stressing the criticism of institutions and beliefs by thought and knowledge and placing central emphasis on the development of the individual toward a possible perfection. In other respects Arnold was a notable critic of much of the liberal thought of his time. He criticized the "stock notions" of nineteenth-century liberalism and was a particularly firm advocate of increased social intervention by the state. He criticized the common liberal conception that progress is merely mechanical and the liberals' preoccupation with material and external improvement, which not only ignored the human results of its materialist emphasis, but also failed to advance any conception of humanity toward the realization of which material progress might be a means. His criticism of the "stock notions" of industry and production as major social ends is of this character. He similarly criticized the standard conception of freedom—"a very good horse to ride, but to ride somewhere." It is the way men use freedom, not merely their abstract possession of it, that for Arnold is really important.

Most liberal thought in his time opposed the state in the name of just this kind of abstract freedom. Arnold argued that the state was simply "the representative acting-power of the nation." To deny its right to act was to deny the possibility of any general action on behalf of the nation as a whole and to reserve the power of action to particular interests and classes. In the England of his time, he distinguished three classes—the aristocracy ("Barbarians"), the middle classes ("Philistines"), and the working class (the "Populace"). Social action by any one of these interests alone merely led to the clash of men's "worst selves." This disorder, or the resultant breakdown of effective government, would be "anarchy." But there existed, within each of these classes, "persons who are mainly led, not by their class spirit, but by a general *humane* spirit, by the love of human perfection." Each member of this human "remnant," maintaining his own "best self" by the process of culture and seeking to awake in others the "best self" now obscured by the "stock notions" and habits of the group, represented the "best self" of society as a whole. It was this "best self" that the state must represent and express.

Arnold never translated these ideas into a coherent political philosophy, but his liberal critique of liberalism was of considerable historical importance. The state, he felt, had to become a "centre of authority and light"; yet it must do this through the existing struggle, or deadlock, between limited interests and classes. Arnold's arguments, at this point, were sometimes vague. In line with his definition of culture as a learning process and with his career as inspector of schools, he stressed not politics, but education. It was in education that the state most needed to intervene, and Arnold acted as a tireless propagandist for a new system of humane state education.

Arnold saw the study of literature as a principal agency of the learning process, that is, of culture. At times, his definitions of criticism and of culture were virtually identical. Criticism was the central way of learning "the *best* that is known and thought in the world." Poetry in particular offered standards for the development of the best life of man.

In the same vein, in *Literature and Dogma* (London, 1873) Arnold offered to "reassure those who feel attachment to Christianity, to the Bible, but who recognise the growing discredit befalling miracles and the supernatural." For any adequate reading of the Bible, after the effects of the Higher Criticism and the scientific controversies, the spirit of culture was indispensable. Only by this approach could the Christian ethic, and its intense expression in the Scriptures as read undogmatically, be preserved in a time of inevitable change. In particular, it was necessary to understand that "the language of the Bible is fluid, passing and literary, not rigid, fixed, and scientific"; its truth had to be verified through reading, rather than merely assumed. The Christian ethic so verified would be stronger than the dogmatic theology that had made the Bible into what it evidently was not.

See also Belief; Literature, Philosophy of; Perfection.

Bibliography

WORKS BY ARNOLD

Poetical Works. Edited by C. B. Tinker and H. F. Lowry. Oxford: Oxford University Press, 1950.

Essays in Criticism. London: Macmillan, 1865.

God and the Bible. London: Smith, Elder, 1875.

Mixed Essays. New York: Macmillan, 1879.

Passages from the Prose Writings of Matthew Arnold, Selected by the Author. London, 1880. Edited by W. E. Buckler. New York: New York University Press, 1963.

Discourses in America. London: Macmillan, 1885.

Essays in Criticism, Second Series. London: Macmillan, 1889.

WORKS ON ARNOLD

Bonnerot, Louis. *Matthew Arnold, Poète.* Paris: Didier, 1947.

Brown, E. K. *Matthew Arnold.* Toronto, 1948.

Eliot, T. S. "Matthew Arnold," in *The Use of Poetry and the Use of Criticism.* Cambridge, MA, 1932.

Trilling, Lionel. *Matthew Arnold.* New York: Norton, 1939.

Raymond Williams (1967)

AROUET, FRANÇOIS-MARIE

See *Voltaire, François-Marie Arouet de*

ART, AUTHENTICITY IN

In the main sense of the term an artwork is "authentic" if it is the artwork it is thought to be—if it has the history of production it is represented as having or gives the impression of having, if it was created where, when, how, and by whom it is supposed or appears to have been created. Thus, a work may be inauthentic in virtue of being a forgery, or a misattribution, or a replica not identified as such. A reproduction (e.g., in an art book) is inauthentic only in a weaker sense: Though not the artwork it reproduces, it does not purport to be and runs no danger of being confused with it.

The chief issue concerning the authenticity of artworks has been the extent to which a work's aesthetic properties, artistic value, and proper appreciation legitimately depend on questions of authenticity in the above sense. The issue is often framed in terms of a challenge: What is wrong with a forgery? or What privileges an original artistically?

Broadly speaking, there are two opposed views on this issue. On one view an artwork is merely a perceivable structure—for example, a constellation of colors and shapes, a set of notes, a string of words, or the like. Furthermore, this structure is the entire source of its aesthetic and artistic properties and is the only thing relevant to its appreciation and evaluation as art. Thus, anything preserving the artwork's perceivable structure, so as to be perceptually indiscernible from it, is equivalent to it artistically and even ontologically. Such a view underlies the formalism of Clive Bell and Roger Fry, the literary stance of the New Critics, and to some extent the aesthetics of Monroe Beardsley. By these lights there is nothing much wrong with a forgery—provided, of course, that it is a perfect one, not detectably different from the original.

On the other view perceivable structure is not the sole determinant of a work's aesthetic complexion or its artistic character. Rather, a work's context of origination, including the problematic from which it issues, partly determines how the work is rightly apprehended and experienced and thus its aesthetic and artistic properties. Aspects of the context or manner of creation arguably enter even into the identity of the work of art, as essential to its being the particular work it is. By these lights there is quite a lot wrong with a forgery. It differs from the original in numerous respects, both aesthetic and artistic, and as a human product—a making, an achievement, an utterance—it is of an entirely different order, however similar it appears on superficial examination.

If the second view sketched above is sound, then any artwork, *pace* Nelson Goodman, can be forged—that is, represented as having a provenance and history other than its own, though how this will be effected differs from art form to art form, especially when one crosses from particular arts (such as painting) to type arts in which structure may be notationally determined (such as music). And this is because, in all art forms, the identity of a work is partly a matter of the historical circumstances of its emergence.

Goodman famously argued, against the aesthetic equivalence of an original painting and an ostensibly perfect forgery, that the possibility of discovering a perceptual difference between the former and the latter constitutes an aesthetic difference between them. Unfortunately, this argument seems to trade on conflating an aesthetic difference and an aesthetically relevant difference between two objects. As suggested above, however, the aesthetic and artistic differences between originals and forgeries, which are ample, rest securely on quite other grounds.

AUTHENTICITY OF ARTWORK INSTANCE

In cases of multiple or type arts an instance of a work—a copy, impression, casting, performance, staging, screening, and so forth—may be denominated authentic or inauthentic insofar as it is or is not a correct or faithful instance of the work. And this, according to different accounts, is a matter of its adequately instantiating and representing the structure thought definitive of the work in question, a matter of its having the right sort of causal or intentional relations to the work in question or of being produced in a certain manner, a matter of its con-

veying the aesthetic or artistic properties believed crucial to the work—or some combination of these.

AUTHENTICITY OF ARTIST

Finally, authenticity is sometimes considered a predicate of the artist, describing laudatorily the artist's characteristic mode of creating or the relation between the artist and the content of the works the artist creates. An authentic artist is one thought, variously, to be sincere in expression, pure in motivation, true to self, honest about medium, rooted in a tradition, resistant to ideology yet reflective of society—or all of these. There seems to be only a passing relation between authenticity in this sense and the authenticity of work or instance canvassed above.

See also Aesthetics, History of; Aesthetics, Problems of; Art, Truth in; Art, Value in; Beardsley, Monroe; Goodman, Nelson.

Bibliography

Bailey, G. "Amateurs Imitate, Professionals Steal." *Journal of Aesthetics and Art Criticism* 47 (1989): 221–228.

Baugh, B. "Authenticity Revisited." *Journal of Aesthetics and Art Criticism* 46 (1988): 477–487.

Currie, G. *An Ontology of Art.* New York: St. Martin's Press, 1989.

Danto, A. *Transfiguration of the Commonplace.* Cambridge, MA: Harvard University Press, 1981.

Davies, S. "Authenticity in Musical Performance." *British Journal of Aesthetics* 27 (1987): 39–50.

Davies, S. "The Ontology of Musical Works and the Authenticity of Their Performances." *Noûs* 25 (1991): 21–41.

Dutton, D., ed. *The Forger's Art.* Berkeley: University of California Press, 1983.

Godlovitch, S. "Authentic Performance." *Monist* 71 (1988): 258–277.

Goodman, N. *Languages of Art,* 2nd ed. Indianapolis: Bobbs Merrill, 1976.

Goodman, N. *Of Mind and Other Matters.* Cambridge, MA: Harvard University Press, 1984. Chap. 4.

Kennick, W. E. "Art and Inauthenticity." *Journal of Aesthetics and Art Criticism* 44 (1985): 3–12.

Levinson, J. "Art, Work of." In *The Dictionary of Art.* London, 1995.

Levinson, J. *Music, Art, and Metaphysics.* Ithaca, NY: Cornell University Press, 1990.

Sagoff, M. "Historical Authenticity." *Erkenntnis* 12 (1978): 83–93.

Sagoff, M. "On Restoring and Reproducing Art." *Journal of Philosophy* 75 (1978): 453–470.

Sartwell, C. "Aesthetics of the Spurious." *British Journal of Aesthetics* 28 (1988): 360–367.

Savile, A. "The Rationale of Restoration." *Journal of Aesthetics and Art Criticism* 51 (1993): 463–474.

Thom, P. *For an Audience: A Philosophy of the Performing Arts.* Philadelphia: Temple University Press, 1993.

Walton, K. "Categories of Art." *Philosophical Review* 79 (1970): 334–367.

Wollheim, R. *Art and Its Objects,* 2nd ed. Cambridge, U.K.: Cambridge University Press, 1980.

Jerrold Levinson (1996)

ART, DEFINITIONS OF

A range of related topics are gathered together under the title "The Definition of Art." These include: (1) metaphysical questions, such as "Is there a set of necessary properties whose possession is conjointly sufficient for a candidate to qualify as an artwork?" and, if so, "What are they?"; and (2) the epistemological issue of how we go about establishing that a candidate is an artwork. Traditionally the default assumption among many philosophers has been that there are necessary and sufficient conditions for classifying things as artworks; that these conditions can be assembled into a real or essential definition of art; and that the application of the aforesaid definition provides us with the means to establish that this or that candidate is an artwork. The trick with this approach is to specify, successfully, the pertinent necessary and sufficient conditions.

Needless to say, this enterprise has turned out to be more challenging than one might have anticipated. And the difficulties encountered in successive attempts to carry off this endeavor have left some philosophers either skeptical or agnostic regarding the prospects of the metaphysical project of defining art. Instead, they have tried more modestly merely to identify the epistemological grounds for classifying candidates as artworks without resorting to real definitions.

The search for a definition of art was not something that taxed ancient philosophers like Plato and Aristotle. For Aristotle, art was skill with respect to any practice or craft. There was an art of poetry and an art of painting, but also the art of medicine, navigation, warfare, and so on. Though Plato, Aristotle, and Horace compared poetry and painting, they did not presume an overarching framework that groups certain arts (in their sense) together in the category that we now call the fine arts or beaux arts, or maybe more simply just Art with a capital *A*—roughly, poetry (literature), painting, sculpture, music, theater, dance, architecture, and, nowadays, photography, film, and video.

The system of the arts was not stably consolidated until the eighteenth century (Kristeller 1992). Thus, it comes as no surprise that Aristotle felt no inclination

toward defining the conditions for membership in the category, though he did analyze the nature of some of the things—like tragedy—that would later be subsumed under the concept of (fine) art. For the ancients, there were arts that were tied to certain functions—quite often religious, political, or otherwise social ones—and these art forms were defined and evaluated in light of that function. For example, Aristotle maintained that the function of tragedy was to educate the emotions by eliciting the catharsis or clarification of pity and fear.

When Aristotle and Plato single out mimesis or imitation as a necessary feature of drama—both tragedy and comedy—and painting respectively, it is immensely unlikely that they were attempting to isolate the essence of art in our sense of fine art. It is more plausible to suppose that they were merely singling out a necessary condition of the relevant art forms that is particularly revelatory of the point and purpose of these practices. If one wants to understand what poetry and painting are about, or wants to know what is appropriate to expect from them, the concept of imitation is central. However, when Plato speaks of mimesis in poetry and painting, he is not offering an analysis or definition of what we mean by (fine) art or even a real or essential definition of poetry or painting. Rather, he is merely pointing to a general feature of these art forms that is especially useful to have in mind, if one hopes to comprehend them and gauge their value.

The pressure to define Art (with a capital *A*) does not arrive on the scene, until the subset of arts mentioned above are separated from the rest and treated as an exclusive confraternity. Perhaps the reason for the emergence of this grouping has to do with the rise of the bourgeoisie who, with leisure time on their hands, turned to these particular arts to fill their hours and days. But, of course, once this grouping took hold, a question arises concerning what property or properties a prospective member needs in order to join the category.

At least initially, it seems that the first gambit for answering this question was that a candidate for membership in the fine arts had, harkening back to Plato and Aristotle, to be an imitation, but, more specifically, an imitation of the beautiful in nature. This view is explicitly advanced in Abbé Charles Batteux's 1746 treatise *The Fine Arts Reduced to the Same Principle* in which the eponymous principle is none other than the imitation of beautiful nature (Beardsley 1966). For something to count as art, then, in the relevant sense of that which the eighteenth century called the fine arts or the beaux arts (and what we now simply call Art with a capital *A*), something had to be the imitation of the beautiful, though it seems that sometimes this requirement was slackened to no more than that the art form in question had to be representational. If the art form in question was representational, then a work made in accordance with this propensity of the pertinent art form was an artwork. That is, a painting that is a picture is, all things being equal, a work of art.

This definition of art—often called the representational theory of art (Carroll 1999, Chap. 1)—was certainly ill-suited for the developments in the arts to come, for example: An abstract expressionist painting is not a representation of anything and especially not an imitation of something beautiful in nature (Carroll 1999). So the definition was fated to be incessantly accosted by counterexamples in the future. But, perhaps more to the point, the representational theory of art was not even viable in its own day.

Dance, for instance, belonged to the system of the fine arts; it had its own muse, Terpsichore. However, not all dance, even in the eighteenth century, was representational. Much dance involved no more than cadenced steps, gracefully executed. In fact, in order to legitimatize a place for dance in the newly anointed system of the arts, choreographers, like Georges Noverre, had to invent the *ballet d'action*—the ballet that told a story. But in cases like this, the definition of art as a matter of representation, in fact, functioned prescriptively rather than descriptively.

But an even greater embarrassment for the representational theory of art than dance was the emergence of absolute music—that is to say, pure orchestral music. When opera and song were the dominant forms of music, music could be counted as implicated in representation because the words that accompanied the notes referred. But once absolute music took pride of place in the order of Calliope, it became very strained to think of the imitation of nature as the essence of art status. Indeed, as absolute music came in the nineteenth century to be praised for its possession of a condition to which all the other arts aspired, it became less and less credible that imitation was a necessary condition for entry into the citadel of art. Though some swatches of Beethoven's *Pastorale* symphony are imitative, most of the rest of his purely musical oeuvre is not. If for no other reason than the ascendancy of absolute music, the representational theory or definition of art was clearly inadequate. Another approach was needed.

Consonant with the reigning artistic movement of the day, Romanticism, one alternative approach to the

representational theory of art was the expression theory—the view that something is an artwork only if it is the expression of an emotion or a feeling (Carroll 1999). Variations of this view have been defended by figures such as Leo Tolstoy (Tolstoy 1996), and R. G. Collingwood (Collingwood 1938).

If the representational theory of art emphasized the representation of the outer world, the expression theory of art stressed art as the presentation of the inner world of the affective life. William Wordsworth asserted that poetry is "the spontaneous overflow of powerful feelings," and this was also thought to be applicable to the other arts. It certainly appeared to fit the absolute music that is now called Romantic. Is that not why it is called Romantic? Moreover, the expression theory seemed to resist counterexamples insofar as it might be supposed that any human artifact would unavoidably carry an expressive trace of the affects of its maker.

Nevertheless, counterexamples appeared in droves starting in the early twentieth century. One source of these counterexamples were various sorts of aleatoric art; the Dadaist Tristan Tzara composed poems by cutting out words from a newspaper, placing them in a hat, and drawing them out randomly—thus thwarting the possibility of any causal connection with what he was feeling. Related chance techniques were mobilized by the surrealists and artists like John Cage and Merce Cunningham. Another kind of counterexample to the expression theory derived from found artworks an ordinary comb presented as an artwork by the likes of Marcel Duchamp, which projects no expressive properties, let alone the trace of anything felt by Duchamp.

Nor could these counterexamples be blocked by appealing to the idea that every human product bears an emotive residue from its maker, for the preceding strategies incontrovertibly sever the emotional link between the artist and the art object. Moreover, the expression theory of art would not only be challenged by the artists of the twentieth century. The theory was undermined by certain forms of art already in existence in the heyday of Romanticism, including art that aspired simply to beauty, as in the case of decorative art, perhaps some absolute music, and art that aimed only to represent the look of the world objectively.

Defenders of the expression theory might attempt to fend off these examples by invoking the claim that there is an inevitable and manifest emotive tie between any artifact and its creator. However, not only does this appear controversial, but if it were so, then the theory would be far too broad to be a satisfactory definition of art because it would fail to differentiate an artwork from any other artifact.

Around the same time that expression theories begin to make their appearance, so, too, do two alternative accounts of art derived from Immanuel Kant's *Critique of Judgment*. These theories can respectively be called the formalist theory of art and the aesthetic theory of art (Carroll 1999). Formalism, as presented by someone such as Clive Bell (1914), maintains that something is an artwork if and only if it is designed primarily to possess a formal design (called significant form) that is worthy of contemplation for its own sake. That is, the form of the work is intended, first and foremost, to afford an aesthetic experience, (which is sometimes called an experience of disinterested pleasure pursuant to contemplating the work's design).

The aesthetic theory of art (Beardsley 1983) is like formalism except that it leaves the object of experience unspecified by making no reference to the form of the work. On this view, something is an artwork if, and only if, it is made primarily with the intention to support an appreciable amount of aesthetic experience (in other words, experience valued for its own sake). Both the formalist theory of art and the aesthetic theory of art make essential reference to intentions in order to differentiate artworks from natural scenes that might give rise to aesthetic experience. With their emphasis on experiences valued for their own sake, both these views may actually articulate the motive behind the modern category of art as a grouping of the things suitable for leisured contemplation and/or diversion.

Neither formalism nor the aesthetic theory of art provides necessary conditions for classifying candidates as artworks. For it is implausible to suppose that most religious artworks were created with the primary intention of abetting experiences valuable for their own sake. Rather, like so many other premodern artworks, they were produced to perform a function. They were created with the primary intention of advancing religious purposes. Paintings of Christ's crucifixion were intended to instill reverence; they were not meant to be occasions for intrinsically valuable experiences of painterly form. And the designs on the shields of the Sepik warriors of New Guinea were not drawn in order to engender experiences valued for their own sake, but with the instrumental aim of frightening the enemy. Nor is experience valued for its own sake a sufficient condition for art status. Games of chess may be said to promote experiences valued for their own sake, but games of chess are not artworks, not even performance artworks.

The successive failures of attempts to define art disposed many philosophers to skepticism about the very venture itself. By the mid-twentieth century, the suspicion, generally encouraged by the writings of Ludwig Wittgenstein, that art could not be defined became popular. Philosophers like Morris Weitz (1956) argued that because art making is an arena in which experimentation, innovation, and novelty are prized, the notion of defining art is incompatible with the practice of art making. For to define art in terms of necessary and sufficient conditions would putatively somehow shackle the essential openness of art to invention and creativity. Philosophers of this ilk, often called neo-Wittgensteinians, maintained that to define art was to contradict the concept of art as that which contained the permanent possibility of art to expand its horizons in new directions. Consequently, neo-Wittgensteinians rejected the metaphysical project of identifying for artworks a set of necessary conditions that were conjointly sufficient. Moreover, with respect to the epistemological question of how it is established that something is an artwork, they suggested that it was a matter of family resemblance; something A is an artwork when it resembles artwork B in some respects, artwork C in other respects, and so on for further paradigmatic artworks.

Though initially quite influential, the spell of the neo-Wittgensteinian brief began to wane by the 1970s. On the one hand, the argument that specifying the conditions according to which a candidate counted as an artwork is inconsistent with the innovative nature of art could be seen to rest on an equivocation. For even if the practice of art is always, in principle, open to innovation and, therefore, supposedly inhospitable to definition, it is not clear why this would stand in the way of defining the concept of an artwork because individual artworks are not typically open to the permanent possibility of change. It just does not follow that if art (in the sense of the practice of art) is an open concept, then art (in the sense of an individual artwork) is an open concept. Moreover, this open concept argument, as it was called, was also challenged by the appearance of definitions of art by people like Arthur Danto (1981) and especially George Dickie (1974), which, though stated in terms of necessary conditions, provided more than ample room for artistic invention, accommodating the entire gallery of works of Dada and its legacy.

Finally, the epistemological wing of neo-Wittgensteinianism also came under fire. Because it relied upon similarity to establish art status and because everything is like everything else in some respect, by means of the family resemblance method one could in fairly short order establish that any candidate is an artwork. For example, Auguste Rodin's *Gate of Hell* and an I-beam about to be shipped from a steel mill are both physical objects, metallic, shaped by human designs, weigh more than 100 pounds, over two feet long, and so on. But all these similarities and more are not enough to warrant calling the I-beam an artwork. Though it may be that in the wake of the found artwork anything can be art, it is not the case that everything *is* art. Nevertheless, the family resemblance method for classifying artworks would appear to force us to conclude that everything is art now.

A common failing of the theories of art as representation, as expression, as form, as well as the family resemblance model for identifying art is that, in each case, art status rests upon some discernible or manifest feature of the object—such as the possession of anthropomorphic or expressive properties, significant form, or similarities with antecedently acknowledged artworks. Perhaps, it was suggested, by Danto and others, that art status rested in some property of art that the eye could not descry. Duchamp's *In Advance of a Broken Arm* and an ordinary snow shovel are putatively indiscernible. Thus, a theory of art that relies on discernible features of artworks cannot hope to cut the difference between them. Rather, the property (or properties) that are constitutive of art status is something perceptually indiscernible.

For Danto (2000), like G. W. F. Hegel, the relevant feature here is *aboutness* in a double sense. Something will be an artwork, on this account, only if: (1) It is about something; and (2) its mode of presentation says something about, makes some comment upon, or advances a point of view concerning whatever it is about. However, this formula is, on the one hand, too exclusive—there are artworks that may be about nothing, but which are simply beautiful or delightful to the senses. On the other hand, Danto's theory may be too inclusive. Though Danto means it to tell us the difference between Andy Warhol's artwork *Brillo Box* and an allegedly indiscernible, though inartistic, one from Proctor and Gamble, surely the ordinary soap pad container in the grocery store meets both of the conditions of Danto's theory of art.

Like Danto, George Dickie is impressed by the thought that the defining features of art might be perceptually indiscernible. This has disposed him to look toward the context that surrounds and frames the work for clues about its status as a work of art. That is, the work does not wear its artistic status on its face; rather, its position in a social framework or institution is the source of its pedi-

gree. This insight has motivated Dickie (1984) to develop a series of what have been referred to as institutional theories of art, the latest version of which he has christened *The Art Circle*. According to Dickie, our concept of art can be captured by five interlocking definitions:

> 1) An artist is a person who participates with understanding in the making of a work of art.
>
> 2) A work of art is of a kind created to be presented to an art world public.
>
> 3) An art world public is a set of persons the members of which are prepared in some degree to understand an object which is presented to them.
>
> 4) The art world is the totality of all art world systems.
>
> 5) An art world system is a framework for the presentation of a work of art by an artist to an art world public.

Even a cursory examination of the preceding set of definitions reveals that it is circular. One needs the concept of an art world to define what counts as a work of art but the concept of a work of art figures in the definition of an art world system, which, in turn, is an element in the definition of an art world. Dickie is aware of this circularity but claims that it is not problematic. Yet it appears to leave the crucial notion of art undefined, though a definition of art was that at which Dickie was aiming.

Dickie's framework does articulate the structure of any communicative practice with its emphasis on mutual understanding. However, what makes art the very communicative practice it is rather than some other, such as philosophy, has not been clarified by Dickie's analysis. Moreover, some, such as Jerrold Levinson, suspect that the model does not even offer a set of necessary conditions for art status because it does not allow for art made by a solitary artist for himself—for example, some Neolithic wanderer who arranges a pile of colored stones in front of his fire because they are delightful to look at as the flames illuminate them variously (Levinson 1979).

Instead of social context, Levinson locates the defining feature of art in the intention of the artist. On Levinson's view, a candidate is an artwork if, and only if, it is created by a person: (1) who has a proprietary right over the work in question; and (2) who nonpassingly intends the work for regard as a work of art (i.e., in one or more of the ways that artworks have been correctly regarded historically [Levinson 1979]). Like Danto and Dickie, Levinson deploys a non-manifest property of the work—a certain kind of intention—as the crux of his definition. Because this intention must be linked to the history of art, Levinson titles his approach *defining art historically*.

It is not clear why Levinson feels compelled to require that artists must have a proprietary right over the work in dispute. Surely if Brancusi constructed a sculpture out of stolen materials, there would be no question that he had created a work of art, even if the ownership of the object was in question. Moreover, the second condition of Levinson's definition is also fraught with difficulties. Though it is called a historical definition, it is historically insensitive. It overlooks the possibility that some historical art regards may become obsolete. For example, appreciating the verisimilitude of a picture was an art regard for centuries, but it is arguably no longer decisive, lest many ordinary family snapshots made with the intention to be appreciated integrally and nonpassingly for their accuracy would, counterintuitively, count as artworks. Unfortunately, Levinson makes no provision for anachronistic art regards.

Like Levinson, Robert Stecker (1997) appeals to history in order to define art. He labels his view *historical functionalism*. It is a disjunctive definition of art. Stecker claims that something is an artwork if, and only if, it is in a central art form at time *t* and it is made with the intention of fulfilling functions standardly or correctly recognized for that form, or it is an artifact that achieves excellence in fulfilling one of the functions of the central art forms at *t*.

This definition seems far too inclusive. According to Pierre Bourdieu, one of the functions of our art form is to produce social capital, or status, or identity. Thus, a Cadillac convertible would be a work of art in virtue of the second disjunct in Stecker's formula. The problem here is that Stecker has not limited the functions he countenances to exclusively artistic functions, but, of course, it is not evident that he can do that readily without inviting circularity.

Historical functionalism is also too exclusive. It cannot assimilate as artworks the initial avant-garde entries of radical art movements, for these works may not belong to a central form of art and they may be designed expressly to repudiate the recognized functions of art at time *t*. Consider the cases of found objects (Duchamp), found music (Cage), and found movement (Yvonne Rainer and Steve Paxton) when they first emerged. They were not obvious examples of a central form and, in any event, they repudiated the functions correctly associated with the forms to which they were related adversarially. Yet certainly any definition of art at this late date must accommodate works such as these.

Perhaps the historical functionalist will attempt to negotiate this shortfall in the theory by saying that once art movements like Postmodern Dance are successful they become—say at time *t+1*—central forms of art with correctly recognized functions; thus, in virtue of the second disjunct of the theory, the originating works of the movement from time *t* can be reclaimed as art. Yet this gambit comes with costs because it has the avant-garde works in question only becoming artworks due to our appreciation of them long after their creators produced them. But surely a dance such as *Satisfyin' Lover* or a composition such as *4'33"* were artworks from the very moment of their inception. And it is their actual creators who imbued them with art status and not some other folks at time *t+1*.

Due to the recurring difficulty with constructing an adequate conceptual analysis of art, some contemporary philosophers are agnostic about the metaphysical prospects of discovering a set of necessary properties that are conjointly sufficient for identifying artworks. Instead they focus their energies upon articulating epistemically satisfactory methods for identifying candidates as artworks which methods are not real definitions. Berys Gaut (2000), mining Wittgenstein again for inspiration, resurrects the notion of a cluster concept, arguing that it is sufficient for classifying a candidate as an artwork that the candidate scores well against the following ten criteria:

1) It possesses positive aesthetic properties.

2) It expresses emotion.

3) It is intellectually challenging.

4) It is complex and coherent.

5) It has the capacity to express complex meanings.

6) It exhibits an individual point of view.

7) It is an original exercise of the imagination.

8) It is the product of skill.

9) It belongs to an established form of art.

10) It is made with the intention to be a work of art.

This is not a real or essential definition of art because none of these properties are necessary conditions for art status. Anything that is a work of art will have at least one of these features; a work that has more and more of these features provides us with more and more reasons to categorize it as an artwork. On this view, a cluster account of a concept is true of that concept just in case it isolates properties whose possession by the work in question necessarily counts toward its belonging to that category. However, though Gaut provides this list of the components of the cluster concept, he does not believe that the cluster concept approach to identifying art stands or falls with his particular sketch of it. He asserts that even if problematic cases for his formulation exist, that should not lead us to distrust in general the cluster concept approach to identifying artworks.

But is Gaut's assertion here convincing? Clearly, there are problem cases with respect to his dissection of the putative cluster concept. I see no reason why a delicious meal made by a master chef to express his devotion to his beloved and to recall their life together by means of culinary references could not instantiate every component of Gaut's list save obviously (9). Indeed, since the preparation of food occasionally figures in certain theatrical works, and especially in examples of performance art, maybe a case could even be made that it satisfies (9), generously construed. It should, therefore, count as a work of art, though this is certainly at least a very controversial case and, for many, a decisive counterexample to Gaut's proposal. But if Gaut's proposal is defeated, why believe that there is some other model of the cluster concept of art that will do the job? If it is inadmissible to maintain that the definitional approach to the concept of art will succeed despite the lack of evidence so far, why should one have faith in the cluster concept approach, when the best version of it so far misses the mark?

Another non-definitional approach to answering the epistemic question of how we might establish that a candidate is an artwork is that we do so by employing historical narratives (Carroll 1993 and 2001). According to what we may call narrativism, establishing that a candidate is an artwork involves telling a certain kind of story about the work in question, namely an accurate historical narrative about the way in which the candidate came to be produced as an intelligible response to an antecedently acknowledged art-historical situation. That is, in order to corroborate the claim that something is an artwork, one standardly mobilizes a narrative explanation of how the work emerged coherently from recognized artistic modes of thinking, acting, composing, decision-making, and so forth already familiar to the practice.

Usually the pressure to establish that something is an artwork arises when there is some dispute over its art status, as frequently occurs with works of the avant-garde. The narrativist observes that these imbroglios are typically managed by recounting art historical narratives that demonstrate the connection between the disputed work and some earlier artworks whose membership in the order of art is uncontested.

If, for example, the distorted figuration of German Expressionist painting is rejected as art properly so-called on the grounds that it departs from the canons of accurate pictorial representation, the narrativist traces its lineal descent from styles of art, such as that of the medieval artist Matthias Grunewald, where distortion was a strategy for signaling the sentiment of the artist toward his subject. Even if German Expressionist art repudiated prevailing styles of realism, the narrativist argues that there is still reason to count the works in question as art because they harken back to earlier forms of art making, discharging functions, such as the expression of feelings, that are abroad, alive, and acknowledged in their contemporary art world.

One objection to narrativism is that it is circular. However, though circularity is a defect in definitions, it is not clear that it raises any problems for narratives. It is also charged that narrativism confronts the same problem that perplexed the family resemblance approach to identifying art. But this is not the case because narrativists do not merely cite similarities between earlier and later works, but also seek to establish a network of causal relations between them. It is not merely that German Expressionist paintings resemble some medieval art that supports their art status; it is also the case that German Expressionist painting was influenced and inspired by the antecedently recognized medieval art.

Insofar as the narrative approach relies upon tracing lines of descent within historically situated artistic practices, the question arises as to how the narrativist intends to identify artworks in alien traditions. A first response is: by tracing the emergence of later works in that tradition from earlier works. But how can the narrativist identify the first works in alien traditions of art—something he needs to do in order to establish the bona fide origin of subsequent artworks from genuine precedents? Here, the narrativist needs to concede that narrativism is not the only way in which artworks may be identified.

With works in alternative traditions of art making, we frequently need to fix the earliest instances of art in those practices by isolating the works that in that culture are meant to perform the same functions—such as representation, expression, symbolization, decoration, signification, and so forth—that the earliest, already recognized artworks execute in our own culture. This, of course, admits that narration is not the only means of identifying candidates as artworks; sometimes we must depend on functional considerations. Moreover, though historical narration may be sufficient for establishing that a candidate is an artwork, it is not a necessary condition for art status, if only because with certain cases of art, notably from ancient and remote civilizations, it may not be possible to retrieve a narrative account of their provenance.

See also Art, Expression in; Art, Ontology of; Art; Representation in.

Bibliography

Beardsley, Monroe. "An Aesthetic Definition of Art." In *What Is Art?*, edited by Hugh Curtler. New York: Haven, 1983.

Beardsley, Monroe. *Aesthetics: from Classical Greece to The Present*. New York: MacMillan, 1966.

Bell, Clive. *Art*. London: Chatto and Windus, 1914.

Carroll, Noël. *Beyond Aesthetics*. Cambridge, MA: Cambridge University Press, 2001.

Carroll, Noël. "Historical Narratives and The Philosophy of Art." *Journal of Aesthetics and Art Criticism* 51 (1993): 313–326.

Carroll, Noël. *Philosophy of Art: A Contemporary Introduction*. London: Routledge, 1999.

Collingwood, R.G. *Principles of Art*. Oxford: Clarendon Press, 1938.

Danto, Arthur. "Art and Meaning." In *Theories of Art Today*, edited by Noël Carroll. Madison: University of Wisconsin Press, 2000.

Danto, Arthur. *The Transfiguration of The Commonplace*. Cambridge, MA: Harvard University Press, 1981.

Davies, Stephen. *Definitions of Art*. Ithaca, NY: Cornell University Press, 1991.

Dickie, George. *Art and The Aesthetic*. Ithaca, NY: Cornell University Press, 1974.

Dickie, George. *The Art Circle*. New York: Havens Press, 1984.

Gaut, Berys. "Art as a Cluster Concept." *Theories of Art Today*, edited by Noël Carroll. Madison: University of Wisconsin Press, 2000.

Kristeller, Paul Oskar. "The Modern System of The Arts." In *Essays on the History of Aesthetics*, edited by Peter Kivy. Rochester, NY: University of Rochester Press, 1992.

Levinson, Jerrold. "Defining Art Historically." *British Journal of Aesthetics* 19 (1979): 232–250.

Stecker, Robert. *Artworks*. University Park, PA: Pennsylvania University Press, 1997.

Tolstoy, Leo. *What Is Art?*. Translated by Almyer Maude. Indianapolis, In.: Hackett, 1996.

Weitz, Morris. "The Role of Theory in Aesthetics." *Journal of Aesthetics and Art Criticism* 15 (1956): 27–35.

Noël Carroll (2005)

ART, EXPRESSION IN

Art is an expressive business, few would deny, but this assertion has meant quite different things to the large number of thinkers who have contemplated the concept of aesthetic expression over the centuries. Certainly, the fact that art has the power to evoke potent emotions has

been noticed since ancient times. Thus Plato, although perhaps more centrally concerned with the imitative or mimetic dimensions of art, worried famously about the power of poetry and tragedy to subvert the control of reason by the arousal of intense emotions (*Republic* 10.605c, *Ion* 535, *Philebus* 47e–50b). Rather more positively, Aristotle argued that one of the beneficial functions of tragic drama is to provide a catharsis of pity and fear in an audience that is emotionally engaged with tragic personae (*Poetics*, Book VIII).

THE AROUSAL THEORY

The power of art to evoke emotional responses is the basis of the"arousal" theory of expression. The core idea is that an artwork expresses *x* if it has the capacity to arouse a feeling or sensation of *x* in the viewer or listener. Sad music, for example, is music that stirs sadness in the listener. The arousal theory has had many proponents, from Francis Hutcheson (1725) to Colin Radford (1989). The British associationist Archibald Alison, as early as 1790, characterized aesthetic experience in general as the employment of the imagination in the creation of a train of ideas that must be "productive of emotions."

Problems arise immediately for this thesis, however. Some writers with "formalist" inclinations flatly reject it. Eduard Hanslick, for example, in his 1891 work, *On the Musically Beautiful*, denied both that the purpose of music is to arouse emotions and that feelings are in any sense the "content" of music. Moreover, it has often been observed that the reactive emotions of the audience are not always those it is most appropriate to say the work expresses. A tragedy expressive of love, jealousy, and hatred may, as Aristotle said, cause feelings of pity and fear in its viewers. Furthermore, it seems possible to recognize the expressive content of a work without undergoing that very emotion or feeling. A sad or elegant artwork need not make the perceiver sad or elegant.

By contrast, Jerrold Levinson (1990) and Aaron Ridley (1995) have argued that music can arouse a truncated version of the emotions it expresses; the emotions or feelings aroused by music lack their usual contexts and intentional objects. Jenefer Robinson (1994) has pointed out that, although the emotions expressed by music are not always identical with what is aroused in the listener, certain "primitive" emotions can be directly aroused by music expressing those same emotions; music that disturbs us, makes us tense, or calms us down *is* disturbing, tense, or calm. However, music, as an extended composition, also expresses more complex emotions, for example, unrequited passion, which are not aroused in us, but which we attribute as true of the piece partly on the basis of the clues given by the more basic emotions aroused in us.

EXPRESSION AND NINETEENTH-CENTURY IDEALISM

Much grander claims for the expressive power of art were made during the period of German idealism, when art was seen as a manifestation of Spirit. Schelling held that art can show what philosophical concepts cannot: the Absolute, the organic unity of the knower and the known. Schopenhauer called music a copy of the will itself—a direct presentation of the will, expressing the essential nature of emotion types. For Hegel, art provides an irreducible form of self-reflection, conveying knowledge of Spirit through a natural sensuous medium. Along with religion and philosophy, art expresses "the Divine, the deepest interests of mankind, the most all-embracing truths of Spirit" (Hegel 1835–1838, vol. I, p. 21).

In his earlier writings, especially *Die Geburt der Tragödie* (1872, later translated as *The Birth of Tragedy*), Friedrich Nietzsche saw art, especially tragedy and music, as expressing the conjunction or synthesis of two strong human impulses, the "Apollonian," a love of order, measure, and formal beauty, and the "Dionysian," the spirit that glories in a state of elation and joyful acceptance of the excitements and pains of life. Later, Nietzsche allied art more closely with the Dionysian solution to the problem of living, presenting the Dionysian in art as an expression of the basic human drive called the "will to power."

THE EXPRESSION THEORY

Romanticism, with its general emphasis on the emotions and its shift away from classicism, embraced and fostered the view that art is a form of expression in the sense of self-reflection or self-discovery. This theory, labeled by Alan Tormey (1971) the "expression theory of art," is a rival to both high-flown idealism and the arousal theory. According to the expression theory, artworks are expressions of the emotional states experienced by the artist during the creative process. In one variation or another, this view has been endorsed in the nineteenth and twentieth centuries by thinkers such as Eugène Véron, Benedetto Croce, R. G. Collingwood, John Dewey, L. A. Reid, and C. J. Ducasse.

Expression theorists see expressive art as a means of articulating the artist's inner life. In fact, the view can perhaps be thought of as romanticism's alternative to the

arousal theory. Very early in the period, Samuel Coleridge observed that "in *Paradise Lost*—indeed in every one of his poems—it is Milton himself whom you see" (1833, p. 250). A systematic development of expression theory can be found in Véron's influential *L'Esthetique* of 1879, but the view reached its zenith in the early twentieth century in the writings of Italian philosopher Benedetto Croce.

Strongly influenced by Hegelian thought as well as by romanticism, Croce proposed that intuition is a kind of nonconceptual awareness of a mental image, and expression is the forming of "artistic intuitions," which are always infused with intense feeling. Artists express these initially inchoate feelings in the process of forming artistic, or "lyrical," intuitions. Indeed, famously and problematically, Croce identified intuition and expression, and defined art in terms of this mental process. "Intuition is truly such because it expresses intense feeling. … Not idea but intense feeling is what confers upon art the ethereal lightness of the symbol" (1965 [1913], p. 25).

Clearly indebted to Croce, R. G. Collingwood took all art to be an expression of individual and unique emotions, but the process is not the mere exhibiting of the symptoms of the emotion. ("The artist never rants"; 1938, p. 22). Rather, expression is the lucid transformation of sensuous-emotional experience by the artist's imagination into an image or idea. True art, unlike the physical crafts accompanying the various arts, is made in the imagination of the artist.

The idealist tendencies seen in Croce and Collingwood are not shared by all expression theorists, perhaps for good reason. If expression is a purely mental or imaginative process, the artist's manipulation of the medium of his or her art appears to be wrongly undervalued. Although agreeing with Croce and Collingwood that expression always involves the artist's "inner" emotions in need of clarification and transformation, American pragmatist John Dewey emphasized that expression is an "outgoing activity" of interaction with the environment, involving the controlled working of a medium (1934, p. 62). In aesthetic expressiveness we find "meanings and values extracted from prior experiences and funded in such a way that they fuse with the qualities directly presented in a work of art" (p. 98).

Perhaps, then, expression theory can be rescued from the common objection that it makes art and the expressive process overly mentalistic, but it is unclear that it can be saved from another, which charges it with committing the "genetic fallacy" of mistaking judgments about the artist, the source of the art, for judgments about the art itself. The presence of expressive properties in an artwork does not entail the occurrence of prior acts of expression, any more than a cruel expression on a face entails that the owner of the face has acted cruelly.

The expression theory is correctly characterized as a theory of expression emphasizing the emotive processes undergone by the artist, but it would be misleading to think that the arousal of emotions in the viewer or audience is not at least acknowledged by most expression theorists. Dewey remarked, "Because the objects of art are expressive, they communicate. I do not say that communication to others is the intent of an artist. But it is the consequence of his work …" (1934, p. 104). He and Collingwood claim that the emotional reaction of the viewer should mirror or reconstruct the artist's expressive process. When elements of the expression and arousal views are conjoined, the result is a kind of "communication" theory of the sort offered by Leo Tolstoy. In *What Is Art?* (*Chto takoe iskusstvo?*) Tolstoy wrote, "To evoke in oneself a feeling one has experienced, and having evoked it in oneself by means of movements, lines, colors, sounds or forms expressed in works, so to transmit that feeling that others experience the same feeling—this is the activity of art" (1960 [1898], p. 55). For Tolstoy, it is essential to the "sincerity" of the art that the artist feel the emotion communicated, and a condition of "success" of the art that the audience is "infected" with the same feeling.

Of course, a theory conjoining the arousal and expression theses inherits the problems of both views. And it does seem quite possible both that an artist can create a passionate artwork without himself being in a passionate state, and that the audience can recognize that the work is passionate without being made to feel passionate themselves. Composer Richard Strauss said, "I work very coldly, without agitation, without emotion, even" (Osborne 1955, p. 162).

Guy Sircello (1972) champions the romantic view that the mind does not merely mirror or represent nonmental reality but is an original source of some of the features of art, and that it thereby infuses art with intentional or anthropomorphic properties. Although Sircello admits a variety of sources for art's expressive properties, he emphasizes that many of the expressive features that we attribute to artworks are true of them because of the "artistic acts" in which the artist is engaged as he or she creates a work. Pieter Brueghel the Elder's painting *Peasant Wedding Dance* (1566) is ironic, Sircello says, because Brueghel views a happy scene ironically. Nicolas Poussin's *Rape of the Sabine Women* (c. 1635–1637) is aloof and detached, even though the scene

is one of violence, because Poussin observes calmly and paints in a detached fashion.

THE EMBODIMENT THEORY

The "embodiment" theory of expression is a reaction to both the expression and arousal theories, and asserts that expressive properties are rightly said to be possessed by, or true of, the artwork itself either in virtue of its form or composition, or as properties that "emerge" in the work due to broader contextual considerations of a cultural, artistic, interpretational, or psychological sort. Whereas the arousal theory focuses on the effects of expressive art, and whereas the expression theory is a theory of the source of art's expressiveness, the embodiment theory is a cognitivist view of our awareness of the expressive properties that are in, or are possessed by, an artwork. A work can be expressive of *x* even if the artist was not experiencing *x* in creating the work, and the audience does not necessarily feel *x* when they appreciate it.

It is worthy of note that American pragmatist George Santayana, although fitting no category very exactly, is closer to the embodiment theory than to the expression theory with which he is sometimes associated. Santayana wrote quite generally about a sense of expressive beauty and did not focus on the artistic process, nor exclusively on art *per se*. His position may be closer to the earlier British "taste" and associationist theories such as those of Archibald Alison and Joseph Addison: A thought or mental image becomes expressive, according to Santayana, when feelings, meanings, or emotive "tones," proper to some past experience, color and reverberate in our present consciousness, indeed become "incorporated" into it (1988 [1896], pp. 121–124).

Although embodiment theories of various sorts gained currency in the second half of the twentieth century, its most common variant, the "resemblance" thesis, has precursors in the eighteenth century. Johann Mattheson (1739), for example, asserted that by resembling the motion and structure of our vital spirits, music, in its structure, comes to bear a resemblance to the "emotive life," and the primary response of the listener is not to feel emotion but to perceive or recognize the emotive content present in the music. A contemporary version of this position can be found in Peter Kivy's theory of musical expressiveness. In most cases, we perceptually recognize music's expressive properties "in virtue of some perceived analogy" (Kivy 1989, p. 167) to the sound of a person's voice or the movements and gestures made by a person who is literally expressing some emotion. But the reason we animate our musical perceptions, so that we cannot but hear the music as expressive, is, Kivy says, "a divine mystery" (p. 258). Stephen Davies (1994) has a similar view. Like Kivy, he says that music's expressive properties or "emotion characteristics in its appearance" depend mainly on a resemblance that we perceive between the dynamic character of the music and human movement, gait, bearing, or carriage. Both Kivy and Davies also allow that some cases of expression are to be explained by the fact that the musical work engages some wider social conventions surrounding the expression of emotions.

Some resemblance views conclude, on the basis of the resemblance, that an expressive artwork is a symbol of, or signifies, what it expresses. Semiotic theory is then seen as a tool for understanding the nature of expression in art. The best-known signification view based on resemblance is that of Susanne Langer. Art, especially music, is, for Langer, a "presentational symbol" of human feeling. Although feelings are not denoted by such symbols (because such symbols are non-discursive and in this respect unlike language), their form is presented to us in the artwork because there is a logical "isomorphism" between the structure of the work and the "morphology" of the feeling state. Artistic form is congruent with the dynamic forms of our direct sensuous, mental, and emotional life. According to Langer, "music is not the cause or the cure of feelings, but their logical expression." (1942, p. 218).

Other theories have also emphasized the semiotic functions of art in their treatment of expression, but have downplayed the resemblance theme. In his extremely influential book, *Languages of Art* (1968), Nelson Goodman, like Langer, treated artworks as symbols but, unlike Langer, defined expression in terms of the semantic relations of reference and denotation. A work expresses φ if and only if the predicate "φ" metaphorically denotes the work, and the artwork, in turn, "refers back" to that predicate. Less nominalistically stated, expression is a form of property exemplification for Goodman. A works exemplifies a property if it not only possesses but "highlights" that property, much as a tailor's swatch highlights the texture and design of the material because of the conventions surrounding its use. Expression, in this view, is the exemplification of properties that an artwork actually, though metaphorically, possesses. Artworks can express more than human emotions, for example, poised power or flashing action.

Although it is unclear whether Alan Tormey's embodiment theory is committed to the resemblance thesis, he does suggest that the relation between an artwork's nonexpressive and expressive properties is analo-

gous to the relation between human behavior and the intentional states of which the behavior is partially constitutive. Tormey (1971) says that expressive properties are those properties of artworks whose names also designate the intentional states of persons. But, since artworks have no mental states, a work's set of nonexpressive properties is wholly constitutive of its expressive properties. In an interesting though puzzling turn, Tormey claims that expressive ambiguity is ineliminable in art, and therefore expressive properties, though wholly constituted by nonexpressive features, are ambiguously so constituted. Within a certain range of compatibility, there is no objective fact whether an artwork has one or another expressive property; only critical choice leads to a unique judgment as to whether Ravel's *Pavane*, for example, is tender, yearning, or nostalgic. The important question of how one comes to perceive the expressive features of art is left largely unanswered by this view.

OTHER VIEWS

Like all philosophical classifications, those of the arousal, expression, and embodiment theories need to be employed with an awareness of the shortcomings of pigeonholing. A case in point is the work of Richard Wollheim. Influenced by Ludwig Wittgenstein, psychoanalytic theory, and the celebrated work of psychologist E. H. Gombrich concerning the cognitive nature of our perception of art, Wollheim proposes that artistic expression involves "expressive perception," a kind of "seeing or hearing as," by which an artwork, because of how it looks or sounds, causes us to project an emotion or feeling onto that which we see or hear (1987, p. 138). Although the artwork does not simply arouse in us an emotion that we associate with its other features, it does arouse in us the process of projection. And, as in the embodiment theory, the expressive property is ascribed to the work, literally projected onto it, and the work is perceived as possessing it. Lastly, like the expression and communicative views, Wollheim's position suggests that correct expressive perception mirrors or recaptures the emotions that, either through direct experience or through contemplation of them, caused the artist to paint, write, or compose as he or she did.

Finally, a number of writers have introduced an imaginary or fictive element into the discussion of expression, especially regarding music. These theories suggest that artistic expression is often best described in terms of the imaginary occurrence of emotion in oneself or in a fictional persona. Bruce Vermazen (1986) thinks of the expressiveness of a musical passage in terms of an inferred ascription of a state of mind to an imagined utterer of the passage that would best explain the passage's features. Kendall Walton (1990) thinks that expressive music can induce listeners to imagine particular instances of properties expressed, such as instances of someone (perhaps oneself) or something's being exuberant, aggressive, uncertain, or resolved. Walton also claims that sometimes one is induced to imagine of one's own auditory experience that it is an expression of, say, anguish or exuberance (1994).

For Jerrold Levinson, the expressiveness of music derives from its "hearability" as a "sui generis" expression, by an imagined persona, of inner states through outer signs (Levinson 1990, 1996). What a passage of music expresses is what it can most readily and spontaneously be imagined to express by "suitably backgrounded" listeners. That is, music invites listeners to hear it, immediately and directly, as an alternate audible mode of behaviorally manifesting emotions by an imagined persona. Levinson argues against resemblance-based accounts, claiming that recognition of a similarity between music and some emotional behavior is not sufficient for hearing the music as expressive. Similarly, Gregory Karl and Jenefer Robinson (1995) claim that what a musical passage expresses can be the mental state ascribed to the imaginary protagonist of the passage that figures in the best overall interpretation of the work. Whether these "fiction-based" views are types of embodiment theory is somewhat difficult to say with confidence since, rather like expression theories, they emphasize the processes underlying expression in the arts rather than the logic and semantics involved in ascribing expressive properties to works of art.

See also Aesthetics, History of; Aesthetics, Problems of; Aristotle; Coleridge, Samuel Taylor; Collingwood, Robin George; Croce, Benedetto; Dewey, John; Ducasse, Curt John; Goodman, Nelson; Hegel, Georg Wilhelm Friedrich; Hutcheson, Francis; Idealism; Langer, Susanne K.; Music, Philosophy of; Nietzsche, Friedrich; Plato; Romanticism; Santayana, George; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Tolstoy, Lev (Leo) Nikolaevich; Wittgenstein, Ludwig Josef Johann; Wollheim, Richard.

Bibliography

Alison, Archibald. *Essays on the Nature and Principles of Taste* (1790). 6th ed. Edinburgh: A. Constable, 1825.

Aristotle. *Politics* (Book VIII). In *The Basic Works of Aristotle*, edited by Richard McKeon. New York: Random House, 1941.

Coleridge, Samuel T. *The Collected Works of Samuel Taylor Coleridge*. Vol. 14: *Table Talk* (1833). Princeton, NJ: Princeton University Press, 1990.

Collingwood, R. G. *The Principles of Art*. Oxford: Clarendon Press, 1938.

Croce, Benedetto. *Guide to Aesthetics* (*Breviario di estetica*, 1913). Translated by Patrick Romanell. Indianapolis: Bobbs-Merrill, 1965.

Davies, Stephen. *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press, 1994.

Dewey, John. *Art As Experience*. New York: Minton, Balch, 1934.

Ducasse, C. J. *The Philosophy of Art*. New York: Dial, 1929.

Gombrich, E. H. *Art and Illusion*. 2nd ed. London: Phaidon, 1962.

Goodman, Nelson. *Languages of Art*. Indianapolis, IN: Bobbs-Merrill, 1968.

Hanslick, Eduard. *On the Musically Beautiful: A Contribution Towards the Revision of the Aesthetics of Music* (1891). 8th ed. Translated by Geoffrey Payzant. Indianapolis, IN: Hackett, 1986.

Hegel, G. W. F. *Philosophy of Fine Art* (*Äesthetik*, 1835–1838). 2 vols. Translated by T. M. Knox. Oxford: Oxford University Press, 1972.

Hutcheson, Francis. *An Inquiry concerning Beauty, Order, Harmony and Design* (1725). Rev. ed. The Hague: Martinus Nijhoff, 1973.

Karl, Gregory, and Jenefer Robinson. "Shostakovich's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions." *Journal of Aesthetics and Art Criticism* 53 (4) (1995): 401–415. Reprinted in *Music and Meaning*, edited by Jenefer Robinson. Ithaca, NY: Cornell University Press, 1997.

Kivy, Peter. *Sound Sentiment*. Philadelphia: Temple University Press, 1989.

Langer, Susanne. *Philosophy in a New Key*. Cambridge, MA: Harvard University Press, 1942.

Levinson, Jerrold. "Hope in the Hebrides." In *Music, Art, and Metaphysics*, edited by Jerrold Levinson. Ithaca, NY: Cornell University Press, 1990.

Levinson, Jerrold. *The Pleasures of Aesthetics: Philosophical Essays*. Ithaca, NY: Cornell University Press, 1996.

Mattheson, Johann. *Der vollkommene Capellmeister*. Hamburg: C. Herold, 1739.

Nietzsche, Friedrich Wilhelm. *The Birth of Tragedy* (*Die Geburt der Tragödie*, 1872). Translated by W. Kaufmann. New York: Vintage Press, 1966.

Nietzsche, Friedrich Wilhelm. *The Will to Power: An Attempted Transvaluation of all Values* (*Wille zue Macht*, 1901). Translated by Anthony M. Lodovici. London: T. N. Foulis, 1914.

Osborne, Harold. *Aesthetics and Criticism*. London: Routledge and Kegan Paul, 1955.

Plato. *The Collected Dialogues*, edited by Edith Hamilton and Huntington Cairns. Princeton, NJ: Princeton University Press, 1961.

Radford, Colin. "Emotions and Music: A Reply to the Cognitivists." *Journal of Aesthetics and Art Criticism* 47 (1989): 69–76.

Ridley, Aaron. *Music, Value, and the Passions*. Ithaca, NY: Cornell University Press, 1995.

Robinson, Jenefer. "The Expression and Arousal of Emotion in Music." *Journal of Aesthetics and Art Criticism* 59 (1) (1994): 13–22.

Santayana, George. *The Sense of Beauty* (1896). Vol. 2 of *The Works of George Santayana*, edited by William Holzberger and Herman Saatkamp Jr. Cambridge, MA: MIT Press, 1988.

Schelling, Friedrich von. *System of Transcendental Idealism* (*System des transzendentalen Idealismus*, 1800). Translated by Albert Hofstadter in *Philosophies of Art and Beauty*, edited by Albert Hofstadter and Richard Kuhns. Chicago: University of Chicago Press, 1964.

Schopenhauer, Arthur. *The World as Will and Representation* (*Welt als Wille und Vorstellung*, 1818). Translated by E. F. J. Payne. New York: Dover, 1969.

Sircello, Guy. *Mind and Art: An Essay on the Varieties of Expression*. Princeton, NJ: Princeton University Press, 1972.

Tolstoy, Leo. *What Is Art*? (*Chto takoe iskusstvo?*, 1898). Translated by Almyer Maude. Indianapolis, IN: Bobbs-Merrill, 1960.

Tormey, Alan. *The Concept of Expression*. Princeton, NJ: Princeton University Press, 1971.

Vermazen, Bruce. "Expression as Expression." *Pacific Philosophical Quarterly* 67 (1986): 196–223.

Véron, Eugène. *Aesthetics* (*L'esthetique*, 1878). Translated by W. H. Armstrong. Philadelphia: J. B. Lippincott, 1879.

Walton, Kendall. "Listening with Imagination: Is Music Representational?" *Journal of Aesthetics and Art Criticism* 52 (1) (1994): 47–61.

Walton, Kendall. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press, 1990.

Wollheim, Richard. *Art and Its Objects*. New York: Harper and Row, 1968.

Wollheim, Richard. *On Art and the Mind*. Cambridge, MA: Harvard University Press, 1974.

Wollheim, Richard. *Painting as an Art*. Princeton, NJ: Princeton University Press, 1987.

John Bender (1996, 2005)

ART, FORMALISM IN

The term *formalism* refers to a number of theses and programs in the philosophy of art and art criticism, all of which assign a priority to the formal elements of works of art.

The doctrine of formalism exists in a number of versions, not all of them compatible with one another, but in general it is a thesis that insists on the importance—either preeminent or exclusive—of the formal features of works of art in determining the value of those works. As such, it is both a topic for philosophical debate and a prescription for critical practice. This brief essay gives a description of the philosophical background of formal-

ism, an indication of formalist commitments in criticism, and a statement of some logical problems besetting formalism.

PHILOSOPHICAL BACKGROUND

The philosophical basis of formalism is often, and typically, traced to Kant, and indeed Kant is a kind of formalist; but a much earlier formalist doctrine is to be found in Aristotle. A central thesis of Aristotle's *Poetics* is that plot is the most important part of tragedy. Aristotle says a tragedy customarily has six parts (plot, character, thought, diction, spectacle, and melody), and, in declaring plot the most important, he seems to be asserting that excellence of its plot contributes more to the overall excellence of a tragedy than does the excellence of any of its other parts.

Aristotle offers a number of arguments in support of his claim of the preeminent importance of plot. Two are of special interest here. One is the assertion that of all the parts, only plot is necessary to something's being a tragedy. The other is the claim that plot has more of a bearing than the other parts of a tragedy on the work's special and proper effect, namely the production of *catharsis*. Thus, although Aristotle himself does not speak in these terms, his arguments are close to a claim that plot is both a necessary and a sufficient condition of tragedy, and his thesis is a kind of essentialism. What makes this essentialism a formalism is Aristotle's conception of plot: a plot, he says, is the "arrangement of incidents." Although Aristotle sometimes uses the term *plot* in something like the modern sense, meaning roughly the "story," the more abstract conception (arrangement of incidents) suggests a structure—a formal entity. And indeed Aristotle identifies plot as the "formal cause" of a tragedy.

There have been attempts to generalize Aristotle's theory. The theory is offered by Aristotle specifically with reference to tragedy, and the obvious question is how to apply it to any other artistic form. Some interpreters have thought that Aristotle would regard the plot as the most important part of any artwork that has a plot, including, for example, an opera or ballet. But it might be a mistake to regard the plot as the most important element of, say, an opera. What an Aristotelian should be looking for is the necessary and sufficient condition of something's being an opera—opera's formal cause—and this may well be its music, as Joseph Kerman has argued in *Opera As Drama*. The incidents whose arrangement is vital will be musical incidents.

Whereas for Aristotle the centrality of form is a metaphysical or ontological matter, having to do with the nature of the objects themselves, for Kant the importance of form is grounded in a quasi-epistemological conviction. A Kantian judgment of taste requires exclusive attention to form because nothing else can underwrite such a judgment's claim to universality. Kant's reasons for thinking this are relatively clear, even if his argument is difficult to formulate.

According to Kant, a judgment of something's beauty is based on the judge's feeling of pleasure in the thing. It is distinguished from other so-called "aesthetic" judgments by its implicit claim to an intersubjective validity. The judgment is thus not parochial because it is in part to some extent a rational judgment, requiring the use of the faculty of concepts. In the exercise of such judgment, according to Kant, attention is restricted to the form of the object. The judge is entitled to suppose that any other judge would also experience pleasure in the object if he judged in the same way—taking pleasure in his contemplation of the mere form of the object. Why does Kant think that everyone judging in this way will experience pleasure? In answering this question, Kant seems to rely on what he claims to have proved in the *Critique of Pure Reason*—namely that states of mind are communicable because unless they were, objective knowledge of the world would not be possible, and he thinks that he demonstrated that such knowledge is possible.

The definition of form is much less clear in Kant than in Aristotle. Kant seems to be thinking of what we might roughly think of as shape, and that seems a reasonable way to understand one of Kant's leading examples, namely the judgment of the beauty of a rose. But it leaves it utterly unclear why Kant has such a low opinion of music, given the entirely plausible conviction that music may well display abstract form more conspicuously and typically than does any other art.

FORMALISM IN THE VARIOUS ARTS

In any art, formalism concentrates on the formal elements in the works it deals with. It is not always clear just which elements are formal, in these theories, and it is not uncommonly clearer which elements do not count as formal than it is how the formal elements are defined.

VISUAL ARTS. In the visual arts, formalism has insisted on a concentration on line and shape. Its early proponents were Clive Bell and Roger Fry, and perhaps its most conspicuous twentieth-century advocate was Clement Greenberg. In its more extreme formulations, formalism in the visual arts has insisted that the value in, say, a painting, is unrelated to its representational features and

is due entirely to the its form, where that form is understood entirely as a generally abstract structure constituted by the lines, shape, and, perhaps, color of the painting.

MUSIC. Formalist theory and criticism of music almost always explicitly refuses to give attention to any "program" associated with the music or even to the sung text in vocal music. Formalism does not always refuse attention to the emotions that may be evoked by music, but it insists that these feelings arise from "music alone" and not from any representational or narrative features, no matter how closely these may be associated with the music. An early statement of this view is given by Eduard Hanslick, and recently one of its most sophisticated exponents has been Peter Kivy.

LITERATURE. Formalist literary theory is somewhat harder to describe than is formalism in the other arts. If formalism, in general, is thought to be a doctrine in which principal or exclusive attention is to be paid to the perceptual elements of a work and to the relations between these elements, then it would seem to require that literary formalists attend only to the shapes and sounds of words, and this requirement is surely incredible. Thus formalism in literature has to be understood more subtly. It is commonly taken to require attention exclusively to "the work itself," where this seems to mean eschewing references to considerations coming from "outside" the work. In particular, formalists have wished to deflect historical, biographical, and psychoanalytical interests, although, of course, even the most severe formalism may have to countenance some historical interests in so far as these are necessary to establish certain features of the work—for instance, the meanings of various words or the references of proper nouns. Furthermore, there have been different species of formalism because of different opinions about which formal features are most important.

PROBLEMS FOR FORMALISM

With it professed interest in works of art themselves, and not to any ancillary features, it is fair to say, with some qualification, that formalism does *not* want attention to representational or narrative features, or to any emotional evocations that result from those things. There are two main problems facing any advocate of formalism. One is to supply some argument in favor of the claim that a work's formal properties are either the only or the most important of its elements; but before that, there is a need to offer some criterion that distinguishes formal from nonformal elements. The latter problem may be more bothersome than it first appears, especially when one asks what formalists mean by *formal*. A useful way of doing this is to ask, "Formal as opposed to what?" When that question is raised, quite different answers are given for various arts. Thus, some procedure or routine must be given that will answer, for any true statement about a work of art *A*, with the form *A* is *F*, whether the property *F* ascribed to *A* is a formal property. This is very difficult to do, and that difficulty often leads to something of a reduced insistence—namely that it be determined, given that *F* is a property of *A*, whether *F* is an *essential* property of *A*. This formulation tends to be more or less agreeable depending upon how favorably one looks at philosophical essentialism.

Supposing it is settled how to tell whether a property is a formal property; the formalist now needs an argument for dealing with this issue: Given that *A* has the property *F*, and also the property *N*, and that *F* is a formal property, whereas *N* is not a formal property, why is *F* a more important property of *A* than *N*, more critical to assessing *A*'s value or importance? Even if it were true that *F* is an essential property, how does it follow that *N* is less important?

Whatever its defensibility as a philosophical thesis, and however vaguely it has to be stated, formalism retains one merit: it has recommended and insisted upon attention to those features of an art work that incontestably are features of the work itself—features often scanted in the assessments of antiformalists.

See also Aesthetic Qualities; Art, Definitions of; Pater, Walter Horatio; Wilde, Oscar Fingal O'Flahertie Wills.

Bibliography

Alperson, Philip. "The Philosophy of Music: Formalism and Beyond." In *The Blackwell Guide to Aesthetics*, edited by Peter Kivy. Oxford: Blackwell, 2004.

Aristotle. *Poetics*. Translated by James Hutton. New York: Norton, 1982.

Batkin, Norton. "Formalism in Analytic Aesthetics." In *Encyclopedia of Aesthetics*, edited by Michael Kelly. New York and Oxford: Oxford University Press, 1998.

Bell, Clive. *Art*. London: Chatto & Windus, 1916.

Davis, Whitney. "Formalism in Art History." In *Encyclopedia of Aesthetics*, edited by Michael Kelly. New York and Oxford: Oxford University Press, 1998.

Greenberg, Clement. *Art and Culture: Critical Essays*. Boston: Beacon Press, 1961).

Isenberg, Arnold. "Formalism." In *Aesthetics and the Theory of Criticism: Selected Essays of Arnold Isenberg*, edited by William Callaghan, et al. Chicago: University of Chicago Press, 1973.

Kant, Immanuel. *Critique of the Powqer of Judgment*. A translation of Kant's *Kritik der Urteilskraft* by Paul Guyer and Eric Matthews. Cambridge, U.K.: Cambridge University Press, 2000.

Kivy, Peter. *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca, Cornell University Press, 1990.

Melville, Stephen. "Greenberg, Clement." In *Encyclopedia of Aesthetics*, edited by Michael Kelly. New York and Oxford: Oxford University Press, 1998.

Strier, Richard. "How Formalism Became a Dirty Word, and Why We Can't Do Without It." In *Renaissance Literature and its Formal Engagement*, edited by Mark David Rasmussen. New York: Palgrave, 2002.

Wollheim, Richard. "On Formalism and Its Kinds." In *On Formalism and Its Kinds/Sobre el formalism I els seus tipus*, edited by Fundacio Antoni Tapies. Barcelona, publisher unknown, 1995.

Ted Cohen (2005)

ART, INTERPRETATION OF

The concept of interpretation is key to our commerce with artworks. For if something is an artwork, then it falls into the category of things that are at least eligible for an interpretation. For example, all things being equal, an ordinary snow shovel is not a candidate for interpretation, but Marcel Duchamp's *In Advance of a Broken Arm* is, despite the fact that it is indiscernible from the other snow shovels produced at the same time, in the same factory.

However, not all the elements or combinations of elements in an artwork merit interpretation. Only those elements or combinations thereof are worthy of interpretation, which somehow mystify, perplex, or elude. The appropriate object of interpretation is that which goes beyond what is given or foregrounded (Barnes 1988).

An interpretation is a hypothesis that accounts for the presence of an element or combination of elements in an artwork where the presence of the relevant elements is not immediately obvious to the interpreter and/or to some target audience. The item may not be obvious in the sense of being unintelligible or enigmatic, or because it is symbolic or allegorical, or because it is understated, barely hinted at, only suggested, or it is in some other way recessive.

The purpose of an interpretation is to enhance our understanding of an artwork. There is something about the artwork that is obscure, ambiguous, apparently incoherent, anomalous, unexpected, inaccessible, perplexing, or latent that invites illumination. The aim of an interpretation is to elucidate the presence of the pertinent elements in the artwork by explaining the contribution they make to the unity, meaning, design, intended effect, and/or structure of the work. Consequently, the work of interpretation presupposes some target audience—to which the interpreter may or may not belong—for whom the significance of some part of the work, or even the artwork as a whole, is elusive, puzzling, obscure, nonmanifest, unfocused, symbolic, or otherwise not immediately apprehensible. The interpretation, then, ideally alleviates that perplexity or gap in the audience's understanding.

Not every element in an artwork calls for an interpretation. Where with respect to a painting such as El Greco's *The Adoration of the Shepherds*, everyone recognizes the subject to be a woman, a child, and two men, then the observation that "this painting represents a woman, a child and two men" is not an interpretation, but a description. Descriptions are nevertheless relevant to interpretations, since sound interpretations must rest upon accurate descriptions.

The literal meaning of many of the words and sentences in literary works are grasped by means of subpersonal routines of processing by literate readers in the language in which the work has been composed (Currie 2004). The literal meaning of the opening line of Kafka's *The Castle*—"It was late in the evening when K. arrived"—does not require an interpretation, insofar as it is obvious to the prepared reader. What might require an interpretation, on the other hand, is its place in the broader design of the novel. Interpretation only pertains to that which is not apparent to some audience. Thus, what is suggested, entailed, or implicated is grist for the interpreter's mill, though not what is spoken outright (although why an author chooses to speak directly rather than obliquely, in certain circumstances, may be a legitimate interpretive question).

That, in a movie, shots of waves pounding on the beach often symbolize intercourse when juxtaposed to shots of lovers may be obvious to the jaded film critic; however, making note of this cinematic figure counts as an interpretation, since there is a target audience for whom it is news. Likewise, a reading of the symbolism of the death's head in a *vanitas* painting is an interpretation, since most people, untutored in art history, are unaware of the association between it and the concept of mortality.

Interpretation is, in general, a holistic enterprise. It strives to isolate the point(s) or purpose(s) of an artwork in order to explain the ways in which the parts cohere or segue with the aims of the whole as contributions to the function and/or meaning of the artwork. The predomi-

nant tendency of interpretation is to show a work to be more and more unified in intent. Of course, in order to build up a conception of the whole, the interpreter must begin with the parts, conjecturing and then adjusting his hypotheses regarding their significance as they arrive before him. The interpreter moves from hypotheses about the part to hypotheses about the whole and then back to the part again. This is sometimes referred to as the *hermeneutic circle* (Gadamer 1975); it underscores the fact that interpretation is a continuous process of reflective equilibrium involving an iterative feedback loop from part to whole and then from whole to part.

The overall direction of interpretation is toward establishing the unity of intent, thought, or design in the artwork. Even an avant-garde work, like Luis Buñuel's *L'age d'or*, which is predicated upon insistently subverting our expectations by a series of what appear to be narrative *non sequiturs*, can be shown by an interpretation to exhibit a sort of second-order unity in virtue of its consistent choice for surrealist purposes of incoherent sequences of events. On the other hand, interpretation can also have a role to play in revealing the disunity in a work. After identifying the intended effect of a novel to provoke a sense of mystery in the audience, the interpreter may then go on to point out that that purpose was ill served by the ineptly transparent way in which the murderer was crudely marked as guilty from his first appearance onwards. Because of its overriding concern with the unity of the artwork, interpretation is intimately related to evaluation, often supplying premises for our judgments of the quality of artworks.

Since interpretation is so involved with exhibiting the unity of artworks, it is often connected to the discovery of meaning, especially in works of narrative, dramatic, and symbolic import. For meaning—in the sense of a theme, a thesis, or an overriding concept—is one of the most frequent ways in which such works may be unified. The theme of the inhumanity of war, for instance, governs *All Quiet on the Western Front*. The interpreter, contemplating the parts of the work, for example its various episodes, hypothesizes this theme and then goes on to show how this concept colligates or unifies Remarque's choice of the incidents he presents to the reader. That is, an interpretation like this isolates the principle of selection—in this case, a concept—that makes a coherent package of the collection of details assembled in the novel.

ANTI-INTENTIONALISM

Meaning of various sorts is so frequently associated with interpretation that many philosophers identify the excavation of meaning as the sole object of interpretation and, for that reason, propose linguistic meaning as the model for understanding interpretation. Linguistic meaning, of course, is highly structured in terms of conventions of semantics and syntax. So on this view, interpreting a work is a matter of discovering its meaning through the rules of the relevant art form. With respect to a poem, for example, it is said, one need only appeal to the public meanings of the words and the traditional practices of figuration; no recourse, for example, to authorial intention is necessary. Because of its reliance upon the conventional meanings of words to the exclusion of authorial intention, this view, which was ably defended by the late Monroe Beardsley, can be called anti-intentionalism.

To the extent that anti-intentionalism depends upon our understanding of linguistic meaning in terms of conventions as a model for the interpretation of works, it cannot, at the very least, be generalized across the arts. For most of the arts do not possess the highly structured meaning conventions that language does. The fact that a stage director chooses to incorporate a swimming pool into the set of her theatrical production of *A Midsummer Night's Dream* is certainly a decision worth pondering in an interpretation of the performance ("What might the director be symbolizing by this?"); but there is no fixed, public meaning attached to the appearance of swimming pools onstage.

And yet even with respect to the literary arts, many of the traditional objects of interpretation are inhospitable to the linguistic model. For example, interpreters often focus upon the significance of plot ellipses or they question why a character possesses a certain set of apparently conflicting attributes. But neither of these recurring objects of interpretation can be referred to pre-existing codes or conventions of decipherment.

Furthermore, literary works often mobilize irony and allusion. The conventions of language will be of no avail with radical cases of irony, since in these instances the author means to say exactly the opposite of what the rules of language entail, while there are no conventions to tell the difference between allusions, properly so called, and coincidental similarities of phrasing. Indeed, even in the case of metaphor, we have no laws to tell us how to proceed in unraveling them interpretively. So it is even controversial whether the anti-intentionalist or conventionalist stance can serve as a comprehensive account of

the arts of language which, on the face of it, would appear to be its most welcoming field of application.

Perhaps an even deeper problem with the linguistic-model version of the conventionalist or anti-intentionalist stance is that it presumes that the object of interpretation is always something construable as a meaning—that is, either as a proposition, an utterance, or a concept. But often the object of interpretation is what the artist has done rather that what he has "said." For example, the art historian may explain to her class that the artist has placed the crucified Christ at the vanishing point of his painting in order to emphasize that it is Christ's death that is the subject of the painting and not, for instance, the Roman soldiers playing dice at the side of the cross. This is a rhetorical or dramaturgical effect that, inasmuch as it may not be apparent to many viewers until it is pointed out, is worthy of interpretive attention. However, it does not involve meaning, linguistically construed. It does not say, "look here"; rather it has the effect of tending to draw the eye of the normal viewer in that direction. Yet, explaining the function of this device in the design of the work as a whole is interpretative because it contributes to disclosing the unity of intent of the work—in effect, to explaining the way in which this strategy reinforces the plan, point, or purpose of the painting.

The limitations of the conventionalist model may encourage us to look elsewhere for a way of understanding interpretation. Moreover, we need not search far afield. For interpretation is not some strange phenomenon that we engage only with respect to rarefied objects like art objects; ordinary human life is shot through with interpretation.

INTENTIONALISM

Barely an hour goes by when most of us are not involved in interpreting the words and deeds, the sayings and doings of our conspecifics. The ability to read the minds of others is an indispensable part of social existence, and those who are extremely deficient at it, such as persons stricken by autism, are typically thought to be disabled. The interpretation of artworks appears simply to be a specialized extension of this natural capacity of the human frame, no different in kind than our interpretation of the behavior, verbal and otherwise, of the family, friends, strangers, and enemies who surround us daily.

Thus, our ordinary practices of interpretation may be expected to shed some light on the interpretation of artworks. In everyday life, interpretation is typically aimed at understanding the intentions of others. We scrutinize the speech and the behavior, often nonverbal, of conspecifics in order to make sense of it by inferring the intentions that gave rise to it. If the behavior takes place against the background of conventions, as speech does, we factor those conventions into our deliberations. However, arriving at our interpretation of an action, including a speech act, rarely involves applying conventional rules to behavior mechanically. We appeal to what we know about the agent, about her beliefs and her desires, about the context of her activity as well as what we know about pertinent conventions to arrive at our interpretations. Why not approach the interpretation of artworks in the same way that we interpret our conspecifics every day? Isn't it very likely that the interpretation of artworks is on a continuum with the interpretative propensities that appear to have been endowed innately by natural selection as a beneficial adaptation for social beings like ourselves?

If it is plausible to answer these questions affirmatively, then the narrow compass of linguistic meaning emphasized by the anti-intentionally disposed conventionalist may be exchanged for the broader notion of sense that is invoked when we speak of making sense of an action—where what makes sense or what renders an action comprehensible is the identification of the coherent intention that lies behind it. Why not suppose that making sense of an artwork is of a piece with making sense of an action? One advantage of this view, in contradistinction to the previous version of anti-intentionalism, is that art forms that are not governed by rules as strict as those of semantics and syntax are still readily interpretable under an intentionalist understanding of interpretation such as this one.

Artworks have a communicative dimension. Consequently, all things being equal, we should try engage them as we do the other communicative behaviors of our fellow humans—as sources of information regarding their intentions. Where interpretation comes into play, its point is arguably to discern the communicative intentions of the creator of the work. An interpretation is successful to the degree that it tracks the intentions of artists. This view, for obvious reasons, we may call intentionalism.

Intentionalism is often rejected because it is thought to force its proponents to the nonsensical position that the preferred interpretation of an artwork is that it has whatever meaning or function its creator says it does. So if a poet says the word "blue" in his poem means "red," then "blue" means "red." But this is absurd. Of course, in a case like this, we may suspect the poet is dissembling about what he truly intends. In the ordinary course of

affairs, we do not allow our interlocutors the last word on their intentions. So it needs to be emphasized that intentionalism is not committed to the view that an artwork means whatever an author merely says it does. Rather, intentionalism is after the actual intention of the artist.

But let us imagine that in this case, we somehow are able to ascertain that the poet really does intend "blue" to mean "red." Surely, we will not accept that this is what the word means, and, moreover, the anti-intentionalist can say why—because it violates the rules of language.

This objection is fatal to the most radical variety of actual intentionalism (Knapp and Michaels 1982). However, there may be more modest forms of actual intentionalism that are capable of dodging this objection. One strategy in this respect is to regard the intentions of the creators of artworks as pertinent to the interpretation of artworks just in case the work itself—including, in this instance, the words and their conventional meanings—can support the putative intention of the artist (Hirsch 1967, Iseminger 1996, Carroll 1999). Where they cannot, isolating the artist's intention will not, the intentionalist concedes, promise a successful interpretation of the work. In this way, the modest actual intentionalist acknowledges the role of both conventional meaning as well as intention in interpretations (Stecker 2003).

Nevertheless, the modest actual intentionalist must surmount further challenges. One charge is that this approach misdirects the interpreter. Instead of focusing on the work, the interpreter is focused on something outside the work, in effect the artist's intention. However, the modest intentionalist notes that since the artwork is the primary source for our evidence about the artist's intention, intentionalism does not beckon us to turn away from the artwork, but to inspect it more closely. Furthermore, the intentionalist contends that it is not quite right to maintain that our interest is in the artwork as if it were an object in nature. Surely, since so many of the critical remarks we lavish on artworks presuppose the notion of *achievement*, our interest in the artwork is in the way intentions are realized in the work. But to appreciate that requires a grasp of the intentions that gave rise to the work.

The intentionalist argues that the interpretation of artworks is on a continuum with our everyday interpretation of our conspecifics. However, critics of intentionalism maintain that once we enter the realm of art, things change. Even if standardly we interpret in order to identify the intentions behind the words and deeds of others, art is not like that. It has purposes above and beyond the practical concern with gathering information from our conspecifics. An essential function of art is to afford aesthetic experience—experience valued for its own sake—by encouraging the imagination of the reader, listener, or viewer of the artwork in lively interpretive play. The claim that the proper aim of interpretation is to attempt to identify the intention of the artist may conflict with this putatively central function of art. Thus, in order to engage artworks appropriately, our normal inclination towards interpreting for intention should be suspended.

On the one hand, the view that a central function of art, one that trumps all the others, is to engender aesthetic experience by abetting the imaginative play of interpretation is, to say the least, controversial. Nor can it be bolstered, without begging the question, by suggesting that the authority of this viewpoint is manifest in the behavior of informed participants in the art world, since one finds that informed participants in the art world indulge in intentionalist interpretations with remarkable frequency.

On the other hand, it is difficult to gainsay that an artwork has at least a communicative dimension—that it is meant as the expression of a thought or a feeling or as a projection of a design for contemplation, or is meant to have some other intersubjectively detectable effect. Moreover, it may be argued, that once we enter a communicative relationship with another, including the creator of an artwork, then it would appear that we are bound by certain moral responsibilities.

That is, we must treat the communiqué of the other fairly, with charity, and with accuracy; we must engage our interlocutor justly and attempt to get at what she intends to communicate. Perhaps the best evidence for this moral commitment is the injustice we ourselves feel when we believe that others are "putting words in our mouths."

But if such moral considerations are germane to interpretation, then it does not seem that the supposed pursuit of aesthetic experience through the free, or, at least intentionalistically independent, play of interpretations trumps all of our other legitimate interests in artworks. Rather the range of acceptable interpretations will be morally constrained by our best hypotheses about what the creator of the artwork intended (Carroll 1991).

HYPOTHETICAL INTENTIONALISTS

Nevertheless, even if it is conceded that the work of interpretation aims at hypothesizing the intention of the creator of the artwork, there is a dispute among intentionalists over what should count as its preferred

interpretation. One side—call them hypothetical intentionalists—claims that the preferred interpretation of the artwork is the one that would be conjectured by an idealized, fully informed audience member, availing herself of all the publicly accessible information surrounding the artwork (including knowledge about the rest of the creator's oeuvre, about the history and practice of the pertinent genre and style of the artwork, about the social context of the work, and even concerning whatever is in the public record of the author's life) (Levinson 1996). The other half of this debate—call them modest actual intentionalists—maintains that the preferred interpretation of the work is whatever the actual intention of the creator was so long as that is supported by the work itself.

Since both hypothetical and actual intentionalists will usually rely upon the same kinds of considerations to arrive at their interpretations—historical context, art history, the rest of the creator's oeuvre, and so forth—in practice the two positions are apt to converge generally on the same interpretations of the work. There is a point at which they clash, however. Since the goal of the modest actual intentionalist is the retrieval of the actual intention of the creator, she is willing to help herself to information—wheresoever it comes from—about what the author really intended, so long as what the creator is thought to intend is consistent with his creation. This includes being prepared to use clues from the private diaries, letters and notes of the creator as well as the reliable testimony of friends of the creator. In contrast, the hypothetical intentionalist believes that the interpreter must be limited in her hypotheses to just what can be found in the public record.

The hypothetical intentionalist defends his viewpoint, in part, by asserting that the aforesaid limitations on the kinds of evidence to which an interpreter has a genuine right are part and parcel of the principles underwriting art world practice. It is a violation of the rules of the game, in other words, to use the private papers of an artist to formulate the preferred interpretation However, it is not clear where the hypothetical intentionalist locates the basis of this alleged rule. It cannot be observed in the actual practice of interpretation, since many critics appear quite happy to use unpublished biographical confidences in their work. Perhaps they are in some violation of some rule, but, since the eclipse of the New Criticism, no one appears to call them on it anymore. Moreover, the notion that such a rule could govern the art world seems unlikely. For when we become interested in an artist and his artworks, we are happy to learn everything we can about him and to incorporate it into our understanding, irrespective of from whence that information originates.

READER-RESPONSE THEORY

Because interpretation is so often involved with the identification of meaning, it is quite natural to suppose that it is connected to intentions. For, the meaning of an utterance—such as "The door is closed"—depends upon whether the speaker intends to be reporting a fact or asking a question (signaled, perhaps, by changing one's intonation at the end of the sentence). However, while agreeing that the meaning of an utterance requires an intention, some may question whether the pertinent intention needs to be that of the author or creator of the artwork. Might not the intention be supplied by the consumers of the work—the readers of the poem, for example?

On this view, which is a variant of reception theory or reader-response aesthetics (Tompkins 1980), the author of the poem supplies his readership with a text—a mere sequence of words whose meanings are to be imputed by the audience, albeit usually within the constraints of the possible dictionary senses of the relevant words and the rules of grammar. In this way, each reader may be thought to construct her own artwork, much as the interpretation of a score by a musician counts as a work of performing art in its own right. That is, in the inevitable process of filling-in the indeterminacies of the text (a sheer sequence of symbols sans fully determinate meaning), the reader putatively creates her own artwork.

Even if this view of interpretation suits some art forms, like literature, it is difficult to generalize across the arts. How exactly would it apply to architecture? It strains language violently to say that each spectator *constructs* his own building, and where, in any event, would those buildings be situated exactly? There would appear to be room for only one Notre Dame cathedral on its present site in Paris; or, Are all those imputed cathedrals immaterial? Surely, such thinking leads to a strange form of architecture.

Another problem with this way of talking is that it would seem to evaporate the relevant category of interpretation entirely. In ordinary language, we countenance at least two notions of interpretation—the notion of a critical interpretation (which has been the topic of this entry) and what might be called a performative interpretation—the sort of interpretation that a musician gives to a piece of music or that an actor gives to a role. These two kinds of interpretations may be related—the actor may produce or consult a critical interpretation of a play

before creating his role through an interpretation/performance. But the two sorts of interpretation are usually thought to be distinct.

However, on the variation of reception aesthetics under discussion, the difference disappears. There is no artwork to be interpreted critically because the interpretation—the performative interpretation—by the reader just is the artwork. There is no conceptual space left over for the critical interpretation to inhabit. Or, in other words, the distinction between the artwork and its (critical) interpretation has disappeared.

Furthermore, if each interpretation, in the sense germane to the reception theorist, amounts to a different artwork, then it is not clear how we will go about comparing different interpretations. What will be the reference point in such comparisons? But we do compare interpretations. Consequently, a theory that makes this impossible is suspicious.

And finally, if audiences create artworks, what is it precisely that artists do? Is it that short-story writers produce texts—strings of symbols without intended meanings? This surely is not what writers think they do, nor does it seem humanly feasible for an author to produce a document on such a scale with no definite utterance meanings in mind. And how would we go about evaluating works constructed on this construal? Would the "text" that generated the most (or the least) reader-response artworks be the best and why? Or, would there be some other criteria.

At the very least, the reception-theory version of interpretation canvassed so far would call for a dramatic overhaul in the way in which we talk and think about art. Before embracing such a view of interpretation, we should require a fuller account of that alternative conceptual framework than any developed so far. On the other hand, it may be an added virtue of modest actual intentionalism that it fits our current interpretive practices as neatly as it does.

See also Hermeneutics; Literature, Philosophy of; Structuralism and Post-structuralism.

Bibliography

Barnes, Annette. *On Interpretation*. Oxford: Blackwell, 1988.

Beardsley, Monroe. *Aesthetics*. Indianapolis, IN: Hackett, 1981.

Carroll, Noël. "Art, Intention and Conversation." In *Intention and Interpretation*, edited by Gary Iseminger. Philadelphia: Temple University Press, 1991.

Carroll, Noël, "Interpretation and Intention: The Debate between Actual and Hypothetical Intentionalism." In *The Philosophy of Interpretation*, edited by Joseph Margolis and Tom Rockmore. Oxford: Blackwell, 1999.

Currie, Gregory. "Interpretation and Pragmatics." In *Arts and Minds*. Oxford: Oxford University Press, 2004.

Gadamer, Hans-Georg. *Truth and Method*. New York: Seabury Press, 1975.

Hirsch, E. D. *Validity in Interpretation*. New Haven, CT: Yale University Press, 1967.

Iseminger, Gary. "Actual Intentionalism versus Hypothetical Intentionalism." In *The Journal of Aesthetics and Art Criticism* 54 (1996): 319–326.

Knapp, Steven, and Walter Benn Michaels. "Against Theory." *Critical Inquiry* 8 (1982): 723–742.

Levinson, Jerrold. "Intention and Interpretation in Literature." In *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996.

Stecker, Robert. *Interpretation and Construction*. Oxford: Blackwell, 2003.

Tompkins, Jane, ed. *Reader-Response Criticism: From Formalism to Post-Structuralism*. Baltimore, MD: Johns Hopkins University Press, 1980.

Noël Carroll (2005)

ART, ONTOLOGY OF

Ontology is concerned with what exists. So one may think the ontology of art is concerned with whether artworks exist. However, most people take the existence of artworks for granted. (See Dilworth 2004 for someone who does not.) The main issue for the ontology of art is what kind or kinds of objects artworks are. A second important issue is about the identity and individuation of works. Concerning both of these issues there is wide disagreement along a variety of parameters.

OBJECTS THAT ARE ARTWORKS

ONE KIND OR MANY. One parameter along which there is disagreement is whether all artworks belong to a single kind or whether they belong to irreducibly different kinds. The second view seems more plausible, at least initially. A painting, such as one made with oils or watercolors, is an entity that has physical properties, such as spatial dimensions, that exists in a single place at a single time, and, for these reasons, may be plausibly taken for a physical object. A novel could be said to exist in many places—wherever there is a copy—or in no place, because no copy or even the original manuscript is the novel. For this reason, novels could not be physical objects. One of a kind sculptures are more like paintings in the respects mentioned above, whereas many musical works are more like novels.

Nevertheless, there are a variety of attempts to argue that artworks belong to a single kind. One strategy for doing this is to argue that all works are types or kinds of some sort, thereby assimilating those, such as paintings, that appear to be physical objects to the category to which novels and musical works more obviously belong. One proposal is that all artworks are structural types. Musical works are sound-structures and literary works are linguistic structures (or possibly, in some instances, plot structures). Paintings also have a structure that could be defined in terms of patterns of colors and shapes, or defined in some other way. This structure is duplicated in a copy of the painting, perfectly duplicated in a perfect copy. Prints and sculpture produced from a model seem to fit this proposal better than paintings, because people currently recognize that prints and sculpture that share a common structure belong to a single work. Current practice does not do this for paintings, no matter how perfect the copy. One can imagine a future time when painters produce a work in two stages. First they paint something. Second, they authorize a certain number of mechanically produced copies to be housed in several different museums or galleries as instances of the work, just as there are now several authorized instances of Henry Moore's sculpture *King and Queen* on different sites in different parts of the world. However, the possibility of imagining this new practice does not show that paintings are really abstract structures. If anything it shows the opposite, because the imaginary practice stands in stark contrast with the actual one. This actual practice does not recognize copies as instances of the work, asserts that the work is deteriorating when the paint applied by artist to canvas deteriorates despite the existence of good copies, and so on. Because painting and some sculptures are not structural types, there are other works that also do not fit the proposal, even though they are not physical objects. Improvisations and happenings are nonrepeatable events. So the strategy of arguing that all artworks are abstract structural types fails.

These considerations also speak against a second proposal: that artworks are action-types (Currie 1988). The type is the discovering of an abstract structure in a specific way (a "heuristic path"). The proposal recognizes a consideration that is discussed at length below: a work's pattern or structure and the context of its making are distinct sources of important artistic properties. However, if the reasoning of the last paragraph is correct, the present proposal has a defect similar to the first proposal in misidentifying the sort of objects that paintings and uncast sculptures are. These are not types of achievement; rather, they are specific concrete objects that are appreciated only in part for what they achieve. Even genuinely abstract works seem to be objects brought about by a type of activity rather than that action-type itself.

A different strategy for arguing that all artworks belong to a single ontological category is to argue that they are all concrete objects of some sort rather than abstract objects. One proposal on the table that fits this approach claims that all artworks are action *tokens*, in particular, the creative activity of artists that bring into existence those objects normally thought of as works of art (D. Davies 2003). On this proposal, the actual work is uniformly the creative activity, the product of that activity being dubbed the work-vehicle and distinguished from the work itself. One may wonder what this accomplishes other than a renaming. Both the creative act and the object are recognized by everyone, and as such, no novel entity is involved in the act-token conception of artworks. So why depart from normal practice and assert that the artwork is not the object produced by the artist's activity, but is the activity itself? Simplifying a complicated argument, the main reason is the importance for art appreciation of reconstructing the artist's creative activity. The claim is that the only way to acknowledge this importance fully is to identify the work with the activity. This claim is unjustified. An object has many relational properties in virtue of its origin and recognizing these properties may be crucial to fully appreciating the object as an artwork. People can accord recognition to the artist's creative activity by understanding that it is in virtue of the creative act, and of the project that gives rise to it—that the work (object) has the relational properties crucial to appreciating it. There is no need to identify the work with the creative activity itself. Hence the renaming that the act-token view proposes is neither necessary nor desirable.

One may conclude that the heterodox view that artworks belong to irreducibly different kinds is not only more plausible initially, but more plausible after reflection as well. Taking this for granted, the next question is how to more sharply define these kinds.

THE ROLE OF INTENTIONS AND CONTEXT. According to the heterodox view, some artworks are abstract types or kinds, others are concrete objects with physical properties, and there are still others that are particular events or processes. One individuating feature of abstract artworks, such as novels, plays, and pieces of music, is that instances of each work share a common structure. Is this sharing of a structure sufficient to individuate a single work? It is clear that this is not always so. Consider the

case of a sculpture that has multiple instances. The structural element here consists of a material such as bronze being shaped in a certain way. Wherever there is a piece of bronze so shaped there is an object that has a structure in common with the sculpture. But this is clearly not sufficient for the object to be an instance of the sculpture. Someone who produced pieces of bronze with shapes identical to those belonging to *King and Queen* would not thereby produce an instance of that sculpture. For the pieces to be such instances, they would have to be produced from the cast Moore supplied to a certain foundry chosen by the artist and be one of a specific number of instances as indicated by Moore. This much is obvious. Controversy arises when one asks why this is so and whether a common structure is equally insufficient to individuate musical and literary works.

One explanation of the insufficiency of structure to individuate works such as cast sculptures and prints appeals to a purported distinction between autographic and allographic works (Goodman 1968). The latter are those that, because they are made in a notational symbol system, are in fact individuated by a shared sequence or structure of symbols. A thought experiment suggesting that musical and literary works are allographic relies on the impossibility of forging a musical or literary work by copying the score of one or the sequence of words of the other. This simply would produce the score of the musical work or a copy of the novel rather than something to be passed off for one of these. If, however, someone copied a cast sculpture by creating a new cast that produced a piece of bronze identical in shape to the sculpture, that would be a forgery. To be the sculpture, even one that has multiple instances, each instance must derive in the right way from the hand of the artist. This is what makes sculpture an autographic art form.

One can accept a version of the autographic/allographic distinction that simply says that some works are made in notational symbol systems and others are not. What this version of the distinction does not imply is that if a work is made in a notational system, it is individuated entirely by notational structure. A different thought experiment suggests that even for works in notational systems, instances sharing a common structure are not necessarily the same work. The experiment revolves around structurally identical items from different periods or cultures. Because of the different historical contexts, the items will have different artistic properties despite sharing, say, the same sequence of words. A well-known version of this thought experiment is the often-cited story by Jorge Luis Borges, "Pierre Menard, Author of the Quixote" (1970). In this story, Borges imagines a late-nineteenth-century writer, Menard, who produces a manuscript word-for-word identical to some chapters of Cervantes' great novel. Borges plausibly proceeds to note the huge differences in style and meaning between the two works. Cervantes' style is colloquial, whereas Menard's is self-consciously archaic. The latter contains allusions to contemporary philosophic thought that the former could not possibly have. Hence even ignoring that Menard's text is identical to only a small part of Cervantes', the two are different works in virtue of different authorial intentions and contexts of creation.

Once one recognizes that intentions and context play roles in individuating works that have the ontological status of types or kinds—whether or not they are also "allographic" in the weak sense noted in the preceding paragraph—one can also recognize that intention and context play similar roles in the case of concrete works such as paintings and uncast sculptures. To recall another famous example from Arthur Danto's *The Transfiguration of the Commonplace* (1981) consider three pieces of canvas covered with red paint by the hand of three different, independent artists. Three "structurally identical" red canvases could form parts of a single work of art, a triptych, if produced with that intention by a single artist or a group working together. That three distinct physical objects are produced in isolation from each other, the product of three different "hands," implies that, if each red canvas is or constitutes a work of art, there are three distinct works. However, when does a red-paint-covered canvas constitute a work of art, and what sort of entity is the art object so constituted? The answer to the first question once again appeals to intentions and context. For a canvas uniformly covered with red paint to be a work of art—a painting—an art-historical context must be in place that permits certain intentions to count as art making ones. In eighteenth-century France such institutions were not present, whereas in twentieth-century America (or France) they were. Second, the art-making intentions must actually exist. If one canvas became red because the artist needed an empty spray paint can and got it by discharging the paint onto this canvas, there is no art-making intention and no artwork. If the canvas became red as the result of an intention to produce a work in the color-field genre, then there is an art-making intention, and thus an artwork.

The second, strictly ontological, question asked above is: What sort of entity is the art object? Is it the painting that results from covering the canvas with red paint? Is it identical to the paint-covered canvas or not?

To grasp the ontological puzzle here, it is easier to turn to a different example: a piece of clay shaped into a human figure. Is the sculpture identical to the lump of clay? Obviously not, because the lump existed before it was shaped to create the sculpture, but the sculpture itself did not exist. An alternative answer to this question is that the sculpture is the human-shaped lump. This entity came into existence when the sculpture did and could be regarded as a phase of the existence of the lump itself. However, even such a "phase" or "time-slice" could be understood as having its shape contingently. That is, if it is possible for it to continue existing as one and the same phase of the lump while radically changing in shape, the phase is not identical to the sculpture because the sculpture would not survive such a radical change in shape. Also, if it is possible for this entity to come into existence independently of any human intention, it could not be identical to the sculpture. A sculpture cannot come into existence exclusively through natural processes. These considerations imply that the entity identical to the sculpture is not simply a lump of material structured in a certain way but such a lump structured to fulfill an artistic function or intention typically made possible by certain institutions or practices. Exactly the same is true of the red painting. It is a canvas covered by red paint to in order fulfill an artistic function or intention made possible by certain institutions or practices (Levinson 1996).

ARE ALL ARTWORKS CREATED? Concrete artworks such as the red paintings and the clay sculpture just discussed are obviously created. Are the abstract works—the novels, musical pieces, and so on—also created? Some scholars, such as Jerrold Levinson (1980), have argued that it is a condition on a satisfactory ontology of art that the ontology accounts for the createdness of abstract artworks. Others, such as Peter Kivy (1993) and Julian Dodd (2000, 2002), have disputed this. Underlying these conflicting views are conflicting intuitions. One intuition is that novels, plays, and musical works are just as much the products of creative activity as are paintings. The other intuition is that abstract objects cannot be created because of the sort of objects that they are.

It may seem that the argument of the preceding section supports the claim that abstract artworks are created. It was argued that these works are not identical to abstract structures per se, but to structures tied to certain intentions and contexts. What could "tied to" mean but created with certain intentions in a certain context? But this raises an important question: What are these entities that are purportedly created? They are not pure abstract structures, because these are really uncreated and eternal, and it has already been denied that they are the artworks. The best known proposal on this score is Levinson's. He claims that they are indicated structures "a structure-as-indicated-by-P-at-T-in-[art]-historical-context-C" (1996, p. 146). The dashes are intended to indicate that this is not a set-like ordered quadruple but something more "unified," a type that comes into existence with the act of musical or literary composition.

There are a variety of objections to Levinson's view. Stefano Predelli asserts that indicating does not in general create new entities (1980). If I point out my favorite house in the neighborhood, I haven't created a new entity: the-house-as-indicated-by-me. So it is implausible that indicating creates one when authors or composers indicate abstract structures. There are two ways of replying to this objection. One reply would be to claim that new entities are always created by indicatings, but people pay no attention to most of them because the indicatings are of no interest them. The house-as-indicated-by-me is an entity that has about as much interest as a scattered object such as a nose-tie consisting of Bill Clinton's nose and a tie he left in a hotel during a visit to Australia. Both nose-ties and indicated-buildings nevertheless exist. The other reply claims that some indicatings are special because they occur within institutions or practices that endow them with special properties and give them special recognition. Sentences can be regarded as abstract syntactic structures, which, when used (when *indicated* by a speaker or hearer) creates a new entity, which has semantic or pragmatic properties not possessed by the abstract sentence type. The ability to convey something distinct from the semantic meaning of the sentence type results from linguistic conventions combined with the intentions of language users and the context of use. Writers are just special cases of people who use (indicate) strings of sentence types to convey something through the creation of a complex literary object. Composers do something similar with abstract sound structures. The two replies are consistent with each other, though the second is available to those who would resist the first.

A second objection is that abstract entities, such as structural types, cannot be created because they cannot enter into causal relations. Being created means being caused to exist and, if an entity cannot enter into a causal relation, it cannot be caused to exist. This claim is said to apply to any abstract type whether it be of the pure unindicated variety or an indicated structural type. A related third objection should also be mentioned at this point. It could be said that even if there are indicated types, they are just as eternal and uncreated as any other

abstract thing. Such types exist just in case a property corresponding to the type does, and all properties exist eternally. Hence the property of being a structure indicated by P at T in C exists eternally. Therefore the indicated structural type does too. Hence it is not created (Dodd 2000, 2002).

Both of these objections are too tendentious to be decisive. The issue of whether types can be caused to exist is not settled by their being abstract; they are abstract because they have instances or tokens. Someone could claim that a type does not exist until at least one token of it does, or instructions for creating a token are present. In either case causing the token (or the instructions for making tokens) to exist in effect causes the type to as well. There are many types that it is plausible to conceive in these terms. Consider artifact types, one example of which is an automobile model. It is plausible that automobile designers bring this type into existence when they create the design for a car model. This plausible claim is deniable. It could be consistently maintained that the type Volkswagen Beetle would exist even if intelligent life had never evolved anywhere in the universe. Though consistently maintainable, the claim is implausible. Saying it is tendentious is perhaps an understatement. If this is true for car models, it would be equally true for literary and musical works. So one may perhaps set aside the second and third objections to the idea that indicated structural types are a kind of entity that can be brought into existence.

INDIVIDUATING ARTWORKS

What has been demonstrated thus far is that indicated structures are distinguishable from unindicated ones, and that the idea that they come into existence—indeed, are brought into existence—is, at least, plausible. However, there is a final set of objections to them that question whether they individuate musical and literary works correctly. Are such works always essentially tied to the precise times they are created, to their creators, and to their context of creation? This is what is denied by the final set of objections.

Look first at authorship. It may be true that Cervantes and Menard (had there been such a person) could not possibly create the same work. But imagine two contemporary writers, composers, or even painters who belong to the same school working at the same time. There are two different scenarios to consider. One occurs when both produce identical works. Suppose Mozart and Haydn had produced, independently of each other, identical scores for a string quartet in the year 1787. Would they both have independently composed the same work? An alternative scenario can be created by supposing a possible world in which Haydn instead of Mozart composed a score identical to the score for Mozart's G major quartet K.387 and in which Mozart produced no such score. Would Haydn have composed in this possible world the same quartet that Mozart composed in actuality? Some people would answer yes to both of these questions; but, if that answer is right for either one, the identity of the artist *may not* be essential to the identity of the work. The first scenario does not raise a problem when it comes to painting because two numerically distinct painted canvases from the hand of different artists are different paintings even if they are indistinguishable. The second scenario, however, raises the same question for painting as it does for music or literature. In a possible world in which Braque rather than Picasso had painted a portrait of Gertrude Stein exactly similar to Picasso's actual painting, would Braque be the artist responsible for the work Picasso actually made (Currie 1988, S. Davies 2001)?

Something similar can be said about the time at which the work is indicated or brought into existence. Is this always an essential property of artworks? Some works seem to be tightly tied to their time of production. Hemingway's fiction is closely tied to the World War I generation. Picasso's *Les demoiselles d'Avignon* seems even more tightly tied to its moment of production. But consider traditional African sculpture from a particular region, some forms of traditional Chinese painting, or the naive work of an amateur artist, all of which may remain unchanged in style over many years. In these cases, it may seem plausible that the same work could be produced many years apart in different possible worlds. However, it seems possible that even those works that seem most closely tied to a moment in time might have been produced at slightly different times or, in special circumstances, very different times. Consider a possible world that duplicates the history of European art, but in which that whole history begins two hundred years earlier than it in fact did. In that world, Picasso paints *Les demoiselles* in the early eighteenth century (D. Davies 2003).

The contextual variable is perhaps immune to considerations such as those just raised about artist and time of production. Works from different eras, traditions, styles, or works made with different intentions will not be the same no matter how superficially similar they appear. This is an important point of the Menard example. Nevertheless, a case may be made for the possibility of the same work in different contexts by appealing to the idea

that slight differences in context in different possible worlds may still result in the same work. This is especially plausible if the specific difference in context would not make a difference to the creator of the work in question (D. Davies 2003).

This set of objections raises two broad concerns for the idea that musical and literary works are indicated structural types and that paintings are contextually identified physical objects. It raises objections to Levinson's specific proposal regarding the individuation of indicated structures, but it also questions whether any general formula appealing to any of the variables under discussion can individuate artworks correctly.

Before concluding that these concerns are correct, one needs to evaluate the objections on which they are based. Do the objections show what they set out to show? One problem with them is that they rely on uncertainties in human modal intuitions about artworks, which point more directly to epistemic rather than metaphysical possibilities. That is, in the face of the sorts of examples considered thus far, many individuals will be uncertain what to think, and so it will be epistemically possible, relative to their beliefs, that a certain principle of individuation is wrong. That, however, falls short of showing it is wrong.

Is there a way of sharpening intuitions to arrive at principles of individuation? Perhaps this can be done by getting clearer about what the Menard example and other similar examples reveal about structurally identical works. Cervantes and Menard had different artistic projects and, in pursuing these, each achieved (did) different things in their respective works. This pair of differences, concerning artistic project and artistic achievement, is crucial in individuating works and in identifying important artistic properties of them. In highlighting these differences, the analogy mentioned earlier between abstract sentence types and utterances—or, more broadly, sentences-in-use—is a helpful one to remember. The language user in question, along with the user's intentions, the time of utterance, and the context of use, all commonly contribute to fixing what the utterance conveys beyond or in distinction from the semantic content of the sentence. However, the precise role each of these items plays may vary in different uses of language. Further, it is possible for different utterances to convey precisely the same thing. Regarding artworks, something similar is true (Stecker 2003): They are individuated by being a specific abstract or concrete structure used by an artist in pursuing such and such a project and achieving so and so. Usually the three variables—identity of artist, time of creation, and artistic context—are crucial in constituting projects and what they achieve, yet their exact role can vary in different art forms and different traditions, as well as for many other reasons. The emphasis on the artistic project and the artist's achievement recalls the idea that artworks are action types or tokens. However, those views identified the work with the wrong entity. As the indicated-structure view emphasizes, the artwork is the product that results from the project and embodies the achievement.

See also Aesthetics, History of; Aesthetics, Problems of; Art, Definitions of; Danto, Arthur; Existence; Ontology; Ontology, History of.

Bibliography

Borges, J. L. "Pierre Menard, Author of the Quixote." In *Labyrinthes*, translated by J. E. Irby, 62–71. London: Penguin, 1970.

Caplan, B., and C. Matheson. "Can Musical Works be Created?" *British Journal of Aesthetics* 44 (2004): 113–134.

Currie, G. *An Ontology of Art*. London: Macmillan, 1988.

Danto, A. *The Transfiguration of the Commonplace*. Cambridge, MA: Harvard University Press, 1981.

Davies, D. *Art as Performance*. Oxford: Blackwell, 2003.

Davies, S. *Musical Works and Performances*. Oxford: Oxford University Press, 2001.

Dilworth, J. *The Double Content of Art*. Amherst, NY: Prometheus Books, 2004.

Dodd, J. "Musical Works as Eternal Types." *British Journal of Aesthetics* 40 (2000): 424–440.

Dodd, J. "Defending Musical Platonism." *British Journal of Aesthetics* 42 (2002): 380–402.

Goodman, N. *Languages of Art*. New York: Bobbs-Merrill, 1968.

Howell, R. "Ontology and the Nature of Literary Works." *Journal of Aesthetics and Art Criticism* 60 (2002a): 67–79.

Howell, R. "Types, Indicated and Initiated." *British Journal of Aesthetics* 42 (2002a): 105–127.

Kivy, P. *The Fine Art of Repetition*. Cambridge, U.K.: Cambridge University Press, 1993.

Levinson, J. "What a Musical Work Is." *Journal of Philosophy* 77 (1980): 5–28.

Levinson, J. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996.

Predelli, S. "Musical Ontology and the Argument from Creation." *British Journal of Aesthetics* 41 (1980): 279–292.

Rohrbaugh, R. "Artworks as Historical Individuals." *European Journal of Philosophy* 11 (2003): 177–205.

Stecker, R. *Interpretation and Construction: Art, Speech, and the Law*. Oxford: Blackwell, 2003.

Wollheim, R. *Art and its Objects.* 2nd ed. Cambridge, U.K.: Cambridge University Press, 1980.

Wolterstorff, N. *Worlds and Works of Art*. Oxford: Clarendon Press, 1980.

Robert Stecker (2005)

ART, PERFORMANCE IN

Some philosophers hold that the creation of art always involves performance, and that artworks are more accurately defined as processes or actions than as objects or events. This entry will consider the more traditional view that only some art forms—drama, music, dance, opera or musical theater, and "performance art"—involve performance.

Performances can be freely improvised. In addition to being judged for their general interest and skill of execution, such performances are rated as well for elements of spontaneity and risk. The performers make and coordinate their activities in real time, without knowing how their performance will continue or end. Though drama can be extemporized, jazz takes improvisation to its highest level. In the paradigm case, however, performance involves the live presentation and interpretation of a previously specified work.

THE LIVE PRESENTATION OF WORKS

Works for performance are often specified through a form of notation, such as a musical score or a script. The notation is addressed to the performer and prescribes what must be done or achieved if the work is to be faithfully performed. It may also contain recommendations that are not work-determinative, and that need not be followed. Features crucial to the work's identity are not always mentioned in the notation, for instance, where they are dictated by practices and conventions that are taken for granted. The performer's first act of interpretation occurs in following and understanding what is instructed in the work's notation, if it has one, along with appreciating the background of performance practices and conventions that it assumes.

In oral traditions, works are transmitted verbally, not by notations. One or more suitably authorized performances are given the status of a model for further instances of the work. Just what in the model is work-specifying and what is merely optional is settled by reference to the work-and-performance traditions and genres within which the relevant piece is located. For example, although the melody in the exemplary performance is elaborately decorated, it might be understood that the manner of decoration is left to the performers' discretion, as long as they respect limits set by the appropriate style. Or it might be understood that the choreography of a sword fight need not be aped in subsequent performance, though appropriate fighting actions will be required.

The actor's, singer's, or dancer's medium is her body, along with costumes, props, and sets. For other musicians, their instruments are their media. When a work is designed for performance, its medium is usually crucial to its identity, since the medium affects and constrains what can be done by the performer. The artist's instructions usually indicate both what is to be achieved and the medium or manner in which this is to be done. To perform a violin concerto, one should play the violin. Merely generating the appropriate sounds on a synthesizer (or a record player) does not qualify as performing the work.

Some works call for media that are not standard. Electronic compositions for live performance involve the use of microphones, sound generators, and the like. One of John Cage's pieces was issued as a vinyl disk with instructions about how the settings of the hi-fi amplifier are to be modified as the disk plays. Hip-hop artists and sound appropriators take the recordings of others as source materials for their own works and, like Cage, turn the record player into an instrument of musical performance.

INTERPRETATION

Works that are for live performance are always ontologically thinner and more abstract than the concrete performances that instance them. If all the artist's work-determinative prescriptions are faithfully followed, many aspects of the performance's detail are not determined. The performer (or conductor/producer) resolves these uncertainties. The playwright might indicate that the actor is to say "Curse the gods," but the choice of facial expressions, gestures, and bodily attitude, along with the tone, inflection, pitch, and volume of the voice, are usually left to the actor. Whether through deliberation or not, the delivery of the line in an actual performance inevitably will display a particular version of all these features.

In adding flesh to the skeleton that is the work, and thereby creating a living performance, the performer interprets the work. It would be misleading, though, to say that interpretation is something added by the performer after he or she has satisfied the work-determinative prescriptions of the artist, or to suggest that interpretation fills the gap between the work's abstractness and its performance's concreteness. The delivery of the work is not prior to or apart from the interpretative contribution, which is crucial at every point or moment. The presentation of the words of a play or the notes of a symphony is not separable from the manner and inflection with which they are presented.

Some works for live performance can be very spare, ontologically speaking. Songs are so, when specified only as a melody, sequence of chords, and verse and chorus. Very different interpretations can be consistent with the faithful presentation of such works. The thinner the work, the more the performer becomes the focus of attention and the more the evaluation of the performance concerns the performer's creativity and vitality, rather than the faithful delivery of the work. But even where works are very detailed, as are Mahler's symphonies or Shaw's plays, the importance of the performer's interpretative contribution cannot be overlooked. Indeed, complex works offer the performer wonderful opportunities for displaying her talents, because their realization is unmistakably demanding and they allow for interpretations that are subtle, rich, and multilayered. Some works, such as instrumental concertos, are intended to draw attention to the virtuosic performances they require.

If all live performances embody interpretations of their works, so do thoughtless, unplanned performances and mechanical ones learned by rote. In the normal case, however, the interpretation is planned by the performer who delivers it and reveals a considered vision of the work. Some performers concentrate on the work's progression from moment to moment, leaving the artist's design to ensure that the whole is satisfying. Other performers structure their efforts in terms of a conception of the work's overall structure. Some performers can describe the ideas that inform their interpretations, while others have a more applied, unarticulated knowledge of what they do.

An interpretation, once mastered, can be repeated. Different performances of a production of a play usually present the same interpretation. Yet a given performer can have more than one way of interpreting a given work. A performer with a long career often adopts a fresh approach when she returns to works she performed previously.

AUTHENTICITY AND INTEGRITY CONDITIONS

The purpose of a performance of a work is to present the work along with an interpretation of it. Such performances therefore presuppose a commitment of faithfulness to the work, or authenticity. Deliberate departures from the work undermine the claims of a performance to be of that work. Accidental errors and slips in performances need not prove fatal to the attempt at performance, however. A performance can instance a work because of its intent, and the work can be recognized in what is produced, despite the imperfection of its representation of that work.

There is disagreement about what faithfulness requires when questions such as the following are debated: Is it necessary to use boys rather than women when performing Shakespeare? Should Scarlatti be performed on the harpsichord only, and can its jacks be made of plastic instead of quills? If an eighteenth-century playwright specifies that his work is set in the present, should we use period costume or the clothes and milieu of the twenty-first century?

Such disagreements can reflect deeper differences of opinion about the ontological character of the works in question. Someone who thinks a musical work is merely a pattern of notes will regard any presentation that reproduces that pattern as faithful, no matter what means are used to produce it. But another who believes the work's instrumentation is also central to its identity will conclude that authentic performances must use instruments of the kind known to and specified by its composer. Differences between their ontological theories lead philosophers to draw the line between performers' legitimate liberty and illegitimate license in contrasting places.

There is another reason for conflict, though. Some people think that authenticity can be traded for interpretative interest. In other words, they do not regard the pursuit of faithfulness to the work as a paramount virtue in performances. As supporting evidence, they may cite the free approach sometimes taken to the interpretation of Shakespeare and of the most famous musical works and operas. It might not be coincidental, however, that works approached in this manner are very familiar to the established audience and that there is a concern to maintain their relevance for future audiences. In other words, the free approach to interpretation in these cases is not necessarily indicative of indifference to or disrespect of the work as such. Provided that audiences are interested in the works being performed, authenticity in performance cannot be reduced merely to another interpretive option.

Stan Godlovitch (1998) specifies the following conditions for the integrity of performances: only one work is performed at a time; its proper sequence is respected, as is the indicated rate of delivery; the performance is continuous, without unjustified breaks; performers comply with the appropriate roles (and do not, for example, swap parts midway through). Also, the audience is in a position to receive the entire performance in its detail. Not all of these conditions are satisfied in all performances. Nevertheless, these conditions are normative in that they indicate what is expected from a performance.

Activities not directed to an audience—practicing, rehearsing, learning, doodling—do not result in performances, according to these conditions. (In many cases, such activities have the goal of preparing performers for performances, however.) Other performance-like activities violate other of the specified conditions and are not exemplary for that reason. Music-minus-one disks and karaoke (as well as new technology, allowing a person to speak one of the roles in a movie) are examples.

STUDIO PERFORMANCES

Not all performances are given live. Some take place in studios and result in recordings or films. Studio performances have their own integrity conditions. The work's segments can be recorded piecemeal and out of order. A single performer can take many different roles in the finished product, as a consequence of multitracking or filming. The performer's inputs can be electronically modified. The projected audience is not present to witness the studio performance.

We accept studio performances of pieces created originally for live performances, such as recordings of classical symphonies or movies of Elizabethan plays. They may use some of the studio's resources, such as the possibility in film of moving seamlessly between different indoor and outdoor locations. But in general, they simulate live performances, and the artists involved are capable of giving live performances.

Some works are designed for studio performances. Rock recordings that sculpt sounds electronically in a fashion that could not be achieved in real time are examples. These are works for performance, but not for live performance. The same song can be recorded by another group, and the result is a new (studio) performance of it, not a different but a derivative work.

Yet other works are not *for* performance of any kind, though they involve studio performances in their creation. Most films rely on the resources of the studios (slow motion, flashbacks, stunts, digital editing, and special effects) and result in works that are for screening, not performance. Similarly, purely electronic musical works are for playback, not for performance, though performers might supply material that is integrated into the work. "Directors' cuts" result in new versions of movies, not in new performances, while remakes result in new but derivative works.

PERFORMANCE ART

During the mid-twentieth century, artists began to challenge traditional conceptions of artworks and the separation of art from life by using their own bodies as the medium for their works. They posed in public or structured some aspect of their lives in terms of an aesthetic goal; they lived in cages or staged happenings. Different strands of the movement featured bodily mutilation, sexual orgies, and primitive rituals, often intended to deliver a political or socio-sexual message. Some feminist artists embraced performance art for its liberating energy, but were sensitive also to the need to subvert the objectifying equation of women with their bodies. Performance artists have often integrated video into their artworks. The works of the French feminist Orlan display many of these features; they are films of the surgical alteration of her face to give it the features of famous art-historical beauties.

See also Aesthetics, History of; Aesthetics, Problems of; Greek Drama; Music, Philosophy of; Tragedy.

Bibliography

Alperson, Philip. "On Musical Improvisation." *Journal of Aesthetics and Art Criticism* 43 (1984): 17–30.

Carlson, Marvin. *Theories of the Theatre: A Historical and Critical Survey from the Greeks to the Present*. Ithaca and London: Cornell University Press, 1984.

Carroll, Noël. "Interpretation, Theatrical Performance, and Ontology." *Journal of Aesthetics and Art Criticism* 59 (2001): 313–316.

Carroll, Noël. "Theater, Dance, and Theory: A Philosophical Narrative." *Dance Chronicle* 15 (1992): 317–331.

Copeland, Roger, and Marshall Cohen, eds. *What Is Dance*? New York: Oxford University Press, 1983.

Davies, David. *Art as Performance*. Oxford: Blackwell, 2003.

Davies, Stephen. *Musical Works and Performances*. Oxford: Clarendon Press, 2001.

Godlovitch, Stan. *Musical Performance: A Philosophical Study*. London: Routledge, 1998.

Hamilton, James R. "Theatrical Enactment." *Journal of Aesthetics and Art Criticism* 58 (2000): 23–35.

Hamilton, James R. "Theatrical Performance and Interpretation." *Journal of Aesthetics and Art Criticism* 59 (2001): 307–312.

Kemal, Salim, and Ivan Gaskell, eds. *Performance and Authenticity in the Arts*. New York: Cambridge University Press, 1999.

Kivy, Peter. *Authenticities: Philosophical Reflections on Musical Performance*. Ithaca: Cornell University Press, 1995.

Mark, Thomas Carson. "On Works of Virtuosity." *Journal of Philosophy* 77 (1980): 28–45.

Mark, Thomas Carson. "Philosophy of Piano Playing: Reflections on the Concept of Performance." *Philosophy and Phenomenological Research* 41 (1981): 299–324.

McFee, Graham. *Understanding Dance*. New York: Routledge, 1992.

Saltz, David. "How to Do Things on Stage." *Journal of Aesthetics and Art Criticism* 49 (1991): 31–45.

Sparshott, Francis. *A Measured Pace: Toward a Philosophical Understanding of the Arts of Dance*. Toronto: University of Toronto Press, 1995.

Sparshott, Francis. *Off the Ground: First Steps to a Philosophical Consideration of the Dance*. Princeton, NJ: Princeton University Press, 1988.

Thom, Paul. *For an Audience: A Philosophy of the Performing Arts*. Philadelphia: Temple University Press, 1993.

Stephen Davies (2005)

ART, REPRESENTATION IN

Pictures form a subset of the artifacts that serve to represent particular things or kinds of thing, real or imagined, in a broad inclusive sense of the term *represent*. Like some of their fellow representations, but unlike others, pictures go on to attribute properties to the things or kinds they represent—properties that thereby constitute their pictorial content. How does this work? What distinguishes the representing done by pictures—*depiction*—from the representing done by various other familiar kinds of representation?

PLATO AND PICTORIAL MIMESIS

Near the start of his case for banishing the poets from his ideally just city (Plato 1992, *Republic* X, 595a–598b), Plato urges that poetry and painting are analogous mimetic activities, structured so as to be able to imitate—approximately replicate—only a superficial and trivial part of the deep and serious things they profess to take as models. The argument employs a three-story metaphysics with Plato's Forms at the top, ordinary three-dimensional worldly particulars in the middle, and appearances (*eidola*, *phainomena*, *phantasmata*) at the bottom. Paradigm cases of appearances are shadows and reflections. Shifty, shimmery, and insubstantial, they owe such limited stability and stable apprehensibility as they to their owners, the three-dimensional worldly particulars from which they derive and to which they bear a real if limited resemblance. They therefore bear to worldly particulars many of the relations that worldly particulars are said to bear to the Forms.

The phrase "what S sees of X here and now" may be taken to refer to another appearance, another insubstantial something owing such limited stability and stable apprehensibility as it possesses to its three-dimensional owner X, the entity it manifests and imperfectly resembles. Such an appearance differs from a reflection or cast shadow in that it is attached to or embedded in its owner. In fact, it may be regarded as literally a part—albeit a dependent and ontologically inferior part—of that owner.

Now painters and poets are mimetic artists, renderers. Painters undertake to render three-dimensional arrangements of physical objects and to do so on a two-dimensional surface, using as their medium line and color. Tragic poets undertake to render human agents engaged in spontaneous morally significant action and to do so on a stage before an audience, using as their medium the rehearsed movements and speeches of actors. One may take these renderers at their word when they say that they are out to replicate important worldly originals to the full extent it is in their power to do so. Still, what extent is that? Given the materials he must work in and the way he must manipulate these materials to count as a painter or poet at all, the most such an imitator can ever accomplish by way of replicating his original is to produce a second worldly particular almost entirely unlike the first except for possessing an exactly similar appearance. His would-be traffic in second-rate entities (worldly particulars) comes to no more than a traffic in third-rate entities (appearances). "Imitation is far removed from the truth, for it touches only a small part of a thing and a part that is itself only an image" (598b).

Add that what meets the eye (or ear) about an important or valuable object seldom if ever includes what makes it behave as it does or what makes it a good or bad thing of its kind, and one will have powerful reason for suspecting that the theoretically and practically decisive aspects of worldly particulars lie beyond the reach of the senses, hence beyond the reach of the particular media that make painters painters and poets poets. Echoes of Plato's reasoning abound in texts as recent as Susan Sontag's *On Photography* (1977).

Thinkers who reject Plato's metaphysics and his deprecatory attitude toward painting nevertheless often agree that depiction consists in the partial replication in a new and alien medium of a certain superficial aspect of the depicted thing's nature, something inherently capable of meeting the eye, call it the depicted thing's outward appearance. Such thinkers have various ways of embracing Plato's account of depiction's workings while avoiding his negative conclusions about painting's value. Sometimes they insist that depictive success is one thing and artistic success is something different and deeper. Sometimes they insist with Oscar Wilde that there is nothing superficial about surfaces.

There have always been dissenters, of course. One is René Descartes, who insists that engravings successfully

portray the things they depict as having lots of properties they could not possibly share with those things. Indeed, when it comes to objects standing at a great distance or whose accurate depiction requires foreshortening, “the perfection of an image often depends on its not resembling its object as much as it might” (Descartes 1985, *Optics*, Discourse IV, AT 113). Descartes thereby prepares his reader for the alarming thought that our most useful and reliable sense-based ideas resemble their originals as little as engravings do theirs.

GOMBRICH AND THE PURSUIT OF ILLUSION

A vast renewal of philosophical interest in depiction begins in the 1950s with the work of the art historian E. H. Gombrich. Like Hermann von Helmholtz, Karl R. Popper, R. L. Gregory, and others, Gombrich holds that the content of visual experience is produced in a kind of unconscious inference by the human visual system to the best available explanation of the available retinal stimuli. (The stimuli on which visual system inferences are ultimately based remain permanently out of introspective reach.) The conceptual resources a visual system may draw on in framing these hypotheses include any and all concepts available to its owner and the standards by means of which it assesses them are sensitive to the full range of beliefs, expectations, and practical priorities its owner brings to the task of seeing what is before his eyes. The beholder’s share in determining the content of his own visual experience is therefore substantial indeed; there is no such thing as *the* appearance a thing can possess when accurately seen from a particular physical viewpoint.

Only one particular kind of image, the *naturalistic* kind, is out to replicate an appearance taken on by a particular object in a particular context for a particular sort of appropriately prepared spectator. Naturalistic image making catches on only in particular cultural traditions at particular times. Images more generally are best conceived as substitutes for the things they depict, standing in for them in various forms of ritual and imaginative activity and sharing with them only the handful of specific properties, visible and otherwise, required for this special purpose. (Think of how a hobby horse stands in for a real horse.) In this sense, *making* (the production of substitutes) comes before and is more generally prevalent than *matching* (the production of objects designed to visually match the things they depict under appropriate objective and subjective conditions).

Consider a naturalistic image maker, out to capture some particular appearance of the particular object she is about to depict. Just as there is no way for her to set aside the effects of past encounters with other objects when it comes to trying to see this one accurately, there is no way for her to set aside the effects of past efforts to depict other objects when it comes to trying to render her depiction of this one appropriately responsive to how she now sees it. Instead she must rely on habits, routines, and formulas inherited from past image-making practice to give her a skeletal generic image of an object of the right general kind, which she then works over in a trial-and-error manner until she finally achieves a convincing likeness of this particular object. Naturalistic image making is a process of *schema and correction*.

Such small-scale explorations contribute to the larger-scale explorations conducted in image-making communities as they invent, refine, and promulgate redeployable techniques for appearance-capturing techniques based on hard-won empirical insights into how the human visual system works. Foreshortening, tonal modeling, and the various perspective systems are major inventions of this sort, but there are countless smaller ones: Think of Rembrandt’s readily imitated trick of suggesting the glint of gold braid with a few loose, broad dots and dashes of yellow paint. When and where the naturalistic project catches on in the first place, the history of art largely consists in the history of such progressive innovation. When and where it does not catch on, art may change over time, but not in ways that possess the large-scale narrative coherence historians demand.

> The history of art … may be described as the forging of master keys for opening the mysterious locks of our senses to which only nature herself originally held the key. … Like the burglar who tries to break a safe, the artist has no direct access to the inner mechanism. He can only feel his way with sensitive fingers, probing and adjusting his hook or wire when something gives way. Of course, once the door springs open, once the key is shaped, it is easy to repeat the performance. The next person needs no special insight—no more, that is, than is needed to copy his predecessor’s master key (Gombrich 1961, pp. 359–360).

Gombrich’s relation to Plato is complex. When properly experienced, a successful naturalistic image partially replicates an appearance the depicted object is capable of taking on. But this appearance is not a superficial part of the object; it is an effect of the object on a particular spec-

tator made possible by the particular concepts and concerns he brings to the act of seeing. The artist devises her own means of achieving *some* of the effects the depicted object would have on a spectator's visual system if it were standing before him. But limitations inherent in her media—the restricted range of lights and darks available from her paints and inks, the manifest flatness of the surface on which she deposits these substances—ensure in advance that her replication of the object's appearance is partial at best. A depiction takes on the appearance the artist intends it to take on only if the spectator actively brings to his inspection of it the highly particular mental set the artist intends for him.

Still, and here Gombrich again sides with Plato, the experience the naturalistic image maker means to induce in the spectator is one in which it is for him as if he were seeing the depicted object face to face. This means he manages to neglect the lack of appropriate color (in drawings), the lack of appropriate binocular disparity (in full-color paintings of nearby objects), and so on. It also means that as the picture takes on its intended appearance for him, the content of his visual experience of the picture has less and less to do with the picture, and more and more to do with the thing depicted. He loses sight of the depiction in favor of the thing depicted, with the result that the specific devices by means of which the image maker induces the intended experience drop from visual awareness at the very moment they achieve their intended effect. Gombrich concludes that naturalistic image makers are inducers of illusion and that illusion obliterates its own conditions.

This illusion will be available to a given spectator only if he can approach the painting with its called-for mental set, hence only if he can readily identify this set and readily assume it without detailed instructions. Pictorial intelligibility is a special case of communicative intelligibility, depending on a rich, historically variable, culturally conditioned stock of expectations, assumptions, and conventions. In order to generate and disappear into an appropriate illusion, a set of marks must first be correctly interpreted as a communicative gesture on the part of the artist. Like Ferdinand de Saussure and Roman Jakobson before him, Gombrich olds that to understand any communicative gesture, one must view it as a choice from among a fixed range of available alternatives, owing its significance in part to the natural significance of certain dimensions of difference (darker tones are naturally taken to signal darker objects, more vigorous gestures to signal greater urgency), in part to the conventional fact that one is tacitly but publicly committed to working within such and such a restricted set of choices (only these tones, only these gestures). To this extent, at any rate, art is a language, a system of signals resting on contingent and mutable conventions that must be internalized and respected by artists and audiences alike. All three main approaches to understanding depiction draw heavily on Gombrich's work, accepting some strands of it while rejecting others.

INTELLIGIBILITY ACCOUNTS

According to the intelligibility approach, pictures are distinctive in virtue of how one's ability as an audience member to make appropriate interpretive sense of them builds on and derives from one's ability as a perceiver to make appropriate visual sense of one's immediate physical surroundings.

J. J. Gibson (1971) holds that perceivers extract certain crucial elements of a picture's content (e.g., depicted recessions in depth) from features of the marked surface (e.g., texture gradients across that surface), using precisely the same methods they use to extract corresponding features of their real visual environment (e.g., actual recessions) from locally available features of the visual stimulus (e.g., texture gradients across one's visual field). Pictorial understanding is just routine environmental feature extraction applied to a special artificially contrived stimulus: a picture's marked surface. The proposal is closely bound up with Gibson's idiosyncratic account of ordinary visual perception, his environmental optics.

Flint Schier (1986) proposes that pictures exhibit a distinctive division of cognitive labor between the mastery of particular pictorial idioms and the ability to visually recognize a particular thing or kind when presented with it face to face. On the one hand, pictorial idioms possess *natural generativity*: every pictorial idiom is such as to make possible a picture so representative of the idiom as a whole that understanding this particular picture would suffice to confer a general competence with the entire idiom. On the other hand, the interpretation of any given picture P redeploys ordinary capacities for face-to-face visual recognition in such a manner that:

> (1) a general competence with the idiom employed by P, and
>
> (2) a capacity to recognize each of the particular things or kinds that P depicts (and each of the particular visually detectable properties and relations that figure in P's pictorial content)

are individually necessary and jointly sufficient for a given spectator to be able to understand P. As it stands,

Schier's proposal makes no allowance for the depiction of particular people that nobody could recognize on sight—Christ, for instance—nor does it allow for the large role collateral information plays in the correct interpretation of many pictures. Still, it feels like a first approximation to an important insight.

SEMIOTIC ACCOUNTS

According to the semiotic approach, pictures are conventional symbols in a richer sense than Gombrich allows. A *symbol system* involves syntactic rules that classify items as tokens of various permanently available symbol types, together with semantic rules determining what an object must be like if it is to *comply with*—be such that it could be accurately symbolized by—tokens of a given type. What differentiates pictures from other conventional symbols are distinctive structural features of the systems to which they belong and from which they derive their pictorial content—*pictorial* symbol systems.

Nelson Goodman (1976) never offers sufficient conditions for a system's being pictorial, but he declares that a system cannot be pictorial unless it is syntactically dense, semantically dense, and relatively replete. The effect of the first condition is to insist that there is no limit to how similar two pictorial symbol tokens can be while remaining tokens of distinct symbol types. The effect of the second is to insist that there is no limit to how similar two objects can be while remaining such that the accurate depiction of one and the accurate depiction of the other would require the deployment of two distinct symbol types, one for each. The effect of the third is to insist that a relatively large range of perceivable features of a given pictorial symbol token are relevant to determining its type. Yet depictions formed in the array of lights on a baseball scoreboard fail to exhibit any of Goodman's three features.

Commonly the most salient parts of a picture are depictions in their own right, depicting parts of the larger whole depicted by the bigger picture and arranged in a manner reflective of the arrangement in this larger whole of those depicted parts. This constitutes an interesting affinity between pictures and such manifestly conventional representations as maps and diagrams. Andrew Harrison (1991) infers from it that maps, diagrams, *and* pictures are conventional symbols, belonging to systems whose (compositional) syntax and semantics relate the part-whole structure of complex symbols to the part-whole structure of compliant objects in an especially simple and uniform manner. Yet while maps and diagrams often come equipped with keys explaining their simplest individual significant components, full-fledged pictures do not and apparently cannot come with anything comparable.

EXPERIENTIAL ACCOUNTS

According to the experiential approach, Gombrich is correct in thinking that pictures operate by inducing a distinctive kind of experience, with the thing depicted figuring in the content of that experience. But he is wrong to attribute an illusionist phenomenology to the experience. Instead we should conceive it as a unitary experience, visual at least in part, whose content involves both the depicted thing and various visible features of the depiction itself. There are three main stories about how this goes.

Experienced resemblance theorists (Peacocke 1987, Budd 1993, Hopkins 1998) hold that when one experiences a picture appropriately, one visually experiences it as resembling the thing or kind of thing it depicts with respect to certain of the visually detectable properties possessed by each. Hopkins's version of the theory centers on a highly relational property known as outline shape. Begin with the cone of rays connecting visible points on the object's facing surface to a given perceiver's point of view. Take the intersection of that cone with a plane perpendicular to the perceiver's line of sight. The shape of this intersection is the object's outline shape for the particular perceiver in question. Hopkins is at pains to argue that despite the arcane way outline shape is defined, people are ordinarily implicitly visually aware of the outline shapes of things around them. He contends that whenever one experiences portion D of marked surface P as depicting object O, one visually experiences the outline shape of P (as seen from where one actually stands) as resembling that of object O (as seen from an appropriate hypothetical place).

The most basic kind of pictorial content accruing to pictures in any given idiom consists of resemblances to parts of the picture surface itself with respect to some fixed list of visually detectable determinable properties renderable in that idiom. The list always includes outline shape; it sometimes includes such further properties as local color and texture. One can call the properties on such lists *visual field properties*. Hopkins maintains that portion D of picture P depicts object O if and only if we are meant to experience D (as seen from here) and O (as seen from some appropriate hypothetical place) as resembling one another with respect to the visual field properties renderable in P's idiom.

Hopkins's indebtedness to the optical approach to picturing running from figures like Euclid in the ancient world to figures like Leon Battista Alberti in the Renaissance is obvious enough. Yet when I inspect the portion of Pablo Picasso's *Guernica* depicting a lantern carrier, I am acutely aware that its outline shape resembles that of a teardrop. (It is by making me aware of this shape and this fact about it that Picasso suggests the haste and strain with which the lantern carrier peers into a scene of carnage from a position outside and behind it.) Still, I do not see the relevant portion of *Guernica* as depicting a teardrop, aware as I am and am meant to be of the just-mentioned resemblance in outline shape. Moreover, I do not experience the lantern carrier's neck and head as having an outline shape resembling that of the portion of *Guernica* by means of which they are depicted—to wit, one very much like that of a teardrop. To do so, I would need to experience the lantern carrier himself as having a flat face and a neck tapering off to nothing— and I do not.

Richard Wollheim (1987) begins by noticing cases in which one's experience of a differentiated flat surface (a muddy wall or a frosty windowpane) involves two distinct aspects:

> (1) a *configurational* aspect, thanks to which one is visually aware (in a manner that is mostly veridical as far as it goes) of the surface itself and its variations in local color; and
>
> (2) a *recognitional* aspect, thanks to which one is visually aware of various robustly three-dimensional things, things that are not and are not believed to be before one's eyes at the time of the experience (battling horsemen, dancers in gauzy dresses).

These two awarenesses are distinguishable but inseparable aspects of a single experience, an experience of *seeing-in: seeing* the relevant three-dimensional things *in* the relevant surface. The configurational aspect can be described on analogy with a veridical simple seeing of a differentiated surface, which it resembles both intrinsically and in its characteristic causal-psychological role. The recognitional aspect can be described on analogy with a face-to-face seeing *of* the things one in fact merely sees *in* the surface. However, one can be aware of a differentiated surface in the particular manner exhibited here only by using the surface to discern absent three-dimensional things, and one can be aware of discerned absent things in the particular manner exhibited here only by being aware of a differentiated surface whose features enable one to discern them in it. In at least this respect, (1) and (2) are inseparable aspects of a single experience. And although they can be described on analogy with the simpler experiences just mentioned, there is a sense in which a detailed point-for-point comparison between them and such simpler experiences is out of the question: seeing-in and the simpler experiences to which it is in various ways analogous are "phenomenologically incommensurate" (1987) Such, Wollheim thinks, is the *twofoldness* involved in seeing-in. A painting depicts a given subject matter when one is inferably meant to see that subject matter in its surface and can indeed do so.

Michael Podro (1998) takes over from Wollheim's early *Art and Its Objects* (1980) the suggestion that a pictorial representation proposes a kind of simile or figurative comparison whose terms are the marked surface D and the subject O. And he adopts from I. A. Richards an interactionist view of figuration, on which any really deep comparison restructures one's thinking about both terms, reshaping one's thoughts about each on the model of one's thoughts about the other.

On the recognitional side of things, Podro insists that for depiction to occur, it is not enough that one's inspection of D activates one's capacity to recognize O in O's acknowledged absence; one must exploit one's recognition of X in a sustained, successful effort to visualize O. On the configurational side, he insists that one's awareness of D is never simply an awareness of how D is differentiated (lighter here, darker there; redder here, greener there); instead, it is framed in terms of how one takes the artist to have made her marks and handled her medium. There are at least two departures from Wollheim here. There is now a difference in kind between the configurational aspect of seeing a subject in a picture and the configurational aspect of seeing a dancer in a frosty windowpane. And configurational awareness is no longer largely veridical; the impressions a painter's marks generate about the manner of their making may be as designedly fanciful as the impressions a dancer's movements generate about the manner of their making. Configurational and recognitional awareness restructure each other repeatedly as one searches O (the represented subject) for real or merely fancied counterparts of what one has already discerned in D (the way the surface has been worked) and vice versa.

If Wollheim views depiction as one of several modes of pictorial meaning, Kendall Walton (1990) views it as lying at the heart of one of several related forms of make-believe pervading the cultural lives of children and adults alike. A *game of make-believe* is a form of individual or collective imaginative activity in which what players are

to imagine comes under the sway of rules or norms of a certain special kind: given the rules of the game in question, what they are to imagine about themselves, the things around them, and reality at large (what is *fictional*) becomes a fixed function of what is actually and discernibly the case about them, the things around them, and reality at large (what is discernibly true). The rules of such games may be maddeningly difficult to state, yet people seem awfully good at playing them and awfully invested in doing so.

Walton proposes in effect that a depiction D of an object O is a prop in a game of make-believe whose role in the game to which it belongs has the following features:

(1) The player is to look at and thereby come to see the object D

(2) He is to imagine about his act of looking at D that it is instead an act of looking at O, and about his resulting experience of seeing D that it is instead an experience of seeing O

(3) He is to manage the foregoing lookings, seeings, and imaginings in such a manner that he imagines *looking at and thereby coming to see* O—and imagines it both (a) vividly and (b) from the inside

(4) The game leaves him free to look at D in any of a wide range of ways, tending to result in a correspondingly wide range of experiences of seeing D

(5) How he is to imagine himself looking at O depends in a richly detailed manner on how he actually ends up looking at D, and the nature and content of the experience of seeing O he is to imagine having as a result depends in a richly detailed manner on the nature and content of the experience of seeing D he actually ends up having

Even the most naturalistic images continue to function as Gombrichian substitutes. When such a game is played, the called-for imaginings are such that they could not take place in the absence of the called-for perceivings, since they are about those perceivings, take those perceivings as their objects. It is equally true that the called-for perceivings could not take place in the absence of the called-for imaginings, since the perceivings involved in the execution of any demanding exploratory project are colored by it, owe their phenomenal character to it, and contain thoughts specifying its goals. This suffices to account for our sense that a spectator's visual experience of a picture is a unitary experience with two different kinds of subject matter: the depicted thing on the one hand, the depiction itself on the other.

THE FUTURE OF DEPICTION THEORY

In the last years of the twentieth century, the perceptual hypothesis account of vision favored by Gombrich lost ground to the modular computational account advocated by David Marr (1982). Many now regard vision as the computation of an accurate spatial model of one's immediate physical surroundings, from raw data about intensity distributions across the visual field, via a fixed set of speedy unconscious algorithms that provably deliver the goods in all but a special and statistically rare set of working conditions—algorithms having no access to the higher recognitional capacities of the person who steers through the world with their help. The various mathematical representations computationalists must appeal to in dividing the task of vision into manageable subtasks turn out to bear striking structural affinities to various familiar kinds of picture. The work of Michael Baxandall (1995), John Willats (1997), and Patrick Maynard (2005) constitute the beginnings of an effort to make principled sense of the whole range of psychologically natural pictorial idioms (and the uses and limitations of each) in a manner informed by emerging computational accounts of vision. What impact these emerging accounts of how pictures differ will eventually have on the best philosophical accounts of how pictures are alike it is too soon to tell.

See also Aesthetics, History of; Aesthetics, Problems of; Art, Authenticity in; Descartes, René; Goodman, Nelson; Helmholtz, Hermann Ludwig von; Plato; Popper, Karl Raimund; Wollheim, Richard.

Bibliography

Baxandall, Michael. *Shadows and Enlightenment*. New Haven, CT: Yale University Press, 1995.

Budd, Malcolm. "How Pictures Look." In *Virtues and Taste: Essays on Politics, Ethics, and Aesthetics*, edited by Dudley Knowles and John Skorupski, 154–175. Oxford, U.K.: Blackwell, 1993.

Descartes, René. "Optics." In *The Philosophical Writings of Descartes*, Vol. 1. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, 152–176. Cambridge, U.K.: Cambridge University Press, 1985.

Feagin, Susan. "Presentation and Representation." *Journal of Aesthetics and Art Criticism*56 (1998): 234–240.

Gibson, J. J. "The Information Available in Pictures." *Leonardo* 4 (1971): 27–35.

Gombrich, E. H. *Art and Illusion: A Study of the Psychology of Pictorial Representation*. 2nd ed. Princeton, NJ: Princeton University Press, 1961.

Gombrich, E. H. "Illusion and Art." In *Illusion in Nature and Art*, edited by R. L. Gregory and E. H. Gombrich, 192–243. New York: Scribner's, 1973.

Goodman, Nelson. *Languages of Art*. 2nd ed. Indianapolis, IN: Hackett, 1976.

Halliwell, Stephen. *The Aesthetics of Mimesis: Ancient Texts and Modern Problems*. Princeton, NJ: Princeton University Press, 2002.
Harrison, Andrew. "A Minimal Syntax for the Pictorial." In *The Language of Art History*, edited by Salim Kemal and Ivan Gaskell, 213–239. New York: Cambridge University Press, 1991.
Hopkins, Robert. *Picture, Image and Experience*. Cambridge, U.K.: Cambridge University Press, 1998.
Isenberg, Arnold. "Perception, Meaning, and the Subject Matter of Art." In *Aesthetics and the Theory of Criticism*, 36–52. Chicago: University of Chicago Press, 1973.
Kennedy, John M. *A Psychology of Picture Perception*. San Francisco: Jossey-Bass, 1974.
Lopes, Dominic. *Understanding Pictures*. Oxford, U.K.: Clarendon Press, 1996.
Marr, David. *Vision*. New York: Freeman, 1982.
Maynard, Patrick. "Depiction, Vision, and Convention." *American Philosophical Quarterly* 9 (1972): 243–250.
Maynard, Patrick. *Drawing Distinctions*. Ithaca, NY: Cornell University Press, 2005.
Merleau-Ponty, Maurice. *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, edited by Galen A. Johnson. Evanston, IL: Northwestern University Press, 1993.
Peacocke, Christopher. "Depiction." *Philosophical Review* 96 (1987): 383–410.
Plato. *Republic*. Translated by G. M. A. Grube. Revised by C. D. C. Reeve. Indianapolis, IN: Hackett, 1992.
Podro, Michael. *Depiction*. New Haven, CT: Yale University Press, 1998.
Schier, Flint. *Deeper into Pictures: An Essay on Pictorial Representation*. New York: Cambridge University Press, 1986.
Walton, Kendall. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press, 1990.
Willats, John. *Art and Representation: New Principles in the Analysis of Pictures*. Princeton, NJ: Princeton University Press, 1997.
Wollheim, Richard. *Art and Its Objects*. 2nd ed. New York: Cambridge University Press, 1980.
Wollheim, Richard. *Painting as an Art*. Princeton, NJ: Princeton University Press, 1987.

David Hills (2005)

ART, STYLE AND GENRE IN

Style and genre are two distinct but related ways in which artworks can be grouped together in the interests of understanding and appreciation. Neither mode of classification is easy to characterize, and much of the philosophical discussion of both genre—predominantly by literary theorists—and style—predominantly by historians and philosophers of the visual arts—has been clarificatory in aim. In the case of genre, there is a tension between structural (e.g., ode, epic, and collage) and functional (e.g., tragedy, romance, and altarpiece) ways of categorizing artworks. But many genres seem to have more to do with subject matter (e.g., bildungsroman and still life)—at least in those art forms that are broadly representational. The diverse bases for generic classification of artworks are reflected in René Wellek and Austin Warren's proposed definition of a literary genre as "a grouping of literary works based, theoretically, upon both outer form (specific metre or structure) and also upon inner form (attitude, tone, purpose—more crudely, subject and audience)" (1949, p. 231).

A much discussed theme in contemporary discussions of literary genre is whether the latter is merely a taxonomic convenience, reflecting the classificatory interests of literary critics and historians, or whether it reflects real differences between works. Certainly, the ascription of a work to a genre sometimes plays a part in explaining puzzling features of that work. For example, the art historian Michael Baxandall (1985) accounts for puzzling features of a Renaissance painting in terms of its belonging to the genre altarpiece. This seems to require an objective basis for genre classification. One can explain features of a work by appeal to genre only if one takes the genre to which the work is ascribed to be causally implicated in its generation—presumably in virtue of the artist's creative activity being guided by a conception of relevant generic constraints.

While genre is predominantly a critic's term for which art historians sometimes have a use, style is traditionally a historian's categorization whose critical and appreciative relevance has been increasingly remarked. Originating etymologically in the Latin term for a writing instrument (*stilus*), and thus applying to styles of writing, the term came to prominence in the writings of eighteenth- and nineteenth-century German art historians such as Johann Joachim Winckelmann. Art historians seek to historically situate individual works in groupings that are open to analogous kinds of explanation and to explain how such works so grouped stand in historical relations to one another. Style serves the first purpose inasmuch as an artistic style is taken to be a manifest feature of works that provides evidence as to their provenance, and the second if one posits an internal or external dynamics to the development of style.

Perhaps the most enduring testament to this tradition in art history is Heinrich Wölfflin's (1950) account of how painting in the High Renaissance differs from Baroque art in its style of pictorial representation. He introduces certain binary distinctions that provide a framework within which one can define different ways in which one might articulate pictorial space. The most

influential of these distinctions is between linear and painterly modes of representation and is defined in terms of a number of interrelated factors such as the way in which form is articulated (through outlines of masses or interplay between masses), the qualities of things through which they are represented (shape or texture), the manner in which relationships between objects are conveyed (atomistically or holistically), and the faculties through which pictorial articulation is primarily grasped (understanding and sensation).

Once one thinks of artworks as being groupable in terms of their styles, it is natural to ask why this is so and why such styles change over time—why, for example, late Renaissance and Baroque paintings differ in the cited ways. Wölfflin (1950) himself posits an internal logic underlying the historical development of artistic styles. His distinctions are taken to capture the various representational possibilities permitted by the artistic medium, and communities of artists are seen as pursuing their artistic goals within this framework of possibilities, which has within it its own dynamic. Wölfflin's account is interestingly but controversially extended to nonrepresentational painting by Clement Greenberg (1962), who sees an oscillation between linearity and painterliness in the postimpressionist tradition. A related idea is found in Arthur Danto's (1964) conception of the "style matrix." Alois Riegl's notion of *Kunstwollen* also manifests a commitment to an internal dynamic in the development of artistic style. On the contrary, James S. Ackerman (1962), reacting against the whiff of stylistic determinism in such accounts, offers an individualistic model where artistic styles change as a result of the attempts of artists to overcome problems arising in the activity of painting. Ernst Gombrich (1968) offers a more materialistic but still broadly individualistic model in which stylistic change is fueled by technological innovation and guided by the social structure of the art world.

As in the case of genre, some question whether the stylistic classifications employed by art historians reflect independent realities of the sort that both call for and furnish explanations, or whether they are taxonomic devices that reflect the culturally inflected interests and purposes of historians. A related concern is that, to the extent that style categories are viewed taxonomically, they are of questionable relevance for one's critical and appreciative engagement with particular artworks. In an influential paper, Richard Wollheim (1979) argues that the concept of style plays two importantly different roles in discussion about visual art. Wollheim distinguishes between general and individual style. The former, which he subdivides into universal style, historical or period style, and school style, is indeed taxonomic in the manner just described. Individual style, however, is what is at issue when one talks of "the style of A" in reference to the work of a given painter A. Those painters whose works are objects of aesthetic interest have "a style of their own," which allows their works to be understood as expressive.

Furthermore, and crucially, individual style is to be understood in generative rather than taxonomical terms: a style description for a painter A picks out elements in A's work that depend on those "processes or operations characteristic of his acting as a painter" that Wollheim terms *style processes* (1979, p. 135). A style process is analyzable in terms of some subset of the pictorial resources available to a painter on which the painter is disposed to act in a rule-like manner. Individual style, so construed, has "psychological reality" in these dispositions of the artist and can be seen as causally operative in the production of the artist's works. Wollheim argues for this generative conception of individual style on the grounds that it is required to make sense of the role played by style descriptions in the explanation of the details of pictures and of the susceptibility of the works of a given artist to grouping in terms of a common style. While Wollheim explicitly restricts his account of individual style to painting, the notion has been extended to literary artworks and, by implication, to artworks in general in two articles by Jenefer Robinson (1984, 1985).

In extending Wollheim's analysis Robinson also insists, in line with an early paper by Kendall Walton (1979), that an artist's individual style is properly identified not with some set of manifest features of the products of the artist's "artistic acts" (Sircello 1975), but with features of those acts themselves, "[Pictorial] style ultimately cannot be defined as a list of pictorial elements but rather as a way of doing certain things, or manipulating pictorial elements" (Robinson 1981, p. 10). In the case of literary works, for example, the relevant artistic acts are "*describing* people, *portraying* landscape, *characterising* personal relationships, *manipulating* rhythms, *organising* patterns of imagery, and so forth" (Robinson 1984, p. 138). Furthermore, these ways of doing things, insofar as they constitute an artist's style, are taken to be characteristic expressions of the mind and personality that the artist appears to have. This agential conception of individual style accords with talk of style in nonartistic contexts and explains both the restriction of style predicates to human actions and their products and the explicitly expressive nature of many style predicates (e.g., a sentimental or witty style). As for those sets of elements

proper to a given art form that are cited as constitutive of style by those who take the latter to reside in the products of artistic activity, Robinson maintains that they are grouped together as stylistic elements only in virtue of being the elements that the artist characteristically uses in performing the relevant artistic acts in a distinctive way.

Part of the significance of ascriptions of individual style for the critical appreciation of particular works of art is said to be their bearing on the expressive and other meaningful properties rightly ascribable to a work. Wollheim (1980), refining a suggestion by Gombrich (1960), argues that it is only through one's grasp of an artist's style that one can determine the precise expressive significance of a configuration of elements in the artist's work, and Robinson (1981) suggests that the same holds for at least some representational and formal properties. But both Walton (1979) and Robinson (1985) insist that what is expressed through the style of a work is not determined by facts about the actual artist, but by facts about the mind or personality of what Walton (1979) terms the *apparent artist*—the mentality or personality the artist appears to have given the stylistic features of the work. However, as Walton recognizes, to ascribe such a role to the apparent artist rather than to the actual artist is far from unproblematic, since what one sees in a configuration of elements in an artistic manifold may reflect ulterior knowledge about the actual artist. One therefore stands in clear need of a principle to delimit when such knowledge is rightly brought to bear in determining the expressivity embodied in the stylistic properties of an artwork.

Even if one thinks of individual styles as ways of doing, it is still through manifest features of the products of those ways of doing that one is able to recognize artists' styles. One question to which one would therefore expect an answer from an adequate philosophical account of style is whether artistic styles admit a univocal characterization in terms of the kinds of manifest features or elements that enter into their expressions. Both Walton (1979) and Robinson (1981, 1984) insist that there can be no checklist of stylistically relevant elements for a given art, since what makes an element part of the expression of an individual style is that it has been used in a particular way in an artistic act that is characteristic of the artist. This allows for all manner of different elements to enter into the styles of different artists and explains why only some elements that figure in an artist's works are mentioned in a style description, why a given element may figure in the style description of one artist but not in that of another who works in the same medium, and why a given element may have different stylistic significance in the works of different artists (Robinson 1985).

Nelson Goodman (1978) also rejects any attempt to distinguish stylistic from nonstylistic elements in terms of such dichotomies as expressive versus nonexpressive, form versus content, intrinsic versus extrinsic, or "the 'how' versus the 'what' of what is said," and so on. A feature of style may be a feature of what a work says, what it expresses, or of its formal or configurational elements. But, according to Goodman, what makes any such feature stylistic is both its contribution to the symbolic functioning of the work and its linking works together in ways that serve to advance appreciation and understanding, "[T]he style consists of those features of the symbolic functioning of a work that are characteristic of author, period, place, or school" (p. 35).

This suggests a way of reconnecting the two conceptions of style distinguished by Wollheim (1979), since both general and individual style might be connected to symbolic functioning in this way. It also suggests how one might bring into dialogue the stylistic interests of the historian and the critic. Goodman (1978) argues that style, as he conceives it, is of interest not only to the historian, who seeks to correctly attribute a history of making to an artwork, but also to the critic, who can use the attributions of the historian to discover further and subtler shared elements of symbolic functioning in the resulting groupings of works. Walton (1979) and Robinson (1985), however, will insist that one's interest in the individual style of works necessarily refers one back to distinctive features of the artistic acts that seem to result in entities capable of so functioning and that the interest of critics, unlike the interest of historians, is in how things appear to have been made rather than in how they were made. The complex relationships between the stylistic concerns of the historian and of the critic have been commented on by Walton (1979) and discussed at some length by Robinson (1981).

See also Aesthetic Qualities; Art, Expression in; Art, Formalism in; Art, Interpretation of.

Bibliography

Ackerman, James S. "A Theory of Style." *Journal of Aesthetics and Art Criticism* 20 (1962): 227–237.

Baxandall, Michael. *Patterns of Intention: On the Historical Explanation of Pictures*. New Haven, CT: Yale University Press, 1985.

Carroll, Noel. "Danto, Style, and Intention." *Journal of Aesthetics and Art Criticism* 53 (3) (1995): 251–257.

Danto, Arthur. "The Artworld." *Journal of Philosophy* 61 (1964): 571–584.
Eck, Caroline van, James McAllister, and Renée van de Vall, eds. *The Question of Style in Philosophy and the Arts*. New York: Cambridge University Press, 1995.
Frye, Northrop. "Rhetorical Criticism: Theory of Genres." In *Anatomy of Criticism: Four Essays*, 243–326. Princeton, NJ: Princeton University Press, 1957.
Gombrich, Ernst. *Art and Illusion: A Study in the Psychology of Pictorial Representation*. New York: Pantheon, 1960.
Gombrich, Ernst. "Style." In *The International Encyclopedia of the Social Sciences*. Vol. 15, edited by David L. Sills, 352–361. New York: Macmillan, 1968.
Goodman, Nelson. *Ways of Worldmaking*. Indianapolis, IN: Hackett, 1978.
Greenberg, Clement. "After Abstract Expressionism." *Art International* 6 (8) (1962): 24–32.
Lang, Berel, ed. *The Concept of Style*. Philadelphia: University of Pennsylvania Press, 1979.
Robinson, Jenefer. "General and Individual Style in Literature." *Journal of Aesthetics and Art Criticism* 43 (1984): 147–158.
Robinson, Jenefer. "Style and Personality in the Literary Work." *Philosophical Review* 94 (2) (1985): 227–247.
Robinson, Jenefer. "Style and Significance in Art History and Art Criticism." *Journal of Aesthetics and Art Criticism* 40 (1) (1981): 5–14.
Sircello, Guy. *A New Theory of Beauty*. Princeton, NJ: Princeton University Press, 1975.
Walton, Kendall. "Style and the Products and Processes of Art." In *The Concept of Style*, edited by Berel Lang, 45–66. Philadelphia: University of Pennsylvania Press, 1979.
Wellek, René, and Austin Warren. *Theory of Literature*. New York: Harcourt, 1949.
Wölfflin, Heinrich. *Principles of Art History: The Problem of the Development of Style in Later Art*. New York: Dover, 1950.
Wollheim, Richard. *Art and Its Objects: With Six Supplementary Essays*. 2nd ed. New York: Cambridge University Press, 1980.
Wollheim, Richard. "Pictorial Style: Two Views." In *The Concept of Style*, edited by Berel Lang, 129–148. Philadelphia: University of Pennsylvania Press, 1979.
Wollheim, Richard. "Style in Painting." In *The Question of Style in Philosophy and the Arts*, edited by Caroline van Eck, James McAllister, and Renée van de Vall. New York: Cambridge University Press, 1995.

David Davies (2005)

ART, TRUTH IN

The question of artistic truth first arises with the ancients. In his *Republic*, Plato argues that fine art and poetry cannot impart truths because they do not give humankind access to the Forms. Just as a mirror can only deliver a reflection of the particulars that themselves are merely reflections of the Forms, so painting and poetry amount to little more than pale images of Platonic Ideas at a third remove. Aristotle, in contrast, defended poetry as a means of obtaining general knowledge about probable courses of human events. One could learn from *Antigone*, for example, the likely turn of affairs when two strong-willed and unbending people, each convinced that he or she is in the right, disagree on matters of principle.

Though this topic is usually referred to in terms of "artistic truth," it is more precisely a concern with knowledge and the question of whether one can derive knowledge—or, even more broadly, cognitive value, from artworks. Truth, of course, comes into the picture, since it is one of the criteria of knowledge. Plato maintained that poets, like Homer, had no knowledge to teach and for that reason should not be esteemed as the educators of the Greeks. Aristotle, on the other hand, argued that poetry, especially tragedy, is akin to philosophy, since it has general truths about life to convey, namely universals about what is necessary or probable in the run of human events.

Throughout most of Western history, the view that art contributes to knowledge held sway. However, with the great advances of modern science and the empiricist philosophies that accompanied it, art began to look as though it had comparatively little, if anything, to offer by way of knowledge. Art, indeed, began to be treated by positivist philosophers as a primarily noncognitive enterprise.

Two types of arguments have been raised in order to challenge the cognitive credentials of art. The first group of arguments can be called epistemic. These allege that artworks cannot educate audiences because what artworks have to offer is not knowledge, properly so called; art is epistemically defective in various ways. The second group of arguments can be called aesthetic. They contend that it is inappropriate to expect artworks to function as sources of knowledge, even if for centuries in Western culture and others, art has been an object of respect for this very service.

EPISTEMIC ARGUMENTS

The epistemic arguments against the cognitive pretensions of art include: the banality argument, the no-evidence argument, and the no-argument argument. The banality argument takes a close look at the kinds of theses for which artworks are so often commended for teaching to their audiences. These are often truisms of the order of "crime doesn't pay" or "the prejudice of first impressions can be misleading." If this is knowledge, the skeptic observes, then it is nevertheless hardly something that we are taught by novels like *Crime and Punishment* or *Pride and Prejudice*. Rather, in order to understand such novels, we probably already need to have some version of

these commonplaces in our cognitive stock. According to the banality argument, there may be truths embodied in artworks, but they are rather paltry, bland, and known by nearly everyone before they encounter the artworks in question. So we cannot be said to learn them from artworks. Indeed, having access to these bromides is often a condition for comprehending the very artworks that contain them. But in any event, these truisms are in no way as revelatory as scientific discoveries are; if they constitute knowledge at all, it is common knowledge.

Whereas the banality argument concedes that there may be knowledge, albeit of a threadbare sort, to be had from artworks, the next two arguments deny this possibility. Of course, one can derive beliefs from artworks; but skeptics charge that it is impossible to gain knowledge from artworks. For knowledge involves not only beliefs, or even true beliefs; those beliefs must also be based upon something—either evidence or argument. And artworks, as a matter of form, it is charged, typically lack these sorts of accompanying justifications.

The no-evidence argument shows the influence of empiricism most clearly. Since Aristotle, it has been claimed that artworks, notably literature, give us knowledge of general truths concerning human life. But, the skeptic retorts, most artworks trade in particulars and one cannot justify a general claim on the basis of a single case. It is not adequate evidence, even if the case is as arresting as that of Antigone versus Creon. Moreover, a very great many of the case studies that are supposed to carry these generalizations about human life are fictional. No claim, general or otherwise, can be supported by a made-up story. Furthermore, most of these fabricated stories are invented precisely to corroborate the point the author wishes to promulgate. So not only is the evidence insufficient; it looks like it is tainted to boot. Thus, the skeptic surmises, artworks cannot be said to afford the kind of general knowledge for which they are so often applauded just because they are evidentially defective.

Of course, not every general claim needs to be supported by empirical evidence. Many philosophical generalizations are not. And the knowledge in which much art is said to traffic is philosophical, concerned, for example, with issues like free will. Since no amount of evidence is going to sway the free will debate one way or the other, the fact that artists do not back up their perspectives on free will with empirical evidence makes them no worse off than philosophers. The no-evidence argument, that is, does not cut against philosophical artworks.

But, the skeptic replies, genuine philosophical theses, even if unaccompanied by a body of empirical evidence, are nevertheless advanced by means of argument and/or analysis. Yet that is something that artworks characteristically have not got. *Nausea* may assert that humans are free; the novel may even be said to illustrate the point. But there is *no argument* to that conclusion in the book. However, if there is no argument, then there is no philosophical knowledge to be had from the text. At best there is unsubstantiated belief.

Nor, the skeptic adds, do commentators on artworks—including even those commentators who speak as though artworks are involved in making philosophical knowledge claims—argue about the truth or falsity of the cognitive theses they excavate from artworks. This lack of concern with argumentation by critics, then, is thought to lend additional credence to the skeptical view that art is not in the knowledge business. If it were, there would be more explicit argumentation both inside the artworks and in the critical estate that surrounds them. The lack of argumentation implies that art is not about securing knowledge, and, be that as it may, sans argumentation it does not do so anyway.

AESTHETIC ARGUMENTS

The epistemic arguments against art propose that what artworks deliver is not worth being called knowledge—it is either too trivial or it is unjustified. As a matter of fact, art just is not a suitable vehicle for the communication of anything robust enough and defensible enough to be counted as knowledge. But another set of arguments worries that knowledge is just the wrong thing to expect from artworks. Even if some artworks could convey knowledge, knowledge is never something we should legitimately expect from artworks. These arguments may be regarded as aesthetic, rather than epistemic, in nature. Three of them are: the common denominator argument, the no-expertise argument, and the mistaken-belief argument.

The common denominator argument points out that even if some artworks appear to provide knowledge—as *Moby Dick* does concerning whaling—many other artworks, like a great many string quartets, do not. Therefore, the expectation that artworks afford knowledge or even that they suggest knowledge claims does not apply to all artworks. Knowledge is not a generic criterion of artistic excellence. Yet if something is a criterion of artistic excellence, it must be relevant to the evaluation of every artwork. Knowledge is not. Consequently, knowledge is an inappropriate expectation to bring to an artwork qua art.

Artists study their craft and the materials that comprise their art form. Painters learn perspective, poets mas-

ter prosody, musicians scales, and so on. Their expertise is with the tools of their trade. They are not psychologists or political scientists or sociologists. They have no special expertise that entitles them to float generalizations about human life. How would a studio-arts education prepare one to discourse on human affairs? This is one of the earliest charges lodged against the attempt to enlist art as a producer of knowledge. Perhaps Plato holds the copyright on the no-expertise argument; Socrates used it to demolish Ion and, by extension, Homer.

Another argument striving to demonstrate the irrelevance of the pursuit of knowledge by art stresses that many artworks have been committed to beliefs that we now regard as obsolete and mistaken, and yet we still esteem the works in question. Indeed, many classic artworks are committed to beliefs that are contradicted by the beliefs recommended by other classic artworks and still, despite these contradictions, we are happy to embrace works on both side of the debate (say free will versus determinism) as canonical. But, it is conjectured, this would not be possible if we thought that truth and knowledge were appropriate standards for art. In that case, artworks associated with false beliefs would have to be demoted in our estimation. That they are not implies that knowledge is not an appropriate concern when it comes to art.

RESPONDING TO THE SKEPTICS

These arguments against the cognitive status of art are longstanding and serious. However, they can also be challenged in various ways. As a group, the epistemic arguments presuppose that if art is cognitive, then it will transmit knowledge to its audiences and this knowledge will take the form of general truths that can be stated in propositional form. Consequently, commentators often seek to outflank the epistemic objections by refusing this presupposition and locating the cognitive contribution of art primarily elsewhere than in the presentation of innovative general truths that can be articulated in propositions.

There are a number of different—nonexclusive and nonexhaustive—alternative candidates here and each suggests a way in which art may be said to make a contribution to cognition, broadly construed. Against the banality argument, it may be said that though artworks often deal in commonplaces, these are commonplaces that we are apt to forget. The cognitive function of art in this regard is to recall to mind the kinds of truths—such as the dangers of indulging a hasty prejudice or refusing to bend when one right is on a collision course with another—that are well known but oft forgot. Artworks, like *Pride and Prejudice* and *Antigone*, are vivid reminders of what we already know, but that of which we are prone to lose sight.

Indeed, artworks—engaging as they do the senses, feelings, emotions, imaginations, and cognitions of their audiences—are especially efficacious instruments for educating peoples in the ethos of their culture, because by mobilizing so many powers of a person at once, artworks deeply embed the common knowledge of a society in its participants in a way that makes it readily accessible for retrieval and use. Arguably, the multidimensional address of the artwork suits it as a means for educating a populace in its ethos in a fashion unrivaled by any other mode of communication.

Epistemic arguments appear to suppose that the only relevant sort of knowledge is *knowing that*. But in addition to propositional knowledge, there is also knowledge by acquaintance. Thus, defenders of the educative power of art maintain that art can provide knowledge by affording the opportunity for audiences to learn about certain experiences from the inside—to acquire, perhaps by simulation or empathy in the process of watching a film, a sense of or a feel for what it would be like, for example, to be a slave.

Moreover, in addition to knowledge by acquaintance, there is also know-how. Artworks can contribute know-how in many ways. For example, many of our concepts of virtue, vice, and other character traits are rather abstract, as are our moral principles. In order to learn how to apply these extremely abstract concepts and maxims, we need practice. Artworks, especially fictions, can provide the opportunity to hone our powers of judgment by giving us particulars, often subtly drawn, that enable us to deepen our faculties of judgment and our skill in deploying them. That is, artworks may enhance cognition by putting it to work in assessing fictional characters and actions in terms of concepts and principles—moral and otherwise (e.g., psychological, political, social)—that we possess abstractly, but which we need to exercise concretely in order to acquire a genuine command over them. Furthermore, inasmuch as a refined sensitivity toward the relevant concepts, like true heroism, plays a role in eliciting appropriate emotional responses, artworks facilitate the education of the emotions.

In addition, artworks may serve the cognitive purpose of orientation; they may help us map our world. Novels present us with crystallizations of various character types—often newly emerging ones, like the radical empiricist in Turgenev's *Fathers and Sons* or the gallery of

social tendencies inventoried in Balzac's *Comedie Humaine*, or the eponymous Sammy in *What Makes Sammy Run?* These character profiles—assembling a significant constellation of attributes—operate like concepts, supplying us with recognizable paradigms of social types which may even help us to navigate everyday life. Such paradigms are not true or false in the manner of a proposition, but apt or fitting. Nevertheless, aptness is as indispensable to cognition as propositional truth. Indeed, Nelson Goodman claims that the ultimate value of art is that it supplies us with apt models of the world.

Furthermore, art can tutor perception. Landscape painting and portraits can teach us how to look at the world. And Goodman has stressed the way in which even abstract paintings exercise and expand the viewer's ability to make fine perceptual distinctions.

One way to deal with the epistemic arguments, then, is to outflank them. But they can also be tackled head-on. Against the no-evidence argument, it is important to remind the skeptic that not all artworks are fictions and therefore cannot be uniformly dismissed as being evidentially empty for that reason. Nor is it only nonfiction literature with which the skeptic must contend. There is also photography, nonfiction motion pictures, and much installation art.

Moreover, even fictions can contain evidence. Thus, there is no grounds for summarily rejecting all fiction as incapable of proffering propositional knowledge. Michael Crichton's novel, *State of Fear*, about environmentalism, includes argumentative theses replete with footnotes to substantiate its case. Whether or not Crichton's book is correct is one question. Nevertheless, it is clear that a novel like it could be written that might succeed in proposing a series of true propositions supported by the appropriate documentary apparatus. This conjecture seems unobjectionable, furthermore, since, though many critics have complained about the quality of *State of Fear*, no one has denied that it is a novel.

Moreover, skeptics are wrong to contend that critics do not initiate charges of falsity accompanied by arguments against fictions. Presently secular humanists in the United States are waging a campaign against horror fictions for fostering superstitious beliefs. Likewise one can bet money that commentators sympathetic to the environmental movement will meet Crichton with the kinds of arguments they would unfurl against any scientist or politician who impugned their theories.

But we need not resort to Crichton to bridle at the no-evidence argument. We need only point out that it sets the bar for communicating knowledge too high. No one denies that the journalism on the op-ed pages of newspapers can convey knowledge. But the beliefs advanced there typically come to us without the kind of evidence it would take to vindicate them in the highest courts of reason. Rather, the author leaves it to the reader to reflect upon her assertions, encouraging us to weigh them against our own experience and to search out further proof of their accuracy. Likewise it may be argued that artists generally play by comparable rules. A novel like *Bonfire of the Vanities* provides a sketch of the 1980s that we are invited to substantiate on our own. Hence, if the aforesaid journalist is allowed into the knowledge game, so should a certain kind of novelist be. Indeed, doesn't the communication of knowledge usually leave some of the work of corroboration up to the reader? Consequently, that artworks encourage readers to test the hypotheses they suggest in what Peter Kivy thinks of as the laboratories of their minds is not an epistemic deficit. It is a recurring feature of the communication of knowledge across the board.

Similar reservations can be brought to bear on the no-argument argument. Not all theses are defended by means of empirical evidence. Most philosophical claims are not. A leading form of argument in defense of philosophical conjectures is the thought experiment—characteristically a narrative fiction predicated upon engaging the mind of the listener in a process of reflective equilibrium leading to a certain conclusion. But if philosophers are entitled to thought experiments as a mode of argument and/or analysis, why should artists be denied equal logical rights?

Many artworks are narrative fictions. Some at least are arguably thought experiments designed to encourage the embrace of certain discoveries, such as insight into the true nature of courage or compassion. That is, artworks may not only enable us through practice to apply concepts with finesse; they may also invite reflection upon the grammar of the concept in question—either by foregrounding a heretofore unappreciated essential criterion of the concept or by reminding us of the kinds of considerations it pays to remember whenever applying the concept. That is, a narrative artwork, functioning as a thought experiment, can engage the mind of the audience in a process of reflective equilibrium that results in propositional knowledge concerning the concept under scrutiny. Moreover, where the artwork is operating as a thought experiment, it is not without argument; the thought experiment, rather, stages the argument in the minds of the audience.

The aesthetic arguments against artistic claims to knowledge are not more decisive than the epistemic ones. The common denominator argument correctly observes that not *all* artworks are such that it is appropriate to evaluate them in terms of the knowledge they impart. They are not all vehicles for communicating knowledge; that is not the kind of thing they all are. So if aesthetic evaluation is keyed to the kind of thing a work is, then knowledge is not the sort of thing to employ in the assessment of, for example, most string quartets.

This much is true. However, the aesthete's argument here is more ambitious; it is that knowledge is never an appropriate measure of an artwork. However, some artworks, given the kinds of things they are essentially, are justifiably expected to bequeath knowledge, even propositional knowledge, to their audiences. This is not only the case with certain nonfiction examples. For instance, realist novels are committed, in virtue of their genre, to the production of various insights including social, psychological and political ones. Fledgling realist authors are instructed to become astute observers just because they are expected to inform readers about psychology and social mores. Moreover, since that is the kind of thing a realist novel is—i.e. B; in effect, the genre to which it belongs—it follows that in such cases disclosing truths figure in artistic evaluation. That the expectation of knowledge in inapposite with respect to many genres does not entail that it is out of bounds for every genre. It is not true that a criterion of artistic excellence must apply globally. Many art forms and genres may possess local standards of excellence given the kinds of artworks they are—the realist novel being a case in point.

The realist novel also indicates what is wrong with the no-expertise argument. Some artists—like realist novelists—are expected to sharpen their powers of psychological and social observation as part of their job description. Furthermore, with many of the things that realist authors have expertise in isolating and explaining—such as the ways of the heart or the claims of social justice—it is not really clear who the better experts are. And, in any event, given the power of artworks to engage simultaneously the whole person—feeling, imagination, memory, perception, cognition, and so forth—it is not evident that there is any more effective way of instilling these truths in recipients than artworks.

Lastly, the mistaken-belief argument is a non-starter. To maintain that knowledge may be a virtue in artworks does not imply that it is the only virtue. Thus, it may be the case that works that contain mistaken, perhaps outmoded, beliefs nevertheless have other merits that dispose us to keep them in the canon. That is also why we may be happy to welcome classics that contradict each other into our pantheon. One virtue that they may possess is that they articulate compellingly the mistaken beliefs they uphold as a work from an archaic culture might. Here, the work gives us knowledge of the past, albeit inadvertently. But at the same time if the work is designed formally in such a way that its theses, however false, are given their best face, then we can appreciate the work aesthetically, despite its cognitive shortcomings. Thus, the fact that palpably false artworks continue to hold our interest does not show that truth and knowledge may not be pertinent to our respect for some other artworks. At best it shows that they are not our only desideratum.

See also Aesthetic Experience; Art, Value in.

Bibliography

Aristotle. *Poetics*. Translated by R. Janko. Indianapolis, IN: Hackett, 1984.

Beardsley, Monroe. *Aesthetics*. Indianapolis, IN: Hackett, 1981.

Carroll, Noël. "The Wheel of Virtue." *Journal of Aesthetics and Art Criticism* 60 (2002): 3–26.

Currie, Gregory. "The Moral Psychology of Fiction." *Australasian Journal of Philosophy* 73 (1995): 250–259.

Goodman, Nelson. *Languages of Art*. Indianapolis, IN: Bobbs-Merrill, 1968.

Hospers, John. *Meaning and Truth in the Arts*. Chapel Hill: University of North Carolina Press, 1949.

Johns, Eileen. "Reading Fiction and Conceptual Knowledge." *The Journal of Aesthetics and Art Criticism* 56 (1998): 331–348.

Kivy, Peter. "On the Banality of Literary Truths." *Philosophic Exchange* 28 (1997–1998): 16–27.

Lamarque, Peter, and Olsen Stein. *Truth, Fiction, and Literature*. Oxford: Clarendon, 1994.

Plato. *Republic*. In *The Dialogues of Plato*. Translated by Benjamin Jowett. Oxford: Oxford University Press, 1953.

Noël Carroll (2005)

ART, VALUE IN

The question of the nature of art, what art is, has been much more widely discussed in philosophy than the question of its value, why art matters. The two issues cannot be completely disentangled, of course, since any account of what makes art art will inevitably isolate features of special importance. In fact, all the main theories of the nature of art have an implicit explanation of its value, but since the question of the value of art has social as well as philosophical significance, it is useful to make

these implicit accounts explicit, and thus expose them to critical scrutiny.

Four lines of thought have emerged as the principal ways in which philosophers and artists have explained the importance of art. These can be given convenient labels: hedonism, aestheticism, expressionism, and cognitivism. Briefly, the first holds that art is valuable for the pleasure derived from it; the second that art is valuable as a source of beauty; the third that art is valuable as a vehicle for expressing emotion; and the fourth that art is a source of knowledge and understanding equivalent to, but distinct from, science and philosophy.

Abstractly stated in this way, it unclear whether any of these theories construe the relationship between art and its value as intrinsic or instrumental. That is to say, is the value in question to be found in art itself, or is art simply a means to it? We might also wonder whether the value resides in the properties of art objects—books, sculptures, paintings, compositions, and so on—or in the experiences that these things give rise to in those who look or listen. These are issues that can be examined at a general level, but in fact the distinctions that they invoke—intrinsic/instrumental and object/experience—are more important in some explanations of the value of art than in others.

HEDONISM

The contention that art is valuable for the pleasure we derive from it is both longstanding and widespread. Indeed, most people, including those engaged in the arts, probably assume its truth without question. Yet as an explanation of the value and importance of art, it faces several difficulties. Before these can be considered directly, one point of clarification is required.

It is natural for people to describe their engagement with the arts in terms of enjoyment, and to express their artistic judgments in terms of liking and disliking. One result is that positive responses to art are usually construed as expressions of enjoyment obtained from encountering things we like. This then leads to the assumption that a favorable view of an artwork is an expression of pleasure. But in fact the conflation of pleasure and enjoyment is a mistake. Enjoyment can arise from other things besides pleasure. While the enjoyment of a good wine or a fine meal is largely, and sometimes exclusively, the result of gastronomic pleasure, a scientific lecture or a television documentary can be enjoyable for their intellectual content. They provide us with interesting material to think about, rather than a pleasurable experience to savor.

It might be replied that intellectual stimulation is a special kind of pleasure. The danger with this response is that it simply collapses the valuable into the pleasurable, and thus converts a substantial claim—that art is valuable *because* it is pleasurable—into an uninformative analytic claim—that to say art is valuable is *the same* as saying that it is pleasurable. In this way the claim about value and pleasure becomes true by definition. It follows that if hedonism about art is to be a substantial theory, we need to distinguish between enjoying something and getting pleasure from it. The fact that we derive pleasure from something is *one* reason for enjoying it and finding it valuable. But it is not the only possible reason.

In the light of this clarification we can now state the three main questions facing hedonism about art. First, is it generally true, as a matter of empirical fact, that the arts generate pleasurable experiences? Second, is it possible to discriminate between major and minor works of art in terms of pleasure? Third, if art is valuable because of the pleasure it gives us, would not other, better sources of pleasure make art redundant?

The first of these questions is a factual matter about which we have to be open-minded. Since probably the majority of people who philosophize about the value and importance of art are themselves art lovers, there is a tendency to assume that art does generally give pleasure. But the statistics of people attending classical concerts, reading serious literature, and making visits to art galleries do not bear this out. Considered solely in terms of the pleasure they give, soap operas, pop music, television shoot-'em-ups, and romantic pulp fiction almost certainly top grand opera, classical music, Shakespeare, and nineteenth-century Russian novels. Indeed, the position of the arts is worse than this. Far more people are bored by Shakespeare than are entertained by him, and to those same people, two hours of Bach or Beethoven is probably a dreadful prospect. Even artworks expressly created for entertaining can, with the passage to time, cease to provide much in the way of pleasure. For example, compared with modern television comedies like *Friends* or *Blackadder*, Restoration comedy is a very poor source of amusement.

The examples chosen to make this point can also be used to elaborate the second of the two difficulties outlined above. An enthusiast for classical music might insist that the principal value of concert going is indeed the pleasure we derive from it. While it is true that tastes differ, this pleasure, for those who find concert going pleasurable, is just as great or even greater than the pleasure of pop music, chiefly because high-quality music gives

pleasure repeatedly. A similar claim might be made on the part of all the arts, but even if we concede that the arts give great pleasure to those who like them, this does not give us any reason to rank them higher than more mundane sources of pleasure, like crossword puzzles, jigsaw puzzles, or board games.

This issue was expressly addressed by the utilitarian philosopher Jeremy Bentham (1748–1832), who was willing to argue that since pleasure is the ultimate source of all value, pushpin (a kind of board game) is as good as poetry. What Bentham did not observe, however, is that subscribing to hedonism raises not only a question about the comparative value of the aesthetic and the nonaesthetic but also a problem within the realm of the aesthetic itself. If the value of an artwork is derived from the pleasure it gives, then major works of art must give more pleasure than minor ones. But have we any reason to believe this? Can pleasure be correlated with estimations of aesthetic merit? Is a piece of music by a major composer like Beethoven, for example, guaranteed to give more pleasure than one by a minor composer like Luigi Boccherini?

This raises a contentious philosophical topic: whether pleasure can be measured or not. But even if it can, it would be difficult to show that the relative amounts of pleasure given by different works of art can be mapped onto the relative artistic merits customarily accorded to them. One suggested solution to this second difficulty is to be found in Bentham's utilitarian successor, John Stuart Mill (1806–1873), who tried to draw a distinction between higher and lower pleasures, a distinction that might be called upon to distinguish the relative merits of poetry over those of pushpin and the merits of Beethoven over those of Boccherini.

How are we to differentiate between higher and lower pleasures? According to Mill, this can only be done in terms of quantity or quality. The former, he thought, will not serve the purpose, since if the only difference between higher and lower pleasures is quantity, then any lower pleasure can equal the value of a higher pleasure provided there is more of it. Lots of pushpin will be equivalent to a little poetry. So the difference between the two must be qualitative. How is this difference in quality to be assessed? Mill's answer is that we should entrust the assessment to a competent judge, defined as someone who has experience of both the pleasures in question. The problem with this proposal is that the deliverances of such a judge cannot in principle be distinguished from mere preferences. Perhaps someone who declares opera to be a higher pleasure than soap opera does indeed detect differing qualities of pleasure arising from each of them. But it could be that there is no more to this "judgment" than a personal preference for opera. And we have no way of telling which is the case.

In any event, there is a further difficulty. If the value of art lies in the pleasure we get from it, and if, as seems obviously true, there are other good sources of pleasure, sports for example, it follows that a world without art would be no worse off than a world with art, provided that it had other sources from which equally pleasurable experiences could be generated. This objection relates to an important distinction drawn at the start, the instrumental versus the intrinsic. Hedonism attributes instrumental value, rather than intrinsic value, to art, and thereby implies that art has no value in and of itself. It is chiefly on this point that an alternative theory of the value of art, aestheticism, is built.

AESTHETICISM

The slogan "Art for art's sake" is a familiar one, and it is intended to capture the thought that art has value that cannot be accessed or realized in any other way. What could this intrinsic value be? One obvious contender is beauty. Since ancient times it has been believed that an important function of the arts is to make beautiful things—paintings, poems, music, buildings, and so on—and that these are to be savored for their beauty alone. Aestheticism holds that, though beautiful things are indeed pleasing, it is in their beauty, and not in the pleasure they give us, that their value lies. Since this beauty is an intrinsic property of the object, it cannot be replaced or substituted for without loss, as the extrinsic effect of giving us pleasure can be.

Now while it is undoubtedly true that many artworks are very beautiful, and valued in large part for this reason, it does not seem plausible to make beauty the ultimate explanation of their value, for two reasons. First, beauty is to be found elsewhere than in art. Second, not all art is, or aims to be, beautiful.

The first point is established by the existence of natural beauty. From the time of the ancient Greeks, human bodies and faces have not only been admired for their beauty, but regarded as templates and standards by which the beauty of pictures and statues is to be measured. Since the eighteenth century, landscapes, skies, and seascapes have also been held up as striking instances of the beautiful. All of these things are natural, not manufactured, and are therefore not works of art (the issue of divine creation aside). But if beauty is all around us in natural forms, a world without art would not be a world without beauty,

and while this fact does not detract from artistic beauty, it does mean that beauty does not give art any special claim to value.

In any case, while some artworks are indeed beautiful, not all are. For some works of art, in fact, the concept of beauty seems hard to apply. There are beautiful speeches in Shakespeare's tragedies, but could *Lear* or *Othello* be called "beautiful" as a whole? Moreover, even in the visual arts and in music, many widely acclaimed works seem expressly to eschew beauty. In Picasso's famous painting *Les demoiselles d'Avignon*, the figures and faces are ugly—deliberately so, it seems. Many modern composers have written music that is harsh and disjointed rather than harmonious and melodic. The pre-Raphaelite painters and Romantic composers of the nineteenth century strove for visual and aural beauty, but the movements that followed them in the twentieth century strove equally vigorously to avoid it. In short, exclusive focus on beauty can at best explain the value only of some art, and even then not uniquely so.

EXPRESSIONISM

These two objections to aestheticism are overcome in a third theory: that the value of art lies in its being an expression of emotion. The difference between natural beauty and beauty in art is that the former is not an expression of anything, whereas the latter is. It is the expression of the artist's emotion or feeling. Conversely, though emotion can be expressed through beauty, it can be expressed in other ways too. What enables us to classify Titian and Picasso, Schubert and Schoenberg together under the label "art" is that these radically different styles are all equally modes of expression.

Expressionism as an explanation of the value of art is almost as widely held as hedonism. Among its best known advocates were the Italian philosopher Benedetto Croce (1866–1952) and the great Russian novelist Leo Tolstoy (1828–1910). But on closer inspection, it too encounters great difficulties. Three are specially important: Whose emotion is it that an artwork expresses? Why is the expression of emotion a good thing? What place does expressionism leave for imagination?

It might seem obvious to answer the first of these questions by saying that the emotion expressed is that of the creator (the author, painter, composer, and so on). But suppose we say of *Othello*, for example, that it is a dramatic expression of jealousy. What reason have we to say that it is *Shakespeare's* jealousy that is expressed? Since we know hardly anything about Shakespeare, still less about the circumstances in which he came to write this play, we have no reason to say this. Something similar is true of a huge number of artworks. We do not know much, if anything, about the psychological or emotional history of their creators, and so we cannot say whether they ever felt the emotions expressed in their works.

An alternative would be to locate the emotion in the audience. Aristotle thought that dramatic works are "cathartic." That is to say, they become the vehicles by which audiences give vent to emotions that are often debilitating when discharged into ordinary life. His examples are fear and pity. By discharging these emotions on imaginary objects, we are less burdened by them in the business of day-to-day living. Aristotle only applied the theory of catharsis to drama, and it is unclear whether it could equally be applied to all the arts. But even if it can, there is this further question: What is so good about the expression of emotion as such? Imagine that a work enables those who watch, listen, or otherwise contemplate it, to give vent to ethnic feeling. Without the work, their racist emotions would never have had such clear definition or powerful expression. But why should that commend the work to us? It seems most plausible to hold that it is the powerful expression of *good* emotions that is to be valued, not the powerful expression of emotion per se. On the contrary, hurtful or destructive emotions ought not to be given powerful expression.

To identify the emotion expressed in a work of art as the audience's, then, carries no positive value; it could as easily be negative. To attribute it to the author means, in a very large number of cases, making unwarranted assumptions about the artist's psychological biography. But a further objection is that, by insisting that the origins of an artwork must lie in its creator's personal history, we seem thereby to deny any influence to the very faculty that seems central to artistic creativity, namely imagination. The great genius of such a major work of art as the novel *Middlemarch* lies in George Eliot's ability to rise above the confines of personal experience and imagine a world of people and events that the author never encountered. The most fundamental objection to expressionism is that it reduces acts of imagination to acts of reporting and recording.

COGNITIVISM

Some exponents of an expressionist account of art, notably R. G. Collingwood (1889–1943), have seen this difficulty and, in their efforts to avoid it, have effectively shifted the center of attention from feeling and expression to imagination and understanding. If there is any value in works that express or depict racist or other neg-

ative feelings, it lies not in the expression itself, but in the extent to which it gives us insight into the minds of those who have such feelings. This idea motivates the move from expressionism to cognitivism, the view that art should be valued as a form of understanding. On this view, the powerful expression of emotion we find in *Othello*, for example, should be valued for enabling us to understand and appreciate the mind of the intensely jealous. From this point of view, the play supplies its audience with material for thought rather than feeling, and it is of no consequence whether Shakespeare ever felt any of the same sort of rage as the character he invented. Indeed, it adds to the critical assessment of the play if he never did, since in that case the play stands as an even more impressive act of imagination.

Aesthetic cognitivism thus overcomes the most important objections to expressionism by construing artworks as acts of imagination rather than autobiography, and valuing them as such. By shifting the focus to imagination, it also circumvents some of the objections brought against aestheticism and hedonism. The products of the imagination can be beautiful, but this is not what makes them works of art. Artworks stand in contrast to natural beauty because natural beauty is not the outcome of imagination. The relative merits of major and minor works lie in the degree to which the understanding they offer us is more or less profound, and this is a judgment quite independent of the pleasure we do or do not derive from them. Relatively shallow works can be attractive and pleasing; much more profound ones rather taxing.

And yet aesthetic cognitivism faces difficulties of its own. First among these is the relation between imagination and reality. If we are to say that works of art enhance our understanding, this implies that they give us insight into the realities of human experience. But how can they do this if the people, places, and events that they depict are all products of the imagination? Must not understanding track how things really are, rather than how someone has imagined them to be? Second, while aesthetic cognitivism may seem plausible with respect to representative art, it seems more implausible when applied to abstract art. Great novels, films, and figurative paintings are easily thought of as giving us insight into life, but how can this be said of abstract painting, instrumental music, or architecture?

These are important questions, and it is by no means clear that they can be answered. But even if they can, there is a further issue. Does cognitivism about the arts not lead to their redundancy somewhat as hedonism does? G. W. F. Hegel (1770–1831) was perhaps the greatest philosopher whose account of art can be broadly described as cognitivist, and he quite explicitly thought that because its value lies in its contribution to the development of human understanding, art must eventually be replaced by philosophy. Even at its best, we might say, art can only gesture toward the sort of understanding that philosophy makes explicit.

The most promising reply to this anxiety lies in stressing the sensuous nature of art, which enables it to enhance our felt experience. An artwork does not tell us about the nature of things, events, or people by formulating general statements about them. Rather, it depicts what has been called a "concrete universal," products of the imagination that give us a sense of what it is like to be present and to undergo the experience of things, people, and events from a particular perspective. In other words, art illuminates the things around us rather than providing us with information about them in the way that science, history, and philosophy do.

Whether this response is ultimately satisfactory is a large topic, but cognitivism's emphasis on the sensuous and on the imagination has the merit of being true to two central aspects of the arts. One further implication of cognitivism is that if the sensual is essential, the late-twentieth-century movement known as *conceptual* art may signal an acknowledgment of the end of art. This is an implication that some philosophers, notably Arthur C. Danto, have endorsed and even welcomed.

See also Aesthetic Experience; Aesthetic Qualities; Art, Interpretation of; Art, Truth in; Beauty; Ugliness.

Bibliography

Aristotle. *Poetics*. Translated by Malcolm Heath. London: Penguin, 1997.

Bell, Clive. *Art*. Oxford, U.K.: Oxford University Press, 1987.

Budd, Malcolm. *Values of Art*. London: Penguin, 1995.

Collingwood, R. G. *The Principles of Art*. Oxford, U.K.: Clarendon Press, 1938.

Croce, Benedetto. *Æsthetic as Science of Expression and General Linguistic*. Translated by Douglas Ainslie. London: Peter Owen, 1953.

Danto, Arthur C. *After the End of Art: Contemporary Art and the Pale of History*. Princeton, NJ: Princeton University Press, 1997.

Gadamer, Hans-Georg. *"The Relevance of the Beautiful" and Other Essays*. Translated by Nicholas Walker. Cambridge, U.K.: Cambridge University Press, 1986.

Goodman, Nelson. *Ways of Worldmaking*. Indianapolis, IN: Hackett. 1978.

Hegel, G. W. F. *Aesthetics*. Translated by T. M. Knox. Oxford, U.K.: Clarendon Press, 1975.

Hume, David. *Essays, Moral, Political, and Literary*. Oxford, U.K.: Oxford University Press, 1963.
Kant, Immanuel. *Critique of Judgment*. Translated by Werner S. Pluhar. Indianapolis, IN: Hackett, 1987.
Langer, Susanne K. *Feeling and Form*. London: Routledge and Kegan Paul, 1953.
Levinson, Jerrold. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996.
Mill, John Stuart. *Utilitarianism*. London: Fontana, 1985.
Nietzsche, Friedrich Wilhelm. *The Birth of Tragedy*. Translated by Shaun Whiteside. London: Penguin, 1993.
Schaper, Eva, ed. *Pleasure, Preference, and Value: Studies in Philosophical Aesthetics*. Cambridge, U.K.: Cambridge University Press, 1983.
Scruton, Roger. *Art and Imagination*. London: Methuen, 1974.
Tolstoy, Leo. *What Is Art?* Translated by Aylmer Maude. Oxford, U.K.: World Classics, 1955.
Walton, Kendall L. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press, 1990.

Gordon Graham (2005)

ARTIFICIAL AND NATURAL LANGUAGES

The only natural languages we know of are human. In addition to such human languages as English, Spanish, Russian, and Chinese, with which we are all familiar, there are many less well-known languages, many of them spoken by hundreds of people. The more marginalized languages are dying out at an alarming rate. Owing to lack of evidence, our information about their origin is limited, but it seems likely that they evolved out of communication systems similar to those used by animals for communication. Living human languages are learned as first languages by infants and are used for face-to-face communication and many other purposes.

Natural languages are influenced by a mixture of unconscious evolutionary factors and conscious innovation and policy making. In most cases, the historical record does not allow us to tell what role these factors played in the development of a given feature, but the difficulty of consciously controlling the language used by a large population suggests that unconscious causes predominate.

The term "artificial language" is often used for humanlike languages that are created either for amusement (like J. R. R. Tolkien's Elvish) or for some practical purpose (Esperanto). Information on such projects can be found in Alan Libert's work (2000).

Artificial languages of a quite different sort are created for scientific and technological reasons, and the design of such languages is closely connected with logical theory. Logic originated with Aristotle in his *Prior Analytics*. Although Aristotle's syllogistic theory used symbols for terms (such as "some," "all," "not") that make up propositions, such symbols and the expressions made up out of them were not generally considered as part of a linguistic system until much later.

Modern logical theory and its connection with artificial languages owes much to the search for a universal language in the seventeenth century (Maat 1999). In Britain, George Dalgarno (1968 [1661]) and John Wilkins (2002 [1668]) promoted the idea of a philosophical language based on rational principles. In retrospect, their ideas seem to be more closely aligned with the goal of designing an improved human language than with the mainstream development of logic and were more concerned to facilitate clear expression of ideas than to serve as a framework for developing a theory of reasoning. Their projects stressed the need for basing a vocabulary on a rational ontology and are more closely connected with later attempts to develop taxonomies and thesauri than with logic per se.

At about the same time, however, G. W. Leibniz attempted to develop a "universal characteristic" based on several ideas central to the later development of logic and artificial-language design. In his "Dissertatio de Arte Combinatoria" (excerpts in Loemker 1956, pp. 117–133), written in 1666 when he was nineteen years old, Leibniz presents a logical program that, in its main proposals, informed his philosophy for the remainder of his life.

Like Dalgarno and Wilkins, Leibniz adopted the goal of a rationally ideal philosophical language, but he differs from them in the stress he lays on reasoning and in the degree to which his account of reasoning is inspired by mathematics. The leading ideas of his program—that truth can be discovered by analysis, or division of concepts into basic constituents; that such analytic reasoning is analogous to combinatory reasoning in mathematics; and that it is facilitated by a language with a clear syntactic structure reflecting the meanings of expressions—have furnished important insights for subsequent work in logic. The stress that Leibniz placed on calculation as part of the reasoning process gives him a well-deserved central place in the history of logic and computation.

The two weak points in Leibniz's program are (1) the assumption that once analysis was achieved in an ideally rational language, testing a proposition for truth should be a relatively trivial matter, and (2) the idea that analysis is appropriate and possible across the entire range of rational inquiry. The first of these weaknesses was cor-

rected late in the nineteenth century, when Gottlob Frege developed a symbolic language for the representation of "pure" or mathematical thought. Frege's "Begriffsschrift," or conceptual calculus, achieves the goal prefigured by Leibniz of a language designed to facilitate reasoning by allowing the relations between concepts to be clearly and unambiguously displayed. And it conforms to the methodological ideal of being completely explicit more than any previous attempt to present an artificial language. Frege's presentation of the *Begriffsschrift* makes it possible to test each constellation of symbols to tell whether it is a well-formed formula (an expression that conforms to the syntactic rules of the system). Although part of proving such a formula in Frege's calculus is a matter of analysis, or the application of explicit definitions, the result of such analysis is a formula that must be proved using logical laws. These laws are explicitly formulated, so that it is also possible to tell whether or not a purported proof conforms to the rules of the system. But *whether there is a proof* of an analyzed proposition need not be a question that can be solved algorithmically. In fact, as the theory of the nature of reasoning systems has shown, we cannot in general expect to have an algorithmic criterion for whether a formula is provable.

The second weakness in Leibniz's program is more difficult to deal with decisively. But many years of experience indicate that we have no reliable methodology for isolating universal atoms of human thought. In many extended attempts to make the rules of reasoning in some domain explicit, it seems more useful to deal with many primitives that are conceptually related by axioms rather than by definitions.

Alonzo Church summarized the results of more than seventy-five years of philosophical and mathematical development of Frege's achievement in section 7, "The Logistic Method," of his *Introduction to Mathematical Logic* (1956). In that and the subsequent two sections, Church sets out the methods logicians had established in the first half of the twentieth century for constructing artificial logical languages (or, to use the usual current term, *formal languages*) and theorizing about them. These methods have changed slightly in the subsequent forty-eight years, the most significant changes having to do with interest in applications other than the explication of deductive reasoning and in the widespread use of formal languages in digital computing. In the beginning of the twenty-first century, it is not essential for formal languages to have a deductive component, and in some cases it may be important to associate implemented computational procedures with a formal language.

What are the essential features of a formal language? First, a formal language must have a *syntax*, a precise definition not only of the vocabulary of the language but also of the strings of vocabulary items that count as well-formed formulas. If other types of complex expressions than formulas are important, for each such *syntactic type* there must be a precise definition of the set of strings belonging to that type. These definitions must be not only precise but *effective*; that is, questions concerning membership in syntactic types must be algorithmically decidable. These syntactic definitions are usually presented as inductive definitions; for instance, the simplest formulas are defined directly, and rules are presented for building up complex formulas from simpler ones. The set of well-formed formulas is not only decidable but usually belongs to a known restricted class of efficiently computable sets of strings. The *context-free* sets of strings are heavily used in computational applications, and are also capable of standing in for large parts of human languages.

Second, if proofs are associated with the language, these too must be precisely defined. Whether or not a list of formulas is a proof must be algorithmically decidable.

Third, the formal language must have a semantic interpretation, which associates *semantic values* or *denotations* with the well-formed expressions of the language. The importance of a semantic component was recognized by Alfred Tarski, who also provided a methodology for placing semantics on a sound mathematical basis and applying it to the analysis of mathematical theories.

A version of Tarskian semantics due to Alonzo Church (1940) starts with a domain of individuals (the objects that the language deals with) and a domain of truth-values (the two values True and False) and constructs possible denotations by taking functions from domains to domains. Sentences, for instance, denote truth-values, and one-place predicates (verblike expressions taking just one argument) denote functions from individuals to truth-values.

In a semantics for deductive reasoning, truth-values are essential. Once the legitimate interpretations (or *models*) of a language are given, the validity of an inference (say of formula B from formula A) can be defined as follows: The inference is valid if every model that assigns A the value True also assigns B the value True.

The theory of any language (natural or artificial) has to be stated in some language. When one language serves as a vehicle for formulating and theorizing about another language, the first is called the *metalanguage* for the sec-

ond, and the second is called an *object language* of the first. Nothing prevents a metalanguage itself from being formalized. When logicians wish to investigate theories of language, they may wish to formalize an object language and its metalanguage. The language in which the theory of both languages is stated would be a *meta-metalanguage*. Since formalization is a human endeavor, the whole enterprise is usually conducted in some human language (typically in some fairly regimented part of a human language, supplemented with mathematical notation), and this language serves as the metalanguage for all the languages developed in the course of the formalization project. In theory, a language can be its own metalanguage, but in such cases we have a situation that can easily lead to paradox.

The use of digital computers has led to the wholesale creation of special-purpose formal languages. Since computer scientists have borrowed the methods for presenting these languages from logic, computational formal languages usually conform to Church's recipe. Sometimes, however, a semantics is not provided. (For instance, mathematical tools for providing semantic interpretations for programming languages only became available years after such languages had been developed and used.) Often it is important to specify the crucial computational procedures associated with such a language. For example, a *query language*, intended to enable a user to present questions to a database, has to provide a procedure for computing an answer to each query that it allows. Sometimes a computational formal language is pointless unless procedures have been implemented to enable computers to process inputs formulated in the language. A programming language is useless without an implemented program that interprets it; a markup language like HTML (Hypertext Markup Language) is useless without browsers that implement procedures for displaying documents written in the language.

These are very natural additions to Church's logistic method. Even in 1956 a semantic interpretation was thought to be desirable but not essential. The methods developed by logicians in the first half of the twentieth century for formalizing languages have not changed greatly since then and are likely to be with us for a long time.

The distinction between natural and formal languages is not the same as the distinction between naturally occurring and artificial languages. Rather, it is the distinction between naturally occurring languages and languages that are formalized, or precisely characterized along the lines suggested by Church. As far as the distinction goes, what prevents a natural language from being formalized is the difficulty (or perhaps impossibility) of actually formalizing a language like English or Swahili. Can natural languages be formalized? Can the grammar of naturally occurring languages be articulated as clearly as the syntax of an artificially constructed language? In assigning denotations to the expressions of a natural language, do we encounter problems that do not arise with artificial languages designed to capture mathematical reasoning?

In fact, there are difficulties. But logical work on formal languages has served as one of the most important sources of inspiration for theories of natural-language syntax, and is by far the most important source of inspiration for semantic theories of natural language. Both types of theories are now primarily pursued by linguists.

The ideal of syntax stated by Church derives from earlier work by David Hilbert, Rudolf Carnap, and other logicians. The essential ideas are an utterly precise description of the syntactic patterns of a language and algorithmic rules specifying how complex expressions are built up out of simpler ones. In essentials, this ideal is also the one that Noam Chomsky proposed in 1957 for the syntax of natural language. It has persisted through the evolution of the theories that Chomsky and his students have created and is also accepted by most of the leading rival approaches. Although there are methodological difficulties associated with the paradigm, they are no worse than the difficulties encountered by other sciences. The idea that natural-language syntax resembles that of formal languages has proved to be a fruitful paradigm for almost fifty years of syntactic research.

Semantics presents a more difficult challenge. Tarski's program addressed the semantics of specialized mathematical languages, and its success seems to depend essentially on certain features of these languages that are not shared by natural human languages: (1) Mathematical notation is designed to be neither ambiguous nor vague, whereas natural languages are both vague and ambiguous. (2) Natural languages have many sorts of *indexical* or context-sensitive expressions, like "I" and "today," whereas mathematical notations tend to use only one kind of indexical expression, variables. (3) *Intensional* constructions like "believe" are not found in mathematics, and they create other difficulties. The verb "believe" does not act semantically on the truth-value of the sentence it modifies. If you know that "Sacramento is the capitol of California" is true, this does not tell you whether "Jack believes that Sacramento is the capitol of California" is true. There are practical difficulties as well

as difficulties in principle. Natural languages are so complex that the task of formalizing them is open-ended and much too large for a single linguist or even for a single generation of linguists.

Richard Montague, a logician who taught at the University of California at Los Angeles until 1971, is primarily responsible for showing how to overcome obstacles that seemed to prevent a semantics for natural languages along the lines advocated by Tarski. His work began a program of research along these lines that is still being pursued. Montague's solution to the problem of ambiguity was to assign denotations to *disambiguated syntactic structures*. With a syntactic structure and a single reading for each word in a sentence, the sentence can have only one meaning. His solution to indexicality was to relativize interpretations to contexts. And his solution to the problem of intensionality, which followed earlier work by Rudolf Carnap, was systematically to assign linguistic phrases two denotations: an *intension* and an *extension*. Montague treated possible worlds as semantic primitives. Intensions, for him, were functions from possible worlds to appropriate extensions. The intension of a sentence, for instance, is a function from possible worlds to truth-values. Montague presented several formal "fragments" of English, the idea being to achieve rigor by focusing on a limited family of natural-language constructions. He also showed how to use higher-order logic to obtain a remarkably elegant and unified semantic interpretation.

This work on natural-language semantics leaves open a number of challenging questions concerning whether natural languages contain elements that somehow resist formalization. For one, Montague did not deal with vagueness, and there are difficulties with his accounts of intensionality and indexicality. These issues have been a major preoccupation of analytic philosophy since the 1970s. Although no philosopher has persuasively argued that the problems are unsolvable, they are certainly more difficult than many people imagined them to be in 1971. While the final question of whether natural languages can be completely formalized remains open, the assumption that this is possible has certainly inspired a fruitful paradigm of research.

See also Semantics; Syntactical and Semantic Categories.

Bibliography

Carnap, Rudolph. *Logische Syntax der Sprache Schriften zur wissenschaftlichen Weltaufiassung*. Vienna: Verlag von Julius Springer, 1934.

Carnap, Rudolph. *Meaning and Necessity*. 2nd ed. Chicago: University of Chicago Press, 1956. First edition published in 1947.

Chomsky, Noam. *Syntactic Structures*. The Hague: Mouton, 1957.

Church, Alonzo. "A Formulation of the Simple Theory of Types." *Journal of Symbolic Logic* 5 (1) (1940): 56–68.

Church, Alonzo. *Introduction to Mathematical Logic*. Vol. 1. Princeton, NJ: Princeton University Press, 1956.

Dalgarno, George. *Ars Signorum, Vulgo Character Universalis et Lingua Philosophica* (1661). Menston, Yorkshire, U.K.: Scholar Press, 1968.

Frege, Gottlob. *Begriffsschrift: Eine der arithmetischen nachgebildete Formalsprache des reinen Denkens*. Halle, Germany: L. Nebert, 1879. Translated in *Frege and Gödel: Two Fundamental Texts in Mathematical Logic*, compiled by Jean van Heijenoort. Cambridge, MA: Harvard University Press, 1970.

Libert, Alan. *A Priori Artificial Languages*. Munich: Lincom Europa, 2000.

Loemker, Leroy E., ed. *Gottfried Wilhelm Leibniz: Philosophical Papers and Letters*. Vol. 2. Chicago: University of Chicago Press, 1956.

Maat, Jaap. "Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz." PhD diss., Institute for Logic, Language, and Computation, University of Amsterdam, Amsterdam, 1999.

Montague, Richard. *Formal Philosophy: Selected Papers of Richard Montague*, edited by Richmond H. Thomason. New Haven, CT: Yale University Press, 1974.

Tarski, Alfred. "The Concept of Truth in Formalized Languages" (1936). In his *Logic, Semantics, Metamathematics*. Oxford, U.K.: Clarendon Press, 1956.

Wilkins, John. *An Essay towards a Real Character and a Philosophical Language* (1668). Bristol, U.K.: Thoemmes, 2002.

Richmond H. Thomason (2005)

ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) tries to enable computers to do the things that minds can do. These things include seeing pathways, picking things up, learning categories from experience, and using emotions to schedule one's actions—which many animals can do, too. Thus, human intelligence is not the sole focus of AI. Even terrestrial psychology is not the sole focus, because some people use AI to explore the range of all possible minds.

There are four major AI methodologies: symbolic AI, connectionism, situated robotics, and evolutionary programming (Russell and Norvig 2003). AI artifacts are correspondingly varied. They include both programs (including neural networks) and robots, each of which may be either designed in detail or largely evolved. The field is closely related to artificial life (A-Life), which aims

to throw light on biology much as some AI aims to throw light on psychology.

AI researchers are inspired by two different intellectual motivations, and while some people have both, most favor one over the other. On the one hand, many AI researchers seek solutions to technological problems, not caring whether these resemble human (or animal) psychology. They often make use of ideas about how people do things. Programs designed to aid/replace human experts, for example, have been hugely influenced by knowledge engineering, in which programmers try to discover what, and how, human experts are thinking when they do the tasks being modeled. But if these technological AI workers can find a nonhuman method, or even a mere trick (a kludge) to increase the power of their program, they will gladly use it.

Technological AI has been hugely successful. It has entered administrative, financial, medical, and manufacturing practice at countless different points. It is largely invisible to the ordinary person, lying behind some deceptively simple human-computer interface or being hidden away inside a car or refrigerator. Many procedures taken for granted within current computer science were originated within AI (pattern-recognition and image-processing, for example).

On the other hand, AI researchers may have a scientific aim. They may want their programs or robots to help people understand how human (or animal) minds work. They may even ask how intelligence in general is possible, exploring the space of possible minds. The scientific approach—psychological AI—is the more relevant for philosophers (Boden 1990, Copeland 1993, Sloman 2002). It is also central to cognitive science, and to computationalism.

Considered as a whole, psychological AI has been less obviously successful than technological AI. This is partly because the tasks it tries to achieve are often more difficult. In addition, it is less clear—for philosophical as well as empirical reasons—what should be counted as success.

SYMBOLIC AI

Symbolic AI is also known as classical AI and as GOFAI—short for John Haugeland's label "Good Old-Fashioned AI" (1985). It models mental processes as the step-by-step information processing of digital computers. Thinking is seen as symbol-manipulation, as (formal) computation over (formal) representations. Some GOFAI programs are explicitly hierarchical, consisting of procedures and subroutines specified at different levels. These define a hierarchically structured search-space, which may be astronomical in size. Rules of thumb, or heuristics, are typically provided to guide the search—by excluding certain areas of possibility, and leading the program to focus on others. The earliest AI programs were like this, but the later methodology of object-oriented programming is similar.

Certain symbolic programs, namely production systems, are implicitly hierarchical. These consist of sets of logically separate if-then (condition-action) rules, or productions, defining what actions should be taken in response to specific conditions. An action or condition may be unitary or complex, in the latter case being defined by a conjunction of several mini-actions or mini-conditions. And a production may function wholly within computer memory (to set a goal, for instance, or to record a partial parsing) or outside it (via input/output devices such as cameras or keyboards).

Another symbolic technique, widely used in natural language processing (NLP) programs, involves augmented transition networks, or ATNs. These avoid explicit backtracking by using guidance at each decision-point to decide which question to ask and/or which path to take.

GOFAI methodology is used for developing a wide variety of language-using programs and problem-solvers. The more precisely and explicitly a problem-domain can be defined, the more likely it is that a symbolic program can be used to good effect. Often, folk-psychological categories and/or specific propositions are explicitly represented in the system. This type of AI, and the forms of computational psychology based on it, is defended by the philosopher Jerry Fodor (1988).

GOFAI models (whether technological or scientific) include robots, planning programs, theorem-provers, learning programs, question-answerers, data-mining systems, machine translators, expert systems of many different kinds, chess players, semantic networks, and analogy machines. In addition, a host of software agents—specialist mini-programs that can aid a human being to solve a problem—are implemented in this way. And an increasingly important area of research is distributed AI, in which cooperation occurs between many relatively simple individuals—which may be GOFAI agents (or neural-network units, or situated robots).

The symbolic approach is used also in modeling creativity in various domains (Boden 2004, Holland et al. 1986). These include musical composition and expressive performance, analogical thinking, line-drawing, painting,

architectural design, storytelling (rhetoric as well as plot), mathematics, and scientific discovery. In general, the relevant aesthetic/theoretical style must be specified clearly, so as to define a space of possibilities that can be fruitfully explored by the computer. To what extent the exploratory procedures can plausibly be seen as similar to those used by people varies from case to case.

CONNECTIONIST AI

Connectionist systems, which became widely visible in the mid-1980s, are different. They compute not by following step-by-step programs but by using large numbers of locally connected (associative) computational units, each one of which is simple. The processing is bottom-up rather than top-down.

Connectionism is sometimes said to be opposed to AI, although it has been part of AI since its beginnings in the 1940s (McCulloch and Pitts 1943, Pitts and McCulloch 1947). What connectionism is opposed to, rather, is symbolic AI. Yet even here, opposed is not quite the right word, since hybrid systems exist that combine both methodologies. Moreover, GOFAI devotees such as Fodor see connectionism as compatible with GOFAI, claiming that it concerns how symbolic computation can be implemented (Fodor and Pylyshyn 1988).

Two largely separate AI communities began to emerge in the late 1950s (Boden forthcoming). The symbolic school focused on logic and Turing-computation, whereas the connectionist school focused on associative, and often probabilistic, neural networks. (Most connectionist systems are connectionist virtual machines, implemented in von Neumann computers; only a few are built in dedicated connectionist hardware.) Many people remained sympathetic to both schools. But the two methodologies are so different in practice that most hands-on AI researchers use either one or the other.

There are different types of connectionist systems. Most philosophical interest, however, has focused on networks that do parallel distributed processing, or PDP (Clark 1989, Rumelhart and McClelland 1986). In essence, PDP systems are pattern recognizers. Unlike brittle GOFAI programs, which often produce nonsense if provided with incomplete or part-contradictory information, they show graceful degradation. That is, the input patterns can be recognized (up to a point) even if they are imperfect.

A PDP network is made up of subsymbolic units, whose semantic significance cannot easily be expressed in terms of familiar semantic content, still less propositions. (Some GOFAI programs employ subsymbolic units, but most do not.) That is, no single unit codes for a recognizable concept, such as dog or cat. These concepts are represented, rather, by the pattern of activity distributed over the entire network.

Because the representation is not stored in a single unit but is distributed over the whole network, PDP systems can tolerate imperfect data. (Some GOFAI systems can do so too, but only if the imperfections are specifically foreseen and provided for by the programmer.) Moreover, a single subsymbolic unit may mean one thing in one input-context and another in another. What the network as a whole can represent depends on what significance the designer has decided to assign to the input-units. For instance, some input-units are sensitive to light (or to coded information about light), others to sound, others to triads of phonological categories … and so on.

Most PDP systems can learn. In such cases, the weights on the links of PDP units in the hidden layer (between the input-layer and the output-layer) can be altered by experience, so that the network can learn a pattern merely by being shown many examples of it. (A GOFAI learning-program, in effect, has to be told what to look for beforehand, and how.) Broadly, the weight on an excitatory link is increased by every coactivation of the two units concerned: cells that fire together, wire together.

These two AI approaches have complementary strengths and weaknesses. For instance, symbolic AI is better at modeling hierarchy and strong constraints, whereas connectionism copes better with pattern recognition, especially if many conflicting—and perhaps incomplete—constraints are relevant. Despite having fervent philosophical champions on both sides, neither methodology is adequate for all of the tasks dealt with by AI scientists. Indeed, much research in connectionism has aimed to restore the lost logical strengths of GOFAI to neural networks—with only limited success by the beginning of the twenty-first century.

SITUATED ROBOTICS

Another, and more recently popular, AI methodology is situated robotics (Brooks 1991). Like connectionism, this was first explored in the 1950s. Situated robots are described by their designers as autonomous systems embedded in their environment (Heidegger is sometimes cited). Instead of planning their actions, as classical robots do, situated robots react directly to environmental cues. One might say that they are embodied production systems, whose if-then rules are engineered rather than programmed, and whose conditions lie in the external

environment, not inside computer memory. Although—unlike GOFAI robots—they contain no objective representations of the world, some of them do construct temporary, subject-centered (deictic) representations.

The main aim of situated roboticists in the mid-1980s, such as Rodney Brooks, was to solve/avoid the frame problem that had bedeviled GOFAI (Pylyshyn 1987). GOFAI planners and robots had to anticipate all possible contingencies, including the side effects of actions taken by the system itself, if they were not to be defeated by unexpected—perhaps seemingly irrelevant—events. This was one of the reasons given by Hubert Dreyfus (1992) in arguing that GOFAI could not possibly succeed: Intelligence, he said, is unformalizable. Several ways of implementing nonmonotonic logics in GOFAI were suggested, allowing a conclusion previously drawn by faultless reasoning to be negated by new evidence. But because the general nature of that new evidence had to be foreseen, the frame problem persisted.

Brooks argued that reasoning shouldn't be employed at all: the system should simply react appropriately, in a reflex fashion, to specific environmental cues. This, he said, is what insects do—and they are highly successful creatures. (Soon, situated robotics was being used, for instance, to model the six-legged movement of cockroaches.) Some people joked that AI stood for artificial insects, not artificial intelligence. But the joke carried a sting: Many argued that much human thinking needs objective representations, so the scope for situated robotics was strictly limited.

EVOLUTIONARY PROGRAMMING

In evolutionary programming, genetic algorithms (GAs) are used by a program to make random variations in its own rules. The initial rules, before evolution begins, either do not achieve the task in question or do so only inefficiently; sometimes, they are even chosen at random.

The variations allowed are broadly modeled on biological mutations and crossovers, although more unnatural types are sometimes employed. The most successful rules are automatically selected, and then varied again. This is more easily said than done: The breakthrough in GA methodology occurred when John Holland (1992) defined an automatic procedure for recognizing which rules, out of a large and simultaneously active set, were those most responsible for whatever level of success the evolving system had just achieved.

Selection is done by some specific fitness criterion, predefined in light of the task the programmer has in mind. Unlike GOFAI systems, a GA program contains no explicit representation of what it is required to do: its task is implicit in the fitness criterion. (Similarly, living things have evolved to do what they do without knowing what that is.) After many generations, the GA system may be well-adapted to its task. For certain types of tasks, it can even find the optimal solution.

This AI method is used to develop both symbolic and connectionist AI systems. And it is applied both to abstract problem-solving (mathematical optimization, for instance, or the synthesis of new pharmaceutical molecules) and to evolutionary robotics—wherein the brain and/or sensorimotor anatomy of robots evolve within a specific task-environment.

It is also used for artistic purposes, in the composition of music or the generation of new visual forms. In these cases, evolution is usually interactive. That is, the variation is done automatically but the selection is done by a human being—who does not need to (and usually could not) define, or even name, the aesthetic fitness criteria being applied.

ARTIFICIAL LIFE

AI is a close cousin of A-Life (Boden 1996). This is a form of mathematical biology, which employs computer simulation and situated robotics to study the emergence of complexity in self-organizing, self-reproducing, adaptive systems. (A caveat: much as some AI is purely technological in aim, so is some A-Life; the research of most interest to philosophers is the scientifically oriented type.)

The key concepts of A-Life date back to the early 1950s. They originated in theoretical work on self-organizing systems of various kinds, including diffusion equations and cellular automata (by Alan Turing and John von Neumann respectively), and in early self-equilibrating machines and situated robots (built by W. Ross Ashby and W. Grey Walter). But A-Life did not flourish until the late 1980s, when computing power at last sufficed to explore these theoretical ideas in practice.

Much A-Life work focuses on specific biological phenomena, such as flocking, cooperation in ant colonies, or morphogenesis—from cell-differentiation to the formation of leopard spots or tiger stripes. But A-Life also studies general principles of self-organization in biology: evolution and coevolution, reproduction, and metabolism. In addition, it explores the nature of life as such—life as it could be, not merely life as it is.

A-Life workers do not all use the same methodology, but they do eschew the top-down methods of GOFAI. Sit-

uated and evolutionary robotics, and GA-generated neural networks, too, are prominent approaches within the field. But not all A-Life systems are evolutionary. Some demonstrate how a small number of fixed, and simple, rules can lead to self-organization of an apparently complex kind.

Many A-Lifers take pains to distance themselves from AI. But besides their close historical connections, AI and A-Life are philosophically related in virtue of the linkage between life and mind. It is known that psychological properties arise in living things, and some people argue (or assume) that they can arise only in living things. Accordingly, the whole of AI could be regarded as a subarea of A-Life. Indeed, some people argue that success in AI (even in technological AI) must await, and build on, success in A-Life.

WHY AI IS A MISLEADING LABEL

Whichever of the two AI motivations—technological or psychological—is in question, the name of the field is misleading in three ways. First, the term *intelligence* is normally understood to cover only a subset of what AI workers are trying to do. Second, intelligence is often supposed to be distinct from emotion, so that AI is assumed to exclude work on that. And third, the name implies that a successful AI system would really be intelligent—a philosophically controversial claim that AI researchers do not have to endorse (though some do).

As for the first point, people do not normally regard vision or locomotion as examples of intelligence. Many people would say that speaking one's native language is not a case of intelligence either, except in comparison with nonhuman species; and common sense is sometimes contrasted with intelligence. The term is usually reserved for special cases of human thought that show exceptional creativity and subtlety, or which require many years of formal education. Medical diagnosis, scientific or legal reasoning, playing chess, and translating from one language to another are typically regarded as difficult, thus requiring intelligence. And these tasks were the main focus of research when AI began. Vision, for example, was assumed to be relatively straightforward—not least, because many nonhuman animals have it too. It gradually became clear, however, that everyday capacities such as vision and locomotion are vastly more complex than had been supposed. The early definition of AI as programming computers to do things that involve intelligence when done by people was recognized as misleading, and eventually dropped.

Similarly, intelligence is often opposed to emotion. Many people assume that AI could never model that. However, crude examples of such models existed in the early 1960s, and emotion was recognized by a high priest of AI, Herbert Simon, as being essential to any complex intelligence. Later, research in the computational philosophy (and modeling) of affect showed that emotions have evolved as scheduling mechanisms for systems with many different, and potentially conflicting, purposes (Minsky 1985, and Web site). When AI began, it was difficult enough to get a program to follow one goal (with its subgoals) intelligently—any more than that was essentially impossible. For this reason, among others, AI modeling of emotion was put on the back burner for about thirty years. By the 1990s, however, it had become a popular focus of AI research, and of neuroscience and philosophy too.

The third point raises the difficult question—which many AI practitioners leave open, or even ignore—of whether intentionality can properly be ascribed to any conceivable program/robot (Newell 1980, Dennett 1987, Harnad 1991).

AI AND INTENTIONALITY

Could some NLP programs really understand the sentences they parse and the words they translate? Or can a visuo-motor circuit evolved within a robot's neural-network brain truly be said to represent the environmental feature to which it responds? If a program, in practice, could pass the Turing Test, could it truly be said to think? More generally, does it even make sense to say that AI may one day achieve artificially produced (but nonetheless genuine) intelligence?

For the many people in the field who adopt some form of functionalism, the answer in each case is: In principle, yes. This applies for those who favor the physical symbol system hypothesis or intentional systems theory. Others adopt connectionist analyses of concepts, and of their development from nonconceptual content. Functionalism is criticized by many writers expert in neuroscience, who claim that its core thesis of multiple realizability is mistaken. Others criticize it at an even deeper level: a growing minority (especially in A-Life) reject neo-Cartesian approaches in favor of philosophies of embodiment, such as phenomenology or autopoiesis.

Part of the reason why such questions are so difficult is that philosophers disagree about what intentionality is, even in the human case. Practitioners of psychological AI generally believe that semantic content, or intentionality,

can be naturalized. But they differ about how this can be done.

For instance, a few practitioners of AI regard computation and intentionality as metaphysically inseparable (Smith 1996). Others ascribe meaning only to computations with certain causal consequences and provenance, or grounding. John Searle argues that AI cannot capture intentionality, because—at base—it is concerned with the formal manipulation of formal symbols. And for those who accept some form of evolutionary semantics, only evolutionary robots could embody meaning (Searle, 1980).

See also Computationalism; Machine Intelligence.

Bibliography

Boden, Margaret A. *The Creative Mind: Myths and Mechanisms.* 2nd ed. London: Routledge, 2004.

Boden, Margaret A. *Mind as Machine: A History of Cognitive Science*. Oxford: Oxford University Press, forthcoming. See especially chapters 4, 7.i, 10–13, and 14.

Boden, Margaret A., ed. *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press, 1990.

Boden, Margaret A., ed. *The Philosophy of Artificial Life*. Oxford: Oxford University Press, 1996.

Brooks, Rodney A. "Intelligence without Representation." *Artificial Intelligence* 47 (1991): 139–159.

Clark, Andy J. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1989.

Copeland, B. Jack. *Artificial Intelligence: A Philosophical Introduction*. Oxford: Blackwell, 1993.

Dennett, Daniel C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Dreyfus, Hubert L. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press, 1992.

Fodor, Jerome A., and Zenon W. Pylyshyn. "Connectionism and Cognitive Architecture: A Critical Analysis." *Cognition* 28 (1988): 3–71.

Harnad, Stevan. "Other Bodies, Other Minds: A Machine Incarnation of an Old Philosophical Problem." *Minds and Machines* 1 (1991): 43–54.

Haugeland, John. *Artificial Intelligence: The Very Idea*. Cambridge, MA: MIT Press, 1985.

Holland, John H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge, MA: MIT Press, 1992.

Holland, John H., Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard. *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, MA: MIT Press, 1986.

McCulloch, Warren S., and Walter H. Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." In *The Philosoophy of Artificial Intelligence*, edited by Margaret A. Boden. Oxford: Oxford University Press, 1990. First published in 1943.

Minsky, Marvin L. *The Emotion Machine*. Available from http://web.media.mit.edu/~minsky/E1/eb1.html. Web site only.

Minsky, Marvin L. *The Society of Mind*. New York: Simon & Schuster, 1985.

Newell, Allen. "Physical Symbol Systems." *Cognitive Science* 4 (1980): 135–183.

Pitts, Walter H., and Warren S. McCulloch. "How We Know Universals: The Perception of Auditory and Visual Forms." In *Embodiments of Mind*, edited by Warren S. McCulloch. Cambridge, MA: MIT Press, 1965. First published in 1947.

Pylyshyn, Zenon W. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex, 1987.

Rumelhart, David E., and James L. McClelland, eds. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* 2 vols. Cambridge, MA: MIT Press, 1986.

Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach.* 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2003.

Searle, John R. "Minds, Brains, and Programs," *The Behavioral and Brain Sciences* 3 (1980), 417–424. Reprinted in M. A. Boden, ed., *The Philosophy of Artificial Intelligence* (Oxford: Oxford University Press 1990), pp. 67–88.

Sloman, Aaron. "The Irrelevance of Turing Machines to Artificial Intelligence." In *Computationalism: New Directions*, edited by Matthias Scheutz. Cambridge, MA: MIT Press, 2002.

Smith, Brian C. *On the Origin of Objects*. Cambridge, MA: MIT Press, 1996.

Margaret A. Boden (1996, 2005)

ASCETICISM

There is a morbid fascination in any survey of the ascetic practices of humankind. Fasting, the virgin priestess, and the mutilation of the body are common features of ancient religions. In monastic Christianity the austere ideals of celibacy, obedience, and poverty have been both practiced and admired. Even today there are many who observe Lent and those for whom fasting and penance are seldom out of season. The most accomplished ascetics have been the wanderers (*sunnyasins*) of ancient India and the anchorites of fourth-century Egypt. One *sunnyasin* held his arms above his head with fists clenched until the muscles in his arms atrophied and the nails grew through his palms. It is said that the anchorite St. Simeon Stylites tied a rope tightly around himself until it ate into his body and his flesh became infested with worms. As the worms fell from his body he replaced them in his putrefied flesh, saying, "Eat what God has given you."

Behind such ascetic practices usually lies the philosophical theory of "asceticism," a theory that demands and justifies this unnatural way of life. Although the term *ascetic* was originally applied to any sort of moral disci-

pline, it has since acquired a narrower and more negative meaning. Asceticism may now be defined as the theory that one ought on principle to deny one's desires. Asceticism may be partial or complete. Partial asceticism is the theory that one ought to deny one's "lower desires," which are usually identified as sensuous, bodily, or worldly and are contrasted with more virtuous or spiritual desires. Complete asceticism is the theory that one ought to deny all desires without exception. Asceticism may also be moderate or extreme. Moderate asceticism is the theory that one ought to repress one's desires as far as is compatible with the necessities of this life. Extreme asceticism is the theory that one ought to annihilate one's desires totally.

HISTORY

The belief that austerities (*tapas*) burn away sin was a product of the non-Aryan tradition of ancient India. This belief persisted, and austerities were recommended by the yogis and the Jains. All orthodox systems of Indian philosophy agreed that the goal of life is liberation (*moksa*) from this world of suffering, and most maintained that the renunciation of worldly desires is necessary for liberation. Although the Buddha tried and rejected austerities, his principle that the cause of suffering is craving led later Buddhists to advocate renunciation and even to practice austerities. The Jains held that liberation is possible only when one has annihilated all passion, because passion attracts karma, believed by this sect to be a subtle form of matter that holds the soul in bondage.

Asceticism seems to have entered Western philosophy from the mystery religions that influenced Pythagoreanism about the end of the sixth century BCE. Although Greek ethics was predominantly naturalistic, Plato sometimes argued that one ought to repress the bodily desires in order to free the soul in its search for knowledge. Some Cynics renounced worldly desires in order to pursue virtue in independence. The early Stoics defined emotion as irrational desire and held up the ideal of the apathetic man in whom all emotions had been annihilated. Plotinus emphasized the ascetic side of Plato's philosophy and claimed that matter is the source of all evil.

This undercurrent of asceticism rose to the surface in medieval philosophy with its emphasis on religious otherworldliness. The foundations of this asceticism were laid by such theologians as St. Athanasius, St. Gregory of Nyssa, St. Ambrose, and even St. Augustine. They believed that the desires of the flesh should be repressed in order to achieve moral virtue and the contemplation of God. Their view molded the monastic institutions that were established in the fourth century. Virtually unchallenged, this asceticism remained a potent influence on religious life until the Renaissance.

Of modern philosophers, only Arthur Schopenhauer has been an important advocate of asceticism; he would have one completely annihilate the will to live in all its manifestations. Jeremy Bentham and Friedrich Nietzsche have each criticized asceticism from very different standpoints.

ARGUMENTS FOR ASCETICISM

The arguments for asceticism fall into three main classes. First, there are those that attempt a direct justification of self-denial. Although some of these arguments might justify a complete asceticism, they have traditionally been used to support only a partial asceticism. (1) We know by some authority that one ought to deny one's lower desires. One authority is the Bible, in which we find both express ascetic commandments and examples like those of the Virgin Mary and the celibate Christ. (2) The sacrament of penance requires the denial of worldly desires. Although one is cleansed of original sin by baptism, subsequent sins must be expiated by penance; the best way to make penance more than a formal ritual is to express repentance in a life of self-denial. (3) By undergoing the suffering of self-denial, one is taking up the cross of Christ. Since Jesus came into this world as a model for all men, all men ought to share in his redemptive suffering. (4) People ought to deny their lower desires to prove their virtue, for the ascetic life is a test of devotion to God, and those who pass the test will win a heavenly reward. (5) The suffering of self-denial is required by our guilt. Since every person has sinned, the retributive theory of punishment requires that every person suffer. By inflicting pain upon oneself, one balances the scales of justice and lifts the guilt from one's soul. (6) Self-denial is valuable because it develops certain character traits such as persistence and self-discipline, which are essential to living well.

The second class of arguments attempts to justify denial indirectly through a criticism of the lower desires. Since these criticisms are aimed only at certain desires, they can support only a partial asceticism. (1) The lower desires cost too much to satisfy. Gratification must be purchased with great effort, and perhaps these desires are insatiable, so that no expenditure of effort will gratify them. (2) The lower desires are misguided, for their objects are really evils or, at best, indifferent things. In either case, no genuine value is realized by fulfilling one's

desires. (3) Although the objects of the lower desires are good, they are much less good than higher values like virtue, knowledge, or heaven. Since an individual's time and energy are limited, one ought not to allow these lower desires to distract from the pursuit of what really matters. (4) The lower desires are intrinsically evil. Since they turn people away from God and his commands toward earthly objects, they are infected with the sin of pride. (5) Although not sinful in themselves, the lower desires do motivate one to sinful actions. Thus greed may tempt a person to steal, and lust can lead to adultery. (6) These lower desires interfere with the pursuit of knowledge, which is essential for the good life. They interfere either by causing an agitation that destroys one's power of reasoning or by fixing one's attention on sensory objects that distract from the transcendent reality.

The third class of arguments also attempts to justify asceticism indirectly through a criticism of desire per se. Since these arguments are aimed at all desires, they support a complete asceticism. (1) Schopenhauer argued that desire, by its very nature, can yield nothing but suffering. Desire springs from a lack and consists in a dissatisfaction. When it meets with hindrances, it produces nothing but frustration, because it cannot attain its object; when it does attain its object, it produces nothing but boredom, because desire ceases with fulfillment and leaves one with an undesired object. Since desire necessarily involves dissatisfaction, frustration, and boredom, the only escape is by the annihilation of all desire. (2) The Buddhists and the Jains maintain that one ought to annihilate desire in order to achieve liberation from this world of pain. A person must destroy all desire because desire is the cause of rebirth into this world. For the Buddhist, desire causes rebirth because, being selfish, it causes selfish actions; these, by the moral law of karma, cause rebirth in painful forms. For the Jain, desire magnetizes the soul so that it attracts karmic matter which, by the physical laws of mechanics, weighs down the soul and causes it to be reborn into this lower world of pain.

ARGUMENTS AGAINST ASCETICISM

It is much harder to classify the traditional arguments against asceticism. Many of them attack some presupposition of the doctrine. (1) Many, but not all, forms of asceticism require a dualism of mind and body. The various philosophical difficulties with metaphysical dualism therefore tend to undercut asceticism. (2) Ascetic practices are often recommended as a means of freeing the soul from the body so that it can contemplate the truth. Actually these practices make knowledge in all its forms impossible because self-denial produces frustration, uneasiness, and pain, which make clear thinking difficult, and self-mutilation destroys the bodily health that is the physiological basis of thought. (3) Asceticism usually assumes that desires are like little animals inside the self that grow when they are fed and wither when they are starved. Freudian psychology, however, reveals that one does not destroy a desire by suppressing it but that the desire continues to exist and to exert itself in new and usually devious ways. Hence ascetic practices may not be an effective means of annihilating or even of controlling desire. (4) Ascetic practices are often thought to be a means to, and even a guarantee of, moral goodness, but in fact they are no protection against vice. The ascetic may become complacent in his confidence in his ascetic practices; he may become proud of his ascetic achievements; and he may even despise others who are less accomplished in asceticism. (5) The religious arguments for asceticism frequently assume that God requires one to renounce available goods and even to inflict harm upon himself, but this is inconsistent with the benevolence of God. (6) There is also a religious argument against the view that bodily desires or worldly objects are essentially evil. Both this world and human nature must be good, because they are creatures of a Creator who is perfectly good.

Another group of arguments is pragmatic in nature. (1) As Bentham pointed out, asceticism cannot be consistently practiced because it runs counter to the basic motives in human nature. Since the function of morality is to guide conduct, asceticism is incapable of becoming a genuine moral standard. (2) To the limited extent that asceticism can be put into practice, its effects are harmful. It obviously increases the amount of suffering in the world. If Freudian psychology is correct, its attempt to suppress natural desires will result in various neuroses. Finally, it stultifies vitality, produces emotional excesses, and fosters the weakling at the expense of the strong man.

Then there are those arguments that attempt to refute asceticism by showing that it has unacceptable implications. (1) Asceticism condemns worldly concerns and natural impulses. This implies that one ought to abandon all social ties and mortify all family affection, which would be immoral. (2) If it is good for one to suffer, it should be better for everyone to suffer. This implies that a person has a duty to inflict pain on others, but not even the hardened ascetic will accept this. (3) If pleasure is really bad, it would seem that pain must be good. This implies the absurd conclusion that the best of all possible

worlds would be the one with the least pleasure and the most pain.

Finally, there is Nietzsche's ad hominem argument. Those who are incapable of living well disguise their impotence and fear by inverting morality in order to excuse their own moral sickness and to restrain the strong men who appear dangerous. Although the ascetic priest condemns the will to power, he uses ascetic ideals as a means of maintaining his own power over the sick herd. Thus an analysis of the psychological origin of asceticism reveals that it is far from a worthy ideal.

Asceticism is the doctrine that one ought to deny his desires. In practice, denial means refraining from the fulfillment of desires and sometimes mortifying the desire by inflicting upon oneself the very opposite of what is desired. This involves abstinence from genuine goods, the frustration of unfulfilled desires, and even self-inflicted pain. Unless one is prepared to accept the view that abstinence, frustration, and pain are intrinsically good, the ascetic life can hardly be defended as an end in itself.

If ascetic practices are to be recommended, they must be a necessary evil, a means to something better. One might regard the ascetic life as a means to liberation from this world of suffering. It would be unrealistic to deny that we all suffer from time to time and that there are those for whom life is mostly suffering. It would be equally unrealistic, however, to deny that for most of us the evils we experience are more than balanced by the genuine values we enjoy. Granted the existence of evil, the obvious expedient is to improve our world rather than to make it even worse by adding the sufferings involved in ascetic practices. If escape were desirable, there is no guarantee that the ascetic life would actually lead to freedom.

One might advocate the ascetic life as a means of pleasing God and winning the eternal bliss of heaven. Asceticism seems most plausible within a religious context. But are its theological presuppositions themselves plausible? Is there really an immortal soul to be rewarded or a God to do the rewarding? Even the believer may reject asceticism on religious grounds. A benevolent deity would hardly have created us with natural desires and then commanded us to deny these very desires and to suffer the consequent evils of frustration and pain.

The ascetic life might be urged as a means to that knowledge which in turn brings the good life. Ascetic practices are supposed to help by freeing the soul from the body. Still, no empiricist would admit that the body, which is the source of all experience, is a hindrance to knowledge, and even a rationalist like Plato concedes that experience reminds reason of the truth. Unless reason is thought of as a disembodied spirit—in which case it is hard to see how the body hinders reason in the first place—it would seem that ascetic practices make one less, rather than more, capable of the clear and sustained reasoning that is required for attaining knowledge.

Finally, the ascetic life might be advanced as a means to virtue. It must be admitted that some desires sometimes cause one to act wickedly, but these same desires also cause one to act virtuously. The sexual desire that can lead to adultery more often leads to conjugal fidelity. Hence there is a double error in regarding sexual desire as evil. It does not always, or even usually, express itself in sinful action; and if adultery is a sin, that is because it does violence to the institution of marriage, which is itself an expression of sex. As this example shows, natural desires are in themselves morally neutral, and to deny desire is to forbid the virtuous act as well as the sin. Instead of being a means to virtue, self-denial is actually a vice. Virtue requires at least prudence and benevolence, but the ascetic is imprudent in abstaining from available goods and in even inflicting harm upon himself. By concentrating on the cultivation of his own soul through suffering, the ascetic tends to become callous toward the suffering of others and to ignore his obligation to work for their welfare.

The ascetic life is not good in itself, nor is it a means to liberation, divine reward, knowledge, or virtue. It does not follow that one must accept the advice of Callicles to attempt gratification of every desire without regard for temperance or justice. Self-discipline is a genuine virtue, but it denies desire only when this is necessary to achieve an inclusive and harmonious satisfaction. Asceticism goes beyond this point to advocate an unnecessary and pointless denial. The logical conclusion is that asceticism should be rejected.

See also Augustine, St.; Bentham, Jeremy; Buddhism; Christianity; Cynics; Gregory of Nyssa; Karma; Liberation in Indian Philosophy; Nietzsche, Friedrich; Pain; Plato; Plotinus; Punishment; Pythagoras and Pythagoreanism; Renaissance; Schopenhauer, Arthur; Stoicism.

Bibliography

Ambrose, St. "On Virgins." In *Ambrose*, translated by Boniface Ramsey. Early Church Fathers. New York: Routledge, 1997.

Ambrose, St. "Letter 59 [63]." In *Letters*, translated by Mary M. Beyenka. The Fathers of the Church, vol. 26. New York: Fathers of the Church, Inc., 1954. Both the treatise and the

letter provide a justification of asceticism on primarily scriptural grounds.

Athanasius, St. "Life of Antony." In *Life of Antony and the Letter to Marcellinus*, translated by Robert C. Gregg. Classics of Western Spirituality. New York: Paulist Press, 1980. The biography of a model ascetic.

Augustine, St. *The City of God Against the Pagans*, edited and translated by Robert W. Dyson. Cambridge Texts in the History of Political Thought. New York: Cambridge University Press, 1998. An attack on the life of the flesh.

Bentham, Jeremy. "Of Principles Adverse to That of Utility." In *An Introduction to the Principles of Morals and Legislation*, edited by J. H. Burns and H. L. A. Hart. London: Athlone Press, 1970.

Blanshard, Brand. *Reason and Goodness*. New York: Macmillan, 1961. A brief, incisive criticism.

Buddhaghosa. "Chapter 17." In *The Path of Purity: Being a Translation of Buddhaghosa's Visuddhimagga*. 3 vols. (1923, 1929, 1931), translated by Pe Maung Tin. Reprint (3 vols. in 1). London: Pali Texts Society, 1975. A brief statement of the view that liberation requires annihilation of desire.

Davids, T. W. Rhys, and C. A. F. Rhys David. "Udumbarikā-Sīhanāda Suttanta." In *Dialogues of the Buddha, Part 3*. Sacred Books of the Buddhists, vol. 4, edited by T. W. Rhys Davids. 1921. Reprint. London: Pali Text Society, 1977. Criticism of austerities.

Diogenes Laërtius. "Diogenes" and "Zeno." In *Lives of Eminent Philosophers, Vol. 2, Books 6–10*. Loeb Classical Library. 1925. Limited but helpful information.

Gregory of Nyssa, St. "On Virginity." In *Ascetical Works*, translated by Virginia W. Callaban. The Fathers of the Church, vol. 58. Washington, DC: Catholic University of America Press, 1967. A defense of the ascetic life as means to virtue.

Lecky, William E.H. *History of European Morals from Augustus to Charlemagne, Vol. 2*, 107–148, 164–194. London: Longmans, Green, 1869. A critical history of the rise of Christian asceticism.

Nietzsche, Friedrich. *On the Genealogy of Morals: A Polemic*, translated by Douglas Smith. Oxford World's Classics (1996). Reprint. New York: Oxford University Press, 1998. Nietzsche's most sustained criticism of ascetic ideals.

Pantañjali. "Chapter II: Practices." In *The Yoga Sutras of Patañjali*, translated by Alfred Scheepers. Amsterdam: Olive Press, 2005. Especially Sutras 32 and 43, which give injunctions to practice austerities.

Plato. *Phaedo*, translated by David Gallop. Oxford World's Classics (1993). Reprint. New York: Oxford University Press, 1999. An important argument for freeing the soul from the body.

Plotinus. *The Enneads*, translated by Stephen Mackenna, abridged by John Dillon (1917–1930, 1956). Reprint. New York: Penguin Books, 1991. An influential condemnation of matter.

Reid, J. S. "Asceticism [Roman]." In *Encyclopedia of Religion and Ethics*, edited by James Hastings. New York: Charles Scribner's Sons, 1913. An informative survey.

Schopenhauer, Arthur. *The World as Will and Representation*. 2nd ed. 2 vols. Translated by E. F. J. Payne. New York: Dover, 1966. A wordy but interesting argument for the annihilation of the will.

Umasvati. *That Which Is: Tattvārtha Sūtra*, translated by Nathmal Tatia. The Sacred Literature Series, edited by Kerry Brown and Sima Sharma. San Francisco: HarperCollins, 1994. The basic document of the Jains.

OTHER WORKS OF INTEREST

Brown, Peter. *The Body and Society: Men, Women and Sexual Renunciation in Early Christianity*. New York: Columbia University Press, 1988. Magisterial survey of early Christian sexual renunciation.

Brown, Peter. "The Rise and Function of the Holy Man in Late Antiquity." *Journal of Roman Studies* 61 (1971):80–101. Explores the role and social function of the ascetic individual within society.

Clark, Elizabeth. *Reading Renunciation: Asceticism and Scripture in Early Christianity*. Princeton, NJ: Princeton University Press, 1999. Incisive study of early Christian asceticizing hermeneutic of scripture.

Dunn, Marilyn. *The Emergence of Monasticism: From the Desert Fathers to the Early Middle Ages*. Malden, MA: Blackwell, 2000. Useful survey of monastic development and transmission from fourth century Egypt to seventh century Europe.

Elm, Susanna. *Virgins of God: The Making of Asceticism in Late Antiquity*. Oxford: Clarendon Press, 1994. On the role of women in the development of ascetic traditions and institutions.

Flood, Gavin. *The Ascetic Self: Subjectivity, Memory and Tradition*. Cambridge, U.K.: Cambridge University Press, 2004. Comparative study of asceticism and identity formation across time and religious traditions.

Francis, James. *Subversive Virtue: Asceticism and Authority in the Second-Century Pagan World*. University Park: Pennsylvania State University Press, 1995. Analysis of the cultural perceptions of asceticism as "deviant" during the period from 121 CE to 217.

Harpham, Gregory Galt. *The Ascetic Imperative in Culture and Criticism*. Chicago: Chicago University Press, 1987. Important theoretical consideration of the implications of asceticism on language, text, and meaning

North, Helen. *Sophrosyne: Self-Knowledge and Self-Restraint in Greek Literature*. Ithaca: Cornell University Press, 1966. Follows the conceptual development of "self-control" from the Greek archaic period into the Patristic centuries.

Silbur, Ilana F. *Virtuosity, Charisma and the Social Order: a Comparative Sociological Study of Monasticism in Theravada Buddhism and Medieval Christianity*. Cambridge, U.K.: Cambridge University Press, 1995. Both a corrective and extension of Weber's pioneering work on religious [ascetic] virtuosi.

Theodoret of Cyhrrus, St. *History of the Monks of Syria*, translated by Richard M. Price. Cistercian Studies Series, no. 88. Kalamazoo, MI: Cistercian, 1985. Contemporary history of the peculiar ascetic behaviors of the early anchorites of Northern Syria.

Valantasis, Richard. "Constructions of Power in Asceticism." *Journal of the American Academy of Religion* 63 (1995): 775–821. On the reconstitution of subjectivity, social dynamics, and symbolic meaning through performative behaviors.

Wimbush, Vincent L., and Richard Valantasis, eds. *Asceticism* (1998). Reprint. Oxford: Oxford University Press, 2002. Very useful reference tool spanning periods, traditions, and religions.

Carl Wellman (1967)
Bibliography updated by Brent A. Smith (2005)

ASSOCIATIONISM

See *Psychology*

ASTELL, MARY

(1666–1731)

Mary Astell was born November 12, 1666, in Newcastle-upon-Tyne, into a family of coal merchants. This fact itself is interesting, since it means that she was a member of the comfortable middle class. Her circumstances became considerably less comfortable when her father died in 1678, when Astell was twelve, leaving her without a dowry, and hence, without prospects. Around 1684, and following the death of her mother, Astell took what seems to be a rather startling step: She transferred herself from Newcastle to London, away from what family she had left, apparently to live alone in a town without family. Scholars are tantalizingly ignorant of the circumstances that prompted this move and of Astell's prospects in London.

While things do not seem to have gone well for her initially, by 1695 she had established herself in Chelsea, enjoying the patronage of Lady Catherine Jones (1672–1740), and surrounded by a circle of intellectually minded women. By this time, moreover, Astell seems to have put herself in a position to make her living by her pen. Scholars are equally ignorant of the circumstances that gave Astell sufficient intellectual confidence to embark on a course such as this. She had, of course, no formal education. A clergyman uncle, Ralph Astell, is often credited with tutoring her, and, since he was known to have attended Emmanuel College, Cambridge during the heyday of the Cambridge Platonists, he is also often assumed to have shaped Astell's philosophical interests. But since he died soon after the death of Astell's father, when she herself was thirteen, her uncle's influence would have had to have been on a very precocious child. That he was removed from his pulpit for drunkenness raises further doubts about his effectiveness as an educator of a young girl.

Astell recently has attracted attention due to the undoubted feminist nature of at least some of her work, on the basis of which she has been hailed as an early feminist. In *A Serious Proposal to the Ladies* (1694), she argues that women's indubitable possession of rational faculties means that they deserve an education, one that would enable them to develop their rational, moral capabilities and so to live a life devoted to the care of their souls. In *Some Reflections upon Marriage Occasion'd by the Duke and Duchess of Mazarine's Case, Which Is Also Consider'd* (1700), Astell develops this theme, arguing that a well-trained mind will enable women to lead a virtuous life, even in the face of a bad marriage.

Astell's interests, however, extended into a number of other areas beyond the defense of her sex. She is the author of several political pamphlets, in which she took up and discussed issues of contemporary moment from a conservative perspective. Her *magnum opus* is a work of Christian theology, *The Christian Religion, as profess'd by a daughter of the Church of England* (1705). In this lengthy work, Astell, critically reacting to an anonymous work called *The Ladies' Religion*, lays out an extensive examination, first of natural, then of revealed religion, and concludes with an examination of Christian practice, including our duty to god, our neighbor and ourselves. There is some interesting philosophical material contained here, most especially in the discussion of the debate between John Locke and Edward Stillingfleet on the possibility of thinking matter. Astell's works that are most predominantly philosophical in nature, however, include her published correspondence with John Norris, *Letters concerning the Love of God* (1695) and *A Serious Proposal to the Ladies, Part II: Wherein a Method is Offer'd for the Improvement of Their Minds* (1697).

Letters concerning the Love of God contains some of the most interesting and tightly argued of Astell's writing. Her role in this correspondence, however, is that of a questioner and a critic. It is not entirely possible, therefore, to derive from the *Letters* an account of Astell's own position on the matters she discusses. She raises two issues with Norris. The first is how to understand God's causal role with respect to pain. If God, as Norris claims, is the sole object of our love as the cause of pleasure, is He not as well the sole object of our aversion, as the cause of pain? While she is prepared to admit that corporeal pain may have a purpose that is good, she is concerned to secure the possibility of a class of evils, that, as sinful, must be the object of aversion. Astell's second worry concerns the consequences for human social relations if God is the only object of our love. While she initially appears

to accept Norris's distinction between loving creatures for our good but not as our good, in a final letter, she raises more substantive questions about Norris's occasionalism. She holds that if bodies are causally inefficacious and do not cause the sensations we have of them, then sensations are irrelevant and God must be said to have created in vain. It is not necessary to the thesis that God is the only object of our love, she points out, to suppose that God acts without instruments, for we never, when receiving a gift, feel gratitude towards the giver's instrument, rather than the giver.

In the second part of *A Serious Proposal*, Astell again adopts a position that reflects some of Norris's approach, while rejecting his occasionalism. The second part of *A Serious Proposal* has a somewhat different project than the first. By the time she wrote it, Astell, who had cherished hopes that she would receive funds to start the educational institution for which she had advocated, had come to realize that these funds would not be forthcoming. Therefore, the second part takes more of a self-help approach to the question of women's education, in which Astell outlines the methods by which a human understanding, as she describes it, may be improved. The argument in favor of improvement is the one she originally put forward, that human action, governed as it is by rationality, requires an informed understanding and a properly directed inclination.

In developing her account, Astell acknowledges a debt to Antoine Arnauld's *Art of Thinking* and to Rene Descartes's *Principles of Philosophy*. She argues that all human endeavor requires the application of right principles, and therefore that anyone, whether doctor or plowman, is concerned with knowledge and with the rules of right reason. These rules are to be induced from right practice, and are not a matter of formal structure. She takes the management of right inclination to be crucial to right conduct and follows Norris is holding that we ought to model our will on God's. She rejects his occasionalism, however, and instead insists that we need to recognize that our minds are united to our bodies. "For if we disregard the Body wholly," she writes, "we pretend to live like Angels whilst we are but Mortals and if we prefer or equal it to the Mind we degenerate into Brutes" (1997, p. 158). Our goal therefore is to harness the passions we feel to the proper goals for human happiness, as discovered by our rational nature, directed to eternal happiness.

See also Arnauld, Antoine; Cambridge Platonists; Descartes, René; Feminist Philosophy; Happiness; Locke, John; Norris, John; Stillingfleet, Edward; Women in the History of Philosophy.

Bibliography

WORKS BY ASTELL

Letters concerning the Love of God, between the Author of the Proposal to the Ladies and Mr. John Norris. Printed for Samuel Manship, 1695. Modern edition: Mary Astell and John Norris, *Letters concerning the Love of God* (Early Modern Englishwoman: A Facsimile Library of Essential Works), edited by Derek Taylor and Melvyn New. Ashgate, 2005.

A Serious Proposal to the Ladies for the Advancement of their True and Greatest Interest. London: printed for R. Wilkin, 1694.

A Serious Proposal to the Ladies, Part II, Wherein a Method Is Offer'd for the Improvement of their Minds. London: printed for Richard Wilkin, 1697. Modern edition: Mary Astell, *A Serious Proposal to the Ladies. Parts I and II*, edited with introduction and notes by Patricia Springborg. London: Pickering and Chatto, 1997.

Some Reflections upon Marriage Occasion'd by the Duke and Duchess of Mazarine's Case; Which is Also Consider'd. London: printed for John Nutt, 1700. Modern edition: (Third edition reproduced) in Astell, *Political Writings*, edited by Patricia Springborg. Cambridge, U.K.: Cambridge University Press, 1996.

The Christian Religion as Profess'd by a Daughter of the Church of England in a Letter to the Right Honorable T.L., C.I. London: printed by S. H. for R. Wilkin, 1705.

WORKS ABOUT ASTELL

Atherton, Margaret, "Cartesian Reason and Gendered Reason." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte Witt. Boulder, CO: Westview Press, 1993.

Broad, Jacqueline. "Mary Astell." In *Women Philosophers of the Seventeenth Century*, 90–113. New York: Cambridge University Press, 2002

Perry, Ruth. *The Celebrated Mary Astell: An Early English Feminist*. Chicago: University of Chicago Press, 1986.

Smith, Florence M. *Mary Astell*. New York: Columbia University Press, 1916.

Springborg, Patricia. *Mary Astell: 1666–1731 Theorist of Freedom from Domination*. New York: Cambridge University Press, 2005.

Squadrito, Kathleen, "Mary Astell's Critique of Locke's View of Thinking Matter." *Journal of the History of Philosophy*, 25.3 (Sept 1987): 433–439.

Margaret Atherton (2005)

ATHEISM

The words *atheist* and *godless* are still frequently used as terms of abuse. Nevertheless, there are relatively few people nowadays in whom the thought of atheism and atheists arouses unspeakable horror. It seems to be agreed that an atheist can be a good person whose oaths and promises are no less trustworthy than those of other people, and in most civilized lands atheists have the same or

nearly the same rights as anybody else. What is more, it appears to be generally realized that some of the world's foremost philosophers, scientists, and artists have been avowed atheists and that the increase in atheism has gone hand in hand with the spread of education. Even spokesmen of the most conservative religious groups in the mid-twentieth century conceded that atheism may well be a philosophical position that is adopted for the noblest of reasons. Thus, in "The Contemporary Status of Atheism" (1965), Jean-Marie Le Blond appealed to his fellow believers for a "truly human and mutually respectful dialogue" with atheists, insisting that a "life without God need not be … bestial, unintelligent, or immoral" and that atheism can be "serene and deeply human." In the previous year Pope Paul VI, in his encyclical *Ecclesiam Suam,* had observed that some atheists were undoubtedly inspired by "greathearted dreams of justice and progress" as well as by "impatience with the mediocrity and self-seeking of so many contemporary social settings."

HOSTILITY TO ATHEISM

It was otherwise in earlier ages. One could fill many volumes with the abuse and calumny contained in the writings of Christian apologists, learned no less than popular. The tenor of these writings is not simply that atheism is mistaken but also that only a depraved person could adopt so hideous a position and that the spread of atheism would be a horrifying catastrophe for the human race. "No atheist as such," wrote Richard Bentley in *Eight Sermons* (1724), "can be a true friend, an affectionate relation, or a loyal subject." In the preface to his *The True Intellectual System of the World* (1678), Ralph Cudworth made it clear that he was addressing himself not to "downright and professed atheists" but to "weak, staggering and sceptical theists." Downright atheists were beyond the pale, for they had "sunk into so great a degree of sottishness" that they evidently could not be reached. Writing almost exactly two centuries later, the Protestant theologian Robert Flint, who readily admitted that he had met atheists of great courage and integrity, nevertheless expressed his extreme concern over the "strenuous propagation" of atheism, especially in the "periodical press." "The prevalence of atheism in any land," he wrote, "must bring with it national decay and disaster." The triumph of atheism in England would "bring with it hopeless national ruin." If once the workers of the large cities became atheists, "utter anarchy would be inevitable" (*Anti-Theistic Theories,* pp. 36–37). All these quotations are from British Protestants. Very similar and frequently more virulent remarks could be quoted from German, French, Italian, and American believers of the same periods.

In France until the Revolution and in most other countries until some time later, it was illegal to publish works in defense of atheism, and in fact real or alleged atheists were subject to dire persecution throughout the times of Christian domination. Some of the world's greatest philosophers were among those who advocated and in some instances actively promoted this persecution. The story antedates Christianity, and persecution of atheists was already advocated in Plato's *Laws.* Plato divided atheists into several groups, all of which must be punished; but whereas the members of some groups required no more than "admonition and imprisonment," those belonging to others deserved punishment exceeding "one death … or two." Thomas Aquinas (*Summa Theologiae,* II, 11, 3 and 4) had no doubt that unbelievers should be "shut off from the world by death." Such a course, he argued, is justified since it surely is "a much more serious matter to corrupt faith, through which comes the soul's life," than it is "to forge money, through which temporal life is afforded." If, as is just, forgers of money and other malefactors are straightaway put to death, it is all the more just that "heretics … be not only excommunicated but also put to death."

John Locke, one of the great pioneers of religious toleration, explicitly exempted Roman Catholics and atheists from the application of the principles he advocated. "Promises, covenants, and oaths, which are the bonds of human society," he wrote, "can have no hold upon an atheist." Moreover, since atheism is not a religion but, on the contrary, a position that is out to "undermine and destroy all religion," it cannot come under the privilege of the toleration that is justly claimed by bona fide religions (*A Letter concerning Toleration*). It may be assumed that Locke did not advocate that atheists be shut off from the world, but that he was merely opposed to the free advocacy of atheism in writing and speech.

After Locke's time, the "shutting off" approach became infrequent, but atheists continued to be the victims of persecution and discrimination in various forms. To give some interesting and far from untypical illustrations: Baron d'Holbach's *The System of Nature* was falsely attributed in its first edition to Jean-Baptiste de Mirabaud, a former secretary of the French Academy who had been dead for ten years. Very shortly after its publication in 1770, it was condemned to be burned by the public hangman after a trial in which the public prosecutor expressed his regret that he could not lay his hands on the unknown real author, adding that the corruption of

morals evident in almost all sections of society was very probably due to the spread of ideas like those contained in the condemned book. When the poet Percy Bysshe Shelley was an undergraduate at Oxford, he published a short and very temperate pamphlet titled *The Necessity of Atheism.* This at once aroused a violent protest that resulted in the burning of all undistributed copies and in the expulsion of Shelley and his friend Thomas Hogg from the university. Some years later Shelley was judicially deprived of the custody of his children on the ground that he was "likely to inculcate the same [atheistic] principles upon them." As late as 1877 Annie Besant, the noted social reformer, was judged to be unfit to take care of her children on the same ground, although the judge admitted that she had been a careful and affectionate mother. Until the passing of the Evidence Amendment Act of 1869, unbelievers in Great Britain were considered incompetent to give evidence in a court of law. Atheists were thus in effect unable to sue when they were the victims of fraud or slander. Charles Bradlaugh, whose efforts were largely responsible for the Act of 1869, was also the main figure in a prolonged battle to secure the right of avowed atheists to sit in the House of Commons. After Bradlaugh was elected, he was found unfit to take his seat. He won the resulting by-election and was again declared unfit to sit in the House, and this merry-go-round continued for several years, until a Conservative speaker found a legal way of securing Bradlaugh's admission. In the United States there has not been similar legal discrimination against atheists, but there is perhaps to this day more de facto discrimination and prejudice than in any other Western country.

A comprehensive entry on atheism would, among other things, trace the history of the persecution of real and alleged atheists, of the changes in public attitudes, and of the gradual repeal of discriminatory legislation. It would also inquire into the psychological sources of the hatred of atheists that is sometimes found in otherwise apparently kindly and sensible men. Because of space limitations, the present entry will, however, be largely confined to what is undoubtedly the most interesting question for philosophers: Is atheism a logically tenable position? What are the arguments for it, what are the arguments against it, and how strong are these, respectively? It will not be possible to deal exhaustively even with these questions, but an attempt will be made to sketch the position of a philosophically sophisticated atheist and to explain why a view of this kind has appealed to many important thinkers in recent times.

DEFINITION OF *ATHEISM*

No definition of *atheism* could hope to be in accord with all uses of this term. However, it would be most confusing to adopt any of several definitions that can only be regarded as eccentric. These would result in classifying as believers many people who would not regard themselves as such (and who would not commonly be so regarded) and in classifying as atheists many people who have not usually been thought of in this way. Thus, Johann Gottlieb Fichte, in denying the charge of atheism, wrote in "Über den Grund unseres Glaubens an eine Göttliche Weltregierung" that the "true atheist" is the person who, instead of following the voice of conscience, always calculates consequences before acting in a moral situation. Friedrich Jodl, who was himself a positivist and an unbeliever, similarly remarked that "only the man without ideals is truly an atheist," implying, no doubt, that, although he did not believe in God, he was not a "true" atheist (*Vom Lebenswege,* 2 vols., Stuttgart and Berlin, 1916–1917, Vol. II, p. 370.). In the twentieth century Paul Tillich defined *atheism* as the view that "life has no depth, that it is shallow." Anybody who says this "in complete seriousness is an atheist"; otherwise, he is not (*Shaking of the Foundations,* New York, 1948, p. 63). Stephen Toulmin, in an article ("On Remaining an Agnostic," *Listener,* October 17, 1957) in which he championed agnosticism as he understood it, distinguishes his own position from that of both believers and atheists in that, unlike them, he does not "find personal attitudes of any sort in Nature-at-large." The believer, according to Toulmin, regards the Cosmic Powers as friendly to man, while the atheist regards the cosmos as indifferent or as "positively callous."

Whatever the point of the definitions just quoted, their paradoxical consequences make them useless in the present context. For our purposes, definitions of *atheism* and corresponding definitions of *God* will be serviceable only if they preserve, at least roughly, the traditional battle lines. Whatever their differences, Augustine, Thomas Aquinas, Locke, George Berkeley, William Paley, Henry Longueville Mansel, J. S. Mill, William James, Paul Tillich, and John Hick should continue to be classified as believers; T. H. Huxley, Leslie Stephen, and Clarence Darrow as agnostics; and Holbach, Ludwig Büchner, Ludwig Feuerbach, Karl Marx, Arthur Schopenhauer, Friedrich Nietzsche, and Jean-Paul Sartre as atheists. The definition proposed in the present entry will, in taking account of certain complexities of the situation, depart in a significant respect from the one that is most popular, but it will not involve reclassification of any of the great philoso-

phers of the past. According to the most usual definition, an *atheist* is a person who maintains that there is no God, that is, that the sentence "God exists" expresses a false proposition. In contrast, an agnostic maintains that it is not known or cannot be known whether there is a God, that is, whether the sentence "God exists" expresses a true proposition. On our definition, an *atheist* is a person who rejects belief in God, regardless of whether or not the reason for the rejection is the claim that "God exists" expresses a false proposition. People frequently adopt an attitude of rejection toward a position for reasons other than that it is a false proposition. It is common among contemporary philosophers, and indeed it was not uncommon in earlier centuries, to reject positions on the ground that they are meaningless. Sometimes, too, a theory is rejected on such grounds as that it is sterile or redundant or capricious, and there are many other considerations that in certain contexts are generally agreed to constitute good grounds for rejecting an assertion. An atheist in the narrower, more popular sense, is ipso facto an atheist in our broader sense, but the converse does not hold.

THEISTIC POSITIONS. Before exploring the implications of our definition any further, something should be said about the different uses of the word *God* and the correspondingly different positions, all of which have been referred to as "belief in God." For our purposes, it will be sufficient to distinguish three of these. All the believers in question have characterized God as a supreme personal being who is the creator or the ground of the universe and who, whatever his other attributes may be, is at the very least immensely powerful, highly intelligent, and very good, loving, and just. While some of them would maintain that the predicates just mentioned—"powerful," "good," and the rest—are used in a literal sense when applied to God, other believers insist that when applied to God, these, and indeed all or almost all, predicates must be employed in "metaphorical," "symbolic," or "analogical" senses. Let us, without implying anything derogatory, refer to the belief that predicates can be applied literally to God as the "anthropomorphic" conception of God and to the belief that predicates can only be applied analogically to God as the "metaphysical" conception of God.

Among professional philosophers, belief in the metaphysical God has been much more common than belief in the anthropomorphic God. This metaphysical position is at least as old as Thomas (and, it may be plausibly argued, as old as Plato). In the early eighteenth century it was championed by Peter Browne, bishop of Cork, who was trying to answer difficulties raised by the infidel John Toland. In the nineteenth century this position was defended by Mansel in his Bampton Lectures, and in the twentieth century it was a key feature of Tillich's philosophy. God, on Tillich's view, "infinitely transcends every finite being"; between the finite and the infinite there is "an absolute break, an 'infinite jump'"; there is here "no proportion and gradation." When we say, for example, "God is Love," or "God is Life," the words *love* and *life* are used symbolically, not literally. They were originally introduced in connection with "segments of finite experience," and when applied to God, they cannot have the same meaning that they have in ordinary human situations.

The anthropomorphic position is by no means confined to unsophisticated believers. It has commanded the support of several eminent philosophers, especially believers who were also empiricists or otherwise opposed to rationalism. Thus, Berkeley emphatically defended the anthropomorphic position against Bishop Browne. In *Alciphron* Berkeley attacked Browne's procedure on the ground that unless "wise" and "good" are used in the same sense for God and man, "it is evident that every syllogism brought to prove those attributes, or (which is the same thing) to prove the being of a God, will be found to consist of four terms, and consequently can conclude nothing." In the nineteenth century J. S. Mill championed anthropomorphic belief as opposed to the metaphysical theology of Hamilton and Mansel; more recently, Miguel de Unamuno y Jugo, who is perhaps best classified as a fideist, indicted the metaphysical God as a "Nothing-God" and a "dead thing." In *The Tragic Sense of Life in Men and in Peoples* he wrote that such a fleshless abstraction cannot be the answer to the cravings of the human heart. Only the anthropomorphic God can ever be "the loving God," the God to whom we come "by the way of love and of suffering."

Among those who believe in an anthropomorphic God, there are two positions to be distinguished. First, there is the more traditional position that allows no limitations upon the extent to which God possesses the various admirable characteristics—on this view, God is all-powerful, all-loving, infinitely good, perfectly just, and so on. Second, there is the somewhat heretical position of those who, while maintaining that God possesses these characteristics to a high degree, allow that he is limited at least in his power or in his goodness. Mill, who believed in such a finite anthropomorphic deity, claimed that regardless of the official pronouncements of the various religions, in actual practice most Western believers adhered to a theory like his own.

Creation. A few words must be said about the possible meanings of *creation* when God is referred to as the creator (or ground) of the universe. Thomas Aquinas, in his *On the Eternity of the World* and elsewhere, makes a distinction between the temporal sense in which God is supposed to have made the universe at a certain moment in time, prior to which it did not exist, and the more sophisticated sense in which it is asserted that the universe is absolutely dependent on God so that it would cease to exist if God were not sustaining it. Thomas himself believed in God's creation of the universe in both senses, but it was only in the second sense that he regarded the theory of divine creation as susceptible of logical proof. Both these senses must be distinguished from the creative activity ascribed to the demiurge of Plato's *Timaeus* or to Mill's God. Here the deity is not, strictly, a creator but merely an arranger of preexisting material. For the purposes of this entry, a person will count as a believer in the creation of the universe by God if he or she makes any of three claims just distinguished.

THE BROADER SENSE OF ATHEISM. Let us now return to our definition of *atheism.* A person is an atheist in our sense who adopts an attitude of rejection toward all three theistic positions previously stated—belief in a metaphysical God, in an infinite anthropomorphic God, and in a finite anthropomorphic God. He or she will count as a believer in God if maintaining that "God exists" expresses a true proposition, where "God" is employed in one of the three ways described. A person will be an agnostic who does not accept any of these three claims but at the same time suspends judgment concerning at least one of them. It will be observed that on our way of drawing the lines, agnosticism and atheism remain distinct positions, since suspension of judgment and rejection are different attitudes.

The broader definition here adopted enables us to classify together philosophers whose attitudes toward belief in God are exceedingly similar, although their detailed reasons may not always coincide. Rudolf Carnap, for example, regards metaphysical theology as meaningless, while treating belief in an infinite as well as a finite anthropomorphic God as "mythology," implying that both are false or probably false. In our sense, he can be classified as an atheist without further ado, and it is doubtful that believers would consider him less hostile than atheists in the narrower sense. It is also worth observing that our broader definition receives a good deal of backing from the actual writings of philosophers and others who regarded themselves as atheists. Many of them were by no means unaware of the fact that the word *God* has a number of uses and that what may be a plausible justification for rejecting one kind of belief in God may be quite inappropriate in the case of another. Charles Bradlaugh, for example, made it very clear that in calling himself an atheist he did not simply maintain that there is no God. In his "Plea for Atheism," he wrote:

> The atheist does not say "there is no God," but he says "I know not what you mean by God; I am without idea of God; the word 'God' is to me a sound conveying no clear or distinct affirmation.… The Bible God I deny; the Christian God I disbelieve in; but I am not rash enough to say there is no God as long as you tell me you are unprepared to define God to me."

The writings of Jean Meslier, Holbach, and other eighteenth-century and nineteenth-century atheists, while certainly containing remarks to the effect that the sentence "God exists" expresses a false proposition, are also full of claims that once we critically examine the talk about a "pure spirit" that supposedly exists timelessly and without a body, we find that words have been used without any meaning. In any event, by using the word *atheism* in the broader sense, it will be possible to discuss certain antitheological considerations of great interest that would otherwise have to be excluded.

TRADITIONAL ATHEISTIC ARGUMENTS

In this section we shall discuss two of the arguments popular among atheistic writers of the eighteenth and nineteenth centuries. In later sections we shall present considerations commonly urged by Anglo-Saxon writers in more recent years. However, in a rudimentary form these more recent reflections are already present in the writings of earlier atheists, just as the older arguments continue to be pressed in current literature.

THE ETERNITY OF MATTER. The first of the two older atheistic arguments is based on the doctrine of the eternity of matter, or, to bring it more in accord with recent physical theory, the eternity of mass-energy. (As far as the basic issues here are concerned, it is not of any moment whether what is said to be eternal is matter or energy or mass-energy, and for the sake of convenience we shall speak only of the eternity of "matter.") There are two steps in this argument. It is claimed, first, either as something self-evident or as a proposition proved by science, that matter is eternal; second, it is asserted that this claim rules out a God conceived as the creator of the material universe. If the physical universe had been created by

God, it would follow that there was a time when the quantity of matter was less than it is now, when it was in fact zero. But physics proves or presupposes that the quantity of matter has always been the same.

Since most ordinary people include "creator of the material universe" in their concept of God, and since they mean by *creation* a temporal act of making something out of nothing, the appeal to the eternity of matter is effective as a popular argument for atheism. A little reflection shows, however, that by itself the argument is of very limited significance. To begin with, regardless of any scientific evidence, the doctrine of the eternity of matter, in all its forms, would be challenged by anybody who accepts any of the causal varieties of the Cosmological Argument. Such a person would presumably argue that while conservation principles may accurately describe a certain feature of the material universe *ever since it began existing,* the material universe itself requires a nonmaterial cause. Hence, any atheistic conclusion in the present context would have to be accompanied by a refutation of the causal forms of the Cosmological Argument. But granting for the moment that the eternity of matter is fully established, this is not incompatible with the theory of divine creation in the sense in which it has been put forward by its philosophically more sophisticated adherents. The eternity of matter is no doubt incompatible with the existence of a God who made the material universe out of nothing and with the kind of activity in which the demiurge is supposed to engage (since bringing order into previously chaotic materials requires the addition of energy); but it is not incompatible with creation in the second of the two senses distinguished by Thomas, in which *creation* means "absolute dependence" and does not refer to any datable act. There may indeed be some difficulty in the notion of a nonphysical entity nonphysically sustaining the universe, and it is tempting to think that this is an intelligible doctrine simply because the words *sustain* and *depend* immediately call up certain pictures in one's mind; but these difficulties raise rather different questions. Finally, in this connection it should be pointed out that the eternity of matter in all its forms is compatible with a belief in God or gods, like those of the Epicureans and Thomas Hobbes (if Hobbes was serious), who are physical beings, or in gods of any kind, as long as it is not claimed that these have created the universe or any aspect of it.

A few words should perhaps be added here about the claim of some writers that the doctrine of the eternity of matter in all its forms has now been refuted by physics and that physics even somehow proves the existence of God. In this connection it should be mentioned, first, that the great majority of scientifically informed philosophers agree that the findings of recent physics do not affect the issues dividing believers and unbelievers, and, second, that even if the doctrine of the eternity of matter were now untenable in all its forms, this would undermine one of the arguments for atheism, but not atheism itself. If there was a time when matter did not exist (assuming this to be a meaningful assertion), it does not automatically follow that matter was created by God. To show that matter was created by God, an appeal to the Cosmological Argument (and not to physics) would be as necessary as ever. As for the theory of continuous creation, advocated by some cosmologists, it does indeed imply that the principle of the conservation of mass-energy is false. However, the basic assumption behind the theory of continuous creation is the so-called perfect cosmological principle, which is in effect an endorsement of the eternity of matter. This principle asserts that the large-scale aspects of the universe are the same at all times and in all places; and this, more specifically, means that the stars and galaxies have always been about as evenly distributed as they are at the present time.

EVIL AND OTHER IMPERFECTIONS. Among the traditional atheistic arguments a second type has generally been regarded as more formidable and still enjoys an undiminished popularity. This type of argument points to some imperfection or defect in the universe and argues that the defect is incompatible with the existence of God insofar as God is defined as a perfect being.

Among the imperfections or alleged imperfections, emphasis has frequently been placed on the enormous waste in nature, especially in matters of reproduction, and on the trial-and-error "method" of evolution. Referring to the process of evolution, G. H. Lewes remarked that "nothing could be more unworthy of a supreme intelligence than this inability to construct an organism at once, without making several tentative efforts, undoing today what was so carefully done yesterday, and repeating for centuries the same tentatives and the same corrections in the same succession." And if the end of this entire process is man, it has been questioned whether it was worth all the pain and tribulations that preceded it. "If I were granted omnipotence, and millions of years to experiment in," writes Bertrand Russell, "I should not think Man much to boast of as the final result of my efforts" (*Religion and Science,* p. 222). Again, it has been suggested by several writers, and not at all facetiously, that if there were a God, then surely he would have provided human beings with clearer evidence of his own existence.

If an omniscient and omnipotent God did not take care that his intentions should be understood by his creatures, asked Nietzsche, "could he be a God of goodness?" Would he not, rather, be a cruel god if, "being himself in possession of the truth, he could calmly contemplate mankind, in a state of miserable torment, worrying its mind as to what was truth?" (*Morgenröte,* Aphorism 91). If a God exists, then, in the words of Charles Bradlaugh, "he could have so convinced all men of the fact of his existence that doubt, disagreement, or disbelief would be impossible."

The most widely discussed of all these arguments from the imperfections of the universe is the argument from evil, and it may be best to restrict our discussion to it. The following is a statement by Brand Blanshard:

> We are told that with God all things are possible. If so, it was possible for him to create a world in which the vast mass of suffering that is morally pointless—the pain and misery of animals, the cancer and blindness of little children, the humiliations of senility and insanity—were avoided. These are … apparently … inflictions of the Creator himself. If you admit that, you deny his goodness; if you say he could not have done otherwise, you deny that with him all things are possible. ("Irrationalism in Theology," in *Faith and the Philosophers,* edited by John Hick, London, 1964, p. 172)

It should be emphasized that the argument from evil, as here stated, is directed against the conclusion of the believer in an infinite anthropomorphic God and is not merely a criticism of his evidence. On occasions, for example in David Hume's *Dialogues concerning Natural Religion,* the argument has been used for the milder purpose of showing that the Design Argument cannot succeed in establishing a maker of the universe who is both omnipotent and perfectly good. It argues from the nature of the world to the nature of its cause, and since the world is a mixture of good and evil, it cannot be established in this way that its creator is perfectly good. The form in which we are concerned with the argument from evil—what we may call its stronger sense—maintains that the evil in the world shows the theological claim to be false. The argument may be construed as comparing the theological assertion to a falsified scientific hypothesis: If the theory that the universe is the work of an all-powerful and all-good being were true, then the universe would not exhibit certain features; experience shows that it does exhibit these features, and hence the theory is false.

The argument from evil has no logical force against belief in a finite God. The evil in the world is perfectly compatible with the existence of a God who is lacking either omnipotence or perfect goodness, or both. In fact, E. S. Brightman and the American personalists and other well-known champions of belief in a finite anthropomorphic God adopted their position precisely in order to reconcile belief in God with the existence of evil. There is also no obvious incompatibility between the existence of the metaphysical God and the evil in the world, since it is not claimed for the metaphysical God either that he is all-powerful or that he is perfectly good in the ordinary senses of these words. Mansel, for example, in *Limits of Religious Thought* openly acknowledged that in the light of the injustice and suffering we find in the world, the moral character of God cannot be represented "after the model of the highest human morality which we are capable of conceiving." His position, Mansel insisted, unlike the position of anthropomorphic believers, to whom Mansel referred as "vulgar Rationalists" in this context, was immune from difficulties like the problem of evil Substantially similar remarks are to be found in the writings of many other members of this tradition.

The most basic objections to metaphysical theology will be discussed in the next section, but perhaps it should be mentioned in passing that according to some critics, philosophers like Mansel have a tendency to revert to the view that God is good in the very same sense in which human beings are sometimes good and, more generally, to anthropomorphic theology. This is not at all surprising since, like other believers, they derive or wish to derive comfort and reassurance from their theology. Such comfort may be derivable from the view that the ultimate reality is good and just in the sense or one of the senses in which we use these terms when we praise good and just human beings. No comfort at all, on the other hand, seems derivable from the statement that God is good and just but that "the true nature and manner of all the divine operation of goodness," in the words of Bishop Browne, "is utterly incomprehensible" or that they differ from human justice and goodness, as Mansel put it, "in kind," not only in degree.

There is a long history of attempts by believers to show that the argument from evil does not really refute the assertion that an infinite anthropomorphic God exists. It has been maintained by some that evil is unreal; by others that, although real, it is of a "privative" rather than a "positive" character; that it is real and positive but that it is the consequence of man's abuse of his gift of free will and that a universe without evil and without free will would be worse than one with both; that the argument is based on a narrow hedonistic conception of good and evil

and that, in any event, the theological position cannot be adequately judged unless it is viewed in conjunction with belief in an afterlife in which the wrongs of the present life will somehow be righted; and many more. Critics have come up with various answers to these rejoinders, and the discussion has been going on with unabated vigor in recent years. There would be little point in reviewing this debate here, but something should perhaps be said about two retorts by believers that have not been adequately discussed by the proponents of the argument from evil.

A Christian rejoinder. One rejoinder to the argument from evil seems to be of considerable value in showing that this argument does not by itself justify rejection of belief in an infinite anthropomorphic God. It has been argued (for example, by Arnold Lunn in his exchange of letters with C. E. M. Joad published in *Is Christianity True?,* London and Philadelphia, 1933) that although the existence of evil cannot be reconciled with the existence of an infinite anthropomorphic God, this is not too serious a problem in view of the powerful affirmative evidence for this position. In other areas too, Lunn reminds us, we do not abandon a well-supported theory just because we meet with some counterevidence. He is not in the least disturbed by "the fact that divine science, like natural science, brings us face to face with apparently insoluble contradictions." This hardly disposes of the argument from evil, as Lunn seems to think. The comparison between the difficulty that a believer faces from the facts of evil and the difficulties besetting a scientific theory for which there is otherwise strong evidence is somewhat tenuous. There are indeed cases answering to this description in science, but they are invariably resolved by further inquiry. Either we come to see that the difficulty or exception was merely apparent or else the original theory is modified or abandoned. In the theological case, several millennia of experience and debate do not seem to have brought us any nearer a resolution. But, assuming that Lunn's comparison fails as a defense of belief in an infinite anthropomorphic God, there can be no question that he would have made out a strong case in favor of agnosticism as opposed to atheism if there were in fact good evidence for the existence of the God in question. If, for example, the Cosmological Argument were, as far as we can judge, free from fallacious transitions, we would have a situation similar to the kind we frequently face in which there is significant and roughly equally impressive evidence both ways (for example, some apparently trustworthy witnesses implicating the defendant in a court case, while other equally trustworthy witnesses exonerate) and in which suspense of judgment is the most rational attitude. The moral for our discussion is that an atheist cannot afford to neglect the arguments for the existence of God. Unless they can be demolished, the argument from evil will not by itself establish the atheist's case, even if none of the answers mentioned earlier are in fact successful.

A fideistic rejoinder. Another rejoinder to the argument from evil has become extremely popular in recent years among existentialist believers and all who maintain that arguments for or against the existence of God are, as it is put, radically beside the point. We are told that one simply either has faith or one has not, one is either "open" to the presence of God or one is not. If one has faith, proofs and reasoning are not needed; if one lacks faith, they are of no avail. A person who has faith is not shaken by absence of evidence or by counterevidence; a person who has no faith will never become a true believer even if intellectually convinced by the arguments of rationalistic theology.

Systematic defenses by those who adopt such a position are exceedingly rare, but in 1964 an article appeared by an existentialist philosopher who seems familiar with contemporary analytic philosophy and whose answer to the argument from evil is representative of this entire approach. In his "On the Eclipse of God" (*Commentary,* June 1964, pp. 55–60), Emil Fackenheim insists that the essential mark of the faith of a person who is "primordially open to God" is certainty, or, specifically, "*the believer's certainty of standing in relation to an unprovable and irrefutable God*" (Fackenheim's italics). It is this "irrefutability" of his faith that, Fackenheim believes, enables him to circumvent the problem of evil. No conceivable experience, he insists, can possibly upset the true biblical faith. If there is good fortune, it "reveals the hand of God." If the fortune is bad and if this cannot be explained as just punishment, the conclusion is that "God's ways are unintelligible, not that there *are* no ways of God." To put it "radically": "*Religious faith can be, and is, empirically verifiable; but nothing empirical can possibly refute it*" (Fackenheim's italics). Fackenheim cites the examples of Jeremiah, Job, and the Psalmist, all of whom encountered tragedy and disaster without losing their faith in the existence of God. Biblical faith, he observes in this connection, "is never destroyed by tragedy but only tested by it," and in the course of such a test, it "conquers" tragedy. To underline the invulnerability of this position, Fackenheim adds that no amount of scientific evidence can "affect" biblical belief any more than "historical tragedy" or "an empty heart" can.

What is to be said in reply to all this, especially to the remarkable claim, made in all seriousness, that although faith is empirically verifiable, nothing can possibly refute it? The answer is surely that there is a confusion here between logical and psychological issues. Fackenheim may well have given an accurate account of faith as a psychological phenomenon, but this is totally irrelevant to the question at issue among believers, agnostics, and atheists—namely, which position is favored by the evidence or lack of evidence. All the words—*destroy, test, conquer, affect,* and *refute*—are used ambiguously in this as in countless similar discussions. They refer on the one hand to certain psychological effects (or their absence) and on the other to the relation between facts and a proposition for or against which these facts are (or fail to be) evidence. If the question at issue were whether tragedy and injustice can produce loss of belief in a person who has the "biblical faith," the answer may well be in the negative, and Fackenheim's examples support such an answer. They have not the slightest bearing, however, on the question of whether the tragedies and the injustices in the world disprove or make improbable or are any kind of evidence against the statement that the world is the work of an all-powerful and all-good God—the statement in which the believers have faith. The first question may be of great psychological and human interest, and if Fackenheim is right, then a person interested in dissuading "biblical" believers would be foolish even to try. It is the second question alone, however, that is of interest to philosophers, and it alone is at issue between believers and unbelievers. By telling his biblical stories, Fackenheim has done nothing whatsoever to circumvent the problem of evil or to show that what the believer has faith in is immune to criticism.

Before leaving this topic, a few words are in order about a certain concession, occasionally made by unbelievers, which does not appear to be warranted. Some atheists are willing to concede that whereas they can come to grips with rationalistic believers, they are powerless when faced with a fideist like Fackenheim. Thus, Ernest Nagel, in his "Defense of Atheism," remarks that such a position is "impregnable to rational argument." Now, if a proposition, *p,* is endorsed on the basis of faith and not on the basis of logical arguments, then indeed a critic cannot undermine any arguments supporting *p,* but may well be in a position to test (and falsify) *p* itself. If a fideist were to maintain, admitting from the outset that there is no evidence for the proposition and that it is based on faith alone, that the *New York Times* sells for 50 cents on weekdays, there is of course no evidence for the proposition that can be attacked, but this would not prevent us from disproving the assertion. Any plea by the fideist there is no evidence or that no evidence can ever move him or her will not have the slightest bearing on the soundness of the refutation. A proponent of the argument from evil would similarly maintain that the assertion of the existence of an infinite anthropomorphic deity has certain publicly testable consequences—that there is no evil in the world or at least not certain kinds of evil—and that experience shows these to be false. It would be to the point to argue either that the assertion of the existence of such a deity does not really have the consequences in question or that experience does not really falsify them; but it is totally beside the point to maintain either that faith in an infinite anthropomorphic God is not, in the case of a particular believer, based on any evidence or that the believer will not abandon his or her position, come what may.

REJECTION OF METAPHYSICAL THEOLOGY

In presenting the case against metaphysical theology, we shall concentrate on the views of Tillich and his disciple, Bishop J. A. T. Robinson, whose *Honest to God* created such a stir among theologians when it was published in 1963. No defender of this position had as much influence in the mid-twentieth century as Tillich. Moreover, his statement of the position is radical and uncompromising and is thus easier to discuss than more qualified versions. At the same time it may well be the case that some of these more qualified versions are not open to quite the same objections. In particular, it might be claimed that the Thomistic doctrine of analogy enables its proponents to escape both the difficulties of straightforward anthropomorphic theology and those besetting Tillich's position.

Tillich and Robinson entirely agree with atheists that belief in any anthropomorphic deity should be rejected. Traditional theism, Tillich writes, "has made God a heavenly, completely perfect person who resides above the world and mankind" (*Systematic Theology,* Vol. I, p. 271). Against such a highest person, he goes on, "the protest of atheism is correct." Elsewhere Tillich repeatedly pours scorn on what he terms "monarchic monotheism" and the theology of the "cosmic policeman." Following Tillich, Bishop Robinson tells us that we must now give up belief in God as somebody "out there," just as Copernican astronomy made people abandon "the old man in the sky." Most believers, he writes, are inclined to think of God as a kind of "visitor from outer space" (*Honest to God,* p. 50). Unlike the "old man in the sky" or the "visi-

tor from outer space," the God of Tillich and Robinson is not another individual entity beside the familiar entities of experience, not even the "most powerful" or the "most perfect" one. He is "being-itself." As such, God is not contingent but necessary, and arguments for his existence are not required. The idea of God, writes Tillich, is not the idea of "something or someone who might or might not exist" (*Systematic Theology,* Vol. I, p. 205). "In making God an object besides other objects, the existence and nature of which are matters of argument, theology supports the escape to atheism.… The first step to atheism is always a theology which drags God down to the level of doubtful things" (*Shaking of the Foundations,* p. 52).

It should be mentioned in passing that to some readers of Tillich and Robinson there appears to be a radical ambiguity in their entire position, specifically in the reasons they give for rejecting the anthropomorphic theory of the God "out there." At times we are told that the old-fashioned believers are mistaken because God is really inside us—insofar as our lives have "depth," insofar as we live "agapeistically." This is what we may call the Feuerbachian tendency in Tillich and his followers. At other times anthropomorphic theology is denounced because God so radically transcends anything we ever experience that the picture of a glorified man cannot possibly do justice to the reality. In the former context, God must not be said to be "out there" because he is really "in here deep down," in the latter context, because he is too removed to be even out there. In the former context, theological sentences become a species of very special psychological statements, and in the latter they are clearly items of transcendent metaphysics. There seems to be a constant oscillation between these two positions, so that at times traditional theology is denounced for not being sufficiently this-worldly, while at other times it is condemned for being too close to the world. The former position is of no interest to us, since it may rightly be dismissed as not being in any accepted sense a theological position at all—it is clearly quite compatible with the most thoroughgoing positivism and atheism. Our discussion will therefore be confined to the latter position exclusively.

As already explained in a previous section, Tillich (that is, Tillich the transcendent metaphysician) regards God as so vastly transcending any finite, familiar entity that predicates taken from ordinary experience cannot be employed in their literal senses when applied to God but must be used symbolically or metaphorically. There is just one statement that we can make about God in which all words are used "directly and properly," namely, that "God as being-itself is the ground of the ontological structure of being without being subject to the structure himself." Tillich expands this statement as follows: "God *is* that structure; that is, he has the power of determining the structure of everything that has being" (*Systematic Theology,* Vol. I, p. 239). If anything is said beyond this "bare assertion," Tillich insists it cannot be regarded any longer as a "direct and proper statement." Although all other predicates must be used symbolically when applied to God, certain symbols are justified or appropriate, while others are unjustified or inappropriate, since the former "point" to aspects of the ultimate reality, while the latter do not. Thus, we are justified in speaking of God, symbolically, as "King," "father," and "healing." These are "pointers to the "divine life."

UNINTELLIGIBILITY OF METAPHYSICAL THEOLOGY. A philosophically sophisticated atheist would object to Tillich's theology not on the ground that it is false or not proven but on the very different ground that it is unintelligible—that it consists of sentences that may be rich in pictorial associations and in expressive meaning but that fail to make any genuine assertions. Tillich's position may indeed be immune to the difficulties of an anthropomorphic theology, but only at the expense of not saying anything about the world. This criticism would almost certainly be offered by anybody who accepts an empiricist criterion of meaning, but it is worth pointing out that it is an objection that has been endorsed, in substance if not in precisely these words, by numerous believers in an anthropomorphic God. Voltaire on occasion objected on such grounds to the theologians who claimed that we must not use words in their familiar senses when applying them to God, and it has already been mentioned that Unamuno dismissed the metaphysical God as a "Nothing" and a "dead thing." Similarly, William James objected to the emptiness of the "universalistic" theology of the Hegelians of his day, preferring what he called a particularistic belief.

Untranslatable metaphors. This criticism might be backed up in the following way: While recognizing that he constantly uses words symbolically or metaphorically, Tillich does not appreciate the difference between translatable and untranslatable metaphors, and he does not see that his own metaphors are untranslatable. Very frequently indeed, especially in ordinary life, when words are used metaphorically, the context or certain special conventions make it clear what is asserted. Thus, the editor of an encyclopedia, when asked why he or she looks so troubled, may reply, "Too many cares are weighing down on me—the pressure is too great." Obviously the words *weighing down* and *pressure* are here metaphorical, yet we

all understand what is being said. Why? Because the metaphorical expressions are translatable—because we can eliminate them, because we can specify in non-metaphorical terms what the sentence is used to assert. If the metaphors could not be eliminated, we would not have succeeded in making any assertion.

A critic would proceed to argue that Tillich's metaphors are of the untranslatable variety and that when he has offered what seem to him translations, he has really only substituted one metaphor for another. Tillich believed that in his basic statement, quoted earlier, all words are used literally, or "properly." But this is open to question. The word *ground,* for example, is surely not used in any of its literal senses when being-itself is said to be the ground of the ontological structure of being. It can hardly be used in the physical sense in which the floor or the grass underneath our feet could be regarded as a "ground," or in the logical sense in which the premises of an argument may be the ground for endorsing the conclusion. Similar remarks apply to the use of *structure, power,* and *determine.* Hence, when we are told that "God is personal" (which is acknowledged to be metaphorical) means "God is the ground of everything personal," or that "God lives" (which is also acknowledged to be metaphorical) means "God is the ground of life," one set of metaphors is exchanged for another, and literal significance is not achieved. Tillich's God, it should be remembered, is so transcendent that not even mystical experience acquaints us with him. "The idea of God," he writes, "transcends both mysticism and the person-to-person encounter" (*The Courage To Be,* p. 178). Consequently, he does not have at his disposal any statements in which God is literally characterized and that could serve as the translations of the metaphorical utterances. The absence of such statements literally characterizing being-itself equally prevent Tillich from justifying the employment of his set of "symbols" as appropriate and the rejection of other symbols as inappropriate.

Unfalsifiability of metaphysical theology. We noted earlier that a metaphysical theology like Tillich's avoids the troublesome problem of evil because it does not maintain that God is perfectly good or, indeed, omnipotent in any of the ordinary or literal senses of these words. This very immunity would, however, be invoked by some critics as a decisive objection and they would, by a somewhat different route, reach the same conclusion—namely, that Tillich's theological sentences do not amount to genuine assertions. The point in question may perhaps be most forcefully presented by contrasting Tillich's position with that of anthropomorphic believers such as John Hick or A. C. Ewing. Hick and Ewing are (theoretically) very much concerned with the problem of evil. They argue that given the nature of man and a world with dependable sequences (or causal laws), evil of certain kinds is unavoidable, and furthermore that (though they do not, of course, claim to be able to prove this) in the next life there will be appropriate rewards and compensations. They admit or imply that their belief would be logically weakened, perhaps fatally so, it if could be shown that there is no afterlife or that in the afterlife injustice and misery, far from vanishing, will be even more oppressive than in the present life, or that the evils which, given the nature of man and a world of dependable sequences, they thought to be unavoidable, could in fact have been prevented by an omnipotent Creator. Tillich, however, need not be (theoretically) concerned about any such contingencies. Even if things in this life became vastly more horrible than they already are, or even if we had conclusive evidence that in the afterlife things are so bad that by comparison, Auschwitz and Belsen were kingdoms of joy and justice, Tillich's theology would be totally unaffected. Being-itself, as Tillich put it, would still be "actual": It is not "something or someone who might or might not exist." God, as Bishop Robinson puts it, is not a "problematic" entity, which might conceivably not have been there." This is true of the anthropomorphic deity, but not of what Tillich in one place terms "the God above God" (*Listener,* August 1961, pp. 169ff.).

In other words, unlike the position of Hick and Ewing, Tillich's theology is compatible with anything whatsoever in this life as well as in the next one; and it is the opinion of many contemporary philosophers, believers as well as unbelievers, that if a putative statement is compatible with anything whatsoever, if it excludes no conceivable state of affairs, then it is not a genuine assertion (it should be noted that "state of affairs" is not used in a narrow way so that much that positivists exclude, for example, happiness or suffering in the next world, could count as conceivable states of affairs). This criterion may, of course, be questioned, but if it is accepted, then Tillich's theology, unlike that of anthropomorphic believers, would have to be condemned as devoid of any assertive force.

We have not here considered other variants of metaphysical theology, but those opposed to Tillich's system for the reasons here outlined would maintain that other forms of this general outlook are bound to be open to some of the same objections: In every case, words would have to be used in a metaphorical way in crucial places, and these metaphors would turn out to be untranslatable;

in every case it would be impossible to justify the employment of one set of metaphors or symbols in preference to another, and in every case the author of the system would be unable to specify what conceivable state of affairs is excluded by his sentences or, if he did do so, the exclusion could be shown to be arbitrary in a way that would not be true of the statements of anthropomorphic believers.

ATHEISM OR AGNOSTICISM?

It is time to discuss a very common challenge to atheists. The challenge is usually issued by agnostics, but it would in general also be endorsed by fideistic believers. "It is admittedly impossible," the critic would reason, "to prove the existence of God, but it is equally impossible to disprove his existence; hence, we must either suspend judgment or, if we embrace some position, we must do so on the basis of faith alone." To avoid misleading associations of the words *prove* and *disprove,* the same point may be expressed by saying that we have no evidence either for or against God's existence. Sometimes the reminder is added that the mere failure of the arguments for the existence of God does not show that there is no God. Anybody who supposed this would plainly be guilty of the fallacy of *argumentum ad ignorantiam.*

If certain of the considerations advanced by atheists that were discussed in previous sections are sound, this agnostic charge would be quite beside the point as far as belief in an infinite anthropomorphic or a metaphysical God is concerned. For in that event, the first theory can be shown to be false (with certain qualifications explained earlier), and the second can be rejected on the ground that it is unintelligible. In the case of an infinite anthropomorphic God, there is evidence against the position; in the case of a metaphysical God, we do not have a coherent position. However, when we turn to the question of a finite anthropomorphic God, the challenge does at first sight seem very plausible. As already pointed out, the argument from evil does not affect this position, and we may, at least provisionally, grant that belief in a finite anthropomorphic God is intelligible because the predicates used in expressing it are applied to this deity in their familiar senses. We shall see, before long, that there are difficulties in regard to the intelligibility of even this position, but waiving all considerations of this kind for the moment, let us inquire how an atheist could reply to this challenge. It is admitted by the challenger that there is no evidence for the existence of such a deity; where, he asks, is the evidence against its existence? If there is none, why should one be an atheist rather than an agnostic? Why is atheism justified if we cannot be sure that there is no God in the sense under discussion?

GROUNDS FOR THE REJECTION OF THEORIES. In justifying his position, an atheist should perhaps begin by calling attention to the fact that the agnostics who suspend judgment concerning God are not also agnostics in relation to the gods of the Greeks or in relation to the devil and witches. Like the majority of other educated people, most agnostics reject and do not suspend judgment concerning the Olympian gods or the devil or witches. Assuming that rejection is the appropriate attitude in these cases, what justifies this rejection?

It will be instructive to look at a concrete example of such a belief that is rejected by agnostics and atheists alike and, incidentally, by most believers in God. Billy Graham is one of the few Protestant ministers who still believe in the devil. The devil is introduced by Dr. Graham as the only plausible explanatory principle of a great many phenomena. He is brought in to explain the constant defeat of the efforts of constructive and well-meaning people, the perverse choices of men who so commonly prefer what is degrading to what is "rich and beautiful and ennobling," the speed with which lies and slander spread in all directions, and also the failure of the world's diplomats. "Could men of education, intelligence, and honest intent," asks Dr. Graham, "gather around a world conference table and fail so completely to understand each other's needs and goals if their thinking was not being deliberately clouded and corrupted?" All such failures are "the works of the devil" and they show that he "is a creature of vastly superior intelligence, a mighty and gifted spirit of infinite resourcefulness." The devil is no "bungling creature" but "a prince of lofty stature, of unlimited craft and cunning, able to take advantage of every opportunity that presents itself" (*Peace with God,* New York, 1954, pp. 59–63).

What reasons could or would be given for rejecting this explanation of diplomatic failures in terms of the devil's cunning ways? Aside from possibly questioning some of Dr. Graham's descriptions of what goes on in the world, that is, of the "facts" to be explained, our reasons would probably reduce to the following: First, we do not need to bring in the devil to explain the failure of diplomats to reach agreement on important international issues. We are confident, on the basis of past experience, that explanations of these failures in terms of human motives, in terms of human ignorance and miscalculation, are quite adequate, although in any particular case we may not be in the possession of such an explanation;

and, second, the devil hypothesis, granting it to be intelligible, is too vague to be of any use. It is hinted that the devil has a body, but what that body is like or where it lives and exactly how it operates, we are not told. If "devil" is construed on the analogy of the theoretical terms of the natural sciences, our complaint would be that no, or none but totally arbitrary, correspondence rules have been assigned to it.

It should be observed that the devil theory is rejected although it has not been tested and, hence, has not been falsified in the way in which certain exploded medical theories have been tested and falsified. There are, in other words, theories that we reject (and which agnostics, like others, believe they have good reason to reject), although they have not been falsified. It is important to distinguish here two very different reasons why a theory may not have been tested and, hence, why it cannot have been falsified. The theory may be sufficiently precise for us to know what would have to be done to test it, but we may be chronically or temporarily unable to carry out any of the relevant tests. This is to be sharply contrasted with the situation in which a theory is so vague that we do not know what we must do to subject it to a test. In the former case, suspension of judgment may well be the appropriate attitude; it does not follow that the same is true in the latter case, and in fact most of us regard rejection as the appropriate attitude in such a situation until and unless the theory is stated with more precision.

An atheist would maintain that we have just as good grounds for rejecting belief in a finite anthropomorphic deity of any sort as we have for rejecting belief in Zeus or in the devil or in witches. It should be noted that the believers in the finite anthropomorphic God usually advance their theory as a hypothesis that is the best available explanation of certain facts. Mill, for example, thought that the Design Argument, in the form in which he advocated it, affords "a large balance of probability in favor of creation by intelligence," although he conceded that new evidence for the Darwinian theory would alter this balance of probability (*Three Essays on Religion,* New York, 1874, p. 174). An atheist would argue that we do not need a finite God to account for any facts any more than we need the devil theory; and, more important, that the theory is too vague to be of any explanatory value. Mill, for example, talks of "creation by intelligence," but he does not tell us in any detail what the "Author of Nature" is like, where he can be found, how he works, and so on. Furthermore, because of its vagueness the theory is totally sterile. It does not lead to subsidiary hypotheses about celestial laboratories or factories in which eyes and ears and other organs are produced. Nor does it help us to interpret fossils or other remains here on earth. It is tempting, but it would be misleading, to say that the accumulation of evidence for the Darwinian theory (or some modified version of it) since Mill wrote on the subject has put the design theory "out of court." This would suggest that the theological explanation was at some time "in court," in the way in which a falsified scientific explanation may once have been a serious contender. It is true, of course, as a matter of history, that informed people cease to bring in God as an explanation for a given set of phenomena once a satisfactory scientific or naturalistic explanation is available. In a more important sense, however, the theological explanations were never serious rivals, just as the devil explanation of diplomatic failures is not a serious rival to psychological explanations. The theological explanations never were serious rivals because of their excessive vagueness and their consequent sterility. We do not at present have anything like a satisfactory scientific explanation of cancer, but no theological theory would be treated as a genuine alternative by a cancer researcher, even a devoutly religious one.

It should be added to all this that believers who, unlike Mill, do not treat their theology as a kind of hypothesis, are not affected by the above objections. Indeed, quite a number of them have strenuously opposed any kind of "God of the gaps." However, some of the very writers who insist that their theology must not be regarded as a scientific hypothesis elsewhere make statements that imply the opposite. They also frequently maintain that certain phenomena—for example, the universal hunger for God or the origin of life—can be explained only, or can be explained best, on the assumption that there is a God, and a God of a certain kind. Whatever they may say on other occasions, insofar as they propose their theology as the only possible, or as the best available, explanation of such phenomena, they are committed to the position that has been criticized in this section.

THE DEMAND FOR A COSMIC BRAIN

There was a good deal of discussion in the late nineteenth century of an antitheological argument that ought to be briefly mentioned here. To many persons, including unbelievers, the argument will seem to be merely grotesque; but in view of the revival in more recent years of several forms of extreme materialism, it deserves some discussion. Moreover, even if it is granted that the argument fails to prove its conclusion, the very grotesqueness of some of its formulations enables a more sophisticated

contemporary atheist to state a challenge in a particularly forceful way.

The two writers chiefly associated with this argument were the German physiologist Emil Du Bois–Reymond and the English mathematician W. K. Clifford, both of whom wrote extensively on philosophical subjects. However, the argument is really much older, and versions of it are found in Meslier and Holbach. The remark attributed to Pierre Simon de Laplace that "in scanning the heavens with a telescope he found no God" may be regarded as an argument belonging to the same family. "Can we regard the universe," asked Clifford in his essay "Body and Mind," "or that part of it which immediately surrounds us, as a vast brain, and therefore the reality which underlies it as a conscious mind? This question has been considered by the great naturalist, Du Bois–Reymond, and has received from him that negative answer which I think we also must give." The student of nature, Du Bois–Reymond had written, before he can "allow a psychical principle to the universe," will demand to be shown "somewhere within it, embedded in neurine and fed with warm arterial blood under proper pressure, a convolution of ganglionic globules and nerve-tubes proportioned in size to the faculties of such a mind" (*Über die Grenzen des Naturerkennens,* p. 37). But, in fact, no such gigantic ganglionic globules or nerve-tubes are discoverable, and, hence, we should not allow a "psychical principle" to the universe. The following would be a more systematic statement of the argument: Experience shows that thinking, volition, and other psychological phenomena do not and cannot occur without a certain physiological basis—more specifically, without a brain and nervous system. Our observations appear to indicate, although this is not a matter of which one can be certain, that no cosmic brain or nervous system exists. Hence, it is probable that no cosmic consciousness exists either.

This argument has been criticized on the ground that it assumes a certain view (or a certain group of views) about the relationship between body and mind that is not self-evidently true and that many believers would deny. It assumes that consciousness can exist only in conjunction with a nervous system and a brain. However, the objector would maintain, the actual evidence on the subject does not warrant such a claim. It is true that within our experience, conscious processes are found only in connection with a highly developed brain, but this does not prove that consciousness may not occur in conjunction with other physical structures or without any physical "attachments" whatsoever. This is a big question about which nothing very useful can be said in a few words. Perhaps all we can do here is point out that if materialism of some kind is true, then the demand to be shown the bodily foundation or aspect of the divine consciousness is not misplaced, while if the opposite view that consciousness can exist independently of a physical structure is correct, the Du Bois–Reymond argument would have no force.

Quite aside from this objection, the argument probably seems to many people, believers and unbelievers alike, to rest on a total, one is almost inclined to say a willful, misunderstanding of the theological position. James Martineau, who replied at some length to Du Bois–Reymond, protested that the "demand for organic centralization" was "strangely inappropriate," indeed quite irrelevant to the question at issue between the believer and the unbeliever. If Du Bois–Reymond himself, wrote Martineau, were "ever to alight on the portentous cerebrum which he imagines, I greatly doubt whether he would fulfill his promise and turn theist at the sight: that he had found the Cause of causes would be the last inference it would occur to him to draw: rather would he look round for some monstrous *creature,* some cosmic megatherium, born to float and pasture on the fields of space" (*Modern Materialism and Its Relation to Religion and Theology,* p. 184). Martineau then likened the argument to Laplace's remark, mentioned earlier, that in looking at the heavens with his telescope, he could nowhere see God and to statements by certain physiologists that in opening the brain, they could not discover a soul. All such pronouncements Martineau regarded as absurd. Although the physiologist finds no soul when he opens up the brain, "we positively know" (by introspection) the existence of conscious thought. Similarly, that "the telescope misses all but the bodies of the universe and their light" has no tendency to prove "the absence of a Living Mind through all." If you take the "wrong instruments" you will not find what you are looking for. "The test tube will not detect an insincerity," nor will "the microscope analyse a grief"; but insincerity and grief are real for all that. The organism of nature, Martineau concludes, "like that of the brain, lies open, in its external features, to the scrutiny of science; but, on the inner side, the life of both is reserved for other modes of apprehension, of which the base is self-consciousness and the crown is religion."

One is strongly inclined to agree with Martineau that there is something absurd in scanning the heavens for God. Étienne Borne, a French Catholic whose discussions are distinguished by fairness and sympathy for the opposition, refers to this approach as "a tritely positivist atheism" that "misses the point of the problem altogether" (*Modern Atheism,* p. 145). One must not expect to find

God or God's body in the heavens because God is not a huge man with huge arms, legs, arteries, nervous system, and brain. Only children think of God as a "king" sitting on his throne in Heaven. Educated grownups do not think of God in any such crude fashion. Du Bois–Reymond, Clifford, and Laplace are all guilty of an enormous ignoratio elenchi.

IS ANTHROPOMORPHIC THEOLOGY INTELLIGIBLE? Let us grant the force of Borne's objection. A critic may nevertheless raise the following questions: What is God like if he is not a grand consciousness tied to a grand body, if he is so completely nonphysical as to make any results of telescopic exploration antecedently irrelevant? If the telescope, as Martineau put it, is the "wrong instrument," what is the right instrument? More specifically, what does it mean to speak of a pure spirit, a disembodied mind, as infinitely (or finitely) powerful, wise, good, just, and all the rest? We can understand these words when they are applied to human beings who have bodies and whose behavior is publicly observable; we could undoubtedly understand these words when they are applied to some hypothetical superhuman beings who also have bodies and whose behavior is in principle observable; but what do they mean when they are applied to a pure spirit? Do they then mean anything at all? In recent years it has come to be widely questioned whether it makes any sense to talk about a disembodied consciousness. It is widely believed, in other words, that psychological predicates are logically tied to the behavior of organisms. This view, it should be pointed out, is not identical with reductive materialism. It does not, or at least does not necessarily, imply that the person is just a body, that there are no private experiences, or that feelings are simply ways of behaving. It makes the milder claim that however much more than a body a human being may be, one cannot sensibly talk about this "more" without presupposing (as part of what one means, and not as a mere contingent fact) a living organism. Anybody who has studied and felt the force of this thesis is not likely to dismiss as facetious or as "trite positivism" the question as to what words such as *wise, just,* and *powerful* can mean when they are applied to an entity that is supposedly devoid of a body. What would it be like to be, for example, just, without a body? To be just, a person has to act justly—to behave in certain ways. But how is it possible to perform these acts, to behave in the required ways, without a body? Similar remarks apply to the other divine attributes.

One may term this the "semantic" challenge to anthropomorphic theology, as distinct, for example, from arguments like the one from evil or from the eternity of matter, which assume the meaningfulness of the position attacked. A proponent of this challenge does not flatly maintain that anthropomorphic theology is unintelligible. For the point—that the predicates in question lose their meaning when applied to a supposedly disembodied entity—would be accompanied by the observation that in fact most anthropomorphic believers do, in an important sense of the word, believe in a god with a body, whatever they may say or agree to in certain "theoretical" moments. If we judge the content of their belief not by what they say during these "theoretical" moments but by the images in terms of which their thinking is conducted, then it seems clear that in this sense or to this extent they believe in a god with a body. It is true that the images of most Western adults are not those of a big king on his heavenly throne, but it nevertheless seems to be the case that, when they think about God unself-consciously (and this is, incidentally, true of most unbelievers also), they vaguely think of him as possessing some kind of rather large body. The moment they assert or deny or question such statements as "God created the universe" or "God will be a just judge when we come before him," they introduce a body into the background, if not into the foreground, of their mental pictures. The difference between children and adults, according to this account, is that children have more vivid and definite images than adults.

This entire point may perhaps be brought out more clearly by comparing it with a similar "semantic" criticism of belief in human survival after death. The semantic critic would maintain that while a believer in reincarnation or the resurrection of the body may be immune from this objection, those who claim that human beings will continue to exist as disembodied minds are really using words without meaning. They do not see this because of the mental pictures accompanying or (partly) constituting their thoughts on the subject. Or, alternatively, they do not see this because, in spite of what they say in certain "theoretical" contexts, in practice they believe in the survival of the familiar *em*bodied minds whom they know in this life. When they wonder whether their friends, enemies, certain historical personages, or, for that matter, anybody did or will go on existing after death, they think of them automatically in their familiar bodily "guises" or else in some ghostly "disguises," but still as bodily beings of some kind. If these images are eliminated on the ground that they are irrelevant or inappropriate because the subject of survival is a disembodied mind, it is not clear that an intelligible statement remains. What, for example, do such words as *love* and *hate* or *hap-*

piness and *misery* mean when they are predicated of a disembodied mind?

It will be seen from all this that the argument of Du Bois–Reymond and Clifford is not without some point. One may incorporate what is of value in their discussion into the following challenge to anthropomorphic theology: Insofar as the believer believes in a god with a body, what he or she says is intelligible; but in that case the available evidence indicates that there is no such body, and the remarks of Du Bois–Reymond and Clifford are to the point; if or insofar as God is declared to be a purely spiritual entity, the observations of Du Bois–Reymond and Clifford become irrelevant, but in that case the predicates applied to God have lost their meaning, and, hence, we no longer have an intelligible assertion.

SUMMARY OF THE ATHEIST'S POSITION. Let us summarize the atheist's case as it has here been presented. A philosophically sophisticated atheist would begin by distinguishing three types of belief in God—what we have called the metaphysical God, the infinite anthropomorphic God, and the finite anthropomorphic God. He will then claim that he can give grounds for rejecting all three, although he does not claim that he can prove all of them to be false. He will try to show that metaphysical theology is incoherent or unintelligible, and, if he can do this, he will certainly have given a good ground for rejecting it. He will also question the intelligibility of anthropomorphic theology insofar as God is here said to be a purely spiritual entity. If and insofar as belief in an infinite anthropomorphic God is intelligible, he will maintain that it is shown to be false by the existence of evil. In the sense in which he will allow the existence of a finite anthropomorphic God to be an intelligible hypothesis, he will argue that it should be rejected because it is not needed to account for any phenomena and, further, because it is too vague to be of any explanatory value. We saw that some of these justifications, even if sound as far as they go, would not establish the atheist's case unless they are accompanied by a demolition of the arguments for the existence of God.

SOME OBJECTIONS TO ATHEISM

If there were reason to believe that any of the arguments for the existence of God are sound or have at least some tendency to establish their conclusions, then they would of course constitute objections to atheism. Since these arguments are fully discussed elsewhere in this encyclopedia, we shall here confine ourselves to objections that are logically independent of them. Some of these objections have been put forward by writers who explicitly reject all the traditional proofs but nevertheless regard atheism as an untenable position.

THE MYSTERY OF THE UNIVERSE. It has been argued by several writers that whatever the objections to the different forms of theology may be, atheism is also unacceptable since it has no answer to the "ultimate question" about the origin of the universe. Thus, the nineteenth-century physicist John Tyndall, after endorsing a thoroughgoing naturalism, proceeded to reject atheism in favor of an agnostic position. In a paper titled "Force and Matter," he tells the story of how Napoleon turned to the unbelieving scientists who had accompanied him to Egypt and asked them, pointing to the stars, "Who, gentlemen, made all these?" "That question," Tyndall comments, "still remains unanswered, and science makes no attempt to answer it." Later he adds that "the real mystery of this universe lies unsolved, and, as far as we are concerned, is incapable of solution" (*Fragments of Science,* pp. 92–93). In much the same vein, the celebrated American freethinker and social reformer Clarence Darrow, after pointing out the weaknesses of the First Cause Argument, observed that the position of the atheist is just as vulnerable. If, he wrote, the atheist answers the question "What is the origin of it all?" by saying that the universe always existed, he has the same difficulty to contend with as the believer has when he is asked the question "Who made God?" To say that "the universe was here last year, or millions of years ago, does not explain its origin. This is still a mystery. As to the question of the origin of things, man can only wonder and doubt and guess" (*Verdicts out of Court,* pp. 430–431).

A philosophically acute atheist could offer a twofold answer to arguments of this kind. First, he would maintain that the question about the "origin of the universe" or the "origin of it all" is improper and rests on the mistaken or doubtful assumption that there is a thing called "the universe." It is tempting to suppose that there is such a thing because we have a tendency to think of the universe as a large container in which all things are located and, perhaps more important, because grammatically the expression functions analogously to expressions like "this dog" or "the Cathedral of Notre Dame," which do denote certain things. Upon reflection, however, it becomes clear, the rejoinder would continue, that "the universe" is not a thing-denoting expression or, putting the point differently, that there is not a universe over and above the different things within the universe. While it makes sense to ask for the origin of any particular thing, there is not a further thing left over, called "the universe" or "it all," into

whose origin one can sensibly inquire. The origin of a great many things is of course unknown to us, but this is something very different from "the ultimate mystery" that figures in the argument under discussion; and there is no reason to suppose that questions about the origin of any individual thing fall in principle outside the domain of scientific investigation.

Furthermore, even if it is granted both that the question concerning the origin of the universe is proper and that we do not and cannot discover the true answer, this is not by itself an argument against atheism. It may well be possible to know that a certain suggested answer to a question is false (or meaningless) without knowing the true answer. All kinds of crimes have never been solved, but this does not prevent us from knowing that certain people did not commit them. An atheist can quite consistently maintain "I have no idea how the origin of the universe is to be explained, but the theological theory cannot be the right answer in view of such facts as the existence of evil." To support his position, the atheist must be able to justify his rejection of theological answers to the question "What is the origin of the universe?" He does not have to be able to answer that question.

ATHEISM PRESUPPOSES OMNISCIENCE. In the popular apologetic pronouncements of liberal believers, it is customary to contrast the agnostic, who is praised for his circumspection, with the atheist, who is accused of arrogant dogmatism and who, like the orthodox or conservative believer, claims to know what, from the nature of the case, no mere human being can possibly know. "The atheist," in the words of Dr. W. D. Kring, a twentieth-century Unitarian, "can be just as closed-minded as the man who knows everything. The atheist just knows everything in a negative direction" (*New York Times,* March 22, 1965).

Reasoning of this kind figured prominently in several influential works by nineteenth-century Protestant theologians. Their favorite argument was the following reductio ad absurdum: Atheism could be known to be true only if the atheist knew everything; but this is of course impossible; hence, atheism cannot be known to be true. For a man to deny God, wrote Thomas Chalmers, "he must be a God himself. He must arrogate the ubiquity and omniscience of the Godhead." Chalmers insists that the believer has a great initial polemical advantage over the atheist. For, he argues, some very limited segment of the universe may provide the believer with strong or even decisive evidence, with an "unequivocal token" of God's existence. The atheist, on the other hand, would have to "walk the whole expanse of infinity" to make out his case (*On Natural Theology,* Vol. I, Book I, Ch. 2). By what miracle, asks John Foster, can an atheist acquire the "immense intelligence" required for this task? Unless he is "omnipresent—unless he is at this moment at every place in the universe—he cannot know but there may be in some place manifestations of a Deity by which even *he* would be overpowered." And what is true of space equally applies to "the immeasurable ages that are past" (*Essays,* 18th ed., p. 35). The atheist could not know that there is no God unless he had examined every part of the universe at every past moment to make sure that at no time was there a trace of divine activity.

According to Robert Flint, who endorsed and elaborated the arguments of Chalmers and Foster, the situation should be clear to anybody who reflects on the difficulty of "proving a negative." If a man landed on an unknown island, any number of traces in almost any spot would be sufficient to show that a living creature had been there, but he would have to "traverse the whole island, examine every nook and corner, every object and every inch of space in it, before he was entitled to affirm that *no* living creature had been there" (*Anti-Theistic Theories,* pp. 9–11). The larger the territory in question, the more difficult it would become to show that it had not a single animal inhabitant. If, then, it is "proverbially difficult to prove a negative," there can surely "be no negative so difficult to prove as that there is no God." This is plain if we reflect that "before we can be sure that nothing testifies to His existence, we must know all things." The territory in this case is "the universe in all its length and breadth." To know that there is no trace of God anywhere in eternal time and boundless space, a man would have had to examine and to comprehend every object that ever existed. This would indeed require omnipresence and omniscience, and Chalmers was there perfectly right when he maintained that the atheist's claim implies that "he is himself God."

Whatever its rhetorical force, this argument is so patently invalid that it can be disposed of in just a few words. We have in preceding sections of this entry presented several of the most widely used arguments and considerations that have been advanced in support of atheism. These may or may not be logically compelling, but none of them in any way imply that the atheist must be omniscient if he is right. To establish that the existence of evil is incompatible with the view that the universe is the work of an all-powerful and all-good Creator, to show that a given theory is too vague to be of any explanatory value, or to call attention to the fact that certain words

have in a certain context lost their meaning—none of these require omniscience.

Writers like Chalmers, Foster, and Flint seem to labor under the impression that as far as its refutability is concerned, "God exists" is on par with a statement like "A hippogriff exists, existed, or will exist in some place at some time." It may be plausible to maintain that our not having found any hippogriffs on earth is no conclusive evidence that such an animal does not exist in some other part of the universe to which we have no access. The same does not at all apply to the question of whether one is or can be entitled to reject the claims of believers in God. For, unlike the hippogriff, God is by some declared to be the all-powerful and all-good Creator of the universe; he is said by most believers to be a mind without a body; and it is asserted by some that predicates taken from ordinary experience can never be applied to God in their literal senses. These features of theological claims may make it possible to justify their rejection although one has not explored every "nook and cranny" of the universe.

ATHEISM, ZEAL, AND GLOOM

In the opening section of this entry we referred to the view, common in previous centuries, that atheism is bound or, at any rate, very likely to lead to immorality, to national ruin, and to other disasters. This warning is no longer taken very seriously among reputable thinkers, but certain other statements about the baleful consequences of unbelief in general and atheism in particular continue to be widely discussed. Thus, it is frequently maintained that if atheism were true or justified, life would be deprived of all meaning and purpose. Again, it has been held that without God the universe becomes "terrifying" and man's life a lonely and gloomy affair. "Old age," wrote William James in his *Varieties of Religious Experience* (New York and London, 1902), "has the last word: a purely naturalistic look at life, however enthusiastically it may begin, is sure to end in sadness." Blaise Pascal, who was particularly concerned about the terror of a "silent universe" without God, observed in a similar vein that "the last act" is always tragic—"a little earth is thrown upon our head, and that is the end forever."

James and Pascal were believers, but very similar statements have frequently come from unbelievers themselves. "I am not ashamed to confess," wrote G. J. Romanes, a nineteenth-century biologist, at the end of his *A Candid Examination of Theism* (a work that was published anonymously in London in 1878 and which caused a commotion at the time), "that with this virtual denial of God, the universe has lost to me its soul of loveliness."

More recently, the anthropologist Bronislaw Malinowski spoke of the state of mind of an unbeliever like himself as "tragic and shattering." Not only does the absence of God, in the opinion of these writers, make the universe "lonely," "soulless," and "tragic," but it also deprives it of love. Only when we have become accustomed to a "loveless" as well as a "Godless universe," in the words of Joseph Wood Krutch, shall "we realize what atheism really means."

Finally, it has been claimed that atheism is fatal to what William James called the capacity of the strenuous mood. James himself had no doubt that the unbeliever is prevented from "getting out of the game of existence its keenest possibilities of zest." Our attitude toward concrete evils, he asserted, "is entirely different in a world where we believe there are none but finite demanders, from what it is in one where we joyously face tragedy for an infinite demander's sake." Religious faith sets free every kind of energy, endurance, and courage in the believer and "on the battlefield of human history," religion will for this reason always "drive irreligion to the wall" (*The Will to Believe,* pp. 213ff.)

Some of these claims seem a great deal more impressive than others. It is not easy to deal with the charge that atheism deprives life of its meaning, chiefly because the word *meaning* in this connection is both ambiguous and extremely vague. However, if what is meant is that an atheist cannot be attached to certain goals that give direction to his life, then the charge is quite plainly false. If what is meant is that although the atheist may, like other men, pursue certain goals, he will not be able to justify any of his activities, then it should be pointed out that most human beings, even believers in God, do not justify the great majority of their acts by reference to God's will. Hence, the justification of these actions, if they ever are justified, could not be affected by the soundness of atheism. It is difficult to see how such activities as engaging in scientific research, assisting people who are in trouble, singing or dancing or making love or eating superb meals, if they ever were worthwhile, would cease to be so once belief in God is rejected. If what is meant by the charge is that the unbeliever will eventually have to fall back, in his justification, on one or more value judgments that he cannot justify by reference to anything more fundamental, this may be true, but it is not necessarily baleful, and it is not a consequence of atheism. Anybody who engages in the process of justifying anything will eventually reach a stage at which some proposition, principle, or

judgment will simply have to be accepted and not referred back to anything else. The unbeliever may, in justifying his acts, regard as fundamental such judgments as "happiness is intrinsically worthwhile" or "the increase of knowledge is good for its own sake," whereas some believers may say that only service of God is intrinsically valuable. If it is a sign of irrationality, which in any normal sense of the word it is not, to accept a value judgment that is not based on another one, then the atheist is not one whit more irrational than the believer.

On the question of zest, it should be observed that neither James nor anybody else has ever offered empirical evidence for the assertion that unbelievers lead less active or strenuous lives than believers. What we know about human temperament suggests that the acceptance or rejection of a metaphysical position has, in the case of the vast majority of men, exceedingly little to do with whether they lead active or inactive lives. The Soviet cosmonauts, who were atheists (to take one relatively recent illustration), appeared to display the same courage and endurance as their American counterparts, who were believers. In general terms, a survey of the contributions of atheists and other unbelievers to science and social progress, often in conditions requiring unusual stamina and fortitude, would seem to indicate that James was in error. The a priori character of James's views on this subject remind one of Locke's conviction, mentioned earlier in this entry, that atheists, since they do not fear divine punishment, cannot be trusted to keep oaths and promises.

As for the "loveless universe" presented by atheism, it must of course be admitted that if there is no God who loves his creatures, there would be that much less love in the world. But this is perhaps all that an atheist would have to concede in this connection. Aside from certain mystics and their raptures, it may be questioned whether a biologically normal human being is capable of feeling any real or deep love for an unseen power; and it hardly seems credible to suppose that a person will cease to love other human beings and animals (if he ever loved them) just because he does not believe them to be the work of God. Perhaps one may hazard a guess that if more human beings grow up in an environment that is free from irrational taboos and repressions (and these, one may add, have not been altogether unconnected with religious belief in the past), there will be more, not less, love in the world—people will be more lovable and will also be more capable of giving love. As far as love is concerned, the record of theistic religions has not been particularly impressive.

The writers whose views we are discussing have probably been on stronger ground when they maintain that atheism is a gloomy or tragic philosophy, but here too some qualifications are in order. To begin with, if atheism implies that life is gloomy, it does so not by itself but in conjunction with the rejection of the belief in life after death. There have been atheists, of whom J. E. McTaggart is probably the most famous, who believed in immortality, and they would deny that their atheism had any gloomy implications. However, since the great majority of atheists undoubtedly reject any belief in survival, this does not go to the root of the matter. It cannot be denied that the thought of annihilation can be quite unendurable; but it may be questioned whether believers, whatever they may be expected to feel, do in fact find the thought of death any less distressing. In the opinion of some observers, this is due to the fact that regardless of his profession, the believer frequently does not really believe that death is the gate to an eternal life in the presence of God. "Almost inevitably some part of him," in the words of Russell, is aware that beliefs of this kind are "myths and that he believes them only because they are comforting" (*Human Society in Ethics and Politics,* p. 207). Russell and Sigmund Freud regard belief in God and immortality as illusions that usually do not work, but they are quick to add that anybody who refuses to be the victim of unworthy fears would dispense with such illusions even if they did work. "There is something feeble and a little contemptible," in Russell's words, "about a man who cannot face the perils of life without the help of comfortable myths." Some years earlier, in an essay titled "What I Believe," Russell had put the point very bluntly:

> I believe that when I die I shall rot, and nothing of my ego will survive. I am not young, and I love life. But I should scorn to shiver with terror at the thought of annihilation. Happiness is nonetheless true happiness because it must come to an end, nor do thought and love lose their value because they are not everlasting.… Even if the open windows of science at first make us shiver after the cozy indoor warmth of traditional humanizing myths, in the end the fresh air brings vigor, and the great spaces have a splendor of their own.

See also Agnosticism; Analogy in Theology; Augustine, St.; Berkeley, George; Blanshard, Brand; Brightman, Edgar Sheffield; Carnap, Rudolf; Clifford, William Kingdon; Cosmological Argument for the Existence of God; Cudworth, Ralph; Du Bois-Reymond, Emil; Epicureanism and the Epicurean School; Evil, The Prob-

lem of; Existentialism; Feuerbach, Ludwig Andreas; Fichte, Johann Gottlieb; Freud, Sigmund; Hamilton, William; Hobbes, Thomas; Holbach, Paul-Henri Thiry, Baron d'; Huxley, Thomas Henry; Immortality; James, William; Jodl, Friedrich; Laplace, Pierre Simon de; Locke, John; Mansel, Henry Longueville; Martineau, James; Marx, Karl; McTaggart, John McTaggart Ellis; Meslier, Jean; Mill, John Stuart; Nagel, Ernest; Nietzsche, Friedrich; Nihilism; Paley, William; Pascal, Blaise; Plato; Popular Arguments for the Existence of God; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Schopenhauer, Arthur; Shelley, Percy Bysshe; Stephen, Leslie; Thomas Aquinas, St.; Tillich, Paul; Toleration; Unamuno y Jugo, Miguel de; Voltaire, François-Marie Arouet de.

Bibliography

The only full-length history of atheism in existence is Fritz Mauthner's four-volume work, *Der Atheismus und seine Geschichte im Abendlande* (Stuttgart: Deutsche Verlagsanstalt, 1920–1923). Although this work contains much interesting information that cannot easily be obtained elsewhere, it is marred by extreme repetitiousness and by a curiously broad use of the word *atheism,* which allows Mauthner to speak of agnostic and even deistic atheists. Probably of greater value are the various works on the history of free thought by J. M. Robertson, chiefly his *A Short History of Free Thought* (New York: Russell and Russell, 1899). Accounts of the struggles of atheists in England in the nineteenth century will be found in H. Bradlaugh Bonner, *Charles Bradlaugh: A Record of His Life and Work* (London: Unwin, 1895); G. J. Holyoake's two-volume *Sixty Years of an Agitator's Life* (London: Unwin, 1892); and A. H. Nethercot, *The First Five Lives of Annie Besant* (London, 1961).

An early defense of atheism is found in Vol. II of Holbach's two-volume *The System of Nature,* translated by H. D. Robinson (Boston: Mendum, 1853) and in his briefer work *Common Sense,* translated by A. Knoop (New York, 1920). Shelley defended atheism in his essays *The Necessity of Atheism* and *A Refutation of Deism,* and in one of the Notes to Canto VII of *Queen Mab,* titled "There is no God." All of these are included in *Shelley's Prose,* edited by D. L. Clark (Albuquerque: University of New Mexico Press, 1954). Charles Bradlaugh's "A Plea for Atheism" was first published in 1864 and reprinted in the Centenary Volume, *Charles Bradlaugh: Champion of Liberty* (London, 1933). Although he rarely used the term *atheism,* Schopenhauer is usually and quite properly classified as an atheist. His fullest discussion of the reasons for rejecting belief in God are found in his "The Christian System" and in his "Religion: A Dialogue." Both of these are available in a translation by T. B. Saunders in *Complete Essays of Schopenhauer* (New York: Willey, 1942). Another nineteenth-century work defending atheism is Ludwig Feuerbach, *The Essence of Christianity* (1841), translated by George Eliot, with an introduction by Karl Barth (New York: Harper, 1957). Of early critical works, special mention should be made of Ralph Cudworth's two-volume *The True Intellectual System of the World* (London, 1678), which is an enormously detailed onslaught on all forms of atheism known to the author, and of Voltaire's article "Atheism" in his *Philosophical Dictionary,* translated by Peter Gay (New York: Basic, 1962). Part II of Voltaire's article is an extended critique of *The System of Nature.*

In more recent years, atheism has been championed in R. Robinson, *An Atheist's Values* (Oxford: Clarendon Press, 1964); in Ernest Nagel, "A Defence of Atheism," which is available in *A Modern Introduction to Philosophy,* edited by Paul Edwards and Arthur Pap (New York: Free Press, 1965), and in Michael Scriven, *Primary Philosophy* (New York: McGraw-Hill, 1966). Rudolf Carnap's position, which is briefly mentioned in the present entry, is presented in his "The Elimination of Metaphysics through Logical Analysis of Language," which is available in a translation by Arthur Pap in *Logical Positivism,* edited by A. J. Ayer (Glencoe, IL: Free Press, 1959). A somewhat similar position is defended by Antony Flew in "Theology and Falsification." This paper is available in various anthologies, perhaps most conveniently in *The Existence of God,* edited by John Hick (New York: Macmillan, 1964). An interesting and unusual defense of theology against contemporary criticisms like those of Carnap and Flew is found in I. M. Crombie's "The Possibility of Theological Statements," in *Faith and Logic,* edited by Basil Mitchell (London: Allen and Unwin, 1957). The comments in the present entry about the attempts of fideists to circumvent the argument from evil and other difficulties are elaborated in Paul Edwards, "Is Fideistic Theology Irrefutable?" in *Rationalist Annual* (1966).

There is a kind of "ontological" argument for atheism proposed by J. N. Findlay in "Can God's Existence Be Disproved?"; this, together with various rejoinders, is reprinted in *New Essays in Philosophical Theology,* edited by Antony Flew and Alasdair MacIntyre (London: SCM Press, 1955). The view that belief in God is not false but self-contradictory and that, hence, atheism is necessarily true is advocated by Jean-Paul Sartre in his *Being and Nothingness,* translated by Hazel Barnes (New York: Philosophical Library, 1956). Bertrand Russell wavered between calling himself an atheist and an agnostic. Many of his publications may plausibly be regarded as defenses of atheism. In this connection special mention should be made of *The Scientific Outlook* (New York: Norton, 1931), *Religion and Science* (New York: Holt, 1935), and *Why I Am Not a Christian and Other Essays on Related Subjects* (New York: Simon and Schuster, 1957), which includes "What I Believe."

What we have been calling metaphysical theology is defended by H. L. Mansel in *The Limits of Religious Thought* (London: Murray, 1858). Mansel's views were vigorously attacked by John Stuart Mill in his *An Examination of Sir William Hamilton's Philosophy* (4th ed., London, 1872); and Mill in turn was answered by Mansel in *The Philosophy of the Conditioned* (London: Strahan, 1866). The version of metaphysical theology on which we concentrated in the present entry is expounded by Paul Tillich in Vol. I of his three-volume *Systematic Theology* (Chicago: University of Chicago Press, 1951–1963), in his *The Courage to Be* (New Haven, CT: Yale University Press, 1952), and in J. A. T. Robinson, *Honest to God* (London, 1963). This position is

criticized in great detail in Paul Edwards, "Professor Tillich's Confusions," in *Mind* 74 (1965): 192–214, and in Dorothy Emmet, "'The Ground of Being,'" in *Journal of Theological Studies* 15 (1964): 280–292. Various reactions to the views of Robinson are collected in *The Honest to God Debate,* edited by D. L. Edwards (London: SCM Press, 1963). The Thomistic doctrine of "analogical predication," which was not discussed in the present entry, is expounded in the *Summa Theologiae,* I, 13, 5, and in the work by Thomas Cajetan available in *On the Analogy of Names and the Concept of Being,* translated by E. A. Bushinski and H. J. Koren (Pittsburgh, 1953). Contemporary expositions of it may be found in G. H. Joyce, *The Principles of Natural Theology* (London: Longmans Green, 1923), and in E. L. Mascall, *Existence and Analogy* (London, 1949). The theory is criticized in Frederick Ferré, *Language, Logic and God* (New York: Harper, 1961), and in W. T. Blackstone, *The Problem of Religious Knowledge* (Englewood Cliffs, NJ: Prentice-Hall, 1963). There is an interesting attempt to state the doctrine with great precision by using the tools of contemporary logic in I. M. Bochenski, "On Analogy," in *Thomist* 11 (1948): 474–497. Tillich's theory, as well as the Thomistic theory, is criticized in Sidney Hook, *The Quest for Being* (New York: St. Martin's, 1960).

Thomas Aquinas's views on the nature of creation and the possibility of proving that the material universe has not always existed are given in *On the Eternity of the World,* translated by Cyril Vollert (Milwaukee: Marquette Univ. Press, 1964), which also contains relevant extracts from the *Summa Theologiae* and the *Summa contra Gentiles.* The argument for atheism based on the eternity of matter is stated in Ludwig Büchner, *Force and Matter* (4th English ed., London 1884; reprinted New York, 1950). The question of whether contemporary theories in physical cosmology have any bearing on the question of the existence of God is discussed in William Bonnor, *The Mystery of the Expanding Universe* (New York: Macmillan, 1964); M. K. Munitz, *Space, Time and Creation* (Glencoe, IL: Free Press, 1957); E. L. Mascall, *Christian Theology and Natural Science* (London: Longmans Green, 1956); and Antony Flew, "Cosmology and Creation," in *Humanist* 76 (May 1961): 34–35. All the writers just mentioned incline to the view that physical cosmology has no bearing on the question of the existence of God. The opposite position is supported by E. A. Milne in *Modern Cosmology and the Christian Idea of God* (Oxford: Clarendon Press, 1952).

The argument for atheism based on the premise that there is no "cosmic brain" is expounded in Emil Du Bois–Reymond, *Über die Grenzen des Naturerkennens* (Berlin, 1873), and by W. K. Clifford in an essay titled "Body and Mind," which is available in Vol. II of Clifford's two-volume *Lectures and Essays,* edited by F. Pollock (London and New York, 1879). It is criticized in James Martineau, *Modern Materialism and Its Relation to Religion and Theology* (London, 1876; New York, 1877). According to Mauthner, op. cit., Vol. III, pp. 439 ff., the remark that "in scanning the heavens with a telescope he found no God" has been falsely attributed to Laplace and occurs in fact in one of the writings of another distinguished astronomer of the same period, Joseph Jérôme de Lalande. Arguments by Indian philosophers, similar to those of Du Bois–Reymond and Clifford, are found in *Slovavartika,* Sec. I, verses 43–59, reprinted in *A Source Book in Indian Philosophy,* edited by Sarvepalli Radhakrishnan and C. A. Moore (Princeton, NJ: Princeton University Press, 1957).

The essay by Tyndall in which he defends agnosticism in contrast to atheism is contained in his *Fragments of Science* (New York, 1871). A similar argument by Clarence Darrow occurs in his lecture "Why I Am an Agnostic," which was first delivered in 1929 and is now available in *Clarence Darrow—Verdicts out of Court,* edited by A. Weinberg and L. Weinberg (Chicago, 1963). Agnosticism is criticized from an atheistic viewpoint in several of the writings of Friedrich Engels. There is a useful collection of all the main discussions of religion by Marx and Engels in Karl Marx and Friedrich Engels, *On Religion* (Moscow, 1957; New York: Schocken, 1964).

The argument that atheism must be untenable since, if it were true, the atheist himself would have to be omniscient, is advanced in Thomas Chalmers's two-volume *On Natural Theology* (New York, 1836); in J. Foster, *Essays* (London, 1844); and in Robert Flint, *Anti-Theistic Theories* (London, 1878). There is a reply to Chalmers and Foster in G. J. Holyoake, *Trial of Theism* (London, 1858). A somewhat similar argument is contained in Paul Ziff, "About 'God,'" in *Religious Experience and Truth,* edited by Sidney Hook (New York: New York University Press, 1961). There is a reply to this in Paul Edwards, "Some Notes on Anthropomorphic Theology," in *Religious Experience and Truth.*

Pascal's horror of a universe without God is expressed in numerous passages in his *Pensées,* translated by W. E. Trotter, with an introduction by T. S. Eliot (New York: Dover, 2003). William James's claims that unbelief is fatal to "the strenuous mood" is contained in his essay "The Moral Philosopher and the Moral Life," which is reprinted in his *The Will to Believe* (New York: Longmans Green, 1897). The view that atheism makes the universe "loveless" is defended by J. W. Krutch in his *The Modern Temper* (New York: Harcourt, Brace, 1929). Malinowski's remarks about the "tragic" nature of life without God are found in his contribution to the BBC symposium *Science and Religion* (New York, 1931). The very different view that there is something liberating in the rejection of belief in God is advocated in J. M. Guyau, *The Non-Religion of the Future,* with an introduction from N. M. Glatzer (New York: Schocken, 1962); in Friedrich Nietzsche, *Die Fröhliche Wissenschaft,* in Vol. II of his three-volume *Werke,* edited by Karl Schlechta (Munich, 1954–1956); and in Sigmund Freud, *The Future of an Illusion,* translated by W. D. Robson-Scott (New York: Liveright, 1927).

In more recent years there have been numerous books and articles by religious thinkers in which the atheist's position is treated with a certain amount of sympathy. The following writings are especially worth mentioning in this connection: James Collins, *God in Modern Philosophy* (Chicago: Regnery, 1959); Henri de Lubac, *The Drama of Atheist Humanism,* translated by E. M. Riley (New York: Sheed and Ward, 1950); Étienne Borne, *Atheism,* translated by S. J. Tester (New York, 1961); Ignace Lepp, *Atheism in Our Time,* translated by Bernard Murchlord (New York: Macmillan, 1963); W. A. Luijpen, *Phenomenology and Atheism,* translated by W. van de Putte (Pittsburgh: Duquesne University Press, 1964); Jacques Maritain, "The Meaning of

Contemporary Atheism," in *Listener* (March 1950): 427–432; Gabriel Marcel, "Philosophical Atheism," in *International Philosophical Quarterly* 2 (1962): 501–514; and Jean-Marie Le Blond, "The Contemporary Status of Atheism," in *International Philosophical Quarterly* 5 (1965): 37–55.

Baggini, Julian. *Atheism: A Very Short Introduction.* Oxford: Oxford University Press, 2003.

Buckley, Michael. *At the Origins of Modern Atheism.* New Haven: Yale University Press, 1987.

Everitt, Nicholas. *The Non-Existence of God.* London: Routledge, 2004.

Flew, Antony. *The Presumption of Atheism, and other Philosophical Essays on God, Freedom and Immortality.* London: Pemberton, 1976.

Herrick, Jim. *Against the Faith: Essays on Deists, Skeptics, and Atheists.* Amherst, NY: Prometheus Books, 1985.

Hunter, Michael, and David Wootton, eds. *Atheism from the Reformation to the Enlightenment.* Oxford: Oxford University Press, 1992.

Le Poidevin, Robin. *Arguing for Atheism: An Introduction to the Philosophy of Religion.* New York: Routledge, 1996.

MacIntyre, Alasdair, and Paul Ricoeur. *The Religious Significance of Atheism.* New York: Columbia University Press, 1969.

Mackie, J. L. *The Miracle of Theism.* Oxford: Oxford University Press, 1982.

Martin, Michael. *Atheism: A Philosophical Analysis.* Philadelphia: Temple University Press, 1990.

McGrath, Alister E. *The Twilight of Atheism: The Rise and Fall of Disbelief in the Modern World.* New York: Doubleday, 2004.

Nielsen, Kai. *Philosophy and Atheism: In Defense of Atheism.* Amherst, NY: Prometheus Books, 1985.

Smart, J. J. C., and John Haldane. *Atheism and Theism,* 2nd ed. Oxford: Blackwell, 2003.

Smith, George H. *Why Atheism?* Amherst, NY: Prometheus Books, 2000.

Wallace, Stan, ed. *Does God Exist?: The Craig-Flew Debate.* Burlington, VT: Ashgate, 2003.

Paul Edwards (1967)
Bibliography updated by Christian B. Miller (2005)

ATHEISMUSSTREIT

Atheismusstreit, a famous controversy in Germany during the closing years of the eighteenth century, concerned the allegedly subversive philosophical views of Johann Gottlieb Fichte (1762–1814) and of the much less well-known Friedrich C. Forberg (1770–1848).

Fichte, who died as a pillar of respectability, had advanced various radical views in his earlier years, and on the nature and reality of God he never became fully orthodox. In 1793, while living as a private tutor in Zürich, Fichte published two political pamphlets titled "Reclamation of the Freedom of Thought from the Princes of Europe" and "Contributions Designed to Correct the Judgment of the Public on the French Revolution" in which he enthusiastically supported the basic principles of the French Revolution, arguing for free expression of opinion as an inalienable human right and subjecting the privileges of the nobility and the church to trenchant criticism. Fichte was at that time already famous, largely as a result of his Kantian work, *Versuch einer Kritik aller Offenbarung* (Essay toward a Critique of All Revelation), which had been published anonymously in Königsberg in 1792. Some reviewers attributed the essay to Immanuel Kant, who thereupon revealed Fichte as the true author, at the same time bestowing high praise on his gifts. In spite of Fichte's reputation as a political radical, he was appointed professor of philosophy at Jena in 1794.

For some time things went fairly smoothly at Jena. Fichte, who was a dynamic lecturer, made numerous converts among both his colleagues and the students, although there were some acrimonious exchanges with the psychologist C. C. E. Schmid and others distrustful of Fichte's speculative bent. There were two violent controversies before the *Atheismusstreit* broke out. One of these concerned a series of public lectures that Fichte had scheduled on Sundays from ten to eleven in the morning. Local clergymen were outraged, and the Over-Consistory (of which no less a man than Johann Gottfried Herder was a member) appealed to the government at Weimar to intervene. One local journal called attention to Fichte's revolutionary politics and asserted that he and his democratic followers were engaging in a deliberate attempt to substitute the worship of reason for the worship of God. The senate of the university and the government of Weimar decided in Fichte's favor, but it was agreed to give the lectures at three in the afternoon. The other controversy involved the university fraternities, which Fichte regarded as unethical and corrupt and whose abolition he publicly recommended. On New Year's Eve of 1795 students belonging to the fraternities attacked Fichte's house, breaking windows and heaping insults upon him and his wife. In the early months of 1795 Fichte felt his life to be in danger and found it necessary to reside outside of Jena until the tempers of the fraternity members had calmed down.

THE OFFENDING ARTICLES

The *Atheismusstreit* itself began in 1798 with the publication in the *Philosophisches Journal,* a periodical of which Fichte was coeditor, of an essay by Forberg titled "The Evolution of the Nature of Religion." Fichte's conservative

English biographer, Robert Adamson, dismisses Forberg's position as an "exaggeration of the dismal rationalism into which the weaker Kantians had drifted." In fact, however, Forberg's paper shows a powerful and independent thinker at work and does not seem dated even now. (Interestingly enough, Hans Vaihinger called attention to the philosophical merits of Forberg's work after almost total neglect for a century, citing him as an early positivistic fictionalist and praising his unusually fine appreciation of the more radical aspects of Kant's philosophy of religion.)

What, Forberg asks, is the foundation of the belief in a moral world order? There are three possible sources—experience, speculation, and conscience. Experience certainly lends no support to such a belief; if anything, it shows an evil deity in conflict with, and more often than not triumphing over, a good one. As for speculation, Forberg briefly and very clearly repeats Kant's objections to the ontological, cosmological, and teleological arguments, adding some critical observations of his own. Accordingly, the foundation of religion must be sought in our conscience. Religion is "purely and solely the fruit of a morally good heart …; it originates entirely from the wish of the good heart that the good in the world should triumph over the evil." To have "genuine religion" is not to have a belief in God; it is to be a partisan of the good, to act as if the kingdom of God, which for Forberg simply means a just and moral world, were attainable. Forberg himself evidently did not believe that such a world was attainable. This belief, however, is no more essential to true religion than is the belief in God. What is essential is the striving in the direction of a moral world whether or not one believes in its attainability. Forberg most emphatically insists that an atheist can be a religious person in his sense of religion. "Practical belief and theoretical unbelief on the one hand and theoretical belief and practical unbelief on the other may very well coexist."

At first sight this position may appear to be a kind of voluntaristic defense of traditional religion and an endorsement of Kant's moral argument, as this has frequently been interpreted. In fact, Forberg is very far removed from any such point of view. He is not saying that since there is no evidence either way, it is as well to believe in a just God or the attainability of a moral world. We are not, according to him, required to believe any such thing, and it does not really matter whether we do. We *are* required to *act as if we believed this.* Forberg was highly critical of the common interpretation of Kant's moral argument as providing cognitive support for belief in God. In his later defense of himself, *Friedrich Carl Forbergs Apologie seines angeblichen Atheismus* (Gotha, 1799), he castigates the "usual, far too theoretical presentation of the notion of a practical belief," adding that it is "an unphilosophical conception which allows people to reintroduce through a back door every kind of nonsense of which theoretical philosophy has rid us with much effort."

In the same issue of the *Philosophisches Journal,* Fichte published an essay, "Concerning the Foundation of Our Belief in Divine Government of the World," which was intended to complement Forberg's paper. In a somewhat patronizing opening Fichte informs the reader that although he agrees with much in Forberg's piece, there are some important questions on which Forberg has not "quite reached" his, Fichte's, position and that since he had not previously had an opportunity to explain himself on these issues, he would do so now. Attempts to infer the existence of God from the world of sense objects, he proceeds, must inevitably fail. From the point of view of common sense and science, the world of sense objects is "absolute" and self-existing, and any attempt to go beyond it is "total nonsense." The assumption of a cosmic intelligence, moreover, would not explain anything, since it is quite unintelligible to talk about the creation of material things out of ideas. Considered from the transcendental viewpoint, the world of the senses is a "mere reflection of our own activity," and as a "nothing" it can hardly require an explanation outside itself.

Our belief in God can be grounded only in the supersensible world, which for Fichte is the only ultimately real world. This is the world of free moral agents, and unlike Forberg, Fichte teaches that the universe is, in fact, moral and just, that "every truly good act must succeed, that every evil one must surely fail, that for those who really love the good all things must turn out for the best." This does not mean that the good necessarily receive rewards in terms of pleasure but the world in which we experience pleasure is not the real world. The world of sense objects exists only as a "stage" on which free agents perform or fail to perform their duty. It has not "the slightest influence on morality or immorality, not the slightest power over our free nature." It is, in fact, nothing more than the "material objectification of our duty; our duty is what is ultimately real, what is the fundamental stuff of all phenomena."

God is identical with the moral world order. A person believes in God insofar as he does his duty "gaily and without concern," without doubts or fears about consequences. The "true atheist" is he who calculates the consequences instead of following the voice of his

conscience; he "raises his own counsel above the counsel of God and thus raises himself to God's position." He who does evil in order to produce good is godless. "You must not lie," Fichte adds by way of illustration, "even if the world were to go to pieces as a consequence"; a moral agent knows, however, that the world could not go to pieces, since "the plan of its preservation could not possibly be based on a lie." Both here and elsewhere Fichte argued that all cognition is based on the existence of the moral world order. The existence of God, which here, of course, simply means the moral world order, is therefore more certain than anything else. It is presupposed in any piece of valid reasoning, and hence it cannot be, nor does it need to be, proved. "It is the ground of all other certainty and the only absolutely valid objective reality."

THE ANONYMOUS PAMPHLETS

Attention was drawn to these essays and their alleged subversion in a pamphlet published late in 1798 under the title "Letters from a Father to His Student-Son concerning the Atheism of Fichte and Forberg." The pamphlet was signed *G* and was at first attributed to D. Gabler, a respectable theologian teaching at Altdorf. Gabler vehemently denied any connection with the pamphlet, however, and publicly expressed his high regard for Fichte. Fichte himself attributed it to one of his enemies at Jena, Gruner, but the authorship remains uncertain. The main argument of the pamphlet followed a simple, popular line: Belief in an ever present "witness and judge" is essential to the moral behavior of human beings; if people were not afraid of punishment in the next world, they would be certain to do evil whenever they expected to escape the secular penalties. As a high school teacher, Forberg in particular is regarded as a most dangerous man. How could such a rector give a "thorough religious education" to the students under his charge? "To sow the seeds of immorality among young people and make belief in God suspect is not a permissible game." When compared to the protector of morality who hunted Bertrand Russell in New York City 150 years later, the attack was conducted with decorum and refinement; however, several later anonymous pamphlets were somewhat less refined. As usually happens in such cases, they contained slanderous comments about Fichte's private life and "sexual philosophy."

FICHTE'S DISMISSAL

The rest of the story does little credit to any of the parties except Fichte and Forberg. Moved by the "Father's Letter," the Saxon government, on November 19, 1798, published a Rescript ordering the universities of Leipzig and Wittenberg to confiscate all copies of the *Philosophisches Journal* because of the atheistic articles contained in it. This was followed by a request to the neighboring German governments to take similar steps. The dukes of Saxe-Weimar were informed that Saxon students would not be allowed to enroll in Jena unless there was an immediate investigation into the conduct of the two offenders. The grand duke of Weimar, a ruler with a genuine respect for scholarship, was free from any trace of religious fanaticism; however, any attempt he might have made to hush up the case was prevented by Fichte's public defenses of himself. In January 1799, Fichte wrote his "Appeal to the Public concerning the Accusation of the Expression of Atheistic Opinions," a copy of which was promptly sent to the grand duke. In March 1799 he wrote the "Juridical Defense against the Accusation of Atheism," which was primarily addressed to the university authorities but a copy of which was also forwarded to the grand duke. In these "defenses" Fichte contended, first, that his philosophical position, although far removed from the anthropomorphic popular religion, could not fairly be regarded as a form of atheism and was, in fact, "true Christianity" and, second, that any punishment inflicted on Forberg or himself would be a gross violation of academic freedom. The case, Fichte insisted, was one of great importance; since the accusation had been public, the verdict should also be public. Fichte's friends regarded this as a most imprudent demand, and rumors were soon current that the Weimar government was about to impose a public censure on Fichte. In the hope of preventing this, Fichte wrote a letter to Privy Councilor Voigt in which he declared that he would under no circumstances submit to censure. In such an event, he said, he would instantly resign. He added that several distinguished members of the Jena faculty shared his opinion that censure would constitute infringement of their academic rights and that they would resign with him. Voigt was told that he was free to show the letter to others, including, presumably, the Weimar authorities, who were about to reach their verdict.

This letter turned out to be Fichte's undoing at Jena. The Weimar government quite improperly treated it as a formal document. It avoided any censure of Fichte (or of his coeditor Niethammer) on the charge of atheism. Instead, both were rebuked in the mildest possible language for their "indiscretion" and advised to exercise greater caution in their selection of articles for the *Philosophisches Journal.* The journal itself was not proscribed, nor was there any mention of what teachers should or should not say in their classrooms. In a post-

script, however, reference was made to Fichte's letter to Voigt, and his threatened resignation in case of censure was noted and accepted. In effect, this amounted to Fichte's dismissal, and two petitions on his behalf by the Jena student body to the duke were of no avail. Johann Wolfgang von Goethe, who a few years earlier had been largely instrumental in securing the Jena chair for Fichte, was one of those in the Weimar council who demanded Fichte's ouster. Fichte's support of the French Revolution was apparently a minor thing, but the language used in the letter to Voigt was unforgivable. "For my own part," Goethe wrote in a letter a few months later, "I declare that I would have voted against my own son if he had permitted himself such language against a government." Forberg was mildly censured by his superiors and did not return to any writings on religion until shortly before his death, when he published his autobiography, in which there is a very full account of the entire episode and a reaffirmation of all his earlier convictions.

THE CHARGE OF ATHEISM

In his "Appeal to the Public," Fichte had vehemently denied the charge of atheism. Using language which is very similar to that employed in the twentieth century by Paul Tillich and Bishop J. A. T. Robinson, he inveighed against the popular "idol-worship" of God as a "substance," as another entity *in* the world, and against the vulgar "eudaemonistic" morality that makes God a giver of "sensuous" rewards for good deeds and "sensuous" punishments for evil deeds. Such a conception—or, indeed, any attribution of personal characteristics to God—constitutes a lowering and limiting of the deity and has to be opposed in the interests of true religion. There is no need to question Fichte's sincerity, and in more senses than one it may be granted that he was a religious man.

At the same time the charge of atheism does not appear to have been totally unjustified. People do not usually mean by God simply the moral world order, and the denial of God as an entity over and above the more familiar objects of experience (including moral human agents) is precisely what is ordinarily meant by atheism. On all these points Fichte had been very explicit in the original essay. "There can be no doubt," he had written, "that the notion of God as a separate substance is impossible and contradictory, and it is permitted to say this plainly." Again, "We need no other god [than the moral world order], and we cannot comprehend another one. There is no rational justification for going beyond the moral world order to a separate entity as its cause."

Granting that there was some basis for the charge of atheism against Fichte, this in no way excuses the behavior of the Weimar authorities or of Fichte's and Forberg's other detractors. Not one distinguished voice was raised anywhere in Germany in defense of the accused men. Kant himself, who was still alive, was moved to a statement in the *Allgemeine Literaturzeitung* (1799, No. 109) in which he emphatically dissociated his philosophy from Fichte's system. "When I compare the state of the German republic of letters of this period with the Enlightenment literature of France a generation earlier, I am overcome with the deepest shame," was the apt comment of the historian Fritz Mauthner.

See also Fichte, Johann Gottlieb.

Bibliography

The main documents of the *Atheismusstreit* have been collected in *Fichte und Forberg. Die philosophischen Schriften zum Atheismusstreit,* edited by F. Medicus (Leipzig, 1910), and *Die Schriften zu J. G. Fichte's Atheismusstreit,* edited by H. Lindau (Munich, 1912).

Critical appraisals of the actions and writings of the participants are found in Vol. IV of Fritz Mauthner's *Der Atheismus und seine Geschichte im Abendlande* (Stuttgart: Deutsche Verlagsanstalt, 1923); F. Paulsen's "G. J. Fichte im Kampf um die Freiheit des philosophischen Denkens," in *Deutsche Rundschau* 99 (1899): 66–76; and H. Rickert's "Fichte's Atheismusstreit und die kantische Philosophie," in *Kant Studien* 4 (1899): 137–166.

There is a discussion of the originality of Forberg as well as of the evidence favoring a "left-wing" interpretation of Kant's moral argument for the existence of God in Appendix A and Appendix B of Hans Vaihinger, *Die Philosophie des Als Ob* (Berlin: Reuther and Reichard, 1911), translated by C. K. Ogden as *The Philosophy of "As If"* (New York: Harcourt Brace, 1924).

OTHER RECOMMENDED TITLES

Di Giovanni, George. "From Jacobi's Philosophical Novel to Fichte's Idealism: Some Comments on the 1798–99 Atheism Dispute." *Journal of the History of Philosophy* 27 (1989): 75–100.

Fabro, Cornelio. "Eine Unbekannte Schrift Zum Atheismusstreit." *Kant Studien* 58 (1967): 5–21.

Paul Edwards (1967)
Bibliography updated by Tamra Frei (2005)

ATOMIC THEORY IN INDIAN PHILOSOPHY

In classical Indian philosophy two Sanskrit words are used for the atom, the smallest impartite physical entity: "aṇu" and "paramāṇu." On the existence of such atoms,

the classical Indian philosophers were divided. Among the orthodox Brahmanic schools, the Nyāya-Vaiśeṣika philosophers were the preeminent defenders of atomism, with the Mīmāṃsā philosophers as allies. On the opposite side, the Vedāntins denied atomism. Among the non-Brahmanic schools, the Jainas were clearly atomists, as were the Hinayana Buddhists. Yogācāra Buddhism, however, was strongly critical of atomism, and so too was Madhyamaka Buddhism.

The division of opinion on the issue thus cuts across the division between the Brahmanic and non-Brahmanic schools. Instead, the range of views about atomism more closely reflects the different schools' commitment to realism. After all, atomism is usually associated with a realist view of the world, in which atoms are taken to be objective, mind-independent entities. Predictably enough, then, we find espousing atomism such staunch philosophical realists as the Naiyāyikas and Mīmāṃsakas, as well as such heterodox realists as the Ābhidharmikas and the Jainas. In contrast, opposition to atomism is led by such antirealists as the Advaitins, the Mādhyamikas, and the Yogācārins.

ATOMISTS

The earliest Indian defenders of atomism may well be the Jainas, with texts defending atomism that date at least as far back as the third century CE. According to Jainism, everything in the world, save for souls and space, is produced from matter, and all matter consists of indivisible atoms (*paramāṇu*), each occupying a single point of space. Matter has two forms: a simple or atomic form and a compound (*skandha*) form. Perceivable material objects are compounds, composed of homogeneous atoms (there are no distinct kinds of atoms corresponding to the four kinds of elements). Impartite atoms are eternal, though this is obviously not true of the partite compounds. Indeed, atoms are supposed to be eternal precisely because they lack parts and are thus incapable of disintegration. But there is nonetheless a sense in which atoms, like compounds, are subject to qualitative change because, though all atoms are indistinguishable in substance, qualities present in an atom can be increased or decreased by many degrees.

To explain how atoms join as they do, the Jainas posit that some atoms are viscid and some dry, which permits aggregation of the two different kinds of atoms (much as particles of barley meal combine to form lumps when drops of water fall upon them). Moreover, they are viscid and dry in various degrees, with no aggregates combining atoms with the lowest degrees of the two properties or equal degrees of the same property.

These Jaina speculations help to highlight three central questions for which the Indian philosophers expected atomic theories to provide answers: What evidence do we have for the existence of atoms? How is it possible for one atom to join with another? Why do atoms come together as they do?

With regard to the first question, the two main Indian arguments for the existence of atoms are both inferential. The first argument rests on the claim that there has to be a lower limit to the scale of diminishing minuteness. Gross objects clearly exist and are divisible. Yet the process of physical division must have a terminal point, and this terminal point to division must, by definition, be indivisible. The second argument attempts a reductio ad absurdum of the denial of such a terminal point: Unless the process of division comes to an end, everything must be equally composed of an infinite number of parts, and hence all comparative ascriptions of unequal magnitude to gross objects are undermined. The mountain and the mustard seed would have to be of equal size!

Of course, even if we are persuaded by these arguments that atoms do exist, any atomic theory still needs to address the second question and offer some explanation of how atoms combine to form partite entities. After all, atoms are supposedly impartite, and yet our only direct experience of conjunction involves partite things. But if we give up the thesis that atoms are truly impartite, we also have to give up one of the main arguments for the existence of atoms.

In reply, the Naiyāyikas utilize their distinctive mereological theory (theory of partition), according to which composite wholes are never reducible to their parts, though wholes inhere in parts. Hence a composite whole is a distinct entity, and not a mere collection of its conjoined parts. Moreover, since the whole is thus distinct from the sum of its parts, it can, unsurprisingly, have properties not possessed by any of its parts. This particular mereological theory, however, is unacceptable to both Buddhists and Advaitins, who object that the idea that wholes inhere in their parts would require a further relation to relate inherence to its relata, and so on ad infinitum. The Buddhists maintain instead that wholes are unreal, being mere conceptual constructions, and only parts are real. Thus for them, all conventional objects are mere aggregates of atoms. The Jaina response is different again: The composite whole is just the parts in a changed state.

Finally, even if we have reason to believe both that there are atoms and that they can combine, a viable atomic theory still needs to offer some sort of explanation of how atoms are brought together. The Jaina explanation in terms of a theory of varying degrees of viscidity and dryness builds on their view that all atoms are homogeneous, with the result that the division into the four elements is derived and secondary. The Nyāya-Vaiśeṣika school denies this, claiming instead that the four elements of earth, water, air, and fire involve four kinds of atoms sufficiently qualitatively different from each other so that the atoms of one element can give rise only to products of that element.

The elaborate Nyāya-Vaiśeṣika theory of how atoms combine to form compound entities seeks to address the issue of how atoms of infinitesimal magnitude can add together to produce a macroscopic object. Their explanation is that when two infinitesimal atoms combine into a dyad, there is a sort of quantum leap, and the new submolecule thus formed has a minute (*hrasva*) magnitude. Dyads then combine into perceptible molecules or triads (composed of three dyads), and there is another quantum leap in magnitude to a gross (*mahat*) quantum. The addition of gross quanta then straightforwardly accounts for the magnitude of macroscopic objects.

The point of this postulated double quantum jump from single atoms to dyads and then from dyads to triads is to insist that the finite magnitude of the triad arises from the infinitesimal atoms as a result of the *number* of the constituent atoms and not as a result of their magnitude, as in gross objects. Unsurprisingly, many Indian philosophers (both atomist and antiatomist) found this part of the Nyāya-Vaiśeṣika atomic theory unconvincing.

Moreover, all of this still leaves unexplained the initial conjunction of two atoms to produce a dyad. Later the Nyāya-Vaiśeṣika school invoked God's agency to help out here: Since all atoms are insentient, the process of combination must be guided by an intelligent divine agent. Other Indian philosophers disagreed, however, as to whether this amounts to a persuasive argument for the existence of God or to just an ad hoc addition to an already unsatisfactory atomic theory.

The Nyāya-Vaiśeṣika school took one advantage of its atomic theory to be that it can avoid the Buddhist theory of universal flux and can explain the identity of a substance through change in terms of the identity of unchanging, eternal atoms. A substance can undergo change without the constituent atoms changing because the qualities of a substance can change while the substance persists. However, consider what happens when we fire a clay pot so that it changes color. The Vaiśeṣikas claimed both that the unfired pot as a whole is replaced by a new pot as a whole, and that the application of heat causes a change of qualities to occur at the level of the individual atoms. But in admitting that change at the level of gross objects involves change at the atomic level, the Vaiśeṣika theory risks collapsing into the Buddhist theory of universal flux. Hence the Nyāya atomic theorists denied that change occurs at the level of the individual atoms, claiming instead that the whole remains intact while the change occurs.

Common to the different atomic theories of both Jainism and Nyāya-Vaiśeṣika are the claims that the atoms are genuinely indivisible, infinitesimal, and eternal. Other Indian atomists deny some of these claims. The Mīmāṃsā school, for instance, is willing to admit that whether entities are gross or minute is only relative. They thus accept as atoms the dust motes visible in a sunbeam (these are triads in the Nyāya-Vaiśeṣika system, the smallest perceivable particles). Although the Mīmāṃsakas do not entirely rule out the Nyāya-Vaiśeṣika conception of an atom as impossible, they criticize it as an overly speculative thesis. Even if the dust mote is theoretically divisible and hence apparently nonatomic, Mīmāṃsakas are only willing to accept such atoms as are established by common experience. There is no purpose served by assuming any atoms beyond these.

In contrast, the Ābhidharmika atomists affirm the existence of atoms smaller than dust motes but deny that they are eternal, since in Buddhism everything is taken to be impermanent. According to Buddhist atomic theory, although atoms are the smallest unit of matter, they never occur alone, but rather occur only as members of an aggregate of at least seven or eight atoms. Hence it is unsurprising that we do not experience individual atoms as separately perceptible. But we do nevertheless perceive the aggregates and, contrary to Nyāya-Vaiśeṣika claims, there are no aggregates distinct from the atoms themselves. Thus our perception that the atoms constituting an aggregate are gross is really an illusion due to the close and collective presence of a multitude of minute atoms.

ANTIATOMISTS

The Vedāntins and the Mahayana Buddhists were the chief representatives of Indian antiatomism, though their objections to atomism are frequently different and their own rival ontologies are significantly distinct. One specifically Vedāntin argument against atomism is that the Hindu scriptures nowhere affirm it. Clearly, this argument is not intended to persuade non-Brahmanic atom-

ists, but it is interesting to note that most Brahmanic atomists too do not feel obliged to respond to it. The mere absence of a Vedic sanction is apparently thought to be obviously insufficient grounds for rejecting a philosophical theory. (A notable exception to this general trend of indifference is the Naiyāyika philosopher Udayana [eleventh century, CE], who goes out of his way to argue that there is indeed a scriptural warrant for atomism.)

The Advaita Vedāntins offered a more straightforwardly philosophical objection to the Nyāya-Vaiśeṣika theory of atomic composition. They argued that ontological parsimony ought to make us reject the Naiyāyikas' posit of dyads as unnecessary, for why can we not just say instead that three atoms directly combine to form a triad, the smallest visible substance. The gross magnitude of the triad will then be explicable not in terms of the magnitude or aggregation of atoms, but in terms of the *number* of atoms.

The main Indian argument that some form of atomism is rationally necessary is, of course, that it is required to explain the existence of gross material objects, which are indisputably partite. Again and again the atomists defended the controversial details of their theories with an argument to the best explanation: that since all agree that there are composite physical objects, one needs to posit atoms to best explain their existence and nature. But this strategy presupposes a common commitment to realism about the external world. The Indian antiatomists did not share this general commitment.

This is particularly obvious when we attend to the antiatomist arguments of the Yogācārin philosopher Vasubandhu (fourth or fifth century CE). Vasubandhu began by explicitly affirming the idealist thesis that everything is mind only. But realism, of course, denies this thesis. Vasubandhu responded by arguing that realism is false because realism implies atomism and atomism is incoherent.

Like the Ābhidharmikas, Vasubandhu rejected the Nyāya-Vaiśeṣika theory of organic wholes as unsupported by experience. But he also rejected the Abhidharma view that material wholes are mere aggregates of atoms, on the ground that for this to be so, the atoms would have to be joined. Such conjunction is either partial or total. If it is partial, the atoms must have parts in contact with one another; if it is total, all the atoms must collapse into the same atom-sized space. Either way, there cannot be a plurality of impartite atoms. Furthermore, an atom cannot be thought of as spatially extended without allowing that it has a front part different from its back part. But if atoms are unextended, then aggregates of them cannot constitute extended gross objects. Thus atomism (and hence realism) is incoherent, and idealism is vindicated.

Yogācāra Buddhism is admittedly a rather peculiar kind of idealism, since it denies the existence of both the objects of consciousness and the subject of consciousness. Ultimately, all that exists is pure consciousness devoid of all subject/object duality. But whether or not Yogācāra thought is best classified as a variety of idealism, it is indubitably a variety of antirealism. Moreover, while other Indian antiatomists, such as the Mādhyamikas and the Advaitins, were certainly not idealists, they also in their various ways shared the Yogācāra thinkers' antirealist doubt of the commonsense assumption of an objective reality populated by ontologically independent entities. These Indian antiatomists are thus all equally unforgiving of the atomists' general strategy of attempting to excuse the anomalies in their various atomic theories by an appeal to atomism as the best explanation of gross external objects. In classical Indian philosophy, the avowed aim of philosophy is liberation (*mokṣa*). For the Indian antirealists, this goal is to be attained not by theorizing about the nature of a supposedly objective external world, but by transcending all such conceptions, including atomism and its presuppositions. In this sense, there is arguably a common antirealist motivation for Indian antiatomism, notwithstanding the very significant philosophical differences among the different antiatomist schools.

See also Causation in Indian Philosophy.

Bibliography

Gangopadhyaya, Mrinalkanti. *Indian Atomism: History and Sources*. Calcutta, India: K. P. Bagchi, 1980. A splendid anthology of Sanskrit primary sources together with annotated English translations. Also contains a very useful historical introduction.

Kapstein, Matthew T. *Reason's Traces*. Boston: Wisdom, 2001. Chapter 7 is an excellent philosophical study of Vasubandhu's arguments against atomism.

Potter, Karl H., ed. *Indian Metaphysics and Epistemology: The Tradition of Nyāya-Vaiśeṣika up to Gaṅgeśa*. Princeton, NJ: Princeton University Press, 1977. An extremely valuable survey of early Nyāya-Vaiśeṣika metaphysics.

Roy W. Perrett (2005)

ATOMISM

Atomism is a doctrine that has a long history in both philosophy and science. For this reason it is not easy to define

its content in such a way as to comprehend all the historical variations and especially the historical development of the doctrine. In a very general sense, however, atomism may be defined as the doctrine that material reality is composed of simple and unchangeable minute particles, called atoms. It holds that all observable changes must be reduced to changes in the configuration of these particles. The multiplicity of visible forms in nature must likewise be based upon differences of configuration. The best way to discuss the variations of this general idea of atomism is to follow the historical development, which shows a gradual shift of emphasis from philosophical to scientific considerations. Consequently, the first part of this article, covering the period from the sixth century BCE to the seventeenth century, will be of a philosophical nature because in this period atomism was considered preponderantly from a philosophic point of view. The second part is concerned primarily with science, for it was in the period after the seventeenth century that atomism evolved in a scientific theory.

THE PHILOSOPHICAL PERIOD

In Greek philosophy we are already confronted with several types of atomism. Atomism in the strict sense, propounded by Leucippus and Democritus (fifth century BCE), should be looked upon as an attempt to reconcile the data of sense experience with Parmenides' thesis that matter is unchangeable. Parmenides rejected the possibility of change on rational grounds; change seemed to be unintelligible. He was convinced that reality must be one, that it must possess unity, and that, being *one* reality, it could not change. It may be remarked that this thesis of Parmenides is a presupposition for all rational science. Without fundamental unity, no universal laws are possible; without fundamental immutability, no laws covering past, present, and future can be valid. Yet, it is clear that Parmenides' approach is one-sided. Science may presuppose unity and immutability, but it also presupposes change. Only by studying changes is science able to discover the immutable laws of nature.

Democritus agreed with Parmenides on the unintelligibility and impossibility of qualitative change. He did not agree on the unintelligibility and impossibility of quantitative change. This type of change is subject to mathematical reasoning and therefore is possible. By the same token, Democritus denied qualitative multiplicity, but accepted multiplicity based on purely quantitative differences. Consequently, he accepted a numeric multitude of original beings, the atoms. These atoms did not differ qualitatively; only their sizes and figures differed. The infinite variety of observable things could be explained by the different shapes and sizes of the atoms that constituted them and by the different ways in which the atoms were combined. Observable changes were based upon a change in combinations of the atoms. During such changes, however, the atoms themselves remained intrinsically unchanged. They did not change their nature, or even their size or figure; they were indivisible (hence their name *'ἄτομος* or indivisible).

Other forms of Greek atomism differed from that conceived by Democritus mainly in two points. First, they did not restrict the differences between the atoms to purely quantitative ones, but also accepted differences in quality. There was even a system that assumed as many qualitatively different atoms as there are different observable substances (Anaxagoras, fifth century BCE). Usually, however, only a few kinds of atoms were assumed, based upon the famous doctrine of the four elements: earth, water, air, and fire (Empedocles, fifth century BCE).

The second point of difference concerned the indivisibility of atoms. It is evident that a system that does not accept the indivisibility of atoms cannot properly be called atomism, but since such systems have played an important role in the history of atomism, we must mention them. For Democritus, the indivisibility of atoms was an absolute indivisibility, being the consequence of an absolute immutability. There were systems, however, that considered the indivisibility and immutability as only relative. The "atoms" could be divided, but they then became "atoms" of another substance; they changed their nature. (Here again an exception must be made for atoms as conceived of by Anaxagoras. These could be divided, but remained of the same kind. Hence they received the name of *homoiomerics,* possessing similar parts.) From the historical viewpoint, the most important system with qualitatively different atoms is that developed by the commentators on Aristotle—Alexander of Aphrodisias (second century CE), Themistius (fourth century) and John Philoponus (sixth century). In their system the atoms are called *elachista* (very small or smallest), the Greek equivalent of the Latin *minima,* which in medieval Latin writings indicates the smallest particles.

That these commentators on Aristotle combined the existence of "atoms" with the possibility of their changing their nature is not surprising. Aristotle was not satisfied by Democritus' atomism and was of the opinion that Democritus went only halfway. Atomism certainly opened up the possibility of explaining some changes that occur in nature, but not all. Nor did it account for all variety. Thus, the first task imposed upon Aristotle was a

careful and critical reexamination of Parmenides' thesis. The result was his matter-form doctrine, stating that every material being is composed of *primary matter* and *form of being.* This composition, however, is not chemical or physical; it goes deeper. The possibility of change presupposes a certain fundamental nonsimplicity, for otherwise it is not possible to account for both aspects that are present in change: the aspect of a certain permanence (matter) and the aspect of something that is really new (form). Matter in the Aristotelian sense is not a substance, but the capacity to receive "forms."

To a certain extent, Democritus followed the same line of thought. Democritus, however, "substantialized" the permanent aspect (the atoms), thus narrowing the possibility of change. For Aristotle the "atoms" too should be subject to change and therefore "composed." Aristotle, however, did not propound a corpuscular theory of his own. Only a few remarks that could have been the starting point are found in a passus (*Physics* I 4, 187B18–34) in which he criticizes Anaxagoras' theory about the infinite divisibility of material things. Somewhere there must be a limit to divisibility. This limit depends on the specific nature of a thing. It was left to Aristotle's Hellenistic, Arabian, and medieval commentators to develop the casual remarks of their master into the *minima naturalia* theory, stating that each kind of substance has its specific *minima naturalia.*

In Greek philosophy there were also transitional theories between qualitative and quantitative forms of atomism. Plato (427–347 BCE), for example, adhered to the doctrine of the four elements; but the differences between the atoms of the respective elements were quantitative. An atom of fire had the form of a tetrahedron; that of air, an octahedron; that of water, an icosahedron; and that of earth, a cube.

When evaluating the importance of Greek atomism in the light of modern atomic theories, it should be borne in mind that in Greek thought philosophy and science still formed a unity. Greek atomism, therefore, was as much inspired by the desire to find a solution to the problem of mutability and plurality in general as by the desire to provide scientific explanations for specific phenomena. Although we meet with some ideas that can rightly be considered as precursors of classical physics and chemistry, the main importance of the old atomistic doctrines to modern science does not lie in these rather primitive scientific anticipations. The greatest achievement of Greek atomism was its general view of nature. The multitude of phenomena must be based upon some unity, and the ever-changing aspects of the phenomena are nevertheless aspects of a fundamentally unchanging world. To this view both the quantitative and the qualitative atomism have contributed—the latter by drawing attention to empirical aspects; the former, to the mathematical.

The history of the two forms of philosophical atomism until the birth of a scientific atomic theory has been rather different. This can easily be explained. Owing to the influence of Plato and Aristotle, Democritus' atomism did not gain preeminence in Greek, Arabian, and medieval thought. Yet that is not the only reason. Much more important is the fact that Democritus' atomism was more or less complete; and his followers, such as Epicurus (341–270 BCE) and the Latin poet Lucretius Carus (96–55 BCE), could confine themselves simply to taking over Democritus' doctrine.

The Aristotelian minima theory, however, existed only in an embryonic state. To Aristotle and his Hellenistic commentators the *minima naturalia* did not mean much more than a theoretical limit of divisibility; they were potentialities rather than actualities. With Averroes, however, we find an important development. According to him, the minima play an important role during chemical reactions. The Latin Averroists followed up this line of thought. Whereas most of the Latin commentators on Aristotle restricted themselves to a more or less systematic treatment of the minima as theoretical limits of divisibility, such Averroists as Agostino Nifo (1473–1538) attributed to the minima a kind of independent actual existence. The minima were considered as actual building stones of reality. The increase or decrease of a quantity of a substance amounts to the addition or subtraction of a certain number of minima. A chemical reaction takes place among the minima.

The fundamental importance of this view to science will be clear. Because the minima had acquired more physical reality, it became necessary to examine how their properties could be reconciled with the specific sensible properties of different substances. A first attempt to do so is found in Julius Caesar Scaliger (1484–1558). Some properties of matter, such as fineness and coarseness, depend on the minima themselves, while others depend on the manner in which the *minima* configurated. Rain, snow, and hail are composed of the same minima; but their densities are different because the minima of these three substances are at smaller or greater distances from one another. As to the chemical reaction, Scaliger remarked: "Chemical composition is the motion of the minima towards mutual contact so that union is effected" (*Exercitationes,* p. 345). Like Aristotle, he was convinced

that Democritus was wrong. In a chemical compound the particles are not just lying close together; they form a real unity. Scaliger, however, was also convinced that the minima play a role in effecting the composition; and for that reason he was not satisfied with the Aristotelian definition of chemical composition as "the union of the reagents," in which the minima are not mentioned.

To sum up our survey of the development of the minima doctrine, and to prove that the opinions of Nifo and Scaliger were no exceptions, we may quote Francis Toletus (1532–1596), one of the best-known sixteenth-century commentators on Aristotle: "Concerning the manner of chemical composition, the opinions of authors vary, but they all agree in this: the reagent substances are divided into minima. In this division the separated minima of one substance come alongside the minima of the other and act upon each other till a third substance, having the substantial form of the compound is generated" (*De Generatione et corruptione* I, 10, 19).

THE SCIENTIFIC PERIOD

The seventeenth century is an important period in the history of atomism. Not only did atomism come to occupy a central position in philosophical discussion, but it also became an inspiring idea for the spiritual fathers of modern science. The philosophic differences between the atomic systems were soon pushed into the background, while the more scientific aspects that were held in common came to the foreground. Daniel Sennert (1572–1657) offers a clear example of this tendency. Basically, his corpuscular theories were derived from the doctrine of *minima naturalia,* but they also contain typically Democritean ideas. In a sense the same could be said of Scaliger; but the difference is that Scaliger discussed the philosophical controversies between Aristotle and Democritus, whereas Sennert showed a pronounced eclectic tendency. He was interested mainly in a chemical theory, and he found that from a chemical point of view the two theories really amount to the same thing. In order to support this opinion, Sennert refused to accept the interpretation that Democritus meant to deny the qualitative differences of atoms. As a chemist, Sennert was convinced that elementary atoms differ qualitatively. His main contribution to the corpuscular theory lies in the clear distinction that he made between elementary atoms and atoms of compounds (*prima mista*). This distinction forced itself upon Sennert through chemical experience. Each chemical substance, elementary or compound, must have its own atoms.

Contrary to Sennert, Pierre Gassendi (1592–1655) faithfully copied Epicurus and therefore Democritus as well. His own contribution consisted of a number of annotations designed to make the original atomic doctrine acceptable to his contemporaries. In order to effect this purpose, two things were necessary. First of all, the atomic system had to be divested of the materialistic interpretation with which it was hereditarily connected. Second, Gassendi had to "adapt" the original atomic theory to the science of his time. Science had reached the stage at which certain definite physical and chemical properties were attributed to the atoms—i.e., the atoms must possess definite natures; they could not be qualitatively equal. For this reason Gassendi stated that from the original atoms certain molecules were formed first; these differed from each other and were the seeds of different things.

While Gassendi's system is basically without any trace of originality, the corpuscular theory of René Descartes (1596–1650) is original in outline and execution. According to Descartes, matter and extension are identical. This thesis of course excludes the idea of indivisible atoms, but not of smallest particles. To the question of how such particles are separate and distinct from each other, Descartes answered that when a quantity of matter moves together, that quantity forms a unit, distinct from other units that have different motions. Along these lines, Descartes succeeded in devising a corpuscular theory in which the corpuscles were characterized by differences in mass, in amount of motion, and other properties that could be expressed in physical terms and treated mathematically. Descartes's corpuscles were endowed with exactly those properties that could be used in contemporaneous mechanics. As we have seen with Sennert, the seventeenth century was less interested in philosophical considerations than in scientifically fruitful ideas. Therefore, a corpuscular theory was judged, first of all, by this standard; and underlying philosophical discrepancies did not much interest the scientist. This explains why, to their contemporaries, Gassendi and Descartes could stand fraternally united as the renovators of the atomic theory.

Robert Boyle (1627–1691), for example, repeatedly confessed how much both Descartes and Gassendi had inspired him. On the other hand, Boyle was too much a chemist to be satisfied with a general idea of atoms or even with atoms endowed only with mechanical properties. Boyle looked for specific chemical properties. In contrast with mechanics, however, chemistry was not yet sufficiently developed to provide the theoretical frame-

work necessary for a satisfactory chemical atomic theory. Boyle was keenly aware of this situation, as his *The Sceptical Chymist* (Oxford, 1661) proves. Neither the traditional theory of four elements nor the three-principle theory current among chemists could be of any use to him. Yet he was convinced that the distinction between elements and compounds was a sound one. This distinction therefore governed his own atomic theory. Theoretically, he adhered to the atoms of Democritus; practically, he did not use them. He was convinced that atoms were associated into so-called primary concretions, "which were not easily dissipable into such particles as composed them." Thus the primary concretions were corpuscles with definite qualities; they corresponded to the smallest particles of elements, and consequently Boyle treated them as such. The primary concretions could combine to form compounds of a higher order that may be compared with Sennert's *prima mista.* Although Sennert's corpuscular theory was based more on the minima theory and Boyle's theory more on the ideas of Gassendi and Descartes, in practice their theories were not very different. Both theories recognized atoms of compounds that are composed of atoms of elements. For Sennert the latter were elements, both theoretically and practically. For Boyle, theoretically they were not elements, but practically they were, because in chemical and physical processes primary concretions are not dissolved.

By combining the relative merits of the minima theory (qualitative atoms) and of Democritus' atomism (open to quantitative treatment), the seventeenth century laid the foundations for the scientific atomic theory of the nineteenth century. The further development of the seventeenth-century atomic theory, however, required better chemical insights, and especially a method of distinguishing elementary from compound substances. This method was found by Antoine Lavoisier (1743–1794), who postulated the conservation of weight as the guiding principle in chemical analysis. For the first time in history, a list of chemical elements could be given, based upon the results of chemical analysis.

The outstanding achievement of John Dalton (1766–1844) was that he connected these chemical results with the atomic theory. His atoms were no longer smallest particles with some general and rather vague physical properties, but atoms endowed with the properties of chemical elements. Dalton himself in *A New System of Chemical Philosophy* stressed the great importance of "ascertaining the relative weights of the ultimate particles, both of simple and compound bodies, the number of simple elementary particles which constitute one compound particle, and the number of less compound particles which enter into the formation of one more compound particle" (2nd ed., p. 213).

The fact that Dalton's theory is primarily a chemical theory does not mean that it has no philosophical implications. It is interesting to note that Dalton conceived the union of atoms in a compound as their simple juxtaposition without their undergoing any internal change. On this point the founder of the chemical atomic theory did not differ from the Democritean tradition. On another point, however, he followed the minima tradition. Dalton's atoms were specifically different for every kind of substance. He did not even think of building these atoms from particles without qualities, as Gassendi and Boyle had done.

After Dalton, the development of the atomic theory was very rapid. Jöns Jakob Berzelius (1779–1848) determined the relative atomic weights with surprising accuracy, guided by the hypothesis that under the same pressure and at the same temperature the number of atoms in all gaseous substances is the same. Since hydrogen and oxygen combine in the constant volume proportion of two to one, Berzelius concluded correctly that two atoms of hydrogen combine with one atom of oxygen. Berzelius also gave to chemistry its modern symbols. Amedeo Avogadro (1776–1856) completed the atomic theory by assuming that compound atoms, or molecules, do not necessarily have to be formed out of atoms of different elements; molecules of elements (H_2; O_2) also exist. According to Avogadro, the law that postulated an equal number of atoms in equal volumes of gas had to be understood as applying to an equal number of molecules. In a short time, the framework for classical chemistry was completed on the basis of Dalton's atomic theory. Chemical reactions were conceived of as a reshuffling of atoms and described by such chemical equations as $2\ H_2 + O_2 \rightarrow 2\ H_2O$.

An important contribution to the development of the atomic-molecular theory came from physics in the form of the kinetic theory of gases. With the aid of the calculus of probability, James Maxwell and Ludwig Boltzmann succeeded in deriving the behavior of gases, as described in the empirical laws of Boyle and Joseph-Louis Gay-Lussac, from the motions of the molecules.

The discovery of the electron, the electric atom, paved the way for a new theory about the nature of chemical compounds and chemical reactions. According to the new theory, a molecule such as NaCl did not consist of an Na atom and a Cl atom, but of an Na ion and a Cl ion; the Na ion was an Na atom minus an electron, and the Cl ion

was a Cl atom plus an electron. Thus the so-called ionic theory revealed the nature of the forces of attraction between the various atoms of a molecule. The Na ion with its positive electric charge was attracted by the Cl ion with its negative charge. As a result of the connection that the theory of electricity established between physics and chemistry, theoretical and experimental materials were available at the beginning of the twentieth century. They led to a new development of the atomic theory that would endeavor to penetrate into the interior of Dalton's atoms.

The atomic model of Niels Bohr (1913) considered every atom as built of a positively charged nucleus around which circled, in fixed orbits as many electrons as were indicated by the charge of the nucleus. This charge corresponded to the place of the element in the periodic system. Bohr's model could explain not only the fundamental chemical properties of the elements, but also such physical properties as the spectrum that is characteristic of each element when it is emitting or absorbing light. Nevertheless, there were also serious difficulties with this model. According to electrodynamics, the moving electrons would ceaselessly emit electromagnetic waves. The atom would not be stable, but would always be losing energy. Hence, the motion of the electrons would gradually decrease and finally cease entirely. In order to save his model, Bohr postulated that emission of energy occurs only when an electron "jumps" from one orbit to another. In other words, the emission of energy is discontinuous. The emitted energy could be only a whole multiple of an elementary quantity of energy.

Thus, following the work of Max Planck, the idea of minima of energy was added to the idea of minima of matter, the traditional basis of atomism. Even light seemed to show an atomistic structure (photon theory). This would have meant a complete victory for the atomistic view if there had not been a complication. This complication was that the reasons which had formerly settled the dispute about the nature of light in favor of Christian Huygens's wave theory against Isaac Newton's corpuscular theory still retained their value. Light showed a dual character. In 1924, it occurred to Louis de Broglie that the same dualism might very well apply to the particles of matter. On the basis of this hypothesis, he could readily explain Bohr's postulate. This resulted in quantum mechanics, a new theory propounded by Erwin Schrödinger and Werner C. Heisenberg, which showed that both the atomic theory and the wave theory were only approximate models and not adequate representations of material reality.

The evolution of the atomic theory in the twentieth century was not limited to these rather startling new theoretical developments; it also gave rise to a new branch of physical science, nuclear physics, which studies the changes that the atomic nucleus is subject to. The first work in this area was in connection with the study of natural radioactivity. It had been observed that through radiation the nucleus of one element changes in charge and mass and thus becomes the nucleus of another element. In 1919 Ernest Rutherford succeeded in effecting an "artificial" transmutation; many others followed. The atoms of chemical elements appeared to be composed like the molecules of chemical compounds. Through nuclear processes a confusingly great number of new elementary particles has been discovered, all of which are subject to transformation under certain conditions. Particles can be changed into other particles and even into radiation. With such transmutations enormous amounts of energy are released.

Thus, twentieth-century science revolutionized many fundamental ideas of the nineteenth century; the atom is not only much more complex than Dalton thought; it is also much more dynamic. Yet Dalton is far from antiquated. Modern chemistry still works along the lines drawn by Dalton and his contemporaries. Can the same be said in relation to his forerunners in the philosophical period of atomism? The answer to this question can be found in the fact that the main mistake of Dalton and other advocates of essentially mechanistic theories lay in the conviction that atoms did not undergo any internal change. Science showed that this assumption was erroneous, but this should not be a de facto statement only. For if we think of the nature of science as experimental, then it is clear that unchangeable atoms would not offer any possibility of being investigated by experimental means. Without change, matter could not respond to experimental questions. Classical science could overlook this simple truth by assuming that it already knew all the relevant features of atoms. This assumption followed from the mechanistic doctrine that, from the seventeenth century onward, formed the philosophical background of the atomic theory and of classical science in general. The mechanistic doctrine points up the fact that classical science originated in a rationalistic climate. The idea of an unchangeable atom endowed with mechanical properties seemed to be in accordance with what an element should be. It satisfied both the imagination and the intellect. The program of science seemed to consist in explaining the forms of nature on the basis of component elements that were already known.

With the development of science, however, increasing knowledge of chemical compounds affected our understanding of elements. The elements, too, became the object of experimental investigation. From this it may be concluded that the mechanistic doctrine was not a real presupposition of the scientific method. In using the experimental method, science presupposed a much more fundamental mutability in nature than traditional mechanism could account for, and the scientific method implied a much more refined view of material reality than the mechanistic interpretations of science suggested. For this reason, the less orthodox forms of atomism were as important to the origin of the scientific atomic theory as were the orthodox. From the point of view of twentieth- and twenty-first-century science, the Greek philosophical discussions about the nature of change remain amazingly modern.

See also Alexander of Aphrodisias; Anaxagoras of Clazomenae; Aristotle; Averroes; Bohr, Niels; Boltzmann, Ludwig; Boyle, Robert; Chemistry, Philosophy of; Descartes, René; Empedocles; Epicurus; Gassendi, Pierre; Heisenberg, Werner; Lavoisier, Antoine; Leucippus and Democritus; Lucretius; Maxwell, James Clerk; Newton, Isaac; Parmenides of Elea; Philosophy of Science, History of; Planck, Max; Plato; Schrödinger, Erwin; Themistius; Toletus, Francis.

Bibliography

Dalton, John. *A New System of Chemical Philosophy.* London: Bickerstaff, 1808; 2nd ed., 1842.

Dijksterhuis, E. J. *The Mechanization of the World Picture.* Oxford: Clarendon Press, 1961. Excellent history of science from antiquity to the seventeenth century.

Hooykaas, R. "Elementenlehre und Atomistik im 17. Jahrhundert." In his *Die Entfaltung der Wissenschaft,* pp. 47–65. Hamburg: Augustin, 1957.

Lasswitz, K. *Geschichte der Atomistik vom Mittelalter bis Newton,* 2 vols. 2nd ed. Leipzig, 1926. A nineteenth-century classic.

Melsen, A. G. M. van. *From Atomos to Atom, the History of the Concept Atom,* 2nd ed. New York: Harper and Row, 1960. Includes references for the primary sources.

Nash, Leonard K. *The Atomic-Molecular Theory.* Cambridge, MA: Harvard University Press, 1950. Discusses the classical chemical theories.

Scaliger, J. C. *Exotericarum Exercitationum Libri XV de Subtilitate ad Hier.* Frankfurt, 1557; 2nd ed., 1607.

Whittaker, E. *History of the Theories of Aether and Electricity,* 2 vols. New York: Philosophical Library, 1951 and 1954. For readers with a good background in science.

Whyte, L. L. *Essay on Atomism: From Democritus to 1960.* Middletown, CT: Wesleyan University Press, 1961. A brief introduction to the idea of atomism and its history.

Yang, Chen Ning. *Elementary Particles, a Short History of Some Discoveries in Atomic Physics.* Princeton, NJ: Princeton University Press, 1962. Gives a general outline of the research done since 1900.

Andrew G. M. van Melsen (1967)

ATOMISM, LOGICAL

See *Analysis, Philosophical; Russell, Bertrand Arthur William; Wittgenstein, Ludwig Josef Johann*

ATTRIBUTE

See *Subject and Predicate; Universals, A Historical Survey*

AUGUSTINE, ST.

(354–430)

St. Augustine, also known as Aurelius Augustinus, was one of the key figures in the transition from classical antiquity to the Middle Ages. He was born at Thagaste, in north Africa, and died as the invading Vandals were closing in on his episcopal city, Hippo. He lived through nearly eighty years of the social transformation, political upheavals, and military disasters that are often referred to as the "decline of the Roman Empire." His life also spanned one of the most important phases in the transition from Roman paganism to Christianity. The old Roman pagan tradition was by no means dead, although the Roman emperors had been Christians since Constantine's conversion some forty years before Augustine was born. Augustine's youth saw the brief rule of Julian the Apostate as well as the last great pagan reaction in the empire, which broke out in the 390s. Nevertheless, it was during this period that the Roman state adopted Christianity as the official state religion. Medieval Europe began to take shape within the framework of the Roman Empire.

Augustine belonged to the world of late Roman antiquity, and its cultural and educational system had a decisive and lasting role in shaping his mind. His education, following the standard pattern of the time, was almost entirely literary, with great stress on rhetoric. Its aim was to enable its recipients to imitate the great literary masterpieces of the past. It tended, inevitably, to encourage a conservative literary antiquarianism. The culture it produced rarely rose above the level of the ster-

ile cult of "polite letters" and generally had little contact with the deeper forces at work in contemporary society. There were many creative minds still at work; but even at their best, their thought was largely derivative. This is especially true of the philosophy of the period. Its stock of learning was in large part contained in compendia, though works of Cicero were still widely read, and those of the Neoplatonist thinkers gave inspiration to both pagans and Christians.

This culture and its educational system were the two sources that supplied the initial impulse for Augustine's thinking. His search for truth and wisdom began with his reading at the age of eighteen of a now lost dialogue by Cicero, the *Hortensius.* The work made an impact that Augustine could not forget and that he often mentions in his later writings. When he recounts the experience in the *Confessions* (III, 4, 7), written in his forties, he tells us that it was this work that changed his interests and gave his life a new direction and purpose: the search for wisdom. The search led him far afield; but looking back on it, Augustine could interpret its start as the beginning of the journey that was finally to bring him back to God.

PHILOSOPHY AND CHRISTIANITY

It was not until 386 that Augustine was converted to Christianity; he was baptized the following year. Meanwhile, his career as a teacher of rhetoric took him from his native Africa to Italy, first to Rome and then to Milan. During this period he was under the spell of the Manichaean religion. Its teachings appeared for a time to offer Augustine the wisdom for which he had been searching, but he became increasingly dissatisfied with it and finally broke with the sect through the influence of his new friends in Milan, Bishop Ambrose and the circle of Christian Neoplatonists around him. In Milan he learned the answers to the questions that had worried him about Manichaean doctrine, and there he encountered a more satisfying interpretation of Christianity than he had previously found in the simple, unintellectual faith of his mother, Monica. There was no deep gulf between the Christianity of these men and the atmosphere of Neoplatonic thought of the time. At this stage of his life Augustine saw no need to disentangle exactly what belonged to Christian and what to Neoplatonic teaching: What struck him most forcibly was how much the two bodies of thought had in common. The blend of Neoplatonism and Christian belief won his adherence, and the moral conflict recounted in his *Confessions* (Books VI–VIII) ended with his baptism.

Even in 400, when he wrote his *Confessions,* he spoke of the teachings of the "Platonists" as preparing his way to Christianity. In a famous passage (VIII, 9, 13–14) he describes Neoplatonism as containing the distinctive Christian doctrines about God and his Word, the creation of the world, and the presence of the divine light; all these he had encountered in the books of "the Platonists" before reading of them in the Scriptures. What he had failed to find anticipated in Neoplatonism were the beliefs in the Incarnation and the Gospel account of the life and death of Jesus Christ. Later in life Augustine came gradually to see a deeper cleavage between philosophy and Christian faith; but he never ceased to regard much of philosophy, especially that of the Neoplatonists, as containing a large measure of truth and hence as capable of serving as a preparation for Christianity.

From Milan he returned to north Africa and retired to live a kind of monastic life with like-minded friends until he was ordained, under popular pressure, to assist the aged bishop of Hippo as a priest. Within four years, in 395, he became bishop of Hippo. From the 390s onward, all of Augustine's work was devoted to the service of his church. Preaching, administration, travel, and an extensive correspondence took much of his time. He continued to lead a quasi-monastic life with his clergy, however, and the doctrinal conflicts with Manichaeans, Donatists, Pelagians, and even with paganism provoked an extensive literary output. Despite this multifarious activity, Augustine never ceased to be a thinker and scholar, but his gifts and accomplishments were turned increasingly to pastoral uses and to the service of his people. The Scriptures took a deeper hold on his mind, eclipsing the strong philosophical interests of the years immediately preceding and following his conversion.

Augustine did not, however, renounce his philosophical interests. He shared with all his contemporaries the belief that it was the business of philosophy to discover the way to wisdom and thereby to show people the way to happiness or blessedness (*beatitudo*). The chief difference between Christianity and the pagan philosophies was that Christianity considered this way as having been provided in Jesus Christ. Christianity could still be thought of as a philosophy, however, in that its aim was the same as that of other philosophic schools. The ultimate source of the saving truths taught by Christianity was the Scriptures, which for Augustine had supplanted the teachings of the philosophers as the gateway to truth. Hence, authority rather than reasoning, faith rather than understanding, came to be the emphasis of "Christian philosophy." For although the pagan philosophers had discovered much of

the truth proclaimed by the Christian Gospel, what their abstract speculation had not, and could not have, reached was the kernel of the Christian faith: the belief in the contingent historical facts that constitute the history of salvation—the Gospel narrative of the earthly life, death, and resurrection of Jesus.

BELIEF AND UNDERSTANDING

Belief in the above facts was the essential first step along the way to saving truth and blessedness, but it was only a first step. Faith, while required of a Christian, was not in itself sufficient for a full realization of the potential rationality of man. For Augustine, an act of faith, or belief, was an act of rational thinking, but of an imperfect and rudimentary kind. In a late work he defined "to believe" as "to think with assent" (*De Praedest. Sanct.* 2, 5). The act of believing is, therefore, itself an act of thinking and part of a context of thought. What distinguishes it from understanding or knowledge is best brought out by Augustine in passages where he contrasts believing with "seeing." By "seeing" Augustine meant either vision, literally, or, metaphorically, the kind of knowledge to which its object is clear and transparent. This kind of knowledge could be acquired only through direct experience or through logical demonstration, such as is possible in mathematics and other forms of rigorous reasoning. Believing, though a necessary and ubiquitous state of mind without which everyday life would be impossible, is therefore a form of knowledge inferior to understanding. Its object remains distant and obscure to the mind, and it is not intellectually satisfying. Faith demands completion in understanding.

In this emphasis on the priority of belief and its incompleteness without understanding, we may see a reflection of Augustine's own intellectual pilgrimage. His tortuous quest for wisdom, with its false trails, had ultimately led him to consider the Christian faith as the object of his search. But this faith offered no resting place, for Augustine never lost his passion for further intellectual inquiry. His faith was only the first step on the way to understanding. He never ceased to regard mere faith as only a beginning; he often returned to one of his most characteristic exhortations: "Believe in order that you may understand; Unless you shall believe, you shall not understand." The understanding he had in mind could be fully achieved only in the vision of God face to face in the life of blessedness; but even in this life, faith could be—and had to be—intensified in the mind by seeking a deeper insight into it. Progress in understanding, founded on faith and proceeding within its framework, was part of the growth of faith itself. After his conversion, then, reasoning and understanding were for Augustine no longer an independent, alternative route to faith. They still had their work, but now within a new setting and on a new foundation.

Some things, like contingent historical truths, could be the objects only of belief; others could be the objects of either belief or understanding (understanding means having an awareness of grounds and logical necessity). For instance, a mathematical theorem can be believed before it is understood. With understanding, however, belief inevitably follows. God, Augustine thought, belongs among the objects that are first believed and subsequently understood. In the process of gaining this understanding, the ordinary human endowments of rational thought, culture, and philosophy have a part to play. They form the equipment of which a Christian may avail himself in the work of seeking deeper insight into the meaning of his faith.

In his *De Doctrina Christiana* Augustine discusses the ways in which the various intellectual disciplines may serve to assist the Christian in understanding the faith he derives from scriptural sources. Philosophy, along with the other branches of learning, is here seen as subordinated to the service of a purpose outside it, that of nourishing and deepening faith; it is no longer to be pursued for its own sake, as an independent avenue to truth. It is also in *De Doctrina Christiana* that Augustine uses the image of the children of Israel, on their way to the Promised Land, spoiling the Egyptians of their treasures at God's bidding: In the same way, Christians are bidden to take from the pagans whatever is serviceable in understanding and preaching the Gospel. Again, we may see here a reflection of Augustine's narrowing of interests and the growing dominance of pastoral concerns in his mind. The theoretical statement of his subordination of secular learning and culture and their consecration to the service of preaching the Gospel (in its widest sense) is contained in the program laid down in the *De Doctrina Christiana.*

Therefore, Augustine is not interested in philosophy, in the modern sense of the word. Philosophical concepts and arguments play a subordinate role in his work; and where they occur, they are usually employed to help in the elucidation of some aspect of Christian doctrine. Typical examples are his use of Aristotle's *Categories* in an attempt to elucidate the notions of substance and relation in the context of Trinitarian theology, especially in his great work *De Trinitate*; his subtle inquiries into human knowledge and emotions, in the second half of the same work, with a view to discovering in man's mind an image

of God's three-in-oneness; and his analysis of the temporal relations "before" and "after," undertaken to elucidate the nature of time in order to solve some of the puzzles presented by the scriptural doctrine of the creation of the world. In all these cases and many more, his purpose would be described today as theological. In Augustine's day the distinction between theology and philosophy did not exist, and "philosophy" could be—and often was—used in a sense so wide as to include what we should call theology.

To study Augustine's thought *as philosophy* is in a sense, to do violence to it: It is to isolate from their purpose and context what he would have regarded as mere techniques and instruments. To focus attention on what Augustine would have regarded as belonging to the sphere of means, however, allows us to see something more than a mere agglomeration of philosophical commonplaces derived, in large measure, from Neoplatonism. Augustine's originality lies not only in his determination to use his inherited philosophical equipment but also in the often slight, but sometimes profound, modification it underwent at his hands. And in the service of Augustine's purpose, many old ideas received new coherence and new power to move. Through his "spoiling of the Egyptians" much of the heritage of late antiquity received a new life in the European Middle Ages.

THE MIND AND KNOWLEDGE

At an early stage of Augustine's intellectual development, the skepticism of the Academic tradition of philosophy appears to have presented him with a serious challenge. His early philosophical dialogues, written in the period immediately after his conversion, are full of attempts to satisfy himself that there are at least some inescapable certainties in human knowledge on which we may absolutely rely. The basic facts of being alive, of thinking, or of simply existing are disclosed in one's immediate awareness of oneself. But Augustine did not limit the range of what was indubitably reliable in one's experience; nor did he seek to build an entire structure of indubitable knowledge on the basis of the absolute certainties of immediate awareness and its strict logical consequences, as René Descartes was to do. He tried instead to vindicate the whole range of human knowledge as being capable of arriving at truth, though also liable to err.

His vindication proceeds on two fronts, according to the fundamental duality of knowledge and of the objects corresponding to it. This duality, like much in his theory of knowledge, is of Platonic origin. Plato is the source of his belief that "there are two worlds, an intelligible world where truth itself dwells, and this sensible world which we perceive by sight and touch" (*C. Acad.* III, 17, 37); and of its corollary, that things can be divided into those "which the mind knows through the bodily senses" and those "which it perceives through itself" (*De Trin.* XV, 12, 21). Although he never departed from this dualistic theory of knowledge, Augustine also always insisted that all knowledge, of either kind, is a function of the mind, or the soul.

He defines the soul as "a substance endowed with reason and fitted to rule a body" (*De Quant. Anim.* 13, 22). Augustine's use of the conceptual framework of the Platonic tradition made it difficult for him to treat man as a single, substantial whole. He did, nevertheless, attempt to stress the unity of body and soul in man as far as his inherited conceptual framework allowed. In a characteristically Platonic formula he defines man as "a rational soul using a mortal and material body" (*De Mor. Eccles.* I, 27, 52). The soul is one of two elements in the composite, but it is clearly the dominant partner: The relation between it and its body is conceived on the model of ruler and ruled, or of user and tool. This conception gave Augustine considerable trouble in his attempt to work out a theory of sense knowledge.

SENSE AND IMAGINATION. It was a basic axiom of Augustine's view of soul and body that while the soul can act on the body, the body cannot act on the soul. This is a consequence of the user-tool model in terms of which he understood their relation. The tool cannot wield its user; the inferior in nature has no power to effect or induce any modification in the higher. Augustine could not, therefore, elaborate a theory of sense knowledge in which the bodily affections would in any way cause or give rise to modifications in the soul; nevertheless, he insisted that even sense perception was a function of the soul, one that it carried out through the bodily sense organs. The mere modification of a sense organ is not in itself sense experience, unless it is in some way noticed by the mind. Augustine's problem was to explain this correlation between the mind's awareness and the modification of the organ without allowing the latter to cause or to give rise to the former.

In an early discussion of this problem, Augustine tried to explain the process of seeing as a kind of manipulation by the mind of its sense organs, much like a blind man's manipulation of a stick to explore the surface of an object (*De Quant. Anim.* 23, 41–32, 69). This is very much in line with his general conception of the relation of the body to the mind as that of an instrument to its user, but

its inadequacy as an explanation of sense perception may have been apparent to Augustine. At any rate, he later came to prefer an account constructed in quite different terms. This account (elaborated in *De Genesi ad Litteram,* Book XII and generally underlying his later views, for instance, those stated in *De Trinitate*) is based on a distinction between "corporeal" and "spiritual" sight. Corporeal sight is the modification undergone by the eyes in the process of seeing and is the result of their encounter with the object seen. Spiritual sight is the mental process that accompanies corporeal sight, in the absence of which the physical process cannot be reckoned as sense experience (since all experience is a function of mind). Spiritual seeing is not, however, caused by corporeal seeing, since the body cannot affect the mind. Indeed, spiritual sight is a separate process that may take place in the mind spontaneously, in the absence of its corporeal counterpart—for instance, in dreaming or imagining. The mental processes involved in sight and in dreaming and imagination are identical; what is before the mind is, in all these cases, of the same nature. What the mind sees in each case is not the object outside it, but the image within it. The difference between sensation and imagination is that in sensation a process of corporeal seeing accompanies the mental process; this is absent in imagination.

Augustine never quite answers the question of how we may know the difference between perception and imagination. The part, however, which he attributes to attention in the process of sense perception is important and gives a clue: It is attention that directs the mind's gaze, and it appears that it is attention that checks the free play of imagery in the mind. Thus, perception and imagination can be distinguished in experience by adverting to the presence of attention; its presence immobilizes the creative imagaination and ensures that the content of the mind has some sort of rapport with the bodily senses and their world. It is difficult to escape the impression that under the guise of "attention" Augustine has introduced what he had begun by excluding—mental process as responsive to bodily change. This is the peculiar difficulty that his two-level theory of man never quite allowed him to escape.

Augustine also speaks of a third kind of sight, one that he calls intellectual. This, the highest kind of sight, is the work of the mind whereby it interprets, judges, or corrects "messages" from the lower kinds of sight. The type of activity Augustine has in mind here is exemplified by any act of judgment on the content of sense perception; for instance, the judgment that an oar partly submerged in water is not actually bent, even though it looks bent. This activity of interpretation and judgment brings us to the second kind of knowledge, that which the mind has independently of sense experience.

REASON AND ILLUMINATION. In his account of sense knowledge, Augustine's Platonic inheritance was a source of difficulty. In the elaboration of his views on reason and intelligence, the reverse is the case: Augustine's account of these is largely an adaptation of the fundamental tenets of the Platonic tradition. Typical instances of knowledge that the mind has independently of sense experience are the truths of mathematics. Here Augustine discovered the universality, necessity, and immutability that he saw as the hallmarks of truth. Although he did not believe that knowledge obtained through the senses possessed these characteristics, Augustine widened the scope of truth considerably beyond the necessary truths of mathematics and logic. He thought that our moral judgments and judgments of value, at least of the more fundamental kind, also shared the character of truth. He did not, however, trace this universality and necessity of such propositions to their logical form or to the nature of the definitions and logical operations involved in them. (He wrote fourteen centuries before Immanuel Kant's distinction between analytic and synthetic judgments.)

Like all his predecessors and contemporaries, Augustine thought that this kind of knowledge was just as empirical as sense experience, and that it differed from the latter only in having objects that were themselves superior to the physical objects of sense experience by being immutable and eternal, and therefore capable of being known with superior clarity and certainty. The knowledge open to the mind without the mediation of the senses was conceived as analogous to sight; indeed, Augustine often speaks of it as sight, sometimes qualifying it as "intellectual sight." Its objects are public, "out there," and independent of the mind that knows them, just as are those of physical sight. In its knowing, the mind discovers the objects; it does not create them any more than the eyes create the physical objects seen by them. Together, the truths accessible to this kind of knowledge form a realm that Augustine, following the whole Platonic tradition of thought, often calls the intelligible world. This he identifies with the "Divine Mind" containing the archetypal ideas of all things. He was not, however, the first to take this step; this identification was the key to all forms of Christian Platonism.

Before Augustine, Plato had already used the analogy between sight and understanding. Its details are worked out in the analogy of the sun in the *Republic.* Here the

intellectual "light" that belongs to the world of intelligible forms is analogous to the visible light of the material world. Like the latter, it renders "visible" the objects seen by illuminating both them and the organ of perception—in this case, the mind. All understanding is a function of illumination by this light. The intellectual light that illuminates the mind and thus brings about understanding is spoken of in various ways by Augustine. Since it is a part of the intelligible world, it is naturally conceived as a kind of emanation from the divine mind or as an illumination of the human mind by the divine. Augustine also refers to it as the human mind's participation in the Word of God, as God's interior presence to the mind, or even as Christ dwelling in the mind and teaching it from within.

Plato had tried to account for the mind's knowledge of the forms in the theory, expressed in the language of myth, that this knowledge was left behind in the mind as a memory of its life among the forms before it was enclosed in an earthly body. After some early flirtation with this theory of reminiscence, Augustine came to reject it; to hold that the mind's knowledge derived from a premundane existence would have raised serious theological difficulties. Therefore, instead of tracing this knowledge to a residue of a past experience, he accounted for it in terms of present experience; it was the result of continual discovery in the divine light always present to the mind. For this reason, too, his conception of *memoria* became so widened as to lose the reference to past experience that *memory* necessarily implies in English. Augustine's *memoria* included what we should call memory; in it, he thought, were preserved traces of past experience, as in a kind of storehouse or a stomach. But *memoria* included very much more than this. He speaks of our a priori mathematical ideas, numbers and their relations, as being contained in it; and in the course of the tenth book of the *Confessions,* in which he devotes a long discussion to the subject, the scope is so widened as to extend to our knowledge of moral and other values, of all truths of reason, of ourselves, and of God. It is, in effect, identified with all the latent potentialities of the mind for knowledge.

Memoria and divine illumination are alternative ways of expressing the basis of Augustine's theory of knowledge. The theory is, in its essence, the belief that God is always intimately present to the mind, whether this presence is acknowledged or not. His presence pervades everything and is operative in everything that happens. To this metaphysical principle the human mind is no exception. The only difference between the human mind, in respect to the divine presence within it, and other things is that unlike these other things, the human mind is able to turn freely toward the light and to acknowledge its presence, or to turn away from it and to "forget" it. Whether the mind is present to the divine light or not, however, the light is present to the mind; on this presence is founded all the mind's ability to know.

The manner of operation of this illumination in the mind and what exactly it produces in the mind have been the subject of much debate. This uncertainty is due partly to the enormous variety of expressions used by Augustine to describe the divine light, but it is also partly the result of approaching Augustine's views with questions formulated in terms of concepts between which he would not have made a distinction. It is clear, at any rate, that Augustine did not think that the divine light in the mind gave the mind any kind of direct access to an immediate knowledge of God. This kind of knowledge was, to him, the result of understanding, a goal to be reached only at the end of a long process—and not this side of the grave. If, however, we ask further what exactly he thought illumination did reveal to the mind, the answer is more difficult. In particular, if we ask whether he conceived illumination primarily as a source of ideas in the mind or, alternatively, as providing the mind with its rules for judgment, the answer is not at all clear. He did not distinguish as sharply as one might wish between the making of judgments and the formation of concepts; he often speaks of both activities in the same breath or in similar contexts, or passes without the least hesitation from one to the other in the course of discussion. Sometimes he speaks of illumination as implanting in the mind an "impressed notion" (*notio impressa*), whether it be of number, unity, wisdom, blessedness, or goodness. Such passages suggest that Augustine thought of illumination primarily as a source of ideas, as providing "impressed notions." It is clear, however, that such "impressed notions" were also to serve as the yardsticks for judging all imperfect participations in individual instances of these notions. And in other passages, again, illumination is spoken of not as supplying any ideas or notions but simply as providing a criterion of the truth or falsity of our judgments.

It was very easy to pass from ideas to judgments in Augustine's way of speaking of illumination. In addition, Augustine's language when he speaks of the mind's judgment made in the light of divine illumination often has further overtones; the judgment he speaks of appears as a kind of foreshadowing of the ultimate divine judgment on all human life and action. The basic reason why Augustine had found Platonic metaphysics so congenial

was that it harmonized so easily with the moral bearings of his own views; and its theories, especially in some of their more imaginative and dramatic expressions, allowed themselves to be exploited to serve Augustine's interests as a moralist. In his discussion of knowledge, as in his discussion of the relation of mind and body, ethical considerations very often play the major part. The central theories of Platonic thought buttressed views held by Augustine primarily on account of their moral bearings.

WILL, ACTION, AND VIRTUE

Morality lies at the center of Augustine's thought. There are many reasons for this, the most noteworthy being his conception of philosophy. As we have seen, philosophy was for Augustine far from being an exclusively theoretical study; and morality itself belonged to its substance more intimately than the discussion and analysis of moral concepts and judgments. Philosophy was a quest for wisdom, its aim being to achieve man's happiness; and this depended on right living as much as on true thinking. Hence the practical orientation of Augustine's thought—an orientation that it shared with most contemporary forms of thinking.

On human conduct and human destiny Augustine's thinking was, of course, molded very largely by the New Testament and by the Christian church's tradition in understanding its conceptions of divine law and commandment, of grace, of God's will, of sin, and of love. Much of this, being specifically theological in interest, lies outside the scope of this presentation of Augustine's thought. What is remarkable is the extent to which Augustine was prepared to read back the characteristic teaching of the Christian church into the works of the philosophers, Plato in particular. Thus he held that Plato had asserted that the supreme good, possession of which alone gives man blessedness, is God. "And therefore," Augustine concluded, Plato "thought that to be a philosopher is to be a lover of God" (*De Civ. Dei* VIII, 8). Rapprochements of this kind helped to reconcile the Christian and the Platonic teachings to each other; in Augustine's treatment of ethical topics the characteristically Christian themes and distinctively Platonic concepts are so closely interwoven that they are often inseparable.

Augustine is able, therefore, to define blessedness itself in terms that make no reference to any distinctively Christian teaching, for instance, when he says that man is blessed when all his actions are in harmony with reason and truth (*cum omnes motus eius rationi veritatique consentiunt—De Gen. C. Man.* I, 20, 31). Blessedness, according to this view, does not consist simply in the total satisfaction of all desires. In another discussion Augustine makes this more explicit: While blessedness is incompatible with unsatisfied desires, the satisfaction of evil or perverse desires gives no ultimate happiness; hence blessedness cannot be identified simply with total satisfaction. "No one is happy unless he has all he wants and wants nothing that is evil" (*De Trin.* XIII, 5, 8; for the entire discussion, see ibid. XIII, 3, 6–9, 12). The only element in all this that is specifically Christian is the insistence that this happiness cannot be attained by man except with the aid of the way revealed by Christ and of God's grace given to men to enable them to follow it.

The dramatic account, given in his *Confessions,* of his own turning to God, though steeped in the language of the Bible and throbbing with the intensity of Augustine's feelings, is, at the same time, an illustration of a central theme in Greek metaphysics. The book opens with a powerful evocation of his coming to rest in God; it ends with a prayer for this rest, peace, and fulfillment. This central theme of longing and satisfaction is a commonplace of Greek thought from Plato's *Symposium* onward. Man, according to the cosmology implicit in this picture, illustrates in his being the forces that are at work in nature in general. Man, like everything else, is conceived as part of a vast nexus of interrelated things within an ordered hierarchy of beings that together form the cosmos. But it is an order in which the components are not stationary but are in dynamic rapport; they are all pursuing their own ends and come to rest only in attaining these ends. Their striving for rest, for completion or satisfaction, is the motive power that drives all things toward their purposes, just as weight, according to this image, causes things to move to the places proper to them in the cosmos—the heavy things downward, the light upward. Augustine thought of the forces that move men as analogous to weight and called them, collectively, love or loves. In a famous passage he wrote, "My weight is my love; by it am I carried wheresoever I am carried" (… *eo feror quocumque feror—Conf.* XIII, 9, 10).

LOVE, LAW, AND THE MORAL ORDER. Man, however, differs from other things in nature in that the forces that move him, his "loves," are very much more complex. Within him there are a great many desires and drives, impulses and inclinations—some of them conscious, others not. The satisfaction of some often involves the frustration of others, and the harmonious satisfaction that forms the goal of human activity appears to be a very distant and scarcely realizable purpose. The reason for this is not only the multiplicity of elements that go into the making of human nature; a further reason is the fact

that these elements have been disordered and deprived of their original state of harmony. Augustine interpreted this aspect of the human condition as a consequence of the sin of Adam and the fall of man.

There is, however, a further respect in which man differs from other things in the way his activity is determined. This lies in the fact that even with his disordered impulses, he is not—at least not entirely—at the mercy of the conflicting forces within him. His activity is not, so to speak, a resultant of them: He is, in some degree, capable of selecting among them, deciding which to resist, which to follow. In this capacity for choice Augustine saw the possibility of what he called voluntary action as distinguished from natural or necessary behavior. He called this human capacity "will." It is a source of some confusion that he used the term *love,* or its plural, *loves,* to designate the sum total of forces that determine a man's actions, whether they are "natural" or "voluntary." As a collective name for natural impulses, "love" is therefore morally neutral; only insofar as the will endorses or approves love of this kind is love morally praiseworthy or blameworthy. Augustine expresses this graphically by distinguishing between loves that ought to be loved and loves that ought not to be loved; and he defines man's moral task in terms of sorting out these commendable and reprehensible loves in himself and putting his loves in their right order.

Augustine's favorite definition of virtue is "rightly ordered love" (as in *De Civ. Dei* XV, 22). This consists in setting things in their right order of priority, valuing them according to their true worth, and in following this right order of value in one's inclinations and actions. The idea of order is central to Augustine's reflections on morals. Before becoming a Christian, he had believed with the Manichaeans that the existence of good and of evil in the world was accounted for by their different origins, respectively from a good and an evil deity. The Neoplatonism of his Christian friends in Milan helped Augustine find an alternative explanation, one that was more in keeping with the Christian doctrine of one world created by one God. According to this theory, evil had no independent, substantial existence in its own right; it existed as a privation, as a distortion or damage within the good. All evil was thus in some sense a breach of the right relation of parts within a whole, a breach of order of some kind. Hence the great emphasis on order in Augustine's thought, from the time of his conversion to the writing of his last works.

Augustine calls the pattern to which human activity must conform "law." Law is, in the first place, the archetypal order according to which people are required to shape their actions and by which their actions are to be judged. Augustine makes it clear that by "law" he means very much more than the actual legal enactments of public authorities. These "human laws" deal only with a part, greater or lesser, of human conduct; they vary from place to place and from time to time; they depend on the vagaries of individual legislators. The true "eternal law" by which all human behavior is judged leaves no aspect of man's life out of its purview; it is the same everywhere and at all times. It is not quite clear how Augustine conceived the relation between divine and human, eternal and temporal, law. His terminology is variable, and although he thought that human law ought to seek to approach the divine, or at least not to contradict it, he does not appear to have denied its claim to being law even when it failed to reflect the eternal law. Also, as we shall see, he appears to have changed his views on this matter in the course of his life.

The "eternal," or "divine," law is in effect the intelligible world or the divine mind (see discussion of reason and illumination above) insofar as it is considered as the pattern that should regulate activity. The language in which Augustine speaks about the divine law is the same as that which he uses in speaking of the eternal truth, and he believed that the achievement of wisdom consisted in pursuing this truth by understanding and then embodying in oneself the order understood. It is clear that there is no significant difference between "eternal law" and "eternal truth"; the two are identical: Eternal law is eternal truth considered under its aspect as a standard of moral judgment. Thus, the problem of how the eternal law is known to men is the same as the problem discussed above of how the eternal truth is known. Here, too, he speaks of the eternal law as being "transcribed" into the human mind or of its "notion" as being impressed on the mind. The deliverance of conscience or reason as manifested in moral judgment is thus no less and no more than the human mind's illumination by the eternal law, or its participation in it; Augustine describes conscience as "an interior law, written in the heart itself" (*lex intima, in ipso … corde conscripta*—*En. in Ps.* 57, 1). He refers to this law, inscribed in man's heart or known to him by reason, as "natural." He can thus speak of law (eternal or natural), reason, and order interchangeably when discussing the ordering of human action to bring about its virtuous disposition.

In defining this order of priority in value, the following of which constitutes virtue, Augustine makes a fundamental distinction between "use" and "enjoyment." These

two forms of behavior correspond to the twofold classification of things according to whether they are valuable for their own sake or as means, for the sake of something else. Things valued for themselves are to be "enjoyed," things valued as means are to be "used"; the inversion of the relation between use and enjoyment is the fundamental perversion of the order of virtue. To seek to use what is to be enjoyed or to enjoy what is to be used is to confuse means with ends. The only object fit for enjoyment, in this sense, is God; he alone is to be loved for his own sake, and all other things are to be referred to this love. In elaborating this theory, Augustine was expressing the traditional view that it behooves people to journey through their lives on Earth as pilgrims and not to regard any earthly goal as a fit resting place. This did not, of course, imply, to Augustine's mind, that nothing but God was a fit object of love; on the contrary, it was a way of stressing the need to put loves in their right order and to love each thing with the kind and degree of love appropriate to it. Although he clearly conceived of love as capable of an endless series of gradations, Augustine is usually content to speak of two kinds of love, which he contrasts: charity (*caritas*) and cupidity (*cupiditas*). The basic distinction is between upright, well-ordered, and God-centered love and perverse, disordered, and self-centered love. A great deal of Augustine's thinking and writing hinges on this distinction.

The individual virtues interested Augustine less than the concept of love. He was content to take over the classical enumeration of the four cardinal virtues. But his own characteristic thoughts on the moral life are always developed in terms of love rather than of any of the virtues. Indeed, as we have seen, he defined virtue in terms of love; similarly, he liked to define the individual cardinal virtues as different aspects of the love of God. This tendency is one of the most important links between what we would distinguish as the theological and philosophical sides of his thought.

THE WORLD AND GOD

Order is a key idea in Augustine's reflections on the morality of human behavior. It also plays a large part in his reflection on the physical universe in its relation to God. The world of nature was not in itself an object of particular interest to Augustine. In cosmological thinking of the kind to be found in Aristotle's *Physics,* for instance, he had little interest. The physical world concerned him only insofar as it was related either to man or to God. Order, then, for Augustine was the expression of rationality. In human action this was something that men should seek to embody in their conduct; in the world of physical and animate nature, which did not share the freedom of human activity, order expressed the divine rationality at work in all natural happenings. To human eyes, however, this order was often glimpsed only in isolated instances, while a great deal of disorder was manifest in the misery, disease, and suffering with which the world is shot through. In part these frustrations of order were held to be due, ultimately, to the initiative of human sin; in part they were held to be merely apparent and capable of being resolved within a perspective larger than that of finite human vision.

Behind the world order stands its author and sovereign ruler, God. All things testify to his presence; the world is full of his "traces" (*vestigia*). God's presence in and behind his creation was, for Augustine, not so much something to be established by argument as it was the premise, taken for granted, of a further argument. This argument, to which Augustine returned on a number of occasions, is particularly well expressed in a chapter of his *Confessions* (X, 6, 9, and 10). He there speaks of putting things to the question in order to allow them to reveal themselves as dependent on their creator. It is clear that what primarily interested Augustine was the questioner's moral attitude: The point of his argument is not so much that the order and beauty of things imply the existence of God, but rather that since God had created them, we must so discipline ourselves as to see things for what they are—his handiwork—and to value them at their true worth and worship only him, their creator—not his handiwork. Again, the moral concern is uppermost in Augustine's mind.

This is not the case with the discussion of the problem of time, in Book XI of the *Confessions.* The problem was forced on Augustine's attention by the scriptural doctrine of creation, but it is clear that it fascinated him and that he pursued it simply because he was interested in it. Manichaean objectors to the Christian doctrine of creation from nothing had raised difficulties about speaking of an absolute beginning. These critics had pointed out that in our ordinary language there is no room for an absolute beginning of the kind envisaged by adherents of the doctrine; we can always ask what happened before something else, even if this was the first of all happenings. Questions of this kind revealed the arbitrariness and absurdity of the belief that God made the world out of nothing: What was God doing before the creation? Why did he create the world when he did and not sooner, or later?

In answer to these difficulties Augustine in effect undertook a critique of the conception of time that underlay them. Such difficulties arise from the fact that time is thought of as having the same kind of being as the events and happenings going on in time; the question "What happened before time?" was thought to be of the same logical form as questions about what happened before any particular events. Augustine denied this assumed logical similarity behind the grammatical similarity of the questions. He pointed out that whereas it makes sense to ask what happened before any particular event, it does not make sense to ask what happened before all events, because time is the field of the relationships of temporal events, and there could ex hypothesi be nothing before the first temporal event. In this argument Augustine in effect rejected the conception of time according to which time has a substantial reality of its own, and he adopted a theory according to which time is the field of temporal relations between temporal events.

He did, however, go further in his reflections on time. Neoplatonic thought had always treated time in close relation to the soul, and Augustine could scarcely avoid discussing this topic. The reality of the past and of the future puzzled him: Can what is not yet but will be, and what is no longer but has been, be said to *be?* If not, then only the present has any reality. But if only the present is real, then reality shrinks to a dimensionless point at which the future is becoming the past. Augustine resolved the whole problem by locating time in the mind and adopting at the end of his discussion, though with hesitation, a definition of time as "extension [*distentio*], I am not sure of what, probably of the mind itself" (*Confessions* XI, 26, 33).

Another question that the doctrine of creation raised for Augustine concerns the natural activity, functioning, and development of creatures. This problem arose from the need to harmonize the story of the creation of the world in seven days or, according to an alternative version, at once, with the fact that some things came into existence only after the creation took place. Augustine's solution of this problem lay essentially in asserting that God created different things in different conditions; some left his hands complete and ready-made, others in a potential or latent state, awaiting the right conditions and environment for their full development. The latter are analogous to seeds, which are thought of as containing in themselves the fully developed plant in potency; and on this analogy, and using the traditional vocabulary, Augustine called these potentialities for later development "seminal reasons" (*rationes seminales,* or *causales*).

Apart from helping him to resolve the apparent contradiction between the belief in a primordial creation and the concept of continued development as a process of natural causality, this theory of "seminal reasons" also prompted Augustine at least to begin to feel his way toward some conception of nature and natural causality. At times, he comes very close to the later medieval distinction between the "First Cause" and the whole range of "second causes," the distinction according to which things depend in different senses both on God (the First Cause) and on their own immediate or distant created causes. Augustine, too, tried to endow the world of created causes with a specific reality of its own, one distinct from the causal activity of God in the world. In this he did not quite succeed. His failure becomes apparent in his treatment of miracles. He did not treat these—as the Scholastics later did as effects of the First Cause (God) produced without the instrumentality of second causes. He allowed the distinction between the two orders of causality (which he had never clearly formulated and which is hinted at, rather than stated, in his writings) to disintegrate during his discussion of miracles. In this context the very idea of "nature" is so widened as to include the miraculous within its scope. Miracles do not contradict the order of nature; they contradict only our idea of this order, an idea based on our restricted view and limited experience. They are not against nature, since nature is God's will; they are only against nature as it is known to us. The distinction between nature and miracle vanishes here, and in his well-known chapter in *The City of God* (X, 12) they become synonymous to the extent that nature itself and man, its crown, become the greatest miracles of all.

INDIVIDUALS IN SOCIETY

Society was not one of the subjects that loomed large in Augustine's earlier thought. Such hints as he gives us of his conception of society in his earlier works (those written before the mid-390s) suggest that he thought that organized human society and the state were part of the worldly dispensation whereby man is assisted to fulfill his destiny. A properly ordered society, like a properly ordered moral life, is a stage on the way to man's ultimate destination in eternity; and as far as Augustine's hints enable us to tell, he expected a properly ordered society to reflect, particularly by means of its legal institutions, the perfection of the eternal, intelligible world.

In step with his theological development, however, his views on human society underwent profound changes, and by the time that society became an impor-

tant theme in his reflection, especially in his great work *The City of God* (written 413–427), these views had been radically transformed. An important factor in the course of this transformation was the increasing stress Augustine had come to lay on the power of sin in human life and in all earthly institutions, on man's need for redemption through Christ, and on his need for grace. In the most general terms Augustine came to see man's destiny and his realization of it more in terms of the scriptural pattern of a redemption-history and less in terms of the Neoplatonic theme of the ascent of the soul. Accordingly, human society came to be understood more in terms of its horizontal, historical relationships within the divine plan for men's salvation and less in terms of what we might call its vertical relationship to the intelligible world.

The first event in the course of the biblical redemption-history, man's fall from grace through Adam's sin, is of decisive importance for Augustine's changed attitude to organized human society. To live in society, according to Augustine, was natural to humans; without society they would not be able to realize fully their human potentialities, and the company of their fellow human beings was necessary to them. This, he held, was as true before man's fall as after; even in his state of primal innocence, in full possession of his nature prior to its distortion by sin, man was a social animal by nature; even the life of the blessed in heaven is a social life. But although Augustine believed that man's nature is social, he did not agree with Aristotle that it is also political. Politically organized society—the machinery of authority, government, and coercion—is, in Augustine's view, not natural to man. It was a useful and necessary arrangement for man in his fallen condition, and indeed the purpose of political society was to remedy at least some of the evils attendant upon man's fallen state. Its function was to check the social disorder and disintegration that followed from the general loss of order at the Fall. The institutions of government, the subjection of governed to government, and the coercive power of political authority over its subjects are thus but one instance of the subjection of man to man, and this was something that, Augustine held, did not exist in man's primal state of innocence. No slavery, servitude, or subjection could exist in that state of natural integrity; these things make sense only if understood as God's punishment for the sin that incurred the loss of integrity and, at the same time, as his dispensation for coping with the needs of man's condition in his new, fallen state.

Augustine used the traditional language of Christian theology to state his view of political society. For reasons to be considered below, he never drew out, at least not explicitly, the full implications of this view. In this view of society, however, the legitimate functions of the state are very much more restricted in scope than in theories according to which man is by nature a political animal. In Augustine's view, the state's sphere is confined to the requirements of social order and welfare; the individual's ultimate welfare and eternal destiny lie outside its realm of competence, whereas they are very much a part of the state's interest if the state is thought of as an ordinance of nature, as an indispensable means of man's realizing his ultimate destiny. In Augustine's estimate, the task of the state in the economy of salvation would be rather to establish the conditions in which men may work out their own salvation in relative peace and security than actively to promote their individual salvation through legislation and coercion.

The state was, for Augustine, synonymous with the Roman Empire; and having revised his ideas on the state in terms of the large categories of the scriptural redemption-history, he had inevitably to take the measure of the state he knew in this same perspective. Here his ideas make sense only if seen as a rejection of views of the empire generally current among Christians during the fourth century, after the adoption of Christianity by the emperors. The empire, represented as eternal ever since Vergil's day, was now widely regarded among Christians as an essential instrument of divine purpose in history, bound up with the possibility of salvation and destined to last until the end of time. It had been taken up into the dimension of the biblical redemption-history. The sack of Rome by the Visigoths in 410 gave a profound shock to this mentality. It led Augustine, whose mind had already moved a long way from the popular picture, to devote his greatest work, *The City of God,* to a reappraisal of the empire's place in the divine providential plan. The upshot was that the empire was no longer allowed an eternal destiny and was removed from the dimension of the redemption-history; the possibility of salvation was not necessarily bound up with it as a means of God's grace. It was simply one of a series of empirical, historic societies. The eternal categories of sin and holiness, of salvation and reprobation, did not apply to it or, indeed, to any other human assembly; they were embodied only in what Augustine called the earthly city and the heavenly city.

The two "cities" consist, respectively, of those predestined to eternal glory and those predestined to eternal torment or, as Augustine also defined them (clearly intending the various definitions to be equivalent), of those who live according to God and those who live

according to man, of the altruistic and the selfish, of those whose love is upright and those whose love is perverse, and so forth. In none of these senses, however, have the two "cities" any discernible reality as communities until their final separation at the Last Judgment. In all discernible human communities they are inextricably intertwined. Here again we may see Augustine's modest estimate of the state's function, for when he discusses it in this context, the realm of the state is identified with the sphere in which the concerns of the two cities overlap. Its task is to secure the temporal peace: the order, security, and material welfare that both the wicked and the righteous cities require during their earthly careers. Its concern is with specifically communal, public matters affecting all its members. Citizens of the heavenly city will not, of course, be content with the welfare and peace thus secured: They will use these things but refer their use to the ultimate enjoyment of a peace beyond the terrestrial.

The general tendency of these views of Augustine's was to undermine the extremely close links that had come to exist between the empire and the Christian church, especially during his own lifetime. He was clearly ill at ease with the current representations of this relationship; but there were considerable pressures working on the minds of his contemporaries to keep them active, and Augustine himself was not exempt from their operation. In the course of the struggle with the Donatist movement in north Africa, a dissenting movement increasingly repressed by the imperial authorities, he came gradually and reluctantly to give his consent to the coercive measures that were being brought into use against the movement. His endorsement of these means of repression ran counter to the most fundamental direction of his thought. Although his endorsement must be regarded as a development in his practical, pastoral, and political attitudes rather than as a reversal of his basic views on the nature of political society, it left deep marks on those views. In later centuries his use of the Gospel phrase "Compel them to come in" (*Coge intrare*—Luke 19:23) and its consecration of repression, persecution, and coercion paved the way to much tragedy. It also helped to obscure the most profound and most original of his contributions to Christian political thinking.

See also Neoplatonism.

Bibliography

The most complete and generally reliable Latin edition of Augustine's works is the edition by the Benedictines of Saint-Maur (1679–1700). It supersedes all earlier editions and was reprinted in J. P. Migne's *Patrologia Latina,* Vols. 32–46 (Paris, 1841–1842), with some unfortunate variants and errors. It is also the basis of the texts of the works now being published in *Bibliothèque augustinienne* (with French translation and useful notes). Modern critical editions of many works exist, mainly in *Corpus Scriptorum Ecclesiasticorum Latinorum* (Vienna) and *Corpus Christianorum.* Details are in E. Dekkers, "Clavis Patrum Latinorum," in *Sacris Erudiri* 3 (1962).

English translations of many works appear in various series, such as Loeb Classical Library, Library of the Fathers, Select Library of Nicene and Post-Nicene Fathers of the Christian Church, Library of Christian Classics, The Fathers of the Church, Ancient Christian Writers, and the Catholic University of America Patristic Studies. A detailed, convenient, and fairly up-to-date list of translations is included in the bibliography by J. J. O'Meara appended to Marrou's *Saint Augustine.*

Of short introductory works, the best is H. I. Marrou, *Saint Augustine* (London: Longmans, 1957), translated from the French in the series Men of Wisdom. It contains a brief biography and a discerning characterization of Augustine's thought by a great scholar, as well as a selection of illustrative texts in translation. R. W. Battenhouse et al., *A Companion to the Study of St. Augustine* (New York: Oxford University Press, 1955), has been superseded by G. Bonner, *St. Augustine of Hippo: Life and Controversies* (London: SCM Press, 1963), as a survey and guide to Augustine's career and literary output. An essential to understanding Augustine in the setting of contemporary education and culture is H. I. Marrou, *Saint Augustine et la fin de la culture antique* (Paris, 1938), completed by his *Retractatio* (Paris, 1949). On Augustine as a bishop, see F. van der Meer, *Augustine the Bishop* (London: Sheed and Ward, 1962).

Among the many books on Augustine's intellectual development and his conversion, P. Alfaric, *L'évolution intellectuelle de saint Augustin* (Paris: Nourry, 1918), stands behind much of the subsequent controversy; C. Boyer, *Christianisme et néo-Platonisme dans la formation de saint Augustin* (Paris: Beauchesne, 1920), is one of the best-balanced replies provoked by it. P. Courcelle, *Recherches sur les Confessions de Saint Augustin* (Paris, 1950), has put the problem on new footing altogether; it is further pursued, with qualifications, by J. J. O'Meara, *The Young Augustine* (London: Longmans Green, 1954).

Of the philosophical aspects of Augustine's thought, the best general account is E. Gilson, *The Christian Philosophy of Saint Augustine,* translated by L. E. M. Lynch (New York: Random House, 1960). See also the survey by R. A. Markus in *Critical History of Western Philosophy,* edited by D. J. O'Connor (New York: Free Press of Glencoe, 1964).

On particular aspects, see C. Boyer, *L'idée de vérité dans la philosophie de saint Augustin* (Paris, 1941); J. Guitton, *Le temps et l'éternité chez Plotin et saint Augustin,* 2nd ed. (Paris, 1955); and E. Dinkler, *Die Anthropologie Augustins* (Stuttgart, 1934), on the topics named in their titles. M. Schmaus, *Die psychologische Trinitätslehre des heiligen Augustinus* (Münster, 1907), is the classic work on Augustine's theory of mind and the Trinitarian speculations based on it. Also valuable on his theory of knowledge are J. Hessen, *Augustins Metaphysik der Erkenntnis,* 2nd printing (Berlin, 1960), and R. Jolivet, *Dieu soleil des esprits* (Paris: Desclée, de Brouwer, 1934).

On ethical topics, see J. Mausbach, *Die Ethik des heiligen Augustins* (Freiburg, 1909); T. Deman, *Le traitement scientifique de la morale chrétienne selon saint Augustin* (Paris, 1957); J. Burnaby, *Amor Dei* (London: Hodder and Stoughton, 1938); and R. Holte, *Béatitude et sagesse: Saint Augustin et le problème de la fin de l'homme dans la philosophie ancienne* (Paris, 1962). On the state and society the least unsatisfactory accounts are Herbert A. Deane, *The Political and Social Ideas of Saint Augustine* (New York: Columbia University Press, 1963) and the short treatment by N. H. Baynes, *The Political Ideas of St. Augustine's De Civitate Dei* (London: G. Bell, 1936).

Many of the most important recent articles on Augustine are to be found in one of the following collections: *Augustinus Magister: Communications et actes du Congrès international augustinien* (Paris: Études Augustiniennes, 1954); *Recherches augustiniennes,* Vols. I and II (Paris: Études Augustiniennes, 1958–1963); *Studia Patristica,* Vol. VI, edited by F. L. Cross; *Texte und Untersuchungen zur Geschichte der altchristlichen Literatur* 81 (1962); and in the quarterly journal *Revue des études augustiniennes.*

The most up-to-date and best-selected bibliography is *Bibliographia Augustiniana,* appended to C. Andresen, ed., *Zum Augustin-Gespräch der Gegenwart* (Darmstadt, 1962), which is Vol. V in the series *Wege der Forschung.* Gilson's book (see above), in which the classified selection of the original is replaced by an alphabetical list in the English translation; B. Altaner, *Patrologie,* 5th printing (1958); and O. Bardenhewer, *Geschichte der altchristlichen Literatur,* Vol. IV (Freiburg, 1924; reprinted 1962), also have good bibliographies. Current literature is surveyed in the bibliographical supplements to *Revue des études augustiniennes.*

R. A. Markus (1967)

AUGUSTINE, ST. [ADDENDUM 1]

St. Augustine continues to elicit scholarly discussions of theological issues, but there is an ever-growing number of studies devoted to historical and philosophical issues in their own right. Recent philosophical work has concentrated on deepening our understanding of his arguments, assessing the adequacy of his positions, and contextualizing them in a historically informed way.

P. Brown, *Augustine of Hippo: A Biography* (London: Faber, 1967), is a masterful work that situates Augustine in his social and historical surroundings. Accessible overviews of Augustine's life and thought are provided in J. J. O'Donnell, *Augustine* (Boston, 1985), and H. Chadwick, *Augustine* (Oxford: Oxford University Press, 1986). Our understanding of Augustine's autobiography has been greatly advanced by the fine commentary given in J. J. O'Donnell, *Augustine: Confessions* (3 vols., Oxford, 1992).

Augustine is seen against the background of classical philosophy in J. Rist, *Augustine: Ancient Thought Baptized* (Cambridge, U.K.: Cambridge University Press., 1994), which provides a guide to Augustine's philosophical views. Another introduction to Augustine as a philosopher is C. Kirwan, *Augustine* (London: Routledge, 1989), which takes up selected topics in detail. The bibliographies of both these works should be consulted as a guide to the literature. Articles on a variety of topics are usefully collected in R. A. Markus (ed.), *Augustine: A Collection of Critical Essays* (London, 1972); there has been no anthology for philosophers as of this writing.

Turning now to particular aspects of Augustine's philosophy, G. J. P. O'Daly, *Augustine's Philosophy of Mind* (London, 1987), and G. Matthews, *Thought's Ego in Augustine and Descartes* (Ithaca, NY: Cornell University Press, 1992), deal with his philosophical psychology. Epistemology and the theory of illumination are the primary focus of R. H. Nash, *The Light of the Mind: St. Augustine's Theory of Knowledge* (Lexington: University Press of Kentucky, 1969); B. Bubacz, *St. Augustine's Theory of Knowledge: A Contemporary Analysis* (New York: E. Mellen Press, 1981); U. Wienbruch, *Erleuchtete Einsicht: Zur Erkenntnislehre Augustins* (Bonn: Bouvier, 1989). Metaphysical problems as well as the issue of Augustine's indebtedness to Plotinus are treated in R. J. O'Connell, *The Origin of the Soul in St. Augustine's Later Works* (New York: Fordham University Press, 1987). Augustine's account of time is analyzed in R. Sorabji, *Time, Creation, and the Continuum* (Ithaca, NY: Cornell University Press, 1983), and subjected to a wide-ranging examination in J. Pelikan, *The Mystery of Continuity: Time and History, Memory and Eternity in the Thought of Saint Augustine* (Charlottesville: University Press of Virginia, 1986). Philosophy of language is discussed by M. Burnyeat, "Wittgenstein and Augustine *de Magistro,*" in *Proceedings of the Aristotelian Society* (1987, supp. vol.). Augustine's ethical theory is the subject of J. Wetzel, *Augustine and the Limits of Virtue* (Cambridge, U.K.: Cambridge University Press, 1992), and discussed in G. R. Evans, *Augustine on Evil* (Cambridge, U.K.: Cambridge University Press, 1982). A recent philosophical study of Augustine's views on freedom, weakness of will, and voluntary action is T. Chappell, *Aristotle and Augustine on Freedom* (New York: St. Martin's Press, 1995). R. Coles, *Self/Power/Other: Political Theory and Dialogical Ethics* (Ithaca, NY: Cornell University Press, 1992), offers a Foucaultian account of Augustine's political philosophy.

A bibliography of works through 1970 is provided in C. Andresen, *Bibliographia Augustiniana* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1973); the next decade of Augustinian studies is covered in T. Miethe, *Augustinian Bibliography, 1970–1980* (Westport, CT: Greenwood Press, 1982).

See also Determinism and Freedom; Metaethics; Philosophy of Language; Plotinus; Time; Time, Being, and Becoming.

Peter King (1996)

AUGUSTINE, ST. [ADDENDUM2]

Augustine thought that what pleases people in beauty is design: "And in design, dimensions; and in dimensions, number" (*De Ordine* ch. 15). Beauty is ultimately a matter of numerical proportion. Rhythm, too, is based on numerical proportions (*De Musica* Book 6). Augustine sees numeric proportions as eternal and divine, yet at the same time he hints that the soul itself may be "the very number by which all things are numbered" (*De Ordine* ch. 15). If so, then he locates the source of all Beauty within the human soul, and this inner Beauty could be one of God's traces in the world.

See also Aesthetic Judgment; Beauty; Number.

Bibliography

Beardsley, Monroe C. *Aesthetics from Classical Greece to the Present: A Short History*. New York: Macmillan, 1966.

Paul Thom (2005)

AUGUSTINIANISM

"Augustinianism" may be described as that complex of philosophical ideas that reflected to a greater or lesser degree the philosophy of Augustine. Many of the philosophers who came after Augustine not only restated his leading ideas but also frequently modified them with their own interpretations. Such interpretations were often the result of the impact of other schools of thought, notably the Avicennian and the Aristotelian. Occasionally doctrines that were only implicit in Augustine—for instance, the plurality of forms and universal hylomorphism—were made explicit and assumed considerable importance. Thus there originated in the medieval period what has been termed the Augustinian tradition, which in the later years of its development was closely identified with the Franciscan order. Such a tradition dominated medieval thought to the time of Thomas Aquinas. After Thomas it gradually disintegrated owing to the impact of Thomism and a resurgent Aristotelianism, and no longer represented a distinctive school or tradition. It continued, however, to be influential to the extent that it inspired or characterized in varying degrees later medieval and modern philosophers. The principal theses of Augustinianism will be discussed under seven headings.

FAITH AND UNDERSTANDING

The relationship between faith and understanding (or reason), with the implications of such a relationship for philosophy and theology, and the conception of Christian wisdom and Christian mysticism are central in the structure of Augustinian philosophy. One of the most influential and significant expressions of the relation between faith and understanding in Augustinian thought is summarized in the famous maxim of Anselm: *Credo ut intelligam* (I believe in order to understand). Peter Abelard similarly expressed the idea of the primacy of faith over understanding in his comments on the function of philosophy: "I do not want to be a philosopher if it is necessary to deny Paul. I do not want to be Aristotle if it is necessary to be separated from Christ. 'For there is no other name under heaven given to men, whereby we must be saved.'" With Roger Bacon the relationship of philosophy and theology is profoundly Augustinian. A conservative theologian despite his enthusiasm for scientific method and experimentation, he was convinced that the highest wisdom is found in Scripture and that philosophy exists only to explicate that wisdom. A similar theme is developed by Bonaventure in his *De Reductione Artium ad Theologiam.* He declared that all the sciences and philosophy should be subordinated to theology, which in turn must be subordinated to faith and the love of God; for faith alone enables man to avoid error and attain a union with God. Other philosophers of the Middle Ages who accepted this primacy of faith over reason and the complete subordination of philosophy to theology were Alexander of Hales, John of La Rochelle, Matthew of Aquasparta, and Roger Marston.

PSYCHOLOGY

The Augustinian psychology is characterized by the definition of man as a soul using a body and the implication of this definition for the relation of soul and body. The soul is regarded as an image of the Trinity and is said to

have a direct knowledge of itself. Hugh of St. Victor is notably Augustinian, not only in his mysticism but also in his identification of the soul with man and his belief that we have a direct knowledge of the soul and its spirituality. The union of soul and body he described as one of "apposition" rather than composition. Similarly, William of Auvergne is Augustinian in his account of man as a soul using the body, his affirmation of the presence of the soul in all parts of the body, and his statement that: "No knowledge is more natural to the soul than the knowledge of its own self." The mysticism of Bonaventure is characterized by the notion of the journey of the soul to God, the presence of the Trinity in the soul of man, and the direct knowledge the soul has of itself. This principle that the soul has a direct knowledge of itself is characteristic of both the Augustinian psychology and the Augustinian theory of knowledge. It has been termed the "principle of interiorization." Augustine expressed it: "For what is so present to knowledge as that which is present to mind? Or what is so present to the mind as the mind itself?" In modern philosophy the principle of interiority was to have significant influence upon writers like René Descartes, Blaise Pascal, Tommaso Campanella, and Maurice Blondel.

EPISTEMOLOGY

The Augustinian theory of knowledge had an extensive influence upon medieval philosophers, but it was frequently compromised with Aristotelianism. This was particularly true with respect to the Augustinian theory that sensation is essentially an act of the soul. However, the theory of the divine illumination, in conjunction with the doctrine of exemplary ideas, and the concept of truth as identified with God and present to, but superior to, all minds had a much stronger influence; but it, too, was often qualified with an Aristotelian theory of knowledge. Anselm held that truth is based on the Divine Ideas that are one with God. William of Auvergne accepted the doctrine of divine illumination but interpreted it as giving us an intuitive knowledge of the intelligible forms. Robert Grosseteste combined the Augustinian theory of the divine illumination with an empirical approach in science; he regarded truth as the conformity of a thing with its divine exemplar. Roger Bacon considered divine illumination as an inspiration, and he compared the divine action in illumination to that of the active intellect. Alexander of Hales combined the theory of divine illumination with an Aristotelian theory of abstraction. John of La Rochelle also combined the two theories of knowledge, especially the notion of the active intellect and the divine illumination. Bonaventure and Matthew of Aquasparta also modified the Augustinian theory of knowledge. The former accepted an Aristotelian account of sense knowledge and abstraction, of the existence of a possible and an active intellect, as well as the Augustinian concept of the necessity of the divine illumination for the attainment of truth. Matthew modified the Augustinian theory of sensation. On the other hand, Roger Marston and Peter Olivi followed closely Augustine's theory of knowledge. Among modern philosophers, the Augustinian doctrine of divine illumination was particularly influential with such philosophers as Nicolas Malebranche, Antonio Rosmini-Serbati, and Vincenzo Gioberti.

RATIONES SEMINALES

The conception of the *rationes seminales* (physical powers or "seeds") that Augustine postulated as potentially present in matter in order to explain the origin of creatures after the creation of the six days reappeared most markedly in the philosophical systems of the Augustinians of the thirteenth century.

HYLOMORPHISM AND PLURALITY OF FORMS

Hylomorphism and plurality of forms were doctrines that were developed from the thought of Augustine. The latter is said to have appeared first in Grosseteste's metaphysics of light and his analysis of bodies as possessing a number of different forms—for instance, the forms of elements, plants, animals. The highest form possessed by any body he held to be light, which was designated as the "form of corporeity." This notion of a plurality of forms was widely accepted by Augustinians after Grosseteste and is particularly prominent in the philosophies of Bonaventure, Raymond Lull, and John Duns Scotus. Generally it appears with its corollary universal hylomorphism, which states that all creatures are composed of matter and form. Thus angelic beings and human souls were said to be composed of a form and a spiritual matter. These doctrines enabled philosophers like Bonaventure and Duns Scotus to maintain more effectively their conception of the completeness of the substantial character of the human soul apart from the body. The Franciscan school strongly supported both doctrines. Robert Kilwardby and John Peckham in particular appealed to the plurality of forms in their vigorous opposition to the Thomistic doctrine of the oneness of man's substantial form.

THE MEANING OF HISTORY

Augustine rejected emphatically the cyclical conception of history as expressed in the Christian revelation and the doctrines of the Incarnation and salvation. History is a part of the divine plan and providence, and reflects the presence of the divine reason. The divine dispensation of grace gives hope to humankind and makes it possible for him to attain his eternal beatitude in the City of God after his pilgrimage in the earthly city. Few medieval philosophers escaped the influence of this Augustinian conception. It is particularly noticeable in the work of Dante Alighieri and in Roger Bacon's idea of a Christian republic. It influenced such later philosophers as Campanella, Jacques Bossuet, and Gottfried Wilhelm Leibniz. And it is indirectly represented in modern secularized versions of the idea of progress and social utopias.

ETHICS OF CHARITY AND SUPERIORITY OF THE WILL

The ethics of charity and the principle of the superiority of the will over the intellect in man as formulated by Augustine were important in the development of religious thought. The former, with its correlative doctrines of grace, election, and predestination, is essentially a religious ethic. It found universal acceptance within the Franciscan school and exerted considerable influence on all medieval theology and ethics. It affected such later thinkers as Martin Luther and John Calvin. The principle of the primacy of the will is reflected in Bonaventure insistence upon the need for moral as well as intellectual illumination. Richard of Middleton held that the will is a faculty that determines itself without being determined by any other faculty. Duns Scotus asserted that the will is free, whereas the intellect is determined by that which is known. The will is the nobler of the two faculties and commands the intellect.

See also Augustine, St.

Bibliography

Cary, Phillip. *Augustine's Invention of the Inner Self.* Oxford: Oxford University Press, 2000.

Cayré, F. "Augustinisme (développement de l')," *Note complémentaire, "Tables générales," Dictionnaire de théologie catholique,* Vol. 2, cols. 317–324. Paris, 1953.

Congrès international augustinien, *Augustinus Magister.* 3 vols. Paris, 1954.

Copleston, F. *History of Philosophy.* Vols. 2 and 3. London: Burns, Oates, and Washbourne, 1950.

Fitzgerald, Alan D. *Augustine through the Ages: An Encyclopedia.* Grand Rapids, MI: Eerdmans, 1999.

Gilson, E. *The Christian Philosophy of St. Augustine.* New York: Random House, 1960.

Gilson, E. *History of Christian Philosophy in the Middle Ages.* New York: Random House, 1955.

Kirwan, Christopher. *Augustine.* London: Routledge, 1989.

Marrou, Henri. *St. Augustine and His Influence through the Ages.* New York, 1958.

Matthews, Gareth B. *Augustine.* Oxford: Blackwell, 2005.

Matthews, Gareth B., ed. *The Augustinian Tradition.* Berkeley: California Press, 1999.

Matthews, Gareth B. *Thought's Ego in Augustine and Descartes.* Ithaca, NY: Cornell University Press, 1992.

Menn, Stephen. *Descartes and Augustine.* Cambridge, U.K.: Cambridge University Press, 1998.

Miscellanea agostiniana: testi e studii. Rome, 1930.

O'Daly, Gerard. *Augustine's Philosophy of Mind.* London: Duckworth, 1987.

Portalié, E. *A Guide to the Thought of St. Augustine.* Chicago: Regnery, 1960.

Rist, John M. *Augustine: Ancient Thought Baptized.* Cambridge, U.K.: Cambridge University Press, 1994.

Rondet, H. *Saint Augustine parmi nous.* Paris, 1954.

Rottmanner, O. *Der Augustinismus.* Munich, 1892.

Stump, Eleonore, and Norman Kretzmann, eds. *The Cambridge Companion to Augustine.* Cambridge, U.K.: Cambridge University Press, 2001.

Wetzel, James. *Augustine and the Limits of Virtue.* Cambridge, U.K.: Cambridge University Press, 1992.

John A. Mourant (1967)
Bibliography updated by Gareth Matthews (2005)

AUREOL, PETER

See *Peter Aureol*

AUSTIN, JOHN
(1790–1859)

John Austin, the most influential English legal philosopher of the analytical school, was born in London; at the age of sixteen he enlisted in the army and served five years, resigning his commission to study law. He was called to the bar in 1818. The following year he married Sarah Taylor, a woman of great intelligence and beauty, to whom many distinguished men of the age were deeply devoted.

The Austins became neighbors of Jeremy Bentham and the Millses and for twelve years remained closely associated with individuals in the Benthamite circle. The practice of law held little appeal for Austin, whose interests were primarily scholarly and theoretical; and after seven years he gave it up. In 1826, on the founding of the

University of London by the Benthamites with whom he had been closely associated for years, he was offered its chair in jurisprudence. He accepted with enthusiasm and immediately began to prepare himself by establishing his family in Bonn, where he taught himself German and studied the newly discovered *Institutes* of Gaius; the *Pandects*; and the works of Gustav Hugo, Anton Friedrich Justus Thibaut, and Friedrich Karl von Savigny. Some of the finest young minds in England—John Stuart Mill, George Cornewall Lewis, Sir John Romilly, and Sir William Erie among them—attended the first series of lectures at London. *The Province of Jurisprudence Determined,* published in 1832, is an expanded version of the first part of these lectures. Apart from this work, Austin published in his lifetime only two articles and a pamphlet attacking reform, *A Plea for the Constitution.* Austin, who once remarked, "I was born out of my time and place—I ought to have been a schoolman of the twelfth century—or a German professor," never again reached the high point of his first year at London. Student interest declined, and the chair, which had been supported by student fees, was given up by Austin in 1832 for financial reasons. His wife tells us that this was "the real and irremediable calamity of his life—the blow from which he never recovered." Plagued by illness and self-distrust, he served a brief and frustrating period, beginning in 1833, on the Criminal Law Commission; and later, with more satisfaction, he served as royal commissioner of Malta. During his remaining twenty years Austin spent some time on the Continent and a final period in Weybridge, not far from London, which proved to be the quietest and most contented part of his life. The second edition of the *The Province* was published in 1861, two years after his death. The first complete edition of *The Lectures on Jurisprudence or The Philosophy of Positive Law,* reconstructed from his notes by his wife, was published in 1863.

Both the nature and the results of Austin's inquiry deserve attention. What are the characteristics of his inquiry? First, his aim was to keep rigorously separate two questions that had formerly been confused, with much practical harm resulting: What is law? And what ought the law to be? Austin wished to lay a solid foundation for answering the second question by clarifying the first. His answer to the second question was along strictly utilitarian lines. Second, his inquiry was analytical rather than empirical. He was concerned with the analysis of concepts, not, for example, with historical or sociological questions. Finally, connected with the preceding analysis, he hoped to provide a general theory of law—"General jurisprudence"—whose concepts would permit us to grasp the essential features of any legal system without describing any particular system; this task of description was reserved for "particular jurisprudence."

What were the results of Austin's inquiry into the nature of law? The province of jurisprudence, the subject matter selected for study, is law "strictly so-called," or positive law, as contrasted, for example, with divine law (related to it by analogy) or physical laws of nature (related to it by metaphor). Positive law is a rule set for subjects by a sovereign in a politically independent society. A major part of *The Province* consists of analyses of the concepts in this explanatory definition. A rule is a species of command; it is a command that obliges the performance of a class of actions. A command is an expression or intimation of a wish that another do or forbear from doing some act, coupled with the ability and intention to inflict harm in case of noncompliance. The command concept, the key to the science of jurisprudence for Austin, encompasses the concept of a sanction (the evil that will probably be incurred in case of noncompliance), the concept of superiority (the power of forcing compliance with one's wishes), and the concept of obligation or duty (sometimes, for Austin, one is "obliged" because one fears the sanction, sometimes when one is "liable" to the sanction). A sovereign is that person or group of persons receiving habitual obedience from most members of a given society but not in turn having a like habit of obedience to a superior. An independent political society is one in which most members of the society have a habit of obedience to some person or group of persons who have no such habit of obedience to another.

Austin addressed his first class at London in these words: "Frankness is the highest compliment … I therefore entreat you, as the greatest favour you can do me, to demand explanation and ply me with objections—turn me inside out." Legal philosophers have paid him this compliment. His method and his results have come in for severe and often valid criticism. The inadequacies of Austin's theory result mainly from his selecting as basic tools of analysis the concepts of a command and habitual obedience. The former cannot account for certain commonly accepted features of law. It fails, first, to explain the varied content of laws, for if we view all law as an order or command backed by threats, we neglect those many laws that do not impose duties but, rather, function in a variety of ways. It also fails to account for the range of persons to whom laws are normally applicable, for orders are addressed to others, whereas most laws bind those who have enacted them as well as those who have not. Next,

orders are deliberate datable events; only with much stretching of meaning and introduction of fictions (the sovereign commands what he permits) can they account for the legal status of customary law and the decisions of the courts. Finally, the concept of a command leads Austin to the erroneous claim that one has a legal obligation because one fears the sanction.

The peculiar deficiency of a concept that links the law to habitual obedience is that serious difficulties are encountered in accounting for either the continuity of legal authority or the persistence of law. With the concept of habitual obedience alone, we should be unable to explain the common legal phenomena of one person's succeeding another in the authority to legislate or of laws that remain obligatory long after the legislator and those who habitually obeyed him are dead. Finally, focusing on coercion as the essence of law prevented Austin from developing sufficiently the connections that law has with morality, connections that make understandable one's moral obligation to obey the law.

In addition to these criticisms, Austin has been charged with lack of originality, even in his fundamental mistakes, for identical views may be found in Thomas Hobbes and Bentham. Bryce commented, "Bentham … drops plenty of good things as he goes along. Austin is barren." It is understandable that we should wonder at Austin's great influence, and his reputation as a great legal philosopher.

First, Austin's positivism, his insistence on separating questions of fact and value, has made legal philosophers sensitive to how easily these questions may be confused and how we may, as a result, delude ourselves into thinking we have answered one of these questions when we have, in fact, answered the other. Even more important, Austin's failures, all associated in some way with his imperativism, have been helpful. He was not alone in feeling the grip of a certain idea, the idea that law is simply the impressing of the will of the stronger upon the weaker. Austin's chief virtue was that he systematically developed, defended, and refined this idea, stripping it of excess philosophical baggage. In doing this he enabled us to focus with greater precision on those features of law that connect it with coercion. More than this, his model presses us to remark upon its limitations, the respects in which viewing law as coercion obscures its complicated role in our lives. After Austin, we understand better what there is in law that connects it with coercion and what there is in law that does not. This is his principal legacy. He provides one more instance in philosophy of our gaining something from a false statement that we might not have gained from a true one.

See also Bentham, Jeremy; Hobbes, Thomas; Legal Positivism; Mill, John Stuart; Philosophy of Law, Problems of; Savigny, Friedrich Karl von.

Bibliography

WORKS BY AUSTIN

Lectures on Jurisprudence, 5th ed., 2 vols. London: Murray, 1885.

The Province of Jurisprudence Determined. Introduction by H. L. A. Hart. London: Weidenfeld and Nicolson, 1954.

WORKS ON AUSTIN

Brown, Jethro. *The Austinian Theory of Law.* London: Murray, 1906.

Bryce, James. *Studies in History and Jurisprudence,* Vol. II. Oxford: Clarendon Press, 1901.

Hart, H. L. A. *The Concept of Law.* Oxford: Clarendon Press, 1961.

Maine, Henry. *Early History of Institutions.* London: Murray, 1875, Chs. XII and XIII.

Mill, J. S. *Dissertations and Discussions,* Vol. III. London: Longmans, 1867.

Morison, W. L. "Some Myths about Positivism." *Yale Law Journal* 68 (1958): 212–233.

Morris, Herbert. "Verbal Disputes and the Legal Philosophy of John Austin." *UCLA Law Review* 7 (1960): 27–56.

Herbert Morris (1967)
Bibliography updated by Philip Reed (2005)

AUSTIN, JOHN LANGSHAW
(1911–1960)

John Langshaw Austin was White's professor of moral philosophy at Oxford from 1952 until his death in 1960. Educated at Shrewsbury School and Balliol College, Oxford, he became a fellow of All Souls College in 1933; in 1935 he moved to Magdalen College, where he taught with conspicuous success until elected to the White's chair. During World War II he served with distinction in the British Intelligence Corps; he attained the rank of lieutenant-colonel and was awarded the OBE and the Croix de Guerre, as well as being made an officer of the Legion of Merit.

In the years before the war Austin devoted a great deal of his time and energy to philosophical scholarship. He made himself an expert in the philosophy of Gottfried Wilhelm Leibniz and also did much work on Greek philosophy, especially Aristotle's ethical works. At this period his own thought, although notably acute and already dis-

tinctive in style, was largely critical and altogether lacked the positive approach that distinguished his postwar work. His one published paper belonging to this early period, "Are There *A Priori* Concepts?" very fairly represents the astringent style and outlook that gave him the reputation of being a rather terrifying person. According to Austin's own statements, it was not until the beginning of the war that he began to develop the outlook on philosophy and method of philosophizing that marked his mature work, and it is of this work alone that an account will be given.

AIMS AND METHODS

The practical exigencies of lecturing and the traditions of paper reading (especially in symposia, to which some of his important papers were contributions) prevented some of the most characteristic features of Austin's preferred methods and aims from being clearly and fully exemplified in his written work. Lecturing is essentially a solo effort, whereas Austin believed that the best way of doing philosophy was in a group, and papers, especially in symposia, are almost inevitably on topics of traditional philosophical interest, whereas Austin preferred to keep the traditional problems of philosophy in the background. We shall therefore start by giving some account of the method and aims that Austin always advocated and practiced, most notably in meetings held regularly on Saturday mornings in the Oxford term with a group of like-minded Oxford philosophers.

LANGUAGE. Austin did not present his aims and methods as the only proper ones for a philosopher; whatever one or two uncautious remarks in his British Academy lecture "Ifs and Cans" may suggest to the contrary, he did not claim more than that his procedures led to definite results and were a necessary preliminary for anyone who wished to undertake other kinds of philosophical investigation. But he certainly considered them so valuable and interesting in their results, and so suited to his own linguistically trained capabilities and tastes, that he never felt it necessary to investigate for himself what else a philosopher might usefully do. What he conceived of as the central task, the careful elucidation of the forms and concepts of ordinary language (as opposed to the language of philosophers, not to that of poets, scientists, or preachers) was, as Austin himself was well aware, not new but characteristic of countless philosophers from Socrates to G. E. Moore. Nor were the grounds for this activity especially novel. First, he claimed, it was only common prudence for anyone embarking on any kind of philosophical investigation, even one that might eventually involve the creation of a special technical vocabulary, to begin with an examination of the resources of the terminology already at one's disposal; clarification of ordinary language was thus the "begin-all," if not the "end-all," of any philosophical investigation. Second, he thought that the institution of language was in itself of sufficient interest to make it worthy of the closest study. Third, he believed that in general a clear insight into the many subtle distinctions that are enshrined in ordinary language and have survived in a lengthy struggle for existence with competing distinctions could hardly fail to be also an insight into important distinctions to be observed in the world around us—distinctions of an interest unlikely to be shared by any we might think up on our own unaided initiative in our professional armchairs.

It is not too soon to remove at this stage some common misconceptions about Austin's aims and methods. First, although he was not concerned with studying the technical terminology of philosophers, he had no objection in principle to such terms; he thought that many such technical terms had been introduced inappropriately and uncritically, as is clear from his discussion, in *Sense and Sensibilia,* of the sense-datum terminology, but he used much of the traditional technical vocabulary of philosophy and added many technical terms of his own invention—as almost any page of *How to Do Things with Words* will bear witness. Second, Austin did not think that ordinary language was sacrosanct; he certainly thought it unlikely that hopelessly muddled uses of languages would survive very long and felt that they were more likely to occur in rather specialized and infrequently used areas of our vocabulary, but there was never any suggestion that language as we found it was incapable of improvement; all he asked was that we be clear about what it is like before we try to improve it.

TECHNIQUE. We have seen that there was nothing essentially novel in Austin's philosophical aims; what was new was the skill, the rigor, and the patience with which he pursued these aims. Here we are dealing with Austin's own personal gifts, which cannot be philosophically dissected. Nor did Austin have any theory of philosophical method; what he had was a systematic way of setting to work, something on a par with a laboratory technique rather than with a scientific methodology. This technique, unlike the skill with which he followed it, was quite public and one that he was willing and eager to employ in joint investigations with others, so we can easily give an account of it.

A philosopher or, preferably, a group of philosophers using this technique begins by choosing an area of discourse in which it is interested, often one germane to some great philosophical issue. The vocabulary of this area of discourse is then collected, first by thinking of and listing all the words belonging to it that one can—not just the most discussed words or those that at first sight seem most important—then by looking up synonyms and synonyms of synonyms in dictionaries, by reading the nonphilosophical literature of the field, and so on. Alongside the activity of collecting the vocabulary one notes expressions within which the vocabulary can legitimately occur and, still more important, expressions including the vocabulary that seem to be a priori plausible but that can nonetheless be recognized as unusable.

The next stage is to make up "stories" in which the legitimate words and phrases occur; in particular, one makes up stories in which it is clear that one can appropriately use one dictionary "synonym" but not another; such stories can also be found ready made in documents. In the light of these data one can then proceed to attempt to give some account of the meaning of the terms and their interrelationships that will explain the data. A particularly crucial point, which is a touchstone of success, is whether one's account of the matter will adequately explain why we cannot say the things that we have noted as "plausible" yet that in fact we would not say. At this stage, but not earlier, it becomes profitable to examine what other philosophers and grammarians have said about the same region of discourse. Throughout (and this is why Austin so much preferred to work in a group) the test to be employed of what can and what cannot be said is a reasonable consensus among the participants that this is so. Such a consensus, Austin found, could be obtained in an open-minded group most of the time; where such agreement cannot be obtained the fact should be noted as of possible significance. Austin regarded this method as empirical and scientific, one that could lead to definitely established results, but he admitted that "like most sciences, it is an art," and that a suitably fertile imagination was all important for success.

It was the lack of thoroughness, of sufficient research before generalization, in previous investigations of language, whether by those who called themselves grammarians or by those who called themselves philosophers, that Austin most deplored. He seriously hoped that a new science might emerge from the kind of investigations he undertook, a new kind of linguistics incorporating workers from both the existing linguistic and the philosophical fields. He pointed to other "new" sciences, such as logic and psychology, both formerly parts of philosophy, as analogues and was indifferent about whether what he was doing "was really philosophy."

So much must suffice as an account of the method of work that Austin advocated. It has been based on a set of notes for an informal talk, characteristically titled "Something about One Way of Possibly Doing One Part of Philosophy." As Austin admitted in those notes, he had said most of this in his papers "A Plea for Excuses" and "Ifs and Cans," and to all who worked with him it was familiar from his practice. Although inevitably, as we have noted, this method could not be followed in writings (it is in any case a method of discovery and not of presentation), its use underlies and can be discerned in his published work. Thus, before writing "Words and Deeds" or *How to Do Things with Words* he went right through the dictionary making a list, which still survives, of all verbs that might be classed as "performative" in his terminology. The art of telling "your story" is amusingly illustrated over and over again in his paper "Pretending" and, indeed, in all his other published writings. His insistence that it is a mistake to dwell only on a few well-examined notions in a field of discourse is illustrated by his concentration on such notions as "mistake," "accident," and "inadvertence" (in "A Plea for Excuses") and on the use of "I can if I choose" (in "Ifs and Cans"), rather than on "responsibility" and "freedom," in his papers that have a bearing on the free-will problem. Similarly, when his Saturday morning group turned its attention to aesthetics Austin betrayed far more interest in the notions of dainty and dumpy milk jugs than in that of a beautiful picture.

WORK

It is not possible to give a systematic account of Austin's "philosophy," for he had none. His technique lent itself rather to a set of quite independent inquiries, the conclusions of none of which could serve as premises for a further inquiry; his discussions of the language of perception (in *Sense and Sensibilia*), the concept of pretending, the notion of truth, and the terminology of excuses were all based on the study of speech in those fields and not on any general principles or theories. Nor would it serve any useful purpose to attempt to summarize his various investigations one by one, since they depend so much for their interest and force on the detailed observations about language that they contain. It will be more useful to discuss, first, what he thought of as his main constructive work—the doctrine of illocutionary forces that arose out of his earlier distinction of performative and constative utterances, contained in *How to*

Do Things with Words—and, second, the application of his technique to the criticism of some traditional theories about perception as found in his *Sense and Sensibilia.*

THEORY OF ILLOCUTIONARY FORCES. Austin's theory of illocutionary forces arose from his observation that a considerable number of utterances, even those in the indicative mood, were such that in at least some contexts it would be impossible to characterize them as being true or false. Examples are "I name this ship the Saucy Sue" (which is part of the christening of a ship, and not a statement about the christening of a ship), "I promise to meet you at two o'clock" (which is the making of a promise and not the report of a promise or a statement about what will happen), and "I guarantee these eggs to be new-laid" (which is the giving of a guarantee and not a report of a guarantee). These utterances Austin called "performative," to indicate that they are the performance of some act and not the report of its performance; he did not speak as some do who purport to discuss his views, of "performative *verbs,*" for the verb *promise* can well occur in reports—for example, "I promised to meet him." To provide the necessary contrast, Austin coined the technical term *constative* to apply to all those utterances that are naturally called true or false; he thought that *statement* and similar words often used by philosophers roughly as he used *constative* had in ordinary use too narrow a meaning to serve the purpose.

For a time Austin appears to have been fairly satisfied with this distinction, which he gave in print in his "Other Minds" article in 1946, using it to illuminate some features of utterances beginning "I know.…" But although the distinction is clearly useful at a certain level, Austin began to doubt whether it was ultimately satisfactory. He found it impossible to give satisfactory criteria for distinguishing the performative from other utterances. The first person of the present indicative, which occurs in the three examples given above, is clearly not a necessary feature; "Passengers are warned to cross the tracks only by the bridge" is an act of warning as much as "I warn you to cross.…" Further, in a suitable context "Don't cross the tracks except by the bridge" may also be an act of warning (as in another context it might be an act of commanding); this makes it necessary to distinguish the primative performative from the explicit performative, the latter, but not the former, making clear what act was being performed in its formulation.

Still more important, the constative seemed to collapse into the performative. Let us consider the four utterances "I warn you that a train is coming," "I guess that a train is coming," "I state that a train is coming," and "A train is coming." The first of these is an act of warning, the second is surely one of guessing, the third apparently one of stating, while the fourth may be any of these as determined by context. Thus, the various forms of constatives—stating, reporting, asserting, and the rest—seem to be merely a subgroup of performatives. It might seem that still one crucial difference remains, that while performative utterances may be in various ways unhappy (I may say "I promise to give you my watch" when I have not got a watch, or am speaking to an animal, or have no intention of handing the watch over), the characteristic and distinctive happiness or unhappiness of constatives is truth and falsehood, to which the other performatives are not liable.

In a brilliant, if not always immediately convincing, discussion (Lecture XI of *How to Do Things with Words*) Austin tried to break down even this distinction. First, we cannot contrast doing with saying, since (in addition to the trivial point that in stating one is performing the act of uttering words or the like) in constative utterances one is stating, describing, affirming, etc., and these acts are on a par with warning, promising, and so on. Second, all constatives are liable to all those kinds of infelicity that have been taken to be characteristic of performatives. Just as I should not promise to do something if I do not intend to do it, so I should not state that something is the case unless I believe it to be so; just as my act of selling an object is null and void if I do not possess it, so my act of stating that the king of France is bald is null and void if there is no king of France; just as I cannot order you to do something unless I am in a position to do so, so I cannot state what I am not in a position to state (I cannot *state,* though I can hazard a guess about, what you will do next year). Further, even if we grant that "true" and "false" are assessments specific to constatives, is not their truth and falsity closely parallel to the rightness and wrongness of estimates, the correctness and incorrectness of findings, and so on? Is the rightness of a verdict very different from the truth of a statement? Further, to speak of inferring *validly,* arguing *soundly,* or judging *fairly,* is to make an assessment belonging to the same class as truth and falsehood. Moreover, it is only a legend that "true" and "false" can always be appropriately predicated of constatives; "France is hexagonal" is a rough description of France, not a true or false one, and "Lord Raglan won the battle of Alma" (since Alma was a soldiers' battle in which Lord Raglan's orders were not properly transmitted) is exaggerated—it is pointless to ask whether it is true or false. It was on the basis of such considerations as these that

Austin felt himself obliged to abandon the distinction between the performative and the constative.

To replace the unsatisfactory distinction of performatives and constatives Austin introduced the theory of illocutionary forces. Whenever someone says anything he performs a number of distinguishable acts, for example, the *phonetic* act of making certain noises and the *phatic* act of uttering words in conformity with grammar. Austin went on to distinguish three other kinds of acts that we may perform when we say something: First, the *locutionary* act of using an utterance with a more or less definite sense and reference, for example, saying "The door is open" as an English sentence with reference to a particular door; second, the *illocutionary* act, which is the act I may perform in performing the locutionary act; third, the *perlocutionary* act, which is the act I may succeed in performing by means of my illocutionary act. Thus, in performing the locutionary act of saying that a door is open I may be performing an illocutionary act of stating, or hinting, or exclaiming; by performing the illocutionary act of hinting I may succeed in performing the perlocutionary act of getting you to shut it. In the same way, by performing the locutionary act of saying "Down with the monarchy" I may succeed in the perlocutionary act of bringing about a revolution, whereas in performing the locutionary act I would be inciting to revolution (successfully or unsuccessfully).

We now see that the constatives, along with performatives, can be construed as members of one particular subclass of illocutionary forces. Thus, in his provisional classification of illocutionary forces Austin had a subclass of *expositives,* which included the "constative" acts. In performing a locutionary act we may be affirming, denying, stating, describing, reporting, agreeing, testifying, rejoining, etc., but in performing a locutionary act we may also perform an act with *commissive* force, as when we promise, bet, vow, adopt, or consent; with *verdictive* force, as when we acquit, assess, or diagnose; with *exercitive* force, as when we appoint, demote, sentence, or veto; or with *behabitive* force, as when we apologize, thank, or curse.

Such is the crude outline of Austin's theory of illocutionary forces. Though his own exposition is of course much more full and rewarding, he said of it (*How to Do Things with Words,* p. 163): "I have purposely not embroiled the general theory with philosophical problems (some of which are complex enough almost to merit their celebrity); this should not be taken to mean that I am unaware of them." We may be permitted to illustrate the philosophical importance of bearing in mind the distinctions Austin made with one example of our own. Very often in recent years philosophers have set out to explain the meaning of the word *good* or of sentences containing the word *good.* Some of them have done so by saying that in such sentences the speaker expresses his own feelings (attitudes) and evokes similar feelings (attitudes in others). It might well seem that here they have set out to give an account relevant to locutionary force and that they have instead given one possible illocutionary force ("In saying that it was good I was expressing my favorable attitude toward it") and, alongside it, one possible perlocutionary force ("By saying that it was good I evoked in him a favorable attitude"). It should be clear in the light of Austin's work that such an account will not do. But Austin said very little about locutionary force in detail, and one of the most pressing general questions that arise from his work is that of the relationship between illocutionary force and locutionary force; while recognizing that they are different, and that locutionary force is in some way prior, can we, for example, conclude that the locutionary force of utterances containing the word *promise* can be explained without reference to the typical illocutionary force of "I promise"? This is far from clear.

CRITICISM OF TRADITIONAL PHILOSOPHY. We have examined in outline an example of Austin's work on a piece of clarification of language without any reference, save incidental, to the traditional problems of philosophy. We shall now turn to *Sense and Sensibilia,* which is emphatically a polemical discussion of one of the central problems of epistemology. But we shall find the essential features of Austin's method still present, the presentation only being different. Austin had recommended that when the method is used as one of inquiry the vocabulary and phrases, natural and odd, that occur to us should be studied and conclusions drawn *before* the conclusions of traditional philosophy are compared with them. Here, however, when he presents results he at each stage presents first the traditional philosophical theses and then shows their errors by confronting them with the actual facts, linguistic and otherwise.

In *Sense and Sensibilia,* Austin examines the doctrine that we never directly perceive material things but only sense data (or ideas, or sense contents, etc.), insofar as that doctrine is based upon the so-called argument from illusion. He maintains that it is largely based on an obsession with a few words "the uses of which are oversimplified, not really understood or carefully studied or correctly described" (*Sense and Sensibilia,* p. 3). With special reference to A. J. Ayer and Price, he shows how illusions are traditionally confused with delusions, are

defined in terms of belief that one sees a material thing when in fact one does not (whereas some illusions, such as one hatched line appearing to be longer than another of equal length, involve nothing of the sort), and are taken to include such phenomena as sticks looking bent in water, which are not illusions at all. A portion of the argument that clearly exhibits his method at work is where he contrasts the actual complexities and differences in our use of "looks," "appears," and "seems" with the traditional confusion of these terms in traditional philosophy. Especially interesting is the discussion of the traditional accounts of "reality"; these he contrasts with the multifarious uses of the word *real,* which takes its significance only from the implied contrast in context with *artificial, fake, bogus, toy, synthetic,* and so on, as well as with *illusory* and *apparent.*

But it is perhaps more important now for us to notice another element in the argument that is very characteristic but that we have as yet given little notice, which is Austin's care to avoid oversimplification and hasty generalization of nonlinguistic, as well as linguistic, fact. The ordinary man does not, as is so often stated or implied in accounts of the argument from illusion, believe that he always sees material things; he knows perfectly well that he sees shadows, mirror images, rainbows, and the like. The number of kinds of things that we see is large and to be settled by scientific investigation, not by philosophy; the question whether the invariable object of perception is a material thing or a sense datum is thus absurd. Again, it is not true that a straight stick in water normally looks like a bent stick out of water, for we can see the water; an afterimage does not look like a colored patch on a wall; a dream is distinguished by the dreamlike quality that occasionally, but only occasionally, we attribute to some waking experience. Again, he points out that situations in which our perception is queer may arise because of defects in sense organs or peculiarities of the medium or because we put a wrong construction on what we (quite normally) see, and it is a mistake to attempt to give a single account of all perceptual error. None of these are linguistic points, and Austin had no purist, theoretical notion that he was prohibited as a philosopher from any attention to nonconceptual issues; he thought that philosophical error did arise from empirical error.

Once again, it would be pointless to attempt to reconstruct the whole argument of *Sense and Sensibilia* here; we must be content with noticing the few points made that perhaps have some bearing on a general understanding of his general position. But it should perhaps be stressed that Austin in these lectures discussed only one theory of perception as based on one particular kind of argument; although one may expect to get help from it in study of other problems in the field of perception, it would be a mistake to suppose that the book contains a full study of all problems of perception or to criticize it because it leaves many difficult problems unanswered.

It is hardly imaginable that anyone would ever deny that Austin displayed a very great talent in the kind of work he chose to do. Some have criticized him on the ground that there are more important things for philosophers to do than this; on that point Austin always refused to argue, simply saying that those who preferred to work otherwise should do so and asking only that they not do what he did in the traditional slipshod way. To those who said that philosophers should work with an improved scientific language he replied flatly that the distinctions of ordinary language were of interest in their own right and that one should not modify what one does not fully understand, but he offered no theoretical objections to such projects. He was content to work in a way which he felt he understood and found rewarding. As for the assertion sometimes made, that Austin's kind of work is private to his own peculiar gifts and that it was therefore a mistake for him to recommend the method to others, time alone can decide.

A final word should be said about Austin's relation to other philosophers. He greatly admired G. E. Moore, but it is a mistake to view his work as an offshoot of Cambridge philosophy. Moore, like Austin and unlike most Cambridge philosophers, had a linguistic and classical background rather than a scientific one. Austin owed no special debt to Bertrand Russell and was far more unlike Wittgenstein than is sometimes recognized. For Ludwig Wittgenstein an understanding of ordinary language was important because he believed that the traditional problems of philosophy arose from misunderstandings of it, but Wittgenstein had in mind gross category mistakes, and he wished to study ordinary language only so far as was essential for eliminating these. Austin was interested in fine distinctions for their own sake and saw the application of his results to the traditional problems of philosophy as only a by-product. He was uninterested in the party conflicts of philosophy, following always his individual bent.

See also Aristotle; Ayer, Alfred Jules; Language; Leibniz, Gottfried Wilhelm; Moore, George Edward.

Bibliography

PRIMARY WORKS

For brevity, first publication of individual papers (collected in *Philosophical Papers*) is omitted.

Foundations of Arithmetic. Oxford: Blackwell, 1950. Translation of Frege's *Grundlagen der Arithmetik.*

Critical notice of J. Łukasiewicz's *Aristotle's Syllogistic: From the Standpoint of Modern Formal Logic Mind* 61 (1952).

"Performatif-constatif" and contributions to discussion in *La philosophie analytique.* Paris: Minuit, 1962.

How to Do Things with Words, 2nd ed. Edited by J. O. Urmson and Marina Sbisà. Cambridge: Harvard University Press, 1975.

Sense and Sensibilia. Edited by Geoffrey J. Warnock. London: Oxford University Press, 1975.

Philosophical Papers, 3rd ed. Edited by J. O. Urmson and Geoffrey J. Warnock. Oxford: Oxford University Press, 1979.

SECONDARY WORKS

Berlin, Isaiah, et al. *Essays on J. L. Austin.* Oxford: Clarendon Press, 1973.

Cavell, Stanley. *Philosophical Passages: Wittgenstein, Emerson, Austin, Derrida.* Bucknell Lectures in Literary Theory, 12. Oxford: Blackwell, 1995.

Di Giovanna, Joseph J. *Linguistic Phenomenology: Philosophical Method in J. L. Austin.* American University Studies. Series V, Philosophy, Vol. 63. New York: Lang, 1989.

Fann, K. T., ed. *Symposium on J. L. Austin.* New York: Humanities Press, 1969.

Furberg, Mats. *Saying and Meaning: A Main Theme in J. L. Austin's Philosophy.* Totowa, NJ: Rowman and Littlefield, 1971.

Graham, Keith. *J. L. Austin.* Atlantic Highlands, NJ: Humanities Press, 1977.

Holdcraft, David. *Words and Deeds.* Oxford: Oxford University Press, 1978.

Rorty, Richard, ed. *The Linguistic Turn: Recent Essays in Philosophical Method.* Chicago: University of Chicago Press, 1967.

Searle, John. *Expression and Meaning: Essays in the Theory of Speech Acts.* Cambridge, U.K.: Cambridge University Press, 1979.

Searle, John. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge, U.K.: Cambridge University Press, 1969.

Warnock, Geoffrey J. *J. L. Austin.* London: Routledge, 1989.

J. O. Urmson (1967)

Bibliography updated by Philip Reed (2005)

AUTHENTICITY REGARDING THE ARTIST AND THE ARTWORK

See *Art, Authenticity in*

AUTHORITY

Three topics have dominated philosophical discussion of authority: the *nature* of authority, the *point* of authority, and the *sources* of authority.

THE NATURE OF AUTHORITY

In providing an account of the nature of authority, the focus in this entry will be on de jure practical authority, though the differences between de facto and de jure authority and between theoretical and practical authority will also be briefly considered at the end of this section.

What is authority? Authority is a relational matter: A has authority over B with respect to some domain D. What follows is first a consideration of items A, B, and D that enter into this relationship, and then of the nature of the relationship among them.

With respect to A: while A need not be a person, it must be something that can have a say-so—that is, it must be the sort of thing about which one can truly assert "A says that B must ϕ." So, A can be a natural person, or a corporate person, or an institution, or a text. With respect to B: while B need not be a natural person, B must be the sort of thing that exhibits agency. B must, that is, be able to act on reasons: B must be capable of ϕ-ing because B has good reason to ϕ. In a relationship of practical authority, then, the authority-bearer must be a speaker, and the person under authority must be an agent.

With respect to D: Authority relationships are characteristically limited to a specific context and are characteristically limited by certain constraints within that context. An employer may have authority over an employee with respect to job-related matters, but may have no authority at all outside that context. And, further, it is not as if an employer has unlimited authority over an employee within the domain of job-related matters: if the employer told the employee to work until the employee dies of exhaustion, one would not take that to be within the range of the employer's authority. With each purported authority relationship there is assumed a domain for that authority, even though that domain is often poorly defined.

What, then, is the relationship between speaker A and agent B that makes for practical authority within a given domain D? Surely it is at least that the speaker has some control over the agent's reasons for action, and control of a specific kind: A's say-so makes a difference to the reasons that the agent B has by giving B a good reason to act a certain way. This good reason is, either in whole or in part, A's say-so—A's say-so produces a reason for

action for B not merely causally but constitutively. So if A has authority over B, then A's telling B to ϕ is itself a reason for B to ϕ. If, for example, parents have authority over their children with respect to household chores, then a parent's telling the child to clean his or her room is itself a reason for the child to clean the room. By telling the child to clean the room, the parent has added to the child's reasons to clean the room; aside from the fact that the current condition of the room may be aesthetically displeasing—and even a health risk to the child—one reason for the child to clean the room is that the parent told him or her to do so.

In de jure practical authority a speaker has constitutive control over an agent's reasons for action. But it is not enough for authority simply that the speaker's say-so constitutes a reason for action for the agent; the speaker's say-so must constitute a particular kind of reason. Reasons of authority play a certain role in the proper decision making of the agent under authority: where there is practical authority, and in the domain in which that authority is effective, the authority's say-so is decisive with respect to the agent's rational action. Authoritative dictates have the function of bringing deliberation to a close by fixing the action selected by the authoritative dictate as the reasonable choice to make.

One prominent way of expressing this idea has been offered by Joseph Raz, whose work has been the most important in explicating the nature and justification of authority. Raz says that the way that authoritative norms fulfill this function is by providing what he calls "protected reasons" (Raz 1979, p. 29). A protected reason to ϕ is both a reason to ϕ and a reason to disregard reasons not to ϕ. Raz claims that the way that authoritative norms fulfill their function of decisively terminating deliberation is not by providing enormously weighty reasons, reasons that compete with and always best any reasons that militate in favor of rival options; rather, authoritative dictates fulfill their function by giving reasons that insulate a course of action from competition. When an authority tells one to ϕ, that is a reason not only to ϕ but also to disregard in deliberation courses of action that preclude ϕ-ing. One might dispute Raz's claim that authoritative norms are always protected reasons, but the more fundamental point is that where there is authority, there is a speaker whose say-so the agent has reason to treat as setting the rationally preferable course of action within some domain.

In sum: if A is a genuine practical authority over B in some domain, then in that domain A's telling B to ϕ is a decisive reason for B to ϕ. Whereas in *practical* authority, speakers have authority over what agents do, in *theoretical* authority, speakers have authority over what agents believe. There are, nonetheless, striking similarities between genuine practical authority and genuine theoretical authority. If A has theoretical authority over B in some domain, then A must be a speaker and B must be one who can believe things for reasons; and if A tells B that it is the case that p, then A's telling B that it is the case that p is a reason to B to believe that p, a reason that is decisive from B's point of view. If an accomplished chef tells a novice that this is not the best way to make a roux, then the novice has decisive reason to believe that this is not the best way to make a roux. Practical authority, being concerned with reasons for action, is an object of investigation within the province of moral philosophy (and political and legal philosophy as well); theoretical authority, being concerned with reasons for belief, is an object of investigation within the province of epistemology.

One can also distinguish between de facto and de jure authority. People often ascribe authority (practical and theoretical) to speakers even without holding that their assertions or commands are reasons for believing or doing anything. Authority is sometimes used as a term of classification and explanation in both the social sciences and in everyday talk without any attempt to evaluate the claims of these putative authorities to give reasons for action. Authority in this sense is de facto, as opposed to de jure, authority. But there is nevertheless a tight connection between them: no speaker can be correctly described as a de facto authority without that speaker's either claiming to be or being widely regarded as a de jure authority.

Here is why. Suppose that one wants to argue that A is a de facto authority over some group simply because as a matter of observable behavior, if A tells members of that group to ϕ, then members of that group, by and large, ϕ. But that would surely be an insufficient basis for ascribing de facto authority; if it were a mere accident that the behavior of the group fell in line with the commands issued by A, one would not say that A bears de facto authority. One might try to complete the case for de facto authority by adding a causal condition: that it is A's commanding the members of that group to ϕ that results in those members' ϕ-ing. But this addition would be insufficient; if it were simply a quirky feature of the individual psychologies of the group's members that believing that A told them to ϕ caused them to ϕ, then one would ascribe a nervous disorder to the group members rather than de facto authority to A. The moral here is that if one wants to appeal to agents' responses to a speaker to establish

that the speaker is a de facto authority, one will have to argue that those agents act in accordance with A's say-so because they take A's say-so to be a reason for them to comply—that is, because they believe A to be a de jure authority.

One can make a similar argument from the side of the speaker: it is not sufficient to treat a speaker as a de facto authority that the speaker can (for example) have people locked up if they fail to obey. What makes the difference between effective kidnappers and de facto state authorities is, inter alia, that only the latter claim that their commands are binding standards, the violation of which justify locking people up—that is, that they are de jure practical authorities. Although there is no doubt a difference between de jure and de facto practical authority, de facto practical authority must be understood in terms of de jure practical authority.

The remainder of the entry will focus on genuine, de jure practical authority. Under what circumstances is it desirable for authority relationships of this sort of exist? And how are such authority relationships to be explained?

THE POINT OF AUTHORITY

What is the point of authority? Is there anything of value realized through such relationships?

Raz has written that the normal way of justifying authority is to show that those subject to it act better on their other reasons for action under authority than they would in the absence of authority. He calls this the "normal justification thesis." Practical authority provides a service—that is, the service of enabling persons to act more reasonably (Raz 1986).

There are several distinct contexts in which practical authority might provide this service. Practical authority might enable one to act more in accordance with what reason—prior to and apart from any authoritative imposition—determinately requires. It may be that a certain regimen of drug treatment is necessary for a person to regain her health. This is just a fact about the world, and given the value of this person's health, it would be unreasonable for her not to follow that regimen. But even if it is perfectly clear to her that this is the regimen to follow—because her doctor prescribes it, and her doctor's views are in line with the consensus of the medical community—she may fail to be motivated adequately by the doctor's theoretical authority alone. She might do better in acting in accordance with what reason requires if she had further reason to go along with the doctor's prescription—perhaps by placing herself under the doctor's authority. Some people do this sort of thing with personal trainers (or career mentors, or spiritual directors): they treat the trainer's (or mentor's or director's) prescriptions not as pieces of advice but as authoritative dictates, and they better reach their health- (or career-, or sanctification-) related goals by having trainers (or mentors or directors) that are not merely dispensers of advice but practical authorities.

Another context in which practical authorities can help persons to act on their other reasons for action is that in which their reasons for action require actions that are, to a significant degree, vague or otherwise indeterminate. The most pressing of such cases are those in which persons need to act in a coordinated way. The standard example is the rule of the road: While there is strong reason for persons to drive on the same side of the road, it is indeterminate which side they should drive on. Although a solution may be reached through trial and error, it would be helpful if there were a party that could set the rule of the road in a clear and determinate way prior to the disasters that can occur on the way to a convention established by trial and error.

There are several ways in which such indeterminacy presents itself. In some cases, there are a number of instrumental means to a single well-defined goal—for example, keeping people from driving into each other—and practical authority's job is to select one such means as a common plan of action. In some cases, there are a number of ways to fill in a vague rule. For example: one should not drive an automobile if one is intoxicated. But: What counts as "intoxicated"? Is it to be fixed by actual level of impairment? By blood alcohol level? And, in either case, at what levels? This is a matter that can be resolved by authoritative imposition: a practical authority can set what counts as intoxication, thus helping persons act in a coordinated way both with respect to their driving behavior and with respect to the claims that they make on one another with respect to their driving behavior. For the rule about drunk driving matters not only when one is deciding whether to drive; it is also important both in deciding whether to make claims on another for the damages that the other has done while driving with alcohol in his or her system and in deciding whether to accept claims made on one by others for damages that one has done while driving with alcohol in one's system.

Raz's normal justification thesis brings out the points that practical authority calls for justification (and thus can fail with respect to this call) and that the usual way that practical authority is justified is by showing that

our other reasons are better served by adding reasons of authority into the mix. Whereas this may indeed be the normal way of establishing the point of authority, it is not the only way. Here is another: Practical authority is not just an ability to change others' reasons for action, it is also a (typically) positive status. Because practical authority is a (typically) positive status, placing someone in a position of practical authority can be a way to honor that person, or a way to give that person what he or she deserves. So one justification for practical authority might make reference not to the ways in which those subject to authority are served by it but by reference to the way that the bearer of authority is appropriately honored or rewarded by it. There may, for example, be in some group no particularly pressing reason to institute structures of authority for decision making, but in light of the need to honor a person who has made great contributions to that group's aims, it would make sense to confer authority on that person.

There are also more tedious reasons for being in an authority relationship. Employers often have limited authority over their employees, and employees enter those authority relationships by contract. From the employer's point of view, the salient reason for the authority relationship is to bind the employee to performance of duties; from the employee's point of view, the salient reason for the authority relationship is that, unless he or she is willing to enter it, he or she will have no job, and no paycheck. This is a far cry from authority's helping an agent to act on his or her preexisting reasons or from authority's being conferred on someone in order to do him or her honor, but it cannot be denied that a number of more-or-less limited authority relationships in which people find themselves are justified in this way.

It is important to note, though, that practical authority has its drawbacks. Recall that what distinguishes the reasons of practical authority is their decisive role in deliberation. If reasons of authority are absent, then something else will have to fill the role that brings deliberation to its conclusion in these cases—often, the agent's own free, rationally underdetermined decision. If one takes the making of free, underdetermined decisions to be a good, then there is something lost by persons who are under practical authority (see Wolff 1970). So it is not as if practical authority is costless. And, furthermore, it should be noted that there may be bad psychological tendencies associated with certain sorts of otherwise worthwhile authority relationships, at least in certain classes of people. Even if practical authority is, properly circumscribed, necessary and valuable, there may be broad types of person who tend to act worse when placed either in such positions of authority or under such authority (see Milgram 1974).

THE SOURCES OF AUTHORITY

Suppose that person X claims to have practical authority over person Y, and Y is rightly curious about the correctness of this claim. When Y challenges X, X's response is: "You are under my authority because I am I, and you are you. You are under my authority because I am X, and you are not." X's case is poor; not only is it unconvincing, but it borders on incoherent. It borders on incoherence because authority relationships are normative matters, and whether a normative fact obtains depends not on irreducibly particular facts but on general ones. So just as claims about one's duties and rights are correct not in virtue of the particular identity of the person to whom those duties and rights are ascribed but in virtue of the general properties instantiated by that person, claims about who holds authority depend on the general properties instantiated by that person. It might fundamentally matter with respect to the presence of authority whether one is a parent, or is morally good, or is powerful, or has a loud voice; it cannot fundamentally matter whether one is Bill Clinton, Bob Dylan, or Bozo the Clown.

How, then, does it come to be the case that one individual is a practical authority over another individual, given that it is not simply as that individual that one is an authority bearer or a person under authority? There are two ways to try to answer this question. The first is to begin with general practical principles, and to show that in certain circumstances those practical principles imply that one party has authority over another. It is crucial that one select a practical principle that stands a chance of generating the crucial features of authority, that is, that at least under some possible set of circumstances it implies that one party's say-so is in some domain a decisive reason for action for another party.

Here is one principle that has been invoked in a number of contexts to explain practical authority: the principle of promising. The principle of promising, stated loosely, is that if one promises to perform an act of ϕ-ing, then one is morally bound to ϕ. One's valid promise is, in standard cases, a reason for the promisor to perform the action promised, and it is a reason of a certain kind: that one validly promised to perform an action is characteristically decisive, again, at least in standard cases. But one can promise to act in accordance with another party's commands: one can promise to obey the personal trainer's commands with respect to one's exercise regi-

men, or to follow the policies of one's immediate superior in the performance of job duties, or to do the bidding of the king. If one has validly promised to act on another person's commands in some domain, then it seems that there is good reason to suppose that authority has been generated.

This way of accounting for practical authority can be called the "top-down" approach. It begins with practical principles of broader application and proceeds to show that in some circumstances the application of that general principle yields a relationship that bears all of the defining conditions of genuine authority. A rival approach, the "bottom-up" approach, takes authority relationships as basic, not to be explained as implications of other practical principles. The guiding idea of this way of proceeding is that relationships of practical authority are no less familiar than—and perhaps may be more familiar than—the general practical principles that users of the top-down approach have employed. The task for the user of the bottom-up approach is simply to describe the general features of the various relationships of genuine practical authority with which we are familiar, and so far as possible to exhibit the unity among them (either the precise features they share, or analogies between them). So, the bottom-up theorist might note (for example) that people commonsensically accept that parents have authority over children, and thus take his or her task to be to define more precisely what counts as the parent/child relationship in which authority exists (is it biological? social? legal? some combination of these?) and what defines the scope of the parent's practical authority (is it over all domestic matters? does it extend beyond that? does the size of its domain remain constant, or not?). He or she may wish to answer similar questions about the authority of the state and of God, and to draw the appropriate connections and disanalogies between these cases of authority.

There are potential drawbacks to both of these approaches. The bottom-up approach seems to be extremely deferential to de facto authorities, offering them the presumption of de jure status. As such, employment of the bottom-up approach is unable to ease the suspicions of the authority skeptic, who is concerned either in particular cases or in a more global way about the existence of de jure practical authority. The top-down approach, by contrast, runs the real risk of failing to hook up with de facto authority relationships in any straightforward way. If there is a clear lesson to be drawn from the history of uses of the top-down approach in investigating questions of practical authority—most attention has been paid to parental, political, and divine authority—it is that it is extraordinarily difficult to generate plausible accounts of the de jure authority of common social institutions from standard applications of widely held general practical principles.

PARENTAL AUTHORITY It is commonly thought, especially among parents, that parents are practical authorities over their children. But employing the top-down approach to explain parents' status as authorities over their children has proven to be a difficult undertaking.

A preliminary difficulty for this undertaking is specifying the window in which this authority is supposed to obtain: at what age does a parent's authority over children begin (babies are not under authority, as they cannot yet act for reasons), and at what age does it cease? Suppose, though, that the time frame in which parents are to be authoritative over children is settled; how is one to explain why parents are authoritative during this stretch of time? A child's requirement of obedience to his or her parents cannot be a matter of voluntary undertaking, as Hobbes erroneously (and even inconsistently with his own view) supposed it to be (Hobbes 1651, ch. 20). So no principles of moral obligation founded on consent or voluntary acceptance of benefits will do. One might appeal to the requirements of gratitude, but this suggestion is rife with difficulties: Why does gratitude, which typically does not require obedience, generate such a requirement in this case? If it is a parent's duty to care for children, why is gratitude owed as a consequence of a mere doing of one's duty? Isn't gratitude characteristically conditioned on free acceptance of benefits, whereas children are typically not free to refuse such benefits from their parents or to seek the benefits elsewhere?

Locke argues, plausibly, that a parent's authority over children is due to the child's deficiencies in reason and choice, and it is only so long as those deficiencies remain that the parent has authority, for it is as a help to remedying those deficiencies that the parent has authority (Locke 1690, §55). These claims seem to be true, but this argument concerns the point, or value, of parental authority, not the explanation of how parents come to be authoritative over children. These are distinct questions. Even those who grant the value of parental authority can find its existence and explanation mysterious indeed. Little progress has been made in providing a top-down account of the authority of parents over their children.

POLITICAL AUTHORITY Political authority—its desirability, its scope, and its explanation—is one of the few

truly perennial problems of political philosophy. The most familiar explanation for political authority is the consent account. On this view, most fully elaborated during the early modern period, citizens characteristically make an agreement, either explicitly or tacitly (Hobbes 1651, ch. 18; Locke 1690, §119; Rousseau 1762, bk. 4, ch. 2), to obey their rulers. It is in virtue of this original agreement that subjects are duty-bound to obey their political authorities. Whereas this view has been tremendously popular, it is subject to overwhelming objections, which are catalogued in David Hume's influential "Of the Original Contract" (1753). Hume convincingly argues that there is little evidence that explicit agreements have generally taken place, and that the tacit agreements that consent theorists find themselves forced to posit to explain state authority are in fact no more than a myth.

Hume's recommended account of political authority is a utilitarian one—that is, that it is simply in virtue of the public benefit brought about by having a political authority in place that the de jure authority of the state is established. As Hume puts it, public authority is necessary for the public good, and public authority cannot be sustained unless subjects pay "exact obedience" to their rulers. But this is unpersuasive: states maintain de facto authority in the face of quite a bit of disobedience, and so Hume has not explained why the need for de facto authority yields the conclusion that states have de jure authority over their subjects.

During the latter half of the twentieth century there was a revival of attempts to provide an account of state authority. Some were attempts to retrieve the old consent view, offering new accounts of the tacit consent that was necessary to bind that vast majority of persons who never explicitly consent. Some were attempts to revive Hume's utilitarian-style argument. H. L. A. Hart (1955) and John Rawls (1964) offered arguments from fairness, holding that the authority of law is based on the fact that those who accept the benefits of legally ordered cooperation would be unfairly free riding on the efforts of others were they not to obey as well. Others appealed to an argument that appears as early as Plato's *Crito*—the idea that citizens owe a debt of gratitude to their political authorities for the goods that they receive through them, and this debt is to be repaid through obedience.

Despite the ingenuity of writers attempting to provide top-down accounts of political authority, the most important writers on political authority at the end of the twentieth century were "philosophical anarchists"—they held, that is, that none of these attempts to account for genuine, de jure political authority is successful, and thus people have reason to reject the view that modern states are genuinely authoritative. Some of these writers, such as Robert Paul Wolff (1970), hold that this is a necessary truth: there cannot be a genuinely authoritative state. But most of them—most prominently, A. John Simmons (1979), Joseph Raz (1979), and Leslie Green (1990)—argue simply that under current political conditions no state holds the wide-ranging authority that it claims for itself.

DIVINE AUTHORITY Recent scholarship employing the top-down approach has had difficulty exhibiting the sources of authority of the two likeliest bearers of wide-ranging authority: parents and political institutions. One might think that even if such human institutions were bound to fall short in this respect, surely divine authority would be an easier matter. After all, whereas God's existence remains a philosophically controverted matter, it is widely accepted by both theists and nontheists that if there is such a being as God then that being is practically authoritative over human beings.

It turns out that accounting for divine authority using a top-down approach raises difficulties that are just as pressing as the difficulties that attend accounting for parental or political authority in the top-down way. One might think that being a practical authority is a logical consequence of traditional divine attributes, such as omniscience, omnipotence, or perfect moral goodness, but this is wrong: omniscience and perfect moral goodness give us reasons only to think of God as a theoretical authority, not a practical authority; and whereas omnipotence of course enables God to control the circumstances in which our reasons for action have application, it does not of itself entail that God's commands constitute reasons for action for humans. One might think that traditional moral principles concerning gratitude for benefits or property in what one has created would yield a moral obligation to obey God, but it turns out that there are severe difficulties in demonstrating that the conditions of application for these principles generate such an obligation of obedience (Murphy 2002). The top-down approach has fared no better in the case of divine authority than in the cases of parental and political authority.

It is unclear what moral should be drawn from the failure of top-down philosophical investigation to generate plausible accounts of parental, political, and divine authority. On the one hand, one might take it simply to be an indication that, given the nature of authority, it is bound to be hard to show that one party has genuine authority over another in nonstylized contexts—that is,

contexts that are not created for the purpose of generating authority relationships (e.g., employer-employee contracts). On the other hand, one might take it to be a reductio ad absurdum of reliance on the top-down approach. One might claim that among people's sturdiest considered moral judgments are the judgments that these authority relationships are genuine. To the extent that distinct general moral principles fail to illuminate the authority present there, this failure gives people reason not to jettison the view that these authority relationships are genuine but only to insist on a bottom-up approach, taking practical authority as a basic feature of the moral world.

See also Civil Disobedience; Cosmopolitanism; Postcolonialism; Republicanism.

Bibliography

For considerations of the nature and justification of authority, see Joseph Raz, *The Authority of Law* (Oxford, U.K.: Oxford University Press, 1979) and *The Morality of Freedom* (Oxford, U.K.: Oxford University Press, 1986). Several important discussions are collected in *Authority*, edited by Joseph Raz (Oxford, U.K.: Blackwell, 1990). For philosophical skepticism concerning the possibility of de jure authority, see Robert Paul Wolff, *In Defense of Anarchism* (New York: Harper & Row, 1970). For some discussion of the deleterious effects of submission to authority, see Stanley Milgram, *Obedience to Authority: An Experimental View* (New York: Harper & Row, 1974).

There is a vast literature on the sources of de jure authority, almost all of it focused on political authority. The classic contract account may be found in Hobbes, *Leviathan*, edited by Edwin Curley (Indianapolis, IN: Hackett, 1994; first published 1651), John Locke, *Second Treatise of Government* (1690), reprinted in *Two Treatises of Government*, edited by Peter Laslett (Cambridge, U.K.; Cambridge University Press, 1988), and Jean-Jacques Rousseau, *On the Social Contract*, in The *Basic Political Writings*, translated by Donald Cress (Indianapolis, IN: Hackett, 1987; first published 1762). David Hume's skeptical treatment of contract views may be found in his "Of the Original Contract" (1753), reprinted in *Political Essays*, and edited by Knud Haakonssen (Cambridge, U.K.: Cambridge University Press, 1994). Accounts justifying political authority in terms of fair distribution of the benefits and burdens of rule following are to be found in H. L. A. Hart, "Are There Any Natural Rights?" (*Philosophical Review* 64 (1955): 175–191), and John Rawls, "Legal Obligation and the Duty of Fair Play," in *Law and Philosophy*, edited by Sidney Hook (New York: New York University Press, 1964). A gratitude view is suggested in Plato's *Crito*, in *The Trial and Death of Socrates*, translated by G. M. A. Grube (Indianapolis, IN: Hackett, 1975). Recent accounts skeptical of all extant theories of political authority may be found in A. John Simmons, *Moral Principles and Political Obligations* (Princeton, NJ: Princeton University Press, 1979), Raz's *The Authority of Law* and Leslie Green, *The Authority of the State* (Oxford, U.K.: Oxford University Press, 1990). There has been less recent work on parental and divine authority. For a discussion of parental authority, see Michael Slote, "Obedience and Illusions," in *Having Children*, edited by Onora O'Neill and William Ruddick (New York: Oxford University Press, 1979); for a discussion of divine authority, see Mark C. Murphy, *An Essay on Divine Authority* (Ithaca, NY: Cornell University Press, 2002).

Mark C. Murphy (2005)

AVEMPACE

See *Ibn Bājja*

AVENARIUS, RICHARD

(1843–1896)

Richard Avenarius, the German positivist philosopher, was born in Paris. He studied at the University of Leipzig, where he became a *Privatdozent* in philosophy in 1876. The following year he was appointed professor of philosophy at Zürich, where he taught until his death. His most influential work was the two-volume *Kritik der reinen Erfahrung* (1888–1890), which won him such followers as Joseph Petzoldt and such opponents as Vladimir Il'ich Lenin.

Avenarius was the founder of empiriocriticism, an epistemological theory according to which the task of philosophy is to develop a "natural concept of the world" based on "pure experience." To obtain such a coherent, consistent view of the world requires a positivistic restriction to that which is directly given by pure perception, together with the elimination of all metaphysical ingredients which man, through introjection, imports into experience in the act of knowing.

There is a close kinship between the ideas of Avenarius and those of Ernst Mach, especially as set forth in Mach's *Analyse der Empfindungen.* The two men never became personally acquainted, and they developed their points of view quite independently of one another; hence, it was only gradually that they became convinced of the profound agreement of their basic conceptions. They held the same fundamental view on the relationship between physical and mental phenomena, as well as on the significance of the principle of the "economy of thought." Above all, both were persuaded that pure experience must be recognized as the sole admissible—and thoroughly adequate—source of knowledge. Thus, the elimination of introjection by Avenarius is only a special

form of that total elimination of the metaphysical which Mach sought.

In addition to Petzoldt and Lenin, others who dealt at length with the philosophy of Avenarius were Wilhelm Schuppe and Wilhelm Wundt. While Schuppe, the philosopher of immanence, agreed with Avenarius on essential points, Wundt criticized the scholastic character of Avenarius's expositions and sought to point out internal contradictions in his doctrines.

COGNITION

The two presuppositions of empiriocriticism are the empiriocritical axiom of the contents of cognition and the axiom of the forms of cognition. The first axiom states that the cognitive contents of all philosophical views of the world are merely modifications of the original assumption that every human being initially assumes himself to be confronted with an environment and with other human beings who make assertions and are dependent on the environment. The second axiom holds that scientific knowledge does not possess any forms and means essentially different from those of prescientific knowledge and that all the forms and means of knowledge in the special sciences are extensions of the prescientific (*Kritik der reinen Erfahrung,* Vol. I, Preface).

Especially characteristic of Avenarius' theory of human cognition was his biological approach. From this biological point of view, every process of knowledge is to be interpreted as a vital function, and only as such can it be understood. Avenarius' interest was directed chiefly to the pervasive relations of dependency between individuals and their surroundings, and he described these relations in an original terminology involving many symbols.

The point of departure for his investigations was the "natural" assumption of a "principal coordination" between self and environment, in consequence of which each individual finds himself facing both an environment with various component parts and other individuals who make assertions about this environment which also express a "finding." The initial principal coordination thus consists in the existence of a "central term" (the individual) and "opposite terms" about which he makes assertions. The encountering individual is represented and centralized in system *C* (the central nervous system, the cerebrum), the basic biological processes of which are nourishment and work.

System *C* is exposed to change in two ways; changes in it are dependent on two "partial-systematic factors": variations in the environment (*R*) or stimuli from the external world (whatever can, as a stimulus, excite a nerve), and fluctuations in metabolism (*S*), or absorption of food (whatever in the environment of system *C* conditions and constitutes its metabolism). System *C* constantly strives for a vital maximum conservation of its strength (*V*), a state of rest in which the mutually opposed processes $f(R)$ and $f(S)$—that is, the variations of system *C* as functions of *R* and *S*—cancel each other out, and the two variations maintain an equilibrium ($f(R) + f(S) = 0$, or $\Sigma f(R) + \Sigma f(S) = 0$). If $f(R) + f(S) > 0$, then there arises in the state of rest or equilibrium state of system *C* a disturbance, a relationship of tension, "a vital difference." The system strives to diminish or cancel out and equalize this disturbance by passing over spontaneously to secondary reactions in order to reestablish its original state (the conservation maximum, or *V*). These secondary reactions to deviations from *V* or to physiological fluctuations in system *C* are the so-called independent vital sequences (the vital functions in system *C*, the physiological processes in the brain), which run their course in three phases: the initial segment (appearance of the vital difference), the middle segment, and the final segment (reappearance of the earlier state). The canceling out of a vital difference is possible, of course, only in the manner and to the extent that system *C* exhibits a readiness for it. Among the changes preparatory to achieving readiness are hereditary dispositions, developmental factors, pathological variations, practice or exercise, and the like. The "dependent vital sequences" (experiences, or *E*-values) are functionally conditioned by the independent vital sequences. The dependent vital sequences, which, like the independent, proceed in three stages (pressure, work, release), are the conscious processes and cognitions ("assertions about contents"). For example, an instance of knowledge is present if in the initial segment the characterization reads "unknown" and in the final segment it reads "known."

Avenarius sought to explain the rise and disappearance of problems in general as follows. A disparity can arise between the stimulation from the environment and the energy at the disposal of the individual either (*a*) because the stimulation is strengthened as a result of the individual's having found anomalies, exceptions, or contradictions in the given, or (*b*) because an excess of energy is present. In the first case, problems arise that can, under favorable circumstances, be solved by knowledge; in the second case, practical-idealist goals arise. The latter are the positing of ideals and values (for example, ethical or aesthetic ideals and values), the testing of them (that is, the forming of new ones), and through them the alteration of the given.

The *E*-values, which depend on the fluctuations in the energy of system *C,* fall into two classes. The first are "elements," or simple contents of assertions—contents of sensation, such as green, hot, and sour, which depend on the objects of sensation or stimuli (whereby the "things" of experience are understood as nothing more than "complexes of elements"). The second are "characters," the subjective reactions to sensations, or the feelinglike modes of apprehension. Three groups of basic characters (kinds of awareness) are distinguished: the "affective," the "adaptive," and the "prevailing." Among the affective characters are the feelings proper (the "affectional," pleasure and aversion) and the feelings in a figurative sense (the "coaffectional," such as anxiety and relief, and the "virtual," such as feelings of movement). The adaptive characters include the "identical" (sameness or "tautote," difference or "heterote"); that is, the "fidential," the "existential" (being, appearance, nonbeing), the "secural" (certainty, uncertainty), and the "notal" (the being known, the being unknown), together with many modifications of these. For example, modifications of the "idential" include, among others, generality, law, whole, and part.

PURE EXPERIENCE AND THE WORLD

Avenarius constructed the concept of pure experience and related it to his theory of the natural concept of the world on the basis of his views on the biology and psychology of knowledge. The ideal of a natural concept of the world of pure experience is fulfilled in the complete elimination of metaphysical categories and of dualistic interpretations of reality, by means of his exclusion of introjection. The basic prerequisite for this is first to acknowledge the fundamental equivalence of everything that is encountered and that can be grasped, regardless of whether it is given through external or internal experience. As a consequence of the empiriocritical principal coordination between self and environment, individuals and environment are encountered in the same fashion, without distinction. "With respect to givenness, I and the environment are on completely the same footing. I come to know the environment in exactly the same sense that I come to know myself—as members of a single experience; and in every experience that is realized the two experience-values, the self and the environment, are in principle coordinated to each other and equivalent" (*Der menschliche Weltbegriff*).

Likewise, the difference between *R*-values and *E*-values is conditional upon the mode of apprehension. Both values are equally accessible to description. They differ only in that the former are interpreted as constituents of the environment, while the latter are conceived of as the content of an assertion of another human individual. In the same way, there is no ontological distinction between the mental and the physical; rather, there is a logical functional relation between them. A process is mental insofar as it is dependent on a change in system *C* and has more than mechanical significance, that is, insofar as it signifies an experience. Psychology has no separate subject matter at its disposal; it is nothing other than the study of experience insofar as experience is dependent on system *C.* Avenarius rejected the usual interpretation of and distinction between mind and body. He recognized neither the mental nor the physical but only a single kind of being.

ECONOMY OF THOUGHT

Of particular importance for the realization of the cognitive ideal of pure experience and for the notion of the natural concept of the world is the principle of the economy of thought. In the same way that thinking in conformity with the principle of least exertion is the root of the theoretical process of abstraction, so knowledge generally orients itself by the degree of exertion required to fulfill experience. Hence, one should exclude all elements of the mental image that are not contained in the given, in order to think about that which is encountered in experience with the least possible expenditure of energy, and thus to arrive at a pure experience. Experience, "cleansed of all adulterating additions," contains nothing but constituents of experience that presuppose constituents of the environment only. Whatever is not pure experience, and thus is not the content of an assertion (an *E*-value) subject to the environment itself, is to be eliminated. What we term "experience" (or "existing things") stands in a certain relationship of dependence to system *C* and to the environment; and experience is pure when it is cleansed of all those contents of assertions that do not depend on the environment.

A world concept relates to the "sum total of the constituents of the environment" and is dependent on the final character of the *C*-system. It is natural if it avoids the error of introjection and is not falsified by animistic "insertions." Introjection transfers the perceptual object into the perceiving person. It splits our natural world into inner and outer, subject and object, mind and matter. This is the origin of metaphysical problems (like immortality and the mind-body problem) and metaphysical categories (like substance). All of these must therefore be eliminated. Introjection, with its unwarranted duplication of reality, must be replaced by the empiriocritical

principal coordination and the natural concept of the world that rests on it. Thus, at the end of its development the world concept returns to that natural form with which it began: a purely descriptive comprehension of the world, with the least expenditure of energy.

See also Cognitive Science; Experience; Lenin, Vladimir Il'ich; Mach, Ernst; Petzoldt, Joseph; Schuppe, Ernst Julius Wilhelm; Wundt, Wilhelm.

Bibliography

WORKS BY AVENARIUS

Philosophie als Denken der Welt gemäss dem Prinzip des kleinsten Kraftmasses. Prolegomena zu einer Kritik der reinen Erfahrung ("Philosophy as Thinking of the World in Accordance with the Principle of the Least Amount of Energy. Prolegomena to a Critique of Pure Experience"). Leipzig: R. Reisland, 1876.

Kritik der reinen Erfahrung ("Critique of Pure Experience"), 2 vols. Leipzig: R. Reisland, 1888–1890.

Der menschliche Weltbegriff ("The Human Concept of the World"). Leipzig: R. Reisland, 1891.

WORKS ON AVENARIUS

Ewald, Oskar. *Richard Avenarius als Begründer des Empiriokritizimus.* Berlin, 1905.

Lenin, V. I. *Materializm i Empiriokritizism.* Moscow, 1909. Translated as *Materialism and Empirio-criticism.* New York: International Publishers, 1927.

Raab, Friedrich. *Die Philosophie von Richard Avenarius.* Leipzig: Meiner, 1912.

Schuppe, Wilhelm. "Die Bestätigung des naiven Realismus." *Vierteljahrsschrift für wissenschaftliche Philosophie* 17 (1893): 364–388.

Suter, Jules. *Die Philosophie von Richard Avenarius.* Zürich, 1910.

Wundt, Wilhelm. "Über naiven und kritischen Realismus." *Philosophische Studien* 12 (1896): 307–408, and 13 (1897): 1–105 and 323–433.

Franz Austeda (1967)
Translated by Albert E. Blumberg

AVERROES

(c. 1126–c. 1198)

Averroes, or ibn Rushd, was the foremost figure in Islamic philosophy's period of highest development (700–1200). His preeminence is due to his own immense philosophical acuity and power and to his enormous influence in certain phases of Latin thought from 1200 to 1650.

Averroes ("ibn Rushd" is a more exact transliteration of the Arabic, while "Averroes" is the medieval Latin version) was born in Córdoba into a family of prominent judges and lawyers; his grandfather, bearing the same name, served as the chief *qāḍī* (judge) of Córdoba, and there is a tradition that his father carried out the same duties. (In Muslim society a *qāḍī*'s professional concepts and practical duties were simultaneously civil and religious. Thus, a "lawyer" had expert knowledge of divine law.)

There are, however, few other specific details about his life and career. Ernest Renan and Salomon Munk mention that he studied under the most learned teachers in theology and law (in the Muslim world the two disciplines are effectively the same). It has been suggested that he studied with such scientists and philosophers as ibn Ṭufayl (d. 1185) and ibn Bajja (or Avempace, d. 1138), but the tenuous evidence would indicate that he became acquainted with the former only when he was past forty and that the death of the latter occurred when Averroes was only eleven or twelve years of age. Thus, significant pedagogical influence by these personalities upon Averroes is doubtful.

There remain, nevertheless, scattered pieces of evidence and suggestions of dates delineating his career. Averroes himself mentions that he was in Marrakech in 1153, on which occasion he observed the star Canope, not visible in Spain at that time. This sighting confirmed for him the truth of Aristotle's claim that the world was round. Some years later he seems to have been associated with the family of the Ibn Zuhr, traditionally physicians and scholars of medicine. He is reported to have been well acquainted with Abū Marwān ibn-Zuhr, perhaps the most outstanding member of the family, and when Averroes composed his medical handbook titled *Kulliyat* (literally, "generalities," which became latinized to *Colliget*), he encouraged Abū Marwān to write a companion text concerned with the details of specific ailments.

Tradition next reports that Averroes came into the favor of the sultan of Marrakech, a notable patron of scholarship and research, through the personal recommendation of his friend and presumed mentor, ibn Ṭufayl. His ready intelligence seems to have pleased the *cālīf,* who, according to a student of Averroes, subsequently encouraged the vast series of commentaries on Aristotle that became known in the West around 1200. It is generally conjectured that the association among ibn Ṭufayl, the *cālīf,* and Averroes can be dated between 1153 and 1169.

Through the *cālīf's* offices, Averroes was appointed *qāḍī* of Seville in 1169, and he began his array of commentaries on Aristotle about that time. In 1171 he returned to Córdoba, probably as *qāḍī,* and eventually

became chief *qādī.* He was, however, continually traveling to Seville and to Marrakech, as the colophons of various of his writings attest. In 1182 he became physician to the *cālīf* of Marrakech, continuing as a court favorite until about 1195. At that time he is supposed to have retired, possibly under a cloud as the result of religious controversy, or perhaps to be protected from conservative theologians, to a village outside Seville; details are not available. In any case, he soon returned to Marrakech, where he died.

His death coincided with the virtual disappearance of the dynamic speculative tradition evidenced in Arabic thinking for the several centuries after 700. Interestingly, it also coincided with the bursting forth of a similarly active tradition in the Latin West, which was greatly stimulated by the translations of Aristotle and Greek science from Arabic and Hebrew manuscripts. All these events—the death of Averroes, the abrupt decline of Arab intellectual dynamism, the translation into Latin of Aristotle (notably the *Metaphysics* and *De Anima* about 1200), and the exponential acceleration of Western philosophizing—occurred virtually within two decades. These are perhaps neither radically causative nor dependent events, but their close association is historically remarkable.

WRITINGS

During the course of his active professional life as *qādī,* physician, scientist, and philosopher, Averroes found time to compose an impressive number of scientific, philosophical, and religious writings. It is possible that some of his appointments may have been, in part, preferments for the purpose of sustaining scholarship. Certainly in the medieval Latin West, many a Sorbonne scholar formally designated "canon of Rheims," for example, could rarely be found at Rheims fulfilling his canonic responsibilities.

Most of Averroes's writings that can be dated fall between 1159 and 1195. There is the medical encyclopedia *Kulliyat* (composed before 1162), along with expositions of and commentaries on such medical writers as the Greek Galen and the Eastern Islamic ibn Sīnā (normally latinized as Avicenna). There are writings on astronomy. In religious philosophy there is the famous reply to the philosopher Muhammad al-Ghazālī's attack on the pretensions of rationalism in matters of divine law (*The Incoherence of the Philosophers*); Averroes's response is titled *The Incoherence of the Incoherence,* in which he strongly affirms the solid adequacy of natural reason in all domains of intellectual investigation. There are many lesser writings, on problems of divine law, on logic, on natural philosophy, and on medicine. Finally, there is the massive set of commentaries on the Aristotelian corpus, which profoundly affected medieval Latin thought—sometimes with official ecclesiastical approbation, sometimes not.

COMMENTARIES ON ARISTOTLE. The commentaries on Aristotle are of three kinds: short, often called paraphrases or epitomes; intermediate; and long, usually meticulous and detailed explications. These different versions may well correspond to stages in the educational curriculum.

The commentaries survive in many forms. For some writings of Aristotle, all three commentaries are available, for some two, and for some only one. Since Aristotle's *Politics* was not accessible to him, Averroes wrote a commentary on Plato's *Republic,* under the assumption that Greek thought constituted a coherent philosophical whole. He believed that the *Republic* contributed to this total philosophical construction. In still a further attempt to complete the presumed integrity of all Greek natural philosophy, Averroes supplemented Aristotle's *Physics* and *De Caelo* with a treatise of his own titled *De Substantia Orbis.*

In supplementing Aristotle in this fashion, Averroes did violence to the original methodology of the Stagirite. For Aristotle the *Physics* and *De Caelo* investigated motions and processes according to two different perspectives—*Physics,* motion as such; *De Caelo,* motion in the particular context of the activities of the heavenly bodies. These investigations were not conceived as standing in any hierarchical order, reflecting any vertical order of being or reality; they were simply different investigations and must not be taken, as did many ancient and medieval commentators, in terms of category and subcategory. Averroes, with methodological dispositions akin to the Platonic, did take them in this way, and thus eventually he found it necessary to provide an all-comprehensive celestial physics—hence, the *De Substantia Orbis.*

TEXTUAL TRADITION. The actual textual tradition of Averroes's works is extremely complex. Some of the commentaries remain in Arabic versions, some in Hebrew translations from the Arabic, some in Arabic texts recorded in Hebrew script, and many in Latin translations. These categories are not mutually exclusive. Beginning in 1472 there appeared numerous printed editions of some, but by no means all, of the commentaries; the format usually consists of a paragraph of Aristotelian text followed immediately by Averroes's comments on and interpretation of that text. This was no doubt an appara-

tus designed for the practical needs of the teaching of natural philosophy in the Western Latin universities, for it is clear that Averroes's analyses had become influential by the first quarter of the thirteenth century, accompanying as they did the translations of Aristotle, and they remained influential in the traditions of the universities well into the seventeenth century.

AVERROES'S PHILOSOPHY

Averroes's own philosophical position can best be characterized as Aristotle warped onto a Platonic frame. He inherited Greek thought as a literary corpus and, like his Islamic philosophical predecessors, viewed this corpus as an intellectually integrated totality. Aristotle, his commentators (such as Alexander of Aphrodisias and Simplicius) and such thinkers as Plotinus and Proclus were all understood as parts dovetailing into a single coherent philosophical system. Al-Fārābī (d. c. 950) is an eminent example of this syncretism: he composed a work titled *The Harmony between Plato and Aristotle,* and Averroes himself, lacking Aristotle's *Politics,* found little difficulty in incorporating Plato's *Republic* within his compass of speculation.

RELIANCE ON NEOPLATONISM. The doctrinal positions of Greek and Alexandrian thinkers were, in fact, often quite divergent and even incompatible, and to complete the final union of their philosophies into a single intellectual system the Arab philosophers made use of a writing called the *Theology.* Late ancient tradition attributed this treatise to Aristotle, but modern scholarship has established that the *Theology* is fundamentally a compendium based on Plotinus's writings. This work was taken uncritically by Arabic philosophers as the capstone of all Greek speculative thought and, as such, was employed by them to effect the unity of ancient philosophy.

"Mystical" knowledge. There were at least two reasons for the eager Islamic approval of the *Theology.* First, it strongly reflected the Neoplatonic emphasis especially evident in Plotinus' *Enneads,* on the culminating "mystical" experience at the apex of human knowledge. This experience involved a passing from a condition of ordinary logical ratiocination over into a condition of nondiscursive (although quasi-rational) grasp of ultimate reality. Such an attitude is strongly sympathetic to the Islamic conception of ultimate religious experience, in which there is an analogous passing from individuality into an impersonal fusion with a Whole or Divine Essence.

Hierarchy of reality. Correlative to its reflection of Neoplatonic "mystical" knowledge, the *Theology* reflected the Neoplatonic methodological conception that is ordered in an organic hierarchy, with interlocking levels indicating superordinate and subordinate dependency. Such relationships involve levels of being and, concomitantly, sources and receivers of being. Such an intellectual structure might be visualized as a series of pyramids successively superimposed, with the preeminent pyramid pointing to an ultimate One that simultaneously comprehends being as such and is the culmination of human reflective experience. This structure is, moreover, dynamic and not static, with a continuing flow of creativity downward and a continuing activity of noetic discovery upward.

ANALYSIS OF THE SOUL. The general methodology described above is evident in many specific places in Averroes's philosophy. In his analysis of the soul, for example, Aristotle's original doctrine undergoes a transformation. Whereas Aristotle's insistence on the physical principle that every form separate from matter is one in species leads to a presumption against the possibility of individual immortality, Averroes takes the obverse: Separate forms or substances can subsist in the general hierarchy of being, and thus immortality, in a purely impersonal sense, is possible.

SCIENTIFIC KNOWLEDGE. The case in natural science is similar to that of the soul. In Aristotle the various sciences are diverse and not necessarily reducible to one another in any formal sense: the *Physics* views natural behavior from one perspective and in accordance with one set of working principles, while the *De Caelo,* in contrast, uses another perspective and another set of principles. Aristotle's natural sciences are irrefragably diversified. In the *Metaphysics* he goes so far as to say that similar terminology is employed in the several sciences; however, this apparent unity of the sciences is qualified by his insistence that the use of the most general metaphysical language is, in disparate domains, only analogous and not semantically equivalent. The particular subject matter that a science encompasses controls the precise significance of the terms and logic used in the analysis and description of that science; the term "being" as it is used in the *Physics* does not possess the same meaning as "being" used in *De Anima.*

For Averroes, however, such differentiations among the sciences were not the case. "Being" had a univocal significance, not equivocal, as it had for Aristotle; and Averroes viewed nature and reality as exhibiting a single

coordinated and coherent structure, proceeding in orderly hierarchical fashion from levels that are lesser (both metaphysically and noetically) to greater and richer levels of being. Aristotle's horizontal and discrete conglomeration of sciences became a harmonious order of vertically structured science with dependent and causative relationships.

ACTIVE AND PASSIVE INTELLECTS. From Aristotle, Averroes understood that the knowing process in man comprised a passive aspect—adumbrant concepts capable of being fully activated—and an active aspect—a power of dynamically activating such concepts. This power, termed during the medieval period the "active intellect," was taken to operate against a "passive intellect" to actualize concepts and thus constituted the thinking activity; and the resulting fusion of function was termed the "acquired intellect." This terminology applicable to the noetic process was based on Aristotle's *De Anima,* and appears, with minor variations, in Greek and Arabic thought down to the time of Averroes. God, as the First Intelligence, provides through the next subordinate level of intelligences—the celestial bodies, upon which he exercises immediate control—activating power for the active intellect controlling man's thought.

The active intellect is not personalized, however, because it is Aristotelian form, and each such form is a species and never an individual. Nor is the passive intellect, in its nonnoetic status apart from participation in the acquired intellect—a further pressing of Aristotle impelled by Platonic dispositions. In Averroes's philosophy, consonant with Muslim theology, it is thus a domain of reality that looks upward to God for its sustaining power and with which individual souls strive to fuse impersonally, in knowledge and ultimately in immortality. Thus Averroes, and certainly his medieval interpreters, believed in the unlikelihood of individual immortality—the active intellect with which man hopes to unite at death being a single undifferentiated form—and the soul, as individuated in this life, cannot subsist without the body.

METAPHYSICS, NATURAL PHILOSOPHY, SCIENCE. Averroes's metaphysics, natural philosophy, and science can be classified as a moderate Platonism, tempered with a profound appreciation of Aristotle. Unlike many of his Islamic predecessors, Averroes accepted Aristotle's rigorous rationalism wholeheartedly, although at various crucial points his renderings of Aristotle's laconic texts are governed by his own Platonic methodological predispositions. Against the latter, he held the principle of the univocality of being, flowing downward from a Supreme Principle. God's existence is established from the *Physics,* in that the eternity of motion demands an unmoved mover, which is in itself pure form. In addition to being the source of motion, such pure form is also Intelligence as such, operating not only as the source of the celestial bodies and all subordinate motions but also as the creative originator and sustaining force behind all lesser intelligences.

THEOLOGY AND NATURAL PHILOSOPHY. In the Christian intellectual environment of the thirteenth century, apparent conflicts between argumentation in natural philosophy and argumentation in matters of theological doctrine became exceptionally acute. The newly introduced writings from the ancients—Greek philosophy and science, accompanied by Arabic and Hebrew commentary—rigorously set forth propositions alien to fundamental dicta of Christian faith: for example, the eternity of the world, the impossibility of individual immortality, and the radical noncontingency of existence as such. Averroes's rendering of the Aristotelian writings contributed heavily to these conflicts. Aristotle was read in the medieval faculties of arts as the staple of natural philosophy and science, and Averroes was read as his primary interpretive adjunct. In fact, in later medieval writings Averroes is merely referred to as "the Commentator." Thus, since he put forward analyses understanding Aristotle to deny the creation of the world in time, personal immortality, and the contingency of existence, such views attained wide currency among masters of arts.

The response from the theological side was early and direct. "Arabic" commentary was forbidden to be read in 1210 and 1215, and permitted only with censoring in 1231, at the University of Paris. Albert the Great published a treatise, *Contra Averroistas,* and Thomas Aquinas wrote about 1269, at a time of great intellectual controversy at Paris, a *Tractatus de Unitate Intellectus Contra Averroistas.*

"Double-truth" doctrine. The replies to Averroes were reasoned and moderate, but they seem to have been accompanied by many contemporary declarations that the "Averroists" were actually maintaining a doctrine of "double truth," according to which conclusions in natural philosophy were said to be true, while simultaneously conclusions affirming the contrary in theological argument were held true—presumably an intolerable intellectual situation. Thus there were official condemnations of "unorthodox" doctrines at the University of Paris in 1270 and 1277, including specific injunctions against two stan-

dards of truth. It is not, however, clear that any philosophers in the thirteenth century explicitly held such a theory of "double truth"; in the writings that survive, philosophers faced with these conflicts take great pains to concede truth itself to the declarations of faith and say of Aristotelian writings only that they have been properly arrived at according to Aristotle's methods.

Averroes himself composed the short treatise *On the Harmony between Religion and Philosophy*; his main effort in this work was to establish that there is but one truth to which there are several modes of access—the rhetorical, open to any man through the persuasions of teachers; the dialectical, available for some to explore the probability of truths of divine law; and the philosophical, to be used only by those few capable of exercising pure ratiocination with the fullest competence. Such a variety of methods ensures for each man, depending on his individual capability, the possibility of grasping ultimate realities. The fact that in this work Averroes distinguishes between such modes of access to truth has, by many historians, been taken to adumbrate the theory of the "double truth," as attributed to many thinkers in the thirteenth century, but this is not probable. First, this work of Averroes was not available to medieval Latin scholars and thus obviously cannot have been directly influential; second, the doctrine of alternative modes of access to truth is hardly the same as that of maintaining incompatible truths in disparate domains.

Thus, the attribution of a doctrine of "double truth" to medievals cannot be sustained by any writings of Aristotle accompanied by Averroistic commentaries, nor can it be justified explicitly from any Christian medieval master. The oppositions between Aristotelian-Averroist argument and basic Christian doctrine constituted a fundamental intellectual dilemma within Christian speculation—one never resolved by the masters of arts in an explicit proclamation of a logical contradiction between two domains of reflection but always by an absolute accession of truth to faith. Averroes did not contribute specifically to the discussion arising from this dilemma, except insofar as his rigorous analysis of Aristotle made necessary certain conclusions in natural philosophy.

Averroes stands as a philosopher in his own right, but his influence was felt essentially in Western Latin philosophy from 1200 to 1650. His commentaries on Aristotle, an integral part of the educational curriculum in the faculties of arts of western European universities, shaped several centuries of Latin philosophy and science. Despite institutional criticism and even formal condemnation, his powerful statements of Aristotelian doctrine were sustained among Latin scholars and thinkers well into the mid-seventeenth century.

See also Albert the Great; Aristotelianism; Aristotle; Averroism; Averroism in Modern Islamic Philosophy; Ibn Bājja Ibn Ṭufayl; Jewish Averroism; Neoplatonism; Plato; Platonism and the Platonic Tradition; Plotinus; Thomas Aquinas, St.

Bibliography

The most important general references are Ernest Renan, *Averroès et l'averroisme* (Paris: A. Durand, 1852; modern ed., Paris, 1949); Salomon Munk, *Mélanges de philosophie juive et arabe* (Paris: A. Franck, 1859), pp. 418–458; Léon Gauthier, *Ibn Rochd* (Paris: Presses universitaires de France, 1948); and G. Quadri, *La philosophie arabe dans l'Europe médiévale* (Paris: Payot, 1947), pp. 198–340. The last two of these studies depend heavily on the first two, which are the unsuperseded (except in occasional detail) classics in the literature on Averroes, although Gauthier properly views some of the indirect traditions with caution. For Averroes's predecessors, mentors, and contemporaries, see George Sarton, *Introduction to the History of Science,* Vol. I (Baltimore: Carnegie Institution of Washington, Williams and Wilkins, 1927) and Vol. II (Baltimore: Carnegie Institution of Washington, Williams and Wilkins, 1931), *passim.* Significant recent interpretations, with varying emphases, can be found in Étienne Gilson, *History of Christian Philosophy in the Middle Ages* (New York: Random House, 1955); A. A. Maurer, *Medieval Philosophy* (New York: Random House, 1962); and D. Knowles, *Evolution of Medieval Thought* (Baltimore: Helicon Press, 1962).

For a detailed catalog of Averroes's writings, see George Sarton, *Introduction to the History of Science,* Vol. II, Part 2, pp. 356–360. Also see Léon Gauthier, *Ibn Rochd,* pp. 12–16, and M. Bouyges, *Notes sur les philosophes arabes connus de Latins au moyen âge,* Vol. IV, *Inventaires des textes arabes d'Averroès* (Beirut, 1922). The latter, a monograph, is in the *Mélanges de l'Université Saint-Joseph* (Beirut: Catholique, 1922), Vol. VIII, Fascicle 1. H. A. Wolfson has meticulously stated the ambitious program for preparing and publishing modern editions of the Aristotelian commentaries in "Plan for the Publication of a Corpus Commentariorum Averrois in Aristotelem," in *Speculum* 6 (1931): 412–427, and "Revised Plan for the Publication of a Corpus Commentariorum Averrois in Aristotelem," in *Speculum* 38 (1963): 88–104. The latter article provides the most reliable listing of the surviving writings. There are other modern editions and translations of some works: for instance, E. I. J. Rosenthal, *Averroës' Commentary on Plato's Republic* (Cambridge, U.K.: Cambridge University Press, 1966); G. F. Hourani, *Averroës on the Harmony of Religion and Philosophy* (London: Luzac, 1961); and Simon Van den Bergh's translation of *The Incoherence of the Incoherence* (London: Luzac, 1954).

Stuart MacClintock (1967)

AVERROES [ADDENDUM]

At the time that Ibn Rushd was working, the philosophical curriculum was largely Neoplatonic, and this is because the Greek tradition of philosophy was transmitted to the Islamic world via the Neoplatonic tradition. But some thinkers like Ibn Rushd were perceptive enough to realize that there were discrepancies between Aristotle—very much his hero—and the Neoplatonists, who were represented at the time by the thought of Ibn Sīnā. In his defense of philosophy in the *Tahāfut al-Tahāfut*, for instance, Ibn Rushd not only seeks to refute al-Ghazālī's attack on philosophy, but he also tries to argue with Ibn Sīnā's particular Neoplatonic philosophy. In fact, he manages to link al-Ghazālī, the critic of philosophy, with Ibn Sīnā, its main representative. Al-Ghazālī argues that causality is nothing more than the way in which people interpret God's bringing things into existence, and therefore they should not think of causal connections as being necessary. Ibn Sīnā does identify necessity and causality, but also, for him, something has to actualize essences. Ibn Rushd criticizes both of these views; he argued that existence is linked with essence—that is, what is meant by cotton is something that bursts into flames when it is touched by fire (other things being equal). The properties of cotton are not just an incidental feature of the cotton; they are an essential aspect of it.

The views of Ibn Rushd came to have a radicalizing influence on European thought when they were translated into Hebrew and Latin. They were often taken to imply that philosophical and religious truths could be in opposition to each other, and yet still both be true. This is not what Ibn Rushd himself argued; he was too good of a thinker to believe that contradictory propositions could both be true. However, he did argue that there are different routes to the truth—routes that are appropriate to different audiences. For those capable of understanding rigorous logical arguments there is philosophy, and for those disinclined or unable to appreciate such arguments there are argument forms of lesser rigor. Ibn Rushd sees the syllogism as being the basis of all uses of language. Thus while the philosopher employs the demonstrative syllogism, the politician will use rhetoric and sophistry, the prophet sometimes uses poetry, and the theologian dialectic. All of these are reasoning processes, but only demonstration—according to Ibn Rushd—reaches the highest standards of reason.

That does not mean, however, that there is anything wrong with the other methods of reasoning; they simply are not so secure as demonstration. The other methods may, nonetheless, be able to express what philosophers can discover through demonstration in ways that are accessible to more people. Because God made everyone different, Ibn Rushd believes it is appropriate that God make everyone capable of understanding some method of argument—although not everyone should be expected to employ the same method.

It is worth pointing to the radical nature of this doctrine. For one thing, Ibn Rushd's doctrine suggests that the philosophers as a group are the best able to understand the language of any text, even difficult scriptural passages. After all, philosophers can operate at the level of demonstration and so are skilled in working with the highest levels of reason. The theologians and lawyers are only used to dialectic, in which they start with propositions that are generally accepted as true, but might not be. Ibn Rushd disparages their efforts as compared with those of the philosophers. Ordinary people are in an even worse position. On the one hand they have to rely on language and on arguments that rely on imagery and persuasion, and thus they are a long way from demonstrative rigor. On the other hand, what they believe is perfectly valid because there is a demonstrative proof for it, but not a proof they themselves can grasp. They do not believe anything false, but they do not appreciate the entire basis of their beliefs. Ibn Rushd gives the analogy of going to a physician or a lawyer with a problem. He suggests that if a person had the expertise of the lawyer and the physician, there would be no need to consult them (even though when lawyer and physician are consulted, what they suggest may not be understood by client or patient). If the advice were understood, there would be no need for the doctor or lawyer in the first place. And yet, there is nothing wrong with people's reliance upon doctors and lawyers because it is assumed that they understand why they make the suggestions they do—and thus if people are wise they will accept and follow those suggestions.

This approach is not a doctrine of double truth, but it is a radical doctrine that relegates religion to a relatively lowly role in the hierarchy of human pursuits. Religion is certainly inferior to reason as a way of finding out truth, because religious language is to be understood primarily by examining it philosophically. One of the features of Ibn Rushd's thought that differentiates him from other Islamic philosophers is his supreme indifference to mysticism. Mysticism, or *taṣawwuf*, was of overriding significance for most of his contemporaries and predecessors—and indeed successors—but not for him. For Ibn Rushd, the meaning of the world is firmly in the world, and not something behind it. In this way he sought to

establish a purified form of Aristotelianism shorn as far as possible from its Neoplatonic accessories and excesses.

It is not surprising that Averroism came to be regarded as a challenging doctrine in the Middle Ages and beyond, and it may well have played a role in displacing traditional religion from its established role in intellectual and social life. Within the Islamic world, Ibn Rushd's views largely disappeared until the Islamic Renaissance, when they reemerged to argue for a division between religious and rational language. In modern times, Averroism has once again been used in the Arab world to argue for a new and enhanced respect for reason as compared with religion. It still appears to be a philosophy for the intellectual elite rather than the religious masses.

See also Averroism; Averroism in Modern Islamic Philosophy; Jewish Averroism.

Bibliography

Some of Ibn Rushd's works now only exist in Hebrew or Latin, and some not at all. The most useful bibliography is "Ibn Rushd: A Catalogue of Editions and Scholarly Writings from 1821 onwards," edited by Phillip Rosemann, *Bulletin de philosophie mediévale* 30 (1988): 153–215.

WORKS BY AVERROES

"Long Commentary on Aristotle's Metaphysics." In *Ibn Rushd's Metaphysics*, edited by Charles Genequand. Leiden, Netherlands: Brill, 1984.

"Tahāfut al-Tahāfut" (The incoherence of the incoherence). In *Averroes' Tahafut al-Tahafut*, edited and translated by S. Van den Bergh. London: Luzac, 1978.

WORKS ABOUT AVERROES

Fakhry, Majid. *Islamic Occasionalism and Its Critique by Averroes and Aquinas*. London: Allen and Unwin, 1958.

Hourani, G., tr. *Averroes on the Harmony of Religion and Philosophy*. London: Luzac, 1976.

Kogan, Barry. *Averroes and the Metaphysics of Causation*. Albany: State University of New York Press, 1985.

Leaman, Oliver. "Averroes." In *Klassiker der Religionsphilosophie*, edited by F. Niewöhner. Munich: C.H. Beck, 1995.

Leaman, Oliver. *Averroes and His Philosophy*. Richmond, Surrey, U.K.: Curzon, 1998.

Leaman, Oliver. "Averroes and the West." In *Averroes and the Enlightenment*, edited by M. Wahba and M. Abousenna. New York: Prometheus, 1996.

Leaman, Oliver. *A Brief Introduction to Islamic Philosophy*. Malden, MA: Blackwell, 2000. Contains material on his philosophical arguments.

Leaman, Oliver. *An Introduction to Classical Islamic Philosophy*. Cambridge, U.K.: Cambridge University Press, 2001. Contains material on his philosophical arguments.

Leaman, Oliver. "Was Averroes an Averroist?" In *Averroismus im Mittelalter und in der Renaissance*, edited by F. Niewöhner and L. Sturlese. Zurich: Spur Verlag, 1994.

Urvoy, Dominique. "Ibn Rushd." In *History of Islamic Philosophy*, edited by S. Nasr and O. Leaman. London: Routledge, 1996.

Urvoy, Dominique. *Ibn Rushd (Averroes)*. Translated by O. Stewart. London: Routledge, 1991.

Oliver Leaman (2005)

AVERROISM

As a designation applicable to a tradition or mode of philosophizing, "Averroism" cannot be used in any account of Arabic thought after the death of Averroes (c. 1198). After that, in a most unusual intellectual situation, Averroes's influence is to be found not in Muslim thought but in Western Latin philosophy between 1200 and 1650, for the dynamic speculative activity vital for five centuries in the Arabic tradition, which was founded in large part on Greek writings in philosophy and science (Aristotle's in particular), disappears after 1200, reappearing almost immediately in Western Latin thought. Throughout the century 1150–1250 a vast number of translations of most of Greek and Alexandrian philosophy and science were made from Arabic and Hebrew into Latin. This literary corpus, which had made its way around the Mediterranean littoral translated from Greek into Syriac and thence into Arabic and Hebrew, caught the attention of Latin scholars and such patrons of scholarship as King Frederick II of Sicily and Archbishop Raymond of Toledo. As a consequence, by about 1200 the indefatigable efforts of many translators working in many locations had made Greek thought, especially that of Aristotle, available to Latin thinkers. The impact of this solid and integrated corpus of natural science on the Western intellectual world was enormous, coming as it did into a climate where for centuries scholars eager for knowledge had had to content themselves with thirdhand encyclopedic compilations of inadequately developed science and scientific methodology.

AVERROES'S COMMENTARIES

The translations of the Greek writings were normally accompanied by many Greek and Arabic commentaries. Commentaries by Alexander of Aphrodisias and by Simplicius were frequent, but those by the Arab Averroes on the Aristotelian works were ultimately the most influential. During a long and varied career as judge, teacher, philosophical and medical adviser to several Muslim rulers, Averroes found time to compose a series of glosses and commentaries on Aristotle's works. These fall into three categories—short (often called epitomes), interme-

diate or middle, and long, a differentiation which probably corresponds to stages in the academic curriculum. The particular argumentation of certain passages of Aristotle presented by Averroes in the mass of commentary had strong appeal for many Western Latin thinkers, and the reflection of his interpretations in their own philosophical analyses gave rise to attitudes which were first termed (by Christian scholars suspicious of their novelties) Arabic and later more specifically called Averroist.

INITIAL IMPACT IN THE WEST. Upon translation the Greek writings, with their attendant commentaries, were rather quickly absorbed into Western Latin scholarship, but not without some formal opposition. These writings were banned at the University of Paris in 1210 and 1215, deemed usable only if corrected in 1231, and not officially introduced into the curriculum until 1255. This literature was nevertheless being intensively read during these years; the philosophical writings of Albertus Magnus (active at least as early as 1230), William of Auvergne (d. 1249), and Alexander of Hales (d. 1245), to name only three prominent examples, reveal an intimate acquaintance with the recently acquired corpus of Greek science. Similarly, in England the philosophy of Robert Grosseteste (bishop of Lincoln, died 1253) shows strong influences derived directly from the newly inherited Greek literature. In Italy, too, the Greek tradition was rapidly assimilated into the scholarly milieu, but the Italian intellectual atmosphere was either medical, as it had been at the University of Salerno for several centuries, or else legal, as at Bologna. There do not appear proscriptions by Italian ecclesiastical authorities as stringent as those made at the University of Paris throughout the thirteenth century, and the possible intellectual conflicts raised by the introduction of these writings into a context of Christian philosophy do not seem to have been seriously felt.

Intellectual conflicts became extremely explicit, however, when the Aristotelian writings were conceived to be in direct confrontation with doctrines of Christian faith. Aristotle asserted, for example, the eternity of the world, the unlikelihood of individual immortality, the possibility of man's attaining ethical perfection in this life, and other theses incompatible with tenets of Christian belief. The appearance of such philosophical conclusions, apparently well reasoned and buttressed by Arabic commentary, occasioned some severe crises for Western Christian philosophy.

The chief agents presenting these, as well as other, renderings of Aristotle were the commentaries of Averroes. For centuries he was called simply the "Commentator" in Latin writings, and his expositions of the Aristotelian corpus were read into the seventeenth century. Cesare Cremonini (d. 1631), the last of the self-proclaimed Averroists, used these commentaries, and even at that late date he was considered unorthodox enough to be included in an array of formal proceedings along with Galileo Galilei himself. Unorthodoxy makes strange bedfellows when the resolute claimant of Aristotelianism and the architect of a scientific rupture with Aristotelian Scholasticism are included in the same condemnatory document.

LATIN AVERROISM

Historically, *Averroism* is a designation applied to certain interpretations of Aristotelian doctrine by Western Latin thinkers. (There are medieval Jewish philosophers holding positions close to these, but the epithet itself does not seem to have been applied to them.) It was originally a term of opprobrium; no one called himself Averroist until possibly John of Jandun (c. 1286–c. 1328), who was followed by Urban of Bologna (fl. 1334) and Paul of Venice (d. 1428). During the thirteenth century Averroists were the object of violent philosophical attack and severe authoritarian action.

Averroes insisted upon, and many scholars in the Western faculties of arts concurred in, the reliable logic of Aristotle's argumentation. Thus, there was clearly the necessity of the purely rational acceptance, given Aristotle's premises, of such "unorthodox" conclusions as have been mentioned. Acceptance is, however, intolerable for serious Christian thinkers, and so such conclusions were taken to be erroneous and thus subversive when pronounced in the schools. When thirteenth-century arts masters taught Aristotle in this fashion, they were awarded (by their opponents) the pejorative title Averroist, and official action often resulted. Siger of Brabant, Boethius of Dacia, and Bernier of Nivelles, masters in the faculty of arts at Paris, were all named in condemnations of the 1270s. This special mention seems to have had limited effectiveness; although these particular masters disappeared from the intellectual scene, countless commentaries on Aristotle dating from the last quarter of the thirteenth century offer similar interpretations and similar caveats as to the logical validity, if not truth, of these interpretations. No recorded disapprovals have been found.

Incidentally, this represents another aspect of the history of intellectual conflict. Explicit authoritarian condemnations were more often the result of a refusal to accept organizational discipline than of a genuine

philosophical error or ideological heresy. This can be illustrated in the careers of Gottschalk (d. c. 868), Peter Abelard (1079–1142), and Roger Bacon (c. 1214/1220–1292), all of whom were subjected to ecclesiastical punishment although little of their thinking was drastically at variance with established or recommended philosophical systems.

THE "DOUBLE TRUTH" PROBLEM

Every exposition of Averroism must examine the problem, arising in the thirteenth century, of the "double truth." The masters of arts, reading Aristotle and following his rigorous logic to conclusions incompatible with certain propositions held by faith, tried to resolve apparent contradictions by including in their commentaries reservations of this nature: "Although this conclusion has been reached according to the method of Aristotle and the Commentator, nevertheless faith and truth declare otherwise." While proclaiming logical rigor and precise validity for Aristotelian arguments, they conceded the final determination of truth itself to the Christian faith.

In this historical context it has often been maintained, both in the thirteenth century and in contemporary scholarship, that such thinkers were actually practicing a system of "double truth," in which a proposition can be true in natural philosophy but contradict a proposition true in theology and conversely. But, as Étienne Gilson and other scholars have convincingly pointed out, no master of arts has yet been found explicitly holding such a radical position. Regardless of the apparent persuasiveness of Aristotelian argument, the truth itself was always the dominant prerogative of Christian faith. In the face of such overwhelming requirements, the limitations and inadequacies of natural reason were recognized by the arts masters.

ATTEMPTED SOLUTIONS

Thus, an intellectual crisis of the first magnitude appeared in Western scholarship in the early thirteenth century. The attempts to deal with this conflict between important arguments in Greco-Arab philosophies and Christian-oriented intellectual systems fall into several main categories.

REASON NOT APODICTIC. First, the masters of arts, whose primary professional obligation was teaching natural philosophy, the core of which was Aristotle and his commentators, resorted to the attitude that although such science was orderly and rigorous, the unreliability of reason and the merely probable nature of its results suggested that conclusions based on such unaided reason must always yield, with respect to truth, to the apodictic proclamations of the faith. Such masters never claimed "truth" for a proposition of natural philosophy in conflict with a proposition of faith; they insisted on its logical validity, however, and conceded the determination of truth-value to faith. In this manner they endeavored to handle an intractable intellectual dilemma and at the same time to avoid subjecting themselves to overt charges of intellectual and ideological inconsistency.

AUGUSTINIANS. Second, masters of theology—for example, Bonaventure, Peter John Olivi, and, in the first decade of the fourteenth century, John Duns Scotus—employed a methodology often termed Augustinian. Their attempt to resolve the difficulties entailed, essentially, an assimilation of Aristotelian natural philosophy into a hierarchical scheme of knowledge. Such a resolution provided a coherent and orderly vertical relation among the several sciences, proceeding from the less perfect to the more perfect, from the less well known to the more surely known, from the less exact to the more exact. Such a structure, culminating in God himself, the ultimate source of perfection, knowledge, and precision, could be coherent and consistent and could accommodate both Christian doctrine and a qualified, because essentially incomplete, natural philosophy. But the achievement of this coherence was purchased at the cost of Aristotle himself, for his scheme of the sciences does not envisage a vertical, or hierarchical, ordering, whereby lesser sciences derive their logic, meaning, and reality from superior sciences. His sciences are basically ordered horizontally, diversified methodologically, and irreducible to any single set of common and univocally meaningful fundamental principles.

THOMAS AQUINAS. Third, the preeminent theologian St. Thomas Aquinas (1224?–1274) attempted a massive resolution maintaining the logical integrity and autonomy of Aristotelian natural philosophy while setting forth a supplementary and compatible structure of Christian theology. The two disciplines run in parallel courses, with differences based on distinctive premises and arguments, but there are many points where the propositions in each discipline are the same and are concluded to be true in both domains. These points were taken by Thomas to ensure the compatibility of Aristotelian natural philosophy and Christian theology, and by this means Thomas sought to sustain a consistent intellectual whole comprehending Greek philosophy and Christian truth.

The carefully poised system of Thomas was not, however, influential in his own time, and most of his immediate successors in the theological faculties preferred to continue in the Augustinian methodology. By the early fourteenth century, moreover, both approaches—the Augustinian assimilative technique and Thomas's sophisticated and delicately poised structure of complementary systems—were abandoned. This becomes explicit in the philosophy of William of Ockham, in whose thought natural science and systematic theology are totally independent domains.

Insofar, then, as the masters of arts, reading Averroes in close conjunction with Aristotle, tended to bring forward the incompatibilities between the two systems, it is possible to affirm the judgment of Gilson that "the rupture of Christianity is from this moment an accomplished fact."

ITALIAN AVERROISM

As a designation Averroism disappeared in the intellectual history of the University of Paris after the first quarter of the fourteenth century, although there are many manuscripts making explicit these crucial difficulties; however, their overt dependence on and acknowledgment of Averroes's commentaries diminish. From about 1300 to 1650 the term *Averroism*—assumed favorably by some thinkers and in a derogatory fashion by others—is found associated with philosophical activity in the Italian universities, Bologna and especially Padua.

Renan wished to establish a dichotomy between Averroist and Alexandrist Aristotelianism in Italy at this time. This distinction was based on alternative interpretations of Aristotle's *De Anima*. The Averroist view emphasized that personal, individual immortality could not be established in Aristotle's writings. In this interpretation the soul, when separated from the body, loses all individuality—a conception congenial to the Muslim doctrine of complete impersonal fusion at the apex of noetic experience. In purely Aristotelian terminology this is known as the theory of the unity of the active intellect—that is, that any form distinct from matter is one in species and never individuated. The Alexandrist analysis likewise denied the possibility of individual immortality but argued against the separate subsistence of the soul under any conditions whatsoever; when the soul-body composite dissolves, nothing remains.

This distinction is an oversimplification of the complexities of Italian Aristotelianism between 1300 and 1650, but it was employed by the scholars themselves and may thus be used with appropriate reservations. However, whether or not these thinkers were designated Averroist or Alexandrist, they all did agree in affirming the logical integrity of Aristotelian natural philosophy, even though some conclusions reached in this philosophy appeared in radical contradiction to dicta of Christian faith.

Although it would be misleading to speak crudely of an Alexandrist tradition in the later Middle Ages, there were eminent philosophers who, though thoroughly convinced of the logical autonomy of Aristotelian thought as such, did not adhere to the letter of Averroes's rather Platonic or Augustinian interpretation. Jean Buridan (d. c. 1358) at Paris and Pietro Pomponazzi (d. 1525) and Jacopo Zabarella (d. 1589), both at Padua, can be taken to fall within the non-Averroist but still naturalistic method of Aristotelian natural philosophy.

Averroism as a term designating a tradition, type, or method of philosophizing is difficult to make precise. Thinkers of varied methodological persuasions—for instance, Siger of Brabant and John of Jandun—have been called Averroist. Averroism can, however, be solidly connected with Latin Aristotelianism where Latin Aristotelianism is taken to include philosophies that agree on the logical rigor and systematic autonomy of natural philosophy as exemplified in Aristotle's writings. Since such arguments appear to lead to conclusions inconsistent with truths of Christian faith, Averroism in its earliest usage was pejoratively employed. But the demands of reason, working with the Aristotelian corpus, were insistent, and by the middle of the fourteenth century philosophers began to proclaim themselves openly Averroist. Gilson has suggested that Averroism was essentially conservative and sterile, but it is clear that it was an integral part of the tradition of Aristotelian scholasticism and that its disappearance in the seventeenth century coincided with the demise of medieval Scholasticism itself.

See also Augustinianism; Averroes; Averroism in Modern Islamic Philosophy; Thomas Aquinas, St.

Bibliography

LATIN AVERROISM

Ernest Renan's pioneering work on Averroism, *Averroès et l'averroïsme*, rev. ed. by H. Psichari (Paris, 1949), originally published in Paris, 1852, is now out of date. See rather *Averroes and the Aristotelian Tradition. Sources. Constitution and Reception of the Philosophy of Ibn Rushd (1126–1198)*, edited by G. Endress and J. A. Aertsen (Leyden-Boston-Köln: Brill, 1999); *Averroismus im Mittelalter und in der Renaissance*, edited by F. Niewöhner, L. Sturlese (Zürich: Spur Verlag, 1994); and *L'Averroismo in Italia, Atti dei*

convegni Lincei, 40 (Rome: Accademia nazionale dei Lincei, 1979). For new manuscript evidences on Hebrew and Latin translations of Averroes's commentaries and their impact in the West, see Carlos Steel and Marc Geoffroy, *La béatitude de l'âme* (Paris: Vrin, 2001); Colette Sirat and Marc Geoffroy, *L'Original arabe du* Grand Commentaire d'Averroès *au* De Anima *d'Aristote* (Paris: Vrin, 2005). On Boethius of Dacia, Siger of Brabant, and the "double-truth" problem, see Luca Bianchi, *Censure et Liberté intellectuelle à l'université de Paris* (Paris: Les Belles Lettres, 1999); J. F. Wippel, "The Condemnations of 1270 and 1277 at Paris," *Journal of Mediaeval and Renaissance Studies* 7 (1977): 169–201. On Aquinas, see *Aquinas against the Averroists, on there being only one intellect*, translated by R. MacInerny (West Lafayette, IN: Purdue University Press, 1993).

JEWISH AVERROISM

Leaman, Oliver. "Jewish Averroism." In *History of Islamic Philosophy*, edited by S. H. Nasr and O. Leaman. London: Routledge, 1996.

Zonta, Mauro. *La filosofia ebraica medievale. Storia e testi.* Roma-Bari: Editori Laterza, 2002.

ARABIC ARISTOTELIANISM

The major figures of Arabic Aristotelianism are al-Kīndī (d. 873), al-Farābī (d. 950), ibn-Sīnā (Latinized Avicenna, d. 1037), and Ibn-Bajja (Latinized Avempace, d. 1138). See *The Cambridge Companion to Arabic Philosophy*, edited by Peter Adamson and Richard C. Taylor (Cambridge: CUP, 2005); *Storia della filosofia nell'Islam medievale*, edited by Cristina D'Ancona Costa (Torino: Einaudi, 2005). These important works provide references to many other valuable and supplementary studies. For a reader, see *Medieval Islamic Philosophical Writings*, edited by Muhammad Ali Khalidi (Cambridge: CUP, 2005). For al-Ghazzali's *Incoherence of the Philosophers* see *The Incoherence of the Philosophers: English Translation of Imam Ghazali's Tahafut al-Falasifa*, translated by Michael E. Marmura (Brigham Young University's Islamic Translation Series), (Brigham Young University Press: 2000). For Averroes' reply, see his *Incoherence of the Incoherence*, translated by Simon Van Den Bergh (Gibb Memorial Trust, 1978). In this debate the term *philosophers* is a translation of the Arabic *falasifa*, which is in turn a transliteration of the Greek *philosophoi*. *Falasifa* thus has the special meaning of "thinkers following the Greek tradition" and not the general sense of philosophers as such.

TRANSMISSION OF GREEK AND ARABIC PHILOSOPHY

Rencontres de cultures dans la philosophie médiévale: traductions et traducteurs de l'antiquité tardive au XIVe siècle, edited by J. Hamesse and M. Fattori. Louvain: Brepols, 1990.

The Introduction of Arabic Philosophy into Europe, edited by C. E. Butterworth and B. A. Kessel. Leiden: Brill, 1994.

Stuart MacClintock (1967)
Bibliography updated by Alain de Libera (2005)

AVERROISM IN MODERN ISLAMIC PHILOSOPHY

Averroes (Ibn Rushd) largely disappeared from the Islamic world after his death in 1198, but returned through the influence of Ernest Renan, who in the nineteenth century presented Averroes as a hero of rationalism and antireligious skepticism. Many Arab intellectuals were educated in France and came into contact with Renan's views, and they played a large part in the Islamic renaissance movement (*al-Nahḍa*). This was designed to combine adherence to religion with a commitment to reason, something that Ibn Rushd was regarded as exemplifying in his life and work. His work has been used to oppose the forces of conservatism and traditionalism in the Arab world. Averroes was in his life also opposed by the local religious authorities, as are his modern supporters in the Arab world.

The *tanwīr*—or enlightenment movement—is more radical than the *Nahḍa* because it often is highly critical of the influence of established religion. Its central text is *Falsafat Ibn Rushd* by Faruḥ Antūn, as well as the books of al-ʿĀṭif al-ʿIrāqī on Averroes. Although of limited influence in the Arab world as a whole, and even in Egypt where it has some presence in the universities of Cairo, this movement has created considerable intellectual discussion among Arab philosophers. Its critics regard it as too aligned with the West and too antagonistic to Islam, but proponents of the *tanwīr* movement argue that only a radical separation of faith and politics can initiate an appropriate degree of modernity into the Arab world.

See also Averroes; Averroism; Islamic Philosophy.

Bibliography

Antun, Faruh. *Falsafat Ibn Rushd*. Beirut, Lebanon: Dar al-Farabi, 1988.

Wahba, Mourad, and Mona Abousenna, eds. *Averroes and the Enlightenment Movement*. New York: Prometheus, 1996.

Oliver Leaman (2005)

AVERROISM, JEWISH

See *Jewish Averroism*

AVICEBRON

See *Ibn Gabirol, Solomon ben Judah*

AVICENNA

(980–1037)

Avicenna, whose full name was Abū ʿAlī al-Ḥusayn ibn ʿAbd-Allāh ibn Sīnā, was the most renowned and influential philosopher of medieval Islam. He was a Persian, born near Bukhara, then the capital of the Persian Samānid dynasty. His father was a partisan of the heterodox Ismāʿīlī sect, whose theology drew on current popularized Neoplatonism. As a boy, Avicenna was exposed to Ismāʿīlī doctrine but found it intellectually lacking. He received some of the basic Islamic religious education, then studied logic, mathematics, the natural sciences, philosophy, and medicine, mastering these subjects by the age of eighteen. A certain al-Nātilī introduced him to logic, geometry, and astronomy, but Avicenna was largely self-taught. He records that he was able to fathom Aristotle's *Metaphysics* only after a chance discovery of a commentary on it by al-Fārābī (Alfarabi). Appointed physician at the Samānid court, he intensified his studies at its excellent library. Thereafter, he states, he added little to his stock of learning but deepened his understanding of what he had acquired.

In 999 Samānid rule disintegrated with the onslaught of the Turkish Ghaznawid dynasty. Avicenna left Bukhara to roam the cities of Transoxania and Iran, serving local warring princes. Between 1015 and 1022 he acted as both vizier and physician to the ruler of Hamadan; after the latter's death he was imprisoned but was released four months later when ʿAlā al-Dawla, the ruler of Isfahan, temporarily occupied the city. Soon afterward, disguised as a dervish, Avicenna left Hamadan for Isfahan, where he spent the rest of his life as physician to ʿAlā al-Dawla. This was a relatively peaceful period of his life, during which he undertook astronomical investigations. A serious interruption occurred in 1030, when the Ghaznawids sacked Isfahan and some of Avicenna's works were pillaged and lost. He died in Hamadan while accompanying his patron on a campaign against that city.

Over a hundred of Avicenna's works have survived, ranging from encyclopedic treatments to short treatises and covering, apart from philosophy and science, religious, linguistic, and literary matters. He wrote some works in Persian, of which the *Dānishnāma-yi ʿAlāʾī* ("The Book of Science Dedicated to ʿAlā al-Dawla") is the most important. Most of his works, however, are in Arabic. His chief medical work is *al-Qanūn fi al-Tibb* ("The Canon of Medicine"), a synthesis of Greek and Arabic medicine which also includes his own clinical observations and views on scientific method. The most detailed philosophical work is the voluminous *al-Shifāʾ* ("The Healing"). *Al-Najāt* ("The Deliverance") is largely a summary of *al-Shifāʾ*, although there are some deviations. *Al-Ishārāt wa al-Tanbīhāt* ("The Directives and Remarks") gives the quintessence of Avicenna's philosophy, sometimes in an aphoristic style, and concludes with an expression of his mystical esoteric views, a part that relates to certain symbolic narratives which he also wrote.

PHILOSOPHY

Avicenna forged a comprehensive philosophical system that owed a great deal to Aristotle, but his system cannot be strictly called Aristotelian. In both his epistemology and his metaphysics he adopted Neoplatonic doctrines but formulated them in his own special way. There were other Greek influences: Plato on his political philosophy; Galen on his psychology; the Stoics on his logic. Nearer home was the influence of Islamic theology and philosophy. The theologians had stressed the contingent nature of things, subjecting Aristotelian causal theory to severe logical and empirical criticism. Avicenna undertook to meet this criticism and attacked the theologians' formulation of the notion of contingency, but he nonetheless was influenced by it. The Islamic philosopher who influenced him most was al-Fārābī; Avicenna adopted al-Fārābī's concept of the identity of divine essence and existence, and developed his dyadic emanative system into a triadic scheme. As both metaphysician and political thinker, Avicenna interpreted the Islamic religion in terms of his own system. Whether this religion remains "Islamic" when so interpreted is a debatable point, but it conditioned the way Avicenna formulated his philosophy.

METAPHYSICS. Although Avicenna's system rests on his conception of the Necessary Existent, God, he held that the subject matter of metaphysics is broader than theology. As distinct from physics, which considers moving things "inasmuch as they move," metaphysics is concerned with the existent "inasmuch as it exists." We arrive at the Necessary Existent by first examining the attributes of the existents. Avicenna undertook such examination in detail, drawing those distinctions which greatly influenced Latin scholastic thought. One such distinction is that between a universal like "horse," by definition predicable of many instances, and a universal like "horseness," in itself outside the category of such predication; considered in itself, horseness is simply horseness, neither one nor many. Related to this is the fundamental distinction between essence and existence.

If we examine any existing species, we find nothing in its essence to account for its existence. In itself, such an existent is only possible: it can exist or not exist. From *what* it is, we cannot infer *that* it exists, although in fact it exists. Something has "specified" it with existence; and this something, argued Avicenna, must be its necessitating cause. If it were not—if it were a cause that may or may not produce its effect—we would have to suppose another cause; and if this cause were not necessitating, yet another; and so on ad infinitum. But an infinity of such causes—even if allowed—would not specify the possible with existence. Hence, such an existent must be necessitated by another, by which Avicenna meant that its existence is the consequence of the essence of another existent. The theory involved here is that of essential causality, where causal action is a necessary attribute of a thing's essential nature and where cause and effect coexist. Existents form a chain of such essential causes; and since these coexist, the chain must be finite. Otherwise it would constitute an actual infinite, which Avicenna deemed impossible. The chain must proceed from an existing essence that does not derive its existence externally. This is God, the Necessary Existent, who, Avicenna attempted to demonstrate, must be eternal, one, and simple, devoid of all multiplicity. Since God, the necessitating cause of all the existents, is eternal, his effect, the world, is necessarily eternal.

The world emanates from God as the consequence of his self-knowledge. Self-knowledge, however, does not imply multiplicity in the knower; nor does multiplicity proceed from God directly. God's act of self-knowledge necessitates the existence of one intellect. Multiplicity proceeds from this intellect which undergoes three acts of awareness, corresponding to the three facts of existence it encounters: (1) God's existence as necessary in itself; (2) the intellect's own existence as necessitated; (3) the intellect's own existence as only possible in itself. These three acts of awareness necessitate the existence of three things—another intellect, a soul, and the first heaven, respectively. The second intellect, in turn, undergoes a similar cognitive process, necessitating another triad; the third intellect, yet another; and so on down to the sphere of the moon. The last intellect thus generated is the Active Intelligence, whose acts of cognition necessitate the world of generation and corruption.

Avicenna's cosmology was oriented toward the Ptolemaic system as modified by some of the Islamic astronomers, who, in order to explain the precession of the equinoxes, added another heavenly sphere beyond that of the fixed stars, and Avicenna inclined toward regarding the number of intellects as ten. He was not dogmatic on this point, however, leaving the question of the number of intellects adjustable to changes in astronomical and cosmological theory. What he insisted on was that the number of intellects should be at least equal to the number of heavens.

In this scheme Avicenna attempted to make precise the relation of the celestial intellects to God, something left uncertain in Aristotle. According to Avicenna, the intellects derive their existence from God and are arranged in an ontological and normative hierarchy corresponding to their proximity to God. God, for him, is not only the prime mover but also the cause of existence. The celestial intellects, in turn, although deriving their existence from God, cause other existents and act as teleological causes. Thus, in each of the triads the heavenly body is moved by its soul through the soul's desire for the intellect. The souls differ from the intellects in that they have a material aspect enabling them to have direct influence over the particulars in the sublunar world and to know them in their particularity. Neither God nor the celestial intellects have this direct influence and know these particulars only "in a universal way."

THE HUMAN SOUL. According to Avicenna, both the human soul and the rational knowledge it acquires are emanations from the Active Intelligence. As such, the body "receives" the soul and the soul "receives" rational knowledge. Certain combinations of formed matter induce the reception from the Active Intelligence of the vegetative soul. Other combinations induce, in addition to this, the reception of the animal soul; and others, in addition to these two, induce the reception of the rational soul, with its practical and theoretical aspects. The human rational soul is an individual, indivisible, and immaterial substance that does not exist as an individual prior to the body—Avicenna denied the theory of transmigration. Further, it is created *with* the body, not "imprinted" on it. The body is no more than the soul's instrument, which the soul must use for perfecting itself through the attainment of theoretical knowledge; this involves complete control of the animal passions. Souls inherently incapable of attaining theoretical knowledge can still control the body and live pure lives by adhering to the commands of the revealed law. With the body's corruption (death), the soul separates to exist eternally as an individual. Souls that have led pure lives and have actualized their potentialities continue in eternal bliss, contemplating the celestial principles. The imperfect souls, tarnished by the body, continue in eternal torment,

vainly seeking their bodies, which once were the instruments of their perfection.

Avicenna denied bodily resurrection but insisted on the Soul's individual immortality. To begin with, he held that the immaterial is incorruptible. Moreover, he was convinced not only of the soul's immateriality but also of its individuality. He argued for both these points simultaneously: When one refers to himself as "I," this cannot be a reference to his body. If a man were to come into being fully mature and rational but suspended in space so that he was totally unaware of his physical circumstances, he would still be certain of one thing—his own existence as an individual self.

Theoretical knowledge consists in the reception of the intelligibles from the Active Intelligence. The primary intelligibles, the self-evident logical truths, are received by men directly, without the need of the soul's preparatory activities on the sensory level. The secondary intelligibles, concepts and logical inferences, whose reception is limited to people capable of demonstrative knowledge, normally require preparatory activities involving the external and internal senses—sensation, memory, imagination, estimation, and cogitation, or imaged thinking. Avicenna assigned special faculties and physiological places to these activities. The human intellect undergoes various stages in its acquisition of the intelligibles. At first it is a material intellect, a pure potentiality analogous to prime matter, ready for the reception of the intelligibles. With the reception of the first intelligibles it becomes the intellect with positive disposition. When it is *in the act* of receiving the secondary intelligibles, it becomes the acquired intellect. When an intellect that receives the secondary intelligibles is not engaged in the act of reception, it is termed "the actual intellect."

POLITICAL AND RELIGIOUS PHILOSOPHY. Avicenna followed al-Fārābī in holding that revealed religion gives the same truths as philosophy but in the symbolic, particular, imaged language that the masses can understand. According to Avicenna, some prophets receive this particular symbolic knowledge directly from the celestial souls. Such reception involves the prophet's imaginative faculty. In a higher form of prophecy that is intellectual, the prophet receives from the celestial intellects not only the first intelligibles, without the need of the soul's preparatory activities, but also the second. Prophetic reception of knowledge thus differs from the philosophical "in manner." It also differs "in quantity." Avicenna suggested that the prophet receives all or most of the intelligibles from the Active Intelligence "all at once." This intellectual revelation is then translated into the language of imagery and divulged to the public. It includes the basic commands of the revealed law, without which man as a political animal cannot survive. Hence, divine goodness must reveal the law at certain moments of discussion through prophets. Prophecy is thus necessary in the sense that it is required for the survival of civilized society and in the sense that it is necessitated by the divine nature. Having argued for the necessity of prophecy, Avicenna proceeded to accommodate Islamic institutions within his philosophical framework.

The high point of Avicenna's religious philosophy is his discussion of mysticism in the *Ishārāt.* In this work he adopted the language of Islamic mysticism (*sufism*) to describe the mystic's spiritual journey to God: Beginning with faith and motivated by desire and love, the mystic undertakes spiritual exercises that first bring him to interrupted glimmerings "of the light of the Truth." These experiences become progressively more frequent and durable until the stage of "arrival" is reached, in which the mystic has a direct and an uninterrupted vision of God. According to Avicenna, there are further stages beyond this, but he declined to discuss them. He also ascribed some of the prophetic qualities to mystics, without implying that all mystics are law-revealing prophets. On the other hand, his language suggests that he held that all prophets are mystics.

LOGIC AND DEMONSTRATIVE METHOD. Avicenna inherited the Aristotelian and Stoic logical tradition as expounded by al-Fārābī and the Baghdadi school of logicians but treated his subject more independently. He found the then current classification of syllogisms into "attributive" (categorical) and "conditional" too narrow. Instead, he classified them as "connective" and "exceptive." Connective syllogisms have the *form* of the categorical, but their premises may consist of combinations of attributive and conditional statements. Similarly, exceptive syllogisms have the form of one of the two types of conditional syllogisms—the conjunctive, corresponding to the *modus ponens* and the *modus tollens,* and the disjunctive in which the logical relation is exclusive—but their premises may consist of attributive statements conditionally related, or combinations of conditional and attributive statements. He attempted the quantification of both conjunctive and disjunctive premises, discussed the temporal aspects of quantification in general, and treated the modality of premises and arguments at length.

Although Avicenna held logic to be merely a tool of knowledge and strove to treat it as distinct from philoso-

phy, his discussion of the epistemic status of premises (which carried him considerably beyond anything in Aristotle) rendered his logic philosophically committed; his discussion of demonstrative premises was committed to his epistemology and metaphysics of causality. He followed Aristotle in his treatment of demonstrative inference, distinguishing between demonstrations that give the reasoned fact and those that give the fact. The former involve inference from cause to effect; the latter, inference from effect to cause. He also included in the latter class inferences from one effect to another. This is possible when it has been established that a single cause necessitates two effects; Avicenna gave a medical example of a disease that has two symptoms.

Avicenna's endorsement of the *Posterior Analytics* extended to much of the *Physics.* He rejected, however, Aristotle's account of falling bodies, substituting for it a theory of acquired force that was a forerunner of the theory of momentum.

Although some Jewish and Islamic philosophers (Maimonides, ibn Bājja [Avempace], Averroes) showed a preference for al-Fārābī, Avicenna's influence overshadowed the latter's in the Islamic world. The mystical side of his philosophy was elaborated in the illuminationist thought of the philosophers of Persia. The orthodox Ashʿarite theologians who condemned his metaphysics adopted his logic, and his medical works continued to dominate the Islamic world until the emergence of the modern university.

In the Latin West his emanative metaphysics and epistemology blended with the Augustinianism of the Franciscan schools as a basic ingredient of their thought. His influence on Thomas Aquinas was considerable, notwithstanding Thomas's rejection of many Avicennian doctrines. He also greatly influenced the development of logic and science, his *Canon of Medicine* remaining an authoritative medical text into the seventeenth century.

See also al-Fārābī; Averroes; Cosmology; Ibn Bājja; Islamic Philosophy; Maimonides; Neoplatonism; Plato; Sufism; Thomas Aquinas, St.

Bibliography

There is no collected edition of Avicenna's works in the original. The closest thing to a collection of Persian works consists of Volumes X–XXV of the series *Silsila-i Intishārāt-i Anjuman-i Āthāri Millī* (Publications of the Society of National Monuments). These volumes, published by the University of Teheran, appeared in 1951 on the occasion of Avicenna's millenary.

Critical editions of parts of the *al-Shifāʾ,* in the original Arabic have been appearing in a series sponsored by the Egyptian ministry of education and supervised by Dr. Ibrahīm Madkūr. The following volumes have appeared: *Al-Madkhal* ("Isagoge," Cairo, 1952); *Al-Khaṭaba* ("Rhetoric," Cairo, 1954); *Al-Burhān* ("Demonstration," Cairo, 1955); *Jawami ʿilm al-Mūsīqā* ("Music," Cairo, 1956); *Al-Safsata* ("Sophistic," Cairo, 1958); *Al-Maqūlāt* ("Categories," Cairo, 1959); *Al-Ilāhīyāt* ("Metaphysics," 2 vols., Cairo, 1960); and *Al-Qiyās* ("Syllogism," Cairo, 1964).

Translations of Avicenna's works include *Avicenna on Theology,* translated by A. J. Arberry (London, 1951); *A Treatise on the Canon of Medicine,* translated by O. Cameron Gruner (London: Luzac, 1930); *Avicennae de Congelatione et Conglutinatione Lapidum,* translated by E. J. Holmyard and D. C. Mandeville (Paris: P. Guethner, 1927); *Die Metaphysik Avicennas,* translated by Max Horten (New York: R. Haupt, 1907); *La métaphysique du Shifāʾ*, translated by M. M. Anawati, mimeographed ed. (Quebec, 1952); *Le livre de science,* translated by Mohammad Achena and Henri Massé, 2 vols. (Paris: Société d'édition "Les Belles Lettres," 1955); *Le livre des directives et remarques,* translated by A. M. Goichon (Paris, 1955); *Psychologie d'Ibn Sīnā (Avicenne) d'après son oeuvre Aš-šifaʾ*, edited and translated by Jan Bakoš (Prague: Editions de l'Académie tchécoslovaque des Sciences, 1956); Ralph Lerner and Muhsin Mahdi, eds., *Medieval Political Philosophy: A Source Book* (New York: Free Press of Glencoe, 1963), pp. 95–121; and *Avicenna's Psychology,* translated by Fazlur Rahman (London: Oxford University Press, 1952).

Studies include S. M. Afnan, *Avicenna, His Life and Works* (London: Allen and Unwin, 1958); M.-T. d'Alverny, "Anniya-Anitas," in *Mélanges offerts à Étienne Gilson* (Paris: J. Vrin, 1959), pp. 59–91; E. G. Browne, *Arabian Medicine* (Cambridge: Cambridge University Press, 1921); Henri Corbin, *Avicenna and the Visionary Recital,* translated by W. R. Trask (New York: Pantheon, 1960); M. Cruz Hernandes, *La metafísica de Avicena* (Granada: University of Granada, 1949); Louis Gardet, *La penséd religieuse d'Avicenne (Ibn Sīnā)* (Paris, 1951); Étienne Gilson, "Les sources gréco-arabes de l'augustinism avicennisant," in *Archives d'histoire doctrinale et littéraire du moyen âge* 4 (1929): 5–149; A. M. Goichon, *La distinction de l'essence et de l'existence d'après Ibn Sīnā (Avicenne)* (Paris: Desclée, de Brouwer, 1937), *Lexique de la langue philosophique d'Ibn Sīnā (Avicenne)* (Paris, 1938), *La philosophie d'Avicenne et son influence en Europe médiévale* (Paris: Adrien-Maisonneuve, 1944), and *Vocabulaires comparés d'Aristote et d'Ibn Sīnā* (Paris, 1939).

See also M. E. Marmura, "Avicenna's Theory of Prophecy in the Light of ash ʿarite Theology," in W. S. McCullough, ed., *The Seed of Wisdom* (Toronto: University of Toronto Press, 1964), pp. 159–178, and "Some Aspects of Avicenna's Theory of God's Knowledge of Particulars," in *American Journal of Oriental Studies* 82 (3) (1962): 299–312;S. H. Nasr, *Islamic Cosmological Doctrines* (Cambridge, MA, 1964), pp. 177–281; Shlomo Pines, "La 'philosophie orientale' d'Avicenne et sa polémique contre les Bagdadiens," in *Archives d'histoire doctrinale et littéraire du moyen âge* 27 (1952): 5–37; Nicholas Rescher, *Studies in the History of Arabic Logic* (Pittsburgh: University of Pittsburgh Press, 1963), pp. 76–86 and 91–105; Djamil Saliba, *Étude sur la métaphysique d'Avicenne* (Paris, 1926); and G. M.

Wickens, ed., *Avicenna: Scientist and Philosopher; A Millenary Symposium* (London: Luzac, 1952).

Bibliographies include G. C. Anawati, *Essai de bibliographie avicennienne* (Cairo: Dar al-Mar'arf, 1950), and Yahya Mehdawi, *Bibliographie d'Ibn Sīnā* (Teheran, 1954).

Michael E. Marmura (1967)

AVICENNA [ADDENDUM]

Avicenna played an important role in Islamic aesthetics. Poetry relies on imagination, he argues, but that does not mean it is entirely without logical structure. On the contrary, one can only understand poetry if it is analyzed in terms of the syllogism. The premises of such a reasoning are statements produced by writers to bring about emotional states in the reader or hearer. This only works if there is some reason to connect the use of words with the emotion, and that reason has precisely to be a logical reason. The conclusion is the pleasure one feels at the bold and striking use of language, and because one is not the only person who can enjoy that use of language, the conclusion is also available to others. It then becomes a general conclusion rather like the conclusion of a strictly demonstrative syllogism. Avicenna follows a similar strategy in discussing music, in that at the end of the reasoning process one undergoes when listening to it, a conclusion is drawn in terms of a pleasure that one can expect others to share.

See also Aesthetics, History of; Logic, Traditional; Music, Philosophy of; Philosophy of Language.

Bibliography

Fakhry, Majid. *Ethical Theories in Islam*. 2nd ed. Leiden, Netherlands: Brill, 1994.

Goodman, Lenn E. *Avicenna*. London: Routledge, 1992.

Gutas, Dimitri. *Avicenna and the Aristotelian Tradition: Introduction to Reading Avicenna's Philosophical Works*. Leiden, Netherlands: Brill, 1988.

Inati, Shams. "Ibn Sina." In *History of Islamic Philosophy*. 2 vols., edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996.

Kemal, Salim. *The Poetics of Alfarabi and Avicenna*. Leiden, Netherlands: Brill, 1991.

Nasr, Seyyed Hossein. "Ibn Sina's Oriental Philosophy." In *History of Islamic Philosophy*. 2 vols., edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996.

Oliver Leaman (2005)

AXIOLOGY

See *Value and Valuation*

AXIOM AND AXIOMATIC METHOD

See *Logical Terms, Glossary of*

AYER, ALFRED JULES

(1910–1989)

Alfred Jules Ayer, the British philosopher, received his education at Eton, where he was a king's scholar, and at Christ Church, Oxford. After graduating in 1932, he spent some time at the University of Vienna familiarizing himself with the logical positivist movement, then little known among English-speaking philosophers. He returned to Oxford in 1933 as a lecturer in philosophy at Christ Church and in 1935 became a research fellow of the college. Army service in World War II kept him from philosophy until 1945, when he went back to university teaching as fellow and dean of Wadham College, Oxford. In the following year he became Grote professor of the philosophy of mind and logic at University College, London, where he remained until his return to Oxford as Wykeham professor of logic in 1959.

Ayer's first book, *Language, Truth and Logic*, was published in 1936. Its combination of lucidity, elegance, and vigor with an uncompromisingly revolutionary position has made it one of the most influential philosophical books of the century. As Ayer explains in the preface, the views he advocates derive from Bertrand Russell and Ludwig Wittgenstein among modern philosophers and from the earlier empiricism of George Berkeley and David Hume and have much in common with the logical positivism of the Vienna circle. But he accepts none of these influences uncritically and clearly puts his own stamp on the position he outlines. He adopts Hume's division of genuine statements into logical and empirical, together with a principle of verification that requires that an empirical statement shall not be counted as meaningful unless some observation is relevant to its truth or falsity. This starting point has drastic and far-reaching results. Metaphysical statements, since they purport to express neither logical truths nor empirical hypotheses, must accordingly be reckoned to be without meaning. Theology is a special case of metaphysics; affirmations of divine

existence are not even false, they are without sense. For the same reason, value statements in ethics or aesthetics fail to attain the status of genuine statements and are exposed as expressions of emotion with imperative overtones. The a priori statements of logic and mathematics are empty of factual content and are true in virtue of the conventions that govern the use of the words that compose them. The tasks left for philosophy after this withdrawal from its traditional boundaries are those of solving by clarification the problems left untouched by the advance of the sciences. Philosophy is an activity of analysis and is seen, in the end, to be identical with the logic of science.

The second edition of the book (1946) contains an introduction that modifies, though it does not retract, the main theses of the first edition. Ayer's attention here is directed chiefly to giving a precise formulation of the principle of verification. His original version is replaced by a much more elaborate and carefully worded formula. Both versions have, however, been shown to be faulty in admitting as meaningful metaphysical statements of precisely the kind that the principle is designed to outlaw. Indeed, there seems to be a weakness of the principle in that, it appears plausible only when its expression is left uncomfortably vague.

The Foundations of Empirical Knowledge (1940) is concerned with two groups of problems, those of perception and those of "the ego-centric predicament" (privacy and publicity in language and in sense experience and the problem of other minds). The most interesting and original feature of the book is Ayer's treatment of the terminology of sense data as a language in which the problems of perception can be most appropriately dealt with rather than as a thesis embodying a discovery about the facts of sense experience. *Thinking and Meaning* (1947) was Ayer's inaugural lecture in the University of London. It is a trenchant application of Ockham's razor to the problems of intentionality and the relations between minds, thinking objects, words, and meaning. This short, powerful essay has so far received less than its due of critical attention. *Philosophical Essays* (1954) is a collection of papers ranging over philosophical logic, the theory of knowledge, and moral philosophy. Half the papers are carefully argued treatments of problems raised in Ayer's first two books; in particular, "The Analysis of Moral Judgements" is a moderate and persuasive restatement of the hints on ethics thrown out in *Language, Truth and Logic*.

In 1956 Ayer published *The Problem of Knowledge*, his most important book since his first was published in 1936. It is a sympathetic and constructive treatment of the various problems of philosophical skepticism. After a short discussion of philosophical method and the nature of knowledge, he discusses at length the pattern of skeptical arguments. He then examines three problems familiar from his earlier work—perception, memory, and other minds—as instances of skepticism at work. It may be that no statement is immune from doubt, but this does not entail that no statement can be known to be true. Where statements cannot, even in principle, be justified, we may conclude not that they are to be rejected but rather that no justification is called for.

The Concept of a Person (1963) is a collection of essays. The most striking, the one that gives the book its title, is a notable survey of some aspects of the problems of body, mind, and personal identity. The outcome can be roughly summarized as follows: To say that I own a mental state *M* is to say that there is a physical body *B* by which I am identified and that a state of *B* causes *M*.

Ayer's Shearman Lectures at the University of London in 1964 were on induction and probability. This was a new field of interest for Ayer, although it was foreshadowed in two papers in *The Concept of a Person*.

Ayer's work is very much of a piece, both in style and attitude. He became more catholic in interest and more cautious and temperate in expression than in his earlier writings. But his arguments were informed by the same principles and set out with the same grace and clarity. He leaned perhaps too heavily on Hume's dichotomy of statements into logical and factual, and he has not so far set himself seriously to meet contemporary criticisms (particularly those of W. V. O. Quine) that have been made of this famous distinction. This is at once a weakness of his present position and, perhaps, a presage of its future development.

Ayer died on June 29, 1989. He was professionally active virtually until the time of his death. In recognition of his accomplishments and public service, Ayer was Knighted in 1968. The following year he published both *Metaphysics and Common Sense*, a set of essays on diverse topics, and also *The Origins of Pragmatism*, an account of the philosophies of William James and Charles Sanders Peirce. In 1970 Ayer presented the William James lectures at Harvard in which he discussed the thought of G. E. Moore and Bertrand Russell. In that same year he gave the John Dewey lectures at Columbia University in which he revisited induction and probability, the topic of his 1964 Shearman lectures at the University of London. Ayer's *The Central Questions of Philosophy* (1974) is regarded by some as a new and refined version of his clas-

sic work *Language, Truth and Logic*. After serving for almost twenty years as Wykeham professor of logic at Oxford, Ayer retired from the position in 1978. Shortly thereafter a festschrift *Perception and Identity* was published in his honor, which contained essays by prominent thinkers and Ayer's replies to them. In 1982 Ayer offered his *Philosophy in the Twentieth Century* as a possible sequel to Bertrand Russell's *History of Western Philosophy*. He published interpretations of Ludwig Wittgenstein in 1985, Voltaire in 1986, and Thomas Paine in 1988. He also wrote two autobiographical volumes: *Part of My Life* (1977) and *More of My Life* (1984). His rather lengthy obituary in *The Times* of London concludes with these words: "Ayer was not a major philosopher like Russell or Wittgenstein, or even, perhaps like Popper and Ryle. But he was a very able philosopher indeed, endowed with particularly sparkling intellectual gifts, an admirable if slightly chilly prose style and unflagging energy. As a philosophical teacher and influence there is no one to compare with him since Russell and Moore."

See also Analytic and Synthetic Statements; Basic Statements; Berkeley, George;Ethics, History of; Hume, David; Logical Positivism; Other Minds; Personal Identity; Private Language Problem; Quine, Willard Van Orman; Russell, Bertrand Arthur William; Skepticism, History of; Verifiability Principle; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY AYER

Books

Language, Truth and Logic. London: Gollancz, 1936; 2nd ed., 1946.

The Foundations of Empirical Knowledge. London: Macmillan, 1940.

Thinking and Meaning. London: Athlone Press, 1947.

Philosophical Essays. London: Macmillan, 1954.

The Problem of Knowledge. London: Macmillan, 1956.

The Concept of a Person. London: Macmillan, 1963.

The Origins of Pragmatism. London: Macmillan, 1968.

Metaphysics and Common Sense. London: Macmillan, 1969.

Russell and Moore: The Analytical Heritage. London: Macmillan, 1971.

Probability and Evidence. London: Macmillan, 1972.

Bertrand Russell. London: Fontana, 1972.

The Central Questions of Philosophy. London: Weidenfeld, 1973.

Part of My Life. London: Collins, 1977.

Hume. Oxford: Oxford University Press, 1980.

Philosophy in the Twentieth Century. London: Weidenfeld, 1982.

Freedom and Morality and Other Essays. Oxford: Clarendon Press, 1984.

More of My Life. London: Collins, 1984.

Ludwig Wittgenstein. London: Penguin, 1986.

Articles and Symposium Pieces

"Jean-Paul Sartre." *Horizon* (1945).

"Albert Camus." *Horizon* (1945).

"Some Aspects of Existentialism." *Rationalist Annual* (1948).

"Logical Positivism—A Debate." Delivered on the BBC June 13, 1949. The participants were Ayer and F. C. Copleston. Published in *A Modern Introduction to Philosophy*, edited by P. Edwards and A. Pap (Glencoe, IL: Free Press, 1957).

"Professor Malcolm on Dreaming." *Journal of Philosophy* (1960): 517–535. Malcolm's reply, with Ayer's rejoinder, *Journal of Philosophy* (1961), 294–299.

WORKS ON AYER

For critical discussion of Ayer, see John Wisdom, "Note on the New Edition of Professor Ayer's *Language, Truth and Logic*," *Mind* 57 (228) (1948): 401–419, reprinted in Wisdom's *Philosophy and Psycho-analysis* (Oxford: Blackwell, 1953); H. H. Price, "Critical Notice of A. J. Ayer's *The Foundations of Empirical Knowledge*," *Mind* 50 (199) (1941): 280–293; H. H. Price, "Discussion: Professor Ayer's Essays," *Philosophical Quarterly* (1955); D. J. O'Connor, "Some Consequences of Professor A. J. Ayer's Verification Principle," *Analysis* (1949–1950); W. V. O. Quine, "Two Dogmas of Empiricism," in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953); M. Lazerowitz, "Strong and Weak Verification I," *Mind* (1939) and "Strong and Weak Verification II," *Mind* (1950), reprinted in Lazerowitz's *The Structure of Metaphysics* (London: Routledge and Paul, 1955).

Austin, J. L. *Sense and Sensibilia*. Oxford: Clarendon Press, 1962.

Church, A. "Review of *Language, Truth, and Logic*." *Journal of Symbolic Logic* (1949): 14:52–53.

Foster, J. *A. J. Ayer*. London: Routledge & Kegan Paul, 1985.

Griffiths, A. P. *A. J. Ayer Memorial Essays*. Cambridge, U.K.: Cambridge University Press, 1991.

Hahn, L. E. *The Philosophy of A. J. Ayer*. La Salle, IL: Open Court, 1992.

Hanfling, O. *Ayer*. London: Routledge, 1999.

Honderich, T. *Essays on A. J. Ayer*. Cambridge, U.K.: Cambridge University Press, 1991.

Macdonald, G., ed. *Perception and Identity*. London: Macmillan, 1979.

Macdonald, Graham, and C. Wright, eds. *Fact, Science, and Morality*. Oxford: Blackwell, 1986.

Martin, R. *On Ayer*. Belmont, CA: Wadsworth, 2000.

Rogers, Ben. *A. J. Ayer: A Life*. London: Grove Press, 2002.

D. J. O'Connor (1967)
Donald M. Borchert (2005)
Bibliography updated by Michael J. Farmer (2005)

B

BAADER, FRANZ XAVIER VON
(1765–1841)

Franz Xavier von Baader, the German philosopher and theologian, was born in Munich. He studied medicine at Ingolstadt and Vienna and practiced for a short time, but soon abandoned this career. While he was in England from 1792 to 1796 studying mineralogy and engineering, he became interested in philosophy and theology. On his return to Germany he formed friendships with Friedrich Heinrich Jacobi and Friedrich Wilhelm Joseph von Schelling. Although Baader later broke with Schelling, the three philosophers continued to exert strong influence on one another. Baader was appointed superintendent of the Bavarian mines and won a prize from the Austrian government for inventing a new method of glass manufacture. He retired in 1820 to devote himself to philosophy.

Baader's two major works are *Fermenta Cognitionis* (Vols. I–IV, Berlin, 1822–1824; Vol. V, Munich, 1825) and *Spekulative Dogmatik* (5 fascicles, Munich, 1827–1828). He was appointed professor of philosophy and speculative theology at the new University of Munich in 1826. He stopped lecturing on theology in 1838, when the Catholic bishop banned the public discussion of theology by laymen, but he continued to lecture on philosophy until his death.

Baader's philosophy is couched in aphorisms, symbols, and analogies, and it is therefore difficult to summarize. He detested David Hume's empiricism, William Godwin's radicalism, and Immanuel Kant's rationalism. He turned the critical method he had learned from Kant against criticism itself, calling for a return to the mystical tradition of Jakob Boehme, Paracelsus, Meister Eckhart, the Cabala, the Neoplatonists, and the Gnostics. He believed that since God is in all things, all knowledge is partly knowledge of God. God is not an abstract being but an eternal process, eternally becoming. As God creates himself, he comes to know himself. The relation between his will and his self-consciousness is the Holy Spirit. The Trinity is an eternal possibility in God and only becomes actual in nature, which is the principle of selfhood eternally produced by God. Nature is God alienated from himself—his shadow, his desire, his want. The purpose of the existence of nature is to afford an opportunity for the redemption of humanity.

Morality is not a matter of inner law, as Kant believed, but apprehension of, and obedience to, God's will. Salvation depends on prayer, faith, and the sacraments as well as on morality and good works. Humans are social

beings under the law of the state, and the subjects owe total subservience to their ruler. But the state is under the law of the church. Any departure from this divinely ordained order leads to the twin modern evils of despotism and liberalism.

Baader sought a theistic, Catholic philosophy reconciling nature and spirit, science and religion, the individual and society. He believed that philosophy had to go back to its sources, from which it had been separated since the time of Descartes. Baader was thus a precursor of the neoscholastic revival, but his own teachings, close to heresy, have no important place in the movement.

See also Boehme, Jakob; Eckhart, Meister; Gnosticism; Godwin, William; Hume, David; Jacobi, Friedrich Heinrich; Kabbalah; Kant, Immanuel; Neoplatonism; Paracelsus; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

The collected works of Franz von Baader were published as *Sämmtliche Werke,* 16 vols. (Leipzig: H. Bethmann, 1851–1860). Vol. XV contains a biography; Vol. XVI, a systematic exposition of Baader's ideas.

Works on Baader are D. Baumgardt, *Franz von Baader und die philosophische Romantik* (Halle: Salle, M. Niemeyer, 1927); J. Claasen, *Franz von Baaders Leben und theosophische Werke,* 2 vols. (Stuttgart, 1886–1887), and *Franz von Baaders Gedanken über Staat und Gesellschaft* (Gütersloh, Germany, 1890); and Kuno Fischer, *Zur hundert jährigen Geburstagsfeier Baaders* (Erlangen, Germany, 1865).

Adam Margoshes (1967)

BACHELARD, GASTON

(1884–1962)

Gaston Bachelard, the French epistemologist and philosopher of science, was born at Bar-sur-Aube. He was a postal employee until 1913, when he gained his *licence* in mathematics and science and became a teacher of physics and chemistry at the Collège of Bar-sur-Aube. In 1927 he received his doctorate of letters and in 1930 became professor of philosophy at the University of Dijon. From 1940 to 1954 he held the chair of history and philosophy of science at the University of Paris.

Bachelard expounded a dialectical rationalism, or "dialogue" between reason and experience. His philosophy was a departure from the view of rational discovery as a process whereby new knowledge is assimilated into a system that changes only insofar as it grows. He rejected the Cartesian conception of scientific truths as immutable elements of a total truth that is in process of being put together like a jigsaw puzzle.

According to Bachelard, experiment and mathematical formulation are mutually complementary. Mathematics is not merely a means of expressing physical laws, nor is it a static realm of ideas; it is "committed." In this context Bachelard talked of "applied rationalism." Bachelard held that the empirical world is not utterly discontinuous and absurd; the confrontation of an isolated, rational human mind with an indifferent and meaningless world postulated by some existentialists is naive. Scientific hypotheses, and even scientific facts, do not present themselves passively to the patient investigator but are created by him. The investigator's reasoning and the natural world on which it operates together constitute a second nature over and above the crudely empirical one.

Bachelard described his conception of this two-way process in which rational organization and experiment are in constant cooperation as a "philosophy of saying no" (*philosophie du non*). It involves negation because the scientific attitude is necessarily "open" or "available" (*disponible*), and the scientist may be obliged at any time to recast his formulation of reality by facts which fail to fit into the old formulation. Since it is frequently mathematical, the reformulation may not necessarily involve the adoption of a new model, but it will often be analogous to a change of structure. At the same time, there will be no jettisoning of truths: The *philosophie du non* destroys nothing, Bachelard held; it consolidates what it supersedes. The framework may be recast and the picture of reality transformed, but only in such a way that the new phenomenon might have been foreseen.

Bachelard did not confine himself to an exclusively rationalist philosophy of science. He saw both technological and imaginative thinking as issuing from reverie and emotion into practical expression. His works on the psychological significance of the four elements, earth, air, fire, and water, illustrate this. He rejected, for example, the common account of the discovery of fire in the rubbing together of two sticks, seeing it rather as the outcome of a kind of symbolical representation of sexual intercourse. Thus passion is no more metaphorical fire than fire is metaphorical passion. Our science and our poetry have a common origin accessible only to psychoanalysis. There is a unity in Bachelard's studies on reason and imagination. In both cases he stressed the projective or creative role of the mind; in art "the subject projects his dream upon things," and in modern science, "above

the *subject,* beyond the immediate object … is the *project.*"

See also Epistemology; Philosophy of Science, History of; Rationalism.

Bibliography

WORKS BY BACHELARD

L'intuition de l'instant. Paris: Stock, Delamain and Boutelleau, 1932.

Le pluralisme cohérent de la chimie moderne. Paris: J. Vrin, 1932.

La dialectique de la durée. Paris, 1933.

Le nouvel Esprit scientifique. Paris: Presses universitaires de France, 1934.

La formation de l'esprit scientifique. Paris: J. Vrin, 1938.

La psychanalyse du feu. Paris: Gallimard, 1938.

Lautréamont. Paris: J. Corti, 1939.

La philosophie du non. Paris: Presses universitaires de France, 1940.

L'eau et les rêves. Paris: J. Corti, 1942.

L'air et les songes. Paris: J. Corti, 1943.

La terre et les rêveries de la volonté. Paris, 1945.

La terre et les rêveries du repos. Paris, 1945.

Le rationalisme appliqué. Paris: Presses universitaires de France, 1949.

La poétique de l'espace. Paris: Presses universitaires de France, 1957.

La poétique de la rêverie. Paris: Presses universitaires de France, 1960.

La flamme d'une chandelle. Paris: Presses universitaires de France, 1961.

WORKS ON BACHELARD

For a complete bibliography of Bachelard's works and for articles on him, see the "Bachelard" issue of the *Revue internationale de philosophie* 19 (1964). See also Jean Hyppolite's article, "Gaston Bachelard ou le romantisme de l'intelligence," *Revue philosophique de la France et de l'étranger* 144 (1954): 85–96.

Colin Smith (1967)

BACHOFEN, JOHANN JAKOB
(1815–1887)

Johann Jakob Bachofen, Swiss jurist, cultural anthropologist, and philosopher of history, studied philology, history, and law at the universities of Basel, Berlin (under Friedrich Karl von Savigny), and Göttingen. After taking his doctorate in 1839 in Roman law, he spent two years at the universities of Oxford, Cambridge, and Paris. In 1841, Bachofen was offered the chair in Roman law at the University of Basel, and a year later he was appointed a judge of the criminal court at Basel. In 1844 he resigned his professorship to devote himself to legal and anthropological research. In 1866 he also gave up his position as a judge. He traveled widely and lived for long periods in Greece, Italy, and Spain.

Bachofen's major works were in the fields of ancient Roman law and Greek antiquity. The work for which he is best known is *Das Mutterrecht. Eine Untersuchung über die Gynaikokratie der alten Welt nach ihrer religiösen und rechtlichen Natur* (Stuttgart, 1861). Following up Herodotus's description of a matriarchal system among the Lycians, Bachofen investigated diverse ancient myths and concluded that both matrilineal descent and matriarchal rule developed out of a state of unregulated promiscuity (*Hetärismus*) by virtue of the difficulty of ascertaining paternity under such conditions. He maintained that the dominant role of the mother in both the economic and political spheres was a phenomenon common to all primitive societies and that this role was inseparably linked to religious beliefs that established the secular primacy of woman on the basis of the cult of a female deity.

There is no element of evolution in Bachofen's theory. His main interest lay in tracing the transmission of social cultures, not in the biological characteristics attending heredity. Bachofen likewise rejected interpretations of myths in terms of individual psychology. The elements that constituted for him the essential ingredients of historical traditions—myths, cults and rituals, customs, law, and folklore—were *shared* characteristics and hence, in his view, objective factors. They embodied a people's collective "spirit," or *Volksgeist,* which, though a persistent continuum in social development, nonetheless operated at a nonrational and subconscious level. According to Bachofen it was the function of the woman and mother to preserve and uphold these nonrational historical forces and thus to exercise a uniting influence, whereas man, representing the progressive and rational forces, exercised a dividing influence over the development of humankind. The historical process consisted in a continuous striving for reconciliation between these opposing tendencies.

Das Mutterrecht encountered considerable skepticism, if not hostility, among contemporary anthropologists. Bachofen was charged with introducing rather fanciful and value-loaded notions into his theory and with confusing matrilineal descent with a matriarchate. But even though some of his theses have been disproved and others continue to be challenged, many of his suggestions have led to fruitful further research into the fam-

ily customs of primitive peoples. Increasingly, too, Bachofen's works have been appraised as a major contribution to the philosophy of history.

Bachofen stressed the continuity of historical sequences and, above all, the close interpenetration of myth and history. In opposition to Georg Wilhelm Friedrich Hegel, Bachofen attached decisive importance to myths and symbols in the shaping of human history, since he accorded to them a far greater and more lasting emotive power than he did to rational concepts. In his stress on the irrational elements in history, as also in his insistence on regarding history as a continuous organic growth, Bachofen shared some of the basic premises of romantic thought. Yet, like Johann Gottfried Herder, the great precursor of romanticism, he never regarded himself as a romantic. Indeed, he explicitly repudiated the nostalgic sentimentality with which a number of romantics approached the study of the past.

Bachofen's political views show an undeniable affinity for the conservatism of the political romantics, but here also he was more directly influenced by Edmund Burke, whom he had assiduously studied during his stay in England. Paradoxically enough, Bachofen has often been associated with L. H. Morgan as one of the founders of a socialist philosophy of history. Bachofen did stipulate a "communist" origin of humankind in that he denied the existence of private property among primitive communities. He also prophesied an ultimate return to communism, understood in this sense. But he viewed such a return as a regression, not as "progress." Bachofen saw in socialism and democracy portents of social and political decay, for he held them to be inherently inimical to harmonious community life. Social and political harmony presupposed, in his view, the willing acceptance of the principle of subordination, for he regarded this principle as the prime source of a naturally and divinely ordered historical process.

Bachofen may have gone too far in the political application of his tradition-centered historicism, just as he probably overstated the role of woman in the development of religion, morals, law, and customs. But he did advance a functional conception of social development, in which social structures are seen as elements of a historical continuum and as constituents of an "idea-system" of nonrational and nonlogical beliefs and symbols, and in so doing he substantially contributed to the understanding of both ancient communities and societies of the modern world.

See also Burke, Edmund; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Morgan, Lewis Henry; Philosophy of History; Philosophy of Social Sciences; Savigny, Friedrich Karl von.

Bibliography

WORKS BY BACHOFEN

Collections

J. J. Bachofens gesammelte Werk. Edited by Karl Meuli. 10 vols. Basel: B. Schwabe, 1943–.

Der Mythus von Orient und Occident. Edited by Manfred Schroeter. 2nd ed. Munich: Beck, 1956. Contains a selection of Bachofen's works, with an introduction by Alfred Baeumler and a bibliography of Bachofen's published works.

Individual Works

Das Naturrecht und das geschichtliche Recht in ihren Gegensätzen. Basel, 1841.

Das römische Pfandrecht. Basel: Schweighauser, 1847.

Ausgewählte Lehren des römischen Civilrechts. Bonn: A. Marcus, 1848.

Versuch über die Gräbersymbolik der Alten. Basel, 1859; 2nd ed., Basel: Helbing and Lichtenhahn, 1925.

Das Mutterrecht. Stuttgart: Krais and Hoffmann, 1861.

Die Unsterblichkeitslehre der orphischen Theologie. Basel, 1867.

Die Sage von Tanaquil. Heidelberg: J.C.B. Mohr, 1870.

Antiquarische Briefe, 2 vols. Strassburg: K.J. Trübner, 1880, 1886.

Römische Grablampen. Basel, 1890.

WORKS ON BACHOFEN

Bernoulli, C. A. *J. J. Bachofen als Religionsforscher.* Leipzig, 1924.

Bernoulli, C. A. *J. J. Bachofen und das Natursymbol.* Basel: B. Schwabe, 1924.

Burckhardt, Max. *J. J. Bachofen und die Politik.* Basel, 1943.

Kerenyi, K. *Bachofen und die Zukunft des Humanismus.* Zürich: Rascher, 1945.

Klages, Ludwig, *Der Kosmogonischen Eros.* Munich, 1922.

Wolf, Erik, article on Bachofen. In *Neue Deutsche Biographie,* Vol. 1, pp. 502–503. Berlin: Duncker and Humblot, 1953. Also contains a detailed bibliography of secondary literature.

Frederick M. Barnard (1967)

BACON, FRANCIS

(1561–1626)

Francis Bacon, Baron Veralum, Viscount St. Albans, gained renown both as an English statesman and a natural philosopher. Bacon was instrumental in the replacement of Aristotelian natural philosophy, effecting a major shift to thinking about the natural world in exclusively empirical and experimental terms, although he remained

entrenched in Aristotelian thought to a significant degree. His achievement was twofold: First, he transformed the discipline of philosophy from something contemplative that focused above all on moral questions into something practical that focused centrally on questions in natural philosophy (what is now called science). Second, his work in the natural sciences resulted in the formulation of precepts that are now regarded as foundation stones of the inductive modern scientific method: moving inferentially from observable effects to deeper underlying causes and eliminating various possible explanations by testing their consequences against experiment or observation.

LIFE

Bacon was born on January, 22, 1561, the eldest son of Sir Nicholas Bacon, lord keeper of the Great Seal, and Ann, second daughter of Sir Anthony Coke, known for her strong Protestant sympathies. Bacon attended Trinity College, Cambridge, from 1573 to 1575, but moved to Gray's Inn in 1575, traveling to France—where he came into contact with Italian republican ideas—in 1576 and remaining there until his father's death in 1579. From that time onward he began a career in law and politics that took him from his first parliamentary seat (1581), admission to the bar (1582), deputy chief steward of the Duchy of Lancaster (1594), solicitor general (1607), attorney general (1613), member of the Privy Council (1616), lord keeper of the Great Seal (1617), to his being created viscount of St. Albans in 1621. In that same year he was impeached and spent the rest of his life in comparative isolation from the court society he had enjoyed for the previous fifteen years. He died on April 9, 1626. His death has traditionally been attributed to his contracting pneumonia as a result of leaving his carriage to test the preserving effects of cold on a chicken, but it is more likely that he died of an overdose of inhaled niter or opiates, self-prescribed to cure a long-running illness.

His intellectual career falls into three stages. From 1592 to1602, his main concern was the reform of English law. From 1602 to around 1620, he worked on a very ambitious project in natural philosophy, advocating a form of atomism and setting out a new method of inquiry in natural philosophy, as well as investigating a huge number of topics in natural history. Around 1620 he began to publish parts of his grand scheme on a systematic basis, although Bacon could never be called a systematic philosopher. His plans for the reform of natural philosophy were not taken seriously by his English contemporaries during his lifetime, but within a few years of his death, critics of contemporary natural philosophy and founders of scientific academies in Italy, France, and England took him as their model, and by the beginning of the eighteenth century his name was linked with Newton's among the founders of modern science.

LAW AND RHETORIC

Bacon's first attempts at reform were in the area of law rather than natural philosophy. The law offered guidance on three questions that would subsequently make it a model for his proposed reform of natural philosophy: the reliability of testimony, what should be concluded from particular testimonies, and how one decided the relevance of particular laws to the case. It was the third of these that he saw most in need of reform, and he set out to investigate how the law might be systematized, how regular records and reviews of legal decisions might be provided, and whether some firm foundations for legal practice might be discovered. What is at issue here is what was referred to as the "discovery" of law. It was a shared premise that the law was structured in accord with reason and that this structure enabled one, in cases where the laws did not give a clear indication of infringements, to appeal to the implicit message of the common law. On the assumption that the law covered every eventuality, the task was to find one's way through its rational structure. The questions to which Bacon directed himself particularly were whether there was an optimal procedure by which to discover that rational structure and what the ultimate source of authority was in the case of dispute.

Bacon's emphasis on the role for discovery in the legal thinking reflects a concern with rhetoric, which plays a crucial role in both his proposed reform of the law and that of natural philosophy. The task of rhetoric was the formulation, organization, and expression of one's ideas in a coherent and compelling way. It was designed to help one find one's way around the comprehensive body of learning built up from antiquity, to recognize where appropriate evidence and arguments might be found, and to provide models that were designed to give one a sense of what was needed if a particular question was to be investigated, or a particular position defended, models that would be shared with those to whom one was expounding or defending one's case. Rhetoric, in Bacon's view, should help to focus the mental powers, to organize one's thoughts in the most economical fashion, and even (in writers like Quintilian) to provide vivid images or representations of situations that enabled one to convince oneself of a case (important especially in acting and in legal argument). It was designed to provide models to

show one how particular kinds of case were best defended, depending on such facts as the availability of and complexity of the evidence and the knowledge, opinions, or prejudices of the audience toward which one was directing one's arguments.

At a general level, rhetoric was deemed to be indifferent to subject matter because comprehensive procedures were recommended that would aid a case or investigation, scientific or a legal, although there would be similarities or analogies (as regards the standing of various kinds of evidence, for example) and dissimilarities (as regards the means of evidence collection, for example) between legal cases and those in natural philosophy. The law, taken in a broad sense, was seen as a paradigm case for rhetorical writers: Rhetorical treatises were often explicitly directed toward lawyers and legislators, and examples were geared to the kinds of problems that arose in law. In light of this, it is only to be expected that using a rhetorical model for knowledge—that is, a model that gives direction on how to collect and assess evidence for a view, how to make a judgment on the basis of that evidence, and how to establish the correctness of judgment, using precepts derived from the study of rhetoric—is in many respects using a legal model.

Rhetoric provided a theoretical foundation for the law, something which, at a practical level, worked with elaborate procedures for the gathering, assessing, and testing of evidence. This was exactly the kind of thing that Bacon had in mind for natural philosophy. What was unusual about his application of precepts learned from rhetoric and law to natural philosophy was that he used them to propose a fundamental reform of philosophy. While Bacon started from a consideration of the law, however, law did not act as a model in its own right. Its importance arose from the fact that (especially once it had been reformed along Baconian lines) it exemplified a rhetorically motivated account of discovery. This holds the key to Bacon's enterprise.

The best way to understand this reform is in terms of the pervasive Renaissance contrast, often drawn in classical terms, between the life of contemplation (*otium*) and the life of practical, productive activity (*negotium*). There had been a decisive shift in favor of the latter in sixteenth-century England. In particular, there was a stress on practical questions and the practical uses of learning; and philosophy—above all Scholastic philosophy—was widely regarded as a useless discipline that fostered argument for its own sake, never getting anywhere and never producing anything of value. Moreover, morality was widely seen as the key philosophical topic (following the Ciceronian model current in Renaissance Europe), and a number of Elizabethan thinkers, most notably the poet Sir Philip Sidney, were arguing that poetry was superior to philosophy because philosophy could only discourse on the nature of goodness, whereas poetry could actually move people to goodness, which was the point of the exercise.

Bacon did two things: He shifted philosophy from *otium* to *negotium*, and he made natural philosophy replace moral philosophy as the center of the philosophical enterprise. The combination of these two (and they are intimately connected) is a radical move that marks a decisive break not only with earlier conceptions of philosophy but also with earlier understandings of the task of the philosopher.

Natural philosophy existed in a number of forms in the sixteenth and seventeenth centuries, and there were two extreme forms. The first was exemplified by alchemy, an esoteric but practical discipline that had little connection with traditional philosophical practice and that suffered, in Bacon's view, from a lack of structure that produced few results, with most of that paucity attributable to chance. At the other extreme was Scholastic natural philosophy, an intensely theoretical discipline that, in Bacon's view, produced nothing at all; despite its great sophistication, it turned out to be almost exclusively verbal.

Bacon wanted something that could deliver the advantages of each of these without any of the disadvantages. He wanted something that would provide a detailed theoretical overview of the natural realm such that natural processes could not only be understood but, more importantly, also transformed on the basis of this understanding; this is the context of his famous dictum "knowledge is power." The ultimate aim was to transform natural processes for the common good (to be decided by the sovereign, on Bacon's view), and it was this, rather than some contemplative understanding of nature, that provided the rationale for natural philosophy and, by extension, philosophy per se.

Bacon himself formulated his project in terms of a politico-religious restoration of human dominion over the natural world, something lost with Adam's expulsion from Eden. Natural philosophy thereby gained a religious imperative, albeit one with little connection with traditional theology.

THE DOCTRINE OF IDOLS

If rhetoric is the first ingredient in Bacon's account of method, the second is a distinctive understanding of why

the need for method arises. Here Bacon's stress on a psychological dimension of knowledge is important: Questions of presentation of knowledge are not only recognized to be important but also have to be understood, where such an understanding is not supplementary to epistemology but actually part of it. At one level, there is nothing new in this, for it is simply part of a long tradition that begins in earnest with the Roman rhetoricians; but although it borrows from Greek writers, it is rather different from the approach to epistemological questions that we find in the classical Greek and Hellenistic philosophers. When one thinks of Bacon's general project in this context, it becomes clear that there is something novel here. For natural philosophy had generally been the preserve of Greek philosophy and had been pursued in a similar way by Scholastic philosophers. The Roman tradition, with the exception of Lucretius, had typically not concerned itself with speculative natural-philosophical questions, dealing instead with practical moral, political, and legal questions. In thinking of persuasion in terms of a psychological theory, of psychological theory as part of epistemology, and of epistemology as being directed primarily toward natural philosophy, Bacon was able to provide himself with some of the resources to recast natural philosophy not as a speculative but as a practical discipline.

This psychological dimension of epistemology is brought out fully in Bacon's doctrine of the "idols of the mind." These idols "do not deceive in particulars, as the others do, by clouding and snaring the judgment; but by a corrupt and ill-ordered predisposition of the mind, which as it were perverts and infects all the anticipations of the intellect." The second part of the "Great Instauration," which aims at the renewal of learning, is devoted to the "invention of knowledge" and has two components, one aiming to rid the mind of preconceptions, the other to guide the mind in a productive direction. These components are interconnected, for until we understand the nature of the mind's preconceptions, we do not know in what direction we need to lead its thinking. In other words, various natural inclinations of the mind must be purged before the new procedure can be set in place. Bacon's approach here is genuinely different from that of his predecessors, as he realizes. Logic or method in themselves cannot simply be introduced to replace bad habits of thought because it is not simply a question of replacement. The simple application of logic to one's mental processes is insufficient.

In his doctrine of the four idols of the mind, Bacon provides an account of the systematic forms of error to which the mind is subject, and this is a crucial part of his epistemology. It is in his treatment of internal impediments, the "idols," that the question is raised of what psychological or cognitive state we must be in to be able to pursue natural philosophy in the first place. Bacon believes an understanding of nature of a kind that had never been achieved since the Fall is possible in his own time because the distinctive obstacles that have held up all previous attempts have been identified, in what is in many respects a novel theory of what might traditionally have been treated under a theory of the passions, one directed specifically at natural-philosophical practice.

The "idols of the tribe" derive from human nature itself, above all from "the homogeneity of the substance of the human mind, or from its preoccupation, or from its narrowness, or from its restless motion, or from an infusion of the affections, or from the incompetence of the senses, or from the mode of impression." (*Works* 1857–1874, vol. 4, p. 58–59). The idols of the tribe affect everyone equally and are manifested in an eagerness to suppose that there is more order and regularity in nature than there actually is; in the tendency to neglect or ignore counterexamples to one's theories; in the tendency to extrapolate from striking cases with which one is familiar to all other cases; in the restlessness of the human mind, which means it is not satisfied with perfectly good fundamental explanations, mistakenly and constantly seeking some more fundamental cause ad infinitum; and in the tendency to believe true what one would like to be true.

The "idols of the cave," we are told, "take their rise in the peculiar constitution, mental or bodily, of each individual; and also in education, habit, and accident" (*Works* 1857–1874, vol. 4, p. 59). They include fascination with a particular subject, which leads to overhasty generalization; the readiness of some minds to focus on differences, and some to focus on similarities and resemblances, while a balance is difficult to attain naturally; and the fact that some minds are overly attracted to antiquity and some to novelty. Finally, there are those who are concerned wholly with material constitution at the expense of structure (the ancient atomists) and those who are concerned wholly with structure at the expense of material constitution.

These examples bring to light a very significant difference between the idols of the tribe and idols of the cave. There seems to be a set of routine procedures one can go through to remedy the situation in the latter case, procedures provided by the positive part of Bacon's doctrine—eliminative induction—whereas the case of idols

of the tribe is, in most cases, much more difficult to remedy.

The third kind of idols, those of the marketplace derive from the fact that we have to express and communicate our thoughts by means of language, which contains systematic deficiencies. One kind of problem with language lies in the fact that words "are commonly framed and applied according to the capacity of the vulgar, and follow those lines of division which are most obvious to the vulgar understanding. And whenever an understanding of greater acuteness or a more diligent observation would alter those lines to suit the true divisions of nature, words stand in the way and resist the change" (*Works* 1857–1874, vol. 4, p. 61). This leads to two kinds of linguistically induced deficiencies. First, language provides names that refer to things that do not exist, such as "Fortune, Prime Mover, Planetary Orbits, Element of Fire, and like fictions that owe their origin to false and idle theories" (*Works* 1857–1874, vol. 4, p. 61). The solution here is simply to get rid of the theories that give rise to these fictitious entities.

The second kind of case is not so straightforward. It arises because words have multiple and/or ill-defined meanings, and this is especially so in the case of terms such as *humid* that have been abstracted from observation. Bacon discerns a gradation in the "degrees of distortion and error" (*Works* 1857–1874, vol. 4, p. 62) of terms, beginning with names of substances, where the degree of distortion is low, proceeding through the names of actions, and finally reaching the names of qualities—he gives the examples of "heavy, light, rare, dense" (*Works* 1857–1874, vol. 4, p. 62)—where the degree of distortion is high.

Finally, the fourth kinds of impediment, the idols of the theater, are innate neither in the mind nor in language but are acquired from a corrupt philosophical culture and its perverse rules of demonstration. Here a general remedy is available, namely following Bacon's positive methodological prescriptions: "The course I propose for the discovery of sciences is such as leaves but little to the acuteness and strength of wits, but places all wits and understandings nearly on a level. For as in the drawing of a straight lines or a perfect circle, much depends on the steadiness and practice of the hand, but if with the aid of a rule or compass, little or nothing; so is it exactly with my plan" (*Works* 1857–1874, vol. 4, p. 62–63).

One of the great values of Bacon's account of the idols is that it allows him to make the case for method in a particularly compelling way. Indeed, never has the need for method been set out more forcefully, for Bacon's advocacy of method is not simply an aid to discovery. He argues that we pursue natural philosophy with seriously deficient natural faculties, we operate with a severely inadequate means of communication, and we rely on a hopelessly corrupt philosophical culture. In many respects, these are beyond remedy. The practitioners of natural philosophy certainly need to reform their behavior, overcome their natural inclinations and passions, but not so that, in doing this, they might aspire to a natural, prelapsarian state in which they might know things as they are with an unmediated knowledge. This they will never achieve. Rather, the reform of behavior is a discipline to which they must subject themselves if they are to be able to follow a procedure which is, in many respects, quite contrary to their natural inclinations, which is at odds with traditional conceptions of the natural philosopher, and which is indeed subversive of their individuality.

ELIMINATIVE INDUCTION

What Bacon is seeking from a method of discovery is something that modern philosophers would deem impossibly strong: the discovery of causes that are both necessary and sufficient for their effects. Why place such strong constraints on causation, so that we call something a cause only when the effect always occurs in the presence of this thing and never in its absence? What Bacon (like Aristotle before him) is after are the ultimate explanations of things, and it is natural to assume that ultimate explanations are unique. Bacon's method is designed to provide a route to such explanations, and the route takes us through a number of proposed causal accounts, which are refined at each stage. The procedure he elaborates, eliminative induction, is one in which various possibly contributory factors are isolated and examined in turn, to see whether they do in fact make a contribution to the effect. Those that do not are rejected, and the result is a convergence on those factors that are truly relevant. The kind of "relevance" that Bacon is after is, in effect, a set of necessary conditions: the procedure is supposed to enable us to weed out those factors that are not necessary for the production of the effect, so that we are left only with those that are necessary.

Bacon provides an example of how the method works in the case of color. We take, as our starting point, some combination of substances that produces whiteness—that is, we start with what are sufficient conditions for the production of whiteness, and then we remove from these anything not necessary for the color. First, we note that if air and water are mixed together in small por-

tions, the result is white, as in snow or waves. Here we have the sufficient conditions for whiteness, but not the necessary conditions, so next we increase the scope, substituting any transparent uncolored substance for water, whence we find that glass or crystal, on being ground, become white, and albumen, which is initially a watery transparent substance, on having air beaten into it, becomes white. Third, we further increase the scope and ask what happens in the case of colored substances. Amber and sapphire become white on being ground, and wine and beer become white when brought to a froth.

The substances considered up to this stage have all been "more grossly transparent than air." Bacon next considers flame, which is less grossly transparent than air, and argues that the mixture of the fire and air makes the flame whiter. The upshot of this is that water is sufficient for whiteness but not necessary for it. He continues in the same vein, asking next whether air is necessary for whiteness. He notes that a mixture of water and oil is white, even when the air has been evaporated from it, so air is not necessary for whiteness; but is a transparent substance necessary? Bacon does not continue with the chain of questions after this point but sets out some conclusions, namely that bodies whose parts are unequal but in simple proportion are white, those whose parts are in equal proportions are transparent, those whose parts are proportionately unequal are colors, and those whose parts are absolutely unequal are black. In other words, this is the conclusion that might be expected of the method of sifting out what is necessary for the phenomenon and what is not, although Bacon himself does not provide the route to this conclusion here.

This being the case, one can ask what his confidence in his conclusion derives from if he has not been able to complete the "induction" himself. The answer is that it derives from the consequences he can draw from his account. There are two ways in which the justification for the conclusions can be assessed: by the procedure of eliminative induction that he has just set out and by the consequences of those conclusions generated by it. In other words, there is a two-way process, from empirical phenomena to first principles, and then from first principles to empirical phenomena. This is a classic Aristotelian procedure. Where Bacon's version of it differs is in how the first step is carried out, and the difference turns on the use of eliminative induction.

Bacon's treatment of heat in *Novum Organum* follows essentially the same route, albeit in a more elaborate way. The first thing to do, he tells us, is to list "instances agreeing in the nature of heat," that is, a list of those cases in which heat is present: the rays of the sun, reflected rays, meteors, thunderbolts, volcanic eruptions, flame, burning solids, natural warm-baths, boiling liquids, hot vapors and fumes, fine cloudless days, air confined underground, wool and down, bodies held near a fire, sparks, rubbed bodies, confined vegetable matter, quick lime sprinkled with water, metals dissolved in acids or alkalis, the insides of animals, horse dung, strong oil of sulfur and of vitriol (i.e. sulphuric acid), oil of marjoram, rectified spirit of wine, aromatic herbs (which are hot to the palate), strong vinegar and acids (which burn those parts of the body where there is no epidermis, such as the surface of the eye), and, finally, intense cold, which can produce a burning effect (*Works* 1857–1874, vol. 4, p. 127–129).

The list makes no claims to completeness, of course, but presumably it does aim to give us some idea of the range of phenomena we have to deal with. Because, on Bacon's view, a cause should not only be present when the effect is present but also absent when the effect is absent, the next step ideally would be to list those cases where the effect was absent, but this is clearly an impossible task, for the list would be infinite. So what Bacon does is to list, in some detail, counterinstances to the items of the first list: cases when heat is absent or at least where there is some doubt. So, for example, the rays of the sun are hot, but those of the moon and the stars are not; the reflections of the sun's rays are usually hot but not in the polar regions; the presence of comets (counting these as a type of meteor) does not result in warmer weather; and so on. The point of this exercise is not simply to record known counterinstances, however, but also to suggest experiments that need to be carried out to discover whether there are counterinstances—for example, in the case of lenses and "burning mirrors," in connection with which he makes several suggestions.

Instances and counterinstances of heat are absolute questions, but we can also discover something of the nature of heat by comparative means, by making a comparison either of its increase and decrease in the same subject, or of its amount in different subjects, as compared one with another. For since the Form of a thing is the very thing itself, and the thing does not differ from the Form except in the way that the apparent differs from the real, or the external from the internal, or the thing in reference to man from the thing in reference to the universe, it necessarily follows that no nature can be taken as the true Form unless it always decreases when the nature in question decreases, and in like manner always

increases when the nature in question increases. (*Works* 1857–1874, vol. 4, p. 137)

This procedure for discovery requires the compilation of a "table of degrees or comparison," in which the instances previously listed are examined in respect of changes in heat. Putrefaction always "contains" heat, for example; inanimate things are not hot to the touch; the heat of lower animals such as insects is barely perceptible, but higher animals are hot to the touch; the heat in animals increases as a result of motion; the heat of celestial bodies is never sufficient to set fire to things on Earth; the sun and the planets give more heat in perigee than in apogee; and so on.

It is at this point that induction comes into play. The various instances must be reviewed with a view to eliminating those natures that can be absent while heat is still found, those natures that are present even though heat is absent, and those where the heat increases or decreases without a corresponding increase or decrease in the nature. Examples of the exclusions are as follows: Because the rays of the sun sometimes warm and sometimes do not, reject the nature of the elements as the explanation for heat; because of ordinary fire and subterranean fires, reject the nature of celestial bodies; because of boiling water, reject light or brightness; and so on. This process is less reliable than it might seem, for the exclusion of some simple natures and the narrowing down to others presupposes that we know what simple natures are, whereas in fact we do not know this; but the procedures followed this far, Bacon believes, do allow us to advance finally to the interpretation of nature, or at least to the first version of that interpretation, which he refers to as "the first vintage."

It is a premise of Bacon's account that the Form that causes an effect must be present in every instance and absent in every counterinstance, but he also points out that it is more evident in some instances than in others. This is particularly so in the case of heat: The tables show that hot things—such as flames and boiling water—are characteristically in rapid motion and that compression puts out a fire. The tables of results show, moreover, that bodies are destroyed or changed radically by heat, indicating that heat causes a change in the internal parts of the body and perceptibly causes its dissolution. Bacon concludes that heat is a species of the general genus of motion, but before examining what marks it out from other species of motion, he removes some ambiguities from the idea of heat. Sensible heat, for example, which is relative to individuals, not to the universe, is not heat properly speaking but the effect of heat upon the animal spirits. Moreover, the communication of heat from one body to another is not to be confused with the Form of heat, for heat itself and the action of heating are two different things. Nor is fire to be confused with the Form of heat, for fire is a combination of heat and brightness.

Having removed these ambiguities, Bacon turns to heat proper. A number of things mark it out as a distinctive species of motion. First, heat is a motion that causes bodies to expand or dilate "towards the circumference"—that is, in all directions—as is evident in the case of vapors or air, liquids such as boiling water, and metals such as iron, which expand when heated. Cold has the opposite effect in all cases. The second distinctive feature is that heat, aside from being a motion to the circumference, is also a motion upward. To determine whether the contrary holds in the case of cold, Bacon proposes an experiment in which a sponge soaked with cold water is placed at the bottom of one heated rod and at the top of another to determine whether one cools faster than the other. He further suggests that the one with the sponge at the top will cool the other end of the rod more quickly.

The third characteristic is that heat comprises a variety of nonuniform motion, whereby small parts of a body are moved in different ways, some motions being checked and others proceeding freely, with the result that the body experiences a constantly subsiding quivering and swelling motion. This third characteristic is evident in flames and in boiling water. Moreover, where the motion is of the whole, such as a gas escaping from confinement at great pressure, we find no heat. Bacon maintains that cooling proceeds like heating, in a nonuniform way, although the absence of great cold on the Earth makes this phenomenon less evident. Finally, the fourth characteristic of heat as a species of motion is that it acts rapidly, for comparison with the effects of age or time on the corruption of bodies shows a similar result, corruption or dissolution of bodies, and the difference must lie in the rate at which the parts of the body are penetrated. The case of cold is not mentioned here, and, unlike the first three characteristics, it is not clear just what Bacon would want to establish in the case of cold. He sums up by drawing two kinds of conclusions from this "first vintage:"

> The Form or true definition of heat … [is that] heat is a motion, expansive, restrained, and acting in its strife upon the smaller particles of bodies. But the expansion is thus modified; while it expands all ways, it has at the same time an inclination upwards. And the struggle in the particles is modified also; it is not sluggish, but hurried and with violence. Viewed with refer-

> ence to operation it is the same thing. For the direction is this: If in any natural body you can excite a dilating or expanding motion, and can so repress this motion and turn it back upon itself, so that the dilation does not proceed equably, but can have its way in one part and is counteracted in another, you will undoubtedly generate heat. (*Works* 1857–1874, vol. 4, p. 155)

This process is only the first stage in induction for Bacon, but it is the one that is both most novel and most problematic. In particular, it is appropriate to ask just how far the process of eliminative induction gets us. After all, to go back to the case of color, it is giant leap, indeed a qualitative leap, from noting that a mixture of oil and water is white to the conclusion that Bacon seeks, namely that those bodies whose parts are in simple proportion are white. Is it plausible to suppose that the continuation of the procedure would in fact get us to the conclusion? More particularly, the "directions" that have been followed to this stage remain wholly at the macroscopic level, yet their continued application is supposed to guide us to the particular microcorpuscular internal structure of a body that makes that body white. This issue prompts two questions: whether eliminative induction generates explanations and whether it genuinely involves a process that converges to a single cause or explanation.

On the first question Aristotelians would have resisted the demand that, in seeking an explanation for a physical phenomenon, they sift through all the possibilities until they have found *the* cause. The question turns on the relation between explanations and causes. Although the Greeks generally did not separate questions of causality and explanation, disputes did arise about which should be given priority. Cause would be given priority if one were seeking to determine or ascribe responsibility for something. Explanation would be given priority if one were trying to provide an account of all the relevant factors concerning how something came about, without necessarily wishing to apportion blame or responsibility. It makes a considerable difference which of these views we take. The Stoics, for example, maintained that the most important thing was to determine responsibility and, as a consequence, they viewed causes as being necessarily active. This view was supported by an analogy with the law, where the person deemed responsible for an offense is the person who had done whatever it was that resulted in the offense being committed.

The physical analogue here is a body: a cause is a body that does something to affect another body in some way. On this construal, an explanation is simply a statement of a cause: cause is prior to explanation. The alternative is to make explanation prior to cause, in which case we might say that a cause is whatever figures in the explanation of an event. Take the legal analogy: if we were seeking an explanation of why an offense occurred rather than simply trying to find out who was to blame, we might look at all kinds of factors, such as the conditions under which offenses of this kind usually occur, whether preventive measures had been taken, what kinds of things motivated people to commit offenses of this kind, and so on. In natural philosophy, Aristotle makes explanations prior to causes. His famous "four causes" are, in fact, four kinds of explanation, the combination of which is designed to yield a complete understanding of the phenomenon. If we know what something is, what it is made from, how it was made, and for what end it was made, we have a complete understanding of the phenomenon. To restrict oneself effectively to efficient causes, as Bacon does, will not yield such an understanding. So Aristotelians might well resist the notion that Bacon's procedure is going to lead to explanations.

Someone who is committed to making explanations prior to causes will argue that there are as many causes of something as there are explanations of that thing, for what will count as a cause will be determined by the kind of explanation one is seeking. Bacon has little in reply to this kind of move. In *Valerius Terminus*, he sets out the error of seeking the causes of particular things, which are "infinite and transitory," as opposed to "abstract natures, which are few and pertinent." Such criticism seems most appropriately leveled against alchemists and others, whom Bacon criticizes for their piecemeal approach, rather than Scholastic natural philosophers, who would agree with his stricture here. But, in fact, Bacon has the Scholastics in mind, telling us that, despite appearances, on closer examination they do not seek abstract natures. This somewhat surprising criticism is possible only because of the very restrictive interpretation he places on "abstract natures," which he compares to "the alphabet or simple letters, whereof the variety of things consisteth; or as the colors mingled in the painter's shell, wherewith he is able to make infinite variety of faces or shapes" (*Works* 1857–1874, vol. 3, p. 243). Clearly, what he really wants is an atomist account of the "abstract natures" of things, something that can be only defended on substantive natural-philosophical grounds. The kind of explanation he is seeking, namely an atomist/corpuscularian one, is without doubt guiding what is going to count as a satisfactory argument here.

This issue brings us to the second question. Is eliminative induction suitable as a method of discovering efficient causes? It is hard to see how it could not help in such a process, but it is far from clear that in itself it could generate an account of such causes. Indeed, it is impossible to see how Bacon's examples of whiteness and heat can be pursued further by eliminative induction to generate a conclusion of the kind he wants. One might admit some degree of convergence, but there is nothing like convergence to a point: things become squared off well before that stage.

TRUTH

Closely tied up with Bacon's account of method is his treatment of the question of truth. Bacon goes through a number of what he considers to be inadequate criteria that have been used to establish truth. He rejects criteria depending on antiquity or authority, those deriving from commonly held views, and those relying upon the internal consistency or the capacity for internal reduction of theories, presumably on the grounds that such criteria do not bear on the question of whether there is any correspondence between the theory and reality. He also rejects "inductions without instances contradictory" that is, inductions that restrict themselves to confirming a theory, as well as "the report of the senses." None of these, he tells us, are "absolute and infallible evidence of truth, and bring no security sufficient for effects and operations." That he ties in evidence for the truth of a theory and its usefulness here is no accident, for these are intimately connected, telling us in *Valerius Terminus* that

> That the discovery of new works and active directions not known before, is the only trial to be accepted of; and yet not that neither, in case where one particular giveth light to another; but where particulars induce an axiom or observation, which axiom found out discovereth and designeth new particulars. That the nature of this trial is not only upon the point, whether the knowledge be profitable or no; not because you may always conclude that the Axiom which discovereth new instances be true, but contrariwise you may safely conclude that if it discover not any new instance it is in vain and untrue. (*Works*, vol. 3, p. 242)

It is unclear here whether Bacon is providing a gloss on truth, maintaining that it has been misconstrued, or saying that something is true, in the ordinary accepted sense, only if it is useful. Whichever, it is a very strong claim on Bacon's part. For there are certainly useless truths, just as there are falsehoods that have practical applications. It is not simply that false premises may lead to true conclusions, but there are cases where approximations, although false, may have more practical value than the truths of which they are the approximation.

The solution becomes clear when we consider that, since antiquity, debates on methods of generating truths had hinged on the question of generating informative truths, the aim being to discover something we did not already know. In particular, there was a concern among Aristotle and his Renaissance followers to show that formal modes of reasoning such as the syllogism were not trivial or circular because, at the start of the inferential process, we have knowledge *that* something is the case, whereas at the end of it we have knowledge *why* it is the case. In particular what they sought to show was that the kind of knowledge of an observed phenomenon we have through sensation is qualitatively different from and inferior to the kind of knowledge we have of that phenomenon when we grasp it in terms of its causes.

This latter kind of knowledge is also what Bacon was seeking. If we think in terms of "informative truths," Bacon's position makes more sense. He is saying that the only way in which we can judge whether something is informatively true is to determine whether it is productive, whether it yields something tangible and useful. And if something does consistently yield something tangible and useful, then it is informatively true. The case of approximations can perhaps be dealt with by saying that these derive their usefulness not from their falsity but from their proximity to the truth, although the cases where the approximation is more useful than the true account cannot be handled so easily.

The question of the practicality of truth turns on how informative it is, but there is another dimension to this question that, although not explicitly mentioned by Bacon, is of importance in understanding his general orientation. In the humanist thought that makes up the source from which Bacon derives much of his inspiration, moral philosophy figures prominently. Now in this philosophy, being virtuous and acting virtuously are the same thing: There is no separate practical dimension to morality. This is all the more interesting because moral philosophy is a cognitive enterprise, one in which the practical outcome is constitutive of the discipline, a point Bacon stresses the *Advancement of Learning*. If moral philosophy is the model for natural philosophy, a natural enough conclusion for a humanist and one that is reinforced in the shift from *otium* to *negotium*, then we may

be able to make a little more sense of the idea that truth is not truth unless it is informative and productive.

If we think of Bacon's project as transforming moral philosophers into natural philosophers, then we might expect some carryover from conceptions of the moral philosopher. Notions that were quite appropriate in moral philosophy but not (at least outside Epicureanism) in natural philosophy remain in the transformation process. And this is exactly what we do find, most strikingly in the idea of truth as productive and informative. For Bacon, the truth of natural philosophy hinges as much on its being informative and productive of works as does the truth of moral philosophy in its way. "In religion," he tells us in *Redargutio Philosophiarum*, "we are warned that faith is to be shown by works" (*Works*, vol. 3, p. 576). And he proposes that the same test that is applied in religion be applied in philosophy: if it produces nothing at all, or, worse, if, "instead of the fruits of the grape or olive, it bear the thistles and thorns of disputes and contentions," then we can reject it.

BACON'S LEGACY

In the early modern era, there emerged in the West a style of doing natural philosophy, a way of thinking about the place of natural philosophy in culture generally, and a way of thinking about oneself as a natural philosopher. Bacon played a key role in this development. He inaugurated the transformation of philosophy into science, for even though the ideas of "science" and "scientist" in their modern sense were only really established in the nineteenth century, their genealogy goes back to Bacon's attempt to effect a fundamental reform of philosophy from a contemplative discipline, exemplified in the individual persona of the moral philosopher, to a communal, if centrally directed, enterprise exemplified in the persona of the experimental natural philosopher. Thanks in large measure to Bacon's exertions, observation and experiment were lifted out of the purview of the arcane and the esoteric and planted firmly in the public realm. As a result, science was transformed: Its tradition of irregular fits of progress alternating with long periods of stagnation gave way to the uninterrupted and cumulative growth that has characterized Western science since then.

In defending natural philosophy, Bacon reshaped it; his establishment of its autonomy, legitimacy, and central cultural role are on a par with Plato's defense of the autonomy and centrality of the "quiet" virtues, such as justice and moderation. Both irreversibly changed the cultures in which they lived and those that followed—above all our own.

See also Aristotelianism; Aristotle; Cicero, Marcus Tullius; Epicureanism and the Epicurean School; Ethics, History of; History and Historiography of Philosophy; Induction; Logic, History of; Lucretius; Naturalized Philosophy of Science; Newton, Isaac; Philosophy of Law, History of; Philosophy of Mind; Philosophy of Science, History of; Philosophy of Science, Problems of; Plato; Psychology; Renaissance; Scientific Method; Semantics, History of; Stoicism.

Bibliography

WORKS BY BACON

The Oxford Francis Bacon, edited by Graham Rees and Lisa Jardine. Oxford: Oxford University Press, 1996 onwards. This will eventually replace the Spedding and Ellis edition, as the volumes appear.

The Works of Francis Bacon. 14 vols, edited by James Spedding, Robert Leslie Ellis, and Douglas Denon Heath. London: Longmans, 1857–1874.

WORKS ABOUT BACON

Anderson, Fulton H. *The Philosophy of Francis Bacon*. New York: Octogan Books, 1971. A comprehensive treatment of Bacon's natural philosophy.

Gaukroger, Stephen. *Francis Bacon and the Transformation of Early-Modern Philosophy*. Cambridge, U.K.: Cambridge University Press, 2001. Looks at the connection between Bacon's natural philosophy and his account of method, and at his account of the shaping of the natural philosopher.

Jardine, Lisa. *Francis Bacon: Discovery and the Art of Discourse*. Cambridge, U.K.: Cambridge University Press, 1974. Deals with those issues in theories of rhetoric and logic that were the key ones in Bacon's formative years and shows their influence on his thinking.

Jardine, Lisa, and Alan Stewart. *Hostage to Fortune: The Troubled Life of Francis Bacon, 1561–1626*. London: Victor Gollancz, 1998. The most comprehensive and up-to-date biography.

Martin, Julian. *Francis Bacon, the State, and the Reform of Natural Philosophy*. Cambridge, U.K.: Cambridge University Press, 1992. The best available treatment of Bacon's thought on legal questions.

Peltonen, Markku, ed. *The Cambridge Companion to Bacon*. Cambridge, U.K.: Cambridge University Press, 1996. General collection of articles on all aspects of Bacon's thought.

Pérez-Ramos, Antonio. *Francis Bacon's Idea of Science and the Maker's Knowledge Tradition*. Oxford: Oxford University Press, 1988. Concentrates on the role of rhetoric in Bacon's work.

Rossi, Paolo. *Francis Bacon: From Magic to Science*. Chicago: University of Chicago Press, 1968. Deals with Bacon's natural philosophy.

Urbach, Peter. *Francis Bacon's Philosophy of Science: An Account and Reappraisal*. La Salle, IL: Open Court, 1987. Provides a comprehensive discussion of Bacon's account of method, with regular comparisons with Popper.

Webster, Charles. *The Great Instauration: Science, Medicine and Reform, 1626–1660*. London: Duckworth, 1975. The best available treatment of the influence of Bacon between the time of his death and the founding of the Royal Society.

Zagorin, Perez. *Francis Bacon*. Princeton, NJ: Princeton University Press, 1998. A general overview of Bacon's thought.

Stephen Gaukroger (2005)

BACON, ROGER

(between 1214 and 1220?–1292)

Roger Bacon, English philosopher and scientist, known as *Doctor Mirabilis,* was probably born between 1214 and 1220 and died in 1292, probably at Oxford. Bacon wrote in 1267 that he had learned the alphabet some forty years before and that his once wealthy brother had been ruined by his support of King Henry III during the barons' revolt. He studied arts at Oxford and then at Paris, where as regent master (c. 1237) he was among the first to lecture on the forbidden books of Aristotle when the ban was lifted. Here he wrote his *Summa Grammatica, Summulae Dialectices, Summa de Sophismatibus et Distinctionibus,* his *Quaestiones* on Aristotle's *Physics, Metaphysics,* and *De Sensu et Sensibili,* and on the pseudo-Aristotelian *De Plantis* and *Liber de Causis*; he also wrote commentaries, now lost, on *De Anima, De Generatione et Corruptione, De Caelo et Mundo,* and *De Animalibus.*

These early lectures reveal a philosopher, immature but of unusual ability, conversant with the new literature of Aristotle and his Arabic commentators. They are of some historical interest, since Bacon was representative of the new breed of masters at Paris who prided themselves on being pure Aristotelians. In fact, however, like Avicenna and Gundissalinus before them, they were still strongly influenced by other traditions (especially Neoplatonism) that dominated such apocryphal works as the *Liber de Causis* and, in Bacon's case, the popular *Secret of Secrets.* This latter work, thought to be Aristotle's esoteric instructions to Alexander the Great, is a study in kingcraft which, in addition to advocating a sound, practical philosophy, gives much astrological advice and hints at the magical virtues of herbs and gems and the occult properties of numbers. From his glosses on the book, it seems that Bacon was most impressed by its vision of a universal science of great practical import that included all the secrets of nature. This unified science, revealed by God to the Hebrews, who passed it on through the Chaldeans and Egyptians to Aristotle, was concealed in figurative and enigmatic language but might be rediscovered by one morally worthy and mentally qualified to receive it. Where the pagans failed, Bacon held, a Christian might succeed. Therefore, around 1247 he left Paris, where he had been pursuing a mastership in theology, and returned to Oxford, where Adam Marsh, Robert Grosseteste's Franciscan associate, introduced him to that great man's work. For two decades, Bacon writes, he studied languages and the sciences, training assistants, cultivating the fellowship of savants, and spending more than £2,000 on "secret books," instruments, and tables.

Sometime during the latter half of this period he must have joined the Franciscans, to whom Grosseteste bequeathed his library. Neither his impoverished brother nor the mendicant friars could provide the experimental equipment Bacon longed to have; nor did the majority of the friars share his views on the importance of his work. Resenting the preference shown to the more orthodox theologians, Bacon became embittered and vented his spite in cutting and often unjust criticisms of some of the best minds of the age. Worse, his childlike credulity with regard to the apocalyptic literature of the times led him to side with the extremist followers of Joachim of Floris. This made his views suspect; he was sent to Paris and forbidden to circulate his writings outside the order. But Pope Clement IV, learning of Bacon's proposed encyclopedia of unified science in the service of theology and unaware that the work was largely in the planning stage, wrote for a secret copy on June 22, 1266. Hoping for papal aid to complete the project, Bacon, in the short space of eighteen months, composed as a preliminary draft his *Opus Maius* (synopsized and implemented by the *Opus Minus* and *Opus Tertium,* the latter rich in biographical detail). With the *Opus Maius,* Bacon sent the pope a copy of his *Multiplicatio Specierum,* a concave lens "made at great expense," and "a precious map of the world." Unfortunately, Clement died in November 1268, before the last of the *opera* arrived.

Bacon probably returned to Oxford; he completed his *Communia Mathematica* and *Communia Naturalium* (two of his most mature works) and wrote Greek and Hebrew grammars and his *Compendium Studii Philosophiae.* The last, intended as a general introduction to his principal writings, degenerated into an emotional diatribe against the evils of the age; these were, according to Bacon, especially manifest in the universities where the two teaching orders (Dominicans and Franciscans) were neglecting his favorite subjects. It also revealed a revival of Joachite interests (Bacon referred to the ridicule his

"logical proof" of the imminence of the Antichrist provoked among the friars).

According to the *Chronicle of the Twenty-four Generals,* written in 1370, the Franciscan minister general, Jerome of Ascoli (later Pope Nicholas IV), imprisoned him for "suspected novelties." This account has been questioned, primarily because nothing could be found in Bacon's scientific or astrological views that had not been endorsed by many reputable theologians of the day, such as Albertus Magnus. More likely, it was a political move to silence the irascible friar, whose caustic views on the morals of the secular masters would do little to ease the strained relations between them and the friars (whose orthodoxy had been seriously compromised by the fanatical Joachite fringe). At any rate, Bacon's confinement could hardly have been rigorous or long enough to inhibit his penchant for frank expression; in 1292 he was writing in the *Compendium Studii Theologiae* on his favorite topics with all his old verve and biting invective. He died, however, before this work was completed.

THOUGHT

The strength and the weakness of Bacon's erratic genius are nowhere more apparent than in the *Opus Maius,* his most characteristic and distinctive work. Both a plea and a plan for educational reform along the study lines pursued by Bacon himself, it is divided into seven parts—the causes of error, philosophy, the study of languages, mathematics, optics, experimental science, and moral philosophy. The first part descries four barriers blocking the road to truth: submission to unworthy authority (for example, crediting living theologians with a prestige due only to the Church Fathers or the Scriptures), the influence of custom, popular prejudice, and concealment of one's ignorance with a technical show of wisdom. Although by far the greatest portion of the book is devoted to mathematics, optics, and moral philosophy (to which, Bacon claimed, all speculative science should be ordered), Bacon's fame until recently rested on this first part and the relatively short section on experimental science. The belief that experimental science was the keystone of Bacon's reform was in part based on the misleading evidence of Samuel Jebb's 1733 edition of the *Opus Maius,* which omitted Part VII. By *scientia experimentalis,* however, Bacon meant any knowledge through experience as opposed to inferential or reasoned knowledge. When he said that nothing can be known with certainty without experience, his use of the term *experience* was twofold. One aspect of experience is based on sense perception and is called human or philosophical; the other aspect is interior and is derived from an illumination of the mind by God (whom Bacon identified with Aristotle's agent intellect). Thus, although sense perception is necessary to knowledge, certainty cannot be attained without divine illumination. Interior experience admits of seven degrees, beginning with that required for certitude in mathematics or the natural sciences and culminating in such mystical or ecstatic states as St. Paul's vision of heaven.

Bacon devoted the most attention, however, to what humans can know about the wonders of nature by sense perception and the first degree of illumination. From the examples cited in Part VI and throughout the work, Bacon seems to have been less an original experimenter and more a propagandist for scientists such as Peter of Maricourt. His contributions to scientific theory, like his empirical research, were confined largely to optics. With the aid of new source material from Alhazen and Abū-Yūsuf Yaʿqūb ibn Ishāq al-Kindī, he was able to develop significantly many of Grosseteste's views concerning the tides, heat, and double refraction and to give the most mature expression to Grosseteste's theory that light (and all physical force generally) is transmitted in pulses like sound waves. Since this "multiplication of species" requires a medium, Bacon argued, the transmission cannot be instantaneous, even though the time interval is imperceptible. His application of the theory to vision and the working of the eye was one of the most important studies done on this subject during the Middle Ages and became the point of departure for developments in the seventeenth century. Bacon seems to have surpassed his teachers both in his knowledge of convex lenses and parabolic mirrors and in his ability to foresee such applications of science as automobiles, motorboats, and aircraft.

If, by continuing the Oxford tradition begun by Grosseteste, Bacon was in advance of his contemporaries, he was also incredibly naive in some of his other views. His uncritical acceptance of what others claimed to have observed is often in violation of his own canons for avoiding error. Much of his stress on the importance of language studies came from his conviction that all knowledge can be found in the Scriptures and "secret books," whose full meaning God reveals by interior illumination only to those whose lives are pure. He held that because of men's sins, God's scientific revelations were obscured by errors—which is one reason for testing empirically what the ancient sages say. Bacon seems to have had little use for abstract reasoning or speculation for its own sake. His interest in mathematics and logic, like his interest in astrology and alchemy, was purely practical. If all physical force, like light, is propagated rectilinearly, it is subject to

geometric analysis. This, together with his conviction that the movement of the planets influences all terrestrial events except free will itself, was his reason for thinking that mathematics is the key to all natural sciences.

Not only was his faith in astrology unwarranted, but his ideas of theology belonged to a bygone age. Even prior to 1250, the Paris Franciscans, impressed by the Euclidean-Aristotelian ideal of a deductive science, were exploring how far the concepts of theology might be analyzed with greater logical rigor and theological propositions formalized in terms of axioms (first principles of reason and philosophy), postulates (the articles of faith), and theses (theological conclusions). Despite his sporadic attendance at theological lectures, Bacon seems to have had no comprehension of what the avant-garde theologians were doing. Perhaps this, more than any insistence on scientific values or the need for experimentation, brought him into conflict with his educated confreres, who apparently considered him, for all his flashes of brilliance and his scientific lore, something of a crank.

See also Albert the Great; Aristotle; Avicenna; Ethics, History of; Grosseteste, Robert; Joachim of Fiore; Neoplatonism.

Bibliography

Fr. Rogeri Bacon Opera Quaedam Hactenus Inedita, edited by J. S. Brewer (London: Longman, Green, Longman and Roberts, 1859), contains *Opus Tertium, Opus Minus, Compendium Studii Philosophiae,* and *Epistola de Secretis Operibus et de Nullitate Magiae.* For supplements to *Opus Tertium,* see Pierre Duhem, *Un Fragment inédit de l'Opus tertium de R. Bacon* (Quaracchi: Ad Claras Aquas, 1909), and A. G. Little, *Part of the Opus Tertium of Roger Bacon* (Aberdeen: University Press, 1912). Also see J. H. Bridges, ed., *Opus Majus,* 3 vols. (Oxford: Clarendon Press, 1897–1900; reprinted Frankfurt am Main: Minerva, 1964), Vol. II of which contains *Multiplicatio Specierum* as an appendix. For the complete text of Part VII, which is incomplete in Bridges, see F. Delorme and E. Massa, *Rogeri Baconis Moralis Philosophia* (Zürich, 1953).

Other writings by Bacon may be found in F. A. Gasquet, "An Unpublished Fragment of a Work by Roger Bacon," in *English Historical Review* 12 (1897): 494–517, a prefatory letter to the *Opus Maius* or *Opus Minus* or both; E. Nolan and S. A. Hirsch, *The Greek Grammar of Roger Bacon and a Fragment of His Hebrew Grammar* (Cambridge, U.K.: Cambridge University Press, 1902); and H. Rashdall, ed., *Fratris Rogeri Bacon Compendium Studii Theologiae* (Aberdeen: Typis Academicis, 1911). Most of Bacon's remaining works have been published by R. Steele (with individual volumes by F. M. Delorme, A. G. Little, and E. Withington) in *Opera Hactenus Inedita Fratris Rogeri Baconis,* 16 fascicles (Oxford: Clarendon Press, 1905–1940). The *Compotus* (Fasc. 6), ascribed to Bacon, is, however, the work of Giles of Lessines. See also S. H. Thomson, "An Unnoticed Treatise by Roger Bacon on Time and Motion," in *Isis* 27 (1937): 219–224, and F. M. Delorme, "Le Prologue de R. Bacon à son traité De influentiis agentium," in *Antonianum* 18 (1943): 81–90.

English translations of Bacon's work are R. B. Burke, *The "Opus Majus" of Roger Bacon,* 2 vols. (Philadelphia: University of Pennsylvania Press, 1928; reprinted, New York, 1962), and T. L. Davis, *Roger Bacon's Letter concerning the Marvelous Power of Art and Nature and concerning the Nullity of Magic* (Easton, PA: Chemical Publishing, 1923). For a bibliography of Bacon's works, see F. Alessio, "Un seculo di studi su Ruggero Bacone (1848–1957)," in *Rivista critica di storia della filosofia* 14 (1959): 81–108.

A discussion of the life and works of Roger Bacon appears in the introduction and appendix of A. G. Little, *Roger Bacon Essays, Contributed by Various Authors on the Occasion of the Commemoration of the Seventh Centenary of His Birth* (Oxford: Clarendon Press, 1914). Also see S. C. Easton, *Roger Bacon and His Search for a Universal Science* (New York: Columbia University Press, 1952), which contains an extensive and annotated bibliography. For a discussion of Bacon's philosophy, see T. Crowley, *Roger Bacon, the Problem of the Soul in His Philosophical Commentaries* (Louvain and Dublin: Éditions de l'Institut supérieur de philosophie, 1950), a good account of hylomorphic theory; D. Sharp, *Franciscan Philosophy at Oxford* (Oxford: Oxford University Press, 1930); and E. Heck, *Roger Bacon. Ein mittelalterlicher Versuch einer historischen und systematischen Religionswissenschaft* (Bonn, 1957). For Bacon's contributions to science, see Lynn Thorndike, *A History of Magic and Experimental Science* (New York: Macmillan, 1929), Vol. II, pp. 616–691, which minimizes Bacon's contributions (a reaction to earlier exaggerations); and A. C. Crombie, *Robert Grosseteste and the Origins of Experimental Science 1100–1700* (Oxford: Clarendon Press, 1953), pp. 139–162, a more balanced account.

Allan B. Wolter, O.F.M. (1967)

BACON, ROGER [ADDENDUM]

Twentieth-century research on Roger Bacon requires some changes to the account above. It is clear that Richard Rufus, and not Bacon, was the first to lecture on the new Aristotle at Paris circa 1235. Bacon responded to the ideas of Rufus in his Parisian *Quaestiones* (c. 1240s). He returned to these topics in his last work *Compendium of the Study of Theology* (1292).

Sometime around 1247, Bacon departed from his teaching at the University of Paris. For the next twenty years he devoted his time to a study of the following works: Ibn al-Haytham *Optics*, the Pseudo-Aristotelian *Secretum secretorum* on statecraft, the *Centiloquium*, the *Commentary on the Centiloquium*, and numerous works on astrology. Most important here was the work of Abu'-

mashar (Albumassar). The *Communia mathematica*, the *Communia naturalium*, and the *Compendium studii philosophiae* were most probably written in Paris.

Research on Bacon since the mid-twentieth century has yielded the following results:

1) Bacon plays a significant role in the history of logic, semantics, and semiotics. Bacon's originality stands out in regard to semiotics, philosophical grammar, quantification, theory of natural sense, univocity, and supposition.

2) The new editions of the *De multiplicatione specierum* (1266) and the *Perspectiva* (1266) have placed these two texts in their proper context as important works in natural philosophy and philosophy of mind.

3) Scholars have gained a greater understanding of Bacon's aims in his knowledge of mathematics, astronomy-astrology, music, experimental science, alchemy, and medicine. Bacon presents himself as an advocate for the experimental science of others such as Petrus Peregrinus of Maricourt. Nevertheless, his account of *Perspectiva* as a model of an "experimental science" is fundamentally important for the later development of optics, perspective, and philosophy of mind, and for methodology in science.

4) Bacon's treatise on *Moralis philosohpia* develops proto-humanist concerns. Overall, in his later post-1266 philosophy, Bacon subordinates his earlier Aristotelianism to a Stoic division of philosophy and to mainly Platonic concerns.

See also Aristotelianism; Aristotle; Logic, History of; Philosophy of Mind; Platonism and the Platonic Tradition; Semantics; Stoicism.

Bibliography

WORKS BY ROGER BACON

Fredborg, Karen Margarita, Lauge Nielsen, and Jan Pinborg, eds. "An Unpublished Fragment of a Work by Roger Bacon." *Traditio* 34 (1978): 75–136.

Lindberg, David C. *Roger Bacon and the Origins of Perspectiva in the Middle Ages: A Critical Edition of the "Perspectiva," with English Translation, Introduction, and Notes*. New York: Oxford University Press, 1995.

Lindberg, David C. *Roger Bacon's Philosophy of Nature: A Critical Edition, with English Translation, Introduction, and Notes, of "De Multiplicatione specierum" and "De speculis comburentibus."* Oxford, U.K.: Clarendon Press, 1983.

Linden, Stanton J. *The Mirror of Alchimy*. New York: Garland, 1997.

Maloney, Thomas S. *Compendium of the Study of Theology*. Leiden: Brill, 1988.

Maloney, Thomas S. *Three Treatments of Universals*. Binghamton: MARTS, 1989.

Molland, George. "Roger Bacon's *Geometria speculative*." In *Vestigia Mathematica: Studies in Medieval and Early Modern Mathematics in Honor of H.H.L. Busard*, edited by M. Folkerts and J. P. Hogendijk. Amsterdam and Atlanta: Editions Rodopi, 1993.

WORKS ABOUT ROGER BACON

Hackett, Jeremiah, ed. *Roger Bacon and The Sciences: Commemorative Essays*. Leiden: Brill, 1997. This volume introduced new work on languages and the natural sciences.

Hackett, Jeremiah, ed. *Vivarium: An International Journal for the Philosophy and Intellectual Life of the Middle Ages and Renaissance* Leiden: Brill, Vol. XXXV, N. 2 (September, 1997). This volume introduces new work on language, physics, metaphysics, and philosophy of mind. It also includes a bibliography.

Rosier-Catach, Irene. *La parole comme acte: Sur la grammaire et la semantique au XIIIe siecle*. Paris: Vrin, 1994. In this work, the author situates Bacon in the context of philosophy of language in the thirteenth century.

Jeremiah Hackett (2005)

BAD FAITH

The most common form of inauthenticity in the existentialism of Jean-Paul Sartre, "bad faith" is paradoxically a lie to oneself. For such self-deception to be possible, the human being must be divided against itself, one level or aspect concealing from the other what it in some sense "knows." The paradox arises from the condition that this operation occurs within the unity of a single consciousness.

The root of Sartrean bad faith is a twofold dividedness of the human being, psychological and ontological. As conscious, humans are prereflectively aware of what they may not reflectively know. Such prereflective awareness or "comprehension," as he will later call it, functions in Sartre's psychology in a manner similar to Sigmund Freud's unconscious, a concept that Sartre notoriously rejected. The project of bad faith—to keep oneself in the dark about certain matters—is itself in bad faith since prereflective consciousness "chooses" not to acknowledge on reflection what it is concealing from reflective consciousness.

There can be an entire Weltanschauung of bad faith: the habits, practices, objects, and institutions that one employs to maintain oneself in a state of "perpetual distraction." Sartre's analysis of Second Empire French society in his work on Gustave Flaubert is a study in collective

bad faith. But the root of the moral responsibility that this term carries lies in the self-translucency of prereflective consciousness: individuals, alone or together, are prereflectively aware of more than they reflectively allow themselves to know.

The ontological basis of bad faith is the dividedness of the human situation. Every human exists in-situation. Situation is an ambiguous mix of facticity (the given) and transcendence (the surpassing of the given by our projects). Bad faith is our way of fleeing the anguish that this ambiguity causes either by collapsing our transcendence into facticity (as in various forms of determinism) or by volatilizing our facticity into transcendence (like the dreamer who refuses to acknowledge the facts of his or her life). Though the details of bad faith are as singular as our self-defining choices, its moral significance is the same in each instance. Bad faith is basically flight from our freedom-in-situation.

As Sartre's concept of situation expanded to include and even place a premium on socioeconomic conditions, the relation between bad faith and class struggle became more pronounced. He later argued that good faith, which in *Being and Nothingness* he dismissed as a form of bad faith, was fostered by socioeconomic equality and that scarcity of material goods made bad faith almost inevitable. The anti-Semite was in bad faith, but so too was his or her liberal assimilationist defender; likewise the neocolonialist and the industrial capitalist, both of whom fled their responsibility for subscribing to and sustaining a system that made exploitation of others "necessary."

Only in his posthumously published *Notebooks for an Ethics* does Sartre discuss the nature and possibility of good faith at any length. This presumes a "conversion" in which one chooses to live one's anguished dividedness while fostering via generous cooperation a situation that enables others to do likewise.

See also Determinism, A Historical Survey; Existentialism; Existential Psychoanalysis; Freud, Sigmund; Sartre, Jean-Paul; Unconscious.

Bibliography

Beauvoir, S. de. *Pour une morale de l'ambiguité.* Paris: Gallimard, 1947. Translated by B. Frechtman as *The Ethics of Ambiguity.* New York: Philosophical Library, 1948.

Fingarette, H. *Self-Deception.* London: Routledge and Kegan Paul, 1969.

Martin, M. W. *Self-Deception and Morality.* Lawrence: University Press of Kansas, 1986.

Mirvish, A. "Bad Faith, Good Faith, and the Faith of Faith." In *Sartre Alive,* edited by R. Aronson and A. van den Hoven. Detroit: Wayne State University Press, 1991.

Morris, P. "Self-Deception: Sartre's Resolution of the Paradox." In *Jean-Paul Sartre: Contemporary Approaches to His Philosophy,* edited by H. J. Silverman and F. A. Elliston. Pittsburgh: Duquesne University Press, 1980.

Santoni, R. E. *Bad Faith, Good Faith and Authenticity in Sartre's Early Philosophy.* Philadelphia: Temple University Press, 1995.

Sartre, J.-P. *Cahiers pour une morale.* Paris, 1983. Translated by D. Pellauer as *Notebooks for an Ethics.* Chicago: University of Chicago Press, 1992.

Sartre, J.-P. *L'etre et le néant.* Paris: Gallimard, 1943. Translated by H. E. Barnes as *Being and Nothingness.* New York: Philosophical Library, 1956.

Thomas R. Flynn (1996)

BAHRDT, CARL FRIEDRICH

(1740 or 1741–1792)

Carl Friedrich Bahrdt, probably the most widely read German theologian except for Martin Luther, was born in Bischofswerda in the electorate of Saxony. He held professorships and lectureships of theology, biblical studies, Christian ethics, classical languages, and many other subjects at the universities of Leipzig, Erfurt, Giessen, and Halle. He was the headmaster of a boys school, or *Philanthropinum,* in Marschlins in Switzerland and established his own *Philanthropinum* in Heidesheim while he was at the same time *Superintendent* (the highest ecclesiastical official) in the domains of Count Carl of Leiningen-Dachsburg. In his last years, he was an innkeeper near Halle. He died at Halle.

Bahrdt was always at the center of a controversy. In his early days he wrote in a fiery orthodox vein, but very soon he seems to have been started on the road to "enlightenment" by suddenly learning that the language of I John 5:7, did not, when subjected to philological scrutiny, constitute proof of the doctrine of the Trinity. He was still further dismayed to learn that the passage was considered by some excellent scholars to be an interpolation. Bahrdt then set out to find undoubted philological support for the orthodox Lutheran system of theology, and instead found that his doubts continued to increase, until by the end of his life he had arrived at a fully rationalistic concept of natural religion.

The high points in Bahrdt's "Rationalist's Progress" are his four-volume paraphrase of the New Testament, *Neueste Offenbarungen Gottes* (Riga, 1773–1774), his

confession of faith, *Glaubensbekenntnis, veranlasst durch ein Kaiserliches Reichshofratsconclusum* (1779), and his fictionalized life of Jesus, *Briefe über die Bibel im Volkston* (Halle, 1782–1783) and *Ausführung des Plans und Zweckes Jesu* (Berlin, 1783–1785). Bahrdt's New Testament paraphrase was up-to-date, intelligible, fluent, and coherent, but it was also a propagandistic vehicle for his heretical views. His enemies were thus enabled to secure, in 1778, a decree barring him from all ecclesiastical offices in the Holy Roman Empire and adjuring him to recant. Bahrdt immediately published his confession of faith, stating in clear and succinct language what he did and did not believe. Through discarding beliefs that he felt could not endure the acid test of rational examination, Bahrdt was left with a Jesus who was a mere product of his life and time. In this almost completely naturalistic view, the teasing question was, "In what way did Jesus obtain his amazing wisdom?" In order to give a hypothetical answer to this question, Bahrdt produced his fictional life of Jesus, the culmination of his development and the first work of its kind. It took the form of a series of weekly letters about the Bible, written in a popular vein, and tried to demonstrate how Jesus might have learned and built up his teachings from the writings of Greek sages, which Providence could have put into his hands through his association with Hellenistic Jews. These first letters were continued in a series on the execution of Jesus' plan and purpose, in which Bahrdt advanced the theory that Jesus founded a kind of Freemasonry to aid him in his purpose to destroy superstition, eliminate all positive religion, restore reason to its rightful rule, and unite people in a rational faith in God, Providence, and Immortality.

See also Luther, Martin; Rationalism.

Bibliography

Bahrdt, C. F. *Dr. Carl Friedrich Bahrdts Geschichte seines Lebens, seiner Meinungen und Schicksale.* Berlin: Bei Friedrich Vieweg, dem Älteren, 1790–1791.

Brewer, John T. "*Gesunde Vernunft" and the New Testament: A Study of C. F. Bahrdt's* Die neuesten Offenbarungen Gottes. PhD diss., University of Texas at Austin, 1962.

Flygt, S. G. *The Notorious Dr. Bahrdt.* Nashville, TN: Vanderbilt, 1963.

Schweitzer, Albert. *Geschichte der Leben-Jesu-Forschung.* Tübingen: J.C.B. Mohr, 1913.

Sten G. Flygt (1967)

BAḤYĀ BEN JOSEPH IBN PAQŪDA

(fl. 11th century)

Baḥyā ben Joseph ibn Paqūda, the Jewish Neoplatonist, was the author of the first systematic philosophic work on ethics in the Jewish tradition. Beyond the fact that he served as a judge (*dayyan*) of the rabbinical court in Saragossa, details of his life are unknown. About 1040 he wrote in Arabic *Al-Hidaja ila Faraid al-Qulub* (Guide to the duties of the heart). This work, as translated into Hebrew about 1160 by Judah ibn Tibbon, under the title *Hoboth Ha-Lebaboth* (*Duties of the Heart*), has achieved great popularity, both in full text and in abridged versions.

Baḥyā's work cites Arabic as well as Jewish philosophers and contains many fine quotations from Arabic literature. There are considerable similarities between his general philosophic orientation and that of the Arabic school of encyclopedists known as the Brothers of Purity. If this relationship is accepted, there is no need to search further for the sources of the somewhat mystical, somewhat ascetic Neoplatonism that moderates the generally Aristotelian character of his position. It has also been suggested that Baḥyā fell under the influence of the Sufi mystics of Islam, chiefly because of his emphasis on the cultivation of self-renunciation and indifference to the goods of the world in the last three books of *Duties of the Heart.*

The distinction between outward and inward obligation, "duties of the limbs" and "duties of the heart," which accounts for the title of the treatise, is a familiar distinction in both Arabic and Hindu religious literature. Baḥyā used the theme to suggest that the rabbis, the leaders of the Jewish community, were overly concerned with the external obligations of men, rather than with the duties of the heart, and that, because of the rabbis' insistence on the duties of the limbs, the masses of the Jewish people remained totally unconcerned about all religious obligations. He tried to correct this deficiency by presenting Judaism as a message of great spiritual vitality and force, directed to the human heart and resting on the threefold base of reason, revelation, and tradition. The fundamental principle upon which the whole structure of Baḥyā's work is based is the wholehearted conviction of God's existence and unity, the subject of the first book of *Duties of the Heart.* From this, he moves to the necessity for apprehending the wisdom, power, and goodness of God by careful study of the larger world in which we live and the smaller world of our own human nature. In this latter

study there emerge the duties of the heart: service of God, trust in God, wholehearted devotion to God, humility in God's presence, repentance, self-communion, and renunciation. In this way, humanity reaches the height of the religious life, the love of God. Despite the superficially rational structure of the book, Baḥyā was not truly a rationalist; rather, he used the techniques of reason to subserve the ends of a contemplative view of life whose method was moral intuition, and whose goal was piety.

An Arabic treatise, *Maʿani al-Nafs* (The attributes of the soul), known only in manuscript until its publication in the early twentieth century, bears the name of Baḥyā on its title page, but this is now generally conceded not to be his work. No other works of Baḥyā are known.

See also Jewish Philosophy; Neoplatonism.

Bibliography

For Baḥyā's work, see *Torath Hoboth Ha-Lebaboth,* 5 vols. (New York: Bloch, 1925–1947). This contains Judah ibn Tibbon's Hebrew translation plus a facing English translation and an introduction by Moses Hyamson.

For discussions of Baḥya, see Isaac Husik, *A History of Mediaeval Jewish Philosophy* (New York: Macmillan, 1916); Jacob B. Agus, *The Evolution of Jewish Thought* (London and New York: Abelard-Schuman, 1959); and Joseph L. Blau, *The Story of Jewish Philosophy* (New York: Random House, 1962). See G. Vajda, *La théologie ascétique de Bahja ibn Paquda* (Paris: Nationale, 1947), for a comparison of Baḥya's doctrines with Islamic ascetic literature.

J. L. Blau (1967)

BAḤYĀ BEN JOSEPH IBN PAQŪDA [ADDENDUM]

Although Baḥyā ben Joseph ibn Paqūda follows the major categories of Sufism in his exploration of human motivation, he also manages to find a social justification for many aspects of Judaism. For example, one of the virtues he discusses is restraint or abstemiousness, the need to resist our desires. He argues that we can pursue this socially by our attitude to others by acquiring a cheerful and calm attitude toward others. A means of being disposed to act thus is the Torah and its laws, for these have the effect of training ourselves to restrain our desires and bring them under the rule of law. For Baḥyā the very private and personal moral rules that we adopt to bring us closer to God have a significant public element. The highest virtue is love of God, and to acquire this we need to practice personal asceticism, together with justice, good manners, and justice. Although the aim of his book on the duties of the heart is to show that Judaism is not only about external actions but has an inner spiritual dimension as well, he does not go to the extreme of denying the significance of law and prayer. On the contrary, he argues that the private and the public aspects of religion complement each other. His book also provides a detailed account of how that works in the case of Judaism.

See also Asceticism; Jewish Philosophy; Justice; Moral Rules and Principles; Sufism.

Bibliography

Fenton, Paul B. "Judaism and Sufism." In *Cambridge Companion to Medieval Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman, 201–217. New York: Cambridge University Press, 2003.

Rudavsky, Tamar. "Medieval Jewish Neoplatonism." In *History of Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman. London: Routledge, 1997.

WORKS BY BAḤYĀ

Al-Hid āya il ā faraʾ iḍ al-Qul ūb (The book of direction to the duties of the heart). Translated by Menahem Mansoor. London: Routledge, 1973.

"Prayer and Faith." In *Jewish Philosophy Reader*, edited by Daniel H. Frank, Oliver Leaman, and Charles H. Manekin. London: Routledge, 2000.

WORKS ON BAḤYĀ

Fenton, Paul B. "Judaism and Sufism." In *History of Islamic Philosophy*. 2 vols., edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996.

Goodman, Lenn E. "Bahya on the Antinomy of Free Will and Predestination." *Journal of the History of Ideas* 44 (1) (1983): 115–130.

Oliver Leaman (2005)

BAIER, ANNETTE

(1929–)

Annette Baier was born in New Zealand in 1929. She received her bachelor of arts and master of arts degrees from the University of Otago, and, in 1954, her bachelor of philosophy degree from Oxford, writing a thesis on precision in poetry under J. L. Austin. After teaching in the United Kingdom, New Zealand, and Australia, Baier moved to the United States, teaching first at Carnegie Mellon and then at the University of Pittsburgh from 1973 until her retirement as Distinguished Service Professor in 1997.

Baier's primary commitment is to naturalism: Human beings are evolved animals and we must under-

stand our capacities, both intellectual and moral, in the light of this natural history. Baier finds philosophers guilty of a kind of willful forgetting of the facts of our embodied existence. We are social animals who experience long periods of dependency in infancy and childhood, and even the more or less symmetric dependencies of maturity are liable to become asymmetric with age or, in some cases, illness. Baier's work charts the implications of our interdependency for epistemology, ethics, and action theory.

Epistemology is a social enterprise. In David Hume, Baier finds the resources to develop a feminist epistemology that recognizes the positive contribution of emotions to knowledge, and that recognizes that all inquiry is fallible and situated, beginning, as it must, from the "prejudices" of tradition and custom. Beliefs, attitudes, and practices that withstand reflective scrutiny merit continued allegiance; those that do not must be abandoned. Baier's account of reflection is distinctive for both its anti-intellectualism and anti-individualism. Reflection is carried out by a *community* of inquirers embracing many differing perspectives and, rather than being the sole province of intellect, reflection uses *all* the capacities of the human mind, including affective capacities such as sympathy. These capacities are capable of being turned on themselves and on our habits and customs and we can come to achieve "reflective self-acceptance, agreement with ourselves" (1994b, p. 277). Reflection reveals the importance of *judgment*. Rules are of limited use in guiding either practical or theoretical judgment; hence Baier's anti-theory stance. In ethics, this anti-theory stance takes the form of suspicion about the possibility of capturing morality in a set of rules. Such systemizing drives are to be replaced by careful exploration of the capacities that enable virtuous action.

In keeping with her emphasis on reflection, Baier proposes a reflective test for evaluating moralities: "a decent morality will *not* depend for its stability on forces to which it gives no moral recognition. Its account books should be open to scrutiny and there should be no unpaid debts, no loans with no prospect of repayment" (1994a, p. 8). Baier argues that liberal morality, with its focus on contractual relations and voluntarily assumed obligations, takes as paradigmatic the interactions between equals or near-equals and so is unable to pass this test. It depends on the unacknowledged moral labor of those producing future moral agents, a labor it cannot itself theorize. Had ethical theory begun from the perspective of those, chiefly women, engaged in such labor, relations between unequals would have come into focus, thus revealing the importance of trust.

Baier's work is largely responsible for the recent upsurge of interest in trust, not just among philosophers, but also among social scientists. She finds trust to bridge the traditional divisions between the cognitive, affective and conative: Trust has a distinctive feel, typically involves a tacit belief in the other's goodwill and competence, and explains the truster's willingness to let others get dangerously near things she cares about. According to Baier, trust, though instrumental to many human goods and a constitutive part of others (for example, friendship), is not a virtue. Nor is untrustworthiness always a vice: Misplaced trust enables exploitation and abuse and sometimes trust is best responded to with judicious betrayals of trust.

Our interdependence also has implications for our understanding of persons and their actions. We are inducted into the "arts of personhood" by others: "Persons essentially are *second* persons who grow up with other persons" (1985, p. 84). It is through being addressed and addressing other second persons—through, that is, coming to master the pronoun "you"—that we come to have self-consciousness. Baier rejects as reductive moves to identify bodily movements or volitions as "basic actions" (actions that are directly done rather than done by doing anything else) and argues that actions can be identified as intentional only given background assumptions of culturally dependent competences. She finds accounts of personhood that focus on a narrow range of properties such as autonomy, dignity, and the capacity to make evaluative judgments guilty of wilfully forgetting our biological nature. She substitutes in their stead a conception of ourselves as "intelligent, talkative, playful mammals" (1991, p. 13) whose personhood comprises many capacities, both cognitive and affective. All these capacities are to be recruited in doing philosophy, which, following Hume, is to use "*all* the capacities of the human mind: memory, passion and sentiment as well as a chastened intellect" (1994b, p. 1). Her own writing style, with its rich use of anecdote, association, playfulness, and irony, enacts as well as argues for a philosophy informed by passion and experience.

See also Analytic Feminism; Austin, John Langshaw; Emotion; Feminist Epistemology; Hume, David; Metaethics; Naturalism; Women in the History of Philosophy.

Bibliography

WORKS BY ANNETTE BAIER

Postures of the Mind: Essays on Mind and Morals. Minneapolis: University of Minnesota Press, 1985.

"A Naturalist View of Persons." Presidential Address Delivered before the Eighty-Seventh Annual Eastern Division Meeting of the American Philosophical Association in Boston, MA. *Proceedings and Addresses of the American Philosophical Association* 65 (1991): 5–17.

Moral Prejudices. Cambridge, MA: Harvard University Press, 1994a.

A Progress of Sentiments: Reflections on Hume's Treatise. Cambridge, MA: Harvard University Press, 1994b.

The Commons of the Mind: Carus Lectures 19. Chicago: Open Court, 1997.

WORKS ABOUT ANNETTE BAIER

Jenkins, Joyce, Jennifer Whiting, and Christopher Williams, eds. *Persons and Passions: Essays in Honor of Annette Baier*. Notre Dame, IN: University of Notre Dame Press, 2005.

Karen Jones (2005)

BAIER, KURT

(1917–)

Kurt Baier was born in Vienna, Austria, in 1917. He had to abandon his law studies at the University of Vienna in 1938, when he went as a refugee to Britain. There he was interned as a "friendly enemy alien" and sent to Australia. He began his study of philosophy in earnest in the internment camp and continued after the war ended. He received his BA (1944) and MA (1947) from the University of Melbourne, and his DPhil (1952) from Oxford University. He taught at the University of Melbourne, the Australian National University, and the University of Pittsburgh. He was a visiting professor at Cornell University, the University of Illinois, the University of Florida, and the University of Otago (New Zealand). He was president of the Eastern Division and chairman of the National Board of Officers of the American Philosophical Association. Annette Baier, whom he married in 1958, was also president of the Eastern Division. After they retired, they moved to New Zealand, which is Annette's native country. They may be the most distinguished philosophical couple in American philosophy, although neither was born in America. Both gave the Paul Carus Lectures, and both were invited to be members of the American Academy of Arts and Sciences. In 2001 Kurt was awarded an Honorary Doctorate of Jurisprudence from the Karl Franzen University of Graz, at a ceremony hosted by the University of Otago.

Baier was one of the most influential philosophers in the field of moral philosophy in the second half of the twentieth century. He is one of the philosophers primarily responsible for returning the field of moral philosophy from an obsession with the language of moral judgments to its traditional concern with describing and justifying guides to moral behavior.

Baier claims that moral rules are meant for everybody. They must be universally teachable, that is, they cannot involve beliefs or concepts not known to all normal adult humans. They cannot be self-frustrating, self-defeating, or morally impossible, that is, impossible or pointless if universally taught. Many moral philosophers after Baier have used these features as necessary conditions for a guide to conduct to count as a morality.

Baier recognizes that these features are merely formal and that moral rules must also have a particular kind of content. Baier describes this content by saying that moral rules must be for the good of everyone alike. However, when he gives examples of these rules (e.g., rules prohibiting killing, cruelty, inflicting pain, maiming, torturing, deceiving, cheating, rape, and adultery), it is quite clear that he means that these rules prohibit causing harm to anyone. He was prescient in recognizing, against both deontologists and utilitarians, that morality does not require doing the optimific act (the act having the best consequences), no matter how one determines what that optimific act is.

Like Thomas Hobbes, whom he acknowledges as a strong influence on his views, Baier put forward the principle of reversibility (a negative version of the Golden Rule), "Do not do unto others as you would not have them do unto you," as summarizing the moral guide to life. Although he does not use the language of natural-law theories, Baier also follows Hobbes in holding that morality has to be known by all those who are held morally responsible for their behavior, that is, moral rules apply to all who can understand the rules and can guide their behavior accordingly.

Baier argues, "It is the very meaning of 'a morality' that it should contain a body of moral convictions which can be true or false, that is, a body of rules or precepts for which there are certain tests" (Baier 1965, p. 89). Baier claims that these tests must involve what he calls "the moral point of view." Although Baier's description of this point of view is not universally accepted, it is acknowledged by all that moral rules must stem from a point of view based on universally shared beliefs and desires.

In addition to providing a plausible and influential account of morality, Baier also put forward an account of rationality that is more acceptable than the standard instrumentalist accounts. He recognizes that it is irrational "when, for no reason at all, we set our hands on fire or cut off our toes one by one" (Baier 1965, p. 158). Unlike many contemporary philosophers, he is aware that there are irrational desires, and hence that it cannot be correct to define a rational action as one that maximizes the satisfaction of a person's desires.

Baier's attempt to use his analyses of the concepts of rationality and morality to arrive at substantive moral conclusions marked the end, in ethics, of a concern with the language of morals that claimed to be morally neutral. By making a distinction between moral judgments and other value judgments, he showed that the terms "right," "ought," "good," and "bad" are primarily related to values, not morality. Recognizing that we offer reasons for choosing and doing many things in addition to those related to morality, Baier convinced many that concentrating on the use of these terms is not likely to be of much help in determining what morality is. Although many contemporary moral philosophers, especially consequentialists, continue to talk of good and bad, right and wrong, it is now generally recognized that these concepts are not identical to the concepts of morally good and morally bad, morally right and morally wrong.

Throughout his work Baier has attempted to show that reason supports acting morally. In his earlier work he distinguished between self-interested reasons, altruistic reasons, and moral reasons; and argued that although self-interested reasons were stronger than altruistic reasons, moral reasons were stronger than self-interested reasons. He showed that anyone picking worlds to live in would pick a world that had this ordering. In his later work, he distinguished between self-interested reasons, self-anchored reasons, and society-anchored reasons, and showed that if a society is to function, its members must accept that society-anchored reasons, particularly moral reasons, overrule both self-interested and self-anchored reasons. Although there is considerable doubt about whether Baier has shown that reason supports morality as he argues for it, his arguments for this view contain many valuable points. Failure to appreciate his distinction between altruistic reasons and moral reasons explains why some people find it difficult to accept that lying to protect a guilty colleague is immoral.

Largely because of Baier's work, moral philosophy no longer is dominated by concerns about the language of ethics. At the beginning of the twenty-first century, moral philosophers are now more likely to put forward substantive ethical views, be they Hobbesian, Kantian, or utilitarian, than they are to view their accounts of morality as having no normative implications. The distinction between concern with analyzing the terms or concepts involved in moral discourse and concern with substantive moral problems has largely disappeared. Even those concerned with analyses of ethical concepts now hold that analyses of these moral concepts may yield substantive moral conclusions. Baier is also primarily responsible for the fact that the central problem of moral philosophy is now showing the relationship between rationality and morality. The mark of a great philosopher is generally thought to lie not in the answers he gives but in the questions he raises. There is no question that on this view Kurt Baier is a great philosopher.

See also Baier, Annette; Ethics, History of; Hobbes, Thomas; Rationalism in Ethics (Practical-Reason Approaches).

Bibliography

WORKS BY BAIER

The Moral Point of View. Ithaca, NY: Cornell University Press, 1958.

The Moral Point of View. Abridged and rev. ed. New York: Random House, 1965.

Values and the Future, edited with Nicholas Rescher. New York: Free Press, 1969.

The Rational and the Moral Order. Chicago: Open Court, 1995.

Reason, Ethics, and Society: Themes from Kurt Baier, with His Responses, edited by J. B. Schneewind. Chicago: Open Court, 1996. Contains a complete bibliography of Baier's publications up until 1995.

Problems of Life and Death: A Humanist Perspective. Amherst, NY: Prometheus Books, 1997.

Bernard Gert (2005)

BAIN, ALEXANDER
(1818–1903)

Alexander Bain, the Scottish philosopher and psychologist, was the son of a weaver. He was mainly self-educated but managed to attend Marischal College, in his native city of Aberdeen. After graduating he assisted the philosophy professor there from 1841 to 1844. A confirmed radical, Bain established close contacts with utilitarian circles in London, helping John Stuart Mill in the revisions of his unpublished *System of Logic* in 1842 and helping Edwin Chadwick with his sanitation reforms from 1848 to 1850. During the next decade, supporting himself

by journalism, he produced his magnum opus in two installments, titled *The Senses and the Intellect* (London, 1855) and *The Emotions and the Will* (London, 1859). Appointed professor of logic and rhetoric at Aberdeen in 1860, he published his *Manual of Rhetoric* (London, 1864) and his *Logic, Deductive and Inductive* (London and New York, 1870). On the proceeds of these and other textbooks he founded *Mind* in 1876, choosing his disciple George Croome Robertson as editor. After Bain's death his *Autobiography* (London, 1904), which gives his personal background and a useful criticism of his own books, was published.

CRITICISM OF ASSOCIATIONISM

Bain was not simply a pedestrian disciple of the two Mills. Fundamentally loyal to associationism, he was as discontented as J. S. Mill with its tenets but more systematic in his criticisms of them. What apparently made Bain uneasy was the narrow combination of introspection and emphasis on facts that characterized the associationistic science of mind. He was attracted by the physiologists' contemporary program of studying mind by a method uniting emphasis on facts with observation rather than introspection. At the same time Bain was interested in the recent efforts of the epistemologists to found a science that, while still introspective, was concerned not with empirical facts but with necessary truths. He had contacts with William Sharpey among the physiologists and James Ferrier among the epistemologists. Physiology and epistemology were interests alien to Mill.

THE WILL

The fusion of diverse tendencies in Bain's philosophy is best seen in the final section of his chief work—the discussion of the will—and especially its last hundred pages, which contain Bain's spirited defense of determinism, his justly famous theory of belief, and his equally interesting, though less known, analysis of consciousness. For Bain the central problem of the will apparently is the question of how I exercise voluntary control over my limbs. From the traditionalist standpoint it seemed an insoluble mystery how the mind knows just what motor nerves to activate when, for instance, expecting a blinding light to be switched on, it causes the eyes to close in advance. Bain's theory swept aside the traditional analogy with the case of first getting information about what is ahead and then operating a lever. The limbs are not inert like levers but possess an inherent spontaneity, and this spontaneity means that the expectation of the painful glare is inseparably associated with preparations to close the eye. The idea is that theory and practice are one. This doctrine of spontaneity, a direct ancestor of pragmatism, Bain rightly considered to be his most original contribution to philosophy, and he both discussed it effectively at the animal level and struggled honestly, in his discussion of effort, with the difficulty of applying it at the human level.

BELIEF

Bain's doctrine of belief arose in the context of his view of will. When he spoke of belief as being inseparable from "a preparation to act," he was envisaging as basic a situation in which one seriously expects alleviation of a present pain from something that is visible but out of reach. In the ensuing action of trying to grasp this thing, the belief is inevitably put to the test: "We believe first and prove or disprove afterwards." The essence of the human situation was thus for Bain a kind of circle of activity in which we inevitably acquire new nonrational beliefs as a direct consequence of practically and experimentally testing those we start with. The point is apparently that our actions have unforeseen consequences.

CONSCIOUSNESS

By an ingenious turn Bain used the pragmatist analysis of belief as a basis for a theory of consciousness inspired by William Hamilton's doctrine of the inverse ratio of sensation and perception. In Bain's version of the theory, a sharp contrast is drawn between the emotive pole of consciousness, where absorption in one's pains or pleasures prevents the objective assessment of one's situation, and the cognitive pole, where pleasures and pains are forgotten in the business of mapping one's world and where emotion appears only in the shock of scientific discovery, as a feeling that, like boredom, is outside the pleasure-pain sphere. The movement from feeling to knowledge in consciousness is linked with the same facts that give human life the character of a passage from belief to self-criticism.

But what, then, is this consciousness that underlies both the emotional side and the intellectual? Inspired by Hamilton and Ferrier, Bain made two points. First, we are unconscious of the undifferentiated. "A constant impression is to the mind a blank"—if temperature were unvarying we would not notice it. Second, we are conscious of the constant only in the midst of variety and difference. The essence of consciousness is thus to be discriminative, and Bain pointed out that of the discriminations involved in consciousness, the most liable to be misunderstood is that implicit in the problem of the external world. Bain argued that although Berkeley was

right in denouncing as meaningless the notion of material objects independent of experience, he overlooked an important point—that a distinction can be drawn within experience between the person sensing and the sensation sensed. Thus Bain, unlike J. S. Mill, conveyed a profound sense of the complexity of the problem of the external world.

Bain was aware that his philosophy was far removed from ordinary associationism. Above all, in the important Note F to the third edition of *The Senses and the Intellect,* he made it clear that for him association presupposed dis-association.

Bain progressively broke away from the heritage of the Mills, in logic as well as in psychology (he ultimately gave up Mill's view of logic for Augustus De Morgan's). At the same time there always survived in him certain tracts of unredeemed associationism. Thus, he retained to the last Mill's peculiar doctrine about the dependence of sight on muscular sense. So, too, his discussions of sympathy and of our knowledge of other minds are very crude examples of associationism.

These weaknesses in Bain have been too much stressed by his critics to the neglect of his merits. Thus, in dealing with the emotions the important role he gave to pure malice, or sadism, as a human motive contrasts refreshingly with the more commonplace views of such critics as Francis Herbert Bradley. Nevertheless, the only part of Bain's work that has been justly appreciated in our time is not his philosophy but his contribution to rhetoric.

See also Berkeley, George; Bradley, Francis Herbert; De Morgan, Augustus; Determinism, A Historical Survey; Ferrier, James Frederick; General Will, The; Hamilton, William; Introspection; Knowledge and Belief; Mill, James; Mill, John Stuart; Psychology; Utilitarianism.

Bibliography

Mental and Moral Science (London: Longmans, Green, 1868) is an abridgment of *The Senses and the Intellect* and *The Emotions and the Will.*

For works on Bain, see W. L. Davidson, "Professor Bain's Philosophy," in *Mind,* n.s., 13 (1904): 161–179; and Howard C. Warren, *A History of the Association Psychology* (New York: Scribners, 1921), pp. 104–117. For Bain's contributions to rhetoric, see Stephen Potter, *The Muse in Chains* (London: J. Cape, 1937).

George E. Davie (1967)

BAKER, LYNNE RUDDER

(1944–)

Lynne Rudder Baker was born in Atlanta, Georgia, received her PhD in philosophy from Vanderbilt University in 1972, and teaches at the University of Massachusetts at Amherst.

Her philosophical work provides a powerful critique of reductive accounts of minds, persons, and artifacts. Her writings in the philosophy of mind are directed against three distinct but related views. The first is that one's meaning something specific by a symbol can be naturalized, that is, reductively explained, in terms of some set of nonsemantic, nonmental, causal properties lawfully instantiated in nature. The second view is that folk psychology is, at best, a second-class prototheory of human behavior that only has instrumental value or, at worst, a discredited theory whose mental posits do not exist. The third view, what Baker calls "the Standard View," shared by dualists, materialists, and functionalists, says that beliefs are states of some proper part of persons, be it material (the brain) or immaterial (the soul).

All three views share two themes. First, we think of ourselves as sentient, sapient agents endowed with states that have referential content and causal efficacy. Second, if this conception is to be correct, it must reductively fit with our best scientific theories of nature, which have the right story (or much of it, at any rate) about things; otherwise, it must be rejected as false or treated as a useful but quaint myth. Baker accepts the first claim but rejects the second in *Saving Belief* (1987) and in *Explaining Attitudes* (1995). In the latter work, she defends practical realism, the view that beliefs are global states of a whole person, not of any proper part of the person. Although beliefs are not entities, they are real (contra eliminative materialists), since they make a genuine causal difference in the world in virtue of their contents (contra epiphenomenalists). Beliefs have an explanatory role, but not in virtue of their being identical to, constituted by, or supervening on brain states, since beliefs do not stand in those relations to any brain state. Rather, their explanatory role is grounded in our shared practice of causally explaining and rationalizing our actions. Baker's practical realism places her squarely in the company of American pragmatists (from William James to Hilary Putnam) and neo-Wittgensteinians.

Baker's third book, *Persons and Bodies* (2000), connects her early writings in the philosophy of mind with her more recent work in metaphysics. In that book she defends the constitution view of human persons, the view

that a human person is a person in virtue of having a first-person perspective and is human in virtue of being constituted by a human body. To have the first-person perspective is to have the ability to think of oneself as oneself in an irreducibly direct way without the mediation of any name or description. Constitution, in turn, is a ubiquitous relation that holds whenever new kinds of things come into existence (e.g., statues, persons), with new causal powers in virtue of other kinds of things (e.g., slabs of marble, human bodies), existing in certain types of circumstances (e.g., the art world, social institutions, and social practices). The things that constitute and the things they constitute have different persistence conditions and natures; hence, they are numerically distinct. Both share many of the same properties and causal powers, although the source of their shared properties and powers may lie with the thing that constitutes and not with the constituted thing, or vice versa. Thus, contrary to immaterialism, human persons are material beings, because they are constituted by their human bodies. However, contrary to animalism, human persons are not identical to the bodies that constitute them.

On Baker's view, although persons are constituted by their bodies and cannot exist without being materially constituted in some way, their identity over time does not depend on the particular bodies that constitute them. Nor does personal identity depend on soul identity, brain identity, or (nonbranching) psychological continuity. Rather, it depends solely on one's having a first-person perspective over time. Facts about one's first-person perspective are not reducible to any nonpersonal fact. Thus, for Baker, one's identity over time is a simple irreducible fact about oneself.

See also Identity; Mental Causation; Personal Identity; Philosophy of Mind.

Bibliography

Baker, Lynne Rudder. *Explaining Attitudes: A Practical Approach to the Mind*. New York: Cambridge University Press, 1995.

Baker, Lynne Rudder. *Persons and Bodies: A Constitution View*. New York: Cambridge University Press, 2000.

Baker, Lynne Rudder. *Saving Belief: A Critique of Physicalism*. Princeton, NJ: Princeton University Press, 1987.

Meijers, Anthonie, ed. *Explaining Beliefs: Lynne Rudder Baker and Her Critics*. Stanford, CA: CSLI Publications, 2001.

Reinaldo Elugardo (2005)

BAKHTIN, MIKHAIL MIKHAILOVICH

(1895–1975)

Mikhail Mikhailovich Bakhtin was a Russian philosopher, philologist, and historian of culture. In opposition to rationalism and, in general, to the modern European (monologic) epistemology, he grounded a personalistic understanding of being as the co-being (event) of interrelations of "I" and the "other" (thou) and developed a corresponding dialogic (and/or polyphonic) approach in the capacity of the uniquely adequate method of the particular humanitarian sciences and—more broadly—of philosophical thought.

THE WORKS

Bakhtin wrote his main works in the period from the 1920s to the beginning of the 1950s, but because of the political conditions of the time, biographical reasons, and the peculiarities of the texts themselves (some of them consisting of unfinished archival manuscripts), they were published (except in one case) either in the final years of the author's life or after his death.

Bakhtin was born in Orel, south of Moscow, and in the second decade of the twentieth century he studied at the historico-philological and philosophy departments first at Novorossisk University and then at Petersburg University. After the Communist Revolution of 1917 he lived in Nevel and Vitebsk, where a circle of like-minded intellectuals was formed (M. I. Kagan, L. V. Pumpiansky, V. N. Voloshinov, P. N. Medvedev et al.). Here, at the beginning of the 1920s, Bakhtin wrote early drafts of philosophical works that remained unfinished, including "*K filosofii postupka*" (Toward a philosophy of the act), first published in 1986, and "*Avtor i geroi v esteticheskoi deiatel'nosti*" (The author and the hero in aesthetic activity), first published in 1979. In 1924 Bakhtin returned to Leningrad, and that same year he wrote the antiformalist essay "*K voprosam metodologii estetiki slovesnogo tvorchestva*" ("On Questions of the methodology of the aesthetics of verbal creation"), first published in 1975.

Bakhtin's first published book (and until the beginning of the 1960s it remained his only published book) was *Problemy tvorchestva Dostoevskogo* (Problems of Dostoevsky's creative works), which appeared in 1929. There exists the assumption (not shared by all scholars or not shared by scholars to an equal degree) that certain other books and essays published in the 1920s and attributed to other authors were to some degree written by Bakhtin. These works include *Freidizm: Kriticheskii ocherk* (Freud-

ism: A critical essay), published in 1927, and *Marksizm i filosofiia iazyka* (Marxism and the philosophy of language), published in 1929, both attributed to V. N. Voloshinov, as well as P. N. Medvedev's *Formal'ni method v literaturovedenii* (Formal method in the study of literature), 1928.

In 1928 Bakhtin was arrested in connection with the affair of the illegal religious organization "Voskresenie." He was sentenced to five years in a concentration camp, but owing to the state of his health this sentence was replaced by a five-year exile in Kazakhstan. (Bakhtin suffered from chronic osteomyelitis, which in 1938 necessitated the amputation of one of his legs.) In accordance with this sentence, after returning from exile, he was prohibited from residing in large cities; and he was thus compelled to move from place to place. In 1945 he obtained a position in Saransk, at the Mordovia Pedagogical Institute, where he first worked as an instructor and then as department chairman. In the 1930s and 1940s he wrote a large study of Rabelais (which in 1946 he defended as his doctoral dissertation). In those years he also wrote a large cycle of works, published only in the 1970s, on the specific characteristics and genesis of the genre of the novel.

Bakhtin retired in 1961. By the middle of the 1960s his name could again be found in official scholarly publications. The second, revised edition of his book on Dostoevsky, *Problemy poetiki Dostoevskogo* (Problems of Dostoevsky's poetics) appeared in 1963; and the book based on his dissertation, *Tvorchestvo Fransua Rable i narodnaia kul'tura srednevekov'ia i Renessansa* (The work of Francois Rabelais and folk culture of the Middle Ages and Rennaissance), appeared in 1965. Bakhtin's ideas become known, particularly in Europe and the United States—first primarily among structuralists, and then, as the archive was published, among scholars with diverse philosophical and philological orientations. At the end of the 1960s Bakhtin moved first to a suburb of Moscow, and then at the beginning of the 1970s to Moscow itself, where he resided until his death.

THE INFLUENCE OF BAKHTIN

Bakhtin was initially subject to the diverse influences, on the one hand, of the development of the problem of the interrelations of "I" and the "other" in German philosophy (Feuerbach, Kierkegaard, Nietzsche, the neo-Kantians Cohen and Natorp) and on the other hand of Russian symbolism (in the version favored by Viacheslav Ivanov, who interpreted the interrelation of "I" and the "other" as a reduced transformation of the interrelations of "I" and "Thou" in the religious mysticism of communion with God). From the status of significant but particular problems of transcendental ethics and aesthetics or positivistic psychology and sociology, Bakhtin translated the interrelation between "I" and the "other" into a fundamental ontological structure of universal character, which determines both the forms of life's being and the forms of thought, language, and cultural meaning as such. In parallel with the legitimate goal, given such an approach, of identifying the universal archetypal forms of the interrelation of "I" and the "other," Bakhtin also posed the problem of exposing the various kinds of distortions of these archetypal forms in the historical types of culture.

Bakhtin did not leave an integral and consistently developed conception. Instead, he formulated several particular theories that are linked by a single personalistic-dialogic teleology but which are sometimes divergent in their outer conceptual contours (in particular, the conceptions of polyphony and carnival). In his early unfinished work "Toward a Philosophy of the Act," Bakhtin sketches out the project of a moral philosophy in which he grounds the constitutive role of the interrelations of "I" and the "other" for the structure of being. (Being is understood here as the co-being of two personal consciousnesses—as the minimum of the "co-being of being"; in order to accomplish the true co-being of the being of "I," which admits the validity of the ethical imperative, one must, according to this project, subject oneself to absolute self-exclusion from the values of the currently given being in favor of imparting these values to the "other.")

In "The Author and the Hero in Aesthetic Activity," Bakhtin gives a typology of different historical forms of the interrelations between author and hero, interpreted as aesthetic transformations of life-interrelations between "I" and the "other" (the author suppresses the hero and the hero suppresses the author; the crisis of the author, the revolt of the hero, etc.). In analyzing the historical types of culture, Bakhtin sees in the majority of them diverse forms of mutual overcoming and suppression of "I" and the "other," which replace their simultaneous mutual outside-locatedness and connectedness in one co-being by surrogates either of their illusory mutual isolation or of their just as illusory unity (physiological, psychological, ideological, national, social). Bakhtin attributed the disharmony of the interrelations between "I" and the "other" to the predominant orientation of the corresponding types of culture toward a unified and universal consciousness (the rationalistic gnoseologism, or monologism, of the modern period). The crisis of the

position of the author shaping the aesthetic co-being of being is advanced as the central aesthetic problem.

In *Problems of Dostoevsky's Poetics*, Bakhtin develops a theory of polyphony as a particular variant (created by Dostoevsky) of the genre of the novel with a specific authorial position that overcomes the crisis, a position that presupposes polyphonic dialogic intersections of the voices of the characters in the absence of the domination of the author's voice (including the narrator and all his other functional variants), which enters into fully equal dialogic relations with the voices of the characters. In the cycle of essays about the novel written in the 1930s and 1940s, Bakhtin complements the polyphonic conception with a general theory of the language of the novel as based on a word with two voices (on the intersection of two personal voices in a formally single utterance); he expounds the theory of the chronotope: the ambivalent relation of the temporal and spatial characteristics of meaning as the inalienable premise of its artistic representation and reception. When united with the spatial-temporal characteristics of the axiological dimension, the chronotope grows for Bakhtin into the analogue of any (not only artistic) point of view concerning meaning in the capacity of a position determined with respect to co-being and person.

In the essay "*Problema rechevykh zhanrov*" (Problem of speech genres) and in the second, revised edition of the book on Dostoevsky, Bakhtin develops a conception of metalinguistics, extending the theory of two voices beyond the word of the novel into the entire sphere of the life of language. In the book on Rabelais, he develops the conception of carnival as a reflection of the ambivalence of the archetypal foundation of folk-comic culture (the fusion without mutual neutralization of serious and comic myths) and, genetically connected with this conception, the conception of cultural meaning that is always constituted by antinomic or, in one respect or another, opposed relations, including dialogic ones.

Bakhtin's fundamental works have been translated into many European and Oriental languages. International conferences devoted to Bakhtin are held regularly, and monographs, collections of articles, and issues of journals devoted to his work are regularly published. Bakhtin's ideas generate much discussion and controversy.

See also Bakhtin Circle, The; Cohen, Hermann; Dostoevsky, Fyodor Mikhailovich; Feuerbach, Ludwig Andreas; Ivanov, Viacheslav Ivanovich; Kierkegaard, Søren Aabye; Natorp, Paul; Neo-Kantianism; Nietzsche, Friedrich; Rationalism; Russian Philosophy.

Bibliography

WORKS ABOUT BAKHTIN

Clark, Katerina, and Michael Holquist. *Mikhail Bakhtin*. Cambridge, MA: Belknap Press of Harvard University Press, 1984.

Dialog. Karnaval. Khronotop [Dialogue. Carnival. Chronotope]. Vitebsk, Belarus, 1992–2004. A periodical quaterly devoted to Bakhtin.

Emerson, Caryl. *The First Hundred Years of Mikhail Bakhtin*. Princeton, NJ: Princeton University Press, 1997.

Gogotishvili, L. A., and P. S. Gurevich, eds. *M. M. Bakhtin kak filosof* [M. M. Bakhtin as philosopher]. Moscow: Nauka, 1992. A collection of articles.

Isupov, K. G. *Mikhail Bakhtin: Pro et contra*. Anthology in two volumes. St. Petersburg: Russian Christian Institute for the Humanities, 2001, 2002.

WORKS BY BAKHTIN

Avtor i geroi v esteticheskoi deiatel'nosti [Author and hero in aesthetic activity]. In *Sobranie sochinenii v 7 tomakh* [Collected works in 7 volumes], 1996–. Moscow: Russkie slovari, 2003. Translated by Vadim Liapunov as *Author and Hero in Aesthetic Activity*. In *Art and Answerability: Early Philosophical Essays by M. M. Bakhtin*, edited by Michael Holquist and Vadim Liapunov (Austin: University of Texas Press, 1990).

"Formy vremeni i khronotopa v romane" [Forms of time and chronotope in the novel]. In *Voprosy literatury i estetiki* [Questions of literature and aesthetics]. Moscow: Khudozhestvennaia literatura, 1975, pp. 234–407. Translated by Caryl Emerson as *Forms of Time and the Chronotope in the Novel*. In *The Dialogic Imagination*, edited by Michael Holquist (Austin: University of Texas Press, 1981).

K filosofii postupka [Toward a philosophy of the act]. In *Sobranie sochinenii v 7 tomakh* [Collected works in 7 volumes], 1996–. vol. 1. Moscow: Russkie slovari, 2003. Translated by Vadim Liapunov as *Toward a Philosophy of the Act*, edited by Michael Holquist and Vadim Liapunov (Austin: University of Texas Press, 1993).

Problemy poetiki Dostoevskogo [Problems of Dostoevsky's poetics]. In *Sobranie sochinenii v 7 tomakh* [Collected works in 7 volumes], 1996–. Moscow: Russkie slovari, 2002. Translated and edited by Caryl Emerson as *Problems of Dostoevsky's Poetics* (Minneapolis: University of Minnesota Press, 1984).

Tvorchestvo Fransua Rable i narodnaia kul'tura srednevekov'ia i Renessansa [The work of François Rabelais and the popular culture of the Middle Ages and the Renaissance]. Moscow: Khudozhestvennaia literatura, 1965. Translated by Hélène Iswolsky as *The Work of Francois Rabelais and Popular Culture of the Middle Ages*. In *Rabelais and His World* (Cambridge, MA: MIT Press, 1968).

Ludmila Gogotishvili (2005)
Translated by Boris Jakim

BAKHTIN, MIKHAIL MIKHAILOVICH [ADDENDUM]

By the time his boom and cult had passed, Mikhail Bakhtin had become a twentieth-century classic and the beneficiary of a huge research industry. Accordingly, the most exciting work shifted from literary or political applications of his famous terms—dialogue, carnival, chronotope—and toward the finer, and far more interesting, arts of historical recuperation: Bakhtin's intellectual debts, and his social and philosophical contexts (see Brandist 2002). For Bakhtin Studies, 1990 was something of a watershed year. It marked, of course, the beginning of the end of Soviet Communism, which made it possible for Russians to pursue pluralistic and de-ideologized scholarship throughout the humanities. For English speakers it was also the year that Bakhtin's writings from the 1920s (combining Kantianism and phenomenology in a distinctive moral philosophy) were published in the excellent Liapunov annotated translations.

It took several years for these difficult early texts to be assimilated, for the received image of Bakhtin in the 1970s and 1980s could not easily be fit back into them. That image, based on several widely (and quickly) translated texts from his middle-to-late period, was polarized between those who wished to see in Bakhtin a pragmatic, systems-shunning liberal humanist and those who preferred a more radical and subversive message. Neither variant had firm documentation (the liberal least of all). The question of Bakhtin's Marxism and his authorship of the "disputed texts" had ended in a draw. Left-wing cultural theorists were faulting "dialogism" for its fascination with process at the expense of justice and for its indifference to power. Bakhtinian ideas permeated every possible discipline (sociology, cultural studies, therapeutic psychoanalysis, history of science, theories of education) but as yet we lacked the luminous renderings of Bakhtin as a spiritual thinker and aesthetician. The biography was still awash in rumor, and influences on him largely conjectural.

By the mid-1990s several Bakhtin scholars, most prominently in Britain and Russia, began to suspect that Bakhtin's ideas were so shockingly famous because we had forgotten, or too thinly investigated, the richness of the historical period of which they were an organic part: the German and then Russian philosophical debates of the 1910s and 1920s. With the appearance of the first volumes of the collected works, *M. M. Bakhtin: Sobranie sochinenii* (Moscow, 1996–) and, in English, of the work of Galin Tihanov, Brian Poole, Ken Hirschkop, David Shepherd, and especially Craig Brandist (2002), it became clear that the "trademark" concepts, painstakingly restored to their appropriate contexts, would have to be retranslated and critically rethought.

To be sure, these concepts had been under revision for some time. *Dialogue*, which insists upon the addressivity, reciprocity, and open-endedness of all relations, had long been reproached for political naiveté, for flattening the epic, for undervaluing poetry (with its tolerance for repetition, symmetry, and formal constraints), and for denying a stable core to the self. Novelistic *polyphony*, which Bakhtin saw exemplified in Dostoevsky, was also controversial. The idea (of surprising appeal to primary authors) that created characters that can act and speak alongside their creators as "equally weighted" consciousnesses has been dismissed by drier and more disciplined critics as a fantasy, as an illusion of the author—or of the readers—who project their own ideas and words on to the text.

Carnival was the term most indiscriminately applied. It had come to mean little more than sassiness, rebellion, or transgression, and as such was applied to every social practice, text, or body that revealed a disruptive, subversive, inverting, or comically grotesque aspect. Sobriety set in here too. Not only was Bakhtin's sunny carnivalesque shown to bear little relation to real, drunken, violent carnival rituals and bodies, but the literary masterwork Bakhtin used to illustrate his theory, Rabelais's *Gargantua et Pantagruel*, was served only partially and rather poorly by so crudely binary and folkloric a filter. Trivially oppositional readings of dialogue and of carnival reinforced each other: The double-voiced word was deployed more often to subvert a perspective than to supplement or enrich it. Of all Bakhtin's famous terms, the *chronotope* proved to be the least contentious. It was also the most "philosophical" of Bakhtin's constructs, a creative extension to narrative of the Kantian time-space matrix.

Thus the age of "applied Bakhtin" gave way to a study of "Bakhtin the philosopher." Problems remained, but they became deeper and more productive. Researchers took seriously Bakhtin's claim that his life's work aimed to present an integrated philosophical worldview rather than to further a strictly "philological" enterprise—that is, a series of literary readings designed to explicate or serve the interests of their respective literary authors. The starting point was Kant and his successors among the German Romantics, a powerful collage of thinkers united by their inquiry into the possibility of human freedom. The end point, arrived at with ample help from Ernst

Cassirer, Max Scheler, Matvei Kagan (1889–1937), and Bakhtin's peer intellectuals of his own circle, was an understanding of freedom not as agency or as initiated deed but as *response*: individuated, concrete, identity-bestowing, and in principle unfinalizable. This mandate applied not only to human responsibility but to thought itself. As Bakhtin's mentor Matvei Kagan wrote in his fragment "Philosophy and Life" (1918–1920), philosophy is a sort of immortal organism, not mechanically logical but also not subject to the constraints of "biologism." Most importantly, philosophy was not obliged to begin with nonbeing. It was always materialized and concrete. "The world is not dying, not being annihilated," Kagan wrote, astonishingly, at the end of the Great War; "*it has not yet completely come to life, but is doing so*." There is no absolute nonbeing, only not-yet-Being, and this "incomplete being is on a constant path of new becoming" (Kagan 2004, p. 311).

This philosophical reorientation promises further shifts and revised shapes for Bakhtin's ideas. The dialogic novel has already begun to be seen as the model site for the "relational self"— cocreated, but not for that reason any less coherent, unified, and authentic (de Peuter in Bell and Gardiner 1998, pp. 30–48). Inspiration for carnival and the grotesque body is being sought in areas as diverse as Ivan Kanaev's research on the regenerative capabilities of the freshwater polyp (Taylor in *The Bakhtin Circle* 2004, pp. 150–166) and in Trinitarian paradigms of Russian Orthodox thought (Mihailovic 1997). Bakhtin's ideas of genre and chronotope, and more recently of "answerability," have been immensely influential on the vast industry of college-level pedagogy in the United States, specifically on the theory of teaching English composition (Halasek 1999, Farmer 2001). And finally, attention is being paid to Bakhtin's fragmentary, somewhat dated, but still robust thoughts on the humanities as the realm of depth and reciprocity over time rather than of scientific precision, the realm of experience rather than experiment. As communication is increasingly pressured to default to the values of speed, here, now, and simultaneity without reflection, the historical embeddedness of Bakhtin's ideas will provide a welcome corrective and relief.

See also Philosophical Anthropology; Russian Philosophy.

Bibliography

WORKS BY BAKHTIN

Art and Answerability: Early Philosophical Essays, edited by Michael Holquist and Vadim Liapunov. Austin: University of Texas Press, 1990.

The Dialogic Imagination: Four Essays by M. M. Bakhtin, edited by Michael Holquist; translated by Caryl Emerson and Michael Holquist. Austin: University of Texas Press, 1981.

Problems of Dostoevsky's Poetics, edited and translated by Caryl Emerson. Minneapolis: University of Minnesota Press, 1984.

Rabelais and His World. Translated by Hélène Iswolsky. Cambridge, MA: MIT Press, 1968. First Midland book edition, Bloomington: Indiana University Press, 1984.

Speech Genres and Other Late Essays, edited by Caryl Emerson and Michael Holquist. Translated by Vern W. McGee. Austin: University of Texas Press, 1986.

M. M. Bakhtin: Besedy V. D. Duvakinym [1973–1974]. Moscow: Soglasie, 2002.

WORKS ABOUT BAKHTIN

Bell, Michael Mayerfeld, and Michael Gardiner, eds. *Bakhtin and the Human Sciences: No Last Words*. Thousand Oaks, CA: Sage, 1998.

Brandist, Craig. *The Bakhtin Circle: Philosophy, Culture and Politics*. Sterling, VA: Pluto, 2002.

Brandist, Craig, David Shepherd, and Galin Tihanov, eds. *The Bakhtin Circle: In the Master's Absence*. Manchester, U.K.: Manchester University Press, 2004.

Clark, Katerina, and Michael Holquist. *Mikhail Bakhtin*. Cambridge, MA: Harvard University Press, 1984.

Coates, Ruth. *Christianity in Bakhtin: God and the Exiled Author*. Cambridge, U.K.: Cambridge University Press, 1998.

Emerson, Caryl. *The First Hundred Years of Mikhail Bakhtin*. Princeton, NJ: Princeton University Press, 1997.

Farmer, Frank. *Saying and Silence. Listening to Composition with Bakhtin*. Logan: Utah State University Press, 2001.

Felch, Susan M., and Paul Contino, eds. *Bakhtin and Religion: A Feeling for Faith*. Evanston, IL: Northwestern University Press, 2001.

Halasek, Kay. *A Pedagogy of Possibility: Bakhtinian Perspectives on Composition Studies*. Carbondale: Southern Illinois University Press, 1999.

Haynes, Deborah J. *Bakhtin and the Visual Arts*. Cambridge, U.K.: Cambridge University Press, 1995.

Hirschkop, Ken. *Mikhail Bakhtin: An Aesthetic for Democracy*. Oxford, U.K.: Oxford University Press, 1999.

Hirschkop, Ken, and David Shepherd, eds. *Bakhtin and Cultural Theory*. 2nd ed. Manchester, U.K.: Manchester University Press, 2001.

Holquist, Michael. *Dialogism: Bakhtin and His World*. London: Routledge, 1990.

Kagan, M. I. *O khode istorii* (Collected writings), edited by V. L. Makhlin. Moscow, 2004.

Mihailovic, Alexandar. *Corporeal Words: Mikhail Bakhtin's Theology of Discourse*. Evanston, IL: Northwestern University Press, 1997.

Morson, Gary Saul, and Caryl Emerson. *Mikhail Bakhtin: Creation of a Prosaics*. Stanford, CA: Stanford University Press, 1990.

Poole, Brian. "Bakhtin and Cassirer: The Philosophical Origins of Bakhtin's Carnival Messianism." *South Atlantic Quarterly* 97 (3/4) (1998): 537–578.

Shepherd, David, ed, *The Contexts of Bakhtin: Philosophy, Authorship, Aesthetics*. London: Harwood, 1998.

Tihanov, Galin. *The Master and the Slave: Lukács, Bakhtin, and the Ideas of Their Time*. Oxford, U.K.: Clarendon Press, 2000.

Caryl Emerson (2005)

BAKHTIN CIRCLE, THE

The Bakhtin Circle was a group of Soviet scholars, including the cultural theorist Mikhail Mikhailovich Bakhtin (1895–1975), the linguist Valentin Nikolaevich Voloshinov (1895–1936), and the literary scholar Pavel Nikolaevich Medvedev (1891–1938). Drawing on a variety of philosophical positions, the group developed a philosophy of the human sciences, language, literary production, and history, and a wide-ranging cultural theory. The group's work combined, in various ways, the neo-Kantianism of the Marburg School (especially Ernst Cassirer), phenomenology (especially Max Scheler and Karl Bühler), Russian Formalism (especially Lev Iakubinskii), Hegelianism, and various types of Marxism current within Soviet scholarship (especially Georg Lukács and "Marrism").

In *K filosofii postupka* (Toward a philosophy of the act; 1993 [written in the mid-1920s]), Bakhtin combines a neo-Kantian idealism, in which ethics is the foundation of the human sciences and jurisprudence its "mathematics," with the phenomenological notion of intentionality to develop an ethics based on the acts of the responsible subject. *Avtor i geroi v esteticheskoi deiatel'nosti* (Author and hero in aesthetic activity; 1990 [written in the mid- to late-1920s]) is a phenomenological investigation into relations between author and hero in narrative fiction based to a considerable extent on the account of intersubjectivity found in Scheler's *The Nature and Forms of Sympathy* (Poole 2001).

Medvedev and Voloshinov had meanwhile been working on developing a sociological approach to poetics and discursive interaction, respectively. Both sought to bring about a meeting of contemporary philosophical trends with the sociological ideas championed by Russian Marxists at the time, particularly Nikolai Bukharin. In his essay *Formal'nyi metod v literaturovedenii* (The formal method in literary scholarship; 1978 [1928]), Medvedev argues that sociological factors shape literature from within and without and that exploration of the category of genre should precede analyses of individual literary devices. In *Marksizm i filosofiia iazyka* (Marxism and the philosophy of language; 1973 [1929]), Voloshinov contends that language is a product of social interaction, emerging in and through dialogue, and, following Bühler, that the utterance constitutes the primary unit of language *in actu*. This phenomenology of social interaction in language is given a sociological form, so that specific styles of language use are the discursive embodiments of the worldviews of specific social groups. Modalities of authorship are also reworked into an analysis of various forms of reported speech in literature whereas literary and extraliterary forms of discourse are all held to have generic characteristics. Bakhtin himself accepted this reworking in his now famous *Problemy tvorchestva Dostoevskogo* (Problems of Dostoevsky's poetics; 1984 [1929, 1963]), where the novelist is held to have produced a "polyphonic" form in which all languages, including that of the narrator, interact on an equal and, indeed, democratic basis.

Whereas the Circle ceased to function as a group after Joseph Stalin's consolidation of power at the end of the 1920s, Bakhtin's own most important work was produced in subsequent years. In a series of essays written in the 1930s and 1940s, Bakhtin drew on the work of, among others, Cassirer and Lukács to develop a radical re-reading of literary history and the place of the novel therein. Recasting Cassirer's idealist dialectic of mythical and critical symbolic forms, Bakhtin argues that the novel has roots in popular and skeptical discursive forms that exploit the social stratification of language (heteroglossia) to undermine the truth claims of official, poetic discourse. This skepticism operates through laughter that, following Cassirer and Henri Bergson, Bakhtin sees as deflating discursive pretension and revealing that knowledge of the empirical world is impossible. In a typically Hegelian move, Bakhtin argues that it is in and through the novel that culture, the totality of discursively embodied perspectives (heteroglossia), becomes aware of itself as its own object. The dogmatic and authoritarian attitude toward another's discourse is termed "monologic" whereas a critical and democratic attitude is termed "dialogic." These essays began to be published in the 1970s and appeared in English under the title *The Dialogic Imagination* (1981).

At the end of the 1930s, Bakhtin develops a theory that the rise of the critical forces of culture represents the reappearance of semantic forms that have survived from preclass society. This theory builds on the theory of "semantic paleontology" developed by the now discredited Soviet archaeologist and linguist Nikolai Marr, who

argued that all languages develop from a primordial gesture language in primitive communism. Marr's position had been reworked and applied to literary material by the classicist Ol'ga Freidenberg, who identified certain primordial "semantic clusters" that reappear in various ways throughout cultural history. In Bakhtin's hands this model became the now-famous theory of carnival, in which forms associated with the popular culture of laughter come to permeate and structure literary works. Symbolic inversions, collective festivity, and mockery relativize the dominant culture, parading its conventionality, pomposity, and claims to discursive adequacy. Carnival on the streets is a licensed and limited rebellion against the ruling symbolic order, but once its features enter "great literature," the critical spirit that motivates it restructures the relationship between official and popular culture, democratizing the former and breaking the isolation of the latter. Bakhtin finds such features throughout the literature of the Renaissance, but he gives special attention to the work of the French novelist François Rabelais in *Tvorchestvo Fransua Rable i narodnaia kul'tura srednevekov'ia i Renessansa* (Rabelais and his world; 1984 [1965]) written at the end of the 1930s. This work was originally Bakhtin's doctoral (*kandidatskaia*) dissertation.

As part of his project dealing with the rise of modern critical culture, Bakhtin also writes important articles on the spatiotemporal characteristics of particular genres, or *chronotopes*, and a special work on the generic features of Johann Goethe's *Bildungsroman* the surviving part of which is known as *Roman vospitaniia vistorii realizma* (The Bildungsroman and its significance in the history of realism, written in the late 1930s). Bakhtin argues that it is in the work of the polymath Goethe that the Renaissance *demythification* of the world reaches its highest point. Following Stalin's denunciation of Marr in 1950, Bakhtin also sought to distinguish between a *human* science of discursive or speech (*rechevoi*) genres and a *natural* science of linguistic structures. In Bakhtin's posthumously published final works, translated as *Speech Genres and Other Late Essays* (1986), this neo-Kantian concern with demarcating the natural and human sciences becomes his central focus. The natural sciences, which adopt a *monologic* approach to their *voiceless* object, deal with questions of causality and determination whereas the human sciences, eschewing all such considerations, are based on a *dialogic* methodology and pursue an ethics of intersubjectivity.

See also Bakhtin, Mikhail Mikhailovich; Bergson, Henri; Cassirer, Ernst; Goethe, Johann Wolfgang von; Hegelianism; Idealism; Intentionality; Lukács, Georg; Marxist Philosophy; Neo-Kantianism; Phenomenology; Russian Philosophy; Scheler, Max.

Bibliography

WORKS

Bakhtin, M. M. *The Dialogic Imagination: Four essays* (*Voprosy literatury i estetiki*). Translated by Caryl Emerson and Michael Holquist; edited by Michael Holquist. Austin: University of Texas Press, 1981.

Bakhtin, M. M. *K filosofii postupka* (c. 1925) (Toward a philosophy of the act). Translation and notes by Vadim Liapunov; edited by Michael Holquist and Vadim Liapunov. 1st ed. Austin: University of Texas Press, 1993.

Bakhtin, M. M. *Problemy tvorchestva Dostoevskogo* (1929) (Problems of Dostoevsky's poetics), edited and translated by Caryl Emerson, 1984.

Bakhtin, M. M. *Rabelais and His World* (*Tvorchestvo Fransua Rable i narodnaia kul'tura srednevekov'ia i Renessansa*). Translated by Hélène Iswolsky. Bloomington: Indiana University Press, 1984.

Bakhtin, M. M. *Speech Genres and Other Late Essays* (*Estetika slovesnogo tvorchestva*). Translated by Vern W. McGee; edited by Caryl Emerson and Michael Holquist. 1st ed. Austin: University of Texas Press, 1986.

Medvedev, P. N. *The Formal Method in Literary Scholarship: A Critical Introduction to Sociological Poetics* (*Formal'nyi metod v literaturovedenii*) (1928).Translated by Albert J. Wehrle. Baltimore, MD: Johns Hopkins University Press, 1978.

Voloshinov, V. N. *Marksizm i filosofiia iazyka* (1929) (Marxism and the philosophy of Language). Translated by Ladislav Matejka and I. R. Titunik. New York: Seminar Press, 1973.

STUDIES

Brandist, Craig. *The Bakhtin Circle: Philosophy, Culture, and Politics*. London: Pluto, 2002.

Brandist, Craig, David Shepherd, and Galin Tihanov, eds. *The Bakhtin Circle: In the Master's Absence*. Manchester: Manchester University Press, 2004.

Hirschkop, Ken. *Mikhail Bakhtin: An Aesthetic for Democracy*. Oxford: Oxford University Press, 1999.

Poole, Brian. "From Phenomenology to Dialogue: Max Scheler's Phenomenological Tradition and Mikhail Bakhtin's Development from 'Toward a Philosophy of the Act' to His Study of Dostoevsky." In *Bakhtin and Cultural Theory*, edited by Ken Hirschkop and David Shepherd. Manchester: Manchester University Press, 2001.

Tihanov, Galin. *The Master and the Slave: Lukács, Bakhtin, and the Ideas of Their Time*. Oxford: Oxford University Press 2000.

Craig Brandist (2005)

BAKUNIN, MICHAEL

See *Bakunin, Mikhail Aleksandrovich*

BAKUNIN, MIKHAIL ALEKSANDROVICH

(1814–1876)

Mikhail Aleksandrovich Bakunin, the anarchist writer and revolutionary leader, was born on the estate of Premukhino in the Russian province of Tver'. His family were hereditary noblemen of liberal political inclinations. His father had been in Paris during the French Revolution and had taken his doctorate of philosophy at Padua. His mother was a member of the Murav'av family; three of her cousins were involved in the earliest Russian revolution, the December rising of constitutionalists in 1825. Bakunin was carefully educated under the supervision of his father, who regarded himself as a disciple of Jean-Jacques Rousseau; later he was sent to the Artillery School in St. Petersburg. He received his commission and went on garrison duty in Lithuania. An awakening taste for literature made him discontent with military life, and in 1835 he obtained his discharge from the army and went to Moscow to study philosophy. There he joined the discussion circle centered on Nicolai Stankevich, which concentrated on contemporary German philosophy.

HEGELIANISM AND REVOLUTION

Bakunin was first influenced by Johann Gottlieb Fichte; his earliest literary task was the translation of that philosopher's writings for Vissarion Belinskii's periodical, the *Teleskop* (The Telescope). Later he transferred his allegiance to G. W. F. Hegel, and he advocated the Hegelian doctrine in its most conservative form with such enthusiasm that when Stankevich left for western Europe, Bakunin became the leader of the Hegelian school in Moscow and challenged the liberalism of the rival group associated with Alexander Herzen, who propagated the ideas of Charles Fourier, Comte de Saint-Simon, and Pierre-Joseph Proudhon.

Bakunin left Russia in 1840 to study German philosophy in Berlin. He still wished to become a professor of philosophy, and assiduously attended the lectures for some time; in his leisure hours he frequented the literary salons in the company of Ivan Turgenev, who used him as a model for the hero of his first novel, *Rudin*.

In 1842 Bakunin moved to Dresden, an intellectual as well as a physical journey. He had made the acquaintance of Arnold Ruge, leader of the Young Hegelians, whose contention that Hegel's dialectical method could be used more convincingly to support revolution than reaction was to influence almost every school of socialist philosophy in mid-nineteenth-century Europe. Bakunin's meeting with Ruge, combined with his reading of Lorenz von Stein's writings on Fourier and Proudhon, effected a change of his viewpoint that had all the strength of religious conversion.

The first manifestation of this change was the essay "Reaction in Germany—A Fragment by a Frenchman," which Bakunin published under the nom de plume of Jules Elysard in Ruge's *Deutsche Jahrbücher für Wissenschaft und Kunst* (October 1842). It puts forward a Young Hegelian view of revolution; before it succeeds, revolution is a negative force, but when it triumphs, it will, by a dialectical miracle, immediately become positive. However, the most striking feature of the essay is the apocalyptic tone in which Bakunin introduces the theme—recurrent in his writings—of destruction as a necessary element in the process of social transformation. "Let us put our trust in the eternal spirit which destroys and annihilates only because it is the unsearchable and eternally creative source of all life. The urge to destroy is also a creative urge."

"Reaction in Germany," with its glorification of the idea of perpetual revolt, was the first step toward Bakunin's later anarchism, but he went through many stages before he reached that destination. At first, in Switzerland, he associated with the German revolutionary communist, Wilhelm Weitling. This drew the attention of the Russian authorities to Bakunin's awakening radicalism, and he was condemned in absentia to indefinite exile with hard labor in Siberia.

PAN-SLAVISM

Meanwhile, Bakunin moved to Paris, where he associated with Karl Marx, Robert de Lamennais, George Sand, and, most important, Proudhon. Only in later years did these discussions bear fruit, when Bakunin became Marx's great enemy and Proudhon's great disciple; for the time being, he was concerned with the liberation of the Poles and other Slav peoples. For his speeches against the Russian government he was expelled to Belgium; he returned to Paris with the February Revolution of 1848. The years of the revolutions in Europe—1848–1849—were the most dramatic period of Bakunin's life. He was an enthusiastic partisan of the uprising in France; later in 1848 he fought on the barricades of Prague, and in March 1849, he took a leading part, with Richard Wagner, in the Dresden revolution. He was captured there and, after periods in Saxon and Austrian prisons and twice being sentenced to death and reprieved, he was handed over to the Russian authorities, who imprisoned him in the Peter and Paul Fortress. Six years there ruined his health. In 1857 he

was sent to exile in Siberia, and in 1861 he escaped, via Japan and the United States, to western Europe.

During the years of action and imprisonment Bakunin produced two important works, the *Appeal to the Slavs*, written in the interval between the Prague and Dresden revolutions, and the *Confession*, which he wrote in prison at the request of Tsar Nicholas II and which was published after the Russian Revolution. The *Appeal to the Slavs* is much more than a statement of Bakunin's Pan-Slavism; in many ways it anticipates his later anarchist attitudes. The social revolution, he declares, must take precedence over the political revolution and, on moral grounds, he claims that the social revolution must be total. "We must first of all purify our atmosphere and transform completely the surroundings in which we live, for they corrupt our instincts and our wills.… Therefore the social question appears first of all as the overthrow of society," by which Bakunin evidently means the overthrow of the contemporary social order. Bakunin further maintains that liberty is indivisible and thus implies the rejection of individualism in favor of the collectivism that becomes explicit in the later development of his anarchist doctrine. The *Confession* is important principally for its account of the early development of Bakunin's revolutionary philosophy.

After his escape to western Europe in 1861, Bakunin resumed the course of Pan-Slavism he had been forced to abandon in 1849 but, after taking part in an abortive Polish attempt to invade Lithuania in 1863, he went to Italy.

ANARCHISM

In 1865 Bakunin founded the International Brotherhood in Naples. Its program—embodied in Bakunin's *Revolutionary Catechism*—was anarchism without the name; it rejected the state and organized religion, advocated communal autonomy within a federal structure, and maintained that labor "must be the sole base of human right and of the economic organization of the state." In keeping with the cult of violence that was part of the romantic revolutionary tradition, Bakunin insisted that the social revolution could not be achieved by peaceful means.

The International Brotherhood was a conspiratorial organization, for Bakunin never outlived his taste for the dark and the secret. Nevertheless, in 1867 he emerged into public life as a figurehead of the short-lived League for Peace and Freedom. This was mainly a body of pacifistic liberals, within which Bakunin led the left wing.

Bakunin was not a systematic writer. He admitted that he had no sense of "literary architecture" and saw himself primarily as a man of action, although his action was rarely successful and his life was punctuated by abortive revolutions. His writings were intended to provoke action; they were topical in inspiration, if not always in content, and it is in pamphlets on current events and in reports written for congresses and organizations that his opinions are scattered. One such report, prepared for the benefit of the central committee of the League for Peace and Freedom, was eventually published as *Federalism, Socialism and Anti-Theologism*. More than any other work, it contains the gist of Bakunin's anarchism.

Bakunin was not a great theoretical originator. The influences in his writings are obvious—Hegel, Auguste Comte, Proudhon, Ruge, Charles Darwin, and even Marx. Original in Bakunin are his insight into contemporary events (he prophesied with uncanny exactitude the way in which a Marxist state would operate) and his power to create a synthesis of borrowed ideas around which the early anarchist movement could crystallize. In *Federalism, Socialism and Anti-Theologism* the view of the structure of a desirable society is almost completely derived from Proudhon's federalism. In one vital respect, however, Bakunin's view differs from Proudhon's: While he follows Proudhon in measuring the consumer's right to goods by the quantity of his labor, he also advocates the collectivization of the means of production under public ownership; Proudhon and his mutualist followers wished to retain individual possession of land and tools by peasants and artisans as far as possible, in order to create a guarantee of personal independence. This difference was regarded as so important that Bakunin's followers were actually described as "collectivists" and did not assume the name of "anarchists" until the 1870s.

In 1868 Bakunin left the League for Peace and Freedom to found the International Alliance of Social Democracy, which was dissolved when he and his followers entered the International Workingmen's Association in 1869. Within the International, Bakunin and the southern European federations challenged the power of Marx. The dispute centered on disagreement over political methods. Marx and his followers held that socialists must seize the state and usher in a transitional dictatorship of the proletariat. Bakunin argued that power seized by workers was no less evil than power in other hands, and a communist state would magnify the evil of other states; he called for the earliest possible destruction of the state and the avoidance of political means toward that end. The workers must win their own liberation by eco-

nomic and insurrectional means. The dispute came to a head at The Hague Congress of the International in 1872, when Bakunin was expelled. The southern federations and those of the Low Countries seceded to form their own federation, and Marx's remnant faded away.

Meanwhile, Bakunin's health declined rapidly. He took part in the Lyons rebellion of 1870 and in the abortive Bologna uprising of 1874. He died, exhausted, two years later at Bern. After his death, the anarchist communism of Pëtr Alekseevich Kropotkin superseded his collectivist anarchism, except in Spain, where the large anarchist movement held his ideas in their purity until 1939.

See also Anarchism; Belinskii, Vissarion Grigor'evich; Comte, Auguste; Darwin, Charles Robert; Fichte, Johann Gottlieb; Fourier, François Marie Charles; Hegel, Georg Wilhelm Friedrich; Hegelianism; Herzen, Aleksandr Ivanovich; Kropotkin, Pëtr Alekseevich; Lamennais, Hugues Félicité Robert de; Marx, Karl; Political Philosophy, History of; Proudhon, Pierre-Joseph; Rousseau, Jean-Jacques; Russian Philosophy; Saint-Simon, Claude-Henri de Rouvroy, Comte de.

Bibliography

WORKS BY BAKUNIN

The Confession of Mikhail Bakunin: With the Marginal Comments of Tsar Nicholas I. Translated by R. C. Howes. Ithaca, NY: Cornell University Press, 1977.

Statism and Anarchy. Translated by M. S. Shatz. Cambridge, U.K., and New York: Cambridge University Press, 1990.

The Basic Bakunin: Writings, 1869–1871. Translated by R. M. Cutler. Amherst, NY: Prometheus, 1992.

Political Thought, edited by M. Rosen and J. Wolff, with the assistance of C. McKinnon. Oxford and New York: Oxford University Press, 1999.

WORKS ON BAKUNIN

Aldred, G. A. *Bakunin*. New York: Haskell House, 1971.

Kelly, A. *Mikhail Bakunin: A Study in the Psychology and Politics of Utopianism*. Oxford: Clarendon Press; New York: Oxford University Press, 1982.

Maximoff, G. P., editor. *The Political Philosophy of Bakunin: Scientific Anarchism*. New York: The Free Press; London: Collier Macmillan, 1953.

McLaughlin, P. *Mikhail Bakunin: The Philosophical Basis of His Theory of Anarchism*. New York: Algora, 2002.

Mendel, A. P. *Michael Bakunin: Roots of Apocalypse*. New York: Praeger, 1981.

Saltman, R. B. *The Social and Political Thought of Michael Bakunin*.Westport, CT: Greenwood Press, 1983.

George Woodcock (1967)
Bibliography updated by Vladimir Marchenkov (2005)

BALFOUR, ARTHUR JAMES
(1848–1930)

Arthur James Balfour, the first earl of Balfour, was born at Whittingehame, Haddington, East Lothian. He was the son of a Scottish landowning family and was connected, through his mother, with the aristocratic house of Cecil. After an education at Eton and Trinity College, Cambridge, where he came under the influence of Henry Sidgwick (later his brother-in-law), he became a Conservative M.P. in 1874 and, despite an early reputation for indolence and frivolity, soon rose, by a combination of influence and ability, to ministerial rank. Having made his name as a courageous and enlightened chief secretary for Ireland during the turbulent period from 1887 to 1891, he became leader of the House of Commons in 1891 and in 1902 succeeded his uncle, Lord Salisbury, as prime minister. Beset by dissensions over tariff reform, his administration fell in 1905; but he remained leader of the Opposition until 1911. He resumed office in the wartime coalition as first lord of the admiralty, later becoming foreign secretary and lord president of the council. In these capacities he played a major part in the postwar negotiations at Versailles and Washington and, by the Balfour Declaration of 1917, in the eventual establishment of the state of Israel. He received the Order of Merit in 1916 and a Garter knighthood, followed by an earldom, in 1922. Among many other distinctions, he was chancellor of both Cambridge and Edinburgh universities, fellow of the Royal Society, president of the British Academy, the British Association, and the Aristotelian Society, and one of the founders of the Scots Philosophical Club. As an elder statesman whose disinterested sagacity was equally valued by both parties, Balfour in his later years enjoyed a unique position in British political life. He died, unmarried, at Woking.

Balfour's intelligence, versatility, and charm were at the service of many causes besides politics. Science and education were among his keenest interests; with his sister, Mrs. Sidgwick, he was a leading figure in the Society for Psychical Research. His leisure was divided equally between the arts and society, on the one hand, and tennis and golf on the other. Philosophy, however, was his main pursuit in private life, and in this sphere also—like his fellow statesman Richard Burton Haldane—he made a definite, if temporary, mark. Aside from having considerable literary merits, his writings are chiefly notable as a vigorous and independent contribution to the literature of the perennial conflict between science and religion.

Balfour had a strong distaste for the evolutionary naturalism of his younger days, and made repeated attempts to expose its pretensions as a prelude to stating the case for a "higher Reason" and the acceptance of Christian belief. To this end he employs skeptical weapons of a type forged by George Berkeley and David Hume and subsequently wielded by Henry Longueville Mansel, while his own defenses owe more than a little to Edmund Burke. If the would-be scientific answers to the problems of knowledge and human existence turn out, on examination, to be at once ungrounded and inconsistent, they supersede neither the time-honored beliefs of common sense nor the equally cherished, albeit unprovable, convictions of religion. Balfour's first book, *A Defence of Philosophic Doubt* (London, 1879), argues derisively against the claims of any prevailing system of thought to justify, let alone criticize, the natural and "inevitable" beliefs in the external world, in the uniformity of nature and, to a lesser extent, in theism. His second book, the widely read *Foundations of Belief* (London, 1895), renews the polemic against John Stuart Mill and Herbert Spencer, dwelling on their inability to account either for the facts of perception or for the appearance of natural law, and still less for the data of ethical and aesthetic experience. So far from being rational, they degrade reason to the status of an evolutionary by-product and ignore the importance of belief. The latter, it is argued in a famous chapter, is founded, not on induction, but on the more enduring basis of "authority"—the climate of traditional opinion, by which all reasonable men live. Where nothing is certain and everything rests on belief, science not only cannot dictate to religion, but even presupposes theism as the basis for its own claims to rationality.

If Balfour's strictures on naturalism were not infrequently mistaken by his opponents for a Tory attack upon science, his defense of the faith tended equally to unnerve the faithful who distrusted its appearances of skepticism. So far as these misunderstandings resulted from his own rather casual employment of such terms as *naturalism, rationalism, theism, reason, authority,* and the like, they were clarified, in part, by his two sets of Gifford Lectures, *Theism and Humanism* (London, 1915) and *Theism and Thought* (London, 1923). These works, however, though readable enough as a restatement of his position, are essentially products of a bygone phase of controversy and have little to add that is new.

See also Berkeley, George; Burke, Edmund; Hume, David; Mansel, Henry Longueville; Mill, John Stuart; Naturalism; Sidgwick, Henry.

Bibliography

Balfour's minor writings are represented in *Essays and Addresses* (London, 1905) and *Essays, Speculative and Political* (London: Hodder and Stoughton, 1920). See also an anthology by his secretary, W. M. Short, *A. J. Balfour as Philosopher and Thinker* (London: Longmans, Green, 1912).

The leading biographies are Mrs. Dugdale (his niece), *Arthur James Balfour,* 2 vols. (London: Hutchinson, 1936), and K. Young, *Arthur James Balfour* (London: G. Bell, 1963). The former has an appraisal of his philosophy by A. Seth Pringle-Pattison, an old friend, whose *Man's Place in the Cosmos* (London: Blackwood, 1897) also contains a useful appreciation of Balfour's earlier point of view.

P. L. Heath (1967)

BALGUY, JOHN

(1686–1748)

John Balguy, the English theologian and moral philosopher, was born in Sheffield and educated at the Sheffield grammar school and at St. John's College, Cambridge. He was admitted to the B.A. in 1706, ordained in the established church in 1710, and granted the living of Lamesley and Tanfield in Durham in 1711. Later he was made a prebendary of Salisbury (1727) and finally vicar of Northallerton, York (1729). He was an associate of Bishop Benjamin Hoadley and was the bishop's defender in the Bangorian controversy. Hoadley was the close friend of Samuel Clarke.

Balguy's first piece of moral philosophy was an attack on the philosophy of Shaftesbury, titled *A Letter to a Deist concerning the Beauty and Excellency of Moral Virtue, and the Support Which It Receives from the Christian Religion* (London, 1726). His most important work was *The Foundation of Moral Goodness* (Part I first published in London in 1728, Part II in 1729). Part I is a criticism of the moral philosophy of Francis Hutcheson and an exposition of Balguy's own views, much influenced by Samuel Clarke. Part II is a set of critical queries with Balguy's answers. A Lord Darcy, an admirer of Hutcheson's philosophy, is said to have proposed the queries.

Hutcheson claimed that we distinguish between virtue and vice by means of the perceptions of a moral sense. These perceptions are kinds of pleasure and uneasiness, and they are invoked to account for our approval of virtue and our abhorrence of vice, as well as our obligation to behave virtuously and to avoid viciousness. Hutcheson believed that our moral sense has been determined by God to operate as it does and that we are

naturally endowed with a benevolence toward our fellow creatures.

Balguy agreed that God has endowed our minds with benevolent affections toward others, but these affections are only helps or incentives to virtue and not the true ground or foundation of it. By making virtuous behavior flow from divinely founded instincts, Hutcheson had made virtue arbitrary. It is compatible with Hutcheson's view that God might have made us different from what we are, even inverting virtue and vice if he pleased. What is more, if God had not given us an instinct for benevolence, it appears that we should be altogether incapable of virtue; and this would be so even if we were possessed of reason and liberty.

Balguy argued that there is something in actions absolutely good (or bad) that is antecedent to both affections and laws. If this were not so, no reason could be given for God's preferring us to act benevolently and disposing us accordingly. For an action to be virtuous, there must be a perception or a consciousness of its reasonableness, or we would have to admit that beasts can be virtuous. Genuine goodness consists in our being determined to do a good action merely by the reason and the right of the thing. This is the purest and most perfect virtue of which any agent is capable. The obligation to perform a virtuous act is to be found in its reasonableness, and for a rational creature to refuse to be reasonable is unthinkable.

Balguy's elucidation of "reasonable" is found in his account of our knowledge of virtue. He argued that our understanding is altogether sufficient for the perception of virtue. Virtue is the conformity of our moral actions to the reasons of things; vice is the contrary. Moral actions are actions directed toward some intelligent being, and Balguy called them moral to distinguish them from other kinds of action. By a moral action's conformity to reason, Balguy meant the agreeableness of the action to the nature and circumstances of the persons concerned and the relations existing between them. Gratitude is an example of what he meant by conformity to reason: "We find … that some actions are agreeable, others disagreeable, to the nature and circumstances of the agent and the object, and the relations interceding between them. Thus, for instance, we find an agreement between the gratitude of A and the kindness of B; and a disagreement between the ingratitude of C and the bounty of D. These agreements and disagreements are visible to every intelligent observer, who attends to the several ideas" (*The Foundation of Moral Goodness*). He likens our perception of such an agreement to our perception of the agreement between the three angles of a triangle and two right ones, or our perception of the agreement between twice three and six. Since we do not require an intellectual sense superadded to our understanding in order to perceive these mathematical agreements, then clearly we do not require a moral sense to perceive the agreement of A's gratitude and B's kindness.

There are difficulties in Balguy's account of virtue as conformity to reason. The agreement between twice three and six is an equality, which is logically necessary. But the agreement of A's gratitude and B's kindness is not a defined equality. How, then, does the agreement come about? One of Balguy's synonyms for "agreement" is "fitting," and it appears to let the proponents of the moral sense in at the back door. For why is gratitude a fitting response to kindness and a lack of gratitude unfitting? What can we say but that we feel gratitude to be fitting and the lack of gratitude unfitting? "Fitting" and "unfitting" are normative terms, and while one can learn such a rule as "Gratitude is the fitting response to kindness," the rule must originally have been given life by someone's feeling that gratitude is the fitting response to kindness. Balguy would treat the rule as an end in itself, because he believed it exhibits some inherent self-consistency. The proponents of the moral sense would argue that the consistency of gratitude and kindness lies not in them but in us who find them to be consistent.

Balguy would agree, of course, that it is we who find gratitude to be the fitting response to kindness. The dispute is only over how we find it to be fitting, and we find it so not by a moral sense as by using our reason or understanding. The final defense for this contention is Balguy's assessment of reason as the noblest of our faculties, superior to any sense. Therefore, reason must be the arbiter of virtue and vice. The question of what faculty assesses the relative superiority of our faculties is never asked.

Balguy also wrote *Divine Rectitude: or a Brief Inquiry concerning the Moral Perfections of the Deity, Particularly in Respect of Creation and Providence* (London, 1730). He argued that God's goodness follows from a regard for a real and absolute order, beauty, and harmony.

See also Clarke, Samuel; Ethics, History of; Hutcheson, Francis; Moral Sense; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Virtue and Vice.

Bibliography

L. A. Selby-Bigge, ed., *The British Moralists* (Oxford: Clarendon Press, 1897), Vol. II, reproduces Part I of *The Foundation of Moral Goodness* and representative selections from Part II.

For critical discussion, see B. Peach, "John Balguy," in *Encyclopedia of Morals,* edited by V. Ferm. (New York: Greenwood, 1956).

Elmer Sprague (1967)

BÁÑEZ, DOMINIC

(1528–1604)

Dominic Báñez, the Spanish theologian, was born at Valladolid and died at Medina del Campo. He studied at the University of Salamanca, where he entered the Dominican order. He first taught courses in philosophy and theology in various houses of study of his order in Spain (Salamanca, Ávila, Alcalá de Henares, Valladolid) and then became a professor at the University of Salamanca, teaching philosophy from 1577 and theology from 1581. He was noted for his role as the spiritual director of St. Teresa of Ávila and for his bitter controversy with the Jesuit Luis de Molina concerning divine grace. Báñez's view on grace and human liberty is called "physical predetermination," which means that man's will is unable to act unless empowered and applied to action by an ultimate principal cause, which is God. Apart from a commentary on Aristotle's treatise *On Generation and Corruption* (1585), Báñez's philosophy is found in his theological work *Scholastica Commentaria in Primam Partem Angelici Doctoris* (Commentary on the first part of the summa of theology; 2 vols., Salamanca, 1584–1588). As a philosopher, Báñez was at his best in interpreting the metaphysics of St. Thomas. Unlike most of his contemporaries, he saw the importance of the act of being (*esse*) as constituting every nature in existence (see L. Urbano, ed., *Scholastica Commentaria,* I, p. 141). In this he anticipated the existential view of Thomistic metaphysics now favored by such thinkers as Jacques Maritain and Étienne Gilson. On the other hand, Báñez interpreted the real distinction of essence and existence as the difference between two individual things (*res*) and then rejected this notion. Moreover, he regarded the limitation of the act of existing by the essence that receives it as an indication that essence may, in this sense, be more noble than existence.

See also Aristotle; Being; Essence and Existence; Gilson, Étienne Henry; Maritain, Jacques; Molina, Luis de; Teresa of Ávila, St.; Thomas Aquinas, St.

Bibliography

Works by Báñez include *Scholastica Commentaria,* newly edited by L. Urbano (Madrid and Valencia, 1934). A later section of the same *Commentary* has also been published as *Commentaria in Primam Secundae,* edited by B. de Heredia, 2 vols. (Madrid, 1942–1944).

For works on Báñez see W. R. O'Connor, "Molina and Báñez as Interpreters of Thomas Aquinas," in *New Scholasticism* 21 (1947): 243–259; and L. Gutiérrez-Vega, "Báñez filósofo existencial," in *Estudios Filosóficos* 3 (1954): 83–114.

Vernon J. Bourke (1967)

BANFI, ANTONIO

(1886–1957)

Antonio Banfi, the Italian philosopher, was born in Milan and studied at the Academy of Science and Letters there and at the University of Berlin. Banfi enjoyed a long acquaintance with Edmund Husserl, who influenced Banfi's thought along with the Marburg Neo-Kantians. Banfi taught at the universities of Florence, Genoa, and Milan. In 1940 he founded the review *Studi filosofici,* which played an important part in the Italian revolt against idealism. Banfi participated actively in political life. In 1925 he adhered to the manifesto of the antifascist intellectuals prepared by Benedetto Croce. After World War II he sat in the Italian Senate as a Communist.

German rather than Italian influences are apparent in Banfi's major work, *Principi di una teoria della ragione* (Principles of a theory of reason; Milan, 1926). According to Banfi philosophical inquiry does not spring from an immediate spontaneity of thought but arises as critical reflection on the cultural heritage of the speculative tradition. By studying the structures of knowledge, reflection grasps the function of reason. Reason is to be understood neither in a psychological sense nor in the metaphysical sense of Hegelianism. Reason, according to Banfi, is the indefinite law of the process of organization or of coordination of experience.

The task of science, Banfi held, is to study experience and resolve it into functional relations or laws. Philosophy continues the work of science in its own manner. It clarifies experience in terms of dialectical antitheses (reality and appearance, matter and form, necessity and liberty, and so on); it resolves the opposition of the antitheses in the unity of an idea; and in the phenomenological conclusion it discloses the rational structure progressively attained in the ordering of experience.

In subsequent works Banfi sought to emphasize the problematic nature of reason as an open system and as the self-ordering of experience. He saw in dialectical materialism the elimination of the mythical moment of knowledge, the affirmation of the unending development of reason, and the liberative function of reason.

See also Croce, Benedetto; Dialectical Materialism; Experience; Hegelianism; Husserl, Edmund; Idealism; Neo-Kantianism; Reason.

Bibliography

Banfi's *L'uomo copernicano* was published in Milan in 1950. His two-volume *La ricerca della realtà* was published in Florence in 1959.

For literature on Banfi, see Giovanni Maria Bertin, *L'idea pedagogica e il principio di ragione in Antonio Banfi* (Rome: A. Armando, 1961); and Fulvio Papi, *Il pensiero di Antonio Banfi* (Florence, 1961).

Eugenio Garin (1967)
Translated by Robert M. Connolly

BARTH, KARL
(1886–1968)

Karl Barth, the Swiss theologian, was born in Basel in 1886. He held professorships at Göttingen, Münster, Bonn, and Basel. His impact on the theological world dates from 1921, with the substantially revised second edition of his *Der Römerbrief* (the first edition was published in 1919). Herein he attacked the prevalent "subjectivism" of Protestant theology, in which he perceived the attempt to fit the Christian revelation into the mold of human preconceptions. After that, though Barth changed and developed many of his ideas, a single main concern ran through all his writings: namely, how to prevent theology from becoming an ideology, that is, a creation of human culture. This was the reason for his early violent attacks on the then fashionable liberal theology, as expounded, for instance, by Adolf von Harnack. According to Barth, the danger of such attempts to formulate a "reasonable" Christianity is threefold: intellectual, ethical, and soteriological. First, there is the danger of identifying human conclusions with the Word of God and thus of destroying the validity of the concept of revelation, which is God's self-manifestation and owes nothing to human initiatives. Second, there is the danger that the church will simply reflect the social and cultural situation, thus losing its power of criticism and its prophetic function. Barth was deeply disturbed by the support given to the kaiser by a number of his liberal theologian teachers in 1914. It is notable that, while at Bonn, he threw his support behind the Confessing church in its opposition to the Nazis, an action that cost him his chair. Third, salvation comes from God alone, and the attempt to identify a human Weltanschauung with God's Word is an instance of the refusal to accept that the only justification is by grace. As Barth wrote: "This secret identification of ourselves with God carries with it our isolation from him."

The principle that theological exposition should be basically independent of human speculations (except insofar as historical and linguistic investigations, etc. are a necessary part of understanding Scripture) was reinforced by Barth's interpretation of the Fall. Not only is the human will vitiated by the Fall, but reason also, in such a way that it is impossible for men to discover the truth about God through their own efforts. Only if God manifests himself can there be any revelation. Thus Barth rejected the whole of natural theology as expounded by, for instance, Aquinas, and in particular its basis in the doctrine of the analogy of being (*analogia entis*), on the ground that it implies some similarity between creatures and God. A strong motif in Barth's theology, therefore, is the transcendence of God (in the sense of his distance from creatures—"the great Calvinist distance between heaven and earth"). Methodologically, all this implies that interpretation of the Bible should not betray the genuine meaning of the text by explaining away or avoiding those hard sayings that are supposedly scandals to modern thought. Nevertheless, Barth is no fundamentalist: The Word of God is not to be identified with the witness to it found in the Bible, and there is no question of using the latter as a "paper pope."

Der Römerbrief was critical rather than constructive, and during the 1920s Barth's theology had the character of being dialectical (to use a term that he later came to reject), that is, it called in question human preconceptions about God, often by denying them in the sharpest terms; but since theology is designed to proclaim what is God-given, it is always necessary to reach out beyond such denials. In this way, there is a constant dialectic between grace and man's religion. The concept that religion itself is under divine judgment, and is a human rather than strictly a divine phenomenon, has had great influence, culminating in Dietrich Bonhoeffer's idea of a "religionless Christianity."

In the late 1920s Barth started on the second main phase of his theological writing, and after what he called his "well-known false start," with the *Prolegomena to a Christian Dogmatics* (*Christliche Dogmatik im Entwurf,*

1927), he began on his many-volumed *Church Dogmatics* (*Die kirkliche Dogmatik,* 1932 and onward). Herein he was influenced by his study of Anselm (expressed in *Fides Quaerens Intellectum,* 1931). The heart of the Ontological Argument is the recognition that theology does not need any metaphysical substructure; it contains within itself its own rationale, namely the unfolding of the inner form of God's Word. Thus dogmatics is systematic in that it presents the material in an orderly way and in that it aims exhaustively to touch on all areas of human concern, but it is not a deduction from some principle or set of principles.

The *Church Dogmatics* is a rich work, though not altogether a consistent one, since Barth's thought was developing in the course of his writing. Its main emphasis is Christocentric. God's revelation is essentially seen in the Christ-event, and Christ is God's Word. However, the God so revealed is trinitarian: "the work of the Son of God includes the work of the Father as its presupposition and the work of the Holy Spirit as its consequence." The first article, the work of the Father, is "to a certain extent the source, the third article, the work of the Holy Spirit, the goal of our path. But the second article, the work of the Son, is the Way upon which we find ourselves in faith. From that vantage we may review the entire fullness of the acts of God." Consequently, such doctrines as creation must be seen from this perspective. The Bible presents no cosmology, but it does contain an anthropology; and thus God's relation to the natural world can only be understood by analogy with his saving revelation to human beings. Notions of a First Cause and Necessary Being, as explaining the existence of the cosmos, are thus beside the point, for they make no use of the concepts of grace and personality as ascribed to God. By contrast, the biblical saga of creation makes it continuous with God's covenant relationship with Israel.

Barth's exposition is controlled throughout by two considerations. First, dogmatics is necessarily church dogmatics, that is, it is an activity that must be carried on within the church, as the place where the preaching or proclamation of the Word occurs. Thus the theologian's continuous concern is to test the doctrine and preaching of the church, which, because it is carried on through human beings, is liable to go astray. Second, the standpoint from which the proclamation is tested is that of Scripture, which is "the document of the manifestation of the Word in Jesus Christ." Dogmatics would become irrelevant if it sacrificed this standard.

The implications of Barth's thesis for the relationship between philosophy and theology are clear. Insofar as philosophy is metaphysical, in the sense of saying something about God or some such substitute as the Absolute, it collides with theology; and it is the theologian's proper task to show how metaphysics has here gone beyond its legitimate limits. Philosophy, as logic, philosophy of science, and so on, is a proper inquiry, but one that is quite separate from theology. Barth does, however, allow (in his *Fides Quaerens Intellectum* and elsewhere) that philosophical concepts may be used in exegesis, so long as they are kept strictly subordinate to the Word of God. But Barth remains insistent that theologians should not make concessions to secular thought; indeed, he holds that such concessions are a principal reason for the contempt that many philosophers have had for "philosophical" theologians. Thus traditional forms of apologetic are ruled out.

Two issues arising from Barth's whole approach are crucial. First, how is one to know that the revelation in Christ is the true one? Or more particularly, how is one to know that the whole doctrine of God as expounded by Barth is true? Second, how can these propositions about God be meaningful if the similarity or analogy between God and human persons is denied? For Barth, the first question is one that virtually does not arise. The Bible, for instance, does not set out to prove God's existence or attributes, rather, it witnesses to his acts. The task of the preacher or theologian is to proclaim this revelation. Theology must be a rational inquiry that is appropriate to its subject matter, namely God's gracious self-revelation; and any attempt to establish the truth of doctrine upon grounds that are extraneous to its subject matter is both irrelevant and dangerous. Thus the Christian message is not to be seen as a religious teaching amid rival teachings, for all religious and metaphysical revelations and conclusions are projections of human wishes (here the influence of Ludwig Feuerbach is apparent). It by no means follows, however, that any particular statement of theology that is consistent with these presuppositions as to the nature of theological inquiry is correct. Barth holds that dogmatics is a continuing process within the church, and it is, of course, a human activity suffering from the defects of human reason. It is therefore necessary to consider the criteria of the worth of a system of dogmatics. These criteria are necessarily derived internally from God's self-revelation (by the former arguments). Barth singles out two. First, theological thinking must be humble: this is a practical test of whether it is refraining from establishing its own claim to truth, i.e., its being in effect an ideology. Second, it must express the doctrine of predestination, which encapsulates the whole of the revelational approach—what man "achieves" in relation to God is due to God. Because of the element of paradox in the first cri-

terion (for the Thomist can be humble in his approach), Barth is at times inclined to speak in a syncretistic way. Imagining a conversation in heaven, he says: "Yes, dear Schleiermacher, I understand you now. You were right, except on some points!" (*Karl Barth's Table Talk*). Further, the notion that theology is dialectical, so that a statement can be balanced by affirming its apparent contradictory, has rendered Barth less rigid than many of the Barthians.

As to the problem of the meaning of theological utterances, Barth holds that revelation is a relational concept, and thus God does not, so to say, reveal himself independently of the human apprehension of his self-manifestation. Consequently, the knowledge of God is itself given by God, through grace. Thus, the *analogia entis* is replaced by the *analogia fidei* (the analogy of faith); faith gives us understanding of the nature of God and is God-given. Thus God is the cause of true theological assertions, as well as their ground.

Barth's influence has been great. This is partly because he has provided the outline of a theology that is powerfully biblical without being fundamentalist and, therefore, can escape the charge of being irrational by being nonrational. The most eminent Europeans who stand close to Barth are Emil Brunner and Oscar Cullmann. The former entered into controversy with Barth in the early 1930s over the question of the fallen character of human reason. Brunner held that in some areas this thesis was obviously false, for example, in the natural sciences; but, nevertheless, in relation to knowledge of God, men are capable of only the most shadowy awareness on their own. One of the most important attempts to apply Barth's theology has been Hendrik Kraemer's *The Christian Message in a Non-Christian World* (1938), which aims to show that all religions, including empirical Christianity, are under the judgment of the revelation in Christ. Thus there is no need to argue for Christianity as an empirical phenomenon as against other religions. But the question remains: If there is no correspondence between the Gospel and empirical Christianity, the church is a sham; and if there is, then the comparison and contrast between empirical Christianity and other faiths is possible, and apologetics unavoidable. This is one illustration of the central problem posed by Barth's theology.

See also Anselm, St.; Brunner, Emil; Creation and Conservation, Religious Doctrine of; Feuerbach, Ludwig Andreas; Harnack, Carl Gustav Adolf von; Ontological Argument for the Existence of God; Revelation; Thomas Aquinas, St.

Bibliography

There is a useful bibliography of Barth's works in T. F. Torrance's *Karl Barth, an Introduction to His Early Theology, 1910–1931* (London: SCM Press, 1962). The main works translated into English are *The Epistle to the Romans* (London: Oxford University Press, 1933); *Church Dogmatics,* edited by G. W. Bromiley and T. F. Torrance (London: T. and T. Clark, 1936; Edinburgh, 1956 onward), 4 vols., comprising 11 separately published sections, one of which is *Church Dogmatics: A Selection with Introduction,* edited by H. Gollwitzer and translated by G. W. Bromiley (Edinburgh: T. & T. Clark, 1961). An outline, by Barth himself, of the thought in this work is *Dogmatics in Outline* (London: Philosophical Library, 1949). Later works include *The Humanity of God* (Richmond, VA: John Knox Press, 1960); and *Anselm: Fides Quaerens Intellectum* (Richmond, VA: John Knox Press, 1960); and *Evangelical Theology: an Introduction* (New York: Holt, Rinehart and Winston, 1963). See also *Karl Barth's Table Talk,* recorded and edited by John D. Godsey (Edinburgh, 1963).

Works expressing related thought are E. Brunner, *Dogmatics,* 2 vols. (London, 1949; 1952); Oscar Cullmann, *Christ and Time* (London, 1951); and H. Kraemer, *The Christian Message in a Non-Christian World* (London: Edinburgh House Press for the International Missionary Council, 1938). Philosophical criticisms of Barth's position can be found in H. D. Lewis, *Morals and Revelation* (London: Allen and Unwin, 1951) and Ronald W. Hepburn, *Christianity and Paradox* (London: Watts, 1958), Ch. 5.

Ninian Smart (1967)

BARTH, KARL [ADDENDUM]

Since his death in 1968, time and distance have provided scholars with space to understand Karl Barth in a larger intellectual and cultural context. He has emerged as one of the most important Christian theologians of the twentieth century—perhaps the most important—while his massive theological oeuvre and the changing shape of his thought have generated a host of alternative interpretations, notably in respect to his understanding of and relationship to Western philosophy. Contemporary theologians have sought to appropriate Barth in several directions, exploring his thought in connection with various postmodern positions.

Barth consistently held that philosophy should not hold sway over theology. In a 1960 essay written in honor of his brother, who was a philosopher, Barth allows that theology and philosophy can coexist in harmony, but he also spells out the important differences between these disciplines. The Christian theologian must be held captive to the Word of God, he contends, for only God's revelation in Christ provides us with the key to under-

standing divinity; biblical theology, not philosophical reasoning, is the basis for Christian theology.

At the same time, Barth's thinking was influenced by European philosophers. The influence of existentialist thinkers (especially Kiekegaard) on Barth has long been acknowledged, even by Barth himself. Barth also read and responded to Heidegger in his own way. In recent years the importance of the Marburg school of Neo-Kantian philosophy has become more clear, especially in Barth's early development. Thus Barth's appreciation for philosophy is more expansive than had been acknowledged in earlier scholarship.

Contemporary scholarship has fruitfully engaged Barth's thought with larger philosophical concerns, bringing him into a larger orbit. Much of this research has brought to light Barth's critique of modernity and his ambivalence toward language as a vehicle for theology. Several so-called postliberal theologians have appropriated Barth as a narrative theologian who sought to read the rest of the world in terms of the biblical story. Here Barth is sometime brought into conversation with the later Wittgenstein, both in terms of an understanding of language and a critique of enlightenment rationalism. Most recently scholars have developed some of the parallels between Barth and postmodern philosophers, especially Derrida. This school seeks to appropriate Barth for postmodern theology, a move roundly criticized by more traditional Barth experts. Thus Barth remains at the center of contemporary theological debate.

See also Derrida, Jacques; Enlightenment; Heidegger, Martin; Kierkegaard, Søren Aabye; Neo-Kantianism; Rationalism; Wittgenstein, Ludwig Josef Johann.

Bibliography

Bush, E. *Karl Barth*. London: SCM, 1976.

Fisher, Simon. *Revelatory Positivism? Barth's Earliest Theology and the Marburg School*. Oxford: Oxford University Press, 1988.

Ford, David. *Barth and God's Story*. Frankfurt: P. Lang, 1981.

Frei, Hans. *Types of Christian Theology*. New Haven, CT: Yale University Press, 1992.

Johnson, W. Stacy. *The Mystery of God*.Louisville: Westminster John Knox, 1997.

MacDonald, Niel B. *Karl Barth and the Strange New World within the Bible*. Carlisle, U.K.: Paternoster, 2000.

Lowe, Walter. *Theology and Difference*. Bloomington: Indiana University Press, 1993.

McCormack, Bruce. *Karl Barth's Critically Realistic Dialectical Theology*. Oxford: Oxford University Press, 1995.

Ward, Graham. *Barth, Derrida and the Language of Theology*. Cambridge, U.K.: Cambridge University Press, 1995.

Alan G. Padgett (2005)

BARTHES, ROLAND

(1915–1980)

Ronald Barthes was a French writer most widely known for declaring "the death of the author." It is ironic, then, in a way Barthes would surely appreciate, that his *Œuvres completes* fill nearly 6,000 pages with the unmistakable observations, distinct voice, and style that shaped the form and content of what came to be known as "cultural studies." He was sixty-five years old in 1980 when a laundry truck struck him down in a street in front of the College de France. He died of his injuries four weeks later.

Barthes was born in November 1915, in Cherbourg. His father died before his first birthday, and he was raised by his mother and paternal grandparents in coastal Bayonne. Normal progress to a university degree was blocked by the onset of tuberculosis. Over the course of ten years convalescing in and out of sanatoria, Barthes earned advanced degrees in Greek and Latin, performed in the Ancient Theater Group, and taught French in Romania and in Egypt where A. J. Greimas introduced him to linguistics. He gained his first regular academic post at the *Écoles practique des haute etudes* in 1962 on the basis of his publications *Le degré zéro de l'écriture* (1953), *Michelet par lui-même* (1954), and *Mythologies* (1957). He gained wider public notice with the publication of *Le plaisir du texte* (1973), a critical erotics of reading pleasures, and *Roland Barthes par Roland Barthes* (1975), an autobiography prefaced, as it were, on the page ordinarily reserved for a dedication with the handwritten remark, "It must all be considered as if spoken by a character in a novel." He was appointed Chair of Literary Semiotics at the College de France in 1976 where he lectured until his death.

Barthes's contributions to philosophy fall under four headings defined, in each case, by pairs of opposed terms: mythology (nature/culture), semeiology (*langue/parole*), structuralism (reading/writing), and hedonism (*plaisir/jouissance*).

MYTHOLOGY

Myth today, according to Barthes, is found in a conflation of nature and culture or, more specifically, in the production and consumption of culture as nature. In his most famous example, it is no accident that the scurrilous

competitor in a professional wrestling match is bested by the fair play of his adversary: his foul play (as the "fairness" of the victor) is fabricated to stage the "natural" and inevitable triumph of "good" over "evil." Again, in Parisian striptease, the *artiste* sheds layers of patently cultural trappings—feathers, furs, and exotic costumes—to reveal her naked body as the "natural" state of woman unnaturally desexualized, in this act, to forgive the voyeur and the culture that condones his voyeurism for their sins. In modern myths, an apparently natural meaning contains the form of a cultural signification whose content discloses the artifice of what is "natural" in appearance only. "Demythologization" was the name given to the critical practice of exposing these myths.

SEMEIOLOGY

Barthes's literary semeiology follows Ferdinand de Saussure's distinction between *la langue*, the syntactic and semantic paradigms that define the language one learns, and *parole*, the series of signifying acts that compose the language one speaks. On this model, meaning is the product of a system of distinctions and conventions, found in *la langue*, which anchor otherwise unruly syntagms of signifying units, articulated as *parole*. The meaning of this sentence, for example, depends on identifying the parts of speech in it and the rules governing their use that define the linguistic system in which the sentence is uttered. Reversing Saussure, semeiology was, for Barthes, a subset of linguistics, a science of the signifying function of language. In his studies of advertising, gastronomy, fashion and Japan, Barthes consistently emphasized the multiplicity and variability of the signifier over the system that governed its significations.

STRUCTURALISM

As Barthes defined it, structuralism studies the rules, norms, and organizing structures that make meaning possible. These structures are the products of cultural practices, which the structuralist uncovers beneath the singular meaning attributed to an image, an artifact, or a text. It is Author who could authorize a Single meaning (the capital letters standing for the "theological" authority supposed by such a concept of signification) who dies in Barthes's analysis of the rules, norms, and organizing structures, of the narrative and social and moral codes that govern the writing (literal and figurative) of a text or any other cultural artifact. In addition, this writing is governed by the rules, norms and structures of reading. So that writing, *écriture*, arranges a meeting of the structures and codes that have formed a writer and a reader and stages the multiplication of meanings sustained by the text a writer and reader share. Barthes calls a text "writerly" which invites the reader to write meanings into it and "readerly" when the text insists on a single authorial intention.

HEDONISM

Our pleasures, in Barthes's writing, are divided along the same lines. There is, on the one hand, *plaisir*, the warmth of sensation that opposes cold abstraction, the contentment, euphoria and delectation that relieve the method, commitment and science of the intellect. It is found in texts of and on pleasure (Gustave Flaubert and Marquis de Sade, for example) and connected to a reading practice that is comfortable and continuous with the culture of the reader and the text. There is, on the other hand, the ecstatic pleasure of *jouissance*, a feeling of enjoyment characterized by a state of loss. It is not centered in the heart (as opposed to the head) but spread sensuously across the entire surface of the body. *Jouissance* is found in a reading practice that "cruises" the text, skipping passages anticipated as "boring," looking up distractedly to consider ideas associated with the body and dissociated from the culture of the reader or the text. *Jouissance* is found in distinctly "writerly" readings and texts that multiply meanings for the sheer pleasure of it.

There is, finally, a distinctive normative orientation in Barthes's writings. While he did not author or advocate an alternative, single meaning of culture, Barthes did license and exhort readers to take ecstatic pleasure in multiplying the meanings of culture and in rewriting the authority of its hegemonic codes.

See also Hedonism; Language; Myth; Structuralism and Post-structuralism.

Bibliography

WORKS BY ROLAND BARTHES

Le degré zéro de l'écriture. Paris, 1953. Translated by Annette Lavers and Colin Smith as *Writing Degree Zero* (New York: Hill and Wang, 1968).

Michelet par lui-même. Paris: Éditions du Seuil, 1954. Translated by Richard Howard as *Michelet* (Berkeley: University of California Press, 1992).

Mythologies. Paris, 1957. Selected translation by Annette Lavers as *Mythologies* (London: Paladan Books, 1973). Translation of remaining essays by Richard Howard as *The Eiffel Tower and Other Mythologies* (New York: Hill and Wang, 1979).

Sur Racine. Paris, 1963. Translated by Richard Howard as *On Racine* (New York: Octagon Books, 1977).

Essais critiques. Paris, 1964. Translated by Richard Howard as *Critical Essays* (Evanston: Northwestern University Press, 1972).

Éléments de sémiologie in Le degré zéro de l'écriture, suivi de: Éléments de sémiologie. Paris, 1965. Translated by Annette Lavers and Colin Smith as *Elements of Semiology* (New York: Hill and Wang, 1968).

Critique et verité. Paris, 1966. Translated by Katrine Kueneman as *Criticism and Truth* (Minneapolis: University of Minnesota Press, 1987).

Système de la mode. Paris, 1967. Translated by Matthew Ward and Richard Howard as *The Fashion System* (New York: Hill and Wang, 1983).

L'empire des signes. Geneva, 1970. Translated by Richard Howard as *Empire of Signs* (New York: Noonday Press, 1989).

S/Z. Paris: Éditions du Seuil, 1970. Translated by Richard Miller as *S/Z* (London: Cape, 1975).

Sade, Fourier, Loyola. Paris: Éditions du Seuil, 1971. Translated by Richard Miller as *Sade, Fourier, Loyola* (Baltimore: Johns Hopkins University Press, 1997).

Nouveaux essais critiques in Le degré zéro de l'écriture suivi de: Nouveaux essais critiques. Paris, 1972. Translated by Richard Howard as *New Critical Essays* (New York: Hill and Wang, 1980).

Le plaisir du texte. Paris, 1973. Translated by Richard Miller as *The Pleasure of the Text* (New York: Hill and Wang, 1975).

Roland Barthes par Roland Barthes. Paris, 1975. Translated by Richard Howard as *Roland Barthes by Roland Barthes* (Berkeley: University of California Press, 1994).

Et la Chine? Paris, 1976.

Fragments d'un discours amoureaux. Paris, 1977. Translated by Richard Howard as *A Lover's Discourse: Fragments* (New York: Hill and Wang, 1978).

Image, Music, Text. Essays selected and translated by Stephen Heath. New York: Hill and Wang, 1977.

Sollers ecrivain. Paris, 1979. Translated by Philip Thody as *Sollers Writer* (London: Athlone, 1987).

La chambre claire: note sur la photographie. Paris, 1980. Translated by Richard Howard as *Camera Lucida: Reflections on Photography* (New York: Hill and Wang, 1981).

Le grain de la voix: entretiens 1962–1980. Paris, 1981. Translated by Linda Coverdale as *The Grain of the Voice: Interviews 1962–1980* (New York: Hill and Wang, 1985).

A Barthes Reader. Edited and with an Introduction by Susan Sontag. New York: Hill and Wang, 1982.

L'obvie et l'obtus. Paris, 1982. Translated by Richard Howard as *The Responsibility of Forms: Critical Essays on Music, Art, and Representation* (New York: Hill and Wang, 1985).

Le bruissement de la langue. Paris, 1984. Translated by Richard Howard as *The Rustle of Language* (New York: Hill and Wang, 1986).

L'aventure sémiologique. Paris, 1985. Translated by Richard Howard as *The Semiotic Challenge* (New York: Hill and Wang, 1988).

Incidents. Paris, 1987. Translated by Richard Howard as *Incidents* (Berkeley: University of California Press, 1992).

Œuvres completes. 5 vols. New, revised, and corrected ed. Paris: Seuil, 2002.

WORKS ABOUT ROLAND BARTHES

Culler, Jonathan. *Barthes: A Very Short Introduction*. Oxford: Oxford University Press, 2002.

Heath, Stephen. *Vertige du déplacement: Lecture de Barthes*. Paris: Fayard, 1974.

Lavers, Annette. *Roland Barthes: Structuralism and After*. Cambridge, MA: Harvard University Press, 1982.

Miller, D. A. *Bringing Out Roland Barthes*. Berkeley: University of California Press, 1992.

John Carvalho (2005)

BASEDOW, JOHANN BERNHARD

(1724–1790)

Johann Bernhard Basedow, the German philosopher, theologian, and educational theorist, was born in Hamburg into the family of a poor wigmaker, whose name, more properly, was Bassedau. A benefactor financed his studies, first at Hamburg under H. S. Reimarus. In 1746 he entered the faculty of theology at Leipzig University, where he studied philosophy under the Pietist philosopher C. A. Crusius. In 1749 he became a private tutor in the family of Herr von Quaalen in Holstein. His experiences as a tutor turned his attention to educational problems, which were the subject of his master's thesis at Kiel University in 1752. On Friedrich Gottlieb Klopstock's recommendation, he was appointed professor of philosophy and rhetoric at the Knightly Academy at Soro, Denmark. A heterodox work, *Praktische Philosophie für alle Stände* (Practical philosophy for all states; Copenhagen, 1758), led to his dismissal. In 1761 he moved to the gymnasium at Altona, but again lost his position, and his writings were prohibited. He left theology and, supported by his benefactor, published his *Vorstellung an Menschenfreunde für Schulen, nebst dem Plan eines Elementarbuchs der menschlichen Erkenntnisse* (Appeal to the friends of mankind about schools, with a plan for an elementary book on human knowledge; Hamburg, 1768), his first significant work on education, which met with a tremendous response. With financial help from several influential people, he published during the following years several textbooks, the most important being his *Methodenbuch für Väter und Mütter der Familien und Völker* (Methodology for fathers and mothers of families and nations; Leipzig, 1770; edited by T. Fritzsch, Leipzig, 1913). Prince Franz Leopold Friedrich of Dessau invited him to organize an experimental school in Dessau. Basedow accepted, and the school, called the *Philanthropin,* opened in 1774.

It was soon imitated by a number of similar institutions in Germany and Switzerland.

By 1776 Basedow had returned to theology, living in Dessau, Leipzig, Halle, and Magdeburg. During this period he published his *Examen in der alten natürlichsten Religion* (Examination of the old most natural religion), which he considered his masterpiece. Basedow's theological ideas, inspired by the English and French deists, aimed at a natural religion, rational and practical, refraining from dogmas and rejecting every kind of orthodox Christianity.

Basedow was one of the "popular philosophers" (*Popularphilosophen*), but his importance as a theoretical philosopher has been underrated by modern historians. His work on theory of knowledge and metaphysics, *Philalethie* (Lübeck, 1764), inspired by Crusius, David Hume, and the French *philosophes,* was one of the most significant books on methodology of its time and influenced Immanuel Kant, Johann Nicolaus Tetens, and others. He supported a moderate skepticism based on common sense and denied the possibility of reaching absolute demonstrative truth in natural philosophy (out of skepticism concerning causation), in rational psychology, or in theology.

Basedow's chief importance lies in his educational theories, which are based on John Amos Comenius, John Locke, and Jean-Jacques Rousseau. He claimed that education should be cosmopolitan, free from any confessional imprint, equal for all classes, and aimed at enabling men to live useful and happy lives as good citizens. Instruction should appeal to the child's sensibility rather than to his understanding and should be encouraged by games and colloquial intercourse. Images (*Zeichen*) are more effective than words.

See also Comenius, John Amos; Crusius, Christian August; Deism; German Philosophy; Hume, David; Kant, Immanuel; Locke, John; Reimarus, Hermann Samuel; Rousseau, Jean-Jacques; Skepticism, History of; Tetens, Johann Nicolaus.

Bibliography

WORKS BY BASEDOW

Ausgewählte Schriften. Edited by Hugo Göring. Langensalza, 1880.

WORKS ON BASEDOW

Basedow, Armin. "Johann Bernhard Basedow." *Friedrich Manns pädagogisches Magazin* 995 (1924).

Diestelmann, Richard. *Johann Bernhard Basedow.* Leipzig, 1897.

Meyer, J. C. *Leben, Charakter und Schriften Basedows,* 2 vols. Hamburg, 1791–1792.

Rammelt, Johannes. *J. B. Basedow, der Philanthropinismus und das Dessauer Philanthropin.* Dessau, 1929.

Rathmann, H. *Beiträge zur Lebensgeschichte Basedows.* Magdeburg, 1791.

Surakoff, K. D. *Der Einfluss der zeitgenössischen Philosophie auf Basedows Pädagogik.* Giessen: Brühl, 1898.

Zimmermann, Hans. *Die Pädagogik Basedows vom Standpunkte moderner Geschichtsauffassung.* Langensalza, 1912.

Giorgio Tonelli (1967)

BASIC STATEMENTS

Any statement of fact is true or false in virtue of some existing state of affairs in the world. In many cases the truth-value of a statement is determined by appealing to the truth-values of certain other statements, but this process must terminate somewhere if the truth-value of any statement of fact is to be assessed at all. An epistemological view according to which the process of verification or falsification terminates with statements of a logically distinct kind is a view to the effect that there is a distinct class of *basic statements*. The principal questions that have been considered are (1) Is there such a class of statements? (2) If there is, what is the relation between these statements and certain nonverbal occurrences called experiences? (3) Are basic statements descriptions of the private experiences of the speaker or of publicly observable events? (4) Are these statements either incorrigible (that is, of such a character that they cannot be false, or cannot be shown to be false) or indubitable (that is, such that they cannot rationally be doubted)? These questions have been much discussed by modern empiricists, especially in connection with the verifiability criterion of meaning. The problems concerning basic statements are not, however, essentially confined to empiricist theories of meaning and truth; they are fundamental in any theory of knowledge.

WITTGENSTEIN

The thesis that there is a class of basic or elementary propositions is powerfully presented in Ludwig Wittgenstein's *Tractatus Logico-Philosophicus* (1921; first English translation, 1922). Wittgenstein argues that if a proposition contains expressions standing for complexes, the sense of the proposition will depend upon the truth of other propositions describing those complexes. This will again be the case if any one of those other propositions contains expressions standing for complexes. Thus, the

determinateness of the sense of the original proposition requires that its analysis should terminate in elementary propositions consisting only of names of simple things (see 2.0211–2.0212, 3.23). An elementary proposition is an arrangement of names that represents a possible arrangement of simple things; it is a logical picture of an elementary state of affairs. Wittgenstein gave no explicit interpretation of "simple things," "names," or "elementary propositions." He is reported as saying that at the time he wrote the *Tractatus* he thought it was not his business, as a logician, to give examples of simple things, this being a purely empirical matter; the *Tractatus* view is that the application of logic decides what elementary propositions there are (5.557).

SCHLICK

Moritz Schlick and some other members of the Vienna circle gave an empiricist interpretation to Wittgenstein's theory. In "Über das Fundament der Erkenntnis" (1934) and other articles, Schlick inquired whether there is a class of statements which provide an "unshakeable, indubitable foundation" of all knowledge. This kind of incorrigibility, he argued (against Otto Neurath and Rudolf Carnap), cannot depend simply upon the coherence of a statement with the existing system of science, nor simply upon someone's decision to accept a statement as true. It is possessed only by the statements a person makes about his own experiences. Schlick called such statements *Konstatierungen* "confirmations" and contrasted these with the "protocol sentences" described by Neurath and Carnap.

Konstatierungen have the following characteristics: (1) They have the form "here, now, so and so"; examples are "here two black points coincide," "Here yellow borders on blue," "Here now pain." (2) In the case of other synthetic statements, understanding their meaning is quite distinct from the actual process of verifying them, and their meaning does not determine their truth-value; but in the case of a *Konstatierung* (since "'this here' has meaning only in connection with a gesture … one must somehow point to reality"), the occasion of understanding it is the same as that of verifying it. Therefore a (significant) *Konstatierung* cannot be false. (3) Unlike "protocol sentences," these statements cannot be written down or recorded at all because of the fleeting reference of the demonstratives that occur in them; but they provide the occasions for the formation of protocol sentences. (4) They are the only empirical statements that are not hypotheses. (5) They are not the starting points of science in either a temporal or a logical sense, but simply the momentary consummations of the scientific process; they are the means by which all scientific hypotheses are confirmed.

The first and most obvious objection to the view that there are *Konstatierungen* (in Schlick's sense) is that it results immediately in a radical form of solipsism. It may also be objected that *Konstatierungen* are either genuine contingent statements, in which case they cannot be of such a nature that they cannot be false, or they are purely demonstrative, in which case they are not statements. Following Wittgenstein's later work, many philosophers would deny the possibility of the essentially private use of demonstratives and descriptions that are supposed to occur in *Konstatierungen*. Further, no adequate account is given of the relation between these private statements and the public protocol sentences to which they give rise. Moreover, if the *Konstatierungen* are meaningful only at the moment at which they are verified, they cannot occur in predictions, and hence it cannot be through them that scientific hypotheses are confirmed.

CARNAP

Rudolf Carnap, in "Die physikalische Sprache als Universalsprache der Wissenschaft" (1931; translated as *The Unity of Science*, 1934) and elsewhere, had at first held that science is a system of statements based upon sentences describing the experiences of scientific observers. These "primitive protocol sentences," Carnap supposed, contain no inferential or theoretical additions; they describe only what is directly given, and hence they stand in no need of any further justification. At this time Carnap left it an open question whether protocol sentences describe the simplest sensations and feelings of the observer (for example, "here now red," "joy now"), or partial or complete gestalts of single sensory fields (for example, "red circle now"), or the total experience of the observer during an instant, or macroscopic material things (for example, "A red cube is on the table"). Later, however, in *Logische Syntax der Sprache* (1934) and other publications, due mainly to the criticisms of Neurath, Carnap held that the question of what protocol sentences describe is not a factual but a linguistic question and that we are free to choose whatever form of language is most convenient for reporting observations in science.

NEURATH

Otto Neurath, in "Soziologie im Physikalismus" (1931/1932; English translation, 1959) and other articles, had argued that sentences cannot be compared with the private experiences of the observer, nor with public mate-

rial things, but only with other sentences. Some sentences are reports of acts of observation, in the sense of being behavioral responses to those acts, and such protocol sentences may have whatever form we find most convenient. In "Protokollsätze" (1932/1933), Neurath maintained that for the purposes of science it must be possible to incorporate the protocol sentences expressed at one time in those expressed at another time, and that comparison of protocols, even with one's own past protocols, requires an intersubjective language. Neurath remarks, "*every* language *as such* is inter-subjective." Carnap later agreed that if protocol sentences were regarded as describing the observer's private experiences, they could be understood, if at all, only solipsistically. Neurath suggested that a convenient form for protocol sentences would be one which contained a name or description of an observer and some words recording an act of observation; he gives as an example "Otto's protocol at 3:17 o'clock [Otto's word-thought at 3:16: (In the room at 3:15 was a table perceived by Otto)]." In this example, it is supposed that the entire sentence is written down by Otto at 3:17, simply as an overt verbal response; the sentence in brackets is Otto's response at 3:16, and the sentence in parentheses is his response at 3:15. The word "Otto" is repeated, instead of using "my" and "me," in order that the components of the protocol may be independently tested, for example, by being found in the protocols of other observers. The protocols of different observers or of the same observer may conflict, and when this happens, one or more of them is to be rejected.

According to Neurath, Carnap, and also Carl Gustav Hempel in "On The Logical Positivists' Theory of Truth" (1934/1935) it is a matter of convenience and decision which of the conflicting protocols should be rejected; hence, no protocol is incorrigible. The aim of science is to build up a coherent system of sentences, but no sentence at any level is sacrosanct; every sentence in science is in the end accepted or rejected by a decision made in the interests of coherence and utility. This view was strenuously opposed—by Schlick, Bertrand Russell, and A. J. Ayer, among others, who argued that (1) on this account protocol sentences are distinguished from others only in respect of their syntactical form; (2) a purely syntactical criterion of truth cannot do the work required of it; and (3) the Neurath–Carnap doctrine is a complete abandonment of empiricism.

RUSSELL

According to Bertrand Russell's early doctrine of knowledge by acquaintance and knowledge by description, "every proposition we can understand must be composed wholly of constituents with which we are acquainted." A person is acquainted with those objects that are directly presented to his mind, and Russell held that sense data and universals are so presented. Later, in *The Analysis of Mind* (1921), Russell maintained that it is not possible to make a distinction between sensation and sense datum and that a sensation is not itself a cognition, although it is a cause of cognitions. This view led to the account of basic propositions that Russell gives in *An Inquiry into Meaning and Truth* (1940). In epistemology, he says, we can arrange our propositions about matters of fact in a certain order such that those that come later are known, if they are known, because of those that come earlier. At the beginning of such an ordering there will be "basic propositions"—those which "on reflection appear credible independently of any argument in their favour."

A basic proposition is one whose utterance is caused as immediately as possible by a perceptual experience. It is known independently of inference but not independently of evidence, since the perceptual experience that causes it to be expressed also gives the reason for believing it. The perceptual experience in question provides the strongest possible evidence for the basic proposition; no previous or subsequent occurrence and no experiences of others can prove that the proposition is false. Nevertheless, according to Russell, a basic proposition is not incorrigible; it cannot be disproved, but it may be false. Since one of the aims of epistemology is to show that all empirical knowledge is based upon these propositions, it is desirable that they should be given a logical form which makes contradiction between them impossible. Russell therefore defines a basic proposition as one "which arises on the occasion of a perception, which is the evidence for its truth, and … has a form such that no two propositions having this form can be mutually inconsistent if derived from different percepts" (*Inquiry into Meaning and Truth*, p. 139). Examples are "there is a canoid (shaped) patch of color," "I am hot," "that is red." Alternatively, "we can consider the whole body of empirical knowledge and define 'basic propositions' as those of its logically indemonstrable propositions which are themselves empirical" (ibid.). Russell believes that this logical definition is extensionally equivalent to his epistemological definition.

AYER

Whether basic propositions are incorrigible or indubitable, and if so in what sense, has been considered at length by A. J. Ayer. In "Basic Propositions" (1950) he defends the view that if a sentence is a direct description

of a private experience, it may be verbally incorrect, but it cannot express a proposition about which the speaker can be factually mistaken. He explains this in the following way. Many descriptive sentences, for example, "That is a table," may be used correctly (that is, in accordance with the rules of the language and on occasions generally agreed to be appropriate for their use), and yet the propositions they express may turn out to be false. But in the case of a sentence which directly describes a present experience, if the sentence is used correctly (that is, in accordance with the speaker's rules), the proposition it expresses cannot turn out to be false. Thus, "the sense in which statements like 'This is green,' 'I feel a headache,' 'I seem to remember——' can be said to be indubitable is that, when they are understood to refer only to some immediate experience, their truth or falsehood is conclusively determined by a meaning rule of the language in which they are expressed" ("Basic Propositions," p. 72).

Later, in *The Problem of Knowledge* (1956) and elsewhere, Ayer argues that language rules may be essentially private and that basic statements may be expressed in a sense-datum terminology, provided that this terminology is translatable into a terminology of seeming. Incorrigibility is not a property belonging to statements as such; "the sentences 'He has a headache,' when used by someone else to refer to me, 'I shall have a headache,' used by me in the past with reference to this moment, and 'I have a headache' all express the same statement; but the third of these sentences alone is used in such conditions as make it reasonable for me to claim that the statement is incorrigibly known" (*The Problem of Knowledge*, p. 58). But Ayer here allows that if he were asked, regarding two lines in his visual field, which looked to him to be the longer, he might very well be uncertain how to answer; and this uncertainty would not be about the meaning of the expression "looks longer than" but about a matter of fact. If anyone can have doubt about such matters of fact, he can presumably come to the wrong decision, that is, he can judge that one of the lines looks to him longer than the other when in fact it does not. No direct test of such a mistake is possible, but there may be various kinds of indirect evidence to show that it has occurred; hence, Ayer concludes, there is no class of descriptive statements which are incorrigible.

POPPER

The requirements made upon basic statements are very often governed by the general nature of the theory of knowledge held by a philosopher. Thus, according to Karl Popper, our experiences cannot justify or establish the truth of any statement; the question for epistemology is not "on what does our *knowledge* rest? … or more exactly, how can I, having had the *experience* S, justify my description of it and defend it against doubt," but rather "how do we test scientific statements by their deductive consequences … *what kind* of consequences can we select for this purpose if they in their turn are to be intersubjectively testable?" (*The Logic of Scientific Discovery*, p. 98). Popper requires a class of basic statements by reference to which it can be decided whether a theory or hypothesis in science is falsifiable. Evidently a theory can be falsified by a basic statement only if the negation of the latter is derivable from the theory. Popper finds that his requirements are met by taking singular existential statements of the form "There is a so-and-so in space-time region *k*" as basic. It follows that the negation of a basic statement is not itself a basic statement (Popper allows some simple exceptions to this in *Conjectures and Refutations*, Addenda, p. 386); it also follows that any conjunction of basic statements which is not a logical contradiction is a basic statement and that the conjunction of a nonbasic and a basic statement may be a basic statement (for example, the conjunction of "There is no pointer in motion at *k*" with "There is a pointer at *k*," which is equivalent to "There is a pointer at rest at *k*". Given a theory t conjoined with a statement of initial conditions r, from which a prediction p can be derived, it follows that $r\cdot\sim p$ will be a falsifier of t and a basic statement—since if $(t\cdot r)\rightarrow p$, then $t\rightarrow(r\rightarrow p)$, that is, $t\rightarrow\sim(r\cdot\sim p)$.

Popper also stipulates that the event referred to in a basic statement should be observable, that is, a basic statement must be intersubjectively testable by observation. He claims that the concept of an observable event can be elucidated either in terms of the experiences of an observer or in terms of macroscopic physical bodies, and hence that his account is neutral regarding the issue between psychologism and physicalism. In Popper's theory, the expression "observable event" is introduced "as an undefined term which becomes sufficiently precise in use: as a primitive concept whose use the epistemologist has to learn." According to Popper, "a science needs points of view and theoretical problems"; hence, in the practice of science we should not accept stray basic statements but only those which occur in the course of testing theories. Every test of a theory must terminate with some basic statement, but every basic statement can itself be subjected to further tests. There are no logical grounds for stopping at any particular basic statement. It is a matter for agreement and decision among those engaged in testing a theory; the process of corroboration or falsification

terminates at the point at which they are satisfied for the time being.

From the preceding selection of views, held by recent and contemporary philosophers, it will be seen that there is no consensus concerning basic statements. The questions listed at the beginning of this article can be answered only in relation to a more general semantic and epistemological theory. Many such theories allow that there is a distinct class of basic statements. It seems that the relation between these statements and certain "experiences" of the speakers who express them must be partly semantic, and perhaps also partly causal, but the correct analysis of this relation is a matter of great difficulty. Many philosophers at the present time deny that there can be a class of statements that describe the *private* experiences of the speaker, on the grounds that there cannot be a language that is essentially private; but this latter view is also strongly contested. Finally, although on some views basic statements are indubitable, it seems that these statements cannot be incorrigible, at least in any sense that implies that they cannot be false. For if basic statements are to play the role assigned to them—namely, of being the terminating points of empirical verification—they must be genuine contingent statements; and a contingent statement is one whose negation is significant and could, as far as logic is concerned, be true.

See also Ayer, Alfred Jules; Carnap, Rudolf; Empiricism; Hempel, Carl Gustav; Knowledge and Belief; Neurath, Otto; Popper, Karl Raimund; Propositions; Russell, Bertrand Arthur William; Schlick, Moritz; Verifiability Principle; Wittgenstein, Ludwig Josef Johann.

Bibliography

Wittgenstein's view in the *Tractatus Logico-Philosophicus* (London: K. Paul, Trench, Trubner, 1922) is the object of his own criticism in *Philosophical Investigations* (Oxford: Blackwell, 1953), especially in Part I, Secs. 1–64.

Schlick's "Über das Fundament der Erkenntnis," in *Erkenntnis* 4 (1934), has been translated by David Rynin in *Logical Positivism*, edited by A. J. Ayer (Glencoe, IL: Free Press, 1959), pp. 209–227. The same volume also contains English translations by Morton Magnus and Ralph Raico of Otto Neurath's "Soziologie im Physikalismus" (which originally appeared in *Erkenntnis* 2 [1931/1932]) on pp. 282–317, and by Frederic Schlick of Neurath's "Protokollsätze" (which originally appeared in *Erkenntnis* 3 [1932/1933]) on pp. 199–208.

For Carnap's views, see "Die physikalische Sprache als Universalsprache der Wissenschaft," in *Erkenntnis* 2 [1931/1932]), translated by Max Black as *The Unity of Science* (London: K. Paul, Trench, Trubner, 1934), and *Logische Syntax der Sprache* (Vienna: Springer, 1934), translated by Amethe Smeaton as *The Logical Syntax of Language* (London: K. Paul, Trench, Trubner, 1937).

The views of Hempel may be found in "On the Logical Positivists' Theory of Truth," in *Analysis* 2 (4) (1934/1935).

Russell's views on basic statements can be found in *The Analysis of Mind* (London: Allen and Unwin, 1921) and *An Inquiry into Meaning and Truth* (London, 1940).

Ayer's contributions to this topic include the following: *Language, Truth and Logic* (London: Gollancz, 1936; 2nd ed., 1946), Ch. 5 and Sec. 1 of introduction to 2nd ed.; "Verification and Experience," in *PAS* 37; *Foundations of Empirical Knowledge* (London: Macmillan, 1940), Ch. 2; "Basic Propositions," in *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950), reprinted in *Philosophical Essays* (London: Macmillan, 1954); and *The Problem of Knowledge* (London: Macmillan, 1956).

Relevant works by Karl Popper are *The Logic of Scientific Discovery* (New York: Basic, 1959), especially Ch. 5, and *Conjectures and Refutations* (New York: Basic, 1962).

Further discussion of Quine's views may be found in "Two Dogmas of Empiricism," in *Philosophical Review* (1951), reprinted in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953); *Methods of Logic* (London: Routledge and Paul, 1952), introduction; and *Word and Object* (New York: Wiley, 1960), Secs. 8–10.

OTHER RECOMMENDED TITLES

Alston, William. *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press, 1993.

Audi, Robert. *The Structure of Justification*. Cambridge, U.K.: Cambridge University Press, 1993.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Chisholm, Roderick. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982.

Chisholm, Roderick. *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice-Hall, 1966; 2nd ed., 1977; 3rd ed., 1989.

DePaul, Michael, ed. *Resurrecting Old-Fashioned Foundationalism*. Lanham, MD: Rowman & Littlefield, 2001.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Fumerton, Richard. *Metaphysical and Epistemological Problems of Perception*. Lincoln: University of Nebraska Press, 1985.

Haack, Susan. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Oxford: Blackwell, 1993.

Moser, Paul K. *Knowledge and Evidence*. New York: Cambridge Unversity Press, 1989.

Putnam, Hilary. "'Two Dogmas' Revisited." In *Realism and Reason, Philosophical Papers.* Vol. 3. Cambridge, U.K.: Cambridge University Press, 1983.

Sellars, Wilfrid. "Givenness and Explanatory Coherence." *Journal of Philosophy* 70 (1973): 612–624.

Sosa, Ernest. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Van Cleve, James. "Foundationalism, Epistemic Principles, and the Cartesian Circle." *Philosophical Review* 88 (1979): 55–91.

Wittgenstein, Ludwig. *On Certainty,* edited by G. E. M. Anscombe and G. H. von Wright. Oxford: Blackwell, 1969.

R. W. Ashby (1967)
Bibliography updated by Benjamin Fiedor (2005)

BATAILLE, GEORGES
(1897–1962)

Georges Bataille is a pivotal thinker in the history of twentieth-century thought, in a literal sense. His work serves as a pivot between any number of significant early twentieth-century trends, and later movements such as postmodernism and deconstruction.

The extremely eclectic Bataille was first, and perhaps most deeply, influenced by the Marquis de Sade. This scandalous thinker had an enormous impact on avant-garde French thought of the post-World War I period, most notably among the surrealists and their followers. Bataille, loosely associated with and against the surrealists, appropriated from Sade the notion of a violent, merciless natural order, and of man as a mimic of the destructive (and hence reconstructive) power of nature through the boundless expression of destructive sexual impulses. Bataille, like Sade, while a proclaimed atheist, nevertheless linked man's necessary violence to the blaspheming of God; in this way God, though denied, is in a strange way revived through the necessity of his transgression. (See early texts by Bataille, such as "The Solar Anus" [1927], and "The Use Value of D. A. F. de Sade [An Open Letter to My Current Comrades]" in *Visions of Excess* [1985]).

Bataille went on in the early 1930s to link Sade with the contemporary French anthropological theories of Marcel Mauss and Emile Durkheim. Both of these early twentieth-century thinkers hoped to find in primitive thought the kind of energy (social effervescence) whose absence led to the anomie, the rootlessness and pointlessness, of modern life. For Durkheim this energy was to be found in *mana*, the enthusiasm of crowds coming together; for Mauss, it was found in the rituals of gift-giving and ritualized destructions (such as the potlatch festivals of Northwest American Indians) of traditional societies. Both thinkers held that the basis for this social ritual was fundamentally rational: the energy of crowds and collective festivals was ultimately based on the peaceful tendency of people to recognize themselves as human. Mauss held that gift-giving, implemented as a major feature of modern economies, could counter the alienating tendencies of self-centered bourgeois economies. Bataille took this model and radicalized it to the extent that he held that gift-giving, crowds, and ritual destruction were energizing to the extent that they were irrational: A person's fundamental tendency was to expend (*dépenser*), and this urge, while making possible the full social experience, nevertheless put in question the stability and comfort of human life, not to mention the sacred integrity of the human person (so beloved by Durkheim). Expenditure, in this sense, was the affirmation of life to the point of the risk of death, and the Sadean affirmation of a "general economy" based not on saving and reinvestment, but on the extravagant squandering of wealth. (See "The Notion of Expenditure" [1932] in *Visions of Excess*, and *The Accursed Share* [1949]).

Throughout the 1930s and 1940s Bataille was at the vanguard of the French reception of Friedrich Nietzsche and G. W. F. Hegel. From Nietzsche, Bataille took the assertion of the death of God as a radical embrace of death, an apocalyptic, erotic moment. Nietzsche for Bataille is a lighthearted leap into the moment when God affirms his own nonexistence: the point at which the sacred is an affirmation not of conservation and reuse (the eternity of divinity), but the night of sacrificial oblivion. Out of this "left-hand" sacred, Bataille evolves a practice of mystical meditation based not on a communion with an ever-present God, but on the ecstatic horror of his definitive absence. For Bataille, this mystical practice is inseparable from the impossible task of writing it: this results in such fragmentary works as *Inner Experience* (1943), *Guilty* (1944) and *On Nietzsche* (1945).

Bataille's reading of Hegel is similarly unusual and arguably mistaken: following and rewriting Alexandre Kojève's Hegel, Bataille's version posits the end of history as a moment in which absolute knowledge turns and tries to incorporate the radical negativity on which it depends (through exclusion) in order to be complete. Rational, recoverable negativity can only be determined as recoverable in opposition to a more fundamental negativity that refuses all use, all constructive effort, all knowledge. Yet to be truly posthistorical, this negativity must finally be (impossibly) appropriated. To be Hegel all the way, one must recognize a negativity that by definition is unrecognizable: what Bataille called "not-knowing." Without this gesture one has not fully attained the "end of History"; with it, one is condemned to a circular agitation in which one's knowledge is incessantly lost in the oblivion of not-knowing. Negativity now, at the end, is a toxic form of *dépense*; at the same time, Hegel is nevertheless maintained to the extent that his philosophy is followed through, mimed, and not so much negated as always-

again affirmed in its loss, its madness (Bataille believed Hegel became mad at the moment he fully realized the consequences of the "end of History"). On this topic, see the "Hegel" section of *Inner Experience* (1943), and Bataille's short novel *Madame Edwarda* (1941).

Finally, in Bataille's writings on eroticism, he comes to see the expenditure of human limits in erotic contact; this "communication," as he called it, entails a community (of lovers) through the risk of the limits of the self. In this way Bataille revised the radical sexualized selfishness put forward by Sade: for Bataille "communication" is above all an act of generosity, if not a moral act. (See *Erotism*, 1957).

Bataille's eclectic rewriting of these major strands of French thought—moving in genres as diverse as sociological essays, mystical meditations, pornographic novels, and economic treatises—has had an enormous influence on French thought of the post-existentialist period. Two examples: Derrida's method of deconstruction, which involves not the refutation of a given work but rather the close following of that work and its steady disarticulation, all the while recognizing that the work of metaphysics cannot be escaped, but only endlessly repeated and deconstructed, owes much to the Bataillean reading of Hegel—indeed Derrida's reading of Bataille's Hegel may be seen as the model of the deconstructive project. Similarly, Foucault's affirmation of a "counter-discourse" in which a full, coherent discourse is destabilized by the discourse it must violently expel in order to constitute itself, clearly owes much to Bataille. Bataille, however, surpasses his recent avatars in his insistence on a political implementation of *dépense*; whereas Derrida, for example, is happy to rewrite Bataille's "general economy" as a "general writing"—thereby shifting the debate to an analysis of largely textual questions—Bataille insists on the need to rethink the future of society in ways that foresee a future economy based not on the profit motive but on the implementation of a global and orgiastic "spending without return."

See also Deconstruction; Derrida, Jacques; Durkheim, Émile; Foucault, Michel; Hegel, Georg Wilhelm Friedrich; Nietzsche, Friedrich; Postmodernism; Violence.

Bibliography

WORKS BY BATAILLE

Erotism. Translated by Mary Dalwood. New York: Marion Boyars, 1987.

Guilty. Translated by Bruce Boone. Introduction by Denis Hollier. Venice, CA: Lapis Press, 1988.

Inner Experience. Translated and introduced by Leslie Ann Boldt. Albany, NY: SUNY Press, 1988.

Madame Edwarda. Paris: Pauvert, 1966.

On Nietzsche. Translated by Bruce Boone. Introduction by Sylvère Lotringer. New York: Paragon House, 1994.

The Accursed Share. Translated by Robert Hurley. New York: Zone Books, 1988.

Visions of Excess: Selected Writings, 1927–1939, edited by Allan Stoekl. Translated by Allan Stoekl, with Carl Lovitt and Donald M. Leslie, Jr. Minneapolis: University of Minnesota Press, 1985.

WORKS ABOUT BATAILLE

Derrida, Jacques. "From Restricted to General Economy: A Hegelianism without Reserve." In *Writing and Difference*. Translated by Alan Bass, 251–277. Chicago: University of Chicago Press, 1978.

Foucault, Michel. "A Preface to Transgression." In *Language, Countermemory, Practice*, edited by Donald F. Bouchard. Translated by Sherry Simon, 29–52. Ithaca, NY: Cornell University Press, 1977.

Foucault, Michel. *The Order of Discourse*. Translated by Ian McLeod. In *Untying the Text: A Post-Structuralist Reader*, edited by Robert Young, 48–78. Boston: Routledge, 1981.

Gemerchak, Christopher M. *The Sunday of the Negative: Reading Bataille, Reading Hegel*. Albany, NY: SUNY Press, 2003.

Hegarty, Paul. *Georges Bataille: Core Cultural Theorist*. London: SAGE, 2000.

Allan Stoekl (2005)

BATTEUX, ABBÉ CHARLES
(1713–1780)

In the history of aesthetic ideas, the abbot Charles Batteux was less of an innovator than an apt synthesizer of prevailing ideas and a late defender of the classical theory of imitation in the new field of taste and aesthetic experience. Nonetheless, *Les beaux-arts réduits à un même principe* (The fine arts reduced to a single principle; 1746/1969) is generally thought to have provided the first modern classification of the fine arts. In all of his undertakings, Batteux sought to submit the fine arts—as opposed to the practical arts, which seek to fulfill various needs—to a single principle, "both simple and wide-reaching" (Foreword, *Les beaux-arts réduits à un même principe*), that could explain all varieties of art. In keeping with the classical theory of poetry and art, this principle is that art should imitate *la belle nature* (beautiful nature, including human actions and passions) to produce aesthetic pleasure. In other words, the *goal* of the fine arts is *pleasure*, their *essential characteristic* is *imitation*, and

their *subject* is *la belle nature*. The *manner* in which this imitation is done makes for the *particular differences* of the various art forms: poetry, painting, sculpture, dance, and music. On this basis, Batteux divided the inquiry conducted in *Les beaux-arts* into three parts. First, he identified the nature of all art forms and their essential differences. Second, he examined the nature of taste as a way of evaluating *la belle nature*. Third, to verify his theory by practice, he proposed a detailed typology of the fine arts.

Batteux first tried to clarify what it means for the fine arts to imitate *la belle nature*. Three aspects of this process deserve to be highlighted: imitation as such, the process of idealization that presides over the production of *la belle nature* in art, and the function of genius in producing works of art. First, art, as the product of genius's activity, works by imitating. Yet all imitation finds its raison d'être and its limits in the model that goes before it. Poetic and artistic invention is therefore not creating per se but rather reproducing what already exists. The function of art is to re-present its subject in a medium. Imitations must nevertheless *appear* to be nature. Perfection in the arts being based on resemblance, falling back on the purely formal (or purely aesthetic) properties of the aesthetic medium seems inadmissible for Batteux.

Second, in the Aristotelian tradition to which Batteux was explicitly connected, what the fine arts imitate is not nature as it truly is, but *la belle nature*, or nature as it *should be* as a result of idealization. In contrast with history, which simply presents the facts and strives to speak the truth, the fine arts present the ideal and strive for verisimilitude. They aspire, through selective representation of the real, to the perfection of the type. Painting and poetry are born with history, but the invention that is their own aims at drawing human actions together in a new and more coherent totality that brings out their meaning.

Third, only an artist of genius in a state of enthusiasm can produce true imitation of *la belle nature*. Far from being an occult faculty, enthusiasm, for Batteux, complements the spirit of observation. It designates the moment when the artist's spirit warms up at the sight of a vivid representation stemming from his imagination.

Although his theory of the imitation of *la belle nature* anchors Batteux's thought in the classical tradition, his theory of taste tends to bring together newer aesthetic tendencies that were forming during his era. Artistic genius is subject not to predetermined rules but to taste, which he defined as the "faculty of appreciating the good, the bad, and the mediocre, and of distinguishing among them" (Batteux [1746] 1969). Far from opposing the intelligence at work in the sciences, taste (which in its largest sense is essentially moral) always presupposes knowledge, to which feeling is added to motivate action or give rise to desire. In strictly artistic taste, sentiment, preceded by a knowledge of the qualities of an object, "tells us if the *la belle Nature* is well or poorly imitated" (Batteux [1746] 1969).

We can see to what extent ethics and aesthetics are intertwined: On the one hand, *la belle nature* that art imitates conforms to principles of taste to move individuals (in other words, it is directly connected with our general moral interests as human beings). On the other hand, it conforms to our cognitive nature, providing our minds with an exercise and movement that widens our sphere of ideas. Batteux considers the spectacle of human actions and human passions to be the primary subject of *la belle nature* represented, or rather engendered, by art. The ideal of artistic imitation associates the good (which corresponds with our universal moral interests), the beautiful (which satisfies our cognitive expectations of variety, uniformity, and novelty in the artistic representation), and the perfection of formal aspects of the work itself.

See also Aesthetics, History of; Aristotelianism; Art, Representation in; Pleasure.

Bibliography

WORKS BY BATTEUX

Les beaux-arts réduits à un même principe (1746). Geneva: Slatkine Reprints, 1969.

Principes de la littérature (5th ed., 1774). Geneva: Slatkine Reprints, 1967.

WORK ON BATTEUX

Saint-Girons, Baldine. *Esthétiques du 18e siècle: Le modèle français*. Paris: Philippe Sers, 1990.

Daniel Dumouchel (2005)

BAUDRILLARD, JEAN
(1929–)

Jean Baudrillard was born in the cathedral town of Reims, France. His grandparents were peasants, his parents became civil servants, and he was the first member of his family to pursue an advanced education. In 1956, he began working as a professor of secondary education in a French high school (Lyceé) and in the early 1960s did editorial work for the French publisher Seuil. Trained as a Germanist, Baudrillard translated German literary works—including Bertolt Brecht and Peter Weiss—

although he turned to the study of sociology and for some decades was a sociology professor at Nanterre.

Baudrillard became renowned for his theorizations of developments in contemporary society, including the trajectories of the consumer society, media and technology, cyberspace and the information society, and biotechnology. He claimed that cumulatively these forces had produced a postmodern rupture with modern culture and society. Whereas modern societies for Baudrillard were organized around production and political economy, postmodern societies were organized around technology and generated new forms of culture, experience, and subjectivities.

Baudrillard's work is extremely hard to categorize because he combines social theory, cultural and political commentary, philosophy, and literary stylistics in his work, crossing boundaries between academic disciplines and fields. Yet in an interview in *Forgetting Foucault* (1987, p. 84) he confessed: "Well, let's be frank here. If I ever dabbled in anything in my theoretical infancy, it was philosophy more than sociology. I don't think at all in those terms. My point of view is completely metaphysical. If anything, I'm a metaphysician, perhaps a moralist, but certainly not a sociologist. The only 'sociological' work I can claim is my effort to put an end to the social, to the concept of the social."

Indeed, beginning in the 1980s, more philosophical themes emerged in his work, although in a highly ironical and paradoxical form. Baudrillard's proliferating metaphysical speculations are evident in *Fatal Strategies* (1990), which can be seen as a turning to a sort of idiosyncratic philosophical musings. This text presented a bizarre metaphysical scenario concerning the triumph of objects over subjects within the obscene proliferation of an object world so completely out of control that it surpasses all attempts to understand, conceptualize, and control it. His scenario concerns the proliferation and growing supremacy of objects over subjects and the eventual triumph of the object.

For Baudrillard, the subject—the darling of modern philosophy—is defeated in his metaphysical scenario and the object triumphs, a stunning end to the dialectic of subject and object that had been the framework of modern philosophy. In *Fatal Strategies* and succeeding writings, Baudrillard seems to be taking theory into the realm of metaphysics, but it is a specific type of metaphysics deeply inspired by the pataphysics developed by Alfred Jarry in "What is Pataphysics" as "the science of the realm beyond metaphysics. … It will study the laws which govern exceptions and will explain the universe supplementary to this one; or, less ambitiously, it will describe a universe which one can see—must see perhaps—instead of the traditional one. …" (1963, p. 131ff.)

Like the universe in Jarry's play *Ubu Roi*, *The Gestures and Opinions of Doctor Faustroll*, and other literary texts, Baudrillard's is a totally absurd universe where objects rule in mysterious ways, and people and events are governed by absurd and ultimately unknowable interconnections and predestination. (The French playwright Eugene Ionesco is another good source of entry to this universe.) Like Jarry's pataphysics, Baudrillard's universe is ruled by surprise, reversal, hallucination, blasphemy, obscenity, and a desire to shock and outrage.

Thus, in view of the growing supremacy of the object, Baudrillard recommends abandoning the subject and siding with the object. Pataphysics aside, it seems that Baudrillard is trying to end the philosophy of subjectivity that has controlled French thought since Descartes by going over to the other side. Descartes's *malin genie*, his evil genius, was a ruse of the subject that tried to seduce him into accepting what was not clear and distinct, but over which he was ultimately able to prevail. Baudrillard's "evil genius" is the object itself that is much more malign than the merely epistemological deceptions of the subject faced by Descartes and which constitutes a "fatal destiny" that demands the end of the philosophy of subjectivity. Henceforth, for Baudrillard, people live in the era of the reign of the object.

Examples of the paradoxical and ironic style of Baudrillard's philosophical musings abound in *The Perfect Crime* (1996). Baudrillard claims that the negation of a higher and transcendent reality in the current media and technological society is a "perfect crime" that involves the destruction of the real. In a world of appearance, image, and illusion, Baudrillard suggests, reality disappears although its traces continue to nourish an illusion of the real. Driven toward virtualization in a high-tech society, all the imperfections of human life and the world are eliminated in virtual reality, but this is the elimination of reality itself, the Perfect Crime. This "post-critical" and "catastrophic" state of affairs render our previous conceptual world irrelevant, Baudrillard suggests, urging criticism to turn ironic and transform the demise of the real into an art form.

Baudrillard has entered a world of thought far from academic philosophy, one that puts in question traditional modes of thought and discourse. His search for new philosophical perspectives has won him a loyal global audience, but also criticism for his excessive irony, word play, and philosophical games. Yet his work stands

as a provocation to traditional and contemporary philosophy that challenges thinkers to address old philosophical problems such as truth and reality in new ways in the contemporary world.

See also Structuralism and Post-structuralism.

Bibliography

Jarry, Alfred. *The Ubu Plays*. New York: Grove Press, 1969.
Jarry, Alfred. "What Is Pataphysics?" *Evergreen Review* 4 (13)(1963): 131–151.

WORKS BY BAUDRILLARD

Simulations. New York: Semiotext(e), 1983.
Forgetting Foucault. New York: Semiotext(e), 1987.
Fatal Strategies. New York: Semiotext(e), 1990.
Symbolic Exchange and Death. London: Sage, 1993.
The Perfect Crime. London: Verso, 1996.

WORKS ON BAUDRILLARD

Genosko, Gary. *Baudrillard and Signs*. London: Routledge, 1994.
Kellner, Douglas, ed. *Jean Baudrillard: A Critical Reader*. Oxford: Basil Blackwell, 1994.
Kellner, Douglas. *Jean Baudrillard: From Marxism to Postmodernism and Beyond*. Cambridge and Palo Alto: Polity Press and Stanford University Press, 1989.

Douglas Kellner (2005)

BAUER, BRUNO
(1809–1882)

Bruno Bauer, the German theologian and historian, studied theology under P. H. Marheineke in Berlin, at the height of Georg Wilhelm Friedrich Hegel's influence there. When Bauer became a docent at the University of Berlin in 1834, he joined Marheineke on the Hegelian right wing. When he transferred to the University of Bonn in 1839, however, he was already reacting theologically against right-wing Hegelianism. D. F. Strauss's *Life of Jesus* (1835–1836) rocked the theological world, but it seemed to Bauer not sufficiently critical, and helped to spur him on to his own investigations of the Gospels.

Bauer began with literary criticism of the Gospel texts themselves, without making any assumptions about the historical life of Jesus or the early church. The fourth Gospel was simply a work of reflective Christian art dominated by Philo's logos concept, impressive as such, but without historical basis (*Kritik der evangelischen Geschichte des Johannes,* Bremen, 1840). The situation was the same with regard to the Synoptic Gospels, except that they were based on the conception of the Messiah (*Kritik der evangelischen Geschichte der Synoptiker,* 3 vols. Leipzig, 1841–1842.) Bauer adopted the conclusion of C. H. Weisse and C. Wilke that only Mark's Gospel was original, but argued further that there was no reason to assume any historical tradition behind this single literary source. Incongruities in Mark's text suggested that Mark had invented the events he related. Mark's story was accepted because it answered the spiritual needs of his age. Jesus was the man in whose consciousness the antitheses between heaven and earth, God and man, were reconciled. His character evoked the Messiah concept, into which his life was absorbed by Mark. Bauer's view seemed to undercut the historical basis of Christianity so sharply that the theological faculties of the Prussian universities were polled (with mixed results) as to whether Bauer should be dismissed from Bonn. Bauer sealed his fate with the article "Theological Shamelessness" (1814), in which he denounced the Christian faith as the source of lies and servile hypocrisy; he was dismissed in March 1842. Ultimately, Bauer denied the historicity of Jesus altogether, holding that Christianity was an amalgam of Stoic and Gnostic ideas in Jewish dress.

Meanwhile, Bauer had written his anonymous *Die Posaune des jüngsten Gerichts über Hegel den Atheisten und Antichristen* (Trumpet of the last judgment on Hegel the atheist and Antichrist, Leipzig, 1841), ostensibly from the standpoint of faith, attempting to show that the real result of Hegelian philosophy was neither the pantheism of Strauss nor the humanism of Ludwig Feuerbach—much less a defense of the Gospel—but Bauer's own out-and-out atheism.

At that time living on a small estate in Rixdorf, near Berlin, Bauer gathered around himself a circle of "free spirits" (including his brother Edgar) who frequented Berlin cafes. Bauer wrote brilliantly ironical "critiques" of recent historical developments in which he announced the downfall of Western philosophy and culture. For a time he collaborated with Arnold Ruge and with other left-wing Hegelians. But Bauer was as contemptuous of their revolutionary programs as he was of the bourgeois establishment. He attacked the inconsistencies and misconceptions of both groups; special class interests, he argued, are blindly one-sided, and the masses are so much dead matter, and inimical to the spirit. Only criticism, without presupposition, reservation, or special pleading, can be pure, can replace blindness with true conceptions, and can bring about the fundamental change in human consciousness that would really be liberating. History will, by its own "logic," bring about the transformation which no deliberate program can institute: what criticism

has destroyed in thought today, history will destroy in fact tomorrow. Bauer justified these views by means of a metaphysic of consciousness, according to which the world is the projection of the ego. Matter is the as yet unclarified aspect of the world; evil social conditions are the product of uncritical and self-alienated principles. Christianity, for example, freed the ego from its thralldom to the material world, but only through an alienation of spirit from matter that had in its turn created a new burden. But Bauer held that once Christianity's historical roots are exposed, its self-alienating power is broken; hence the importance of criticism. The same must be done with other forms of human bondage: revolutionary programs which do not reach to the roots of consciousness are futile.

Accordingly, Bauer attacked various reform movements as insufficiently radical. Jewish agitation for political rights, for example, was based on the separate religious identity of the Jew, and could never be defended on those grounds against those whose religious prejudices took a different form; the Jew could become free only by ceasing to be religious. Karl Marx answered this argument in his essay "On the Jewish Problem" (1844), and attacked Bauer as "St. Bruno" in *The Holy Family: Critique of the Critical Critic, against Bruno Bauer and Consorts* (1845). The real problem, according to Marx, was economic class behavior, and not the religious projections of that behavior. Bauer's view that social conditions could be changed by changing men's minds was a vestige of idealist-theological error, and the practical result of Bauer's theoretical radicalism would be political reactionism.

Bauer did in fact become a defender of Prussian conservatism, on the radical grounds that limited reform movements seemed to him to do more harm than good. But after 1850 his influence waned; though he continued to write prodigiously, his views were generally too eccentric to be relevant.

See also Conservatism; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Hegelianism; Marx, Karl; Philo Judaeus; Strauss, David Friedrich.

Bibliography

ADDITIONAL WORKS BY BAUER

Vollständige Geschichte der Parteikämpfe in Deutschland während der Jahre 1842–1846, 3 vols. Charlottenburg, Germany: E. Bauer, 1847.

Die bürgerliche Revolution in Deutschland seit dem Anfang der deutsche-katholischen Bewegung bis zur Gegenwart. Berlin: G. Hempel, 1849.

Russland und das Germanentum. Berlin: n.p., 1853.

Die Hegelsche Linke. Edited by Karl Löwith. Stuttgart and Bad Cannstatt: F. Frommann, 1962. Includes *Die Posaune* and selections from *Russland und das Germanentum.*

Christus und die Cäsaren, der Ursprung des Christentums aus dem römischen Griechentum. Berlin: E. Grosser, 1877.

WORKS ON BAUER

Hertz-Eichenrode, Dieter. *Der Junghegelianer Bruno Bauer im Vormärz.* Berlin: n.p., 1959.

Hook, Sidney. *From Hegel to Marx.* New York: Reynal and Hitchcock, 1936. Pp. 89–125.

Löwith, Karl. *Von Hegel zu Nietzsche,* 4th ed. Stuttgart, 1958. Pp. 120–125; 322–324; 366–374. For an extensive bibliography see pp. 432–433.

Schweitzer, Albert. *Geschichte des Leben-Jesu-Forschung.* Tübingen, 1926. Pp. 141–161. Translated from the first German edition, *Von Reimarus zu Wrede* (1906), by W. Montgomery under the title *The Quest of the Historical Jesus.* London, 1910. Pp. 137–160. Reprinted, New York: Macmillan, 1950.

Stephen D. Crites (1967)

BAUMGARTEN, ALEXANDER GOTTLIEB
(1714–1762)

Alexander Gottlieb Baumgarten, the German Wolffian philosopher and aesthetician, was born in Berlin. He was the son of an assistant to the Pietist theologian and pedagogue August Hermann Francke; his brother was the famous divine and church historian Sigmund Jakob. Baumgarten studied philosophy and theology at Halle. After receiving a master's degree in 1735, he was appointed a teacher at Halle and in 1738 became extraordinary professor. While teaching there, Baumgarten, in reaction against the Pietism dominant at Halle after the expulsion of Christian Wolff in 1723, reintroduced Wolffian philosophy. In 1740 he was appointed full professor at Frankfurt an der Oder, where he remained until his death.

Baumgarten's Latin handbooks on metaphysics, ethics, and practical philosophy were widely used in German universities both in his time and after his death, and his influence was extraordinary. Kant considered him to be one of the greatest metaphysicians of his time and adopted his *Metaphysics* and *Practical Philosophy* as textbooks for his own lectures at Königsberg. With the exception of his works on aesthetics, Baumgarten in general kept very close to Wolff's teachings, although he dissented

from Wolff on several special points. For instance, he adopted a middle position in the controversy over the problem of the interaction of substances by reconciling Wolff's theory of the "preestablished harmony" of the soul and body with the theory of physical influence supported by the Pietists. Baumgarten, as a supporter of Leibnizian panpsychism, applied his solution to the connections among all substances. Wolff, to the contrary, distinguished very sharply between spiritual and material substances. Baumgarten was thus less Leibnizian than Wolff in accepting physical influence and more Leibnizian in his panpsychism.

Baumgarten made his most important contributions in the field of aesthetics, expanding a subject that had been summarily treated by Wolff and going far beyond Wolff in developing it. In this field he collaborated so closely with his pupil G. F. Meier (1718–1777) that it is difficult to establish the real authorship of many doctrines. There is a very close connection between Baumgarten's *Meditationes Philosophicae de Nonnullis ad Poema Pertinentibus* and his unfinished *Aesthetica* and Meier's *Anfangsgründe aller schönen Künste und Wissenschaften* (3 vols., Halle, 1748–1750). Baumgarten introduced the term *aesthetics* to designate that section of empirical psychology which treats of the inferior faculty, that is, the faculty of sensible knowledge. The problem of beauty was only one part of this subject. Even in Kant, *aesthetics* referred both to sensible knowledge in general and to knowledge of beauty and the sublime in particular. Only later was it restricted to the field of beauty and sublimity. Aesthetics and logic together composed, in Baumgarten's view, a science that he called *gnoseology,* or theory of knowledge.

According to Baumgarten, the foundations of poetry and the fine arts are "sensitive (*sensitivae*) representations," which are not simply "sensual" (*sensuales*), but are connected with feeling (and therefore are pertinent both to the faculty of knowledge and to that of will). A beautiful poem is a "perfect sensitive discourse," that is, a discourse that awakens a lively feeling. This requires a high degree of "extensive clarity," which is different from "intensive (or intellectual) clarity." This means that an aesthetic representation must have many "characteristics," that is, it must be characterized by many different traits or particular elements, rather than by a few well-differentiated characters. Beauty must be "confused" and, therefore, excludes "distinctness," the main property of intellectual representations. Distinctness is reached by rendering clearly each of the characteristics of the characteristics of a representation. Establishing these characteristics presupposes intensive clarity and leads to a further abstraction of the concept of representations. This abstraction is obnoxious to aesthetic liveliness and leads to pedantry.

The artist is not an imitator of nature in the sense that he copies it: He must add feeling to reality, and thereby he imitates nature in the process of creating a world or a whole. This whole is unified by the artist through a coherent "theme," which is the focus of the representation.

This does not mean that the artist should prefer fiction to truth; on the contrary, knowledge of the beautiful is, at its best, sensible knowledge of truth made perfectly lively. This is a main point of divergence between Wolff and Baumgarten. Baumgarten held that, since rational knowledge of several orders of facts or of many facts in general is impossible, it must be replaced or supplemented by "beautiful knowledge," that is, reliable sensible knowledge of things that cannot be known rationally; such knowledge is as reliable as rational knowledge; typical aesthetic elements of the cognitive process are inductions and examples. By stressing the importance and relative independence of the inferior faculty (which Wolff held to be only an imperfect stage of knowledge, to be superseded by intellect and reason), Baumgarten foreshadowed Immanuel Kant's doctrine of the peculiar and independent function of sensibility in knowledge.

See also Aesthetics, History of; Aesthetics, Problems of; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Meier, Georg Friedrich; Panpsychism; Pietism; Wolff, Christian.

Bibliography

WORKS BY BAUMGARTEN

Meditationes Philosophicae de Nonnullis ad Poema Pertinentibus. Halle, 1735. Translated by K. Aschenbrunner and W. B. Hoelther, eds., as *Reflections on Poetry.* Berkeley, CA, 1954.

Metaphysica. Halle, 1739.

Ethica Philosophica. Halle, 1740.

Aesthetica, 2 vols. Frankfurt an der Oder, 1750–1758.

Initia Philosophiae Practicae Primae. Halle, 1760.

Acroasis Logica. Halle, 1761.

Ius Naturae. Halle, 1765.

Sciagraphia Encyclopaediae Philosophicae. Edited by J. C. Förster. Halle, 1769.

Philosophia Generalis. Edited by J. C. Förster. Halle, 1769.

WORKS ON BAUMGARTEN

Abbt, Thomas. *A. G. Baumgartens Leben und Charakter.* Halle: C.H. Hemmerde, 1765.

Bergmann, Ernst. *Die Begründung der deutschen Aesthetik durch A. G. Baumgarten und G. F. Meier.* Leipzig, 1911.

Cassirer, Ernst. *Die Philosophie der Aufklärung.* Tübingen: Mohr, 1932. Translated by F. C. A. Koelln and James D. Pettegrove as *The Philosophy of the Enlightenment.* Princeton, NJ, 1954. Pp. 338–357.

Meier, G. F. *A. G. Baumgartens Leben und Schriften.* Halle, 1763.

Poppe, B. *A. G. Baumgarten, seine Bedeutung und seine Stellung in der Leibniz-Wolffschen Philosophie und seine Beziehung zu Kant.* Münster, 1907.

Riemann, A. *Die Ästhetik A. G. Baumgartens.* Halle: M. Niemeyer, 1928.

ON BAUMGARTEN'S RELATION TO KANT

Bäumler, A. *Kants Kritik der Urteilskraft.* Halle, 1923.

Tonelli, Giorgio. "Kant, dall'estetica metafisica all'estetica psicoempirica." *Memorie della Accademia delle Scienze di Torino* series 3, Vol. 3, Pt. 2.

Giorgio Tonelli (1967)

BAYES, BAYES' THEOREM, BAYESIAN APPROACH TO PHILOSOPHY OF SCIENCE

The posthumous publication, in 1763, of Thomas Bayes's "Essay Towards Solving a Problem in the Doctrine of Chances" inaugurated a revolution in the understanding of the confirmation of scientific hypotheses—two hundred years later. Such a long period of neglect, followed by such a sweeping revival, ensured that it was the inhabitants of the latter half of the twentieth century above all who determined what it was to take a "Bayesian approach" to scientific reasoning.

Like most confirmation theorists, Bayesians alternate between a descriptive and a prescriptive tone in their teachings: They aim both to describe how scientific evidence is assessed and to prescribe how it ought to be assessed. This double message will be made explicit at some points, but passed over quietly elsewhere.

SUBJECTIVE PROBABILITY

The first of the three fundamental tenets of Bayesianism is that the scientist's epistemic attitude to any scientifically significant proposition is, or ought to be, exhausted by the subjective probability the scientist assigns to the proposition. A subjective probability is a number between zero and one that reflects in some sense the scientist's confidence that the proposition is true. (Subjective probabilities are sometimes called degrees of belief or credences.)

A scientist's subjective probability for a proposition is then more a psychological fact about the scientist than an observer-independent fact about the proposition. Roughly, it is not a matter of how likely the truth of the proposition actually is, but about how likely the scientist thinks it to be. Thus *subjective*—though in hindsight, *psychological* might have been a better term.

Unlike every other approach to confirmation theory, Bayesianism has no use for the notion of theory acceptance: There is no amount of evidence sufficient to induce a qualitative shift in a Bayesian's epistemic attitude from not accepting to accepting a theory. Learning from the evidence is always a matter of a quantitative adjustment, of changing your subjective probability for a hypothesis to reflect the latest evidence. At any time, the most favored theories are simply those with the highest subjective probabilities.

To found its first tenet Bayesianism must establish that it is plausible to suppose or reasonable to require that scientists have a subjective probability for every proposition that figures in their inquiry. Ramsey proposed that to have a subjective probability for a proposition is to have a certain complex disposition to act, a disposition that can be measured at least tolerably well in many cases by assessing betting behavior, as follows. The higher your subjective probability for a proposition, the lower the odds, all other things being equal, you will be prepared to accept in betting on the truth of that proposition. To be precise, given a subjective probability p for the proposition, you will accept odds of up to p: $(1 - p)$ on its truth—you will avoid just those bets, in other words, where you have to pay in more than p for every dollar you stand to win, so that for example if your subjective probability for the proposition is 0.3 then you will pay no more than $3 to play per game in which you win $10 just in case the proposition is true. Ramsey thought it likely that we have appropriately stable behavioral dispositions of this sort, accessible to measurement using the betting test, with respect to just about any proposition we understand, and so that we have subjective probabilities for all these propositions.

The Bayesian's principal tool is mathematical argument, and the mathematics in question is the probability calculus—the standard mathematics of probability—to which all subjective probabilities are assumed to conform. Conformance to the axioms is Bayesianism's second fundamental tenet.

Here the Bayesian argument tends to take a prescriptive turn. Having established that scientists have, as a matter of psychological fact, subjective probabilities for

all propositions that matter, the next step is to show that scientists ought to—whether they do or not—arrange their probabilities so as to satisfy the axioms of the probability calculus.

Typically this is done by way of a Dutch Book argument, an argument that shows that, if you do not adhere to the calculus, there is a certain set of bets on the truth of various propositions that you are committed in principle to accepting, but that will lead to a certain loss however things turn out. The details of the argument are beyond the scope of this entry, but an example may help. The first axiom of the probability calculus requires that the probability of a proposition and that of its negation sum to one. Suppose you violate this axiom by assigning a probability of 0.8 both to a certain proposition *h* and to its negation. Then you are committed in principle to accepting odds of 4 : 1 on both *h* and ¬*h*, which means a commitment to playing, at the same time, two games, in one of which you pay $8 and win $10 (i.*e*., your original $8 plus $2 "profit") if *h* is true, and in one of which you pay $8 and win $10 if *h* is false. Whether *h* is true or false you pay $16 but win only $10—a certain loss. To play such a game is irrational; thus you should conform your subjective probabilities to the probability calculus. Objections to the Dutch Book argument typically turn on the vagueness of the idea that you are "committed in principle" to accepting the bets in question; replies to these objections attempt to make the nature of the commitment more precise without leavening its evident undesirability.

BAYESIAN CONDITIONALIZATION

The third of Bayesianism's three fundamental tenets is Bayes' conditionalization rule, which instructs you on how to update your subjective probabilities as the evidence arrives. There are four steps to Bayes' rule. The first step is to define prior and posterior subjective probability. These notions are relative to your receipt of a piece of evidence: Your prior probability for a hypothesis is your subjective probability for the hypothesis immediately before the evidence comes in; your posterior probability for the hypothesis is your subjective probability immediately after the evidence (and nothing else) comes in. Bayes' rule gives you a formula for calculating your posterior probabilities for every hypothesis given your prior probabilities and the nature of the evidence. In so doing it offers itself as the complete story as to how to take evidence into account. In what follows, prior subjective probabilities are written as $C(\cdot)$, and posterior subjective probabilities as $C^{+}(\cdot)$.

The second step towards Bayes' rule is the introduction of the notion of conditional probability, a standard notion in probability mathematics. An example of a conditional probability is the probability of obtaining a four on a die roll, given that an even number is obtained. This probability is ⅓, since there are three equally probable ways for a die roll to be even, one of which is a four. Formally the probability of a proposition *h* conditional on another proposition *g* is written $C(h|g)$; it is usually defined to be $C(hg)/C(g)$. (Alternatively conditional probability may be taken as a primitive, as explained in the entry on Probability and Chance.)

The third step is to make the following simple posit about conditionalization: when you receive a piece of evidence *e*, you should update your probability for any given hypothesis *h* so that it is equal to your prior probability for *h* given *e*. That is, on learning that *e* is true, you should set your posterior probability $C^{+}(h)$ equal to your prior probability $C(h|e)$. This is Bayes' rule in its simplest form, but one further step will produce a more familiar, and revealing, version of the rule.

The fourth and final step is to notice a simple mathematical consequence of the definition of conditional probability, confusingly called Bayes' theorem (confusing because Bayes' theorem and Bayes' rule are two quite different propositions). According to Bayes' theorem,

$$C(h|e) = \frac{C(e|h)}{C(e)} C(h).$$

Combine Bayes' theorem and the simple form of Bayes' rule and you obtain the more familiar version of Bayes' rule:

$$C^{+}(h) = \frac{C(e|h)}{C(e)} C(h).$$

The effect of the application of Bayes' rule then—or as philosophers usually say, the effect of Bayesian conditionalization—is, on receipt of *e*, to multiply the old probability for *h* by the factor $C(e|h)/C(e)$. Call this factor the Bayesian multiplier.

What justification can be offered for Bayesian conditionalization? Since the notion of conditional probability is introduced by definition, and Bayes' theorem is a simple consequence of the definition, this amounts to the question why you ought, on learning *e*, to set your posterior probability for a hypothesis *h* equal to the prior probability $C(h|e)$.

Various arguments for conditionalizing in this way exist in the literature, often based on Dutch book considerations that invoke the notion of a conditional bet. The consensus is that none is entirely convincing. It is important to note that mathematics alone cannot settle the question: The probability calculus relates only different probabilities that are part of the same overall distribution, whereas Bayes' rule relates probabilities from two quite different distributions, the prior and posterior distributions.

Two further remarks on Bayesian conditionalization. First Bayes' rule assumes that the subjective probability of the evidence e goes to one when it is acquired, therefore that when evidence arrives, its content is exhausted by a proposition that comes to be known for sure. A natural extension of the rule, called Jeffrey conditionalization, relaxes this assumption. Second you may wonder whether background knowledge must be taken into account when conditionalizing. In fact it is automatically taken into account: Background knowledge has subjective probability one, and for any proposition k with probability one, $C(h|k) = C(h)$; thus, your subjective probability distribution always has your background knowledge in every respect "built in."

Now to discuss the implications of Bayesianism for confirmation. (Further implications will be considered below.)

The impact of evidence e on a hypothesis h is determined, recall, by the Bayesian multiplier, $C(e|h)/C(e)$, which when multiplied by the prior for h yields its posterior. You do not need any great mathematical expertise to see that, when $C(e|h)$ is greater than $C(e)$, the probability of h will increase on receipt of e, while when it is $C(e)$ that is greater, the probability of h will decrease.

When the receipt of e causes the probability of h to increase, e is said to confirm h. When it causes the probability of h to decrease, it is said to disconfirm h. This may look like a definition, but it is in fact a substantive philosophical thesis: The Bayesian claims that the preexisting notions of confirmation and disconfirmation can be given a satisfactory Bayesian analysis. (Or at least the Bayesian usually makes this claim: They also have the option of interpreting their definition as a piece of revisionism, not intended to capture our actual notion of confirmation but to replace it with something better.)

Two remarks. First to say that a hypothesis is confirmed is only to say that its probability has received some kind of upward bump. The bump may be small, and the resulting posterior probability, though higher than that prior, may be almost as small. The term *confirmed* has, in philosophical usage, a different sense from a term such as *verified*.

Second since whether or not a piece of evidence confirms a hypothesis depends on a subjective probability distribution, confirmation is in the first instance a relative matter. More on this in The Subjectivity of Bayesian Confirmation below.

One further definition: The quantity $C(e|h)$ is called a likelihood, specifically the likelihood of h on e (not to be confused with the probability of h given e, though there is a close relationship between the two, spelled out by Bayes' theorem).

The significance of the Bayesian multiplier can now be stated in natural language: A piece of evidence confirms a hypothesis relative to a particular subjective probability distribution just in case the likelihood of the hypothesis on the evidence is greater than the subjective probability for the evidence.

Consider a special case, that in which a hypothesis h entails the evidence e. By a theorem of the probability calculus the likelihood of h on e, that is, $C(e|h)$, is in any such case equal to one. Suppose that e is observed to be true. Assuming that $C(e)$ is less than one (which will be true unless all viable hypotheses predict e), then the likelihood will be greater than $C(e)$, and so h will be confirmed. Ignoring the parenthetical qualification, a hypothesis is always confirmed by its predictions. Further the more surprising the prediction, in a sense—the lower the prior probability of e—the more h will be confirmed if e is in fact observed.

The significance of this observation is limited in two ways. First some hypotheses predict evidence only with a certain probability less than one. Second hypotheses tend to make observable predictions only in conjunction with other, "auxiliary" hypotheses. The Bayesian response will be considered in the next section.

THE BAYESIAN MACHINE

Suppose you want to know whether a certain coin is fair, that is, biased neither towards "heads" nor "tails." You toss the coin ten times, obtaining exactly five "heads" and five "tails." How to conditionalize on this evidence? You will need three subjective probabilities: The prior probability for the hypothesis h that the coin is fair, the prior probability for the evidence e, and the likelihood of h on e. A good Bayesian is committed to adopting definite values for these subjective probabilities one way or another. If necessary, they will be set "by hand," that is, by some sort

of reflective process that is constrained only by the axioms of the probability calculus. But a great part of the appeal of Bayesianism is that the vast majority of subjective probabilities can be set "mechanically," that is, that they will have their values fully determined once a few special probabilities are set by hand. In the case of the coin, once the prior probability for h and its rivals is set by hand, a little philosophy and mathematics of probability will take care of everything else, mechanically fixing the likelihood and the probability for the evidence.

Begin with the likelihood, the probability of getting exactly five "heads" in ten tosses given that the coin is fair. Since the fairness of the coin entails (suppose) both a physical probability for "heads" of 0.5 and the independence of the tosses, the hypothesis that the coin is fair assigns a definite physical probability to your observed outcome of five "heads"—a probability of about 0.25, as it happens. Intuitively it seems right to take this as the likelihood— to set your subjective probability $C(e|h)$, that is, equal to the physical probability that h assigns to e. In its sophisticated form this intuition is what is sometimes known as Miller's Principle or the Principal Principle; call it the Probability Coordination Principle or PCP for short. Bayesians normally take PCP on board, thus relieving you of the effort of setting a value by hand for the likelihood in a case such as this.

Now consider the probability of the evidence. A theorem of the probability calculus, the total probability theorem, looks (in one of its forms) like this:

$$C(e) = C(e|h_1)C(h_1) + C(e|h_2)C(h_2) + \cdots.$$

where the hypotheses $h_1, h_2, \ldots$ form a mutually exclusive, exhaustive set, in the sense that one and only one of them must be true. In many cases the set of hypotheses among which you are trying, with the help of e, to decide form such a set (though see below). Thus if you have set values for the likelihoods $C(e|h_i)$ and prior probabilities $C(h_i)$ for all your rival hypotheses, the probability calculus gives you a unique correct subjective probability to assign to e.

To sum up: If your rival hypotheses assign definite physical probabilities to the evidence e and form a mutually exclusive, exhaustive set then by an independent principle of rationality, PCP, and a theorem of the probability calculus, total probability, the Bayesian multipliers for all of the hypotheses are completely determined once their prior probabilities are fixed.

As a consequence, you need only assign subjective probabilities by hand to a relatively small set of propositions, and only once in your life: At the beginning, before any evidence comes in, you will assign subjective probabilities to every possible scientific hypothesis. These assignments made, everything you need for Bayesian conditionalization is decided for you by PCP and the probability axioms. In this sense, Bayesian confirmation runs like a well-conditioned machine: You flip the on switch, by assigning initial prior probabilities to the different hypotheses that interest you, and then sit back and enjoy the evidential ride. (Conditionalization is also machine-like without PCP and total probability, but in that case flipping the on switch involves assigning values to $C(e|h_i)$ and $C(e)$ for every possible piece of evidence e.)

There are two obstacles to the smooth functioning of the Bayesian machine. First it may be that some or all of the rival hypotheses do not, on their own, assign a determinate physical probability to the evidence. In such cases the likelihood must either be fixed by hand, without the help of PCP or (more usually in the quantitative sciences) by supplementing the hypothesis with an auxiliary hypothesis in conjunction with which it does fix a physical probability for the evidence. In the latter case, PCP can be applied but complications arise when, as is typical, the truth of the auxiliary hypothesis is itself not known for sure. The conjunction of original and auxiliary hypothesis may be confirmed or disconfirmed mechanically, but the implication for the original hypothesis on its own—whether it is confirmed, and if so by how much—will continue to depend on handcrafted likelihoods such as $C(e|h)$. This is the Bayesian's version of confirmation theory's QuineDuhem problem. Strevens offers a partial solution to the problem. (The application of PCP will also fall through if the evidence is "inadmissible.")

Second, even when the likelihoods are fixed mechanically, the theorem of total probability may not apply if the rival hypotheses are either not mutually exclusive or not exhaustive. Lack of exhaustiveness is the more pressing worry, as it would seem to be the norm: Exhaustiveness implies that you have thought of every possible theory that predicts e to any extent—an unlikely feat. A simple fix is to include a residual hypothesis in your set to the effect that none of the other hypotheses is correct. Such a hypothesis will not however determine a definite physical probability for the evidence, so its likelihood and therefore the probability for the evidence will after all have to be fixed by hand.

BAYESIANISM AND THE PROBLEM OF INDUCTION

Does the Bayesian theory of confirmation solve the problem of induction? The case for an affirmative answer: Adherence to the tenets of Bayesianism can be justified a

priori (by Dutch book arguments and the like, or so some philosophers believe). And this adherence alone is sufficient to turn you into an inductive reasoner: Once you have settled on priors for all the hypotheses, the Bayesian machinery tells you what sort of things to expect in the future given your experience of the past.

Suppose for example that you wish to predict the color of the next raven. You have various theses about raven color: All ravens are blue; ravens are green with 50% probability, otherwise black; all ravens are black, and so on. In your life to date you have observed a number of ravens, all of them black. This evidence rules out altogether some of the raven color theses, such as the thesis that all ravens are blue. (The likelihood of the blue thesis on this evidence is zero, so the multiplier is zero: Observation of a black raven therefore causes your subjective probability for the blue thesis to drop to zero.)

Other theses have their probability shifted around by the evidence in other ways. The more they probabilify the evidence, the greater their likelihoods on the evidence and so the higher their Bayesian multipliers. Observing many black ravens has the effect then of moving your subjective probability away from hypotheses that do not probabilify blackness and towards theses that do. As a result, the observation of many black ravens in the past increases your subjective probability that the next raven will be black. Thus you have an a priori argument—the argument for accepting Bayesianism—that justifies inductive behavior.

The case for a negative answer as to whether Bayesianism solves the problem of induction can be made in two ways: By arguing that the a priori arguments for adopting the Bayesian apparatus fall through, or by arguing that Bayesianism does not, after all, underwrite inductive behavior. The second approach is the more illuminating.

Return to the ravens. The theses listed above have the uniformity of nature as a consequence: If any is true then the future will be, with respect to raven color, like the past. Once some non-uniform theses are thrown into the mix, everything changes. Consider for example the following thesis, reminiscent of Goodman's grue puzzle: All ravens observed until now are black, the rest green. The Bayesian multipliers for this thesis and the thesis that all ravens are black remain the same as long as all observed ravens are black, which is to say, up until this point in time. Just as probability has been flowing to the latter hypothesis, it will have been flowing to the former. It turns out then that the probability flow is not only towards theses that predict blackness for future ravens but also toward many others. Since the multipliers for these theses have been the same until now, your predictions about the color of ravens will favor blackness only if your initial prior probabilities—the probabilities you assigned to the different theses before any evidence came in—already favored the thesis that all ravens are black over the grue-like thesis, which is to say, only if you yourself already favored uniformity over diversity.

Many Bayesians have made their peace with Bayesianism's open-minded policy on natural uniformity. Howson argues for example that the Bayesian approach should not be considered so much a positive theory of confirmation—of how evidence bears on hypotheses—as a framework for implementing any theory of confirmation you like.

THE SUBJECTIVITY OF BAYESIAN CONFIRMATION

Suppose that the Bayesian machine is in good working order: You choose your prior probabilities for the rival hypotheses and then let the evidence, in conjunction with PCP and the total probability theorem, do the rest. Even then, with your personal input limited to no more than an assessment of the initial plausibility of the rival hypotheses, there is an unsettling element of subjectivity to the process of Bayesian confirmation, which is perhaps best brought out by the following observation: Two scientists who agree on the physical probabilities that a hypothesis h assigns to evidence e, and who follow PCP, so assigning the same value to the likelihood $C(e|h)$, may disagree on whether e confirms or disconfirms h.

To see why: e confirms h if the Bayesian multiplier is greater than one, and disconfirms it if the multiplier is less than one. The question then is whether $C(e|h)$ is greater than or less than $C(e)$. The scientists agree on $C(e|h)$, but they may have different values for $C(e)$: A scientist who assigns higher prior probabilities to hypotheses that assign higher physical probabilities to e will have a higher value for $C(e)$. It is quite possible for the two scientists priors for e to fall on either side of $C(e|h)$, in which case one will take e to confirm, the other to disconfirm, h.

A radical personalist denies that this is a problem: Why should two scientists agree on the significance of the evidence when one was expecting the evidence much more than the other? In the extreme, personalism of this sort approaches the view that Bayesian confirmation theory provides no guidance at all on assessing the significance of evidence, other than by establishing a standard of consistency; see also the discussion of induction above.

There is some objectivity underlying Bayesianism's subjectivity, however. The two scientists above will, because they agree on the likelihoods, agree on the ordering of the Bayesian multipliers. That is they will agree on which of any two hypotheses has the higher Bayesian multiplier, even though they may disagree on the size of the multipliers.

An important consequence of this agreement is a result about the convergence of opinion. When hypotheses assign physical probabilities to the evidence, as assumed here, it can be shown that as time goes on, the subjective probability distributions of any two scientists will with very high physical probability converge on the truth, or rather to the class of hypotheses empirically equivalent to the truth. (Even when the likelihoods are purely subjective, or at least only as objective as the probability calculus requires, a convergence result, albeit more limited, can be proved.)

Many Bayesians regard this convergence as ameliorating, in every important way, the subjective aspect of Bayesianism, since any disagreements among Bayesian scientists are ephemeral, while agreement lasts forever. Indeed, that Bayesianism makes some, but not too much, room for scientific dissent may not unreasonably be seen as an advantage, in both a descriptive and a prescriptive light.

Now consider a contrary view: While dissent has its place in science, it has no place in scientific inference. It is fine for scientists to disagree, at least for a time, on the plausibility of various hypotheses, but it is not at all fine that they disagree on the impact of the evidence on the hypotheses—agreement on the import of the evidence being the *sine qua non* of science. In Bayesian terms scientists may disagree on the priors for the rival hypotheses, but they had better not disagree on the Bayesian multipliers. But this is, for a Bayesian, impossible: The priors help to determine the multipliers. The usual conclusion is that there is no acceptable Bayesian theory of confirmation.

A less usual conclusion is that Bayesianism is still viable, but only if some further principle of rationality is used to constrain the prior probabilities in such a way as to determine uniquely correct values for the Bayesian multipliers. This is objectivist Bayesianism. Just as PCP is used to determine definite, objective values for the likelihoods, the objectivists suggest, so another rule might be used to determine definite, objective values for the prior probabilities of the hypotheses themselves, that is, for the subjective probabilities $C(h)$.

What principle of rationality could possibly tell you, before you have any empirical evidence whatsoever, exactly how plausible you ought to find some given scientific hypothesis? Objectivists look to the principle of indifference for the answer. That principle, discussed more fully in the entry on Probability and Chance, is in one of its guises intended to specify a unique probability distribution over a set of propositions, such as hypotheses, that reflects complete ignorance as to which of the set is true. Thus the fact that you have no evidence is itself taken to commit you to a particular assignment of prior probabilities—typically, a probability distribution that is uniform in some sense. Jaynes (1983) has done the most to develop this view.

The objectivist envisages all Bayesian reasoners marching in lock-step: They start with precisely the same priors; they apply (thanks to PCP and total probability) precisely the same Bayesian multipliers; thus they have the same subjective probabilities at all times for everything.

There are various powerful objections to the most general forms of the principle of indifference. Even its most enthusiastic supporters would shy away from claiming that it determines a uniquely correct prior for absolutely any scientific hypothesis. Thus the lock-step picture of Bayesian inference is offered more as an ideal than as a realistic prospect. To be a modern objectivist is to argue that parts of science, at least, ought to come close to realizing the ideal.

THE PROBLEM OF OLD EVIDENCE

Among the many achievements of Newton's theory of gravitation was its prediction of the tides and their relation to the lunar orbit. Presumably the success of this prediction confirmed Newton's theory, or in Bayesian terms, the observable facts about the tides e raised the probability of Newton's theory h.

But the Bayesian it turns out can make no such claim. Because the facts about the tides were already known when Newton's theory was formulated, the probability for e was equal to one. It follows immediately that both $C(e)$ and $C(e|h)$ are equal to one (the latter for any choice of h). But then the Bayesian multiplier is also one, so Newton's theory does not receive any probability boost from its prediction of the tides. As either a description of actual scientific practice, or a prescription for ideal scientific practice, this is surely wrong.

The problem generalizes to any case of "old evidence": If the evidence e is received before a hypothesis h

is formulated then *e* is incapable of boosting the probability of *h* by way of conditionalization. As is often remarked, the problem of old evidence might just as well be called the problem of new theories, since there would be no difficulty if there were no new theories, that is, if all theories were on the table before the evidence began to arrive. Whatever you call it, the problem is now considered by most Bayesians to be in urgent need of a solution. A number of approaches have been suggested, none of them entirely satisfactory.

A recap of the problem: If a new theory is discovered midway through an inquiry, a prior must be assigned to that theory. You would think that, having assigned a prior on non-empirical grounds, you would then proceed to conditionalize on all the evidence received up until that point. But because old evidence has probability one, such conditionalization will have no effect. The Bayesian machinery is silent on the significance of the old evidence for the new theory.

The first and most conservative solution to the problem is to take the old evidence into account in setting your prior for the new theory. In doing this you are entirely on your own: You cannot use conditionalization or any other aspect of the Bayesian apparatus to weigh the evidence in a principled way. But because you are free to choose whatever prior you like, you are free to do so in part on the basis of the old evidence.

A second solution requires a radical revision of Bayesian conditionalization, so as to allow conditionalization using not the actual probability of the old evidence, but using a (now) counterfactual probability such as your prior for the evidence immediately before you learned it. This provides a natural way to use conditionalization to weigh the old evidence, but the difficulties involved in choosing an appropriate counterfactual prior and in justifying conditionalization on the false prior, rather than the actual prior, have not unreasonably scared most Bayesians away.

The third and perhaps most popular solution suggests that, although conditionalization on old evidence *e* has no effect on the prior probability of a new theory *h*, conditionalizing on the fact that *h* predicts *e* (for simplicity's sake, assume that it entails *e*) may have an effect. The idea: Until you formulate *h*, you do not know that it entails *e*. Once *h* is formulated and assigned a prior, you may conditionalize on the fact of the entailment; learning that *h* entails *e* will have much the same impact on the probability of *h*, it is supposed, as learning *e* would have had if it were not already known.

There are two difficulties with this suggestion. The first is that facts about entailment (either of *e* itself, or of a physical probability for *e*) are logical truths, which ought according to the probability calculus to be assigned probability one at all times—making the logical facts as "old" as the evidence itself. Proponents of the present approach to old evidence argue not unreasonably that a sophisticated Bayesianism ought to allow for logical learning, so that it is the requirement that subjective probabilities conform to the probability calculus in every respect that is at fault here, for imposing an unreasonably strict demand on flesh-and-blood inquirers.

The second (and related) difficulty is that the theory of conditionalization on logical facts is not nearly so nicely developed as the theory of orthodox Bayesian conditionalization. A case can be made that conditionalizing on *h*'s entailment of old evidence will increase the probability of *h*, but the details are complicated and controversial.

BAYESIANISM ACCESSORIZED

Two notable additions to the Bayesian apparatus are ever under consideration. First is a theory of acceptance, that is, a theory that dictates, given your subjective probabilities, which hypotheses you ought to "accept." Conventional Bayesianism has no need of acceptance: Your subjective probabilities are taken to exhaust your epistemic attitudes to the hypotheses, and also to determine, along with your preferences in the usual decision-theoretical way, the practical significance of these attitudes.

Some philosophers argue that there is, nevertheless, work for a notion of acceptance to do, and hold either a simple view on which hypotheses with high subjective probability are to be accepted, or a more sophisticated view on which not only probability but the consequences for science, good and bad, of acceptance must be taken into account.

Second is a theory of confirmational relevance, that is, a theory that dictates, given your subjective probabilities, to what degree a given piece of evidence confirms a given hypothesis. Conventional Bayesianism has no need of confirmational relevance: Your subjective probabilities are taken to exhaust your epistemic attitudes to the hypotheses, and so the dynamics of confirmation are exhausted by the facts about the way in which the subjective probabilities change, which are themselves fully determined, through conditionalization, by the values of the subjective probabilities themselves. Nothing is added to the dynamics of probability change—nothing could be added—by finding a standard by which to judge whether

certain evidence has a "large" or "small" impact on the hypotheses; however you talk about probability change, it is the change that it is. (A pure-hearted Bayesian need not even define *confirms* and *disconfirms*.)

Many different measures of relevance have, nevertheless, been proposed. The simple difference measure equates the relevance of *e* to *h* with the difference between the prior and posterior probabilities of *h* after conditionalization on *e*, or equivalently, with $C(h|e) - C(h)$. The likelihood measure equates the relevance of *e* to *h* with $C(e|h)/C(e|\neg h)$. It should be noted that all popular Bayesian measures render relevance relative to background knowledge.

There is no doubt that scientists sometimes talk about accepting theories and about the strength of the evidence—and that they do not talk very much about subjective probability. The degree to which you see this as a problem for unadorned Bayesian confirmation theory itself measures, perhaps, your position on the spectrum between prescriptive and descriptive.

See also Confirmation Theory; Decision Theory; Goodman, Nelson; Induction; Newton, Isaac; Probability and Chance; Ramsey, Frank Plumpton; Rationality.

Bibliography

Earman, John. *Bayes or Bust?* Cambridge, MA: MIT Press, 1992.

Fitelson, Branden. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (1999): S362–S378.

Glymour, Clark. *Theory and Evidence*. Princeton, NJ: Princeton University Press, 1980.

Howson, Colin. *Hume's Problem: Induction and the Justification of Belief*. Oxford: Oxford University Press, 2001.

Howson, Colin, and Peter Urbach. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court, 1993.

Jaynes, Edwin T. *Papers on Probability, Statistics, and Statistical Physics*, edited by Roger Rosenkrantz. Dordrecht: Reidel, 1983.

Jeffrey, Richard. *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press, 1983.

Levi, Isaac. *Gambling with the Truth*. Cambridge, MA: MIT Press, 1967.

Ramsey, Frank. "Truth and Probability." Reprinted in *Philosophical Papers*. Edited by D. H. Mellor. Cambridge, U.K.: Cambridge University Press, 1931.

Strevens, Michael. (2001). "The Bayesian Treatment of Auxiliary Hypotheses." *British Journal for the Philosophy of Science* 52 (2001): 515–538.

Michael Strevens (2005)

BAYLE, PIERRE

(1647–1706)

Pierre Bayle, the most important and most influential skeptic of the late seventeenth century, was born in Carla (now Carla-Bayle), a French village near the Spanish frontier, where his father was the Protestant pastor. He grew up during the religious persecutions under Louis XIV that culminated in the revocation of the Edict of Nantes (1685) and the outlawing of Protestantism in France. Bayle was sent first to a Calvinist school and then to a Jesuit college at Toulouse, where after studying the controversial literature and hearing the dialectical arguments of some of the professors, he converted to Catholicism. The intellectual considerations that led him to Catholicism, after further examination, soon led him back to Calvinism. He became technically a *relaps*, a person who has returned to heresy after having abjured it, and under French law he was therefore subject to severe penalties.

He left France for Geneva, where he completed his philosophical and theological studies. In 1674 he returned to France incognito and became a tutor in Paris and Rouen. The next year he obtained the philosophy professorship at the Protestant academy of Sedan as the protégé of Pierre Jurieu, a superorthodox theologian who was to become Bayle's greatest enemy. Bayle taught at Sedan until the school was closed in 1681. He and Jurieu went to Holland; they became members of the *École illustre* of Rotterdam and of the French Reformed church there. Bayle brought with him his first work, a letter concerning the comet of 1680, which he published under a pseudonym. This volume, like many of those to follow, attacked superstition, intolerance, and poor philosophy and history. The work was immediately successful and was soon followed by others, including an answer to Father Maimbourg's history of Calvinism and a collection of defenses of Cartesianism.

During these early years in Rotterdam, Bayle apparently made some fundamental personal decisions that affected the rest of his life. The first was not to marry but to devote himself to the solitary life of the dedicated scholar seeking truth. The second was to refuse any important professorship to carry on his work in Rotterdam (where he lived almost continuously for the rest of his life). Lastly, after his father and his brothers died in France as a result of the religious persecutions, Bayle apparently committed himself to both the cause of Calvinism and the cause of toleration.

From 1684 until 1687 he edited the *Nouvelles de la république des lettres*, one of the first learned journals of modern times, in which he reviewed works in many fields. His critical appraisals soon made him a major figure in the learned world and brought him in contact with the leading lights of his day, among them Antoine Arnauld, Robert Boyle, Gottfried Wilhelm Leibniz, John Locke, and Nicolas Malebranche.

TOLERATION

In 1686 Bayle published in Amsterdam his *Commentaire philosophique sur ces paroles de Jésus-Christ "Constrains-les d'entrer"* (Philosophical commentary on the words of Jesus "constrain them to come in"), a brilliant argument for complete religious toleration. Starting with the problem raised by Louis XIV's persecutions, Bayle developed a defense of toleration for Jews, Moslems, Socinians (Unitarians), Catholics, and even atheists, extending its scope far beyond Locke's not yet published *Essay on Toleration*.

Enmity had begun to develop between Bayle and Jurieu, who conceived of himself as the chief spokesman for Calvinist orthodoxy, opposed all kinds of deviation as heresy and atheism, and advocated political victory over Louis XIV. As Jurieu became a violent political radical and religious bigot, Bayle drifted away from the views and company of his former mentor. According to Jurieu the disaffection reached the breaking point with the publication of Bayle's "Philosophical Commentary." Bayle had tried to hide his authorship, but Jurieu soon guessed the truth and realized that they disagreed completely about almost everything. He saw his colleague as a menace to true religion and a secret atheist. Bayle intensified the quarrel by ridiculing Jurieu, attacking his intolerance and his political plans. Throughout the quarrel, Bayle insisted that he was a true follower of John Calvin and that he had imbibed his orthodoxy from Jurieu's antirational theology.

When Bayle began to publish his views, the Protestant liberals thought that he was on their side. But Bayle quickly employed his dialectical and critical skill to decimate their contentions and to show that there was no way of making the rational and scientific world compatible with the basic claims of Christianity, as they in part believed it to be. As a result, various liberal Protestants spent years defending themselves against Bayle's sharp criticisms, while Bayle alternately joined them in attacking Jurieu and Jurieu in attacking them.

Between 1690 and 1692 the argument between Bayle and Jurieu reached fever pitch, especially concerning whether or not Bayle was the author of the notorious "Advice to the French Refugees," a work criticizing the romantic optimism and hopes of the Protestant exiles. (Bayle so confused the evidence that even present-day scholars are unwilling to state positively that he did write it.) These controversies with Jurieu led in 1693 to Bayle's dismissal from his teaching post, an event that allowed him time to carry on his many controversies and to complete his great *Dictionnaire historique et critique* (A general dictionary, historical and critical; first published in two volumes in Rotterdam in 1695 and 1697), a work in which Jurieu is constantly attacked.

HISTORY AND COMPOSITION OF THE *DICTIONARY*

Bayle had conceived the basic idea of the *Dictionary* long before its composition. For many years he had been assembling collections of errors uncovered in various historical works. As early as 1675, Bayle's letters show, he was actively interested in skeptical thought. In the lectures Bayle gave at Rotterdam he criticized every possible theory. The *Dictionary* brought his critical and skeptical sides together. Originally, Bayle planned only to write a dictionary that would list the mistakes in all other dictionaries and in particular the one by Louis Moréri. A sample portion of this project was printed in 1692 to test public interest. The negative reaction led to a change of plan; the dictionary became a historical and critical one, dealing principally with persons and mainly with those who were not treated fully or at all (usually because of their obscurity or insignificance) in Moréri's opus. The result was two folio volumes full of articles on little-known or totally unknown figures, omitting significant figures like Plato, Michel Eyquem de Montaigne, and Cardinal Richelieu.

The *Dictionary* was composed in Talmudic style. Relatively brief biographical articles appeared at the top of the page, while all sorts of digressive notes on factual, philosophical, religious, or other matters appeared below, with notes on notes appearing in the margins. The biography of some extremely little-known personage, like Rorarius, would provide the stage for profound discussions of the nature of man and beasts, the mind-body problem, and the new metaphysical theory of Leibniz. Other subjects would provide forums for discussing the problem of evil; the immorality of great figures, especially Old Testament ones; the irrationality of Christianity; the problems of Locke's, Isaac Newton's, Malebranche's, Aristotle's, or anyone else's philosophy; or some salacious tale about a famous theologian, Catholic or Protestant, or a famous political figure of almost any age. There was little

relation between the official subject of an article and its real content. But there were several major themes and threads that ran through many or most of the articles, themes that amounted to a massive onslaught against almost any religious, philosophical, moral, scientific, or historical view that anyone held. (Once Bayle explained that he was a Protestant in the true sense of the term and that he opposed everything that was said and everything that was done.)

The *Dictionary* was an instant success and immediately led to criticism and condemnation, both by the French Reformed church of Rotterdam and by the French Catholic church. The latter group banned the work, while the former demanded that the author revise or explain his views about the good moral character of atheists, the inability of Christians to answer the Manichaean views about the nature of evil, the strength of Pyrrhonian skepticism, the immoral character of King David, and why so many obscenities appeared in the work.

Bayle promised the congregation of the French Reformed church that he would revise the article "David" and would offer explanations of the other matters. Almost as soon as the first edition of the *Dictionary* appeared, he began work on the second, revising the article "David" and adding many additional articles, plus a set of clarifications. This final edition appeared in Rotterdam in 1702 and consisted of 7 to 8 million words. After this monumental effort, the rest of Bayle's career was devoted to carrying on various controversies, defending some of the claims in the *Dictionary*, and fighting a growing list of opponents. He died on December 28, 1706, while completing his *Entretiens de Maxime et de Thémiste* (Conversations between Maxime and Themiste; Rotterdam, 1707), a final reply to the liberal Protestants.

Replies to Bayle kept appearing, written by such figures as Leibniz, Bishop William King, and Jean-Pierre Crousaz; and the avant-garde spirits of the Enlightenment found much ammunition in Bayle's folio columns with which to attack the ideological and theological ancien régime. François-Marie Arouet de Voltaire, David Hume, Edward Gibbon, Denis Diderot, and many others found intellectual nutrition in Bayle's skeptical and critical efforts. Thomas Jefferson recommended the *Dictionary* as one of the hundred basic books with which to start the Congressional Library. Poets and writers of fiction like Alexander Pope, Henry Fielding, and Herman Melville found inspiration and plots in some of Bayle's spicy tales. Ludwig Feuerbach (1967), in the nineteenth century, saw Bayle as a major figure in the rise of modern thought and devoted a whole volume to him.

The *Dictionary* was enormously influential during the eighteenth century, both for its spirit and for its wealth of information. Though it was written in the form of a reference work, its lopsided contents, overloaded with lives of obscure theologians and figures of French political history, made it difficult for the *Dictionary* to maintain its character as a guide to research and scholarship. Efforts to improve it by adding and updating articles were only temporarily successful. The editors of the 1734–1741 English edition put in hundreds of articles on English and Arab figures, plus some "correctives" to what they regarded as outlandish in Bayle's original. In 1740 Jacques-Georges de Chaufepié translated many of the English articles into French, adding a great many more on Bayle's opponents, and put out a four-volume folio supplement. However, the type of critical and careful research Bayle had fostered gave birth to projects that would forever make his *Dictionary* obsolete as a reference work. *La Grande Encyclopédie* and the *Encyclopaedia Britannica*, which replaced it, were continuing team efforts, rather than one man's appraisal of the whole intellectual world. Thus, Bayle's work became a victim of its own offspring. It gradually disappeared as an important element in the intellectual world and was superseded by the works of leaders of the Enlightenment who had imbibed at least part of Bayle's spirit.

PHILOSOPHICAL ASPECTS OF THE *DICTIONARY*

The discussions in the *Dictionary* that had the greatest philosophical impact were those dealing with the problem of evil, with the independence of morality from religion, and with the unintelligible nature of the physical and mental world, especially when analyzed in terms of the categories of the "new science" and the "new philosophy." With a dialectical skill unknown to earlier skeptics Bayle dissected every theory and showed that it was unsatisfactory. Instead of merely utilizing the classical epistemological arguments of Sextus Empiricus, slightly modernized by the Montaignians, Bayle employed primarily the method of one of his heroes, the "subtle Arriaga" (Roderigo Arriaga, the last of the Spanish scholastics, who died in 1667), a method that Bayle had probably learned from the Jesuits at Toulouse.

The technique consisted in exposing the weakness of every rational attempt to make sense of some aspect of human experience. Bayle, like Arriaga before him, repeatedly exhibited man's sorry intellectual plight. All human rational efforts are always their own undoing and terminate in theories that are "big with contradiction and

absurdity." Bayle concentrated on a few shocking illustrations of this thesis. In a series of articles, "Manichaeans," "Marcionites," "Pauilicians," and "Rufinus," he contended that the Manichaean or dualistic theory of two gods, one good and one evil, could not be refuted by orthodox Christian theology, that it was a better explanation of human experience of evil, but that it was ultimately repugnant to sound reasoning. (Leibniz's *Theodicy* was largely an attempt to refute Bayle on Manichaeanism and the problem of evil.)

RELIGION AND MORALITY

Throughout his writings, from his letter on the comet to the *Dictionary* and its various defenses, Bayle argued the then scandalous thesis that a society of atheists could be moral and a society of Christians immoral. He tried to show that people's moral behavior is not a consequence of their beliefs but is rather the result of many irrational factors, such as education, custom, passion, ignorance, and the trace of God. In the article "Jupiter" he pointed out that Greek mythology was absurd and immoral, but the Greeks lived moral lives nonetheless. In his "Clarification on Atheism" he stated that he could find no case of a classical atheist, or a modern one like Benedict (Baruch) de Spinoza, who lived a wretched, morally degenerate life. Instead, the cases he found all seemed to be ones of highly moral people, who also happened to be atheists.

Additionally, Bayle knew of myriad cases—from a biblical one to leading Catholic and Protestant clergy of his day—of religious heroes who were immoral and whose behavior seemed to have been influenced by the most irreligious factors. Among many articles dealing with the sexual aberrations of different religious fanatics, early reformers, and Renaissance popes, the long one on "David" brought this point out most forcefully. David was introduced as the most holy figure in the Old Testament, and a series of notes outlined and analyzed his immoral conduct. This massive assault on any alleged rational or necessary connection between religious belief and moral behavior greatly influenced the Third Earl of Shaftesbury (Anthony Ashley Cooper; who lived and argued with Bayle for a while), and Bernard Mandeville (who was apparently one of Bayle's students at Rotterdam), and through them many of the eighteenth-century British moralists.

METAPHYSICS

In metaphysics Bayle employed his dialectical skill to show that theories about the nature of matter, space, time, motion, mind, and mind-body relationships, when thoroughly analyzed, are contradictory, inadequate, and absurd. Starting with Zeno of Elea's paradoxes and the sections in Sextus against metaphysics, Bayle attacked all sorts of ancient and modern forms of atomism, Platonism, and Aristotelianism, as well as the modern substitutes offered by René Descartes, Thomas Hobbes, Spinoza, Malebranche, Leibniz, Locke, Newton, and many others. He showed the weird, incredible conclusions that would follow from each of these theories. (Bayle's article "Rorarius" was the first public examination of, and attack on, Leibniz's theories of preestablished harmony and of monads.) In the articles "Pyrrho" and "Zeno of Elea" (which greatly influence George Berkeley and Hume) Bayle brilliantly challenged the distinction between primary and secondary qualities, so fundamental in the theories about reality of all the "new philosophers."

SKEPTICISM

Bayle repeatedly showed that the many attempts by human beings to explain or understand their world were all just "highroads to Pyrrhonism," since they only made every supposition more perplexing, absurd, and dubious. Rational activity, no matter what problem it is directed at, leads to complete skepticism, since reason invariably leads one astray. In the article "Acosta" Bayle compared reason to a corrosive powder that first eats up errors, but then goes on to eat up truths, "When it is left on its own, it goes so far that it no longer knows where it is, and can find no stopping place."

FAITH

Each time Bayle reached this point he would proclaim that in view of the inability of reason to arrive at any complete and adequate conclusion about anything, man should abandon the rational world and seek a different guide: faith. This claim was forcefully stated in the articles "Bunel, Pierre," "Charron," "Manichaeans," "Pomponazzi," "Pyrrho," and the "Clarification on the Pyrrhonians." Bayle's dwelling on the theme that reason makes men perplexed and so requires that they look for another guide suggests, perhaps, that his purpose was something like that of Maimonides in *The Guide of the Perplexed*, one of Bayle's favorite works.

REVELATION

In various discussions (such as the articles "Pyrrho," "Simonides," and the "Clarification on the Pyrrhonians") Bayle insisted that the rational and the revealed worlds are in complete conflict, because the latter is based on

claims that are in direct opposition to the principles that appear most evident to reason. Starting with the first line of Genesis, the world of faith contains claims that are rationally unintelligible and unacceptable. According to Bayle the principle that reason finds the most evident and certain is that nothing comes from nothing, whereas faith reveals that God created the world *ex nihilo*. Similarly, the most acceptable rational moral principles are at complete variance with the revealed accounts of the behavior of the heroes of the faith, the leading figures of the Old Testament. In this total opposition between reason and revelation, faith is man's only refuge. Bayle insisted that his irrational fideism was the traditional orthodox position from St. Paul and Quintus Septimius Florens Tertullian down to Calvin and Jurieu. (In fact, some passages of Bayle sound like Søren Aabye Kierkegaard and other more fideistic theologians.)

BAYLE'S RELIGIOUS POSITION

No matter how often Bayle claimed that he was advocating the faith and was merely restating what orthodox Christians had always said, his opponents, especially Jurieu and some of the liberals, insisted that Bayle was actually an unbeliever trying to destroy the faith by making it sound as ridiculous and irrational as possible. Certainly, some of Bayle's passages have such a ring. And none of his statements of the fideistic message have the anguish of Blaise Pascal or Kierkegaard, or even the despair of the truth seeker unable to find satisfaction in either the rational world or in revealed truths.

However, this may not necessarily be a sign that Bayle was insincere. Bayle himself offered an alternative possibility in a discussion in the longest article in the *Dictionary*, that on Spinoza. In note M Bayle described two kinds of people, those who have religion in their minds, but not in their hearts, and those who have religion in their hearts, but not in their minds. The first kind are convinced of the truth of religion, but their consciences are not affected by the love of God. The second kind lose sight of religion when they seek it by rational means and are lost in the wilderness of the pros and cons; but when they listen only to their feelings, conscience, or education, they find that they are convinced of religion and regulate their lives accordingly, within the limits of human frailties. If Bayle had religion in the heart in this sense (rather than Pascal's), it was an emotionless religion, which became confused and perplexing whenever he tried to explain or comprehend it. When he abandoned the attempt to be rational about it, then it became a calm guide for a life of pious study.

In the article "Bunel, Pierre" Bayle presented this fervorless religion as almost a testimonial of faith. Bunel, an obscure Renaissance pedant from Toulouse (who accidentally had an enormous influence on the development of modern skepticism by giving Raimond Sebond's *Natural Theology* to Montaigne's father) is one of the few genuine heroes of Bayle's *Dictionary*. He was pictured as a perfect Christian, in contrast to myriad imperfect ones (including Jurieu), because he rejected all worldly goals and devoted himself solely to the life of the pure scholar, harming no one and seeking truth. Bayle's own life was much like Bunel's. Beyond this, Bayle's religion seems to have had little or no content, though he always claimed to be a Calvinist Christian.

The lack of content in Bayle's religion may account for his important doctrine of toleration of the rights of the erring conscience. In many works Bayle insisted that man's ultimate appeal for justification of his beliefs and actions was his own conscience and that man had no further ultimate standard to employ to determine if his conscience was correct. Therefore, each man could act only as he saw fit, and no one was justified in trying to compel another to act contrary to the dictates of his conscience, erring or otherwise.

Though Bayle continually presented his appeal to faith, and his own faith, in tranquil and colorless terms, a fundamental problem remains of determining what Bayle did in fact believe and what his arsenal of doubts was intended to achieve. Shaftesbury, who knew Bayle well, called him "one of the best of Christians." Jurieu was sure he was an atheist. The Enlightenment leaders saw him as one of them, perhaps a deist, but definitely a scoffer at all historical religions. The biographical data would suggest that, barring some strange private joke, Bayle was committed to some aspects of the French Reformed church. He persisted in belonging to it, attending it, and proclaiming his sincere adherence to it, no matter how much he was abused by Jurieu and others. He could have lived and prospered in Holland either in a more liberal church or as a complete independent. In tolerant Holland it was extremely unlikely that he would have been punished or have had his works censored, no matter what he said or believed.

Coming from a family that suffered inordinately from persecution for its Calvinism, Bayle may have felt a need and desire to maintain his original tradition. His last message to a friend as he knew his life was ending was, "I am dying as a Christian philosopher, convinced of and pierced by the bounties and mercy of God, and I wish you a perfect happiness." Elisabeth Labrousse (1963) points

out that this is a most minimal Christian testament, since Jesus is not mentioned, nor any Christian doctrine, nor anything about Bayle's church. In his writings Bayle rarely discussed religion without making Manichaeanism or Judaism seem either more plausible or more significant than Christianity; and he occasionally (as in the article "Takiddim") even called Judaism the true religion. Bayle may have been either a Christian in his own sense or actually a Manichaean or Judaizer or both, working out an enormous defense of his cause by undermining the rational and moral foundations of other possibilities.

Until it is possible to ascertain Bayle's actual beliefs, it will remain extremely difficult to determine his aims and whether the impact he had was the intended one. Bayle undermined all the philosophical positions of the great seventeenth-century metaphysicians and posed basic problems that Berkeley, Hume, Voltaire, and others were to use to establish other approaches and alternatives. He provided an enormous amount of argument and ridicule for the Enlightenment to use in destroying the intellectual ancien régime and in launching the Age of Reason. But even Voltaire and Hume were aware that Bayle was much more given to doubt and destructive criticism than they considered themselves to be. At times, they believed they had found new ways of overcoming Bayle's doubts. Perhaps they were both too far removed from Bayle's calm religious haven to be able to entertain his complete doubt about everything without utter dismay and horror.

Bayle seems to have lived in a different world from that of the Enlightenment that he helped produce. Though he may not have been "the greatest master of the art of reasoning," as Voltaire called him, he was one of the best. He was a genius at seeing how to attack and destroy theories about almost anything and a master at determining what the facts in the case were. Bayle would turn his attacks against everyone and everything, modern, ancient, scientific, rationalistic, or religious. He did not, apparently, see a new and better world emerging from his critique, nor see the need for one. The havoc he was wreaking seemed to leave him completely tranquil. It was for subsequent generations to discover the problem of living in a world in which all is in doubt and in which the solution proffered by Bayle seems meaningless or unattainable.

Some scholarship focuses on Bayle's last writings after the *Dictionary*. Gianluca Mori (1999) and others believe that they have found that Bayle was evolving more positive views in his last few years.

See also Aristotelianism; Aristotle; Arnauld, Antoine; Berkeley, George; Boyle, Robert; Calvin, John; Cartesianism; Descartes, René; Diderot, Denis; Enlightenment; Feuerbach, Ludwig Andreas; Fideism; Gibbon, Edward; Hobbes, Thomas; Hume, David; Jefferson, Thomas; Kierkegaard, Søren Aabye; Leibniz, Gottfried Wilhelm; Locke, John; Maimonides; Malebranche, Nicolas; Mandeville, Bernard; Mani and Manichaeism; Montaigne, Michel Eyquem de; Newton, Isaac; Pascal, Blaise; Plato; Pope, Alexander; Religion and Morality; Renaissance; Sextus Empiricus; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Skepticism, History of; Spinoza, Benedict (Baruch) de; Tertullian, Quintus Septimius Florens; Voltaire, François-Marie Arouet de; Zeno of Elea.

Bibliography

WORKS BY BAYLE

Oeuvres diverses. 4 vols. The Hague: n.p., 1727. A collection of Bayle's writings other than the *Dictionary*.

A General Dictionary, Historical and Critical. 10 vol. Translated by John Peter Bernard, Thomas Birch, and John Lockman. London: J. Bettenham, 1734–1741. This was published in French under the title *Dictionnaire historique et critique* in 1740.

Historical and Critical Dictionary, Selections. Translated by Richard H. Popkin, with Craig Brush. Indianapolis, IN: Bobbs-Merrill, 1965. Contains selections from forty articles, plus Bayle's clarifications.

Correspondance de Pierre Bayle, edited and annotated by Elisabeth Labrousse et al. Oxford, U.K.: Voltaire Foundation, 1999–2004.

Bayle: Political Writings, edited by Sally L. Jenkinson. New York: Cambridge University Press, 2000.

WORKS ON BAYLE

Barber, W. H. "Bayle: Faith and Reason." In *The French Mind: Studies in Honour of Gustav Rudler*, edited by Will Moore, Rhoda Sutherland, and Enid Starkie, 109–125. Oxford, U.K.: Clarendon Press, 1952. Suggests that Bayle was sincerely religious.

Bracken, Harry M. "Bayle Not a Sceptic?" *Journal of the History of Ideas* 25 (1964): 169–180. Attempts to clarify the sense in which Bayle was a skeptic and a fideist.

Bracken, Harry M. "Pierre Jurieu: The Politics of Prophecy." In *Millenarianism and Messianism in Early Modern European Culture. Vol. 4: Continental Millenarians: Protestants, Catholics, Heretics*, edited by John Christian Laursen and Richard H. Popkin, 85–94. Boston: Kluwer Academic, 2001.

Courtines, Léo. *Bayle's Relation with England and the English*. New York: Columbia University Press, 1938. Deals with Bayle's contacts with and influence on English philosophers, theologians, and writers, including Berkeley and Hume.

Dibon, Paul, ed. *Pierre Bayle, le philosophe de Rotterdam*. Amsterdam: Elsevier, 1959. An important collection of articles in French and English reevaluating Bayle's views.

Feuerbach, Ludwig. *Pierre Bayle: Ein Beitrag zur Geschichte der Philosophie und Menschheit*. Berlin: Akademie-Verlag, 1967. Bayle seen as caught between rationalism and the irrationality of Christianity.

Hasse, Erich. *Einführung in die Literature des Refuge*. Berlin: 1959. A monumental study of the French Protestant refugees. Places Bayle in the context of this group.

Hazard, Paul. *The European Mind, 1680–1715*. Translated by J. Lewis May. London: Hollis and Carter, 1953. Originally published in French under the title *La Crise de la conscience européenne, 1680–1715* in 1935. The intellectual climate of Bayle's time.

James, E. D. "Scepticism and Fideism in Bayle's Dictionnaire." *French Studies* 16 (1962): 307–324. Challenges the views of Popkin and others regarding Bayle's religious views.

Kemp Smith, Norman. *The Philosophy of David Hume: A Critical Study of Its Origins and Central Doctrines*. London: Macmillan, 1941. Contains analyses of sections in Bayle that influenced Hume, especially in his *Treatise of Human Nature*.

Labrousse, Elisabeth. *Pierre Bayle. Vol. 1: Du pays de foix à la cité d'Erasme*. The Hague: Nijhoff, 1963. First biography of Bayle in recent times, based on monumental researches.

Labrousse, Elisabeth. *Pierre Bayle. Vol. 2: Hétérodoxie et rigorisme*. The Hague: Nijhoff, 1964. A study of Bayle's theology by the leading authority today. Contains a massive bibliography.

Laursen, John Christian. "Bayle's Anti-millenarianism: The Dangers of Those Who Claim to Know the Future." In *Millenarianism and Messianism in Early Modern European Culture. Vol. 4: Continental Millenarians: Protestants, Catholics, Heretics*, edited by John Christian Laursen and Richard H. Popkin, 95–106. Boston: Kluwer Academic, 2001.

Mason, H. T. *Pierre Bayle and Voltaire*. London: Oxford University Press, 1963. A comparison, with an attempt to assess what Voltaire borrowed from Bayle.

Mason, H. T. "Pierre Bayle's Religious Views." *French Studies* 17 (1963): 205–217. Defends interpretation of Bayle as an irreligious thinker.

Mori, Gianluca. *Bayle philosophe*. Paris: Champion, 1999.

Norton, David. "Leibniz and Bayle: Manicheism and Dialectic." *Journal of the History of Philosophy* 2 (1964): 23–36. An attempt to see how Bayle might have dealt with Leibniz's *Theodicy* and a new analysis of Bayle's dialectic.

Paganini, Gianni, ed. *The Return of Scepticism from Hobbes and Descartes to Bayle*. Dordrecht, Netherlands: Kluwer Academic, 2003. Includes several articles on Bayle and eighteenth-century skepticism.

Popkin, Richard H. *The High Road to Pyrrhonism* edited by Richard A. Watson and James E. Force. San Diego, CA: Austin Hill Press, 1980. Contains articles on Bayle in the context of the late seventeenth century and on his legacy for the eighteenth century.

Rétat, Pierre. *Le "Dictionnaire" de Bayle et la lutte philosophique au XVIIIe siècle*. Paris: Les Belles Lettres, 1971.

Rex, Walter. *Essays on Pierre Bayle and Religious Controversy*. The Hague: Nijhoff, 1965.

Sandberg, K. C. "Pierre Bayle's Sincerity in His Views on Faith and Reason." *Studies in Philology* 61 (1) (1964): 74–84. Interprets Bayle as a sincere Calvinist. Shows Bayle had no need to fear censorship.

Whalen, Ruth. *The Anatomy of Superstition: A Study of the Historical Theory and Practice of Pierre Bayle*. Oxford, U.K.: Voltaire Foundation, 1989.

Richard Popkin (1967, 2005)

BEARDSLEY, MONROE

(1915–1985)

Monroe C. Beardsley published in several areas of philosophy but is best known as an aesthetician. He is arguably the most important figure of twentieth-century analytic aesthetics. His *Aesthetics: Problems in the Philosophy of Criticism* (1958) was a watershed book, furnishing an organization aesthetics had lacked. Beardsley's careful discussions of almost all of the field's questions provided an aesthetic education for his and succeeding generations.

Two ideas shaped all Beardsley's work: his view of the philosophy of art criticism (called "metacriticism") and his aestheticism. Metacriticism's task is the analysis of art criticism's central concepts. Aestheticism is the view that aesthetic characteristics (e.g., unity, delicacy) alone are the proper objects of art criticism; thus, aesthetic features become the sole focus of criticism and the basis for artistic value. Beardsley acknowledged that artwork can have nonaesthetic, referential characteristics, and he does not deny that these features are important. He does, however, deny that referential features are relevant to aesthetic experience and, thus, to artistic value.

Aesthetics begins with the metacritical task of discussing objects of criticism, designating them "*aesthetic* objects"; a hard-and-fast connection is, thus, forged at the book's beginning between metacriticism and aestheticism with the contents of the objects of criticism identified as aesthetic features. This identification sets the stage for Beardsley's view of artistic value. He claims that artworks are instrumentally valuable because their aesthetic characteristics can produce (valuable) aesthetic experience. Aesthetic experience, as he conceives of it, is the foundational notion of Beardsley's book.

John Dewey is the primary source of Beardsley's notion of aesthetic experience. Dewey conceived of aesthetic experience as an experience that coheres to such an extent that it is set off, although not detached, from the flow of experience. Beardsley, however, was also influenced by aesthetic-attitude theorists. Consequently, unlike Dewey, he claimed aesthetic experience is detached from ordinary experience. But whereas the aesthetic-attitude theorists claim various mental mechanisms—

such as "psychical distancing"—detach aesthetic experience from ordinary life, Beardsley maintained that aesthetic experience's internal coherence detaches it from the flow of experience. And it is the detachedness of aesthetic experience that blocks artworks' referential characteristics (names, descriptions, portrayals, etc.) from referring to anything outside of ongoing aesthetic experience of artworks. On his view, only aesthetic, nonreferential characteristics of an artwork can cause aesthetic experience and, thereby, be the focus of artistic criticism and the basis for artistic value.

Beardsley argued that artistic value is an instrumental value (an objective value) because it can cause valuable aesthetic experience. To provide an objective basis for the value of aesthetic experience, Beardsley contended that aesthetic experience is in turn instrumentally valuable, being productive of human welfare. As an aspect of his account of artistic value, Beardsley argued that there are principles of art criticism involving the potential of three aesthetic features (unity, intensity, and complexity) for producing aesthetic experience, thus, opposing the conventional wisdom that there are no such principles. Present-day accounts of critical principles have their beginnings in Beardsley's work.

Throughout his career Beardsley continued to defend aestheticism and the inherent detachment of the aesthetic from ordinary life. In 1978 he argued against Nelson Goodman's view that artworks' referential features produce artistic value. In the second edition of *Aesthetics*, Beardsley wrote, "I think distance or detachment—withdrawal from practical engagement—in some form … is a factor in aesthetic character" (1981, p. lxii).

The only major question not discussed in *Aesthetics* is the nature of art. Finally in 1979, responding to the art theories that developed in the wake of Arthur Danto's "The Artworld" (1964), Beardsley sketched a theory of art in the midst of discussing aesthetic value; he wrote, "… an artwork can be usefully defined as an intentional arrangement of conditions for affording experiences with marked aesthetic character" (1979, p. 729). Beardsley's theory of art was determined by his aestheticism.

In 1946 Beardsley and William Wimsatt had coauthored "The Intentional Fallacy" and initiated a polarizing debate by arguing that artists' intentions are irrelevant to the interpretation and evaluation of their artworks. Beardsley also defended anti-intentionalism in *Aesthetics*, "The Authority of the Text" in *The Possibility of Criticism* (1970), and "Intentions and Interpretations: A Fallacy Revived" in *The Aesthetic Point of View* (1982). On his anti-intentionalist account, artworks are severed from their creators' actions when they are objects of criticism and of aesthetic experience. According to both his anti-intentionalism and his aestheticism, artworks as objects of aesthetic experience and criticism are detached—on the one hand, from their creators and, on the other, from their referents. Thus, in an aesthetic experience, audiences and critics savor only the aesthetic features of artworks.

Anti-intentionalism has been debated on grounds other than those used in *Aesthetics*, making use of arguments from the philosophy of language. Beardsley himself participated in this later controversy and produced additional arguments against intentionalism in "The Authority of the Text" and in "Intentions and Interpretations: A Fallacy Revived." In the first article, he argued that three different kinds of texts created without any authorial intent have specific meanings, namely, some randomly generated computer texts, some poetic lines with a word that has come to have a different meaning than it had at the time it was composed, and texts that reveal meanings of which their authors were unconscious. Unfortunately, Beardsley's argument merely contradicts the intentionalists' claim that such texts cannot have meaning and, therefore, will not persuade them.

In the second article, Beardsley applies J. L. Austin's distinction between locutionary and illocutionary acts to fictional discourse, claiming that illocutionary acts in fiction are *representations* of illocutionary acts and thus not actual illocutionary action of the text's author. Unfortunately, this argument is limited to fiction and the dispute is about texts generally, not just fiction. Furthermore, the dispute is really about locutionary meaning rather than illuctionary meaning.

The controversy over intentionalism continues.

See also Aesthetics, History of.

Bibliography

WORKS BY MONROE BEARDSLEY

Aesthetics: Problems in the Philosophy of Criticism (1958). Indianapolis: Hackett, 1981.

"The Authority of the Text." In *The Possibility of Criticism*, 16–37. Detroit: Wayne State University Press, 1970.

"In Defense of Aesthetic Value." *Proceedings and Addresses of the American Philosophical Association* 52 (1979): 723–749.

"Intentions and Interpretations: Fallacy Revived." In *The Aesthetic Point of View*, edited by Michael C. Wreen and Donald M. Callen, 188–207. Ithaca, NY: Cornell University Press, 1982.

WORKS ABOUT MONROE BEARDSLEY

Dewey, John. *Art as Experience*. New York: Minton, Balch, 1934.

Dickie, George, and W. Kent Wilson. "The Intentional Fallacy: Defending Beardsley." *The Journal of Aesthetics and Art Criticism* 53 (1995): 233–250.

George Dickie (2005)

BEATTIE, JAMES
(1735–1803)

James Beattie was born in Laurencekirk, Scotland, on October 25, 1735. He received an MA at Marischal College, Aberdeen, in 1753, became schoolmaster at the Fordoun Parish Church, and in 1760 was appointed Professor of Moral Philosophy and Logic at Marischal College. He was a member of the Aberdeen Philosophical Society with Thomas Reid and other notable Scottish writers. Beattie was known internationally as both a philosopher and poet. His principal philosophical contribution is *An Essay on the Nature and Immutability of Truth* (1770), for which he was awarded a yearly pension of £200 by King George III. His relentless attacks on David Hume in that work sparked a controversy that permanently linked his name with Hume's. He was ill much of his life and endured the progressive insanity of his wife and the early death of his children. He died on August 18, 1803.

Beattie's *Essay* is an interesting critique of modern metaphysics as well as an important assault on Hume. The crux of his position is this: Truth is that which common sense "determines me to believe," and skeptical metaphysicians have erred by ignoring commonsense intuitions. He discusses eight types of human reasoning that are grounded in common sense: mathematics, external sensation, internal sensation such as moral approval and personal identity, memory, causality, induction, analogy, and testimony. He acknowledges that merely having a commonsense belief does not guarantee that such a belief is true, since one can never be in a privileged position to compare one's commonsense beliefs to absolute reality. Like René Descartes, though, Beattie argues that one can trust that God has not deceived one in giving one faulty commonsense intuitions (2000). He argues further that denying the truth of common sense leads to absurd consequences.

Beattie takes issue with the skeptical trend of modern philosophers since Descartes who begin, he holds, with a few presumably factual general principles and deduce from these a range of noncommonsensical conclusions that call into doubt one's senses, the external world, free will, memory, and any of the previously mentioned eight types of reasoning. Skeptical metaphysics, he argues, is loathsome and harmful to normal affairs of life. About one-fourth of the *Essay* is a criticism of Hume's views of personal identity, ideas and impressions, necessary connection, the broad scope of the virtues, the natural inferiority of blacks, and other issues. His rhetoric against Hume is harsh, and in a 1771 postscript to the *Essay* he states that this treatment is necessary for placing the absurdity of skeptics' views in perspective and to combat the danger that skeptics pose to morality. He writes, "Let opinions then be combated by reason, and let ridicule be employed to expose nonsense."

In addition to his polemical *Essay* Beattie published *Dissertations Moral and Critical* (1783) on the subjects of memory, imagination, and language, *Evidences of the Christian Religion* (1786), and a collection of his philosophy lectures titled *Elements of Moral Science* (1790–1793). One of his more provocative pieces is the allegorical short story "The Castle of Scepticism," which he circulated among friends but that remained unpublished for almost 200 years. It describes how, after falling asleep, he was led on a journey to a surreal land of skeptics who defied commonsense beliefs. During and shortly after his life, Beattie's *Essay* was defended by Thomas Blacklock (1720–1791) and Dugald Stewart, and criticized by Joseph Priestley, James Steuart (1712–1790), and Thomas Cogan (1736–1818), in writings all of which are reprinted in *Early Responses to Reid, Oswald, Beattie, and Stewart* (2000).

See also Common Sense.

Bibliography

WORKS BY BEATTIE

The Works of James Beattie. 10 vols., edited by Roger J. Robinson. Bristol, U.K.: Routledge/Thoemmes Press, 1996.

An Essay on the Nature and Immutability of Truth, in Opposition to Sophistry and Scepticism, edited by James Fieser. Bristol, U.K.: Thoemmes Press, 2000.

"The Castle of Scepticism." In *Early Responses to Hume's Life and Reputation*. 2 vols., edited by James Fieser. Bristol, U.K.: Thoemmes Press, 2003.

WORKS ABOUT BEATTIE

Fieser, James, ed. *Early Responses to Reid, Oswald, Beattie, and Stewart*. Bristol, U.K.: Thoemmes Press, 2000.

Forbes, Margaret. *Beattie and His Friends*. Westminster, U.K.: Archibald Constable, 1904.

Forbes, William. *An Account of the Life and Writings of James Beattie: Including Many of His Original Letters*. 2 vols. Edinburgh, Scotland: Archibald Constable, 1806.

Robinson, Roger J., ed. *The Correspondence of James Beattie*. 4 vols. Bristol, U.K.: Thoemmes Continuum, 2004.

James Fieser (2005)

BEAUTY

Until the eighteenth century, "beauty" was the single most important idea in the history of aesthetics. One of the earliest works in the literature of aesthetics, the *Hippias Major* (probably by Plato), was addressed to the question, "What is beauty?" Around this question most of later thought revolves. The treatment of the other major concept, *art,* when it is not ancillary to that of beauty, lacks comparable generality, for it is often restricted to a single artistic form or genre, or its theoretical status is equivocal, because art is taken as identical with craft or skill. The modern notion of the *fine arts* did not appear until the eighteenth century and, more important, it was then too that the concept of *aesthetic experience* was first formulated systematically. As a consequence, beauty lost its traditional centrality in aesthetic theory and has never since regained it.

Our survey of these historical developments will be selective. Specific theories will be singled out because they are paradigms of the major kinds of theory of beauty. Thus, where beauty is taken to be a property, we will be less concerned with what, on some particular proposal, this property is, more with the logical relations of beauty, so construed, to the other properties of beautiful things and to the conditions of its apprehension. Where it is not so construed, the chief alternative meanings for beauty will be illustrated. *Beautiful* is used to esteem or commend and therefore to make a claim that is honored in the processes of criticism. Throughout this article, accordingly, the implications of the major kinds of theory for evaluation of the object will be traced.

CLASSICAL AESTHETICS

The concluding section of Plato's *Philebus* is the prototype of the dominant ways of thinking about beauty prior to the eighteenth century. This will be shown by unpacking its major theses, which, whether they were taken over or whether they became the focuses of dispute, made up the framework of classical theory and defined its preoccupations.

The discussion of beauty in the *Philebus,* as in other dialogues, arises in the course of discussion of a larger question not itself aesthetic, namely, whether pleasure or knowledge is the supreme good for humankind. Socrates wished to distinguish "pure" from "mixed" pleasures, and among the examples that he gives of the former are the pleasures evoked by objects that are "beautiful intrinsically." He cited simple geometrical shapes, single colors, and musical notes (50E–52B).

The first thing to see is that Plato took beauty to be a property ingredient in things. It is nonrelational twice over, for its existence is not dependent upon, or affected by, perceiving it; and whereas "relative" beauty exists only by virtue of comparison with things that are of a lesser degree of beauty or simply ugly, "intrinsic" beauty does not. This view can be specified in two different ways, both of which appear to be suggested by Plato: Either the property of beauty is identified with, and defined by, certain properties of the object, here the determinate ordering or "measure" of the whole (64E), or beauty is itself indefinable, but supervenes upon a further, distinct property, the internal unity of the parts, which is the condition of its existence (66B).

On the former theory, whether a thing is beautiful is decided just by finding whether it does or does not possess the salient property. In the *Philebus,* the success of such inquiry, even on Plato's rigorous conception of knowledge, is assured by the markedly intellectualist character of measure. It is a formal or structural property and therefore cognate with the nature of intelligence (59B–C, 65D), unlike matter which is opaque to mind. It is no accident that, having illustrated intrinsic beauty by objects produced by the "carpenter's rule and square," Socrates later eulogized carpentering for its cognitive exactness (55D–56E). This insistence on the clarity and knowability of beauty (shared by Aristotle in *Metaphysics* 1078b) is also reflected in the choice of sight and hearing, the senses most appropriate to rational cognition, as the sole avenues of the perception of beauty (cf. *Phaedrus* 250D).

The nondefinist theory is, for the reasons to be cited in later philosophers, more plausible but considerably more complicated. This theory is that, given unity in variety in a thing, beauty is also necessarily present. It will still be true that whether a thing is beautiful can be decided by showing that it possesses internal unity if—but this proviso is crucial—we can be certain that the two properties do, in all instances, exist together. Hence we must be able to apprehend beauty in its own right. Yet to say that beauty is indefinable is to say that what it is cannot be identified conceptually and therefore in commonly understandable terms. The cognitive assurance and stability of definist theory may be lost as a result. Plato was amply aware of the possibility of uncertainty and dis-

agreement among judgments of beauty (*Laws* Bk. II). The account of intrinsic beauty in the *Philebus* guards against these dangers. Things are beautiful intrinsically precisely because they are "always beautiful in their very nature" (51C–D). Though the objects cited by Socrates are empirical—"the surfaces and solids which a lathe, or a carpenter's rule and square, produces from the straight and the round"—they nevertheless enjoy the self-identity, unaffected by adventitious or contextual factors, that is also characteristic of the Platonic Ideas. Unlike objects of relative beauty, they resemble the ideal beauty described in the *Symposium* (211–212), which cannot be "fair in one point of view and foul in another" (cf. *Republic* 479). Socrates held that they will necessarily arouse in the beholder a kind of pleasure that is peculiar to intrinsic beauty (51D). That the apprehension of such beauty will be veridical is further assured in the *Philebus* by the notion of "pure" pleasures, that is, those unmixed with pain. Pain warps or falsifies judgment (36C et seq.), but it is never present in the appreciation of intrinsic beauty. The related concepts of the *intrinsic* and the *pure* are used to guarantee the stability of the experience of beauty. They lead, however, to a severe delimitation of the class of beautiful objects. Paintings and living creatures are excluded as relative, tragedy and comedy (50A–B) because they are impure. Human significances are hostile to beauty because they encourage error and diversity in our responses to it.

In its analysis of the concrete phenomena of beauty, the *Philebus* is distinguished from the mythic and metaphysical approaches of the *Phaedrus* and *Symposium* and the social moralism of the *Republic* and *Laws.* Even here, however, the beautiful does not constitute a distinct and autonomous subject matter. It is treated as a "form" or mode of goodness in general, and the term *beautiful* is used, as it was by the Greeks generally, interchangeably with *excellent, perfect,* and *satisfying.* It is also worthy of note that the concept of art enters in hardly at all. Painting and literature are mentioned only so that they may be excluded. By contrast, Aristotle's *Poetics* devotes itself to a single art form, tragedy, making only a casual reference to beauty —measure is a necessary condition (VII). Later treatments of beauty and art are even less congenial to our modern conception of aesthetics, which led the historian Bernard Bosanquet to speak of a centuries-long "intermission" in aesthetics between the Greco-Roman and the modern eras. The metaphysic of Plotinus, which derived from Plato, is spiritualist and Idealist; and here, as in later philosophy, the bias of such thought is to encourage regard for, and insight into, the experience of beauty. The soul is said to strive toward beauty, which is a manifestation of the spiritual force that animates all of reality. It is just because of the vitality and moving appeal of beauty that Plotinus rejected the identification of beauty with a merely formal property. The living face and the dead face are equally symmetrical, but only the former stirs us. Hence "beauty is that which irradiates symmetry rather than symmetry itself" (*Enneads* VI; VII, 22). Further, some simple, sensory objects lacking internal structure are beautiful, and, finally, symmetry is present in some ugly things as well (I; VI, 1). Plotinus's critique of formalism effectively made the larger point that beauty cannot be identified with any single element of the object, form or any other. It is the total object, the whole of form and expressiveness and what the form is of, that possesses beauty. If, on the other hand, beauty is thought to be a global quality that "irradiates" this object and moves us, it is difficult or impossible, in a definition, to specify conceptually the nature of this quality. Moreover, Plotinus's argument cast doubt on the possibility of finding even the conditions of beauty. A formal property such as symmetry is the most likely candidate, because it can be shared by objects that are otherwise highly diverse, artistic or natural, abstract or representational, sensory or mathematical. Yet if the negative instances cited by Plotinus show that this property is not even a universal concomitant of beauty, then a fortiori it cannot be the necessitating ground of beauty.

Still, the effort to enunciate a set of conditions for beauty is persistent in Western thought, because it answers to the desire for a criticism whose verdicts will be certifiable and authoritative. The high noon of such criticism was the neoclassical period, particularly the sixteenth and seventeenth centuries, when the conditions were detailed and formalized, and endowed with the institutional sanctions of the new "Academies." A multiplicity of treatises were devoted to particular arts or genres, each of which was taken to be subject to "rules," inherent in its specific nature and function, which can be rationally known (e.g., Castelvetro, Palladio). The treatises borrowed heavily from their Greek and Roman antecedents—Aristotle, Horace, Vitruvius. The "lawmakers of Parnassus" thereby invested their claims to speak on behalf of "reason" and "nature" with the authority of antiquity. Given that beauty is an objective property, attainable artistically and knowable critically, by reference to the rules, the question of the percipient's response to it was scanted. As in the *Philebus,* beauty can be expected to arouse the appropriate response, which was referred to briefly and loosely as "pleasure," or "delight."

THE EIGHTEENTH CENTURY

The rebellion against the rules, in the name of the spectator's felt response—"the taste is not to conform to the art, but the art to the taste" (Addison)—intimates, in art criticism, the larger and more profound reconstruction of thought that took place in aesthetic theory. In the eighteenth century, indeed, aesthetics first established itself as an autonomous philosophical discipline. It defined a subject matter that is not explicable in terms of any of the other disciplines and is therefore taken out of the metaphysical and moral context of much traditional aesthetics, to be studied in its own right. The pioneer work is to be found in the prolific and assiduous writings of the British who, throughout the century, carried out the inquiry that Addison, at its beginning, justly described as "entirely new."

The century was a Copernican revolution, for instead of looking outward to the properties of beauty or the art object, it first examined the experience of the percipient, to determine the conditions under which beauty and art are appreciated. The decisive condition is disinterestedness, that is, perception directed upon an object without, as in practical or cognitive activity, any purpose ulterior to the act of perception itself. In aesthetic theory so conceived, beauty is no longer the central concept. It now stands for just one kind of aesthetic experience among others, and it can be defined and analyzed only by reference to the logically more basic concept of *aesthetic perception.*

The introspective examination of our "ideas," stimulated by John Locke's *Essay,* discloses experiences that differ significantly, in their felt quality, from that of beauty. This century distinguished a great many other "species" of aesthetic response, but the most important was that of sublimity. Sublimity is profoundly unlike beauty, for whereas the latter arouses "joy" and "cheerfulness," the feeling of the sublime is "amazement" and awe. Still, most of the British hold that the two can coexist and that the experience of both is pleasurable. The most drastic distinction was drawn by Edmund Burke (1757), who argued that beauty and sublimity are, conceptually, mutually exclusive and, existentially, antithetical. He at the same time limited the range of beauty severely and pushed back the boundaries of the aesthetic to include a radically different kind of experience, which cannot be accommodated in the traditional category. Indeed Burke clearly considered the experience of sublimity to be the more valuable of the two. Both Moses Mendelssohn and Immanuel Kant read Burke and were greatly affected by him, and through their influence Burke's critique of beauty made a lasting impression on Continental thought.

Burke granted that a beautiful object arouses pleasure, but he argued that a sublime object, that is, one that is "terrible," even though it is apprehended disinterestedly, arouses "some degree of horror." Beauty "relaxes," but the experience of sublimity is of great emotional intensity. The two experiences are therefore incompatible with each other. Moreover, the properties that Burke attributed to sublime objects are just the opposites of those that the *Philebus* had enshrined in the classical conception of aesthetic value. Against clarity and lucidity, Burke urged that we are moved most greatly by what is "dark, uncertain, confused." In place of formal ordering, Burke eulogized what is "vast" and "infinite." The sublime therefore renders beauty "dead and unoperative." When beauty had been taken as the sole value category, ugliness, its contradictory, had necessarily been excluded from aesthetic value. Burke went so far as to suggest that even the ugly can be an object of aesthetic appreciation. In all this, he is pointing the way to the nineteenth-century and twentieth-century concept of *expression,* which, more catholic by far than classical beauty, admits a limitless diversity of subject matter, treatment, and form, if only the work of art be moving and powerful.

A comparable challenge to the classical values of order and serenity came from another direction. The historical study of art, pioneered by Johann Joachim Winckelmann (1764), disclosed that these values are found only in relatively limited epochs and styles, even, indeed, of Greek art itself. Later research emboldened the protest against the once unchallenged arbiters of classical and neoclassical criticism that they had identified selected stylistic properties of Greek and High Renaissance art with what is beautiful "naturally" and universally.

In the eighteenth century, also, the "logic" of beauty underwent a profound sea change. Francis Hutcheson (1725) announced a new locus for beauty: "Let it be observed, that in the following papers, the word *beauty* is taken for *the idea raised in us.*" It follows that any object whatever that does in fact excite this idea must be judged to be beautiful. But this invites the possibility of diverse and conflicting judgments that, if subjective response is the sole and decisive test, must all be accepted as equally valid. Are there, however, any properties peculiar to beautiful objects, which can be pointed to, to legitimate certain judgments and whose absence will show others to be mistaken? Hutcheson thought that there was—the classical property of "uniformity in variety." Yet to be consistent with the definition of beauty with which he began, he

had to guarantee that things possessing this property would uniquely and universally arouse the appropriate idea. It can be said summarily that he failed to do so, and his failure is instructive. It points up the tension between the old and the new ways of thinking, between taking beauty to be an inherent, nonrelational property and using *beauty* to refer to the capacity of things to evoke a certain experience. A capacity is not, however, an observable property in things like uniformity. It must be interpreted as either a very different sort of property or else it is not a property at all. David Hume drew out the radical implications of Hutcheson's initial meaning for beauty with the acute remark that Euclid described all of the properties of a circle, but beauty is not among them ("The Sceptic").

In general, the later British aestheticians did not take *beautiful* to denote a property. Necessarily, therefore, the logical status of the properties that they attribute to beautiful objects—proportion, utility, and so on—is correspondingly altered. Such properties are no longer, as in the *Philebus,* either identical with, or the conditions of, a property of beauty. They are, rather, causes of the experience of beauty. Even so considered, however, the traditional formulas of beauty were brought under fire throughout the eighteenth century. Since the attribution of causes can be justified only by the evidence of their effects in experience, the British, arguing from the things that people do in fact find beautiful, showed that none of these properties are shared by all these things. There was also the more subtle and damning criticism that the traditional formula of "unity in variety" is simply devoid of meaning, because it applies indiscriminately to any object whatever. By the close of the century, Alison (1790) concluded that any attempt to find properties common and peculiar to beautiful objects is "altogether impossible." Finally it was suggested that "beautiful" is just "a general term of approbation" (Payne Knight, 1805).

The British thereby generated the problem that is central to Kant's *Critique of Judgment* (1790): How, if the aesthetic judgment arises from subjective feeling and predicates nothing of the object, can it claim to be more than an autobiographical report and can, indeed, claim to be universally binding?

THE NINETEENTH AND TWENTIETH CENTURIES

The most novel development in this period has been the attempt at a scientific approach to aesthetics. This has taken two forms, generally, and the status of beauty in each is worth noting. Psychological aesthetics applies experimental methods to aesthetic experience in an effort to work out "laws" of appreciation. These are to be derived from the consensus of pleasure and displeasure reported by the laboratory subject in the face of various objects. When beauty is used at all in speaking of these objects, as it was by Gustav Theodor Fechner (1876), it is a loose, omnium-gatherum term. The objectivist-formalist connotations of the word have made it increasingly unsatisfactory to later psychologists. Either they have stipulated that it refers to certain psychological responses (e.g., O. Külpe, 1921), or they have abandoned it in favor of the more apt "liberal and comprehensive" (E. Bullough, 1907) concept of "aesthetic value." The last decades of the nineteenth century also saw the rise of *Kunstwissenschaft,* which may be rendered as "the sciences of art," for it comprises historical, anthropological, and other empirical studies of art as a cultural product. One of the impulses to the development of this field was a pervasive dissatisfaction with beauty, either because it is too limited, if interpreted on the classical model, and cannot therefore encompass, for example, primitive art, or too vague, if it is not. Art, by contrast, is a concrete, institutional phenomenon that is tractable to science. Thus *Kunstwissenschaft,* which is at present one of the most thriving and fruitful branches of aesthetics, defines itself by opposition to the concept of beauty.

The distinction between the meaning of beauty when it is synonymous with aesthetic value generally and when it stands for one class or kind of such value has been commonly remarked in recent aesthetics. In the former sense, it is often used to signalize the characteristic excellence of a work of art or an aesthetic object. Thus *beautiful* does not denote a property such as symmetry but also it is more than just a "term of approbation." It makes a claim on behalf of the object, which must be supported by appealing to the relevant value criteria. These criteria need not, however, be the same for two different artistic media or even for two works in the same medium. They are, perhaps indefinitely, plural; they are of different weight in different cases, and no one of them can be said to be a necessary condition for the use of *beautiful.* Their relevance is determined by the unique character of each work. In its second meaning, *beauty* generally connotes a relatively high degree of value, in contrast to, for example, the *pretty,* a fairly orthodox style or genre, pleasure unmixed with pain and the absence of bizarre or discordant elements. But this is just why so much of recent aesthetics and ordinary discourse finds the word awkward or even irrelevant for evaluation. It will do for Wolfgang Amadeus Mozart, but not the later Ludwig van Beethoven, for Raphael, but not Francisco de Goya. In the

Philebus, Socrates had, for his own purposes, narrowed the range of beauty severely, but it was just this narrowness that made it impossible for later thought to preserve *beauty* as the sole, or perhaps even the major, concept of aesthetic value.

See also Addison, Joseph; Aesthetic Experience; Aesthetic Judgment; Aesthetic Qualities; Aesthetics, History of; Aesthetics, Problems of; Aristotle; Art, Value in; Burke, Edmund; Fechner, Gustav Theodor; Feminist Aesthetics and Criticism; Hume, David; Hutcheson, Francis; Kant, Immanuel; Locke, John; Mendelssohn, Moses; Plato; Plotinus; Properties; Tragedy; Ugliness; Winckelmann, Johann Joachim.

Bibliography

Bosanquet, Bernard. *A History of Aesthetic.* New York: Meridian, 1957. Fairly difficult but thoughtful and acute; sees the history of aesthetics as a movement from the concept of "unity in variety" to "significance, expressiveness."

Bullough, Edward. "The Modern Conception of Aesthetics." In *Aesthetics,* edited by E. Wilkinson. Stanford, CA: Stanford University Press, 1957. A vigorous polemic, by an able psychologist-philosopher, against the "sterile" effort to find a definition of "beauty"; the "modern conception" is the "study of aesthetic consciousness."

Carritt, E. F. *The Theory of Beauty.* London, 1931. A readable account, by a leading Crocean, of many of the major historical theories.

Kainz, Friedrich. *Vorlesungen über Ästhetik.* Vienna, 1948. Translated by Herbert Schueller as *Aesthetics the Science.* Detroit: Wayne State University Press, 1962. A valuable conspectus of the major tendencies in more recent Continental aesthetics; examines the relation between beauty and the aesthetic attitude.

Morpurgo-Tagliabue, Guido. *L'esthétique contemporaine.* Milan, 1960. An extremely comprehensive and well-documented study of aesthetics since 1800, weighted heavily on the twentieth century. Stresses the distinction between the theories of art and beauty.

Osborne, H. *Theory of Beauty: An Introduction to Ethics*. New York: Philosophical Library, 1953. Lucid and informed criticism of some of the major historical theories.

Stolnitz, Jerome. "'Beauty': Some Stages in the History of an Idea." *Journal of the History of Ideas* 22 (2) (April–June 1961): 185–204. An analysis of the history of the concept in eighteenth-century British aesthetics. An extensive bibliography includes works by authors referred to above.

OTHER RECOMMENDED WORKS

Brand, Peg Zeglin, ed. *Beauty Matters*. Bloomington: Indiana University Press, 2000.

Danto, Arthur C. *The Abuse of Beauty: Aesthetics and the Concept of Art* (Paul Carus Lectures). Chicago and LaSalle: Open Court, 2002.

Devereaux, Mary. "Beauty and Evil: The Case of Leni Riefenstahl's 'Triumph of the Will.'" In *Aesthetics and Ethics: Essays at the Intersection*, edited by Jerrold Levinson. Cambridge, U.K.: Cambridge University Press, 1998, pp. 227–256.

Donoghue, Denis. *Speaking of Beauty*. New Haven, CT: Yale University Press, 2004.

Eco, Umberto, et al. *History of Beauty*. Rizzoli International, 2004.

Guyer, Paul. *Values of Beauty: Historical Essays in Aesthetics*. Cambridge, U.K.: Cambridge University Press, 2005.

Hickey, Dave. *The Invisible Dragon: Four Essays on Beauty*. Los Angeles: Art Issues Press, 1993.

"Beauty" (several entries). In *Encyclopedia of Aesthetics*, edited by Michael Kelly. Oxford: Oxford University Press, 1998, pp. 237–251.

Kieran, Matthew. *Revealing Art*. New York: Routledge, 2004.

Lorand, Ruth. *Aesthetic Theory*, Vol. 5: *A Philosophy of Beauty and Art*. New York: Routledge 2000.

Mattick, Paul. "Beautiful and Sublime: 'Gender Totemism" in the Constitution of Art." In *Feminism and Tradition in Aesthetics*, edited by Peggy Zeglin Brand and Carolyn Korsmeyer. University Park: Pennsylvania State University Press, 1995, pp. 27–48.

Mothersill, Mary. *Beauty Restored*. Oxford: Clarendon Press, 1984.

Prettejohn, Elizabeth. *Beauty and Art*. New York: Oxford University Press, 2005.

Sartwell, Crispin. *Six Names of Beauty*. New York: Routledge 2004.

Scarry, Elaine. *On Beauty and Being Just*. Princeton, NJ: Princeton University Press, 1999.

Zangwell, Nick. "Beauty." In *The Oxford Handbook of Aesthetics*, edited by Jerrold Levinson. Oxford: Oxford University Press, 2003, pp. 323–343.

Jerome Stolnitz (1967)
Bibliography updated by Mary Devereaux (2005)

BEAUVOIR, SIMONE DE
(1908–1986)

Simone de Beauvoir, French existentialist feminist, was born in Paris in 1908 and died in 1986, after a prolific career as a philosopher, essayist, novelist, and political activist. Her writings were, by her own accounts, heavily influenced by the philosophy of Jean-Paul Sartre, her intellectual companion for half a century—a fact that led some critics to dismiss her as philosophically unoriginal. Even de Beauvoir, in a 1979 interview, said that she did not consider herself to be a philosopher. In her view, however, "a philosopher is someone like Spinoza, Hegel, or like Sartre, someone who builds a grand system" (quoted in Simons, 1986, p. 168), a definition that would exclude most contemporary professional philosophers. Furthermore, as several recent commentators have argued, de Beauvoir seems to have underestimated her

influence on philosophy in general and on Sartre in particular. While she incorporated Sartre's ideas, such as his existentialist conception of freedom, in her ethical and political writings, her critiques of Sartre's work in progress also helped shape his philosophy, which she then extended and transformed in significant ways.

In *The Ethics of Ambiguity* (1948), de Beauvoir attempted to develop an existentialist ethics out of the ontological categories in Sartre's *Being and Nothingness*. In Sartre's view, there is no God and therefore no God-given human nature. Nor is human nature determined by biological, psychological, economic, cultural, or other factors. People are "condemned to be free," and in the course of existing and making choices, they construct their own natures (which are continually revisable). Although human consciousness is being-for-itself (the being of free and transcendent subjects), it vainly tries to turn itself into being-in-itself (the being of objects, things trapped in their immanence). De Beauvoir called this doomed attempt to synthesize the for-itself and the in-itself the "ambiguity" of the human condition, and she argued that ethics is both possible and required because of this inability of human beings to "coincide with" themselves. She attempted to ground ethics in individual freedom, asserting, "To will oneself free is also to will others free" (1948, p. 73), but her defense of this claim appears to slip Kantian and Hegelian presuppositions about human nature into a philosophy that denies that there is such a thing as human nature.

In *The Ethics of Ambiguity*, de Beauvoir moved beyond Sartrean existentialism in acknowledging certain constraints on freedom, including political oppression and early socialization, that Sartre did not recognize until much later. In her memoirs (1962), de Beauvoir recalled conversations she had with Sartre in 1940 about his account of freedom as an active transcendence of one's situation. She maintained that not every situation offered the same scope for freedom: "What sort of transcendence could a woman shut up in a harem achieve?" Sartre had insisted that even such a limiting situation could be lived in a variety of ways, but de Beauvoir was not persuaded. To defend her view, though, she would "have had to abandon the plane of individual, and therefore idealistic, morality," from which Sartre and de Beauvoir developed their philosophies (1962, p. 346).

In *The Second Sex* (1953) de Beauvoir continued to move away from a purely metaphysical view of freedom in developing an account of how the oppression of women limits their freedom. In arguing, "One is not born, but rather becomes, a woman," de Beauvoir applied the existentialist tenet that "existence precedes essence" to the situation of women, but she was also influenced by Marxist accounts of the material constraints on our freedom to create ourselves. In addition, she described how the socialization of girls and the cultural representations of women perpetuate the view of woman as other, thereby limiting women's potential for transcendence.

Critics of de Beauvoir's feminism have pointed out tensions between her existentialist premises and her account of the relation between embodiment and oppression. Although, according to existentialism, anatomy is not destiny (nor is anything else), de Beauvoir's discussion of female sexuality at times suggests that women's reproductive capacities are less conducive than men's to achieving transcendence. De Beauvoir has also been criticized for advocating in 1949 (1953) that women assume men's place in society, although in interviews in the 1970s and 1980s she urged a transformation of both men's and women's roles.

Even de Beauvoir's critics acknowledge her enormous impact on contemporary feminism. Her analysis of what has become known as the sex/gender distinction set the stage for all subsequent discussions. In drawing on philosophy, psychology, sociology, biology, history, and literature in *The Second Sex* and other essays, she anticipated the interdisciplinary field of women's studies. Her concern with autobiography, with self-revelation as "illuminating the lives of others" (1962, p. 8), prefigured the preoccupation of feminism with the personal as political. She also drew on a philosophical tradition as old as Socrates; her relentless scrutiny of herself and others exemplified, to an extent unmatched by any other twentieth-century philosopher, the maxim that "the unexamined life is not worth living."

In her fiction as well as in her essays and memoirs, de Beauvoir discussed numerous philosophical themes—for example, freedom, choice, responsibility, and the other—and she also explored the political issues and conflicts of the day, so much so that she has been described as "witness to a century." But she was more than a mere chronicler of events; she was a powerful social critic and an internationally known "public intellectual," whose influence will continue to be felt for a long time.

See also Existentialism; Sartre, Jean-Paul.

Bibliography

Francis, Claude, and Fernande Gontier. *Les écrits de Simone de Beauvoir*. Paris, 1979. A comprehensive bibliography of the publications of de Beauvoir from 1943 to 1977.

WORKS BY DE BEAUVOIR

Essays

Pyrrhus et Cinéas. Paris: Gallimard, 1944.

The Ethics of Ambiguity. Translated by Bernard Frechtman. New York: Philosophical Library, 1948. Originally published as *Pour une morale de l'ambiguïté*. Paris: Gallimard, 1947.

The Second Sex. Translated by Howard M. Parshley. New York: Knopf, 1953. Originally published as *Le deuxième sexe*. Paris: Gallimard, 1949.

The Coming of Age. Translated by Patrick O'Brien. New York: Putnam, 1972. Originally published as *La vieillesse*. Paris: Gallimard, 1970.

Fiction

She Came to Stay. New York: Norton, 1954. Translated by Yvonne Moyse and Roger Senhouse. Originally published as *L'invitée*. Paris: Gallimard, 1943.

The Mandarins. Translated by Leonard M. Friedman. New York: Norton, 1956. Originally published as *Les Mandarins*. Paris: Gallimard, 1954.

Memoirs

Memoirs of a Dutiful Daughter. Translated by James Kirkup. New York: Harper and Row, 1959. Originally published as *Mémoires d'une jeune fille rangée*. Paris: Gallimard, 1958.

The Prime of Life. Translated by Peter Green. New York; Paragon, 1962. Originally published as *La force de l'âge*. Paris: Gallimard, 1960.

Force of Circumstance. Translated by Richard Howard. New York: Putnam, 1965. Originally published as *La force des choses*. Paris: Gallimard, 1963.

All Said and Done. Translated by Patrick O'Brien. New York: Putnam, 1974. Originally published as *Tout compte fait*. Paris: Gallimard, 1972.

WORKS ON DE BEAUVOIR

Bauer, Nancy. *Simone de Beauvoir, Philosophy, and Feminism*. New York: Columbia University Press, 2001.

Brosman, Catharine Savage. *Simone de Beauvoir Revisited*. Boston: Twayne Publishers, 1991.

Card, Claudia. *The Cambridge Companion to Simone de Beauvoir*. Cambridge, U.K.: Cambridge University Press, 2003.

Dietz, Mary G. "Introduction: Debating Simone de Beauvoir." *Signs* 18 (1992): 74–88.

Gatens, Moira. "Woman as Other." In her *Feminism and Philosophy*. Cambridge, U.K.: Polity, 1991.

Le Doeuff, Michèle. "Simone de Beauvoir and Existentialism." Translated by Colin Gordon. *Feminist Studies* 6 (1980): 277–289.

Mackenzie, Catriona. "Simone de Beauvoir: Philosophy and/or the Female Body." In *Feminist Challenges: Social and Political Theory*, edited by Carole Pateman and Elizabeth Gross. Boston: Northeastern University Press, 1986.

McCall, Dorothy Kaufmann. "Simone de Beauvoir, *The Second Sex*, and Jean-Paul Sartre." *Signs* 5 (1979): 209–223.

Simons, Margaret A. "Beauvoir and Sartre: The Philosophical Relationship." *Yale French Studies* 72 (1986): 165–179.

Tong, Rosemarie. "Existentialist Feminism." In her *Feminist Thought*. Boulder, CO: Westview Press, 1989.

Whitmarsh, Anne. *Simone de Beauvoir and the Limits of Commitment*. Cambridge, U.K.: Cambridge University Press, 1981.

Susan J. Brison (1996, 2005)

BECCARIA, CESARE BONESANA
(1738–1794)

Cesare Bonesana Beccaria, the Italian criminologist and economist, was born in Milan of aristocratic parents. His formal education began at the Jesuit college in Parma and ended with his graduation from the University of Pavia in 1758. After graduation Beccaria came under the intellectual influence of two brothers, Pietro and Alessandro Verri, who had gathered around themselves the young Milanese intelligentsia to form a society known as the "academy of fists," committed to promoting reforms in political, economic, and administrative affairs.

Beccaria was prompted by Pietro Verri to read the then prominent philosophies of the Baron de Montesquieu, Claude-Adrian Helvétius, Denis Diderot, David Hume, and the Comte de Buffon. At the suggestion of his friends, Beccaria wrote and published his first treatise, *Del disordine e de' rimedi delle monete nello Stato di Milano nell'anno 1762* (Lucca, 1762). It was also through the encouragement of the Verri brothers that Beccaria composed his most important work, *Dei delitti e delle pene* (translated by H. Paolucci as *On Crimes and Punishments,* New York, 1963). Through Alessandro Verri, who was an official of the prison in Milan, Beccaria visited that institution and saw the conditions that furnished information and moral stimulus for his writing. Pietro, who had already begun writing a history of torture, in many conversations on the errors of criminal law and administration provided Beccaria with new arguments and insights for the treatise. In the end, the work was almost a collaboration by the three men, for Beccaria until that time had been relatively uninformed about crime and punishment. Begun in March 1763 and completed in January 1764, the book was published anonymously at Livorno out of fear of reprisals because of its devastating attack on the legal and judicial system then in operation. But anonymity was soon dropped when it became clear that the Milanese authorities were receptive and when the essay drew the attention and respect of the Parisian intelligentsia.

Beccaria held a chair in political economy in the Palatine School of Milan from 1768 to 1770, and his lec-

tures during this period were published posthumously in 1804 under the title *Elementi di economia pubblica.* His economic ideas on the division of labor and the determination of wages have been compared to those of Adam Smith (who wrote the *Wealth of Nations* seven years after publication of Beccaria's economic views). In economics Beccaria espoused a form of mercantilism based on some of the ideas of the physiocrats, expressed the belief that agriculture was the most productive enterprise, advocated commercial freedom within a nation and the abolition of guilds, and displayed a Malthusian concern with the relation of population growth to the means of subsistence. He also held a series of minor public offices through which he aided his friends in securing reforms in taxation, currency, and the corn trade.

On Crimes and Punishments was a protest against the use of torture to obtain confessions, secret accusations, the arbitrary discretionary power of judges, the inconsistency and inequality of sentencing, the influence of power and status in obtaining leniency, the lack of distinction in treatment of the accused and the convicted, and the use of capital punishment for serious and even minor offenses.

The concepts that Beccaria employed—rationalism, the social contract, utility, and hedonism—were current among the intellectuals of his time. The application of these ideas to crime and punishment, and the style of writing, were his own. Building upon Rousseau's social-contract philosophy, he argued that each person willingly sacrifices to the political community only so much of his liberty as "suffices to induce others to defend it." Laws are only the necessary conditions of this contract, and punishments under the law should have no other purpose than to defend the sum of these sacrificed shares of liberty "against private usurpations by individuals." Punishments for any other reason are unnecessary and unjust.

Beccaria declared that the law should be clear in defining crimes and that judges should not interpret the law but simply ascertain whether a person has or has not violated the law. He also held that punishment should be adjusted in severity to the seriousness of the crime. The primary purpose of punishment, Beccaria argued, is to ensure the existence of society, and the seriousness of the crime, therefore, varies according to the degree to which the transgressor's act endangers that existence. Treason and other acts against the state are most harmful, followed by injuries to the security of person and property and finally, by acts which are disruptive of public harmony and peace, such as rioting or inciting to disorder.

To ensure the continuance of society, punishment should aim at deterrence, that is, at preventing offenders from doing additional harm and others from committing crimes. To be effective as a deterrent to crime, punishment should be swift and certain; it is the certainty rather than the severity of punishment that deters. Life imprisonment is sufficient to deter: The death penalty is not necessary, nor is it legitimate, for individuals did not under the social contract relinquish the right to their lives. Corporal punishment is bad, and torture as part of a criminal investigation makes the suffering of pain rather than evidence the test of truth. Crimes against property should be punished by fines or, when fines cannot be paid, by imprisonment.

Beccaria's classic conclusion—the principles of which were adopted almost in their entirety by the revolutionary National Assembly of France in 1789 as Article VIII of the "Declaration of the Rights of Man and of the Citizen"—read in part as follows: "In order for punishment not to be, in every instance, an act of violence of one or of many against a private citizen, it must be essentially public, prompt, necessary, the least possible in the given circumstances, proportionate to the crimes, dictated by the laws."

Beccaria's essay became famous almost overnight. It was translated into French in 1766 by the Abbé Morellet, passed through six editions within eighteen months, one of which was embellished by a laudatory comment by Voltaire, and was thereafter translated into every important language. The Church of Rome placed the treatise on the Index in 1766, but the Austrian government, which controlled Milan, defended and honored Beccaria. Maria Theresa of Austria, Leopold II, grand duke of Tuscany, and Catherine the Great of Russia announced their intentions to be guided by Beccaria's principle in the reformation of their laws. The essay both paved the way for, and was the guiding force in, the major penal reforms that took place for two centuries afterward.

See also Buffon, Georges-Louis Leclerc, Comte de; Diderot, Denis; Hedonism; Helvétius, Claude-Adrien; Hume, David; Montesquieu, Baron de; Philosophy of Law, History of; Rationalism; Smith, Adam; Social Contract; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY BECCARIA

Opere. Edited by Pasquale Villari. Florence, 1854. A more recent edition was edited by S. Romagnoli, 2 vols. (Florence, 1958), with a bibliography, Vol. 2, pp. 917–918.

Scritti e lettere inediti raccolti ed illustrati da Eugenio Landry. Edited by Eugenio Landry. Milan, 1910.

WORKS ON BECCARIA

Cantù, C. *Beccaria e il diretto penale.* Florence: G. Barbèra, 1862.

Monachesi, E. "Cesare Beccaria." In *Pioneers in Criminology,* edited by H. Mannheim. London: Stevens, 1960. Pp. 36–50.

Paolucci, Henry. "Introduction" to his translation of Cesare Beccaria, *On Crimes and Punishments.* New York, 1963. Pp. ix–xxiii.

Phillipson, Coleman. *Three Criminal Law Reformers: Beccaria, Bentham, Romilly.* London: Dent, 1923.

Schumpeter, J. A. *History of Economic Analysis.* New York: Oxford University Press, 1954.

Marvin E. Wolfgang (1967)

BECK, JAKOB SIGISMUND

(1761–1840)

Jakob Sigismund Beck, the German Kantian philosopher, was born in Marienburg. He studied mathematics and philosophy in Königsberg with P. Krause and Immanuel Kant, completing his studies in 1783. In 1791 he became a teacher at the gymnasium in Halle and, in 1796, extraordinary professor of philosophy at Halle University. He was called to Rostock as professor of metaphysics in 1799 and remained there until his death.

Purporting to defend the "true" Kantian position against "dogmatic" misinterpretations, Beck called attention to problems concerning the role of the thing-in-itself in Kant's theory of perception. Beck rejected any positive role for the thing-in-itself and argued that the object affecting our senses must be phenomenal. Kant's theory of affection is to be understood not in the transcendent sense, as the working of an unknowable thing-in-itself on an unobservable "I"-in-itself, but only in the empirical sense: A phenomenal body in phenomenal space affects the "I" of inner sense.

But this "I" and this body, according to Beck, are themselves the products of an original activity of the understanding. The synthetic activity of "representing" (*vorstellen*) is presupposed by our viewing sense data as given *by* something objectively outside ourselves. Beck therefore objected to Kant's definition of sensibility as an immediate relation to an affecting object. The intuitions of sense say nothing about their own objectivity or source. Not until they are subjected to the categories of the understanding do they become objective, for only then can we invoke the notion of external objects and speak of intuitions as given to our senses by such objects. The order of exposition of the *Critique of Pure Reason* is therefore misleading. One ought not to begin with sensibility, but with the synthetic unity or "original activity" (*ursprüngliche Beilegung*) of the understanding, the unique a priori act of combination (*Zusammensetzung*).

In philosophy of religion, Beck held that God is a symbol created by man, a symbol of man's ethical conscience. Piety consists simply in obedience to the commands of conscience.

In letters to Beck (1792) Kant complimented him for investigating "what is just the hardest thing in the *Critique,*" approved Beck's reorganization of the Critical Philosophy, and said that he himself planned to write a work on metaphysics that would utilize the order of exposition that Beck had suggested. Kant's *Opus Postumum* shows the extent of Beck's influence, particularly in Kant's manuscript on the progress of metaphysics since Gottfried Wilhelm Leibniz and Christian Wolff.

Some of Kant's followers classed Beck with Johann Gottlieb Fichte and accused Beck of making the understanding the creator of objects. Beck did write: "Reality is itself the original act of representing, from which the concept of objects subsequently derives." But although he spoke of the original act as object-generating, he told Kant that he did not mean that the understanding *creates* objects. Beck granted the existence and importance of the given in knowledge while he attempted to bridge the dualism of sense and intellect and to insist that neither the given nor the notion of "things" could be taken as epistemologically primary.

See also Kant, Immanuel.

Bibliography

WORKS BY BECK

Erlaüternder Auszug aus den kritischen Schriften des Herrn Prof. Kant, auf Anrathen desselben. Riga, 1793–1796. Volume III of this work, *Einzig möglicher Standpunkt aus welchem die kritische Philosophie beurteilt werden muss* (Only possible standpoint from which the critical philosophy must be judged), contains Beck's most important ideas.

Grundriss der kritischen Philosophie. Halle, 1796.

Kommentar über Kants Metaphysik der Sitten. Halle, 1798.

Lehrbuch der Logik. Rostock, 1820.

Lehrbuch des Naturrechts. Jena, Germany, 1820.

WORKS ON BECK

Dilthey, Wilhelm. "Die Rostocker Kanthandschriften." *Archiv für Geschichte der Philosophie* II (1889): 592–650. Discusses Beck's place among Kant's disciples and critics.

Durante, G. *Gli epigoni di Kant.* Florence: Sansoni, 1943.

Meyer, Thomas L. "Das Problems Eines Höchsten Grundsatzes der Philosophie bei Jacob Sigismund Beck." *Analogia Filosofica* 5 (2) (1991): 205–208.

Potschel, W. *J. S. Beck und Kant.* Breslau, 1910.

Vleeschauwer, H. J. de. *L'evolution de la pensée kantienne.* Paris, 1939. Translated by A. R. C. Duncan as *The Development of Kantian Thought.* London: Nelson, 1962. Discusses Beck's influence on Kant's last reworking of the *Critique of Pure Reason.*

Arnulf Zweig (1967)
Bibliography updated by Tamra Frei (2005)

BEHAVIORISM

Traditional notions of the mind have tended to treat mental states as "private" and "subjective," not accessible to the public and objective methods of science. With the failure of an "introspectionist" psychology in the early twentieth century, the only recourse seemed to be either to deny that mental states had any role to play in any serious science, or to try to find a way to understand talk of mental states that was entirely objective. The first option is called the "eliminativist" strategy, and *Radical behaviorism* was a monumental effort to realize it. The eliminativist strategy proposed to explain all human and animal behavior in terms of physically specified stimuli, responses, and reinforcements. It is to be distinguished from the second, "reductionist" strategy, which attempts not to *eliminate* mental phenomena, but rather to *save* mental phenomena by identifying them with some or other existing physical phenomena. *Analytical behaviorism* was the specific reductionist view that mental phenomena could be identified in one way or another with dispositions to overt behavior. Both Radical behaviorism and Analytical behaviorism dominated Anglo-American philosophy, and especially psychology, from roughly 1920 through 1970.

Although the two views are similarly motivated, they are independent. As will be seen in section one, Radical behaviorism is a specific scientific hypothesis, to be assessed according to the usual scientific criteria of how well it predicts and explains its intended range of phenomena. Analytical behaviorism is essentially a semantic, or philosophical hypothesis, to be assessed according to how well it captures the mental notions it purports to analyze (sec. 2). A person could subscribe to one and reject the other: Strict radical behaviorists might be skeptical of semantic proposals of analytical behaviorists; and many analytical behaviorists might reject the scientific proposals of Radical behaviorism.

There is also a third view, *methodological behaviorism*, according to which the only *evidence* for any mental phenomena must be behavioral. As a claim about evidence, this is actually independent of both the other views, although it often accompanied them. Indeed, one of the lasting positive contributions of the entire behaviorist movement was a much higher standard of evidence than had been observed previously, discouraging the kind of reliance on empathic intuitions that was characteristic, for example, of clinical psychotherapeutic claims. Unlike Radical behaviorism and Analytical behaviorism, methodological behaviorism survives in some quarters to this day, although some problems for methodological behaviorism are raised at the conclusion of section three.

RADICAL BEHAVIORISM

THE LAW OF EFFECT. Since this is a philosophical encyclopedia, the treatment of Radical behaviorism will perforce be brief (for a more thorough discussion in which references can be found to the experiments cited here see Gleitman et al. 2004, Gallistel 1990, and Rey 1997). However, the treatment of Radical behaviorism is not philosophically irrelevant since a substantial number of twentieth-century philosophical views often relied on it, most famously those of the American philosopher W. V. Quine.

Radical behaviorism emerged from the work of the Edward Thorndike (1874–1949), John Watson (1878–1958), and Ivan Pavlov (1849–1936), receiving its most energetic development in the work of B. F. Skinner (1904–1990) and attaining considerable precision in the work of Clark Hull (1884–1952). It has its source in traditional empiricist theories of the mind, according to which the mind at birth is a tabula rasa, or blank tablet on which experience forms sensory impressions. Ideas are derived from experience and are welded together to form complex ideas by a process of association, which closely tracks the presentation of those experiences in reality. Thus, certain sights, sounds, and tactile sensations become associated in experience to form the idea of a material object, and certain associations of "contiguity, succession and constant conjunction" form the idea of causation (Hume 1734).

This traditional suggestion, though regarded by radical behaviorists as right in spirit, suffered from a major defect—namely, a reliance upon peculiar private entities, ideas, and impressions, which did not seem to radical behaviorists to be proper objects of scientific inquiry. To remedy this situation, they proposed studying not associations among ideas but among physically characterizable

stimuli to sense organs and responses of the motor system. The specific law that linked stimuli and responses was Thorndike's Law of Effect, which for purposes of this entry may be stated thus:

> The Law of Effect: *The probability of a response R following a stimulus S is increased/decreased if pairs ⟨R,S⟩ have been followed by positive/negative reinforcements, F, in certain patterns (e.g., intermittently) in the past.*

For example, should a particular movement like pressing a paw on a lever (=R) when a light is on (=S) be followed intermittently by the presentation to a hungry animal of a food pellet (=F), then the probability of the animal pressing its paw on the lever when the light is on in the future will be increased. Such are rewards. Negative reinforcements are either the absence of positive reinforcements, or actual punishments, which also reinforce, but in the opposite direction: the probability of the R given S is reduced if pairs of S and R have been followed by punishment in the past. Radical behaviorism is essentially the bold hypothesis that all intelligent human and animal behavior can be explained by the Law of Effect.

As Skinner frequently stressed, the Law of Effect is nearly the biological principle of natural selection, extended now beyond the persistence of traits that are genetically inherited to the persistence of acquired behaviors in individual animals. Just as from a random generation of genetic mutations certain ones are selected by virtue of meeting an environmental test of "survival of the fittest," so from an essentially random generation of responses in an animal certain ones are selected by virtue of being reinforced when they occur after certain stimuli. The responses that are selected in this way Skinner (1938) called "operants," since they involved ways that an animal "operated" on an environment that secured reinforcement. This process of "operant conditioning" was Skinner's distinctive contribution over "classical conditioning," where the response was *elicited* (e.g., salivation by hunger in Pavlov's dogs), rather than being spontaneously *emitted*.

How could such a simple law as the Law of Effect possibly stand a chance of explaining the full range of animal behavior? The central idea was an extension of the associationist strategy of building complex ideas from simpler ones, only now it was a matter of building not complex ideas, but complex responses. These could be built up out of simpler responses by "response chaining," whereby stimuli associated with a reward themselves become ("secondary") reinforcers, and so available for the conditioning of further responses. Thus, a pigeon conditioned to peck a lever on hearing a bell could now be conditioned by the sound of the bell itself to produce further responses given further stimuli, say, doing a little dance on seeing a red light, which is then followed by the bell, which is then followed by food if the pigeon pecks again at the lever. Discrimination of complex stimuli would similarly be built from discrimination of simpler stimuli, through either a chain of discriminations of simpler stimuli, or by "stimulus generalization," whereby novel stimuli are treated as "of the same kind" as earlier ones.

The Law of Effect is likely true of some animal behavior. Skinner achieved remarkable successes using it to train animals to engage in all manner of curious behavior: for example, rats to run mazes, pigeons to play Ping-Pong, and pigs to push shopping carts around supermarkets. And the Law of Effect seems to play a role in explaining a variety of persistent behavioral patterns, such as gambling and drug addiction, as well as in extinguishing them, as in "behavior modification therapy." For the purposes of this entry, the issue is not whether such applications occur or are a good idea, but whether they offer a theoretically adequate paradigm for understanding the full range of intelligent animal behavior.

INADEQUACIES OF THE LAW OF EFFECT. Problems with the Law of Effect emerge in the first instance from the radical behaviorist experiments themselves. Contrary to popular belief, it is not only human behavior that resists radical behavioristic explanation; the theory does not even really work for the rats. The main problem is that the probability of a response can be increased in ways other than by the Law of Effect. There are at least four classes of phenomena that the law has trouble explaining: latent learning, passive learning, spontaneous alteration, and improvisation.

Latent learning occurs when an animal learns without reinforcement. Rats that were well sated with, for example, food and water were allowed to run around in a maze for ten days without any reward, sometimes being placed in the maze at arbitrarily different points. Subsequently, when they were hungry again they were introduced into the maze and were able to find the food much faster than rats not previously exposed. Similarly, Harry Harlow showed that monkeys presented with a complex hinge, requiring the undoing of several pins and bolts to free it, learned to undo it with no special reward other than "the fun of it." Further, indigo buntings learn something about the position of the stars while still in the nest, despite not using this information for navigation (and so,

a fortiori, not for any reward) until they are much older. In all of these cases the probability of the animal producing the appropriate response was greater than that of animals that had not been previously exposed to such stimuli, but without any history of reinforcement. In a related vein, passive learning occurs when an animal learns without antecedently producing the requisite response. Thus, rats can learn a maze merely by being pulled through it in a transparent trolley car, not executing anything like the responses that will take them through the maze when they are tested later.

Not only can rats learn without reinforcement or response, they can sometimes respond in ways that defy their conditioning history. In "spontaneous alteration," an animal actually avoids emitting the response that has recently been reinforced. After having found food at a particular location, for example, hummingbirds will go somewhere else to find more food. Rats presented with a number of paths of equal length to a goal will vary their routes, although invariably in ways that advance their approach to the goal. The phenomenon is most dramatically displayed by rats in a "radial maze," consisting of eight pathways radiating out in all directions from a central location, with baits placed at the end of each arm. The Law of Effect should predict that the rats should return to an arm in which they have found food. What they do instead, however, is to *avoid* an arm they have already visited until they had—at random—visited all the others. That is (as we might put it mentalistically), they seem not to be matching responses to stimuli, but "keeping track" of "where the baits are," and, knowing they had consumed one, no longer "expected" it to be there. The Law of Effect seems not only inadequate to account for such cases; it actually seems to be disconfirmed by them

Animals also produce appropriate behaviors that have not even previously been produced, much less reinforced. Thus, rats trained to take a circuitous route to a goal box will immediately take a shortcut if it is made available. Indeed, animals apparently refuse to be tied to specific physical responses: Rats will swim a flooded maze after being taught to run it, and—moving to the wild, outside the confines of a structured maze—desert ants will forage in a winding path up to one hundred meters from their nests, and then, once they find food, will take a beeline home.

Passing beyond issues of navigation, it has been noted with regard to latent learning that monkeys presented with a novel, complex hinge, figure out how to undo several pins and bolts to free it. Köhler demonstrated even more remarkable improvisation in chimpanzees: They would use sticks as rakes to secure food that was outside a fence; they would then use these sticks as poles, which they would climb up in order to snatch food that was out of reach, grasping the food just as the stick toppled over. In all of these cases, the responses—that is, the sequence of muscular motions required to execute the acts—are by no means physically type-identical between prior and test trials. So the animals must have learned something other than merely to repeat certain physically typed responses.

Of course, these inadequacies with the Law of Effect become even more glaring in the human case. Picking up on an example of Skinner's (1957, p. 38), Daniel Dennett (1975) provides an apt and amusing discussion of the difficulties besetting a radical behaviorist attempting to explain why someone mugged in New York hands over his wallet: Why doesn't the person instead do any number of things that were more likely to have been previously reinforced with the stimulus "Your money or your life," such as giggling, or yawning? Of course, it is not impossible that there is a story of prior threat stimuli and responses of the requisite sort. But the burden is squarely on the radical behaviorist to supply it.

Radical behaviorists, of course, did not take challenges to the inadequacies of the Law of Effect lying down. They often made ingenious replies to them involving elaborate emendations of the theory—for example, by Clark Hull (1943). But these emendations were then subject to further tests, showing animals to be more ingenious than the Law of Effect allowed. A consensus began to emerge that what animals learn is not any mere sequence of responses to stimuli, but rather to the development of what Edward Tolman called an "internal map" (1948). Such talk of "insight" and "maps" of course, begins to imperil any eliminativist ambitions of Radical behaviorism: such a map would be an inner representation, involving an internal mental state.

STRUCTURED RESPONSES AND LANGUAGE. An important problem in principle for Radical behaviorism was raised by Karl Lashley (1951): Serial responses like those involved in tying shoes or riding a bicycle seemed to be *structured* in a way that it did not seem possible to explain by local response-chaining alone. A domain of behavior that exhibits particularly striking structure is *language*. Skinner (1957) tried to sketch an account of linguistic behavior, but it was soon subject to a devastating review by the then young linguist Noam Chomsky (1964). Among other things, he pointed out:

(i) along lines indicated by Lashley, language is *structured* in units that cannot be captured by response chaining. For example, a sentence of the form "Either … or … ," or "If … then … " can involve waiting indefinitely for novel items to be inserted in the blanks;

(ii) language is *creative*: most of the sentences people encounter and produce are constantly *novel*—it is why people bother to converse, read, and write—all contrary to the Law of Effect's commitment to a prior history with the stimuli and responses;

(iii) language is *productive*: in grasping a grammar, even small children know how to produce a potential infinity of novel, structured sentences, as in "This is the house that Jack built," "This is the rat that lived in the house … ," "This is the cat that chased the rat … ," without any history of conditioning each component in this way; indeed:

(iv) the complex set of rules that constitute the grammar is acquired effortlessly by practically all human children by the age of three, without (and sometimes despite) any efforts at explicit instruction.

(For more detailed discussion see Chomsky 1972 and Pinker 1994).

Although Chomsky's review was (to many minds) a definitive blow by itself, what really led to the end of Radical behaviorism was the spectacular positive research program that he and others (e.g., Fodor 1968, 1975) had begun to develop, what has come to be called the *cognitive revolution*, associated with computational-representational theories of mental processes.

Often recognizing the difficulty of avoiding mentalisms in the explanation of animal behavior, radical behaviorists sometimes allowed mentalisms to creep into their explanations, postulating "exploratory" and "curiosity" drives, or "drives to perceive" or "know." Of course, if the theory was to remain true to its goal to avoid reference to subjective mental states that it regarded as unscientific, it would be obliged to define these postulations in terms of overt behavior. It was in this way that Radical behaviorism invited Analytical behaviorism, to which we now turn.

ANALYTICAL BEHAVIORISM

Analytical behaviorism was motivated by two related philosophical trends of the twentieth century that persist into the twenty-first century: the well-known *verifiability theory of meaning* (or *verificationism*) and the less well-known doctrine that might be called *irreferentialism*. Because the latter serves as something of a background for the former, it will be considered first.

IRREFERENTIALISM. Irreferentialism is a novel suggestion that arose from Bertrand Russell's (1905) famous theory of definite descriptions, according to which expressions like "the present king of France" should not be construed as *referring* to any (in this case) nonexistent entity, but as rather shorthand for some logically complex expression, only some of the most basic parts of which manage to refer. Perhaps the most obvious deployment of such a strategy is the in the case of a sentence such as "The average American family has 2.5 children," which, of course, does not entail that there is some family somewhere in America that has a half of a child. A proper analysis of the grammar of the claim reveals that it is simply a way of expressing the ratio between American families and their children.

The view begins to be applied as a claim about mental expressions in the work of the later Wittgenstein (1953) and Gilbert Ryle (1949). They argued that philosophers too often think about the phenomena that people introspect in their "inner mental worlds" on the model of the objects in the familiar "outer" one. The temptation to this analogy arose, Wittgenstein and Ryle claimed, from an excessively referential conception of the functioning of the human mental vocabulary, treating words like "belief," "thought," and "sensation" as referring to "inner," "private" objects, in the way that words like "cat" and "rock" refer to outer, public ones. It is not that, like "Zeus," they do not *happen* to refer to anything; rather, like "the average American," they do not even *purport* to refer to anything. The view is perhaps best known from Ryle's attack on the idea of the "ghost in the machine": A mind is not some sort of *thing* that *could* be a ghost, or any other *thing*. Not surprisingly, this irreferentialism was often associated with an antipathy one finds in Wittgenstein (1953) and Ryle (1949) toward a psychology that suggests a "promise of hidden discoveries yet to be made of the hidden causes of our actions and reactions" (Ryle 1949, p. 325).

Irreferentialism is an essentially negative thesis about the analysis of mental terms. Understandably, many people might want to hear something more positive: if the analysis of mental terms does not involve the postulation of mental entities, what does it involve? For Wittgenstein and his followers, in particular, this question (like, in their view, most philosophical questions) was the wrong one to

ask: it exhibited a somehow inappropriate "craving for generality" about the nature of thought and language. That may in the end be so, but many wanted to see a greater effort made toward some systematic account than he and his followers provided. The use of mental language does not seem entirely capricious and chaotic, and, if it is not, then it is not unreasonable at least to ask what the principles might be that guide its use.

VERIFICATIONISM. According to verificationism, the meaning of a claim consists in the method by which it could be tested (see Ayer 1952 for a classic statement). For example, claims about something's being an acid might be defined in terms of its turning litmus paper red. Or claims about the existence of material objects might be analyzed as logical construction of claims about those sense experiences that people ordinarily take to confirm such claims (e.g., that people would have certain experiences of color, shape, and resistance to touch). Hypotheses such as those of the possibility of human lives being a dream, or other people having radically different mental lives were to be ruled out as "meaningless" if there really were in fact no evidence in principle that could make a difference to their truth or falsity.

There are myriad problems with verificationism: It is by no means obvious how to apply it to the claims of logic, mathematics, ethics, or aesthetics, or even to itself (what is the test for assessing where it is true?). Even in the supposed parade cases of natural science verificationism did not fare well: Scientists often do not know how to seriously test a hypothesis (as in contemporary string theory in physics) and often change their tests as their theories evolve (as new tests are devised for a disease). But the most serious problem is *confirmation holism*, or the fairly obvious observation that claims are not tested by experiment individually, but only as parts of whole theories (see Quine 1960). As will be seen, Analytical behaviorism offered a vivid case in point.

ANALYTICAL BEHAVIORIST PROPOSALS. Analytical behaviorism was largely motivated by verificationism and the observation that the vast majority of human mentalistic claims are tested by observing overt behavior (of course, this does not seem to be true in the case of first-person reports, which were always a problem for Analytical behaviorism, although these represented a small minority of claims). Moreover, it did seem that those ascriptions were by and large indifferent discoveries that might be made about the actual physical aetiology of mental life. If one were to open up the heads of familiar people and discover that they were empty or full of sawdust, one would not conclude that these familiar people did not have the mental states that seem to be constituted by the familiar behavior observed. Consequently, it seemed reasonable to suppose that mental claims should be understood as equivalent to various sorts of dispositional or conditional claims about how an agent would behave if she were in such and such circumstances. One particular model that impressed behaviorists was that of *dispositional* claims that arise elsewhere: "Salt is soluble" presumably means something like, "If salt is put into water in certain normal conditions, then it dissolves"; "Glass is fragile," something like, "If struck in normal circumstances, it breaks." Analogously, *wanting* something should be taken to mean something like "*trying to get it, if the occasion were to arise*."

Successfully spelling out the appropriate dispositions in the case of mental terms turned out, however, to be none too easy. Ryle was never precise, but he offered a strategy, exemplified by the following characterization of belief:

> To believe that the ice is dangerously thin is to be unhesitant in telling oneself and others that it is thin, in acquiescing in other people's assertions to that effect, in objecting to statements to the contrary, in drawing consequences from the original proposition, and so forth. (Ryle 1949, 134–135)

ACTION VS. "COLORLESS MOVEMENT." A formidable problem arises, however, as soon as one considers the "behavior" on which people normally rely, namely of distinguishing *actions* from *mere movement*: To take a famous example from Wittgenstein, it is the difference between raising one's arm and one's arm rising. The rising of an arm might occur as a result of, say, some machine moving the arm up and down; it is only the raising of an arm if it was the result of the person whose arm it is *intending* to raise it. Ryle may think that he is describing mere behavior in talking about someone being "unhesitant" and "acquiescing," "objecting" to certain "assertions," but a moment's reflection reveals these as only slightly covert mentalisms: *hesitation*, *acquiescence*, and the possible involvement of any of an indefinite variety of bodily movements (or none); all that is crucial is that whatever the agent does or does not do is a result of a certain psychological attitude.

This point was often most seriously missed by the radical behaviorists who, as has been noted already, often resorted to mentalisms to deal with the apparent counterexamples to their Law of Effect. Thus, Skinner wrote:

> The artist … is reinforced by the effects his works have upon … others … [his] verbal behavior … reach[ing] over centuries or to thousands of listeners or readers at the same time. The writer may not be reinforced often or immediately, but the reinforcement may be great. (Skinner 1957, pp. 206, 224)

But, as Chomsky (1964) noted, the term *reinforcement* here has degenerated to only a ritual function, being used as a cover term for "X wants Y," "X likes Y," "and X wishes that Y were the case."

INSUPERABLE PROBLEMS. For every thought experiment arguing for Analytical behaviorism there are compelling ones against it as well. Consider not only people with sawdust heads, but people who turn out to be robots cleverly controlled by radio waves produced by some ingenious scientists at MIT: Such creatures would seem to have no more of a mental life than do marionettes. Or consider a race of "Super-Spartans" who, as matter of training and principle, refuse to flinch or complain in response to even the most excruciating pain and are inarticulate about an enormous range of their psychological repertoire (Putnam 1975). Surely it is possible that these Super-Spartans have by and large the full range of psychological states of the more expressive and articulate.

All of these objections become more evident when one considers an underlying technical difficulty noticed by Roderick Chisholm (1957): Every effort to define most mental states by behavior seems to require citation of other mental states. Typically, any particular mental state causes a particular behavior only in conjunction with (an often large number of) other mental states. Beliefs, hopes, and expectations issue in behavior only in conjunction with (at least) desires; desires issue in behavior only in conjunction (at least) with beliefs and expectations. To take a proposed example from Tolman, suppose a person tried to define a rat's expectation that there is food at L in terms of the rat's moving toward L: This only if the rat *wants* food; and the rat's *wanting* food can be defined in terms of its moving toward L only if it *expects* there is food at L. Insofar as this is true, the prospects of a definition of a single informational or single directional state in terms of behavior seem dim. This problem is an instance of the aforementioned confirmation holism emphasized by Quine. Indeed, a philosophically influential (and disconcerting) way of understanding this and related difficulties with Analytical behaviorism is provided by Quine's (1960, ch. 2) "thesis of the indeterminacy of translation," according to which there is no fact of the matter about the content of mental states, a thesis that has influenced later philosophers such as Donald Davidson, David Lewis, and Dennett.

METHODOLOGICAL BEHAVIORISM

In the twenty-first century few, if any, philosophers or psychologists would be prepared to defend either Radical behaviorism or Analytical behaviorism. However, there is a weaker view that survives, called "methodological behaviorism," according to which the only permissible evidence for a psychological claim can be behavioral. It is not, like Radical behaviorism, an explanatory scientific hypothesis, but neither does it, like Analytical behaviorism, offer analyses of mental terms, although it is motivated by vaguely verificationist concerns like those that motivated Analytical behaviorism. Methodological behaviorism is perhaps best expressed by Wittgenstein's famous dictum, "An inner process stands in need of outward criteria," (1953, sec. 580) but without any of the analytical behaviorist commitment to defining an inner process in terms of some specific criteria. It has most recently been defended by Dennett (1993), who describes its motivation as not a "village" but an "urbane verificationism" that is merely trying to avoid "epiphenomenalism, zombies, conscious teddy bears, [and] self-conscious spiders" (1993, p. 461). Indeed he "unhesitatingly endorse[s] the claim that necessarily, if two organisms are behaviorally exactly alike, they are psychologically exactly alike" (1993, p. 922).

For all methodological behaviorism's urbanity, however, it is hard to find a convincing argument for it. Why shouldn't psychologists avail themselves of evidence that may go beyond ordinary overt behavior, as they indeed seem increasingly to do when they investigate the finer structure of the brain? Consider, for example, the nice case Dennett (1991, pp. 395–396) discusses of the familiar plight of the adult beer-drinker who wonders whether in coming to like it since childhood, it is his *experiences* or his *preferences* that have changed. It can seem obscure what further considerations should settle the matter, and it is not implausible to suppose that current behavioral discriminations or even introspection would not suffice. Dennett concludes that that there is in such a case "no fact of the matter" about "the way the beer tastes" to such a person.

But suppose it turns out that children have more taste buds than adults. One might have independent evidence that both children and adults have the same preferences for bitter titillation, but that consequently children reach a painful threshold sooner with the same quantity

of a bitter substance. It tastes differently because, arguably, more intense sensation is caused by their tongues and/or gustation subsystems. However, such cases would clearly transcend mere behavioral evidence.

So, in the end, even methodological behaviorism seems problematic given the increasingly rich conceptual and evidential resources of cognitive science, and especially of a computational/representational theory of thought. Behaviorism in all its forms seems a heroic theory that was ultimately defeated by the high standards of theory and experimentation that it encouraged. Eliminativist strategies survive in ambitions to replace mental talk by neurophysiological descriptions, and reductionist strategies abound, along either neurophysiological lines or computational ones. But the effort to replace mental talk with behavioral talk, or reduce it to it, can safely be said to have passed with the twentieth century, in which it first appeared.

See also Chomsky, Noam; Fodor, Jerry A.; Functionalism; Mind-Body Problem; Physicalism; Quine, Willard Van Orman; Reductionism in the Philosophy of Mind; Verifiability Principle.

Bibliography

Ayer, A. *Language, Truth and Logic*. New York: Dover, 1952. A classic statement of the "Verificationist Theory of Meaning," which provided the general semantic view underlying Analytical behaviorism.

Chisholm, R. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.

Chomsky, N. *Language and Mind*. New York: Harcourt Brace Jovanovich, 1972. Chomsky's earliest and most accessible statement of the relation of his theory of grammar to theories of the mind generally.

Chomsky, N. "Review of Skinner's *Verbal Behavior*." In *The Structure of Language: Readings in the Philosophy of Language*, edited by J. Fodor and J. Katz, 547–578. Englewood Cliffs, NJ: Prentice Hall, 1964. The devastating review of Skinner's effort to incorporate linguistic behavior into Radical behaviorism.

Dennett, D. *Brainstorms*. Cambridge, MA: MIT Press, 1975. Several essays ("Skinner Skinner," "Why the Law of Effect Won't Go Away") provide excellent discussions of the attractions and limitations of Radical behaviorism.

Dennett, D. *Consciousness Explained*. New York: Little Brown, 1993. A recent defense of methodological behaviorism (which he calls "urbane verificationism").

Fodor, J. *The Language of Thought*. New York: Crowell, 1975. The classic statement of the computational/representational theory of thought that, he argues, is presupposed by virtually all work in cognitive and perceptual psychology.

Fodor, J. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press, 1983. The classic statement of the now highly influential view of perception and language processing as consisting of computational "modules."

Fodor, J. *Psychological Explanation: An Introduction to the Philosophy of Psychology*. New York: Random House, 1968. An early, powerful philosophical critique of both Radical behaviorism and Analytical behaviorism.

Gallistel, C. *The Organization of Learning*. Cambridge, MA: MIT Press, 1990. An important non-behaviorist, computational theory of animal navigation, bringing together extensive studies of animals in the wild.

Gleitman, H., A. Fridlund, and D. Riesberg. *Psychology*. 6th ed. New York: Norton, 2004. An excellent, up-to-date text in psychology. Chapter four provides a detailed account of Radical behaviorism.

Hull, C. *Principles of Behavior: An Introduction to Behavior Theory*. New York: Appleton-Century-Crofts, 1943. A serious, concerted effort to state the laws of behavior in terms of "colorless movement" in ways designed to accommodate many of the experimental problems Radical behaviorism faced.

Hume, D. *A Treatise of Human Nature*, (1739), edited by L. A. Selby-Bigge. 2nd edition revised by P. H. Nidditch. Oxford: Clarendon Press, 1975. The classic statement of an associationist psychology, of which Radical behaviorism was an instance.

Lashley, K. S. "The Problem of Serial Order in Behavior." In *Cerebral Mechanisms in Behavior*, edited by L. A. Jeffress, 112–136. New York: Wiley, 1951. Widely regarded as raising a crucial problem for Radical behaviorism by calling attention to highly structured behaviors that could not be the result of mere response chaining.

Pinker, S. *The Language Instinct: How the Mind Creates Language*. New York: Harper and Row, 1994.

Putnam, H. "Brains And Behavior." In *Mind, Language And Reality: Philosophical Papers*. Vol. 2. Cambridge, U.K., and New York: Cambridge University Press, 1975, 325–341.

Quine, W. V. *Word and Object*. Cambridge, MA: MIT Press, 1960. Chapter two presents his influential discussion of the "indeterminacy of translation," which he draws as an interesting consequence of Radical behaviorism.

Rey, G. *Contemporary Philosophy of Mind: A Contentiously Classical Approach*. Oxford: Blackwell, 1997. Chapters four and five present detailed discussions of Radical behaviorism and Analytical behaviorism along lines of the present article. Chapters seven and ten present arguments against methodological behaviorism (called "superficialism" there).

Russell, B. "On Denoting" *Mind* 14 (1905): 479–493.

Ryle, G. *The Concept Of Mind*. London: Hutcheson, 1949.

Skinner, B. F. *The Behavior of Organisms: An Experimental Analysis*. New York, Appleton-Crofts, 1938. The official statement of Skinner's famous version of Radical behaviorism.

Skinner, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, 1957. Skinner's effort to provide a radical behaviorist framework for understanding language.

Tolman, E. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (1948): 189–208.

Georges Rey (2005)

BEING

Philosophy proceeds in part by the asking of large, imprecise, and overgeneral questions. In the attempt to answer them, the questions themselves come to be reformulated with greater clarity, and one large question often comes to be replaced by several smaller ones. The history of pre-Socratic philosophy is the best example of this process, and Being first appeared on the philosophical scene as part of it. To the question "What is Being?" the Parmenidean answer that there is Being and nothing else besides Being appears to have the merit of truth, even if it is tautological truth. What is, is; and what is not, is not. But what Parmenides' question in fact contains is a non-tautological demand for the characteristics of what is, to which the answer that Being is one, unchanging, and eternal is appropriate. Since the objects we perceive are many, changing, and transient, they do not belong to the realm of Being. Parmenides thus fathered in broad outline a doctrine of Being from which philosophers as diverse as Aristotle, Georg Wilhelm Friedrich Hegel, and John Dewey have tried to rescue us. This is the doctrine that Being is a name.

"BEING" AS A NAME

"Being" may be thought to name a property possessed by everything that is. Or it may be thought to name an object or a realm beyond, above, or behind the objects of the physical world; in this case, physical objects somehow exist by virtue of their relationship to "Being." Or again, "Being" may be the name of the genus to which everything that is belongs in virtue of the possession of the property of Being or of standing in relation to Being. The doctrine that "Being" is a name implies some kind of dualism, according to which the realm of Being is contrasted with that of the merely phenomenal. Variations on this doctrine are general enough to be put to a number of different uses in the attempt to solve quite different problems. Nevertheless, the basic doctrine is founded on a false assumption, for it obscures the facts that the verb "to be" has a number of different uses and that in its central and commonest use it does not ascribe a property, a relation, or class membership in any way. "Being" is normally a participle, not a noun. To break with normal usage without special justification is to be gratuitously liable to confusion. We can investigate the type of confusion generated by the acceptance of "Being" as a name, and also the type of clarification that came to be needed, by considering what Plato and Aristotle make of Being.

PLATO AND ARISTOTLE

Plato was anxious to mark the distinction between properties and objects that possess properties. He located the former in the realm of Being and the latter in the realm of the transient. One reason for this distinction was that Plato accepted the identification of Being with the unchanging (in this case, the unchanging meanings of predicate, the Forms). As a consequence, he was forced to deny that physical objects "are"—they belong to a stage intermediate between Being and Not-Being, that of becoming. This is not the only paradox in Plato's analysis of the subject: The Form of the Good, which exists at a higher level than that of the other Forms, cannot just "be," either; it must exist "beyond being."

Thus, we can see in Plato one of the characteristic results of treating Being as either a special kind of object or a special kind of attribute, namely, that all sorts of ordinary uses of the verb "to be" must be qualified or rewritten. The outcome of the attempt to make what is mystifying clear is to make what was clear mystifying. The author who first attacked this kind of mystification was, of course, Plato himself. In the *Sophist,* the problem of negative judgment is handled in such a way that it is no longer possible to make Parmenides' mistake of supposing that when one speaks of what is not, one is speaking of what does not exist. Moreover, it is scarcely proper to speak casually of confusion and mistake at this stage in the development of philosophy. The first steps toward producing a logical grammar of the verb "to be" perhaps necessarily involved assimilating the different senses and uses of the words, and of consequently becoming caught up in paradox and learning how to free oneself. When Aristotle, in Book I of the *Metaphysics,* clarified earlier errors, he was able to do so only because he had learned from the efforts and missteps of Parmenides and Plato.

Aristotle made three crucial points about the study of Being as Being. The first is that the special sciences may make use of the concept of Being and of other similar fundamental concepts, but these concepts are not the objects of their inquiries—only philosophy has such fundamental concepts as the proper object of its studies. The second point is that to inquire about Being as Being is to attempt to isolate the unifying strand of meaning in the multifarious senses in which the word "is" is used. The third point is that this inquiry can be carried on only as an inquiry into a whole range of closely related fundamental concepts, in which the different species of cause and the notions of unity and plurality are foremost.

Aristotle recognized that we use "is" to deny as well as to affirm, and to ascribe properties as well as to ascribe

existence; and in various passages he makes use of these distinctions to clarify conceptual points. He recognized, as did the Scholastics, that in ascribing properties to a subject we sometimes imply the existence of that subject and we sometimes do not. But in his willingness to recognize the diversity of uses of "is," Aristotle almost too easily accepted the view that we can speak of abstract entities as well as of physical objects without allowing the former "separate" existence. Aristotle said very little, in fact, about the common thread that binds together the various uses of "is."

SCHOLASTIC PHILOSOPHERS

The non-Aristotelian medieval writers who insisted on a single meaning for "is" unintentionally provided a reductio ad absurdum proof of the correctness of the Aristotelian approach. Both nominalists and realists, at least in their extreme and consistent versions, asserted that properties and objects exist in the same way: properties for the nominalists were merely collections of objects, and objects for the realists were merely properties of properties. For the nominalist Eric of Auxerre, "Being" was simply the collective name of all the individuals that exist taken together and was logically equivalent to "this and this and this …," while for the realist Odo of Tournai, individuals were accidents of properties that are substances, and the realm of Being was a realm only of properties.

Abelard to some extent reasserted the Aristotelian distinctions (and suggested some new ones of his own), but it was Thomas Aquinas who returned to the pure Aristotelian tradition. Thomas refuted once again the view that Being can be either a genus or a property.

In his Commentary on Aristotle's *Metaphysics,* Thomas diagnosed Parmenides' mistake and applied his conceptual insights to related problems, notably in his refutation of Anselm's Ontological Argument. But Thomas's position necessarily has a complexity lacking in some other writers who have been equally careful, for although he could not accept Anselm's view that to know what God is is to know that he is, he also could not reject the identification of God's Being with his essence. According to Thomas, with all finite creatures it is the case that what they are—their essence—is one thing, and that they are—their existence—is another. But God simply is Being—*Esse Ipsum Subsistens.* Because this is so, Thomas was obliged to agree with Anselm that *if* God exists, he exists necessarily. But from this it does not follow that God does exist. That there is such a being, who *is* Being, is shown, according to Thomas, by a posteriori proofs. And of course in Thomist terms it is improper to think of God as just *a* being, one entity among others. The difficulty here, however, is derived from difficulties that are implicit in the notion of the God of monotheism and not from difficulties in the notion of Being itself.

CENTRAL QUESTIONS

We are now in a position to discriminate different kinds of questions about Being raised by the Greeks and the Scholastics.

IS EXISTENCE A PREDICATE? How should we characterize the difference between ascribing existence to a subject and ascribing a property to a subject? Is "is" ever a predicate? If it is, what sort of predicate? Later writers who have discussed this problem include René Descartes, in his version of the Ontological Proof; Gottlob Frege, with his clarification of the nature of predicates; G. E. Moore, with his argument that "existence" is not a predicate because we cannot, for example, significantly replace "growl" with "exist" in all the quantified and negated forms of "Tame tigers growl"; and W. V. Quine, with his analysis of Being as "to be is to be the value of a variable." This list of names points up the fact that these questions are susceptible of solution only within the philosophy of logic, and the solution depends upon an adequate characterization of names, predicates, variables, functions, and so on. It is also clear that it is of primary importance to discriminate the metaphysically noncommittal "is," formalizable by means of the existential quantifier, from other uses of the verb "to be" that are far more committed in their implications. Noncommittal uses of the verb appear in ordinary language in such expressions as "There is a prime number between six and eight," "There are three basic colors," "There is a mountain more than 29,000 feet high." Other uses of the verb "to be," however, are far more committed. For example, in the statement "Rachel wept for her children because they were not," "to be" is equivalent to "to be alive." Clearly, however, if I say "There is such-and-such a prime number," there is no such implication; hence, this sense of "there is" must be different.

One finds that all analyses of existential assertions that treat them as predicative are generally unsatisfactory. Briefly, the reason for this is that predicates refer to properties, and properties are what discriminate individuals from each other and enable us to pick out similarities and dissimilarities, and hence to classify. But Being cannot be a property in this sense, for it is not something that it is logically possible for two objects either to have or not to

have in common. Two objects cannot be said to resemble each other in virtue of their both being, and since existence is not a shared property, it cannot characterize a class of objects. For this reason, Being can be neither a property nor a genus.

Of course some philosophers—Gottfried Wilhelm Leibniz, for example—have talked as though Being were a property shared by actual objects but not possessed by *possibilia.* There is no objection to talking like this, provided that it is noticed that the word *property* is not now being used to refer to distinguishable characteristics of real things. Hence, the assertion by such philosophers that Being is a property is not compatible with the Thomas Aquinas–Moore view that it is not, given the two different senses in which the word is used.

ABSTRACT ENTITIES. How do we characterize the status of abstract entities, numbers, possibilities, fictions? These are all different problems, each of them complex. They are envisaged as part of the problem of Being, partly because of our ordinary use of "There is/are" in, for example, "There are two possibilities" or "There is a prime number between six and eight," and partly because of a misunderstanding involved in describing certain possibilities by such terms as "real." When we apply the adjective "real" both to possible states of affairs and to actual states, we suggest that there is a realm of reality wider than the merely existent. This is one source of the belief that there is a genus Being, of which the existent and the nonexistent (such as the possible) are species. Everything called real belongs to the realm of Being. The mistake lies in not seeing the difference between the way in which "real" functions as an adjective and the way in which "reality" functions as a noun. If I call a dollar bill real, I contrast "real" with "counterfeit." If I call a painting "a real Vermeer," I contrast it with a copy. But I do not ascribe to dollar bill and painting the common property of "being real," in virtue of which they belong to the same realm, that of "reality." To say that there is a kind of Being in which both what exists and what does not exist can share is obviously to commit the same mistake. But at this point we have returned to the question of whether Existence and Being can be properties, which belongs to our first group of questions.

THE CHARACTERIZATION OF BEING-AS-SUCH. Can we find any characteristic that belongs to everything that is and that may therefore be said to characterize Being-as-such, rather than individual objects? Here again, one must distinguish two kinds of questions. Aristotle pointed out that of any object whose existence I affirm, I shall also be able to say that it is *one,* that it is *an* object. That is, by picking out something for the purpose of saying that it is, or that it is such and such, I pick it out as an individual. But just because this is so, individuality or unity is not something that it is logically possible for a given object to possess or not to possess any more than existence is; hence, they are not properties any more than existence is. The Aristotelian question of what concepts must be applicable to anything that exists must not, therefore, be identified with the question of whether there are any properties that belong to everything that exists.

There might, of course, have been some property that belonged to everything that existed just as a matter of contingent fact. The world might have been such that everything was green or cubic, or made of blancmange. But this would be philosophically uninteresting (quite apart from the fact that in most such worlds there would be no philosophers). It has been held, however, that it is necessary on, for example, metaphysically epistemological grounds that everything that is shall be of a certain character. Hence Plato's view in his middle period that only Forms exist, and hence Leibniz's view that there are only monads, and George Berkeley's view that to be is always either to be percipient or to be perceived.

ABSOLUTE BEING. Is there a being who exists without the limitations of finite beings and who may therefore just be said to be? This is the question of God's existence.

REALM OF BEING-AS-SUCH. Is there—beyond, over, and above the being of individual objects—a realm of Being-as-such? If so, what is its character? The belief that there is such a realm has always haunted metaphysics. The notion that Aristotle held such a belief has pervaded the history of metaphysics. This misinterpretation of Aristotle has similarly been foisted upon Thomas, and a neo-Thomist myth of the history of philosophy has been constructed in which the four questions that have already been distinguished, all of which are genuine questions, are merged into this fifth question, whose character is much more dubious. It then becomes possible to suggest that there is a single problem: "What is Being?" to which different philosophers have given rival answers. The kind of metaphysics to which reference is being made can be found in Jacques Maritain's *Preface to Metaphysics,* where Maritain is ostensibly expounding Thomas. In order to treat Being as a subject matter, however, Maritain invokes what he calls the intuition of Being, a notion that cannot be found anywhere in Thomas. Thomas, as we have already seen, never treated "Being" as the name of an

independent subject matter and thus had no reason to suggest any means of becoming aware of the existence of such a subject matter.

The kind of history of metaphysics to which reference is being made can be found in D. A. Drennan's *A Modern Introduction to Metaphysics,* which asserts that to the question "What is Being?" Parmenides replied that it was One; Plato, that it was One and Many; Aristotle, that it was Substance; Descartes, that it was Substance in the modes of thought and extension; and so on. Nevertheless, an awareness of the nonexistence of the single question of Being rids us of the misleading idea that we have here a set of competing answers to a single question.

The temptation to see the history of metaphysics in this light seems often to be provoked by an espousal of the metaphysics that makes "Being" a name. We can illustrate this point by considering two sequences in the history of modern philosophy. Hegel argued that Being is the most fundamental of concepts because the most elementary forms of judgment must involve some assertion of existence, no matter how bare. But, he continued, the notion of Being by itself is the emptiest of all notions. Merely to say of something that it is, is to say nothing at all about it; hence, the notion of Being merges into that of its apparent opposite, Nothing. It is not necessary to follow through the Hegelian scheme of categories to see that Hegel is, in fact, extremely cautious at this point. His extreme antidualism always led him to assert that there is nothing else beyond what we confront in experience. The Hegelian Absolute is the rational culmination of historical experience, not a power beyond and outside it. Similarly, for Hegel, Being is a concept expressed in our judgments of experience at a certain level, not the name of a realm beyond all judgments about experience.

In Nicolai Hartmann's philosophy, however, we find a misreading of Hegel parallel to the neo-Scholastic misreading of Thomas and Aristotle. In *Grundzüge einer Metaphysik der Erkenntnis,* Hartmann begins by stating a set of antinomies between, for example, the nature of consciousness as consciousness of what is other than itself and the nature of consciousness as self-contained, so that whatever consciousness is aware of is part of consciousness. That is, Hartmann describes consciousness in two ways that appear incompatible and then inquires how he may reconcile these two descriptions. Instead of asking whether the incompatibility is perhaps only apparent, however, he suggests that the problem arises, and is soluble, because both the knowing, conscious subject and the known object exemplify modes of Being, although different modes. Clearly, it is true that both knower and known *are,* but equally clearly—for reasons given earlier—this is not a property that is open to further study and that has strange characteristics that enable us to resolve antinomies. This, however, was Hartmann's conclusion, and he attributed it to Hegel. He merged Hegel's classification of different subject matters and his scheme of concepts in order to read him as a metaphysician who understood Being as having different grades and modes.

Just as Maritain misreads Thomas and Hartmann misreads Hegel, so Martin Heidegger has misread the pre-Socratics. Heidegger's own views have a mixed ancestry. Søren Kierkegaard, one of the important influences on him, in the *Concept of Dread* writes of dread as an experience whose object is Nothing. Usually in Kierkegaard this sort of statement appears to be a dramatically effective and logically innocent way of characterizing dread as objectless, but at times it seems as if Kierkegaard is no longer saying that dread has no object. Rather, he gives it a particular object whose name is "Nothing," thus making—but not as a joke—the mistake of the Red King in *Through the Looking-Glass,* who thought that if Nobody had passed the messenger on the road, Nobody should have arrived first. To treat "Nothing" as a name is like treating "Something" as a name and easily becomes a counterpart to treating "Being" as a name, as it does with Heidegger. Heidegger takes up Leibniz's question, "Why is there something rather than nothing?" He objects that this question does not take seriously the fact that Being and Nothing necessarily exist together as contrasted and opposed powers. Heidegger allows that he is using "Being" and "Nothing" as names and is therefore involved in treating "Nothing" as if it were the name of something. He even allows that this is "unscientific," but he concludes that this is so much the worse for science and so much the better for philosophy and poetry. Being and Nothing are not objects, and Being is indeed sharply contrasted with beings. Logic presupposes Being and Nothing, but they lie beyond the grasp of logic. Heidegger treats what others have written of the indeterminateness of the concepts as evidence of the elusiveness of Being and Nothing.

Heidegger extends his metaphysics into the history of philosophy by finding his views anticipated in the thought of Heraclitus and Parmenides. The evidence for this claim depends partly on a set of unreliable etymologies that Heidegger thinks he has found for key Greek words, but even when Heidegger is plausible in his interpretation at the linguistic level, he is at the least anachronistic in his view of the kind of problem the pre-Socratics

confronted. They progressively recognized as paradoxical, and therefore as needing reformulation, those very forms of utterance that to Heidegger are and remain fundamental.

If the philosophy of Being has bred not merely rival doctrines but rival views of the history of philosophy, it has also bred rival diagnoses of the errors involved in treating "Being" as a noun. A. J. Ayer has suggested that a misuse of the verb "to be" is the root of the error. This would imply, however, that standard forms of grammar embodied in ordinary usage are somehow philosophically normative—and this appears to get matters upside down. Linguistic distortion is certainly liable to breed confusion, but there is, in fact, nothing grammatically wrong with forming a verbal noun such as "Being" as an analogy with, for example, "riding." "Riding" is used as the name of an activity; why, then, should "Being" not be made into a name? It is surely because of the logical and metaphysical confusion involved that we want to criticize the linguistic construction and not because the linguistic construction itself is an error.

John Dewey diagnosed a twofold root of errors about Being. They are partly a survival from religious modes of thought, the retention of belief in a realm free from change and decay and separate from the realm of sense perception. This is explained by the fact that although mythological thought has been discredited, the impulses behind it still need satisfaction. Also, belief in changeless Being is a consequence of man's habit of abstracting truths from the contexts of practice and activity in which they were acquired (and where alone they have meaning) and treating them instead as belonging to a timeless realm in which they wait upon our apprehension. Dewey's diagnosis, however, while it may explain how we come to hold and retain confused views of Being, does not embody an explanation of why the views are confused, except perhaps to those who are already convinced in general of the truth of Dewey's pragmatism.

In order to clarify the issue, we must, in fact, make the sort of analysis of concepts that Aristotle used in the *Metaphysics.* We may expect any analysis of the concept of Being to vary with the general framework of concepts within which it is considered. Aristotelians, Hegelians, and Quineans will not all agree, but any analysis that fails to discriminate the different questions involved, and that fails to identify the confusion that results from merging them into a single question, will be doomed to conceptual error and very likely to a misreading of the history of philosophy as well.

See also Absolute, The; Anselm, St.; Aristotle; Ayer, Alfred Jules; Descartes, René; Dewey, John; Existence; Frege, Gottlob; Hartmann, Nicolai; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Heraclitus of Ephesus; Kierkegaard, Søren Aabye; Maritain, Jacques; Medieval Philosophy; Moore, George Edward; Nothing; Ontological Argument for the Existence of God; Ontology; Parmenides of Elea; Plato; Quine, Willard Van Orman; Time, Being, and Becoming; Universals, A Historical Survey; Thomas Aquinas, St.

Bibliography

TEXTS

Aristotle. *Metaphysics.* Translated by H. Tredennick. Cambridge, MA, 1945.

Aristotle. *Metaphysics.* Edited by W. D. Ross. Oxford translation. Oxford, 1948.

Cajetan. *Commentaria in De Ente et Essentia.* Edited by M. H. Laurent. Turin, 1934.

Cornford, F. M. *Plato and Parmenides.* London, 1935; New York: Liberal Arts Press, 1957. A translation of the *Parmenides.*

Dewey, John. *The Quest for Certainty.* New York: Minton, Balch, 1929.

Duns Scotus. *Opera Omnia,* 12 vols. Paris, 1891–1895. Vol. III, *Quaestiones Subtillissimae Super Libros Metaphysicorum Aristotelis.*

Hartmann, Nicolai. *Grundzüge einer Metaphysik der Erkenntnis.* Berlin: de Gruyter, 1925.

Hegel, G. W. F. *The Science of Logic.* Translated by W. H. Johnston and L. G. Struthers. New York, 1929.

Heidegger, Martin. *Existence and Being.* Translated by D. Scott, R. Hall, and A. Crick. Chicago: H. Regnery, 1949. Includes various works.

Kant, Immanuel. *Critique of Pure Reason.* Translated by Norman Kemp Smith. London: Macmillan, 1934.

Kirk, G. S., and J. E. Raven, eds. *The Presocratic Philosophers.* Cambridge, 1960. Ch. 10, Parmenides texts. Also contains translation and commentary.

Plato. *Platonis Opera.* Edited by J. Burnet. Oxford, 1899–1906. The *Theaetetus* and the *Sophist.* Translated by F. M. Cornford as *Plato's Theory of Knowledge.* London: K. Paul, Trench, Trubner, 1935; New York: Liberal Arts Press, 1957.

Quine, W. V. *From a Logical Point of View.* Cambridge, MA, 1953.

Thomas Aquinas. *De Ente et Essentia.* Edited by M. D. Roland-Gosselin. Le Saulchoir, 1926. Translated by Armand Maurer as *On Being and Essence.* Toronto, 1949.

Thomas Aquinas. *In Metaphysicorum Aristotelis Commentaria.* Edited by M. R. Cathala. Turin, 1935.

Thomas Aquinas. *Summa Theologica.* Ottawa, 1941–1945. Translated by the Dominican Fathers of the English Province, 3 vols. New York: Benziger, 1947.

Wolff, Christian von. *Philosophia Prima Sive Ontologia.* Frankfurt and Leipzig: Officina libraria Rengeriana, 1736.

DISCUSSIONS OF TEXTS

Gilson, Étienne. *L'Ètre et l'essence.* Paris: J. Vrin, 1947.

Gilson, Étienne. *Being and Some Philosophers.* Toronto: Pontifical Institute of Mediaeval Studies, 1949.

Martin, Gottfried. *Immanuel Kant: Ontologie und Wissenschaftstheorie.* Cologne: University of Cologne, 1951. Translated by P. G. Lucas as *Kant's Metaphysics and Theory of Science.* Manchester, U.K.: Manchester University Press, 1955.

Owens, Joseph. *The Doctrine of Being in the Aristotelian Metaphysics.* Toronto: Pontifical Institute of Mediaeval Studies, 1951.

ARTICLES

Geach, P. T. "Form and Existence." *PAS* 55 (1954/1955): 251–272.

Geach, P. T., A. J. Ayer, and W. V. Quine. "Symposium: On What There Is." *Aristotelian Society Supplement* 25 (1951): 125–160.

Moore, G. E. "Is Existence a Predicate?" *Aristotelian Society Supplement* 15 (1936): 175–188.

Quine, W. V. "On What There Is." *Review of Metaphysics* 2 (1948–1949): 21–38.

Weiss, P. "Being, Essence and Existence." *Review of Metaphysics* 1 (1947–1948): 69–92.

Alasdair MacIntyre (1967)

BELIEF

Beliefs are a species of propositional attitude distinguished by their having the mind-to-world direction of fit.

Propositional attitudes are psychological states characterized by a psychological mode, Ψ, and a propositional content, *P*, schematically: $\Psi(P)$. My belief that the earth moves has *belief* as its psychological mode, and *that the earth moves* as its propositional content. A desire that the earth move has the same propositional content, but a different psychological mode, *desire*. Within a psychological mode, propositional attitudes are distinguished by their contents. I could not have two beliefs with the content that the earth moves. Many, though not all, propositional attitudes admit of a bivalent evaluation. Beliefs are true or false. Desires are satisfied or unsatisfied. Intentions are carried out or not carried out. Propositional attitudes with a bivalent evaluation have either the mind-to-world direction of fit or the world-to-mind direction of fit (Searle 1983, chapter 1). Its direction of fit expresses the basic function of a propositional attitude in our mental economy. Beliefs aim to represent how the world is independently. They aim at truth. The belief that Solomon was wise is true if and only if (iff) its content matches the world, that is, iff Solomon was wise. Belief's aim to represent how the world is independently is reflected in its being irrational to retain a belief when one sees that it does not match the world. Thus, beliefs have the mind-to-world direction of fit. A desire, in contrast, seeks not to match how the world is independently, but for the world to come to match its content. Desires seek satisfaction. The desire that a toothache go away is satisfied iff the world comes to match its content, that is, iff the toothache goes away. It is not irrational to retain desires known to be unsatisfied, for seeking satisfaction gives them their point. Desires thus have the world-to-mind direction of fit. Beliefs and desires, in virtue of their opposite directions of fit, have an interlocking role in the production and explanation of action.

BELIEF, SENSATION, EXPERIENCE, AND CONCEPT

Beliefs (and other propositional attitudes) must be distinguished from sensations, sensory images, and experience, on the one hand, and concepts, on the other. The classical British empiricists of the seventeenth and eighteenth centuries—John Locke (1632–1704), George Berkeley (1685–1753), and David Hume (1711–1776)—were unable to provide an adequate account of belief because they assimilated all of these to sensations or sensory images, like the taste of apple or a toothache. But sensations are not an adequate model for belief, or for other propositional attitudes. Sensations are not true or false, or satisfied or unsatisfied. They do not admit of a bivalent evaluation, as propositional attitudes do. They do not have propositional contents. Their differences are differences of qualitative feel. These differences are not variations in psychological function, as are the psychological modes of belief and desire. In particular, propositional contents are required to make sense of the logical relations that obtain between beliefs, and which are crucial to understanding the role of beliefs in reasoning and action. For example, understanding the validity of the inference from the belief that gold is a metal, and the belief that this ring is gold, to the belief that this ring is metal, requires seeing the logical connections between the propositional contents of the first two beliefs and the last, and their shared elements. Similarly, as explained below, the logical relations between the contents of belief and desire are crucial to understanding rational action (see the discussion of the practical syllogism below). Since sensations lack propositional contents, treating beliefs as a subclass of sensations obliterates distinctions needed to understand the role of beliefs in thought and action.

Perceptual experiences, unlike sensations, can be veridical or nonveridical. A hallucination of a rhinoceros on the sofa is a nonveridical perceptual experience. It represents what is not so. Perceptual experiences are like beliefs in admitting of bivalent evaluation and having the mind-to-world direction of fit. But though many beliefs are based on perceptual experience, they are a different coin. Perceptual experience is a fieldlike representational medium and makes use of the qualitative features of modes of consciousness associated with different sensory modalities (visual, auditory, tactile, etc.) to represent how things are around us. Beliefs, in contrast, do not in the same way make use of qualitative features of modes of consciousness to represent. Their mode of representation is purely propositional. Beliefs are to perceptual experiences as statements are to maps. Perceptual beliefs, those based directly on perceptual experience, in effect abstract from the richer representational content of perceptual experience.

Just as beliefs, sensations, and experiences must be distinguished from each other, they must also be distinguished from the shared elements, concepts, in different attitude contents. The concept of gold, for example, is shared between the belief that gold is a metal and the belief that this ring is gold. It is a constituent of both contents. Concepts are neither true nor false, though they apply or fail to apply to things. Sensations neither have such constituents nor are identical with them, for belief contents do not have sensations or images as constituents. Experiences involve concepts, much as beliefs do. A visual experience *as of* a baseball represents a spherical object as a baseball. The experience is distinct from the concept of a baseball, but represents something visually presented as falling under the concept. Without the concept, there could be no such visual experience.

Someone who believes that gold is a metal *possesses* the concept of gold and the concept of metal. Possessing a concept requires having beliefs expressing the simplest conceptual connections that the concept enters into. For example, someone who possesses the concept of a gun must believe that guns are weapons, that they are physical objects, that they can be aimed, and the like. Perhaps no precise set of beliefs is required, but if a person lacks a network of beliefs articulating the connections of a concept with other concepts, he does not possess the concept. This shows that we can make sense of attributing a belief to someone only by locating it in a pattern of beliefs, and that the other propositional attitudes and perceptual experience presuppose belief.

BELIEFS AND THE EXPLANATION OF ACTIONS

Beliefs, because they aim at truth, play a central role in theoretical reasoning (reasoning about what is so), and hence in practical reasoning (reasoning about what to do).

Theoretical reasoning is central to practical reasoning because we get what we want by doing something that promotes it. We therefore need to know what we can do, and how what we can do is related to what we want. When seeking knowledge of these things, we seek true beliefs about them. Thus, what we do is conditioned by what we believe, whether our aim at truth hits the mark or not. Accordingly, when one explains an action, it is not enough to cite a desire that the action satisfies. One must also cite a belief that the action increased the likelihood of satisfying the desire. If I want my rival to come to grief and an idle comment of mine leads to his downfall, my desire that he come to grief does not help explain my bringing about his downfall if I did not think that my idle comment would have that as a consequence.

An action explanation provides the materials for a practical syllogism that justifies the action from the point of view of the agent. Suppose that I lifted my finger because I wanted to signal you and believed that my lifting my finger, *F*, would constitute signaling you. One can construct the following argument in favor of this action, drawing the evaluative premise from the desire and the factual premise from the belief:

My signaling you is desirable (has a desirable aspect).
F, being lifting of my finger, is a signaling of you.

F is desirable (has a desirable aspect).

This does not show that the action is desirable all things considered, but only that it has a desirable aspect. Actions have many consequences and properties, some of which one may want and others of which one may not. In deciding all things considered what to do, the agent must rank his desires and take into account his degrees of confidence in desired outcomes attending certain actions, or, in a common phrase, his degrees of belief in outcomes, given the actions. Candidate actions are evaluated on the basis of the value of their results and one's degree of confidence in their having those results. If the chance is low but the value great, the undertaking may still be judged best overall. The notion of degree of belief in a proposition is often treated as a generalization of the ordinary notion of belief. Typically, on buying a ticket, one does not believe that one will win the lottery, though one does

have a nonzero degree of belief that one will. However, it may be that degrees of belief in a proposition can be assimilated to beliefs about its probability.

BEHAVIORIST THEORIES OF BELIEF

Logical behaviorism, a form of materialism, holds that ascribing beliefs and desires and other mental states to an agent is just a compendious way of describing a complex pattern of his actual and potential behavior. On this view, there are no inner mental states or events—no "ghost in the machine," in Gilbert Ryle's memorable phrase. If logical behaviorism is true, action explanation is not causal explanation, but rather functions by locating some particular behavior in a broader pattern of behavior. Logical behaviorism was championed by the logical positivists of the 1930s (see Ayer), and in more subtle forms by Ryle and Ludwig Wittgenstein in the 1940s and 1950s. A central motivation for the logical positivists was their verificationist criterion of meaning, according to which the meaning of a statement should be sought in the conditions that verify or falsify it.

Logical behaviorism fell from fashion after the Second World War (Block). One reason was disenchantment with the verificationist criterion of meaning. A second was the failure to provide necessary or sufficient conditions for attributing psychological states in purely behavioral terms. This failure is connected with the interlocking roles of belief and desire in explaining behavior, and carries an important lesson about the relation of belief (and other propositional attitudes) to behavior.

Let us try to say what behavior is characteristic of the belief that there is an apple pie in the pantry. It is clear that what behavior we expect will depend on what the agent wants and what else she believes. If she is not hungry, we expect no tendency to investigate the pantry. If she likes apple pie but wants to save it for the guests more than to indulge herself, we expect a delayed advance on the pantry. If she believes it is poisoned, we expect her to dispose of it. And so on. The important point is that the behavior we expect from someone who has a particular attitude is conditioned by the other psychological states that we think he has. This makes it impossible to state, in purely behavioral terms, what it is to believe, for example, that there is a pie in the pantry.

Once one understands the role of desire, in particular, in action, one can see that it will play the spoiler for any behaviorist program. For desires can be about behavior usually taken to be a sign of a given sort of psychological state. In particular, one may want to exhibit misleading behavior. One can pretend to believe or want things one does not, and one can suppress behavior that expresses what one wants. It seems plausible, as Hilary Putnam argued in "Brains and Behavior," that the limited deceptions with which we are familiar could take forms that would preclude any behavioral expression of some of an agent's psychological states throughout his life.

BELIEFS AS CAUSES OF BEHAVIOR

The moral is that beliefs issue in behavior only in conjunction with appropriate other propositional attitudes. Desires motivate behavior, but beliefs guide it. Each is impotent without the other. To vary a phrase of Kant's, desire without belief is blind; belief without desire is empty. We cannot read back from behavior to the motivating belief and desire, because any given behavior may issue from different sets of beliefs and desires. Behavior is related to belief and desire as symptoms to a disease. The symptoms may be expressed in the absence of the disease, and the disease may be present without being expressed by any symptoms, or by the usual symptoms. The reason is that the disease is a cause of the symptoms. Its giving rise to the usual symptoms depends on the usual background conditions being present. This analogy suggests that, as with disease and background conditions, beliefs and desires are interlocking causes of behavior. This would explain the failure of the behaviorist programs, since, as Putnam put it, "*causes* are not logical constructions out of their *effects*" (p. 27).

This conclusion is bolstered by an argument by Donald Davidson in his seminal paper "Actions, Reasons, and Causes." We sometimes have multiple reasons (belief and desire pairs that potentially explain an action) for doing something, but we act on only one of them. I may believe that if I do not obey the speed limit when driving, I will likely receive a ticket, and I may wish to avoid receiving a ticket. I may believe also that obeying the speed limit is the right thing to do, and wish to do the right thing. I may obey the speed limit for the first reason rather than the second, or for the second rather than the first. In either case, each reason would justify what I do, but only one would explain it. We can make sense of this being so if we hold that the reasons for which I act are those that cause my action.

VOLUNTARISM ABOUT BELIEF

Beliefs play a role in guiding and explaining action, but can they be the products of actions? Can one come to believe something by choosing or deciding to believe it? According to voluntarism about belief, the answer is yes.

René Descartes (1596–1650), "the father of modern philosophy," apparently endorsed belief voluntarism. In his explanation in the *Meditations* (pp. 37–43) of how we fall into error despite God's supreme benevolence, he assumed that when one lacks adequate evidence one can choose to believe something and hence fall into error through the exercise of free will, and thereby absolve God of responsibility for the error. With his famous wager in *Pensées* (pp. 151–153), Blaise Pascal (1623–1662) argued that the infinite utility that attaches to believing truly that God exists means that no matter how small the probability, one is rationally compelled to belief. This argument likewise seems to assume belief voluntarism. In "The Will to Believe," William James (1842–1910), like Pascal, argued that it is not only possible but sometimes rationally required that we make decisions about belief when evidence underdetermines choice.

One may, of course, bring oneself to believe something indirectly. I may pay another to brainwash me. I may adopt the outer forms of religious faith in the hope that inner conviction will follow. But the issue is not whether I can bring myself to believe something by doing something else that brings it about, but whether I can do this without doing anything else to bring it about.

This seems not to be something that is typically within our power. I cannot just decide now to believe that I do not have hands or a head, or that I am flying, or fabulously wealthy. Religious belief, which is less engaged with the practical, is a more difficult case and has historically been the focus of the debate about belief voluntarism. One can try to inculcate religious belief indirectly, but can one simply decide to believe that God exists, while also believing one lacks adequate evidence? One might answer yes because it can be reasonable to continue to believe that God exists in the face of doubts. But this falls short of what is required. For this is not deciding to believe, but rather deciding not to give up a belief one already has.

Still, we are sometimes faced with an unavoidable practical choice where evidence bearing on a crucial fact leaves the fact, and hence the choice, undetermined. Must we not then make a choice about what to believe despite not having reasons to believe one way or the other? No. We must make the practical choice about what to do. But this does not imply belief. When some action is better than none, we take a chance and hope for the best, without belief.

Is the situation arguably different when a belief itself has a practical value? It is dubious that a belief itself having a practical value makes it easier to conceive of choosing to believe. If someone were to offer me a million dollars to believe that there is life on the Sun, I might wish to do so, but I would be baffled about how to comply.

The purpose of belief is to represent the world accurately. Therefore, belief serves its role only if the formation, retention, and revision of belief are sensitive to what one takes to be one's evidence. This does not mean that we always believe in proportion to our evidence. We make mistakes of assessment and reasoning; we are lazy; we fail to attend to evidence we have. Nor does it mean that what we take to be evidence is evidence, or that all our opinions were entrenched with the spade of reason. Further, it does not mean that belief is never influenced by desire. Otherwise, wishful thinking—believing what one wants to be true—would not be possible. Rather, it means that belief, by its nature, aims at truth, that its function is undermined if one lets belief formation be sensitive to anything other than what one takes to be evidence. Where there is uncertainty, one may be conservative and persist in a belief. But not to give up a belief in the face of strong contrary evidence is irrational. Worse still is to believe where no evidence bears, as in the case of wishful thinking. Even when belief has a practical benefit—as when believing one will win a race increases one's chances, but not enough to warrant the belief—rationality is at best at war with itself. Belief voluntarism thus seems to be something that can occur, at best, only on the fringes of rationality—in the shadow regions of self-deception.

See also Action; Ayer, Alfred Jules; Behaviorism; Berkeley, George; Davidson, Donald; Descartes, René; Hume, David; James, William; Locke, John; Materialism; Pascal, Blaise; Propositional Attitudes: Issues in Philosophy of Mind and Psychology; Ryle, Gilbert; Sensa; Voluntarism; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ayer, Alfred Jules, ed. *Logical Positivism*. New York: Free Press, 1959.

Block, Ned. "Psychologism and Behaviorism." *Philosophical Review* 90 (1) (1981): 5–43.

Davidson, Donald. "Actions, Reasons, and Causes" (1963). In his *Essays on Actions and Events*. New York: Clarendon Press, 2001.

Descartes, René. "Meditations on First Philosophy" (1641). In *The Philosophical Writings of Descartes*. Cambridge, U.K.: Cambridge University Press, 1984.

James, William. "The Will to Believe" (1897). In his *The Will to Believe and Other Essays in Popular Philosophy*. Cambridge, MA: Harvard University Press, 1979.

Pascal, Blaise. *Pensées* (1670). Harmondsworth, U.K.: Penguin, 1966.

Price, H. H. *Belief: The Gifford Lectures Delivered at the University of Aberdeen in 1960*. London: Allen and Unwin, 1969.

Putnam, Hilary. "Brains and Behavior." In *Readings in Philosophy of Psychology*, Vol. 1, edited by N. Block. Cambridge, MA: Harvard University Press, 1980.

Ryle, Gilbert. *The Concept of Mind*. New York: Barnes and Noble, 1969.

Searle, John. *Intentionality: An Essay*. Cambridge, U.K.: Cambridge University Press, 1983.

James, William. "The Will to Believe" (1896). In his *The Will to Believe and Other Essays in Popular Philosophy*. Cambridge, MA: Harvard University Press, 1979.

Kirk Ludwig (2005)

BELIEF ATTRIBUTIONS

"Belief attributions" are uses of sentences of the form *N* believes that *s* (where *N* is a noun phrase, *s* a sentence). Their semantic and logical properties have been debated under the assumption that an account of "believes" will carry over to other propositional attitudes such as desire, knowledge, and fear. Most of the debate focuses on two issues: Does "believe" pick out a relation, and how do so-called *de re* and *de dicto* attributions differ?

IS "BELIEVES" RELATIONAL?

The obvious hypothesis is that in

(1) Maggie believes that Twain lives.

"believes" has the semantic status of a transitive verb, picking out a relation between a believer and something (a proposition) provided by the verb's complement,

(2) that Twain lives.

Grammatical evidence suggests this: "believes" can be followed by names and demonstratives ("I believe Church's thesis," "she believes that") as well as expressions that behave like (nominal) variables ("whenever the pope says something I believe it").

Gottlob Frege and Bertrand Russell, whose work inspires most subsequent debate about belief attribution, agreed on the obvious hypothesis. Frege held that expressions embedded within "believes that" shift their reference to a way of thinking, or sense, of what they refer to unembedded. Russell held that no such semantic shift occurs; the proposition "that *s*" is determined by what *s*'s parts pick out when used unembedded.

Since "Twain" and "Clemens" refer to the same author, the Russellian approach seems committed to the identity of the propositions, that Twain lives and that Clemens does, and thus to (1)'s implying

(3) Maggie believes that Clemens lives.

Russell would avoid this by saying that "Twain" and "Clemens" typically function as truncated definite descriptions. This last suggestion is widely thought to have been discredited by Saul Kripke.

One problem Fregean views face is that sense is idiosyncratic: Different people associate with a name different ways of thinking of the referent. It is implausible that when I utter (1) I speak truly only if Maggie thinks of Twain as do I. But if (2) in (1) named Maggie's sense for "Twain lives," the argument "Maggie believes that Twain lives; Seth believes what Maggie does; so Seth believes that Twain lives" would be invalid.

Contemporary Russellians such as Nathan Salmon and Scott Soames hold that to believe a proposition involves grasping or representing it and its constituents; thus, belief is a three-place relation among a believer, a Russellian content, and a representation. Salmon and Soames nonetheless hold that (1) tells us only that Maggie believes ("under some representation") the Russellian proposition that Twain lives; the appearance that (1) and (3) may disagree in truth value results from mistaking a conversational or pragmatic implicature, about the representation under which a belief is held, for part of what a belief attribution, strictly speaking, says.

John Perry and Mark Crimmins have suggested that a belief attribution involves implicit reference to the Russellian's representations or modes of grasping: the complement of "believes" determines a Russellian proposition, but the verb has an "implicit argument place" for representations. A use of (1) makes a claim along the lines of "Maggie believes the Russellian proposition that Twain lives under representation *r*," with the representation referred to differing across occasions of use. A problem with this view is that it renders the argument mentioned two paragraphs above invalid.

Some think belief attributions implicitly quotational. The simplest version of such a view sees *that s* as a quotation name of *s*, "believes" naming a relation to sentence types. To this it may be objected that different uses of "Seth thinks I am sad" may have different truth values. Another view sees a "that" clause as picking out a fusion of linguistic items with their interpretations—for example, the result of combining a sentence with the semantic values of its expressions.

Mark Richard's version of this view has *that s* pick out a fusion of the sentence *s* and its Russellian content.

In belief attribution, such fusions are offered as "translations" of the believer's thoughts, where a thought is the result of combining a representation that realizes a belief with its Russellian content: (1) is true if the "that" clause provides a translation of a thought of Maggie's. Standards of translation shift from context to context: "Twain" may represent a representation of Maggie's in some but not all contexts. Thus, on this view, the truth of (1) does not demand that of (3).

Donald Davidson denies that (2) is a semantically significant part of (1). "Believes" is a predicate whose second argument is the demonstrative "that"; its referent is the ensuing utterance of "Twain lives." The overall force of (1) is roughly some belief state of Maggie's agrees in content with that utterance. (Davidson made such a proposal for "says" but clearly intended to generalize.) Yet more radical views deny that "believes" is a predicate. Arthur N. Prior took "believes" to combine with a name and sentence to form a more complex sentence; W. V. O. Quine has entertained the idea that "believes that Twain lives" is a predicate without semantically significant structure. A problem for Quine is to explain how infinitely many (semantically unstructured) belief predicates acquire their meanings; Prior thought little useful could be said on such issues.

DE RE AND *DE DICTO*

There seem to be two ways of interpreting such sentences as

(4) Sam believes that Melinda's husband is unmarried. Sam believes that some Frenchman is not French.

One interpretation attributes to Sam necessarily false beliefs; the other, suggested by

(4') Of Melinda's husband, Sam believes he is unmarried.

Of some Frenchman, Sam believes he is not French, does not. Note that (4') ascribes to Sam beliefs in some sense about particular individuals, while this is not true of the interpretation of (4).

The interpretations seem to correspond to different scopes that may be assigned to the quantifier phrases "Melinda's husband" and "some Frenchman." In a *de re* attribution, an expression functioning as a variable within the scope of "believes" is bound by a quantifier outside its scope (and the scopes of other verbs of propositional attitudes). Interpreting the sentences in (4) as in (4') is *de re* attribution: "he" and "she" are bound to "Melinda's husband" and "some women," which are not in the scope of "believes." An attribution that is not *de re* is *de dicto.* If we accept a relational account of "believes," we will say that a *de dicto* interpretation of "*N* believes that *s*" attributes to *N* a belief in the proposition expressed by *s.* (An attribution might also count as *de re* if it has a term anaphoric on a name outside of the attribution, as in the natural understanding of

(5) Twain was an author, but Seth believes that he was president.)

Not everyone would characterize the *de re–de dicto* distinction as above. Quine held that it is impossible for a quantifier to bind a variable that occurs opaquely—that is, inside a construction, like "believes," which causes failures to substitutivity. If Quine were correct, some other account of the two understandings of (4) is needed. (Quine himself suggested that "believes" is ambiguous.) Quine's view is not widely shared. (See Kaplan, 1986, for discussion.)

The relations between *de re* and *de dicto* attributions are of interest in good part because *de re* attributions are anomalous on some views. A *de re* attribution identifies a belief in terms of the objects it is about, not in terms of how those objects are conceptualized. For a Russellian this is the norm: All there is to belief attribution is identifying the state of affairs believed to obtain. For a Fregean, (4') is at best an aberration, lacking information about sense, which belief attribution is supposed to convey. *De re* belief attributions provide a focus for the debates among Russellians, Fregeans, and others.

See also Causal or Conditional or Explanatory-Relation Accounts; Content, Mental; Davidson, Donald; Epistemology; Frege, Gottlob; Knowledge and Belief; Kripke, Saul; Prior, Arthur Norman; Quine, Willard Van Orman; Russell, Bertrand Arthur William.

Bibliography

Crimmins, M. *Talk about Beliefs.* Cambridge, MA: MIT Press, 1992.

Davidson, D. "On Saying That." In *Inquiries into Truth and Interpretation.* Oxford: Clarendon Press, 1984.

Frege, G. "Uber Sinn und Bedeutung." *Zeitschrift fur Philosophie and Philosophische Kritik* 100 (1892): 25–50. Translated by P. Geach and M. Black as "On Sense and Reference," in *Translations from the Philosophical Writings of Gottlob Frege,* edited by Geach and Black. New York: Philosophical Library, 1952.

Higginbotham, J. "Belief and Logical Form." *Mind and Language* 6 (1991): 344–369.

Kaplan, D. "Opacity." In *The Philosophy of W. V. Quine,* edited by L. Hahn and P. Schilpp. La Salle, IL: Open Court, 1986.

Kaplan, D. "Quantifying In." In *Words and Objections,* edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1969.
Kripke, S. "A Puzzle about Belief." In *Meaning and Use,* edited by A. Margalit. Dordrecht, 1979. Also in Salmon and Soames, 1988.
Perry, J., and M. Crimmins. "The Prince and the Phone Booth." *Journal of Philosophy* 86 (1989): 685–711.
Prior, A. N. *Objects of Thought.* Oxford: Clarendon Press, 1971.
Quine, W. V. O. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53 (1956): 177–187.
Quine, W. V. O. *Word and Object.* Cambridge, MA: MIT Press, 1960.
Richard, M. *Propositional Attitudes.* Cambridge, U.K.: Cambridge University Press, 1990.
Salmon, N. *Frege's Puzzle.* Cambridge, MA: MIT Press, 1986.
Salmon, N., and S. Soames, eds. *Propositions and Attitudes.* Oxford: Oxford University Press, 1988.

Mark Richard (1996)

BELINSKII, VISSARION GRIGOR'EVICH
(1811–1848)

Vissarion Grigor'evich BelinskiI (Belinsky), the Russian literary critic, was an early leader of the Russian intelligentsia and a major representative of German Absolute Idealism, as well as of the subsequent reaction against it, in nineteenth-century Russian philosophy.

Belinskii was born in Sveaborg, Russia (now Finland), the son of a provincial physician. He entered the University of Moscow in 1829 but was expelled after three years, perhaps for the radical criticism of serfdom in a romantic drama he wrote; his subsequent education was self-acquired. He began a journalistic career in 1833 and soon became the chief critic for a succession of literary journals in Moscow and (after 1839) in St. Petersburg, principally *Otechestvennyye Zapiski* (Annals of the Fatherland). His brilliant, philosophically oriented critical essays, including perceptive early appreciations of Nikolay Gogol, Mikhail Lermontov, and Feödor Dostoevsky, won him great renown but little material reward; he died in St. Petersburg after a short life filled with poverty and illness.

Belinskii's intellectual development typifies that of the early Russian "Westernizers," or admirers of Western progressive ideas and institutions, whose leader he became: He passed from the romantic extremes of German Absolute Idealism through G. W. F. Hegel to a mature position representing the influence of the French socialists and Ludwig Feuerbach. In Belinskii's case, the doctrinal changes were magnified and accelerated by a mercurial personality, while their expression was often clouded by the pressures of journalistic writing under tsarist censorship. Belinskii published no systematic theoretical works, and his voluminous critical essays and private correspondence leave room for divergent interpretations of his views.

Belinskii's earliest writings (1831–1836) show the clear influence of Friedrich Schiller and Friedrich von Schelling. Basing his views on Schelling's nature philosophy and philosophy of art, Belinskii glorified art and the creative process, and emphasized man's inner aesthetic and moral experience in rising above empirical reality to the "eternal Idea."

In 1837, after a brief enthusiasm for Johann Gottlieb Fichte, Belinskii was introduced by his friend and mentor, Mikhail Bakunin, to the thought of Hegel. Belinskii found in the Hegelian formula "all that is real is rational" a summons to a "reconciliation with reality" that turned his attention from man's subjective world to the objective reality around him and led him to praise Russian autocracy, to view the state as sacred, and to regard society as metaphysically and ethically superior to the individual. He expressed a Hegelian conception of art as "thinking in images" and as reproducing rational reality.

Belinskii's Hegelianism, however, did not extinguish the regard for human individuality that in some degree had always marked his thinking and had been manifested most explicitly during his brief Fichtean period. By 1841 he repudiated Hegel's subordination of the individual and thenceforth turned from Absolute Idealism to an ethical personalism that emphasized the supreme value of the individual personality. At the same time, he abandoned the attempt to show the rationality of the tsarist order: He became acquainted with the writings of Comte de Saint-Simon and other French socialists, and called increasingly for radical social reforms in the direction of democracy and socialism. His mature view of art stressed art's moral and political functions in expressing socially progressive ideas, for which reason he is generally regarded as the founder of the dominant tradition of social or "civic" criticism in Russia.

Belinskii's socialism remained individualistic in inspiration, and there is evidence that toward the end of his life he moved to a more moderate liberal position, advocating the development of a middle class in Russia. His reformist enthusiasm and generally enlightened outlook were well expressed in a famous "Letter to Gogol" (1847), which set a moral tone for the Russian intelligentsia for generations. The "Letter" illustrates the antiecclesiasticism and positivist leanings of Belinskii's final

period, if not the outright atheism and materialism attributed to him by Soviet interpreters.

See also Hegel, Georg Wilhelm Friedrich; Hegelianism; Russian Philosophy.

Bibliography

WORKS BY BELINSKII

Polnoye Sobraniye Sochineni (Complete Works). 13 vols. Moscow, 1953–1959.

Selected Philosophical Works. Westport, CT: Hyperion, 1981.

Russian Philosophy. Vol. 1, edited by James Edie, et al. Knoxville: The University of Tennessee Press, 1987, pp. 281–320.

WORKS ON BELINSKII

Bowman, Herbert E. *Vissarion Belinski, 1811–1848: A Study in the Origins of Social Criticism in Russia*. Cambridge, MA: Harvard University Press, 1954.

Ranall, F. B. *Vissarion Belinskii*. Newtonville, MA: Oriental Research Partners, 1987.

Terras, V. *Belinksij and Russian Literary Criticism: The Heritage of Organic Aesthetics*. Madison: University of Wisconsin Press, 1974.

Zenkovskii, V. V. *A History of Russian Philosophy*. 2 vols. Translated by G. Kline. New York and London: Routledge, 2003.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

BELL, JOHN, AND BELL'S THEOREM

John Stewart Bell (1928–1990), a truly deep and serious thinker, was one of the leading physicists of the twentieth century. He became famous for his discovery that quantum mechanics implies that nature is nonlocal, that is, that there are physical influences between events that propagate faster than light.

From 1960 until his death Bell worked at the Conseil Européen pour la Recherche Nucléaire (CERN; European Laboratory for Particle Physics) in Geneva, Switzerland, on the physics of particle accelerators, making a number of important contributions to high-energy physics and quantum field theory. Noteworthy was his discovery in 1969, together with Roman W. Jackiw, of the so-called "Bell-Jackiw-Adler" anomaly (discovered independently by Stephen Adler), a mechanism explaining physical effects such as neutral pion decay (which are unexplainable on the basis of the symmetries of the classical field Lagrangian), in terms of an "anomalous" term arising from the renormalization of quantum field theory. Since then this mechanism has become an important cornerstone of quantum field theory. Another important contribution was the argument he gave in 1967 for why weak interactions should be described using a gauge theory.

John Bell was one of the leading experts—perhaps *the* leading expert—on the foundations of quantum mechanics. The book collecting his articles on this subject, *Speakable and Unspeakable in Quantum Mechanics* (1987), is unsurpassed for clarity and depth and it is still the best reference for whoever wishes to learn about the field.

Bell strongly opposed the "Copenhagen interpretation" of quantum physics, according to which macroscopic objects, such as chairs and planets, do exist out there, but electrons and other microscopic particles do not. According to the Copenhagen view, the world is divided into two realms, macro and micro, "classical" and "quantum," logical and contradictory—or, as Bell put it in one of his essays, into "speakable" and "unspeakable." Along with Albert Einstein, Erwin Schrödinger, Louis de Broglie, and David Bohm, Bell was one of the few physicists compelled by his conscience to reject the Copenhagen interpretation.

Bell emphasized that the empirical facts of quantum physics do not at all force us to renounce realism. There is, in fact, a realist theory (Bohmian mechanics, also known as the de Broglie–Bohm theory) that accounts—insofar as the nonrelativistic theory is concerned—for all of these facts in a most elegant way. This theory describes a world in which electrons, quarks, and the like are point particles, always having positions that move in a manner dictated by the wave function. It should be taught to students, Bell insisted, as a legitimate alternative to the prevailing orthodoxy. After GianCarlo Ghirardi, Alberto Rimini, and Tullio Weber succeeded in formulating in 1986 a second kind of realist theory, Bell encouraged the further development of this theory as well (1987). He thought that such a theory contained the seeds of a reconciliation of quantum mechanics with fundamental Lorentz invariance, and thus a resolution of the tension between quantum mechanics and relativity that arose from his own work on quantum nonlocality.

In 1964, Bell proved that any serious version of quantum theory (regardless of whether or not it is based on microscopic realism) must violate locality. He showed that if nature is governed by the predictions of quantum theory, the "locality principle," precluding any sort of instantaneous (or superluminal) action-at-a-distance, is simply wrong, and the world is nonlocal. The theoretical

analysis leading to such a conclusion is commonly known as *Bell's theorem*.

Bell's theorem involves two parts. The first part is the Einstein-Podolsky-Rosen (1935) argument applied to the simplified version considered by David Bohm (1951), the EPRB experiment: a pair of spin one-half particles, prepared in a spin-singlet state, are moving freely in opposite directions. Measurements are made, say by Stern-Gerlach magnets, on selected components of the spins of the two particles. The spin-singlet state has the following property: whenever the component of the spin σ_1 in any direction α is measured for one of the two particles, a measurement of the same component of the spin σ_2 of the other particle will give with certainty the opposite value. For such a state the assumption of locality implies the existence of what are often called noncontextual hidden variables. More precisely, it implies, for the spin-singlet state, the existence of random variables $Z^i_\alpha (= Z_{\alpha \cdot \sigma_i})$, i = 1, 2, which can be regarded as corresponding to preexisting values of all possible spin components of the two particles. In particular, focusing on components in only three directions a, b, and c for each particle, locality implies the existence of six random variables Z^i_α , i = 1, 2, $\alpha = a, b, c$ such that

(1) $$Z^i_\alpha = \pm 1$$

(2) $$Z^1_\alpha = -Z^2_\alpha$$

and, more generally,

(3) $$\text{Prob}(Z^1_\alpha \neq Z^2_\beta) = q_{\alpha\beta},$$

where the $q_{\alpha\beta} = (1 + \alpha \cdot \beta)/2 = \cos^2(\theta/2)$ are the corresponding quantum mechanical probabilities, with θ the angle between α and β.

The argument for this conclusion can be expressed as follows: The existence of such random variables amounts to the idea that measurements of the spin components reveal preexisting values (the Z^i_α). Assuming locality, this is implied by the perfect quantum mechanical anticorrelations:

> Now we make the hypothesis, and it seems one at least worth considering, that if the two measurements are made at places remote from one another the orientation of one magnet does not influence the result obtained with the other. Since we can predict in advance the result of measuring any chosen component of σ_2, by previously measuring the same component of σ_1, it follows that the result of any such measurement must actually be predetermined. (Bell 1964, p. 195; reprinted in Bell 1987, p. 14)

Otherwise, the result would have, at least in part, been produced by the remote measurement, just the sort of influence that Bell's locality hypothesis precludes. One may also note that if the results had not been predetermined, the widely separated correlated residual innovations thereby implied would be an instance of nonlocality.

Observe that, given locality, the existence of such variables is a consequence rather than an assumption of Bell's analysis. In his writing, Bell repeatedly stressed this point (by determinism Bell here means the existence of the preexisting values that would determine the results of the corresponding measurements):

> It is important to note that to the limited degree to which *determinism* plays a role in the EPR argument, it is not assumed but *inferred*. What is held sacred is the principle of "local causality"—or "no action at a distance." …
>
> It is remarkably difficult to get this point across, that determinism is not a *presupposition* of the analysis (1987, p. 143).

and

> Despite my insistence that the determinism was inferred rather than assumed, you might still suspect somehow that it is a preoccupation with determinism that creates the problem. Note well then that the following argument makes no mention whatever of determinism.… Finally you might suspect that the very notion of particle, and particle orbit … has somehow led us astray. … So the following argument will not mention particles, nor indeed fields, nor any other particular picture of what goes on at the microscopic level. Nor will it involve any use of the words "quantum mechanical system," which can have an unfortunate effect on the discussion. The difficulty is not created by any such picture or any such terminology. It is created by the predictions about the correlations in the visible outputs of certain conceivable experimental set-ups (1987, p. 150).

The second part of the analysis, which unfolds the "difficulty … created by the … correlations," involves only very elementary mathematics. Clearly,

$$\text{Prob}(\{Z^1_a = Z^1_b\} \cup \{Z^1_b = Z^1_c\} \cup \{Z^1_c = Z^1_a\}) = 1$$

since at least two of the three (2-valued) variables Z^1_α must have the same value. Hence, by elementary probability theory,

$$\text{Prob}(Z^1_a = Z^1_b) + \text{Prob}(Z^1_b = Z^1_c) + \text{Prob}(Z^1_c = Z^1_a) \geq 1,$$

and using the perfect anticorrelations (2) one has that

(4)

$$\text{Prob}(Z^1_a = -Z^2_b) + \text{Prob}(Z^1_b = -Z^2_c) + \text{Prob}(Z^1_c = -Z^2_a) \geq 1.$$

(4) is equivalent to the celebrated *Bell's inequality*. It is incompatible with (3). For example, when the angles between *a*, *b*, and *c* are 120°, the three relevant quantum correlations $q_{\alpha\beta}$ are all ¼, implying a value of ¾ for the left hand side of (4).

Let P be the hypothesis of the existence of noncontextual hidden variables for the EPRB experiment, that is, of preexisting values Z^i_α for the spin components relevant to this experiment. Then Bell's nonlocality argument, just described, has the following structure:

(5) Part 1: quantum mechanics + locality ⇒ P

(6) Part 2: quantum mechanics ⇒ not P

(7) Conclusion: quantum mechanics ⇒ not locality

For this argument, what is relevant about "quantum mechanics" is merely the predictions concerning experimental outcomes corresponding to (1–3) (with Part 1 using in fact only (2)). To fully grasp the argument it is important to appreciate that the content of P—what it actually expresses, namely the existence of the noncontextual hidden variables—is of little substantive importance for the argument. What is important is the fact that P is incompatible with the predictions of quantum theory.

The content of P is, however, of great historical significance: It is responsible for the misconception that Bell proved that (i) hidden variables are impossible, a belief until recently almost universally shared by physicists, and, more recently, for the view that Bell proved that (ii) hidden variables, while perhaps possible, must be nonlocal. Statement (i) is plainly wrong, since a hidden-variables theory exists and works, as mentioned earlier. Statement (ii) is correct, significant, but nonetheless rather misleading. It follows from (5) and (6) that *any* account of quantum phenomena must be nonlocal, not just any hidden-variables account. Bell's argument shows that nonlocality is implied by the predictions of standard quantum theory itself. Thus, if nature is governed by these predictions, then *nature is nonlocal*. (That nature is so governed, even in the crucial EPRB-correlation experiments, has by now been established by a great many experiments, the most conclusive of which is perhaps that of Alain Aspect, Jean Dalibard, and Gérard Roger [1982].)

Concerning the wrongness of statement (i), some historical facts should be recalled. John von Neumann, one of the greatest mathematicians of the twentieth century, claimed to have mathematically proven that Einstein's dream, of a deterministic completion or reinterpretation of quantum theory (i.e., a hidden-variables theory), was mathematically impossible. Von Neumann's claim was almost universally accepted among physicists and philosophers of science. But Bohmian mechanics is a counterexample, so something had to be wrong with von Neumann's argument. Precisely what was wrong was elucidated by Bell in 1966. Nonetheless, many physicists continued to rely on von Neumann's proof and in recent years more commonly on Bell's inequality to support their rejection of the possibility of hidden variables.

The following is how Bell himself reacted upon learning of Bohmian mechanics:

> But in 1952 I saw the impossible done. It was in papers by David Bohm. Bohm showed explicitly how parameters could indeed be introduced, into nonrelativistic wave mechanics, with the help of which the indeterministic description could be transformed into a deterministic one. More importantly, in my opinion, the subjectivity of the orthodox version, the necessary reference to the "observer," could be eliminated. …
>
> But why then had Born not told me of this "pilot wave"? If only to point out what was wrong with it? Why did von Neumann not consider it? More extraordinarily, why did people go on producing "impossibility" proofs, after 1952, and as recently as 1978? … Why is the pilot wave picture ignored in textbooks? Should it not be taught, not as the only way, but as an antidote to the prevailing complacency? To show us that vagueness, subjectivity, and indeterminism, are not forced on us by experimental facts, but by deliberate theoretical choice? (1987, p. 160)

In fact, Bell's examination of Bohmian mechanics led him to his nonlocality analysis. In the course of his investigation of Bohmian mechanics, he observed that

> … in this theory an explicit causal mechanism exists whereby the disposition of one piece of

> apparatus affects the results obtained with a distant piece. …
>
> Bohm of course was well aware of these features of his scheme, and has given them much attention. However, it must be stressed that, to the present writer's knowledge, there is no *proof* that *any* hidden variable account of quantum mechanics *must* have this extraordinary character. It would therefore be interesting, perhaps, to pursue some further "impossibility proofs," replacing the arbitrary axioms objected to above by some condition of locality, or of separability of distant systems. (1966, p. 452; reprinted in Bell 1987, p. 11)

In a footnote, Bell added that "Since the completion of this paper such a proof has been found." This proof was presented in his 1964 EPR-nonlocality paper discussed here. (The 1966 paper was in fact written before the 1964 paper, but its publication was delayed.)

Physicists' misconceptions notwithstanding, Bell did not establish the impossibility of a deterministic reformulation of quantum theory, nor did he ever claim to have done so. On the contrary, over the course of several decades, until his untimely death in 1990, Bell was the prime proponent, for a good part of this period almost the sole proponent, of the very theory, Bohmian mechanics, that he is supposed to have demolished.

See also Bohm, David; Bohmian Mechanics; Einstein, Albert; Neumann, John von; Philosophy of Physics; Quantum Mechanics; Realism; Relativity Theory; Schrödinger, Erwin.

Bibliography

Aspect, Alain, Jean Dalibard, and Gérard Roger. "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers." *Physical Review Letters* 49 (25) (1982): 1804–1807.

Bell, John Stewart. "On the Einstein-Podolsky-Rosen Paradox." *Physics* 1 (3) (1964): 195–200. Reprinted in Bell 1987.

Bell, John Stewart. "On the Problem of Hidden Variables in Quantum Mechanics." *Reviews of Modern Physics* 38 (3) (1966): 447–452. Reprinted in Bell 1987.

Bell, John Stewart. *Speakable and Unspeakable in Quantum Mechanics*. New York: Cambridge University Press, 1987.

Bohm, David. *Quantum Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1951.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777–780.

Detlef Dürr, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghì (2005)

BELLARMINE, ST. ROBERT

(1542–1621)

St. Robert Bellarmine, an Italian cardinal and controversialist, was born at Montepulciano in Tuscany and died at Rome. Educated in the Jesuit order, of which he became a member, he taught philosophy and theology at the University of Louvain (1570–1576), then at the Roman (Jesuit) College, where he later served as rector. After Bellarmine was created a cardinal in 1599, much of his time was devoted to the administrative and diplomatic affairs of the Roman Catholic Church, in which he is now venerated as a saint. His chief published work is the *Disputations on Controversial Matters (Disputationes de Controversiis),* in which Book III (*De Laicis*) treats questions of political and social philosophy. Another treatise in political philosophy is the *Defense of His Reply to King James I of England* (*Apologia Bellarmini pro Responsione Sua ad Librum Jacobi Magnae Britanniae Regis,* reprinted in Giacon's *Scritti politici*), concerning the theory of the divine right of kings.

In general, Bellarmine's philosophic thought is Thomistic. His lectures at Louvain covered all of Thomas Aquinas's *Summa Theologiae* and are now preserved in the Vatican Archives, though they have not been printed. As a result, little is known of his metaphysical and psychological views, except for occasional explanations given in his more practical writings. It is assumed that he had a very sound understanding of the speculative thought of Thomas Aquinas, however, and the publication of the Louvaine lectures is a desideratum. In ethics and philosophy of law, Bellarmine is a strong opponent of the view that the source of justice is the will of God; instead, he argues that man's awareness of moral law derives from his understanding of the nature of man and his environment, and that ultimately the command (*imperium*) of God's law is intellectual, stemming from the divine wisdom. Thus, he is opposed to voluntarism and defends intellectualism in morals and jurisprudence.

Bellarmine's political theories developed in part from opposition to King James's claim that both spiritual and temporal power belong to the civil monarch. In defending the autonomy of ecclesiastical authority, Bellarmine strongly supported the distinction and separation of the powers of church and state. In chapter 13 of the *Apologia,* he argued that, though the ultimate source of both powers is divine, the civil power is conferred on rulers, *mediately,* through the people as a medium. Thus, with Francisco Suárez, Bellarmine is one of the most

prominent Catholic advocates of the "translation theory" of political sovereignty.

Bellarmine was firmly convinced of the importance of the individual citizen and the dignity of every person. His social and political thinking is reminiscent of the fourteenth-century views of Marsilius of Padua. There is a possibility that Bellarmine's arguments influenced British antimonarchist thinking and, through John Locke, the founders of American democracy. He also recognized something of the investment value of money and helped to modify the older Catholic theory that all taking of interest on loans was to be condemned as usury. In a treatise on the power of the pope (*De Summo Pontifice,* I, 9), Bellarmine favored the idea of a world state but admitted that a plurality of national states regulated by international law might be more practical.

About Bellarmine's role in the prosecution of Galileo Galilei it is hard to be precise; in 1616 he seems to have warned Galileo to discuss the Copernican theory merely as a "mathematical supposition," but he almost certainly did not enjoin him from "teaching or discussing Copernicanism in any way," as was charged after Bellarmine's death. Galileo's publication of the *Dialogue of the Two Chief World Systems,* in 1632, caused him to be prosecuted for heresy on the grounds that he had thereby violated the supposed stricter warning.

See also Thomism.

Bibliography

WORKS BY BELLARMINE

Disputationes de Controversiis Christianae Fidei Adversus Huius Temporis Haereticos. 3 vols. Ingolstadt: Ex Officina Typographica Davidis Sartorii, 1586–1593. Venice, 1596 (4 vols.). The Venice edition is the definitive Latin text.

Opera Omnia, edited by J. Fèvre. 12 vols. Paris, 1870–1874.

De Laicis or the Treatise on Civil Government. Translated by K. E. Murphy. New York: Fordham University Press, 1928.

Scritti politici, edited by C. Giacon. Bologna: Zanichelli, 1950. Selected politico-social writings in Latin.

WORKS ON BELLARMINE

Brodrick, J. *Robert Bellarmine: Saint and Scholar.* Westminster, MD: Newman Press, 1961.

Davitt, T. *The Nature of Law,* 195–218. St. Louis: B. Herder, 1951.

Kuntz, Paul G. "The Hierarchical Vision of St. Roberto Bellarmine." In *Jacob's Ladder and the Tree of Life,* edited by Marion Leathers. New York: Peter Lang, 1987.

Pera, Marcello. "The God of Theologians and the God of Astronomers: An Apology of Bellarmine." In *The Cambridge Companion to Galileo,* edited by Peter Machamer. New York: Cambridge University Press, 1998.

Riedl, J. "Bellarmine and the Dignity of Man." In *Jesuit Thinkers of the Renaissance,* edited by G. Smith, 193–226. Milwaukee: Marquette University Press, 1939.

Springborg, Patricia. "Thomas Hobbes and Cardinal Bellarmine: Leviathan and the 'Ghost of the Roman Empire.'" *History of Political Thought* 16 (4) (1995): 503–531.

Vernon J. Bourke (1967)
Bibliography updated by Tamra Frei (2005)

BENEKE, FRIEDRICH EDUARD

(1798–1854)

Friedrich Eduard Beneke, the German philosopher and psychologist, was born in Berlin and after his gymnasium education studied theology and philosophy, first at Halle and then at Berlin. He became university lecturer (*Privatdozent*) at the University of Berlin in 1820 and, despite Georg Wilhelm Friedrich Hegel's power and official connections, managed to have a considerable number of students.

His first books were *Erkenntnislehre nach dem Bewusstsein der reinen Vernunft* (Theory of knowledge according to the consciousness of pure reason) and *Erfahrungsseelenlehre als Grundlage alles Wissens* (Experiential theory of the soul as foundation of all knowledge). Both were published in Jena in 1820. Two years later, he published in Berlin *Grundlegung zur Physik der Sitten* (Foundations of the physics of morals), a work that found disfavor among the entrenched Absolute Idealists and resulted in his being forbidden to lecture. Beneke was accused of Epicureanism, although the objections given by Minister von Altenstein, a Hegelian who opposed Beneke's attempted application of science to ethics, were that the book was not so much wrong on particular points as that it was *unphilosophisch* in its totality because it did not attempt to derive everything from the Absolute. Beneke's anti-Hegelian position led to further difficulties. An offer of a position at the University of Jena was overruled by the authorities in Berlin, who managed to find a state law to support this move. Beneke moved to Göttingen, where his reception was more cordial, and remained there until 1827, when he received permission to resume his lectures in Berlin. After Hegel's death, Beneke managed to advance to the rank of "extraordinary professor." Although he was active in teaching and writing, his later years were plagued by illness. In 1854, under unexplained circumstances, his body was found in a Berlin canal.

Along with Johann Friedrich Herbart and some others, Beneke represented a reaction against the Fichte-Schelling-Hegel phase of German philosophy. He insisted that psychology, which ought to be established inductively, is the necessary presupposition of all disciplines in philosophy. Logic, ethics, metaphysics, and especially the philosophy of religion should be based on it. Beneke's psychology is a form of associationism, and shows the influence of both Immanuel Kant and the British empiricists, especially John Locke, whose disciple Beneke claimed to be. The senses give us only a mediated knowledge of the external world and of ourselves. Nevertheless, we can obtain an immediate, fully adequate knowledge of our own mental acts by means of inner perception. Starting from this perception, we infer the inner nature of other beings by analogy with our own. The result of this inference is a picture of reality as containing an uninterrupted series of minds or "faculties of representation" (*Vorstellungsfähigkeit*), extending downward from man. The soul consists of a system of powers or forces; it is a "bundle" but, contrary to Hume, not a bundle of perceptions.

Beneke used the language of faculty psychology, although he did not intend "powers" or "faculties" to be viewed as hypostatized concepts. All psychological processes, he claimed, can be traced back to four basic ones: (1) the process of stimulus appropriation (*Reizaneignung*), in which the mind creates sensations and perceptions out of externally caused impressions; (2) the process of formation of new "elementary faculties" (*Urvermögen*) by means of the assimilation of received stimuli; (3) the process of transmission (*Übertragung*) and equalization (*Ausgleichung*) of stimuli and powers, whereby a systematic connection is formed between our becoming conscious of one idea and our becoming unconscious of another idea; (4) the process of mutual attraction and "blending" (*Verschmelzung*) of ideas of the same sort.

Beneke's attempt to explain the mind's activities in terms of their genesis is reminiscent of Herbart. Unlike the latter, however, he assumed that philosophy must proceed from what is immediately given in consciousness. We have no alternative to this starting with inner experience, he believed, because our own soul is the only thing that we know as it is in itself. We recognize it as a nonspatial and therefore an immaterial entity. At least we have no reason to suppose it to be material, since it is not perceived through outer sense. The soul, however, cannot be simple, as Herbart had maintained. It has, as we have noted, specific powers or capacities for receiving and organizing stimuli; these powers must be underivative, since stimuli of different kinds can be received even at the outset of our experience. Each of our senses is supposed to include several of these *Urvermögen.* But the soul must also be capable of forming new *Urvermögen,* in order to be receptive to new sorts of stimuli.

Beneke thus conceived the mental life as compounded of active impulses (*Triebe*) that are activated by external stimuli. The seemingly substantial unity of mind is explained by the persistence of traces (*Spuren*) of ideas that have become unconscious and by the mutual adjustment of faculties that produce new impulses.

See also Empiricism; Epicureanism and the Epicurean School; Hegel, Georg Wilhelm Friedrich; Herbart, Johann Friedrich; Hume, David; Kant, Immanuel; Locke, John; Psychologism; Psychology.

Bibliography

ADDITIONAL WORKS BY BENEKE

Neue Grundlegung zur Metaphysik. Berlin, 1822.

Psychologische Skizzen, 2 vols. Göttingen: Vandenhoeck and Ruprecht, 1825–1827.

Das Verhältniss von Seele und Leib. Göttingen: Vandenhoeck and Ruprecht, 1826.

Kant und die Philosophische Aufgabe unserer Zeit. Berlin, 1832.

Lehrbuch der Psychologie als Naturwissenschaft. Berlin, 1833.

Die Philosophie in ihrem Verhältnisse zur Erfahrung zur Spekulation und zum Leben. Berlin, 1833.

Grundlinien des näturlichen Systems der praktischen Philosophie. Berlin, 1837. Beneke regarded the last part of the *Grundlinien,* which contains his theory of morals, as his best work.

Metaphysik und Philosophie der Religion. Berlin, 1840.

WORKS ON BENEKE

Benner, H. *Benekes Erkenntnistheorie.* Halle: Druck von Wischan and Wettengel, 1902.

Gargano, V. *L'Etica di Beneke.* Catania, Sicily, 1912.

Gramzow, O. *Benekes Leben und Philosophie.* Bern: Buchdr. Steiger, 1899.

Murtfeld, R. "Vergeblischer Kampf gegen den Idealismus: Friedrich Eduard Beneke." In *Zeitschrift für Geschichte der Erziehung und des Unterrichts,* pp. 1–48. Berlin: Weidmannsche, 1923.

Samuel, E. *Die Realität des Psychischen bei Beneke.* Berlin, 1907.

Wandschneider, A. *Die Metaphysik Benekes.* Berlin, 1903.

Arnulf Zweig (1967)

BEN GERSHON, LEVI

See *Gersonides*

BENJAMIN, WALTER
(1892–1940)

Walter Benjamin, philosopher, literary and social critic, and aesthetic theorist of the modernist period, was born to a liberal, middle-class Jewish family on July 15, 1892, in Berlin. He died in 1940 by suicide, having failed to cross the border from France to Spain. Many of his writings were published posthumously. He lived mainly in Germany, but spent his last years in exile in Paris barely surviving as an independent writer. Although close friends such as Gershom Scholem and Theodor W. Adorno encouraged him to move to Israel or to New York, he chose fatefully to stay in Europe.

Benjamin had a seminal impact, especially after the 1960s, on critical theory, art history, and aesthetics; on political philosophy and the philosophy of language and history (in the continental vein); on linguistics, literature, and criticism; on communications, technology, and mass media; even, later, on anthropology, cultural studies, and postcolonial and feminist theory.

Benjamin developed his central concept of *critique* from his extensive reading in philosophy, poetry, and literature, especially of Immanuel Kant, Johann von Goethe, Friedrich Hölderlin, and Friedrich Schlegel. Critique was a concept or, better, a philosophical approach to establishing the parameters of knowledge and experience (*Erfahrung*). Lifelong, Benjamin attempted to move beyond the limits that he saw the neo-Kantians to have imposed far more dogmatically than Kant himself. He saw in the Enlightenment concept of experience the gradual movement toward "scientism" and with this toward an ever more severe limitation and impoverishment of its promise. Experience, he argued, ought not to be reduced to the "object realm" of science.

His earliest work on educational reform was influenced by Gustav Wyneken's Youth Movement. Again, as that movement became more dogmatic, the more Benjamin distanced himself from it. He resisted partisan thinking all his life, given his unwillingness to compromise either "the life of the spirit" or the claims of the early Karl Marx's historical materialism. Similarly, though instructed by well-known professors, he showed himself to be as anti-academic and anti-programmatic as he was anti-partisan. He was wary of any well-trodden path or anything that smacked of matter-of-factness.

Benjamin became not only a philosophical thinker but also a writer who would sharply oppose those who aimed in thought and language simply to stipulate the principles of method. He wondered how a writer could release the truth in a world that acts as if it would rather have a "higher," "absolute," or "certain" truth imposed upon it. He thought about how one writes "against the grain" or how one writes oneself out of restrictions by which one, as a writer, is historically and socially conditioned. He wrote against the dominant positivist idea or myth of progress which, he maintained, far more concealed than brought truth to appearance.

Benjamin was wary of traditional forms of philosophical argument. He used literary and visual images to develop a language he regarded as more appropriate or truthful for modern times. He wrote sometimes in fragments, sometimes with quotations or aphorisms, in part to demonstrate his interest in what he called constellations or dialectical images. He experimented with both the long and the short form of the (literary) essay. He was particularly interested in story-telling as still historically able to transmit genuine experience.

In his "On Language as Such and on the Language of Mankind" (1916), he argued against the idea that writing was either transparent or merely a vehicle for the communication of an independently existing meaning. Meaning was, rather, contained and usually concealed within language, a view that necessitated entirely rethinking the task of translation. Given a secularized Messianic myth of the fall of humanity and of humanity's entry into history, Benjamin maintained that the more a society misuses its language the more the language (like society) falls into decay. The aim of critique was double-sided: to describe language's decay or the loss of meaning under present social or historical conditions at the same time that one seeks to bring that meaning back to presence. Critique as retrieval was no straightforward matter: It demanded different modes of extreme and explosive, but also fragile and experimental, thought.

Although influenced both by classicism and early romanticism, he explored in modernist terms the complex relations between the truth and deception of language, sign, and image. Between 1919 and 1920 he completed his doctorate in Switzerland with "The Concept of Criticism in German Romanticism." In 1928 the University of Frankfurt refused him his *Habilitation* for his *Origin of German Tragic Drama*. This dissertation, written largely through quotation and focused on a distinction between allegory and symbol, explored the modern form of tragedy. Benjamin often described modernity in terms of ruins: to modernity he liked to attach the terms of meaninglessness, mortification, and fragmentation. Allegory, as one critic has put it, was "a poetic response to the degradation that language undergoes

in the instrumental conception that modernity gives to it" (Rochlitz 1996, p. 99). However, though Benjamin so described modernity, he did not engage merely in a conservative lament about how the world once was. Instead, through allegory, he sought the redemptory, and at times also the revolutionary, promise of the new languages, images, and cultural forms as given from the temporal perspective of the "here and now." Benjamin's work on allegory later proved most influential on the thinking of Jacques Derrida and Paul de Man.

Benjamin refused to treat works of art, literature, or philosophy as if one were attending a funeral, as if the works were situated merely in the monumentalized context of a dead past. As he argued in his "Theses" on the philosophy of history, so he argued in his work on literary criticism, that the critic should aim to keep works alive by showing how their meaning, described as belonging to the past, was still present or available to us albeit in enigmatic or allegorical form.

In his writing on history, he argued against the dominant teleological, progressivist and perfectibility visions which saw the world as ordered, rational, and, purposive. He rather presented a view of the past as radically fragmented and of history as something that narrates a story far less of victory and inclusion and far more of failure and exclusion. Inspired by a painting of Paul Klee, he pictured the historian as an angel (the "angel of history") who, though propelled by progressive forces toward the future, would prefer rather to look backward. He described the historian as a guardian of the past, as one who desires to subvert future catastrophe by awakening the dead in an attempt to make whole again what has been destroyed.

Comparably, in his literary criticism, he argued that meaning does not reside in works as if fixed, saturated, or completed; it exists rather as possibility or as a suggestion still flickering in the flames of the coherence the world once had. To retrieve the meaning present in a work is to retrieve that which critics of antiquarian tendency have allowed to fall into oblivion. For Benjamin, art cannot do without this act of retrieval, as he demonstrated in his early mammoth essay on Goethe's *Elective Affinities*. Here, Benjamin distinguished critique from commentary. Whereas the latter focuses on *material content*, the realities more immediately apparent to the eye, the former focuses on the concealed, but historically gradually to be revealed, *ideal- or truth-content* of the work; his point was always to demonstrate the intricate relation between the two.

Increasingly influenced by Baudelaire, Benjamin exposed the contradictory or antagonistic structures of modernist, urban, bourgeois or capitalist life in the metropolis: Berlin, Moscow, Paris. In his late and unfinished *Arcades Project*, he traced the allegorical significance of advertising slogans and neon signs attached to the aging architectural structures of the Parisian Arcades. He looked at the postures of prostitutes, mannequins, and gamblers, and at the movements of the trains and the stock exchange. He looked at the speed of pedestrian traffic and at the exhibition in the shoppers of their boredom, idleness, desire, and satisfaction. Stamps, toys, newspaper headlines revealed the city in its smallest details. No detail and no commodity were treated as trivial or insignificant.

Influenced by the Parisian surrealists, he described fantasies and dreams as collective forms of social experience; he experimented with opium to gain access to new forms of experience. He wrote about dialectical images, which, while structured by capitalist relations of production, nonetheless contained a redemptory potential that would appear to the viewer in momentary or disoriented experiences. He investigated the psychical processes (influenced by art, writing, and drugs) that would crack habitual forms of life or break through the apparent fixity of social forms. To interpret the world was to reorder (change) the world through profane illumination. With André Breton, it was to release the world from its chains or to allow the uncanny dimensions of experience, suppressed under the social construction of ordered appearance, to emerge.

Benjamin's work is often distinguished by earlier and later periods, by decisive transitions from his more esoteric and elitist interest, inspired by Jewish Messianic thought, to his late, often Bertold Brecht–inspired, more revolutionary work in (Marxist) historical materialism. As, however, his last writings on history and art show, there are significant continuities across these transitions. In perhaps his best-known essay, "The Work of Art in the Age of its Technological Reproducibility," he argued that how we receive or view art is changed not only by radical alteration in conditions of production, but also, more esoterically, by how the art, in its experimentation, may surrealise (transform or shatter through creative disorientation) how we think and feel. He showed how the mediation in art between aesthetic concepts and technological conditions matters both for the sake of art and for that of politics.

Crudely, to speak of the "aestheticization of the political" was to speak with the Fascists or the totalitarian

thinkers, of how new forms of technology were being regressively employed to sustain outdated aesthetic concepts. He described how concepts of aura and aesthetic absorption had come increasingly to be employed to produce political rather than merely artistic forms, hence, the use of aesthetic concepts and artistic techniques in modern war and propaganda. Contrarily, to speak of the "politicization of art" encouraged a production of art that would more truthfully adapt to currently existing conditions, a production that would rather help liberate a people, so Benjamin argued at his most committed revolutionary moment, than be used to promote knowingly false political illusions.

Benjamin juxtaposed his concrete and diverse reflections on mass art, film, photography, epic theater, and spectatorship with reflections on violence, war, and militarism. In turn, his reflections on violence and fragility were inseparable from his own thinking and writing on the modern condition and possibility of experience.

See also Adorno, Theodor Wiesengrund; Aesthetics, History of; Critical Theory; Derrida, Jacques; Enlightenment; Goethe, Johann Wolfgang von; Historical Materialism; Hölderlin, Johann Christian Friedrich; Kant, Immanuel; Marx, Karl; Neo-Kantianism; Political Philosophy, History of; Schlegel, Friedrich von.

Bibliography

WORKS BY WALTER BENJAMIN

Illuminations. Translated by Harry Zohn; edited and with an introduction by Hannah Arendt. New York: Schocken Books, 1969.

The Origin of German Tragic Drama. Translated by John Osborne. Introduction by George Steiner. London: New Left Books, 1977.

Charles Baudelaire: A Lyric Poet in the Era of High Capitalism. Translated by Harry Zohn. London: New Left Books (Verso), 1983.

Understanding Brecht. Translated by A. Bostock. London: New Left Books (Verso), 1983.

The Arcades Project. Translated by Howard Eiland and Kevin McLaughlin. Cambridge, MA: Harvard University Press, 1999.

Selected Writings. 4 vols edited by Michael W. Jennings and Howard Eiland. Cambridge, MA: Harvard University Press, 2002.

WORKS ABOUT WALTER BENJAMIN

Benjamin, Andrew, ed. *Walter Benjamin and Art*. London and New York: Continuum, 2005.

Bolz, Norbert, and Willem Van Reijen. *Walter Benjamin*. Translated by L. Mazzarins. Atlantic Highlands, NJ: Humanities Press, 1996.

Buck-Morss, Susan. *The Dialectics of Seeing: Walter Benjamin and the Arcades Project*. Cambridge, MA: MIT Press, 1989.

Hanssen, Beatrice, and Andrew Benjamin, eds. *Walter Benjamin and Romanticism*. New York and London: Continuum, 2002.

Jennings, Michael W. *Dialectical Images: Walter Benjamin's Theory of Literary Criticism*. Ithaca, NY, and London: Cornell University Press, 1987.

McCole, John. *Walter Benjamin and the Antinomies of Tradition*. Ithaca, NY, and London: Cornell University Press, 1993.

Richter, Gerhard, ed. *Benjamin's Ghosts: Interventions in Contemporary Literary and Cultural Theory*. Stanford, CA: Stanford University Press, 2002.

Rochlitz, Rainer. *The Disenchantment of Art: The Philosophy of Walter Benjamin*. Translated by Jane Marie Todd. New York and London: The Guilford Press, 1996.

Smith, Gary, ed. *Benjamin: Philosophy, Aesthetics, History*. Chicago: University of Chicago Press, 1989.

Smith, Gary, ed. *On Walter Benjamin: Critical Essays and Recollections*. Cambridge, MA: MIT Press, 1991.

The Walter Benjamin Research Syndicate. Available from http://www.wbenjamin.org.

Wolin, Richard. *Walter Benjamin: An Aesthetic of Redemption*. Berkeley, CA: University of California Press, 1994.

Lydia Goehr (2005)

BENN, GOTTFRIED
(1886–1956)

Gottfried Benn, the German poet and critic, was born in Mansfeld in Westprignitz, of mixed Prussian and Swiss-French parentage. After studying philosophy and philology at the universities of Marburg and Berlin, he received a military scholarship to the Kaiser Wilhelm Academy of Berlin, from which he was graduated as doctor of medicine in 1912. Commissioned as a medical officer in the German Imperial Army, he served briefly in 1912 and then again after the outbreak of the war in 1914. A close friendship with the poet Else Lasker-Schüler ended in 1913, and in July 1914 he married the actress Eva Brandt. From 1917 to 1935 he practiced in Berlin as a specialist in venereal and skin diseases. After his wife's sudden death in 1922, he befriended Ellen Overgaard, a Danish woman, who adopted his daughter.

Benn collaborated with Paul Hindemith on the oratorio *Das Unaufhörliche,* which was performed in 1931. Extensive contact with representative writers of the Weimar Republic led to his election, in 1932, into the German Academy of Arts (whose president, Heinrich Mann, the brother of Thomas, Benn had eulogized in an essay in 1931). A somewhat sordid period of jockeying for positions in the new Reich ended in 1935 with Benn's los-

ing the post of municipal medical specialist, and in 1938 all his writings were banned. He rejoined the army in 1935, coining for this move the much-publicized term *innere Emigration,* in contrast to the actual emigration of his former friends. In 1938 he married his secretary, Herte von Wedemeyer; she committed suicide in 1945, when the Russian armies were approaching the village to which she had been evacuated. After the war Benn's writings were banned, but the publication of *Statische Gedichte* in Switzerland (1948) marked the beginning of a new creative phase. In 1946 he married Ilse Kaul, a young dentist. Benn gave up his medical practice in 1953. Through his decision to remain in Berlin, he became something of a spokesman for the intelligentsia of the city. At his death he was hailed as the greatest German poet since Rainer Maria Rilke; his influence on the styles and themes of contemporary German poetry, certainly, is second to none.

Benn always insisted on the hermetic nature of his poetry and prose; nevertheless, his work faithfully reflects both the historical events and the intellectual turmoil of his age. His first collection of poems, *Morgue* (1912), achieved notoriety and success because of its ruthless exploitation of the phenomena of physical decay and disease. The stark naturalism of such a poem as "Man and Woman Walking through a Cancer Ward" lies both in its rhythmically weak form and in the direction of its argument, typical of much of Benn's later work: the poem attempts to designate some bedrock of "reality" that will withstand contemporary skepticism. The "reality" that emerges from behind the clinical details is a representation of life as impersonal, merely physical or biological, and bereft of all spirit.

The major German poets of the twentieth century have expressed an acute consciousness of their historical situation, a consciousness that derives from Friedrich Nietzsche's critique of the historical imagination and from Oswald Spengler's *Decline of the West.* Benn, in the wake of these works, described the age after the defeat of 1918 as "postnihilistic." In the face of national collapse he set out to formulate an "absolute aesthetic," the aim of which was to "transcend" the actual situation by means of the idea of a "pure poem," the poem of "absolute expressiveness" (as opposed to the poem of communication or opinion with didactic intent). In Benn's poetry, however, there are elements of self-disclosure that seem not to be consistent with his concept of the "pure poem." And his doctrine that art should be exclusively concerned with "style, not truth," raises more questions than it answers.

Benn's ideas on the role of art in life varied. He was able to speak of art as "historically ineffective, without practical consequences," but also to define it (in the wake of Nietzsche) as "the only valid vindiction of life." The "biologism" of Benn's earlier poetry had been morally indifferent, and he had nothing but contempt for every form of social organization and democratic politics, especially those of the Weimar Republic. It is therefore not surprising that after March 1933 he emerged as the most important of those German poets who convinced themselves that national socialism offered an answer to their search for a valid artistic ideology—or, rather, for valid poetic symbols. Benn discerned in Adolf Hitler's regime the rule of "a new biological type … [and] the victory of the national idea, the victory of genuine human values, in perfect harmony with the logic of history." His courtship with national socialism was brief, yet even in 1950 (in his embarrassing autobiographical apologia, *Doppelleben*) his main criticism of the Hitler regime was that it "lacked style." "Style" was for Benn the product and the justification of an image-making faculty that conforms to certain "absolute" laws; these laws are "autonomous" in the sense of being indifferent to the demands of personal experience and social reality alike. Questions of personal expediency apart, Benn's astonishing expectations for Hitler's regime seem to have sprung from that contemptuous disregard of political realities that had been characteristic of an important section of the German cultural scene for many years. He saw no contradiction in asserting the hermetic nature of poetry while claiming that the heroic virtues of the new regime would be more propitious for its creation. The historicism he cultivated served Benn (as it did Martin Heidegger in 1933) as justification for his collaboration, but it did not lead him to a clear understanding of the total claim of Hitler's dictatorship.

Benn is the only major German poet who felt, albeit briefly, that his vision was realized in the National Socialist ideology, even though his poems soon proved to be incompatible with the party line in art. The elements that form his best poems derive from the cosmopolitan expressionist school that flourished in Germany in the 1920s as much as from French and Italian imagism; even his invocation of chthonic and instinctual values (in his praise of "Quaternary man" and his values) has its parallels in Ezra Pound, T. E. Hulme, and Julian Benda. His poetic style is clipped, paratactic, full of laconic allusions to the natural sciences. Memories are imaged by means of strong and complex sense perceptions; striking physical details are selected, often for their sound values; all mention of "you" and "we" is rhetorical, the solipsistic circle

hardly ever being breached; and the situations invoked are almost always related to a self whose isolation is, if anything, underlined by an appeal to primordial memories.

See also Heidegger, Martin; Hermeticism; Historicism; Nietzsche, Friedrich; Rilke, Rainer Maria (René).

Bibliography

WORKS BY BENN

Gesammelte Werke. Wiesbaden: Limes, 1958–1961. The most important aesthetic statements are to be found in "Züchtung I" and "Züchtung II," Vol. I; "Roman des Phänotyps" and "Der Ptolemäer," Vol. II; "Doppelleben" and "Ausdruckswelt," Vol. III; and "Autobiographische und vermischte Schriften," Vol. IV.

Statische Gedichte. Zürich: Arche, 1948. These two volumes contain representative selections of Benn's poetry.

Trunkene Flut, 2nd ed. Wiesbaden: Limes, 1952.

WORKS ON BENN

Hamburger, M. *Reason and Energy.* London: Routledge and Paul, 1957.

Holthusen, H. E. *Das Schöne und das Wahre.* Munich: R. Piper, 1958.

Jens, W. *Statt einer Literaturgeschichte.* Tübingen, 1958.

Lohner, E. *Gottfried Benn.* Wiesbaden: Limes, 1958. Bibliography.

Loose, G. *Die Ästhetik Gottfried Benns.* Frankfurt, 1961.

Wellershoff, D. *Gottfried Benn: Phänotyp dieser Stunde.* Cologne: Kiepenheuer and Witsch, 1958.

Wodtke, F. W. *Gottfried Benn.* Stuttgart: J. B. Metzler, 1962. Biography.

J. P. Stern (1967)

BENNETT, JONATHAN
(1930–)

Born in 1930 and educated in New Zealand and at the University of Oxford, Jonathan Bennett taught philosophy at the University of Cambridge for twelve years before taking up professorial positions in Canada (at the University of British Columbia) and the United States (at Syracuse University). He is a Fellow of the American Academy of Arts and Sciences and of the British Academy. Now retired, he continues to write from his home on an island near Vancouver, British Columbia.

Bennett's work covers a wide range of issues in analytic philosophy and the history of philosophy, especially the early modern period. His first book, *Rationality* (1964), explored the differences between human intelligence and the intellectual capacities of other animals, and the role of language in these differences. Subsequently influenced by Paul Grice's seminal work on meaning and communicative intentions, he significantly modified his views about such matters in a later book, *Linguistic Behaviour* (1976), which incorporates an account of convention building on but also differs in certain respects from David Lewis's ground-breaking theory.

Bennett's interest in the nature of intentional behavior connects his work in philosophical psychology and the philosophy of language with his work in the metaphysics of actions and events. His major contribution to the latter topic is *Events and their Names* (1988), in which he explores the distinction between events and facts through an examination of the semantics of everyday language, focusing on the differences between two kinds of sentence nominals, exemplified by the pair *Quisling's betrayal of Norway/Quisling's betraying Norway.* In this book Bennett addresses the important question of whether facts or events should properly be regarded as the things related by causal relations; he contends that both may be but that fact-causation statements and event-causation statements require different kinds of analysis, whether in terms of counterfactual conditionals or in terms of causal laws. Bennett concludes, however, that the language of event-causation, though useful, is impoverished compared with that of fact-causation and that the former must be analysed in terms of the latter. He also offers an analysis of the "by" locution employed in action sentences of the form S *did such-and-such* by *doing so-and-so.*

In a later book on the theme of agency, *The Act Itself* (1995), Bennett discusses in depth the moral dimension of human action, including the thorny question of whether a morally significant distinction can be drawn between doing something and letting something happen: for example, between killing someone and letting someone die. He makes it clear, early in the book, that he is a moral nonrealist, denying that moral judgements are answerable to independent moral facts and hence denying that they have, in that sense, truth values.

Closely connected with Bennett's work on actions and events is his important contribution, over a period of more than thirty years, to philosophical debate on the semantics of conditional statements. This work culminated in his book *A Philosophical Guide to Conditionals* (2003), perhaps the most comprehensive and authoritative treatment of the subject available. On a number of key issues in this debate, Bennett has shifted his position over the years, notably on the question of whether there

is a significant distinction to be drawn between counterfactual and indicative conditionals. Reversing his earlier opinion, formed under the influence of the work of V. H. Dudman, he now thinks that there is and that these two classes of conditionals demand radically different analyses: the former a possible-worlds analysis along the lines proposed by David Lewis and the latter a probabilistic analysis of the sort pioneered by Ernest Adams. As a consequence, he holds that indicative conditionals, unlike counterfactuals, lack truth conditions and hence truth values. At the same time, he tries to explain why, despite their radically different analyses, there are close similarities between the logics of the two kinds of conditionals and why it is often correct to move from asserting an indicative conditional at one time to asserting a corresponding counterfactual at a later time.

Bennett's work in the history of philosophy has centred on the core texts of the British Empiricists—Locke, Berkeley and Hume—and those of certain eminent continental philosophers of the seventeenth and eighteenth centuries, especially Spinoza and Kant. *Kant's Analytic* (1966) was followed eight years later by its sequel, *Kant's Dialectic* (1974), with *Locke, Berkeley, Hume: Central Themes* (1971) appearing in between. Bennett's next major project of this kind was *A Study of Spinoza's Ethics* (1984); at about the same time he collaborated with Peter Remnant to produce an important new edition and translation of *Leibniz's New Essays on Human Understanding* (1981).

The culminating synthesis of Bennett's thoughts about the major philosophers of the early modern period is provided by his magisterial two-volume magnum opus, *Learning from Six Philosophers* (2001). The first volume treats Descartes, Spinoza, and Leibniz and the second Locke, Berkeley, and Hume. Bennett has always been clear about his own approach to the writings of the great philosophers of the past: although he does not ignore their historical context, he is concerned chiefly with the ideas and arguments to be found in them—not merely as illustrative of the philosophical thought of their times, but for their own sake and for the light that they can shed on present-day philosophical debate. Inevitably, this sort of approach has attracted criticism from certain quarters, especially from historians of philosophy who are skeptical about the very notion of *philosophia perennis*—the idea that there are perennial philosophical problems and arguments that transcend cultural and historical boundaries. But whatever the rights and wrongs of this dispute might be, it is manifest that Bennett's approach is motivated not least by his concern, as a teacher of philosophy, to keep the seminal texts of past philosophers alive for succeeding generations of students.

See also Berkeley, George; Conditionals; Counterfactuals; Descartes, René; Empiricism; Event Theory; Grice, Herbert Paul; History and Historiography of Philosophy; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lewis, David; Locke, John; Ontology; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY BENNETT

Rationality: An Essay Towards Analysis. London: Routledge & Kegan Paul, 1964.

Kant's Analytic. Cambridge, U.K.: Cambridge University Press, 1966.

Locke, Berkeley, Hume: Central Themes. Oxford: Clarendon Press, 1971.

Kant's Dialectic. Cambridge, U.K.: Cambridge University Press, 1974.

Linguistic Behaviour. Cambridge: Cambridge University Press, 1976.

A Study of Spinoza's Ethics. Cambridge, U.K.: Cambridge University Press, 1984.

Events and their Names. Oxford: Clarendon Press, 1988.

The Act Itself. Oxford: Clarendon Press, 1995.

Learning from Six Philosophers. Oxford: Clarendon Press, 2001.

A Philosophical Guide to Conditionals. Oxford: Clarendon Press, 2003.

E. J. Lowe (2005)

BENTHAM, JEREMY
(1748–1832)

Jeremy Bentham, English philosopher and reformer, was born in Houndsditch, London, on February 15, 1748. His father was a solicitor, with wealthy and important clients in the City of London. Of his siblings, only one younger brother, Samuel (1757–1831), survived into adulthood, becoming a prominent naval architect and engineer. His mother died on January 6, 1759. In 1760 his father entered him, at the age of twelve, into the University of Oxford, where he attended the lectures of William Blackstone (later published as *Commentaries on the Laws of England*, 1765–1769). He graduated in 1764, having been obliged to subscribe to the Thirty-nine Articles of the Church of England, the statement of its dogma and discipline.

Having entered Lincoln's Inn in 1763, he was admitted to the bar in 1769. He did not, as his father wished, practice law, but decided instead to devote himself to its reform. Bentham thought of himself as "the Newton of

legislation" (Milne 1981, p. 169); just as Isaac Newton (1642–1727) had brought order to the physical sciences, so would he to the moral sciences. Bentham adopted the principle of utility (an action was judged to be morally right to the extent that that it promoted the greatest happiness of the greatest number) as a critical standard by which to test the value of existing practices, laws, and institutions, and to suggest reform and improvement. He set about composing a comprehensive penal code, to which his best-known work, *An Introduction to the Principles of Morals and Legislation* (abbreviated as *IPML*), which was printed in 1780 and published in 1789, was intended to form a preface.

After returning from a visit to his brother in Russia from 1785 to 1788, his career was dominated by his attempt to build a panopticon prison in London. When the scheme effectively collapsed in 1803 Bentham was left embittered by what he regarded as the bad faith of successive ministries, and he became increasingly committed to political radicalism. In 1809 he began to write on parliamentary reform, and in 1822 he embarked on *Constitutional Code*, in which he advocated the establishment of a representative democracy. Having lived in Lincoln's Inn from 1769 to 1792, he had then inherited his father's home in Queen's Square Place, Westminster, where he died on June 6, 1832.

Bentham's contemporary reputation was founded on the five recensions of his works produced in elegant French between 1802 and 1828 by his Genevan translator and editor, Étienne Dumont (1759–1829). Bentham met Dumont in or around 1788, when both were members of the Bowood Circle that gathered at the country house of William Petty (1737–1805), second Earl of Shelburne and first Marquis of Lansdowne. Dumont's recensions were not literal translations of Bentham's writings, but lucid distillations of his central ideas. The first and most influential was *Traités de législation civile et pénale* (The Theory of Legislation; 1802). To those who wished to introduce political and legal reform, but who faced resistance from entrenched interests such as the privileged nobility and the church, the rational, secular, reforming programme offered by Bentham carried great appeal. While profoundly critical of the legal institutions and practices that he found in existence, he was at the same time optimistic about what could be achieved by law. As he had announced in *IPML*, his enterprise was "to rear the fabric of felicity by the hands of reason and of law" (Burns 1970. p. 11). Bentham's vision of the law as an instrument of reform and improvement had considerable impact in an age that viewed ignorance, prejudice, and superstition as the main barrier to human progress.

BENTHAM'S ACHIEVEMENTS

Bentham's achievements, only some of which are noticed in detail here, were immense. He was the founder of classical utilitarianism, which inspired the movement known as philosophic radicalism in which the young John Stuart Mill (1806–1873) played a leading role, and which has remained one of the most influential doctrines in political philosophy. His method of calculating the potential utility of actions forms the basis of cost benefit analysis in economics. Distinguishing sharply between law as it is and law as it ought to be, he inspired the proponents of the doctrine of legal positivism. In his extensive and detailed writings on judicial procedure, he produced the most comprehensive theory of evidence in the Anglo-American tradition. He developed a theory of punishment and reward which emphasized deterrence, proportionality, and rehabilitation of the offender, and which went far beyond, in terms of rigor and coherence, that associated with Cesare Beccaria (1738–1794).

In politics he produced, in 1789, the earliest utilitarian defense of political equality (at one point even advocating women's suffrage), and later, in *Constitutional Code*, produced a sophisticated and detailed blueprint for representative democracy. His essay on *Political Tactics* was the first systematic treatise on the organization of a political assembly. He put forward a scheme to promote peace between nations, advocating an international court of arbitration and a proportional reduction of armed forces. Indeed, the word "international" was coined by Bentham. His proposals for dealing with poverty provided the intellectual basis for the Poor Law Amendment Act of 1834, and for the welfare state more generally. His educational ideas, based on "useful learning" and access for all regardless of religion or gender (in contrast to the Universities of Oxford and Cambridge, where students had to be Anglican and male) inspired the founders of the University of London in the mid-1820s.

LANGUAGE

The starting point for Bentham's thought was his understanding of the way in which the human mind perceived the physical world, and the way in which language was used to describe that world. The fundamental distinction in language was between the names of real entities, which represented objects existing in the physical world (e.g., an apple), and the names of fictitious entities, objects that were spoken about as if they did exist, and about which it

made sense to talk as though they existed, but to which it was not intended to ascribe physical existence (e.g., the property of a physical object, such as the sweetness of an apple, or an abstraction, such as a law). In order to make sense, language had to refer, either directly or indirectly, to physical objects. The difficulty lay in finding a method by which the names of fictitious entities could be related to their "real source" in the physical world. The names of fictitious entities were not capable of exposition by means of representation, where a specific object was produced and its assigned name pronounced, for there was no such object to produce. Nor was it possible to define a fictitious entity by means of the Aristotelian method of definition *per genus et differentiam*. Definition by this means was possible where the object belonged to a nest of aggregates, and was not the highest object in the nest, but was not possible where the word had no superior genus.

Bentham's solution consisted in the complementary techniques of paraphrasis and phraseoplerosis. The operation of phraseoplerosis, the filling up of the phrase, was logically prior to that of paraphrasis. Discourse often contained ellipses, which needed to be "filled in" by inserting the omitted words. Thereupon, the operation of paraphrasis could be undertaken, whereby a sentence in which the name of the fictitious entity appeared was translated into another sentence in which the words were either real entities, or were more nearly related to real entities. Take the word "duty." A person (X) had a legal duty when someone else (Y) had a right to have him (X) made to perform it, in which case X had a duty toward Y, and Y a right against X; what Y had a legal right to have X be made to do was that for which X was legally liable, upon a requisition made on Y's behalf, to be punished for not doing. The definition or exposition had "resolved" the notion of duty into its simple, or more simple, elements: namely the prospect of suffering a punishment (a term which itself would require further exposition), upon the forbearance to perform some action, when required to do so by the person invested with the corresponding right. However, if an exposition by paraphrasis proved to be impossible, then the fictitious entity in question belonged to the class of nonentities, the noun substantive by which it was represented was merely a sound, and any proposition in which it occurred was nonsensical.

PRINCIPLE OF UTILITY

Bentham's critical standard, the principle of utility, was a fictitious entity, and had to be expounded by relating it to the physical entities that formed its "real source." As Bentham explained in *IPML*, the "real source" in question consisted in the sensations of pain and pleasure: "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do." The "sovereign masters" of pain and pleasure not only accounted for human motivation, "govern[ing] us in all we do, in all we say, in all we think," but also provided "the standard of right and wrong" (p. 11). Psychology and ethics were both founded on, and therefore linked by, their relation to pleasure and pain.

In relation to psychology, the desire for pleasure and the aversion to pain formed the basis for all motivation, both in humans and sentient creatures generally. An individual had a motive to perform an action—or put another way, had an interest in performing an action—if he or she expected to gain some pleasure or avert some pain from doing so; and the greater or more valuable the pleasure experienced or pain averted, the stronger the motive or greater the interest. The value of a pleasure or pain was determined by its quantity, which, in the case of a single individual, was a product of its intensity, duration, certainty or uncertainty, and propinquity or remoteness. Where the value of a pleasure or pain was considered in relation to more than one person, then in addition to these circumstances, the circumstance of extent, that is the number of persons affected by it, had also to be taken into account. At this point, a statement of psychological fact has become a statement of moral science. An act was morally good if, after calculating all the pains or pleasures produced in the instance of every individual affected, the balance was on the side of pleasure, and morally evil if on the side of pain. Bentham's method of determining the value of pleasure and pain is known as the "felicific calculus," though this was not a phrase that he appears to have used himself.

An adherent of the principle of utility would approve of any action that increased the overall happiness (understood in terms of a balance of pleasure over pain) of all the individuals affected by the action in question, where more than one individual was affected. An adherent of the principle of utility would likewise approve of any action that increased the happiness of a particular individual where no other individual was affected by the action in question. In the former instance the extent was equal to the total number of individuals in question, and in the latter instance to one. It was only when extent was taken into account that an action could be judged to be ethically right or wrong. The question of right and wrong was a question of fact—an account of the value, understood in terms of quantity, of the pleasures and pains that

had been brought into existence by the act in question. In order for the utilitarian legislator to accomplish his objective of promoting the greatest happiness of the greatest number, he had to use sanctions (punishments and rewards), themselves composed of pain and pleasure, to discourage actions detrimental to the happiness of the community, and (to a lesser extent) to encourage those that were beneficial.

NATURAL LAW

Bentham's adoption of the principle of utility—with its "real source" in the feelings of pain and pleasure experienced by sentient creatures—as a critical standard of morality led him to distinguish between "law as it is" and "law as it ought to be." This distinction provided the basis both for his strategy of reform, and for his attack on natural law. In *A Fragment on Government* (1776), which took Blackstone's *Commentaries* for its target, Bentham distinguished two approaches that the legal commentator might adopt: the first was that of the expositor, whose task was to describe what had been done by legislators and judges (law as it is); the second was that of the censor, whose task to show what they ought to do in future (law as it ought to be).

Blackstone, by not only describing but also attempting to justify the laws of England, had confounded the two approaches. He had, moreover, failed to adopt the principle of utility as his standard of morality, but had appealed to the doctrine of natural law, claiming that human (positive) law was valid insofar as it did not contradict the natural law. Bentham condemned Blackstone both for linking the validity of positive law to a particular substantive content, and for thinking that the natural law could supply the content in question. The natural law did not exist (it was a nonentity), hence any appeal to the law of nature in order to validate a positive law was nonsense, and in practice reflected the mere subjective approval of the supporter of the positive law in question. Blackstone had stated that where there was law, there was some superior who made it. Bentham drew out the corollary: if there was no maker, there was no law. The same problem of nonexistence bedevilled a further device adopted by Blackstone, the original contract. Having accepted the criticisms of the doctrine made by David Hume (1711–1776), Bentham went on to argue that, even if one assumed its historical existence, the original contract, like any promise, had binding force only if adherence to it would promote utility. The original contract was, therefore, superfluous, since the question as to whether to obey or resist government should be based directly on considerations of utility.

NATURAL RIGHTS

Bentham deployed similar arguments against a doctrine closely related to that of natural law, namely the doctrine of natural rights. In the French Declaration of Rights of 1789 it was asserted that the end of every political association was the preservation of the natural and imprescriptible rights of man, and that these natural rights could not be abrogated by government. The purpose of establishing government was to protect preexisting natural rights, and any government that failed to do so lacked legitimacy. In "Nonsense upon Stilts" (known as "Anarchical Fallacies" until the publication of the authoritative text in *Rights, Representation, and Reform* [Schofield, Pease-Watkin, and Cyprian Blamires 2002, pp. 317–401]) Bentham argued that there were "no such things as natural rights—no such things as rights anterior to the establishment of government—no such things as natural rights opposed to, in contradistinction to, legal" (p. 329). The notion of a state of nature, where men lived without government, was perfectly comprehensible, but in such a state there were no rights, and consequently no property and no security. Such rights might be desirable, but it was fallacious to assume that because a certain thing was desirable, that the thing in question existed. Furthermore, if natural rights did not exist, they could not be abrogated. To say that they were imprescriptible was to mount one nonsensical statement upon another: "Natural rights is simple nonsense: natural and imprescriptible rights, rhetorical nonsense, nonsense upon stilts" (p. 330). The purpose of declaring the existence of imprescriptible rights was to incite resistance to law and insurrection against government. To claim that no government could abrogate natural rights was "Terrorist language," whereas those who spoke the "language of reason and plain sense" judged whether a right should or should not be established or abrogated on the basis of whether or not it was for the advantage of society to do so (p. 330).

In *A Fragment on Government* Bentham's concern was with the distinction between the censor and the expositor, while in "Nonsense upon Stilts" it was with that between the censor and the anarchist. The anarchy that Bentham associated with the French Revolution was closely related to the conservatism he associated with Blackstone. The latter had claimed to be describing the laws of England, but had attempted to justify those laws on no other ground than that they existed. His approach confused what existed with what ought to exist. A similar

confusion characterized the anarchist, who, in claiming to describe natural rights, was making prescriptions. The difference was that while Blackstone assumed that existing law was consistent with the natural law, and therefore valid, the anarchist assumed that existing law was inconsistent with natural rights, and therefore invalid. To the extent that both were appealing to a nonexistent standard in justification of their respective claims, both were talking nonsense.

In Bentham's view, only the principle of utility provided any rational ground for resolving moral, political, and legal disputes, while talk of justice, right reason, natural rights, or moral sense was merely a cover to give respectability to, or to endow with persuasive force, the likes and dislikes of the speaker. The doctrines of natural law and natural rights were grounded on the delusive properties of language, and in particular the confusion involved in taking the name of a fictitious entity to be the name of a real entity. The use of the noun-substantive "rights" had given rise to the opinion that rights as such did actually exist. Now to talk of rights established by law did make sense, since they might be shown to have their "real source" in the will of a sovereign legislature. To talk of natural rights, with their source in natural law or a supernatural being, was to talk nonsense. The techniques of exposition that Bentham had developed in his theory of language were at the root of his attacks on natural law and natural rights.

CODIFICATION OF THE LAW

In the early 1780s Bentham concluded that the most effective means of promoting the happiness of the community would be through the introduction of a complete code of laws, or a "pannomion." Bentham's commitment to codification arose from a profound dissatisfaction with the English common law, which he characterized as corrupt, unknowable, incomplete, and arbitrary. It could not perform the minimum purpose for which law was instituted, namely to guide conduct. Still less was it able to afford protection to those basic interests of the individual—his person, property, reputation, and condition in life—which constituted his security, and hence a major component of his well-being. Security was closely related to the notion of expectations, for it involved both the present possession and the future expectation of possessing the property or other subject-matter in question. Without security, and thus the confidence to project oneself and one's plans into the future, there could be no civilized life. Security was a product of law, resulting from the imposition of rules on conduct. To an extent it did not matter which set of rules were imposed, so long as some set of rules were imposed, and these rules were known and certain. The crux of the problem with the common law was that those subject to it did not, and could not, know what it ordained, and this created insecurity. Expectations could either not be formed or were constantly liable to be disappointed.

The solution lay in codification. In his writings on the subject in the 1810s and 1820s Bentham explained that the pannomion should be "all-comprehensive" and "rationalized." This meant that the law would be logically complete, in that all legal terms would be defined consistently and related to some superior genus (where one existed), and that each provision would be followed by the reasons that justified it. At the apex of the pannomion was the civil code, concerned with the distribution of rights and duties. The purpose of the civil law was to maximize the four sub-ends of utility—namely subsistence, abundance, security, and equality. The purpose of the penal law was to give effect to the civil law, by means of attaching punishment to certain actions which, on account of their tendency to diminish the greatest happiness, were classified as offenses.

The constitutional code was also, at least in part, distributive in character, being concerned with the powers, rights, and duties of public officials, and their modes of appointment and dismissal. As with the civil law, the penal law would give effect to the relevant parts of the constitutional law. The penal, civil, and constitutional law together formed the substantive law, which was itself given effect by the adjective law, or the law of judicial procedure. The chain was completed by the law of the judicial establishment, the purpose of which was to give effect to the adjective law, and thence to the substantive law. In other words, the civil code, and to some extent the constitutional code, would contain the "directive rules" by which rights and duties were distributed, while the penal code would contain the sanctions which would enforce observance. For instance, the penal code would forbid and sanction interference with property without title, while the civil code would explain what events constituted a valid claim to title.

Bentham offered his services as a codifier to a variety of countries, including Scotland in 1808, the United States in 1811, and Russia in 1814. In April 1822 he received the invitation for which he had been longing: the Portuguese Cortes formally accepted his offer to draw up civil, penal, and constitutional codes. He immediately began to compose *Constitutional Code*, but long before even the first volume of this work had been printed in

1827, the liberal regime that had accepted Bentham's offer had been swept away. In the 1820s Bentham also devoted time and attention to Spain, Tripoli, Greece, and the emerging states of Latin America, as well as becoming fully involved in the movement to reform and codify English law. By this time he enjoyed an international reputation as the doyen of liberal legal philosophers and political reformers. José del Valle (1776–1834), for instance, the Guatemalan lawyer, economist, and politician, wrote to Bentham hailing him as "the legislator of the world."

PANOPTICON

The panopticon design was the brainchild of Bentham's brother Samuel, when employed in the 1780s on the estates of Prince Grigoriy Aleksandrovich Potemkin (1724–1791) at Krichev, in Russia. He found that by organizing his workforce in a circular building, with himself at the center, he could supervise its activities more effectively. Visiting his brother and seeing the design, Bentham immediately appreciated its potential. Enshrining the principle of inspection, the design was applicable to mental asylums, hospitals, schools, poor houses, factories, and, of course, prisons.

The prison building would be circular, with the cells, occupying several stories one above the other, placed around the circumference. At the center of the building would be the inspector's lodge, with an open space between the lodge and the cells. Each cell would have a window to the outside of the building, which would, from the perspective of the lodge, backlight the cell in daytime, while lamps, placed outside the lodge with a reflector behind them, would light the cells at night. The lodge would be so constructed, with appropriate partitions and blinds, that the inspector would always be capable of seeing into the cells, while the prisoners would be unable to see whether they were being watched. The activities of the prisoners would be transparent to the inspector; his actions, insofar as the prisoners were concerned, were hid behind a veil of secrecy. On the other hand, it was a cardinal feature of the design that the activities of the inspector and his officials should be laid open to the general scrutiny of the public, who would be encouraged to visit the prison. Bentham did not succeed in building a panopticon in London, despite gaining parliamentary approval in 1794, and the scheme was effectively quashed in 1803 (a half-hearted attempt to revive it in 1811–1812 failed). Several so-called panopticons have since been built, but none which has been particularly faithful to Bentham's own vision.

Michel Foucault in *Discipline and Punish* (1977) has described Bentham's panopticon as a paradigm of the modern state, hence placing Bentham at the center of debates about what it means to be modern. What Foucault overlooked in Bentham's case (whatever might be the case with the modern state) is that Bentham was concerned not only with the ability of officials to gain knowledge of the community subject to them (which was, of course, critical if they were to rule well), but also with the ability of the people to monitor the conduct of their rulers. The panopticon prison would be open to inspection from the public at large, just as the actions of officials would be under *Constitutional Code*. Publicity was the means of securing responsibility, and the most effective antidote against corruption.

POLITICAL REFORM

By the 1820s Bentham was convinced that the only regime with an interest in enacting good legislation was a representative democracy. Scholars disagree over precisely when Bentham committed himself to political radicalism. One view is that Bentham was a political radical from the time of the French Revolution, when, for a short period in late 1789, he advocated democracy for France. Another view, which is based on a coincidence of dates, is that Bentham became a political radical in 1808–1809, having come into contact with James Mill (1773–1836). The most plausible view, however, is that the crucial development took place around 1804 with the emergence in Bentham's thought of the notion of sinister interests, that is the systematic development of the insight that rulers wished to promote not the happiness of the community, but their own happiness. There was no point in showing rulers what the best course of legislation might be unless they had an interest in adopting it. Only a legislature elected by a democratic suffrage had such an interest.

If the key episode is the emergence of sinister interests, then the panopticon prison becomes significant. Bentham devoted many years of his life, large sums of his money (which he eventually recovered in a compensation settlement), and considerable energy, on the scheme. He was never so bitter or so despondent as when the plan was quashed in 1803. He became convinced that nothing worthwhile could be achieved through the existing political structure in Britain, or through similar regimes elsewhere. Having concentrated on questions of law reform from 1803, he was in the summer of 1809 prompted to compose material on political reform, eventually bearing fruit in *Plan of Parliamentary Reform* (1817).

In this work Bentham called for universal manhood suffrage (subject to a literacy test), annual parliaments, equal electoral districts, payment of members of parliament, and the secret ballot. Bentham then went a stage further and drew up a blueprint for representative democracy that would have abolished the monarchy, the House of Lords and any other second chamber, and all artificial titles of honor, and would have rendered government entirely open and, he hoped, fully accountable. These proposals were developed in astonishing detail in the magisterial *Constitutional Code*, the work he began in 1822 upon learning that the Portugueze Cortes had accepted his codification offer.

For Bentham the key principle of constitutional design was to ensure the dependence of rulers on subjects. Instead of the traditional theory of the separation of powers, he proposed lines of subordination, based on the ability of the superior to appoint and dismiss (in Bentham's terminology to locate and dislocate) the inferior, and to subject the inferior to punishment and other forms of vexation. The supreme power or sovereignty in the state would be vested in the people, who held the constitutive power. Immediately subordinate to the people would be the legislature, elected by universal manhood suffrage, and subordinate to the legislature would be the administrative (i.e., the executive) and judicial powers. The system of representative democracy was not an end in itself—the end was the greatest happiness—but was an indispensable means to that end, in that it was only under such a constitution that effective measures could be implemented to secure the good behavior (appropriate aptitude) of officials and minimize the expense of government. The securities for official aptitude, otherwise termed securities against misrule, included the exclusion of factitious dignities (titles of honor), the economical auction (whereby officials made bids for the salary attached to the office), subjection to punishment at the hands of the legal tribunals of the state, the requirement to pass an examination, and, most importantly, publicity.

Bentham went to great lengths to ensure that government would be open to public scrutiny, and thence subject to the force of the moral or popular sanction operating through the public opinion tribunal, which consisted in all those who commented on political matters, and of whom newspaper editors were the most important. Bentham saw the freedom of the press as a vital bulwark against misrule: hence his proposal to encourage the diffusion of literacy by making the suffrage dependent on a literacy test. These measures were intended to ensure that rulers would be so situated that the only way they could promote their own interest was by promoting the interest of the community.

RELIGION

Bentham offered a secular vision of society, where the standard of rectitude would be founded not on theology, or natural law, or right reason, or precedent, or sheer prejudice, but on observation and experience. Knowledge of society (and of the individuals who composed it) enshrined in a "political science" (for Bentham's use of the term see, for instance, *Official Aptitude Maximized* [Schofield 1993, p. 191]) would be the basis for the art of legislation, the practical measures that an enlightened legislator would introduce in order to promote the greatest happiness of the community. Bentham was committed to freedom of expression in religion, as in other areas. While it may be too quick to conclude that he was an atheist, he did ally himself from an early period in his life with those who were sceptical, if not of religious belief, certainly of organized religion, and he never wavered in his outright opposition to religious establishments. As early as the mid-1770s, he drew attention to the potential mischiefs associated with what he termed the religious sanction. The expectation of a future state amounted to the expectation of the distribution of pains and pleasures, but did not in itself entail any rules specifying in what way such pains and pleasures would be distributed. If this distribution was to be random, then the expectation of them could not have any influence in encouraging good conduct or restraining bad. Given that the idea of God might provide motives, but could not provide direction, it was better that the moralist and legislator had nothing to do with it.

In the 1810s Bentham launched a sustained attack against established religion. He argued that religious belief was used to further the particular and sinister interest of the priesthood and those linked with it. The Anglican Church was an instrument in the hands of rulers to oppress and extort resources from subjects. It extracted large sums of money from the population generally, in order to provide income for rulers, without providing any useful service in return. The state supported the Church with its coercive force, while the Church manufactured delusive arguments in support of the state. Indeed, the scale of abuse in the Church was not only greater than that in the political and legal establishments, but acted as a bulwark against reform elsewhere. Bentham was particularly critical of the role of the Church in education, both in schools and in the Universities of Oxford and Cambridge. In relation to the poor, its policy was to exclude

from the benefits of education those unwilling to declare their belief in Anglican doctrine, and to pervert the morals and intellects of those who were willing.

Bentham's resentment at being forced to subscribe to the Thirty-nine Articles while at Oxford led him to insist that the provision of education should be divorced from the profession of belief. He recommended the "euthanasia" of the Anglican Church, whereby, as livings and other offices became vacant, they would be abolished. The present possessors would retain their incomes and thereby not suffer the pain of disappointment, while the expense of the religious establishment to the state, and thus to the people generally, would gradually diminish, and the additional income derived would be used to reduce taxation. Those people who wished to receive religious instruction could continue to do so at their own expense.

AUTO-ICON

Bentham was not buried, but his body transformed into what he termed an auto-icon. He had left instructions in his will that his body should be used in a series of anatomical lectures, and thereafter his skeleton "put together in such manner as that the whole figure may be seated in a Chair usually occupied by me when living in the attitude in which I am sitting when engaged in thought" (Crimmins 2002, p. 8). The operation was entrusted to Bentham's surgeon, Thomas Southwood Smith (1788–1861), who created the auto-icon—the combination of skeleton, wax head, clothes, and stuffing—which now resides in University College London.

See also Aristotelianism; Beccaria, Cesare Bonesana; Democracy; Foucault, Michel; Hume, David; Legal Positivism; Mill, James; Mill, John Stuart; Newton, Isaac; Pleasure; Property; Punishment; Utilitarianism.

Bibliography

WORKS BY JEREMY BENTHAM

The Bentham Project, University College London, is preparing a new authoritative edition of *The Collected Works of Jeremy Bentham*, which, it is estimated, will run to sixty-eight volumes. The twenty-fifth appeared in March 2002. The new edition is based on two main sources: first, texts printed during Bentham's lifetime; and second, Bentham's original manuscripts, of which around 55,000 folios are deposited in University College London Library, and around 10,000 in the British Library. Of the volumes in *The Collected Works*, the following have been most extensively drawn on for the entry above, and in-text citations correspond to these editions:

An Introduction to the Principles of Morals and Legislation, edited by J. H. Burns and H. L. A. Hart. London: Athlone Press, 1970.

A Comment on the Commentaries and A Fragment on Government, edited by J. H. Burns and H. L. A. Hart. London: Athlone Press, 1977.

Constitutional Code, Vol., edited by F. Rosen and J. H. Burns. Oxford: Clarendon Press, 1983.

The Correspondence of Jeremy Bentham. Vol. 4, *October 1788 to December 1793* edited by A. T. Milne. London: Athelone Press, 1981.

First Principles Preparatory to Constitutional Code, edited by Philip Schofield. Oxford: Clarendon Press, 1989.

Official Aptitude Maximized; Expense Minimized, edited by Philip Schofield. Oxford: Clarendon Press, 1993.

"Legislator of the World": Writings on Codification, Law, and Education, edited by Philip Schofield and Jonathan Harris. Oxford: Clarendon Press, 1998.

Rights, Representation, and Reform. Nonsense upon Stilts and Other Writings on the French Revolution, edited by Philip Schofield, Catherine Pease-Watkin, and Cyprian Blamires. Oxford: Clarendon Press, 2002.

Where works have not appeared in *The Collected Works*, the standard source is the so-called Bowring edition: *The Works of Jeremy Bentham*, published under the superintendence of his executor, John Bowring. 11 vols. Edinburgh: William Tait, 1843. For Bentham's economic thought see *Jeremy Bentham's Economic Writings*. 3 vols., edited by Werner Stark. London: George Allen & Unwin, 1952–1954.

There are two convenient collections of essays on Bentham: *Jeremy Bentham: Critical Assessments*. 4 vols., edited by Bhikhu Parekh. London: Routledge, 1993; and *Bentham: Moral, Political, and Legal Philosophy*. 2 vols., edited by Gerald J. Postema. Aldershot: Ashgate, 2002.

WORKS ABOUT JEREMY BENTHAM

Bahmueller, Charles F. *The National Charity Company: Jeremy Bentham's Silent Revolution*. Berkeley: University of California Press, 1981.

Ben Dor, Oren. *Constitutional Limits and the Public Sphere: A Critical Study of Bentham's Constitutionalism*. Oxford: Hart Publishing, 2000.

Boralevi, Lea Campos. *Bentham and the Oppressed*. Berlin: Walter de Gruyter, 1984.

Crimmins, James E. *Bentham's Auto-Icon and Related Writings*. Bristol: Thoemmes Press, 2002.

Crimmins, James E. *Secular Utilitarianism: Social Science and the Critique of Religion in the Thought of Jeremy Bentham*. Oxford: Clarendon Press, 1990.

Dinwiddy, John. *Bentham: Selected Writings of John Dinwiddy*, edited by William Twining. Stanford, CA: Stanford University Press, 2004.

Engelmann, Stephen G. "Imagining Interest." *Utilitas* 13 (3) (2001): 289–322.

Foucault, Michel. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. Harmondsworth: Allen Lane, 1977.

Halévy, Elie. *The Growth of Philosophic Radicalism*. Translated by Mary Morris. London: Faber & Faber, 1928.

Harrison, Ross. *Bentham*. London: Routledge & Kegan Paul, 1983.

Hart, H. L. A. *Essays on Bentham: Jurisprudence and Political Theory*. Oxford: Clarendon Press, 1982.

Hume, L. J. *Bentham and Bureaucracy*. Cambridge, U.K.: Cambridge University Press, 1981.

Kelly, P. J. *Utilitarianism and Distributive Justice: Jeremy Bentham and the Civil Law*. Oxford: Clarendon Press, 1990.

Lieberman, David. "Economy and Polity in Bentham's Science of Legislation." In *Economy, Polity, and Society: British Intellectual History, 1750–1950*, edited by Stefan Collini, Richard Whatmore, and Brian Young. Cambridge, U.K.: Cambridge University Press, 2000.

Lieberman, David. *The Province of Legislation Determined: Legal Theory in Eighteenth-Century Britain*. Cambridge, U.K.: Cambridge University Press, 1989.

Lobban, Michael. *The Common Law and English Jurisprudence 1760–1850*. Oxford: Clarendon Press, 1991.

Long, Douglas G. *Bentham on Liberty: Jeremy Bentham's Idea of Liberty in Relation to His Utilitarianism*. Toronto: University of Toronto Press, 1977.

Lyons, David. *In the Interest of the Governed: A Study in Bentham's Philosophy of Utility and Law*. Rev. ed. Oxford: Clarendon Press, 1991.

Mack, Mary P. *Jeremy Bentham: An Odyssey of Ideas 1748–1792*. London: Heinemann, 1962.

Pitts, Jennifer. "Legislator of the World? A Rereading of Bentham on Colonies." *Political Theory* 31 (2) (2003): 200–234.

Postema, Gerald J. *Bentham and the Common Law Tradition*. Oxford: Clarendon Press, 1986.

Postema, Gerald J. "Fact, Fictions, and Law: Bentham on the Foundations of Evidence." In *Facts in Law*, edited by William Twining. Wiesbaden: Franz Steiner Verlag, 1983.

Rosen, Frederick. *Bentham, Byron, and Greece: Constitutionalism, Nationalism, and Early Liberal Political Thought*. Oxford: Clarendon Press, 1992.

Rosen, Frederick. *Classical Utilitarianism from Hume to Mill*. London: Routledge, 2003.

Rosen, Frederick. *Jeremy Bentham and Representative Democracy: A Study of the Constitutional Code*. Oxford: Clarendon Press, 1983.

Schofield, Philip. "Jeremy Bentham's 'Nonsense upon Stilts.'" *Utilitas* 15 (1) (2003): 1–26.

Schofield, Philip. "Jeremy Bentham, the French Revolution and Political Radicalism." *History of European Ideas* 30 (2004): 381–401.

Schofield, Philip. "Jeremy Bentham, the Principle of Utility, and Legal Positivism." In *Current Legal Problems 2003*. Vol. 56, edited by M. D. A. Freeman. Oxford: Oxford University Press, 2004.

Semple, Janet. *Bentham's Prison: A Study of the Panopticon Penitentiary*. Oxford: Clarendon Press, 1993.

Twining, William. "Imagining Bentham: A Celebration." In *Current Legal Problems 1998*. Vol. 51: *Legal theory at the End of the Millennium*, edited by M. D. A. Freeman. Oxford: Oxford University Press, 1998.

Twining, William. *Theories of Evidence: Bentham and Wigmore*. London: Weidenfeld & Nicolson, 1985.

Warke, Tom. "Multi-Dimensional Utility and the Index Number Problem: Jeremy Bentham, J. S. Mill, and Qualitative Hedonism." *Utilitas* 12 (2) (2000): 176–203.

Philip Schofield (2005)

BERDYAEV, NIKOLAI ALEKSANDROVICH

(1874–1948)

Nikolai Aleksandrovich Berdyaev, a Russian religious philosopher, was born in Kiev in a family of the old nobility. He attended the Kiev military school. In 1894 he enrolled in St. Vladimir's University of Kiev as a natural sciences student, but after a year transferred to the department of law. Infatuation with Marxism and participation in the social-democratic movement led to his arrest, exclusion from the university (in 1898), and a three-year exile to Vologda. This represented a break with the aristocratic environment to which he had been accustomed, a break that he later called a fundamental fact of his biography, not only of his external biography but also of his inner one.

Berdyaev's Marxist period did not last long; in a short period of time he underwent an evolution that was characteristic for many Russian thinkers of the beginning of the twentieth century—from Marxism to idealism to the search for God. Berdyaev was one of the initiators of three collections of essays that became famous and provoked much heated argument: *Problemy idealizma* (Problems of idealism; 1902), *Vekhi* (Landmarks; 1909), and *Iz glubiny* (De Profundis, Out of the depths; 1918). Berdyaev greeted the fall of the monarchy in February 1917 with great enthusiasm, but he assessed the October Revolution differently—as the triumph of the destructive principle in the Russian revolution. He participated in the work of the Vladimir Sergeevich Solov'ëv (Solovyov) Religious-Philosophical Society and was the founder of the Free Academy of Spiritual Culture (1918–1922), which became a non-Marxist spiritual center and continued the traditions of the Russian Silver Age after the Bolshevik coup. In 1919 Berdyaev was elected as a professor of Moscow University. Despite the fact that Berdyaev was remote from actual political struggle, in 1922 he and other outstanding figures of Russian culture were forcibly deported from Soviet Russia to Germany.

In 1922 Berdyaev founded the Religious-Philosophical Academy in Berlin, and in 1923 he became the dean of the Russian Scholarly Institute, established in Berlin to educate the Russian émigré youth. Also in 1923 he

became a member of the council of the Russian Student Christian Movement, in which he participated until 1936. In 1924 he moved to France, where he edited the religious-philosophical journal *Put'* (The way; 1925–1940). The Religious-Philosophical Academy that he had founded also moved to Paris, and there he read lecture courses on "The Problems of Christianity," "The Fate of Culture," "Man, the World, and God," and so on. Berdyaev was one of the few Russian émigré thinkers who did not confine himself in the émigré milieu. During his lifetime he wrote a great many books that were published not only in Russian but also in other languages. His religious existentialism found a response among a number of West European thinkers; his philosophical ideas were esteemed highly by such figures as Jacques Maritain, Gabriel Marcel, Ernst Bloch, and Karl Barth. Berdyaev had a particular influence on the philosophical circles gathered around the journal *Esprit*, which was founded by Emmanuel Mounier in 1932 and inaugurated French personalism. In 1947 Cambridge University awarded Berdyaev the title "Honoris causa." Berdyaev died in 1948 in a suburb of Paris.

METAPHYSICS OF FREEDOM

Berdyaev's religious-philosophical doctrine was greatly influenced by the ideas of Solov'ëv, Immanuel Kant, Fëdor Mikhailovich Dostoevsky, and the seventeen-century German mystic Jakob Boehme. According to Berdyaev the distinguishing characteristic of philosophy consists in the fact that it is not reducible to a system of concepts, but that it rather represents a knowledge that speaks in the language of symbols and myths. In his own philosophy the central role belonged to freedom and creativity. Berdyaev (like Boehme) bestowed an ontological status on freedom; he believed that freedom has primacy in relation to natural and human being and that it is independent of God's being. Berdyaev often used Boehme's term *Ungrund* (groundlessness or bottomlessness) to describe such pre-ontic freedom. God expresses only the light or radiant side of this freedom, and the world created by him could also be radiant and good. But God cannot compel the world to be good, and one's free choice is not always in favor of the good (such was Berdyaev's interpretation of the biblical myth of the fall of man). That is how evil arises in the world. One has difficulty understanding why God did not create a world without sin, sicknesses, children's tears, and suffering. The answer is simple: Such a world would not have freedom, which lies at the foundation of the universe and which God does wish to limit and cannot limit.

Berdyaev traced the paradoxical and tragic dialectic of the good and freedom: on the one hand, it is obvious that one cannot be compelled to be good, but on the other hand, the freedom of the good also presupposes the freedom of evil in the world. Like Dostoevsky, Berdyaev rebelled against compulsory harmony imposed on human beings from outside. Without the freedom of sin, evil, trial, and suffering, one cannot understand harmony or the kingdom of God. Because of this tragic dialectic the world has to undergo the "trial by freedom" so that its choice in favor of the good will be free; and the fate of the world coincides, in the final analysis, with the fate of freedom in the world. The thesis that freedom has an uncreated and pre-ontic character is foundational for Berdyaev's philosophy, for if one supposes that freedom was created character, then God himself would turn out to be responsible for the evil of the world. However, for Berdyaev, God is revealed to humans, and humans, through their freely followed destiny, are revealed to God; and Revelation is thus a mutual process.

Berdyaev's Christianity was tragic and not fully orthodox. He had an acute sense of the presence of evil in the world and the substantiality of evil. This led him to pose the problem of theodicy, to attempt to understand the causes why evil is permitted in the world. If the first stage of Berdyaev's spiritual evolution was Marxist and the second idealistic in character, the third stage begins precisely with posing the problem of theodicy. It can be described as Berdyaev's Christian period.

PERSONALISM

In Berdyaev's worldview freedom and spirit are opposed to unfreedom and necessity, to the material "world of objects." For him these are two kinds of realities, interacting with each other. The world in which one lives is fallen precisely because it is dominated not by freedom but by necessity. In the reality that surrounds one, all things are regulated by law and unfree. (Here, Berdyaev's position converges with that of the other existentialists.) Reason and rational knowledge cannot help one free oneself from the necessity externally imposed on one, since reason and rational knowledge signify only adaptation to the world of objects.

Free people find themselves in a world dominated by necessity. And naturally they strive to escape from the power of the lower reality, where all things are regulated by law and are predictable. But they can escape only through creative activity, which is always a free expression

of their selves. In a creative act people once again feel themselves to be a godlike being, not constrained by the laws of the material world. People are called to creative activity, to the continuation of the creation of the world, for the world is fundamentally unfinished. The primacy of freedom over being also determines the meaning of human life: the goal of people is not salvation, but creative activity; the creative act has intrinsic value.

Berdyaev proclaimed that the purpose of creative activity is not to accumulate cultural values, but to bring to an end the fallen world of necessity. For Berdyaev the social reality is only an objectification (symbolization or materialization) of the subjective personal spirit. He reinterprets Kant in his own manner, concurring with Kant's recognition of another reality that is more profound and hidden behind the objectified world.

For Berdyaev, social problems (e.g., hunger, poverty, and inequality) are secondary in comparison with spiritual problems. The elimination of hunger and poverty will not liberate people from the mystery of death, love, and creative activity. Furthermore, the conflicts between the individual and society, humans and the cosmos, history and eternity are only made more acute in the case of a more rationally ordered society. People are called to creative activity, but all creative activity is inevitably a failure, since the results of such activity are objectified and participate in the enslavement of man. "The ardent creative spirit" cannot recognize itself in works of art, books, or theories—in its products. The results of creative activity are alienated from the creator. According to Berdyaev creative activity is "ascent out of the world," but a total break with the world is impossible; and this constitutes the tragic character of human existence.

HISTORIOSOPHY

According to Berdyaev every person lives not in one time, but in at least three times: Since people are simultaneously natural, social, and spiritual beings, there also exist three times for them: cosmic, historical, and existential. Berdyaev uses geometrical figures to describe these three times: the circle, the line, and the point. Cosmic time follows the natural and regular logic of circular motion; this time operates not with days and years but with epochs and millennia. By contrast, historical time follows a straight line and operates with smaller temporal categories. However, the most significant events occur in existential time; it is precisely in the latter that creative acts and free choice take place. For existential time the duration of an event is relative: sometimes for a person a day is longer and more significant than a decade, whereas sometimes a year slips by imperceptibly. A person's earthly time itself is only a phase, a period within eternity; it is rooted in eternity. The eternal is made incarnate in time; it invades time (just as heavenly history invades earthly history), and history becomes the history of the battle of the eternal against the temporal. But the forces are not equal. The eternal will triumph over all that is corruptible and fleeting: The objectified world will perish. All creative activity represents an escape from the chain of cause and effect, which is why every creative act shakes the foundations of cosmic necessity. Berdyaev's vantage point is an eschatological one; he believes that the meaning of history is in its end, in the triumph of the free spirit over objectification. Earthly history is the path to the other world; this history is too narrow and limited for the incarnation of the ideal; the problem of history can be solved only beyond the limits of earthly history, in eternity.

In trying to understand the tragic experience of the Russian revolution and the tendencies of European development, Berdyaev proclaimed that the areligious, humanistic epoch had reached its completion and that humankind had entered the sacral epoch of new Middle Ages, characterized by a religious renaissance and religious conflicts. Berdyaev claimed that, in the twentieth century, all significant ideas inevitably acquired a religious meaning. This goes also for communist ideology: using Soviet Russia as an example, he showed that this country had entered the epoch of new Middle Ages, for he considered Russian Marxism to be a type of religious faith with its savior (the proletariat), prophets (Marx, Friedrich Engels, and Vladimir Il'ich Lenin), "doctrine of man's fall" (the history of the emergence of private property), paradise (communism), and so on. Russia was at the leading edge of the historical process, as it were; and after the revolution the Russian idea had acquired a universal significance.

Berdyaev identified six fundamental stages of world history. The first stage was that of antiquity, when people were submerged in the depths of natural necessity. Berdyaev associated the second stage with the fate of the Jewish nation, with its messianic consciousness, thanks to which the static ancient was replaced by the historical approach to reality. The third stage was that of the overcoming of the two preceding stages by Christianity, which introduced the idea of eschatology into the human consciousness. The fourth stage was the epoch of the Renaissance, when humanism was born and people's falling away from God began. The reaction to this was the Reformation, the fifth stage, when, in counterweight to the Renaissance spirit, people's independence was denied and

their total dependence on divine providence was proclaimed. The sixth stage, according to Berdyaev's conception, was associated with socialism, with the attempt to realize the kingdom of God on earth.

By the will of the fates, Russia, without having experienced some of these historical stages, became humanity's testing ground for the realization of the totalitarian-socialistic ideal. But, in Berdyaev's opinion, Russian socialism also became the sign of the transition to the seventh stage, to the new Middle Ages, a period of religious-social synthesis. Berdyaev proposed his own version of socialism, which resembled its Marxist counterpart in only one thing: a fundamental antibourgeois attitude. For Berdyaev, socialism has a dual nature: it can create either a new free society or a new slavery. Berdyaev himself was a proponent of personalistic Christian socialism.

See also Barth, Karl; Bloch, Ernst; Boehme, Jakob; Communism; Dostoevsky, Fyodor Mikhailovich; Engels, Friedrich; Evil; Existentialism; Humanism; Idealism; Kant, Immanuel; Lenin, Vladimir Il'ich; Marcel, Gabriel; Maritain, Jacques; Marxist Philosophy; Marx, Karl; Personalism; Reformation; Renaissance; Russian Philosophy; Socialism; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY BERDYAEV

The Meaning of History. Translated by George Reavey. London: G. Bles, 1936.

The Origin of Russian Communism. Translated by R. M. French. London: G. Bles, 1937.

The Russian Idea. Translated by R. M. French. London: G. Bles, 1947.

The Meaning of the Creative Act. Translated by Donald A. Lowrie. London: V. Gollancz, 1955.

WORKS ABOUT BERDYAEV

Lowrie, Donald Alexander. *Rebellious Prophet: A Life of Nicolai Berdyaev*. New York: Harper, 1960.

Vallon, Michel Alexander. *An Apostle of Freedom: Life and Teachings of Nicolas Berdyaev*. New York: Philosophical Library, 1960.

Olga Volkogonova (2005)
Translated by Boris Jakim

BERGERAC, CYRANO DE

See *Cyrano de Bergerac, Savinien de*

BERGMANN, GUSTAV
(1906–1987)

Gustav Bergmann came to the United States in 1938 from Vienna, Austria, where he had earned a JD and a PhD in mathematics. He had also been a junior member of the Vienna Circle.

In 1939 he became a faculty member at the University of Iowa, retiring in 1976. He held a joint appointment in the Departments of Philosophy and Psychology. He regularly taught a course on the history and philosophy of psychology. Bergmann became well known as an apologist for behaviorism. Significantly, he distinguished between methodological and metaphysical behaviorism, embracing the former and rejecting the latter. Bergmann never wavered in his ontological commitment to the mental.

Bergmann also published in mathematics, the philosophy of physics, the history of philosophy, and the philosophy of law. His *Philosophy of Science* (1957) is an elegant and still useful work. He was, however, first and foremost a philosopher, an ontologist to be exact. The central question is what exists. His method for answering that question, the ideal-language method, was to design a formalism into which one could transcribe all empirical statements of the natural language and which formalism could be used to account for the difference between the necessary and the contingent statements of the natural language. The ontology of the world would be revealed by the difference in the kinds of basic, undefined sign of the formalism.

The necessary-contingent distinction was relatively easy to handle. What is necessary and contingent is a given. One needs merely to transcribe the necessary statements into sentences of the formalism, the truth values of which sentences are a matter of form, and the contingent ones to sentences the truth values of which are not a matter of form. The idea is a classical one; the only difference being that the classical philosophers spoke of thoughts as truth bearers whereas the ideal-language philosophers spoke of sentences of the formal language as truth bearers. Relatedly, for the classical philosophers the truth makers were either features of the thought or of something beyond, the thought, whereas for the formalist the truth bearers were either features of sentences or something beyond the sentence.

Determining what kinds of signs are basic was difficult to handle. Bergmann began as a positivist: The only existents were the entities stood for by the subjects and predicates of atomic sentences, entities with which one

had to be acquainted. He was, then, a phenomenalist. In time, he acknowledged that the operators were not eliminable; they had to stand for entities that had ontological status. A distinction was thus made between existents and subsistents. Logical entities subsist; empirical, sensuous ones exist. The latter presented their own problems. Each entity was of a kind, particular or universal. Thus, a simple entity was a complex of sorts, a form and a content. Unlike the early Ludwig Josef Johann Wittgenstein, Bergmann insisted on according the forms ontological status. Forms subsist. That put pressure on the use of the Principle of Acquaintance, sufficient pressure to force Bergmann to replace it with a Principle of Presentation, a principle that cast a wide net indeed.

In his last phase Bergmann became sensitive to the criticism that he was a mere formalist and that all his ontological claims were transcendental ones, his talk of acquaintance and presentation being mere talk. His last work, *New Foundations of Ontology* (1992), published posthumously, is rich in talk about "phenomenological bedrock." Bergmann's fate was a curious one. His commitment to particulars, universals, forms, and whatever else was dictated by the needs of the formalism and by his conception of the difference between eliminable and ineliminable terms rather than by the need to solve such problems as that of individuation and universals. The issue of whether the basic entities are "experienced" was an afterthought, a most nettlesome one.

Bergmann's devotion to the method was never shaken; and in the context of the method he made two brilliant moves. First, in the mid-1950s he found a way to render in the formalism an analysis of mental acts. As act was a particular with two properties, one for the kind of act it was (a remembering, a doubting, or whatever) the other for the content of the act (that the moon is blue, that the ball is red, or whatever). (One would benefit from comparing Bergmann's analysis with René Descartes's third-meditation discussion of the use of the term *idea*.) Regarding the content-carrying property, Bergmann ran into a problem. He wanted it to be simple but had to have it complex, the reason being that the property had to serve as a truth bearer and for that need to be satisfied the property had to have within it a mark that would indicate the truth maker for it. The alternative would be to introduce an objectionable state of affairs that would show that the content property was related to a possibility that itself would contain a mark of its truth maker. The move, brilliant though it was, failed; but its failure provides one with something deeply instructive about the "make true" talk.

Second, ontology is about the kinds of entity that exist. Most formalisms need to give significance to the order of the signs in a relational sentence. There is an important difference between, say, Othello loving Desdemona and Desdemona loving Othello. The order of the terms flanking the relation sign contributes to the meaning of the sentences. Bergmann's last work was in part an attempt, as he liked to express it, to delinearize the language. He introduced dyads, a dyad being a pair of entities combined by a nexus that is other than exemplification, the tie that tied, say, two particulars and a relation into a fact. Accordingly, "aRb" was replaced by "aR{ab}," and "bRa" by "bR{ab}." Order makes no difference. The two relational facts are different in virtue of different entities. The disposing of order comes with a heavy price: nonsimple entities that are not facts require a tie, cannot exist independently of facts, and are treated by the syntax as if they were simple terms. Again, a brilliant move fails; and for a reason rather like the reason for the first failure. A nonfact complex is needed where one wants desperately to have a simple one.

Notwithstanding the failures, Bergmann's philosophical work is deep and probing and unfailingly illuminating. It has much to teach about not only the use of formalisms in doing ontology but also about the classical tradition.

See also Behaviorism; Descartes, René; Logical Positivism; Ontology, History of; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY BERGMANN

The Metaphysics of Logical Positivism. New York: Longmans, Green, 1954.

The Philosophy of Science. Madison: University of Wisconsin Press, 1957.

Meaning and Existence. Madison: University of Wisconsin Press, 1959.

Logic and Reality. Madison: University of Wisconsin Press, 1964.

Realism: A Critique of Brentano and Meinong. Madison: University of Wisconsin Press, 1967.

New Foundations on Ontology, edited by William Heald. Madison: University of Wisconsin Press, 1992.

Collected Works, edited by Erwin Tegtmeier. Ontos Verlag, 2003–2004.

WORKS ABOUT BERGMANN

Gram, M. S., and E. D. Klemke, eds. *The Ontological Turn: Studies in the Philosophy of Gustav Bergmann*. Iowa City: University of Iowa Press, 1974.

Hochberg, Herbert. *The Positivist and the Ontologist: Bergmann, Carnap, and Logical Realism*. Amsterdam, Netherlands: Rodopi, 2001.

Edwin Allaire (2005)

BERGSON, HENRI

(1859–1941)

Henri Bergson, the French philosopher of evolution, was born in Paris of Anglo-Polish parentage. During a lifetime of teaching, lecturing, and writing, he gained an international reputation as the author of a new and distinctive philosophical outlook presented in a succession of books whose fluent, nontechnical style gave them a wide appeal. In 1900 Bergson became professor of philosophy at the Collège de France, a post he held until 1921, when ill health obliged him to retire. He received many honors, including election to the French Academy and in 1927 the Nobel Prize for literature. After World War I, Bergson devoted much attention to international affairs, in the hope of promoting peace and cooperation among nations. But World War II had begun and France had been occupied by the armies of Nazi Germany at the time of his death.

Despite the novelty of his outlook, Bergson owed much to his predecessors in the European, and especially in the French, philosophical tradition, primarily to thinkers whose ideas supported his opposition to materialism and mechanism; he was convinced that neither of these doctrines is philosophically tenable. Thus, he was influenced by the idea of Maine de Biran that we sense the "flow" of life as a primary inner experience; by the contentions of Felix Ravaisson that philosophic thought should be focused on the directly intuited, concrete individual, and that mechanism is the external form of an inner spiritual activity; by the contention of Alfred Fouillée that there is an intrinsic freedom in human action; and by the teaching of Émile Boutroux that there exists a radical contingency in nature. His obligation to ancient thought was chiefly to Plotinus, whose mysticism became increasingly congenial to Bergson in the later years of his life. The theory of biological evolution, in both Charles Darwin's scientific formulation and Herbert Spencer's speculative formulation deeply influenced him. He was once "very much attached to the philosophy of Spencer" (*The Creative Mind*, p. 93), but broke away because of its unsatisfactory treatment of evolution and of time.

TWO KINDS OF TIME

Of central importance in Bergson's outlook is his distinction between the time that occurs in the theories of natural science and the time that we directly experience. Scientific time is a mathematical conception, symbolized in physical theory by the letter *t* and measured by clocks and chronometers. Because these measuring instruments are spatial bodies, scientific time is represented as an extended, homogeneous medium, composed of standard units (years, hours, seconds). Most of man's practical life in society is dominated by these units. But time thus represented neither "flows" nor "acts." It exists passively, like a line drawn on a surface. When we turn to our direct experience, Bergson urged, we find nothing that corresponds to this mathematical conception. What we find, on the contrary, is a flowing, irreversible succession of states that melt into each other to form an indivisible process. This process is not homogeneous but heterogeneous. It is not abstract but concrete. In short, it is "pure time" or "real duration" (*durée reelle*), something immediately experienced as active and ongoing. If we try to represent it by a spatial image, such as a line, we only generate abstract, mathematical time, which is at bottom an illusion. The great weakness of mechanistic modes of thought is that they consider this illusion to be a reality.

DETERMINISM AND FREEDOM

In *Time and Free Will* Bergson undertook to show that the recognition of real duration provides a basis for vindicating human freedom and disposing of determinism. The determinist, according to Bergson, holds that freedom of choice does not exist. He supports his view by picturing the situation in which one confronts an ostensible choice as being like arriving at a point on a line where a branching occurs, and taking one of the branches. The determinist then contends that the particular branch taken could not *not* have been taken. He further holds that, given full knowledge of the antecedent states of mind of the agent, the branch taken could have been predicted beforehand.

The force of this argument, according to Bergson, derives from misrepresenting the situation of choice by using an abstract, spatialized conception of time. At best the determinist's image of the line symbolizes the choice already made, not the choice in the making. In acting we do not move along a path through time. Deliberating about a choice is not like being at a point on a line and oscillating in space between various courses confronting us. Deliberation and choice are temporal, not spatial, acts. Moreover, the determinist makes the associationist's mis-

take of supposing that the mind of the agent consists of a succession of atomic states that determine how he will act. The associationist's mechanistic interpretation of the mind produced a fallacious picture upon which determinism was superimposed.

Freedom of action, according to Bergson, is something directly experienced. Man feels himself to be free as he acts, even though he may be unable to explain the nature of his freedom. Nevertheless, we are free only when our act springs spontaneously from our *whole* personality as it has evolved up to the moment of action. If this spontaneity is absent, our actions will be simply stereotyped or mechanical responses. In such cases we behave like automata. Hence, freedom is far from being absolute. Indeed, for most people free acts are the exception, not the rule. To this extent the determinists are right.

BODY AND MIND

Direct experience not only establishes the reality of time and of freedom; it also testifies that each of us is a body, subject to the same laws as all other bits of matter. Bergson's dualism emerges clearly in *Matter and Memory*. Bodies are there interpreted as "images"; that is, objects perceived in space. Among these images is one that I know from the outside by perception and from the inside by sensation or affection. This is my own body, which I also know to be a center of action.

What is the relation between the body and the mind? Materialism holds that mind, or consciousness, is either identical with brain activity or existentially dependent on brain activity. Bergson rejected both positions because, he claimed, there is vastly more in a given occasion of consciousness than in the corresponding brain state. The attempt to substantiate this claim led him to reject the doctrine that a parallelism exists between the series of conscious states and the series of brain states. The considerations to which he appealed came mainly from an examination of memory.

TWO KINDS OF MEMORY

Living organisms, unlike nonliving objects, retain their past in the present. This phenomenon is manifested, according to Bergson, in two kinds of memory. One kind consists of sensory-motor mechanisms or "habits" fixed in the body of the organism and designed to ensure adaptation to a present situation. When an appropriate stimulus arises, one of these mechanisms "unwinds" as a response. The other kind of memory, which humans alone possess, records in the form of memory images all the events of daily life as they occur in time. These images provide the content of occasions of recalling. This is "pure" memory, which is wholly spiritual. "Consciousness signifies, before everything, memory."

To defend his view of pure memory, Bergson argued against any correlation of memory images with hypothetical memory traces stored in the brain. Physiologically, the brain consists of a vast number of neurons, synapsing with each other and with afferent and efferent nerves. It resembles a telephone exchange, not a storage device. There is no evidence that memories are located spatially within it. Moreover, if a visual recollection of an object were dependent on a brain trace, there would have to be thousands of traces, corresponding to all the variations due to different points of view from which the object has been perceived. But what we actually have in each case is one practically invariable memory image of an object, not a large class of different images. This, Bergson thought, constitutes proof that something quite distinct from mechanical registration is involved. Finally, there are facts associated with loss of word memory and its restoration which point to the conclusion that the recollective process is independent of brain traces. It follows that materialism and psychoneural parallelism are untenable doctrines.

How, then, is pure memory related to the brain? Bergson's answer is derived from his contention that pure memory retains the whole of our past. If this is the case, something must prevent all our memories from being simultaneously present to consciousness, since we do in fact recall only one or two things at a time. The brain must therefore act as a filter for our memories, allowing only those that are practically useful to emerge on a given occasion. In other words, the brain is a mechanism invented by nature to canalize and direct our attention toward what is about to happen, in order to assist our actions. It is designed not so much to promote remembering as to promote forgetting. By bringing pure memory into contact with practical actions, it also establishes a link with habit memory, since most of our everyday actions tend to be habitual and routine. In this way the two kinds of memory are united.

Although he would not countenance the idea that memory traces are stored in the brain, Bergson allowed for the storage of images in pure memory. He asserted that pure memory retains all our conscious states "in the order in which they occur." This view led him to accept the conclusion that part of the mind is unconscious or subconscious. It is erroneous to suppose that the existence of psychical states depends on their apprehension by consciousness. To suppose this is to vitiate the concept

of mind by casting an artificial obscurity over the idea of the unconscious. The significance of pure memory can be understood only by supposing that past psychological states have a real, though unconscious, existence.

It is now possible to explain the relation between the body and the mind. Here, as elsewhere, there has been a strong temptation to think in spatial terms, envisaging two separate substances that have to be connected. But the relation between body and mind must be understood in temporal, not spatial, terms. The point becomes clear when we unite the insight derived from our consciousness of real duration with the recognition that the body is a center of action, for on an occasion of action, body and mind are related by a convergence in time. No spatial representation of this convergence can be adequate. It can be grasped only by noting what takes place whenever we act. A familiar example is our perception of the external world.

PERCEPTION AND THE EXTERNAL WORLD

The discussion of this question forms an integral part of *Matter and Memory*. In considering perception, traditional realism and idealism have, according to Bergson, made two unjustified assumptions. First, they have assumed that perception is a kind of photographic process that yields a picture of what is perceived. The mind is envisaged as a *camera obscura* inside which images are generated. Second, they have regarded perception as a cognitive function whose aim is to provide pure knowledge. Bergson contended that perception cannot possibly be a photographic process, for images are not inside the mind but are part of the spatially extended world. Moreover, perception does not generate images, but selects those images that have a possible bearing on actions. Nothing remotely akin to pure knowledge is involved at the perceptual level. Once these assumptions are discarded, the dispute between realism and idealism can be resolved.

In supporting this idea Bergson used biological considerations. Biologists are agreed that there has been an evolution of the structure and the functions of the central nervous system in living organisms. This evolution has proceeded from relatively simple types of organization toward greater and greater complexity, through a series of minute, adaptively significant changes. In simple organisms the rudiments of perception are to be found in mechanical responses to external stimulation. Direct contact with bodies, such as we experience in tactile perception, belongs to this stage. The role of the rudimentary nervous system is to facilitate action. What occurs is a reflex activity, not a "representation" of things. The sole difference between this stage and much later ones is that voluntary action became possible as a result of the evolution of the higher brain centers. But the difference is not one of kind, but only one of complication. Accordingly, since the nervous system is constructed from one end of the evolutionary scale to the other as a utilitarian device, we must conclude that perception, whose evolution is regulated by the evolution of the nervous system, is also directed toward action, not toward knowledge.

If that is so, why is human perception a conscious process, and why does everything happen as if consciousness were a product of brain activity? The reason is that human perception is normally "impregnated with memory images." It is possible to form a metaphysical concept of "pure perception" free from any admixture of memory. It is even possible, Bergson thought, to have such a pure perception, which he spoke of as an "intuition." But most of the time our perceptions are interlaced with memories; conversely, a memory becomes actual by being embodied in some perception. The convergence that takes place accounts for the fact that perceptual images (objects perceived) have a "subjectivity." We become conscious of them. This phenomenon has a biological significance, for in humans, and in higher organisms generally, perception is predominantly directed toward distant objects spread over a wide field. These objects have a great many potential effects on action. One way an organism has of adapting to this situation is to anticipate the effects by "reflecting" possible lines of action from its body to the distant objects. This gives the organism a biological advantage by putting it in a position where it can select a course of action that will serve its needs. Thus the world is consciously perceived by us; but it is not a different world from the one that antedated our perception. It is the same world related to our needs and intentions.

Body and mind, then, are united in the selective act of perception. The body contributes perceptive centers that respond to the influences of environing bodies. The mind contributes appropriate memory images that give to what is perceived a completed, meaningful form. There is no "constructing" of the external world out of subjective impressions; no "inferring" of the existence of that world from ideas in the mind; no positing of things in themselves that are beyond the limits of possible experience. By interpreting physical things as images, Bergson was able to regard the material world as directly perceivable. Traditional idealism was therefore repudiated. Yet a partial concession to idealism was made by calling things

"images." This term implies a rejection of the realist's view that things consist only of material particles, or of primary qualities, or of some hidden substance. Things have all the qualities they are perceived to have. A partial concession to realism was made by admitting that the totality of perceived things, past, present and future, must always be a small fragment of material reality. The upshot is a doctrine, intermediate between idealism and realism, that combines, Bergson contends, what is sound in each and discards what is unsound.

Body and mind are above all united in real duration, for perception is an event in the concrete present, and the present is no geometrical point or "knife edge" separating past from future. It is a continuous flowing, an "invisible progress of the past gnawing into the future." Perceptual acts are intrinsically temporal and dynamic. Yet the world we come to know by means of them is not a flux. It has a relative stability. Our concepts often refer to things that remain much the same for long periods. These things may have fixed position, sharp outlines, and clearly marked qualities. In view of what has been said about perception, how are such facts accounted for? The reply involves Bergson's conception of the intellect and its functioning.

THE INTELLECT AND THINGS

The evolution of the human species gave rise to the capacity for conceptual or rational thought. This capacity is traditionally referred to as the intellect. Its origin, Bergson contended, was conditioned by several circumstances. First, man is one of the social animals, and effective action in human societies requires some use of rational thought. Second, man is a tool-using and tool-making animal. These activities could not advance far without fostering conceptualization. Third, man is an animal who invents and uses language. This powerful instrument of communication stimulated the development of intellect, and was in turn profoundly influenced by it. Here again the aim was to promote community of action. Thus, both in origin and in function, the intellect is a practical capacity. It is no more speculative than is perception.

By using his intellect, civilized man has produced a vast body of knowledge about the world. Is not much of this knowledge speculative, in the sense of being a cognitive reflection of the world as it really is? Bergson held that this is not so. Since the intellect is practical, its products must be instrumental to action, not mirrorlike reflections. Concepts, even when they belong to advanced theories in the sciences, are still pragmatic devices. For scientific knowledge is directed toward prediction and control of events, being in this respect an extension of commonsense knowledge. The technological triumphs of modern man provide the clue to the proper understanding of his intellectual powers.

Because of its practical orientation, the intellect functions in a characteristic way. It treats whatever it deals with in spatial terms, as if the latter were a three-dimensional body. Ordinary language is pervaded by spatial metaphors; and scientific theories, especially those of physics, make great use of geometrical models. The operations of our intellect, especially in science, "tend to geometry, as to the goal where they find their perfect fulfilment" (*Creative Mind*, Introduction II). Again, the intellect has an inherent tendency to break up whatever it deals with into homogeneous units. A whole can be understood only by analyzing it in terms of uniform parts. This tendency is reflected in the predominance of measuring operations and instruments, such as clocks, scales, and yardsticks, in civilized societies. Furthermore, the intellect is at home only when dealing with what is static, fixed, immobile.

Hence, in seeking to understand the phenomenon of motion, the intellect has recourse to immobile units, such as points of space or instants of time, out of which motion is reconstructed. Bergson spoke of "the cinematographical method" of the intellect, likening it to a movie camera that translates motion into a series of static "frames." An important consequence of this is that the intellect is committed to the use of formal logic and mathematics, both of which supply unchanging structures for thought. Finally, when something comes into existence or ceases to exist, the intellect interprets what happens as a rearranging of constituent elements. This means that the arising of something absolutely new, the creation of novelty, cannot be admitted by rational thought. Even growth and evolution must be understood as new arrangements of old parts.

It is now possible to explain why the world external to us consists of relatively stable, discrete things. The intellect, functioning in its characteristic way, is responsible. It "breaks up," "cuts up," or "carves up" matter into distinct and separate objects so as to promote the interests of action. Presumably, the operation requires the collaboration of perception, although Bergson did not make the point clear. He also failed to make clear whether the intellect is perfectly free in carving out individual things, or whether it has to follow certain lines of cleavage in the intrinsic structure of matter. Sometimes he talked as if the external world of things had been "fabricated" by the

intellect's imposing form on a featureless, material flux. At other times, he implied that the intellect "carves nature at the joints," following "the lines which mark out the boundaries of real bodies or of their real elements." In one place he even stated that "matter is primarily what brings division and precision" into things; but this can hardly be construed as an acceptance of the doctrine that matter is the principle of individuation. Despite these obscurities, Bergson's position entails that the intellect is necessary, if not sufficient, for the "individuating" of things in space.

This requirement is relevant, of course, only to things of which we have conceptual knowledge. What is its bearing on the knowledge each of us has of his own body? Here a further obscurity arises. Bergson declared that we know our body in two ways, externally by perception and internally by affection. But since at the level of affection the intellect is not involved, it would appear to follow that the object known cannot be a separate, individual thing. Nevertheless, Bergson did speak of the central image, "distinct from all others," that each of us identifies as his body. What determines its distinct individuality? In *Matter and Memory* he remarked that "our needs … carve out, within this continuity [of the perceptible world], a body which is to be their own." This is a puzzling remark, because often the body is what has the needs, and hence it can scarcely be "carved out" by them. It may be that the living human body, unlike inanimate bodies, has an individuality that does not depend on the functioning of the intellect. Or it may be that the obscurity here originates in Bergson's doctrine about what the intellect knows and what can be known only by intuition.

INTUITION AND INTELLECT

Alongside the capacity for conceptual thought, there exists in humans a capacity that Bergson called "intuition." Both capacities are the result of evolution, but the second is derived from instinct, the type of biological activity most elaborately manifested in the social insects. Instinctive activity has consciousness "slumbering" within it, and evolution has awakened the consciousness in humankind. Intuition for Bergson is "instinct that has become disinterested, self-conscious, capable of reflecting upon its object and of enlarging it indefinitely." Since it is disinterested, the capacity is detached from the demands of action and of social life. It is like a painter's power of seeing the world just as it is presented to him in pure perception. But instead of yielding an aesthetic experience, intuition yields knowledge. Hence, it is of profound importance for the philosopher.

In his *Introduction to Metaphysics*, Bergson emphasized the immediate, nonconceptual character of intuition, envisaging it as a direct participation in, or identification with, what is intuited. In the case of the external world, intuition is an act "by which one is transported into the interior of an object in order to coincide with what there is unique and consequently inexpressible about it." In the case of the self, intuition is an immersion in the indivisible flow of consciousness, a grasping of pure becoming and real duration. The result is "knowledge which is contact and even coincidence." Unlike the intellect, which remains outside what it knows, requires symbols, and produces knowledge that is always relative to some viewpoint, intuition enters into what it knows, dispenses with symbols, and produces knowledge that is absolute.

Bergson subsequently modified this doctrine in certain respects. He came to emphasize the cogitative character of intuition instead of its immediacy, and even spoke of it as a mode of thinking. As such, it is not a spontaneous flash of insight but an act that is engendered by mental effort. To achieve an intuition, we must turn our attention away from its natural concern with action. This act demands concentration of thought. Even when we are successful, the results are impermanent. Yet the intellect can effect a partial communication of the results by using "concrete ideas," supplemented by images. "Comparisons and metaphors will here suggest what cannot be expressed." Consequently, the knowledge attained by intuition is not altogether ineffable. Nor is it, in the strict sense, absolute, for intuition is a progressive activity that can widen and deepen its scope indefinitely. Its limits cannot be fixed a priori. These modifications were related to changes in Bergson's conception of the roles of metaphysics and the natural sciences.

THE NATURAL SCIENCES AND METAPHYSICS

The natural sciences are for Bergson a typical achievement of the intellect, and they therefore reflect a limitation in the intellect's functioning. This limitation emerges when the sciences form their conceptions of time and motion. In each case a static abstraction is produced. Time is conceived as what clocks measure in spatially discrete units. Motion is conceived as a succession of fixed positions on a linear path. Both abstractions are practically useful, but they falsify the nature of time and motion as concretely experienced by ignoring the crucial element of becoming. This falsification is inherent in the intellect's way of working. By its very nature, the intellect

is equipped to handle only what is repetitive and routine; real becoming baffles it. Hence the sciences have a severe disability built into them. Moreover, as the ancient philosopher Zeno of Elea first pointed out, conceptual thought runs into contradictions or "paradoxes" whenever it tries to give a thorough analysis of motion. These paradoxes, although designed by Zeno for a different purpose, show, according to Bergson, that the scientific concept of motion is basically incoherent. The conclusion must be that the sciences can never provide a complete and adequate account of the universe. They need to be supplemented by some other discipline.

An obvious choice would seem to be metaphysics, but classical metaphysics is equally a creation of the intellect and suffers from the same disability as the sciences. Metaphysicians, with a few exceptions like Heraclitus, have misconstrued change and failed to give it the priority it actually has in the world. They have regarded being as ultimate, and becoming as derivative. Accordingly, metaphysical theories have been based on such concepts as the indestructible atoms of Democritus, the eternal forms of Plato, or the fixed categories of Immanuel Kant. These concepts illustrate the intellect's addiction to unchanging units that are mechanically combined or separated according to the rules of logic. Neither time nor change can be understood when so approached. The constructions of metaphysics are as inadequate here as those of science, without the latter's usefulness.

Classical metaphysics has also mistakenly supposed that an all-embracing "system" can be constructed, bringing within its scope not only what is actual but also what is possible. This idea rests on a fallacious assumption that there is a "realm of possibility" over and above the realm of actuality. The belief in possibles that would be realized by acquiring existence is an illusion of the intellect, designed to exclude the notion of absolute novelty. "Let us have done," Bergson urged, "with great metaphysical systems embracing all the possible and sometimes even the impossible!"

By following this course, we shall automatically get rid of a number of pseudo problems that classical metaphysicians have generated. They have asked, for instance, why something exists rather than nothing. This has seemed a sensible question because they could always add, "There could be nothing." Bergson replied that the sentence "There could be nothing" has no meaning. "'Nothing' is a term in ordinary language which can only have meaning in the sphere proper to man, of action and fabrication." The term designates the absence of what we are seeking in the world around us. It can be properly used only because many things already exist. To oppose "nothing" in an absolute sense to existence is to embrace a pseudo idea and engender pseudo problems.

These criticisms do not imply that metaphysics is to be rejected, for Bergson proposed to redefine metaphysics and provide it with a new method. Instead of employing the intellect, it is to employ intuition. This is the theme of the *Introduction to Metaphysics*. In elaborating it, Bergson sometimes seemed to be saying that since intuition alone provides knowledge of the real, the intellect is restricted to knowledge of appearances. It would follow from this that metaphysics is a discipline superior to the natural sciences. Indeed, from a philosophical standpoint the sciences are cognitively worthless because they can say nothing about reality. The impression was thus created that Bergson's outlook was "antiscientific." In later writings he endeavored to correct this impression by urging that metaphysics and the sciences must be coordinate and equal in value. Both are concerned with the real, the sciences with the domain of matter, metaphysics with the domain of spirit. Moreover, the knowledge that each gains is capable of indefinite expansion, and can approach completeness as an ideal limit. It was in this connection that Bergson seems to have revised his doctrine of intuition, closing the gap between it and the intellect without obliterating the distinction between the two. His objective was to formulate a philosophy that would submit to the control of science and that could in turn enable science to progress. The disciplines would then have a common frontier. In adopting the method of intuition, metaphysics is able to supplement the sciences by giving a true account of duration, of becoming, and even of evolution.

MECHANISTIC AND CREATIVE EVOLUTION

Bergson was born in the same year that *The Origin of Species* was published, and the revolutionary implications of this work permanently affected his thought. He accepted the historical reality of evolution, but rejected attempts to explain it in mechanistic or materialistic terms. Hence he criticized Darwin's explanation, and also the less influential explanations of the Chevalier de Lamarck, Theodor Eimer, and Spencer. In place of them he advanced a doctrine that owed much to the tradition of European and especially French vitalism, and at the same time drew inspiration from Plotinus. The result was a vision of the cosmos going far beyond the facts of biology, though purportedly based on them. These matters

were presented in *Creative Evolution*, Bergson's most famous book.

Darwin explained the evolutionary process by supposing that in every population of organisms there occur random variations that have different degrees of adaptive value. The variations having maximum value for the survival and reproduction of the organisms are "naturally selected"; that is, they are preserved and transmitted to subsequent generations, while the other variations are eliminated.

Bergson argued that this explanation failed to account for a number of facts. A multicellular animal, or an organ like the vertebrate eye, is a functional whole made up of coordinated parts. If just one or a few of the parts happened to vary independently of the rest, the functioning of the whole would be impaired. Since evolution has occurred, we must suppose that at each stage all the parts of an animal and of its complex organs have varied contemporaneously so that effective functioning was preserved. But it is utterly implausible to suppose, as Darwin did, that such coadapted variations could have been random, for then their coadaptation would remain a mystery. Some agency other than natural selection must have been at work to maintain continuity of functioning through successive alterations of form.

Another fact that Darwinism failed to explain is why living things have evolved in the direction of greater and greater complexity. The earliest living things were simple in character and well adapted to their environments. Why did the evolutionary process not stop at this stage? Why did life continue to complicate itself "more and more dangerously"? To appeal to the mechanism of selection for an answer was, Bergson thought, insufficient. Something must have driven life on to higher and higher levels of organization, despite the risks involved.

Darwin's predecessor Lamarck avoided the idea of random variations by supposing that variations were caused by the "effort" exerted by individuals in adapting to the environment. Bergson considered this a more adequate explanation than the Darwinian. Yet it involved accepting the principle that acquired characteristics are transmitted from one generation to the next, and empirical evidence is heavily against this. Furthermore, the Lamarckian notion of a conscious "effort" is too limited to serve as an explanatory device. It could perhaps operate in the case of animals but hardly in the case of plants or microorganisms. To make the notion work, it must be broadened and deepened. Similarly, Eimer's appeal to orthogenesis; that is, to an inner principle that directs the course of evolution, has merit if interpreted nonmechanistically, but not if interpreted, as Eimer did, in physicochemical terms.

The synthetic philosophy of Spencer also had merit in so far as it sought to extend the evolutionary conception to the universe at large. Yet because Spencer relied exclusively on the intellect, and because he subscribed to the false idea that philosophy can be a super science, Spencer failed to do justice to real duration and to the creation of novelty. He held that evolution is due to combinations of matter and motion. This makes his philosophy a thinly disguised version of mechanical materialism, which reconstructs evolution "with fragments of the evolved."

To obtain a true understanding of the evolutionary process, the findings of biology must be supplemented, Bergson thought, by the findings of metaphysics. The chief clue is found in what intuition reveals of our own inner nature as living beings; we are typical constituents of the universe, and the forces that work in us also work in all things. When we focus upon what intuition discloses of ourselves, we find not only continuous becoming and real duration, but also a consciousness of a vital impetus (*élan vital*), of our own evolution in time. We are thus led to the idea of "an *original impetus* of life" (*un élan original de la vie*) that pervades the whole evolutionary process and accounts for its dominant features. Accordingly, the history of life is to be understood in creative, not mechanistic, terms.

THE VITAL IMPETUS AND EVOLUTION

Bergson's doctrine of the vital impetus is speculative, although often formulated as if it were a report of an established fact. The impetus is declared to be "a current of consciousness" that has penetrated matter, given rise to living bodies, and determined the course of their evolution. The current passes from one generation to the next by way of reproduction—in bisexual organisms, by way of the reproductive cells. The vital impetus is the cause of variations that accumulate and produce new species. It coordinates the appearance of variations so as to preserve continuity of functioning in evolving structures. And it carries life toward ever higher complexity of organization. Strictly speaking, the impetus does not generate energy of its own, over and above that already present in matter. What it does is "to engraft on to the necessity of physical forces the largest possible amount of indetermination." This indetermination is evident in the contingency and creativity that have characterized the history of life. At every stage the impetus has been limited by recalcitrant matter. Hence, it is always seeking to transcend the

stage it has reached and always remains inadequate to what it tries to produce.

The earliest living things were physicochemical systems into which the vital impetus "insinuated itself." Its potentialities could be realized only minimally in these systems. Consequently, it divided so that life moved forward in several quite different directions. One direction was taken by the plants, another by the insects, and a third by the vertebrates. The three directions illustrate respectively the predominance of stability, instinct, and intelligence. No predetermined plan or purpose was involved in all this. Bergson expressed as much opposition to the doctrine of radical finalism as he did to mechanism. Both doctrines deny that there has been an unforeseeable creation of forms, that these forms involve discontinuous "leaps," and that real duration is a cumulative, irreversible flow. Yet although the vital impetus is not finalistic, it does engender progress. A perfecting of functions has occurred through successive stages. An increasing realization of consciousness has also occurred.

This last contention made it difficult for Bergson to maintain an opposition to finalism, for it is in man that consciousness has been most fully realized. Here the vital impetus has found its most adequate expression as intelligence. It has likewise achieved genuine freedom by at last making matter its instrument. There was in fact "a sudden leap from the animal to man." Hence in *Creative Evolution* Bergson said that man might be considered the reason for the existence of the entire organization of life on our planet. He immediately qualified this statement by adding that it is "only a manner of speaking." We should not think that humanity was "prefigured" in the evolutionary process from the beginning.

By the time he wrote the essay that became the "Second Introduction" to *The Creative Mind*, Bergson was more forthright. He there stated categorically that the appearance of humans is the *raison d'être* of life on the earth. In *The Two Sources of Morality and Religion* he also contended that it is humankind, "or some other being of like significance, which is the purpose of the entire process of evolution." This contention seems very close to finalism. Nevertheless, Bergson continued to insist that the appearance of man was in no sense predetermined, though "it was not accidental, either." Terrestrial evolution might have produced some other being "of the same essence." Such beings have doubtless arisen elsewhere, for Bergson thought that the vital impetus animates innumerable planets in the universe. The impetus is thus not limited to the earth; creative evolution is a cosmic process.

This contention is not argued for in any detail. As so often in his writings, Bergson tried to make the contention acceptable by means of analogies. He likened the vital impetus to steam escaping at high pressure through the cracks in a container. Jets gush out unceasingly, the steam condenses into drops of water, and the drops fall back to the source. Each jet and its drops represent a world of matter animated by life. A small part of the jet remains uncondensed for an instant, and makes an effort to raise the drops that are falling. But it succeeds at most in retarding their fall. So the vital impetus achieves a moment of freedom at its highest point, in humans. It might be inferred from this analogy that matter is not something sui generis, but is rather the lowest form assumed by the outpouring of spirit. Matter and spirit, however, were repeatedly described by Bergson as coexistent and interdependent.

GOD AND THE MYSTICS

The religious aspect of Bergson's outlook became increasingly pronounced toward the close of his life. Even in *Creative Evolution* he had spoken of the vital impetus as a "supra-consciousness" to which the name "God" might be attached. But this is very different from the conception of traditional Western theology. For if God is identical with the vital impetus, then he is pure activity, limited by the material world in which he is struggling to manifest himself. He is neither omnipotent nor omniscient. God "has nothing of the already made," but is ceaselessly changing. In *The Two Sources of Morality and Religion*, Bergson moved somewhat closer to the Christian position; he affirmed that God is love and the object of love. There is also a divine purpose in the evolutionary process. Evolution is nothing less than God's "undertaking to create creators, that He may have, besides Himself, beings worthy of His love."

The discovery of this purpose and of the reality of God cannot be made by the intellect. It can be made only by the sort of intuition that is the mystical experience. For the vital impetus, Bergson contended, is communicated "in its entirety" to exceptional persons. These are the mystics who achieve contact and partial coincidence with the creative effort that "is of God, if it is not God Himself." This experience does not terminate in passivity, but leads to intense activity. The mystics participate in God's love for humankind. They are therefore impelled to advance the divine purpose by helping to complete the development of man. They want to make of humanity what it would straightway have become if humanity had been able to reach its final form without the aid of humans

themselves. The spirit of the mystics must become universal in order to ensure man's future evolution.

Bergson acknowledged that the biggest obstacle to the spread of the mystical spirit is the ceaseless struggle that most people must wage against the material conditions of life. Yet he did not believe that these conditions could be ameliorated by programs of political and economic reform devised by the intellect. Consequently, the most we can hope for at present is that the spirit of the mystics will be kept alive by small groups of privileged souls, "until such time as a profound change in the material conditions imposed on humanity by nature should permit, in spiritual matters, of a profound transformation." The mystics, through their experience of love, will keep open a trail along which the whole of humanity can eventually pass.

CLOSED AND OPEN SOCIETIES

Since man is a social animal, his future evolution will be accelerated or retarded by the sort of group in which he lives. Bergson discussed this question in *The Two Sources of Morality and Religion*, where he drew a distinction between a society that is "closed" and one that is "open," describing in each case corresponding types of religion and of morality.

A closed society is one dominated by the routine and mechanical. It is resistant to change, conservative, and authoritarian. Its stability is achieved by increasing its self-centeredness. Hence, conflict with other self-centered groups, often involving war, is a condition of its preservation. Internal cohesiveness is secured by a closed morality and a closed religion. Bergson's analysis was influenced by the sociological doctrines of Émile Durkheim. Closed morality is static and absolutistic; closed religion is ritualistic and dogmatic. Both institutions exert pressure on individuals to accept the standard practices of the community. Spontaneity and freedom are reduced to a minimum. Conformity becomes the prime duty of the citizen. There is an obvious analogy between such a society and the repetitive mechanisms dealt with by the intellect. Indeed, Bergson regarded closed societies as in large measure the intellect's products.

The existence of a multiplicity of closed societies on the earth is an obstacle to human evolution. Accordingly, the next development in humankind requires the establishment of an open society. Instead of being limited, it will embrace all humankind; instead of being static, it will be progressive; instead of demanding conformity, it will encourage the maximum diversity among individuals. Its moral and religious beliefs will be equally flexible and subject to growth. Religion will replace the stereotyped dogmas elaborated by the intellect with the intuition and illumination now achieved by the mystics. The spread of the mystical spirit must ultimately create an open society whose freedom and spontaneity will express the divine élan which pervades the universe.

Bergson's outlook had a marked influence on the thought and literature of Europe. His gifts as a writer, his ingenuity in constructing vivid analogies, and his flair for describing the subtleties of immediate experience—"true empiricism," as he called it—contributed to the popularity of his work, as did the impressive use that he made of the biological and psychological ideas of his time. On the other hand, critics have contended that many of his doctrines are vague and ill-supported by arguments. Too often, it is said, rhapsodic formulations are offered where there ought to be sustained logical analysis. There is, for instance, no clear statement of how real duration, the flow of consciousness, and the vital impetus are related. Are these separate processes, or just distinguishable aspects of one process? Does matter have an independent status, or is it simply a "devitalized" form of the *élan vital*? Such questions are difficult, if not impossible, to answer. Many critics have also deplored the encouragement that Bergson's doctrine of the intellect gave to the advocates of irrationalism and the cruder versions of pragmatism. Yet when all these criticisms have been made, the Bergsonian heritage remains an important element in twenty-first-century philosophy.

See also Aesthetic Experience; Continuity; Darwin, Charles Robert; Darwinism; Determinism and Freedom; Durkheim, Émile; Evolutionary Theory; Fouillée, Alfred; Idealism; Intuition; Irrationalism; Kant, Immanuel; Lamarck, Chevalier de; Leucippus and Democritus; Maine de Biran; Materialism; Memory; Metaphysics, History of; Mind-Body Problem; Mysticism, History of; Nothing; Philosophy of Language; Philosophy of Science, History of; Plato; Plotinus; Ravaisson-Mollien, Jean Gaspard Félix; Realism; Time; Vitalism; Zeno of Elea.

Bibliography

WORKS BY BERGSON

Quid Aristoteles de Loco Senserit. Paris: F. Alcan, 1889. Bergson's doctoral thesis. Translated by Robert Mossé-Bastide as "L'Idée de lieu chez Aristote," in *Les Études bergsoniennes*. Paris: A. Michel, 1949. Vol. II.

Essai sur les donnés immédiates de la conscience. Paris, 1889. Translated by F. L. Pogson as *Time and Free Will: An Essay on the Immediate Data of Consciousness*. New York: Macmillan, 1910.

Matière et mémoire. Paris, 1896. Translated by Nancy Margaret Paul and W. Scott Palmer [pseud.] as *Matter and Memory*. New York: Macmillan, 1912.

Le rire. Paris, 1900. Translated by Cloudesley Brereton and Fred Rothwell as *Laughter. An Essay on the Meaning of the Comic*. New York: Macmillan, 1911.

"Introduction à la métaphysique." *Revue de la métaphysique et de morale* 11 (January 1903): 1–36. Translated by T. E. Hulme as *Introduction to Metaphysics*. New York: Putnam, 1912.

L'évolution créatrice. Paris: F. Alcan, 1909. Translated by Arthur Mitchell as *Creative Evolution*. New York: Henry Holt, 1911.

L'énergie spirituelle. Paris: F. Alcan, 1919. Translated by H. Wildon Carr as *Mind-Energy*. New York: Henry Holt, 1920. A collection of essays.

Durée et simultanéité. Paris: F. Alcan, 1922; 2nd ed. with 3 appendices, 1923. On aspects of the theory of relativity; not included in centenary edition of Bergson's works.

Les deux sources de la morale et de la religion. Paris: F. Alcan, 1932. Translated by R. A. Audra and Cloudesley Brereton as *The Two Sources of Morality and Religion*. New York: Henry Holt, 1935.

La pensée et le mouvant. Paris: F. Alcan, 1934. Translated by Mabelle L. Andison as *The Creative Mind*. New York: Philosophical Library, 1946. A collection of essays.

Écrits et paroles. Edited by R. M. Mossé-Bastide. 3 vols. Paris, 1957–1959. Preface by Édouard LeRoy.

Oeuvres. Édition du centenaire. Paris: Presses universitaires de France, 1959. Annotated by André Robinet, introduction by Henri Gouhier.

WORKS ON BERGSON

English

Carr, Herbert Wildon. *The Philosophy of Change*. New York, 1912.

Chevalier, Jacques. *Henri Bergson*. New York: Macmillan, 1928.

Hanna, Thomas, ed. *The Bergsonian Heritage*. New York: Columbia University Press, 1962. Articles on Bergson's thought by various scholars.

Höffding, Harald. *Henri Bergson's Filosofi. Karacteristik ag Kritik*. Copenhagen, 1914. Translated by Alfred C. Mason in *Modern Philosophers and Lectures on Bergson*. London, 1915.

Huxley, Julian. *Essays in Popular Science*. New York: Knopf, 1927. Claims that Bergson's vitalism was based on dubious factual material.

LeRoy, Édouard. *Une Philosophie nouvelle: Henri Bergson*. Paris, 1912. Translated by Vincent Benson as *The New Philosophy of Henri Bergson*. New York: Henry Holt, 1913.

Lindsay, A. D. *The Philosophy of Henri Bergson*. London, 1911.

Ruhe, Algot, and Nancy Margaret Paul. *Henri Bergson*. London: Macmillan, 1914.

Russell, Bertrand. *Our Knowledge of the External World*. London, 1914. Ch. 1. Highly critical.

Russell, Bertrand. *The Philosophy of Bergson*. Cambridge, U.K.: Bowes and Bowes, 1914. Highly critical.

Santayana, George. *Winds of Doctrine*. New York: Scribners, 1913. Highly critical.

Scharfstein, Ben-Ami. *Roots of Bergson's Philosophy*. New York: Columbia University Press, 1943.

Stephen, Karin. *The Misuse of Mind*. London: K. Paul, Trench, Trubner, 1922.

Stewart, J. McK. *A Critical Exposition of Bergson's Philosophy*. London: Macmillan, 1911.

French

Delhomme, Jeanne. *Vie et conscience de la vie: Essai sur Bergson*. Paris, 1954.

Les études bergsoniennes. Paris: A. Michel, 1948–1959. Vols. I–V.

Husson, Léon. *L'intellectualisme de Bergson*. Paris: Presses universitaires de France, 1947.

Jankélévitch, Vladimir. *Henri Bergson*. Paris, 1959.

Marietti, Angèle. *Les formes du mouvement chez Bergson*. Paris, 1957.

Maritain, Jacques. *La philosophie bergsonienne*. Paris: M. Rivière, 1930.

OTHER WORKS OF INTEREST

Ansell-Pearson, Keith. *Philosophy and the Adventure of the Virtual: Bergson and the Time of Life*. London; New York: Routledge, 2002.

Bergson, Henri. *Durée et simultanéité*. 7th ed. Paris: Presses universitaires de France, 1968.

Bergson, Henri. *An Introduction to Metaphysics: The Creative Mind*. Totowa, NY: Littlefield, Adams, 1975.

Bergson, Henri. *Mind-Energy: Lectures and Essays*. Westport, CT: Greenwood Press, 1975,.

Bergson, Henri. *The Two Sources of Morality and Religion*. Westport, CT: Greenwood Press, 1974.

Bergson, Henri, and André Robinet. *Correspondances*. Paris: Presses universitaires de France, 2002.

Bergson, Henri, and André Robinet. *OEuvres*. 5th ed. Paris: Presses universitaires de France, 1991.

Capek, Milic. *Bergson and Modern Physics. A Reinterpretation and Re-evaluation*. Dordrecht: Reidel, 1971.

Deleuze, Gilles. *Bergsonism*. New York: Zone Books, 1988.

Fink, Hilary L. *Bergson and Russian Modernism, 1900–1930*. Evanston, IL: Northwestern University Press, 1999.

Gallagher, Idella J. *Morality in Evolution: The Moral Philosophy of Henri Bergson*. The Hague: Nijhoff, 1970.

Gunter, P. A. Y., comp. *Bergson and The Evolution of Physics*. Knoxville: University of Tennessee Press, 1969.

Gunter, P. A. Y. *Henri Bergson: A Bibliography*. 2nd rev. ed. Bowling Green, OH: Philosophy Documentation Center, Bowling Green State University, 1986.

Kolakowski, Leszek. *Bergson*. Oxford, Oxfordshire; New York: Oxford University Press, 1985.

Lacey, A. R. *Bergson*. London; New York: Routledge, 1989.

Maritain, Jacques. *Bergsonian Philosophy and Thomism*. New York: Greenwood Press, 1955, 1968.

Mullarkey, John. *Bergson and Philosophy*. Notre Dame, IN: University of Notre Dame Press, 2000.

Mullarkey, John. *The New Bergson*. Manchester; New York: Manchester University Press, 1999.

Pilkington, Anthony Edward. *Bergson and His Influence: A Reassessment*. Cambridge; New York: Cambridge University Press, 1976.

T. A. Goudge (1967)
Bibliography updated by Michael J. Farmer (2005)

BERKELEY, GEORGE
(1685–1753)

George Berkeley, the Irish philosopher of English ancestry, and Anglican bishop of Cloyne, was born at Kilkenny, Ireland. He entered Trinity College, Dublin in 1700 and became a fellow in 1707. In 1709 he published his first important book, *An Essay towards a New Theory of Vision.* This was well received, and a second edition appeared in the same year. The following year *A Treatise concerning the Principles of Human Knowledge,* Part 1, was published. This is the work in which Berkeley first published his immaterialist philosophy, and although it made him known to some of the foremost writers of the day, its conclusions were not taken very seriously by them. In 1713 Berkeley went to London and there published the *Three Dialogues between Hylas and Philonous,* a more popular statement of the doctrines of the *Principles.* While in London, Berkeley became acquainted with Joseph Addison, Jonathan Swift, Alexander Pope, and Richard Steele and contributed articles to Steele's *Guardian,* attacking the theories of the freethinkers. He traveled on the Continent in 1713–1714 (when he probably met and conversed with Nicolas Malebranche) and again from 1716 to 1720. During this tour he lost the manuscript of the second part of the *Principles,* which he never rewrote. Toward the end of the tour, he wrote a short essay, in Latin, titled *De Motu,* published in London in 1721, criticizing Isaac Newton's philosophy of nature and Gottfried Wilhelm Leibniz's theory of force. In 1724 Berkeley was made dean of Derry.

About this time, Berkeley began to prepare a project for establishing a college in Bermuda, at which not only the sons of American colonists but also Indians and Negroes were to receive a thorough education and be trained for the Christian ministry. Having obtained promises of subscriptions from many prominent people, Berkeley promoted a bill, which was passed by Parliament, providing for considerable financial help from the government. In 1728, before the money was forthcoming, Berkeley, who had just married, left for Rhode Island, where he intended to establish farms for supplying food for the college. He settled in Newport, but the grant never came; and in 1731, when it was clear that the government was diverting the money for other purposes, Berkeley had to return home. While in Newport, however, Berkeley had met and corresponded with the Samuel Johnson who later became the first president of King's College, New York (now Columbia University). Johnson was one of the few philosophers of the time to give close attention to Berkeley's philosophical views, and the correspondence between him and Berkeley is of considerable philosophical interest. While he was in Newport, Berkeley also wrote *Alciphron,* a series of dialogues in part developed from the articles he had written for the *Guardian,* directed against the "minute philosophers," or freethinkers. This was published in 1732.

Berkeley was in London from 1732 to 1734 and there wrote *The Analyst* (1734), a criticism of Newton's doctrine of fluxions and addressed to "an infidel mathematician." This and *A Defence of Free-Thinking in Mathematics* (1735) aimed at showing that the mathematicians so admired by freethinkers worked with concepts that could not withstand close scrutiny, so that the confidence given to them by "the philomathematical infidels of these times" was unjustified. It is not surprising that Berkeley was made bishop of Cloyne, Ireland, in 1734.

Berkeley carried out his episcopal duties with vigor and humanity. His diocese was in a remote and poor part of the country, and the problems he encountered there led him to reflect on economic problems. The result was *The Querist* (1735–1737), in which he made proposals for dealing with the prevailing idleness and poverty by means of public works and education. He also concerned himself with the health of the people and became convinced of the medicinal value of tar water. In 1744 he published *A Chain of Philosophical Reflexions and Inquiries concerning the Virtues of Tar-Water, and divers other Subjects connected together and arising from one another.* When the second edition appeared in the same year, the title *Siris,* by which the book is now known, was added. Much of the book is concerned with the merits of tar water, but Berkeley passed from this subject to the causes of physical phenomena, which, he held, cannot be discovered in the phenomena themselves but must be sought for in the Divine activity. This is in line with his earlier views, but some readers, on the basis of his admiring references to Plato and the Neoplatonists, have considered that by this time he had considerably modified his original system. The *Siris* was Berkeley's last philosophical work. He died suddenly in Oxford nine years later.

An account of Berkeley's life and writings would be inadequate without some reference to his *Philosophical Commentaries.* A. C. Fraser discovered a series of notes by Berkeley on all the main topics of Berkeley's philosophy and published them in 1871 in his edition of Berkeley's works, under the title of *Commonplace Book of Occasional Metaphysical Thoughts.* It was later noticed that these notes had been bound together in the wrong order, and it has now been shown that they were written by Berkeley, probably in 1707–1708, while he was thinking out his *New Theory of Vision* and *Principles.* This work makes it

clear that Berkeley was already convinced of the truth of immaterialism before he published the *New Theory of Vision,* in which that view is not mentioned. The *Philosophical Commentaries* throw valuable light upon Berkeley's sources, bugbears, prejudices, and arguments.

MAIN THEMES OF BERKELEY'S PHILOSOPHY

Since the word *idealism* came into use in the eighteenth century, Berkeley has been known as a leading exponent of idealism, and even as its founder. He himself referred to his main view as "the immaterialist hypothesis," meaning by this that he denied the very possibility of inert, mindless, material substance. This description has some advantage over idealism in that it brings out Berkeley's radical opposition to materialism; whereas the opposite of idealism is realism, and there are grounds for doubting whether Berkeley intended to deny the realist contention that in perception people become directly aware of objects that persist unchanged when they cease to be perceived. Berkeley's fundamental view was that for something to exist it must either be perceived or else be the active being that does the perceiving. Things that are perceived he called "sensible things" or "sensible qualities," or, in the terminology he had borrowed from John Locke, "ideas." Sensible things or ideas, he held, cannot exist except as the passive objects of minds or spirits, active beings that perceive and will. As he put it in the *Philosophical Commentaries,* "Existence is *percipi* or *percipere,*" and he added "or *velle* i.e. *agere*"—existence is to be perceived or to perceive or to will, that is, to be active. Thus there can be nothing except active spirits on the one hand and passive sensible things on the other, and the latter cannot exist except as perceived by the former. This is Berkeley's idealism or immaterialism.

CRITICISM OF CONTEMPORARY SCIENCE. The above account of Berkeley's writings emphasizes their apologetic intent, an intent that can be seen in the subtitles of his major writings—that of the *Principles* is typical: *Wherein the chief causes of error and difficulty in the sciences, with the grounds of scepticism, atheism and irreligion, are inquired into.* It will be seen that "the chief causes of difficulty in the sciences" are also prominent. Berkeley considered that in the mathematics and natural sciences of his day insufficient attention was given to what experience reveals to us. Apart from Newton, the mathematicians were, he wrote in the *Philosophical Commentaries,* "mere triflers, mere Nihilarians." For example, they conceived of lines as infinitely divisible, but this is not only absurd, it could be maintained only by men who "despised sense." Thus Berkeley regarded himself as protesting against the excesses of uncontrolled rationalism. Hence he put forward a most antirationalistic view of geometry, although he never developed its implications very far. Similarly he thought that the natural philosophers deluded themselves with words when they tried to explain the physical world in terms of attractions, forces, and powers. Natural science, as he understood it, was descriptive rather than explanatory and was concerned with correlations rather than with causes. He thus sketched out a view of science that was revived and developed by nineteenth-century and twentieth-century positivists.

SENSIBLE QUALITIES ARE THE SIGNS OF GOD'S PURPOSE. Berkeley's positivism, however, was confined to his account of natural science. The order of phenomena, he held, was willed by God for the good of created spirits. In deciphering the conjunctions and sequences of our sense experience we are learning what God has decreed. Thus sensible qualities are the language in which God speaks to us. In the third and fourth editions (1732) of the *New Theory of Vision* Berkeley said that the objects of sight are a divine visual language by which God teaches us what things are good for us and what things are harmful to us. In the *Alciphron,* published that same year, he argued that "the great Mover and Author of Nature constantly explaineth Himself to the eyes of men by the sensible intervention of arbitrary signs, which have no similitude or connexion with the things signified." We learn that certain visual ideas are signs of certain tactual ones, certain smells signs of certain colors, and so on. There is no necessity about this, any more than things necessarily have the names that convention assigns to them. Just as some sensible qualities are signs of others, so sensible qualities as a whole are signs of the purposes of God who "daily speaks to our senses in a manifest and clear dialect."

Thus, taken as a whole, Berkeley's philosophy is a form of immaterialism combined with an extreme antirationalist theory of science. The regularities between phenomena are regarded as evidence for, and as signs of, God's purposes. Just as a man's words reveal his thoughts and intentions by means of the conventional signs of language, so the sensible order reveals God's will in phenomena that could have been ordered quite differently if he had so decided.

THE NEW THEORY OF VISION

Although Berkeley did not mention his immaterialism in *An Essay towards a New Theory of Vision,* this work throws important light upon his quarrel with the mathematicians and his rejection of the rationalist point of view. It contains, too, an interesting statement of what Berkeley then thought about geometry. Furthermore, the *Essay* helps us to see, from what Berkeley said about the objects of vision, how he came to the view that sensible qualities cannot exist "without the mind." Among the main contentions of the book is the claim that distance or "outness" is not immediately perceived by sight; it is "suggested" in part by the sensations we get in moving our eyes but mainly by association with the ideas of touch. According to Berkeley, we see the distance (and size) of things only in the sense in which we see a man's shame and anger. We see his face, and the expression on it suggests to us how he is feeling. In themselves, shame and anger are invisible. Similarly, we see shapes and colors, which are signs of what we would touch if we were to stretch out our hands, but distance itself is no more seen than anger is. In expounding this view, Berkeley developed the thesis that the objects of sight and touch are utterly disparate, so that no feature of the one can have more than a contingent connection with any feature of the other.

DESCARTES'S THEORY OF THE PERCEPTION OF DISTANCE. Consideration should first be given to Berkeley's criticisms of an important geometrical account of how distance is perceived and assessed, the account given by René Descartes in his *Dioptrics* (1637). In this work Descartes referred to six "qualities we perceive in the objects of sight," namely, light, color, shape, distance, magnitude, and situation. Descartes argued that one of the ways in which men ascertain the distance of objects is by means of the angles formed by straight lines running from each of their eyes and converging at the object seen. He illustrated this by reference to a blind man with a stick (the length of which he does not know) held in each hand. When he brings the points of the sticks together at the object, he forms a triangle with one hand at each end of the base, and if he knows how far apart his hands are, and what angles the sticks make with his body, he can, "by a kind of geometry innate in all men" know how far away the object is. The same geometry would apply, Descartes argued, if the observer's eyes are regarded as ends of the base of a triangle, and straight lines from them are regarded as converging at the object. The more obtuse the base angles formed by the lines running from this base and converging at the object, the farther away the object must be; the more acute these angles, the nearer the object must be. Berkeley put the matter somewhat differently from Descartes, pointing out that according to the latter's view the more acute the angle formed at the object by the lines converging from the eyes, the farther away it must be; the more obtuse this angle, the nearer the object must be. It is important to notice that this "must" is the "must" of mathematical necessity. From what Descartes said, it is necessarily the case that the more acute this angle is, the farther away the object is; the more obtuse the angle, the nearer the object. "Nearer" and "farther" logically depend upon the obtuseness or acuteness of the angle. In criticizing this view, therefore, Berkeley was criticizing the view that distance is known a priori by the principles of an innate geometry according to which we know that the distance of the object must vary in accordance with the angle made at the object by straight lines converging there from the eyes of the observer.

BERKELEY'S CRITICISM OF DESCARTES. Against Descartes's view Berkeley brought a complex argument that for purposes of exposition, is here broken up into three parts. The first is that people who know nothing of the geometry of the matter can nevertheless notice the relative distance of things from them. This is not very convincing, for Descartes obviously thought that the geometry he regarded as "innate in all men" might be employed by them without their having reflected on it. The second argument used by Berkeley is that the lines and angles referred to by Descartes "have no real existence in nature, being only an hypothesis framed by the mathematicians." This argument is of interest in showing how Berkeley thought that mathematicians were inclined to deal in fictitious entities, but it is unlikely that Descartes was deceived by them in this way.

Berkeley's third and main argument was based upon a theory that he expressed in the words, "distance, of itself and immediately, cannot be seen." William Molyneux, from whose *Dioptrics* (1692) Berkeley borrowed this theory, had supported it by the argument that since distance is a line or length directed endwise from the object seen to the eye, it can reach the eye at only one point, which must necessarily remain the same however near or far away the object is. If this argument is accepted, then distance could not possibly be seen, and could only be judged or, as Berkeley believed, "suggested."

DISTANCE IS SUGGESTED BY WHAT IS SEEN. What, then, according to Berkeley, is seen? The answer is not altogether clear, but it would seem that he thought that the immediate object of vision is two-dimensional, con-

taining relations of above and below and of one side and the other, with no necessary connection with a third dimension. Hence the relation between what is immediately seen on the one hand and the distance of objects on the other must be contingent and cannot be necessary. Distance, then, must be ascertained by means of something that has only a contingent relationship with what is seen. Berkeley mentioned the sensations we have when we adjust our eyes, the greater confusedness of objects as they come very close to the eyes, and the sensations of strain as we try to see what is very near. But he mainly relied on the associations between what a man has touched and what he now sees. For example, when a man now sees something faint and dim, he may, from past experience, expect that if he approaches and touches it he will find it bright and hard. When he sees something at a distance, he is really seeing certain shapes and colors, which suggest to him what tangible ideas he would have if he were near enough to touch it. Just as one does not hear a man's thoughts, which are suggested by the sounds he makes, so one does not directly see distance, which is suggested by what is seen.

SIGHT AND TOUCH. Berkeley's view that distance is not immediately perceived by sight is rejected by some writers, for instance by H. H. Price, in his *Perception* (1932), on the ground that it is plainly contradicted by experience. We just do see visual depth, it is held, so that it is idle to deny this fact on the basis of an argument purporting to prove that we cannot. Again, some critics, such as T. K. Abbott in *Sight and Touch* (1864) have argued not only that we do get our idea of distance from sight, but also that touch is vague and uninformative by comparison with sight, and hence less effective in giving knowledge of the material world. This discussion need not be developed, however, since, although he said in the *Essay* that by touch we get knowledge of objects that exist "without the mind" (§55), Berkeley's real view was that no sensible thing could so exist. It cannot be denied that on occasion Berkeley's language was imprecise. A crucial example of this occurs in his discussion of the question of whether a man born blind would, on receiving his sight, see things at a distance from him. According to Berkeley, of course, he would not; but to such a man, the most distant objects "would all seem to be in his eye, or rather in his mind" and would appear "(as in truth they are) no other than a new set of thoughts or sensations, each whereof is as near to him as the perceptions of pain or pleasure, or the most inward passions of his soul" (*Essay,* §41). It will be noticed how readily Berkeley passed from "in his eye" to "in his mind," and how he assimilated such very different things as sensations and thoughts. Indeed it is hard not to conclude that he thought that whatever was not seen at a distance must appear to be in the mind. If this is true, then one of the objects of the *Essay* was to show that the immediate objects of vision must be in the mind because they are not seen at a distance.

GEOMETRIES OF SIGHT AND OF TOUCH. As already seen, an extremely important thesis of the *Essay* is that the objects of sight and the objects of touch are radically different from one another. We see visible objects and we touch tangible objects, and it is absurd to suppose that we can touch what we see or see what we touch. According to Berkeley, it follows from this that tangible shape and visible shape have no necessary connection with one another. Geometers certainly supposed themselves to be concerned with shapes in abstraction from their being seen or touched, but Berkeley did not allow that this is possible. A purely visual geometry would necessarily be confined to two dimensions, so that the three-dimensional geometry that we have must be fundamentally a geometry of touch. He reinforced this strangely pragmatic view with the observation that a sighted but disembodied being that could not touch or manipulate things would be unable to understand even plane geometry, since without a body it would not understand the handling of rulers and compasses and the drawing of lines and the placing of shapes against one another.

ARGUMENTS FOR IMMATERIALISM

The arguments now to be considered are set out in the *Principles* and in the *Three Dialogues.* They are largely concerned with what Berkeley called "ideas," "ideas or sensations," "sensible things," or "sensible qualities." The very use of the word *idea* itself and, even more, its use in apposition with *sensation* had the purpose of indicating something that does not exist apart from the perception of it. Pains and itches are typical sensations, and no one supposes that they could exist apart from a being that experiences them. Rocks do not suffer, and water does not itch. When, therefore, sensible things such as colors, sounds, tangible shapes, tastes, and smells are called ideas, they are assimilated with sensations and hence relate to the perceiving beings that have them. It is now necessary, therefore, to examine the arguments with which Berkeley justified this.

SEVENTEENTH-CENTURY MATERIALISM. Berkeley's arguments for immaterialism can be understood only if we first consider the sort of view it was intended to refute. When Berkeley was forming his views, the natural sci-

ences had been so far advanced by the work of such men as Galileo Galilei, Andreas Vesalius, William Harvey, Robert Boyle, and Newton as to have given rise to a scientific view of the world. Such a view had been elaborated, in its philosophical aspects, by Locke in his *Essay concerning Human Understanding* (1690). Space and time were, so to say, the containers within which material things were situated. The movements and relations of material things could be explored by experiments and characterized in mathematical formulae.

Explanation in terms of particles in motion. The features of the world, thus revealed as fundamental, were those of place, shape, size, movement, weight, and the like; and it was in terms of these that heat and cold and color and sound found their explanation. Heat was thought to be due to the rapid movement of atomic particles, color to the transmission of particles or to the spreading of waves, and sound to the movement of the air between the emitting object and the ear. Whereas solid, shaped, moving objects, and the air and space within which they existed, were regarded as basic features of nature, the colors we see, the heat we feel, and the sounds we hear were held to be the effects that substances possessing only the basic characteristics produced in creatures with sense organs. If all creatures with sense organs and consciousness were removed from the world, there would no longer be any experienced sounds, but only pulsations in the air; particles would increase or decrease their speed of movement, but no one would feel hot or cold; light would be radiated, but there would be no colors as we know them. In such a world colors and sounds, heat and cold, would exist, as Boyle put it, in his *Origins of Forms and Qualities* (Oxford, 1666), only "dispositively," that is, those primary things would be there that would have given rise to the secondary ones if creatures with the requisite sense organs and minds had been there too.

Primary and secondary qualities. In this way a distinction was made between the primary qualities of things, which are essential and absolute, and their secondary qualities, which are those among the primary ones that give or would give rise to heard sounds, seen colors, and felt heat. It was an important element of this view that nothing could be perceived unless it acted upon the sense organs of the percipient and produced in his mind an idea. What was immediately perceived was not the external object but an idea representative of it. Locke had made people familiar with this theory, and had maintained that whereas the ideas we have of heat and cold and of color and sound correspond to nothing like themselves in the external world; for all that exists in the external world are solid bodies at rest or in movement, the ideas we have of the solid, shaped, moving bodies, that is, our ideas of primary qualities are like their sources or archetypes outside us. According to the view, then, that Berkeley was considering, material objects are perceived mediately or indirectly by means of ideas, some of which, the ideas of primary qualities, are like their originals; others, the ideas of secondary qualities, are relative to percipients and are unlike anything that exists in the external world.

MATERIALISM LEADS TO SKEPTICISM. Berkeley had two objections to the view that material objects are perceived mediately by means of ideas. One is that since it is held that we never perceive material things directly, but only through the medium of ideas, then we can never know whether any of our ideas are like the qualities of material substances since we can never compare our ideas with them; for to do so we should require direct or immediate acquaintance with them (*Principles,* §18). Indeed, if we accept Locke's position, then the very existence of material substances is in doubt, and we are constantly under the threat of skepticism (*Principles,* §86). Thus Berkeley argued that Locke's theory was in fact, although not by intention, skeptical, and that it could be remedied only by the elimination of material substances that could never be directly apprehended.

DISTINCTION BETWEEN PRIMARY AND SECONDARY QUALITIES UNTENABLE. Berkeley's second objection is that there can be no distinction between ideas of primary qualities and ideas of secondary qualities such as to make secondary qualities relative to the mind in a way in which primary qualities are not. In the *Three Dialogues* Berkeley elaborated the arguments, already used by Locke, to show that the ideas we have of secondary qualities are relative to the percipient and are what they are by reason of his condition and constitution. Things have no color in the dark; the same water can feel hot or cold to different hands, one of which has been in cold water and the other in hot; heat and cold are inseparably bound up with pain and pleasure, which can only exist in perceiving beings; and so on. But Berkeley then went on to argue that just as heat, for example, is inseparably bound up with pleasure and pain, and can therefore, no more than they can, exist "without the mind," so extension is bound up with color, speed of movement with a standard of estimation, solidity with touch, and size and shape with position and point of view (*Principles,* §§10–15). Thus Berkeley's argument is that nothing can have the primary

qualities without having the secondary qualities, so that if the latter cannot exist "without the mind," the former cannot so exist either.

ALL SENSIBLE QUALITIES MUST BE EITHER PERCEIVED OR PERCEPTIBLE. The preceding argument, however, is only a hypothetical one to the effect that if secondary qualities cannot exist "without the mind," primary qualities are in like case. What must now be considered are the reasons for holding that secondary qualities and, indeed, all sensible qualities can exist only in the mind so that their being is to be perceived. Berkeley, as already indicated, stated and elaborated well-known arguments to show that heat and cold, tastes, sounds, and the rest are relative to the percipient. Perhaps the most persuasive of these are those that purport to establish an indissoluble connection between heat, taste, and smell on the one hand, and pain or pleasure or displeasure on the other. Since no one denies that pain and pleasure can exist only if felt, then this applies to heat so intense as to be painful and to lesser degrees of heat as well. But in the *Principles,* his systematic treatise on the subject, Berkeley did not make use of these arguments, but said that "an intuitive knowledge may be obtained of this, by any one that shall attend to what is meant by the term *exist* when applied to sensible things" (§3). His view here is that "sensible things" are by their very nature perceived or perceivable. He supported this by asserting that to say there was an odor is to say that it was smelled, to say that there was a sound is to say that it was heard, to say that there was a color or shape is to say that it was seen or touched. According to Berkeley, unsmelled odors, sounds unheard, colors unseen, and shapes unseen or untouched are absurdities or impossibilities; brown leaves could not rustle on a withered tree in a world where life was extinct and God was dead. The very notion is absurd or impossible. Can more light be shed on the matter than is provided by the assertion that we have "intuitive knowledge" of it?

It must be remembered, in the first place, that Berkeley was contrasting the sounds we hear, for example, with the movements in the air, which men of science sometimes call sounds. Sounds in the latter sense, he said, "may possibly be *seen* or *felt,* but never *heard*" (*Three Dialogues,* 1). From this it may be seen that Berkeley looked upon sensible qualities as each the object of its own mode of perception, so that sounds are heard but not seen or touched, colors seen but not heard, heat felt but not seen, and so on. Hence colors require a viewer, sounds a hearer, and heat someone who feels it; and this is one reason why the being of sensible things is held to be their being perceived. The various modalities of sense are distinguished from one another by the mode of perception peculiar to each one, and in making these distinctions it is implied that perception is essential to them all. It is well known, of course, that Berkeley's critics accuse him of failing to distinguish between the object perceived and the perceiving of it. The perceiving of it, they say, can only be an act of a percipient without whom it could not exist, but the perceived object, whether it be a sound or a color or a shape, is distinct from the perceiving and could conceivably exist apart from it.

Whatever may be thought of this argument, it should not be used against Berkeley as if he had not thought of it. In fact he put it into the mouth of Hylas in the first of the *Three Dialogues* and rejected it on the ground that in perception we are passive and so are not exerting an act or activity of any kind. It should also be noticed that when Berkeley discussed sensation in detail he stated that sensible things or sensible qualities are perceived *immediately,* that is, without suggestion, association, or inference. We say that we hear vehicles and that we hear sounds. According to Berkeley, we hear sounds immediately, but vehicles, if they are out of sight, are suggested by or inferred from what we do hear, and so are heard only mediately or by means of the sounds immediately heard. Thus the sound we hear immediately is neither suggested nor inferred, but is heard just as it is. For this to be so, it must be before the mind; for if it were not before the mind, it would have to be inferred or suggested. Thus sensible qualities, as immediately perceived, must be objects of perception; their being is to be perceived.

Inconceivability of a sensible object existing unperceived. A very famous argument is now to be considered: It is inconceivable that anything should exist apart from, or independent of, mind. This argument was put forward by Berkeley in similar terms both in the *Principles* (§§22, 23) and in the *Three Dialogues* (1) and takes the form of a challenge to the reader to conceive of something—e.g., a book or a tree—existing absolutely unperceived. Berkeley argued that the attempt is impossible of fulfillment, since in order to conceive of a tree existing unperceived we who conceive of it, by the very fact of doing so, bring it into relation to our conception and hence to ourselves. As Hylas admits, in recognizing the failure of his attempt, "It is a pleasant mistake enough. As I was thinking of a tree in a solitary place, where no one was present to see it, me-thought that was to conceive a tree as existing unperceived or unthought of, not considering that I myself conceived it all the while." This is an argument that was later accepted as fundamental by idealists of such different persuasions as Johann Gottlieb Fichte and Francis

Herbert Bradley, who held that it shows that mind or experience is essential to the universe.

Sensible objects are complex ideas. Berkeley's example of a tree makes it necessary to consider how trees and other things in nature are related to ideas, sensible qualities, sounds, colors, shapes, and so on. According to Berkeley, such things as trees, books, and mountains are groups of ideas or sensible qualities and are hence as much within the mind as the latter are. Indeed, in his view, books, trees, and mountains are ideas, though complex ones. He admitted (*Principles,* §38) that this use of the word *idea* for what is ordinarily called a *thing* is somewhat odd, but held that, the facts being as they are, *idea* is better than *thing.* A tree is a group of ideas touched, seen, and smelled; a cherry, a group of ideas touched, seen, smelled, and tasted. The sensible qualities or ideas, without which we should have no conception of a tree or cherry, do not belong to some unseen, untouched, untasted substance or substratum, for the very conception of such a "something I know not what" (as Locke had called it) is incoherent, and rests upon the false view that we can conceive something in complete abstraction from ideas of sense.

Sensible objects, as ideas, are perceived directly. Berkeley therefore concluded that it is his theory that conforms with common sense, not that of the materialists or the dualists. For according to Berkeley we perceive trees and cherries directly by seeing, touching, and tasting them, just as the plain man thinks we do, whereas his opponents regard them as perpetually hidden from us by a screen of intermediaries that may be always deceiving us. Berkeley considered that by this view he had refuted skepticism of the senses, for, according to his theory, the objects of the senses are the things in the world: the trees, houses, and mountains we live among. But trees, houses, and mountains, as compounded of sensible qualities or ideas, cannot exist "without the mind."

SENSIBLE OBJECTS NOT COPIES OF MATERIAL ARCHETYPES. Berkeley's arguments showing that all sensible qualities or ideas exist only as perceived and that, therefore, things in nature, being groups of such ideas, cannot exist "without the mind" have now been expounded. It is now necessary to complete this account of Berkeley's arguments for immaterialism with his argument to show that not only must sensible qualities or ideas exist in the mind, but also that nothing like them can exist outside it. For anyone reluctant to accept immaterialism is likely to fall back on the view that our ideas, although in our minds, are copies of material archetypes. Berkeley's objection to this in the *Principles* (§8) is that "an idea can be like nothing but an idea," which he illustrated by saying that a color or shape can only be like another color or shape. In the *Three Dialogues* (1) he expanded the argument in two ways. Ideas, he said, are regarded by some as the perceived representatives of imperceptible originals, but "Can a real thing in itself *invisible* be like a color; or a real thing which is not *audible,* be like a *sound?*" His other reason for holding that ideas cannot be like any supposed external originals is that ideas are "perpetually fleeting and variable," and "continually changing upon every alteration in the distance, medium or instruments of sensation," while their supposed originals are thought to remain fixed and constant throughout all changes in the percipient's organs and position. But something that is fleeting and relative cannot be like what is stable and absolute, any more than what is incapable of being perceived can be like what is essentially perceptible.

SUMMARY. The following are Berkeley's central arguments in favor of immaterialism. They arose out of his exposure of the weaknesses and inconsistencies in the then current scientific view of the world, with its distinction between primary and secondary qualities and its theory of representative perception. According to Berkeley, since primary qualities cannot exist apart from secondary qualities, and since secondary qualities, and indeed all sensible qualities, cannot exist "without the mind," the independent material world of the then current scientific view was a conceptual absurdity. This was supported by the argument that our ideas cannot be likenesses of an external material world, since there is nothing conceivable they could be likenesses of except mind-dependent existences of their own type. The theory of representative perception was held to be essentially skeptical, and Berkeley claimed that his own theory, according to which we directly perceive ideas and groups of ideas that exist only as perceived, eliminates skepticism and accords with common sense.

METAPHYSICS AND THEOLOGY

In section 3 of the *Principles,* where Berkeley stated that we have intuitive knowledge of the fact that for sensible qualities to exist they must be perceived, he also stated that when we say that the table is in the room that we have left we mean that if we were to return there we could perceive it "or that some other spirit actually does perceive it." This shows that Berkeley was concerned with the problem of giving an account, within the terms of his immaterialism, of the continued existence of things that

are not being perceived by any human being. It also shows that he considered two ways of dealing with this problem. One way was to extend the doctrine that the existence of sensible things is their being perceived into the doctrine that the existence of sensible things is their being perceptible. The other way was to argue that when sensible things are not being perceived by human beings they must be perceived by "some other spirit."

BERKELEY NOT A PHENOMENALIST. The first way points in the direction of the modern theory of phenomenalism, the theory according to which, in John Stuart Mill's happily chosen words, material objects are "permanent possibilities of sensation." But might not anything, even material substances possessing only primary qualities, be perceptible, even if not actually being perceived? Some twentieth-century upholders of phenomenalism argued that the world was perceptible before there was any life or mind, in the sense that if there had been gods or human beings they would have perceived it. This could not be possible on Berkeley's theory, however, since, as we have seen, he held that only ideas or sensible things can be *like* ideas or sensible things, so that what is perceptible is limited by what is perceived.

PERCEPTIBLE OBJECTS PERCEIVED BY GOD. The perceptible, therefore, is limited to the mind-dependent, and, for Berkeley, the very notion of something that might be perceived, but is not, is unacceptable. Thus it seems that Berkeley was forced to supplement his phenomenalist account of unperceived objects with the view that whatever is not being actually perceived by human beings, but is only perceptible by them, must be an object of perception by "some other spirit." He used this same expression in section 48 of the *Principles,* where he denied that "bodies are annihilated and created every moment, or exist not at all during the intervals between our perception of them." In the *Three Dialogues* (2) he argued that since sensible things do not depend on the thought of human beings and exist independently of them "*there must be some other mind wherein they exist.*" This other mind is God; and thus, according to Berkeley, the existence of sensible things when not being perceived by finite spirits is a proof of the existence of an infinite spirit who perceives them always. Indeed, Berkeley considered it a merit of immaterialism that it enables this brief and, as he thought, conclusive proof to be formulated.

OUR IDEAS COME FROM GOD. In the *Principles* Berkeley put forward another proof of the existence of God, this time a proof based upon God as the cause of our ideas. As has been shown, Berkeley held that ideas are passive and that the only active beings are minds or spirits. Now some of our ideas, namely, ideas of imagination, we ourselves produce, but others, the ideas of sense, come to us without our willing them. "There is therefore some other will or spirit that produces them" (*Principles,* §29). That this is God may be concluded from the regular order in which these ideas come to us. The knowledge we have of God is analogous to the knowledge we have of other men. Since people are active spirits, we do not have ideas of them, but only of their expressions, words, and bodily movements. Through these we recognize them as possessors of minds and wills like those we know ourselves to have. Similarly, God reveals himself to us in the order of nature: "every thing we see, hear, feel, or in any wise perceive by sense, being a sign or effect of the Power of God."

ACTIVE SPIRITS AND PASSIVE IDEAS. These, then, are the elements of Berkeley's metaphysics. There are active spirits on the one hand and passive ideas on the other. The latter could not exist apart from the former, but the ideas in the minds of human beings are caused in them by God and sustained by him when they are not perceiving them. Regularly recurring groups of ideas are called bodies, and the ideas that form them are arbitrarily connected together and might have been connected quite differently. Thus there is no natural necessity or internal reason about the laws of nature, but the regular sequences of ideas reveal to us a single infinite being who orders things for our benefit. Active spirits and passive ideas are of different natures. The mind is not blue because the idea of blue is in it, nor is the mind extended because it has an idea of extension. Ideas are neither parts nor properties of minds. Berkeley seems to have thought that the relationship is sui generis, for he said that sensible qualities are in the mind "only as they are perceived by it, that is, not by way of *mode* or *attribute, but only by way of idea*" (*Principles,* §49).

GOD'S IDEAS AND OUR IDEAS. As already seen, Berkeley held that God was both the cause of the ideas in the minds of embodied finite spirits and also the Mind in which these ideas continued to exist when embodied finite spirits were not perceiving them. Berkeley was thus faced with the problem of how the ideas in finite minds are related to the ideas in God's mind. If we recall Berkeley's claim that he was on the side of common sense against the skeptics, then we should expect the ideas that continue to exist in God's mind to be identical with those that had been in the minds of the embodied finite spirits who had formerly perceived them.

However, he found that there were difficulties in this view. Humans perceive ideas of sense by means of sense organs, and their ideas vary in accordance with their position and condition, but God does not have sense organs. Furthermore, some ideas—for example, those of heat and cold, and sensations of smell and taste—are inseparable from sensations of pain and pleasure, but God is impassible, that is, not subject to feeling or emotion; hence he cannot be supposed to perceive ideas of this nature. In the *Three Dialogues* (3), therefore, Berkeley concluded that "God knows or hath ideas; but his ideas are not conveyed to Him by sense, as ours are." From this it is natural to conclude that the ideas that God perceives are not identical with the ideas that embodied finite spirits perceive. Berkeley was obviously thinking along these lines when, in the same *Dialogue,* he said that the things that one perceives, "they or their archetypes," must, since one does not cause them, have an existence outside one's mind. Elsewhere in this *Dialogue* he distinguished between what is "ectypal or natural" and what is "archetypal and eternal." Thus Berkeley's arguments and the language he used combine to suggest that the ideas in God's mind are not the same ideas as those in the minds of embodied percipients.

This point was taken up by the Samuel Johnson referred to earlier, in his correspondence with Berkeley. Johnson suggested that Berkeley's view is that "the real original and permanent existence of things is archetypal, being ideas *in mente Divina,* and that our ideas are copies of them." Johnson was too polite to press the point, but it follows that what we directly perceive are copies or representatives of divine originals, so that Berkeley's claim to have reinstated the direct, unmediated perception of common sense, in place of the representative and skeptical theory of the philosophers and scientists, cannot be substantiated. In his reply, Berkeley hardly met this point when he stated that material substance is an impossibility because it is held to exist apart from mind, whereas the archetypes in the divine mind are obviously inseparable from God's knowledge of them.

PHILOSOPHY OF NATURE

Berkeley carried on a persistent battle against the tendency to suppose that mere abstractions are real things. In the *New Theory of Vision* he denied the possibility of "extension in abstract," saying "A line or surface which is neither black, nor white, nor blue, nor yellow, etc., nor long, nor short, nor rough, nor smooth, nor square, nor round, etc., is perfectly incomprehensible" (§ 123). In the introduction to the *Principles,* his most explicit discussion of the matter, he quoted Locke's account of the abstract idea of a triangle "which is neither oblique nor rectangle, neither equilateral, equicrural, nor scalenon, but all and none of these at once," and pointed out that any actual triangle must be one of these types and cannot possibly be "all and none" of them. What makes any idea general, he held, is not any abstract feature that may be alleged to belong to it, but rather its being used to represent all other ideas that are like it in the relevant respects. Thus if something that is true of a triangle of one of these types is not true of it because it is of that one type, then it is true of all triangles whatever. Nothing exists but what is particular, and particular ideas become general by being used as representatives of others like them. Generality, we might say, is a symbolic device, not a metaphysical status. Thus Berkeley's attack on abstractions is based on two principles: (1) that nothing exists but what is particular, and (2) that nothing can exist on its own except what can be sensed or imagined on its own. If we accept the first principle, then abstract objects and Platonic forms are rejected, and if we accept the second, then possibility is limited to the sensible or imaginable.

SPACE, TIME, AND MOTION. We have already seen how Berkeley applied the above two principles to the abstract conception of unperceived existence, and to the abstract conception of bodies with only the primary qualities. It must now be shown how he applied them to some of the other elements in the scientific worldview he was so intent on discrediting. Chief among these were the current conceptions of absolute space, absolute time, and absolute motion. According to Berkeley, all these are abstractions, not realities. It is impossible, he held, to form an idea of pure space apart from the bodies in it. We find that we are hindered from moving our bodies in some directions and can move them freely in others. Where there are hindrances to our movement there are other bodies to obstruct us, and where we can move unrestrictedly we say there is space. It follows that our idea of space is inseparable from our ideas of movement and of body (*Principles,* §116).

So too our conception of time is inseparable from the succession of ideas in our minds and from the "particular actions and ideas that diversify the day"; hence Newton's conception of absolute time flowing uniformly must be rejected (*Principles,* §§97, 98).

Newton had also upheld absolute motion, but this too, according to Berkeley, is a hypostatized abstraction. If there were only one body in existence there could be no idea of motion, for motion is the change of position of

two bodies relative to one another. Thus sensible qualities, without which there could be no bodies, are essential to the very conception of movement. Furthermore, since sensible qualities are passive existences, and hence bodies are too, movement cannot have its source in body; and as we know what it is to move our own bodies, we know that the source of motion must be found in mind. Created spirits are responsible for only a small part of the movement in the world, and therefore God, the infinite spirit, must be its prime source. "And so natural philosophy either presupposes the knowledge of God or borrows it from some superior science" (*De Motu,* §34).

CAUSATION AND EXPLANATION. The thesis that God is the ultimate source of motion is a special case of the principle that the only real causes are spirits. This principle has the general consequence, of course, that inanimate bodies cannot act causally upon one another. Berkeley concluded from this that what are called natural causes are really signs of what follows them. Fire does not cause heat, but is so regularly followed by it that it is a reliable sign of it as long as "the Author of Nature always operates uniformly" (*Principles,* §107). Thus Berkeley held that natural laws describe but do not explain, for real explanations must be by reference to the aims and purposes of spirits, that is, in terms of final causes. For this reason, he maintained that mechanical explanations of movements in terms of attraction were misleading, unless it was recognized that they merely recorded the rates at which bodies in fact approach one another (*Principles,* §103). Similar arguments apply to gravity or to force when these are regarded as explanations of the movements of bodies (*De Motu,* §6). This is not to deny the importance of Newton's laws, for Newton did not regard gravity "as a true physical quality, but only as a mathematical hypothesis" (*De Motu,* §17). In general, explanations in terms of forces or attractions are mathematical hypotheses having no stable being in the nature of things but depending upon the definitions given to them (*De Motu,* §67). Their acceptability depends upon the extent to which they enable calculations to be made, resulting in conclusions that are borne out by what in fact occurs. According to Berkeley, forces and attractions are not found in nature but are useful constructions in the formulation of theories from which deductions can be made about what is found in nature, that is, sensible qualities or ideas (*De Motu,* §§34–41).

PHILOSOPHY OF MATHEMATICS

We have already seen that when he wrote the *New Theory of Vision,* Berkeley thought that geometry was primarily concerned with tangible extension, since visual extension does not have three dimensions, and visible shapes must be formed by hands that grasp and instruments that move. He later modified this view, an important feature of which has already been referred to in the account of Berkeley's discussion of Locke's account of the abstract idea of a triangle. A particular triangle, imagined or drawn, is regarded as representative of all other triangles, so that what is proved of it is proved of all others like it in the relevant respects. This, he pointed out later in the *Principles* (§126), applies particularly to size. If the length of the line is irrelevant to the proof, what is true of a line one inch long is true of a line one mile long. The line we use in our proof is a representative sign of all other lines. But it must have a finite number of parts, for if it is a visible line it must be divisible into visible parts, and these must be finite in length. A line one inch long cannot be divided into 10,000 parts because no such part could possibly be seen. But since a line one mile long can be divided into 10,000 parts, we imagine that the short line could be divided likewise. "After this manner the properties of the lines signified are (by a very usual figure) transferred to the sign, and thence through mistake thought to appertain to it considered in its own nature." Thus it was Berkeley's view that infinitesimals should be "pared off" from mathematics (*Principles,* §131). In the *Analyst* (1734), he brought these and other considerations to bear in refuting Newton's theory of fluxions. In this book Berkeley seemed to suggest that the object of geometry is "to measure finite assignable extension" (§50, Q.2).

Berkeley's account of arithmetic was even more revolutionary than his account of geometry. In geometry, he held, one particular shape is regarded as representative of all those like it, but in arithmetic we are concerned with purely arbitrary signs invented by men to help them in their operations of counting. Number, he said, is "entirely the creature of the mind" (*Principles,* §12). He argued, furthermore, that there are no units and no numbers in nature apart from the devices that men have invented to count and measure. The same length, for example, may be regarded as one yard, if it is measured in that unit, or three feet or thirty-six inches, if it is measured in those units. Arithmetic, he went on, is a language in which the names for the numbers from zero to nine play a part analogous to that of nouns in ordinary speech (*Principles,* §121). Berkeley did not develop this part of his theory. However, later in the eighteenth century, in various works, Étienne Bonnot de Condillac argued in detail for the thesis that mathematics is a language, and this view is, of course, widely held today.

CONCLUDING COMMENTS

Berkeley's immaterialism is a strange and unstable combination of theses that most other philosophers have thought do not belong together. Thus he upheld both extreme empiricism and idealism, both immaterialism and common sense, and both subjectivism (as it would seem) and epistemological realism (as it would also seem). Are these mere skillful polemical devices in the war against the freethinkers, or can they be regarded as elements in a distinctive and reasonably coherent metaphysics?

It is odd that Berkeley had so much to say about the relativity of each particular sense and so little to say about our perception of the physical world. He referred to perspectival distortions and the like in the course of defending his view that the existence of sensible qualities is their being perceived, but he did not seem to realize the difficulties they made for his view that perception is direct. Indeed, when, in the *Three Dialogues* (3) he mentioned the case of the oar that looks bent in the water when in fact it is straight, he said that we go wrong only if we mistakenly infer that it will look bent when out of the water. There is something seen to be straight, something else seen to be crooked, and something else again felt to be straight. We go wrong only when we expect that when we see something crooked we shall feel something crooked. But this implies that our perceptions of such things as oars, as distinct from our perceptions of colors and pressures, are not direct as common sense supposes. This reinforces the criticism we have already mentioned, that the ideas perceived by finite spirits with sense organs are different from, and representative of, the ideas in the mind of God. Berkeley was farther from common sense and closer to the views that he was criticizing than he was ready to admit.

It is obvious enough that Berkeley's immaterialism is not in accord with common sense. What place, then, must be given to his empiricism? He certainly rejected the Cartesian conception of a natural world that deceives the senses and is apprehended by the reason. He denied that mathematics reveals the ultimate necessities of things and anticipated to some extent the linguistic theory of mathematics. In arguing that causes are not to be found in nature, and in maintaining that the sciences of nature are primarily concerned with predicting human experiences, he formulated views that Ernst Mach and his modern-day followers have advocated. Furthermore, although he did not himself adopt it, he briefly formulated the theory of the physical world known as phenomenalism, the theory that consistent empiricists have adopted in order to avoid postulating objects that transcend sense experience. But, in spite of all this, Berkeley was an idealist rather than an empiricist. He held that sensible qualities or ideas are not independent or substantial existences and that minds or spirits are. On this most important matter, he was in agreement with his great contemporary, Leibniz. Furthermore, Berkeley's antiabstractionism, as we may call it, was constantly leading him toward the conclusion that the universe is a concrete unity in which an infinite mind is manifesting itself. If we look at his writings as a continuing and developing critique of abstraction, then we shall see that the *Siris* is not an aberration or a recantation but, as Henri Bergson said in his lectures on Berkeley, 1908–1909, a natural continuation of Berkeley's earlier views (*Écrits et paroles,* 2, p. 309).

See also Touch.

Bibliography

LIFE AND PRINCIPAL EDITIONS OF WORKS

Fraser, A. C. *The Works of George Berkeley.* 4 vols. London: Clarendon Press, 1871. New (first complete) edition, 1901.

Luce, A. A. *The Life of George Berkeley, Bishop of Cloyne.* London: Nelson, 1949.

Luce, A. A., and T. E. Jessop, eds. *The Works of George Berkeley, Bishop of Cloyne.* 9 vols. London and New York: Nelson, 1948–1957. The introduction and notes in this definitive edition are of great value.

Rand, Benjamin. *Berkeley and Percival.* Cambridge, U.K.: Cambridge University Press, 1914.

MAIN THEMES OF BERKELEY'S PHILOSOPHY

Hicks, G. Dawes. *Berkeley.* London: Benn, 1932.

Leroy, A. L. *George Berkeley.* Paris, 1959.

Warnock, G. J. *Berkeley.* London, 1953.

Wild, John D. *George Berkeley: A Study of His Life and Philosophy.* New York: Russell and Russell, 1936, 1962.

THE NEW THEORY OF VISION

Abbott, T. K. *Sight and Touch.* London: Longman, 1864.

Armstrong, D. M. *Berkeley's Theory of Vision.* Melbourne: Melbourne University Press, 1960.

Bailey, S. *A Review of Berkeley's Theory of Vision.* London, 1842.

Turbayne, C. M. "Berkeley and Molyneux on Retinal Images." *Journal of the History of Ideas* 16 (1955).

Vesey, G. N. A. "Berkeley and the Man Born Blind." *PAS* 61 (1960–1961).

ARGUMENTS FOR IMMATERIALISM

Bracken, H. M. "Berkeley's Realism." *Philosophical Quarterly* 8 (1958).

Broad, C. D. "Berkeley's Denial of Material Substance." *Philosophical Review* 43 (1954).

Laird, J. "Berkeley's Realism." *Mind,* n.s., 25 (1916): 308ff.

Luce, A. A. "The Berkeleyan Idea of Sense." *PAS,* supp. 27 (1953).

Luce, A. A. "Berkeley's Existence in the Mind." *Mind,* n.s., 50 (1941): 258ff.

Mates, Benson. "Berkeley Was Right." In *George Berkeley, Lectures Delivered before the Philosophical Union of the University of California.* Berkeley: University of California Press, 1957.

Moore, G. E. "Refutation of Idealism." In his *Philosophical Studies.* London: Routledge, 1922.

Sullivan, Celestine J. "Berkeley's Attack on Matter." In *George Berkeley, Lectures Delivered before the Philosophical Union of the University of California.* Berkeley: University of California Press, 1957.

METAPHYSICS AND THEOLOGY

Bracken, H. M. "Berkeley on the Immortality of the Soul." *Modern Schoolman* 38 (1960–1961).

Davis, J. W. "Berkeley and Phenomenalism." *Dialogue. Canadian Philosophical Review* 1 (1962–1963): 67–80.

Fritz, Anita D. "Berkeley and the Immaterialism of Malebranche." *Review of Metaphysics* 3 (1949–1950).

Gueroult, M. *Berkeley. Quatre études sur la perception et sur Dieu.* Paris, 1956.

Luce, A. A. *Berkeley and Malebranche.* London: Oxford University Press, 1934.

Myerscough, Angelita. "Berkeley and the Proofs for the Existence of God." In *Philosophy and the History of Philosophy.* Vol. 1, edited by John K. Ryan. Washington, DC: Catholic University of America Press, 1961.

Sillem, E. A. *George Berkeley and the Proofs for the Existence of God.* London: Longmans Green, 1957.

PHILOSOPHY OF NATURE AND PHILOSOPHY OF MATHEMATICS

Mach, Ernst. *The Analysis of Sensations.* Chicago: Open Court, 1914.

Myhill, John. "Berkeley's *De Motu*—An Anticipation of Mach." In *George Berkeley, Lectures Delivered before the Philosophical Union of the University of California.* Berkeley: University of California Press, 1957.

Popper, K. R. "A Note on Berkeley as Precursor of Mach." *British Journal for the Philosophy of Science* 3 (1953–1954).

Strong, Edward W. "Mathematical Reasoning and Its Object." In *George Berkeley, Lectures Delivered before the Philosophical Union of the University of California.* Berkeley: University of California Press, 1957.

Whitrow, G. J. "Berkeley's Critique of the Newtonian Analysis of Motion." *Hermathena* 82 (1953).

Whitrow, G. J. "Berkeley's Philosophy of Motion." *British Journal for the Philosophy of Science* 3 (1953–1954).

Wisdom, J. O. "Berkeley's Criticism of the Infinitesimal." *British Journal for the Philosophy of Science* 3 (1953–1954).

H. B. Acton (1967)

BERKELEY, GEORGE [ADDENDUM]

George Berkeley believed that there are only minds and ideas. The existence of minds (or spirits or souls), Berkeley contended, consists in perceiving whereas the existence of ideas (including sensations) consists in being perceived. Minds, which are the only substances, are active, and ideas are passive. The existence of physical objects consists in their being perceived. This is so because such objects consist of their qualities, and qualities are sensations. Thus Berkeley endorsed the idealist view that the physical world is kept in existence by being perceived. It depends upon the mind and cannot exist apart from perception. Consequently there is no need to presuppose material substance. Indeed, the very concept of material substance is incoherent. God is the source of our sensations. Hence we are in intimate contact with God, and we ought therefore always to be assured of God's existence and to be thankful to God.

The foregoing claims are central to Berkeley's thought. However, questions remain about the meaning and implications of some of these claims and about other aspects of his philosophy.

IDEAS AND OBJECTS

Berkeley sometimes espouses the view that physical objects are just collections or families of sensations that are produced by God in the minds of finite perceivers. But he explains the continued existence of objects that are not currently perceived by us by appealing to God's perceptions and to God's volitions. In addition he says that the ideas we perceive exist apart from the minds of finite perceivers at all times at which they exist. But if physical objects can exist qua divine ideas or volitions, or if they have any sort of existence independent of our sense perception, then such objects are not just collections of sensations in our minds.

Further, God perceives a great deal more than we perceive. For example, God presumably perceives all perceivable objects. Perhaps God perceives all such objects from all angles at once and perceives the interiors of physical objects whose surfaces alone we can see. Moreover, whatever form God's perception may take, it is not limited to the few senses that are our lot. Hence, if objects consist of our ideas along with God's ideas, our ideas are in danger of being second-class counterparts of God's. Our sensations seem to be relatively insignificant constituents of the familiar objects of our experience. It would not follow that the real objects are solely in God's mind. What would follow is that each object is largely in God's mind.

Nor would it follow that our perception of objects is indirect or lacking inimmediacy. If objects consist, or even partly consist, of ideas of sense,we can perceive them

immediately and directly by perceiving our ideas ofsense that are among their constituents. We perceive physical objects byperceiving them in part.

THE EXISTENCE OF GOD

Berkeley thought that we can know that God exists because our sensations both come to us from an external source and display a coherence, beauty, scale, and variety that bespeak a wise and benevolent source. He also thought that because some physical objects continue to exist while unperceived by us, there must be some other mind that perceives them while we do not perceive them. And Berkeley also presents this line of thought as the basis of a case for God's existence. Further, he thought that either the ideas we perceive or their archetypes exist independent of our perception in some other mind that exhibits them to us. Hence there must be such a mind.

Does Berkeley intend to offer three distinct arguments for God's existence (one that appeals to the source of our ideas, a second that appeals to the continued existence of unperceived objects, and a third that appeals to the independent existence of our ideas or their archetypes) or are these best understood as three strands in a single argument? However he may have conceived of the connections among these lines of thought, a case can be made for regarding the appeal to continuity as subsidiary to the appeal to the independent existence of our ideas or their archetypes. For if, at all times at which they exist, the objects we perceive by sense exist in another mind, by whom they are exhibited to us, then the fact that they exist when we do not perceive them seems fairly incidental. That is, their existence at times when we are not perceiving them is just a function of the fact that they have an independent existence, an existence that they have both while we perceive them and at times at which we do not perceive them but during which they exist.

Yet another argument for God's existence derives from Berkeley's thought that visual sensations are a language—for example, they tell us what other sensations we may receive, and our sensations are often combined in complex patterns. The use of a language requires a mind.

MINDS AND BODIES

How did Berkeley conceive of the relationship between the mind and the body? A human body, like any other physical object, is—at least is in part—a set of sensations. If the sensations that constitute, or partly constitute, physical objects are bestowed on finite minds by God, then when, say, I perceive your arm moving, one set of ideas produced in me by God is followed by another such set. Yet Berkeley says that we move our own limbs and that on this issue he differs from Malebranche. But how can he account for our moving our limbs or for our being able to move anything else by moving our limbs, or in any other way? Motion is always motion of some sensible body: it is inseparably united with other sensible qualities. There seems to be little room in Berkeley's theory for an account of motion without sensations. And if he is unable to account for our being able to move our limbs, on what basis does he think that one finite mind may reasonably conclude that there are other such minds?

Perhaps the claim that we are able to move our limbs is to be reduced to the view that certain volitions or non-sensory ideas that we produce serve as the occasion for God to grant us certain sensations. Or perhaps this claim is to be reduced to the view that certain sensations (such as those that constitute, or partially constitute, states of affairs that we wish to obtain) can be thought of as being produced by us, without any suggestion that this is indeed the case.

On these readings Berkeley would be "speaking with the vulgar and thinking with the learned" on the various occasions on which he says that we are able to move our limbs. However, he gives no indication that this is so. His treatment of this issue is quite different from that of physical causation. In the latter case, unlike the former, he is willing to say that fire heats and that all manner of causal connections obtain in the world even though, strictly speaking, he believes that this is not so.

One alternative reading is that we are able to produce some sensations. If God provides us with sensations on a great and wonderful scale, can we do so on a small scale? If the coherence, regularity, and so forth, of nature as a whole is good reason to conclude that God is its source, then perhaps the presence of discrete portions of the whole that manifest their own coherence and regularity is good reason to conclude that finite beings with limited powers are the sources of some of the sensations we receive.

If, in order to make a difference to the sensations you receive, I have to make a change in a mind-independent object, the suggestion that I should directly affect the sensations you receive without changing the world, including my body, seems absurd. If, on the other hand, as Berkeley avers, your sensations (and the sensations of all perceivers) constitute (or partly constitute) the physical world, to say that I can directly affect your sensations is just to say that I can make (or contribute to making) a change in the physical world.

IDEAS AND THE PERCEPTION OF IDEAS

How did Berkeley conceive of the relationship between ideas and their perception? Are there two things, an object and an act, that stand in a certain relationship to each other? Are there, at any rate, an object and a process in the mind to be related? For Berkeley says that in sense perception the mind is passive, which incidentally is a view that needs to be reconciled with his idea of the mind as an active, indivisible entity. Berkeley says that the existence of an idea is identical with its being perceived. His model for the relation between an idea and its perception is the relation between a pain and its perception. That relation is one of numerical identity. If an idea is identical with its being perceived, and if the perception of an idea is a private event in the mental life of an individual, it follows that an idea is something private to the mind in which it occurs.

At the same time it is natural to think of the qualities of objects, such as the redness of an apple, as something public that different people can perceive. Berkeley would want to preserve this commonsense belief. Yet if qualities are ideas, and an idea is identical with its being perceived, how can different perceivers perceive the same quality? Perhaps Berkeley should say that different people may perceive numerically the same quality even though they may not perceive numerically the same idea, thereby abandoning his identification of ideas and qualities.

ABSTRACTION

Berkeley devotes most of the Introduction of *A Treatise concerning the Principles of Human Knowledge*, Part 1, to a refutation of the Lockean belief in abstract ideas. Berkeley believed himself to have shown that abstract ideas are impossible. For any abstract idea would contain at least one inconsistency. Consider the abstract idea of man. This is supposed to contain only what is common to all men and to leave out what is distinctive of each. Yet, being the idea of a man, it must have some determinate human features. For example, it must be of a man with a particular size and shape. So an abstract idea of man must both lack and have such features. Again, since it must contain only what is common to all motion, the abstract idea of motion can't be of fast or slow, straight or curved motion. Yet being an idea of motion it must be of motion that is either fast or slow, either straight or curved. Therefore there can be no such idea.

Berkeley denies the possibility of a certain sort of precision or mental separation that allegedly gives rise to abstract ideas. He considers it to be impossible to separate mentally from our perception of an object that has color, extension, and motion an idea that consists of, say, extension alone or motion alone. Motion cannot exist except in something moving and we cannot separate mentally what can not exist separately. Moreover, he thought that, having noticed that particular motions have something in common, namely their motion, it is impossible for us to separate mentally what they have in common, thereby forming, once again, an idea that consists of motion alone.

In addition to being both incoherent and the alleged product of a process that actually is nonexistent, abstract ideas are quite unnecessary. General terms have meaning without signifying any such idea. For example, *triangle* has meaning in virtue of signifying indifferently a vast number of ideas of particular triangles.

ABSTRACTION AND IMMATERIALISM

Berkeley apparently understood his case against abstraction to be central to his case for thinking that physical objects cannot exist apart from perception. This is suggested by the fact that he devoted the introduction of what is probably his most important work to opposing abstract ideas. Indeed he says that the belief in abstraction has led to numerous errors and difficulties in nearly every area of inquiry, including the error of distinguishing the existence of sensible things from their being perceived. But it is not immediately obvious how he conceived of the connections between exposing the bogus character of abstract ideas and arguing that sensible objects can not exist unperceived.

One strand in his thinking may be this: It is impossible to believe to exist apart things that are incapable of existing apart. Physical objects are incapable of existing apart from perception. Hence it is impossible to believe them to do so. Or perhaps the point is that we cannot conceive of, or have an idea of, *a* and *b* as existing apart if *a* and *b* are incapable of existing apart; and since there cannot be existence apart from perception, we are unable to conceive of, or have an idea of, existence without perception.

On neither of these very similar readings does the case against abstraction contribute to the argument against mind-independent existence. Instead it has to be shown independently that there cannot be existence apart from perception. At most the case against abstraction illuminates the sort of error that is involved in believing that there are unperceived objects. Or at least it is a diagnosis of the sort of error involved in thinking that one is believing in mind-independent existence, because we are

told that it is impossible to so believe. But that believing in mind-independent existence actually involves this error is something that has to be established independently. Indeed, if the idea of unperceived existence is contradictory, as Berkeley insists, then even if there were abstract ideas, we would still be incapable of conceiving of existence apart from perception, just as we would still be incapable of conceiving of married bachelors or round squares or any other manifestly contradictory concepts.

Perhaps the focus should instead be on the idea of a material substratum and on an argument along the following lines. The idea, or rather alleged idea, of a material substratum involves the idea of being in general. The idea of being in general would be an abstract idea. And because there are no abstract ideas, there is no idea of being in general. Hence the very idea of a material substratum is unthinkable.

BERKELEY'S PHILOSOPHY OF LANGUAGE

We have already mentioned the central aspect of how general terms get their meaning: They signify a range of particular ideas. Berkeley makes some additional comments that bear on this topic, but the connections between the various themes he pursues are not always clear.

One question concerns the relationship between, on the one hand, thinking that *triangle* may be used to signify indifferently any and every idea of a triangle and, on the other hand, having in mind when one talks of triangles the idea of a particular triangle that stands proxy for other triangles, the latter being another theme that Berkeley mentions. Berkeley says that ideas, like terms, become general while remaining particular by fulfilling a general function. The latter of the two themes just mentioned is a matter of a general term getting its meaning by signifying a general idea. But the former seems rather different.

Another question concerns the role of selective attention. Berkeley says that we can consider a triangle solely as a triangle, ignoring all of its other features. It is not clear exactly what this involves. For example, if I selectively attend to the color of the red apple before me, do I concentrate on its redness while also being aware of its other properties, or perhaps while merely being aware that it has other properties? Whatever exactly selective attention may amount to, it—or something like it—is presupposed by the idea of one particular standing proxy for others that resemble it. For there will be some crucial respect in which those other things resemble it, and the idea that stands proxy for those other things will emphasize or single out in some way that crucial aspect.

Berkeley's account of what is involved in meaningful use of language has additional aspects. Words, whether they are particular or general, are sometimes used without the ideas they signify being brought to mind. Once the meaning of a term has been fixed by habitual association with one or more ideas, we often use it meaningfully without bringing those ideas to mind. Thus we can talk meaningfully about triangles without having any idea of a triangle in mind. We also become habituated to the association between certain terms or expressions and the arousal of passions such as fear, love, hatred, or admiration. Originally this process of arousal required an intervening stage at which ideas would render the relevant use of language meaningful, with the ideas in turn giving rise to various passions. But when the path is well trodden and the connections have become familiar, the mediating stage is omitted. This point about arousing passions exemplifies an important theme for Berkeley, namely that language has a number of uses. It can be used to communicate ideas to others. It can also be used to influence others—for example, by causing them to feel a certain way or by leading them to behave in a certain way.

So in virtue of the prior establishment of a word-idea connection, language can come to be used meaningfully without our having the relevant ideas before our minds. This assumes that, at least initially, a word is rendered meaningful by signifying one or more ideas. Berkeley may have held that there are also uses of language that are meaningful in the complete absence of all word-idea connections. At the very least the connection with ideas is further weakened. He seems to have thought this to be so in theology, in science, and indeed in some everyday parlance.

There are scriptural passages that are largely beyond our grasp but that we must nevertheless accept on faith. And Christians ought to be believe in, value, and pursue eternal happiness in heaven even though they lack any determinate ideas of the pleasures of heaven. Theological terms such as *grace* and *original sin* derive much of their significance from their influence on our passions and conduct, perhaps eliciting in us hope or gratitude or charity and in turn the actions that bespeak these attitudes.

Likewise, scientific terms such as force and gravity are convenient theoretical fictions that can help us to make accurate predictions and hence have practical value even though we lack a distinctive idea in each case. And there are many everyday terms (*myself*, *will*, *memory*, *love*,

and *hate*, for example) that we understand perfectly well but that do not suggest any distinct ideas to us. Also we talk meaningfully of minds even though these, being active, are not such that there could be an idea of them.

It is not clear in some of these cases whether Berkeley thought that we have no idea rather than no distinct and precise idea. If the latter were the case we might still have a vague and imprecise idea. Indeed, sometimes he seems to aim to show only that we have no relevant *abstract* idea.

See also Idealism; Ideas; Locke, John; Malebranche, Nicolas; Mind-Body Problem; Perception; Philosophy of Language; Sensa; Volition.

Bibliography

IDEAS AND OBJECTS

Foster, John. "Berkeley on the Physical World." In *Essays on Berkeley: A Tercentennial Celebration*, edited by John Foster and Howard Robinson. Oxford: Clarendon, 1985.

McKim, Robert. "Berkeley on Perceiving the Same Thing." In *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, edited by Phillip D. Cummins and Guenter Zoeller. Atascadero, CA: Ridgeview, 1992.

THE EXISTENCE OF GOD

Ayers, M. R. "Divine Ideas and Berkeley's Proofs of God's Existence." In *Essays on the Philosophy of George Berkeley*, edited by Ernest Sosa. Dordrecht: D. Reidel, 1987.

Kline, A. David. "Berkeley's Divine Language Argument" In *Essays on the Philosophy of George Berkeley*, edited by Ernest Sosa. Dordrecht: D. Reidel, 1987.

MINDS AND BODIES

Taylor, C. C. W. "Action and Inaction in Berkeley." In *Essays on Berkeley: A Tercentennial Celebration*, edited by John Foster and Howard Robinson. Oxford: Clarendon, 1985.

Tipton, Ian. *Berkeley*. London: Methuen, 1974, 302 ff.

IDEAS AND THE PERCEPTION OF IDEAS

McCracken, Charles J. "Berkeley on the Relation of Ideas to the Mind." In *Minds, Ideas, and Objects: Essays on the Theory of Representation in Modern Philosophy*, edited by Phillip D. Cummins and Guenter Zoeller. Atascadero, CA: Ridgeview, 1992.

ABSTRACTION

Atherton, Margaret. "Berkeley's Anti-Abstractionism." In *Essays on the Philosophy of George Berkeley*, edited by Ernest Sosa. Dordrecht: D. Reidel, 1987.

Winkler, Kenneth. *Berkeley: An Interpretation* Oxford: Clarendon, 1989, Ch. 2.

ABSTRACTION AND IMMATERIALISM

Brandt Bolton, Martha. "Berkeley's Objection to Abstract Ideas and Unconceived Objects." In *Essays on the Philosophy of George Berkeley*, edited by Ernest Sosa. Dordrecht: D. Reidel, 1987.

Pappas, George S. "Abstract Ideas and the esse is percipi thesis." In *George Berkeley: Essays and Replies*, edited by David Berman. Dublin: Irish Academic Press, 1986.

PHILOSOPHY OF LANGUAGE

Pitcher, George. *Berkeley*. London; Boston: Routledge & Kegan Paul, 1977, pp. 78–90.

Warnock, G. J. *Berkeley* London: Penguin, 1953. Chapter 4.

OTHER WORKS

Atherton, Margaret. *Berkeley's Revolution in Vision*. Ithaca: Cornell University Press, 1990.

Berman, David. *George Berkeley: Idealism and the Man*. Oxford: Clarendon, 1994.

Dancy, Jonathan. *Berkeley: An Introduction*. Oxford: Basil Blackwell, 1987.

Grayling, A. C. *Berkeley: The Central Arguments*. La Salle, IL: Open Court, 1986.

Jesseph, Douglas. *Berkeley's Philosophy of Mathematics*. Chicago: University of Chicago Press, 1993.

Muehlmann, Robert. *Berkeley's Ontology*. Indianapolis: Hackett, 1992.

Muehlmann, Robert. *Berkeley's Metaphysics: Structural, Critical and Interpretive Essays*. University Park: Pennsylvania State University Press, 1995.

Turbayne, Colin M., ed. *A Treatise concerning the Principles of HumanKnowledge: George Berkeley with Critical Essays*. Indianapolis: Bobbs-Merrill, 1970.

Robert McKim (1996, 2005)

BERLIN, ISAIAH
(1909–1997)

Latvian born, English educated, and a cosmopolitan in the world of ideas, Isaiah Berlin was both a prolific public intellectual and a distinguished academic, concluding his career as Oxford University's Chichele Professor of Social and Political Theory. After publishing some early essays in analytical philosophy, Berlin soon turned to more historical studies. While favoring the essay form, he published an important book-length study of Marx (1939) that was critical of Marx's historical determinism in ways that anticipated his later critiques of theories of historical inevitability. During the Second World War, Berlin worked for the British government in the United States, after which he returned to teaching at Oxford University, with occasional sojourns in London and the United States. His practical political involvements lent a spirit of engagement to his writings, whatever the subject.

Berlin championed political theory at a time when it was distinctly unfashionable in professional philosophy. To dismiss political reflection because of its rough-hewn character, he maintained, is to misconstrue the nature of the subject and leave oneself at the mercy of uncriticized

political prejudices. But Berlin's major importance as a political thinker rests in the vision of liberalism that he articulated in the post–World War II decades. In his seminal essay "Two Concepts of Liberty," he developed an influential distinction between negative freedom (to act without interference) and positive freedom (to be one's own master), and expressed special concern about the totalitarian dangers lurking in the latter.

While Berlin clearly privileged negative freedom over positive freedom, his distinction is more nuanced than is often acknowledged. He made no fetish of liberty, and reminded readers that communities in conditions of dire poverty cannot give much thought to formal freedoms. What he most bemoaned in positive freedom was the ideal of self mastery projected onto classes, peoples, or the whole of mankind. His championing the liberal commitment to rights, as demarcating individual spheres of autonomy, has had a deep impact on all subsequent liberal theory, including John Rawls's political liberalism and Richard Rorty's pragmatic liberalism. He wrote, "There are frontiers, not artificially drawn, within which men should be inviolable," frontiers so secure that their observance "enters into the very conception of what it means to be a human being" (1969, 165).

Berlin also argued for identifying liberalism with an ethic of pluralism, for which ultimate good as postulated by determinist views of historical development, does not exist. "To assume that all values can be graded on one scale … seems to me to falsify our knowledge that men are free agents" (1969, p. 171). Liberal society is one in which values are always in conflict, and such conflicts cannot be resolved by metaphysical fiat but must instead be addressed by the arduous patient work of practical negotiation. Thus conceived, the liberal outlook is intrinsically opposed to the totalitarian impulse in all its forms. It rests on the acceptance of moral uncertainty as our epistemological fate, and tolerance as our political imperative.

Of his many contributions to the history of ideas, Berlin's studies of Giambattista Vico, Johann Gottfried Herder, Johann Georg Hamann, and Romanticism were of special importance to philosophy. His discussions of Romantic "expressivism" were instrumental to the English-language revival of studies in the philosophy of Georg Wilhelm Friedrich Hegel, starting in the 1970s. They helped shape the understanding of the Romantic background that Hegel both appropriated and criticized. Berlin's writings on Romanticism intertwined with his long interest in modern nationalism, which he regarded more sympathetically than many other post–World War II liberals. Berlin also wrote widely on Russian novelists and thinking, translating Ivan Turgenev (1818–1883) and other classic writers into English.

Berlin wrote for popular as well as academic audiences and received much acclaim throughout his long life. He was awarded the Jerusalem Prize, the Erasmus Prize, the Angelli Prize, and the Lippincott Prize, among others. He was knighted in 1957 and received the Order of Merit in 1971. He died in Oxford, U.K., at the age of 88, having once remarked, "I don't mind death. … but I find dying a nuisance" (*New York Times*, November 7, 1997).

See also Determinism in History; Hamann, Johann Georg; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Ideas; Liberalism; Marx, Karl; Pluralism; Rawls, John; Rights; Romanticism; Rorty, Richard; Vico, Giambattista.

Bibliography

WORKS BY BERLIN

Marx: His Life and Environment. London: T. Butterworth, 1939.

Four Essays on Liberty. London: Oxford University Press, 1969.

The Hedgehog and the Fox: An Essay on Tolstoy's View of History. New York: Simon and Schuster, 1970.

Russian Thinkers. New York: Viking Press, 1978.

Against the Current. New York: Viking Press, 1980.

The Crooked Timber of Humanity. London: John Murray, 1990.

The Proper Study of Mankind. London: Chatto and Windus, 1997.

The Roots of Romanticism. Princeton, NJ: Princeton University Press, 1999.

Vico and Herder: Two Studies in the History of Ideas. Princeton, NJ: Princeton University Press, 2000.

WORKS ON BERLIN

Gray, John. *Isaiah Berlin*. Princeton, NJ: Princeton University Press, 1996.

Ignatieff, Michael. *Isaiah Berlin: A Life*. New York: Metropolitan Books, 1998.

Ryan, Alan, ed. *The Idea of Freedom: Essays in Honour of Isaiah Berlin*. Oxford, U.K.: Oxford University Press, 1979.

Cheyney Ryan (2005)

BERNARD, CLAUDE
(1813–1878)

Claude Bernard, French physiologist, was born in Saint-Julien (Rhône). He received his M.D. in 1843 and became a professor at the Sorbonne in 1852, taking the new chair in physiology in 1854. The following year he was appointed professor of experimental medicine at the Col-

lège de France and in 1868 became professor of general physiology at the Museum of Natural History in Paris. He was elected a member of the Academy of Sciences in 1854 and of the Académie Française in 1868; in 1869 he became a senator.

Bernard early gave up any idea of clinical practice in favor of experimental physiology. He made a number of important contributions in this field (on the chemistry of digestion, the production of sugar in animals, the nervous system, poisons, and anesthetics), many of which were awarded scientific prizes. After a period of ill health, while not ceasing laboratory work, he turned to more general and programmatic questions of scientific method and published, in particular, his famous *Introduction à l'étude de la médecine expérimentale* (Paris, 1865; translated by H. C. Green as *An Introduction to the Study of Experimental Medicine,* New York, 1927).

In the *Introduction,* Bernard based his conclusions as much as possible on his own scientific experiences, since he believed that proper procedure cannot be legislated for scientists from without but must be developed from the nature and needs of science itself. He distinguished the mature experimental method from empiricism, which is merely its first step. Bernard identified crude empiricism, which observes and experiments at random, not only with his own teacher, François Magendie, but also, mistakenly, with Francis Bacon, regarding himself rather in the tradition of Descartes, despite the fact that he insisted on constant laboratory experimentation and criticism and had a low opinion of the application of mathematics to biological problems. His hostility to the use of statistical methods in biology derived from the one article of faith he regarded as necessary to any scientist: belief in the operation of a determinism without exceptions, such that a set of conditions (a cause) will invariably produce the same phenomenon (an effect). This determinism he called an absolute principle, in contrast to theories and hypotheses, which are always provisional and subject to revision or abandonment because of the discovery of incompatible facts. But theories and hypotheses, the products of human reason, are on the other hand the necessary guides for rational experimentation.

Bernard saw no difference in principle between scientific method as applied to living beings and to inorganic matter, although results were more difficult to achieve in physiology because of the far greater complexity of the phenomena. He believed in a fundamental unity among all forms of life, the higher forms being distinguished by their greater independence of the external environment and a correspondingly greater dependence on their "internal environment" (above all, the blood). He also held that the phenomena taking place in living beings are ultimately reducible to physicochemical processes. Efforts to enlist Bernard in the cause of vitalism are wide of the mark. Equally mistaken is the attempt to affix a positivist label. He strenuously advocated scientific doubt and self-criticism, and was opposed to all philosophical systems, including the positivist, while not denying the usefulness of the work of philosophers in their own sphere. Bernard's critical method was closer to twentieth-century methods based on the principle of falsifiability, used by Karl Popper and others, than to those of many of his contemporaries.

See also Bacon, Francis; Descartes, René; Empiricism; Popper, Karl Raimund; Positivism; Scientific Method; Vitalism.

Bibliography

Bernard's general works also include *La science expérimentale* (Paris: J. B. Baillière, 1878) and *Principes de médecine expérimentale* (Paris: Presses universitaires de France, 1947).

For works on Bernard, consult J. M. D. Olmsted and E. H. Olmsted, *Claude Bernard and the Experimental Method in Medicine* (New York: H. Schuman, 1952); Robert Clarke, *Claude Bernard et la médecine expérimentale* (Paris, 1961); and Paul Foulquié, *Claude Bernard* (Paris: Éditions de l'école, 1954).

W. M. Simon (1967)

BERNARD OF CHARTRES

(died c. 1124–1130)

Bernard of Chartres, a Breton and elder brother of Theodoric of Chartres, was a master at Chartres at periods during the second and third decades of the twelfth century and became chancellor at least by 1119. He is no longer to be confused with Bernard Silvestris of Tours. To Bernard of Chartres belongs much of the credit for bringing the intellectual life of Chartres to its apogee, and his pupils included Gilbert of Poitiers, William of Conches, and Richard the Bishop. No complete writing by Bernard has survived, although he is known to have written philosophical verse and to have expounded Porphyry's *Isagoge.* Nevertheless, John of Salisbury learned of the character of Bernard's literary and philosophical teaching through William and Richard, and in John's writings we find a sympathetic portrait of Bernard as a real lover of learning and a leading grammarian, the most abounding spring of letters and the most finished Platonist of those days. John

eulogizes the "old Chartrain" as an excellent teacher of Latin language and literature, whose aim was to produce well-lettered and well-spoken students by means of an unhurried, cultured, humanist education, firmly based upon a groundwork of grammar. Bernard's love of the ancients was expressed in a famous simile of the moderns as dwarfs who can see farther than the ancients because they are perched upon the shoulders of giants.

Bernard was a philosopher with a taste for speculative grammar and for Platonism. He held opinions that the more Aristotelian John did not entirely share. We know only one of Bernard's grammatical speculations, namely, that the relationship of a quality-word (e.g., whiteness) to its derivatives (e.g., to whiten, white) resembles the relationship of the Platonic Ideas to the things in which they participate. As a Platonist, Bernard held that true reality is found in the eternal Ideas, which are the models of all perishable things. Particular sensible things, being unstable and ephemeral, cannot properly be said to be. Bernard's contribution to the disputes of his time over the nature of universals was to equate universals with Ideas; hence universals, in his view, were real beings. Guided by Boethius, Bernard and his school also labored to reconcile the differences between Plato and Aristotle.

Under the influence of the ninth-century thinker John Scotus Erigena, Bernard also sought to reconcile the teaching of Plato's *Timaeus* with that of the Bible by reexamining the relationships between the three categories of true being: God, matter, and the Ideas. He adhered to patristic teaching in accepting the view that matter was created by God. He also held that the eternal Ideas are in some way posterior to God. The Ideas are assimilated with God's mind or the divine providence; but although they are immanent in the mind of God, they are also a created effect. They are eternal, but not, in Bernard's view, coeternal with God. Only the three persons of the Trinity are both coequal and coeternal.

On the other hand, Bernard also attempted to show that the Ideas were not directly mixed with sensible objects. He distinguished between Ideas that subsist in the mind of God and the copies of these Ideas that are concreated with matter. To the latter Ideas he gave, under Boethian influence, the name of native forms (*formae nativae*).

Essentially, Bernard sought to affirm the transcendence of God over the Ideas and to avoid pantheism by the theory of native forms, which allowed no confusion of God with creation. Insofar as we can judge his motives, Bernard was adapting the Platonism that he knew to Christianity, just as he modified this Platonism in the light of Aristotelianism. His teaching was promoted by other Chartrains, especially by Gilbert of Poitiers.

See also Aristotelianism; Aristotle; Bernard of Tours; Boethius, Anicius Manlius Severinus; Chartres, School of; Erigena, John Scotus; Gilbert of Poitiers; Ideas; John of Salisbury; Plato; Platonism and the Platonic Tradition; Theodoric of Chartres; William of Conches.

Bibliography

PRIMARY SOURCES

John of Salisbury. *Metalogicon.* Edited by C. Webb. Oxford, 1929. I, v, xi, xxiv; II, xvii; III, ii; IV, xxv. Translated under the same title by D. D. McGarry, Berkeley: University of California Press, 1955.

John of Salisbury. *Policraticus.* Edited by C. Webb. Oxford, 1909. VII, 13.

Otto of Freising. *Gesta Friderici.* Edited by G. Waitz. 3rd ed. Hanover, 1912. I, xlvii; lii.

SECONDARY SOURCES

Clerval, A. *Les Écoles de Chartres au moyen âge.* Paris: A. Picard, 1895. Pp. 158–163.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages.* New York: Random House, 1951. Pp. 619–620.

Gilson, Étienne. "Le Platonisme de Bernard de Chartres." *Revue néo-scolastique de philosophie* 25 (1923): 5–19.

Gregory, T. *Anima Mundi. La filosofia di Guglielmo di Conches.* Florence: G. C. Sansoni, 1955. Pp. 76–79.

Parent, J. *La Doctrine de la création dans l'école de Chartres.* Paris: J. Vrin, 1938. Pp. 45–48, 84–85.

Poole, R. L. *Illustrations of the History of Medieval Thought,* 2nd ed. London: Society for Promoting Christian Knowledge, 1920. Pp. 100–107.

David Luscombe (1967)

BERNARD OF CLAIRVAUX, ST.

(1090–1153)

St. Bernard of Clairvaux, the monastic reformer and theologian, was born of a noble family at Fontaine, France, near Dijon. He became a Cistercian at Cîteaux in 1112 and founding abbot of Clairvaux in 1115. Throughout his life he was a tireless founder, reformer, preacher, and writer who, as friend or opponent, made contact with almost every notable in western Europe. His influence as a simple abbot on high ecclesiastical affairs is without parallel in the history of the Western church, and his spiritual teaching has been a living force to the present day. Though he was a professed enemy of secular culture (he

"raided" the schools of Paris on a celebrated occasion in 1140) and was lacking in scholastic training, Bernard was a literary genius of the first order, and no mean theologian. His treatises *De Diligendo Deo* (On the love of God; 1126) and *De Gratia et Libero Arbitrio* (On grace and free will; 1127), though based on St. Augustine, also show the influence of Origen, Gregory of Nyssa, and the Pseudo-Dionysius, as do also some of his longer letters. In the history of thought he is remembered for his controversies with Peter Abelard and Gilbert de La Porrée. He distrusted contemporary dialectic, partly because of a justified apprehension of the dangers in the formulas of both his opponents, but most of all because his approach to theological truth was by way of meditation and intuitive penetration, whereas theirs was by way of logical expression and analysis. His influence restrained theological improvisation and methodical virtuosity, and left the field clear for the great scholastics of the next century.

His most valuable contribution to thought was in the realm of mystical theology. He was a medieval pioneer of the analysis and explanation of mystical experience. His teaching, ostensibly based on St. Augustine, was in many respects new, and was followed by that of the Victorines and others, though later rivaled and eclipsed by the Dionysian-Thomist school of Rhineland Dominicans. Bernard's mysticism was one of love. Man, by recognizing his own nothingness, turns to God with humility and love, and man's will, with divine help, can reach perfect accord with the divine will. The divine Word can then teach him (infused knowledge) and move him (infused love) in an intimate union sometimes momentarily experienced as ecstasy. Thus Bernard differs, in expression at least, from the intellectual mysticism of Neoplatonism reflected in both Augustine and Dionysius. In his *Sermons on the Canticle,* Bernard was also a pioneer in the clear description of his own mystical experience, which in many ways resembled that of St. Teresa of Ávila.

See also Abelard, Peter; Augustine, St.; Gregory of Nyssa; Love; Medieval Philosophy; Mysticism, History of; Neoplatonism; Origen; Pseudo-Dionysius; Teresa of Ávila, St.

Bibliography

Works by Bernard are to be found in *Patrologiae Cursus Completus, Series Latina.* Edited by J. P. Migne, Vols. 182–185 (Paris, 1844–1864).

For biography, see E. Vacandard, *Vie de saint Bernard,* 2 vols. (Paris, 1895), often reprinted. Also useful are articles on Bernard by E. Vacandard in *Dictionnaire de théologie catholique* (Paris, 1910) and by J. Canivez in *Dictionnaire d'histoire et de géographie écclesiastique* (Paris, 1935).

For Bernard's contribution to mystical theology, see Étienne Gilson, *Théologie mystique de saint Bernard* (Paris: J. Vrin, 1934), translated by A. H. C. Downes as *The Mystical Theology of Saint Bernard* (London: Sheed and Ward, 1940), C. Butler, *Western Mysticism: The Teaching of SS. Augustine, Gregory and Bernard on Contemplation and the Contemplative Life,* 2nd ed. (London: Constable, 1951). Dom J. Leclercq is preparing a critical edition of St. Bernard's works.

David Knowles (1967)

BERNARD OF TOURS

(d. after 1167)

Bernard of Tours was a humanist who taught at Tours and was known as Bernardus Silvestris. He is uncertainly identified with Bernard, chancellor of Chartres circa 1156 and bishop of Quimper from 1159 to 1167. Very little else is known of his life except that he taught the art of writing and wrote an *Ars Versificatoria,* which has not been found. He also wrote a moralizing allegorical commentary on part of Vergil's *Aeneid* that displays leanings toward a naturalistic ethic. He translated into Latin an Arabic treatise on geomancy, the *Experimentarius,* and, inspired by Quintilian, composed the *Mathematicus,* a poem about an astrological prediction.

His most famous work, dedicated to Theodoric of Chartres in about 1150, is the *De Mundi Universitate,* an allegory in prose and verse on the origin of the world and man. The theme is Nature's appeal to Nous (mind), the providence of God, to end the chaos of *hylē* (matter), the primordial matter of the megacosmos. In Nous exist the exemplary forms of creation. Nous separates four elements out of *hylē* and informs the world with a soul ("entelechy," the Aristotelian εντελεχια). Nous next sends Nature to find Urania and Physis. Urania, queen of the stars, and Physis, in the lower world, use the remains of the four elements, in collaboration with Nature, to form man (the microcosmos). The sources of Bernard's inspiration were the Latin version of Plato's *Timaeus* with the commentary of Chalcidius, and also Ovid, Claudian, Macrobius, Boethius, and Augustine. There is, in addition, a marked biblical and a Hermetic influence.

The humanism of this work is more profane than Christian; the world is that of the *Timaeus* rather than that of Genesis. But the paganism, even unorthodoxy, of Bernard should not be exaggerated. Thus, Bernard was silent about a divine creation of matter, but his concern was to depict the organization of matter into the universe. There is no consistent dualism of God and matter;

hylē is preexistent to the ordering work of Nous, but the problem of its eternity is not broached. One should not conclude from the emanation of a world soul from Nous that Bernard was a pantheist. We cannot, in fact, extract from this often nebulous work a unified view of Bernard's thought. Bernard's purpose was imaginative rather than strictly philosophical. Nonetheless, Bernard reflects the speculative interests of his time, particularly those of the Chartrains; he reflects their desire for a more rational explanation of the universe and of biblical cosmology with the aid of Greek ideas.

See also Augustine, St.; Boethius, Anicius Manlius Severinus; Chartres, School of; Hermeticism; Humanism; Medieval Philosophy; Nous; Plato; Theodoric of Chartres.

Bibliography

WORKS BY BERNARD OF TOURS

Commentum Super Sex Libros Eneidos. Edited by G. Riedel. Greifswald, 1924.

De Mundi Universitate. Edited by Carl Sigmund Barach and Johann Wrobel. Innsbruck: Wagner'schen Universitäts-Buchhandlung, 1876.

"Experimentarius." Edited by M. B. Savorelli. *Rivista critica di storia di filosofia* 14 (1959): 283–342.

Mathematicus. Edited by B. Hauréau. Paris, 1895.

WORKS ON BERNARD OF TOURS

Curtius, E. R. *European Literature and the Latin Middle Ages.* London, 1953. Pp. 108–113. Translated from the German edition of 1948 by W. R. Trask.

Faral, E. "Le manuscrit 511 du Hunterian Museum." *Studi medioevali,* n.s., 9 (1936): 69–88.

Gilson, E. "La cosmogonie de Bernardus Silvestris." *Archives d'histoire doctrinale et littéraire du moyen âge* 3 (1928): 5–24.

Gregory, T. *Anima Mundi. La filosofia di Guglielmo di Conches,* Florence: G. C. Sansoni, 1955. Pp. 64–67.

Silverstein, T. "The Fabulous Cosmogony of Bernardus Silvestris." *Modern Philology* 46 (1948/1949): 92–116.

Thorndike, Lynn. *A History of Magic and Experimental Science.* New York: Macmillan, 1929. Vol. II, pp. 99–123.

David Luscombe (1967)

BERTALANFFY, LUDWIG VON

(1901–1972)

Ludwig von Bertalanffy, one of the chief exponents of the "organismic" standpoint in theoretical biology, was born in Austria in 1901 and educated at the universities of Innsbruck and Vienna. Until 1948 he taught at the University of Vienna, first as an instructor and later as professor of biology in the medical school. He emigrated to Canada in 1949 and held academic posts at the University of Ottawa and the University of Alberta, where he was appointed professor of theoretical biology in 1962. Von Bertalanffy's writings are voluminous, amounting to more than two hundred items. These include scientific papers in such fields as animal growth, cell physiology, experimental embryology, and cancer research. His two best-known books on philosophical biology are *Kritische Theorie der Formbildung* (Berlin, 1928; translated by J. H. Woodger as *Modern Theories of Development,* London, 1933) and *Das biologische Weltbild* (Bern, 1949; translated by the author as *Problems of Life,* New York, 1960). Since 1950 he had been active in promoting an interdisciplinary field called "General System Theory." The society associated with this enterprise has issued several yearbooks.

Von Bertalanffy contended that neither classical mechanism nor vitalism provides an adequate model for understanding organic phenomena. Vitalism is intellectually sterile because it appeals to a mysterious élan vital, entelechy, or psychoid to account for the properties of living things. Mechanism, von Bertalanffy declared, involves three mistaken conceptions: (1) the "analytical and summative" conception, according to which the goal of biological inquiry is the analysis of organisms into fundamental units and the explaining of organic properties by a simple adding up of these units; (2) the "machine-theoretical" conception, which regards the basis of vital order as a set of preestablished structures or "mechanisms" of a physicochemical kind; and (3) the "reaction-theoretical" conception, according to which organisms are automata, reacting only when subjected to stimulation and otherwise quiescent. These conceptions, von Bertalanffy argued, cannot yield a well-grounded explanatory theory of life.

In place of them he proposed an organismic model on which such a theory can be built. The model represents organisms as wholes or systems that have unique system properties and conform to irreducible system laws. Organic structures result from a continuous flow of processes combining to produce patterns of immense intricacy. Far from being passive automata, living things are centers of activity with a high degree of autonomy. Biological systems are stratified. There is a hierarchy of levels of organization from living molecules to multicellular individuals and supraindividual aggregates. The whole of nature is "a tremendous architecture in which

subordinate systems are united at successive levels into ever higher and larger systems."

Von Bertalanffy sought to show that this conception illuminates such matters as embryonic development, genetic processes, growth, self-regulation, metabolism, and evolution. Thus, in embryology it is no longer necessary to take sides in the old contest between preformationism and epigenesis, if we adopt the hypothesis that a fertilized ovum is a system whose development is determined by internal system conditions. Similarly, the ostensible purposefulness manifested by this development is an illustration of the unique property of "equifinality," which marks the behavior of organisms as "open" systems. These systems differ in important respects from the closed systems dealt with by physics. The thermodynamic principles that apply to the two cases are by no means the same. Nevertheless, von Bertalanffy believed that "there are general principles holding for all systems, irrespective of their component elements and of the relations or forces between them." These principles, he thought, can be studied through General System Theory, whose function is to bring about the unity of science.

The organismic conception of life is presented by its author as an intellectual breakthrough that "may well be set beside the great revolutions in human thought." Critics have found this claim extravagant in view of the sketchy and programmatic character of von Bertalanffy's presentation. They contend that the organismic conception has no right to be called "revolutionary" until its merits have been shown in detailed and extensive biological analysis. Nevertheless, von Bertalanffy has called attention to issues of major importance for the future of theoretical biology.

See also Organismic Biology; Philosophy of Biology; Vitalism.

Bibliography

ADDITIONAL WORKS BY BERTALANFFY

"An Outline of General System Theory." *British Journal for the Philosophy of Science* 1 (1950): 134–165.

"Problems of General System Theory." *Human Biology* 23 (1951): 302–311.

Bertalanffy, Ludwig von, and A. Rapoport, eds. *General Systems Yearbook.* Published yearly since 1956.

WORKS ON BERTALANFFY

Buck, R. C. "On the Logic of General Behavior Systems Theory." In *Minnesota Studies in the Philosophy of Science,* edited by H. Feigl and M. Scriven. Minneapolis: University of Minnesota Press, 1956. Vol. I, pp. 223–238.

Hempel, Carl G. "General System Theory and the Unity of Science." *Human Biology* 23 (1951): 313–327.

Jonas, Hans. "Comment on General System Theory." *Human Biology* 23 (1951): 328–335.

Medawar, P. B. Review of *Problems of Life. Mind,* 43 (1954): 105–108.

T. A. Goudge (1967)

BIBLIOGRAPHIES OF PHILOSOPHY

See "*Philosophy Bibliographies" in Volume 10*

BIEL, GABRIEL

(c. 1410–1495)

Gabriel Biel, the Ockhamist philosopher and theologian, was born at Speyer, Germany, and died at Einsiedel (Schönbuch). He studied philosophy and theology at Heidelberg and Erfurt, joined the Brethren of the Common Life, and became a professor of theology (1484) at the newly founded University of Tübingen, where he taught the "modern way," that is, according to the nominalist position of William of Ockham. Biel's "Commentary on the Sentences" (*Epithoma Pariter et Collectorium Circa IV Sententiarum Libros,* Tübingen, 1495) is a skillful summary of Ockham and a collection of the views of other medieval thinkers from Anselm to John Duns Scotus. Widely read in the German universities, Biel exerted a strong influence on Martin Luther (see P. Vignaux, *Luther, Commentateur des Sentences,* Paris, 1935).

As a philosopher, Biel was quite ready to criticize and to offer his own developments of Ockham's nominalism. Basically a theory of knowledge, his thought had some influence in ethics and political philosophy. For Biel formal logic displaced metaphysics because he considered universals to be but names (*nomina*) arbitrarily applied to classes; he considered all existents to be completely individual in character. Essence and existence are not really distinct principles in things but are merely distinguished in thought.

Biel's psychology was, like Ockham's, close to Augustinianism: the powers of the soul are not distinct faculties; intellect is the soul understanding, will is the soul desiring and loving. Biel was a psychological voluntarist; for him the most important psychic activity of man was will-

ing. He taught that all man's conscious activities entailed some use of will. Man was viewed as a volitional rather than rational animal.

In practical philosophy, he considered moral goodness to consist in volitional conformity to God's will. The obligatory force of law has no basis in the nature of created things but is solely due to the fact that God has willed a certain action to be right. This is moral and legal voluntarism. "God could command that a man deceive another through a lie," wrote Biel, "and he would not sin" (*Epithoma,* II, 38, q. 1, G).

See also Anselm, St.; Augustinianism; Duns Scotus, John; Essence and Existence; Luther, Martin; Ockhamism; Psychology; Voluntarism; William of Ockham.

Bibliography

Bonke, E. *Doctrina Nominalistica de Fundamento Ordinis Moralis apud Guillelmum de Ockham et Gabriel Biel.* Rome, 1944. Pp. 57–83.

Davitt, T. *The Nature of Law.* St. Louis: B. Herder, 1951. Pp. 55–68.

Feckes, C. *Die Rechtfertigungslehre des Gabriel Biel und ihre Stellung innerhalb der nominalistischen Schule.* Münster, 1925.

Vernon J. Bourke (1967)

BILFINGER, GEORG BERNHARD
(1693–1750)

Georg Bernhard Bilfinger was the German philosopher who coined the expression Leibniz-Wolffian philosophy for the view he expounded. Bilfinger, whose family name was also spelled Buelffinger, was born in Kannstadt, Württemberg. He studied theology at Tübingen, and mathematics and philosophy at Halle under Christian Wolff. He was appointed extraordinary professor of philosophy at Tübingen in 1721, but after Wolff's expulsion from Halle in 1723, Bilfinger was accused of atheism and deprived of his positions. On Wolff's recommendation he was appointed professor of philosophy and academician in St. Petersburg. His growing reputation as a natural philosopher caused Duke Eberhard Ludwig of Württemberg to recall him to Tübingen as professor of theology. In 1735 the new Duke Karl Alexander of Württemberg called Bilfinger to his capital, Stuttgart, as a member of the privy council. Bilfinger became president of the Consistorium, a council for ecclesiastical and educational affairs, and in this capacity permitted Pietism to be taught in Württemberg.

Although Bilfinger's doctrines are quite close to Wolff's, he showed a certain originality, discussing Wolff's doctrines critically and frequently accepting them only with reservations. In an early work he held, against John Locke, the view that there are innate ideas in the human mind, identifying them with axioms. In psychology he did not accept the distinction, introduced by Wolff, between empirical and rational psychology, but proceeded in a more traditional manner. In his later writings, Bilfinger referred less frequently to Wolff.

The most independent part of Bilfinger's system was his theory of possibility, expounded in his main work, *Dilucidationes Philosophicae de Deo, Anima Humana, Mundo et Generabilis Rerum Affectionibus* (Tübingen, 1725). He asserted that the notion of possibility is more fundamental than the principles of identity and contradiction. Possible things are not absolute beings in an independent realm of ideas, but they depend for their existence on God's understanding (not on his will). It is a part of God's essence to think possible things as they are, but they are, only insofar as God thinks them.

See also Leibniz, Gottfried Wilhelm; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY BILFINGER

De Harmonia Animi et Corporis Humani Maxime Praestabilita ex Mente Illustris Leibnitii, Commentario Hypothetica. Frankfurt and Leipzig, 1723.

"De Viribus Corpori Moto Insitis et Illarum Mensura." *Commentarii Academiae Petropolitanae,* Vol. 1 (1728). A famous essay on the measurement of forces.

Praecepta Logica, edited by C. F. Vellnagel. Jena, Germany, 1739.

Varia in Fasciculos Collecta. 3 vols. Stuttgart, 1743.

WORKS ON BILFINGER

Kapff, P. "G. B. Bilfinger als Philosoph." *Württembergischen Vierteljahrshefte für Landesgeschichte,* n.f. (1905).

Liebing, H. *Zwischen Orthodoxie und Aufklärung. Das philosophische und theologische Denken G. B. Bilfingers.* Tübingen: Mohr, 1961.

Wahl, Richard. "Professor Bilfingers Monadologie und prästabilirte Harmonie in ihrem Verhältniss zu Leibnitz und Wolff." *Zeitschrift für Philosophie und philosophische Kritik* 85 (1884): 66–92 and 202–231.

Giorgio Tonelli (1967)

BINET, ALFRED
(1857–1911)

Alfred Binet, the French psychologist, was born at Nice. The son of a doctor and an artist, Binet studied at the Sorbonne, qualifying in 1878 in both law and science. He embarked immediately on a doctorate under Edouard Balbiani, embryologist and professor at the Collège de France, whose daughter Binet married in 1884. In the same year he submitted an article on the fusion of images to *La revue philosophique*. The editor, Théodule Ribot, persuaded him in due course to devote his energies to psychology. Through Charles Féré, Binet came to work with Jean Charcot at the Salpêtrière hospital.

Binet is known mainly for his work, with his younger colleague Théodore Simon, in devising tests for assessing children's intelligence. The Binet-Simon scale, published in 1905 and revised in 1908 and 1911, constituted the first systematic, effective, and widely accepted attempt to devise sets of simple verbal and nonverbal tasks, performance on which could be quantified with a fair degree of objectivity, and on which norms for different age groups in the school population were carefully worked out. The principal American versions were produced, revised, and restandardized by L. M. Terman and his colleagues at Stanford University in 1916 and 1937. It was, however, Binet and Simon's careful studies that showed the necessity of valid data to ascertain the intellectual skills and concepts normally to be expected of children at each age before any assessment of a child's retardation can fairly be made. The revised tests are still employed for research and clinical purposes, although increasing use is now being made of the Wechsler tests.

Binet himself was well aware that cultural factors have a bearing on test performance and that interestingly different patterns of results on various subtests might be shown by children achieving similar overall scores. Hence, the conception of an intelligence quotient (IQ) as

$$\frac{\text{mental age}}{\text{chronological age}} \times 100$$

popularly linked with Binet's name, in fact runs counter to his stress on studying and appreciating individual differences.

The practical utility of the Binet-Simon scale has overshadowed to a large extent the rich background of inquiries from which the tests were developed. A man of wide theoretical and practical interests, Binet wrote in lucid and lively French a dozen books and some 250 articles, many of which appeared in *La revue philosophique* and in *L'année psychologique*, of which he was the editor. Seven of his books and a few articles appeared in English, which Binet wrote and spoke fluently. The *Psychologie des grands calculateurs et joueurs d'échec* (Paris, 1894), and *L'étude expérimentale de l'intelligence* (Paris, 1903), the latter reporting studies of his own children, remain neglected classics of French psychology. Both works provided evidence of individual differences in imagery and evidence that images could be less important in thinking than the associationists supposed. Furthermore, these studies, especially the former, showed that the subsequent line of thought was affected by the nature and presentation of the problem a thinker was asked to solve, by the mental set induced by that problem, and by his attitudes in other respects. The studies of his young daughters illustrate Binet's patient, systematic mode of inquiry into children's thought processes, and they enhance understanding of the developmental approach to psychology to which Jean Piaget was the heir.

Chronological scrutiny of his writing shows Binet's work on intelligence to have been the practical outcome of prolonged theoretical and experimental study of the nature of thought processes—subnormal, normal, outstanding, and abnormal. These investigations were carried out in hospitals, notably the Salpêtrière, in schools, and in the psychological laboratory at the Sorbonne, of which Binet became director. Influenced by Hippolyte Taine in France and by the British empirical tradition (including J. S. Mill, Alexander Bain, and Francis Galton), Binet had started as a narrowly orthodox associationist. His evidence for conceptual processes not involving visual imagery anticipated some of the Würzburg experimental findings on "imageless thought." This evidence and that found by Binet and his collaborators for central factors, for unconscious processes, and for attitudes influencing a train of thought led Binet slowly to change his standpoint. In doing so, he moved from treating thinking by analogy with visual inspection to emphasizing the affinities of thought and action and to stressing the importance of developmental studies. Such an approach has proved more acceptable in the 1960s than when Binet died, unfortunately leaving his own research and theory incomplete.

See also Bain, Alexander; Empiricism; Mill, John Stuart; Piaget, Jean; Psychology; Scientific Method; Taine, Hippolyte-Adolphe.

Bibliography

WORKS BY BINET

"Mental Imagery." *Fortnightly Review* 52 (1892): 95–104.

"The Mechanism of Thought." *Fortnightly Review* 55 (1894): 785–799. Except for this and the preceding reference, all of Binet's works are listed in the Varon monograph (see below).

"L'intelligence des imbéciles." *L'année psychologique* 15 (1909): 1–147. Written with Théodore Simon. This and the following article are of salient importance for understanding Binet's later treatment of thinking.

"Qu'est ce qu'une émotion? Qu'est ce qu'un acte intellectuel?" *L'année psychologique* 17 (1911): 1–47.

WORKS ON BINET

Bertrand, F. L. *Alfred Binet et son oeuvre*. Paris, 1930.

Groot, A. D. de. *Thought and Choice in Chess*. The Hague: Mouton, 1965.

Reeves, Joan Wynn. *Thinking about Thinking*. London, 1965. Ch. 7.

Varon, Edith J. *The Development of Alfred Binet's Psychology*. Psychological Monographs, Vol. 46 (No. 207). Princeton, NJ, and Albany, NY: Psychological Review, 1935. Includes a full bibliography of Binet's works except for the first two cited above.

Wolf, Theta A. "An Individual Who Made a Difference." *American Psychologist* 16 (5) (1961): 245–248.

Joan Wynn Reeves (1967)

BINSWANGER, LUDWIG

(1881–1966)

Ludwig Binswanger, the Swiss psychiatrist whose school of *Daseinsanalyse*, or existential analysis, is the most extensive attempt to relate the philosophies of Edmund Husserl and Martin Heidegger to the field of psychiatry, was born in Kreuzlingen, Thurgau, Switzerland, into a family line of eminent physicians and psychiatrists. After attending the universities of Lausanne, Heidelberg, and Zürich, he received his medical degree from Zürich in 1907. In 1910 he succeeded his father, Dr. Robert Binswanger, as chief medical director of the Sanitorium Bellevue, an institution founded by his grandfather at Kreuzlingen. He relinquished his directorship in 1956.

Daseinsanalyse is an original amalgam of phenomenology, Heideggerian existentialism, and psychoanalysis, the goal of which is to counter the tendency of scientific psychology to view man's being as solely that of a natural object. However, the school does not seek spheres of human existence that argue against the explanatory power of psychoanalysis. Binswanger complained of the overreductionism of natural science as applied to humankind, but in doing so he was not questioning science's ability to explain; he was, rather, urging that that which is being explained be kept in mind in its full phenomenal reality. Binswanger is a phenomenologist in that he demands a presuppositionless discipline in which the investigator can apprehend the world of the patient as it is experienced by the patient. To this end he limits his analysis to that which is actually present (or immanent) in the patient's consciousness. He seeks the essential structure of these phenomena without relying on reductive theory, his aim being to allow the phenomena to speak for themselves. As an existentialist he views the essential structures that the phenomena reveal on their own terms as "universals with power." That is, he sees them as the matrix within which the individual's world and self—his essence—are determined. He seeks in each patient a general context of meaning within which the patient exists. He calls this meaning-context the transcendental category of that patient's world design.

This notion of a general existential meaning-context must be understood as that which expresses with equal validity all aspects of the patient's life and world. The criterion of a complete expression is based on Heidegger's ontology of man and includes his orientation in space, his mode of being in time, his relation to his bodily life and to his fellow man, his way of thinking, and his fears and anxieties. For example, a universal such as continuity is equally understandable and expressive in reference to time (continuity of events versus the sudden and unexpected), space (contiguity), relationships with others (for example, oedipal ties or bonds), and the individual's own world ("inner" continuity, continuity of feelings or of affections). But such explanatory categories as aggression or libidinal energy emphasize one aspect of man's being as most real and are therefore rooted in a one-sided ontology of human existence.

What psychoanalysis takes as conditioning factors—such as instinct or childhood sensations—are regarded by Binswanger as already being representations of a basic world design. It is not that Binswanger wants to push back the causal chain beyond instincts or childhood sensations, but rather that the causal chain itself, as described in scientific depth analysis, must be viewed as a whole, without any a priori privileged reference point in terms of which all else is to be explained. Explanation in terms of a privileged reference point presupposes a theory, and a theory assumes a world outlook—in this case the world outlook of natural science. Binswanger does not, therefore, use the past to account for the present. He sees the past of a patient as existing in the present in that the entire world design—within which a particular event in

the past "conditioned" a present neurosis—*is* the patient. Therefore, the present, or the conscious, or the manifest content of dreams and the manifest verbal expressions, all point to a unity or category(ies) that is the basis of the patient's world. In other words, because the self cannot experience a "pure" event outside of a meaning-context, even if the self be that of a child, it is that source meaning-context which Binswanger seeks to apprehend.

Binswanger does not offer his approach as a substitute for psychoanalysis; insofar as the goal of psychiatry is intervention in the patient's life—manipulation of or change in it—only a scientific approach, such as psychoanalysis or clinical psychiatry, is adequate. For Binswanger, phenomenology and reductive explanation are two complementary aspects of the *Geisteswissenschaften*, including psychology. Phenomenology can provide us with an essential description of the data, and phenomenological existentialism can provide a full dynamic understanding of the individual's life on his own terms. But if we are willing and find it necessary to transform and control phenomena, natural science is at present our major tool. However, whereas in the natural sciences we confer meanings, in the *Geisteswissenschaften* the phenomena under investigation are themselves meanings to a self, and it becomes necessary phenomenologically to receive these meanings on their own terms.

See also Heidegger, Martin; Husserl, Edmund.

Bibliography

PRINCIPAL WORKS BY BINSWANGER

Wandlungen in der Auffassung und Deutung des Traumes von der Griechen bis zur Gegenwart (Changes in understanding and interpretation of the dream from the Greeks to the present). Berlin: Springer, 1928.

Über Ideenflucht (On the flight of ideas). Zürich, 1933.

Einführung in die Probleme der allgemeinen Psychologie (Introduction to the problems of general psychology). Berlin, 1942.

Grundformen und Erkenntnis menschlichen Daseins (Basic forms and cognition of human existence). Zürich: Niehans, 1942; 2d rev. ed., Zürich, 1953.

Ausgewählte Vorträge und Aufsätze (Selected lectures and essays). 2 vols. Bern: Francke, 1947–1955.

Drei Formen Missglückten Daseins (Three forms of unsuccessful dasein). Tübingen: Niemeyer, 1956.

Schizophrenie. Pfullingen, 1957. Contains Binswanger's well-known clinical studies of Ilse, Ellen West, Jürg Zünd, Lola Voss, and Suzanne Urban.

Melancholie und Manie. Pfullingen: Neske, 1960.

Traum und Existenz (with an introduction by M. Foucault). Bern: Gachnang & Springer Verlag, 1992.

Ausgewählte Werke in vier Bänden. Heidelberg: R. Asanger Verlag, 2004.

ENGLISH TRANSLATIONS

May, Rollo, Ernest Angel, and Henri F. Ellenberger, eds. *Existence*. New York: Basic, 1958. This and the following work contain some of Binswanger's works in translation.

Needleman, Jacob. *Being-in-the-World*. New York: Basic, 1963.

WORKS ON BINSWANGER

Keen, E. *A Primer in Phenomenological Psychology*. New York: Holt, Reinhart & Winston, 1975.

Needleman, Jacob. *Being-in-the-World*. New York: Basic, 1963. A philosophical critique of Sigmund Freud and Binswanger, taking account of the background of Immanuel Kant and Heidegger.

Sonnemann, Ulrich. *Existence and Therapy*. New York: Grune and Stratton, 1954. Brilliant and unreadable.

Speigelberg, H. *Phenomenology in Psychology and Psychiatry*. Evanston: Northwestern University Press, 1992.

Van Den Berg, J. H. *The Phenomenological Approach to Psychiatry*. Springfield, IL: Thomas, 1955. Simple and clear.

Jacob Needleman (1967)
Bibliography updated by Thomas Nenon (2005)

BIOETHICS

Bioethics is an interdisciplinary field of study dealing with practical ethical issues roughly at the intersection of morality, medicine, and the life sciences. Within philosophy, bioethics is one of several different areas of applied ethics, a domain within general normative ethics. The term "bioethics" was coined in 1970, but the development of bioethics as a discipline may be dated to the late 1960s or early 1970s, depending on which historical markers are used.

The scope of bioethics as a discipline is not entirely fixed, it is important to note. At least three competing visions are available. On the most restricted view, bioethics simply reduces to biomedical ethics, which encompasses ethical issues relating to the practice of medicine broadly understood and the pursuit of medical research. Even on this restricted view of bioethics, the scope extends to the ethics of our use of nonhuman animals in biomedical research, for example. On the second understanding, bioethics encompasses, in addition to biomedical ethics, ethical issues related to the life sciences and technologies. On this understanding, also included is consideration of environmental issues, for example, issues such as genetic modification of plants or the use of cloning technologies to revive extinct species of animals or plants. According to the widest view, bioethics includes the biological aspects of environmental ethics, issues related to nonhuman-animal use, and biomedical ethics. On this understanding, the ethical dimensions of vegetarianism and how global warming affects biotic communi-

ties are also bioethical issues. Interestingly, this widest understanding of the term is closest to the meaning given by biochemist Van Rensselaer Potter, who originally coined the term. However, it also offers the least common understanding of the term within the discipline. The second sense probably offers the most common understanding within the discipline, while most people working in the field of bioethics work on issues in biomedical ethics, the first sense. This entry will explore issues related to all three senses of bioethics.

LOCATING BIOETHICS WITHIN PHILOSOPHICAL ETHICS

Ethics is the philosophical study of morality. It is to be distinguished from the empirical study of moral norms and practices. This second area of investigation is sometimes also called simply "ethics," but to distinguish it from philosophical ethics, it may more felicitously be called descriptive ethics. As a philosophical discipline, ethics may be further divided into metaethics, general normative ethics, and various areas of applied ethics. Metaethics is concerned primarily with reflections on ethics itself. Some issues within metaethics include the meaning of moral terms like "ought," "right," and "virtue"; the metaphysical status of moral norms; the proper grounds for justifying moral claims; and the nature of moral knowledge.

Normative ethics, by contrast is concerned in general with positive guidance to living morally. Normative ethics concerns questions about how to act, what kind of character to develop, and what values to live by. Within normative ethics generally, various ethical theories have been developed as guides in answering these questions. Normative moral theories lay out the structure for particular fundamental sources of normative moral value. Examples include utilitarianism (a type of consequence-based ethics), deontology (a duty-based ethics), and virtue ethics (a character-based ethics). However, normative ethics may also proceed without any particular theoretical structure and may engage directly with the various issues at stake in practical moral living.

Applied ethics is normative ethics at the level of engagement with various specific topics in practical moral life. As such, applied ethics may proceed by following some more general normative theory, by following some methodology or theory particular to the area of study, or without following any specific theoretical or methodological underpinning. The term "*applied* ethics" implies that general normative theories are simply interpreted in light of specific moral problems to generate practical moral answers; however, this is seldom actually the case. A better term might be "practical ethics."

Bioethics, then, is a type of practical ethics. It is on the same philosophical level as business ethics, environmental ethics generally (unless understood as a subset of bioethics), cyberethics, and a host of other specific fields dealing with particular areas of complex lived morality. The division into areas of practical ethics may not be particularly neat. As already discussed, bioethics may be understood as distinct from, or inclusive of, environmental ethics. Other areas of practical ethics, such as professional ethics, overlap with various other fields (for example, business ethics, legal ethics, and medical ethics).

HISTORY AND SOCIAL CONTEXT

Many of the specific issues addressed in bioethics have historical roots. The issue of physician-patient confidentiality was addressed in the Hippocratic oath. The bioethics of how we treat animals has roots in the work of such historical figures as Porphyry (232–309) in his treatise *On Abstinence from Animal Food* (discussed in Sorabji 1993). More recently, the requirement that human subjects voluntarily consent to medical research was spelled out in the Nuremberg Code (International Military Tribunal 1949), following the "doctors trial" for atrocities committed as part of the holocaust.

Despite these and multiple other sources of historical precedence, the discipline of bioethics coalesced only in the very late 1960s to early 1970s. The social forces behind this formation into a specific disciplinary field include a growing social awareness of issues of medical paternalism and some unethical practices in medical experimentation; the consumer, feminist, and civil rights movements; the increased institutionalization of health care and with it a growing concern for issues of allocation; advances in biotechnology and biomedicine; growing awareness of issues of sustainable economic growth and environmental impact; and rising awareness of the conditions of animals in newly evolving factory farms.

Since its inception, bioethics has taken deep root in academia, professional education, public policy, the law, and public deliberation. Bioethics courses are offered to undergraduates as part of humanities and science curricula and to medical and other health-professional students. A variety of centers for bioethics research have been established, and numerous government commissions dealing with bioethical issues have been formed and have had varying impact on public policy. A U.S. commission of high impact was the National Commission for the Protection of Human Subjects of Biomedical and

Behavioral Research, the source for the *Belmont Report* (1979), spelling out the ethical principles guiding experimentation on human subjects. Those working in bioethics have extensively scrutinized legal cases, such as the 1989 U.S. Supreme Court case regarding Nancy Beth Cruzan, whose parents sought to withdraw life support after Ms. Cruzan fell into a persistent vegetative state. The law has also been influenced by bioethical analysis, for example by the inclusion of those working in bioethics as expert witnesses in trials. Finally, as biomedical ethical issues and advances in the life sciences have received more attention by the mass media, public arenas for debate over these issues has also grown.

THE DISCIPLINE

Bioethics as a discipline crosses over other disciplinary boundaries, both within and outside of philosophy. Here the focus is on the philosophical aspects of the discipline. Nonetheless, some general remarks about the discipline as a whole are necessary in order to view the philosophical area of bioethics in context. This is important also because there is controversy within the field of bioethics, as has been seen, about what counts as bioethics, but also about the extent to which it is a unified discipline, and about exactly which general methodologies and areas of expertise are relevant.

It is uncontroversial that moral philosophy plays a central role within bioethics, and also that other areas of philosophical study are implicated by the topics relevant to bioethics. For example, political philosophy is central to issues of distributive justice in access to health care and to public-health measures affecting human health, as well as in adjudicating questions of public, institutional, and governmental decision-making about controversial bioethical issues such as use of stem cells and cloning of human somatic cells. In addition to addressing general issues of the relationship between the law and practical morality, philosophy of law is also relevant to determining how case law and legislative law relate to ethical decision making. Finally, although the connection is less widely recognized, philosophy of science is crucial to the investigation of some central conceptual issues in bioethics, for example, the nature of the scientific facts that often play a central role in practical ethical decisions and the meaning of the concept of human health often invoked in such ethical distinctions as that between using genetic technologies for enhancements versus using them therapeutically. Even the question of what counts as a biological kind or species is central in determining whether legitimate ethical distinctions can be made between human beings and other nonhuman animals for how we treat them, and in dealing with ethical issues related to the transhuman movement for improved humans.

While philosophy is the right place to look for expertise in clarifying the issues involved in answering moral questions and for theoretical structures intended in part to answer those questions, it is not clear that philosophy is the right place to look for positive answers to specific practical moral questions. We might then delineate two kinds of ethics expert: academic and directive. The broader issue is whether expertise in academic bioethics provides any expertise in directive ethics. There is no clear consensus within the discipline on this issue.

Outside of philosophy, disciplines relevant to bioethics include the social sciences, law, and medicine, as well as those within what has been termed the medical humanities. With such a wide range of participants, it is not surprising that there are disagreements over the scope of bioethics, the relevance of different fields to the discipline, the training necessary to qualify as a member of the discipline, the kind of expertise that such members have, and the legitimacy of the very term "discipline" in describing this diverse range of fields, methodologies, and topics of interest. What is certain, however, is that enough overlapping consensus exists to create a discipline identifiable to its members and to the general public.

METHODS IN BIOETHICS

How do moral philosophers and others working in bioethics go about dealing with the complex moral issues at the heart of many bioethical issues? The answer to this question is quite complex. Three basic approaches are available. The first approach applies established general normative theories to particular issues in bioethics. The second embraces one or several methodologies specifically developed for bioethical issues. The third method either avoids or rejects outright specific methodologies outside of basic philosophical and critical-thinking skills. These methods will be discussed in turn.

GENERAL NORMATIVE THEORIES

Perhaps surprisingly, the least common approach to bioethics is the approach of applied ethics, that is, the application of general normative theories to specific moral problems in bioethics. Nevertheless, there is some substantial work in bioethics that proceeds roughly along these lines. Examples include Peter Singer's utilitarian approach to the ethics of how we treat animals (2002), Tom Regan's deontological approach to that same issue

(1983), and Rosalind Hursthouse's essay on abortion from a virtue-ethical perspective (1991). Of work that follows this general model, the most commonly appealed-to theories are the three just mentioned: utilitarianism, deontology, and virtue ethics. These theories are actually umbrella categories under which fall a number of specific theories.

Even the work in bioethics that falls into this first category is not really simply straightforward application of normative theories to particular bioethical problems. Rather, some particular vision of a general theoretical structure is rendered, and the bioethical problem is interpreted in light of that theoretical structure. This might involve, among other things, modifying the theory to fit the issue and arguing for some particular interpretation of the practical implications of the theory over others.

Some theories lend themselves more readily than others to application to specific moral problems. However, even the application of these theories requires extensive interpretation. For example, according to hedonistic act utilitarianism, the right action is the one that maximizes pleasure (or minimizes pain) for all those affected. In principle, then, the answer to the question "Which action is right?" is a matter of calculation of hedonistic utility output. Yet we still need to know which outcomes count as pleasures and pains, of what strength and type, and for what range of beings. (Does the calculation include sentient nonhuman animals? Future persons?)

Despite the term "applied ethics," then, normative moral theories are at a level of abstraction not conducive to straightforward application to particular moral problems. Making matters more difficult, normative theories conflict with one another, sometimes in ways that imply different practical recommendations. In such cases one has to decide which theoretical approach is the best before tackling the moral problem at hand. Moral theories are still helpful in making practical moral decisions, since they provide essential analysis of basic moral values, coherent frameworks for understanding moral issues, and general justificatory strategies for particular approaches to morality. Yet it is not even clear that a proper goal of normative moral theory is to generate specific moral directives. On a virtue-ethical view, for example, moral guidance in specific practical contexts flows from a virtuous character, not from abstract theoretical principles.

METHODS SPECIFIC TO BIOETHICS

In part because of the problems associated with the application of general normative theories of morality to problems in bioethics, a number of methodologies specific to bioethics have been developed since the inception of the discipline. It is important to note at the outset that these methods have been developed largely for biomedical ethics rather than for bioethics in the broader sense embraced in this entry. Among these methods, the most well known is the principles-based approach developed largely by Tom Beauchamp and James Childress in successive editions of *Principles of Biomedical Ethics* (2001) but also inspired by the more general approach to ethics favored by earlier philosophical figures such as W. D. Ross (1930). The principles approach relies on a variety of prima facie norms, the most prominent of which are four principles: beneficence, nonmaleficence, justice, and respect for autonomy. The source of these principles is supposedly common morality. No single principle is a trump principle, since each may be overridden by considerations deriving from the others in specific contexts. How the principles are spelled out in specific situations and which one(s) hold sway in case of conflict is determined by a process of specification and balancing.

Another influential approach is the casuistic approach revitalized from ancient and medieval roots by Albert Johnson and Stephen Toulmin in *The Abuse of Casuistry: A History of Moral Reasoning* (1988). For casuists, general ethical principles stem only from the analysis of paradigm cases. These paradigm cases offer clear moral outcomes and create a set of initial presumptions about how to resolve other cases. These presumptions hold sway unless we come across exceptional circumstances. When such exceptional circumstances arise, we must go through a process of analogical reasoning, which includes identifying the ethical values at issue, the alternate courses of action, the morally relevant ways in which cases of the sort at issue can differ (the casuistic factors), and the relevant paradigm case(s) (Strong 2000).

A third approach, narrative bioethics, is a relative newcomer but has close ties to antitheoretic trends in normative moral philosophy generally. Insofar as narrative bioethics is not a single approach, it is hard to specify exactly what it amounts to methodologically. However, a couple of themes can be drawn out to give a flavor for this type of approach. First is the ethical significance of the various narrative voices involved in ethically complex situations. In opposition to casuistry, which relies on some single paradigm case and thereby a neutral voice, narrative bioethics focuses on telling the story from the viewpoints of all the participants. In this way we can see that the neutral voice in the paradigm cases (the physicians, lawyers, and/or judges) may in fact be the most powerful

voice. Further, narrative bioethics focuses not on principles as a way of solving ethical quandaries, but rather on the different ways of telling the story and the comparative choices supported by these ways. Thus the stories themselves have normative impact.

While none of these approaches offers a general normative theory for bioethics, the principles-based approach comes closest, whereas the narrative and casuist approaches offer methods to deal with bioethical issues but eschew theory. A final approach offers a general theory of practical morality with particular focus on bioethics issues. This is the theory offered by Bernard Gert, Charles Culver, and K. Danner Clouser in *Bioethics: A Return to Fundamentals* (1997). On their view, morality is a public system whose purpose is to minimize the amount of evil suffered by those protected by it. In opposition to the principles-based view of bioethics, this view has more specific moral rules as fundamental touchstones.

Despite the availability of these various methods, much work in bioethics actually proceeds without a specified method or by a piecemeal approach. This kind of no-method method can be criticized for its ad hoc nature and for its lack of any specific justificatory framework. Even without specific appeal to some particular methodological framework, bioethical analysis at its best avails itself of a number of useful tools in approaching areas of ethical conflict, including gathering and sifting through morally relevant factual information; providing conceptual clarity on the moral concepts at issue; engaging in casuistic reasoning (without necessarily embracing casuistry as a methodology); and offering analysis, critique, construction, and revision of moral arguments.

THEMES IN BIOETHICS

As already discussed, bioethics as understood here includes at least biomedical ethics and ethical issues related to advances in the life sciences and life-science technologies, but may be broadened to include environmental ethics and ethics related to our treatment of nonhuman animals generally. The specific topics in bioethics are numerous and change in character and focus over time as the field advances. Each anthology of bioethics lists and groups the topics somewhat differently. A small sampling of these topics from four well-known anthologies includes justice in access to health care, mother-fetus relations, research involving human subjects, reproductive technologies, eugenics, genetics, health-care policy, physician-assisted suicide and euthanasia, medical confidentiality, the physician-patient relationship, informed consent, research involving animal subjects, definitions of death, human cloning and stem-cell research, and organ donation. Most of these topics include subtopics and may also be subsumed under more general headings. Developing any additional list of specific topics in bioethics here would be unhelpful; more useful is to focus on a few general philosophical issues at the overlap of a variety of topics in bioethics.

MORAL STATUS

A key issue in several central topics in bioethics is the moral status of various animal species, the environment, and human beings in various life stages. The issue of moral status is in part a question about how far the moral community extends, that is, what the scope is of those entities considered to have direct moral value. An answer to the scope question does not resolve the issue entirely, since there may be different degrees of moral status. For example, we might think that both pigs and adult human beings have some direct moral status, but still that the adult human being has more moral status.

One way in which the issue of moral status has been addressed is through the concept of personhood. This concept introduces a normative category for those kinds of beings with full moral status, namely persons. In the philosophical debate, persons are not simply all and only human beings. Rather, it is normally assumed that some capacities are required to attain the status of a person. These capacities must be judged by their moral relevance, and not simply along species lines. Some morally relevant capacities might be the ability to feel pain and to have pleasure, the ability to engage emotionally with others, the ability to act intentionally, and the ability to make rational choices. If the level of capacity required for personhood is drawn at the more basic abilities, then the category of persons will include many animals and most human beings. Alternately, if the line is drawn at the higher abilities, for example at the capacity for autonomous actions and choices, then many human beings and most animals will not be persons.

Within biomedical ethics, the issue of moral status is of crucial significance for topics such as abortion and for issues at the beginning and end of life, for example, issues of the moral acceptability of discontinuing life support for severely impaired newborns or humans in persistent vegetative states. For those wanting to extend full moral status to fetuses and severely impaired postnatal human beings, a problem arises of how to ground that moral status. If it is grounded in the particular capacities or potential capacities of those beings, then a relevant question is

whether the same moral status should also be extended to some nonhuman animals.

A central question in environmental ethics is the moral status of the environment in general and of particular ecosystems or other entities. Do two-hundred-year-old oak trees have sufficient moral status that it is wrong to cut them down independently of their effect on human beings or animals with moral status? Establishing the source of such moral status has been a source of difficulty within the field. Yet reflecting on questions like this may bring to light some of the limitations of a capacities-based approach to moral status. Moreover, even if oak trees, nonhuman animals, and human fetuses do not have individual moral status, it does not follow that we can ignore their well-being. Indeed, through such ethical resources as virtuous habits of character, the relationships of persons to other beings, and the effects of our treatment of such beings on other persons, we can establish a wide range of protections for nonpersons.

DISTRIBUTIVE JUSTICE

A very different core topic within bioethics focuses on questions of distributive justice, that is, what the proper distribution is of social resources and burdens. Looking more closely at this issue gives a sense of the different types of theoretical resources brought to bear in bioethics and also shows how the discipline has developed. Issues of distributive justice are usually approached not from the perspective of general normative moral theory, although they may be, but from theories in political philosophy. Examples of theories constructed to deal generally with issues of distributive justice are John Rawls's *A Theory of Justice* (1971) and the libertarian theory found in Robert Nozick's *Anarchy, State, and Utopia* (1974). However, as with normative moral theories, these general approaches to the distribution of social resources and burdens are often not well suited to the specific practical issues involved in bioethics.

To deal with issues of distributive justice at a level directly relevant to bioethics, a number of specific views have been developed. Many of these are varieties of egalitarianism, a general view in distributive justice focusing on the moral foundations for an equal distribution of social goods and resources. A major question for egalitarian theories in bioethics is what should be equally distributed: health care, health outcomes, satisfaction with health? Other views of the ethically proper way to allocate health care rely on formal mechanisms, such as cost-effectiveness analysis, which has theoretical roots in utilitarianism. The least common approach to issues of distributive justice in bioethics is libertarianism, although this view has had some supporters.

The evolution of the particular topics involved in distributive justice in bioethics gives a sense of how the discipline has changed over time. Initially, there was little focus on issues of distributive justice except for the discussion of the just distribution of the burdens of research on human subjects. While the issue of research on human subjects has retained significance, with the growth of patient activism and the perception of promising new interventions the distributive focus shifted from protection from the burdens of research to assurance of equitable access to research protocols. As the ramifications of institutionally centered health care and various health-insurance mechanisms grew, issues of distributive justice in health care became focused on questions of access. Such questions as whether there is a right to health care came to the fore. These questions have remained significant, particularly in the United States, where the number of uninsured persons continues to rise along with the costs of health care.

To complicate matters further, an additional twist has been added into the mix, which is that health inequalities appear to be tied less to health-care access than to relative social and economic status. While providing equitable access to health care may retain moral significance as a matter of distributive justice, providing such access may make a relatively small dent in the problem of health inequality.

OTHER THEMES

In addition to the problem of moral status and topics in distributive justice, a number of other philosophical issues lie at the core of various specific topics in bioethics. While it is impossible to discuss all these topics here, several significant questions should be noted. One set of issues focuses on the science side of bioethics. First, what is the role of scientific facts in moral decision making? In this area, relevant questions might be, "How significant is an understanding of the biological developmental stages of human fetuses to the morality of abortion?" And, "What difference does it make to issues of distributive justice whether some genetic predispositions to disease significantly lower some persons' life expectancies?" Second, how do advances in the life sciences affect ethical issues? In this area the main issue is whether advances in the life sciences actually create the need for new ethical concepts and models or whether they simply create an opportunity to reinterpret established ethical debates. Fields where this question is especially relevant include

human-somatic-cell cloning and assisted-reproduction technologies. A general underlying question with regard to these and other issues at the overlap of science and ethics is the moral relevance of naturalness.

Another set of core issues has to do with the development of role ethics as a way of understanding the specific obligations of physicians, other health-care workers, researchers, and scientists to particular populations. Here the main philosophical issue is whether individuals incur some obligations simply by occupying particular social roles or whether all obligations are versions of more general social obligations. The question is not whether physicians, for example, have a specific duty to protect the privacy of some medical information. All agree that this is the case, in addition to agreeing on a number of other specific duties that physicians have to their patients. The question is rather whether this duty is the product simply of occupying a specified social role or whether it is a duty that anyone with the requisite expertise in the particular relationship would also have. If there are obligations that are completely role dependent, then one can expect that some of these may conflict with other moral obligations, thus creating the potential for moral dilemmas. By contrast, if role obligations are contextual renderings of general moral obligations, then no such conflict can be expected.

Bioethics is a relatively young discipline that has already had a dramatic impact on academic curricula, public policy, public awareness of ethical issues, health care practices, and health sciences research. Continued advances in the life and health sciences on the one hand and continued disparities in health and health care on the other hand make it likely that bioethics as a discipline will continue both to grow as a field and to evolve in focus and methodology.

See also Environmental Ethics; Medical Ethics; Science, Research Ethics of.

Bibliography

Beauchamp, Tom L., and James F. Childress. *Principles of Biomedical Ethics*. 5th ed. New York: Oxford University Press, 2001.

Beauchamp, Tom L., and LeRoy Walters. *Contemporary Issues in Bioethics*. 6th ed. Belmont, CA: Wadsworth Press, 2003.

Buchanan, Allen E. "The Right to a Decent Minimum of Health Care." In *Securing Access to Health Care*, Vol. 2, by the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Washington, DC: U.S. Government Printing Office, 1983.

Casarett, David J., Frona Daskal, and John Lantos. "Experts in Ethics? The Authority of the Clinical Ethicist." *Hastings Center Report* 28 (6) (1998): 6–11.

Charon, Rita, and Martha Montello, eds. *Stories Matter: The Role of Narrative in Medical Ethics*. New York: Routledge Press, 2002.

Clarke, Stanley G., and Evan Simpson. *Anti-theory in Ethics and Moral Conservatism*. Albany, NY: SUNY Press, 1989.

Crosthwaite, Jan. "Moral Expertise: A Problem in the Professional Ethics of Professional Ethicists." *Bioethics* 9 (5) (1995): 361–379.

Daniels, Norman. "Equality of What: Welfare, Resources, or Capabilities?" *Philosophy and Phenomenological Research* 50, suppl. (1990): 273–296.

Daniels, Norman. *Just Health Care*. New York: Cambridge University Press, 1985.

DeGrazia, David. *Taking Animals Seriously: Mental Life and Moral Status*. New York: Cambridge University Press, 1996.

Engelhardt, H. Tristram, Jr. *The Foundations of Bioethics*. 2nd ed. New York: Oxford University Press, 1996.

Gert, Bernard, Charles M. Culver, and K. Danner Clouser. *Bioethics: A Return to Fundamentals*. New York: Oxford University Press, 1997.

Gruen, Lori, and Dale Jamieson, eds. *Reflecting on Nature: Readings in Environmental Philosophy*. New York: Oxford University Press, 1994.

Hargrove, Eugene C., ed. *The Animals Rights/Environmental Ethics Debate: The Environmental Perspective*. Albany, NY: SUNY Press, 1992.

Hursthouse, Rosalind. "Virtue Theory and Abortion." *Philosophy and Public Affairs* 20 (3) (1991): 223–246.

International Military Tribunal. "Permissible Military Experiments." In *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10*, 2: 181–184. Washington, DC: U.S. Government Printing Office, 1949.

Johnson, Albert R., and Stephen Toulmin. *The Abuse of Casuistry: A History of Moral Reasoning*. Berkeley: University of California Press, 1988.

Kuhse, Helga, and Peter Singer, eds. *Bioethics: An Anthology*. Oxford, U.K.: Blackwell Publishers, 1999.

Mappes, Thomas A., and David DeGrazia, eds. *Biomedical Ethics*. 5th ed. Boston: McGraw-Hill, 2001.

Nelson, Hilde Lindemann, ed. *Stories and Their Limits: Narrative Approaches to Bioethics*. New York: Routledge Press, 1997.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Potter, Van Rensselaer. "Bioethics: The Science of Survival." *Perspectives in Biology and Medicine* 14 (1970): 127–153.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Regan, Tom. *The Case for Animal Rights*. Berkeley: University of California Press, 1983.

Rhodes, Rosamond, Margaret P. Battin, and Anita Silvers, eds. *Medicine and Social Justice: Essays on the Distribution of Health Care*. New York: Oxford University Press, 2002.

Ross, William David. *The Right and the Good*. Oxford, U.K.: Oxford University Press, 1930.

Singer, Peter. *Animal Liberation*. New York: Ecco, 2002.

Sorabji, Richard. *Animal Minds and Human Morals: The Origins of the Western Debate*. Ithaca, NY: Cornell University Press, 1993.

Spielman, Bethany J., and Ben A. Rich, eds. *Expert Testimony: Bridging Bioethics and Evidence Law. Law, Medicine, and Ethics* 33 (2) (2005): 194–278.

Steinbock, Bonnie, John D. Arras, and Alex John London, eds. *Ethical Issues in Modern Medicine*. 6th ed. Boston: McGraw-Hill, 2003.

Strong, Carson. "Specified Principlism: What Is It, and Does It Really Resolve Cases Better Than Casuistry?" *Journal of Medicine an Philosophy* 25 (3) (2000): 323–341.

Rebecca L. Walker (2005)

BIOLOGY, PHILOSOPHY OF

See *Philosophy of Biology*

BLACK, MAX

(1909–1988)

The American analytic philosopher Max Black was born in Baku, Russia (now Baky, Azerbaijan). He read mathematics at Cambridge and, after he earned his BA in 1930, received a fellowship for research at Göttingen, Germany, where he wrote *The Nature of Mathematics* (1933). Returning to Britain, he was awarded a doctorate by the University of London for his dissertation *Theories of Logical Positivism* (1939) and held the position of lecturer and tutor in its Institute of Education from 1936 to 1940. After he emigrated to the United States in 1940, he was appointed to the faculty at the University of Illinois. In 1946 he moved to Cornell University, in Ithaca, New York, where, in 1954, he became Susan Linn Sage Professor and later helped found both the Society for the Humanities and the Program on Science, Technology, and Society. He was president of the Eastern Division of the American Philosophical Association; president of the International Institute of Philosophy; Tarner Lecturer at Trinity College (Cambridge), Guggenheim Fellow; Fulbright Fellow; and visiting fellow at Oxford, Cambridge, Princeton, Palo Alto, and Canberra. He also lectured on contemporary American philosophy in Japan (1957) and India (1962). He died in Ithaca in 1988.

Black's early years in Cambridge—where he attended classes of G. E. Moore, Frank Ramsey, Susan Stebbing, and Ludwig Wittgenstein—influenced his later teaching and writings. Along with his analytic orientation of C. D. Broad and Ramsey, Black retained a wide range of scientific and humanistic interests and a careful regard for the commonsensical approach that marked the writings of Moore and Stebbing; but the influence of Wittgenstein was the most profound. Black's first work, *The Nature of Mathematics*, was an exposition of the logicist, formalist, and intuitionist conceptions of mathematics; it paralleled Wittgenstein in declining to embrace any of the three theories or to propose a new one, and his subsequent doctoral study of logical positivism required coming to grips with Wittgenstein's *Tractatus*. His abiding interest in that work culminated in *A Companion to Wittgenstein's* Tractatus (Cambridge and Ithaca, 1964), a work some six times as long as the text it analyzed; it was posthumously reprinted, and admired—even imitated—for its astute and engaging combination of exegesis, explication of sources, and critical comment.

Black's commitment to philosophical analysis involved constructive work on small, well-defined problems with expository and critical discussion; hence the bulk of Black's contributions are essays rather than books. The exceptions are the two noteworthy books already mentioned and a logic text, *Critical Thinking* (1951). Other volumes were published, to be sure, but they are collections of essays rather than treatises. Black published some twenty books (including those edited and/or translated) and more than 200 essays and reviews.

Many of Black's essays take up problems or themes from Wittgenstein's later works, generally concentrating on the issues, especially meaning, rather than on Wittgenstein's texts. Black did not fret about the metaphysical status of meanings, since (as for the later Wittgenstein) explanations of meanings are explanations of how words are used, and it is a mistake to suppose that there are "such things as meanings to be categorized." One aspect of explanations of meaning involves formulating rules for the use of words, and Black (again following Wittgenstein) shows how seemingly necessary propositions often serve as surrogates for rule formulations. Black is aware that a certain vagueness or "looseness" pervades these rules governing ordinary usage, and he explores this dimension in several essays. One of his conclusions is that we normally presuppose that the looseness does not matter. This calling attention to presuppositions of linguistic acts is characteristic of Black. In other essays he calls attention to the contrasting presuppositions of definitions and assertions, and he gives a detailed comparison of presupposition and implication, with special reference to controversies about denoting phrases.

Black's conception of philosophy emphasizes the everyday practicality of linguistic analysis: "philosophical clarification of meaning is … as practical as slum clear-

ance and as empirical as medicine"—hence the title essay of one of his last books: "The Prevalence of Humbug." This all-too-prevalent humbug consists not only in logical fallacies but also in overvaluing speculation, ignoring or minimizing induction, and, at times, misplaced logical rigor. Therefore Black deplores not only broad-brush dismissals of rationality and science but also the excesses of pettifogging rationalism and scientism; he lacks sympathy with Hume's criticism of induction and philosophical complaints about vagueness. Here we see Black's respect for common sense, which he learned in part from Moore. The vagueness of ordinary language works partly because normal usage has roots in truth: "To say that a word is correctly used in accordance with normal usage, in certain circumstances, is to say that a certain sentence containing the word is, in those circumstances, true." This remark works in defense of the much-criticized paradigm-case argument, because the circumstances envisaged will be a paradigm case for that word. In other essays, Black augments references to paradigm cases by discussing models and metaphors, both of which also occur in ordinary language but exemplify different sorts of justified vagueness. One later essay, "Reasoning with Loose Concepts," (1963) argues that we can be sure of clear cases without knowing at what point cases cease to be clear. Paradigm cases, however, do not provide a road from language to metaphysics: "The conception of language as a mirror of reality is radically mistaken."

As an undergraduate, Black heard Moore deliver the Tarner Lectures at Trinity College in 1929, so he was delighted to receive an invitation to deliver them in 1978. His topic, "Models of Rationality," conformed to his customary piecemeal pattern of output in yielding not in a book but a series of essays that were incorporated into several later publications. One theme running through the lectures is that models of rationality cannot eliminate the need for judgment; hence formal schemes, such as those employed by economists in decision theory and choice theory, are bound to remain heuristic rather than definitive.

Black's interests had an Enlightenment breadth; the topics of his essays range from formal logic to poetry. In the philosophy of science, he argued eloquently and persuasively for induction and commented on perception, cosmology, decision theory, aesthetics, and sociology, all while retaining his early interest in mathematics. His work in philosophy of language included reviews of the work of many of his contemporaries, including Russell, Dewey, Wittgenstein, Korzybski, Carnap, Tarski, Morris, Whorf, Chomsky, and Skinner. His writing is remarkably free of specialized terminology or jargon. The range and the freshness of his writing help to account, no doubt, for his continuing appeal and relevance.

See also Analysis, Philosophical; Broad, Charlie Dunbar; Carnap, Rudolf; Chomsky, Noam; Decision Theory; Dewey, John; Enlightenment; Induction; Logical Positivism; Metaphor; Moore, George Edward; Paradigm-Case Argument; Philosophy of Language; Ramsey, Frank Plumpton; Russell, Bertrand Arthur William; Skinner, B. F.; Stebbing, Lizzie Susan; Tarski, Alfred; Wittgenstein, Ludwig Josef Johann.

Bibliography

In addition to the works mentioned, Black's publications include eight volumes of essays: *Language and Philosophy* (Ithaca, NY: Cornell University Press, 1949), *Problems of Analysis* (Ithaca, NY: Cornell University Press, 1954), *Models and Metaphors* (Ithaca, NY: Cornell University Press, 1962), *Margins of Precision* (Ithaca, NY: Cornell University Press, 1970), *The Labyrinth of Language* (New York: Praeger, 1968), *Caveats and Critiques* (Ithaca, NY, 1975), *The Prevalence of Humbug and Other Essays* (Ithaca, NY: Cornell University Press, 1983), and *Perplexities* (Ithaca, NY: Cornell University Press, 1990).

Volumes translated or edited include *R. Carnap*, *The Unity of Science* (London, 1934), *Philosophical Analysis* (Ithaca, NY, 1950), *Translations from the Philosophical Writings of Gottlob Frege* (with P. T. Geach; Oxford, 1960), *The Social Theories of Talcott Parsons* (Englewood Cliffs, NJ: Prentice-Hall, 1961), *The Importance of Language* (Englewood Cliffs, NJ: Prentice-Hall, 1962), *Philosophy in America* (Ithaca, NY: Cornell University Press, 1965), *The Morality of Scholarship* (Ithaca, NY: Cornell University Press, 1967), and *Problems of Choice and Decision* (Ithaca, NY: The Program, 1975).

A reasonably complete bibliography to1985 is appended to Black's autobiographical sketch in Volume 12 of *Philosophers on Their Own Work* (Bern/Frankfurt/New York: H. Lang, 1985), pp. 9–41.

Newton Garver (1967, 2005)

BLACK HOLES

A black hole (the term was coined by John Archibald Wheeler in 1967) is a closed surface through which gravity prevents light from propagating. Insofar as relativity prohibits anything from traveling faster than light, it follows that nothing can escape through the surface of a relativistic black hole. That said, in general relativity the notion of energy is problematic, and energy and hence mass can be extracted by classical and quantum processes. Classically the interior of a black hole contains a *singularity*: Along certain paths physical quantities

become ill-behaved (e.g., the gravitational field may become infinite). While nothing can pass back through the surface of a black hole, it is possible in certain models to travel into other universes. All of these properties have philosophically disturbing implications that have strongly influenced the development of physics, especially since there are solid theoretical and experimental reasons to believe that black holes are not merely hypothetical, but actually exist.

HISTORY

Black holes (henceforth BHs) arise in the general theory of relativity (GTR). However something similar is possible in Newtonian physics. John Michell (1784) and Pierre Simon Laplace (1796) pointed out that, as a ball thrown upwards with insufficient speed will eventually fall back to Earth, if a star of a given mass were smaller than a certain size (in modern parlance, its *critical radius*) then even light corpuscles, emitted from the surface at the speed of light, would eventually be pulled back to its surface. If such a star were a sufficient distance away it would not be directly visible (though faster or accelerating bodies could escape).

Karl Schwarzschild discovered the first exact solution of GTR in 1916, before the Einstein field equations of the theory were cast in their final form. The model has a point mass *M* at its center and in radial co-ordinates (two angles, θ and Φ, giving the latitude and longitude of a point, its distance *r* from the center and time *t*) the line element (the distance between two infinitesimally separated points) is:

$$ds^2 = -(1 - 2GM/c^2 r)dt + \frac{dr^2}{1 - 2GM/c^2 r} + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

where *G* is Newton's gravitational constant and *c* the speed of light. The idea in GTR is that *M* determines ds^2, which determines the geometry of spacetime, which explains the effect of gravity. Inspection of the second and first terms reveals ds^2 diverges at $r = 2GM/c^2$—the "Schwarzschild radius"—and $r = 0$—the location of the point mass, respectively. The singularity at $r = 0$ is genuine, though one would suspect (wrongly) that it would not occur if it were not for the idealization of a point mass—see below. The divergence at $r = 2GM/c^2$ is not physical, but merely an artifact of the co-ordinates used to describe the solution, a point that was not properly appreciated until the late 1930s. By way of analogy, if we used $x'=1/(x-1)$, $y'=y$ as co-ordinates for the x-y plane, then $ds^2=d\,x'^2\,x'^4 + dy'^2$. Along $x=1$ the plane is perfectly smooth but ds^2 is singular since $x'=\infty$, a reflection of the "poor" choice of co-ordinates.

In the Schwarzschild solution the singularity reflects not a geometric irregularity but the existence of a sphere of radius $2GM/c^2$ (named the "horizon" by Wolfgang Rindler in the 1950s) from which no light can escape (a point first made by Johannes Droste in 1916). Clearly if a body is smaller than $2GM/c^2$ then light cannot escape its horizon, so a star's Schwarzschild radius is its critical radius: The horizon forms a BH around any star smaller than its Schwartzschild radius. Other solutions for BHs were discovered by Hans Reissner (1916) and Gunnar Nordström (1918) for a charged BH, by Roy Kerr (1964) for a spinning BH, and by Ted Newman (1965) for a charged and spinning BH. It is important to emphasize that the nature of the horizon and hence of the BH (and hence of the early solutions) was not properly understood until the mid-1960s. Remarkably a so-called "No Hair Theorem" shows that the exterior (but not the interior) of any BH is completely characterized by its mass, charge, and spin and not on any other details of its composition or formation: The exterior of every possible BH is described by one of the four models mentioned here.

The Schwarzschild solution was quickly accepted as the description of gravity outside a (stationary) star, where the mass could be treated as located at the star's center—that is, providing that the star was larger than its Schwarzschild radius (18.5km for the Sun). The early pioneers of GTR did not properly understand the horizon (they worried about its possible singular nature and the fact that bodies approaching it would apparently take bodies an infinite time to reach it) and tried to argue that they could not occur in nature. However the question arose of what would happen to a star after it exhausted its fuel supply and began to cool and contract. By 1925 it was apparent that such stars could shrink under their own gravity to form *white dwarfs* 1,000s of times denser than the Sun, but in the early 1930s Subrahmanyan Chandrasekhar showed that white dwarfs of masses more than 1.5 the mass of the Sun ($M_\odot$) were not stable against gravity. In 1933 Walter Baade and Fritz Zwicky proposed that stars could further implode to form *neutron stars* as dense as atomic nuclei (the gravitational energy released by the implosion explaining supernovae), but in 1938 Robert Oppenheimer and George Volkoff argued that neutron stars heavier than a few times $M_\odot$ (the contemporary value is $2M_\odot$) would be unable to resist their own gravity, and in 1956 Wheeler and Masami Wakano demonstrated that there were no denser stable objects than neutron

stars. In 1939 Oppenheimer and Hartland Snyder provided the first model of a star collapsing through its Schwarzschild radius, a model vindicated in the early 1960s by simulations based on hydrogen bomb research (subsequent models show that stars below around $20M_{\odot}$ will eject sufficient matter during implosion to avoid complete collapse). In other words, it became clear not only that BHs were quite possible but also that their formation was likely. One philosophically significant aspect of this scientific revolution is how it was affected by physicists' changing intuitions about what mathematical models were physically realistic.

In the first years of the twenty-first century the astronomical evidence for BHs was strong and rapidly evolving; as of this writing, there are some fifty known candidates, half of them strong candidates. One class of candidates occurs in binary systems in which a star orbits a heavy body that strips material from it. The speed of the star can be calculated from the Doppler shift of its emission spectrum and the mass of the heavy body derived from that: If it is above $2M_{\odot}$ it is too heavy to be a neutron star and is presumably a BH. Typical BHs of this type have $5–15M_{\odot}$. A second kind of BH candidate is the *supermassive* BH, thought to occur at the center of galaxies. For instance it is believed that a BH over $3 \times 10^6 M_{\odot}$ lies at the heart of our galaxy in the constellation Sagittarius. Observational work has been done to verify that these candidates are not some unknown, denser objects, for example by looking for nuclear reactions that can occur only on material surfaces and not on horizons.

BH INTERIOR

When objections to the physical possibility of a horizon were overcome, the question became whether real BHs, like the known solutions, contained singularities. According to Oppenheimer and Snyder's model, stellar matter collapses to a point to form a singularity, but their work was not definitive because it assumed an unrealistically symmetric distribution of collapsing matter. However in the late 1960s Roger Penrose and then Steven Hawking proved Singularity Theorems showing that singularities must arise under very general conditions, including those believed to hold in BH formation, while in 1969 Vladimir Belinsky, Isaac Khalatnikov, and Evgeny Lifshitz found a singularity that would form if stellar collapse was only approximately symmetric.

Ordinarily a singularity in a function means that it diverges somewhere in its domain. The situation is more complicated in GTR because space has no existence separately from the fields: GTR is the theory of the metric field (the coefficients of ds^2), which describes the geometry of space. Intuitively speaking, when the fields of GTR become singular the very notion of spatial points fails, and space can contain a singularity even though the fields do not diverge anywhere. Singularities potentially raise several philosophically interesting issues. One is that singularities are associated with failures of determinism: The problem is roughly analogous to that of calculating the propagation of an electromagnetic wave in a space with a hole removed. Another is that Physicists have thus postulated various forms of *Cosmic Censorship*: That singularities cause at most localized failures of determinism (e.g., only inside a BH).

One may be struck by the similar (mistaken) initial reaction to the horizon (also by the "transcendental" nature of this move—without censorship and determinism, physics of a certain kind is impossible), though censorship has been shown in a range of cases. If uncensored "naked singularities" are possible, then it would be possible to use them to complete "supertasks" in a finite time relative to a distant observer. Important to remember is that our discussion so far has been in the context of classical GTR (utilizing some quantum properties of matter), but around a singularity, quantum gravitational effects likely become important. Physicists generally expect that singularities will not occur in a quantized version of GTR. If so, classical singularities may offer no philosophically important lessons after all. However John Earman's *Bangs, Crunches, Whimpers, and Shrieks* argues that classical singularities may pose no physical problem, so physicists need not demand that quantum gravity banish them.

Everything that enters the Schwarzschild BH eventually reaches its singularity, but in the other models it is possible to avoid the singularity altogether and travel on to a flat region of spacetime: The other BHs contain "worm holes" or "Einstein-Rosen bridges" to a "new" universe or to a region of space far from the BH (in the latter case spacetime would contain "closed timelike curves," paths that allow one to travel to one's past). However even in classical GTR the models assume unrealistic symmetries, and so the interior parts cannot be trusted: As of this writing, while it has not been shown that more realistic classical BHs do not contain worm holes, it is widely assumed that they do not. The situation in quantum gravity is even less clear, though Lee Smolin (1992) has speculated that a new universe is created inside whenever a BH forms, with laws of nature that vary in a small, random way from their parent universe, so that all possible physics eventually comes into existence.

BLACKNESS

Consider what would happen as an astronaut entered the horizon of a BH while he was watched from the outside. It is useful to think of the region around the BH as analogous to a deep, sloped, hole in the ground: As one gets nearer the center, the distance up the walls and along the ground to the outside grows rapidly. Analogously, as the BH is approached, light has to travel up ever steeper "walls" of curved spacetime to escape: Since the speed of light is constant, it follows that light takes an increasingly long time to reach the outside. Just as showing movie frames at increasing intervals makes a scene appear to slow, so the astronaut will appear to decelerate. In a BH, however, the time for light to reach the outside becomes infinite at the horizon (though space there is perfectly smooth). The effect is that the astronaut appears to decelerate indefinitely and from the outside can never be seen to enter the BH, as if the movie were slowed until frozen on a single frame. It is crucial to appreciate that this phenomenon is entirely optical: The astronaut himself measures only a finite amount of time until he is inside the BH. (That a collapsed massive star would thus be seen frozen at its horizon was an impediment to understanding BH formation.)

Just as it would take light infinitely long to escape from the horizon of a BH, nothing localized inside the horizon can pass through it (ignoring subtleties concerning the speed of light as a cosmic speed limit): The BH is opaque to its exterior. However it is theoretically possible to extract energy and hence mass, via the relativistic equivalence of mass and energy, from a spinning BH by classical processes that slow the spin. It is rather surprising that this extraction is achieved by throwing matter into the hole, but not surprising that a BH stores energy in its rotation. In other words part of the BH's mass arises from its interaction with the spacetime outside it, so no energy has to leave the interior. More surprisingly, Hawking (1974) showed quantum effects mean that even a non-spinning BH would radiate energy (and mass) with a temperature inversely proportional to it mass.

This effect is philosophically important for two reasons. First it confirmed Jacob Bekenstein's (1972) speculation that BHs obeyed the laws of thermodynamics and hence possessed entropy; among other things BHs are relevant to an arrow time based on the Second Law of thermodynamics. Second what happens when the mass of the BH "evaporates" to zero? One issue is the possibility of a naked singularity. Another is the "loss of information paradox": Physical theories typically allow an earlier state to be retrodicted from a later one, so that no information about the earlier state is lost over time. However, in Hawking's calculation, radiation carries no information about how a BH was formed, so that information remains inside the BH and is lost: Once a BH evaporates, retrodiction is impossible. For these reasons BH radiation became an important issue since theories of "quantum gravity" can be judged according to how they treat BH entropy and the loss of information paradox. In particular, BH entropy is connected to the powerful idea of "holography," which connects the physics of any region to the physics of the surface bounding it.

See also Earman, John; Laplace, Pierre Simon de; Philosophy of Physics; Quantum Mechanics; Space; Time.

Bibliography

Bekenstein, Jacob D. "Blackholes and Information Theory." *Contemporary Physics* 45 (2003): 31–43. Available from http://www.arXiv.org/abs/quant-ph/0311049.

Belot, Gordon, John Earman, and Laura Ruetsche. "The Hawking Information Loss Paradox: The Anatomy of a Controversy." *British Journal for the Philosophy of Science* 50 (1999): 189–229.

Carroll, Sean. *Spacetime and Geometry: An Introduction to General Relativity*. New York: Pearson Addison Wesley, 2003.

Earman, John. *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. Oxford: Oxford University Press, 1995.

Israel, Werner. "Dark Stars: The Evolution of an Idea." In *Three Hundred Years of Gravitation*, edited by Steven S. Hawking and Werner Israel. Cambridge, U.K.: Cambridge University Press, 1987.

Smolin, Lee. *The Life of the Cosmos*. Oxford: Oxford University Press, 1997.

Taylor, Edwin F., and John Archibald Wheeler. *Exploring Black Holes: Introduction to General Relativity*. New York: Benjamin Cummins, 2000.

Thorne, Kip S. *Black Holes & Time Warps: Einstein's Outrageous Legacy*. New York: Norton, 1994.

Unruh, William. "Black Holes, Dumb Holes, and Entropy." In *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, edited by Craig Callender and Nick Huggett. Cambridge, U.K.: Cambridge University Press, 2001.

Ziólkowski, Janusz. "Black Hole Candidates." arXiv.org e-print Archive. Available from http://www.arXiv.org/abs/astro-ph/0307307.

Nick Huggett (2005)

BLAKE, WILLIAM

(1757–1827)

William Blake was an English poet, painter, and engraver. He was born in London, the second of five children in the

family of a retail hosier. His social status precluded university education, and he was apprenticed to an engraver. Apart from that training and a few months at the Royal Academy, Blake was self-educated. Most of his pictorial work took the form of illustrations for books, biblical subjects forming the largest group. His painting and engraving were thus primarily related to literature, and the interdependence of poetry and painting is a central principle of all his work. He lived in London nearly all his life, very frugally, sometimes in poverty, and constantly dependent on patrons. He met William Wordsworth, Samuel Taylor Coleridge, and Charles Lamb, and was admired by the last two; but he died practically unknown as a poet, although he had been writing poetry since the age of twelve. After one volume of juvenile verse (*Poetical Sketches,* 1783) was published through the efforts of friends, Blake determined to produce his poetry by engraving the text himself and accompanying it with illustrations. Practically all his later poetry, except what was left in manuscript, took the form of a text and designs etched on copper, stamped on paper, and then colored by hand. Most of his lyrics are in two collections: *Songs of Innocence* (first engraved in 1789) and *Songs of Experience* (1794). Others are longer poems, generally called prophecies, which are sequences of plates. The "prophecies" include *The Book of Thel* (1789), *The Marriage of Heaven and Hell* (1793), *America* (1793), *Europe* (1794), *Milton* (about 1808, in 50 plates) and *Jerusalem* (about 1818, in 100 plates).

THOUGHT

The prophecies are symbolic poems in which the characters are states or attitudes of human life. This means that these poems embody religious and philosophical concepts as well as poetic imagery. These concepts are mainly concerned with Blake's sense of the relevance and importance of the arts and of the creative faculty of man, and seem to have been derived mainly from a negative reaction to the British empirical tradition of thought. He tells us that he had read John Locke and Francis Bacon in his youth and had decided that they mocked inspiration and vision. Blake's attitude would be better understood if it were thought of as anti-Cartesian, although he is unlikely to have read René Descartes, and his attitude embodies many elements that would now be called existential.

IMAGINATION. According to Blake, man is a working or constructing imagination—the creative artist is normative man. In this context there is no difference between human essence and human existence, for the imagination is the human existence itself and is also essential human nature. Works of art are neither intellectual nor emotional, motivated neither by desire nor by reason, neither free nor compelled: all such antitheses become unities in them. Even more important, the imagination destroys the antithesis of subject and object. Man starts out as an isolated intelligence in an alien nature, but the imagination creates a world in its own image, the world of cities and gardens and human communities and domesticated animals.

INTERPRETATION OF THE BIBLE. For Blake, the Bible is a definitive parable of human existence, as it tells how man finds himself in an unsatisfactory world and tries to build a better one—one which eventually takes the form of a splendid golden city, the symbol of the imaginative and creative human community. God in Blake's work is the creative power in man (here Blake shows the influence of Emanuel Swedenborg, with his emphasis on the unity of divine and human natures in Jesus), and human power is divine because it is infinite and eternal. These two words do not mean endless in time and spaces; they mean the genuine experience of the central points of time and space, the now and the here. Many features of Blake's anti-Lockean position remind us of George Berkeley, especially his insistence that "mental things are alone real"; but this doctrine of God takes Blake far beyond the subjective idealism and nominalism of Berkeley.

In Blake's reading of the Bible, "the creation"—the alien and stupid nature that man now lives in—is part of "the fall" and is the world man struggles to transcend. The objective world is the anticreation, the enemy to be destroyed. Blake says that man has no body distinct from his soul. He does oppose mind and body, but as contrasting attitudes to nature, not as separate essential principles. The "corporeal understanding," or perverted human activity, contemplates nature as it is (as a vast, objective, subhuman body) and tries to overcome the alienation of the subject by identifying the subject with nature as it sees nature. Nature is controlled, apparently, by automatic laws like the law of gravitation and by a struggle to survive in which force and cunning are more important than love or intelligence. Perverted human life imitates nature by continually waging war and by maintaining a parasitic class. Perverted religion, or natural religion, as Blake calls it, invents harsh and tyrannical gods on the analogy of nature. Perverted thought exposes itself passively to impressions from the external world and then evolves abstract principles out of these impressions that attempt to formulate the general laws of nature. These are the operations known as sensation and reflection in Locke. The abstracting tendency is perverted because it is not a

genuine effort to understand nature, but is a step toward imitating the automatism of nature by imposing a conforming morality on human life. The principle of this conformity is the acceptance of injustice and exploitation as inescapable elements of existence. The end of this perverted process is hatred and contempt of life, as expressed in the deliberate efforts at self-annihilation that Blake saw as beginning with the Napoleonic wars in his own time.

PROPHETIC BOOKS. The action in Blake's prophecies is concerned with the conflict of these creative and perverted states in human life. The sense of conservatism, of accepting things as they are, is symbolized by Urizen, who is associated with old age and the sky. When conservatism deepens into hatred of life itself, Urizen is replaced by Satan. The force that struggles against Urizen is the revolutionary impulse in man, called Orc or Luvah, who is associated with youth and sexual desire. Orc cannot achieve a permanent deliverance from Urizen; that is possible only for the creative power itself, called Los. The central theme of the prophecies is the effort of humanity, called Albion, to achieve through Los the kind of civilization that is symbolized in the Bible as Jerusalem and thus to reach the integration of human and divine powers represented in Christianity by Jesus.

See also Bacon, Francis; Berkeley, George; Coleridge, Samuel Taylor; Descartes, René; Existentialism; Imagination; Locke, John; Romanticism; Swedenborg, Emanuel.

Bibliography

The most convenient edition of Blake's literary work is the one-volume *The Poetry and Prose of William Blake*, edited by Geoffrey Keynes (London: Random House, 1927).

Works on Blake include Bernard Blackstone, *English Blake* (Cambridge, U.K.: Cambridge University Press, 1949); Foster S. Damon, *William Blake: His Philosophy and Symbols* (Boston: Houghton Mifflin, 1924); David V. Erdman, *Blake: Prophet against Empire* (Princeton, NJ: Princeton University Press, 1954); Northrop Frye, *Fearful Symmetry: A Study of William Blake* (Princeton, NJ: Princeton University Press, 1947); and Mark Schorer, *William Blake: The Politics of Vision* (New York: Henry Holt, 1946).

Northrop Frye (1967)

BLANCHOT, MAURICE

(1907–2003)

Maurice Blanchot was first and foremost a literary theorist, and his work included a number of essay collections, among them *The Space of Literature* (1982), *The Book to Come* (2003), and *Friendship* (1997). He also wrote powerful but rather hermetic novels such as *Thomas the Obscure* (1973), *Death Sentence* (1978), *Aminadab* (2002), and *The Most High* (1996), in addition to aphoristic works such as *The Writing of the Disaster* (1986). Blanchot's work has profound implications for the practice of philosophy. His influence therefore stretches from practitioners of the New Novel to philosophers such as Jacques Derrida and Michel Foucault.

Blanchot's strategy, which is meant to reconceive literature and to carry out a thoroughgoing critique of the possibility of language, is in major part derived from a critique of the Hegelian notion of the sign, by way of Martin Heidegger. In Blanchot's version of Hegel—as seen in his essays "The Experience of Mallarmé" and "Literature and the Right to Death"—the word, by isolating things and representing them in their absence, "gives me the being, but it gives it to me deprived of being"; or, put another way, the word makes the world appear and disappear in a moment (Blanchot, 1995, p. 322). The givenness of the thing (or person) in and through the word is also its radical removal, its distance from simple subjectivity or objectivity, its mortality. That one's words represent a thing that—or person who is—absent means that it or she or he can be absent (can be removed, destroyed, can be dead). The word thus represents things or persons in the act of constituting them and indicating their disappearance, their death. This is a negation that has nothing to do with the patient work of the dialectic; it is of the instant, an instant that cannot be recaptured in any constructive movement. Named things—people—are always already dead, and their life is an infinitely repetitious death. "Pure language," as Blanchot calls it, entails a nomination where this "neutral" action of language is recognized; this (impossible) recognition in turn characterizes true literature (Blanchot, 1995). Put another way, the world of work recognizes and uses the negating power of the word; true literature, however, recognizes this negation as so thorough that it penetrates and radically negates beings and things—including, of course, literature itself—in the moment of their constitution. Blanchot in fact compares literature to the Terror, where beings are called forth in a repetitious movement that both constitutes them as revolutionary subjects and kills them. Thus Blanchot can write, as he did in "The Experience of Mallarmé," that "the fulfillment [*accomplissement*] of language coincides with its disappearance" (1982). This fulfillment is literature.

This "neuter" (*le neutre*) of Blanchot bears an obvious connection with the Heideggerian *Dasein* or the Levinasian *Il y a*. But there are crucial differences. The work of death is so thorough for Blanchot that it is hard to see how any notion of authenticity or foundational ethics could be carried out through it. This incessant, unproductive constituting and destroying death erodes all philosophical systems, all coherent models of subjectivity (Descartes), all constructive movements of negation in time, and all doctrines based on force and power. As Gerald Bruns puts it, "The 'nocturnal' experience of words, in which the cognitive or speaking subjectivity is deprived of its sovereignty and its power, reduced to the passivity of its fascination, is one of the most important events in Blanchot's thinking" (1997, p. 77). One could add that the sovereignty of space and time are emptied out as well, because the subject moving in them and making the world is not making a coherent entity but rather is caught up in the emptying out of the possibility of all relation, all mediation between (dead) self and (dead) world.

Blanchot narrativizes this nocturnal relation in myths and fictions. "The Gaze of Orpheus" provides an excellent example: Orpheus would bring Eurydice from death to a realm of the disclosure of truth and beauty. But halfway through the journey he must see her as she is, as death, as radical concealment. This demand, to see and to speak "the most certain masterpiece," is inevitably fatal to Orpheus's aesthetic and philosophical hopes. Seeing constitutes, but it is also inseparable from, radical loss, "the movement of writing" (1982, pp. 103–104). Truth and beauty can be grasped, but only in the night in which such certainties are immediately and incessantly lost. The strangeness of the Heideggerian "thing" is rewritten as an impossible interpersonal relation, between man and woman, that is also an allegory of the necessity and impossibility of language at its limit (literature).

Many of Blanchot's fictions work out this interpersonal relation between man and woman, or self and other, in the mode of the radical death of both the subject and signification. The Blanchotian hero is thus a figure obsessed with a negativity that pervades all things, a personage for whom the only relation is a repetitive recognition of the impossibility of the recognition of a stable relation, as in *Death Sentence* and *The Most High*.

Blanchot's conception of language is literary in that its radicality is seen to characterize literature at the highest level. Blanchot, however, does not limit its implications exclusively to what is conventionally conceived as the literary realm. Language as understood by Blanchot invests philosophy, making its movement possible and at the same time undermining every one of its certainties. In this way Blanchot's version of literary language and of language in general clearly anticipates Jacques Derrida's analyses of writing. But language also conditions Blanchot's version of the community.

The problem of the enthusiastic community, so central to French social thought since (at least) the Revolution, is rewritten by Blanchot in *The Unavowable Community* as the fundamental relation of those who have "nothing in common," an inescapable and unmediated relation of reading in which nothing is knowable, nothing endures, a moment that constitutes nothing in a coherent movement of time. Similarly, in *The Writing of the Disaster*, the Holocaust is seen as a "disaster" in the Blanchotian sense. Not a "word, not the name of anything … but always an entire complex or simple sentence, where the infinitude of language … seeks … to fall outside language—without, however, ceasing to belong to it" (1986, p. 84). From writing on literature and literary language, then, Blanchot has moved to a larger conception of the word, and language. Essential language leads nowhere, guarantees nothing, and only has "its end in itself" ("The Experience of Mallarmé" 1982), yet its negativity is fundamental to an understanding of society and its limit term, the moment in which it grasps itself as the radically ungraspable: the Holocaust. "Literature," in Blanchot's sense, therefore resists any easy containment.

See also Death; Derrida, Jacques; Descartes, René; Foucault, Michel; Heidegger, Martin; Literature, Philosophy of; Philosophy of Language.

Bibliography

WORKS BY BLANCHOT

Essay Collections

The Space of Literature. Translated by Ann Smock. Lincoln: University of Nebraska Press, 1982. Includes the essays "The Gaze of Orpheus" and "The Experience of Mallarmé."

The Writing of the Disaster. Translated by Ann Smock. Lincoln: University of Nebraska Press, 1986.

The Unavowable Community. Translated by Pierre Joris. Barrytown, NY: Station Hill, 1988.

The Step Not Beyond. Translated by Lynette Nelson. Albany: State University of New York Press, 1992.

The Work of Fire. Translated by Charlotte Mandell. Stanford, CA: Stanford University Press, 1995. Includes the essay "Literature and the Right to Death."

Friendship. Translated by Elizabeth Rottenberg. Stanford, CA: Stanford University Press, 1997.

The Instant of My Death. With Jacques Derrida as *Demeure: Fiction and Testimony*. Translated by Elizabeth Rottenberg. Stanford, CA: Stanford University Press, 2000.

Faux pas. Translated by Charlotte Mandell. Stanford, CA: Stanford University Press, 2001.

The Book to Come. Translated by Charlotte Mandell. Stanford, CA: Stanford University Press, 2003.

Novels

Thomas the Obscure. Translated by Robert Lamberton. New York: D. Lewis, 1973.

Death Sentence. Translated by Lydia Davis. Barrytown, NY: Station Hill, 1978.

The Most High. Translated by Allan Stoekl. Lincoln: University of Nebraska Press, 1996.

Aminadab. Translated by Jeff Fort. Lincoln: University of Nebraska Press, 2002.

WORKS ABOUT BLANCHOT

Bruns, Gerald L. *Maurice Blanchot: The Refusal of Philosophy*. Baltimore: Johns Hopkins University Press, 1997.

Haase, Ullrich, and William Large. *Maurice Blanchot*. New York: Routledge, 2001.

Herr, Deborah. *Politics and Literature: The Case of Maurice Blanchot*. New York: Peter Lang, 1999.

Allan Stoekl (2005)

BLANSHARD, BRAND

(1892–1987)

Brand Blanshard was an American philosopher whose task is best described in his own words as the "vindication of reason against recent philosophical attacks." Blanshard was thus a critic—a critic of all those who, he alleged, reject rationality—but at the same time he tried to exhibit the credentials that reason can show in its own right.

Blanshard was educated at the University of Michigan, Columbia, Oxford, and Harvard—where he received his PhD. He taught at the University of Michigan, at Swarthmore College, and at Yale—where he was Sterling professor of philosophy and chairman of the department. The multitude of honors he received during his career precludes their enumeration here.

Blanshard's first major work was *The Nature of Thought* (London, 1939), in two volumes, each divided into two books. The first volume is largely concerned with a subject matter common to both philosophy and psychology. The stated goal is to discover a theory of perception (Book I) and a theory of ideas (Book II) that will simultaneously satisfy the psychologist, who views percepts and ideas as contents of the mind, and the philosopher, who views them as potential items of knowledge. Various theories are examined and rejected—most notably the traditional empiricist approach—and it is finally argued that only a theory along the lines developed by Francis Herbert Bradley, Bernard Bosanquet, and Josiah Royce is able to meet this double demand. The universal, Blanshard maintained, is present in all thought, even in the most rudimentary forms of perception; and it is the presence of the universal that is the most important feature of thought. This conclusion exhibits a theme that recurs throughout both volumes: the use of doctrines drawn from the idealist tradition in dealing with contemporary problems.

In the second volume of *The Nature of Thought,* the subject matter becomes more specifically philosophical. The main task of Book III (titled "The Movement of Reflection") is to answer the epistemological problem: What is the test and the nature of truth? Once more, after examining and rejecting alternatives, Blanshard turns to the idealist tradition for his answer, adopting a version of the coherence theory of truth. His exposition of the coherence theory has a number of distinctive features. Foremost is the clarity, rigor, and persuasiveness of the presentation; in this respect Blanshard has only Royce as a rival. Furthermore, he develops the theory independently of metaphysical doctrines that are for the most part now repudiated. Finally, he develops the theory in full cognizance of contemporary criticisms and attempts to offer direct answer to them.

In Book IV (titled "The Goal of Thought") Blanshard moves from epistemology into metaphysics. Still operating within the framework of idealism, he accepts the connected notions of internal relations, concrete universality, and concrete necessity. But he does not, as do most idealists, give these doctrines a gratuitous theological turn, nor does he attempt to secure the foundation of the entire system through an a priori proof that the completed, fully articulated system must exist. He does introduce the conception of a transcendent end for thought, which he considers a necessary postulate for knowledge, but he admits that it is possible (though unlikely) that this postulate is mistaken.

Some two decades after the publication of *The Nature of Thought,* and upon retirement from Yale, Blanshard began a projected three-volume sequence that would bring together material originally presented in his Carus and Gifford lectures. *Reason and Analysis* (London, 1962), the second of the three volumes, is his most polemical work. It is in large measure a systematic and unremitting attack upon the analytic tradition as it emerged in various forms during the twentieth century. Some of the arguments presented are refinements of those used in *The Nature of Thought,* but *Reason and Analysis* is not a mere echo of the earlier work. On the constructive side, many of the earlier idealistic doctrines,

although not silenced, seem decidedly muted. If philosophies are to bear labels, this later position might better be called rationalism than idealism.

The first work in the sequence, *Reason and Goodness* (London, 1961), introduces another aspect of Blanshard's thought. In this work he traces out the dialectical interplay between the demands of reason and the demands of feeling throughout the history of ethical theory. Not surprisingly, Blanshard rejects any theory that will not provide a place for reason in the account of human values, and he thus offers elaborate critiques of subjectivism, emotivism, and related theories.

In developing his own ethical position Blanshard does not turn, at least primarily, to the idealist tradition but rather to the works of Henry Sidgwick, G. E. Moore, H. A. Prichard, and W. D. Ross. Throughout his career Blanshard favored teleology in ethics, and for a time he was attracted by Moore's ideal utilitarianism. He came to reject this position largely because of the difficulties associated with Moore's conception of nonnatural properties. In *Reason and Goodness* Blanshard rejects Moore's critique of naturalism and argues that goodness is characterized by the joint properties of satisfaction and fulfillment. The idea of fulfillment is associated with the idealist tradition, but as Blanshard uses it, it carries no suggestion of loss of individuality and is thus quite different from the idea of fulfillment as employed by Bradley and most other idealists. By including both satisfaction and fulfillment in the definition of goodness, Blanshard hopes to provide for feeling on one hand and reason on the other and, in this way, to resolve the dialectical tension outlined earlier in the work.

Reason and Belief was not yet published as of this writing, but from Blanshard's lectures it may be assumed that in this work he will challenge the religious irrationalism that is currently fashionable in some quarters. What positive doctrines he will espouse is more a matter of speculation.

See also Bosanquet, Bernard; Bradley, Francis Herbert; Coherence Theory of Truth; Idealism; Relations, Internal and External; Moore, George Edward; Rationalism; Reason; Ross, William David; Royce, Josiah; Teleology.

Bibliography

Blanshard, Brand. *Reason and Belief.* New Haven, CT: Yale University Press, 1974.

Additional works by Blanshard are "Current Strictures on Reason," in *Philosophical Review* 54 (1945): 345–368, and *On Philosophical Style* (Bloomington: Indiana University Press, 1954).

A work on Blanshard is Ernest Nagel, "Sovereign Reason," in his *Sovereign Reason* (Glencoe, IL: Free Press, 1954), pp. 266–295.

Robert J. Fogelin (1967)

BLOCH, ERNST
(1885–1977)

Ernst Bloch, the German Marxist philosopher, was born at Ludwigshafen. Influenced by late German expressionism and by the atmosphere of Munich after World War I, Bloch's style and thought reveal contradictory and uncertain trends. He began his career at the University of Leipzig by publishing *Von Geist der Utopie* in 1918. This work was followed in 1922 by a study of Thomas Münzer in which mystical and eschatological ideas blend with dialectic elements of Marxist-Hegelian origin. *Spuren* followed in 1930 and *Erbschaft dieser Zeit* in 1933. In the latter work the various elements of Bloch's thoughts are for the first time clearly placed within a Marxist framework showing revisionist tendencies.

In 1933 Bloch left Germany, eventually reaching the United States, where he created his major work, *Das Prinzip Hoffnung*, a huge work that has been called "a monstrous essence of his thoughts." After World War II Bloch, like Bertolt Brecht, went to East Germany, where from 1948 until his retirement in 1957 he was professor at the University of Leipzig. At first, Bloch's political and intellectual influence in East Germany was limited, but nevertheless, he was never fully appreciated by party authorities. His winning the *Nationalpreis* of the German Democratic Republic in 1955 stirred controversy, and Bloch's views had changed considerably during his sojourn there. His ideas, which were carefully watched by party authorities, became the center of many discussions. In 1953, after the publication of *Subjekt-Objekt, Erläuterung zur Hegel* and *Avicenna und die Aristotelische Linke*, Bloch became editor of the *Deutsche Zeitschrift für Philosophie*. But the journal's comparative independence led to a series of arrests and trials of its collaborators and editors. Wolfgang Harich, Günther Zehm, and Manfred Hertwig were sentenced to prison, and Richard Lorenz and Gerhard Zwerenz fled to the Federal Republic. Although Bloch was only slightly involved, he was forbidden to publish, and in 1957 his works were officially condemned. When Bloch tardily made a declaration of loyalty, it was vague and noncommittal.

Although he was finally permitted to publish the third volume of his *Das Prinzip Hoffnung* in 1959, Bloch asked for political asylum during a visit to the Federal Republic in 1961, where he became a visiting professor at the University of Tübingen. Bloch is generally known in the West as a major Marxist philosopher, but he drew on a far wider heritage that includes classical German thought, Christian and Jewish mysticism, Neoplatonism, and even the esoteric speculation of the *Zohar*. His major work, *Das Prinzip Hoffnung*, gives the impression that Bloch, although claiming that the economic element is fundamental, relegates it to a secondary level and focuses his attention on what Marxist theory regards as only a superstructure, the problem of intellectual culture.

According to Bloch, all reality is "mediation," or the subject-object relation, a dynamic relation that tends ultimately toward the final goal (*Endziel*) of the reunion of subject and object. The *Urgrund*, the primordial stuff prior to the distinction between subject and object, matter and spirit, is moved by an obscure immediate cosmic impulse, which Bloch terms "hunger" and contrasts with Sigmund Freud's libido. After subject and object have been distinguished, Bloch claims, this hunger remains essential to both subject and object. Thus the reality of both subject and object is in the future, and the category of possibility comes to play a central role in his thought.

SUBJECT

In humans, the primordial hunger becomes desire, or hope. Hope presents itself as utopia, as a vision of a possibility that might be realized. Hope is tension toward the future, toward the new. It moves from a mere state of mind (*Stimmung*) to a representation, and then to knowledge. Although hope is founded on the will, in order to be hope that understands (*begriffene Hoffnung, docta spes*), it must draw strength from something real that will survive even when hope itself is completely satisfied. This residue makes hope something more than a project of reason and puts it in relation to what is objectively possible. The future possibility is not just a dream, even if it is heralded in dreams.

POSSIBILITY

The relations between subject, object, reality, and possibility are complex. The nature of the real is a tendency toward, or anticipation of, the future, and thus its reality is the reality of something in the future. But the future is already real as objective possibility. Bloch distinguishes between objective possibility, which (because the object as object is not real) is merely theoretical, and real possibility, which is practically connected with the future. What is really possible is concretely connected with utopia. Reality always contains elements of possible change, possibilities not yet actually existing. Utopias are concerned with these possibilities and thus have an essential function in human consciousness. On the other hand, these possibilities must have a foundation in the object because thought can represent in imagination infinitely many possible objects in infinitely many relationships.

If an event were completely conditioned, it would be "unconditionally certain." Therefore, what can possibly come into existence is possible only insofar as it is not conditioned. What is objectively possible, therefore, is so only insofar as it is not constrained by predetermined conditions. Bloch distinguishes between two senses of objective possibility. One sense concerns the thing and is the thing's "behavior," or the appearance of the thing as an object of knowledge. The other sense concerns our knowledge of the thing. The objectivity (*Sachlichkeit*) of the thing concerns only our knowledge of it, while its factuality (*Sachhaftigkeit*) concerns only the object of knowledge.

MATTER

The distinction between objectivity and factuality leads Bloch to claim that Marxism is only a partial outlook on reality and needs completion, even though the reconciliation of the real and the possible is achieved in historical materialism, which retains, in its complete immanentism, an element akin to the doctrine of salvation of the great religions. According to Marxism, historical changes arise out of precise historical socioeconomic conditions, and physical movement arises out of contradiction, the clash of opposites. But just as Bloch supplements the claims of historical materialism with his concept of hope, so he supplements the claims of dialectical materialism. In the object, or matter, the primordial hunger becomes a motive force (*agens*). But even though Bloch affirms that this force is completely immanent in matter, it is doubtful whether his view is still materialistic. His hostility toward all forms of mechanism and his inclination toward organic solutions weaken the materialistic features of Marxism to the point of nonexistence. The innate drive that he ascribes to matter has meaning only from the point of view of the final goal. Matter is not predetermined, since it has the capacity not only to express itself in existence but also to do so in forms that are always new. Nevertheless, the teleological doctrine of a final goal for the entire world process is not an extension of a psychological category or historical principle to nature. Rather,

it is the cosmic unity of the subject process and object process when being finally becomes thinking and thinking finally becomes being. The historical process of society is thus related to the world process and ultimately to matter.

Thus Bloch identifies dialectical matter with real possibility, but its being in process is not material and contradicts the fundamental Marxist tenet that matter is an independent reality that cannot enter into a relation with anything. Several critics have remarked that Bloch's conception of matter has its sources in the romantic *Naturphilosophie* of G. W. F. Hegel and Friedrich von Schelling; on this view, Bloch belongs among the idealist critics of natural science.

UTOPIA

The reconciliation of subject and object comes through utopia. In utopia, romantic *Sehnsucht*—the nostalgic regret that our dream of rationally conquering the world is blocked by a limit that we try unceasingly but perhaps vainly to overcome—is united with messianic expectancy. Utopia foresees the "kingdom of the children of God" of Thomas Münzer, the kingdom of freedom in which the exploitation of man by man ceases. At this time will come that unification, the identification of subject and object, which Bloch claims Karl Marx foresaw when he spoke of the future historicization of nature and naturalization of man. It is thus from man that the world expects its realization, and the realization of the world process coincides with the self-realization of humankind. The Marxist epistemological theory of reflection will no longer be needed, since knowledge itself will be overcome by hope and the object as object will disappear; it will no longer be the having-become (*Gewordenes*) but rather pure process, the becoming (*Werdendes*), the not-yet (*noch nicht*).

Block's thought is very far from Marx's historical outlook and perhaps not too far from the early views of Georg Lukács. In his conflict with the schematicism and dogmatism of orthodox Marxism, Bloch belongs with such idealist and existentialist revisionists as Lukács, Antonio Gramsci, and Jean-Paul Sartre. Bloch's attempt to save Marxist theory from ossifying has wider implications than their attempts, however, for it is related to the problem of how Marxism is to make use of a cultural heritage, especially the heritage of classical German philosophy and, at least for Bloch, the heritage of the great religions of salvation. Bloch's solution has been to develop one vast comprehensive vision of reality, combining the original intuitions of the Old Testament and apocalyptic literature with the dynamic and messianic elements in Marxism. Bloch's very language reveals this mixture of ancient and modern. Difficult and intense, it echoes both recent expressionism and the language of the Bible and of mystical literature. The past is for Bloch not something fixed in an unreachable dimension, its cultural wealth to be discarded in order to start anew, but a dynamic field of research still of use to man.

See also Marxist Philosophy.

Bibliography

WORKS BY BLOCH

Vom Geist der Utopie. Munich: Duncker and Humblot, 1918.

Thomas Münzer als Theologe der Revolution. Munich: Wolff, 1921.

Spuren. Berlin: Cassirer, 1930.

Erbschaft dieser Zeit. Zürich: Oprecht and Helbling, 1935.

Freiheit und Ordnung, Abriss der Sozial-Utopien. New York: Aurora, 1946.

Subjekt-Objekt: Erläuterung zur Hegel. Berlin: Aufbau, 1951.

Avicenna und die Aristotelische Linke. Berlin: Rütten and Loening, 1951.

Das Prinzip Hoffnung. 3 vols. Berlin: Aufbau, 1954–1959.

Gesamtausgabe, 17 vols. Frankfurt: Suhrkamp, 1959–1978.

Naturrecht und menschliche Würde. Frankfurt: Suhrkamp, 1960. Vol. VI of *Gesamtausgabe*.

Philosophische Grundfragen, zur Ontologie des Noch-Nicht-Seins. Frankfurt: Suhrkamp, 1961. Vol. I of *Gesamtausgabe*.

Fromm, Erich, ed. *Socialist Humanism: An International Symposium*. New York: Doubleday, 1965. Contains an essay by Bloch.

Ernst Bloch zu ehren; Beiträge zu seinem Werk. Edited by Siegfried Unseld. Frankfurt am Main: Suhrkamp Verlag, 1965.

Bildung und Konfessionalität;. Frankfurt am Main: M. Diesterweg, 1967.

Religion im Erbe. Eine Auswahl aus seinen religionsphilosophischen Schriften. Edited by Jürgen Moltmann. München u. Hamburg, Siebenstern-Taschenbuch-Verlag, 1967.

Widerstand und Friede; Aufsätze zur Politik. Frankfurt am Main: Suhrkamp, 1968.

Uber Karl Marx. Frankfurt am Main: Suhrkamp, 1968.

Atheismus in Christentum; zur Religion des Exodus und des Reichs. Frankfurt am Main: Suhrkamp, 1968.

"Man as Possibility." In *Hope: A Symposium*. West Nyack, NY: Cross Currents, 1968.

Christian Thomasius, ein deutscher Gelehrter ohne Misere. Frankfurt am Main: Suhrkamp, 1968.

Philosophische Aufsätze zur objektiven Phantasie. Frankfurt am Main: Suhrkamp, 1969.

Die Kunst, Schiller zu sprechen: Und andere literarische Aufsätze. Frankfurt am Main: Suhrkamp, 1969.

Karl Marx und die Menschlichkeit. Utopische Phantasie u. Weltveränderung. Reinbek b. Hamburg: Rowohlt, 1969.

Man on His Own; Essays in the Philosophy of Religion. New York: Herder and Herder, 1970.

Uber Methode und System bei Hegel. Frankfurt am Main: Suhrkamp, 1970.
A Philosophy of the Future. New York: Herder and Herder, 1970.
Bloch, Ernst. Frankfurt am Main: Suhrkamp, 1970.
Politische Messungen, Pestzeit, Vormärz. Frankfurt am Main: Suhrkamp, 1970.
Marx und die Revolution. Frankfurt am Main: Suhrkamp, 1970.
Das Materialismusproblem, seine Geschichte und Substanz. Frankfurt am Main: Suhrkamp, 1972.
Vorlesungen zur Philosophie der Renaissance. Frankfurt am Main: Suhrkamp -Taschenbuch-Verl., 1972.
Erbschaft dieser Zeit. Frankfurt am Main: Suhrkamp Verlag, 1973.
Ästhetik des Vor-Scheins. Frankfurt am Main: Suhrkamp, 1974.
Zur Philosophie der Musik. Frankfurt am Main: Suhrkamp, 1974.
Experimentum mundi: Frage, Kategorien d. Herausbringens, Praxis. Frankfurt am Main: Suhrkamp, 1975.
Ernst Bloch zum 90. Geburtstag: Es muss nicht immer Marmor sein: Erbschaft aus Ungleichzeitigkeit. Berlin: K. Wagenbach, 1975.
Zwischenwelten in der Philosophiegeschichte: Aus Leipziger Vorlesungen. Frankfurt am Main: Suhrkamp, 1976.
Aesthetics and Politics. London: NLB, 1977.
Literarische Aufsätze. Frankfurt am Main: Suhrkamp, 1977.
Experimentum mundi: Frage, Kategorien des Herausbringens, Praxis. Frankfurt am Main: Suhrkamp, 1977.
Tendenz, Latenz, Utopie. Frankfurt am Main: Suhrkamp, 1978.
Aesthetics and Politics. London: Verso, 1980.
"Denken heisst überschreiten": In Memoriam Ernst Bloch. Frankfurt: Ullstein, 1982, 1978.
Essays on the Philosophy of Music. Translated by Peter Palmer. Cambridge; New York; Cambridge University Press, 1985.
Werkausgabe. Frankfurt am Main: Suhrkamp, 1985, (c)1969.
Natural Law and Human Dignity. Cambridge, MA: MIT Press, 1986.
The Utopian Function of Art and Literature: Selected Essays. Cambridge, MA: MIT Press, 1988.
Heritage of Our Times. Cambridge, U.K.: Polity Press, 1991.

WORKS ON BLOCH

Baumgart, J. "E. Bloch, Erbschaft dieser Zeit." *Neue Rundschau* 2 (1963).
Buhr, M. "Der religiöse Ursprung und Charakter der Hoffnungsphilosophie." *Deutsche Zeitschrift für Philosophie* 6 (1958): 576–598.
Bütow, H. G. *Philosophie und Gesellschaft im Denken Ernst Blochs*. Wiesbaden, 1963.
Eucken-Erdsiek, E. "Prinzip ohne Hoffnung. Kritische Betrachtungen zum Hauptwerk von E. Bloch." *Philosophische Jahrbuch* 70 (1962): 147–156.
Holz, H. H. "Der Philosoph E. Bloch und sein Werk 'Das Prinzip Hoffnung.'" *Sinn und Form* 3 (1955).
Kurella, A. "Zur Theorie der Moral. Eine alte Polemik mit E. Bloch." *Deutsche Zeitschrift für Philosophie* 6 (1958): 599–621.
Ley, H. "Ernst Bloch und das Hegelsche System." *Einheit* 3 (1957).
Lorenz, K. "Hoffnung als Wissenschaft. Die Philosophie Ernst Blochs." *Deutsche Universitätszeitung* 22 (1957): 9–11.
Rühle, Jürgen, "Philosopher of Hope: Ernst Bloch." In *Revisionism*, edited by Leopold Labedz, 166–178. New York: Praeger, 1962.
Strolz, W. "Der Marxist und die Hoffnung. Einige Überlegungen zu dem Werke Ernst Blochs." *Wort und Wahrheit* (1960).
Tjaden, K. M. "Zur Naturrechts Interpretation Ernst Blochs." *Archiv für Rechts- und Sozialphilosophie* 56 (1962): 573–584.
Zehm, G. A. "Ernst Bloch." *Der Monat* 158 (1961).
See also *Ernst Bloch zum 70. Geburtstage; Festschrift* (Berlin: Deutscher Verlag der Wissenschaften, 1955); and *Ernst Blochs Revision des Marxismus* (Berlin, 1957), an anthology with an introduction by J. H. Horn.

Franco Lombardi (1967)
Bibliography updated by Michael J. Farmer (2005)

BLONDEL, MAURICE
(1861–1949)

Maurice Blondel is considered one of the foremost French Catholic philosophers of the twentieth century. Blondel was born at Dijon. He studied at the local *lycée,* and in 1881 entered the École Normale Supérieure, where he was taught by Léon Ollé-Laprune. Because of pragmatic tendencies in his thought, Blondel's name was associated for a time with the modernist movement. He was, however, essentially orthodox, and his work has been increasingly influential among those Catholic thinkers who look for an alternative to Thomism.

Through Ollé-Laprune, Blondel was influenced by John Henry Newman's theory that belief is a matter of will as well as of logical demonstration. Blondel was far from being a thoroughgoing pragmatist or vitalist and showed none of the naturalism of thinkers like Henri Bergson and James, yet he held that truth is to be reached not only through the intellect but through the whole range of experience, and to this extent he departed from the emphasis on rational demonstration found in traditional Catholic philosophy. Most of Blondel's teaching was done at the University of Aix-en-Provence, where he taught from 1896 until his death.

THOUGHT

An extended statement of Blondel's philosophy is found in the book *L'Action,* first published in 1893 and revised near the end of his life. This book should not be confused with another of the same title, published in 1937.

The claim of Blondel's early work is that philosophy must take its impetus from action rather than from pure

thought. The expression "action" is used in a wide sense to refer to the whole of our life, thinking, feeling, willing. Blondel tells us that it is to the whole man in his concreteness that philosophy must look in its quest for truth. One must turn from abstract thought to actual experience in all its fullness and richness. It is indeed this experience itself that motivates the philosophical quest, for man by his nature must act, and then he cannot help questioning the meaning of his action. Blondel anticipated the ideas later developed in existentialism when he pointed out that although we have not chosen to live and know neither whence we come nor even who we are, we are continually taking action and engaging ourselves in chosen policies.

Blondel rejected any nihilistic attempt to set aside the question of the meaning of action, and he had an ingenious argument to show that we cannot be content to say that action has no meaning. He claimed that to affirm nothing is really to affirm being. The very idea of nothing can be formed only by conceiving something positive and then denying it. There is something positive and affirmative underlying the denials of the nihilist, and even from his pessimistic view of life he derives a certain satisfaction. Blondel argued that the nihilist's nothing is his all. The very extent of what he denies reveals the greatness of what he wishes, for he cannot prevent affirmative ideas and aspirations from asserting themselves in the midst of his denials. Therefore, Blondel claimed, the problem of action and of its meaning must have a positive solution.

This solution is to be sought by means of a kind of phenomenology of action, though a phenomenology that is meant to show that we must pass beyond the phenomena to the discovery of the "supraphenomenal." We are impelled to this solution by reason of an immanent dialectic in action itself, made clear by a phenomenological description.

The basis of the dialectic is the gap between action and its realization. Man cannot in his action equal what he himself demands, and so there is in life a permanent dissatisfaction set up by the contrast between action and the realization at which it aims. This impels man to further action, and in the effort to close the gap, Blondel visualized the expansion of action in terms of an ever-wider outreach. Self-regarding action passes over into various forms of social action, and these in turn come to their limit in the highest type of moral action—that which aims at the good of all humankind.

But although this process partially overcomes the contrast between action and its realization, it never does so completely, and the gap reappears at each stage. There is no immanent solution to the problem of action. But we have seen already that an affirmative solution is demanded, and Blondel claimed that the demands of action itself point us from the immanent to the transcendent or supraphenomenal. The Catholic dimensions of Blondel's philosophy become fully apparent at this point, for it is essentially a philosophy of grace. God is immanent within man, in the sense that human action is already directed beyond the phenomenal order. To will all that we do will is already to have the action of God within us. Yet this quest for realization would be a frustrating one were it not that God in turn moves toward us in his transcendence, and human action is supported and supplemented by divine grace.

Since action is concrete, the beliefs that arise out of action and the experience of acting are not abstract formulations. It is in action that we apprehend God, but if we try to imprison him in a proposition or prove his existence by a logical demonstration, he escapes us.

In *La pensée* and subsequent writings, Blondel gave a more prominent place to thought and modified some of the anti-intellectualist tendencies that characterized his earlier period. At the same time, he reduced the differences that had separated him from traditional Catholic philosophy. But it must not be supposed that he departed in any essential respect from his philosophy of action. Thought and action were never rival principles for Blondel, but were at all times to be taken together. Action is no blind drive, but always includes thought; thought can attain its philosophical goals only as it remains closely associated with action. Thus, in his later phase, when he reconsidered the rational proofs of theism, he claimed that these proofs are possible only on the basis of a prior affirmation of God that has arisen out of our experience as active beings.

See also Action; Bergson, Henri; Dialectic; Existentialism; James, William; Modernism; Newman, John Henry; Nothing; Thomism.

Bibliography

WORKS BY BLONDEL

L'action: Essai d'une critique de la vie et d'une science de la pratique. Paris, 1893; rev. ed., 1950.

La pensée, 2 vols. Paris, 1934.

L'être et les êtres. Paris: Presses universitaires de France, 1935.

L'action, 2 vols. Paris: F. Alcan, 1937.

La philosophie et l'esprit chrétien, 2 vols. Paris: Presses universitaires de France, 1944–1946.

Exigences philosophique du christianisme. Paris: Presses universitaires de France, 1950.

WORKS ON BLONDEL
Dumery, H. *Blondel et la religion.* Paris, 1954.
Dumery, H. *La philosophie de l'action.* Paris: Aubier, 1948.
Lefèvre, F. *L'itinéraire philosophique de Maurice Blondel.* Paris, 1928.
Taymans d'Eypernon, F. *Le Blondélisme.* Louvain, 1935.
Tresmontant, Claude. *Introduction à la métaphysique de Maurice Blondel.* Paris: Éditions du Seuil, 1963.

John Macquarrie (1967)

BLOUNT, CHARLES
(1654–1693)

Charles Blount was an English deist, freethinker, and controversial writer on religion and politics. He was born at Upper Holloway, and was educated under the supervision of his father, Sir Henry Blount, traveler and author of *Voyage to the Levant* (1636). A disciple of Lord Herbert of Cherbury ("father of English deism") and of Thomas Hobbes, Blount is commonly regarded as the second English deist. Although not particularly original, he was the first popularizer of deistic thought. By artful writing—associating himself not only with Lord Herbert and Hobbes but also with John Dryden, Dr. Thomas Sydenham, Bishop Thomas Burnet, and Sir Thomas Browne—and by family influence, Blount was able to steer clear of prosecution under the Licensing Act and the blasphemy laws.

In 1679 Blount began a career of publication with *Anima Mundi: or an Historical Narration of The Opinions of the Ancients concerning Man's Soul after this Life: According to Unenlightened Nature*, a collection from pagan writers concerning disbelief in immortality. This was shortly followed by *The Last Sayings, or Dying Legacy of Mr. Thomas Hobbs of Malmsbury, who departed this Life on Thursday, December 4th, 1679* (1680). This work is a compilation of some of Hobbes's rationalistic (deistic) passages on religion: For example, "To say he [man] speaks by supernatural inspiration, is to say he finds an ardent desire to speak, or some strong opinion of himself, for the which he can alledge no natural reason"; "He that believes a thing only because it may be so, may as well doubt of it, because it may be otherwise."

Also in 1680 Blount published an oblique attack on priestcraft in *Great Is Diana of the Ephesians, or the Original of Idolatry, Together with the Politick Institution of the Gentiles Sacrifices*. In the same year there appeared his ironic survey of a sham pagan miracle-maker in *The Two First Books of Philostratus concerning the Life of Apollonius Tyaneus, written originally in Greek, with philological notes upon each chapter*. In 1683 Blount published *Religio Laici*, "Written in a Letter to John Dryden, Esq.," whose poem of the same name had appeared the previous year. Blount's work, long supposed to have been derived from Lord Herbert's prose tract of 1645 also titled *Religio Laici*, is now known to be much more closely related to a similarly titled manuscript of Lord Herbert's, unpublished until 1933. In his tract, Blount, under the guise of defending universal or natural religion, attacked by indirection the whole concept of a particular revelation. Attributed to Blount (by Antony a Wood) was the free translation (1683) of Chapter VI of Benedict de Spinoza's *Tractatus Theologico-Politicus* (in Latin, 1670; in English, 1689), under the title of *Miracles No Violations of the Laws of Nature*, which emphasized the Spinozistic interpretation of biblical miracles as natural phenomena or metaphorical or exaggerated language.

The appearance of Bishop Thomas Burnet's *Archaeologiae Philosophicae* (Latin and English versions in 1692) gave Blount the welcome opportunity to "vindicate" the pseudoscientific and allegorical attempts of the writer to explain certain delicate problems in the early chapters of Genesis. Writing in the form of a letter to Charles Gildon, Blount cited the authority of Sir Thomas Browne that "there are in Scripture stories that do exceed the Fables of Poets" and proceeded to ridicule Burnet's amiable rendition of the conversation between Eve and the Serpent, and his handling of such questions as "how out of only one rib a woman's whole body could be built" and "what language Adam spoke in the first hour of his nativity in naming the animals." This work, edited by Gildon, appeared in 1693, the year of Blount's death, in *The Oracles of Reason*. Another letter in the same collection from Blount to Dr. Thomas Sydenham is prefixed to *A Summary Account of the Deist's Religion*, wherein the worship of God by means of images and sacrifices or through a mediator is impugned and worship by imitation of God's perfections is upheld.

Blount, a Whig, was also active on the political front. Derived from John Milton's *Areopagitica*, his *A Just Vindication of Learning, And the Liberty of the Press*, and *Reasons humbly offered for the Liberty of Unlicensed Printing* were published in 1693. A third work of the same year, written under the pseudonym "Junius Brutus," was a master stroke demonstrating the futility of licensing. It was titled *King William and Queen Mary Conquerors: Or, A Discourse Endeavouring to prove that their Majesties have on Their Side, against the Late King, the Principal Reasons that make Conquest a Good Title*, and Blount duped the Tory licenser, Edmund Bohun, into granting

permission to publish. By order of the House of Commons the work was burnt by the common hangman, and Bohun was dismissed in disgrace (Thomas Macaulay makes much of the incident in Chapter 19 of his *History of England*).

In this year of triumph Blount let emotionalism get the better of rationalism and committed suicide over hopeless love for his deceased wife's sister, who would not agree to a marriage deemed illegal by the Church of England.

See also Deism.

Bibliography

The nearest approach to a collection of Blount's works is *Miscellaneous Works* (London, 1695), edited by Charles Gildon with a life of Blount and a justification of his suicide. But see J. S. L. Gilmour, "Some Uncollected Authors XVII: Charles Blount," in *Book Collector* 7 (1958): 182–187.

Modern studies of Blount include Harold R. Hutcheson, "Lord Herbert and the Deists," in *Journal of Philosophy* 43 (1946): 219–221; Eugene R. Purpus, "Some Notes on a Deistical Essay Attributed to Dryden," in *Philological Quarterly* 30 (1950): 342–349; George F. Sensabaugh, "Adaptations of *Areopagitica*," in *Huntington Library Quarterly* 13 (1950): 201–205; J. A. Redwood, "Charles Blount (1654–93), Deism, and English Free Thought." in *Journal of the History of Ideas* 35 (1974): 490–498.

See also the general bibliography under the *Deism* entry.

Ernest Campbell Mossner (1967)

Bibliography updated by Tamra Frei (2005)

BODIN, JEAN

(1530–1596)

Jean Bodin, the French philosopher, statesman, and early writer on economics, is known chiefly for four major systematic works: *Method for the Easy Comprehension of History* (*Methodus ad Facilem Historiarum Cognitionem,* Paris, 1566); *Six Books of the Republic* (*Six Livres de la république,* Paris, 1576); *Universae Naturae Theatrum* (The Theater of Nature; Lyons, 1596); and *Heptaplomeres Sive Colloquium de Abditus Rerum Sublimium Arcanus* (Dialogue of Seven Wise Men; Schwerin, 1857).

Although Bodin's life is only imperfectly known, he was probably born in Anjou into a Catholic family who sought social promotion through service to the king and in clerical charges. Through the help of his bishop, Bodin was admitted at an early age to the Carmelite friars of Angers, who sent him to their school in Paris. While in Paris he probably later studied under the *lecteurs royaux* instituted by Francis I, who personified for Bodin the ideal sovereign. Bodin was probably imprisoned for some time, but later released, on charges of professing Lutheran views. He later studied in Toulouse and was an assistant in the faculty of law there. He participated enthusiastically in the Renaissance ferment at Toulouse, which at that time was a great center of international learning, in close contact with Germany, Switzerland, Italy, Spain, and the papacy at Avignon. Bodin kept in touch with all foreign publications on religion and history, which benefited his lectures on the *Pandects.* He envisaged for a short time the career of a humanist historian in the capacity of headmaster of the Collège de l'Esquille, to which idea we owe a superb discourse of 1559, *Oratio de Institutenda in Republica Juventate.* In addition to a panoramic picture of the French Renaissance inspired by Francis I, the discourse presents a complete humanist pedagogical system.

The failure of his local ambitions and the expectation that the approaching religious wars would engulf Toulouse induced Bodin to leave for Paris, where he found a position as advocate of the Parliament of Paris, a favorable post for receiving any nomination of significance in the king's service. In his work in parliament, Bodin found a type of practical law far superior to the exegesis of ancient texts. He broke with the writers of such exegeses in the preface to his first systematic work, the *Method of History.* The history of the title is the history of knowledge and is similar in conception to that which René Descartes later presented in the preface to his *Principles.* For Bodin the three main branches of knowledge are human history, or anthropology; natural history, or physics; and divine history—theology or religion. The *Method* is a general outline of Bodin's whole system; his other three major works are each devoted to one of the three branches. The *Method* itself, though it outlines the entire system, covers in detail only Bodin's anthropology and discusses nearly all of the topics of the later *Republic.*

SOCIAL THEORY

The *Republic* itself, though it partly owes its genesis to Bodin's entire scheme, is also an outcome of a serious French political crisis of the period, which engaged Bodin's attention for many years. The book is a defense of the theory of the French monarchy, as Bodin conceived it, against Machiavellians in the Court and against various rebellious groups. The book seeks to demonstrate that monarchy, and the French monarchy in particular, is the best of all possible regimes.

The state, the republic, is a lawful government of the several households comprising it. The state arises when

each head of a household, each *pater familias,* acts in concert with the others. These men are the citizens of the republic. Private property is an inalienable right of the family. At the head of this group of households is the sovereign, the administrator of the republic, whose task is the proper government of the households composing the state.

SOVEREIGNTY. Bodin's whole political philosophy rests on the doctrine of sovereignty. Sovereignty is defined in the *Republic* as "the absolute and perpetual power of a Republic, that is to say the active form and personification of the great body of a modern State."

In Bodin's conception of sovereignty two different traditions, that of Roman law and that of French monarchy, converge. The former brought with it the notion of *majestas,* which gave supreme authority established above all magistrates, however important they might be, to an absolute power of which they were but a reflection. The tradition of French monarchy, in order to demonstrate the autonomy of the French king in relation to the emperor, had been concerned chiefly with cataloguing the privileges acknowledged as the king's by the pope; these were regarded as so many proofs of the king's sovereign authority. Of these *insignia pecularia,* one list contains no fewer than 208 items.

Bodin reinterprets this twofold juridical trend and attempts to synthesize it. In the *Method* he therefore retains only five marks of sovereignty: the power of appointing higher magistrates and delineating their offices, the power of promulgating or repealing laws, the power of declaring war and concluding peace, the power of judicial review, and the power of life or death even when the law requires death.

When he wrote the *Republic,* Bodin had realized that the essential mark of sovereignty was that of making and repealing laws and that the others were dependent on this right. This right of the sovereign cannot be restricted by custom; the sovereign sanctions customary law by allowing it to continue in force. "Thus, all the force of civil laws and custom lies in the power of the Sovereign Prince." All legislative and judicial power is concentrated in the sovereign, but the sovereign is conceived as the incarnation of a principle and cannot be regarded as having a personal will at variance with the interests of the state. Against the medieval theory, reaffirmed in France in Bodin's day, of the *Politie*—a state in which supreme authority was shared among the prince, an aristocracy based on birth and office, and the representatives of the people—Bodin contends that, if sovereignty is absolute, it is therefore indivisible, wherever it resides. There can be monarchies, aristocracies, or democracies, but never a mixed state.

In a given system of government, different modes of rule are possible. An aristocracy may be governed monarchically, as in Germany, or more or less democratically, as in Venice. But a monarchy, in which the king guarantees all liberty, is the best of regimes.

The state that Bodin depicts—a complex of families and of corporations, classes, and heterogeneous provinces—is enriched by the differences and interactions of its components. They all obey the sovereign, their sole arbiter and the personification of a public weal that is also the weal of its parts. Thus the absolute power of the sovereign transcends that of the paterfamilias, but is conceived in the latter's image. Though the authority of the sovereign is absolute with respect to the other elements of the state, the source of this authority lies in social law, as is clear from the long history of the French state, with its hereditary monarchy subject to a higher law. Though sovereignty is not limited by custom, it is limited by the requirements of justice: Authority is acknowledged as belonging only to a just government—a regime that gives every person, even the wicked, his chance. Sovereignty is also limited externally through the recognition of the legitimacy of other sovereignties, even of conflicting types. The sovereign is further obliged to collaborate with neighboring countries, so that M. J. Basdevant was enabled to see in Bodin one of the founders of modern international law. Bodin's thought is very close to the concept of peaceful coexistence that today forms one of the norms of international law.

THE THEORY OF CLIMATES. Besides outlining the structure of his ideal republic, a monarchy, Bodin also examines the diversity of states offered by experience. On the one hand he follows the pattern of the Greek philosophers, tracing historically the degradation of this ideal prototype and the manner in which are successively engendered the various forms, sound and pathological, of political organization— tyranny, democracy, aristocracy, and so on. But Bodin also studies the modes of a state's adaptation to its territory. In this investigation, which is known as the theory of climates from a later similar exposition by Montesquieu, Bodin seeks to define more precisely the ways through which geography influences human societies: "the nature of Northern and Southern peoples as well as that of the Eastern and Western ones, then, the influence of the various places, either mountainous, marshy, windswept or sheltered" (*Method of His-*

tory, Ch. 5). He gives a rather circumstantial account agreeing in many respects with modern human geography and ethnic psychology. He describes northerners as unequaled in wars and industry and southerners as unequaled in the contemplative sciences, but the inhabitants of the median region are in a particularly fit position for the blossoming of arts and laws.

In the *Method,* Bodin uses anthropogeography as a critical weapon to detect errors committed by outstanding historians in their assessment of facts, and to build a solid framework relating human history to natural history. In the *Republic* his point of view becomes more dogmatic, though his individual observations are more perspicacious. And he makes the important observation that, whatever the ontological superiority of monarchy over other forms of government may be, for a given state the most appropriate regime is the one that answers best to the people and the geography of the place. "One of the greatest and perhaps the chief foundation of Republics is to adapt the state to the citizens' nature, and the edicts and ordinances to the character of places, persons, and times."

Bodin's defense of the French monarchy in the *Method* and his vast culture and philosophical wisdom won him the confidence of the royal family, and in 1571 he entered the service of the duke of Alençon, the brother of the future Henri III, who, after his coronation in 1574, befriended Bodin. But in 1576, at a meeting of the States-General, Bodin delivered a speech in which he succeeded in defeating the king's request for the financial means necessary to suppress the French Protestants. By this speech Bodin temporarily diverted the civil war, but lost the king's favor and was relegated to a humble post in Laon, where he took advantage of the relative calm to write in 1578 the Latin version of the *Republic* (published Paris, 1586) and the *Demonomanie des sorciers* (Paris, 1580). The latter work, which went through some ten editions, advocates the repression of witchcraft and contains as well a complete demonology, in great part taken from the Bible.

NATURAL HISTORY

Upon his return to Laon from trips to the Court of Queen Elizabeth I and to Belgium on missions with the duke of Alençon, Bodin returned to work on the second part of his system, his physics. The *Amphiteatrum Naturae* is in the form of a dialogue in which a "mystagogue" expounds to a "theoretician" a complex and obscure philosophy that attempts to reconcile Neoplatonic idealism with Aristotelian naturalism and also with important religious attitudes derived from the Hebrew tradition. Living beings are explained in terms of Platonic forms, but the nature of the explanation and of the forms remains obscure. The soul is corporeal and is the form of the body. It is separable from the body both in life and at death. It possesses unity, and its function is to vivify the extended matter of the body. The powers of the soul, including sensation and appetite, are seen as modeled on the will: They act directly upon the body with no need of an intermediary. Angels, too, are material, and the human soul is inhabited not only by a good angel and a bad angel, but also by a large number of spirits, each in charge of a special gift. But Bodin is constrained from scrutinizing too closely the mysteries of nature by his awareness of the abyss that separates the Creator from the world of creatures. The *Amphiteatrum Naturae* thus fails, in the end, on a level where Bodin's contemporaries could not question its failure, the religious level.

THEOLOGY. A similar failure is evident in the *Heptaplomeres Sive Colloquium de Abditus Rerum Sublimium Arcanus,* a work composed during the last years of Bodin's life and published in part in 1841 and completely in 1857. This work is on the third of Bodin's three branches of knowledge, theology. The seven sages of the title represent three branches of Christianity, Judaism, Islam, natural religion, and skeptical materialism. Despite fertile discussion and a generous courtesy to one another, they cannot arrive at a common foundation for religious matters. In the progress of the discussion, it becomes apparent that in almost every instance the majority agrees with the doctrine of the Jews and that all might accept the Decalogue, looked upon as a spiritualization of the natural law and as embodying such fundamental principles. (Bodin had in an earlier work made a comparative study of the institutions of the most diverse countries, from the ancient empires to the recently discovered nations of Africa and America. From this study he had conceived the idea of replacing Roman law with a synthetic and universal law that allowed for different modes of application depending on the place, the era, and the geographic or economic conditions.)

But from the historical standpoint, which is so significant for Bodin, only the Christian faiths can contend for victory. Among these, the discussion goes badly for the Protestants, who cannot rationally justify their conservatism, their innovations, or their contradictions. The Catholic Church, since it possesses the most elaborate body of doctrine, is subjected to the most criticism; but the fact that the Catholic Church remains the religion of the state, and is relatively stable in the midst of uncer-

tainty, is for Bodin to some degree a vindication of the faith of its partisans. The book proposes, therefore, that the church is to be believed, as the Catholic prelate has held successfully throughout the dialogue.

This justification of the Catholic Church is in line with Bodin's support of the Catholic League during his last years, a support that was not dictated simply by the instinct of self-preservation. But Bodin was not fully trusted by the members of the League and was more or less confined to his house, where he spent most of his time in contemplation and the education of children, for whom he wrote a catechism in the spirit of the *Amphiteatrum Naturae.* Bodin died as a Christian and was buried in the choir of a church.

Bodin's work enjoyed outstanding renown until the middle of the seventeenth century but was totally disregarded in the eighteenth, and without a famous article in Pierre Bayle's *Dictionary,* it would never have recovered from this neglect. Bodin's work was restored to favor in 1853 through Henri Baudrillart's *Jean Bodin et son temps,* and in the twentieth century he resumed his place among the acknowledged great political philosophers of all time. Bodin also merits consideration as one of the most representative spirits of the Renaissance, and one of the first to formulate historical laws in each of the three realms—divine, natural, and human—that he considered.

See also Aristotelianism; Bayle, Pierre; Descartes, René; Idealism; Naturalism; Neoplatonism; Philosophy of Law, History of; Renaissance; Sovereignty.

Bibliography

WORKS BY BODIN

Le théâtre de la nature. Translated by François de Fougerolles. Lyons, 1597.

The Six Books of a Commonweale. Translated by Richard Knolles. London, 1606. Modern edition, edited by K. D. McRae. Cambridge, MA, 1962.

Johannis Bodini Colloquium Heptaplomeres. Edited by L. Noack. Paris and London, 1857.

La réponse à M. de Malestroict, edited by Henri Hauser. Paris, 1932. One of Bodin's economic works.

La méthode de l'histoire. Translated by Pierre Mesnard. Algiers, 1941.

Method for the Easy Comprehension of History. Translated by Beatrice Reynolds. New York: Columbia University Press, 1945.

Oeuvres philosophiquès, edited by Pierre Mesnard. Paris: Presses Universitaires de France, 1951. Vol. I.

On Sovereignty: Four Chapters from "The Six Books of the Commonwealth," edited and translated by Julian H. Franklin. New York: Cambridge University Press, 1992.

Selected Writings on Philosophy, Religion, and Politics, edited by Paul L. Rose. Geneva: Librairie Droz, 1980.

WORKS ON BODIN

Blair, Ann. *The Theater of Nature: Jean Bodin and Renaissance Science.* Princeton, NJ: Princeton University Press, 1997.

Brown, L. S. *The Methodus ad Facilem Historiarum Cognitionem.* Washington, DC: Catholic University of America Press, 1939.

Chauvire, R. *Jean Bodin l'auteur de la République.* Paris, 1914.

Engster, Dan. "Jean Bodin, Scepticism and Absolute Sovereignty" *History of Political Thought* 17 (4) (1996): 469–499.

Franklin, Julian H. "Bodin and Locke on Consent to Taxation: A Note and Observation." *History of Political Thought* 7 (1986): 89–91.

Franklin, Julian H. *Jean Bodin and the Rise of Absolutist Theory.* London: Cambridge University Press, 1973.

Franklin, Julian H. *Jean Bodin and the Sixteenth-Century Revolution in the Methodology of Law and History.* New York: Columbia University Press, 1963.

Garosci, A. *Jean Bodin.* Milan: Corticelli, 1934.

Harding, Alan. "The Origins of the Concept of the State." *History of Political Thought* 15 (1) (1994): 57–72.

King, Preston. *The Ideology of Order: A Comparative Analysis of Jean Bodin and Thomas Hobbes.* New York: Barnes and Noble, 1974.

Kuntz, Marion D. "Harmony and the Heptaplomeres of Jean Bodin." *Journal of the History of Philosophy* 12 (1974): 31–41

Lloyd, Howell A. "Sovereignty: Bodin, Hobbes, Rousseau." *Revue Internationale de Philosophie* 45 (179) (1991): 353–379.

Mayer, Jacob P., ed. *Fundamental Studies on Jean Bodin.* New York: Arno Press, 1979.

Mesnard, Pierre. *L'essor de la philosophie politique au XVIième siècle.* 2nd ed. Paris: Vrin, 1952.

Mesnard, Pierre. *Jean Bodin en la historia del pensamiento.* Translated by José Antonio Maravall. Madrid: Instituto de Estudios Políticos, 1962.

Mesnard, Pierre. "The Psychology and Pneumatology of Jean Bodin." *International Philosophical Quarterly* 2 (May 1962).

Parker, David. "Law, Society and the State in the Thought of Jean Bodin." *History of Political Thought* 2 (1981): 253–285.

Reynolds, Beatrice B. *Proponents of Limited Monarchy in Sixteenth Century France: Francis Hotman and Jean Bodin.* New York: AMS Press, 1968.

Salmon, John Hearsey McMillan. "The Legacy of Jean Bodin: Absolutism, Populism or Constitutionalism?" *History of Political Thought* 17 (4) (1996): 500–522.

Smith, Constance I. "Jean Bodin and Comparative Law." *Journal of the History of Ideas* 25 (1964): 417–422.

Pierre Mesnard (1967)

Bibliography updated by Tamra Frei (2005)

BODY-MIND PROBLEM

See *Mind-Body Problem*

BOEHME, JAKOB
(1575–1624)

Jakob Boehme, the Lutheran contemplative, was born at Alt Seidelberg near Görlitz in Silesia and lived there nearly all his life, working chiefly as a cobbler. Among his mystical experiences, the seminal one occurred in 1600, when he glanced at a pewter dish that reflected the sunlight and in a rapt state saw "the Being of Beings, the Byss and the Abyss, the eternal generation of the Trinity, the origin and descent of this world, and of all creatures through the Divine Wisdom" (*Second Epistle,* §6). Though not formally educated, Boehme read rather widely and was influenced by, among others, Paracelsus (1493–1541) and Valentin Weigel (1538–1588), the Lutheran mystic. The above quotation, however, hints at most of the main features of Boehme's *Weltanschauung,* which he first expressed in his *Aurora, oder die Morgenröte im Aufgang* (1612) and then in other works (from 1618 onward—he did not write in the intervening period because of ecclesiastical pressure). The "Abyss" is God considered as the *Ungrund*—the undifferentiated Absolute that is ineffable and neither light nor darkness, neither love nor wrath. The "eternal generation of the Trinity" occurs because the *Ungrund* contains a will to self-intuition. This will (identified with the Father) finds itself as the "heart" (the Son). Emanating from these is the "moving life" (the Spirit). This eternal process toward self-knowledge and outgoing dynamic activity generates the inner spiritual world, which is the prototype of the visible universe. With differentiation, conflict of wills becomes possible; and Satan, in severing himself from the "heart," falls. Sometimes Boehme writes as if evil were necessary, at others as though it were a contingent spoiling of the cosmic harmony. Indeed, Boehme in general shifted his position, and no single metaphysical theory fits all his writings.

This was partly because, in addition to his doctrine of the Trinity considered in itself, Boehme also enunciated a theory of seven qualities or energies in nature; and the fluidity of his metaphysics results from different ways of coordinating these two main aspects of his thought. The seven qualities divide into two triads, a higher and a lower, between which there is the crucial energy he called "the flash" (*Blitz*). The lower triad is (1) contraction (whereby substances become individuated), (2) diffusion (whereby things gravitate to one another), and (3) rotation or oscillation (the tension produced by the interplay of the forces of contraction and diffusion). The higher triad is in effect the lower triad transformed: It is (1) love, (2) expression, and (3) eternal nature or the Kingdom of God, through which there is achieved a harmony between the material and spiritual worlds.

The meaning of this evolutionary scheme is that the Trinity considered in itself is merely formal or ideal. The abysmal will needs a real object to arouse self-knowledge. Thus the Father differentiates himself through the first (lower) triad into material nature. An obstacle is thereby created to the abysmal will, which can be overcome, not by abolition, but only by transformation. The flash is the collision, as it were, between the absolute will and nature. Herein the Spirit reveals in its light the higher triad, identified with the Son as the incarnation of spirit in matter. This is the goal of the divine operation, whereby the opposition is overcome and made into a harmony.

Psychologically, the flash reveals to man his choices. He can remain at the level of anguish implicit in the welter of sensation represented by the oscillation of nature; or he can "die" unto self, and identify himself with the abysmal will—which also has to negate itself in order to achieve victory. Thus the mystical life is an imitation of Christ's suffering and triumph.

Boehme's doctrines brought him into conflict with church authorities. He was critical of the bibliolatry he detected in contemporary Protestantism, of a formalistic doctrine of election, and of crude notions of heaven (for Boehme, heaven is not a place). In England, William Law and the Behmenists (Boehme's disciples), who merged with the Quakers, were strongly influenced by him. And German Romanticism owed something to him—especially Friedrich von Schelling, notably in his later writings.

See also Evil; Law, William; Mysticism, History of; Paracelsus; Romanticism; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

WORKS BY BOEHME

The German edition of Boehme is *Jakob Böhmes sämmtliche werke,* 7 vols. (Leipzig: Barth, 1832–1860); translated by J. Ellistone and J. Sparrow as *The Works of Jacob Behmen,* 4 vols. (London, 1644–1662; reedited, London: Richardson, 1764–1781; reprinted, 1909–1924). *The Signature of All Things* and two other works are to be found in the Everyman edition, edited by Emert Rhys, introduction by Clifford Bax (London: Dent, 1912). See also W. S. Palmer, ed., *The Confessions of Jacob Boehme* (London: Methuen, 1920).

WORKS ON BOEHME

A good introduction is Rufus M. Jones, *Spiritual Reformers in the 16th and 17th Centuries* (London: Macmillan, 1914),

Chs. 9–11. See also H. H. Brinton, *The Mystic Will* (London: Allen and Unwin, 1931) and H. L. Martensen, *Jacob Boehme: His Life and Teaching,* translated by T. Rhys Evans (London: Hodder and Stoughton, 1885).

Other Recommended Titles

Brown, Robert F. *The Later Philosophy of Schelling: The Influence of Boehme on the Works of 1809–1815.* Lewisburg, PA: Bucknell University Press, 1977.

Dourley, John P. "Jacob Boehme and Paul Tillich on Trinity and God: Similarities and Differences." *Religious Studies* (4) (1995): 429–445.

Koenker, Ernest. "Potentiality in God: Jacob Boehme." *Philosophy Today* 15 (1971): 44–51.

O'Regan, Cyril. *Gnostic Apocalypse: Jacob Boehme's Haunted Narrative.* Albany: State University of New York Press, 2002.

Walsh, David. *The Mysticism of Innerworldly Fulfillment.* Gainesville: University Press of Florida, 1983.

Ninian Smart (1967)

Bibliography updated by Tamra Frei (2005)

BOETHIUS, ANICIUS MANLIUS SEVERINUS

(c. 480–524)

Anicius Manlius Severinus Boethius, late Roman statesman and philosopher, was born into the ancient Anician family in Rome, the son of a distinguished father who was consul in 487 and twice prefect of the city. Carefully educated in the liberal arts and philosophy—possibly in Athens—and precocious in genius, he entered public life at an early age under Theodoric the Ostrogoth, the Arian king of Italy from 493 to 526, who made use of Romans and the traditional administrative methods in his government.

Boethius became consul in 510 and for many years was Theodoric's principal minister (*magister officiorum*). In 522 his two sons became consuls; shortly thereafter Boethius was arrested on a charge of treason that cannot now be defined but that he denounced as a calumny. It has been suggested that he wished to exalt the Roman senate and to negotiate with Byzantium; it is also possible that as a Catholic he was distasteful to Theodoric. Condemned to exile and then to death, he was imprisoned for a year at Pavia and executed in 524. His father-in-law Symmachus and Pope John II were similarly put to death in 525 and 526.

Boethius's cult at Pavia, apparently resting on a confusion with Severinus of Cologne, won him popular canonization as a martyr. In recent centuries, however, his Christian allegiance has been questioned because of the absence of religious themes in his *De Consolatione* and the doubtful authenticity of his theological writings. The question was settled when definite proof of his authorship of these pieces was provided by H. Usener in 1877. Many readers have felt it strange that Boethius, faced with death, should have found his principal stay in Stoic and Neoplatonist philosophy, but such an attitude is not without parallels in the cultured circles of late Roman society. We may note that the readers of Boethius in the ages of faith seem to have felt no uneasiness on this count.

WRITINGS

The literary fecundity of Boethius is astonishing, especially in view of his family life and exacting official duties. He wrote on education, science, philosophy, and theology, but he was above all a logician, a translator, and a commentator. His *Elements of Arithmetic, Elements of Music,* and *Elements of Geometry* (written 500–510) all summarize existing works by Nicomachus of Gerasa and by Euclid. Of theological works attributed to him, four are now recognized as authentic: *On the Trinity* and *On the Person and Two Natures in Christ, Against Eutyches and Nestorius,* and two smaller tracts. The treatise *On the Catholic Faith* is of doubtful authenticity.

In philosophy Boethius set himself the task of translating and commenting upon all the works of Plato and Aristotle, with a view to a final harmonization of their teachings.

TRANSLATIONS. As part of his ambitious program, Boethius produced the following translations: the *Introduction* (*Isagoge*) of Porphyry and the *Categories* of Aristotle (the so-called old logic); the *Prior Analytics* and *Posterior Analytics,* the *Sophistic Arguments* and the *Topics* of Aristotle (the so-called new logic). It is questionable whether the Boethian translations are still extant among the various primitive translations that were supplanted by versions by Gerald of Cremona and others.

COMMENTARIES. Boethius produced two commentaries on the *Introduction* of Porphyry, one for beginners and the other, his chief philosophical work, for advanced students (composed 507–509); one on the *Categories* (510); on Victorinus's translation of the *Introduction* (before 505); and on the *Topics* of Cicero. In addition, he wrote several short treatises on logic.

Finally, there is Boethius's masterpiece, *On the Consolation of Philosophy,* written while he was in prison at Pavia, a dialogue in prose and verse between the writer and Philosophy personified, in which the just man unjustly suffering is confirmed in his conviction that hap-

piness and fortitude may be found in adversity. The arguments used are in part Stoic and in part Neoplatonic, but the sentiment throughout is religious, though not explicitly Christian.

Boethius lived during a period of considerable intellectual activity in Rome. Cassiodorus was his colleague, and among his elder contemporaries were the great popes Gelasius I and Hormisdas, and the canonist and chronologist Denis the Little. By his early death he escaped the disasters that befell Italy during Justinian's attempt to recapture the peninsula for the Byzantine Empire and the ravages of the Goths.

The sack and evacuation of Rome in 546 may with some assurance be taken as the dividing line in Italy between the ancient and the medieval cultures. Standing thus at the very end of a civilization, Boethius may rightly be called an eminent founder of the Middle Ages and a figure of supreme importance in the history of Western thought. Himself one of the "last of the Romans," he was also the last Western thinker to whom the works of Plato and Aristotle were familiar in Greek and to whom ancient thought in all its fullness was still comprehensible. His translations and commentaries, though neglected for centuries, stimulated and fed the minds of those who brought about the revival of dialectic in the eleventh century, and gave to medieval speculation the dialectical bent and the Aristotelian color that it never lost. Moreover, his approach to theological issues, though consciously reflecting the procedure of Augustine, was in fact more technical and dialectical in method than that of any of his predecessors. He professedly used the human power of reasoning to penetrate and explain the dogmas of Christianity and regarded the effort of reason (*ratio*) to support and discuss authority (*auctoritas*) as a principal means in the elucidation of revealed truth. On the technical level of a translator he had a genius second only to that of Cicero for exact reproduction of terms of art in his native language. Many of these terms became current coin in the Middle Ages, and a number of his definitions—those of nature, substance, person, eternity, providence, and beatitude—were accepted and stereotyped by Aquinas and others.

Boethius's influence upon the thinkers of the early scholastic period (1000–1150) can scarcely be exaggerated. It was the Boethian age as surely as the next age was Aristotelian. It was his commentary on Porphyry, in which he gave the answers of Plato and Aristotle to the "problem of universals" that initiated the great controversy on universals in the eleventh century. The early Scholastics' concentration of interest upon logic gave to the whole fabric of medieval thought from Roscelin to William of Ockham, and to the form and content of academic teaching, that preoccupation with method rather than with matter which characterized Scholastic thought, giving it accuracy and subtlety but also tending to divorce it from life and to substitute logic for discovery.

THE "CONSOLATION"

In another realm, the *Consolation of Philosophy* was one of the two or three books of universal appeal throughout the Middle Ages. Philosophically it is notable for containing a long discussion of the eternity of God, defined as the full and perfect possession of endless life always present in its entirety, and the "aeviternity" of the created universe, without beginning or end but existing in the ever-changing succession of time. On the basis of this definition, Boethius tried to solve the problem raised by God's prevision of free human acts. God in eternity has a simultaneous *vision* of all temporal reality, and he sees free acts as free. Here Boethius also made the valuable and influential distinction between that which is (*id quod est*)—for instance, the totality of parts of an individual compound substance—and that by which a substance is what it is, its being (*quo est, esse*). He identified the latter with the "form" of the whole, an important metaphysical declaration rendered classical by Thomas Aquinas. Boethius, who was engaged in distinguishing God from all other things, went on to remark that in creatures the form (*esse*) is mentally separable from the substance (*id quod est*), whereas in God his being is identical with "that which is." This is not, as has sometimes been stated, a first enunciation of the celebrated Thomist distinction between essence and existence—it is, rather, the distinction between a substance and its metaphysical cause—but it was a step on the journey, inviting further progress. The mingled melancholy, resignation to divine providence, and sense of the supreme value of the good in life in the *Consolation* appealed powerfully to the experience of those confronting the risks and disasters of medieval life, and it was to them, rather than to monks or theologians, that the work of Boethius brought comfort. It was translated into Anglo-Saxon by King Alfred the Great (c. 890), into German by Notker (c. 1000), and into French by Jean de Meung (c. 1300). It was favorite reading of Dante Alighieri, Giovanni Boccaccio, and Geoffrey Chaucer, and inspired numerous imitators.

See also Aristotelianism; Aristotle; Augustine, St.; Cicero, Marcus Tullius; Dante Alighieri; Logic, History of; Medieval Philosophy; Neoplatonism; Plato; Porphyry;

Roscelin; Stoicism; Thomas Aquinas, St.; William of Ockham.

Bibliography

Critical editions of all of Boethius's translations and of his own treatise *De divisione* have now been published as well as new editions of his *Theological Treatises*, *Consolation of Philosophy*, *De differentiis topicis*, *De hypotheticis syllogismis*, and *De institutione arithmetica*.

WORKS BY BOETHIUS

Boethius's "De Topicis Differentiis." Translated by Eleonore Stump. Ithaca, NY: Cornell University Press, 1978.

Boethius's "In Ciceronis Topica." Translated by Eleonore Stump. Ithaca, NY: Cornell University Press, 1988.

The Consolation of Philosophy. Translated with an introduction and explanatory notes by P. G. Walsh. Oxford: Clarendon Press, 1999.

De divisione. Edited and translated by John Magee. Leiden: Brill, 1998

De hypotheticis syllogismis. Edited by Luca Obertello. Brescia: Paideia, 1969.

De topicis differentiis. In *De topicis differentiis kai hoi buzantines metafraseis tou Manouel Holobolou kai Prochorou Kudone*, edited by Dimitrios Nikitas. Athens: Academy of Athens, 1969.

Opuscula Sacra and *De consolatione Philosophiae*. Edited by Claudio Moreschini. Munich: K. G. Saur, 2000.

Porphyry's *Isagoge* and Aristotle's *Categories*, *de Interpretatione*, *Prior Analytics*, *Topics*, *Sophistical Refutations*. *Arisoteles Latinus*. Vols. I–III, V, VI. Brussels: Desclée De Brouwer; Paris: Leiden, 1961–1975.

The Theological Tractates. Translated by H. F. Stewart, E. K. Rand, and S. J. Tester. Cambridge, MA: Harvard University Press, 1973.

WORKS ABOUT BOETHIUS

Chadwick, Henry. *Boethius: The Consolations of Music, Logic, Theology, and Philosophy*. Oxford: Oxford University Press, 1981.

Courcelle, Pierre. *La Consolation de Philosophie dans la tradition littéraire*. Paris: Études Augustiniennes, 1967.

Magee, John. *Boethius on Signification and Mind*. London: Brill, 1989.

Marenbon, John. *Boethius*. New York: Oxford University Press, 2002.

Martin, Christopher J. "The Logic of Negation in Boethius." *Phronesis* 36 (1991): 277–304.

Martin, Christopher J. "Non-Reductive Arguments from Impossible Hypotheses in Boethius and Philoponus." *Oxford Studies in Ancient Philosophy* 17 (1999): 279–302.

Sorabji, Richard, ed. *Aristotle Transformed: The Ancient Commentators and Their Influence*. London; Duckworth, 1990.

Sorabji, Richard. *Time, Creation, and the Continuum*. London: Duckworth, 1983.

David Knowles (1967)

Bibliography updated by Christopher J. Martin (2005)

BOETHIUS, ANICIUS MANLIUS SEVERINUS [ADDENDUM]

Boethius's position in the history of philosophy is curious. He is at best a competent representative of the Neoplatonic commentary tradition of late antiquity. His decision, however, to make that tradition available in Latin led to his having a deep and lasting influence on the development of philosophy.

It was Boethius's answer to the question left open by Porphyry that provided the basic material for later disputes over universals. Boethius argues that no extramental thing can be present entire in each of many individuals. He offers, without apparently noticing a difference, two accounts of universal concepts that are not obviously compatible. In one he maintains that mind is able to separate, or abstract, from an individual that which makes it the kind of thing it is: its species. In the other, inductive, account, he claims that the mind collects "likenesses" from many individuals to obtain their species. On either account the result is a universal. Boethius concludes that species and genera are incorporeal things that are universal in the mind and singular in sensible individuals. He does not, however, explain how this multiplied singularity is to be reconciled with his own argument against an extramental unity common to many and so leaves open to medievals the full range of positions on universals from extreme antirealism to extreme realism.

In his commentary on Porphyry's *Isagoge*, Boethius develops a distinction made there between separable accidents, such as being asleep, which a subject can cease to have without ceasing to exist, and "inseparable" accidents such as the blackness of crows, which, he claims, can be mentally but not actually separated—we can conceive of a crow that is not black but one cannot exist.

A related distinction is made by Aristotle between numerical separability and separability in account, or definition. Features included in the definition of something are conceptually inseparable from it. Inseparable accidents and properties are conceptually but not actually separable. This distinction seems to be invoked by Boethius in *De hypotheticis syllogismis*, in which he allows that we may suppose to be so what is actually impossible in order to see what follows. An example of such a nonreductive hypothesizing of an impossibility is found in his theological treatise *Quomodo substantiae*. There Boethius

posits the impossibility that god does not exist in order to explore the nature of the goodness of created things. Similar thought experiments are found in Philoponus and, through Boethius, were transmitted to the Middle Ages, during which they played a crucial role in theology and were formalized in logic textbooks as the *obligatio* of impossible *positio*.

In *Quomodo substantiae* Boethius proposes to derive his conclusion about the goodness of created beings from a set of principles that, he claims, are recognized as true by everyone or at least by the learned. These principles provided a terminology for the description of the ontological structure of created beings and God that became the standard one in Middle Ages. Boethius characterizes here as "*id quod est*" (that which is) what he refers to in his commentaries on Porphyry and Aristotle as a substance. That which makes a creature the kind of thing that it is is, according to *Quomodo substantiae*, its *esse* (being). For creatures *id quod est* and *esse* are distinct, but for God they are one and the same.

In another of the *Theological Treatises*, *Contra Eutychen*, Boethius makes a different but historically equally important distinction. Here *esse* refers to any kind of being, that is to individuals and their species and genera in the Aristotelian category of substance or any of the accidental categories. Subsistences (*subsistentia*) are beings that are not accidents and do not require accidents in order to exist—that is to say, individuals and species from the category of substance. Finally substances are the individuals but not the species in this category because, according to Boethius, a substance is a being which is the subject of accidents.

Probably the most influential of Boethius's definitions, however, was that which he gave in *Contra Eutychen* of a person as an individual substance with a rational nature. The problem that Boethius began to tackle here and that exercised theologians for rest of the Middle Ages was that of showing how God may be three persons without at the same time being three distinct substances.

In *De hypotheticis syllogismis* Boethius distinguishes two sorts of conditional propositions. Accidental conditionals such has "*If fire is hot, then the heavens are spherical*" hold merely because it is impossible for the antecedent to be true when the consequent is false. The condition is also satisfied by natural conditionals, such as "*If something is human, then its an animal*," but in these there is a connection between the antecedent and consequent.

The distinction between two forms of conditional was identified in the twelfth century with that between actual inseparability and inseparability in account and for reasoning about impossibilities only the latter were allowed. It provided the basis for Peter Abaelard's development of a unified theory of inference from Boethius's remarks on topical inference in *De differentiis topicis* and on the conditional in *De hypotheticis syllogismis*.

Boethius's own account of the hypothetical syllogism did not survive long into the twelfth century because he had no understanding of propositional negation. He thus allows inferences such as "*If (if it's an A, it's a B), then it's a C, but its not a C; therefore if it's an A, then its not a B*," which later logicians, possessing the notion of propositional negation, were able to make little sense of.

The treatment of the reconciliation of divine foreknowledge and human freedom in Books 4 and 5 of the *Consolation of Philosophy* provided the Middle Ages with one of its standard solutions. Boethius makes a distinction between absolute necessity—such as that in virtue of which a human being is an animal—and conditioned necessity—the necessity, for example, that Socrates is sitting given that he is known to be sitting. This latter necessity, he claims, is compatible with Socrates freely having chosen to sit.

Boethius argues that God's knowledge of the past, present, and future history of the world determines it only with conditioned necessity and so is compatible with human freedom. What he does not offer, however, is an account of the possibility, corresponding to that of Socrates' not choosing to sit, of the history of the world being other than it will be. Rather, he maintains, it is the expression, as fate, of the divine providential plan. Again medieval thinkers were left with a problem as much as with a solution. Of Boethius's works on the quadrivium, only two—*De institutione arithmetica* and *de Institutione Musica*—survived into the twelfth century, but they became textbooks for the rest of the Middle Ages.

See also Logic, History of; Medieval (European) Logic.

Christopher J. Martin (2005)

BOETIUS OF DACIA
(c. 13th century)

Boetius of Dacia was an Aristotelian and Averroist philosopher of the thirteenth century, sometimes called Boetius of Sweden, after the country of his birth. Born during the first half of the century, he was probably a sec-

ular cleric and canon of the diocese of Linköping. He was an associate of Siger of Brabant as a teacher of philosophy in the faculty of arts at Paris and, as a leader of the Averroist movement, condemned in 1277 by Stephen Tempier, bishop of Paris. With Siger, Boetius fled the city after the condemnation and appealed to the pope. After detention at the pontifical curia at Orvieto, Boetius joined the Dominican order as a member of the province of Dacia. The date of his death is unknown.

Boetius wrote works on logic, natural philosophy, metaphysics, and ethics. Some of these are lost; only a few have been edited. A complete edition of his extant works is now in progress.

Boetius philosophized in a rationalistic spirit, defending his right as a philosopher to discuss any subject falling within the competence of reason and to come to whatever conclusions reason dictated, even though they might contradict Christian faith. He taught, for example, that philosophizing is the most excellent human activity, that philosophers alone are the wise men of this world, that creation ex nihilo is impossible, that the world and the human species are eternal, and that there can be no resurrection of the dead. His treatise *On the Highest Good, or On the Life of the Philosopher* contains one of the most glowing and optimistic descriptions of the life of pure reason written in the Middle Ages. Setting aside the teachings of faith, Boetius inquires what reason tells us about the ultimate purpose of human life. Following Aristotle, he defines man's supreme good as the philosophical contemplation of truth and virtuous living according to the norms of nature. The philosopher alone, he concludes, lives rightly and achieves the ultimate end of human life.

Despite his rationalism, Boetius did not abandon his Christian faith but sought an ultimate reconciliation with it. Philosophy, in his view, is the work of human reason investigating the natural causes and principles of the universe, whereas the Christian religion rests on supernatural revelation and miracles of God. Because the teachings of faith have a higher source than those of philosophy, in cases of conflict the latter must give way to the former. Human reason is fallible and often comes to only probable conclusions. Even when its conclusions seem necessary, if they are contrary to revealed doctrine they are not true. In these cases truth is on the side of revelation and not on the side of reason. For example, the philosophical conclusion that the world is eternal must give way to the revealed truth that the world was created in time.

Boetius was condemned for speaking as though there were a double truth, one of faith and another of philosophy. But he carefully avoided calling true a philosophical conclusion contrary to faith.

See also Aristotelianism; Averroism; Logic, History of; Medieval Philosophy; Rationalism; Siger of Brabant.

Bibliography

Doncoeur, P. "Notes sur les averroistes latins. Boèce le Dace." *Revue des sciences philosophiques et théologiques* 4 (1910): 500–511.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages.* New York: Random House, 1955.

Grabmann, M. "Die Opuscula De Summo Bono sive De Vita Philosophi und De Sompniis des Boetius von Dacien." *Archives d'histoire doctrinale et littéraire du moyen âge* 6 (1931): 287–317.

Grabmann, Martin. *Mittelalterliches Geistesleben,* 2nd ed. Munich, 1936. Vol. II.

Grabmann, Martin. *Neuaufgefundene Werke des Siger von Brabant und Boetius von Dacien.* Munich, 1924.

Mandonnet, P. "Note complémentaire sur Boèce de Dacie." *Revue des sciences philosophiques et théologiques* 22 (1933): 246–250.

Maurer, Armand A. *Medieval Philosophy.* New York: Random House, 1962.

Sajó, G. "Boèce de Dacie et les commentaires anonymes inédits de Munich sur la physique et sur la génération attribués à Siger de Brabant." *Archives d'histoire doctrinale et littéraire du moyen âge* 33 (1958): 21–58.

Sajó, G. "Boetius de Dacia und seine philosophische Bedeutung." In *Miscellania Mediaevalia, Vol. II, Die Metaphysik im Mittelalter.* Berlin: De Gruyter, 1963.

Sajó, G. *Tractatus de Aeternitate Mundi.* Berlin, 1964.

Armand A. Maurer (1967)

BOHM, DAVID
(1917–1992)

David Bohm was a major twentieth-century physicist, and one of the world's leading authorities on quantum theory and its conceptual foundations. He was born in Wilkes-Barre, Pennsylvania, on December 20, 1917, and died on October 27, 1992, in London.

A student of J. Robert Oppenheimer, Bohm received his doctorate from the University of California at Berkeley in 1943. While still a graduate student, he discovered a particular collective movement of electrons in a plasma, now called Bohm-diffusion. At Princeton University in 1950, he completed the first of his six scientific books, *Quantum Theory*, which became the definitive exposition of the orthodox (Copenhagen) interpretation of quantum mechanics, the development of which was led by the Danish physicist Niels Bohr between 1925 and 1930. Here

Bohm presented his reformulation of the paradox of Albert Einstein, Boris Podolsky, and Nathan Rosen (EPR) concerning the possibility of simultaneous values of position and momentum for a pair of separated particles.

Bohm's version of the EPR analysis, involving components of spin in place of position and momentum, has been the basis of the enormous expansion of research on the foundations of quantum theory, focusing on nonlocality and the possible incompleteness of the quantum description (the question of "hidden variables"), that has occurred during the past several decades. Bohm and Yakir Aharonov, in 1957, made the first major step in this research when they demonstrated the existence of a "rather strange kind of correlations in the properties of distant things" (p.1072). This work was a forerunner of the seminal work of John Bell on quantum nonlocality (Bell's theorem).

In 1951 Bohm accomplished what physicists at the time regarded as impossible: He constructed, as an alternative to the prevailing observer-oriented Copenhagen interpretation of quantum theory, an objective, fully deterministic account of nonrelativistic quantum phenomena in terms of a theory describing a motion of particles under an evolution choreographed by the wave function (Bohmian mechanics). The theory Bohm proposed was in fact a rediscovery of Louis de Broglie's 1927 pilot-wave model, of which Bohm had been unaware. However, unlike de Broglie, Bohm fully appreciated the significance of the model. In particular, he showed how the predictions of the quantum measurement formalism, involving a non-noncommutative algebra of operators as observables, could be entirely explained.

In 1959 at Bristol, England, Bohm again collaborated with Aharonov, this time on a paper concerned with a very different sort of nonlocality. The result was the Aharonov-Bohm effect: In quantum mechanics a magnetic field can influence the behavior of electrons confined far away from the field, a phenomenon incompatible not only with classical physics but with the spirit of the Copenhagen interpretation of quantum theory as well. The Aharonov-Bohm effect remains, some four decades after its discovery, a subject of intense research.

Bohm was a person of extraordinary commitment to principle, both moral and scientific. He refused in 1951 to testify against colleagues before the House Un-American Activities Committee, an act that led to his indictment for contempt of Congress and his banishment from Princeton and, indeed, from all of American academia. During most of his last forty years he was engaged in an often lonely pursuit of scientific truth, showing little regard for prevailing fashion or orthodoxy.

Bohm's interests were not confined to physics. In particular, he was profoundly concerned with philosophical issues, ranging from the philosophy of science and the philosophy of mind to ethics and moral philosophy. Late in his life he was also inspired by mysticism. He saw an all-encompassing unity in the world and thought that quantum physics was but a manifestation of a deeper underlying wholeness of nature, an idea that he developed in his 1980 book *Wholeness and the Implicate Order*.

Shortly after his death Bohm's last book, *The Undivided Universe*, was published. Written in collaboration with Basil J. Hiley, his long-time colleague at London's Birkbeck College, where Bohm had for three decades been a professor, the book provided an exposition of his 1951 pilot-wave theory, together with later developments including his thoughts on the implicate order.

Bohmian mechanics in the early twenty-first century is an area of increasingly active research. However, very few scientists working in this field see an operational connection between Bohmian mechanics and Bohm's ideas on the implicate order. Nonetheless, these ideas remain an inspiration for many others.

See also Bell, John, and Bell's Theorem; Bohmian Mechanics; Quantum Mechanics.

Bibliography

Bohm, David, and Yakir Aharonov. "Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky." *Physical Review* 108 (1957): 1070–1076.

Peat, F. David. *Infinite Potential: The Life and Times of David Bohm*. Reading, MA: Addison-Wesley, 1997.

Detlef Dürr, Sheldon Goldstein, Roderich Tumulka, Nino Zanghì (2005)

BÖHME, JAKOB

See *Boehme, Jakob*

BOHMIAN MECHANICS

While quantum mechanics as presented in physics textbooks provides us with a formalism, it does not attempt to provide a description of reality. The formalism is a set of rules for computing the probability distribution of the outcome of essentially any experiment (within the realm

of quantum mechanics). A description of reality, by contrast, would tell us what processes take place on the microscopic level that lead to the random outcomes that we observe and would thus explain the formalism. While the correctness of the formalism is almost universally agreed upon, the description of the reality behind the formalism is controversial. It has also been doubted whether a description of reality needs to conform to ordinary standards of logical consistency, and whether to have such a description is desirable at all. Indeed it has often been claimed that quantum theory forces us to reject the reality of an external world that exists objectively, independently of the human mind.

BOHMIAN MECHANICS AND QUANTUM MECHANICS

Bohmian mechanics, which is also called the de Broglie-Bohm theory, the pilot-wave model, and the causal interpretation of quantum mechanics, is a version of quantum theory discovered by Louis de Broglie in 1927 (de Broglie 1928) and rediscovered by David Bohm in 1951 (Bohm 1952). It is a theory providing a description of reality, compatible with all of the quantum formalism and all of ordinary logic. In Bohmian mechanics a system of particles is described in part by its wave function, evolving according to Schrödinger's equation, the central equation of quantum theory. However the wave function provides only a partial description of the system. This description is completed by the specification of the actual positions of the particles. The latter evolve according to the "guiding equation," which expresses the velocities of the particles in terms of the wave function. Thus in Bohmian mechanics the configuration of a system of particles evolves via a deterministic motion choreographed by the wave function. In particular, when a particle is sent into a two-slit apparatus, the slit through which it passes and where it later arrives on a screen are completely determined by its initial position and wave function.

As such, Bohmian mechanics is a counterexample to the claim that quantum theory is incompatible with the reality of an objective external world. It is a "realistic quantum theory," and, since its formulation makes no reference to observers, it is also a "quantum theory without observers." For historical reasons it has been called a "hidden-variables theory." The existence of Bohmian mechanics shows that many of the radical epistemological consequences usually drawn from quantum mechanics by physicists and philosophers alike are unfounded. It shows that there is no need for contradictory notions such as "complementarity," that there is no need to imagine a particle as somehow being in two places at the same time or physical quantities as having unsharp values, and that there is no need to assume that human consciousness intervenes in physical processes (by, e.g., collapsing wave functions). Bohmian mechanics resolves all of the paradoxes of quantum mechanics, eliminating its weirdness and mystery.

THE MEASUREMENT PROBLEM

The most commonly cited of the conceptual difficulties that plague quantum mechanics is the measurement problem or, what amounts to more or less the same thing, the paradox of Schrödinger's cat. The problem is as follows: Suppose that the wave function of any individual system provides a complete description of that system. When we analyze the process of measurement in quantum mechanical terms we find that the after-measurement wave function for system and apparatus arising from Schrödinger's equation for the composite system typically involves a superposition over terms corresponding to what we would like to regard as the various possible results of the measurement—for example different pointer orientations. It is difficult to discern in this description of the after-measurement situation the actual result of the measurement—for example some specific pointer orientation. By contrast if, like Einstein, one regards the description provided by the wave function as incomplete, the measurement problem vanishes: With a theory or interpretation like Bohmian mechanics, in which the description of the after-measurement situation includes, in addition to the wave function, at least the values of the variables that register the result, there is no measurement problem. In Bohmian mechanics pointers always point.

THE EQUATIONS OF BOHMIAN MECHANICS

Bohmian mechanics is the minimal completion of Schrödinger's equation, for a nonrelativistic system of particles, to a theory describing a genuine motion of particles. For Bohmian mechanics the state of a system of N particles is described by its wave function $\psi = \psi(\mathbf{q}_1, \dots, \mathbf{q}_N) = \psi(q)$, a complex- (or spinor-) valued function on the space of possible configurations q of the system, together with its actual configuration Q defined by the actual positions $\mathbf{Q}_1, \dots, \mathbf{Q}_N$ of its particles. The theory is then defined by two evolution equations: *Schrödinger's equation*

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

for $\psi = \psi_t$, the wave function at time t, where H is the nonrelativistic (Schrödinger) Hamiltonian, containing the masses of the particles and a potential energy term, and a first-order evolution equation, *the guiding equation*

$$\frac{d\mathbf{Q}_j}{dt} = \frac{\hbar}{m_j} \operatorname{Im} \frac{\psi^* \nabla_j \psi}{\psi^* \psi} (\mathbf{Q}_1, \ldots, \mathbf{Q}_N),$$

for $Q = Q(t)$, the configuration at time t, the simplest first-order evolution equation for the positions of the particles that is compatible with the Galilean (and time-reversal) covariance of the Schrödinger evolution. Here $\hbar$ is Planck's constant divided by 2π, m_j is the mass of the j-th particle, and ∇_j is the gradient with respect to the coordinates of the j-th particle. If ψ is spinor-valued, the products in the numerator and denominator should be understood as scalar products. If external magnetic fields are present, the gradient should be understood as the covariant derivative, involving the vector potential. For an N-particle system these two equations (together with the detailed specification of the Hamiltonian, including all interactions contributing to the potential energy) completely define the Bohmian mechanics.

It is perhaps worth noting that the guiding equation is intimately connected with the de Broglie relation $\mathbf{p} = \hbar\mathbf{k}$, proposed by de Broglie in late 1923, the consideration of which quickly led Schrödinger to the discovery of his wave equation in late 1925 and early 1926. The de Broglie relation connects a particle property, momentum $\mathbf{p} = m\mathbf{v}$, to a wave property, the wave vector $\mathbf{k}$ of a plane wave $\psi(\mathbf{q}) = e^{i\mathbf{k}\cdot\mathbf{q}}$. From this one can easily guess the guiding equation as the simplest possibility for an equation of motion for Q for the case of a general wave function ψ.

Bohmian mechanics inherits and makes explicit the nonlocality implicit in the notion, common to just about all formulations and interpretations of quantum theory, of a wave function on the configuration space of a many-particle system. It accounts for all of the phenomena governed by nonrelativistic quantum mechanics, from spectral lines and scattering theory to superconductivity and quantum computing. In particular the usual measurement postulates of quantum theory, including collapse of the wave function, probabilities given by the absolute square of probability amplitudes constructed from the wave function, and the role of self-adjoint operators as observables emerge from an analysis of the two equations of motion—Schrödinger's equation and the guiding equation.

QUANTUM RANDOMNESS

The statistical significance of the wave function was first recognized in 1926 by Max Born, just after Schrödinger discovered his famous wave equation. Born postulated that the configuration Q of a quantum system is random, with probability distribution given by the density $|\psi(q)|^2$. Under the influence of the developing consensus in favor of the Copenhagen interpretation, $|\psi(q)|^2$ came to be regarded as giving the probability of *finding* the configuration Q were this to be measured, rather than of the configuration actually *being* Q, a notion that was supposed to be meaningless. In accord with these quantum probabilities, quantum measurements performed on a system with definite wave function ψ typically yield random results.

For Bohmian mechanics the $|\psi(q)|^2$-distribution has a particularly distinguished status. As an elementary consequence of Schrödinger's equation and the guiding equation, it is *equivariant*, in the sense that these equations are compatible with respect to the $|\psi(q)|^2$-distribution. More precisely this means that if, at some time t, the configuration $Q(t)$ of a Bohmian system were random, with distribution given by $|\psi_t(q)|^2$, then this would also be true for any other time. This distribution is thus called the *quantum equilibrium distribution*.

A Bohmian universe, though deterministic, evolves in such a manner that an *appearance* of randomness emerges, precisely as described by the quantum formalism. To understand how this comes about one must first appreciate that in a world governed by Bohmian mechanics, measurement apparatuses too are made of Bohmian particles. In a Bohmian universe tables, chairs, and other objects of our everyday experience are simply agglomerates of particles, described by their positions in physical space and whose evolution is governed by Bohmian mechanics.

Then, for the analysis of quantum measurements, the following observation is crucial: To the extent that the result of any quantum measurement is registered configurationally, at least potentially, the predictions of Bohmian mechanics for the result must agree with those of orthodox quantum theory (assuming the same Schrödinger equation for both) provided that the configuration Q (of the largest system required for the analysis of the measurement, with wave function ψ) is random, with probability density in fact given by the quantum equilibrium distribution, the quantum mechanical prediction for the distribution of Q.

To justify this quantum equilibrium hypothesis is a rather delicate matter, one that has been explored in con-

siderable detail (Dürr, Goldstein, and Zanghì 1992). It can be shown that the probabilities for positions given by the quantum equilibrium distribution $|\psi(q)|^2$ emerge naturally from an analysis of "equilibrium" for the deterministic dynamical system defined by Bohmian mechanics, in much the same way that the Maxwellian velocity distribution emerges from an analysis of classical thermodynamic equilibrium.

TYPICALITY

Thus, with Bohmian mechanics, the statistical description in quantum theory indeed takes, as Einstein anticipated, "an approximately analogous position to the statistical mechanics within the framework of classical mechanics" (1949, p.672). A key ingredient for appreciating the status and origin of such a statistical description is the notion of *typicality*, a notion that, historically, goes back to Ludwig Boltzmann's mechanical analysis of the second law of thermodynamics. In Bohmian mechanics, a property P is typical if it holds true for the overwhelming majority of histories $Q(t)$ of a Bohmian universe. More precisely, suppose that Ψ_t is the wave function of a universe governed by Bohmian mechanics; a property P, which a solution $Q(t)$ of the guiding equation for the entire universe can have or not have, is called *typical* if the set $S_0(P)$ of all initial configurations $Q(0)$ leading to a history $Q(t)$ with the property P has size very close to one,

$$\int_{S_0(P)} |\Psi_0(q)|^2 dq = 1 - \varepsilon \qquad 0 \le \varepsilon \ll 1,$$

with "size" understood relative to the $|\Psi_0|^2$ distribution on the configuration space of the universe. For instance, think of P as the property that a particular sequence of experiments yields results that look random (accepted by a suitable statistical test), governed by the appropriate quantum distribution. One can show, using the *law of large numbers*, that P is a typical property; see Dürr, Goldstein, and Zanghì (1992) for a thorough discussion.

OPERATORS AS OBSERVABLES

It would appear that because orthodox quantum theory supplies us with probabilities for a huge class of quantum observables and not merely for positions, it is a much richer theory than Bohmian mechanics, which seems exclusively concerned with positions. In this regard, as with so much else in the foundations of quantum mechanics, the crucial remark was made by Bell (1987 p. 666): "[I]n physics the only observations we must consider are position observations, if only the positions of instrument pointers. It is a great merit of the de Broglie-Bohm picture to force us to consider this fact. If you make axioms, rather than definitions and theorems, about the 'measurement' of anything else, then you commit redundancy and risk inconsistency."

In Bohmian mechanics, the standard quantum observables, represented by self-adjoint operators, indeed arise from an analysis of quantum experiments, as "definitions and theorems": For any quantum experiment, take as the relevant Bohmian system the combined system that includes the system upon which the experiment is performed as well as all the measuring instruments and other devices used in performing the experiment (together with all other systems with which these have significant interaction over the course of the experiment). The initial configuration is then transformed via the guiding equation for the big system into the final configuration at the conclusion of the experiment. With the quantum equilibrium hypothesis, that is, regarding the initial configuration of this big system as random in the usual quantum mechanical way, with distribution given by $|\psi|^2$, the final configuration of the big system, including in particular the orientation of instrument pointers, will be distributed according to $|\psi|^2$ at the final time.

If the experiment happens to be "measurement-like," and the outcomes of the experiment are calibrated by an assignment of numerical values to the different pointer orientations, then the induced probability distributions of these results will be given by the familiar quantum measurement postulates—that is, by the spectral measure, relative to the wave function of the system upon which the experiment is performed, of a self-adjoint operator A associated with the experiment (Dürr, Goldstein, and Zanghì 2004), in which case we speak, in orthodox quantum theory, of a "measurement of A."

The Stern-Gerlach experiment provides an illuminating example: By means of a suitable interaction (with a magnetic field), the parts of the wave function that lie in different eigenspaces of the relevant spin operator become spatially separated, and the result ("up" or "down") is thus a function of the final, detected position of the particle, concerning which we can only predict that it is random and distributed according to $|\psi|^2$ at the final time. By calibrating the outcomes of the experiment with numerical values, e.g., +1 for upper detection, and –1 for lower detection, it is not difficult to see that the probability distribution for these values can be conveniently expressed in terms of the quantum mechanical spin operators—for a spin-1/2 particle given by the Pauli spin matrices.

CONTEXTUALITY AND NAÏVE REALISM ABOUT OPERATORS

Since the result of a Stern-Gerlach experiment depends upon, not just the initial position and the initial wave function of the particle, but also on a choice among several magnetic fields that could be used to perform a Stern-Gerlach measurement of the same spin operator, this experiment is not a genuine measurement in the literal sense, that is, it does not reveal a preexisting value associated with the spin operator itself. In fact there is nothing the least bit mysterious or even nonclassical about the nonexistence of such values associated with operators. Thus the widespread idea that in a realistic quantum theory all quantum observables should possess actual values, which is in fact impossible by the Kochen-Specker theorem, was from the outset not as reasonable at it may have appeared but rather was based on taking operators as observables too seriously—an attitude, almost implicit in the word "observable," that can be called "naïve realism about operators."

Another consequence concerns *contextuality*, the notion that the result of an experiment depends not just on "what observable the experiment measures" but on more detailed information that conveys the "context" of the experiment. Contextuality is often regarded as deep, mysterious, and even close to Bohr's complementarity. However in Bohmian mechanics it boils down to the trivial insight that the result of an experiment depends on the experiment.

COLLAPSE OF THE WAVE FUNCTION

According to the quantum formalism, performing an ideal quantum measurement on a quantum system causes a random jump or "collapse" of its wave function into an eigenstate of the observable measured. But while in orthodox quantum theory the collapse is merely superimposed upon the unitary evolution of the wave function, without a precise specification of the circumstances under which it may legitimately be invoked—and this ambiguity is nothing but another facet of the measurement problem—Bohmian mechanics consistently embodies both the unitarity evolution and the collapse of the wave function as appropriate. Concerning the evolution of the wave function Bohmian mechanics is indeed formulated in terms of Schrödinger's equation alone. However, since observation implies interaction, a system under observation cannot be a closed system but rather must be a subsystem of a larger system that is closed, for example, the entire universe. And there is no reason a priori why a subsystem of a Bohmian universe should itself be a Bohmian system, even if the subsystem happens to be "closed." Indeed, it is not even clear a priori what should be meant by the wave function of a subsystem of a Bohmian universe.

The configuration Q of this larger system, this universe, naturally splits into X, the configuration of the subsystem, and Y, the configuration of its environment. Suppose the universe has wave function $\Psi = \Psi(q) = \Psi(x, y)$. According to Bohmian mechanics, this universe is then completely described by Ψ, evolving according to Schrödinger's equation, together with X and Y. Thus there is a rather obvious choice for what should be regarded as the wave function of the subsystem, namely the *conditional wave function* $\psi(x) = \Psi(x,Y)$, obtained by plugging the actual configuration of the environment into the wave function of the universe. Moreover, taking into account the way that the conditional wave function $\psi_t(x) = \Psi_t(x,Y(t))$ depends upon time, it is not difficult to see that it obeys Schrödinger's equation for the subsystem when that system is suitably decoupled from its environment and, using the quantum equilibrium hypothesis, that it randomly collapses according to the usual quantum mechanical rules under precisely those conditions on the interaction between the subsystem and its environment that define an ideal quantum measurement.

UNCERTAINTY

It follows from the quantum equilibrium hypothesis and the definition of the conditional wave function that when the (conditional) wave function of a subsystem is ψ, its configuration must be random, with distribution $|\psi(x)|^2$, even if its full microscopic environment Y—itself grossly more than what we could conceivably have access to—were taken into account. In other words, the (conditional) wave function ψ of a subsystem represents maximal information about its configuration X. Thus, in a universe governed by Bohmian mechanics there are sharp, precise, and irreducible limitations on the possibility of obtaining knowledge, limitations which can in no way be diminished through technological progress leading to better means of measurement. This *absolute uncertainty* is in precise agreement with Heisenberg's uncertainty principle. The fact that knowledge of the configuration of a system must be mediated by its (conditional) wave function may partially account, from a Bohmian perspective, for how orthodox physicists could identify the state of a quantum system—its complete description—with its (collapsed) wave function without encountering any practical difficulties.

OBJECTIONS

A great many objections have been and continue to be raised against Bohmian mechanics. Most of these objections have little or no merit. The most serious is that Bohmian mechanics does not account for phenomena such as pair creation and annihilation characteristic of quantum field theory. However this is not an objection to Bohmian mechanics per se but merely a recognition that quantum field theory explains a great deal more than does nonrelativistic quantum mechanics, whether in orthodox or Bohmian form. It does however underline the need to find an adequate, if not compelling, Bohmian version of quantum field theory, and of gauge theories in particular, a problem that is pretty much wide open.

A related objection is that Bohmian mechanics cannot be made Lorentz invariant, by which it is presumably meant that no Bohmian theory—no theory that could be regarded somehow as a natural extension of Bohmian mechanics—can be found that is Lorentz invariant. The main reason for this belief is the manifest nonlocality of Bohmian mechanics. But nonlocality, as John Bell has argued and the experiments have shown, is a fact of nature. Moreover, concerning the widespread belief that standard quantum theories have no difficulty incorporating relativity while Bohmian mechanics does, there is much less here than meets the eye. On the one hand, one should keep in mind that the empirical import of orthodox quantum mechanics relies on both the unitary evolution of the state vector (or the equivalent unitary evolution of the operators in the Heisenberg representation) and the collapse or reduction of the state vector (or any other equivalent device that incorporates the effect of observation or measurement). But the Lorentz invariance of this part of the theory has rarely been considered in a serious way—most of the empirical import of standard relativistic quantum mechanics is in the so-called "scattering regime." But if this were done, arguably, the tension between Lorentz invariance and quantum nonlocality would soon become manifest. On the other hand, a variety of approaches to the construction of a Lorentz invariant Bohmian theory have in fact been proposed, and some toy models formulated.

WHAT IS A BOHMIAN THEORY?

Finding a satisfactory relativistic version of Bohmian mechanics and extending Bohmian mechanics to quantum field theory are topics of ongoing research and we shall not attempt to give an overview here. (Some remarks, however, are given in the next section.) Rather we shall briefly sketch what we consider to be the general traits of any theory that could be regarded as a natural extension of Bohmian mechanics. Three requirements seem essential to us: 1. The theory should be based upon a clear ontology, the *primitive ontology* representing what the theory is fundamentally about—the basic kinds of entities (such as the particles in Bohmian mechanics) that are to be the building blocks of everything else, including tables, chairs, and measurement apparatuses. 2. There should be a quantum state vector, a wave function, that evolves according to the unitary quantum evolution and whose role is to somehow generate the motion for the variables describing the primitive ontology. 3. The predictions should agree (at least approximately) with those of orthodox quantum theory—at least to the extent that the latter are unambiguous. Note that we do not regard as essential either the deterministic character of the dynamics of the primitive ontology or that the latter should be given by particles described by their positions in physical three-dimensional space—a field ontology, or a string ontology would do just as well.

In short a "Bohmian theory" is merely a quantum theory with a coherent ontology. But when the theory is regarded in these very general terms, an interesting philosophical lesson emerges: In the structure of a Bohmian theory one can recognize some general features that are indeed common to all "quantum theories without observers," that is, to all precise formulations of quantum theory not based on such vague and imprecise notions as "measurement" or "observer"—such as Ghirardi-Rimini-Weber-Pearle's "dynamical reduction" models or Gell-Mann and Hartle's "decoherent histories" approach. One essential feature is the primitive ontology of the theory—what the theory is fundamentally about. The other very general and crucial feature is the sort of explanation of physical phenomena the theory should provide: an *explanation based on typicality*. Not just for a Bohmian theory, but for any physical theory with probabilistic content, the physical import of the theory must arise from its provision of a notion of typical space-time histories, specified for example via a probability distribution on the set of all possible histories of the primitive ontology of the theory.

HISTORY AND PRESENT STATUS

In 1951 Bohm rediscovered de Broglie's 1927 pilot-wave model and showed that the quantum measurement formalism, based on non-commuting operators as observables, emerged from the basic principles of de Broglie's theory. Since then Bohmian mechanics has been developed and refined: Noteworthy are Bell's clarification of the axioms of the theory and the analysis of the status of

probability and the role of typicality (Bell 1987; Dürr, Goldstein, and Zanghì 1992), as well as the investigations of quantum non-equilibrium (Valentini 2002). Several ways of extending Bohmian mechanics to quantum field theory have been proposed. One (Bohm 1952) for bosons (i.e., force fields) is based on an actual field configuration on physical three-dimensional space that is guided by a wave functional according to an infinite-dimensional analogue of the guiding equation (see also Bohm and Hiley 1993; Holland 1993). Another proposal (Dürr, Goldstein, Tumulka, and Zanghì 2004) relies on seminal work by Bell (1987 p. 173) and ascribes trajectories to the electrons or whatever sort of particles the quantum field theory is about; however, in contrast to the original Bohmian mechanics, this proposal involves a stochastic dynamics, according to which particles can be created and annihilated.

See also Bell, John, and Bell's Theorem; Bohm, David; Boltzmann, Ludwig; Einstein, Albert; Quantum Mechanics; Realism.

Bibliography

Bell, John. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge, U.K.: Cambridge University Press, 1987.

Bohm, David. "A Suggested Interpretation of Quantum Theory in Terms of 'Hidden' Variables, Parts I and II." *Physical Review* 85 (1952): 166–193.

Bohm, David, and Basil Hiley. *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge & Kegan Paul, 1993.

Broglie, Louis de. *Electrons et Photons: Rapports et Discussions du Cinquième Conseil de Physique tenu à Bruxelles du 24 au 29 Octobre 1927 sous les Auspices de l'Institut International de Physique Solvay*. Paris: Gauthier-Villars, 1928.

Dürr, Detlef, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghì. "Bohmian Mechanics and Quantum Field Theory." *Physical Review Letters* 93 (2004): 090402.

Dürr, Detlef, Sheldon Goldstein, and Nino Zanghì. "Quantum Equilibrium and the Origin of Absolute Uncertainty." *Journal of Statistical Physics* 67 (1992): 843–907.

Dürr, Detlef, Sheldon Goldstein, and Nino Zanghì. "Quantum Equilibrium and the Role of Operators as Observables in Quantum Theory." *Journal of Statistical Physics* 116 (2004): 959–1055.

Einstein, Albert. "Reply to Criticisms." In *Albert Einstein, Philosopher-Scientist*, edited by Paul Arthur Schilpp. New York: Tudor, 1951.

Holland, Peter. *The Quantum Theory of Motion*. Cambridge, U.K.: Cambridge University Press, 1993.

Valentini, Antony. "Subquantum Information and Computation." *Pramana-Journal of Physics* 59 (2002): 269–277.

Detlef Dürr, Sheldon Goldstein, Roderich Tumulka, Nino Zanghì (2005)

BOHR, NIELS
(1885–1962)

Quantum physics is often credited with far-reaching metaphysical and epistemological implications, including the denial of causality and determinism and the existence of strict limits on what can be known about natural systems. One of the main figures whose work has been used—and often misused—in support of such conclusions is the Danish physicist Niels Bohr. Bohr is rightfully viewed as one of the major figures in the history of quantum physics and is widely known both for his extraordinary contributions to the development of quantum theory and for his philosophically oriented work, which focused on the task of interpreting the quantum mechanics. Bohr's interpretation centers on his notion of complementarity, which he developed in 1927, two years after the development of quantum mechanics by Heisenberg, Born, Jordan, and Schrödinger and shortly after the publication of Heisenberg's famous uncertainty paper.

Bohr's interpretive approach attracted many followers but also many critics. Most notable among the latter was Einstein, whose public critique of quantum mechanics and Bohr's interpretation began in 1927 and culminated with his 1935 "EPR" paper, written with Podolsky and Rosen. Bohr's response to Einstein's criticisms, and part of his general interpretive approach, was that quantum mechanics is a complete theory the statistical indeterminacies of which neither need be nor could be overcome with a more foundational theory.

While Bohr is most philosophical after the introduction of complementarity, the overarching theme of much of his earlier work was also associated with certain clear philosophical ideas about the nature of physical theories and the appropriate method for developing a theory in a new realm, and complementarity can be seen as an application of these ideas to the new quantum mechanical formalism.

QUANTUM THEORY

Bohr's famous 1913 model of the hydrogen atom, with which he explained the hydrogen spectrum, marks the beginning of the quantum theory of the atom. Because classical electrodynamics had dictated that the oscillation of electrons is accompanied by the emission of electromagnetic radiation, the theory could account neither for the stability of the atom nor for the discreteness of the spectrum of frequencies emitted by excited hydrogen gas. Bohr's model solved this puzzle by suggesting that the electron orbits the nucleus in stable stationary states and

that the emission of radiation occurs not during that orbit but rather in sudden transitions between the states; the radiation carries the difference in energy between the states according to a quantum frequency rule (based on work by Planck and Einstein) that correlates energy with frequency. Bohr eventually presented the rationale for his model in terms of the quantization of angular momentum, and that is how it is often presented in texts. However, Bohr's original rationale, and arguably the one closest to his actual approach to physics in the years afterwards, is that he read the existence of independent stationary states off the Balmer formula of the hydrogen spectrum by interpreting the spectrum with the quantum-frequency rule. That is, the discrete stationary states were not hypothesized but rather were inferred from an empirical generalization.

THE CORRESPONDENCE PRINCIPLE

Bohr eventually expanded this general approach of inferring atomic properties from empirical generalizations or phenomena with the development of his correspondence principle. The principle, first implicitly used in a general form in 1918 and named as a specific principle by Bohr in 1920, is a claim about the relationship between classical and quantum theory, and in particular about classical descriptions of empirical evidence and quantum models of the atom. As Bohr sometimes stated it—the way in which it is most often quoted—the new quantum theory ought to recapture classical electrodynamics in some limit—that is, the old theory ought to be shown to be an approximation that in retrospect is roughly accurate in the realms where quantum effects are negligible. In the hydrogen atom, according to Bohr's principle, that will occur when the quantum number is high, where the difference in energy between stationary states becomes small in comparison with the energies of the states themselves.

While it is tempting to understand the correspondence principle as a requirement for the rationality of the progression of theories, that is at best only one aspect of Bohr's approach with the principle. For Bohr, the correspondence principle was an intratheory claim, not an intertheory one, and it was important because the developing quantum theory had no account of the relation between the motions of the electrons within their orbits and the empirical phenomena of the atomic spectra, whereas classical theory had had such an account. Bohr consistently insisted that we need a stable description of observations from which we can infer atomic properties, and he emphasized that generalizations about atomic spectra—about the frequencies of radiation emitted or absorbed by atoms—are essentially claims about wave phenomena, because measurements of radiation frequencies with spectroscopy equipment unavoidably assume wave theory. Thus, even though the quantum theory might seem to call into question the wave nature of electromagnetic radiation (at least according to the light-quantum concept implied by the photoelectric effect, and later by the Compton effect), scientists still must use wave electrodynamics to provide evidence about atomic properties, so a link or coordination between the theories is needed.

The agreement in the limit between the theories was therefore not the goal of the correspondence principle but only a means of allowing the linkage of claims within the new theory. In particular, it let Bohr relate periodic motion within the atom to periodic aspects of the radiation in the spectrum. This principle both gave empirical content to parts of the model that previously had had none and allowed the inference of properties of certain atomic processes—for example, selection rules for quantum transitions—for which there was no other method of determination. For Bohr the principle was a way to relate observable, empirical phenomena with the quantum mechanisms (such as they were) "behind" the empirical phenomena.

Two related aspects of the correspondence principle were very important for Bohr's work after the development of quantum mechanics. First, although Bohr had been able to apply it only imprecisely and often only qualitatively, it inspired Heisenberg's approach in developing what was to become quantum mechanics, and Bohr claimed that quantum mechanics embodied the correspondence principle. Second, the general approach of incorporating independent, classically based descriptions of empirical phenomena within quantum theory became the foundation for his own interpretation of that quantum mechanics.

COMPLEMENTARITY AND THE INTERPRETATION OF QUANTUM MECHANICS

Bohr's interpretation is notoriously difficult to pin down, but the core ideas are that our descriptions of the properties of quantum systems must be based on classical concepts, that these concepts are restricted in scope to a particular experimental context, that different concepts are appropriate for different contexts, that the different contexts make the use of certain pairs of mutually exclusive concepts, and that those concepts do not fully capture the nature of quantum systems. Bohr used the word

"complementarity" to describe this complex of ideas that together were meant to address interpretive problems posed by quantum mechanics.

Quantum mechanics, especially in Heisenberg's formulation, had retained some aspects of the old quantum theory but had abandoned the definite electron orbits of that theory and had substituted abstract, formal methods for calculating "observable" properties. Heisenberg's uncertainty paper had given a further argument for thinking in these terms by deriving equations that described a reciprocal relationship between the precisions with which certain pairs of properties (for example, position and momentum) could be measured. Although there is some indication in Heisenberg's paper that he might have thought of the tradeoffs in precision in terms of disturbance (every measurement of one property disturbs a specific other one in a way that prevents us from knowing simultaneously both properties to arbitrary precision), Bohr associated the uncertainty relations with his notion of complementarity and claimed that the uncertainty or indeterminacy described by the relations reflect not merely a lack of knowledge of the values of metaphysically definite properties of a system, but rather a degree to which our concepts just do not and cannot be made to apply to the system. Complementarity claims that, although we cannot simultaneously give both normal space-time and causal descriptions of the same quantum phenomenon and although neither description fully captures the nature of the phenomenon, we nevertheless have no other way to describe phenomena besides through these causal and spatiotemporal pictures.

Although Bohr's philosophy is sometimes called the Copenhagen interpretation, there are important distinctions between Bohr's actual views and what is often meant by that name. The name is sometimes used to describe what might better be called the standard interpretation, which is perhaps inspired by Bohr but is really based on von Neumann's work and includes the collapse of the wave packet, which had no part in Bohr's philosophy. Otherwise, it is used to describe a set of views held by Bohr and a number of his former students and associates from Copenhagen, especially Heisenberg and Pauli, but there are disputes regarding how much their views really had in common.

Central to Bohr's interpretation is a sort of holism that we can now understand in terms of entanglement. This holism is clear in Bohr's work starting in 1929 and certainly by 1935. Bohr then explicitly states that it is misleading to think that observation disturbs properties because that would imply the existence of preexisting complete sets of properties. Bohr emphasized that the novel and interpretively challenging aspect of quantum effects is not the discreteness of, say, the exchange of energy but rather the apparent mathematical and theoretical fact that quantum mechanical processes generally cannot be broken down in a way that allows us accurately to describe them in terms of an interaction between component systems such as a measuring instrument and a measured system. In order to describe or analyze an experiment, scientists nevertheless must treat measurement in this way, and the consequence is that descriptions of measured properties of subsystems of a larger whole system at best misconstrue the true quantum mechanical state or phenomenon. And it is precisely in this misconstrual that the statistical nature of quantum mechanical predictions arise.

Though not all interpreters of Bohr agree, this explicit emphasis in his later work did not represent a drastic change in his interpretation. Indeed, it is plausible to argue that complementarity is and was always for Bohr a conclusion based on his correspondence approach and the discovery of noncommutativity and the holism of entanglement. Bohr thought that although one can give an abstract mathematical representation of a quantum mechanical system independent of classical conceptualizations of the phenomena, the symbols used to represent quantum properties have empirical meaning only when they can be associated or put into correspondence with observable phenomena. Doing this requires first establishing independent theoretical descriptions of the observations, and for this it is necessary to use classical concepts to describe the measurement context. Complementarity is, then, an expression of the limitations that noncommutativity places on the degree to which different quantum symbols can be given empirical meaning.

Although Bohr was a realist about the entities described by quantum mechanics and he seems to have believed that quantum mechanics does describe the true nature of quantum-mechanical systems, the foregoing features of his work suggest certain antirealist aspects to his interpretation, especially with respect to the way the meaning and applicability of our concepts about quantum properties depend somehow on the context in which those properties are measured.

This tension is evident in Bohr's response to the EPR paper. That paper questioned the completeness of quantum mechanics precisely on the grounds of the quantum relations of entangled systems; EPR claimed that the ability to predict the properties of one of an entangled pair of particles after the measurement of the other, over dis-

tances and within times that preclude a causal interaction on relativistic grounds, indicates that quantum mechanics must assume that the prediction concerns a real, preexisting property that is independent of the other measurement. Bohr's response does not explicitly deny realism but says that any descriptive account of quantum reality is good only within the conditions of applicability of the concepts used in measurement and prediction and that the effect on the distant particle is not a causal, physical one but rather an effect on those conditions; this suggests, perhaps, that disentanglement is only conceptual.

Although in later years Bohr began to discuss complementarity in increasingly broad terms and as applied to other fields, especially biology, it is his philosophical work closest to physics that has had the greatest impact in both philosophy and physics. In the early twenty-first century theorems about the impossibility of certain kinds of hidden variable theories can be seen as a vindication of many of the intuitions in that work, intuitions that remain evident in the pragmatic approach to quantum mechanics assumed by many working physicists.

See also Bell, John, and Bell's Theorem; Bohm, David; Copenhagen Interpretation; Einstein, Albert; Heisenberg, Werner; Quantum Mechanics.

Bibliography

GENERAL

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777–780. Reprinted in Wheeler and Zurek.

Heisenberg, Werner. "Über den ansclauchichen Inhalt der quantentheoretischen Kinematik und Mechanik." *Zeitschrift für Physik* 43 (1927): 172–198. Translated as "The Physical Content of Quantum Kinematics and Mechanics" in Wheeler and Zurek.

Wheeler, John Archibald, and Wojciech Hubert Zurek. *Quantum Theory and Measurement*. Princeton, NJ: Princeton University Press, 1983.

WORKS BY NIELS BOHR

"On the Constitution of Atoms and Molecules." *Philosophical Magazine* 26 (1913): 1–25, 476–502, 857–875.

"The Quantum Postulate and the Recent Development of Atomic Theory." *Nature* 121 (supplement) (1928): 580–590.

"Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 48 (1935): 696–702.

"Discussion with Einstein on Epistemological Problems in Atomic Physics." In *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp. Cambridge: Cambridge University Press, 1949. Reprinted in Wheeler and Zurek.

Atomic Theory and the Description of Nature, reprinted as *The Philosophical Writings of Niels Bohr*. Vol. 1. Woodbridge, CT: Ox Bow Press, 1934/1987.

Essays 1932–1957 on Atomic Physics and Human Knowledge. Reprinted as *The Philosophical Writings of Niels Bohr*. Vol. 2. Woodbridge, CT: Ox Bow Press, 1958/1987.

Essays 1958–1962 on Atomic Physics and Human Knowledge. Reprinted as *The Philosophical Writings of Niels Bohr*. Vol. 3. Woodbridge, CT: Ox Bow Press, 1963/1987.

Causality and Complementarity, edited by Jan Faye and Henry Folse as *The Philosophical Writings of Niels Bohr*. Vol. 4. Woodbridge, CT: Ox Bow Press, 1998.

WORKS ABOUT NIELS BOHR

Beller, Mara. *Quantum Dialogue: The Making of a Revolution*. Chicago: University of Chicago Press, 1999.

Faye, Jan. *Niels Bohr: His Heritage and Legacy: An Anti-Realist View of Quantum Mechanics*. Dordrecht: Kluwer Academic Publishers, 1991.

Faye, Jan, and Henry J. Folse, eds. *Niels Bohr and Contemporary Philosophy*. Boston Studies in the Philosophy of Science 153. Dordrecht: Kluwer Academic Publishers, 1994.

Folse, Henry J. *The Philosophy of Niels Bohr: The Framework of Complementarity*. Amsterdam: North-Holland, 1985.

Jammer, Max. *The Conceptual Development of Quantum Mechanics*. New York: McGraw-Hill, 1966.

Pais, Abraham. *Niels Bohr's Times: in Physics, Philosophy, and Polity*. Oxford: Clarendon Press, 1991.

Tanona, Scott. "Uncertainty in Bohr's Response to the Heisenberg Microscope." *Studies in History and Philosophy of Modern Physics* 35 (2004): 483–507.

Scott Tanona (2005)

BOILEAU, NICOLAS

(1636–1711)

Nicolas Boileau, also known as Boileau-Despréaux, has retrospectively been raised to the rank of emblematic figure of French classicism. He has been described as the "lawgiver of Parnassus" (a reference to his being an arbiter of taste), the champion of poetic rationalism, and a chief apologist for the ancients in their quarrel with the moderns. At the beginning of the twenty-first century, specialists of the era consider the truth about Boileau to be more nuanced. Boileau was first and foremost a poet engaged in the literary life of his time. After having written his *Satires*, a vigorous denunciation of the faults and mistakes commonly made in the literary world of his days, he attempted, in his *Art poétique* (1674), to determine the rules that should govern the creation and reception of art in most literary genres.

Published during the same year, his translation of Longinus's *Peri hypsous* (*On the Sublime*, first cent.) contributed to popularizing this work all over Europe. In 1677 he became, along with Jean Racine, the historiographer of Louis XIV. This noticeably slowed down his literary production. From 1687 on, as defender of the

ancients, he was Charles Perrault's main adversary in the first of two disputes between the ancients and the moderns that divided the field of classical aesthetics in France. His nine *Réflexions critiques sur quelques passages du rhéteur Longin* (Critical reflections on several passages of the orator Longinus; 1694) are explicit arguments in favor of the advocates of the ancients.

Boileau's position was not simply the result of a general nostalgic or conservative attitude, but rather followed from his very strict conception of literature. His aim is to look at the ancients' masterworks in order to find examples of perfection to stimulate the creativity and imagination of contemporaries, and models to provide the distance necessary to avoid the relativist pitfalls, not to mention the conceit, that threatened modernist partisans. According to Boileau, the criterion by which one can attest to the merit of the great artworks of the past is that they have passed the test of time. Far from being an illegitimate prejudice, imitation of the ancients is the source of the true rules of art, which reason can use as its guide.

Two aspects of Boileau's thought are of interest to the historian of philosophical aesthetics. First, there is his formulation of classical doctrine, of which *Art poétique* provides a synthesis. Far from displaying the merely theoretical attitude of an arbiter, Boileau reflects the aesthetic consensus obtained during the decades from 1630 to 1670 on the basis of a precarious balance between reason and sentiment, freedom and norms. Second, there is his clarification of the role of the sublime in poetry. In discussing the sublime, Boileau tried to cast light on the causes of the legitimate and enduring admiration we have for authors of merit, whether ancient or modern.

Art poétique, where Boileau provided a synthesis of classical doctrine, explicitly draws from the tradition inherited from Aristotle and Horace. It is divided into four cantos written in verse. The first canto gives authors general advice on poetry. The second canto deals with minor genres: the eclogue, sonnet, ode, satire, elegy, epigram, and the like. The third canto tackles major genres: tragedy, comedy, and epic. The fourth canto gives rules for writing, insisting on the edifying function of poetry, on the writer's disinterestedness, and on the need for the writer to surround himself with friends whose sound judgment will help him improve himself.

In the course of the four cantos, Boileau simply reaffirmed, without ever analyzing, all the principles of classical aesthetics. If genius, as a natural gift, is necessary to write poetry, only art, polishing of the work under the guidance of reason and judgment, can lead to perfection. Thus, although it is not a source of inspiration, the light of reason must nonetheless accompany the conception of thoughts, their arrangement, and their expression. As far as tragedy is concerned, Boileau reinforced the classical interpretation of the Aristotelian theory held by his contemporaries. Tragic art was said to provide an idealizing imitation of the terrifying in which pain is transformed into pleasure. The purpose of tragedy is to please and move the spectator by producing a "pleasant terror" and a "delightful pity." To produce such effects, however, reason must be respected.

Thus Boileau advocated absolute respect for the three unities of action, time, and place, even though Aristotle confined himself to the unity of action. Also, the representation ought to be submitted to the principle of *verisimilitude*, since what is historically true but not credible will not produce any emotion in the spectator. Verisimilitude also requires the writer to respect the rules of propriety (Horace's decorum), whether from an external point of view (agreement between the represented action and the public's expectations and customs) or from an internal one (internal coherence among characters and the language ascribed to them).

For Boileau, the sublime constitutes the supreme perfection of poetic discourse. He saw a nonrhetorical conception of the sublime at work in Longinus's treatise, one that makes possible the distinction between the really sublime (what "strikes us in a discourse, elevates, ravishes and transports us" (*On the Sublime*, first cent) and the sublime *style* (the lofty style that traditional rhetoric thought best adapted to the expression of noble ideas). The sublime can thus be found in a single thought or turn of phrase, an excellent example being God's command "Let there be light," in Genesis. The sublime reconciles grandeur and conciseness in accordance with the demands of simplicity and naturalness imposed by the aesthetics of classicism.

In his last three reflections on Longinus, published posthumously in 1713, Boileau added that the perfectly sublime—that which has the property of elevating the soul and making us participate in the greatness that we perceive—unites the grandeur of the thought with the nobility of the sentiment driving the person expressing it, the splendor of the words, and the harmony of the expression. The sublime is, paradoxically, the summit of Boileau's aesthetics. On the one hand, the "energic littleness of the words" (*Réflexions* X) manifests the sublime in the density of meaning sought by classicism. On the other hand, favoring the sublime introduces tension in a system of thought governed by the ideal of reason and clarity. The significant role of the sublime sufficiently demon-

strates that classicism, far from being a sterile formalism, is in fact a constantly renewed demand for equilibrium between judgment and inspiration, lucidity and emotion, conciseness and grandeur.

See also Aesthetics, History of.

Bibliography

WORKS BY BOILEAU

Œuvres complètes, edited by Antoine Adam and Françoise Escal. Paris: Gallimard, 1966.

WORKS ON BOILEAU

Beugnot, Bernard, and Roger Zuber. *Boileau: Visages anciens, visages nouveaux, 1665–1970*. Montreal: Presses de l'Université de Montréal, 1973.

Brody, Jules. *Boileau and Longinus*. Geneva: Droz, 1958.

Génetiot, Alain. *Le classicisme*. Paris: Presses universitaires de France, 2005.

Pocock, Gordon. *Boileau and the Nature of Neo-classicism*. Cambridge, U.K.: Cambridge University Press, 1980.

Wood, Theodore E. B. *The Word "Sublime" and Its Context, 1650–1760*. The Hague: Mouton, 1972.

Daniel Dumouchel (2005)

BOLINGBROKE, HENRY ST. JOHN
(1678–1751)

Henry St. John Bolingbroke, the English Tory statesman, orator, man of letters, friend of the Augustan wits, libertine, and deist, was born at Battersea, the son of Sir Henry St. John and Lady Mary Rich, daughter of the second earl of Warwick. After early schooling by his paternal grandmother, he was educated at Eton and, putatively, at Christ Church, Oxford, for in 1702 he was made an honorary doctor of Oxford. He had made the customary dissipated grand tour, 1698–1699, but he also mastered several languages and studied the history and customs of the lands he visited. In 1701 he became M.P. for the family borough of Wootton Bassett in Wiltshire. His eloquence and brilliance soon made him a leader of the Tory party. With the help of Robert Harley, he became secretary at war in 1704, but resigned in protest over the dismissal of Harley in 1708. The growing unpopularity of the "Whiggish" War of the Spanish Succession brought Harley back into power in 1710, and Bolingbroke joined the new Tory ministry as secretary of state. Two years later he was created Viscount Bolingbroke and was one of the negotiators of the Treaty of Utrecht signed in 1713. Following the accession of George I in 1714, Bolingbroke and the other Tory ministers were dismissed from office. In 1715 he fled to France to take political asylum for alleged Jacobitism. In 1723 he was pardoned, and he spent the remainder of his life living variously in England and in France.

WORKS

Some of Bolingbroke's political writings appeared in the Tory periodical the *Craftsman* between 1726 and 1736; but most others, including the philosophical, were published posthumously in 1754 by David Mallet in an edition of five quarto volumes. This publication elicited Dr. Johnson's famous attack on this "blunderbuss against religion and morality." David Hume's reaction is less well known but more pertinent:

> Lord Bolingbroke's posthumous Productions have at last convinc'd the whole World, that he ow'd his Character chiefly to his being a man of Quality, & to the Prevalence of Faction. Never were so many Volumes, containing so little Variety & Instruction: so much Arrogance & Declamation. The Clergy are all enrag'd against him; but they have no Reason. Were they never attack'd by more forcible Weapons than his, they might for ever keep Possession of their Authority.

POLITICAL AND HISTORICAL WORKS. Bolingbroke's contributions to the *Craftsman* exhibit much vigorous political writing, including *Remarks on the History of England* and *Dissertation on Parties.* Other tracts, political and historical, are *On the True Use of Retirement and Study, On the Spirit of Patriotism,* and *Letters on the Study and Use of History,* the last of which made famous the maxim, "History is philosophy teaching by examples." *The Idea of a Patriot King* also became famous because of its use in the education of the future George III. Matthew Arnold was to lament that Bolingbroke's historical writings were unduly neglected. Unfortunately, the neglect of his philosophical writings is less to be regretted.

PHILOSOPHICAL WRITINGS. Bolingbroke made much of the antithesis between nature and art; that is, the alleged superiority of a pure state of nature over the evils of civil society. Edmund Burke, who wrote his *Vindication of Natural Society* (1756) as an imitation of Bolingbroke's style and as an ironic refutation of this antithesis, asked rhetorically in *Reflections on the Revolution in France* (1790): "Who now reads Bolingbroke? Who ever read him through?" The long-held myth of Voltaire's great indebtedness to Bolingbroke has been completely disproved by N. L. Torrey. A similar claim of Alexander

Pope's great indebtedness has been vigorously challenged by Maynard Mack, who presents evidence that Bolingbroke's *Fragments or Minutes of Essays* were composed later than the *Essay on Man.* There is, however, no question that Pope discussed many matters with his "Guide, Philosopher, and Friend." With the single exception of Peter Annet, Bolingbroke was the last of the distinguished group of English deists beginning with Lord Herbert of Cherbury; but he proves somewhat of a disappointment to students of the history of ideas. Scrappy and unsystematic in his presentations, he is replete with contradictions. Despite recent attempts, especially by D. G. James and W. McMerrill, to take Bolingbroke's philosophy more seriously than has been customary, candor demands the conclusion that, although his style is more eloquent than that of most other deists, he contributed little or nothing original to the movement. This is not, however, to accuse him of plagiarism; for his ideas were part and parcel of the Augustan climate of opinion.

Despite frequent use of the name of John Locke (a device used by many deists), Bolingbroke was an unmitigated but curiously inconsistent rationalist. At one moment he asserts that the existence of Deity can and must be proved empirically, and at the next he asserts that only Right Reason can demonstrate the existence of Deity. He wrote *Reflections concerning Innate Moral Principles* to prove that compassion or benevolence is founded on reason alone. Unlike many of the deists, he was a metaphysical optimist, explaining away the evils of the universe and arguing that it is for man the best of all possible worlds despite the sufferings of individuals. He did not, however, believe that immortality and a future state of rewards and punishments can be proved by reason; and, although he accepted God as spirit, he was a materialist insofar as man is concerned.

He believed that there is no separation between soul and body and that at death man is annihilated; even in life, there is no communication between divine spirit and human matter.

Bolingbroke's concept of Natural Religion was essentially the same as the Common Notions of Lord Herbert of Cherbury. Yet with all his insistence on a priori reason, he lamented time and again that reason is fallible and must be corrected by a return to the primitive religions, particularly those of China and Egypt. Like all the deists, he was contemptuous of priestcraft and, despite his rationalism, of metaphysics. His criticism of Christian revelation is much like Matthew Tindal's, and the insinuation is that any revelation that is not universal is unnecessary.

In sum, Bolingbroke was more the orator than the philosopher. There is, however, considerable truth in his statement that "There is no reason … to banish eloquence out of philosophy; and truth and reason are no enemies to the purity, nor to the ornaments of language." He considered Plato, Nicolas Malebranche, and George Berkeley as poets, not philosophers, and his own best defense is the eloquence he admired.

See also Annet, Peter; Arnold, Matthew; Berkeley, George; Burke, Edmund; Deism; Herbert of Cherbury; Hume, David; Johnson, Samuel; Locke, John; Malebranche, Nicolas; Plato; Pope, Alexander; Tindal, Matthew; Voltaire, François-Marie Arouet de.

Bibliography

PRIMARY WORKS

A Collection of Political Tracts, 2nd ed. London: B. Francklin, 1748.

Letters on the Study and Use of History. London: T. Cadell, 1770.

The Works of Lord Bolingbroke, 4 vols. New York: A. M. Kelley, 1967.

Political Writings. Edited by David Armitage. Cambridge, U.K.: Cambridge University Press, 1997.

SECONDARY WORKS

Dickinson, H. T. *Bolingbroke.* London: Constable, 1970.

Harkness, Sir Douglas. *Bolingbroke: The Man and His Career.* London: Staples Press, 1957.

James, D. G. *The Life of Reason: Hobbes, Locke, Bolingbroke.* London: Longmans, Green, 1949.

Kramnick, Isaac. *Bolingbroke and His Circle: The Politics of Nostalgia in the Age of Walpole.* Cambridge, MA: Harvard University Press, 1968.

Mansfield, Harvey C., Jr. *Statesmanship and Party Government: A Study of Burke and Bolingbroke.* Chicago: University of Chicago Press, 1965.

McMerrill, Walter. *From Statesman to Philosopher: A Study in Bolingbroke's Deism.* New York: Philosophical Library, 1949.

Nadel, George H. "New Light on Bolingbroke's *Letters on History.*" *Journal of the History of Ideas* 23 (1962): 550–557.

Petrie, Sir Charles. *Bolingbroke.* London: Collins, 1937.

Pope, Alexander. *An Essay on Man.* Edited by Maynard Mack. London: Methuen, 1951.

Sichel, Walter. *Bolingbroke and His Times,* 2 vols. London: J. Nisbet, 1901.

Torrey, Norman L. "Chubb and Bolingbroke: Minor Influences." In *Voltaire and the English Deists.* New Haven, CT: Yale University Press, 1930.

Ernest Campbell Mossner (1967)

Bibliography updated by Philip Reed (2005)

BOLTZMANN, LUDWIG

(1844–1906)

Ludwig Boltzmann was born in Vienna, where he received his education. Boltzmann's major contribution to physics and, indirectly, to philosophy, was his profound work in the theory that grounded the phenomenological theory of heat, temperature, and the transformations of internal energy at the macroscopic level—that is to say thermodynamics—in the theoretical description of the underlying mechanical behavior of the basic constituents of a system, such as the molecules of a gas. Boltzmann also contributed directly to the ongoing philosophical discussions about the nature of scientific theories as a member of the group of outstanding physicist-philosophers concerned with such issues in the latter half of the nineteenth century, a group including Pierre Duhem, Ernst Mach, Wilhelm Ostwald, and Heinrich Hertz. During his career he held chairs at Graz, Munich, and Vienna.

After a long career as distinguished researcher and teacher whose influence through popularizing works extended beyond the narrow confines of academic scientists, Boltzmann tragically fell into a terminal depression ending in his suicide.

PHILOSOPHY OF SCIENCE

It would probably be a mistake to seek for a single, coherent, and fully developed account of the nature of scientific theories in Boltzmann's work. One must extract his views from a large number of short discussions, marginal remarks, and views expressed in correspondence with his colleagues. Nonetheless, certain themes are constant and clear and one can gain some understanding of what Boltzmann was after when one considers the scientific and philosophical context in which his remarks on the nature of theories were made.

Boltzmann's central scientific work posits that a macroscopic piece of matter, such as the volume of gas in a box, is composed of innumerable components—the molecules of the gas—too small to be observed in any direct manner. Following a long development from John Bernoulli, John Herepath, John Waterston, August Krönig, and Rudolf Clausius, and working in parallel with James Clerk Maxwell, Botzmann developed the kinetic theory of heat in which the dynamics of molecules moving more or less independently of one another—except for collisions and short-range interactions with one another and with the walls of a confining box—was used to explain the well-known laws of macroscopic thermodynamics.

It is important to understand just how indirect the evidence was for the genuine existence of molecules at this time. Their existence had been hypothesized in a resurrection of ancient atomic theory by chemists such as John Dalton to explain the combining laws of weight and volume in chemistry. The partial success of kinetic theory also provided indirect evidence of their existence. But the kinds of rich and more direct evidence available now for this particulate view of matter was then nonexistent.

A kind of radical empiricism was popular among the physicist-philosophers with whom Boltzmann associated. Duhem, Mach, and Ostwald shared the view that the aim of science was the production of simple and elegant lawlike regularities among the observable features of matter. They also shared deep skepticism toward any science that hypothesized unobservable entities as real explanatory components of the world. This skepticism included a negative attitude toward any theory positing "unobservable" molecules or atoms. Naturally such a position would be uncongenial to Boltzmann.

Boltzmann sought a view about theories that would legitimate inference to the existence of molecules, but that would not fall prey to empiricist skepticism about any scientific belief that rests upon "mere hypothesis" and that leaps beyond the observable features of the world to the postulation of unobservable entities and properties. Boltzmann's position seems close to that adopted by Hertz.

Theoretical beliefs do, indeed, rest upon hypotheses. New concepts for describing the world arise out of the scientist's imagination and are not all presented to one's direct sensory experience. There is no certainty in theoretical beliefs; they are certainly not derivable by any a priori reasoning, nor can they be established by "induction" from experience. They are hypotheses, guesses, invoked by humans to explain the observable phenomena. Such explanations consist in deductions of the observable phenomena from the hypothesized theory.

Only theories built on such hypotheses and invoking the unobservable will provide truly useful explanations in science. There is no hope of reconstructing science as a set of regularities that range only over the directly observable features of the world. But one must always remember that such hypothesized theories are merely pictures (*Bilder*) constructed by humans to fit the observable order into a coherent, deductive scheme. And one must always contemplate the possibility that alternative

schemes—alternative pictures—may be available. These may present a different picture of the unobservable world, but insofar as they are as empirically adequate as the theories people have adopted, they are equally satisfactory from a scientific point of view.

That the deepest theories rest upon idealization is another reason—in addition to the belief in these theories resting only upon hypothesis—for Boltzmann to retreat from a fully realist position with regard to fundamental physical theories.

Boltzmann's views may perhaps be best understood as a kind of instrumentalism and pragmatism with regard to theories, but with the insistence that physics could not do without such hypothesized theories in its attempts to account for the observable data. Although people must be wary of taking theoretical inferences too realistically, they must not put any of their hopes in a reconstructed physics that eschews the use of concepts and laws invoking the unobservable altogether.

THEORETICAL PHYSICS

Boltzmann's great contribution to physics was in kinetic theory and the beginnings of what later was called statistical mechanics. Here his work paralleled that of Maxwell. The two great scientists often came up with similar results independently, but each also found great inspiration in the work of the other.

Maxwell had found a velocity distribution for the molecules of a gas at equilibrium by a curious argument that utilized results from the theory of errors. Boltzmann generalized this distribution to allow for external forces acting on the molecules. In studying the problem of approach to equilibrium, Maxwell derived his so-called "transfer equations." Independently Boltzmann derived his kinetic equation of how the velocity distribution changes with molecular collisions, the famous Boltzmann Equation.

It was easy to show that the Maxwell-Boltzmann equilibrium distribution would be a stationary solution of this equation, hence appropriate for equilibrium that is an unchanging thermodynamic state. To show that this was the only possible such state, Boltzmann invented a quantity "H" as a function of the distribution. He showed that according to his equation this quantity must decrease unless the distribution is the standard equilibrium distribution. Hence the standard distribution is the only one possible for equilibrium.

Boltzmann developed a new method of thinking about the equilibrium as well. Divide a space in which points represent the position and momentum of a single molecule into boxes macroscopically small but in which one expects to find many molecular states. Boltzmann considered all of the ways in which molecules could be permuted among these boxes. He then showed that the combination (number of molecules in specific boxes) corresponding to the largest number of possible ways of permuting the molecules among the boxes (subject to conservation of total energy of the molecules) was that corresponding to the standard equilibrium distribution. One could then think of the numbers of permutations corresponding to a combination as the "probability" of that combination and argue that equilibrium was the overwhelmingly most probable state of the gas. And one could identify thermodynamic entropy as a measure of such probabilities.

Considerations of these results by Maxwell, Boltzmann, and such critics as Samuel Burbury, Edward Culverwell, and later Ernst Zermelo, led Boltzmann to a long process of reinterpretation of his work. Maxwell, considering the possibilities of mechanisms that would molecule-by-molecule subvert the approach to equilibrium (Maxwell's Demon) spoke of the kinetic equation as only describing probable changes in the gas. Considerations of the dynamical reversibility of the system at the molecular level, and of recurrence results for dynamical systems discovered by Henri Poincaré, also forced Boltzmann to modify the initial view of the equation as describing the inevitable behavior of a system.

Reflection revealed that in deriving his equation Boltzmann had used a time-asymmetric hypothesis about the numbers of collisions of molecules of specified kinds that would occur over a given time interval (the *Stosszahlansatz*). Both Maxwell and Boltzmann began to frequently invoke probabilistic language in their interpretations of their results. What were such "probabilities"? Boltzmann expressed the view that whereas Maxwell thought of them as frequencies with which states would occur in a large collection of similarly prepared systems, he, Boltzmann, thought of them as frequencies with which states would occur over long periods of time for an individual system.

Maxwell and Boltzmann also discovered another approach to calculating equilibrium values, in which these values could be calculated as average values of functions of the microscopic dynamical state of the system in question, where one used (1) a collection of all possible such microscopic states compatible with the macroscopic constraints, and (2) an easily discovered probability distribution over these states, to calculate the mean values.

Both Maxwell and Boltzmann introduced dynamical postulates (the Ergodic Hypotheses) to justify this method. The nature of this justification was made much clearer later by the work of Paul and Tatiana Ehrenfest. Although one can show the Ergodic Hypothesis in its Ehrenfest version false, this work led to later, sounder formulations of this approach by means of correct ergodic theorems and important work on the specific dynamics of idealized molecular systems.

Boltzmann, pushed by insightful criticism, realized that invoking probability by itself would not solve all his interpretive problems. His kinetic equation was time asymmetric, but the underlying dynamics was time symmetric. Because for each molecular motion going from nonequilibrium to equilibrium there was one going from equilibrium to nonequilibrium, it was hard to see how one could argue that the equation even characterized "most probable" evolutions of systems. (Although there are current interpretations of the Boltzmann equation that revert to this way of thinking.)

Boltzmann's later interpretation of the whole scheme resorted to cosmological considerations. One thinks of probabilities of states as given by Boltzmann's method. Equilibrium is then the overwhelmingly most probable state. Why is the world in nonequilibrium then? Boltzmann's assistant Dr. Schuetz suggested that maybe the cosmos is in equilibrium overall, but that humans live in a "small" part of it temporarily in a nonequilibrium fluctuational condition. Boltzmann added to this the "anthropic" argument that people must find themselves in such a region because equilibrium regions could not support life-forms. Finally Boltzmann added the argument that what is meant by the "future" direction of time is just the direction of time in which entropy is increasing in this local, nonequilibrium patch of the universe. He draws a deep analogy here with the fact that what people take as "down" is just the local spatial direction of the gravitational force. In equilibrium regions of the cosmos there would be two time directions, but neither could be thought of a "past" or as "future," just as in gravitation-free regions there is no "up" and no "down."

The Ehrenfests later provided a deep interpretation of the kinetic equation and its solutions consonant with this later Boltzmannian interpretation. The solutions to the equation describe neither the inevitable not the most probable behavior of a system, but rather the "concentration curve" that describes the state of most of the systems of a collective of systems started in common nonequilibrium at any later moment of time. But at different times different members of the original collection are making up this majority that is approaching equilibrium.

BOLTZMANN'S CONTINUED INFLUENCE

Boltzmann's methodological thoughts about theories remain provocative and worthy of reflection when one reflect's now on the still problematic status of foundational physical theories. His introduction of probabilistic reasoning into physics was seminal. His work on kinetic theory and statistical mechanics is a rich source of problems for the philosopher of physics interested in probabilistic explanation in physics and in the relationship between phenomenological macroscopic theories and their microscopic, atomistic underpinnings. Boltzmann's invocation of cosmology (still done in current theories of statistical mechanics but within an entirely different cosmological background) also opens up a wide range of important questions for methodologists concerned with how people can construct their fundamental physical explanations. And his views on the "direction of time" remain fundamental for anyone discussing the origin and nature of ideas of the asymmetric nature of past and future.

See also Philosophy of Statistical Mechanics.

Bibliography

Blackmore, John. *Ludwig Boltzmann His Later Life and Philosophy 1900–1906*. Dordrecht, Netherlands: Kluwer, 1995.

Broda, Englebert. *Ludwig Boltzmann: Man, Physicist, Philosopher*. Woodbridge, CT: Ox Bow Press, 1983.

Brush, Stephen. *The Kind of Motion We Call Heat.* Vol. 2: *Statistical Physics and Irreversible Processes*. Amsterdam: North-Holland, 1976.

Brush, Stephen. *Kinetic Theory.* Vol. 2: *Irreversible Processes*. Oxford: Pergamon Press, 1966.

Brush, Stephen. *Statistical Physics and the Atomic Theory of Matter, from Boyle and Newton to Landau and Onsager*. Princeton, NJ: Princeton University Press, 1983.

Ehrenfest, Paul and Tatiana. *The Conceptual Foundations of the Statistical Approach in Mechanics*. Ithaca, NY: Cornell University Press, 1959.

Sklar, Lawrence. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge, U.K.: Cambridge University Press, 1993.

Lawrence Sklar (2005)

BOLZANO, BERNARD

(1781–1848)

Bernard Bolzano, a philosopher, theologian, logician, and mathematician, was born in Prague, where his father, an Italian art dealer, had settled; his mother was a German merchant's daughter. Bolzano studied mathematics, philosophy, and theology in Prague and defended his doctor's thesis in mathematics in 1804; he was ordained a Roman Catholic priest the following year. Shortly thereafter he was appointed to a temporary professorship in the science of religion at Karlova University in Prague and two years later was given a newly established chair in this field. Some time later he was accused of religious and political heresy and was removed from his teaching position in December 1819. Bolzano spent much of his time thereafter with the family of his friend and benefactor, A. Hoffmann, at their estate in southern Bohemia. He had difficulty getting his later publications through the Metternich censorship. Some of his books were put on the Index, and many appeared only posthumously. Some manuscripts are yet to be published; the most important of these are in the National Museum and the University Library in Prague, others are in the Österreichische Nationalbibliothek in Vienna. In December 1848, Bolzano died of a respiratory disease from which he had suffered for most of his life.

MATHEMATICS

Bolzano's mathematical teachings were not quite understood by his contemporaries, and most of his deep insights into the foundations of mathematical analysis long remained unrecognized. A famous theorem in the early stages of a modern presentation of the calculus is known as the Bolzano-Weierstrass theorem, but another masterful anticipation (by more than forty years) of Karl Theodor Wilhelm Weierstrass's discovery that there exist functions that are everywhere continuous but nowhere differentiable remained buried in manuscripts until the 1920s. But perhaps more important than Bolzano's actual discoveries of new theorems was the meticulousness with which he endeavored to lay new foundations for the *Grössenlehre,* the science of quantity—which was how Bolzano, using a very broad interpretation of "quantity," designated mathematics. In particular, his insistence that no appeal to any intuition of space and time should be acknowledged for this purpose and that only "purely analytical" methods were to be recognized put him in opposition to the then current Kantian ways of thinking and back into the Leibnizian tradition.

Bolzano's most famous posthumously published work is *Paradoxien des Unendlichen* (F. Prihonsky, ed., Leipzig, 1851; translated by D. A. Steele as *The Paradoxes of the Infinite,* London, 1950), in which he anticipated certain basic ideas of set theory, developed only a generation later by Georg Cantor, who fully acknowledged his indebtedness to Bolzano in this respect. This anticipation should, however, not be overrated. Bolzano was not quite able to rid himself of all the prejudices of his time and was, therefore, unable to reach a clear and fruitful conception of equivalence between infinite sets.

ETHICS AND PHILOSOPHY OF RELIGION

Bolzano was, in his time, much more influential as a theologian and social moralist than as a mathematician. An advocate of the Bohemian Catholic enlightenment, he lectured on religion and moral philosophy with strong pacifistic and socialistic overtones. He used the pulpit to proclaim before hundreds of impressed students a kind of utopian socialism. In his sermons he tried to prove the essential equality of all human beings, attacked private property obtained without work, and exhorted his listeners to sacrifice everything in their struggle for human rights. These sermons served him as a preparation for what he regarded as his most important book, *Von dem besten Staate,* which he finished in 1837 but was unable to publish. It first appeared in Prague in 1932.

Bolzano's philosophy of religion is presented in the books *Athanasia oder Gründe für die Unsterblichkeit der Seele* (Sulzbach, 1827) and *Lehrbuch der Religionswissenschaft* (4 vols., Sulzbach, 1834), the latter being a revised version of his lectures at the Prague university. He tried to prove that Catholicism is in full harmony with common sense. To this end he either disregarded or interpreted allegorically all mystical elements of Catholicism.

Bolzano derived his utilitarian ethics from a "highest ethical principle": "Of all actions possible to you, choose always the one which, weighing all consequences, will most further the good of the totality, in all its parts" (*Lehrbuch der Religionswissenschaft,* Vol. I, Sec. 87). This reminds one, of course, of Jeremy Bentham. "The most important idea of mankind" Bolzano took to be the "essential" equality of all human beings, which he tried to prove from historical, rational, and ethical considerations.

LOGIC AND EPISTEMOLOGY

It is as logician, methodologist, and epistemologist that Bolzano, after a long period of neglect, regained philosophical attention in the twentieth century. Mainly in order to combat radical skepticism, he found it necessary to base his teachings in these fields on certain ontological conceptions. He was convinced that there exist truths-in-themselves (*Wahrheiten an sich*) prior to and independent of language and man. These truths he carefully distinguished from truths expressed in words and conceived truths. The set of truths-in-themselves is a subset of the set of propositions (in-themselves) (*Sätze an sich*), again to be distinguished from propositions expressed in words and conceived propositions. Propositions consist of terms (ideas-in-themselves, *Vorstellungen an sich*). These are likewise to be distinguished, on the one hand, from the words or word sequences by which they are denoted and, on the other, from subjective ideas that occur in our mind. Although linguistic entities and conceived entities exist concretely, terms, propositions, and truths do not. Terms were equally carefully distinguished from their objects, whether or not these objects themselves existed concretely. Though Bolzano was a Platonist (in the modern sense), his ontology was rather remote from that of Plato or, for that matter, from that of Immanuel Kant, in spite of the common *an sich* terminology.

Beyond these negative determinations, Bolzano had little positive to say on the ontological status of terms and propositions except that they are the matter (*Stoff*) or sense (*Sinn*) of their correlates in language and thought.

Terms can be either simple or complex and either empty (*gegenstandslos*) or nonempty (*gegenständlich*); if nonempty, they are either singular or general. Examples of empty terms are –1, 0, Nothing, Round Square, Green Virtue, and Golden Mountain; absolutely simple terms are Not, Some, Have, Be, and Ought, but Bolzano was uncertain about others. Simple, singular terms he called intuitions (*Anschauungen*).

Propositions are composed of terms and are perhaps best regarded as ordered sequences of terms, while the content (*Inhalt*) of a proposition is the (unordered) set of the simple terms out of which the terms constituting the proposition are composed. The content of a complex term is similarly defined. The terms 3^5 and 5^3 are different, though they have the same content. The terms 2^4 and 4^2 are different, though they have not only the same content but even the same object. With this conception of content, the traditional doctrine of the reciprocity between the extension of a term (the set of objects falling under it) and the content of a term can easily be seen to be invalid.

Among Bolzano's many idiosyncratic convictions, perhaps the most interesting, but also the most strange to the modern mind, was his belief that each branch of science has a unique, strictly scientific presentation, which for him meant not only a unique finite axiom system (a belief he shared with many) but also an essentially unique entailment (*Abfolge*) of each theorem of this science by the axioms, a belief which might well be unique to Bolzano.

This relationship of entailment, as presented by Bolzano, is very peculiar and obscure. Bolzano was never quite sure that he understood it himself, though he was convinced that there objectively must exist some such relationship, that each science must have its basic truths (*Grundwahrheiten*) to which all other truths of that science stand in the peculiar relation of consequence (*Folge*) to ground (*Grund*). Bolzano was constantly struggling to differentiate this relation of entailment from the relation of derivability (*Ableitbarkeit*), which was the basic relation of his logic. Though he did not succeed in putting his theory of entailment into consistent and fruitful shape—and could not possibly have done so, in view of the chimerical character of his goal—his acumen, mastery of the contemporary logical and methodological literature, intellectual honesty, and lifelong self-criticism more than made up for his numerous shortcomings. Bolzano remains a towering figure in the epistemology, logic, and methodology of the first half of the nineteenth century.

See also Bentham, Jeremy; Cantor, Georg; Kant, Immanuel; Logic, History of; Propositions, Judgments, Sentences, and Statements.

Bibliography

ADDITIONAL WORKS BY BOLZANO

Bolzano's masterwork is his *Wissenschaftslehre,* 4 vols. (Sulzbach, 1837; edited by Wolfgang Schultz, Leipzig: F. Meiner, 1929–1931). *Grundlegung der Logik* (Hamburg, 1964) is a very useful selection by Friedrich Kambartel from the first two volumes of the *Wissenschaftslehre,* with summaries of omitted portions, an excellent introduction, and a good index.

WORKS ON BOLZANO

Bolzano's philosophical work was virtually disregarded until Edmund Husserl called attention to it at the start of the twentieth century. Hugo Bergmann's monograph, *Das philosophische Werk Bernard Bolzanos* (Halle: M. Niemeyer, 1909), increased the revived interest in Bolzano's ideas. Heinrich Scholz's articles, especially "Die Wissenschaftslehre Bolzanos," in *Abhandlungen des Fries'schen Schule,* n.s, 6

(1937): 399–472, reprinted in *Mathesis Universalis,* pp. 219–267 (Basel: B. Schwabe, 1961), presented Bolzano's contributions to logic, semantics, and the methodology of the deductive sciences in a modernized form. The best recent study in English of Bolzano as a logician is J. Berg's *Bolzano's Logic* (Stockholm: Almqvist and Wiksell, 1962). D. A. Steele's historical introduction to his translation of Bolzano's *Paradoxien des Unendlichen* is useful. Among other secondary works the most important are Eduard Winter's *Bernard Bolzano und sein Kreis* (Leipzig: J. Hegner, 1933), Günter Buhl's *Ableitbarkeit und Abfolge in der Wissenschaftstheorie Bolzanos* (Cologne: Cologne University Press, 1961), and (from a Marxist viewpoint) A. Kolman's *Bernard Bolzano* (in Russian, Moscow, 1955; in Czech, Prague, 1957; and in German, Berlin, 1963).

Yehoshua Bar-Hillel (1967)

BONALD, LOUIS GABRIEL AMBROISE, VICOMTE DE

(1754–1840)

Louis Gabriel Ambroise, Vicomte de Bonald, the French publicist and philosopher, was born in the château of Le Monna, near Millau (Aveyron). He emigrated in 1791, during the Revolution, to Heidelberg, moving later to Constance, and joined the circle of royalist writers who in 1796 published a number of books attacking the Revolutionary Party and defending the monarchy. His own contribution to the propaganda was his famous *Théorie du pouvoir politique et religieux* (3 vols., Constance, 1796), the first of a long series of volumes expressing the ultramontane position, the political supremacy of the papacy, absolute monarchy, and traditionalism.

The basic premise of Bonald, as far as his philosophy was concerned, was the identity of thought and language. Against the usual eighteenth-century idea that language was a human invention, he revived Jean-Jacques Rousseau's argument that since an invention requires thought and thought is internal speech, language could not have been invented. Consequently, he argued, it must have been put into the soul of man at creation. By means of certain philological investigations, Bonald was able to convince himself that there was a basic identity in all languages, as indeed there is in the Indo-European.

But language is a social, not an individual, phenomenon. It binds individuals together into groups and expresses an interpersonal set of ideas. These ideas are tradition. The unity of tradition may be disrupted, as it was during the Revolution, but nevertheless humankind will have to return to it if they have any hope of regaining social health. When this return occurs, people will cooperate in a single political system and a single set of religious beliefs. The former will be absolute monarchy, the latter Roman Catholicism, both having single and omnicompetent heads. Thus, just as the universe is created and governed by one God, so both the church and state must preserve administrative unity. But since the church is the direct channel of communication between God and his creatures, the state and its subjects must be governed in moral affairs by the church.

The ultramontanism of Bonald was as extreme as logically possible. He maintained that the arts, for instance, flourished only in an absolute monarchy, and hence saw nothing to praise in Greek art. In fact, he had nothing good to say about anything Greek, since Greece was given to democracy, though he made an exception of the Spartans. He was opposed to the legalization of divorce and to equal rights for women. He accepted capital punishment, since God would see to it that the innocent would not suffer in the afterlife. He supported general censorship and denounced freedom of the press. And since he was a man of Stoic morals, he did not worry much about human dissatisfaction or unhappiness.

Bonald was a philosopher who never changed his views. In each of his numerous works he repeated the same fundamental theses. His influence was restricted to men of the extreme right, in spite of his ingenuity in argument and logical rigor. His ideas survived in France in *L'action française* and even in the nonpolitical writings of Charles Maurras, through whom they passed in diluted form to T. S. Eliot.

See also Eliot, Thomas Stearns; Language and Thought; Rousseau, Jean-Jacques; Traditionalism.

Bibliography

The collected works of Bonald appeared as *Oeuvres complètes,* edited by Abbé Migne (Paris, 1859).

For works on Bonald, see G. Boas, *French Philosophies of the Romantic Period* (Baltimore: Johns Hopkins Press, 1925), ch. 3; Harold Laski, *Authority in the Modern State* (New Haven, CT: Yale University Press, 1919), which follows Moulinié; H. Moulinié, *De Bonald* (Paris: Alcan, 1916); C. Sainte-Beuve, *Causeries du lundi* (Paris: Garnier, 1851–1862), Vol. IV, 426; Émile Faguet, *Politiques et moralistes du dix-neuvième siècle,* lère série (Paris, 1890).

OTHER RECOMMENDED TITLES

Dejnozka, Jan. "Origins of the Private Language Argument." *Dialogos* 30 (66) (1995): 59–78.

Reedy, W. Jay. "Language, Counter-Revolution and the 'Two Cultures': Louis de Bonald's Traditionalist Scientism." *Journal of the History of Ideas* 44 (1983): 579–598.

Reedy, W. Jay. "The Traditionalist Critique of Individualism in Post-Revolutionary France: The Case of Louis de Bonald." *History of Political Thought* 16 (1) (1995): 49–75.

George Boas (1967)
Bibliography updated by Tamra Frei (2005)

BONATELLI, FRANCESCO

(1830–1911)

Francesco Bonatelli, an Italian spiritualist philosopher, was born in Iseo, Brescia. He studied at the University of Vienna and taught philosophy at the universities of Bologna (1861–1867) and Padua (1867–1911). Bonatelli belonged to the tradition of Catholic spiritualism. He was one of the principal editors of *Filosofia delle scuole italiane,* a review founded in 1870 by Terenzio Mamiani to defend a Platonizing position, but he resigned in 1874 when the Platonist Giovanni Maria Bertini published criticisms of Catholicism that Bonatelli considered too bold. Bonatelli introduced the analytic method of German psychological research into Italy.

Bonatelli attempted to distinguish consistently between the unity of the ego and the multiplicity of psychic events. In his first work, *Pensiero e conoscenza* (Thought and consciousness; Bologna, 1864), Bonatelli distinguished two ways of life for the soul, one that is subject to the laws of fate and another that, although it recognizes these laws, is able to rise above them and use them as tools.

The conscious subject can be aware of other things only if it is capable at one and the same time of being modified and of remaining identical with itself, or inalterable. The solution of this apparent contradiction might lie in distinguishing between consciousness, understood as thought or pure mentality, and sensibility. In his most important work, *La coscienza e il meccanismo interiore* (Consciousness and the internal mechanism; Padua, 1872), Bonatelli insisted that consciousness neither is changed by the object nor changes it. The act of consciousness detaches the psychic event from its matrix in reality and thinks its possible essence or its "possibility or quiddity or whatever you wish to call it." Bonatelli investigated both consciousness itself and the relation between the psychic mechanism external to consciousness and consciousness, between the existing object and the object thought in its "quiddity."

He regarded consciousness as thought turned back upon itself and almost creating itself, but also as freely accepting the "yoke of logic." If consciousness were not of this nature, it would be reduced to a "logical machine," whereas it is free reflection on itself, grasping itself by directing itself toward objects. However, although the distinctive essence of consciousness is its infinite turning back upon itself (*la riflexione infinita degli atti,* "the infinite reflection of acts"), this reflection is not an infinite succession in which consciousness would lose itself in an endless postponement but rather a completed penetration of self, the fullness and richness of the activity of thought.

See also Consciousness.

Bibliography

ADDITIONAL WORKS BY BONATELLI

La filosofia dell'inconscio di Edoardo von Hartmann. Rome, 1876.

Elementi di psicologia e logica. Padua, 1892.

WORKS ON BONATELLI

Alliney, Giulio. *Francesco Bonatelli.* Brescia, 1947.

Gentile, Giovanni. *Le origini della filosofia contemporanea in Italia.* Messina: G. Principato, 1917. Vol. I, pp. 220–227.

Scatturin, Umberto. "Francesco Bonatelli." *Filosofia* 3 (1952): 433–439.

Varisco, Bernardino. *Francesco Bonatelli.* Chieri, 1912.

See also "In onore di Francesco Bonatelli," by various authors. *La cultura filosofica* 4 (2) (1910). Contains a bibliography.

Eugenio Garin (1967)
Translated by Tessa Byck

BONAVENTURE, ST.

(c. 1217–1274)

St. Bonaventure, the Italian Scholastic philosopher, was known as the Seraphic Doctor. Bonaventure, whose real name was John of Fidanza, was born in Bagnorea, in Tuscany. After obtaining a master of arts degree at Paris, Bonaventure joined the Franciscan friars (probably in 1243) and studied theology under their masters, Alexander of Hales and John of La Rochelle. After their deaths in 1245, he continued his studies under Eudes Rigaud and William of Meliton. He also came under the influence of the Dominican Guerric of Saint-Quentin and the secular master Guiard of Laon. In 1248 as a bachelor of Scripture he began lecturing on the Gospel of St. Luke and then on other books of Scripture (not all of these commentaries have survived). His monumental "Commentary on the *Sentences* of Peter Lombard," perhaps the most perfect example of this form of medieval literature, was composed between 1250 and 1252.

In 1253 he was licensed by the chancellor of the University of Paris and functioned as regent master of theology until 1257. During this time he composed four sets of *Quaestiones Disputatae,* of which the *De Scientia Christi* (On Christ's knowledge) is important for his theory of illumination; *De Mysterio Trinitatis* (On the mystery of the Trinity) contains the best exposition of his proofs of God's existence; and *De Caritate et de Novissimis* (On charity and the last things) contains sections taken over literally by Thomas Aquinas.

Bonaventure's formal reception into the masters' guild was delayed until October 1257 by the controversy between the mendicant friars and the secular masters. By that time, however, he was no longer actively teaching; in February 1257 he had been elected minster general of the Franciscan order and had resigned his chair at the university to devote himself to the administration of that post. Although often absent on business for the order or church, he continued to make Paris his general headquarters and was largely responsible for the friars' being so active in academic pursuits. He himself preached frequently at the university, touching on many of the religious and philosophical troubles that disturbed faculty and students.

It was during these years that he composed the *Breviloquium* (1257), or brief compendium of speculative theology, which was a departure from the usual scholastic method of presentation; *De Reductione Artium ad Theologiam* (On the reduction of the arts to theology), whose exact date of composition is unknown; and *Itinerarium Mentis in Deum* (The journey of the mind to god; 1259). All of these are important for understanding his general system of thought and the particular role of philosophy in it. Even more important in this connection are the three sets of *Collationes*—a series of informal evening conferences given during Lent to the faculty members and students in the Paris friary—including *De Decem Praeceptis* (On the ten commandments; 1267), *De Septem Donis Spiritus Sancti* (On the seven gifts of the Holy Spirit; 1268), and *In Hexaemeron Sive Illuminationes Ecclesiae* (On the six Days of creation or enlightenments of the church; 1273). All of these reflect the Averroistic tendencies in the arts faculty and Bonaventure's reaction to them. The last of these *Collationes* was left unfinished when Bonaventure was called from Paris and made cardinal bishop of Albano by Pope Gregory X, with whom he worked in organizing the Second Ecumenical Council of Lyons. He died shortly before the council closed and was buried there in the presence of the pope.

SPIRIT OF BONAVENTUR'S PHILOSOPHY

Bonaventure's fame rests primarily on his reputation as a theologian rather than as a philosopher. In both Dante Alighieri's *Paradiso* and Raphael's "Disputà" he appears as the equal of St. Thomas, and in the field of mystical theology he has been considered without peer. It is more difficult, however, to isolate the philosophical components of his system. This is partly due to the fact that all Bonaventure's extant works postdate his entrance into the Franciscan order and the beginning of his career as a theologian and ascetical writer. The chief reason, however, for the prevalence of theological interests in all of his writings was his understandable reaction against the rationalism rampant in the arts faculty at Paris that threatened the very raison d'être of speculative theology and led to the condemnations of 1270 and 1277 by Stephen Tempier, bishop of Paris. Among the 219 items listed as theological errors in the second of these condemnations, for example, are such statements as

(a) The most exalted of all vocations is that of the philosopher.

(b) There is no subject he is not competent to discuss and settle.

(c) One gains nothing in the way of knowledge by knowing theology.

(d) Only the philosophers deserve to be called wise; the speech of the theologian is founded on fables.

In the face of such views, it is understandable why Bonaventure, who believed in the validity of Christian revelation, should have stressed the inability of philosophers in general and of Aristotle in particular to learn the full truth about man's existential situation. Conversely, Bonaventure tried to show the continuity between the aims of philosophy and those of theology. He maintained that philosophy has a genuine, albeit limited, autonomy; the knowledge it yields is a stage in the overall ascent of the human mind to true wisdom, the culmination of which in this life is found in quasi-experiential knowledge of God, achieved by such mystics as Francis of Assisi.

Part of the great literary charm of Bonaventure's style is his ability to play upon words. Throughout his later works, particularly his sermons and *Collationes,* he continually gives a deliberately theological twist to technical philosophic terms, with the result that he has frequently been unjustly accused of confusing theology with philosophy either in principle or in practice. The truth of the matter is that while he was eminently able to conduct a purely philosophical discussion and often did so in his

university lectures, he preferred to limit himself to particular topics. He never formed a complete system from his philosophical analyses, but he put them into the service of his overall theological synthesis.

BONAVENTURE'S METAPHYSICS

Bonaventure's linguistic sophistication and his idea of the continuity between philosophy and theology are perhaps best represented in his discussion of metaphysics in the *In Hexaemeron.* Christ, the Son of God, not Aristotle, is the "metaphysician" par excellence.

> As the Son said: "I came forth from the Father and have come into the world; again I leave the world and go to the Father" [John 16:28], so anyone may say: "Lord, I came forth from you, the All High; I go to you, the All High, and by means of you, the All High." Here is the metaphysical medium leading us back. And this is the whole of our metaphysics: it concerns emanation, exemplarity, and consummation [that is, being illumined by spiritual rays and led back to the All High]. It is in this way you become a true metaphysician. (*Collatio* I, No. 17; in *Opera,* Vol. V, p. 332)

EMANATION. Bonaventure uses the term *emanation* to designate the general theory of how creation proceeds from God. With its Plotinian overtones, however, "emanation" suggested more specifically the thesis of al-Fārābī, Avicenna, and Averroes that all creatures, by an inevitable and eternal process, spring from the creative mind of God through a chain of intermediary causes of continually diminishing perfection. This thesis was designed to reconcile Aristotle's eternal world with the creation concept of the Qurʾan. Bonaventure, however, wished to reconcile "emanation" with Christian theology. His counterthesis is summarized in the *Breviloquium*: "The whole of the cosmic machine was produced in time and from nothing, by one principle only who is supreme and whose power, though immense, still arranges all according to a certain weight, number and measure" (Book II, Part 1, in *Opera,* Vol. V, p. 219). It is to be noted that he rejects the concepts of the eternity of the world, of the eternity of matter, of a dual principle of good and evil, and of the existence of intermediary causes.

His description of the supreme principle implies that a perfect power must be free to create varying degrees of perfection, in contrast with the Arab belief that direct creation by a perfect power could only result in perfect effects. Also, the use of Augustine's triad of weight, number, and measure suggests the seal of the Blessed Trinity stamped on every creature. This becomes clearer if we consider the next and most characteristic feature of Bonaventure's metaphysics.

EXEMPLARISM. Emanation concerns natural philosophy as much as metaphysics. God, as final cause and ultimate goal of man's quest for happiness, is the concern of the moral philosopher as well as the metaphysician. But only the metaphysician can understand God as exemplar cause. And it is in analyzing this aspect of the science of causes and first principles that man is most truly a metaphysician.

Though this metaphysical pursuit begins with reason, it can be successfully terminated only by a person with faith. Comparing the two greatest pagan philosophers, Aristotle and Plato, Bonaventure maintained that Plato, the master of wisdom, erred in looking only upward to the realm of eternal values, of the immutable ideas, while Aristotle, the master of natural science, looked only earthward to the everyday sensible world that Plato neglected. But Aristotle's was the greater sin, for in rejecting the Platonic ideas in toto, he closed the door to a full understanding of the universe in terms of its causes. Bonaventure saw Augustine as the model of Christian wisdom because he combined the science of Aristotle with Plato's wisdom (*Christus Unus Omnium Magister,* Nos. 18–19, in *Opera,* Vol. V, p. 572). As a Christian he could complete what Plato could only begin. Not only did he demonstrate that Plato's archetypal Ideas are the exemplar causes or models that God used in creating the universe, a point that a philosopher alone could establish, but he also showed further that these Ideas are associated in a special way with the second person of the Trinity, an insight only divine revelation could help one discover. Bonaventure, following Augustine, explained that since the Father begets the Son by an eternal act of self-knowledge, the Son may also be called the wisdom of the Father and expresses in his person all of God's creative possibilities. As such, the Son is the Word or Logos adumbrated in the writings of the philosophers but fully revealed only at the beginning of the Gospel of John, where he appears as the one through whom all things are made (that is, as exemplar cause) and who "enlightens every man who comes into the world" (an allusion to Augustine's theory that only some illumination by divine ideas can account for man's knowing immutable truths). "From his [magisterial] chair in heaven Christ teaches us interiorly," wrote Bonaventure. "If as the Philosopher [Aristotle] says, the knowable qua knowable is eternal, nothing can be known except through that Truth which is

unshaken, immutable and without limit" (*In Hexaemeron, Collatio* I, No. 13; in *Opera,* Vol. V, p. 331).

Averroes had written of Aristotle: "I believe this man to be nature's model, the exemplar which nature found to reveal the ultimate in human perfection" (*De Anima* III, 2). Bonaventure maintained that Christ, not Aristotle, is God's model for humanity. The Word is not only God but also a perfect man. He gives us "the power of becoming the sons of God," and he is the "one master of all the sciences" (*Sermo IV*; in *Opera,* Vol. V, p. 567); to know him fully is to know all that can be known.

Bonaventure held that Plato's theory of Ideas was a first philosophical approximation to this theological insight, and Aristotle's rejection of this view led to his errors about God. For if God lacked the exemplar ideas, he would know only himself and nothing of the world. He would be, as Aristotle claimed, related to the world only as final cause and not as creator. Moreover, in Aristotle's world, since chance clearly does not explain the cyclic changes of the cosmos, the universe must be ruled by determinism, as the Arabic commentators claim. But then man would no longer be a responsible agent; he would deserve neither reward nor punishment, and divine providence would be a myth.

With the recognition of exemplarism, on the other hand, the whole of creation takes on a sacramental character— that is, it becomes a material means of bringing the soul to God. Nature becomes the "mirror of God," reflecting his perfections in varying degrees. Although we see only a shadowy likeness (*umbra*) or trace (*vestigium*) of the creator in inorganic substances and the lower forms of life, the soul of man is God's image (*imago*) and the angel his similitude (*similitudo*).

The recognition of God in nature begins in philosophy, but it is continued and perfected in theology. In *De Mysterio Trinitatis* Bonaventure argued that philosophers know that secondary beings imply a first; dependent beings imply an independent being; contingent things imply some necessary being; the relative implies an absolute; the imperfect, something perfect; Plato's participated beings imply one unparticipated being; if there are potential beings, then pure act must also exist; composite things imply the existence of something simple; the changeable can only coexist with the unchangeable. Pagan philosophers, knowing that these ten self-evident conditionals have their antecedents verified in the corporeal world, learned much about God (*De Mysterio Trinitatis* I, 1; in *Opera,* Vol. V, pp. 46–47).

More can be learned, however, by the soul reflecting upon itself. In his other works Bonaventure went on to suggest that the soul, possessed of memory, intelligence, and will, is an image of God, not only mirroring his spiritual nature but adumbrating the Trinity itself. Memory, which creates its own thought objects, resembles the Father who begets the Son or Logos (intelligence) as an intellectual reflection of himself, and the two through their mutual love (will—the active principle of "spiration") breathe forth the Holy Spirit. But although a philosopher can discover a spiritual God as the ultimate object of the soul's search for truth and happiness, only a man of faith like Augustine can find the Trinity manifest throughout creation.

CONSUMMATION OR ENLIGHTENED RETURN. The third aspect of Bonaventure's metaphysics concerns a creature's fulfillment of its destiny by returning to God. This return (called technically a *reductio*) in the case of the lower creation is achieved in and through man (who praises God for and through subhuman creation). Man's return is made possible in turn by Christ. For man returns to God by living an upright life—that is, by being rightly aligned with God—and this can be accomplished only through the grace of Christ. Man's mind is right (*rectus*) when it has found truth, and above all, eternal truth. His will is right when it loves what is really good, his exercise of power is right when it is a continuation of God's ruling power. Through original sin or the Fall, man lost this triple righteousness. His intellect, lured by vain curiosity, has enmeshed itself in interminable doubts and futile controversies; his will is ruled by greed and concupiscence; in his exercise of power he seeks autonomy. But although man lost the state of original justice, he still hungers for it. This longing for the infinite good is revealed in his ceaseless quest for pleasures. Through faith and love (grace), man can find his way back.

Since knowledge is involved at every stage of the return, *reductio* is also a quest for wisdom and hence, in an extended theological sense, it is metaphysical. It is an enlightened return, because every branch of learning is a gift from above, from the "Father of lights" (Epistle of St. James, 1. 17), and can be put into the service of theology (this is the theme of Bonaventure's *De Reductione Artium*). Although man's return begins with the natural light of reason reflecting first on the external world and then turning inward in an analysis of the soul, it is perfected initially by a natural illumination of the divine ideas and then by varying additional degrees of supernatural illumination which culminate in the experiential cognition of God through mystical union (the theme of

the *Itinerarium*). This experience is not the same as the clear vision of the blessed in heaven but is the "learned ignorance" referred to by the mystical writers—a union of the soul with God in darkness, granted to saints like Francis before death.

OTHER DOCTRINES

The elements of Bonaventure's philosophy are woven into his religiously oriented system. Like all the Parisian thinkers of this period, Bonaventure developed a basically Aristotelian philosophy, but he included a larger admixture of Neoplatonic and Augustinian elements than we find in St. Thomas, for instance, who studied Aristotle somewhat later and more thoroughly under Albert the Great.

THEORY OF KNOWLEDGE. Bonaventure believed that the mind has no innate ideas, not even in the sense postulated by the authors of the *Summa Theologica* (ascribed to Alexander of Hales), who argued that ideas are latent in the agent intellect but are actually acquired only when the light of the agent intellect illumines the possible intellect. Bonaventure rejected this, holding with Aristotle that the mind at birth is a tabula rasa. It needs sensory stimulation before it can acquire any notions about the external world of objects. However, Bonaventure did use the Augustinian theory of illumination to explain how the mind passes judgment on sensible things in terms of their values. For when the mind judges something to be, for example, good or beautiful, there must be an implicit awareness of what beauty and goodness are in themselves; and this requires that the human mind have some knowledge of the divine ideas. Obviously this is not a clear or intuitive knowledge of God such as the angels or the blessed in paradise possess. Yet just as one can see by sunlight without looking into the sun itself, so one can have knowledge of the divine ideas. At the same time, Bonaventure rejected the interpretation (also found in the *Summa* of Alexander) that we attain these ideas only in terms of the residual effects of the divine action—effects which remain in the soul like habitual or buried memories. Bonaventure claimed that in some mysterious way (which he called *contuition* but which he never fully explained), when we know a created object, our mind is simultaneously enlightened so that it is moved to judge correctly about the object and is hence in accord with God's own mind on the subject.

Although Bonaventure agreed with Aristotle that our knowledge of the external world is sense-dependent, he did not fully subscribe to Aristotle's principle that "nothing is in the intellect that was not first in the senses." He held that the intellect can turn inward, reflecting on the soul and its tendencies. In analyzing the precise nature of the object of these tendencies, the mind discovers God and itself as his image. The reasoning process involved is neither deductive nor inductive in the usual meaning of these terms, but is called technically a "reduction" and seems to resemble in some respects the "abduction" of Charles S. Peirce. Reasoning proceeds by progressively deepening insights into what the desire for truth and perfect happiness involve. If the reduction remains imperfect and does not go on to completion, God is not discovered and one may err about his nature or even his existence. Although at times Bonaventure, following the authority of John of Damascus, Ancius Manlius Severinus Boethius, or Augustine, spoke of the existence of God as a truth implanted by nature in the human mind, he meant this to be interpreted as referring immediately to man's natural desire for knowledge, truth, happiness, or goodness—all of which need explication before man realizes they have God as their ultimate object (*De Mysterio Trinitatis,* I, 1; in *Opera,* Vol. V, p. 49).

COSMOLOGY. In his analysis of material creation, Bonaventure introduced extraneous elements into Aristotle's theory of matter and form. Thus, for instance, he adopted Avicebron's theory of the hylomorphic composition of spiritual as well as corporeal creatures. The argument here is that since creatures have some measure of potentiality (only God is pure actuality), they must have some kind of matter, for according to Aristotle matter is the principle and source of potentiality. This spiritual matter, found both in the angel and in the human soul, is never separable from its spiritual form; hence, such spiritual substances are not subject to change—they cannot die or disintegrate like terrestrial bodies, nor can they be perfected by a hierarchy of forms, as can corporeal matter.

In *Breviloquium,* Book II, Bonaventure, in explaining the visible universe, made use of the theories of light developed by Robert Grosseteste and the Oxford school. He distinguished light (*lux*), luminosity (*lumen*), and color. The first is the most basic of substantial forms; it enables both terrestrial and celestial bodies to subsist and is the root source of whatever internal dynamism they possess. *Lumen* is the invisible radiation which has its origin especially in celestial bodies like the sun but exists in the intervening transparent medium. It is described by Bonaventure as being both an active power (*virtus activa*) and something substantial in itself but only accidentally related to the transmitting medium through which it

flows continually and instantaneously by a self-generative process called multiplication. Being neither an accidental nor a substantial form properly speaking, it is not educed from the potentialities of matter as are other corporeal forms, with the exception of *lux.* Yet it requires some material medium or body and coexists with such without changing it substantially. Not only does it penetrate the bowels of the earth, where it governs the formation of minerals, but in virtue of its purity and similarity to the spiritual, this substantial radiation disposes bodies to receive the life form and acts as a sort of intermediary between soul and body. It is active in the reproduction of animals, functioning as one of the external agents that educes the higher forms from the matter where they exist as "seminal reasons."

This theory of seminal reasons was adopted on the authority of Augustine, but Bonaventure interpreted it within the framework of the general Aristotelian formula that forms are educed from the potency of matter. Unlike St. Thomas, Bonaventure interpreted these "potencies" as active powers rather than passive potentialities. They are really latent forms existing in matter in an inchoate or germinal state. External agents only cooperate with these powers, in much the way that a gardener cultivates a rosebush or a seedbed so that it bears flowers or germinates (*Commentarium in Librum II Sententiarum,* Dist. 7, in *Opera,* Vol. II, p. 198). All forms, except the primary light form and the human soul, which are directly created by God, arise through the cooperation of seminal powers and external agents, under the influence of light.

Bonaventure, unlike Thomas, believed that creation in time (in contrast with Aristotle's belief in the eternity of the world) is demonstrable from reason, using Aristotle's own principles (*Commentarium in Librum I Sententiarum,* Dist. 1, in *Opera,* Vol. II, pp. 20–22). His arguments, although interesting, are based on a medieval concept of number and infinity and on the presupposition that the immortality of the human soul is a purely rational truth.

As his name implies, Bonaventure's character seems to have represented all that the medieval Christian regarded as ideal. Born at a critical period in the history of his church, his order, and of speculative theology, he saw himself cast in a mediating role. As a bachelor of theology, trained in the arts, he sought to put the new philosophy into the service of theology. As a master of theology he tried not only to defend the new mendicant orders against the attacks of the secular masters but also to heal their differences. As minister general he took a middle position between the extreme factions of the Franciscan order, who differed on the subjects of evangelical poverty and the pursuit of studies. Bonaventure's works, such as *De Reductione Artium ad Theologiam* and *Itinerarium Mentis in Deum* were not only theoretical expressions of his gift for synthesis but also served the practical purpose of silencing the anti-intellectual friars who claimed that the academic life was incompatible with the ascetical aims of a follower of St. Francis. As cardinal, Bonaventure played a major role at the Council of Lyons in healing the rift between Greek and Latin Christendom. Under the aegis of Augustine, he consolidated theological opposition to the cult of Aristotle and Averroistic rationalism. Although this led eventually to the Parisian condemnations of 1270 and 1277, in which even theses of St. Thomas were included, it also bore fruit in a renewed interest in Augustine's contributions to philosophy by Matthew of Acquasparta, Roger Marston, John Peckham, and others of the Augustinian school.

See also Albert the Great; Alexander of Hales; al-Fārābī; Aristotle; Augustine, St.; Augustinianism; Averroes; Averroism; Avicenna; Boethius, Ancius Manlius Severinus; Dante Alighieri; Determinism, A Historical Survey; Grosseteste, Robert; Ibn Gabirol, Solomon ben Judah; John of Damascus; John of La Rochelle; Marston, Roger; Matthew of Acquasparta; Medieval Philosophy; Mysticism, History of; Peckham, John; Peter Lombard; Plato; Rationalism; Revelation; Thomas Aquinas, St.

Bibliography

The collected Latin works of St. Bonaventure were published as *S. Bonaventurae Opera Omnia.* 10 vols. (Quaracchi, 1882–1902). The two works not contained in that collection are *Questions disputées "De Caritate," "De Novissimis,"* edited by P. Glorieux (Paris, 1950), and a second redaction of *Collationes in Hexaemeron,* edited by F. Delorme (Quaracchi, 1934). *Commentary on the Sentences* has been reprinted in 4 vols. (Quaracchi, 1934–1949) without the critical notes. Also see *Tria Opuscula,* 5th ed. (Quaracchi, 1938), which includes *Breviloquium, Itinerarium Mentis in Deum,* and *De Reductione Artium ad Theologiam.*

English translations of Bonaventure's work include *St. Bonaventure's De Reductione Artium ad Theologiam,* translated by E. T. Healy (St. Bonaventure, NY, 1955), in Latin and English; *St. Bonaventure's Itinerarium Mentis in Deum,* translated by Philotheus Boehner (St. Bonaventure, NY, 1956), in Latin and English; *Breviloquium by St. Bonaventure,* translated by E. E. Nemmers (St. Louis and London: Herder, 1946); J. de Vinck, *The Work of Bonaventure* (Paterson, NJ: St. Anthony Guild Press, 1960–1970); and four questions from the "Commentary on the Sentences" in *Selections from Medieval Philosophers,* edited by Richard McKeon (New York: Scribners, 1930), Vol. II, pp. 118–148.

A Latin and Spanish translation is L. Amoros, B. Aperribay, M. Oromi, and M. Oltra, eds., *Obras de S. Buenaventura,* 6 vols. (Madrid, 1945–1949), Vol. I of which contains an extensive bibliography and new data on authentic works not in *Opera Omnia.*

General bibliographies for Bonaventure can be found in F. Ueberweg and B. Geyer, *Grundriss der Geschichte der Philosophie,* 12th ed. reprint (Basel, 1951), Vol. II, pp. 735–738, and Étienne Gilson, *History of Christian Philosophy in the Middle Ages* (New York: Random House, 1955), pp. 685–686. For studies after 1953, see the annotated *Bibliographia Franciscana* (Rome: Istituto Storico dei Cappuccini, 1962–), Vol. XI ff. The classic introduction to Bonaventure's thought is Étienne Gilson, *La philosophie de saint Bonaventure.* 2nd ed. (Paris, 1953; English translation of 1st ed., London: Sheed and Ward, 1938). Critical evaluations of this and alternate views are B. A. Gendreau, "The Quest for Certainty in St. Bonaventure," in *Franciscan Studies* 21 (1961): 104–227, and J. G. Bugerol, *Introduction à l'étude de S. Bonaventure* (Paris, 1961). Bonaventurean themes in Raphael are discussed in H. B. Gutman, "The Medieval Content of Raphael's 'School of Athens,'" in *Journal of the History of Ideas* 2 (1941): 420–429, and "Raphael's 'Disputà,'" in *Franciscan Studies* 2 (1942): 35–48. For Bonaventure's own theory of art, see E. J. M. Spargo, *The Category of the Aesthetic in the Philosophy of St. Bonaventure* (St. Bonaventure, NY, 1953). Also see R. P. Prentice, *The Psychology of Love according to St. Bonaventure,* 2nd ed. (St. Bonaventure, NY, 1957), which is a comparison of Bonaventure and Max Scheler, and A. Schaefer, "The Position and Function of Man in the Created World according to St. Bonaventure," in *Franciscan Studies* 20 (1960): 261–316; and 21 (1961): 233–382.

Allan B. Wolter, O.F.M. (1967)

BONHOEFFER, DIETRICH

(1906–1945)

Dietrich Bonhoeffer was a German theologian and religious leader during the period when National Socialism dominated. He was active in the resistance to Hitler; and his anti-Nazi activities led to his death in a concentration camp. The heroism of his end served to call attention to his life and thought, but by itself the drama of his life does not account for the continuing interest Bonhoeffer has aroused in twenty-first century theological circles. He has been read eagerly for the substance of his thought, his example of resistance to oppression, and his provocative portrayal of the secular settings that provide the context for much theological inquiry. The Nazi milieu prevented Bonhoeffer from making a sustained impact on the academic world during his lifetime; he was then recognized chiefly for his involvement in the nascent ecumenical movement, for his leadership of a clandestine seminary at Finkenwalde and, of course, for his part in the resistance to Hitler. (Thanks to the work of theologians such as John de Gruchy, Bonhoeffer's thought inspired much South African resistance to apartheid, and he has been invoked elsewhere by critics of oppressive political orders.)

PHILOSOPHY AND THEOLOGY

Only one of Bonhoeffer's works, *Akt und Sein*, is wholly devoted to formal questions concerning the relation of philosophy to theology. *Akt und Sein* was his inaugural dissertation, and it is marked by a certain pretentiousness and heavy-handed systematic theological concern. At times its jargon obscures the author's line of thought. It is doubtful whether the work possesses any great worth in isolation from Bonhoeffer's life. However, because it anticipates many of the themes that he later elaborated without explicit philosophical reference, it is of interest.

In *Akt und Sein* Bonhoeffer carried on a veiled polemic, on the one hand, against those who wished to reduce Christianity either to a philosophy of transcendence (*Akt*) or of being (*Sein*), and on the other hand against those who believed that Christian theology could be expressed independently of philosophical concerns. His own interests were in many ways synthetic. Critical of philosophical attempts to account for or exhaust the meaning of Christian revelation, Bonhoeffer admitted the general necessity of relating theology and philosophy. He appreciated the Kantian *Akt*-philosophy, which stresses the thinker or the knower "in relation to" the known, but he criticized its lack of interest in the problem of the known, as in the mundane world. He turned with some interest to the *Sein*-philosophies, which focus on God as the known but which may lack a proper corollary interest in the concrete historical events in which believers find God to be revealed. These philosophies Bonhoeffer categorized repeatedly throughout his career as "theologies of glory" that seek to explicate the nature of the Divine on a philosophical basis. He advocated mainly what in his Lutheran theological lineage has always been called "a theology of the Cross" because it accented an event in history, specifically in the crucifixion of Jesus Christ.

If the corpus of Bonhoeffer's most important literary work is to be related to philosophy, it must be categorized as a philosophy of history. In all his writings he shows an active and positive interest in the concrete character of Divine revelation. He often voiced an agnostic position on the possibility of making meaningful statements about God apart from revelation in Jesus Christ. In lectures on Christology delivered in 1932 and available in the form of published classroom notes, he concentrated consistently on the historical, concrete, and conditioned

character of revelation in Jesus Christ and the church over against philosophies of transcendence.

ETHICS

Bonhoeffer's *Ethics* is his most systematic work (although it survives only in fragments from the concentration camp years). Whereas it profits from philosophical debate, *Ethics* is largely a rejection of philosophical ethics. In this book Bonhoeffer takes a negative view of Roman Catholic ontological ethics, which moves from general abstract ethical statements to specific Christian principles. He was more closely identified with existentialism, but he regarded that philosophy also as an abstraction from revelatory events in Jesus Christ. Bonhoeffer has been accused, along with his teacher Karl Barth, of presenting an overly Christological philosophy and ethic, a critique that would not have disturbed him at all.

LATER THOUGHT

During his final imprisonment before his execution, Bonhoeffer's thought took a surprising—some would say a radical—turn. Pondering the collapse of continental humanist traditions at the hands of Nazis and other totalitarians, he focused on the blithe ways many of his contemporaries shrugged off inherited traditions of piety, even though some remained Christian. In his eyes, they joined free-spirited nonbelievers as they left behind preoccupations with guilt and modes he associated with conventional religion. He has come to be best remembered for his interpretation of modern history, developed on the basis of these observations and his study of the Bible during his imprisonment. From the Christian point of view he regarded secularization as a largely positive process. In a celebrated historical analysis he perceived that the "god of explanation" was gradually disappearing from European history; and disappearing with it was what he called "the religious *a priori*" (Bonhoeffer 1953). By this term he referred to the idea that a person must adopt a specific metaphysics, a specialized view of transcendence, or a particular form of piety and churchly existence before becoming a Christian. All of these, Bonhoeffer claimed, belonged to the spiritual adolescence of humans.

Contemporary humans, Bonhoeffer thought, reckoned less and less with a transcendent and hypothetical deity located outside the circle of the empirical. He cherished those Biblical texts and those aspects of theological tradition that spoke of transcendence located in the center of human affairs, particularly in the history of Jesus Christ. In this historical context, Bonhoeffer pointed out, the role of philosophy had become increasingly secularized as it focused on human autonomy (Bonhoeffer 1953).

In his eyes, René Descartes had begun to see the world as a mechanism. Benedict de Spinoza was a pantheist. Immanuel Kant, in Bonhoeffer's view, was close to the deists in his reluctance to deal philosophically with God as the known, in his revelation in history. He commented on the ways in which Johann Gottlieb Fichte and Georg Hegel had also developed brands of pantheism that drew them away from the historical involvement of God with the secular world.

All of these developments, he claimed in letters he wrote from prison, revealed the "growing tendency to assert the autonomy of man and the world" (*Prisoner for God*, Bonhoeffer 1954, p. 163). He came to be seen as a forerunner of a school of antimetaphysical theologians who insisted that Christian life and language were most free when they were not based on a philosophy of being or the expression of transcendence. Some of their writings became best-sellers in the 1960s and 1970s, when elements of Bonhoeffer's thought appeared in the controversial *Honest to God* (1963) by Bishop John A. T. Robinson and in a number of radical theological works, some of them momentarily associated with the concept of "the death of God."

Subsequently, cultural changes in Europe, wherein non-Christians and many Christians came to rediscover the potency of myths and symbols, which Bonhoeffer had earlier come to minimize, found significant figures resorting to new languages touting spirituality. In this context, a later generation of those influenced by Bonhoeffer reexplored those sources in his thought that were not exhausted by his witness to a "world come of age" and to the existence of a Christian church that was engaged, in almost carefree ways, with secular philosophies.

Part of this reexploration led some theologians to revisit the long-overlooked influence of Martin Heidegger on the young Bonhoeffer who wrote *Akt und Sein*. In that work, only Martin Luther was referenced more frequently than Heidegger. The most elaborate statement of this engagement was written by the American Charles Marsh in *Reclaiming Dietrich Bonhoeffer* (1994). Recognizing that Heidegger, a devotee of National Socialism, and Bonhoeffer, who was to give his life opposing it, were poles apart in politics, and that Bonhoeffer seldom quoted Heidegger after that early work, Marsh did discern some revisitations of the themes of transcendence that showed the influence of the philosopher. In Marsh's terms: "In an attempt to shape reflection in a way that is

not determined by the totality of the self-reflective subject but emerges from a source prior to and external to the individual, Bonhoeffer finds certain themes in Heidegger's fundamental ontology congenial to his theological purposes. Bonhoeffer subjects these themes to Christological redescription," and so does not stay with existential analysis (1994, p. 112). Nonetheless, Marsh argued, "Heidegger's notions of potentiality-for-being, authenticity, and being-with others push[ed] Bonhoeffer in his thinking about human selfhood and sociality to recognize specific social-ontological distinctions and concepts critical to his developing Christology" (Marsh 1994, p. 112).

Needless to say, such a view of connections and influence does not go unchallenged. Thus German theologian Ernst Feil presented anew what Marsh called "the conventional wisdom." In it, Heidegger's "concept of existence, derived from the human and not from revelation, was, for Bonhoeffer, theologically unusable" (1985, p. 31) agrees that Bonhoeffer finally did reject Heidegger's fundamental ontology on theological grounds, but awareness of this rejection "should not obscure Bonhoeffer's' admiration for Heidegger's *Being and Time's* attempt to 'destrue' or destructure the history of ontology," which captured Bonhoeffer's imagination in a decisive way (1994, p. 31). Yet even this self-described "reclamation" project by thinkers such as Marsh, while showing early dependence on Heidegger, does not serve to limit the imagination with which Bonhoeffer "revisited" Christological themes in a milieu he described as "a world come of age" (Bonhoeffer 1953, p. 327).

See also Barth, Karl; Descartes, René; Existentialism; Heidegger, Martin; Kant, Immanuel; Philosophy of History; Religion; Religious Language; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY BONHOEFFER

Nachfolge. München: C. Kaiser, 1937. Translated by Reginald H. Fuller as *The Cost of Discipleship* (New York: Macmillan, 1959).

Ethik. München: C. Kaiser, 1949. Translated by Neville Horton Smith as *Ethics* (New York: Macmillan, 1955).

Widerstand und Ergebung, edited by Eberhard Bethge. München: C. Kaiser: 1951. Translated by Reginald H. Fuller as *Letters and Papers from Prison*. London: SCM Press, 1953 and New York: Macmillan, 1971; and as *Prisoner for God* (New York, Macmillan, 1954.

Gesammelte Schriften. 4 vols. Munich: C. Kaiser, 1958–1961.

Akt und Sein. Translated by Bernard Noble as *Act and Being* (New York: Harper, 1962).

Sanctorum Communio. Translated as *The Communion of Saints* (New York: Harper & Row, 1963).

WORKS ON BONHOEFFER

DeGruchy, John W. *Bonhoeffer and South Africa: Theology in Dialogue*. Grand Rapids, MI: Wm. B. Eerdmans, 1984.

Feil, Ernst. *The Theology of Dietrich Bonhoeffer*. Translated by Martin Rumscheidt. Philadelphia: Fortress, 1985.

Marsh, Charles. *Reclaiming Dietrich Bonhoeffer: The Promise of His Theology*. New York: Oxford, 1994.

Robinson, John A. T. *Honest to God*. London: SCM Press, 1963.

Martin E. Marty (1967, 2005)

BONNET, CHARLES
(1720–1793)

Charles Bonnet, the Swiss naturalist, "religious cosmologist," and philosopher, was born and died in Geneva. An original if eccentric thinker, Bonnet was widely read and influential. He was early attracted to natural history, and especially to entomology, by René Réaumur's work and by the Abbé Pluche's apologetic, *Spectacle de la nature* (1732). At the age of twenty, he discovered that the aphis can reproduce for several generations without mating, and that animals other than the "polyp" (hydra) can regenerate themselves. He treated these and other matters in his *Traité d'insectologie* (1745). When his eyesight became severely weakened from microscopic work, he turned to botany and philosophy. In *Recherches sur l'usage des feuilles dans les plantes* (1754), he outlined a vitalistic concept of plant behavior in relation to physical environment. In the *Essai de psychologie* (1754) and the *Essai analytique sur les facultés de l'âme* (1760), he followed Étienne Bonnot de Condillac by using the device of the imaginary statue to illustrate the genetic method of explaining the development of the personality. The personality arises from memory, which grows out of sensations. Especially concerned with the body-mind relation, Bonnet accepted David Hartley's theory of association of ideas. He defined freedom as the power of the soul to follow necessary motives; but in granting man a substantial mind, he denied mechanical determinism. He held that the relation between mind and body indicates that the mind must operate in a physical organism, but survives it—an idea that was to be developed in his cosmic speculations.

With the *Considérations sur les corps organisés* (1762) and the popular *Contemplation de la nature* (1764–1765), Bonnet approached the general problems that were crucial in the biology of his time. In the *Considérations* he espoused the preformation theory (which he also needed for his cosmological speculations), using the work of

Albrecht von Haller and Lazzaro Spallanzani. In the *Contemplation,* he developed the traditional idea of the chain of beings, temporalizing it as a process rather than as a static creation. Bonnet's cosmic philosophy received full development in his *Palingénésie philosophique, ou Idées sur l'état passé et sur l'état futur des êtres vivants* (1770), a work that Arthur O. Lovejoy termed "one of the most extraordinary speculative compounds to be found in the history of either science or philosophy." Bonnet looked to biology as a support for his religious beliefs, and used both biology and religion to build a view of cosmic evolution.

Bonnet's theory held, essentially, that the immortal soul ("the ethereal machine") is a "subtle matter" (as distinguished from "gross matter") in the pineal gland. The ethereal machine is the germ of the resurrected body. All possible beings, all individuals, were created at once, according to the principle of plenitude. They exist in germ until released by the death of other individual organisms. The lower souls of animals are perfectible, and the universe is one in which all things tend to perfection. The principal changes occur as the result of catastrophes. Earth has passed through a series of epochs, each terminated by a cataclysm that destroyed all organic life except the immortal germs, allowing the germs to take on different forms, all foreseen in the original creation and all ascending to higher levels. Ontogenesis is a proof of this. Thus, every germ will reappear in a succession of higher embodiments, the soul of each waiting until the proper state of Earth evokes its next and higher incarnation. The entire creation is moving upward; man will become angel, and apes and elephants will take man's place. There is also life on other worlds, more or less advanced in perfection than on Earth.

This theory cannot be called one of organic evolution (as is sometimes erroneously affirmed), since species, according to Bonnet, have no natural history within a single world epoch. Species do not evolve from lower forms in the way modern biology conceives this process; their history is predetermined and fully inscribed in the germ at the moment of the original creation. The germ bears the form of all it will ever be. Nevertheless, Bonnet's universe is self-differentiating arid progressive.

Bonnet considered finalism in organisms an incontrovertible argument against atheism. An optimist, he maintained there is greater good than evil in the universe, and that created things necessarily have a lesser degree of perfection than their creator. Man is superior to animals in his sensual apparatus, brain, and speech organs; but he is part of the general, unfolding order of nature. Man knows a Natural Law that is virtual in him but develops by experience; however, he is moved by self-love and by passions, which may be beneficent or may be destructive and cruel. In considering the inherited organization more determining than education (experience), Bonnet was closer to the "man-machine" school of Julien Offray de La Mettrie than to the sensationist theories of Claude-Adrien Helvétius.

See also Condillac, Étienne Bonnot de; Descartes, René; Fichte, Johann Gottlieb; Hartley, David; Hegel, Georg Wilhelm Friedrich; Helvétius, Claude-Adrien; La Mettrie, Julien Offray de; Lovejoy, Arthur Oncken; Mind-Body Problem; Organismic Biology; Philosophy of Biology; Psychology.

Bibliography

The best work on Bonnet is Max Offner, *Die Psychologie Charles Bonnets* (Leipzig, 1893). See also Georges Bonnet, *Charles Bonnet* (Lac, 1929); Edouard Claparède, *La psychologie animale de Charles Bonnet* (Geneva: Georg, 1909); Jacques Roger, *Les sciences de la vie dans la pensée française du XVIII siècle* (Paris, 1963). For further bibliographical information, see D. C. Cabeen, A *Critical Bibliography of Eighteenth Century French Literature* (Syracuse, NY: Syracuse University Press, 1951), pp. 294–296.

OTHER RECOMMENDED TITLES

Anderson, Lorin. *Charles Bonnet and the Order of the Known.* Boston: Kluwer, 1982.

Anderson, Lorin. "Charles Bonnet's Taxonomy and Chain of Being." *Journal of the History of Ideas* 37 (1976): 45–58.

Kaitaro, Timo. "Ideas in the Brain: The Localization of Memory Traces in the Eighteenth Century." *Journal of the History of Philosophy* 37 (2) (1999): 301–322.

O'Neal, John C. *The Authority of Experience: Sensationist Theory in the French Enlightenment.* University Park: Pennsylvania University Press, 1996.

L. G. Crocker (1967)
Bibliography updated by Tamra Frei (2005)

BOOLE, GEORGE
(1815–1864)

George Boole, an English mathematician and logician, is regarded by many logicians as the founder of mathematical logic. He could be called the Galileo of logic in that he definitively established the mathematical nature of logic—assuming that it was Galileo Galilei who did this for physics, rather than, say, Archimedes. He is considered to be among the five greatest logicians, the others being the Greek philosopher Aristotle, the German mathemati-

cian Gottlob Frege, the Austrian mathematician Kurt Gödel, and the Polish mathematician Alfred Tarski.

Like Aristotle, he never had the opportunity to take a course in logic. His parents' economic circumstances precluded the usual formal education. He never took a college course and, thus, never received a bachelor's degree. Nevertheless, he taught many college courses as a professor of mathematics and he received honorary doctoral degrees from such distinguished institutions as Trinity College of Dublin and Oxford University. These are among the many surprises, ironies, and paradoxes surrounding Boole's life and work.

His ambition, energy, originality, and dedication were evident even when he was a boy. By the age of twenty-six he had published the first of many articles in mathematics journals. By twenty-nine, for his 1844 article "On a General Method in Analysis," he had won the Royal Society's gold medal first prize recognizing "the most significant contribution to mathematics" submitted between 1840 and 1844. At thirty-four he was appointed Professor of Mathematics at Queen's University. In 1864, when he died tragically just before the age of fifty, he was one of the most celebrated figures on the British intellectual scene.

In his lifetime he was known almost exclusively for his work in mathematical analysis, a specialty that includes traditional algebra, differential equations, the calculus of finite differences, and, of course, differential and integral calculus. In this field he wrote several articles and two books, both still in print: *Treatise on Differential Equations* (1859) and *Treatise on the Calculus of Finite Differences* (1860). During his lifetime few knew his logic at all, and of them few appreciated it. Today, his work in mathematical analysis is largely unknown; his fame rests entirely on his logic. Boolean algebra, the branch of modern mathematics named in his honor, derives from Boole's logic, not from his other mathematics.

REVOLUTIONARY LOGICIAN WHO NEVER INTENDED TO REVOLUTIONIZE LOGIC

His work in logic still retains a vigor and freshness; it continues to be read and enjoyed by many people including professional mathematicians and logicians. In 2003 Prometheus Books brought out a new reprint edition of his most mature and influential book: *An Investigation of the Laws of Thought on Which Are Founded the Mathematical Theories of Logic and Probabilities*—better known by its shortened title *Laws of Thought*—which was originally published at his own expense in 1854. The non-mathematical passages in this book are lucid and unusually well written—a testament to Boole's humanistic learning, to his confidence in his own theories, and to his desire to contribute to the advancement of knowledge. Besides the logic, Boole's 1854 book applies logic to probability theory.

Unlike other revolutionary logical innovators, Boole's greatness as a logician was recognized almost immediately. In 1865, hardly a decade after his 1854 *Laws of Thought* and not even a year after his death, his logic was the subject of the Harvard University lecture "Boole's Calculus of Logic" by Charles Sanders Peirce, America's most creative native logician. Peirce opened his lecture with these prophetic words: "Perhaps the most extraordinary view of logic which has ever been developed with success is that of the late Professor Boole. His book … *Laws of Thought* … is destined to mark a great epoch in logic; it contains a conception which in point of fruitfulness will rival that of Aristotle's *Organon*" (Peirce, pp. 223f.).

Even though Boole is thought of today as the initiator of a radical revolution that conclusively and irrevocably overthrew the Aristotelian paradigm then reigning in the domain of logic, he never thought of himself as opposing Aristotle. He admired Aristotle's logic—as far as it went. He never criticized any of the positive features that Aristotle instituted; he accepted as valid every argument that was valid according to Aristotle—including those with "existential import," deducing existential conclusions from universal premises. On the contrary, Boole's goals included revealing the mathematical nature of Aristotle's logic, something that he felt Aristotle had failed to clarify, broadening Aristotle's logic by showing that it could be made to do much more than was envisaged by Aristotle's followers, and deepening it by penetrating beyond Aristotle's analysis to the "ultimate" fine structure of the reasoning process—thereby providing it with what he called a mathematical foundation and showing that it had much more in common with mathematics than had previously been thought and thus justifying it. From Boole's point of view Aristotle's faults were all faults of omission, not of commission. Ironically, Boole's unquestioning acceptance of certain details of Aristotle's system, for example, existential import, may have been one of the things that led to Boole's unfortunate mistaken implementation of his own sound ideas.

BOOLE'S FULLY SYMBOLIC LOGIC

In the process of extending and deepening Aristotle's logic Boole brought many radical ideas into logic. Where Aristotle had represented propositions by a kind of formalized phonetic Greek, Boole represented them by purely ideographic algebraic equations—giving rise to the first successful formalized language in the modern sense. Where Aristotle's propositions were limited to exactly two basic nonlogical elements, one being the "subject" and one the "predicate," Boole's propositions had no limitation of that kind—they could involve any finite number of basic elements, which Boole represented with the letters familiar from algebra: x, y, z, and so on. In fact, by introducing for the first time in history the two logical elements—1 for "everything" or the universe of discourse and 0 for "nothing" or the empty class—he was able to express propositions of pure logic that were devoid of nonlogical elements, another historical first. It was Boole who coined the expression "universe of discourse," which is ubiquitous in modern logic, and it was Boole who first suggested the possibility of reinterpreting a formal language by changing the universe of discourse and the meanings of the nonlogical symbols.

Where for Aristotle the elements were represented by the Greek words having fixed meanings—for *human*, *animal*, and other substantives, Boole's letters were reinterpretable. Each of Aristotle's formal sentences expressed exactly one proposition whether true or false, but for Boole any single formal sentence was capable of expressing indefinitely many propositions not necessarily all true (as $x(1 - x) = 0$) or all false (as $x(1 - x) = 1$). Those that expressed only truths he said were "true in virtue of form," perhaps coining this expression also. This innovation was eventually to play a crucial role in modern logic. For example, with the multiplication sign or juxtaposition representing "logical term-conjunction" (the Boolean "and"), with x for *human* and y for *animal*, Boole thought he had expressed Aristotle's "Every human is an animal" by $xy = x$.

SOLVING LOGICAL EQUATIONS, DISCOVERING PROPOSITIONAL LOGIC, TRANSFORMING AN ORGAN INTO AN AXIOMATIC SCIENCE

These innovations opened the way to Boole's most radical, totally unexpected and unprecedented insight: that a fully interpreted equation expressing a proposition, whether true or false, could be considered as an equation with one element regarded as an "unknown" to be solved for in terms of the others. Where Aristotle's focus in formal logic had been exclusively with determining logical validity and invalidity of premise-conclusion arguments, that is, with what has been called formal epistemology, Boole's broader focus included, besides a much expanded formal epistemology, several new concerns, two of which were his wholly new theory of logical equation-solving and his formal ontology concerned with axiomatizing logical truths, which he called by the expression "laws of thought." Moreover, Boole explicitly recognized, as Aristotle had not, that "class logic," even in its expanded form, could not treat the arguments now dealt with in truth-functional proposition logic. To meet this deficiency, he proposed an ingenious reinterpretation of his "class logic" that, in his view, transformed it into a propositional logic. In the process he discovered the key ideas now incorporated into laws of modern truth-function logic, establishing himself as the first modern figure in any history of propositional logic.

Before Boole, logic had been thought of as an *organon* or general instrument necessarily presupposed by any axiomatic science, not as an axiomatic science; Boole proposed regarding logic itself as subject to axiomatic treatment. Boole believed that his logic transcended, included, explained, and thus replaced Aristotle's in much the way that Isaac Newton's mechanics transcended, included, explained, and thus replaced Johannes Kepler's.

See also Aristotle; Frege, Gottlob; Galileo Galilei; Gödel, Kurt; Kepler, Johannes; Logic, History of: Modern Logic; Mathematics, Foundations of; Newton, Isaac; Peirce, Charles Sanders; Probability and Chance; Tarski, Alfred.

Bibliography

WORKS BY BOOLE

The Mathematical Analysis of Logic: Being an Essay Towards a Calculus of Deductive Reasoning. Cambridge, U.K.: Macmillan, Barclay, and Macmillan, 1847.

"The Calculus of Logic." *Cambridge and Dublin Mathematical Journal* 3 (1848): 183–198.

Laws of Thought (1854). Introduction by John Corcoran. Amherst, NY: Prometheus Books, 2003.

WORKS ABOUT BOOLE

Boole, Mary Everest. *Collected Works*. Vol. 1, edited by E. M. Cobham. London: C. W. Daniel, 1931. Originally published under the title "The Home Side of a Scientific Mind" in *The University Magazine* in 1878.

Corcoran, John. "Aristotle's *Prior Analytics* and Boole's *Laws of Thought*." *History and Philosophy of Logic* 24 (2003): 261–288. This is an extensive, but incomplete, point-by-

point comparison of Boole's logic with the Aristotelian system it was intended to perfect.

Gasser, James, ed. *A Boole Anthology: Recent and Classical Studies in the Logic of George Boole*. Dordrecht, Netherlands: Kluwer Academic, 2000.

Grattan-Guinness, Ivor, and Gérard Bornet, eds. *George Boole: Selected Manuscripts on Logic and Its Philosophy*. Boston: Birkhäuser Verlag, 1997.

Hailperin, Theodore. *Boole's Logic and Probability: A Critical Exposition from the Standpoint of Contemporary Algebra, Logic, and Probability Theory*. Amsterdam: North-Holland, 1976.

MacHale, Desmond. *George Boole: His Life and Work*. Dublin, Ireland: Boole Press, 1985. This includes a complete bibliography of Boole's publications.

Peirce, Charles. *Writings of Charles S. Peirce*. Vol. I. Bloomington: Indiana University Press, 1982.

Rhees, R., ed. *Studies in Logic and Probability*. London: Watts, 1952.

John Corcoran (2005)

BOSANQUET, BERNARD

(1848–1923)

Bernard Bosanquet, the English philosopher, was born at Altwick and educated at Harrow and at Balliol College, Oxford. He taught ancient history and some philosophy at Oxford from 1871 to 1881, when he left Oxford for London. In London he edited translations of Rudolf Hermann Lotze's *Logic* and *Metaphysics*, played an active part in the London Ethical Society, worked with the Charity Organisation Society, and did some teaching in the adult education movement. In 1895 he married Helen Dendy, who had been employed by the Charity Organisation Society and who later wrote much on social problems and became a member of the important Royal Commission on the Poor Law of 1909. From 1903 to 1908 he held the chair of moral philosophy at St. Andrews. He died in London.

Bosanquet's first important philosophical work is an essay titled "Logic as the Science of Knowledge" in *Essays in Philosophical Criticism* (A. Seth and R. B. Haldane, eds., London, 1883), a collection of papers in memory of T. H. Green. In *Knowledge and Reality* (London, 1885) he criticized F. H. Bradley's *Principles of Logic* for divergences from the central and, as Bosanquet thought, correct course charted in that book. In 1888 Bosanquet's *Logic or the Morphology of Knowledge* (2 vols., London) was published. Bosanquet had earlier translated the introduction to G. W. F. Hegel's *Philosophy of Fine Art* (London, 1886), and his own *History of Aesthetics* appeared in London and New York in 1892. His Gifford lectures were published as *The Principle of Individuality and Value* (London, 1912) and *The Value and Destiny of the Individual* (London, 1913). Bosanquet was a prolific writer who contributed to discussion in all branches of philosophy and also took part in some social controversy. He was two years younger than Bradley and, like him, came to the Idealist point of view partly through the influence of T. H. Green and partly through reading Hegel. Bradley's *Ethical Studies* influenced him, but Bradley, in his turn, learned from Bosanquet's writings, especially from those on logic. Although both were Idealists, and both were called Absolutists, Bosanquet was more Hegelian and less of a skeptic than Bradley.

LOGIC

In the essay "Logic as the Science of Knowledge," which appeared in the same year as Bradley's *Logic* and seems to be independent of it, Bosanquet set out the main lines of his 1888 *Logic*. In this preliminary essay he argued that truth is comprehensible only within systems of knowledge, and that although truth is correspondence with fact, such correspondence is conceivable only within systems because "the facts by which we test conclusions are not simply given from without," and they are not available for judgment until they are "organised into knowledge." He also argued that judgment and inference are not fundamentally distinct, but that judgment is inference not yet made explicit and inference is explicit judgment. A further feature of this striking essay is that in it the forms of judgment are not regarded as fixed and rigid but as "elastic" in their application, so that a form of sentence best suited to express one form of judgment can in fact be used to express many others.

In *Knowledge and Reality* Bosanquet suggested that Bradley had, in spite of his "essential and original conceptions" as to the general nature of judgment and inference and their connection with each other, fallen into some of the errors of "reactionary logic." Bradley said, for example, that categorical judgments state facts, whereas hypothetical judgments (and with them universal ones) do not. By an ingenious choice of examples, Bosanquet shows that such a contrast cannot be sustained and that there is no contrast between being a fact and being a universal. Bosanquet's method is to cite intermediate cases that make impossible the acceptance of sharp distinctions between forms of judgment. He thinks that Bradley was inclined to isolate his examples from their contexts and to lose sight of the subtleties and complexities of language. An instance of this part of Bosanquet's argument is his discussion of Bradley's example "the sea-serpent exists."

Bosanquet points out that it is far from clear what this means in the abstract and that "'exist' is a formal predicate which receives material interpretation from context."

In *Logic or the Morphology of Knowledge* these views are worked out in systematic form. The first volume is concerned with judgment and the second with inference, but the two parts are very closely linked. Bosanquet did not think that, in actual and advancing thought, form and subject matter could be separated. Thus he regarded formal logic not as the standard of thought but as a highly specialized and idealized, and somewhat subsidiary, type of thinking. The forms of judgment and inference with which he concerns himself, therefore, are those that he regards as operative in the actual advancement of knowledge. Judgment is concerned with truth, and mere interjections do not claim to be true; but there are rudimentary judgments of quasi-interjectional type, such as "How ugly!" or "Oh, horrible!" Such impersonal judgments as "It rains" take us still further along the road of developing thought, and demonstratives take us still further. "This" is always so by relation to "that," so that demonstratives lead on to comparison; and as comparison is made more exact, it leads on to proportion and measurement.

At this point, according to Bosanquet, the series diverges, one route being that taken by what he calls "the concrete or categorical series" and the other by what he calls "the abstract or hypothetical series." Along the first route there are singular judgments and those he calls generic judgments, in which a kind is regarded as real, as when we say "Man is mortal" or "Water boils at 212 degrees Fahrenheit." Along the second line of development there are the various types of abstract judgment, such as "Heat is a mode of motion" or "7 + 5 = 12," in which the emphasis is on necessary connection rather than on concreteness. The two series converge again in the hypothetical judgment, and the whole culminates in the disjunctive judgment, which Bosanquet regards as the most adequate form. His reason for this is that it combines the concreteness of the categorical series with the necessity of the hypothetical series. The various disjuncts, in this view, reveal a system in which every member has its distinct place.

Bosanquet illustrates this by such examples as "The triangle is either scalene, isosceles, or equilateral." In the *Essentials of Logic* (London and New York, 1895), he refers to functions within a social order of the sort which, if an individual exercises one of them he does not and cannot exercise any of the others: if a person is king, he is not subject; if he is judge, he is not prosecutor. In his account of inference, Bosanquet also lays great stress on intermediate and transitional forms. Furthermore, just as he minimizes the difference between judgment and inference, so he minimizes the difference between deduction and induction. He holds that knowledge advances neither by generalization from particulars nor by the elimination of hypotheses. Inference, in his view, depends upon the existence of systematic connections, and neither mere counting nor mere discarding can reveal these to us. What is needed is "depth and complexity of insight into a sub-system of the world," and the word "induction" is used when our points of contact with the real world are "isolated perceptions, occurrences or qualities." But the aim of all inquiry is to break down this isolation and to show how the elements of a system must be what they are. Thus, as knowledge advances, the aspect of contingency is less prominent, mere facts or mere observations play a vanishing part, and we come to see that things must be as they are.

METAPHYSICS

For Bosanquet, as for Hegel, there is no sharp division between logic on the one hand and epistemology and metaphysics on the other. Indeed, although logic is concerned with the forms of judgment and inference, the study of these forms leads to the conclusion that reality is systematic. If facts were distinct and isolated, it would be impossible to infer from one to another. Since inferences can be made, facts are not isolated but are "implicated" with one another and "transcend" themselves. The possibility of inference points to the metaphysical fact of "self-transcendence."

Bosanquet's metaphysical system is outlined in his *Principles of Individuality and Value* and given more detailed application in *The Value and Destiny of the Individual*. These titles indicate Bosanquet's concern with individuality and individuals. His view is that individuals are concrete universals. He contrasts (as Bradley had done) abstract universals, such as redness, with concrete universals, such as Julius Caesar. Abstract universality is the repetition of an identical quality in many instances, whereas concrete universality is the realization of the same individual in its various interrelated acts or manifestations. The many red things are extremely diverse, whereas the actions of an individual are more or less systematically connected with one another. According to Bosanquet, "there can be only one individual, and that, *the* individual, the Absolute." When people are called individuals, it is in a "secondary sense," insofar as they are regarded as relatively independent, stable, and unique.

But this uniqueness is not some internal, private, inaccessible feature of them. The "inwardness" of persons is not something private, not "the banishment of all that seems outward, but the solution of the outward in the circulation of the total life."

McTaggart complained that everything Bosanquet says about mind and body "might have been written by a complete materialist," and Bosanquet himself in *Knowledge and Reality* had written that "a consistent materialist and a thorough idealist hold positions that are distinguishable only in name." Bosanquet rejects both psychophysical interactionism and the view that mind is an effect of matter. He holds that mind is a perfection of the organism and that an organism possesses more or less of it as the organism selects from, and adapts itself to, the circumstances of its world. He rejects the possibility of a mind independent of matter, and draws ethical conclusions from this. Without things, he says, there would be no problems for men. If there were nothing but disembodied persons, there would be nothing to do.

In bringing these general principles to bear upon aspects of experience, Bosanquet comes to some surprising conclusions. His view of individuals as concrete universals might have been expected to lead to a respect for historical knowledge, as it has done with other Idealists. But, according to Bosanquet, history is "a hybrid form of experience," "the doubtful story of successive events." His view is that the spatiotemporal contingencies of human life must, as knowledge grows, become absorbed into a fuller understanding of society, art, philosophy, and religion. These, he says, are "concrete and necessary living worlds." Bosanquet also rejects the view, advocated by Thomas Carlyle, James Anthony Froude, and Bradley, that human conduct and discovery cannot be predicted. He argues that this thesis depends upon the false assumption that individuals cannot overlap, and he holds that such facts as "anticipatory" inventions that have to be "reinvented" are evidence to the contrary. Thus, in *The Value and Destiny of the Individual* he concluded that "intelligences must overlap" and stigmatized as "the pathos and bathos of sentimentalism" the view that selves are essentially withdrawn and alone.

SOCIAL PHILOSOPHY

From what has already been said about Bosanquet's metaphysics, it follows that societies are individuals to a fuller degree than individuals can be. In the *Philosophical Theory of the State*, he treats the relation between the individual and the state as that of microcosm to macrocosm. The individual world and the social world are held to be correlated with one another in such a way that for every element in the one there is some corresponding element in the other. Like Aristotle and Jean-Jacques Rousseau, he emphasizes the civilizing influence of the state on the individual. He rejects the commonsense, pluralistic metaphysics that he thinks misdirects the social philosophy of Jeremy Bentham and John Stuart Mill. "All individuals," he writes, "are continually reinforced and carried on, beyond their average immediate consciousness, by the knowledge, resources, and energy which surround them in the social order." "The common self or moral person of society," he holds, "is more real than the apparent individual."

Hence, like Rousseau, he regards coercion by the state as coercion exercised by the social aspect of the individual upon the recalcitrant and less real aspects of his being. According to classical liberalism, the individual is free when he is left alone to do what he wants. According to Bosanquet, this is a metaphysical as well as a practical impossibility, so he develops the conception of freedom as self-mastery. But since selves are not exclusive atoms, self-mastery, social control, and freedom are held to coincide. Bosanquet accepts T. H. Green's view that action under compulsion has less value than action freely willed, thus recognizing that state enforcement can lead to mere external conformity. But just as he regarded nature as the necessary complement of mind, so he regarded force, habit, and tradition as the necessary complements of creative choice. Thus, although punishment acts on the "lower self" by means of threats, it can also stimulate the "higher self" by producing a shock that forces attention to legitimate social demands. Still, the function of the state is forcibly to "hinder hindrances to the best life or common good," and the very notion of promoting morality by force is "an absolute self-contradiction."

Thus, although Bosanquet minimizes and even denies the reality of individual men, he does not advocate totalitarian or even socialistic measures. Indeed, just as Bastiat, the publicist of laissez-faire, considered that society as a whole was moved by an impersonal reason, so Bosanquet believed that intelligence is manifested in society to a greater degree than it ever could be in any particular person. He has been criticized for failing to distinguish between society and the state and for suggesting that the state can do no wrong. There is justice in the former criticism, even though we may agree that force is inevitable if developed societies are to continue in existence. As to the second, Bosanquet's main philosophical point was that theft, murder, and such are concepts that apply to people within a society, and that war, conquest,

confiscation, and such are concepts of a different type, applying to beings of a different type.

Bosanquet's account of what makes them different types is very complex. He points out that many crimes committed on behalf of the state result from the desire of some individual agent of the state to take a short cut or to save trouble and hence are not imputable to it. Furthermore, the state cannot commit wrongs of the sort that are the consequences of individual selfishness or sensuality. On the other hand, a state that ordered the killing of a hostile statesman would rightly be criticized, not on the ground of murder but "by the degree of its failure to cope with the duties of a state." Bosanquet seems to mean that when a state is rightly criticized, it is compared with more adequate specimens of its own type but is not blamed or punished as are individuals who break the law. Bosanquet holds that states are morally responsible beings, but that they cannot do wrong in the way that individual persons can and do. States fall short rather than do wrong. Furthermore, he repudiates the idea that individuals are guilty of murder when a state wages war or of theft when it annexes or confiscates; any moral criticism, he holds, should be directed against the morally responsible agent, the state itself, and such criticism must relate to the general level of life it sustains and promotes. At the end of World War I Bosanquet opposed such popular appeals as "Hang the Kaiser" and "Punish the Germans," and although he said that the League of Nations was "the hope and refuge of mankind," he believed that individual members should no more submit themselves unreservedly to this organization than individual men should submit themselves unreservedly to their own governments.

See also Aristotle; Bentham, Jeremy; Bradley, Francis Herbert; Carlyle, Thomas; Green, Thomas Hill; Hegel, Georg Wilhelm Friedrich; Idealism; Knowledge and Belief; Logic, History of; Lotze, Rudolf Hermann; Macrocosm and Microcosm; McTaggart, John McTaggart Ellis; Mill, John Stuart; Punishment; Rousseau, Jean-Jacques; Society; State.

Bibliography

On Bosanquet's life, see Helen Bosanquet, *Bernard Bosanquet* (London: Macmillan, 1924) and J. H. Muirhead, ed., *Bernard Bosanquet and His Friends: Letters Illustrating the Sources and Development of His Philosophical Opinions* (London: Allen and Unwin, 1935).

Apart from the works mentioned in the text, Bosanquet's books include: *A Companion to Plato's Republic* (New York: Macmillan, 1895); *The Psychology of the Moral Self* (London and New York: Macmillan, 1897); *Three Lectures on Aesthetics* (London: Macmillan, 1915); *Some Suggestions in Ethics* (London: Macmillan, 1918); *Implication and Linear Inference* (London: Macmillan, 1920); *What Religion Is* (London: Macmillan, 1920); *The Meeting of Extremes in Contemporary Philosophy* (London, 1921); and other shorter works.

There are also the following collections of essays and lectures: *Essays and Addresses* (London, 1889); *The Civilisation of Christendom* (London S. Sonnenschein, 1893); *Social and International Ideals* (London: Macmillan, 1917); *Science and Philosophy and Other Essays* (London: Allen and Unwin, 1927).

For discussions of Bosanquet's views, see: J. M. E. McTaggart's critical notice of *The Principle of Individuality and Value* in *Mind*, n.s., 21 (1912): 416–427; H. B. Acton, "The Theory of Concrete Universals," in *Mind*, n.s., 45 (1936): 417–431 and n.s., 46 (1937): 1–13; and F. Houang, *Le Néo-Hégélianisme en Angleterre: la philosophie de Bernard Bosanquet* (Paris, 1954).

On Bosanquet's social philosophy, see: L. T. Hobhouse, *The Metaphysical Theory of the State* (London: Allen and Unwin, 1918) and A. J. M. Milne, *The Social Philosophy of English Idealism* (London: Allen and Unwin, 1962), Ch. VII.

OTHER RECOMMENDED WORKS

Bosanquet, Bernard, ed. *Aspects of the Social Problem*. New York: Kraus Reprint, 1968.

Bosanquet, Bernard. *Croce's Aesthetic* (1919). Folcroft, PA: Folcroft Library Editions, 1974.

Bosanquet, Bernard. *The Distinction between Mind and Its Objects* (1913). Bristol: Thoemmes, 1990.

Bosanquet, Bernard, tr. *The Education of the Young in The Republic of Plato* (1917). Folcroft, PA: Folcroft Library Editions, 1973.

Bosanquet, Bernard. *Three Chapters on the Nature of Mind* (1923). Edited by Helen Dendy Bosanquet. New York, Kraus Reprint, 1968, 1923.

H. B. Acton (1967)
Bibliography updated by Michael J. Farmer (2005)

BOSCOVICH, ROGER JOSEPH
(1711–1787)

Roger Joseph Boscovich (or Rudjer Josip Bošković) was a Jesuit scientist whose originality and advanced views have only recently been appreciated. A natural philosopher, mathematician, physicist, astronomer, geodesist, engineer, and poet, Boscovich was, in the words of the physicist John Henry Poynting, "amongst the boldest minds humanity has produced." Boscovich published about one hundred books and papers, most of them in Latin. These works display an unusual combination of enthusiasm and logic as well as a passionate conviction that simple fundamental assumptions and precise reasoning can lead

to the understanding of natural phenomena. The French astronomer Joseph Jérôme Le Français de Lalande said that in each of these works there are ideas worthy of a man of genius.

Boscovich was born at Ragusa (now Dubrovnik, Croatia) of Serb and Italian parentage. He entered the novitiate of the Society of Jesus in Rome in 1725 and the Collegium Romanum in 1727. At the Collegium stress was laid on clear logical thought and on the development of a way of thinking that combined religious convictions with the results of science. Boscovich devoted himself chiefly to mathematics and physics and published his first scientific paper in 1736. He became professor of mathematics at the Collegium in 1740, and in 1744 he took his vows as a priest. Since his gifts were scientific, Boscovich was left free to apply himself to teaching, research, and tasks designated by the religious authorities. In 1734 Pope Benedict XIV appointed him, with others, as a technical adviser concerned with cracks in the dome of St. Peter's, and in 1750 commissioned him with Christopher Maire, an English Jesuit, to measure an arc of the meridian through Rome. Later, Boscovich was designated to arbitrate a dispute between the Republic of Lucca and Austrian Tuscany over the drainage of a lake. This task took him to Vienna, where he already enjoyed a high reputation as a scholar and a diplomat. From 1759 on, Boscovich was engaged in extensive travels as far away as Constantinople. In 1760 he met Benjamin Franklin and many other leading personalities in London and Cambridge, and he was elected a fellow of the Royal Society in 1761. He became professor of mathematics at Pavia in 1765, but his health was failing and he grew restless. A chair was created for him at Milan in 1769, and he pursued studies at the Brera observatory. In 1775 Boscovich was appointed director of naval optics for the French navy and went to Paris, where he was made a subject of France by Louis XV. He returned to Italy in 1783. During his last years he suffered from melancholia.

Despite these activities Boscovich continued to publish. Each of his numerous works in pure and applied mathematics presented either a new method for or a survey of some branch of mathematical inquiry. Among the topics he discussed were spherical trigonometry, the cycloid, conic sections, infinitely great and infinitely small quantities, the accuracy of astronomical observations, the telescope, sunspots, eclipses, the determination of the sun's rotation and of the orbits of planets and comets, the aurora borealis, the transit of Mercury, the shape of Earth, the variation of gravity, the center of gravity, and optical problems. His last major publication was a five-volume work on optics and astronomy, *Opera Pertinentia ad Opticam et Astronomiam,* published at Venice in 1785.

Boscovich's masterpiece, and his work of greatest interest to philosophers, is *Philosophiae Naturalis Theoria Redacta ad Unicam Legem Virium in Natura Existentium* (A theory of natural philosophy reduced to a single law of the actions existing in nature), published in Vienna in 1758 and, in an improved edition, at Venice in 1763. In this work Boscovich presented an atomic theory on which he had been working for fifteen years. The importance of this theory was widely recognized, especially in Britain, where the *Encyclopaedia Britannica* devoted fourteen pages to it in 1801. Boscovich had been the first supporter in Italy of Isaac Newton's theory of gravitation, and the *Theoria* was looked upon in Britain as an interesting speculative extension of the Newtonian system.

Boscovich's atomic theory arose, as he himself stated, from an attempt to build a comprehensive physics based on the ideas of Newton and Gottfried Wilhelm Leibniz but going beyond both to obtain new results. Boscovich developed the idea that all phenomena arise from the spatial patterns of identical point particles (*puncta*) interacting in pairs according to an oscillatory law that determines their relative acceleration. This view of matter is akin to that of recent physics in that it is relational, structural, and kinematic. It contains three original features:

(1) Material permanence without spatial extension: Quasi-material point-centers of action are substituted for the rigid finite units of matter of earlier atomists.

(2) Spatial relations without absolute space: Internal spatial coordinates (the distances between the two members of pairs of *puncta*) are used instead of external coordinates.

(3) Kinematic action without Newtonian forces: In modern dimensional terms, Boscovich's theory is kinematic rather than dynamical. It uses only two-dimensional quantities (length and time) rather than the three (mass, length, and time) used by Newton. Since all particles are identical, the number of particles in a system, which is an integral pure number obtained by counting, is employed in place of Newtonian mass.

Although all of these features are of interest, the first is most important, for by it Boscovich helped emancipate physics from naive atomism's uncritical assumption that the ultimate units of matter are small, individual, rigid pieces possessing shape, size, weight, and other properties. The alternative point atomism assumes that the ultimate units are persistent quasi-material points, all identical, which form stable patterns or interact to pro-

duce changes of pattern and relative motion. Between 1710 and 1760 such other thinkers as Giambattista Vico, Leibniz (whose theory of monads and relational conception of space influenced Boscovich), Emanuel Swedenborg, John Michell, and Immanuel Kant had produced atomic theories based on points, but Boscovich was the first scientist to develop a general physical theory using point particles.

Boscovich preferred the concept of *puncta* to that of rigid pieces of matter because they were simpler and, since they avoided the awkward discontinuity at the surface of a piece of matter, were better adapted to mathematical treatment. His law of oscillatory change from attraction to repulsion enabled him to posit points of stable equilibrium at finite distances and thus to account for the finite extension of gross matter, as Kant did also. The complexity of the world, according to Boscovich, arises from two factors: the varied arrangement of different numbers of particles, and the parameters determining the law of oscillation.

To a modern reader, the impressive feature of the *Theoria* is Boscovich's interpretation of the universe as a three-dimensional structure of patterns in equilibrium or change determined by points and their mutual distances. There is no distinction between occupied and empty space, for space is only the relation between *puncta.* Space, time, and motion are all relative; the *puncta* form a vast variety of stable patterns; the laws of the universe are simple, but their consequences are complex; the laws contain several natural units of length, as do the laws of modern physics since the introduction of Planck's constant; there is a pervasive continuity in nature permitting inference from the macroworld to the microworld; geometry is in part a creation of the human mind and can to some extent be chosen at will; the ability of atomism to account for the forms and processes of the natural universe is unlimited, and even organic forms are easy to understand, because complex patterns of particles will adhere to one another in figures of certain shapes.

As a speculative vision of a universe of changing structure supported by an appropriate philosophy of physics, Boscovich's system is brilliant, but as a scientific theory it is incorrect because it does not allow for the highly complex properties of the wave-particles of present-day physics. No data concerning the atomic world were available to provide a quantitative basis for Boscovich's theory, and he was able to give only a qualitative description of simple mechanical and physical properties. The physical world is more complex than the world Boscovich created from his imagination. Nevertheless, his philosophy of physics was in some respects near the truth, for he predicted—a century and a half before the facts were known—that matter is penetrable by high-speed particles and that relative motion affects the measurement of space and time. Moreover, these predictions were necessary consequences of his mathematical conception of three-dimensional structure. Boscovich's standard of simplicity remains a challenge to physics, and only a future, fully unified, particle theory will be able to show precisely where his assumptions were mistaken. Boscovich postulated that there is only one fundamental particle; we do not yet know how many must be assumed. Modern conceptions of molecular structure have much in common with Boscovich's ideas, but since the development of the physical concept of a field, it can be seen that the Boscovichian particle is inadequate even to account for electromagnetic processes.

It is not certain how far the *Theoria* influenced the development of atomic theory. It was widely studied, and Michael Faraday, Sir William Hamilton, James Clerk Maxwell, and Lord Kelvin (to mention only English scientists) stressed the theoretical advantages of the Boscovichian atom over rigid atoms. In any case, Boscovich's work marked an important stage in the history of our ideas about the universe, and his system will remain the paradigm of the theory of point particles.

See also Faraday, Michael; Franklin, Benjamin; Hamilton, William; Kant, Immanuel; Laws, Scientific; Leibniz, Gottfried Wilhelm; Maxwell, James Clerk; Newton, Isaac; Philosophy of Physics; Swedenborg, Emanuel; Vico, Giambattista.

Bibliography

The second edition of the *Theoria* was republished in a Latin-English edition, with the English translation by J. M. Child, as *Theory of Natural Philosophy* (Chicago and London: Open Court, 1922).

For literature on Boscovich, see L. L. Whyte, ed., *Roger Joseph Boscovich, S.J., F.R.S., 1711–1787: Studies of His Life and Work on the 250th Anniversary of His Birth* (London: Allen and Unwin, 1961; New York: Fordham University Press, 1961), which contains a biographical essay by Hill and eight papers on aspects of Boscovich's work by English, American, and Yugoslav scholars. Its extensive bibliography of works by and about Boscovich does not, however, cover Yugoslav studies. See also H. V. Gill, *Roger Joseph Boscovich, S.J., 1711–1787: Forerunner of Modern Physical Theories* (Dublin, 1941), and L. Pearce Williams, *Michael Faraday* (New York: Basic, 1965).

Lancelot Law Whyte (1967)

BOSSUET, JACQUES BÉNIGNE
(1627–1704)

Jacques Bénigne Bossuet was born in Dijon, the son of a lawyer. At the age of thirteen he was a boy canon of Metz. After a period in Paris, where he became known in the salons and distinguished himself as a theologian, he was ordained priest in 1652 (having been prepared by Vincent de Paul) and began his ministry at Metz. Friends in high places secured his recall to Paris in 1659, and he soon established a reputation as preacher and spiritual director. Contemporaries agree that he had the ability, and presumably the desire, to please everyone; and his early reputation for moderation may reflect tactics more than convictions. Winning favor at Court, he was rewarded in 1669 with the see of Condom and was appointed tutor to the dauphin, Louis XIV's son, in 1670. He is most famous for the series of funeral orations he delivered as Court preacher (1666–1687), of which the last and finest commemorates the great Condé. Besides these set (and published) pieces, he preached numerous sermons for all occasions, often using the feast of a particular saint for an exposition of his own views on a contemporary question, such as the relations between church and state, lucidly discussed in the panegyric of St. Thomas of Canterbury (Becket). Some 200 sermons survive, mostly as notes on which he usually improvised, and it is easier to establish his main ideas than to reconstruct his mastery of the spoken word.

On completion of the tutorial task, he was transferred in 1681 to Meaux, conveniently near Paris, where he remained until his death. His influence at Court gave him more effective power than his hierarchical superiors, and in 1682 he composed and presented the Gallican Articles as spokesman for the whole French church. His last years were marred by quarrels, especially with his former protégé François Fénelon, whose condemnation for quietism he secured only by resorting to methods so ignoble that formal victory was bought at the cost of moral defeat. Despised at Court and broken in health, he ended his life among relatives of notoriously unedifying character.

All Bossuet's thinking was deeply influenced by St. Augustine and characterized by a peculiar emphasis on authority. In his eyes, obedience and discipline are the highest virtues. The supreme authority of the church and the divine right of kings are inseparable and constantly recurrent themes in his work. In the *Politique tirée de l'ecriture sainte* (Politics Drawn from Scripture), written for the dauphin, he is heavily in favor of the absolute monarch, chosen by God and responsible to him alone (distinguished, however, from the arbitrary monarch, a tyrant who merely gratifies his own whims). The *Traité de la connaissance de Dieu et de soi-même* (Treatise concerning the Knowledge of God and Oneself) combines Thomist and other standard teaching with a marked sympathy for the reassuringly authoritarian side of Cartesianism, with its insistence on order and certainty, although Bossuet elsewhere denounced the dangers of encouraging individual reason and inquiry. The unfinished *Discours sur l'histoire universelle* (Discourse on Universal History) was intended to teach the dauphin not so much what had happened as why. Though later editions made some concessions to currently changing views on the chronology of ancient times, history was primarily interpreted as showing the ways of God to man, especially as revealed in the Bible. In tracing the fortunes of empires down to Charlemagne (and to Louis XIV, if he had completed his plan), Bossuet emphasized moral and religious development, regarding freedom as a prime cause of decadence.

Similarly, the *Histoire des variations des églises protestantes* (History of the Variations of the Protestant Churches) attributes to Protestant reliance on individual liberty of conscience a disunity amounting to near anarchy. Bossuet naturally regarded heresy and sedition as twinned evils; and in his orations on Henrietta Maria and Henrietta Anne (widow and daughter of Charles I), he adduces the recent revolution in England to prove his contention that social equality is an impious chimera. He was curiously ambivalent in his relations with Protestants, converting many individuals (including the vicomte de Turenne) and courteously corresponding with Gottfried Wilhelm Leibniz in an attempt to effect a reconciliation, while greeting the revocation of the Edict of Nantes, followed as it was by brutal persecution, with an embarrassingly effusive eulogy of Louis's piety.

Bossuet earns his place in history above all as a public figure, "the eagle of Meaux." In the *grand siècle* Bossuet was the church, just as Louis was the state.

See also Augustine, St.; Authority.

Bibliography

WORKS BY BOSSUET

Oeuvres complètes. 31 vols, edited by F. Lachat. Paris, 1862–1866.

Discourse on Universal History, edited by Orest Ranum; translated by Elborg Forster. Chicago: University of Chicago Press, 1976.

Politics Drawn from the Very Words of Holy Scripture, edited by Patrick Riley. New York: Cambridge University Press, 1990.

WORKS ON BOSSUET

Adam, A. *Histoire de la littérature française au XVIIe siècle.* Paris, 1956, Vol. V, Ch. 4.

Calvet, J. *Bossuet, l'homme et l'oeuvre.* Paris: Boivin, 1941.

Hazard, P. *La crise de la conscience européenne.* 2nd ed. Paris: Fayard, 1961. Translated by J. L. May as *The European Mind.* London: Hollis and Carter, 1953.

Kearns, Edward J. *Ideas in Seventeenth Century France.* New York: St. Martin's Press, 1979.

Lebarcq, J. *Histoire critique de la prédication de Bossuet.* Lille, 1888.

Rebelliau, A. *Bossuet.* Paris, 1900.

A. J. Krailsheimer (1967)
Bibliography updated by Tamra Frei (2005)

BOSTRÖM, CHRISTOPHER JACOB
(1797–1866)

Christopher Jacob Boström, an Swedish Idealist philosopher, studied and also taught at Uppsala University, where he was assistant professor of "practical philosophy" (the philosophy of morals, law, and religion) from 1828 to 1833. After an interlude as tutor to the royal princes in Stockholm from 1833 to 1837, he resumed his academic teaching, and from 1842 to 1863 he held the chair in practical philosophy. His "rational idealism" is a spiritualistic metaphysics, combining traits from Plato's theory of ideas, Gottfried Wilhelm Leibniz's monadology, and George Berkeley's immaterialism. With arguments, some of which are reminiscent of Berkeley's, he tried to show that nothing but minds and their perceptions exist.

Two of his more original, though hardly very convincing, arguments were these: (1) Truth means agreement between the perception and the perceived object. Perfect truth, therefore, is perfect agreement; and perfect agreement is the same as identity. Hence, the object of any perfectly true perception is identical with that perception; in other words, any object, when perceived with perfect truth, is itself a perception. (2) "Outside" has a meaning only when it refers to space. Since a mind is not in space, nothing can be outside a mind. Hence, everything exists inside a mind.

Particular minds and particular perceptions are forms of "self-consciousness," which can be likened to "a substance or stuff of which everything ultimately consists." With this spiritualistic position Boström combined the Leibnizian-Kantian distinction between a thing as it is in itself (essence) and a thing as it appears to us (phenomenon). The spatiotemporal world of experience is merely phenomenal. Or, more correctly, the spatiotemporal world of a person's experience is merely the way in which the things-in-themselves appear to that person because of the imperfection of his particular perceptive faculty. The things-in-themselves, which underlie the appearances, are purely rational minds whose existence is nonspatial and nontemporal. Boström usually called them "ideas," the word being borrowed from Plato rather than from British empiricism. These ideas form a series that, according to him, is similar to the series of natural numbers—except that it contains a maximal idea, God. In this series each idea contains and perceives all the preceding, but none of the succeeding, ones. On this point, however, he was apparently not quite consistent. Simultaneously he asserted that every idea perceives the entire system of ideas but with varying perfection and clarity. God alone has a perfect perception of the whole system. Because every idea that is not God perceives the system imperfectly, the system presents a phenomenal appearance to that idea.

Boström's system contains several other apparent inconsistencies. Although each mind is a purely rational, nonspatial, and nontemporal idea, Boström also taught that each mind other than God has a double existence. Besides existing as a rational idea, it also exists as a temporal mind with a mixed rational and sensual nature. Each mind even has a whole (temporal?) sequence of such mixed and temporal manifestations. (Boström himself points to the analogy between this doctrine and the Hindu belief in reincarnation.) He was thinking primarily of human beings in this context, but the doctrine of double existence is also supposed to apply to such "moral personalities" as the state, the "people," and each one of the four estates.

Boström was aware of the nonintellectual motives that attracted him to this view of the world and once asserted that no philosopher would ever embrace a system that was repugnant to his feelings. Simultaneously, however, he made excessive claims concerning the provability of his own doctrine, to which he attributed the same kind of certainty that has traditionally been ascribed to mathematics.

From the vantage point of his rather fantastic metaphysics, Boström took an active part in public debate in Sweden. In religious questions he was, on the whole, a liberal, vigorously attacking many of the dogmas of Lutheran orthodoxy, especially the dogma of eternal damnation. On political questions, on the other hand, he

took an ultraconservative stand. He was one of the staunchest opponents of the parliamentary reform that took place in 1866, soon after his death, and that replaced the four estates by a two-chamber system. His metaphysics might seem to indicate a mystical strain, but his very systematic, precise, and dry mode of writing does not corroborate this impression. The dominant traits in his philosophic temperament would seem to be a strong, puritanical, moral pathos, an unorthodox but firm religious belief, a love of neat systematics, and a rather naive private dogmatism. Boström's philosophy represents the culmination of the idealistic tradition that dominated Swedish philosophy through the entire nineteenth century. In the 1860s, 1870s, and 1880s, Boströmianism and Hegelianism reigned supreme in Swedish academic philosophy. At the turn of the twentieth century a strong neo-Kantian current set in.

See also Berkeley, George; Hegelianism; Idealism; Leibniz, Gottfried Wilhelm; Plato.

Bibliography

WORKS BY BOSTRÖM

Skrifter av Christopher Jacob Boström. Edited by H. Edfeldte and J. G. Keijser. Vols. I and II, Uppsala, Sweden: V. Roos, 1883; Vol. III, Norrköping, Sweden, 1901. Collected works.

C. J. Boströms Förelasningar i Religionsfilosofi. Edited by S. Ribbing. Stockholm, 1885. Lectures on philosophy of religion.

Prof. C. J. Boströms Förelasningar i Etiken. Edited by S. Ribbing. Uppsala, Sweden, 1897. Lectures in ethics.

C. J. Boströms Förelasningar i Religionsfilosofi II. Edited by G. J. Keijser. Vol. I, Stockholm, 1906; Vol. II, Stockholm, 1910; Vol. III, Stockholm, 1913. Second series of lectures in the philosophy of religion.

Prof. C. J. Boströms Förelasningar i Etik Vårterminen 1861. Edited by G. Klingberg. Uppsala, Sweden: Akademiska bokhandeln, 1916. Boström's lectures in ethics of the spring term of 1861.

Translations

Grundlinien eines philosophischen Systems. Translated by R. Geijer and H. Gerloff. Leipzig, 1923. German translation of various writings.

Philosophy of Religion. Translated by Victor E. and Robert N. Beck. New Haven, CT: Yale University Press, 1962.

WORKS ON BOSTRÖM

Larsson, H. *Minnesteckning över C. J. Boström.* Stockholm, 1931. Memorial oration.

Morin, Harald. *Om Dualismen i Boströms Definitiva Filosofi.* Uppsala, Sweden: Almqvist and Wiksells, 1940. On the dualism in Boström's definitive philosophy.

Nyblaeus, A. *Den Filosofiska Forskningen i Sverige,* 4 vols. Lund, Sweden, 1873–1897.

Rodhe, S. E. *Boströms Religionsfilosofiska Åskådning.* Göteborg, Sweden: Elanders, 1950. Boström's views in the philosophy of religion.

Wedburg, A. *Den Logiska Strukturen hos Boströms Filosofi.* Uppsala, Sweden, 1937. Logical structure of Boström's philosophy.

GENERAL BACKGROUND

Ueberweg, F., and M. Heinze. *Grundriss der Geschichte der Philosophie,* 12th ed. Berlin, 1928. Vol. 5. Excellent survey of Swedish philosophy up to the beginning of the twentieth century.

A. Wedburg (1967)

BOULAINVILLIERS, HENRI, COMTE DE

(1658–1722)

The historian, philosopher, astrologer, and savant Henri, Comte de Boulainvilliers, or Henry, Comte de Boulainviller, as he preferred to spell his name, was born at Saint-Saire, Normandy. From 1669 to 1674 he was educated at the Oratorian school at the College of Juilly, where Richard Simon taught rhetoric and philosophy. Boulainvilliers took up military service, as befitted a member of an old aristocratic family, proud of his lineage. After leaving the army, he developed an interest in history, first studying his own family tree and then the social and political institutions of the Middle Ages. He approved of feudalism, which he envisaged as a kind of federal republic governed by distant and independent aristocratic families, whom he considered to be the inheritors of the Franks who had conquered the Gauls. He deplored the increase in the power of the central authority—the king—and in the liberties of the people as encroachments on the rights of the nobles. He favored a patriarchal society. Many of his reforms, submitted to the regent, recommended the fostering of trade, proportional taxation, the suppression of tax collectors, and the calling of the États Généraux. The count had access to Court circles; he was connected with d'Argenson, president of the council of finance, to whom it is thought he passed on a number of clandestine philosophical tracts. He also frequented the home of the maréchal, duc de Noailles, where he met César Dumarsais, a disciple of Bernard Le Bovier de Fontenelle, future author of articles for the *Encyclopédie* and probable author of *La religion chrétienne analysée* and *Examen de la religion*; Nicolas Fréret, a devotee of Pierre Bayle; and Jean-Baptiste de Mirabaud, the *secrétaire perpétuel* of the *Académie Française.*

For a time Boulainvilliers was the center of much intellectual activity, and in the history of free thought his coterie antedates by fifty years the better-known *côterie holbachique.* Voltaire in his *Dîner du comte de Boulainvilliers* (1767) has given us an insight into this milieu, which certainly disseminated a surprisingly large number of clandestine manuscripts and seems to have provided the only organized center for the compiling, copying, and distribution of philosophical tracts. Boulainvilliers is best known as the probable author of parts of the *Essai de métaphysique,* which was published in 1731 under the title *Réfutation des erreurs de Benoît de Spinoza.* He became interested in Benedict de Spinoza through reading the *Tractatus Theologico-Politicus,* which he annotated copiously, and also the *Ethics,* which he read in 1704. The first part, or *Vie de Spinoza,* of the *Essai de métaphysique* has been attributed to J. M. Lucas. The second part, or *Esprit de Spinoza,* has been attributed by I. O. Wade and others to Boulainvilliers. Both parts are commonly coupled together in the manuscripts and in the editions under the title *La vie et l'esprit de Spinoza.* Boulainvilliers correctly presents Spinoza's doctrine that God and the universality of things are one and the same, then proceeds to argue that Spinoza's "attributes" are in fact "modes"; that is, "modes" of something he terms existence.

In this work, he has evolved an original philosophy. Starting from the Cartesian principle that he knows himself to be a thinking being, he infers that other beings exist, some endowed with thought, others only with feeling, and others without feeling or thought. All beings, whether living or nonliving, thinking, feeling, or merely extended, have one property in common: existence. From such premises, he proceeds to a universal Idea or Being more all-embracing than matter. He stresses the degrees of being, and claims that sensations are the source of all experience. He concludes by asserting that at death the body returns to universal matter while the soul remains as an idea in the infinite mind and is, therefore, capable of being restored to the body. It is clear that Boulainvilliers's exposition of Spinoza is curiously based on the Cartesian assertions and incorporates ideas borrowed from John Locke.

He strove to harmonize the notion of a single substance with a sensationalist psychology and a naturalistic ethics. He believed in a "chain of being," in the capacity of animals to think, and in evidence (as opposed to judgment) as the only criterion of truth; he also helped to discredit Christian revelation. In an *Abrégé d'histoire ancienne* he expressed his belief in the primacy of natural laws, denying the possibility of miracles. These points were later taken up by Denis Diderot in the article "Certitude" of the *Encyclopédie.*

DE TRIBUS IMPOSTORIBUS

Figuring as part of the *Essai de métaphysique,* sometimes titled *L'esprit de Spinoza,* is to be found a treatise commonly known as the *Traité des trois imposteurs,* under which title it was published in 1719 (2nd ed., 1721; numerous others throughout the century). Since printed copies were commonly impounded and consequently hard to find, manuscript copies continued to circulate both before and after publication. Polemic and concise, it provided freethinkers with valuable ammunition. Its aggressive title helped to ensure its success and may have been chosen by the Dutch printers as the last and profitable stage of an elaborate hoax. It is an allusion to a lost treatise, *De Tribus Impostoribus* (1230), supposedly written by Frederick II for the edification of his friend Othon. Interest in this Latin work, evidenced in *Theophrastus Redivivus* (1659), had been revived at the close of the seventeenth century and the beginning of the eighteenth.

The author of the *Traité des trois imposteurs,* believed by Voltaire to be Boulainvilliers, launched a virulent attack on the prophets and apostles; he expressed his disbelief in heaven or hell, rewards or punishments, his faith in natural law as enshrined in the hearts of men, and in the soul as the expression of the principle of life. The system of religion is, according to him, the work of false legislators, among whom are Moses, Christ, and Muḥammad. Moses was nothing more than a magician and a charlatan; Christ, who may be likened to Genghis Khan, was a casuist in his discussions with the Philistines and in claiming to be the son of a god; his religion owes much to Greek mythology and his ethics compare unfavorably with those of Epictetus and Epicurus. Muḥammad differs from the other two impostors in having recourse to violence in the establishment of his kingdom. Voltaire, among others, seized on these points to bolster his polemics against the church. He, too, saw the advantage of an oblique attack on the church by an onslaught against Islamic fanaticism, coupled with the claim that all religions are equal. The treatise marks an early, if crude, attempt to consider religion from the comparative standpoint.

Boulainvilliers is best remembered as a confirmed "spinoziste," and his views on the subject of nature and matter, the relationship of matter and thought, and the origin and nature of government won him a place as a forerunner of the philosophes.

See also Clandestine Philosophical Literature in France; Spinozism.

Bibliography

WORKS BY BOULAINVILLIERS

État de la France, etc., avec des mémoires historiques sur l'ancien gouvernement de cette monarchie jusqu' à Hugues Capet …. 3 vols. London: T. Wood and S. Palmer; W. Roberts, 1727.

Histoire de l'ancien gouvernement de la France, avec 14 lettres historiques sur les parlements ou états généraux. 3 vols. The Hague and Amsterdam: Aux dépends de la compagnie, 1727.

Mémoires présentés au duc d'Orléans, régent de France, contenant les moyens de rendre ce royaume très puissant et d'augmenter considérablement les revenus du roi et du peuple. 2 vols. The Hague, 1727.

La vie de Mahomed. London: Humbert, 1730. Unfinished.

Essai de métaphysique dans les principes de Spinoza. Brussels, 1731. Published under the title *Réfutation des erreurs de Benoît de Spinosa,* par M. de Fenelon, par le P. Lami et par M. le Comte de Boullainvilliers, avec *La vie de Spinosa* écrite par M. Jean Colerus.…

Histoire des Arabes. 2 vols. Amsterdam: Humbert, 1731.

Mémoire pour la construction d'un nobiliaire général and *Mémoire sur la noblesse* (1753). Both unpublished; Mss. at Angoulême.

WORKS ON BOULAINVILLIERS

Buranelli, V. "The Historical and Political Thought of Boulainvilliers." *Journal of the History of Ideas* 18 (4) (1957).

Dosher, Harry R. "Henri, Comte de Boulainvilliers, Historian of the French Aristocracy, 1658–1722." MA thesis. University of North Carolina, Chapel Hill, 1960.

Gargallo di Castel Lentini, Gioacchino. *Boulainvilliers e la storiografia dell'Illuminismo francese.* Giannini-Naples, 1954.

Levy, Neil. "History as Struggle: Foucault's Genealogy of Genealogy." *History of the Human Sciences* 11 (4) (1998): 159–170.

Simon, R. *Henry de Boulainviller, historien, politique, philosophe, astrologue.* Paris, 1939.

Spink, J. S. *French Free-Thought from Gassendi to Voltaire.* London: Athlone Press, 1960.

Torrey, N. L. "Boulainvilliers: The Man and the Mask." *Travaux sur Voltaire et le XVIIIe siècle* 1 (1955): 159–173.

Wade, I. O. *The Clandestine Organization and Diffusion of Philosophic Ideas in France from 1700 to 1750.* Princeton, NJ: Princeton University Press, 1938.

Robert Niklaus (1967)
Bibliography updated by Tamra Frei (2005)

BOWNE, BORDEN PARKER

(1847–1910)

Borden Parker Bowne, an American Personalist philosopher, spent his scholarly life, that is, from 1876 to 1910, at Boston University, where he taught in the liberal arts college and the school of theology, and where he became the first dean of the graduate school. In many articles and in seventeen books, Bowne expounded his Personalism, or Personalistic Idealism, which held that the Creator-Person, God, and created persons constitute the real.

Bowne was constantly concerned with taking full account of every dimension of human experience, be it the logical, the emotional, the moral, or the religious. Each dimension should be given full value and not be arbitrarily explained away by pontifical claims made in the name of such doctrines as Christian supernaturalism, psychological associationism and materialism, or ethical utilitarianism. For Bowne, reason is the criterion of truth. This means that for him reasoning discovers the real by interweaving and interpreting the different dimensions of experience.

The presupposition of thought and action is a unified, thinking self, or person. Were the person unable to will freely (granted limitations) and to choose in accordance with moral and intellectual ideals, there could be no trustworthy science or philosophy and no significance to moral and religious living. It is in the nature and experience of this self-identical, thinking, willing, and feeling person, who may not be reduced either to a mode of matter or to a mode of divinity, that Bowne finds his clue to, and his model of, reality.

Persons, however, do not create themselves, or each other. They could not communicate with each other were they not bound by the same laws of reason and subject to a common world. Each knower is bombarded by a flux of discontinuous sense impressions and responds as constructively as possible in accordance with his or her own dynamic categories, such as time, space, quality, quantity, cause, substance, and purpose. Thus the "common world" is the phenomenal world as organized by knowers who interact with, and ultimately depend upon, the structure of the real world independent of them. The phenomenal world is not a mask of the real world; it is the real world as related to the cognitive nature and purposes of finite knowers.

Bowne argues that the real world is neither nonmental nor independent of persons. For in knowing, and in interacting with an order other than itself, the mind must meet not only the conditions of its own nature but those of some agency or agencies independent of it. Since knowledge exists, and yet is not imported into a passive mind, the realist's contention that the real is unaffected by knowing is unintelligible. The fact must stand that minds, in following their own natures, can know with reasonable assurance the reality in which they live and can construct

a common world of thought and action, even though they are not identical with the real in knowing.

Furthermore, minds in their theoretical and practical action are clearly alien neither to each other nor to the reality that is the source of their experiences. The world as known is the world persons construct, following the nature of their own theoretical interests, on the basis of the reality beyond their thought. Why, then, hold that any reality beyond finite things is nonmental if such cooperative interaction is possible?

Bowne granted that the case against nonmental "material being" is not proved beyond a shadow of a doubt. But he argued that what we do know about the relation of mind to nature is more economically explained if we think of nature as the energizing of a cosmic Person. Nature is God willing in accordance with rational principles, hence nature dependably supports the orderly common world our finite reasons construct in response to it. God, however, is not identical with the natural world. He is transcendent as well as immanent in relation to it. He is the unified, dynamic ground of nature, and he uses it for his purposes, inclusive of his interaction with finite persons.

How, then, are finite persons related to God? Finite persons are created by God and have relative, delegated autonomy. The real world, whose structure maintains and guides the constructive cognitive adjustments of persons, does not force their moral and appreciative responses. But when persons do not treat each other as persons in a realm that is morally purposeful, they fall short of what their own natures in God's world can be. God created man free, to work out the content of his freedom in a world order that at once limits and gives him opportunity for fulfillment. Human freedom could effect nothing in a world without order, for persons do not create the rational or moral principles by which they guide their thought and action in the given ultimate order.

For Bowne, then, the natural world as known by persons is the objectification of the orderly interaction between finite wills and cosmic Will. The ethical world is the objectification of the orderly, chosen, interaction among free, finite persons in the natural world God makes possible. Bowne's universe is not (like Benedict de Spinoza's) a unity with many finite modes. It is a realm of persons united both by God's purposive action in nature and by the further moral unity created as persons freely respond to the reason, will, and love of the cosmic Person.

See also Idealism; Personalism; Spinoza, Benedict (Baruch) de.

Bibliography

Among the most interesting philosophical works by Borden Parker Bowne are *The Theory of Thought and Knowledge* (New York: Harper, 1897); *Metaphysics* (New York: Harper, 1898); *Theism* (New York: American, 1902); and *Personalism* (Boston: Houghton, Mifflin, 1908).

Works about Bowne include E. S. Brightman, "Personalism and the Influence of Bowne," in *Proceedings of the Sixth International Congress of Philosophy,* edited by E. S. Brightman (New York: Longmans, Green, 1927); and A. C. Knudson, *The Philosophy of Personalism* (New York: Abingdon Press, 1927).

For bibliography, see F. J. McConnell, *Borden Parker Bowne* (New York: Abingdon Press, 1929).

Peter A. Bertocci (1967)

BOYLE, ROBERT
(1627–1691)

Robert Boyle, the English natural philosopher, was the fourteenth child of Richard Boyle, the first earl of Cork, who by judicious marriages and land purchases had made himself the most influential man in Ireland and the richest in England. The political and financial fortunes of the earl of Cork fluctuated considerably during his son's lifetime, but ultimately Robert Boyle inherited a considerable income, which greatly facilitated his scientific researches.

In October 1635, Boyle entered Eton, which with Sir Henry Wotton as provost was a notable center of culture and learning. As a result of a change of teachers, Boyle left Eton in 1638 to be privately tutored. In 1639 he went to Geneva, where he studied mathematics; his devotion to religion, so he tells us in his fragment of an autobiography, *An Account of Philaretus during His Minority,* dates from this same period. A visit to Florence in 1641–1642 introduced him to Galileo Galilei's ideas and confirmed him in his hostility to Roman Catholicism. His return to England was delayed by a crisis in his father's affairs. When Boyle was free to return to England in 1644, his father was dead and he had inherited the manor of Stalbridge in Dorsetshire.

Boyle stayed at first in London with his favorite sister, Lady Ranelagh, whose house was a center of intellectual life. There he met Samuel Hartlib (d. 1670?), an enthusiastic educator and intellectual middleman, through whom Boyle was brought in touch with the burgeoning scientific activities of London. In Boyle's correspondence with Hartlib there are several references to their membership in an "Invisible College"; this has generally been identified by biographers with the Gresham's

College group out of which the Royal Society was to develop. The "Invisible College" Boyle referred to, however, would seem rather to have been an independent group centering on Hartlib and having an interest in social and educational reform as well as in science.

From 1645 until 1652 Boyle lived in retirement at Stalbridge, remote from the political upheavals of the times. He was still essentially a dilettante, interesting himself— but not too seriously—in chemistry, writing theological tracts of a highly moral character, and composing what was perhaps the first religious novel, *Seraphic Love* (1648). In 1652–1653 he visited his Irish estates; unable to obtain materials for chemical experiments, he studied anatomy under William Petty. The interest in biological processes thus engendered remained with him. In bad health from early manhood, he was particularly interested in the application of chemical methods to the cure of disease and was a diligent collector of prescriptions.

The Commonwealth had appointed a number of London scientists to posts at Oxford; in 1654 Boyle accepted an invitation from John Wilkins to make his home there. Now his serious career as a scientist began. He built a laboratory and employed a number of research assistants, in particular Robert Hooke (1635–1703), later to be curator of experiments at the Royal Society. With Hooke's help, Boyle constructed a greatly improved air pump, experiments with which provided the groundwork for Boyle's first and most important scientific work: *New Experiments Physico-Mechanical touching the Spring of the Air and Its Effects* (1660). Following up the work of Galileo and Evangelista Torricelli, Boyle demonstrated that air has both weight and elasticity and that the phenomena that had traditionally been ascribed to an anthropomorphically conceived "horror of a vacuum" were, in fact, a product of the air's elasticity.

His conclusions created an immediate stir but were not universally accepted. Boyle was criticized on philosophical grounds by Thomas Hobbes, Henry More, and the Jesuit Franciscus Linus (1595–1675), to all of whom he replied in detail. In the course of his reply to Linus, Boyle formulated what is known as Boyle's law. (On the Continent it is called Mariotte's law, Mariotte having confirmed it in 1676.) In the years that followed, Boyle took part in the meetings of the embryonic Royal Society at Oxford, conducted and published a great many experiments, corresponded voluminously with most of the leading thinkers of Europe, studied Oriental languages, actively supported the distribution of the Bible in foreign parts—becoming for that purpose a governor of the Corporation for the Spread of the Gospel to New England and a director of the East India Company—and wrote a considerable number of scientific, philosophic, and theological treatises. After the Restoration most of his scientific friends returned to London; Boyle left Oxford for London in 1668 and lived in Lady Ranelagh's household until her death. He died a week later.

SCIENCE AND PHILOSOPHY

Boyle was profoundly influenced by Francis Bacon's conception of science; much of his published work consists of what Bacon called "histories"—systematic accounts of such qualities as color, firmness, and coldness as they appear under a variety of circumstances. His *Spring of the Air* was the first scientific paper of the modern type. He encouraged scientists to write relatively brief experimental "essays" rather than general treatises. His *Animadversions upon Mr. Hobbes' Problemata de Vacuo* (published in Boyle's *Tracts*, 1674) emphasized the fruitlessness of a priori philosophical reasoning—what Boyle called "book philosophy"—about issues that could be settled only by experiment.

But it is wrong to suppose that Boyle was an opponent of theorizing. He discusses the place of theory in science in his proemial essay to *Certain Physiological Essays and other Tracts* (1661). Scientists, he says, should "set themselves diligently to make experiments and collect observations, without being over forward to establish principles and axioms." Theories ought never to be taken as final; they should be thought of as "the best we have but capable of improvement." Nevertheless, it is the scientist's task to develop theories that are as clear, as simple, and as comprehensive as possible—a point that particularly emerges in Boyle's essay "About the Grounds of the Mechanical Hypothesis" (published in *The Excellency of Theology*, 1674).

Indeed, it was Boyle's main object "to beget a good understanding between the chemists and the mechanical philosophers, who have hitherto been too little acquainted with each other's learning." The corpuscular theory, which Pierre Gassendi had revived, suffered, Boyle thought, in the eyes of practical chemical experimentalists because so little had been done to test it. Theorists had been accustomed to illustrate their theories rather than to test them. On the other side, the work of the chemists had been ignored by physical theorists, largely because it had been associated with theories of a totally inadequate kind.

DOCTRINE OF MATTER

Boyle's *The Sceptical Chemist* (1661) is mainly concerned with demonstrating the unsatisfactory character of the standard chemical theories. It is written in the form of a dialogue in which the main speaker, Carneades, attacks not only the traditional theory of elements but also the alchemical theories that had been proposed by Paracelsus and Jan van Helmont. None of these theories, Boyle argued, can be reconciled with experiment, unless they are interpreted in so vague and symbolic a manner as to make them scientifically worthless. As an alternative, he set up the corpuscular theory. It is sometimes said that he also so redefined "elements" as to prepare the way for the modern doctrine of elements; but that is a mistaken interpretation. Indeed, what his chemistry lacked was precisely this modern conception of elements. That is why he was still able to believe in the possibility of alchemical transmutations. In 1689 he secured the repeal of Henry IV's statute against "multiplying gold."

In a sense, however, Boyle's work was too advanced theoretically. Not enough was known about chemical substances to enable the corpuscular theory to be effectively applied in chemistry. Although, in trying to bring together physics and chemistry and chemistry and biology, Boyle anticipated the long-range development of science, the program that he laid down for chemistry was one that for the moment no one knew how to fulfill; the immediate effect may well have been to hold back the development of chemistry. Boyle conceded, it is true, that explanations referring to perceptible properties rather than to the behavior of corpuscles are, at a certain level, perfectly satisfactory; but the general effect of his work was to discourage explanations of the only sort that chemists were actually in a position to offer. His own writings abound in interesting theoretical suggestions—in his *General History of the Air* (1692), for example, he anticipated the kinetic theory of gases—but for a very long time they had to remain no more than suggestions. Although Boyle's actual contributions to science are very few in number, the range of his anticipations is remarkable. He had set out to make chemistry respectable; he had succeeded, many chemists thought, only at the cost of turning it into physics.

PRIMARY AND SECONDARY QUALITIES

Boyle exerted an important influence on philosophy by lending the authority of a practicing scientist to the corpuscular theory of matter and the associated doctrine of primary and secondary qualities. In *The Experimental History of Colours* (1663), Boyle sets out to demonstrate that color is a "secondary quality" (his own terminology). Objects give rise to sensations of color, he tries to show, not because they are themselves colored but because the structure of their corpuscles modifies light in a special way. The word *color* is most properly applied, he argues, to the modified light that "strikes upon the organ of sight and so causes that sensation we call colour"; if we say that bodies themselves are colored, this can mean no more than that, by virtue of "a certain disposition of the superficial particles," they are capable of refracting or reflecting light.

This thesis is generalized in *The Origin of Forms and Qualities according to the Corpuscular Philosophy* (1666), in which the theory of qualities, which John Locke was to rely upon in his *Essay concerning Human Understanding*, is set forth in detail and contrasted with the Scholastic doctrine of substantial forms. The qualities of a material object, Boyle argues, consist of "the size, shape and motion or rest of its component particles, together with that texture of the whole which results from their being so contrived as they are." These primary qualities of objects, operating upon the "peculiar texture" of a sensory organ, "occasion ideas in us."

SCIENCE AND RELIGION

The corpuscular philosophy had generally been associated with atheism. Boyle sets out to show that "by being addicted to experimental philosophy a man is rather assisted than indisposed to be a good Christian" (*The Christian Virtuoso,* 1690). His views about the relation between God and Nature, however, are by no means clear. In "An Hydrostatical Discourse Occasioned by Some Objections of Dr. Henry More," included in *Tracts* (1672), Boyle strongly opposes More's view that mechanical principles cannot explain the phenomena of pressure or any other physical phenomena. We do not need, he says, to have recourse to More's "incorporeal creatures"; mechanism is enough. Yet, at the same time, in *Forms and Qualities* he argues against René Descartes that we cannot account for the behavior of living organisms by supposing that they consist of particles on which God bestowed motion. We have to suppose, Boyle says, that the Creator not only set the world moving but also introduced into it "seminal seeds" that are responsible for the growth and propagation of animal organisms.

Again, in *A Disquisition about the Final Causes of Natural Things* (1688), he expresses his disagreement with those who would reject final causes completely, although he also argues that the scientist, in his day-to-

day work, need pay no attention to anything except the size, shape, texture, and motion of particles. At times, indeed, as in *The Excellency of Theology, or the Pre-eminence of the Study of Divinity above That of Natural Philosophy,* Boyle's anxiety about the contemporary tendency to abandon theology in favor of scientific inquiries leads him into a skepticism about science. If theology has its obscurities, he argues, they are as nothing to the obscurities inherent in the scientific account of continuity or of the relation between mind and body. Revelation can tell us far more about the place of man in nature than can science. But the example of Boyle the scientist was more influential than the precepts of Boyle the theologian. His last gesture in favor of Christianity was to leave in his will a sum sufficient to endow lectures for the defense of Christianity against its opponents; his intellectual legacy, however, was the mechanical interpretation of the world that deism took as its starting point.

See also Atheism; Bacon, Francis; Carneades; Colors; Deism; Descartes, René; Galileo Galilei; Gassendi, Pierre; Hobbes, Thomas; Locke, John; Matter; More, Henry; Paracelsus; Physicotheology; Primary and Secondary Qualities; Scientific Method; Scientific Theories.

Bibliography

For a bibliography of Boyle's voluminous works, see John Farquhar Fulton, *A Bibliography of the Honourable Robert Boyle,* rev. 2nd ed. (Oxford: Clarendon Press, 1961). The only complete collected edition of Boyle's *Works,* edited by Thomas Birch, and with a still useful life by Birch that incorporates Boyle's *Account of Philaretus,* was published in London in five volumes in 1744 and in six volumes in 1772. *The Sceptical Chemist* is in the Everyman Library (London, 1911), and *The Weight and Spring of the Air* is in Vol. I of *Harvard Case Histories in Experimental Science* (Cambridge, MA: Harvard University Press, 1957), an anthology edited by James B. Conant.

For works on Boyle, see Edwin A. Burtt, *The Metaphysical Foundations of Modern Physical Science* (London: K. Paul, Trench, Trubner, 1925; rev. ed., 1932), Ch. 6, "Gilbert and Boyle"; Louis T. More, *The Life and Works of the Hon. Robert Boyle* (New York and London: Oxford University Press, 1944); Mitchell S. Fisher, *Robert Boyle, Devout Naturalist* (Philadelphia: Oshiver Studio Press, 1945); Marie Boas, *Robert Boyle and Seventeenth Century Chemistry* (Cambridge, U.K.: Cambridge University Press, 1958)—an excellent book; Richard S. Westfall, "Unpublished Boyle Papers relating to Scientific Method," in *Annals of Science* 12 (1) (March 1956): 63–73, and 12 (2) (June 1956): 103–117.

John Passmore (1967)

BRADLEY, FRANCIS HERBERT
(1846–1924)

The English idealist philosopher Francis Herbert Bradley was born in Clapham and educated at University College, Oxford; in 1870 he was elected to a fellowship at Merton College, Oxford, terminable on marriage. Since he never married and the terms of the fellowship did not require him to teach, he was able to devote himself entirely to philosophical writing. His first published work was a pamphlet titled *The Presuppositions of Critical History* (Oxford, 1874). There followed *Ethical Studies* (London, 1876), *Principles of Logic* (London, 1883), and *Appearance and Reality* (London, 1893), as well as many articles in philosophical journals, some of which were published in *Essays on Truth and Reality* (Oxford, 1914) and others in *Collected Essays* (Oxford, 1935).

Like Bernard Bosanquet, Bradley was influenced by T. H. Green. Like Bosanquet, too, he read and admired G. W. F. Hegel, but was less in sympathy with Hegelianism than Bosanquet was. Bosanquet was active in social reform, as Green had been, whereas Bradley was a Tory who hated liberalism and sometimes thought along the lines of Thomas Carlyle's later writings. Bradley was, and intended to be, a highly polemical writer. His *Ethical Studies* and *Principles of Logic* are a sustained attack on the utilitarianism and empiricism of John Stuart Mill and his followers and upon the positivist outlook of the times. Later in his career, Bradley crossed swords with William James (who, however, greatly influenced Bradley's views on existence and reality) and with Bertrand Russell. His views were at their maximum influence during the first decade of the twentieth century, and the philosophical analysis of Russell and G. E. Moore arose largely in the attempt to refute them. Bradley's literary style has been much admired, notably by T. S. Eliot, who, as a graduate student at Harvard, studied Bradley in detail and wrote a thesis about him. Few if any other works on logic have been written with the verve, eloquence, and exuberant clarity of Bradley's *Principles*, but *Appearance and Reality* is less varied, and, from a stylistic point of view, much less successful.

ETHICS

Bradley's *Ethical Studies* is the most Hegelian of his writings. There is much criticism in it of Mill and some criticism of Immanuel Kant. There are amusing skirmishes with Matthew Arnold and with Frederick Harrison, the English positivist. Running through the book is the idea

that it is not for the moral philosopher to tell people what to do, but rather to dispel false views of the nature of morality and to provide an analysis of morality that can stand up to philosophical criticism. Thus he starts with an analysis of the moral concepts of the plain man, which, he holds, are not consistent with utilitarian views on punishment and responsibility. He goes on to criticize hedonism, largely on the ground that since pleasure is an "infinite perishing series," it cannot be the object of a rational pursuit. (The influence of Hegel's doctrine of the False Infinite is apparent here.) As to utilitarianism, Bradley holds that in the light of the Greatest Happiness Principle *any* course of conduct *might* conceivably be right, and "this is to make possible, to justify, and even to encourage, an incessant practical casuistry; and that, it need scarcely be added, is the death of morality."

Like Hegel, Bradley considered Kantian ethics to be formal and abstract, and, again like Hegel, he endeavored to supplement Kant's theories by a more concretely social view of ethics. In the study "My Station and its Duties" he developed the concept that Hegel had called "social morality" (*Sittlichkeit*). According to this view, duties are determined by the agent's place and functions in society. Bradley argued, furthermore, that men themselves are what they are because the society in which they are born and bred is what it is. The "individuals" of liberal and utilitarian social theory do not exist. The community is not, as the liberals assumed, a mere collection of individuals who are logically prior to it, but is a real being "and can be regarded (if we mean to keep to facts) only as the one in the many."

This language shows that Bradley regarded communities as both real and as concrete universals, and individual men as factually and logically dependent upon them, a view that was to achieve logical status in the *Principles of Logic*. Bradley wrote of morality as "self-realization," and some writers have therefore classed him as an ethical egoist. But the self that realizes itself is, according to Bradley, a socialized self that expresses and develops itself in making its contribution to the whole. It should be noted (and here again he is following Hegel) that Bradley did not regard "my station and its duties" as the culmination of morality. He held that on the basis of social morality other forms are developed. In pursuing science or in producing works of art, people are not confined to any particular station, and they also set themselves ideals that go beyond what mere duty would require of them. Perhaps humankind is the beneficiary in such cases, but humankind is not a being or community (this is in criticism of the positivists) in the way that a state or a nation is. Thus, on the basis of "the objective world of my station and its duties" ideals of social and of nonsocial perfection are constituted. These various spheres and duties often clash with one another, but the moral philosopher cannot formulate rules (as the utilitarians thought they could) that would enable the clashes to be avoided or settled. Conflict and failure are inseparable from morality, which could not exist without them.

The *Ethical Studies* are impressive today by virtue of the anticipations in them of twentieth-century views on socialization and the formation of conscience. But Bradley's position is different from that of present-day sociologists in that he thought that the plain man's views on responsibility are superior to any utilitarian reformulation of them and that they presuppose a nonatomistic metaphysics. The facts of moral judgment and of moral action, he held, force the philosopher to a monistic view of social life and to a metaphysics of the self as a being that can be itself only by transcending itself.

LOGIC

In his *Principles of Logic*, Bradley endeavored to refute false views of the subject without going thoroughly into questions of epistemology and metaphysics. The main objects of his attack were: the traditional subject-predicate, syllogistic, formal logic; the inductive logic with which, since the appearance of Mill's *Logic*, this traditional logic had been supplemented; and the confusion he claimed to see in the current empiricist logic between logical and psychological problems.

Bradley thought that the traditional logic was inadequate and incomplete. For example, in treating all judgments as of the subject-predicate form it omitted relational judgments, and the doctrine of the syllogism failed to take account of relational arguments. He maintained, too, that universal affirmative judgments are not categorical but hypothetical, since they do not necessarily assert that there are members of the subject class. These are theses that subsequent logicians have accepted.

Bradley denied that the advance of knowledge was from particulars to universals, or from particulars to particulars as Mill had suggested. Hence he denied the existence of induction as understood by Mill and the writers of textbooks who followed him. The great mistake of the empiricists, Bradley argued, was to suppose that thought could possibly get started with knowledge of separate and independent particulars. Such particulars, in his view, could be known only after a preceding condition of vagueness, ambiguity, and generality. This, however, is a historical, not a logical, consideration. Bradley's main

argument is that inference is possible only on the basis of universals and hence cannot be a procession from particulars to particulars or from particulars to universal. Inference presupposes judgments and ideal contents, and these, in their turn, presuppose generality and universality. It is only legitimate to argue from some to all if it is known or surmised that the particulars share some universal character. Bradley supported this by a detailed examination of Mill's inductive methods, an examination that owes something, as Bradley acknowledged, to William Whewell's criticism of them in his *Philosophy of Discovery*. The main point is that the facts or particulars from which the induction is alleged to start must already be ordered and defined in terms of some sort of theory, and hence in terms of a universal, if they are to give rise to an advance in knowledge. Both premises and conclusion must be organized around the central concept in a system of related concepts.

The empiricists subordinated logic to psychology. David Hume's account of thought was in terms of ideas that, by the very fact of being described as "fainter" than impressions, were regarded as a sort of mental image. Based on Hume's views, there had grown up a theory that knowledge advanced by the association of ideas. Bradley set out to refute this view, which today is known as psychologism. He argued that logicians are not concerned with ideas as psychical facts, but with ideas as meanings. As meanings, ideas do not have dates and histories, but are "ideal contents" and hence abstract. The real distinction between subject and predicate, he argued, is not to be found in the relation of one ideal content to another but in the relation of a complex ideal content to the reality to which it is referred.

In judgment, therefore, an ideal content is referred to a reality existing beyond the act of judgment. The real subject of a judgment is thus often quite different from the grammatical subject of the sentence, as can be seen in such an example as "A four-cornered circle is an impossibility," where the real subject is not a four-cornered circle, for there could be no such reality, but the nature of space. (This distinction between the grammatical form and the logical form was later to play an important part in analytic and linguistic philosophy.) If this view is accepted, then psychological accounts of inference fare no better than psychological accounts of judgment, since it is meanings, not psychical occurrences, that are relevant. There could not be any association between particular mental occurrences since they perish as they pass, and past ones would have somehow to be revived or re-created if they were to be associated with those existing in the present. Thus similarity and reproduction presuppose universals, just as inference itself does.

We have said that in his *Logic* Bradley tried to avoid being drawn into epistemological and metaphysical discussions. It is not surprising that he failed in this. Part of his attack on the "School of Experience" consisted in his bringing to light the untenable atomistic metaphysics that he regarded as basic to it. This is a parallel operation to his assault on utilitarianism. The claim that scientific knowledge is based on a prior knowledge of facts or particulars he rejected on the ground that from atomistic particulars no inference could be made. No inference could be valid apart from identities or universals linking one fact with another. It is clear, therefore, that Bradley thought that the fact of inference invalidated metaphysical pluralism, as the facts of morality went against it too. At this point Bradley has some important things to say about universals. He takes the view that what is essential to universality is identity in difference. Identity in difference can take two main forms. It can be abstract, as with such adjectives as "red" or "hard," which require substances in which to inhere. Or it can be concrete, as with an individual man, who is identical throughout his many actions, or a community, which persists through many generations of inhabitants. Abstract universals, therefore, are dependent, insubstantial, unreal, whereas concrete universals are (relatively) independent, substantial, and real. If what is real is individual, then concrete universals are individuals. Bradley ends this part of the discussion with the words: "It might be urged that if you press the enquiry, you will be left alone with but a single individual. An individual which is finite or relative turns out to be no individual; individual and infinite are inseparable characters." He does not pursue this in the *Logic*, but says that such a "revision" (an interesting choice of words) "must be left to metaphysics." So it is to his metaphysics that we now turn.

METAPHYSICS

Bradley's metaphysics, apart from the glimpses of it given in the *Ethical Studies* and the *Logic*, is set out in *Appearance and Reality* and in *Essays on Truth and Reality*. The main argument of *Appearance and Reality* is quite simple. It is divided into two books. The first and shorter one is titled "Appearance" and is about the contradictory character of mere appearances. Book II is titled "Reality" and is about the Absolute.

In Book I, certain commonsense concepts, such as relation, cause, space, time, thing, and self, and certain philosophical concepts, such as the thing-in-itself and the

distinction between primary and secondary qualities, are declared to be self-contradictory and are in consequence "degraded to the rank of mere appearances." In Chapters 2 and 3 of Book I, titled, respectively, "Substantive and Adjective" and "Relation and Quality," Bradley argues that the very notion of a relation is self-contradictory and that this inconsistency is alone sufficient to condemn "the great mass of phenomena," since space, time, causation, the self, all imply relations.

In Chapter 2, in considering the suggestion that all things are groups of related attributes, Bradley argues that if *A* and *B* stand in relation to *C*, then *C* must be related to *A* and *B* by another relation *D*, and this by a third relation *E*, and so on indefinitely. In Chapter 3 he argues that if simple qualities are to be conceived, they must be conceived as related to one another; but if *A* is related to *B*, then there must be the independent aspect of *A* and the aspect in which it is related to *B*, and hence it cannot be simple; but if *A* is not simple, then the independent aspect and the aspect in which it is related to *B* must be related to one another, so that there is set up in each of them a further plurality of aspects generating what Bradley calls "a principle of fission which conducts us to no end."

In Book II, it is argued that if it is being self-contradictory that degrades mere appearances, then reality must at least be not self-contradictory, but consistent and harmonious. Furthermore, reality must also be of the nature of experience, for what is not experience cannot be conceived of without self-contradiction. Finally, it is clear that reality must be comprehensive and include all that is. If reality is a consistent and harmonious and all-inclusive experience, then it cannot be a plurality of independent reals, for whatever is related to anything else must be to some extent dependent on it. "Plurality and relatedness are but features and aspects of a unity." Furthermore, the sort of unity that reality or the Absolute must have may be understood by analogy with feeling or immediate experience, for here there is diversity without relatedness.

According to Bradley, our experience of related things arises out of a prior immediate experience in which there are felt differences but no distinct qualities, and therefore no conception of things with different qualities in relation with one another. In passing from the primitive harmonious vagueness to a knowledge of related things, we pass from what might be called the state of precognitive innocence to the flawed world of contradiction. Wherever there is thought, there is the distinction between the what and the that, between ideal content and reality, between adjective and substantive; and hence wherever there is thought, there is contradiction. Thus reality, or the Absolute, must transcend thought, and thought always points beyond itself to something in which "mere thinking is absorbed." The Absolute must be conceived as analogous to immediate experience but transcending thought rather than falling short of it.

It is clear that contradiction, error, and evil are not harmonious and hence are not real, but it is equally clear that they are not nothing. How then must they be considered in the light of the Absolute? To this question Bradley gives a very interesting answer. He says that although error and evil are discordant and hence not real, it is possible that they contribute to the harmony of the whole, and if this is possible then we must conclude that it is so even though we do not know how it is possible. "For what is *possible*," he says, "and what a general principle compels us to say *must be*, that certainly *is*" (*Appearance and Reality*, Ch. 16). In this way, he protects himself against demands to show exactly how appearances are self-contradictory, unreal, not nothing, and yet are elements in the total harmony. Even so, he does make some attempts to show how all this is possible. In Book I, for example, time is condemned as self-contradictory, but in Book II Bradley says that although it is not real it nevertheless exists.

In explaining what he means by existence, he says it consists in being an event in time, in being a fact, in being directly perceived. In a later essay he says that what exists is what is continuous with our waking body. Existence, therefore, is the mode of being of the phenomenal world. But this would seem to bring us back to the point from which we started. Bradley also says that the real, the Absolute, must appear in what exists, that it cannot remain unmanifested. But he also attempts to mitigate the dualism between harmonious reality and self-contradictory appearance-existence by sketching a scheme in which reality permits of degrees. At the bottom of the scale, there are sheer contradictions and the abstract being of lifeless matter. Organic matter has more reality and is higher in the scale, and mind is higher still, for in mind the whole is immanent in its manifestations and the manifestations express the whole.

It is in mind that we see how the real must appear. But insofar as mind is thought, it suffers the disruption into the what and the that, which we have already considered. Perhaps, then, reality is to be found in mind as practical. This is rejected on the ground that practice essentially contains the distinction between reality as it is and reality as it will be when altered. Reality cannot be found in aesthetic experience either, for art entails pleas-

ure, pleasure is an experience of selves, and selves, Bradley has argued, cannot be ultimately real. "The Absolute," Bradley concludes, "is not personal, nor is it moral, nor is it beautiful or true." Yet in spite of all this he ends *Appearance and Reality* with the words: "… the more that anything is spiritual, the more is it veritably real."

The weakest part of *Appearance and Reality* is Book I. The amount of space and care given in it to the task of discrediting the whole of common sense and much of the philosophy of the past is trifling compared with the magnitude of the desired result. Bradley seems almost to take the reader's agreement for granted and to hasten on to the more congenial, yet only slightly more constructive, task of showing what the Absolute must be. A good part of the argument of Book I assumes that predication is identity, in accordance with "the old puzzle how to justify the attributing to a subject something other than itself." After all, Bradley had refuted this view of predication in his *Logic*. Perhaps then he is arguing dialectically, in order to bring out the unhappy consequences of working with this "logic of identity." But if this were so, then relation, space, time, the self, etc. would only be self-contradictory if looked at in the light of a false logic, and might be reinstated if the true logic were brought to bear on them. The doctrine of degrees of reality goes some way towards meeting this difficulty. But in Book I there is no indication that the self is more real or less self-contradictory than space and time. As A. S. Pringle-Pattison put it in his review of *Appearance and Reality:* "Mr. Bradley has the aim of swallowing at a gulp in Book II what he had choked over in the successive chapters of Book I."

As to Book II there are two main defects. One is that the Absolute described in it seems to be without any definite features but is an amorphous refuge into which appearances are "fused," "transformed," "transmuted," or "dissolved." The other is that in the course of developing the doctrine of degrees of reality Bradley unwittingly reverts on occasion to the arguments of Book I, as when he says that aesthetic experience cannot be or reveal the Absolute since it involves pleasure and selves and selves are self-contradictory. Bradley here seems to be reverting to the logic of identity that in Book II he had been moderating. On the other hand, there is much excellent discussion of details. The account of time is particularly good. Bradley holds that we should not think in terms of one time series only, but in terms of several or many. Just as the events of one fiction are not temporally related to the events in another fiction, so there may be various time series in which what is past in one may be yet to come in another.

What Bradley said about time and about existence and reality greatly exercised G. E. Moore who, in various writings, notably "The Conception of Reality" (1917–1918), endeavored to make clear what it is to say that something exists. Moore argued that Bradley's view that time, although unreal, must exist, depended upon his assuming that whatever can be thought of must somehow exist in order to be thought of. But Moore rejected this assumption. Bradley, he thought, was deceived into making it because he did not notice that although "unicorns are objects of thought" is of the same *grammatical* form as "lions are objects of the chase," it is of quite a different *logical* form. Moore's reason for this was that if lions are to be hunted there must be lions, whereas unicorns can be thought of although there are no unicorns. Thus Moore used against Bradley the distinction between logical and grammatical form that Bradley had formulated in 1883. A weapon that Bradley had himself devised was employed against him by a philosopher who had improved its range and sophistication.

See also Absolute, The; Analysis, Philosophical; Appearance and Reality; Arnold, Matthew; Bosanquet, Bernard; Carlyle, Thomas; Eliot, Thomas Stearns; Ethics, History of; Green, Thomas Hill; Hegel, Georg Wilhelm Friedrich; Hegelianism; Hume, David; Idealism; James, William; Kantian Ethics; Kant, Immanuel; Logic, History of; Logic, Traditional; Mill, John Stuart; Moore, George Edward; Pringle-Pattison, Andrew Seth; Psychology; Russell, Bertrand Arthur William; Whewell, William.

Bibliography

Bradley's main writings and the dates of publication have been indicated in the body of the article. A second edition of *Ethical Studies* (Oxford: Clarendon Press, 1927) contains corrections and additional notes that Bradley left at his death. A second edition of *Principles of Logic* (London: Oxford University Press, 1922) contains the original unaltered text with an extensive commentary, which owes much to Bosanquet, at the end of each chapter and a set of "Terminal Essays" at the end of the book. The second edition of *Appearance and Reality* (London, 1897) contains an appendix occasioned by criticisms of the first edition. See also *Aphorisms* (Oxford: Clarendon Press, 1930), a few of which appeared in the preface to *Appearance and Reality*. Collections of his works include *Collected Essays* (Freeport, NY: Books for Libraries Press, 1967) and *Essays on Truth and Reality* (Oxford: Clarendon Press, 1968).

For a detailed critical study of Bradley's philosophy, see Richard Wollheim, *F. H. Bradley* (Harmondsworth, U.K.: Penguin, 1959). This contains further bibliographical references. See also C. A. Campbell, *Scepticism and Construction* (London: Allen and Unwin, 1931); R. W.

Church, *Bradley's Dialectic* (London: Allen and Unwin, 1942); and T. S. Eliot, *Knowledge and Experience in the Philosophy of F. H. Bradley* (New York: Farrar, Straus, 1964), the thesis mentioned above.

A. S. Pringle-Pattison, "A New Theory of the Absolute," reprinted in *Man's Place in the Cosmos* (Edinburgh and London: Blackwood, 1897, 1902), is probably the best criticism of Bradley's metaphysics. G. E. Moore, *Some Main Problems of Philosophy* (London: Allen and Unwin, 1953), Chs. 11, 12, 16, contains a detailed discussion of parts of Bradley's metaphysics.

Criticism of Bradley's view that the notion of relation is self-contradictory is contained in J. Cook Wilson, *Statement and Inference* (Oxford: Clarendon Press, 1926), Vol. I, p. 255, Vol. II, pp. 692–695. See also W. H. Walsh, "F. H. Bradley," in *A Critical History of Western Philosophy*, edited by D. J. O'Connor (New York: Free Press of Glencoe, 1964).

OTHER RECOMMENDED WORKS

Bradley, F. H. *Selected Correspondence, January 1905–June 1924*, edited by Carol A. Keene. Bristol, U.K.: Thoemmes Press, 1999.

Bradley, F. H. *Collected Works of F. H. Bradley*, edited by Carol A. Keene. Bristol, U.K.; Sterling, VA: Thoemmes Press, 1999.

Horstmann, Rolf-Peter. *Ontologie und Relationen: Hegel, Bradley, Russell und die Kontroverse über interne und externe Beziehungen*. Athenäum: Hain, 1984.

Ingardia, Richard. *Bradley: A Research Bibliography*. Bowling Green, OH: Philosophy Documentation Center, 1991.

Mander, W. J. *An Introduction to Bradley's Metaphysics*. Oxford: Clarendon Press; New York: Oxford University Press, 1994.

Mander, W. J. *Perspectives on the Logic and Metaphysics of F. H. Bradley*. Bristol: Thoemmes, 1996.

Manser, Anthony Richards. *Bradley's Logic*. Totowa, NJ: Barnes & Noble Books, 1983.

Manser, Anthony Richards, and Guy Stock. *The Philosophy of F.H. Bradley*. Oxford; New York: Clarendon Press, 1984.

Shrivastava, S. N. L. *Sámkara and Bradley: A Comparative and Critical Study*. Delhi: Motilal Banarsidass, 1968.

Sprigge, Timothy L. S. *James and Bradley: American Truth and British Reality*.Chicago, IL: Open Court, 1993.

H. B. Acton (1967)
Bibliography update by Michael J. Farmer (2005)

BRADWARDINE, THOMAS

(c. 1300–1349)

Thomas Bradwardine studied arts at Balliol College and theology at Merton College, Oxford. In September 1337, he was appointed chancellor of Saint Paul's in London. From 1346 to 1348, as a king's clerk, he enjoyed a prominent position in the household of Edward III. In June 1349 he was elected archbishop of Canterbury; soon afterwards, in October, he died of the Black Death.

Like many Mertonians, Bradwardine was a logician and a mathematician. He wrote a treatise *De insolubilibus* (an *insolubile*is a self-referential sentence, such as the "liar paradox"), *a Geometria speculativa*, and a treatise *De continuo*. In his *Tractatus de proportionibus velocitatum in motibus* (1328) he attempted to introduce mathematic functions into Aristotelician physics. His masterpiece, however, is a voluminous theological and philosophical work, *De causa Dei contra Pelagium*, divided into three books (1344). It originates from lectures he had given in Oxford and London and, more radically, from a deep spiritual change he had experienced in his youth: "When I was applying myself to philosophy … Pelagius's opinion seemed to me nearer to truth.… But afterwards (I was not yet a theological student) … I thought I saw from afar the grace of God preceeding all merits in time and in nature, in the same way that in all movements He is the first Mover." (bk. I, ch. 35, p. 308). This conversion induced Bradwardine to fight for God's cause against "the new Pelagians, " a group of post-ockhamists theologians that included Richard Fitzralph, Adam Wodeham, and Robert Holcot.

To these thinkers the issues of chief concern were grace and merit, future contingents, prescience, and predestination. On the first point, Bradwardine, as an ardent Augustinian, strongly reasserts that grace is a mere gift, not a retribution: in no way man can merit it, and, moreover, without God's special help man cannot act right.

Concerning future contingents, the new Pelagians' opinion stressed the contrast between the necessity—that is, the fixity—of the past and the contingency of the future. This view could hardly be reconciled with the idea of an immutable and truthful God: If God or a prophet were to reveal a future event, is it possible that it would not happen? If it is possible, then God can deceive and lie. Countering this opinion, which he had first rejected in his question, *De futuris contingentibus*, Bradwardine closely examines the notions of contingency and necessity; he argues they are founded on the power of the will. Aristotle wrote, "What is, necessarly is, when it is. (*De interpretation*, ch. 9). But Duns Scotus observed that when man wills *A* at time *t*, he has the power not to will *A*, not only before or after *t*, but also at time *t*. Therefore a kind of necessity, the "consequent" necessity of present, is compatible with contingency. Regarding God, Bradwardine extends this conclusion to all times: For God, past, present, and future are equally contingent and equally necessary. Consequently He can undo any past event (in an improper meaning of *undo*), not because He could alter it (this would be a contradiction), but because at each instant of time He is yet freely willing the past event. In this way, there is no longer antinomy between the necessity of the prophecy and the contingency of the future event.

The same argument about contingent causality clears up the most famous tenet of Bradwardine's teaching, the assertion of "antecedent necessity": Since God's will is the first cause of everything and cannot be thwarted, everything happens by necessity in relation to His will. That is the proper definition of theological determinism. But again, according to Bradwardine, when man is willing something, though his act is determined by God, he does not lose the power to do the opposite act at the same time. So it seems there is in Bradwardine's doctrine an original attempt to conciliate God's predetermination and human freedom of will.

See also Determinism and Freedom; Duns Scotus, John; Pelagius and Pelagianism; Precognition.

Bibliography

WORKS BY BRADWARDINE

De causa Dei contra Pelagium. Franfurt am Main: Minerva, 1964. This is a reprint of a work that was originally edited by Henry Savile in London and published in 1618.

De proportionibus velocitatum in motibus. In *Thomas of Bradwardine: His* Tractatus de Proportionibus*—Its Significance for the Development of Mathematical Physics*, edited by Lamar H. Crosby Jr. Madison: University of Wisconsin Press, 1955.

"Insolubilia." In "La problématique des propositions insolubles au xiiiesiècle et au début du xive, suivie de l'édition des traités de W. Shyreswood, W. Burleigh et Th. Bradwardine, " edited by Marie-Louise Roure. *Archives d'histoire doctrinale et littéraire du moyen âge*, 37 (1970): 205–326.

"De futuris contingentibus." In "Le *De futuris contingentibus* de Thomas Bradwardine, " edited by Jean-François Genest. *Revue des études augustiniennes*, 14 (1979): 249–336.

WORKS ON BRADWARDINE

Leff, Gordon. *Bradwardine and the Pelagians: A Study of His "De Causa Dei" and Its Opponents*. Cambridge, MA: Cambridge University Press, 1957.

Oberman, Heiko *A. Archbishop Thomas Bradwardine: "A Fourteenth Century Augustinian."* Utrecht: Kemink & Zoon, 1958.

Genest, Jean-François. *Prédétermination et liberté créée à Oxford au xive siècle. Buckingham contre Bradwardine*. Paris: J. Vrin, 1992.

Sylwanowicz, Michael. *Contingent Causality and the Foundations of Duns Scotus' Metaphysics*. Leiden: E. J. Brill, 1996.

Jean-François Genest (2005)

BRAHMAN

The origin and meaning of the term *brahman*are shrouded in mystery. Using the verbal root √*bṛh*, Western Indological scholars derive such meanings as "sacred magical power" (Hermann Oldenberg), "form, formulation" (Paul Thieme), "priestly utterance," "energy that is expressed in paradoxical terms" (Louis Renou), and "the live connection that holds the cosmos together" (Jan Hesterman) The meanings of *brahman* in the ancient "heard texts" (*śrutis*) and later Indian philosophical systems are not unrelated to these meanings. For example, the Vedic understanding of the *brahman* survives in Bhartrhari's concept of the "*śabda brahman*." Likewise, the ideas of power, energy, and cosmic unity among opposites are taken up in the Vedāntic notion of the *brahman* as absolute reality. The notion of the *brahman* as the sacred power within a priest may have contributed to an identification of the brahman with the inner spirit (*ātman*). This transformation of a much older notion into a discursively idealized philosophical concept resembles the way the concept of *logos* was transformed into "logic" "*Vernunft*," and "language."

Etymologically, the word *brahman* has two constituent components: the verbal root √*bṛh* and the suffix *matup*. The verbal root √*bṛh* means "to grow" and "the great," and together with the suffix provides two allied meanings: "the greatest" and "the root of all things." In the Vedic hymns the term *brahman* not only refers to the power contained in the words recited but also to the mysterious power present in the utterances of the Vedic hymns. Though the idea of *brahman* as the ground of all things is not entirely absent in the Vedas, the primary goal was to search for the power connecting the microcosm with the macrocosm.

BRAHMAN IN THE UPANIṢADS

This sense of power continues in the Upaniṣads (e.g., *Kaṭha Upaniṣad)*, which say that the various *devas* (gods; literally, "the shining ones") each carry out their respective jobs for fear of the brahman (6.3); *Kena Upaniṣad* states that the various *devas* have no power outside the power of *brahman* residing in them. The *brahman* of the Upaniṣads is much more than a power; it is the cause of the origination, sustenance, and destruction of the world (*Taittiīya Upaniṣad*, 3.1.1). In the *Bṛhadāraṇyaka Upaniṣad*, when Yājñavalkya is questioned about the number of gods, he initially responds by saying that 3,306 gods were simply manifestations of thirty-three gods, and then successively reduces the number to six, three, two, one and a half, and then one. This god is none other than the *brahman*, and all other gods of the Vedas, the senses, and the mind are said to be the various powers of *brahman* (*Bṛhadāraṇyaka Upaniṣad*, 3.9.1–10). The central

question in the Upaniṣads is framed as follows: "What is that by knowing which all else becomes known?" (*Muṇḍaka Upaniṣad*, 1.13) The answer given is "*brahman*." If *brahman* is the source of everything, then *brahman* is also the core of each individual being, and this core is called *ātman*. In many places in the Upaniṣads the two terms *brahman* and *ātman* are used synonymously. The *Chāndogya Upaniṣad* asks: "What is *ātman*? What is *brahman*?" (5.11.1, *ko nu ātmā, kiṃ brahmeti*?) When the inquiry pertains to the source of the universe, the word *ātman* is used, and in other cases when the inquiry is regarding the true self of a human being the word *brahman* is used. To the Upaniṣadic seers the *brahman* and the ātman signified the same reality, one within, and the other without.

The Upaniṣads describe *brahman* both negatively and positively. It is described as neither gross, nor subtle, nor short, nor long, nor red, nor adhesive, without shadow, darkness, air, space, attachment, taste, smell, eyes, ears, speech, mind, light, breath, mouth, and measure, and without inside and outside (*Bṛhadāraṇyaka Upaniṣad*, 3.8.8), and that who consists of mind, whose body is life, whose form is light, whose conception is truth, whose soul is space, containing all works, desires, odors, and tastes, and encompassing the whole world, the speechless and the calm (*Chāndogya Upaniṣad,* 3.14.2).

BRAHMAN IN VEDĀNTA

The exegetical effort to construe these different groups of sentences to resolve any apparent contradiction shaped the understanding of *brahman* in the Upaniṣadic schools. Two hermeneutic principles were applied: accord priority and finality to the positive texts, or since negation implies a prior affirmation that is then negated, the final import of the Upaniṣads may be taken to be expressed in the negative texts, the positive ones simply preparing the ground for it. The latter hermeneutical principle is adopted by Śaṃkara, the most well-known exponent and defender of the school known as Advaita Vedānta (nondualistic Vedānta); and the former by Rāmānuja, the founder of Viśiṣṭādvaita (qualified nondualism), and Madhva in his Dvaita Vedānta.

ADVAITA VEDĀNTA

From the perspective of Śaṃkara's nondualistic Vedānta, *brahman* is one without a second; the world is false (*māyā*, in a rather technical sense of "presented appearance") and the finite individual and the *brahman* are nondifferent. The *brahman* can neither be comprehended by rational minds, nor be expressed or literally referred to in the language, nor be an object of knowledge. It does not have qualities (because all determination is negation), and so it cannot be described or defined. While using language to refer to *brahman* is natural, it does not achieve its goal. Language refers to an object either by its direct power of meaning (*abhidhā*), or as its suggested meaning (*lakṣaṇā*). Normally, the suggested meaning is sustained and supported by its relationship to the literal meanings, but in the case of language referring to the brahman, the meaning may be said to be "only the meaning function, but not an actual meaning" (Bhattacharyya 1930). There is a pointing, as one points to something with one's finger, toward a small, almost invisible star, accompanied by a series of descriptions each one of which is then canceled, leading the listener to identify, even in the absence of an identifying description of, what is being pointed at. *Brahman* is described as *saccidānanda*, that is, as *sat* (existence), *cit* (consciousness), and *ānanda* (bliss), with reference to its essence (*svaūpa lakṣaṇā*), whereas *brahman* as the cause, sustainer of the universe, and so on with reference to its accidental attributes (*taṭastha lakṣaṇā).*

It is important to keep in mind that from a strictly Advaita point of view no positive description can be intrinsic when the thing being described lacks any positively determining qualities. Nevertheless, Advaita Vedānta describes the *brahman* as existence, pure consciousness, and bliss. These three are not qualities or qualifying attributes of the *brahman*. Advaita Vedānta holds that these familiar terms must be understood in their negative implications, not as referring to what *brahman* is, but rather as pointing to what the *brahman* is not. *Sat* points to the fact that the *brahman* is not *asat* (nonexistent); cit suggests that the *brahman* is not *acit*, that is, *jaḍa* (insentient matter); and ānanda points to that, in the experience of the *brahman*, there is no *duḥkha*, no unsatisfied desire. The negative statements in this regard more closely approach the intrinsic nature of the *brahman*. In this light one can say that in Advaita Vedānta, Benedict (Baruch) de Spinoza's principle "all determination is negation" holds good of the infinite: no determination of it is possible. Underlying this account are a theory of meaning and a theory of language that are of particular importance, and, that possibly, found their first systematic exposition in the Buddhist theory of *apoha* (the negative theory of meaning).

The thesis of Advaita Vedānta is logically substantiated (1) by a critique of difference (*bheda nirodha*) and (2) hermeneutically by an exegesis of the *śrutis*. To these one may add (3) a phenomenological and experiential

dimension that would consist in showing that in its search for freedom from all suffering, human subjectivity passes through levels of ordinary experience: the waking-bodily, the dreaming-psychic, and the dreamless sleep (the seemingly total inaction and quietude). Finally, there is the experience of the *brahman*, which goes beyond the distinction of the subject and the object and which is articulated in such famous *mahāvākyas* (great sentences) of the Upaniṣads as "I am he" and "thou art that." Knowing the *brahman*, according to the tradition of Advaita Vedānta, is to become the *brahman*. It is not knowing an object, however large and great in its dignity, that stands over against one as an other; rather, it is an experience in which all otherness is canceled, and one discovers that within oneself nothing else remains to be achieved. When there is no duality between the subject and the object, there is no duḥkha or fear. A modern account of the phenomenological stages of a path to freedom is found in Krishna Chandra Bhattacharyya's *The Subject as Freedom* (1930).

VIŚIṢṬĀDVAITA VEDĀNTA

In his Viśiṣṭādvaita, that is, "one reality (*brahman*) with qualifications," Rāmānuja holds that all knowledge necessarily involves distinctions and differentiations. It is impossible to know an object in its undifferentiated form; therefore, both pure identity and pure difference are false. The *brahman* as God possesses cit (matter) and acit (self); all three are real (*brahman*, cit, and acit). Though real, the last two are dependent on the *brahman*. Consciousness presupposes the self of which it is an essential attribute (*dharmabhuta jñāna*). Perhaps the most original aspect of Rāmānuja's philosophy is the rejection of the principle that to be real means to be independent. Although soul and matter are substances in themselves, in relation to the *brahman* they are attributes. They are God's body and he is their soul. The self is substance and quality, an organ and organism of the *brahman*. Rāmānuja's theory of *apṛthaksiddhi viśeṣaṇa*, that is, the adjectival theory of inseparability, explains this relation. Just as qualities are real and cannot exist apart from the substances in which they subsist, similarly matter and soul are parts of the brahman and cannot exist without the *brahman*.

Rāmānuja used the same Upaniṣadic texts that Śaṃkara used, but arrived at a different conception of the *brahman*. Rāmānuja holds that the Upaniṣadic texts such as *neha nānā asti kiṃcena* (there is no multiplicity here) do not really deny the multiplicity of objects, names, and forms, but asserts that these objects do not have any existence apart from the *brahman*. Thus, all negative texts of the Upaniṣads, which assert that none of this is the brahman, are construed to mean that none of it in its presumed independence is the *brahman*. However, the positive sentences, for example, "all this is the *brahman*," mean that everything belongs to the *brahman* as the ultimate totality. Whereas for Śaṃkara the *brahman* is pure consciousness devoid of any distinctions, a pure identity without any difference (*nirguṇa*), Rāmānuja's brahman is identity-in-difference. The *brahman* creates the world out of acit by an act of will, so creation is a real act of will. Ignorance (māyā or *avidyā*) in this system is no longer creative of illusory world, and the finite individuals are not illusory. It is indeed true, Rāmānuja concedes, that some Upaniṣadic texts articulate the *brahman* as wielder of a magical power (*māyā*). However, *māyā* for Rāmānuja properly understood is the unique power of God by which God creates the wonderful world of objects. He vehemently criticizes Śaṃkara's theory of the world as false appearance. The created world, for Rāmānuja, is as wonderful as the *brahman* himself.

If someone were to ask how the one contains the many, Rāmānuja would respond with the grammatical principle of *sāmānādhikaṇya* (coordination). According to this rule, the words in a sentence with different meanings can denote one and the same thing. Rāmānuja's interpretation of the classical text "this is that Devadattaḥ" explains this rule clearly. For Rāmānuja, Devadattaḥ of the past and the Devadattaḥ of the present cannot be entirely identical, because the person seen at the present and the person seen in the past are different, have different meanings, and yet both refer to the same person. Similarly, unity and diversity, the one and the many, can coexist; they are not contradictions and they can be reconciled in a synthetic unity. Thus, he does not deny the many: the many, on the contrary, characterizes the one. *Mokṣa* comes about with the knowledge of the *brahman* together with devotion (*bhakti*).

DVAITA VEDĀNTA

Madhva carries much further the protest against the nondualism of Śaṃkara than Rāmānuja. Whereas for Śaṃkara the texts teaching difference have a practical value in that they steer one in the right direction and lead one to the real teaching of the Upaniṣads, that is, the teaching of nondifference, for Madhva the texts teaching difference convey the true import of the Upaniṣads. Substance is one of the ten categories that Madhva accepts. Out of the twenty substances that Madhva enumerates, he accepts, like Rāmānuja, three as the most important: brahman or God, matter, and selves.

Bheda (difference) is the central category in Madhva's philosophy. This is another way of saying that each object is unique; each object possesses its own nature, which accounts for one object's difference from another object. The *brahman* or God is the only independent reality. God has a divine body and is transcendent. However, since God is the inner controller of all souls, he is also immanent. God creates the world by his will and brings into existence the world of objects and selves. Objects and selves, though real, eternal, and irreducible to each other, are dependent on the first. At the time of the dissolution of the world, God transforms material objects into undifferentiated matter and selves into disembodied intelligences. It is important to note in this context that even in the state of dissolution God, matter, and selves remain distinct. Unlike Rāmānuja, for Madhva no two souls are alike. Thus, whereas Rāmānuja advocates qualitative monism and quantitative pluralism of souls, Madhva advocates both qualitative and quantitative pluralism of souls. Since the immediate cause of bondage is ignorance of the real nature of the *brahman* or God, the soul must acquire the knowledge of the real nature of God to attain mokṣa. It is important to remember in this context that knowledge by itself does not and cannot remove ignorance; knowledge is only a qualification for release, which in the final analysis depends on God's will. No matter how hard an aspirant may try, he or she cannot gain such an immediate knowledge, unless God chooses to reveal himself to him or her.

ŚAIVISM

Finally, apart from the previously discussed classical understandings of the brahman, there is another nondualistic school known as Śaiva Siddhānta. Of its many representative schools and thinkers, Kāshmir Śaivism of Abhinavagupta is the most well known. In his nondualism Abhinavagupta argues that *brahman* alone *is*. He painstakingly attempts to bridge the gulf between the one and its many phenomenal differences by positing many levels of consciousness, descending from the one to the many. *Māyā* or *avidyā* is now construed as a *śakti* or the power of the *brahman*-consciousness; consciousness is not a mere *prakāśa* (illumination) but also *śakti* (force). Indeed, the two in their difference are also one. While, on the one hand, Abhinavagupta wishes to preserve both the one and the many in the being of brahman, he makes it a graded dynamic process instead of using a static set of categories like the part and the whole. The one *brahman*-consciousness or pure *cit* objectifies itself into "I," and this power of self-objectification is called *vimarṣa śakti* (the reflective power), from which arises the power of referring to intentional objects that lie concealed within it. This process yields a domain of seemingly independent objects. Kāshmir Śaivism has been a major influence in the shaping of the concept of the "integral *brahman*" of Sri Aurobindo's philosophy.

Other systems of Indian philosophy do not advance a concept of the *brahman*. Although Sāṃkhya-Yoga seems to have had a theistic form in addition to the better-known atheistic form, it does not develop a concept of the *brahman*, nor do the Nyāya-Vaiśeṣika schools. The latter systems come perhaps closest to such a project when they substantiate their concept of God as an infinite self, all knowing, omnipresent, which is called *Iśavara* in the school. Despite the fact that the authors may cite many texts of the Upaniṣads in support of their theism, one misses in these schools any attempt to take into account the *śruti* texts in their totality.

To sum up: The *brahman* is the absolute reality in the school of Vedānta. The relationships of the one with the many preoccupied its thinkers, leading them to postulate a fundamental category to explain the connection. These categories range from pure identity (*tādātmya* having that as its self), *apṛthakasiddhi* (the relation of no separate existence), pure difference, and a progressive self-differentiation through self-objectification and intentionality. In the nondualistic Vedānta, the *brahman*, in the words of Georg Wilhelm Friedrich Hegel, is (Spinozistic) "the substance becoming spirit," bringing together two different concepts of monism into one, resulting in the position that the only reality is the spirit. The following crucial issues remained: How does the one spirit become many? How to understand self-differentiation? Where to locate the power of creativity? Does it belong to the *cit* or consciousness as an inalienable aspect, or is it an "other" to consciousness? In the latter case, the basic otherness is not a product of ignorance. However, can one escape this problem by saying as nondualist Vedānta says, that ignorance is not a real other and not a nonreal other? Is not this nonreality itself a creation of ignorance? Thus, dialectic of one and many seems to have had an interminable hold on Indian metaphysical theories.

See also Causation in Indian Philosophy; God in Indian Philosophy; Mind and Mental States in Indian Philosophy; Self in Indian Philosophy.

Bibliography

Abhinavagupta. *Parātriśikā-vivarana: The Secret of Tantric Mysticism*. Translated by Jaidev Singh; edited by Bettina Bäumer. Delhi, India: Motilal Banarsidass, 1988.

Apte, V. M., tr. *Brahma-sūtra Shānkara-Bhāsya*. Bombay, India: Popular Book Depot, 1960.

Bhattacharyya, Krishna Chandra. *Studies in Vedantism*. Calcutta, India: Calcutta University Press, 1909.

Bhattacharyya, Krishna Chandra. *The Subject as Freedom*. Bombay, India: G.R. Malkani, 1930.

Carman, John Braisted. *Theology of Rāmānuja: An Essay in Interreligious Understanding*. New Haven, CT: Yale University Press, 1974.

Hesterman, J. "Brahmin, Ritual, and Renouncer." *Weiner Zeitschrift Deutschen Morgenländischen* 8 (1964): 1–31.

Hume, Robert Ernest, tr. *The Thirteen Principal Upaniṣads*. 2nd ed. Madras, India: Oxford University Press, 1965.

Mahadevan, T. M. P. *The Philosophy of Advaita, with special reference to Bhāratītīrtha-Vidyāranya*. New Delhi, India: Arnold-Heinemann, 1976.

Murti, T. R. V. "The Two Definitions of Brahman in the Advaita." In *Studies in Indian Thought: Collected Papers of Prof. T. R. V. Murti*, edited by Harold G. Coward. Delhi, India: Motilal Banarsidass, 1983.

Oldenberg, Hermann. *Der Religion des Veda*. Stuttgart, Germany: J. G. Cotta, 1923.

Ranade, R. D. *A Constructive Survey of Upaniṣadic Philosophy: Being an Introduction to the Thought of the Upaniṣads*. 2nd ed. Bombay, India: Bharatiya Vidya Bhavan, 1968.

Sharma, B. N. K. *A History of the Dvaita School of Vedānta and Its Literature: From the Earliest Beginnings to Our Own Time*. 2nd ed. Delhi, India: Motilal Banarsidass, 1981.

Sivaraman, Krishna. *Śaivism in Philosophical Perspective: A Study of the Formative Concepts, Problems, and Methods of Śaiva Siddhānta*. Delhi, India: Motilal Banarsidass, 1973.

Thibaut, George, tr. *The Vedānta-Sūtras: With the Commentary by Śankarākārya*. 3 vols. Oxford, U.K.: Clarendon Press, 1890–1904.

Bina Gupta (2005)

BRAITHWAITE, RICHARD BEVAN

(1900–1990)

Richard Bevan Braithwaite, an English philosopher, was educated at King's College, Cambridge, where he studied physics and mathematics before turning to philosophy. Braithwaite was Knightbridge Professor of Moral Philosophy at Cambridge University. He served as the president of the Mind Association (1946) and of the Aristotelian Society (1946–1947). In the philosophy of science he made significant contributions on the nature of scientific theories and explanation, theoretical terms, models, foundations of probability and statistics, the justification of induction, and teleological explanations. He also wrote on subjects in moral and religious philosophy.

SCIENTIFIC THEORIES

Braithwaite defended the view that a scientific theory consists of a set of initial hypotheses, with empirically testable generalizations that follow deductively. To explain a generalization is to show that it is implied by higher level generalizations in the theory. Often, especially in the physical sciences, the initial postulates will contain so-called theoretical terms, such as *electron* or *field*, that refer to items not directly observable. To understand the meaning of such terms, as well as the logical structure of the theory, one must begin by considering the theory as a formal calculus; that is, as a set of uninterpreted formulas. A calculus designed to represent a specific theory will have to be interpreted, but not all at once and not completely: Meanings are directly given only to those formulas representing the lower order empirical generalizations, rather than to initial formulas containing theoretical terms. The latter are indirectly and partially interpreted by the former.

Braithwaite's major contribution here consisted in the detailed attention he devoted to the nature of the initial or "theoretical" postulates. He divided these postulates into Campbellian hypotheses, which contain only theoretical terms, and dictionary axioms, which relate theoretical terms to observational ones. The latter include identificatory axioms, which identify single observational terms with theoretical terms—for example, a color word with expressions referring to wavelengths of light. Braithwaite argued that the advantage of systems containing theoretical terms over those whose initial postulates are entirely observational is that the former can more readily be extended to new situations than can the latter. However, Braithwaite held there is no special advantage to Campbellian hypotheses, because, at least for certain systems, the same testable consequences can be derived from identificatory axioms.

Scientific models are to be construed as alternative interpretations of a theory's calculus where the theoretical concepts in the original theory (such as molecules) are interpreted as designating more familiar and intelligible items (such as billiard balls). Accordingly, the theory and the model are to be distinguished; and while a model is not essential, it can sometimes be of help in extending a theory and clarifying its concepts.

PROBABILITY AND INDUCTION

Braithwaite proposed a novel finite-frequency theory of probability. Consider the statement (P), "The probability of a child being born a boy is 0.51," and the observed data that among 1,000 children 503 are boys. Such a situation is to be understood by imagining 1,000 sets of children, each containing 100 children of whom 51 are boys, and a selection of 1 child from each of the 1,000 sets, of whom 503 are boys. Since P is logically consistent with any observed data, the problem is to decide when to reject P. For this purpose it is necessary to have a rule specifying that a probability statement is to be rejected if the observed relative frequency differs from the probability postulated by more than a specified amount. This amount is determined by extralogical considerations involving the purpose for which the hypothesis is to be used and the value attached to possible consequences of its adoption. Such a rejection rule, Braithwaite claimed, is what gives empirical meaning to probability statements considered as constituents of theoretical systems. But suppose there are alternative probability hypotheses not rejected by the evidence in accordance with this rule. How is one to choose among them? Here again considerations of value must be invoked, and Braithwaite outlined a "prudential policy" of choosing the probability hypothesis that maximizes the minimum mathematical expectation of value.

Braithwaite also provided an original defense of Charles Sanders Peirce's solution to the problem of justifying induction. The problem was formulated by Braithwaite as follows: What warrant does one have for adopting the policy of accepting a hypothesis on the basis of many positive instances (the policy of "induction by simple enumeration")? The proposed answer consists of the following argument (where π is the principle of induction by simple enumeration): The policy of using π has been effective in many instances in the past; therefore (using π as the rule of inference) π will continue to be effective. Such an argument was traditionally dismissed as viciously circular, and Braithwaite undertook to prove this charge unjustified. The argument can be deemed valid and hence free from circularity, he claimed, because it enables one to pass from a mere belief in the general effectiveness of using π as a rule of inference, with a reasonable belief in π's past effectiveness, to a reasonable belief in π's general effectiveness. It would be viciously circular only if one were required to have an initial reasonable belief in π's general effectiveness. Since this requirement is unnecessary, the argument is not invalidated.

MORAL AND RELIGIOUS PHILOSOPHY

Many of the conclusions and techniques of the philosophy of science were applied by Braithwaite in areas of moral and religious philosophy. Thus, just as one can defend the adoption of a particular scientific hypothesis by appeal to an inductive policy, so one can justify a particular action, such as returning a book, by reference to a moral policy, such as promise-keeping. Both sorts of policies are in turn justified by reference to the ends they subserve. Braithwaite showed how the mathematical theory of games, which he invoked in his discussion of hypothesis selection, can also be used to shed light on such notions as prudence and justice in situations involving human choices and cooperation between individuals. Finally, just as a moral assertion is to be construed as an expression of an intention to act in accordance with a certain policy, so a religious assertion must be understood, according to Braithwaite, as a declaration of adherence to a system of moral principles governing "inner life" as well as external behavior. The major difference between religious and moral assertions is that the former, being associated with empirical narratives, have a propositional element lacking in the latter.

See also Explanation; Game Theory; Induction; Knowledge and Belief; Peirce, Charles Sanders; Philosophy of Science, History of; Teleology.

Bibliography

Black, Max. Review of *Theory of Games as a Tool for the Moral Philosopher*. *Philosophical Review* 66 (1957): 121–124.

Hirst, R. J. Review of *Scientific Explanation*. *Philosophical Quarterly* 4 (1954): 351–355.

Russell, L. J. Review of *Scientific Explanation*. *Philosophy* 20 (1954): 353–356.

WORKS BY BRAITHWAITE

"Propositions about Material Objects." *Proceedings of the Aristotelian Society* 38 (1937–1938): 269–290.

"Teleological Explanation." *Proceedings of the Aristotelian Society* 47 (1946–1947): 1–20.

"Moral Principles and Inductive Policies." *Proceedings of the British Academy* 36 (1950): 51–68.

Scientific Explanation: A Study of the Function of Theory, Probability, and Law in Science. Cambridge, U.K.: Cambridge University Press, 1953.

An Empiricist's View of the Nature of Religious Belief. Cambridge, U.K.: Cambridge University Press, 1955a.

Theory of Games as a Tool for the Moral Philosopher. Cambridge, U.K.: Cambridge University Press, 1955b.

"Probability and Induction." In *British Philosophy in the Mid-Century: A Cambridge Symposium*, edited by Cecil Alec Mace. London: Allen and Unwin, 1957.

"Axiomatizing a Scientific System by Axioms in the Form of Identifications." In *The Axiomatic Method with Special*

Reference to Geometry and Physics, edited by Leon Henkin, Patrick Suppes, and Alfred Tarski, 429–442. Amsterdam: North-Holland, 1959.

"Models in Empirical Science." In *Logic, Methodology, and Philosophy of Science*, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski, 224–231. Stanford, CA: Stanford University Press, 1962.

WORKS ABOUT BRAITHWAITE

Coburn, Robert C. "Braithwaite's Inductive Justification of Induction." *Philosophy of Science* 28 (1961): 65–71.

Nagel, Ernest. "A Budget of Problems in the Philosophy of Science." *Philosophical Review* 66 (1957): 205–225.

Peter Achinstein (1967, 2005)

BRANDT, R. B.

(1910–1997)

Richard Booker Brandt was born on October 17, 1910, in Wilmington, Ohio. He graduated from Denison University in 1930 and received a second BA from Trinity College in Cambridge, U.K. After receiving a PhD in philosophy at Yale University in 1936, Brandt taught at Swarthmore College and then at the University of Michigan, where he was named Roy Wood Sellars Distinguished College Professor of Philosophy. Brandt was a fellow of the Guggenheim Foundation and of the Center for Advanced Studies in the Behavioral Sciences at Stanford University; he was also a senior fellow of the National Endowment for the Humanities and a member of the American Academy of Arts and Sciences.

For more than fifty years, Brandt addressed a broad spectrum of theoretical and applied issues in ethics, drawing on the natural and social sciences to enrich our understanding of morality. His empiricist orientation displayed itself early on in his *Hopi Ethics* (1954), which recorded his own anthropological studies undertaken during three summers spent on a Hopi reservation in the 1940s. It found full expression in *A Theory of the Good and the Right* (1979), which presented his mature metaethical ideas and reflected his close study of work in behavioral psychology.

Brandt was a prominent exponent of utilitarianism, the view that morally correct action is action that maximizes utility. His ideas about what utility is changed over the years. In *Ethical Theory* (1959), he adopted a pluralistic view that included pleasure, knowledge, virtue, and equality of welfare as intrinsic values. Soon, however, he came to see happiness and desire-satisfaction theories as the real contenders. He briefly defended a desire theory, but by the time he wrote *A Theory of the Good and the Right*, he had evidently come to favor a happiness theory.

Brandt's most important contribution to normative ethics was his formulation and defense of an ideal rule utilitarianism, or "ideal moral code" theory. According to ideal rule utilitarianism, an act is right if and only if it would not be prohibited by the ideal moral code for a society. By an "ideal moral code," Brandt meant a code, the currency of which in a society would produce at least as much welfare or good per person as that of any other code. A moral code has currency in a society when a high proportion of adults in that society subscribe to its rules and recognize that those rules are accepted. The rules an ideal code comprises must be ones that people can learn and apply, so they cannot be too complex or too numerous. And because any set of rules will exact costs—in training, in guilt for noncompliance, and in restrictions on freedom—the rules should pertain only to important matters. Brandt recognized that the institutional rules of a society can give rise to obligations, even when existing institutions are less than optimal, and so institutional setting partly determines which moral code would produce the most good in the long run.

Brandt argued that ideal rule utilitarianism was distinct from act utilitarianism, because it need contain no higher-order rule prescribing that people maximize utility when lower-level rules conflict. So the two theories will differ in at least some of their implications. He also argued that whereas act utilitarianism seemingly implies that various immoral acts, like murdering one's aged parent, would be permissible if they could be kept secret, a moral code that sanctioned secret murders, say, would not maximize utility. Finally, because an ideal moral code would contain rules akin to W. D. Ross's prima facie duties, ideal rule utilitarianism can capture the personal character of morality, which Ross alleged that act utilitarianism misses.

Critics have questioned whether Brandt's ideal rule utilitarianism escapes the standard problems for rule utilitarianism, among them, that it is internally inconsistent, that it collapses into act utilitarianism, and that it cannot handle cases where others are not behaving as they ought. Critics have also questioned whether Brandt's theory can allow for moral appraisal of unique situations not covered by the rules, and whether moral rules lacking actual currency can plausibly provide the criterion of right acts. But at least one defender of an ideal-code consequentialism, Brad Hooker (2000), argues that a properly formulated theory can withstand such objections.

Later in his career, Brandt famously worked to resuscitate the metaethical theory known as ethical naturalism. He advocated a "method of reforming definitions," which involves redefining our ordinary vague ethical words in terms sufficiently clear and precise to render the traditional questions of moral philosophy empirically tractable. Following his proposed method, Brandt defined "rational" to refer to desires, actions, and moral systems that would survive maximal criticism and correction by facts and logic. He defined "good" to mean rationally desired, treating rational desires as those that would survive or be produced by "cognitive psychotherapy," a process in which persons represent to themselves repeatedly, in an ideally vivid way and at the appropriate time, all available relevant information. He defined "morally wrong" and "morally right" relative to the idea of a moral code that all fully rational persons would tend to support for a society in which they expected to spend a lifetime. Brandt argued that fully rational persons would opt for a broadly welfare-maximizing moral code, and that fully rational persons, insofar as they are benevolent, would seek to maximize not desire satisfaction but net lifetime enjoyment or happiness.

Brandt's critics have argued that his definitions foreclose important normative questions, such as whether it is rational to smoke even when the desire to smoke survives cognitive psychotherapy. They have questioned the coherence and empirical adequacy of appeals to full information, though such appeals remain common. They have also argued that his method begs the question against moral realism, and that his theory, like earlier forms of ethical naturalism, fails to capture the normativity of ethics. Ethicists continue to debate whether naturalism and moral realism are tenable. Whatever one concludes about Brandt's own views, his work played a crucial part in the late-twentieth-century revival of metaethics.

See also Consequentialism; Ethical Naturalism; Hedonism; Metaethics; Utilitarianism.

Bibliography

WORKS BY BRANDT

Hopi Ethics. Chicago: University of Chicago Press, 1954.

Ethical Theory. Englewood Cliffs, NJ: Prentice-Hall, 1959.

"Toward a Credible Form of Utilitarianism." In *Morality and the Language of Conduct*, edited by Hector-Neri Castañeda and George Nakhnikian. Detroit, MI: Wayne State University Press, 1965.

"Some Merits of One Form of Rule-Utilitarianism." In *Readings in Contemporary Ethical Theory*, edited by Kenneth Pahel and Marvin Schiller. Englewood Cliffs, NJ: Prentice-Hall, 1970.

A Theory of the Good and the Right. Oxford, U.K.: Clarendon Press, 1979.

Morality, Utilitarianism, and Rights. Cambridge, U.K.: Cambridge University Press, 1992.

Facts, Values, and Morality. New York: Cambridge, 1996.

WORKS ON BRANDT

Diggs, B. J. "A Comment on 'Some Merits of One Form of Utilitarianism.'" In *Readings in Contemporary Ethical Theory*, edited by Kenneth Pahel and Marvin Schiller. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Hooker, Brad. *Ideal Code, Real World: A Rule Consequentialist Theory of Morality*. Oxford, U.K.: Clarendon Press, 2000.

Hooker, Brad, ed. *Rationality, Rules, and Utility: New Essays on the Moral Philosophy of Richard B. Brandt*. Boulder, CO: Westview Press, 1993.

Loeb, Don. "Full-Information Theories of Individual Good." *Social Theory and Practice* 21 (1995): 1–30.

Rosati, Connie S. "Brandt's Notion of Therapeutic Agency." *Ethics* 110 (2000): 780–811.

Rosati, Connie S. "Persons, Perspectives, and Full-Information Accounts of the Good." *Ethics* 105 (1995): 296–325.

Sobel, David. "Full Information Accounts of Well-Being." *Ethics* 104 (1994): 784–810.

Sturgeon, Nicholas. "Brandt's Moral Empiricism." *Philosophical Review* 91 (1982): 389–422.

Velleman, J. David. "Brandt's Definition of *Good*." *Philosophical Review* 97 (1988): 353–371.

Connie S. Rosati (2005)

BRENTANO, FRANZ

(1838–1917)

Franz Brentano, a German philosopher and psychologist, was the nephew of the poet Clemens Brentano and of the author Bettina von Arnim. He taught at Würzburg and at the University of Vienna. As a teacher he exerted extraordinary influence upon his students, among whom were Alexius Meinong, Edmund Husserl, Kasimierz Twardowski, Carl Stumpf, Tomas Masaryk, Anton Marty, Christian Ehrenfels, and Franz Hillebrand. Brentano became a Roman Catholic priest in 1864, was involved in the controversy over the doctrine of papal infallibility, and left the church in 1873. At his death he left behind voluminous writings and dictation (he was blind during the last years of his life) on almost every philosophical subject. Some of this material has since been published.

The most important of Brentano's works published during his lifetime is *Psychologie vom empirischen Standpunkt* (Leipzig, 1874). The two-volume second edition (Leipzig, 1911) includes revisions and supplementary material; the third edition, edited by Oskar Kraus, was

published in Leipzig in 1925. The second edition includes *Von der Klassifikation der psychischen Phänomene*, which had also been published separately (Leipzig, 1911). The posthumously published *Vom sinnlichen und noetischen Bewusstsein*, also edited by Kraus (Leipzig, 1928), is referred to as Volume III of the *Psychologie*.

OBJECTS OF MENTAL PHENOMENA

Brentano took the mental to comprise such phenomena as hearing, seeing, sensing, thinking, judging, inferring, loving, and hating. He held that what is common to mental phenomena and what distinguishes them from the physical is "intentional inexistence," which he also described as "reference to a content" and "direction upon an object." Mental phenomena, he said, may be defined as phenomena that "include an object intentionally within themselves." He did not mean to imply, however, that when, for example, a person thinks of a horse, there is a duplicate of the horse, a mental simulacrum, existing within the mind. The essential point, as he later emphasized, is that a person could think of a horse even if there were no horse. In the second edition of the *Psychologie*, he contrasted strict relations with mental relations. *A* and *B* cannot be related in the strict sense of the term *relation* unless *A* and *B* exist; if one tree is to the left of another, then both trees exist. "But in the case of psychical relations the situation is entirely different. If someone thinks of something, then, although there must be a thinker, the thing that he thinks about need not exist."

Reference or "direction upon something" (*Gerichtetsein*) thus is common and peculiar to what is mental, and Brentano classified mental phenomena in terms of the different ways in which they may refer to, or be directed upon, their objects. There are three ways in which one may be "intentionally" related to any object *A*. (1) One may think of *A*, or, as we sometimes say, have it "before the mind" or "present to consciousness." (2) One may take an intellectual stand with respect to *A*; this stand will consist either of accepting *A* or of rejecting *A*. (3) One may take an emotional stand with respect to *A*: This is a matter of loving or hating *A*, in a very broad sense of these terms. It is a matter of pursuit or avoidance, or, as one might now say, a matter of having a "pro-emotion" or an "anti-emotion" with respect to *A*. Brentano identified these three types of phenomena with (1) *Vorstellungen* (ideas, thoughts, or presentations); (2) judgments; (3) "emotive phenomena," or "phenomena of love and hate," a category including both emotions and volitions.

Ideas, or thoughts, are basic in that the other two types of mental phenomena presuppose them. In judging that there is food, or in wanting it, one has ipso facto the thought of food. Nevertheless, judging is not simply a matter of "combining ideas"; if we combine the idea of gold and the idea of a mountain, we obtain not a judgment but another idea—that of a golden mountain. The members of the third class of mental phenomena, the "phenomena of love and hate," are like judging—and unlike the mere having of an idea—in that they involve an "opposition of intentional relation." We adopt toward the object of our idea an attitude of liking or disliking, love or hate.

There is still another respect in which the third class of phenomena is like the second and unlike the first. This is stated in Brentano's *Ursprung sittlicher Erkenntnis* (The origin of our knowledge of right and wrong; 1889).

> Concerning acts of the first class, none can be called either correct [*richtig*] or incorrect. In the case of the second class, on the other hand, one of the two opposed modes of relation, affirmation and rejection, is correct and the other incorrect. The same naturally holds good of the third class. Of the two opposed modes of relation, love and hate, being pleased and being displeased, one of them in every case is correct and the other incorrect.

This significant thesis is basic to Brentano's theory of knowledge and to his moral philosophy.

To judge, then, is to take an intellectual stand with respect to an object, and the object of the judgment is the same as the object of the idea that the judgment presupposes. If one judges that there are horses, the object of one's judgment is simply the object *horse*, which one thereby accepts, affirms, or acknowledges (*erkennt*); if one denies that there are horses, the object of one's judgment is again the object *horse*, which this time one denies or rejects (*leugnet*). In neither case does the judgment take as its object either a proposition or state of affairs or the type of entity that other philosophers have attempted to designate by such phrases as "the being of horses," "the nonbeing of horses," and "that there are horses."

This nonpropositional theory of judgment, which is fundamental to Brentano's theory of truth and his theory of categories, may be put schematically, in slightly oversimplified form, as follows. To judge that there are *A*'s is to accept (or affirm) *A*'s. To judge that there are no *A*'s is to reject (or deny) *A*'s. To judge that some *A*'s are *B*'s is to accept *AB*'s (*A*'s that are *B*'s), and to judge that no *A*'s are *B*'s is to reject *AB*'s. To judge that some *A*'s are not *B*'s, therefore, is to accept *A*'s that are non-*B*'s, and to judge

that all *A*'s are *B*'s is to reject them. (Brentano noted, however, that the sentence "All *A*'s are *B*'s" is normally used to express a twofold judgment: the acceptance of *A*'s that are *B*'s and the rejection of *A*'s that are non-*B*'s.)

Brentano attempted to extend his theory to apply to so-called compound judgments. "He judges that there are *A*'s and *B*'s" presents no difficulty, since, according to Brentano's theory of categories, if *A* is a concrete object and *B* is a concrete object, then the collective consisting of just *A* and *B* is also a concrete object. The object of this conjunctive judgment is simply *A*-and-*B*, which the judger is said to accept. Brentano suggests two interpretations of "He judges that if there are *A*'s, then there are *B*'s." According to the first interpretation, the judger is said simply to reject *A*'s-without-*B*'s. The second interpretation is more complex, making use of the terms *true* and *apodictic*. (The latter term designates a mode of judgment. To reject *A* "apodictically" is, in effect, to reject the possibility of *A*; but Brentano explicated "possibility" in terms of "apodictic rejection," and not conversely.) If by "a correct *A*-acceptor" we mean a man who accepts *A* truly, or correctly, then the hypothetical judgment becomes: "He apodictically rejects judgers who are both correct *A*-acceptors and correct *B*-rejectors." The disjunctive judgment "He judges that either there are *A*'s or there are *B*'s" could then become "He apodictically rejects judgers who are both correct *A*-rejectors and correct *B*-rejectors."

The philosophical consequences of this nonpropositional theory of judgment are far-reaching. One consequence is an interpretation of Immanuel Kant's dictum that "existence" is not a predicate. According to Brentano, when we say that *A* exists, "it is not the conjunction of an attribute of 'existence' with '*A*,' but '*A*' itself which we affirm." The word *exists* is a synsemantic term that is used to express the act of judgment.

All of the doctrines set forth above fall within the province of what Brentano called descriptive psychology. Unlike experimental psychology—including genetic and physiological psychology—descriptive psychology, according to Brentano, is an exact science, capable of arriving at laws that hold true universally and not merely "for the most part." It is the basis for all philosophy and is even capable of providing a *characteristica universalis* of the sort that Gottfried Wilhelm Leibniz had conceived. Descriptive psychology is closely related to what Husserl was to call phenomenology. Husserl had studied with Brentano in Vienna from 1884 to 1886, when Brentano used the expression *beschreibende Phänomenologie* (descriptive phenomenology) as an alternative name for descriptive psychology. (Husserl later wrote that without Brentano's doctrine of intentionality, "phenomenology could not have come into being at all.") Brentano's conception of psychology has led some of his critics to accuse him of what Gottlob Frege and Husserl called psychologism. This accusation, however, does not take into account Brentano's theory of evidence and his moral philosophy, both of which he took to be branches of descriptive psychology.

MORAL PHILOSOPHY

Brentano's ethical views are set forth in *Ursprung sittlicher Erkenntnis* (Leipzig, 1889; 3rd ed., edited by Oskar Kraus, 1934), translated by Cecil Hague as *The Origin of Our Knowledge of Right and Wrong* (London, 1902), and in *Grundlegung und Aufbau der Ethik* (The basis and structure of ethics; edited by F. Mayer-Hillebrand, Bern, 1952). Brentano based his ethics upon the assumption that the members of the third class of mental phenomena, loving and hating, may be said to be correct or incorrect, just as judgments may be said to be correct or incorrect. To say that something, *A*, is good is to say that it is impossible to love *A* incorrectly; that is, it is apodictically to reject incorrect lovers of *A*. Analogously, to say that *A* is bad is apodictically to reject incorrect haters of *A*.

The only way to grasp the concept of correct emotion, according to Brentano, is to contrast actual cases of emotions that are "qualified as correct" with cases of emotions that are not. This is analogous to the way in which we understand, for example, what it is to be red and what it is to be colored. Thus we learn that knowledge is good, joy is good (unless it is joy in what is bad), every enrichment within the realm of ideas is good, love of the good is good, love of the bad is bad, and the right end in life is to choose the best among all attainable ends.

The correctness of loving and hating, like that of judging, is objective in that it is impossible for anyone to love correctly what anyone else hates correctly or to love incorrectly what anyone else hates incorrectly. Ethics must make use of the comparative concept *better than*, for which there is no analogue in the theory of knowledge. "*A* is better than *B*," according to Brentano, means that it is correct to prefer *A*, as an end, to *B*.

EVIDENCE AND TRUTH

Brentano's views on evidence and truth may be found in the posthumously published *Wahrheit und Evidenz* (edited by Oskar Kraus, Leipzig, 1930). The distinction between judging on the basis of evidence and judging "blindly" is not to be described in terms of instinct, feel-

ings, degree of conviction, or impulse to believe. We arrive at the general concept of being evident, according to Brentano, in the same way we arrive at the concept of a correct emotion: by contemplating actual instances of the concept, in this case actual instances of evident judgments and of blind judgments.

Every evident judgment is either directly or indirectly evident; if a judgment is indirectly evident, its evidence is conferred, ultimately, by judgments that are directly evident. Directly evident judgments are of two kinds. First, there are the judgments of "inner perception," such as the judgments that I am now judging in a certain way, that I seem to see such-and-such, that I think I remember so-and-so. Second, there are judgments of reason or insights (*Einsichten*), such as the judgments that two things are more than one thing; that that which is red is, as such, other than that which is green; that there cannot be a triangle with four sides; or that a whole cannot exist if its parts do not exist.

Every judgment that is evident is true, but not every judgment that is true is evident. Most judgments of "outer perception" (of the external world), Brentano believed, are true, but all of them are "blind"; they are not evident. He argued, however, that the hypothesis of a three-dimensional external world, with its familiar details concerning physical bodies, has an "infinitely greater probability" than any of its alternatives. Judgments based on memory, too, are "blind"; but many of them confirm each other, and they are worthy of our confidence.

In *Wahrheit und Evidenz* Brentano characterized truth by reference to evidence: "Truth pertains to the judgment of the person who judges correctly … hence it pertains to the judgment of one who asserts what the person who judges with evidence would assert" (p. 139). In addition, to say that *A* exists is to say that anyone who judged about *A* with evidence would accept *A*, and to say that *A* does not exist is to say that anyone who judged about *A* with evidence would reject *A*. The "measure of all things," then, is the man who judges with evidence.

These statements, however, relating truth to evidence, do not give us the whole of Brentano's theory of truth. "Evident" is said to be predicate in the strict sense of the term, but "true" and "exists" are not, being only synsemantic. This brings us to Brentano's theory of categories.

THEORY OF CATEGORIES

The basic theses of Brentano's theory of categories may be stated as (1) there is nothing other than concrete particular things, and (2) every judgment is either the acceptance or the rejection of some concrete particular thing. "Concrete" must be taken as the opposite of "abstract" and not as a synonym for "physical." Human souls and God, according to Brentano, are concrete things but not physical things.

Our language seems to make reference to a great variety of *irrealia*—entities that are not concrete things. In fact, however, "the objects of our thought are never anything other than concrete things," and therefore for every sentence that is true and that seems to mention some nonconcrete thing, "one can form an equivalent in which the subject and predicate are replaced by something referring to a real thing" (*Psychologie*, Vol. II, p. 163). For example, "There is a lack of gold" becomes "There is no gold" (a rejection of gold); "He believes that there are horses" becomes "He accepts (affirms) horses"; and "Red is a color" becomes "A red thing is, as such, a colored thing." This latter translation is more effective in German—*Das Rotes ist als solches ein Farbiges*—where adjectives are readily transformed into nouns.

Many philosophically troublesome words, such as "exists," "good," "impossible," and "true," are synsemantic; their linguistic function is not that of referring to concrete things. "Exists" in "God exists," as we have noted, is used to express acceptance of God; "does not exist," analogously, is used to express rejection. "*A* is good" expresses an apodictic rejection of incorrect lovers of *A*. "*A* is impossible" expresses an apodictic rejection of evident acceptors of *A*—of judgers who accept *A* with evidence.

A true judgment, according to Brentano, is a judgment that cannot contradict an evident judgment. Thus "true," in "It is true that God exists," may be used to express apodictic rejection of evident rejectors of God. "It is not both true and false that God exists" may express apodictic rejection of collectives consisting of evident acceptors and evident rejectors of God. (He also noted that "true" may be used to express agreement and that, at times, it is simply redundant.) Brentano could thus be said to have an expressive theory of truth, but one that involves an objective—and not merely expressive—theory of evidence. His theories of existence and of the nature of goodness may be similarly described. Brentano's theory of categories contains important material on substance and accident, wholes and parts, the theory of relations, causation, and time and space that cannot be summarized here.

LOGIC

Brentano proposed the following revision of the theory of the syllogism on the basis of his theory of judgment. He

wrote "All *S* are *P*" (A) as "There is no *S* which is a non-*P*"; "No *S* are *P*" (E) as "There is no *S* which is a *P*"; "Some *S* are *P*" (I) as "There is an *S* which is a *P*"; and Some *S* are not *P*" (O) as "There is an *S* which is a non-*P*." Since in this account both A and E are denials, and both I and O affirmations, Brentano was able to say that no affirmative judgment is universal and no negative judgment is particular. Barbara is written as "There is no *M* which is a non-*P*; there is no *S* which is a non-*M*; hence there is no *S* which is a non-*P*." And Ferio is written as "There is no *M* which is a *P*; there is an *S* which is an *M*; hence there is an *S* which is a non-*P*." Brentano was then able to formulate the doctrine of the syllogism in three rules, which may be confirmed by the two examples just cited.

> (1) Every categorical syllogism contains *four* terms, two of which are opposed to each other and the other two of which occur twice. (2) If the conclusion is negative, then each premise is negative and has a term in common with the conclusion. (3) If the conclusion is affirmative, then one premise will share its quality and contain one of its terms, and the other premise will have the opposite quality and contain the opposite of one of its terms. (*Psychologie*, Vol. II, p. 78)

The so-called weakened and strengthened moods, according to this account, are invalid. The subaltern inferences from A to I and from E to O fail, but all four propositions, if written in Brentano's notation, may be simply converted.

OTHER WRITINGS

Vom Dasein Gottes (On the existence of God; edited by Alfred Kastil, Leipzig, 1929), is a systematic theodicy in which Brentano appealed to the fact of contingency and the principle of sufficient reason, a principle that he believed to be logically necessary, in order to prove that there is a Necessary Being. He appealed to the evidence of design in order to prove that this Being is intelligent and good. Here, and in *Religion und Philosophie* (edited by F. Mayer-Hillebrand, Bern, 1954), he attempted to show that the soul is both spiritual and immortal. The subject of consciousness is said to be a nonspatial substance, forming no part of the physical body but capable of acting upon and being affected by the brain; it is created ex nihilo at the time of the conception of the body. Brentano defended the concept of creation ex nihilo by noting that whenever one calls an image to mind, one creates ex nihilo.

In *Versuch über die Erkenntnis* (Inquiry into the nature of knowledge; edited by Alfred Kastil, Leipzig, 1925) and *Grundlegung und Aufbau der Ethik*, Brentano argued that the assumption that there can be absolute chance is self-contradictory and that the thesis of indeterminism is incompatible with the existence of human responsibility. But we have "freedom of the will" in that we are able to bring about some of the things we desire to bring about and are able to deliberate and then to decide accordingly. Moreover, we can "will to will" in that, at any given time, there are things we can do that will affect our volitions at some later time.

According to *Die vier Phasen der Philosophie* (edited by Oskar Kraus, Leipzig, 1926), those periods in which philosophy flourishes tend to be followed by three phases of decline: the first phase is characterized by a transfer of interest from the theoretical to the practical, the second by a tendency toward skepticism, and the third by a relapse into mysticism. This was the pattern of Greek philosophy; in modern philosophy the period of John Locke, René Descartes, and Leibniz was followed by the Enlightenment, then by the skepticism of David Hume, and finally, according to Brentano, by the obscurities of Kant and the idealists who followed him.

See also Descartes, René; Ehrenfels, Christian Freiherr von; Enlightenment; Ethics, History of; Existence; Frege, Gottlob; Hume, David; Husserl, Edmund; Idealism; Intentionality; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Logical Terms, Glossary of; Marty, Anton; Masaryk, Tomáš Garrigue; Meinong, Alexius; Propositions; Psychology; Stumpf, Karl; Twardowski, Kazimierz.

Bibliography

Brentano's historical writings include the following works on Aristotle: *Von der mannigfachen Bedeutung des Seienden nach Aristoteles* (Freiburg, 1862; republished Darmstadt, 1960), an important work that is the source of much of Brentano's later thought; *Die Psychologie des Aristoteles* (Mainz: Kirchheim, 1867); *Aristoteles Lehre vom Ursprung des menschlichen Geistes* (Leipzig: Veit, 1911); and *Aristoteles und seine Weltanschauung* (Leipzig: Quelle and Meyer, 1911). His *Geschichte der griechischen Philosophie*, edited by F. Mayer-Hillebrand (Bern: Francke, 1963), is compiled from the notes for his university lectures.

Brentano's other writings include *Untersuchungen zur Sinnespsychologie* (Leipzig: Dunker and Humblot, 1907); *Die Lehre Jesu und ihre bleibende Bedeutung*, edited by Alfred Kastil (Leipzig, 1922); *Grundzüge der Ästhetik*, edited by F. Mayer-Hillebrand (Bern: Franck, 1959); and *Aenigmatias*, 5th ed. (Bern, 1962).

Certain portions of the *Psychologie* are translated in *Realism and the Background of Phenomenology*, edited by R. M.

Chisholm (Glencoe, IL: Free Press, 1960); other translations are being prepared.

The most informative works on Brentano are Alfred Kastil, *Die Philosophie Franz Brentanos: Eine Einführung in seine Lehre* (Bern: Francke, 1951) and Oskar Kraus, *Franz Brentano: Zur Kenntnis seines Lebens und seiner Lehre* (Munich: Beck, 1919). The latter contains "Erinnerungen an Franz Brentano," by Carl Stumpf and Edmund Husserl. See also G. E. Moore, "Review of Franz Brentano, *The Origin of the Knowledge of Right and Wrong*," in *International Journal of Ethics* 14 (1903): 115–123.

Works published since this original entry was written in 1967 include the following:

The True and the Evident. Edited by Oskar Kraus. English ed. edited by Roderick M. Chrisholm. Translated by Roderick M. Chrisholm, Ilse Politzer, and Kurt R. Fischer. London: Routledge & K. Paul; New York: Humanities Press, 1966.

Die vier Phasen der Philosophie und ihr augenblichlicher Stand, nebst Abhandlungen über Plotinus, Thomas con Aquin, Kant, Schopenhauer und Auguste Comte, mit Anmerkungen, edited by Oskar Kraus. Hamburg: Meiner, 1968.

Uber die Zukunft Philosophie; nebst den Vorträgen: Uber die Gründe der Entmutigung auf philosophischem Gebiet, Uber Schellings System, sowie den 25 Habilitationsthesen, edited by Oskar Kraus and Paul Weingartner. Hamburg: F. Meiner, 1968.

Vom sinnlichen und noetischen Bewusstsein. Aussere und innere Wahrnehmung, Begriffe. 2nd ed., edited by Oskar Kraus. Hamburg: Felix Meiner, 1968.

Kategoriënlehre. Mit Einleitung und Anmerkungen hrsg. von Alfred Kastel, edited by Alfred Kastel. Hamburg: Meiner, 1968.

The Origin of our Knowledge of Right and Wrong, edited by Oskar Kraus. English ed. edited by Roderick M. Chisholm. Translated by Roderick M. Chisholm and Elizabeth H. Schneewind. New York: Humanities Press, 1969.

The Foundation and Construction of Ethics. Compiled from His lectures on Practical Philosophy, edited by Franziska Mayer-Hillebrand. English ed. edited and translated by Elizabeth Hughes Schneewind. London: Routledge and Kegan Paul, 1973.

Psychology from an Empirical Standpoint, edited by Oskar Kraus. Translated by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister. English ed. edited by Linda L. McAlister. London: Routledge and Kegan Paul; New York: Humanities Press, 1973.

The Philosophy of Brentano, edited by Linda L. McAlister. London: Duckworth, 1976.

Philosophische Untersuchungen zu Raum, Zeit und Kontinuum. Hamburg: Meiner, 1976.

The Psychology of Aristotle: In Particular His Doctrine of the Active Intellect: With an Appendix concerning the Activity of Aristotle's God. Berkeley: University of California Press, 1977.

Aristotle and His World View, edited and translated by Rolf George and Roderick M. Chisholm. Berkeley: University of California Press, 1978.

Die Philosophie Franz Brentanos: Beiträge zur Brentano-Konferenz Graz, 4-8. September 1977, edited by Rocerick M. Chisholm and Rudolf Haller. Amsterdam: Rodopi, 1978.

Aristoteles Lehre vom Ursprung des menschlichen Geistes, edited by Rolf George. Hamburg: Meiner, 1980.

Geschichte der mittelalterlichen Philosophie im christlichen Abendland, edited by Klaus Hedwig. Hamburg: F. Meiner, 1980.

Sensory and Noetic Consciousness: Psychology from an Empirical Standpoint III, edited by Oskar Kraus. English ed. edited by Linda L. McAlister. Translated by Margarete Schättle and Linda L. McAlister. London: Routledge & Kegan Paul; New York: Humanities Press, 1981.

The Theory of Categories. The Hague; Boston: Martinus Nijhoff, 1981.

Deskriptive Psychologie, edited by Roderick M. Chisholm and Wilhelm Baumgartner. Hamburg: Meiner, 1982.

Brentano and Meinong Studies. Atlantic Highlands, NJ: Humanities Press, 1982.

Brentano and Intrinsic Value. Cambridge, U.K; New York: Cambridge University Press, 1986.

On the Existence of God: Lectures Given at the Universities of Würzburg and Vienna, 1868–1891, edited by Susan F. Krantz. Dordrecht; Boston: M. Nijhoff, 1987.

Brentano Studien: Internationales Jahrbuch der Franz Brentano Forschung / Franz Brentano Forschung; Franz Brentano Foundation. Würzburg: Röll, 1988.

Grundzüge der Ästhetik. 2nd ed., edited by Franziska Mayer-Hillebrand. Hamburg: F. Meiner, 1988.

Uber Ernst Machs "Erkenntnis und Irrtum": mit zwei Anhängen, Kleine Schriften über Enrst Mach, Der Brentano-Mach-Briefwechsel. Amsterdam: Rodopi, 1988.

Philosophical Investigations on Space, Time, and the Continuum. London: New York: Croom Helm, 1988.

Clemens Brentano: Briefe 1803–1807: Textedition und Kommentierung. München: s.n., 1989.

Descriptive Psychology, edited and translated by Benito Müller. London; New York: Routledge, 1995.

The Four Phases of Philosophy. Amsterdam; Atlanta, GA: Rodopi, 1998.

Roderick M. Chisholm (1967)
Bibliography updated by Michael J. Farmer (2005)

BRIDGMAN, PERCY WILLIAM
(1882–1962)

An American physicist and professor of mathematics and natural philosophy at Harvard, Percy William Bridgman was awarded the Nobel Prize for physics in 1946 for his work on the properties of matter under extremely high pressures. He wrote at length on the philosophical implications of the discoveries of modern physics, particularly Albert Einstein's revolutionary special theory of relativity, and on the analysis of scientific concepts. To Bridgman it seemed that Einstein's theory arose chiefly from the application of sound conceptual analysis based on what Bridgman called the "operational point of view." In his

opinion, Einstein had not shown "something new about nature"—he was "merely bringing to light implications already contained in the physical operations used in measuring time." Bridgman held that analysis shows that there exists no answer to the question of what we should do, what operations we could perform, in order to determine whether or not two distant events occurred simultaneously. Therefore, it is meaningless to speak of the two events as having or not having occurred simultaneously.

According to Bridgman, then, Einstein's work dramatically highlighted an important feature of scientific methodology, the determination to link all scientific concepts to experimental procedures. From the operationalist views implicit in the practices of working scientists, we should learn to undertake a rigorous analysis of all scientific concepts, cleansing science of operationally undefinable elements.

Bridgman disclaimed all intention of founding a new philosophical school, yet his name has become linked inseparably with operationalism. Many scientists have hailed Bridgman's ideas as indispensable to the correct understanding of modern science, and some, particularly psychologists, have urged the inauguration of an extensive program of analysis of scientific concepts along the lines laid down by Bridgman. Others have regarded Bridgman's philosophy as not only wrong, but also harmful—if it were imposed on science, it could stifle creative inquiry. Bridgman later claimed that each concept need not be completely definable in terms of performable instrumental operations, but that it is sufficient that a concept should be one "indirectly making connection with instrumental operations."

The controversy over operationalism diverted attention from Bridgman's numerous other ideas, many of which are original and provocative. Perhaps the most interesting is his view that discoveries in physics may help us to deal with problems in quite different domains. In his opinion, the great achievements in physics are discoveries of new ways in which our minds can master problems, discoveries about our conceptual makeup.

Through relativity physics, we have learned how apparent contradictions may arise through inadvertently admitting into science meaningless propositions that cannot stand up to operational analysis. Similarly, in human affairs seemingly irreconcilable demands of different groups may be eliminated by showing that some of the basic tenets on which the demands rest are meaningless. The methodology of the social sciences no doubt can learn much from the methodology of physics, but Bridgman's suggestion as to how human conflicts may be resolved will strike many as overly optimistic and somewhat naive.

See also Einstein, Albert; Operationalism; Philosophy of Physics; Relativity Theory.

Bibliography

WORKS BY BRIDGMAN

The Logic of Modern Physics. New York: Macmillan, 1927.

The Nature of Physical Theory. Princeton, NJ: Princeton University Press, 1936.

Reflections of a Physicist. New York: Philosophical Library, 1950.

The Nature of Some of Our Physical Concepts. New York: Philosophical Library, 1952.

SECONDARY SOURCES

Cornelius, B. A. *Operationalism.* Springfield, IL, 1955.

Frank, Philipp. *The Validation of Scientific Theories.* Boston: Beacon, 1957.

G. Schlesinger (1967)

BRIGHTMAN, EDGAR SHEFFIELD

(1884–1953)

Edgar Sheffield Brightman was the leading American advocate of personalism. At Boston University he studied under Borden Parker Bowne, the first philosopher in America to develop the personalistic position. Brightman taught at Nebraska Wesleyan University (1912–1915), Wesleyan University (1915–1919), and from 1919 at Boston University, occupying the chair of Borden Parker Bowne professor of philosophy from 1925 until his death. He was president of the Eastern Division of the American Philosophical Association in 1936.

Brightman conceived of personalism as a mediating position in philosophy. As such, it for him superseded William James's pragmatism and Josiah Royce's absolute idealism, to each of which, in turn, he had been attracted early in his career. Brightman also held that personalism could resolve the impasse between supernaturalism and naturalism. Furthermore, although he criticized positivism for being too restricted an empiricism and although he eschewed much in existentialism, Brightman's personalism can be understood as an attempt to combine the surface experience (sense) of positivism and the depth dimension (value) of existentialism in a concept of the total person.

EPISTEMOLOGY

Brightman held firmly to an epistemic dualism of the "shining present" (immediate experience) and "the illuminating absent" (the referent). Constantly emphasizing that all primary data were present experiences, he advocated a radically empirical method; that is, a method that considers whatever is, at any time, present in consciousness. Since knowledge involves reference, it is always hypothetical and tentative. Brightman accepted this as a healthy probabilism (and not a destructive skepticism), because he found in the principle of coherence an adequate test of reference (or criterion of truth). Deeply influenced by Hegelian dialectic, he viewed coherence not as formal consistency but as a principle for interpreting experience: a statement or a set of statements is true to the extent that it organizes and orders experience.

ONTOLOGY

The metaphysical perspective that emerges is a pluralistic idealism. Reality is a society of persons: the ultimate (uncreated) Person and finite (created) persons. Reality is thus not nature but history. The natural order does not have ontological identity "outside" the ultimate Person; rather, this order is his "behavior." The laws of logic do not have privileged priority; they are constitutive of the supreme mind. In philosophy of religion, this position is idealistic theism (not theological dualism). God is a conscious Person who creates finite persons and cooperates with them in the cosmic endeavor. A human person is a context of experience capable of self-consciousness, reason, and ideal values.

EVIL

Brightman is probably most widely known for his controversial treatment of the problem of evil. He argued that the power of God is limited by nonrational conditions (the Given) within the divine nature that God's will neither created nor approves. God maintains constant and growing, though never complete, control of the Given.

See also Bowne, Borden Parker; Evil, The Problem of; James, William; Personalism; Royce, Josiah.

Bibliography

Brightman's chief works are *Introduction to Philosophy* (New York: Henry Holt, 1925; revised editions, 1951 and 1963, the latter edited by Robert N. Beck); *The Problem of God* (New York: Abingdon Press, 1930); *The Finding of God* (New York: Abingdon Press, 1931); *Moral Laws* (New York: Abingdon Press, 1933); *A Philosophy of Religion* (New York: Prentice-Hall, 1940); *Nature and Values* (New York: Abingdon-Cokesbury Press, 1945); *Person and Reality* (edited by Peter A. Bertocci in collaboration with Jannette E. Newhall and Robert S. Brightman [New York: Ronald Press, 1958]). A selected bibliography of his philosophical writings, including some 200 monographs and articles in addition to books, may be found in *Person and Reality,* pp. 367–370, or in the Brightman Memorial issue of *Philosophical Forum* 12 (1954): 22–28.

For references to discussions of Brightman's influence, see Peter A. Bertocci, "Edgar S. Brightman—Ten Years Later," in *Philosophical Forum* 20 (1962/1963): 3–10.

John H. Lavely (1967)

BROAD, CHARLIE DUNBAR

(1887–1971)

Charlie Dunbar Broad, the English epistemologist, historian of philosophy, moral philosopher, philosopher of science, and writer on the philosophical aspects of psychical research, was born at Harlesden, now a suburb of London. The only child of middle-class parents in comfortable circumstances, he received a good education at Dulwich College. With his special interest and ability in science and mathematics he won, in 1905, a science scholarship to Trinity College, Cambridge, with which Broad's philosophical career was to be chiefly associated. Despite success in his work at Cambridge, he became convinced that he would never be outstanding as a scientist and turned to philosophy, in which he took first-class honors with special distinction in 1910. A year later he was elected to a fellowship at Trinity because of a dissertation that became his first book, *Perception, Physics, and Reality* (Cambridge, U.K., 1914).

From 1911 to 1920 Broad was at the University of St. Andrews, first as assistant to G. F. Stout, the professor of logic and metaphysics, then as a lecturer at Dundee. During World War I, he combined his lecturing duties with work for the Ministry of Munitions in a chemical laboratory. He followed C. Lloyd Morgan in the chair of philosophy at the University of Bristol in 1920, but after a few years he returned to Trinity College to succeed J. M. E. McTaggart as lecturer in moral science. In 1933 Broad somewhat reluctantly became Knightbridge professor of moral philosophy. Until his retirement in 1953, Broad had not traveled outside Great Britain except for periodic visits to Scandinavia, in particular to Sweden, a country whose people, life, and language had long attracted him. Broad's encouragement of Swedish philosophers and phi-

losophy led to his being generously honored by the academicians of that country. In Britain his services to philosophy were recognized by bestowal of most of the honors available to a don so secluded from public activity.

At Cambridge, Broad was most influenced by his teachers, McTaggart and W. E. Johnson, and by Bertrand Russell and G. E. Moore. These four men, with the important additions of Stout and A. E. Taylor at St. Andrews, represent in the diversity of their thought something of the extraordinary range of Broad's own interests. Among British philosophers of this century, no one, including Russell, published so much on so many different philosophical topics. The largest part of Broad's writing falls within the theory of knowledge and the philosophy of science—provided that we assign some of the problems of traditional metaphysics to these two fields—although he also wrote extensively, if less systematically, on ethics and on the life and thought of such scattered figures as Francis Bacon, Isaac Newton, Butler, and Immanuel Kant.

The ample scope and scale of Broad's work were displayed early in his career. Within his first three years of serious publication, he had produced almost two dozen reviews of widely different books, essays on "The Doctrine of Consequences in Ethics" (*International Journal of Ethics* 24 [April 1914]: 293–320) and "Lord Hugh Cecil's 'Conservatism'" (*International Journal of Ethics,* 23 [July 1913]: 396–418), a critical notice of Meinong's *Über Annahmen* (*Mind*, n.s., 22 [January 1913]: 90–102), and his own first volume, which discussed the relation between causation and perception. This catholicity of interest remained apparent for the next fifty years, despite Broad's confession in the autobiographical chapter of *The Philosophy of C. D. Broad* that some time after his acceptance of the Knightbridge chair he gave up philosophy in all but title and routine: "I no longer believed in the importance of philosophy, I took little interest in its later developments, and I knew very well that I at least had shot my bolt and had nothing further of value to contribute."

The most curious feature of this confession is that it makes the development of ennui coincide with a period of considerable publication by Broad. The 800 pages of the second volume of his *Examination of McTaggart's Philosophy* (Cambridge, U.K., 1933–1938) were written at this time, as were his essays on John Locke (*Hibbert Journal* 31 [January 1933]: 249–267) and Henry Sidgwick (ibid., 37 [October 1938]: 25–43), his inaugural lecture on determinism, a number of papers given to the Aristotelian Society, and a spate of notes on psychical phenomena. Broad's changed attitudes and feelings toward his chosen field had little substantial effect on the work he contributed to it.

THEORY OF KNOWLEDGE

Broad's writings on perception and knowledge, like the rest of his work, form neither a system nor a set of unequivocal answers to a group of related questions. For every philosophical position there were always reasons pro and con; and on any given issue Broad often found it difficult to decide where the weightier reasons lay.

SENSE DATA. Thus, following Stout, and ultimately Locke, in distinguishing between the odors, noises, and colored patches that we sense and the physical objects like coins and books that we perceive, Broad gave rather cautious support to a version of the causal theory of perception. There are, he thought, two kinds of particulars involved in perception—persistent substances (bodies) with properties like shape, size, inertial mass, and spatial position; and the "sense-qualified occurrents" of which we are immediately aware in sensing, as when we see the upper surface of a dinner plate. Broad argued that visual sense data, or *sensa* as he called them, at least are never, in fact, identical with, or parts of, the surface of the physical object that is seen. If we recall that the sense data obtained by a given person in looking at the same surface from different positions and angles form a continuous series, and that the velocity of light is finite, it is reasonable to believe that at least some of the properties of sense data must be different from those of their correlated bodies, that a penny, for example, retains the same size and shape while our sense data of it change in these respects as we alter position. The greater the distance between our eyes and the body seen, the more obvious it is that the properties of the body and of our sense data must differ.

It is likewise reasonable that if this difference sometimes holds, it must always hold; for there is no gap in the continuity of conditions in which we obtain sense data of a particular surface that would allow us to identify only some of the sense data with that surface. As underpinning for this sharp distinction, Broad tried to establish that a sense datum must have all the properties that it is sensed as having, although it may also have unnoticed properties; that unsensed sense data can exist; and that the word *sensation* refers both to bodily feelings and to "genuine sensations," the former of which are not, although the latter are, analyzable into an act of sensing and its object, the sense datum.

In general, Broad treated these claims about the existence and properties of sense data as being empirical

ones, and so was led to a similar treatment of such questions as: Are sensa qualitatively mind-dependent? Can two people sense the same sensum? How long can a sensum last? Do we infer from the properties of our sensa to the properties of physical objects? How much resemblance is there between the properties of sensa and the properties of physical objects? In his "Reply to Critics," written late in his career, Broad indicated that he did not feel the force of the view, made familiar by G. A. Paul and A. J. Ayer, that these questions can be answered only by decisions in particular cases or else are misconceived, since the sense-data theory is simply an elaborate terminological proposal for dealing with the argument from illusion. Nor did he recognize the radical criticism that this view offered of his own attempts to deal with sense data as private objects interposed between human observers and the unobservable physical world. The latter is the "remote causal ancestor" of our sensations, he thought, and the kind of isomorphism one must postulate between the properties of sense data and the properties of "the hypothetical system of physical things and events" he was "willing to leave to experts to decide."

THE MIND-BODY PROBLEM. In his discussion of the mind-body problem, Broad set out to produce a theory that would account for the apparent fact that brain events are a necessary condition of mental events, and also leave open the possibility that some mental events occur after the death of their associated bodies. He suggested that minds are the result of two components—a nervous system, and a "psychogenic factor," which is modified by experience and capable of persisting after bodily death. Since no other properties are assigned to the psychogenic factor, nor is its relation to the brain described, the factor remains unobservable, either directly or indirectly; and the parent theory is obviously ad hoc. Broad would have welcomed a theory that was more open to experimental testing; although he distinguished metaphysical from scientific theories by the latter's susceptibility to such testing. He was thus in the position of answering the philosophical question, How are bodies related to minds? with what was, by his own criteria, an inadequate scientific theory. Just as he took sense data to be private objects whose properties could be investigated by introspection, so he took the mind-body relation as being similar to the relation between a visible body and an invisible one—a relation open in theory, if not in practice, to empirical investigation.

GENERAL EXPLANATORY PRINCIPLES. Closely related to this treatment of philosophical problems was Broad's attempt, throughout his writings, to isolate a set of very general principles that would be both necessarily true and genuinely explanatory of the most pervasive and important features of the world. Broad was not convinced either that every necessarily true statement is analytic or that every synthetic statement is testable by means of perceptual experience. He thought that there might well be propositions, such as "The cause of any change contains a change as an essential factor," which are synthetic—informative about the world—but necessarily true. The denial of this proposition is not self-contradictory, so the proposition cannot be analytic; yet a counterexample is impossible to imagine, so the proposition, rather than being an ordinary empirical one, is self-evidently true. Propositions as general as this, Broad half suggested, are the appropriate axioms of metaphysical theories, theories whose results he compared unfavorably to the "beautiful and surprising consequences" deduced from the premises of geometry and such physical premises as the "entropy principle." Broad's pessimism about the utility of deductive metaphysics seems to have been the outcome of a desire to treat speculative philosophy as a suprascience, one that accounted for our most general concepts, such as cause, substance, potentiality, and actuality, in much the same way that physics accounted for such less general concepts as velocity, mass, simultaneity, and the atom.

A PRIORI CONCEPTS. This distinction between the concepts dealt with by the sciences and those more general ones dealt with by philosophy has its parallel, and perhaps its source, in the distinction drawn by Broad between empirical and a priori (nonempirical) concepts. He believed that the simplest empirical concepts, for example, the ideas of red or yellow, are formed by our contrasting and comparing many different red or yellow objects. Eventually, we abstract the required quality from all other qualities and from any particular substance in such a way that we are able to think of the quality in the absence of any instance or image of it. In thus accepting the traditional story of the genesis of empirical concepts, Broad hesitated between the two equally ancient views of how we form a priori concepts. The first view is that we have innate dispositions to form specific ideas like those of cause, substance, and rightness as the result of having certain kinds of experiences. The second is that we have "a general power of non-perceptual intuition," distinct from our ability to have sense perceptions and to introspect, which allows us to intuit such relations as causation and rightness whenever we have the appropriate kinds of experiences to stimulate the power.

A standard criticism of these theories of concept formation is that the story about abstraction is logically circular; and that the accounts of a priori concepts apply equally well or little to empirical ones, so that Broad's distinction between the two cannot be drawn. The abstraction story is circular because in order to compare and contrast one color with another we must already have the ability to recognize and distinguish those colors. Yellow objects that are to be compared must be seen as yellow before the suggested procedure can begin. Hence, we can rightly claim that innate ideas or nonempirical intuitions are needed for the concept of yellow as they are for concepts like that of substance.

However, thinking of an absent quality yellow is not the intellectual analogue of sensing a yellow patch, for thinking of yellow is not a matter of "contemplating the characteristic" yellow, as Broad once assumed it was. Noting the logically necessary relations between concepts, for example, that all yellow things must be colored, is not like having a sense datum and noting that in it a red patch borders on a yellow patch. Granting these two points, as Broad did in his "Reply to Critics," would make it less plausible to hold that some synthetic propositions may be necessarily true. For once we abandon the sense-datum picture of logical necessity, there is little temptation to appeal to self-evidence (the intellectual sensing of universal connections) in support of metaphysical principles.

PSYCHICAL RESEARCH

Broad often urged philosophers to take something of his own keen interest in psychical research. He claimed that no one could answer the question as to whether any person actually has the power of paranormal precognition without having made a careful study of the available evidence; but most philosophers obviously considered this to be a scientific task for psychologists. In the absence of any encouragement from scientists, few philosophers would join Broad in discussing the further question, which chiefly interested him, How does the existence of supernormal precognition affect such philosophical topics as causation, the mind-body problem, immortality, and sense perception? Suppose we took seriously the suggestion that each person has an extended but intangible and invisible body as well as his ordinary body and that the invisible body puts forth pseudopods that touch and affect external objects. The existence of such a body would certainly alter a number of our views on topics like causation and the mind-body problem. But exactly how they were altered would depend on such factors as the degree of control we could exert over our invisible bodies, whether they survived our corporeal bodies, and what sort of knowledge we could have of our intangible bodies.

Thus until there is scientific agreement on what has been established concerning paranormal cognition, it is difficult to say how its existence would affect philosophical discussion. What can undoubtedly be done is to consider whether the notion of supernormal precognition is logically coherent. Broad thought that it is and tried to rebut arguments that it is self-contradictory to speak of precognizing something that does not yet exist as well as arguments that paranormal precognition makes an effect precede its cause—correctly guessing a card symbol would be influenced by what is to be known later about the card. However, showing that paranormality is logically possible does nothing to advance its claims over alternative hypotheses in the explanation of unlikely experimental data, data that may be unlikely because of selective sampling alone.

PROBABILITY AND INDUCTION

Although Broad's two papers titled "Induction and Probability" gave what will probably be a definitive expression to their point of view, they were overshadowed by the simultaneous appearance of J. M. Keynes's *A Treatise on Probability*. In much the same way, Broad's *Scientific Thought* (London, 1923)—perhaps his best book—was neglected after the publication, a few years later, of Russell's *The Analysis of Matter*. Broad argued that the degree of belief we give to well-established inductions cannot be justified "by any known principle of probability unless some further premise about the physical world be assumed." Yet this premise is notoriously difficult to state. If induction is to be a rational procedure, nature must consist of a few kinds of substances that combine in various lawlike ways and thus produce variety in a finite world. In brief, we need Keynes's principle of limitation of independent variety. Without such a principle we cannot make use of inductive analogies, for they assume that future cases will resemble past cases, or in other words, that no one object has an infinite number of independent qualities or is producible by an infinite number of different causes. In "The Principles of Problematic Induction" (*PAS*, n.s., 28 [1927–1928]: 1–46), Broad went on to consider, and answer affirmatively, the question whether we can know that nature has this desirable structure.

Thus, Broad held that the problem of justifying inductive inferences is a genuine one. He thought that the two questions, What is meant by calling this inductive belief well-supported? and What makes induction a valid procedure? have similar answers. Each question requires

us to state the criteria by which we can distinguish sound from unsound inferences, and these criteria will enable us to provide the necessary and sufficient conditions for well-grounded inferences. Such conditions must in turn be based on fundamental principles that will serve as general premises in every sound inductive inference. This last step of Broad's claim has been much criticized as confusing two quite different issues. The first concerns the empirical statement, for which there is ample evidence, that nature is so organized that in the future at least some of our inductive beliefs will be correct. The second concerns the logically necessary truth that induction is a rational procedure; for we could not have an inductive policy that was both successful and irrational, that is, not supported by good evidence. What we mean by "rational inductive procedure" is one that is well supported by evidence. It is this support that "justifies" the policy in the only permissible sense of "justify." The structure of nature is known inductively and so cannot itself be referred to for support of the inductive procedure; nor is there any need to do so. The only justification we require is the success of the policy, and that we already have.

ETHICS

On the problems of ethics, Broad showed a cautious hesitancy to commit himself. Two of his late papers, "Some Reflections on Moral-Sense Theories in Ethics" (*PAS*, n.s., 45 [1944–1945]: 131–166) and "Some of the Main Problems of Ethics" (*Philosophy* 21 [July 1946]: 99–117), have been widely read; but they provide only hints as to Broad's own views. As in the early chapters of *Five Types of Ethical Theory* (London 1920), on such writers as Benedict de Spinoza and David Hume, Broad classified types of ethical theories, exposed their assumptions, and drew out their logical implications, without committing himself. For example, in his paper on moral-sense theories he distinguished three analyses of "That act is right": The sentence does not express the speaker's judgment, but his emotions or desires or commands; what is expressed is a judgment about "certain human experiences, certain sensations or emotions or desires," that is, a "moral feeling"; and a judgment is made that ascribes a property like "what it is fitting to approve" or "conducive to social stability," properties independent of the speaker's opinions, desires, or feelings.

In his "Reply to Critics" Broad said that theories of the second and third types must admit the existence both of nonempirical concepts of moral attributes and of synthetic a priori propositions like "any act of promise-breaking tends as such to be wrong." Since he was not convinced that there were no such concepts and propositions, he was able to sympathize with theories of these types, as well as with theories of the first type. But to the question, does "That act is right" express a judgment, a feeling, or a command? Broad could only reply, "I have no definite opinion." He was similarly undecided on the question whether ethical terms such as *wrong* and *duty* stand for properties, and if so, exactly what sort of properties these might be. His attitude here, as to many other philosophical problems, resembled that of a prudent scientist awaiting further evidence before coming to a decision.

Broad had no "philosophy" in the sense of a deeply original way of interpreting and dealing with the issues of his field. He was a scientist manqué who took up philosophical problems much as he found them, leaving them classified and more manageable but not transformed. His impressive ability to understand and recast the most difficult arguments, the elegance of his writing, his unrivaled thoroughness and lucidity, were placed at the service of other people's questions rather than his own.

See also Ayer, Alfred Jules; Bacon, Francis; Ethics, History of; Hume, David; Induction; Innate Ideas; Kant, Immanuel; Keynes, John Maynard; Locke, John; McTaggart, John McTaggart Ellis; Meinong, Alexius; Mind-Body Problem; Moore, George Edward; Newton, Isaac; Parapsychology; Precognition; Probability and Chance; Russell, Bertrand Arthur William; Sensa; Sidgwick, Henry; Spinoza, Benedict (Baruch) de; Stout, George Frederick; Taylor, Alfred Edward.

Bibliography

Broad's other books include *The Mind and Its Place in Nature* (London: Kegan Paul, 1925), his most characteristic work, and *Lectures on Psychical Research* (London: Routledge, 1963). Some of his essays have been collected in two volumes, *Ethics and the History of Philosophy* (London: Routledge, 1952) and *Religion, Philosophy, and Scientific Research* (London: Routledge, 1953). His two papers titled "Induction and Probability" appeared in *Mind* 27 (1918): 389–404 and 29 (1920): 11–45. There is a complete bibliography up to 1959 in *The Philosophy of C. D. Broad*, edited by P. A. Schilpp (New York: Tudor, 1959), which also contains 21 essays on his work by various philosophers, Broad's "Reply to Critics," and his "Autobiography." A critical examination of Broad's theory of perception is given in Martin Lean, *Sense Perception and Matter* (London: Routledge and Kegan Paul, 1953).

OTHER RECOMMENDED WORKS BY BROAD

The Nature of Existence (1921), edited by John McTaggart Ellis McTaggart. Northampton: John Dickens, 1968.

Induction, Probability, and Causation: Selected Papers. New York: Humanities, 1968.

Broad's Critical Essays in Moral Philosophy, edited by David Ross Cheney. London: Allen and Unwin; New York: Humanities Press, 1971.

Ethics. Edited by Casimir Lewy. Dordrecht; Boston: M. Nijhoff; Hingham, MA: Kluwer, 1985.

Robert Brown (1967)
Bibliography updated by Michael J. Farmer (2005)

BROUWER, LUITZEN EGBERTUS JAN
(1881–1966)

Luitzen Egbertus Jan Brouwer, the founder of mathematical intuitionism, was born in 1881 in Overschie, near Rotterdam, the Netherlands. After attending schools in Medemblik, Hoorn, and Haarlem, he studied mathematics at the Municipal University of Amsterdam. He obtained his doctorate in 1907 for his thesis, *Over de Grondslagen der Wiskunde* (Amsterdam and Leipzig, 1907). He became *privaat-docent* at Amsterdam in 1909 and served as professor there from 1912 until his retirement in 1955. In the year that he became a professor he was elected to the Royal Dutch Academy of Sciences.

Besides contributions to the foundations of mathematics, Brouwer made major contributions to other areas of mathematics, in particular to topology, in which his most important publications date from the period 1909–1913. Combinatorial or algebraic topology came into being through discoveries of Henri Poincaré in the 1890s. A fundamental technique of Poincaré was to analyze figures into combinations of simple figures and to represent the topological structure of the figures by algebraic properties of the combination. Brouwer extended and deepened this technique, particularly in relation to questions of the existence of mappings and fixed points. He proved such classic results as the topological invariance of dimension, which implies that there is no bicontinuous one-to-one mapping of Euclidean m-dimensional space onto Euclidean n-dimensional space, for $m \neq n$.

Although he was primarily a mathematician, Brouwer was always preoccupied with general philosophy and had elaborated a highly individual philosophical vision. Indeed, the most remarkable feature of Brouwer's work in the foundations of mathematics was the boldness and consistency with which, starting from his own philosophical position, he questioned the principles on which the mathematics he inherited was based, down to so elementary a principle as the law of excluded middle, and then proceeded to criticize these principles in detail and to begin to reconstruct mathematics on a basis he regarded as sound.

Although he later presented them more systematically, the essentials of Brouwer's philosophy were already present in his thesis of 1907 and, in certain respects, in *Leven, Kunst, en Mystiek* (Delft, 1905). These works antedate the decisive steps in the development of mathematical intuitionism. In effect, Brouwer argued in his thesis that logic is derivative from mathematics and dependent for its evidence on an essentially mathematical intuition that rests on a basis close to Immanuel Kant's notion of time as the "form of inner sense." Intellectual life begins with "temporal perception," in which the self separates experiences from each other and distinguishes itself from them. Brouwer describes this temporal perception as "the falling apart of a life moment into two qualitatively different things, of which the one withdraws before the other and nonetheless is held onto by memory" ("Weten, Willen, Spreken," 1933). This perception, however, belongs to an attitude (which Brouwer earlier termed "mathematical consideration") that the self adopts to preserve itself; the adoption of this attitude is an act of free will, in a broad sense that Brouwer probably derived from Arthur Schopenhauer. The fundamental intuition of mathematics is this structure of temporal perception "divested of all content"; in mathematics one sees that the process of division and synthesis can be iterated indefinitely, giving rise to the series of natural numbers. In the temporal order thus revealed, one can always imagine new elements inserted between the given ones, so that Brouwer could say that the theories of the natural numbers and of the continuum come from one intuition, an idea that, from his point of view, was made fuller and more precise by his theory of free choice sequences, although one might argue that it was made superfluous by that theory.

Brouwer's constructivism was developed in this context. His constructivism was probably motivated less by an insistence on absolute evidence and a rejection of hypotheses (which might have led to "finitism" in David Hilbert's sense of the term or even to a still narrower thesis) than by Brouwer's subjectivism and his insistence on the primacy of will over intellect. On these grounds, mathematics should consist in a constructive mental activity, and a mathematical statement should be an indication or report of such activity. Brouwer credited this

way of looking at mathematics to the inspiration of his teacher, Gerrit Mannoury.

In his thesis Brouwer limited himself to criticizing alternative theories of the foundations mathematics and to criticizing Cantorian set theory, but in "De Onbetrouwbaarheid der Logische Principes" (1908), perhaps urged on by Mannoury, Brouwer raised doubts about the validity of the law of excluded middle, although he still regarded the question as open. In *Intuitionisme en Formalisme* (1912) Brouwer did not say flatly that the law of excluded middle is false, but he gave an instance of his standard argument, an example like that presented in the section on intuitionism in the entry titled "Mathematics, Foundations of," which also gives a fuller exposition of constructivism.

In a number of publications beginning in 1918 and extending through the 1920s, Brouwer developed intuitionist mathematics and worked out in detail his critique of classical mathematics, determining for different branches of mathematics which of their theorems are intuitionistically true. In "Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten," Brouwer undertook to develop an intuitionist set theory, on which a theory of the continuum could be based. In this work he introduced his concept of set (*Menge*; later, in "Points and Spaces," 1954, he called it "spread") and therefore the idea of an arbitrary infinite sequence as generated by successive free choices. He also introduced the notion of species, which led to his own version of a predicative hierarchy of classes. The principle that the value of a function everywhere defined on a spread must, for a given sequence as argument, be determined by a sufficiently large finite number of its terms is already present in "Begründung der Mengenlehre." This "continuity axiom" is the first of the two distinctive principles of intuitionist analysis.

In "Beweis, dass jede volle Funktion gleichmässig stetig ist" (1924), Brouwer announced a proof that a function everywhere defined on the closed unit interval is uniformly continuous. In this proof Brouwer used two fundamental assertions about spreads, later called the bar and fan theorems. The bar theorem, or an equivalent assertion, constitutes the other distinctive principle of intuitionist analysis. Brouwer's proof was presented in full in "Über Definitionsbereiche von Funktionen" (1927) and reworked, in a more general setting, in "Points and Spaces."

After World War II Brouwer published a long series of short papers in which he developed a new type of counterexample to classical theorems, based on another new principle.

Brouwer's philosophy is not limited to what is relevant to the foundations of mathematics. Mathematical consideration has a second phase, which he called causal attention. In this phase "one identifies in imagination certain series of phenomena with one another," an operation by which one can pick out objects and postulate causal rules. (The relation between temporal perception and causal attention is analogous to that between Kant's mathematical and dynamical categories.) The whole point of mathematical consideration lies in the fact that it makes possible the use of means: One produces a phenomenon that will be followed in a certain repeatable series by a desired phenomenon that cannot be directly reproduced. This makes the pursuit of instinctual satisfaction more efficient.

Especially in *Leven, Kunst, en Mystiek* and in "Consciousness, Philosophy, and Mathematics" (1948), Brouwer regards this "mathematical action" as a kind of fall from grace, whose results are uncertain and ultimately disappointing. With this view he couples a pessimism about society. Society is based on communication, which is itself a form of mathematical action. What is ordinarily called communicating one's thoughts actually amounts to influencing the actions of another, although sometimes a deeper communication of souls is approached. Brouwer, however, was not always aloof from all efforts at social reform, as is shown by his participation, immediately after World War I, with the poet Frederik van Eeden, Mannoury, and others, in the Signific Circle, whose original goal, inspired by the abuses of propaganda during the war, was a far-reaching reform of language.

See also Hilbert, David; Intuitionism and Intuitionistic Logic; Kant, Immanuel; Mathematics, Foundations of; Number; Poincaré, Jules Henri; Schopenhauer, Arthur.

Bibliography

ADDITIONAL WORKS BY BROUWER

"De Onbetrouwbaarheid der Logische Principes." *Tijdschrift Voor Wijsbegeerte* 2 (1908): 152–158. Reprinted in *Wiskunde, Waarheid, Werkelijkheid.*

"Die Theorie der endlichen kontinuerlichen Gruppen." *Mathematische Annalen* 69 (1910): 181–202.

"Beweis der Invarianz der Dimensionenzahl." *Mathematische Annalen* 70 (1911): 161–165.

"Über Abbildung von Mannigfaltigkeiten." *Mathematische Annalen* 71 (1912): 97–115.

Intuitionisme en Formalisme. Amsterdam, 1912. Reprinted in *Wiskunde, Waarheid, Werkelijkheid.* Translated by A. Dresden as "Intuitionism and Formalism." *Bulletin of the American Mathematical Society* 20 (1913): 81–96. Reprinted, with other writings by Brouwer, in *Philosophy of Mathematics: Selected Readings,* edited by Paul Benacerraf and Hilary Putnam. Englewood Cliffs, NJ: Prentice-Hall, 1964.

"Über den naturlichen Dimensionsbegriff." *Journal für die reine und angewandte Mathematik* 142 (1913): 146–152.

"Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten." *Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen,* Series A, 12 (1918–1919), Nos. 5 and 7.

Wiskunde, Waarheid, Werkelijkheid. Groningen: P. N. Noordhoff, 1919.

"Besetzt jede reele Zahl eine Dezimalbruchentwicklung?" *Mathematische Annalen* 83 (1921): 201–210.

"Begründung der Funktionenlehre unabhängig vom logischen Satz ausgeschlossenen Dritten." *Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen,* Series A, 13 (1923), No. 2.

"Über die Bedeutung des Satzes vom ausgeschlossenen Dritten in der Mathematik, insbesondere in der Funktionentheorie." *Journal für die reine und angewandte Mathematik* 154 (1924): 1–7. Translated in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931,* edited by John van Heijenoort. Cambridge, MA: Harvard University Press, 1965.

"Beweis, dass jede voile Funktion gleichmässig stetig ist." *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings.* Series A, 27 (1924): 189–194 (cf. the remarks in ibid., 644–646).

"Zur Begründung der intuitionistischen Mathematik." *Mathematische Annalen* 93 (1924): 244–257; 95 (1926): 453–473; 96 (1927): 451–489.

"Über definitionsbereiche von Funktionen." *Mathematische Annalen* 97 (1927): 60–75. Translated in van Heijenoort.

"Weten, Willen, Spreken." *Euclides* 9 (1933): 177–193.

"Synopsis of the Signific Movement in the Netherlands." *Synthese* 5 (1946): 201–208.

"Address to Prof. G. Mannoury." *Synthese* 6 (1947): 190–194.

"Consciousness, Philosophy, and Mathematics," in *Proceedings of the Tenth International Congress of Philosophy.* Amsterdam, 1949. Vol. II, pp. 1235–1249. Last section of this paper reprinted in Benacerraf and Putnam.

"Historical Background, Principles, and Methods of Intuitionism." *South African Journal of Science* 49 (1952): 139–146.

"Points and Spaces." *Canadian Journal of Mathematics* 6 (1954): 1–17.

WORKS ON BROUWER AND INTUITIONISM

Heyting, Arend. *Les Fondements des mathématiques. Intuitionnisme. Théorie de la démonstration.* Paris: Gauthier-Villars, 1955. Expanded version of German original, *Mathematische Grundlagenforschung. Intuitionismus. Beweistheorie.* Berlin: J. Springer, 1934.

Heyting, Arend. *Intuitionism: An Introduction.* Amsterdam: North-Holland, 1956. This work by Heyting, and the one above, contain the most comprehensive bibliographies on intuitionism.

Charles Parsons (1967)

BROUWER, LUITZEN EGBERTUS JAN [ADDENDUM]

Luitzen Egbertus Jan Brouwer was one of the first to clearly distinguish between language and metalanguage, as well as to distinguish between mathematics and metamathematics. Published in his dissertation from 1907—written in Dutch—these ideas did not disseminate quickly, although they soon found fertile ground when Brouwer explained them to David Hilbert on the beach of Scheveningen, Holland, in 1909.

Brouwer also had conversations with Ludwig Wittgenstein (Vienna, 1928), Edmund Husserl (Amsterdam, 1928), and Kurt Gödel (Princeton, New Jersey, 1953). These conversations may equally have been of philosophical interest, but little seems to be known about their actual contents.

Throughout his life, Brouwer explored a deep interest in the history and practice of mysticism, yet this had no effect on the content of intuitionistic mathematics. In fact, in Brouwer's view, not to engage in even the simplest mathematics is a necessary condition for obtaining mystical insight, and the other way around. He reasoned that, whereas mathematics comes into being with the perception of a move of time, abolishing that perception is a necessary condition for the return of consciousness to its "deepest home" (Brouwer 1975, p. 480).

A note by Brouwer in the margin of an offprint (kept in the Brouwer archive, Utrecht, Holland) of his 1928 lecture "The Structure of the Continuum" (Brouwer 1975, pp. 429–440) shows that he held that the introduction of choice sequences did not make the intuitive continuum dispensable: "the continuum is still the result of the Ur-intuition." One can make do with just the choice sequences as far as the mathematical modeling of the continuum is concerned, but from a philosophical point of view, its intuitive givenness remains: Continuity and discreteness are mutually dependent and irreducible polarizations of the fundamental intuition of mathematics.

While from the beginning of his career Brouwer had combined mathematical and philosophical work, it was upon the theft of his mathematical notebook from a tram

in Brussels, Belgium, in 1929, that he came to despair of the continuation of his mathematical research. Perhaps the difficulties he foresaw in reconstructing these notes made him want to concentrate on philosophy instead. However, there were two other setbacks around the time that, in the long run, proved so devastating to Brouwer as to thwart whatever career plans he may have had. One setback was his conflict with Hilbert over the German journal *Mathematische Annalen*, on the editorial board of which they both served. The direct outcome of this dispute was—as intended by Hilbert—Brouwer's expulsion from the journal's board. The other setback was Brouwer's priority conflict with Karl Menger over the correct definition of *dimension*. Through the emotional and mental toll these fights took from someone of Brouwer's constitution, his creative work was, for the most part, brought to a halt (he would resume his work after 1945).

See also Gödel, Kurt; Hilbert, David; Husserl, Edmund; Intuitionism and Intuitionistic Logic; Mathematics, Foundations of; Mysticism, History of; Wittgenstein, Ludwig Josef Johann.

Bibliography

The Brouwer Archive is kept at the Department of Philosophy at Utrecht University, the Netherlands.

WORKS BY BROUWER

Collected Works. Vol. 1: *Philosophy and Foundations of Mathematics*, edited by Arend Heyting. Amsterdam: North-Holland, 1975.

Collected Works. Vol. 2: *Geometry, Analysis, Topology and Mechanics*, edited by Hans Freudenthal. Amsterdam: North-Holland, 1976.

Brouwer's Cambridge Lectures on Intuitionism, edited by Dirk van Dalen. Cambridge, U.K.: Cambridge University Press, 1981.

Intuitionismus, edited by Dirk van Dalen. Mannheim, Germany: BI-Wissenschaftsverlag, 1992.

"Life, Art and Mysticism." Translated by Walter van Stigt. *Notre Dame Journal of Formal Logic* 37 (3) (1996): 389–429. Introduction, also by Walter van Stigt, pp. 381–387.

WORKS ON BROUWER AND INTUITIONISM

Dummett, Michael. *Elements of Intuitionism*. Oxford: Clarendon Press, 1977. 2nd, rev. edition, 2000.

Hesseling, Dennis. *Gnomes in the Fog. The Reception of Brouwer's Intuitionism in the 1920s*. Basel: Birkhäuser, 2003.

Heyting, Arend. *Intuitionism. An Introduction*. Amsterdam: North-Holland, 1956. 2nd, rev. edition, 1966. 3rd, rev. edition, 1971.

Heyting, Arend. *Mathematische Grundlagenforschung, Intuitionismus, Beweistheorie*. Berlin: Springer, 1934. Expanded French translation: *Les Fondements des mathématiques. Intuitionnisme. Théorie de la démonstration*. Paris: Gauthiers-Villars, 1955.

Largeault, Jean. *Intuition et Intuitionisme*. Paris: Vrin, 1993.

Mancosu, Paolo, ed. *From Hilbert to Brouwer: The Debate on the Foundations of Mathematics in the 1920s*. Oxford: Oxford University Press, 1998.

Placek, Tomasz. *Mathematical Intuitionism and Intersubjectivity: A Critical Exposition of Arguments for Intuitionism*. Dordrecht, Holland: Kluwer, 1999.

Tieszen, Richard. *Mathematical Intuition. Phenomenology and Mathematical Knowledge*. Dordrecht, Holland: Kluwer, 1989.

Tieszen, Richard, ed. *Perspectives on Intuitionism*. Special issue, *Philosophia Mathematica* 6 (2–3) (1998).

Tieszen, Richard, ed. *Phenomenology and Mathematics*. Special issue, *Philosophia Mathematica* 10 (2) (2002).

Troelstra, Anne. *Choice Sequences: A Chapter of Intuitionistic Mathematics*. Oxford: Clarendon Press, 1977.

van Atten, Mark. *On Brouwer*. Belmont, CA: Wadsworth, 2004.

van Dalen, Dirk. *Mystic, Geometer, and Intuitionist*. Oxford: Clarendon Press, 1999.

van Dalen, Dirk, and Anne Troelstra. *Constructivism in Mathematics*. 2 vols. Amsterdam: North-Holland, 1988.

van Stigt, Walter. *Brouwer's Intuitionism*. Amsterdam: North-Holland, 1990.

Mark van Atten (2005)

BROWN, THOMAS
(1778–1820)

Thomas Brown, a philosopher on the periphery of the common sense school, was born at Kirkmabreck in Scotland. Radically opposed eighteenth-century traditions met in him. He shared with the common sense school, which derived from Thomas Reid, a number of its metaphysical doctrines and its appeal to intuitive truths; and he was also Reid's tireless critic. Philosophy, for Brown, was very largely "analysis": analysis of what he regarded as darkened notions, designed to exhibit their character free from spurious mystery and complication; analysis of the genuinely complex into its elementary constituents and of the deceptively simple into its real complexity. He saw Reid as a great resister of analysis. In the procedure of analysis Brown was influenced by French empiricism in the line of descent from Étienne Bonnot de Condillac.

During the course of his studies at the University of Edinburgh, Brown attended the lectures given by Dugald Stewart, Reid's close adherent. He subsequently graduated in medicine. In 1798 he published a criticism of the *Zoonomia* of Erasmus Darwin and in 1804 a defense of David Hume's account of causal relations (enlarged in 1806 and again in 1818, when it appeared under the title *Inquiry into the Relation of Cause and Effect*). Brown was

among the first of the contributors to the *Edinburgh Review* (he attacked Immanuel Kant in the second number of the *Review*). In 1810 he was appointed conjoint professor of moral philosophy with Stewart and took over the teaching duties of the chair. His lectures were a dazzling success; they were published after his death and went through many editions in a few years.

CAUSE AND EFFECT

Brown's views on causation typically combined an empiricist analysis with what he called a principle of intuitive belief. He defined a cause as "that which immediately precedes any change, and which, existing at any time in similar circumstances, has been always, and will be always, immediately followed by a similar change" (*Cause and Effect,* p. 13). Brown thought that if we reflect with sufficient patience and imagination, we can see that this definition exhausts the notion of a cause. To suppose that a cause is something more than the antecedent of an invariable consequent is to suppose that we might know all the unfailing regularities of nature and yet have no conception of a causal connection. Material and volitional agents, Brown argued in detail, do not differ in agency; all agency is the same. The omnipotence of God resides simply in the fact that whenever he wills anything, his will is "immediately and invariably *followed* by the existence of its object" (p. 103).

In tracing the sources of the complex illusion which, he thought, hangs over the relation of cause and effect, Brown emphasized the power of metaphor to mislead. Thus, things that are connected or bound together dependably go together; from this circumstance various figurative expressions enter the language and their figurative character is unnoticed. No bond or connection between causally connected events ever presents itself; yet unless we shift our attention from words to things, we shall easily suppose that it must be insensibly present. Experience (coupled with a kind of negative insight) enables us to see that the causal relation is merely one of sequence; but on what authority do we import the notion of invariableness into this sequence? Brown maintained that we are intuitively certain that the same antecedents will always be followed by the same consequents.

THE WILL

Under Brown's analysis, mystery vanished from the will: will is an amalgam of desire and the belief that one has it in one's power to realize the desire; there is no further, indefinable operator in our voluntary actions. Brown was not impressed by denials of the identity of will and desire on the ground that there can be opposition between them—Reid had said, "We may desire what we do not will, and will what we do not desire." When the types of situation referred to are looked at more carefully, Brown said, the opposition is seen to lie between desire and desire, and to be terminated by the desire upon which action immediately depends.

CONSCIOUSNESS

The examination of consciousness that provides data for the philosophy of mind is not, in Brown's opinion, conducted by consciousness. Once again, he saw entities as having been multiplied beyond necessity and, in this case, beyond possibility. He maintained that consciousness is not, as some philosophers have supposed, a surveyor of the mind's various states as they occur; rather, it is constituted by them. To suppose that "the same indivisible mind" could exist at the same time in two different states, one of them an object to the other, is "a manifest absurdity" (*Philosophy of the Human Mind,* Lectures XI and XII). What is thought of as an introspective examination of mental phenomena is therefore, strictly speaking, retrospective.

Below the phenomena of the mind, analysis encounters metaphysical bedrock. Let us imagine, Brown said, a man born with fully matured powers and a completely blank mind. Let him now be allowed a single sensation. This will be his total consciousness. Let a second sensation be added and let him be made to recall the first. He will then come to a recognition of something different from either—of himself as their common subject. The conviction that we exist with an "absolute" identity through time is intuitive and irresistible; only the circumstances in which it arises afford matter for inquiry. This identity is the prerogative of our minds; "some sort of identity of the body" is associated with it in our ordinary ideas about "sameness of person" (Lecture XII).

PERCEPTION

Brown's most subtle analyses occur in his theory of perception. His general problems were to explain how we come to know of the existence of an external, physical world and to specify the precise content of this knowledge. He was very conscious of the danger of question-begging assumptions; he maintained that at every turn we take externality for granted, and that all our language implies it. ("There is no distinct vocabulary of scepticism." Lecture XXII). Brown considered that our original awareness of things in their externality—their independence of our perception—is brought about by means of

sensations commonly but inaccurately ascribed to touch. The sensations belonging to other senses acquire an external reference by association with these.

Brown proceeded first to reductive simplification: the various tangible qualities were maintained to be various modifications of either extension or resistance. He then went on to disclose and systematize the complexity of sensations involved in our tactual relations with things. He argued that sensations of mere touch do not primitively inform us of extension and externality. We derive the notion of spatial extension from our repeated experience of the temporal succession of muscular feelings in the movements of arms and fingers. When a familiar series of these feelings is interrupted by feelings of resistance to muscular effort—as, for example, our fingers closed around an object—we become aware for the first time of something separate from ourselves and learn something of its dimensions. Physical objects were, for Brown, essentially extended, resisting objects; but before his argument has ended, extension and resistance seem to have become merely phenomenal and, in their unperceived existence, to have disappeared into their unknown causes.

MORAL THEORY

Brown's zeal for simplification is nowhere more conspicuous than in his moral theory. The distinctions, for example, between the obligatoriness, rectitude, and merit of an action are simply a matter of tense: contemplated before performance, the action is "obligatory"; in performance, it is "right"; and it is "meritorious" afterward. And what makes it so is the "emotion" of approval it arouses in us when we are in a fit state of mind to form a moral judgment—an emotion in no way arbitrary, for as morally definitive it proceeds from constitution of human nature. The strength and elevation of Brown's moral sentiments assisted his great, brief reputation.

See also Common Sense; Condillac, Étienne Bonnot de; Darwin, Erasmus; Hume, David; Kant, Immanuel; Reid, Thomas; Stewart, Dugald.

Bibliography

Principal works by Thomas Brown are *Inquiry into the Relation of Cause and Effect* (Edinburgh, 1818); and *Lectures on the Philosophy of the Human Mind* (Edinburgh, 1820). Selections from Brown appear in Daniel Sommer Robinson, *The Story of Scottish Philosophy* (New York: Exposition Press, 1961).

For literature on Brown, see David Welsh, *Account of the Life and Writings of Thomas Brown* (Edinburgh: W. and C. Tait, 1825), and François Réthoré, *Critique de la philosophie de Thomas Brown* (Paris, 1863). Sir William Hamilton, *Discussions on Philosophy and Literature,* 2nd ed. (Edinburgh: Longman, Brown, Green and Longmans, 1853), pp. 39–99, attacks Brown in defense of Reid. Brown is defended against Hamilton by John Stuart Mill in *An Examination of Sir William Hamilton's Philosophy* (London: Longman, Green, Longman, Roberts and Green, 1865), Ch. 10.

There are chapters on Brown in James McCosh, *The Scottish Philosophy* (New York: R. Carter, 1875), and in Henry Laurie, *Scottish Philosophy in Its National Development* (Glasgow: J. Maclehose, 1902). An indication of Brown's influence in America can be found in Terence Martin, *The Instructed Vision* (Bloomington: Indiana University Press, 1961).

S. A. Grave (1967)

BROWNSON, ORESTES AUGUSTUS

(1803–1876)

Orestes Augustus Brownson, a Transcendentalist philosopher and journalist, was born in Stockbridge, Vermont. He had little formal education. Until 1822 he belonged to the Congregationalist Church; he then joined the Presbyterians but was quickly repelled by their depreciation of human reason and by the Calvinist doctrine of predestination. In 1824 he became a Universalist, being ordained a minister in 1826. Three years later he abandoned Christianity and joined the socialist sect of Robert Dale Owen and Fanny Wright; at this time he wrote in behalf of the Workingmen's Party. He was reconverted to the Christian religion in 1832, when he joined the Unitarians.

Brownson was introduced to philosophy in 1833, through the works of Victor Cousin, whose disciple he remained for ten years. Cousin was warm in his praise of Brownson as a philosopher. Though Brownson later criticized Cousin's philosophy for its eclecticism and psychologism, he always remained under its influence. His reading of Immanuel Kant and the Italian idealist Vincenzo Gioberti were major factors in shaping his mature philosophy. For a while he was a member of the Transcendentalist group that met in Boston and at Brook Farm, but he considered their thinking poorly grounded and undisciplined.

In 1838 he founded the *Boston Quarterly Review,* which in 1842 was merged with the *U.S. Magazine and Democratic Review of New York.* In 1844, he was received into the Catholic Church. The same year he founded *Brownson's Quarterly Review,* which he published, except

for the years 1865–1872, until 1875. Most of Brownson's numerous articles and reviews appeared in this publication. His most important book was *The American Republic: Its Constitution, Tendencies, and Destiny.*

Although Brownson was a deeply religious thinker, he insisted that philosophy should begin not with authority or faith, but with data of reason. He criticized the notion of Christian philosophy proposed by the *Annales de philosophie chrétienne* for failing to do justice to the rational nature of philosophy.

Like Cousin, he made the starting point of philosophy the analysis of thought, stressing, in opposition to Cousin, its objective, rather than its subjective, side. All thought, he maintained, presupposes the presence of an object that can be analyzed into three elements: the ideal, the empirical, and the relationship between them. The ideal is the a priori element in all thought; it is that which makes any experience intelligible. The ideal is not a Kantian category, which Brownson interpreted to be a subjective form, but a necessary aspect of the object of knowledge. Since the object must be real in order to present itself to thought, its ideal, or content, must also be real. Further analysis revealed that this content includes both necessary and contingent "being," which Brownson identified respectively with God and creatures. God is a necessary and independent being; creatures are dependent existences, so called because they stand outside (*exstare*) their cause. Hence Brownson adopted the "ideal formula" of Gioberti: "Being creates existences" (*Ens creat existentias*). Accordingly, creative being is present to the mind in all its thinking; it alone makes thought possible.

Brownson defended himself against the charge of ontologism, which was condemned by Rome in 1861, on the ground that he did not teach that we have an immediate intuition of God, but only of being. Though being is God himself, we discover this only by rational analysis.

In his early days, Brownson believed in the divinity of humanity and the infallibility of the popular will. Political experience in later life convinced him of the absurdity of these notions. He rejected the idea that government and law have a purely human origin. Only in a qualified sense did he admit that governments derive their powers from the assent of the governed. All power ultimately comes from God; he alone has absolute sovereignty. Brownson thought the American constitution more nearly perfect than others because it recognizes the existence of the Creator and of God-given rights of individuals, which the government is bound to respect and protect.

See also Being; Cousin, Victor; Gioberti, Vincenzo; Kant, Immanuel; New England Transcendentalism.

Bibliography

WORKS BY BROWNSON

The Works of Orestes A. Brownson, 20 vols. Edited by H. F. Brownson. Detroit: T. Nourse, 1822–1907.

WORKS ON BROWNSON

Brownson, H. F. *Orestes A. Brownson's Early, Middle, and Latter Life,* 3 vols. Detroit: H. F. Brownson, 1898–1900.

Cook, T. I., and A. B. Leavelle. "Orestes A. Brownson's *The American Republic.*" *Review of Politics* 4 (1) (January 1942): 77–90; 4 (2) (April 1942): 173–193.

Farrell, Bertin. *Orestes Brownson's Approach to the Problem of God.* Washington, DC: Catholic University of America Press, 1950.

Fitzsimons, M. A. "Brownson's Search for the Kingdom of God: The Social Thought of an American Radical." *Review of Politics* 16 (1) (January 1954): 22–36.

Maynard, Theodore. *Orestes Brownson, Yankee, Radical, Catholic.* New York: Macmillan, 1943.

McMahon, F. E. "Orestes Brownson on Church and State." *Theological Studies* 15 (1954): 175–228.

Parry, S. J. "The Premises of Brownson's Political Theory." *Review of Politics* 16 (2) (April 1954): 194–211.

Raemers, Sydney A. *America's Foremost Philosopher.* Washington, DC: St. Anselm's Priory, 1931.

Schlesinger, A. M., Jr. *Orestes A. Brownson: A Pilgrim's Progress.* Boston: Little, Brown, 1939.

Armand A. Maurer (1967)

BRUNNER, EMIL
(1899–1966)

Emil Brunner was a Swiss theologian. He was educated in Switzerland and served in the Swiss army in 1914. Later he became a pastor and then professor of theology at Zürich. He participated extensively in the work of the World Council of Churches and also for a time in the Moral Re-Armament movement. He lectured on theology in many countries, notably in the United States, Japan, and Scotland.

Brunner's earliest theological positions were typical of Swiss and German Protestantism before 1914. He accepted the liberal theological emphasis on the social and ethical aspects of the Gospel, as well as its stress upon the rational alliance between philosophy and theology. Even in his earliest theological writings he exhibited his personal interest in philosophy in a well-informed discussion of Edmund Husserl, *Das Symbolische in der*

religiösen Erkenntnis (Tübingen, 1914). But after World War I he embarked upon a critique of liberalism that at first seemed to make him the natural ally of Karl Barth. His *Die Mystik und das Wort* (Tübingen, 1924) is a hostile discussion of Friedrich Schleiermacher's attempt to find a basis for Christianity in the general form of religious experience. Against this, Brunner poses the distinctive claims of Christian revelation, a revelation that cannot be discovered or appropriated through the use of criteria derived from natural theology or private experience.

The adjective much used of Brunner's (and also Barth's) concept of revelation was "dialectical." Theology is dialectical in that its attempts to grasp revelation necessarily involve the use of concepts that in purely philosophical discourse would cancel each other out. So the contradictions that arise, for example, in combining belief in divine goodness and omnipotence with an acknowledgment of the occurrence of physical evil are taken by the dialectical theologian to be simply manifestations of the necessarily paradoxical character of theological concepts. Contradiction is not a sign of intellectual failure, but of the inadequacy of our intellects before the splendor of divine revelation. Thus, if we try to use our ordinary criteria of consistency, we shall fail to grasp revelation at all. The major reason for this is that we shall be at fault if we try to understand revelation as consisting in a set of propositions. When God reveals himself, he does so as a person. Revelation is the act of a person, not the setting out of a doctrine.

It is for this reason that philosophy must necessarily limit its aspirations. The god of whom philosophy speaks is not the God of Christian revelation for at least two reasons. First, he is an inferred entity; and second, he is an object. It is not always clear whether Brunner believes that what philosophy says about God is false or simply inadequate. At times it seems clear that it is the former, yet Brunner is unlike Barth in the stress he puts upon the positive contribution that philosophy can make to theological thinking. Philosophy's role is to be critical, in the Kantian sense. It is to exhibit the limitations of human reason, and so to prevent speculative reason from attempting to occupy territory that belongs by right to revelation.

In revelation, God encounters man as person to person; man cannot argue his way to God by philosophy or discover God apart from the biblical revelation, yet when God calls, man at least can answer. Even this minimal concession to human powers brought Brunner into conflict with Barth. Barth's position, which he outlined in the short, bitter pamphlet *Nein! Antwort an Emil Brunner* (1934), is that man, totally corrupted by the Fall, cannot advance an inch toward God by means of his natural powers. Grace has to supply even the capacity of responding to God's initiative. Brunner, who always feared the depiction of men as mere puppets, laid great stress on the natural man's capacity for speech and for elementary rationality as a precondition of any response to God.

The contrast between Brunner's theology and both liberalism on the one hand and Barthianism on the other is most marked in Brunner's ethics and social philosophy. Unlike Barth, Brunner believes that the basis for a natural ethics, even if a very limited one, exists. He revives the idea, which is found in Luther, of orders of creation. An order of creation is a social institution or practice of ordinary human origin, not derived from revelation, but shown by biblical evidence to have divine authorization. So Christ blessed monogamy in his appearance at the wedding at Cana and in his utterances about marriage; so he expressed the divine source of the state's authority when he said, "Render unto Caesar …" These orders supply human beings with norms to whose validity revelation itself testifies, but for knowledge of which revelation is not necessary. Such norms have the negative function of restraining sin, rather than any positive role. Brunner differs from liberal theology in his belief that no secular morality can hope to provide a satisfactory way of life, but is bound to founder on the sinfulness of human nature.

The key way in which sin manifests itself in human life is in the failure of men, both in theory and in practice, to understand themselves as persons. (It should be noted that it is not clear how far Brunner uses the word *person* in the same sense when he speaks of God as a person and men as persons. He speaks of God as the "original" person and of men as "derivative" persons, and says that before the Fall men were persons as God is a person. Some analogy is intended, but we are not told how strong the analogy is.) Brunner makes the position of philosophy in respect to human beings parallel to that which he gives it in respect to the knowledge of God. Philosophy as philosophy cannot grasp men as persons, but only as objects and inferred entities. The ghost of the view of both God and the self as Kantian noumena haunts his thought at this point. But it is not only in philosophy that the secular view of man is inadequate. In practice, too, men continually reject their status as persons.

They do this by seeking to be autonomous. The will, as the center of man's rebellion against God, seeks continually to be its own master. The ideal of the self-sufficient individual is one human ideal that must be rejected.

Its counterpart, the concept of man in the mass—collective man—is equally subhuman. But secular thought provides us with no adequate basis for rejecting these alternatives. Only revelation can do this, for it is only in revelation that we discover not only God as a person but also ourselves as persons. This is where Brunner's doctrine of atonement finds its place. What Jesus Christ showed us in his life, death, and resurrection was a love that alone can break our rebellious self-will and that alone can provide us with a model for goodness. Secular ethics can at best exhibit the kind of goodness that can defeat depersonalization as a hypothetical possibility. The revelation of Christ alone makes it actual. Revelation, however, does not provide us with a code that we can then detach from its origin and live by. We must return continually to revelation for renewal. This is in part because of the character of human sin, but it is also in part because we must reassert the personal character of social life in new contexts.

According to Brunner, the depersonalization that is a consequence of technology is distinctive of the contemporary context. Men are degraded to the status of tools and means. The social incarnation of this process is the totalitarian state. For Brunner, totalitarianism is the category ultimately opposed to that of true community, and both Nazism and communism are forms of it. This political judgment took Brunner into further public argument with Karl Barth, on the grounds that Barth's theological views obliterate the moral differences between rival political systems by insisting on the sinfulness of human nature as such.

See also Barth, Karl; Human Nature; Husserl, Edmund; Revelation; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

ADDITIONAL WORKS BY BRUNNER

"Religionsphilosophie evangelischer Theologie." In *Handbuch der Philosophie* (Munich: R. Oldenbourg, 1927). Translated by A. J. D. Farrer and B. L. Woolf as *Philosophy of Religion from the Point of View of Protestant Theology.* Edinburgh, 1937.

Der Mittler. Tübingen, 1927. Translated by Olive Wyon as *The Mediator.* New York: Macmillan, 1934.

Das Gebot und die Ordnungen. Tübingen, 1932. Translated by Olive Wyon as *The Divine Imperative.* London: Lutterworth Press, 1937.

Der Mensch im Widerspruch. Berlin: Furche, 1937. Translated by Olive Wyon as *Man in Revolt.* London: Scribners, 1939.

Christianity and Civilisation. London: Scribners, 1948–1949. Parts I and II.

CRITICAL STUDY

Kegley, Charles W., ed. *The Theology of Emil Brunner.* New York: Macmillan, 1962.

Alasdair MacIntyre (1967)

BRUNO, GIORDANO

(1548–1600)

Giordano Bruno, the most famous of the Italian philosophers of the Renaissance, was born at Nola, near Naples. At an early age he entered the Dominican order and became an inmate of the Dominican convent in Naples. In 1576 he was accused of heresy and fled, abandoning the Dominican habit. Thereafter he wandered through Europe. After visiting Geneva, and lecturing on the *Tractatus de Sphaera Mundi* of Sacrobosco at Toulouse, Bruno reached Paris in 1581. Here he gave public lectures that attracted the attention of King Henri III, and published two books on the art of memory that reveal him as greatly influenced by that textbook of Renaissance magic, the *De Occulta Philosophia* of Henry Cornelius Agrippa, from which he quotes lists of magic images of the stars, incantations, and other occult procedures. Bruno as a Renaissance magus, in line of descent from the learned philosophical magic inaugurated by Marsilio Ficino, is already present in these books. The title of one of them, *De Umbris Idearum* (Shadows of Ideas), is taken from the necromantic commentary on the *Sphere* of Sacrobosco by Cecco d'Ascoli, whom Bruno mentions admiringly in other works. It may be inferred that the lectures at Toulouse were probably based on this commentary.

Early in 1583 Bruno went to England with letters of recommendation from Henri III to the French ambassador in London. He lived in the French embassy during the two years he spent in England, and the ambassador protected him from the tumults aroused by his writings, which were clandestinely printed in London. These included the *Triginta Sigilli* (Thirty seals), an extremely obscure work on his magic art of memory; those who manage to reach the end of it find an advocacy of a new religion based on love, art, magic, and mathesis. It is dedicated to the vice-chancellor and doctors of the University of Oxford in high-sounding terms in which Bruno announces himself as "the waker of sleeping souls, tamer of presumptuous and recalcitrant ignorance, proclaimer of a general philanthropy."

In June 1583 the Polish prince Albert Alasco (Laski) visited Oxford and was entertained with public disputations. Bruno was in his train, and, according to a recently

discovered account by George Abbot, afterward archbishop of Canterbury, Bruno returned to Oxford after the party had left and delivered, uninvited, lectures that were largely a repetition of Marsilio Ficino's work on astral magic, the *De Vita Coelitus Comparanda* (On drawing down the life of heaven), although he also maintained Nicolas Copernicus's opinion "that the earth did go round and the heavens did stand still." Abbot says that Bruno was induced to discontinue the lectures when the plagiarism from Ficino was pointed out to him.

While in England, Bruno published five dialogues in Italian. In *La cena de le ceneri* (The Ash Wednesday supper; 1584) he defends his version of the Copernican theory against Oxford "pedants," a reflection of his visit to Oxford. In *De la causa, principio e uno* (1584) he apologizes for the storms aroused by his attack on Oxford, but makes matters worse by defending the friars of pre-Reformation Oxford, whom he prefers to their Protestant successors. The *De l'infinito, universo e mondi* (1584) is an exposition of his vision of an infinite universe and innumerable worlds. The *Spaccio de la bestia trionfante* (The expulsion of the triumphant beast; 1584) envisages a universal moral and religious reform and is dedicated to Sir Philip Sidney. The *Cabala del cavallo pegaseo* (Cabal of the horse Pegasus; 1585) indicates Bruno's adaptation of the Jewish kabbalah. The *De gli eroici furori* (On heroic enthusiasms; 1585) also dedicated to Sidney, is in the form of a sonnet sequence with commentaries expounding the philosophical and mystical meanings of the poems. It is upon this series of most striking and brilliant works, in which Bruno appears as the propagator of a new philosophy and cosmology, a new ethic and religion, that his fame largely rests. They are all full of Hermetic influences and are bound up with a complex religious, or politico-religious, mission for which Bruno believed he had the support of Henri III, and which cannot have been uncongenial to the French ambassador, Michel de Castelnau de Mauvissière, to whom three of the books are dedicated. Sidney's reactions to Bruno are unknown.

Late in 1585 Bruno returned to Paris, where he delivered an address on his philosophy in the Collège de Cambrai, arousing strong opposition, and where he had a curious controversy with Fabrizio Mordente about the compass that Mordente had invented. Paris was in a disturbed state, on the eve of the wars of the League, and Bruno's activities added to the "tumults," from which he fled in 1586 and began his travels through Germany. He was favorably received at the University of Wittenberg, and during his stay there he wrote a number of works, particularly on the art of Ramón Lull, to which he attached great importance and which he believed he understood better than Lull himself. From Wittenberg he went to Prague, where he tried to obtain the favor of Emperor Rudolph II with his *Articuli Adversus Mathematicos* (1588), in which he states that he is strongly against mathematics, which he regarded as a "pedantry" lacking in deep magical insight into nature. His objection to Copernicus as a "mere mathematician" had been on similar lines. The work is illustrated with magical diagrams, representing what he called his mathesis, and its preface outlines a movement of tolerance and general philanthropy that is to replace sectarian bitterness. He next spent some time at Helmstedt, where he enjoyed the favor of the reigning duke, Henry Julius of Brunswick-Wolfenbüttel, and made a speech in praise of the late duke in which he outlined his program of moral reform in language similar to that used in the *Spaccio de la bestia trionfante.* It was probably while at Helmstedt that Bruno wrote the *De Magia* and other works on magic unpublished in his lifetime.

With the money Henry Julius gave him for the oration, Bruno went to Frankfurt to have printed the Latin poems he had written during his wanderings. These were the *De Innumerabilibus, Immenso et Infigurabili,* the *De Triplici Minimo et Mensura,* and the *De Monade Numero et Figura,* all of which were printed by John Wechel in 1591. In these Latin poems, written in a style imitating Lucretius, Bruno expresses his philosophical and cosmological speculations in their final form. Like the Italian dialogues on these themes, the Latin poems are full of Hermetic influences, particularly of the mathesis, or magical numerology, which Bruno had been further developing during his travels. He also published the last of his books on his magical arts of memory at Frankfurt.

TRIAL AND DEATH

In August 1591, Bruno returned to Italy at the invitation of a Venetian nobleman who wished to learn the secrets of his art of memory. There can be little doubt that Bruno was encouraged to take this step by the hopes of greater religious toleration aroused by the conversion of Henri IV of France. Bruno had in his baggage the manuscript of a book he intended to dedicate to Pope Clement VIII. It is strange that one who had stated in his published works that Christ was a magus and that the magical religion of the Egyptians was better than Christianity should have felt that he could place himself with impunity within reach of the Inquisition. Bruno seems, however, always to have sincerely believed that his religious and moral reform could take place within a Catholic framework. He

was arrested in Venice and thrown into the prisons of the Inquisition. At the end of the Venetian trial he recanted his heresies, but was sent to Rome for another trial. Here he remained in prison for eight years, at the end of which he was sentenced as a heretic (having refused, this time, to recant) and was burned alive on the Campo de' Fiori.

Although the actual *processo* stating on what grounds he was condemned is not extant, it seems most probable that Bruno was burned as a magician, as an "Egyptian" who had been propagating throughout Europe some movement the nature of which remains mysterious, although it may well be connected with the origins of Rosicrucianism and of Freemasonry. His philosophical views in themselves can have had little to do with the condemnation, unless insofar as they, too, were associated with the movement.

LATER INTERPRETATION

In the seventeenth century there was a conspiracy of silence about Bruno and his reputation. Where the silence was broken, he usually appeared in the character of a diabolical magician. It was rumored that he had made a speech in praise of the devil at Wittenberg (Pierre Bayle and Gottfried Wilhelm Leibniz heard this story). In the eighteenth century he was interpreted by John Toland as a deist. The nineteenth century rediscovered Bruno and read its own beliefs and attitudes into his works. It was then that he appeared as the martyr for modern science and the Copernican theory, and statues were erected in his honor by anticlericals in Italy. The crudity of this approach was modified in later philosophical studies of Bruno, but the attempt to isolate a philosophy or a metaphysics from his works and to discuss his thought in a context of straight history of philosophy meant that large areas in his writings must be disregarded as unimportant or unintelligible. Moreover, no coherent philosophical system could be extracted in this way, as Leonardo Olschki saw when he criticized Bruno as a confused thinker. But when Bruno is placed in the context of the Renaissance Hermetic tradition, his philosophy, his magic, and his religion can all be seen as forming part of an outlook on nature and on man which, however strange, is nevertheless perfectly coherent within its own premises.

HERMETIC PHILOSOPHY

The extraordinary prestige of the Hermetica in the Renaissance was encouraged by the belief that they were the writings of Hermes Trismegistus, an Egyptian sage who foretold Christianity and whose wisdom had inspired Plato and the Platonists. The Hermetic core in Renaissance Neoplatonism was an important factor in the revival of magic. Christian magi, like Ficino and Giovanni Pico della Mirandola, used some caution in their approach to the magical passages in the Hermetic *Asclepius,* which is the basis of the astral magic described by Ficino in his *De Vita Coelitus Comparanda.* These safeguards were largely abandoned by the magician Cornelius Agrippa and totally abandoned by Bruno, who adopted the position that the Hermetic magical religion was the true religion, the religion of nature in contact with its powers. The cure for the wars, persecutions, and miseries of contemporary Europe was a return to the magical religion of the Egyptians—hence the long quotations in the *Spaccio de la bestia trionfante* from the passages in the *Asclepius* describing the religious practices of the Hermetic pseudo Egyptians, ecstatically interpreted by Bruno as their worship of "God in things," and as a "profound magic" by which they were able to draw down cosmic powers into the statues of their gods. The lament for the Egyptian religion in the *Asclepius* was interpreted by Bruno as a lament for a better religion, destroyed by Christianity. Since Augustine had condemned these passages as referring to the wicked demon worship of the Egyptians, it is easy to see how Bruno's "demonic" reputation arose. Bruno's "Egyptian" religion included belief in metempsychosis, which he also derived from the Hermetic writings.

Bruno's views on religion are organically related to his philosophy, for the philosophy of the living Earth moving around the divine sun and of the innumerable worlds, moving like great animals with a life of their own in the infinite universe, is the animist philosophy of a magus who believes he can establish contact with the divine life of nature. The sun is frequently mentioned in the Hermetic writings as a god, and it is the chief of the astral gods worshiped in the religion described in the *Asclepius.* Ficino's use of the astral magic of the *Asclepius* was chiefly directed toward the sun, whose beneficent influences he sought to draw down through solar talismans and incantations.

BRUNO'S COPERNICANISM

That Bruno thought of the Copernican sun in the context of the magic of Ficino's *De Vita Coelitus Comparanda* is indicated in the report of his lectures at Oxford, in which he is said to have repeated the Ficinian text while also maintaining the opinion of Copernicus. This report fits in with passages in Bruno's works in which the sun appears in a magical context, and particularly with his

defense of the Copernican opinion against the Oxford doctors in *La cena de le ceneri,* where he describes Copernicus as "only a mathematician" who has not seen the true meaning of his discovery as he, Bruno, has seen it. When a speaker in these dialogues asks what is the cause of Earth's movement around the sun, the reply is an almost verbatim quotation from *Corpus Hermeticum* XII, in which Hermes Trismegistus explains that the energy of life is movement and that therefore nothing in the living universe is immobile, not even Earth. Bruno applied these words as an explanation of the cause of Earth's movement around the sun. The Copernican opinion had, for him, confirmed the "Egyptian" philosophy of universal animation. He also repeated from the same Hermetic treatise one of his most characteristic doctrines: that there is no death in nature, only change.

Thus Bruno's acceptance of Copernican heliocentricity did not rest on Copernicus's mathematical arguments. On the contrary, Copernicus as a mere mathematician was despised by him as a superficial person who had not understood the true meaning of his discovery. Bruno was always "against" mathematics. Although he had some acquaintance with the scientific basis of the Copernican theory, it was not on mathematical grounds that Bruno defended Copernicanism from reactionary Aristotelians, but on animist and magical grounds. In fact, when the passages on the sun in the different works are compared, it becomes apparent that Copernican heliocentricity was for Bruno a kind of celestial portent of the approaching return of "Egyptian" philosophy and religion. "Aristotelianism" was for Bruno a symbol of all that is dead and dry—or, as he would say, "pedantic"—in philosophy and religion (the two were for him inseparable), compared with his own philosophy and religion—in contact, so he believed, with living, divine nature.

NEW VISION OF THE UNIVERSE

The essence of the Hermetic writings is that they give a religious impulse toward the world. It is within the setting of the universe, not through any divine mediator, that the Hermetic gnostic achieves his religious experience. The closest parallel to Bruno's imaginative leap upward through the spheres is the description in the Hermetic *Pimander* of how man "leant across the armature of the spheres, having broken through their envelopes." So did Bruno break through the spheres in his ecstatic ascent to his new vision of the universe. The immediate source of his vision of infinite space and innumerable inhabited worlds was Lucretius's poem *De Rerum Natura,* But Bruno transformed the Epicurean and Lucretian notions by imparting animation to the innumerable worlds—a feature totally absent from Lucretius's universe—and by imparting the function of being an image of the infinite divinity to the infinite. The godless universe of Lucretius turns in the Brunian vision into a vast extension of Hermetic gnosis; in order to receive this within himself, man, that "great miracle," as he is defined in the *Asclepius,* must expand himself infinitely. The *magnum miraculum est homo* passage is quoted from Trismegistus near the beginning of the *De Immenso* as a preliminary to the new vision of the world to be revealed in the poem.

This infinitely extended All was nevertheless One. The unity of the All in the One is a basic theme of the Hermetic writings and also of Bruno's. The unity of the All in the One is for Bruno "a most solid foundation for the truths and secrets of nature. For you must know that it is by one and the same ladder that nature descends to the production of things and the intellect ascends to the knowledge of them; and that the one and the other proceeds from unity and returns to unity" (*De la causa, principio e uno,* in *Dialoghi italiani,* edited by Giovanni Aquilecchia, p. 329).

This is the philosophy conducive to magic—that the magus can depend on the ladders of occult sympathies running through all nature. When this philosophy is not only a magic but also a religion, it becomes the religion of the Hermetic pseudo Egyptians who, as Bruno says in the *Spaccio de la bestia triofante,* "with magic and divine rites … ascended to the height of the divinity by that same scale of nature by which the divinity descends to the smallest things by the communication of itself" (*Dialoghi italiani,* p. 777). Bruno's philosophy and religion are one and the same, and both are Hermetic. This accounts for the main aspects of his philosophy, his panpsychism and his monism, and also for the magic and the references to magical practices with which his books are filled.

Like all Renaissance magi, Bruno was a syncretist and drew from his vast reading many philosophies which had accreted to the Hermetic core. The pre-Socratics, Plato and the Platonists, the Scholastics (Bruno revered Thomas Aquinas as a great magus), Nicholas of Cusa—all were incorporated into the central theme. Bruno's chief textbook of magic was Agrippa's *De Occulta Philosophia*; he also used the conjuring books of Trithemius and admired, and perhaps practiced, the Paracelsian medicine.

ART OF MEMORY

The side of Bruno's work that he regarded as the most important was the intensive training of the imagination

in his occult arts of memory. In this he was continuing a Renaissance tradition that also had its roots in the Hermetic revival, for the religious experience of the Hermetic gnostic consisted in reflecting the universe within his own mind or memory. The Hermeticist believed himself capable of this achievement because he believed that man's *mens* is in itself divine and therefore able to reflect the divine mind behind the universe. In Bruno, the cultivation of world-reflecting magic memory becomes the technique for achieving the personality of a magus, and of one who believes himself to be the leader of a religious movement. Strange though these beliefs and practices are, Bruno had some profound things to say in his books on memory concerning the imagination, which he made the sole cognitive power (sweeping away the divisions of the Aristotelian faculty psychology by a kind of inner anti-Aristotelianism), and on the mental image in relation to the psychology of the "inspired" personality. When the magical aspect (which includes such practices as the use of talismans or images of the stars as mental images) is discounted or allowed for, Bruno's bold explorations of the inner world may become important to the historian of psychology.

SIGNIFICANCE AND INFLUENCE

The emphasis on the Hermetic and magical side of Bruno's thinking does not discredit his significant contribution to the history of thought. He exemplifies the Hermetic religious impulse as a motive force behind the imaginative formulation of new cosmologies. From within his own frame of reference, this highly gifted man made guesses that may have given hints to seventeenth-century thinkers. A notable example is his transformation of the Democritean atoms, of which he read in Lucretius, into magically animated monads; this may well have been a stage leading to Leibniz's monadology, and there are other curious links between Bruno and Leibniz. Although Bruno was obviously not in the line leading to the mathematical advances, his extraordinary vision of an immensely expanded universe, ruled by the laws of magical animism, may be said to prefigure, on the Hermetic plane, the new cosmology of the seventeenth century. Drained of its animism, with the laws of inertia and gravity substituted for the psychic life of nature as the principle of movement, Bruno's universe would turn into something like the universe of Isaac Newton, moving under laws placed in it by a God who is not a magician but a mathematician and a mechanic. In the Hermetic phase of European thought, which was the immediate prelude to the seventeenth-century revolution, Bruno is an outstanding figure. Regarding him in this light, the old legend of the martyrdom of the advanced thinker becomes almost true again, although not in the old sense.

See also Hermeticism.

Bibliography

Additional works by Bruno in Italian are *Dialoghi italiani,* edited by Giovanni Gentile and revised by Giovanni Aquilecchia (Florence, 1957), which contains all the Italian dialogues in one volume, and *Il candelaio,* a comedy, edited by Vincenzo Spampanato (Bari, 1923). In Latin, see *Opera Latine,* edited by Francisco Fiorentino, Vittorio Imbriani, C. M. Tallarigo, Felice Tocco, and Girolamo Vitelli, 3 vols. (Naples: Morano, 1879–1891), issued in eight parts; there is also a facsimile reprint (Naples and Florence, 1962).

Translations include *The Heroic Enthusiasts,* translated by L. Williams (London: Redway, 1887); *Des Fureurs héroïques,* translated by P.-H. Michel (Paris, 1954); *Giordano Bruno's The Heroic Frenzies,* translated by P. E. Memmo Jr. (Chapel Hill: University of North Carolina Press, 1964); *The Expulsion of the Triumphant Beast,* translated by A. D. Imerti (New Brunswick, NJ: Rutgers University Press, 1964); "On the Infinite Universe and Worlds," in D. W. Singer, *Giordano Bruno, His Life and Thought* (New York: Schuman, 1950), pp. 227ff.; and "Concerning the Cause, Principle and One," translated by D. W. Singer in S. Thomas Greenberg, *The Infinite in Giordano Bruno* (New York: King's Crown Press, 1950), pp. 77ff.

Documentary sources on Bruno's life are Vincenzo Spampanato, ed., *Documenti della vita di Giordano Bruno* (Florence, 1933), and Angelo Mercati, ed., *Il sommario del processo di Giordano Bruno* (Vatican City: Biblioteca Apostolica Vaticana, 1942). A biography is Vincenzo Spampanato, *Vita di Giordano Bruno* (Messina: Principato, 1921). On his trial, see Luigi Firpo, "Il processo di Giordano Bruno," in *Rivista storica italiana* 60 (1948): 542–597, and 61 (1949): 5–59.

For a bibliography of Bruno's works and of books and articles on him up to and including 1950, see Virgilio Salvestrini and Luigi Firpo, *Bibliografia di Giordano Bruno 1582–1950* (Florence, 1958).

Studies of Bruno include Felice Tocco, *Le opere latine di Giordano Bruno* (Florence: Sansoni, 1889), and *Le fonti più recenti del Bruno* (Rome, 1892); J. L. McIntyre, *Giordano Bruno* (London: Macmillan, 1903); Giovanni Gentile, *Giordano Bruno e il pensiero del Rinascimento* (Florence, 1920); Leonardo Olschki, *Giordano Bruno* (Halle, 1924), which has also been translated into Italian (Bari, 1927); Ernst Cassirer, *Individuum und Kosmos in der Philosophie der Renaissance* (Leipzig: Teubner, 1927), translated by Mario Domandi as *The Individual and the Cosmos in Renaissance Philosophy* (New York: Harper, 1963); Antonio Corsano, *Il pensiero di Giordano Bruno* (Florence: Sansoni, 1940); Eugenio Garin, *La filosofia* (Milan, 1947); Alexandre Koyré, *From the Closed World to the Infinite Universe* (Baltimore: Johns Hopkins Press, 1957); Paolo Rossi, *Clavis Universalis* (Milan, 1960); P.-H. Michel, *La cosmologie de G. Bruno* (Paris: Hermann, 1962); and Paul O. Kristeller, *Eight Philosophers of the Italian Renaissance* (Stanford, CA:

Stanford University Press, 1964). Also see Frances A. Yates, *The Art of Memory* (Chicago: University of Chicago Press, 1966).

OTHER RECOMMENDED TITLES

Byrum, Stephen C. "A Proposal for Considering Giordano Bruno's Influence on Descartes." *Philosophy Research Archives* 9 (1983): 303–336.

Calcagno, Antonio. *Giordano Bruno and the Logic of Coincidence: Unity and Multiplicity in the Philosophical Thought of Giordano Bruno.* New York: Peter Lang, 1998.

Decker, Kevin S. "The Open System and Its Enemies: Bruno, the Idea of Infinity, and Speculation in Early Modern Philosophy of Science." *American Catholic Philosophical Quarterly* 74 (4) (2000): 599–620.

Dynnik, M. A. "Man, Sun, and Cosmos in the Philosophy of Giordano Bruno." *Soviet Studies in Philosophy* 6 (1967): 14–21.

Gatti, Hilary. *Giordano Bruno and Renaissance Science.* Ithaca, NY: Cornell University Press, 1999.

Gatti, Hilary, ed. *Giordano Bruno: Philosopher of the Renaissance.* Aldershot, U.K.: Ashgate, 2002.

Gatti, Hilary. "The Natural Philosophy of Giordano Bruno." *Midwest Studies in Philosophy* 26 (2002): 111–123.

Granada, Miguel A. "Aristotle, Copernicus, Bruno: Centrality, the Principle of Movement and the Extension of the Universe." *Studies in History and Philosophy of Science* 35a (1) (2004): 91–114.

Maiorino, Giancarlo. "The Breaking of the Circle: Giordano Bruno and the Poetics of Immeasurable Abundance." *Journal of the History of Ideas* 38 (1977): 317–327.

Massa, Daniel. "Giordano Bruno's Ideas in Seventeenth-Century England." *Journal of the History of Ideas* 38 (1977): 227–242.

Paris, Paul-Henri. *The Cosmology of Giordano Bruno.* Translated by R. E. W. Maddison. Ithaca, NY: Cornell University Press, 1973.

Paterson, Antoinette Mann. *The Infinite Worlds of Giordano Bruno.* Springfield, IL: Thomas, 1970.

Voise, Waldemar. "Bruno on the Morality of the Inhabitants of the Infinite Universe and on the Cognitive Passion of Copernicus." *Dialectics and Humanism* 6 (1979): 115–123.

White, Michael. *The Pope and the Heretic: The True Story of Giordano Bruno, the Man Who Dared to Defy the Roman Inquisition.* New York: Morrow, 2002.

Yates, Frances A. *Giordano Bruno and the Hermetic Tradition.* New York: Routledge, 1999.

Frances A. Yates (1967)
Bibliography updated by Tamra Frei (2005)

BRUNSCHVICG, LÉON

(1869–1944)

Léon Brunschvicg, the French idealist philosopher, was born in Paris and educated at the Lycée Condorcet, where he won awards in science as well as in classics and philosophy. He received both the *licence ès lettres* and the *licence ès sciences* from l'École Normale Supérieure in 1891. During the following nine years he taught philosophy at lycées in Lorient, Tours, and Rouen. His doctoral thesis, *La modalité du jugement*, was presented to the Sorbonne in 1897, and published in Paris the same year. In 1900 he returned to Paris to teach at his old lycée, later moving to the Lycée Henri IV and l'École Normale de Sèvres. In 1909 he was named professor of general philosophy at the Sorbonne. Except for the period 1914–1918, when he served in the armed forces auxiliary and as adviser to the government on educational reform, Brunschvicg held various chairs at the Sorbonne until the German occupation of Paris in 1940. He then settled in Aix-en-Provence and finally in Aix-les-Bains until his death.

Brunschvicg was one of the founders of the *Revue de Métaphysique et de Morale* (1893) and of the Société française de Philosophie (1901). In 1919 he was elected to the Académie des Sciences morales et politiques, serving as president in 1932. A prolific writer, editor of Blaise Pascal, and well known for his studies of René Descartes and Benedict de Spinoza, Brunschvicg was a major figure in French intellectual life for nearly half a century.

The "critical idealism" of Brunschvicg primarily recalls Immanuel Kant's analysis of the conditions of knowledge, but Brunschvicg's method was historical rather than deductive: He wished to grasp the mind's activity as it has revealed itself in the history of mathematics, science, and philosophy. In general perspective, Brunschvicg may be seen as heir to two currents in nineteenth-century French philosophy: the tradition of epistemological idealism descending through Charles Renouvier from Kant and Antoine Cournot, and the metaphysical idealism of Maine de Biran, Félix Ravaisson, Jules Lachelier, and Jules Lagneau.

For Brunschvicg, the goal of philosophical reflection was to disclose intellectual activity tending toward self-consciousness as it progressively constitutes knowledge. He therefore frequently characterized history as "the progress of consciousness" (*le progrès de la conscience*). The double meaning of this expression—the progress of conscience as well as of consciousness—also suggests the moral dimension of Brunschvicg's monistic idealism. Viewed subjectively, the process is a conversion from naive acceptance of reality as external to an affirmation of the primacy of the mind as it provides intelligibility. Brunschvicg equated this with recognition of the supremacy of intelligence in a moral sense, which is to say that self-knowledge progresses toward refinement of conscience and moral autonomy. According to Brunschvicg, personal conversion reflects an absolute historical devel-

opment undetermined in form but immanently oriented toward spiritual values (of which Unity is highest) and self-knowledge on the part of humanity as a whole.

The critique of this process, Brunschvicg insisted, cannot depend on a priori assumptions, nor can it hope to specify categories or functions of thought; such analysis would only falsify the mind's essential freedom and inventiveness. The emphasis on creative spontaneity suggests a relationship with Henri Bergson that Brunschvicg was proud to acknowledge, but not to the extent that he wished to embrace Bergson's intuitionism. Although Brunschvicg preferred the general terms *mind* and *intelligence* to *thought* and *reason*, this does not imply a commitment to nonintellectual modes of understanding. At the heart of his work lay studies in the history of science and of mathematics. Brunschvicg regarded scientific progress not only as a triumph of intellect but also as an exemplification of humankind's growing self-understanding. In this way, he defended a moral or "spiritual" conception of science as opposed to positivistic and conventionalistic theories. In his view, the truth of a theory essentially depends on the creative vitality of the mind as it assimilates what is given as nonmental, and as it judges, in turn, the adequacy of this synthesis.

In *La modalité du jugement*, Brunschvicg attempted to delineate the mind's developing accord with being or the real in a theory that classifies judgments according to the forms of "internality" and "externality." Brunschvicg took judgment, rather than the concept or category, as fundamental because he saw it as a synthesizing or unifying act, combining form and content. The form of "externality" was interpreted (evidently following Johann Gottlieb Fichte) as a restraining activity that the mind imposes dialectically on its own creative freedom or "internality."

In *Les étapes de la philosophie mathématique* (Paris, 1912), Brunschvicg examined the highest expression of "internality," mathematical judgment, which he regarded as uniquely appropriate to science because it is at once a free creation—not to be justified through physical interpretation—yet inseparable from experience in its origin and in its "collaborative" task of assimilating being to the understanding. Brunschvicg substantiated this theme in *L'expérience humaine et la causalité physique* (Paris, 1922), which further revealed an implicit dualism and a reluctance to employ categories or principles of analysis, however provisional.

Brunschvicg's last decade was marked by works of a religious nature, following a comprehensive history of philosophy, *Le progrès de la conscience dans la philosophie occidentale* (Paris, 1927), intended to bear witness to humanity's spiritual unification. "Our destiny is to tend toward unity." Religious value apparently attaches to a particular dimension of the "progress of consciousness": The assimilation of being to consciousness insofar as the process is regarded as immanently guided by the value of unity. In this assimilation, humankind moves toward self-identification through the communion of shared intelligence.

Although it appears likely that Brunschvicg felt a moral or spiritual ideal to be predominant in his career, he will perhaps be best remembered as an interpreter of the French philosophical tradition and as a leading spokesman for the life of reason and the value of science.

See also Bergson, Henri; Cournot, Antoine; Descartes, René; Fichte, Johann Gottlieb; Idealism; Kant, Immanuel; Lachelier, Jules; Maine de Biran; Ravaisson-Mollien, Jean Gaspard Félix; Renouvier, Charles Bernard; Spinoza, Benedict (Baruch) de.

Bibliography

ADDITIONAL WORKS BY BRUNSCHVICG

Introduction à la vie de l'esprit. Paris, 1920.

Spinoza et ses contemporains. Paris: Alcan, 1923.

Les âges de l'intelligence. Paris: Alcan, 1934.

La raison et la religion. Paris: Alcan, 1939.

Écrits philosophiques. 3 vols. Paris: Presses Universitaires de France, 1951, 1954, 1958.

WORKS ON BRUNSCHVICG

Boirel, R. *Brunschvicg, sa vie, son oeuvre avec un exposé de sa philosophie*. Paris: Presses Universitaires de France, 1964.

Cochet, M. A. *Commentaire sur la conversion spirituelle dans la philosophie de Léon Brunschvicg*. Brussels, 1937.

Dagognet, F. "Brunschvicg et Bachelard." *Revue de Métaphysique et de Morale* 70 (1965): 43–54.

Deschoux, Marcel. *La philosophie de Léon Brunschvicg*. Paris: Presses Universitaires de France, 1949. Contains a complete bibliography.

Gex, M. "L'idéalisme critique de Léon Brunschvicg." *Revue de Théologie et de Philosophie* 19 (1969): 145–164.

Messaut, J. *La philosophie de Léon Brunschvicg*. Paris: Presses Universitaires de France, 1938.

Bernard Elevitch (1967)
Bibliography updated by Thomas Nenon (2005)

BUBER, MARTIN

(1878–1965)

Martin Buber, the religious existentialist, was born in Vienna and spent his childhood in L'viv, Galicia, at the

home of his grandfather Solomon Buber, a businessman and well-known scholar of rabbinic literature. From 1896 to 1900 he studied philosophy and art history at the universities of Vienna, Leipzig, Berlin, and Zürich. He was early active in the Zionist movement, especially in its cultural and religious aspects, and in 1901 he was appointed editor of the Zionist journal *Die Welt.* Instrumental in the founding of the publishing house Jüdischer Verlag in 1902, in 1916 he founded the German Jewish monthly *Der Jude,* which, until it ceased publication in 1924, was the most respected and literate voice of German Jewry. From 1924 until 1933 Buber was professor of the philosophy of Jewish religion and ethics at Frankfurt-am-Main University, the only chair in Jewish religion at any German university. In 1920 he and Franz Rosenzweig founded the Freies Jüdisches Lehrhaus, an institute for adult Jewish education; and with Adolf Hitler's coming to power Buber devoted his energy to strengthening the religious and spiritual resources of German Jewry in the face of the unprecedented challenge posed to it. Buber continued in the institute until 1938, when he left for Palestine, where he was appointed professor of sociology of religion at the Hebrew University. With Y. L. Magnes he led the Yihud movement, devoted to Arab-Jewish understanding and to the creation of a binational state. In 1952 and 1957 he traveled widely in the United States, lecturing at many universities and to diverse student groups. While his acceptance of various German awards in the postwar period led to criticism from some Jewish quarters, Buber remained steadfast in his encouragement of those German circles that realize the magnitude of the Nazi crimes against the Jews and seem genuinely repentant. He died in Jerusalem.

Buber's basic insight, an insight that runs through all of his work and that determines his approach to everything he touches, is the realization that there is a basic difference between relating to a thing or to an object that I observe, and to a person or a "Thou" that addresses me and to whose address I respond. In its simplest form, this is the difference between the way people usually relate to inanimate things on the one hand and to living persons on the other. Inanimate objects are watched, while persons are spoken to. However, the distinction cannot be drawn simply on this basis. A person as well as an inanimate thing can be viewed as a thing, or, in Buber's terminology, an "It." Whenever we take an "objective" attitude toward a person, whenever we view him as part of the world and caught in its causal chain, we are in an "I–It" relationship, even though the object happens to be a person. The "I–It" relationship is characterized by the fact that it is not a genuine relationship because it does not take place between the I and the It. When another person is an It to me, I am, first of all, perfectly alone. I gaze at him and view him from every possible direction, I observe his place in the scheme of things, and I find elements that he has in common with other persons and things and elements that distinguish him from them. All of this, however, takes place within me; *I* am judging and *I* am observing, and the external world is relevant only to the extent that it enters my being.

It is otherwise in the "I–Thou" relationship. Here the relationship is genuine because it is between me and the Thou that addresses me. This Thou is no longer one thing among other things of the universe; the whole universe is seen in the light of the Thou, and not the Thou in the light of the universe. In fact, it is not only the object in the "I–Thou" relationship that is different from that in the "I–It" situation; the very "I" is different in the two situations. There is no "I" that sometimes relates to a Thou and sometimes to an It. If that were the case, both the It and the Thou would be objects that float into the I's field of vision and then out of it, leaving the I essentially unaffected. Instead, Buber argues, the I of the I–It is a different I from that of the I–Thou because it is not the I as such that has preeminent reality, but the relations I–It and I–Thou. The I appears and is shaped only in the context of some relationship with either an It or a Thou and can never be viewed independently of such a relationship.

Buber further states that the I–It relationship is maintained with only part of ourselves in it. There is always a part of us that remains outside the relationship and views it from some vantage point. In the I–Thou relationship, on the other hand, our whole being must be involved. Should I attempt to hold back any part of myself, I will find myself in an I–It situation because there will be a part of me that is not participant but spectator, a sure sign of the I–It. This means that the I–Thou relationship carries with it much greater risk than the I–It, since there is no withholding of the self possible, as in the I–It. In the I–It situation the part of the self that remains outside the relationship cannot be injured by the other party because he cannot reach it. In the I–Thou relationship there is no such security because the Thou of the I is addressed with the whole of the I, and any response elicited necessarily pertains to this total I. In the I–Thou relationship, therefore, everything possible is risked without any defensive position being left to which the I can withdraw in case of need. However, this is not the only risk involved in the I–Thou situation. The Thou who is addressed cannot be viewed in the context of any causal, deterministic framework. He must be encountered in the

full freedom of his otherness, an otherness that is addressed and that responds in the total unpredictability of human freedom. The moment the responses of the Thou are calculated, the moment the I asks itself what impression its speech and being will make on the Thou, it is relating to an It instead of to a Thou.

Because of this, Buber tells us, there is never a present for the I–It relationship, only a past. This is so because all objective knowledge about a human being is knowledge about his past, of what he has been rather than of what he is. If the present moment is to have genuine novelty, if it is not perfectly determined by the events of the past, then it must be possible for the present to produce a break with the past in the form of a response that could not have been calculated from a knowledge of the past. In the I–Thou relationship we are therefore genuinely living in the present because we are prepared for any and every response to our address, the expected as well as the unexpected—and it is this that constitutes genuine listening. The difference between a pseudo listening and a genuine listening is that while in the pseudo listening situation the listener pretends to listen, what he hears is determined by his past knowledge of the person he is listening to or by his theories concerning the nature of man. Genuine listening does not know ahead of time what it will hear; in the full uniqueness of the present it listens to the speech of the other without filtering what it hears through the screen of its own prejudgments. The purpose of genuine listening is therefore really to hear what the other is saying, constantly being aware that he is saying something that is new and not just a revelation of his nature, which the hearer has already identified and which is fixed as the other's "psychology."

It is in the religious context that the significance of Buber's distinctions emerges most clearly. In contrast to much of mysticism that aims at the obliteration of the abyss between the self and the Absolute in the ecstasy of mystical union, the essence of biblical religion, as conceived by Buber, is the dialogue between man and God in which each is the other's Thou. "The extended lines of relations meet in the eternal Thou," writes Buber in the opening sentence of the final portion of *I and Thou.* Life is an endless transition from the Thou to the It and back to the Thou. Sooner or later, the time comes when even the most cherished Thou recedes, when a spiritual tiredness overtakes the most authentic I–Thou relationship and turns it into the I–It. There is one Thou, argues Buber, who by his very nature cannot become an It. A man may hate God and curse him, he may turn away from him when the suffering of human destiny becomes unbearable; but no man can reduce God to the status of a thing who no longer addresses him and who becomes one object among others in the world for him. Much of traditional theology, for Buber, errs in dealing with God as if he could be turned into an It. Time and again, however, man turns from thinking about God to addressing him, and it is then that he communicates with the living God, as distinct from merely giving intellectual assent to the God of the philosophers. This is true even when the Absolute Thou addressed is not called God. "But when he, too, who abhors the name, and believes himself to be godless, gives his whole being to addressing the Thou of his life as a Thou that cannot be limited by another, he addresses God."

In the course of his long career Buber applied these basic ideas to a diversity of fields. In a number of works devoted to biblical interpretation, he developed in detail his view of the Bible as the record of Israel's dialogue with God. He wrote a definitive work on the relation between Christian and Jewish faith. In this work he distinguishes between the Jewish *emunah* and the Greek *pistis,* the former of which, according to Buber, is faith in the sense of trust while the latter is faith in the sense of belief in the truth of propositions. Jewish faith, as found in the Hebrew Bible, is Israel's trust in the faithfulness of God's word as that word is spoken in dialogue. The faith of the New Testament, particularly in its Pauline version, is heavily influenced by Greek philosophical elements that are reflected in the emphasis on salvation as resulting from belief in the truth of propositions concerning the divinity and resurrection of Jesus. In Paul, Buber thus sees a profound departure from the Hebrew biblical spirit, a departure that is no more than partial and implicit in the Gospels.

In his later years Buber's interest to some extent turned to psychotherapy, in which he emphasized the necessity for the therapist not to hide behind the teachings of his school and not to forget that psychotherapy is above all dialogue in which therapist and patient speak to each other. When seen in this light, the therapist encounters the patient for the individual he is and is ready for the unexpected that the theoretical categories of his discipline do not prepare him for. Similarly, in the field of social philosophy Buber contrasted Marxist socialism, with its centralized control and allegiance to impersonal and inevitable historical forces, with the socialism of the community in which the authenticity of the I–Thou relationship is the foundation on which the living community is built and to which it must return, again and again, for renewal. In the Israeli kibbutz Buber saw an exempli-

fication of the communal or "Utopian" socialism for which he stands.

See also Absolute, The; Existentialism; Jewish Philosophy; Philosophical Anthropology; Rosenzweig, Franz.

Bibliography

Publication of the collected works of Buber in German, *Werke,* was begun in 1962 by Kösel Verlag in Munich. The first three volumes appeared by 1964.

Buber's most important work is *Ich und Du* (Berlin, 1922), translated by R. G. Smith as *I and Thou* (New York: Scribners, 1958). *Die Frage an den Einzelnen* (Berlin: Schocken, 1936), translated by R. G. Smith in *Between Man and Man* (Boston: Beacon, 1955), develops the basic themes in some detail. *Der Glaube der Propheten* (Zürich, 1950), translated from the Hebrew by C. Witton-Davies as *The Prophetic Faith* (New York: Macmillan, 1949), is one of Buber's best biblical studies. *Paths in Utopia,* translated by R. F. C. Hull (London: Routledge, 1949), is Buber's study of social philosophy; *Two Types of Faith,* translated by N. P. Goldhawk (London: Routledge and Paul, 1951) is his study of Judaism and Christianity.

Other writings that have been translated into English are *Eclipse of God; Studies in the Relation between Religion and Philosophy,* translated by Maurice Friedman et al. (New York: Harper, 1952) and *Bilder von Gut und Bose* (Cologne, 1952), translated by R. G. Smith and M. Bullock as *Good and Evil; Two Interpretations* (New York: Scribners, 1953); *Pointing the Way: Collected Essays,* translated and edited by Maurice Friedman (New York: Harper, 1957); and *Martin Buber, Writings,* a selection edited and introduced by Will Herberg (New York: Meridian, 1956).

Maurice Friedman's *Martin Buber: The Life of Dialogue* (Chicago: University of Chicago Press, 1955; New York: Harper, 1960) is a thorough secondary work with an extensive bibliography.

Edwards, Paul. *Buber and Buberism: A Critical Evaluation.* Lawrence: University of Kansas Press, 1970.

Friedman, Maurice. *Martin Buber's Life and Work: The Early Years, 1878–1923.* New York: Dutton, 1981.

Friedman, Maurice. *Martin Buber's Life and Work: The Middle Years, 1923–1945.* Reprint ed. Detroit: Wayne State University Press, 1988.

Friedman, Maurice. *Martin Buber's Life and Work: The Later Years, 1945–1965.* New York: Penguin, 1986.

Moonan, Willard. *Martin Buber and His Critics: An Annotated Bibliography of Writings in English through 1978.* New York: Garland Publishing, 1981.

Schilpp, Paul, and Maurice Friedman, eds. *The Library of Living Philosophers, Vol. 12: The Philosophy of Martin Buber.* La Salle, IL: Open Court, 1967.

Wood, Robert. *Martin Buber's Ontology.* Evanston, IL: Northwestern University Press, 1969.

Michael Wyschogrod (1967)

Bibliography updated by Christian B. Miller (2005)

BUCKLE, HENRY THOMAS

(1821–1862)

Henry Thomas Buckle, the English historian, was the son of a prosperous businessman who left him sufficient money to devote his life to private study and writing. In common with a number of other dominant thinkers of the Victorian age—such as J. S. Mill, Herbert Spencer, and T. H. Huxley—he was largely self-educated. As he was a delicate child, it was thought unwise for him to undertake work involving much intellectual effort or strain, with the consequence that he was (as he put it) "never much tormented with what is called Education, but allowed to pursue my own way undisturbed … whatever I may be supposed to know I taught myself." Thus he was taken from school, at his own request, at the age of fourteen, never went to a university, and conducted his subsequent reading and research (which by any standards were vast) in the absence of all external supervision or direction. Buckle expressed no regret at not having gone to Oxford or Cambridge, considering both universities to be in a contemptible condition and believing himself in any case to be equipped with natural aptitudes and talents that more than compensated for the lack of a rigorous academic training. Certainly his gifts were far from negligible. He had an excellent memory, he could express himself both in writing and in conversation with great fluency and eloquence, and he was a first-class linguist (by the age of thirty he could read eighteen foreign languages and speak six); he possessed, moreover, an immense capacity for methodical work, together with an intense intellectual curiosity and a meticulous eye for detail.

Buckle led a comparatively uneventful life, his energies being to a large extent absorbed by the ambitious project of writing a history of civilization, to which, from his early twenties, he had decided to dedicate his career. But though the preparation of this enormous enterprise always remained his chief concern, he was not without other interests. He was, for example, a brilliant chess player, achieving an international reputation; he traveled widely, in Europe and beyond; and by the end of his life he had established a wide circle of acquaintances, including William Makepeace Thackeray, Charles Kingsley, Charles Darwin, and John Stuart Mill. For Mill in particular he had the highest admiration, and in 1859 he wrote a long review in *Fraser's Magazine* praising Mill's essay "On Liberty"—a review that created considerable stir at the time, since in it Buckle drew public attention to the fantastic sentence of twenty-one months' imprisonment recently passed upon a man for inscribing on a gate

words offensive to Christianity. Although Buckle never married, he liked feminine society and secretly kept a mistress; when, after his death, the truth ultimately leaked out, it caused consternation and dismay among some of his close friends and relatives.

SIGNIFICANCE OF THE *HISTORY*

Buckle died at the age of forty while touring the Middle East. Only two volumes of his *History of Civilisation in England* had appeared, and these represented no more than an introduction to the vast work he had envisaged writing. Yet they had been sufficient to achieve for their author sensational fame, not merely in his own country but also throughout Europe and in the United States; Darwin applauded the work's brilliance and originality; and an influential American writer, Theodore Parker, attributed to it an importance in the history of thought comparable to that of Francis Bacon's *Novum Organum.* Buckle's reputation has since suffered a heavy decline, and many of the claims made on behalf of his work at the time of its publication seem grotesquely exaggerated today. Even so, what he wrote represents (as Henry Sidgwick pointed out) the first major attempt on the part of a thinker versed in the tradition of British empiricism and inductivism to enter the treacherous field of historical speculation, and to offer a comprehensive and detailed theory of historical development of the type that previously only Continental philosophers had ventured to provide. For this reason alone it preserves a certain interest and is still worth studying.

BUCKLE'S INTENTIONS

Buckle was fully aware of what had been done by some of his predecessors in Germany and France; and references to their works, particularly to those of Johann Gottfried Herder and Auguste Comte, are to be found scattered among the footnotes that abound throughout his own volumes. Like Herder, he was eager to connect the facts of human history with the conditions imposed by different forms of natural and geographical environment; like Comte, he wished to present the course of history as exemplifying a fundamental pattern of progress and improvement. But he rejected the tendency to revere past ages, and to exalt imagination at the expense of rational and scientific modes of thinking, that often manifested itself in Herder's writings; and he equally distrusted the strain of aprioristic dogmatism and respect for authoritarian methods of social control that he detected in Comte's historical system, calling the latter's theory of government "monstrously and obviously impracticable." Buckle's allegiance lay chiefly with the ideals set out by English radicals and Utilitarians early in the nineteenth century, and it was these that finally determined the valuations embodied in his conception of social and historical progress.

HUMAN ACTIONS SUBJECT TO LAWS

Early in his book Buckle raised the question, "Are the actions of men, and therefore of societies, governed by fixed laws, or are they the result either of chance or of supernatural interference?" He supposed these possibilities to represent exhaustive alternatives, and argued that either variety of the latter hypothesis was plainly unacceptable.

So far as the theory of supernatural interference was concerned, this, together with the associated theological doctrine of predestination, must remain a "barren hypothesis," since no conceivable experience could count for or against its truth. On the other hand, the view that what occurs in the realm of human affairs is the product of chance was demonstrably false; it had, however, been given an aura of spurious respectability by metaphysical philosophers who had carried the principle in question over into the sphere of individual human psychology. There it emerged as the famous doctrine of free will, according to which a mysterious, undetermined power of free choice is held to be directly vouched for by the evidence of the introspective consciousness. But in Buckle's opinion it is precisely such blind reliance upon the findings of individual introspection that has been the besetting sin of "metaphysicians," leading them to construct their impressive-looking, though nonetheless mutually incompatible, systems in accordance with a radically mistaken procedure.

By contrast, in order to achieve a realistic conception of the nature and workings of the human psyche it is necessary to adopt an external and general view of human behavior analogous to that taken by natural scientists in the investigation of nonhuman phenomena: From this altered standpoint it can indeed be seen that the actions of men are subject to regularities as strict and mathematically exact as those that operate in other spheres of scientific inquiry. As a conclusive demonstration of his thesis, Buckle cited the evidence afforded by large-scale statistical surveys concerning the numbers of marriages contracted, and of murders and suicides committed, in particular countries and towns during successive years; the relative uniformity of the results obtained would, he held, be unintelligible on any other assumption than that

there are certain social laws capable of keeping the level constant.

When discussing this topic, Buckle on occasions fell into confusions; he did not, for example, always distinguish between the necessary and the sufficient conditions of an occurrence, and was prone to disregard the difference between causal laws and statistical frequencies. In consequence he sometimes interpreted the statistical data in a misleading way, suggesting that the sole effective determinant of individual actions was what he called "the general condition of society." He also spoke as if the mere existence of a proportional average, observed to hold over a period of time, necessitated, with a kind of irresistible momentum, the commission of a particular number of crimes in any given year. As a result, a picture is presented wherein human beings appear as the helpless victims of social forces over which they can exert no effective influence or control—a conclusion in no way entailed by the premises from which Buckle initially proceeded.

ORIGIN AND DEVELOPMENT OF CIVILIZATION

Be this as it may, it is noticeable that when Buckle approached his principal theme—the genesis and development of civilization—he made little further reference to precise numerical regularities or frequencies; although he still spoke of "laws," it was the broad, indeterminate, and sometimes very doubtful generalizations concerning the factors influencing the evolution of human societies that he chiefly appealed to in providing his explanations. Thus, the fundamental agents of social growth were deemed to be material or, to use his term, "physical," and were listed as being "Climate, Food, Soil, and the General Aspect of Nature." These—and not, as some previous theorists had alleged, innate racial characteristics or mysterious "national spirits"—originally determine the divergent forms of organization and progress achieved by different historical cultures.

FOOD SUPPLY AND CIVILIZATION. Buckle believed that the degree of civilization attained by a society depended upon its wealth and upon the manner in which this wealth was distributed; such factors were in turn dependent upon the population of the country concerned, and the size of the population was determined by its food supply. In countries where cheap food was plentiful, the population increased in a fashion that led to the labor market's becoming overstocked; as a consequence there was unemployment and also poverty, since there is an inevitable tendency in societies where there is a surplus of labor for laborers to be underpaid and for immense economic inequalities to develop. He cited such examples as Egypt, Peru, Mexico, and India, where riches were concentrated in very few hands and where the vast majority of the inhabitants lived in a miserable and depressed condition: "Among nations subjected to these conditions, the people have counted for nothing, they have had no voice in the management of the state, no control over the wealth their own industry created."

EUROPEAN CONDITIONS IDEAL. Buckle, in fact, considered that the ideal conditions for the development of civilization were to be found in Europe. Here the food supply was not so abundant as to lead to overpopulation, nor was it so scanty as to make the accumulation of wealth and the enjoyment of leisure (on which intellectual progress depends) impossible. Here, also, the temperate climate was favorable to enterprise and the energetic exploitation of natural resources; moreover, the aspect that nature presented to human beings was of a less extreme and unpredictable character than in other parts of the world. Thus, men did not regard it with superstitious awe as a terrifying and insuperable power, but saw it instead as something that obeys regular laws and is therefore capable of being tamed and utilized for their purposes. It followed (he thought) that Europe could be distinguished from all other centers of human society by the circumstance that it was human rather than natural or physical factors that had determined the course taken by its history and progress. Man was here the master of nature, and consequently the key to the development of European culture lay in the influence exercised by "the laws of the human mind."

KNOWLEDGE DETERMINED DIRECTION OF CULTURE

It might be expected that Buckle would go on to state what these laws of the human mind were, using them to explain patterns of social change in European history in a fashion comparable to that suggested by Mill in Book VI of his *System of Logic* when he spoke of the possibility of deriving principles governing historical development from the "ultimate" laws of human psychology. Buckle can scarcely be said, however, to have adopted this procedure, perhaps because he believed that the psychological and historical data available at his time were insufficient to make it practicable. Instead, he contented himself mainly with trying to show that it was the advance and diffusion of knowledge, and particularly of scientific knowledge, that had in the last analysis given European

history its characteristic overall direction—"the progress Europe has made from barbarism to civilization is entirely due to its intellectual activity."

Other factors were considered, but only to be ruled out. Thus Buckle claimed—as if (rather surprisingly) it were a self-evident truth—that men's moral opinions had remained essentially unaltered for thousands of years: How then could these have been responsible for the far-reaching transformations that had overtaken European nations like England and France in the course of their historical evolution? Likewise, he rejected the claims of religion, literature, and government to be "prime movers of human affairs." Acceptance of a particular religious creed is a symptom rather than a cause of the condition in which a given society finds itself. The literature of a country merely reflects and serves to fix the degree of civilization already attained; it does not initiate further achievement. So far as the influence of government is concerned, Buckle maintained that the rulers of a nation were only "creatures of the age, never its creators." Enlightened legislation occurs only as a consequence of the pressure exerted by changes in the climate of opinion, these being due in the first instance to the efforts of "bold and able thinkers" who belong to the intellectual, and not the governing, classes; nor will such legislation be effective unless the ground has been prepared for it and "the age is ripe."

POLITICAL THOUGHT

Writing very much as an exponent of the principles of laissez-faire radicalism, Buckle displayed an intense distrust of governmental interference and "protectionism," which tended to be identified in his mind with the suppression of free opinion and free trade. Accordingly, he argued that most beneficial legislation is negative in character, taking the form of repealing the bad enactments passed by earlier generations; and, generally speaking, he restricted the legitimate functions of government to such things as the maintenance of order and the preservation of public health. The moral drawn is that the ineluctable laws of historical development should be permitted to take their course freely and without impediment; unlike many other philosophers of history, Buckle did not try to combine a doctrine of historical inevitability with a comprehensive positive program of political action and social reconstruction.

BUCKLE'S SIGNIFICANCE

There is much that is intellectually naive in Buckle's theory of history, and it is easy to find inconsistencies and non sequiturs among his arguments; Leslie Stephen's gibe that Buckle's "mental fibre was always rather soft" is not wholly beside the mark. His uncritical use of vague abstractions like "intellectual progress" and the "spirit of a time" often led him into treating vacuous truisms as significant discoveries, and the collectivist conception of historical change that pervades much of his work contrasts oddly with the influence he ascribes to individual scientists and economists in promoting social advance. Nevertheless, the impact of his ideas upon his age was undeniably great, and his criticisms of previous and current historiography were not without important long-term effects. Like Karl Marx, though with far less insight and imagination, he helped to turn the eyes of historians away from the political surface of events, making them look more closely at the technological and economic realities of human life that lie beneath; at the same time, through his determinism, he provided a corrective to the tendency toward excessive moralizing that his contemporaries exhibited in their treatment of the past. And, by enlarging the perspective of historical study to include cultures and societies remote in time or space from his own, he made a definite, if limited, contribution to widening the horizons and counteracting the provincialism of future students of human affairs.

See also Bacon, Francis; Comte, Auguste; Darwin, Charles Robert; Herder, Johann Gottfried; Huxley, Thomas Henry; Marx, Karl; Mill, John Stuart; Parker, Theodore; Sidgwick, Henry; Spencer, Herbert; Utilitarianism.

Bibliography

PRIMARY SOURCES

Buckle, H. T. *The History of Civilisation in England,* 2 vols. London, 1857–1861.

Taylor, Helen, ed. *Miscellaneous and Posthumous Works of Henry Thomas Buckle,* 3 vols. London: Longmans, Green, 1872.

STUDIES AND COMMENTARIES

Huth, Alfred Henry. *Life and Writings of Henry Thomas Buckle,* 2 vols. London: S. Low, Marston, Searle, and Rivington, 1880.

Robertson, John M. *Buckle and His Critics.* London: n.p., 1897.

St. Aubyn, Giles. *A Victorian Eminence.* London: Barrie, 1958.

Stephen, Leslie. *The English Utilitarians,* Vol. III. London: Duckworth, 1900.

Patrick Gardiner (1967)

BUDDE, JOHANN FRANZ
(1667–1729)

The German philosopher, theologian, and historian Johann Franz Budde, or Buddeus, was born in Anklam, Pomerania. He entered the University of Wittenberg in 1685 and became an assistant there in 1689. Budde was appointed professor of moral philosophy at Halle in 1693, full professor of theology at Jena in 1705, and church councilor at Gotha in 1715. Although he insisted on his independence from all schools and considered himself an eclectic, he was close to Pietist thought and to the philosophy of Christian Thomasius, his colleague at Halle.

Budde's most significant work in theoretical philosophy was his *Institutiones Philosophiae Eclecticae* (2 Teile, Halle, 1703). In the first section, in which he expounded his logical doctrines and the intent was chiefly methodological, the influences of John Locke and Thomasius are apparent. Budde derived error from original sin and prescribed means for restoring the "good health" of the mind. He regarded ontology as a part of logic and as consisting in a simple explanation of basic metaphysical terms. According to Budde, these terms had a purely instrumental value because he refused to confer upon metaphysics the rank of independent and universal science. Rather, he interpreted it as the science of the most general nouns used in theology and philosophy.

In the second section of the *Institutiones,* Budde first discussed natural philosophy in a phenomenalistic manner, holding that we cannot know the real nature of things, but only their appearances and effects. He attempted to reconcile the physical animism or spiritualism typical of Pietist natural philosophy with mechanism. He frequently appealed to the Bible and gave an important place to final causes. At the end of this section he discussed spirits and God, whose existence he demonstrated by rational proofs.

In practical philosophy (*Elementae Philosophiae Practicae,* Halle, 1697) Budde followed Hugo Grotius, Samuel von Pufendorf, and Thomasius. He completely denied human freedom, referring the possibility of good actions to God's grace and restricting accountability to a narrow and extrinsic sphere of material liberty. He devoted much space to discussions of practical psychology and prudence, for he believed that such practical psychology was a better means than abstract instruction of healing the human will from sin. However, revelation is essential to this healing process.

As with the Pietists, practical philosophy is central to Budde's thought. He also agreed with the Pietists in stressing the will's independence of the intellect, in his emphasis on psychology in practical philosophy and on spiritualism in cosmology, and in the importance he placed on revelation. However, Budde was much more systematic than Thomasius, who was likewise very much influenced by Pietism. Budde joined the Pietists in their fight against Christian Wolff, and in 1723 he wrote a pamphlet attacking Wolff.

Although in practical philosophy Budde agreed with the Pietists, in theology he tried to reconcile the views of orthodoxy and Pietism. Because he held that man has an original religious impulse, he gave an important position to natural religion. He presented cosmological, physicotheological, and historical proofs of God's existence, and tried to refute atheism by argument.

Budde was one of the most learned men of his time. His writings on the history of Jewish philosophy (*Introductio ad Philosophiam Ebraeorum,* Halle, 1707), on general history of philosophy, and on the history of theology (*Historia Theologiae Dogmaticae et Moralis,* Frankfurt, 1725) were excellent in their time and are still valuable for the information they contain.

See also Determinism, A Historical Survey; Grotius, Hugo; Jewish Philosophy; Locke, John; Pietism; Pufendorf, Samuel von; Revelation; Thomasius, Christian; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY BUDDE

Selecta luris Naturae et Gentium. Halle, 1704.

Institutiones Theologiae Moralis. Leipzig, 1711.

Institutiones Theologiae Dogmaticae. Leipzig, 1723.

WORKS ON BUDDE

Hirsch, E. *Geschichte der neuren evangelischen Theologie.* Vol. II, 319–340. Gütersloh, 1960.

Stolzenburg, A. F. *Die Theologie des Joh. Fr. Buddeus und des Chr. Pfaff.* Berlin, 1926.

Wundt, Max. *Die deutsche Schulphilosophie im Zeitalter der Aufklärung,* 63–75, 242–243. Tübingen: Mohr, 1945.

Giorgio Tonelli (1967)

BUDDHISM

Buddhism derives its name from the Sanskrit word *buddha* (awakened, wise, or learned), which was one of the many epithets given to Siddhārtha Gautama (c. 563–c.

483 BCE), the founder of the set of theories and practices that are now called Buddhism. Traditional accounts of Gautama's life are more inspirational and hagiographical than historical in nature, and any attempt to extract a historical record from them is likely to prove frustrating, although the attempts of such authors as Hans Wolfgang Schumann (1989) and Michael Carrithers (1983) to find a credible story of Gautama's life are well worth reading.

According to traditional accounts Gautama left his wife and newborn child to seek his liberation from suffering and followed various teachers who ultimately failed to satisfy his needs. He then set out on his own and found the liberation he sought through meditation and self-discipline. At first disinclined to teach, because he felt his teachings would appeal to few people, he finally decided to tell others what he had discovered. Soon after his death, his disciples met and repeated all they could remember being taught by him, and these recollections were committed to memory. All the rules he had set down for the community of his disciples were collectively known as the *vinaya*. The collections of his other teachings on good character, contemplative exercises, and the theory behind them were known collectively as *sutras*. The vinaya and sutras supposedly collected shortly after Gautama's death became a closed canon for some Buddhists; other Buddhists eventually accepted as canonical a large corpus of other literature. Although there is a great deal of agreement between what is found in both the closed and extended canon, there is also a good deal of difference. In what follows, an attempt will be made to make note of where there is agreement and where there is divergence of opinion among Buddhists.

The epithet Buddha emphasizes Gautama's claim to have awakened, as if from a slumber, to seeing things as they really are. Another epithet commonly given to Gautama is *jina* (conqueror), which emphasizes his having overcome his internal enemies, the passions. In much of the Buddhist canonical literature Gautama refers to himself as *Tathāgata*, an epithet that has been explained in various ways by later Buddhists; one possible interpretation is that the Tathāgata knew the truth or understood things as they really are. Traditionally being a follower of Buddhism consists in going for refuge to the Buddha, the *dharma* (the goal of Buddhist practice), and the *sangha* (the community of virtuous people). In what follows, each of these terms will be discussed with reference to how understanding of them has changed down through the centuries.

THE BUDDHA

During the time when the Buddha Gautama was alive, going for refuge to him meant becoming his disciple and agreeing to follow his teachings and the rules of his community. After his death, however, the meaning of going for refuge to the Buddha changed. The action came to mean making an effort to cultivate in oneself the virtues associated with buddhas in general, for the claim attributed to Gautama was that he was the most recent in a series of buddhas, all of whom had taught the same thing to the people of their generation and all of whom had had the same set of virtues. The set of virtues associated with buddhas are called the factors of awakening. Canonical texts always talk about thirty-seven mental factors that are required for awakening. These factors are the sum of seven different lists of wholesome mental qualities. When all redundant terms are eliminated from these lists, however, there are just ten different factors: wisdom, courage, concentration, mindfulness, inner joy, mental and emotional flexibility, equanimity, faith, right resolve, and good moral habit.

Wisdom is explained as understanding and discrimination, and it includes awareness of one's own body and mind, reflections on the inevitability of death, and recognition that all complex beings change and therefore are not worth striving for. Wisdom also entails realizing that no one is fully in control of one's own body, mind, or personality and that therefore these things are not really one's self, and none of them really belongs to anyone; rather, everything that comes into being is an essentially impersonal event. Because the factors conditioning any one event are beyond reckoning, no one can be in control of all of them; since it is possible to alter some of the conditions in one's life, however, discipline and practice are not in vain, however difficult they may be.

Courage consists in having the resolve and energy to do virtuous and wholesome actions that benefit oneself and others. Concentration is defined as having a healthy mentality focused on a single topic at a time. Mindfulness is defined as good memory, and especially recalling the importance of virtue in all situations and remembering to cultivate it. Inner joy is described as zest and enthusiasm for being virtuous and helping others do the same. Flexibility is defined as workability and pliability of one's thoughts and emotions, which are the opposite of intellectual and emotional rigidity, obsessiveness, and the tendency to pass judgment on others. Equanimity means indifference to pleasure and pain, and impartiality with respect to people. Faith is described as conviction and trust in the teachings of the Buddha as a result of experi-

encing the initial benefits of practicing what he taught. Right resolve consists in the resolve to cultivate wholesome and to eliminate unwholesome mental states. Good moral habit includes thinking, speaking, and acting in ways that conduce to the well-being of oneself and others, and it manifests in earning one's livelihood in ways that minimize damage done to other living beings and to their environment.

Even though it was considered possible for a person to cultivate all these virtues while living an ordinary family life, it was said to be much easier to succeed if one first renounced family life and lived alone or in a community of like-minded friends. For this reason, the ideal setting for Buddhist practice has nearly always been seen to be a monastery.

For the first several centuries in the history of Buddhism, the Buddha was venerated as a man who had been born an ordinary man and who had struggled to discover and eliminate the root causes of rebirth and its inevitable difficulties. After a long life of teaching others how to eliminate their own causes of rebirth, he died a serene death, knowing that he had helped many others to become awakened and liberated from their suffering. He likened himself to a physician who had studied the symptoms of a disease, made a diagnosis, and prescribed a course of treatment. Like a physician, he could only encourage his patients to take the necessary course of treatment; he could not do their work for them. After some five or six centuries, however, this description of the Buddha's career lost its appeal to many people, and new movements evolved within Buddhism that portrayed buddhas in importantly different ways.

One of the most influential of these new portrayals of what a buddha is appeared in a Mahayana Buddhist text, probably written in the second or third century CE, known as the Sutra of the White Lotus of the True Doctrine, commonly referred to simply as the Lotus Sutra. This complex and highly polemical text repudiates the earlier Buddhist doctrine that the Buddha was born, lived, and died, never again to be reborn in any form anywhere. The Lotus Sutra puts forth the view that all particular buddhas, including Gautama, are but manifestations of an eternal entity known as Shakyamuni Buddha, who is omniscient, omnipotent, and perfectly compassionate but otherwise beyond human comprehension. Shakyamuni Buddha, being transcendent, can be known to human beings only by taking human form. All the buddhas of the past, present, and future should be known to be manifestations of this cosmic buddha.

Moreover, the most important teaching of all these manifestations of Shakyamuni Buddha is that every sentient being in the universe is destined to become a fully enlightened buddha, for all beings, and not just those who are known now to be buddhas, are essentially one with the enlightened mind called Shakyamuni Buddha. Announcing this message in various ways, the Lotus Sutra pronounces harsh condemnation of those who teach that the goal of Buddhism is to attain nirvana, if that is understood as the end of the cycle of rebirths. Monks who teach that Buddha Gautama was an ordinary human being who achieved extraordinary things and that he eventually died never to be reborn, are denounced in the Lotus Sutra as charlatans and pseudo-Buddhists whose teachings could prevent others from attaining perfect enlightenment. The immediate destiny of such monks is a long and painful stay in hell, but even they, assures the Lotus Sutra, will eventually realize full and perfect enlightenment.

A second sort of new portrayal of a buddha figure is found in a genre of literature that has come to be known as Pure Land sutras. The term *pure land* is a translation of a Chinese expression that is in turn an interpretation of a Sanskrit expression that means "a happy land" or "a land of ease." The principal innovation in this genre of sutras is the notion that there are buddhas who attained buddhahood only after amassing an incalculable amount of merit through austerities and good works. After attaining buddhahood these buddhas used their merit to create realms in which all the distractions posed by hardships are unknown so that inhabitants of these realms could devote all their energy to cultivating virtue and striving for nirvana. People from our burdensome world are said to be able to gain entry in one of these realms of ease simply by calling on the name of the buddha who created it.

By far the most popular of these buddhas was Amitābha, whose name means "he whose light is unmeasured." The invocation of Amitābha's name became one of the most common practices among Buddhists in East Asia. In some places, and especially in Japan, some followers of the Lotus Sutra held Amitābha (Amida in Japanese) practitioners in contempt because of the latter's reluctance to regard Amitābha as a manifestation of Shakyamuni. The various views that Buddhists have held on what exactly the nature of a buddha is have been described in detail and with considerable philosophical refinement by Paul J. Griffiths in *On Being Buddha: The Classical Doctrine of Buddhahood* (1994).

THE DHARMA

Those who followed the closed canon of Buddhist texts teach that the dharma to which a Buddhist goes for refuge is nirvana, which is seen as the ultimate goal of all Buddhist practice and theory. Ultimately, nirvana is defined as the cessation of rebirth after one's present life comes to an end, but the term also refers to the cessation of psychological afflictions while one is still alive. The principal afflictions discussed in Buddhist teachings are greed, hatred, and delusion. Greed is understood broadly as all craving for material possessions, physical and psychological comforts, physical and psychological pleasures, celebrity, approval, and anything that one regards as desirable. Hatred includes irritation, resistance, anger, and any sort of aversion or wish for dissociation from something. Delusion includes any sort of misunderstanding or misjudgment that could result in unsuccessful action.

These three root afflictions are said to be the principal causes of all distress and discontent. Eliminating them results in being content with whatever situation that may present itself. In many Buddhist texts it is said that contentment arises not merely from the absence of afflictions but from the presence of their opposites. Thus, when greed is replaced by generosity, hatred by love, and delusion by wisdom, then one is truly contented, and when these replacements are permanent, then one has attained liberation from suffering in this life.

While the dharma to which a Buddhist goes for refuge is nirvana, the term *dharma* also refers to virtue in general and to anything, such as teachings and practices, that help one to cultivate virtue. The most important of the virtues is wisdom, since it plays a role in the cultivation of all other virtues. Wisdom is said to arise in three stages. The first stage consists in learning what wise people have said and how they have acted. The second consists in reflecting on what one has learned through study. And the third consists in realizing in one's own life what the wise people of the past have discussed. This third stage includes a variety of contemplative exercises that have been designed to improve a person's mentality. For each of the virtues discussed earlier as those associated with buddhas, specific meditative exercises have been designed.

In canonical Buddhism the attainment of nirvana is usually described as incremental. The analogy most frequently used is that one's mentality is like gold ore, which is a mixture of precious metal and various unwanted minerals. Refining ore to get pure gold requires a gradual elimination of the unwanted minerals through various chemical and mechanical processes. Similarly, one's mentality is a mixture of wholesome and unwholesome habits that mute the effectiveness of the wholesome traits. Refining one's character requires the gradual elimination of bad habits through study, reflection, and cultivation, and the culmination of all this refinement is nirvana. In some forms of later Buddhism, however, a different conception of nirvana arose. In this new view nirvana, understood not as the mere absence of affliction but as the constant presence of tranquillity, lucidity, and bliss, is the fundamental nature of all things. Thus, all consciousness is fundamentally calm, lucid, and contented, and the so-called afflictions are temporary obscurations of that lucidity. The most common analogy for this view of nirvana is that of the sun, which shines all the time but is sometimes temporarily obscured by clouds. In this view of consciousness the condition of enlightenment is innate and permanent.

Nirvana, therefore, is not the cessation of existence but the realization that consciousness is beginningless and endless and constantly tranquil. In some forms of this doctrine it is said that ultimately there is only one mind, namely, the Buddha's; all apparently individual minds are but episodes of this one Buddha mind. Since the Buddha's mind can only be wholesome, it follows that all those who are apparently individuals are also wholesome, and all mental events, including those called unwholesome or vicious are in fact virtuous. Delusion, then consists in a failure to recognize the innate wholesomeness of all existence. In some formulations of this position delusion consists in thinking in terms of oppositions at all; thus, it is delusional to think in terms of the contrast between virtue and vice, wholesomeness and unwholesomeness, delusion and wisdom, liberation and bondage, buddha and ordinary person, and so forth.

THE SANGHA

The word *sangha* means "community." The community to which a Buddhist goes for refuge is the so-called noble (*ārya-sangha*) community, which comprises all those who have reached one of the four stages leading to and including nirvana. Since it is seen as nearly impossible for an individual to make the necessary refinements in character that lead to nirvana, it is considered almost essential for one to keep company with virtuous people who will understand and support one's resolve to cultivate virtue and attain nirvana. In the hopes of providing a community of people dedicated to virtue and thus providing a noble sangha, the Buddha Gautama founded a monastic community as well as a community of lay disciples. Ide-

ally, these formal communities should include enough members of the noble community to be of benefit to the world as a whole, and so to help these visible communities not only to be virtuous but also to be seen to be virtuous, Gautama set forth various sets of precepts. Lay disciples are expected to refrain from five harmful activities: killing, stealing, sexual transgressions, lying, and intoxication. Novitiates seeking ordination into the monastic community are expected to refrain from ten harmful actions and thoughts: the first four of the five expected of the laity plus refraining from harsh speech, gossip, frivolous speech, attachment, anger, and false views. Monastics are expected to observe more than two hundred vows, depending on which monastic order they belong to. Four of those vows are considered so important that any person who breaks them is dismissed from the monastic order for the rest of his or her life; these four vows are refraining from killing a human being, from the theft of anything that human beings regard as property, from any kind of sexual intercourse with any other being living or dead, and from making false claims about one's spiritual attainments.

The study of the monastic rules (vinaya) suggests that the principal purposes of the monastic community were twofold: to provide an ideal environment for individuals to cultivate virtue and to serve as a visible community that demonstrated to the society at large that a life of material simplicity dedicated to the cultivation of virtue and self-contentment is far more satisfactory than a life of material acquisitiveness dedicated to seeking possessions and the approval of others. Taking monastic vows is not seen as necessary for the attainment of nirvana, but is seen rather as the taking on of responsibilities to be of service to society at large. Some scholastics favor the view, based on passages in canonical texts, that, while it is not necessary to be a monastic to attain nirvana, it is impossible for anyone who has attained nirvana to remain a lay person for more than one day. Others, however, take the view that renunciation is itself a kind of attachment and that a liberated person would be able to live a normal lay life without becoming either attached to or afraid of its pleasures. This latter attitude can be found in many Buddhist movements within East Asia, and especially Japan, and in some movements in Tibetan Buddhism.

The Buddha Gautama made several observations about statecraft. He made these observations by telling stories, which often had a satirical edge. One attitude that emerges in these stories is that government came into human society at a time when morality was breaking down, and, since government was devised by people living in morally broken down cultures, government is itself as likely to exacerbate the problem as alleviate it. That notwithstanding, those whose task is to provide governance can sometimes benefit by the counsel of wise people, although not all governments are equally willing to heed wise counsel.

In his own instructions to kings, Gautama urged them, above all else, to provide to all citizens the means to earn their own livelihoods. This could best be achieved by taxing the wealthy and distributing resources to the needy and by educating the unskilled. A king who fails to do these things, said Gautama, is most likely to bring about a society in which the poor have no means of living other than stealing from the wealthy, and the wealthy then hire guardians to protect their wealth. This situation in turn leads to both the thieves and the mercenary guardians arming themselves to protect themselves against one another, and it leads to the wealthy seeking ever stricter laws and more severe punishments, until nearly everyone is armed and afraid. As fear and suspicion grows, violence increases, and as violence increases, the life expectancy of people declines. Eventually, said Gautama, the decline will become so dramatic that most people will die only shortly after reaching the age of reproduction, and children will be left to raise themselves, and morality will become so rare that people will have forgotten even the word *virtue*, let alone know what it stands for. All this can be avoided by governments that are more interested in protecting the poor than in serving the wealthy, said Gautama, and such governments are more likely to occur if wise and learned men and women remain actively engaged in society. Even monks who have renounced the family life should take an interest in providing wise counsel to governments. The ideal of providing selfless service to one's society was particularly emphasized in some of the Mahayana sutras that came into prominence in the first several centuries CE.

The entire philosophy of Buddhism is traditionally summarized in a formula called the four noble truths: (1) all forms of existence involve some suffering, (2) suffering arises because of unrealistic expectations, (3) suffering can be eliminated by eliminating unrealistic expectations, and (4) there is a method to be followed to eliminate them. The method itself is summarized in the formula: "Do what is beneficial, avoid doing harm, and keep the mind pure."

See also Buddhism—Schools: Chan and Zen; Buddhism—Schools: Dge-lugs; Buddhism—Schools: Hua

yan; Buddhism—Schools: Madhyamika; Buddhism—Schools: Yogacara; Buddhist Epistemology.

Bibliography

Carrithers, Michael. *The Buddha*. New York: Oxford University Press, 1983.

Conze, Edward, tr. *The Large Sutra on Perfect Wisdom, with the Divisions of the Abhisamayālankāra*. Berkeley: University of California Press, 1975.

Griffiths, Paul J. *On Being Buddha: The Classical Doctrine of Buddhahood*. Albany: SUNY Press, 1994.

Harvey, Peter. *An Introduction to Buddhism: Teachings, History, and Practices*. New York: Cambridge University Press, 1990.

Keown, Damien. *The Nature of Buddhist Ethics*. New York: St. Martin's Press, 1992.

Lopez, Donald S., ed. *Buddhist Scriptures*. London: Penguin, 2004.

Lopez, Donald S. *The Story of Buddhism: A Concise Guide to Its History and Teachings*. San Francisco: HarperSanFrancisco, 2001.

Schumann, Hans Wolfgang. *The Historical Buddha: The Times, Life, and Teachings of the Founder of Buddhism*. London: Arkana, 1989.

Sizemore, Russell F., and Donald K. Swearer, eds. *Ethics, Wealth, and Salvation: A Study in Buddhist Social Ethics*. Columbia: University of South Carolina Press, 1990.

Swearer, Donald K, ed. *Me and Mine: Selected Essays of Bhikkhu Buddhadāsa*. Albany: SUNY Press, 1989.

Richard P. Hayes (2005)

BUDDHISM—SCHOOLS

This composite entry is composed of the following sub-entries:

CHAN AND ZEN
DGE-LUGS
HUA YAN
MADHYMAKA
YOGĀCĀRA

CHAN AND ZEN

Zen is the latest Japanese development in a number of similar Buddhist traditions known as "Chan" in China, "Seon" in Korea, and "Thiên" in Vietnam, all supposedly having origins in India. It is an open question whether there is a sufficient degree of homogeneity to label this multifarious tradition "Chan" or "Zen." A safe alternative would be to treat each of the regional variations as distinct traditions, or even to handle the numerous subcategories in each of the East Asian regions as not necessarily connected with each other, at least not in the sense of a historical continuity.

One factor that makes these traditions especially complex is their emergence at various times and in various settings without being submitted to a central religious authority that would have defined their identity, their doctrine, and their structure as a religious unity. The multifarious nature of these traditions does not mean that Zen institutions did not participate in games of power; they certainly did. The vicissitudes of these lineages result from influences that cannot be reduced to institutional fates and orientations, or to their connections with political contingencies. This is because in most cases their self-proclaimed criterion for religious authority was spiritual realization.

ZEN AGENDAS

While there is no unified tradition, this presentation uses the word "Zen" to indicate the fuzzy field comprising all the traditions mentioned above. For the sake of simplicity it is convenient to adopt the widely used Japanese pronunciation, except when referring to a specific geographical area.

Since many Zen lineages and most Buddhist schools seek to disentangle the nexus of our projections even on sacred matters, awareness of our own hermeneutic circle is a necessary prerequisite for examining the possible confluence between traditional and philosophical approaches to Zen. One of the sources that have shaped the understanding of Zen is the agendas of those who first introduced it to Western audiences and readers. Fortunately, a growing array of sharp studies is now available to facilitate the deconstruction and subsequent understanding of how missionaries, apologists, and romantics contributed to fabricating a contemporary notion of Zen. These studies examine why, for instance, Daisetsu Suzuki (1870–1966) in his own time and context chose to present Zen as the finest product of "Japanese spirituality," and even to claim, "As I conceive it, Zen is the ultimate fact of all philosophy and religion" (1961, p. 268). Works by Faure, McRae, and Wright provide an insightful analysis of this crucial dimension and some of the necessary antidotes. One of their achievements is to reveal contradictions inherent in the discourse of apologists who denied their own historicity.

THE CONCEPT OF MEDITATION

"Zen," pronounced "Chan" in Chinese, has an interesting linguistic background. The Chinese compound "channa"

was used to phonetically transcribe the Pali terms "jhāna" or "jhān," from which it derives. "Dhyāna" being the Sanskrit equivalent for the Pali "jhāna," popularizers often simplify this etymology by explaining "chan" as if it derived from "dhyāna." Eventually "chan," the first half of the compound, became a word in itself, retaining some of its original implications.

Indian Buddhists chiefly used the word "jhāna" as a generic term for meditation (singular) and as a technical term for particular meditative states (plural). For example, in the *Sutta-nipāta*, "jhāna" is always singular, and appears in contexts such as "One who possesses the strength of wisdom, born of the moral precepts and restraints, who is tranquil in mind and delights in *meditation*, who is mindful, free from attachment, free from fallowness of mind and the Intoxicants, is called a sage by the wise" (I.12, verse 212; Saddhātissa 1985, p. 22; italics added). Here "meditation" is apt, as long as the English word is understood in its pseudo-etymological sense of (*re*)*centering* the mind, an approximation for one of the definitions of "jhāna," the mind "focused on one point" (Skt., *ekagrātā*), and as long as the object of this concentration is understood as being nondiscursive.

The technical usage of "jhānas" in the plural refers to particular meditative states, often translated as "absorptions" or "enstasis." In the Buddhist canon the *jhānas* gradually were systematized to include four stages. An even more elaborate description of these stages in the canon mentions how the practitioner moves through these four successive absorptions, then enters the four "attainments" (*samāpattis*), which culminate in the ninth stage with cessation of perception and feeling (Pali, *saññāvedayitanirodha*), better known as the attainment of cessation (Skt., *nirodha-samāpatti*).

Despite the importance of these nine meditative states, no Indian Buddhist school ever focused exclusively on the practice of absorptions or the practice of meditation. Such developments in the Chinese cultural sphere constitute a huge semantic leap and a complete reinterpretation of the tradition. (See Griffiths [1993] on *jhānas*.)

THE EMERGENCE OF CHAN AS A DISTINCT MOVEMENT

Details of how Buddhism entered East Asia around the first century CE and gradually spread within the Chinese cultural sphere remain surprisingly ill defined. At a certain point after the end of this transmission process, in some circles, meditation ceased to be considered as only one of the three central methods of self-cultivation (morality, concentration, and insight), and groups of practitioners started identifying themselves as adepts of Chan, understood in the sense of "meditation."

When did Chan Buddhism emerge as a distinct movement, historically and geographically? Here again caution is required, because those seeing themselves as spokespersons for Chan largely defined their identity in contrast with other Buddhist schools prevalent at that time. If we adopt the scheme proposed by John McRae (2003), this phase began with proto-Chan around 500–600 CE, which coincides with the growing success of the rival Tiantai lineage. In the following stage,

> at the beginning of the eighth century the self-described successors to this community exploded on the national scene, and in the process they described themselves as an identifiable religious movement using the lineage model. No matter how diverse and multifaceted the Chan movement was at this point in time, no matter how fuzzy the boundaries were between it and other realms of Chinese religious life, from this point onward Chan had achieved a significant level of sectarian identity. (McRae, p. 121)

Yet in the Chinese context it would be inaccurate to speak of members of an organized "meditation school." Even in the ninth century, Guifeng Zongmi (780–841) included in his *Chan yuan zhu quanji duxu* (Preface to the collected writings on the source of Chan) a list of Chan teachers that included the Tiantai patriarch Zhiyi (538–597) (Gregory 1991). Put differently, "Chan" in the sense of meditation never exclusively belonged to the Chan School. For one, it was part of the Buddhist legacy and played a central role in the practice of other lineages, such as the Tiantai School. For another, as John McRae convincingly argues, the organization of Chinese Buddhism never implied a sectarian-centered administrative system. Despite a heavy bureaucratic apparatus and the government-sponsored system that emerged in the Song dynasty (960–1279), monasteries were mostly administrated in rotation by the different lineages, and in China sectarian borders never became as strictly delimited as in premodern and modern Japan.

Sectarian developments took a further turn in Japan during the Tokugawa period (1600–1867) and evolved into the present rigid structures at the beginning of the Meiji era (1868–1912). Yet even in Japan until at least the eighteenth century, the expression "Chanzong" (Jpn., *Zenshū*) meant the Chan lineage or the principle of Chan,

and by no means referred to the Zen School or any institutionalized sect. The latter connotation emerged in Japan after the Meiji Restoration, when in 1872 the new Ministry of Doctrine created a single school labeled "Zen sect" including the Rinzai, Sōtō, and Ōbaku denominations. However, this forceful attempt to centralize Buddhist institutions met such strong clerical opposition that the government quickly stepped back. It recognized the independence of the Rinzai and Sōtō schools in 1874, then the autonomy of the Ōbaku School in 1876.

Geographically, where did Chan emerge? Saying that it emerged in the Chinese cultural sphere, rather than in China, aims at avoiding the easy assumption that China (understood as the modern nation) was the one and only cradle in which the Chan tradition grew up. This point is still controversial. Thich Nhât Hanh, a leading representative of the Vietnamese Thiên tradition, claims that the area of Jiaozhou, a colony of southern China from 111 BCE until 939 CE corresponding to present Thuan Thanh in northern Vietnam, saw the emergence of such a tradition at a much earlier time. He argues, "Buddhism was first introduced to Vietnam from India via the sea trade routes, beginning around the first century CE. Many people think that Buddhism came to Vietnam through China, but in fact it arrived first in Vietnam from India and was subsequently introduced to southern China from Vietnam" (2001, p. 4). This idea is appealing, especially to demonstrate that the Vietnamese Buddhist tradition is older than that of its former Chinese oppressor, but the additional suggestion that Buddhism was introduced to southern Vietnam (Cham at the time) from Jiaozhou seems difficult to support. Further, Nhât Hanh's presenting Kuong Tang Hôi (Chin., Kang Senghui; d. 280) as the first patriarch of Zen in Vietnam is questionable. Unfortunately, Nhât Hanh's ambiguous use of the word "Zen" and his agenda to demonstrate the antiquity of the Vietnamese tradition with a candidate who predates proto-Chan by more than two centuries undermine the reliability of his perspective. (For a balanced evaluation of the construction of Vietnamese orthodoxy, see Nguyen's [1997] study.)

THE PHILOSOPHICAL TURN

There are contemporary philosophers who seek inspiration in Zen or Buddhism, and there is a philosophical endeavor within the tradition itself. The former case stretches from intellectual curiosity to the commitment of Nishida Kitarō (1870–1945), whose Zen practice laid the basis for a major part of his philosophical work. Yet even Nishida claimed not to formulate a Zen philosophy, but only to reflect about universal philosophical problems in the light of his personal understanding of Buddhism. In any case, Nishida and his philosophical project must be appreciated within the context of the Japanese industrial revolution. Japan was engaged in importing techniques and culture from the West at a high pitch. To compensate for the unbalance caused by this new situation, Japanese thinkers sought to highlight the unique aspects of Japanese culture. This desire found expression in efforts by Nishida and others to demonstrate the compatibility or superiority of the alleged intuitive thinking of the East with the newly imported Western rationality. (About Nishida, see Heisig 2001, Tremblay 2000, and Yusa 2002.)

The philosophical articulation of the tradition itself is a more difficult subject. The difficulty stems not from the alleged absence of rationality in the East Asian Buddhist tradition, a critique overcome by the dedicated work of a generation of scholars. Rather, it results from the absence of a clear demarcation between Zen philosophy and Buddhist philosophy. Kasulis observes that, despite a huge literary production, traditional Zen accounts fail to justify their distrust of verbal distinctions or dualistic formulations "simply because it has already been offered by traditions influential in the very emergence of Zen Buddhism" (1981, p. 15).

Here the term "Zen Buddhism" (an oddity coined by Daisetsu Suzuki that should be avoided in scholarly contexts) confirms the suspicion that philosophical questioning cannot be confined to Zen alone, insofar as it constitutes a subcategory of the Buddhist worldview. Past attempts to present Zen as special and unfathomable are now better understood for what they were: sectarian proselytism. This observation does not prevent one from asking whether, after all, the Zen traditions have a specific philosophical perspective to offer.

SPECIFIC FEATURES OF ZEN

The quest to discover the real self, with subtle nuances sometimes labeled as "awakening" or "seeing one's true nature," is not a uniquely Zen feature. No doubt it occupies a central place in the Zen traditions, but it equally belongs to all Buddhist schools, being precisely what makes them Buddhist. Yet each particular Buddhist approach definitely displays a different flavor. For instance, Pure Land practices favor more devotional attitudes, while the Tiantai or Huayan traditions tend to privilege a more intellectual apprehension of the Buddhist path. Specific features can be found in the style of teaching, in the emphasis on particular types of cultiva-

tion, in doctrinal formulations or textual records, in rituals, and in the interactions with distinctive sociohistorical contexts. The use of vernacular Chinese in the Chan literature to duplicate or imagine dialogical encounters also constitutes a new genre.

The special features of the various Zen lineages did not pop up from some transhistorical background. Scholars now unanimously agree that most of the above-mentioned features of Chan in the Chinese cultural sphere found some degree of standardization during the Song dynasty. This means that even philosophical investigations into the spiritual cultivation of past Zen practitioners must cope with a double-layered filter: the construction of various orthodoxies in the Song period and subsequent interpretations by proselytes, which often replicate or amplify the first filter. With this in mind, let us nevertheless examine an example of a Zen teaching device where the context appears sufficiently explicit.

CRITICAL VOICE OR RHETORICAL DEVICE

One of the literary monuments of the Song period is the *Blue Cliff Record* (*Biyan lu*), a Chan anthology with commentaries by Yuanwu Keqin (1063–1135). The following dialog is provided here as it stands as case 11 in the *Hundred Cases of Xuedou*, the older version containing only the cases selected by Xuedou Zhongxian (980–1052) without Yuanwu's commentary:

> Huangbo taught the Sangha saying: "All of you people are stuffing yourselves with wine lees! If you keep roaming this way, how could this [decisive] moment [ever] arrive? Are you aware of the *nonexistence* of Chan teachers in the whole Tang China?"
>
> At this point, a monk emerged [from the crowd] saying: "What about all those [like you] who help students and lead the Sangha?"
>
> Huangbo.—"I didn't say Chan is nonexistent, only teachers are nonexistent." (Taishō 48: 151b11–b16)

Previous translations used the expression "gobblers of dregs," which sounds good in English and has the advantage of evoking lowlifes, but remains unsatisfactory. The provocation at the beginning of the passage refers to wine lees to make listeners aware that just as eating wine lees leads to intoxication, so depending on teachers and repeating teachings without personal insight is delusive.

Another overlooked dimension of this passage is the allusion of this metaphor to the Buddhist canon. The *Nirvana Sutra* tells an elaborate story about an ignorant king debating with a wise physician. The physician describes a marvelous medication that counteracts the effects of poison. This drug is actually a particular type of milk produced under strict conditions:

> If the cows don't eat *wine lees*, smooth grasses, or barley chaff, their calves will choose the good [path]. For grazing they will neither stay in the highlands nor come down to swamps. They will drink in limpid streams and won't be forced to run. They will not gather in herd with the bulls. Their drinking and eating will be adjusted; they will fit walk or immobility and find their place. Milk [produced] this way perfectly eliminates all ailments. (Taishō 12: 378c04–07)

Should this metaphor remain obscure, several commentators offer keys to understanding it. Huiyuan (523–592) of Jingying Temple provides a straightforward explanation: "If the cows represent the bodhisattvas … , *wine lees* represent ignorance, smooth grasses represent avidity, and barley chaff represents anger" (Taishō 37: 651a19–21).

In this light, the utterance attributed to Huangbo (d. c. 850) is far from simple rudeness to his audience or a dismissive critique of contemporary teachers. This teaching is a rhetorical device pointing at the auditors' fundamental ignorance and need to rediscover true autonomy. Whether Huangbo really uttered these words is best answered by Wright's careful statement: "The Huang Po texts available for our reading should be attributed not to any one creative individual or mind, but rather to the Zen tradition in China as it took shape over many centuries" (1998, p. 18).

The above excursion into the maze of intertextuality serves three purposes. First, it illustrates the immaturity of most Chan translations. Second, it shows the interdependence of Chan texts and Buddhist classics, and the need for further integration of the two fields. Third, it allows one to envision these dialogs in the context of monastic practice.

THE PRIORITY OF SOTERIOLOGICAL CONCERNS

If a common thread binds together the different Zen lineages, it may seem to be their uncompromising emphasis on awakening, based on the premise that the means and the end ultimately are not separate. In his characterization of Buddhism in general, Guy Bugault notes the primacy of the soteriological dimension over theoretical

constructs, saying, "Accurately speaking, Buddhism at its original stage was neither a religion nor a philosophy, but a psychosomatic discipline including three elements: morality (*śīla*), concentration (*samādhi*), and intellectual discernment or *acies mentis* (*prajñā*). None of them can function without the other" (1994, p. 43, translated from the French).

As with the poisoned arrow representing existential dis–ease (Pali, *dukkha*), the most urgent task is to remove it, speculations about its nature or shape being no more than delusive thought. Acute intellectual discernment is required to remove the arrow. The subtle boundary separating concentration and intellectual discernment is itself a theme worthy of examination, from both the Zen and philosophical perspectives. If there is a philosophical aspect specific to the Zen traditions, it is not so much in their striving to remove the arrow than in their emphasis on going beyond it, aiming at removing all traces of the operation.

See also Buddhism; Buddhism—Schools: Hua yan; Dogen; Jinul.

Bibliography

Bugault, Guy. *L'inde pense-t-elle?* Paris: Presses universitaires de France, 1994.

Buswell, Robert E., Jr. *The Zen Monastic Experience: Buddhist Practice in Contemporary Korea*. Princeton, NJ: Princeton University Press, 1992.

Faure, Bernard. *Chan Insights and Oversights: An Epistemological Critique of the Chan Tradition*. Princeton, NJ: Princeton University Press, 1993.

Faure, Bernard. *Double Exposure: Cutting across Buddhist and Western Discourses*. Translated by Janet Lloyd. Stanford, CA: Stanford University Press, 2003.

Foulk, T. Griffith. "Myth, Ritual, and Monastic Practice in Sung Ch'an Buddhism." In *Religion and Society in T'ang and Sung China*, edited by Patricia Buckley Ebrey and Peter N. Gregory. Honolulu: University of Hawai'i Press, 1993.

Gregory, Peter N. *Tsung-mi and the Sinification of Buddhism*. Princeton, NJ: Princeton University Press, 1991.

Griffiths, Paul J. "Indian Buddhist Meditation." In *Buddhist Spirituality*. Vol. 1: *Indian, Southeast Asian, Tibetan, and Early Chinese*, edited by Yoshinori Takeuchi and Jan Van Bragt. New York: Crossroad, 1993.

Heine, Steven, and Dale S. Wright, eds. *The Kōan: Texts and Contexts in Zen Buddhism*. New York: Oxford University Press, 2000.

Heisig, James W. *Philosophers of Nothingness: An Essay on the Kyoto School*. Honolulu: University of Hawai'i Press, 2001.

Heisig, James W., and John C. Maraldo, eds. *Rude Awakenings: Zen, the Kyoto School, and the Question of Nationalism*. Honolulu: University of Hawai'i Press, 1994.

Hoffman, Frank J. *Rationality and Mind in Early Buddhism*. Delhi: Motilal Banarsidass, 1987.

Hori, Victor Sōgen. *Zen Sand: The Book of Capping Phrases for Kōan Practice*. Honolulu: University of Hawai'i Press, 2003.

Huiyuan of Jingying Temple. *Dapan niepanjing yiji*. Taishō 37 no. 1764, pp. 613–903.

Kasulis, Thomas P. "The Zen Philosopher: A Review Article on Dogen Scholarship in English." *Philosophy East and West* 28 (3) (1978): 353–373.

Kasulis, Thomas P. *Zen Action, Zen Person*. Honolulu: University Press of Hawaii, 1981.

Kirchner, Thomas Yūhō. *Entangling Vines: Zen Koans of the Shūmon Kattōshū*. Kyoto: Tenryu-ji Institute for Philosophy and Religion, 2004.

McRae, John R. *Seeing through Zen: Encounter, Transformation, and Genealogy in Chinese Chan Buddhism*. Berkeley: University of California Press, 2003.

Mohr, Michel. "Hakuin." In *Buddhist Spirituality*, Vol. 2: *Later China, Korea, Japan, and the Modern World*, edited by Yoshinori Takeuchi, J. W. Heisig, P. L. Swanson, and J. S. O'Leary. New York: Crossroad, 1999.

Mohr, Michel. "Emerging from Non-duality: Kōan Practice in the Rinzai Tradition since Hakuin." In *The Kōan: Texts and Contexts in Zen Buddhism*, edited by Steven Heine and Dale S. Wright. New York: Oxford University Press, 2000.

Mohr, Michel. "Imagining Indian Zen: Tōrei's Commentary on the *Damoduoluo chanjing* and the Rediscovery of Early Meditation Techniques during the Tokugawa." In *Zen Classics*, edited by Steven Heine and Dale S. Wright. New York: Oxford University Press, 2005.

Nguyen, Cuong Tu. *Zen in Medieval Vietnam: A Study and Translation of the Thien Uyen Tap Anh*. Honolulu: University of Hawai'i Press, 1997.

Nhât Hanh, Thich. *Master Tang Hôi: First Zen Teacher in Vietnam and China*. Berkeley, CA: Parallax Press, 2001.

Nirvana Sutra. Translation attributed to Dharmaksema, Taishō 12 No. 374 pp. 365–604.

Nishida Kitarō. *L'éveil à soi*, edited by Jacynthe Tremblay. Paris: CNRS éditions, 2003.

Saddhātissa. *The Sutta-Nipāta*. London: Curzon Press, 1985.

Sharf, Robert H. *Coming to Terms with Chinese Buddhism: A Reading of the Treasure Store Treatise*. Honolulu: University of Hawai'i Press, 2001.

Suzuki, Daisetz Teitaro. *Essays in Zen Buddhism*. First series (1927). New York: Grove Press, 1961.

Tremblay, Jacynthe. *Nishida Kitarô: Le jeu de l'individuel et de l'universel*. Paris: CNRS éditions, 2000.

Williams, Duncan Ryūken. *The Other Side of Zen: A Social History of Sōtō Zen Buddhism in Tokugawa Japan*. Princeton, NJ: Princeton University Press, 2005.

Wright, Dale S. *Philosophical Meditations on Zen Buddhism*. New York: Cambridge University Press, 1998.

Yusa, Michiko. *Zen and Philosophy: An Intellectual Biography of Nishida Kitarō*. Honolulu: University of Hawai'i Press, 2002.

Yuanwu Keqin. *The Blue Cliff Record*. Translated by Thomas Cleary. Berkeley, CA: Numata Center for Buddhist Translation and Research, 1998.

Yuanwu Keqin, and Xuedou, Zhongxian. *Biyan lu*. Taishō 48 no. 2003, pp. 139–225.

Michel Mohr (2005)

DGE-LUGS

The Dge-lugs (pronounced "geluk") tradition of Tibetan Buddhism, the tradition followed by the Dalai Lamas, traces its origins to the towering Tibetan philosopher and monastic reformer Tsong kha pa (1357–1419) and his two closest disciples, Rgyal-tshab (pronounced "gyelt-sap") (1364–1432) and Mkhas grub (pronounced "kay-drup") (1358–1438), whose views have come to represent orthodoxy for the tradition. According to traditional hagiographies, Tsong kha pa studied with more than sixty of the greatest scholars in Tibet during his early life and went on to compose numerous treatises and commentaries on the entire spectrum of Buddhist thought and practice, leaving a set of collected works that numbers nineteen volumes. His philosophical works address virtually all the major topics in Buddhist philosophical discourse, including issues of ontology, metaphysics, epistemology, logic, soteriology, and hermeneutics, among others. Aided by historical and political conditions Tsong kha pa's works, those of his disciples, and the monastic and educational systems he initiated made the Dge-lugs tradition the dominant philosophical tradition in Tibet. Indeed, Tsong kha pa was such a powerful intellectual force in Tibet that all subsequent philosophers, including those who disagreed with him, felt compelled to acknowledge and address Dge-lugs-type criticisms that they anticipated their views might incur.

While there is much in common among Dge-lugs philosophers in terms of their philosophical positions and methods, it would be misleading to suggest that the Dge-lugs tradition and its notable philosophers are univocal in their philosophical presentations. Many lively debates and polemic directed at fellow members of the Dge-lugs tradition can be found in the works of the great thinkers of the tradition, including Tsong kha pa's direct disciples Rgyal-tshab and Mkhas grub, as well as later thinkers such as 'Jamdbyangs bzhad pa (1648–1721), Rje btsun Chos kyi 'gyal mtshan (1469–1544), and Lcang skya rol baï rdo rje (1717–1786), among others. Despite this marked diversity of opinion, there is nonetheless a relatively standard Dge-lugs philosophical presentation that those in the tradition generally agree on. The intra-tradition debates tend to focus on lofty and quite subtle points, while the mainstream Dge-lugs philosophical worldview accepted across the tradition remains as the foundation for debates about such subtle points of contention.

Many significant features of Dge-lugs philosophy stand in contrast with other Buddhist traditions. Among the most significant is the marked emphasis Dge-lugs philosophers place on the study of the Indian Buddhist philosophical tradition they inherited and on what they understand to be the correct interpretation of that tradition. Thus any discussion of Dge-lugs philosophy must be approached through an examination of how the earliest Dge-lugs masters interpreted and represented Indian Buddhist philosophical history.

DGE-LUGS MADHYAMAKA

While the works of many Indian philosophers have impacted Dge-lugs philosophy, the Dge-lugs tradition traces its intellectual lineage most significantly through two important Indian philosophers: Nāgārjuna (c. first century C.E.) and Candrakīrti (c. 600–650). Nāgārjuna, author of the *The Fundamental Wisdom of the Middle Way* (*Mūlamadhyamakakārika*), among other texts, is considered the founder and systematizer of the school of Mahayana philosophical thought known as Madhyamaka or the Middle-Way School. Virtually all Tibetan Buddhist schools consider themselves to be Mādhyamikas, followers of Nāgārjuna's tradition in one form or another and the Dge-lugspas are no exceptions in this regard. The central idea that guides the thought of Nāgārjuna and the Madhyamaka School is the notion of emptiness (*śūnyatā*). When Mādhyamikas such as Tsong kha pa use the term "emptiness," they mean that an object lacks a fixed or unchanging nature. To say that a pot, for example, is empty (metaphysically empty) is to say that it lacks a permanent nature or essence, an independent, intrinsic identity.

The Dge-lugs presentation of the middle way owes much to their reading of the history of their Indian Madhyamaka predecessors. When Dge-lugs philosophers and scholars study the history of Indian Madhyamaka, they begin by recognizing that Nāgārjuna and his student Āryadeva are considered authoritative by all subsequent commentators and interpreters of Madhyamaka thought. According to Tsong kha pa's assessment of the history of Indian Madhyamaka, an important philosophical split occurred in Madhyamaka discourse several centuries after Nāgārjuna when Buddhapālita (c. 470–540?) wrote a commentary on Nāgārjuna's *Fundamental Wisdom of the Middle Way*. This was followed by a criticism of that treatise by Bhāvaviveka (c. 500–570?) and a subsequent defense by Candrakīrti of Buddhapālita's position against those criticisms leveled by Bhāvaviveka. Much of this debate in the Indian tradition revolved around the appropriate form of reasoning to be utilized by Madhyamaka philosophers. With this point in mind, later Tibetans such as Tsong kha pa distinguished subschools of Indian Mad-

hyamaka philosophy, in part on the basis of the form of reasoning that their proponents utilize.

Buddhapālita's commentary, simply titled (in English) *Buddhapālita's Commentary on [Nāgārjuna's] "Fundamental Wisdom of the Middle Way"* is a lucid exposition of Nāgārjuna's text that utilizes a method known as consequentialist argument (*prasaṅga*). Much as in Nāgārjuna's text, Buddhapālita's form of argumentation examines the positions of philosophical rivals to reveal the absurd consequences of holding such positions. In other words, consequentialist arguments attempt to reduce the philosophical positions of opponents to absurdities. All philosophical opponents of Mādhyamikas maintain that some things are not empty, have a true nature or essence, and have independent, permanent existence. For all such contemporary opponents, Buddhapālita, like Nāgārjuna before him, attempted to reveal what he saw to be the absurd positions entailed by their various positions asserting true existence. Though the logical innovations of Dignāga (c. 480–540) were beginning to make headway into Mahayana discourse, Buddhapālita avoided these innovations in logic, such as the use of independent argument (*svatantrānumāna*), thus avoiding commitment to a counterposition when criticizing his opponents. Dge-lugspas have tended to presume that Buddhapālita was simply and faithfully following the method of Nāgārjuna.

Bhāvaviveka, in contrast, argued that Mādhyamikas must assert their own philosophical position. Simply to criticize others without establishing one's own position, the emptiness view, is inadequate. And to establish one's own position, Bhāvaviveka argued, one must use independent inferences. Thus, in *Prajñā-pradīpaḥ: A Commentary on the Madhyamaka Sūtra*, his commentary on Nāgārjuna's *Fundamental Wisdom of the Middle Way*, Bhāvaviveka offers a pointed criticism of Buddhapālita's failure to establish a Madhyamaka thesis, as well as an exposition of the need to use independent argument (*svatantrānumāna*) to fulfill that task. Accordingly, Dge-lugspas categorized Bhāvaviveka's particular brand of the middle way as Svātantrika-Madhyamaka. In contrast, Dge-lugs doxographers describe Buddhapālita and his defender Candrakīrti (described below) as Prāsaṅgika-Mādhyamikas, because they insist on primarily using consequentialist arguments (*prasaṅga*).

Candrakīrti (c. 600–650) is the third important figure in this Indian Madhyamaka debate, according to Dge-lugs authors. Candrakīrti composed several philosophical texts, two which are important to Dge-lugs philosophers on the central issue of the appropriate form of reasoning for proponents of Madhyamaka views: his *Introduction to the Middle Way* (*Madhyamakāvatāra*) and *Lucid exposition of the middle way* (*Prasannapadā Madhyamakavṛtti*. In these texts Candrakīrti philosophically defended Buddhapālita against the criticisms leveled by Bhāvaviveka. Candrakīrti argued not only that Buddhapālita was correct to use only consequentialist arguments, but also that using independent arguments are incompatible with Madhyamaka tenets.

In the Dge-lugs reading of this debate, there is a fundamental ontological problem with using independent arguments. Such usage implies acceptance of an inherent, absolute, or unchanging nature in phenomena, and this implication is utterly contrary to the most basic Madhyamaka tenet—that all phenomena are empty of just such a nature or essence. Dge-lugs philosophers such as Tsong kha pa argued that because one characteristic of an independent inference is a commonly appearing subject in the inference, the inference implies that the subject must have some sort of absolute existence. For example, in the independent inference "This book is impermanent because it is produced," the subject, this book, must appear in precisely the same way, in a way which is common to both the proponent and opponent of the argument. If it does not, then the two debaters are just talking past each other. If it does have a precise and common mode of appearance to both the proponent and opponent, then it must have some absolute mode of existence, some intrinsic identity, some sort of inherent nature.

Thus, the mere use of independent arguments runs utterly contrary to the Madhyamaka view, according to Tsong kha pa and his Dge-lugs followers. Although Tsong kha pa considered Buddhapālita to be a Prāsaṅgika-Mādhyamika from his views and method, he considered Candrakīrti to be the "founder" of the Prāsaṅgika-Madhyamaka view, since he was the first clearly to articulate the importance of using consequentialist arguments and the contradictions involved when Mādhyamikas use independent arguments.

An interesting feature of Tsong kha pa's middle way is that though he recognized the limits of language, he still insisted on rationality and the laws of logic in his investigations and conclusions concerning the ultimate. In this sense, notes Georges Dreyfus in *The Sound of Two Hands Clapping*, Tsong kha pa ought to be considered a realist. Śāntarakṣita, an eighth-century scholar who was a key figure in the early dissemination of Buddhism in Tibet, was a late Indian Mādhyamika who incorporated the logico-epistemological (*pramāṇavāda*) innovations of Dignāga and, more prominently, Dharmakīrti (seventh

century) into his particular brand of the middle way. Though Śāntarakṣita was considered to be a Svātantrika-Mādhyamika with whom he took issue on several counts, Tsong kha pa nevertheless preserved, and even intensified, Śāntarakṣita's emphasis on the role of reason in his Madhyamaka method. The particularities of how Tsong kha pa integrated the logico-epistemological tradition into Madhyamaka analysis are central to the insights that made his thought unique.

To turn now to the Tibetan Madhyamaka tradition, for Dge-lugs philosophers, an issue central to all Madhyamaka philosophical analysis and inseparably tied to the issue of an appropriate logic is the issue of the two types of truth: ultimate truth (*don dam bden pa* [Tibetan], *paramārthasatya* [Sanskrit]) and conventional truth (*kun dzob bden pa*, *saṃvṛtisatya*). Truths in this context are objects of knowledge. Hence it makes sense to talk of truths existing. Since its earliest formulation in the works of Nāgārjuna, Madhyamaka thinkers have used a presentation of the two types of truths as a primary marker against which they have delineated their positions on central Buddhist topics in ontology and epistemology.

Dge-lugs philosophers illuminated the distinctions they drew between ultimate and conventional truths by contrasting their positions as Prāsaṅgika-Mādhyamikas with the position of their Madhyamaka rivals, the so-called Svātantrika-Mādhyamikas, such as Bhāvaviveka. This takes place in the treatises of Tsong kha pa and his direct disciples and also in a genre of philosophical literature prominent in monastic study and known as tenet-system texts or doxographies. In this literature, Dge-lugs authors present major systems of non-Buddhist and Buddhist philosophical thought in a hierarchically organized fashion. Each of the tenet systems (or at least the Buddhist systems) are presented in terms of a host of philosophical categories of analysis, such as the two truths, definitions of consciousness and objects of consciousness, delineation of the path, delineation of the fruits of the path, and so on. Consistency in analytic categories across the presentation of schools of thought facilitates easy comparisons between systems and usefully allows one easily to ascend a hierarchy of philosophical positions in a dialectical fashion by contrasting the present system with the less subtle and less accurate system just below it.

For example, one can easily compare all four Buddhist schools' definitions of ultimate truths, conventional truths, consciousness, and the like, by seeing that school *x* defines a conventional truth in one way, school *y* in another, and school *z* in yet another. Often the views presented in this literature do not precisely mirror those of any single Indian author, but rather are amalgamations of the views of several presumably like-minded thinkers and of unstated positions considered to follow logically from other stated positions. As mentioned above, for Dge-lugs thinkers, the highest and most accurate Buddhist philosophical description of the nature of reality is the Prāsaṅgika-Madhyamaka. Just below that position in the hierarchy is the Svātantrika-Madhyamaka view. Thus, the Svātantrika view is most commonly contrasted with the Prāsaṅgika-Madhyamaka position to help illuminate the Dge-lugs-Prāsaṅgika view.

When Dge-lugs authors discuss the issue of the two types of truths, they employ a number of key technical terms. Jeffrey Hopkins mentions sixteen terms in his book *Meditation on Emptiness*, of which the six most commonly used are the following:

- Ultimately established existence (*don dam par grub pa*)
- Truly established existence (*den par grub pa*)
- Existence established in reality (*yang dag par grub pa*)
- Existence established by way of the intrinsic identity/characteristics of an object (*rang gi mtshan nyid kyis grub pa*)
- Existence established by way of the inherent nature of an object (*rang bzhing gyis grub pa*)
- Existence established from its own side (*rang ngos nas grub pa*)

According to Dge-lugs philosophers such as Tsong kha pa, all Mādhyamikas, including the Prāsaṅgikas and the Svātantrikas, held that the first three terms on the list accurately describe ultimate truths, since such truths lack (are empty of) ultimately established existence, truly established existence, and existence established in reality. An example of an ultimate truth for either a Svātantrika-Mādhyamika or a Prāsaṅgika-Mādhyamika would be the lack of any ultimately established existence or truly established existence in this book, for example. The lack of ultimately or truly established existence refers to the absence of any objective existence, any independent absolute mode of being, any fixed independent essence, within this book. Thus far, according to Dge-lugs thinkers, both subschools of Madhyamaka thought are in agreement.

The disagreement between the two subschools concerns their positions on the ontological status of conventional truths. According to Dge-lugs thinkers, while all Mādhyamikas, when presenting ultimate truths, argue

that phenomena lack an ultimate nature, the Svātantrika-Mādhyamikas accept that conventional truths exist in the latter three ways listed above; that is, they exist by way of their own intrinsic identity, by way of inherent nature, and from their own side. This, according to Dge-lugspas, is how the Svātantrika-Mādhyamikas could view their position as maintaining a middle ground between absolute permanence and absolute nonexistence, or nihilism. They avoid the extreme of permanence by saying that phenomena ultimately lack true existence. They avoid the extreme of nihilism by claiming that phenomena conventionally exist by way of their own characteristics, by way of their intrinsic nature, or from their own side.

In relation to the logical issues discussed above, because phenomena conventionally exist from their own side or by way of their own intrinsic identity/characteristics, objects such as books and tables can serve as commonly appearing subjects in independent inferences. An inherent nature or intrinsic characteristics on the side of the book, for example, cause an ignorant, unenlightened consciousness to recognize that object and correctly impute the conventional designation "book" on the basis of a nondefective conventional cognition. Such an imputation has a referent as object, to which it correctly points with a conventional designation ("book").

Dge-lugspas found this position, which they attributed to Svātantrika-Mādhyamikas, highly problematic. They argued that all six technical terms mentioned above to describe the ontological status of things are coextensive. If an object can be described in one of the six ways, it can be described in all six ways. Dge-lugspas thus argued that ultimate truths *and* conventional truths do not exist in any of the ways described above. They argued that if objects are established by way of their own intrinsic identity, by way of some sort of inherent nature of their own, or from their own sides, even conventionally, then objects must have some sort of ultimate nature. Dge-lugspas would criticize their Madhyamaka opponents by arguing that although they claim that some objects exist only conventionally, if they assert that the objects inherently possess some nature of their own in any way, even conventionally, this is really just a masked way of continuing to cling to some independent essence or nature in the objects ultimately. For an object to cause a conventional consciousness to correctly recognize and label it, there must be something true or absolute in the object. Thus in the Prāsaṅgika-Madhyamaka position held by Dge-lugspas, both ultimate and conventional truths lack all six criteria (sixteen as listed by Hopkins) of ultimate and conventional truths.

While Svātantrikas accept that conventional truths exist inherently, by way of their own characteristics, and from their own sides, Dge-lugspas, such as the Prāsaṅgikas, reject the idea that even conventional truths are established in this way. Conventional truths are actually falsities. There is nothing true about how minds under the sway of ignorance conceptualize these falsities. This is not to say that the world does not exist out there. It is just to say that we are utterly deluded when we engage with the world because we impose fixed essences in things when no such essences exist. And this is what Svātantrikas are doing when they claim that even mere conventional truths inherently exist. For Dge-lugspas, such as the Prāsaṅgika-Mādhyamikas, conventional truths are true only for a consciousness for which the actual nature of reality is obscured. They do not exist as they appear to a conventional consciousness. Dge-lugspas such as Tsong kha pa held that they avoided the extreme of nihilism by accepting the functionality of conventional phenomena despite the falsity of their appearances.

It is important to keep in mind that this is a standard Dge-lugs presentation of this history and these ideas. While Dge-lugs authors associated specific Indian Madhyamaka thinkers with these subschools of Madhyamaka thought, there does not appear to be evidence in Indian sources before the twelfth century of any explicitly named subschools of Madhyamaka thought. Prior to this time, the thinkers discussed here and labeled "Prāsaṅgika-Madhyamaka" or "Svātantrika-Madhyamaka" in the Dge-lugs literature identified their own views as simply Madhyamaka.

THE DGE-LUGS CURRICULUM AND SCHOLASTIC METHODS

Any discussion of Dge-lugs philosophy must move beyond ideas and also discuss the curriculum and methods employed in Dge-lugs institutions, which direct a significant amount of their focus to philosophical study. Tsong kha pa initiated a scholastic approach to Buddhism that, although presented to lesser degrees before him in Tibet, marked a significant departure from the mystical gnosis of individuals as sources of authority for the tradition before him. Tsong kha pa emphasized reasoning, which could be learned, in time, only in monastic universities, thus advancing a shift in authority from individuals to institutions and a wholesale reform of monastic culture, which he saw as deteriorating in Tibet during his time. As time went on and the monastic colleges grew, the

degree of scholasticism grew with it. Key figures from the Dge-lugs monastic colleges began to compose textbooks (*yig cha*) as manuals for study in attempt to present coherent, consistent, totalizing systems of thought immune to critique, especially internal contradiction. This move toward scholasticism certainly reinforced institutional authority, but the importance of mystical gnosis of expert scholar-adepts was far from lost in the Dge-lugs tradition, though the reins on the careers of independent yogis were significantly tightened in this tradition.

The Dge-lugs tradition maintains a large monastic component that includes three major monastic seats—Sera, Drepung, and Ganden—and several colleges within those monastic seats. Traditionally, the monastic seats housed between 5,000 and 10,000 monks in each, with good portions of the monks engaged in the philosophical curriculum of one of the monastic colleges. Even in exile in south India, Sera and Drepung each had more than 5,000 monks in residence in 2005. The colleges of the three monastic seats all have a similar curriculum that culminates in a degree known as "geshe." A geshe degree is somewhat akin to a doctorate in Buddhist philosophy. It generally takes somewhere between fifteen and twenty-five years to complete the curriculum, which includes study and memorization of all the major philosophical texts of the tradition, extensive periods of debate (usually four to six hours a day, six days per week), and study of the major commentaries and textbooks of the college, which serve as the interpretive frame through which to engage the major treatises of the Indian and Dge-lugs traditions. Though most monks at Sera, Drepung, and Ganden begin the geshe training, only a small percentage successfully complete the degree because of the difficulty of the subject matter and the rigors of the curriculum, again, much like a doctoral program in the West.

Each of the monastic colleges cover the same basic subjects, though they use different monastic textbooks (*yig cha*), commentaries that present the interpretive frameworks of their particular colleges. Here in the monastic textbooks one begins to find differences in interpretation on subtle philosophical points between authors within the Dge-lugs tradition. Often scholars from the different monastic colleges will take great pride in the monastic textbooks of their particular colleges and the interpretive framework they employ for understanding the philosophical views of Tsong kha pa, Candrakīrti, and other great philosophers. Within the curriculum there are preliminary subjects covering the basics of topics such as the forms of reasoning and debate, soteriological grounds, and paths; types of minds/consciousnesses according to the Buddhist tradition; the philosophical tenet systems of the four Indian Buddhist schools; and so on. Once these preliminary subjects are successfully completed, the Dge-lugs scholar progresses on to the five subjects of the geshe curriculum, which include the perfection of wisdom, maplike descriptions of states of consciousness as the practitioner removes obstacles to enlightenment and progresses along the Buddhist path, logic and epistemology, Madhyamaka philosophy, cosmology, and monastic ethics. These five subjects include topics on ethics, metaphysics, ontology, hermeneutics, karma, and personal identity among others. For each of these subjects, years are dedicated to primary philosophical texts, which are memorized and then studied intensively with a teacher, who gives the students informed oral explanation on the texts. Students then debate the ideas and fine-tune their understanding in the debate courtyards. Progress exams are given regularly, and the final exam includes a multi-day public debate with top scholars in the tradition.

Many of those who complete this geshe curriculum successfully go on for a sort of postdoc for one to three years at one of the two major tantric colleges, Gyume or Gyuto. Here they study the theory and practice of the major tantric meditational cycles in the Dge-lugs tradition. Completing all these requirements usually qualifies the student as a teacher. Thus, authority is granted primarily through institutions, though this curriculum ideally cultivates—and indeed, was constructed to cultivate—experiential/gnostic authority as well.

The Dge-lugs tradition of Tibetan Buddhism is perhaps the most scholarly and philosophical of all the world's Buddhist traditions. As a living Buddhist tradition, it makes for a fascinating area of investigation, not only for those interested in the history of Buddhist philosophy in general, but also and particularly for those interested in understanding the ideas and structure of a living Buddhist tradition, and understanding how philosophy and philosophical study are integrated with a larger human path.

See also Buddhism; Buddhism—Schools: Madhyamika; Buddhism—Schools: Yogācarā.

Bibliography

Blumenthal, James. *The Ornament of the Middle Way: A Study of the Madhyamaka Thought of Śāntarakṣita*. Ithaca: Snow Lion Publications, 2004.

Cabezón, José Ignacio. "The Canonization of Philosophy and the Rhetoric of Siddhānta in Tibetan Buddhism." In *Buddha*

Nature: A Festschrift in Honor of Minoru Kiyota, edited by Paul J. Griffiths and John P. Keenan. San Francisco: Buddhist Books International, 1990.

Cabezón, José Ignacio. *A Dose of Emptiness: An Annotated Translation of the sTong thun chen mo of mKhas grub dGe legs dpal bzang*. Albany, NY: SUNY Press, 1992.

Cabezón, José Ignacio. "The Prāsaṅgikas' Views on Logic: Tibetan Dge-lugs pa Exegesis on the Question of Svatantras." *Journal of Indian Philosophy* 16 (1988): 217–224.

Cozort, Daniel. *The Unique Tenets of the Middle Way Consequence School*. Ithaca, NY: Snow Lion, 1998.

Dreyfus, Georges. *The Sound of Two Hands Clapping: The Education of a Tibetan Buddhist Monk*. Berkeley: University of California Press, 2003.

Dreyfus, Georges. "Tibetan Scholastic Education and the Role of Soteriology." *Journal of the International Association of Buddhist Studies* 20 (1) (1997): 31–62.

Hopkins, Jeffrey. *Emptiness in the Mind-Only School of Buddhism*. Berkeley: University of California Press, 1999.

Hopkins, Jeffrey. *Maps of the Profound: Jam-yang-shay-ba's Great Exposition of Buddhist and Non-Buddhist Views on the Nature of Reality*. Ithaca, NY: Snow Lion Publications, 2003.

Hopkins, Jeffrey. *Meditation on Emptiness*. London: Wisdom Publications: 1983.

Hopkins, Jeffrey. *Reflections on Reality: The Three Natures and Non-natures in the Mind-Only School*. Berkeley: University of California Press, 2002.

Hopkins, Jeffrey. "The Tibetan Genre of Doxography: Structuring a Worldview." In *Tibetan Literature: Studies in Genre*, edited by José Ignacio Cabezón and Roger R. Jackson. Ithaca, NY: Snow Lion Publications, 1996.

Jinpa, Thupten. "Delineating Reason's Scope for Negation: Tsongkhapa's Contribution to Madhyamaka's Dialectical Method." *Journal of Indian Philosophy* 26: 275–308, 1988.

Jinpa, Thupten. *Self, Reality, and Reason in Tibetan Philosophy: Tsongkhapa's Quest for the Middle Way*. London: RoutledgeCurzon, 2002.

Klein, Anne. *The Path to the Middle: Oral Mādhyamika Philosophy in Tibet; The Spoken Scholarship of Kensur Yeshey Tupden*. Albany, NY: SUNY Press, 1994.

Lopez, Donald S. *A Study of Svātantrika*. Ithaca, NY: Snow Lion Publications, 1987.

Napper, Elizabeth. *Dependent Arising and Emptiness: A Tibetan Buddhist Interpretation of Mādhyamika Philosophy Emphasizing the Compatibility of Emptiness and Conventional Phenomena*. Boston: Wisdom Publications, 1989.

Ruegg, David Seyfort. *Two Prolegomena to Madhyamaka Philosophy*. Vienna: Arbeitskreis für Tibetische und Buddhistische Studien, Universität Wien, 2002.

James A. Blumenthal (2005)

HUA YAN

The Hua yan school of Buddhism developed in China between 600–1000 CE, flourishing at the end of the Tang dynasty. It relies for much of its doctrine on exegesis of the Mahayana Buddhist scripture known as the *Hua yan Jing*. The name *Hua yan* (Japanese: *Kegon*) is intended to be the Chinese translation of the Sanskrit *Avataṃsaka*, which means "flower garland." The term is ostensibly the title of a Sanskrit sutra, the *Mahāvaipulya Buddhāvataṃsaka Sūtra*. The Hua yan school developed a *panjiao* (system of classification of Buddhist doctrines), which takes the *Hua yan Jing* to be the most profound of all the Buddhist sutras. This is because it was, according to legend, spoken by the Buddha while in the throes of his awakening experience.

CENTRAL TEXTS

The term *vaipulya* in the title indicates that the text is a composite one, cobbled together from several other texts of various lengths and origins. Some parts of the text, for example the *Daśabhūmika* and the *Gandavyūha*, do exist in a Sanskrit original. In addition, some parts of the text are laden with Chinese phoneticizations of Sanskrit terms, which also indicate a likely Indian origin. The rest is more or less likely to be of indigenously Chinese origin, passed off as or uncritically taken to be translations of Sanskrit originals. For this reason, the origins of the Hua yan tradition are linked to the evolution of a fully sinicized Buddhism.

This is complicated by the fact that many of the key, pivotal translators and advocates of these materials were not indigenously Chinese but in fact were from Central Asia. China and India were kept culturally autonomous for a long time because of the daunting obstacle presented by the Himalayas, so early contact actually was more likely to take place in areas of easy access to the Silk Road. This complicated matters because of the cultural homogenization that also followed along with such developments. Since the latter part of the twentieth century, there has been much study about the extent to which the flow of ideas from many cultures along the Silk Road influenced the development of the uniquely Chinese forms of Buddhism.

There are two arguably complete versions, or translations, of the text in Chinese. The earliest consists of sixty chapters, produced by Buddhabhadra in about 420. Traditionally, this has been used by Hua yan writers as the standard text. In 699 a version in eighty chapters was produced by Śikṣananda. The only complete English translation of the Hua yan sutra, in three volumes, was produced by Thomas Cleary in the late 1980s. For reasons he does not explain, Cleary translates Śikṣananda's version, although it is not as historically important as Buddhabhadra's text.

In addition, there do exist various Chinese versions of parts of the sutra, such as the *Gandavyūha*, for which is there is a Sanskrit original.

Another text of crucial importance to the development of the Hua yan tradition is the *Dasheng Qixin Lun* (Mahayana awakening of faith). This text is also arguably an apocryphal text, written in Chinese but taken as a translation of a nonexistent Sanskrit text ostensibly titled *Mahayānaśraddhotpāda*. This text is cited by all the prominent Hua yan writers and is thus granted a substantial authority. This text has been linked to the ontologization of Buddhism as it developed in the Chinese context, perhaps due to Central Asian and Silk Road influences. Ideas that take shape in this text include such metaphysical notions as buddha nature and *tathāgatagarbha* (womb of buddhahood), which some scholars take to be countertheoretical to basic Indian Buddhist premises of the pointlessness of metaphysical assertions and speculations. In fact, within modern Japanese Zen traditions, there are those who suggest that East Asian Buddhism in general is not Buddhism. These critical Buddhists point precisely to the type of foundational *tathāgatagarbha* thinking that can be directly linked to the *Awakening of Faith* and its influence as topical, non-Buddhist elements that encroach on the central insights.

PATRIARCHAL LINEAGE

According to the retrospective view of Zongmi (780–841), there are four patriarchs or lineage figures in the Hua yan tradition, and he styles himself as the fifth patriarch. This comes to be seen as the orthodox lineage by the subsequent tradition. This standard list of patriarchs includes Dushun, Zhiyan, Fazang, Chengguan, and Zongmi. This is a retrospective lineage, which means that it is not at all clear that Dushun and Zhiyan saw themselves as members of a Hua yan school. This attribution is applied after the fact, as the tradition comes to consider the sources of its own emphases.

Dushun is said to have lived from 558 to 640. Although apparently prominent as an adept and miracle worker in his time, he is most influential as the purported author of a text known as the *Hua yan Fajie Guanmen* (Meditative approaches to the Hua yan Dharmadhātu). This text introduces the Four *Dharmadhātu* model that will be discussed later on, and thus provides a solid basis for the later developments in Hua yan thought.

Zhiyan (602–668), the second patriarch, is not as well known. His most prominent contribution to the discourse is the so-called Ten Mysteries. These are basically a series of metaphors for interpenetration and mutual causation, and many of them are in fact redundant. Regardless, this language persists in the work of Fazang, perhaps the grand systematizer of Hua yan thought.

Although attributed as the third patriarch, Fazang (643–712) may have been the first to think of himself as founding or joining a specific school of thought. Fazang's family was of Central Asian origin, in Samarqand, a prominent center on the Silk Road. A prolific writer, he wrote somewhere between sixty and one hundred works on various topics, the most important being commentaries on the *Hua yan Jing* and the *Mahayana Awakening of Faith*. He rose to prominence at the court of the empress Wu, after a series of performances in which he used such examples as a room of mirrors to demonstrate Hua yan principles of interpenetration and nonobstruction. Fazang's school stood in contrast to the school of Xuanzang, who had gone to India to learn Sanskrit and translate scores of Buddhist texts into Chinese. This conflict can be seen as being between the Indic and the sinicized forms of Buddhism. Ultimately, Fazang's view prevails, for a variety of philosophical, cultural, and political reasons. This may be an early and important stage in the sinicization of Buddhism.

Chengguan, the fourth patriarch, lived from 738 to 840. The lineage is somewhat obscure here, as Fazang's actual disciple, Huiyan, was understood by the later tradition to have corrupted the teaching. Chengguan, who was born after Fazang died, was nevertheless seen as the fourth patriarch in the sense that he is believed to have restored the integrity of Fazang's teachings. He did seem to have led a renewed interest in the school on the part of the ruling class and the scholars.

The last of the orthodox patriarchs is Zongmi (780–841). Zongmi is best known for his syncretic concerns, including his interest in sorting out the various schools of Buddhism, especially Chan Buddhism. Because of his interest in *panjiao*, his works are a treasure house of historical information about the schools of Buddhism active at his time. What is perhaps most significant about Zongmi is his concern with reconciling and synthesizing Hua yan and Chan Buddhism. In fact, Zongmi is sometimes attributed with lineage roles in both the Chan and Hua yan traditions, though these claims cannot be accepted uncritically. This leads to an oversimplification expounded by the famous Japanese Zen scholar Daisetz Suzuki, who argues that Hua yan is theoretical and establishes the principle behind Zen that is practical. However, this is too polemic a description of the situation, since Chan and Zen have a long textual and theoretical history, while Hua yan does provide practices of its

own, for instance the meditation on the Four *Dharmadhātus* discussed later in this entry.

Besides the so-called orthodox lineage just discussed, there are also a number of figures who belong to what might be called heterodox lineages in the sense that they follow exegetical lines of reasoning not adopted by the later traditions. These include, as mentioned, Fazang's student, Huiyuan, and the iconoclastic Li Tongxuan.

THE FOUR *DHARMADHĀTUS*

Perhaps the most fundamental concept in all of Hua yan Buddhist thinking is the synonymy of emptiness and dependent arising. Emptiness (Sanskrit: *śūnyatā*; Chinese: *kong*) is a traditional Buddhist notion that refers to the absence of self-being in all things and events. It does not mean that things do not exist—it means that all things that exist do so in dependence on other things, which is the meaning of dependent arising (Sanskrit: *pratītyasamutpāda*; Chinese: *yinyuan*). Hua yan, consistent with characteristic Chinese attitudes, placed focus on the positive side of this formulation, that even though empty, things actually do exist.

This is perhaps most clearly expressed in the model of the Four *Dharmadhātus* as initially formulated in Dushun's seminal text, *Meditative Approaches to the Hua yan Dharmadhātu*, and subsequently developed further by Chengguan. The term *dharmadhātu* is a way of referring to the realm of all *dharmas* (events). In other words, the *dharmadhātu* is the world in the most comprehensive sense. This model of the world is represented sometimes, especially in the work of Fazang, in terms of the metaphor of Indra's jeweled net. This net consists of many-faceted gems, each of which reflects every other gem, and reflects itself reflected in every other gem.

The formula of the Four *Dharmadhātus* is proposed as a support for meditation practices. Although they are often rendered in such a way as to suggest that there are four separate realms, they more properly represent four types or orders of perspectives on experience. The first is the tacit, uncritical commonsense lower-order perspective, and the others are higher-order or meditative perspectives. The goal seems to be a type of perspectival flexibility, which corrects the obsessive-compulsive tendency to identify with a single perspective by acknowledging the multiplicity of perspectives available and by adopting higher-order perspectives that reconcile the inconsistencies present between lower-order perspectives. This is like standing in a hallway with two people on either end. I can see one or the other, because of my limited perspective, but I cannot see both simultaneously. If I were to stand above the hallway somehow and look down on it, I might be able to see both at once. Higher-order perspectives similarly circumscribe and sustain perspectives that appear incompatible at the surface level.

The first of these types of perspectives is termed *shi*, often rendered as "phenomenon" or "event." This is the tacit, ordinary, conventional perspective adopted and identified with by most people most of the time. It takes events at more or less face value—it does not raise questions about metaphysics or ontological or epistemological status. There is virtually an infinite set of possible perspectives at this level. Garma C. C. Chang, in *The Buddhist Teaching of Totality* (1971), offers the example of a glass of water. The water is seen by a chemist as H_2O, or a universal solvent. It is seen by a firefighter as something to extinguish flames. It is seen by a thirsty person as something to drink. It is in fact all these things, potentially, though at any given time it may function in one or another way. The problem with this perspective arises when it is universally applied, even in cases when other perspectives seem to conflict with it. Although admittedly a silly example, if a firefighter were dying of thirst but could only see the water as a means of extinguishing fires, then he might die of thirst before he would think to drink the water. An obstinate application of disjunctive perspectives is counterproductive and causes frustration or suffering, the elimination of which is the goal of Buddhism in general.

The second type of perspective is represented by the word *li*, which translates as "rule" or "underlying or abstract principle." In that general sense, *li* is what all *shi* have in common. To shift perspective to the *li* is to resolve all distinctions into some commonality. For example, one can either see coffee and tea as separate things, which would be the level of *shi*, or one can see them as all being water, which is the level of *li*. However, in the case of Hua yan metaphysics, the *li* is *śūnyatā* (emptiness). What all things have in common is that they all lack self-causation or causal autonomy. Everything depends on everything else. The Buddhist texts warn, however, not to ontologize emptiness and make it into a thing. It is the nature of things, which is not a thing in itself. So whereas in the first *dharmadhātu* things are seen as distinct things, in the second they are all seen as empty of self-being.

The third *dharmadhātu* is called *lishi wuai* (nonobstruction of *li* and *shi*). From this perspective, the emptiness of things does not interfere with the thingness of things. This would be experience things as in some sense distinguishable while simultaneously experiencing them as indistinguishably empty.

This does not, however, yet constitute full accomplishment. The final *dharmadhātu* is *shishi wuai* (nonobstruction between phenomena and other phenomena). By realizing that the emptiness of things does not interfere with the thingness of things, one is then able to realize that the specific nature of any one thing does not interfere with the specific nature of any other one thing. As Zongmi says in his commentary to Dushun's text, "all distinct phenomenal dharmas interfuse and penetrate in all ways" (Fox 1988, p. 299). In terms of the example used earlier in the description of the first dharmadhātu, the potability of the water does not interfere with the fire extinguishing properties of the water, which does not interfere with the solvency of water. All these manifestations are all potential manifestations of the same phenomenon. This is how the Buddha sees the world according to the Hua yan tradition, as omnipotentially present in a world of infinitely fractal possibilities. This is a liberation from the fixation on a single, lower-order perspective.

To put this model using modern concepts, one might look at a baseball as a baseball, intended for a certain use in a certain game according to certain rules. One would not be wrong in doing so, but one can also see the baseball as more basically composed of atoms. One would also not be wrong, of course. When one sees the baseball as a baseball, one sees what makes it different from everything else. When one sees the baseball as atoms, one sees what the baseball has in common with everything else, that is, one overlooks the distinctions between things. At the level of the third *dharmadhātu*, one is able to see that the phenomenal and atomic natures of the object do not interfere with each other. It is both atoms and a baseball. Meanwhile, the fourth level encourages one to see the baseball in either its phenomenal or atomic sense as overlapping with every other ostensible object in the universe. This is not far fetched. Phenomenally, one might point out that a baseball would not exist if there was not a game and a population to play it, and so is not entirely separable from those other events. Atomically, one notes that objects share ions with their environments in such a way as to constitute overlapping. It would not even make sense to suggest that an atom could exist in complete isolation, since in fact the atom is made of parts as well, which are made of parts, possibly ad infinitum, as modern string theorists suggest.

Fazang is particularly famous for a couple of metaphors used to demonstrate this principle of nonobstruction and mutual penetration. He is said to have made a huge impression on the empress Wu with these demonstrations, attracting much in the way of imperial support for his writing and translation projects. In one case he is said to have had constructed a room with mirrors on all four walls, as well as in the corners, floor, and ceiling. A torch and statue of the Buddha were placed in the center, and the result was reflections within reflections, each mirror reflecting the other mirrors reflecting itself. This suggested to Fazang a way of explaining how everything can simultaneously be the cause and the effect of everything else. As Chang notes in the *The Buddhist Teaching of Totality*, Fazang is said to have exclaimed that "[t]he principle of the simultaneous arising of different realms is so obvious here that no explanation is necessary" (1971, p. 24). Fazang is also known for using a golden statue of a lion to illustrate a similar principle. Although from one point of view the lion has distinguishable hairs and claws and limbs and teeth, from another point of view the lion is entirely and homogeneously gold.

It is worth pointing out that such an omnicausal model conflates the various types of causal relations that Aristotle, for example, distinguishes, such as efficient, material, final, contiguous, and other types of causal relations. By contrast, the purpose of the model is not to distinguish causal subtleties but to stimulate contextual and perspectival flexibility.

In general, the practice of Hua yan can be described as the attempt to deconstruct one's typically logocentric preoccupation with a fixed perspective, by engaging in a series of exercises that cultivate perspectival flexibility. This is seen to liberate one from the oppression of identifying with a single perspective, which leads to conflict and frustration.

There are many possible parallels between Hua yan thought and Western philosophers and philosophies. For instance, Alfred North Whitehead's process philosophy has been compared to Hua yan's emphasis on the actualization of events out of potentiality, an idea that is also present in modern quantum mechanics. Gestalt and other forms of cognitive psychologies share with Hua yan an emphasis on the importance of perspectival flexibility. In particular, contemporary phenomenological approaches have much in common with Hua yan's concern with the phenomenon qua phenomenon, and both share an emphasis on the importance of experience and perspective that renders metaphysical and absolute statements speculative and counterproductive.

See also Buddhism; Buddhism—Schools: Chan and Zen.

Bibliography

Chang, Garma C. C. *The Buddhist Teaching of Totality: The Philosophy of Hwa Yen Buddhism*. University Park: Pennsylvania State University Press, 1971.

Cleary, Thomas, tr. *Entry into the Inconceivable: An Introduction to Hua-yen Buddhism*. Honolulu: University of Hawaii Press, 1983.

Cleary, Thomas, tr. *Entry into the Realm of Reality: The Guide*. Boulder, CO: Shambhala, 1989.

Cleary, Thomas, tr. *The Flower Ornament Scripture*. 3 vols. Boulder, CO: Shambhala, 1984–1987.

Cook, Francis. *Hua-yen Buddhism: The Jewel Net of Indra*. University Park: Pennsylvania State University Press, 1981.

Fox, Alan. *Elements of Omnicontextual Thought in Chinese Buddhism: Annotated Translations of Gui Feng Zong Mi's Preface to Collection of Various Writings on the Chan Source and His Commentary on "Meditative Approaches to the Hua Yen Dharmadhatu."* Ann Arbor, MI: University Microfilms International, 1988.

Gimello, Robert. *Chih-yen and the Foundations of Hua-Yen Buddhism*. Ann Arbor, MI: University Microfilms International, 1976.

Gimello, Robert, and Peter Gregory, eds. *Studies in Ch'an and Hua-yen*. Honolulu: University of Hawaii Press, 1983.

Gregory, Peter. *Inquiry into the Origin of Humanity*. Honolulu, Hawaii: Kuroda Institute, 1995.

Gregory, Peter. *Tsung-mi and the Sinification of Buddhism*. Princeton, NJ: Princeton University Press, 1991.

Liu, Ming-wood. *The Teaching of Fa-tsang: An Examination of Buddhist Metaphysics*. Ann Arbor, MI: University Microfilms International, 1979.

Odin, Steve. *Process Metaphysics and Hua-Yen Buddhism: A Critical Study of Cumulative Penetration vs. Interpenetration*. Albany: SUNY Press, 1982.

Suzuki, Daisetz Teitaro. *Essays in Zen Buddhism, Third Series*. New York: Samuel Weiser, 1971.

Takakusu, Junjiro. *The Essentials of Buddhist Philosophy*. Delhi, India: Motilal Banarsidass, 1956.

Verdu, Alfonso. *Dialectical Aspects in Buddhist Thought: Studies in Sino-Japanese Mahayana Idealism*. Lawrence: Center for East Asian Studies, University of Kansas, 1974.

Wright, Dale. *Emptiness and Paradox in the Thought of Fa-Tsang*. Ann Arbor, MI: University Microfilms International, 1980.

Alan Fox (2005)

MADHYAMAKA

Madhyamaka is one of the two major schools of Mahāyāna Buddhist philosophy (the other being Yogācāra. It traces its origins to the work of the South Indian philosopher Nāgārjuna (c. 150 CE), who first gave systematic philosophical expression to insights articulated in the earliest Mahāyāna sūtras, the *Prajñāpāramitā* literature. Central to those texts was the claim that all things thought to be ultimately real are in fact "empty" or devoid of intrinsic nature. The Madhyamaka school arose out of efforts to defend this claim and explore its consequences. The Madhyamaka understanding of the concept of emptiness, and the dialectical strategies used to establish its validity, played central roles in the development of Mahāyāna thought in India and subsequently in Tibet and East Asia.

EMPTINESS AS LACK OF INTRINSIC NATURE

When the Mādhyamikas say that all things are empty (*śūnya*), what they mean is that nothing bears an intrinsic nature (*svabhāva*). To understand this claim, one must consider the concept of intrinsic nature as it was developed in the scholastic Abhidharma phase of Buddhist philosophy. It is a basic teaching of Buddhism that suffering is caused by one's ignorance of the truth of nonself: that one does not have a separately existing self and that what one thinks of as an enduring person just consists in a causal series of impermanent, impersonal physical and mental events. Philosophers of the Abhidharma schools sought to buttress this conclusion by arguing that all partite entities (wholes made up of parts) are conceptually constructed and thus not ultimately real. This would enable them to claim that the person is conceptually constructed out of the psychophysical elements making up a causal series and so is not itself objectively real.

The general argument is that a partite thing such as a chariot borrows all its properties from the properties of its parts: There is no fact about a chariot that cannot be explained strictly in terms of facts about its parts and their relations. This is taken to show that positing the chariot as an additional entity is superfluous, something one is inclined to do only because of facts about one's interests and cognitive limitations. Since this holds as well for the person, as a whole made up of the elements in a causal series, it follows that one's view of oneself as an enduring substance reflects a failure to distinguish between a mere useful fiction and what is ultimately or mind-independently real.

This reductionist line of thought in the Abhidharma rests on the assumption that there are entities that are ultimately real. To say that persons and other partite things are not ultimately real because they are conceptually constructed is to assume that there are those ultimately real things out of which partite things are constructed. Now conceptually constructed things were held to borrow their properties from the properties of their parts. So Abhidharma thinkers concluded that ultimately real things must have their natures intrinsically. Only that is ultimately real, they claimed, that "is found

under analysis," that is conceptually irreducible. The Madhyamaka claim that all things are empty is meant to contest the Abhidharma view that there could be such things. Through a wide variety of arguments the Mādhyamikas seek to demonstrate the absurd consequences that would follow if it were held that there are entities with intrinsic nature.

THE ARGUMENT FROM CAUSATION

One such argument concerns the causal relation. It is a fact of one's experience that existing things are impermanent, and this would seem to hold for whatever is ultimately real. But it is also a fact of one's experience that things do not come into or go out of existence in an utterly random way. There seem instead to be patterns of regular succession. So an adequate account of the nature of reality seems to require that ultimately real things be said to arise and cease in accordance with causal laws. At this point the Mādhyamikas raise a simple question: Are cause and effect identical or distinct? Consider the first possibility. Certain Indian philosophers held that the effect is identical with the cause—that causation represents just the manifestation of what already exists in the cause in unmanifest form. But this view is readily dismissed. For it requires that there already exist something with the intrinsic nature of the effect before the effect is produced. And in that case one must wonder why one would set about trying to produce the effect. One might build a fire because one is cold and wants the heat of fire. But if the fire already existed in its cause, the fuel, then its heat should already be present there, so it would be pointless to build a fire.

If, on the contrary, cause and effect were distinct things, two difficulties would follow. First, if these are genuinely distinct things, some account must be given as to why things of the first sort regularly give rise to things of the second sort. Why should fuel give rise to fire and not, say, to cheese? The stock answer to this question is that fuel possesses the causal power to produce fire. But now it must be asked whether this causal power is a third thing that is distinct from both cause and effect or is rather identical with one or the other. If it is distinct from the cause, one may then ask why this sort of cause should be conjoined with just this sort of causal power. This quickly leads to an infinite regress: A second causal power will be required to account for the occurrence of the first, a third for that of the second, and so on. But if the causal power is identical with the cause, then no answer has been given to the original question, and likewise if the causal power is identical with the effect.

The second problem for the view that cause and effect are distinct things is that it is then unclear when the cause produces the effect. To call one thing the cause of another is to say that the first produces the second, so surely there must be some time when this productive activity takes place. There are three possibilities here: when the effect already exists, when the effect does not yet exist, and when the effect is coming into existence. The first is clearly ruled out, since production of something that already exists would be redundant. The second is likewise wrong, for something may be said to be productive only if there is some actually existing product. And with respect to things that are ultimately real, there could be no third time during which they are coming into existence. With respect to partite things like chariots it makes sense to speak of a process of assembly during which the entity is undergoing production. But this is possible only because the chariot is made of parts. Something impartite that bore its nature intrinsically could only be said to be either existent or nonexistent; a third intermediate time is ruled out for such a thing. The upshot of all this is that it appears impossible to account for the causal relations that should obtain among things with intrinsic natures.

THE ARGUMENT FROM THE PROPERTY-BEARER RELATION

A second Madhyamaka argument for emptiness involves examining the relation between an ultimately real thing and its intrinsic nature. Either these are distinct or they are identical. If they are distinct, a number of difficulties follow. First, there is the problem of saying what the entity itself is like apart from its intrinsic nature. Since the notion of a pure propertyless substrate seems incoherent, this problem is likely to prove intractable. But there is also the difficulty that then the entity's acquiring its nature will depend on causes and conditions. Such dependence seems incompatible with calling its nature intrinsic; it then seems more appropriate to say that the entity borrows its nature from other things.

Suppose then that the entity and its intrinsic nature are identical—that one's distinction between the thing and its nature merely reflects the concepts one uses. In that case an occurrence of what one calls fire is really just the occurrence of heat (the property of being hot). But then the question arises how fires are to be individuated. Suppose there are two distinct fires of equal intensity. Each fire is just its heat, and the two heats are identical in nature. Ordinarily, one would say that the two occurrences of heat are distinct because each occurs in a distinct particular (the fire whose heat it is). But on the

hypothesis under scrutiny there are no particulars over and above the property of heat; the occurrence of what one judges to be a particular fire just is the occurrence of heat. One might then suppose that each fire is individuated in terms of the discrete space that it occupies. But then the question arises what makes two spaces discrete? Suppose the intrinsic nature of a space is its nonresistance. Since one is now supposing that the existence of a space just is the occurrence of a certain nonresistance, it is not clear what will make two spaces distinct, unless it is their being occupied by distinct entities, such as two fires. But now one has come full circle. So it looks as if the hypothesis that entity and intrinsic nature are identical does not hold up to critical scrutiny either. It appears that no adequate account can be given of how something could have an intrinsic nature.

MADHYAMAKA AS NIHILISM

A host of similar arguments against things with intrinsic nature was developed by Mādhyamika philosophers such as Nāgārjuna, Āryadeva (170–270), Buddhapālita (c. 500), Bhāvaviveka (500–570), and Candrakīrti (c. 600–650). Nāgārjuna's targets were chiefly views held by Ābhidharmikas, but Āryadeva extended the scope of Mādhyamika dialectics to include the views of non-Buddhist Indian philosophers, a practice that becomes systematic in Bhāvaviveka's *Tarkajvālā*. Suppose that these arguments succeed in showing that nothing could bear an intrinsic nature. Suppose also that the Ābhidharmikas were correct in concluding that only something with an intrinsic nature could be ultimately real. What would then follow? What should one make of the Mādhyamika doctrine of emptiness? Modern scholars have put forward a wide variety of interpretations, but there is also some difference of opinion among classical Indian authors. One modern reading, that of Thomas E. Wood (1994) and David Burton (1999), that is also the common view of the Mādhyamikas' ancient Indian critics is that the doctrine of emptiness is tantamount to metaphysical nihilism, the thesis that reality is ultimately devoid of existing things. The stock characterization of the Mādhyamikas that one finds in the writings of their classical opponents is that the Mādhyamikas believe nothing whatever exists.

Of course, the thesis of metaphysical nihilism is virtually self-refuting: If nothing whatever existed, the thought could not occur that it might be true. Still, attributing this thesis to the Mādhyamikas might not seem unfair. If there is reason to believe that only things with intrinsic nature could be ultimately real, then demonstrating the incoherence of the concept of a thing with intrinsic nature seems equivalent to showing that ultimately nothing whatever is real. One major difficulty with this interpretation, however, is that it is explicitly argued against by the Mādhyamikas. Thus, Nāgārjuna points out that to understand the thesis that no ultimately real things exist, one must first understand what it would mean for there to be ultimately real things. But an ultimately real thing would have to be something with intrinsic nature. Since the Mādhyamikas claim there can be no such things, they would say one cannot understand the thesis that ultimately nothing exists. So perhaps they should not be interpreted as seeking to establish metaphysical nihilism.

DO MĀDHYAMIKAS AFFIRM CONTRADICTIONS?

Of the remaining interpretations of emptiness found in the modern scholarship, only some find support in the original sources. (Of course, the lack of such support need not detract from the philosophical significance of an interpretation.) For instance, Graham Priest and Jay L. Garfield (2002) claim Nāgārjuna as perhaps the first exponent of dialetheism, the view that there are true contradictions that arise at the limits of thought. As evidence, they cite his assertion that it cannot be ultimately true that all things are empty (*Madhyamakakārikā* chapter 22, verse 11). The notion of ultimate truth at work here derives from the Abhidharma distinction between two kinds of truth: conventional and ultimate. Only statements concerning ultimately real things can be said to be ultimately true; statements concerning such mere conceptual fictions as chariots and persons can only be conventionally true. For Abhidharma, then, the set of ultimately true statements would give the complete account of all those things with intrinsic natures; it would be a complete description of the ultimate nature of reality.

Now the Mādhyamikas claim to have shown that the only statement that can truly be made about those things that are thought to be ultimately real is that they are empty. But in the verse in question Nāgārjuna says that this cannot be ultimately true. Indeed, he says it is not ultimately true that all things are empty, or that they are nonempty, or both or neither. The reason for this is that emptiness is itself a mere conceptual fiction. So any statement about emptiness could at best be conventionally true. Priest and Garfield take Nāgārjuna to be thereby asserting both that the ultimate truth cannot be characterized and that it can be characterized (namely as being

uncharacterizable). But this is not how Mādhyamika commentators have understood the verse. Instead, they assimilate it to the Buddha's treatment of the so-called indeterminate questions (*avyākṛta*). When, for instance, the Buddha was asked whether the enlightened person survives death, does not survive death, both survives and does not survive, or neither survives nor does not survive death, the Buddha rejected all four possibilities. One can consistently do this, they explain, because all share an implicit presupposition—that there ultimately is such a thing as an enlightened person—and this presupposition should be rejected. By the same token, the commentators say, Nāgārjuna should be understood as rejecting the presupposition that there is such a thing as the ultimate truth. In that case he asserts neither that the ultimate truth is uncharacterizable nor that it can be characterized. He does not hold that a contradiction is true.

MADHYAMAKA AS SKEPTICISM

Other interpreters of the doctrine of emptiness, such as Thomas McEvilley (1982) and Bimal Krishna Matilal (1986), see it as a form of skepticism. This reading is suggested by the Mādhyamika response to objections coming from Indian epistemologists. The thrust of these objections is that since the Mādhyamikas hold all things to be empty, they must hold that all means of knowledge are empty. But in that case it cannot be known that all things are empty, so the Mādhyamika claim is a mere dogmatic assertion. Part of the Mādhyamika response involves calling into question the epistemologist's project of determining which are the means of knowledge. For instance, they argue that a given procedure can be known to be a means of knowledge—a reliable cause of veridical belief—only if one already possesses some means of knowing which beliefs are true. Thus, any attempt to determine which are the means of knowledge either is circular or else leads to an infinite regress.

An argument of this sort might be used to support the skeptic's claim that one can never know which, if any, of one's beliefs amount to knowledge. But this is not how the Mādhyamikas themselves see such arguments. For one thing, the skeptical conclusion requires the additional assumption that one can only know some statement p if one knows that one knows p—an assumption that neither the Mādhyamikas nor their opponents seem to have held. Second, nowhere do the Mādhyamikas appeal to the sorts of error possibilities that are the skeptic's stock in trade, such as perceptual illusions, hallucinations, dreams, and the like. Indeed, the Mādhyamikas do not deny that, conventionally, certain procedures can count as means of knowledge. What they deny is just that anything could ultimately be a means of knowledge, that anything could have the intrinsic character of reliably causing veridical beliefs as part of its mind-independent essential nature. The Mādhyamika epistemological stance seems to be that something can be a means of knowledge only through its relations to other things that are themselves equally empty of intrinsic nature. The resulting view may have its affinities with some forms of skepticism (particularly Pyrrhonian skepticism). But its chief concern is not to call into question the possibility of knowledge, but to deflate the pretensions of a certain sort of epistemological realism.

THE MĀDHYAMIKAS AS MYSTICS OR AS QUIETISTS?

Two interpretations of emptiness seem more firmly grounded in the self-understanding of the Mādhyamika tradition. The first sees emptiness as leading to a kind of mystical silence. Mādhyamika arguments are said to demonstrate that no set of concepts can ever adequately represent the world. This realization is said to then usher in a nondiscursive grasping of the nature of reality (perhaps through a kind of intuition that is cultivated in meditation). On this interpretation, emptiness serves to point to an ultimate reality that lies beyond the reach of philosophical rationality. The second of the two, by contrast, sees emptiness not as pointing to an ineffable ultimate, but as indicating that the very idea of an ultimate nature of reality is incoherent. The exercise of philosophical rationality leads not to the silence of the beyond, but back to the conventional. For Mādhyamika dialectic reveals the error in the notion of an ultimate truth that represents how things are independently of all facts about the cognizer. This shows that truth can only be transactional, a matter of what facilitates interactions among creatures like us. The notion of a truth that potentially outstrips all our conceptual resources is revealed to be no more than a useful fiction.

The "mystical silence" reading of emptiness has been championed by T. R. V. Murti (1955) and David Seyfort Ruegg (1977) among others. The second reading is commonly called a quietist interpretation, since it grows out of the attempt by Frederick J. Streng (1967) to read elements of the later Ludwig Josef Johann Wittgenstein into Nāgārjuna. But as developed by Tom J. F. Tillemans (2002), it has clear affinities with both antirealist and minimalist conceptions of truth. Both readings may be seen as seeking to explicate the claim that insight into emptiness results in a kind of nondual awareness.

Mahāyāna Buddhist texts commonly claim that final liberation from suffering requires a kind of seeing that transcends all problematic dualities. On the mystical silence interpretation it is the dualism fostered by conceptualization that is to be overcome, for without concepts one cannot make such invidious distinctions as that between cognizing subject and object. On the quietist reading it is the dualism of ultimate and conventional truth that is erased through knowledge of emptiness. Presumably, this duality is problematic because the notion of ultimate truth as correspondence to mind-independent reality fosters a subtle form of belief in a self, namely that expressed through attachment to metaphysical theories.

Each reading is not without its own difficulties. For the mystical silence view, the chief problem is to explain how Madhyamaka then differs from other views that posit an ineffable ultimate, such as the absolute monisms of Advaita Vedānta and Parmenides of Elea. For the quietist there is the difficulty of explaining how there can be truth without there being such a thing as how the world is anyway. This problem is sometimes addressed by claiming that what emptiness really shows is just that no entity has a nature that is independent of its relations to other things. But to the extent that this addresses the problem of grounding truth in mind-independent reality, it contravenes the quietist claim to be showing a way out of the trap of metaphysical theories.

THE SVĀTANTRIKA-PRĀSAṄGIKA DISTINCTION

Modern studies of the Mādhyamikas have profited enormously from contact with the Tibetan tradition, a tradition for which Mādhyamika thought continues to play a crucial role. But there are cases in which reliance on Tibetan doxographical categories has led to distortion of the Indian Mādhyamika sources. A case in point is the alleged distinction between two schools of Madhyamaka: Svātantrika and Prāsaṅgika. This distinction was invented by Tibetan doxographers, and it is a matter of some dispute to what extent it reflects substantive differences in the views of the Indian thinkers covered by the classification. It is in any event clear that Indian Mādhyamikas did not see themselves as falling into two camps to which these labels could be applied.

Those who accept the distinction identify a dispute between Bhāvaviveka and Candrakīrti as its point of origin. The dispute concerns the proper methodology for a Mādhyamika. The arguments of Bhāvaviveka's Mādhyamika predecessors were usually expressed in the *reductio ad absurdum* (*prasaṅga*) style: The hypothesis to be refuted (e.g., that something with intrinsic nature could be an effect) is considered and then shown to lead to some absurd result (e.g., that its intrinsic nature is actually extrinsic). Employing the methods of the Buddhist logician Diṅnāga, Bhāvaviveka sought to convert such *reductios* into independent arguments (*svatantra anumāna*). Thus, one would have:

> It is not the case that ultimately an entity arises from distinct causes and conditions.Because of depending on them for its nature.

Whatever depends on other things for its nature is not ultimately real, like the chariot. Candrakīrti disagrees, claiming that the Mādhyamikas may only use *reductios*. But since the two types of argument turn out to be formally equivalent once the *reductio* has been fully spelled out, it may not be clear what the dispute is actually about.

The difference Candrakīrti sees between them is this: In giving a *reductio* one need not assert anything to be the case oneself; the proponent merely shows the opponent the inconsistency in his or her view, thereby impelling the opponent to withdraw assent from his or her thesis. In the case of an independent argument, on the contrary, both the proponent and opponent must agree about such things as the subject (in this case, an entity), the pervasion (that what is dependently originated is not ultimately real), and the example (the chariot). But the Mādhyamika proponent holds that entities can only exist conventionally, while the opponent thinks some entities are ultimately real, so the two sides do not agree about the subject. And likewise for the other elements of the argument that require consensus. From the perspective of the Mādhyamikas, the opponent is simply, hopelessly wrong about everything. So there can be no common framework for resolving their disagreement. Instead, the Mādhyamikas should just give their opponents the rope with which to hang themselves.

SYNCRETISM IN MADHYAMAKA

One may wonder if the opponent will be so obliging toward a proponent who seems to speak a different (and perhaps unintelligible) language. But there may be a deeper point here. Those Tibetan commentators such as Tsong-kha-pa (1357–1419) who align themselves with the Prāsaṅgika allege that Svātantrikas have not fully realized emptiness, since they continue to posit intrinsic natures, albeit at just the conventional level. And it is true that those Mādhyamikas who are identified as Svātantrikas exhibit a tendency toward syncretism, seeking to incorporate the views of overtly metaphysical Buddhist schools within an overall Mādhyamika framework.

This tendency is especially clear in Śāntarakṣita (eighth century), who embraces Dharmakīrti's formulation of the Yogācāra school's subjective idealist ontology and epistemology. But it can already be seen in Bhāvaviveka, who champions the Sautrāntika school's realism about physical objects and its associated representationalist theory of sense perception. In neither case is the other school's view identified as anything more than the best way of representing conventional truth. As Mādhyamikas, Śāntarakṣita and Bhāvaviveka remain committed to the position that the only ultimate truth is that there is no ultimate truth (i.e., that all things, including emptiness, are empty). Still, they do take a position on the question whether external objects exist conventionally.

Candrakīrti does as well. He sides with Bhāvaviveka in rejecting the idealist view. But his reasons are different. Where Bhāvaviveka tries to answer Yogācāra arguments against the existence of physical objects, Candrakīrti simply dismisses the arguments. For him such arguments can only show that physical objects are ultimately empty—something a Mādhyamika already knows. But by the same token mental states (which the idealist thinks are real) are equally empty. So the availability of philosophical arguments against the conventional belief in external objects cannot show that they are not conventionally real. While Bhāvaviveka thinks the use of philosophical rationality can lead to improvements in one's conventional account of the world, Candrakīrti thinks it can only lead to the ultimate truth of emptiness. Conventional truth neither needs nor can it sustain either refinement or defense at the hands of philosophers. It is just simply that which is given through the everyday practices of ordinary people.

Given this difference in attitude, one can see why Svātantrikas might be described by their critics as positing conventionally real intrinsic natures. It is, after all, philosophical analysis that gives rise to the demand for things with intrinsic nature. So if philosophical rationality is allowed to play a role in shaping one's conventional worldview, the resulting theory will be committed to there being such things, the things at which analysis stops. And to Candrakīrti, Bhāvaviveka's demand that the Mādhyamikas give independent arguments and not mere *reductios* looks like a requirement that the Mādhyamikas construct a philosophically defensible version of the conventional truth. This will inevitably lead in the direction of syncretism, and with it the danger that the Mādhyamikas will become ensnared in metaphysical theories. The Prāsaṅgika side in this dispute is not without its dangers as well though. For on its account, conventional truth does not allow of progressive improvement, it can only be utterly overthrown. The result would seem to be a strong form of relativism about conventional truth. And an opponent could always use this to turn back the Prāsaṅgika's *reductio* arguments, in effect saying to the Mādhyamikas, "We simply disagree about whether there is an inconsistency in my position, and in such matters there is no right and wrong." What this dispute brings out, then, is a tension that seems inherent in the concept of truth, a tension that is also reflected in current debates between realists and antirealists.

Indian Madhyamaka came to an end in the late twelfth century, when all Buddhist philosophical activity ceased in India following the Turkish invasion. Madhyamaka has continued to play a prominent role in Tibetan Buddhism to this day. It also enjoyed some popularity among Chinese Buddhist philosophers, playing an important role in the development of the Huayan school. Perhaps a case might even be made for its having had a profound impact on Chan Buddhism. Chan formally eschews the study of precisely those sorts of doctrinal texts that form the core of Mādhyamika practice. But it does make extensive use of paradox in some of the methods it has devised for helping the adept attain enlightenment. Analysis of the structural features of those paradoxes and their uses might reveal more than merely superficial resemblances with the dialectical strategies of Madhyamaka.

See also Buddhism; Buddhism—Schools: Dge-lugs.

Bibliography

PRIMARY SOURCES AND TRANSLATIONS

Āryadeva. "Catuḥśataka." In *Āryadeva's Catuḥśataka*, edited and translated by Karen Lang. Copenhagen: Akademisk Vorlag, 1986.

Bhāvaviveka. *Madhyamakahṛdayakārikā Tarkajvālā*. In *Reason and Emptiness: A Study in Logic and Mysticism*. Translated by Iida Shōtarō. Tokyo. Hokuseido Press, 1980.

Nāgārjuna. "Mūlamadhyamakakārikā." In *Mūlamadhyamakakārikās (Mādhyamikasūtras) de Nāgārjuna avec la Prasannapadā, Commentaire de Candrakīrti*, edited by Louis de La Vallée Poussin. Osnabrück, Germany: Biblio Verlag, 1970. Also in *The Fundamental Wisdom of the Middle Way: Nāgārjuna Mūlamadhyamakakārikā*. Translated by Jay L. Garfield. New York: Oxford University Press, 1995.

Nāgārjuna. "Vigrahavyāvartanī." In *The Dialectical Method of Nāgārjuna: Vigrahavyāvartinī*. Translated by Kamaleswar Bhattacharya and edited by E. H. Johnston and Arnold Kunst. Delhi: Motilal Banarsidass, 1978.

Śāntarakṣita. "Tattvasaṅgraha." In *The Tattvasaṅgraha of Śāntarakṣita, with the Commentary of Kamalaśīla*, edited by Embar Krishnamacharya. Baroda, India: Oriental Institute,

1984. Also in *The Tattvasaṅgraha of Śāntarakṣita, with the Commentary of Kamalaśīla*. Translated by Ganganatha Jha. Delhi: Motilal Banarsidass, 1986.

Śāntideva. "Bodhicāryāvatāra." In *The Bodhicāryāvatāra of Ārya Śāntideva with the Commentary Pañjika of Śrī Prajñākaramati*, edited by Dwārikādās Śāstrī. Varanasi, India: Bauddha Bharati, 2001. Also in *The Bodhicaryāvatāra*. Translated by Kate Crosby and Andrew Skilton. New York: Oxford University Press, 1993.

SECONDARY LITERATURE

Burton, David. *Emptiness Appraised: A Critical Study of Nāgārjuna's Philosophy*. Richmond, U.K.: Curzon, 1999.

Dreyfus, Georges B. J., and Sara McClintock, eds. *The Svātantrika-Prāsaṅgika Distinction: What Difference Does a Difference Make?* Boston: Wisdom Publications, 2002.

Eckel, Malcolm David. *Jñānagarbha's Commentary on the Distinction between the Two Truths: An Eighth Century Handbook of Madhyamaka Philosophy*. Albany, NY: SUNY Press, 1987.

Hayes, Richard P. "Nāgārjuna's Appeal." *Journal of Indian Philosophy* 22 (1994): 299–378.

Huntington, C. W., Jr., with Geshé Namgyal Wangchen. *The Emptiness of Emptiness: An Introduction to Early Indian Mādhyamaka*. Honolulu: University of Hawaii Press, 1989.

Lopez, Donald S., Jr. *A Study of Svātantrika*. Ithaca, NY: Snow Lion, 1987.

Matilal, Bimal Krishna. *Perception: An Essay on Classical Indian Theories of Knowledge*. New York: Oxford University Press, 1986.

McEvilley, Thomas. "Pyrrhonism and Mādhyamika." *Philosophy East and West* 32 (1) (1982): 3–35.

Murti, T. R. V. *The Central Philosophy of Buddhism: A Study of the Mādhyamika System*. London: Allen and Unwin, 1955.

Oetke, Claus. "Remarks on the Interpretation of Nāgārjuna's Philosophy." *Journal of Indian Philosophy* 19 (3) (1991): 315–323.

Potter, Karl H., ed. *Encyclopedia of Indian Philosophies, Vol. 8: Buddhist Philosophy from 100 to 350 A.D.* Delhi: Motilal Banarsidass, 1999.

Potter, Karl H., ed. *Encyclopedia of Indian Philosophies, Vol. 9: Buddhist Philosophy from 350 to 600 A.D.* Delhi: Motilal Banarsidass, 2003.

Priest, Graham, and Jay L. Garfield. "Nāgārjuna and the Limits of Thought." In *Beyond the Limits of Thought*, edited by Graham Priest. New York: Oxford University Press, 2002.

Robinson, Richard H. *Early Mādhyamika in India and China*. Madison: University of Wisconsin Press, 1967.

Ruegg, David Seyfort. "The Uses of the Four Positions of the Catuṣkoṭi and the Problem of the Description of Reality in Mahāyāna Buddhism." *Journal of Indian Philosophy* 5 (1977): 1–71.

Ruegg, David Seyfort. *The Literature of the Madhyamaka School of Philosophy in India*. Wiesbaden, Germany: Harrassowitz, 1981.

Siderits, Mark. "The Madhyamaka Critique of Epistemology I." *Journal of Indian Philosophy* 8 (1980): 307–335.

Siderits, Mark. "Thinking on Empty: Madhyamaka Anti-realism and Canons of Rationality." In *Rationality in Question: On Eastern and Western Views of Rationality*, edited by Shlomo Biderman and Ben-Ami Scharfstein. Leiden, Netherlands: Brill, 1989

Streng, Frederick J. *Emptiness: A Study in Religious Meaning*. Nashville, T.N.: Abingdon Press, 1967.

Tillemans, Tom J. F. *Materials for the Study of Āryadeva, Dharmapāla, and Candrakīrti: The Catuḥśataka of Āryadeva, chapters XII and XIII, with the commentaries of Dharmapāla and Candrakīrti*. 2 vols. Vienna: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien, 1990.

Tillemans, Tom. "Metaphysics for Mādhyamikas." In *The Svātantrika-Prāsaṅgika Distinction: What Difference Does a Difference Make?* edited by Georges B. J. Dreyfus and Sara McClintock. Boston: Wisdom Publications, 2002.

Williams, Paul. "On the Interpretation of Madhyamaka Thought." *Journal of Indian Philosophy* 19 (1991): 191–218.

Wood, Thomas E. *Nāgārjunian Disputations: A Philosophical Journey through an Indian Looking-Glass*. Honolulu: University of Hawaii Press, 1994.

Mark Siderits (2005)

YOGĀCĀRA

The origins of Yogācāra Buddhism—one of the two mainstream schools of Indian Mahāyāna Buddhism (the other being Madhyamaka)—are obscure, but tradition credits its founding to two half-brothers, Asaṅga and Vasubandhu (c. fourth century). Many of Yogācāra's distinctive terms and models, such as eight consciousnesses, three self-natures, and *vijñapti-mātra* (explained later in this entry), had already appeared in certain Mahāyāna scriptures such as the *Saṅdhinirmocana Sūtra* (Sūtra elucidating the hidden connections), but the expansion of those ideas in the prolific writings of Asaṅga and Vasubandhu gave the school its foundation. Yogācāra attempted to absorb the full range of teachings and practices that had arisen over the centuries since the time of the Buddha—from the elaborate scholastic schemas of the Abhidharma schools to the reformulation of Buddhist doctrine in terms of emptiness (*śūnyatā*) posed by early Mahāyāna literature—to fashion a detailed, systematic, coherent, step-by-step path to unsurpassable complete awakening (*anuttarā-samyak-saṃbodhi*) and nirvāṇa.

Since Buddhism identified the root problem as ignorance, Yogācāra devised methods for uncovering and correcting the cognitive errors inherent in the way the mind works. Yogācāra's sustained attention to cognitive issues such as consciousness, perception, psycholinguistic conditioning, and epistemology, coupled with claims such as "external objects do not exist," has led some to misinterpret Yogācāra as a form of metaphysical idealism. For Yogācāra, however, consciousness is not an eternal sub-

stance or immutable substrate. Rather, individual consciousnesses arise and cease each moment because of everchanging causes and conditions, these discrete moments of consciousness linked in sequential causal chains, giving the illusion of a continuous self-identity or selfhood. Overcoming ignorance meant first eliminating this false view of self and subsequently seeing things as they truly are. Yogācāra developed perhaps the most sophisticated examination and description in all of Buddhism of how the mind works—in psychological, epistemological, logical, emotional, cognitive, meditative, developmental, and soteric modes.

HISTORICAL OVERVIEW

Though the historical details of the lives of the early Yogācāras have been obscured by later legends—some so unreliable that a few scholars swayed by miscues theorized that tradition had conflated two different Vasubandhus who lived at different times, a theory no longer accepted—enough of their prolific writings has survived (though not always in the original Sanskrit) for us to appreciate the depth and complexity of their thinking. Legend holds that Asaṅga, after years of fruitless meditation, was about to give up when the future Buddha, Maitreya, appeared to him, instructing him in hitherto unknown scriptures. Some of his writings are ascribed to Maitreya, others to Asaṅga himself (the Chinese and Tibetan traditions differ on which texts to ascribe to which). Most important among his works are the encyclopedic *Yogācārabhūmi* (Stages of yoga practice), *Mahāyānasaṃgraha* (Compendium on Mahāyāna), *Abhidharmasamuccaya* (Abhidharma compilation), and *Madhyānta-vibhāga* (Differentiating the middle from the extremes). Vasubandhu at first studied Vaibhāṣika Buddhism at its headquarters in Kashmir, composing a detailed summary of its doctrines titled *Abhidharmakośa* (Treasury of Abhidharma). Under Asaṅga's influence, Vasubandhu became a Yogācāra, adding a commentary to his *Kośa* that critiqued the Vaibhāṣika positions, incorporating ideas and phraseology found in Asaṅga's works. Along with many commentaries on Mahāyāna scriptures (most no longer extant), his most important works are his commentary on *Madhyānta-vibhāga*, *Triṃsikā* (Thirty verses), *Viṃśatikā* (Twenty verses), and four of the earliest Buddhist treatises on logic.

Yogācāra subsequently split into two wings: (1) an Abhidharmic wing primarily engaged in commentary writing and doctrinal exposition, its main figures being Sthiramati (sixth century), Dharmapāla (sixth century), and Xuanzang (seventh century); and (2) an epistemological-logic wing that for centuries became the vanguard of Indian epistemology and logic, its main figures including Dignāga (fifth century), Dharmakīrti (seventh century), and Ratnakīrti (c. eleventh century). By the seventh century Yogācāra dominated the leading Indian Buddhist centers at Nālandā and Vālabhi. Texts like the *Laṅkāvatāra Sūtra* blended Yogācāra with *tathāgatagarbha* (Buddhahood-potentiality) thought.

Translators introduced Yogācāra writings to China in the early fifth century, where it dominated for the following two centuries. It became influential in Korea and Japan in the seventh century, and though it eventually was overshadowed by other forms of East Asian Buddhism that themselves were offshoots of Yogācāra-*tathāgatagarbha* hybrids, periodically East Asians have renewed interest in Yogācāra. Yogācāra was also influential during the formative years of Tibetan Buddhism, and has remained part of the Tibetan Buddhist curriculum until the present.

VIJÑAPTI-MĀTRA

The core of Yogācāra doctrine is expressed by the term *vijñapti-mātra*, usually translated "consciousness-only" or "representation-only." Despite repeated strenuous denials in Yogācāra texts that *vijñapti-mātra* means that only consciousness exists or that consciousness alone is real, over the centuries many non-Yogācāras have interpreted it that way. Since consciousness (*vijñāna*) is the domain in which all contemplation, examination, theorization, and knowledge about reality occurs, its facticity is undeniable, though, for Yogācāra, consciousness is neither ultimate reality nor the solution to life's problems. Rather, consciousness itself is the problem. The grammatically causative term *vijñapti* means "what makes known," signifying that consciousness actively constructs the appearances it apprehends and appropriates. Since to appear within a perception or cognition means to appear within an act of consciousness, we are usually not directly aware of anything outside of consciousness. Whatever one is aware of or thinks about necessarily occurs to one only within consciousness. *Vijñapti-mātra* means that we confuse our imaginary projections for the world itself. Since this confusion pervades ordinary mental operations, it ends only when those operations cease.

Yogācāra explains that each individual is a distinct consciousness stream or mental continuum (*citta-santāna*) that, like a river, changes moment by moment in accord with causes and conditions, giving the illusion of a continuity of identity despite the perpetual reconfiguring of the water. It has no fixed, invariant identity or self.

The stream flows from moment to moment and from life to life. These distinct consciousness streams can affect and communicate with other streams, learning from and teaching each other, and mutually influencing each other. Hence, Yogācāra rejects both solipsism and the idea of an overarching universal mind. If, as solipsists claim, each mind is closed off from others, how could Buddhas and others teach anyone anything? If we cannot learn from others, then Buddhism itself becomes superfluous and untenable. If everything shares a single mind, then that mind would have to be either deluded or enlightened. If deluded, then enlightenment for individuals would be impossible; if enlightened, then either unenlightened individuals should be impossible, or else they are already enlightened just as they are, which again would render individual practice and Buddhism meaningless.

Vijñapti-mātra is not the denial of anything real outside an individual mind. Even *rūpa* (sensorial materiality) is accepted, since physicality is known through the senses and cognition; sensations should not be confused, however, with abstract concepts or theories about materiality. Sensation is real (Asaṅga calls the five senses pure); conceptualization is not real in the same way, especially when it imports notions of selfhood or substantialism, or posits appropriational entities. That would be the sort of error the term *vijñapti-mātra* is designed to caution against.

Everything we know, conceive, imagine, or are aware of, we know through cognition, including the notion that entities might exist independent of our cognition. Cognitive objects appear within acts consciousness; Yogācāra never denies that. By definition, they cannot appear elsewhere. What Yogācāra denies in the term *external object* is the concept of externality, not the object itself. Although the mind does not create the physical world, it generates the interpretative categories through which we know, classify, and interact with the physical world, and it does this so seamlessly that we mistake our interpretations for the world itself. Those interpretations—conditioned conventionalisms expressed as desires, preferences, and anxieties—become obstructions (*āvaraṇa*) that prevent us from seeing what is actually the case. In simple terms, we are blinded by our own self-interests, our own prejudices, our desires. We think like others because we have undergone similar conditioning and reinforce that conditioning by congregating with those who are like-minded. Such consensus is bred from tautology, not universality.

Unenlightened cognition is an appropriative act. Yogācāra texts do not speak about subjects and objects; instead, perception is analyzed in terms of sentient beings and cognitive fields, or, more often, graspers (*grāhaka*) and what is grasped (*grāhya*). The Buddhist notion of karma is intimately connected to the notion of appropriation (*upādāna*). As explained in the earliest Buddhist texts, suffering and ignorance are produced by karma. Karma is defined in Buddhism as any intentional activity of body, speech, or mind. Intention is the crucial factor, and, since intention is a cognitive condition, whatever is noncognitive must be also nonkarmic and nonintentional. Thus, by definition, whatever is noncognitive can have no karmic implications or consequences. Intention means to direct one's attention toward some thing or purpose, to desire something. Physically, linguistically, or mentally, we try to "get it."

Put another way, only cognitive acts can have karmic repercussions. Cognitive acts include meaningful bodily gestures that communicate intentions (such gestures are also called *vijñapti*). Thus, to overcome ignorance and suffering by eliminating karmic conditioning, Buddhists need only focus on what occurs within the domain of cognitive conditions (*citta-gocara*).

Categories such as external object and materiality (*rūpa*) are cognitive constructions. *Materiality* is a word for the colors, textures, sounds, and so on that we cognize in acts of perception, and it is only to the extent that they are perceptually apprehended and ideologically grasped, thereby becoming objects of attachment, that they have karmic significance. Materialism is not the problem. There is nothing intrinsically good or bad about gold, for example; rather, our ideas about gold's value and uses, which we project and then act upon, lead to good or bad consequences. The incessant propensity (*anuśaya*) to appropriate what consciousness projects is the problem. These projections are not just things, but moral qualities, status, ideals, religious and national doctrines and identities, the holding of opinions, whatever we can make our own, or make ourselves to be. For Buddhism, attachment to ideas and theories is much deeper and more problematic than attachment to physical things, since the latter is rooted in and merely an expression of the former.

A deceptive trick is built into the way consciousness operates at every moment. Consciousness constructs a cognitive object in such a way that it disowns its own creation, pretending the object is "out there," to render that object capable of being appropriated. Even while what we cognize is occurring within our act of cognition, we cognize it as if it were external to our consciousness. This is called *abhūta-parikalpa*, imagining something exists in a locus in which it is absent. Realizing *vijñapti-mātra* means exposing this trick at play in every act of consciousness, catching it in the act, as it were, and thereby

eliminating it. Consciousness engages in this deceptive game of projection, dissociation, and appropriation because there is no self. The most deep-seated erroneous view to which sentient beings cling, according to Buddhism, is *ātmadṛṣṭi*, the view that a permanent, eternal, immutable, independent self exists. No such self exists, and deep down we know that. This makes us anxious, since it entails that no self or identity endures forever. To alleviate that anxiety, we attempt to construct a self, to fill the anxious void, to do or acquire something enduring. Projecting objects and ideas that one can appropriate and cling to is the way consciousness contributes to this project. If I own things (e.g., ideas, theories, identities, and material objects), then "I am." If there are eternal or universal objects that I can possess, then I, too, must be eternal or universal. To undermine this erroneous appropriative grasping, Yogācāra texts say: Negate the object, and the self is also negated (e.g., *Madhyānta-vibhāga*, 1:4, 8).

Intentional acts also have moral motives and consequences. Since effects are shaped by their causes, an act with a wholesome intent would tend to yield wholesome fruits, while unwholesome intentions produce unwholesome effects.

THREE NATURES

Yogācāra devised a model of three self-natures (*trisvabhāva*) to explain *vijñapti-mātra* more concisely. The pervasive mental constructions that obstruct our view of what truly is the case are called *parikalpita* (imaginative construction). The actual webs of causes and conditions at play are called *paratantra* (dependent on other [causes]). Other-dependence is so-called to emphasize that no thing exists independently, self-caused, eternal, invariant; everything arises dependent on causes and conditions other than itself, in the absence of which it ceases to be. Ordinarily, *paratantra* is infested with *parikalpita*. *Pariniṣpanna* (consummation) is the removal of *parikalpita* from *paratantra*, leaving only purified *paratantra*.

Since the notion of self-nature is itself a parikalpic idea that presumes selfhood, it, too, must be eliminated. Thus, the three self-natures are actually three nonself-natures (*tri-niḥsvabhāva*). *Parikalpita* is devoid of self-nature since it is unreal by definition. *Paratantra* lacks self-nature, since other-dependence precludes self-nature. *Pariniṣpanna*—the Yogācāra counterpart to the Madhyamaka notion of *śūnyatā* (emptiness), which signifies the lack of self-nature in everything—is defined as the absence of self-nature. Thus, the three self-natures are ultimately understood as three nonself-natures.

EIGHT CONSCIOUSNESSES

According to Buddhism, consciousness arises as a by-product of the contact of a sense organ with its corresponding sphere of sense objects. The eye contacting visibles (e.g., colors and shapes) produces visual consciousness; likewise for the remaining four senses (hearing, smell, taste, and touch). The mental organ, *manas*, operates similarly. Coming into contact with the sphere of mental objects (*mano-dhātu*), mental consciousness (*mano-vijñāna*) arises. Hence, there are six sense organs and six corresponding sense realms, which, combined with the six types of resultant consciousnesses, makes eighteen factors altogether. Yogācāra accepted these eighteen factors but found them inadequate to explain several issues that had become important for Buddhists, including the sense of selfhood, appropriative propensities, continuity of experience, and how projection worked. To address these issues, Yogācāra expanded the mental level, resulting in eight rather than six types of consciousness. *Mano-vijñāna* became the sixth sense organ, a kind of empirical consciousness that discerns mental objects as well as the activities of the five senses; *manas* became the seventh consciousness, responsible for appropriating experience as "mine" and thus infesting experience with a sense of selfhood (thus also called *kliṣṭa-manas*, "defiled mind"). The eighth consciousness, *ālaya-vijñāna*, was a novel innovation.

Yogācāra used a seed metaphor to describe the process of karmic conditioning. Experience engenders a seed that is planted out of sight (unconsciously retained in the *ālaya-vijñāna*), where it remains latent until catalytic conditions bring it to fruition (karmic result, *vipāka*), engendering new seeds of the same type. This was a powerful metaphor in agrarian societies. As a warehouse (*ālaya*) to these seeds, the *ālaya-vijñāna* was called the all-seeds consciousness (*sarva-bījāka-vijñāna*). Since it was the conduit and repository of their fruitions, it was also called *vipāka-vijñāna* (karmic requital consciousness). Since the *ālaya-vijñāna* always operates, even when the other seven consciousnesses temporarily cease (e.g., in deep sleep), it was also called foundational consciousness (*mūla-vijñāna*). Although it stores karmic seeds and engenders their projection, the *ālaya-vijñāna* is a karmically neutral mechanical process (*anivṛta*, *avyākṛta*). *Manas* appropriates the activities of the other consciousnesses, thinking they are "my" experience, and appropriates the *ālaya-vijñāna* as a "self."

Karmic continuity ceases by overturning the basis (*āśraya-parāvṛtti*), in which the *ālaya-vijñāna* and the other consciousnesses cease to function. The conscious-

nesses (vi*jñāna*) lose their discriminative (*vi-*, compare the English prefix *dis-*) projective propensities and become direct cognitions (*jñāna*). *Ālaya-vijñāna* becomes the "great mirror cognition" (*mahādarśa-jñāna*), no longer appropriatively storing or engendering new seeds; instead, like a mirror, it reflects everything impartially without attachment. *Manas* loses its self-prejudicial nature and becomes the equalizing immediate cognition (*samatājñāna*), equalizing self and other. *Mano-vijñāna*, which discriminates cognitive objects, becomes immediate cognitive mastery (pratyavekṣamā-jñāna), and properly discerns the particular and general characteristics of things. The five sense consciousnesses, now unhindered by the mental obstructions of the sixth and seventh consciousness, become immediate cognitions that accomplish what must be done (kṛtyānuṣṭhāna-jñāna), thereby engaging the world effectively. Yogācāra texts differ on which overturning occurs at which stage of practice, but they agree that full enlightenment entails accomplishing all of them.

ASAṄGA ON LANGUAGE AND NONLINGUISTIC THINGS

In his texts—notably the *Madhyānta-vibhāga* and the *tattvārtha* ([relation of] referents and real things) chapter of the *Yogācārabhūmi*—Asaṅga challenges the Madhyamaka claim that emptiness (*śūnyatā*) is the ultimate and final position, the true Middle Way, not by denying the importance and validity of emptiness, but by taking the analysis one extra step. For Asaṅga, emptiness is a tool for eliminating false views, especially the false view of selfhood attributed to beings or things. But once these views are emptied, something remains, namely reality understood as emptied of false conceptualizations.

A quick summary of the *tattvārtha* chapter of Asaṅga's *Yogācārabhūmi* may illustrate how he refashioned rival teachings, in this case redefining emptiness (*śūnyatā*) and the Middle Way, while providing a useful summary of his philosophy. For Madhyamaka the Middle Way (*madhyamā-pratipad*, from which Madhyamaka derives its name) entails that all things are empty (*śūnya*)—meaning they are devoid of self-essence or own-being (*svabhāva*)—because they are dependently arisen from causes and conditions that are themselves empty. Thus, existent things are conventionally real but ultimately empty.

Asaṅga responds by describing four types of people, each experiencing a different phenomenological sphere of reals (*tattva*) and conceptual-linguistic referents (*artha*): (1) ordinary people, (2) philosophers, (3) Hīnayāna adepts, and (4) Mahāyāna adepts, the latter denoting accomplished Bodhisattvas and Buddhas. The first are naive realists, immersed in a cognitive field of compulsive presuppositions (*niścitādhimukti-gocara*), who accept things as established by convention. What appears to be real to an ordinary person has been conceptually and linguistically constructed from one's own discriminative imaginings (*vikalpa*) and remains unquestioned. Philosophers apply rational epistemological methods to logically investigate things and accept as real what has been logically proven by an articulate, discursive person. Hīnayāna adepts who have eliminated the affective obstructions (*kleśāvaraṇa*) realize there is no real referent corresponding to the notion self. They see a person as the five aggregates only (*skandha-mātra*; the five are form, hedonic tone, linguistic conceptualizing, embodied karmic conditioning, and consciousness), conditionally arisen, devoid of an imagined self.

By seeing that not only people, but all things lack selfhood, Mahāyāna adepts eliminate all obstructions to knowable realities (*jñeyāvaraṇa*). Asaṅga then enters a detailed discussion on the relation between the linguistic ideational sphere (nominal reality, *prajñapti*) and its cognitive basis (nonarticulable, nonconceptual things, *vastu*), providing numerous reasons for why they are not reducible to each other, nor entirely separable from each other. For Asaṅga emptiness signifies cleansing cognition of erroneous conditioning and views, so that reality is cognized nonerroneously. Emptiness is not a final state, but a purificatory, antidotal process that eliminates erroneous conceptualizations; once they are eliminated what remains is reality. Since this remainder is nonconceptual and therefore nonlinguistic, it cannot be adequately rendered in words without re-reducing it to the conceptual sphere.

Put simply, to perceive something blue is nonlinguistic (and hence indescribable to a blind person), though one can conceive of it as "something blue." The concept *blue* is neither the same nor different from the perception. Without *vastus*, referential articulations (*abhilāpya*) would have no basis; without such articulations, the nature of *vastus* could not be defined or intellectually understood. To think that *vastus* are merely nominal realities is more pernicious than believing in self, Asaṅga argues, since believing in self is to be mistaken about only one type of knowable, whereas to reject all *vastus* is to be mistaken about everything. Not holding the extreme views that (1) nonexistent things (like self) exist or (2) that all cognizables are nonexistent is, for Asaṅga, the true

Middle Way. Neither *prajñapti* nor *vastu* is rejected completely or accepted naïvely.

In his *Madhyānta-vibhāga*, implicitly deploying the theory of three natures to explain Buddhist practice, Asaṅga illustrates how emptiness and cultivating positive insight (*pariniṣpanna*) act as an antidote (*pratipakṣa*) to the pervasive false mental constructions (*parikalpita*) one projects as lived experience, resulting in reality being experienced just as it is (purified *paratantra*). In the *Mahāyānasaṃgraha* Asaṅga asserts that bondage and liberation cannot be explained coherently without reference to the *ālaya-vijñāna*, since it conveys the seeds and habits (*vāsanā*) that make bondage and liberation possible. Even brief contact with true Buddhist teachings (*saddharma*) may instill a propensity (*śruta-vāsanā*), outside one's conscious awareness, toward enlightenment and Buddhahood. Asaṅga claims this propensity, called *mano-jalpa* (mental murmuring), is utterly different from and irreducible to the *ālaya-vijñāna*; it gradually destroys the *ālaya-vijñāna* from within, like a germ infecting a host. Eliminating the *ālaya-vijñāna* results in Buddhahood. To label Asaṅga an idealist would be a gross mischaracterization.

VASUBANDHU'S *TWENTY VERSES*

In the *Viṃśatikā* (Twenty verses) Vasubandhu, following Asaṅga's lead, refutes the realism of naive and philosophical realists. The realists assert that the objects we perceive exist outside of consciousness, which is the reason that these objects remain stable through (1) time and (2) space; (3) different people can have differing perceptions of a thing and yet reach a consensus about it; and (4) the objective world operates by determinate causal principles, rather by than imaginary, ineffective fantasies.

Vasubandhu addresses each of these four claims with numerous counterarguments, including an analogy to dreams. In dreams seemingly external objects appear as if in time and space, even though no actual external object is present to cause them, thus proving that while consciousness is a necessary and sufficient condition for objects to appear in perception, the presence of actual external objects is neither necessary nor sufficient. For Vasubandhu, as for Asaṅga, the perceptions of ordinary people are like a dream, a mental projection based on conditioned predispositions. That different beings have differing perceptions of the purportedly same thing proves this. Updating Vasubandhu's example, that flies and humans perceive and react to excrement in radically different ways, demonstrates that what each perceives is a projection based on its own conditioning, or its own mental seeds (*bījās*) acquired from past experiences (perhaps in past lives). Moreover, karma (action) is collective, in that we gravitate toward beings or types who perceive as we do, erroneously justifying the seeming universality of our group perspective.

Thus, the varying perception argument supports rather than undermines the Yogācāra position. Note that the dream example and the varying perception example not only neglect to disprove that something outside the activities of consciousness may play a role in its perceptions; on the contrary, both require that there be such a thing for the examples to make sense at all. The observation that dreams imitate waking perceptions minus the presence of actual objects requires that we appreciate the contrast; the object in contention between flies and humans is obviously not reducible to the perceptual projections of either.

Vasubandhu is not arguing for either a subjectivism or a metaphysical relativity, but he is pointing out that we mistake our imaginings for reality, obstructing our view of things as they are. Projective imaginings blocking our vision can have powerful karmic consequences, as Vasubandhu shows in his response to the realist's claim about causal efficacy. He uses the example of a wet dream. Though the erotic cognitive object is a mental construction, without an actual external or physical corresponding object present, the imaginative act causes actual seminal emission, a physical effect produced outside the dream and recognized as such on awakening. The monastic vow of celibacy treats wet dreams as an infraction of the monastic code. Even though dreams are only fantasies, they have real karmic consequences. The deluded mind produces real effects that can only be fully known after awakening, once delusion has ceased. Awakening means enlightenment—the term *bodhi* (awakening) can also mean "enlightenment"—the cessation of the deluded mind. Even though we act in a collective deluded world of our own construction, our actions have real causal consequences.

The realist objects that objects perceived while awake seem stable in time and space, whereas objects in dreams do not. Vasubandhu replies that objects and events seem less clear and consistent in dreams because one's mind is overcome by sleepiness so one is not "thinking clearly." Furthermore, one does not know that the objects in a dream are only dream-objects until one awakens. Vasubandhu's reply to the question of whether we can know other minds extends the dream analogy: Even our own minds are opaque to us since our mental capacities are dim and sleepy. However, one who is awakened (the

literal meaning of *Buddha*) can know other minds more clearly than we know our own. Not only can we know other minds (if we awaken), but we constantly influence each other for better and for worse (though we may not notice that within our individual dreams). Thus, karma is intersubjective. Moreover, since the more awake one is the more causally effective one's mind becomes, sages and buddhas can exert powerful effects on the world, including devastating destruction, and even life and death.

THE FIVE STAGES

Precise details of the stages in which the mental stream is purified of pollutants (*āsrava*), filtering out karmically unwholesome seeds while nourishing and fortifying the wholesome ones, vary across different Yogācāra texts. A five-stage model is found in several foundational texts and has become the standard account:

> (1) During the provisioning stage (*saṃbhārāvasthā*) one gathers and stocks up on "provisions" for the journey. The provisions consist of orienting oneself toward the pursuit of the path and developing the proper character, attitude, and resolve to accomplish it. This stage commences at the moment the aspiration for enlightenment (*bodhicitta*) arises.
>
> (2) Next is the experimental stage (*prayogāvasthā*), in which one converts Buddhist theory into praxis. *Prayoga* also means "intensifying effort," or applying oneself with increasing vigilance. While increasing meditative abilities, one begins to suppress the grasper-grasped relation and commences on a careful and detailed study of the relation between things, language, and cognition.
>
> (3) The next stage is deepening understanding (*prativedhāvasthā*), also called the Path of Corrective Vision (*darśana-mārga*). One works on realizing the emptiness of self and *dharmas* while reducing the obstructions (*kleśāvaraṇa* and *jñeyāvaraṇa*). This stage ends when one acquires some insight into nonconceptual cognition (*nirvikalpa-jñāna*), that is, cognition devoid of interpretive or imaginative overlay.
>
> (4) In the Path of Cultivation (*bhāvanā-mārga*), nonconceptual cognition deepens. The grasper-grasped relation is utterly eliminated, as are all cognitive obstructions. This path culminates in the full Overturning of the Basis, or enlightenment.
>
> (5) In the final stage (*niṣṭhāvasthā*) one abides in unsurpassable complete awakening and engages the world through the four immediate cognitions (mirror cognition and so on). All of one's activities and cognitions are postenlightenment (*pṛṣṭhalabdha*), and one compassionately endeavors to alleviate the suffering and ignorance of others.

See also Buddhism.

Bibliography

Anacker, Stefan, trans. and ed. *Seven Works of Vasubandhu: The Buddhist Psychological Doctor*. Delhi, India: Motilal Banarsidass, 1984.

Cook, Francis H., trans. *Three Texts on Consciousness Only*. Berkeley, CA: Numata Center, 1999.

Ganguly, Swati. *Treatise in Thirty Verses on Mere-Consciousness*. Delhi, India: Motilal Banarsidass, 1992.

Griffiths, Paul. *On Being Mindless: Buddhist Meditation and the Mind-Body Problem*. La Salle, IL: Open Court, 1986.

Griffiths, Paul, et al. *The Realm of Awakening: Chapter Ten of Asaṅga's Mahāyānasaṃgraha*. New York: Oxford University Press, 1989.

Hayes, Richard P. *Dignāga on the Interpretation of Signs*. Dordrecht, Netherlands: Kluwer Academic, 1988.

Hopkins, Jeffrey. *Reflections on Reality: The Three Natures and Non-natures in the Mind-Only School*. Berkeley: University of California Press, 2002.

Kochumuttom, Thomas. *A Buddhist Doctrine of Experience: A New Translation and Interpretation of the Works of Vasubandhu the Yogacarin*. Delhi, India: Motilal Banarsidass, 1982.

Lamotte, Étienne, trans. *Saṃdhinirmocana-sūtra*. Louvain, Belgium: Université de Louvain and Adrian Maisonneuve, 1935.

La Vallée Poussin, Louis de, trans. *Vijñaptimātratāsiddhi*. 2 vols. Paris: Geuthner, 1928.

Lusthaus, Dan. *Buddhist Phenomenology: A Philosophical Investigation of Yogācāra Buddhism and the Ch'eng wei-shih lun*. London: RoutledgeCurzon, 2002.

Nagao, Gadjin. *Mādhyamika and Yogācāra*. Translated by Leslie Kawamura. Albany: SUNY Press, 1991.

Powers, John. *Wisdom of the Buddha: The Saṃdhinirmocana Mahāyāna Sūtra*. Berkeley, CA: Dharma, 1995.

Powers, John. *The Yogācāra School of Buddhism: A Bibliography*. Metuchen, NJ: Scarecrow Press, 1991.

Rahula, Walpola, trans. *Le Compendium de la Super-Doctrine d'Asaṅga (Abhidharmasamuccaya)*. Paris: Publications de l'École Française d'Extrême Orient, 1971. Translated by Sara Webb-Boin as *Abhidharmasamuccaya: The Compendium of the Higher Teaching*. Freemont, CA: Asian Humanities Press, 2001.

Schmithausen, Lambert. *Ālayavijñāna*. Tokyo, Japan: International Institute for Buddhist Studies, 1987.

Shih, Heng-ching, and Dab Lusthaus, trans. *A Comprehensive Commentary on the Heart Sutra: Translated from the Chinese of K'uei-chi*. Berkeley, CA: Numata Center, 2001.

Sparham, Gareth, trans. *Ocean of Eloquence: Tsong kha pa's Commentary on the Yogācāra Doctrine of Mind*. Albany, NY: SUNY Press, 1993.

Tat, Wei, trans. *Ch'eng Wei-Shih Lun: The Doctrine of Mere Consciousness*. Hong Kong: Ch'eng Wei-Shih Lun Publication Committee, 1973.

Tatz, Mark. *Asaṅga's Chapter on Ethics with the Commentary of Tsong-kha-pa: The Basic Path to Awakening, the Complete Bodhisattva*. Lewiston, NY: Edwin Mellen, 1986.

Willis, Janice Dean. *On Knowing Reality: The Tattvārtha Chapter of Asaṅga's Bodhisattvabhūmi*. New York: Columbia University Press, 1979.

Dan Lusthaus (2005)

BUDDHIST EPISTEMOLOGY

For Buddhist thinkers philosophy should aid one in eliminating suffering and obtaining happiness. They maintain that to achieve those ends, one must eliminate ignorance (*avidyā*), a fundamental mental flaw that is suffering's basic cause. Although variously construed ignorance inevitably involves the mistaken belief that a fixed, unchanging personal essence, or *ātman*, lies at the core of each person's identity. Hence, to eliminate ignorance one must eradicate that belief, and to do so Buddhist philosophers stress the importance of seeing things as they are (*yathābhūtadarśana*), a corrective cognitive state through which one knows that persons necessarily lack essence. The need to give an account of such a state leads to a concern with epistemology in Buddhist thought from its earliest period (500 BCE–100 CE) in South Asia.

Although early Buddhism evinces a nascent epistemology, precise and sophisticated accounts of knowledge do not begin until adequate tools are developed by South Asian philosophers, primarily non-Buddhists, starting no later than the first century CE. The Buddhist theorist Vasubandhu initially appropriates these tools, but Dignāga first employs them in a manner that reflects all the issues addressed by later Buddhist epistemologists. Finally, Dharmakirti modifies and expands Dignāga's work in such a manner that all subsequent Buddhist epistemologists in India and Tibet cast their work as interpretations of Dharmakirti's philosophy. Hence, for the purposes of this entry, *Buddhist epistemology* refers to the thought of Dharmakirti and his subsequent interpreters in both India and Tibet, where epistemological works continue to be composed. In their voluminous writings Buddhist epistemologists express a variety of competing views developed in distinct historical contexts. Nevertheless, they largely agree on the following central theories and principles.

MODEL OF KNOWING

Buddhist epistemologists examine knowledge in terms of a knowledge-event or act of knowing (*pramiti*). Their account rests on the claim that the mind consists of a series of causally related, instantaneous mental moments, each of which is ontologically irreducible. Thus, as a mental event the act of knowing is ontologically identical to a mental moment. The act of knowing occurs when the mind comes into a direct or indirect causal relation with an object such that, with other conditions in place, the next mental moment contains an image (*ākāra*) of the object. Due to the ontological unity of a mental moment, the notion that the mental moment contains an image of the object is metaphorical; in fact, the image is ontologically identical to that mental moment itself. Nevertheless, from a phenomenal standpoint the act of knowing presents itself with two images, the aforementioned object-image (*grāhyākāra*) and a subject-image (*grāhakākāra*). The latter accounts for the sense of subjectivity in the act of knowing, whereas the former accounts for the content of the cognition.

On the Buddhist theory of mind all cognitions must have an object, which is to say that all cognitions have an object-image. Not every cognition, however, is an act of knowing. Instead, only two types of cognitions—perception (*pratyakṣa*) and inference (*anumāna*)—can be acts of knowing because only they can satisfy two criteria: they are reliable (*avisaṃvāda*) and they are motivators of action (*pravartaka*). Reliability concerns the justification of knowledge. The fact of being a motivator of action is a psychological feature that reflects teleological and ontological concerns.

RELIABILITY

For Buddhist epistemologists, an act of knowing—whether it be a perception or an inference—is reliable in that it directs one to an object with the desired telic efficacy (*arthakriyā*). On this criterion an act of knowledge is distinguished from an unreliable cognition in one of two ways: Either it directs one to an object that can fulfill a particular goal, or it presents itself as the fulfillment of that goal. Suppose, for example, that one is cold, and that one seeks to warm one's hands at a fire. Because the hearth contains a fire that is capable of fulfilling one's goal, the perception of a fire in the hearth is deemed reliable. When one reaches the hearth, the sensation of heat on one's hands is itself the fulfillment of one's goal. Thus, that cognition of heat is also reliable.

By grounding reliability in telic efficacy Buddhist thinkers seek to justify beliefs by interpreting them as

descriptions of their objects' causal characteristics. Hence, the ultimate arbiter of a cognition's reliability is the way in which it presents its objects in causal terms. If it presents the object's causal characteristics such that the object is capable of functioning in the expected fashion, then the cognition is reliable; otherwise, it is not. In some cases the evidence for the desired functionality is given with the cognition itself: for example, the sensation of warmth requires no other cognition to verify that one is feeling warm. Such cognitions are said to be intrinsically (*svataḥ*) reliable, and this applies to all inferences and some perceptions. In other instances of perception another cognition must verify the cognition's content. One may only glimpse the fire from a corner of the room, and one must appeal to inferential evidence (such as smoke) or a subsequent perception to verify that one was indeed seeing fire. A perception that requires such confirmation is said to be extrinsically (*parataḥ*) reliable.

PURPOSE AND MOTIVATION

Arguments for a cognition's reliability generally serve to justify a belief. Thus, one's belief that "there is a fire in the hearth" is true inasmuch as the cognition that includes that belief reliably represents the causal characteristics of the object in question. For that cognition to be an act of knowing, however, that cognition must include other dispositions. Of prime importance is the desire to know (*jijñāsa*) without which the cognition could not arise: it may be true that "there is a fire in the hearth," but without some purpose one will not have a cognitive event in which that belief occurs. Thus, for Buddhists the account of knowledge as justified true belief is inadequate if that account ignores the role played by cognitive dispositions, especially those related to purpose.

In appealing to dispositions related to purpose Buddhist epistemologists hold that the reliability of a belief shifts according to the purpose to which it is tied. One might believe, for example, that the object on one's table is an unbreakable vase, although it is in fact fragile. Relative to the purpose of containing a bouquet, the cognition in which that belief occurs is reliable, since the vase can function so as to hold flowers. But relative to the aim of cracking a walnut's shell, a cognition in which that belief occurs would not be reliable, since the vase lacks the causal capacity to crack open a nut. By thus evaluating complex beliefs within various teleological contexts, Buddhist thinkers can accept some philosophical claims in one context, while rejecting them in another—a strategy that is central to Buddhist soteriology.

In relating reliability to purpose Buddhist epistemologists argue that an act of knowing must not only be reliable but must also be a motivator of purposeful action. Frequently, this assertion is formulated as a requirement for novelty, whereby an act of knowing reveals a previously unknown object (*ajñātārthaprakāśa*). On either version—motivation or novelty—this requirement points not only to the role of purpose but also to the notion that an act of knowing reduces doubt. That is, the cognition must pass a threshold whereby the person, usually idealized as judicious (*prekṣāvant*), is willing to act on a particular goal based on the content of that cognition. The early epistemologist Dignāga appears less concerned with the utter removal of doubt, but Dharmakirti and most subsequent thinkers maintain that an act of knowing grants apodictic certainty, even if certainty must sometimes be supplied by a subsequent cognition.

Finally, the notion that an act of knowing must motivate action is also tied to ontological issues. The chief concern here is to eliminate the possibility that universals could be the objects of perception. As will be evident in the following text, the Buddhist strategy is to make perception the actual motivator of action, while relegating the determinate content of perception to a subsequent judgment, which is not strictly speaking the motivator.

PERCEPTION AND ILLUSION

As one of the two types of cognitions that are both reliable and motivate action, a perception is an act of knowledge. The Buddhist model of perception is causal and eidetic: an object interacts with a sense-organ such that, with other factors in place, the next moment of mind occurs with an image or simulacrum (*sādṛśya* or *sārūpya*) of the object. Unlike inference, in perception the image is produced directly by the object, and the reliability of perception is based on this direct causal relation.

As a mental moment, a perception is causally conditioned by the previous mental moment, including all the dispositions and physiological conditions that contribute to its occurrence. In a perception, however, not only the previous mental moment but also the perceived object is contributing causally to the occurrence of the perception. Hence, the causal character of the mental moment that is a perception is restrained (*niyata*) by the causal characteristics of the object to which it is in relation through the sense organ. Thus, a perception is reliable—it accurately reflects the object's causal characteristics—because the causal constraints imposed by the object on the perception's contents are indicative of that object's causal characteristics. To put it another way, the perception of blue is

a reliable indicator of its object's causal characteristics because when that content—an image of blue—is the undistorted effect of an object, it can only be produced by an object with the causal capacity to produce a blue image.

This appeal to a causal relation between perceptual content and object compels Buddhist epistemologists to face the problem of illusion. A favorite Tibetan example is the "blue snow mountain": When one looks at a snowy Himalayan peak on a clear day, the snowcap appears blue because it reflects the sky's color. Here, the cognition is a spurious perception (*pratyakṣābhāsa*) because it lacks reliability, in that snow is not blue. But since the perceptual content—the image—is distorted by causal factors not given with the object, the content itself does not provide any basis for recognizing that distortion. Instead, some other perception or inference would need to reveal that distortion. Still, as noted earlier, some perceptions are alleged to be intrinsically reliable, such that they do not require confirmation by a subsequent act of knowing. What then would distinguish those perceptions such that, unlike the sight of "blue snow," they could never be spurious?

Buddhist epistemologists do not provide an easy answer to this question, but their theory of perceptual judgment provides a partial response. On their view perception itself is indeterminate in that it involves no conceptual or linguistic operation. A purely indeterminate cognitive event, however, cannot be either reliable or unreliable because it conveys no knowledge about the causal characteristics of its object in relation to one's goal. Hence, the reliability of a perception consists in that it leads to an immediately subsequent perceptual judgment (*tatpṛṣṭḥalabdhaniścaya*) that does provide that knowledge. Strictly speaking, only the judgment is reliable or unreliable, in that it only describes the object in a determinate fashion. Nevertheless, since the form of that judgment is causally constrained by the image presented by indeterminate perception, the perception itself is considered reliable.

Returning, then, to the problem of illusion, the theory of perceptual judgment means that an uninterpreted perception could not itself be an act of knowing because, lacking any depiction of its object's causal characteristics, it could not be reliable. But when the subsequent judgment describes the object, it must do so in relation to a particular goal. One explicit outcome of this in theory is that a perception may only be partially reliable in that it can lead to correct judgments in regard to one goal, but not in regard to some other goal. For example, the perceptual content interpreted as "blue snow" might be unreliable in regard to one's need to identify a blue object, and yet it may still be reliable in regard to the need to identify snow. Although the implications of this claim are left covert, it seems likely that for Buddhist epistemologists one factor in the intrinsic reliability of some perceptions is that the goals in question are such that the perceptual content could never be erroneously interpreted. In other words the teleological context constrains the perceptual judgment such that incorrect interpretations of the perceptual content cannot occur in those cases.

PERCEPTUAL JUDGMENT AND ONTOLOGY

Besides its role in intrinsic reliability, the theory of perceptual judgment is also closely allied to Buddhist ontological concerns. For Buddhist epistemologists to exist is to be knowable (*jñeya*), and since knowledge is a causal process, an existent entity must therefore be causally efficient; likewise, any causally efficient entity must exist. The paradigmatic case of an entity's causal efficiency is its capacity to produce an image of itself in a perceiver's mind, and it is for this reason that Dharmakirti remarks, "To exist is to be perceived" (*sattvam upalabdhir eva*). Moreover, since any object of perception must exist, Buddhists are careful to exclude the possibility of perceiving any metaphysically objectionable entity, such as a fixed personal essence. Largely because a personal essence is considered a special case of a universal, Buddhists likewise reject the existence—and hence the perception—of universals. Instead, only particulars (*svalakṣaṇa*) truly exist, and particulars alone are the objects of perception because only particulars are causally efficient.

Perception cannot include universals, and linguistic or conceptual cognitions must include universals. Hence, perception must be a sheer apprehension of an object that is not linguistic or conceptual in character. But as noted earlier, the criterion of reliability requires a determinate cognition, which is necessarily conceptual or linguistic in form. Hence, on the one hand, perception must be the immediate apprehension of a particular through a nonconceptual image in the mind and, on the other hand, to be reliable and to motivate action, that nonconceptual content must be interpreted by a determinate cognition. The solution is to relegate the determinate aspect of a perception to an immediately subsequent judgment, and in doing so Buddhists avoid the notion that linguistic or conceptual entities—that is, universals—are the objects of perception.

INFERENCE AND THE PROBLEM OF REFERENCE

Besides perception, inference is considered an act of knowing. As with perception, inference is a cognitive event in which an image of the object appears. Unlike perception, however, the image in an inference is not directly produced by the object. Instead, it bears an indirect causal relation to the object in two ways, namely, by way of the relations on which an inference relies and by way of the process of constructing universals.

The Buddhist approach to universals is central to their theory of inference because inferences are conceptual or linguistic acts of knowing. Moreover, their theory arises in response to the way their non-Buddhist rivals address the problem of reference. In short, these rivals claim that, for words to successfully refer to their proper referents, they must always have a relation to those referents and only to those referents. The word *cow*, for example, should refer only to a cow, and not to something different, such as a horse. Each individual cow, however, is different from every other cow. Hence, if the word *cow* were to stand in a direct relation to one individual cow, it should always refer only to that individual. Such would be the case because the word *cow* should never refer to something that is different from its proper referent, and if the proper referent of the word *cow* were a particular cow, then by referring to some other cow, the word *cow* would be referring to something different from its proper referent. And if the word *cow* can refer both to its proper referent and something other than its proper referent, why should it not refer to a horse?

Most South Asian thinkers solve this familiar problem in the philosophy of language by positing the existence of real universals (technical terms for which include *sāmānya*, *jāti*, and *ākṛti*). On this model, the word *cow* does not have a direct relation to any particular cow. Instead, it is directly related to the universal "cowness." Nevertheless, the word *cow* still refers successfully to each individual cow because the universal cowness is necessarily instantiated in each individual cow. A word such as *cow* thus refers to each particular cow by virtue of the universal cowness to which both the word and each particular are related. On this view one can thus say that all cows are the same not because each individual is identical, but because each individual instantiates that one universal cowness.

This model is problematic for Buddhists because it would justify the false belief in a personal essence. That is, just as cowness is present in each different cow in time and space, so, too, a personal essence would be present in all the different spatiotemporal instances of what people consider to be one person. To avoid this outcome, Buddhist epistemologists therefore deny the ultimate reality of universals as things in the world. Thus, for them the universe is populated by spatiotemporally unique particulars, and nothing more. All cows are in fact unique; one only thinks that they are the same because one constructs a universal or sameness (*sāmānya*) for them. So too, each spatiotemporal instance of a person is actually unique. "John" at birth and "John" at forty-five are actually different. When one constructs a sameness that warrants one's use of the label "John," one falsely believes that the sameness is not constructed, but real.

THE EXCLUSION THEORY

Although Buddhist epistemologists deny the ultimate existence of universals, they nevertheless adopt their rivals' approach to reference. They are therefore obliged to formulate a theory that, while denying the ultimate reality of universals, accounts for the way that universals may be contingently constructed so that words may refer to their referents. Buddhists develop a model known as the exclusion theory (*apohavāda*), and to do so they once again resort to causality. In brief, the sameness required to construct the universal cowness is formulated by appealing to the causal characteristics of the individuals in question. More specifically, even though all individual cows are in fact utterly unique and distinct, one may ignore the differences among them and focus instead on the way they are different or excluded from all other entities. That difference or exclusion from other entities is a matter of causality: All cows are the same in that they are all equally different from those entities that are not capable of the causal functionality that one expects when one uses the word *cow*.

Here as well, the paradigmatic case for causal functionality is a perceptual image. Thus, all cows are the same in that they produce the same effect, namely, the same perceptual image in the mind. The problem, however, is that just as each cow is a unique individual, so, too, each perceptual image should also be a unique mental particular. Hence, Buddhist epistemologists must argue that all those images are the same, and to do so they use the same reason: Those images are the same because they all have the same effect, which in this case is a second-order determination of sameness (*ekapratyavamarśajñāna*). The obvious question here is: What warrants the sameness of all those determinations? If one again asserts that they all have the same effect, then the argument ends in an infinite regress. Well aware of this

problem, Buddhist epistemologists follow Dharmakirti's argument: The sameness of those second-order determinations is not constituted by the fact that they all produce the same effect; rather, they are counted as the same because they are phenomenally presented in that fashion. In short, each instance of the judgment, "That is a cow," just seems the same.

Dharmakirti's answer to the problem of infinite regress may seem ad hoc, but it probably reflects a subtle approach to conventionality. In brief, Dharmakirti apparently holds that some conventions—including causality—are so stable that they may be treated as invariable when they are used in nomological arguments about the interpretation of perceptual content. Most Buddhist epistemologists, however, do not pursue this controversial aspect of Dharmakirti's thought and instead leave such concerns to philosophers of the *Madhyamika* or Middle Way school.

RELATIONS IN INFERENCE

The exclusion theory and the attendant problem of infinite regress may leave several questions unasked, but Buddhists seem satisfied with its use, perhaps because it so greatly simplifies the theory of inference. On their view, all inferences take this basic form: "S is P because S is E," where S is the subject of the proposition to be proven, P is the predicate, and E is the evidence. A common example would be: "The mountain is a locus of fire because it is a locus of smoke." The success of the inference depends on the pervasion (*vyāpti*), which by the time of Dharmakirti is understood as a necessary relationship between evidence and predicate. Dharmakirti formulates this relation as a necessary rule of unaccompanied nonarising (*avinābhāvaniyama*). In other words the evidence cannot occur if it is not accompanied by the proximate occurrence of the predicate, or to put it another way the predicate is necessarily predicable of any subject to which the evidence is correctly predicated.

Buddhist epistemologists describe this invariable relation between evidence and predicate as being of only two kinds: either the evidence is the effect of the predicate, or else the evidence stands in a relation of identity (*tādātmya*) to the predicate. The causal relation is operative in the inference of fire from smoke; the identity relation is operative in an inference such as, "This is a tree because it is an oak."

Both in the case of the causal relation and the identity relation the success of the Buddhist analysis of inference depends heavily on the exclusion theory of meaning and reference. For example, when one infers the presence of fire from seeing smoke, the inference succeeds precisely because of the meaning of the concept smoke. That is, an instance of smoke is excluded from all those other entities that do not have the causal characteristics of smoke. One of those characteristics is central to the inference: namely, that any entity properly called smoke is necessarily caused by an entity that can be properly described as fire. Hence, if one's perceptual content has been correctly interpreted, the identification of the object as smoke already gives one the information needed to infer the presence of fire. The same type of account holds true in the identity relation: the concept or term *oak* can only be properly applied to an entity that also has all the causal characteristics that make it suitable to be called a tree. In this way the inferential process is a matter of recognizing the relation between concepts, sometimes through the help of empirical examples.

The exclusion theory thus provides a seemingly analytical relation between the concepts employed in an inference, and inferences are therefore treated as intrinsically reliable. This suggests that inference is largely a matter of understanding the conventions that govern the use of concepts. The problem, however, is determining whether those conventions accurately depict the causal characteristics of real things. How does one determine, for example, that smoke is necessarily produced by fire? Here, one encounters the general problems of induction, and while Buddhist epistemologists propose various empirical means of overcoming such problems, it would be difficult to argue that they have fully succeeded.

See also Buddhism; Buddhism—Schools: Dge-lugs; Buddhism—Schools: Madhyamaka; Epistemology, History of; Illusions; Mind and Mental States in Buddhist Philosophy; Perception; Reference; Universals, A Historical Survey; Vasubandhu.

Bibliography

Dreyfus, Georges B. J. *Recognizing Reality: Dharmakīrti's Philosophy and Its Tibetan Interpretations*. Albany: SUNY Press, 1997.

Dunne, John D. *Foundations of Dharmakīrti's Philosophy*. Boston: Wisdom Publications, 2004.

Frauwallner, Erich. "Landmarks in the History of Indian Logic." *Wiener Zeitschrift für die Kunde Süd und Ostasiens* 5 (1961): 125–148.

Hayes, Richard P. *Dignāga on the Interpretation of Signs*. Dordrecht: D. Reidel, 1987.

Katsura, Shoryu. "Dharmakīrti's Theory of Truth." *Journal of Indian Philosophy* 12 (3) (1984): 215–235.

Matilal, Bimal Krishna. *Perception: An Essay on Classical Indian Theories of Knowledge*. Oxford, U.K.: Clarendon Press, 1986.

Siderits, Mark. "Apohavāda, Nominalism and Resemblance Theories." In *Dharmakīrti's Thought and Its Impact on Indian and Tibetan Philosophy: Proceedings of the Third International Dharmakīrti Conference, Hiroshima, November 4–6, 1997*, edited by Shoryu Katsura. Vienna: Verlag der Österreichische Akademie der Wissenschaften, 1999.

Siderits, Mark. "Deductive, Inductive, Both or Neither?" *Journal of Indian Philosophy* 31 (1–3) (2003): 302–321.

Steinkellner, Ernst. "Once More on Circles." *Journal of Indian Philosophy* 31 (1–3) (2003): 323–341.

Tillemans, Tom J. F. *Scripture, Logic, Language: Essays on Dharmakirti and His Tibetan Successors*. Boston: Wisdom Publications, 1999.

John D. Dunne (2005)

BUFFON, GEORGES-LOUIS LECLERC, COMTE DE

(1707–1788)

The French naturalist and author Georges-Louis Leclerc, Comte de Buffon, enjoyed international acclaim for the artistic expression of his own grandiose, often brilliant theories and for presenting in similar fashion the discoveries of leading contemporaries, particularly in the field of natural science.

LIFE

Born at Montbard, son of an upper middle-class magistrate, Buffon was first educated by the Jesuits of Dijon. Details about his personal life are sparse and uncertain. It is generally believed that, after studying law and despite a marked proclivity for mathematics, he went to Angers at the age of twenty-two to study medicine while indulging in botany and horsemanship. His stay ended abruptly when, presumably having killed an opponent in a duel for no verifiable reason, he set out on travels through France and Italy with the irresponsible young duke of Kingston. His mother's death in 1731 recalled him to Montbard where, as heir to her wealth, he turned the family manor into a château. Assuming the name of de Buffon, he adroitly enlarged his estates, which, in due course, were raised to an earldom.

The rest of his long life was divided between Montbard and Paris; no evidence has yet appeared supporting the belief that he also spent a year in England. When only twenty-six, he was, through influence in high places, elected to the Academy of Science after having presented a paper on mathematical probability. He was soon engaged in silviculture and publishing experiments on the means of preserving and strengthening wood, and his reputation as a scientist was further enhanced by a translation in 1735 of Stephen Hales's *Vegetable Staticks* and, five years later, of Isaac Newton's *Method of Fluxions,* for which he wrote a much admired preface on the history of calculus.

From 1739 until his death he was curator of the Jardin du Roi in Paris, which, under his direction, expanded greatly and became an important scientific center. By 1740 he had begun work on his monumental forty-four-volume *Histoire naturelle,* the most ambitious and comprehensive history of natural science until recent times. Buffon was aided in this enormous task by reports from correspondents scattered throughout the world and by a team of highly specialized collaborators at home.

The first three volumes of the *Natural History,* including *Theory of the Earth* and *History of Man,* appeared in 1749. Published by the royal press, they were exempt from censorship. Almost immediately, however, they incurred the wrath of the Sorbonne for the bold views that ran counter to the book of Genesis. Out of deference to religious authority, Buffon penned an act of submission, only to proceed serenely in the same audacious manner.

Along with the volumes on quadrupeds (1753–1767), birds (1770–1783), and minerals (1783–1788) were the so-called *Supplements* (1774–1779), which included his justly famous work on Earth's geological periods, *The Epochs of Nature* (1778). After Buffon's death the vast project was brought to a close by B. G. E. Lacépède, with eight volumes on oviparous quadrupeds, snakes, fish, and whales.

Buffon's *Discourse on Style,* delivered upon the occasion of his admission to the French Academy in 1753, remains the best known of his shorter pieces. It contains the celebrated dictum: "The style is the man himself," the meaning of which has often been simplified to the point of misinterpretation.

THOUGHT

Buffon's death in Paris shortly before the French Revolution was mourned by the leading journals of Europe as the passing of one of the great figures of the century. His place in the history of ideas has since been undergoing a gradual reassessment still far from settled; certain areas of agreement have, nevertheless, been established. It is generally accepted that while he often engaged in scientific investigation, either through personal observation or through wide reading, his true inclination was for generalization. Influenced especially by Bacon, Newton, Gott-

fried Wilhelm Leibniz, and John Locke, he held seminal views that frequently inspired others to push his inquiries to fruitful conclusions. He rejected the popular conception of God as the Great Clockmaker and, instead of final causes, he looked for natural causes to explain the world about him. He insisted, and the stand was unusual for the day, that religion and science should be strictly separated. Thus, he evolved the theory that our planetary system had resulted from the glancing blow of a comet against the sun's molten surface. Perhaps the most original contribution of Buffon's cosmogony to science was to have introduced a new concept of the vast expanses of geological time. His published calculation of Earth's age as some 80,000 years, rather than the traditional estimate of 6,000, was in itself a generous concession to the prevailing spirit of the day; in his unpublished manuscripts he deals with figures that run into the millions.

Not an evolutionist in the modern sense, he nevertheless persistently stressed change at least in varieties, if not in species, of animal life. This and similar propositions or speculations led Charles Darwin to acclaim Buffon as the first author in modern times to have treated transformism in a scientific spirit. Moreover, in biology he rightly opposed epigenesis to the more widely accepted preformation theory of generation, though his ideas on "inner molds," "organic molecules" and spontaneous generation have long since fallen into disrepute. "He may be said to have asked all the questions which were to be answered in the course of the succeeding century," the oft-quoted comment of Henry Fairchild Osborn, perhaps remains the best generalization to date on Buffon's contribution to posterity.

See also Bacon, Francis; Darwin, Charles Robert; Evolutionary Theory; Leibniz, Gottfried Wilhelm; Locke, John; Newton, Isaac; Scientific Method.

Bibliography

WORKS BY BUFFON

Histoire naturelle, générale et particulière. 44 vols. Paris, 1749–1804. Translated and edited by William Smellie as *Natural History, General and Particular.* 20 vols. London, 1812. More recent edition edited by Jean Piveteau. Paris, 1954.

Oeuvres complètes de Buffon, edited by J.-L. Lanessan. Vols. 13 and 14: *Correspondance inédite.* Paris, 1885.

WORKS ON BUFFON

Cherni, Amor. "Brute Matter and Organic Matter in Buffon." *Graduate Faculty Philosophy Journal* 22 (1) (2000): 87–106.

Duchesneau, François. "The Role of Hypotheses in Descartes's and Buffon's Theories of the Earth." In *Problems of Cartesianism,* edited by Thomas M. Lennon. Montreal: McGill Queen's, 1982.

Fellows, Otis. "Buffon's Place in the Enlightenment." *Studies on Voltaire and the Eighteenth Century* 25 (1963): 603–629.

Flourens, Pierre. *Des Manuscrits de Buffon.* Paris, 1860.

Gayon, Jean. "The Individuality of the Species: A Darwinian Theory?—From Buffon to Ghiselin and Back to Darwin." *Biology and Philosophy* 11 (2) (1996): 215–244.

Goodman, David. "Buffon's *Histoire Naturelle* as a Work of the Enlightenment." In *The Light of Nature,* edited by J. D. North and J. J. Roche. Dordrecht: Kluwer, 1985.

Heim, Roger et al. *Buffon.* Paris: Museum National d'Histoire Naturelle, 1952.

Wilkie, J. B. "The Idea of Evolution in the Writings of Buffon." *Annals of Science* 12 (1) (1956): 45–62; 12 (3) (1956): 212–227; 12 (4) (1956): 255–266.

Wohl, Robert. "Buffon and His Project for a New Science." *Isis* 51 (2) (1960): 186–199.

Otis Fellows (1967)
Bibliography updated by Tamra Frei (2005)

BULGAKOV, SERGEI NIKOLAEVICH
(1871–1944)

Sergei Nikolaevich Bulgakov, a Russian economist, philosopher, and theologian, was a leading twentieth-century religious philosopher in the tradition of Vladimir Solov'ëv. Bulgakov was born in Livny, Russia, the son of a priest. He attended a church school in Livny and spent four years in a theological seminary before enrolling in the faculty of law at the University of Moscow in 1890. He was graduated in 1894 and began teaching political economy at the Moscow Technical School in 1895. From 1898 to 1900 he traveled in western Europe and Great Britain, gathering material for his master's dissertation, *Kapitalizm i zemledelie* (Capitalism and agriculture; 2 vols., St. Petersburg, 1900). Through this and other writings on economic and social questions he soon acquired a national reputation. After teaching in Kiev for five years, he returned to Moscow in 1906 to become professor of political economy at the Moscow Institute of Commerce; in the same year he was elected to the second state Duma as a Constitutional Democrat. In 1912 he received a doctorate from the University of Moscow, and in 1917 he was named professor of political economy at that institution.

Although Bulgakov was a leading "legal Marxist" in the 1890s, he even then acknowledged the philosophical supremacy of Immanuel Kant and soon began to depart from orthodox Marxism on socioeconomic issues as well. In his master's dissertation he argued that Karl Marx's the-

ory of the centralization of production is inapplicable to agriculture, where small-scale production is more stable and viable than large-scale. When, in the early years of the twentieth century, Bulgakov underwent a religious crisis, he abandoned Marxism completely, first for the idealistic position represented in his book of essays, *Ot Marksizma k idealizmu* (From Marxism to idealism; St. Petersburg, 1903), and subsequently for a mystical, "Sophiological" interpretation of the Russian Orthodox faith showing the direct and extensive influence of Solov'ëv and Pavel Florenskii and the ultimate influence of Plato and Friedrich Wilhelm Joseph von Schelling. In 1909 Bulgakov contributed to the celebrated miscellany, *Vekhi* (Landmarks), in which ex-Marxist Russian intellectuals, including Nikolai Berdiaev and Petr Struve, criticized the radical intelligentsia. Bulgakov first outlined his positive religious philosophy in his doctoral dissertation, *Filosofiia khoziaistva* (The philosophy of the economy; Moscow, 1912) and over the years 1911–1916 he composed the work in which this philosophy received its fullest expression, *Svet nevechernii* (The unfading light; Moscow, 1917).

During the same period Bulgakov studied for holy orders, and in 1918 he was ordained a priest in the Russian Orthodox Church. He moved to the Crimea, where he became professor of political economy and theology at the University of Simferopol', but in 1921 he lost this position because he was a member of the clergy. At the end of 1922 he was expelled from Russia along with many other non-Marxist scholars and writers. He settled first in Prague and lived from 1925 in Paris, where he took part in founding the Orthodox Theological Institute, serving as its dean and professor of dogmatic theology until his death. During these years Bulgakov wrote extensively on theological subjects and took an active part in ecclesiastical conferences in many countries, becoming an internationally known church figure. Some of his later theological works, particularly *Agnets Bozhii* (The lamb of God; Paris, 1933) and *Nevesta Agntsa* (The bride of the lamb; Paris, 1945) also carried further the development of his distinctive philosophical outlook.

Basic to this outlook is a cosmology that, although marked in its expression by obscurities and progressive modifications, centered consistently on the following themes: (1) The world, or cosmos, is an organic whole animated by a "world soul" or entelechy that is revealed in the structure, function, and connection of its parts. (2) God, or the Absolute, in creating the cosmos "out of nothing," created it not as something external or alien to him (for then it would limit the Absolute, which is impossible), but as an emanation of his own nature; the world is God as becoming, the divine nature fused with nothingness. (3) Mediating between the Absolute and the cosmos, uniting them both within itself, is a "third being"—Sophia, the principle of divine wisdom. As the world of Platonic Ideas, Sophia is the ideal basis of the cosmos; as the object of divine love, purely receptive and conceiving everything within herself as the womb of being, Sophia is "eternal femininity"; as the principle of the Divine within the created, she is the "world soul," or entelechy; as a participant with the Trinity in the generation of the cosmos, she is a kind of "fourth hypostasis" in God. In his later works Bulgakov distinguished between the "divine Sophia" in God and the "created Sophia" in the cosmos, but he still emphasized their ultimate metaphysical identity and thus the consubstantiality of God and the cosmos.

Bulgakov resisted the pantheistic implications of his position, preferring to call it a form of panentheism, and strove to provide solutions to the chief philosophical problems it raised, such as the problems of evil and human freedom. He attributed evil to the nothingness or nonbeing that is the substratum of the cosmos: Through the willfulness of created beings, nothingness is actualized as a chaotic force erupting into the created world, which in itself is not evil but simply incomplete. He provided for human freedom through a doctrine of self-creation: man is free even in the act by which he comes into existence, for God allows man to collaborate in his own creation; at the same time, however, Bulgakov also asserted that Sophia guides history by a kind of necessity.

Like Florenskii, Bulgakov laid great stress on the antinomic character of rationality and looked to divine revelation through religious experience for knowledge of the highest truths, but his epistemological views in general received no thorough, original development or synthesis; the same is true of his scattered treatments of ethical questions and of his aesthetic reflections—the latter appearing principally in *Tikhie dumy* (Quiet meditations; Moscow, 1918). The work Bulgakov himself regarded as his most strictly philosophical product—*Filosofiia imeni* (Philosophy of the name)—was written in 1919 but first published posthumously in Paris in 1953. It is an exhaustive study of language, with particular application to theology, in which Bulgakov argued that words are not mere outward signs of meanings but are internally related to them as animate symbols.

Bulgakov's later works abounded in imaginative theological conceptions, including a doctrine of universal salvation and original treatments of the Incarnation and of the theological differences between Roman Catholicism and Orthodoxy. Some of his theological views, par-

ticularly his Sophiology, were severely censured in the early 1930s by the Moscow patriarchate, which affirmed that the doctrine of Sophia is incompatible with the Trinitarian nature of God and that it falsely introduces a distinction between masculine and feminine principles into the divine essence.

See also Absolute, The; Berdyaev, Nikolai Aleksandrovich; Florenskii, Pavel Aleksandrovich; Kant, Immanuel; Marxist Philosophy; Marx, Karl; Plato; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Solov'ëv (Solovyov), Vladimir Sergeevich; Sophia.

Bibliography

ADDITIONAL WORKS BY BULGAKOV

The Orthodox Church. Crestwood, NY: St. Vladimir Seminary Press, 1988.

Pain, J., and Nicholas Zernov, eds. *A Bulgakov Anthology*. Philadelphia: Westminster Press, 1976.

Philosophy of Economy: The World as Household. Translated by C. Evtuhov. New Haven, CT: Yale University Press, 2000.

Poole, R., ed. *Problems of Idealism: Essays in Russian Social Philosophy*. Translated by R. A. Poole. New Haven, CT: Yale University Press, 2003.

Sophia, the Wisdom of God: An Outline of Sophiology. Hudson, NY: Lindisfarne Press, 1993.

Williams, R., ed. *Sergii Bulgakov: Towards a Russian Political Theology*. Edinburgh: T&T Clark, 1999.

WORKS ON BULGAKOV

Evtuhov, C. *The Cross and the Sickle: Sergei Bulgakov and the Fate of Russian Religious Philosophy*. Ithaca, NY: Cornell University Press, 1997.

Kindersley, R., *The First Russian Revisionists: A Study of Legal Marxism in Russia*. Oxford: Clarendon, 1962.

Lossky, N. O., *History of Russian Philosophy*, Ch. 15. New York: International Universities Press, 1951.

Zenkovskii, *A History of Russian Philosophy*. 2 vols. Translated by G. Kline. New York and London: Routledge, 2003.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

BULLOUGH, EDWARD

(1880–1934)

Edward Bullough was a British aesthetician and literary scholar. He taught modern languages at Cambridge University, holding University lectureships in German and then Italia. He never held a philosophy fellowship or chair, but he gave the first lectures on aesthetics at Cambridge, beginning in 1907, and was widely read in aesthetics. He also conducted psychological research on aesthetic responses in collaboration with Cambridge psychologists.

Bullough is known in aesthetics primarily on the basis of a single article, "'Psychical Distance' as a Factor in Art and an Aesthetic Principle," originally published in *The British Journal of Psychology* in 1912, in which he maintains that aesthetic experience depends on a distancing from "our practical, actual self," thereby "permitting only such reactions on our part as emphasise the 'objective' features of the experience." By means of such distance, we can escape what is merely idiosyncratic but still experience a "*personal* relation, often highly emotionally coloured," to the object, whether work of art or nature.

Bullough's proposal stands in the long tradition of theories of disinterestedness dating back to Shaftesbury and Hutcheson and of the "aesthetic attitude" dating back to Schopenhauer. Bullough's innovation in this tradition is to treat aesthetic distance as a factor in both the creation and reception of art (it is obviously not involved in the creation of nature), and as a variable in the sense that different degrees of distance are appropriate for different kinds of objects, artists, and audiences. The experience of a particular object could thus suffer from either underdistancing or overdistancing. Bullough's emphasis on the variability of distance was suggested at least in part by his experiments on different aesthetic responses to colors, reported in a series of articles from 1907 to 1910, and led him to emphasize the variability rather than uniformity of indidual tastes. In this regard he distanced himself from the traditional theory of disinterestedness.

The 1912 article does not, however, fully explain the value of distancing oneself from objects. For Bullough's fuller account of the "aesthetic consciousness" that can be produced by the proper degree of distance from an object and its value to us, one must turn to his lectures on aesthetics, which were posthumously published in 1957. Here, after a thorough review of the problems of previous approaches to aesthetics, Bullough argues that "the aesthetic attitude is neither scientific nor ethical … neither explanatory nor final, but *contemplative*," giving "a plasticity and relief to objects and experiences which they inevitably lose" in ordinary scientific or practical contexts, an experience we obviously enjoy (Bullough 1957, p. 75). Here again Bullough stands in a long tradition, going back at least to Kant, but his view is distinguished by his emphasis on the availability of "aesthetic consciousness" in both quotidian and artistic contexts, and on its role in the creation and the reception of works of art. Bullough also stresses in a distinctive way in which "aesthetic consciousness" is a form of "objectivity,"

although it is distinct from both scientific objectivity and from the "egotistical *subjectivity* of practical consciousness."

Bullough's conceptions of aesthetic distance and aesthetic consciousness were harshly criticized by later analytical philosophers, such as George Dickie. But his use of these concepts to characterize a valued form of experience and to ground the possibility of a general "aesthetic culture" going beyond the specific realm of art save his reflections from the critique that these concepts cannot yield a satisfactory definition of art proper, which was not Bullough's goal.

See also Aesthetic Experience; Aesthetics, History of.

Bibliography

Bullough, Edward. *Æsthetics: Lectures and Essays*, edited with an introduction by Elizabeth M. Wilkinson. Palo Alto, CA: Stanford University Press, 1957.

Dickie, George. "Psychical Distance: In a Fog at Sea." *British Journal of Aesthetics* 13 (1973): 17–29.

Paul Guyer (2005)

BULTMANN, RUDOLF
(1884–1976)

Rudolf Bultmann, the biblical historian and theologian, was born in Wiefelsted, Oldenburg, Germany. He studied at Marburg, Tübingen, and Berlin and taught first at Marburg and then at Breslau and Giessen. In 1921 he became professor of New Testament studies at Marburg, where he remained until his retirement in 1951.

Bultmann's work and the controversies it has generated are of undoubted importance for the philosophy of religion. His ventures in "demythologizing" the New Testament and in reinterpreting its content "existentially" have raised (and have tried to answer) crucial questions about the logical status of religious language and the nature of Christian belief.

CHRISTIAN FAITH

Bultmann's thought was inspired by his keen sense of the remoteness and unacceptability of the thought forms of New Testament Christianity to most people of the twentieth and twenty-first centuries. We do not and cannot see our world as a theater of conflict between supernatural powers, the demonic seeking to possess and destroy us, and God intervening to secure our salvation. Moreover, miracle stories lie at the very heart of New Testament belief: "If Christ be not raised, your faith is vain" (1 Corinthians 15:17). Thus, the critical question is: Must a man, in order to be a Christian, commit himself simultaneously to two mutually incompatible world pictures—that of twenty-first-century science and that of first-century prescientific speculation? According to Bultmann, to attempt this is to make Christian belief unnecessarily difficult. It is equally unrewarding to view Christianity as a strictly and objectively "historical" religion and anxiously to sift all the evidence for and against the recorded events of the life of Jesus. The evidence is substantial enough to show that Jesus indeed lived and that he made a quite extraordinary impact upon certain contemporaries. But if religious faith is to stand or fall with the historicity of, say, the birth stories or the Easter narratives, if its degree of assurance must rationally be tempered with the historical probabilities, the assurance will be pitifully uncertain, and faith will almost certainly fall.

To these perplexities Bultmann offers a bold remedy. The Christian may properly grant that a very large part of the New Testament message is couched in mythical language and does not record objective history. This mythical material is not, however, an embarrassment, and it need not be discarded. It can be interpreted as indirect description not of the cosmos but of the conditions and possibilities of human existence. Historical studies derive their real seriousness not from sheer factuality but from what they discover about viable ways of life and viable options for human decision. Among such options, the Christian gives preeminence to that displayed in the accounts of the cross and the resurrection. For it is through these that God makes available a distinctively "authentic" and free mode of existence to all humanity.

INFLUENCE OF HEIDEGGER

"Authentic" is Martin Heidegger's term. It is only one of Bultmann's many borrowings from *Sein und Zeit*. There is a prima-facie oddness here—a Christian theologian reinterpreting the New Testament teachings in terms of concepts drawn from atheist existentialism. Nevertheless, the concepts are undeniably relevant and, within limits, illuminating. There are clear and suggestive analogies between Heidegger's general picture of inauthenticity and the New Testament's accounts of life "in" and "after" the flesh, the life of the "natural man" who is alienated from God. In both views humans are uneasy, anxious, and guilty over their condition. If to Heidegger Angst reveals that man is "not at home" in the world, the New Testament affirms that here we have no continuing city

but seek one to come. To both we are strangers and pilgrims.

On the "authentic" type of existence, there are both marked similarities and differences in the views of Heidegger and Bultmann. Heidegger's account centers upon a total acceptance of the fundamental conditions of our life. This involves, for any man, a realization of his own death, not as some vague, unpleasant, but indefinite future event, but as something whose constant presence, in possibility, should modify his sense of his own existence at every moment. Christianity, too, speaks of renouncing the world and a life entangled with the world, of "dying" to the life of self. It has, however—or ought to have—some very different things to say about life eternal.

Heidegger's authentic man sees and accepts the limitations on his freedom imposed by the given circumstances of his life as so far lived ("facticity"); he sees the present moment as the locus of decision, and it is in the future that he will work out those authentic possibilities of existence for which he decides. The Judeo-Christian tradition also has a dualism of facticity and freedom: It claims both that man was created "out of the dust of the ground," stressing the given factuality of human existence, and that God "breathed into his nostrils the breath of life," endowing him with freedom to pursue his diverse possibilities.

How can we discover our authentic possibilities? In answering this question both Heidegger and Bultmann point to the thoroughly temporal, historical nature of human life. History discloses human possibility. For Bultmann the Christian is he who, in R. G. Collingwood's term, "incorporates" the essentials of the New Testament story in his present thought and action.

Bultmann's account of the human situation is, therefore, an "existential" analysis, and to call it that is to contrast it both with the findings of empirical psychology and with a philosophical analysis of nonpersonal structures. Far from being based on empirical investigations, existential analysis tries to uncover the concepts that are, and have to be, employed in any such researches—the fundamental concepts of personal existence.

But there are complexities to be noted here. Although to Bultmann the New Testament has much to say about the general human predicament, we must not analyze its discourse exhaustively as delineating permanent and universal human possibilities. The authentic life, crucially, is available to a man only by virtue of divine grace and through his appropriating the Word revealed in Christ.

DEMYTHOLOGIZING

There is, however, an uneasy duality in Bultmann's thought. Almost everything in the New Testament is to be understood as describing modes of personal existence, but not so the central claim of the kerygma itself, the claim that God decisively acted in Christ. This contains a reference to God that cannot be eliminated. Yet it must be noted that although Bultmann refuses to "dekerygmatize," others (Fritz Buri, for instance) have tried to do just that. They have been unable to stop at what looks to them like a halfway house and have taken the kerygma too as material for existential analysis.

Other theologians have offered various arguments to show that Bultmann's position is too extreme. They claim that he has underestimated the importance of objective history, that he has made too many concessions to twentieth-century skepticism, that his existentialist concepts cannot express the full meaning, the nuances, the complex mesh of associations of the biblical writings, that the myth must be kept intact.

It is not surprising, therefore, that the controversy over demythologizing has been intense and involved. This entry shall single out for brief discussion only a few of the most crucial issues, beginning with the question of Bultmann's existentialism.

CRUCIAL ISSUES. (1) Without doubt, Heidegger's existential analysis has provided Bultmann with a valuable nonmythical vocabulary, able to express an important part of the New Testament message. However, there are certainly some points at which his analyses appear to clarify the Christian position but in fact tempt a theologian to distort it seriously. For example, if Christianity were no more than a philosophy of life, then matters of objective history would not be crucial to it. So long as we knew that someone had lived roughly the sort of life Jesus allegedly lived, we could at least take the "imitation of Christ" as an ideal for human living. "Possibility," in this rather weak sense, would be enough. But if we want to go beyond that (as Bultmann certainly does) and claim that God was actually imparting himself in a quite distinctive and decisive way in the events of Jesus' life, then it is a matter of immense seriousness to learn what these events were. We cannot have a historical religion, in that strong sense, without historical vulnerability. For all its subtlety (most likely because of its subtlety), the existential analysis of historicity deflects attention from this uncomfortable fact.

One should not conclude, however, that Bultmann has never stated a coherent and clear position on his-

toricity and Christian belief. In *History and Eschatology* (1957) he expressed himself much more lucidly in alternative terms derived from Collingwood. But the link between his position in this book and traditional Christian theology has become very tenuous indeed. Whatever the impression we receive from other writings of Bultmann, in *History and Eschatology* the Gospel seems to be about human self-understanding from first to last; dependence on objective historicity has receded to the vanishing point.

(2) Several important and difficult New Testament concepts seem to yield very readily to existential analysis; yet these concepts remain philosophically problematic. The concept of "body" has clear existential meaning—related to Heidegger's concept of what it is to "exist-in-a-world." Likewise, "eternal life," in the New Testament, characterizes a manner, or quality, of living. Yet even if much of the meaning of these expressions is translatable into existentialist language, there surely remains a vital part that is not. The existential analysis by itself cannot answer such a question as "Does our existence end with our bodily death?" Nor does it help solve the problems of meaning and logic (particularly problems of personal identity) that arise over concepts like life after death and the resurrection of the dead.

(3) Because the life and personality of Jesus play so muted a part in this theology, and because the summons to authentic existence tends to be rather individualistic in its emphasis, it is very difficult to build up an adequate account of Christian discipleship and Christian love on Bultmann's foundations. The quality of the Christian ethical life has always been determined by the believer's response not simply to the bare proclamation that a new life has been made available to him, but to the concrete particularities of the life and teaching of Jesus. One guesses that a theology like Bultmann's can succeed in expressing this quality only through implicit dependence on a more conservative view of the New Testament that is still secretly operative in the religious imagination.

(4) From the philosopher's point of view, perhaps the most urgent need is for Bultmannian theology to construct a much more precise logical map of its key concepts, myth, mythology, and analogy. "Mythology," Bultmann wrote, "is the use of imagery to express the other worldly in terms of this world and the divine in terms of human life, the other side in terms of this side." But Bultmann does not want to conclude that discourse about God is always, and necessarily, mythological. To speak mythologically is to represent God as a kind of superentity, observably acting upon and interacting with natural entities. However, Bultmann has claimed (in *Kerygma and Myth*) that it is possible to speak of God's "acts" analogically, and to do so with the help of concepts borrowed once again from the field of human personal existence.

Bultmann is here in pursuit of what may well be a valuable distinction, but it has not been at all clearly articulated. The different modes of discourse about God are not rigorously defined, and thus a good deal of uncertainty is left about appropriate tests for sense and nonsense, truth and falsity, in claims about God. It is by no means obvious, for instance, whether one can really think through those existential, "analogical" utterances about God without implicitly relying upon a mythological picture of God as a superperson and superentity. Further, since both mythological discourse and analogical discourse are indirect or oblique, we need to ask whether any direct, literal talk about God is possible, or whether it is necessarily all oblique. If it must all be oblique, the problem of how we can refer to God and relate the myths and analogies to him surely becomes unmanageable. If it is not all oblique, then we still need to discover what, and how much, can be affirmed directly and literally about God. The temptation is to resort to theological makeshifts—to analyze virtually all talk about God in terms of human self-understanding, but to rely, devotionally and pastorally, upon an unanalyzed transcendent remainder, of which, however, no clear account is given in a systematic theology.

All these puzzling instabilities in Bultmann's thought are not careless or stupid blunders of reasoning. They are illuminating, disturbing indications of how immensely hard it is to steer between, on the one hand, a wholly secularized Christianity, a humanism, and, on the other, a religion of the supernatural and the miraculous.

See also Christianity; Collingwood, Robin George; Existentialism; Heidegger, Martin; Philosophy of Religion; Philosophy of Religion, History of; Religious Language.

Bibliography

OTHER RECOMMENDED WORKS

Barth, Karl, Rudolf Karl Bultmann, and Bernd Jaspert. *Karl Barth-Rudolf Bultmann, Briefwechsel, 1922–1966*. Zürich: Theologischer Verlag, 1971.

Barth, Karl, Rudolf Karl Bultmann, Bernd Jaspert, and Beoffrey William Bromiley. *Karl Barth-Rudolf Bultmann Letters, 1922–1966*. Grand Rapids, MI: Eerdmans, 1981.

Bultmann, Rudolf Karl. *Das Evangelium des Johannes*. Gottingen, Vandenhoeck & Ruprecht, 1968.

Bultmann, Rudolf Karl. *Das Urchristentum in Rahmen der antiken Religionen*. Reinbek b. Hamburg: Rowohlt, 1969.
Bultmann, Rudolf Karl. *Das Verhältnis der urchristlichen Christusbotschaft zum historischen Jesus*. Heidelberg: C. Winter, 1965.
Bultmann, Rudolf Karl. *Der zweite Brief an die Korinther*. Göttingen: Vandenhoeck und Ruprecht, 1987.
Bultmann, Rudolf Karl. *Die drei Johannesbriefe* (1967). Göttingen: Vandenhoeck & Ruprecht, 1969.
Bultmann, Rudolf Karl. *Die Geschichte der synoptischen Tradition*. Göttingen: Vandenhoeck & Ruprecht, 1967.
Bultmann, Rudolf Karl. *Die zweite Brief an die Korinther*. Göttingen: Vandenhoeck & Ruprecht, 1976.
Bultmann, Rudolf Karl. *The Gospel of John; A Commentary*. Philadelphia: Westminster Press, 1971.
Bultmann, Rudolf Karl. *The History of the Synoptic Tradition* (1963). New York, Harper & Row, rev. ed. 1968.
Bultmann, Rudolf Karl. *The Johannine Epistles: A Commentary on the Johannine Epistles*. Philadelphia: Fortress Press, 1973.
Bultmann, Rudolf Karl. *Primitive Christianity in Its Contemporary Setting* (1956). New York: Meridian Books, 1965.
Bultmann, Rudolf Karl. *What Is Theology?*. Minneapolis: Fortress Press, 1997.
Bultmann, Rudolf Karl, and Erich Dinkler. *The Second Letter to the Corinthians*. Minneapolis: Augsburg, 1985.
Bultmann, Rudolf Karl, Friedrich Gogarten, and Hermann Götz Göckeritz. *Briefwechsel, 1921–1967*. Tübingen: Mohr Siebeck, 2002.
Bultmann, Rudolf Karl, Friedrich Gogarten, and Hermann Götz Göckeritz. *Rudolf Bultmann-Friedrich Gogarten: Briefwechsel, 1921–1967*. Tübingen: Mohr Siebeck, 2002.
Bultmann, Rudolf Karl, and Bernd Jaspert. *Rudolf Bultmanns Werk und Wirkung*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1984.
Bultmann, Rudolf Karl, and Roger A. Johnson. *Rudolf Bultmann*. Collins, 1987.
Bultmann, Rudolf Karl, and Roger A. Johnson. *Rudolf Bultmann: Interpreting Faith for the Modern Era*. London; San Francisco, CA: Collins, 1987.
Bultmann, Rudolf Karl, Eberhard Jüngel, and Klaus W. Müller. *Theologische Enzyklopädie*. Tübingen: Mohr, 1984.
Bultmann, Rudolf Karl, Karlis Kundzins, and Frederick C. Grant, ed. and tr. *Form Criticism; Two Essays on New Testament Research. The Study of the Synoptic Gospels*. New York: Harper, 1966.
Bultmann, Rudolf Karl, and Andreas Lindemann. *Neues Testament und christliche Existenz: Theologische Aufsätze*. Tübingen: Mohr Siebeck, 2002.
Bultmann, Rudolf Karl, and James McConkey Robinson, ed. *The Future of Our Religious Past; Essays in Honour of Rudolf Bultmann*. New York, Harper & Row, 1971.
Bultmann, Rudolf Karl, Rudolf Zingel, and Otto Kaiser. *Gedenken an Rudolf Bultmann*. Tübingen: Mohr, 1977.
Malet, André, and Rudolf Karl Bultmann. *Bultmann et la mort de Dieu; Présentation, choix de textes, biographie, bibliographie.* Paris: Seghers, 1968.

Ronald W. Hepburn (1967)
Bibliography updated by Michael J. Farmer (2005)

BURCKHARDT, JAKOB
(1818–1897)

The Swiss cultural historian Jakob Burckhardt was born in Basel, the son of a Protestant minister. He began his university education as a theology student, but lost his faith in orthodox Christianity comparatively early and turned instead to history. He spent part of his formative years in liberal and freethinking circles in Germany; it was in Germany, too, that he discovered and worked under Leopold von Ranke, probably the most potent and lasting influence upon his future career as a historian. On his return to Switzerland in the 1840s, Burckhardt was at first attracted to the political and religious dissensions that he found there. The violence to which they subsequently led, however, was repulsive to his temperament; and he retired to Italy, having, in his own words, "given up political activity forever." Some time later he finally settled in Basel, dedicating himself, as professor of history and history of art, to the routine of teaching and lecturing that was to occupy him continuously up to the last years of his life.

Burckhardt's chief writings were all published before he was fifty: *The Age of Constantine the Great* (1852), *Cicerone* (1855), *The Renaissance in Italy* (1860), and *The History of the Renaissance* (1867). In addition to these major works, he also gave a number of lectures between 1868 and 1871 on the general study of history, the notes for which were preserved and eventually published posthumously under the title of *Weltgeschichtliche Betrachtungen* (Reflections on world history). These are remarkable, not only for the prophetic insight they display in their analysis of contemporary trends, but also for the many subtle and individual observations they contain concerning the purposes of historiography and the theoretical problems it poses. They were attended by Friedrich Nietzsche, who at the time was professor of classics at Basel and whose later essay, *The Use and Abuse of History,* bears the impress of some of Burckhardt's ideas.

Burckhardt did not regard his lectures as representing a contribution to "philosophy of history" in the then current sense. Indeed, he made it clear at the outset that he was profoundly suspicious of fashionable schemes and systems that attempted to exhibit the course of historical development as conforming to a rationally ordered pattern, and referred with special scorn to the Hegelian conception of history as the "inevitable march of the world spirit." For him such projects were the manifestation of a crude and vulgar "optimism"; they sprang from the arrogant and egotistical assumption that "our time is the con-

summation of all time" and tended to "justify" the crimes and disasters of previous ages as necessary to the promotion of what came afterward. Burckhardt thought that the role of moral judgment in history could not be spirited away in this complacent manner; but neither, on the other hand, should the historian allow his view of the past to be distorted by moral predilections peculiar to his own time and society. What was above all requisite for true historical understanding was a contemplative, disinterested sense of the abiding and tragic aspects of human existence. Only through such detachment from prevailing concerns and preoccupations could the historian transcend the barriers that separate the mental life of one age from that of another.

Burckhardt admired Arthur Schopenhauer, and he tended to extend to the historian a position in some ways similar to that which the German philosopher had reserved for the artist. It was not merely that works of art and culture provided the historian with his most fertile material for the interpretation of previous phases of human experience; history itself was (or should be) a form of art. The mechanical piling up of the results of specialized research, dear to so-called scientific historians, was not enough; there must also be "intuition," an imaginative ability to re-create the vision of life underlying the relics left by former times. To see the past in these terms was to see it as the expression of the inexhaustible creative power of the human mind—great individuals, great artistic achievements, great moments of civilization, all exemplified in different ways its potentialities. Scholarship, painstaking investigation, were indeed essential, but they must be properly used and directed. Only thus could a particular source or authority throw light on the character of a person, the significance of a style, the pervasive atmosphere of a period.

Ultimately, Burckhardt claimed, the subject of historical study was man himself, not the hypostatized abstractions of the philosophers of history. These philosophers, by implying that the historical process followed a fixed and predetermined course, betrayed a fundamental blindness to its most striking feature, the revelation of individual originality and creativity. Likewise, their "astrological impatience" to set limits to its future by talk of world plans and metaphysical goals was not only unwarranted; it failed to respect the very conditions of uncertainty and suspense that make human achievement possible. From this point of view, and insofar as the development of humankind is concerned, "a future known in advance is an absurdity."

Toward the close of the nineteenth century the tide of historical speculation began to recede. Philosophers, rather than continuing to offer sweeping interpretations of the human past, turned their attention toward examining the distinctive characteristics of historical thought and inquiry. In retrospect, Burckhardt can be seen to occupy an interesting position in this development. Though not a philosopher himself, he nonetheless anticipated in his own reflections on historical procedure some of the ideas that later found philosophical expression in the writings of Wilhelm Dilthey and Benedetto Croce.

See also Croce, Benedetto; Dilthey, Wilhelm; History and Historiography of Philosophy; Nietzsche, Friedrich; Renaissance; Schopenhauer, Arthur.

Bibliography

WORKS BY BURCKHARDT

Die Zeit Constantins des Grossens. Basel, 1852. Translated by Moses Hadas as *The Age of Constantine the Great.* New York: Pantheon, 1949.

Der Cicerone. Basel, 1855. Translated by Mrs. A. H. Clough as *Cicerone,* rev. ed. London: J. Murray, 1879.

Die Cultur der Renaissance in Italien. Basel: Schweighauser, 1860. Translated by S. G. C. Middlemore from 15th German ed. as *The Civilization of the Renaissance in Italy,* 2nd ed. London, 1890.

Geschichte der Renaissance in Italien. Stuttgart, 1867.

Weltgeschichtliche Betrachtungen. Edited by J. Deri, 2nd ed. Berlin, 1910. Translated by J. H. Nichols as *Force and Freedom: Reflections on History.* New York: Pantheon, 1943.

Gesammelte Werke, 7 vols. Basel, 1957.

WORKS ON BURCKHARDT

Duerr, E. *Freiheit und Macht bei Jacob Burckhardt.* Basel: Helbing and Lichtenhahn, 1918.

Heller, E. "Burckhardt and Nietzsche." In *The Disinherited Mind.* Cambridge, U.K.: Bowes and Bowes, 1952, Ch. 3.

Joel, K. *Jacob Burckhardt als Geschichtsphilosoph.* Basel, 1910.

Martin, A. W. O. von. *Burckhardt und Nietzsche philosophieren über Geschichte.* Krefeld, 1948.

Meinecke, F. "Ranke and Burckhardt." In *German History: Some New German Views,* edited by Hans Kohn. London: Allen and Unwin, 1954.

Trevor-Roper, H. R. "The Faustian Historian: Jacob Burckhardt," in *Men and Events.* New York: Harper, 1957, Ch. 40.

Patrick Gardiner (1967)

BURIDAN, JOHN
(c. 1300–1361)

John Buridan, or Johannes Buridanus, was a philosopher and arts master at the University of Paris. Little is known

about his early life other than that he hailed from Picardy in the north of France, most likely from the town of Béthune. As a young man he studied at the Collège Lemoine in Paris, where he was awarded a benefice or stipend for needy students, and then at the University of Paris, where he earned the degree of master of arts and received his license to teach in the 1320s. He spent his entire academic career at the University of Paris, twice serving as its rector. He was a respected figure who was often asked to settle jurisdictional disputes and assist in other matters of academic governance.

Two features of Buridan's career are distinctive. The first is that he remained a teaching master in the faculty of arts without ever moving on to take a higher, doctoral degree in theology, which was the more typical career track for philosophers at the time. Why he decided not to join the more prestigious ranks of the theologians he does not say, but given his philosophical talent and stature at the University, it is safe to assume that he had his reasons for remaining where he was. One possibility, which is suggested by some of his remarks about the relation between philosophy and theology, is that he believed philosophy to be an essentially secular enterprise, which he would have to abandon if he became a theologian. Whether this represents an important first step in the direction of modernity awaits further investigation, but at the very least, Buridan was passionately committed to the autonomy of philosophy as a discipline proper to the faculty of arts, not theology.

The other distinctive feature of Buridan's academic career is that he remained a secular cleric rather than joining a religious order such as the Dominicans or Franciscans. The popularity of these orders in the thirteenth century had revitalized the study of theology, raising it to speculative heights it has not seen since. But as the larger orders began to institutionalize the training of their novices outside the university and develop their own intellectual traditions—with Thomas Aquinas being championed by the Dominicans and Bonaventure and John Duns Scotus by the Franciscans—serious disputes arose not only within religious orders but between them, a phenomenon that led to the development of different schools of philosophy: Thomistic, Scotistic, and so on (hence the term *Schoolmen*). As a secular cleric, Buridan could safely ride above these disputes, without being obliged to defend or explain the authorities of any particular tradition. This theoretical independence can be seen in the occasionally eclectic character of his remarks.

Most of Buridan's writings are in the form of commentaries on Aristotle, whose texts were the primary subject of study in the medieval arts curriculum. These commentaries survive in two forms: *expositiones* or literal commentaries and *quaestiones* or question commentaries, both of which have their origins in the way Buridan actually taught. He would begin by giving his students a line-by-line exposition of a portion of Aristotle's text and follow this up with a problem or question raised by the passage although not explicitly discussed in it, such as whether the intellect has the capacity to recall previous thoughts, analogous to the power of memory in the sensitive part of the soul (see Aristotle, *De Anima* III.5, 430a24). Arguments for and against would be inventoried, after which Buridan would give his own—sometimes lengthy—resolution of the question, with responses to arguments on the opposite side. A similar method was used by Thomas Aquinas in composing the *Summa Theologiae*.

Buridan wrote commentaries on all of the major works of Aristotle. But because he lectured more than once on a given text over the course of his long career, some commentaries exist in more than one version, and the evolution in his thinking about a particular issue can occasionally be seen in these different versions. In addition to the commentaries, he wrote a massive logic textbook, the *Summulae de Dialectica* (Compendia of dialectics), as well as a number of shorter, independent treatises on controversial topics such as the *Tractatus de relationibus* (Treatise on relations)], *Tractatus de universalibus* (Treatise on universals), and *Tractatus de consequentiis* (Treatise on consequences). He was by any measure a prolific author.

Buridan's influence is immediately evident in the work of his younger contemporaries at Paris: Albert of Saxony, Marsilius of Inghen, and Nicole Oresme. But his commentaries and his *Summulae de Dialectica* continued to be read and commented on for several generations. Manuscripts and early printed editions of his writings were carried by his students and followers to the new universities in Heidelberg, Kraków, Prague, and Vienna, where they served as primary texts in courses on logic and Aristotelian philosophy. In this way, the *via Buridani* continued to influence European thought well into the early modern period.

LOGIC

Buridan's view of logic is best conveyed by the opening line of Peter of Spain's *Summulae Logicales* (Compendia of logics), the thirteenth-century textbook on the basis of which Buridan prepared his logical masterwork, the *Summulae de Dialectica*: "*Dialectica est ars artium, ad omnium*

methodorum principia viam habens (Dialectic is the art of arts, having access to the principles of all other inquiries)." More than just a method, logic is the grammar of philosophical discourse, the discipline whose procedures govern rational inquiry in virtually every field investigated by the arts master, from metaphysics and cosmology to natural philosophy and ethics. Buridan composed the nine treatises of his *Summulae* so that they exhibit an orderly progression of teachings based on the proposition, beginning with propositions themselves (I), moving down to the significance and referential function of their component terms (II–IV), then back up to propositions again, considered as parts of more complex patterns of reasoning: syllogisms (V), topics (VI), fallacies (VII), and demonstrations (VIII). The work closes with a series of logical exercises (IX). The order of the *Summulae* reflects Buridan's assumptions about the semantic character of human understanding, which is in turn a reflection of the metaphysical structure of creation.

Buridan is usually classed as a terminist logician. The terminists (sometimes referred to as the *moderni* or *moderns*) were a diverse group of thirteenth- and fourteenth-century philosophers who regarded the semantic properties of terms (literally, the "ends [*termini*]," or subjects and predicates, of propositions) as the primary unit of logical analysis. His main contribution was to modernize and systematize the old logic of Aristotle and Boethius using the newer techniques of the terminists, though in the process he offered innovative solutions to traditional problems in the philosophy of logic. His solutions to logical paradoxes such as the liar are still being discussed today. Consider, for example, the sentence, "Every proposition is false," assuming "that all true propositions are annihilated while the false ones remain, and then Socrates propounds only this: 'Every proposition is false' " (*Summulae* 9.8, seventh sophism). Is Socrate' proposition true or false? Buridan argues that it is false, and his reasoning shows his mastery of the semantic nuances of the question. "Every proposition," he says, "virtually implies another proposition in which the predicate 'true' is affirmed of the subject that supposits for [the original proposition]" (*Summulae* 9.8, seventh sophism). Thus, for the truth of any proposition P, it is required not only (1) that the subject and predicate terms of P stand for the same thing or things, but also (2) that P imply another proposition, "P is true," which must also be true—otherwise there would be a true antecedent and a false consequent. Accordingly, the constituent terms in the proposition uttered by Socrates—"Every proposition" and "false"—stand for the same things, since in the posited case, "all true propositions are annihilated and the false ones remain, and then Socrates propounds only this: 'Every proposition is false'." So the first condition is satisfied. But the implied proposition, "P is true" (where P is the name of "Every proposition is false"), is false because its constituent terms, "Every proposition is false" and "true," do not stand for the same thing, since *ex hypothesi*, P stands for the antecedent proposition "Every proposition is false," not for things that are true. But this gives us a true antecedent and a false consequent, and so the consequence does not hold. Therefore, the sophism is false.

METAPHYSICS

Buridan viewed metaphysics as the highest form of philosophical inquiry, yet his *Questions on Aristotle's Metaphysics* is among the shortest of his commentaries. There appear to be two reasons for this. First, he is not optimistic about the possibility of humans coming to know the ultimate nature of reality in this life because he doubts whether people are ever in a position to be acquainted with the natures or essences of things as such. Most of the time one must make do with inferences based on sense, memory, and experience, and the latter experience shows that even the firmest empirical conviction is subject to revision. Second, Buridan is adamant that metaphysics belongs to philosophy, not to theology, and hence that it cannot take its principles or starting points from Scripture or religious doctrine: "metaphysics differs from theology in the fact that although each considers God and things that pertain to divinity, metaphysics considers them only as regards what can be proved and implied, or inductively inferred, by demonstrative reason. But theology has for its principles articles [of faith], which are believed quite apart from their evidentness, and further, considers whatever can be deduced from articles of this kind" (*Questions on Aristotle's Metaphysics*, I.2). This leads him to assert the autonomy of philosophers—and implicitly of the arts masters as well—in a rather striking way: metaphysics, or philosophical wisdom, cannot be ordained by theology because its methods, which are rooted in its principles, are different. Philosophy is accordingly not inferior to theology, just different. This was an important step toward the modern view of philosophy as a secular enterprise.

Buridan was also a nominalist, though it is better to think of late-medieval nominalism as a parsimonious way of doing philosophy than as a commitment to denying the existence of real or Platonic universals. The method in Buridan's metaphysics is his logic. He tries

wherever possible to apply the *Summulae*'s analytical techniques to the interpretation of Aristotle, and his approach is critical in that it tends to view traditional questions in metaphysics as based on confusions of logic or language. Thus, when asked whether universals really exist outside the soul, he replies by clarifying the meaning of the common term *universal* with respect to its correlative terms, *individual*, *particular*, and *singular*. His rejection of realism is expressed in the same fashion: universal terms have no ultimate significate, nothing outside the soul they can make known as such. What such terms mean is other terms: the primary signification of *universal* is "predicable of many," which makes it a term of second intention, or a term of terms, since only terms are predicable. Likewise, when the term *universal* occurs in a proposition, it signifies not a *what* but a *how*, that is, how one conceives of something—in this case, that the term so designated is "indifferent to many supposits," or individuals.

Clearly, Buridan thinks that the careful and systematic analysis of language is the best way of dealing with such metaphysical problems. The trouble usually begins with untutored persons who think that each and every substantive term must correspond to a thing, or that true predication must involve the real inherence of attributes in subjects rather than making the more modest assumption that the subject and predicate terms simply stand for the same thing(s).

NATURAL PHILOSOPHY AND ETHICS

Buridan's natural philosophy and ethics are also shaped by the methods of the *Summulae*. Thus, his treatment of infinite magnitudes in his *Questions on Aristotle's Physics* focuses on clarifying the different senses of the term *infinite*: nothing is infinite if by that one means an actually existing infinite magnitude, although one can always imagine a magnitude greater than the one being considered, and do so without limit. The concept of infinity is thereby redeemed for natural science as a mode, or way of thinking.

Buridan also played a key role in the demise of the Aristotelian picture of the cosmos in the later Middle Ages. His major contribution was to develop and popularize the theory of impetus, or impressed force, to explain projectile motion. Rejecting the Aristotelian idea of *antiperistasis*—according to which the tendency of a moving projectile to continue moving (think of a ball after it has left the hand of a thrower) is due to a proximate but external moving cause (the air surrounding it, in this case)—Buridan argued that only an internal motive force, transmitted from the mover to the projectile, could explain its continued motion. The theory did not originate with Buridan, but he is perhaps the first to have seen that a force of this kind need not be self-dissipating: "[A]fter leaving the arm of the thrower, the projectile would be moved by an impetus given to it by the thrower," he says, "and would continue to be moved as long as the impetus remained stronger than the resistance, and would be of infinite duration were it not diminished and corrupted by a contrary force resisting it or by something inclining it to a contrary motion" (*Questions on Aristotle's Metaphysics*, XII.9). This is a long way from Aristotle, and not all that far from Galileo.

Despite its revolutionary implications, Buridan did not use impetus to transform the science of mechanics. He remained unapologetically Aristotelian in other respects, continuing to hold, for example, that motion and rest are contrary states of bodies. He should instead be thought of as someone who tried hard to reshape Aristotelian physics in the face of an increasingly mechanistic worldview.

Buridan's method in natural science is empirical in the sense that it emphasizes the evidentness of appearances, the reliability of a posteriori modes of reasoning, and the application of naturalistic models of explanation—such as the concept of impetus—to natural phenomena. Purely theological considerations are dismissed as irrelevant: "[O]ne might assume that there are many more separate substances than there are celestial spheres and celestial motions, viz., great legions of angels [*magnae legiones angelorum*], but this cannot be proved by demonstrative arguments originating from sense perception" (*Questions on Aristotle's Metaphysics,* II.9). Buridan concedes that an omnipotent God could deceive people in ways they could never detect, but this is tempered by his confidence, for which he cites empirical evidence, that people's ordinary powers of perception and inference are sufficiently reliable to make "the comprehension of truth with certitude possible for us" (*Questions on Aristotle's Metaphysics*, II.1). He had little patience for skeptical arguments (such as those he believed were advanced by his Parisian contemporary, Nicholas of Autrecourt), objecting that it is absurd to demand that all knowledge be demonstrable by reduction to the principle of noncontradiction. Natural philosophy is about what happens for the most part, assuming the common course of nature.

Despite Buridan's prolific output, stellar reputation, and profound influence on later thinkers, most philosophers know of him only in connection with Buridan's

Ass, the traditional example in which a donkey starves to death because it has no reason to choose between two equidistant and equally tempting piles of hay. This is doubly unfortunate because this example is nowhere to be found in Buridan's writings, though there are versions of it going back at least to Aristotle (see *De Caelo* 295b32). The best explanation of its association with Buridan is that it began as a parody of his account of free choice by later critics, who found absurd his idea that the will's freedom could consist in inaction, or more specifically, in its ability to defer or send back for further consideration any practical judgment that is not absolutely certain. But Buridan's Ass, which is apparently possessed of reason, would have surely seen the good in ceasing to deliberate once his hunger or thirst became too acute, and would have permitted his sensory appetite to lead him to whichever appeared first.

See also Impetus; Logic, History of: Medieval (European) Logic; Universals, a Historical Survey.

Bibliography

WORKS BY JOHN BURIDAN

Iohannis Buridani Tractatus de consequentiis. Edited by Hubert Hubien. Louvain, Belgium: Publications universitaires, 1976.

John Buridan on Self-Reference: Chapter Eight of Buridan's Sophismata: An Edition and Translation with an Introduction and Philosophical Commentary. Edited and translated by G. E. Hughes. Cambridge, U.K.: Cambridge University Press, 1982.

Iohannes Buridanus Quaestiones in Praedicamenta. Edited by Johannes Schneider. Munich, Germany: Beck, 1983.

"Jan Buridan, Kommentarz do *Isagogi* Porfiriusza." Edited by Ryszard Tatarzyński. *Przeglad Tomistyczny* 2 (1986): 111–195.

"Johannis Buridani, Tractatus de differentia universalis ad individuum." Edited by Sławomir Szyller. *Przeglad Tomistyczny* 3 (1987): 137–178.

John Buridan's "Tractatus de Infinito". Edited by J. M. M. H. Thijssen. Nijmegen, Netherlands: Ingenium, 1991.

Le traité de l'âme de Jean Buridan [De prima lectura]. Edited by Benoît Patar. Louvain-la-Neuve, Belgium: Editions de l'Institut supérieur de philosophie; Longueuil, Quebec: Editions du Préambule, 1991.

Johannes Buridanus, Questiones Elencorum. Edited by Ria van der Lecq and H. A. G. Braakhuis. Nijmegen, Netherlands: Ingenium, 1994.

"Buridan, On Aristotle's *Ethics*, Book X." Edited by R. J. Kilcullen. Edition available online at: http://www.humanities.mq.edu.au/Ockham/ (1996).

Ioannis Buridani, Expositio et Quaestiones in Aristotelis "De Caelo". Edited by Benoît Patar. Louvain-la-Neuve, Belgium: Éditions de l'Institut supérieur de philosophie; Paris: Peeters, 1996.

Summulae de Dialectica. Translated by Gyula Klima. New Haven, CT: Yale University Press, 2001.

WORKS ABOUT JOHN BURIDAN

Biard, Joël. *Logique et théorie du signe au XIV^e^ siècle*. Paris: Vrin, 1989.

Klima, Gyula. "Buridan's Logic and the Ontology of Modes." In *Medieval Analyses in Language and Cognition*, edited by Sten Ebbesen and Russell L. Friedman, 473–495. Copenhagen: Royal Danish Academy of Sciences and Letters—C. A. Reitzels Forlag, 1999.

Thijssen, J. M. M. H., and Jack Zupko, eds. *The Metaphysics and Natural Philosophy of John Buridan*. Leiden-Boston-Köln: Brill, 2001.

Walsh, James J. "Buridan on the Connection of the Virtues." *Journal of the History of Philosophy* 24 (1986): 453–482.

Willing, A. "Buridan and Ockham: The Logic of Knowing." *Franciscan Studies* 45 (1985): 47–56.

Zupko, Jack. "John Buridan." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Winter 2003 Edition. Available from http://plato.stanford.edu/.

Zupko, Jack. *John Buridan: Portrait of a Fourteenth-Century Arts Master*. Notre Dame, IN: University of Notre Dame Press, 2003.

Zupko, Jack. "Substance and Soul: The Late Medieval Origins of Early Modern Psychology." In *Meeting of the Minds: The Relations between Medieval and Classical Modern European Philosophy*, edited by Stephen F. Brown, 121–139. Turnhout, Belgium: Brepols, 1999.

Jack Zupko (2005)

BURKE, EDMUND
(1729–1797)

Edmund Burke, the British statesman and political philosopher, was born in Ireland to a family of modest means. His mother's family was Catholic, his father's Protestant. He was raised a Protestant and educated at a Quaker school and at Trinity College, Dublin, where he took the equivalent of a first-class honors degree in classics. He went to London to read law but was never called to the bar. He devoted most of his time to authorship and literary journalism. Robert Dodsley, a leading London bookseller of the time, loyally backed him; by 1757, Dodsley had published two books by Burke, *A Vindication of Natural Society* (1756) and *Philosophical Inquiry into the Origin of Our Ideas on the Sublime and the Beautiful* (1756), had given him employment as editor of *The Annual Register,* and had contracted to pay him £300 for an *Abridgement of the History of England.*

A Vindication of Natural Society is a satire on the views of Henry St. John Bolingbroke. It claimed to be a recently discovered work by Bolingbroke and was designed to ridicule the idea that the rise of civilized society is attended by misery and suffering. The parody was written with such conviction, however, that many

assumed it was in fact the work of Bolingbroke, and even when it was known that Burke was the author, some critics still thought it was a sincere expression of his true opinion.

Burke's book *On the Sublime and the Beautiful* is more important; indeed, it might well be said to signalize the point at which aesthetic taste in England changed from the classical formalism of the earlier years of the eighteenth century to the romanticism of the later years. Burke attacked the rationalist, classicist notion that clarity is an essential quality in great art. He argued, on the contrary, that what is greatest and noblest is the infinite, and that the infinite, having no bounds, cannot be clear and distinct. He argued that the imagination, moreover, is most strongly affected by what is suggested or hinted at and not by what is plainly stated. Burke also maintained that fear plays a large part in our enjoyment of the sublime. Such fear is diminished by knowledge, but sharpened by veiled intimations. Obscurity, not clarity, is the property of the most powerfully moving art; and, Burke added, "It is our ignorance of things that causes all our admiration and chiefly excites our passions."

Both of Burke's first two works were well received, but neither set him on the road to any further achievement. The *Annual Register* was a success, although Burke regarded it as mere hackwork. He never finished the projected *History of England.* Burke's growing interest in questions of ethics and politics provided him, in time, with an escape from the frustrations of Grub Street. He entered the House of Commons at the age of thirty-seven, and this new life brought him satisfactions he had never known in his earlier career. He became an outstanding parliamentarian; what distinguished him and made him a philosopher among politicians, however, was his capacity to look beyond the matters of the day and to articulate general principles in terms of which he believed the problems of the day should be judged.

A diligent study of Burke's letters and manuscripts brings home the extent to which his approach to politics was a religious one. What is often spoken of as his "empiricism" appears in this light to be better described as Christian pessimism. As a Christian, Burke believed that the world is imperfect; he regarded his "enlightened" contemporaries' faith in the perfectibility of man as atheistical as well as erroneous. Thus, whereas the fashionable intellectuals of his time looked for the progressive betterment of the world through the beneficent influence of Reason and Nature, Burke maintained that the moral order of the universe is unchanging. The first duty of rulers and legislators, he argued, is to the present, not to the future; their energies should be devoted to the correction of real ills, not to the promotion of an ideal order that exists only in the imagination.

Burke put great faith in the inherited wisdom of tradition. He held that the moral order of the temporal world must necessarily include some evil, by reason of original sin. Men ought not to reject what is good in tradition merely because there is some admixture of evil in it. In man's confused situation, advantages may often lie in balances and compromises between good and evil, even between one evil and another. It is an important part of wisdom to know how much evil should be tolerated. To search for too great a purity is only to produce fresh corruption. Burke was especially critical of revolutionary movements with noble humanitarian ends because he believed that people are simply not at liberty to destroy the state and its institutions in the hope of some contingent improvement. On the other hand, he insisted that people have a paramount duty to prevent the world from getting worse—a duty to guard and preserve their inherited liberties and privileges.

These considerations explain the so-called inconsistencies often attributed to Burke, who supported the movement for the independence of Ireland and the rebellion of the American colonists against the English government, but bitterly opposed the French Revolution. The reason for this seeming inconsistency was that Burke regarded the Irish movement and the American rebellion as actions on behalf of traditional rights and liberties that the English government had infringed on. The French Revolution was quite different, he argued, because it was designed to introduce a wholly new order based on a false rationalistic philosophy. Burke did not object to a resort to force as such; it was the aims of the French revolutionists to which he objected. Similarly, Burke approved of the English Revolution of 1688 because he saw it as designed to restore the rights of Englishmen and to secure the hereditary succession to the throne. The French Revolution, on the contrary, was intended to establish the so-called rights of man and the republican ideals of liberty, equality, and fraternity at the expense of personal property, religion, and the traditional class structure of a Christian kingdom.

In one of his most celebrated works, *Reflections on the Revolution in France* (1790), Burke attacked those of his contemporaries who made an abstraction of liberty, and who invited people to seek liberty without any real knowledge of what they meant by it. He claimed that he himself loved "a manly, moral, regulated liberty as well as any gentleman in France," but he would not "stand for-

ward and give praise" to an "object stripped of all concrete relations" and standing "in all the solitude of a metaphysical idea." As for equality, Burke insisted that it was contrary to nature and therefore impossible to achieve; its advocates, moreover, did "great social harm," for by pretending that real differences were unreal, they inspired "false hopes and vain expectations in those destined to travel in the obscure walk of laborious life." Burke dismissed talk of fraternity as so much "cant and gibberish"; such splendid words were simply the pretexts of the French revolutionists; the causes of the French revolution, however, were "men's vices—pride, ambition, avarice, lust, sedition."

Burke's view of the ancien régime in France was in many ways a romantic one; he was certainly no less a "man of feeling" than was Jean-Jacques Rousseau, whom he detested. But Burke was essentially a religious man living in a rationalistic age. Although he often spoke the language understood by that age—the language of calculation, expediency, utility, and political rights—he had a mind that his contemporaries, and many others, could not readily comprehend. Burke was conscious, above all things, of the reality and unavoidability of evil, and was thus led to claim that the only hope for humankind was to cling to safeguards that had stood the test of time. His hopes for bliss lay in heaven; on earth, his policy was to defend the tolerable, and sometimes the bad, against the immeasurably worse.

Until recently Burke was considered too unsystematic, too empirical, too "unphilosophical," and too much of a theorist to deserve serious attention. His conservative views were uncongenial to left-wing historians, such as Harold J. Laski and Richard Wollheim, who found him inconsistent. In 1948, however, the Sheffield Public Library (Yorkshire, England) acquired the Wentworth Woodhouse manuscripts, and the largest known collection of Burke's private papers became available to scholars for the first time since the writer's death. The study of these papers did much to enhance Burke's reputation as a political philosopher of signal importance and originality.

See also Aesthetics, History of; Bolingbroke, Henry St. John; Political Philosophy, History of; Rousseau, Jean-Jacques; Social and Political Philosophy; Traditionalism.

Bibliography

WORKS BY BURKE

*Works,*16 vols. Edited by F. Lawrence and W. King. London: F. C. and J. Rivington, 1803–1827.

Correspondence of Edmund Burke, 10 vols. Edited by Thomas W. Copeland et al. Chicago: University of Chicago Press, 1958–1978.

The Writings and Speeches of Edmund Burke, 12 vols. Edited by Paul Langford et al. Oxford: Clarendon Press, 1981–.

WORKS ON BURKE

Cobban, A. *Edmund Burke and the Revolt against the Eighteenth Century.* London: Allen and Unwin, 1929. Adds new thought on the resemblances between Burke's thought and that of Rousseau.

Cone, C. B. *Burke and the Nature of Politics.* Lexington: University of Kentucky Press, 1957. A valuable introductory study from a modern standpoint.

Copeland, T. W. *Our Eminent Friend, Edmund Burke: Six Essays.* New Haven, CT: Yale University Press, 1949. Essays by a literary historian and leading Burke scholar.

Lock, F. P. *Edmund Burke,* 2 vols. Oxford: Clarendon Press, 1998–.

MacCunn, J. *The Political Philosophy of Burke.* London: Arnold, 1913. A useful traditional reading of Burke's philosophy.

Magnus, P. *Edmund Burke.* London: Murray, 1939. A reliable short biography.

O'Brien, Conor Cruise. *The Great Melody: A Thematic Biography and Commented Anthology of Edmund Burke.* Chicago: University of Chicago Press, 1992.

Parkin, Charles. *The Moral Basis of Burke's Political Thought.* Cambridge, U.K.: Cambridge University Press, 1956. Stresses the importance of religion in Burke's political philosophy.

Stanlis, Peter J. *Edmund Burke: The Enlightenment and Revolution.* New Brunswick, NJ: Transaction, 1991.

Maurice Cranston (1967)
Bibliography updated by Philip Reed (2005)

BURLEY, WALTER

(c. 1274–c. 1345)

Walter Burley, renowned logician, natural philosopher and theologian, was born in 1274 or 1275, perhaps at Burley-in-Wharfedale or Burley, near Leeds, in Yorkshire, England. He studied and taught both at Oxford (c. 1294–c.1309) and at the University of Paris (c.1309–1327). Based in England from 1327–1341, he perhaps spent his last years in retirement in southern France and Italy (1341–1344).

OXFORD

Burley was a master of arts by 1301 and is mentioned as a fellow of Merton College in 1305. He appears to have heard John Duns Scotus lecture on the *Sentences*, probably in the academic year 1298–1299, and adopts some Scotistic positions in later works: that being *qua* being is the primary and adequate object of the intellect, and that the intellect understands the singular as singular. If Bur-

ley began to study theology at Oxford, he and William of Ockham, whose studies began c.1307–1308, may have been fellow students. Burley's writings from this period, as Jan Pinborg (1937–1982) has rightly observed, "comprise an almost complete course of logic," including *Quaestiones in librum Perihermeneias* and *Quaestiones super librum Posterior Analytics*, as well as treatments of specific topics, *De suppositionibus* and *De consequentiis*. There are commentaries on Aristotle's natural philosophy as well, including *Questions on the De anima of Aristotle*, Book 3.

PARIS

Burley's career in Paris, assuming some prior study of theology, could be reconstructed as follows. Between 1309 and 1314 he was an *auditor* of lectures on the scriptures and the *Sentences* of Peter Lombard, from 1314 to 1317 a *biblicus* (lecturer) on the scriptures, and from 1317 to 1318 a *sententiarius* although his lectures on the *Sentences* are lost. The *Tractatus Primus*, however, recounts a controversy on accidental form with his master, Thomas Wilton, which arose out of his *principium* on Book IV. Its argumentation exhibits a layering of logic and physics in a way that makes Burley a precursor of the Oxford *calculators*, such as Richard Swineshead and John Dumbleton (fourteenth century). In support of his claim that contrary forms, such as hot and cold, belong to the same ultimate species, he argues first *from logic* that things equidistant from an extreme are of the same species. Then, *from Aristotle in natural philosophy*, he argues that if a cooled body is immediately reheated, at some instant, *B*, preceding the first instant the body is cold, *A*, it will have a degree of heat, and at some instant, *C*, succeeding *A*, it will have a degree of cold, both of which degrees will be formally equidistant from maximum heat and thus in the same species. This argument also reflects contemporary debates over first, the *latitude* of forms, the intensive range of possible degrees that an instance of a species of quality may possess; and second, the *first and last instants* of change, the subject also of his *disputatio* at Toulouse, *De primo et ultimo instanti* of the same period.

In 1321, now a priest, he received his last leave of absence for two years of study and had completed his studies by the end of 1323 at the latest. He is referred to as *doctor of sacred theology* in 1324. His teaching career was short since he had left Paris by the beginning of 1327.

BURLEY AND OCKHAM

Perhaps in the same year (1317–1318) that Burley was lecturing on the *Sentences* at Paris, William Ockham was doing likewise at Oxford. It is clear that from his first exposure to Ockham's *Sentences* commentary, Burley found it necessary to oppose him on a number of important issues in logic and natural philosophy. It was not a one-sided engagement. Ockham borrows from Burley's *Tractatus primus* (before 1324) in his *Quaestiones* on the *Physics*, which Burley in turn uses and criticizes in his own final commentary on the *Physics*, the first six books of which were written after 1324–7. In the *Summa logicae* Ockham both uses and attacks Burley's *De suppositionibus*. Burley counterattacks in his second version (after 1323) of *De puritate artis logicae*. While Ockham's *Logic* is organized in the traditional way around terms, propositions, and arguments, Burley's is organized around the general rules of consequences, thus giving priority to propositional logic.

Burley's explanation of the supposition of terms differs from Ockham's, who holds that first, universals do not exist in re, and second, that they are not constitutive parts of the essence of individuals. On the contrary, Burley holds that universals do exist in re although not apart from singulars. Therefore, according to Burley, when the term *human* in a sentence has *simple supposition* or stands for what is common or universal, it stands for what it primarily signifies: the humanness in Socrates or Plato. For Ockham, however, when *human* has simple supposition, it stands for a common concept, humanness in the mind. The only thing a term can signify or refer to is the individual, for instance when *human* supposits *personally* for Socrates, Plato, and so on. Burley eventually ceded ground to Ockham on the issue of universals as constitutive parts, holding that the universal form merely discloses the individual's essence (for instance, human). Ockham's position that universals are only general concepts implies that science, which is of the universal, must be about spoken, written, and mental propositions while for Burley, science is founded on *real propositions*, that is, propositions whose subjects and predicates are real entities, either singular or universal, but whose copulas are purely mental.

As well as resisting Ockham's reduction of res to singular things, Burley objects to Ockham's reduction of Aristotle's categories to substance and quality. In his *De formis*, (c. 1324–1326), he holds that quantity is a form separate from the quantified body, and he also argues that motion is a form over and above the body in motion, increased and decreased by a succession of specifically distinct forms (*De intensione et remissione formarum*, written after 1323). This explanation, which can be calls a *succession* theory, extends to all changes in the degree of a qual-

ity a thing may possess: how the just person comes to have more justice, or that something cold becomes somewhat hot. Every increase in justice or heat, every acceleration of motion, results from the acquisition of a new, more perfect form and the loss of the old, less perfect form.

ENGLAND

Burley's departure from Paris was coincident with the coronation of Edward III (1312–1377), who sent him with a deputation in February 1327 to the papal court in Avignon and again in 1330, now as one of the king's *beloved clerks*, men in the royal service, usually of humble beginnings, who were often the king's agents on diplomatic missions. Again, from September 1338 until Easter 1339, Burley went "beyond the seas on the king's service" (*Calendar of Patent Rolls*, 1338–1340, p. 123).

Burley's academic career ended when he left Paris, and it seems that he had no significant scholarly projects in hand during the next seven years. However, some time after Richard Bury was enthroned as bishop at Durham in 1334, Burley became a member of his household. Bury's patronage and the intellectual energy of the circle he gathered around him would fuel Burley's renewed career as a scholar.

Between 1334 and 1337 Burley completed a commentary on Books 1–6 of the *Ethics*, added Books 7 and 8 to his final commentary on the *Physics*, and revised his commentary on the *Ars vetus*. He began to revise the commentary on *Ethics* 1–6 and add a commentary on 7–10 in 1338–1339. In the commentaries on the *Physics* and *Ars vetus* are found Burley's references to the *moderni*, those thinkers encountered first during his Paris years, who threaten the purity of the font of all philosophy: Aristotle. The doctrines that Burley identifies as being those of the *moderni* are not confined to any single philosophical discipline, and appear, by Burley's account, to form a systemic threat to philosophy itself. His commentary on Aristotle's *Politics*, begun in 1338–1339, is, along with his *Ethics* commentary, heavily dependent on Thomas Aquinas's expositions of those works (written between 1269–1272). Nevertheless, they contain doctrines original with Burley, for example, in the *Politics*, that of the "co-rulership" of kings with those who are "their friends and the friends of the government" (fol. 186*r*) and doctrinal divergences from Aquinas, for example, in the *Ethics*, the role of the speculative intellect in understanding the precepts of natural law (1500, fol. 103*r*).

Upon completion of the four expositions of Aristotle (c.1340), Burley, who was now in his mid-sixties, appears to have sought some disengagement from the rigors and antagonisms of scholarly life, which may have led to his journey to Italy, probably in 1341.

SOUTHERN FRANCE AND ITALY

In 1341 Burley engaged in a *disputatio de quolibet* in the arts faculty at Bologna, an event that has been connected with his supposed Averroism. Burley was not an Averroist, however, if this term implies someone who adopts positions contrary to the Christian faith on the authority of Averroes. This is clear from the beginning of his career in his questions on *De anima*, Book 3, where he concludes that "neither is the material intellect one in all, nor also the agent intellect" (3.44). Then in Paris, where his master was the Averroist Wilton, his short work *De potentiis animae* reiterates this position.

The *De vita et moribus philosophorum* was long thought to have been the fruit of Burley's retirement in southern Europe. However, large sections from it are found in a manuscript dated 1326, when Burley was in Paris, which, together with the claim that no attribution of the work to him is earlier than the fifteenth century, has led to a presumption against Burley's authorship of this immensely popular work. Nevertheless, this evidence is not conclusive, and given his habits of appropriating large amounts of text from other authors and frequently reworking his own texts, it is not impossible that the *De vita et moribus philosophorum* passed through Burley's hands at some point in its history.

On 23 November 1343, Burley was in Avignon to present a copy of his commentary on the *Politics* to his old Parisian acquaintance Pierre Roger, now Clement VI (1291–1352). This gift, complete with an elegant letter and a miniature showing the presentation, could have been both in appreciation and expectation of further patronage. Indeed, Burley obtained the rectory at Great Chart, Kent, on 19 June 1344, the last date he is known to have been alive.

Walter Burley exerted considerable influence both on his contemporaries and on philosophical thought into the sixteenth century, to which the number of early printed editions of his commentaries on Aristotle testify. This influence may be attributed, firstly, to the originality and the clarity of the positions he maintained in the controversies of his day, both in logic and natural philosophy. He contributed significantly to the debates concerning supposition theory, consequences, *obligationes*, and *sophismata*. In natural philosophy his theory of the *first and last instants* of change, which distinguishes between *permanent* and *successive* things or states, becomes a stan-

dard view, and the *succession* position, which he defends in his classic work, *On the Intension and Remission of Forms*, is frequently cited, being both opposed and defended, into the sixteenth century.

His skill at the traditional exercise of commentary on Aristotle was also acknowledged. In glossed Latin manuscripts of Aristotle and Averroes, he is one of the commentators most frequently cited, especially in connection with the *Ethics*, *Politics*, *Physics*, and logical works of Aristotle. In addition, manuscripts of Burley's commentaries on these works had a wide circulation. Early printed editions of an important collection of *auctoritates* of Aristotle and other philosophers carry his textual comments, along with those of Averroes, Robert Grosseteste, Albert the Great, and Thomas Aquinas. A revival of interest in Burley's thought, particularly his logic and natural philosophy, was underway by the 1960s, and earlier assessments of him as an unworthy opponent of Ockham have not survived a closer study of his work, which has revealed its originality and depth.

See also Albert the Great; Aristotle; Averroes; Averroism; Duns Scotus, John; Grosseteste, Robert; Peter Lombard; Swineshead, Richard; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS BY BURLEY

De formis, edited by F. J. D. Scott. Bayerische Akademie der Wissenschaften, 1970.

Commentary on *Ethics*. Venice, 1481, 1500.

De intensione et remissione formarum. Venice, 1496.

Commentary on *Logica Vetus*. Venice, 1497.

Commentary on *Physics*. Venice, 1509, 1524.

De vita et moribus philosophorum, edited by H. Knust. Stuttgard: Tübingen, 1866.

De puritate artis logicae tractatus longior, with a revised edition of the *Tractatus brevior*, edited by P. Boehner. St. Bonaventure, NY: Franciscan Institute Publications, 1955.

De primo et ultimo instanti, edited by H. and C. Shapiro, *Archiv fur Geschichte der Philosophie* 47 (1965): 157–173.

De potentiis animae, edited by M. J. Kitchel. *Mediaeval Studies* 33 (1971): 85–113.

De suppositionibus, edited by S. F. Brown. *Franciscan Studies* 32 (1972): 15–64.

Quaestiones in librum Perihermeneias, edited by S. F. Brown. *Franciscan Studies* 34 (1974): 200–295.

De consequentiis, edited by N. J. Green-Pedersen. *Franciscan Studies* 40 (1980): 102–166.

Questions on the De anima of Aristotle, edited by E. A. Synan. Leiden: E. J. Brill, 1997.

On the purity of the Art of Logic: The Shorter and the Longer Treatises. Translated by P. V. Spade. New Haven, CT: Yale University Press, 2000.

Quaestiones super librum Posterior Analytics, edited by M. Sommers. Toronto: Pontifical Institute of Mediaeval Studies, 2000.

WORKS ON BURLEY

Conti, A. D. "Ontology in Walter Burley's Last Commentary on the *Ars vetus*." *Franciscan Studies* 50 (1990): 120–176.

De Rijk, L. M. "Burley's So-Called *Tractatus primus*, with an edition of the additional quaestio *Utrum contradictio sit maxima oppositio*." *Vivarium* 24 (1996): 61–91.

Krieger, Gerhard. "Studies on Walter Burley 1989–1997." *Vivarium* 37 (1) (1999): 94–100.

Martin, C. "Walter Burley." In *Oxford Studies Presented to Daniel Callus*, Oxford Historical Society, 194–230. Oxford: Clarendon Press, 1964.

Nederman, C. J. "Kings, Peers, and Parliament: Virtue and Corulership in Walter Burley's *Commentarius in VIII libros politicorum Aristotelis*." *Albion* 24 (3) (1992): 391–407.

Ottman, Jennifer, and Rega Wood. "Walter Burley: His Life and Works." *Vivarium* 37 (1) (1999): 1–23.

Pinborg, J. "Walter Burley on Exclusives." *English Logic and Semantics*, Acts of the 4th European Symposium on Medieval Logic and Semantics. Leiden: Nijmegen, 23–27 June 1979. Reprinted in J. Pinborg. *Medieval Semantics: Selected Studies on Medieval Logic and Grammar*. London: Variorum, 1984.

Sylla, E. D. "Infinite Indivisibles and Continuity in Fourteenth-Century Theories of Alteration." In *Infinity and Continuity in Ancient and Medieval Thought*, edited by N. Kretzmann, 231–257, 322–330. Ithaca: Cornell University Press, 1982.

Sylla, E. D. "Medieval Concepts of the Latitude of Forms: The Oxford Calculators." *Archives d'Histoire Doctrinale et Littéraire du Moyen Âge* 40 (1973): 223–283.

Sylla, E. D. "Ockham and the Mertonians." *History of the University of Oxford*. Vol. 1: *Early Oxford Schools*, edited by J. I. Catto, 607–558. Oxford: Clarendon Press, 1984–2000.

Sylla, E. D. "*Repertorium Mertonense*." *Mediaeval Studies* 31 (1969): 185–208.

Weisheipl, J. A. "Ockham and Some Mertonians." *Mediaeval Studies* 30 (1968): 174–188.

Wood, Rega. "Studies on Walter Burley, 1968–1988." *Bulletin de Philosophie Médiévale* 30 (1988): 233–250.

Mary Sommers (2005)

BURTHOGGE, RICHARD

(c. 1638–c. 1698)

Richard Burthogge, the English physician and idealist philosopher, was born in Plymouth. After taking an arts degree at Lincoln College, Oxford, he studied medicine at the University of Leiden and returned to his native country to practice near Totnes in Devonshire. Of pacific and conciliatory disposition, he seems to have wavered in the religious controversy between Catholicism and Puritanism, and in philosophy, between Lockean sensationalism and Cambridge Platonism. He distinguished between

heresy and error, maintaining that the former "must be eradicated," but the latter tolerated for humanity's sake. His life is obscure, and little is known of it beyond that information revealed in his writings, which have a certain importance as anticipations of Immanuel Kant.

We know the world, according to Burthogge, only through our own ideas, and these do not give us its real nature. On the contrary, our ideas transform the nature of things into qualities that are purely subjective. Similarly, our values are our own; and such relative judgments as those involving categories of cause and effect, or whole and part, are arrived at through the constitution of our minds, not discovered embedded *in rerum natura.* The things themselves, though remaining unknowable, nevertheless cause ideas to arise in our minds. Here Burthogge foreshadowed Kant's paradox of the relation between *noumena* and *phenomena.* Burthogge's view that the human mind projects relations into the external world exemplifies his Neoplatonic streak. However, this strain was accompanied by a Lockean one which led him to assert that no confidence could be placed in an idea contradicted by sensation. Burthogge thus seems to have accepted John Locke's theory of two kinds of ideas, those of sensation and those of reflection.

For Burthogge, there were also two kinds of truth—metaphysical and logical. Metaphysical truth is found in the conformity between our ideas and those in the mind of God; logical truth, in the conformity between our ideas and the things of which they are ideas. We cannot apprehend the former kind of truth; but since the latter involves knowing the unknowable, logical truth is reduced to consistency. Burthogge would not accept the doctrine of innate ideas, because if we had such ideas, we would be able to discover truth through introspection alone. He asserted dogmatically that there is a coherent system of ideas, duplicating the system of things, even though no individual possesses it. This system, he maintained, exemplifies God's ideas.

In his treatise on the soul of the world, Burthogge supported the Neoplatonic concept of a plastic nature permeating the universe and accounting for its "harmony." This is breathed into things by God himself but is not to be identified with God. If nothing else, this treatise is valuable as an example of the philosophy of nature which was acceptable to learned men of the time.

Burthogge, in sum, is one of the anomalies of the history of philosophy. He advanced startlingly "modern" ideas, side by side with fantasies no longer taken seriously.

See also Cambridge Platonists; Error; Idealism; Ideas; Kant, Immanuel; Locke, John; Neoplatonism; Sensationalism.

Bibliography

WORKS BY BURTHOGGE

Organum Vetus et Novum. London, 1678.

An essay upon Reason, and the Nature of Spirits. London: Dunton, 1694.

Of the Soul of the World, and of Particular Souls. London, 1699.

Landes, Margaret W., ed. *The Philosophical Writings of Richard Burthogge.* Chicago and London: Open Court, 1921. Contains reprints of the three above works by Burthogge as well as a valuable introduction, notes, and bibliography.

WORKS ON BURTHOGGE

Boas, George. *Dominant Themes of Modern Philosophy,* pp. 253–259. New York: Ronald Press, 1957.

Cassirer, Ernst. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit.* Vol. 1, pp. 464–473. Berlin: Cassirer, 1906.

Grünbaum, Jacob. *Die Philosophie Richard Burthogges.* Unpublished PhD diss., Bern, 1939.

L[ee], S. L. "Burthogge, Richard." In *Dictionary of National Biography.* Vol. 7, p. 453. London: Smith Elder, 1885–1900.

Lovejoy, A. O. "Kant and the English Platonists." In *Essays Philosophical and Psychological in Honor of William James,* pp. 265–302. New York, 1908.

George Boas (1967)

BUSINESS ETHICS

Discussions of ethics and business trace back to the writings of Plato and Aristotle and persist in the modern philosophical writings of Karl Marx, John Rawls, and others. Although business ethics as a specialized field of study did not emerge until the 1970s, it has grown sharply since. Philosophers, political scientists, business academics, and social psychologists have written systematically about a variety of issues such as the moral status of the corporation, the ethical foundations of the market, fairness in advertising, bribery, corporate governance, human rights and multinational corporations, and business obligations to the environment. During that time, rival theories for interpreting business ethics have emerged and been debated.

Traditional philosophers such as Plato, Aristotle, Aquinas, and Kant discuss issues of the right and wrong in economic activity. They sometimes examine specific business ethics puzzles, including the ethics of the profit motive, just price in trade, usury in lending, and ethics in negotiation. Thomas Aquinas writes at length about the

question raised first by Cicero of whether a grain merchant carrying grain to a community stricken by famine is obliged to reveal to the townspeople that other merchants behind him are bringing more grain. (Aquinas concludes that, contra Cicero, the merchant is not so obliged because no businessperson has an obligation to make a prediction which, if it turned out to be false, would rob him of a "just" price.) Moreover, questions about broad economic design are ubiquitous in the history of philosophy. For example, the issue of the communal ownership of property (in modern terms, communism and socialism) was first brought into sharp relief by Plato, was critiqued by Aristotle, and has been the subject of bitter controversy ever since.

For convenience, it is helpful to conceive business ethics as having three parts, where each part corresponds to the level of entity being analyzed: namely,

1. Individual businesspersons: including employees, entrepreneurs, investors, traders, and consumers

2. Business systems, including economic systems, cultural norms, and regulatory and judicial systems.

3. Business organizations, including corporations, trade associations, and international financial organizations such as the WTO, the World Bank, and the IMF.

Each of these three entities gives rise to both questions of right and wrong (normative issues) and to questions of fact (empirical issues). Because empirical issues are not, properly speaking, philosophical ones, and despite the fact that a large and important empirical literature now exists (authored by sociologists, economists, and business academics), this article will not attempt to analyze and explain that empirical literature.

INDIVIDUAL BUSINESSPERSONS

Philosophers have debated the issue of the individual's pursuit of money and profit for centuries. Plato famously denied top-status positions of ruler or guardian in his ideal state, the Republic, to business people (indeed to all owners of property) out of fear that their pursuit of wealth would corrupt their political virtue. It remained for the eighteenth-century philosopher, Adam Smith, author of the *Wealth of Nations*, to make the pursuit of profit at least moderately respectable: "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard of their own interest. We address ourselves not to their humanity, but to their self-love and never talk to them of our own necessities, but of their advantage." (p. 13.)

Smith meant to draw attention to the fact that efficient economic transactions frequently rely on self-interested or profit-oriented motives rather than more noble motives such as benevolence. In his view, then, our shared goal of achieving a healthy, efficient economy justifies a significant amount of profit-seeking and self-interested activity in business. His well known "invisible hand" provides a metaphor for explaining how free markets seem to direct the inevitable, if regrettable, self-interest of businesspersons toward the common good.

One's ethical evaluation of profit-seeking by businesspersons may be influenced by one's antecedent commitments to ethical theory. Smith's invisible hand relies heavily on consequential considerations: Individual acts and motives are judged ethically through their consequences. For Adam Smith, then, we should sometimes tolerate darker, self-interested motives in business so long as the consequences produce social benefits. Yet a nonconsequential approach to ethics—one placing more emphasis on the quality of the motive or the principle of the individual's action—lacks appeal directly to such a practical justification. A nonconsequential approach must justify profit-seeking, if at all, by nesting the profit motive under other, less selfish motives, such as attempting to benefit one's family, one's community, or society by way of pursuing profit.

Critics have objected to a broad, self-centered view of business because it appears to presume selfishness or, at the very least, psychological egoism. The focus in much of modern economics is upon developing increasingly sophisticated conceptual mechanisms to maximize the achievement of economic goods such as money, market share, or profits, all of which seem to exclude the pursuit of "higher" interests such as benevolence, social welfare, and environmental integrity. Even well-known economists such as Amartya Sen have asserted that the rational economic man, *homo economicus*, is dangerously close to being a "rational fool." Opposing economists respond, however, that the maximization of individual preferences can easily include the satisfaction of other-oriented preferences such as helping the poor or protecting the environment. A businessperson may simply prefer saving the environment to maximizing his income. Whether such other-oriented preferences can be subsumed comfortably within the mathematically inclined methods that dominate modern-day economics remains hotly debated.

BUSINESS SYSTEMS

Disputes are common about the extent to which self-interested motives are acceptable in economic behavior. These disputes overlap with others about the desirability of forms of business systems. Just as Adam Smith did, modern economists often stress the societal benefits of free, self-interested market activity. They note that markets free from government interference encourage free exchanges among individuals, and, in turn, business productivity. A realm of perfectly free exchanges, indeed, is often said to establish a condition called "Pareto Optimality": a state in which no one can be made better off without someone being made worse off.

Not surprisingly, then, debates in business ethics have frequently centered on the assumptions of traditional economic theory. Microeconomic theory (which constitutes a part of what is sometimes called neoclassical economic theory) views market participants as rational agents seeking to maximize their own utility. In more recent economic writings "utility" is interpreted to mean the maximal satisfaction of one's individual preferences.

Whether economic theory contains an embedded bias towards selfishness or not, most economists agree that market participants can encounter situations where a businessperson's rational self-interest collides with the social welfare. One of the most notable of these situations is the "prisoner's dilemma" discussed by game theorists, wherein rational self-interest leads each player to defect in certain contexts where cooperation is clearly the best long-term strategy for all. Because prisoner's-dilemma situations are believed to arise frequently in business transactions, it follows that even fully self-interested businesspersons should have an interest in developing techniques of cooperation, both for themselves and others. Indeed, some philosophers have even argued that nearly all morality can be derived from such rational pursuit of self-interest through cooperation.

Others theorists argue that business ethics is simply impossible so long as market freedom is the dominant value. They assert that, in addition to problems such as the prisoner's dilemma, persistent discrimination, sexual harassment, environmental pollution, false advertising, financial scandals, child labor, and bribery require a more of a "visible hand" (usually government's) than an "invisible" one.

Nonetheless, even defenders of heaver regulation of business grant that often law is relatively impotent in ensuring business ethics. For example, law tends to lag behind the knowledge emerging in an industry, so that it often comes too late to correct abuse. Scientists in the asbestos industry in the United States knew about the dangers of asbestos long before laws could be drafted to regulate asbestos harm. Moreover, laws tend to apply to the jurisdiction from which they emanate. Hence, U.S. or German law is nearly powerless to control multinational corporations operating in host countries. This point has special force in many developing host countries where laws are unsophisticated and poorly enforced.

Conflicting cultural values can frustrate ethical decisions. For example, in countries where "grease" payments are common, are businesspersons justified in paying customary bribes to government officials? Or consider issues of human rights. In countries where educational opportunities are inadequate, is it acceptable to hire a fourteen-year-old for full-time employment? Does it make a difference that, as sometimes happens, the majority opinion among adults in a given country holds that child labor is ethically acceptable? Business ethicists have proposed a variety of theories to help solve such dilemmas. Most deny that all employment conditions between the home and host countries of the corporations must be comparable; if that were true, it is argued, employees would, for example, receive exactly the same pay (or at least the same pay adjusted for cost-of-living differences) for the same work. But such wage parity would freeze out almost all foreign investment by multinational corporations in the developing world. Instead, the dominant approach has been to specify a floor of rights that apply to labor conditions and that all corporations must respect.

BUSINESS ORGANIZATIONS

Some disagree that a corporation can ever be "responsible" or "irresponsible." They note that corporations have exceedingly narrow personalities; they are chartered for the purpose of making money for their investors. They have, in the words an English jurist, "no pants to kick or soul to damn." Can such organizations be said to have a conscience or moral responsibility? A few theorists regard the corporation as analogous to a large bureaucratic machine and for this reason hold it to be misleading to speak of a corporate "conscience." In turn, they reject the very idea of corporate ethical responsibility. Only individual businesspersons, not corporations, are the true bearers of ethical responsibility. They thus deny moral agency to the corporation, denying that a corporation attains the status of an actor for which such moral predicates as "is responsible" and "is blameworthy" are appropriate. In contrast, theorists who see the corporation as either a large, abstract "person" (the corporation in most

legal systems is regarded as a *persona ficta*, a fictional person) or an organization possessing a decision-making structure capable of rational deliberation are called moral-agency theorists. They believe that corporations *are* capable of behaving responsibly or irresponsibly.

Assuming, then, that the corporation is even the kind of thing that can behave responsibly or irresponsibly, the question next arises about what a corporation's "being responsible" means. Three major approaches to this question have been offered. These may be labeled: the Classical Framework, the Stakeholder Framework; and the Social Contract Framework.

THE CLASSICAL FRAMEWORK

The "classical" framework asserts that the moral responsibility of the corporation is nothing other than maximizing profits for its investors. This approach is associated with modern economic theory and writings of Frederich Hayek and Milton Friedman. The view holds that the sole moral responsibility of the corporation, and in turn of the managers who serve as agents for the shareholders, is to enhance the interests *only* of the owners of the corporation, the shareholders. The corporation is often seen by its classical defenders as a nexus of contracts among free-acting individuals whose peculiar advantage lies in its ability to reduce transaction costs among participants by, for example, offering organizational remedies in lieu of expensive, individual contracts among individuals.

Critics of the this approach are quick to point out that corporate executives are not publicly elected officials and as such are poor choices for shouldering decision-making promoting the common good. Indeed, often corporate executives have been associated with bad choices, as when large U.S. companies in Chile decades ago helped unseat the country's democratically elected president. Do we really want, these critics ask, to entrust corporate officials with the common good?

THE STAKEHOLDER FRAMEWORK

On the stakeholder theory, managers have obligations primarily to shareowners but also have certain ethical obligations to other groups called "stakeholders"—those who have a stake in the corporation's activity, including customers, stockholders, employees, and people who live in areas affected by the corporation. Disagreements exist about precisely who should be included as stakeholders, but almost all theorists agree that three principal groups of stakeholders are customers, employees, and stockholders. Hence, the stakeholder framework agrees with the classical framework in assigning special importance to the interests of stockholders. The difference between the stakeholder view and the classical view, however, is that stakeholder theorists do not limit the responsibilities of corporate managers entirely to satisfying stockholder interests. Managers, in turn, must make tradeoffs among the interests of the corporation's stakeholders if they are to manage well. Some stakeholder theorists argue that by working to enhance the interests of all stakeholders, the company will also maximize the long-run interests of the stockholders. But other theorists disagree, arguing that some stakeholders must inevitably receive less in order for the stockholder to achieve a maximum return on his investment.

THE SOCIAL CONTRACT/SOCIAL CONTRACTS FRAMEWORK

This view construes corporate and managerial obligations in terms of implicit "contracts" that exist in and among companies, industries, political units, and other relevant economic communities. For example, it has been argued that an implicit "social contract" exists between corporations and society requiring that corporations refrain from exploiting their workers or from destroying the environment; in return for the special favors it receives from society—unlimited longevity (because in most legal systems a corporation is a "persona ficta" or fictional person, it never dies) and limited liability (investors in corporations are responsible for the actions and debts of the corporation only up to the extent of their invested money). In a similar vein, it has been argued that an implicit social contract exists in most societies requiring that jobs and advancements allocated by a consideration of the qualifications of the applicant rather than his or her gender or race. Beginning in the 1990s, the idea of a social contract was extended by some to include the possibility of a multiplicity of social contracts, interpreted as the implicit set of agreements that exist within and among communities of economic actors, including corporations, trade associations, unions, industries, and professional associations.

Other business ethics issues arise for for-profit corporations. One of these is the factual question of whether a corporation that has better ethics will make more money in the long run than a corporation with worse ethics. Scores of empirical studies on this topic have been conducted, although the answer remains elusive. There is also the question of how a good corporation should be structured. What form of corporate governance should a corporation adopt? Should it include employees on its

board of directors? Should employees participate in the management of the corporation, and should they perhaps be given automatic status as shareholders?

Lurking in the backdrop of many discussions of corporate ethics is the issue of what power, if any, managers should have in making ethical decisions. Suppose, for example, that competitive market forces eclipse any moral "space" that managers might have. In such an instance, the entire notion of "business ethics" seems irrelevant. If "ought implies can" and if business managers are captive to the dictates of the market, then how can one say that they "ought" to behave well? On this view, the only way to reform business behavior is to change the surrounding market or regulatory environment—that is, to force business to recognize that its self-interest lies in ethical behavior. Most business ethicists, however, agree that corporations have at least some discretionary space. The empirical debate centers on how much.

See also Applied Ethics; Aristotle; Cicero, Marcus Tullius; Engineering Ethics; Ethics and Economics; Kant, Immanuel; Marx, Karl; Philosophy of Economics; Plato; Rawls, John; Sen, Amartya K.; Smith, Adam; Thomas Aquinas, St.

Bibliography

Arrow, Kenneth J. "Social Responsibility and Economic Efficiency." *Public Policy* XXI.3 (1973).

Coase, R. H. "The Nature of the Firm." In *The Nature of the Firm: Origins, Evolution, and Development*, edited by Oliver E. Williamson and Sidney G. Winter, 18–33. New York: Oxford University Press, 1991.

Donaldson, Thomas. *The Ethics of International Business*. New York: Oxford University Press, 1989.

Donaldson, Thomas, and Thomas Dunfee. *Ties That Bind: A Social Contracts Approach to Business Ethics*. Cambridge, MA: Harvard Business School Press, 1999.

Frank, Robert H. "Can Socially Responsible Firms Survive in a Competitive Environment?" In *Codes of Conduct: Behavioral Research into Business Ethics*, edited by David M. Messick and Ann E. Tenbrunsel, 86–103. New York: Russell Sage Foundation, 1996.

Freeman, R. Edward. *Strategic Management: A Stakeholder Approach*. Pitman Series in Business and Public Policy. Boston: Pitman, 1984.

Jackall, Robert. *Moral Mazes: The World of Corporate Managers*. New York: Oxford University Press, 1988.

Mitchell, R. K., B. R Agle, and D. J. Wood. "Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts." *Academy of Management Review* 22 (1997): 853–886.

Preston, Lee E., and Duane Windsor. *The Rules of the Game in the Global Economy: Policy Regimes for International Business*. 2nd ed. Norwell, MA.: Kluwer Academic Publishers, 1997.

Sen, Amartya. "Does Business Ethics Make Economic Sense?" *Business Ethics Quarterly* 3 (1993): 45–54.

Sen, Amartya Kumar. *On Ethics and Economics*. New York: Blackwell, 1987.

Smith, Adam, and D. D. Raphael. *The Wealth of Nations*. New York: Knopf, 1991.

Solomon, Robert C. *Ethics and Excellence: Cooperation and Integrity in Business*. New York: Oxford University Press, 1992.

Trevino, L. K., and G. R. Weaver. "Business Ethics/Business Ethics: One Field or Two?" *Business Ethics Quarterly* 4 (1994): 13–128.

Thomas Donaldson (2005)

BUTLER, JOSEPH
(1692–1752)

Though he has not left us a complete philosophical system, Joseph Butler produced a moral philosophy that is still held in the highest esteem, and a philosophical theology of considerable long-term value. Butler was the eighth child of a prosperous draper. His father enrolled him in a dissenting academy, but he decided to join the established church and entered Oriel College, Oxford, in 1714. While still at school he had engaged in a philosophical correspondence with Samuel Clarke and at Oxford was befriended by Edward Talbot, son of the Bishop of Salisbury. Clarke and Talbot's father were instrumental in Butler's being appointed, after graduation, as Preacher at the Chapel of the Rolls. A selection of his sermons there was published in 1726 under the title *Fifteen Sermons Preached at the Rolls Chapel*. In 1729 a second edition appeared, with an important new preface. Bishop Talbot's patronage continued with Butler's entering the living of Haughton, and later that of Stanhope, in Talbot's later diocese of Durham. While at Stanhope Butler wrote his other major work, of which the full title is *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*. This appeared in 1736, and appeared in a second edition in the same year. By then Butler had entered royal circles. His school friend Robert Secker had drawn him to the attention of Queen Caroline, who appointed him Clerk of the Closet in 1736, conversed with him frequently on theological and philosophical matters, and received the sacrament from him on her deathbed in 1737. The king promised her that he would advance Butler and made him Bishop of Bristol in 1738. There is an unsubstantiated story that he was offered the see of Canterbury in 1747 and declined it. In 1751 he

became Bishop of Durham but was not destined to preside there for long because his health rapidly declined. He died in 1752, and was buried in Bristol. He never married.

BUTLER'S AIMS AND METHODS

Butler's personal history shows that he was, in C. D. Broad's words, "a thoroughly unworldly man whom the world treated very well." His integrity and intellectual prowess were widely recognized, and the patronage he received merely ensured that he did not suffer for them. His writings are often hard reading (and the sermons must often have been hard listening), not because they are unclear but because Butler aims at clarity exclusively and often sacrifices elegance in pursuit of it.

Butler is a Christian priest who seeks his readers' spiritual welfare. So, although his theoretical skills are considerable, they are wholly subordinated to his practical concern for the exercise of virtue and the proper consideration of the claims of religion. In urging these, however, he does not appeal to revelation. Nor does he use the a priori arguments in ethics and theology employed by Clarke, although he says he agrees with these. Butler's own methods are empirical ones. His ethical arguments are designed to show that the exercise of virtue is the expression of our true human nature and that vice violates it. His religious apologetic is based on the same appeal to probability that he thinks necessary for prudent conduct in everyday life. His famous attacks on selfish and hedonistic theories of human nature are designed to remove what he sees as the morally dangerous influence of faulty philosophy and are not intellectual explorations undertaken for their own sake.

ETHICS

In the Rolls Sermons, Butler seeks to encourage his worldly-wise hearers to practice virtue by arguing that to do so is to live in accordance with our nature. Virtue is the natural form of life for us, and vice is unnatural. He assumes, as his hearers would also have done, at least nominally, that the motives and capacities in our nature are placed there by God for our good, and he maintains that a realistic attention to those motives and capacities will show that living virtuously represents their natural exercise.

His argument has two main stages. The first stage is an account of the components of human nature, and the second is a claim about its structure and about the implications of that structure for our conduct. He argues that our nature is misrepresented by those (particularly Hobbes) who think that we are always selfish and by those who hold that we are always motivated by the desire for pleasure. If either of these theories were true, genuinely virtuous action would be impossible. Butler holds instead that our nature contains within it several distinct principles. There are, first, the "particular passions, appetites, and affections" such as the desires for food or possessions, or the emotions like joy or anger. There is, next, the "general affection of self-love," which is the desire for one's own long-term interest or happiness (which Butler interprets as the proper satisfaction or expression of one's own particular passions). It is self-love that causes us to restrain our present appetites in the interest of our long-term health, for example; and Butler clearly thinks of it as requiring rational calculation. Thirdly, there is the "natural principle" of benevolence. Butler uses this term as a general name to include all those desires we have for the good of others. (Scholars disagree over whether he also thinks of it as a rational principle in the same way that self-love is.) He identifies it with the love of one's neighbor.

Finally, and most important, our nature includes conscience. He describes this as "a principle of reflection in men, by which they distinguish between, approve and disapprove their own actions." Its judgments pronounce actions and motives to be "in themselves just, right, good" or "evil, wrong, unjust," and when it makes such judgments it "magisterially exerts itself." So conscience judges actions and motives in an intuitive manner and judges them as being of certain kinds, not as having good or bad consequences.

In defending his account of the components of our nature, Butler appeals primarily to our common experience. He also produces classic arguments against Hobbesian and other theories that say our motives are always selfish or are always directed toward pleasure. Experience seems to show us many examples of actions done from benevolence, and only a priori commitment to theory can incline us to doubt that our motives are often as they seem. Furthermore, self-love is only the motive for some actions and not for all. And although we do indeed gain pleasure from the successful pursuit of objects we desire, it is these objects themselves, and not the pleasure we derive from them, that we are pursuing.

But our nature is not merely one in which all these principles are to be found. It is one in which they form a system or constitution in which there is a built-in order of superiority and subordination. When we act in accordance with this order, we act naturally and so virtuously; when we violate it, we act unnaturally and so viciously. Butler introduces this claim with reference to the natural superiority of self-love to particular desires. If an animal

enters a baited trap in pursuit of food, it acts naturally because it follows the desire that is the strongest. But if a human knowingly satisfies a desire at the expense of his or her long-term good, then he or she acts unnaturally by ignoring the proper superiority of self-love to the ruinous desire. There is, therefore, a crucial distinction to be made in human nature between the strength of some motivating principle and its authority. In prudent behavior they coincide; in imprudent behaviour they clash.

Butler's key ethical doctrine is that of the natural supremacy not of self-love but of conscience. To live virtuously is to do what conscience approves and avoid what it disapproves. This does not mean that Butler identifies virtue with acting from duty (or conscientiousness); for conscience may well add its approval to actions that are already motivated by desire or by self-love. But when we are inclined to do something conscience rejects, or fail to desire what it enjoins, it may well have to supply its own motivating influence.

Butler thinks we usually have no difficulty in identifying right actions. He also thinks that these very largely coincide with the promptings of benevolence. But in the "Dissertation on Virtue" appended to the *Analogy*, he firmly rejects the utilitarian suggestion that virtue and benevolence can be identified. We lack the detailed knowledge of consequences for this to be true, so virtue consists rather in doing those acts that conscience approves—that is, acts of the right kind. That such acts will lead to the general good must be left to providence. He also thinks that providence must ensure that following conscience will not prove to be at odds with the demands of self-love and that benevolence (or love of neighbor) and self-love will also prove, in the end, to coincide.

Butler's case for the supremacy of conscience is therefore based on four related claims: that conscience has a natural authority, that is manifest in the way it makes its judgments; that to disregard it is to behave unnaturally; and that doing what conscience tells us is in the end for our good, even though we may not immediately discern this. These arguments are designed to persuade those who feel they know well enough what conscience tells them to do, but are still inclined to ask whether this is a compelling reason to do it. He tells them that if they recognized the place conscience has in their natures, they would see that it is.

The Rolls Sermons are notable for Butler's shrewdness, theoretical acumen, and wise moral psychology. They contain interesting and durable treatments of themes such as compassion, resentment, forgiveness, and self-deception.

PHILOSOPHICAL THEOLOGY

Butler's ethical sermons are still widely read, and their arguments have not dated. His religious apologetic has fared less well, even though it was better known in the century after his death. The reason for its present lack of influence is the fact that the debates to which it was intended to be a contribution have long since ceased. Butler's intent in the *Analogy of Religion* was to respond to the attacks on Christian orthodoxy made by the Deists. The Deists believed that the rational order of the cosmos revealed by science shows that our world had a creator, but they rejected Christian claims to revelation, maintaining that we only have need of "natural religion"—that is, the moral guidance of conscience and a vague general reverence for God. A deity who is rational in the way the design in nature shows God to be would have no need of special revelation, miracle, or priest craft to instruct us. Butler sees it as his task to restore the traditional connection between belief in God and openness to revelation in the face of this criticism.

Butler wants to encourage his readers, whom he assumes accept the reality of God, to pay close attention to the claims of Christianity and not dismiss them. He thinks that these claims have strong evidence in their favor; but his aim in the *Analogy* is less to show this than to persuade those who doubt it that they would still be prudent to examine them with care. He repeatedly stresses the importance of the claims that Christianity makes and the rashness of disregarding them. Probability, he tells us famously, is the guide of life. This assertion, though it is not accompanied by any philosophical analysis of the concept of probability, has two implications in Butler's thought. First, just as we have, in daily life, to base decisions on likelihoods rather than certainties, so in religious matters we must recognize our intellectual limitations and base our faith on what experience and reflection teach us is likely to be true rather than demand an unattainable certainty. Second, just as in life we often have to base decisions on the fact that there is a small chance of events that it would be foolish not to be prepared for, so in religious matters we should take the claims of revealed religion seriously as along as they have some degree of probability, even if it is a very modest one.

Butler opens the first part of the *Analogy*, on natural religion, with a case for a future life, a case that makes no appeal to providence. The key argument that he uses rests on a distinction between a person's possession of powers and the possession of means for their exercise. Although physical death clearly removes all sign of the capacity to exercise our powers, we cannot assume that it destroys

those powers themselves; just as there are many examples in nature of radical transformation in the history of living creatures, so we can reasonably expect the continuance of human powers hereafter. (In a well-known appendix to the *Analogy*, "Of Personal Identity," he further argues that our consciousness reveals to us that we are identical beings in the "strict and philosophical" sense—that is, fundamentally unchanging spiritual substances.)

In the remainder of Part I, Butler draws an analogy between the early and mature stages of human life on the one hand and the present life and the future life taken together on the other. He argues that we can discern clear signs that God teaches us the value of prudent and moral behavior in the early years of life in order to equip us to make good choices in our adult years and that we can reasonably infer that the exercise of virtue in the present life should be viewed as a training that fits us to enter the next. We are, he says, in a state of moral probation—a concept that partially anticipates John Hick's "soul-making" theodicy of the mid-twentieth century. God's government of the world is a "scheme imperfectly comprehended"; our ignorance of it, which Butler repeatedly emphasizes, is nevertheless only partial.

Part II defends revealed religion against deist criticisms. There should be no general presumption against miracles, because occasional divine violations of natural law might still be manifestations of "general laws of wisdom" and thus teach us, even though we could not predict them; and even though biblical prophecies may not have involved foresight on the part of their writers, if one thinks of God as the ultimate author of the book in which they are recorded, they can still reveal a divine purpose. Butler's basic defense, however, is that the recognition of our limitations should deter us from supposing that we know enough of God's purposes to dismiss the claims of revelation without careful study and that the overwhelming importance of Christian claims, if they are true, makes it frivolous and imprudent not to consider them with care, even if their probability may not at first seem high. He insists that with our limitations we should not expect more certainty in religious matters than we do in comparable secular ones, where our knowledge is also often merely partial—a form of argument that anticipates later demands by Christian apologists for philosophers to accord intellectual parity to the claims of religion. He also tells us that the claims of Christianity should be considered as a whole rather than piecemeal and that the case for its acceptance must be a cumulative one.

Butler's theology suffers in retrospect because Hume has made us question whether we can properly draw the analogy between this life and another on which the arguments of Part I of the *Analogy* depend, because only one of the terms of this analogy has been an object of experience. It has also seemed dated because the assumption of divine government that Butler and the Deists shared are no longer current. But many features of his reasoning can be detached from these two handicaps. His emphasis on our intellectual limitations, his doctrine of probation, and his insistence that the case for Christianity is a cumulative and probable on, all have present-day counterparts, and his detailed defenses of revealed religion are easily detachable from their contexts.

See also Clarke, Samuel; Conscience; Deism; Egoism and Altruism; Ethical Egoism; Ethics, History of; Evil; Hobbes, Thomas; Hume, David; Moral Motivation; Revelation; Self-Interest.

Bibliography

There is no complete edition of Butler's works in print at present. The best available is *The Works of Joseph Butler*, edited by J. H. Bernard, 2 vols. London: Macmillan, 1900. Next best is *The Works of Bishop Butler*, edited by W. E. Gladstone, 2 vols., Oxford, 1897. (Gladstone breaks up Butler's paragraphs, which Bernard numbers but leaves intact; and Bernard's notes are often helpful.) Butler, Joseph, *Fifteen Sermons*and *Dissertation upon the Nature of Virtue*, edited by T. A. Roberts, London: S.P.C.K., 1970, is a good recent edition of the Rolls Sermons. The most accessible reprint of the sermons it contains, and of the Dissertation, is *Five Sermons Preached at the Rolls Chapel and a Dissertation upon the Nature of Virtue*, edited by Stephen L. Darwall, Indianapolis: Hackett, 1983. There is no extant edition of the *Analogy of Religion*, although libraries and bookstores often have nineteenth-century reprints of it because it was required reading for Anglican ordinands for many years. As of this writing, a new edition of Butler's sermons is expected from Oxford University Press (U.K.) in 2006, edited by David McNaughton.

There are two general books about Butler and his times that are valuable: W. A. Spooner, *Bishop Butler*, London: Methuen, 1901; and Ernest Campbell Mossner, *Bishop Butler and the Age of Reason*, New York: B. Blom, 1971. A recent collection that contains fine essays on all aspects of Butler's life and work is Christopher Cunliffe, ed., *Joseph Butler's Moral and Religious Thought: Tercentenary Essays*, Oxford: Clarendon Press, 1992. Terence Penelhum, *Butler*, London: Routledge, 1985, covers his ethical teachings and attempts to rescue his philosophical theology from undeserved neglect. Austin Duncan-Jones, *Butler's Moral Philosophy*, Harmondsworth: Penguin, 1952, is a high-quality treatment of the ethics.

There are many good essays and chapters about Butler. On the ethics, the best beginning is Chapter 3 of C. D. Broad's *Five Types of Ethical Theory*, London: Routledge, 1930. A severe critique of Butler that argues his doctine of the naturalness of virtue leads to incoherence is Nicholas L. Sturgeon, "Nature and Conscience in Butler's Ethics," *Philosophical Review* 85 (1976), 316–356. On this and the issue of the

supremacy of conscience, see Stephen Darwall, "Conscience as Self-Authorizing in Butler's Ethics" in Cunliffe (1992), 209–242. The notion of the naturalness of virtue is explored with originality in Alan Millar, "Butler in God and Human Nature" in Cunliffe, 293–315. Another excellent essay is Jerome Schneewind, "The Divine Corporation and the History of Ethics," in *Philosophy in History*, edited by R. Rorty, J. B. Schneewind, and Q. Skinner, 173–192, Cambridge, U.K.: Cambridge University Press, 1984.

For a long time the only good treatment of Butler's theology was C. D. Broad, "Bishop Butler as Theologian," in C. D. Broad, *Religion, Philosophy, and Psychical Research*, London: Routledge, 1953, 202–219. The situation improved with Anders Jeffner's *Butler and Hume on Religion*, Stockholm: Diakonistyrelsens Bokforlag, 1966. Penelhum, *Butler* (above), carries the debate further. See also especially the essays by David Brown ("Butler and Deism," 7–28), Basil Mitchell ("Butler as a Christian Apologist," 977–1116), and T. A. Roberts ("Butler and Immortality," 169–188) in Cunliffe. The importance of Butler's thinking as a stimulus to Hume is explored in Paul Russell, "Butler's 'Future State' and Hume's 'Guide of Life,'" *Journal of the History of Philosophy* 42 (2004): 425–448. It is clear that Section 11 of Hume's *Enquiry concerning Human Understanding*, "Of a Particular Providence and Of a Future State," is for the most part an attempt to undercut the use of analogical reasoning found in Butler's *Analogy*. John Hick's "soul-making" theodicy is to be found in his *Evil and the God of Love*, London: Macmillan, 1966.

The best place to begin study of Butler's religious thought, however, is Sermon 15 of the Rolls Sermons, "Of the Ignorance of Man," which prefigures much of the prudential apologetic that the *Analogy* develops in detail.

Terence Penelhum (2005)

BUTLER, SAMUEL

(1835–1902)

The English writer and critic Samuel Butler was the author of the satirical novels *The Way of All Flesh, Erewhon,* and *Erewhon Revisited,* as well as several discussions of philosophical biology and the theory of evolution. He was the son of the Reverend Thomas Butler, whom he depicted as a domestic tyrant in *The Way of All Flesh.* Butler was sent to Cambridge by his father in the hope that he would become a clergyman, but after graduating he refused to take orders because of doubts about the Christian creed. In 1859 he emigrated to New Zealand, where he became a successful sheep farmer and for a time a convert to Darwinism. Returning to England in 1864 with enough money to live on, he began a career as an author, painter, and musician. The subject of evolution continued to occupy his mind for many years. It forms the substance of several essays and four books: *Life and Habit* (London, 1878), *Evolution, Old and New* (London, 1879), *Unconscious Memory* (London, 1880), and *Luck or Cunning?* (London, 1887). These works reflect a mounting hostility to the ideas of Charles Darwin and a desire to champion those of Erasmus Darwin and the Chevalier de Lamarck. This hostility first made its appearance in *Erewhon* (London, 1872).

EVOLUTION

Butler was neither a scientist nor a philosopher. His discussions of evolution are the work of a literary man with strong intellectual interests but little capacity for exact thought. He was at his best when giving scientific and philosophical ideas an original twist that often put them in quite a new light. To many fellow Victorians he seemed an irreverent skeptic or even an atheist; but in fact, he wanted to retain religion while discarding the Christian creed and to discard Darwin while retaining evolution. This outlook pervades all his major writings.

The central weakness of Darwinism, according to Butler, was its failure to identify the cause of the variations on which selection was said to operate. They were described as random or accidental, which would mean that the course of evolution has been a matter of luck. The older evolutionists, such as Erasmus Darwin and Lamarck, were far sounder in their views, for they attributed the cause of variations to the activity of organisms and to the inherited effects of the use or disuse of their various functions. Not luck, they claimed, but cunning displayed by organisms in coping with their environment lies at the basis of evolution. Hence, the activity of organisms is profoundly purposive. The great mistake of Charles Darwin was to dismiss teleology from the domain of living things, for they then become indistinguishable from machines.

In an essay of 1865 Butler toyed with the idea that machines are adjuncts to organisms, like extra, though inferior, limbs, by means of which organisms have become more highly evolved. Hence, "a leg is only a much better wooden leg than anyone can manufacture." This led Butler to consider the problem of how living things have come to produce their natural organs and to equip themselves with adaptive habits. The answer, he asserted, is that the individual plant or animal must "know" at the start what to do. A fertilized ovum possesses the knowledge it needs to make itself into an embryo and subsequently into an adult organism. This knowledge is really a remembering of what its ancestors did in the past. Hence, we must postulate an "unconscious memory" at work in all living things, binding successive generations and providing the basis for the transmission of acquired characteristics.

Butler then leaped to two sweeping conclusions. First, consciousness and intelligence exist throughout the whole organic world. "For the embryo of the chicken, we claim exactly the same kind of reasoning power and contrivance which we claim for the amoeba, or for our own intelligent performances in later life." Second, since evolution involves a continuous process of derivation, there must be an "identity" between parents and offspring: the latter are not different individuals but *are* the parents at a later evolutionary stage. "Birth has been made too much of." A newborn infant is simply part of an unbroken biological process, not an utterly separate individual. Accordingly, there is a deep unity of all life, so that it constitutes "in reality, nothing but one single creature, of which the component members are but, as it were, blood corpuscles or individual cells."

With the aid of these conclusions, Butler sought to justify an idealistic and religious interpretation of evolution. In *Unconscious Memory* he contended that his earlier separation of the organic from the inorganic was unwarranted. "What we call the inorganic world must be regarded as up to a certain point living, and instinct with consciousness." Hence, "all space is at all times full of a stuff endowed with a mind," and "both stuff and mind are immaterial and imperceptible, so long as they are undisturbed, but the moment they are disturbed, the stuff becomes material and the mind perceptible." Evolution is therefore the life history of this primordial world stuff, "to which no name can be so fittingly applied as 'God.'"

Many of Butler's criticisms of Darwinism have been made irrelevant by the rise of the science of genetics. Yet he was justified in urging those criticisms at the time and in calling attention to vacillations in Darwin's thought on basic issues. If Butler had been more scrupulous in his own thinking and less facile with his pen, his works on philosophical biology might have had greater survival value.

THEOLOGY

Butler's rather unusual theology is set forth in three essays, posthumously published as *God the Known and God the Unknown* (London, 1909). He there contended that an adequate concept of God requires him to be a living person with a material body. To regard God as *merely* a spirit is tantamount to atheism. At first Butler held that the divine body is just the totality of life, the "one single creature" whose unconscious memory is part of the divine mind. When he rejected the distinction between the organic and the inorganic, his view shifted from a "panzoistic" conception of God to pantheism. He intended to rewrite his theology in the light of this shift, but never managed to do so. One odd belief he expressed was that the grand design of the cosmos points to the existence of "some vaster Person who looms out behind our God, and who stands in the same relation to him as he to us. And behind this vaster and more unknown God there may be yet another, and another, and another." This pyramiding of deities was one of the many items with which Butler enlivened the Victorian scene.

SOCIAL THOUGHT

Despite the barbs he directed at the institutions of his day, Butler's social outlook was conservative. He took the position that those who are rich and successful are the highest types thus far produced in the evolutionary process. Poor men are biological misfits; hence, the sooner they disappear and leave room for those better able to take care of themselves, the better. In the imaginary society of Erewhon, "if a man has made a fortune of over £20,000, they exempt him from all taxation, considering him a work of art and too precious to be meddled with." Butler's account of this society is not so much a blueprint of utopia as a device for satirizing the beliefs and practices of middle-class Englishmen by inverting accepted values. Thus, in Erewhon bodily illness was considered a punishable crime, whereas moral failings deserved sympathy and were given therapeutic treatment. Instead of fostering machinery, the Erewhonians, after a long struggle, destroyed it when they realized that machines, like organisms, were evolving and would soon acquire a mastery over men. In *Erewhon Revisited* (London, 1901), Butler depicted a community showing signs of degeneration, as if to underline the conclusion that a social order is an impermanent evolutionary product and inevitably alters. Yet here again no consistent point of view was worked out.

See also Consciousness; Darwin, Charles Robert; Darwin, Erasmus; Darwinism; Evolutionary Theory; Lamarck, Chevalier de; Pantheism; Philosophy of Biology; Teleology.

Bibliography

Furbank, P. N. *Samuel Butler.* Cambridge, U.K., and New York: Cambridge University Press, 1948.

Jones, H. Festing. *Samuel Butler, Author of Erewhon (1835–1902): A Memoir,* 2 vols. London: Macmillan, 1919. The standard biography.

Kingsmill, Hugh. *After Puritanism.* London: Duckworth, 1929.

Muggeridge, Malcolm. *The Earnest Atheist: A Study of Samuel Butler.* London: Eyre and Spottiswoode, 1936.

Willey, Basil. *Darwin and Butler, Two Versions of Evolution.* New York: Harcourt, Brace, 1960.

T. A. Goudge (1967)

BYZANTINE PHILOSOPHY

The age of the Byzantine Empire stretches from the end of late antiquity to the fall of Constantinople in 1453. During Byzantine times scholars who copied and studied or even lectured on the texts of ancient philosophers are known and praised chiefly for their efforts to transmit and to keep alive the philosophical traditions of antiquity. To take the obvious case of Plato's and Aristotle's works, there are more than 260 Byzantine manuscripts of dialogues by Plato and at least 1,000 Aristotelian texts. This does not mean, however, that all Byzantine scholars should be regarded as mere copyists. There were among them important figures who, being philosophers themselves, not only carefully studied and commented on ancient philosophical works but also wrote their own treatises on central philosophical problems.

How did the Byzantines conceive of philosophy and of themselves as philosophers? John of Damascus (*Dialectica* 1:56), for instance, gives six complementary definitions of philosophy:

(1) the knowledge of beings as beings;

(2) the knowledge of things divine and human;

(3) a preparation for death;

(4) the assimilation of man to God as far as humanly possible;

(5) the art of arts and the science of sciences;

(6) the love of wisdom.

These six definitions, which were often cited by other Byzantine philosophers too, can also be found in the works of the Neoplatonists of the Alexandrian school (for example, David, *Prolegomena* 20.27–31). They are clearly derived from Aristotelian (1, 5), Stoic (2), and Platonic (3, 4) conceptions of philosophy, attesting thus to the Byzantines' solid knowledge and eclectic use of the different traditions in ancient philosophy.

However, the Byzantines were by no means unanimous about the importance of ancient philosophy, or of "the wisdom from without," as they called pagan philosophy in contrast to Christian theology, which they called "the wisdom from within." Some, under the influence of St. Paul and authors like Tatian, considered ancient philosophy useless and dangerous because it corrupts the Christian view of things and leads to heresies. Others, under the influence of Basil the Great and Gregory of Nyssa, claimed that ancient philosophy, if used in a cautious way, could be a preparation for the true faith, help in its elucidation, and serve as a dialectical weapon against heresies. Moreover, Byzantine philosophers like John Italos and Barlaam of Calabria undertook the task, in some cases at high personal risk, of defending ancient philosophy in its own right, but also as a means for a better understanding of Christian dogma.

The term *philosophy* could also be used in Byzantium in a much wider sense to include encyclopedic knowledge, including mathematical sciences such as astronomy. Sometimes, following some of the Church Fathers, the term could be used to refer to a life of contemplation as exemplified by Christian monasticism. But that philosophy was partly understood as the Christian way of contemplative life does not necessarily mean that philosophy collapsed into theology. On the contrary, the borders between philosophy and theology were reasonably clearly defined in Byzantium.

The view expressed by some Church Fathers, for instance by Clement and by Origen, that philosophy is the handmaiden of theology (*philosophia theologiae ancilla*), was not the dominant position in the Byzantine East. Byzantine philosophy seems to have managed to preserve its autonomy. Even though many of the problems with which Byzantine philosophers were concerned, like that of divine providence, did indeed arise in the context of a Christian theological tradition, these problems nonetheless constitute genuine philosophical issues that would be of interest to any philosopher, even one who did not believe in Christian dogma. For example, the following are some of the issues that profoundly and systematically occupied many Byzantines philosophers: the creation or origin of the world, the existence of God, the ontological status of universals, the character of the perceptible world, the problem of evil and human free will, the relation between soul and body, the necessary requirements for a good life, the possibility of a just state, the connection between faith and reason, the skeptical challenge to knowledge.

But did the Byzantine philosophers express original views in discussing these issues? There is no doubt, of course, that Byzantine philosophical writings are quarries of information about earlier philosophical doctrines, which would have been otherwise completely lost or only meagerly documented. Besides, whatever attitude the Byzantines took towards ancient philosophy, it was impossible for them to escape altogether from its influence. It was ancient philosophy that clearly provided them with a well-articulated theoretical framework and with the philosophical language that served as the basis for their own philosophical discourse. At the same time,

however, the Byzantine philosophers offered in their commentaries and treatises numerous clarifications, developments, criticisms, and modifications of ancient doctrines, some of which are philosophically interesting and remarkably subtle.

Even when they simply paraphrased or briefly commented on ancient philosophical texts, the Byzantines presented different degrees of independent thinking; sometimes they gave a slightly different argument to support an established position, sometimes they made a small but interesting addition to an ancient doctrine, and sometimes they considerably diverged from the view generally accepted in antiquity. But this should not be understood as suggesting that the Byzantine philosophers were interested in being original; like most of their late ancient predecessors they would have firmly rejected such a suggestion.

Nevertheless, Byzantine philosophy as a whole exhibits a distinctive character that differentiates it from the previous period in the history of philosophy. For it is clear that many of the views and doctrines presented by the Byzantines originated in their aim to reconcile their Christian tradition with ancient philosophy. For instance, they taught Aristotle's logic as generally useful, but mainly as a preparation for more theoretical studies; they disagreed, however, with his doctrine of the eternity of the world and his understanding of God as the first unmoved mover who moves the heavens but exerts no providence on the details of the sublunary world, including individual human beings. Instead, Byzantine philosophers considered Plato's metaphysics to be closer to the Christian worldview, especially on issues like the immortality of the soul and the creation of the world; still, for doctrinal reasons they could not accept the Platonic theory of metempsychosis and the separate existence of eternal ideas or forms.

Hence, Byzantine philosophers seem to have followed the eclectic tradition of late antiquity and combined aspects of Plato's and Aristotle's theories, although always strongly influenced by Neoplatonic philosophers like Proclus. The Byzantines also engaged in a limited dialogue with the other schools of ancient philosophy; for instance, they were interested in criticizing elements of Epicurean or Stoic doctrine, and they critically examined the implications of the Skeptics' views on the possibility of human knowledge. This is the picture at least up to the fifteenth century, when the leading intellectuals of the time, George Gemistos Pletho and George Scholarios Gennadios, started emphasizing the contrast between ancient philosophers and believed that they should take sides, presenting themselves either as Platonists or as Aristotelians.

BYZANTINE PHILOSOPHERS

In Byzantium there were no institutions of higher education in which philosophers could be trained as philosophers. The main purpose of institutional higher studies was to train civil servants. The figure of the Byzantine philosopher, therefore, emerges as somewhat of a polymath and an erudite scholar, who, moreover, might make use of his knowledge and rhetorical skill to play an active role in the political life of his times. Philosophical instruction was mainly private, but it sometimes received support from the Emperor and the Church, as in the case of the so-called University of Constantinople, which was founded in 1045 by Constantine Monomachos. Such support, however, also meant occasional intervention by the secular or ecclesiastical authorities, as when John Italos was put on trial and condemned for advocating the systematic use of philosophical analysis in clarifying theological issues.

In general, the philosophical curriculum would start with Aristotle's logic, considered as the instrument of all sciences (Porphyry's *Isagoge*, Aristotle's *Categories*, *De interpretatione*, and *Prior Analytics* 1.1–7); then ethics, teaching a rationally ordered moral life of the soul as joined to the body; and finally, through physics and the *quadrivium* (arithmetic, geometry, astronomy, and harmonics), to Platonic or, more precisely, to Neoplatonic metaphysics, which is the highest philosophical science because it has to do with knowledge of first principles and brings the soul nearer to assimilation to the divine.

The genres of philosophical writing in Byzantium are quite diverse. For teaching purposes the Byzantine scholars produced marginal notes and explanatory paraphrases on ancient philosophical works, but also extended commentaries, sometimes in question-and-answer form, small handbooks, or large surveys of philosophy. They also wrote small treatises on specific topics or longer works, occasionally in dialogue form, with the aim of rebutting the views of their opponents and to explain and defend their own theories. To all these we should further add their letters and orations, which frequently made reference to philosophy.

The real starting point of Byzantine philosophy is usually placed in the ninth and tenth century, when the so-called Byzantine humanists, men like the Patriarch Photios, Arethas, or Leo the Mathematician, started again studiously to read, edit, and comment on the works of ancient philosophers. Having said that, however, the dis-

tinctive character of Byzantine philosophy undoubtedly owes a lot to the influence of the previous period, which was dominated by the thought of the Church Fathers such as Basil the Great, Gregory of Nyssa, Pseudo-Dionysius, Maximus the Confessor, and John of Damascus.

Photios (820–891), who is famous mainly for his *Bibliotheke*, a vast compilation of ancient Greek literature, also taught Aristotelian logic and wrote, for this purpose, comments on Aristotle's *Categories*. In addition, he composed a number of small treatises in which he criticized both Plato's and Aristotle's views, especially their theories on universals; he himself claimed that universals have no independent existence but are conceived by God and are instruments of God's will. Arethas (c. 850–944) also commented on Aristotle's *Categories* and Porphyry's *Isagoge*, but he is better known for having been instrumental in the transmission of ancient texts, in particular the Platonic corpus. He commissioned the transcription of a complete copy of Plato's works, to which he added marginal notes; the first part of his Plato text is extant as the famous Clarkianus 39 manuscript in the Bodleian Library of Oxford. Unfortunately, we know little about Leo the Mathematician (c.790–869), who seems to have taught philosophy at the so-called Magnaura School in Constantinople.

There is a significant development from the humanistic Photios and Arethas interests to the way the Byzantines in the eleventh and twelfth century, the period of the Comneni, viewed the philosopher as someone with a hard-earned and unsurpassed knowledge in all branches of learning, and especially as someone who formed his own views on the philosophical topics discussed by the ancients. Michael Psellos (1018–1078) was one of the most erudite and intriguing figures of the Byzantine Middle Ages. He was given the honorific title "first among the philosophers" and taught all branches of philosophy. He commented on Aristotle's logic (*Categories*, *De interpretatione*, *Prior Analytics*) and his physics, and he wrote a large number of short treatises discussing particular problems raised by his pupils; he also compiled a short encyclopaedia with the title *De omnifaria doctrina*. He was greatly influenced by Proclus, whom he considered as an authority among ancient authors. In his attempts to advance philosophical learning he was often attacked concerning his theological orthodoxy, so that he often had to be careful to distance himself from heretical doctrines, as in his writings on the *Chaldaean Oracles*.

John Italos (c.1025–1082), a pupil of Psellos, who was condemned by the Church of Constantinople for his extensive use of logical reasoning in theological matters, wrote treatises discussing the Aristotelian categories and commented on Aristotle's logic (*De interpretatione*, *Topics*). Eustratios of Nicaea (c.1050–1120) and Michael of Ephesus (c.1050–1129) belonged to the intellectual circle around Anna Comnena and took part in her project to produce commentaries on Aristotle's works.

Eustratius wrote commentaries on Aristotle's *Posterior Analytics* and *Nicomachean Ethics*, whereas Michael of Ephesus commented on Aristotle's metaphysics, logic (*Sophistici elenchi*), ethics, and natural philosophy (*Parva naturalia*, *De partibus animalium*, *De generatione animalium*, *De motu animalium*, and *De incessu animalium*). Their work, in which they followed ancient commentaries (some of which are now lost) but also added their own insightful remarks, was instrumental in the transmission and revolutionary rediscovery of Aristotelian thought in the Latin West. Finally, Nicholas of Methone (d. 1165) wrote at the same time a detailed refutation of Proclus's *Elements of Theology*. During the short period after the Latin conquest of Constantinople in 1204, when the center of Byzantine intellectual life moved to Nicaea in Asia Minor, the main intellectual figure was Nikephoros Blemmydes (1197–1272), who wrote a much-used handbook of physics and logic that also was translated in Latin.

Lastly, the final centuries of the Byzantine empire, which are known as the Palaeologan period, saw a renewal of interest in the sciences, particularly in mathematics and astronomy. George Pachymeres (1242–1310) composed a summary of Aristotelian philosophy and wrote Neoplatonic commentaries, supplementing Proclus's commentary on Plato's *Parmenides*. Theodore Metochites (1270–1332) criticized Aristotelian physics and metaphysics in debate with Nikephoros Choumnos (c. 1250–1327), who in turn attacked the orthodoxy of Neoplatonic psychology. Sophonias and Leo Magentinos paraphrased works of Aristotle; Sophonias paraphrased Aristotle's *Categories*, *Sophistici elenchi*, and *De anima*, while Leo Magentinos paraphrased Aristotle's *De interpretatione*, *Prior Analytics*, and *Sophistici elenchi*.

Moreover, three important intellectuals of the fourteenth century, namely Nikephoros Gregoras (1290/3–1358/61), Barlaam of Calabria (c. 1290–1348), and Gregory Palamas (c.1296–1359), got involved in a fierce dispute over the use of logical reasoning in theology. Gregoras claimed that logical studies should be regarded as completely useless and should be therefore altogether dismissed, whereas Barlaam and Palamas adopted a more complex attitude toward logic. They both stressed that logic is indeed useful in defending Christian dogma

against pagans and heretics, but they disagreed about the limits of the use of logical reasoning in clarifying or establishing the truth of Christian belief; whereas Barlaam argued that logical methods can be used to prove the Christian beliefs, Palamas insisted that logical arguments are of no help in our attempt to acquire knowledge of God and of his attributes. The controversy between Gregoras, Barlaam, and Palamas extended to a second stage, known as the Hesychast debate, which centered on the method of prayer and contemplation of the Byzantine monks, who claimed to be able to achieve communion with God through inner quietude and silence.

In the fifteenth century, around the time of the fall of Constantinople, a main focus of Byzantine philosophers was, as mentioned above, stressing the differences between Plato and Aristotle and determining the superiority of the one over the other. George Gemistos Plethon (c. 1360–c. 1453) is famous for his renewal of Proclus's Neoplatonism as a theological and political alternative to Christianity. In his treatise *De Platonis et Aristotelis philosophiae differentia* he argued for the superiority of Plato over Aristotle; in his *Laws* he presented an utopia based primarily upon Plato and the Neoplatonists. George Scholarios Gennadios (c. 1400-–1424) thought that Pletho's utopia was heretical and should be consigned to the flames. He defended Aristotle's works and was more favourable to Latin scholasticism. He commented on Aristotle's logic (*Categories*, *De interpretatione*), natural philosophy (*Physics* 1–3, *Parva naturalia*), and Aristotle's *De anima*. He also translated part of Petrus Hispanus's *Summulae logicales* and works by Thomas Acquinas, for instance the *De fallaciis* and his commentary on the *Posterior Analytics*. Bessarion (1403–1472), who had studied under Pletho, tried to mediate the dispute between Pletho and Scholarios, and gave a sympathetic summary of Plato's philosophy, which he thought reconcilable with Aristotelianism. He, like Pletho, greatly helped to bring works of Plato and Aristotle to the attention of Italian humanists.

From the second half of the thirteenth century onward, there were translations into Greek of Western Latin texts, especially logical texts: Manuel Holobolos (fl. 1267) translated Boethius's *De topicis differentiis* and *De hypotheticis syllogismis*; Maximos Planudes (c. 1255–c. 1305) translated Boethius's *De consolatione philosophiae* and Augustine's *De trinitate*; Demetrios Kydones (c. 1324–97/8) and his brother Prochoros Kydones (c. 1333–69/70) translated Augustine, Anselm, and Thomas Aquinas. But it was only in the fifteenth century that Byzantine and Western philosophers actually began to talk to one another, to read one another's books, and to be influenced by others' traditions and views. Still, although the Byzantine scholars like John Argyropoulos—who went to Italy and worked there as teachers of Greek, editors of Greek texts and translators, and as teachers of philosophy—exerted a fertile influence on the West, Byzantium itself in general remained closed to Western scholasticism.

THE STUDY OF BYZANTINE PHILOSOPHY

Byzantine philosophy remains a little-explored field. Most of the writings of Byzantine philosophers are yet unpublished or are available only in old and often quite inadequate editions. The nineteenth-century Berlin series *Commentaria in Aristotelem Graeca*, which was supposed to include all commentaries on Aristotle's works, actually includes a very small selection of Byzantine commentaries. Translations of and commentaries on Byzantine philosophical works are hardly ever available. In addition, there are important unresolved issues about the authorship of many Byzantine philosophical texts, and we often have no reliable information concerning their sources. But even when we do have careful editions of the philosophical works of Byzantine thinkers, their philosophical contribution for the most part still needs to be critically assessed. Being regarded either as mere scholars or as religious thinkers, Byzantine philosophers have not been studied on their own merit, and their works have hardly been scrutinized as works of philosophy.

The interest of the scholars of the nineteenth and early twentieth century, who worked with great care on some Byzantine philosophical texts, was not primarily philosophical. Philosophers, on the other hand, understandably were discouraged both by the rhetorical style of the Byzantine writings and by the theological interests displayed in much of Byzantine philosophy. Therefore, although distinguished historians have in the past tried to reconstruct the intellectual life of the Byzantine period, we still lack even the beginnings of a systematic understanding of the philosophical works produced in Byzantium. It is particularly telling that there is no adequate recent monograph even on the most prominent Byzantine philosopher, Michael Psellos.

After World War II, however, we see significant changes in the study of Byzantine philosophy. These changes clearly are connected with the rediscovery and philosophical reappraisal of the Western medieval philosophical tradition and of certain areas in ancient philosophy, such as the works of the Neoplatonists and of the

ancient commentators. Critical editions of texts are appearing regularly, in particular in the series *Corpus Philosophorum Medii Aevi—Philosophi Byzantini* (The Academy of Athens) and in the Bibliotheca Teubneriana. Moreover, books and articles are now being published that investigate the teaching of philosophy in Byzantium and the original philosophical contributions of Byzantine philosophers in a philosophically more adequate and serious way. Nevertheless, much more work is required to achieve a reliable overview of Byzantine thought. Following the rising interest of the last decades of the twentieth century, it now seems important to encourage further the systematic study and critical assessment of the individual works of Byzantine thinkers. Most importantly, we need to take their works seriously as philosophical writings; putting aside our prejudices and misconceptions, we need to make a renewed effort to reconstruct and to do justice to Byzantine philosophy.

See also Analysis, Philosophical; Anselm, St.; Aristotelianism; Aristotle; Augustine, St.; Boethius, Anicius Manlius Severinus; Clement of Alexandria; Determinism and Freedom; Evil, The Problem of; Gregory of Nyssa; John of Damascus; Medieval Philosophy; Neoplatonism; Origen; Patristic Philosophy; Plato; Platonism and the Platonic Tradition; Pletho, Giorgius Gemistus; Proclus; Pseudo-Dionysius; Stoicism; Thomas Aquinas, St.

Bibliography

Arethas of Caesarea. *Scholia on Porphyry's Isagoge and Aristotle's Categories*. Edited by M. Share. Athens: The Academy of Athens, 1994.

Gregoras, Nikephoros. *Antirrhetika I*. Edited by H.-V. Beyer. Vienna: Verl. d. Osterr. Akad. D. Wiss., 1976.

Gregoras, Nikephoros. *Fiorenzo o Intorno alla Sapienza*. Edited by P. Leone. Naples: University of Naples, 1975.

Italos, Joannes, and Leon Magentinos. *Byzantinische Kommentatoren der aristotelischen Topik*. Edited by Sofia Kotsabassi. Thessaloniki, Greece: Vanias, 1999.

John of Damascus. *Schriften*. Edited by B. Kotter. Berlin: De Gruyter, 1969.

Nicholas of Methone. *Refutation of Proclus' Elements of Theology*. Edited by A.A. Angelou. Athens: The Academy of Athens, 1984.

Pachymeres, George. *Commentary on Plato's Parmenides*. Edited by T. Gadra et al. Athens: The Academy of Athens, 1989.

Pachymeres, George. *Philosophia. Buch 10. Kommentar zur Metaphysik des Aristoteles*. Edited by Eleni Pappa Athens: The Academy of Athens, 2002.

Photios. *Epistulae et Amphilochia*. Edited by L. G. Westerink. Leipzig: Teubner, 1986.

Pletho, George Gemistos. *Contra Scholarii pro Aristotele objections*. Edited by E. V. Maltese. Leipzig: Teubner, 1988.

Pletho, George Gemistos. *Oracles Chaldaïques. Recension de Georges Gémiste Pléthon*. Edited by Brigitte Tambrun-Krasker. Athens: The Academy of Athens, 1995.

Pletho, George Gemistos. *Traité des vertus*. Edited by Brigitte Tambrun-Krasker. Athens: The Academy of Athens, 1987.

Psellos, Michael. *De omnifaria doctrina*. Edited by L. G. Westerink. Utrecht: J. L. Beuers N.V., 1948.

Psellos, Michael. *Philosophica minora I*. Edited by J. M. Duffy. Leipzig: Teubner, 1992.

Psellos, Michael. *Philosophica minora II*. Edited by D. J. O'Meara. Leipzig: Teubner, 1989.

Scholarios Gennadios, George. *Oeuvres complètes*. Edited by L. Petit, X. A. Sidéridès, and M. Jugie. Paris: Maison de la bonne presse, 1936.

SUGGESTED READING

Benakis, Linos. "Commentaries and Commentators on the Logical Works of Aristotle in Byzantium." In *Gedankenzeichen. Festschrift Klaus Oehler*. Edited by R. Clausen and R. Daube-Schackat. Tübingen, Germany: Stauffenburg, 1988: 3–12.

Benakis, Linos. "Bibliographie internationale sur la philosophie byzantine (1949–1990)." In *Bibliographie byzantine publiée à l'occasion du XVIIIe Congrès international d'Etudes byzantines*. Athens: Assoc. intern. D. Etud. Byz., 1991.

Benakis, Linos. "Commentaries and Commentators on the Works of Aristotle (Except the Logical Ones) in Byzantium." In *Historia philosophiae Medii Aevi. Festschrift Kurt Flasch*. Edited by B. Mojsisch and O. Pluta. Amsterdam: B. R. Grüner, 1991. Vol. I: 45–54.

Bydén, Börje. *Theodore Metochites' Stoicheiosis Astronomike and the Study of Natural Philosophy and Mathematics in Early Palaiologan Byzantium*. Göteborg, Sweden: Acta Universitatis Gothoburgensis, 2003.

Hunger, Herbert. *Die hochsprachliche profane Literatur der Byzantiner*. Munich: C.H. Beck, 1978. Vol. 1:3–62.

Ierodiakonou, Katerina, ed. *Byzantine Philosophy and Its Ancient Sources*. Oxford: Oxford University Press, 2002.

Kapriev, Georgi. *Philosophie in Byzanz*. Wurzburg, Germany: Konigshausen & Neumann, 2005.

Lemerle, Paul. *Le premier humanisme byzantin*. Paris: Presses Univeritaires de France, 1971.

Oehler, Klaus. *Antike Philosophie und byzantinische Mittelalter*. Munich: C. H. Beck, 1969.

Podskalsky, Gerhard. *Theologie und Philosophie in Byzanz*. Munich: C. H. Beck, 1969.

Ševčenko, Ihor. *Études sur la polémique entre Théodore Métochite et Nicéphore Choumnos*. Brussels: Byzantion, 1962.

Tatakis, Basil. *La philosophie byzantine*. Paris: Presses Univeritaires de France, 1949. (Modern Greek translation with bibliography for 1949–1976 by L. Benakis, Athens: Etaireia Spoudon, 1977; English translation by N. J. Moutafakis, Indianapolis/Cambridge: Hackett, 2003.)

Wilson, Nigel. *Scholars of Byzantium*. London: Duckworth, 1983.

Katerina Ierodiakonou (2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

APPENDIX: ADDITIONAL ARTICLES
THEMATIC OUTLINE
BIBLIOGRAPHIES
INDEX

10
volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

appendix

ALBERT THE GREAT
(before 1200–1280)

According to the near-contemporary testimony of Tolomeo of Lucca (*Historia Ecclesiastica* [1317], 22.19) and confirmed by other, later sources, Albert the Great (Albertus Magnus) was more than 80 years old when he died on November 15, 1280, establishing the turn of the thirteenth century as the *terminus ante quem* of his birth. He was born in the town of Lauingen in Schwaben in the diocese of Augsburg, at the time a part of Bavaria, the son of a knight in the service of the counts of Bollestadt. He was already a student in the *studium litterarum* at Padua when, in 1223, Jordan of Saxony came in search of recruits to the Dominican Order among the young men in residence at the new university. Albert received the habit from Jordan sometime around Easter of 1223 and was sent to Cologne for his novitiate. By 1228 he had become a lecturer (*lector*), and he served in that office in Dominican communities at Heldesheim, Freiberg, Regensburg, and Strassburg. In 1243 or 1244 he was sent to Paris by John of Wildeshausen, where he became a master of theology in 1245 and lectured on Peter Lombard's *Sententiarum* (Sentences).

In the fall of 1245 Thomas Aquinas was sent to Paris, also at the direction of John of Wildeshausen, and in 1248 he and probably other Dominicans accompanied Albert to Cologne, where Albert was to establish the first *studium generale* (or liberal-arts college) in Germany. He served as Provincial of Teutonia from 1254 to 1257, during which time he was summoned before the papal curia to defend the Dominican Order against the attacks of William of Saint-Amour. He was well received by the curia, and his lectures and debating were found to be extraordinary. In January of 1260 Pope Alexander IV appointed him bishop of Regensburg, but he served less than two years before submitting his resignation, after instituting many reforms in his diocese. Although retired, he was directed by Pope Urban IV, in 1263, to preach to the Germans a crusade to the Holy Land, and this he did, until Urban's death in 1264.

It is said that after the death of Thomas Aquinas, Albert traveled to Paris one last time to defend the views of his former student, but this story, related at the canonization proceedings for Aquinas in 1319, is not fully consistent with other known facts about Albert's final years and, indeed, appears to interpret the events in Paris in 1277 in a manner that places far too much importance on the connection, if any, between Aquinas and the doctrines that were being formally condemned. The complete absence of any official correspondence after August 18, 1279, in the face of a full and active participation in the life of the Church and his order right up until that date, has suggested to some that Albert's memory, and perhaps other aspects of his mental life, had begun to fail him at that time, but there is no good reason to suppose, as some have done, that this decline began as early as 1277. Whether he was already in decline or not, he and his Dominican brothers were apparently not unprepared when death finally took him away on November 15, 1280.

WRITINGS

Albert was committed to the preservation and propagation of the philosophical ideas of antiquity, in particular the philosophy of Aristotle, which he saw himself as introducing to the Latin west. Like Aristotle, he produced a body of philosophical work that spanned the discipline in both breadth and depth. As in the case of Aristotle, some of the works attributed to Albert in his corpus are

not actually from his hand, and other works known to have been written by him have yet to be found. Little is known with any certainty about the chronology of the corpus, but there are good reasons for thinking that the bulk of his philosophical writings, in particular, his Aristotelian paraphrases, were completed between the years 1250 and 1270.

His corpus can be divided into three main categories: philosophy (nine treatises in logic, five in metaphysics, and three in ethics), theology (thirty treatises), and what we would call natural science but what throughout the medieval period was known as natural philosophy (twenty-two treatises). His method in most of his writings is the paraphrastic style employed by Avicenna (ibn Sīnā), as opposed to the line-by-line commentary characteristic of the works of Averroes (ibn Rushd), and his logical works in particular are deeply influenced by the work not only of Avicenna but also of al-Fārābī and Robert Kilwardby. Although Aristotle's scientific writings had been condemned in 1210 by Innocent III and the University of Paris established a commission to purge the Aristotelian corpus of heretical ideas in 1231, Albert encountered no difficulty in making use of Aristotelian ideas when he began to work on his *Summa de creaturis* (Treatise on creatures), before 1246, and his commentary on the *Sententiarum* of Peter Lombard, completed in 1249. It was probably not until the condemnation of 1277 that Aristotelianism as such encountered any serious resistance at the universities.

PHILOSOPHY

Part of what was at issue in the condemnation of 1277 was the relation between philosophy and theology, which the so-called Latin Averroists argued were separate disciplines corresponding to entirely distinct objects of knowledge, and hence governing different sorts of truths. The truths of theology were grounded in divine revelation and prophecy, while those of philosophy were grounded in human reason, and the mendicant orders were concerned to keep the two disciplines separate, on the grounds that philosophy, an inherently skeptical discipline, might intrude itself into theology in an unwarranted way, calling into question conclusions drawn in a domain in which it had no authority. In this context, Albert's insistence on the importance of knowing and understanding the philosophy of the ancient Greeks is striking and serves to illustrate his intellectual integrity.

Albert's approach to ancient philosophy has been criticized by late-twentieth-century historians of philosophy as an unrealistic syncretism of Aristotelianism and Neoplatonism. The complaint is that the two systems are philosophically and philologically incompatible, and any attempt to reconcile them is not only doomed to failure but is also methodologically misguided. It is worth noting, however, that this view is itself grounded in historical research based upon certain a priori assumptions about the relation between Plato's philosophical system and Aristotle's. Albert's Neoplatonism was essentially the Neoplatonism of the Greek commentators on Aristotle, which was itself an attempt to syncretize Plato and Aristotle, and it is fair to say that in antiquity the disparities between the two systems were not viewed as they have been by modern commentators. In fact, Albert, in offering a Neoplatonic harmonization of the two systems, is simply following the example, not only of his Arabic sources, but of a tradition that extends back to the Hellenistic period. The view that the systems are beyond harmonizing is of rather recent vintage and is subject to modification.

METAPHYSICS

Albert's metaphysics focused primarily on a theory of causation that can be traced to such sources as Aristotle, Avicenna, Pseudo-Dionysius, and the *Liber de causis* (*The Book of Causes*). He adapted the Neoplatonic notion of emanation of form, but in his system the causation is by attraction rather than by pure emanation from the One. He preferred attraction to pure emanation because he identified the One with the Good, and the Good, by its very nature, is diffusive of itself and of being (*diffusivum sui et esse*), that is, it causes other things to be by means of a kind of "calling to resemblance." (Albert here treats the word for good, "bonum," as cognate with the verb "boare" [to call]. This appeal to homespun etymology was also common in antiquity, particularly in Plato but also in Aristotle.) By virtue of this "calling to resemblance," the Good is not merely the first mover, as Aristotle's unmoved mover is, but is also the first producer, that is, the Creator—a role for the First Cause that is not found in Aristotle's *Metaphysics* (bk. ?), but rather is drawn from the *Liber de causis*, which Albert regarded as Aristotelian in provenance.

LOGIC

Albert's logical works consist, for the most part, of paraphrases of the treatises of the Organon (from Gr. "organōn," instrument, tool), so-called in the medieval period because logic was viewed not as a part of philosophy but rather as an implement that is necessary for the advancement of philosophy. The Organon consisted of

Aristotle's *Categories*, *De interpretatione* (On interpretation), *Topics* (including the *De sophisticis elenchis* [On sophistical refutations]), *Prior Analytics*, and *Posterior Analytics*. Yet Albert moved beyond Aristotle in a number of areas, most notably in his treatment of universals, which was grounded on the notion of form found in Plato and Aristotle. Aristotle had objected to the separability of the Platonic form and argued that forms are immanent in particulars. Drawing again upon Aristotle's Greek commentators, Albert argued that the universal must be analyzed into three *modi essendi*, or modes of being. Although a universal is a metaphysical unity, it may be considered under three aspects: as an entity in its own right, really existing separately from a particular, as in the mind of God (*ante rem*); as an entity that informs a particular, causing it to be the thing it is (*in re*); or as an entity in human thought (*post rem*). The distinction between the universal *in re* and the universal *post rem* is grounded in the Aristotelian notion of abstraction, which is discussed in more detail below under the heading of "Natural Science." Although Albert achieves here another notable syncretism, it is worth noting that he does not treat universals as substantial forms, as Plato and Aristotle both do.

NATURAL SCIENCE

Albert's interest in the natural world was driven by his belief that all knowledge is interconnected, and he pursued scientific questions with such intensity that critics, such as Henry of Ghent (*De scriptoribus ecclesiasticis* 2.10) suggested that he neglected theology and philosophy. Of particular interest with regard to his scientific writings is his attitude toward the distinction between rationalism and empiricism, a distinction that had been of great interest in antiquity but that had faded during the early medieval period as a consequence of both the ascendancy of rationalism under the influence of Neoplatonism and the decline in scientific investigations during periods of social and political upheaval. Working against the grain of the prevailing rationalism, Albert's attitude towards work in the natural sciences was decidedly empiricist: *experimentum solum certificat in talibus* ("Experience alone gives certainty in such matters" (*De vegetabilibus et plantis*, VI, 2.1). Although "experimentum" (here translated "experience") is reminiscent of our word "experiment," the modern concept of scientific experiment, in which a hypothesis is tested against observational data for confirmation or falsification, was unknown at this time.

For Albert, as for his contemporary Roger Bacon, the other great experimentalist of the thirteenth century, scientific "experiment" consisted in the gathering of observational data only, not the comparative analysis of data against hypotheses with controlled variables (The Latin word "experimentum" is cognate with the Greek word "empeiria" [experience], from which we get the English word "empiricism.") As in Aristotle's treatises on nature, observational data served only to illustrate or confirm a priori hypotheses, never as a means of hypothesis formation. But Albert is not a strict Aristotelian in this matter. For natural philosophers in the Aristotelian tradition, such as Aquinas, experience must be understood in terms of an inductive process leading from sense perception of particulars to the formation of general concepts in the soul, as described in Aristotle's *Metaphysics* (A.1) and *Posterior Analytics* (B.19).

In this account, the specific features of particulars are the proper objects of sense perception, but memory functions to gather together the perceptual information from similar particulars into what Aristotle calls an *empeiria* (experience) of the natural kind involved, and the rational faculty called *nous* in Greek (variously translated into English as either intellect or understanding) abstracts from *empeiria* an intelligible object, which then resides in *nous* and is a likeness (*homoiōma*) of the immanent form present in the particulars. Since these intelligible objects are different in kind from the perceptual objects that are the proper objects of the perceptual faculties, Aristotle is properly regarded not as an empiricist but as a rationalist. Nonetheless, experience clearly plays an essential role in the acquisition of knowledge of universals.

For Albert, although scientific knowledge is of the universal, the mechanism by which the universal comes to reside in the soul is by the "calling to resemblance" of the emanation of the intelligences. Intelligences illuminate the human rational faculty in accord with the doctrine of causation by attraction, and universal concepts form in the soul not because of the capacity of human intellect to abstract them but because the First Cause uses the intellect in its causal process. In Albert's and Bacon's reliance on experience, though different in kind from later notions of experience, we see the beginnings of the movement that would, by the time of the Renaissance, establish empiricism as the dominant scientific attitude, an attitude that, in time, would drive a wedge between natural philosophy and first philosophy and separate the natural sciences from philosophy.

See also al-Fārābī; Aristotelianism; Aristotle; Avicenna; Bacon, Roger; Liber de Causis; Neoplatonism; Peter Lombard; Pseudo-Dionysius; Thomas Aquinas, St.

Bibliography

WORKS BY ALBERT

Alberti Magni opera omnia, edited by the Institutum Alberti Magni Coloniense. Münster, Germany: Monasterium Westfalorum, 1951–.

Alberti Magni opera omnia. 21 vols., edited by Petri Jammy. Lyon, France: 1651.

Alberti Magni opera omnia. 38 vols., edited by Auguste Borgnet and E. Borgnet. Paris: L. Vivès, 1890–1899.

WORKS ON ALBERT

Bianchi, Luca. *Il vescovo e i filosofi: La condanna pariginia del 1277 e l'evoluzione dell'aristotelismo scolastico*. Bergamo, Italy: 1990.

Craemer-Ruegenberg, Ingrid. *Albertus Magnus*. Munich, Germany: Beck, 1980.

D'Ancona Costa, Cristina. *Recherches sur le "Liber de causis."* Paris: J. Vrin, 1995.

Hoenen, Maarten, and Alain de Libera. *Albertus Magnus und der Albertismus: Deutsche philosophische Kultur des Mittelalters*. Leiden, Germany: Brill, 1995.

Kovach, Francis J., and Robert W. Shahan, eds. *Albert the Great: Commemorative Essays*. Norman: University of Oklahoma Press, 1980.

Libera, Alain de. *Albert le Grand et la Philosophie*. Paris: J. Vrin, 1990.

Pegis, Anton. "St. Albert the Great and the Problem of Soul as Substance." In his *St. Thomas and the Problem of the Soul in the Thirteenth Century*, chap. 3. Toronto: St. Michael's College, 1934.

Weisheipl, James, ed. *Albertus Magnus and the Sciences: Commemorative Essays, 1980*. Toronto: Pontifical Institute of Mediaeval Studies, 1980.

Zimmermann, Albert. *Albert der Große: Seine Zeit, sein Werk, seine Wirkung*. Berlin: de Gruyter, 1981.

Scott Carson (2005)

COUNTERFACTUALS

A conditional is a sentence, statement, proposition, or thought of the form

If A then C

"A" is called the *antecedent* of the conditional and "C" the *consequent*. Philosophers have traditionally divided conditionals into two main groups, *indicative*, which can be symbolized as [A→C], and subjunctive ([A□→C]). The so-called *counterfactual* conditionals that have been the subject of so much discussion in analytic philosophy are subjunctive conditionals of the form

If it were to be the case that X then it would be the case that Y (if X were to happen, then Y would happen)

and

If it had been the case that X, then it would have been the case that Y (if X had happened, then Y would have happened)

Subjunctive conditionals of the form "If she be gone, he is in despair" are not at issue.

It is because the antecedents of such subjunctive conditionals usually state something that is not in fact the case or "contrary-to-fact," or is at least assumed not to be the case by the thinker or utterer of the conditional, that they have come to be known as *counterfactuals*.

It is not clear that there is any interesting difference between present and future tense indicative and subjunctive conditionals. It is not clear, for example, that there is any important semantic difference between one saying "If it were raining they would not be playing" and "If it's raining, then they're not playing." Nor is it clear that there is any important semantic difference between one saying "If she goes to the party, he will not go" and "If she were to go, he would not go," or between one saying "If salt is mixed with water it dissolves (will dissolve)" and "If salt were to be mixed with water it would dissolve." The idea that there is an important difference here is perhaps an artifact of the empiricist outlook dominant in analytic philosophy in the last century, which endorsed the "regularity theory of causation" and the associated idea that laws of nature could be adequately expressed by the "material conditional" of standard first-order logic.

However that may be, the difference between indicative and subjunctive conditionals seems clearer in the case of past-tense conditionals. Consider

If Georges Agniel and his friends did not discover the Lascaux caves, then someone else did

and

If Georges Agniel and his friends had not discovered the Lascaux caves, then someone else would have

The difference of meaning is immediately apparent and is sufficiently shown by the fact that although one takes the first to be true, one has no reason to believe the second.

The commonly used labels ("indicative," "subjunctive," and "counterfactual") do not, however, perspicuously mark out the set of conditionals that concern philosophers when they discuss counterfactuals. The indicative/subjunctive distinction is purely syntactical and simply fails to pick out the right set of conditionals. On the one hand, "If the Palestinians declared statehood now, the Israelis would retaliate" is a counterfactual that is not grammatically subjunctive. On the other hand, one

can utter a subjunctive conditional of the form "If X had happened, then Y would have happened" without having any intention to assert or imply the falsity of the antecedent. Suppose I am a detective who suspects that a criminal did A although none of my colleagues believe me. I note that the criminal did something peculiar, that is, B, and remark truly that if she had done A, she would have had to have done B in support of my case, without in any way implying that the state of affairs specified in the antecedent is not the case (alternatively, I may say this before dispatching someone to find out whether she did B). Again, I may set you a puzzle, asking you to work out what I have done, and give you clues, pointing out that if I had done X then this would have happened, that if I had done Y then this other thing would have happened, without ever asserting or implying that I did not do X or Y. Again, I may truthfully assert both "If I had come to the party I would have got drunk" and "If I had not come to the party I would have got drunk" without for a moment thinking or implying, inconsistently, that both these antecedents are false.

The purely syntactical criterion is no good, then, and blanket use of the term "counterfactual" to cover all the subjunctive conditionals that concern philosophers is no better. It remains true, nevertheless, that when one asserts a subjunctive conditional one almost invariably suggests that the state of affairs specified in the antecedent is not in fact the case. This entry will therefore use the traditional term "counterfactual" in this discussion, and contrast counterfactuals generally with indicatives in spite of the difficulties just noted.

THEORIES OF CONDITIONALS

Any theory of counterfactuals will be part of a general theory of conditionals, and the question arises as to what form a general theory of conditionals should take. Many favor a truth-conditional approach, that is, one that analyzes conditionals by offering an account of the conditions under which statements of the form "If A then C" are true or false (possible-worlds and metalinguistic accounts of conditionals are examples of truth-conditional approaches). Others seek to analyze conditionals by reference to the conditions under which they can be justifiably asserted or accepted as true (e.g., see Edgington 1986). An attractive alternative is John L. Mackie's (1973) condensed argument/supposition account, according to which conditionals are condensed arguments or suppositions and so not strictly true or false at all.

A central issue for any theory of conditionals is whether indicatives and counterfactuals should receive a uniform treatment, that is, one that uses the same theoretical apparatus across the board. David K. Lewis (1973, 1976) and Frank Jackson (1977, 1979) both reject this idea, offering nonuniform theories that fix the truth-conditions of indicatives and counterfactuals in different ways. Mackie (1973), by contrast, offers a uniform account of all conditionals in terms of the single basic notion of suppositions, and Robert C. Stalnaker (1968), having given an account of all conditionals in terms of possible worlds, accounts for the intuitive difference between indicatives and counterfactuals by appeal to pragmatic considerations.

Central to this debate is the question whether one bases one's account of indicative conditionals on the material conditional of standard first-order logic, often symbolized as "A⊃C," which is true just in case its antecedent is false or its consequent is true (the truth-value of the whole is determined in a purely truth-functional way by the truth-values of the parts). Lewis and Jackson are among those who think that the material-conditional approach can give an adequate account of all indicative conditionals (others think that it can only provide a necessary and not a sufficient condition), but a unified material-conditional account of both indicatives and counterfactuals seems a nonstarter. The material-conditional account, for example, classifies

If the moon had been made of cheese, I would be immortal

as just as surely true as

If this apple had been made of copper, it would have conducted electricity

simply on the ground that the antecedent is false. But one is much more discriminating about the truth-values of counterfactual conditionals than this account allows. That is why Lewis and Jackson, having accepted the material-conditional theory for indicatives, adopt a nonuniform general theory of conditionals, Lewis (1973) offering a possible-worlds account of counterfactuals and Jackson (1977) a causal account.

A further issue concerns whether one can give a uniform account of the logic of indicatives and counterfactuals. The following inference patterns

(I1) If A then C, therefore, if not-C then not-A (contraposition)

(I2) If A then B, if B then C, therefore, if A then C (hypothetical syllogism)

(I3) If A then B, therefore, if A and C then B (strengthening the antecedent)

are valid for the material conditional, but are widely agreed not to hold for counterfactual conditionals (e.g., consider the failure of (I3), in the move from the true claim "If he had walked on the ice, it would have broken" to the false claim "If he had walked on the ice and had been holding a large bunch of helium balloons, the ice would have broken"). While a nonuniform account can allow that these inference patterns hold for indicatives but fail for counterfactuals (see Lewis [1973] and Jackson [1979], who attempts to explain away apparently invalid indicative cases like "if he has made a mistake, then it is not a big mistake, therefore, if he has made a big mistake, he has not made a mistake" in terms of failure of assertibility), a uniform account must hold that if they fail for counterfactuals then they also fail for indicatives (see Stalnaker 1968).

THEORIES OF COUNTERFACTUALS

Turning now to counterfactuals, one finds three main approaches. The metalinguistic account initiated by Nelson Goodman in 1947 (see also Chisholm 1955, Mackie 1973, Tichy 1984) analyses counterfactuals in terms of an entailment relationship between the antecedent plus an additional set of statements or propositions, and the consequent. The causal approach offered by Jackson in 1977 (see also Kvart 1986) is closely related but deserves a separate category because it appeals essentially to causal concepts in its analysis of counterfactuals, thereby ruling out the popular strategy of using counterfactuals in an analysis of causation (one of the first to do this was Hume 1748/1975, p. 76; see also Lewis 1986b). Finally, there is the possible-worlds approach initiated by William Todd (1964), Stalnaker (1968), and Lewis (1973), which analyses counterfactuals in terms of similarity relations between worlds. This entry will consider them in turn, after hereby putting aside, as unimportant to the present concerns, all counterfactuals that are true (or false) as a matter of logic or *a priori* necessity, such as

If Q had been P it would have entailed P (Q)

If this number had been 2 it would have been even (odd)

If this circle had been square it would have had fewer than (more than) seven sides

THE METALINGUISTIC APPROACH. According to Goodman's (1947) metalinguistic approach a counterfactual asserts a certain connection or consequential relation between the antecedent and the consequent. Since in the case of the counterfactuals that concern this discussion the antecedent does not entail the consequent as a matter of logic or *a priori* necessity, certain other statements, including statements of laws and existing particular conditions, must be combined with the antecedent to entail the consequent. These counterfactuals, then, are true, if true at all, only if (and if) the antecedent combined with a set of statements S that meets a certain condition ϕ entails the consequent as a matter of law. The theory is metalinguistic because counterfactuals are treated as equivalent to metalinguistic statements of the relevant entailments.

A notorious difficulty for this theory has been to give an adequate specification of condition ϕ. Consider $[A \Box\!\!\rightarrow C]$. Given that the assumption, in the case of a counterfactual, is that A is false, one may reasonably assert ~A. However, if ~A were admissible into S, then with A one would get the contradiction [A&~A], and since it is generally accepted that anything can be inferred from a contradiction, anything could be inferred from the conjunction of A and S, including C. All counterfactuals would therefore turn out to be true (*a priori* false counterfactuals have been excluded). To prevent this trivialization, the statements that constitute S must be (logically) compatible with A. This excludes ~A. A further requirement noted by Goodman is that the statements that constitute S must be compatible with ~C; for if they were not, C would follow from S itself, and A and the laws would play no role in the inference to C.

With this in hand Goodman offers the following analysis: "A counterfactual is true if and only if (iff) there is some set S of true sentences such that S is compatible with C and with ~C, and such that [A&S] is self-compatible and leads by law to C; while there is no set S' compatible with C and with ~C and such that [A&S] is self-compatible and leads by law to ~C" (Goodman 1947, p. 120; for a discussion of this last condition, see Bennett 2003; Parry 1957). Restricting S with the notion of compatibility does not seem to be enough, however, for counterfactuals that clearly seem false still threaten to turn out true. Consider

(1) If match m had been struck, it would have flared

and

(2) If match m had been struck, it would not have been dry

Despite the restrictions on S, one gets the unacceptable result that (1) and (2) both turn out true. To see this, assume that it is a law that (L) when oxygen is present, dry matches flare when struck. Start with the situation of the dry match (D), the presence of oxygen (O), and suppose that the match has not been struck (~S) and has not flared (~F). O, D, and L are compatible both with S and with ~F, and with S, they imply F. Thus, (1) is true. Now, however, suppose ~F: that in fact the match has not flared. ~F, O, and L are compatible both with S and with D, but with S they imply ~D. Thus, (2) is true.

To eliminate this unwanted consequence, Goodman (1947) suggests that the relevant conditions in S must be cotenable with the antecedent. A is cotenable with B if it is not the case that B would have been false if A were true. ~F is thus compatible with S but not cotenable with it, because if the match had been struck (S), it would have flared (F). So (1) is true and (2) is false. However, this solution results in a circular definition or a regress, for counterfactuals are defined in terms of cotenability and cotenability is defined in terms of counterfactuals. Goodman proposed no solution to this problem (for a short discussion, see Bennett 2003, pp. 310–312).

THE CONDENSED ARGUMENT-SUPPOSITIONAL APPROACH. Closely related to the metalinguistic account is Mackie's (1973) condensed argument or suppositional account according to which all conditionals, including all counterfactuals, are condensed or abbreviated arguments that leave certain auxiliary premises unstated. Generally, to assert [A□→C] is to assert C within the scope of the supposition A (Mackie replaced the notion of a condensed argument by that of a supposition in an attempt to cover certain atypical conditionals that do not readily expand into arguments, e.g., "If that's a Picasso I'm a Martian").

There are two central ways in which Mackie's (1973, 1974) account differs from Goodman's (1947). First, Mackie abandons any metalinguistic element. In fact, according to Mackie, this feature of Goodman's account is the reason to reject it. Mackie argues that it simply "does not ring true" that when one asserts counterfactuals one is performing a higher-level linguistic act whose subject is a lower-level linguistic act. If-sentences are about the world, not about what is said about the world.

Second, Mackie relaxes the cotenability requirement on A and S. One does not need to provide an exact criterion of cotenability. All that one needs is the idea that the speaker assumes the cotenability of A and S and a notion of cotenability that can, he claims, be elucidated simply in terms of it being reasonable to combine a belief that S with A.

This suggestion is closely in line with what are sometimes called third-parameter views of counterfactuals (see Tichy 1984, who attributes this view to Chisholm 1955; Mill 1868; Ramsey 1931). According to this view, when a speaker asserts a counterfactual, he or she implicitly assumes a set of propositions. The counterfactual is true just in case the antecedent of the counterfactual and the assumed propositions entail the consequent and the implicitly assumed propositions are true. Since the implicitly assumed propositions depend on the attitudes of the speaker, no analysis of these propositions can be given and so the cotenability problem does not arise.

One point strongly in favor of such views is their ability to deal with ambiguous counterfactuals. Consider

> If Caesar had been in command in Korea, he would have used the atom bomb

> If Caesar had been in command in Korea, he would have used catapults

Although both counterfactuals can plausibly be asserted, they make different predictions about what would have happened. By introducing a third parameter this ambiguity can be located in the set of implicitly assumed propositions. The first counterfactual is asserted by someone who is assuming that Caesar was alive during the actual Korean War, and the second counterfactual is asserted by someone who is assuming that Caesar was involved in a war in Korea during Caesar's actual lifetime.

Jonathon Bennett (2003, pp. 305–308) objects to Chisholm's (1955) version of this solution to the cotenability problem, arguing that it implausibly requires that the asserter of [A□→C] have the assumed propositions in mind, although one can, for example, be sure that the lights would have gone off if one had turned the oven on again without knowing about the faulty electrical wiring in one's kitchen. He further argues that there are no limits to what a speaker could assume in asserting [A□→C], and that this lets in unwanted counterlogical conditionals like "if that piece of cast iron were gold some things would be malleable and not malleable."

THE CAUSAL APPROACH. Another theory closely related to the metalinguistic approach is Jackson's (1977) causal theory of counterfactuals, so-called because of the central role that causality plays in it. To determine the truth-value of a counterfactual one takes the causal laws at the actual world at a time. These determine the state of the world at later times. One then takes the state of the

world at the antecedent time, changes it as little as possible to make the antecedent true, and determines whether the causal laws predict subsequent states that make the consequent true.

More formally, [A□→C] is true at all the A-worlds satisfying the following:

(i) Their causal laws are identical with ours at the time of the antecedent and after

(ii) Their antecedent time-slices are the most similar to ours in particular facts

(iii) They are identical in particular fact to our world prior to the time of the antecedent

Sequential counterfactuals assert that if something had happened at one time, something else would have happened at a later time, and one difficulty for the theory is presented by asequential counterfactuals like:

If I had had a coin in my pocket, it would have been a Euro.

If Flintoff had not taken the winning wicket, Harmison would have (where this is understood as meaning that sooner or later one of them would have taken the winning wicket)

Jackson (1977) proposes to analyze asequential counterfactuals in terms of sequential counterfactuals. For example, one asserts the counterfactual about Flintoff and Harmison when one thinks that if Flintoff had failed to take the final wicket, events would have ensured Harmison's taking it (they were the only bowlers left and Australia was batting so poorly).

Jackson's account appeals to similarities between worlds. Does that mean that he is really giving a possible-worlds account of counterfactuals? Although he no longer objects to being classified as a possible-worlds theorist, in 1977 he drew a sharp division between his causal account and the possible-worlds account. He argued that a causal theorist about counterfactuals could avoid ontological commitment to possible worlds because the relevant similarities were things like the mass of an object or the magnitude of a force, similarities that could be characterized by reference to features of the actual world without any appeal to possible worlds.

THE POSSIBLE WORLDS APPROACH. In asserting a counterfactual one is of course standardly considering possibilities, how things would or might have been if certain other things had not been as they were, how things would or might be if things were not as they are, and the most influential treatment of counterfactuals has been the possible-worlds approach, which proposes to analyze counterfactuals by giving a rigorous account of their truth conditions and logical behavior using possible-worlds semantics. Stalnaker (1968) and Lewis (1973) are the most influential proponents of this view, and the basic idea is that the counterfactual [A□→C] is true just in case the closest possible A-worlds (worlds where A is true) are C-worlds (worlds where C is true), and the central notions are those of a possible world and the closeness relation. Both Stalnaker and Lewis introduce the idea of a "logical space," which is, roughly, a space of possible worlds. They locate the actual world in a "similarity structure" in such a logical space and make use of this similarity structure to determine the truth-values of counterfactuals.

More formally, for Stalnaker (1968)

[A□→C] is true iff A is impossible or C is true at $f(A, w^*)$

where f is a "selecting" function that takes the antecedent A and the actual world w^* as arguments and delivers a unique possible world as a value. The counterfactual is true if C is true at the possible world that f delivers as the value.

How exactly does the selection function select? The informal answer is that the selection is based on an ordering of possible worlds with respect to their similarity or resemblance to the actual world. More formally, for Lewis (1973)

[A□→C] is true iff either there is no A-world or some [A&C] world is more similar to the actual world than any [A&~C] world

It is convenient to represent Lewis's truth conditions in this way, with direct reference to similarity, although in his original presentation the ordering relation is explicated in terms of a system of spheres of worlds (for any possible world, all other possible worlds can be placed on spheres centered on that world, the sizes of the spheres representing how close those worlds are to that world. All worlds on a given sphere are equally close to the centered world, and inner spheres are closer to the centered world than outer spheres).

Lewis (1973) and Stalnaker (1968) agree that if the antecedent of a counterfactual is impossible than the counterfactual is trivially true. For Lewis, this is because there is no such A-world; for Stalnaker, function f selects the impossible world in which every statement is true. (It is not however clear that all impossible counterfactuals are alike in respect of truth. There is, intuitively, a difference between "If Picasso had been a sonnet, he would

have had fourteen lines" and "If Picasso had been sonnet, he would have had compound eyes," and Daniel Nolan [1997] and others argue that impossible worlds, like possible worlds, can be ranked with respect to comparative similarity to the actual world.) Lewis and Stalnaker also agree that inference patterns like contraposition, hypothetical syllogism, and strengthening the antecedent ((I1) to (I3) earlier) are invalid for counterfactuals. However, they disagree about the conditional excluded middle: [[A□→C] ∨ [A□→~C]] for all A and C. Stalnaker accepts it because according to his account there will always be one closest possible world, whereas Lewis accepts ties among closest possible worlds and so the principle is not universally true.

Stalnaker and Lewis also agree in analyzing the "closeness" relation in terms of similarity between worlds. However, what makes one world more similar to the actual world than another world? Kit Fine (1975) and Bennett (1974) object that Lewis's (1973) theory does not provide the correct truth conditions if closeness of worlds is understood in terms of our everyday intuitive notion of similarity. Intuitively, the counterfactual

If Nixon had pushed the button, there would have been a nuclear holocaust

seems true, and yet it is false by the lights of one commonsense notion of similarity, according to which a world in which a nuclear holocaust does not occur although Nixon presses the button is much more similar to our unholocausted world than a world where a nuclear holocaust does occur.

Lewis responds to this objection in "Counterfactual Dependence and Time's Arrow" (1979), claiming that a possible-worlds theory of counterfactuals does not need to appeal to any everyday notion of overall similarity. It is rather up to the theorist to work out a way of weighing factors relevant to overall similarity that will deliver the right truth-values for counterfactuals. Lewis offers the follows systems of weights:

> [i] It is of the first importance to avoid big, widespread, diverse violations of law
>
> [ii] It is of the second importance to maximize the spatiotemporal region throughout which perfect match of particular fact prevails
>
> [iii] It is of the third importance to avoid even small, localized, simple violations of law
>
> [iv] It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly (Lewis 1979, p. 473)

According to this system of weights, the Nixon counterfactual turns out true. Consider a world in which Nixon pushes the button and there is no nuclear holocaust; rather, events proceed in such a way as to match those in our world with perfect similarity. The trouble with claiming that this is the most similar world is that Nixon's pressing the button would have numerous effects (including the button's warming slightly, the subsequent state of Nixon's memory, and so on), and only a large miracle could wipe out all these changes. The worlds closest to ours are the ones that agree with our actual world until Nixon presses the button and then continue on in accordance with the laws of the actual world. (However, for a reformulation of the Nixon objection in the light of this reply, see Tooley 2003).

Many philosophers shy away from the apparent metaphysical commitments of the possible-worlds approach. For what is a possible world? Lewis's (1986c) answer that possible worlds are concrete entities, each as real as the actual world, seems to most hopelessly implausible, but there are many other views. Stalnaker's (1968) and Bennett's (2003) possible worlds, for example, are maximally consistent sets of propositions; Saul Kripke's are stipulations; and others hold that possible worlds are combinatorial constructions out of elements of the actual world.

Whatever one's view, and whether or not one wishes to appeal to possible worlds, counterfactual conditionals are the vehicles of two of the most fundamental forms of thought: "What if?" and "If only." They are central to imagination and invention, essential to curiosity and regret, essential, along with conditionals in general, to the fundamental capacities for debating, supposing, speculating, and hypothesizing that constitute the heart of one's intelligence.

See also Bennett, Jonathan; Chisholm, Roderick; Conditionals; Goodman, Nelson; Hume, David; Kripke, Saul; Lewis, David; Mackie, John Leslie; Modality, Philosophy and Metaphysics of; Response-Dependence Theories; Semantics.

Bibliography

Adams, Ernest W. "Subjunctive and Indicative Conditionals." *Foundations of Language* 6 (1) (1970): 89–94.

Bennett, Jonathan. "Counterfactuals and Possible Worlds." *Canadian Journal of Philosophy* 4 (1974): 381–402.

Bennett, Jonathan. *A Philosophical Guide to Conditionals*. New York: Oxford University Press, 2003.

Chisholm, Roderick M. "Law Statements and Counterfactual Inference." *Analysis* 15 (1955): 97–105.

Dudman, Vic H. "Three Twentieth-Century Commonplaces about 'If.'" *History and Philosophy of Logic* 22 (2001): 119–127.

Edgington, Dorothy. "Do Conditionals Have Truth-Conditions?" *Critica* 28 (52) (1986): 3–30.

Fine, Kit. "Review of Lewis's Counterfactuals." *Mind* 84 (335) (1975): 451–458.

Goodman, Nelson. "The Problem of Counterfactual Conditionals." *Journal of Philosophy* 44 (1947): 113–128.

Grice, Herbert Paul. "Logic and Conversation." In *Conditionals*, edited by Frank Jackson, 155–176. New York: Oxford University Press, 1991.

Hume, David. *Enquiries concerning Human Understanding and concerning the Principles of Morals* (1748), edited by L. A. Selby-Bigge. Oxford, U.K.: Clarendon Press, 1975.

Jackson, Frank. "A Causal Theory of Counterfactuals." *Australasian Journal of Philosophy* 55 (1977): 3–21.

Jackson, Frank, ed. *Conditionals*. New York: Oxford University Press, 1991.

Jackson, Frank. "On Assertion and Indicative Conditionals." *Philosophical Review* 88 (1979): 565–589.

Kvart, Igal. *A Theory of Counterfactuals*. Indianapolis, IN: Hackett, 1986.

Lewis, David K. "Causation." In *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986a. This was originally published in the *Journal of Philosophy* 70 (1973).

Lewis, David K. "Counterfactual Dependence and Time's Arrow." In *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986b. This was originally published in *Noûs* 13 (1979)

Lewis, David K. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

Lewis, David K. *On the Plurality of Worlds*. Oxford, U.K.: Blackwell, 1986c.

Lewis, David. K. *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986d.

Lewis, David K. "Postscripts to 'Counterfactual Dependence and Time's Arrow.'" In *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986e.

Lewis, David. K. "Probabilities of Conditionals and Conditional Probabilities." *Philosophical Review* 85 (1976): 297–315.

Mackie, John L. *The Cement of the Universe*. Oxford, U.K.: Clarendon, 1974.

Mackie, John L. *Truth, Probability, and Paradox*. Oxford, U.K.: Clarendon, 1973.

Mill, John S. *System of Logic*. London: Longmans, 1868.

Nolan, Daniel. "Impossible Worlds: A Modest Approach." *Notre Dame Journal of Formal Logic* 38 (4) (1997): 535–572.

Parry, William T. "Reexamination of the Problem of Counterfactual Conditionals." *Journal of Philosophy* 54 (1957): 85–94.

Pollock, John L. *Subjunctive Reasoning*. Dordrecht, Netherlands: D Reidel, 1976.

Ramsey, Frank P. "General Propositions and Causality." In *The Foundations of Mathematics*. New York: Humanities, 1931.

Sanford, David. *If P, then Q: Conditionals and the Foundations of Reasoning*. New York: Routledge, 1989.

Stalnaker, Robert C. "A Theory of Conditionals." *Studies in Logical Theory, American Philosophical Quarterly* 2 (1968): 98–112.

Tichy, Pavel. "Subjunctive Conditionals: Two Parameters vs. Three." *Philosophical Studies* 45 (1984): 147–179.

Todd, William. "Counterfactual Conditionals and the Presuppositions of Induction." *Philosophy of Science* 31 (1964): 101–110.

Tooley, Michael. "The Stalnaker-Lewis Approach to Counterfactuals." *Journal of Philosophy* 100 (7) (2003): 321–327.

Michelle Montague (2005)

EUDAIMONIA

Strictly speaking, the term "eudaimonia" is a transliteration of the Greek word for prosperity, good fortune, wealth, or happiness. In philosophical contexts the Greek word "eudaimonia" has traditionally been translated simply as "happiness," but a number of contemporary scholars and translators have tried to avoid this rendering on the grounds that it can suggest unhelpful connotations in the mind of the uncritical reader. (For example, it does not refer to an affective state, nor is it coextensive with the classical utilitarian conception of happiness, though both of these notions may, in some thinkers, count as aspects of *eudaimonia*.) Since the word is a compound of the prefix "eu-" (well) and the noun "daimōn" (spirit), phrases such as "living well" or "flourishing" have been proposed as possible alternatives. But the consensus appears to be that "happiness" is adequate if the term is properly understood within the philosophical context of antiquity.

Aristotle wrote that all agree that *eudaimonia* is the chief good for humans, but that there is considerable difference of opinion as to what *eudaimonia* consists in (*Nicomachean Ethics* I.2, 1095a15–30). The portrait of Socrates presented in Plato's early, Socratic dialogues has Socrates endorsing the view that *eudaimonia* consists in living a just life, which requires knowledge in the form of a kind of foresight (see especially *Gorgias*). In his later works (for example, the *Republic*), Plato continued to argue that virtue is sufficient for happiness, and that nonmoral goods do not add to *eudaimonia* (the so-called sufficiency thesis).

As is well known, Aristotle agreed that virtue is a necessary condition for *eudaimonia* but held that it is not sufficient (the so-called necessity thesis). On his account, "eudaimonia" is most properly applied not to any particular moment of a person's life, but to an entire life that has been well lived. While virtue is necessary for such a

life, Aristotle argued that certain nonmoral goods can contribute to *eudaimonia* or detract from it by their absence. There is some controversy among scholars as to how Aristotle finally characterized the happy life, the life marked by *eudaimonia*. Throughout the first nine books of the *Nicomachean Ethics*, he appears to think that a happy life is a life that centrally involves civic activity. The virtues that mark the happy person are themselves defined as states of the soul that arise out of certain interactions taking place in social relations. But in book X, Aristotle's argument appears to be that a life of contemplating the theoretical (*theoria*) is the happiest sort of life, and that civic involvement can actually detract from this sort of activity (though the private life of contemplation appears to presuppose the public life, since without the public life to produce goods and services, the philosopher is incapable of living in isolation).

Where Socrates, Plato, and Aristotle agreed was in the objective nature of *eudaimonia*, which set them sharply apart from the popular morality of their day. In a famous passage from the *Gorgias* (468e–476a), Socrates shocks Polus by arguing that a wrongdoer is actually worse off than the person whom he wrongs, and that any wrongdoer is bound to be unhappy until he is punished. The person who has been wronged, by contrast, may be happy in spite of whatever physical suffering he may undergo at the hands of the wrongdoer. The *Gorgias* concludes with a myth about the fate of the human soul after death that makes it clear that only the state of the soul, not the physical state of the body, determines whether one is happy or unhappy.

Although Aristotle did not agree that happiness cannot be diminished at all by physical suffering, it is not because he thought that feelings are decisive for happiness. On the contrary, he argued for an objective standard of human happiness grounded in his metaphysical realism. In *Nicomachean Ethics* (I.7), he argued that human excellence ought to be construed in terms of what ordinarily characterizes human life (the so-called function or *ergon* argument). This argument is clearly grounded in his doctrine of causation, according to which any member of a natural kind is characterized by four causes: a formal cause, a material cause, an efficient cause, and a final cause. The final cause is inextricable from the formal cause: To be a certain kind of thing is just to function in a certain way, and to have a certain sort of function is just to be a certain kind of thing. The human function (*ergon*) is to be found in the activity of our rational faculties, particularly practical wisdom (*phronēsis*) and learning (*sophia*). Since the activity of both of these faculties is ordered not by subjective considerations but by the formal constraints of reason itself, human excellence is objectively determined: To live well is to live a life characterized by the excellent use of one's rational faculties, and this excellence is marked by successfully applying general rules for virtuous living to particular situations calling for moral deliberation.

Aristotle rejected alternative accounts of happiness as falling short of his ideal in some way (*Nicomachean Ethics* I.5, 1095b14–1096a10). The life of political honor, for example, reduces happiness to the degree to which one is esteemed by others, thus disconnecting happiness from the operation of one's own proper function. A more popularly held view equated happiness with pleasure, a view that Aristotle quickly dismissed as failing to distinguish humans as a natural kind from other animals that also feel pleasure and that rely on it as a motivating force in their daily quest for survival. For Aristotle, as for Plato before him, the hedonistic view overlooks the essential function of human rationality: to order and control human appetites and desires, channeling them into activities that, in the long run, best ensure human flourishing. Indeed, it is this very order and control that distinguishes human society from all other forms of life, so that there is an intimate connection between human excellence and the political life. This connection is subject to a certain tension, however, since both Plato, in the *Republic*, and Aristotle, in his life of theoretical contemplation, make social order a necessary condition for human excellence while simultaneously arguing that personal happiness in some sense involves disconnecting oneself from the community at large.

The Stoics agreed that happiness is our ultimate end, for which all else is done, and they defined this as consistently living in accordance with nature. By this they meant not only human nature but the nature of the entire universe, of which we are a part, and the rational order that both exhibit. Practical reason thus requires an understanding of the world and our place in it, along with our resolute acceptance of that role. Following nature in this way is a life of virtue and results in a "good flow of life," with peace and tranquility.

The Epicureans also took *eudaimonia* to be the end for humans, but they defined "eudaimonia" in terms of pleasure. Yet many of the things we take pleasure in have unpleasurable consequences, which on balance disrupt our lives, and so do not provide us with the freedom from concerns (*ataraxia*) and the absence of physical pain (*aponia*) that characterize true happiness. These traits, they believed, must be secured through the exercise of

moderation, prudence, and the other virtues, yet they are not valued for their own sakes but as instrumental means to a life of pleasure and happiness.

This form of hedonistic eudaemonism is to be contrasted with the hedonism of the Cyrenaics, the main exception to Aristotle's statement that all agree that the highest good is *eudaimonia*. Sketchy accounts of the elder Aristippus suggest that his hedonism involved giving free reign to sensual desires (Xenophon, *Memorabilia* 11.1.1–34), so as always to be capable of enjoying the moment, making use of what was available (Diogenes Laertius 11.66). Later Cyrenaics refined this position as seeking to enjoy sensual pleasure to the full without sacrificing autonomy or rationality. Their conception of pleasure emphasized bodily pleasures, understood as either a kind of movement (*kinēsis*) or the supervening state of the soul (*pathos*). Because they regarded such transient states as the highest good, the Cyrenaics rejected the view that *eudaimonia*, a comprehensive and long-term type of fulfillment, is the end that should govern all our choices.

See also Aristotle; Cyrenaics; Epicurus; Phronêsis; Plato; Socrates; Sophia; Stoicism.

Bibliography

Ackrill, J. L. "Aristotle on *Eudaimonia*." *Proceedings of the British Academy* 60 (1974): 339–359.

Aristotle. *Nicomachean Ethics*. Translated by Christopher Rowe. Oxford, U.K.: Oxford University Press, 2002.

Annas, Julia. *The Morality of Happiness*. New York: Oxford University Press, 1993.

Broadie, Sarah. *Ethics with Aristotle*. New York: Oxford University Press, 1991. See especially chapter 1, "Happiness, the Supreme End," and chapter 7, "Aristotle's Values."

Cooper, John M. "Contemplation and Happiness: A Reconsideration." In his *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 212–236. Princeton, NJ: Princeton University Press, 1999.

Cooper, John M. "Intellectualism in the Nicomachean Ethics." In his *Reason and Human Good in Aristotle*, 144–182. Cambridge, MA: Harvard University Press, 1975.

Diogenes Laertius. *Lives of Eminent Philosophers*. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1972.

Gosling, J. C. B., and C. C. W. Taylor. "Epicurus." In their *The Greeks on Pleasure*, 345–364. Oxford, U.K.: Clarendon, 1982.

Irwin, Terence. "Socrates: From Happiness to Virtue." In his *Plato's Ethics*, 52–64. New York: Oxford University Press, 1995.

Plato. *The Collected Dialogues of Plato*. Princeton, NJ: Princeton University Press, 1963.

Vlastos, Gregory. "Happiness and Virtue in Socrates' Moral Theory." In his *Socrates, Ironist and Moral Philosopher*, 200–232. Ithaca, NY: Cornell University Press, 1991.

Xenophon. *Memorabilia*. Translated by Amy L. Bonnette. Ithaca, NY: Cornell University Press, 1994.

Scott Carson (2005)

EXPERIMENTATION AND INSTRUMENTATION

Experiment, William Herschel wrote, is a matter of "putting in action causes and agents over which we have control, and purposely varying their combinations, and noticing what effects take place" (Herschel 1966, p. 76). In this sense, the earliest recorded scientific experiments appeared in biological and medical contexts. In the second century CE, the physician Galen performed detailed animal experiments to find out about the functions of various organs. In the sixteenth century, Andreas Vesalius, pioneer in dissection, carried out elaborate experiments; and William Harvey, notwithstanding his Aristotelian orientation, supported his discovery of the circulation of the blood with painstaking experimental arguments. It is highly plausible that the practice of alchemy also served as an early source of experimentation. From the thirteenth century on, alchemists used laboratory equipment in order to create new agents and were arguing against the overly narrow interpretation of the art-nature divide in Aristotelian philosophy.

A third area where experimentation took place before the scientific revolution was supplied by Ptolemy's optics. Ptolemy, active in the second century CE, formulated an experimental, quantitative law of the refraction of light at the boundary of air and water and performed experiments to investigate binocular vision. In continuing this tradition in the early eleventh century CE, the Arab Ibn al-Haytham (Alhazen) wrote an impressive experimental treatise on optics in which he related in a mathematically demanding way the physics and geometry of light to the anatomy of the eye. Al-Haytham's work was translated into Latin in the thirteenth century and decisively influenced later optical research for a long time. Because of this and similar developments, Crombie saw experimental science of the modern world created by thirteenth-century philosophers of the West transforming Greek geometrical method and uniting it with the experimental habit of the practical arts.

All these different attempts of probing nature through experimental trials certainly contributed to the final emergence of experimentation in the seventeenth century as a self-conscious, methodically controlled and systematically used form of scientific experience. Galileo's

new conception of motion, which was based on experiment and measurement from about 1604 on, played an instrumental and decisive role in this (Schmitt 1969). In the second half of the century, scientific academies devoted themselves to the new science and became the primary centers of experimental activity.

From the seventeenth century on, experimentation increasingly meant the implementation of new or improved scientific instruments. Following a suggestion of Thomas S. Kuhn, we can group these instruments mainly into two categories according to their origin in the classical or the Baconian tradition of physical science (Kuhn 1976). The classical sciences comprise those mathematical disciplines like astronomy, geometrical optics, statics, harmonics, and geometry itself, which were first constituted in classical antiquity and experienced their major developments already then. With the exception of harmonics, the close connections of these fields with each other lasted way up into the nineteenth century. The instruments belonging to this tradition were often called "mathematical instruments" and are of a restricted variety: ruler and compass, balance, clock, and geometrical-astronomical devices. They served as aids to "mixed mathematics," which allowed for certain physical attributes in addition to the abstract mathematical ones. To experiment with them mostly meant to confirm a belief that was established beforehand by rational considerations, or to detail a fully established theory in a special respect. Many experiments performed in this tradition proved to be in reality only thought experiments—mental constructions of possible experimental situations whose results were thought to be predictable already from everyday experience. Even Galileo participated sometimes in this attitude.

The second tradition to which we can attribute many of the new instruments of the period is the Baconian one whose disciplines owe their status as sciences mainly to the experimental movement of the seventeenth century and to the practice of "natural histories," including those of the different practical arts that experienced a tremendous re-evaluation at the time. The barrier between the craft and scholarly traditions, which had so far separated the mechanical from the liberal arts, began to break down. To the Baconian sciences belong the studies of heat, electricity, magnetism, chemistry, metallurgy, glass making, and the like. The instruments of these fields were used to investigate nature under previously unobserved or non-existent conditions and were often called "philosophical instruments." During the next decades, the Baconian movement brought forth the telescope, the microscope, the thermometer and the barometer, the air pump, electric charge detectors, the Leyden jar, and many other contrivances. It is interesting to see that these instruments were primarily used in a qualitative way and that a strictly quantitative application came only very late, mainly at the end of the eighteenth or during the early nineteenth century when the two traditions, the classical and the Baconian, started to merge with each other. From about the middle of the seventeenth century on, the Baconian movement had adopted some form of the atomic or corpuscular philosophy and became the official "experimental philosophy" of the Royal Society.

PHILOSOPHICAL ASSESSMENTS OF EXPERIMENTATION AND INSTRUMENTATION

In the second book of his *Physics*, Aristotle had developed a contrast between "physis" and "techne," that is, between natural entities that have an innate principle of change—like plants, animals and humans, but also stones and clouds—and those that are artificially constructed, like bedsteads and clothes. Until the scientific revolution, Aristotelians used this nature-artifact divide as an argument against the epistemological relevance of experimentation. In order to understand nature, they claimed, one must not intervene with her order. Intervention would either invalidate nature's innate principles or play her a trick with mechanical contrivances, but would not lead to any genuine knowledge of natural reality. Instead, one must let nature pursue her own course and purposes and gain knowledge of her principles by closely observing them. The fact that techne or art is declared by Aristotle to be able to complete nature's unfinished processes or to imitate her does not change this state of affairs. To complete nature in regard to the behavior of a natural entity meant to remove all obstacles that might have come in its way; and to imitate nature denoted the general maxim to bring form and matter of an entity in an intricate union as nature does it with her beings.

It seems that the major author in providing a *philosophical* bridge over the art-nature divide was Francis Bacon (1561–1626). This justifies Kuhn's choice of using Bacon's name for a whole new tradition of experimentation. Bacon argued that art was only a special way of arranging a state of affairs in which nature herself will then produce an intended result. He redefined Aristotle's concept of form and took it as the key to the operational features of a natural being, leaving out the teleological dimension. The discovery of operational rules of an entity can now be identified with the true form or real

essence of relations among its simple natures. Consequently, Bacon rejected Aristotle's three other causes besides the formal one and took forms as "nothing more than those laws and determinations of absolute actuality which govern and constitute any simple nature, as heat, light, weight, in every kind of matter and subject that is susceptible of them" (*Nov. Org.* ii, XVII).

As a result, knowledge of our world cannot, according to Bacon, be read off from its surface, so to say. We can work our way through to the "viscera naturae," or nature's intestines, only by methodical and experimental procedures of induction. Perhaps Bacon's major insight was that simple enumerative induction, as taught by Aristotle, that is, induction without experiment and without the method of exclusion, is not enough to tell essential correlations from accidental ones.

Bacon's procedure of induction was taken as a valuable method of creating new empirical theories and laws way up into the twentieth century. The Baconian tradition culminated during the nineteenth century in John Stuart Mill's elaboration and refinement of Bacon's and Herschel's inductive rules. There is, however, a tendency visible in Mill to take experiment not quite with the same force as Bacon had taken it. For Bacon, experiment is inevitable if one wants to snatch secrets from nature—they never show up by themselves. Yet for Mill, situations are conceivable where observation can serve the same purpose as experiment: "For the purpose of varying the circumstances [in order to find out the real laws] we may have recourse … either to observation or to experiment; we may either *find* an instance in nature suited to our purposes, or, by an artificial arrangement of circumstances, *make* one. The value of the instance depends on what it is in itself, not on the mode in which it is obtained: its employment for the purposes of induction depends on the same principles in the one case and in the other, as the uses of money are the same whether it is inherited or acquired. There is, in short, no difference in kind, no real logical distinction, between the two processes of investigation" (*System of Logic*, III, vii, 2).

The belief that there is no "logical distinction" between observation and experiment became a matter of course for almost all the schools of philosophy of science of the entire twentieth century until the 1980s. It is interesting to see how an excellent nineteenth-century experimentalist, Hermann von Helmholtz, resisted this tendency, although he followed Mill in many other and important respects. His reasons, however, were different from Bacon's: If *I* can vary the conditions of an event in different respects, he argued, I can be sure that *my intervention* is the cause of observed change because I know of my will's impulse. If, however, I can only passively *observe* correlations without any help from me, I can never be sure whether these make up genuine causal relations or only accidental covariation (Helmholtz 1903). Whereas for Bacon it is the coyness of nature that compels humans to experiment, for Helmholtz it is the epistemological limitation of the passive mind that forces them to intervene in nature's course.

One of the strongest and most influential anti-inductive texts ever written is a chapter in Pierre Duhem's *Aim and Structure of Physical Theory* of 1906, titled "Physical Theory and Experiment." In order to show the general inadequacy of inductivism, Duhem picked the "Newtonian method" to pieces, as it appeared both in the hands of Newton himself as well as with Ampère's electrodynamics. He brilliantly showed that there is no question in Newton's celestial mechanics of any extraction of hypothesis by induction from experimenting, as Newton himself required in the *General Scholium*, nor in Ampère's mathematical theory of electrodynamic phenomena of any deduction "only from experiment," as stated already in the title of Ampère's treatise of 1827.

As a logical consequence, Duhem concluded that "in the course of its development, *a physical theory is free to choose any path it pleases provided that it avoids any logical contradiction; in particular, it is free not to take account of experimental facts*." It has to take account of them only "when the theory has reached its complete development" (Duhem 1974, p. 206; Duhem's emphasis). In order that experiment can unfold its true function—the testing of theories— it must be preceded by theory. Duhem intensified the priority of theory when he demanded that "this test by facts should bear exclusively on the conclusions of a theory, for only the latter are offered as an image of reality; the postulates serving as points of departure for the theory and the intermediary steps by which we go from the postulates to the conclusions do not have to be subject to this test."

Duhem's criticism was later taken up and continued by Karl Popper. In exactly the same spirit as Duhem, Popper decreed that "the theoretician puts certain definite questions to the experimenter, and the latter by his experiments tries to elicit a decisive answer to these questions and to no others" (Popper 1959, p. 107). For Popper therefore, it is only the theoretician who shows the experimenter the way, and never the other way around. The only function left for experiment is to liberate us from sterile and false theories. With Popper, experiment has altogether become the handmaiden of theory.

Duhem had even gone one step further than Popper in questioning the capability of experiment to fulfill this critical task of refuting theories as well. Even if a theory is mature enough to be tested, experiment cannot mechanically decide between it and its rival. "An experiment in physics can never condemn an isolated hypothesis but only a whole theoretical group" (p. 183). And it is hardly ever possible to decide trenchantly which of the many assumptions of a theoretical system is doubtful and responsible for the experimental contradiction. "The physicist concerned with remedying a limping theory resembles the doctor and not the watchmaker" (p. 188). A watchmaker, Duhem maintained, can take the broken watch apart and examine each component separately until he finds the defective one. The doctor, however, cannot dissect the patient to find out the problem, but has to guess its seat by inspecting disorders affecting the whole body. And even if all the assumptions of a theoretical group were known to be true except one, the rival group would not have been established as superior. This would be shown only if every possible alternative were conclusively eliminated. But we never know of course what alternatives remain to be discovered.

All these considerations led Duhem to explicitly condemn Bacon's idea of a "crucial experiment." Bacon had suggested that there do exist experiments that conclusively decide between competing theories. They do this in the way of *instantiae cruces* or "fingerposts" that are set up at crossroads to indicate the several directions. In 1951, W. V. O. Quine joined Duhem in rejecting crucial experiments. He generalized Duhem's argument to *all* of our empirical tenets. An unexpected unsuitable empirical observation does not only contradict a theoretical system, as Duhem had told us, Quine argued, but *all* our beliefs and theories: "Our statements about the external world face the tribunal of sense experience not individually but only as a corporate body. … The unit of empirical significance is the whole of science" (Quine 1961, p. 41f.). Quine used this claim for a searching critique of logical empiricism. One consequence of this is that any assumption apparently refuted by observation can be retained as true, so long as we are willing to make appropriate changes elsewhere in the system of our beliefs. This holistic argument for the underdetermination of theories by experience has become known as the "Duhem-Quine thesis."

The series of philosophical arguments to denigrate the role of experiments continued further into the 20th century. The logical empiricist Hans Reichenbach coined the influential distinction between "context of discovery" and "context of justification" which had been developed earlier by the philosophers Alois Riehl, Gottlob Frege and others under different names (Reichenbach 1951). According to this dichotomy, all the actual historical and social circumstances of the creation of a scientific theory, including its experimental generation, if there was one, cannot be used as reasons to justify it. Experiment can be good as a heuristic guide to hit upon a useful theory, but it is neither necessary nor sufficient for the validity of its results. As a result of Reichenbach's division all attention focused on the epistemology of theory and none on discovery and the possibilities of experiment.

Although Thomas S. Kuhn is routinely regarded as major critique of both logical empiricism with its forerunner Duhem and of Popper's critical rationalism, he was surprisingly enough in large agreement with his predecessors as far as the subordinate role of experiment is concerned—at least in his central work *The Structure of Scientific Revolutions* of 1962/1970. Unlike Reichenbach, however, Kuhn wanted to overcome the separation of discovery and justification, but the admissible discovery part of his logic considered the founding of theories again in overarching paradigms, but not in experiments. In this he followed his teacher Alexandre Koyré and others, who saw the success of modern science in the superiority of mathematically oriented Platonism over Aristotelianism with its "brute, common-sense experience" and over all other experimentally and technologically oriented historical endeavors. For Koyré as for Kuhn a scientific revolution is foremost an "intellectual mutation" (Koyré 1943, p. 400), i.e. a revolution of thought and not of momentous experimental innovation. Paradigms have priority over theories "in their conceptual, observational, and instrumental applications" (Kuhn 1970, p. 43). True experimental research is only possible, if questions to nature are posed in a suitable mathematical language. According to such a view, a history of experimentation could not only be a contingent epiphenomenon of the development of paradigms and would not have much explanatory value. (The contrary view is defended by deSolla Price 1984.) Only when in his later work he began to appreciate the Baconian sciences as an autonomous movement did Kuhn start to appreciate the possibility of a meaningful history of scientific experimentation (Kuhn 1976).

In retrospect, the discussion of experiment in philosophy of science from the late nineteenth century until the 1980s appears as a series of increasingly negative results: We know more and more what experiments don't accomplish and we understand better and better where earlier

epistemic pretensions of experimentation find their limits. As a result, we can diagnose an "invisibility of experiment." In the same way as scientific revolutions of a field are, according to Kuhn, normally invisible to the scientific profession of the present, so experiments and their development remain largely invisible to philosophy of science because their exclusive role of testing theories seems ingrained in the ideology of its practitioners.

THE NEW EXPERIMENTALISM

Since the early 1980s, however, a change has taken place in the attitude of the study of science toward experiment. One can detect a growing awareness of the rich history of experimentation and of the vast variety of its (nondemonstrative) functions. This swing of appreciation is primarily due to detailed work of historians and sociologists of science. It is true that historiography never ceased to deal with experiment, but it had rarely put it into the center of its interest. Socio-historical analysis has now come to concentrate much more on the microstructure of experiment than before and has started to consider all kinds of other sources besides official reports, like diaries and laboratory notebooks. Especially rich sources are Faraday's laboratory notebooks and letters, Ampère's "dossier" in the archive of the Académie des Sciences and Hans Krebs' laboratory diaries and interview protocols (Gooding 1990, Steinle 2005, Holmes 1993, Graßhoff 2000). Historians even went so far as to replicate historical experiments with rebuilt apparatus and to hereby bring to light neglected or otherwise hidden dimensions of experimentation (Heering 2000). Sociologists tried to show that the formulation of experimental results requires special structures of communication in the scientific community and that there is a good deal of negotiation involved until an experimental result is considered as achieved (Shapin and Shaffer 1985, Licoppe 1996; for a discussion see Holmes 1992). The variety of fields from where these case studies come from raise hopes that the traditional concentration on physics in relation to experiment will soon be done with once and for all.

It was Ian Hacking's *Representing and Intervening* that set the ball rolling in philosophy of science. There are two phrases from Hacking's book that became the slogans of "new experimentalism": "If you can spray them, then they are real" and "Experimentation has a life of its own" (Hacking 1983, pp. 23, 150). The first catchphrase stands for a novel argument in favor of scientific realism. The philosopher's favorite theoretical entity is the electron—never given directly to our senses, but central to modern particle physics. There is an endless debate between scientific realists and their opponents whether explanatory success of a theory is ground for belief in the reality of its theoretically postulated entities. Hacking does not think very highly of this "inference to the best explanation," on which the ordinary scientific realist bases her belief in the reality of the electron. He rather sets high hopes in the fact that if you spray, say, a niobium ball with electrons, it makes a difference in the world: it decreases the charge of the niobium ball. "From that day forth," Hacking confesses, "I've been a scientific realist." In a way, Hacking's argument is a version, adapted to scientific antirealism, of Dr. Johnson's refutation of Bishop Berkeley's metaphysical antirealism concerning matter by kicking a stone. "It is not thinking about the world but changing it that in the end must make us scientific realists."

With the second catchphrase Hacking opposes the alleged theory-domination of experimentation: There actually exists experimental practice, he argues, that is *not* subordinate to theory and this practice actually proves to be very important. This claim is backed up with many intriguing examples. But liberating experiment from permanent condemnation to the role of theory's handmaiden does not automatically show what other roles it can take on and what the principles of their variations are. About this, Hacking does not say very much. The only other role he addresses in detail is, as he says, the experiment's "chief role": the "creation of phenomena." Some aspects of this role have been brought to light in Steinle's concept of "exploratory" experiments or in Heidelberger's notion of "productive" instruments (Steinle 2005; Heidelberger 1998, 2003).

All in all, Hacking seems to be largely content with a "Baconian fluster of examples of many different relationships between experiment and theory" (Hacking 1983, p. 66). This has surely proven to have been enough to initiate a "Back-to-Bacon movement, in which we attend more seriously to experimental science" (p. 150) as it had been Hacking's intention. But, if neo-Baconianism is sound, it is not enough as an explanation of what happens or should happen with other theoretical commitments of general philosophy of science, like, for example, the theory-ladenness of observation. This doctrine—dear to many philosophers of science for other reasons—comes, at least *prima facie*, into conflict with Hacking's faith in the priority of experiment.

In the wake of renewed interest in experiment, several substantial studies and edited volumes have appeared. Many of them are divided over the philosophical issue whether experiment can decide between competing theories and thus have an objective meaning or

whether social and political factors are in the end responsible for scientific development. There is, for example, Pickering's sociological history of particle physics or Collins's study of gravity wave detection maintaining the social construction of scientific evidence whereas Franklin and Mayo argue for the existence of strategies that secure reliable experimental outcomes and thus of rational belief. It would be wrong, however, to perpetuate the polarization between history, sociology, and philosophy of science. One of the results of taking experiment more seriously is precisely the insight that these dichotomies have to be transcended. An attempt into this direction has been made by Rheinberger who takes "experimental systems" as functional research units, especially of the life sciences (see Hagner and Rheinberger 1998 for a programmatic overview.) They are made up of research objects, theories, experimental arrangements, instruments, as well as disciplinary, social, cultural and institutional constellations that for some time crystallize in a certain stable configuration.

EXPERIMENTATION AND THEORY-LADENNESS

The idea of theory-ladenness of experience enabled a powerful and effective criticism of logical empiricism. This is the view already encountered with Popper that there are no theory-neutral data and that the meaning of observational terms fundamentally depends upon the theoretical context in which they occur. This view can easily be strengthened to serve as the cornerstone of a constructivist and anti-empiricist account of science: The categories in terms of which we carve up our experience are not read off from the external world but follow from prior theoretical or other commitments of its observers, either individually or socially.

The implications of theory-ladenness for a view of scientific experimentation are straightforward: If observations are theory-laden and if experimentation involves observation of results, then experimentation has to be theory-laden too. Since experiments, according to this view, make sense only in relation to some theoretical background, they cannot play a role that is independent from theory.

Now, the question arises: If new experimentalism is right, do we have to give up the idea of theory-ladenness? It is difficult to imagine a straightforward "yes" as an answer, because the general spirit in which the idea of theory-ladenness has been formulated is largely the same as that of the idea that experimentation has a life of its own. It is the spirit addressed by Hacking at the beginning of his book in which philosophers finally realized that they "long made a mummy of science"—the same spirit which, in the face of history and the reality of the laboratory, denies the "Popper/Carnap common ground." To deny theory-ladenness would to some extent feel like a return to logical empiricism and thereby of mummification, even if the autonomy of experiment is the reward.

Before some kind of dénouement of this question is formulated, let us have a closer look at theory-ladenness as it appeared in the work of its most important originators. One of the first propagators of this outlook was Pierre Duhem who wrote: "An experiment in physics is the precise observation of phenomena accompanied by an *interpretation* of these phenomena; this interpretation substitutes for the concrete data really gathered by observation abstract and symbolic representations which correspond to them by virtue of the theories admitted by the observer. … The result of an experiment in physics is an abstract and symbolic judgment" (Duhem 1974, p. 147). It would not be enough for an experimental report to state, as a layman would express it, that a piece of iron carrying a mirror oscillates. Instead it should read that the electrical resistance of a coil is measured. This shows that the physicist draws conclusions from experiment only in abstract and symbolic terms "to which you can attach no meaning if you do not know the physical theories admitted by the author." In sciences less advanced than physics like physiology or certain branches of chemistry "where mathematical theory has not yet introduced its symbolic representations" and where causal explanation reigns instead of a causally neutral description, the experimenter can reason "directly on the facts by a method which is only common sense brought to greater attentiveness" (p. 180).

This kind of theory-ladenness by theoretical interpretation, as we can call it, is very often confounded with another sort which was provided by Norwood Russell Hanson in 1958 and which can be called "theory-ladenness by prior belief or knowledge." "Seeing an object *x*," Hanson wrote, "is to see that it may behave in the ways we know *x*'s do behave" (Hanson 1958, p. 22). As a result of this, Tycho and Kepler watching the sun at dawn would literally see different things: Tycho who believes in the geocentric theory sees the sun beginning its diurnal circuit, whereas Kepler as defender of heliocentrism sees the earth spinning back into the light of the sun. "Analogously," Hanson wrote, "the physicist sees an X-ray tube, not by first soaking up reflected light and then clamping on interpretations, but just as you see this page before you."

In addition, theory-ladenness in science means "causality-ladenness" for Hanson, being loaded with causal meaning. He does not exclude theory-neutral talk after all, but it only happens in the oculist's office or like circumstances but not in scientific observation or experimentation. This shows that Hanson rejects all of Duhem's points: (1) Seeing an experimental result is not interpreting it; (2) both the layman and the physicist have prior beliefs and therefore both their seeing is theory-laden; and (3) physical theory (as well as common beliefs about the world) is causal theory and not just causally neutral description. Whereas for Hanson any injection of causality into the mere registering of facts is bound to render them theoretical, for Duhem, theory begins with the representation of (causal) relations in an abstract, causally neutral structure.

In Thomas Kuhn's work we find several different conceptions of theory-ladenness that are not always separated clearly. The most frequently used is similar to Hanson's, except that it is not prior knowledge that shapes perception, but paradigm and that it stresses and utilizes the psychology of perception even more than in Hanson: "Something like a paradigm is prerequisite to perception itself. What a man sees depends both upon what he looks at and also upon what his previous visual-conceptual experience has taught him to see" (Kuhn 1970, p. 113).

In order to exhibit his other uses of theory-ladenness, let us have a look at Kuhn's treatment of scientific discovery. Kuhn admits the possibility of "fundamental novelties of fact," that go *against* a well-established paradigm. Without this possibility, as he himself realizes, science could only develop in a theoretical manner and never by adjustment to facts. "Discovery commences with the awareness of anomaly, i.e., with the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science" (Kuhn 1970, pp. 52–53).

Where, according to Kuhn, does a violation of the paradigm-induced expectations come from? Does it come from a causal process that violates the received view or from a new theoretical interpretation that makes old facts appear in a new light? It seems that in Kuhn, it is almost always the *theoretical interpretation*, the assimilation to theory, that is decisive for discovery and hardly ever any causal experience. "Assimilating a new sort of fact demands a more than additive adjustment of theory, and until that adjustment is completed—until the scientist has learned to see nature in a different way—the new fact is not quite a new fact at all." That sounds more as if new facts and causal processes were created by new paradigms than the other way around. Lavoisier, we are told, for example, was enabled through his new paradigm "to see in experiments like Priestley's a gas that Priestley had been unable to see there himself" and was "to the end of his life" unable to see p. 56).

The only case where Kuhn explicitly admits that discovery has been effected by a genuinely novel causal experience appears to be the case of the X-rays. "Its story opens on the day that the physicist Roentgen interrupted a normal investigation of cathode rays because he had noticed that a barium platino-cyanide screen at some distance from his shielded apparatus glowed when the discharge was in process" (p. 57). Although Kuhn seems to consider this observation theory-laden, I maintain that, in Duhem's sense, it is not. If it were, Roentgen, by definition of theory-ladenness, would have been able to interpret it in light of the theories of physics he had at his disposal. But here it is exactly the point that his theories deserted him and he could *not* find a place for this new experience in his customary theoretical structure. For this reason he interrupted his investigation and asked himself why the screen had come to glow. Yet the novel observation is certainly theory-laden in the sense of Hanson, because Roentgen immediately looked for a causal relationship between his apparatus and the glowing of the screen, although this went completely against all his expectations!

Kuhn seems to say that Roentgen would never have paid attention to the glowing screen if he had not disposed of deeply entrenched theories of physics that *prohibited* such a phenomenon. If this is true then we have here a third sense of the notion of theory-ladenness before us. It frames a psychological hypothesis about the ease with which a phenomenon is detected or paid attention to in the light of a *contradicting* paradigm: An observation is theory-laden in this sense if it were improbable that an observer would have made it (that an observer would have noticed it or would have attributed any importance to it) without her holding a theory beforehand that created expectations *to the contrary*. It would be better to drop the term "theory-ladenness" for this case altogether and instead call it "theory-guidance" because the experimental result made sense to Roentgen as an observation in its simple causal structure already *without* the theoretical background of the theory that guided it or any other one. "Theory-guidance" refers to a psychological disposition how well one is prepared to notice a particular phenomenon in certain situations.

After Roentgen had noticed the anomaly, he conducted various experiments in order to explore the *cause*

of the incident: "Further investigations—they required seven hectic weeks during which Roentgen rarely left the laboratory—indicated that the cause of the glow came in straight lines from the cathode ray tube, that the radiation cast shadows, could not be deflected by a magnet, and much else besides. Before announcing his discovery, Roentgen had convinced himself that his effect was not due to cathode rays but to an agent with at least some similarity to light" (Kuhn 1970, p. 57). This is perhaps the only place in his book where Kuhn uses the term "cause" (or an equivalent) in relation to an experimental investigation. The quotation shows vividly that Roentgen did not conduct his experiments in order to test a theory but to expand our knowledge of causal connections in relation to the scientific instruments and devices involved.

What does our discussion suggest therefore as the most adequate description of Roentgen's early experiments? They were certainly theory-guided in the sense of Kuhn and they were, or immediately became, causality-laden in the sense of Hanson, but not (or not yet) theory-laden in the sense of Duhem (which Kuhn also shares). Kuhn is right when he suggests that only after the phenomena had received an abstract and symbolic representation can we speak of a "discovery" of X-rays. Yet before this interpretation has taken place, we can say that an anomaly has occurred and that it can be replicated in certain ways; not more, but also not less.

If the case of the X-rays is in this way correctly understood, then Kuhn can give in to Hacking without loosing anything essential and admit that experimentation can be, and very often is, autonomous and free from theory. The lesson to learn is to distinguish between two kinds of experiments: those that are causal, but not (yet) embedded in a theoretical structure and those that presuppose the knowledge of such a framework. This emphasis of an autonomous "lower level" in experimentation is not a relapse into positivist observation statements and protocol sentences allegedly giving meaning to theory. The claim rather is that two types of experimentation should conceptually be kept apart: experimentation at the causal level, where the manipulation of instruments and objects under scrutiny takes place, and experimentation taking place at the theoretical level, where the results at the causal level are represented in a theoretical superstructure.

See also Ampère, André Marie; Aristotelianism; Aristotle; Bacon, Francis; Berkeley, George; Carnap, Rudolf; Duhem, Pierre Maurice Marie; Faraday, Michael; Frege, Gottlob; Galen; Galileo Galilei; Harvey, William; Helmholtz, Hermann Ludwig von; Herschel, John; Johnson, Samuel; Kepler, Johannes; Kuhn, Thomas; Lavoisier, Antoine; Logical Positivism; Mill, John Stuart; Newton, Isaac; Philosophy of Science, History of; Platonism and the Platonic Tradition; Popper, Karl Raimund; Priestley, Joseph; Quine, Willard Van Orman; Realism; Reichenbach, Hans; Riehl, Alois; Scientific Method; Thought Experiments in Science; Underdetermination Thesis, Duhem-Quine Thesis.

Bibliography

Bacon, Francis. *Novum Organum*. London: 1620.

Batens, Dederik, and Jean P. van Bendegem, eds. *Theory and Experiment: Recent Insights and New Perspectives on Their Relation*. Dordrecht: Reidel, 1988.

Buchwald, Jed Z. *Scientific Practice: Theories and Stories of Doing Physics*. Chicago: University of Chicago Press, 1995.

Collins, Harry M. *Changing Order: Replication and Induction in Scientific Practice*. Beverly Hills, CA, and London: Sage, 1985.

Crombie, Alistair C. *Robert Grosseteste and the Origins of Experimental Science, 1100–1700*. Oxford: Clarendon, 1953.

deSolla Price, Derek J. "Of Sealing Wax and String." *Natural History* 93 (1) (1984): 48–57.

Duhem, Pierre. *The Aim and Structure of Physical Theory*. New York: Atheneum, 1974. Translated by Philip P. Wiener from the 2nd ed. Paris: Chevalier & Rivière, 1914.

Franklin, Allan. *Experiment, Right or Wrong*. Cambridge, U.K.: Cambridge University Press, 1990.

Franklin, Allan. *The Neglect of Experiment*. Cambridge, U.K.: Cambridge University Press, 1986.

Galavotti, Maria Carla, ed. *Observation and Experiment in the Natural and Social Sciences*. Dordrecht: Kluwer, 2003.

Galison, Peter. *How Experiments End*. Chicago: University of Chicago Press, 1987.

Gooding, David. *Experiment and the Making of Meaning: Human Agency in Scientific Observation*. Dordrecht: Kluwer, 1990.

Gooding, David, Trevor Pinch, and Simon Schaffer, eds. *The Uses of Experiment: Studies in the Natural Sciences*. Cambridge, U.K.: Cambridge University Press, 1989.

Graßhoff, Gerd, R. Casties, and Kärin Nickelsen. *Zur Theorie des Experimentes. Untersuchungen am Beispiel der Entdeckung des Harnstoffzyklus*. Bern: Bern Studies, 2000.

Hacking, Ian. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge, U.K.: Cambridge University Press, 1983.

Hacking, Ian. "The Self-Vindication of the Laboratory Sciences." In *Science as Practice and Culture*, edited by Andrew Pickering, 29–64. Chicago: University of Chicago Press, 1992.

Hagner, Michael, and Hans-Jörg Rheinberger. "Experimental Systems, Objects of Investigation, and Spaces of Representation." In *Experimental Essay—Versuche zum Experiment*, edited by Michael Heidelberger and Friedrich Steinle, 355–373. Baden-Baden: Nomos, 1998.

Hanson, Norwood Russell. *Patterns of Discovery*. Cambridge, U.K.: Cambridge University Press, 1958.

Heering, P., Falk Rieß, and C. Sichau, eds. *Im Labor der Physikgeschichte. Zur Untersuchung historischer Experimentalpraxis*. Oldenburg: BIS-Verlag, 2000.

Heidelberger, Michael. "Die Erweiterung der Wirklichkeit im Experiment." In *Experimental Essays—Versuche zum Experiment*, edited by Michael Heidelberger and Friedrich Steinle, 71–92. Baden-Baden: Nomos, 1998.

Heidelberger, Michael. "Theory-Ladenness and Scientific Instruments in Experimentation." In *The Philosophy of Scientific Experimentation*, edited by Hans Radder, 138–151. Pittsburgh: University of Pittsburgh Press, 2003.

Heidelberger, Michael, and Friedrich Steinle, eds. *Experimental Essays—Versuche zum Experiment*. Baden-Baden: Nomos, 1998.

Helmholtz, Hermann von. "Die Tatsachen in der Wahrnehmung [1878]." In *Vorträge und Reden*, 213–247. Braunschweig: Vieweg, 1903.

Herschel, John F. W. *Preliminary Discourse on the Study of Natural Philosophy* (1830). New York: Johnson Reprint Corporation, 1966.

Holmes, Frederic L. "Do We Understand Historically How Experimental Knowledge Is Acquired?" *History of Science* 30 (1992): 119–136.

Holmes, Frederic L. *Hans Krebs: Architect of Intermediary Metabolism 1933–1937*. Oxford: Oxford University Press, 1993.

Holmes, Frederic L. "The Logic of Discovery in the Experimental Life Sciences." In *Biology and Epistemology*, edited by Jane Maienschein, 167–190. Cambridge, U.K.: Cambridge University Press, 2000.

Holmes, Frederic L., Jürgen Renn, and Hans-Jörg Rheinberger, eds. *Reworking the Bench: Research Notebooks in the History of Science, New Studies in the History and Philosophy of Science and Technology*. Dordrecht: Kluwer, 2003.

Holmes, Frederic L., and Trevor H. Levere, eds. *Instruments and Experimentation in the History of Chemistry*. Cambridge, MA: MIT Press, 2000.

Koyré, Alexandre. "Galileo and Plato." *Journal of the History of Ideas* 4 (4) (1943): 400–428.

Kuhn, Thomas S. "Mathematical vs. Experimental Traditions in the Development of Physical Science." *Journal of Interdisciplinary History* 7 (1) (1976): 1–31. Reprinted in *The Essential Tension: Selected Studies in Scientific Thought and Change*, 31–65. Chicago: University of Chicago Press, 1977.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press, 1970.

Latour, Bruno, and Steve Woolgar. *Laboratory Life: The Social Construction of Scientific Facts*. 2nd ed. Princeton, NJ: Princeton University Press, 1986.

Le Grand, Homer E., ed. *Experimental Inquiries: Historical, Philosophical and Social Studies of Experimentation in Science*. Dordrecht: Kluwer, 1990.

Licoppe, Christian. *La formation de la pratique scientifique: le discours de l'expérience en France et en Angleterre* (1630–1820). Paris: Editions La Decouverte, 1996.

Mayo, Deborah G. *Error and the Growth of Experimental Knowledge, Science and Its Conceptual Foundations*. Chicago: University of Chicago Press, 1996.

Mill, John Stuart. *A System of Logic, Ratiocinative and Inductive*. London: Parker, 1843.

Pickering, Andrew. *Constructing Quarks: A Sociological History of Particle Physics*. Chicago: University of Chicago Press, 1984.

Pickering, Andrew. *The Mangle of Practice: Time, Agency, and Science*. Chicago: University of Chicago Press, 1995.

Pickering, Andrew, ed. *Science as Practice and Culture*. Chicago: University of Chicago Press, 1992.

Popper, Karl. *The Logic of Scientific Discovery*. London: Hutchinson, 1959.

Quine, Willard Van Orman. *From a Logical Point of View*. New York: Harper & Row, 1963.

Radder, Hans, ed. *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 2003.

Reichenbach, Hans. *The Rise of Scientific Philosophy*. Berkeley: University of California Press, 1951.

Schmitt, Charles B. "Experience and Experiment: Comparison of Zabarella's View with Galileo's in *De Motu*." *Studies in the Renaissance* 16 (1969): 80–138.

Shapin, Steven, and Simon Schaffer. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press, 1985.

Steinle, Friedrich. *Explorative Experimente. Ampère, Faraday und die Ursprünge der Elektrodynamik*. Stuttgart: Franz Steiner, 2005.

Michael Heidelberger (2005)

MODALITY AND LANGUAGE

Modality is a category of linguistic meaning having to do with the expression of possibility and necessity. A modalized sentence locates an underlying or prejacent proposition in the space of possibilities (the term *prejacent* was introduced by medieval logicians). *Sandy might be home* says that there is a possibility that Sandy is home. *Sandy must be home* says that in all possibilities Sandy is home. The counterpart of modality in the temporal domain should be called temporality, but it is more common to talk of tense and aspect, the prototypical verbal expressions of temporality. Together, modality and temporality are at the heart of the property of displacement (one of Charles F. Hockett's design features of human language) that enables natural language to talk about affairs beyond the actual here and now.

There are numerous kinds of expression that have modal meanings, the following is just a subset of the variety one finds in English:

(1) Modal auxiliaries
Sandy must/should/might/may/could be home.

(2) Semimodal verbs
Sandy has to/ought to/needs to be home.

(3) Adverbs
Perhaps, Sandy is home.

(4) Nouns
There is a slight possibility that Sandy is home.

(5) Adjectives
It is far from necessary that Sandy is home.

(6) Conditionals
If the light is on, Sandy is home.

It is traditional to use English modal auxiliaries or semimodal verbs as the primary source of illustrative examples. This is in spite of the fact that these elements have a rather curious set of grammatical properties. Indeed, it appears that modal meanings are part of a natural logical vocabulary and thus elements with modal meanings easily become part of the inventory of grammatical or functional morphemes, which are typically associated with idiosyncratic, nonproductive grammatical characteristics (for a cross-linguistic survey of this process, compare Bybee, Perkins, Pagliuca 1994).

KINDS OF MODAL MEANING

One can distinguish different kinds of modal meaning. Alethic modality (Greek: *aletheia*, meaning "truth"), sometimes logical or metaphysical modality, concerns what is possible or necessary in the widest sense. It is in fact hard to find convincing examples of alethic modality in natural language, and its inclusion in this list is primarily for reason of historical completeness. The following categories, however, are of primary importance in the study of natural language. Epistemic modality (Greek: *episteme*, meaning "knowledge") concerns what is possible or necessary given what is known and what the available evidence is. Deontic modality (Greek: *deon*, meaning "duty") concerns what is possible, necessary, permissible, or obligatory, given a body of law or a set of moral principles or the like. Bouletic modality, sometimes boulomaic modality, concerns what is possible or necessary, given a person's desires. Circumstantial modality, sometimes dynamic modality, concerns what is possible or necessary, given a particular set of circumstances. Teleological modality (Greek: *telos*, meaning "goal") concerns what means are possible or necessary for achieving a particular goal. In the descriptive literature on modality, there is taxonomic exuberance far beyond these basic distinctions.

FLEXIBILITY OF MEANING

Many modal expressions can be used to express many or all these kinds of modal meaning. Witness the English semimodal *have to* in the following set of examples:

(7) It has to be raining. [after observing people coming inside with wet umbrellas; epistemic modality]

(8) You have to go to bed in ten minutes. [stern father; bouletic]

(9) Visitors have to leave by six p.m. [hospital regulations; deontic]

(10) I have to sneeze. [given the current state of one's nose; circumstantial]

(11) To get home in time, you have to take a taxi. [teleological]

Some modal expressions are more specialized in what kind of meanings they can carry. The English auxiliary *might* is most comfortable expressing epistemic modality.

(12) It might be raining.

Some modals only occur in specialized environments. The modal *need* with a bare infinitive complement can only occur in negative environments:

(13) a. You need not worry.
b. *You need worry.

(14) Nobody need worry.

Such negative polarity modals occur in other languages as well (compare the Dutch *hoeven* and the German *brauchen*).

POSSIBLE WORLDS SEMANTICS

In technical work on natural language semantics, modality is analyzed with the machinery of possible worlds semantics, developed by logicians for the artificial language of modal logic. The most influential incarnation of this idea is found in the work of the semanticist Angelika Kratzer (1981, 1991).

The starting tenet is that modal expressions express quantification over possible worlds—regardless of what those might be (most practitioners have few ontological scruples). Possibility modals correspond to existential quantification, while necessity modals correspond to universal quantification. Different kinds of modal meaning correspond to different choices of sets of possible worlds as the domain of quantification. These sets of possible worlds are assigned to the world in which the complex sentence is evaluated (the evaluation world) by an accessibility relation.

The accessibility relation underlying epistemic modality delivers as the domain of quantification for the modal those worlds that are compatible with what is known, with the available evidence in the evaluation world. Similarly, deontic modality quantifies over worlds that satisfy the relevant body of law or principles. Bouletic modality quantifies over worlds that conform to what the relevant person desires.

Actually, Kratzer (1981, 1991) argues that modal meaning does not just rely on an accessibility relation but also on an ordering of the accessible worlds. The clearest argument for this complication of the semantics comes from deontic cases. Imagine a city whose traffic bylaws outlaw the practice of double parking at any time for any reason. The bylaws further specify that anyone who is found guilty of double parking must pay a considerable fine. Robin has been found guilty of double parking, so the following sentence seems to be true:

(15) Robin must pay a fine.

Notice, however, that in all the worlds that conform to the traffic bylaws there never occurs any double parking, since that is against the law. Therefore, in none of those worlds does Robin pay a fine for double parking. Thus, the simple possible worlds analysis incorrectly predicts the sentence to be false.

Kratzer's (1981, 1991) analysis makes modal expressions doubly relative: they need to be interpreted relative to (1) a set of accessible worlds (modal base), and (2) an ordering of those worlds. For the case in hand, the accessible worlds would be those where Robin's actions hitherto are what they are (double parking occurs) and that from then on develop in many conceivable ways. The ordering would be that induced by the traffic bylaws, which would favor among the accessible worlds those where Robin pays a fine. The truth-conditions of this example are then that in all the favored worlds among the accessible worlds Robin pays a fine. The sentence could be made false either if Robin did not in fact double park or if the traffic bylaws do not in fact require a fine.

The surface variety of modal meanings is thus a product of the interplay of three factors: (1) the quantificational strength (possibility, necessity, and shadings in between, e.g. slight possibility), (2) the modal base, and (3) the ordering source.

Epistemic modality has an epistemic modal base and either no ordering or an ordering based on plausibility or stereotypicality. Deontic modality has a circumstantial modal base (because one may have to abstract away from one's knowledge that the right thing will not be done) and an ordering source based on a body of law or principles. Bouletic modality again has a circumstantial modal base and an ordering source based on a relevant person's desires.

There is much detailed research remaining to be done on the fine distinctions between different modal expressions. Consider, for example, the fact that *ought to* and *have to* somehow differ in strength in their deontic use:

(16) You ought to call your mother, but of course you don't have to.

Or, consider the fact (explored by Ninan 2005) that deontic *should* and deontic *must* differ whether one can admit that the right thing will not happen:

(17) I should go to confession, but I'm not going to.

(18) #I must go to confession, but I'm not going to.

There is also an interesting literature on fine details of epistemic meaning. Work by Ian Hacking (1967), Paul Teller (1972), and Keith DeRose (1991) shows that there is much additional complexity and context-dependency behind the phrases *what is known* or *the available evidence*, which are typically used to characterize epistemic accessibility. In particular, the context may specify whose knowledge or evidence base is relevant to the claim made with an epistemically modalized sentence.

CONTEXT-DEPENDENCY AND LEXICAL SPECIALIZATION

Kratzer (1981, 1991) argues that rather than treating the multitude of modal meanings as a case of (accidental) polysemy, it should be seen as the outcome of context-dependency. In other words, modal expressions have in of themselves a rather skeletal meaning and it is only in combination with the background context that they take on a particular shade of meaning (such as epistemic or deontic). She points to ways of making explicit what the intended conversational background is:

(19) According to the hospital regulations, visitors have to leave by six p.m.

(20) Considering the evidence before us, it has to be raining.

In the absence of such explicit markers, natural language users need to rely on contextual clues and reasoning about each other's intentions to determine what kind of modal meaning a particular sentence is intended to express in its context of use.

As seen earlier, some modals are not entirely subject to the whims of context but impose their own preferences as to what kind of modal meaning they would like to express. English *might* likes to be epistemic (with some interesting exceptions, such as the use in *You might try to put the key into this slot*, which has the force of a suggestion). This kind of behavior is not uncommon for expressions that are context-dependent: pronouns refer to contextually furnished individuals but may include restrictions on what the context can furnish, for example, the gender marking on *she* requires that the context furnish a female individual.

It has been shown that there is a recurring historical development where a modal expression that initially has a nonepistemic meaning only (something that for opaque reasons is often called a root modal) develops over time into an expression that also has epistemic meanings (e.g., Nordlinger and Traugott [1997] document this development for the case of English *ought to*).

THE ARGUMENT STRUCTURE OF MODALS

So far, this entry has been presupposing that modality concerns the possibility or necessity of a prejacent proposition. There is, however, an ancient and persistent doctrine that another kind of modality concerns the possible or necessary existence of a relation between a subject or agent and a predicate. For example, one finds the claim that deontic modality can at least sometimes concern what an agent is permitted or obliged to do.

(21) Sandy ought to call his mother.

The propositional analysis has it that the sentence expresses the necessity of the prejacent proposition that Sandy calls (will call) his mother, relative to the current circumstances and a body of ethics, for example. The predicate-level analysis has it that the sentence expresses that the agent Sandy and the property of calling his mother stand in a certain modal relation. Some authors call this the *ought to be* versus *ought to do* distinction. Certain sentences are clearly cases of propositional-level *ought to be* modality:

(22) There ought to be a law against double parking.

For sentences with an agentive subject, it is an open question, debated in the technical literature, whether a predicate-level or propositional-level analysis is correct. Whatever one's position in this debate is, one has to admit that some sentences with human subjects still do not express an obligation imposed on that subject:

(22) Jimmy ought to go in his crib now. [said of a six-month-old baby]

FURTHER AND RELATED CATEGORIES

At the outset, this entry listed a set of expressions that have modal meanings. The list was far from complete. Here, some other types of expressions that may fall under the general category of modality or at least belong to adjacent categories will be added.

A closely related category, perhaps subsumable under modality, is evidentiality. Various languages regularly add markers, inflectional or otherwise, to sentences that indicate the nature of the evidence that the speaker has for the prejacent proposition. A typical evidential system might centrally distinguish between direct and indirect evidence. The latter concept might be further subdivided into indirect reasoning from direct evidence or conclusions based on hearsay or the like. The standard European languages do not have elaborate evidential systems but find other ways of expressing evidentiality when needed. The English adverb *apparently* seems to prefer indirect evidence:

(24) Kim has apparently been offered a new job.

The German modal *sollen* has a hearsay interpretation:

(25) Kim soll einen neuen Job angeboten bekommen haben.
Kim soll a new job offered get have
"Kim has supposedly been offered a new job."

Another important category is mood, an inflectional marking on the main verb of a sentence, which expresses some kind of modal meaning. English has only a rudimentary mood system, if that. However, Romance languages, for example, productively use mood. In Italian, the complement clause of a verb like *say* occurs in the indicative mood, while the complement of *believe* appears in the subjunctive mood. There are attempts at analyzing the mood selection in such cases as depending on technical properties of the possible worlds semantics of the embedding verb. The research topic remains active and thriving.

Propositional attitude constructions are also related to modality. Consider the near equivalence of the following two sentences:

(26) Robin suspects that the butler is guilty.

(27) Given Robin's evidence, the butler might be guilty.

Jaako Hintikka (1969) proposed to analyze propositional attitudes with the same possible worlds machinery that was originally applied to modals, thus making the

relation between the two categories explicit in their semantics.

Expressions of illocutionary force are also within or close to the field of modality. Consider in particular attenuating speech act markers, as explored in pioneering work by J. O. Urmson (1952):

(28) The butler is, I suspect, guilty.

The difference between attenuated assertion of a proposition and categorical assertion of a modalized proposition is small, one suspects.

One particular kind of expression deserves attention: the modal particles that are rampant in some languages, such as German:

(29) Kim hat ja einen neuen Job.
Kim has JA a new job
"Kim has a new job, as you may know already"

The gloss here is only approximate, the meaning of the modal particles is elusive and under active investigation.

Modality is a pervasive feature of natural language and sometimes it clearly appears in the semantics of an expression without a clear syntactic or morphological exponent. Such hidden modality can be detected, for example, in infinitival relatives in English (for extensive discussion, see Bhatt 2005):

(30) When you have computer trouble, Sandy is the person to talk to. [≈ Sandy is the person one ought to talk to]

Sometimes the source for the modality can be identified but its etymology and nature remains opaque:

(31) What Arlo is cooking has garlic in it.

(32) Whatever Arlo is cooking has garlic in it. [epistemic modality triggered by *-ever*: speaker does not know what precisely Arlo is cooking]

The range of modal expressions is a rich domain for language-internal and cross-linguistic investigations.

MODALITY WITHOUT CONTENT?

So far, this entry has assumed that modalized sentences express complex propositions with a possible worlds-based quantificational meaning built on top of a prejacent unmodalized proposition. While this is indeed the standard analysis in formal natural language semantics, it is not the standard assumption in descriptive and typological linguistics.

The most common analysis in descriptive work treats modality as an expression of the speaker's attitude toward the prejacent proposition, rather than giving rise to a complex proposition with its own distinct content. The prevalence of this conception can perhaps be traced back to the influence of Immanuel Kant, who wrote in his *Critique of Pure Reason* that "the modality of judgments is a very special function thereof, which has the distinguishing feature that it does not contribute to the content of the judgment" (1781, p. 74). This idea seems to have influenced both practicing linguists and a subset of logicians, including Gottlob Frege, who wrote in his *Begriffsschrift* that "[b]y saying that a proposition is necessary, I give a hint about the grounds for my judgment. But, since this does not affect the conceptual content of the judgment, the form of the apodictic judgment has no significance for us" (1879, p. 5).

It may be that scholars have typically adopted one of the two conceptions without much reflection. Within the descriptive literature, there is rarely any argumentation for the speaker's comment analysis. And the formal semantic literature rarely addresses the issue either, basically ignoring the preponderance of the speaker's comment analysis in the descriptive literature.

One rather straightforward prediction of the speaker's comment analysis is that modalized sentences should not be easily embeddable. This prediction seems to be false for at least some standard modal expressions:

(33) It might be that visitors have to leave by six p.m. [epistemic modality embedding a deontic modality]

Such iterated modality is unexpected from the point of view of the speaker's comment analysis. Better cases for a comment analysis come from speech act markers:

(34) #If yesterday, I suspect, was the worst day of the year, the market is in good shape.

The suspicion arises that some modal expressions have a comment-type meaning, while others contribute to the propositional content of the complex sentence. There is here, it seems, the opportunity for empirical and theoretical debate on this issue. It should be noted that the question here is related but not identical to the issue of whether a modal element expresses subjective or objective modality (these terms are discussed by Lyons 1977).

Independently of these ideas from descriptive linguistics, there are proposals that would give modals a meaning that goes beyond truth-conditions. In dynamic semantics, epistemic modals are treated as particular operations on an information state, see, for example, Veltman (1996). Finally, at least for deontic modals, it has

been suggested that they can be used with performative force, whether or not they also have propositional content. Kamp (1973, 1978) and Lewis (1979) explore the idea that deontic 'may' is used to grant permission, while Ninan (2005) explores the idea that deontic 'must' is used to issue commands.

COMPOSITIONAL INTERACTIONS

As the examples of iterated modality in the previous section showed, at least some, if not most, modal expressions can compositionally interact with other expressions. Interactions with negation, quantifiers, and tense are particularly interesting.

The combination of modals with negation is a fountain of idiosyncratic facts. Consider that English *may* scopes under negation when read deontically, but scopes above negation when read epistemically:

(35) He may not have any cake. [deontic, "not allowed"]

(36) He may not be home. [epistemic, "possible that not"]

Or, consider that English *must* scopes above negation (in either reading) while German *müssen* scopes under negation:

(37) a. He must not have any cake. ["obligatory that not"]
b. He must not be home. ["evident that not"]

(38) Er muss nicht zuhause bleiben.
He must not at-home remain
"He doesn't have to stay home."

Lastly, note that while *can* does not easily allow an epistemic reading, negated *cannot* does have an epistemic reading:

(39) a. Sandy can be home. [?]
b. Sandy cannot be home. [epistemic]

Most of these facts have resisted systematic explanation and remain mysterious.

Sentences containing both modals and quantificational noun phrases are often ambiguous:

(40) Most of our students must get outside funding …
a. for the department budget to work out.
b. the others have already been given university fellowships.

In some of the literature, this ambiguity is assimilated to the distinction between *de dicto* and *de re* interpretations, probably inappropriately. In any case, it has been observed that not all sentences show this ambiguity. For example, epistemic modals seem to resist having quantifiers scope over them (for an exploration, see von Fintel and Iatridou 2003):

(41) Most of our students must be home by now.
a. must ≻ most of our students
b. *most of our students ≻ must

Again, this kind of fact remains mysterious, it may be an idiosyncratic syntactic fact without any grounding in semantics.

The interaction of modality and temporality is intricate and ill understood. One should first note that the aspectual nature of the prejacent sentence has a strong influence on what kind of meaning a modal sentence can carry. A nonstative prejacent typically gives rise to deontic readings, while a stative prejacent is compatible with both epistemic and deontic readings:

(42) He has to be in his office. [epistemic/deontic]

(43) He has to see his doctor this afternoon. [nonepistemic]

While modal auxiliaries do not inflect for tense (the fact that *might* may be a past-tense inflected form of *may* has reasons in the mist of history), other expressions do allow such inflection.

(44) He had to be in his office.

It is not always obvious whether what is happening here is that the modal sentence is located in the past or whether the modal has scope over a past-tense prejacent. The preceding sentence, when read epistemically, is plausibly ambiguous, reporting a past deduction about a simultaneous state of affairs or a present deduction about a past state of affairs. Finally, some modals in embedded positions seem not to express any modal meaning of their own but occur in "agreement" or "harmony" with a higher modal or mood. One relevant case is "I am convinced that it must be raining." See Portner (1997) for discussion.

CONDITIONALS

An interaction of modals with other expressions that is of paramount importance is their appearance in conditional constructions. It has been noticed again and again that for sentences of the form *if p, modal q* it is hard to find a compositional interpretation that treats the *if*-construction as expressing some kind of conditional meaning, while the modal in the consequent expresses its usual modal meaning.

Consider, for example, the following conditional:

(45) If Robin double parked her car, she must pay a fine.

A tempting idea is that the conditional construction introduces universal quantification over epistemically accessible worlds and says that the consequent is true in all epistemically accessible worlds where Robin double parked her car. The consequent in turn is true in an evaluation world if in all worlds circumstantially accessible from that world and favored by the deontic ordering source Robin pays a fine. However, now assume that one knows that Robin is invariably law abiding. She would never do anything that contravenes any law. So, among the epistemically accessible worlds there are none where she double parks against the law, so if she double parked, that must be consistent with the law. Hence, the above sentence would come out false. However, this seems wrong. The sentence does not make a claim about what the law must be like if Robin double parked her car. What it claims is that the actual law is such that double parking necessitates a fine.

The conclusion drawn from this and many parallel examples with other modal operators is that it is a mistake to analyze such structures as involving two-layered operators: a conditional construction embedding or embedded in a modal construction. Rather, the idea has been to say that in such sentences, the *if*-clause does not supply its own operator meaning but serves as a restriction on the modal base of the modal operator. The proper analysis of the previous sentence is that it says that among those circumstantially accessible worlds where Robin double parked her car, the ones favored by the law as it is in the actual world are all worlds where Robin pays a fine.

After surveying a number of such cases, Kratzer summarizes the thesis as follows, "[T]he history of the conditional is the story of a syntactic mistake. There is no two-place *if … then* connective in the logical forms of natural languages. *If*-clauses are devices for restricting the domains of various operators" (1986).

What about bare conditionals such as *If Sandy's light is on* she is home? Here, there is no modal operator for the *if*-clause to restrict. Should one revert to treating *if* as an operator on its own? Kratzer (1986) proposes that one should not and that such cases involve covert modal operators—in this case, possibly a covert epistemic modal. This entry has nothing to say about that here.

This entry has shown that the topic of modality is characterized by rich empirical detail, considerable cross-linguistic variation, and intriguing theoretical issues. The following bibliography can serve as a start for further reading and exploration.

See also Artificial and Natural Languages; Conditionals; Hintikka, Jaakko; Kant, Immanuel; Modality, Philosophy and Metaphysics of; Philosophy of Language; Possibility; Propositional Attitudes: Issues in Semantics; Semantics.

Bibliography

van der Auwera, Johan. "On the Typology of Negative Modals." In *Perspectives on Negation and Polarity Items*, edited by Jack Hoeksema et al., 23–48. Amsterdam, Netherlands: Benjamins, 2001.

van der Auwera, Johan, and Vladimir A. Plungian. "Modality's Semantic Map." *Linguistic Typology* 2 (1998): 79–124.

Bhatt, Rajesh. *Covert Modality in Non-finite Contexts*. de Gruyter, 2005.

Bybee, Joan L., Revere Perkins, and William Pagliuca. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: Chicago University Press, 1994.

Condoravdi, Cleo. "Temporal Interpretation of Modals: Modals for the Present and for the Past." In *The Construction of Meaning*, edited by David I. Beaver et al. Stanford, CA: CSLI Publications, 2002.

DeRose, Keith. "Epistemic Possibilities." *Philosophical Review* 100 (4) (1991): 581–605.

Farkas, Donka. "On the Semantics of Subjunctive Complements." In *Romance Languages and Modern Linguistic Theory*, edited by Paul Hirschbühler and Konrad Koerner, 69–104. Amsterdam, Netherlands: Benjamins, 1992.

von Fintel, Kai, and Sabine Iatridou. "Epistemic Containment." *Linguistic Inquiry* 34 (2) (2003): 173–198.

Frege, Gottlob. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: L. Nebert, 1879.

Hacking, Ian. "Possibility." *Philosophical Review* 76 (2) (1967): 143–168.

Hintikka, Jaako. "Semantics for Propositional Attitudes." In *Philosophical Logic*, edited by J. W. Davis, D. J. Hockney, and W. K. Wilson, 21–45. Dordrecht, Netherlands: D. Reidel, 1969.

Hockett, Charles F., and Stuart A. Altmann. "A Note on Design Features." In *Animal Communication: Techniques of Study and Results of Research*, edited by Thomas A. Sebeok, 61–72. Bloomington: Indiana University Press, 1968.

Kamp, Hans. "Free Choice Permission." *Proceedings of the Aristotelian Society, New Series*, 74: 57–74, 1973.

Kamp, Hans. "Semantics versus Pragmatics". In *Formal Semantics and Pragmatics for Natural Languages*, edited by Franz Guenthner and S. J. Schmidt, pp. 255–288. Dordrecht: Reidel, 1978.

Kant, Immanuel. *Critik der reinen Vernunft*. Riga: Johann Friedrich Hartknoch, 1781.

Karagjosova, Elena. *Modal Particles and the Common Ground: Meaning and Functions of German ja, doch, eben/halt, and auch*. Amsterdam, Netherlands: Benjamins, 2003.

Kiefer, Ferenc. "Modality." In *The Encyclopedia of Language and Linguistics*, edited by Ronald E. Asher, 2515–2520. Oxford, U.K.: Pergamon, 1994.

Kiefer, Ferenc. "On Defining Modality." *Folia Linguistica* 21 (1) (1987): 67–94.
Kratzer, Angelika. "Conditionals." *Chicago Linguistics Society* 22 (2) (1986): 1–15.
Kratzer, Angelika. "Modality." In *Semantics: An International Handbook of Contemporary Research*, edited by Arnim von Stechow and Dieter Wunderlich, 639–650. Berlin: de Gruyter, 1991.
Kratzer, Angelika. "The Notional Category of Modality." In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, edited by H. J. Eikmeyer and H. Rieser, 38–74. Berlin: de Gruyter, 1981.
Lewis, David. "A Problem about Permission". In *Essays in Honour of Jaako Hintikka: On the Occasion of His Fiftieth Birthday on January 12, 1979*, edited by Esa Saarinen, Risto Hilpinen, Ilkka Niiniluoto, and Merril Provence Hintikka, pp. 163–175. Reidel, 1979.
Lyons, John. *Semantics*. New York: Cambridge University Press (1977).
Ninan, Dilip. "Two Puzzles about Deontic Necessity." In *New Work on Modality*, edited by Valentine Hacquard et al. MIT Working Papers in Linguistics, no. 52. Department of Linguistics and Philosophy, MIT, Cambridge, MA, 2005.
Nordlinger, Rachel, and Elizabeth Traugott. "Scope and the Development of Epistemic Modality: Evidence from *ought to*." *English Language and Linguistics* 1 (1997): 295–317.
Palmer, Frank Robert. *Mood and Modality*. 2nd ed. New York: Cambridge University Press, 2001.
Plungian, Vladimir A. "The Place of Evidentiality within the Universal Grammatical Space." *Journal of Pragmatics* 33 (2001): 349–357.
Portner, Paul. "The Semantics of Mood, Complementation, and Conversational Force." *Natural Language Semantics* 5 (2) (1997): 167–212.
Stowell, Tim. "Tense and Modals." In *The Syntax of Time*, edited by Jacqueline Guéron and Jacqueline Lecarme, 621–636. Cambridge, MA: MIT Press, 2004.
Teller, Paul. "Epistemic Possibility." *Philosophia* 2 (1972): 302–320.
Urmson, J. O. "Parenthetical Verbs." *Mind* 61 (1952): 192–212.
Veltman, Frank. "Defaults in Update Semantics". *Journal of Philosophical Logic* 25(3) (1996): 221–261.
Willett, Thomas. "A Cross-linguistic Survey of the Grammaticalization of Evidentiality." *Studies in Language* 12 (1) (1988): 51–97.
von Wright, Georg Henrik. *An Essay in Modal Logic*. Amsterdam, Netherlands: North-Holland, 1951.

Kai von Fintel (2005)

PHRONÊSIS

Often translated as "practical wisdom," the Greek word *phronêsis* derives from the verb *phronein*, meaning "to have understanding," or "to be wise or prudent." In its earliest uses the word is normative only in the sense that it signifies a correct cognitive grasp of some kind; only gradually does it come to be used in ethical contexts for a correct grasp of what ought to be done. For Plato and the other Socratics, *phronêsis* represents that aspect of our rational faculty that derives genuine knowledge about values and norms, that is, about the virtues (see especially *Protagoras, Gorgias*). The famous debate between the Socratics and their critics, such as the orator Isocrates, turned on the possibility of demonstrative knowledge in the sphere of virtue. Plato had attacked oratory on the grounds that its aim is not to discover what is morally right, but merely to persuade, and he offered in its place the Socratic method of dialectic, a cooperative search for the truth by means of hypothesis formation, critical examination and refutation, and hypothesis modification. Isocrates had characterized Socratic dialectic as mere eristic (*Against the Sophists* 1; *Antidosis* 261) or argument for argument's sake—probably for this reason, Plato is especially careful to distinguish the Socratic method from mere eristic in his *Euthydemus*—and referred to the Socratics as "disputers." But Plato devotes much argument to showing how the careful examination of various conceptions of the virtues can lead inexorably to a recovery of their essential nature, which resides in the soul of every person from birth.

Aristotle's treatment of *phronêsis* (*Nicomachean Ethics* VI.5 1140a24–b30; cf. 1141b8–1143a5) is similar in many respects to Plato's, but in his account the knowledge that we obtain of virtue is not the equivalent of scientific (demonstrative) knowledge (*episteme*): unlike *episteme*, which is concerned with necessary truths, *phronêsis* is always concerned with contingent truths. Aristotle defines *phronêsis* by reference to something more concrete and familiar, namely, the practically wise person, *ho phronimos*, someone who has *phronêsis*. It is the mark of the practically wise person, he says, to be able to deliberate well about what is good and advantageous for himself not merely in one area, such as health or strength, but as a means to human flourishing in general. The operation of *phronêsis* in Aristotle's account of the rational faculties appears to hinge on the application of general rules for right conduct (the *orthos logos*) to the particular circumstances of a given situation so as to result in action that will generally tend toward human flourishing. The *phronimos* is the person whose life is characterized by such applications of *phronêsis* and who, as a result, tends to flourish throughout his life. Such a person is said to be *eudaimôn* or "happy."

In contrast, the Stoics characterize *phronêsis* as a kind of scientific knowledge (*episteme*), namely, of what should be done or not. Although they differ amongst

themselves about the precise relationship, the Stoics regard the other virtues as this sort of knowledge in more specific domains: justice concerns what should be done or not with regard to deserts, courage with regard to what should be endured, and moderation with regard to what should be chosen or avoided. But given the Stoics' conception of a good life as one lived in agreement with nature, knowledge of what should be done will depend on knowledge of both human nature and nature as a whole, and above all our role within the latter. *Phronêsis*, therefore, has a considerably larger scope for the Stoics than for Aristotle, and is possessed only by the Stoic ideal of the wise person.

For Epicurus, *phronêsis* has more to do with prudential reasoning. It is what enables us to assess the consequences of every choice and so calculate its overall value. It is thus crucial for leading a happy life—in fact, Epicurus regards it as even more precious than philosophy itself. In particular, he believes, it reveals that virtue and pleasure are inseparable: It is impossible to live pleasantly without living virtuously or, for that matter, to live virtuously without living pleasantly.

See also Aristotle; Dialectic; Epicureanism and the Epicurean School; Eudaimonia; Gorgias of Leontini; Plato; Protagoras of Abdera; Socrates; Stoics; Virtue and Vice; Wisdom.

Bibliography

Broadie, S. "Practical Wisdom." In *Ethics with Aristotle*. New York: Oxford University Press, 1991.

Gigon, O. "Phronesis und Sophia in der *Nicomach. Ethik des Aristoteles*." In *Mélanges de C. de Vogel*, Assen, 91–104. 1975.

Hardie, W. "Practical Wisdom." In *Aristotle's Ethical Theory*. 2nd ed. New York: Oxford University Press, 1980.

Hueffmeier, F. "Phronesis in den Schriften des Corpus Hippocraticum." *Hermes* 89 (1961): 51–84.

Kraut, R. "Function, Virtue, and Mean." In *Aristotle on the Human Good*. Princeton, NJ: Princeton University Press, 1989.

Menn, S. "Physics as a Virtue." *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 11 (1995): 1–34.

Scott Carson (2005)

QUANTIFIERS IN NATURAL LANGUAGE

Quantifiers in natural language correspond to words such as *every, some, most, few*, and many others.

THE SEMANTICS OF DETERMINERS

What is the semantics of expressions like *every* and *most*? An answer to this question emerged in the early 1980s, in work of Jon Barwise and Robin Cooper (1981), James Higginbotham and Robert May (1981), Edward L. Keenan and Jonathan Stavi (1986), Johan van Benthem (1986), Dag Westerståhl (1985), and many others.

The basic idea of how to interpret quantified expressions comes from Gottlob Frege (1879). Frege observed that the familiar quantifiers ∀ (*everything*) and ∃ (*something*) can be thought of, in Frege's terms, as second-level concepts. Let us call whatever gives the interpretation of an expression its semantic value. Assuming an extensional and set-theoretic framework, we my assign predicates sets of individuals as their semantic values. Frege's idea can then be recast as saying that the semantic values of ∀ and ∃ are sets of sets. $\exists xFx$ (*something is F*) is true if the semantic value of *F* is in the interpretation of ∃, which happens just in case the semantic value of *F* is nonempty. More generally, quantifiers have as semantic values sets of the values of predicates which result in true sentences when the quantifiers are applied.

In logic, this idea was later investigated by Andrzej Mostowski (1957) and then Per Lindström (1966). But it does not apply to natural language without an important modification. Consider:

(1) Most students attended the party.

In this, *most* does not tell us something about a single predicate. Rather, it compares the students with the people attending the party. In particular, it compares the size of the set of students with the size of the set of people attending the party.

This binary or relational character of quantification in natural language is extremely widespread (as is demonstrated by the extensive list of examples in Keenan and Stavi 1986). It is also no accident. Rather, it reflects a fundamental feature of the syntax of natural languages. Simplifying somewhat, sentences break down into combinations of noun phrases (NPs) and verb phrases (VPs). Noun phrases also break down, into combinations of determiners (DETs) and common nouns (CNs) (or more complex construction with adjectival modifiers like *small brown dog*). Quantifier expressions occupy the determiner positions in noun phrases, as in:

(2) $[_{S}\ [_{NP}\ [_{DET}$ most $]\ [_{CN}$ students $]\ ]\ [_{VP}$ attended the party$]\ ]$

(See any current syntax text for a more thorough presentation of this material, or the handbook discussions of

Bernstein [2001] and Longobardi [2001]. For some interesting cross-linguistic work, see Matthewson [2001] and the papers in Bach et al. [1995].)

Quantifier expressions, such as *every* and *most*, are determiners. Their semantic values must be relations between sets of individuals, representing the semantic values of CNs and VPs in simple syntactic configurations like (2). Using some set theory, we may give examples of the semantic values of determiners explicitly. For instance, for a universe of discourse M and sets of individuals $X, Y \subseteq M$:

(3) a. $\mathbf{every}_M\,(X,Y) \longleftrightarrow \Xi \subseteq Y$

b. $\mathbf{most}_M(X,Y) \longleftrightarrow |X \cap Y| > |X \setminus Y|$

(Here the boldface $\mathbf{every}_M$ is the semantic value of the expression *every*.) This characterization of the semantic values of determiners as relations between sets is often called the relational theory of determiner denotations.

As the semantic values of determiners are relations between sets, the semantic values of noun phrases built out of determiners (or most determiners) are interpreted as sets of sets, along Fregean lines. For instance, the semantic value of *most boys* is $\boldsymbol{most}_M$ **boys** $= \{Y \subseteq M: |\mathbf{boys} \cap Y| > |\mathbf{boys} \setminus Y|\}$. We may use the term 'quantifier' for either sort of semantic value. The latter are often called unary or simple quantifiers. Quantifiers taking more than two arguments are well documented in natural language, and have been investigated by a number of authors, including Filippo Beghelli (1994) and Edward L. Keenan and Lawrence S. Moss (1984). Quantifiers taking as inputs relations rather than sets, called polyadic quantifiers, have also been investigated, by authors including Higginbotham and May (1981), May (1989) and van Benthem (1909), though their place in natural language remains controversial. The survey by Keenan and Westerståhl (1997) is a good place to look for an introduction to these issues.

PROPERTIES OF QUANTIFIERS

The relational theory of determiner denotations has been applied to a number of issues in logic, philosophy of language, and linguistics. Many of these are discussed in the surveys by Keenan (2002), Keenan and Westerståhl (1997), and Westerståhl (1989). These applications rely on some important properties of quantifiers, of which two examples are given here.

RESTRICTED QUANTIFIERS. Quantifiers in natural language appear to be restricted quantifiers. Whereas ∀ and ∃ range over the entire universe, a quantifier like $\mathbf{most}_M$ ranges over its first input, corresponding to the CN position in an NP. *Most boys are happy* expresses $\mathbf{Most}_M$(**boys**, **happy**). Whether this holds or not depends on the properties of the **boys**, and not anything about the rest of the universe.

The mere presence of the CN argument is not enough to show that it functions as the domain of quantification. But the CN does play an important role, which is brought out by the following pattern:

(4) a. i. Every student attended the party.

ii. Every student is a student who attended the party.

b. i. Most students attended the party.

ii. Most students are students who attended the party.

In each of these, (i) and (ii) are equivalent.

The pattern we see in (4) is called conservativity:

(5) (CONS) $\mathbf{Q}_M(X,Y)$ is conservative if and only if for all $X,Y \subseteq M$, $\mathbf{Q}_M(X,Y) \longleftrightarrow \mathbf{Q}_M(X,X \cap Y)$.

Conservativity expresses the idea that the interpretation of a sentence with a quantified noun phrase only looks as far as the CN, so the CN restricts the domain of quantification.

One of the striking facts about natural languages, observed in Barwise and Cooper (1981) and Keenan and Stavi (1986), is that the semantic values of all natural language determiners satisfy CONS. This is a proposed linguistic universals: a non-trivial empirical restrictions on natural languages.

Conservativity has proved an extremely important property. The space of conservative quantifiers is much more orderly than the full range of relations between sets. This is brought out most vividly by the conservativity theorem due initially to Keenan and Stavi (1986), further investigated by van Benthem (1983, 1986) and Keenan (1993). The key insight is that the class of conservative quantifiers can be build up inductively, from a base stock of quantifiers and some closure conditions. Let M be a fixed finite universe and let $CONS_M$ be the collection of conservative quantifiers on M. We will build up a class of quantifiers $D\text{-}GEN_M$ on M as follows. $D\text{-}GEN_M$ contains $\mathbf{every}_M$ and $\mathbf{some}_M$. We also assume that each set of members of M is definable by a predicate, and that $D\text{-}GEN_M$ is closed under Boolean combination and predicate restrictions. The latter assumes that if $\mathbf{Q}_M(X,Y)$ is in $D\text{-}GEN_M$, so is $\mathbf{Q}_M(X \cap C,Y)$ for $C \subseteq M$. This amounts to closure under (intersective) adjectival restriction in an NP.

The conservativity theorem says that for each M:

(6) $CONS_M = D\text{-}GEN_M$

This tells us that the domain of natural language determiners is far more orderly than it might have appeared. Some logical properties extending CONS have been studied, by van Benthem (1983, 1986) and Westerståhl (1985, 1989). These appear to strengthen the proposed universal as well.

LOGICALITY. Quantified NPs are often described as expressions of generality. One way to articulate the relevant notion of generality is that it requires the truth of a sentence to be independent of exactly which individuals are involved in interpreting a given quantifier. This can be captured formally by the constraint of permutation invariance. A permutation π of M is a 1-1 onto mapping of M to itself, which can be thought of as a rearranging of the elements of M. The constraint of permutation invariance then says:

(7) (PERM) Let π be a permutation of M. Then $\mathbf{Q}_M(X,Y) \longleftrightarrow \mathbf{Q}_M(\pi[X],\pi[Y])$.

(Here $\pi[X] = \{\pi(x): x \in X\}$.) PERM, or some strengthening of it, is commonly assumed in the mathematical literature, and is built into the definitions of quantifier in Lindström (1966) and Mostowski (1957). The semantic values of most natural language determiners satisfy PERM. (At least, the values of most syntactically simple determiners do.) There remain some hard cases, such as possessive constructions (as well as proper names, which can be interpreted as unary quantifiers not satisfying PERM). As these may not be examples of genuine determiners, the hypothesis that all natural language quantifiers satisfy both CONS and PERM is commonplace.

SEMANTIC COMPOSITION

The relational theory of determiner denotations does not explain how quantifiers interact with the rest of syntax and semantics. The way the values of determiners combine with other semantic values provides an example of such interaction.

According to the relational theory, the semantic values of quantified NPs are sets of sets, while the values of VPs are sets. How do these combine? When we have a quantified NP in subject position, the semantics of composition is given by set membership. For a quantified NP value α:

(8) $[_S [_{NP}\ \alpha\ ]\ [_{VP}\ \beta\ ]\ ]$ is true if and only if $\beta \in \alpha$.

This simple story does not always work. Transitive verbs with quantified NPs in object position provide one sort of problem. A transitive verb will be interpreted as a relation between individuals. Now, consider an example like:

(9) a. John offended every student.

b. $[_S [_{NP}$ John $]\ [_{VP} [_V$ offended $]\ [_{NP}$ every student$]\ ]\ ]$

The value of *offended* is a relation between individuals, while the value of *every student* is a set of sets. We have no way to combine these to give us a set of individuals, which the value of the VP must be.

The theory of determiner denotations does not help solve this problem. Instead, some more apparatus is needed, either in the semantics or in the syntax. One approach is to posit underlying logical forms for sentences which are in some ways closer to the ones used in the standard formalisms of logic.

The goal is to replace the quantified NP *every student* with a variable that can occupy the argument position of a VP, that is, a variable over individuals. This variable is then bound by the quantifier. We thus want a structure that looks something like:

(10) $[\ [_{NP}$ every student$_x\ ]\ [_S$ John offended $x\ ]\ ]$

In fact, many theories (following May 1977, 1985) argue that a structure like (10) is the underlying logical form of a quantified sentence. This is a substantial empirical claim about natural language, which holds that syntactic structures like (10) provide the input to semantic interpretation. Typically, such theories also hold that a syntactic process of movement produces a syntactic structure with initial quantifiers, and variables in the argument positions those quantifiers originally occupied. (For a survey of ideas about logical form in syntactic theory, see Huang 1995.)

Providing a structure like (10) does not by itself explain the semantics of binding: It does not explain semantically how the quantified NP binds the variable in the VP. The theory of the semantic values of determiners does not explain this either. Some separate account is needed.

The semantic operation that corresponds to binding is one of forming the right set to be the input of the semantic value of the determiner. Hence, even though we think of the syntactic structure *John offended x* as sentence-like (with the variable functioning like a pronoun), its interpretation needs to wind up being $\{x$: John offended $x\}$. Once we have this, we can say the sentence is true if this set is in the semantic value of the quantified

NP *every student*. Hence, binding is carried out by the appropriate form of set abstraction (as in Barwise and Cooper 1981). Many current presentations are embedded in the framework of the typed lambda-calculus, which treat sets as functions from individuals to truth values. In such a framework (Büring 2005, Heim and Kratzer 1989), set abstraction is replaced by lambda-abstraction. Other approaches use similar syntactic structures to (10), but offer a more Tarskian account of binding (Higginbotham 1985; Larson and Segal 1995). Finally, there are approaches that avoid positing syntactic structures like we see in (10), including early work of Cooper (1983), and type shifting approaches (Hendriks 1993, Jacobson 1999, Steedman 2000, van Benthem 1991). There is also an approach that seeks to explain semantic composition via a generalized account of the semantic values of determiners (Keenan 1992).

See also Artificial and Natural Languages; Frege, Gottlob; Semantics.

Bibliography

Bach, Emmon, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee, eds. *Quantification in Natural Languages*. Dordrecht: Kluwer, 1995.

Barwise, Jon, and Robin Cooper. "Generalized Quantifiers and Natural Language." *Linguistics and Philosophy* 4 (1981): 159–219.

Beghelli, Filippo. "Structured Quantifiers." In *Dynamics, Polarity, and Quantification*, edited by Makoto Kanazawa and Christopher J. Piñón. Stanford, CA: CSLI Publications, 1994.

Bernstein, Judy B. "The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain." In *Handbook of Contemporary Syntactic Theory*, edited by Mark Baltin and Chris Collins. Oxford: Blackwell, 2001.

Büring, Daniel. *Binding Theory*. Cambridge, U.K.: Cambridge University Press, 2005.

Cooper, Robin. *Quantification and Syntactic Theory*. Dordrecht: Reidel, 1983.

Frege, Gottlob. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Nebert, 1979. Translated by Stefan Bauer-Mengelberg as *Begriffsschrift, a Formal Language, Modeled Upon That of Arithmetic, for Pure Thought* in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by Jean van Heijenoort. Cambridge, MA: Harvard University Press, 1967.

Heim, Irene, and Angelika Kratzer. *Semantics in Generative Grammar*. Oxford: Blackwell, 1989.

Hendriks, Herman. *Studied Flexibility*. Amsterdam: ILLC Publications, 1993.

Higginbotham, James. "On Semantics." *Linguistic Inquiry* 16 (1985): 547–593.

Higginbotham, James, and Robert May. "Questions, Quantifiers and Crossing." *Linguistics Review* 1 (1981): 41–79.

Huang, C.-T. James. "Logical Form." In *Government and Binding Theory and the Minimalist Program*, edited by Gert Webelhuth. Oxford: Blackwell, 1995.

Jacobson, Pauline. "Towards a Variable-Free Semantics." *Linguistics and Philosophy* 22 (1999): 117–184.

Keenan, Edward L. "Beyond the Frege Boundary." *Linguistics and Philosophy* 15 (1992): 199–221.

Keenan, Edward L. "Natural Language, Sortal Reducibility and Generalized Quantifiers." *Journal of Symbolic Logic* 58 (1993): 314–325.

Keenan, Edward L. "Some Properties of Natural Language Quantifiers: Generalized Quantifier Theory." *Linguistics and Philosophy* 25 (2002): 627–654.

Keenan, Edward L., and Lawrence S. Moss. "Generalized Quantifiers and the Expressive Power of Natural Language." In *Generalized Quantifiers in Natural Language*, edited by Johan van Benthem and Alice ter Meulen. Dordrecht: Foris, 1984.

Keenan, Edward L., and Jonathan Stavi. "A Semantic Characterization of Natural Language Determiners." *Linguistics and Philosophy* 9 (1986): 253–326. Versions of this paper were circulated in the early 1980s.

Keenan, Edward L., and Dag Westerståhl. "Generalized Quantifiers in Linguistics and Logic." In *Handbook of Logic and Language*, edited by Johan van Benthem and Alice ter Meulen. Cambridge, MA: MIT Press, 1997.

Larson, Richard, and Gabriel Segal. *Knowledge of Meaning*. Cambridge, MA: MIT Press, 1995.

Lindström, Per. "First Order Predicate Logic with Generalized Quantifiers." *Theoria* 2 (1966): 186–195.

Longobardi, Giuseppe. "The Structure of DPs: Some Principles, Parameters, and Problems." In *Handbook of Contemporary Syntactic Theory*, edited by Mark Baltin and Chris Collins. Oxford: Blackwell, 2001.

Matthewson, Lisa. "Quantification and the Nature of Crosslinguistic Variation." *Natural Language Semantics* 9 (2001): 145–189.

May, Robert. "The Grammar of Quantification." PhD diss. Massachusetts Institute of Technology, 1977.

May, Robert. "Interpreting Logical Form." *Linguistics and Philosophy* 12 (1989): 387–435.

May, Robert. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press, 1985.

Mostowski, Andrzej. "On a Generalization of Quantifiers." *Fundamenta Mathematicae* 44 (1957): 12–36.

Steedman, Mark. *The Syntactic Process*. Cambridge, MA: MIT Press, 2000.

van Benthem, Johan. "Determiners and Logic." *Linguistics and Philosophy* 6 (1983): 447–478.

van Benthem, Johan. Essays in *Logical Semantics*. Dordrecht: Reidel, 1986. Includes revised versions of van Benthem (1983, 1984).

van Benthem, Johan. *Language in Action*. Amsterdam: North-Holland, 1991.

van Benthem, Johan. "Polyadic Quantifiers." *Linguistics and Philosophy* 12 (1989): 437–464.

van Benthem, Johan. "Questions about Quantifiers." *Journal of Symbolic Logic* 49 (1984): 443–466.

van Heijenoort, Jean, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press, 1967.

Westerståhl, Dag. "Logical Constants in Quantifier Languages." *Linguistics and Philosophy* 8 (1985): 387–413.

Westerståhl, Dag. "Quantifiers in Formal and Natural Languages." In *Handbook of Philosophical Logic*. Vol. 4, edited by Dov M. Gabbay and Franz Guenthner. Dordrecht: Kluwer, 1989.

Michael Glanzberg (2005)

QUESTIONS

All too often when philosophers talk and write about sentences, they have in mind only indicative sentences, that is, sentences that are true or false and that are normally used in the performance of assertions. When interrogative sentences are mentioned at all, it is usually either in the form of a gesture toward some extension of the account of indicatives or an acknowledgment of the limitations of such an account. For example, in the final two sentences of his influential paper "Truth and Meaning" (1967), Donald Davidson remarks, "And finally, there are all the sentences that seem not to have truth values at all: the imperatives, optatives, interrogatives, and a host more. A comprehensive theory of meaning for a natural language must cope successfully with each of these problems." Nonindicatives are an embarrassment to Davidson's program of identifying meaning with truth conditions. They are equally an embarrassment for the old identification of meanings with verification conditions, as well as the newer identification of meanings with inferential roles. Nonindicatives in general, and interrogatives in particular, have neither truth conditions nor verification conditions, nor do they function naturally or principally as the premises or conclusions of inferences. Yet they are no less meaningful than indicatives. And they are certainly no less important. As Nuel Belnap has observed, following David Harrah, "[We] will not assert anything ever, nor profit from the assertions of others, without at least the traces of such interests as can be expressed by interrogatives" (1990, p. 16).

Why have philosophers felt comfortable in virtually ignoring interrogatives and the other nonindicatives? Probably because of the persistent yet rather inchoate idea that indicatives and assertion are somehow fundamental to language and meaning, and that the other forms of sentences and speech acts are secondary or derivative, perhaps even unnecessary. J. L. Austin railed against this idea in *How to Do Things with Words* (1962). Austin's pioneering work gave birth to the field of speech-act theory, which found its fullest development in the work of his student John Searle. Speech-act theory is one of the few areas in philosophy that pays due attention to uses of language other than assertion. But even here one finds a residue of the tendency to subordinate the nonassertive to the assertive. We will return to this issue a bit later on.

Outside of speech-act theory, the idea that interrogatives and the other nonindicatives are secondary survives in a number of forms. The aforementioned identification of meaning with truth conditions is a primary example. One sometimes hears philosophers defend this idea by observing that everything that can be done with language can be done with just assertions. One can ask what time it is by asserting, "I wish to know what time it is"; one can command another to lower a weapon by asserting, "You will lower your weapon"; and so on. In the opposite direction, any assertion can be performed by way of a question or an order. For any *p*, one can assert that *p* by asking "Did you know that *p*?" or by commanding "Be aware that *p*." Just as questions and orders can be performed indirectly by way of assertions, assertions can be performed indirectly by way of questions and orders.

There is also the widespread view that the shared contents of all sentences and speech acts are propositions, which are nonlinguistic representations that are true or false and are the objects of belief and assertion. For example, it is thought that, in addition to its interrogative mood, the interrogative sentence "Did Martha shoot Henry?" expresses the proposition that Martha shot Henry, the same proposition expressed by the indicative sentence "Martha shot Henry." Similarly, in asking whether Martha shot Henry, a speaker expresses the very same proposition as when asserting that Martha shot Henry. The difference between these speech acts is located in what is called their illocutionary forces, not in their shared propositional content. The study of questions thus becomes a branch of the theory of force and not part of semantics proper, which is concerned with propositions and truth conditions. This provides some excuse for the philosophical focus on the truth-conditional areas of language at the expense of the vast non-truth-conditional areas.

FREGE AND WITTGENSTEIN ON QUESTIONS

The distinction between the propositional content of a sentence or speech act and its mood or force is associated

with Gottlob Frege, for whom this distinction was a recurring theme. It is not often noticed, however, that Frege changed his mind about this distinction with regard to interrogatives. In his important paper "On Sense and Reference" (1970), Frege's view was that interrogative sentences do *not* express propositions (Frege's word for propositions was "thoughts"). Rather, interrogatives express what Frege called questions, where a question is not a proposition but something that "stands on the same level" as a proposition. In his later paper "Thoughts" (1984), he reversed himself, arguing, "An interrogative sentence and an assertoric one contain the same thought; but the assertoric sentence contains something else as well, namely assertion. The interrogative sentence contains something more too, namely a request" (p. 355). In other words, the sentences "Martha shot Henry" and "Did Martha shoot Henry?" express the same truth-conditional proposition. The difference is that the indicative sentence includes the force of assertion in the form of the indicative mood and the interrogative sentence contains the force of request in the form of the interrogative mood. (On imperatives, in contrast, Frege, in "Thoughts," did not reverse his earlier position. He held throughout that these sentences express commands, that is, contents that are like thoughts yet lack truth-values. Also, it must be noted that in "On Sense and Reference" Frege was discussing embedded questions, e.g., the "whether" clause in "Nancy knows whether Martha shot Henry," whereas in "Thoughts" he was concerned with stand-alone questions, e.g., "Did Martha shoot Henry?" If Frege held that the indirect reference of an embedded question should differ from the sense of its stand-alone counterpart, which seems unlikely, then we need not read him as having changed his mind.)

Ludwig Wittgenstein clearly rejected Frege's later account in *Philosophical Investigations*:

> Frege's idea that every assertion contains an assumption, which is the thing that is asserted, really rests on the possibility found in our language of writing every statement in the form: "It is asserted that such-and-such is the case." … We might very well also write every statement in the form of a question followed by a "Yes"; for instance: "Is it raining? Yes!" Would this show that every statement contained a question? (Sec. 22)

One of the ideas in this passage is a criticism of Frege's arbitrary identification of the contents of interrogatives with propositions. One could hold instead that the shared content of "Martha shot Henry" and "Did Martha shoot Henry?" is an interrogative content, something akin to a proposition except that it has interrogative-satisfaction conditions, that is, conditions of being properly answered, instead of truth conditions. Then one could say that the indicative contains this interrogative content along with an element of affirmation ("Yes!"). Wittgenstein's point is not that this alternative is preferable to Frege's, but rather that both accounts are arbitrary and should be rejected. In other words, indicatives and interrogatives have distinct kinds of contents. Of course, this was the view that Frege held in his earlier work "On Sense and Reference."

QUESTIONS IN SPEECH-ACT THEORY

Despite Wittgenstein's objections, many philosophers now accept Frege's later view that propositions are the shared contents of indicatives and interrogatives. This idea is the foundation of Searle's theory of speech acts. With a few exceptions (e.g., greetings), Searle analyzes speech acts on the basis of his schema $F(p)$, where "F" stands for force and "p" for propositional content. A consequence of this is that, aside from greetings and a few other speech acts, most speech acts have propositions as their contents (a circumstance that is a residue of subordinating the nonassertive to the assertive). The distinctive feature of questions is their interrogative force, which Searle takes to be a species of request. For Searle, asking a question is a request for an answer. Questions thus fall into Searle's more general category of directives, the paradigms of which are orders and commands. The defining feature of directives is that they are attempts by speakers to get hearers to do something. So on Searle's account, a question is essentially an attempt by a speaker to get the hearer to provide an answer.

Another important feature of directives is that they have what Searle calls "world-to-words" direction of fit (1979, p. 14). This means that for a directive speech act to be satisfied, the world must come to match the proposition expressed in the performance of the speech act. When I order Martha to shoot Henry, I express the proposition that Martha will shoot Henry with the force of an order. My order is satisfied just in case Martha acts to make this proposition true. This is the sense in which the order is satisfied if the world comes to fit the words used in the order. This position, however, leads to a problem when applied to questions. When I ask whether Martha shot Henry, my question is satisfied, that is, answered, just in case the hearer provides an answer. Yet the propositional content of my question is just that Martha shot Henry; it is not that the hearer will provide

an answer to the question of whether Martha shot Henry. There is no sense in which my question is satisfied when Martha shoots Henry. Another way to bring out this problem is to note that speech acts with world-to-words direction of fit require that their propositional contents describe future events or states of affairs. There is obviously no such restriction on the propositional contents of questions. The upshot of this is that questions do not fit neatly into Searle's category of directives. The fact that natural languages have a separate syntactic category of interrogative sentences, distinct from that of imperatives, further suggests that questions are not simply a variety of directives but rather constitute their own distinct category of speech acts.

THE HAMBLIN POSTULATES

The growing interdisciplinary cooperation between philosophers of language and linguists provides reason for hope that the philosophical neglect of interrogatives is coming to an end. Interrogative expressions have always occupied a central place in linguistics. For example, the behavior of so-called "wh-" words, for example, "who" and "what," provided an important source of data for early work on Chomsky's theory of transformational grammar, and the phenomenon of "wh-" movement continues to be a rich topic for linguists working on the syntax of natural language.

Interrogatives have also received a great deal of attention from linguists working in semantics. Much of this work has been guided by a set of postulates about questions and answers first laid down by the philosopher and logician C. L. Hamblin in his paper "Questions" (1958):

1. To know the meaning of a question is to know what counts as an answer to that question.

2. An answer to a question is a complete sentence or proposition.

3. The possible answers to a question form an exhaustive set of mutually exclusive possibilities.

(Hamblin's ordering and wording of these postulates is slightly different.) The first postulate is the analog for interrogatives of the idea that to know the meaning of an indicative is to know what the world would be like if it were true, that is, that to know the meaning of an indicative is to know its truth conditions. This idea is the intuitive ground for the identification of the meaning of an indicative with its truth conditions. The first Hamblin postulate plays a similar role for interrogatives. It is the intuitive motivation for the identification of the meaning of an interrogative with its answers. This first postulate is thus fundamental to semantic approaches to interrogatives.

Like the corresponding principle for indicatives, the first Hamblin postulate for interrogatives has been challenged. It seems possible to understand an interrogative without having any idea of what would count as an answer to it. The linguist Jonathan Ginzburg provides the example "What is the word for 'relaxation' in Chukotian?" (1996, p. 400). Working in the semantic framework known as situation theory, Ginzburg has developed a semantic account in which the contents of interrogatives are fine-grained structures that determine answers but are not identical with answers. This approach bears affinities to semantic accounts in which the contents of indicatives are structured propositions. Another range of counterexamples to Hamblin's first postulate derives from the work of the philosopher of science Sylvain Bromberger, who has argued that the search for answers to "why" questions for which we cannot formulate any answers is essential to the enterprise of science.

The first Hamblin postulate is also implicitly rejected by paraphrase theories of interrogatives, which analyze interrogatives by paraphrasing them into noninterrogative forms. In the theories of David Lewis and Max Cresswell, interrogatives are paraphrased as performatives. For example, "Did Martha shoot Henry?" is paraphrased as "I hereby ask you whether Martha shot Henry." A basic problem for these theories is that the interrogative reappears in the analysis in embedded form, in the example, "whether Martha shot Henry," which renders the analysis circular. In the epistemic-imperative approach of Lennart Åqvist and Jaakko Hintikka, "Did Martha shoot Henry?" is analyzed as the imperative "Bring it about that I know whether Martha shot Henry." The remaining embedded "whether" clause is then eliminated in terms of "that" clauses. "I know whether *p*," for example, is analyzed as a conjunction of conditionals: "If *p*, then I know that *p*, and if not *p*, then I know that not *p*." This account has some plausibility in this case, but as Lauri Karttunen has pointed out, it falls apart when applied to other uses of "whether" clauses. "I wonder whether *p*" is clearly not synonymous with the possibly ungrammatical "If *p*, then I wonder that *p*, and if not *p*, then I wonder that not *p*." And it is not clear even how to apply this account to a sentence like "Martha's mental health depends on whether she takes her prescriptions."

The second and third Hamblin postulates concern the nature of answers. These two postulates combine to form a conception of answers that differs from what can

count as an answer in ordinary discourse. For example, the second postulate is in apparent conflict with the fact that one can often answer a question with something less than a complete sentence. For example, the proper name "Alexander Hamilton" seems like a perfectly good answer to the question "Who was the first U.S. Secretary of the Treasury?" The point of the second postulate is that, despite appearances, answers are always complete sentences or propositions, in this case, the sentence "Alexander Hamilton was the first U.S. Secretary of the Treasury" or the proposition expressed by this sentence. This postulate is motivated by the idea that a correct answer must be true, and being true is a property of sentences or propositions. Furthermore, answers always convey information, and information comes in sentences or propositions.

Despite these considerations, the second Hamblin postulate has not been universally accepted. So-called categorial theories, such as that of Roland Hausser, take seriously the surface grammatical forms of answers. On these approaches, answers can be of various categories, for example, names, common nouns, sentences, set designations, and predicates, which denote respectively individuals, objects, propositions, sets, and properties.

The third Hamblin postulate requires first that the set of answers to an interrogative be exhaustive. This is related to the fact that many interrogatives carry presuppositions. To use Hamblin's example, consider the question "In which continent is Honolulu?" (1958, p. 163). This question falsely presupposes that Honolulu is in a continent. According to one position, for the set of answers to this question to be exhaustive, it must include an answer that denies the presupposition, that is, "Honolulu is in no continent." Alternatively, one might hold that the presuppositions of a question restrict the range of possibilities to just those in which the presuppositions hold. A set of answers would then be exhaustive if it exhausts this restricted range of possibilities. On this alternative, the denial of the presupposition of a question is not an answer but rather a rejection of the question.

The third Hamblin postulate also requires that answers are mutually exclusive. This is intended to capture the idea that genuine answers are complete, in the following sense. Consider the question "Who ran the marathon?" where the candidate runners are Martha, Henry, George, and Nancy. A complete answer will indicate both who ran and who did not. For example, the proposition that only Martha and Henry ran and no one else ran is complete, whereas the proposition that Martha and Henry ran is not complete, since it leaves unspecified whether George or Nancy ran. A consequence of this is that the proposition that Martha and Henry ran is at best a *partial* answer. The fact that answers can be merely partial is what motivates the requirement that answers be mutually exclusive. Allowing partial answers requires a contrasting criterion of completeness, which is provided by the notion that answers be mutually exclusive. (Incidentally, the above example illustrates how "wh-" words are context-sensitive, as are quantifier expressions. Intuitively, a speaker who asks "Who ran the marathon?" is not asking about everyone who has ever lived but rather about some contextually determined set of candidate runners. Parallel remarks apply to someone who asserts "Everyone ran the marathon." In each case, a range of values for "who" and "everyone" is determined by features of the context of utterance. This is one of many similarities between "wh-" words and quantifiers.)

If answers are mutually exclusive, then there cannot be more than one complete and true answer to a question. This runs into problems with so-called mention-some questions. Suppose that Martha, who is new in town, asks Henry "Where can I buy an Italian newspaper?" (This example is due to Jeroen Groenendijk and Martin Stokhof.) On the most natural reading, Martha is only asking Henry to mention some place where she can buy an Italian newspaper. If so, Henry has available any number of complete and true answers, for example, "At the train station," or "At the bookstore downtown." Another sort of problem case, raised by Belnap, consists in choice questions, for example, "What are two cities that host marathons?" Intuitively, a complete answer mentions two cities that host marathons, and the choice of which two to mention is left up to the hearer. Thus, many complete and true answers are available, such as "Boston and New York host marathons," "Chicago and Los Angeles host marathons," and so on.

THREE SEMANTIC APPROACHES TO INTERROGATIVES

This section sketches three prominent approaches to the semantics of interrogatives, all of which are set in the framework of Montague semantics, also variously known as intensional semantics, model-theoretic semantics, or possible-worlds semantics. In this framework, expressions are assigned both intensions and extensions. Intensions are functions from possible worlds to entities of various kinds. The extension of an expression at a possible world is the value of its intension with respect to that world. For example, the intension of a complete indicative sentence is a function from possible worlds to truth-values. The intensions of indicatives essentially divide the

set of possible worlds into two subsets: those possible worlds in which the indicative is true and those in which it is false. The proposition expressed by an indicative is normally identified either with its intension or, more simply, with the set of worlds in which the intension has the value true. This identification of propositions with sentence intensions or with sets of possible worlds is a notoriously problematic feature of the possible-worlds framework. It has the consequence that all necessarily true sentences express the same proposition. As we will see later on, a related problem arises for interrogatives.

On C. L. Hamblin's approach in his "Questions in Montague English" (1973), the intension of an interrogative is a function from possible worlds to sets of answers, where answers are propositions. The extension of an interrogative at a possible world is thus a set of propositions. This set is determined compositionally from the parts of the interrogative. For Hamblin, the extension of "who" at a possible world is a set of individuals. For example, suppose that the extension of "who" in a possible world *w* is the set {Martha, Henry, George, Nancy}. The extension of "Who runs?" in *w* is then the set of propositions {⟨Martha runs⟩, ⟨Henry runs⟩, ⟨George runs⟩, ⟨Nancy runs⟩}. (Remember that each of these propositions is itself an indicative sentence intension or a set of possible worlds.) Hamblin is aware that this approach is a departure from his own third postulate, since there is no requirement here that sets of answers be exhaustive nor that answers themselves be mutually exclusive. The extension of the yes/no interrogative "Is it the case that *p*?" in a world *w* is the set consisting of the proposition that *p* and its negation. For example, the extension of "Does Martha run?" in *w* is {⟨Martha runs⟩, ⟨Martha does not run⟩}.

Perhaps the best-known approach to interrogatives is due to Lauri Karttunen. Karttunen's account is similar to Hamblin's except that Karttunen requires that each member of the extension of an interrogative be true. In other words, on Karttunen's approach, the intension of an interrogative is a function from possible worlds to sets of true answers. Suppose that in *w* only Martha and Henry run. For Karttunen, the extension of "Who runs?" in *w* is the set {⟨Martha runs⟩, ⟨Henry runs⟩}. Similarly, the extension of "Does Martha run?" is the singleton set {⟨Martha runs⟩}. Karttunen argues that the advantage of his approach over Hamblin's is that' his approach provides a simpler account of the semantics of question-embedding verbs like "knows," as in sentences such as "Nancy knows who runs." It is widely assumed that the content of the embedded interrogative "who runs" is identical with the content of its stand-alone counterpart "Who runs?" Very roughly, Karttunen's idea is that "Nancy knows who runs" is true in *w* just in case in *w* Nancy knows each of the propositions in the extension of "who runs." The advantage of Karttunen's approach is that this extension includes only true propositions, which accords with the fact that one cannot know something false.

A third prominent approach to interrogatives is due to Jeroen Groenendijk and Martin Stokhof (1997). Unlike Hamblin and Karttunen, Groenendijk and Stokhof accept the third Hamblin postulate. On their account, the sets of answers to interrogatives are exhaustive, and each answer is mutually exclusive. A consequence of this position is that, on their view, the intension of an interrogative is a function from possible worlds to single propositions, that is, the unique, complete answers in each world. Suppose that in *w* only Martha and Henry run. Then the extension of "Who runs?" in *w* is the single proposition that Martha runs and Henry runs and no one else runs. Groenendijk and Stokhof's approach is sometimes called a partition theory. This is because on their view the intension of an interrogative partitions the set of possible worlds into jointly exhaustive, nonoverlapping subsets, one for each possible complete answer. One advantage of this model is that it captures the apparent fact that if Nancy knows who runs, she knows both who runs and who does not run. For example, if George does not run, and Nancy does not know it, then it seems that Nancy does not know who runs, even if she knows that Martha and Henry run. For Groenendijk and Stokhof, this is captured by the fact that "Nancy knows who runs" is true just in case Nancy knows the complete answer to the question "Who runs?" For Karttunen, all that is required for the truth of "Nancy knows who runs" is that Nancy knows all the true propositions of the form ⟨X runs⟩. She need not know any of the true propositions of the form ⟨X does not run⟩.

A feature shared by all three approaches is that they assign contents to interrogatives that are distinct from those for indicatives. The content of an expression is its intension. This means that for Hamblin, Karttunen, and Groenendijk and Stokhof, the contents of interrogatives are *not* propositions. Rather, they are functions from possible worlds to sets of propositions (Hamblin, Karttunen) or single propositions (Groenendijk and Stokhof). These functions can be thought of as *properties* of propositions. Thus, for Hamblin, the content of an interrogative is the property of being an answer to that interrogative (where answers can be incomplete), for Karttunen it is the prop-

erty of being a true (possibly incomplete) answer, and for Groenendijk and Stokhof it is the property of being a complete and true answer.

As noted earlier, the framework of Montague semantics faces difficulties arising from its identification of propositions with sets of possible worlds. Because they are set within this framework, all three of these accounts of interrogatives face similar problems. For example, the contents of "Does 5 + 7 = 12?" and "Is first-order logic undecidable?" turn out to be identical on all three accounts. Philosophers have responded to the problems for possible-worlds accounts of propositions by searching for more fine-grained entities, such as structured propositions, to serve as the contents of indicatives. Whether or not similarly fine-grained interrogative contents can be found is a question that is currently being explored.

See also Aristotle; Carnap, Rudolf; Explanation; Mackie, John Leslie; Non-Truth-Conditional Meaning; Presupposition; Prior, Arthur Norman; Propositions; Schlick, Moritz; Strawson, Peter Frederick; Why.

Bibliography

Åqvist, Lennart. *A New Approach to the Logical Theory of Interrogatives*. Uppsala, Sweden: Filosofiska Föreningen, 1965.

Austin, J. L. *How to Do Things with Words*. Cambridge, MA: Harvard University Press, 1962.

Belnap, Nuel. "Declaratives Are Not Enough." *Philosophical Studies* 59 (1990): 1–30.

Belnap, Nuel. "Questions and Answers in Montague Grammar." In *Processes, Beliefs, and Questions*, edited by Stanley Peters and Esa Saarinen, 165–198. Dordrecht, Netherlands: D. Reidel, 1982.

Bromberger, Sylvain. *On What We Know We Don't Know*. Chicago: University of Chicago Press, 1992.

Cresswell, Max. *Logics and Languages*. London: Methuen, 1973.

Davidson, Donald. "Truth and Meaning." *Synthese* 17 (1967): 304–323.

Ginzburg, Jonathan. "Interrogatives: Questions, Facts, and Dialogue." In *The Handbook of Contemporary Semantic Theory*, edited by Shalom Lappin, 385–422. Oxford, U.K.: Blackwell, 1996.

Groenendijk, Jeroen, and Martin Stokhof. "Questions." In *Handbook of Logic and Language*, edited by Johan van Benthem and Alice ter Meulen, 1055–1124. Cambridge, MA: MIT Press, 1997.

Hamblin, Charles L. "Questions." *Australasian Journal of Philosophy* 36 (1958): 159–168.

Hamblin, Charles L. "Questions in Montague English." *Foundations of Language* 10 (1973): 41–53.

Harrah, David. *Communication: A Logical Model*. Cambridge, MA: MIT Press, 1963.

Hausser, Roland. "The Syntax and Semantics of English Mood." In *Questions and Answers*, edited by Ferenc Kiefer, 97–158. Dordrecht, Netherlands: D. Reidel, 1983.

Hintikka, Jaakko. "The Semantics of Questions and the Questions of Semantics." *Acta Philosophica Fennica* 28 (4) (1976): 1–200.

Frege, Gottlob. "On Sense and Reference." In *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black, 56–78. Oxford, U.K.: Blackwell, 1970.

Frege, Gottlob. "Thoughts." In his *Collected Papers on Mathematics, Logic, and Philosophy*, edited by Brian McGuinness, 351–372. Oxford, U.K.: Blackwell, 1984.

Karttunen, Lauri. "Syntax and Semantics of Questions." *Linguistics and Philosophy* 1 (1977): 3–44.

Lewis, David. "General Semantics." *Synthese* 22 (1970): 18–67.

Searle, John. *Speech Acts*. Cambridge, U.K.: Cambridge University Press, 1969.

Searle, John. *Expression and Meaning*. Cambridge, U.K.: Cambridge University Press, 1979.

Wittgenstein, Ludwig. *Philosophical Investigations*. New York: Macmillan, 1953.

Peter W. Hanks (2005)

REDUCTIONISM IN THE PHILOSOPHY OF MIND

Reduction can be understood in a loose or in a strict sense. In the loose sense, entities (or expressions) of a given type are reduced if they refer to "nothing over and above" other entities (expressions) that we consider well established. This is consistent with the conclusion that the reduced entities are among the posits of a mistaken world view and thus have no place in our ontology, and it is also consistent with the conclusion that the reduced entities are conserved among other accepted, better established or understood entities. In the first case we have *elimination*, and proposing this for entities of a given kind makes us eliminativists about those entities. In the second case we have *reduction in the strict sense*, and proposing this for a given kind makes us reductionists (sometimes called "conservative" or "retentive" reductionists). Reductionist projects can also be semantic or theoretical. A semantic reduction attempts to show that items belonging to a certain class of expressions are semantically equivalent to—that is, definable in terms of—another class of expressions. A theoretical reduction aims at showing that a given scientific theory can be fully subsumed under (that is, derivable from) another more basic theory.

TYPES OF MIND-BODY REDUCTIONISM

In the philosophy of mind, reductionist projects come in all formats. A reductionist effort will typically be directed against the claim that the mental has some real, independent status. But this claim has a range of versions that go from the mind being a nonphysical/biological object, to mental properties constituting a level of *sui generis* properties of organisms that is in some sense autonomous vis-à-vis the physical/biological properties, to mental expressions possessing meanings that cannot be accounted for in purely behavioral/physical terminology.

The substance dualist assertion influential until the twentieth century—that the (human) mind is an immaterial object or substance—has faced widespread philosophical criticism of an eliminativist type: "Immaterial mind" or "soul," like "élan vital," "elf," or "chupacabras," are ghostly expressions that come from mistaken frameworks or conceptions and do not refer to anything. An influential formulation of this view is Gilbert Ryle's claim that the immaterial entity posited by substance dualism is the result of a category mistake in which we reify our mental activities by placing a ghost in charge of our body. Another major reason for the eliminativist consensus about nonmaterial substances is the inability of a nonphysical substance to causally interact with the physical world, because of conservation of energy considerations and because of the difficulty of making sense of bridging mechanisms between the two ontologically diverse realms. Absent causal interaction, the argument goes, postulating souls seems pointless if not absurd.

Eliminating mental substances, however, does not directly lead to a reductive view of the mental. In the twentieth century substance materialism or physicalism has been the orthodoxy in tune with modern science, but "the reducibility of mind" has remained as a philosophical issue of first importance. It is only that the focus of the debate has now shifted to the ontological or semantic autonomy of mental *properties* or *predicates*. The first systematic attempt to fully reduce the mental to the physical comes from logical behaviorism, a position championed by Rudolf Carnap, Carl Hempel, and Gilbert Ryle in the 1930s and 1940s. The view has doctrinal connections to methodological behaviorism, the dominant methodology of psychology in the first half of the twentieth century.

Based on the logical positivist's verification criterion according to which the content of an expression is just the expressions' verification conditions and on the assumption that these conditions have to be publicly observable, logical behaviorism argues that in order for sentences including mental expressions to be meaningful they have to be translatable without loss of content into sentences including just behavioral and other physical expressions. This implies that mental expressions should be defined in terms of behavioral and other physical expressions. Following the model of definitions of dispositional properties in the natural sciences, these definitions standardly include conditional sentences showing dispositions to behave under given environmental circumstances including stimuli. So logical behaviorism is a form of semantic reduction of the mental.

Logical behaviorism has been largely abandoned for several reasons, one of them being its inability to meet the positivist standards in its own reductionist strategy. Most mental terms cannot be associated with a single behavioral disposition; there is no single behavioral manifestation of, say, "believing in God" or "loving one's country." If mental terms denote behavioral dispositions, these dispositions must be "multitracked," and this would make behavioral definitions of mental terms enormously complex. This makes the behaviorist project of defining mental terms a highly dubious project.

Moreover, it has been convincingly argued that even in simple cases a purely behavioral definition just is not possible—unless one uses some mentalistic term in the definition, which of course undermines the behaviorist enterprise. The fall of behaviorism as the accepted reductive view led to a different reductionist approach. In the 1950s U. T. Place, J. C. C. Smart, and Herbert Feigl proposed the mind-body identity theory, a simple and appealing view in line with the surge of neural research. According to the view, while there is no meaning equivalence between mental and neural terminology (thus no semantic reduction) mental states are just states of the brain or the nervous system. The claim is one of numerical identity between types of states or properties and as such it involves ontological reduction in the strict sense.

A main line of argument for the identity theory is based on ontological simplicity, a standard strategy for ontological reduction. Once we have observed a pervasive set of systematic correlations between mental occurrences and neural events, the argument goes, we should conclude that the mental and the neural are identical. For while mind-brain correlations are compatible with a range of views about the mind, simplicity dictates that we should not multiply entities that are not going to enhance our explanatory power. The view is also supported by considerations of theoretical reduction. The history of science offers countless cases of predicates of everyday frameworks being reduced to predicates of explanatorily

richer scientific frameworks (a standard example is the reduction of temperature [of gases] to molecular kinetic energy). Given the advances in the neurosciences we have good reason to think a neural reduction of mentality is going to be one more item in a chain of successful theoretical reductions. This theoretical reduction would proceed by establishing "bridge laws" between mental and neural predicates and then reducing all generalizations involving the mental to the more encompassing laws of neuroscience.

Of the many attacks raised against the identity theory, two have aimed at the core of its reductive stance. Donald Davidson has argued against type-identification by claiming that there cannot be laws connecting the mental and the physical (this is called anomalism of the mental, an essential part of Davidson's nonreductive view discussed below). Mental states, in particular intentional states such as beliefs and desires, are governed by principles of rationality without which attribution of mentality would be impossible. Laws connecting the physical and the mental would constrain the mental by the principles of physical theory and thereby undermine its own peculiar rationality constraints.

Another highly influential argument against the identity theory is the "multiple realization" argument initially developed by Hilary Putnam. The identity theory requires a single physical property be the reduction base for each mental state. But surely the same mental state can occur in organisms with diverse neurophysiological structures. Nonhuman animals can be in pain and we can conceive of noncarbon based species and perhaps even artificial creatures being in pain. Mental states, Putnam argues, can be implemented or "realized" in widely diverse physical/chemical structures and so there is no unifying reduction base or structure for them. (This multiple realization objection is also at the core of the nonreductive functionalist approach discussed below.)

An alternative, eliminativist stance was defended in the 1960s by Richard Rorty and Paul Feyerabend and has as more recent versions the views of Patricia Churchland, Paul Churchland, and Stephen Stich. Learning from the failure of the identity theory to establish type-type identities between mental and neuro-chemical properties, eliminativism claims that the mental expressions used in our everyday psychological talk have no more reality or significance than "phlogiston" and "caloric fluids," terms of superseded and discarded scientific theories. It is highly unlikely that these concepts of vernacular psychology could be sharpened into concepts that will be useful to the sciences and do not correspond to the concepts of the sciences (neuroscience or cognitive science) that have the task of explaining human behavior. This radical view proposes to eliminate mental terminology for the purposes of scientific theorizing and can go as far as predicting that a full replacement is possible even for everyday purposes. The analogy with concepts in the history of science that were found to be fully misguided and therefore replaced plays an important role in the argumentation in favor of eliminativism. This view has been found by most philosophers to be unacceptably extreme since it means that an essential component of our conceptual framework has to be given up. Also, some have argued that the view is incoherent since the view cannot be expressed without the very (mental) concepts it rejects (since in the very act of affirming their view, the eliminativist is expressing a *belief*, something that, according to their view, does not exist).

TYPES OF MIND-BODY ANTI-REDUCTIONISMS AND THE REDUCTIVISTS' REACTIONS

Starting in the late 1960s, the problems plaguing reductive views let to the establishment of nonreductive physicalism as a reigning orthodoxy in the philosophy of mind. Its two most salient versions are anomalous monism and functionalism. Functionalism in fact has been the predominant view into the twenty-first century.

Davidson's anomalous monism is a physicalist view that eschews reduction. From the principles that every singular causal relation needs to be backed by strict laws (nomological character of causation) and that there are no "strict" laws about mental properties (mental anomalism), together with the assumption that at least some mental events causally interact with physical events, Davidson concludes that mental events must be identical with physical events. According to Davidson, this provides causal efficacy to mental events, even though there are no strict psychological laws governing them, and it also leads to a nonreductive view of the mental because there are no laws connecting mental properties with physical properties.

Many critics have argued that Davidson's view leaves the mental with no causal role to play. Davidson is entitled to affirm that a mental event causes a physical event (by being identical to a physical—probably neural—event). Now, an event instantiates a law—required for causation—in virtue of some of its properties, or, in other words, in virtue of falling under some event-type. Since anomalism entails that there are no laws involving mental properties or event-types, it is the physical (neural)

properties of the cause event that are efficacious in the production of the effect. The fact that the cause event falls under a mental type, or the fact that the event has mental properties, is completely irrelevant for the event's causing the effect. Thus, critics conclude, Davidson's anomalous monism renders the mental epiphenomenal, making it an easy target for elimination.

The functionalist view of the mental defended by Putnam and Jerry Fodor, among others, starts with the anti-reductivist stance included in the multiple realization argument. Its positive view includes the claim that mental properties are functional properties, rather than physical/neural properties as claimed by the identity theory. On the functionalist view, for something to have a mental property M is for it to instantiate some physical property P that has the right causal connections with inputs, behavioral outputs and other mental states. Thus, a mental property is a second-order property of having a (first-order) property that fulfills a certain specified causal specification. A first-order property meeting the causal specification is called a "realizer" or "realizing property" of the second-order functional property. For any given mental property there will likely be indefinitely many realizing properties satisfying its causal specification.

The reductionist can challenge the functionalist by suggesting that the mental property be identified with the disjunction of realizers. Settling this challenge would require a metaphysical discussion on the nature of disjunctive properties. A more powerful challenge raised by Kim is the claim that since having the functional mental property implies having one of its realizing properties and since the casual powers of the instance of a functional property must be considered to be inherited from the causal powers of the realizing property, mental properties have no autonomous causal powers and so are epiphenomenal. To the reply that it is the mental kind and not the instance that has its own causal powers Kim answers that the sheer heterogeneity and diversity of the realizers of a functionally conceived mental property deprives the property of the kind of causal-nomological unity required for nomological and causally efficacious properties.

All versions of nonreductive physicalism (including anomalous monism and traditional functionalism) are targets of the exclusion argument initially put forth by Norman Malcolm and developed by Jaegwon Kim. Physicalists, even those in the nonreductive camp, accept the primacy of the physical not only in terms of substance monism but also in terms of physical properties being primary vis-à-vis mental properties. This commitment includes, according to Kim, accepting the causal closure of the physical and accepting a strong sense of dependence of the mental upon the physical. Thus, every physical event, including human behavior, has to have a complete physical cause. The mental event that is supposed to be the cause of behavior is preempted of its causal role by the physical state upon which it depends and which is the required physical cause of behavior. The upshot is that we cannot attribute a causal role to the mental unless it is identified with the physical, transforming nonreducible mental properties into epiphenomena. And epiphenomena, Kim thinks, should be cut from our ontology because they serve no purpose.

A common theme across several discussions so far has revolved around whether the mental, on one view or another, has autonomous causal powers. It is not obvious whether causal reduction or elimination implies full ontological reduction or elimination, that is, whether putative entities that are causally inefficacious or epiphenomenal can still be *bona fide* entities. To achieve full reduction we need the extra assumption that independent causal powers are necessary for the very reality of an entity. This view has been explicitly defended by Kim and Sidney Shoemaker, among others, and is largely the orthodox view. A negative answer (supported for instance by Elliott Sober and Marcelo Sabatés) makes room for epiphenomenalism as a nonreductive option about the mind.

In the first decade of the twenty-first century reductionism has gained some momentum. Kim has developed an influential functionalist view of reduction with ties to the version of functionalism defended by David Lewis in the 1970s. Kim's position, in agreement with his criticism of traditional functionalism á la Putnam, claims that "functionalizing" a property provides a form of theoretical reduction that does not require bridge laws and fully explains its reductive relationship on the base property. The view implies that on account of its multiple diverse realizability, no mental property has sufficient causal/nomological homogeneity to count as a genuine, projectible property useful in science. Instead, it proposes that we eschew talk of mental *properties* in favor of mental *predicates or concepts* that at most we get a pragmatically useful mental *predicate*. In making this move, functional reductionism appears to turn itself into a form of eliminativism with regard to mental properties.

See also Alexander, Samuel; Anomalous Monism; Broad, Charlie Dunbar; Davidson, Donald; Eliminative Materialism, Eliminativism; Emergence; Frege, Gottlob; Knowledge Argument; Logic, History of; Metaphysics, History of; Mind-Body Problem; Moral Realism; Mor-

gan, C. Lloyd; Multiple Realizability; Phenomenalism; Philosophy of Mind; Philosophy of Science, History of; Philosophy of Science, Problems of; Physicalism; Properties; Qualia; Reduction; Russell, Bertrand Arthur William; Set Theory; Supervenience.

Bibliography

Block, Ned. "Antireductionism Slaps Back." *Philosophical Perspectives* 11 (1997): 107–132.

Carnap, Rudolf. "Psychology in Physical Language" (1931). Translated in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

Churchland, Patricia. *Neurophilosophy*. Cambridge, MA: MIT Press, 1986.

Churchland, Paul. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78 (1981).

Davidson, Donald. "Mental Events." In *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.

Fodor, Jerry. "Special Sciences, or the Disunity of Science as a Working Hypothesis." *Synthese* 28 (1974): 97–115.

Fodor, Jerry. "Special Sciences: Still Autonomous after All These Years." *Philosophical Perspectives* 11 (1997): 149–163.

Hempel, Carl G. "The Logical Analysis of Psychology" (1935). In *Readings in Philosophy of Psychology*, vol. 1, edited by Ned Block, 14–23. Cambridge, MA: Harvard University Press, 1980.

Jackson, F. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982).

Kim, Jaegwon. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

Kim, Jaegwon. *Supervenience and Mind*. Cambridge, U.K. and New York: Cambridge University Press, 1993.

Lewis, David. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (1972): 249–258.

Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47 (1956): 44–50.

Putnam, Hilary. "The Nature of Mental States." In *Mind, Language and Reality*. Cambridge, U.K.: Cambridge University Press, 1978.

Rorty, Richard. "Mind-Body Identity, Privacy and Categories." *Review of Metaphysics* 19 (1965).

Ryle, Gilbert. *The Concept of Mind*. New York: Harper and Row, 1949.

Sabatés, Marcelo. "Being without Doing." *Topoi* 22 (2003).

Shoemaker, Sidney. "Causality and Properties." In *Identity, Cause and Mind*, edited by S. Shoemaker. Cambridge, U.K.: Cambridge University Press, 1984.

Smart, J.C.C. "Sensations and Brain Processes." *Philosophical Review* 68 (1959): 141–156.

Sober, Elliot. "A Plea for Pseudo-Processes." *Pacific Philosophical Quarterly* 66 (1985).

Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9 (1984).

Stich, Stephen. *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press, 1983.

Marcelo H. Sabatés (2005)

SOPHIA

The Greek word *sophia* properly refers to cleverness or skill in handicraft and the productive arts, such as carpentry, music, singing, poetry, chariot driving, medicine, and even divination. In short it tends to pick out the sort of excellence in a particular domain that derives from experience and expertise. In early applications of the term to "wise men," for example the Seven Sages, the term referred primarily to the sorts of skills that would make for expertise in matters of common life and so was virtually synonymous with practical wisdom or prudence (*phronêsis*). By the late fifth century BCE, however, the term was coming to have a more specialized meaning having to do with technical skill and the expertise derived from expert training and experience; that is, it encompassed both a knowledge base and an intimate familiarity with the applications of that knowledge base. The Sophists in particular claimed to have this sort of knowledgeable expertise in many different areas, from medicine to mathematics, oratory, and political science. Indeed, the name "sophistēs" simply means someone who makes a profession of the practice and teaching of such sorts of knowledge.

In Plato, "sophia" clearly has more philosophical connotations. Already in the early, Socratic dialogues we find an attempt to draw a distinction between the kinds of "expertise" that Sophists had and the sort of genuine reflective wisdom modeled by Socrates. For Plato, the former is clearly mere logical chicanery used to generate linguistic puzzles for the purpose of winning debates (see, for example, Socrates' line of reasoning in the *Gorgias* 464b–465e). By the time Plato wrote the *Theaetetus*, he had clearly settled on an antisophistic conception of knowledge and expertise that takes the life and methodology of Socrates as its model, though even in that arguably late dialogue there is no clear line of demarcation drawn between *sophia* and *epistēmē* (knowledge). Since, for Plato, all knowledge, whether of mathematical objects or normative concepts such as the virtues, involves cognitive grasp of purely formal entities, there is less demand in his epistemology for a clear and concise differentiation between the two types of mental states and their proper objects.

Aristotle, by contrast, drew rather sharp distinctions not only between *epistēmē* and *sophia*, but also among those rational faculties and *phronêsis* (practical wisdom), *technē* (art, skill), and *nous* (intelligence, understanding). Yet the relation of *sophia* to the other rational faculties is somewhat specialized. In the *Nicomachean Ethics* (VI.7,

1141a9–b3), Aristotle began by noting the traditional use of the word "sophia" to denote those who have mastered their craft (*technē*) in a most exacting way, but added that it was also used to denote those who are "wise in general and not in one department," and he gave this as his reason for thinking that *sophia* is the "most perfect of the modes of knowledge." Thus *sophia* is associated with both *technē* and *epistēmē*, but it marks off a superlative kind of knowledge in which the knower not only fully understands the consequences of the principles of his craft but also fully understands the natures of the principles themselves. There is thus a sense in which *sophia* encompasses both the necessary truths that follow from demonstrations (the domain of *epistēmē*) and the necessary truths that are the first principles of the demonstrative sciences (the domain of *nous*). In the *Metaphysics* (981b28), this controlling wisdom is said to have the causes and first principles of all the other intellectual faculties as its proper objects, and so it is the highest form of wisdom.

The Stoics likewise took *sophia* as the perfection of human understanding (Seneca, *Epistulae* 89.4), and as consisting in a fully comprehensive and systematic grasp of the rational order in the universe. They characterized *sophia* as "knowledge of the divine and the human," with some adding "and their causes" (von Arnim, 2.35; Seneca, *Epistulae* 89.5). They also regarded this understanding as the crucial underpinning for the goal of leading a moral life and hence considered it a virtue.

In later antiquity, *sophia* held an even more elevated place. In the early Christian theologies of Philo Judaeus and Origen, it is associated with *logos* (word) and thus with the daughter or son of God, respectively. A central feature of the various Gnostic movements was the personification of *sophia* as a salvation figure. In some systems there were two sorts of *sophia*, Wisdom from above and Wisdom from below, representing the female, or noumenal, world and the male, or material, world, respectively. This dualism of *sophia* came in varying degrees. In Marcionism, a heretical doctrine of the second through fifth centuries and the most dualistic system of all, salvation consisted of accepting the wisdom that comes from the Good God and rejecting whatever comes from the Demiurge.

See also Aristotle; Gnosticism; Origen; Philo Judaeus; Phronêsis; Plato; Sophists; Stoicism.

Bibliography

Aristotle. *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Gigon, Olof. "Phronesis und Sophia in der *Nikomachischen Ethik* des Aristoteles." In *Kephalaion: Studies in Greek Philosophy and Its Continuation Offered to Professor C. J. de Vogel*, edited by J. Mansfeld and L. M. de Rijk, 91–104. Assen, Netherlands: Van Gorcum, 1975.

Gladigow, Burkhard. *Sophia und Kosmos: Untersuchungen zur Frühgeschichte von Sophos und Sophia*. Hildesheim, Germany: Olms, 1964.

Hankinson, R. J. "Natural Criteria and the Transparency of Judgment: Philo, Antiochus, and Galen on Epistemological Justification." In *Assent and Argument: Studies in Cicero's Academic Books*, edited by Brad Inwood and Jaap Mansfeld, 161–216. Leiden, Netherlands: Brill, 1997.

Menn, Stephen. "Physics as a Virtue." *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 11 (1997): 1–34.

Motte, André. "Cicerón et Aristote: A propos de la distinction entre la sophia et la phronèsis." In *Aristotelica: Mélanges offerts à Marcel de Corte*, edited by André Motte and Christian Rutten, 263–303. Brussels: Éditions Ousia, 1985.

Plato. *The Collected Dialogues of Plato*. Princeton, NJ: Princeton University Press, 1963.

Seneca. *Letters from a Stoic. Epistulae morales ad Lucilium [by] Seneca*. Selected and translated by Robin Campbell. Harmondsworth: Penguin, 1969.

Von Arnim, Hans. *Stoicorum veterum fragmenta*. Leipzig, Germany: B. G. Teubneri, 1903–1924.

Woodruff, Paul. "Plato's Early Theory of Knowledge." In *Epistemology*, edited by Stephen Everson, 60–84. Cambridge, U.K.: Cambridge University Press, 1990.

Scott Carson (2005)

SÔPHROSUNÊ

Sôphrosunê is the Greek virtue of self-control, or temperance, a virtue that Aristotle says lies between self-indulgence (*akolasia*) on the one hand and insensibility (*anaisthêsia*) on the other. In its earliest uses (Homer) the word means "soundness of mind," "prudence," "discretion," and is related to the verb *sôphronein*, combining *sôs*, safe, and *phronein*, to think, a verb related to *phrên*, an archaism for mind (literally, "midriff," "heart," "the seat of thought," according to the Greeks).

Although Plato dedicated an entire dialogue (*Charmides*) to a discussion of the meaning of *sôphrosunê*, the notion of self-mastery is central to his ethical theory and he invokes it in many contexts, ranging from the *Gorgias* to the *Republic* to the *Laws*. Plato's central claim is that self-mastery is more than the mere abstention from certain forms of physical pleasure—that was the popular and sophistic characterization of the virtue—he "exalts" it (*semnunôn, Laws* 710a5) by equating it with *phronêsis*, practical wisdom. Already in the so-called "early" or "Socratic" dialogues (among which the *Charmides* may be counted) Plato had spoken not only of self-control but of

all the virtues as reducible, in some way, to knowledge of one kind or another. Like the other "early" dialogues, the *Charmides* ends in *aporia*, puzzlement, about what *sôphrosunê* "really" is, but the suggestion is quite clear that it has to do with knowledge of what is the objectively best way for one to live. When, at *Gorgias* 491e, Callicles scorns self-control as a mere convention valued only by stupid, foolish people (*êlithious*), Socrates mounts an argument to show that those who cannot master their own desires and inclinations cannot master anything, a theme he takes up again in the *Republic*.

Aristotle regards temperance as moderation regarding pleasures and pains, and he loosely associates this virtue with courage as the two virtues of the non-rational (*alogon*) part of the soul (*Nicomachean* Ethics II.7, 1107b5–8; cf. III.10–12 1117b23–1119b10). Aristotle notes that temperance applies more to physical pleasures and pains than mental, and rather more to pleasure than to pain. On Aristotle's account, the temperate person does not crave pleasures more than is right, nor does he crave the wrong sorts of pleasures. The self-indulgent, by contrast, will crave either greater quantities of physical satisfaction than is right, for example, more food than he needs for healthy sustenance, or else he will crave the wrong sorts of physical satisfaction. Aristotle maintains that the other vice opposed to temperance, insensibility, is not merely rare but quite unnatural in humans as well as other animals. The point of both temperance and self-indulgence is the satisfaction of desire, in the one case correctly achieved in the pursuit of human flourishing, in the other a disordered pursuit of pleasure for its own sake rather than for one's natural end. Insensibility, by contrast, is an outright denial of one's basic physical needs and, by extension, a contravention of one's natural end.

Post-Aristotelian philosophy is quite heterogeneous in its treatment of ethical issues. The central conception of the virtue of self-control still has to do with controlling one's desires, though in certain cases (see, for example, *SVF* 1.200–201) it is connected more directly to the foregoing of pleasures. For the Stoics, *sôphrosunê* was counted among the cardinal virtues along with courage, prudence, and justice. Since their highest good was a life lived in accordance with nature (*kata phusin*) the wise person is one whose understanding of nature and his place in it leads him to a kind of unity with nature, and they defined *sophrosynê* very generally as practical wisdom concerned with choice and avoidance (Plut. *Stoic. rep.* 1034ce). The Epicureans, according to Cicero (*De finibus* 1.14.47–8), associated self-control with peace of mind and harmony, by freeing us from the disruptions and consequences of an unbridled pursuit of pleasure. This has value, according to them, not in itself, but because it secures greater pleasure over the long run.

See also Aristotle; Hellenistic Thought; Homer; Pain; Plato; Pleasure; Virtue and Vice.

Bibliography

Macintyre, Alisdair. "*Sôphrosunê*: How a Virtue Can Become Socially Disruptive." *Midwest Studies in Philosophy* 13 (1988): 1–11.

North, Helen. *Sophrosyne and Self-Restraint in Greek Literature*. Ithaca, NY: Cornell University Press, 1966.

Rademaker, Adriaan. *Sophrosyne and the Rhetoric of Self-Restraint: Polysemy and Persuasive Use of an Ancient Greek Value Term*. Leiden: Brill, 2005.

Santas, Gerasimos. "Socrates at Work on Virtue and Knowledge in Plato's *Charmides*." *Exegesis and Argument: Studies in Greek Philosophy Presented to Gregory Vlastos. Phronesis* suppl. vol. 1 (1973): 105–132.

Young, C. "Aristotle on Temperance." *Philosophical Review* 97 (1988): 521–542.

Scott Carson (2005)

TENSE

Tense is a grammatical category by means of which some natural languages express the temporal location of the event described by the sentence in which the grammatical tense occurs. (This definition assumes a distinction between grammatical and lexical categories. For the technically inclined, lexical categories are part of the lexicon of a language and are open classes [classes that allow new vocabulary through compounding, derivation, coining, and borrowing]. They become inflected, and do not contract, affix, or cliticize. Examples of lexical categories are nouns, adjectives, adverbs, verbs. Grammatical categories are part of the grammatical system of a language and are closed classes [classes that do not allow additions]. They may contract, affix, or cliticize. Examples include inflectional and derivational morphemes and function words, such as prepositions, determiners, conjunctions, and pronouns.) An instance of a tensed language is English. In the English unembedded sentence "Bill called," the grammatical tense "-ed" conveys the information that Bill's call happened before the time of speech. Similarly, in the English sentence "Bill will call," the grammatical tense "will" contributes the information that Bill's call occurs after the time of speech. When a language does not have grammatical tenses, as in the case of Chinese, the temporal

information may be conveyed by lexical categories, such as adverbs.

The mapping between the grammatical tenses of a natural language and the expression of temporal location is very complex, and one of the goals of linguistic semantics is to investigate the relation between grammatical tenses and the expression of time. To achieve this goal, scholars in both linguistics and philosophy have proposed different theories of tense.

One type of theory, beginning with the work of the logician Arthur Prior, analyzes tenses as temporal operators. Prior (1957, 1967) treated the past and future tenses as sentential operators meaning "it was the case that" and "it will be the case that," respectively. The sentence "Bill called" is translated into *P*(^*p*) and is true in a world *w* at a time *t* if and only if "Bill calls" is true in *w* at a time *t*' *t* (^*p* is the intension of *p*, and "⟨" means "earlier than"). In his intensional system, Montague (1974) adopted Prior's tense logic by introducing tense operators for the past and future tenses, with the time parameter of the intensional expression embedded in the tense operator.

A different approach to the analysis of tense is that proposed by Reichenbach (1947). According to Reichenbach, tense is not a temporal operator but a complex structure built from a small set of primitives: the event time (*E*), the speech time (*S*), the reference time (*R*), and two relations that can hold between these times, simultaneity (symbolized with a comma) and anteriority (symbolized with an underscore). One of these relations holds between *S* and *R*, and one relation holds between *R* and *E*. The relation between *S* and *E* is not represented but is inferred from the first two. With this small set of primitives, Reichenbach was able to define the set of possible English tenses. For example, the simple past, future, and present tenses have the structures [*E*, *R*_*S*], [*S*_*R*, *E*], [*S*, *R*, *E*], respectively. The contribution of *R* becomes crucial in the analysis of complex tenses, such as the future and past perfect (which Reichenbach called "anterior future" and "anterior past"), where *R* overlaps neither *E* nor *S*. For example, the past perfect in "At 3:00 p.m., John had (already) called" has the structure [*E*_*R*_*S*], where the calling time *E* precedes the reference time *R* (3:00 p.m.), which in turn is before *S*.

The case of the future perfect is a little more complex. Take the sentence "By 3:00 p.m., John will have called." Our intuition is that, while 3:00 p.m. must follow the speech time, the time of John's calling must be before 3:00 p.m. but does not have to follow the speech time. The availability of the reference time *R* allows Reichenbach to account for this intuition easily: *R* must be future relative to *S*, and *E* must be past relative to *R*, but the relation between *S* and *E* is left unspecified, leaving open the following three possibilities: [*S*_*E*_*R*] or [*S*, *E*_*R*] or [*E*_*S*_*R*].

A third family of theories views tenses as temporal predicates expressing relations between times (or events). Zagona (1995), Stowell (1996), and Higginbotham (2002) are the main proponents of this view. According to these authors, tenses express temporal relations, such as anteriority, posteriority, and simultaneity, between two events (or times). However, unlike Reichenbach's theory, events (or times) are not introduced by the tenses but by verbs and adjectives instead. This view is also different from the operator analysis of tense since tenses are not operators shifting evaluation parameters

The operator theory of tense has been very influential and has inspired semantic analyses where tense is an existential quantifier binding the time argument in the predicate. Versions of the quantificational theory of tense have been proposed by David Dowty (1979), Arnim von Stechow (1995), Toshi Ogihara (1996), and Dorit Abusch (1997), among others (see Kuhn and Portner 2002 for an overview on tense logics for natural languages). Barbara Partee (1973, 1984) has observed that existential-quantifier theories are problematic when we consider some occurrences of tense in natural language. Her famous example is

(1) I didn't turn off the stove

uttered as the speaker is driving down the freeway. According to the existential analysis of tense, the sentence can be interpreted either as "There is no past time at which I turned off the stove" or as "There is a past time at which I did not turn off the stove," depending on the scope of negation with respect to the temporal quantifier. However, neither interpretation correctly captures the meaning of the sentence in the context we are considering. Clearly, the speaker did not mean to negate the existence of *any* time at which she turned off the stove, nor did she mean to assert the existence of *some* time at which she did not turn off the stove. She merely meant to assert that she did not turn off the stove at a *contextually salient past time*.

To solve this problem, Partee proposed a *referential* analysis of tense, in which tenses are linguistic devises by which we refer to times salient in the previous discourse. This analysis treats English tenses analogously to how Hans Kamp (1981) and Irene Heim (1988) treated pronouns and nominal anaphora. Variants of this idea have also been proposed by Enç (1986), Heim (1994), and

Kratzer (1998). However, there are occurrences of tenses that are not about particular times, as in the sentence

(2) Einstein visited Princeton

where a quantificational analysis of tense seems more apt. Both quantificational and referential theories of tense need to account for the occurrences of tense in (1) and (2). One possibility is to analyze the past tense as a restricted quantifier, just like ordinary nominal quantifiers. In this analysis, (1) would assert that there is no time within a contextually salient past interval at which the individual turned off the stove. The indeterminate reading of (2) would arise when the restriction of the quantifier is Einstein's entire life span.

The discussion so far has been about the meanings of the English tenses, and we have been silently assuming that there is a one-to-one correspondence between grammatical tenses and these meanings. While this is generally true in simple clauses, there are exceptions. For example, (3) illustrates a use of the grammatical present tense with the so-called *futurate* meaning.

(3) The 4:00 o'clock train leaves in five minutes.

Example (4) from Enç 1996 illustrates a mismatch between the future tense morpheme will and the semantics of the future, since Pat's sleeping is understood to be overlapping the speech time. Similarly, example (5) illustrates a mismatch between the past tense morpheme on was and the semantics of the past, since the past tense is allowed to occur with the future adverb tomorrow.

(4) Pat will be sleeping now.

(5) Pat was leaving tomorrow.

The idea of a one-to-one correspondence between tense morphology and tense meanings turns out to be even more problematic when we consider subordinate clauses and the phenomenon of sequence of tense. Consider the following sentence, where the matrix verb and the embedded verb both occur in the past tense.

(6) Bill thought that Sue was pregnant.

There are two possible readings of (6). According to the first reading, the content of Bill's thought was that Sue was pregnant at some time before the time at which Bill was having the thought. This is the so-called *shifted* reading. According to the second reading, the content of Bill's thought is that Sue was pregnant at the time when Bill was having the thought. This is the so-called *simultaneous* reading. The possible simultaneous reading, where the embedded past morpheme is not interpreted as a past tense, seems problematic for a theory in which the morpheme "-ed" is always interpreted as a semantic past. The simultaneity relation, generally expressed in English by the absence of either past or future morphemes, is expressed in sequence of tense with the past morpheme. Furthermore, notice that when we actually embed a grammatical present tense under a grammatical past tense, as in (7), we obtain not a simultaneous reading but yet a third reading, the so-called *double-access* reading. As pointed out by Enç (1987), in (7) Bill's thought is that Sue's pregnancy extends over a period of time including both the time at which Bill had the thought and the time at which (7) was uttered.

(7) Bill thought that Sue is pregnant.

Let us go back to (6). Operator theories of tense try to reconcile the occurrence of an embedded past-tense morpheme with the simultaneous reading by proposing accounts where, at the level of semantic interpretation, the embedded past tense is deleted (Ogihara 1989, 1995; von Stechow 1995) or is semantically bound by the matrix past tense (Abusch 1997) and its temporal features are deleted (von Stechow 2003).

Within the referential theories of tense, Enç (1987) proposed that the simultaneous reading of (6) is obtained when the embedded past tense is coindexed with the matrix past tense, and thus bound by it. Therefore, in her account, the embedded past tense refers to the past time referred to by the matrix past tense. Abusch (1988) points out that already in Kamp and Rohrer (1984) we can find some evidence against the claim that the morphological past tense in an embedded clause is interpreted as a semantic past tense. Abusch provides (8) as an example illustrating the fact that, the most embedded past tense, that associated with "were," cannot refer to any past time since, in the intended reading, the event of having their last meal together is understood as overlapping John's saying event.

(8) John decided a week ago that in ten days at breakfast he would say to his mother that they were having their last meal together.

Among referential theories of tense, a slightly different approach has been taken by Kratzer (1998). Kratzer's proposal, while inheriting several elements from Abusch's (1997) analysis of sequence of tense, is based on Irene Heim's observation that in some occurrences, pronouns have features that are not interpreted. For example, the second occurrence of "I" in Heim's example (9) is interpreted not as an indexical but as a bound variable in the so-called strict reading.

(9) Only I got a question that I understood.

According to Kratzer, the simultaneous reading of (6) arises when the embedded past tense is interpreted as a bound variable, just as the second occurrence of the first-person pronoun in (9) is interpreted as a bound variable, rather than as an indexical. The features on both the embedded "I" in (9) and the embedded past tense in (6) are not "interpretable" (in the sense of Chomsky 1995), that is, they do not contribute to the LF (logical form) representations of these sentences. They are zero pronouns, or zero tenses, whose morphological and phonological features probably derive from agreement with their antecedents and do not carry any semantic information. Kratzer's parallel between zero pronouns and sequence-of-tense tenses expands Partee's original insight about an analogy between pronouns and tenses. The parallel between pronouns and tenses is also at the center of recent work by Schlenker (2003) and von Stechow (2003).

The discussion of sequence-of-tense phenomena above has been concerned with sequences of tenses where th e matrix tense is a past. Hornstein (1990) claims that the availability of the simultaneous reading in sequence of tense is not restricted to the past tense but applies to all tenses. Enç (1996) challenges this claim on the basis of examples like (10), where, according to her judgment, only the shifted interpretation is possible:

(10) Mary will say that she will be tired.

Furthermore, Enç points out that the double-access reading is not forced by embedding the present tense under the future—a fact that thus sets the future tense apart from the past tense. In (11) the only reading is that Mary is upset at the time of John's assertion.

(11) John will say that Mary is upset.

On the basis of these asymmetries between the future and the past and on the basis of the observation that future-oriented modals behave like "will" with respect to sequence of tense, as in (12), Enç suggests that the future morpheme "will" is not a tense but a modal.

(12) John must claim that he is sick.

This last point raises the question of the relation between tense and two other grammatical categories: aspect and mood. Tense, aspect, and mood are intimately related, since they all contribute some information about the event that a given sentence is about: Tense, as mentioned, conveys information about the time of the event; aspect conveys information about the beginning, duration, completion, or repetition of the event; finally, mood conveys information about whether the sentence is about a possible or actual event. It is common to assume, however, that these categories are distinct, even though their boundaries are not always clear. (An example is the debate over the semantics of the present perfect in English and other languages. For a general overview of the topic, see Alexiadou, Rathert, and von Stechow 2003 and the references cited there.) Further comparative studies across Indo-European and non-Indo-European languages will, it is hoped, shed light on these intricate issues.

See also Artificial and Natural Languages; Chomsky, Noam; Intensional Transitive Verbs; Language; Montague, Richard; Prior, Arthur Norman; Quantifiers in Natural Language; Reichenbach, Hans; Semantics.

Bibliography

Abusch, Dorit. "Sequence of Tense and Temporal De Re." *Linguistics and Philosophy* 20 (1997): 1–50.

Abusch, Dorit. "Sequence of Tense, Intensionality, and Scope." In *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, edited by Hagit Borer, 1–14. Stanford, CA: CSLI, 1988.

Alexiadou, Artemis, Monika Rathert, and Arnim von Stechow, eds. *Perfect Explorations*. Berlin: Mouton de Gruyter, 2003.

Chomsky, Noam. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

Dowty, David. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Grammar and Montague PTQ*. Dordrecht, Netherlands: D. Reidel, 1979.

Enç, Mürvet. "Anchoring Conditions for Tense." *Linguistic Inquiry* 18 (1987): 633–657.

Enç, Mürvet. "Tense and Modality." In *The Handbook of Contemporary Semantic Theory*, edited by Shalom Lappin, 345–358. Oxford, U.K.: Blackwell, 1996.

Enç, Mürvet. "Towards a Referential Analysis of Temporal Expressions." *Linguistics and Philosophy* 9 (1986): 405–426.

Heim, Irene. "Comments on Abusch's Theory of Tense." In *Ellipsis, Tense, and Questions*, edited by Hans Kamp, 143–170. Amsterdam: University of Amsterdam, 1994.

Heim, Irene. "The Semantics of Definite and Indefinite Noun Phrases." PhD diss. University of Massachusetts, Amherst, 1982. Published by Garland, New York, 1988.

Higginbotham, James. "Why Is Sequence of Tense Obligatory." In *Logical Form and Language*, edited by G. Preyer and G. Peter. New York: Oxford University Press, 2002.

Hornstein, Norbert. *As Time Goes By: Tense and Universal Grammar*. Cambridge, MA: MIT Press, 1990.

Kamp, Hans. "A Theory of Truth and Semantic Representation." In *Formal Methods in the Study of Language, Part I*, edited by J. Groenendijk, T. Janssen, and M. Stokhof, 277–322. Amsterdam: Mathematisch Centrum, 1981.

Kamp, Hans, and Christian Rohrer. "Indirect Discourse." Manuscript. University of Texas, Austin, and University of Stuttgart, 1984.

Kratzer, Angelika. "More Structural Analogies between Pronouns and Tenses." In *Proceeding of Semantics and*

Linguistic Theory, VIII, edited by Devon Strolovitch and Aaron Lawson. Ithaca, NY: CLC Publications, 1998.

Kuhn, S., and P. Portner. "Tense and Time." In *The Handbook of Philosophical Logic*. 2nd ed., edited by D. Gabbay and F. Guenthner, 6: 277–346. Dordrecht, Netherlands: D. Reidel, 2002.

Montague, Richard. *Formal Philosophy: Selected Papers of Richard Montague*, edited by Richmond H. Thomason. New Haven, CT: Yale University Press, 1974.

Ogihara, Toshiyuki. "Double-Access Sentences and References to States." *Natural Language Semantics* 3 (1995): 177–210.

Ogihara, Toshiyuki. "Temporal Reference in English and Japanese." PhD diss. University of Texas at Austin, 1989.

Ogihara, Toshiyuki. *Tense, Attitudes, and Scope*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1996.

Partee, Barbara. "Nominal and Temporal Anaphora." *Linguistics and Philosophy* 7 (1984): 243–286.

Partee, Barbara. "Some Structural Analogies between Tenses and Pronouns in English." *Journal of Philosophy* 18 (1973): 601–609.

Prior, Arthur N. *Past, Present, and Future*. Oxford, U.K.: Oxford University Press, 1967.

Prior, Arthur N. *Time and Modality*. Oxford, U.K.: Oxford University Press, 1957.

Reichenbach, Hans. *Elements of Symbolic Logic*. New York: Macmillan, 1947.

Schlenker, Philippe. "A Plea for Monsters." *Linguistics and Philosophy* 26 (2003): 29–120.

Stowell, Tim. "The Phrase Structure of Tense." In *Phrase Structure and the Lexicon*, edited by Johan Rooryck and Laurie Zaring, 277–291. Dordrecht, Netherlands: Kluwer, 1996.

Von Stechow, Arnim. "Feature Deletion under Semantic Binding: Tense, Person, and Mood under Verbal Quantifiers." In *Proceedings of the North East Linguistic Society 33*, edited by Makoto Kadowaki and Shigeto Kawahara. Amherst: GLSA, University of Massachusetts, 2003.

Von Stechow, Arnim. "On the Proper Treatment of Tense." In *Proceeding of Semantics and Linguistic Theory, V*, edited by Mandy Simons and Teresa Galloway. Ithaca, NY: CLC Publications, 1995.

Zagona, Karen. "Temporal Argument Structure: Configurational Elements of Construal." In *Temporal Reference, Aspect, and Actionality*. Vol. 1: *Semantic and Syntactic Perspectives*, edited by Pier Marco Bertinetto, Valentina Bianchi, James Higginbotham, and Mario Squartini. Turin, Italy: Rosenberg e Sellier, 1995.

Michela Ippolito (2005)

XENOPHON [ADDENDUM]

The central concern regarding Xenophon since the mid-1960s has been his place in the so-called Socratic problem, the question of to what extent our knowledge of the historical Socrates is accurate and on the basis of what sources we may have any confidence in the portrait of him that has come down to us. Although Xenophon's Socratic writings have been criticized on the grounds that their philosophical acumen does not compare with that of Plato, scholarship since antiquity has tended to regard them as important sources of information about the life and character of the historical Socrates. But Xenophon's portrait of Socrates has received mixed reviews. Scholars continue to debate whether the Socrates that we encounter in the early, Socratic dialogues of Plato is the historical man himself, a Platonic fiction, or something in between, and the portrait of Socrates that we find in Aristophanes is clearly something of a caricature, in which Socrates appears to serve virtually as a stock character for the ridicule of philosophers generally.

This has prompted some to claim that Xenophon is our best hope for piecing together the real life of the man. Others, however, argue that Xenophon shows no real sophistication in his writings and hence cannot be relied upon to produce an accurate portrait of such a central figure in the history of philosophy, and that if we compare Xenophon's portrait of Socrates with those of other writers of *Sōkratikoi logoi* (stories about Socrates), a genre that grew up among the followers of Socrates shortly after his death in 399 BCE, we find that we have no compelling reason to prefer his portrait to any other, including Plato's. Plato himself mentions the views of Xenophon only once (*Laws* 694c), and only to criticize an element of the political education of Cyrus as portrayed in the *Cyropaedeia*. Plato has nothing to say about Xenophon's portrait of Socrates. The other writers of *Sōkratikoi logoi* (Antisthenes, Phaedo, Eucleides, Aristippus, Aeschines, and Plato), were actively writing memoirs of Socrates as early as the 390s and 380s, but Xenophon did not begin to write his *Sōkratikoi logoi* until the 360s, and some scholars see in him a repository of recycled information, with at least one scholar suggesting that Xenophon's own youthful memories of Socrates were "filtered through the Socratic literature that had been published in the meantime" (Kahn 1996, p. 30).

Another area of scholarly attention since the mid-1960s has focused on Xenophon as comparative biographer. Even if we accept the view that his portrait of Socrates is as accurate than that of any other Socratic, some scholars maintain that we may nevertheless see in his accounts of Socrates and Cyrus an attempt at comparative biography that has value in its own right. This judgment must be weighed against the view of other scholars who argue that Xenophon's imagination is not on a par with those of the other Socratics, nor is his philosophical acumen up to the task of drawing and compar-

ing such lives with anything like the skill that one finds in, for example, the writings of Plutarch.

See also Antisthenes; Aristippus of Cyrene; Plato; Plutarch of Chaeronea; Socrates.

Bibliography

WORKS BY XENOPHON

Xenophontis opera omnia, edited by E. C. Merchant. Oxford, U.K.: Clarendon, 1950–1954.

Mémorables, edited by Michele Bandini. Translated by Louis-André Dorion. Paris: Belles lettres, 2000–.

WORKS ON XENOPHON

Anderson, J. K. *Xenophon*. Bristol, U.K.: Bristol Classical, 2001.

Cooper, John M. "Notes on Xenophon's Socrates." In his *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 3–28. Princeton, NJ: Princeton University Press, 1999.

Dorion, Louis-André, and Michele Bandini. "Introduction générale." In *Mémorables*, by Xenophon. Vol. 1. Paris: Belles lettres, 2000.

"Les écrits socratiques de Xénophon." Special issue, *Études philosophiques* 2 (2004).

Gray, Vivienne J. *The Framing of Socrates: The Literary Interpretation of Xenophon's "Memorabilia."* Stuttgart, Germany: F. Steiner, 1998.

Gray, Vivienne J. "Xenophon's Image of Socrates in the *Memorabilia*." *Prudentia* 27 (1995): 50–73.

Kahn, Charles H. "*Sokratikoi logoi*: The Literary and Intellectual Background of Plato's Work." In his *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*, 29–35. Cambridge, U.K.: Cambridge University Press, 1996.

Morrison, Donald R. *Bibliography of Editions, Translations, and Commentary on Xenophon's Socratic Writings, 1600–Present*. Pittsburgh, PA: Mathesis Publications, 1988.

Morrison, Donald R. "On Professor Vlastos' Xenophon." *Ancient Philosophy* 7 (1987) 9–22.

Morrison, Donald R. "Xenophon's Socrates on the Just and the Lawful." *Ancient Philosophy* 15 (1995): 329–347.

Vander Waerdt, Paul A. *The Socratic Movement*. Ithaca, NY: Cornell University Press, 1994.

Vlastos, Gregory. "The Evidence of Aristotle and Xenophon." In his *Socrates: Ironist and Moral Philosopher*, 81–106. Ithaca, NY: Cornell University Press, 1991.

Scott Carson (2005)

thematic outline of contents

The Thematic Outline of Contents has been constructed to assist readers who wish to explore a number of entries in a specific time period or in a distinct subfield of philosophy. Entries have, accordingly, been grouped under two general headings: "Historical Periods" and "Subfields of Philosophy."

The personal entries and a few of the subject entries in the *Encyclopedia* can be placed in one of the following five historical time periods.

(1) Ancient Philosophy—from Homer in the 8th century BCE to Augustine in the 4th–5th century CE

(2) Medieval Philosophy—from Augustine to Thomas Aquinas in the 13th century CE

(3) Modern Philosophy—from Thomas Aquinas to Georg W. F. Hegel in the 18th–19th century CE

(4) Nineteenth Century

(5) Twentieth Century

To group philosophers in this fashion may give the false impression that ancient philosophy began in the time of Homer and ended with Augustine and that medieval philosophy began in the time of Augustine and ended with Aquinas after whom we get modern philosophy. The developing story of philosophy does not lend itself to such rigidly defined temporal boundaries. Indeed, Augustine was importantly influenced by ancient Platonic thought even as Aquinas was significantly indebted to ancient Aristotelian thought. Moreover, the philosophical traditions springing from Augustine and Aquinas—Augustinianism and Thomism—are very much alive in our world today. If we respect the insight that philosophical thinking develops throughout the centuries by virtue of philosophers engaging in dialogues not only with their contemporaries but also with their predecessors, then we will avoid such false notions as ancient philosophy ending at such and such a time, and we will be safe in talking about certain temporal periods of philosophy—which we are doing in this Thematic Outline of Contents.

Personal entries in the Nineteenth Century section have been divided into two groups: "major" to signify large entries containing more than 2,500 words and "minor" to indicate smaller entries. It must be noted, however, that a personal entry in the category of "minor" may discuss an enormously influential philosopher. The size of a personal entry is not always an indication of the importance of that person in a philosophical tradition. Indeed, in preparing personal entries for very influential philosophers from the current scene, our standard word allocation was between 1,000 and 1,500 words.

Personal entries in the Twentieth Century section are also divided into two groups: "early" referring to scholars whose major work was done in the first half of the twentieth century, and "recent" referring to scholars whose major work takes place in the last half of the twen-

tieth century continuing frequently into the twenty-first century.

Some of the subject entries can be placed within Ancient Philosophy (such as Aretē and Demiurge), Medieval Philosophy (such as Liber de Causis and Scotism) and Modern Philosophy (such as Atheismusstreit and Jansenism). Most subject entries, however, defy allocation to one time frame because they relate to questions that have engaged philosophers for many centuries. Accordingly, most of the subject entries have been grouped within the following twenty-one philosophical subfields that appear in the following order: Epistemology; Philosophy of Mind and Cognitive Science; Metaphysics; Philosophy of Science; Logic, Philosophy of Logic, and Philosophy of Mathematics; Philosophy of Language; Continental Philosophy; Feminist Philosophy; Ethics; Applied Ethics; Social and Political Philosophy; Philosophy of Law; Aesthetics and Philosophy of Art; Philosophy of Religion; Buddhist Philosophy; Chinese Philosophy; Indian Philosophy; Islamic Philosophy; Japanese Philosophy; Jewish Philosophy; and Russian Philosophy. Two additional categories—Philosophical Perspectives and Movements; and Special Topics—complete the Thematic Outline of Contents.

The list of philosophical subfields into which entries have been placed is not exhaustive. Indeed, there are many additional subfields that frequently bear the title "Philosophy of . . . " such as Philosophy of Education, Philosophy of Medicine, Philosophy of Sex, Philosophy of Technology, etc. Many of these subfields have entries dedicated to them in the *Encyclopedia* and they are listed in the Thematic Outline of Contents under Special Topics.

Because most personal entries describe scholars who make contributions in more than one subfield of philosophy (such as Plato, Aristotle, Augustine, Aquinas, Descartes, Hume, Kant, Hegel, Bertrand Russell, and Hilary Putnam to mention only a few), it would be overly cumbersome to list each of them in all the subfields in which they worked. Some persons, however, can be reasonably associated with one particular subfield, such as Monroe Beardsley with "Aesthetics and Philosophy of Art," Georg Cantor with "Logic, Philosophy of Logic, and Philosophy of Mathematics," Confucius and Mencius with "Chinese Philosophy," and Mohammed Arkoun and Seyyed Hossein Nasr with "Islamic Philosophy." Such scholars are listed not only in one of the five historical periods but also within their distinctive subfields.

Clearly, the Historical Periods and the Philosophical Subfields of this Thematic Outline of Contents reflect the influence of the Western philosophical tradition originating with the ancient Greeks. Equally evident is the fact that the *Encyclopedia* contains articles devoted to non-Western philosophical traditions such as the African, Buddhist, Chinese, Indian, Japanese, and Korean. These traditions together represent a host of entries in the *Encyclopedia.* Inserting these entries into the Thematic Outline of Contents might suggest at first glance that the integrity of these traditions is being violated because they are being forced into a Procrustean Western mold. Yet it is important to remember that contemplative people from diverse cultural traditions have pondered some of the same perennial human questions for centuries. Philosophy begins with wonder, and the West has no monopoly on wonder. Human beings from diverse cultures have wondered about such things as truth, knowledge, logic, morality, and the nature of the human and also the transcendent. While the way questions are posed and answers are given may vary significantly from culture to culture, the topics of philosophy are truly multicultural. Admittedly, three of the Historical Periods used in this Thematic Outline of Contents—Ancient, Medieval, and Modern—have employed Western thinkers—Homer, Augustine, Aquinas, and Hegel—as temporal markers. Non-Western traditions would no doubt use other markers. Dividing human development into distinct periods has an element of unavoidable arbitrariness. The point to be emphasized by utilizing such divisions, however, is that philosophical thinking is a growing concern and that stages of growth are usually recognizable. The entries covering diverse philosophical traditions will, it is hoped, display that growth and also the commonality of human wonder.

HISTORICAL PERIODS IN PHILOSOPHY

Ancient Philosophy

Aenesidemus
Agent Intellect
Agrippa
Aitia
Alcinous
Alcmaeon of Croton
Alexander of Aphrodisias
Alexander of Hales
Anaxagoras of Clazomenae
Anaximander
Anaximenes
Ancient Aesthetics
Ancient Skepticism
Antiochus of Ascalon
Antiphon
Antisthenes
Apeiron/Peras
Apologists
Arcesilaus
Archē
Archytas of Tarentum
Aretē/Agathon/Kakon
Aristippus of Cyrene
Aristo of Chios
Aristotelianism
Aristotle
Arius and Arianism
Atomism
Carneades
Celsus
Chrysippus
Cicero, Marcus Tullius
Cleanthes
Clement of Alexandria
Confucius
Cosmos
Cratylus
Cynics
Cyrenaics
Demiurge
Dikē
Diodorus Cronus
Diogenes Laertius
Diogenes of Apollonia
Diogenes of Sinope
Dogma
Dong Zhongshu
Empedocles
Epictetus
Epicureanism and the Epicurean School
Epicurus
Eternity
Eudaimonia
Eusebius
Galen
Gongsun Long
Gorgias of Leontini
Greek Academy
Greek Drama
Gregory of Nazianzus
Gregory of Nyssa
Guo Xiang
Han Fei
Hen/Polla
Heraclitus of Ephesus
Hippias of Elis
Hippocrates and the Hippocratic Corpus
Homer
Hui Shi
Hypatia
Iamblichus
Impetus
Inner Senses
Kalon
Katharsis
Laozi
Leucippus and Democritus
Logos
Longinus (Pseudo)
Lucian of Samosata
Lucretius
Mani and Manichaeism
Marcion
Marcus Aurelius Antoninus
Megarians
Melissus of Samos
Mencius
Mimesis
Moira/Tychē/Anankē
Mozi
Musonius Rufus
Nāgārjuna
Nemesius of Emesa
Neoplatonism
Nomos and Phusis
Nous
Numenius of Apamea
Origen
Orphism
Ousia
Panaetius of Rhodes
Parmenides of Elea
Pelagius and Pelagianism
Peripatetics
Phantasia
Philodemus
Philo Judaeus
Philolaus of Croton
Philo of Larissa
Philo of Megara
Phronêsis
Plato
Plotinus
Plutarch of Chaeronea
Pneuma
Porphyry
Posidonius
Pre-Socratic Philosophy
Proclus
Prodicus of Ceos
Protagoras of Abdera
Psychē
Pyrrho
Pythagoras and Pythagoreanism
Seneca, Lucius Annaeus
Sextus Empiricus
Simon Magus
Socrates
Sophia
Sophists
Sôphrosunê
Stoicism
Strato and Stratonism
Tertullian, Quintus Septimius Florens
Thales of Miletus
Themistius
Theophrastus
Thucydides
Timon of Phlius
Valentinus and Valentinianism
Vasubandhu
Wang Bi
Wang Chong
Xenophanes of Colophon
Xenophon
Xunzi
Yang Xiong
Yang Zhu
Zeno of Citium
Zeno of Elea
Zhuangzi

Medieval Philosophy

Abelard, Peter
Ailly, Pierre d'
Albert of Saxony
Albert the Great
Albo, Joseph
al-Fārābī
al-Ghazālī, Ahmad
al-Ghazālī, Muhammad
al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq
Anselm, St.
Augustine, St.
Augustinianism
Averroes

Modern Philosophy

Nineteenth Century: Major Personal Entries

Nineteenth Century: Minor Personal Entries

Early Twentieth Century: Personal Entries

Recent Twentieth Century: Personal Entries

Metaphysics

Philosophy of Science

Logic, Philosophy of Logic, Philosophy of Mathematics

Philosophy of Language

Continental Philosophy

Feminist Philosophy

Ethics

Applied Ethics

Social and Political Philosophy

Philosophy of Law

Aesthetics and Philosophy of Art

Philosophy of Religion

Buddhist Philosophy

Chinese Philosophy

Indian Philosophy

Islamic Philosophy

Japanese Philosophy

Jewish Philosophy

Russian Philosophy

Philosophical Perspectives and Movements

bibliographies

The First Edition of the Encyclopedia of Philosophy *included bibliographical essays dealing with philosophy dictionaries and encyclopedias, philosophy journals, and philosophy bibliographies. To preserve and enhance these essays, they have been reproduced in this Second Edition along with detailed updates. The updates to the bibliographies cover material published between 1965 and mid-2005. All of the references appear in OCLC's WorldCat bibliographic database and are thus available either in mid- to large-size academic libraries, or through interlibrary loan. While the bibliographies are extensive, they are not exhaustive. This is especially true in the case of the journal bibliography, where less readily available non-English-language journals have been excluded, as have journals published for short periods of time. Accessibility was deemed to be more important than exhaustive coverage. The subject coverage includes both general philosophical works and works from the major sub-domains of philosophy. The bibliographic lists show that philosophy is a vital, worldwide discipline. A perusal of the journal bibliography will show that new journals are appearing every year, and the dictionary and encyclopedia bibliography identifies publications in fifty different languages. The constant stream of new journals and the accumulation of philosophical resources in so many languages are indicators of a truly vibrant discipline.*

PHILOSOPHY BIBLIOGRAPHIES

Lists of philosophers and the titles of their works were for the most part provided only *en passant* by ancient and medieval writers and scholars, as in the brief citations scattered through the first book of Aristotle's *Metaphysics* and throughout Aquinas' *Summa Theologica.* It is true that Diogenes Laërtius' listing was somewhat more systematic, but philosophical bibliographies fully worthy of the name date from more recent times.

Modern philosophy has been well supplied with bibliographies in the general sense of the term, as will be noted in the present survey, but it has been weak in a special variety of bibliographical literature, namely, journals of abstracts. The two main journals containing abstracts of current work in philosophy—the *Bibliographie de la philosophie* and the *Bulletin signalétique: Philosophie, sciences humaines*—have done and are doing a good job as far as they go, but the scope of each is limited: the first covers books only, and the two-line précis in the second are enough only to whet a desire for more.

Modern bibliographies of philosophy are of four kinds: general bibliographies; those covering a specific region or country; those covering a particular period, movement, or philosopher; and those covering a specific philosophical discipline.

GENERAL BIBLIOGRAPHIES

BIBLIOGRAPHICAL BOOKS AND PAMPHLETS. One of the earliest of the general bibliographies of philosophy is the *Bibliotheca Philosophorum Classicorum Authorum Chronologica; in qua Veterum Philosophorum Origo, Successio, Aetas, & Doctrina Compendiosa, ab Origine Mundi, Usq. ad Nostram Aetatem, Proponitur; Quibus Accessit Patrum, Ecclesiae Christi Doctorum a Temporibus Apos-*

tolorum, Usque ad Tempora Scholasticorum ad An. Usq. Do. 1140, Secundum Eandem Temporis Seriem, Enumeratio, by Johann Jacob Fries (Zürich, 1592, 110 pages), with about 2,500 entries. Three of its significant successors in the next three hundred years are the *Bibliotheca Realis Philosophica*, by Martin Lipen (2 vols., Frankfurt, 1679), with about 40,000 entries, some on subjects no longer regarded as philosophical in a strict sense; the *Bibliotheca Philosophica*, by B. G. Struve (Jena, 1704; 5th ed., 2 vols., 1740), containing about 4,000 entries; and the *Systematisch-alphabetischer Hauptkatalog der Königlichen Universitätsbibliothek zu Tübingen; Erstes Heft; A. Philosophie* (Tübingen, 1854, 63 pages), with about 3,000 entries and with annual supplements to 1880.

Of the four pre-twentieth-century bibliographies mentioned, all are available at the Library of Congress in Washington and at the British Museum in London. The last-named item is also available at the New York Public Library and at the Library of the University of Illinois.

In the twentieth century four main general philosophical bibliographies have been compiled. The first is the *Bibliography of Philosophy, Psychology, and Cognate Subjects*, by Benjamin Rand (2 vols., New York, 1905), which has about 70,000 entries and is a major work of scholarship. It was published as the two-part Volume III of the *Dictionary of Philosophy and Psychology*, edited by James M. Baldwin (3 vols., New York, 1901–1905). Part I of the two-part *Bibliography* covers histories of philosophy and works by and about philosophers from Abel to Zwingli, and Part II is systematic.

Second among the main general bibliographies of the present century is the *Bibliographische Einführung in das Studium der Philosophie*, edited by I. M. Bocheński, which consists of 20 fascicles (24 to 85 pages each) published at Bern from 1948 to 1950 and which covers philosophy in certain periods (ancient and medieval philosophy), countries (modern Italian, French existentialist, and American philosophy), religious and ethnic groups (Buddhist, patristic, Jewish, and Arabic philosophy), systems and disciplines (philosophy as a whole, symbolic logic, and logical positivism), and individuals (Plato, Aristotle, Augustine, Aquinas, and Kierkegaard).

The third principal source of this kind is Gilbert Varet's *Manuel de bibliographie philosophique* (2 vols., Paris, 1956), which contains about 25,000 entries, Volume I being historical and Volume II systematic.

Finally, there is Wilhelm Totok's *Handbuch der Geschichte der Philosophie* (Frankfurt, 1964–), of which the first volume, *Altertum* (400 pages), covers works on Indian, Chinese, Greek, and Roman philosophy, with an introduction listing works on the methodology of research in philosophy and on the general history of philosophy, dictionaries of philosophy, introductions to philosophy, and works on the philosophical disciplines. Articles from over 400 periodicals are cited.

BIBLIOGRAPHICAL SERIALS. Apparently the earliest general serial covering works in philosophy was the *Allgemeines Repertorium der Literatur; . . . philosophische Literatur*, by J. S. Ersch (Jena and Weimar, one volume each for 1785–1790, 1791–1795, and 1796–1800). Partly overlapping it in time was the *Lehrbuch der Geschichte der Philosophie und einer kritischen Literatur derselben,* by J. G. Buhle (Göttingen, one volume for each year from 1796 to 1804). After a gap of 87 years, the *Critical Review of Theological and Philosophical Literature*, edited by S. D. F. Salmond, was published at Edinburgh, covering the years 1891 to 1904. It was succeeded by the *Review of Theology and Philosophy*, edited by Allan Menzies, also at Edinburgh, which covered 1905/1906 to 1914/1915.

Meanwhile, in 1895 at Louvain a periodical was begun which was entitled the *Sommaire idéologique des ouvrages et des revues de philosophic*. After a number of changes (and with no volumes published from 1915 to 1933 and from 1941 to 1945), this periodical is now entitled the *Répertoire bibliographique de la philosophic*. It is issued four times a year and is one of the three general bibliographical serials now being published in the field of philosophy; it covers both books and periodical articles. (It is reproduced *in toto*, with Dutch headings replacing the French headings, in the *Tijdschrift voor Filosofie*, published quarterly at Louvain.) A second of the three leaders in this category is the *Bibliographie de la philosophie,* begun in 1937 as a semiannual by the International Institute of Philosophy, continued (with the omission of the years 1939 to 1945) until 1953, and issued since 1954 four times a year by the International Federation of Philosophical Societies; it covers books only, with a summary of each.

The third is the *Bulletin signalétique: Philosophie, sciences humaines* (entitled the *Bulletin analytique: Philosophie* from 1947 to 1955), published quarterly at Paris by the Centre de Documentation du Centre Nationale de la Recherche Scientifique; it is the only world-wide source of its kind which not only covers both books and periodicals but also contains a succinct abstract of each entry.

Remaining to be mentioned, as regards serial bibliographies of philosophy, are a number of sources which

are either limited in scope in one way or another or are no longer issued.

A general world-wide serial no longer issued but useful for works published in the period in which it appeared is *Philosophic Abstracts*, published for the most part quarterly at New York from 1939 to 1954, with an index covering 1939 to 1950. It contains abstracts of books and lists of periodical articles.

There are two important serials, of a quasi-bibliographical character, devoted exclusively to critical reviews of philosophical books: The *Philosophischer Literaturanzeiger*, published eight times a year at Meisenheim am Glan (begun in 1949 at Schlesdorf am Kochelsee), which contains about 15 reviews in each issue; and *Philosophical Books*, issued quarterly since 1960 at Leicester, England, which contains about a dozen reviews in each issue, written largely from the viewpoint of analytical philosophy. Also deserving of mention, as regards coverage of books only, is *Scripta Recenter Edita*, issued ten times a year since 1959 at Nijmegen, the Netherlands, which is a list of books on philosophy and theology (each issue containing about 400 entries with emphasis on theology), designed especially for use by acquisitions officers of libraries.

Periodicals. It may be added, as regards serial bibliographies, that selective lists or reviews (and, in a few cases, abstracts) of current philosophical books, plus lists of periodical articles in some cases, are published either in each issue or annually or from time to time in many philosophical periodicals, and the coverage is in some cases fairly comprehensive. (For the names of periodicals in this field, see *Philosophical Periodicals, An Annotated World List*, by David Baumgardt, Washington, 1952, 89 pages, 489 entries; the list, with 157 entries, which appears under the heading "Philosophy" in *Ulrich's Periodicals Directory*, 10th ed., New York, 1963, 667 pages; and the article "Philosophy Journals" in this volume.) Especially strong in book reviews and abstracts are the German philosophical periodicals.

Of the currently published annual bibliographies in philosophical periodicals, mention may be made of the one which appears in the *Deutsche Zeitschrift für Philosophic*, published in East Berlin. Although generally global in coverage, it emphasizes works on dialectical materialism written in Eastern Europe.

Finally, topical, regional, or other summaries and evaluations of current philosophical literature (as distinguished from lists, reviews, or abstracts) appear regularly or occasionally in *The Hibbert Journal* (world-wide), *Cross Currents* (world-wide), *Philosophy* (selected countries), the *Revue philosophique de la France et de l'étranger* (selected countries), and the *Revue des sciences philosophiques et théologiques* (world-wide).

BIBLIOGRAPHICAL SECTIONS OF BOOKS. Many of the standard histories of philosophy contain bibliographical sections. The most important source of this kind is the voluminous bibliographical material in the *Grundriss der Geschichte der Philosophie*, by Friedrich Ueberweg and others (12th ed., 5 vols., Berlin, 1923–1938). The handiest is the series of lists of philosophers preceding each main part of the *History of Philosophy*, by Wilhelm Windelband, translated by James H. Tufts (2 vols., New York, 1958, paperback reprint of the rev. ed. of 1901). Also useful for the history of philosophy are the bibliographical lists (usually divided into "Fonti" and "Studi") at the ends of the chapters of the *Guida storico-bibliografica allo studio della filosofia*, by Carmelo Ferro (Milan, 1949?).

In addition, many introductory works on philosophy contain bibliographical guides. An outstanding example is the discussion of philosophical books, periodicals and dictionaries in Louis de Raeymacker's *Introduction to Philosophy*, translated by Harry McNeill (New York, 1948, 297 pages), on pp. 196–258.

NATIONAL OR REGIONAL BIBLIOGRAPHIES

BIBLIOGRAPHICAL BOOKS AND PAMPHLETS. A convenient list of the bibliographies of philosophy which are national in scope, covering some twenty countries or groups of countries, will be found in *A World Bibliography of Bibliographies*, by Theodore Besterman (4th ed., 4 vols., Lausanne, 1965–1966), Volume III, Columns 4809–4827. Outstanding among these country guides are the *Manuel de la recherche documentaire en France; . . . Philosophie*, by Raymond Bayer (Paris, 1950, 410 pages), with about 6,000 entries; the *Repertorium der Nederlandse Wijsbegeerte*, by J. J. Poortman (Amsterdam, 1948, 404 pages), with about 20,000 entries and a 168-page supplement published in 1958; the *Bibliografia filosofica italiana del 1900 al 1950* (4 vols., Rome, 1950–1957), with about 50,000 entries; the *Bibliografia filosófica española e hispanoamericana* (1940–1958), by Luis Martínez Gómez (Barcelona, 1961, 524 pages), 10,166 entries; and the anonymous *Philosophie und Grenzgebiete, 1945–1964* (Stuttgart, 1964, 434 pages), covering philosophical works in the German language, with a list of periodicals. Also deserving of mention, as regards French philosophy, are the fascicles entitled "Logique et philosophie des sciences," by Robert Blanché (1959, 54 pages), and "Morale

et philosophie politique," by Georges Bastide (1961, 92 pages), in the *Bibliographie française établie à l'intention des lecteurs étrangers* (Paris).

Two volumes of a *Bibliografia Filozofii Polskiej*, covering 1750–1830 and 1831–1864, were published at Warsaw by the Polska Akademia Nauk in 1955 and 1960 (1,241 and 3,771 entries, respectively). The first volume of a *Bibliographie der sowjetischen Philosophie* (listing the articles which appeared in the Soviet periodical *Voprosy Filosofii* from 1947 to 1956; 906 entries) was compiled under the direction of I. M. Bocheński and published in 1959 by the Ost-Europa Institut at the University of Fribourg, Switzerland; four subsequent volumes, published from 1959 to 1964, covered books of 1947 to 1960 and articles of 1957 to 1960.

BIBLIOGRAPHICAL SERIALS. Serials (mostly annuals) devoted to philosophical works issued in particular countries include the following:

Abstracts of Bulgarian Scientific Literature; Philosophy and Pedagogics (Sofia; one volume for each year since 1958).

Bibliografia filosofica italiana (Milan; one volume for each year since 1949).

Bibliography of Current Philosophical Works Published in North America, issued as a supplement to certain issues of *The Modern Schoolman* (St. Louis, Mo.) and covering mainly the United States.

Die deutschen Universitätsschriften zur Philosophie und ihre Grenzgebieten, edited by Kurt Gassen (published annually at Erfurt from 1924 to 1930).

Literarische Berichte aus dem Gebiete der Philosophie, edited by Arthur Hoffman (published semi-annually at Erfurt from 1923 to 1932), which covered current German periodical publications, with special retrospective bibliographies on Hegel, Nietzsche, and others.

"Thèses de doctorat concernant les sciences philosophiques et théologiques soutenues en France," published each year since 1954 in a spring or summer issue of the *Revue des sciences philosophiques et théologiques* and covering the preceding year.

The annual *Handbook of Latin American Studies* (published since 1935, originally and now again at Gainesville, Fla.) regularly contains a chapter on philosophical studies. A "Scandinavian Bibliography," covering philosophical works published in Denmark, Finland, Norway, and Sweden, appears once a year in *Theoria* (Lund, Sweden). The *Heythrop Journal* (Oxford, quarterly) regularly contains a "select list of British books on philosophy and theology."

BIBLIOGRAPHICAL SECTIONS OF BOOKS. Many of the standard historical, critical, or documentary treatments of philosophy in particular countries or regions (American, British, French, German, Indian, etc.; and Latin American, Anglo-American, European, Scandinavian, Western, Oriental, etc.) include extensive bibliographical sections, either at the end of the book or at the end of each chapter. Examples are the bibliographies in the introductions to the several parts of the anthology *The Development of American Philosophy*, edited by W. G. Muelder and others (2d ed., Cambridge, Mass., 1960), with about 500 entries, and the bibliography at the end of Chandradhar Sharma's *Indian Philosophy* (New York, 1962, paperback reprint of the Benares edition of 1952), with about 300 entries.

PERIOD OR MOVEMENT BIBLIOGRAPHIES

BIBLIOGRAPHICAL BOOKS AND PAMPHLETS. Noteworthy among the philosophical bibliographies which cover a particular period are one on antiquity, one on the Renaissance, one on an 11-year period of the twentieth century, and one on the twentieth century as a whole:

Guía Bibliografia de la Filosofía Antigua, by Rodolfo Mondolfo (Buenos Aires, 1959, 102 pages), which is a worthy extension of the author's many substantive contributions to philosophical scholarship.

A Catalogue of Renaissance Philosophers (1350–1650), by John O. Riedl and others (Milwaukee, 1940, 179 pages), dealing with about 2,000 philosophers, with lists of writings in some cases.

Bibliographia Philosophica, 1934–1945, by G. A. de Brie (2 vols., Brussels and Antwerp, 1950–1954), Volume I historical and Volume II systematic; 48,178 entries.

Bibliografia filosofica del siglo XX.; Catalogo de la Exposición Bibliografica Internacional de la Filosofia del Siglo XX. (Buenos Aires, 1952, 465 pages), with 4,011 entries.

A period bibliography which is specialized in two senses (limited with respect to the period when the items were published and to the period with which the items deal) is the *Thomistic Bibliography, 1920–1940*, by Vernon J. Bourke (St. Louis, Mo., 1945; supplement to Vol. 21 of

The Modern Schoolman), with about 5,700 entries. It lists a number of earlier bibliographies of scholastic philosophy.

Illustrative of bibliographies covering philosophical movements is the "Bibliographic der Geschichte der idealistischen Philosophic," in *Idealismus; Jahrbuch für die idealistische Philosophie* (Zurich), Vol. I (1934), pp. 217–256 (about 350 entries). Bibliographies covering philosophical movements in particular countries include V. E. Harlow's *Bibliographical and Genetic Study of American Realism* (Oklahoma City, Okla., 1931, 132 pages), with some 700 entries, and Vito A. Belleza's "Bibliografia italiana sull'esistentialismo," in *Archivio di filosofia*, Vol. 15 (1946), 171–217, with over 700 entries. Works dealing with problems of philosophy and the history of philosophy from the standpoint of Marxism are listed in *O Marxistickej Filozofii a Vedeckom Komunizme*, compiled at the University of Bratislava (Bratislava, 1962, 146 pages), with over 400 entries.

Bibliographies covering individual philosophers are very numerous. They are listed in the appropriate sections of the general bibliographies mentioned earlier. For contemporary philosophers, the comprehensive bibliographies in the volumes of the Library of Living Philosophers, edited by Paul A. Schilpp (now published in La Salle, Ill.), are especially worthy of mention; the series covers C. D. Broad, Rudolf Carnap, Ernst Cassirer, John Dewey, Albert Einstein, Karl Jaspers, G. E. Moore, Sarvepalli Radhakrishnan, Bertrand Russell, George Santayana, and A. N. Whitehead, and volumes on others are in preparation.

BIBLIOGRAPHICAL SERIALS. The main bibliographical serial covering a specific period or movement in philosophy is the annual *Bibliographia Patristica; Internationale patristische Bibliographie*, by Wilhelm Schneemelcher (Berlin, begun with a volume for 1956 published in 1959), with about 1,000 entries in each volume.

BIBLIOGRAPHICAL SECTIONS OF BOOKS. Many of the standard works on the philosophy of a particular period or movement include extensive bibliographical sections either at the end of the volume or at the end of each chapter. As regards books on particular periods, mention may be made, for example, of Maurice de Wulf's *History of Mediaeval Philosophy*, 3d English ed., based on the 6th French ed., translated by E. C. Messenger (2 vols., London, 1935–1938; reprinted 1952); it contains (1) in Volume I an introductory chapter entitled "General Bibliography," with sections on research methods, auxiliary sciences, dictionaries and encyclopedias, collections, monographs on problems, etc. (totaling over 500 entries), and (2) at the end of each major section in each chapter a bibliographical discussion (for example, about 25 entries on John Scotus Erigena).

As regards books on particular movements, mention may similarly be made, for purposes of illustration, of *Logical Positivism* (Glencoe, Ill., 1959), edited by A. J. Ayer, which contains on pp. 381–446 a section entitled "Bibliography of Logical Positivism" (over 2,000 entries), covering not only logical positivism strictly interpreted but also "all types of analytical philosophy." Ayer's book is part of the series entitled Library of Philosophical Movements; the other books in the series (on existentialism, Scholasticism, "realism and the background of phenomenology," etc.) also contain extensive bibliographies.

BIBLIOGRAPHIES OF SPECIFIC DISCIPLINES

Among the bibliographies covering specific philosophic disciplines are the following:

I. M. Bocheński's bibliography of the history of formal logic in his *Formale Logik* (Fribourg, 1956), pp. 531–605 (over 2,000 entries), which was reproduced photographically in the English translation by Ivo Thomas, *A History of Formal Logic* (Notre Dame, Ind., 1961), on pp. 460–534, with English section headings substituted for the German headings and 34 additions to the bibliography given on p. 567.

Alonzo Church's "A Bibliography of Symbolic Logic," in *Journal of Symbolic Logic*, Vol. 1 (1936), 121–218 (about 1,800 entries), which is supplemented by abstracts of books and periodical articles on symbolic logic in each issue of the *Journal of Symbolic Logic*. Vol. 3 (1938), 178–212, contained the section "Additions and Corrections," applicable to the basic bibliography.

William A. Hammond's *A Bibliography of Aesthetics and of the Philosophy of the Fine Arts from 1900 to 1932* (rev. ed., New York, 1934, 205 pages, 2,191 entries), which also has a continuing supplement in the "Selective Current Bibliography for Aesthetics and Related Fields," now published annually in June in the *Journal of Aesthetics and Art Criticism* and originally published quarterly, under the title "Quarterly Bibliography of Aesthetic Theory, Criticism, and Psychology of Art," from the beginning of the issuance of the periodical in 1941.

Ethel M. Albert and Clyde Kluckhohn's *A Selected Bibliography on Values, Ethics, and Esthetics in the Behavioral Sciences and Philosophy, 1920–1958* (Glencoe, Ill., 1959, 342 pages), which contains 600 items in Chapter 6, "Philosophy."

John C. Rule's *Bibliography of Works in the Philosophy of History, 1945–1957* (The Hague, 1961, 87 pages, 1,307 entries), which excludes Marxist interpretations of history in the expectation of covering them separately later.

Amedeo G. Conte's "Bibliografia di logica giuridica (1936–1960)," in *Rivista internazionale di filosofia del diritto*, Vol. 38 (1961), 120–144 (about 250 entries). Addenda appeared in Vol. 39 (1962), 45–46.

For a discussion of some of the bibliographies mentioned in this article, from a librarian's standpoint, see Wilhelm Totok, "Die bibliographische Situation auf dem Gebiet der Philosophie," in *Zeitschrift für Bibliothekswesen und Bibliographie,* Vol. 5 (Frankfurt, 1958), 29–43; and his *Bibliographischer Wegweiser der philosophischen Literatur* (Frankfurt, 1959, 36 pages). See also the section on bibliographies of philosophy in Jean Hoffmans, *La Philosophie et les philosophes; ouvrages généraux* (Brussels, 1920, 395 pages).

William Gerber (1967)

PHILOSOPHY BIBLIOGRAPHIES [ADDENDUM]

The English- and non-English-language citations in this update are combined and are organized chronologically by year of publication. The citations appear within the individual year listings alphabetically by author's last name.

1965

Chan, Wing-tsit. *An Outline and an Annotated Bibliography of Chinese Philosophy.* Supp. New Haven, CT: Far Eastern Publications, 1965.

Higgins, Charles L. *The Bibliography of Philosophy; A Descriptive Account.* Ann Arbor, MI: Campus Publishers, 1965.

1966

Bibliography of Philosophy. Paris: Cultural Center of the French Embassy, 1966.

Jessop, T. E. *A Bibliography of David Hume and of Scottish Philosophy from Francis Hutcheson to Lord Balfour.* New York: Russell & Russell, 1966.

1967

Alston, R. C. *Logic, Philosophy, Epistemology, Universal Language.* Bradford, U.K.: E. Cummins, 1967.

Chan, Wing-tsit. *Chinese Philosophy, 1949–1963; An Annotated Bibliography of Mainland China Publications.* Honolulu: East-West Center Press, 1967.

Draper, John William. *Eighteenth Century English Aesthetics: A Bibliography.* New York: Octagon Books, 1967.

Hammond, William A. *A Bibliography of Aesthetics and of the Philosophy of the Fine Arts from 1900–1932.* Rev. and enl. ed. New York: Russell & Russell, 1967.

McLean, George F. *An Annotated Bibliography of Philosophy in Catholic Thought, 1900–1964.* New York: F. Ungar, 1967.

McLean, George F. *A Bibliography of Christian Philosophy and Contemporary Issues.* New York: F. Ungar, 1967.

1968

Baxandall, Lee. *Marxism and Aesthetics: A Selective Annotation Bibliography; Books and Articles in the English Language.* New York: Humanities Press, 1968.

Hessop, T. E. *A Bibliography of George Berkeley.* New York: B. Franklin, 1968.

Hoffmans, Jean. *La Philosophie et les Philosophes; Ouvrages Généraux.* New York: Burt Franklin, 1968.

Leroux, Emmanuel. *Bibliographie Méthodique du Pragmatisme Américain, Anglais et Italien.* New York: Burt Franklin, 1968.

Robert, Jean Dominique. *Philosophie et Science, Éléments de Bibliographie.* Paris: Beauchesne, 1968.

1969

Bibliography of Philosophy. New York: Cultural Center of the French Embassy, 1969.

Chan, Wing-tsit. *An Outline and an Annotated Bibliography of Chinese Philosophy.* Rev. ed. New Haven, CT: Far Eastern Publications, 1969.

1970

Adickes, Erich. *German Kantian Bibliography.* New York: B. Franklin, 1970 (reprint of the 1893–96 ed.).

Jones, Joe R., and Terry Louis White, eds. *Analytic Philosophy and Religious Language: A Bibliography.* Dallas, TX: Perkins School of Theology, Southern Methodist University, 1970.

Kraav, Marju. *Guide to Research in Philosophy.* Hamilton, ON: McMaster University Library Press, 1970–1979.

Martin, Mary Anne. *The Bibliography of Philosophy; A Guide to Basic Sources,* 1970.

Matczak, Sebastian A. *Philosophy; A Select, Classified Bibliography of Ethics, Economics, Law, Politics, Sociology.* Leuven, Belgium: Editions Nauwelaerts, 1970.

Smith, Marilynn K. *A Bibliography of Philosophy: A Partial List of Holdings in the USMA Library.* West Point, NY: U.S. Military Academy, 1970.

1971

De George, Richard T. *A Guide to Philosophical Bibliography and Research.* New York: Appleton-Century-Crofts, 1971.

1972

Barr, Mary-Margaret H. *A Century of Voltaire Study; A Bibliography of Writings on Voltaire, 1825–1925.* New York: B. Franklin, 1972 (reprint of the 1929 ed.).

Bochenski, Joseph M. *Guide to Marxist Philosophy; An Introductory Bibliography.* Chicago: Swallow Press, 1972.

Redmond, Walter Bernard. *Bibliography of the Philosophy in the Iberian Colonies of America.* The Hague: Nijhoff, 1972.

1973

Jasenas, Michael. *A History of the Bibliography of Philosophy.* Hildesheim: G. Olms, 1973.

Mitcham, Carl, and Robert Mackey, eds. *Bibliography of the Philosophy of Technology.* Chicago: University of Chicago Press, 1973.

1974

Galama, I. J. K., and A. F. Sanders. *Logic, Epistemology, and Analysis of Religious Language: A Select Bibliography.* Groningen, Netherlands: Theologisch Instituut van de Rijksuniversiteit te Groningen, 1974.

Shields, Allan. *A Bibliography of Bibliographies in Aesthetics.* San Diego: San Diego State University Press, 1974.

1975

Bochenski, Joseph M., and Kazunori Kunishima. *Marx Syugi Tetugaku: Kenkyu no hoho to Bunken.* Tokyo: Koronsha, 1975.

Inter-University Board of India. *Philosophy, Religion: A Bibliography of Doctoral Dissertations Accepted by Indian Universities, 1857–1970.* New Delhi: The Board, 1975.

Moss, Michael. *A Bibliography of Logic Books.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1975.

Negley, Glenn Robert. *Utopian Literature: A Bibliography with a Supplementary Listing of Works Significant in Utopian Philosophy.* Durham, NC: Friends of the Library, Duke University, 1975.

Peacocke, Christopher A. B., Dana S. Scott, and Martin K. Davies, eds. *A Selective Bibliography of Philosophical Logic.* 2nd ed. Oxford: Sub-Faculty of Philosophy, University of Oxford, 1975.

Seuren, Pieter A. M. *A Selective Bibliography of Philosophy of Language.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1975.

Zylstra, Bernard. *Bibliography of the Philosophy of the Cosmonomic Idea and Related Materials: English, French, and German Titles.* Toronto: Institute for Christian Studies, 1975.

1976

Lomax, Harvey. *A Contemporary Bibliography in Political Philosophy and in Other Areas.* Dallas: Lomax, 1976.

1977

Baker, John Arthur. *A Select Bibliography of Moral Philosophy.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1977.

Cortright, S. A. *A Classified Bibliography in the Philosophy of Religion.* South Bend, IN: University of Notre Dame, Center for the Study of the Philosophy of Religion, 1977.

Harre, Rom, and John Hawthorn. *A Selective Bibliography of Philosophy of Science.* 2nd ed. Oxford: Sub-Faculty of Philosophy, University of Oxford, 1977.

Lawford, Paul. *Marxist Aesthetics: A Short Bibliography of Works in English, with a Supplement on Russian Formalism, Structuralism, Semiotics.* Keele, Staffordshire, U.K.: Dept. of Sociology and Social Anthropology, University of Keele, 1977.

Lindley, Richard Charles, and J. M. Shorter. *The Philosophy of Mind: A Bibliography.* Part 1, *The Self.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1977.

1978

The Classical World Bibliography of Philosophy, Religion, and Rhetoric. New York: Garland, 1978.

Wainwright, William J. *Philosophy of Religion: An Annotated Bibliography of Twentieth-Century Writings in English.* New York: Garland, 1978.

Weber, Nancy. *Women in Philosophy, Twentieth Century: A Selectively Annotated Bibliography.* Rohnert Park, CA: Sonoma State College, 1978.

1980

McClendon, John H. *Afro-American Philosophers and Philosophy: A Selected Bibliography.* Urbana: University of Illinois, 1980–1985.

Steinhauer, Kurt, and Gitta Hausen. *Hegel Bibliography: Background Material on the International Reception of Hegel within the Context of the History of Philosophy.* Munich: Saur, 1980.

1981

Geldsetzer, Lutz. *Bibliography of the International Congresses of Philosophy: Proceedings, 1900–1978.* Munich: Saur, 1981.

1982

Gabel, Gernot U. *Canadian Theses on German Philosophy, 1925–1975: A Bibliography.* Cologne: Edition Germini, 1982.

1983

Blackwell, Richard J. *A Bibliography of the Philosophy of Science, 1945–1981.* Westport, CT: Greenwood Press, 1983.

Hernandez, Justin. *A Bibliography for the History of Medieval Philosophy.* Conception, MO: Conception Abbey, Seminary College, 1983.

Vance, Mary A. *Architecture-Philosophy: A Bibliography.* Monticello, IL: Vance Bibliographies, 1983.

1984

Nawabi, Mahyar. *Religion, Philosophy & Science.* Tehran: Mu'assasa-i mutala at wa-tahqiqat-i farhangi, 1984.

Vance, Mary A. *Aesthetics: Monographs.* Monticello, IL: Vance Bibliographies, 1984.

1986

Gombocz, Wolfgang L., Norbert Henrichs, and Rudolf Haller. *International Bibliography of Austrian Philosophy, 1974–1975.* Amsterdam: Rodopi, 1986.

1987

Chappell, V. C., and Willis Doney. *Twenty-Five Years of Descartes Scholarship, 1960–1984.* New York: Garland, 1987.

Hurley, S. L., and Jeff McMahan. *A Select Bibliography of Moral and Political Philosophy.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1987.

Swinburne, Richard. *A Selective Bibliography of the Philosophy of Mind.* New ed. Oxford: Sub-Faculty of Philosophy, University of Oxford, 1987.

Warren, Karen J. *Philosophy and Feminism, A Selected Annotated Bibliography.* 1987.

1989

Mathien, Thomas. *Bibliography of Philosophy in Canada: A Research Guide.* Kingston, ON: R. P. Frye, 1989.

1990

Brown, David. *A Selective Bibliography of the Philosophy of Religion.* Oxford: Sub-Faculty of Philosophy, University of Oxford, 1990.

1991

Bell, Albert A., and James B. Allis. *Resources in Ancient Philosophy: An Annotated Bibliography of Scholarship in English, 1965–1989.* Metuchen, NJ: Scarecrow Press, 1991.

1992

Kaylor, Noel Harold. *The Medieval Consolation of Philosophy: An Annotated Bibliography.* New York: Garland, 1992.

1993

Burr, John Roy. *World Philosophy: A Contemporary Bibliography.* Westport, CT: Greenwood Press, 1993.

Chakraborty Ganguly, Krishna. *A Bibliography of Nyaya Philosophy.* Calcutta: National Library, 1993.

1995

Navia, Luis E. *The Philosophy of Cynicism: An Annotated Bibliography.* Westport, CT: Greenwood Press, 1995.

1996

Cogswell, Robert Elzy. *Process Philosophy and Process Theology: An Annotated Bibliography of Introductory Texts.* Austin, TX: Library, Episcopal Theological Seminary of the Southwest, 1996.

Kellerwessel, Wulf. *A Bibliography on Reference and Some Related Topics in Analytical Philosophy.* New York: P. Lang, 1996.

1997

Hutchings, Noel, and William D. Rumsey. *The Collaborative Bibliography of Women in Philosophy.* Bowling Green, OH: Philosophy Documentation Center, Bowling Green State University, 1997.

Radice, Roberto. *Artistotle's Metaphysics: Annotated Bibliography of the Twentieth-Century Literature.* New York: Brill, 1997.

1998

Jones, Ward E., W. Newton-Smith, and Samir Okasha. *A Selective Bibliography of the Philosophy of Science.* New ed. Oxford: Sub-Faculty of Philosophy, University of Oxford, 1998.

Wolf, Robert G. *Analytic Philosophy of Religion: A Bibliography, 1940–1996.* Bowling Green, OH: Philosophy Documentation Center, Bowling Green State University, 1998.

1999

Daiber, Hans. *Bibliography of Islamic Philosophy.* Leiden, The Netherlands: Brill, 1999.

2000

Fieser, James. *A Bibliography of Scottish Common Sense Philosophy.* Bristol, U.K.: Thoemmes Press, 2000.

2001

Bretzke, James T. *Bibliography on East Asian Religion and Philosophy.* Lewiston, NY: E. Mellen Press, 2001.

Meissner, Werner. *Western Philosophy in China, 1993–1997: A Bibliography.* New York: P. Lang, 2001.

Stagaman, David J., James Kraft, and Kristin Sutton. *International Bibliography of Austrian Philosophy, Wittgenstein and Religion: A Bibliography of Articles, Books, and Theses in the Twentieth Century that Relate the Philosophy of Ludwig Wittgenstein to the Study of Religion and Theology.* Quezon City, Philippines: Ateneo de Manila University, 2001.

2003

Internationale Bibliographie zur österreichischen Philosophie 1993/1994. Amsterdam: Rodopi, 2003.

2004

Henrici, Peter. *A Practical Guide to Study; with a Bibliography of Tools of Work for Philosophy and Theology.* Rome: Editrice Pontificia Universita Gregoriana, 2004.

Slater, John G. *Bibliography of Modern British Philosophy.* Bristol, U.K.: Thoemmes Continuum, 2004.

Michael J. Farmer (2005)

PHILOSOPHY DICTIONARIES AND ENCYCLOPEDIAS

Aristotle compiled the first dictionary of philosophy. Other outstanding philosophers who either wrote such works or made slight beginnings in that direction include Avicenna, Leibniz, Voltaire, and Dewey. Kant lectured on *philosophische Enzyklopädie,* but his topic was really the encyclopedic scope of philosophy; Hegel wrote an "encyclopedia" of philosophy which was not an encyclopedia in the ordinary sense. Indeed, what constitutes a dictionary or encyclopedia of philosophy deserves discussion. First, it will be helpful to inspect early examples of such works as well as what might be called embedded dictionaries—the philosophical articles, alphabetically arranged but separated by nonphilosophical material, in general encyclopedias.

In Book V of Aristotle's *Metaphysics* each section consists of a definition and discussion of a philosophical

concept. The various sections begin, for example, "Beginning means . . .," "Cause means . . .," "Element means . . .." He thus covered 29 topics in this first dictionary or quasi dictionary of philosophy: beginning, cause, element, nature, necessity, one, being, substance, sameness and difference, limit, that in virtue of which (or reason why), disposition, priority and posterity, potency, quantum, quality, relation, completeness (or perfection), state, being affected, privation, possession, derivation, part, whole, mutilation, genus, falsity, and accident. The rationale for the order of topics can only be conjectured.

After Aristotle dictionary-type or encyclopedic compendiums were produced by Alexandrian, Roman, and Byzantine lexicographers and doxographers, covering, for the most part, philosophy among other domains of knowledge, not philosophy exclusively. Many of these compendiums were arranged in an order other than alphabetical. Thus, in his *Bibliotheca*, or *Myriobiblion,* Photius (c. 850) summarized, in no special order, some 280 philosophical and nonphilosophical books, including works by Philo Judaeus, Justin Martyr, Origen, and Gregory of Nyssa but none by Plato or Aristotle, although he mentions having read books by Timaeus, Boëthus, and Dorotheus on Plato's use of words.

By contrast Suidas' *Lexicon* (c. 950) is arranged alphabetically. It contains articles on Aristotle (about 150 words), Zeno of Elea (about 75 words), and numerous other philosophers, as well as many topical entries, such as those on *physis, physikos*, and related terms (about nine hundred words in this group). After Suidas, however, through the rest of the medieval period and the Renaissance, most of the summaries of knowledge reverted to the nonalphabetical arrangement.

In modern times the alphabetical arrangement has been dominant in general compendiums of knowledge, and useful philosophical articles have frequently been included in them. It will be instructive, before examining the separately published dictionaries of philosophy, to survey the embedded dictionaries of philosophy.

PHILOSOPHICAL ARTICLES IN GENERAL ENCYCLOPEDIAS

From the standpoint of embedded philosophical material four French, six English, and seven other encyclopedias are especially worthy of comment. In addition, readers may note (*a*) the interest of various prominent philosophers in general encyclopedias, as illustrated by Leibniz' proposal to Louis XIV around 1675 that a group of learned persons "extract the quintessence of the best books, add the unwritten observations of experts, and thus build systems of knowledge based upon experience and demonstrations"; (*b*) the role of the *philosophes* in the work on the *Encyclopédie*; and (c) Giovanni Gentile's role in the Italian encyclopedia of 1929–1939.

FRENCH GENERAL ENCYCLOPEDIAS.

Moreri, Bayle, Diderot, and Larousse are the key figures in the history of French encyclopedias. Of these four Louis Moreri and Pierre Bayle each produced an entire encyclopedia singlehandedly.

Moreri. Moreri's *Le Grand Dictionnaire historique* (1st ed., Lyon, 1674, 1,346 pages; 20th ed., 10 vols., 1759) was translated twice into English and at least once into German, Italian, and Spanish. Reprintings and supplements continued to be published until 1845. By contrast with many dictionaries of philosophy which cover only topics, not individual philosophers, Moreri, in his articles on philosophy, covered many of its practitioners but offered no separate treatments of philosophical domains, problems, schools, or technical terms. Moreover, his articles on the philosophers are so thoroughly oriented toward biography that little attention is paid to doctrines.

Bayle. Bayle's *Dictionnaire historique et critique* (1st ed., 2 vols., Rotterdam, 1697; 5th ed., 5 vols., 1734; annotated ed., 16 vols., Paris, 1820–1824), two editions of which were translated or paraphrased into English, contains some basic facts plus philosophical or critical (usually impish and skeptical) comments for each entry. The comments on both the philosophical and the nonphilosophical topics support atheism, hedonism, and skepticism. As professor of philosophy at Sedan, France, and at Rotterdam, Bayle possessed the necessary technical equipment with which to support his trenchant skepticism. Acknowledging the roar of disapproval which greeted the first edition, Bayle made some revisions in the articles, but the second edition was no less outspoken than the first.

New English translations of selected articles from the *Dictionnaire* were published at Princeton in 1952, edited by E. A. Beller and M. du P. Lee, Jr., and at Indianapolis in 1965, edited by Richard H. Popkin.

"Encyclopédie." The third French general encyclopedia with significant philosophical articles was the one called simply, by common consent, the *Encyclopédie.* Its full title was *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers, par une société de gens de lettres*, edited by Denis Diderot and Jean d'Alembert.

The *Encyclopédie* had a stormy history. It was originally conceived by André F. Le Breton as merely a translation of Ephraim Chambers' *Cyclopaedia* of 1728

(described below), but the character of the project changed, especially after Diderot was put in charge. A corps of contributors was rapidly enlisted which included men of the caliber of Rousseau and Voltaire.

Among the vicissitudes which followed were the periodic banning of the work as irreligious or politically dangerous after the publication of the early volumes and the discouraged resignation of d'Alembert from the project. In 1764, while the manuscript for the final volumes was being edited, Diderot learned to his consternation that Le Breton was toning down the language in order to obviate further prosecution; some of Le Breton's most extensive changes were made in Diderot's own article "Pyrrhonienne ou sceptique philosophie," containing Diderot's most cherished ideas. The original proofs, showing Le Breton's changes and deletions, were discovered in 1933.

The *Encyclopédie* contains no articles on philosophers as such. Among its main articles dealing with philosophical schools or otherwise of philosophical interest are those on Socratic philosophy, Aristotelianism, Epicureanism, and skepticism. The spirit of the philosophical and ethical articles in the *Encyclopédie,* many of which were written by Diderot himself, was antidogmatic, but it was not atheistic or consistently skeptical. Voltaire's 40-odd articles, written in this vein, included 3 in the *E*'s ("Élégance," "Éloquence," and "Esprit"), 21 in the *F*'s ("Félicité," "Finesse," "Fornication," and so on), 11 in the *G*'s ("Goût," "Grandeur," and so on), 5 in the *H*'s ("Heureusement," "Histoire," and so on), "Idolatrie" in the *I*'s, and "Messie" (Messiah) in the *M*'s.

Rousseau wrote the articles on economics (in which he laid the groundwork for his *Contrat social*) and music. Baron de Montesquieu declined the invitation to write on democracy and despotism but promised an article on taste; the portion of it which he had finished before his death in 1755 at the age of 66 was published in Volume VII immediately after Voltaire's article on the same subject.

Eight articles from the *Encyclopédie* on ethical subjects (calumny, unhappiness, and the like) were translated by Ivan Vanslov into Russian and published in 1771 at St. Petersburg by the Imperial Academy of Science as a 21-page dictionary of ethics.

(For a full discussion of the purpose, influence, and philosophic content of the *Encyclopédie,* see the entry **Encyclopédie**.)

Larousse. The excitement aroused by Diderot's original *Encyclopédie* and by the revised editions which followed it eventually subsided, and a calm period in this field ensued. The fourth main French encyclopedia, Larousse's, had its birth in the second half of the nineteenth century. Several encyclopedias bear the name Larousse, beginning with the 15-volume *Grand Dictionnaire universel du XIXe siècle* sponsored by Pierre Larousse (Paris, 1865–1876; 2-vol. supp., 1878–1890) and extending through *Larousse de XXe siècle*, compiled by Paul Augé and published in Paris by the Librairie Larousse (6 vols., 1928–1933; supp., 1953), and the *Grand Larousse encyclopédique*, also published by the Librairie Larousse (10 vols., 1960–1964).

In the *Grand Dictionnaire* the article on philosophy, which covers only the history of philosophy, is curiously followed (perhaps to compensate for the lack of topical discussion) by 51 extensive articles on books with *philosophie* as the first or principal word of the title, such as "Philosophie (Principes de)," by Descartes; "Philosophie morale (Principes de)," by Shaftesbury; "Philosophie premiere, ou Ontologie," by Wolff; "Philosophie de la vie," by Schlegel; and "Philosophie de l'art," by Taine. This is hardly the best way to cover philosophy in an encyclopedia.

The current *Grand Larousse encyclopédique* contains numerous philosophical articles, both topical and biographical, which, although pithy, are excessively brief; for example, Bergson is covered in eight hundred words and logic in nine hundred. The space devoted to the separate articles "Logique (Grande), ouvrage de Friedrich Hegel," "Logique déductive et inductive (Système de), par John Stuart Mill," and "Logique de Port-Royal ou Art de penser" (after the fashion of the nineteenth-century edition) could have been far better used in the article on logic.

ENGLISH GENERAL ENCYCLOPEDIAS. Of the numerous English-language encyclopedias mention may be made of Harris' and the two Chambers', which are mainly of historical interest, and the *Britannica, the Americana*, and *Collier's*, which are influential today.

Harris. The *Lexicon Technicum, or an Universal English Dictionary of Arts and Sciences*, by the clergyman John Harris (1st ed., London, 1704; 5th ed., 2 vols., 1736; supp. vol., 1744), is called by the *Encyclopedia Britannica* the first alphabetical encyclopedia in English, although there seem to be other claimants to this honor. Harris wrote in the Preface, "In Logick, Metaphysicks, Ethicks, Grammar, Rhetorick, &c. I have been designedly short; giving usually the bare meaning only of the Words and Terms of

Art, with one or two instances to explain them, and illustrate them."

The book contains no articles on individual philosophers, and the articles on philosophical topics show a popular rather than a technical understanding (or misunderstanding) of the subject. For example, the article "Logick" (32 lines, mainly laudatory and, curiously, ascribing to logic our ability to explain why we dislike a painting) refers the reader, for details, to the articles "Apprehension" (7 lines), "Discourse" (5 lines defining the term as if it were a synonym of "inference"), "Judgment" (12 lines), and "Method, or Disposition" (40 lines, outlining Descartes's four methodological precepts, with condescending comment) but does not refer the reader to the articles "Conditional Propositions" (8 lines) or "Definition" (19 lines). There is no article on fallacy or syllogism.

Chambers' "Cyclopaedia." A quarter of a century after the appearance of the Harris volume Ephraim Chambers published the *Cyclopaedia, or an Universal Dictionary of Arts and Sciences* (1st ed., 2 vols., London, 1728; 5 other eds., 2 vols., London, 1739–1751/1752, and another 2-vol. ed., Dublin, 1742); supplements were published at various times from 1738 to 1753. Later editions were reportedly used in an unpublished French translation by the writers of the French *Encyclopédie*. Chambers was a freethinker, but many of his articles repeat superstitions and preposterous medical marvels as fact. The *Cyclopaedia* contains succinct articles on essence, ethics, God, knowledge, logic, metaphysics, philosophy, Sophists, truth, and will, as well as on Academic, Cartesian, Epicurean, Platonic, Pyrrhonian, Socratic, and Stoic philosophy, among others. It does not cover individual philosophers.

"Chambers's Encyclopaedia." The so-called *Chambers's Encyclopaedia, a Dictionary of Universal Knowledge for the People* (10 vols., London and Edinburgh, W. & R. Chambers, 1860–1868; rev. eds. issued periodically to 1935) was not a new edition of Ephraim Chambers' *Cyclopaedia* but a new work, written by over one hundred contributors and influenced greatly by the 15-volume tenth edition of the *Conversations-Lexikon* published from 1851 to 1855 by F. A. Brockhaus at Leipzig. The philosophical articles in *Chambers's Encyclopaedia* are uneven. Anaximander is allotted ten times as much space as Anaximenes. The article on the Gnostics is scholarly (although the author wrongly says that they "feigned a naive surprise" at not being accepted as Christians), whereas other articles are more popular in style. The article on Pascal is wholly biographical, but the one on Plotinus covers both his life and his teaching.

The current *Chambers's Encyclopaedia* (15 vols., London, George Newnes, 1950; rev. ed., 1959) is a successor of the 1860–1868 work, not of Ephraim Chambers'. Its advisers on philosophy were John Laird and A. C. Ewing. The articles on Greek philosophy incorporate recent scholarship; the one on Antisthenes, for example, avoids the error, embodied in many earlier treatments, of calling him the first Cynic. It seems odd, however, to find the intellectual work of Mohandas Gandhi and Sarvepalli Radhakrishnan discussed in A. B. Keith's article on Hinduism, which is concerned mainly with the Hindu religion, rather than in S. N. Dasgupta's article on Indian philosophy. The index volume contains a useful classified list of the philosophical articles: 29 on philosophy, metaphysics, and epistemology; 32 on logic; 8 on ethics; 41 on systems and schools; and over 200 on individual thinkers.

"Encyclopaedia Britannica." The last edition of Ephraim Chambers' *Cyclopaedia* was published in the 1750s, and the French *Encyclopédie* had appeared in the 1750s and 1760s. In the middle or late 1760s William Smellie, a printer, historian, and naturalist, wrote most of the articles for a new compendium, the *Encyclopaedia Britannica, or Dictionary of Arts and Sciences* (1st ed., 3 vols., Edinburgh, A. Bell and C. Macfarquhar, 1768–1771). It was issued in installments beginning in December 1768, and subsequent editions, some with supplements, were issued by various publishers. The numbering of the editions was discontinued after the fourteenth edition, which appeared in 1929. The *Britannica* is now published, with continuous revisions, in Chicago by William Benton.

The most famous (and on some topics the most scholarly and comprehensive) edition of the *Britannica* is the eleventh (29 vols., London and New York, 1910–1911). It was sharply attacked by Willard H. Wright (better known by his pseudonym S. S. Van Dine, under which he wrote best-selling murder mysteries) in *Misinforming a Nation* (New York, B. W. Huebsch, 1917), which made several points in Chapter XI, "Philosophy." The Britannica is provincial, he claimed, as in its description of Locke as "typically English in his reverence for facts"; dogmatic, as in the statements that Berkeley "once for all lifted the problem of metaphysics to a higher level" and that Hume "determined the form into which later metaphysical questions have been thrown"; and patronizing, as in the statement that Condillac's thought "was by no means suited to English ways of thinking." Wright also pointed out that the eleventh edition contained no arti-

cles on Bergson, Bradley, Dewey, Royce, or Santayana, and only 1 column on Nietzsche, as compared to 3 on Samuel Clarke, 5 on Spencer, 7 on Fichte, 11 on Cousin, 14 on Hume, 15 on Hegel, 15 on Locke, and 19 on Newton.

Edmund Husserl's article on phenomenology, first published in the 14th edition (1929), was included in the various printings through 1955. It was also reproduced in *Realism and the Background of Phenomenology* (Roderick M. Chisholm, ed., Glencoe, Ill., Free Press, 1961). In subsequent printings of the *Britannica* the article on phenomenology was written, at first, by J. N. Findlay and, currently, by Herbert Spiegelberg.

Many of the philosophical articles in the *Britannica* were rewritten around 1957. Some of the topical articles reflect the current Oxford philosophy. Of the current revision (1966), which for the most part reproduces the recently rewritten articles, the editors and advisers for articles on philosophy are Alonzo Church of Princeton, W. C. Kneale and W. H. Walsh of Oxford, and Sarvepalli Radhakrishnan, president of India. Contributors near the beginning of the alphabet include A. J. Ayer, Max Black, and Brand Blanshard and near the end I. A. Richards, Gilbert Ryle, A. E. Taylor, Wilbur M. Urban, and Abraham Wolf.

In a later revision Thomas E. Jessop is lively as well as scholarly on Hume. The article on Plato, by A. E. Taylor and Philip Merlan, is a comprehensive monograph of the highest value; the bibliography of over 125 items covers manuscripts, editions, commentaries, translations, and analyses. The article on aesthetics, by Thomas Munro, and "Aesthetics, History of," by Helmut Kuhn, which refer to each other, overlap somewhat; for historical data one should consult both. In his article on metaphysics Gilbert Ryle presents a penetrating survey of the status of metaphysics from the origin of the term through the twentieth-century attacks on the discipline; he predicts that the term may "come back into ordinary or pedagogic use" when the motives which generate synoptic world views swing once more into prominence.

"Encyclopedia Americana." Another major English-language encyclopedia, the *Encyclopedia Americana*, edited by Francis Lieber and Edward Wigglesworth (13 vols., Philadelphia, Carey, Lea and Carey, 1829–1833), was originally in large part a translation of the seventh edition (1827–1829) of the *Conversations-Lexikon* published by Brockhaus. Subsequent unnumbered editions, some with supplements, have been issued by various publishers.

The 1996 edition of the *Americana* has Morton G. White of Harvard University as the philosophy member of its editorial advisory board. Among the principal contributors are Brand Blanshard on idea and idealism, Richard B. Brandt on duty and ethics, Herbert Feigl on the Vienna circle, Carl G. Hempel on meaning, Walter Kaufmann on Nietzsche, C. I. Lewis on philosophy, Kingsley Price on fine arts, and Donald C. Williams on conceptualism, free will, innate ideas, mechanism, and pluralism. The article on logic, by Ernest Nagel; "Logic, Symbolic," by W. V. Quine; and the "Logic Glossary," by Arthur Danto, excel in covering a broad range of technical data briefly but comprehensibly. Some of the articles need updating; for example, the death of G. E. Moore, which is mentioned in Volume XIX, has not yet occurred in the article on common sense in Volume VII. The unsigned article on Santayana is philosophically weak.

"Collier's Encyclopedia." *Collier's Encyclopedia* (20 vols., New York, P. F. Collier & Son, 1950–1951; rev. ed., 24 vols., 1962), is published by Crowell Collier and Macmillan. It has T. V. Smith as its adviser on philosophy. Among its American contributors in the field of philosophy are Max Black, Brand Blanshard, George Boas, Roderick M. Chisholm, Raphael Demos, C. J. Ducasse, Marvin Farber, Carl Hempel, Sidney Hook, C. I. Lewis, Ernest Nagel, and Herbert W. Schneider. There are also philosophical articles by such eminent foreigners as T. M. P. Mahadevan and John Passmore. *Collier's* is stronger on the philosophical disciplines than on the schools. It contains first-class articles on aesthetics, by Van Meter Ames; epistemology, by Roderick M. Chisholm; history of ethics, by R. A. Tsanoff; logic, by I. M. Copi; metaphysics, by Blanshard; and philosophy, by a group including Blanshard, Demos, and C. W. Hendel. However, there is no article on realism, the one on naturalism has 1 paragraph, the one on monism 2 paragraphs, and the one on pragmatism 3 paragraphs. Existentialism, however, has 12 paragraphs. The bibliography of philosophy in the final volume lists over four hundred books.

OTHER GENERAL ENCYCLOPEDIAS. Of the numerous other modern encyclopedias, mention may be made of seven—three in German, one each in Italian and Spanish, and two in Russian—which are perhaps the most prominent.

German. The *Grosses vollständiges universal Lexicon,* edited by Johann Heinrich Zedler and Carl G. Ludovici (64 vols., Halle and Leipzig, 1732–1750; reprinted, 1959), was the first encyclopedia compiled on a cooperative basis. The number of its collaborators, nine, was meant to

correspond to the number of the Muses. The articles display an orthodox and partly medieval point of view, acknowledging the existence of the devil and of miracles, accepting astrology ("the influence of the planets must be conceded"), and stressing the scientific contributions of Roger Bacon and Albert the Great.

Der grosse Brockhaus (16th ed., 12 vols., Wiesbaden, F. A. Brockhaus, 1952–1957; supp. vol., 1958) is the current progeny of the Brockhaus-sponsored *Conversations-Lexikon.* It is especially strong on bibliography. The bibliographical sections of some of the philosophical articles, especially those on individual philosophers, constitute one-third or more of the entire text. The bibliographical section of the article on philosophy contains seven subsections, including one on dictionaries, which lists 12 items (9 German, 2 English, and 1 French).

Of the series of encyclopedias begun by Joseph Meyer as *Das grosse Conversations-Lexicon für die gebildeten Stände* ("The Great Encyclopedia for the Educated Classes," 38 vols. in 46, Philadelphia and Hildburghausen, Germany, Bibliographisches Institut, 1840–1853; 6-vol. supp., 1853–1855) the various editions, most of which were published at Leipzig and Vienna, included, for the most part, very creditable articles on philosophers and philosophical topics. The eighth edition, called *Meyers Lexikon* (Leipzig, Bibliographisches Institut, begun 1936; Vol. XII, an atlas, published 1936), was abandoned in 1942 with the ninth volume, covering *R* and *S*. This edition showed decided Nazi influence, using, for example, the exclamation point of sarcasm in noting, in a discussion of Jewish thought, Spinoza's doctrine that God and nature (and substance also, according to the author of the article) are "identisch(!)" and in referring, in the article on Salomon Maimon, "Philosoph, Ostjude," to the baleful influence of his "Ghetto-Intellekt" on Neo-Kantianism.

Italian. Giovanni Gentile was a director and later a vice-president of the organization which produced the *Enciclopedia italiana di scienze, lettere ed arti* (36 vols., Milan and Rome, Istituto Giovanni Trecanni, later the Istituto della Enciclopedia Italiana, 1929–1939; supp. vol., 1938; 2-vol. supp., 1938–1948). The philosophical articles often include special features. For example, the one on Socrates offers a detailed analysis and appraisal of the sources, the one on Aristotle contains a section on medieval legends about Aristotle and Alexander, the one on Bruno discusses *la libertà filosofica*, and the one on *filosofia* (almost 100,000 words) quotes from a large number of writers on the nature of philosophy. Mussolini was the author of the article on fascism.

Spanish. In the Spanish *Enciclopedia universal illustrada europeo-americana* (70 vols., in 72, Bilbao, Spain, Espasa–Calpe, 1905–1930; 10-vol. appendix, 1930–1933; supp., usually biennially) many of the articles on philosophical schools or positions—materialism, utilitarianism, and so on—are usefully divided into two sections, exposition and criticism. In the article on pragmatism, for example, the sections on Anglo-American pragmatism and French pragmatism are each so divided.

Russian. The outstanding encyclopedia of prerevolutionary Russia was the *Entsiklopedichesky Slovar'* ("Encyclopedic Dictionary"), edited by Ivan E. Andreyevsky and others (43 vols. in 86, St. Petersburg and Leipzig, F. A. Brockhaus–I. A. Ephron, 1890–1907). Its philosophy articles were edited by Vladimir S. Solovyov, one of Russia's greatest philosophers, until his death in 1900 and then by Ernest L. Radlov, author of a philosophical dictionary published in 1911 (mentioned below). Solovyov himself wrote the articles on actuality, Campanella, cause, Comte, Duns Scotus, eternity, freedom of the will, Gorgias, Hartmann, Hegel (22 columns), Indian philosophy, Kant, Lully, Maine de Biran, Malebranche, metaphysics, nature, optimism, pessimism, Plato (28 columns), Plotinus, space, time, Vedānta, world process, and others.

The first edition of the *Bol'shaya Sovetskaya Entsiklopediya* ("Great Soviet Encyclopedia") was published in Moscow from 1926 to 1947 in 66 volumes. The second edition, whose chief editor was S. I. Vavilov, was published in Moscow by the Soviet Encyclopedia Publishing House from 1950 to 1958 in 53 volumes. Stalin's death during the course of publication of the second edition led to a change in the tone in the later volumes, where, for example, the cult of personality is rejected. In 1964 *Pravda* announced plans for a third edition.

The philosophical articles in both editions of the "Great Soviet Encyclopedia" are characteristically Marxist in viewpoint. Thus, Rudolf Carnap's philosophy is branded as "a typical example . . . of subjective idealism under the new labels adopted by the ideologists of the imperialist bourgeoisie in the struggle against the scientific materialist world view." In the allocation of space Hegel gets 5 pages, Kant 4, Spinoza 2, Plato 1, and G. E. Moore none; dialectical materialism gets 19 pages, philosophy 17 pages, and pragmatism half a page.

Some of the philosophical articles of the "Great Soviet Encyclopedia" were translated into German and issued in separate brochures (one each on Aristotle, Hegel, Voltaire, and idealism; one covering Bacon, Berkeley, and Bruno; and one covering Helvétius, Heraclitus,

Hobbes, and Holbach) in a series entitled *Grosse Sowjet-Enzyklopädie: Reihe Geschichte und Philosophie* (Berlin, Aufbau-Verlag, 1953–1955).

ENCYCLOPEDIC DICTIONARY. Well deserving of mention is the fact that Charles S. Peirce wrote the definitions of terms in metaphysics, logic, mathematics, and other subjects and Lyman Abbott was responsible for those in theology in *The Century Dictionary; An Encyclopedic Lexicon of the English Language*, edited by William D. Whitney (8 vols., New York, Century, 1891; issued, together with *The Century Encyclopedia of Names* and an atlas, as a 10-vol. work entitled *The Century Dictionary and Encyclopedia* in various years, with revisions, to 1911; issued in condensed form as *The New Century Dictionary*, 2 vols., D. Appleton–Century, 1943 and later years). According to the Preface, "Though it has not been possible to state all the conflicting definitions of different philosophers and schools," nevertheless, ". . . the philosophical wealth of the English language has, it is believed, never been so fully presented in any dictionary." Peirce's fine hand is evident not only in the choice of illustrative quotations but also in the breakdown of terms into subcategories; for example, the article on being includes definitions of actual being, accidental being, being in itself, connotative being, and so on.

SEMIGENERAL ENCYCLOPEDIAS. The following constitute bridges between the dictionaries of philosophy embedded in general encyclopedias and the separate dictionaries of philosophy.

"Cyclopedia of Education." Articles on 114 philosophers or groups of philosophers "whose systems have educational significance" and on 29 "philosophic views bearing on the nature of education" (atomism, determinism, dualism, empiricism, and so forth) appear in *A Cyclopedia of Education*, edited by Paul Monroe (5 vols., New York, Macmillan, 1919). John Dewey was the departmental editor for philosophy of education, and he wrote the articles on determinism, positivism, and many others. Other contributors include John Burnet, Paul Carus, Morris R. Cohen, Arthur O. Lovejoy, I. Woodbridge Riley, Frank Thilly, and Frederick J. E. Woodbridge. Cohen's article on philosophy is one of the best sources for the history of the teaching of philosophy; its bibliography contains 45 painstakingly assembled entries on philosophy in American, British, and Continental colleges, on philosophy in the secondary school, and so on.

"Encyclopaedia of the Social Sciences." The *Encyclopaedia of the Social Sciences*, edited by Edwin R. A. Seligman and Alvin Johnson (15 vols., New York, Macmillan, 1930–1935; reissued in part or in whole in various years), of which John Dewey was the advisory editor for philosophy, had over a dozen philosophers among its editorial consultants, including Morris R. Cohen, Benedetto Croce, Arthur O. Lovejoy, Ralph Barton Perry, Herbert W. Schneider, and T. V. Smith. This encyclopedia contains some extraordinarily illuminating articles on philosophical subjects.

Among the contributors of philosophical articles were George Boas on Berkeley; Léon Brunschvicg on Pascal and on Plato and Platonism; Ernst Cassirer on Kant; Cohen on atheism, belief, Bradley, Descartes, fictions, Hegel, and scientific method; Dewey on human nature, logic, and philosophy; Sidney Hook on Engels, Feuerbach, materialism, and violence; Hu Shih on Confucianism; Horace M. Kallen on behaviorism, James, modernism, morals, pragmatism, and radicalism; Lovejoy on academic freedom; Richard McKeon on Albert the Great, Anselm, Averroës, and Peter Lombard; C. R. Morris on Locke; M. C. Otto on hedonism; J. H. Randall, Jr., on Copernicus and on deism; F. C. S. Schiller on humanism; Herbert Schneider on Christian socialism, ethical culture, and transcendentalism; T. V. Smith on common sense, duty, ethics, and honor.

Dewey's article on human nature (ten columns) sets forth with clarity and force the principal meanings of the term *human nature*, the basic questions which may be asked about human nature, and the history of the understanding of human nature; his 21-item bibliography begins appropriately with his own *Human Nature and Conduct* (1922). Cassirer's article on Kant (eight columns) highlights Kant's significance for social thought and succinctly traces his impact through Fichte, the Hegelians, and the socialists; the 39-item bibliography begins with Cassirer's ten-volume edition of Kant's *Werke.* McKeon's article on Anselm brings out Anselm's little-known contribution to the problem of church–state relations.

"International Encyclopedia of the Social Sciences." The *International Encyclopedia of the Social Sciences*, a completely new encyclopedia, edited by David L. Sills, is a lineal descendant of the *Encyclopaedia of the Social Sciences.* This encyclopedia is devoted primarily to the fields of anthropology, economics, political science, psychology, sociology, and statistics. However, many of its articles are of direct relevance to philosophy; others describe the relevance of philosophical concepts to the social sciences. There are also many biographical articles on philosophers

who have made significant contributions to the social sciences.

EARLY DICTIONARIES OF PHILOSOPHY

What is to count as a dictionary of philosophy and the difference between a dictionary of philosophy and an encyclopedia of philosophy are largely matters of definition. Two definitions seem most useful for the present purpose. First, a dictionary of philosophy is an expository work setting forth information about philosophical ideas in an arrangement which either is alphabetical (as in the embedded dictionaries of philosophy already mentioned and in most of those described below) or is based on key words or concepts (as in Aristotle's "dictionary" mentioned above and the first few of those mentioned below) rather than on a systematic division of philosophy into its disciplines or parts. Second, an encyclopedia of philosophy is a comprehensive dictionary of philosophy in which various articles are monographic in scope.

Dictionaries of philosophy range from those which are purely factual through those which are partly interpretive or evaluative to those, such as Voltaire's, which present rhapsodic or satirical reflections on key general topics. Divergences from this broad range of varieties also occur—for example, a "dictionary" which merely lists philosophical terms in one language with equivalents in other languages, a "dictionary" which presents for each important philosophical term a suggested usage rather than a statement of actual usage, and an anthology of philosophical quotations arranged alphabetically by topic. Over one hundred dictionaries of philosophy of one sort or another have been published. Most have been soon forgotten, but some have gone through multiple editions over many decades.

MIDDLE AGES. Of the medieval works which may be counted as dictionaries of philosophy perhaps those of Isaac Israeli and Avicenna are most worthy of note. Israeli (c. 855–c. 955), the first Jewish Neoplatonist, wrote, in Arabic, *Kitāb al-Hudūd wal Rusum* ("Book on Definitions and Descriptions"), later translated into Latin and Hebrew. This work contains definitions, with comments thereon, of topics grouped roughly as intellect, soul, vital spirit, and so on; reason, knowledge, opinion, memory, deliberation, and so on; division, syllogism, demonstration, truth, falsehood, necessary, impossible, and so on; imagination, estimation, and sense perception; love, passion, and desire; innovation, creation, coming to be, passing away, and so on; time, eternity, and perpetuity; and other topics. The influence of al-Kindī in some 20 of the 56 sections has been noted by the latest editors of Israeli's work, A. Altmann and S. M. Stern.

Avicenna's *Kitāb al-Hudūd* ("Epistle on Definitions") contains, after an introduction on the pitfalls of the process of defining, definitions—extracted in part from Avicenna's other works—of accident, body, cause, continuous, creation, definition, form, individuality, intelligence, limit, motion, nonbeing, place, prime matter, priority, rest, soul, substance, time, universe, and other subjects. Terms having obvious mutual relations are grouped. The definitions are close to Aristotelianism in tenor.

An anonymous *Compendium Philosophiae*, based mainly on Aristotle and Albert the Great, written (probably in France) about 1327 and as yet only partly edited and published (Paris, 1936), was one of the last medieval dictionaries of philosophy. In topical groups it contains, in Books I to V, brief discussions of God, the physical features of the world, plants, animals, and man and, in Books VI to VIII, scholastic-type discussions of accident, actuality, art, becoming, being and nonbeing, cause, fate, free will, identity, language, law, motion, names, necessity, perfection, philosophy, place, potentiality, quality, quantity, relation, science, substance, time, truth and falsity, virtue and vice, and wisdom, as well as other subjects.

SIXTEENTH CENTURY. After the revival of learning and the invention of printing there appeared a number of compendiums of philosophical information. Apparently, the first formal dictionary of modern times devoted exclusively to philosophy was Giovanni Baptista Bernardo's *Seminarium Totius Philosophiae* (3 vols., Venice, Damian Zenarius, 1582–1585; 2nd ed., 3 vols. in 2, Geneva, Jacob Stoer and Franc. Faber, 1599–1605), later referred to as the *Lexicon Triplex.* In separate alphabetical dictionaries the three volumes cover, respectively, Aristotelian, Platonic, and Stoic philosophy in the writings not only of Aristotle, Plato, and the early Stoics but also of other philosophers, Greek, Roman, Christian, and Arabic. Thus, the first volume contains articles on Aristotelian philosophy from "Abstractio," "Accidens," "Actus," and other topics in the A's to "Zeleucus," "Zephirus," "Zodiacus," and other topics in the Z's.

The article "Definitio" in Volume I contains 333 paragraphs summarizing or quoting specific passages on the subject in Aristotle, Ammonius, Alexander of Aphrodisias, Themistius, Simplicius, Boethius, Averroës, Alexander of Hales, Albert the Great, Thomas Aquinas, and others. A similar approach—abstracting specific pas-

sages—is used throughout the three volumes. Accordingly, the work is essentially useful as a thorough guide to the sources but not as a synthesis.

SEVENTEENTH CENTURY. The seventeenth century provided nine principal dictionaries of philosophy, all in Latin.

1610—Nicolaus Burchardi. Buchardi's *Repertorium Philosophicum, Quo Omnes in Universa Philosophia Subinde Occurrunt Termini Perspicue Traduntur* (Leipzig, 573 pages) appeared in 1610. It was also issued at Grimma in 1613 and at Gera in 1614, 1615, and 1616. Only two copies of this work are known to exist, having been located, after many fruitless searches elsewhere, in the Universitätsbibliothek in Marburg, West Germany (the 1614 printing), and in the Sächsische Landesbibliothek in Dresden, East Germany (the 1616 printing). A microfilm copy of the 1616 printing was procured and is filed in the Public Library of the District of Columbia.

The main part of the book is not arranged alphabetically. It treats exactly one hundred topics, from philosophy, logic, metaphysics, art, nature, and word, near the beginning, to infinite soul, theology, and God, at the end. The articles are superficial in their analysis but reflect wide reading in the classic sources. An alphabetical index of topics (*abstractum, ars,* and so on) appears at the beginning, and the book ends with an alphabetically arranged index of themes discussed in the articles—for example, *abstracta saepe ponuntur pro concretis* ("the abstract is often substituted for the concrete") and *amicitia honesta cur rara* ("why true friendship is rare").

1612—Henri Louis Chasteigner. Chasteigner's *Celebriorum Distinctionum turn Philosophicarum tum Theologicarum Synopsis* (Poitiers, A. Mesner, 1612, 71 pages; subsequent eds. or reprints, various places, 1616, 1617, 1619, 1623, 1635, 1645, 1651, 1653, 1657, 1658, 1659, and 1667) made a beginning in the provision of syntheses that Bernardo's work lacked. Thus, absolute is explained as in one sense opposite to relative; in another, to dependent; and in still another, to restricted. Abstraction is broken down into real (when the thing abstracted can exist separately) and rational; rational abstraction, into negative (or divisive) and precise (or simple); and precise abstraction, finally, into physical, mathematical, and so on. A prefatory alphabetical list names 48 authors—Alexander of Hales, Aristotle, Bonaventure, Buridan, Duns Scotus, Suárez, Thomas Aquinas, William of Ockham, and so on—whose writings were chiefly used in compiling the work.

1613—Rudolf Goclenius the elder. The *Lexicon Philosophicum* (Frankfurt, Matthias Becker, 1613, 1,143 pages; additional printings or eds., Marburg, 1613, 1615 and Frankfurt, 1633, 1634; Frankfurt 1613 ed. reissued in facsimile, Hildesheim, 1964) opens with four tributes to Goclenius (Rudolf Goeckel) in Latin verse. There follow articles on terms beginning with the vowels—*absolutum, existentia, idea, obligatio, unitas,* and the like—and then articles on terms beginning with the consonants—*beatitudo, causa,* and so on. The articles are informative, presenting standard scholastic breakdowns and definitions. As has been noted by José Ferrater Mora, Goclenius, although he was the first to use the term *ontologia* (in Greek letters), did not make significant use of the term. Goclenius is cited for support on a particular point in a work as late as Eisler's *Wörterbuch der Philosophie* (1899).

1626—Johann H. Alsted. Alsted's *Compendium Lexici Philosophici* (Herborn, Germany, Georg Corvin and J. G. Muderspach, 1626, 720 pages) is a group of dictionaries on about thirty separate disciplines—anatomy, arithmetic, astronomy, and so on in nonalphabetical order—including ten on philosophy covering ethics, logic, metaphysics, philosophical "archelogy" (basic terms), philosophical didactics (teaching of philosophy), philosophical "hexilogy" (mental faculties involved in philosophy), philosophical method, pneumatics (study of spiritual beings), poetics, and politics. Some parts of the dictionaries are alphabetical; others are not. Of the ten philosophical dictionaries, the one on logic, which is 26 pages long, is perhaps the best, but most of the material in it is not arranged alphabetically and is therefore difficult to follow.

1629—George Reeb, S.J. Reeb's *Distinctiones Philosophicae* (Ingolstadt, Germany, Gregory Haenlin, 1629, 167 pages; 2nd ed., Cologne, 1630) was reprinted in 1653, 1657, and 1658 in the same volume as Chasteigner's *Synopsis*. With Reeb's *Axiomata Philosophica* it was reissued under the editorship of J. M. Cornoldi, S.J. (Bressanone, Italy, 1871 and Paris, 1873, 1875, 1891), under the title *Thesaurus Philosophorum, seu Distinctiones et Axiomata Philosophica*. Reeb's work, written from a scholastic viewpoint, discusses, as philosophical topics, such adverbial opposites as absolutely and dependently, in act and in potency, artificially and naturally, collectively and distributively, concretely and abstractly, and so on.

1653—Johann Micraelius. The *Lexicon Philosophicum Terminorum Philosophis Usitatorum* of Johann Micraelius (Jena, Jeremiah Mamphras, 1653, 667 pages; 2nd printing, 1662) contains explanations of the terms used in philosophy, broadly understood; a 51-page

appended outline, by discipline, of the topics covered; a 30-page index of Greek terms; and 17 pages of illustrations, mostly geometric figures. Many articles begin with what Aristotle said on the subject and continue with the scholastic elaborations of what Aristotle said. The article "Deus," however, begins by saying flatly that Aristotle was right in calling God the prime mover but was wrong in denying God's creation of the world, God's omniscience, and so on.

1658—Johann Adam Scherzer (Schertzer) and others. Scherzer and others' *Vade Mecum, Sive Manuale Philosophicum Quadripartitum* (Leipzig, Christian Kirchner, 1658) has four parts separately paged but bound as one volume. Part I, by Scherzer, entitled *Definitiones Philosophicae*, is a scholastic-type alphabetical dictionary, with definitions, for example, under *necessarium*, of absolute, hypothetical, physical, moral, and logical kinds of necessary thing. Part II consists of Chasteigner's *Synopsis* and Reeb's *Distinctiones Philosophicae.* Part III, by Scherzer, entitled *Axiomata Resoluta*, presents a system of rules of thought (a thing cannot be and not be, a proposition must be true or false, and so on). Part IV, by Scherzer, entitled *Aurifodina Distinctionum* ("Gold Mine of Distinctions"), discusses selected distinctions in an alphabetical arrangement (for example, intrinsic and extrinsic accidents among the A's and remote and proximate cause among the C's). Scherzer's project was ambitious, but the resulting complex was too cumbersome for convenience.

1675—Heinrich Volckmar. The *Dictionarium Philosophicum, Hoc Est Enodatio Terminorum ac Distinctionum* of Heinrich Volckmar (Frankfurt, Jacob Gottfried Seyler, 1675, 798 pages; 2nd printing, 1676) is in Latin, but the author sprinkles a little German here and there. Thus, in citing the tenet "Credo quod Deus creavit me" under "Creatio," he translated it (as if it were difficult Latin) "Ich gläube dass mich Gott geschaffen hat." In an epilogue he asked the reader to ascribe any omissions not to negligence but to the enormity of the field to be covered, and he named as predecessors Chasteigner, Goclenius, Reeb, Micraelius, and Scherzer but not Alsted.

1692—Étienne Chauvin. In Chauvin's *Lexicon Rationale, Sive Thesaurus Philosophicus Ordine Alphabetico Digestus* (Rotterdam, P. van der Slaast, 1692, 756 pages; 2nd ed., entitled *Lexicon Philosophicum*, Leeuwarden, Netherlands, Franciscus Halma, 1713, 719 pages) philosophy includes natural science. Thus, there are articles, in their Latin equivalents, on acceleration, fire, meteors, and the stomach, as well as on Aristotle, Descartes (a particularly laudatory article), other philosophers, and *cognitio, simplicitas, subsistentia*, and other philosophical concepts. Cartesian influence is apparent in many of the articles.

"Nondictionaries." Mention may also be made of an unalphabetical "lexicon" of this period by Pierre Godart. The second edition of his *Totius Philosophiae Summa* (Paris, L. Billaine, 1666, 245 pages) was entitled *Lexicon Philosophicum* (2 vols. in 1, Paris, J. and R. I. B. de La Caille, 1675) although it was not really a dictionary. After an introduction on philosophy and its divisions, the philosophical schools, and some principles of philosophy the book discusses being, causes, properties, and species; physics, including matter, motion, soul, sensation, and so on, with an attack on Cartesian philosophy; economics and politics; and logic. The alphabetical index in the second edition is 47 pages long.

Wolter Schopen's *Alphabetum Philosophicum* (Nissa, John Joseph Krembsl', 1696, 105 pages), although sometimes referred to as a dictionary of philosophy, is, like Godart's *Lexicon*, not a dictionary. It is a straight exposition of twenty-odd philosophical topics, such as what a definition is, what conversion and opposition of propositions are, and how many kinds of syllogism there are, each topic being designated by a letter of the alphabet (*A, B, C,* and so on).

LEIBNIZ AND AFTER. Among the fragments of Leibniz edited in 1903 by Louis Couturat and assigned to the period 1670–1704 are two which show an interest in the Alsted work mentioned above and several which consist of lists of definitions of terms, as if Leibniz were thinking of compiling a dictionary of philosophy apart from the general encyclopedia which he had discussed with Louis XIV. One of these lists of definitions, for example, is headed "Introductio ad Encyclopaediam Arcanam." It contains definitions of *conceptus clarus, conceptus distinctus, conceptus adaequatus, conceptus primitivus,* and the like. Another, untitled, contains definitions of *amor* (love), *sapientia* (wisdom), *laetitia* (joy), *perfectio* (perfection), and so on. Illustrative are his definitions of love, the emotion by which it happens that the good or evil of another is considered part of our own, and of wisdom, the science of happiness. If Leibniz had completed a dictionary of philosophy along these lines, it would probably have constituted a vade mecum to his own philosophy rather than an exposition of historical viewpoints in philosophy.

In 1716, the year of Leibniz' death, there appeared the last Latin dictionary of philosophy before the first modern-language dictionaries. It was the *Lexicon Philo-*

sophicum; Sive Index Latinorum Verborum Descriptionumque ad Philosophos & Dialecticos Maxime Pertinentium (The Hague, Henri du Sauzet, 322 pages), of which the author is listed on the title page as Plexiacus ("Auctore Plexiaco"). Plexiacus has been identified as Charles Du-Plessis d'Argentre or Michèle Toussaint Chrétien Duplessis (or du Plessis), but the best scholars attribute the work to one Michel Brochard. Following an extended systematic treatment of argumentation, definition, words and things, distinctions, and so on, the author presents, in an alphabetical arrangement, numerous philosophical terms and their definitions. The systematic treatment in the first part of the book, which leans heavily on the writings of Cicero, is more interesting than the somewhat routine definitions in the lexicon proper.

A Latin quasi dictionary of philosophy that may deserve mention here is the book *Philosophia Definitiva, Hoc Est Definitiones Philosophicae*, by Frederick Christian Baumeister (Wittenberg, Germany, J. J. Ahlfeld, 1738, 252 pages; 7th ed., 1746; enlarged ed., 1767), which contains definitions, grouped according to subject, of 329 logical terms, 233 terms in ontology, 95 terms in cosmology, 264 in psychology, 53 in natural theology, 182 in ethics, 69 in political philosophy, and 35 in physics, with a consolidated alphabetical index. The definitions, based in large part on the philosophy of Christian Wolff, are useful but not profound.

FIRST MODERN-LANGUAGE DICTIONARIES

In 1715 there appeared a work by J. H. (Johann Hübner) entitled *Compendieuses Lexicon Philosophicum* (Frankfurt and Leipzig, B. P. C. Monath, 208 pages; 2nd ed., 1717). The title of the second edition, varying slightly from that of the first, was *Compendieuses Lexicon Metaphysicum, zum besondern Nutzen aller Studierenden, vornemlich abet der politischen Wissenschaften befliessenen zusammen getragen* ("Compendious Metaphysical Lexicon, for Special Uses by All Students, but Chiefly Those Specializing in Political Sciences Taken as a Whole"). Although the work is in German, it discusses only Latin philosophical terms in nonalphabetical order. It begins with *ens* (a being) and among other things points out, with German examples, the distinctions among *ens, res* (a thing), and *reale* (a real thing). Other terms discussed include *verum* and *bonum* (true and real), *ubi* and *quando* (where and when), and the four causes. An alphabetical index at the end contains over four hundred entries, including about fifty under *causa—efficiens, in sensu juridico, necessaria, proxima*, and so on. The treatment is elementary, the analyses are not sharp, and the work has only historical interest today.

The first alphabetically arranged dictionary of philosophy in a modern European language appears to be Hubert Gautier's *La Bibliothèque des philosophes, et des sçavans, tant anciens que modernes* (2 vols., Paris, André Cailleau, 1723). Chauvin had treated philosophy as including the natural sciences; Gautier treated it as including the natural sciences and the humanities. Thus, his book contains articles on Alexander the Great, Copernicus, and La Fontaine, as well as on Avicenna, Descartes, Porphyry, and many others, plus a smaller number of topical articles, such as those on the Académie Royale des Sciences, *homme* (man), and *terre* (earth). Each volume has a topical index. Today, the work has interest mainly as a curiosity rather than for the information it provides.

Strictly speaking, the first dictionary of philosophy in a modern language appears to be the *Philosophisches Lexikon*, by Johann Georg Walch (Leipzig, 1726, 3,048 cols.; 2nd ed., 1733; 3rd ed., 1740; 4th ed., 2 vols., 1775), which set a new standard of comprehensiveness and scholarship for works of this kind. It reflects in part the ideas of Leibniz and Wolff, quoting or citing them in various articles as authorities. Among the more intriguing articles in this *Lexikon* are those on atheism (16 cols.), discussing arguments derived from the existence of evil, the eternity of the world, the sufficiency of nature as an explanation of events, the anthropomorphic character of our idea of God, and so on; self-knowledge, knowledge of others, knowledge of nature, and knowledge of God; fate, with summaries of the views of Parmenides, Democritus, Plato, Aristotle, the Stoics, the Epicureans, the Chaldeans and other Oriental peoples, Sextus Empiricus, Leibniz, and others; and freedom of thought (25 cols.), discussing the ipse dixit principle, freedom of interpretation, freedom of belief, the role of reason, the fate of Spinoza, the right to know the truth, and other aspects of the topic.

An appendix covers philosophers from Abelard, Albinus, and others at the beginning to the two Zenos and Zoroaster at the end. These biographical sketches are of decidedly less interest than the vivid expositions in the topical articles. Many of the biographical sketches begin, repetitiously, ". . . one of the most famous philosophers of" such-and-such a country.

In 1963 the Stuttgart firm of Friedrich Frommann Verlag was planning to issue a facsimile reprint of the fourth edition of Walch's *Philosophisches Lexikon.*

Walch's work was followed by one which originated the exact title used shortly thereafter by Voltaire. This was

the *Dictionnaire philosophique portatif, ou Introduction à la connoissance de l'homme*, by Didier Pierre Chicaneau de Neuville (London, J. M. Bruyset, 1751, 381 pages; 2nd ed., Lyon, J. M. Bruyset, 1756; Italian translation of 2nd ed., Venice, 1756; 3rd ed., Paris, 1764). In de Neuville's pioneering French philosophical lexicon many of the articles are, or begin with, dictionary-type definitions, but the further explanatory material (including quotations from Boileau, Pope, Rousseau, and the early writings of Voltaire) is sometimes piquant.

VOLTAIRE AND AFTER

On September 28, 1752, Voltaire and other intellectual companions of Frederick the Great were dining with the king at Potsdam. Someone, perhaps Frederick himself, mentioned the idea of producing a philosophical dictionary on which men of letters, including Frederick, would collaborate. Voltaire began work on the project the next day and soon showed the article "Abraham" to Frederick, who considered it good and asked Voltaire to set up a list of proposed articles for the work. Voltaire instead quickly produced articles on *âme* (soul), *athéisme, baptême*, and so on, and Frederick commented that the whole book would soon be finished. Voltaire, however, interrupted the project some months later, when he left Potsdam following his break with Frederick, and he presently became involved in preparing articles for Diderot's *Encyclopédie.*

Early in 1760 Voltaire resumed work on his own dictionary. He wrote to the marquise du Deffant on February 18, "I am absorbed in rendering an alphabetical account to myself of everything that I think about this world and the other, entirely for my own use, but (perhaps after my death) for that of honest people."

In the summer of 1764 the *Dictionnaire philosophique portatif*, which was 344 pages long and contained 73 articles, was printed anonymously at Geneva with London given as the place of publication. There was a second printing later in the year. The book was banned by the Parlement of Paris on March 19, 1765, and was placed on the Index of prohibited books by the pope on July 8, 1765. Voltaire denied authorship of the book in 68 letters between July 1764 and February 1768.

A second edition was published at London in 1765 in four printings (varying from 308 to 364 pages), with eight additional articles. Three of these four printings were subsequently counted as the second, third, and fourth editions. A printing which was counted as the fifth edition was issued at Amsterdam in 1765 in two volumes with 15 additional articles. An edition specifically labeled "Sixième Édition," with 34 additional articles, was published at London in 1767 in two fascicles bound as one volume. Another edition, also called the "Sixième Édition," was printed at Geneva in 1769 under the title *La Raison par alphabet*, with further additions, in two volumes.

Subsequent editions continued to appear both during and after Voltaire's lifetime under various titles, sometimes including the articles prepared by Voltaire for the *Encyclopédie*; Voltaire's *Questions sur l'Encyclopédie*, an alphabetically arranged set of comments; *L'Opinion par alphabet*, a manuscript found after Voltaire's death; or a combination of the foregoing. One of the most useful editions was edited by Julien Benda (2 vols., Paris, Garnier Frères, 1936). Of the English versions, complete or abridged, the first appeared in 1765; a noteworthy successor appeared in 1824 (6 vols., London, J. and H. L. Hunt), comprising about three-fourths of the original, the remainder being, according to the anonymous translator, repetitive. In 1901 an "unabridged and unexpurgated" edition was translated by William F. Fleming (10 vols., London, E. R. DuMont): the latest edition, translated by Peter Gay with a preface by André Maurois, was published in 1962 (2 vols., New York, Basic Books).

Voltaire's dictionary covers primarily topics, almost totally excluding individual philosophers; among the few philosophers accorded separate treatment are Arius and "Julien le Philosophe." The topical articles are largely in the nature of discursive essays, occasionally in dialogue form, rather than directly informative expositions, but they nevertheless reflect extensive research and critical analysis. In the article on miracles Voltaire made such points as the following: if a miracle is an event to be marveled at, then everything is a miracle; if a miracle is a violation of an eternal (inviolable) law, then it is a contradiction in terms; it is a strange God who is so incapable of achieving his purposes through his own laws of nature that he must resort to changing his own "eternal" ways.

The topics covered are, for the most part, in the field of popular philosophy or religious controversy, such as Adam, apocalypse, *tout est bien* (all is good), confession, *enfer* (hell), inquisition, and so on. A few touch on technical philosophy; examples are those on *âme* (soul), *beauté* (beauty), *chain des êtres créés* ("great chain of being"), *destin* (fate), and *nécessaire* (necessary). Of the articles in his own dictionary Voltaire submitted only the one on idolatry intact to Diderot for inclusion in the *Encyclopédie*. It was reprinted there without change.

Various literary scholars have studied the sources of Voltaire's dictionary. Although Voltaire acknowledged his

indebtedness to Bayle's *Dictionnaire historique et critique* and the title of his dictionary is identical with that of the one de Neuville published in 1751, it appears that he owed more to the English deists and the early French deists. As André Maurois has observed, the ideas in Voltaire's dictionary "were clichés in its epoch. Gassendi, Fontenelle, Bayle, had said all that." But the form in which the ideas are clothed in Voltaire's dictionary is inimitably adroit, vivid, chatty, anecdotal, and essentially consistent in its rough humaneness and urbanity though inconsistent in details.

REACTION TO VOLTAIRE. Reacting with indignation to the religious skepticism of the *Dictionnaire philosophique portatif* and without knowing that Voltaire was the author of the work, Louis M. Chaudon published—also anonymously—the *Dictionnaire anti-philosophique, pour servir de commentaire & de correctif au Dictionnaire philosophique & aux autres livres qui ont paru de nos jours contre le christianisme* (Avignon, 1767, 451 pages; 4th ed., 2 vols. in 1, Avignon, La Veuve Girard, 1775). Among the approximately 150 articles in the first edition are those on soul, atheism, Bayle, *Encyclopédie*, faith, hell, miracles, natural law, and reason; new articles in subsequent editions include those on deists, Spinoza, suicide, theater; and tyrannicide. Some of the articles are in two sections, presenting the orthodox view of the subject and replying to the skeptics' objections. After the alphabetical part of the work is a summary headed "Résultat des réflexions répandues dans ce *Dictionnaire*." Chaudon's defense of religion in general and of Christianity in particular was spirited and literate.

OTHER EIGHTEENTH-CENTURY DICTIONARIES. Between Voltaire and Chaudon and the end of the eighteenth century six dictionaries of philosophy appeared—three in French, two in German, and one in English—plus a number of works which have promising titles but are not dictionaries of philosophy.

French. In 1772, eight years after the first appearance of Voltaire's dictionary, a work comparable in outline, *La Petite Encyclopédie, ou Dictionnaire des philosophes*, by Abraham J. de Chaumeix, a Frenchman, was published anonymously and posthumously (Antwerp, Jean Gasbeck, 136 pages). It contains only topical articles, none on philosophers, and the articles are popular rather than strictly philosophical in tenor. The motto at the end of the book is a misquotation from Virgil, "Heu! Ubi prisca fides?" ("Alas! Where now is your former faith?").

The other two of the three French dictionaries were parts of a 166-volume rearrangement, by disciplines, of the material in the Diderot *Encyclopédie*. The rearrangement, entitled *Encyclopédie méthodique* (Paris, C. J. Panckoucke and others, 1782–1832), consisted of about fifty separate dictionaries. One of these was *Logique, métaphysique et morale*, edited by Pierre L. Lacratelle (4 vols., 1786–1791). The Lacratelle work started out to cover only logic and metaphysics, and a complete alphabetical arrangement of topics in those two disciplines was presented, from absolute (in logic, 2 cols.) and abstraction (19 cols.) at the beginning of Volume I to sensation (230 cols.) and systems (41 cols.) near the end of Volume II; however, the scope was then changed to include ethics, and the remainder of Volume II and Volumes III and IV contain an alphabetical series of articles on ethics. Volume III was the first volume to include ethics on the title page.

Immediately adjacent to the Lacratelle work in the *Encyclopédie méthodique* is *Philosophie, ancienne et moderne*, edited by Jacques A. Naigeon, an atheist who considered himself Diderot's successor (3 vols., 1791–1793). The topics range from Academics (352 cols.) and Academy (2 cols.) in Volume I to Zend-Avesta (10 cols., Diderot's article on the subject transplanted intact from the *Encyclopédie*) in Volume III. The third volume also contains, on pages 767–945, articles omitted from the first two volumes.

German. Various giants in the history of philosophy—Aristotle, Leibniz, and Voltaire—have thus far entered this record as contributors to the development of dictionaries of philosophy. Another giant—Kant—enters the record by a quirk of terminology. Kant lectured on the subject *philosophische Enzyklopädie* ten times from 1767/1768 to 1781/1782 and advertised lectures on this subject for 1785/1786 and 1787, although these did not materialize. A set of his lecture notes on *philosophische Enzyklopädie*, probably dating from the winter semester of 1781/1782, was edited by the Deutsche Akademie der Wissenschaften zu Berlin and published for the first time in 1961 in East Berlin. But the work actually deals with what might suitably be called philosophy as an encyclopedic discipline rather than philosophy expounded in encyclopedic form. It presents a structured (not alphabetical) outline of philosophy in its broadest ramifications, based on J. H. Feder's *Grundrisz der philosophischen Wissenschaften* ("Foundation of the Philosophical Sciences," Coburg, Germany, J. C. Findeisen, 1767; 2nd ed., 1769).

Thus, Kant did not write a dictionary of philosophy. However, his admirer Salomon Maimon did. Maimon was the author of *Philosophisches Wörterbuch, oder Beleuchtung der wichtigsten Gegenstände der Philosophie, in alphabetischer Ordnung* ("Philosophical Dictionary, or Illumination of the Most Important Themes of Philosophy, in Alphabetical Order," Berlin, Johann F. Unger, 1791, 222 pages). This work is an impressionistic presentation of various philosophical topics, in substance less iconoclastic than Voltaire's dictionary but just as unconventional stylistically. One of the articles, for example, includes separate vehement apostrophes, each beginning "Meine Herren!," to "die Dogmatiker oder Antikantianer" ("dogmatic philosophers or anti-Kantians") and "die kritischen Skeptiker oder Kantianer" ("critical skeptics or Kantians").

Another German dictionary of philosophy in this period, also impressionistic, was Carl Ludwig Friedrich Rabe's *Gedanken und Urtheile über philosophische, moralische und politische Gegenstände, aus guten Schriften gezogen, alphabetisch geordnet* ("Thoughts and Judgments on Philosophical, Ethical and Political Themes, Deduced From Reliable Publications, Alphabetically Arranged," Stendal, Germany, D. C. Franzen and J. C. Grosse, 1789–1790, 2 vols.). This work is even rarer than the Burchardi book of 1610. The Royal Library at Copenhagen possesses what may be the sole extant copy of it, located after the trail had run dry in many other directions. A microfilm of the Copenhagen copy is now available at the Public Library of the District of Columbia.

Volume I of Rabe's *Gedanken* contains reflections on topics with initial letters from A to Z, and Volume II likewise begins at the beginning of the alphabet and goes through to Z. Among the topics discussed are antiquity, art, business, culture, death, despotism, freedom of the press, God, guilt, happiness, language, man, religion, republic, *Schmerz,* science, soul, and time. Articles range from one or two lines to three or four pages in length. The one on freedom of the press reads, in translation, "Without freedom of the press, the soul is crippled. Freedom to think, without freedom to say, is no better than being in a straitjacket." The article on *Held* (hero) reads, "Ein Held wird nicht geformt, er wird geboren" ("A hero is born, not made").

Toward the end of the eighteenth century, there appeared two other documents like Kant's with titles that sound relevant to the story of dictionaries of philosophy but which turn out to have no relevance to the subject. The first of these was Johann Georg Büsch's *Encyclopädie der historischen, philosophischen und mathematischen Wissenschaften* (2 vols. in 1, Hamburg, Heroldsche Buchhandlung, 1775), which presents its material not in an alphabetical but in a systematic arrangement, Volume I covering history and philosophy and Volume II mathematics. The section on philosophy stresses the contributions of Descartes and Wolff and discusses philosophy in general, logic, theology, philosophical psychology, ethics, politics, economics, and related topics.

The second was the *Encyclopädische Einleitung in das Studium der Philosophie,* by Karl Heinrich Heydenricks (Leipzig, Weygandsche Buchhandlung, 1793, 249 pages), which is a systematic, nonalphabetical exposition of the nature of philosophy, systems of philosophy, the bearing of philosophy on other disciplines and on life, and the way to study philosophy.

English. In 1786, *The Philosophical Dictionary, or The Opinions of Modern Philosophers on Metaphysical, Moral, and Political Subjects*, by François Xavier Swediaur, was published (4 vols., London, G. G. J. and J. Robinson, 1786) with "F. S******r, M.D." at the end of the Preface as the only indication of the author or compiler. Many of the articles bear at the end the name of an author (Gibbon, Helvétius, Hume, Locke, Rousseau, Voltaire, and others) from whose writings the article was adapted. Swediaur did not show much understanding of or sympathy for technical philosophy. His article "Ancient Greek Philosophy" mentions Hesiod and Theognis but not Socrates, Plato, or Aristotle.

NINETEENTH CENTURY

Dictionaries of philosophy, or works purporting to be such, appeared in German, English, French, Italian, Latin, and Russian in the nineteenth century.

GERMAN. Initiating the contributions of the century to the library of dictionaries of philosophy, J. C. Lossius published *Neues philosophisches allgemeines Real-Lexikon* (4 vols., Erfurt, Germany, J. E. G. Rudolph, 1803–1805). It contains no articles on individual philosophers. Many of the articles are written from a Kantian point of view. The topics treated include not only such philosophical concepts as *angebohrne Begriffe* (innate ideas) and *cogito ergo sum* but also concepts in anthropology, mathematics, and other disciplines.

Lossius' four-volume work was followed soon after by two other works, both of which were left incomplete.

The first of these, Georg S. A. Mellin's *Allgemeines Wörterbuch der Philosophie, zur Gebrauch für gebildete Leser* ("General Dictionary of Philosophy, for Use by

Educated Readers," 2 vols., Magdeburg, Germany, Ferdinand Matthias, 1806–1807), covers the letters *A* and *B*; no more volumes were published. The work is thoroughly Kantian, as is evidenced particularly in such articles as those on apperception, on the various aspects of *Begriff* (concept), and on the various kinds of concepts.

The other, Gottfried Immanuel Wenzel's *Neues vollständiges philosophisches Real-Lexikon* ("New Complete Philosophical Encyclopedia," 2 vols., Linz, Austria, Akademische Buchhandlung, 1807–1808), was planned in four volumes, but the author died before the work was completed, and only two volumes (covering *A* to *H*) appeared. The quaint subtitle gives an adequate, if overstated, description of the work. Literally translated, the subtitle reads: "In Which the Materials and Technical Terms Appearing in All Parts of Recent and Most Recent Philosophy Are Explained, Being Developed From History Where Necessary; Disagreements of Philosophers Are Expounded and Analyzed, Many Propositions Thereof Being Corrected, Made Precise, or Expanded; Obscurities Are Lifted; New Contributions to the Stock of Philosophical Knowledge Are Presented; and Higher Pedagogy and the Science of Intellectual Excellence [*Klugheitslehre*] Are Similarly Treated."

Original works of encyclopedic scope. Each of three German works published in the subsequent years of the nineteenth century, although denominated an encyclopedia of philosophy, presented its material nonalphabetically. The works are Gottlob E. Schulze's *Enzyklopädie der philosophischen Wissenschaften, zum Gebrauche für seine Vorlesungen* ("Encyclopedia of the Philosophical Sciences, for Use With the Author's Lectures," Göttingen, Vandenhoeck und Ruprecht, 1814, 150 pages; 2rd ed., 1818; 3rd ed., 1823, 1824); Georg Friedrich Hegel's *Encyklopädie der philosophischen Wissenschaften zum Grundrisse* ("Encyclopedia of the Philosophical Sciences in Outline," Heidelberg, A. Oswald, 1817, 288 pages; 2nd and 3rd eds., 1827, 1830; 4th ed., 3 vols., Berlin, issued by Hegel's students with their lecture notes and other materials, 1840–1845); and Johann F. Herbart's *Kurze Encyklopädie der Philosophie aus praktischen Gesichtspuncten entworfen* ("Short Encyclopedia of Philosophy Designed From the Practical Standpoint," Halle, Germany, C. A. Schwetschke und Sohn, 1831, 405 pages; 2nd ed., 1841), which is reprinted in the various editions of Herbart's collected works.

Other dictionaries. German dictionaries of philosophy, more properly so designated, were written after the earliest years of the century by Krug, Furtmair, Hartsen, No-ack, and Kirchner (as well as by Eisler, who wrote a landmark work described in a section below). Of the works referred to the first four are of mainly historical interest.

The first is Wilhelm T. Krug's *Allgemeines Handwörterbuch der philosophischen Wissenschaften, nebst ihrer Literatur und Geschichte, nach dem heutigen Standpuncte der Wissenschaft* ("General Concise Dictionary of the Philosophical Sciences, Including Their Literature and History, From the Present Standpoint of Science," 4 vols., Leipzig, F. A. Brockhaus, 1827–1829, plus supp., 1829; 2nd ed., 4 vols., 1832–1833, plus supp., 1838). Krug succeeded Kant in the chair of philosophy at Königsberg. Among the more interesting and unusual articles of Krug's book, all competently written, are "Aegyptische Weisheit" (Egyptian wisdom), "Baccalaureus der Philosophie" (Ph.B. degree), "Freund und Freundschaft" (friend and friendship), "Immoralität" (immorality), "Ontologischer Beweis für's Dasein Gottes" (Ontological Proof of God's existence), Schöne Kunst" (fine art), and "Supernaturalismus" (supernaturalism). The collaborators who produced the Adolphe Franck dictionary of 1844–1852 mentioned below and Pierre Larousse of the French encyclopedia firm criticized Krug more sharply than seems warranted for working, as far as they could see, without plan or method, for giving more emphasis to the history of philosophy than to philosophy itself, and for showing, in their opinion, insufficient gravity in his style.

Another dictionary was Max Furtmair's *Philosophisches Real-Lexikon* (4 fascicles in 1 vol., Augsburg, Karl Koll-manschen Buchhandlung, 1853–1855). The third and fourth fascicles were prepared with the collaboration of Johann N. Uschold. The author, inviting attention to his title, said that his aim was to clarify not words but things. What he presented, however, is indistinguishable from the contents of lexicons with more modest pretensions. His heavy indebtedness to Krug, which he acknowledged, is evidenced by, among other things, his inclusion of the articles "Aegyptische Weisheit," "Baccalaureus der Philosophie," "Freundschaft," and others on topics suggested by Krug's work.

In 1877 appeared Frederik A. Hartsen's *Ein philosophisches Wörterbuch* (Heidelberg, Carl Winter, 45 pages). The terms defined in this work are generally philosophical expressions rather than single terms. An example is *Betrachten etwas* (*A*) *als etwas* (*B*) ("considering something [*A*] as something [*B*]"). In some cases the definitions are of the dictionary type, with little philosophical depth.

There is also Ludwig Noack's *Philosophiegeschichtliches Lexikon* (Leipzig, Erich Koschny, 1879, 936

pages). This work emphasizes individual philosophers and is especially useful for little-known Renaissance and early modern thinkers. Although some topics—the Academy, eclectics, French philosophy, Cabala—are covered, there are no articles on the philosophical disciplines—ethics, logic, metaphysics, and so on. In 1963 the Stuttgart firm of Friedrich Frommann Verlag was planning to issue a facsimile reprint of this work.

Friedrich Kirchner, author of philosophical monographs and textbooks, including a history of philosophy which went into several editions and was translated into English, wrote a *Wörterbuch der philosophischen Grundbegriffe* (Heidelberg, G. Weiss, 1886, 459 pages), which also appeared in second and third editions (1890 and 1897), by Kirchner, and in fourth, fifth, and sixth editions (1903, 1907, 1911), revised by Carl Michaëlis. The first fascicle, 96 pages, of a projected Russian translation was published at St. Petersburg by Brockhaus–Ephron in 1913. Kirchner's work contains no articles on individual philosophers. The articles are scholarly but not penetrating; the one on logic, for example, is mainly historical and biographical.

ENGLISH. The four English dictionaries of philosophy published in the nineteenth century are now outmoded.

The first one, Isaac Taylor's *Elements of Thought, or First Lessons in the Knowledge of the Mind* (London, B. J. Holdsworth, 208 pages), appeared in 1822. With some changes in the subtitle this work went through 11 British editions (11th ed., 1866) and two American editions (2nd American ed., New York, 1851). Part II contains an exposition, in alphabetical order, of about ninety topics—analysis, argument, art, axiom, being, belief, cause, and so on—bearing upon "the nature and operation of the intellectual powers."

In 1857 William Fleming's *The Vocabulary of Philosophy, Mental, Moral, and Metaphysical, With Quotations and References, for the Use of Students* (London and Glasgow, Richard Griffin, 560 pages) was published. Subsequent editions included the second (1858), an American edition, edited by Charles P. Krauth (Philadelphia, 1860; reissued 6 times, 1866–1873); a third, edited by Henry Calderwood (1876); another American edition edited and entitled *A Vocabulary of the Philosophical Sciences* by Krauth (1878; reissued, 1879); another American edition edited by Calderwood (New York, 1887, 1890); and a work by Calderwood entitled *Vocabulary of Philosophy and Student's Book of Reference, on the Basis of Fleming's Vocabulary* (1894). The illustrative quotations in the various articles are taken mainly from English writers such as Berkeley, Hume, Jeremy Taylor, Sir William Hamilton, and J. S. Mill, but there are quotations from Kant (in English) in the article "A Priori," from Cicero (in Latin) in "Faculty," and from other foreign thinkers in other articles.

In *A Dictionary of English Philosophical Terms* (London, Rivington, 1878, 161 pages) Francis Garden undertook to present a more general and less technical account of philosophical topics than had appeared in Fleming's work. Like Fleming, however, he leaned heavily on Hamilton for arguments, illustrations, and even topics, including, for example, the article "Worse Relations" (that is, more distant relations) in logic, which is written chiefly according to Hamilton's views.

Edwin S. Metcalf's *Olio of Isms, Ologies and Kindred Matter, Defined and Classified* (Chicago, L'Ora Queta P. and J. Co., 1899, 158 pages) is an elementary and popular manual. In the section "Doctrinal and Sectarian Isms" it has articles on agnosticism, antinomianism, Arminianism, and the like; the section "Civic Isms" has articles on topics like anarchism and collectivism; "Ologies" deals with such topics as aetiology and cosmology. A section headed "Miscellany" treats altruism, analogy, art, and so forth, and "Divination" has articles on aruspicy (art or practice of divination), bibliomancy, and similar topics.

The work entitled *A Dictionary of Philosophy in the Words of Philosophers,* compiled by John R. Thomson (London, R. D. Dickinson, 1887, 479 pages; 2nd ed., 1892), is not a dictionary. Its material is arranged according to a strange outline the logic of which leaves much to be desired. In some cases it is not clear whether the material presented is in Thomson's words or in those of the philosopher who is under discussion.

FRENCH, ITALIAN, AND LATIN DICTIONARIES. Adolphe Franck, a disciple of Victor Cousin, and more than fifty collaborators, including A. A. Cournot, Paul Janet, and Ernest Renan, produced the *Dictionnaire des sciences philosophiques* (6 vols., Paris, Librairie Hachette, 1844–1852; 2nd ed., 1 vol., 1875; 3rd ed., 1 vol., 1885); the second and third editions had an analytical guide to the alphabetical articles. In matters touching on religion the authors of the articles, as pointed out by Pierre Larousse in 1865, showed restraint and circumspection; indeed, in the Preface they acknowledged reverence as one of their key principles. The work is still useful today for its extensive articles on less well-known philosophers. It contains, for example, individual articles on 12 Sophists, 11 Cyrenaics, 6 Pyrrhonists, 13 Greek Stoics, 15 Roman Stoics, and 21 members of the school of Leibniz and Wolff.

Of the other French-language dictionaries of philosophy published in the nineteenth century, one was a Belgian product, and three were Parisian.

The Belgian work, (Louis J. A.) de Potter's *Dictionnaire rationnel des mots les plus usités en sciences, en philosophie, en politique, en morale et en religion* (Brussels and Leipzig, August Sehnée, 1859, 348 pages), began as a glossary at the end of the author's *La Réalité déterminée par le raisonnement* (Brussels, 1848). The glossary was reprinted under the title *A, B, C de la science sociale* (Brussels, 1848) and was then extensively elaborated into the *Dictionnaire rationnel*. The author defended middle-class conservatism in religion, politics, morals, and economics. He decried the intellectual elite and the democratic masses, the philosophical skeptics and the radical innovators.

In 1877 Bernard Pérez wrote the 16-page *Petit dictionnaire philosophique* (Paris, A. Morant). This work, intended for baccalaureate candidates, contains mostly two-line to four-line definitions or explanations of technical terms (plus identifications of a few philosophers), from *acatalepsie, actuel, and animisme* to *vitalisme*, Xenocrate, and *zététique* (persistent skepticism). Pérez also produced a similar work, *Dictionnaire abrégé de philosophie* (Paris, Félix Alcan, 1893, 90 pages).

Pages 483–521 of Henri Marion's *Leçons de psychologie appliquées à l'éducation* (Paris, Armand Colin, 1882, 538 pages; 13th ed., 1908) contained a "Vocabulaire des noms propres et des expressions philosophiques." This vocabulary covers topics in philosophy and other fields, including art, religion, and science.

Alexis Bertrand's *Lexique de philosophie* (Paris, P. Delaplane, 1892, 220 pages) has had at least four printings. This work covers topics only, on an elementary level, but the explanations are not always clear.

There were one Italian and two Latin works of this kind published in the nineteenth century.

The first Latin work was J. A. Albrand's *Lexicon Philosophicum, Quo Verba Scholastica Explicantur*, a work 68 pages long printed on pages 557–624 of Volume IV of Albrand's edition of the *Theologia Dogmatica*, by Thomas ex Charmes (4 vols., Paris, Louis Vivès, 1856–1857). The articles, explicating *absolutum, beatitudo, esse,* and so on, provide, in prosy Latin, the standard scholastic definitions of the regular scholastic philosophical terms. The *Lexicon* was intended for the use of theological students, especially those trying to understand the system of the eighteenth-century theologian Thomas ex Charmes (also called Thomas a Charmes). Of the several reprints of Albrand's edition of Thomas' *Theologia* (6, 7, or 8 vols.) some do and some do not include Albrand's *Lexicon Philosophicum.*

The Italian work was Luigi Stefanoni's *Dizionario filosofico* (2 vols, in 1, Milan, Natale Battezzati, 1873–1875). Some of the articles—for example, those on immaculate conception, matrimony, molecule, pope, and Shakers—are a bit unusual in dictionaries of philosophy, but the articles on technical philosophical subjects are useful and contain a significant amount of detail. A pro-Catholic bias is evident in the articles on theological subjects.

The second Latin work, Niceto A. Perujo's *Lexicon Philosophico-theologicum* (Valencia, Spain, Friedrich Domenech, 1883, 352 pages), had a scholastic orientation. It contains 1,364 articles, including explanations not only of terms but also of such common philosophical propositions as "Dato uno absurdo, sequitur aliud" ("If one absurdity is granted, another follows"). Some of the explanations are supported by extensive quotations from Aquinas, Bonaventure, and others.

RUSSIAN. A number of notable dictionaries of philosophy were written in Russia in the nineteenth century.

In 1819 appeared Alexander I. Galich's *Opyt Filosofskogo Slovaria* ("Toward a Philosophical Dictionary," St. Petersburg), the second fascicle of a larger work on the history of philosophical systems. This dictionary contains 217 articles, from "Absolute" to "Theurgy." The topic headings are given in the Latin alphabet—for instance, "Absolutum"—and the explanations in Russian. Special attention is paid to new philosophical terms.

Alexander I. Galich's *Leksikon Filosofskikh Predmetov* (Vol. I, No. 1, St. Petersburg, Tip. Imp. Akad. Nauk, 1845, 298 pages) is the first fascicle of a proposed set of nine (three volumes with three numbers in each). It covers about 170 terms beginning with *A* or *B* in aesthetics, ethics, logic, and metaphysics. The project was discontinued when the author's notes were destroyed in a fire.

S. S. Gogotsky's monumental work *Filosofsky Leksikon* (4 vols., Kiev, University of Kiev and other publishers, 1857–1873; 2nd ed., 1 vol., St. Petersburg, I. I. Glazunov, 1859) contains about twelve hundred articles. The articles on philosophical method, such as those on analogy, classification, dialectic, dogmatism, and method in general, are especially noteworthy. In 1876 Gogotsky produced *Filosofsky Slovar'* (Kiev, Tip. Red. "Kievsk Telegrafa," 146 pages), a one-volume condensation of his lexicon, containing approximately the same number of articles.

EISLER. Rudolf Eisler produced his *Wörterbuch der philosophischen Begriffe und Ausdrücke* ("Dictionary of Philosophical Concepts and Expressions," Berlin, E. S. Mittler und Sohn, 1899, 956 pages), which, following the setup of the *Wörterbuch* by Friedrich Kirchner, has no articles on individual philosophers. Of the three volumes of the fourth edition, whose title was shortened to *Wörterbuch der philosophischen Begriffe* (Berlin, E. S. Mittler und Sohn, published with the cooperation of the Kant-gesellschaft, 1927), the second and third were edited with the assistance of Karl Roretz, Eisler having died after the work on the first volume was completed.

This is perhaps the best technical dictionary of philosophy produced up to its time. Even now, it is probably one of the ten best available dictionaries of philosophy, ranking along with the better works of the twentieth century. Its articles contain terse definitions and are rich not only in relevant quotations in the original languages, including English, but also in bibliographical citations. On Oriental subjects the articles were weak in the first edition (Sāmkhya being dismissed with the statement that it is the system of the Indian thinker Kapila) but were strengthened somewhat in subsequent editions. The later editions, although expanded in coverage, contain fewer quotations in languages other than German.

In 1964 the Basel firm of Benno Schwabe had in preparation a new edition of the *Wörterbuch* under the editorship of Joachim Ritter.

For use by students Eisler summarized the main articles of his large dictionary in the *Handwörterbuch der Philosophie* (Berlin, E. S. Mittler und Sohn, 1913, 801 pages), of which a second edition, supervised by Richard Müller-Freienfels, was issued not only as a regular book in 1922 but also as a "microbook" (Düsseldorf, Microbuch- und Film Gesellschaft, 1922, 785 pages on 88 sides).

Eisler also produced the *Philosophen-Lexikon: Leben, Werke und Lehren der Denker* ("Dictionary of Philosophers: Lives, Works and Doctrines of the Thinkers," Berlin, E. S. Mittler und Sohn, 1912, 889 pages) to make up for the lack of treatment of individuals as such in his *Wörterbuch*. The *Philosophen-Lexikon* was the first modern biographical dictionary of philosophers. Although its articles are shorter, more numerous, and alphabetically arranged, it recalls the useful work of Diogenes Laërtius. From Anathon Aall of Norway to Ulrich Zwingli, the Reformation figure, some four thousand philosophers are identified and, when appropriate, discussed, with their main writings and writings about them listed. Eisler could perhaps be excused for according some emphasis to German philosophers, and it is not strictly fair to criticize comparative comprehensiveness on the basis of lines of print, especially since most of Eisler's allocations of space seem right; nevertheless, one may perhaps with some warrant complain that Kant gets 33 pages, Wundt 16, Spinoza 11, Plato (as well as Hegel and Leibniz) 10, and Aristotle 9 and that Hermann Cohen gets more space than Augustine, Fichte more than Descartes, Herbart more than Hume, Lotze more than Locke, Maimon more than Maimonides, and Meinong more than Bentham.

EARLY TWENTIETH CENTURY

In 1901 an important dictionary was published, and an important dictionary was begun. The early twentieth century also saw the publication of dictionary-type or supposedly encyclopedic treatments of philosophical topics by Lalande, Windelband, and less well-known writers.

GOBLOT. Edmond Goblot issued *Le Vocabulaire philosophique* (Paris, Librairie Armand Colin, 1901, 513 pages; 6th ed., 1924), in which he tried not only to record the actual meanings of terms but in part to correct confused usages by suggesting, for example, separate meanings for *général* and *universel*; for *particulier, individuel,* and *singulier;* and for *mémoire* and *souvenir.* But philosophers being the individualists that they are in the use of words, their degree of acceptance (if any) of his commendable suggestions is not perceptible. Spanish translations of this work were published at Barcelona in 1933 and at Buenos Aires in 1942 and 1945.

BALDWIN. The other important work of 1901 was Baldwin's. James M. Baldwin, a psychologist, edited, with the collaboration of an international board of advisers and contributors that included Bosanquet, Dewey, William James, Janet, Lloyd Morgan, Moore, Münsterberg, Peirce, Pringle-Pattison, Royce, Sidgwick, Stout, and Urban, the *Dictionary of Philosophy and Psychology* (3 vols. in 4, New York, Macmillan, 1901–1905; reprinted with corrections several times, in part or in entirety, by the same firm, in some cases with the designation "New Edition"; also reprinted by Peter Smith twice, partly at New York and partly at Gloucester, Mass., 1940s, 1950s). Volume III, in two parts, is a bibliography of philosophy and psychology, by Benjamin Rand, to which there were annual supplements in the *Psychological Index* from 1901 to 1908.

In the Preface the editor stated that a dictionary of terms used in Greek and scholastic philosophy "is much needed: but we have not attempted it." The dictionary does, however, include articles on Greek terminology (8

pages, by Royce) and Latin and scholastic terminology (11 pages, by Royce), as well as on analogy, nous (mind), and other special terms. Moreover, the editor aimed "to present science—physical, natural, moral—with a fullness and authority not before undertaken in a work of this character." Thus, there are articles on anthropology, brain, case law, hybrid, money, peace, pupa, and others. Like Goblot, Baldwin futilely suggested that his readers follow the recommendations made in some of the articles for preferred philosophical usage. For many entries German, French, and Italian equivalents are recommended. In addition, at the end of Volume II there is an index of Greek, Latin, German, French, and Italian terms, including those covered by separate articles on the terms as such and those merely mentioned as recommended equivalents.

Philosophically, the articles in the Baldwin dictionary are of uneven value. Some, especially the biographical articles, are too short, and there are no articles at all on Maine de Biran, Renan, and Saint-Simon. Others are broken down too minutely into terms rarely encountered, including Peirce's articles on particulate, *parva logicalia,* philosopheme, predesignate, and prosyllogism. In others there is cavalier treatment of the philosophical aspects of a subject, as in the psychologically oriented article on the self. Some articles, however, are excellent, especially the longer ones by Dewey—for example, those on nature, pluralism, and skepticism; those by Moore on cause and effect, change, nativism, quality, real, reason, relation, relativity of knowledge, spirit, substance, teleology, and truth; and the longer ones of the approximately 180 written by Peirce, including his 23 columns on syllogism, 10 columns on uniformity, and 10 on matter and form. Peirce's articles (the preparation of which, from 1901 to 1905, constituted his last steady employment) were mainly fragments of a book on logic which he never finished; only about half of these articles were reprinted in the Harvard *Collected Papers* of Peirce. Moore's 12 articles, which he later, with undue modesty, called crude, have not been reprinted.

LALANDE. With the collaboration of others André Lalande, a professor at the Sorbonne, issued the *Vocabulaire technique et critique de la philosophie* (21 fascicles, Paris, Félix Alcan, 1909–1922; revision of fascicle covering *A* in *Bulletin* of Société Française de Philosophie, 1923; 2nd ed., 2 vols., 1926; 3rd ed., 2 vols., 1928; 4th ed., 3 vols., 1931, reissued in 1932, Vols. I and II reissued, 1938; 5th–9th eds., 1 vol., 1947, 1950, 1956, 1960, 1962; 5th ed. translated into Spanish, Buenos Aires, 1953, with 2nd ed., Buenos Aires, 1964, 1,502 pages). Lalande was 95 years old when the ninth edition of the *Vocabulaire* was published. At the bottom of most of the pages appear the comments of members of the Société Française de Philosophic, including Peano and Russell among the foreign members, on the articles. The emphasis of the articles is on clarifying the meanings of terms and the usage of expressions rather than on the imparting of historical or technical information.

ORIGINAL WORKS OF ENCYCLOPEDIC SCOPE. Just as, early in the nineteenth century, the works of Schulze, Hegel, and Herbart were published as encyclopedias of the philosophical sciences, so early in the twentieth century three works of this kind were published or begun. The first "nonencyclopedia" was a series of works, edited by H. Renner and published at Charlottenburg, Germany, by O. Günther beginning in 1907, under the general title *Encyklopädie der Philosophie*. It included, for example, an introduction to philosophy and volumes on the philosophy of Rudolf Stammler and Rudolf Eucken.

Second of the three nonencyclopedias was August J. Dorner's *Encyklopädie der Philosophie* (Leipzig, Verlag der Durr'schen Buchhandlung, 1910, 334 pages); in Kantian fashion it dealt with phenomenological investigations, the construction of empirical science, and similar topics.

The third was a proposed *Encyklopädie der philosophischen Wissenschaften,* of which the first volume, *Logik*, was published in 1912, edited by Wilhelm Windelband and Arnold Ruge (Tübingen, Germany, J. C. B. Mohr, 275 pages), containing expositions of the principles of logic by Windelband, Josiah Royce (translated from English), and Louis Couturat (translated from French); of the task of logic by Benedetto Croce (translated from Italian); of the problems of logic by Federigo Enriques (translated from Italian); and of the bearing of the concepts of consciousness on logic by Nicholas Lossky (translated by Lossky himself from the original Russian).

An English edition of the Windelband–Ruge encyclopedia was projected under the editorship of Sir Henry Jones, and the first volume, *Logic,* was published in 1913 (London, Macmillan, 269 pages). For the English edition Royce's English version was available, Couturat's article was done into English from the original French rather than from the published German version, and the German of Lossky's article was his own; therefore, as the translator, B. Ethel Meyer, pointed out, only Croce's and Enriques' articles "suffered a double process of translation."

The onset of war in 1914 and the death of Windelband in 1915 resulted in the abandonment of the project. Windelband's contribution to the first volume, issued separately in German in 1913, was republished in English years later as *Theories in Logic* (New York, Philosophical Library, 1961, 81 pages). Royce's contribution was also published separately, as *The Principles of Logic* (New York, Philosophical Library, 1961, 77 pages).

OTHER WORKS. The works of the early twentieth century by less well-known writers in Italian, French, German, English, Russian, and Japanese were numerous.

In Italian there was Cesare Ranzoli's *Dizionario di scienze filosofiche* (Milan, Ulrico Hoepli, 1905, 683 pages; 2nd ed., 1916, 1,252 pages; 3rd ed., 1926, 1,207 pages; 4th ed., Maria P. Ranzoli, ed., 1943, 1,360 pages; 5th ed., Maria Ranzoli, ed., 1952, 1,313 pages). Covering only topics, not individual philosophers, the book contains articles on Pyrrhonism and Pythagoreanism (and later editions cover existentialism), but there is none on Platonism. The articles are of high quality.

In 1906 appeared Élie Blanc's *Dictionnaire de philosophie ancienne, moderne et contemporaine* (Paris, P. Lethielleux, 1,248 cols.; supp., for 1906–1907 and 1906–1908; consolidated ed., 1909). Blanc also published a vocabulary of scholastic and contemporary philosophy, presented at the beginning of his *Traité de philosophie scolastique* (3 vols., Lyon, Emmanuel Vitte, 1889; 3rd ed., Paris, 1909), and the *Dictionnaire universel de la pensée, alphabetique, logique et encyclopédique* (2 vols., Lyon, Emmanuel Vitte, 1899), which was a thesaurus-type classification of words, ideas, and things. In the *Dictionnaire de philosophie* his Catholic viewpoint is evident in many places; indeed, his starting point, he said, is moderate dogmatism.

In Germany Rudolf Odebrecht produced the *Kleines philosophisches Wörterbuch; Erklärung der Grundbegriffe der Philosophie* (Berlin, Buchverlag der "Hilfe," 1908, 83 pages; 6th ed., Leipzig, Felix Meiner, 1929). The choice of topics in this highly condensed wordbook was in some cases injudicious. There are entries on *heliozentrisch* (heliocentric) and *Hypnose* (hypnosis) but none on the Academy, Epicureanism, or Taoism.

A pocket volume, about 2½ inches by 4 inches, one of a series of about fifty covering literary terms, commercial terms, art terms, and so on was edited by Arthur Butler, *A Dictionary of Philosophical Terms* (London, G. Routledge and Sons, and New York, E. P. Dutton, 1909, 114 pages). The *Dictionary of Philosophical Terms* depends heavily on Kant, who is cited in ten of the first fifty articles. Among the topics treated are a number of German terms, such as *Anschauung* (outlook), *Begriff* (concept), and *Ding an sich* (thing-in-itself).

In 1909 also appeared Arturo Mateucci's *Vocabolaretto di termini filosofici* (Milan, Casa Editrice Sonzogno, 63 pages; 2nd ed., 1925). Intentionally elementary in its treatment, in many cases this work contains little more than dictionary definitions of the concepts covered. Some 75 percent of the articles consist of only one, two, or three lines.

Fritz Mauthner edited the *Wörterbuch der Philosophie* (2 vols., Munich, G. Muller, 1910–1911; 2nd ed., 3 vols., Leipzig, Felix Meiner, 1923). Mauthner was a literary critic and nonacademic philosopher who contributed pioneering insights on the question of what, if anything, ordinary language reveals about the world, whether the distinction between analytic and synthetic propositions is tenable, and so on. His *Wörterbuch*, after a rambling introduction of 96 pages, presents a mixture of very odd items and very useful, though informal, ones. The odd items include the articles "Babel," "Bacon's Ges'pensterlehre" (Bacon's study of ghosts), "Form" (40 cols., with only a passing reference to Aristotle), and "Graphologie." The more useful ones include "Geschichte" (history, 68 cols.), "Natur" (nature, 29 cols.), "Nichts" (nothing, 14 cols.), and "Spinoza's 'Deus'" (Spinoza's "God," 19 cols.); even these, however, should be used with caution, for they contain some questionable material.

Ernest L'vovich Radlov's *Filosofsky Slovar'* (St. Petersburg, Brockhaus–Ephron, 1911, 284 pages; 2nd ed., Moscow, G. A. Leman, 1913) covers aesthetics, ethics, logic, psychology, and the history of philosophy. It is of only limited usefulness.

Tetsujiro Inouye, Yujiro Motora, and Rikizo Nakashima edited the *Dictionary of English, German, and French Philosophical Terms, With Japanese Equivalents* (Tokyo, Maruzen Kabushiki–Kaisha, 1912, 205 pages), written in English. This is the definitive edition of the *Dictionary of Philosophical Terms* first brought out by Inouye and others in 1881 and issued in a second edition in 1884. For topical entries, including some in Arabic, Greek, and Latin besides the languages listed in the title, only the Japanese equivalents are given; the personal entries also provide identifying information.

Julius Reiner's *Philosophisches Wörterbuch* (Leipzig, Otto Tobies, 1912, 295 pages) is an elementary work in which, for example, the article on *Ambiguität* (ambiguity) consists of one word, *Zweideutigkeit* (having two meanings), and the article on *Intellekt* (intellect) consists of two

words, *Geist, Verstand* (spirit or mind, understanding). Other articles, however, such as those on *Darwinismus* and *Ethik* (Darwinism and ethics), go more deeply into the subject.

Another German work was Heinrich Schmidt's *Philosophisches Wörterbuch* (Leipzig, Alfred Kröner, 1912, 106 pages; 8th ed., 1930). This was republished in the United States in 1945 by authority of the alien property custodian and went through several editions; the tenth edition (1943) was reprinted in the United States without the authority of the alien property custodian; the sixteenth edition appeared in 1961. The editions which appeared after the death of the author in 1935 were supervised by various editors. The numerous editions of this work had a vast circulation in all German-language areas. Indeed, it is perhaps the most widely used philosophical dictionary in any language at any time, the Eisler work being its main rival for this distinction. In the ninth edition (1934), while Schmidt was still alive, some pro-Nazi and anti-Jewish comments were included, and in the tenth edition (1943) the desecration of scholarship was compounded with obsequious compliments to insignificant Nazis and truly monstrous articles on Bergson, Freud, Husserl, and others. Recent editions bend over backward to rectify these aberrations.

Paul Thormeyer's *Philosophisches Wörterbuch* (Leipzig, B. G. Teubner, 1916, 96 pages; 4th ed., 1930) is an uncommonly useful short reference work. It is well organized and was up-to-date at the time it was issued.

THE NINETEEN-TWENTIES

ANGLO-SAXON SILENCE. In the 1920s 12 dictionaries of philosophy appeared or were begun—4 in German and 1 each in Hungarian, Swedish, Dutch, French, Spanish, Hebrew, Japanese, and Chinese. Not one was published in the United States or Great Britain. Indeed, the only English-language work deserving of mention here published between Butler's *Dictionary* of 1909 and Runes's *Dictionary* of 1942 was a quasi encyclopedia, the *International Encyclopedia of Unified Science*, begun in 1939. The Anglo-Saxon silence can only be recorded here. The explanation of it requires more data than are readily at hand.

GERMAN. Of the German works published in the 1920s three were published in 1923. The *Systematisches Wörterbuch der Philosophie*, by Karl W. Clauberg and Walter Dubislav (Leipzig, Felix Meiner, 1923, 565 pages), is systematic to a fault, many of the articles being broken down into standard subdivisions—for example, definition, statement, addition, and example—in a somewhat rigid fashion. Dubislav, who was a professor of philosophy at the University of Berlin, had a continuing interest in the clarification of concepts. He was close to logical empiricism and wrote the comprehensive *Die Definition* (Leipzig, Felix Meiner, 1931, 160 pages); he also made notable contributions to the philosophy of method, mathematics, and science.

In Rudolf Wagner's *Philosophisches Wörterbuch* (Munich, Rösl, 1923, 148 pages) articles range in length from one-word or two-word definitions or identifications to the six-page article on the history of philosophy, which consists mainly of a five-page outline taken from Wilhelm Wundt's *Einleitung in die Philosophie* (1914); individual philosophers are not accorded separate treatment.

In most dictionaries of philosophy that cover both topics and persons, the articles on topics are far more numerous than those on people; in Alfred Sternbeck's *Führer durch die Philosophie; Philosophenlexikon und philosophisches Sachwörterbuch* (Berlin, Globus Verlag, 1923, 306 pages), however, those on people almost equal the topical articles in number. Moreover, whereas some of the topical articles are elementary, containing little more than dictionary definitions, the biographical articles are more substantial.

Two years later, there was published the last of the German works of the 1920s, *Klare Begriffe! Lexikon der gebräuchlicheren Fachausdrücke aus Philosophie und Theologie*, by Theodor Mönnichs, S.J. ("Clear Concepts! Dictionary of the Most Common Technical Terms of Philosophy and Theology," Berlin, Ferdinand Dümmlers Verlag, 1925, 170 pages; 2nd ed., 1929). This work was written, according to the author, from the standpoint of *philosophia perennis* and Catholic theology. The longest article is the sixty-line one on religion. The pervasive scholastic emphasis in the book is indicated by the fact that many articles begin with the Latin equivalent of the term being covered, and the second edition contains, as an appendix, a 20-page alphabetical list of Latin philosophical terms with their German equivalents.

HUNGARIAN. The Hungarian work of the 1920s was *Philosophiai Szótár*, by Enyvvári Jenö (family name Enyvvári), published at Budapest by Franklin-Társulat (1923, 187 pages). The articles in this work show a creditable familiarity with West European scholarship. The titles of many of the articles are in languages other than Hungarian—for example, "Élan vital," and "Moral Insanity." Appended are a list of philosophers and a competent discussion of philosophical bibliographies.

SWEDISH. Sweden contributed the *Filosofiskt Lexikon*, edited by Alf Ahlberg (Stockholm, Bokförlaget Natur och Kultur, 1925, 207 pages; 3rd ed., 1951). In this work Swedish philosophers were given fuller treatment than others—C. J. Boström, 15 cols.; E. G. Geijer, 10 cols.; Aristotle and Plato, 6 cols. each.

DUTCH. The Dutch work of the period was C. J. Wijnaendts Francken's *Koort Woordenboek van Wijsgeerige Kunsttermen* ("Short Dictionary of Philosophical Terms," Haarlem, D. H. Tjeenk Willink & Zonen, 1925, 157 pages). It covers topics only, in a fairly popular style, and the choice of topics is liberal, making room for such terms as *kosmopolitisme, opportunisme*, and *sarcasme*, along with more technical philosophical terms.

FRENCH. In France appeared Armand Cuvillier's *Petit Vocabulaire de la langue philosophique* (Paris, Librairie Armand Colin, 1925, 109 pages; 13th ed., 1953). It was subsequently translated into Turkish (Ankara, 1944) and Portuguese (São Paulo, Brazil, 1961). This work was intended by its author to be at once *élémentaire* and *précis*. In large measure it succeeded in achieving both objectives.

SPANISH. Begun in Spain was the *Diccionario manual de filosofía* by Marcelino Arnáiz and B. Alcalde (Madrid, Talleres Voluntad, 1927–). Volume I, "Vocabulario Ideario" (659 pages), is rich in bibliography, and many of the articles contain sound historical data in addition to the conceptual explanations which the volume was essentially intended to provide. A projected second volume, covering the history of doctrines, biographies, and bibliography, was not published.

EASTERN LANGUAGES. In the 1920s dictionaries of philosophy appeared in three Eastern languages, apparently for the first time (aside from translations).

Hebrew. The Hebrew dictionary of philosophy begun in the 1920s was the *Otsar ha-Munahim ha-Filosofiyim ve-Antologiyah Filosofit* ("Thesaurus of Philosophical Terms and Philosophical Anthology"), by Jacob Klatzkin (4 vols., Leipzig, August Pries, 1928–1933); an introductory volume, published in Berlin by "Eschkol" Verlag in 1926, contains an anthology of Hebrew philosophy. Each of the four regular volumes has, as an added Latin title, *Thesaurus Philosophicus Linguae Hebraicae et Veteris et Recentioris*; Volumes III and IV had M. Zobel as coeditor. The dictionary articles are on topics only, not philosophers or schools of philosophy. Many of the articles contain the German or Latin equivalent of the title of the article; indeed, the purpose usually seems to be to explain the use of terms rather than to convey historical information on the topic as a topic, although the usage of historical writers on the subject is often indicated.

Japanese. A 1,026-page work entitled *Tetsugaku dai-Jisho* ("Dictionary of Philosophy") was published at Tokyo in 1924 by Dai Nippon Hyakka Jisho (Japanese Encyclopedia). The eighth edition (1928) consists of three volumes of text, an index volume, and a supplement. In the text volumes and in the supplement each article begins with the title in Japanese, followed usually by English, German, and French equivalents of the title. Thus, the first article in the first volume is headed, after the Japanese title, "Love. Liebe. Amour." The next several articles deal with patriotism, agape (listed alone after the Japanese title), affection, love and hate (with the Greek equivalents, φιλότηζ and νεικοζ), Aitareya Upanishad, idealism, vaguedualism, pity, and Augustine. Some of the articles, including the one on religion, are extensive, and many include references to European works.

The index volume of this Japanese dictionary has a title page in German ("Encyclopaedia Japonica, *Enzyklopädische Wörterbuch der Philosophie . . . Register . . .* Tokyo: Dobunkwan"). In addition to a Japanese index, it contains English, French, German, Latin, Pali, Sanskrit, and Chinese indexes and a *Namenregister* (index of names). In the English index approximately 35 of the first 100 entries are strictly philosophical—absolute, abstract, Academy, accident, actual, and so on; most of the others pertain to psychology. In the *Namenregister*, too, about 35 of the first 100 entries are standard names in philosophy—Abelard, Aenesidemus, Albert the Great, al-Fārābī, and so on.

Chinese. In *Chê Hsüeh Tz'ŭ Tien* ("Dictionary of Philosophy"), by Fan Ping-ch'ing (Shanghai, Commercial Press, 1926, 1,110 pages; 2nd ed., 1935; 3rd ed., 1961), the title of each article is given in Chinese, English, French, and German. The dictionary begins with an article on monism and continues with articles on monotheism, Monophysites, the seven liberal arts, the seven wise men, dualism, dilemma, antinomy, ethnology, subconscious, Albert the Great, major term, minor term, asymmetry, *credo quia absurdum*, medieval philosophy, Pascal, Parmenides, and so forth. The content is scholarly, but there are numerous errors in the Western languages. The work closes with an alphabetical index of names (in which Abelard has 8 references, Aristotle 45, Kant 28, and Marx 5) and an alphabetical index of topics from abiogenesis (1 reference) to *Zwecksystem* (1 reference).

THE NINETEEN-THIRTIES

In the 1930s there appeared four Italian and two Russian works. During this period a number of works in other languages were also published.

GERMAN. Germany began the decade with Max Apel's *Philosophisches Wörterbuch* (Berlin and Leipzig, W. de Gruyter, 1930, 155 pages). The fifth edition, which was revised by Peter Ludz, appeared in 1958, and a Spanish translation was published at Mexico City in 1961. Editions of Apel's work published since World War II are pro-Soviet.

DUTCH. In the Netherlands appeared the *Encyclopaedisch Handboek van het Moderne Denken*, edited by Willem Banning and 41 collaborators (2 vols., Arnhem, Van Loghum Slaterus, 1930–1931; 2nd ed., 1 vol., 1942; 3rd ed., 1 vol., 1950). Although the third edition emphasizes such modern ideas as anarchism, Gestalt theory, phenomenology of worship, quantification of the predicate, and the United Nations, the work does not neglect such standard philosophical ideas as category, natural law, and thing.

ENGLISH. A United States contribution, a quasi encyclopedia, in the 1930s was the inauguration of the *International Encyclopedia of Unified Science*, by Otto Neurath, Rudolf Carnap, and Charles Morris in 1936/1937 at the University of Chicago. This work, carried on after Neurath's death in 1945 by the Institute for the Unity of Science in Boston under the joint editorship of Carnap and Morris, consists thus far of 15 fascicles, of which Volume I, Number 1 (1938), contained articles by Niels Bohr on analysis and synthesis in science, by Carnap on logical foundations of the unity of science, by John Dewey on unity of science as a social problem, by Morris on scientific empiricism, by Neurath on unified science as encyclopedic integration, and by Bertrand Russell on the importance of logical form. The other 14 are monographs by individual authors. To each of these a volume and a number are assigned. The latest numerically is Volume II, Number 9 (1951), a study by Jørgen Jørgensen on the development of logical empiricism. The latest chronologically, Volume II, Number 2 (1962), is a monograph by Thomas S. Kuhn on the structure of scientific revolutions.

Thus, this "encyclopedia," like Hegel's, Herbart's, Contri's (see below), Windelband–Ruge's and the *Nouvelle Encyclopédie philosophique*, is a compendium but it is not alphabetical. The announced topics of the volumes are foundations of the unity of science, Volumes I and II; theories, induction, probability, and so on, Volume III; logic and mathematics, Volume IV; physics, Volume V; biology and psychology, Volume VI; social and humanistic science, Volume VII; and history of the scientific attitude, Volume VIII. This project, inspired by logical positivism and designed by Neurath to show that all the sciences speak the same language—essentially, physicalism—was overambitious.

FRENCH. France's contribution in the 1930s was Jean B. Domecq's *Vocabulaire de philosophie* (Tours, Alfred Cattier, 1931, 208 pages), which has separate alphabetical arrangements of topics for logic, ethics, and metaphysics and a consolidated index at the end. The author was an abbot, and the work has a Catholic orientation.

Mention may also be made of a series of monographs inaugurated in Paris in 1934 by the Presses Universitaires de France, *Nouvelle Encyclopédie philosophique,* which do not constitute an encyclopedia in the strict sense. Among the monographs published thus far are, for example, Louis Lavelle's *Introduction à l'ontologie* (No. 41, 1947) and Robert Blanché's *Les Attitudes idéalistes* (No. 45, 1949).

ITALIAN. Four Italian dictionaries of philosophy appeared or were begun in this period. The first was Giovanni Semprini's *Piccolo dizionario di coltura filosofica e scientifica* (Milan, Edizioni Athena, 1931, 502 pages). This was revised as *Nuovo dizionario di coltura filosofica e scientifica* (Turin, Società Editrice Internazionale, 1952, 470 pages). The work covers topics and individuals in philosophy, science, and education.

In 1933, Antonio Bettioli's *Il pensiero filosofico attraverso i secoli* (Urbino, Editoriale Urbinate, 234 pages) was published. The articles are grouped into schools and systems of philosophy—for example, the Academy, eclectics, idealism—and individual philosophers—113 names, including Dante, Feuerbach, Goethe, Leonardo, Swedenborg, and Tolstoy but not Bergson, Dewey, Husserl, Origen, Philo, or Proclus. The book is of limited value.

An elementary work with little penetration, Francesco Varvello's *Dizionario etimologico filosofico e teologico* (Turin, Società Editrice Internazionale, 406 pages), appeared in 1937, with a second edition in 1938. Fascism is lauded as the opposite of various false forms of government. According to the author, Marx (described as a Jew) rejected the idea that man does not live by bread alone. The articles on religion are pro-Catholic.

There was also Emilio Morselli's *Piccolo dizionario filosofico* (Milan, Carlo Signorelli, 1938, 104 pages). In

this book the author aimed to help young readers who encounter in the classics of thought special philosophical expressions, expressions whose meanings differ not only from what they are in ordinary discourse but also from period to period.

An Italian work of the 1930s which called itself an encyclopedia of philosophy but which was not arranged alphabetically was Siro Contri's *Piccola enciclopedia filosofica* (Bologna, Costantino Galleri, 1931), of which only the first volume, on logic and the philosophy of science, was published.

PORTUGUESE. In Brazil appeared Renato Kehl's *Bioperspectivas; dicionário filosófico* (Rio de Janeiro, Livraria Francisco Alves, 1938, 187 pages), which is a series of Voltairian musings on art, the categorical imperative, civilization, death, education, free will, God, history, intelligence, original sin, personality, philosophy, politics, progress, work, and other subjects.

RUSSIAN. The first of the two Soviet contributions of the 1930s was Timofei S. Ishchenko's *Kratky Filosofsky Slovar'* (Moscow, Moskofsky Rabochy, 1931, 200 pages), which gave more space to Stalin (four cols.) than to Plato, Aristotle, Kant, Hegel, or Marx. Other Marxist topics, such as dictatorship of the proletariat, were accorded correspondingly disproportionate treatment with the usual positive bias. The three items in the bibliography on Aristotle are by Marx, Engels, and Stalin, respectively.

The second was a work by Mark M. Rozental' and Pavel F. Yudin, likewise entitled *Kratky Filosofsky Slovar'* (Moscow, 1939; 2nd–4th eds., 1940, 1951, 1954, each of which was reprinted the following year). A new edition appeared in 1963 with the title modified by the omission of the first word, which means "short," although the 1963 edition of 544 pages is actually shorter than the previous edition, which had 567 pages. The encyclopedia was translated into Spanish in 1945, Bulgarian in 1947, English in 1949, Ukrainian in 1952, Hebrew in 1954, and Chinese, French, Polish, and Rumanian in 1955. Reportedly, 2 million copies of the Russian original were sold in the first ten years after publication, and the press run of one of the printings in the 1950s was 500,000. The English version, adapted and translated by Howard Selsam (New York, International Publishers, 1949, 128 pages), stated in the Preface that the volume reflects Marxist partisanship (for materialism and for socialism) as contrasted with the lack of a "common approach" and the "alphabetic disorder" of other dictionaries of philosophy.

Illustrative of the topical entries in the English version are those in the *E*'s: "Eclecticism," "Economic Bases of Society," "Economic Determinism," "Economics and Politics," "Empiricism," "Empirio-criticism," "Energism (metaphysical)," "Epistemology," "Equality," and "Equilibrium, Theory of." The men treated in the *S*'s are Saint-Simon, Schelling, Spencer, Spinoza, and Stalin, and the article on Stalin is the longest of these.

The article on Kant in the English version dutifully quotes from Lenin, and those on Campanella and dualism, among others, drag in quotations from Stalin. Many of the articles on individual philosophers vapidly make a point of recounting what Marx, Engels, Lenin, or Stalin thought of the philosopher or even reverently disinter a colorless quotation from Stalin summarizing what Marx or Lenin thought of the philosopher. The article "Partisanship of Philosophy" states that the class struggle is always behind the scenes in the open struggle of philosophical opinions.

According to Alexander Philipov, a former professor of philosophy at the University of Kharkov who later emigrated in the United States, for the English version Selsam watered down two features of the original—its invective and its extravagant praise of Lenin and Stalin—in order to make the edition less offensive to Western readers.

A significant feature of the original is the fact that the article on Stalin in the fourth edition (1954) ended with a sentence which may be translated "The immortal name of Joseph Stalin will live forever in the minds and hearts of the Russian people"; that sentence vanished without a trace in the 1955 printing of the same edition. In the 1963 edition, of which 400,000 copies were printed and which had about 160 collaborators (including most of the important figures in current Soviet philosophy), there is no article on Stalin, and the Preface acknowledges the "enormous harm" resulting from the cult of Stalin. The 1963 edition is stronger than its predecessors in coverage of linguistic philosophy, logical positivism, and logic.

LITHUANIAN. Lithuania's contribution to the history of philosophical dictionaries is a 97-page article entitled "Bendroji Filosofijos Terminija" ("General Terminology of Philosophy"), by Stasys Šalkauskis; it constituted an entire issue of the periodical *Logos; Filosofijos Žurnalas* (Kaunas), 1937. The article listed some fifteen hundred Lithuanian terms useful in philosophical discussions, with their equivalents in French, German, and Russian. The list was supplemented by a discussion of synonyms of various philosophical terms in Lithuanian. In a 1938 issue of the same periodical Šalkauskis presented a list of

over fifteen hundred German philosophical terms with their Lithuanian equivalents.

HEBREW. In Palestine, Zvi Hirsch Rudy produced the *Leksikon le-Filosofiyah* (Tel Aviv, Dvir, 1939, 816 cols.), with an added title page in Latin, *Philosophiae et Scientiarum Propinquarum Lexicon Hebraicum*. This work is generous with Latin terms, as the titles of articles—for example, "Actus purus," in Hebrew transliteration; as the Latin equivalents of the Hebrew titles of topical articles—for example, "Natura Naturans" as the equivalent of "Teva Tovei"; and as the titles of works cited—for example, works by Abelard and Augustine cited in the articles on those thinkers. Contemporary writers, such as Dewey and Meyerson, and topics of current interest, such as absurd and *élan vital*, are also included. The articles lack penetration. The Bibliography at the end is erratic in including, along with students' handbooks, a poorly balanced small selection of specialized monographs.

CHINESE. In 1934 appeared a new Chinese dictionary, not so strictly confined to philosophy as was the 1926 Chinese work. This was the *Ssu Hsiang Chia Ta Tz'ŭ Tien* ("Dictionary of Great Thinkers"), by P'an Nien-chih (Shanghai, Shih Chieh, 1,062 pages), which contains over five hundred articles on philosophers, writers, artists, musicians, and others. Mo Tzu quite properly is accorded 12 columns, but in the modern period Kant and Mill get only 5 columns each while Mussolini rates 6. Many names are misspelled.

THE NINETEEN-FORTIES

The 1940s saw six philosophical dictionaries in Spanish, five in English, five in German, two in Italian, two in French, and one each in Hungarian and Turkish.

FERRATER MORA. José Ferrater Mora began the decade by producing the *Diccionario de filosofía* (Mexico City, Editorial Atlante, 1941, 598 pages; 2nd ed., 1944; 3rd–4th eds., Buenos Aires, Editorial Sudamericana, 1951, 1958; 5th ed. in preparation). It is one of the most useful dictionaries published in the twentieth century. From the technical standpoint it may be mentioned that the author used a sensible system of cross references which eliminates the need for an index; he chose as topics for articles units which are neither too large nor too small. The bibliographical citations provided at the ends of some articles are judiciously selected.

The writing shows a philosophical understanding decidedly above the average for writers of philosophical dictionaries. Ferrater Mora was equally strong in his knowledge of modern logic and positivism and in the more traditional philosophical trends and developments associated with Continental metaphysics. The comprehensiveness of his scholarship and the soundness of his judgment have combined to create a monumental one-man contribution to the library of dictionaries of philosophy.

OTHER LATIN AMERICAN WORKS. In the same year, 1941, two other dictionaries were published in Latin America. One was Martín T. Ruiz Moreno's *Vocabulario filosófico* (Buenos Aires, Editorial Guillermo Kraft, 1941, 156 pages; 2nd ed., 1946, 302 pages). Among the articles of special interest in it are "Angustia" (anguish), which sets forth the viewpoints of Kierkegaard and Heidegger, and "Cosa" (thing), which distinguishes the philosophical, the (Argentine) juridical, and the economic uses of the term.

The other dictionary was César A. Guardia Mayorga's *Léxico filosófico* (Arequipa, Peru, 1941, 138 pages). A second edition was published in Arequipa in 1949 under the title *Terminología filosófica*. This work allots more space to Oriental subjects than does Ruiz Moreno's.

A work of the 1940s described as a dictionary of Argentine thought—Florencio J. Amaya's *Diccionario político, sociológico y filosófico argentino* (Mendoza, Argentina, Editorial Cuyo, 1946, 520 pages)—is more general than its title indicates. The philosophical articles are mainly subjective reflections (in the manner of Voltaire but more conservative) with occasional references to historic positions. The author's declared intention to produce sequels 6 and 12 years later (described on the title page of this book as Volumes II and III) was not carried out.

In 1947 appeared the anonymous *Pequeño diccionario de filosofía* (Buenos Aires, 156 pages), issued by Ediciones Centurión for use in conjunction with Emilio Gouiran's *Historia de la filosofía* (Buenos Aires, 1947), published by the same house. The *Pequeño diccionario* consists of two parts, one on philosophers from Peter Abelard to Xavier Zubiri, with indications of their dates and their principal works, and the other on philosophical terms, from Academia (the Academy) to *univoco* (univocal), with explanations ranging from 1 to 29 lines.

SPANISH—SPAIN. In José M. Rubert Candau's *Diccionario manual de filosofía* (Madrid, Editorial Bibliográfica Española, 1946, 658 pages) the main topics of philosophy are dealt with in extensive articles or groups

of articles, and the less important topics are given merely as entries with references to the main articles where they are treated. Thus, there are articles on being (5 cols.), supreme modes of being (21 cols.), and transcendental properties of being (27 cols.); the entry "Categorías supremas" refers the reader to the articles on supreme modes of being and on predicables and predicaments. This work deserves to be better known for its clear and systematic exposition of complex subject matter, especially on topics where its Catholic orientation is not a factor.

ITALIAN. Alfredo Galluccio's *Dizionarietto dei principali vocaboli filosofici* (Cava de' Tirreni, Italy, Editore Coda, 1942, 23 pages; 3rd ed., Naples, 1952) covers only topics. Most of the eight hundred articles in the third edition are only a few lines long and are intended to identify unfamiliar terms which students may encounter in their philosophical reading.

Another miniature dictionary is Paolo Rotta's *Dizionarietto filosofico* (Milan, Carlo Marzorati, 1944, 125 pages; 5th ed., 1953), which likewise covers only topics, including concepts, problems, and movements. Many of the almost five hundred articles in the fifth edition present Kant's ideas on the subject at hand.

FRENCH. Régis Jolivet, dean of the faculty of philosophy of the Catholic University of Lyon, produced the French contribution of the 1940s, *Vocabulaire de la philosophie* (Lyon, Emmanuel Vitte, 1942, 207 pages; 2nd–4th eds., 1946, 1951, 1957; Spanish translation, Buenos Aires, 1953). The articles are brief (4 lines for "Thomisme" but 53 for "Liberté" and 52 for "Nature"). A 17-page appendix presents a "tableau historique des écoles de philosophie," showing, in conventional groupings, the dates and (in 1–11 lines) the "écoles et doctrines" of about 250 philosophers from Zoroaster to Wittgenstein.

A book described in its foreword as a "dictionnaire abrégé" is Georges Barbarin's *L'Ami des heures difficiles; un consolateur et un guide* (Paris, Éditions Niclaus, 1946, 173 pages). The author presents conventional advice, constituting a popular philosophy or a popular psychology, on more than 130 problems of life—adversity, anxiety, despair, humiliation, injustice, pain, remorse, scandal, and seduction, among others. A seduced and betrayed woman is advised to look inward and find the Divine Friend in her own soul. The friend (*Ami*) mentioned in the title is not the book but God.

ENGLISH. The *Dictionary of Philosophy* (343 pages), edited by Dagobert Runes, was published at New York by the Philosophical Library in 1942. The list of 72 contributors included some outstanding American philosophers plus a few noted Europeans. When the work was published, 13 of the contributors—C. A. Baylis, A. C. Benjamin, E. S. Brightman, Rudolf Carnap, Alonzo Church, G. W. Cunningham, C. J. Ducasse, Irwin Edman, Hunter Guthrie, Julius Kraft, Glenn R. Morrow, Joseph Ratner, and J. R. Weinberg—declared their disapproval of it. Their statement, published in various periodicals including the *Philosophical Review* and *Mind*, read in part: "We objected to the publication of the work in its present form, and some of us made vigorous efforts to persuade Mr. Runes to delay publication until it had been very materially revised. These efforts were to no avail." They added that their own articles had been altered without their consent and that although they were listed as associate or contributing editors, they "feel obliged to make a public disavowal of any editorial responsibility for it."

Despite the important defects of this work, chiefly imbalance, there are many pithy, useful identifications, descriptions, and discussions in it, especially those by Church on topics in logic. Indeed, the collection of Church's contributions to the dictionary and their issuance in a separate volume on issues and methods in logic would be a worth-while project.

A new edition of the Runes dictionary has been issued every few years (16th ed., 1960); these are, however, essentially reprints, containing only minor variations from the first edition. At least one edition, or reprint, was issued overseas (Bombay, Jaico Publishing House, 1957).

Runes also edited *Who's Who in Philosophy*, Vol. I, *Anglo-American Philosophers* (New York, Philosophical Library, 1942, 193 pages), a biographical dictionary of over five hundred living thinkers, covering not only Americans and Britons but also Indians, Europeans who came to the United States or England during Hitler's regime, and others. A contemplated second volume, for other parts of the world, was not issued. An unusual feature of the work is the listing of numerous periodical articles, as well as the major books, written by the philosophers included. Thus, the entry on Dewey runs to over 650 lines, listing over 50 books and over 250 articles.

In 1946, Father William D. Bruckmann published the third of the four American dictionaries of philosophy of this period, a volume entitled *Keystones & Theories of Philosophy* (New York, Benziger Brothers, 230 pages). This work includes comprehensive explanations—from the standpoint of Catholic philosophy—of concepts from *abstractio* (abstraction), to *voluntas* (will), of theories

from absolutism to voluntarism, and of technical terms from *ab intrinseco–ab extrinseco* (from the intrinsic–from the extrinsic) to *ut sic* (as such). It also lists chronologically 121 philosophers with very brief indications of their view-points. The bulk of the work is devoted to concepts, only 19 pages being given to the individual philosophers.

Finally, John Dewey and Arthur F. Bentley, in an article in the *Journal of Philosophy* (Vol. 44, 1947, 421–434), "Concerning a Vocabulary for Inquiry Into Knowledge," presented what may, by a broad interpretation, be counted as a dictionary of philosophy. It is an array of ninety terms in alphabetical order, from *accurate, action, activity, actor, application,* and *aspect* near the beginning to *thing, trans* (as a prefix), *transaction, true, truth,* and *word* near the end. Although the entry for *mental* begins "This word is not used by us" and continues that the word usually "indicates a hypostatization arising from a primitively imperfect view of behavior," the remainder of the entry sanctions the use of the word for "emphasizing an aspect of existence." The entry for *real* reads: "Its use is to be completely avoided when not a recognized synonym for genuine as opposed to sham or counterfeit." The other entries show a similar striving for clarity and rigor.

A British dictionary of philosophy published in the 1940s is *A Rationalist Encyclopaedia: A Book of Reference on Religion, Philosophy, Ethics, and Science* (London, Watts, 1948, 633 pages; 2nd ed., 1950), by Joseph McCabe, a former priest. McCabe debunks Aquinas as bracketing "serfs and animals," Aristotle as having had almost no influence for several centuries and then a deleterious influence on science, Augustine as writing poor Latin, Avicenna as sensual and dissipated, Bacon as hypocritical, Bergson as using largely inaccurate scientific material, Buddha as unoriginal, and so on. He generally lauds philosophers who were agnostics or deists. Some of the topical articles, while equally tendentious, contain useful criticism.

GERMAN. The Kirchner work of 1886 as revised by Michaëlis in 1903 was the basis of the *Wörterbuch der philosophischen Begriffe*, by Johannes Hoffmeister (Leipzig, Felix Meiner, 1944, 776 pages; 2nd ed., 1955, 687 pages). The 1944 edition shows the influence of Adolf Hitler's regime. For example, the article "Volk" (folk) in the 1944 edition includes a lyrical exposition of the meaning of membership in a tight ethnic group and cites Hitler's *Mein Kampf* and Alfred Rosenberg's *Der Mythus des 20. Jahrhunderts*, but in the 1955 edition that exposition and those citations have vanished. The 1944 article "Rassenbiologie" (racial biology) does not appear in the later edition. The 1944 article "Demokratie" (democracy) says that pure democracy is impossible to achieve because it falsely assumes the equality of individuals; that statement is omitted in the 1955 edition. The article "Relativitätstheorie" in the 1944 edition refers to "der jüd. Gelehrte Einstein," but in the 1955 edition it says simply "Einstein"; the articles "Marxismus," "Spinozismus," and others show the same difference in the two editions.

In 1945 the Zurich firm of Rudolf Schaltegger published the first of the new German-language dictionaries of the decade, the *Ruscha Fachwörterbuch der Philosophie* ("Ruscha Dictionary of Technical Terms in Philosophy," 147 pages), in which the entries are, for the most part, a few lines long. The book would be of use to only the most elementary students.

Three years later Erwin Metzke published *Handlexikon der Philosophie* (Heidelberg, F. H. Kerle Verlag, 1948, 457 pages; 2nd ed., 1949). The wealth of topics it covers may be noted, for example, in the *L*'s, where one finds the articles "Leben" (life), with four meanings distinguished, three of them broken down into submeanings; "Lebensanschauung" (outlook on life), two meanings; "Lebensform" (form of life), two meanings; "Lebensgefühl" (feeling toward life), three meanings; "Lebenskraft" (vigor), two meanings, with cross references to "Vitalismus" (vitalism) and "Vitalität" (vitality); and "Lebensphilosophie" (philosophy of life), six meanings. A 138-page appendix consists of 1-line to 34-line identifications or brief accounts of almost two thousand philosophers, many of them living, with Americans well represented.

Walter Brugger, S.J., is the principal author of the *Philosophisches Wörterbuch,* prepared with the collaboration mainly of his colleagues at the Berchmans-Kolleg near Munich (Vienna, Herder Verlag, 1948, 532 pages). This work went into 11 editions published in various years to 1964; it was also translated into Italian (Turin, 1959) and Spanish (4th ed., Barcelona, 1964). Many of the more than two thousand articles contain bibliographical references, mostly to German works. The Catholic viewpoint from which the book was prepared is not conspicuous, and the topics are treated factually, with a minimum of controversial interpretation. An appendix of over one hundred pages (including an index of about two thousand names) presents an outline history of philosophy.

Six fascicles, covering *A* to *J*, of the *Philosophen-Lexikon* were issued in 1936–1937 by various publishers in Berlin, having been prepared under the editorship of Eugen Hauer, Werner Ziegenfuss, and Gertrud Jung. The

completed work was issued in 1949–1950 by Ziegenfuss, with the collaboration of Gertrud Jung, under the title *Philosophen-Lexikon: Handwörterbuch der Philosophie nach Personen* (2 vols., Berlin, Walter de Gruyter). Most of the articles contain biographical data about the individual covered, an indication of his contribution to philosophical thought, the titles (and years of publication) of his principal works or the principal collections of his works, and the titles of selected writings about him. Some articles, such as those on von Hartmann, Friedrich Schiller, and Unamuno, present significant quotations from their writings. For Karl Barth there are, atypically, only 3 lines of text, followed by a 24-line bibliography of his writings and a 12-line list of writings about him.

The two volumes of the Ziegenfuss work are remarkably comprehensive. They are also accurate and relatively cosmopolitan. Germans, it is true, get more space than others—for example, 5 pages for Benno Erdmann, who was Gertrud Jung's teacher, and 6 pages for Fechner, compared with 1 for Democritus and 3 for Socrates. A few Marxists also get disproportionate coverage—4 pages for Lenin and 5 for Marx—and contemporaries likewise are given some preference—for example, 6 pages for Berdyaev, compared with 1 for Bentham. One is surprised to see 5 pages devoted to the racist Houston Stewart Chamberlain. But Americans are given fairly good coverage—1 page for Peirce, 3 for Emerson, 3 for James, 2 for Dewey, and 1 for Royce.

A few of the articles in the Ziegenfuss work (for example, those on Nicolai Hartmann, 17 pages; P. A. Sorokin, 3 pages; and Erich Rothacker, 7 pages) were written by the subjects themselves.

HUNGARIAN. Volume I ("Aall" to "Avicebrón") of Pal Sandor's *Filozofiai Lexikon* (Budapest, Faust Kiadás, 64 pages) appeared in 1941. No further volumes seem to have been published. This is a biographical dictionary of philosophers with some emphasis on nineteenth-century and twentieth-century thinkers—Erich Adickes, four men named Adler, Samuel Alexander, and so on—and with considerable space devoted to selected great figures of the past—Anselm, Antisthenes, Aquinas, Aristotle (32 cols.), and others.

TURKISH. The *Felsefe ve Gramer Terimleri* ("Dictionary of Philosophy and Grammar," Istanbul, Cumhuriyet Basimevi, 1942, 318 pages), prepared by the Türk Dil Kurumu (Turkish Language Society), contains a series of alphabetical three-language lists of equivalent terms (Turkish, Osmanli, French; Osmanli, French, Turkish; and French, Osmanli, Turkish) and three corresponding lists of grammatical terms. (Osmanli is a Turkish dialect.) The philosophical lists usefully include over one thousand terms in cosmology and metaphysics—*causality, demiurge,* and so on; ethics—*altruism, deontology,* and so on; logic—*amphibology, contraposition,* and so on; and other domains of philosophy, plus terms in psychology—*abulia, claustrophobia,* and so on.

THE NINETEEN-FIFTIES

The flowering begun in the 1940s continued in the 1950s. Where the previous decade saw 22 new dictionaries of philosophy that have come to the writer's attention, 24 were published in the 1950s. Nine languages were represented: English, Gaelic, German, Dutch, French, Italian, Spanish, Portuguese, and Turkish. The great landmark of the 1950s is the monumental four-volume Italian encyclopedia of philosophy written by scholars at Gallarate.

ENGLISH. A philosophical dictionary vastly different from most is *The Great Ideas: A Syntopicon of Great Books of the Western World,* compiled under the direction of Mortimer J. Adler (Chicago, Encyclopaedia Britannica, 1952), comprising Volumes II and III of the publisher's 54-volume "Great Books of the Western World." It covers 102 "great ideas," including art, being, cause, chance, change, democracy, eternity, form, God, good and evil, idea, knowledge, logic, love, matter, metaphysics, mind, nature, necessity and contingency, one and many, reasoning, sense, sign and symbol, soul, space, time, truth, will, wisdom, and world.

For each idea the work presents an analytical and expository introduction, followed by a list of elements of the idea with a series of references to pertinent passages in the great books for each element. There is also a list of related great ideas and finally a list of additional readings on the subject in classics which are not included in the "Great Books" collection. At the end of the second volume of the *Syntopicon* there are a bibliography consolidating the lists of additional readings, a discussion of "syntopical construction" (which lists, among the ideas originally considered for inclusion but rejected, becoming, belief, deduction, doubt, essence, probability, purpose, reality, self, spirit, substance, value, and many others), and an "inventory" (index) of eighteen hundred terms.

A more self-conscious book could scarcely be imagined. Virtually every portion of the book is preceded by an explanation of why that portion was formed in the way in which it was formed and not otherwise. Critics are

answered before they have a chance to formulate criticisms. The reader is everywhere shown the scaffolding, and his attention is invited to a close inspection of its features.

Nevertheless, the book is highly useful. For the elements of the idea of form, for example, the reader is referred to specific passages in Plato, Aristotle, Lucretius, Augustine, Aquinas, Bacon, Descartes, Locke, Berkeley, Kant, Hegel, William James, and others. The analytical and expository introductions are for the most part general rather than technical, but they go as deeply into a subject as a thoughtful, educated reader may desire. All in all, this unique work was decidedly worth undertaking and was competently executed.

The only other English dictionary of philosophy published in the 1950s was Michael H. Briggs's *Handbook of Philosophy* (New York, Philosophical Library, 1959, 214 pages). It is difficult to see the usefulness of the article "Future," which reads, in its entirety, "Those events that will happen in time to come," or of the opening definition of the article "Change"—namely, "A constant alteration of states of the universe so that specific combinations of events do not persist." Several other articles in this handbook are equally unenlightening.

GAELIC. The *Focloir Fealsaimh* ("Vocabulary of Philosophy"), by Colmán O Huallacháin, O.F.M. (Dublin, An Clóchomhar, 1958, 169 pages), begins with a preface in French by Monsignor Louis de Raeymaeker of the University of Louvain. The book presents brief Gaelic descriptions or explanations of about two thousand Gaelic terms in philosophy and related humanistic disciplines, with the equivalent terms in German, English, French, and Latin. At the end of the book are four reciprocal word lists—German, English, French, and Latin—with the Gaelic equivalent of each word. The English word list includes not only such specifically philosophical terms as *Absolute, actual, aesthetics, agnostic,* and *aseity* but also such terms as *abnormal, acoustics, agoraphobia, anthropology*, and *atavism*.

GERMAN. In Germany and Switzerland five works were produced or begun, not counting a nonalphabetical so-called encyclopedia published in 1959. First, Carl Decurtins produced the *Kleines Philosophen-Lexikon* (Affoltern am Albis, Switzerland, Aehren Verlag, 1952, 312 pages), containing biographical sketches of three hundred individuals, among whom are not only the main figures in the history of philosophy strictly conceived but also Helena P. Blavatsky, Karl von Clausewitz, Lenin, Mussolini, the racists Chamberlain, Gobineau, and Alfred Rosenberg, as well as Jesus Christ, Dostoyevsky, Emerson, and Omar Khayyám. Chamberlain gets more space than Jesus Christ.

In 1954, Franz Austeda wrote the *Kleines Wörterbuch der Philosophie* (Frankfurt, Humboldt-Verlag, 188 pages; 2nd ed., entitled *Wörterbuch der Philosophie*, Berlin and Munich, Verlag Lebendiges Wissen, 1962, 270 pages). This work contains over eighteen hundred articles, including eight hundred which are biographical. It is a highly sensible and sound short reference work, with a reasonable proportion of space allotted to each of the standard topics in philosophy and the principal philosophers of the past and the present, as well as topics in less standard fields, such as Oriental philosophy, disciplines close to philosophy, and even old saws like Terence's "Homo sum; humani nihil a me alienum puto" ("A man am I; nothing human do I consider alien to me").

On behalf of the Kommission für Philosophie der Akademie der Wissenschaften und der Literatur zu Mainz, Erich Rothacker undertook a series of volumes under the general title *Archiv für Begriffsgeschichte; Bausteine zu einem historischen Wörterbuch der Philosophie* ("Archive for History of Concepts; Building Stones for a Historical Dictionary of Philosophy," Bonn, H. Bouvier, 1955–). Among the volumes which have appeared are Volume II (Part 2), *Kosmos* (1958, 168 pages), by Walther Kranz; Volume III, *Gewohnheit* ("Custom," 1958, 606 pages), by Gerhard Funke; Volume IV (1959, 239 pages), containing discussions by eight writers regarding various concepts or suggested texts of articles for the *Wörterbuch*; Volume V (1960, 718 pages), containing, under the headings "Absolut," "Abstrakt, Abstraktion," and "Aktivität, aktiv-passiv," the *Bibliographie deutscher Hochschulschriften von 1900–1955*, by Hans Flasche and Utta Wawrzinek; Volume VII (1962, 325 pages), containing discussions by a number of writers on such concepts as the Kantian *Analytik* and *Dialektik*; and Volume VIII (1963, 398 pages), by Karl Otto Apel, on the idea of language in the humanistic tradition from Dante to Vico. This is an ambitious and useful undertaking. Although it may not eventuate in an actual dictionary of philosophy, future writers of such dictionaries should feel obliged to utilize its findings.

In 1958, Max Müller and Alois Halder produced the paperback *Herders kleines philosophisches Wörterbuch* (Freiburg, Verlag Herder, 204 pages; 7th ed., 1965), with a bibliographical appendix citing various histories of philosophy and journals of philosophy and nine earlier dictionaries of philosophy. Portraits of Aristotle, Plato,

Augustine, Aquinas, Descartes, Leibniz, Kant, Hegel, Husserl, Bergson, Heidegger, and Jaspers appear on the back cover. The articles on medieval, modern, and contemporary thinkers are especially useful; Nicholas of Cusa is given 76 lines, Unamuno 34 lines, and Buber 28 lines.

The last of the German-language contributions of the decade is Volume II of *Das Fischer Lexikon, Enzyklopädie des Wissens*, a compilation entitled *Philosophie*, edited by Alwin Diemar and Ivo Frenzel (Frankfurt, Fischer Bucherei, 1958, 376 pages). This paperback book was reprinted in 1959 and 1960, and an English version translated by Salvatore Attanasio and prepared under the direction of James Gutmann was published as *Philosophy—A to Z* (New York, Grosset and Dunlap, 1963) in hardback and paperback editions. The collaborators consisted of 15 German authorities plus Paul K. Feyerabend of the United States. The work presents a small number of comprehensive articles—26—on such broad topics as anthropology, aesthetics, and Chinese and Japanese philosophy rather than a multitude of short ones. Historical information is given where necessary, but the emphasis is on concepts and problems. The articles show originality and penetration.

A nonalphabetical so-called encyclopedia was *Die Philosophie im XX. Jahrhundert: Eine Enzyklopädische Darstellung ihrer Geschichte, Disziplinen und Aufgaben*, edited by Frederick H. Heinemann ("Philosophy in the Twentieth Century; An Encyclopedic Presentation of Its History, Disciplines and Formulations," Stuttgart, Ernst Klett Verlag, 1959, 600 pages; 2nd edition, 1963). Heinemann begins with a discussion of the term *encyclopedia* which de-emphasizes the alphabetical order of topics, and he continues with chapters, written by himself or others, on Oriental, ancient, medieval, and modern philosophy; on movements in twentieth-century philosophy; and on epistemology, logic, philosophy of mathematics, metaphysics, philosophy of nature, and other philosophical disciplines. The treatment of the topics is mainly interpretive and constructive, rather than purely expository, especially in the chapters on the philosophical disciplines.

DUTCH. The Dutch work of this decade was edited by Johan Grooten and G. Jo Steenbergen. It is *Filosofisch Lexicon* (Antwerp, Standaard-Boekhandel, 1958, 331 pages), written by 32 collaborators, of whom the best known are perhaps E. W. Beth and Louis de Raeymaeker. The book begins with an explanation of how the topics are broken down, what type of spelling is used, how to find medieval names, and how the cross references are shown. The articles themselves are scholarly and well balanced.

FRENCH. Armand Cuvillier's *Nouveau Vocabulaire philosophique* (Paris, Librairie Armand Colin, 1956, 203 pages; 3rd ed., 1958) is a worthy successor to his *Petit Vocabulaire*, which went through 13 editions from 1925 to 1953. The new work includes a number of terms borrowed from other languages, such as *Erlehnis* (experience), *Dasein* (existence), and *pattern*. A number of articles, à la Goblot and Baldwin, set forth more than one meaning and then discourage the use of the term in one of the senses. For example, under "Empirique" (Empirical), the third meaning is "fondé sur l'expérience en général . . ." ("founded on experience in general"), but the author comments, "impropre au sens 3; dire expérienciel" ("improper in sense 3; say experiential"). A Spanish translation, entitled *Diccionario de filosofia*, was published at Buenos Aires in 1961.

J. Claude Piguet's *Le Vocabulaire intellectuel* (Paris, Centre de Documentation Universitaire et S.E.D.E.S. Réunis, 1957, 112 pages; reprinted, 1960, backstrip title, *Vocabulaire de philosophie*) disclaims being a dictionary in the sense of a list of pat definitions. It aims, instead, to stimulate students' thinking, partly by provocative opposition. For many terms an antonym is given, or two or more "opposites" are cited; for example, the article on absolute contrasts absolute with relative, and the article on duty contrasts duty not only with moral indifference but also with right. The book is probably of use mainly to students specializing in subjects other than philosophy.

ITALIAN. Of the seven Italian works of the period, three were published in 1951. Eustachio P. Lamanna and Francesco Adorno produced the *Dizionario di termini filosofici* (Florence, Felice le Monnier, 1951, 104 pages; 9th ed., 1960), in which the articles are brief, ranging from 1 line for "Verbo, (il)," ending with a cross reference to "Logos," to 47 lines for "Intelleto."

Giovanni Semprini compiled the *Nuovo dizionario di coltura filosofica e scientifica* (Turin, Società Editrice Internazionalc, 1951, 470 pages), which chiefly has articles on philosophical subjects, with errors in various articles on British and American philosophy, but also covers topics in the empirical sciences; for example, there are articles on anesthesia, clan, geology, and Mesmer.

Mario A. Boccalaro's *Dizionario filosofico* (Bologna, Licinio Cappelli, 1951, 91 pages) covers topics only. Its articles, generally a few lines long, are carefully and accurately phrased.

In 1952, Vincenzo Miano and 12 Italian collaborators produced the *Dizionario filosofico* (Turin, Società Editrice Internazionale, 1952, 693 pages), written with a Thomistic approach. Only topics are treated, but the appended "Schema della storia della filosofia" shows the name of the article in which each important philosopher is discussed; over 150 thinkers are included in the list.

Umberto Cantoro's *Vocabulario filosofico* (Bologna, Casa Editrice N. U. Gallo, 1955, 283 pages) began with an introduction on the philosophical disciplines and continued with an alphabetically arranged vocabulary which purportedly emphasized terms in common usage that have a special meaning in philosophy—for example, *absolute, concrete, and criticism*—but actually devoted most of its pages to the usual philosophical terms—*agnosticism, ambiguity, anguish, free will*, and the like. Psychology was taken by the author to be a philosophical discipline.

The *Dizionario di filosofia*, edited by Andrea Biraghi with contributions by 29 Italian collaborators (Milan, Edizioni di Comunità, 1957, 787 pages), is not strictly a dictionary since the materials in its two parts (on the history and problems of philosophy, respectively) are arranged in a nonalphabetical order, but it contains, as appendixes, three features which put it in the broad stream of dictionaries of philosophy: a dictionary of Greek terms, a dictionary of German terms, and a comprehensive alphabetical index.

The Gallarate landmark. In 1957 a group of Italian scholars in the Centro di Studi Filosofici di Gallarate, together with a few foreign collaborators, produced the *Enciclopedia filosofica* (4 vols., Venice and Rome, Istituto per la Collaborazione Culturale for the Ministry of Public Education and the Giorgio Cini and the Enrico Lossa foundations), which for the first time in half a century outshone the Baldwin work in comprehensiveness and up-to-date scholarship. The directing committee aimed to produce not "un mero *dizionario* filosofico" but a true encyclopedia of philosophy which would go beyond the dry explanation of the usages of terms and would present deeper analyses of the elements and implications both of individual problems and ideas and of more general points of view.

Each volume contains a number of full-page illustrations (mostly portraits of philosophers), and many of the articles contain bibliographical references at the end. This colossal work, totaling some 6 million words, is a basic landmark in the field of philosophical reference works, far outstripping its nearest competitors in magnitude. Physically, also, it is outstanding; the print and the 233 illustrations are not only tasteful but in some ways sumptuous. The work contains about twelve thousand articles, of which seven thousand are historical (on individual philosophers, movements, and the like) and five thousand are analytical (on concepts, problems, and the like). There are, for example, over 130 articles on past and present Russian philosophy, 82 on individual philosophical journals, over 80 on twentieth-century American philosophy, 74 on Indian philosophy, and 55 on subtopics of deduction and induction.

The contributors are mainly professors in Italian universities. Their contributions are factual, reliable, and broad in scope. The article on Aristotle (27 cols., with a full-page glossy reproduction of Raphael's head of Aristotle in the "School of Athens") is followed by articles on Pseudo-Aristotle (1 col.), Aristotle in Latin (2 cols.), and Aristotelianism (6 cols.), all of them rich in content and based on vast learning. There are worthwhile articles on neoclassicism, neocriticism, neo-empiricism, Neo-Guelphism, Neo-Hegelianism, Neo-Lutheranism, Neo-Malthusianism, Neo-Pythagoreanism, Neoplatonism, neopositivism (16 cols.), neorealism, Neo-Scholasticism, and neo-humanism.

There is some bias toward religious and idealistic positions in philosophy. Moreover, more Italian twentieth-century philosophers are treated in separate articles than either French or British. G. E. Moore gets only a column, which is less than the space assigned to Bernardo Varisco or Michele F. Sciacca, and a number of eminent American philosophers—Brand Blanshard, C. I. Lewis, Arthur O. Lovejoy, and R. W. Sellars—get less than a column.

The encyclopedia also goes far afield in including material on economics, pedagogy (with articles on *scoutismo*—the boy scouts—and on coeducation), and literary art (with articles on Joseph Addison, Sir Philip Sidney, and Jonathan Swift). Moreover, there are many minute articles which could profitably have been combined into more meaningful longer articles. However, weighing the encyclopedia's many merits against its few shortcomings, one must conclude that the work represents a highly laudable achievement, destined to be useful over a prolonged period.

SPANISH. Of the two Spanish-language dictionaries of philosophy produced in the 1950s, the first was published in Argentina and the second in Spain. Julio Rey Pastor and Ismael Quiles directed five editors and ten collaborators in the production of the *Diccionario filosófico* (Buenos Aires, Espasa–Calpe Argentina, 1952, 1,114

pages), in which the material is arranged according to a systematic outline of topics in 18 chapters instead of in alphabetical order. The 18 chapters are headed "Introducción á la historia de la filosofía"; "Lógica"; "Teoría del conocimiento"; "Epistemología y teoría de la ciencia"; "Logística, Lógica Simbólica o Lógica Matemática"; "Ontología," with 19 subheads, including "Ser," "Ente," "Existencia," and "Esencia"; "Metafisica general" (nature and structure of being and individuality); "Metafísica especial" (matter, life, mind, and spirit); "Filosofía de los valores"; "Filosofía de la religión"; "Ética"; "Estética"; "Filosofía del arte y poética"; "Psicología"; "Antropología filosófica"; "Concepción del mundo"; "Sociología"; and "Filosofía del derecho." At the end are the 45-page "Vocabulario filosófico," alphabetically arranged, and the 17-page "Equivalencias idiomáticas" (German–Spanish, English–Spanish, French–Spanish, and Italian–Spanish).

Juan Zaragüeta Bengoechea, director of the Luis Vives Institute of Philosophy in Madrid, is the author of the *Vocabulario filosófico* (Madrid, Espasa–Calpe, 1955, 571 pages), in which almost every article begins with the German, French, English, and Italian equivalents of the term being discussed. The terms are defined and explained from a scholastic point of view, generally without historical references. The articles are weak on contemporary philosophy, the one on *logística*, for example, merely setting forth in 20 lines what symbolic logic is about.

PORTUGUESE. Three Portuguese-language dictionaries of philosophy were published or were begun and dropped in the decade of the 1950s. Volume I (*A–D*) of the *Dicionário de filosofia*, by Orris Soares, was published at Rio de Janeiro in 1952 by the Instituto Nacional do Livro of the Ministério da Educação e Saúde. No other volumes have appeared. At the beginning of many of the articles are the equivalents of the term being covered in one or more of the following languages—Greek, Latin, French, Italian, English, and German. The article on Aristotle runs to more than 25 columns, with subtopics arranged alphabetically (for instance, "Aristóteles e a alma" and "Aristóteles e a astronomia"). To take the *D*'s for an example, there are useful articles on Dalton, Dante, Darwin, Descartes (15 cols.), Diogenes (four persons so named), Driesch, Duhem, Dühring, Duns Scotus, Durkheim, and others but none on Dewey.

Published at São Paulo were the first fascicle, covering the letter A, of the *Dicionario de filosofia*, by Luís Washington Vita, reprinted from the *Revista do Arquivo Municipal* (1950, 48 pages), and the *Vocabulário filosófico*, by Carlos Lopes de Mattos (Ediçôes Leia, 1957, 387 pages). Both cover only topics but include among the topics the philosophy of some individuals, in the articles on Aristotelianism, Averroism, and so forth. Vita modestly ascribes any errors which may appear in his work (of which no more has been published) to the fact that his is the first dictionary of philosophy in the Portuguese language; thus, he does not count the Voltairian 1938 work of Renato Kehl as a true dictionary of philosophy. Vita includes and Mattos excludes fields akin to philosophy. For many of his terms Mattos gives the equivalents in Esperanto, French, German, Greek, Italian, Latin, and Spanish and enumerates in the Bibliography 17 earlier dictionaries of philosophy.

TURKISH. Of Cemil Sena's *Büyük Filozoflar Ansiklopedisi* (Istanbul, Negioğlu Yayinevi, 1957–), only one volume, covering *A* to *D*, appeared. This work is a dictionary of philosophers which ranges from technical philosophers like Anaxagoras (12 cols.) to popular philosophers like Angelus Silesius and Will Durant, natural scientists like Ampère, and sociologists like Durkheim. The articles—some of them illustrated—are well balanced between biography and doctrine. Appended to Volume I are a glossary of Turkish philosophical terms with their French equivalents and an index of persons mentioned, showing, for example, 130 pages of the 642 pages in Volume I as containing references to Plato.

URMSON, ABBAGNANO, AND AFTER

ENGLISH. *The Concise Encyclopedia of Western Philosophy and Philosophers* (New York, Hawthorn Books, 1960, 431 pages), edited by James O. Urmson, contains over 150 articles on individual philosophers and about 65 articles on philosophical topics and schools. It also includes over one hundred full-page illustrations, mostly portraits of philosophers, of which eight are in color. It closes with an 11-page bibliography. Many outstanding contemporary British and American philosophers are among the 48 contributors. In the Preface the editor set forth his principles. Where it was difficult to summarize the views of a philosopher briefly, he was to be given enough space to make his position intelligible (six thousand words for Kant). More generally, it was deemed better to have fewer and longer articles than many short ones of doubtful utility. Philosophy was interpreted narrowly, excluding such popular topics as the philosophy of life. Eastern thinkers were excluded because, according to Urmson, they are philosophers in the popular, and not in the technical, sense. In recapitulation, however, he gives as the reason

for their omission the fact that "their achievement is not closely related to that of western philosophers." Exception could be taken to the former of these justifications for the omission of Oriental philosophy, but the addition of the latter makes it hard to object.

Although the articles in the Urmson work are not signed, the authorship of some has become known—for instance, the article on epistemology is by Gilbert Ryle, on ethics by R. M. Hare, on Heidegger by Walter Kaufmann, and on logic by D. J. O'Connor. The articles on epistemology and ethics display a freshness seldom found in encyclopedias; they are readable, free of academicism, informative, and challenging. Many other articles are also both brilliant and original. However, the article on Heidegger not only, with some justification, makes much of his welcome of Hitlerism but also, with less warrant, dismisses the fabric of his thought as comparable to the nonexistent clothes of Hans Christian Andersen's fairytale emperor.

Urmson's choice of topics is questionable. Although topics outside technical philosophy were to be excluded, Karl Marx is covered in an article of fifty-three hundred words, of which the first sentence is "Marx was not primarily a philosopher." Many of the contributors are themselves the subjects of articles, but one does not find any article on Gödel, Tarski, or, among thinkers of the past, Bayle or Voltaire. Among the topical articles one does not find any on belief, causation, error, existence, identity, necessity, philosophy of history, negation, self, or vitalism.

Another English work of the 1960s was Henry Thomas' *Biographical Encyclopedia* (New York, Doubleday, 1965, 286 pages). This is a work for the general reader, not for the specialist in philosophy. For example, the more than four hundred thinkers covered include a generous selection of poets (Horace, Omar, Byron, Shakespeare), social commentators (Benjamin Franklin, Oliver Wendell Holmes, Jr.), and theosophists (Annie Besant, Helena P. Blavatsky) but not Ayer, Carnap, Jaspers, Lovejoy, Meinong, Moore, Reichenbach, Ryle, or Schlick. The expositions and evaluations are likewise on a popular level.

Another popular biographical work is Thomas Kiernan's *Who's Who in the History of Philosophy* (New York, Philosophical Library, 1965, 185 pages). The expositions of the doctrines of some of the philosophers covered are naive. For example, Aquinas is said to have redirected Aristotelianism "towards truth and away from doubt," and Mill's inductive methods are said to be based on his "advocacy of the law of the uniformity of nature."

RUSSIAN. The year 1960 saw the first volume (A to "Diderot") of a new Russian dictionary of philosophy, the *Filosofskaya Entsiklopediya*, edited by F. V. Konstantinov and others (Moscow, "Soviet Encyclopedia" Publishing House). The second volume (covering "Disjunction" to "The Comic") of the four projected volumes was published in 1962. Volume I includes four articles—"Democracy," by L. Denisova; "Dialectics," by P. Kopnin; "Humanism," by L. Denisova; and "Dialectical Materialism," by A. G. Spirkin—which are available in English, the first three having been translated by William Mandel in the quarterly *Soviet Studies in Philosophy* (Vol. I, Spring 1963) and the fourth having been translated for *Russian Philosophy: A Book of Readings*, edited by James M. Edie and others (3 vols., Chicago, Quadrangle Books, 1965).

The article on democracy attempts to show that bourgeois democracy is dictatorship of the capitalist class, with illusory freedoms, whereas socialist democracy is dictatorship of the proletarian class, with genuine freedom of speech, the press, assembly, and demonstration. The truth is also labeled elsewhere in the encyclopedia, as in the article on absolute idealism, which is described as based on "the false assumption of the existence of an absolute idea." On the positive side may be mentioned the numerous good articles on logic, the broad coverage of both topics and persons (except that Bukharin and some other heretic Marxists are omitted), and the many halftone cuts. Such sociological topics as marriage are included.

Karl G. Ballestrem's *Russian Philosophical Terminology* (Dordrecht, D. Reidel, 1964, 116 pages) contains a glossary of about one thousand philosophical terms in Russian, with English, French, and German equivalents. Emphasis is placed on terms having a special use in Soviet philosophy.

ITALIAN. Nicola Abbagnano published the *Dizionario di Filosofia* (Turin, Unione Tipografico, 1961, 905 pages; Spanish translation, Mexico City and Buenos Aires, 1963) with the collaboration of Giulio Preti on topics in the field of logic. Abbagnano is a distinguished figure in contemporary philosophy and philosophical scholarship. His dictionary, covering only topics, shows vast erudition and commendable acumen in appraising tendencies and movements in philosophy. It gives, for example, a fair and thoroughly knowledgeable treatment to contemporary Anglo-American and positivistic philosophy. In the admiring words of Urmson, who noted a few inaccuracies in the Abbagnano work in a review in *Mind* (Vol. 71,

1962, 425), Abbagnano "refers as readily to the latest numbers of American journals as to the works of Plato."

Topics for which the standard name is in a language other than Italian—for example, *Erlebnis* (living experience), *Gegenstandtheorie* (object theory), and *Weltanschauung* (world outlook)—are treated by Abbagnano or cross-referenced in their regular alphabetical order. For many of the Italian words he also gives equivalents in Greek, Latin, English, French, and German. This work is one of the outstanding dictionaries of philosophy of our time. An English translation is scheduled to be published by the University of Chicago Press.

DUTCH. K. Kuypers is the editor of a Dutch work, *Elseviers Kleine Filosofische en Psychologische Encyclopedie* (Amsterdam, Elsevier, 272 pages), that appeared in 1960. Short but useful articles are presented on obscure as well as prominent thinkers and topics. Some topics—for instance, the Gifford lectures—are not often found in dictionaries of philosophy. Appended are a 15-page historical outline showing the schools or other groupings of over five hundred philosophers; a bibliography; and a selected list of philosophical journals and organizations.

DANISH. A work of this period is Henrik Thomsen's *Hvem Taenkte Hvad; Filosofiens Hvem-Hvad-Hvor* (Copenhagen, Politikens Forlag, 1961, 390 pages), with an introductory note by Justus Hartnack. The book contains a thumbnail history of philosophy from the pre-Socratics to Husserl, Wittgenstein, and Russell; numerous illustrations and two maps; a who's who of philosophy with illustrations of Augustine reading and Heidegger hiking; a dictionary of technical terms; and a bibliography.

GERMAN. Joseph Münzhuber wrote the *Kleines Wörterbuch der Philosophie, zum Gebrauch an Schulen* (Düsseldorf, Pädagogischer Verlag Schwann, 1962, 45 pages). This work contains about 135 articles ranging from the 2-line "Transintelligibel" to the 48-line "Existenzphilosophie." Among the more unusual articles are "In-der-Welt-Sein" (being-in-the-world) and "Unschärferelation" (Heisenberg's uncertainty relation).

Anton Neuhäusler wrote *Grundbegriffe der philosophischen Sprache: Begriffe viersprachig* (Munich, Ehrenwirth Verlag, 1963, 276 pages). The length of the article on any topic covered by Neuhäusler is based not on the topic's importance but on its "Klärungs-bedurfigkeit und schwierigkeit" ("need and difficulty of explanation"). Each entry includes the English, French, and Italian equivalent of the term; an indication of the origin of the term (if this is relevant); a sophisticated but clear discussion of the use of the term; and a brief bibliography. An appendix presents a decimal classification of philosophical concepts—for example, 1 for philosophy itself; 11 for metaphysics; 11.1 for ontology; 111.11 for existence, *Dasein*, and reality; 19 for history of philosophy; 2 for theology.

In 1964 there appeared another *Philosophisches Wörterbuch* (Leipzig, VEB Bibliographisches Institut, 650 pages; reprinted 1965), edited by Georg Klaus and Manfred Buhr. It was a joint project of the Institute for Philosophy of the German Academy of Sciences in Berlin and the professorial chair for philosophy of the Institute for Economics of the Central Committee of the German Socialist Unity party. The Marxist–Leninist slant is sometimes blatant, as in the article "Demokratie," where bourgeois democracy is characterized as a form of government in which everything is subordinated to profit. Among the examples presented to illustrate the use of "is" in the article "Kopula" are (in translation): "Marx is the author of *Capital*" and "Marx is one of the greatest thinkers of mankind."

FRENCH. In 1962, Paul Foulquié, with the collaboration of Raymond Saint-Jean, produced the *Dictionnaire de la langue philosophique* (Paris, Presses Universitaires de France, 1962, 776 pages), which, as the Preface states, is heavily indebted to Lalande's work. Since Foulquié's is a dictionary of concepts, there are no articles on schools or viewpoints, such as Aristotelianism and Eleaticism. Although the basic arrangement is alphabetical, related concepts are in some cases grouped around a generic term—for example, *étant, entité, essence, exister*, and *existentialisme* around *être*. Many of the articles quote texts to support the definitions presented.

The anonymous *Dictionnaire des philosophes* (Paris, Collection Seghers, 1962, 383 pages; binder's title, *Dictionnaire illustré des philosophes*) contains biographical statements regarding approximately six hundred standard Western philosophers and philosophic thinkers, such as Ruth Benedict, Karen Horney, and Kurt Lewin, in allied fields. There follow references to about thirty Oriental thinkers and a vocabulary of some five hundred terms, most of them defined in a few lines. Scattered in the book are 64 portraits.

According to Didier Julia, the purpose of his *Dictionnaire de la philosophie* (Paris, Librairie Larousse, 1964, 320 pages) is the disclosure of eternal truths as being applicable to daily life. In keeping with that purpose, the illustrations are popular: an abstract painting, a scene in Paris after the explosion of a plastic bomb, a child peer-

ing through curtains (illustrating "Attention"), a Buddhist immolating himself at Saigon, and others. Marx gets more space than anyone else, and Trotsky gets more than Aristotle. Maimonides and Peirce are among the omissions. It is doubtful that the announced purpose of the work was achieved.

SPANISH. Paul Henri Boyer's *Diccionario breve de filosofía* (Buenos Aires, Club de Lectores, 1962, 187 pages) has some material of questionable validity. There is only one article on Oriental philosophy, on nirvana, which is wrongly defined as negation of the will to live. The spelling of non-Spanish names in the work is not reliable.

ORIENTAL LANGUAGES. Three Asian countries—nationalist China, Japan, and, most notably, Korea—have made significant contributions in the 1960s.

Chinese. The Chinese dictionary is *Chê Hsüeh Ta Tz'ŭ Tien* ("Comprehensive Dictionary of Philosophy," Taipei, Ch'i Ming Shu Chû, 1960, 464 pages), containing about one thousand five hundred articles, each printed with the equivalent of the term in at least one Western language. The first entry is on monism, and the last is on "ideal-realism." The rest cover the standard philosophical and psychological topics and personalities plus such unusual topics as dilemmatic proposition and "summists" (authors of works entitled *Summa*). Two indexes in Western languages (and roman type) list topics and personal names.

Japanese. Naomichi Takama's *Tetsugaku Yogo No Kiso Chishiki* ("Philosophical Terminology," Tokyo, Seisun Shuppan Sha, 242 pages), a Japanese work, was published in 1961. The title of each article is given with English and German equivalents. There are articles on patriotism, happiness, justice, human nature, freedom of the will, suicide, space, time, dialectical materialism, scholastic philosophy, and many other popular and technical topics. Some of the articles show an undue influence of Marxism.

Another Japanese work was edited by Yasumasa Oshima—*Shin Rinri Jiten* ("Dictionary of Ethics," Tokyo, Sobun Sha, 1961, 472 pages). The scope of this work is broader than its title indicates. Some of the articles are on ethical subjects, including agape, evil, ataraxia, will, Epicureanism, and human rights, but others transcend the domain of ethics, including those on atman, Aristotle, either–or, a priori, causality, Eleatics, entelechy, and *Dasein.* In general, this is the more scholarly of the Japanese works.

Korean. One hundred and four Korean scholars worked on the *Dictionary of Philosophy: Ch'ŏrhak Taesajŏn* (Seoul, Hagwŏnsa, 1963, 1,376 pages). A monumental job of scholarship and printing, this work contains, for many entries, the Korean expression followed by equivalents in other pertinent languages, the article in Korean with romanized transliterations where needed, and a bibliography. The field covered includes philosophy, psychology, and sociology, and the articles are of exceptionally high quality. Among the added features are about four hundred pictures of philosophers; other illustrations, including Wittgenstein's duck-rabbit and four full-page maps; a uniquely rich year-by-year chronology of philosophy, showing, for example, 1905 as the year of the inauguration or publication of specific works by 22 philosophers; and an index of about five thousand terms in Western languages.

A "NONENCYCLOPEDIA." From time to time we have paused to poke a curious finger into works which are called dictionaries or encyclopedias of philosophy but which are not arranged alphabetically. The latest of these is Ramón Conde Obregón's *Enciclopedia de la filosofía* (Barcelona, De Gassó Hernanos, 1961, 363 pages). The first four parts of the book are on philosophy in general, prephilosophy, Western philosophy, and Oriental philosophy; the fifth is headed "Conclusion." Conde's work will probably not be the last, in the march of philosophical exposition, to exploit the perennial intellectual magnetism of the term *dictionary or encyclopedia.*

DICTIONARIES OF SPECIAL PHILOSOPHICAL TOPICS

There are dictionaries which cover one or more philosophical disciplines, periods, and schools, as well as individual philosophers. The listings presented here are merely illustrative; complete coverage is not attempted.

DISCIPLINES. Some dictionaries cover a single discipline, such as aesthetics, ethics, logic, or theology; others cover a combination, such as ethics and theology or logic and philosophy of science.

Aesthetics. Among the dictionaries of aesthetics is Ignaz Jeitteles' *Aesthetisches Lexikon: Ein alphabetisches Handbuch zur Theorie der Philosophie des Schönen und der Schönen Künste* ("Dictionary of Aesthetics: An Alphabetical Handbook of the Theory of the Philosophy of Beauty and the Fine Arts," 2 vols., Vienna, Carl Gerold, 1835–1837). This is a capably written reference work, covering numerous topics in architecture, the dance,

drama, drawing, music, painting, poetry, rhetoric, sculpture, and other arts, as well as topics applicable to natural beauty or to more than one of the arts. An 84-page appendix reviews the classic literature on aesthetics.

In 1946 Roger Caillois produced the *Vocabulaire esthétique* (Paris, Éditions de la Revue Fontaine, 141 pages). In addition to whole chapters on nature and art, this work contains articles on art for art's sake, authority, image, order, originality, sincerity, and other topics in nonalphabetical order. Each article is a discursive essay rather than a systematic treatment.

A curiosity among dictionaries of aesthetics is Paolo Mantegazza's *Dizionario delle cose belle* (Milan, Fratelli Treves, 1891, 346 pages; German translation, 2 vols., Jena, 1891–1892). After an introduction on elements of beauty (color, symmetry, and so on) constituting about a third of the book, the author presents over one hundred articles in alphabetical order on "beautiful things"—alabaster, eagle, gazelle, jasmine, lark, lion, moon, snow, stars—with rhapsodic comments on each.

Ethics. Among the dictionaries of ethics, mention may be made of two in particular. The first is *Dictionnaire des passions, des vertus, et des vices* (2 vols., Paris, Chez Vincent, 1769), published anonymously by Antonio F. Sticotti and Antoine Sabbatier. Discussing such topics as abasement, abominable, admiration, and adultery near the beginning of the alphabet and urbanity, utility, vivacity, and volition near the end of the alphabet, the authors epitomized the comments of famous writers—Aristotle, Bacon, Confucius, Diderot, Locke, Pascal, Voltaire, and others—on these topics.

In 1956 Vergilius Ferm's *Encyclopedia of Morals* (New York, Philosophical Library, 682 pages) appeared. The contributors to this scholarly and well-balanced volume include Lewis White Beck on Nicolai Hartmann; William K. Frankena on Ross, Sidgwick, and moral philosophy in America; Lucius Garvin on major ethical viewpoints; Walter Kaufmann on Freud, Goethe, Hammurabi, and Nietzsche; George L. Kline on current Soviet morality; Clyde Kluckhohn on Navaho morals; Swami Nikhilananda on Hindu ethics, and Frederick Sontag on Socrates, Plato, and Aristotle. Most of the articles are of substantial length and rich in content; some are a bit pedestrian.

Logic. In logic there is a Spanish *Vocabulario de Lógica*, by Baldomero Diez y Lozano (Murcia, Spain, Imp. Lourdes, 1925, 198 pages; 2nd ed., 1928), which contains about five hundred articles covering not only topics in traditional logic, such as absurd, affirmation, a fortiori, but also topics in related philosophical fields, such as change, causality, phenomena, tree of Porphyry. Given the brevity of the articles, the treatment is necessarily superficial, but the identifications of the more obscure terms are useful.

Theology. Dictionaries of theology are fairly numerous. Among them, some warrant special mention.

From 1908 to 1914 was published a work edited by Samuel M. Jackson and others, *The New Schaff–Herzog Encyclopedia of Religious Knowledge* (13 vols., New York and London, Funk and Wagnalls; reprinted, Grand Rapids, Mich., Baker Book House, 1949–1950). This work was based on the nineteenth-century works in this field edited by Philip Schaff and Johann J. Herzog. The Preface lists numerous preceding Catholic, Protestant, Anglican, Jewish, Muslim, and other theological dictionaries. More of the articles are on individuals—prophets, religious leaders, and theologians—than on topics. Most articles of philosophical interest, such as those on dualism, duty, ethics, freedom of the will, gnosticism, philosophy of religion, positivism, probabilism, Stoicism, utilitarianism, and others, as well as on individual philosophers, were written by specialists in religion; a few, however, such as those by Troeltsch on British moralists, deism, the Enlightenment, idealism, and so on, are philosophically penetrating. The 13-volume work was condensed and brought up to date in the *Twentieth Century Encyclopedia of Religious Knowledge*, edited by Lefferts A. Loetcher (2 vols., Grand Rapids, Mich., Baker Book House, 1955).

Joseph Bricout edited the *Dictionnaire pratique des connaissances réligieuses* (7 vols., Paris, Letouzey et Ané, 1925–1933). In this Catholic-sponsored work the articles of philosophic interest—prepared mostly by professors at the Séminaire des Missions located at Vals in southern France—include those on aesthetic sense, agnosticism, atheism, belief, categories, criteria of truth, deism, doubt, efficient cause, empiricism, and others, plus about 230 articles on philosophers, theologians, and schools of thought. The articles on non-Catholic viewpoints are factual and fair.

Joseph Höfer and Karl Rahner edited the *Lexikon für Theologie und Kirche* (10 vols., Freiburg, Verlag Herder, 1957–1965), a revision of the work of the same title, edited by Michael Buchberger (10 vols., 1930–1938), which was itself referred to as the second edition of Buchberger's two-volume *Kirchliches Handlexikon* (Munich, Allgemeines Verlags-Gesellschaft, 1907–1912).

The work on the philosophical articles was coordinated by Bernhard Welte of Freiburg. The Catholic view-

point is supported throughout, but the presentation of other viewpoints is informative.

The dictionary edited by Everett F. Harrison, *Baker's Dictionary of Theology* (Grand Rapids, Mich., Baker Book House, 1960, 566 pages), includes articles on movements of theological thought—for example, Calvinism, Lutheranism, and Thomism—but none on individual thinkers as such. Only those philosophical topics which are theological in a strict sense are dealt with. The orientation is that of sophisticated fundamentalism.

In 1962 was begun *A Catholic Dictionary of Theology* (Edinburgh, Thomas Nelson and Sons), edited by Monsignor H. Francis Davis and others. One volume of the four projected volumes has been issued thus far. Very Reverend Ivo Thomas is among the editors, and Father F. C. Copleston is among the better-known contributors. Instead of the usual prosaic and often uninspired articles on individual thinkers, Volume I contains articles on special features, such as Augustine and his influence, Berkeley and Catholicism, and the system of Boscovich. The writing is lively, and the authors do not hesitate to propound new theories.

Dictionaries or encyclopedias of specific religions and denominations are also available and contain articles on theological and even general philosophical topics. Several of these sectarian dictionaries of philosophy are outstanding.

The Jewish Encyclopedia, edited by Cyrus Adler and others (12 vols., New York and London, Funk and Wagnalls, 1901–1906; reprinted in various years), contains rewarding articles on Aristotle in Jewish literature, the influence of Arabic philosophy on Judaism, Maimonides (21 cols.), Spinoza (17 cols.), ethics, theology, and numerous other topics of philosophical relevance.

The Universal Jewish Encyclopedia, edited by Isaac Landman (10 vols., New York, Universal Jewish Encyclopedia, 1939–1943), had significant contributions by Isaac Husik, perhaps the greatest historian of medieval Jewish philosophy. This encyclopedia is a worthy successor to *The Jewish Encyclopedia*.

The Catholic Encyclopedia, edited by Charles G. Herbermann and others (16 vols., New York, Robert Appleton, 1907–1912; reprinted, 1913; supp., 1917, 1922, 1954), contains over five hundred articles on cosmology, theology, metaphysics, epistemology, logic, ethics, and individual philosophers. The articles expound these topics with clarity and vigor. Noteworthy contributors include Émile Bréhier, Pierre Duhem, and Maurice de Wulf. Comparable works exist in French, German, and Italian.

Of projected works the *New Catholic Encyclopedia* being edited at the Catholic University of America, Washington, D.C., will devote about 1 million of the total of 14 million words to subjects pertinent to philosophy. *Je sais, je crois: Encyclopédie du catholique au XXeme siècle*, edited by Henri Daniel-Rops (Paris, Librairie A. Fayard, 1956–), is scheduled to comprise 150 volumes (more than 130 have been published as of 1965); it is being translated into English as *The Twentieth Century Encyclopedia of Catholicism* (New York, Hawthorn Books). It is arranged by topic rather than alphabetically. Among the volumes of philosophical interest are Claude Tresmontant's *Les Origines de la philosophic chrétienne*, Vol. XI (1962), Philippe Delhaye's *La Philosophie chrétienne au moyen âge*, Vol. XII (1959), and Régis Jolivet's *L'Homme métaphysique*, Vol. XXXV (1958).

Theology and ethics. Of the dictionaries that cover two philosophical disciplines, chief among those covering theology and ethics is the *Encyclopedia of Religion and Ethics*, edited by James Hastings and others (13 vols., Edinburgh and New York, T. and T. Clark and Charles Scribner's Sons, 1908–1926; reprinted in whole or in part in various years). This is one of the great encyclopedias of all time. In conception it is original and imaginative; in execution, apt. The choice of topics is sagacious; the research has weathered the test of time; the analyses are thorough and penetrating. Among the philosophical contributors are John Burnet on the Academy, skeptics, and Socrates; C. D. Broad on reality and time; A. F. R. Hoernlé on solipsism; the Reverend William R. Inge on logos and Neoplatonism; Rufus M. Jones on mysticism; John Laird on will; J. M. E. McTaggart on personality; John H. Muirhead on ethics and rights; Josiah Royce on axiom, error and truth, mind, monotheism, negation, and order; F. C. S. Schiller on humanism, pragmatism, spiritualism, and values; A. E. Taylor on identity and theism; Erust Troeltsch on idealism and Kant; Frederick J. E. Woodbridge on Hobbes, Hume, and pluralism; and Maurice de Wulf on aesthetics and beauty. The orientation in the articles on religion is generally that of liberal Protestantism, but opposing points of view are presented fairly. The bibliographies are compact and useful.

Also deserving of mention as covering both theology and ethics is *A Dictionary of Religion and Ethics*, edited by Shailer Mathews and Gerald B. Smith (New York, Macmillan, 1921, 513 pages), which had as contributors Franz Boas, Edgar J. Goodspeed, Rufus Jones, Eugene Lyman, George Herbert Mead, Roscoe Pound, James B. Pratt, James H. Tufts, and others. For less important topics the articles present dictionary-type definitions or

identifications and little more. Imbalance in some of the articles may be illustrated by the fact that the 800-word article on Aristotle presents only one sentence on his ethics. There is a bibliography at the end, containing almost two thousand items.

Logic and philosophy of science. Major topics of another pair of philosophical disciplines—logic and the philosophy of science—are covered, though inadequately, in the *Harper Encyclopedia of Science*, edited by James R. Newman (4 vols., New York, Harper and Row, 1963), which had Ernest Nagel as its consultant on the philosophy and history of science. Among the contributors to the Newman work besides Nagel were Max Black, Irving M. Copi, Arthur C. Danto, and Milton K. Munitz. However, the philosophical articles are for the most part excessively brief. Exceptions include those on logic (four thousand words) and logical empiricism (almost five hundred words).

PERIODS. There are dictionaries covering the philosophy of specific periods, including, for example, the *Lexicon Philosophicum Graecum*, by Rudolf Goclenius the elder (Marburg, Rudolf Hutwelcker, 1615, 390 pages; 2nd ed., Frankfurt and Paris, S. Celerius, 1634), in which the terms defined are in Greek and the definitions and explanations are in Latin. Sources used by the author include the Greek philosophical classics, the New Testament, and the writings of the Greek Fathers of the Church.

The *Index zu philosophischen Problemen in der klassischen griechischen Literatur*, by Georg T. Schwarz (Bern, Francke Verlag, 1956, 109 pages), is a list of about 280 topics, such as being, definition, democracy, good, idea, life, love, philosophy, and reason, with an indication of where and how each one is discussed in pre-Aristotelian Greek literature and philosophy. Its limited objective is well carried out.

The Dictionnaire de philosophie et de théologie scolastique, ou Études sur l'enseignement philosophique et théologique au moyen âge, by Frédéric Morin, is included in the *Encyclopédie théologique*, edited by J. P. Migne (168 vols. in 170 in 3 series, Paris, 1844–1866), as Volumes XXI and XXII (1856–1857) of the third series. This dictionary covers adequately the medieval Scholastics, the main Arabic thinkers (but no Jewish philosophers), and the more important topics, problems, and movements of medieval philosophy. (The Migne encyclopedia is an unsystematic collection of dictionaries of aspects of religion—the Bible, church history, liturgy, saints, and so on.)

SCHOOLS. Movements or schools in philosophy are covered by various works. Among these is *A Biographical Dictionary of Modern Rationalists*, by Joseph McCabe (London, Watts, 1920, 934 pages). Rationalists are defined here as those who "uphold the right of reason against the authority of Church or tradition." Included are biographies of philosophers—for example, Bergson, Bradley, Lovejoy, and Moore; statesmen—for example, John Adams and Clemenceau; writers—for example, Balzac and Keats; musicians, artists, scientists, inventors, historians, sociologists, and so on.

Another school is covered in the *Dictionary of Scholastic Philosophy*, by Bernard Wuellner (Milwaukee, Wis., Bruce Publishing Co., 1956, 138 pages). Many of the articles are merely definitions. For example, the article on belief consists simply of the synonym *faith* and the article on faith gives only dictionary-type definitions of *faith* and *divine faith*, with references to two works of Aquinas. However, the book contains 33 interesting diagrams and charts, which show the subdivisions of act and potency, the categories of being, the kinds of evil, and the like.

A Concise Dictionary of Existentialism, edited by Ralph B. Winn (New York, Philosophical Library, 1960, 122 pages), contains quotations from six thinkers—Kierkegaard, Heidegger, Jaspers, Marcel, Sartre, and de Beauvoir—on anguish, being, boredom, choice, encounter, and other topics. Some of the quotations are epigrammatic; others are more extensive.

PHILOSOPHERS. Dictionaries devoted to the thought of individual philosophers are numerous. Aquinas, Aristotle, Bonaventure, Kant, Hegel, Maimonides, Plato, Russell, Schopenhauer, Spinoza, Teilhard de Chardin, and Wolff are among the main figures having special dictionaries devoted to their work. Aristotle, for example, is covered by four works.

First was Hermann Bonitz' *Index Aristotelicus* (Berlin, G. Reimer, 1870, 878 pages), which was reprinted from Volume V of the Academia Regia Borussica edition of Aristotle (5 vols., Berlin, G. Reimer, 1831–1870), with Greek texts edited by Immanuel Bekker. The index was reprinted in 1955 by the Akademie-Verlag in East Berlin. It is a complete concordance, indispensable to Aristotle scholars working with the original Greek.

Matthias Kappes' *Aristoteles-Lexikon* (Paderborn, Germany, Ferdinand Schöningh, 1894, 70 pages) contains a discussion in German of about four hundred Greek words used by Aristotle, with references to the main passages where those words play a part in his philosophy. On the basis of the 11-volume Oxford translation of Aristo-

tle, Troy W. Organ's *An Index to Aristotle in English Translation* (Princeton, N.J., Princeton University Press, 1949, 181 pages) covers about four thousand English words, from Abdera, abdomen, and abortions to Zeno, Zeus, and zodiac, with references to the passages where they significantly occur.

In 1962 there appeared the *Aristotle Dictionary*, edited by Thomas P. Kiernan (New York, Philosophical Library, 524 pages), which has passages from Aristotle's writings, translated by H. E. Wedeck and others. It begins with a 161-page summary of the individual writings of Aristotle and continues with quotations under alphabetically arranged topic headings. The quotations chosen are not always apt; for example, the five sentences quoted under "Form" do not represent Aristotle's philosophy of form.

Plato, Aquinas, and Kant are similarly covered by three or more dictionaries each; one of the Kant dictionaries is in Russian.

CONCLUSION

In the past it was possible for a scholar to encompass in a lifetime of learning the whole of a broad domain of human interest, such as philosophy. It was possible for one person to read all the important sources, major interpretations, and critiques of the sources. One could then write a thorough, well-balanced, and accurate dictionary of philosophy for his less knowledgeable colleagues.

However, with the democratization of education and the spread of intellectual activity the philosophical sources and the critical works have become too voluminous for a single individual to master. The truly comprehensive study of what philosophers have thought and said has therefore necessarily become a cooperative venture. Although some commendable dictionaries of philosophy have been produced by great scholars singlehandedly in the twentieth century, the scholarship of a single individual is, after all, limited.

Periodically, therefore, the need arises for expert summaries and appraisals of the philosophical books and articles that rush from the presses. Thus, cooperative summings up have appeared with some regularity. This *Encyclopedia of Philosophy* is intended to provide a new, more inclusive treatment of a wide variety of philosophical topics and to be a repository of up-to-date, detailed scholarship for the use of researchers and creative philosophers alike.

Bibliography

ON GENERAL ENCYCLOPEDIAS

Especially useful are the articles on encyclopedias which appear in the *Encyclopaedia Britannica*, 11th ed., 29 vols. (Cambridge and New York, 1910–1911); in the *Encyclopedia Americana*, 30 vols. (New York, 1966); and in *Chambers's Encyclopaedia*, 15 vols. (London, 1959).

OLDER LISTS

Some 20 dictionaries of philosophy, of ethics, or of individual philosophers are listed in Johann A. Fabricius, *Abriss einer allgemeinen Historie der Gelehrsamkeit*, 3 vols. (Leipzig, 1752–1754), Vol. I, p. 422. Shorter, evaluative lists of dictionaries of philosophy appear in the Preface to the first edition of Franck's *Dictionnaire des sciences philosophiques*, 6 vols. (Paris, 1844–1852), and in the Preface to Larousse's *Grand Dictionnaire universel du XIXe siècle*, 15 vols. (Paris, 1865–1876), Vol. I, pp. xli–xlii.

TWENTIETH-CENTURY LISTS

The section "Dictionnaires de philosophic" in Jean Hoffmans, *La Philosophie et les philosophes: Ouvrages généraux* (Brussels, 1920), pp. 1–4, lists, along with 11 general dictionaries of philosophy, about 80 specialized dictionaries, on topics like ethics and aesthetics. See also the lists of philosophical dictionaries in the Alcalde *Diccionario* (Madrid, 1927), Vol. I, pp. 12–15; the section on dictionaries in the *Allgemeine philosophische Bibliographie*, by I. M. Bocheński and Florenzo Monteleone (Bern, 1948), pp. 32–33; the section on dictionaries in Carmelo Ferro's *Guida storico-bibliografica allo studio della filosofia* (Milan, 1949), pp. 187 ff.; and the section "Philosophie" in the *Index Lexicorum*, by Gert A. Zischka (Vienna, 1959), pp. 40–43.

DISCUSSIONS

André Lalande, in "Les Récents Dictionnaires de philosophie," *Revue philosophique de la France et de l'étranger*, Vol. 56 (1903), 628–648, and Frederick H. Heinemann, in "Die Aufgabe einer Enzyklopädie des XX. Jahrhunderts," pp. 1–22 of his compilation *Die Philosophie im XX. Jahrhundert*, 2nd ed. (Stuttgart, 1963), provide thoughtful comments. Also provocative is Benedetto Croce's "Un Vocabolario della lingua filosofica italiana," *La Voce*, Vol. 1 (1909), 42; in it he urged the need for historical and analytical work on philosophical terminology which would not be a mere "dizionario filosofico," with its dismembered alphabetical order; he suggested, rather, a work like "una enciclopedia filosofica," having some of the attributes of Hegel's *Encyklopädie*. The editors of the Gallarate encyclopedia, in offering more than "un mero *dizionario filosofico*" may have had Croce specifically in mind.

William Gerber (1967)

PHILOSOPHY DICTIONARIES AND ENCYCLOPEDIAS [ADDENDUM]

The English-language citations in this update are organized alphabetically by book title. The non-English language citations are grouped initially as either Asian or European, and are listed within those two categories by specific language; the citations appear alphabetically by book title within each specific language's listing.

ENGLISH LANGUAGE

Adler's Philosophical Dictionary. Adler, Mortimer Jerome. Norwalk, CT: Easton Press, 1995.

American Philosophers Before 1950. Dematteis, Philip Breed, and Leemon B. McHenry. Detroit: Gale Group, 2003.

American Philosophers, 1950–2000. Dematteis, Philip Breed, and Leemon B. McHenry. Detroit: Gale Group, 2003.

Biographical Dictionary of Twentieth Century Philosophers. Brown, Stuart C., Diané Collinson, and Robert Wilkinson. London; New York: Routledge, 1996.

Biographical Encyclopedia of Philosophy. Thomas, Henry. Garden City, NY: Doubleday, 1965.

The Blackwell Dictionary of Western Philosophy. Bunnin, Nicholas, and Jiyuan Yu. Malden, MA: Blackwell Pub., 2004.

British Philosophers, 1500–1799. Dematteis, Philip Breed, and Peter S. Fosl. Detroit: Gale Group, 2002.

British Philosophers, 1800–2000. Dematteis, Philip Breed, Peter S. Fosl, and Leemon B. McHenry. Detroit: Gale Group, 2002.

The Cambridge Dictionary of Philosophy. Audi, Robert. Cambridge, U.K.: Cambridge University Press, 1995.

The Cambridge Dictionary of Philosophy. 2nd ed. Audi, Robert. Cambridge, U.K.: Cambridge University Press, 1999.

Christian Philosophy A–Z. Hill, Daniel J. Edinburgh: Edinburgh University Press, 2005.

Companion Encyclopedia of Asian Philosophy. Carr, Brian, and Indira Mahalingam. London: Routledge, 1997.

Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences. Grattan-Guinness, I. ed. London: Routledge, 1994.

A Companion to Aesthetics. Cooper, David Edward, Joseph Margolis, and Crispin Sartwell. Oxford: Blackwell Reference, 1995.

A Companion to Metaphysics. Kim, Jaegwon, and Ernest Sosa, eds. Oxford: Blackwell Reference, 1995.

A Concise Dictionary of Indian Philosophy: Sanskrit-English. Grimes, John A. Dr. S. Radhakrishnan Institute for Advanced Study in Philosophy. Madras: University of Madras, 1988.

A Concise Dictionary of Indian Philosophy: Sanskrit Terms Defined in English. Grimes, John A. New and rev. ed. Albany, NY: State University of New York Press, 1996.

A Concise Encyclopedia of Early Buddhist Philosophy: Based on the Study of the Abhidhammatthasangahasarupa. Varma, Chandra B. Delhi: Eastern Book Linkers, 1992.

The Concise Encyclopedia of Western Philosophy. 3rd ed. Rée, Jonathan., and J. O. Urmson. London: Routledge, 2005.

The Concise Encyclopedia of Western Philosophy and Philosophers. Urmson, J. O., and J. L. Ackrill, eds. New York: Hawthorn Books, 1965.

The Concise Encyclopedia of Western Philosophy and Philosophers. 2nd ed. Rev. Urmson, J. O., ed. London: Hutchinson, 1975.

The Concise Encyclopedia of Western Philosophy and Philosophers. Urmson, J. O., and Jonathan Rée. Rev. ed. London: Unwin Hyman, 1989.

The Concise Encyclopedia of Western Philosophy and Philosophers. 3rd ed. Urmson, J. O., and Jonathan Rée. London: Routledge, 2004.

Dictionary of American Philosophy. Nauman, St. Elmo, Jr. New York: Philosophical Library, 1972.

Dictionary of Asian Philosophies. Nauman, St. Elmo, Jr. New York: Philosophical Library, 1978.

Dictionary of Cognitive Science: Neuroscience, Psychology, Artificial Intelligence, Linguistics, and Philosophy. Houdé, Olivier. New York: Psychology Press, 2004.

A Dictionary of Common Philosophical Terms. Pence, Gregory E. Boston: McGraw-Hill, 2000.

The Dictionary of Eighteenth-Century German Philosophers. Kuehn, Manfred, and Heiner Klemme. Bristol, U.K.: Thoemmes, 2004.

The Dictionary of Important Ideas and Thinkers. Rohmann, Chris. London: Arrow, 2002.

Dictionary of Indian Philosophical Concepts. Singh, B. N. Varanasi: Asha Prakashan, 1988.

Dictionary of Irish Philosophers, A–Z. Duddy, Thomas, David Berman, and M. A. Stewart. Bristol, U.K.: Thoemmes Continuum, 2004.

A Dictionary of Muslim Philosophy. 2nd ed. Sheikh, M. Saeed. Lahore: Institute of Islamic Culture, 1981.

The Dictionary of Nineteenth-Century British Philosophers. Mander, W. J., and Alan P. F. Sell. Bristol, U.K.: Thoemmes, 2002.

Dictionary of Oriental Philosophy. Reyna, Ruth. New Delhi: Munshiram Manoharlal, 1984.

A Dictionary of Philosophical Concepts. Peterson, Russell Arthur. Lake Mills, LA: Graphic Pub. Co., 1977.

Dictionary of Philosophy. Angeles, Peter Adam. New York: Barnes & Noble Books, 1981.

Dictionary of Philosophy. Angeles, Peter Adam. HarperPerennial ed. New York: HarperPerennial, 1991.

A Dictionary of Philosophy. Flew, Antony. New York: St. Martin's Press, 1979.

A Dictionary of Philosophy. 2nd ed. Flew, Antony. London: Macmillan, 1984.

A Dictionary of Philosophy. 2nd ed., rev. Flew, Antony. New York: Gramercy Books, 1999.

A Dictionary of Philosophy. 3rd ed. Flew, Antony, and Stephen Priest. London: Pan, 2002.

A Dictionary of Philosophy. Lacey, A. R. New York: Scribner, 1976.

A Dictionary of Philosophy. 2nd ed. Lacey, A. R. London: Routledge & Kegan Paul, 1986.

A Dictionary of Philosophy. 3rd ed. Lacey, A. R. London: Routledge, 1996.

A Dictionary of Philosophy. Mautner, Thomas. Oxford; Cambridge, MA: Blackwell Publishers, 1996.

A Dictionary of Philosophy. Rozental, M. M., and Pavel Fedorovich Iudin. Moscow: Progress Publishers, 1967.

Dictionary of Philosophy. 15th ed., rev. Runes, Dagobert D., ed. Totowa, NJ: Littlefield, Adams, 1965.

Dictionary of Philosophy. Rev. ed. Runes, Dagobert D., ed. New York: Philosophical Library, 1983.

Dictionary of Philosophy. Saifulin, Murad, and Richard R. Dixon. New York: International Publishers, 1984.

A Dictionary of Philosophy. Srinivas, K. New Delhi: Ashish Pub. House, 1993.

Dictionary of Philosophy and Religion: Eastern and Western Thought. Reese, William L. Atlantic Highlands, NJ: Humanities Press, 1980.

Dictionary of Philosophy and Religion: Eastern and Western Thought. New and enl. ed. Reese, William L. Atlantic Highlands, NJ: Humanities Press, 1996.

Dictionary of Religion and Philosophy. MacGregor, Geddes. New York: Paragon House,1989.

A Dictionary of Scholastic Philosophy. 2nd ed. Wuellner, Bernard J. Milwaukee, WI: Bruce Pub. Co., 1966.

The Dictionary of Seventeenth-Century British Philosophers. Pyle, Andrew. Bristol, U.K.: Thoemmes, 2000.

The Dictionary of Seventeenth-Century French Philosophers. Foisneau, Luc. Bristol, U.K.: Thoemmes, 2005.

The Dictionary of Seventeenth and Eighteenth-Century Dutch Philosophers. Bunge, Wiep van. Bristol, U.K.: Thoemmes, 2003.

Dictionary of Technical Terms in Philosophy: English-Kannada-Hindi. Mahadevappa, N. G. Dharwad: Extension Service and Publications, Karnataka University, 1979.

Dictionary of Theories. Bothamley, Jennifer. New York: Barnes & Noble, 2004.

Dictionary of Twentieth-Century British Philosophers. Brown, Stuart C. Bristol, U.K.: Thoemmes Continuum, 2005.

Dictionary of World Philosophy. Iannone, A. Pablo. London: Routledge, 2001.

Encyclopaedia of Language and Linguistics. Vol. 9: Philosophy of Linguistics. Riley, Brian T. New Delhi: Cosmo Publications, 1999.

Encyclopaedia of Upanisads and Its Philosophy: An Exposition of the Fundamental Concepts, History, Philosophy, Teachings, Doctrines, and the System of Upanisads. Kapoor, Subodh. New Delhi: Cosmo Publications, 2002.

Encyclopedia of Asian Philosophy. Leaman, Oliver. London: Routledge, 2001.

Encyclopedia of Chinese Philosophy. Cua, A. S. New York: Routledge, 2003.

The Encyclopedia of Eastern Philosophy and Religion: Buddhism, Hinduism, Taoism, Zen. Fischer-Schreiber, Ingrid, and Stephan Schuhmacher. Boston: Shambhala, 1989.

The Encyclopedia of Indian Philosophies. Bhattacharya, Sibajiban, Karl H. Potter, Karl H. Larson, and Gerald James. Delhi: Motilal Banarsidass, 1970.

Encyclopedia of Indian Philosophies. Bhattacharya, Ram Shankar, and Karl H. Potter. Princeton, NJ: Princeton University Press, 1977.

The Encyclopedia of Indian Philosophies. Potter, Karl H. Princeton, NJ: Princeton University Press, 1970.

Encyclopedia of Indian Philosophies. Rev. ed. Potter, Karl H. Delhi: M. Banarsidass, 1983.

Encyclopedia of Indian Philosophies. 3rd rev. ed. Potter, Karl H. Delhi: Motilal Banarsidass Publishers, 1995.

The Encyclopedia of Indian Philosophies: Advaita Vedanta up to Samkara and His Pupils. Potter, Karl H. Delhi: Motilal Banarsidass, 1981.

Encyclopedia of Indian Philosophies, Bibliography. Potter, Karl H. Seattle: University of Washington, Center for Advanced Research Technology in the Arts and Humanities, 1999.

The Encyclopedia of Indian Philosophies: The Philosophy of the Grammarians. Coward, Harold G., and Raja K. Kunjunni. Delhi: Motilal Banarsidass, 1990.

Encyclopedia of Nineteenth Century Thought. Claeys, Gregory. London: Routledge, 2003.

Encyclopedia of the Philosophical Sciences in Outline, and Critical Writings. Hegel, Georg Wilhelm Friedrich, Ernst Behler, Arnold V. Miller, et al. New York: Continuum, 1990.

Encyclopedia of Philosophy. 2nd ed. Borchert, Donald M., ed. Detroit: Macmillan Reference USA, 2005.

Encyclopedia of Philosophy. Edwards, Paul, ed. New York: Macmillan, 1967.

Encyclopedia of Philosophy: Supplement. Borchert, Donald M., ed. New York: Macmillan Reference USA, Simon & Schuster Macmillan, 1996.

An English-Chinese Dictionary of Chinese Traditional Philosophy. Dankowski, John. Taipei: Chinese News & World Report, 1977.

English-Slovak, Slovak-English Dictionary of Philosophy. Stekauer, Pavol and L. Stekauerová. Presov: ManaCon, 1997.

Great Thinkers A–Z. Baggini, Julian, and Jeremy Stangroom. New York: Continuum, 2004.

The Greek Philosophical Vocabulary. Urmson, J. O. London: Duckworth, 1990.

The HarperCollins Dictionary of Philosophy. 2nd ed. Angeles, Peter Adam. New York: HarperPerennial, 1992.

Historical Dictionary of Descartes and Cartesian Philosophy. Ariew, Roger. Lanham, MD: Scarecrow Press, 2003.

Historical Dictionary of Hegelian Philosophy. Burbidge, John W. Lanham, MD: Scarecrow Press, 2001.

Historical Dictionary of Kierkegaard's Philosophy. Watkin, Julia. Lanham, MD: Scarecrow Press, 2001.

Historical Dictionary of Schopenhauer's Philosophy. Cartwright, David E. Lanham, MD: Scarecrow Press, 2005.

Historical Dictionary of Wittgenstein's Philosophy. Richter, Duncan. Lanham, MD: Scarecrow Press, 2004.

Indian Metaphysics and Epistemology: The Tradition of Nyaya-Vaisesika up to Gangesa. Potter, Karl H. Princeton, NJ: Princeton University Press, 1977.

Indian Philosophy A–Z. Bartley, C. J. Edinburgh: Edinburgh University Press, 2005.

Jewish Philosophy A–Z. Hughes, Aaron W. Edinburgh: Edinburgh University Press, 2005.

Key Concepts in Eastern Philosophy. Leaman, Oliver. London: Routledge, 1999.

Logic from A to Z. Detlefsen, Miael, John Bacon, and David Charles McCarty. London: Routledge, 1999

New Encyclopedia of Philosophy. Grooten, Johan, and G. J. Steenbergen, eds. New York: Philosophical Library, 1972.

Nineties Knowledge. Crystal, David. Edinburgh; Chambers, 1992.

The Oxford Companion to Philosophy. Honderich, Ted. Oxford: Oxford University Press, 2003.

The Oxford Companion to the Mind. 2nd ed. Gregory, R. L. Oxford: Oxford University Press, 2004.

The Oxford Dictionary of Philosophy. Blackburn, Simon. Oxford: Oxford University Press, 1994.

The Penguin Dictionary of Philosophy. Rev. ed. Mautner, Thomas. London: Penguin Books, 1997.

The Penguin Dictionary of Philosophy. Rev. ed. Mautner, Thomas. London: Penguin Books, 2000.

The Penguin Dictionary of Philosophy. 2nd ed. Mautner, Thomas. London: Penguin, 2005

The Philosopher's Dictionary. Martin, Robert M. Peterborough, ON: Broadview Press, 1991.

The Philosopher's Dictionary. 2nd ed. Martin, Robert M. Peterborough, ON: Broadview Press, 1994.

The Philosopher's Dictionary. 3rd ed. Martin, Robert M. Peterborough, ON: Broadview Press, 2002.

Philosophical Dictionary. Brugger, Walter, and Kenneth Baker, S. J. eds. Spokane, WA: Gonzaga University Press, 1972.

Philosophical Dictionary. Enl. ed. Bunge, Mario Augusto. Amherst, NY: Prometheus Books, 2003.

Philosophy and Connectionist Theory. Ramsey, William, and Stephen P. Stich. Hillsdale, NJ: L. Erlbaum Associates, 1991.

The Philosophy of the Grammarians. Coward, Harold G., and K. Kunjunni Raja. Princeton, NJ: Princeton University Press, 1990.

Philosophy of Religion A–Z. Quinn, Patrick. Edinburgh: Edinburgh University Press, 2005.

Quiddities: An Intermittently Philosophical Dictionary. Quine, W. V. Cambridge, MA: Belknap Press of Harvard University Press, 1987.

The Rider Encyclopedia of Eastern Philosophy and Religion: Buddhism, Hinduism, Taoism, Zen. Fischer-Schreiber, Ingrid, Stephan Schuhmacher, and Gert Woerner. London: Rider, 1989.

Sacramentum Mundi: An Encyclopedia of Theology. Rahner, Karl, Cornelius Ernst, and Kevin Smyth. London: Search Press, 1968.

Spinozistic Glossary and Index Dedicated to Spinoza's Insights. Spinoza, Benedictus de; Joseph B. Yesselman, ed. Fairfax, VA: Joseph B. Yesselman, 1999.

Stanford Encyclopedia of Philosophy. Zalta, Edward N., ed. http://plato.stanford.edu. Stanford, CA: Stanford University, Metaphysics Research Lab.

The Thinker's Dictionary: A Handbook for Philosophy and Similar Intellectual Endeavors. 2nd ed. Van de Mortel, J. A. New York, NY: McGraw-Hill, 1995.

A World of Ideas: A Dictionary of Important Theories, Concepts, Beliefs, and Thinkers. Rohmann, Chris. New York: Ballantine Books, 1999.

ASIAN LANGUAGES

ARABIC

Istilah Usuluddin dan falsafah Islam, bahasa Arab-bahasa Malaysia. Kuala Lumpur: Dewan Bahasa dan Pustaka, Kementerian Pendidikan, 1991.

Mafahim wa-al-alfaz fi al-falsafah al-hadithah. 2nd ed. Siddiq, Yusuf. Libya and Tunis: al-Dar al-ʿArabiyah lil-Kitab, 1980.

Al-Mawsuʿah al-falsafiyah al-ʿArabiyah. Ziyadah, Maʾan. Beirut: Maʿhad al-Inmaʾ al-ʿArabi, 1986–1997.

Al-Mawsuʿah al-Muyassarah fi al-Fikr al-Falsafi wa-al-Ijtimaʿi: ʿArabi-Injilizi. Hajj, Kamil. Beirut: Maktabat Lubnan Nashirun, 2000.

Al-Mawsuʿah al-naqdiyah lil-falsafah al-Yahudiyah. Hifni, ʿAbd al-Munʿim. Cairo: Maktabat Madbuli, 1980.

Mawsuʿat al-falasafah. ʿAbbas, Faysal. Beirut: Dar al-Fikr al-ʿArabi, 1996.

Mawsuʿat al-falsafah. Badawi, ʿAbd al-Rahman. Beirut: al-Muʾassasah al-ʿArabiyah lil-Dirasat wa-al-Nashr, 1984.

Mawsuʿat mustalahat al-falsafah ʿinda al-ʿArab. Jihami, Jirar. Beirut: Maktabat Lubnan Nashirun, 1998.

Mawsuʿat mustalahat al-fikr al-ʿArabi wa-al-Islami al-Hadith wa-al-Muʿasir. Dughaym, Samih, Rafiq ʿAjam, and Jirar Jihami. Beirut: Maktabat Lubnan Nashirun, 2002.

Mawsuʿat mustalahat al-Imam al-Ghazzali. ʿAjam, Rafiq. Beirut: Maktabat Lubnan Nashirun, 2000.

Mawsuʿat mustalahat al-Imam Fakhr al-Din al-Razi. Dughaym, Samih. Beirut: Maktabat Lubnan Nashirun, 2001.

Mawsuʿat mustalahat Ibn Khaldun wa-al-Sharif ʿAli Muhammad al-Jurjani. ʿAjam, Rafiq. Beirut: Maktabat Lubnan Nashirun, 2004.

Mawsuʿat mustalahat Ibn Rushd al-Faylasuf. Jihami, Jirar. Beirut: Maktabat Lubnan Nashirun, 2000.

Mawsuʿat mustalahat Ibn Sina: al-shaykh al-raʾis. Jihami, Jirar. Beirut: Maktabat Lubnan Nashirun, 2004.

Mawsuʿat mustalahat Sadr al-Din al-Shirazi. Dughaym, Samih. Beirut: Maktabat Lubnan Nashirun, 2004.

Muʿjam al-falasifah. Tarabishi, Jurj. Beirut: Dar al-Taliʿah, 1987.

Muʿjam al-falasifah al-muyassar. Aubral, François, and Jurj Saʿd. Beirut: Dar al-Hadathah lil-Tibaʿah wa-al-Nashr wa-al-Tawziʿ, 1993.

Muʿjam al-falsafah. Yaʿqubi, Mahmud. Algiers: Maktabat al-Sharikah al-Jazaʾiriyah, 1979.

Al-Muʿjam al-falsafi. Karam, Yusuf, Murad Wahbah, and Yusuf Shallalah. Cairo: Maktab Yulyu, 1966.

Al-Muʿjam al-falsafi. 2nd ed. Wahbah, Murad, Yusuf Karam, and Yusuf Shallalah. Cairo: Dar al-Thaqafah al-Jadidah, 1971.

Al-Muʿjam al-falsafi. 3rd ed. Wahbah, Murad. Cairo: Dar al-Thaqafah al-Jadidah, 1979.

Al-Muʿjam al-falsafi. Majmaʿ al-Lughah al-ʿArabiyah. Cairo: Jumhuriyat Misr al-ʿArabiyah, 1979.

Al-Muʿjam al-falsafi: ʿArabi-Injilizi-Faransi-Almani-Latini. Hifni, ʿAbd al-Munʿim. Cairo: al-Dar al-Sharqiyah, 1990.

Al-Muʿjam al-falsafi bi-al-alfaz al-ʿArabiyah wa-al-Faransiyah wa-al-Inkliziyah wa-al-Latiniyah. Saliba, Jamil. Beirut: Dar al-Kitab al-Lubnani, 1971, 1973.

Muʿjam alfaz al-qiyam al-akhlaqiyah wa-tatawwuruha al-dalali bayna lughat al-shiʿr al-Jahili wa-lughat al-Qurʾan al-karim. Zarzur, Nawal Karim. Beirut: Maktabat Lubnan Nashirun, 2001.

Muʿjam al-mustalahat al-falsafiyah: ʿArabi-Faransi-Inklizi. Khalil, Khalil Ahmad. Beirut: Dar al-Fikr al-Lubnani, 1995.

Muʿjam al-mustalahat al-falsafiyah: Faransi-ʿArabi. Hulw, ʿAbduh. Beirut: al-Markaz al-Tarbawi lil-Buhuth wa-al-Inmaʾ, 1994.

Mulhaq Mawsuʿat al-falsafah. Badawi, ʿAbd al-Rahman. Beirut: al-Muʾassasah al-ʿArabiyah lil-Dirasat wa-al-Nashr, 1996.

Mustalahat al-falsafah fi al-taʿlim al-ʿamm (Injilizi-Faransi-ʿArabi). Maktab al-Daʿim li-Tansiq al-Taʿrib fi al-ʿAlam al-ʿArabi. Casablanca: Dar al-Kitab, 1977.

Mustalahat al-Faylasuf al-Kindi: bahth tahlili. Sayf, Antuwan. Beirut: al-Jamiʿah al-Lubnaniyah, 2003.

Nahwa muʿjam lil-falsafah al-ʿArabiyah: mustalahat wa-shakhsiyat. ʿIraqi, Muhammad ʿAtif. Alexandria, Egypt: Dar al-Wafaʾ li-Dunya al-Tibaʿah wa-al-Nashr, 2001.

Sharh al-mustalahat al-falsafiyah. Bunyad-i Pizhuhishha-yi Islami; Guruh-i Kalam va Falsafah. Mashhad, Iran: al-Majmaʿ, 1993.

Sharh mustalahat al-falsafiyah. Bunyad-i Pizhuhishha-yi Islami; Guruh-i Kalam va Falsafah. Mashhad, Iran: al-Majmaʿ, 1995.

ARMENIAN

P'ilisop'ayakan bararan. Rozental', Mark Moiseevich. Yerevan, 1975.

AZERBAIJANI

Falsafa: ensiklopedik lughati. Rustamov, I. A., and Ismaiyl Omaroghlu. Baku: Azarbaijan Ensiklopediiasy, 1997.

BENGALI

Bharatiya darsana kosha. Bhattacarya, Srimohana, and Dinesh Chandra Bhattacharya. Kalikata: Samskrta Kaleja, 1978–.

Darsana o manobidya paribhashakosha. Abdul Hai, Saiyed. Dhaka: Bamla Ekademi, 1978–.

CHINESE

Beijing da xue fa xue bai ke quan shu: Zhongguo fa lü si xiang shi, Zhongguo fa zhi shi, wai guo fa lü si xiang shi, wai guo fa zhi shi [Peking University Encyclopedia]. Rao

Xinxian, Pu Jian, Wang Zhe, and You Rong. Beijing: Beijing da xue chu ban she, 2000.

Chang yong zhe xue ming ci ci dian. Huang Ming. Nanning: Guangxi ren min chu ban she; Guangxi xin hua shu dian fa xing, 1985.

Che hsüeh hsiao tz'u tien: wai kuo che hsüeh shih pu fen. Shang-hai "che hsüeh hsiao tz'u tien" pien hsieh tsu. [Beijing?], 1976.

Che hsüeh tz'u tien. Liu Yanbo. [Ch'ang-ch'un]: Chi-lin jen min ch'u pan she; Chi-lin sheng hsin hua shu tien fa hsing, 1983.

Ci hai. Shanghai: Shanghai ci shu chu ban she; Xin hua shu dian Shanghai fa xing suo fa xing, 1980.

Ci hai. 2nd ed. Shanghai: Shanghai ci shu chu ban she, 1986.

Da ci hai. Zhe xue juan. Da ci hai bian ji wei yuan hui. Shanghai: Shanghai ci shu chu ban she, 2003.

Dang dai xi fang si chao ci dian [A Dictionary of Contemporary West Trends of Thought]. Wang Miaoyang, and Zhang Huajin. Shanghai: Hua dong shi fan da xue chu ban she, 1995.

Dang dai xi fang zhe xue xin ci dian. Cheng Zhimin, and Jiang Yi. Changchun: Jilin ren min chu ban she, 2003.

Deng Xiaoping li lun ci dian. Yu Yuanpei. Shanghai: Shanghai ci shu chu ban she, 2004.

Dong xi fang zhe xue da ci dian [A Dictionary of Eastern-Western Philosophy]. Jiang Yongfu, Wu Ke, and Yue Changling. Nanchang: Jiangxi ren min chu ban she, 2000.

Han Ying zhe xue chang yong ci hui [Chinese-English Useful Terminology in Philosophy]. Zhang Junfu. [Beijing]: Beijing shi fan da xue chu ban she; Xin hua shu dian zong dian ke ji fa xing suo fa xing, 1989.

Han Ying Zhongguo zhe xue ci dian [A Dictionary of Chinese Philosophy with English Annotations]. Guo Shangxing, and Wang Chaoming.Kaifeng: Henan da xue chu ban she, 2002.

Jian ming xian dai xi fang zhe xue ci dian. Tao Yinbiao, Wu Bin, and Lü Chao. Chengdu: Sichuan ren min chu ban she; Sichuan sheng xin hua shu dian fa xing, 1988.

Jian ming xi fang zhe xue ci dian. Tao Yinbiao. Shenyang: Liaoning ren min chu ban she; Liaoning sheng xin hua shu dian fa xing, 1985.

Jian ming ying yong zhe xue ci dian. Wang Lianfa, and Huang Changjun. Beijing: Zhongguo guang bo dian shi chu ban she; Xin hua shu dian zong dian Beijing fa xing suo jing xiao, 1991.

Jian ming zhe xue bai ke ci dian. "Jian ming zhe xue bai ke ci dian" bian xie zu. Beijing: Xian dai chu ban she; Xin hua shu dian zong dian Beijing fa xing suo fa xing, 1990.

Jian ming zhe xue ci dian. Meng Xianhong. [Wuhan]: Hubei ci shu chu ban she; Xin hua shu dian Hubei fa xing suo jing xiao, 1987.

Jian ming zhe xue ci dian. Rozental', Mark Moiseevich, and Pavel Iudin. Beijing: Sheng huo, du shu, xin zhi san lian shu dian: Xin hua shu dian fa xing, 1973.

Jian ming zhe xue xiao ci dian. Zhou Lin, Li Shijia, and Yuan Youwen. Chengdu: Sichuan ren min chu ban she; Sichuan sheng xin hua shu dian fa xing, 1986.

Jianqiao zhe xue ci dian. [The Cambridge Dictionary of Philosophy]. Audi, Robert, and Lin Zhenghong. Taipei: Mao tou ying chu ban, 2002.

Kong xue zhi shi ci dian. Dong Naiqiang. Beijing: Zhongguo guo ji guang bo chu ban she; Xin hua shu dian jing xiao, 1990.

Lun heng ji jie. Wang Chong, Liu Pansui, Liang Yuandi, XieZhangting, Xiao Yi, and Jinlouzi. Taipei: Shi jie shu ju, 1966.

Makesi zhu yi ji ben yuan li jian ming jiao xue ci dian. Jun dui di fang shi yi suo yuan xiao "Makesi zhu yi ji ben yuan li jian ming jiao xue ci dian" bian xie zu. Shenyang: Liaoning jiao yu chu ban she; Liaoning sheng xin hua shu dian jing xiao, 1988.

Makesi zhu yi zhe xue ci dian. Li Shikun. [Beijing]: Zhongguo guang bo dian shi chu ban she; Xin hua shu dian zong dian Beijing fa xing suo jing xiao, 1990.

Makesi zhu yi zhe xue da ci dian. Jin Binghua. Shanghai: Shanghai ci shu chu ban she, 2003.

Makesi zhu yi zhe xue da ci dian. Wang Guoyan, Wei Xiaozhou, and Li Guozhen. Beijing: Zhongguo guang bo dian shi chu ban she, 1993.

Makesi zhu yi zhe xue quan shu. Li Huaichun. Beijing: Zhongguo ren min da xue chu ban she; Jing xiao Xin hua shu dian, 1996.

Makesi zhu yi zhe xue shi ci dian. Zhuang Fuling, and Xu Lin. Beijing: Beijing chu ban she; Xin hua shu dian Beijing fa xing suo jing xiao, 1992.

Ouzhou zhe xue shi ci dian. Ma Xiaoyan. Kaifeng: Henan da xue chu ban she; Henan sheng xin hua shu dian fa xing, 1986.

Qi meng yun dong bai ke quan shu. Reill, Peter Hanns, Ellen Judy Wilson, Liu Beicheng, and Wang Wanqiang. Shanghai: Shanghai ren min chu ban she, 2004.

Ren lei si xiang de zhu yao guan dian: xing cheng shi jie de guan nian [Guide to Human Thought]. McLeish, Kenneth. Beijing: Xin hua chu ban she, 2004.

Ru jia wen hua ci dian. Xu Xinghai, and Liu Jianli. Zhengzhou: Zhongzhou gu ji chu ban she; Xin hua shu dian jing xiao, 2000.

Si shu wu jing ming ju jian shang ci dian. Tian Ren. Huhehaote: Neimenggu ren min chu ban she, 1999.

Wai guo zhe xue da ci dian. Feng Qi and Xu Xiaotong. Shanghai: Shanghai ci shu chu ban she, 2000.

Xian dai xi fang zhe xue ci dian. Ge Li. Beijing: Qiu shi chu ban she, 1990.

Xian dai xi fang zhe xue ci dian. Xia Jisong, and Zhang Jiwu. [Hefei]: Anhui ren min chu ban she; Anhui sheng xin hua shu dian fa xing, 1987.

Xian xue ci dian. Dai Yuanchang. Taipei: Chen shan mei, 1978.

Xi fang zhe xue ci dian. Tan Xintian. Jinan: Shandong ren min chu ban she, 1992.

Xi fang zhe xue Ying Han dui zhao ci dian [Dictionary of Western Philosophy: English-Chinese]. Bunnin, Nicholas, and Yu Jiyuan. Beijing: Ren min chu ban she, 2001.

Xin bian jian ming zhe xue ci dian. Blauberg, Igor' Viktorovich, and Igor' Konstantinovich Pantin. Changchun: Jilin ren min chu ban she, 1983.

Xin bian Lun heng. Wang Chong and Xiao Dengfu. Taipei: Taiwan gu ji chu ban gong si, 2000.

Xin bian zhe xue ci dian. Wang Wenqing, Yu Dunhua, and Luo Chi. [Wuhan]: Hubei jiao yu chu ban she; Xin hua shu dian Hubei fa xing suo jing xiao, 1989.

Xin bian zhe xue da ci dian. Huang Nansen and Yang Shoukan. Taiyuan: Shanxi jiao yu chu ban she: Shanxi sheng xin hua shu dian fa xing, 1993.

Xin bian zhe xue da ci dian. Sun Yun and Sun Meiyao. [Harbin, China]: Ha'erbin chu ban she; [Beijing]: Xin hua shu dian shou du fa xing suo fa xing, 1991.

Xin Ying Han zhe xue ci dian [A New English-Chinese Dictionary of Philosophy]. Xu Changming. Chengdu: Sichuan da xue chu ban she; Sichuan sheng xin hua shu dian jing xiao, 1991.

Xi yang zhe xue ci dian. Brugger, Walter. Taipei: Guo li bian yi guan, 1976.

Ying Han zhe xue ci hui. Fu dan da xue, Xian dai zhe xue yan jiu suo, "Ying Han zhe xue ci hui" bian xie zu. Nanchang: Jiangxin ren min chu ban she; Jiangxi sheng Xin hua shu dian fa xing, 1987.

Ying Han zhe xue shu yu ci dian. Zhongguo she hui ke xue yuan; "Zhe xue yi cong" bian ji bu. Beijing: Zhong gong zhong yang dang xiao chu ban she; Xin hua shu dian jing xiao, 1991.

Zhe xue bai ke xiao ci dian. Liu Wenying, Qi Fenzhong, Zeng Xiangli, and Yang Keli. Lanzhou: Gansu ren min chu ban she; Gansu sheng xin hua shu dian fa xing, 1987.

Zhe xue ci dian. Angeles, Peter Adam. Taipei: Mao tou ying chu ban she gu fen you xian gong si, 2000.

Zhe xue ci dian. 4th ed. Fan Bingqing. Taipei: Taiwan shang wu yin shu guan, 1976.

Zhe xue ci dian. Frolov, Ivan Timofeevich, and Lin Shubo. [Canton]: Guangdong ren min chu ban she; Guangdong sheng xin hua shu dian jing xiao, 1989.

Zhe xue ci dian. Liu Yanbo. [Changchun]: Jilin ren min chu ban she; Jilin sheng xin hua shu dian fa xing, 1983.

Zhe xue da ci dian. Feng Qi. Shanghai: Shanghai ci shu chu ban she: Shanghai ci shu chu ban she fa xing suo fa xing, 2001.

Zhe xue da ci dian. Makesi zhu yi zhe xue juan. "Zhe xue da ci dian Zhongguo zhe xue shi juan" bian ji wei yuan hui. Shanghai: Shanghai ci shu chu ban she; Shanghai ci shu chu ban she fa xing suo fa xing, 1990.

Zhe xue da ci shu. Luo Guang and Li Zhen. Taipei Xian Xinzhuang: Fu ren da xue chu ban she, 1993–.

Zhe xue da ci shu. Zhe xue da ci shu bian shen wei yuan hui. Taipei Xian Xinzhuang: Fu ren da xue chu ban she, 1993–.

Zhe xue gai nian bian xi ci dian. Huang Nansen, Li Zongyang, and Tu Yinsen. Beijing: Zhong gong zhong yang dang xiao chu ban she; Jing xiao Xin hua shu dian, 1993.

Zhe xue jian ming ci. [Nanjing]: Jiangsu ren min chu ban she; Jiangsu sheng xin hua shu dian fa xing, 1985.

Zhe xue xiao bai ke. Xing Fensi and Zhao Fengqi. [Beijing]: Zhongguo qing nian chu ban she; Xin hua shu dian Beijing fa xing suo fa xing, 1986.

Zhe xue xiao ci dian. Chao yang chu ban she. Xianggang: Chao yang chu ban she, 1974.

Zhe xue xiao ci dian. Liu Yanbo. [Changchun]: Jilin ren min chu ban she; Jilin sheng xin hua shu dian fa xing, 1983.

Zhe xue xiao ci dian. Ma Quanmin and Su Houzhong. [Beijing]: Ren min chu ban she; Xin hua shu dian jing xiao, 1990.

Zhe xue xiao ci dian. Wan Zhonghang. Shanghai: Shanghai ci shu chu ban she, 2003.

Zhe xue xiao ci dian: Bian zheng wei wu zhu yi he li shi wei wu zhu yi bu fen. Shanghai Zhe xue xiao ci dian bian xie zu. Shanghai: Ren min chu ban she; Xin hua shu dian Shanghai fa xing suo fa xing, 1975.

Zhe xue xiao ci dian: Ru fa dou zheng shi bu fen. Shanghai Zhe xue xiao ci dian bian xie zu. Shanghai: Ren min chu ban she; Xin hua shu dian Shanghai fa xing suo fa xing, 1974.

Zhe xue xiao ci dian xu bian. Wai guo zhe xue shi bu fen. Chao yang chu ban she; Bian ji bu. Xianggang: Chao yang chu ban she, 1976.

Zhe xue xin gai nian ci dian. Shi Lei, Cui Xiaotian, and Wang Zhong. Ha'erbin: Heilongjiang ren min chu ban she; Heilongjiang sheng xin hua shu dian fa xing, 1988.

Zhe xue zhi shi quan shu. Han Shuying and Zhang Yongqian. Lanzhou: Gansu ren min chu ban she; Gansu sheng Xin ha shu dian fa xing, 1989.

Zhe xue zi dian. Luo Guang. Taipei: Fu ren da xue chu ban she, 1990.

Zhe xue zi dian. Zhe xue da ci shu bian shen wei yuan hui. Zhonghua min guo Taiwan Sheng Taibei Xian Xinzhuang: Fu ren da xue chu ban she, 1990.

Zhongguo da bai ke quan shu: zhe xue. Zhongguo da bai ke quan shu zong bian ji wei yuan hui, Li xue bian ji wei yuan hui. Beijing: Zhongguo da bai ke quan shu chu ban she; Shanghai: Xin hua shu dian Shanghai fa xing suo fa xing, 1987.

Zhongguo da bai ke quan shu: zhe xue. Zhongguo da bai ke quan shu zong bian ji wei yuan hui, Zhe xue bian ji wei yuan hui. Beijing: Zhongguo da bai ke quan shu chu ban she, 1992.

Zhongguo li shi da ci dian. Si xiang shi. Zhongguo li shi da ci dian si xiang shi juan bian zuan wei yuan hui. Shanghai: Shanghai ci shu chu ban she; Shanghai ci shu chu ban she fa xing suo fa xing, 1989.

Zhongguo ru xue ci dian. Zhao Jihui and Guo Hou'an. Shenyang: Liaoning ren min chu ban she; Liaoning sheng xin hua shu dian fa xing, 1988.

Zhongguo shen mi wen hua bai ke zhi shi. Jin Liangnian. Shanghai: Shanghai wen hua chu ban she, 1994.

Zhongguo zhe xue ci dian. Wei Zhengtong. Taipei: Da lin chu ban she, 1977.

Zhongguo zhe xue ci dian. Wei Zhengtong. Taipei: Da lin chu ban she, 1980.

Zhongguo zhe xue ci dian. Wei Zhengtong. Taipei: Da lin chu ban she, 1981.

Zhongguo zhe xue ci dian. Wei Zhengtong. Taipei: Shui niu tu shu chu ban shi ye you xian gong si, 1986.

Zhongguo zhe xue ci dian. Wei Zhengtong. Beijing: Shi jie tu shu chu ban gong si Beijing gong si chong yin; Xin hua shu dian Beijing fa xing suo fa xing, 1993.

Zhongguo zhe xue ci dian da quan. Wei Zhengtong. Taipei: Shui niu chu ban she, 1983.

Zhongguo zhe xue ci dian da quan. Wei Zhengtong. Taipei: Shui niu chu ban she, 1988.

Zhongguo zhe xue da ci dian. Fang Keli. Beijing: Zhongguo she hui ke xue chu ban she, 1994.

Zhongguo zhe xue shi xiao ci dian. Zhao Shulian. [Zhengzhou]: Henan ren min chu ban she; Henan sheng xin hua shu dian fa xing, 1986.

Zhongguo zhe xue xiao bai ke quan shu. Zheng Wan'geng, Wang Deyou, and Li Shen. Beijing: Zhongguo da bai ke quan shu chu ban she, 2001.

Zhonghua ru xue tong dian. Wu Feng, Song Yifu, Huo Baozhen, and Song Yanzong. Haikou: Nanhai chu ban gong si; Jilin sheng xin hua shu dian fa xing, 1992.

Zhu zi bai jia da ci dian. Feng Kezheng and Fu Qingsheng. Shenyang: Liaoning ren min chu ban she, 1996.

Zhu zi bai jia da ci dian. Huang Kaiguo. Chengdu: Sichuan ren min chu ban she, 1999.

Zhu zi bai jia da ci dian. Liu Guancai. Beijing: Hua ling chu ban she, 1994.

Zhu zi bai jia da ci dian. Liu Guancai. Taipei: Jian hong chu ban she, 2000.

Zhu zi bai jia ming ju jian shang ci dian. Tianren. Huhehaote: Nei Menggu ren min chu ban she, 1999.

Zhu zi bai jia ming pian jian shang ci dian. Ma Zhenduo, Mao Huijun, Tang Qinfu, Zhang Liangyi, and Liao Jianhua. Shanghai: Shanghai ci shu chu ban she, 2003.

GEORGIAN

K'art'vel p'ilosop'ost'a lek'sikoni: personalia. Buachidze, Tamaz. Tbilisi: Gamomc'emloba "Oazisi," 2000.

GUJARATI

Paribhashika kosa-tattvajñana. Ravala, S. V. Amadavada, India: Yunivarsiti Grantha Nirmana Borda, Gujarata Rajaya, 2001.

HINDI

Bharatiya-darsana-brhatkosa. 2nd ed. Jñana, Bacculala Avasthi. Dilli: Sarada Pablising Hausa, 2004.

Bharatiya darsana kosha. Misra, Lakshmi Kanta. Dilli: Indiyana Yunivarsiti Presa ke lie Himalaya, 1973.

Bharatiya darsana paribhasha kosa. Sukla, Dinanatha. Dilli: Pratibha Prakasana, 1993.

Jaina samskrti kosa [Encyclopaedia of Jainism]. Jain, Bhagchandra. Nagapura: Sanmati Pracya Sodha Samsthana; Varanasi: Kala evam Dharma Sodha Samsthana, 2002.

INDONESIAN

Ensiklopedi agama dan filsafat. Effendy, Mochtar. [Palembang]: Penerbit Universitas Sriwijaya, 2000–2001.

Kamus populer filsafat. Hartoko, Dick. Jakarta: Rajawali, 1986.

JAPANESE

Furansu tetsugaku shiso jiten. Kobayashi, Michio. Tokyo: Taishukan Shoten, 1999.

Gendai shiso ki wado jiten. Washida, Koyata. Tokyo: San'ichi Shobo, 1993.

Gendai tetsugaku jiten. Yamazaki, Masakazu, and Hiroshi Ichikawa. Tokyo: Kodansha, 1970.

Hikaku shiso jiten. Nakamura, Hajime, and Hideo Mineshima. Tokyo: Tokyo Shoseki, 2000.

Iwanami shojiten tetsugaku. Awata, Kenzo, and Yoshishige Kozai. Tokyo: Iwanami Shoten, 1968.

Iwanami tetsugaku, shiso jiten. Hiromatsu, Wataru. Tokyo: Iwanami Shoten, 1998.

Iwanami tetsugaku shojiten. Awata, Kenzo, and Yoshishige Kozai. Tokyo: Iwanami Shoten, 1979.

Jiten tetsugaku no ki. Nagai, Hitoshi. Tokyo: Kodansha, 2002.

Kan'i tozai tetsugaku shiso jiten. Hara, Tomio, and Shinpen sekai tetsugaku shojiten. Tokyo: Sanshin Tosho, 1983.

Rarusu tetsugaku jiten. Julia, D., Hisaaki Katayama, Yorihiro Yamagata, and Kiyokazu Washida. Tokyo: Taishukan Shoten, 1998.

Seiyo shisoshi jiten. Matsunami, Shinzaburo. Tokyo: Tokyodo Shuppan, 1973.

Sekai shiso kyoyo jiten. Kashiyama, Kinshiro. Tokyo: Tokyodo Shuppan, 1965.

Sekai shiso kyoyo jiten. Kashiyama, Kinshiro. [Tokyo]: Tokyodo Shuppan, 1978.

Sekai shiso kyoyo jiten. Nihon, Toyo hen. Nakamura, Hajime, and Sumeru Yamada. Tokyo: Tokyodo Shuppan, 1965.

Shinpan tetsugaku, ronri yogo jiten. Shiso no Kagaku Kenkyukai. Tokyo: San'ichi Shobo, 1995.

Shoshi hyakka no jiten. Ezure, Takashi. Tokyo: Taishukan Shoten, 2000.

Tetsugaku chujiten. Oda, Seiji, Ichiro Kobayashi, and Isao Kondo. Sendai-shi: Shogakusha, 1983.

Tetsugaku jii: tsuketari Shinkoku onpu. Inoue, Tetsujiro. Tokyo: Meicho Fukyukai, 1980.

Tetsugaku jii yakugo sosakuin. Hida, Yoshifumi. Tokyo: Kasama Shoin, 1979.

Tetsugaku jiten. Tokyo: Heibonsha, 1966.

Tetsugaku jiten. Tokyo: Heibonsha, 1969.

Tetsugaku jiten. Tokyo: Heibonsha, 1971.

Tetsugaku jiten. Mori, Koichi. Tokyo: Aoki Shoten, 1973.

Tetsugaku jiten. 4th ed. Mori, Koichi. Tokyo: Aoki Shoten, 1985.

Tetsugaku jiten. Mori, Koichi. Tokyo: Aoki Shoten, 1995.

Tetsugaku jiten. Mori, Koichi, and Yoshishige Kozai. Tokyo: Aoki Shoten, 1971.

Tetsugaku jiten. 2nd ed. Mori, Koichi, and Yoshishige Kozai. Tokyo: Aoki Shoten, 1972.

Tetsugaku jiten. Quine, Willard Van Orman. Tokyo: Hakuyosha, 1994.

Tetsugaku jiten. Shitanaka, Kunihiko, ed. Tokyo: Heibonsha, 1967.

Tetsugaku kihon jiten: tetsugaku nyumon. 2nd ed. Satomi, Gunshi, and Shuichi Hida. Tokyo: Fuji Shoten, 1993.

Tetsugaku, ronri yogo jiten. Ofuchi, Kazuo. Tokyo: San'ichi Shobo, 1975.

Tetsugaku shiso kopasu jiten. Alchive Co. Tokyo: Nihon Jitsugyo Shuppansha, 1987.

Tetsugaku shojiten. Komatsu, Setsuro. Kyoto: Horitsu Bunkasha, 1970.

Tetsugaku shojiten. Miwatari, Yukio. Tokyo: Kyodo Shuppan, 1974.

Tetsugaku yogo jiten. Muraji, Yoshinari. Tokyo: Tokyodo Shuppan, 1974.

Toyo tetsugaku kiwado jiten [Key Concepts in Eastern Philosophy]. Leaman, Oliver, and Hiromi Ogino. Tokyo: Seidosha, 2000.

KAZAKH

Filosofiialyq sozdik. Nurghaliev, Rymghali. Almaty: "Qazaq entsiklopediiasy" Bas redaktsiiasy, 1996.

KOREAN

Ch'oesin ch'orhak sajon. Ch'orhak Sajon P'yonch'anhoe (Korea). Seoul: Ilsinsa, 1986.

Ch'orhak sajon. Choson Minjujuui Inmin Konghwaguk Sahoe Kwahagwon; Ch'orhak Yon'guso. Tokyo: Hagu Sobang, 1971.

Ch'orhak sajon. Im, Sok-chin. Seoul: Isak, 1983.

Ch'orhak sajon. Im, Sok-chin. Seoul: Toso Ch'ulp'an Ch'ongsa, 1998.

Ch'orhak sajon. Im, Sok-chin, and Se-yon Hwang. Seoul: Chungwon Munhwa, 1987.

Ch'orhak sajon. Sahoe Kwahagwon (North Korea); Ch'orhak Yonguso; Sahoe Kwahak Ch'ulp'ansa. [P'yŏngyang]: Sahoe Kwahak Ch'ulp'ansa, 1970.

Ch'orhak sajon. Sahoe Kwahagwon (North Korea); Ch'orhak Yonguso; Sahoe Kwahak Ch'ulp'ansa; Ch'orhak Sajon P'yonjipcho. [P'yŏngyang]: Sahoe Kwahak Ch'ulp'ansa, 1985.

Ch'orhak sajon. Sahoe Kwahagwon (North Korea); Ch'orhak Yonguso. Seoul: Toso Ch'ulp'an Him, 1988.

Ch'orhak sosajon. Chongno Sojok Chusik Hoesa; P'yonjippu. Seoul: Chongno Sojok, 1989.

Ch'orhak sosajon. Han'guk Ch'orhak Sasang Yon'guhoe. Seoul: Tongnyok, 1990.

Ch'orhak sosajon. Rozental', Mark Moiseevich, and Chin Yu. Pukkyong: Minjok Ch'ulp'ansa; Sinhwa Sojom parhaeng, 1981.

Ch'orhak taesajon. Han'guk Ch'orhak Sasang Yon'guhoe. Seoul: Tongnyok, 1989.

Ch'orhak taesajon [Encyclopedia of Philosophy]. Yi, Kwang-mo. Seoul: Sin T'aeyangsa, 1991.

Segye ch'orhak taesajon [The World Dictionary of Philosophy]. Seoul: Kyoyuk Ch'ulp'an Kongsa; chon'guk ch'ongp'an Taeyong Sorim, 1985.

Segye ch'orhak taesajon [The World Dictionary of Philosophy]. Kang, Yong-son. Seoul: Songgyun Sogwan, 1977.

Segye ch'orhak taesajon. Kang, Yong-son. Seoul: Kyoyuk Ch'ulp'an Kongsa, 1980.

Segye sasang sajon. Segye Sasang P'yonjiphoe. Seoul: Yangudang, 1986.

Togyo sasang sajon. Kim, Sung-dong. Pusan: Pusan Taehakkyo Ch'ulp'anbu, 2004.

Tongyang sasang sajon. Yu, Chong-gi. Taejŏn: Umundang Ch'ulp'ansa, 1965.

Yang-su eui sae-kae. Sae-kae sa-sang kyo-yang sa-jun. Ahn, Choon-keun, comp. [Seoul]: Eul-yoo moon-wha Sa, 1969.

Yugyo taesajon. Yugyo Sajon P'yonch'an Wiwonhoe. Seoul: Pagyongsa, 1990.

MARATHI

Marathi tattvajñana-mahakosa. Vadekara, Devidasa Dattatreya. Pune: Marathi Tattvajñana-mahakosa Mandala, 1974.

MONGOLIAN

Filosofiin ukhaany Oros Mongol ner tom'ëo. Luvsantseren, G., B. Sambuu, S. Norovsambuu, and C. Dogsuren. Ulaanbaatar: BNMAU-yn Shinzhlekh Ukhaany Akademi, 1985.

PAHLAVI

A Philosophical Lexicon in Persian and Arabic. Afnan, Soheil Muhsin. Beirut: Dar el-Mashreq, 1969.

PERSIAN

Alifba-yi falsafah-i jadid: (da'irat al-maʿarif-i falsafi) Inglisi bi-Farsi. Javadi, Zabih Allah. Tehran: Intisharat-i Ibn Sina, 1969.

Dictionary of Philosophy and Social Sciences, English-Persian. Pizhuhishgah-i ʿUlum-i Insani (Iran). [Tehran]: Center, 1976.

Farhang-i falsafah va aʿlam-i vabastah. Ardabili, ʿAli ʿIlmi. Mashhad: Intisharat-i imamat, 1981.

Farhang-i falsafi. Saliba, Jamil. [Tehran]: Hikmat, 1987.

Farhang-i ʿilmi va intiqadi-i falsafah. Lalande, André. Tehran: Muʾassasah-i Intisharati-i Firdawsi-i Iran, 1998.

Farhang-i istilahat-i falsafah: Inglisi-Farsi. Babayi, Parviz. Tehran: Muʾassasah-i Intisharat-i Nigah, 1995.

Farhang-i istilahat-i falsafah va ʿulum-i ijtimaʿi: Ingilisi-Farsi. Birijaniyan, Mari. Tehran: Muʾassasah-ʾi Mutalaʿat va Tahqiqat-i Farhangi, 1992.

Farhang-i istilahat-i falsafi-i Mulla Sadra. Sajjadi, Jaʿfar. Tehran: Sazman-i Chap va Intisharat-i Vizarat-i Farhang va Irshad-i Islami, 2000.

Farhang-i ʿulum-i ʿaqli: shamil-i istilahat-i falsafi, kalami, mintaqi. Sajjadi, Jaʿfar. Tehran: Anjuman-i Islami-i Hikmat va Falsafah-ʾi Iran, 1982.

Istilahat-i falsafi va tafavut-i anha bayakdigar. Karaji, ʿAli. Qom: Daftar-i Tablighat-i Islami-i Hawzah-i ʿIlmiyah-i Qum, Markaz-i Intisharat, 1996.

Istilahʾnamah-ʾi falsafah-ʾi Islami. Yaʿqubʾnizhad, Muhammad Hadi. Qom: Muʿavinat-i Mutaliʿati va Ittilaʿ-i Rasani, Markaz-i Mutaliʿat va Tahqiqat-i Islami "Pazhuhishgah," 1997.

Sih sunnat-i falsafi: guzarishi az falsafah-i Hindi, Chini va Yahudi. Smart, Ninian, Shlomo Pines, Wing-tsit Chan, and Paul Edwards. Qom: Markaz-i Intisharat-i Daftar-i Tablighat-I Islami, Hawzah-i ʿIlmiyah-i Qum, 1999.

Vazhahʾnamah-ʾi falsafah va ʿulum-i ijtimaʿi: Farsi, ʿArabi, Faransah, Ingilisi, Almani, Latin. Saliba, Jamil. [Tehran]: Shirkat-i Sihami-i Intishar, 1991.

Vazhahʾnamah-ʾi falsafi: Farsi-ʿArabi. Afnan, Soheil Muhsin. Beirut: Dar al-Mashriq, 1969.

Vazhahʾnamah-ʾi falsafi: Farsi-ʿArabi-Inglisi-Faransah-Pahlavi-Yunani-Latin. 2nd ed. Afnan, Soheil Muhsin. [Tehran]: Nashr-i Nuqrah, 1984.

Vazhahʾnamah-ʾi tarikh-i falsafah dar Islam. Bakhtar, Zhalah-iʾ. Tehran: Markaz-i Nashr-i Danishgahi, 1994.

PUNJABI

Angrezi-Pañjabi takaniki shabadawali, falasafa. 2nd ed. Punjab State University Text-Book Board. Chandigarh: Pañjaba Sateta Yuniwarasiti, Taikasata-buka Borada, 1970s.

SANSKRIT

Nyayakosah: sakalasastropakarakanyayadisastriyapadarthaprakasakah. Abhyankara, Vasudevasastri. 4th ed. Punyapattana: Bhandarakarapracyavidyasamsodhanamandiradhikrtaih [1978?].

Nyayakosa, or, Dictionary of Technical Terms in Indian Philosophy. Jhalakikar, Bhimacarya, and Vasudevasastri

Abhyankara. 4th ed. Pune: Bhandarkar Oriental Research Institute, 1996.

TAMIL

Caiva camayak kalaik kalañciyam. Tiruccirrampalam, Civa. Chennai: Rajesvari Puttaka Nilaiyam, 2002.

Camayac collakarati. Cacivalli. Chennai: Ulakat Tamilaraycci Niruvanam, 1987.

TELUGU

Tattvasastranighantuvu. Rajagopalaravu, Em. Hyderabad: Telugu Akadami, 1978.

THAI

’Athibai sap pratchaya kanmu’ang læ sangkhom [Dictionary of Political and Social Philosophy]. 2nd ed. Witthayakon Chiangkun. Bangkok: Samnakphim Saithan, 2004.

Photchananukrom pratya: ’Angkrit-Thai. Chetsada Thongrungrot. Bangkok: Bo Dæng, 2004.

Sapthanukrom pratya waduai chittaniyom-watthuniyom. Methi ’Iamcharœn. Bangkok: Somchai Kanphim, 1976.

TURKISH

Ansiklopedik felsefe sözlügü. Tokatli, Attilâ. Istanbul: Bilgi Yayinevi, 1973.

Bir felsefe dili kurmak: modern felsefe ve bilim terimlerinin Türkiye’ye girisi. Kara, Ismail. Istanbul: Dergâh Yayinlari, 2001.

Ezoterik sözlük. Werner, Helmut, Bülent Atatanir, and Özgü Çelik. Istanbul: Omega Yayinlari, 2005.

Felsefe ansiklopedisi: kavramlar ve akimlar. Hançerlioglu, Orhan. Istanbul: Remzi Kitabevi, 1976–1980.

Felsefe ansiklopedisi: kavramlar ve akimlar. Hançerlioglu, Orhan. Istanbul: Remzi Kitabevi, 2000.

Felsefe okullari sistemleri: ülke, konulariyla ünlü kisiler. Göktepe, Salahattin. Izmir: Yeniyol Matbaasi, 1976.

Felsefe sözlügü. 3rd ed. Hançerlioglu, Orhan. Istanbul: Remzi Kitabevi, 1975.

Felsefe sözlügü. 6th ed. Hançerlioglu, Orhan. Istanbul: Remzi Kitabevi, 1982.

Felsefe sözlügü. Ulas, Sarp Erk, and A. Bâki Güçlü. Ankara: Bilim ve Sanat, 2002.

Felsefî doktrinler sözlügü. 2nd ed. Bolay, Süleyman Hayri. Istanbul: Ötüken, 1981.

Felsefî doktrinler ve terimler sözlügü. Bolay, S. Hayri. [Ankara]: Akçag, 1997.

Felsefî doktrinler ve terimler sözlügü. Bolay, S. Hayri. [Ankara]: Akçag, 1999.

Islâm felsefesi sözlügü. Vural, Mrhmet. Ankara: Elis, 2003.

UIGHUR

Chat’al pailasopliri qisqicha lughiti. Ahmidi, Ibrahim. Ürümqi, China: Shinjang Khalq Nashriyati, 1989.

URDU

Kashshaf-i istilahat-i falsafah: Urdu-Angrezi. Kadir, Kazi A. [Karachi]: Shuʻbah-yi Tasnif o Talif o Tarjumah, Karaci Yunivarsiti, bah ishtirak-i mali, Muqtadirah-yi Qaumi Zaban, 1994.

UZBEK

Filosofiiadan qisqacha lughat. Azarov, Nikolai Ivanovich, and M. M. Khairullaev. Tashkent: “Uzbekiston” Nashriëti, 1973.

Filosofiia lughati. Abdurahmonov, Rustam. Tashkent: “Uzbekiston” Nashriëti, 1976.

VIETNAMESE

Tu’ dien triet hoc. Cung, Kim Tien. Hanoi: Nhà xuat ban Van hóa-thông tin, 2002.

Tu’ dien triet hoc. Rozental’, Mark Moiseevich. Mát-xco’-va: Tien bo, 1986.

Tu’ diên triêt hoc. Rozental’, Mark Moiseevich, and Pavel Fedorovich Iudin. Hanoi: Su’ Thât, 1976.

Tu’ dien triet hoc gian yeu: có doi chieu tu’ Nga, Anh, Đu’c, Pháp. Hu’u Ngoc, Phú Hiep Du’o’ng, and Hu’u Tang Lê. Hanoi: Đai hoc và Trung hoc chuyên nghiep, 1987.

Tu’-dien và danh-tu’ triêt-hoc. Trân-van Hiên-Minh. Saigon: Tu-Sách Ra-koi, 1966.

EUROPEAN LANGUAGES

ALBANIAN

Fjalor i filozofisë. Shkolla e Partisë "V. I. Lenin." Tiranë: Shtëpia Botuese "8 Nëntori," 1981.

BULGARIAN

Anglo-bulgarski ucheben rechnik-minimum za studenti po filosofiia. Dodova, L., and R. Mukharska. Sofia: Sofiiski universitet "Kliment Okhridski," 1984.

Filosofite: kratuk rechnik. Radev, Radi, Ivan Stefanov, and Aleksandur Lichev. Sofia: Izd-vo "Khristo Botev," 1992.

Filosofski rechnik. Buchvarov, Mikhail Dimitrov, Mincho Draganov, Stoiu G. Stoev, and Mark Moiseevich Rozental'. Sofia: Partizdat, 1977.

Filosofski rechnik. 2nd ed. Buchvarov, Mikhail Dimitrov, Mincho Draganov, and Stoiu G. Stoev. Sofia: Partizdat, 1985.

Filosofski rechnik suvremenni filosofi XIX–XX vek shkoli, napravleniia. Filatov, Vladimir Petrovich, and V. S. Malakhov. Sofia: Izd-vo GAL-IKO, 1993.

CROATIAN

Covjek, odgoj, svijet: mala filozofijskoodgojna razlozba. Polic, Milan. Zagreb: Kruzak, 1997.

Filozofija. 2nd ed. Grlic, Danko. Zagreb: Panorama, 1965.

Filozofija: uvod u filozofsko mišljenje i rjecnik. Bošnjak, Branko. Zagreb: Naprijed, 1973.

Filozofija: uvod u filozofsko mišljenje i rjecnik. 2nd ed. Bošnjak, Branko. Zagreb: Naprijed, 1977.

Filozofija: uvod u filozofsko mišljenje i rjecnik. 3rd ed. Bošnjak, Branko. Zagreb: Naprijed, 1982.

Filozofijski rjecnik. Filipovic, Vladimir, and Branko Bošnjak. Zagreb: Matica hrvatska, 1965.

Filozofijski rjecnik. 2nd ed. Filipovic, Vladimir, and Branko Bošnjak. Zagreb: Nakladni zavod Matice hrvatske, 1984.

Filozofijski rjecnik. 3rd ed. Filipovic, Vladimir, and Branko Bošnjak. Zagreb: Nakladni zavod Matice hrvatske, 1989.

Filozofski recnik. Maric, Svetislav. Belgrade: Dereta, [1991?].

Leksikon filozofa. Grlic, Danko. Zagreb: Naprijed, 1968.

Rjecnik filozofskih pojmova. Misic, Anto. Split: Verbum, 2000.

CZECH

Filosofický slovník. 2nd ed. Blecha, Ivan. Olomouc: Nakl. Olomouc, 1998.

Filozofický slovník. Iudin, Pavel, and Mark Moiseevich Rozental', eds. [Bratislava]: Vydavatel'stvo politickej literatury, 1965.

Filozofický slovník. Javurek, Zdenek, E. Marešová, and M. Landová. Prague: Svoboda, 1976.

Filozofický slovník. Klaus, Georg, and Manfred Buhr. Prague: Svoboda, 1985.

Filozofický slovník pro samouky: neboli Antigorgias. 2nd ed. Neff, Vladimír. Prague: Mladá fronta, 1993.

Filozovický slovník. 4th ed. Frolov, Ivan Timofeevich. Bratislava: Pravda, 1989.

Slovník ceských filozofu. Gabriel, Jirí. Brno: Masarykova univerzita v Brne, 1998.

Slovník filosofických pojmu soucasnosti. Olsovský, Jirí. Prague: Erika, 1999.

Strucný prehled významných filosofu a filosofických pojmu. Hrubes, Jaromír. Ostrava: Ostravská univerzita, Pedagogická fakulta, 2002.

Strucny filosofický slovník. [Prague]: Svoboda, 1966.

DANISH

Filosofisk ordbog. Ord og udtryk fra oldtidens til dagens tankeverden. Hanneborg, Bente, Knut Hanneborg, and Carl Henrik Koch. Copenhagen: Høst, 1971.

DUTCH

Elseviers filosofische en psychologische encyclopedie. Kuypers, K. Amsterdam: Elsevier, 1970.

Encyclopedie van de filosofie. Kuypers, K. Amsterdam: Elsevier, 1977.

Encyclopedie van de filosofie. Waldram, Joop, and T. van Kooten. Baarn: Tirion, 1991.

Klein wijsgerig woordenboek. Schilfgaarde, Paul van. Wassenaar: Servire, 1968.

Kritisch denkerslexicon. Achterhuis, Hans. Alphen: Samsom, 1986–.

Prisma van de filosofie. Groen, Nico, and Robbert Veen. Utrecht: Het Spectrum, 1990.

Woordeboek filosofie. Willemsen, Harry. Assen: Van Gorcum, 1992.

ESTONIAN

Filosoofia leksikon. 4th ed. Frolov, Ivan Timofeevich. Tallinn: Eesti Raamat, 1985.

Lühike filosoofia leksikon. Lumi, M. Tallinn: Eesti Raamat, 1975.

FRENCH

Abécédaire de l'engagement. Benasayag, Miguel, and Béatrice Bouniol. Paris: Bayard, 2004.

L'altérité. Groux, Dominique, and Louis Porcher. Paris: Harmattan, 2003.

Bibliographie de la philosophie: glossaire [Bibliography of Philosophy: Glossary]. Unesco, Centre national de la recherche scientifique (France), and International Institute of Philosophy. Paris: J. Vrin, 1995.

Les 50 mots-clés de la philosophie contemporaine. Mantoy, Jacques. [Toulouse]: privately printed, 1971.

Dico de philosophie. Vergely, Bertrand. Toulouse: Milan, 1998.

Dictionnaire abrégé des philosophes médiévaux. Patar, Benoît. [Longueuil, Québec]: Presses philosophiques, 2000.

Dictionnaire critique à l'usage des incrédules. Memmi, Albert. Paris: Félin, 2002.

Dictionnaire de la langue philosophique. 2nd ed. Foulquié, Paul. Paris: Presses Universitaires de France, 1969.

Dictionnaire de la langue philosophique. 3rd ed. Foulquié, Paul. Paris: Presses Universitaires de France, 1978.

Dictionnaire de la philosophie. Comte-Sponville, André. Paris: Encyclopaedia Universalis, A. Michel, 2000.

Dictionnaire de la philosophie. Rev. ed. Julia, Didier. Paris: Larousse, 1970.

Dictionnaire de la philosophie. Rev. ed. Julia, Didier. Paris: Larousse, 1984.

Dictionnaire de la philosophie. Rev. ed. Julia, Didier. Paris: Larousse, 1991.

Dictionnaire de la philosophie. Rev. ed. Julia, Didier. Paris: Larousse, 1995.

Dictionnaire de la philosophie. Rev. ed. Julia, Didier. Paris: Larousse, 1998.

Dictionnaire de philosophie. Baraquin, Noëlla. Paris: A. Colin, 1995.

Dictionnaire de philosophie. 2nd ed. Baraquin, Noëlla. Paris: A. Colin, 2000.

Dictionnaire de philosophie. Durozoi, Gérard, and André Roussel. Paris: Nathan, 1987.

Dictionnaire de philosophie. Durozoi, Gérard, and André Roussel. Paris: Nathan, 1990.

Dictionnaire de philosophie. Durozoi, Gérard, and André Roussel. Paris: Nathan, 1997.

Dictionnaire de philosophie. Godin, Christian. Paris: Fayard, 2004.

Dictionnaire de philosophie. Legrand, Gérard. Paris: Bordas, 1973.

Dictionnaire de philosophie. Legrand, Gérard. Paris: Bordas, 1983.

Dictionnaire de philosophie. Russ, Jacqueline, and Clotilde Badal-Leguil. Paris: Bordas, 2004.

Dictionnaire de philosophie: les concepts, les philosophes, 1850 citations. Russ, Jacqueline. Paris: Bordas, 1991.

Dictionnaire de philosophie politique. Raynaud, Philippe, and Stéphane Rials. Paris: Presses Universitaires de France, 2003.

Dictionnaire des auteurs et des thèmes de la philosophie. Auroux, Sylvain, and Yvonne Weil. Paris: Hachette, 1975.

Dictionnaire des auteurs et des thèmes de la philosophie. Auroux, Sylvain, and Yvonne Weil. Paris: Hachette, 1984.

Dictionnaire des grandes philosophies. Jerphagnon, Lucien. Toulouse: privately printed, 1973.

Dictionnaire des idées contemporaines. Mourre, Michel, ed. Paris: Éditions Universitaires, 1966.

Dictionnaire des philosophes. 2nd ed. Baraquin, Noëlla, and Jacqueline Laffitte. Paris: A. Colin, 2000.

Dictionnaire des philosophes. Comte-Sponville, André. Paris: Encyclopaedia Universalis, A. Michel, 1998.

Dictionnaire des philosophes. Huisman, Denis, and Jean-François Braunstein. Paris: Presses Universitaires de France, 1984.

Dictionnaire des philosophes. 2nd ed. Huisman, Denis, and Marie-Agnès Malfray. Paris: Presses Universitaires de France, 1993.

Dictionnaire des philosophes antiques. Goulet, Richard. Paris: Editions du Centre national de la recherche scientifique, 1989–.

Dictionnaire des philosophes antiques. Goulet, Richard. Paris: Editions du Centre national de la recherche scientifique, 1994–.

Dictionnaire du Moyen âge: littérature et philosophie. Favier, Jean. Paris: Encyclopaedia Universalis, A. Michel, 1999.

Dictionnaire philosophique de citations. Rev. ed. Grateloup, Léon-Louis. Paris: Hachette, 1990.

Dictionnaire philosophique en langue arabe: avec index des termes français, anglais et latins. Saliba, Djemil. Beirut: Dar al-Kitab al-Lubnani, 1978–1979.

Dictionnaire Sartre. Noudelmann, François, and Gilles Philippe. Paris: Champion, 2004.

Dictionnaire Wittgenstein. Glock, Hans-Johann. [Paris]: Gallimard, 2003.

Encyclopédie philosophique universelle. Jacob, André. Paris: Presses Universitaires de France, 1989–1998.

Grand dictionnaire de la philosophie. Blay, Michel. Paris: Larousse, 2003.

Les grandes notions de la philosophie. Warin, François. Paris: Ellipses, 2001.

Histoire et mémoire. Le Goff, Jacques. [Paris]: Gallimard, 1988.

Lexique de philosophie. Stirn, François, and Hervé Vautrelle. Paris: A. Colin, 1998.

Lexique philosophique: français-arabe. Gaïd, Tahar. Algiers: Office des Publications Universitaires, 1990.

Les Notions philosophiques: dictionnaire. Auroux, Sylvain, and André Jacob. Paris: Presses Universitaires de France, 1990.

Nouveau vocabulaire des études philosophiques. Auroux, Sylvain, and Yvonne Weil. Paris: Hachette, 1975.

Nouveau vocabulaire philosohique. 13th ed. Cuvillier, Armand. Paris: A. Colin, 1967.

Petit dictionaire de la philosophie. Huisman, Denis, and André Vergez. Paris: F. Nathan, 1971.

Petit dictionnaire philosophique. Rozental', Mark Moiseevich, and Pavel Iudin, eds. Paris: E. Varlin, 1977.

Petit lexique philosophique de l'anarchisme: de Proudhon à Deleuze. Colson, Daniel. Paris: Livre de poche, 2001.

Le Petit Retz des nouvelles idées en philosophie. Huisman, Denis, and Serge Le Strat. Paris: Retz, 1987.

Philosopher au cégep: guide d'apprentissage, de vocabulaire et d'histoire. Clavet, Jean-Claude, and G.-Magella Hotton. Sainte-Foy, Quebec: Éditions Le Griffon d'argile, 1993.

Philosophes de l'humanité (Grèce, Rome, Inde, Chine, Perse, Islam, Europe médiévale et moderne). Koribaa, Nabhani. Paris: Publisud, 1995.

La Philosophie. Verviers, Belgium: Gérard, 1972.

La Philosophie. Paris: Le Livre de poche, 1977.

La Philosophie. Akoun, André. Paris: Centre d'étude et de promotion de la lecture, 1977.

La Philosophie. Noiray, André, ed. Paris: Centre d'étude et de promotion de la lecture, 1969.

La Philosophie: encyclopédie du monde actuel. Favrod, Charles-Henri. Paris: C. H. Favrod, 1977.

Philosophie, religions. Paris: Bordas. Brussels: ASEDI. Lausanne: Spes. Montréal: CEC, 1968.

Philosophie, religions. Rev. ed. Caratini, Roger, and Françoise Caratini. Paris: Bordas, 1976.

Philosophie, religions. Pascal, Georges. Paris: Bordas, 1974.

Pierre Bayle, pour une histoire critique de la philosophie: choix d'articles philosophiques du Dictionnaire historique et critique. Bayle, Pierre, Jean-Michel Gros, and Jacques Chomarat. Paris: Champion, 2001.

Pour une encyclopédia des auteurs classiques. Colli, Giorgio, Jean Christophe Bailly, Jean-Paul Manganaro, Danielle Dubroca, and Giuliana Lanata. N.p.: Christian Bourgois Editeur, 1990.

Le taureau de Phalaris: dictionnaire philosophique. Matzneff, Gabriel. Paris: Table Ronde, 1987.

Le taureau de Phalaris: dictionnaire philosophique. Matzneff, Gabriel. Paris: Table Ronde, 1994.

Vocabulaire Bordas de la philosophie. Legrand, Gérard. Paris: Bordas, 1986.

Vocabulaire Bordas de la philosophie. Legrand, Gérard. Paris: Bordas, 1987.

Vocabulaire Bordas de la philosophie. Legrand, Gérard. Paris: Bordas, 1993.

Vocabulaire de la philosophie et des sciences humaines. Morfaux, Louis Marie. Paris: A. Colin, 1980.

Vocabulaire de la philosophie; suivi d'un Tableau Historique des Écoles de Philosophie. 6th ed. Jolivet, Régis. Lyon: Emmanuel Vitte, 1966.

Vocabulaire des études philosophiques. Rev. ed. Auroux, Sylvain, Yvonne Weil, and Yvon Belaval. Paris: Hachette, 1993.

Le vocabulaire des philosophes. Zarader, Jean-Pierre. Paris: Ellipses, 2002.

Vocabulaire européen des philosophies: dictionnaire des intraduisibles. Cassin, Barbara. Paris: Le Robert, Seuil, 2004.

Le vocabulaire grec de la philosophie. Gobry, Ivan. Paris: Ellipses, 2000.

Vocabulaire philosophique arabe (avec tables françaises et anglaises). Karam, Yusuf, Murad Wahbah, and Youssef Chellalah. Cairo: Tsoumas 1966.

Le vocabulaire philosophique: français-arabe. Hulw, ʿAbduh. Beirut: Centre de recherche et de développement pédagogique, Librairie du Liban, 1994.

Vocabulaire technique et critique de la philosophie. 10th ed. Lalande, André. Paris: Presses Universitaires de France, 1968.

Vocabulaire technique et critique de la philosophie. 11th ed. Lalande, André. Paris: Presses Universitaires de France, 1972.

Vocabulaire technique et critique de la philosophie. 16th ed. Lalande, André. Paris: Presses Universitaires de France, 1988.

Vocabulaire technique et critique de la philosophie. Lalande, André. Paris: Presses Universitaires de France, 1992.

GERMAN

Chinesisch-Deutsches Lexikon der Chinesischen Philosophie. Geldsetzer, Lutz, and Han-Ding Hong. Aalen: Scientia, 1986.

Chinesisch-Deutsches Lexikon der Chinesischen Philosophie. Geldsetzer, Lutz, and Han-Ding Hong. Aalen: Scientia, 1995.

Chinesisch-Deutsches Lexikon der Klassiker und Schulen der Chinesischen Philosophie. Geldsetzer, Lutz, and Han-Ding Hong. Aalen: Scientia, 1991.

Enzyklopädie der Philosophie: von der Antike bis zur Gegenwart; Denker und Philosophen, Begriffe und Probleme, Theorien und Schulen. Augsburg: Gruppo, 1992.

Enzyklopädie Philosophie und Wissenschaftstheorie. Stuttgart: Metzler, 1995–1996.

Enzyklopädie Philosophie und Wissenschaftstheorie. Blasche, Siegfried, Gereon Wolters, and Jürgen Mittelstrass. Mannheim: Bibliographisches Institut, 1980–1996.

Enzyklopädie Philosophie und Wissenschaftstheorie. Blasche, Siegfried, Martin Carrier, Gereon Wolters, and Jürgen Mittelstrass. Stuttgart: Metzler, 1995–.

Europäische Enzyklopädie zu Philosophie und Wissenschaften. Sandkühler, Hans-Jörg, and Arnim Regenbogen. Hamburg: F. Meiner, 1990.

Grundbegriffe der altchinesischen Philosophie: ein Wörterbuch für die Klassische Periode. Unger, Ulrich. Darmstadt: Wissenschftliche Buchgesellschaft, 2000.

Grundbegriffe der analytischen Philosophie. Prechtl, Peter, and Ansgar Beckermann. Stuttgart: Metzler, 2004.

Grundbegriffe der antiken Philosophie: ein Lexikon. Bächli, Andreas, and Andreas Graeser. Stuttgart: P. Reclam, 2000.

Grundbegriffe der philosophischen Sprache. 2nd ed. Neuhäusler, Anton. Munich: Ehrenwirth, 1967.

Handbuch philosophischer Grundbegriffe. Krings, Hermann, Hans Michael Baumgartner, and Christoph Wild, eds. Munich: Kösel-Verlag, 1973.

Handlexikon zur Wissenschaftstheorie. Seiffert, Helmut, and Gerard Radnitzky. Munich: Ehrenwirth, 1989.

Handwörterbuch Philosophie. Rehfus, Wulff D. Göttingen: Vandenhoeck & Ruprecht; [Stuttgart]: UTB, 2003.

Herders kleines philosophisches Wörterbuch. 8th ed. Müller, Max, and Alois Halder. [Basel]: Herder-Bücherei, 1966.

Historisches Wörterbuch der Philosophie. Ritter, Joachim, and Rudolf Eisler. Wörterbuch der philosophischen Begriffe. Darmstadt: Wissenschaftliche Buchgesellschaft, 1971–.

Historisches Wörterbuch der Philosophie. Ritter, Joachim, Karlfried Gründer, Gottfried Gabriel, and Rudolf Eisler, eds. Wörterbuch der philosophischen Begriffe. Basel: Schwabe, 1971–.

Historisches Wörterbuch der Philosophie. Ritter, Joachim, Karlfried Gründer, and Rudolf Eisler, eds. Wörterbuch der philosophischen Begriffe. Basel: Schwabe, 1992.

Jugendlexikon Philosophie: Geschichte, Begriffe und Probleme der Philosophie. Delf, Hanna. Hamburg: Rowohlt, 1988.

Kleines philosophisches Wörterbuch. Müller, Max, and Alois Halder, eds. Freiburg, Basel, and Vienna: Herder, 1971.

Kleines philosophisches Wörterbuch. 7th ed. Müller, Max, and Alois Halder, eds. Freiburg, Basel, and Vienna: Herder, 1979.

Kleines Wörterbuch der angewandten Philosophie. Mosch, Annemarie. Zürich: Haffman, 1996.

Kleines Wörterbuch der marxistischen Philosophie. Buhr, Manfred, and Alfred Kosing. Berlin: Das Europäische Buch (Dietz), 1967.

Kleines Wörterbuch der marxistisch-leninisten Philosophie. 3rd ed. Buhr, Manfred, and Alfred Kosing. Berlin: Dietz, 1975.

Kleines Wörterbuch der marxistisch-leninisten Philosophie. 4th ed. Buhr, Manfred, and Alfred Kosing. Berlin: Dietz, 1979.

Kleines Wörterbuch der marxistisch-leninistischen Philosophie. Buhr, Manfred. Berlin: Dietz, 1966.

Kleines Wörterbuch der marxistisch-leninistischen Philosophie. 2nd ed. Buhr, Manfred, and Alfred Kosing. Berlin: Dietz, 1974.

Kleines Wörterbuch der marxistisch-leninistischen Philosophie. 5th ed. Buhr, Manfred, and Alfred Kosing. [Berlin]: Das Europäische Buch, 1981.

Kleines Wörterbuch der marxistisch-leninistischen Philosophie. 7th ed. Buhr, Manfred, and Alfred Kosing. Berlin: Dietz, 1984.

Kleines Wörterbuch der Wissenschaftstheorie. Schwarz, Gerhard, and Walter Adolf Jöhr. Saint Gall, Switzerland: Hochschule St. Gallen, 1974.

Kleines Wörterbuch zum Verständnis asiatischer Weltanschauung. Erklärungen von Worten und Begriffen. Mangoldt, Ursula von. Weilheim and Oberbayern: Barth, 1966.

Die Kunst, selber zu denken: ein philosophischer Dictionnaire. Sommer, Andreas Urs. Frankfurt am Main: Eichborn, 2002.

Lexikon der Erkenntnistheorie und Metaphysik. Ricken, Friedo. Munich: C. H. Beck, 1984.

Lexikon der Esoterik. Iwersen, Julia. Düsseldorf: Artemis & Winkler, 2001.

Lexikon der Philosophie. 5th ed. Austeda, Franz. Vienna: Hollinek, 1979.

Lexikon der Philosophie. 6th ed. Austeda, Franz. Vienna: Hollinek, 1989.

Lexikon der Philosophischen Begriffe. Ulfig, Alexander. Wiesbaden: Fourier, 1997.

Lexikon der Philosophischen Begriffe. 2nd ed. Ulfig, Alexander. Wiesbaden: Fourier, 1999.

Lexikon des Kritischen Rationalismus. Niemann, Hans-Joachim. Tübingen: Mohr Siebeck, 2004.

Marxistisch-Leninistisches Wörterbuch der Philosophie. Klaus, Georg, and Manfred Buhr, eds. Reinbek: Rowohlt, 1972.

Marxistisch-Leninistisches Wörterbuch der Philosophie. Rev. ed. Klaus, Georg, Manfred Buhr, and Karlheinz Barck, eds. Reinbek: Rowohlt, 1975.

Metzler Lexikon jüdischer Philosophen: Philosophisches Denken des Judentums von der Antike bis zur Gegenwart. Kilcher, Andreas B., Otfried Fraisse, and Yossef Schwartz. Stuttgart: Metzler, 2003.

Metzler Philosophie Lexikon: Begriffe und Definitionen. Prechtl, Peter, and Franz-Peter Burkard. Stuttgart: Metzler, 1996.

Metzler Philosophie Lexikon: Begriffe und Definitionen. 2nd ed. Prechtl, Peter, and Franz-Peter Burkard. Stuttgart: Metzler, 1999.

Onomasticon philosophicum: Latinoteutonicum et Teutonicolatinum. Aso, Ken. Tokyo: Tetsugaku-Shobo, 1989.

Philosophenlexikon. Lange, Erhard, and Dietrich Alexander. Berlin: Das Europäische Buch, 1982.

Philosophenlexikon. Lange, Erhard, and Dietrich Alexander. Berlin: Dietz, 1984.

Philosophie. Diemer, Alwin, and Ivo Frenzel. [Frankfurt am Main]: Fischer Bücherei, 1967.

Philosophie. 11th ed. Diemer, Alwin, and Ivo Frenzel. [Frankfurt am Main and Hamburg]: Fischer Bücherei, 1969.

Philosophie. 4th ed. Fiedler, Frank, and Günter Gurst. Leipzig: Bibliographisches Institut, 1986.

Philosophie. 6th ed. Fiedler, Frank, and Günter Gurst. Leipzig: Bibliographisches Institut, 1988.

Philosophie in der Geschichte ihres Begriffs: Historisches Wörterbuch der Philosophie, Sonderdruck. Gründer, Karlfried. Basel: Schwabe, 1990.

Der Philosophische Flohmarkt: Kleines Lexikon der Philosophie. Moser, Friedhelm. Frankfurt am Main: Eichborn, 1995.

Philosophisches Wörterbuch. 6th ed. Apel, Max, and Peter Christian Ludz. Berlin and New York: de Gruyter, 1976.

Philosophisches Wörterbuch. 14th ed. Brugger, Walter, ed. Freiburg, Basel, and Vienna: Herder, 1976.

Philosophisches Wörterbuch. 15th ed. Brugger, Walter, ed. Freiburg, Basel, and Vienna: Herder, 1978.

Philosophisches Wörterbuch. Halder, Alois, and Max Müller. Freiburg: Herder, 1988.

Philosophisches Wörterbuch. Rev. ed. Halder, Alois, and Max Müller. Freiburg: Herder, 1993.

Philosophisches Wörterbuch. Rev. ed. Halder, Alois, and Max Müller. Freiburg: Herder, 2000.

Philosophisches Wörterbuch. 2nd ed. Klaus, Georg, and Manfred Buhr, eds. Leipzig: Bibliographisches Institut, 1965.

Philosophisches Wörterbuch. 6th ed. Klaus, Georg, and Manfred Buhr, eds. Leipzig: VEB Bibliographisches Institut; Berlin: Das Europäische Buch, 1969.

Philosophisches Wörterbuch. 7th ed. Klaus, Georg, and Manfred Buhr, eds. Berlin: Das Europäische Buch; Leipzig: VEB Bibliographisches Institut, 1970.

Philosophisches Wörterbuch. 8th ed. Klaus, Georg, and Manfred Buhr, eds. Leipzig: VEB Verlag Enzyklopädie; Berlin: Das Europäische Buch, 1972.

Philosophisches Wörterbuch. 10th ed. Klaus, Georg, and Manfred Buhr, eds. Leipzig: Bibliographisches Institut, 1974.

Philosophisches Wörterbuch. 12th ed. Klaus, Georg, and Manfred Buhr, eds. Berlin: Das Europäische Buch; Leipzig: VEB Bibliographisches Institut, 1976.

Philosophisches Wörterbuch. 13th ed. Klaus, Georg, and Manfred Buhr, eds. Berlin: Das Europäische Buch, 1985.

Philosophisches Wörterbuch. 14th ed. Klaus, Georg, and Manfred Buhr, eds. Berlin: Das Europäische Buch, 1987.

Philosophisches Wörterbuch. 9th ed. Schmidt, Heinrich, ed. Stuttgart: A. Kröner, 1974.

Philosophisches Wörterbuch. 17th ed. Schmidt, Heinrich, and Georgi Schischkoff, eds. Stuttgart: A. Kröner, 1965.

Philosophishces Wörterbuch. 18th ed. Schmidt, Heinrich, and Georgi Schischkoff, eds. Stuttgart: A. Kröner, 1969.

Philosophisches Wörterbuch. 19th ed. Schmidt, Heinrich, and Georgi Schischkoff, eds. Stuttgart, A. Kröner, 1974.

Philosophisches Wörterbuch. 20th ed. Schmidt, Heinrich, and Georgi Schischkoff, eds. Stuttgart: Kröner, 1978.

Philosophisches Wörterbuch. 22nd ed. Schmidt, Heinrich, and Georgi Schischkoff, eds. Stuttgart: A. Kröner, 1991.

Philosophisches Wörterbuch. Stockhammer, Morris. [Cologne]: Kölner Universitäts-Verlag 1967.

Philosophisches Wörterbuch. Stockhammer, Morris. Essen: Magnus, 1980.

Philosophisches Wörterbuch. Unter Mitwirkung der Professoren des Berchmanskollegs in Pullach bei München und anderer. 12th ed. Brugger, Walter, ed. Freiburg: Herder, 1965.

Philosophisches Wörterbuch. Unter Mitwirkung der Professoren des Berchmanskollegs. 13th ed. Brugger, Walter, ed. Freiburg, Basel, and Vienna: Herder, 1967.

Schülerduden, die Philosophie. Kwiatkowski, Gerhard. Mannheim, Vienna, and Zürich: Dudenverlag, 1985.

Die sowjetische "Philosophische Enzyklopädie"; Dokument eines Politischen Umbruchs. Marko, Kurt. Cologne: Bundesinstitut für Ostwissenschaftliche und Internationale Studien, 1973.

Terminologie der frühen Philosophischen Scholastik in Indien: ein Begriffswörterbuch zur altindischen Dialektik, Erkenntnislehre und Methodologie. Oberhammer, Gerhard, Ernst Prets, and Joachim Prandstetter. Vienna: Österreichischen Akademie der Wissenschaften, 1991–.

Wissenschaftstheoretisches Lexikon. Braun, Edmund, and Hans Radermacher. Graz, Vienna, and Cologne: Styria, 1978.

Wörterbuch der antiken Philosophie. Horn, Christoph, and Christof Rapp. Munich: C. H. Beck, 2002.

Wörterbuch der Marxistisch-Leninistischen Philosophie. Kosing, Alfred. Berlin: Dietz, 1985.

Wörterbuch der Philosophie. Hegenbart, Rainer. Munich: Humboldt-Taschenbuchverlag, 1984.

Wörterbuch der Philosophie. Kosing, Alfred. Berlin: Das Europäische Buch, 1985.

Wörterbuch der Philiosophie. Pachnicke, Erich, and Richard Hönigswald. Göttingen: K. Pachnicke, 1985.

Wörterbuch der philosophischen Begriffe. Regenbogen, Arnim, and Uwe Meyer. Hamburg: F. Meiner, 1998.

GREEK

Hermeneutiko lexiko tes Indikes philosophias kai tou Gionka. Pantouvas, Theodoros. Athens: Ekdoseis Kardamitsa, 1989.

Lexikon tes prosokratikes philosophias [Lexicon of Presocratic Philosophy]. Despototpoulos, Konstantinos I. Athens: Akademia Athenon and Kentron Ereunes tes Hellenikes Philosophias, 1988.

Lexikon tes prosokratikes philosophias. Theodoracopoulos, Ioannes. Athens: Akademia Athenon, 1998.

Lexikon vasikon ennoion. Sarales, I. A. Athens, 1974.

Lexiko philosophikon horon. Dokas, Agesilaos S. Athens: Ekdot, Oikos "Aster," 1981.

Lexiko vasikon ennoion tou hylikou-technikou, pneumatikou kai ethikou politismou: themeliodeis gnoseis ennoion gia tis ektheseis. Diamantopoulos, D. P. Athens: Patake, 1984.

Neo philosophiko lexiko. Gkikas, Sokrates. Athens: Ekdoseis Savvalas, 2002.

Philosophiko lexiko. 3rd ed. Gkikas, Sokrates. Athens: Ekdoseis Pheleke, 1982.

HUNGARIAN

Filozófiai kislexikon. Szigeti, Györgyné. Budapest: Kossuth Könyvkiadó, 1970.

Filozófiai kislexikon. 2nd ed. Szigeti, Györgyné, and Györgyné Vári, eds. Budapest: Kossuth Könyvkiadó, 1972.

Filozófiai kislexikon. 3rd ed. Szigeti, Györgyné, Györgyné Vári, and Árpád Volczer. [Budapest]: Kossuth Könyvkiadó, 1973.

Filozófiai kislexikon. 4th ed. Tomori, Lajos. [Budapest]: Kossuth Könyvkiadó, 1976.

Idegen szavak a filozófiában: az arkhétól a zoón politikonig. Rathmann, János. [Budapest]: Kossuth, 1988.

ITALIAN

Antidizionario filosofico. Dizionario critico di filosofia. Testa, Aldo. Urbino: AGE, 1974.

Breve dizionario filosofico. Del Re, Raffaello. Rome: Edizioni dell'Ateneo & Bizzarri, 1979.

Concetti e problemi di filosofia. Garulli, Enrico. [Lanciano]: Editrice Itinerari, 1983.

Concetti fondamentali della filosofia. Krings, Hermann, Hans Michael Baumgartner, and Christoph Wild. Brescia: Queriniana, 1981–1982.

Dizionario critico di filosofia: rivisto dai membri e dai corrispondenti della Società francese di filosofia e pubblicato con le loro correzioni e osservazioni. 3rd ed. Lalande, André, and Mario Dal Pra. Milan: ISEDI: Mondadori, 1980.

Dizionario dei filosofi. Centro di studi filosofici di Gallarate. Florence: G. C. Sansoni, 1976.

Dizionario dei filosofi del Novecento. Centro di studi filosofici di Gallarate. [Florence]: L. S. Olschki, 1985.

Dizionario delle idee. Centro di studi filosofici di Gallarate. Florence: G. C. Sansoni, 1977.

Dizionario di filosofia. Abbagnano, Nicola. [Turin]: Unione tipografico-editrice torinese, 1968.

Dizionario di filosofia. 2nd ed. Abbagnano, Nicola. Turin: Unione tipografico-editrice torinese, 1971.

Dizionario di filosofia. 3rd ed. Abbagnano, Nicola, and Giovanni Fornero. Turin: UTET libreria, 2001.

Dizionario di filosofia. Rossi, Paolo. Florence: La nuova Italia, 1996.

Dizionario di filosofia. Runes, Dagobert David, ed. Milan: Arnoldo Mondadori, 1972.

Dizionario di filosofia: con un prospetto storico delle scuole di filosofia. Jolivet, Régis, and Luigi Castiglione. Brescia: Morcelliana, 1966.

Dizionario di filosofia: 2500 voci, 700 bibliografie, tavole cronologiche. Milan: Rizzoli, 1976.

Dizionario di filosofia: 2500 voci, 700 bibliografie, tavole cronologiche. 3rd ed. Milan: Rizzoli, 1977.

Dizionario di termini filosofici. 16th ed. Lamanna, Eustachio Paolo, and Francesco Adorno. Florence: F. Le Monnier, 1968.

Dizionario di termini filosofici. 17th ed. Lamanna, Eustachio Paolo, and Francesco Adorno. Florence: F. Le Monnier, 1969.

Dizionario filosofico. Piebe, Armando. Padua: R.A.D.A.R., 1969.

Dizionario filosofico. Zamboni, Giuseppe, and Ferdinando L. Marcolungo. Milan: Vita e pensiero, 1978.

Un dizionario filosofico ebraico del XIII secolo: l'introduzione al "Sefer De'ot ha-filosofim" di Shem Tob ibn Falaquera. Zonta, Mauro, and Shem Tov ben Joseph Falaquera. Turin: S. Zamorani, 1992.

Dizionario interdisciplinare di scienza e fede: cultura scientifica, filosofia e teologia. Tanzella-Nitti, G. and Alberto Strumia. Vatican City: Urbaniana University Press; Rome: Città Nuova, 2002.

Enciclopedia della filosofia contemporanea. Papi, Fulvio. Milan: Teti, 1979.

Enciclopedia filosofica. Rev. ed. Centro di studi filosofici di Gallarate.Florence: G. C. Sansoni, 1967.

Enciclopedia filosofica. Rev. ed. Centro di studi filosofici di Gallarate. Rome: Edipem, 1979.

Enciclopedia filosofica. Rev. ed. Centro di studi filosofici di Gallarate. [Rome?]: Lucarini, 1982.

Filosofia. Apel, Karl-Otto, and Carlo Sini. Milan: Jaca, 1992.

Il linguaggio dei filosofi: Kant, Fichte, Schelling, Hegel. Campagna, Nunzio. Naples: La città del sole: Istituto italiano per gli studi filosofici, 1998.

Parole che contano: da amicizia a volontà, piccolo dizionario politico-filosofico. Bencivenga, Ermanno. Milan: Mondadori, 2004.

Il pensiero quotidiano: piccolo sillabario filosofico per tutti. Guarini, Ruggero. Milan: BUR, 1993.

Significato, metafora e interpretazione. Welby, Lady Victoria, and Susan Petrilli. Bari: Adriatica, 1986.

Termini della filosofia contemporanea. Plebe, Armando. Rome: A. Armando, 1966.

La trama e l'ordito: storia della filosofia occidentale. Maranini, Letizia, and Ester Dolce. Milan: Hoepli, 1990.

Voci di scienze umane: dal Dizionario critico di filosofia. Testa, Aldo. Rome: Biblioteca del dialogo, 1974.

LATVIAN

Filozofijas vardniica A–Z. Rozental', Mark Moiseevich. Riga: Izdevnieciba "Liesma," 1974.

LITHUANIAN

Filosofijos zodynas. Rozental', Mark Moiseevich, and Romanas Pleckaitis, eds. Vilnius: Mintis, 1975.

MOLDAVIAN

Diktsionar de filozofie. Frolov, Ivan Timofeevich. Kishinev: Red. princlipale a Enchiklopedìei Sovietiche Moldovenesht', 1985.

NORWEGIAN

Filosofisk ordbok: [en oppslagsbok fra Tanum]. Hanneborg, Bente, and Knut Hanneborg. Oslo: Tanum, 1975.

POLISH

Filozofia a nauka: zarys encyklopedyczny. Cackowski, Zdzislaw, and Maria Izewska. Wrocław: Zaklad Narodowy im. Ossolinskich, 1987.

Filozofowie wspólczesni: leksykon. Szmyd, Jan. Kraków: Wydawn. Krakowskie, 2003.

Idee w Rosji: leksykon rosyjsko-polsko-angielski [Ideas in Russia]. Lazari, Andzhei. Warsaw: Semper, 1999–.

Leksykon filozoficzny dla mlodziezy. Delf, Hanna. Warsaw: Wiedza Powszechna, 1996.

Leksykon filozofii klasycznej. Herbut, Józef. Lublin: Tow. Nauk. Katolickiego Uniwersytetu Lubelskiego, 1997.

Leksykon filozofii: postaci i pojecia. Andrzejewski, Boleslaw. Poznań: Oficyna Wydawnicza De Facto, 2000.

Leksykon filozofów wspólczesnych. Jasinski, Boguslaw. Warszaw: Wydawn. Ethos, 1999.

Logika dla prawników: slownik encyklopedyczny. Lewandowski, Slawomir, Andrzej Malinowski, and Jacek Petzel. Warsaw: Wydawn. Prawnicze LexisNexis, 2004.

Maly slownik terminów i pojec filozoficznych, dla studiujacych filozofie chrzescijanska. Podsiad, Antoni, and Zbigniew Wieckowski. Warsaw: Instytut Wydawniczy Pax, 1983.

Podstawowe kategorie i pojecia filozofii. 2nd ed. Hull, Zbigniew, Józef Staranczak, and Witold Tulibacki. Olsztyn: [AR-T], 1976.

Slownik filozoficzny. Lacey, Alan Robert, and Roman Matuszewski. Poznań: Zysk i S-ka Wydawn, 1999.

Slownik filozofii marksistowskiej. Jaroszewski, Tadeusz M. Warsaw: Wydawn. Wiedza Powszechna, 1982.

Slownik filozofów: filozofia powszechna. Andrzejewski, Boleslaw. Poznań: Dom Wydawniczy Rebis, 1995.

Slownik filozofów polskich. Andrzejewski, Boleslaw, and Roman Kozlowski. Poznań: Wydawn, 1999.

Slownik pojec filozoficznych. Krajewski, Wladyslaw, and Ryszard Banajski. Warsaw: Wydawn, 1996.

Slownik terminów i pojec filozoficznych. Podsiad, Antoni. Warsaw: Instytut Wydawniczy Pax, 2000.

Szkolny slownik filozoficzny. Syjud, Jerzy, and Krzysztof Hermann. Katowice: Videograf II, 2000.

PORTUGUESE

Dicionário básico de filosofia. Japiassu, Hilton, and Danilo Marcondes de Souza Filho. Rio de Janeiro: J. Zahar, 1990.

Dicionário de filosofia. Abbagnano, Nicola; translated and revised by Alfredo Bosi. São Paolo: Editôra Mestre Jou, [1970?].

Dicionário de filosofia. Ferrater Mora, José. Lisbon: Publicações Dom Quixote, 1977.

Dicionário de filosofia. Legrand, Gérard. Lisbon: Edições 70, 1991.

Dicionário de filosofia. Lobo, António. Lisbon: Paralelo Editora Limitada, 1982.

Dicionario de filosofia e ciências culturais. 4th ed. Santos, Mário Dias Ferreira dos. São Paulo: Editôra Matese, 1966.

Dicionário de filosofia portuguesa. Gomes, Jesué Pinharanda. Lisbon: Publicações Dom Quixote, 1987.

Dicionário do pensamento contemporâneo. Carrilho, Manuel Maria, João Sàágua, and Diogo Pires Aurélio. Lisbon: Publicações Dom Quixote, 1991.

Dicionário universal das idéas. Castanho, César Arruda. São Paulo: Editora Meca, [1978–1986?].

Enciclopédia filosófica. Corbisier, Roland. Petrópolis, Brazil: Editora Vozes, 1974.

Logos: enciclopédia luso-brasileira de filosofia. Lisbon: Verbo, 1989–1992.

Pequeno dicionário filosófico. Pugliesi, Márcio, and Edson Bini. [São Paulo]: Hemus, 1977.

ROMANIAN

Dictionar de filosofie Indiana. Filip, Adrian. Iaşi: Glasul Bucovinei, 1996.

Dictionar de filozofie. Apostol, Pavel. Bucharest: Editura Politica, 1978.

Dictionar de termeni filosofici ai lui Lucian Blaga: introducere prin concepte. Diaconu, Florica, and Marin Diaconu. Bucharest: Univers Enciclopedic, 2000.

Dictionar filosofic: termeni, conceptii, orientari. Bujdoiu, Nicolae. Braşov: Romprint, 2002.

Dictionarul operelor filozofice românesti. Ianosi, Ion, and Vasile N. Morar. Bucharest: Humanitas, 1997.

Mic dictionar filozofic. 2nd ed. Apostol, Pavel. Bucharest: Editura Politica, 1973.

O enciclopedie a filosofiei grecesti. Vladutescu, Gheorghe. Bucharest: Paideia, 2001.

O enciclopedie a filosofiei grecesti: filosofi, filosofii, concepte fundamentale. Vladutescu, Gheorghe. Bucharest: Paideia, 1994–.

Tratat de enciclopedia dreptului. Vallimarescu, Alexandru. Bucharest: Lumina Lex, 1999.

RUSSIAN

Bukvar': Nauka, Filosofiia, Religiia. Moscow, 2001.

Chastotnyi Anglo-Russkii Slovar'. Pod'iazyk Filosofii. Razdel "Etika i estetika." Elizarova, E. F., L. K. Chernetsovskaia, N. B. Karachan, and Leonid Pavlovich Stupin. Saint Petersburg: Izd-vo Leningradskogo universiteta, 1985.

Chelovek: Entsiklopedicheskii Slovar'. Volkov, Iurii Grigor'evich, and Vitalii Semenovich Polikarpov. Moscow: Gardariki, 1999.

Entsiklopediia vysokogo uma. Taranov, Pavel Sergeevich. Moscow: Izd-vo AST, 1997.

Filosofiin tol'. Balkhaazhav, T. Ulaanbaatar: Ulsyn Khevleliin Gazar, 1990.

Filosofskaia mysl' Vostochnykh Slavian: Biobibliograficheskii Slovar'. Ogorodnik, Ivan Vasil'evich, and Leonid Vasil'evich Guberskii. Kiev: Parlamentskoe izdatel'stvo, 1999.

Filosofskii entsiklopedicheskii slovar'. 2nd ed. Averintsev, Sergei Sergeevich. Moscow: Sov. Entsiklopediia, 1989.

Filosofskii entsiklopedicheskii slovar'. Gubskii, E. F., G. V. Korableva, and V. A. Lutchenko. Moscow: INFRA-M, 1997.

Filosofskii entsiklopedicheskii slovar'. Gubskii, E. F., G. V. Korableva, and V. A. Lutchenko. Moscow: INFRA-M, 1998.

Filosofskii entsiklopedicheskii slovar'. Il'ichev, L. F. Moscow: Sov. Entsiklopediia, 1983.

Filosofskii slovar'. 4th ed. Ado, Anatolii Vasil'evich, and Ivan Timofeevich Frolov. Moscow: Izd-vo polit. lit-ry, 1980.

Filosofskii slovar'. 5th ed. Frolov, Ivan Timofeevich, and Anatolii Vasil'evich Ado. Moscow: Izd-vo polit. lit-ry, 1986.

Filosofskii slovar'. 6th ed. Frolov, Ivan Timofeevich, and Anatolii Vasil'evich Ado. Moscow: Izd-vo polit. lit-ry, 1991.

Filosofskii slovar'. 7th ed. Frolov, Ivan Timofeevich. Moscow: Izd-vo Respublika, 2001.

Filosofskii slovar'. 2nd ed. Rozental', Mark Moiseevich, and Pavel Fedorovich Iudin. Moscow: Politizdat, 1968.

Filosofskii slovar'. 3rd ed. Rozental', Mark Moiseevich, and Pavel Fedorovich Iudin. Moscow: Politizdat, 1972.

Filosofskii slovar'. 3rd ed. Rozental', Mark Moiseevich. Moscow: Izd-vo polit. lit-ry, 1975.

Filosofskii slovar' Vladimira Solov'eva. Solov'ëv, Vladimir Sergeevich. Rostov-na-Donu: Feniks, 1997.

Filosofy Rossii XIX–XX stoletii: biografii, idei, trudy. 4th ed. Alekseev, Petr Vasil'evich. Moscow: Akademicheskii proekt, 2002.

Gita: al'ternativa vybora: entsiklopedicheskii sbornik: filosofiia, religiia, istoriia, poeziia. Kulaichev, A. P. Moscow: Trivola, 1999.

Glavnyi trud Kanta: k 200-letiiu vykhoda v svet "Kritiki chistogo razuma." Abramian, Lev Arutiunovich. Yerevan, Armenia: Izd-vo Aiastan, 1981.

Kitaiskaia filosofiia: entsiklopedicheskii slovar'. Titarenko, Mikhail Leont'evich, and Nikolai Viacheslavovich Abaev. Moscow: Mysl', 1994.

Kratkaia filosofskaia entsiklopediia. Gubskii, E. F., G. V. Korableva, and V. A. Lutchenko. Moscow: Progress, 1994.

Kratkii Anglo-Russkii filosofskii slovar. Tsarev, P. V. Moscow: Izd. Mosk. un-ta, 1969.

Kratkii filosofskii slovar'. Alekseev, Aleksandr Petrovich. Moscow: Prospekt, 1998.

Kratkii Russko-Tuvinskii filosofskii slovar'. Gavrilova, Marina Mongushevna, and Maksim Shyyrapovich Artaev. Kyzyl: Tuvinskoe knizhnoe izd-vo, 1996.

Kratkii slovar' po filosofii. Blauberg, Igor' Viktorovich, Pavel Vasil'evich Kopnin, and Igor' Konstantinovich Pantin, eds. Moscow: Izd.-vo polit. lit-ry, 1966.

Kratkii slovar' po filosofii. 2nd ed. Blauberg, Igor' Viktorovich, Pavel Vasil'evich Kopnin, and Igor' Konstantinovich Pantin, eds. Moscow: Politizdat, 1970.

Kratkii slovar' po filosofii. 3rd ed. Blauberg, Igor' Viktorovich, and Igor' Konstantinovich Pantin, eds. Moscow: Politizdat, 1979.

Kratkii slovar' po filosofii. 4th ed. Blauberg, Igor' Viktorovich, Igor' Konstantinovich Pantin, and Nikolai Ivanovich Azarov, eds. Moscow: Izd-vo polit. lit-ry, 1982.

Nemetsko-russkii i Russko-Nemetskii filosofskii slovar'. Zaitseva, Zoia Nikolaevna. Moscow: Izd-vo Moskovskogo universiteta, 1998.

Novaia filosofskaia entsiklopediia: v chetyrekh tomakh. Stepin, Viacheslav Semenovich. Moscow: Mysl', 2000–2001.

Noveishii filosofskii slovar'. 2nd ed. Gritsanov, Aleksandr Alekseevich. Minsk: Interpresservis, 2001.

Proektivnyi filosofskii slovar': novye terminy i poniatiia. Tul'chinskii, Grigorii L'vovich, and Mark Naumovich Epshtein. Saint Petersburg: Aleteiia, 2003.

Russkaia filosofiia: malyi entsiklopedicheskii slovar'. Abramov, A. I. Moscow: Nauka, 1995.

Russkaia filosofiia: slovar'. Maslin, Mikhail Aleksandrovich. Moscow: Izd-vo Respublika, 1995.

Russkaia filosofiia: slovar'. Maslin, Mikhail Aleksandrovich. Moscow: Terra, 1999.

Russkie filosofy: spravochnik. Kornilov, Sergei Vladimirovich. Saint Petersburg: Lan', 2001.

Russkie mysliteli vtoroi poloviny XIX-nachala XX veka: opyt kratkogo biobibliograficheskogo Slovaria. Emel'ianov, Boris Vladimirovich, and Viacheslav Vladimirovich Kulikov. Yekaterinburg: Izd-vo Ural'skogo universiteta, 1996.

Russko-Kirgizskii slovar' geograficheskikh terminov. Amaniliev, Batyrbek. Bishkek, Kyrgyzstan: Ilim, 1967.

Slovar'. Kruglov, Aleksandr. Moscow: Gnosis, 1994–.

Slovar'-spravochnik "Chelovek i obshchestvo": Filosofiia: dlia uchashchikhsia vsekh form srednei stupeni obrazovaniia. Korotets, I. D., L. A. Shtompel', and O. M. Shtompel'. Rostov-na-Donu: Feniks, 1996.

Sovremennaia filosofiia: slovar' i khrestomatiia. Zharov, L. V. Rostov-na-Donu: Feniks, 1995.

Sovremennaia zapadnaia filosofiia: slovar'. Lektorskii, V. A., V. S. Malakhov, and Vladimir Petrovich Filatov. Moscow: Izd-vo polit. lit-ry, 1991.

Sovremennaia zapadnaia filosofiia: slovar'. 2nd ed. Malakhov, V. S., and Vladimir Petrovich Filatov. Moscow: TON, 2000.

Sovremennyi filosofskii slovar'. Kemerov, Viacheslav Evgen'evich. Moscow: Izd-vo Odissei, 1996.

Sovremennyi filosofskii slovar'. 2nd ed. Kemerov, Viacheslav Evgen'evich, and S. N. Dergachev. London: Izd-vo Panprint, 1998.

Sovremennyi filosofskii slovar'. 3rd ed. Kemerov, Viacheslav Evgen'evich. Moscow: Akademicheskii proekt, 2004.

Sto Russkikh filosofov: biograficheskii slovar'. Sukhov, Andrei Dmitrievich. Moscow: Mirta, 1995.

Sto sueverii: kratkii filosofskii slovar' predrassudkov. Bochenski, Joseph M. Moscow: Progress, 1993.

Velikie filosofy: slovar'-spravochnik. 2nd ed. Blinnikov, Leontii Vasil'evich. Moscow: Logos, 1997.

SPANISH

Breve diccionario de filosofía. Müller, Max, and Alois Halder. Barcelona: Editorial Herder, 1976.

Breve diccionario de filosofía. Müller, Max, and Alois Halder. Barcelona: Editorial Herder, 1986.

Breve diccionario filosófico. 2nd ed. Quintanilla, Miguel A. Navarra, Spain: EVD, 1996.

Breve vocabulario filosófico. Salazar Bondy, Augusto. Lima: [Editoral Universo], 1967.

Breve vocabulario filosófico. 3rd ed. Salazar Bondy, Augusto. Lima: Editoral Arica, 1974.

Conceptos fundamentales de filosofía. Krings, Hermann, Hans Michael Baumgartner, and Christoph Wild. Barcelona: Editorial Herder, 1977–1979.

Diccionario básico de filosofía. Lorenzo Lizalde, Carlos, and Andrés Plumed Allueva. Zaragoza: Mira Editores, 1992.

Diccionario bio-bibliográfico de filósofos. Menchaca, José A. Bilbao: El Mensajero del Corazón de Jesús, 1966–.

Diccionario de escuelas de pensamiento o ismos. Arroyo Fernández, Miguel. Madrid: Alderabán, 1997.

Diccionario de filosofía. Moscow: Editorial Progreso, 1984.

Diccionario de filosofía. 2nd ed. Abbagnano, Nicola. Mexico: Fondo de Cultura Económica, 1966.

Diccionario de filosofía. 2nd ed., rev. Abbagnano, Nicola. Mexico: Fondo de Cultura Económica, 1974.

Diccionario de filosofía. 2nd Spanish ed. Abbagnano, Nicola. Mexico: Fondo de Cultura Económica, 1980.

Diccionario de filosofía. Albornoz, Hernán. Valencia, Venezuela: Vadell Hermanos Editores, 1990.

Diccionario de filosofía. Blauberg, Igor' Viktorovich. Mexico, D.F.: Quinto Sol, 1986.

Diccionario de filosofía. 2nd ed. Brugger, Walter, ed. Barcelona: Editorial Herder, 1969.

Diccionario de filosofía. 10th ed. Brugger, Walter, ed. Barcelona: Editorial Herder, 1983.

Diccionario de filosofía. Durozoi, Gérard, and André Roussel. Barcelona: Editorial Teide, 1994.

Diccionario de filosofía. 5th ed. Ferrater Mora, José. Buenos Aires: Editorial Sudamericana, 1965.

Diccionario de filosofía. 6th ed. Ferrater Mora, José. Madrid: Alianza, 1979.

Diccionario de filosofía. Rev. ed. Ferrater Mora, José, and Josep-Maria Terricabras. Barcelona: Editorial Ariel, 1994.

Diccionario de filosofía. González García, Juan Carlos. [Madrid]: Edaf, 2000.

Diccionario de filosofía. 2nd ed. Herrera, Hermogenes. Quito: Santo Domingo, 1968.

Diccionario de filosofía. Julia, Didier. Mexico: Diana, 1983.

Diccionario de filosofía. Rozental', Mark Moiseevich, and Pavel Iudin. Madrid: Akal, 1975.

Diccionario de filosofia. Runes, Dagobert David. Barcelona: Ediciones Grijalbo, 1969.

Diccionario de filosofía. 7th ed. Runes, Dagobert David. Mexico: Editorial Grijalbo, 1981.

Diccionario de filosofía, abreviado. Ferrater Mora, José. Buenos Aires: Sudamericana, 1970.

Diccionario de filosofía abreviado. 2nd ed. Ferrater Mora, José. Buenos Aires: Editorial Sudamericana, 1972.

Diccionario de filosofía abreviado. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Barcelona: EDHASA-Sudamericana, 1976.

Diccionario de filosofía abreviado. 3rd ed. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Barcelona: EDHASA-Sudamericana, 1978.

Diccionario de filosofía abreviado. 5th ed. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Barcelona: EDHASA; Buenos Aires: Sudamericana, 1980.

Diccionario de filosofía abreviado. 19th ed. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Buenos Aires: Editorial Sudamericana, 1991.

Diccionario de filosofía abreviado. 24th ed. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Buenos Aires: Editorial Sudamericana, 2002.

Diccionario de filosofía abreviado. 25th ed. Ferrater Mora, José, Eduardo García Belsunce, and Ezequiel de Olaso. Buenos Aires: Editorial Sudamericana, 2004.

Diccionario de filosofía: con autores y temas latinoamericanos. Bogotá: Editorial El Buho, 1986.

Diccionario de filosofía: con autores y temas latinoamericanos. 2nd ed. Bogotá: Editorial El Buho, 1994.

Diccionario de filosofía contemporánea. Quintanilla, Miguel A. Salamanca: Sígueme, 1976.

Diccionario de filosofía contemporánea. 2nd ed. Quintanilla, Miguel A. Salamanca: Sígueme, 1979.

Diccionario de filosofía de bolsillo. Ferrater Mora, José, and Priscilla Cohn. Madrid: Alianza, 1983.

Diccionario de filosofia de bolsillo. Ferrater Mora, Jose, and Priscilla Cohn. Madrid: Alianza, 2001.

Diccionario de filosofía: ilustrado: autores contemporáneos, lógica, filosofía del lenguaje. 3rd ed. Martínez Echeverri, Leonor, and Hugo Martínez Echeverri. Santa Fe de Bogotá, Colombia: Panamericana Editorial, 1998.

Diccionario de filosofía: ilustrado: autores contemporáneos, lógica, filosofía del lenguaje. 4th ed. Martínez Echeverri, Leonor, and Hugo Martínez Echeverri. Bogotá: Panamericana Editorial, 2000.

Diccionario de filosofía: ilustrado: autores contemporáneos, lógica, filosofía del lenguaje. 6th ed. Martínez Echeverri, Leonor, and Hugo Martínez Echeverri. Bogotá: Panamericana Editorial, 2000.

Diccionario de filosofía latinoamericana. Cerutti Guldberg, Horacio, Mario Magallón Anaya, Isaías Palacios Contreras, María del Rayo Ramírez Fierro, and Sandra Escutia Díaz. Toluca, Mexico: Universidad Autónoma del Estado de México, 2000.

Diccionario de filosofía oriental. Wolpin, Samuel. Buenos Aires: Kier, 1993.

Diccionario de filosofía: seguido de un cuadro histórico de las escuelas de filosofía. Jolivet, Régis. Buenos Aires: Club de lectores, 1978.

Diccionario de filosofía y sociología marxista. Iudin, Pavel, ed. Buenos Aires: Editorial Seneca, 1965.

Diccionario de filósofos. Centro de Estudios Filosóficos de Gallarte. Madrid: Ediciones Rioduero, 1986.

Diccionario de grandes filósofos, 2: (K–Z). Ferrater Mora, José. Madrid: Alianza, 2002.

Diccionario de las mil obras clave del pensamiento. Huisman, Denis. Madrid: Tecnos, 1997.

Diccionario del lenguaje filosófico. Foulquié, Paul. Barcelona: Editorial Labor, 1967.

Diccionario del lenguaje filosófico, dirigido. Foulquié, Paul. Barcelona: Editorial Labor, 1967.

Diccionario de pensamiento contemporáneo. Moreno Villa, Mariano. Madrid: San Pablo, 1997.

Diccionario de términos filosóficos. Robert, François, and José Manuel Revuelta. Madrid: Acento, 1994.

Diccionario filosófico. Ezcurdia Hijar, Agustín, and Pedro Chávez Calderón. Mexico: Limusa: Noriega Editores, 1994.

Diccionario filosófico. Hernández, Pablo María. [Dominican Republic: Universidad Autónoma de Santo Domingo], 1982.

Diccionario filosófico. 2nd ed. Hernández, Pablo María. [Dominican Republic: Universidad Autónoma de Santo Domingo], 1984.

Diccionario filosófico. Rozental', Mark Moiseevich, and Pavel Fedorovich Iudin, eds. Montevideo: Ediciones Pueblos Unidos, 1965.

Diccionario filosófico. Rozental', Mark Moiseevich, and Pavel Fedorovich Iudin, eds. [Argentina]: Ediciones Universo, 1973.

Diccionario filosófico. Rozental', Mark Moiseevich. [Havana]: Edición Revolucionaria, 1984.

Diccionario filosófico. Savater, Fernando. Barcelona: Planeta, 1995.

Diccionario filosófico. 2nd ed. Savater, Fernando. Barcelona: Planeta, 1999.

Diccionario filosófico abreviado. Rev. ed. Rozental', Mark Moiseevich, and Pavel Fedorovich Iudin. Mexico: Ediciones Quinto Sol, [1970–1979?].

Diccionario filosófico: manual de materialismo filosófico: una introducción analítica. García Sierra, Pelayo, and Gustavo Bueno. Oviedo: Fundación Gustavo Bueno, 2000.

Diccionario marxista de filosofía. Blauberg, Igor' Viktorovich, ed. Mexico: Ediciones de Cultura Popular, 1979.

Enciclopedia concisa de filosofía y filósofos. Urmson, J. O. Madrid: Cátedra, 1979.

Enciclopedia concisa de filosofía y filósofos. 2nd ed. Urmson, J. O. Madrid: Cátedra, 1982.

Enciclopedia concisa de filosofía y filósofos. 3rd ed. Urmson, J. O. Madrid: Cátedra, 1994.

Enciclopedia de la filosofía. Barcelona: Ediciones B, 1992.

La filosofía: las ideas, las obras, los hombres. Noiray, André, ed. Bilbao: Ediciones Mensajero, 1974.

Filosofía y religión. Madrid, 1991.

Los filósofos. Aubral, François. Madrid: Acento Editorial, 1994.

Grandes profecías. Madrid: Nueva Lente, 1986.

Invitación a la fenomenología. Husserl, Edmund. Barcelona: Paidós, 1992.

Léxico filosófico. Millán-Puelles, Antonio. Madrid: Rialp, 1984.

Nueva enciclopedia temática Planeta: arte y filosofá. Gutiérrez Ducóns, Juan Luis, and Javier Sánchez Almazán. Barcelona: Planeta, 1993.

Pequeña filosofía para no filósofos. Jacquard, Albert, and Huguette Planès. Mexico, D.F.: Debolsillo/Random House Mondadori, 2004.

Vocabulario técnico y crítico de la filosofía. 2nd ed. Lalande, André. Buenos Aires: El Ateneo, 1967.

UKRAINIAN

Anhlo-ukraïns'kyi filosofs'kyi slovnyk. Polishchuk, Nina Pavlovna, and Vasyl Semenovych Lisovy. Kiev: Lybid', 1996.

Filosofs'ka dumka v Ukraïni: biobibliohrafichnyi slovnyk. Tkachuk, Maryna, and Vilen Serhiiovych Hors'kyi. Kiev: Pul'sary, 2002.

Filosofs'kyi entsyklopedychnyi slovnyk. Shinkaruk, Vladimir Illarionovich. Kiev: Abrys, 2002.

Filosofs'kyi slovnyk. Shinkaruk, Vladimir Illarionovich. Kiev: Gol. Red. Ukraïns'koï Rad. Entsyklopediï, 1973.

Filosofs'kyi slovnyk. 2nd ed. Shinkaruk, Vladimir Illarionovich. Kiev: Gol. Red. Ukraïns'koï Rad. Entsyklopediï, 1986.

Michael J. Farmer (2005)

PHILOSOPHY JOURNALS

The learned journal was one of the major cultural innovations of the seventeenth century. Of the pioneering scholarly journals inaugurated during that century, the earliest one that regularly presented philosophical material is, remarkably, still being published, but, unhappily for our story, it now deals mainly with philology and related fields. This patriarch of professional periodicals, 300 years old and still lively, is the *Journal des savants* (Paris, 1665–), issued quarterly, with variations. The title was *Journal des sçavans* from 1665 to 1792. Publication was suspended from 1797 to 1816. The journal was devoted originally to book reviews, bibliographies, and news notes on philosophy, science, and literature.

In the same century similar learned journals, commenting on new books in philosophy and other fields, were issued for various periods in a number of cultural centers besides Paris. Indeed, on October 22, 1668, Leibniz wrote to Emperor Leopold I, taking note of the fact that the rival French nation had inaugurated the *Journal des sçavans* and declaring that Germany needed a similar medium of intellectual communication; Leibniz asked for a license to issue such a periodical, and the issuance of *Acta Eruditorum* beginning some 14 years later (see below) may have been the result. Prominent among the early learned journals issued outside of France which covered philosophy among other subjects were:

1668–1690. *Giornale de'letterati* (Parma), monthly. Suspended 1679–1686. Periodicals with the same title were also published in other Italian cities, including Rome, for various periods.

1681–1683. *Weekly Memorials for the Ingenious* (London), weekly.

1684–1718. *Nouvelles de la république des lettres* (Amsterdam), originally issued monthly, later issued six times a year. Founded by Pierre Bayle during his exile from France and edited by him from 1684 to 1687. Suspended 1689–1698 and 1711–1715.

1688–1690. *Freymüthige lustige und ernsthaffte Monats-Gespräche* (Halle), monthly.

In Latin, philosophical and other material appeared in *Acta Eruditorum* (Leipzig, 1682–1776), issued monthly. The title was *Nova Acta Eruditorum* from 1732 to 1776. This periodical, founded by Otto Mencke, Leibniz' friend, contained many contributions by and about Leibniz, and some authorities even say it was founded by Leibniz; he probably at least had a hand in Mencke's establishment of it. In this vehicle Leibniz first gave the world his notions respecting the differential calculus, and in it raged the controversy, beginning in 1699, over whether Leibniz or Newton first discovered the principles of the calculus.

Beginning in the eighteenth century a number of learned journals were devoted exclusively or largely to philosophy. These were the earliest instances of philosophical journals in a strict sense.

DEFINITION OF A PHILOSOPHICAL JOURNAL. A philosophical journal, for the purposes of this article, is a publication that fulfills the following criteria:

(1) It is devoted to the whole field of philosophy (and nothing else) or, more narrowly, to a part of the field of philosophy (for example, symbolic logic or Thomism) or, more broadly, to the whole field of philosophy plus one or two other fields of interest (philosophy and psychology, philosophy and theology, and so forth). The specification "part of the field of philosophy" is taken strictly, thereby excluding philosophy of education and pure theosophy, but it includes political and social philosophy. Magazines of popular philosophy or popular morals—such as Addison and Steele's *Spectator* (founded 1710), *Der Leipziger Diogenes* (founded 1723), and *Der Dresdnische Philosoph* (founded 1737)—are excluded. Student journals, such as the *Graduate Review of Philosophy* (Minneapolis), are also excluded. One theosophical journal, however, *The Aryan Path*, which contains many strictly philosophical articles, is included here.

(2) It is issued at stipulated intervals of less than a year. Thus, the intent reader will notice that this account

does not mention bibliographical yearbooks, annual collections of studies, annual proceedings of philosophical societies, or irregular collections of articles (such as the quasi journal *Polemic*, of which eight issues were published at London at irregular intervals, from 1945 to 1947). In a few cases, however, a publication which, although it had no stipulated frequency, was actually issued (say) four times a year for a period of years, is counted as a regular journal.

(3) It has survived longer than a year. This requirement leads to the exclusion of, for example, *Symposion; Philosophische Zeitschrift für Forschung und Aussprache* (Erlangen), edited by Ernst Cassirer, Hans Driesch, and others, since only four issues were published, in 1926. In the case of periodicals inaugurated just prior to the completion of this article, however, the requirement of more than a year's duration is relaxed.

Strict adherence to the second criterion listed above has led to the exclusion of at least one vitally important medium of philosophical discussion—*Proceedings of the Aristotelian Society* (London), issued annually (referred to in this Encyclopedia as *PAS*). Since its founding in 1891 this periodical has presented numerous important articles, including papers by Bertrand Russell, Gilbert Ryle, J. L. Austin, Ludwig Wittgenstein, G. E. Moore, and many others.

In our account of significant journals devoted to the whole of philosophy, we have attempted to be comprehensive. It is probable, however, that a considerable number of obscure, though worthy, journals have slipped through the net and that on these strict criteria a few borderline semiprofessional journals are omitted which may have merited inclusion. Therefore, the statistics offered from time to time in this article (for example, that so many journals originated in a certain period) are based on its author's particular standards and should be taken as approximate. For fairly complete details about philosophical journals devoted to the whole of philosophy which are still being issued, the reader is referred to the *International Directory of Philosophy and Philosophers,* edited by Gilbert Varet and Paul Kurtz (New York, 1966).

For journals devoted to a part of philosophy or to philosophy and other disciplines, the present list is definitely incomplete. For example, there have been over a hundred theological journals, past and present, but only some fifty are mentioned as outstanding examples.

Regarding each journal mentioned in this article, four facts are ordinarily presented: the year it began publication (and, if it is no longer issued, its last year of publication), its title, its place of publication, and the frequency of its issuance. Other facts, such as historic figures who were editors, periods of suspension, and changes of title, are sometimes noted, but these additions are illustrative rather than complete.

In the case of journals that have changed their titles, the latest title is usually given as the main entry, and earlier titles are noted. Where one journal has succeeded another with some definite contact or relationship between them, they are considered as a single journal with a changed title. Thus, *Ratio* is considered as continuous with its earlier incarnation and appears below as one of the two oldest living philosophical journals.

In several instances some outstanding articles that the periodicals have published or striking facts about their influence in the philosophical world are set forth. Other periodicals, such as *Philosophy* (1926–), *Voprosy Filosofii* (1947–), and some German and Italian periodicals, could also have been appropriately singled out for such an exposition, had space permitted.

STATISTICAL CONSPECTUS. From the eighteenth century to the present, approximately 70 philosophical journals have been born with more or less fanfare, have survived for a period, and have given up the ghost. About 180 others, however, are still alive, some flourishing, some bravely keeping their heads above water, some pitifully gasping for breath.

Of the philosophical journals published today, two are more than a century old, and four others are over 90 years old; their average life span, however, is about 28 years. Of those which no longer exist, the longest-lived at the time of its death was 81 years old: the *Zeitschrift für Philosophie und philosophische Kritik*, founded by the younger Fichte in 1837 and discontinued in 1918.

More births of philosophical journals occurred in the 1950s (55) and the 1940s (49) than in any other decades, but likewise more deaths of philosophical journals occurred in the 1950s (19) and the 1940s (17) than in any other decades. During World War I about 12 per cent of the philosophical journals in existence in 1914 ended their lives; during World War II about 15 per cent of those in existence in 1939 were terminated.

EARLY JOURNALS

EARLY QUASI JOURNALS. Publications devoted to philosophy that were intended to be issued from time to time (more than once a year) but not at uniform intervals may be denominated "quasi journals of philosophy" since they

do not conform to the requirement of a set frequency of issuance. A number of these quasi journals came into being from about 1715 on, especially in Germany, and lasted for varying periods. The following may serve as examples:

1715–1726. *Acta Philosophorum* (Halle). In German. Covered books on the history of philosophy. Probably the earliest quasi journal of philosophy.

1741–1744. *Philosophische Büchersaal* (Leipzig). Eight issues were published.

1789–1790. *Neues philosophisches Magazin; Erläuterungen und Anwendungen des Kantischen Systems bestimmt* (Leipzig). Two volumes, each with four issues, appeared.

1790–1850. Of seven genuine philosophical journals that saw the light before 1850, two—the *Theologische Quartalschrift* and *Ratio*—still survive, but the continuity of *Ratio* with its origin is tenuous. Chronologically, the seven pre-1850 journals fall into two groups. Those in the first group are:

1794–1807. *Revue philosophique, littéraire et politique* (Paris), issued three times a month, with variations. Title was *Décade philosophique, littéraire et politique* from 1794 to 1804 and became *Revue, ou Décade philosophique* late in 1804. Merged in 1808 with the *Mercure de France* (Paris, 1672–1820) and at that point may be considered to have lost its standing as a philosophical journal.

1795–1800. *Philosophisches Journal einer Gesellschaft teutscher Gelehrten* (Neustrelitz, 1795–1796; Jena and Leipzig, 1797–1800), monthly. J. G. Fichte was coeditor from 1797 to 1800. An article that Fichte published in the *Journal* in 1798 regarding the grounds of our belief in a divine government of the universe (defining God as the moral order of the universe) caused a cry of atheism to be raised and led to the suppression of the *Journal* in all the German states except Prussia, as well as to Fichte's resignation in 1799 from his teaching position at the University of Jena (see Atheismusstreit).

1802–1803. *Kritisches Journal der Philosophie* (Tübingen), issued five times in 1802 and once in 1803. Editors, F. W. J. von Schelling and G. W. F. Hegel. Included a number of articles by Hegel.

1819–. *Theologische Quartalschrift* (Tübingen; later Ravensburg; now Stuttgart), quarterly. Suspended 1945.

The following journals belong to the second pre-1850 group:

1832–1852. *Zeitschrift für Philosophie und katholische Theologie* (Cologne, 1832–1836; Coblenz, 1836–1839; Cologne, 1840–1841; Bonn, 1842–1852), quarterly, with variations.

1837–1918. *Zeitschrift für Philosophie und philosophische Kritik; Vormals Fichte-Ulricische Zeitschrift* (Bonn, 1837–1842; Tübingen, 1843–1846; Halle, 1847–1890; Leipzig, 1891–1918), quarterly, with variations. Title was *Zeitschrift für Philosophie und spekulative Theologie*, from 1837 to 1846; subtitle varied. Founded by I. H. von Fichte (son of J. G. Fichte); later edited by him and Hermann Ulrici. Supported Christian and Hegelian views.

1847–. *Ratio* (Oxford and Frankfurt; formerly Göttingen), semiannual, with variations. Title was *Abhandlungen der Fries'schen Schule* from 1847 to 1936. Suspended 1850–1903, 1915–1917, 1919–1928, and 1937–1956. Now issued in English and German editions.

1850–1900. Of the decades from 1850 to 1899, the first produced 1 new journal of philosophy, the second and third a total of 11, and the fourth and fifth a total of 19. The lone philosophical journal born in the 1850s was *La Revue philosophique et religieuse* (Paris, 1850–1858), issued monthly.

The 1860s. In the 1860s seven philosophical journals were begun—three in Germany and one each in Belgium, France, Switzerland, and the United States. The first four journals to appear in this decade, including the first English-language journal of philosophy, are now defunct.

1861–1914. *Zeitschrift für Philosophie und Pädagogik* (Leipzig; later Langensalza), monthly, with variations. Title was *Zeitschrift für exakte Philosophie im Sinne des neuern philosophischen Realismus* (and the journal was, for the most part, a quarterly) from 1861 to 1896; suspended 1876–1882. Merged in 1896 with the *Zeitschrift für Philosophie und Pädagogik*, which had been issued since 1894; the combined publication took the title of the latter *Zeitschrift*.

1862–1864. *Athenäum; Philosophische Zeitschrift* (Munich), quarterly.

1867–1913. *L'Année philosophique* (Paris), issued annually and therefore not a "periodical" in the required sense, from 1867 to 1869 and again from 1890 to 1913, but it was a weekly (with variations)

from 1872 to 1885 and a monthly from 1885 to 1889. Title was *La Critique philosophique* when the publication was issued weekly or monthly, from 1872 to 1889. C. B. Renouvier was a coeditor from 1890 to 1900.

1867–1893. *Journal of Speculative Philosophy* (St. Louis, Mo., 1867–1880; New York, 1880–1893), quarterly, with variations. Apparently the first philosophical journal in the English language. Founded by William T. Harris. Organ of the St. Louis Philosophical Society. Served as the vehicle for the first published writings of James, Royce, and Dewey. Its motto was "Philosophy can bake no bread, but she can procure for us God, freedom, and immortality."

The three philosophical journals of this period which have survived are German, Swiss, and Belgian, respectively:

1868–. *Archiv für Geschichte der Philosophie* (Berlin; previously Leipzig and Heidelberg), quarterly, with variations. Title was *Philosophische Monatshefte* from 1868 to 1887, *Archiv für Geschichte der Philosophie* from 1888 to 1894, *Archiv für Philosophie* from 1895 to 1926 (in this period the periodical was issued in two parts, *Archiv für Geschichte der Philosophie* and *Archiv für systematische Philosophie*), title was *Archiv für Philosophie und Soziologie* from 1927 to 1930 (again issued in two parts, *Archiv für Geschichte der Philosophie und Soziologie* and *Archiv für systematische Philosophie und Soziologie*), title reverted to *Archiv für Geschichte der Philosophie* in 1931; suspended 1933–1959. Original editor was Ludwig Stein, with the collaboration of Hermann Diels, Wilhelm Dilthey, Benno Erdmann, and Eduard Zeller. Paul Natorp became coeditor of the combined publication in 1895 and editor of the systematic part. Has contained articles in English, French, German, and Italian since 1895.

1868–. *Revue de théologie et de philosophie* (Lausanne), issued six times a year from 1868 to 1920, quarterly since 1921. Title has varied.

1869–. *Nouvelle Revue théologique* (Louvain, Belgium), monthly. A Jesuit organ.

THE 1870S. The 1870s are remembered as the decade which produced the *Revue philosophique de la France et de l'étranger and Mind.* However, two other journals were inaugurated during the decade:

1870–. *Rivista di filosofia* (Turin; previously Bologna-Modena, Florence, Forli, Genoa, Pavia, Rome, and Milan), quarterly, with variations. Title was *La filosofia delle scuole italiane* (Florence; then Rome), 1870–1885; title was *Rivista italiana di filosofia* (Rome) from 1886 to 1898; became two separate periodicals, *Rivista di filosofia e scienze affini* (Bologna), 1899–1908, and *Rivista filosofica* (Pavia), 1899–1908; combined under the title *Rivista di filosofia* in 1909. Suspended 1922. In 1963 it absorbed *Il pensiero critico,* a quarterly published at Milan since 1950.

1877–1916. *Vierteljahrsschrift für wissenschaftliche Philosophie und Soziologie* (Leipzig), quarterly. Title was *Vierteljahrsschrift für wissenschaftliche Philosophie* from 1877 to 1901. Coeditors at various times included Richard Avenarius, Ernst Mach, and Wilhelm Wundt.

"Revue Philosophique." The *Revue philosophique de la France et de l'étranger* (Paris, 1876–) was originally issued monthly and later issued six times a year; it is now issued quarterly. The *Revue*'s first editor, Théodule Ribot, served for 40 years, until his death in 1916. Under his direction the *Revue* gave primary emphasis to articles on psychology. Philosophy began to gain predominance under Ribot's successor, Lucien Lévy-Bruhl, who was also a long-lived editor, conducting the periodical for 23 years until his death in 1939. The editors in succeeding years, when philosophy was fully established as the main arena of discussion in the *Revue*, were Émile Bréhier and Paul Masson-Oursel (1940–1952), Masson-Oursel and Pierre-Maxime Schuhl (1952–1956), and Schuhl alone (since 1956).

Even now the *Revue*'s strongest contribution is represented not so much by publication of original hypotheses as by careful analysis and criticism of old and new viewpoints. Useful articles, for example, have been published on Leibniz, Hume, and English linguistic philosophy. An entire issue was devoted to Lévy-Bruhl in 1957 on the one hundredth anniversary of his birth. The coverage of philosophy "de l'étranger" has consisted in large part of some translations from English and German, extensive critical reviews of books, and summaries of periodical articles.

Among the more original contributions in the *Revue* have been C. S. Peirce's "La Logique de la science" (1878–1879), Étienne Gilson's "Essai sur la vie intérieure" (1920), Raymond Ruyer's "Ce qui est vivant et ce qui est mort dans la matérialisme" (1933), and Georges Gurvich's "Le Problème de la sociologie de la connaissance" (1957–1958). Famous contributors have included Rudolf

Hermann Lotze, Herbert Spencer, J. S. Mill, Wilhelm Wundt, Henri Bergson, and Georges Sorel.

"Mind." *Mind; A Quarterly Review of Psychology and Philosophy* (originally London, later Edinburgh, now Oxford, 1876–), is issued quarterly.

In 1874 Alexander Bain broached the idea of establishing the first British philosophical journal to his pupil George C. Robertson, who suggested the title *Mind.* Bain appointed Robertson editor and supported the journal financially, sinking almost £3,000 into it in 15 years, until Robertson resigned in 1891.

Robertson, on laying down his mantle as editor, lamented that the journal had attracted more attention from "the lay student" than from those "whose regular business is with Philosophy." G. F. Stout, when he succeeded Robertson in 1892, wrote that "what is of prime importance is that our pages shall be filled with genuine work to the exclusion of merely dilettante productions." The implication here is curious when one considers that among the contributors to *Mind* during Robertson's stewardship were philosophers of the caliber of Samuel Alexander, A. W. Benn, Bernard Bosanquet, F. H. Bradley, T. H. Green, William James, C. Lloyd Morgan, Hastings Rashdall, Josiah Royce, Henry Sidgwick, and John Venn.

Sidgwick, who succeeded Bain as the financial "angel" of *Mind*, died in 1900. It was then that, pursuant to a suggestion made by Sidgwick in 1899, the Mind Association was formed (with Edward Caird as the first president) to support the journal. Meanwhile, G. E. Moore and Bertrand Russell had published their earliest contributions in *Mind* in the 1890s, and the periodical was well on its way to becoming what it is now, one of the dozen most influential journals of philosophy in the world. Stout relinquished the editorship in 1920; his successors were Moore, 1921–1947, and Gilbert Ryle since 1948.

Over the decades *Mind* has published many highly influential articles, such as Moore's "The Refutation of Idealism" (1903), Russell's "On Denoting" (1905), and H. A. Prichard's "Does Moral Philosophy Rest on a Mistake?" (1912). During Moore's editorship the journal set a particularly high standard, publishing such papers as W. T. Stace's "The Refutation of Realism" (1934), A. J. Ayer's "Demonstration of the Impossibility of Metaphysics" (1934), C. L. Stevenson's "Persuasive Definitions" (1938), Norman Malcolm's "Are Necessary Propositions Really Verbal?" (1940), John Wisdom's eight articles entitled "Other Minds" (1940–1943), and Frederick Will's "Will the Future Be Like the Past?" (1947). More recently (under Ryle's editorship), *Mind* has presented such important articles as J. N. Findlay's "Can God's Existence Be Disproved?" (1948), R. M. Hare's "Imperative Sentences" (1949), Paul Edwards' "Bertrand Russell's Doubts About Induction" (1949), P. F. Strawson's "On Referring" (1950), A. J. Ayer's "Individuals" (1952), G. E. M. Anscombe's "Aristotle and the Sea Battle" (1956), Nelson Goodman's "About" (1961), and many papers—written by Wittgenstein's disciples—that helped to establish Wittgenstein's reputation before the posthumous publication of his books.

A public controversy occurred when *Mind* declined to publish a review of Ernest Gellner's *Words and Things* (London, 1959), which was critical of the ordinary-language school. Bertrand Russell, in a letter to the London *Times*, on November 5, 1959, protested against *Mind*'s decision.

Ryle's policy as editor has been to give some preference to philosophers who have not previously appeared in print. This policy, while testifying to the kindness of the editor and his concern for providing needed encouragement to tomorrow's leading spirits, has made it difficult to maintain the Olympian level of quality to which readers became accustomed during Moore's period as editor.

THE 1880S. Of the nine journals of philosophy generated in the 1880s, five are still functioning, *The Monist* being perhaps the best known. One of the nine, among the oldest Italian philosophical journals, is *Divus Thomas*; another, a Swiss product, has had *Divus Thomas* as its subtitle or (for a time) as its main title. Japan and Russia gave birth to journals of philosophy in this decade, and the Japanese entry is still in the field.

American and British. Paul Carus was associated with two of the three English-language journals begun in this decade.

1886–1915. *Review of Theology and Philosophy* (Edinburgh), quarterly. Title was *Theological Review and Free Church College Quarterly* from 1886 to 1890 and *Critical Review of Theological and Philosophical Literature* from 1890 to 1904.

1887–1936. *The Open Court* (Chicago), issued every other week from 1887 to 1888; weekly from 1888 to 1896; monthly from 1897 to 1933; quarterly from 1934 to 1936. Successor to *The Index* (published weekly at Toledo, Ohio, and later at Boston, 1870–1886), organ of the Free Religious Association. *The Open Court* was founded by Paul Carus. Devoted to the establishment of ethics and religion on a sci-

entific basis, it was more clearly a philosophical journal than was *The Index.*

1888–. *The Monist; An International Journal of General Philosophical Inquiry* (La Salle, Ill.; previously Chicago), quarterly. Suspended 1937–1963. International editorial board. Each issue now devoted to a specific topic. Edited by Paul Carus from 1888 to his death in 1919. Contributors have included Peirce, Dewey, Bosanquet, and Russell.

European. Two of the journals first issued in the 1880s are products of Italy and two are German-language publications.

1880–. *Divus Thomas; Commentarium de Philosophia et Theologia* (Piacenza), issued six times a year. Articles in English, French, Italian, and Latin. Suspended 1906–1923.

1881–1900. *Rivista speciale di opere di filosofia scientifica* (Milan), monthly. Title was *Rivista di filosofia scientifica* from 1881 to 1891.

1886–. *Freiburger Zeitschrift für Philosophie und Theologie* (Fribourg, Switzerland), quarterly, with variations. Title was *Jahrbuch für Philosophie und spekulative Theologie* (Paderborn; later Vienna) from 1886 to 1922, with the subtitle *Divus Thomas* from 1914 to 1922; title was *Divus Thomas* from 1923 to 1953.

1888–. *Philosophisches Jahrbuch* (Munich; previously Fulda), semiannual, with variations. Title has varied. Catholic-oriented.

Japanese and Russian. The Japanese journal begun in the 1880s is still being issued. The Russian journal, *Voprosy Filosofii i Psikhologii,* died in 1917, but a new *Voprosy Filosofii*, as will be noted later, arose from its ashes 30 years later, in 1947.

1887–. *Tetsugaku Zasshi; Journal of Philosophy* (Tokyo), quarterly, with variations. Journal of the Philosophical Society of Tokyo University. Titles of articles in English.

1889–1917. *Voprosy Filosofii i Psikhologii* (Moscow), issued six times a year by the Moskovskoe Psikhologicheskoe Obshchestvo.

THE 1890S. The 1890s constituted the fertile decade of *Ethics* and *The Philosophical Review*, of the *Revue de métaphysique et de morale*, of the *Revue philosophique de Louvain*, and of *Kant-Studien*, all of which are on the scene today, plus the oldest Indian and Polish philosophical journals, which also continue to appear.

Louvain was the parent of a pair of French-language journals, both flourishing today:

1894–. *Revue philosophique de Louvain* (Louvain), quarterly. Founded by Cardinal Mercier. Published by the Société Philosophique de Louvain. Neo-Scholastic. Suspended 1915–1918 and 1941–1944. Title was *Revue néo-scolastique* from 1894 to 1910, *Revue néo-scolastique de philosophie* from 1910 to 1933, and *Revue néoscolastique de philosophie* from 1934 to 1945. A *Répertoire bibliographique* has been published as an adjunct of the *Revue* since 1895; since 1938 it has been published separately, and since 1949 it has been administratively separate. Some articles in English; others in French with English summaries.

1895–. *Répertoire bibliographique de la philosophie* (Louvain), quarterly. Title was *Sommaire idéologique des ouvrages et des revues de philosophie* (with variations) from 1895 to 1914. Suspended 1915–1933 and 1941–1945. Reproduced *in toto*, with Dutch headings, in the *Tijdschrift voor Filosofie* (Louvain, 1939–).

Two of the journals that started publication in the 1890s have ethics as their subject matter, ethics alone in one case and ethics plus metaphysics in the other:

1890–. *Ethics; An International Journal of Social, Political and Legal Philosophy* (Chicago), quarterly. Established, under the title *The International Journal of Ethics,* as an outgrowth of *The Ethical Record*, organ of the Ethical Societies; responsibility assumed by the University of Chicago in 1923; name changed to *Ethics* in 1938.

"Revue de métaphysique et de morale." The *Revue de métaphysique et de morale* (Paris, 1893–) has been issued quarterly since 1920 (previously issued six times a year). This *Revue*, now the principal French philosophical journal, was established by Xavier Léon, with the collaboration of Élie Halévy. The title of the publication reflected not only a reaction against positivism but also, affirmatively, a belief that the conclusions of speculative philosophy could have a practical value. Léon (who also founded the Société Française de Philosophie in 1901 and organized various international congresses of philosophy) served as editor of the *Revue* until his death in 1935, when he was succeeded by Dominique Parodi. Parodi died in 1955, and Jean Wahl (who had assisted Parodi on the *Revue*) took over.

Until World War II special numbers of the *Revue* were occasionally devoted to a single topic. For example,

issues were devoted to Kant (1904, the centennial of his death), Rousseau (1912, the bicentennial of his birth), American philosophy (1922, with articles by John Dewey, W. E. Hocking, C. I. Lewis, R. B. Perry, and others), Pascal (1923, the tercentenary of his birth), Hegel (1931, the centennial of his death), and Descartes (1937, the tercentenary of the *Discourse on Method*).

The journal's contributors have included all French philosophers of note as well as many eminent foreigners, such as Bertrand Russell, A. N. Whitehead, and Bernard Bosanquet; Benedetto Croce and Giovanni Gentile; Miguel de Unamuno; and Edmund Husserl. Among the articles of more than ordinary interest which have appeared in the journal are Henri Poincaré's "La Logique de l'infini" (1909), Henri Bergson's "L'Intuition philosophique" (1911), Étienne Gilson's "Art et métaphysique" (1916), Gabriel Marcel's "Existence et objectivité" (1925), Léon Brunschvicg's "Religion et philosophie" (1935), José Ferrater Mora's "Philosophie et architecture" (1955), and Wahl's "Physique atomique et connaissance humaine" (1962).

"The Philosophical Review." Two of the major philosophical journals now on the scene had their origin in the *fin de siècle* decade. One of these was American, *The Philosophical Review* (Ithaca, N.Y., 1892–), which is now issued quarterly. It was previously issued six times a year. It is published by the Sage School of Philosophy at Cornell University.

The Philosophical Review was relatively undistinguished until the late 1940s. Among the few important articles that preceded the late flowering of the journal were C. I. Lewis' "Experience and Meaning" (1934) and Moritz Schlick's reply, "Meaning and Verification" (1936), which is commonly regarded as Schlick's most telling contribution to contemporary philosophy.

The *Review*'s recent burgeoning took place under the guidance of Max Black and his colleagues on the philosophy staff at Cornell. Significant contributions in this period include W. V. Quine's "Two Dogmas of Empiricism" (1951), Black's "Definition, Presupposition, and Assertion" (1952), Gilbert Ryle's "Ordinary Language" (1953), H. P. Grice and P. F. Strawson's "In Defense of a Dogma" (1956), Stuart Hampshire's "On Referring and Intending" (1956), various articles by John Rawls on justice (1958–1963), J. J. C. Smart's "Sensations and Brain Processes" (1959), Norman Malcolm's "Anselm's Ontological Arguments" (1960), and Richard Taylor's "Fatalism" (1962).

Other prominent philosophers whose work has appeared in *The Philosophical Review* include George Boas, R. M. Chisholm, P. T. Geach, Nelson Goodman, Arthur E. Murphy, Ernest Nagel, and J. A. Passmore.

"Kant-Studien." The other famous journal founded in the last decade of the nineteenth century was German: *Kant-Studien; philosophische Zeitschrift* (Hamburg and Leipzig; later Berlin; now Cologne, 1896– —), issued quarterly, with variations. The title was originally spelled *Kantstudien*. This journal has been the organ of the Kant-Gesellschaft since 1904. It was suspended twice, from 1937 to 1942 and from 1945 through 1953.

Hans Vaihinger founded *Kant-Studien* and was its editor, alone or with one or two coeditors, until 1922; Max Scheler was coeditor in 1902/1903. The periodical has published articles in English, French, German, and Italian by outstanding scholars and philosophers, including Erich Adickes, Émile Boutroux, Edward Caird, Ernst Cassirer, Rudolf Eucken, and Norman Kemp Smith. Although the periodical is specifically oriented toward Kant, it includes in its purview current thought on questions raised by Kant, pre-Kantian philosophy as part of the background of Kantianism, and other liberal extensions of the frame of reference.

For many years *Kant-Studien* published abstracts of new books in philosophy on a unique basis: the abstracts were written by the authors themselves. As another special feature, 86 separate monographs have been issued under the auspices of the journal.

During the short-lived revival of *Kant-Studien* in 1942–1944, the journal was not uninfluenced by Nazism. In 1954, however, on the 150th anniversary of Kant's death, the periodical was reborn in a new setting (Cologne), and since then it has regained its international reputation.

Other journals. Additional journals which arose in the 1890s were the following:

1893–. *Revue thomiste; Revue doctrinale de théologie et de philosophie* (originally Brussels; now Toulouse), quarterly. Founded by the Dominican order.

1895–. *The Vedanta Kesari* (Madras), monthly, with variations. Title was *The Brahmavadin; A Fortnightly Religious and Philosophical Journal* from 1895 to 1914. Organ of the world-wide Ramakrishna order.

1896–1915. *Neue metaphysische Rundschau* (Berlin), monthly. Title was *Metaphysische Rundschau* from 1896 to 1897.

1897–1949. *Przegląd Filozoficzny* (Warsaw), quarterly. English summaries of articles in some issues; table of contents also printed in French. Suspended 1940–1946. Issuance stopped by the government at the beginning of the period of militant Marxist domination of Polish philosophy (first half of the 1950s). Replaced (not succeeded) by the periodical now called *Studia Filozoficzne; Kwartalnik* (1951– —).

PREWAR PERIOD

The vital statistics for the period from 1900 to 1914 show 19 journals born, 12 of which have survived. This is the period of *The Hibbert Journal, The Journal of Philosophy, The Harvard Theological Review*, and the first of various journals called *Logos*. Czechoslovakia, Ireland, the Netherlands, and Spain are represented for the first time in this period.

THEOLOGY. The prewar period was exceptionally rich in journals which emphasized theology or the philosophy of religion:

1900–1939. *Revue de philosophie* (Paris), issued six times a year. Thomist-oriented.

1902–. *The Hibbert Journal; A Quarterly Review of Religion, Theology and Philosophy* (London), quarterly; originally issued monthly. Treats religious and humanistic questions from a philosophical or cultural point of view. Contributors have included Henri Bergson, Sarvepalli Radhakrishnan, Bertrand Russell, Rabindranath Tagore, and Leo Tolstoy.

1905–1910. *Rivista storico critica delle scienze teologiche* (Rome), monthly.

1907–. *Revue des sciences philosophiques et théologiques* (Étiolles, Soisy-sur-Seine), quarterly, with variations. Founded by the French Dominicans of the Facultés de Philosophie et de Théologie du Saulchoir. Suspended 1915–1919 and 1943–1946.

1908–. *The Harvard Theological Review* (Cambridge, Mass.), quarterly.

1909–. *Rivista di filosofia neo-scolastica* (Milan), issued six times a year, with variations. Organ of the Istituto de Filosofia, Università Cattolica del Sacro Cuore.

1910–. *La ciencia tomista* (originally Madrid; now Salamanca), issued six times a year from 1910 to 1949, quarterly since 1950. Edited by the Spanish Dominicans.

SOCIAL PHILOSOPHY AND AESTHETICS. Journals of ethics, social philosophy, philosophy of culture, and aesthetics were fostered in the prewar period in Germany and Italy:

1906–1926. *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft* (Stuttgart), quarterly. Max Dessoir, editor.

1907–. *Archiv für Rechts- und Sozialphilosophie* (Neuwied; previously Munich), quarterly. Title has varied. Contains articles mainly in German and English.

1906–. *Rivista rosminiana di filosofia e di cultura* (Milan; formerly published at Pallanza and other Italian cities), quarterly, with variations. Edited by Giuseppe Morando from 1906 to his death in 1914; edited by his son Dante since 1937. Combats positivism and subjectivism.

1910–1941. *Zeitschrift für deutsche Kulturphilosophie; Neue Folge des Logos* (Tübingen), issued three times a year. Title was *Logos; Internationale Zeitschrift für Philosophie der Kultur* from 1910 to 1933. In 1934, when the journal was completely Nazified, Richard Kroner was replaced as editor in chief (a post which he had held since 1910) and Ernst Cassirer, Edmund Husserl, Friedrich Meinecke, and Rudolf Otto were summarily removed from the roll of collaborating editors.

GENERAL PHILOSOPHY. Several of the prewar journals were general in their philosophical coverage. Two were Italian:

1903–1951. *Quaderni della critica* (Naples), issued six times a year, with variations. Founded by Benedetto Croce. Title was *La critica; Rivista di letteratura, storia e filosofia* from 1903 to 1944. Contained many articles by Croce and by Gentile.

1908–1925. *Bollettino filosofico; Organo della Biblioteca Filosofica di Firenze* (Florence), monthly, with variations. Suspended 1910, 1913–1915, and 1917–1923.

France and the Netherlands gave rise to two others:

1900–. *Bulletin de la Société Française de Philosophie* (Paris), quarterly. Contributors have included Bergson, Louis de Broglie, Brunschvicg, Croce, Einstein, and Russell.

1907–. *Algemeen Nederlands Tijdschrift voor Wijsbegeerte en Psychologie* (Assen; formerly Amsterdam), issued five times a year. Organ of the

Algemene Nederlandse Vereniging voor Wijsbegeerte. Title was *Tijdschrift voor Wijsbegeerte* from 1907 to 1934. From 1934 to 1938 this periodical included, as a separate section, *Annalen der Critische Philosophie* (Assen, 1931–1938), which was also published separately and was succeeded by *Annalen van het Genootschap voor Wetenschappelijke Philosophie* (Assen, 1939–1959), which was likewise published separately in addition to being included in this periodical. Suspended 1944–1946.

One periodical begun in this period originated in what is now Czechoslovakia, and one in what is now Poland:

1902–1937. *Česká Mysl; Casopis Filosofický* (Prague), quarterly, with variations.

1911–. *Ruch Filozoficzny* (Torun; previously Lvov), originally monthly; now quarterly. Was a supplement to *Przegląd Filozoficzny* (Warsaw, 1897–1949) from 1911 to 1914. Suspended 1915–1919, 1939–1947, and 1951–1957 (the third period being one of militant Marxist domination of Polish philosophy). Organ of the Polskie Towarzystwo Filozoficzne.

Two of the general prewar products were English-language journals. One was an Irish intellectual quarterly:

1912–. *Studies; An Irish Quarterly Review of Letters, Philosophy & Science* (Dublin), quarterly. Title on individual issues is now *Studies; An Irish Quarterly Review*, but the annual title page for bound volumes continues to use the full title.

The other English-language philosophical periodical of a general character has been associated from the start with Columbia University.

"The Journal of Philosophy." *The Journal of Philosophy* (New York, 1904–), issued fortnightly, was founded by Frederick J. E. Woodbridge and Wendell T. Bush. The title was *Journal of Philosophy, Psychology, and Scientific Methods* from 1904 to 1920. Provocative articles by William James, Arthur O. Lovejoy's "The Thirteen Pragmatisms" (1908), the "First Platform and Program of the New Realists" (1910), and numerous other notable articles have appeared in its pages. Dewey was a frequent contributor, and his philosophy has been analyzed and appraised in the *Journal* from many angles.

A few of the other important articles, on a variety of subjects, that have appeared in the *Journal* are C. I. Lewis' "A Pragmatic Conception of the A Priori" (1923), Herbert Feigl and Albert Blumberg's "Logical Positivism, A New Movement in European Philosophy" (1931), which introduced the term "logical positivism," Ernest Nagel's "Impressions and Appraisals of Analytic Philosophy in Europe" (1936), W. V. Quine's "Designation and Existence" (1939), C. G. Hempel's "The Function of General Laws in History" (1942), Nelson Goodman's "The Problem of Counterfactual Conditionals" (1947), and Norman Malcolm's "Knowledge of Other Minds" (1958). Also noteworthy are Nagel's penetrating reviews, which were frequently featured in the *Journal* in the 1930s and 1940s.

From 1933 to 1936 the *Journal* published annual world-wide bibliographies of philosophy. In more recent years it has carried texts of papers presented at the annual meetings of the Eastern Division of the American Philosophical Association. In 1963/1964 the *Journal* was involved in a minor *cause célèbre* when, after publishing an article by one of its editors on the discussion between non-Soviet and Soviet philosophers at the Thirteenth International Congress of Philosophy (Mexico City, 1963), it declined to provide equal space, although it offered some space, for an article giving a contrary view of the same discussion.

WORLD WAR I TO 1928

During World War I two new philosophical journals were begun in Europe and one each in Argentina and Japan. Only the Japanese journal is still being published. One of the European journals was a new *Logos*.

1914–1943. *Logos; Rivista trimestrale di filosofia e di storia della filosofia* (Perugia; later Naples and Florence; then Rome), quarterly. International board of editors. Suspended 1916–1919. Title was *Logos*, without the subtitle, from 1914 to 1938.

1915–1929. *Revista de filosofia, cultura, ciencias, educación* (Buenos Aires), issued six times a year, with variations.

1916–. *Tetsugaku Kenkyu; Journal of Philosophical Studies* (Kyoto), monthly. Organ of the Philosophical Society of Kyoto University. Contributors have included Heidegger and Jaspers.

1918–1943. *Blätter für deutschen Philosophie; Zeitschrift der Deutschen Philosophischen Gesellschaft* (Erfurt; later Berlin), quarterly. Title was *Beiträge zur Philosophie des deutschen Idealismus* from 1918 to 1927.

The years from 1919 to 1928 saw 32 new journals of philosophy roll off the presses, the largest quota in any ten-year period up to that time. Included were *The Per-*

sonalist, the first Chinese philosophical journals, another *Logos, The Australasian Journal of Philosophy, Philosophy*, and *The New Scholasticism*. Of the total of 32, 7 were Italian (5 survive), 5 German (2 survive), 5 French (all survive), and 4 American (all survive); the rest were scattered among China (3, none surviving), Czechoslovakia (2, none surviving), and Australia, Great Britain, India, Lithuania, the Netherlands, and Poland.

ITALIAN. Three of the Italian journals of the postwar decade have Latin names and concern theological matters chiefly:

1920–. *Gregorianum* (Rome), quarterly. Published by the Università Gregoriana di Roma. Articles in English, French, German, Italian, Latin, and Spanish.

1924–. *Angelicum* (Rome), quarterly. Journal of the Faculty of Theology, Canon Law, and Philosophy, Pontificium Athenaeum Angelicum. Articles in French, German, Italian, and Latin.

1926–. *Antonianum; Periodicum Philosophico-theologicum Trimestre* (Rome), quarterly. Published by the Athenaeum Antonianum de Urbe. Articles mainly in Latin; those in other languages are summarized in Latin

The other Italian philosophical journals of the period cover various fields:

1920–1923. *Rivista trimestrale di studi filosofici e religiosi* (Perugia), quarterly.

1920–. *Giornale critico della filosofia italiana* (Florence; previously Messina, Milan, Rome, and elsewhere), quarterly, with variations. Founded by Giovanni Gentile and edited by him until his assassination in 1944.

1921–. *Rivista internazionale di filosofia del diritto* (Milan), issued six times a year, with variations.

1924–1945. *L'idealismo realistico; Rivista di filosofia mazziniana* (Rome), monthly.

GERMAN. *Erkenntnis* (see below) was the most important journal of the postwar decade, but four other German journals also merit attention.

1919–1924. *Grundwissenschaft; Philosophische Zeitschrift der Johannes-Rehmke-Gesellschaft* (Leipzig), quarterly. Subtitle varied.

1923–1932. *Literarische Berichte aus dem Gebiete der Philosophie* (Erfurt), semiannual, with variations. Title was *Literarische Berichte der Deutschen Philosophischen Gesellschaft* from 1923 to 1924.

The two German journals begun in this period that are still on earth are concerned with heavenly matters:

1923–. *Neue Zeitschrift für systematische Theologie und Religionsphilosophie* (Berlin; originally Gütersloh), issued three times a year; formerly quarterly (irregular 1956–1959). Title was *Zeitschrift für systematische Theologie* from 1923 to 1958, and *Neue Zeitschrift für systematische Theologie* from 1959 to 1962. Suspended 1944–1949.

1926–. *Scholastik; Vierteljahresschrift für Theologie und Philosophie* (Frankfurt; previously Freiburg im Breisgau), quarterly. Published by the Jesuits of the faculties of philosophy and theology, Hochschule St. Georg, Frankfurt, and Berchmanskolleg, Pullach-am-Main. Suspended 1941–1943; combined with the *Theologische Quartalschrift* (1819–) for one year, 1944; suspended 1945–1948.

"*Erkenntnis.*" *The Journal of Unified Science (Erkenntnis)* (Leipzig; later The Hague and Chicago, 1919–1940) was issued six times a year, with variations. Its title was *Annalen der Philosophie, mit besonderer Rücksicht auf die Probleme der als ob Betrachtung* from 1919 to 1923, *Annalen der Philosophie und philosophischen Kritik* from 1924 to 1930, and *Erkenntnis, zugleich Annalen der Philosophie* from 1930 to 1939. Hans Vaihinger was coeditor from 1919 to 1930. From 1930 to 1940 the editors were Rudolf Carnap and Hans Reichenbach (but Carnap alone in 1937/1938).

In the 1930s *Erkenntnis* was perhaps the most influential philosophical periodical ever published. The Vienna circle of logical positivists took over the journal, then entitled *Annalen*, in 1930 (Vaihinger, its coeditor, was then 78 years old), renamed it *Erkenntnis*, and transformed it into a medium—which struck sparks of fire in the philosophical world—for the discussion and propagation of the circle's theses. The first issue of *Erkenntnis* contained Moritz Schlick's "Die Wende der Philosophie" ("The Turning Point in Philosophy") as the opening article and also Carnap's "Die alte und die neue Logik" ("The Old and the New Logic"). In quick succession, in the early 1930s, the periodical published Carnap's "Überwindung der Metaphysik durch logische Analyse der Sprache" ("Elimination of Metaphysics Through Logical Analysis of Language"), probably his most famous paper; Schlick's "Positivismus und Realismus" ("Positivism and Realism") and "Über das Fundament der Erkenntnis" ("On the Foundation of Knowledge"); Otto Neurath's "Pro-

tokollsätze" ("Protocol Sentences"); and Ernest Nagel's "Measurement."

Other notable articles which appeared in *Erkenntnis* are Hans Reichenbach's "Wahrscheinlichkeitslogik" ("Logic of Probability," 1935) and others by him on probability theory, Max Black's "Relations Between Logical Positivism and the Cambridge School of Analysis" (1939), and articles by Niels Bohr and other famous scientists and mathematicians, not all of whom were logical positivists. Various issues of *Erkenntnis* contained the proceedings of the Tagung für Erkenntnislehre der Exakten Wissenschaften (1929–1930), and of the International Congress for the Unity of Science (1934–1938).

Many of the articles published in *Erkenntnis* were translated into English and other languages and published in collections of the foundation papers of the logical positivist movement. Indeed, the journal had its greatest impact on philosophers in England and the United States rather than on those in continental Europe, many of whom had fallen under the spell of Martin Heidegger's *Dasein*.

FRENCH. Two of the French periodicals of the first postwar decade are religiously oriented.

1921–. *Revue d'histoire et de philosophie religieuse* (Strasbourg), quarterly. Published by the Facultéde Théologie Protestante de I'Université de Strasbourg.

1924–. *Bulletin thomiste* (Étiolles, Soisy-sur-Seine), quarterly, with variations. Organ of the Société Thomiste.

The other three are secular and humanistic.

1923–. *Archives de philosophie* (Paris), quarterly, with variations. Suspended 1953–1954.

1926–. *Les Études philosophiques* (Paris), quarterly, with variations. Founded by Gaston Berger.

1927–. *Revue des sciences humaines* (Lille and elsewhere), quarterly, with variations. Title was *Revue d'histoire de la philosophie* from 1927 to 1931 and *Revue d'histoire de la philosophie et d'histoire générale de la civilisation* from 1933 to 1946. Suspended 1932, 1940–1941, and 1945.

EASTERN EUROPEAN. During the 1920s Prague was the birthplace of two philosophical journals and Kaunas and Cracow of one each (including another *Logos*):

1920–1939. *Ruch Filosoficzký* (Prague), issued six times a year, with variations.

1921–1938. *Logos; Filosofijos Laikraštis* (Kaunas), semiannual.

1922–1950. *Kwartalnik Filozoficzny* (Cracow), quarterly. Published by the Polskiej Nakladem Akademii Umiejętności. Suspended 1934 and 1940–1945. Editor in the last years of the periodical was Roman Ingarden. Emphasis on phenomenology and conceptual analysis.

1927–1929. *Filosofie* (Prague), issued ten times a year. Published under the auspices of the Ministerstvo Školstvi a Národni Osvěty of Czechoslovakia.

AMERICAN. In the United States a personalistic magazine and three religious journals were founded and are still being issued:

1920–. *The Personalist* (Los Angeles), quarterly. Issued by the University of Southern California.

1923–. *The Modern Schoolman* (St. Louis, Mo.), quarterly.

1926–. *Thought; A Review of Culture and Idea* (New York), quarterly. Founded by the Jesuit periodical *America*; directed since 1940 by Fordham University. Subtitle was *A Quarterly of the Sciences and Letters* from 1926 to 1939.

1927–. *The New Scholasticism* (Washington), quarterly. Organ of the American Catholic Philosophical Association. This periodical is one of the two best sources of philosophical news (teaching appointments, publication projects, congresses, etc.), the other being the *Revue philosophique de Louvain* (1894–).

ASIAN AND AUSTRALASIAN. The three Chinese journals of philosophy that were introduced in the 1920s are:

1921–1927. *Chê Hsüeh* [”Philosophy”] (Peking), issued six times a year, with variations. Generally referred to as *Chê Hsüeh Tsa Chih* ("Philosophical Journal").

1926–1930. *Chê Hsüeh Yüeh K'an* [”Philosophical Monthly”] (Peking), monthly, with variations.

1927–1944. *Chê Hsüeh P'ing Lun* [”Philosophical Review”] (Peking), issued six times a year, with variations.

A journal published in India continues to be active:

1925–. *Philosophical Quarterly* (Calcutta; later Amalner), quarterly. Organ of the Indian Institute of Philosophy and the Indian Philosophical Congress.

"The Australasian Journal." One journal published in Australia merits a pause for special comment: *The Australasian Journal of Philosophy* (Glebe, New South Wales, Australia, 1923–), issued quarterly from 1923 to 1937 and three times a year since 1938. It is the organ of the Australasian Association of Psychology and Philosophy. The title was *The Australasian Journal of Psychology and Philosophy* from 1923 to 1946.

The *Journal* announced in its first issue that some of its articles would be technical and addressed to a few experts, whereas others would treat of "topics of universal interest, ranging from the high metaphysical quest of the secret of the Absolute, to concrete problems of social and political ethics." It undertook not to "scorn the old fogey in Philosophy, or disdain the new faddist." Bertrand Russell helped the *Journal* get off to a flourishing start by publishing in its first volume (second issue) a little-known but important article of his, "Vagueness."

In 1935 John Anderson, a controversial philosopher of Scottish origin, became the editor of the *Journal*. He thereafter exerted a strong influence not only on the *Journal* but also on the thinking of philosophers in his part of the world. The legislators of New South Wales, shocked by Anderson's militant atheism, unsuccessfully demanded his removal from his teaching post at the University of Sydney.

The current editor of *The Australasian Journal of Philosophy* is A. K. Stout, son of G. F. Stout, former editor of *Mind*. This is the third notable case in which a son followed his father's trade as editor of a philosophical journal, the other such families being the Fichtes, who respectively edited the *Philosophisches Journal*, 1795 ff., and the *Zeitschrift für Philosophie*, 1837 ff.; and the Morandos, who were editors at different times of the *Rivista rosminiana*, 1906 ff.

Among the challenging and widely discussed papers that Anderson on his pluralistic, positivistic realism; the last pieces by the elder Stout; some of the most celebrated articles by J. N. Findlay and others in the early 1940s on the philosophy of Wittgenstein; J. A. Passmore's three articles entitled "Logical Positivism" (1943, 1944, and 1948) and his "Christianity and Positivism" (1957); J. J. C. Smart's "The Reality of Theoretical Entities" (1956); A. N. Prior's "The Autonomy of Ethics" (1960); and Keith Lehrer's "Doing the Impossible" (1964). The *Journal's* influence reached a particularly high level in the period beginning about 1955.

BRITISH AND DUTCH. The remaining examples of journals begun in the first postwar decade have had London and Hilversum as their headquarters:

1926–. *Philosophy* (London), quarterly. Orgán of the British Institute of Philosophy. Title was *Journal of Philosophical Studies* from 1926 to 1931. Contributors have included Samuel Alexander, George Dawes Hicks, and Bertrand Russell.

1926–1944. *Denken en Leven; Wijsgeerig Tijdschrift* (Hilversum), issued six times a year.

1929–1938

From 1929 to 1938, 25 new journals of philosophy sought subscribers. Of these, 8 have fallen by the wayside, 6 being casualties of World War II. Among the new journals of this period were a Yugoslav quarterly, the first journals covering the philosophy of science and symbolic logic, and *Analysis*. Italy produced the most new philosophical journals (10); Germany produced none.

IDEALISTIC, RELIGIOUS, AND MYSTIC. Two publications on nonworldly philosophy were established in India:

1930–1935. *Review of Philosophy and Religion* (Poona), semiannual. Organ of the Academy of Philosophy and Religion.

1930–. *The Aryan Path* (Bombay), monthly. Popular ethics and mysticism, with emphasis on Indian philosophy.

The remaining examples of this kind of journal had their homes in Belgium, Italy, the Netherlands, and the United States, respectively:

1929–. *Recherches de théologie ancienne et médiévale* (Louvain), quarterly.

1934–. *Doctor Communis; Acta et Commentationes Pontificiae Accademiae Sanctae Thomae Aquinatis* (Rome; previously Turin), issued three times a year. Title was *Acta Pontificiae Accademiae Sanctae Thomae Aquinatis* from 1934 to 1947. Articles mainly in Latin; those in other languages are summarized in Latin.

1938–. *Bijdragen van de Philosophische en Theologische Faculteiten der Nederlandsche Jezuieten* (Maastricht), issued three times a year, with variations. Title has varied.

1938–. *Vedanta and the West* (Hollywood, Calif.), issued six times a year. Emphasis on mysticism.

Sponsored by the Vedanta Society of Southern California.

LOGIC AND RELATED DISCIPLINES. Balancing the inaugurations of religious periodicals were those of periodicals on logic, philosophy of science, and language analysis. The two most influential were *Analysis* and the *Journal of Symbolic Logic*.

"Analysis." *Analysis* (Oxford, 1933–) is issued six times a year, with variations. The journal was suspended from 1940 to 1947. This periodical was founded by a number of younger philosophers under the influence of G. E. Moore, Bertrand Russell, and Ludwig Wittgenstein. It was intended mainly as a medium for short analyses and discussions. A group of supporters pledged to pay £5 each if the venture should so require, but the journal paid its way. In 1936 an Analysis Society was formed, also aimed at guaranteeing the financial stability of the journal, but it went out of existence a few years later; some of the papers read at its meetings were published in *Analysis.*

Max Black, in America, was closely associated with the journal from its foundation, and Rudolf Carnap, Carl Hempel, and Moritz Schlick, of the Vienna circle, contributed articles to early issues. Among the memorable articles that *Analysis* has published are A. J. Ayer's "The Genesis of Metaphysics" (1934), Schlick's "Facts and Propositions" (1935), Margaret Macdonald's "Necessary Propositions" (1940), Black's "The Semantic Definition of Truth" (1948), Friedrich Waismann's six articles entitled "Analytic–Synthetic" (1949–1953), P. T. Geach's "Russell's Theory of Descriptions" (1950), Alonzo Church's "On Carnap's Analysis of Statements of Assertion and Belief" (1950), Gilbert Ryle's "Heterologicality" (1951), Karl R. Popper's "A Note on the Body–Mind Problem" (1955), Yehoshuah Bar-Hillel's "New Light on the Liar" (1957), Peter Achinstein's "The Circularity of a Self-supporting Inductive Argument" (1962), and Keith Gunderson's "Interview With a Robot" (1963).

Many highlights from the journal were reprinted in *Philosophy and Analysis* (New York, 1954), edited by Margaret Macdonald, who was editor of *Analysis* from 1948 to her death in January 1956. For a time in the 1950s, Analysis conducted "competitions" and published the best short answers to such questions as "Does it make sense to say that death is survived?"

Especially in the early years of *Analysis*, its pages crackled with iconoclasm, terseness, and wit. Currently, some of the articles are longer than the average of the early years, and a supplement containing extended articles is now issued annually.

"Journal of Symbolic Logic." The *Journal of Symbolic Logic* (Providence, R.I.; previously Menasha, Wis., and Baltimore, Md., 1936–), issued quarterly, publishes articles in English, French, and German. It is the organ of the Association for Symbolic Logic.

This journal was the first one to be devoted exclusively to its field. In April 1934, Paul Weiss called attention to the fact that papers on logic were scattered in heterogeneous periodicals, and (without specifically proposing a new periodical) he suggested the formation of a logic association. Later in the year, C. J. Ducasse and C. A. Baylis explicitly urged the establishment of a journal of symbolic logic, to be supported by an association for symbolic logic. The response was encouraging, and the venture was undertaken.

Financing the *Journal* was a problem in the early years, and it was uncertain, after the publication of the third issue, whether the publication could continue. Happily, subventions were obtained from a number of universities, and dues payments accumulated sufficiently to enable the *Journal* to meet its bills.

Aside from the high quality of many of the articles, the *Journal* is noted for an exceptionally useful section devoted to reviews and abstracts of current literature. These reviews and abstracts purport to cover all pertinent books and articles which have come to the attention of the editors; the frame of reference of publications pertinent to symbolic logic is interpreted broadly. The reviews and abstracts constitute a continuation of Alonzo Church's nonpareil bibliography of symbolic logic from 1666 to 1935 which appeared in the issue of December 1936, with a supplement in the issue of December 1938.

The well-deserved international reputation of the *Journal* derives in large part from the vast knowledge and logical acumen of Church, who is the principal editor. Among the many articles of enduring worth which have appeared in the *Journal* are Church's "A Note on the *Entscheidungs-problem*" (1936), Barkley Rosser's "Extensions of Some Theorems of Gödel" (1936), W. V. Quine's "On the Theory of Types" (1938) and his "On Universals" (1947), Carl G. Hempel's "A Purely Syntactical Definition of Confirmation" (1943), Rudolf Carnap's "Modalities and Quantification" (1946), Wilhelm Ackermann's "Begründung einer strengen Implikation" (1956), Gordon Matheson's "The Semantics of Singular Terms" (1962), and Frederic B. Fitch's "A Logical Analysis of Some Value Concepts" (1964).

Other journals. Other journals on logic, analysis, and so forth, were published in Poland, the United States, and the Netherlands:

1934–. *Studia Logica* (Warsaw), semiannual, with variations; formerly an annual. Suspended 1937–1952. Sponsored since 1953 by the Komitet Filozoficzny, Polska Akademia Nauk. Articles in English, French, German, Polish, and Russian, each with summaries in two other languages. In 1953 it absorbed the irregularly published *Studia Philosophica* (Warsaw, 1935–1951, four volumes).

1934–. *Philosophy of Science* (East Lansing, Mich.), quarterly. Organ of the Philosophy of Science Association.

1936–. *Synthese; An International Quarterly for the Logical and the Psychological Study of the Foundations of Science* (Dordrecht; previously Utrecht), quarterly, with variations. Subtitle has varied. Suspended 1940–1945 and 1964–1965. Articles in English, French, and German (originally, mainly in Dutch). Various issues have included a section (sometimes separately paged) entitled "Communications of the Institute for the Unity of Science" or "Unity of Science Forum."

SOCIAL AND MORAL PHILOSOPHY. Italy produced three, and the United States one, of the social and moral periodicals that started in this prewar period:

1932–1943. *Archivio della cultura italiana* (Rome), quarterly. Title was *Archivio di storia della filosofia italiana* from 1932 to 1938.

1935–1941. *Rassegna di morale e di diritto* (Rome), quarterly.

1935–1942. *Journal of Social Philosophy & Jurisprudence; A Quarterly Devoted to a Philosophic Synthesis of the Social Sciences* (New York), quarterly. Title was *Journal of Social Philosophy* (with the same subtitle as later) from 1935 to 1941.

1935–. *Rivista internazionale di filosofia politica e sociale* (Genoa; formerly Padua), quarterly. Suspended 1944–1963.

PHILOSOPHY, HISTORY, AND LETTERS. Italy fathered three journals linking history and literature with philosophy:

1929–1943. *Civiltà moderna; Rassegna bimestrale di critica storica, letteraria, filosofica* (Florence), issued six times a year.

1929–. *Convivium; Rivista di lettere, filosofia e storia* (Turin), issued six times a year. Suspended 1944–1946. Subtitle has varied.

1931–. *Ricerche filosofiche; Rivista di filosofia, storia e letteratura* (Messina), semiannual, with variations. Since 1948 it has been the organ of the Società Filosofica Calabrese, founded in that year.

GENERAL. Seven regular academic or professional periodicals devoted to philosophy in general were begun in this period:

1931–1959. *Annalen van het Genootschap voor Wetenschappelijke Philosophie* (Assen), issued five times a year. Title was *Annalen der critische Philosophie* from 1931 to 1938. In 1959 absorbed into *Algemeen Nederlands Tijdschrift voor Wijsbegeerte en Psychologie* (Amsterdam and later Assen, 1907–), after having been published both separately and as a section of that periodical from 1934 to 1959.

1931–. *Archivio di filosofia* (Rome), issued three times a year, with variations. Originally the organ of the Società Filosofica Italiana; more recently the organ of the Istituto di Studi Filosofici and the Associazione Filosofica Italiana. Suspended 1943–1945.

1933–. *Sophia; Rassegna critica di filosofia e storia della filosofia* (Rome; formerly Palermo, Naples, and Padua), quarterly. Became international in 1935. Subtitle has varied. Contains articles in English, French, German, Italian, and Spanish, with subtitles in these languages.

1935–1940. *Bollettino filosofico* (Rome), quarterly.

1935–. *Theoria* (Lund, Goteborg, and Copenhagen; previously Goteborg), issued three times a year. Contains articles in English, French, and German (before 1937, in Danish, Norwegian, and Swedish).

1936–1940. *Philosophia; Philosophorum Nostri Temporis Vox Universa* (Belgrade), quarterly, with variations. Contained articles in English, French, and German.

1938–. *Revue internationale de philosophie* (Brussels), quarterly, with variations. Suspended 1939–1948. Each issue is devoted to a movement, problem, or philosopher, with a comprehensive bibliography.

WORLD WAR II

In the seven years from 1939 to 1945, 21 journals of philosophy came into being. Fully 16 of these have survived,

and they include a number of today's outstanding philosophical journals.

NORTH AMERICAN. Canada, the United States, and Mexico produced a total of eight philosophical journals during World War II. Canada provided a new medium for discussions of theology and philosophy, *Laval théologique et philosophique* (Quebec, 1945–), issued semiannually. This journal is published by the Facultés de Théologie et Philosophie de l'Université Laval de Québec.

In Mexico, for 17 years, a university review of philosophy and letters was published: *Filosofía y letras* (Mexico City, 1941–1957), issued quarterly, with variations. It was the organ of the Facultad de Filosofía y Letras, Universidad Nacional Autónoma.

In the United States six periodicals, varying widely in their character and in their topical focus, began in the period from 1939 to 1945. Five of these were:

1939–1954. *Philosophic Abstracts* (New York), quarterly, with variations.

1939–. *The Thomist; A Speculative Quarterly Review* (Washington; formerly New York), quarterly. Edited by the Dominican Fathers of the Province of St. Joseph.

1940–. *Journal of the History of Ideas* (New York), quarterly.

1941–. *Journal of Aesthetics and Art Criticism* (Baltimore Md.), quarterly. Organ, since 1945, of the American Society for Aesthetics. Contributors have included Croce, Dewey, and Santayana.

1943–. *Etc.: A Review of General Semantics* (San Francisco; formerly Bloomington, Ill.), quarterly. Organ of the International Society for General Semantics. Anthology volumes, consisting of selections from *Etc.*, were published in 1954 and 1959.

"Philosophy and Phenomenological Research." The most influential journal begun during World War II was *Philosophy and Phenomenological Research* (Buffalo; then Philadelphia; now Buffalo again, 1940–), which is issued quarterly. This journal is an outgrowth of the *Jahrbuch für Philosophie und phänomenologische Forschung* (Halle, 1913–1930), which was founded by Edmund Husserl.

Husserl died in 1938. In the following year the International Phenomenological Society was formed in New York City to further the understanding, development, and application of phenomenological inquiry as inaugurated by Husserl. The Society's journal, *Philosophy and Phenomenological Research*, although taking Husserl's philosophy as "the point of departure," announced at the outset that it would represent "no special school or sect." Its editor for a quarter of a century, Marvin Farber, has kept the journal's pages open to diverse points of view.

Philosophy and Phenomenological Research published the proceedings of the First Inter-American Conference of Philosophy (held at Yale University in 1943) and several stimulating symposia. The symposia dealt with meaning and truth, with articles by C. A. Baylis, C. J. Ducasse, Felix Kaufmann, C. I. Lewis, Ernest Nagel, R. W. Sellars, Alfred Tarski, W. M. Urban, A. Ushenko, and John Wild (1943–1945); probability, with articles by Gustav Bergmann, Rudolf Carnap, Kaufmann, Richard von Mises, Nagel, Hans Reichenbach, and Donald Williams (1945–1946); Russian philosophy and psychology, educational philosophy, "philosophy of freedom," and the philosophy of Arthur O. Lovejoy (various years in the 1940s and 1963); and "logical subjects and physical objects," with articles by Wilfrid Sellars and P. F. Strawson (1957).

Among the memorable individual articles in the journal were three little-known papers by Husserl entitled "Notizien zur Raumkonstitution" (Nos. 1 and 2, 1940), "Phänomenologie und Anthropologie" (1941), and "Persönliche Aufzeichnungen" (1956). Others include Paul Weiss's "The Meaning of Existence" (1940), Ernst Cassirer's "The Concept of Group and the Theory of Perception" (1944), Arthur Pap's "Logical Nonsense" (1948), Richard Mc-Keon's "Dialogue and Controversy in Philosophy" (1956), Lewis S. Feuer's "The Bearing of Psychoanalysis Upon Philosophy" (1959), Nagel's "Determinism in History" (1960), and Nicholas Rescher's "On the Logic of Presupposition" (1961). The journal publishes Spanish abstracts of its articles.

SOUTH AMERICAN. In 1944 two philosophical periodicals were established in Argentina:

1944–. *Stromata: Ciencia y fé* (Buenos Aires), issued quarterly by the Facultades de Filosofía y Teología, Colegio Máximo de San José, San Miguel. Title was *Ciencia y fé* from 1944 to 1964. Considered to be the successor to *Fascículos de biblioteca* (1937–1943) and *Stromata* (1938–1943).

1944–. *Philosophia* (Mendoza, Argentina), semiannual. Issued by the Instituto de Filosofía y Disciplinas Auxiliares, Universidad Nacional de Cuyo.

WESTERN EUROPEAN. Despite the atmosphere of war or preparations for war, new journals for philosophical discussion were begun in Belgium and France and in Spain and Portugal:

1939–. *Tijdschrift voor Filosofie* (Louvain), quarterly. Articles in English, Dutch, French, and German, with English, French, or German summaries of the articles in Dutch. Editors are chosen from Netherlands universities and Dutch-language universities of Belgium.

1942–. *Revista de filosofía* (Madrid), issued three times in 1942, quarterly since 1943. Organ of the Instituto de Filosofía Luis Vives. Scholastic. Some foreign contributors.

1945–. *Pensamiento; Revista de investigación e información filosófica* (Madrid), quarterly. Organ of the Facultades de Filosofía, Compañía de Jesús en España. Strong on the bibliography of Spanish and Latin American philosophy.

1945–. *Revista portuguesa de filosofia* (Braga; formerly Lisbon), quarterly. Organ of the Faculdade Pontifícia de Filosofia of Braga, a branch of the Society of Jesus.

1945–1955. *Dieu vivant; Perspectives religieuses et philosophiques* (Paris), quarterly, with variations.

CENTRAL AND SOUTHERN EUROPEAN. Contributions of Italy and neutral Switzerland were:

1940–1943. *Bollettino dell'Istituto di Filosofia del Diritto dell'Università di Roma* (Rome), issued six times a year.

1940–1949. *Studi filosofici; Problemi di vita contemporanea* (Milan), quarterly. Pro-Marxist from 1946 to 1949. Subtitle varied.

1945–. *Methodos; Linguaggio e cibernetica* (Milan; previously Rome), quarterly, with variations. Title was *Analisi; Rassegna di critica della scienza* from 1945 to 1947 and *Sigma; Conoscenza unitaria* from 1947 to 1948. Subtitle has varied. Contains articles in various languages. International editorial board. Organ, since 1959, of the Centro di Cibernetica e di Attività Linguistiche, Università di Milano, and of the Consiglio Nazionale delle Ricerche.

1945–. *Theologische Zeitschrift* (Basel), issued six times a year.

BULGARIAN AND ISRAELI. In Bulgaria and Israel the following journals came into being:

1945–. *Filosofska Mis'l* (Sofia), issued six times a year. Table of contents also in English, French, German, and Russian; summaries in English and Russian. Issued since 1952 by the Institut po Filosofia, Bulgarska Akademiia na Naukite.

1945–. *Iyyun* (Jerusalem), quarterly. Irregular 1945–1948; suspended 1949–1950. Contains English summaries.

POSTWAR PERIOD

In the early postwar years philosophical journals were founded at an unprecedented pace. They numbered 11 in 1946 (of which 9 have survived); 8 in 1947 (6 still alive); 5 in 1948 (4 still alive); and 7 in 1949 (3 still alive). Among them was another *Logos*.

1946. Three products of the first postwar year had humanistic titles:

1946–. *Teoresi; Rivista di cultura filosofica* (Catania; formerly Messina), quarterly, with variations. Emphasizes the synthesis of idealism and realism.

1946–. *Sapientia* (Buenos Aires), quarterly. Organ of the Facultad de Filosofía, Universidad Católica Argentina. Thomist. International contributors.

1946–. *Humanitas* (Brescia, Italy), monthly. In four parts, of which the part on philosophy is edited by Michele Federico Sciacca.

Four journals, including two from Japan, had standard, traditional titles:

1946–1949. *Tetsugaku Hyôron; Philosophical Review* (Tokyo), monthly.

1946–1949. *Tetsugaku Kikan* ["Quarterly Review of Philosophy"] (Kyoto), quarterly.

1946–. *Giornale di metafisica* (Turin), issued six times a year. Founded and edited by M. F. Sciacca. From 1946 to 1948 published by the University of Pavia; since then, by the University of Genoa. Has been described as following the Plato–Augustine–Rosmini tradition. Contributors include Maurice Blondel, Gabriel Marcel, and Jacques Maritain.

1946–. *Zeitschrift für philosophische Forschung* (Meisenheim am Glan, Germany; formerly Wurzach), quarterly.

The others cover a variety of fields:

1946–. *Otázky Marxistickej Filozofie* (Bratislava, Czechoslovakia; formerly Prague), issued six times a year, with variations. Title was *Philosophica Slovaca* from 1946 to 1949 (issued annually); *Filozofický Sborník* from 1950 to 1952 (issued annually); *Filo-*

zofický Časopis from 1953 to 1955 (quarterly); and *Slovenský Filozofický Časopis* from 1956 to 1960 (quarterly). Issued by the Slovenská Akadémie Vied. Table of contents also in English, German, and Russian. Emphasis on historical materialism.

1946–. *Rassegna di scienze filosofiche* (Naples; previously Bari and Rome), quarterly. Title was *Noesis; Rassegna internazionale di scienze filosofiche e morali* in 1946. Suspended 1947. Neo-Scholastic.

1946–. *Rivista critica di storia della filosofia* (Milan), quarterly. Title was *Rivista di storia della filosofia* from 1946 to 1949.

1946–. *Nederlands Theologisch Tijdschrift* (Wageningen), issued six times a year.

1947. Two of the 1947 products expired within 3 to 11 years:

1947–1949. *Tetsugaku* ["Philosophy"] (Tokyo), quarterly.

1947–1958. *Wiener Zeitschrift für Philosophie, Psychologie, Pädagogik* (Vienna), semiannual.

The ones that are still alive include two that are general in their scope:

1947–. *Archiv für Philosophie* (Stuttgart), quarterly. Not to be confused with the *Archiv für Geschichte der Philosophie* (Berlin, 1868–), which was entitled *Archiv für Philosophie* from 1895 to 1926. Some issues of the Stuttgart periodical, beginning in the late 1940s, incorporated issues of the irregularly published *Archiv für mathematische Logik und Grundlagenforschung*.

1947–. *Voprosy Filosofii* (Moscow), monthly, with variations. Issued by the Institut Filosofii, Akademiia Nauk SSR. Contains summaries in English and titles in English, French, German, and Spanish.

"Review of Metaphysics." The *Review of Metaphysics* (New Haven, 1947–), published quarterly, is one of the major media of discussion of the perennial problems of metaphysics. In addition, it publishes annual lists of doctoral dissertations accepted by philosophy departments in the United States and Canada, of professors who have become emeritus in philosophy, and of visiting philosophy professors from abroad. Beginning with December 1964, each issue contains abstracts of articles in certain philosophical periodicals, written (as in the case of the book abstracts formerly published in *Kant-Studien*) by the authors of the articles themselves. In earlier years the *Review* conducted competitions, comparable to those in *Analysis* (1933–), for the best short answers to piquant questions, such as why there has never been a great woman philosopher.

Outstanding among the many important articles that have appeared in the *Review* are Paul Weiss's "Being, Essence and Existence" (1947), W. V. Quine's "On What There Is" (1948), Charles Hartshorne's "The Immortality of the Past" (1953), Nathan Rotenstreich's "The Genesis of Mind" (1962), and Wilfrid Sellars' "Abstract Entities" (1963). The discussion section of the *Review* has also provided a large number of valuable contributions to current thought.

Two of the 1947 periodicals concern the philosophy of science or the unity of the sciences, and one is bibliographical:

1947–. *Dialectica; International Review of Philosophy of Knowledge* (Neuchâtel, Switzerland; and Paris), quarterly. Emphasis on philosophy of science.

1947–. *Studium Generale; Zeìtschrift für die Einheit der Wissenschaften im Zusammenhang íhrer Begriffsbildungen und Forschungsmethoden* (Berlin), monthly. Articles in English, French, and German.

1947–. *Bulletin signalétique: Philosophie, sciences humaines* (Paris), quarterly. Title was *Bulletin analytique: Philosophie* from 1947 to 1955. Contains abstracts of books and articles on philosophical subjects. Published by the Centre de Documentation du Centre Nationale de la Recherche Scientifique.

1948. Three journals begun in 1948 were founded on the European continent:

1948–. *Revue d'esthétique* (Paris), quarterly.

1948–. *Sapienza; Rivista di filosofia e di teologia dei Domenicani d'Italia* (Naples), issued six times a year. Since 1956 the organ of the Centro Italiano di Studi Scientifici, Filosofici e Teologici. Subtitle has varied.

1948–. *Roczniki Filozoficzne* (Lublin), quarterly.

The others were issued in South America:

1948–1950. *Revista colombiana de filosofía* (Bogotá), issued six times a year. Emphasis on Thomism and phenomenology.

1948–. *Filosofía, letras y ciencias de la educación* (Quito), semiannual. Published by the Facultad de Filosofía, Letras y Ciencias de la Educación, Universidad Central, Quito. Title has varied.

1949. Another *Logos* appeared in 1949, along with two periodicals called "philosophical studies" (in German and in English), and four other journals:

1949–1951. *Logos* (Mexico City), quarterly. Published by the Mesa Redonda de Filosofía, Facultad de Filosofía y Letras, Universidad Nacional Autónoma de Mexico.

1949–1952. *Philosophische Studien* (Berlin), quarterly, with variations.

1949–. *Philosophical Studies* (Minneapolis), issued six times a year. Brief articles. Many distinguished contributors.

1949–1953. *Revista de filosofía* (Santiago, Chile), quarterly. Organ of the Sociedad Chilena de Filosofía and the Universidad de Chile.

1949–1954. *Notas y estudios de filosofía* (Tucumán, Argentina), quarterly.

1949–. *Philosophischer Literaturanzeiger* (Stuttgart; formerly Schlesdorf am Kochelsee, then Stuttgart, then Meisenheim am Glan), issued eight times a year.

1949–. *Analele româno-sovietice; Filozofie* (Bucharest), quarterly, with variations. Table of contents also in Russian. From 1949 to 1951 it was a part of *Analele româno-sovietice; Seria istorie-filozofie* (quarterly; issued six times in 1951; title also in Russian), which itself had been a part, from 1946 to 1949, of *Analele româno-sovietice* (issued irregularly; title also in Russian).

THE NINETEEN-FIFTIES

The decade of the 1950s saw 11 new English-language journals, 13 Spanish-language journals, 11 Italian, 4 Portuguese, 4 French, 3 German, 2 Dutch, and 1 each in Hungarian, Rumanian, Polish, Serbo-Croat, Russian, Chinese, and Japanese. As in two earlier periods, Italy was the leading or a leading producer of new philosophical journals.

ENGLISH. In continental United States and Hawaii the following journals were introduced:

1951–. *Philosophy East and West* (Honolulu), quarterly. Emphasizes Oriental and comparative thought. Suspended from 1964 to 1966.

1957–. *Philosophy Today* (Celina, Ohio), quarterly. Mainly contains reprints or translations of articles appearing elsewhere. Religious emphasis.

In Scotland are published a journal for the philosophy of science and a quarterly which has the same title as a living Indian journal begun in 1925:

1950–. *British Journal for the Philosophy of Science* (Edinburgh), quarterly.

1950–. *The Philosophical Quarterly* (St. Andrews, Scotland), quarterly. Published for the Scots Philosophical Club.

The Commonwealth countries of Canada, India, and Pakistan produced the following periodicals:

1953–. *Diogenes; An International Journal for Philosophy and Humanistic Studies* (Montreal; formerly New York), quarterly. Published under the auspices of the International Council for Philosophy and Humanistic Studies with the assistance of UNESCO.

1953–. *Journal of the Philosophical Association* (Amraoti, India; later Nagpur), quarterly. Organ of the Indian Philosophical Association. Contributors outside India have included P. T. Geach, Elizabeth Anscombe, and A. N. Prior.

1956–. *Indian Philosophy and Culture* (Vrindaban, India), quarterly. Issued by the Vaishnava Research Institute.

1957–. *Pakistan Philosophical Journal* (Lahore), quarterly.

1959–. *The Indian Journal of Philosophy* (Bombay), quarterly; formerly issued three times a year. Published for the Association for Philosophical Research.

From the Netherlands and Norway come the following:

1956–. *Phronesis; A Journal for Ancient Philosophy* (Assen), semiannual.

1958–. *Inquiry; An Interdisciplinary Journal of Philosophy and the Social Sciences* (Oslo), quarterly. Emphasis on analytic philosophy.

SPANISH. In South America five periodicals sprang to life, including one which repeated the title (*Humanitas*) of an Italian journal begun in 1946:

1950–1954. *Revista de filosofía* (La Plata, Argentina), quarterly. Issued by the Instituto de Filosofía, Universidad Nacional de La Plata.

1951–1954. *Ideas y valores* (Bogotá), quarterly. Issued by the Facultad de Filosofía y Letras de la Universidad Nacional. Title varied slightly.

1952–. *Arkhé Revista americana de filosofía sistemática y de historia de la filosofía* (Córdoba, Argentina), semiannual (formerly issued three times a year). Suspended 1955 to mid-1964. Title was originally *Arqué* subtitle varied.

1953–. *Filosofía; Revista semestral* (Quito), semiannual. Organ of the Sección de Ciencias Filosóficas y de la Educación de la Casa de la Cultura Ecuatoriana.

1953–. *Humanitas; Revista de la Facultad de Filosofía y Letras, Universidad Nacional de Tucumán* (San Miguel de Tucumán), issued three times a year, with variations.

In Central America and the Caribbean, two university *Revistas* appeared:

1956–1958. *Revista dominicana de filosofía* (Ciudad Trujillo, now called Santo Domingo), semiannual, with variations. Organ of the Facultad de Filosofía of the Universidad de Santo Domingo.

1957–. *Revista de filosofía de la Universidad de Costa Rica* (San José), semiannual.

In Spain itself six periodicals arose, including one which repeated the title (*Convivium*) of a journal begun at Turin in 1929:

1951–. *Estudios filosóficos; Revista de investigación y crítica* (Las Caldas de Besaya, Spain), issued three times a year. Organ of the Spanish Dominicans.

1951–. *Archivum; Revista de la Facultad de Filosofía y Letras, Universidad de Oviedo* (Oviedo), semiannual, with variations.

1952–. *Espíritu; Cuadernos del Instituto Filosófico de "Balmesiana"* (Barcelona), semiannual, with variations.

1954–. *Crisis; Revista española de filosofía* (Madrid), quarterly. Emphasizes Christian existentialism.

1956–1957. *Convivium; Estudios filosóficos* (Barcelona), semiannual. Issued by the Facultad de Filosofía y Letras, Universidad de Barcelona.

1956–. *Augustinus* (Madrid), quarterly. Many foreign contributors.

ITALIAN. Three of the births of Italian philosophical journals took place at Milan: one each in 1950, 1951, and 1952.

1950–1962. *Il pensiero critico* (Milan), quarterly. In 1963 absorbed into the *Rivista di filosofia* (Milan, 1870–).

1951–. *Aut Aut; Rivista di filosofia e di cultura* (Milan), issued six times a year. Title is based on the Kierkegaardian *Either/Or*

1952–. *Bollettino della Società Filosofica Italiana* (Milan), quarterly.

Three births also occurred at Rome, including that of a journal with a Latin title which contains articles in Italian and other languages:

1952–. *Rassegna di filosofia* (Rome), quarterly. Organ of the Istituto di Filosofia, Universitá di Roma.

1955–. *La nuova critica; Studi e rivista di filosofia delle scienze* (Rome; formerly Florence), semiannual. Articles mostly in Italian, but with some in English and French. International board of editors. The title may reflect a desire for association with Croce's Naples journal *La critica* (1903 ff.), which, under a slightly different title, had died in 1951.

1958–. *Aquinas; Ephemerides Thomisticae* (Rome), issued three times a year, with variations. Subtitle has varied. Now issued by the Faculty of Philosophy, and the Patristic–Medieval Institute "Joannes XXIII," of the Pontificia Universitas Lateranensis. Articles in English, French, Italian, Latin, and Spanish.

The locale of two births was Turin; of two others, Padua; and of one, Bologna:

1950–. *Filosofia* (Turin), quarterly.

1951–. *Il saggiatore; Rivista di cultura filosofica e pedagogica* (Turin), quarterly.

1954–. *Studia patavina; Rivista di filosofia e teologia* (Padua), issued three times a year; formerly a quarterly.

1956–. *Rivista di estetica* (Turin; formerly Padua), issued three times a year.

1957–. *Il dialogo* (Bologna), quarterly, with variations.

PORTUGUESE. The Portuguese-language journals which were brought into being in the 1950s were:

1951–1959. *Revista filosófica* (Coimbra, Portugal), issued three times a year, with variations.

1951–. *Revista brasileira de filosofia* (Sâo Paulo), quarterly. Organ of the Instituto Brasileiro de Filosofia. Chiefly in Portuguese, with some articles in English, French, Italian, Spanish, and other languages.

1954–. *Filosofia; Revista do Gabinete de Estudos Filosóficos* (Lisbon), quarterly. Subtitle has varied.

1959–. *Organon; Revista da Faculdade de Filosofia da Universidade do Rio Grande do Sul* (Pôrto Alegre), quarterly, with variations.

FRENCH. Four new journals of philosophy in the French language appeared in the 1950s, including two published in Belgium (one with articles in English, French, and German) and one published in the Saar (with articles in French and German), which are included here among the French journals, since the titles of two are in French, and the title of the third is in Latin and French:

1951–. *Morale et enseignement* (Brussels), quarterly, with variations. Published by the Institut de Philosophic, Université de Bruxelles.

1951–. *Revue de l'enseignement philosophique* (Paris), issued six times a year, with variations. Organ of the Association des Professeurs de Philosophie de l'Enseignement Public.

1952–. *Annales Universitatis Saraviensis; Philosophie–lettres* (Saarbrücken), quarterly, with variations. Published since 1957 by the Philosophische Fakultät, Universität des Saarlandes. Articles in English, French, and German.

1954–. *Logique et analyse* (Louvain), quarterly, with variations. Articles in English, French, and German. Organ of the Centre National (Beige) de Recherches de Logique; issued only to members from 1954 to 1957 under the title *Bulletin intérieure.*

GERMAN AND DUTCH. Three new journals of philosophy in the German language appeared during the 1950s:

1950–. *Philosophia Naturalis; Archiv für Naturphilosophie und die philosophischen Grenzgebiete der exakten Wissenschaften und Wissenschaftsgeschichte* (Meisenheim am Glan), quarterly, with variations.

1953–. *Philosophische Rundschau* (Heidelberg), quarterly, with variations. Reviews of current books. Concerned largely, in its early years, with surveys of new philosophical literature, this became a general philosophical journal. Contains occasional articles in English.

1953–. *Deutsche Zeitschrift für Philosophie* (East Berlin), monthly, with variations (quarterly, 1953–1954; issued six times a year, 1955–1959). Table of contents also in English, French, Russian, and Spanish.

The Dutch-language journals of the 1950s include one with a Dutch title and one with a Latin title:

1959–. *Dialoog; Tijdschrift voor Wijsbegeerte* (Antwerp), quarterly.

1959–. *Scripta Recenter Edita* (Nijmegen), issued ten times a year. Contains a list of books on philosophy and theology, with emphasis on theology.

RUMANIAN, HUNGARIAN, AND SLAVIC. The period produced one Rumanian and one Hungarian organ, each issued for the most part four times a year:

1954–. *Cercetári filozofice* (Bucharest), quarterly, with variations. Table of contents also in French and Russian; summaries in French or German and in Russian.

1957–. *Magyar Filozófiai Szemle* (Budapest), quarterly, with variations. Table of contents, and summaries, in English, German, and Russian. Organ of the Magyar Tudományos Akadémia Filozófiai Intézetének Folyóirata.

Of the Slavic languages, Polish, Serbo-Croatian, and Russian are represented once each in the new philosophical journals of the 1950s.

1951–. *Studia Filozoficzne; Kwartalnik* (Warsaw), quarterly, with variations. Title was *Myśl Filozoficzna* from 1951 to 1955 (issued six times a year). Sponsored from 1952 to 1955 by the Komitet Filozoficzny, Polska Akademia Nauk. Suspended 1956. Published by the Instytut Filozofii i Socjologii, Polska Akademia Nauk. Table of contents and summaries of articles in English and Russian. This periodical replaced *Przegląd Filozoficzny* (1897–1949) at the beginning of the period of militant Marxist domination. According to an article in a 1963 issue of *Studia Filozoficzne*, it was Lenin who first solved Zeno's antinomy of the arrow in flight.

1953–1958. *Filozofski Pregled* (Belgrade), issued three times a year, with variations.

1958–. *Nauchnye Doklady Vysshei Shkoly; Filosofskie Nauki* (Moscow), issued six times a year; originally issued quarterly. Often cited as *Filosofskie Nauki*, without the series title ("Scientific Reports of the Higher School") represented by the first four words.

JAPANESE AND CHINESE. Also begun in the 1950s were *Bigaku; Aesthetics* (Tokyo, 1950–), issued quarterly, and *Chê Hsüeh Yen Chiu* ["Philosophical Research"] (Peking, 1955–), issued six times a year; formerly quarterly.

THE NINETEEN-SIXTIES

The early years of the 1960s were fruitful in the production of new journals of philosophy, but not as fruitful as the record year of 1946 (11 journals). The year 1960 brought forward 9; 1961, 4; 1962, 6; 1963, 5; 1964, 3; 1965, 6; and 1966, 1 (as of the time of the completion of this article).

1960. Three philosophical journals which were started in 1960 had their origin in England:

1960–. *The Heythrop Journal; A Quarterly Review of Philosophy and Theology* (Oxford), quarterly. Issued by the Jesuit Faculties of Philosophy and Theology, Heythrop College, Oxford.

1960–. *The British Journal of Aesthetics* (London), quarterly. Published for the British Society of Aesthetics.

1960–. *Philosophical Books* (Leicester, England), originally a quarterly; now issued three times a year.

Three had their origin in the United States:

1960–. *Notre Dame Journal of Formal Logic* (Notre Dame, Ind.), quarterly.

1960–. *Studies in Philosophy and Education* (Toledo, Ohio; previously New Brunswick, N.J.), quarterly, with variations.

1960–. *Journal of Existentialism* (New York), quarterly. Title was *Journal of Existential Psychiatry* from 1960 to 1964.

Amsterdam, Madrid, and Rome fathered one philosophical journal each in 1960:

1960–. *Wijsgerig Perspectief op Maatschappij en Wetenschap* (Amsterdam), issued six times a year. Each issue devoted to a specific topic.

1960–. *Noesis; Revista de filosofía y arte* (Madrid), quarterly. Suspended 1962–1963. *Noesis* had previously been the title of a philosophical journal in Italy in 1946.

1960–. *Filosofia e vita; Quaderni trimestrali de orientamento formativo* (Turin; previously Rome), quarterly.

1961. Two journals of philosophy were inaugurated in the United States, and one each in India and the Netherlands, in 1961:

1961–. *Journal for the Scientific Study of Religion* (New Haven), semiannual.

1961–. *International Philosophical Quarterly* (New York and Heverlee–Louvain), quarterly. Edited by the department of philosophy of Fordham University and the professors of philosophy, Berchmans Philosophicum, Heverlee, Belgium.

1961–. *Darshana* (Moradabad, India), quarterly. International board of consultants.

1961–. *Studies in Soviet Thought* (Dordrecht), quarterly. Published by the Institute of East-European Studies, University of Fribourg, Switzerland. Articles in English, French, and German.

1962. Two more journals were inaugurated in the United States, and one each in Argentina, Canada, Italy, and Australia, in 1962:

1962–. *Pacific Philosophical Forum* (Stockton, Calif.), quarterly. Each issue devoted to a specific subject, with a set format (thesis and countertheses).

1962–. *Soviet Studies in Philosophy* (New York), quarterly. Contains translations from Soviet publications, mainly Soviet periodicals.

1962–. *Cuestiones de filosofía* (Buenos Aires), quarterly.

1962–. *Dialogue; Canadian Philosophical Review; Revue canadienne de philosophie* (Montreal), quarterly. Articles in English and French. Sponsored by the Canadian Philosophical Association.

1962–. *De Homine* (Rome), quarterly. Issued by the Centro di Ricerca per le Scienze Morali e Sociali, Istituto di Filosofia, Università di Roma.

1962–. *Sophia; A Journal for Discussion in Philosophical Theology* (Melbourne), issued three times a year. An Italian *Sophia* began publication in 1933.

1963. As in 1961 and 1962, two journals of philosophy were inaugurated in the United States in 1963; in addition, two were inaugurated in India and one in the Netherlands:

1963–. *Southern Journal of Philosophy* (Memphis, Tenn.), quarterly.

1963–. *Journal of the History of Philosophy* (Berkeley), semiannual.

1963–. *Indian Journal of Philosophic Studies* (Hyderabad), semiannual. Published for the Andhra Pradesh Philosophical Society by the department of philosophy of Osmanian University, Hyderabad.

1963–. *Research Journal of Philosophy and Social Sciences* (Meerut, Uttar Pradesh, India), semiannual, with variations. International editorial board. Each issue contains about 200 pages on a particular subject.

1963–. *Vivarium; A Journal for Mediaeval Philosophy and the Intellectual Life of the Middle Ages* (Assen), semiannual.

1964. Three new contributions appeared in the year 1964:

1964–. *American Philosophical Quarterly* (Pittsburgh, Pa.), quarterly. International board of consultants. Articles only; no book reviews.

1964–. *The Philosophical Journal* (Edinburgh), semiannual. Issued by the Royal Philosophical Society of Glasgow. Although mainly concerned with scientific matters, the *Journal* also contains some valuable philosophical articles.

1964–. *Documentación crítica iberoamericana de filosofía y ciencias afines* (Seville), quarterly.

1965. The following journals began publication in 1965:

1965–. *Concilium; An International Review of Theology* (London), issued ten times a year.

1965–. *Foundations of Language; International Journal of Language and Philosophy* (Dordrecht, Netherlands), issued quarterly.

1965–. *Information aus dem philosophischen Leben der Deutschen Demokratischen Republik* (East Berlin), issued quarterly.

1965–. *Religious Studies* (London), semiannual. Articles on philosophy of religion and history of religion.

1965–. *Transactions of the Charles S. Peirce Society* (Amherst, Mass.), semiannual.

1965–. *Revue universitaire de science morale* (Geneva), issued three times a year.

1966. One philosophical journal began publication in 1966 before the present article was completed:

1966–. *The Bulletin of Philosophy* (Washington), issued eight times a year. Contains news of interest to philosophers.

The expansion in the twentieth century of the number of currently published journals of philosophy has roughly paralleled the growing interest in philosophy as an academic discipline.

Bibliography

Four authors have studied philosophical journals in general. In chronological order, their reports on this field are Friedrich Medebach, "Die philosophische Fachzeitschrift," in *Zeitungswissenschaft*, Vol. II (Berlin, 1936), 210–214; David Baumgart, *Philosophical Periodicals; An Annotated World List* (Washington, 1952); Augusto da Silva, *Revistas de filosofia* (Braga, Portugal, 1955); and Tóth Ilona Kovácsné, "A Magyar Közkönyvtarakban Megtaláható Kurrens Filozófia Periodikák" ("Current Philosophical Periodicals Available in Public Libraries in Hungary"), in *Magyar Filózofiai Szemle,* Vol. 8 (Budapest, 1964), 574–601.

Three sources cover philosophical journals in particular countries: Paul Feldskeller, "Das philosophische Journal in Deutschland," in *Reichls philosophischer Almanach* (Darmstadt, 1924), pp. 302–458; Enrico Zampetti, *Bibliografia ragionata delle riviste filosofiche italiana del 1900* (Rome, 1956), and the highly knowledgeable passages on periodicals in Max Rieser, "Polish Philosophy Today," in *Journal of the History of Ideas*, Vol. 24 (1963), 423–432.

Short articles on 83 periodicals appear in the *Enciclopedia filosofica*, 4 vols. (Venice and Rome, 1957). Other pertinent sources are the list, published annually in the *Répertoire bibliographique de la philosophie*, of periodicals covered by the *Répertoire*, and the list headed "Philosophy" in *Ulrich's Periodicals Directory* (New York, 1932; 10th ed., 1963).

The titles appearing in the philosophy category in the monthly *New Serial Titles; Classed Subject Arrangement* (Washington, 1955–) are useful, as is also the record of births and deaths of periodicals (as well as of libraries which possess complete or partial sets of the periodicals) in the *Union List of Serials in Libraries of the United States and Canada*, edited by Edna Brown Titus, 3rd ed., 5 vols. (New York, 1965), continued as *New Serial Titles*, which was begun by the Library of Congress in 1953 and is published monthly.

William Gerber (1967)

PHILOSOPHY JOURNALS [ADDENDUM]

The English-language journal citations in this update are organized chronologically by year of first publication, and appear within the individual year listings alphabetically by journal title. The non-English-language journal citations directly follow the English-language list; they too are organized chronologically by year of first publication, and appear within individual year listings alphabetically by journal title.

ENGLISH LANGUAGE

1965

Foundations of Language. Dordrecht, The Netherlands; Boston: D. Reidel. 1965–1976 bimonthly (formerly quarterly).

International Directory of Philosophy and Philosophers. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1965–.

Transactions of the Charles S. Peirce Society. Amherst, MA: University of Massachusetts Press. 1965– quarterly.

1966

Apeiron. University of Alberta, Dept. of Classics and Monash University, Dept. of Classical Studies. Edmonton: Academic Print. & Pub. 1966–.

The Teilhard Review. Pierre Teilhard de Chardin Association of Great Britain and Ireland (London). Teilhard Centre for the Future of Man. 1966–1981.

1967

Conceptus. Innsbruck: J. Zelger. 1967–.

Noûs. Wayne State University, Dept. of Philosophy. Detroit: Wayne State University Press. 1967– quarterly.

The Philosopher's Index. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1967– quarterly.

Royal Institute of Philosophy Lectures. Royal Institute of Philosophy. London: Macmillan. New York: St. Martin's. 1967–1990; semiannual, 1987–1990.

1968

American Philosophical Quarterly. Monograph Series. University of Pittsburgh, Dept. of Philosophy. Oxford: Basil Blackwell. 1968–1978 irregular.

Kinesis. Carbondale, IL: Southern Illinois University, Dept. of Philosophy. 1968– semiannual.

Man and World. State College, PA: I.P.R. Associates.1968–1997 quarterly.

The Philosopher's Index. Richard H. Lineback. Bowling Green, OH: Philosopher's Information Center. 1968– cumulative ed.

Philosophy and History. Tubingen, Germany: Institut fur wissenschaftliche Zusammenarbeit. 1968–1991 semiannual.

Philosophy and Rhetoric. University Park, PA: Pennsylvania State University Press. 1968– quarterly.

The Philosophy Forum. DeKalb, IL: Northern Illinois University. 1968–1980 quarterly.

1969

Chinese Studies in Philosophy. White Plains, NY: International Arts and Sciences Press, M. E. Sharpe, Inc. 1969–1997.

The Journal of Critical Analysis. Bemidiji, MN: National Council of Teachers for Critical Analysis. 1969– quarterly.

The Owl of Minerva: Quarterly Journal of the Hegel Society of America. Villanova, PA: Hegel Society of America. Villanova University, Philosophy Dept., and Florida State University. 1969– semiannual.

Studies in Philosophical Linguistics. William L. Todd. Evanston, IL: Great Expectations. 1969.

The Undergraduate Journal of Philosophy. Oberlin, OH: Oberlin College, Philosophy Dept. 1969–1977.

1970

Auslegung. Lawrence, KS: Dept. of Philosophy, University of Kansas. 1970s– semiannual.

International Journal for Philosophy of Religion. Dordrecht, The Netherlands; Boston: Martinus Nijhoff. 1970– four issues per year; 1983– six issues per year.

The Journal of Philosophical Linguistics. William Todd. Evanston, IL: Great Expectations. 1970–1971.

Metaphilosophy. Oxford: Basil Blackwell for the Metaphilosophy Foundation. 1970–1999; 2000– five issues per year.

Philosophic Exchange. Brockport, NY: State University of New York College at Brockport, Center for Philosophic Exchange, College of Arts and Science. 1970–.

PSA; Proceedings of the Biennial Meeting of the Philosophy of Science Association. East Lansing, MI: Philosophy of Science Association. 1970–1994 biennial.

The Southwestern Journal of Philosophy. Norman, OK: Southwestern Philosophical Society. 1970–1980.

Studies in History and Philosophy of Science. Oxford; New York: Pergamon Press. 1970– quarterly since 1995.

Theory and Decision. Dordrecht, The Netherlands: D. Reidel. 1970– eight issues per year.

1971

Canadian Journal of Philosophy. Edmonton: Canadian Association for Publishing in Philosophy. 1971–.

Graduate Faculty Philosophy Journal. New York: New School for Social Research, Philosophy Dept. 1971– semiannual.

Idealistic Studies. Worcester, MA: Clark University Press. 1971– three issues per year.

1972

Aitia. Farmingdale, NY: State University of New York at Farmingdale. 1972– 1992 three issues per year.

Journal of Philosophical Logic. Association for Symbolic Logic. Dordrecht, The Netherlands; Boston: Kluwer Academic. 1972– bimonthly.

Paideia. Buffalo, NY: State University College at Buffalo; University of New York College at Brockport. 1972.

Philosophical Linguistics. William Todd. Evanston, IL: Great Expectations. 1972–1973.

Philosophical Papers. Dept. of Philosophy, Rhodes University, and University of the Witwatersrand. 1972– three issues per year.

Philosophy in Context. Cleveland, OH: Cleveland State University, Dept. of Philosophy. 1972–1990 annual.

Radical Philosophy. Radical Philosophy Group (Great Britain). Canterbury: Radical Philosophy Group. 1972– bimonthly.

Second Order. Ile-Ife, Nigeria: University of Ife Press. 1972— semiannual.

Thêta-pi. Leiden, The Netherlands: E. J. Brill. 1972–1974 semiannual.

1973

CIRPHO. Montreal: International Society for Computer Research in Philosophy. 1973–1976.

Gnosis. Montreal: Sir George Williams University, Dept. of Philosophy. 1973–.

Indian Philosophical Quarterly. Pratap Centre of Philosophy. Amalner, India: University of Poona, Dept. of Philosophy. 1973– quarterly.

Radical Philosophers' Newsjournal. Somerville, MA: Radical Philosophers' Newsjournal. 1973–1990s.

Revolutionary World. Amsterdam: B. R. Grüner Pub. Co. 1973–1982 five issues per year.

1974

Indian Journal of Philosophic Studies. Osmania University, Dept. of Philosophy. Hyderabad: Osmania University 1974–.

International Studies in Philosophy. State University of New York at Binghamton. Torino, Italy: Filosofia. 1974–1979 annual, 1980–1981 semiannual, 1982– three issues per year.

Journal of the Philosophy of Sport. Philosophic Society for the Study of Sport, and the International Association for the Philosophy of Sport. Champaign, IL: Human Kinetics Publishers. 1974–2000 annual, 2001– semiannual.

Lias. Amsterdam: Holland University Press. 1974– two issues per year.

1975

Canadian Journal of Philosophy. Supp. vol. Guelph, ON: Canadian Association for Pub. in Philosophy. 1975–.

Journal of the Department of Philosophy. University of Calcutta, Dept. of Philosophy. Calcutta: University of Calcutta. 1975–.

Philosophy and Medicine. Spicker, Stuart F., and H. Tristram Engelhardt. Dordrecht, The Netherlands; Boston: Reidel. 1975–.

Philosophy Research Archives. American Philosophical Association, and Canadian Philosophical Association. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1975–1981 annual.

Poznan Studies in the Philosophy of the Sciences and the Humanities. Amsterdam: Rodopi. 1975– quarterly.

Teaching Philosophy. Cincinnati, OH: [s.n.]. 1975– quarterly, formerly semiannual.

1976

Midwest Studies in Philosophy. Morris, MN: University of Minnesota, Morris. 1976– annual.

Philosophical Studies in Education. Ohio Valley Philosophy of Education Society. Terre Haute, IN: School of Education, Indiana State University. 1976– annual.

Philosophy and Literature. University of Michigan, Dearborn; Whitman College. Baltimore, MD: Johns Hopkins University Press. 1976–.

Southwest Philosophical Studies. Lubbock, TX: Texas Tech Press. 1976 annual, 1982 Spring, 1988 three issues per year, 1988– annual.

1977

Aletheia. Irving, TX: International Academy of Philosophy Press. 1977– irregular.

A Directory of Women in Philosophy. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1977–1982.

The Independent Journal of Philosophy. Vienna: G. E. Tucker. 1977–.

Linguistics and Philosophy. Dordrecht, The Netherlands: Kluwer Academic. 1977– bimonthly.

Reports on Philosophy. Uniwersytet Jagiellonski. Warsaw: Polish Scientific Publishers. 1977–.

Review Journal of Philosophy & Social Science. Meerut, India: Anu Prakashan. 1977–.

Towards. Northridge, CA: C. Monks. 1977– semiannual.

1978

Bulletin of the Evangelical Philosophical Society. San Bernadino, CA: Evangelical Philosophical Society. 1978–.

Eidos. Waterloo, Ontario: University of Waterloo, Philosophy Graduate Student Association. 1978–.

Human Studies. Society for Phenomenology and the Human Sciences. Dordrecht, The Netherlands: Martinus Nijhoff. 1978– quarterly.

Lonergan Workshop: Collected Essays. Fred Lawrence. Missoula, MT: Scholars Press for Longergan Workshop. 1978–.

Milltown Studies. Dublin: Milltown Institute of Theology and Philosophy. 1978– semiannual.

The Philosopher's Annual. Center for the Study of Language and Information (U.S.). Totowa, NJ: Rowman and Littlefield. 1978– annual.

Philosophical Inquiry. Aristoteleio Panepistemio Thessalonikes. Athens: [s.n.]. 1978– quarterly.

Philosophical Investigations. Oxford: Basil Blackwell. 1978– quarterly.

Philosophy & Social Criticism. Chestnut Hill, MA: [s.n.]. 1978–1994 quarterly, 1995–2003 bimonthly, 2004– seven issues per year.

Research in Philosophy & Technology. Society for Philosophy & Technology (U.S.). Greenwich, CT: JAI Press. 1978–.

1979

The Etienne Gilson Series. Pontifical Institute of Mediaeval Studies. Toronto: Pontifical Institute of Mediaeval Studies. 1979–.

Nature and System. Tucson, AZ: Nature and System. 1979–.

1980

American Journal of Theology and Philosophy. American Society for Social Philosophy and Philosophical Theology. West Lafayette, IN: American Journal of Theology and Philosophy. 1980–.

Analytic Teaching. Texas Wesleyan College and Viterbo College. Fort Worth: Texas Wesleyan College. 1980s–.

Concordia. Valencia: Concordia. 1980s– two issues per year.

The Journal of Mind and Behavior. Institute of Mind and Behavior. New York: The Journal. 1980– quarterly.

Logos: Philosophic Issues in Christian Perspective. Santa Clara, CA: University of Santa Clara, Dept. of Philosophy. 1980– annual.

The Objective Forum. New York: TOF Publications. 1980–1987 bimonthly.

Pacific Philosophical Quarterly. Los Angeles: School of Philosophy, University of Southern California. 1980– quarterly.

Southwest Philosophy Review: Papers Presented at the Annual Meeting of the Southwestern Philosophical Society. Southwestern Philosophical Society. Conway, AR: The Society. 1980s– two issues per year.

1981

Philosophical Topics. Southwestern Philosophical Society, and University of Arkansas, Fayetteville, Dept. of Philosophy. Denver, CO: Philosophical Topics. 1981– two issues per year, 1986– three issues per year plus two supplements.

World Futures. New York: Gordon and Breach. 1981– semiannual.

1982

Contemporary German Philosophy. University Park: Penyslvania State University Press. 1982–1984 annual.

Law and Philosophy. Dordrecht, The Netherlands; Boston: D. Reidel. 1982– six issues per year.

Monograph Series. Bundoora, Victoria: Australasian Association of Philosophy. 1982– semiannual.

Philosophy Research Archives: PRA. American Philosophical Association and Canadian Philosophical Association. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1982–1989 annual.

South African Journal of Philosophy. Foundation for Education, Science, and Technology, South Africa. Pretoria: Bureau for Scientific Publications of the Foundation for Education, Science, and Technology. 1982–.

Spindel Conference: Proceedings. Memphis: Dept. of Philosophy, Memphis State University. 1982– annual.

The Teilhard Review and Journal of Creative Evolution. Teilhard Centre for the Future of Man. London: Teilhard Centre. 1982–1988 three issues per year.

The Thoreau Quarterly. Minneapolis: The Thoreau Quarterly. 1982–1985 quarterly.

Topoi. Dordrecht, The Netherlands; Boston: D. Reidel. 1982– semiannual.

1983

Afro American Journal of Philosophy: AAJP. New York: Afro American Philosophy Association. 1983– quarterly.

Cognito: An International Journal for Philosophy, Society, and Politics. Quezon City, Philippines: Cogito. 1983– quarterly.

Hypatia: A Journal of Feminist Philosophy. Oxford; New York: Pergamon. 1983–.

Krisis. International Circle for Research in Philosophy. Houston: International Circle for Research in Philosophy. 1983– two issues per year.

Method. Loyola Marymount University, Boston College, Lonergan Institute, and Institute for Integrative Studies in Los Angeles. Los Angeles: Method. 1983– semiannual.

Philosophy in Science. Specola Vaticana, Center for Interdisciplinary Studies, Papieska Akademia Teologiczna (Krakow, Poland), and Pachart Foundation. Tucson: Pachart Pub. House. 1983– irregular.

Social Philosophy and Policy. Bowling Green State University, Social Philosophy and Policy Center. Oxford: Basil Blackwell, 1983– semiannual.

Theoretical Medicine. Dordrecht, The Netherlands; Boston: D. Reidel. 1983–1997 four issues per year.

1984

Explorations in Knowledge. Eastleigh, U.K.: Sombourne Press. 1984–1997 semiannual.

Faith and Philosophy: Journal of the Society of Christian Philosophers. Society of Christian Philosophers. Wilmore, KY: The Society, 1984– quarterly.

Grantees' Reports. American Philosophical Society. Philadelphia: The Society. 1984–1986 annual.

History of Philosophy Quarterly: HPQ. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1984– quarterly.

Irish Philosophical Journal. Belfast: Queen's University of Belfast, Dept. of Scholastic Philosophy. 1984– semiannual.

Journal of Applied Philosophy. Society for Applied Philosophy. Abingdon, Oxfordshire, U.K.: Carfax Pub. Co. 1984– three issues per year.

Rabindra Bharati Journal of Philosophy. Rabindra Bharati University. Calcutta: Rabindra Bharati University. 1984–.

Research in Philosophy and Technology. Suppl. Society for Philosophy and Technology. Greenwich, CT: JAI Press. 1984–.

1985

CC AI. Communication and Cognition (Firm). Ghent: Communication and Cognition. 1985– quarterly.

Economics and Philosophy. Cambridge, U.K.: Cambridge University Press. 1985– semiannual.

1986

Biology and Philosophy. Dordrecht, The Netherlands; Boston: D. Reidel. 1986– five issues per year.

Current Philosophy. Randolph, MA: Honor Publications. 1986– bimonthly.

Philosophy, Theology. Milwaukee: Marquette University Press. 1986– semiannual. Formerly *International Studies in the Philosophy of ISPS.* London: Routledge and Kegan Paul.

1987

African Philosophical Inquiry. African Society for Philosophical Research. University of Ibadan, Dept. of Philosophy. Ibadan: Ibadan University Press. 1987– annual, with each issue containing two numbers.

Cognito. Cognito Society; University of Bristol. Bristol, U.K.: The Society. 1987–1999 three issues per year.

Discussions in Contemporary Culture. Dia Art Foundation and Dia Center for the Arts. New York: Bay Press. 1987–1995 irregular.

Journal of Scientific Exploration: A Publication of the Society for Scientific Exploration. New York: Pergamon. 1987– quarterly.

Medical Humanities Review. University of Texas Medical Branch at Galveston; Institute for the Medical Humanities. Galveston, TX: The Institute. 1987– semiannual.

Philosophical Perspectives. James E. Tomberlin. Atascadero, CA: Ridgeview. 1987– annual.

Quest. University of Zambia, Dept. of Philosophy, Rijkuniversiteit te Groningen, Centre for Development Studies, and University of the North (South Africa). Lusaka, Zambia: Quest. 1987– semiannual.

Science in Context. Cambridge, U.K.: Cambridge University Press. 1987– quarterly.

1988

International Journal for the Semiotics of Law. Eric Landowski. International Association for the Semiotics of Law. Merseyside, U.K.: D. Charles Publications, and Holmes Beach, FL: Wm. W. Gaunt. 1988– four issues per year.

International Journal on the Unity of the Sciences. International Cultural Foundation and International Conference on the Unity of the Sciences. New York: The Conference. 1988–1992 quarterly.

Oxford Studies in Ancient Philosophy. Supp. vol. Oberlin Colloquium in Philosophy. Oxford: Clarendon Press, and New York: Oxford University Press. 1988–.

Philosophical Psychology. Abingdon, Oxfordshire, U.K.: Carfax Pub. Co. 1988– quarterly.

1989

Bridges. Columbia, MD: [s.n.]. 1989– semiannual.

Journal of Philosophy and the Visual Arts. London: Academy Editions, and New York: St. Martin's Press. 1989–.

The Teilhard Review. Teilhard Centre for the Future of Man. London: The Centre. 1989–1994 three issues per year.

The Westminster Tanner-McMurrin Lectures on History and Philosophy of Religion at Westminster. Salt Lake City: Westminster College of Salt Lake City. 1989–1992.

1990

American Catholic Philosophical Quarterly: Journal of the American Catholic Philosophical Association. Washington, DC: The Association. 1990– quarterly.

Behavior and Philosophy. Cambridge Center for Behavioral Studies. Cambridge, MA: The Center. 1990– semiannual.

Guide to Graduate Programs in Philosophy. American Philosophical Association. Newark, DE: The Association. 1990s– biennial.

Imodoye: A Journal of African Philosophy. Lagos, Nigeria: University of Lagos, Dept. of Philosophy, Faculty of Arts. 1990–.

Journal for General Philosophy of Science. Dordrecht, The Netherlands: Kluwer Academic. 1990– two issues per year.

Journal of Philosophical Research: JPR. Bowling Green State University, Philosophy Documentation Center; American Philosophical Association; Canadian Philosophical Association; University of Nebraska; and University of Iowa. Bowling Green, OH: Bowling Green State University, Philosophy Documentation Center. 1990– annual.

Pli, Warwick Journal of Philosophy. Coventry, U.K.: University of Warwick, Dept. of Philosophy. 1990– semiannual.

1991

Dialogue and Humanism: The Universalist Quarterly. Polska Akademia Nauk; International Society for Universalism. Warsaw: Warsaw University, World Order and Universalism Research Program. 1991–1993 quarterly, 1994 bimonthly.

The Harvard Review of Philosophy. Cambridge, MA: The Harvard Review of Philosophy (Organization). 1991– annual.

Philosophical Issues. Sociedad Filosofica Ibero Americana. Atascadero, CA: Ridgeview Pub. Co. 1991– annual.

Royal Institute of Philosophy. Suppl. Royal Institute of Philosophy. Cambridge, U.K.: Cambridge University Press. 1991– semiannual.

1992

Russian Studies in Philosophy. Armonk, NY: M. E. Sharpe. 1992– quarterly.

Skeptic. Skeptics Society. Altadena, CA: The Society. 1992– quarterly.

1993

The American Philosophical Association's Guide to Graduate Programs in Philosophy. Newark, DE: American Philosophical Association. 1993– annual.

Angelaki: A New Journal in Philosophy, Literature, and the Social Sciences. Oxford: Angelaki. 1993– three issues per year.

British Journal for the History of Philosophy (BJHP): The Journal of the British Society for the History of Philosophy. Bristol, U.K.: Thoemmes Press. 1993–2001 two or three issues per year; 2002– quarterly.

Epoche. Provo, UT: Brigham Young University, Dept. of Philosophy. 1993–1998 two issues per year.

European Journal of Philosophy. Oxford: Blackwell. 1993– three issues per year.

Humanities Aitia. State University of New York College of Technology at Farmingdale. Farmingdale, NY: SUNY. 1993– three issues per year.

IJPS. London: Routledge. 1993– semiannual.

The Journal of Political Philosophy. Cambridge, MA: Blackwell. 1993–.

Perspectives on Science: Historical, Philosophical, Social. Chicago: University of Chicago Press. 1993– quarterly.

Philosophia Mathematica. Philosophy of Mathematics, Its Learning, and Its Application. Canadian Society for the History and Philosophy of Mathematics. Toronto: University of Toronto Press. 1993– semiannual.

Routledge History of Philosophy. G. H. R. Parkinson and Stuart Shanker. London: Routledge. 1993–1999 irregular.

Studies in East European Thought. Dordrecht, The Netherlands: Kluwer Academic. 1993– quarterly.

Tekhnema: Journal of Philosophy and Technology. American University of Paris. Paris: The University. 1993– annual.

1994

European Review of Philosophy. Stanford: CSLI Publications, Center for the Study of Language and Information. 1994–.

Film and Philosophy. Society for the Philosophic Study of the Contemporary Visual Arts. Portsmouth, OH: The Society. 1994– annual.

Philosophia Christi: Journal of the Evangelical Philosophical Society. Orlando: The Society. 1994– annual.

Philosophy in the Contemporary World: An International Journal Sponsored by the Society for Philosophy in the Contemporary World. Morehead, KY: The Society. 1994– quarterly.

Philosophy, Psychiatry and Psychology: PPP. Baltimore: Johns Hopkins University Press. 1994– quarterly.

1995

Dialogue and Universalism. Warsaw: Warsaw University, Centre of Universalism. 1995– monthly.

Journal of Philosophy and Development. Ago-Iwoye, Nigeria: Ogun State University, Dept. of Philosophy. 1995–.

Making the Rounds in Health, Faith, and Ethics. Chicago: Park Ridge Center. 1995–1996 semimonthly, except July and August.

Philosophy of Education. Urbana, IL: Philosophy of Education Society. 1995– annual.

Res Publica: A Journal of Legal and Social Philosophy. Association for Legal and Social Philosophy (Great Britain). Liverpool: D. Charles Publications. 1995– semiannual.

Studies in History and Philosophy of Science. Studies in History and Philosophy of Modern Physics. Oxford: Pergamon. 1995– three issues per year (April, August, December).

1996

Archimedes: New Studies in the History and Philosophy of Science and Technology. Dordrecht, The Netherlands: Kluwer Academic. 1996– annual.

Ethics and the Environment. Greenwich, CT: JAI Press. 1996– semiannual.

Philosophical Writings. Durham, U.K.: University of Durham, Dept. of Philosophy. 1996– three issues per year.

Studies in History and Philosophy of Modern Physics. Oxford: Pergamon. 1996– quarterly.

1997

Contemporary Chinese Thought: Translations and Studies. M. E. Sharpe, Inc. Armonk, NY: Sharpe. 1997– quarterly.

Crossings. Binghamton, NY: State University of New York at Binghamton. 1997– semiannual.

The Philosophers' Magazine. London: [s.n.]. 1997– quarterly.

Philosophy and Geography. Society for Philosophy and Geography. Lanham, MD: Rowman and Littlefield. 1997–.

Philosophy in Review. Edmonton: Academic Print and Pub. 1997– six issues per year.

1998

APA Newsletters. Newark, DE: American Philosophical Association. 1998– semiannual.

Biomedical Sciences. Oxford: Pergamon. 1998– four issues per year.

Continental Philosophy Review. Dordrecht, The Netherlands; Boston: Kluwer Academic. 1998– quarterly.

Critical Review of International Social and Political Philosophy. London: Frank Cass. 1998–.

Theoretical Medicine and Bioethics. Dordrecht, The Netherlands: Kluwer Academic. 1998– six issues per year.

1999

Aristoi: An Interdisciplinary Journal of Philosophy. Broome Community College, Division of Liberal Arts. Binghamton, NY: The Division. 1999–.

Foundations of Chemistry. Dordrecht, The Netherlands: Kluwer Academic. 1999– eight issues per year.

2002

Epoche. Villanova, PA: Villanova University, Philosophy Dept. 2002– two issues per year.

Politics, Philosophy, and Economics: PPE. Gerald F. Gaus and Jonathan Riley. Murphy Institute of Political Economy. London: Sage. 2002– three issues per year.

Think: A Periodical of the Royal Institute of Philosophy. London: The Royal Institute of Philosophy. 2002– three issues per year.

2004

Journal of Moral Philosophy. London: Continuum. 2004– three issues per year.

Oxford Studies in Metaphysics. Oxford: Clarendon Press. 2004– irregular.

NON-ENGLISH LANGUAGE

1965

Cuyo. Mendoza, Argentina: Universidad Nacional de Cuyo, Instituto de Filosofía, Sección de Historia del Pensamiento Argentino. 1965–1983 annual.

Al-Fikr al-muʿasir. Cairo: al-Dar al-Misriyah al-ʿAmmah lil-Taʾlif wa-al-Nashr. 1965–[?] monthly.

Praxis. Zagreb, Croatia: Hrvatsko filosofsko drustvo. 1965–1974 quarterly.

Stromata. Colegio Máximo de San José, Facultad de Filosofía, Facultad de Teología; Universidad del Salvador, Facultad de Filosofía, Facultad de Teología. San Miguel, Argentina: Universidad del Salvador, Filosofía y Teología. 1965–.

Studia metodologiczne. Uniwersytet Poznanski. Poznań, Poland: Uniwersytet im. Adama Mickiewicza. 1965–.

Studia philosophiae christianae. Warsaw: Akademia Teologii Katolickiej. 1965– semiannual.

Voprosy filosofii i psikhologii. Leningradskii gosudarstvennyi universitet, Filosofskii fakul'tet, Fakul'tet psikhologii. [Leningrad]: Izd-vo Leningradskogo universiteta. 1965–[?].

1966

A Magyar Tudományos Akadémia Filozófiai és Történettudományi Osztályának közleményei. Magyar Tudományos Akadémia, Filozófiai és Történettudományok Osztálya. Budapest: Akadémiai Kiadó. 1966–1973.

Anales del Seminario de Metafísica. Universidad Complutense de Madrid, Seminario de Metafísica. Madrid: El Seminario. 1966–1997 annual.

Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio philosophica et sociologica. Eötvös Loránd Tudományegyetem. Budapest: Universita. 1966–1992 irregular (formerly annual).

Azarbaijan SSR Elmlar Akademiiasynyn khabarlari. Izvestiia Akademii nauk Azerbaidzhanskoi SSR. Seriia istorii, filosofii i prava. Tarikh, falsafa va hugug seriiasy. Baku, Azerbaijan: Azarbaijan SSR Elmlar Akademiiasy Nashriiiaty. 1966–1990 quarterly.

Cahiers pour l'analyse. Paris: Le Cercle d'épistémologie de l'école normale supérieure. 1966–1969 irregular.

Dijalektika. Belgrade: Univerzitet u Beogradu. 1966–.

Filozofia. Slovenská akadémia vied, Filozofický ústav, ústav filozofie a sociológie. Bratislava, Slovakia: Vydavatelstvo Slovenskej akadémie vied. 1966–1991 six issues per year, 1992– ten issues per year.

Problemy filosofiï: mizhvidomchyi naukovyi zbirnyk. Kyïvs'kyi derzhavnyi universytet im. T. H. Shevchenka. Kiev, Ukraine: Vyd-vo Kyivs'kogo universytetu. 1966– three issues per year.

Raison présente. Paris: Editions Rationalistes. 1966–.

Rassegna bibliografica di storia della filosofia. Università di Parma, Istituto di filosofia. Padua: Liviana editrice. 1966–[?] annual.

Studia Universitatis Babes-Bolyai. Series Philosophia. Cluj-Napoca, Romania: [Universitatea Babes-Bolyai]. 1966–1974.

Theologie und Philosophie. Freiburg im Breisgau, Germany: Herder. 1966–[?] four issues per year.

Vestnik Moskovskogo universiteta. Seriia VIII: Filosofiia. Moskovskii gosudarstvennyi universitet im. M. V. Lomonosova. [Moscow]: Izd-vo Moskovskogo universiteta. 1966–1976 bimonthly.

1967

Bibliographie Philosophie. Berlin: Akademie für Gesellschaftswissenschaften beim ZK der SED, Institut für Marxistisch-leninistische Philosophie, Zentralstelle für philosophische Information und Dokumentation. 1967–1987 quarterly.

Bibliographie Philosophie. Beiheft. Berlin: Institut für Gesellschaftswissenschaften, Zentralstelle für philosophische Information und Documentation. 1967–.

Crítica: revista hispanoamericana de filosofía. Mexico: Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México. 1967–.

Études philosophiques et littéraires. Jamʿiyat al-Falsafah bi-al-Maghrib. Casablanca: Dar el Kitab. 1967–1981 irregular (formerly two issues per year).

Filosofija. Jugoslovensko udruzenje za filozofiju, Filozofsko drustvo Srbije. [Belgrade]. 1967–1900s quarterly.

1968

L'age de la science. Paris: Dunod. 1968–1970 quarterly.

Anuario filosófico. Universidad de Navarra, Facultad de Filosofía y Letras. Pamplona, Spain: Universidad de Navarra. 1968–.

Bibliografía filosófica mexicana. [Mexico, D.F.]: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas and Instituto de Investigaciones Filosóficas. 1968– annual.

Estudio agustiniano. Valladolid, Spain: Estudio teológico agustiniano. 1968– three issues per year.

Gnozis [Gnosis]. New York. 1968–.

Problemos. Lithuania, Aukstojo ir specialiojo vidurinio mokslo ministerija. Vilnius, Lithuania: Mintis. 1968–[?] two issues per year.

Science et esprit. Jesuits, Province du Canada français, Faculté de philosophie, Faculté de théologie. Montreal: Les Éditions Bellarmin. 1968–.

Wiener Jahrbuch für Philosophie. Vienna: W. Braumüller. 1968– annual.

1969

Analele Universitatii Bucuresti. Filozofie. Bucharest: Universitatea din Bucuresti. 1969–1973 semiannual.

Annales de l'Institut de Philosophie. Université libre de Bruxelles, Institut de Philosophie, Institut de Sociologie. [Brussels]: Editions de l'Institut de Sociologie. 1969–1978 annual.

Anthropos. Drustvo psihologov Slovenije, Slovensko filozofsko drustvo. Ljubljana, Slovenia. 1969–.

Contributi dell'Istituto di filosofia. Università cattolica del Sacro Cuore, Istituto di filosofia. Milan: Società Editrice Vita e Pensiero. 1969–[?] irregular.

Eidos; revista de filosofía. Universidad Nacional de Córdoba, Instituto de Filosofía. Córdoba, Argentina: Instituto de Filosofía. 1969–.

Filosofskie voprosy logicheskogo analiza nauchnogo znaniia. P'ilisop'ayut'yan ev Iravunk'i Institut (Haykakan SSH Gitut'yunneri Akademia), Akademiia nauk Armianskoi SSR, Baku, Institut Filosofii i Prava. Yerevan, Armenia: IZD-vo AN Armianskoi SSR. 1969–.

Philosophische Perspektiven. Frankfurt am Main: V. Klostermann. 1969–1973 annual.

Studi internazionali di filosofia. Turin: Filosofia. 1969–1973 annual.

Voprosy filosofii i sotsiologii. Leningradskii gosudarstvennyi universitet imeni A. A. Zhdanova, Filosofskii fakul'tet. Saint Petersburg: Izd-vo Leningradskogo universiteta. 1969– annual.

1970

Actualidad bibliográfica de filosofía y teología. Barcelona: Facultades de Filosofía y Teología San Francisco de Borja. 1970– semiannual.

Algemeen Nederlands tijdschrift voor wijsbegeerte. [Assen, Netherlands]: Van Gorcum. 1970– quarterly.

Auslegung. Lawrence: University of Kansas, Department of Philosophy. 1970s– semiannual (formerly three issues per year).

Discurso. Universidade de São Paulo, Departamento de Filosofia. São Paulo, Brazil: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. 1970– irregular.

Zeitschrift für allgemeine Wissenschaftstheorie [Journal for General Philosophy of Science]. Wiesbaden, Germany: F. Steiner. 1970–1989.

1971

Bollettino del Centro di Studi Vichiani. [Naples: Centro di studi vichianii]. 1971–.

Filosofskie nauki. Qazaqtyng S.M. Kirov atyndaghy memlekettik universiteti. Almaty, Kazakhstan: Kazakhskii gos. universitet. 1971–.

Neue Hefte für Philosophie. Göttingen, Germany: Vandenhoeck & Ruprecht. 1971–1995 irregular.

Philosophia. Athens: Kentron Ereunes tes Hellenikes Philosophias. 1971– annual.

Philosophia. Universitat Bar-Ilan, Mahlakah le-filosofyah. Ramat-Gan, Israel: Bar-Ilan University. 1971– quarterly.

Philosophica Gandensia. Meppel, Netherlands: Boom Pers. 1971–1973.

Quellen und Studien zur Philosophie. Berlin and New York: de Gruyter. 1971– irregular.

Teorema. [Valencia, Spain]: Universidad de Valencia, Departamento de Lógica y Filosofía de la Ciencia, Departamento de Historia de la Filosofía. 1971–.

Zhe xue lun ping. Taipei: Guo li Taiwan da xue zhe xue xi. 1971–[?], 2004– semiannual.

1972

Aitia. Farmingdale: State University of New York at Farmingdale. 1972–1992 three issues per year.

Godishnik na Sofiiskiia universitet, Filosofski fakultet. Sofiiski universitet, Filosofski fakultet. Sofia, Bulgaria. 1972–1979 irregular.

Philosophie. Toulouse: Université de Toulouse-Le Mirail. 1972–1989 annual.

Prace filozoficzne. Uniwersytet Jagiellonski. Kraków: Panstwowe Wydawn. Nauk. 1972–1976 irregular.

Verifiche. [Trent, Italy]. 1972– four issues per year.

Zhe xue lun ji. Fu ren da xue (Hsin-chuang shih, Taiwan), zhe xue yan jiu suo. Taipei: Fu ren chu ban she. 1972–[?] semiannual.

1973

Beijing da xue xue bao: Zhe xue she hui ke xue ban. Beijing da xue, Qing hua da xue (Beijing, China). Beijing: Beijing ren min chu ban she. 1973– bimonthly.

Diotima. Hellenike Hetaireia Philosophikon Meleton. [Athens]: Ekdosis Hellenikes Hetaireias Philosophikon Meleton. 1973–.

Obshchestvennye nauki v SSSR. Seriia 3: Filosofskie nauki. Institut nauchnoi informatsii i fundamental'naia biblioteka po obshchestvennym naukam, Institut nauchnoi informatsii po obshchestvennym naukam (Akademiia nauk SSSR). Moscow: Akademiia nauk SSSR, In-t nauch. informatsii po obshchestvennym naukam. 1973–1976 quarterly, 1977–1991 six issues per year.

Protokoly … Vsemirnogo Kongressa po Filosofii [Proceedings of the … World Congress of Philosophy]. World Congress of Philosophy. Sofia, Bulgaria. 1973–[?] quinquennial.

Realitas. [Madrid]: Sociedad de Estudios y Publicaciones. 1973–.

Revista venezolana de filosofía. Universidad Simón Bolívar, Departamento de Filosofía, Sociedad Venezolana de Filosofía. [Caracas]: Universidad Simón Bolívar, Sociedad Venezolana de Filosofia. 1973– semiannual.

Sprawozdania-Poznanskie Towarzystwo Przyjaciól Nauk, Wydzial Filologiczno-Filozoficzny. Poznanskie Towarzystwo Przyjaciól Nauk, Wydzial Filologiczno-Filozoficzny. Poznań, Poland: Poznanskie Towarzystwo Przyjaciól Nauk. 1973–1993 annual.

1974

Analele Universitatii Bucuresti. Filosofie, istorie, drept. Universitatea din Bucuresti. [Bucharest: Tipografia Universitatii Bucuresti]. 1974–1976.

Análisis. Bogotá, Colombia: Universidad de Santo Tomás. 1974–[?] semiannual.

Argument-Sonderbände. Berlin: Argument-Verlag. 1974–.

Cuadernos salmantinos de filosofía. [Salamanca, Spain: Universidad Pontificia de Salamanca]. 1974–.

Filosofiia i nauchnyi kommunizm. Belaruski dziarzhauny universitet imia Ul. I. Lenina. Minsk: Izd-vo BGU. 1974–1989.

Philosophica. Rijksuniversiteit te Gent. [Ghent: Rijksuniversiteit]. 1974– two issues per year.

Philosophiques. [Montreal]: Bellarmin. 1974–.

Revista de filosofía. Universidad del Zulia, Centro de Estudios Filosóficos, Revista. Maracaibo, Venezuela: Universidad del Zulia, Facultad de Humanidades y Educación. 1974– three issues per year.

Transformação. Universidade Estadual Paulista, Faculdade de Filosofia, Ciências e Letras de Assis, Departamento de Filosofia. São Paulo, Brazil: Universidade Estadual Paulista. 1974–2002 annual, 2003– semiannual.

Zhe xue yu wen hua. Taipei: Zhe xue yu wen hua yue kan bian ji wei yuan hui. 1974– monthly.

1975

Erkenntnis. Dordrecht, Netherlands: Reidel. 1975– bimonthly (formerly three issues per year).

Grazer philosophische Studien. Universität Graz, Institut für Philosophie, Forschungsstelle für Österreichische Philosophie. Amsterdam: Rodopi. 1975–1976 annual, 1977–2002 at least two issues per year, 2003– annual.

Javidan-i khirad: nashriyah-i Anjuman-i Shahanshahi-i Falsafah-i Iran. Anjuman-i Shahanshahi-i Falsafah-ʾi Iran. Tehran: Anjuman. 1975–[?] semiannual.

Perspektiven der Philosophie. Amsterdam: Rodopi. 1975–.

Revista de filosofía latinoamericana. [San Antonio de Padua, Argentina: Ediciones Castañeda]. 1975–1979.

Revista latinoamericana de filosofía. [Buenos Aires, Argentina: Centro de Investigaciones Filosóficas]. 1975–.

Studia Universitatis Babes-Bolyai. Philosophia. Cluj-Napoca, Romania: [Universitatea Babes-Bolyai]. 1975–1976 annual, 1977–1982 two issues per year, 1983–1985 annual, 1986– two issues per year.

1976

Allgemeine Zeitschrift für Philosophie. Allgemeine Gesellschaft für Philosophie in Deutschland. [Stuttgart]: Frommann-Holzboog. 1976–.

Annales Universitatis Mariae Curie-Sklodowska. Sectio I, Philosophia-sociologia. Lublin, Poland: Nakl. Uniwersytetu Marii Curie-Sklodowskiej. 1976– annual.

Dialéctica. Puebla, Mexico: Universidad Autónoma de Puebla, Escuela de Filosofía y Letras. 1976– three issues per year.

Theoria: casopis Filozofskog drustva Srbije. Filozofsko drustvo Srbije. Belgrade: Drustvo. 1976– four issues per year.

1977

Analele Universitatii Bucuresti. Filosofie. Bucharest: Universitatea din Bucuresti. 1977– annual.

Kennis en methode. Amsterdam: Boom Meppel. 1977–1996.

Lituanistika v SSSR: Filosofiia i psikhologiia. Lietuvos TSR Mokslu akademija, Visuomenes mokslu informacijos sektorius, Visuomenes mokslu informacijos centras (Lietuvos TSR Mokslu akademija). Vilnius, Lithuania: Akademiia nauk Litovskoi SSR, In-t filosofii, sotsiologii i prava, Sektor nauch. informatsii po obshchestvennym naukam. 1977–[?].

Manuscrito. [Campinas, Brazil]: Universidade Estadual de Campinas, Centro de Lógica, Epistemologia e História da Ciência. 1977– semiannual.

Unabhängige Zeitschrift für Philosophie [The Independent Journal of Philosophy]. Vienna: [G. E. Tucker]. 1977–.

Vestnik Moskovskogo universiteta. Seriia VII: Filosofiia. Moskovskii gosudarstvennyi universitet im. M.V. Lomonosova. [Moscow]: Izd-vo Moskovskogo universiteta. 1977– bimonthly.

Zborník Filozofickej fakulty. Philosophica. Univerzita Komenského v Bratislave, Filozofická fakulta. Bratislava: Slovenské pedagogické nakl. 1977–, 1982– annual.

1978

Epistemologia. Genoa, Italy: Tilgher-Genova. 1978.

Escritos de filosofía. Academia Nacional de Ciencias de Buenos Aires, Centro de Estudios Filosóficos. Buenos Aires, Argentina: El Centro. 1978– semiannual.

Filosofia oggi. Bologna: Tip. editrice compositori. 1978– quarterly.

Humanitas. Instytut Filozofii i Socjologii (Polska Akademia Nauk). Wrocław, Poland: Zaklad Narodowy im. Ossolinskich. 1978–1989.

Studi filosofici. Naples: Istituto universitario orientale. 1978–.

Zhe xue yan jiu. Zhexue yanjiu. [Beijing: Ren min chu ban she]. 1978–1989 monthly, 1990 bimonthly, 1991– monthly.

1979

Annales de l'Institut de philosophie et de sciences morales. Université libre de Bruxelles, Institut de philosophie et de sciences morales. [Brussels]: Editions de l'Université de Bruxelles. 1979– annual.

Annali dell'Istituto di filosofia. Università di Firenze, Istituto di filosofia. Florence: L. Olschki. 1979–1984 annual.

1980

Actas. Fundación para el Estudio del Pensamiento Argentino e Iberoamericano; Jornadas del Pensamiento Filosófico Argentino. Buenos Aires, Argentina: Ediciones FEPAI. 1980s–[?].

Anhui da xue xue bao. Zhe xue she hui ke xue ban. Anhui da xue. [Hefei, China]: Anhui da xue xue bao bian ji wei yuan hui. 1980–1994 quarterly, 1995– bimonthly.

Annales de philosophie. Beirut: Université Saint-Joseph, Faculté des lettres et des sciences humaines. 1980–2001 annual.

Cahiers du Séminaire d'histoire des mathématiques. Paris: Université Pierre et Marie Curie, Laboratoire de mathématiques fondamentales, École pratique des hautes études, 1ère section, sciences mathématiques. 1980– annual.

Concordia. Valencia, Spain: Concordia. 1980s– two issues per year.

De philosophia. Ottawa, ON: University of Ottawa, Department of Philosophy, Student Association. 1980– annual.

Dialektik: Beiträge zu Philosophie und Wissenschaften. Universität Leipzig, Zentrum für Höhere Studien, Istituto italiano per gli studi filosofici. Cologne: Pahl-Rugenstein. 1980–1989 annual, 1991–1999 three issues per year, 2000– semiannual.

Ehu. Taipei: Ehu yue kan za zhi she. 1980s– monthly.

Godishnik na Sofiiskiia universitet "Kliment Okhridski," Filosofski fakultet. Sofia, Bulgaria: Universitetska pechatnitsa. 1980–1985 annual.

Slagmark. Århus, Denmark: Slagmark. 1980s–[?].

Vestnik. Institut za marksisticne studije (Slovenska akademija znanosti in umetnosti). Ljubljana, Slovenia: SAZU. 1980–1988 two issues per year.

1981

Acta Universitatis Lodziensis. Folia philosophica. Łódź, Poland: Uniwersytet Łódźki. 1981– irregular.

Agora. Universidad de Santiago de Compostela, Sección de Filosofía, Departamento de Filosofía e Antropoloxía Social, Departamento de Lóxica e Filosofía da Ciencia-Filosofía do Dereito, Moral e Política. Santiago de Compostela, Spain: Universidad de Santiago, Sección de Filosofía. 1981–1991 annual, 1992–[?] semiannual.

Análisis filosófico. Buenos Aires, Argentina: SADAF. 1981– semiannual.

Epistemens: revista del Institute de Filosofía. Universidad Central de Venezuela, Instituto de Filosofía, Facultad de Humanidades y Educación. [Caracas]: Ediciones de la Facultad de Humanidades y Educación. 1981–.

Recherches sur la philosophie et le langage: cahier du Groupe de recherches sur la philosophie et le langage. Université des sciences sociales de Grenoble, Groupe de recherches sur la philosophie et le langage. Grenoble: Institut de philosophie et sociologie. 1981–.

Revista de la Sociedad Argentina de Filosofía. Córdoba, Argentina: La Sociedad Argentina de Filosofía. 1981–.

Tetsugaku shiso ronshu. Ibaraki, Japan: Tsukuba Daigaku Tetsugaku Shiso Gakukei. 1981–.

Uchenye zapiski kafedr obshchestvennykh nauk vuzov Leningrada. Filosofskie i sotsiologicheskie issledovaniia. Russian S.F.S.R.; Ministerstvo vysshego i srednego spetsial'nogo obrazovaniia. Saint Petersburg: Izd-vo Leningradskogo universiteta. 1981–.

Xi bei da xue xue bao [Journal of Northwest University]. Xian Shi, China: Xi bei da xue xue bao bian ji bu. 1981– quarterly.

Zhe xue nian kan [Bulletin of the Association of Philosophy of Republic of China]. Zhongguo zhe xue hui. Taipei: Gai hui. 1981–[?] annual.

1982

Anuario de filosofía. San Miguel de Tucumán, Argentina: Universidad Nacional de Tucumán, Facultad de Filosofía y Letras, Departamento de Filosofía. 1982– annual.

Chuan xi lu. Dong wu da xue, Zhe xue xi. Taipei: Dong wu da xue. 1982–1988 annual.

Revista Universidad Pontificia Bolivariana. Medellín, Colombia: La Universidad Pontificia Bolivariana. 1982–1990.

Revue sénégalaise de philosophie. [Dakar, Senegal]: Nouvelles éditions africaines. 1982– semiannual.

Studi urbinati. B2, Filosofia, pedagogia, psicologia. Università di Urbino. Urbino: Università degli studi. 1982–1987.

Suid-Afrikaanse tydskrif vir wysbegeerte [South African Journal of Philosophy]. Foundation for Education, Science, and Technology, Bureau for Scientific Publications, Philosophical Society of Southern Africa. Pretoria: Foundation for Education, Science, and Technology, Bureau for Scientific Publications. 1982–.

Topoi. Dordrecht, The Netherlands: Reidel. 1982– semiannual.

Zhongguo zhe xue nian jian. Zhongguo she hui ke xue yuan, Zhe xue yan jiu suo. Shanghai: Zhongguo da bai ke quan shu chu ban she. 1982– annual.

1983

Bao kan zi liao suo yin. Di 1 fen ce, Zhe xue, she hui xue. [Beijing]: Zhongguo ren min da xue shu bao zi liao she. 1983– annual.

Beseda. Saint Petersburg. 1983–1993.

Cahiers du séminaire de philosophie. Université des sciences humaines de Strasbourg, Centre de documentation en histoire de la philosophie. [Strasbourg, France]: Centre de documentation en histoire de la philosophie. 1983–1996.

Ch'orhak sasang ui che munje. Kyŏnggi-do Songnam-si, Korea: Han'guk Chongsin Munhwa Yon'guwon. 1983–1986 annual.

Dilthey-Jahrbuch für Philosophie und Geschichte der Geisteswissenschaften. Göttingen, Germany: Vandenhoeck & Ruprecht. 1983–2000 annual.

Filosofiia i sotsiologiia nauki i tekhniki. Nauchnyi sovet po filosofskim i sotsial'nym problemam nauki i tekhniki (Akademiia nauk SSSR). Moscow: Nauka. 1983– annual.

Revue philosophique de Kinshasa. Kinshasa, Democratic Republic of the Congo: Faculté de théologie catholique de Kinshasa, Département de philosophie et religions africaines. 1983–[?] semiannual.

Yunnan min zu xue yuan xue bao [Journal of Yunnan University of the Nationalities]. Yunnan Sheng min zu shi wu wei yuan hui, Yunnan min zu xue yuan. Kunming, China: Yunnan min zu xue yuan xue bao (she zhe ban) bian ji bu. 1983– bimonthly.

1984

Acta Universitatis Wratislaviensis. Prace filozoficzne. Historia filozofii. Uniwersytet Wrocławski im. Boleslawa Bieruta. Wrocław, Poland: Wydawn. Uniwersytetu Wrocławskiego. 1984–1992 irregular.

Análise. Lisbon: GEC Publicacões. 1984– semiannual (formerly three issues per year).

Cuyo: anuario de filosofía argentina y americana. [Mendoza, Argentina]: Universidad Nacional de Cuyo, Facultad de Filosofía y Letras, Instituto de Filosofía Argentina y Americana. 1984–.

Darshan-manjari: The Burdwan University Journal of Philosophy. Barddhaman, India: University of Burdwan. 1984– annual.

Ezhegodnik Filosofskogo obshchestva SSSR. Filosofskoe obshchestvo SSSR. Moscow: Nauka. 1984–1990 annual.

Folia philosophica. Uniwersytet Slaski w Katowicach. Katowice, Poland: Uniwersytet Slaski. 1984– annual.

Karunungan [Sophia]. Philippine Academy of Philosophical Research. Manila: De La Salle University Press. 1984– annual.

Philosophie. Paris: Les Editions de Minuit. 1984– quarterly.

Prometeo: revista latinoamericana de filosofía. Guadalajara: Universidad de Guadalajara, Facultad de Filosofía y Letras. 1984–[?] three issues per year.

Rabindra Bharati Journal of Philosophy. Calcutta: Rabindra Bharati University. 1984–.

Rivista di storia della filosofia. Milan: Franco Angeli. 1984– quarterly.

Supplemente zu den Sitzungsberichten der Heidelberger Akademie der Wissenschaften, Philosophisch-Historische Klasse. Heidelberger Akademie der Wissenschaften, Philosophisch-Historische Klasse. Heidelberg, Germany: C. Winter. 1984–1996 irregular.

1985

Annali del Dipartimento di filosofia. Università di Firenze, Dipartimento di filosofia. Florence: L. Olschki. 1985– annual.

Annuario filosofico. Milan: Mursia. 1985– annual.

Archiwum historii i filozofii medycyny. Polskie Towarzystwo Historii Medycyny i Farmacji. Wrocław, Poland: Zaklad Narodowy im. Ossolinskich. 1985– quarterly.

Boletín mexicano de historia y filosofía de la medicina. [Mexico, D.F.: Sociedad Mexicana de Historia y Filosofía de la Medicina]. 1985– irregular.

Le cahier. Collège international de philosophie. Paris: Editions Osiris. 1985–1990 annual.

Ethernité. Paris: Editions de la Différence. 1985–.

Études maritainiennes [Maritain Studies]. Ottawa, Ontario: Association canadienne Jacques Maritain [Canadian Jacques Maritain Association]. 1985– annual.

Investigación humanística. Universidad Autónoma Metropolitana, Unidad Iztapalapa, Departamento de Filosofía, Unidad Azcapotzalco, Departamento de Humanidades, Unidad Xochimilco, Departamento de Política y Cultura. Mexico, D.F.: Universidad Autónoma Metropolitana. 1985– three issues per year.

Istoriia KPSS, nauchnyi kommunizm, filosofiia, pravo. Leningradskii gosudarstvennyi universitet imeni A. A. Zhdanova. Saint Petersburg: Izd-vo Leningradskogo universiteta. 1985– quarterly.

Lexicon philosophicum. Centro per il Lessico intellettuale europeo. Rome: Edizioni dell'Ateneo. 1985– annual.

Mathesis: filosofía e historia de las matemáticas. Universidad Nacional Autónoma de México, Departamento de Matemáticas, Grupo de Filosofía e Historia de las Matemáticas. [Mexico]: Universidad Nacional Autónoma de México, Departamento de Matemáticas, Facultad de Ciencias. 1985– quarterly.

Radovi. Razdio filozofije, psihologije, sociologije i pedagogije. Sveuciliste u Splitu, Filozofski fakultet Zadar. Zadar, Croatia: Fakultet. 1985– annual.

Studies in Logic and Theory of Knowledge. Lublin, Poland: [Katolickiego Uniwersytetu Lubelskiego, Towarzystwo Naukowe]. 1985–.

1986

Filosofia. Rome: Laterza. 1986–1995 annual.

Istoriko-filosofskii ezhegodnik. Institut filosofii (Akademiia nauk SSSR). Moscow: Nauka. 1986– annual.

Pensamiento hondureño. Tegucigalpa, Honduras: Imprenta Calderón. 1986–1987 semiannual.

Revista de filosofía latinoamericana y ciencias sociales. Buenos Aires, Argentina: Asociación de Filosofía Latinoamericana y Ciencias Sociales. 1986–2000.

Sédiments. LaSalle, Quebec: Hurtubise HMH. 1986– triennial.

Synthesis philosophica. Hrvatsko filozofsko drustvo, Savez filozofskih drustava Jugoslavije. Zagreb, Croatia: Croatian Philosophical Society. 1986– semiannual.

Vestnik Leningradskogo universiteta. Seriia 6, Istoriia KPSS, nauchnyi kommunizm, filosofiia, pravo. Leningradskii gosudarstvennyi universitet imeni A. A. Zhdanova. Saint Petersburg: Izd-vo Leningradskogo universiteta. 1986–1990 four issues per year.

1987

Analogía. Dominicans, Provincia de Santiago de México, Centro de Estudios. Xochimilco, Mexico: Centro de Estudios de la Provincia de Santiago de México de la Orden de Predicadores. 1987–.

La Balsa de la medusa. Madrid: Ediciones Antonio Machado. 1987– quarterly.

Filosofia politica. Bologna: Il Mulino. 1987– semiannual.

Filosofskie nauki. Soviet Union, Ministerstvo vysshego i srednego spetsial'nogo obrazovaniia. [Moscow: Izd-vo Vysshaia shkola]. 1987– twelve issues per year.

Filosofskie osnovaniia teorii mezhdunarodnykh otnoshenii. Moscow: Institut nauchnoi informatsii po obshchestvennym naukam (Akademiia nauk SSSR). 1987– irregular.

Signos. Universidad Autónoma Metropolitana, Unidad Iztapalapa, Departamento de Filosofía. Iztapalapa, Mexico: Universidad Autónoma Metropolitana. 1987–1996 annual.

Zborník Filozofickej fakulty Univerzity Komenského. Logica et methodologica. Bratislava: Univerzita Komenského v Bratislave, Filozofická fakulta. 1987–.

Zhe xue dong tai. Zhongguo she hui ke xue yuan, Zhe xue yan jiu suo. Beijing: Zhe xue yan jiu za zhi she. 1987– monthly.

1988

Analele stiintifice ale Universitatii "Al. I. Cuza" din Iaşi. Filosofie. Iaşi, Romania: Universitatea "Al. I. Cuza" din Iaşi. 1988–[?] annual.

Athenäums monografien. Philosophie. Frankfurt am Main: Athenäum. 1988–1993 irregular.

Metalogicon: rivista internazionale di logica pura e applicata di linguistica e di filosofia. Naples: L.E.R. 1988– semiannual.

Philosophia perennis: annales française de philosophie stylistique. Association française de philosophie stylistique. Lausanne, Switzerland: Editions l'Age d'homme. 1988.

Ratio juris. Università di Bologna. Oxford, and New York: Basil Blackwell for the University of Bologna. 1988–1995 three issues per year, 1996– quarterly.

1989

Bulletin de la Société américaine de philosophie de langue française. Northern Illinois University, Société américaine de philosophie de langue française. DeKalb, IL: Société américaine de philosophie de langue française. 1989–2000 quarterly.

Daimon: revista de filosofía. Murcia, Spain: Universidad de Murcia, Departamento de Filosofía y Lógica. 1989– semiannual.

Filosofs'kaia i sotsiologicheskaia mysl'. Instytut filosofiï (Akademiia nauk Ukraïns'koï RSR). Kiev, Ukraine: Naukova dumka. 1989– monthly.

Filosofs'ka i sotsiolohichna dumka. Instytut filosofiï (Akademiia nauk Ukraïns'koï RSR). Kiev, Ukraine: Naukova dumka. 1989–1990s monthly.

Filozofski vestnik. Filozofski institut (Slovenska akademija znanosti in umetnosti). Ljubljana, Slovenia: Akademija. 1989– three issues per year.

1990

Filosofija, sociologija. Lietuvos Mokslu akademija. Vilnius, Lithuania: Mokslas. 1990– three issues per year.

I castelli di Yale. Florence: Vallecchi editore. 1990s– annual.

Ìmodòye: A Journal of African Philosophy. [Lagos, Nigeria: University of Lagos, Faculty of Arts, Department of Philosophy]. 1990–.

Isegoría: revista de filosofía moral y política. Instituto de Filosofía (Consejo Superior de Investigaciones Científicas). Madrid: Instituto de Filosofía. 1990– semiannual.

Kairos. Université de Toulouse-Le Mirail, Faculté de philosophie. Toulouse: Presses Universitaires du Mirail. 1990– irregular.

Novoe v zhizni, nauke, tekhnike. Filosofiia i zhizn'. Moscow: Izd-vo Znanie. 1990–1991 monthly.

Die Philosophin. Tübingen, Germany: Edition Diskord. 1990– semiannual.

Revista de filozofie si drept. Institutul de Filozofie, Sociologie si Drept (Akademiia nauk Respubliki Moldova). Chişinău, Moldova: Izd-vo Shtiintsa. 1990s– three issues per year.

Revue roumaine de philosophie et logique. Academia Româna. Bucharest: Editura Academiei. 1990– quarterly.

Zeitschrift für allgemeine Wissenschaftstheorie [Journal for General Philosophy of Science]. Dordrecht, Netherlands, and Boston: Kluwer Academic. 1990– two issues per year.

1991

Filosofskii ezhegodnik. Yekaterinburg, Russia: Izd-vo Ural'skogo un-ta. 1991– annual.

Jahrbuch des Forschungsinstituts für Philosophie Hannover. Forschungsinstitut für Philosophie Hannover. Hildesheim, Germany: Bernward. 1991–1993 annual.

Nachala: organ Filosofskogo obshchestva SSSR. Filosofskoe obshchestvo SSSR, Filosofskoe obshchestvo, Assotsiatsiia prepodavatelei gumanitarnykh nauk, Vserossiiskii blagotvoritel'nyi fond kul'tury, nauki i iskusstva "Ros'." Moscow: Obshchestvo. 1991–1998 four issues per year.

Novye idei v filosofii: ezhegodnik Filosofskogo obshchestva SSSR. Filosofskoe obshchestvo SSSR. Moscow: Nauka. 1991– annual.

Philosophical Issues. Sociedad Filosófica Ibero Americana. Atascadero, CA: Ridgeview. 1991– annual.

Répertoire bibliographique de la philosophie [International Philosophical Bibliography]. Louvain-la-Neuve, Belgium: Université catholique de Louvain, Editions de l'Institut supérieur de philosophie. 1991– quarterly.

Silentium: neperiodicheskoe izdanie Filosofsko-kul'turologicheskogo issledovatel'skogo tsentra "Eidos." Saint Petersburg: Izd-vo Filosofsko-kul'turologicheskii issledovatel'skii tsentr "Eidos." 1991–[?] irregular.

Stupeni. Sovetskii fond kul'tury, Leningradskii filial "Grifon," Tvorcheskoe ob'edinenie "Stupeni." Saint Petersburg: Stupeni. 1991–1998 three issues per year.

Teoria. Pisa: ETS. 1991– semiannual.

Vestnik Leningradskogo universiteta. Seriia 6, Filosofiia, politologiia, teoriia i istoriia sotsializma, sotsiologiia, psikhologiia, pravo. Leningradskii gosudarstvennyi universitet. Saint Petersburg: Izd-vo Leningradskogo universiteta. 1991– quarterly.

1992

Dong wu zhe xue chuan xi lu. Dong wu da xue, Zhe xue xi. Taipei: Dong wu da xue. 1992–1995 annual.

Filosofski alternativi. Institut za filosofski nauki (Bulgarska akademiia na naukite). Sofia, Bulgaria: Institut po filosofski nauki pri BAN. 1992– bimonthly.

Obshchestvennye nauki v Rossii. Seriia 3, Filosofiia. Moscow: Institut nauchnoi informatsii po obshchestvennym naukam (Rossiiskaia akademiia nauk). 1992– six issues per year.

Put'. Moscow: Progress. 1992–.

Spisy Masarykovy univerzity v Brne, Filozofická fakulta. Brno, Czech Republic: Masarykova univerzita. 1992– irregular.

Vestnik Sankt-Peterburgskogo universiteta. Seriia 6, Filosofiia, politologiia, sotsiologiia, psikhologiia, pravo. Sankt-Peterburgskii gosudarstvennyi universitet. Saint Petersburg: Izd-vo Sankt-Peterburgskogo universiteta. 1992– quarterly.

Zhe xue za zhi. Taipei: Ye qiang chu ban she. 1992– quarterly.

1993

Axiomathes: quaderni del Centro studi per la filosofia mitteleuropea. Centro studi per la filosofia mitteleuropea. Trent, Italy: Il Poligrafo. 1993– three issues per year.

Comenius-Jahrbuch. Deutsche Comenius-Gesellschaft. Sankt Augustin, Germany: Academia Verlag. 1993– annual.

Cuadernos del sur. Filosofía. Bahía Blanca, Argentina: Universidad Nacional del Sur, Departamento de Humanidades. 1993–.

Filosofiia kul'tury. Samarskii gosudarstvennyi universitet, Kafedra filosofii gumanitarnykh fakul'tetov. [Samara, Russia]: Samarskii universitet. 1993–.

Ludus vitalis. Mexico, D.F.: Centro de Estudios Filosóficos, Políticos y Sociales Vicente Lombardo Toledano. 1993– semiannual.

Nemetsko-russkii filosofskii dialog. Institut nauchnoi informatsii po obshchestvennym naukam (Rossiiskaia akademiia nauk), Institut filosofskikh issledovanii. Gannovera. Moscow: INION RAN. 1993–.

Peterburgskie chteniia po teorii, istorii i filosofii kul'tury. Komitet po kul'ture i turizmu Merii Sankt-Peterburga, Filosofsko-kul'turologicheskii issledovatel'skii tsentr "Eidos." Saint Petersburg: Izd-vo Filosofsko-kul'turologicheskii issledovatel'skii tsentr "Eidos." 1993– annual.

Sotsial'nye i gumanitarnye nauki. Seriia 3, Filosofskie nauki Otechestvennaia i zarubezhnaia literatura. Institut nauchnoi informatsii po obshchestvennym naukam (Rossiiskaia akademiia nauk). Moscow: INION RAN. 1993–1994 four issues per year.

Studia philosophica. Tartu Ulikool, Filosoofia Osakond. Tartu, Estonia: Ulikool. 1993–.

Tekhnema: Journal of Philosophy and Technology. Paris: American University of Paris. 1993– annual.

1994

Eon: al'manakh staroi i novoi kul'tury. Institut nauchnoi informatsii po obshchestvennym naukam (Rossiiskaia akademiia nauk), Laboratoriia teorii i istorii kul'tury, TSentr gumanitarnykh nauchno-informatsionnykh issledovanii, Otdel kul'turologii. Moscow: INION RAN. 1994–.

Sfinks. Tvorcheskoe ob'edinenie "Stupeni." Saint Petersburg: Alga-Fond. 1994– irregular.

Sotsial'nye i gumanitarnye nauki. Seriia 3, Filosofiia. Otechestvennaia i zarubezhnaia literatura. Institut nauchnoi informatsii po obshchestvennym naukam (Rossiiskaia akademiia nauk). Moscow: INION RAN. 1994– four issues per year.

1995

Dialogue and Universalism. Uniwersytet Warszawski, Centrum Uniwersalismu, Polska Akademia Nauk, International Society for Universalism. Warsaw: Warsaw University, Centre of Universalism. 1995– monthly.

Edith Stein Jahrbuch. Stein, Edith, and Teresianum. Würzburg, Germany: Echter. 1995– annual.

Fa zhi bo lan [Legality Vision]. Gong qing tuan, Shanxi Sheng wei, Shanxi sheng Qing shao nian fan zui yan jiu hui. Taiyuan, China: Fa zhi bo lan bian ji bu. 1995– monthly.

Filosofiia nauki. Moscow: Institut filosofii (Rossiiskaia akademiia nauk). 1995– annual.

Hyle. Universität Karlsruhe, Institut für Philosophie. Karlsruhe, Germany: Hyle. 1995– annual.

Rubezhi. Demokraticheskaia al'ternativa. Moscow: Agentstvo "Vremia." 1995–1998 monthly.

1996

Dong Wu zhe xue xue bao [Soochow Journal of Philosophical Studies]. Taipei: Dong Wu da xue chu ban she. 1996–2002 annual, 2003– biennial.

Forum philosophicum: studia a Facultate Philosophica Societatis Jesu Cracoviae edita. Towarzystwo Jezusowe w Krakowie, Wydzial Filozoficzny. Kraków: WAM Press. 1996– annual.

Naukovi zapysky. Natsional'nyi universytet. Kiev, Ukraine: Kyievo-Mohylians'ka Academia. 1996–.

Pleroma. Ivano-Frankivs'k, Ukraine: Proekt "Delos'kyi nyrets'." 1996–.

Postizhenie kul'tury. Moscow: Rossiiskii institut kul'turologii. 1996– annual.

Variaciones Borges: revista del Centro de Estudios y Documentación "Jorge Luis Borges." Århus, Denmark: Århus universitet, Jorge Luis Borges Center for Studies and Documentation, Romansk institut. 1996– semiannual.

1997

Dukh i litera. Natsional'nyi universytet "Kyievo-Mohylians'ka akademiia." Kiev, Ukraine: FAKT. 1997–.

Istoriia filosofii. Moscow: Institut filosofii (Rossiiskaia akademiia nauk). 1997– irregular.

Metafizicheskie issledovaniia. Sankt-Peterburgskii gosudarstvennyi universitet, Laboratoriia metafizicheskikh issledovanii. Saint Petersburg: Aleteiia. 1997– four issues per year.

Mysl': ezhegodnik Peterburgskoi assotsiatsii filosofov. Peterburgskaia assotsiatsiia filosofov, Sankt-Peterburgskii gosudarstvennyi universitet. Saint Petersburg: Izd-vo Universiteta. 1997– annual.

Recherches de théologie et philosophie médiévales. Universität zu Köln, Thomas-Institut, Centre de Wulf-Mansion. Leuven, Belgium: Peeters. 1997– semiannual.

Vestnik Rossiiskogo filosofskogo obshchestva. Moscow: Rossiiskaia akademiia nauk, Rossiiskoe filosofskoe obshchestvo. 1997– quarterly.

1998

Alpha omega: rivista di filosofia e teologia dell'Ateneo Pontificio Regina Apostolorum. Ateneo pontificio Regina Apostolorum. Rome: L'Ateneo. 1998– three issues per year.

Bibliographie de la philosophie [Bibliography of Philosophy]. International Institute of Philosophy, Unesco, Centre national de la recherche scientifique (France). Paris: J. Vrin. 1998– annual.

Denkwege. Tübinger Gesellschaft für Phänomenologische Philosophie. Tübingen, Germany: Attempto. 1998–.

Pensamiento de los confines. Universidad de Buenos Aires, Programa de Estudios de Cultura y Pensamiento Contemporáneo, Centro de Estudios de Profesores Universitarios. Buenos Aires, Argentina: Universidad de Buenos Aires, Argentina, Diótima. 1998– semiannual.

Philosophiegeschichte und logische Analyse [Logical Analysis and History of Philosophy]. Paderborn, Germany: F. Schöningh. 1998– annual.

2000

Cités. Vendôme, France: Presses Universitaires de France. 2000– quarterly.

2002

Antropologiia kul'tury. Moskovskii gosudarstvennyi universitet im. M. V. Lomonosova, Institut mirovoi kul'tury. Moscow: OGI. 2002–.

2003

Studia antyczne i mediewistyczne. Warsaw: Wydawn. Instytut Filozofii i Socjologii (Polska Akademia Nauk). 2003– annual.

Michael J. Farmer (2005)

index

Terms are alphabetized on a word-by-word basis. Page references to entire articles are in ***boldface.***

A

B

C

D

E

G

H

I

J

K

L

M

N

O

Q

R

S

T

U

V

W

X

Y

Z

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

CABANIS – DESTUTT DE TRACY

2 volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT

Editor in Chief

MACMILLAN REFERENCE USA

An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

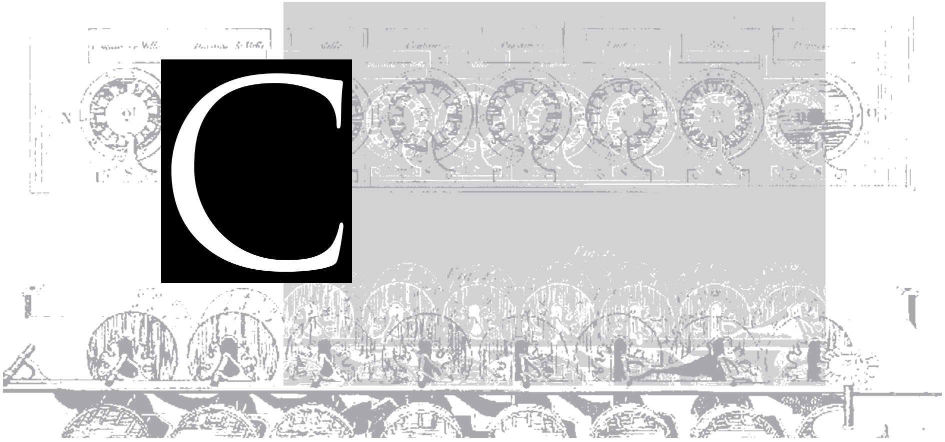

C

CABALA

See *Kabbalah*

CABANIS, PIERRE-JEAN GEORGES

(1757–1808)

Pierre-Jean Georges Cabanis was, with Comte Antoine Louis Claude Destutt de Tracy, the leader of the Idéologues. A precocious student of philosophy and of the classics, he chose medicine as a career, but he never practiced. As a protégé of Claude-Adrien Helvétius's widow, he frequented the company of Étienne Bonnot de Condillac, Baron d'Holbach, Benjamin Franklin, and Thomas Jefferson. When Voltaire disparaged his poetry in 1778, Cabanis turned to physiology and philosophy. During the Revolution, he collaborated with Mirabeau on public education and was an intimate of Marquis de Condorcet. Later, he backed the Directory and Napoleon Bonaparte's coup d'état of 18 Brumaire. Although Napoleon made him a senator, Cabanis opposed his tyrannical policies. Bitter and scornful, Napoleon dubbed Cabanis's group "Idéologues." Cabanis wrote on medical practice and teaching, but his fame and influence derive from one book, *Rapports du physique et du moral de l'homme* (12 memoirs written between 1796 and 1802, published in 1802).

The Idéologues (who also included Constantin Volney, Condorcet, Antoine Lavoisier, and Pierre de Laplace) were often scorned in their time, and later, as belated *philosophes* and purveyors of visionary speculations. In the rising tide of metaphysical idealism, their positivistic approach was held in disfavor. They suffered from the influence of the religious revival and the spell exercised by François René de Chateaubriand's *Le génie du Christianisme,* as well as from the popularity of "Illuminist" fads derived from Masonic practices. Their political activity during the Revolution also worked against them, and Napoleon's suppression of their movement left them without an outlet for publication.

Cabanis, like the others, sought a mechanistic explanation of the universe, nature, and human behavior—an approach later continued by Auguste Comte and Hippolyte-Adolphe Taine. Matter alone is real and eternal in its many transitory forms. As Lavoisier had applied analysis to chemistry, so—Cabanis declared—it could be applied to ideas, which could thereby be reduced to the original sensations whence they spring. Self-interest, the

pursuit of happiness and pleasure, and self-preservation are the only motives of action. These notions, already advanced by the eighteenth-century materialists, were systematically developed by Cabanis and Destutt de Tracy. The study of man, they held, must be reduced to physics and physiology. Man must be observed and analyzed like any mineral or vegetable. The medical expert, said Cabanis, should play the part formerly taken by the moralist (an idea that harks back to René Descartes and Julien Offray de La Mettrie). "Physiology, analysis of ideas, and morals are three branches of one science which may be called the science of man." Consequently, Cabanis and his fellow theorists refused to recognize notions not based on phenomena or sensations, that is, not susceptible of exact knowledge and (ultimately, at least) of mathematical notation. An understanding of the "mechanism of language" was considered essential to the understanding of the "mechanism of the intellect" and to the meaning of ideas. Language itself, however, had to be illumined by analysis of the sensations which constitute an idea an by the functioning of the intellect.

In his preface to the *Rapports du physique et du moral de l'homme,* Cabanis insisted that both the moralist and the physician are interested in the whole man; that is, in the physical and the moral, which are inseparable, and incomprehensible taken separately. The moral sciences must be placed on a physical basis. The union of mind and body is the theme of the first "Mémoire." Sensation is the necessary cause of our ideas, feelings, needs, and will. Since sensitivity is the connection between biological life and mind, the mental is only the physical considered from a certain point of view. Cabanis makes a famous comparison between the brain and the stomach: As the latter is a machine for digesting food, so the former is a machine for digesting impressions, by "the secretion of thought." He then develops a genetic analysis of sensations and ideas. There are no causes except those which can act on our senses, no truths except in relation to "the general way of feeling" of human nature, which varies with such positive factors as age, sex, disposition, health, climate, and so on. Thus the state of the abdominal viscera may influence the formation of ideas.

The second "Mémoire" is a "physiological history of sensations." Cabanis defines life as feeling and, following the work of Albrecht von Haller and La Mettrie, discusses the difference between sensitivity and irritability. The latter, he maintains, is only a result of the former, which is the basic biological phenomenon; since both depend on the nerves, they are essentially the same. Voluntary movements come from perceptions, which arise from sensations. Involuntary movements are caused by the organs' sensitivity, which produces the unconscious (autonomic) impressions that determine many of our ideas and decisions. The action of the nervous system, moreover, is only a specialized application of the laws of physical motion, which are the source of all phenomena. The third "Mémoire" develops a theory of the unconscious. The nervous system is affected by internal changes, that is, by memory and imagination; thus within man exists "another internal man" in constant action, the effects of which are noticeable in dreams. The fourth "Mémoire" explores the influence of age on ideas and "moral affections." The organs, like all else in nature, are in constant motion, and are therefore involved in decomposition and recomposition. Consequently, variations in the cellular tissue produce physical and psychic changes due to chemical action. The fifth "Mémoire" takes up sexual differences. The generative organs are essentially glandular, and their secretions influence the brain and the whole body. Unknown primitive "dispositions" (structures), which cause the embryo to be male or female, are also the cause of sexual differences, both physical and psychic. The fact that women can be forced to reproduction and men only excited to it produces vast differences in habits and mental outlook. What the sexes have in common constitutes human nature.

The sixth "Mémoire" treats the influence of "temperament," that is, the determining effects of the inherited physical constitution. Thus a large heart and lungs produce an energetic character, small ones an intellectual character. Because of heredity, the human race could be improved by hygienic methods. Believing in the inheritance of acquired characteristics and in improvement of species through crossbreeding, Cabanis pleads for a program of eugenics that will do for the human species what human beings have done for dogs and horses. In the seventh "Mémoire" Cabanis explores emotional and mental perturbations caused by diseases. For instance, weakness and irritability of the stomach produce muscular enervation and rapid alternations between excitement and depression. The eighth "Mémoire" discusses such effects of diet, air pressure, humidity and temperature, as excitation and sedation. Cabanis analyzes the effects of different foods and drinks, but his information and conclusions are rather fantastic.

Climate is the subject of the ninth "Mémoire." Man, the most modifiable animal, responds to heat and cold with differences in sexual and physical activity, and consequently in mental and moral habits. The tenth "Mémoire" is the longest. It explores the phenomena of

animal life, including sensitivity, instinct, sympathy, sleep, dreams, and delirium. The forces that cause matter to organize (a natural tendency) are unknown, and will always remain so. Nevertheless these forces are only physical, and life is only organization. Cabanis believed in spontaneous generation. Species have evolved through chance mutations ("fortuitous changes") and planned mutation ("man's experimental attempts"), which change the structures of heredity. Cabanis does not, however, develop a general theory of evolution. The eleventh "Mémoire" concerns the influence of the "moral" (mental) on the physical, which is merely the action of the brain on the body. The last "Mémoire," on "acquired dispositions," treats the influence of habituation and experience in general.

As a positivist, Cabanis was willing to renounce ultimate explanations. He was interested only in cause and effect on the level of phenomena. Unlike the other Idéologues, he was much influenced by La Mettrie and the man-machine school. He opposed the psychological method of Condillac and the sensationists, which was limited to external sensations. He preferred the physiological approach, which emphasized hereditary dispositions, the state of the organs, dreams, and automatic or unconscious impulses. These factors were more significant for him than experience (sensation) in determining the individual's behavior; for the tabula rasa concept ignored what the child or adult brings to experience. For the same reason, Condillac's statue is only an unreal abstraction from the reality of the unified, total, active organism. Cabanis was interested in the moral and social improvement of humankind, which he considered possible through an understanding of physiology—a science that he thought would eventually influence even positive law.

Cabanis and the Idéologues were one moment of a tradition that extends from Epicurus to the contemporary logical positivists (whose interest in linguistic analysis was prefigured by the Idéologues). Cabanis, like the others, has frequently been accused of impoverishing human experience by reducing it to the physical and mechanical level, and by denying the possibility of transcending internal and external sensations. On the other hand, the Idéologues considered man to be his own justification and the master of his own destiny. They had faith in his capacity to progress indefinitely by means of his own resources.

Bibliography

WORKS BY CABANIS

The *Oeuvres complètes* (Paris, 1823–1825) of Cabanis was edited by P. J. G. Thurot.

Cabanis, Pierre-Jean Georges. *On the Relations between the Physical and Moral Aspects of Man,* edited by George Mora. Translated by Margaret Duggan Saidi. Baltimore: Johns Hopkins Press, 1981.

WORKS ON CABANIS

The best study of the Idéologues (although it ignores Cabanis's connection with La Mettrie and the man-machine outlook) is Emile Cailliet, *La tradition littéraire des Idéologues* (Philadelphia: American Philosophical Society, 1943). See also Charles H. Van Duzer, *The Contribution of the Idéologues to French Revolutionary Thought* (Baltimore: Johns Hopkins Press, 1935), and the more apologetic F. Picavet, *Les idéologues* (Paris: Alcan, 1891).

Moravia, Sergio. "From 'Homme Machine' to 'Homme Sensible': Changing Eighteenth-Century Models of Man's Image." *Journal of the History of Ideas* 39 (1978): 45–60.

Other Recommended Titles

Richards, Robert J. "Influence of Sensationalist Tradition on Early Theories of the Evolution of Behavior." *Journal of the History of Ideas* 40 (1979): 85–105.

Staum, Martin S. *Cabanis: Enlightenment and Medical Philosophy in the French Revolution.* Princeton, NJ: Princeton University Press, 1980.

Wright, John P., and Paul Potter, eds. *Psyche and Soma: Physicians and Metaphysicians on the Mind Body Problem from Antiquity to Enlightenment.* Oxford: Clarendon Press, 2000.

L. G. Crocker (1967)
Bibliography updated by Tamra Frei (2005)

CAIRD, EDWARD
(1835–1908)

Edward Caird, a leading Scottish Hegelian, was born in Greenock, the fifth of seven boys. His eldest brother, John Caird, became well known as a preacher and theologian, and exercised considerable influence on the young Edward. Educated at Greenock Academy and Glasgow University (with a brief interlude at St. Andrews), Edward Caird went to Balliol College, Oxford, gaining first-class honors in Classical Moderations and in "Greats." From 1864 to 1866 he was a fellow and tutor of Merton, leaving to take the chair of moral philosophy at Glasgow, which he held until 1893. He then returned to Oxford to succeed Benjamin Jowett as master of Balliol. He resigned because of ill health in 1907, and died the year after.

Caird had a profound influence on his students, who regarded themselves as his disciples and included such distinguished philosophers as Henry Jones, J. H. Muir-

head, J. S. Mackenzie, and John Watson. "The greatest theme of modern philosophy," Caird held, "is the problem of the relation of the human to the divine" (*The Evolution of Theology in the Greek Philosophers,* 1904). Many of his Glasgow students were destined for the church, and his liberalizing influence on religion was widely transmitted through them beyond the classroom.

Caird's philosophy was a form of speculative idealism, based on Immanuel Kant but going beyond him. It was essentially a philosophy of reconciliation. The need for philosophy, he held, arises from the apparently irreconcilable opposition between different elements in our spiritual life—between subject and object, religion and science, freedom and determination, reason and desire. Unless we reconcile these antagonisms in a higher unity, we cannot achieve the spiritual harmony without which the highest achievements of humanity are impossible.

Kant, he was convinced, had found the key to the problem, but had failed to grasp the implications of his own doctrine. Caird had first to clear away what he thought was a common misinterpretation of Kant and then to go further along the Kantian road, with G. W. F. Hegel as his guide. Kant had been held, according to Caird, to teach that the material of knowledge is given in sense perception and that the mind then goes to work on it, ordering it by concepts supplied by itself. But, in fact, for Kant there *are* no objects until thought has done its work. Thought enters into the very constitution of experience. And further, the process of knowing is dominated by an "idea of the Reason," which drives the mind to seek a form of experience in which all differences are seen as elements in a single system.

But instead of insisting that the larger the part played in knowledge by the mind's synthetic activity, the more adequate that knowledge is, Kant took the view that this activity confines us to appearances and bars us from things-in-themselves. He should have shown, Caird argued, that our knowledge of objects will be imperfect insofar as we fail to recognize that they are only partial aspects of the ideal whole toward which reason points.

Caird's ethical theory had close affiliations with that of his lifelong friend, T. H. Green. His main problem centered on the opposition of inclination and duty, and his solution lay in establishing the power of human beings to determine their conduct by reference to the self, as a permanent center, as distinct from its relatively isolated and transient desires. A self-conscious being seeks *self*-satisfaction, not just the satisfaction of this or that desire. And in this power of determining conduct by reference to the self lies human freedom.

The principle of evolution, Caird recognized, was of great value in reconciling differences, and in his Gifford Lectures, *The Evolution of Religion* (1891–1892), he traced the development of a single religious principle through its varied manifestations in the main religions of the world.

See also Green, Thomas Hill; Hegel, Georg Wilhelm Friedrich; Hegelianism; Kant, Immanuel;.

Bibliography

WORKS BY CAIRD

A Critical Account of the Philosophy of Kant. Glasgow: J. Maclehose, 1877.

Hegel. Edinburgh: Blackwood, 1883.

The Social Philosophy and Religion of Comte. New York: Macmillan, 1885.

The Critical Philosophy of Immanuel Kant. Glasgow: J. Maclehose, 1889.

Essays on Literature and Philosophy. Glasgow: J. Maclehose, 1892.

The Evolution of Religion. Glasgow, 1893: J. Moclehose.

The Evolution of Theology in the Greek Philosophers. Glasgow: J. Maclehose, 1904.

WORKS ON CAIRD

Jones, Sir Henry, and J. H. Muirhead. *The Life and Philosophy of Edward Caird.* Glasgow: Maclehose, Jackson, 1921.

Mackenzie, J. S. "Edward Caird as a Philosophical Teacher." *Mind* 18 (1909): 509–537.

Watson, John. "The Idealism of Edward Caird." *Philosophical Review* 18 (1909): 147–163, 259–280.

A. K. Stout (1967)

CAIRNS, DORION
(1901–1973)

Thomas Dorion Cairns was born on July 4, 1901. His father was a Methodist pastor. Cairns studied phenomenological theory of value with Winthrop Bell at Harvard in 1923 and 1924, used a traveling fellowship to study with Edmund Husserl for two years, returned later for over another year, and received his doctorate with *The Philosophy of Edmund Husserl* in 1933. After temporary positions in New York, Cairns taught psychology as well as philosophy at Rockford College from 1938 to 1950. During World War II, he won a Bronze Star as a prisoner of war interrogator in the Air Corps. He was invited to the New School for Social Research in 1954 by Alfred Schutz, taught there with Aron Gurwitsch during the 1960s, retired in 1969, and died on January 4, 1973. All who heard him considered him a brilliant teacher, but he pub-

lished little. However, his translations of Husserl's *Cartesian Meditations* (1960) and *Formal and Transcendental Logic* (1969) played an important role. His *Conversations with Husserl and Fink* (1976), *Guide for Translating Husserl* (1973), and a dozen essays from his Nachlass have appeared posthumously. The editing of the manuscripts of his New School lecture courses began in 2003.

SOME RESULTS OF CAIRNS'S INVESTIGATIONS

Cairns's original project was to bring Husserl's earlier work up to the level of *Cartesianische Meditationen* (1931), but from attempting to repeat the investigations, he came to propose at least seven major revisions.

(1) Like many in modern philosophy, Husserl pursued a first philosophy that seeks grounds in consciousness for everything else. Hence, the positive sciences are grounded in a primal science called transcendental phenomenology. This first philosophy is transcendental because it refrains from accepting the intramundane status of consciousness in order to avoid trying to ground the world in part of itself. Cairns always accepted the transcendental *epochē* and agreed with his master that it was Husserl's chief contribution.

Husserl's publications emphasize the theory of science (*Wissenschaftstheorie*), especially the theory of logic, although there are remarks about valuation and action. Cairns revised Husserl so that the goal of phenomenological philosophy became not merely knowledge, but the integration of critically justified willing, valuing, and believing.

(2) There is a considerable shift in emphasis when Cairns follows his revision of Husserl's goal by affording value theory and theoretical ethics as much attention as epistemology within his presentation phenomenological first philosophy.

(3) Although many stop after defining intentionality (which Cairns came to call "intentiveness") as directedness toward objects, Cairns followed Husserl in using the concept of synthesis to make this insight fruitful—for example, a synthesis of intentive processes constituting an object as self-identical and different from other objects.

Although Husserl saw intentiveness more clearly than anybody previously, Cairns believed that Husserl still tended to reify the noema (i.e., the thing-as-intended-to in an intentive process), which is easy to do if one conceives of intentionality as a relation, whereas intentiveness is actually a property.

(4) Husserl held that there were sensuous hyletic data immanent in the stream of consciousness. These moments are themselves not intentive and no distinction was needed between sensing and sensa for Husserl, but for Cairns that distinction must be carefully maintained and sensa are transcendent of consciousness.

(5) Cairns held that Husserl left much to be done on the emotions and advanced the account by showing above all how emotion can be critically justified by the evidencing of objects valued in it. By contrast, rationality for most philosophers is wholly a matter of propositions conforming to the norms of logic.

(6) Cairns went beyond Husserl in developing the idea of ethics as a theory of critically justified willing (i.e., a theory of practical reason).

(7) Cairns's most radical revision of Husserl concerns the theory of the other. He objected to the reduction of the sphere of ownness introduced in the latter's Fifth Cartesian Meditation because the procedure described as a suspending acceptance of a noema without a suspending acceptance of the noesis is impossible to perform. Instead, Cairns asserted that a series of noetico-noematic strata of transcendental consciousness must be reflectively suspended through "unbuilding" (*Abbau*). Fields of sensa are ultimately reached. Through "building up" (*Aufbau*), one allows founded strata to be motivated once again, and thereby can reflectively observe how the intersubjective world is constituted.

A fundamental distinction for most European and North American philosophers holds between inanimate physical nature and the stratum of animate nature. A course in Indian philosophy with James Houghton Woods at Harvard in 1923 prepared Cairns to recognize that when the sense "animate body" is transferred from one's own body it transfers not to some but to all sensuous objects—rocks, trees, and sky included—and that animism follows. In class, Professor Cairns would say that chairs were rather stupid animals who stood in one place unless moved by somebody else. The distinction between inanimate and animate is then secondary, and may be recast as a distinction between animals with evident organs of sensation and locomotion and those without them. And phenomenology is clearly not merely about human consciousness.

In an era when practically all soi-disant phenomenologists devote themselves entirely to the interpretation of texts, Dorion Cairns is among the few who made a strict distinction between what may be called scholarship, which includes translation as well as interpretation of

texts, and what may be called investigation, which is concerned not with texts, but with the "things themselves" in the signification whereby anything is a "thing." Like Husserl, Cairns regularly offered methodological reflections: he not only described the things reflectively observed, but also described how he had been able to analyze them, emphasizing reflection, analysis, "seeing," and description.

Furthermore, Cairns often began by describing the *psychological* phenomenological epochē and reduction—a methodological step whereby consciousness remains intramundane but is abstracted from other mundane things—before contrasting it with the specifically *transcendental philosophical* epochē and reduction that refrains from accepting the intramundaneity of consciousness and makes the grounding of the world and all sciences of it possible. Although investigation, methodology included, predominates overwhelmingly in the writings of Husserl, it may be hoped that the posthumous publications of his arguably closest critical continuer will also help phenomenologists remember what phenomenology is.

See also Consciousness; Consciousness in Phenomenology; Husserl, Edmund; Phenomenology.

Bibliography

WORKS BY CAIRNS

"An Approach to Phenomenology." In *Essays in Memory of Edmund Husserl*, edited by Marvin Farber, 3–18. Cambridge, MA: Harvard University Press, 1940.

"The Ideality of Verbal Expressions." *Philosophy and Phenomenological Research* 1 (1941): 453–462.

"Phenomenology." In *A History of Philosophical Systems*, edited by Vergilius Ferm, 353–364. New York: Philosophical Library, 1950.

"The Many Senses and Denotations of the Word *Bewusstsein* ('Consciousness') in Edmund Husserl's Writings." In *Life-World and Consciousness: Essays for Aron Gurwitsch*, edited by Lester E. Embree, 19–31. Evanston, IL: Northwestern University Press, 1972.

Guide for Translating Husserl. The Hague: Martinus Nijhoff, 1973.

"My Own Life," edited by Lester Embree. In *Phenomenology: Continuation and Criticism: Essays in Memory of Dorion Cairns*, edited by Frederick I. Kersten and Richard M. Zaner. The Hague: Martinus Nijhoff, 1973.

"Perceiving, Remembering, Image-Awareness, Feigning Awareness." In *Phenomenology: Continuation and Criticism. Essays in Memory of Dorion Cairns*, edited by Frederick I. Kersten and Richard M. Zaner, 251–262. The Hague: Martinus Nijhoff, 1973.

Conversations with Husserl and Fink, edited by Husserl-Archives. The Hague: Martinus Nijhoff, 1976.

"Philosophy as Striving toward Universal *sophia* in the Integral Sense," edited by Lester Embree. In *Essays in Memory of Aron Gurwitsch*, edited by Lester Embree, 27–43. Washington, DC: Center for Advanced Research in Phenomenology and University Press of America, 1984.

"Reason and Emotion," edited by Lester Embree, Fred Kersten, and Richard M. Zaner. *Husserl Studies* 17 (2000): 21–33.

"Theory of Intentionality in Husserl," edited by Lester Embree, Fred Kersten, and Richard M. Zaner. *Journal of the British Society for Phenomenology* 22 (2) (2001): 116–124.

"The First Motivation of Transcendental Epoch&emacron;," edited by Lester Embree, Fred Kersten, and Richard M. Zaner. In *One Hundred Years of Phenomenology*, edited by Dan Zahavi and Frederik Stjernfelt, 219–231. Dordrecht, Netherlands: Kluwer, 2002.

WORKS ABOUT CAIRNS

Embree, Lester. "Dorion Cairns: The Last Lecture Course on Ethics." In *Phenomenological Approaches to Moral Philosophy*, edited by John J. Drummond and Lester Embree, 139–160. Dordrecht: Kluwer, 2002.

Kersten, Fred. *Phenomenological Method: Theory and Practice*. Dordrecht, Netherlands: Kluwer, 1989.

Lester Embree (2005)

CAJETAN, CARDINAL

(1469–1534)

Cajetan (Tommaso de Vio), the most influential Renaissance Thomist, studied and taught in Italy, early distinguishing himself in teaching, commentaries, and debates as a philosopher and theologian. Rising to the leadership of the Dominican Order and becoming prominent in ecclesiastical politics, he was made cardinal in 1517. In 1518–1519 he disputed with Martin Luther.

Cajetan's works number more than a hundred titles. His later writing was primarily devoted to biblical exegesis; his primary contributions to Thomistic philosophy and theology are due to his earlier commentaries and treatises, most notably his commentary on St. Thomas Aquinas's *De Ente et Essentia* (On being and essence, 1495), his treatise *De Nominum Analogia* (On the analogy of names, 1498), and his formidable commentary on Aquinas's *Summa Theologiae* (1507–1522), which is printed with the pontifical (Leonine) edition of Aquinas's work. Other significant philosophical works include commentaries on Porphyry's *Isagoge* and on Aristotle's *Categories*, *Posterior Analytics*, *De Anima*, *Physics*, and *Metaphysics* (these last two have never been published), and a treatise on economics.

The *De Ente et Essentia* commentary is a sophisticated defense of Aquinas's metaphysics, loosely organized in question format, clarifying (inter alia) the Thomistic

theses that being is the first object of cognition, that matter is the principle of individuation, and that essence and existence are really distinct in creatures. Sensitively attending to language, the work, with the *Categories* commentary, is also an important source for Cajetan's realist semantics.

De Nominum Analogia teaches a threefold classification and hierarchy of analogical signification. *Analogy of inequality* only counts as analogy from the metaphysician's perspective; logically, it is a form of univocation (as *body* is predicated equally of, though realized differently in, plant and stone). *Analogy of attribution* is Aristotle's *pros hen* equivocation; a term naming primarily one thing is extended to others by virtue of their relation to the first, as *healthy* denominates animal (intrinsically, as *subject* of health) and medicine (extrinsically, as *cause* of the animal's health). *Analogy of proportionality* is based not on a relation, but on a similarity of relations (as the body's ocular vision is proportional to the soul's intellectual vision). When proper and not merely metaphorical, denomination here is always intrinsic. Cajetan regards this as the most genuine form, a true mean between univocation and equivocation, and the majority of his treatise explores the implications (for abstraction, judgment, and reasoning) of proportionally similar concepts.

Cajetan's writings are shaped by the polemical context of Renaissance Thomism. Concerned to address the objections of humanists (such as Count Giovanni Pico Della Mirandola, whom he debated in 1495), Italian Averroists, and especially Scotists (foremost Anthony Trombetta, his contemporary at Padua and primary dialectical target of the *De Ente* commentary), Cajetan does not simply repeat formulas from Aquinas, he rearticulates Thomistic ideas in sometimes novel terminology. Despite this, and notwithstanding apparent departures from Aquinas on particular points (e.g., whether the soul's immortality is demonstrable), Cajetan was long regarded as a definitively authoritative expositor of Aquinas. When the twentieth-century Thomistic revival, distinguishing the historical Aquinas from longstanding scholastic traditions, emphasized differences between Cajetan and Aquinas, Étienne Gilson and others criticized Cajetan, especially on the topics of abstraction and existence. On analogy some scholars challenged whether the elements of Cajetan's comprehensive, systematic theory—especially the discussion of extrinsic versus intrinsic denomination, the preference for proportionality, and the threefold classification itself—are warranted from Aquinas's rather more dispersed and occasional reflections on the subject. Whether Cajetan's distinct philosophical vocabulary is a departure from the mind of his master, or a legitimate development of authentic Thomism in light of the innovations of the intervening centuries, remains a question, but the forcefulness of his mind has never been doubted.

See also Aristotle; Humanism; Thomas Aquinas, St.; Thomism.

Bibliography

WORKS BY CAJETAN

The Analogy of Names, and the Concept of Being. Translated by Edward A. Bushinski and Henry J. Koren. Pittsburgh, PA: Duquesne University Press, 1953.

Commentary on Being and Essence. Translated by Lottie H. Kendzierski and Francis C. Wade. Milwaukee, WI: Marquette University Press, 1964.

WORKS ABOUT CAJETAN

Reilly, John P. *Cajetan's Notion of Existence*. The Hague: Mouton, 1971.

Riva, Franco. *Analogia e univocità in Tommaso de Vio "Gaetano."* Milan, Italy: Vita e Pensiero, 1995.

Joshua P. Hochschild (2005)

CALDERONI, MARIO

(1879–1914)

Mario Calderoni ranks next to his teacher Giovanni Vailati as an Italian "Peircean pragmatist." He was graduated in law from the University of Pisa in 1901, and later lectured on the theory of values at the universities of Bologna and Florence.

Calderoni engaged in analyses of human behavior. These began with the interpretation of voluntary acts, which he regarded as the only nonmetaphysical problem of free will. In everyday life we all possess as good a criterion as is necessary to distinguish between voluntary and involuntary acts. To find out whether an act is to be called voluntary or not, we must modify the circumstances in which it usually occurs. If it still occurs in any case, we call it "involuntary"; if not, we call it "voluntary." The difference rests on the "plasticity" of voluntary acts, on their liability to modification by certain influences. A voluntary act "is liable not to be performed if the actor … is given some new information on its consequences." What determines his acting is some expectation, which we can modify "either by changing one of the actor's beliefs by means of persuasion or reasoning, or, so to say artificially, by adding to the consequences the act would bring about if it were performed" (*Scritti,* vol. 2, pp. 25–26.). This criterion would hold good even if it were proved that all our

acts are subject to the principle of causality. In Calderoni's hands, it became an empirical, perfectible tool applied to the analysis of moral and legal responsibility.

In *Disarmonie economiche e disarmonie morali* (Florence, 1906) Calderoni viewed moral life as a "wide market where some men … make determinate demands on other men who oppose such demands with more or less resistance and claim in their turn … some sort of reward." Moral acts are judged not according to their total value, but according to their marginal or comparative value. We tend to confer the highest moral value not on common acts but on acts so rare that we would be obliged to repress them if their normal production increased. The moral value of actions is therefore related to their supply.

See also Peirce, Charles Sanders; Vailati, Giovanni; Value and Valuation.

Bibliography

ADDITIONAL WORKS BY CALDERONI

I postulati della scienza positiva ed il diritto penale (The postulates of positive science and penal law). Florence, 1901. His doctoral dissertation, containing his main ideas in brief.

Scritti. Edited by Odoardo Campa. 2 vols. Florence: Società anon. editrice "La voce," 1924. Preface by Giovanni Papini.

For joint papers by Calderoni and Giovanni Vailati, see the entry on Vailati.

WORKS ON CALDERONI

Bozzi, Paolo. "Il pragmatismo italiano: Mario Calderoni." *Rivista critica di storia della filosofia* 12 (3) (1957): 293–322.

Renauld, J.-F. "L'oeuvre inachevée de Mario Calderoni." *Revue de Métaphysique et de Morale* 23 (1918): 207–231.

Santucci, Antonio. *Il pragmatismo in Italia.* Bologna: Il Mulino, 1963. Ch. 5, 216–262.

Ferruccio Rossi-Landi (1967)

CALVIN, JOHN
(1509–1564)

John Calvin, the Protestant reformer and theologian, was born at Noyon, France. The son of middle-class parents of considerable local importance, Calvin was early directed toward an ecclesiastical career. From 1523 to 1528 he studied theology in Paris, there becoming acquainted with both the scholastic and humanist trends of his day. When he had achieved the master of arts degree, Calvin, in response to his father's wishes, left Paris to study law at Orléans, finishing his doctorate there by early 1532.

By 1534 Calvin had decisively broken with his Catholic heritage and had joined the Protestant reform movement in France. From this time on, all his efforts were devoted to the cause of the Reformation, and most of the remainder of his life was spent preaching, teaching, and writing in Geneva. He carried on a voluminous correspondence with thinkers and reformers all over Europe, and he had a powerful voice in the political and educational, as well as the ecclesiastical, institutions of Geneva.

Calvin's major work was the *Institutes of the Christian Religion,* first published in 1536 and originally addressed to King Francis I of France in defense of the French Protestants. It was extensively revised several times, and the last edition, published in 1559, provides a systematic presentation of virtually all the lines of thought found in Calvin's other mature works.

KNOWLEDGE OF GOD AND SELF

"Nearly all the wisdom we possess," wrote Calvin in the opening of the *Institutes,* "consists of two parts: the knowledge of God and of ourselves." The overarching question in the *Institutes* is how we acquire this twofold knowledge, and the answers to this question have proved to be the most influential part of Calvin's thought.

Thomas Aquinas had taught that the theologian should start with God and then consider creatures insofar as they relate to God as their beginning and end. Calvin broke decisively with this approach in claiming that knowledge of God is so interrelated with knowledge of ourselves that the one cannot be had without the other. He taught that when we accurately reflect on ourselves, we realize the excellence of our natural gifts; but we also realize that our exercise of these gifts yields "miserable ruin" and unhappiness, and that "our very being is nothing but subsistence in the one God." Without this realization of our misery and dependence—especially of our misery—none of us comes, or even tries to come, to a knowledge of God. On the other hand, there is also no knowledge of self without a knowledge of God. Without a standard by which to measure ourselves, we invariably yield to pride, overestimating the worth of our natural gifts and overlooking the corruption that has resulted from the exercise of those gifts. Calvin readily allowed that "the philosophers," without knowing God, can give us much accurate and worthwhile information concerning man's faculties and constitution (I, XV). Philosophy, however, cannot yield a true estimate of our worth and condition.

In any discussion of Calvin's views on how we can come to know ourselves and God, it is very important to

understand what he meant by knowing God, for his views on this point are both original and subtle. The Scholastics tended to equate knowing God with knowing truths about God. Calvin invariably regarded this as inadequate. He did not deny, indeed he insisted, that knowing God presupposes knowing *about* God. But in addition to this he always maintained that an essential aspect of our knowledge of God is our acknowledgment of his attitude toward us, especially his attitude of benevolence and love. Again, Calvin never equated acknowledging God's benevolence toward us with believing *that* God is benevolent toward us. Rather, acknowledging God's benevolence presupposes worshiping and obeying him. Thus, as Calvin uses the concept "knowing God," there is no knowledge of God apart from worship of, and obedience to, him. For this reason E. A. Dowey (1952) said that Calvin conceived of knowledge of God as existential. It may be added that Calvin held, as did many of the Scholastics, that what can be known about God is never his nature (*quid est*), but only what he is like (*qualis est*); and more specifically, what he is like toward us.

How is knowledge of God to be achieved? Calvin always held that knowledge of God can, in principle, be achieved by nourishing one's subjective awareness of deity and its will, with reflection on the structure of the objective world.

"There is," he said, "within the human mind, and indeed by natural instinct, an awareness of divinity [*sensus divinitatis*]" (I, iii, 1). Although this concept of a sense of divinity played a significant role in Calvin's thought, he spent little time elucidating it. Apparently he thought of it as yielding a rudimentary conviction of dependence on some Maker, as well as a numinous awareness of the glory and majesty of the Creator. In support of his conviction that this sense is universal in humankind, Calvin frequently quoted Cicero. It is this universally innate sense of divinity in humankind that, according to Calvin, accounts for the universality of religion in human society. It is a seed of religion (*semen religionis*). Religion is intrinsic to human life; it was not "invented by the subtlety and craft of a few to hold the simple folk in thrall" (I, iii, 2).

In Calvin's thought, conscience (*conscientia*), as a subjective mode of revelation, was closely related to the sense of divinity. Conscience too, he said, is part of the native endowment of all men, written "upon the hearts of all." Typically he spoke of it as a sort of knowledge whose object is God's will; or, equivalently, the difference between good and evil, the law of God, or the law of nature. Thus it is by virtue of conscience that man is aware of his responsibility—aware of the moral demands to which he is subject with respect to God and man. Calvin did not state with any exactitude the actual principles that all men know by virtue of conscience. He did say, however, that "that inward law … written, even engraved, upon the hearts of all, in a sense asserts the very same things that are to be learned from the [Decalogue]" (II, viii, 1); and he said that what the Decalogue requires is perfect love of God and of our neighbor.

The subjective awareness of divinity and of its will can be supplemented, Calvin taught, by reflecting on the structure of the external world and the pattern of history. "[God has] not only sowed in men's minds that seed of religion of which we have spoken but revealed himself and daily discloses himself in the whole workmanship of the universe. As a consequence, men cannot open their eyes without being compelled to see him" (I, v, l). At various times Calvin called the universe at large a book, a mirror, and a theater for the display of God's attributes—preeminently for the display of his goodness to us but also of his glory, wisdom, power, and justice. In the course of expounding his view that God can be known through his works, Calvin explicitly opposed the view that God can be known by speculation concerning his essence. It is by nourishing his sense of divinity and his conscience, with the contemplation of God's works, that man can in principle arrive at a knowledge of God.

SIN. It was Calvin's persistent teaching, however, that in fact no one does come to know God in the manner described above. The positive demands placed on all men by God's internal and external revelation are rejected, and this rejection results in an endless series of spurious religions. This resistance to God's demands is what Calvin identified as sin. Thus sin is not primarily ignorance about God; although such ignorance, or blindness, as Calvin often called it, will always be a consequence. Rather, Calvin viewed sin as an active willful opposition to God, as a positive refusal to acknowledge his demands of worship and obedience and as a deliberate alienation from him. Its prime characteristic is perversity, and its root is ordinarily pride and self-love.

Thus, being in sin is just the opposite of knowing God. Calvin, however, was quite willing to allow that a person who does not know God because he refuses to worship and obey him can still know or believe a variety of propositions about God that happen to be true. This explains what has, to so many readers, proved to be such an infuriating feature of Calvin's thought—his insistence, sometimes in adjacent sentences, that the pagans do not

at all know God but are not wholly ignorant of him. For example, Calvin, speaking of man's natural ability to know God, said, "the greatest geniuses are blinder than moles." In the very next sentence he said, "Certainly I do not deny that one can read competent and apt statements about God here and there in the philosophers" (II, ii, 18).

Not only was Calvin insistent that knowing or believing "competent and apt" propositions about God was not sufficient for knowing God; he was also profoundly convinced that man's proud refusal to worship and obey God leads him to resist acknowledging the truth about God. Sin, although primarily a matter of the will, infects man's reason as well. Perversity leads to blindness and distortion. Immediately after saying that the philosophers make competent and apt statements about God, Calvin added, "but these always show a certain giddy imagination… . They [the philosophers] saw things in such a way that their seeing did not direct them to the truth, much less enable them to attain it." Thus the consequence of man's willful alienation from God is not merely that he does not know God but also that his views about God are now so incomplete and distorted that nothing at all can be built on them. This is Calvin's judgment on natural theology.

It must be added that Calvin regarded the effects of sin as far more pervasive than have yet been indicated. Not only does sin disrupt man's relation to God; it thereby spreads corruption throughout the whole of human life. Of course, it does not impair our natural faculties as such. Calvin typically spoke of reason and will as man's chief faculties, and he held that the man in sin may be as intelligent and as capable of making decisions as the man who knows God. The corruption is to be found, rather, in the *use* we make of our native capacities.

Calvin maintained that if we are to state accurately what sin does to man's use of his native talents, we must distinguish between man's supernatural gifts, his abilities concerning heavenly things, and his natural gifts, his abilities concerning earthly things (II, ii, 12–13). The supernatural gifts comprise man's ability to know God, to worship him properly, and to obey him inwardly as well as outwardly. We have, however, been stripped of these gifts. The natural gifts pertain to matters of the present life, such as government, household management, all mechanical skills, and the liberal arts. Concerning these, said Calvin, our abilities have certainly not been destroyed. Not only are ancient law, medicine, and natural philosophy worthy of the highest admiration (II, ii, 15); but man, even in his estrangement from God, retains some sense of the laws that must be obeyed if human society is to be preserved. Man "tends through natural instinct to foster and preserve society. Consequently, we observe that there exist in all men's minds universal impressions of a certain civic fair dealing and order… . And this is ample proof that in the arrangement of this life no man is without the light of reason" (II, ii, 13). Calvin immediately added, however, that although man's abilities concerning earthly things have not been destroyed, they have been profoundly corrupted. In opposition to what he understood as the teaching of the Greek philosophers, he held that both reason and will have been gravely wounded; the mind "is both weak and plunged into deep darkness. And depravity of the will is all too well known" (II, ii, 12).

If man's natural gifts are to be healed and his supernatural gifts restored, his sin must be overcome; he must come to know God. We have already seen that for this purpose man's conscience, his sense of divinity, and his awareness of God's revelation in the objective world are all inadequate. Thus, if human life was to be renewed, it was necessary that God should choose some special means. This he did by revealing himself with special clarity in the history of the Jewish people, culminating in the life and words of Christ. When God leads man to respond to this revelation with faith, then man again knows God. Indeed, faith, consisting as it does in a clear knowledge about God coupled with proper worship and true obedience, is a certain sort of knowledge of God—that sort which focuses on Christ as interpreted in the Scriptures. Thus, in Calvin's thought there is never a contrast between faith in God and knowledge of God; rather, given man's prior perversity, faith is the only kind of knowledge of God available to men. Also, faith, in Calvin's teaching, is never understood in scholastic fashion as an assent to divinely revealed propositions. Rather, the object of faith is God as revealed in Christ.

SOCIAL AND POLITICAL TEACHINGS

Calvin's social and political theory has also proved most influential. Man, according to Calvin, is a creature of fellowship, created with tendencies that find their fulfillment in a variety of natural groupings, each concerned with a certain facet of man's life in society. One of these groupings is the church, another the state. Church and state are differentiated primarily by reference to their different tasks. The concern of the church is the spiritual realm, the life of the inner man; the concern of the state is the temporal realm, the regulation of external conduct. In regulating external conduct, the general aim of the state, in Calvin's view, is to insure justice or equity in society at large. This equity has two facets. Obviously the state

must enforce restrictive justice, but Calvin also believed that the state should secure distributive justice, doing its best to eliminate gross inequalities in the material status of its members.

It is the duty of the church to seek the welfare of the state, but equally it is the duty of the state to seek the welfare of the church. Thus, part of the state's duty is to promote piety; and Calvin, along with most of his contemporaries, regarded blasphemy as a civil crime. It was Calvin's view, however, that church and state ought to be structurally independent of each other. Church officials are not, by virtue of their office, to have any official voice in the state; and state officials are not, by virtue of their office, to have any official voice in the church.

Although he thought that the best form of government would vary with circumstances, Calvin quite firmly believed that the ideal government would be a republic in which those of the aristocracy who are competent to rule are elected by the citizenry, and in which power is balanced and diffused among a number of different magistrates. The magistrate has his authority from God. In a sense his authority is God's authority; for magistrates, Calvin said, are ministers of Divine justice, vicegerents of God. Thus the duty of the magistrate is to apply the law of God, implanted on the hearts of all and clarified in the Scriptures, to the affairs of civil society. To what extent and under what circumstances Calvin regarded civil disobedience as justified is a matter of debate. What is clear is that Calvin regarded the law of nature as in some sense a standard by which the decisions of the magistrate are to be judged, and at the same time he regarded revolutions which rip apart the entire fabric of human society as not to be condoned.

INFLUENCE

Both the theological and social views of Calvin have had an enormous influence throughout history. The Reformed, churches of the Continent and the Presbyterian churches of England adhered fundamentally to his thought, and the dominant theological thought of the American colonies was Calvinistic. In the eighteenth and nineteenth centuries the impact of Calvinism on society and theological thought suffered a decline, but the twentieth century saw a resurgence in Calvin's influence. In the early part of the century in the Netherlands, Abraham Kuyper led a revival of Calvinism in politics and education as well as in theology. And the so-called neoorthodox theology, represented by such figures as Karl Barth and Emil Brunner, not only was accompanied by a renewed interest in the writings of Calvin but also in large measure marked a return to the main patterns of Calvin's theological thought.

See also Barth, Karl; Brunner, Emil; Thomas Aquinas, St.

Bibliography

The standard edition of Calvin's works is that by J. W. Baum, E. Cunitz, E. Reuss, et al., in the *Corpus Reformatorum,* 59 vols. (Brunswick, Germany, 1863–1900). Most of these were translated by the Calvin Translation Society as *Works,* 48 vols. (Edinburgh: Calvin Translation Society, 1843–1855). An especially fine annotated translation is J. T. McNeill, ed., *Calvin: Institutes of the Christian Religion,* translated by F. L. Battles, 2 vols. (Philadelphia; Fortress Press, 1959). The classic study of Calvin's life is Émile Doumergue, *Jean Calvin; les hommes et les choses de son temps,* 7 vols. (Lausanne: G. Bridel, 1899–1927).

For discussions of the sources of Calvin's thought, see François Wendel, *Calvin: Sources et évolution de sa pensée religieuse* (Paris: Presses Universitaires de France, 1950), translated by Philip Mairet as *Calvin* (New York: Harper & Row, 1963); and Josef Bohatec, *Budé und Calvin* (Graz, Austria, 1940). A good discussion of Calvin's theology as a whole is Wilhelm Niesel, *Die Theologie Calvins* (Munich: C. Kaiser, 1938), translated by Harold Knight as *The Theology of Calvin* (Philadelphia: Westminster Press, 1956). The most adequate discussion of Calvin's views on the knowledge of God is E. A. Dowey, *The Knowledge of God in Calvin's Theology* (New York: Columbia University Press, 1952). Calvin's doctrine of man is treated in Thomas Forsyth Torrance, *Calvin's Doctrine of Man* (London: Lutterworth, 1949). Calvin's social and political thought is well discussed in André Biéler, *La pensée économique et sociale de Calvin* (Geneva: Librairie de l'universityé, 1959); and Josef Bohatec, *Calvins Lehre von Staat und Kirche* (Breslau: M. and H. Marcus, 1937). The finest survey of the history of Calvinism is J. T. McNeill, *The History and Character of Calvinism* (New York: Oxford University Press, 1954).

Nicholas Wolterstorff (1967)

CALVIN, JOHN [ADDENDUM]

During the past few decades much scholarly work has been done on John Calvin by theologians, historians, and others. Some of this work has recognized the ways in which Calvin, despite his rejection of Scholasticism and his ostensibly purely scriptural approach to theology, does in fact use philosophical argument in his work and does engage implicitly with philosophical issues even in his decisions not to proceed philosophically (see Helm 2004). But the context in which philosophers are most likely to have encountered Calvin's ideas since the early 1980s has been that of so-called Reformed epistemology.

This is an approach to the philosophy of religion pioneered mainly, though not exclusively, by philosophers associated with the Reformed (i.e., Calvinist) tradition. It is noteworthy for combining some of Calvin's ideas on the understanding of God with the antifoundationalism that has become more or less orthodox in the mainstream of secular epistemology since the 1950s and 1960s.

The Reformed epistemologists start with a rejection of evidentialism—the claim that one is only justified in holding a belief if one can provide a rational justification for it. Reformed epistemologists such as Alvin Plantinga and Nicholas Wolterstorff (1983) repudiate evidentialism in epistemology generally, and the epistemology of religion in particular. One cannot refute a skeptic by giving a nonquestion-begging proof of the reality (or even probability) of the external world or of other minds; but there is no rational requirement that one should do so. These beliefs are "properly basic" (Plantinga 1981); one does not form them on the basis of argument and is only rationally required to defend them if good reasons for doubt are given in some particular case. Similarly, according to the Reformed epistemologists, with belief in God.

This account has been worked out most elaborately by Plantinga (1993). He argues that what is needed to turn true belief into knowledge is warrant, an externalist notion that he explicates in terms of proper function. A belief is warranted if it is formed by the proper functioning of a subject's cognitive apparatus. The internalist notion of justification is given only a secondary role; one is justified in holding a belief if one can defend it against specific claims that it is false or unreasonable. Applying this account to religious belief, Plantinga (2000) draws heavily on Calvin's notion of the *sensus divinitatis*. People have been so created that their minds, when functioning properly, are naturally led to a belief in God. This is not through argument, any more than their belief in other minds or physical objects is formed by argument. The obvious disanalogy is that religious skepticism is a live issue in a way that other forms of skepticism are not. Here, Plantinga turns again to Calvin, to his doctrine of sin and its noetic effects. Those who disbelieve in God (or who have inadequate, confused, or half-hearted beliefs) do so because, ultimately, they are repressing or distorting the operations of the *sensus divinitatis* in themselves. (Plantinga compares this with the error theories of religion advanced by Marxists and Freudians, who argue that religious beliefs are self-deceiving evasions of reality.) This tendency to repression is universal; those who escape from it do so through the operations of divine grace. Calvin is again the main source for Plantinga's account of how the "internal instigation of the Holy Spirit" is necessary for one to be brought to belief in the specifically Christian doctrines of sin and redemption and thus to a true belief in God, which the sin-damaged *sensus divinitatis* cannot now achieve alone. Hence, Plantinga, while seeing non-Christian religions as evidence of the universality of the *sensus divinitatis*, rejects the idea that they can give their adherents a true or adequate knowledge of God.

It is striking that what is perhaps the most discussed late twentieth/early twenty-first-century development in religious epistemology is so deeply indebted to a theologian often thought of as nonphilosophical (although Plantinga's interpretation of Calvin has itself been questioned, for example, see Jeffreys [1997]). Plantinga denies that his account is Calvinist in any narrowly denominational sense, and indeed appeals to St. Thomas Aquinas as well as to Calvin. But as a Catholic commentator notes (Zagzebski 1993), the Reformed epistemologists' characteristic externalism, and their focus on the beliefs of individuals rather than of communities, are both, for better or worse, deeply rooted in the thought and sensibility of the Reformed tradition.

Bibliography

Helm, Paul. *John Calvin's Ideas*. New York: Oxford University Press, 2004.

Jeffreys, Derek. "How Reformed Is Reformed Epistemology? Alvin Plantinga and Calvin's *Sensus Divinitatis*." *Religious Studies* 33 (1997): 419–431.

Plantinga, Alvin. "Is Belief in God Properly Basic?" *Nous* 15 (1981): 41–51.

Plantinga, Alvin. *Warrant and Proper Function*. New York: Oxford University Press, 1993.

Plantinga, Alvin. *Warranted Christian Belief*. New York: Oxford University Press, 2000.

Plantinga, Alvin, and Nicholas Wolterstorff, eds. *Faith and Rationality: Reason and Belief in God*. Notre Dame, IN: University of Notre Dame Press, 1983.

Zagzebski, Linda. "Religious Knowledge and the Virtues of the Mind." In *Rational Faith: Catholic Responses to Reformed Epistemology*, edited by Linda Zagzebski. Notre Dame, IN: University of Notre Dame Press, 1993.

Anthony Rudd (2005)

CAMBRIDGE PLATONISTS

The Cambridge Platonists were a group of seventeenth-century thinkers, associated with Cambridge University, who drew on the neoplatonic tradition and contemporary philosophical developments in order to combat vol-

untarism, materialism, and determinism, and promote a tolerant and inclusive understanding of Christianity.

The core members of this school were active from the late 1630s through the 1680s, and were associated either with Emmanuel or Christ colleges. The central thinkers in the movement were Ralph Cudworth (1617–1688), Henry More (1614–1687), John Smith (1618–1652), and Benjamin Whichcote (1609–1683), their founding figure. Other close associates at Cambridge included Peter Sterry (1613–1672), John Worthington (1618–1680), George Rust (1626–1670), and Nathanael Culverwell (1618–1651). Beyond Cambridge, thinkers with connections to the school include John Norris (1657–1711), Joseph Glanvill (1636–1680), and Anne Conway (1631–1679). Leading latitudinarian divines, including Simon Patrick (1626–1707), John Tillotson (1630–1694), Gilbert Burnet (1643–1715), and Edward Stillingfleet (1635–1699), can also be considered disciples of the Cambridge Platonists.

Although the movement was centered in Emmanuel College, long a stronghold for Calvinistic Puritanism, it constituted a repudiation of what the Cambridge Platonists took to be a central feature of Calvinist thought, its voluntaristic understanding of morality as a creation of the divine will. Against this voluntarism, which the Cambridge Platonists perceived as offering an unacceptable account of God as arbitrary tyrant, they argued for a form of moral realism. Good and evil are "eternal and immutable"; moral distinctions are ontologically real and unchanging. Influenced by Renaissance neoplatonism (and thus interpreting Plato through the lens of Plotinus and later Christian Platonism), the Cambridge Platonists conceived of God as the Good, the form of forms. The goodness that God wills is an expression of God's own nature. Thus, while what is good is not good by virtue of being willed by God, eternal moral distinctions also do not serve as constraints on God's will.

The Cambridge Platonists declared themselves opposed to any separation of the realms of reason and faith, of the rational and the spiritual. By this they meant most fundamentally to assert that God's ways are fair, and in this sense reasonable. Rejecting the doctrine of predestination, they insisted that God's decrees are not arbitrary or unfathomable but are objectively just. The Cambridge Platonists were staunch defenders both of freedom of the will and freedom of conscience. If God is just in holding us responsible for our actions, then these actions must be up to us and freely chosen. Furthermore, faith is reasonable, and reason must be persuaded; it cannot be forced. On matters that reason cannot determine, the Cambridge Platonists advocated tolerance of a diversity of opinion. They worked for a policy of broad comprehension in the Church of England, minimizing core doctrines and emphasizing moral truths. Their theology thus resembles that of the Dutch Arminians, although arrived at independently.

Reason served for the Cambridge Platonists, as for so much of Renaissance neoplatonism, as a substantial link between the human and divine natures. Whichcote often wrote of reason as the "candle of the Lord." Discounting the impact of the Fall on human nature, the Cambridge Platonists were optimistic about the capacity of human persons to know God and eternal moral truths through reason. Human knowledge of various moral goods is a participation in God's own self-knowledge. Although there is a mystical aspect to the Cambridge Platonists' assertion that God is present within human persons through reason, they were critical of claims to private communications from God, which they condemned as "enthusiasm." Despite their emphasis on access to divine truth through reason, the Cambridge Platonists did not seek to undermine the authority of revealed truths. They did, though, tend to blur the boundaries between reason and revelation. So, for instance, they entertained the possibility that Plato's wisdom derived from Moses or other ancient Hebrews, and thus that pagan wisdom was indebted to revelation. But they also argued that pagan anticipations of revealed doctrines, including the trinity, might have derived from the powers of reason, God within.

If Puritan theology was the target against which the Cambridge Platonist movement took shape, the Platonists (particularly Cudworth and More) soon took on new foes, notably Thomas Hobbes. Like the Calvinists, Hobbes was a voluntarist, who made morality dependent on will. That for Hobbes morality was dependent on the will of the human sovereign rather than the will of God rendered his thought no less problematic in their eyes. Hobbes was also attacked for his materialism, which the Cambridge Platonists regarded as a dangerous form of atheism.

Initially, the Cambridge Platonists perceived René Descartes as a valuable ally against both materialism and the old scholastic Aristotelianism. The Cambridge Platonists were among the first English thinkers to read Descartes, and More carried on an extensive correspondence with him. Like Descartes, the Cambridge Platonists were dualists and they regarded a dualism of spirit and matter as indispensable for their defense of the spiritual realm against materialistic reduction. (More's friend and

pupil Anne Conway, author of *Principles of the Most Ancient and Modern Philosophy* [1692], parted ways with the Cambridge Platonists on this point, moving in the direction of a metaphysical monism).

The Cambridge Platonists came to think, though, that Descartes carried mechanistic explanations of the natural world too far. Arguing that matter is essentially passive and incapable of accounting for complex and orderly natural phenomena, they argued for a spiritual presence mediating between God and the physical universe. More termed this a *Spirit of Nature* or *Hylarchic Principle*, whereas Cudworth spoke of *Plastic Nature*. The eagerness to demonstrate the reality of immaterial substance reinforced in More and Glanvill a belief in witchcraft and a fascination with purported spiritual phenomena. Once seen as evidence of their credulity and backwardness, this feature of their thought is now understood as a further reflection of their support for the new experimental science. The Cambridge Platonists were familiar not only with the work of Hobbes, Descartes, and Benedict Spinoza, but also with Francis Bacon, Robert Boyle, and the Royal Society.

Whichcote's sermons were published by Anthony Ashley Cooper, third Earl of Shaftesbury. John Locke, Richard Price, and Thomas Reid were also indebted to the Cambridge Platonists, particularly Cudworth. Gottfried Wilhelm von Leibniz read both Cudworth and More, and Pierre Bayle critiqued Cudworth's *Plastic Nature*. Jonathan Edwards, John Wesley, Samuel Taylor Coleridge, and Matthew Arnold all admired the Cambridge Platonists. The lasting significance of the Cambridge Platonists resides in their success in carrying forward the insights of the tradition of Christian Platonism through a creative rapprochement with the philosophical revolution underway during their time. Within the heavily empiricist cast of English philosophy, they introduced a distinctive form of idealism.

See also Cudworth, Ralph; Culverwel, Nathanael; More, Henry; Whichcote, Benjamin.

Bibliography

Beiser, Frederick C. "Cambridge Platonism." In *The Sovereignty of Reason: The Defense of Rationality in the Early English Enlightenment*. Princeton, NJ: Princeton University Press, 1996.

Hutton, Sarah. "The Cambridge Platonists." In *A Companion to Early Modern Philosophy*, edited by Steven Nadler. Oxford, U.K.: Blackwell, 2002.

Patrides, C. A., ed. *The Cambridge Platonists*. Cambridge, MA: Havard University Press, 1970.

Rogers, G. A. J., J. M. Vienne, and Y. C. Zarka, eds. *The Cambridge Platonists in Philosophical Context: Politics, Metaphysics and Religion*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1997.

Jennifer A. Herdt (2005)

CAMPANELLA, TOMMASO

(1568–1639)

Tommaso Campanella, a Renaissance philosopher and scholar, was born at Stilo, in Calabria, Italy. At an early age he entered the Dominican order and devoted himself to the study of philosophy. In 1599 he was arrested by order of the Spanish government on charges of heresy and conspiracy. Although he never confessed to either charge, he was considered to be a dangerous subject and was kept in prison at Naples for twenty-seven years. Released in 1626, he was arrested again and arraigned before the Holy Office in Rome to stand trial for certain suspect propositions found in his works. After regaining his freedom, he spent some time at the Dominican monastery of Minerva in that city. In 1634, fearing further persecution, because of the suspicion that he might be involved in a new conspiracy, he followed the advice of Pope Urban VIII and fled to France, where he was befriended by Cardinal Richelieu and King Louis XIII. He died in the quiet of the Dominican monastery of Rue St. Honoré in Paris.

Campanella wrote a great number of books dealing with subjects ranging from grammar and rhetoric to philosophy and theology, from apologetics to politics, and from medicine to magic and astrology. He conceived of philosophy as an all-embracing science to which all other sciences must be referred as their ultimate source and foundation. No subsidiary science deals with all things as they are, but only as they appear, whereas philosophy, and especially metaphysics, deals with all things as they are and insofar as they are. Philosophy is an inquiry after the truth of both human and divine things, based on the testimony of God, who reveals himself either through the world of created things or by direct teaching. Consequently, nature and the Scriptures are the two codes on which philosophy must be built.

EPISTEMOLOGY

In his actual approach to philosophy, Campanella discussed first the possibility and reality of knowledge, thus anticipating a common trend among later thinkers. He was the first philosopher (antedating René Descartes) to assert the need of positing a universal doubt at the begin-

ning of his system and to state the principle of self-consciousness as the basis of knowledge and certitude. He distinguished between innate and acquired knowledge. Innate knowledge (*notitia innata*) is cognition through self-presence and belongs to the very essence of the soul; acquired knowledge (*notitia illata*) is the soul's cognition of external things. Innate knowledge is superior to, and more certain than, acquired knowledge; for the soul cannot be mistaken about what belongs to its nature. Knowledge of the external world can be obtained either by intuition or by abstraction. By intuition one grasps a thing immediately in its concrete reality, so that nothing of the object escapes the penetrating and all-embracing act of the intellect. By abstraction, one obtains only an indistinct and confused image of a thing. This image is what Campanella called the Aristotelian universal and is the object of both sense and intellect. The Platonic universal, on the contrary, is the idea as the formal cause of a thing and can be grasped exclusively by the intellect.

As to the essence and process of knowledge, Campanella gave a twofold explanation. A first explanation is contained in his early works and developed along the general lines of Bernardino Telesio's system. It represents his empirical approach to knowledge, which he reduced mainly to sensation and explained in terms of partial assimilation of the object known. This assimilation is made by contact between the knower and the sensible species of the object known. These species are neither the intentional species of the Aristotelians nor the corporeal images of Democritus. Although they may assume as many different forms as there are sensations, they are always something material that impinges on the senses and represents to a certain extent the external object.

A second and more advanced explanation of knowledge is what may be called the metaphysical approach from the standpoint of the soul as an essentially knowing nature. Here we meet Campanella's characteristic doctrine that to know is to be (*cognoscere est esse*). In this new approach, knowledge is still called sensation and assimilation, but the assimilation is carried so far as to mean a real transformation of the knower into the object known. This doctrine that to know is "being" or "to be" must not be understood in the idealistic sense of the absolute identity of object and subject. Campanella introduced a distinction between knowledge that a person has of himself in virtue of his own nature and knowledge that a person acquires from outside himself. Campanella called this the distinction between "innate" and "illate" knowledge. Both types of knowledge are said to belong to "being": But the former refers to knowledge of the original being of the knower, and the latter refers to the knowledge of being that is inferred by reasoning and is formally distinct from the being of the knower. In the first case, knowledge *is* the *esse*; in the second case, it becomes intentionally the *esse* in the possession of the extramental reality.

METAPHYSICS

For Campanella the object of metaphysics is "being," namely, whatever exists either within or outside our mind. He denied a real distinction between essence and existence in creatures, but admitted a real distinction between essence and extrinsic existence, or that type of existence that corresponds to the particular circumstances and environment wherein an essence happens to be in the physical world. All things, whether spiritual or material, consist ultimately, although in different degrees, of power, knowledge, and love as their transcendental principles. These are called "primalities" and are found in creatures as well as in God, of whom creatures are faint imitations. Whereas God is pure and infinite being, creatures are composites of finite being and infinite nonbeing. Being and nonbeing concur in making up finite things, not as physical components but as metaphysical principles. Just as a creature is essentially and necessarily a particular and limited entity, so it also is essentially and necessarily the nonbeing of all other things and of God himself.

PSYCHOLOGY

In psychology Campanella accepted the trichotomic theory, according to which man is a composite of three substances, body, spirit, and mind or *mens.* The spirit or sensitive soul is the corporeal principle that animates the body and serves as a link between body and mind. The mind or intellective soul is created and infused by God into the body already organized by the spirit; it is a spiritual substance and the form of the whole man. With the Platonists, Campanella defended the doctrine of a world soul, and developed the theory of universal animation by endowing all things with some kind of sensation.

PHILOSOPHY OF NATURE

Campanella was greatly influenced by Telesio's *De Rerum Natura,* which he defended against the attacks of G. A. Marta (1559–1628). He conceived of space as a primary and incorporeal substance having the capacity to receive all bodies. Space is the substratum of all things. In this space God placed matter, a body that is formless and inactive but capable of being molded into many forms, just as wax is acted upon by a seal. Matter is not pure

potency, as Aristotle taught, but has a reality of its own distinct from the form. This, in turn, is not a substantial principle of material beings and is only improperly called an act. In short, Campanella dismissed the Aristotelian hylomorphic theory and substituted for it Telesio's naturalistic doctrine of heat and cold as the active principles and matter as the passive principle of all material beings. He also rejected Aristotle's notion of time as measure of movement and claimed that time is not something ideal and subjective, but something real. Time is the successive duration of things having a beginning and an end. Or, more concretely, time is the thing itself considered in its successive duration through change.

ETHICS

Following Telesio, Campanella taught that man's supreme good consists in self-preservation. However, this must not be understood in a purely egoistic sense, but rather as the conservation of one's existence in God in the next life. Whereas God is his own supreme good and does not look to another being outside himself for his preservation, so that to be and to be happy are for him one and the same thing, man depends entirely on God for his own preservation. God is therefore the supreme good toward which man must direct all his acts and operations.

POLITICAL THEORY

Campanella advocated a universal monarchy with the pope as its supreme temporal and spiritual ruler. This ambitious but hardly realistic plan is described in the *Monarchia Messiae* (The Messiah's Monarchy) and represented the dream of his entire life. *Civitas Solis* (*The City of the Sun*), on the other hand, contains the scheme of a state modeled after Plato's *Republic* and Sir Thomas More's *Utopia,* where people, who live in the pure order of nature, organize themselves into an ideal society ruled by philosophers and share everything. Many of the ideas expressed in this work have some practical value, inasmuch as they contain the germs of social, political, and educational reforms that would be beneficial to the state. In this respect, Campanella may be considered as an original thinker and a forerunner of various modern theories and practices.

Bibliography

PRINCIPAL WORKS BY CAMPANELLA

Philosophia Sensibus Demonstrata, in defense of Telesio. Naples, 1591.

Monarchia Messiae. Jesi, 1633.

Atheismus Triumphatus. Paris, 1636.

Disputationum in Quatuor Partes Suae Philosophiae Reales Libri Quatuor. Paris, 1637.

Philosophiae Rationalis Partes Quinque. Paris, 1638.

Universalis Philosophiae, seu Metaphysicarum Rerum Iuxta Propria Dogmata Partes Tres, Libri 18. Paris, 1638.

Del senso delle' cose e della magia, edited by A. Bruers. Bari, 1925.

Epilogo Magno, edited by C. Ottaviano. Rome: Reale Accademia d'Italia, 1939.

Theologicorum Libri XXX, edited by R. Amerio. Florence and Rome, 1949–.

Civitas Solis. Frankfurt, 1623. Translated by W. J. Gilstrap as *The City of the Sun.* New York, 1952.

WORKS ON CAMPANELLA

Amerio, Romano. *Campanella.* Brescia: "La Scuola" Editrice, 1947.

Bonansea, Bernardine M. "Campanella as Forerunner of Descartes." *Franciscan Studies* 16 (1–2) (1956): 37–59.

Bonansea, Bernardine M. "The Concept of Being and Non-being in the Philosophy of Tommaso Campanella." *New Scholasticism* 31 (1) (1957): 34–67.

Bonansea, Bernardine M. "Knowledge of the Extramental World in the System of Tommaso Campanella." *Franciscan Studies* 17 (2–3) (1957): 188–212.

Bonansea, Bernardine M. "The Political Thought of Tommaso Campanella." In *Studies in Philosophy and the History of Philosophy,* edited by John K. Ryan, Vol. II, pp. 211–248. Washington, DC: Catholic University of America Press, 1963.

Bonansea, Bernardino M. *Tommaso Campanella, Renaissance Pioneer of Modern Thought.* Washington, DC: Catholic University of America Press, 1969.

Corsano, Antonio. *Tommaso Campanella.* Bari: Laterza, 1961.

Di Napoli, Giovanni. *Tommaso Campanella, filosofo della restaurazione cattolica.* Padua: CEDAM, 1947.

Headley, John M. "On the Rearming of Heaven: The Machiavellism of Tommaso Campanella." *Journal of the History of Ideas* 49 (1988): 387–404.

Headley, John M. *Tommaso Campanella and the Transformation of the World.* Princeton, NJ: Princeton University Press, 1997.

Szentpeteri, Marton. "Il Transilvano: The Image of Zsigmond Bathory in Campanella's Political Thought." *Bruniana and Campanelliana* 9 (1) (2003): 217–225.

Bernardine M. Bonansea, O.F.M. (1967)
Bibliography updated by Tamra Frei (2005)

CAMPBELL, NORMAN ROBERT
(1880–1949)

Norman Robert Campbell, the English physicist and philosopher of science, was educated at Eton. From Eton he went as a scholar to Trinity College, Cambridge, and became a fellow there in 1904. From 1903 to 1910 he also

worked as a research assistant at the Cavendish Laboratory, whose director, the celebrated J. J. Thomson, became the most important inspiration of his scientific work. In 1913 he became an honorary fellow for research in physics at Leeds University, but he left this post after the war and from 1919 to 1944 was a member of the research staff of the General Electric Company.

The writers who seem to have influenced him most are Ernst Mach and Henri Poincaré, apart from classical authors such as William Whewell, John Stuart Mill, and W. S. Jevons. On the other hand, such philosophers as Bertrand Russell and Alfred North Whitehead came too late to have much effect on him; the main outlines of his thought developed during the first decade of the century, and there are only occasional references to their writings.

Campbell exhibited the very rare combination of competence in both physics and philosophy, but while he preferred to think of himself primarily as an experimental physicist, it is as a philosopher of science that he made his mark. This point is brought out in the writings of F. P. Ramsey, R. B. Braithwaite, and Ernest Nagel, although these concentrate largely on the formal parts of Campbell's doctrines and pay scant attention to the more valuable contributions that he made to certain methodological ideas, particularly that of analogy. These philosophical views, shaped by Campbell's actual experiences and ideas as a physicist and expositor of physical theories, were meant to be construed as answers to intellectual pressures and problems that confronted him in the years that saw the rise of the twentieth-century atomic theory on the one hand and relativity and quantum mechanics on the other. In philosophy of science, his most important contributions were in the fields of the logic of theory construction and (to a lesser extent) the principles of physical measurement.

PHILOSOPHY OF THEORY CONSTRUCTION

Campbell's views were stated in systematic form for the first time in a popular book, *The Principles of Electricity.* Thereafter they were developed, with minor changes of emphasis and greater attention to the nature of "mathematical theories," in *Physics: The Elements.* In contrast with the usual textbook approach, his views were deeply interwoven with, and at times even explicitly discussed in, his more formal scientific treatises.

CONCEPTS AND IDEAS. Campbell distinguishes sharply between the laws and theories of a science. In the case of laws, the constituent terms (Campbell calls them concepts) designate entities whose magnitudes may be determined more or less directly by instrumental means; they are not unlike what later came to be called operational concepts. The explanatory part of theories, the hypotheses, involve terms that Campbell calls ideas. These lack the instrumental relations of concepts, for a variety of reasons that Campbell does not always clearly distinguish.

Sometimes the ideas refer to the unobservable infrastructure of a physical system, as in the case of the atoms and electrons of modern electrical theory or, more properly (as Campbell points out), to their adjectival aspects, such as their mass, velocity, and momentum. At other times, the ideas pertain to such interstructural devices as Michael Faraday's lines of force, or the carriers of the transfer of electrical and optical phenomena, such as light waves, light corpuscles (photons) or even the "aether," considered the substantival carrier of electromagnetic energy. (Infrastructural entities are unobservable in a sense different from interstructural ones, but the question is controversial.) A third case in which theories are said to involve unobservables is that of geological and evolutionary theories. And there is yet another case, for Campbell denominates certain notions "ideas" because they involve an amount of idealization and abstraction to which no physical entities could correspond. The most frequent and important cases are those ideas which involve infinitesimals, such as the differential coefficients in James Clerk Maxwell's equations or François Marie Charles Fourier's theory of heat.

It follows from the nature of ideas that the hypotheses in which they occur are not directly testable. Their function consists merely in systematically relating a set of corresponding laws, and, through extensions of the theory, in foreshadowing further laws and experiments. This foreshadowing is sometimes negative, for when the ideas are too narrowly framed, they demand not only extension but also the formulation of additional concepts and theories.

"DICTIONARY" OF A THEORY. Since the ideas of the hypotheses lack operational meaning, and since their deductive development can, in the first place, yield only statement forms containing either ideas or combinations of them, it is necessary to add certain rules (a kind of "dictionary") that will coordinate the ideas with those operational concepts which occur in the laws to be explained. Of course, not all ideas need dictionary entries. In the beta-ray theory, for instance, the velocity, *v*, of the hypothetical electrons *means* "the quantity that is

defined by the relation $F = e[X+(v \cdot H)]$." This expression, however, is a hypothesis in Campbell's sense because v never occurs either alone or in combination in the testable derivations at all.

"MATHEMATICAL" THEORIES. All this provided Campbell with a means of distinguishing so-called mathematical theories from nonmathematical ones. In the former, each and every idea is separately coordinated with a corresponding concept by means of a dictionary entry. It follows that whether a theory is of the mathematical type depends partly on historical accidents: Maxwell's theory *became* a mathematical theory only *after* Heinrich Rudolf Hertz's experiment had demonstrated the existence of the displacement current.

Nonetheless, ideas so far have no meaning apart from their use in hypotheses and their coordination with concepts. In the mathematical cases this is often forgotten, but in the nonmathematical cases this fact is more difficult to overlook. Because of the lack of independent significance of ideas, Campbell held that a theory is not a real explanation unless certain additional requirements are satisfied. One of his reasons for this view was that it is always possible to construct an indefinite number of hypotheses that would account for a set of laws. In the case of mathematical theories, the additional element of consolidation that Campbell suggests is the regulative feature of simplicity and aesthetic elegance—for instance, through symmetrical arrangements of the parts of a theory. (Thus, it was the introduction of Maxwell's displacement current into the original equations of André Marie Ampère and Faraday that produced a symmetrical set of equations regarding the relations between the electrical and magnetic phenomena for the case of open circuits.) Furthermore, the hypotheses are not entirely arbitrary because their ideas mirror the corresponding concepts of the laws. There is, according to Campbell, a sort of analogy between ideas and concepts (*Physics: The Elements,* p. 141).

ANALOGY. Analogy plays a more central role in the case of the nonmathematical theories. As we have seen, their ideas frequently cannot be clarified at all by the concepts that occur in the laws. According to Campbell, it is an analogy of the hypotheses and their ideas with corresponding laws and concepts of some testable field of science that imparts the missing element of significance and logical strength to the theory. It follows that analogies are not merely aids to the establishment of theories; "they are an utterly essential part of theories, without which theories would be completely valueless and unworthy of the name" (ibid., p. 129).

Campbell's point is that "a theory is not a law" (ibid., p. 130); that hypotheses are, from the nature of the case, never directly testable; and, hence, that their addition to the corpus of scientific knowledge would make no difference to science at all if it were not for some additional features that make the hypotheses significant. He dismisses the fact that they supply a systematic relation between the laws of the theory on the grounds that an infinity of such hypotheses can be constructed.

Campbell's positive grounds for the necessity of analogies are of various kinds. The fundamental reason is that since hypotheses are not directly testable but are only instruments for deductive development, possessing a purely formal content, they lack the sort of meaning required for genuine explanatory power: Only analogy can supply this. Another ground of a more heuristic nature is that analogies aid in the extension of theories, especially when a new field is grafted onto the dictionary of an existing theory (as when optical conceptions were added to Maxwell's generalization of the electrical theories of Ampère and Faraday).

As mentioned, however, analogy must be supplemented by additional criteria, which are clearly needed for dealing with mathematical theories. These criteria are largely derived from Campbell's actual experience with the theories with which he had been dealing in his physical textbooks. In addition to simplicity and aesthetic elegance, there is "simplification in our physical conceptions," such as was produced by the early theories of Faraday, Thomson, and Hendrik Antoon Lorentz. Campbell insists on the importance of such regulative conceptions precisely because "scientific propositions are [not] capable of direct and irrefutable proof." An additional criterion is the "anticipative force" of a theory—for instance, the suggestiveness of Faraday's lines in the direction of the existence of electromagnetic radiation, of a motion that is displaced in time, with a given velocity, in empty space.

Finally, another regulative criterion is that of importance, or depth, of the ideas involved. This is invoked particularly in those cases where analogy is barely a relevant consideration, as in such mathematical theories as Maxwell's, or Albert Einstein's special theory of relativity.

METHODOLOGICAL CONTRIBUTIONS

Campbell's clear account of the logical structure of a theory, with its hypotheses, laws, and dictionary, offers an

elegant means of formalizing the place of ideas (theoretical concepts) within theories. He emphasizes also the logical gap between hypotheses and laws even in cases where its existence had previously been practically overlooked—the mathematical theories. He uses this fact to question Mach's preference for such theories (called phenomenological by Mach), on the grounds that they employ hypotheses and hypothetical ideas just like any other theory. (Whether this does sufficient justice to the difference between the two types of theories must be left an open question.) The theoretical nature of such substantival entities as atoms and electrons seems to differ from that of lines of force on the one hand and, say, from the entropy functions on the other, in deeper ways not caught by Campbell's criteria of ideas.

The fact that the systematizing power of hypotheses is an insufficient criterion of their truth or explanatory power introduces the remaining feature of Campbell's doctrine—such regulative notions as the existence of a strong analogy, of simplicity, symmetry, anticipative force, and, finally, of importance. The most interesting of these is analogy, which in the end emerges as a metaphysical device in terms of which to formulate the special aspect of those theories that involve unobservables. The "absolute necessity" for an analogy is the result of the emasculation of the semantic power of hypotheses, coupled with the consideration that this emasculation entails the introduction of a special constraint that prevents such hypotheses from being mere arbitrary formulas.

THEORY OF MEASUREMENT

The second part of *Physics: The Elements* is a detailed discussion of the principles of physical measurement; this, like most of Campbell's ideas, was already contained in embryo in *The Principles of Electricity* (Ch. 2). His interest in measurement is not altogether removed from his main philosophical preoccupations mentioned so far. Just as he was concerned with a clear delineation of laws from theories, he was equally firm in stating the differences as well as the relations between laws and definitions. In *Measurement and Calculation* Campbell defines measurement "as the assignment of numerals to present properties in accordance with … laws." Thus, every measurable property must have a definite order; the systems to be measured must be capable of "addition," but what operation is considered "addition" must be carefully specified in a given situation; and whether the resultant quantities yield consistent measurements is a matter for lawlike experience. Campbell points out that the specification in question is usually tacitly adopted ab initio and is, indeed, often suggested by theory and the relevant analogy. Hence, he believes that "no new measurable quantity has ever been introduced into physics except as the result of the suggestion of some theory" (*The Principles of Electricity,* p. 41).

See also Ampère, André Marie; Braithwaite, Richard Bevan; Einstein, Albert; Faraday, Michael; Fourier, François Marie Charles; Hertz, Heinrich Rudolf; Jevons, William Stanley; Mach, Ernst; Maxwell, James Clerk; Mill, John Stuart; Nagel, Ernest; Poincaré, Jules Henri; Quantum Mechanics; Ramsey, Frank Plumpton; Relativity Theory; Russell, Bertrand Arthur William; Whewell, William; Whitehead, Alfred North.

Bibliography

Campbell's works include *Modern Electrical Theory* (Cambridge: Cambridge University Press, 1907; 2nd ed., 1913); *The Principles of Electricity* (London, 1912); *Physics: The Elements* (Cambridge, U.K.: Cambridge University Press, 1920), which was reprinted as *Foundations of Science* (New York: Dover, 1957); *Measurement and Calculation* (London: Longmans, Green, 1928); and *Photoelectric Cells,* written with Dorothy Ritchie (London: Pitman, 1929; 3rd ed., 1934).

As supplementary chapters of *Modern Electrical Theory,* the following monographs by Campbell were published in the Cambridge Physical Series: *Series Spectra,* Suppl. Ch. 15 (Cambridge, 1921); *Relativity,* Suppl. Ch. 16 (Cambridge, 1923); and *The Structure of the Atom,* Suppl. Ch. 17 (Cambridge, 1923).

For works with detailed references to Campbell, see Brian D. Ellis, *Basic Concepts of Measurement* (Cambridge: Cambridge University Press, 1966), Chs. 4, 5, and 8; Mary B. Hesse, *Models and Analogies in Science* (London, 1963), Ch. 1; and George Schlesinger, *Method in the Physical Sciences* (London: Routledge and Paul, 1963), Ch. 3, Sec. 5.

Gerd Buchdahl (1967)

CAMUS, ALBERT

(1913–1960)

Albert Camus, the French novelist and essayist, was born in Mondovi, Algeria, and was educated at the University of Algiers. From 1934 to 1939 he was active writing and producing plays for a theater group he had founded in Algiers. About the same time he began his career as a journalist, and in 1940 he moved to Paris. During the German occupation of France, Camus was active in the resistance movement, and after the liberation of Paris he became the editor of the previously clandestine newspaper *Combat*. His literary fame dates from the publication

in 1942 of his first novel, *L'étranger* (*The Stranger*), and an essay titled *Le mythe de Sisyphe* (*The Myth of Sisyphus*). During the immediate postwar period Camus was deeply involved in political activity, and his name was for a time closely associated with that of Jean-Paul Sartre and with the existentialist movement. In 1947 he published a second major novel, *La peste* (*The Plague*), and, in 1951, *L'homme revolté* (*The Rebel*), an essay on the idea of revolt. The latter book provoked a bitter controversy between Camus and Sartre, which ended with a severance of relations between them. In 1957 Camus was awarded the Nobel Prize for literature. His last major work was *La chute* (*The Fall*), a novel that appeared in 1956. In 1960 Camus was killed in an automobile accident.

Although Camus studied philosophy for a number of years at the University of Algiers, he was not a philosopher in any technical or academic sense. Nevertheless, virtually all his literary work was deeply influenced by philosophical ideas, and in two major essays, *The Myth of Sisyphus* and *The Rebel*, he undertook a more or less systematic exposition and defense of the moral attitudes that had in each case found expression in his novels and plays. *The Myth of Sisyphus* can thus be regarded as in some sense a philosophical commentary on *The Stranger*, and *The Rebel* has clear affinities with *The Plague*. There can be no doubt that there are profound differences between the views set forth in these two essays. Camus's philosophical career was essentially a movement away from the nihilism of *The Myth of Sisyphus* toward the humanism of *The Rebel*. Ideas that had been present in his work from the beginning, in one form or another, were to retain their place there; but he progressively revised his views of their relative importance within the moral life.

Although Camus's name is often associated with contemporary European phenomenology and existentialism, there is no evidence that he was ever deeply influenced by, or very much interested in, the doctrines of Edmund Husserl or Martin Heidegger or even Sartre; and on occasion he expressed himself as having distinct reservations with respect to existentialism as a philosophy. In fact, his philosophical thought was formed on much more traditional models. His deepest interest was in those great figures in the Western philosophical tradition—among them Socrates, Blaise Pascal, Benedict de Spinoza, and Friedrich Nietzsche—whose lives and personalities were all reflected in their philosophizing. If he came, as he did, to reject the exaggerated claims that philosophers have made for human reason and subscribed to many of the criticisms that contemporary existentialists have made of the classical tradition, he continued to regard the striving of the great thinkers of the past to achieve a total conception of reality and of the human relation to the world as reflecting one of the deepest human aspirations and to view its inevitable failure as marking a crisis in man's relation to himself.

On the other hand, Camus does not appear to have had any theoretical interest in the analysis of philosophical problems. His interest in philosophy was almost exclusively moral in character; when he had come to the conclusion that none of the speculative systems of the past could provide any positive guidance for human life or any guarantee of the validity of human values, he found himself in the situation that he describes in *The Myth of Sisyphus*. This essay is ostensibly a consideration of the problem of suicide, which Camus describes as the only serious philosophical problem. The question he asks is whether it makes any sense to go on living once the meaninglessness of human life is fully understood and assimilated. Camus gives a number of somewhat different formulations of what this meaninglessness or "absurdity" comprises. At bottom, it is the failure of the world to satisfy the human demand that it provide a basis for human values—for our personal ideals and for our judgments of right and wrong.

It is very important for an understanding of Camus's point of view to see how closely he thought ordinary moral attitudes are dependent upon metaphysical belief in some kind of congruence between human values and the nature of reality. The external supports on which the validity of moral distinctions rested in the past were, of course, primarily religious in character; but Camus held, as do many others, that with the decline of religious belief in the modern period a number of secular religions—in particular, Hegelian and Marxist historicism—have attempted to tie values to reality by means of a postulated schedule of historical development that guarantees their eventual realization. In *The Myth of Sisyphus*, Camus presupposes, without very much argument, that none of these interpretations of reality as value-supporting can survive critical scrutiny; the tenability of any purposive or evaluative attitude on the part of human beings—the only moral beings—is thus called into question. It is this isolation of the human being as an evaluative and purposive being in a world that affords no support to such attitudes that Camus calls the absurdity of the human condition.

Camus maintained that suicide cannot be regarded as an adequate response to the experience of absurdity. The reason he gives is that suicide deals with absurdity simply by suppressing one of the two poles—the human

being and the "world"—that together produce the tension described above. Suicide is thus an admission of incapacity, and such an admission is inconsistent with that human pride to which Camus openly appeals. Indeed, he goes so far as to say that "there is nothing equal to the spectacle of human pride." Only by going on living in the face of their own absurdity can human beings achieve their full stature. For Camus, as for Nietzsche, whose influence at this stage of Camus's thought is very marked, the conscious espousal of the metaphysical arbitrariness of human purpose and action transforms nihilism from a passive despair into a way of revolting against and transcending the world's indifference to man.

It is evident that in *The Myth of Sisyphus* Camus believed that absurdity, in the sense of recognition and acceptance of the fact that there are no metaphysically guaranteed directives for conduct, could by itself generate a positive ethic. In particular, the ideal of human fraternity was connected with Camus's heroic nihilism on the grounds that to accept oneself as the sole guarantor of one's own values would necessarily involve accepting a principle of respect for other human beings. It is here, however, that Camus encountered a very serious difficulty. He found it necessary to show by means of examples just what the specific implications for conduct of his doctrine of absurdity are and also make it plausible that these implications are consistent with the humanistic ideal to which he as an individual is clearly devoted. In *The Myth of Sisyphus*, however, the specimens that are offered of the mode of life appropriate to the "absurd" man bear only a rather remote affinity to that ideal or, for that matter, to any general social ethic. Camus did not demonstrate satisfactorily either that the kind of life that followed from an acceptance of nihilism bore any clear relation to his own moral ideals or that a life dedicated to these ideals could be adequately motivated by an acceptance of absurdity.

What is clear is that Camus, from the beginning, regarded certain responses to absurdity as morally unacceptable. In his "Letters to a German Friend" (1943–1944), he interpreted Nazism as one reaction to the very nihilistic vision of the world that he himself had come to accept. He then went on to condemn it in the severest terms for its denial of human fraternity. Even at this stage in the development of his thought, Camus insisted that an authentic revolt against the human condition had to be a revolt in the name of the solidarity of man with man.

In the character of Meursault, the "hero" of *The Stranger*, this tension between Camus's nihilistic vision and his ethical demands becomes particularly clear. Meursault is presented as a man characterized by the moral equivalent of achromatic vision. Although he is not at all given to philosophical reflection, he views the whole conventional human apparatus of moral distinctions, of justice and of guilt, as a kind of senseless rigmarole with no basis in reality. He stands, in fact, outside the whole moral world in a peculiar state that Camus describes as "innocence," apparently because in a world that affords no transcendental sanction for human judgments of right and wrong there can be no real guilt. His relationship to his mother and to his mistress are devoid of feeling, and he eventually kills an Arab for no particular reason. But at the very end of the novel, after Meursault, facing execution, has burst into a rage against a priest who tries to persuade him to accept the reality of his guilt and the possibility of redemption, there is a long semipoetic passage in which he declares his love of the world and its sensuous immediacy and speaks tenderly and almost lovingly of his fellow men and of their common fate, which he shares. As a number of critics have noted, there is nothing in the novel that prepares one for this passage. Camus, however, clearly wishes to persuade us that these two aspects of Meursault's character are not just consistent but intimately related to one another; but again he experienced difficulty in showing how a positive ethic of human fraternity can be generated by a nihilistic attitude toward all values.

There can be little doubt that in the years immediately following the publication of *The Stranger* and *The Myth of Sisyphus* Camus substantially revised his view of the moral significance of value-nihilism. Increasingly, it was the injustice and cruelty of man to man that aroused Camus to action; by comparison with the hideous but remediable evils of human society, the cosmic injustice of the human condition seems to have lost some of its obsessive hold on his mind. Like many of the existentialists, Camus still tried to present these two revolts—the revolt against the human condition and the revolt against human injustice—as essentially continuous with one another. Nevertheless, he came to feel that the relationship between these two revolts had been misconceived and that this misconception was at the heart of twentieth-century totalitarianism, to which he was as resolutely opposed in its communistic as in its Nazi version. Camus gradually came to believe that the reason for the extraordinary miscarriage of the Soviet revolution was that the revolutionary tradition had its roots in a revolt against the human condition as such, and that such a revolt can never lead to human fraternity but leads instead to a new enslavement of man by man. This radical revision of his

earlier views found its full expression in Camus's second main philosophical essay, *The Rebel*.

The Rebel begins with a consideration of the problem of murder or, more exactly, with the problem of political justification for the killing of human beings. For Camus, political action is essentially violent revolt, and it thus inescapably raises the question of whether one has the right to take the life of another human being. Camus's answer is that taking a human life is inconsistent with true revolt since, as he now makes clear, that revolt involves the implicit assertion of a supraindividual value, the value of human life. It is not altogether clear how this rejection of violence is to be interpreted, but it is interesting to note the approval that Camus expresses in his play *The Just* (1950) of the Russian terrorist Kaliaev who murders the Grand Duke Serge but insists that he himself pay for his act with his life in order to affirm the moral inadmissibility of murder. In any case, the revolt that Camus still advocates in *The Rebel* is presented there as ethically inspired from its inception. He rejects, however, what he now calls "metaphysical revolt," which he sees as a radical refusal of the human condition as such, resulting either in suicide or in a demonic attempt to depose God and remake the world in the image of man. Its deepest motive is not a love for humankind but a desire to destroy the world as it is. The order it attempts to impose on the new world it constructs is informed by no ethically creative principle because, as Camus now declares, nihilism can yield no such principle. A nightmare state of power for power's sake is the ultimate fruit of metaphysical revolt.

In order to substantiate this thesis, Camus reviews the intellectual history of the past two hundred years and discusses in detail a number of poets, philosophers, and practicing revolutionaries whom he regards as the chief fomentors of metaphysical revolt. Among them are the Marquis de Sade, Max Stirner, Nietzsche, le Comte de Lautréamont, Baron de Saint-Just, and Sergei Nechaiev, to mention only a few. G. W. F. Hegel and Karl Marx are assigned a central role in the construction of a view of history and of the state that exempts man from all moral controls and that proposes as the only valid ideal man's total mastery of his own fate. The two political revolutions that Camus thinks were inspired by the ethos of metaphysical revolt are the French and the Russian, although the Nazi "revolution" represents some of the same tendencies in even purer form. Camus considers none of the modern revolutions that did not eventuate in political terrorism, and he makes no attempt to evaluate or even consider other kinds of explanation of the revolutions that he does discuss. As many critics have remarked, the apocalyptic character of the historical tableau that he presents is in good part due to a principle of selection that seems to reflect a personal predilection for extreme or crisis situations rather than any objective assessment of the real influence that the representatives of metaphysical revolt may have had on the course of events.

Camus's novel *The Plague*, which appeared four years before *The Rebel*, gives clear indications of his reevaluation of nihilism. The plague that descends on Oran symbolizes not just the Nazi occupation of France or even totalitarianism as a political system but all of the many forms that injustice and inhumanity can assume. A variety of reactions to this "plague" is presented; but it is Dr. Rieux, the organizer of the "sanitation squads" that fight the plague, who represents Camus's ideal of moral action. Rieux is not inspired by any dream of a total conquest of evil. Instead, his conception of himself is modest and limited; throughout the struggle he retains his sense of humanity and his capacity for love and for happiness. The doctor is in fact what many have said Camus aspired to be, a kind of "saint without a God."

If *The Rebel* and *The Plague* represent—as they seem to do—Camus's mature position, it would appear that this position differs from traditional nonreligious humanism mainly by virtue of the terminology of revolt that Camus retained even after he had so thoroughly moralized his conception of revolt as to make most of the normal connotations of that term inapposite. As he himself says in *The Rebel*, the true significance of nihilism is negative; it clears the ground for new construction but by itself provides no principle of action. As such it survives in Camus's view of the moral world mainly as a prophylactic against the kind of mystification, religious or metaphysical, by which a man tries to rid himself of his radical contingency and confer upon himself a cosmic status that makes it easier for him to be a human being. Camus was a pitiless critic of all such forms of shamming, and he was convinced that their general tendency was to enable their practitioners to evade the responsibility that goes with moral self-ownership and to confirm them in their inhumanity to their fellow men. Nihilism would seem, in Camus's final view, to be a kind of immunizing experience, although one with very considerable dangers of its own, by virtue of which one is enabled to grasp the ideal of human fraternity in its pure form without the entanglements of ideology and doctrine by which it has so often been disfigured. Camus's attitude toward life is thus, at bottom, simply a stubborn moral integrity and a deep sympathy with his fellow men, to which the somewhat meretricious rhetoric of revolt adds very little. At

the same time, however, it must be conceded that the absence or unavailability of absolute values, whatever these might be, remains for Camus anything but trivial, and it pervades the atmosphere of the humanistic ethic that he erected in their place.

The work of Camus's last years reinforces one's impression that an essentially nonmetaphysical and strongly moralistic humanism was his final view of life. He drew away more and more from direct political action; his refusal to side unambiguously with the Algerian rebels brought him the bitter reproaches of many former associates, among them Sartre. In 1960 in *Réflexions sur la peine capitale* ("Reflections on the Guillotine"), Camus argued that society does not have the right to put its criminals to death, and one wonders in what circumstances Camus would have regarded war as morally defensible. Finally, in *The Fall*, he seems to have abandoned political and social action entirely in favor of a conception of evil that no longer situates it in unjust social institutions or in the terms on which man is permitted to exist but in the very heart of man himself. The protagonist, Clamence, is a man whose interior corruptness is concealed from the world—and for a long time from himself—by a life of philanthropy and active sympathy for his fellow men. He is, in fact, a sort of monster whose ultimate self-knowledge leads him to create a sense of guilt and unworthiness in others by advertising his own corruption. In this way he again feeds his obsessive need for superiority, which was the real motive of his earlier philanthropy. It is not justifiable to impute the unrelieved pessimism of this novel to Camus personally, or to suggest, as some have, that he had accepted the doctrine of original sin; but there can be little doubt that his treatment of the character of Clamence is indicative of a further shift in the locus of the struggle between good and evil. The shift, broadly speaking, is one that emphasizes our inner complicity with evil and our lack of the kind of innocence that Camus had always claimed for humanity. Whether this strain would have been developed further in Camus's thought if he had lived longer is a question to which there can be no answer.

See also Ethics, History of; Existentialism; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Humanism; Husserl, Edmund; Life, Meaning and Value of; Literature, Philosophy of; Marx, Karl; Nietzsche, Friedrich; Nihilism; Pascal, Blaise; Phenomenology; Sartre, Jean-Paul; Socrates, Spinoza, Benedict (Baruch) de; Stirner, Max; Suicide.

Bibliography

WORKS BY CAMUS

L'etranger. Paris: Gallimard, 1942. Translated by S. Gilbert as *The Stranger*. New York: Knopf, 1946.

Le mythe de Sisyphe. Paris: Gallimard, 1942. Translated by J. O'Brien as *The Myth of Sisyphus*. New York: Knopf, 1955.

Lettres à un ami allemand. Paris: Gallimard, 1945. Translated by J. O'Brien as "Letters to a German Friend," in *Resistance, Rebellion and Death*. New York: Knopf, 1961.

La peste. Paris: Gallimard, 1947. Translated by S. Gilbert as *The Plague*. New York: Knopf, 1948.

Les justes. Paris: Gallimard, 1950. Translated by S. Gilbert as "The Just." In *Caligula and Three Other Plays*. New York: Knopf, 1958.

L'homme revolté. Paris: Gallimard, 1951. Translated by A. Bower as *The Rebel*. New York: Knopf, 1954.

La chute. Paris: Gallimard, 1956. Translated by J. O'Brien as *The Fall*. New York: Knopf, 1957.

Réflexions sur la peine capitale. Paris, 1960. Translated by J. O'Brien as "Reflections on the Guillotine." In *Resistance, Rebellion and Death*. New York: Knopf, 1961.

Carnets. Paris: Gallimard, 1962. Translated by P. Thody as *Notebooks 1935–42*. New York: Knopf, 1963.

BIOGRAPHIES

Lebesque, Morvan. *Portrait of Camus: An Illustrated Biography*. New York: Herder and Herder, 1971. Originally published in French, 1963.

Lottman, Herbert R. *Albert Camus: A Biography*. New York: Doubleday, 1979.

Todd, Oliver. *Albert Camus. A Life*. New York: Knopf, 1997.

CRITICAL STUDIES OF CAMUS

Brée, G. *Camus*. New Brunswick, NJ: Rutgers University Press, 1959.

Cruickshank, J. *Albert Camus and the Literature of Revolt*. London: Oxford University Press, 1959. Contains a detailed bibliography.

Knapp, Bettina L., ed. *Critical Essays on Albert Camus*. Boston: G.K. Hall, 1988.

McBride, Joseph. *Albert Camus: Philosopher and Littérateur*. New York: St. Martin's Press, 1992.

Rhein, Philip H. *Albert Camus*. Boston: Twayne, 1989.

Sprintzen, David. *Camus: A Critical Examination*. Philadelphia: Temple University Press, 1988.

Thody, P. *Albert Camus: A Study of His Work*. London: Hamilton, 1957.

Frederick A. Olafson (1967)
Bibliography updated by Thomas Nenon (2005)

CAN

What can be true or can be done varies with the meaning of "can." As far as philosophy is concerned, the important senses of this word ("could," past indicative) fall into five major groups. For convenience these groups, most of which are distinguished in *Webster's Third New Interna-*

tional Dictionary, may be singled out as the "can" of ability, of right, of inclination or probability, of opportunity, and of possibility.

"CAN" OF ABILITY

The "can" of ability has at least three subsenses: (1) to have the skill—"He can speak five languages or paint lifelike portraits"; (2) to have the requisite mental or physical power—"He can solve difficult problems, invent remarkable machines, or foretell the future" or "He can swim a mile or do one hundred push-ups"; (3) to have the requisite strength of character—"He can resist anything but pleasure, pass up a free drink, or bear criticism of his books."

"CAN" OF RIGHT

The "can" of right, which is often used interchangeably with "may," has at least four subsenses: (1) logically or axiologically can—"Equivalent formulas can be interchanged, *salva veritate,* in any extensional context" or "From this we can reasonably infer …"; (2) can in virtue of custom, agreement, law, and so on—"One can be prosecuted for saying that" or "An ambulance can disregard traffic lights"; (3) permission-giving "can"—"You can borrow my car if you'd like"; (4) be permitted by conscience or feeling—"I can condone no willful act of destruction" or "I can accept electrocution but not hanging."

"CAN" OF INCLINATION OR PROBABILITY

Examples of the "can" of inclination or probability are "I was so angry that I could have killed him" and "That car could hardly have made a trip across the desert."

"CAN" OF OPPORTUNITY

"He could have played chess had he known how," "Come in here where we can talk," and "The traffic was so heavy that I could not cross" illustrate the "can" of opportunity.

"CAN" OF POSSIBILITY

The "can" of possibility has at least five subsenses: (1) consistency with knowledge—"For all that I know, Jones could have been the one"; (2) whether it is possible for someone (compare with the "can" of opportunity)—"Can you get away for lunch?" (3) the "can" of physical possibility—"If such-and-such has to happen, then it cannot fail to happen" or "A man, properly equipped, can survive indefinitely in outer space"; (4) the "can" of logical possibility (compare with the logical or axiological use of the "can" of right)—"Nothing can be red all over and green all over at the same time"; (5) conditional possibility (logical or physical)—"If the conclusion of a valid argument is false, not all of the premises can be true" or "In a deterministic system everything that can occur is necessitated by something else."

CAN AND FREE WILL

Because the field of philosophical perplexity is virtually limitless, any one of the "cans" listed above is a possible source of trouble to the philosopher. Nevertheless, several of them (especially the "cans" of ability, opportunity, and possibility), have proved exceptionally potent in bewitching the philosophical imagination, mainly in connection with the age-old problem of free will. This problem is partly generated by the conviction that a man can be said to perform an action freely only if he did not have to perform it but he could have done something else instead. A conviction of this kind tends to generate a problem because if the metaphysical thesis of determinism is intelligible, tenable, and applicable to human actions, it becomes doubtful whether it is ever true that a man can do anything other than what he does do, at least in one of these three basic senses of "can."

"CAN" OF ABILITY. How the ability senses of "can" bear on the free will issue has received perhaps the largest share of attention in the recent literature, possibly because questions about a man's abilities are often so crucially relevant in moral contexts. Yet the decisive points about abilities in this connection are easily stated. In all of the subsenses of the "can" of ability, there is an essential distinction between the possession of an ability and the exercise of that ability. To show that a person lacks an ability is more complicated than to show that he does not exercise it. A failure to perform a certain action implies that a man lacks the corresponding ability only if both he wants, wills, intends, or chooses to perform that action and his failure to perform it occurs in relevantly normal conditions. This fact has tempted philosophers (for instance, P. H. Nowell-Smith) to analyze "He can" (in the sense of ability) as meaning "He will if … ." Important difficulties with such hypothetical analyses have been pointed out by Austin and others, but it has not been shown that there is anything wrong with the line of thought that prompted these analyses—namely, that our use of "can" in this sense is built on the idea that a man need not do what he can do and that in order to find out what he can do, we must find out what he will do if, in relevantly normal conditions, he wants, wills, intends, and

so forth to do certain things. This line of thought is not, moreover, inconsistent with determinism, since determinism does not imply that if, under appropriate conditions, I wanted and were to try to perform an alternative action, I should certainly fail. On the contrary, it is presumably only because a measure of determinism does hold that my trying, in certain circumstances, to perform a particular action is likely to meet with consistent success.

"CAN" OF OPPORTUNITY. Although the truth of determinism does not imply that if a man performs a certain action, he could not (in the sense of the "can" of ability) have done otherwise, it might still be claimed that he would not, under these conditions, have the opportunity to do otherwise and, thus, that he could not do otherwise in the sense of the "can" of opportunity. But this claim is simply false, since in the ordinary sense of "opportunity" one can be said to have the opportunity to do many things that one is not presently doing, whether or not determinism holds. As the examples of the "can" of opportunity indicate, "having the opportunity to do *X*" does not mean anything like "being in a situation in which nothing physically essential for one's performance of *X* is lacking," which the claim in question seems to suppose (for more on this point see Taylor, *Metaphysics*). On the contrary, to have the opportunity to do something requires only that one be in a situation such that if, roughly speaking, one wanted to do it, it would be reasonable to expect that one would be successful in doing it if one were able to do it (that is, could do it in the sense of ability). And such a situation would normally be lacking in many things essential, in the required sense, to one's performing that action. Not only might it lack the essential interest or even ability on one's own part, but it might also fail to involve the means that one would have to take in order to accomplish that action if it were at all complex—for instance, walking across the room in order to grasp the vase that one "has the opportunity" to break, throwing it toward the floor with sufficient force, and so on.

"CAN" OF POSSIBILITY. In spite of all this, it still seems possible to argue that, given determinism, a man cannot do other than what he does do in the sense that any alternative action on his part is physically impossible. A claim of this sort is, however, false if taken literally, since what is physically possible *simpliciter* need be consistent only with the laws of nature, not consistent with the laws of nature *and* certain initial conditions. If, however, the claim is to be taken in a slightly different way—namely, that it is *conditionally* physically impossible for the man to perform some other action—then it is entirely unexceptionable if the thesis of determinism is tenable and applicable to human actions. The reason for this is simply that the notion of conditional possibility is a technical one, definable by reference to determinism: Roughly, "*A* is conditionally physically possible" is by definition equivalent to "Nothing has happened that physically determines non-*A*."

Because one is to make sense of "conditional physical possibility" by reference to determinism or something like it, it is clear that the hard-fought question whether determinism rules out human freedom is not the question whether determinism rules out the conditional possibility of a man's doing other than what he does do. There is, in fact, little that is controversial about the last question; it gets an analytic "Yes." What is controversial is the question whether the sense of "can" involved in the morally relevant query "Can he do otherwise?" is to be understood as the "can" of conditional possibility. For if, as both libertarians and sophisticated fatalists seem to think, this "can" *is* of basic moral significance, then free actions are possible only if determinism is false, untenable, or inapplicable to human actions. If, on the other hand, this sense of "can" is *not* the one that does concern us or should concern us when in a moral context we wonder whether a man can do other than what he does do—the opinion of the "reconcilers" of the empiricist tradition—then there is, perhaps, no incompatibility between determinism and human freedom after all.

NORMATIVE "CAN." How is this basic question about the "can" in the morally crucial use of "He can do otherwise" to be resolved? Only a few, admittedly feeble, hints can be given here. First, the idea that this "can" is that of conditional possibility seems extremely dubious, since this sense of the word is pretty clearly a contrived one, not mentioned even in unabridged dictionaries and thus hardly one that, like the "can" of ability and opportunity, is likely to be used in the familiar, everyday, morally compelling assessment of free, responsible actions. Second, the less heavy-handed and therefore far more tempting claim—that it is at any rate naive or unreasonable to describe an action as free if it is conditionally impossible for the agent to have done otherwise—seems very unsatisfactory when it is carefully pressed. For one thing, to think of free actions as differing from unfree ones in being conditionally undetermined is to make the very notion of a free action practically useless, since any question that might arise about the freedom of a given act would presumably then have to be settled by a fairly

hopeless hunt for causes in the jungles of neurology. For another thing, to conceive of free actions in this way is to sever their ties with those complex principles of personal responsibility that incline us to excuse, rather than emphatically condemn, the kindly old parson who (we might imagine) suddenly, spontaneously, and without cause wills to, and does, brain the infant he is baptizing. The last point really seems to go to the heart of the matter: To conceive of free actions as conditionally physically indeterminate actions is to conceive of them in too naturalistic a way. After all, the very identity of an action—think of promising or murdering—is determined not just by the physical movements involved but also by a complex system of rules, laws, and so forth. Since it is the application of such rule concepts that distinguishes actions involving the same physical movements—murder and defensive or punitive acts—the basic vocabulary of action descriptions is essentially normative to a very large extent. (Actually, the vocabulary of action description is "intentional" in a way in which "scientific" language presumably is not.) Because the "can" in the morally crucial claim "He can do otherwise" plainly belongs to the family of words specifically used in connection with human actions, there is an inescapable force to the claim, made by many contemporary philosophers, that to identify this sense of the word with "conditional physical possibility" is to confuse a practical, largely normative "can" with an aseptic, scientific, theoretical one and thus to misconceive drastically the purpose, point, and import of the familiar, nontechnical statement "His action was done freely."

See also Determinism and Freedom; Possibility.

Bibliography

Aune, Bruce. "Abilities, Modalities, and Free Will." *Philosophy and Phenomenological Research* 23 (1963): 397–413.

Austin, J. L. "Ifs and Cans." In *Philosophical Papers,* edited by J. O. Urmson and G. J. Warnock. Oxford: Clarendon Press, 1961.

Baier, Kurt. "Could and Would." *Analysis* Supp. Vol. (1963): 20–29.

Edwards, Paul, and Arthur Pap, eds. *A Modern Introduction to Philosophy.* Glencoe, IL: Free Press, 1957. See Ch. 1; this book should also be consulted for more detailed bibliographical information.

Hook, Sidney, ed. *Determinism and Freedom in the Age of Modern Science.* New York: New York University Press, 1958.

Melden, A. I. *Free Action.* London, 1961.

Morgenbesser, Sidney, and James Walsh, eds. *Free Will: A Book of Readings.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

Nowell-Smith, P. H. *Ethics.* London: Penguin, 1954.

Pears, D. F., ed. *Freedom and the Will.* New York: St. Martin's Press, 1963.

Raab, F. V. "Free Will and the Ambiguity of 'Could.'" *Philosophical Review* 64 (1955): 60–77.

Taylor, Richard. *Metaphysics.* Englewood Cliffs, NJ: Prentice-Hall, 1963. See Chs. 4–5.

Bruce Aune (1967)

CANTOR, GEORG
(1845–1918)

Georg Cantor, a mathematician who created set theory and a corresponding theory of transfinite numbers, revolutionized mathematics at the end of the nineteenth century with his ideas about the infinite, which were to be of profound significance not only for mathematics but for philosophy and many allied disciplines as well.

He was born on March 3, 1845, in St. Petersburg, Russia, to Georg Woldemar Cantor, a successful merchant and the son of a Jewish businessman from Copenhagen, and Maria Anna Böhm, who came from a family of notable musicians and was a Roman Catholic. But Cantor's father, raised in a Lutheran mission, was a deeply religious man and passed his own strong convictions on to his son. Later in life, Cantor's religious beliefs would play a significant role in his steadfast faith in the correctness of his controversial transfinite set theory, just as his mother's Catholicism may have made him particularly amenable to the substantial correspondence he undertook with Catholic theologians over the nature of the infinite from a theological perspective.

EARLY MATHEMATICAL STUDIES

Cantor received his doctorate in 1868 from the University of Berlin, where he had studied with Leopold Kronecker, Ernst Eduard Kummer, and Karl Weierstrass. His dissertation was devoted to number theory, as was his *Habilitationsschrift*. When Cantor began teaching as an instructor at the University of Halle, among his colleagues there was Eduard Heinrich Heine. Heine had been working on problems related to trigonometric series, and he urged Cantor to take up the challenging problem of whether or not, given an arbitrary function represented by a trigonometric series, the representation was unique. In 1870 Heine had established the uniqueness of such representations for almost-everywhere continuous functions, assuming the uniform convergence of the trigonometric series in question. Cantor succeeded in establishing increasingly general versions of the uniqueness theorem in a series of papers he published between 1870 and 1872, the most remarkable of which showed that even if an infi-

nite number of exceptional points for the representation were allowed, the uniqueness could still be shown if such infinite sets of "exceptional" points were distributed in a particular way. Such sets of exceptional points constituted what Cantor called sets of the first species.

An infinite set of points P was said to be of the first species if its set of limit points P' was finite; if not, then P' must contain an infinite number of points and also have a derived set, the second derived set of P, P''. If for some finite number v the v^{th} derived set P^v contains only a finite number of points, then its derived set will be empty, that is, $P^{v+1} = \emptyset$. It was for such first-species sets that he was able to establish the uniqueness of trigonometric series representations, even though there were an infinite number of exceptional points. Transfinite set theory would arise from Cantor's later consideration of point sets of the second species, all of whose derived sets were infinite. From these Cantor would eventually generate an endless hierarchy of what he came to call transfinite ordinal, and later their corresponding cardinal, numbers.

THE REAL NUMBERS

Cantor realized that to define the structure of point sets of the first species unambiguously required a rigorous definition of the real numbers, which he approached in terms of fundamental, convergent sequences of rational numbers in his last paper on trigonometric series of 1872. In the same year Richard Dedekind introduced his own rigorous definition of the real numbers in terms of "Dedekind cuts." Both approaches are concerned with the continuity of the real numbers in general, a subject that was to haunt Cantor for the rest of his life. In particular, he succeeded in proving just a few years later, in 1874, that the set of all real numbers was in fact nondenumerably infinite, that is, of a distinctly higher order of infinity than denumerably infinite sets like the whole, rational, or algebraic numbers. This fact soon led to the articulation of one of Cantor's most famous problems: his continuum hypothesis, that the infinite set of real numbers R is the next higher order of infinite sets following denumerably infinite sets like the set of all natural numbers N. Cantor became especially interested in the question of whether or not point sets of two and higher dimensions might furnish examples of increasingly infinite orders of infinity, something he answered negatively in 1877. This was another of Cantor's important early results, his proof (though faulty) of the invariance of dimension; the first correct proof was published by L. E. J. Brouwer in 1911.

Between 1879 and 1883 Cantor wrote a series of articles that culminated in an independently published monograph devoted to the study of linear point sets, *Grundlagen einer allgemeinen Mannigfaltigkeitslehre: Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen* (Foundations of a general theory of sets: A mathematico-philosophical investigation into the theory of the infinite). In addition to introducing such concepts as everywhere-dense sets, he showed that whereas everywhere-dense sets were necessarily of the second species, first-species sets could never be everywhere-dense.

TRANSFINITE NUMBERS

In his series of papers on linear point sets, and in the *Grundlagen*, Cantor introduced his new concept of transfinite numbers. At first, these were limited to the transfinite ordinal numbers that he generated from the point sets of the second species that he had introduced in 1872. Considering the entire sequence of derived sets P^v, none of which was empty (i.e., every derived set P^v contained an infinite number of limit points): $P', P'', \ldots, P^v, \ldots$, Cantor defined the intersection of all these sets as P^∞. This was an infinite set that in turn led to the next derived set $P^{\infty+1}$. If this set were infinite, and in fact every derived set thereafter, this led to an endless hierarchy of further infinite derived sets: $P', P'', \ldots, P^v, \ldots, P^\infty, P^{\infty+1}, \ldots, P^{\infty+v}, \ldots, P^{2\infty}, \ldots$

At first, Cantor only regarded the superscripts as "infinite symbols," but early in the 1880s he began to distinguish these indexes as numbers independently of point sets of the second species. By the time he wrote the *Grundlagen* in 1883, these infinite symbols had become transfinite ordinal numbers.

CONTROVERSY AND CRITICISM

Cantor understood that his new ideas would be controversial, and his work had already met with criticism, especially from Kronecker, his former teacher at the University of Berlin. Cantor was so concerned about the possible objections to his new ideas that he undertook a detailed analysis of the subject historically, which served his strategy in the *Grundlagen* to present a detailed analysis of the foundations of transfinite set theory from both a philosophical and theological perspective. It was in the *Grundlagen* that he made one of his most famous statements, that "the essence of mathematics lies precisely in its freedom" (1996, p. 182). As Cantor later confided to the mathematician David Hilbert, this statement was inspired by the negative criticism Kronecker had made of set theory and was a call for open-mindedness among mathematicians, especially in dealing with new and novel ideas proposed by younger mathematicians. But the

opposition mounted by Kronecker served a useful purpose in stimulating Cantor's own philosophical reaction and his determination to provide the soundest possible foundations, both mathematically and philosophically, for transfinite set theory.

What Cantor did in the *Grundlagen* was to present the transfinite ordinal numbers as a direct extension of the real numbers. But because he generated these infinite real numbers as abstractions from sets of points, he rejected the possibility of there being actually infinitesimal numbers. He also knew that an important property of the transfinite ordinal numbers was their noncommutativity, that is:

$$2+\omega = (1, 2, a_1, a_2, \ldots, a_n, a_{n+1}, \ldots) \neq (a_1, a_2, \ldots, a_n, a_{n+1}, \ldots, 1, 2) = \omega+2,$$

$$2\omega = (a_1, a_2, a_3, \ldots; b_1, b_2, b_3, \ldots) \neq (a_1, b_1, a_2, b_2, a_3, b_3, \ldots) = \omega 2.$$

Such distinctions brought new insights to the differences between finite and infinite sets. For finite sets and their corresponding ordinal numbers, addition and multiplication were commutative; infinite sets were more interesting because their corresponding ordinal numbers and transfinite arithmetic were not commutative. Cantor expected that understanding such differences would not only explain the seemingly paradoxical nature of the infinite but would also answer some of the long-standing objections to the infinite that historically had been so persuasive to mathematicians and philosophers alike.

TRANSFINITE CARDINALS AND CANTOR'S ALEPHS

Although the *Grundlagen* offered a systematic presentation of Cantor's transfinite ordinal numbers, there was no mention of his best-known innovation: the transfinite cardinal numbers, or alephs. Indeed, nowhere in the *Grundlagen* was there any indication that the power of an infinite set was to be equated with the concept of a transfinite cardinal number, a step he first took in a lecture he delivered at Freiburg in September 1883. Over the next decade he used a number of different notations for transfinite cardinal numbers, but did not decide on a definite symbol until Giulio Vivanti, an Italian mathematician who was writing an introductory monograph on set theory, asked Cantor about notation. Only then did he finally choose the Hebrew aleph for the transfinite cardinal numbers. In "Beiträge zur Begründung der transfiniten Mengenlehre" (Contributions to the founding of the theory of transfinite numbers) he designated the least transfinite cardinal number as $\aleph_0$.

It was also in "Beiträge" that Cantor offered an algebraic interpretation of his continuum hypothesis, based on his proof of 1891 that given any infinite set *P*, the set of all its subsets was of a higher power than *P*. Since the cardinality of the set of all real numbers could be written as $2^{\aleph_0}$, and if $\aleph_1$ was the next largest cardinal following $\aleph_0$, then the continuum hypothesis could now be expressed as $2^{\aleph_0} = \aleph_1$. Cantor hoped that with this new algebraic formulation of the hypothesis, he would soon manage to produce a proof that the power of the real numbers was indeed equal to $\aleph_1$. He never succeeded in doing so, for reasons that only became apparent in the twentieth century, thanks to the results of Kurt Gödel (who established that the continuum hypothesis was consistent with the basic axioms of Zermelo-Fraenkel set theory) and Paul Cohen (who showed, on the contrary, that the continuum hypothesis was independent of the same axioms), which meant that it was possible to conceive of consistent set theories in which Cantor's continuum hypothesis did not hold.

Cantor's last major publication appeared in two parts in the journal *Mathematische Annalen* in 1895 and 1897. "Beiträge" not only offered a complete account of both his transfinite ordinal and cardinal numbers but also his theory of order types, which investigated in detail the different properties of the sets of natural, rational, and real numbers, respectively. The well-ordered set of integers, taken in their natural order, he designated (ω; the set of rational numbers in their natural order, which were everywhere-dense but not continuous, he designated η; sets like the real numbers that were continuous he designated by the order-type θ. But the result he hoped to achieve in "Beiträge" but failed to produce, namely, proof of his continuum hypothesis, remained illusive.

CANTOR'S MANIC DEPRESSION

Much has been written about Cantor's unfortunate history of mental illness, which some writers have linked with the heavy criticism of Cantor's transfinite set theory from Kronecker. But recent studies suggest that what Cantor suffered from was manic depression, which would have afflicted him regardless of the controversies surrounding his mathematical work (see Grattan-Guinness 1971, Dauben 1979, Charraud 1994). Whereas the earliest serious breakdown occurred in 1884, as Cantor was encountering his first disappointments in trying to prove the continuum hypothesis (for a detailed account of what happened, see Schoenflies 1927), the manic depression

became more serious as he grew older, and after 1900 he spent increasingly long periods under professional care, often at the Nervenklinik in Halle. Also, following the first attack in 1884, Cantor began to take up interests other than mathematics, including the idea that Francis Bacon was the real author of writings attributed to William Shakespeare and that Joseph of Arimathea was the natural father of Jesus. Cantor also began an extensive correspondence with Catholic theologians, and even wrote to Pope Leo XIII directly, in hopes that a correct understanding of the infinite mathematically, in terms of his transfinite set theory, would help the church avoid making any incorrect pronouncements on the subject, especially where the absolutely infinite nature of God was concerned, which Cantor took to be consistent with but wholly different from the concepts of transfinite set theory.

The mathematician Eric Temple Bell (1986) offers a Freudian analysis of Cantor's relationship with his father, whose initial opposition to Cantor's wish to become a mathematician Bell takes to be the source of his son's later mental problems; more recently, Nathalie Charraud (1994), a French psychoanalyst, examined the records of Cantor's treatment at the neurological clinic in Halle and offers a different, Lacanian assessment of the role Cantor's father played in his son's life. Equally important in understanding Cantor's tenacious defense of his controversial set theory is the role that religion played with respect to the transfinite numbers, which he took to have been communicated to him from God directly. For details of how his religious convictions and periods of manic depression may actually have played important, supportive roles in the battle to establish transfinite set theory as a fundamental part of modern mathematics, see Joseph Warren Dauben (2005).

One final aspect of Cantor's career as a mathematician deserves brief mention, because he was primarily responsible for the creation of the Deutsche Mathematiker-Vereinigung (German Mathematical Society), of which Cantor was elected its first president in 1891. He was also instrumental in promoting the idea of the first International Congresses of Mathematicians, beginning with Zürich in 1897, and then Paris in 1900 (Dauben 1979, pp. 163–165).

THE PARADOXES OF SET THEORY

To conclude with an assessment of Cantor's significance for philosophy, he was above all responsible for making the infinite a central part of modern mathematics. From the time of the Greeks, Zeno's discovery of the paradoxes of motion and Aristotle's opposition to the concept of completed infinities (as opposed to the potential infinite) led most mathematicians to avoid using the infinite in their work. Cantor faced the subject head-on and argued that there was nothing inherently contradictory in considering actually infinite collections of point sets or the infinite sets of integers, rational, and real numbers as unified, completed objects of thought. His contemplation of these eventually led to his development of transfinite set theory, transfinite arithmetic, and his fundamental concepts of transfinite ordinal and cardinal numbers. His greatest contribution was understanding the roles these played in establishing a proper foundation for mathematics, which he approached essentially on formalist terms. Consistency, for Cantor, was the only test a new mathematical theory needed to pass before he considered it legitimate as a subject for study and application.

When Cantor himself first realized the contradictions inherent in trying to decide the ordinal number of the set of all transfinite ordinal numbers, or the cardinality of the set of all transfinite cardinal numbers, his solution was to simply ban such "collections" from mathematics, saying they were too large to be considered legitimately as "sets." But as others like Cesare Burali-Forti and Jules Richard began to consider the antinomies of set theory, Bertrand Russell discovered a logical paradox at the heart of set theory involving the set of all sets that are not members of themselves. One solution to this dilemma was advanced by Ernst Zermelo, who sought to axiomatize set theory in such a way that the paradoxes would be excluded. Further developments along such lines were made by Russell and Alfred North Whitehead in their monumental *Principia Mathematica*; alternative axiomatizations were also advanced by Abraham Fraenkel and John von Neumann, among others.

By the end of his life, Cantor was a mathematician honored by the Royal Society with its Copley Medal for his outstanding contributions to mathematics. He was also granted an honorary degree by the University of St. Andrews (Scotland). Today, the highest award conferred by the German Mathematical Society is a medal honoring its first president, Georg Cantor.

See also Infinity in Mathematics and Logic; Set Theory.

Bibliography

Cantor, Georg, and Richard Dedekind. *Briefwechsel Cantor-Dedekind*, edited by E. Noether and J. Cavaillès. Paris: Hermann, 1937.

WORKS BY CANTOR

Gesammelte Abhandlungen mathematischen und philosophischen Inhalts, edited by Ernst Zermelo. Berlin: Springer, 1932.

"Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen." In *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, edited by Ernst Zermelo, 115–118. Berlin: Springer, 1932. This was originally published in the *Journal für die reine und angewandte Mathematik* in 1874.

"Foundations of a General Theory of Manifolds: A Mathematico-Philosophical Investigation into the Theory of the Infinite." Translated by W. B. Ewald. In *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, edited by W. B. Ewald, 2:878–920. New York: Oxford University Press, 1996. This was originally published under the title "Grundlagen einer allgemeinen Mannigfaltigkeitslehre: Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen" in 1883.

"Beiträge zur Begründung der transfiniten Mengenlehre." In *Contributions to the Founding of the Theory of Transfinite Numbers*. Translated by Philip E. B. Jourdain. Chicago: Open Court, 1915.

WORKS ABOUT CANTOR

Bell, Eric Temple. "Paradise Lost: Georg Cantor." In *Men of Mathematics*, 555–579. New York: Simon and Schuster, 1986.

Charraud, Nathalie. *Infini et inconscient: Essai sur Georg Cantor*. Paris: Anthropos, 1994.

Dauben, Joseph Warren. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Cambridge, MA: Harvard University Press, 1979.

Dauben, Joseph Warren. "The Battle for Cantorian Set Theory." In *Mathematics and the Historian's Craft. The Kenneth O. May Lectures*, edited by Micahel Kinyon and Glen van Brummelen. New York: Springer Verlag, Canadian Mathematical Society Books in Mathematics, 2005.

Fraenkel, Abraham A. "Georg Cantor." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 39 (1930): 189–266.

Grattan-Guinness, Ivor. "Towards a Biography of Georg Cantor." *Annals of Science* 27 (1971): 345–391.

Hallett, Michael. *Cantorian Set Theory and Limitation of Size*. Oxford, U.K.: Clarendon Press, 1984.

Lavine, Shaughan. *Understanding the Infinite*. Cambridge, MA: Harvard University Press, 1994.

Meschkowski, Herbert. *Probleme des Unendlichen: Werk und Leben Georg Cantors*. Braunschweig, Germany: Vieweg and Sohn, 1967.

Purkert, Walter, and Hans Joachim Ilgauds. *Georg Cantor*. Leipzig, Germany: Teubner, 1985.

Purkert, Walter, and Hans Joachim Ilgauds. *Georg Cantor, 1845–1918*. Basel, Switzerland: Birkhäuser, 1987.

Schoenflies, Arthur. "Die Krisis in Cantor's mathematischem Schaffen." *Acta Mathematica* 50 (1927): 1–23.

Joseph W. Dauben (2005)

CAPREOLUS, JOHN

(c. 1380–1444)

John Capreolus, a French Dominican theologian, was born in Rodez. He studied at the University of Paris, receiving the magistrate in theology in 1411. Later he taught in Dominican houses of study at Toulouse and Rodez and came to be recognized as the "Leader of the Thomists" (*Princeps Thomistarum*). His chief work is *Defensiones Theologiae D. Thomae* (Defenses of the theology of St. Thomas). This is the first commentary that considers the *Summa Theologiae* more important than Thomas Aquinas's *Commentary on the Sentences,* a view which has persisted in later Thomism. The *Defensiones* is historically useful for its information on scholastic philosophical controversies of the fourteenth century and the views of John Duns Scotus, John of Ripa, Peter Aureolus, and Durandus of Saint-Pourcain. Capreolus' contributions to philosophy are in the field of metaphysics. On the then central question of the relation between essence and existence, he taught that they are distinguished as two different beings (an extreme real distinction) and used the terminology of Giles of Rome (*esse essentiae* and *esse existentiae*) to express his position. Capreolus regarded essences as eternal and uncreated entities, not efficiently produced by God but subject only to divine formal causality. On the other hand, he stressed the importance of existence in treating personality (divine and human), teaching that personality is the very subsistence of the act of existing (*esse actualis existentiae,* see *Defensiones,* Vol. V, pp. 105–107). Where other thinkers required some sort of formal or modal constituent of the person, Capreolus demanded nothing more than the act of existing as an intelligent individual nature. He taught that the intrinsic principle that individuates bodies is matter marked by quantity (*materia signata*), as did Thomas, but Capreolus insisted that the quantification must be actual (under definite dimensions) and not indeterminate (*Defensiones,* Vol. III, pp. 200–241).

See also Aristotelianism; Duns Scotus, John; Durandus of Saint-Pourçain; Peter Aureol; Thomas Aquinas, St.

Bibliography

WORKS BY CAPREOLUS

Defensiones Theologiae D. Thomae. Edited by Paban-Pègues, 7 vols. Turin, 1900–1908. There are no known English translations of Capreolus's work.

WORKS ON CAPREOLUS

Grabmann, M. "J. Capreolus, O.P., der Princeps Thomistarum und seine Stellung in der Geschichte der Thomistenschule." *Divus Thomas* (Freiburg) 22 (1944): 85–109, 145–170.

Wells, N. J. "Capreolus on Essence and Existence." *Modern Schoolman* 38 (1960): 1–24.

Vernon J. Bourke (1967)

CARD, CLAUDIA

(1940–)

Claudia Card, an American philosopher, has published work on a wide range of philosophical topics but is best known for her contributions to ethics and feminist philosophy. Card began her academic career at Harvard University, where she received a PhD for her dissertation on theories of punishment. Currently the Emma Goldman Professor of Philosophy and Senior Fellow at the Institute for Research in Humanities, Card has been a professor of philosophy at the University of Wisconsin at Madison since 1966. Card is also an affiliate professor in women's studies, environmental studies, Jewish studies, and lesbian, gay, bisexual, and transgendered studies. Extraordinarily active in various philosophical societies, Card was named the Distinguished Woman Philosopher of the Year in 1996 by the Eastern Society for Women in Philosophy. The author of numerous scholarly books and journal articles, Card has also given a number of media interviews and served on many editorial boards. Her research interests include feminist philosophy, lesbian ethics, and evil.

Card's work is striking not just for the range of areas of philosophy to which it makes a contribution but also for the connections it draws between them. In ethics Card's work is notable for showing how questions in mainstream moral philosophy are tied to pressing political issues. In *The Unnatural Lottery: Character and Moral Luck* (1996), for example, Card defends the concept of moral luck and explores how a person's opportunity to act morally is affected by such variables as gender, race, social class, and sexual identity. Card asks about the implications of moral luck for attributions of moral responsibility and in the course of her examination discusses the problems faced by survivors of childhood sexual abuse. Another notable feature of Card's contribution to philosophy is her attraction to difficult, troubling, and important questions. Her work on moral luck falls into this category, as does her later work on evil. Card's monograph *The Atrocity Paradigm: A Theory of Evil* (2004) explores the relationship between evil and other concepts/practices such as forgiveness, toleration, and hate. Card asks what distinguishes evils from ordinary wrongs. The theory of evil developed in the book is applied to such practices as war rape and violence against intimates. She also addresses the moral "gray zone," in which persons can occupy the dual role of the victim and the perpetrator of evil.

Within feminist philosophy Card argues that feminism is not a single, unified worldview, but rather a lively debate characterized by the belief that women's subordination is wrong and that one should pay close attention to women's lived experiences. She stresses the importance of enabling women to describe their experiences in their own terms and cautions against the tendency to gloss over the unique experiences of nonwhite and poor women. Card's work urges one to be alert to the dangers of internalized oppression and adaptive preferences, and, in particular, to the ways in which oppression can compromise the integrity of its victims. Under oppressive circumstances victims are often motivated to ease their burdens by collaborating with their oppressors and/or uncritically adopting oppressive practices. In such cases, she contends that the oppressed cannot elude all responsibility; bad luck, for Card, does not necessarily subvert moral culpability and she advises that "[w]e need to be alert to the dangers of becoming what we despise" (1991, p. 26).

Card believes that opposition to real evils, such as, for example, domestic violence, should be given priority to opposition to gender inequalities, such as pay equity for tenured professors. Equality feminism, she says, trivializes the feminist movement and takes attention away from the graver problems that women face. On its own, Card thinks that care ethics is ill equipped to handle real evils. She also impugns care ethics for the way in which it dichotomizes the values of justice and care. Justice, she says, is not only about impartiality and universal principles of fairness, but also about giving people what they deserve, including compassionate, caring responses such as gratitude, trust, loyalty, and forgiveness.

Rejecting the idea that there is an essential lesbian identity, Card believes that there are, nonetheless, some historical commonalities among the experiences of lesbians. In *Lesbian Choices* (1995) Card attempts to articulate a lesbian ethics, understood as a ethics that grows of out the histories and experiences of lesbians and draws on paradigms and archetypes common in lesbian culture.

See also Ethics; Feminist Ethics; Feminist Philosophy.

Bibliography

Card, Claudia, and Robert R. Ammerman, eds. *Religious Commitment and Salvation: Readings in Secular and Theistic Religion*. Columbus, OH: Merrill Press, 1974.

WORKS BY CARD

Feminist Ethics. Lawrence: University Press of Kansas, 1991.

Adventures in Lesbian Philosophy. Bloomington: Indiana University Press, 1994.

Lesbian Choices. New York: Columbia University Press, 1995.

The Unnatural Lottery: Character and Moral Luck. Philadelphia: Temple University Press, 1996.

On Feminist Ethics and Politics. Lawrence: University Press of Kansas, 1999.

The Cambridge Companion to Simone de Beauvoir. New York: Cambridge University Press, 2003.

The Atrocity Paradigm: A Theory of Evil. New York: Oxford University Press, 2004.

Amanda Porter (2005)
Samantha Brennan (2005)

CARLYLE, THOMAS

(1795–1881)

Thomas Carlyle, the essayist, historian, and philosopher of culture, was born in Ecclefechan, Scotland, the eldest son of a stern, puritanical stonemason. There can be little doubt that the often-hysterical extravagances of Carlyle's later social doctrines had a direct emotional origin in the Calvinism of his childhood. In 1809 he became a divinity student at Edinburgh University, but he soon stopped attending the university courses and read widely on his own in modern literature. After leaving Edinburgh in 1814, he taught school, at the same time broadening his already impressive span of reading. In addition to imaginative literature and German philosophy, Carlyle's serious interests at this time extended to Voltaire and François Fénelon, as well as to the scientific works of Isaac Newton and Benjamin Franklin. A reading of Edward Gibbon in 1817 immediately precipitated Carlyle's rejection of the Bible as a historical record and gave impetus to his growing interest in history and social institutions.

Convinced that he could never become a minister, he returned to Edinburgh in 1819 and began his literary career as a freelance journalist. The next three years were the most miserable in a generally agonized life. He was unknown; he was socially, ideologically, even stylistically antipathetic to the fashionable literary world. He was also very poor, desperately lonely, and because of his irregular eating habits, almost permanently dyspeptic. Religious doubts quickly darkened into unbelief, and in 1822 he experienced the spiritual crisis later hieroglyphically recorded in *Sartor Resartus* (1833–1834). Like the hero of *Sartor,* Diogenes Teufelsdröckh, Carlyle found a new (if decidedly secular) faith in the moral efficacy of work: "Doubt of any sort cannot be removed except by Action," extols Teufelsdröckh. Conviction is worthless until it is converted into activity, mere speculation being "endless, formless, a vortex amid vortices." Therefore, one must "Do the Duty which lies nearest thee … Work while it is called To-day; for the Night cometh wherein no man can work." Here, in a language persuasively familiar to his readers, Carlyle expressed the chief psychotherapeutic discovery of his youth—one which was more widely disseminated in the writings of Thomas Arnold, John Ruskin, John Henry Newman, and particularly the later prophetic Carlyle himself, and was to become a leitmotif of mid-Victorian culture. Soon Carlyle found a role in which his genuine talents could emerge. His translation of Johann Wolfgang von Goethe's *Wilhelm Meister* in 1824 and his *Life of Schiller,* which was published as a book in 1825, established him as the first interpreter of German literature to the British public.

Carlyle's marriage in 1826 to Jane Baillie Welsh, an attractive, high-strung, and unusually intellectual twenty-five-year-old woman, ended his loneliness without in any way soothing the more creative ontological anxieties upon which his work depended. Carlyle's long years of isolated reading now bore fruit in a series of remarkable articles published in the Great Reviews.

LITERARY CRITICISM

Carlyle's early essays, especially "Jean Paul Friedrich Richter" (1827), "The State of German Literature," "Goethe," "Burns" (1828), "Voltaire," and "Novalis" (1829), are masterpieces of literary and ideological exegesis. However, his critical method, which was uncompromisingly didactic even for its day, was much more a criticism of life than any technical analysis of words on a page; in effect, it was essentially romantic criticism. Carlyle viewed literature as a form of self-revelation and literary criticism as a heightened confrontation of personalities engaged in the quest for moral truth. He stressed the primary need for the "transposition of the critic into the author's point of vision," which is the prerequisite of all historical and biographical as well as literary studies. Like Samuel Taylor Coleridge before him, Carlyle recognized Germany as the great contemporary source of spirituality and inwardness. For Carlyle, however, Goethe rather than Immanuel Kant was Germany's spiritual leader. More than any other writer, Goethe tri-

umphed over all doubts and denials and manifested the freedom of belief and activity. In this respect Carlyle believed that there was a significant contrast to be made between Goethe and Voltaire. In the essay "Voltaire," Carlyle argued that despite Voltaire's intellectual adroitness, his power of rapid, perspicuous arrangement of scientific and historical data, his humanity, and his universal susceptibility of mind, his real claim to greatness was that he "gave the death-stab to modern superstition." Such an achievement was, however, too negative: For Carlyle, Voltaire remained essentially a mocker, "the greatest of all Persifleurs," his chief fault being a terrible lack of earnestness.

This contrast between Voltaire and Goethe—between the pragmatic values of the eighteenth century and those of a new age of belief which was, if not actually beginning, at least imminent—ran through Carlyle's works in ever-widening applications. Moreover, it is symptomatic of the type of thinker Carlyle was that most of his later ideas were already contained embryonically in his very earliest writings (for example, in his first original publication in 1822 in the *New Edinburgh Review,* which was significantly a critique of Goethe's *Faust*). Had he stuck to literature and written more about the English classics, Carlyle would today no doubt be placed between Coleridge and Matthew Arnold as one of the major British literary critics of his age. But his interest in literature was only a steppingstone to a more vital concern with history and social diagnosis. He never really methodologically distinguished between criticism, biography, and historical and philosophical analysis. They were all used as media through which the current *crise de conscience* was to be more clearly seen and diagnosed. In this respect Carlyle may be thought of, in his early works, as an amateurish practitioner of *Geisteswissenschaften* (or "human studies"), in roughly the sense given to that term by Wilhelm Dilthey.

EARLY SOCIAL CRITICISM

"Signs of the Times" (1829), "On History" (1830), and particularly "Characteristics" (1831) were Carlyle's earliest communications in the self-assumed role of Victorian prophet. The early nineteenth century, he claimed, was a mechanical age, both externally and internally, its chief symptom being an excessive self-consciousness. With its inheritance of the largely negative contributions of the Enlightenment, it was an age of inquiry and doubt rather than of meditation and faith. Outwardly, social mechanization was more prized than individual vitality. Inwardly, morality no longer sprang from belief in a transcendental authority but arose out of prudential feeling grounded on mere calculation of consequences. The most grievous mistake of bourgeois liberalism was its doctrine that social welfare can be promoted solely through external politico-economic legislation, whereas, in truth, all human progress that is genuine ("dynamical") must emerge from the moral culture of individual men. According to Carlyle, although the present time is thus out of joint, there is nevertheless strong hope for the future. History is a cyclical but progressive (perhaps spiral) unfolding of human capabilities, and borrowing freely from Johann Gottfried Herder and the Saint-Simonians, he affirmed that the modern period is the end of a critical phase. Even as the darkest hour heralds the dawn, so the springtime of organic rebirth is now at hand.

As it happened, Carlyle was not the only British subscriber to this philosophy of history in the early 1830s. J. S. Mill's papers on "The Spirit of the Age," which appeared in the *Examiner* for 1831, propounded very similar views. These papers, which immensely impressed Carlyle, led to the formation of his somewhat precarious friendship with Mill. Doubtless the chief obstacle for Mill was Carlyle's blatantly authoritarian concept of morality and his notorious views on liberty and democracy, three notions that were soon to be dramatically embodied in Carlyle's theory of the hero.

THE HERO AND HISTORY

In the *French Revolution* (1837), Carlyle stereoscopically visualized the events between the death of Louis XV and the appointment of Napoleon Bonaparte as commander in chief of the Army of the Interior in 1795 as the accumulated result not so much of economic or social, but of moral and, in the last analysis, theological causes. The French Revolution, he sometimes seemed to suggest, was an upheaval ordained by the Creator to punish the sins of the world. Yet at the same time, and importantly for Carlyle's anthropomorphic imagination, it was an exhibition of individual personalities (of Honoré Gabriel Riqueti, Comte de Mirabeau, Georges-Jacques Danton, Maximilien-Francois-Marie-Isidore de Robespierre, etc.) in their most intense form. "History," he had written in 1830, "is the essence of innumerable Biographies." Biography, which is based on insight into human personality, is the foundation of all historical inquiry; hence, the true history of an age is the biography of its great men. Carlyle's main interest in history (as in literature) was in the moral psychology of specific individuals who seemed to him endowed with certain admirable traits of character that

he felt to be chronically lacking in the contemporary *Zeitgeist.*

The lectures he delivered in 1840, *On Heroes, Hero Worship, and the Heroic in History,* blended mythology with metaphysics to produce an image of the ideal type of individual needed as the savior of humankind. The hero can take many forms: He can be a god (Odin), a prophet (Muḥammad), poet (Dante Alighieri and William Shakespeare), priest (Martin Luther and John Knox), a man of letters (Samuel Johnson, Jean-Jacques Rousseau, Robert Burns), or a political ruler (Oliver Cromwell and Napoleon). In fact the hero can be "what you will, according to the kind of world he finds himself born into": His ever-varying persona results from the deeper needs of society. He is directed not by the "mechanical" needs of men, but by their "dynamical," unseen, mystical needs. Thus, all heroes have discerned "truly what the time wanted" and have led it "on the right road thither." In this sense, the hero is a gift from heaven, or as Carlyle otherwise puts it, a force of nature; his essential quality is "Original Insight" into the "primal reality of things." Because of the hero's firm contact with the "great Fact of Existence," he cannot lie. "He is heartily in *earnest*"; an unconscious sincerity emanates from him turning his acts or utterances into "a kind of 'revelation'" which the ordinary, unheroic man is morally obliged to recognize and obey. For "all that is *right* includes itself in this of co-operating with the real tendency of the World." Indeed, the proper feelings of ordinary men toward the heroes of their age are loyalty (which is "akin to religious Faith"), reverence, admiration, and "an obedience which knows no bounds." Hero worship, Carlyle significantly concludes, is a basic and indestructible tendency of human nature: It is "the one fixed point in modern revolutionary history, otherwise as if bottomless and shoreless."

As with Friedrich Nietzsche's *Übermensch,* there was a tendency in the twentieth century to view Carlyle's theory of the hero far too much in terms of contemporary political experience—that is, to think of the hero as a direct ancestor of fascism. But Carlyle, like Nietzsche, was essentially a philosopher of culture, not a political theorist. The hero concept is best understood as a rather curious and obsessional example of a spiritual phenomenon that reached something of a climax in the nineteenth century, most notably in the thought of Ludwig Feuerbach, Auguste Comte, Karl Marx, and Nietzsche—namely, the uneasy substitution of purely secular objects of veneration for the traditional transcendental one. Worship of God gave way to worship of man and human society.

AFTER 1840

Beginning with *Chartism* (1839), and more disastrously in *Past and Present* (1843) and the *Latterday Pamphlets* (1850), Carlyle explicitly incorporated the hero concept within the central tenets of his early social criticism to produce not only a renewed attack upon the materialistic spirit of industrial society but also an indictment of political liberty and democracy. Once more he protested against laissez-faire, the irresponsible pursuit of wealth in which "cash payment" has become the "sole nexus" between men, thus displacing the traditional ties of obligation. But social justice, he now paradoxically asserted, can be achieved only through the enforcement of social inequality. Members of the aristocracy and those heroes of the business world, the "Captains of Industry," must assume their responsibilities as rulers of the masses: Freedom consists in "the right of the ignorant man to be guided by the wiser." In this instance, as in nearly all of Carlyle's writing after about 1840, it seems that genuine social criticism was lost sight of in an increasingly pathological obsession with power: Nothing could have been further from the spirit of Mill's *On Liberty* (1859) and *Representative Government* (1861). In *Oliver Cromwell's Letters and Speeches, with Elucidations* (1845) and the *History of Frederick the Great* (1858–1865), Carlyle tried to give some historical backing to his by now hopeless moral aberrations for which he ultimately received the Prussian Order of Merit in 1874.

It is impossible to exaggerate Carlyle's impact, for better and worse, upon all aspects of Victorian culture, ranging from the development of the novel (particularly as evidenced in the work of Charles Dickens), to the formation of social policy. Nietzsche described him as a man constantly misled by a craving for a strong faith that he lacked the necessary capacity to experience. But it was hardly the capacity Carlyle lacked; rather, like Nietzsche himself, he needed something to have faith in. In the absence of his father's God, he chose what seemed to him the best substitute—the hero.

See also Arnold, Matthew; Coleridge, Samuel Taylor; Comte, Auguste; Dante Alighieri; Dilthey, Wilhelm; Fénelon, François de Salignac de la Mothe; Feuerbach, Ludwig Andreas; Franklin, Benjamin; Gibbon, Edward; Goethe, Johann Wolfgang von; Johnson, Samuel; Kant, Immanuel; Luther, Martin; Marx, Karl; Mill, John Stuart; Newman, John Henry; Newton, Isaac; Nietzsche, Friedrich; Rousseau, Jean-Jacques; Ruskin, John; Social and Political Philosophy; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY CARLYLE

The Centenary Edition, edited by H. D. Trail, 30 vols. (London, 1896–1899; New York, 1896–1901) is the most complete edition of Carlyle's *Works.* Carlyle's correspondence, most notably with Goethe, Ralph Waldo Emerson, Mill, and John Stirling, is essential reading for anyone interested in Carlyle's thought or, indeed, in nineteenth-century intellectual history in general. See also *Correspondence of Emerson and Carlyle,* edited by J. Slater (New York: Columbia University Press, 1964).

WORKS ON CARLYLE

To date there exists no full-scale analysis of the whole range of Carlyle's thought. The best short account is given by Ernst Cassirer in *The Myth of the State* (New Haven, CT: Yale University Press, 1946), Chs. 15–16. J. A. Froude, *Thomas Carlyle 1795–1835,* 2 vols. (London: Longmans Green, 1882) and *Thomas Carlyle 1834–1881,* 2 vols. (London: Longmans Green, 1884) are still indispensable. Among the many important specialized studies are B. H. Lehmann, *Carlyle's Theory of the Hero* (Durham, NC: Duke University Press, 1928); C. F. Harrold, *Carlyle and German Thought: 1819–1834* (New Haven, CT: Yale University Press, 1934); René Wellek, "Carlyle and the Philosophy of History," *Philological Quarterly* 23 (1944): 55–76; and G. Holloway, *The Victorian Sage* (London: Macmillan, 1953), Chs. 1–3.

OTHER RECOMMENDED WORKS

Rigney, Ann. "Sublimity: Thomas Carlyle and the Aesthetics of Historical Ignorance." In *Imperfect Histories: The Elusive Past and the Legacy of Romantic Historicism*. Ithaca, NY: Cornell University Press, 2001.

Sussman, Herbert L. *Fact into Figure: Typology in Carlyle, Ruskin, and the Pre-Raphaelite Brotherhood*. Columbus: Ohio State University Press, 1979.

Vida, Elizabeth Maximiliana. *Romantic Affinities: German Authors and Carlyle: A Study in the History of Ideas*. Toronto: University of Toronto Press, 1993.

Michael Moran (1967)
Bibliography updated by Desirae Matherly Martin (2005)

CARNAP, RUDOLF
(1891–1970)

Rudolf Carnap was the philosophically most articulate member of the Vienna Circle in the 1920s and 1930s, and later of the movement that came to be known in the United States as logical empiricism. During his lifetime, he was respected among analytic philosophers as the proponent of a number of ambitious language projects, especially, in his later years, a system of inductive logic. The philosophical agenda underlying these technical projects, however, remained largely implicit; only disconnected fragments of this agenda, often reduced to superficial slogans, gained some currency. Subsequent generations, quite reasonably, discarded these fragments with some contempt. The coherent and powerful view that Carnap actually held (and partly articulated), of which the ambitious technical projects were manifestations and illustrations, but not explicit statements, has only begun to be unearthed. As a result, the view of Carnap held during his lifetime and since his death is under revision.

INFLUENCES AND EARLY AMBITIONS

Carnap was born on May 18, 1891, in the German town of Wuppertal At this time the region ("Bergisches Land") was known for its pietistical, even mystical, brand of Lutheranism, and the Carnap family was strongly imbued with this local tradition. Carnap's mother's family was more intellectual, in the German tradition of *Bildung*. Carnap's grandfather, Friedrich Wilhelm Dörpfeld, was a leading educational thinker and writer who championed the ideals of Johann Friedrich Herbart. When Carnap was eight, his father died. His mother taught him and his sister at home, following her father's educational program. Teaching was restricted to one hour a day, and the children were encouraged to work out the implications of what they had learned for themselves. His mother also emphasized the arbitrary nature of moral and linguistic conventions.

Carnap's mother was evidently the strongest influence on Carnap's early mental development. In many ways this influence probably derived from the religious and educational views of her father, of whom she wrote a biography. She was herself a highly literate person, at home in the German classics, who took a particular interest in the philosophical and religious writings of Theodor Gustav Fechner, the founder of psychophysics. When Carnap began to doubt the religious doctrines he had been brought up with, he turned first to Fechner's mystical pantheism as a more explicit and detailed version of the worldview embodied in the writings of Johann Wolfgang von Goethe. The ethical and practical convictions associated with the religion of his childhood never changed. Though he abandoned it, first for pantheism, then for atheism, this change was very gradual and took a long time. And it was an entirely doctrinal change; it did not affect his values. The pragmatist streak in educational Herbartianism allowed room for the replacement of its religious props by a scientific humanism of the kind Carnap imbibed from the popular writings of Hermann von Helmholtz, Ernst Haeckel, Wilhelm Ostwald, and others.

Ostwald in particular appears to have been an important early influence. A Nobel Prize–winning chemist, he had sketched in his popular writings a consis-

tent and comprehensive worldview firmly anchored in the nineteenth-century positivist tradition of Auguste Comte and Ernst Mach. His wide-ranging interests also encompassed the German classics, the history of science, visual art, politics, and much else. He was perhaps the archetypal embodiment, during the first decade or so of the new century, of a thoroughly and uncompromisingly "scientific worldview." He was unusually cosmopolitan for a German of his generation and had traveled widely, including to the United States, where he was instrumental in establishing physical chemistry as a discipline.

Carnap advocated pacifism and internationalism, and campaigned for the use of an international language such as Esperanto, both among scientists and more widely. Many of these aspects appealed to Carnap; he even became an Esperantist in his teens, while still at school. He was especially influenced by Ostwald's conception of a "system of science" (*System der Wissenschaft*), modeled on Comte's system of unified science. In this conception, there was no fundamental distinction between human and physical sciences, of the kind that the German idealists and neo-Kantians had advocated. All knowledge was part of a single system, whose basic concepts were those of physics. This system was needed as a blueprint, Ostwald thought, for optimizing the hitherto rather aimless and chaotic output of the scientific disciplines; it would give them coherence and enable them to cooperate.

When Carnap studied at the University of Jena, from 1909, he encountered Gottlob Frege and learned modern logic from him. Though he did not immediately see the wider philosophical applications of Frege's logic, he was enthused by Frege's Leibnizian ideal of a universal language that could tie all knowledge together and display its deductive interrelations. Comte and Ostwald, like other nineteenth-century positivists, had been vague about the precise nature of the relations among the various sciences in their proposed "system of science." In Frege's logic, Carnap saw a tool for making these relations completely transparent and explicit, and making the "system of science" into something much more than a vague ideal. Logic could serve as a central discipline for bringing order to the rather chaotic and spontaneous growth of knowledge. This "system of science" could thus be a tool for coordinating and *organizing* knowledge-production on a large scale, in just the way Ostwald had envisaged.

But Jena also subjected Carnap to a quite different kind of influence, one much more at odds with anything in his background up to that time: the German Youth Movement (*Jugendbewegung*). This was a Romantic, back-to-nature rebellion of middle-class German teenagers against the materialistic and socially complacent values of their parents. There was a strong emphasis on a healthy life, especially long walks in the wilderness and avoidance of the "bourgeois" drugs (alcohol, tobacco, caffeine), as well as an idealization of peasant life and the customs of premodern times. The movement took many different local forms. In Jena, the publisher Eugen Diederichs organized the "Sera Circle," a group of university students and other young people who undertook outings with elaborate, medieval-style costumes and rituals, some improvised and some traditional. The annual midsummer celebration was a high point, when the group, with its banners, costumes, and pageantry walked in procession to a mountain some distance from town, accompanied by horse-drawn carriages. There they built a bonfire, danced, feasted, sang, and jumped over the flames two by two until sunrise. In the years just before the First World War, Carnap became very active in organizing these and similar events.

For Carnap, the lasting effect of this involvement was to give him the sense that the basic forms of human life are within human control; they do not have to be accepted from tradition or from existing conventions. This attitude of "voluntarism" would prove to be of fundamental importance to Carnap's philosophy through all its phases. And though the Youth Movement "did not leave any externally visible achievements," Carnap later wrote, "the spirit that lived in this movement, which was like a religion without dogmas, remained a precious inheritance for everyone who had the good luck to take an active part in it. What remained was more than a mere reminiscence of an enjoyable time; it was rather an indestructible living strength which forever would influence one's reactions to all practical problems of life" (Carnap 1956/7, pp. B34–B35). Moreover, it was something he missed throughout his subsequent life:

> After the war … the same spirit was still alive in the life of my newly founded family and in the relationships with friends. When I went to Vienna, however, the situation was different. I still preserved the same spirit in my personal attitude, but I missed it painfully in the social life with others. None of the members of the Vienna Circle had taken part in the Youth Movement, and I did not feel myself strong and productive enough to transform single-handedly the group of friends into a living community, sharing the style of life which I wanted. Although I was able to play a leading role in the

> philosophical work of the group, I was unable to fulfill the task of a missionary or a prophet. Thus I often felt as perhaps a man might feel who has lived in a strongly religious [and] inspired community and then suddenly finds himself isolated in the Diaspora and feels himself not strong enough to convert the heathen. The same feeling I had in a still greater measure later in America, where the power of traditional social conventions is much stronger than it was in Vienna and where also the number of those who have at least sensed some dissatisfaction with the traditional forms of life is smaller than anywhere on the European continent. (Carnap 1956/7, p. B35)

Into this idyllic dawn of a new world erupted the unheralded disaster of August 1914 and the Great War. Carnap and his Sera friends dutifully enlisted and were not even unwilling combatants, at first. Only when they witnessed the scale of the slaughter did doubts arise. Like Wittgenstein on the eastern front, Carnap participated in many of the bloodiest engagements on the western front. Both young philosophers were profoundly alienated by the culture of the officer corps. Both were wounded and were decorated for bravery. But their reactions could not have been more different. Wittgenstein withdrew into an inner life of mystical contemplation, inspired by Leo Tolstoy. Carnap, in contrast, came to appreciate that it was precisely an over-emphasis on the contemplative life, and a lack of interest in public life, that had made the German intelligentsia complicit in the bloodshed, and had allowed it to stand idly by while the political elites had started a world war. The only answer, he now decided, was active involvement in politics. Accordingly, he joined the antiwar independent socialist party, sent clandestine circular letters to friends with excerpts from the foreign press, and wrote well-informed articles about world government for underground newsletters.

The general conception behind this new commitment was a natural extension of the positivist idea of a "system of science" inherited from Comte and Ostwald, combined with the voluntarism Carnap derived from the Youth Movement. For the human race to survive and avoid disasters like the Great War, Carnap thought, it needed to take its fate into its own hand. Conflicts among nations and classes could not be left to an anarchic state of nature, but had to be subordinated to consciously chosen forms of civic cohabitation. These, of course, required highest-level conceptual planning and organization of knowledge; this too was part of the "politics" Carnap now regarded himself as involved in. For all the various social functions to work together, it was essential to arrive at a "structure of community" (*Gemeinschaftsgestalt*) that could serve to coordinate them so as "to remove [these tasks] from the realm of chaotic whim and subordinate them to goal-oriented reason" (Carnap 1918, p. 18).

Carnap's intention immediately after the war was to realize this ambition through teaching and direct political involvement. Before the war he had intended to become a physicist; now his first priority was to obtain the teaching certificate for secondary schools. The papers he wrote to qualify for the certificate show him at work, both within physics itself and in philosophical reflection about the foundations of geometry, on the construction of an Ostwaldian-Comtean "system of science" with Fregean logicist tools. In the course of these projects, he evidently came to realize that his vision of a "system of science" was anything but obvious. Though there had been much talk, among positivists (like Mach) and some systematic philosophers (like Richard Avenarius) of the reducibility of all knowledge to an empirical starting point, much work was still to be done. Like Comte in response to an earlier revolution, Carnap now realized that the reconstruction of society along the lines he had in mind, with its *Gemeinschaftsgestalt* to coordinate all productive activities within it, required the reconstruction of *knowledge* as the first and indispensable step. Though social reform could go ahead meanwhile, it could not be put on a genuinely rational basis until a "system of science" was developed, a conceptual system that was adequate to the scientific and conceptual revolutions of the past decades *and* that afforded a vantage point from which the whole of knowledge could be surveyed and organized, allowing individual claims or theories to be rationally judged. It was to the development of such a conceptual system that Carnap now single-mindedly devoted himself.

EARLY WRITINGS AND PROJECTS

This change in priorities also brought with it a change in career plans. Carnap now decided to pursue an academic career after all, but was faced with the quandary that the kind of work he planned fell between academic stools. The first project he chose for a dissertation topic was, like his 1920 paper on space and geometry, intended to work out a partial "system of science" for a subrealm of knowledge. This time it was to be an axiomatization of relativistic space-time kinematics, and the question Carnap particularly had in mind was much discussed then: Precisely what is the *empirical content* of general relativity, and precisely what parts of it were conventional? Even

before the war, Carnap had read Henri Poincaré. Now he also encountered the "radical conventionalist" Hugo Dingler, who rejected relativity on the Poincaréan grounds that all the observations involved could be accommodated without giving up Euclidean geometry, whose axioms are much simpler. Carnap disagreed; the simplicity of the system as a whole should be maximized, he said, not just the simplicity of the axioms, though he admitted that this was itself a conventional decision.

However, his proposed project was rejected by the physics department in Jena as too philosophical, and the philosophers thought it was too scientific. So instead, he reworked his 1920 paper on space and geometry, and this was accepted. The result was Carnap's doctoral dissertation and first philosophical publication, *Der Raum* (Space; 1922). Here too the central question was the status of the empirical basis (*Tatbestand*) within our conceptions of space. The answer, Carnap said, depends on whether we have mathematical, intuitive, or physical space in mind. Formal or mathematical space, Carnap said, can be constructed from logic alone, in the way Bertrand Russell had suggested in *Principles of Mathematics*, and so it has no empirical content. Intuitive space is not constructed in this logicist way, but derives from axioms based on a pure phenomenological essence-perception (*Wesensschau*) of our spatial experience. These axioms concern not the metrical properties of space, as Immanuel Kant had thought, but only its topological properties. Physical space, finally, adds the empirical basis, which, however, as Carnap argued with the aid of an extended example, underdetermines the choice of metrical geometry (it fixes the choice only up to topological assumptions).

During this period, Carnap framed the basic epistemological questions in terms of an "idealistic conception" deriving from the "positivist idealism" of Hans Vaihinger, a neo-Kantian philosopher whose book *The Philosophy of As If* had generated a great deal of discussion after its publication in 1911. Vaihinger took an extreme positivist view of what we actually know: It is only the "chaos" of our immediately present sensations that we can rely on for certain. The "reality" we construct on this basis, whether in science or in everyday life, is not genuine knowledge but a tissue of useful *fictions* that we purposefully invent to get things done in the world and to serve our mental and social needs. These fictions include not just Kant's synthetic a priori propositions (the axioms of arithmetic, geometry, and mechanics, as well as the principles of causality and of the uniformity of nature), but also, for example, the fictions of religion, of natural justice and equal citizenship, of free will and moral reasons. This was essentially a pragmatist position, as Vaihinger himself recognized, though he thought William James wrong to make utility a *standard* of truth. There is genuine truth, Vaihinger maintained, however limited in scope, while the fictions, though useful, are *not* true. They are to be judged by practical results, not by cognitive standards.

Carnap sought to pursue his dream of a system of knowledge within the framework of such an "idealistic conception." He tried various ways of *deducing* aspects of physical "reality" from the "chaos" of experience, even using a makeshift fuzzy logic at one point, but these efforts led nowhere. It seemed impossible to break out of the phenomenal "chaos" convincingly. But amidst all his other projects, the preoccupation with this overall system did not let him go. "I worked on many special problems, always looking for new approaches and improved solutions," Carnap wrote of this period "But in the background there was always the ultimate aim of the total system of all concepts. I believed that it should be possible, in principle, to give a logical reconstruction of the total system of the world as we know it" (Carnap 1956/7, p. E4).

THE *AUFBAU* PROJECT AND VIENNA

In the winter of 1921/1922 Carnap read a book that showed him how to overcome the main obstacle to his project of a "total system of all concepts," Russell's *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. This book gave Carnap the crucial hint that the way to get from the chaos of experience to a "reality" was not by *analysis* of experience, but by *construction*, using what Russell called a "principle of abstraction": "When a group of objects have that kind of similarity which we are inclined to attribute to possession of a common quality, the principle [of abstraction] shows that membership of the group will serve all the purposes of the supposed common quality, and that therefore, unless some common quality is actually known, the group or class of similar objects may be used to replace the common quality, which need not be assumed to exist" (Russell, pp. 44–45). Experiences could be gathered into equivalence classes. For example, a series of experiences of "red," at a certain position in the visual field, could be defined as equivalent. For the purposes of constructing a "real" world, this class can be regarded as an objectand used in place of the quality. No actual quality, transcending momentary experience, need figure in subsequent steps to a "reality." The evanescence of "chaotic" experi-

ence is no longer a constraint. The problem of forcing the fluid character of lived experience into the straightjacket of deductive relations disappears.

Russell's principle also solved another problem. According to Vaihinger, the "chaos" of subjective experience has no structure; nothing is "given" but the undifferentiated chaos itself. No distinguishable "elements" present themselves as naturally discrete or isolable from the chaos, available unambiguously in themselves, without calling on externally imposed fictions. A somewhat less extreme version of this holistic starting point had just been articulated by a new school of "Gestalt" psychologists. Russell's principle of abstraction—his method of substituting "logical constructions for inferred entities" (such as qualities)—solved this problem as well. Instead of trying to isolate specific elements within the undifferentiated "chaos," Carnap could obtain the elements he sought by partitioning the entire "chaos" into just two sectors, which he called the "living" and "dead" parts of experience, corresponding essentially to David Hume's "impressions" and "ideas." This one distinction allowed Carnap to arrange experiences into a temporal sequence ("ideas" belong to the past; "impressions" are present), and thus made it possible to identify holistic "temporal cross-sections" of experience, in which the total experience of a given specious present remains intact as a momentary whole.

This chronological sequence of experiential time slices gave Carnap the basic framework he needed for identifying qualities as cross-temporal equivalence classes of particular aspects within certain time slices. The holistic time-slices of experience did not need to be *analyzed*. Rather, qualities and qualitative relations could be *constructed* by defining equivalence classes of sufficiently "similar" experience aspects (e.g., approximations to "red" at certain coordinates of the visual field) across a series of time slices. ("Similarity" could be defined as precisely as needed.) The result of this procedure—with "quality classes" standing in for qualities, and so on—was therefore essentially what empiricists (like Hume, John Stuart Mill, and Mach) had always hoped to achieve by analysis, but it was accomplished *without* analysis. Carnap called it "quasi-analysis." Once qualities had been constructed, physical objects could be constructed as classes of spatial relations among qualities, and the path to a "reality" was clear.

Carnap still followed Vaihinger in distinguishing sharply between the direct, genuine, first-hand knowledge of the "chaos" and the fictive, constructed nature of "reality." But he put the boundary between them in a different place. Phenomenology, Carnap thought, offered an escape route from Vaihinger's completely undifferentiated chaos. It gave certain basic distinctions *within* the chaos (such as that between "living" and "dead" experience) a degree of objectivity. These distinctions, then, were *not* "fictional" but actually extended the range of what could be genuinely *known*, even *without* fictions, just from the "chaos" itself. So Carnap put the boundary between the "chaos" and the fictions further out than Vaihinger had done. But fictions were still needed to get from this immediately known *primary world* (of "chaos" supplied with a minimal, phenomenologically justified structure) to a fictive *secondary world* of "reality"—be it the everyday world of physical objects and forces, the abstract scientific world of fields and space-time coincidences, or some other construction.

Carnap thought at this point that he could show on phenomenological grounds that the primary world was two-dimensional, in all sense modalities. So the stepping-off point from the fixed primary world up to a freely choosable secondary world was located at the point of ascent from two to three dimensions. Within the primary world, the construction proceeded entirely by explicit definition, beginning from the qualities obtained by quasi analysis. Secondary worlds are not uniquely determined by the one given primary world, so the construction of a secondary world proceeds by optimizing its "fit" to whichever fictions are chosen to guide the construction, subject to the constraint of the primary world.

Regarding the choice among fictions to guide this ascent, Carnap remained as radically pragmatist as Vaihinger. The choice of fictions was entirely a matter of what was practically useful for some purpose. To obtain the scientific secondary world, Carnap suggested, we need adopt only two fictions, corresponding roughly to Kant's categories of cause and substance: (1) a principle of induction or uniformity of nature and (2) a principle of "continuity" (as Mach had called it), the principle that a certain cluster of perceptions grouped into a physical object, say, remains constant while we are not perceiving it if it remains sufficiently similar (by defined standards) before and after the interruption.

It seemed then that the problems facing Carnap's dream of a "total system of all concepts" had been overcome. He could now go public with his grand plan to revolutionize the conceptual framework of knowledge. He immediately wrote up a sketch of the new "total system of all concepts" that he gave the Vaihinger-inspired title *Vom Chaos zur Wirklichkeit* (From the chaos to reality). He organized a conference for the following year (1923) to

discuss it—the first conference of "scientific philosophy." The participants, who previously had each been working alone, became a like-minded community. Carnap also talked to Hans Reichenbach and others about starting a new journal to propagate the new ethos. The program of "conceptual politics" was well under way.

Carnap continued to work on his "total system of concepts" and in 1928 published *Der logische Aufbau der Welt* (*The Logical Construction of the World*), which became the programmatic bible of the Vienna Circle (Carnap had joined it in 1926, when he became a junior lecturer at the University of Vienna). The *Aufbau* exemplified the Vienna Circle's goal of "rational reconstruction," the replacement of vague, informal concepts by precise ones defined within a standard logical language in which all of knowledge could be expressed. The concept rationally reconstructed in the *Aufbau* was that of "empirical content" (or "empirical meaning"), which had long been of central importance for empiricists but had never been made logically precise.

Though the germ of the *Aufbau* is already contained in "From the Chaos to Reality," there were also some important changes. In the 1922 system, three components had worked somewhat uneasily together: (1) the *basis* of momentary time-slices of total experience, distilled from a chaotic primary world by phenomenological reflection; (2) the *fictions* that guided the construction of a secondary world from the primary world; and (3) the *logic* that connected the constructional steps. As Carnap worked on the system after 1922, these three parts came to seem less compatible with each other. Though he had greatly reduced the number of fictions from Vaihinger's heterogeneous jumble, the two he had chosen still seemed somewhat ad hoc. And phenomenological reflection, though also a kind of "thought," did not operate mechanically, without mental assistance, as the logical system of Frege and Russell did. Logic and phenomenology seemed to be fundamentally different kinds of constructional procedure that could not be reduced to each other. If Carnap was to take seriously Russell's dictum that "logic gives the method of research," then everything that *could* be done by logic alone *had* to be done by logic alone. Accordingly, by 1925 Carnap gave up the distinction between "primary" and "secondary" worlds (between a single determinate "given" reality and optional constructed "realities"). Instead, he extended the logical construction *downwards* as far as possible to perform the tasks that had previously been left to phenomenology.

This displacement of phenomenology by logic led Carnap to minimize the number of relations required for the construction. By 1925 the number of basic relations had been reduced to five, and in the published *Aufbau* there is only a *single* basic relation—that of "remembered similarity" of qualitative aspects across temporal slices of experience. Indeed, the imperative to eliminate the subjective element altogether and make the construction *entirely* logical led Carnap to the extreme of suggesting that even this one remaining basic relation might be eliminated if we define it "implicitly," that is, define it simply as "whatever basic relation leads to our existing body of scientific knowledge" (1928/2003, sec. 153).

Carnap did not, however, give up Vaihinger's pragmatist orientation. To make the fictions of cause and substance that guided the construction less ad hoc, Carnap suggested that they could be deduced from some "highest principle of constitution," which might in turn be deducible from "whatever it is that knowledge contributes to the more comprehensive context of life purposes" (1928/2003, sec. 105). And he emphasized that the *Aufbau* construction was not the only possible one, but that quite different approaches might be appropriate for different purposes.

The *Aufbau* construction gave the Vienna Circle a standard by which to judge any statement and determine whether it has meaning. Carnap gave a popular lecture around this time in which he depicted human intellectual history since the Greeks as a struggle between "critical intellect" and "poetic imagination." In the ancient world, he said, critical intellect had dealt poetic imagination a major blow with its concept of a single, all-encompassing physical space. In response to any mythical creature or entity the imagination might dream up, critical intellect could now ask, "Where is it located in space?" or, "Tell me exactly how I can get there from here." Imagination took to hiding its goblins and spirits in remote, inaccessible places, but this was only a stopgap. Eventually, imagination struck back more forcefully by inventing *metaphysics*. It hit on the idea of a *nonmaterial* God and other nonmaterial entities. This was plausible, Carnap explained, because we often refer, quite legitimately, to nonmaterial items like numbers, relations, and so on. Many thinking people were taken in. But now, he said, critical intellect has found a tool to combat this maneuver. Just as the ancients had hit upon the idea of an all-encompassing *physical* space, so now we, here in Vienna, have developed a single, all-encompassing *conceptual* space: the *Aufbau* system. This system puts the burden on the poetic imagination to specify exactly how to get to any supposed non-material entity from "here"—from my own immediate experience. This was how the *Aufbau* sys-

tem provided the basis for the Vienna Circle's campaign against metaphysics and traditional obscurantism, and exemplified the circle's project of "rational reconstruction"—the piecemeal replacement of traditional, vague concepts by more precise and useful ones.

WITTGENSTEIN

When Carnap went to Vienna in 1926, the *Aufbau* was substantially complete. He assumed that its construction of physical objects and theoretical entities would all be of a piece, so that concrete and theoretical objects could also be cashed out again in terms of subjective experience. In 1926 he published the booklet *Physikalische Begriffsbildung* (Physicalconcept formation), in which he argued for the completely seamless intertranslatability of subjective experiences and the sets of 14-tuples of numbers in which, he said, the world could, against a set of background theories, be exhaustively described.

But on arriving in Vienna, Carnap was confronted with a new influence that disrupted this harmony. The Vienna Circle was just in the process of reading Wittgenstein's *Tractatus Logico-Philosophicus* line by line, and Carnap came to share their appreciation of it. The *Tractatus* solved what historically had been the severest problem for empiricism: its inability to account for mathematics. Frege's critique of empiricist efforts (by Mill, for instance) to found arithmetic on empirical generalizations had convinced members of the circle that a different approach was needed. But they also rejected Frege's and Russell's view that logic and mathematics were essentially like laws of nature, only of much greater generality, governing *everything*. Wittgenstein argued, rather, that logic and mathematics are about *nothing*; they are empty. They convey no information about the world, as they are "tautological" artifacts of the language itself and neither make nor exclude any assertions about anything that is or is not the case.

What gives a sentence meaning, Wittgenstein said, is that it is a logical "picture" of a fact. So all meaningful sentences have to be built up out of "atomic" sentences, picturing simplest facts, by truth-functional connectives. Since the number of observation sentences supporting a physical law can only ever be finite, this meant, to the Vienna Circle, that a *universal* law cannot, strictly speaking, have meaning. So in Wittgenstein's framework, a law could be nothing more than the body of evidence for it. This made theoretical science as it had been done for the past few centuries impossible, and it broke the seamless continuity Carnap had previously assumed between subjective experience and theoretical concepts. This was bad enough, but Wittgenstein's conception of meaning raised another problem for the circle. The very sentences expressing that conception fell victim to their own consequences. Wittgenstein confirmed this in the final sentences of the *Tractatus*, where he declared his own book meaningless. So although the Vienna Circle regarded the *Tractatus* as indispensable, they also realized that to do the job they relied on it to do, its conception of language would somehow have to be expanded to admit physical laws and metalinguistic "elucidations."

Carnap's first task, in this project, was an attempt to fit *axiomatic* concepts within Wittgenstein's constraints. During his first few years in Vienna, this was his main focus; he worked until 1930 on a large manuscripthe called *Untersuchungen zur allgemeinen Axiomatik* (Investigations in general axiomatics). Its main point was to show that David Hilbert's use of a "metamathematics" to prove the consistency of merely formal axiom systems, of which most mathematics consists, was ultimately not essential, but that only a single basic language would suffice. In the *Axiomatics*, Carnap takes a "foundation system" of logic, arithmetic, and set theory as the starting point, and stipulates that all axiom systems must be expressed in it; they derive their meaning from being anchored in this absolute system. Where does this "foundation system" itself come from? Carnaop gave a preliminary answer in a sketch entitled "Neue Grundlegung der Logik" (New foundation of logic), where he tried to expand the repertoire of what can be regarded as meaningful (and tautological) within Wittgenstein's picture theory by experimenting with arbitrarily long truth tables.

All this effort came to naught in early 1930 when Alfred Tarski visited the Vienna Circle. In private conversations, he convinced Carnap that the single-language approach of the *Axiomatics* did not really capture the metamathematical concepts that Carnap had wanted to account for in a single language. Later that year a young student of Carnap's, Kurt Gödel, showed that arithmetic was incomplete—that it contained sentences that, although true, could not be proved from its axioms. This contradicted one of the central theorems Carnap had arrived at in his *Axiomatics*.

By the end of 1930, then, the program of rational reconstruction had run aground. The efforts to expand Wittgenstein's restrictive conception of language to allow universal laws and axiomatic mathematics had come to nothing. And much of the damage had been done by mathematicians like Alfred Tarski and Kurt Gödel, who were using metalanguages in very precise ways, appar-

ently flouting Wittgenstein's claim that it was impossible to speak about language *in* language. Could the Vienna Circle's program somehow be rescued?

SYNTAX

On January 21, 1931, Carnap came down with a bad flu. He hardly slept that night. As he lay awake an idea came to him, in a flash, that solved all his problems. The Wittgensteinian conception of meaning went overboard. We can forget about meaning, he realized, at least in our statements *about* the scientific language—our metalinguistic "elucidations." Though the scientific language itself had empirical meaning (in a way that remained to be clarified—this became the subject of the "protocol sentence debate"), in our elucidations of it we are not talking about anything extralinguistic; we are talking *always and only about language*. In these metalinguistic elucidations, we must be careful not to talk about "facts" or about "things," but always confine ourselves to talking rather about "sentences" or "thing names." As Carnap would soon put it, we should in principle restrict ourselves to the "formal mode of speech" (sentences and names) and indulge in the "material mode of speech" (facts and things) only if we are sure we can translate our statements into the formal mode. Carnap adopted the metalinguistic viewpoint of Hilbert, Tarski, and Gödel, and applied this hitherto purely mathematical method to the whole of knowledge. Philosophy itself was to be reconstructed in the formal mode of speech. What remained of philosophy was the metalinguistic "logic of science" (*Wissenschaftslogik*) that could be expressed in the formal mode.

Carnap immediately threw himself into creating the language for the formal mode of speech. Taking his cue from Hilbert's metamathematics, Carnap sought to strip this standard metalanguage of all problematic assumptions. It would consist simply of strings of dots on a page, and the basic laws of arithmetic would arise unambiguously in the metalanguage from the immediately evident patterns of dots (the commutative law, for instance, is immediately evident from the perceptible equivalence of the number of dots counted from the left and from the right). A few months later, when he was preparing to present his new ideas to the Vienna Circle in June of 1931, Carnap found that he could not express certain essential concepts in this limited language, and turned instead to a more usual axiomatized arithmetic. This also had the advantage that, by using Gödel's trick of arithmetizing syntax, Carnap could now express the syntax of the language (its logic) in the language itself. So the syntactic metalanguage collapses into its object language, and there is after all only one language again.

Though some details still needed working out, Carnap was convinced he had what he needed: a canonical language for the formal mode of speech. This gave him a new and different way of eliminating metaphysics, superceding the previous, Wittgensteinian way. The previous criterion had been a criterion for *meaning*. The new criterion was not. It required that any statement either be straightforwardly factual or be translatable into the formal mode of speech. In other words, an acceptable sentence had to be statable in a "correct" language—the canonical language or an equivalent. Assuming that the kinks in his canonical language could be ironed out, Carnap thought it would be capable of expressing the entire language of physics, as well as its own syntax in a sublanguage. Since the Vienna Circle's "unity-of-science" program held that all knowledge was expressible in the language of physics, Carnap put his canonical language forward as a *universal* language (though not as *the* universal language) for all knowledge. So another way of putting the new criterion was this: An acceptable statement must be expressible in the language of physics. The new ideas of January 1931 flowed into the stream of Carnap's discussions in the circle, particularly with Otto Neurath, to produce this new doctrine of *physicalism*.

But the demands on the "correct" language were exorbitant. Though Carnap had wanted to keep it weak and uncontroversial, it also had to be capable of expressing all the mathematics needed for physics. On the other hand, its arithmetized syntax had to be capable of expressing the basic concept of "analytic truth," or there would be no way of saying whether a formal-mode statement "holds." Gödel had shown that provability was not a sufficient criterion for mathematical or logical truth; there are true sentences that are not provable. So a different criterion was needed, one that would identify the logically true sentences solely by means of the formation and transformation rules of the language. Carnap did attempt such a criterion for "analyticity" in the first draft of his syntax book *Logische Syntax der Sprache* (*Logical Syntax of Language*), written between late 1931 and the spring of 1932. He sent the typescript to Gödel, who pointed out that the new criterion was defective, and that it is *impossible* to give a correct definition of analyticity or logical truth in *any* metalanguage that can be faithfully represented in the object language (e.g., by arithmetization). (This is the indefinability of truth we now associate with Tarski.) So it turns out that Carnap's single-language approach will not work after all.

But although Carnap, with Gödel's assistance, would later develop a new definition of analyticity, in a metalanguage, this definition no longer enjoyed the privileged status that one in the *same language* (had it been possible) could have claimed. And indeed, there is no basis for singling out any particular metalanguage as more "suitable" or "natural" than any other. One option may turn out to be more *useful* than another, but there is no basis for privileging one of the many possible candidates as "correct." So the new definition of analyticity hardly seemed to matter any more. Carnap was more impressed with the *language relativity* of any definition of truth or analyticity. The disputes about protocol sentences within the circle merged in his mind with the disputes among intuitionists, logicists, and formalists in the philosophy of mathematics. All these disputes, it suddenly seemed to him in October 1932, really just revolved around the question how to set up the *language*, and there was no right or wrong answer to such questions. He no longer saw any basis for choosing one solution as "correct." One could only try out different ways and see which ones worked better. This new attitude, which completed Carnap's "linguistic turn" and first appeared in his reply to Neurath about protocol sentences in late 1932, received its definitive statement in the "principle of tolerance," enunciated in *Logische Syntax der Sprache* (The Logical Syntax of Language) in 1934.

SEMANTICS, LATER PROJECTS, AND THE IDEAL OF EXPLICATION

Carnap's syntax period was characterized by two successive major ideas. The first, from January 1931, had been the rejection of Wittgenstein's picture theory of meaning and its replacement by (a) a sharp distinction between a language (a calculus or purely formal symbol system) and its interpretation, and (b) the requirement that a language be entirely specified by explicit rules. The second major idea, from October 1932, had been the principle of tolerance: No language is inherently definitive or "correct"; there is no logical "reality" for a language to "correspond to." In the published *Logical Syntax of Language*, these two ideas were enmeshed with a third idea: the restriction to the "formal mode of speech" and the avoidance of meaning. But within a year of the book's publication, that third idea was dropped; Carnap accepted Tarski's new semantical accounts of designation and truth. The first two major "syntax" ideas (those of January 1931 and October 1932), however, survived unscathed, though, for the rest of Carnap's career (so it is actually a bit misleading to call them "syntactic"). What did *not* survive was the overreaction against "meaning" that accompanied the original insight. In distinguishing between a language and its interpretation, Carnap's first response had been to restrict extra-linguistic interpretation to the object language (and there to physicalistic interpretation), and dispense with it entirely in the "elucidatory" metalanguage. But this restriction was loosened when he saw that interpretation could be completely specified by explicit rules (governing satisfaction, designation, and truth).

The remaining thirty-five years of Carnap's career were largely occupied with technical work on a number of of not very successful language projects, of which the best known were these: (1) He tried, in a series of semantic works, to develop a general definition of "analyticity" that would distinguish analytic from synthetic sentences in a natural and obvious way. The shortcomings of these successive attempts were pointed out by W. V. O. Quine, and were often taken to undermine other parts of Carnap's view, for example, the principle of tolerance itself. (2) Carnap also tried unsuccessfully to specify a strict logical relation between observation sentences and theoretical sentences. After he abandoned the *Aufbau* effort to construct theories directly from subjective experience, he offered a series of progressively looser definitions of "empirical content" or "empirical reducibility." These attempts were also subjected to searching criticism, above all by C.G. Hempel. The lesson derived from this failure has generally been to abandon the question altogether, instead of confining the pessimism to Carnap's particular approach. (3) The last three decades of Carnap's life were largely devoted to the creation of an inductive logic. This was intended as a tool for practicing scientists, to give them a way of measuring the objective probability of a theory with respect to the available evidence. It was intended to make precise the informal usage, in everyday *and* scientific life, by which the evidence is taken to "make" one hypothesis "more likely" than another. Carnap's proposals attained some currency in the 1950s and 1960s and were considered by R. B. Braithwaite, for instance, to be the most promising route to a fundamental justification of John Maynard Keynes's theory of probability. But with a few exceptions, Carnap's work on probability has not been in the mainstream of discussion since the 1980s.

Even if these language projects are written off as failures, though, this would not discredit the larger vision or ideal of explication and language engineering that guided Carnap after 1935. He devoted little time to making this ideal explicit, so it must be gleaned indirectly from his

approach to the various language projects and from occasional statements, like the famous paper "Empiricism, Semantics, and Ontology" (1950), his replies to critics in the Schilpp volume, as well as unpublished papers and notes.

The basis of this ideal is the utopian conception of highest-level "conceptual politics" that never left him after 1918. He believed that those who are fortunate enough to be able to devote their lives to thought and reflection have a responsibility to devise conceptual frameworks for human knowledge (as a whole) that will maximize the usefulness of that knowledge for the human species—not for some particular use, but for the full spectrum of uses to which humans put knowledge, especially for the purpose of enlightenment, or liberation from unreflective tradition and conformity. In devising such frameworks we are constrained by certain obvious human limitations, but we should not allow ourselves to be overly constrained by the past—the languages handed down to us by our ancestors. Those give us a starting point, certainly, but we should not treat the puzzles and contradictions embedded in natural languages, or in historical languages of philosophy, with any undue reverence. In fact, we should liberate ourselves from them as far as possible when planning new and better frameworks of thought. Certainly our habitual ways of thinking and talking are deeply entrenched, and are hard for us to abandon, but in Carnap's view this is no reason to be constrained by them when we envision new ones.

In Carnap's mature conception, there are three levels of language engineering and language study: *Syntax* considers languages in isolation from anything extralinguistic that they might be thought of as indicating; *semantics* considers languages in relation to an extralinguistic world, but still in isolation from the actual uses of those languages by (human or machine) users; and *pragmatics* considers languages in relation to their use contexts and their users. Each of these three (syntax, semantics, pragmatics) can be considered as *engineering* activities (the creation or discussion of new or improved languages) or as *empirical* studies (of existing languages). The engineering activity Carnap called "pure" syntax, semantics, or pragmatics, while the empirical study he called "descriptive" syntax, semantics, or pragmatics. Linguists generally engage in the *descriptive* syntax, semantics, and pragmatics of already existing natural languages, while logicians generally engage in the *pure* syntax and semantics of constructed languages. Among the traditional sectors of philosophy, epistemology and methodology belong to pragmatics, while whatever remains of metaphysics and ontology belongs to semantics—though this now becomes a matter of *deciding* which entities to make fundamental to a language framework, given existing scientific knowledge, rather than *finding out* what those entities are or might be.

This voluntarist orientation remained fundamental for Carnap. The notion that something beyond the scope of science might actually *be the case* seemed to him a back door to the re-admission of traditional prejudices and conformities of all kinds. Certainly we need to make assumptions, he acknowledged, but we can *decide* on these and spell them out; they are not "out there" for us to *find*. On these grounds he deprecated Quine's preoccupation with ontology. It makes no sense to talk about "what there is," Carnap said, without specifying the language framework in which this is asserted; any such claim is intelligible only relatively to a language framework. It makes perfectly good sense to ask, *within* a framework that includes, say, the Zermelo-Frankel axioms for set theory, whether there are infinite numbers. Such "internal" questions have determinate answers. But it makes no sense, *outside* such a framework, to ask "just in general" whether "there are" infinite numbers. Not only is there no determinate answer, but there is no way to give such an "external" question itself any clear meaning. What we *can* ask instead is the *practical* question whether it is better (e.g., for use in science) to choose a linguistic framework that has infinite numbers or one that does not. But this is not a question of ontology or semantics; this is a question of pragmatics, a question of *which language* we *want*.

The process by which the human species upgrades its messy and imprecise inherited languages to newly built and more precise ones Carnap called *explication*. He acknowledged that this is a piecemeal process, not a revolutionary one. Humanity replaces its concepts a few at a time. Even the people working at the frontier of knowledge have to use a vernacular, a derivative of ordinary language, to discuss the application of the more precise calculi in which they frame their theories. Their vernacular will, of course, be cleaner and more precise than the vernacular of the society at large. In the scientific vernacular, all concepts used are intended in their scientifically rigorous meanings.

But many concepts even in this tidied-up vernacular have no such precise meanings. They may go on being used for generations before they are made precise. The mathematical concept of the derivative of a function, for instance, was put to good use for nearly two centuries before it was given a precise meaning by the work of Cauchy and Weierstrass. Another example Carnap often

cited was the replacement of our vague, subjective, intuitive sense of "hot" and "cold" by the precise, quantitative concept of temperature, which we can define intersubjectively by reference to measurement devices. This concept not only takes the place of the former vague concepts for many purposes; it also gave us many capabilities the vague concepts lacked. For instance, it can provide an outside, objective framework or standard against which to judge subjective feelings; instead of just saying "I feel hot" or "I feel feverish," I can take my temperature and find out exactly how much higher it is than its ordinary level. So explication also provides a framework of objectivity that enables us to escape from a merely subjective view of the world. But the replacement of the vague, informal worldview by a framework of more objective concepts is iterative and never complete; temperature is not an ultimate constituent of our theory of nature.

Explication, which in Carnap's view is the main task of conceptual engineering, consists in the *replacement* of a vague concept in need of explication—the *explicandum*—by a more precise one, the *explicatum*. The first step is the *clarification* of the explicandum, the establishment of some basic agreement among those using the vague concept what they mean by it. The next step is a proposal for its replacement, a proposed *explicatum*. This should have the most important uses agreed on in the clarification stage, but need not have all of them. It should, if possible, be expressed in a language framework that makes clear its relation to a wide range of other concepts. Above all, it should be more precise and more useful than the explicandum. The (provisional) acceptance of an explicatum is just its use by the specific community to which it has been proposed and, ultimately, its wider use by the community of those who use the tidied-up scientific vernacular.

Explication differs in one critical respect from the previous Vienna Circle program of "rational reconstruction." Rational reconstruction was a one-way street; vernacular concepts were to be replaced, piece by piece, with more precise ones. It was assumed that there was a single, definitive logical language in which this reconstruction could be done. But under the new regime of tolerance, there is no longer a single correct language. There is an infinity of possible languages for the community to choose from. Explication is therefore *dialectical*, as Howard Stein, a student of Carnap's, has pointed out, in a way that rational reconstruction was not. Knowledge has obvious and far-reaching effects on our practical life (more and more so, it seems, as history advances). It can tell us, among other things, about the likely consequences of various value systems and courses of action, far more than we could have known a few centuries ago. On the other hand, the way we represent our knowledge to ourselves is language-relative. We can only know what we know in a particular language, and the form in which it presents itself to us is relative to that language. The choice *among* languages, though, is not a choice we make *within* a given language framework. It is a practical choice, involving values (as is the choice among explications for a given explicandum, at the local, piecemeal level.). These are *external* questions, in Carnap's terms. So knowledge and values are in a constant feedback relation to each other, in this dialectical ideal of explication; knowledge shapes values and values shape knowledge.

See also Analysis, Philosophical; Logical Positivism; Positivism; Quine, Willard Van Orman.

Bibliography

WORKS BY CARNAP

P. A. Schilpp, ed., *The Philosophy of Rudolf Carnap*, Vol. XI in the Library of Living Philosophers (La Salle, IL: Open Court, 1963), is indispensable. This 1,100-page volume contains Carnap's illuminating intellectual autobiography, 26 critical essays on various aspects of his philosophy, his replies to the critics, and an exhaustive bibliography through 1963. The manuscripts of the original, longer version of the autobiography is held by the Carnap collection of the Young Research Library, UCLA and cited above as "Carnap 1956/7. Another manuscript item cited above (as "Carnap 1918") is a newspaper editorial Carnap planned to publish in October 1918; it is held by the Carnap Collection of the Archive of Scientific Philosophy, University of Pittsburgh.

Carnap's major works include *Der Raum: Ein Beitrag zur Wissenschaftslehre*. Berlin: Reuther & Reichard, 1922; *Physikalische Begriffsbildung*. Karlsruhe, Germany: Braun, 1926; *Der Logische Aufbau der Welt* (1928); English translation by R. George as *The Logical Structure of the World*, 2nd ed. LaSalle, IL: Open Court, 2003; "Überwindung der Metaphysik durch die logische Analyse der Sprache." *Erkenntnis* 2 (1931–1932): 219–241, translated by Arthur Pap as "The Elimination of Metaphysics through Logical Analysis of Language." In A. J. Ayer, ed., *Logical Positivism*. Glencoe, IL: Free Press, 1959: 60–81; "Die physikalische Sprache als Universalsprache der Wissenschaft," in *Erkenntnis* 2 (1931–1932): 432–465, translated by Max Black as *The Unity of Science* (London: Routledge, 1934) *Logische Syntax der Sprache*. Vienna: Springer, 1934, translated by Amethe Smeaton as *Logical Syntax of Language*, 2nd edition. LaSalle, IL: Open Court 2002; "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–471 and 4 (1937): 1–40; *Foundations of Logic and Mathematics*, Vol. I, no. 3, of *International Encyclopedia of Unified Science*. Chicago: University of Chicago Press, 1939; *Introduction to Semantics*. Cambridge, MA: Harvard University Press, 1942; *Meaning and Necessity: A Study in Semantics and Modal Logic*, 2nd ed. Chicago: University of

Chicago Press, 1956; "Empiricism, Semantics, and Ontology." *Revue internationale de philosophie* 4 (1950), 20–40, reprinted in *Meaning and Necessity*, 2d ed., op. cit., 205–221; *Logical Foundations of Probability*, 2nd edition (Chicago: University of Chicago Press,1962); and "The Methodological Character of Theoretical Concepts." In Herbert Feigl et al., eds., *Minnesota Studies in the Philosophy of Science*, Vol. I, Minneapolis: University of Minnesota Press, 1956: 38–76; "The Aim of Inductive Logic." In Ernest Nagel, Patrick Suppes, and Alfred Tarski, eds., *Logic, Methodology, and Philosophy of Science*, Stanford, CA: Stanford University Press, 1962: 303–318.

WORKS ON CARNAP

The best overview of Carnap's work is the collection of essays *Carnap Brought Home: The View from Jena*, edited by Steve Awodey and Carsten Klein (LaSalle, IL: Open Court, 2004). Also essential is Michael Friedman, *Reconsidering Logical Positivism* (Cambridge, U.K.: Cambridge University Press, 1999), which focuses on the early works. A. W. Carus, *Carnap in Twentieth-Century Thought: Explication as Enlightenment* (Cambridge, U.K.: Cambridge University Press, 2006) cites a range of unpublished sources.

The *Aufbau* was subject to close technical analysis by Nelson Goodman, *The Structure of Appearance*, 3rd ed. Boston, MA: Kluwer, 1977, and from a very different viewpoint, focused on its neo-Kantian origins, by Alan Richardson, *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*. Cambridge, U.K.: Cambridge University Press, 1998.

The *Syntax* has attracted a great deal of attention since 1990. Among the most important contributions are Thomas Ricketts, "Carnap's Principle of Tolerance, Empiricism, and Conventionalism." In *Reading Putnam*, edited by P. Clark and B. Hale. Oxford: Blackwell, 1994: 176–200; Warren Goldfarb and Thomas Ricketts, "Carnap and the Philosophy of Mathematics." In *Science and Subjectivity: The Vienna Circle and Twentieth-Century Philosophy*. Berlin: Akademie-Verlag, 1992: 61–78.

The debate between Carnap and Quine has also generated much secondary literature. The classic statements are in the Schilpp volume, as well as Quine's famous "Two Dogmas of Empiricism." Richard Creath has edited their correspondence, together with some very interesting documents, in *Dear Carnap—Dear Van: The Quine-Carnap Correspondence and Related Work*. Berkeley: University of California Press, 1990, with an introduction that gives a lucid overview of the controversy. Important contributions to the debate since then include Howard Stein, "Was Carnap Entirely Wrong, After All?" *Synthese* 93 (1992): 275–295; G.H. Bird, "Carnap and Quine: Internal and External Questions," *Erkenntnis* 42 (1995): 41–64; and Richard Creath "The Linguistic Doctrine and Conventionality: The Main Argument in 'Carnap and Logical Truth,'" In *Logical Empiricism in North America*, edited by Gary L. Hardcastle and Alan Richardson. Minneapolis: University of Minnesota Press, 2003: 234–256.

Among other topics in the vast literature, the posthumously published critique of Carnap by Kurt Gödel ("Is Mathematics Syntax of Language?" in Kurt Gödel, *Collected Works, vol. III: Unpublished Essays and Lectures*, Oxford: Oxford University Press, 1995: 324–363, with an introductory note by Warren Goldfarb) has aroused much interest; see Goldfarb and Ricketts (above); also Steve Awodey and A.W. Carus, "How Carnap Could Have Replied to Gödel." In Awodey and Klein (above). Finally, Richard Jeffrey addresses a central feature of Carnap's philosophy, in the context of the inductive logic, in his paper on "Carnap's Voluntarism," in *Logic, Methodology, and Philosophy of Science IX*, edited by D. Prawitz, B. Skyrms, and D. Westerståhl. Amsterdam: Elsevier, 1994: 847–866.

A.W. Carus (2005)

CARNEADES

(214–129/8 BCE)

Carneades became scholarch of the Academy (Plato's school) sometime before 155 BCE, when he was sent to Rome along with the leaders of the Stoa and the Peripatos (Aristotle's school) to represent the interests of Athens before the senate. It was during the embassy to Rome that the most notorious episode in his life took place. According to tradition, Carneades delivered public lectures on succeeding days, defending justice on the first and arguing that it is a form of folly on the second day.

He was renowned in antiquity above all for the argumentative virtuosity that he displayed in the skeptical examination of views of other philosophers. For this he was indebted to the example of Arcesilaus, who had inaugurated the skeptical turn in the Academy in the third century BCE, which saw the examination of other schools' theories, especially the Stoa's, replace the elaboration of its own positive doctrines as the school's principal occupation. By common consent, Carneades brought this practice to its highest level. Until the dissolution of the school, which probably occurred under the scholarch Philo of Larissa, who left Athens for Rome in 88 BCE, philosophy in the Academy and among the philosophers in its orbit was dominated by Carneades and his legacy. He also stimulated Stoics such as Antipater of Tarsus to modify and refine their positions.

CARNEADES AND THE ACADEMY

Like Arcesilaus and Socrates before him, Carneades wrote nothing, but exerted an influence on his students and contemporaries through his teaching and in-person practice of philosophical debate. What is known of him depends ultimately on works written by those who were in a position to observe him, especially Clitomachus, his student and, after an interval, successor as head of the Academy. None of these works have survived, but they

were mined extensively by authors such as Cicero and Sextus Empiricus, whose books are available.

Carneades was credited in antiquity with founding the third or New Academy, which succeeded the second or Middle Academy of Arcesilaus and the old Academy of Plato and his immediate followers. Two new characteristics appear to set Carneades apart from his middle Academic predecessors. Ancient philosophers and modern historians of philosophy have credited him with a less skeptical attitude toward the possibility of well-founded beliefs, if not of certain knowledge. And the evidence shows that he tackled and sometimes defended views about a wider range of issues—not only epistemology, but logic, ethics, natural philosophy, and theology as well. If the first of these is correct, the second comes as no surprise. A moderation of the Academy's skepticism would have opened the way for the suitably circumspect adoption of views in ethics, natural philosophy, and other areas.

Caution is in order, however. The Academics' arguments were in the first instance dialectical. They aimed to deduce conclusions unwelcome to an opponent from assumptions to which that opponent was committed, either because they were already explicitly incorporated in the opponent's theories or because they were for some other reason difficult for the opponent to reject. Without committing their authors to a position themselves, such arguments expose difficulties within the opponent's position and show that the opponent's claims to knowledge were not secure.

Carneades's practice of defending positive views, which at first appears to be a departure from the Academic tradition of dialectical argument, may instead be viewed as a continuation of it by other means. Arguments between the Academy and other schools often reached an impasse. The powerful case brought by the Academics against Stoic epistemology, for instance, elicited a formidable response. If the burden of proof belonged to the Academy, it had not proved its case; the Stoics were not obliged to concede all the premises of the Academy's arguments on pain of self-contradiction. On the other hand, by rejecting those premises, the Stoics often committed themselves to theses that were highly disputable and implausible. And they were not content merely to exhibit the consistency of their theories; they claimed that these theories were true, and that open-minded and intelligent auditors could be persuaded of this.

To this end, the Stoics now argued that the consequences of rejecting their position were unacceptable and that no alternative could do justice to the relevant considerations. If an argument of this kind were successful, the Stoics' opponents would be compelled to reevaluate their doubts. At a minimum, Carneades's positive proposals served to counter arguments of this kind by showing that there remained alternatives that his opponents were not in a position to exclude. Thus, although they were his in the sense of being his creations, Carneades's proposals need not have been his in the sense of expressing his convictions. Some of his theories seem to have been meant only to serve polemical purposes, others were considerably more substantial and deserve to be taken seriously in their own right. It is obvious that some of Carneades's successors did adopt positions of his; It is obvious that some of Carneades's successors adopted some of his theories as their own positions; it is less clear whether Carneades committed himself to these or any other theories.

CARNEADES'S SKEPTICAL ARGUMENTS

Like his Academic predecessors, Carneades argued for the two epistemological propositions for which ancient skepticism is most famous: that nothing can be known and that one ought therefore to suspend judgment about all matters. Strictly speaking, they argued that there are no cognitive impressions. The cognitive impression (*kataleptikē phantasia*), the Stoics' criterion, is a perceptual impression that arises in conditions that both ensure that it is true and impart to it a clarity and distinctness that belong only to impressions produced in this way. By confining one's assent to cognitive impressions, one can avoid accepting any false perceptual impressions. Because this is a necessary condition for knowledge according to the Stoa, if there are no cognitive impressions, it follows for anyone who accepts Stoic epistemological views that nothing can be known. The Academy made its case by arguing that the special character of clarity and distinctness allegedly peculiar to the cognitive impression was not, in fact, confined to impressions that had arisen in the ideal conditions specified by the Stoa, but could in fact also belong to false impressions, which were therefore indistinguishable from impressions with the required truth-guaranteeing origin.

Carneades probably added to the stock of skeptical arguments that he had inherited, but the contribution to the debate for which he is best known came in response to the Stoics' counterarguments. In answer to their contention that, without cognitive impressions, human beings would be deprived of a basis for rational action as well as the possibility of wisdom, he developed a theory of probable impressions (from *probabilis*, that which

invites approval, Cicero's Latin for the Greek *pithanos*, persuasive). The theory describes how one may discriminate among impressions by checking to see whether an initially persuasive impression agrees with one's other impressions or if there is anything about the conditions in which it arose that casts doubt on it. Depending on the amount of time available and the importance of the matter at issue, one may perform more or fewer such checks. No amount of checking is sufficient to eliminate the possibility of error, but it will be possible to achieve the degrees of confidence required in different circumstances to make rational action and theoretical inquiry possible. The theory is an early instance of fallibilism.

This account of probable impressions is behind the views that Carneades defended about assent. Sources reveal that he sometimes argued that the wise person will withhold assent, but be able to act and inquire by going along with probable impressions in a way that does not amount to assent; whereas on other occasions, Carneades maintained that the wise person will assent and so form opinions, but with the proviso that he may be wrong. The first view, championed by Clitomachus, is the classical skeptical stance that influenced the other ancient school of skeptics, the Pyrrhonists. The second, which was favored by Philo of Larissa among others, gave rise to a form of probabilism, which is the other legacy of the New Academy.

ETHICS

In ethics Carneades was famous for describing a framework that allegedly classified not only all the views about the goal of life that had been held, but also all those that could be held. He starts with the assumption that practical wisdom must have an object, and one toward which human beings have a natural impulse. He identifies three possibilities: pleasure, freedom from pain, and natural advantages such as health and strength. The principle of virtue is to act with a view to obtaining one of these. There are six simple views, depending on whether the goal is merely to act with a view to obtaining one of the three candidate objects or actually to obtain it. Three further combined views take the goal to be a combination of virtue and actually obtaining the corresponding object. The Stoic position, that virtue is the only good, appears third on the list as the view that the goal is acting with a view to obtaining the natural advantages whether one obtains them or not. At different times Carneades defended the view that the goal is actually to obtain the natural advantages or the view that it is a combination of virtue and pleasure. His aim seems to have been to challenge the Stoics by showing that the considerations captured by the framework do not all point to their view. Carneades's division influenced his successors and through Cicero the understanding of Hellenistic ethical theory.

Other issues that attracted Carneades's attention include Stoic and Epicurean views about fate and moral responsibility and Stoic theology, against the last of which he used a series of Sorites arguments to show that the Stoics could not consistently set any bounds to the divine, so that on their view everything threatened to become divine.

See also Ancient Skepticism; Arcesilaus; Greek Academy.

Bibliography

TEXTS OF CARNEADES

Mette, H. J. "Weitere Akademiker heute Von Lakydes bis zu Kleitomachos." *Lustrum* 27 (1985): 39–148.

STUDIES OF CARNEADES

Algra, Keimpe. "Chrysippus, Carneades, Cicero: The Ethical *Divisiones* in Cicero's *Lucullus*." In *Assent and Argument: Studies in Cicero's Academic Books*, edited by Brad Inwood and Jaap Mansfeld. Utrecht, Netherlands: Brill, 1997.

Allen, James. "Academic Probabilism and Stoic Epistemology." *Classical Quarterly*, n.s., 44 (1994): 85–113.

Allen, James. "Carneadean Argument in Cicero's Academic Books." In *Assent and Argument: Studies in Cicero's Academic Books*, edited by Brad Inwood and Jaap Mansfeld. Utrecht, Netherlands: Brill, 1997.

Bett, Richard. "Carneades' Distinction between Assent and Approval." *Monist* 73 (1990): 3–20.

Burnyeat, Myles. "Gods and Heaps." In *Language and Logos: Studies in Ancient Greek Philosophy Presented to G. E. L. Owen*, edited by Malcolm Schofield and Martha Craven Nussbaum. Cambridge, U.K.: Cambridge University Press, 1982.

Frede, Michael. "The Sceptic's Two Kinds of Assent and the Question of the Possibility of Knowledge." In *Philosophy in History*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge, U.K.: Cambridge University Press, 1984. Reprinted in Michael Frede's *Essays in Ancient Philosophy*. Minneapolis: University of Minnesota Press, 1987.

Frede, Michael. "Stoics and Skeptics on Clear and Distinct Impressions." In *The Skeptical Tradition*, edited by Myles Burnyeat. Berkeley: University of California Press, 1983.

Schofield, Malcolm. "Academic Epistemology." In *The Cambridge History of Hellenistic Philosophy*, edited by Keimpe Algra, Jonathan Barnes, Jaap Mansfeld, and Malcolm Schofield. Cambridge, U.K.: Cambridge University Press, 1999.

Striker, Gisela. "Sceptical Strategies." In *Doubt and Dogmatism: Studies in Hellenistic Epistemology*, edited by Malcolm Schofield, Myles Burnyeat, and Jonathan Barnes. Oxford: Oxford University Press, 1980.

James Allen (2005)

CAROLINGIAN RENAISSANCE

The reign of Charlemagne (768–814) ended the long period of cultural decay and intellectual stagnation that had begun over three centuries before with the barbarian invasions of Western Europe. Despite the disintegration of the Carolingian Empire under Charlemagne's successors, the cultural revival that he inspired continued until the Vikings put an end to it, and even then something of the achievement of the eighth and ninth centuries survived to foster the renaissance of the eleventh and twelfth centuries.

The "Carolingian Renaissance" was dominated by two practical interests, ecclesiastical reform and social progress. Since Charlemagne depended on churchmen to implement his educational policy, the religious motives and ecclesiastical achievements—liturgical reform, monastic renewal, advancement of clerical education—inevitably predominated. Literary sensibility and intellectual curiosity were not, however, wholly lacking in the churchmen of the age, and some charming poems and substantial doctrinal treatises remain to testify to their intellectual versatility.

The chief agent, though not the finest mind, of the Carolingian Renaissance was the Englishman Alcuin (735–804). The Irishman John Scotus Erigena (c. 810–c. 877), the Lombard Paul Warnefrid (d. c. 800), the Spaniard Theodulf of Orleans (d. 821), the Frenchman Remigius of Auxerre (d. c. 908), and the German Rabanus Maurus (d. 856) exemplify the cosmopolitan character of the movement.

The centers of the revival were cathedral and monastic schools established by legislation throughout the Frankish dominions. In addition to a theology consisting mainly of traditional biblical exegesis, their curriculum included the seven liberal arts—the trivium of grammar, rhetoric, and logic and the quadrivium of arithmetic, geometry, astronomy, and music. The assimilation of ancient learning was stressed, and little original work was done; the chief forms of academic literature were commentaries and handbooks.

In philosophy the arts curriculum did not go beyond logic. Several scholars are known to have touched on the question of universal ideas, but the issue does not seem to have been widely debated. The Carolingian Renaissance produced very little speculative philosophy; the great exception, the work of Erigena, stands alone both in its systematic character and in its Neoplatonic inspiration. The few philosophically interesting ideas of the age emerged more or less incidentally in the course of theological reflection and debate.

Perhaps the most important single fragment of philosophical theology to survive from the ninth century is the *Dicta Candidi de Imagine Dei,* attributed to the monk Candidus, schoolmaster at Fulda in 822, which includes the earliest known dialectical demonstration of God's existence by a medieval author. The principle of the proof is the idea of the scale of perfection. Moving from that which simply exists through that which exists and lives and that which exists, lives, and possesses intelligence, the writer argues that the scale would be incomplete without the omnipotent intelligence which is God.

Another small work of some philosophical interest was obviously inspired by consideration of the problem of universals. Fredegisus of Tours (died 834), in his *Epistola de Nihilo et Tenebris,* assumes that every term has some real entity corresponding to it. He concludes that the "nothing" (*nihil*) of the orthodox Christian doctrine of creation "out of nothing" must be conceived as a preexistent, undifferentiated stuff out of which God created everything, including human souls and bodies. Fredegisus was evidently an early instance of a theological dialectician who found difficulty in reconciling the results of his logical analysis of the meaning of terms with doctrinal orthodoxy; the problem was not widely recognized as urgent until the eleventh century.

The outstanding intellectual issue of the Carolingian Renaissance was unquestionably the problem of predestination. The German monk Gottschalk (d. c. 868) was accused of teaching that from eternity God has infallibly predestined some men to salvation and others to damnation; that God therefore does not in any sense will the salvation of all men; that Christ's atoning sacrifice was offered only for the elect; and that each man's will is irresistibly determined either to good or to evil. The authority of Augustine and of his great disciples Fulgentius of Ruspe and Prosper of Aquitaine was invoked by Gottschalk and others in favor of these ideas. In opposition to this intransigent Augustinianism, Erigena expounded a libertarian doctrine, inspired by Greek thought; others sought a middle way within the Augustinian tradition. The controversy was long and heated, and its terms were not always clearly defined, but it is obvious that the crucial issue was the relation between divine immutability and omnipotence, on the one hand, and human freedom and moral responsibility, on the other. After a series of conflicting synodical decisions, the moderate Augustinians were officially vindicated, but the

debate was to be repeatedly renewed in the later Middle Ages and the Reformation and Counter-Reformation.

A second vigorous controversy of the period had to do with the presence of Christ in the Eucharist. Paschasius Radbertus (d. c. 860), in his *De Corpore et Sanguine Domini,* the first technical elaboration of Eucharistic doctrine in theological history, asserted the identity of the sacramental elements with the historical body of Jesus crucified and glorified. Although he insisted at the same time on the spiritual and mystical manner of Christ's presence, some of his statements could be interpreted in a crudely materialistic sense, and Ratramnus (d. 868), in his *De Corpore et Sanguine Domini,* opposed an ostensibly symbolist doctrine to the realism of Radbertus; owing to vagueness of definition, however, it remains uncertain how far and in precisely what way the two doctrines were incompatible. The debate is significant primarily because it eventually issued in the definition of the dogma of transubstantiation by the Fourth Lateran Council (1215) and in the subtle metaphysical elaboration of that dogma in the theology of Thomas Aquinas.

See also Alcinous; Augustine, St.; Augustinianism; Determinism, A Historical Survey; Erigena, John Scotus; Libertarianism; Reformation; Thomas Aquinas, St.

Bibliography

TEXTS

Alcuin

Duemmler, E., ed. *Monumenta Germaniae Historica, Epistolae.* Vol. 4. Berlin: Weidmann, 1895.

Duemmler, E., ed. *Monumenta Germaniae Historica, Poetae Latini Aevi Carolini.* Vol. 1, 160–351. Berlin: Weidmann, 1881.

Migne, J.-P., ed. *Patrologiae cursus completus. Series Latina.* Paris, 1844–. Vols. 100–101.

Candidus: Dicta Candidi de Imagine Dei

Hauréau, B. *Histoire de la philosophie scolastique.* Vol. 1, 134–137. Paris, 1872.

John Scotus Eriugena

Sheldon-Williams, I. P. ed., and L. Bieler, col. *Johannis Scotti Eriugenae Periphyseon (De divisione naturae).* Dublin: Dublin Institute for Advanced Studies, 1968–.

Fredegisus of Tours: Epistola de Nihilo et Tenebris

Fredegisus. *Epistola de Nihilo et Tenebris.* In J. P. Migne, *Patrologia Latina.* 221 vols. Paris, 1844–1864. Vol. CV, cols. 751–756.

Gottschalk

Strecker, K., ed. *Monumenta Germaniae Historica, Poetarum Latiorumi Medii Aevi.* Vol. 6, 86–106. Munich, 1951.

Hrabanus Maurus

Duemmler, E., ed. *Monumenta Germaniae Historica, Poetae Latini Aevi Carolini.* Vol. 2, 154–258. Berlin: Weidmann, 1884.

Paschasius Radbertus

Paschasius Radbertus. Works Vol. CXX. Berlin: Weidmann, 1884.

Paulus, B. ed. *Pascasius Radbertus De corpore et sanguine Domini: cum appendice Epistola ad Fredugardum.* Turnhout, Belgium: Brepols, 1969.

Traube, L., ed. *Monumenta Germaniae Historica, Poetae Latini Aevi Carolini.* Vol. 3, 38–53. Berlin: Weidmann, 1896.

Paul the Deacon

Duemmler, E., ed. *Monumenta Germaniae Historica, Poetae Latini Aevi Carolini.* Vol. 1, 35–86. Berlin: Weidmann, 1881.

Waitz, G., ed. *Monumenta Germaniae Historica, Scriptores Rerum Langobardicarum et Italicarum.* Hannover: Weidmann, 1878.

Ratramnus

Migne, J.-P., ed. *Patrologiae cursus completus. Series Latina.* Paris, 1844–. Vol. 121, cols. 1–346, 1153–1156.

Ratramnus. Works. Vol. CXXI.

Remigius of Auxerre

Edwards, B., ed. *Remigii Autissiodorensis Expositio super Genesim.* Turnhout, Belgium: Brepols, 1999.

Lutz, C. E., ed. *Commentum in Martianum Capellam.* 2 vols. Leiden: Brill, 1962–1965.

Silk, E. T., ed. *Saeuli noni auctoris in Boetii Consolationem philosophiae commentarius.* Rome: American Academy in Rome, 1935.

Theodulf of Orleans

Duemmler, E., ed. *Monumenta Germaniae Historica, Poetae Latini Aevi Carolini.* Vol. 1, 437–581. Berlin: Weidmann, 1881.

STUDIES

Colish, M. L. *Medieval Foundations of the Western Intellectual Tradition, 400-1400.* New Haven, CT: Yale University Press, 1997.

Colish, M. L. *The Mirror of Language; A Study in the Medieval Theory of Knowledge.* New Haven, CT: Yale University Press, 1968.

Marenbon, J. *Early Medieval Philosophy (480–1150): An Introduction.* New York: Routledge, 1983.

Marenbon, J. *From the Circle of Alcuin to the School of Auxerre: Logic, Theology, and Philosophy in the Early Middle Ages.* Cambridge, U.K.: Cambridge University Press, 1981.

McCracken, G. E., ed. *Early Medieval Theology.* Philadelphia: Westminster Press, 1957.

McKitterick, R., ed. *Carolingian Culture: Emulation and Innovation.* Cambridge, U.K.: Cambridge University Press, 1994.

Moran, D. *The Philosophy of John Scottus Eriugena: A Study of Idealism in the Middle Ages.* Cambridge, U.K., and New York: Cambridge University Press, 1989.

Sullivan, R. E. "The Carolingian Age: Reflections on Its Place in the History of the Middle Ages." *Speculum: A Journal of Medieval Studies* 64 (1989): 267–306.

Waddell, Helen. *The Wandering Scholars,* 6th ed. London, 1932.

Eugene R. Fairweather (1967)
Bibliography updated by Joseph Pucci (2005)

CARROLL, LEWIS

(1832–1898)

Lewis Carroll is the pen name of Charles Lutwidge Dodgson. The eldest son of a large clerical family, he was born at Daresbury, Cheshire, was educated at Rugby School, and entered Christ Church, Oxford, in 1850. On obtaining first-class honors in mathematics in 1854, he was appointed student and mathematical lecturer of the college, and remained on its foundation until his death. In many ways an archetype of the pernickety bachelor don, Dodgson had a wholly uneventful academic career. Hampered by a stammer, he shone neither as lecturer nor as preacher (he took deacon's orders in 1861). He embroiled himself—often amusingly, although usually without effect—in academic politics, was for a time curator of the college common room, and visited Russia in 1867. His leisure was spent in gallery-going and theatergoing; in photography, at which he was an expert; in the writing of light verse; and in the patronage of an interminable succession of small girls. The last peculiarity has endeared him to psychoanalytical biographers, who would seem, however, to have enriched the literature of nonsense on the subject more often than they have been able to explain it.

Dodgson the mathematician published a number of books and pamphlets, none of any lasting importance. The best known is *Euclid and His Modern Rivals* (London, 1879); the most useful, probably his edition of *Euclid I & II* (London, 1882); and the most original, his contributions to the mathematical theory of voting, to which attention was drawn by D. Black in his *Theory of Committees* (Cambridge, U.K., 1958). Dodgson's mathematical outlook was, in general, conservative and provincial, aiming no higher than the improvement of elementary teaching or routine calculation. His talent found greater scope in the construction of puzzles contained in *A Tangled Tale* (London, 1885) and *Pillow Problems* (London, 1893), which at times show depth as well as ingenuity. The same can be said of his dabblings in symbolic logic, which otherwise make little advance on the work of Augustus De Morgan and John Venn. His *Game of Logic* (London, 1887) and *Symbolic Logic, Part I* (London, 1893) present logic merely as a mental recreation devoted to the solution of syllogistic problems by means of a square diagram and colored counters. His logical output was completed by nine papers on elementary logic and by two short pieces in *Mind* (n.s., 3, 1894 and n.s., 4, 1895). His influence is to be seen mainly in the attempts of later logicians to imitate the elegant absurdity of his examples. Their failure merely emphasizes the rarity of his own peculiar gift.

Needless to say, that gift finds its happiest exercise in his writings for children. *Alice in Wonderland* (London, 1865), *Through the Looking-Glass* (London, 1871), and *The Hunting of the Snark* (London, 1876) and, to a lesser extent, the two parts of *Sylvie and Bruno* (London, 1889 and 1893), are the only works that keep his name alive—or deserve to do so. Apart from *Pickwick,* and perhaps *Waverley,* they seem also to be the only works of fiction generally known to philosophers, and have been constantly pillaged for quotations. All five are dream narratives or have episodes depicting dreams, whose aberrant logic is responsible for much of their philosophic interest and fun. *Alice in Wonderland* exploits the idea of sudden variations in the size of the heroine; its sequel, the conception of a world in which time, space, and causality are liable to operate in reverse. The characters—a bizarre medley of nursery and proverbial figures, animals (fabulous or otherwise), plants, playing cards, and chessmen—are all much addicted to argument; and their humor, where it does not rely upon puns, is largely a matter of pursuing logical principles to the point of sophistry or absurdity. The frog, who supposes that an unanswered door must have been asking something, is a simple case in point. The King of Hearts and the White King, who both take "nobody" for a person, are victims of the same error and have often been cited as a warning to less venial, because less nonexistent, hypostatizers of the null class.

These books are further remarkable for their echoes—and pre-echoes—of philosophic controversy. Tweedledum and Tweedledee are Berkeleian metaphysicians, and the latter has notions of logic that bespeak the influence of Gottfried Wilhelm Leibniz. Alice herself, on the road to their house, is a step ahead of Gottlob Frege in discovering the difference between *Sinn* and *Bedeutung.* Humpty Dumpty has been taken, on anatomical grounds, for a Hegelian; but his ascription of fixed meaning to proper names and denial of it to general terms, plus his confident philology and shaky mathematics, proclaim him beyond doubt an early, if eccentric, linguistic analyst. The White Knight's reactionary views on the mind-body question give no hint of the metalinguistic virtuosity he later displays in the announcement of his song. The distinctions there enunciated have been formalized by Ernest Nagel in

"Haddocks' Eyes" (in J. R. Newman, *The World of Mathematics,* New York, 1956, Vol. III, pp. 1886–1890). They would not have troubled the Duchess, another adroit logician, although her primary interest is in morals. Her cat, on the other hand, although adept enough at defying the principle that an attribute must inhere in a substance, offers a regrettably invalid proof of its own madness, as does the pigeon of Alice's serpentinity. The Hatter, March Hare, and Dormouse are sounder reasoners; whatever their troubles with time, they know a fallacy of conversion when they see one, and it is no great wonder that Messrs. Bertrand Russell, George Edward Moore, and John McTaggart, who were supposed to resemble them, should have been known at one time as the "Mad Tea Party of Trinity."

Not even Nobody, in his senses, would venture to identify that other and more formidable trio, the Queen of Hearts and her chessboard cousins. The former's principle of government by decapitation scarcely ranks as a political theory; but the White Queen is respected by philosophers both for her abilities in believing the impossible and for her success in proving, for the special case of jam at least, that the future *will* resemble the past, if not the present. The Red Queen is no less celebrated, among physicists, for her anticipations of the theory of relativity. In this, however, she meets competition from the Bellman in the *Snark,* who has been acclaimed, on the strength of his map, as the first general relativist and is, in any case, the undisputed inventor of an interesting three-ply version of the semantic theory of truth ($\vdash$p. $\vdash$p. $\vdash$p ≡ "p" is true). Of his crew members, the Baker, with his lost identity and Heideggerian premonitions of impending *Vernichtung,* has been plausibly represented as a protoexistentialist; but the other protagonists still abide the conjecture of commentators, as do the quest and the quarry itself. The Snark has been taken for everything from the Tichborne inheritance to the North Pole, and from a business depression to the atom bomb. F. C. S. Schiller's interpretation of it in *Mind!* (1901, pp. 87–101) as the Absolute is elaborately argued, and doubtless finds an echo in the *Oxford Dictionary*'s definition of the creature as a "chimerical animal of ill-defined characteristics and potentialities"; but its fondness for bathing machines is not really explained thereby, and the theory founders completely on the Bellman's explicit assertion, confirmed by the Baker's uncle, that Snarks are Many and not One. Nobody, it is true, has been more successful than Schiller on this point, and his views have been generally accepted; but the opinions of nonentities have no place in a grave work of learning such as the present, so neither use nor mention of them is appropriate here.

See also Berkeley, George; De Morgan, Augustus; Frege, Gottlob; Heidegger, Martin; Leibniz, Gottfried Wilhelm; Logic, History of; McTaggart, John McTaggart Ellis; Moore, George Edward; Nagel, Ernest; Russell, Bertrand Arthur William; Schiller, Ferdinand Canning Scott; Venn, John.

Bibliography

Apart from the standard *Life and Letters of Lewis Carroll* by his nephew, S. D. Collingwood (London: T.F. Unwin, 1899), the soberest accounts of Carroll's life are Derek Hudson, *Lewis Carroll* (London: Constable, 1954), and Roger Lancelyn Green, *Story of Lewis Carroll* (New York: Henry Schuman, 1950) and *Lewis Carroll* (London: Bodley Head, 1960).

The least incomplete version of Lewis Carroll's works is *The Complete Works of Lewis Carroll* (London and New York, 1939). The most philosophical editions are *The Annotated Alice* (New York: C.N. Potter, 1960) and *The Annotated Snark* (New York: Simon and Schuster, 1962), edited by Martin Gardner.

The pioneer work of logical investigation in this field is P. E. B. Jourdain, *The Philosophy of Mr B*rtr*nd R*ss*ll* (Chicago: Open Court, 1918). Further light on the subject may be obtained, inter alia, from R. B. Braithwaite, "Lewis Carroll as Logician," in *Mathematical Gazette* 16 (1932): 174–178; P. Alexander, "Logic and the Humour of Lewis Carroll," in *Proceedings of the Leeds Philosophical and Literary Society* 6, Part 8 (1951): 551–566; and, despite some inaccuracies, from W. Weaver, "Lewis Carroll, Mathematician," in *Scientific American* 194 (4) (1956): 116–128; and R. W. Holmes, "The Philosopher's *Alice in Wonderland,*" in *Antioch Review* 19 (2) (1959): 133–149.

P. L. Heath (1967)

CARROLL, LEWIS [ADDENDUM]

The success of the "Alice" books established Charles L. Dodgson's reputation as a gifted writer of children's literature. His admirers expected humor in everything he wrote from then on, an attitude that affected the reception of his serious pieces and prevented his work from contributing to the development of their subjects. For example, the more amusing *Euclid and His Modern Rivals* (1879) overshadowed his more important book, *Curiosa Mathematica. Part 1. A New Theory of Parallels* (1888).

Dodgson made significant contributions to linear algebra in *An Elementary Treatise on Determinants* (1867), a book that though marred by odd notation and unusual terminology, contains the first written proof of a standard theorem connecting the rank of a matrix with the existence of solutions to certain linear systems (chap-

ter 4, proposition II). One of his techniques, condensation, was used in an early step of the solution to the alternating sign matrix problem (Bressoud 1999).

In the field of cryptology his five cipher systems, based on the three cipher paradigms of his time (Vigenère, Beaufort, Variant Beaufort) are not well known. These were: Key-Vowel, Matrix, Alphabet, Telegraph, and Memoria Technica. The first two (1858, unpublished) were unbreakable from a practical point of view. The third and fourth (1868) were secure by the standards of his time for ordinary telegrams and mailed postcards. The last (1875), directly tied to word games, was the most literary.

His publications on the theory of voting consisted of four pamphlets, three written between 1873 and 1876, and *The Principles of Parliamentary Representation* (1884). The pamphlets of the 1870s, an outcome of Dodgson's involvement with college and university affairs, reflect his independent rediscovery of Condorcet's cyclical majorities and include the first application of game theory to sophisticated voting. The argument of the 1884 pamphlet, written to influence the outcome of two electoral reforms, a goal it did not accomplish, is based on the zero-sum game. Dodgson was the first to treat formally apportionment (allocating seats to districts) and proportional representation (assigning seats to political parties) together.

Dodgson's contributions to logic have been widely recognized since William Warren Bartley, III's edition of *Lewis Carroll's Symbolic Logic* (1977) which includes the unpublished manuscript of part 2 of Dodgson's *Symbolic Logic*. Dodgson developed a formal logic where he set down intuitively valid rules for making inferences. A comparison of the two parts reveals the progress he made toward an automated approach to the solution of multiple connected syllogistic problems, many being humorous puzzle problems. The most important of his techniques, the method of trees, foreshadowed modern concepts and techniques in automated reasoning that were developed from the 1950s. Dodgson's use of existential import, abandoned in modern logical usage, marred the reception of part 1 of his book. He developed a method of diagrams as a visual proof system for syllogisms that he introduced in *The Game of Logic* (1887). Like his tree test, which is a proof system for soriteses, it is sound and complete. His self similar diagrams (invariant under a change of scale) are capable of handling existential statements and are easily extended to any number of sets using a linear iterative process. In this regard, they are superior to the diagrams described by John Venn in 1880.

See also Logic Diagrams; Logic, History of; Logic, Traditional; Venn, John.

Bibliography

Abeles, Francine F. "Lewis Carroll's Ciphers: The Literary Connections." *Advances in Applied Mathematics* 34(2005): 697–708.

Abeles, Francine F., ed. *The Mathematical Pamphlets of Charles Lutwidge Dodgson and Related Pieces*. New York: Lewis Carroll Society of North America; Charlottesville: Distributed by the University Press of Virginia, 1994.

Abeles, Francine F., ed. *The Political Pamphlets and Letters of Charles L. Dodgson and Related Pieces*. New York: Lewis Carroll Society of North America; Charlottesville: Distributed by the University Press of Virginia, 2001.

Bressoud, David M. *Proofs and Confirmations. The Story of the Alternating Sign Matrix Conjecture*. Washington, D.C.: The Mathematical Association of America; Cambridge, U.K.: Cambridge University Press, 1999.

Cohen, Morton N. *Lewis Carroll. A Biography*. New York: Knopf, 1995.

Francine F. Abeles (2005)

CARTESIANISM

According to one panoramic view of modern philosophy, René Descartes is the father and Cartesianism an inherited characteristic or family trait. With no disparagement intended of this assessment of Descartes's influence, the term *Cartesianism* will be used here in a less contentious way to refer to the multifarious, more or less self-conscious efforts on the part of his contemporaries and immediate successors to supply what they found lacking in his ambitious attempt to reconstitute human knowledge. Three directions of their activities can be distinguished and, corresponding to them, three particular applications of the term *Cartesianism.*

(1) It was evident that Descartes's project of a universal and all-encompassing science of nature was not fully realized. His intended *summa philosophiae, Principia Philosophiae* (*Principles of Philosophy,* Amsterdam, 1644), lacked the proposed parts on plants and animals and man; and his posthumously published and widely read *Traité de l'homme* (*Treatise on Man,* Paris, 1664) ended abruptly. Moreover, in his *Discours de la méthode* (*Discourse on Method,* Leiden, 1637) and in the letter prefacing the French translation of the *Principles* (Paris, 1647), he asked for assistance in carrying out his program for the sciences, suggesting that cooperative endeavor in the acquisition of *expériences* would be necessary to decide among equally possible explanations of the more partic-

ular facets of nature. His early admirers, attracted as much—and often far more—by his physics than by his metaphysics, accepted the invitation, and, working within the framework of his methodological prescriptions and cosmologic theory, distinguished themselves not only from their scholastic opponents of the academic establishment but also from other non-Aristotelian scientists of the time whose work went against views they had inherited. In the seventeenth century, *les cartésiens* were predominantly Descartes's followers in physics; and the term *Cartesianism* has acquired some of its less favorable associations from its application to this maligned movement in the history of science.

(2) A second line of development can be traced from Descartes's novel use of the term *idea* in presenting what has sometimes been considered the characteristically Cartesian view that knowledge is attained by way of ideas. These "as it were images of things" (*tanquam rerum imagines, veluti quasdam imagines*), as they were introduced in the *Third Meditation,* were variously described in his works, and a host of questions arose about their origin and nature. "Orthodox" Cartesians differed in their interpretations of Descartes's answers to these questions, while the more independently minded, accepting the thesis that knowledge is attained by way of ideas, produced deviant answers of great subtlety and originality. Since John Locke and his followers accepted Descartes's general thesis although they disagreed on the subject of innate ideas, Cartesianism, in a second application of the term, has been taken to cover a considerable domain, including family squabbles among rationalists and empiricists as well as more recent disputes, such as that about the genesis and status of sense data. (It should be noted that this use of "Cartesianism" to refer to the "way of ideas" differs from another use, in which "Cartesianism" and "rationalism" are roughly coextensive and connote a view or views about innate ideas or principles.)

(3) When Descartes was presented with objections to his metaphysics framed in terms of traditional categories and distinctions, a number of thorny problems became apparent; notably, concerning the substantiality and causal efficacy of his seemingly formless and inert corporeal things and concerning the union in man of a body and a soul, or mind, that is alleged to be really distinct from the body. In these sensitive areas, Descartes's teachings were interpreted and developed in various ways; and those who chose to follow the natural light rather than Descartes came to conclusions far removed from, and incompatible with, his. Yet, because of a common view concerning the distinction of mind and matter, Nicolas Malebranche and Benedict (Baruch) de Spinoza, as well as some less celebrated metaphysicians, have been called Cartesians; and Cartesianism, in a third acceptation of the term, comprises various monist, pluralist, and occasionalist variations on a common metaphysical theme. Within the limits of this general survey of Descartes's influence, Cartesianism will be mapped in each of the three general areas to which the term has been applied.

PHYSICS AND DERIVATIVE SCIENCES

Like Descartes, the Cartesians attracted to his program for the sciences thought of themselves as possessing a powerful method for investigating nature; and, though they disagreed with him and among themselves on particular applications, they accepted a general theory in physics, salient features of which were the laws of motion in Part II of the *Principles*; the theory of vortices in Part III; and the doctrine of subtle matter that underlies explanations of various phenomena, both celestial and terrestrial, in Parts III and IV of the *Principles.* Although Descartes's laws of motion became increasingly troublesome—Malebranche accepted them at first but was later forced to modify them beyond recognition—the cosmogonic picture of which they were part was altered but not effaced. It was an integral feature of the picture that Earth, like the other planets, was transported in a whirlpool that centered about the sun; and, while Descartes took pains in the *Principles* to distinguish his view from that of Nicolas Copernicus and to point out that, in his view and according to his definitions, Earth, though indeed a planet, was, strictly speaking, at rest, his followers were less concerned to establish a difference. They, too, rejected the possibility of unoccupied space or a vacuum, and claimed that apparently empty spaces—the heavens, the "pores" of bodies, and experimentally produced vacuums—were actually filled with subtle matter. Like Descartes, they made free use of the adaptable particles of subtle matter in their jigsaw-puzzle explanations of the workings of nature. There was some question as to what they conceived the vaunted "true" method to be, as evidenced by Gottfried Wilhelm Leibniz' skeptical queries. Nonetheless, some general characteristics of their practice were apparent.

Following the rule of evidence in the *Discourse,* they understood Descartes's injunctions against preconception and precipitancy as condemnations of merely accepted opinion and of idle speculation; and contrary to a popular conception of their apriorism, they were keenly interested in the detailed observation of nature and in experiments, thinking of themselves as countering the

bookish physics of the Scholastics and the wanton practices of alchemists, astrologers, and the like. Lenses, Torricellian tubes, and sundry apparatus were much in evidence; and, like Descartes, many of them took pleasure in anatomical and physiological investigations. To what use they put their observations and experiments is one thing; their cult of *expériences,* another—and an indisputable fact. The requirement of clear and distinct ideas was met in the doctrine that matter is extension and the corollary that change is local motion, or *translatio.* The methodological implications of these complex views were manifold. Negatively, they ruled out explanations involving qualitative entities or "real" qualities, such as light, heat, and weight, in physics, and substantial forms, such as vegetative and sensitive souls, in biology. Also banished were final causes, including natural place, gravitation, and attraction; faculties, virtues, and powers as causes of change; and sensible qualities supposed to inhere in bodies and to be mysteriously purveyed to us by intentional species. Distinctly conceived, bodies were geometrical solids occupying parts of space and were subject to alteration by the crowding, or impact and pressure, of their neighbors. A vacuum, or void, was thought impossible, as were, at least for the "orthodox" Cartesians, indivisible particles or atoms. Sharing corpuscular and mechanistic assumptions with other nonscholastic scientists, they showed the mark of the master in their geometrical notions of—or, as some would have it, their lack of concepts of—mass and force. Quantity of matter was volume; weight was a centripetal reaction in a vortex of bodies of a certain size. Force, as effort or action on the part of bodies, was as suspect as were the powers and virtues of the Scholastics. Distinctly conceived, it was derived from a principle of inertia, and the force of a body in motion was reckoned as the product of mass (volume) and velocity.

HOLLAND: REGIUS AND CLAUBERG. During Descartes's long expatriation in Holland, he made a number of converts to his program for the sciences; and despite outbreaks of official opposition, Cartesianism made an impression on academic life that it did not make in France.

Regius. Of special note is Descartes's sometime friend and disciple Henry de Roy, or Regius (1598–1679), professor of medicine at the University of Utrecht, who typified Cartesian scientists in following the master more or less closely in physics and the derivative sciences while departing from his views in metaphysics. His *Fundamenta Physices* (Amsterdam, 1646), which appeared two years after the *Principles,* recapitulated the physics of Parts II, III, and IV, to which were added views from the earlier *Meteors* and *Dioptric* and also from unpublished work. Regius's physics, unlike Descartes's in the *Principles,* was not represented as derived from metaphysical principles. Moreover, in the concluding chapter on man, adverting to issues concerning the soul, he presented views to which Descartes could only take exception. In the preface to the French translation of the *Principles* (1647), Descartes disowned both the physics and the metaphysics of his disciple; and Regius in turn circulated a defense of his metaphysical theses, arguing for an empiricist view of the origin of ideas and against the necessity of a real distinction of mind and body. Descartes's reply to Regius, his *Notae in Programma* (1648), contained the prototype of later defenses of innate ideas against empiricist incursions. Innate ideas, he maintained, need not be *actually* present in the mind. Moreover, certain ideas—for example, of God—differ in kind from "adventitious" ideas; and even the latter do not, strictly speaking, come to us from the senses, that is, the sense organs.

Clauberg. From Holland, Cartesianism was taken to Germany by Johannes Clauberg, who attempted to explain and defend both Descartes's physics and his metaphysics. Working out apparent implications of the metaphysics in *De Cognitione Dei et Nostri …* (Duisberg, 1656), he too came to hold a deviant view of the relation of mind and body (though not Regius's), a view linking him with the occasionalists. Clauberg also faced the problem of the relation of traditional logic and Cartesian methodology, and his work in logic anticipated the more famous *Logique, ou L'art de penser* (*Port-Royal Logic,* 1662) of Antoine Arnauld and Pierre Nicole, which was the chief contribution of the Cartesians (Leibniz, of course, excluded) to logic.

FRANCE: ROHAULT AND RÉGIS. In France, Cartesianism, though it was not received in the universities and was, in effect, interdicted in 1671, flourished in extra-academic circles. Dissemination of Descartes's unpublished works and letters was in the hands of his devoted admirer Claude Clerselier (1614–1684), while leadership of his scientific enterprise devolved upon Jacques Rohault.

Rohault. The most gifted of the Cartesian scientists, Rohault devised ingenious experiments for his popular weekly meetings and presented the results of his work in his influential *Traité de physique* (Paris, 1671; translated by John Clarke as *System of Natural Philosophy,* London, 1723). Like Regius, he was inclined to separate Descartes's

physics from his metaphysics; and, in line with this, he developed Descartes's notion of hypothesis or *supposition,* eliminating, however, any qualification to the effect that hypotheses were to be accepted for lack of something better.

Régis. Pierre-Sylvain Régis succeeded Rohault as leader of the Cartesian school. In his *Système de philosophie …* (Paris, 1690), a comprehensive work containing sections on logic, metaphysics, and moral philosophy as well as his extensive physics, he assimilated work that had been done since Descartes's death. The apogee of the Cartesian movement in physics has been set at about the time of Régis's *Système* and of Bernard Le Bovier de Fontenelle's imaginative exploration of the vortices in his *Entretiens sur la pluralité des mondes* (Paris, 1686).

CRITICAL RECEPTION. While receiving acclamation, the Cartesians were simultaneously threatened—and eventually discredited—by discoveries, such as that of the finite velocity of light, that contravened crucial parts of their system and by the objections and strictures of Leibniz and of Isaac Newton and his followers. These adverse judgments have been generally accepted. It is commonplace (and true) that Newton showed beyond the shadow of a doubt the incompatibility of the theory of vortices and Johannes Kepler's laws, while Leibniz neatly proved the inconsistency of Descartes's laws of motion with Galileo Galilei's. Citing Leibniz' derogatory characterization of the Cartesians, the not unsympathetic historian Charles Adam has reiterated comments on the paucity of equations in their work and on the uncontrolled play of their imagination in assigning jobs to the ubiquitous particles of subtle matter. His verdict was that Descartes's physics threatened to become as harmful to the progress of science as Aristotle's had been.

Yet, more recently, some less disparaging comments have been made. The picture is considerably brightened when Malebranche and especially Christiaan Huygens (1629–1695) are, by virtue of obvious influences, included among the Cartesians (as in Paul Mouy's [1934] account.) It has also been suggested that the attempted geometrization of physics was premature rather than perverse (Mouy; Max Jammer, *Concepts of Force,* Cambridge, U.K., 1957) and that the unstable and indeterminate particles of the Cartesians, not the billiard-ball atoms of the opposition, were in line with things to come (Geneviève [Rodis-] Lewis, *L'individualité selon Descartes,* Paris, 1950). Nonetheless, Descartes's followers in physics and the derivative sciences, Malebranche and Huygens aside, have not, on the whole, enhanced his reputation.

THEORY OF KNOWLEDGE

Proposing, in the *Third Meditation,* the term *idea* for those of his thoughts that are the "as it were images of things," Descartes proceeded to classify ideas according to their apparent origin—as innate or adventitious or made by him. He introduced distinctions bearing on their nature—between formal and material truth or falsity, and between objective and formal reality. Discussions generated by these passages concerned both Descartes's intent and the tenability of the views attributed to him. Four main problems can be distinguished, two relating to the tentative classification of ideas according to origin and two having to do with the distinctions bearing on their nature.

INNATE IDEAS. The contratraditional notion of innate ideas—that is, of ideas not derived in some way from the senses but instead having their source in the mind itself—presented an obvious difficulty; namely, how could such an idea, taken to be the form of a thought, exist or preexist in a person's mind if he did not in fact have the thought or indeed never had it? It seemed that Descartes's metaphor of a treasure house in which these ideas were stored needed to be cashed—a process that he attempted and that was carried out in various ways, in the face of some formidable difficulties, by supporters of his doctrine of innate ideas.

ADVENTITIOUS IDEAS. It was evident that ideas provisionally classified as adventitious—for instance, of a sound, the sun, or a fire—could not, strictly speaking, come to us from external objects; for, in Descartes's view, there was nothing in the objects or in the sense organs exactly like these ideas, or at least like many of them. Although these ideas could, in some sense, be said to be caused by external objects, they could not, strictly speaking, originate there; and some other cause or source more in keeping with their nature seemed to be necessary. Descartes suggested that the mind had the faculty or power of forming these ideas on the occasion of motions in the brain and that ideas seeming to come to us from without were in fact innate. Both suggestions were explored by his successors.

MATERIALLY FALSE IDEAS. Noting that falsity (formal falsity) was to be found in judgments and not in ideas, Descartes added that nonetheless certain ideas—for instance, the idea of cold—might be materially false; that is, if cold were a privation, then the idea of cold, representing a privation or what is not a thing, as if it were a thing would be materially false. The implications to be

drawn from this remark were that, in his view, ideas of sensible qualities—of heat as well as cold, of sounds, colors, and the like—were materially false; and questions arose as to whether the notion of a materially false idea (literally, an idea misrepresenting what is not a thing) made sense, and whether sensations of heat, cold, and the like were, in a strict sense of the term, ideas. Two models seemed to be at work in Descartes's account of sense perception, and a problem bequeathed to his followers was that of specifying the latent distinction between the nonrepresentational and the representational elements—sensations and ideas properly so called—that were supposed to be ingredients of sense experience.

IDEAS OF EXTENDED THINGS. There was also a problem concerning ideas of extended things derived from the dual reality—objective and formal—accorded them. As representations, it seemed that they must have something in common with, or be in some respect like, the extended things they represented. Nevertheless, it was taken to follow from their formal reality as modes of thought that they were totally unlike extended things. A dilemma presented itself: Either ideas of extended things were totally unlike extended things, in which case they could not represent them; or, if they were in some respect like extended things, then they could not be accommodated in the mind.

MALEBRANCHE. Malebranche, among others, addressed himself to these problems; and, in his elaborate discussions of the nature and origin of ideas and in the numerous polemics to which they gave rise, various answers were surveyed and the major lines of development of Descartes's theory of knowledge were represented. Regarding the problem of materially false ideas and the difficulty concerning ideas of extended things, Malebranche, in the numerous editions of *De la Recherche de la vérité* (first published 1674–1675) and in the *Éclaircissements* added to them, drew a sharp distinction between the perception of heat, color, and the like and the perception of objects as extended. The former consisted in sensations or feelings (*sentiments*), nonrepresentational modifications of the mind conceived on the analogy of feelings of pain, and did not, in his precise use of the term, involve ideas (*idées*). The latter required ideas, which were distinguished from the mind's awareness of them and were not, in his view, modifications of the soul. Approaching the problem of the location or status of these ideas, Malebranche investigated a number of possibilities suggested by Descartes's tripartite classification (adventitious, made by the mind, and innate). Finding difficulties in the suggested sources, he concluded that ideas of extended things were neither adventitious nor made by the mind nor innate. The arguments against these possibilities served as indirect evidence for his own thesis: that these ideas were (as in a medieval use of the term) archetypes of created things in the Divine Understanding and that the human mind, intimately united with God, perceived created, extended things by way of ideas in him. Because, in this theory, ideas of extended things were not modifications of the human mind, the problem of their existence in an unextended mind did not arise, though, as became evident in the ensuing controversies, there was a related problem about the possibility of their existence in God.

FOUCHER. Two of the polemics were especially revealing. In his *Critique de la recherche de la vérité …* (Paris, 1675) and subsequent writings, Simon Foucher, though he misunderstood parts of Malebranche's tortuous theory, raised problems worthy of serious consideration. First, he urged that, if ideas of extended things had to have something in common with what they represented, they could not be, as he at first wrongly interpreted Malebranche, modifications of the mind or—as Malebranche in fact believed—inhabitants of the divine understanding. Second, granted that ideas of extended things were not modifications of the human mind but were divinely situated, could they be immediately perceived? The basis of the question was that, if immediate perception were tied to Descartes's views about indubitability and the *cogito,* then we could not be immediately aware of anything outside or apart from the mind. Third, he also questioned the distinction (to use Locke's terms) of primary and secondary qualities along lines that were continued by Pierre Bayle and George Berkeley, noting what, in Malebranche's distinction of sensation and idea, seemed to require explanation: that, when we perceive an object, we are aware of one uniform appearance of something having both shape and color. Unfortunately, Malebranche was inclined to dismiss Foucher's criticisms on the ground of misinterpretation, but Dom Robert Desgabets (d. 1678), in his *Critique de la Critique de la recherche de la vérité …* (Paris, 1675), attempted to defend Cartesian views (though not Malebranche's peculiar versions of them) against this attack.

ARNAULD. The most interesting controversy was with Arnauld, who, in *Des Vrayes et des Fausses Idées* (Cologne, 1683), attacked Malebranche's view of ideas as entities distinct from the mind's perception of them by tracing the source of this view to a misconceived analogy with

ocular vision and a confusion of presence in the mind with local presence. For Arnauld, as for Descartes, ideas were modes of thought; and, as Descartes was content to explain the objective presence of objects in the mind as the way they were wont to be there, so Arnauld took it to be the nature of thought or mind, requiring no explanation of the kind Malebranche proffered, to represent objects—near or at a distance, present or absent, real or imaginary. Though Malebranche was not moved by this attempt to impugn his theory as the answer to a pseudo problem, in the course of the controversy he was forced to articulate his view that we perceive extended things in God, not by way of individual archetypes but by way of infinite, intelligible extension, which is the common archetype of all extended things, actual or possible.

LOCKE AND LEIBNIZ. A significant event in the annals of the Cartesian theory of knowledge was the publication of Locke's *An Essay concerning Human Understanding* (London, 1690). Locke's attack on innate ideas and principles and Leibniz' defense in his *Nouveaux essais sur l'entendement* (published posthumously, Amsterdam and Leipzig, 1765) are a long story, that cannot be told here. Suffice it to say that, in this division of Cartesianism into empiricism and rationalism, Leibniz used arguments like Descartes's in the *Notae in Programma* and, on this question, represented the orthodox Cartesian point of view.

METAPHYSICS

The occasionalist, monist, and pluralist developments included in the third application of the term *Cartesianism* were foreshadowed in Descartes's views about corporeal substance.

OCCASIONALISM. In the *Principles* (II, 36), maintaining that God was the primary and universal cause of motion, Descartes explained that, when God created matter or extension, he created it with motion and rest; and Descartes implied that, but for God's imparting motion to matter, it would have been motionless and undifferentiated, and that motion and rest, and the resulting differentiation of matter, did not follow necessarily from its nature or essence. He further explained that, in conserving matter from moment to moment, God preserved the same quantity of motion that He originally introduced; and it seemed to follow that God's continuing to impart motion to matter was a necessary condition of the continued existence of motion and that bodies of themselves did not have the power of remaining in motion or of producing motion in other bodies. The conclusion toward which Descartes was drawn was that, although motion (*translatio*) was a characteristic or mode of bodies, the moving force of bodies was not in bodies themselves but in God. He did not, however, draw this conclusion. In a letter to Henry More, he noted that he was reluctant to discuss the question of the moving force (*vis movens*) of bodies in his published works, for fear that his view might be confused with that of God as *anima mundi*; and the view that he apparently wished to maintain was that, though the moving force of bodies was from God and in a sense was in God, it was also a characteristic or mode of bodies.

The occasionalists, taking the views that matter was inert and that the motion ascribed to bodies was simply change of position, did not hesitate before the conclusion that the force required to move bodies was not in bodies themselves but in the primary and universal Cause of motion, God. According to their conclusion, when a billiard ball that was in motion came in contact with a second ball that was at rest, there was no power or force in the first ball capable of moving the second, and the movement of the second ball required the action of God, who, on the occasion of impact, moved the second ball in accordance with rules that he had established for the motion of bodies. By virtue of the uniformity of God's action, the first ball could be called the cause—the particular or occasional cause—of the second ball's moving; but, without God's action, it was inefficacious, and the primary and universal cause of motion, that is, God, was the effectual cause of the second ball's moving. The occasionalists took it to be true a fortiori that bodies of themselves lacked the power of producing, as in sense perception, changes in the mind; and they offered a number of arguments to show that the mind in turn lacked the power, as in volition, of moving the body. The true cause of both sensations and voluntary movements was God, who instituted laws for the union of mind and body and acted accordingly in particular instances.

The originators of the occasionalist movement were Louis de La Forge and Géraud de Cordemoy.

La Forge. In the *Traité de l'esprit de l'homme* (Paris, 1666), La Forge represented himself as continuing work that Descartes had left unfinished in his *Treatise on Man* and undertook to explain and develop the notion of a mind or soul distinct from, yet united to, the body. Facing problems concerning the possibility of the body acting on the mind and vice versa, he noted that these problems were not isolated and that there was a related problem concerning the possibility of one body acting on another. In his discussion of these problems, La Forge did not deny that bodies acted on one another or on the mind, or that

the mind acted on the body; on the contrary, he insisted that God in his omnipotence could delegate the power of acting to created things. Yet, distinguishing two senses of "cause," he denied that created things were unambiguously the causes of the effects attributed to them and called them the "occasional" or "equivocal" causes.

Cordemoy. In *Le discernement du corps et de l'âme* (Paris, 1666), Cordemoy, unlike La Forge, was not concerned with presenting views necessarily in harmony with Descartes's, and he denied outright the action of bodies on one another or on the mind and the action of the human mind on the body. In his formally presented proof that God was the true cause of the movement of bodies, he made use of principles that Descartes would have accepted but drew conclusions from them that it would be safe to say would have greatly disturbed Descartes. Descartes had written of a motion in the brain as giving occasion (*donnera occasion*) to the soul to have a certain sensation or thought, and Cordemoy may have had these passages in mind in employing the expression *cause occasionelle* to refer to what, as in the case of a motion in the brain, might be thought to be the true cause of an event. But, unlike Descartes, he denied that the occasion or occasional cause was, strictly speaking, the cause of the event and maintained that the true cause was God.

Geulincx. Arnold Geulincx apparently developed his version of occasionalism independently of La Forge and Cordemoy. Illustrating the lack of causal relation between mind and body, he used the analogy of synchronized clocks, which was later taken up by Leibniz; and, to prove a lack of genuine causation, he made use of the principle that nothing can be done unless there is knowledge on the part of the putative agent or cause of how it is done.

Malebranche. Malebranche, the most celebrated of the occasionalists, was familiar with the work of Cordemoy and adapted, for his own purposes and with great originality, the theory of causation he found in Cordemoy. He added powerful arguments, extended the view to cover volitions not pertaining to bodily movements (such as the volition to form an idea), and presented it as an integral part of his theocentric vision of the universe.

MONISM AND PLURALISM. It has been argued that the dualisms and pluralism found in Descartes's statements about substance—of uncreated and created substance, corporeal and spiritual substance, and individual substances—contradicted his own definitions and principles and that Spinoza's doctrine of the unity of substance was the consistent and pure form of Cartesianism. It has also been maintained that Spinoza's monism and Leibniz' pluralism were the opposite poles to which philosophers accepting a notion of substance like that of Descartes were inescapably driven. Discussions of these views and of Spinoza's and Leibniz's metaphysics of substance is beyond the limits of this article, though it need hardly be added that the historical and logical relations of Descartes's assertions about substance and those of Spinoza and Leibniz have figured importantly in discussions of Cartesianism and that the essence of Cartesianism has sometimes been located in a common notion of, or presupposition about, substantiality.

It may be noted, however, that Descartes's assertions about corporeal substance also gave rise to conflicting theories among less renowned students of his metaphysics. On the one side, Geulincx, following Descartes's inclination to think of particular bodies as portions of a common stuff or substance, contended that "body itself" (*corpus ipsum*) was primary and substantial and that particular bodies were limitations or modes of corporeal substance. On the other side, Cordemoy, sharing Descartes's inclination to think of particular bodies as objects really distinct from one another, came to the unorthodox conclusion that body in general, or matter, was an aggregate and that the parts of which it was composed were indivisible extended substances, or atoms.

See also Arnauld, Antoine; Bayle, Pierre; Berkeley, George; Clauberg, Johannes; Copernicus, Nicolas; Cordemoy, Géraud de; Descartes, René; Desgabets, Robert; Dualism in the Philosophy of Mind; Fontenelle, Bernard Le Bovier de; Foucher, Simon; Geulincx, Arnold; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Monism and Pluralism; Newton, Isaac; Nicole, Pierre; Régis, Pierre-Sylvain; Regius, Henricus (Henry de Roy); Rohault, Jacques.

Bibliography

Balz, Albert G. A. *Cartesian Studies.* New York: Columbia University Press, 1951.

Belaval, Yvon. *Leibniz critique de Descartes.* Paris: Gallimard, 1960.

Berthé de Besaucèle, Louis. *Les cartésiens d'Italie.* Paris: A. Picard, 1920.

Bohatek, Josef. *Die cartesianische Scholastik in der Philosophie und reformierte Dogmatik des 17. Jahrhunderts.* Leipzig, 1912.

Bordas-Demoulin, Jean Baptiste. *Le cartésianisme,* 2nd ed. Paris, 1874.

Bouillier, Francisque. *Histoire de la philosophie cartésienne,* 3rd ed. Paris: C. Delagrave, 1868. A standard work, to be used with caution.

Cousin, Victor. *Fragments philosophiques,* 5th ed. Paris, 1866.

Damiron, Jean Philibert. *Essai sur l'histoire de la philosophie en France au XVIIe siècle.* Paris: Hachette, 1846.

Dibon, Paul. *La philosophie néerlandaise au siècle d'or.* Paris: Elsevier, 1954.

Dijksterhuis, E. J. et al. *Descartes et le cartésianisme hollandais.* Paris: Presses universitaires de France, 1950.

Lemaire, Paul. *Le cartésianisme chez les bénédictins; Dom Robert Desgabets, son système, son influence et son école.* Paris: Alcan, 1901.

Monchamp, Georges. *Histoire du cartésianisme en Belgique.* Brussels: F. Hayez, 1886.

Mouy, Paul. *Le développement de la physique cartésienne, 1646–1712.* Paris: J. Vrin, 1934. An invaluable account—intelligent, informed, and judicious—of Cartesianism in physics.

Prost, Joseph. *Essai sur l'atomisme et l'occasionalisme dans la philosophie cartésienne.* Paris, 1907.

Sortais, Gaston. *Le cartésianisme chez les jésuites français au XVIIe et au XVIIIe siècle.* Paris: G. Beauchesne, 1929.

Thijssen-Schoute, Caroline Louise. *Nederlands cartesianisme.* Amsterdam: Noord-Hollandsche, 1954.

Vartanian, Aram. *Diderot and Descartes.* Princeton, NJ: Princeton University Press, 1953.

Willis Doney (1967)

CARTESIANISM [ADDENDUM]

Apparently, it was the Cambridge Platonist Henry More who introduced the term *Cartesianism*—from the Latin *Cartesius*—into the English language. The term itself now denotes either the views of René Descartes or the various defenses and developments of these views in the writings of les cartésiens, an eclectic group of seventeenth- and eighteenth-century European intellectuals.

SCIENCE AND THEOLOGY

Descartes is perhaps best known in the early twenty-first century both for his epistemological "method of doubt" and for his metaphysical doctrine of mind-body dualism. However, he was known in the early modern period primarily for his attempt to systematically displace explanations of natural phenomena, deriving from the work of Aristotle, that were then predominant in both Catholic and Protestant schools on the Continent. In *Principles of Philosophy* (1644) Descartes proposed as an alternative for Aristotelian explanations in terms of prime matter, substantial forms, and final causality his own more austere explanations in terms of extension, its modifications, and purely mechanistic laws. There were other critics of the Aristotelianism of the schools, most notably Pierre Gassendi and the Gassendists. Nevertheless, Descartes's followers proved to be more adept than the Gassendists at packaging the new mechanistic science. Even so, it is understandable that Cartesian science is not as prominent today given the decisive refutation of Descartes's particular brand of physics in the work of More's greatest student, Isaac Newton.

Theological issues also dominated discussions of Descartes's system in earlier centuries in a way that they no longer do today. Such issues were of immediate practical concern to Descartes himself, who encountered fierce theological resistance not only in France but also in the United Provinces (now Holland), where he lived for most of his adult life. He failed in his attempt to infiltrate the Catholic universities in France at least partly because Aristotelian traditionalists saw his system as a threat to the Catholic dogma of the miraculous conversion in the Eucharist of the substance of bread and wine into the body and blood of Christ.

Descartes did fare somewhat better in the Calvinist United Provinces, where his writings received an audience in the academy during his residence there. Even in this region, however, orthodox Calvinists urged that his insistence on the real distinction between mind and body conflicts with the Aristotelian position that the soul bears a natural relation to a certain body in virtue of being its substantial form. These critics emphasized the threat that his system posed to Christian doctrines such as the resurrection of the body and the unity of the incarnated Christ. Moreover, before and after Descartes's death critics attempted to gain an advantage over Cartesianism by linking it to heterodox theological views. In the United Provinces the connection was typically to the doctrinally tolerant Dutch Remonstrant Calvinists, who deviated from Reformed Orthodoxy in insisting on one's freedom to accept or reject divine grace. After his death, however, Descartes was linked in France to a different group, the rigoristic French Jansenists, who set themselves in opposition to a Jesuit theology that emphasizes the dependence of one's salvation on the activity of one's undetermined free will. That the Jansenists were linked to Descartes bespeaks the influence of Antoine Arnauld, who was a prominent defender of both Cartesianism and Jansenism.

CARTESIANISM AND AUGUSTINIANISM

There was a strong inclination among French Cartesians to counter theological objections by invoking the authority of St. Augustine. There were roughly two general approaches, which were reflected in the distinction of the

scholar Henri Gouhier (1978) between "Cartesianism augustinized" and "Augustinianism cartesianized." The augustinized Cartesians, including Claude Clerselier, Descartes's literary executor, and the physician Louis de la Forge, were concerned to bolster Cartesian natural philosophy by stressing the ways in which Descartes's proofs of the existence of God and of the immateriality of mind complement Augustinian spiritualism. The defense of a cartesianized theology was pursued with disastrous consequences by the Benedictine Robert Desgabets, whose development of Descartes's account of the Eucharist provided the impetus for the official censorship of Cartesianism in France two decades after Descartes's death.

The cartesianized Augustinians tended to emphasize not Descartes's infrequent forays into theology, but his more common insistence that theological issues are outside of his jurisdiction insofar as their treatment requires recourse to revelation. This insistence allowed theologians such as Arnauld to appeal to Descartes to safeguard against Jesuit intrusion a "positive" or dogmatic theology devoted to providing a philosophical explication (or, for critics of the Jesuits, misrepresentation) of Augustinian views on matters of faith. Dutch Cartesians also attempted to insulate Cartesian philosophy from theology, though for them the concern was less to promote Augustinian purity in theology than to honor the distinction of the disciplines in the universities. This interest in making Descartes fit for the schools also explains the emphasis in the work of these Cartesians on the similarities between Aristotle and Descartes. It is this "scholasticized" Cartesianism that was exported from the United Provinces to Germany soon after Descartes's death by Dutch-trained Cartesians such as Johannes Clauberg.

MALEBRANCHE AND HIS CRITICS

The reception of Descartes was conditioned by the work of Nicolas Malebranche, a member of the Oratory in Paris. Malebranche attempted with other French Cartesians to link Descartes to Augustine. In Malebranche's case the result was a synthesis that stressed the dependence of creatures on God's rational activity. His system included the view, anticipated in the work of La Forge and others, that bodies serve as the noncausal occasion for God to distribute motion by means of the most economical laws. Malebranche further extended this sort of view to theology, arguing that God distributes grace in accord with simple general laws.

Malebranche's theological views upset Arnauld, his former ally, who took them to be an illustration of the dangers of philosophical incursions into theology. Nonetheless, the opening salvo in his protracted and increasingly bitter dispute with Malebranche was his critique of Malebranche's philosophical doctrine that "we see all things in God," that is, that one knows the bodies one sees through the idea of extension in God that represents them. Arnauld appealed to Descartes in defense of the alternative position that representative ideas are merely modes of one's soul. The French Cartesian Pierre-Sylvain Régis, who had earlier published a popularization of Cartesianism in his *System of Philosophy* (1690), defended Arnauld's account of ideas in a polemical exchange with Malebranche during the mid-1690s. Unlike Arnauld, however, but like Desgabets, whom he admired, Régis challenged Malebranche's claim that eternal essences that serve as the ground for eternal truths are identical to uncreated ideas in the divine reason. As Malebranche himself recognized, such a claim undermines Descartes's doctrine of the creation of the eternal truths. Régis and Desgabets were both concerned to defend this doctrine by claiming that eternal truths concerning creatures derive not from uncreated ideas in God, but from features of the world that God created with complete indifference.

See also Clauberg, Johannes; Descartes, René; Desgabets, Robert; Régis, Pierre-Sylvain; Regius (Henri de Roy); Rohault, Jacques.

Bibliography

Armogathe, Jean-Robert. *Theologia cartesiana: l'explication physique de l'Eucharistie chez Descartes et dom Desgabets*. The Hague: Nijhoff, 1977.

Clarke, Desmond M. *Occult Powers and Hypotheses: Cartesian Natural Philosophy under Louis XIV*. Oxford, U.K.: Clarendon Press, 1989.

Des Chene, Dennis. *Physiologia: Natural Philosophy in Late Aristotelian and Cartesian Thought*. Ithaca, NY: Cornell University Press, 1996.

Gaukroger, Stephen. *Descartes: An Intellectual Biography*. Oxford, U.K.: Clarendon Press, 1995.

Gouhier, Henri. *Cartésianisme et augustinisme au XVIIe siècle*. Paris: J. Vrin, 1978.

Lennon, Thomas M. *The Battle of the Gods and Giants: The Legacies of Descartes and Gassendi, 1655–1715*. Princeton, NJ: Princeton University Press, 1993.

Marion, Jean-Luc. *Sur la théologie blanche de Descartes: Analogie, création des vérités éternelles, et fondement*. Paris: Presses Universitaires de France, 1981.

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Verbeek, Theo. *Descartes and the Dutch: Early Reactions to Cartesian Philosophy, 1637–1650*. Carbondale: Southern Illinois University Press, 1992.

Tad M. Schmaltz (2005)

CARTWRIGHT, NANCY

(1944–)

Nancy Cartwright, as of 2005, held several academic positions, including professor of philosophy in the Department of Philosophy, Logic, and Scientific Method at the London School of Economics (since 1991); director of the LSE Centre for Philosophy of Natural and Social Science (since 1993); and professor of philosophy in the Department of Philosophy at the University of California at San Diego (since 1998). She had also served on the faculty at the University of Maryland (1971–1973) and Stanford University (1973–1991). She is the recipient of a MacArthur Fellowship and is a Fellow of the British Academy.

Cartwright first became widely known for the radical thesis, presented in her landmark 1983 collection of essays *How the Laws of Physics Lie*, that the fundamental laws of physics did not state truths about the world. The thesis is radical because philosophers have generally assumed that there is some set of underlying physical laws which, ultimately, describe all natural events. This is probably still a majority opinion among philosophers of science, though a much more controversial one than when Cartwright wrote these essays. At the same time she also proposed (along with Ian Hacking) a cautious realism about theoretical entities, which did not depend on people's ability to formulate true laws about them.

Cartwright's argument is based on a distinction between phenomenological and theoretical—or fundamental—laws. Phenomenological laws are, unsurprisingly, the laws that apply to actually observable phenomena. Their application is generally tightly circumscribed by detailed specification of the situations to which they apply. While fundamental laws may play an essential role in the formulation of phenomenological laws, the former are not themselves true. This is because they abstract from all the detailed *ceteris paribus* conditions that give phenomenological laws a chance, at least, of being true within their specific domains.

In her most recent book, *The Dappled World* (1999), Cartwright continues her attack on fundamentalism, the idea—from realism—that there is one unique set of laws applying to everything. The attack on fundamentalism, however, is now more uncompromising, as she has become increasingly skeptical about the usefulness of fundamental laws for deriving phenomenological laws. At the same time, a positive theme that she has developed throughout her career is increasingly emphasized: The conception of science not as searching for laws at all, but as constructing models. For models, the question of truth does not arise. They may more or less adequately represent parts of reality, and they may be more or less useful in providing understanding, explanation, and prediction.

Another theme more strongly emphasized in the later book is the disunity of science. Whereas a majority of philosophers of science accept a disunified science in the sense that laws in different domains are not reducible to laws of a more fundamental science, a majority of these philosophers see this as a consequence only of practical problems of complexity or the limited cognitive capacities of humans. Cartwright, on the other hand, is a leading advocate of a more radical position: that the autonomy of theories is indicative of what there is to know about the world. The world itself does not have a unitary underlying lawlike pattern. Its nomological structure is dappled.

The other related topic to which Cartwright has been among the most prominent contributors is the nature of causality. The decentering of fundamental laws from the vision of science naturally engenders skepticism about the Humean program of reducing causes to instances of laws. A project introduced in her first book—and developed in detail in her 1989 work, *Nature's Capacities and their Measurement*—is that an understanding of causality in terms of laws should be replaced with one in terms of capacities. In parallel with the emphasis on models, this move contributes to doubts as to whether laws are needed at all. The central thesis of this book is that science cannot be understood without assuming real capacities in the world. As is well known, Hume argued that positing capacities violated a proper empiricism. Cartwright, a committed empiricist, insists that capacities are as empirically accessible as laws and more specifically, that their measurement is a defining activity of science. In a further anti-Humean move, she argues that singular rather than generic causes are fundamental. A paradigm for Cartwright of causal knowledge is that aspirins have the capacity to cure headaches. Yet the canonical evidence for this claim is that on some specific occasions an aspirin actually does cure a headache.

This also connects to a central topic of her earliest work: probabilistic causality. This topic arises because capacities are to be thought of as being displayed only under specific circumstances, so that the relation between a capacity and its exercise is typically probabilistic. Conversely, Cartwright explores the question whether probabilistic relations can provide evidence for causes. Her answer is that they can, but only on the assumption that the effects are indications of real capacities in objects.

This entry has described some main themes from Cartwright's work in fairly abstract terms, but it should be

emphasized that she has been a leader of the move to focus philosophy of science on detailed examination of exemplary cases of scientific work. For the earlier part of her career most of this work was addressed to physics. From the late 1980s she increasingly switched her attention to examples from economics, and is now a leading figure in the philosophy of economics. Perhaps surprisingly to those who see the sciences as hierarchically arranged with physics secure at the top of the heap, Cartwright finds many themes in common to physics and economics. A central idea linking the two is her interest in machines, which can also be seen as concrete instantiations of models. A paradigm from her earlier work is the laser. The moral of this example is that the laser concretely embodies the *ceteris paribus* clauses emphasized in her critical discussion of fundamental laws by a range of actual mechanisms that ensure the proper conditions for the exercising of the crucial capacity—in this case the capacity for inversion in a population of atoms. Central to Cartwright's work on economics is the idea of a socioeconomic machine. As an example, she considers the mechanism by which a central bank increases the money supply. Like the laser, this does not reflect a law of nature, but a capacity of a certain kind of money, under properly controlled conditions, to have an important economic effect.

Cartwright claims as a philosophical hero Otto Neurath, a founding member of the Vienna Circle. Her admiration is of his commitment to seeing in science the capacity to change the world. A concern with the social impact of science and philosophy of science, while often beneath the surface, has been discernible in much of Cartwright's work.

See also Laws, Scientific; Scientific Realism.

Works by Cartwright

How the Laws of Physics Lie. Oxford: Oxford University Press, 1983.

Nature's Capacities and their Measurement. Oxford: Oxford University Press, 1989.

The Dappled World: A Study of the Boundaries of Science. Cambridge, U.K.: Cambridge University Press, 1999.

John Dupré (2005)

CARUS, CARL GUSTAV

(1789–1869)

Carl Gustav Carus, a German physician, biologist, and philosopher, was born in Leipzig and studied chemistry and then medicine at the University of Leipzig. In 1811 he became the first person to lecture there on comparative anatomy. Two years later he became director of the military hospital at Pfaffendorf and, in 1814, professor of medicine at the medical college of the University of Dresden, where he remained to the end of his life. He was appointed royal physician in 1827 and privy councilor in 1862.

Carus was widely known for his work in physiology, psychology, and philosophy, and was one of the first to do experimental work in comparative osteology, insect anatomy, and zootomy. He is also remembered as a landscape painter and art critic. He was influenced by Aristotle, Plato, Friedrich Wilhelm Joseph von Schelling, and Johann Wolfgang von Goethe, about whom Carus wrote several works, the most important of which is *Goethe dessen seine Bedeutung für unsere und die kommende Zeit* (Vienna, 1863). Carus's philosophical writings were more or less forgotten until the German philosopher and psychologist, Ludwig Klages, resurrected them.

Carus's philosophy was essentially Aristotelian in that it followed the unfolding or elaboration of an idea in experience from an unorganized multiplicity to an organized unity. This universal, unfolding unity or developing multiplicity within unity Carus called God. God, or the Divine, is not a being analogous to human intelligence; rather, it is the ground of being revealed through becoming, through the infinitely numerous and infinitely varying beings or organisms that come into being through the Divine in space and time.

Carus called his theory of a divine or creative force "entheism." The unknown Divine is revealed in nature through organization, structure, and organic unity. As the ground of being, it is outside space and time, unchanging, and eternal. As thought or insight, it is the God-idea of religion, found everywhere in life and the cosmos. As life, it is the sphere, the basic form taken by living cells and the heavenly stars. As matter, it is the ether exfoliating in infinitely varied things.

According to Carus, the body cannot be separated from the soul. Both are soul, but we speak of "body" when some unknown part of the soul affects the known part; and we speak of "soul" when the known part affects the unknown part.

Carus's metaphysics, and his important contribution to psychology, is a theory of movement from unconsciousness to consciousness and back again. Whatever understanding we can have of life and the human spirit hinges upon observation of how universal unconscious-

ness, the unknown Divine, becomes conscious. Universal unconsciousness is not teleological in itself; it achieves purpose only as it becomes conscious through conscious individuals. Consciousness is not more permanent than things; it is a moment between past and future. As a moment, it can maintain itself only through sleep or a return to the unknown.

See also Aristotle; Goethe, Johann Wolfgang von; Klages, Ludwig; Plato; Schelling, Friedrich Wilhelm Joseph von; Unconscious.

Bibliography

PRINCIPAL WORKS BY CARUS

Psyche: zur Entwicklungsgeschichte der Seele. Pforzheim, 1846; 3rd ed., Stuttgart, 1860.

Physis: zur Geschichte des leiblichen Lebens. Stuttgart: Scheitlin, 1851.

Symbolik der menschlichen Gestalt. Leipzig, 1853.

Natur und Idee. Vienna: W. Braumüller, 1861.

Lebenserinnerungen und Denkwürdigkeiten, 4 vols. Leipzig: F. A. Brockhaus, 1865–1866.

Vergleichende Psychologie. Vienna: W. Braumüller, 1866.

WORKS ON CARUS

Bernouilli, Christoph. *Die Psychologie von Carl Gustav Carus.* Jena, Germany: E. Diederichs, 1925.

Kern, Hans. *Carus: Personlichkeit und Werk.* Berlin, 1942.

Rubin Gotesky (1967)

CARUS, PAUL
(1852–1919)

Paul Carus, a philosopher and monist, was born at Ilsenburg, Germany, and died in La Salle, Illinois. After receiving his Ph.D. at Tübingen, in 1876, and completing his military service, he taught in Dresden. Censure of religious views he had expressed in pamphlets led him to leave Germany for England. He then went to New York, where in 1885 he published *Monism and Meliorism.* This book aroused the interest of a German chemist in La Salle, Illinois, Edward Carl Hegeler, who had started a periodical, the *Open Court.* He invited Carus to take over the editorship. In 1888 another and more technical journal, the *Monist,* was founded, and Carus became its editor. Carus also published a series of philosophical classics, edited by leading professors of philosophy, which are still widely used in classrooms. The Carus family operated the Open Court Publishing Company until 1996. Open Court publishes the volumes of the Carus Lectures, which are given at meetings of the American Philosophical Association. The *Monist* was revived in 1962 under the editorship of Eugene Freeman.

For the *Monist,* Carus chose articles on the history and philosophy of religion, archaeology, biblical criticism, and especially the philosophy of science, both philosophy for the scientifically minded and philosophy about the sciences. He invited contributions from France and Germany and arranged for their translation. Important articles by Bertrand Russell, Ernst Mach, David Hilbert, Jules Henri Poincaré, John Dewey, and Charles Sanders Peirce appeared in the *Monist.* Carus frequently published articles of his own in criticism of his contributors, but the debates seem not so much to have modified his own monistic philosophy as to have led him to explain in detail how it differed from other monisms, such as Ernst Haeckel's.

Monism, for Carus, was the doctrine that all the things that are—however varied, diverse, and independent of each other they may appear to be—are somehow one. What makes them one are certain eternal laws that reside in things and are discovered, not created, by the investigator. These laws of nature are asserted to be dependent on a single law, which Carus identified with God.

Carus viewed his metaphysics as a speculative generalization from the view of mathematics that he had learned from Hermann Grassmann, his teacher at the Stettin Gymnasium. Alfred North Whitehead, too, acknowledged the influence of Grassmann, in his *Universal Algebra.* Some of the similarities between the metaphysics of Carus and Whitehead may have resulted from this common influence.

Carus can be called a realist inasmuch as he rejected the notion that the laws of nature depend on the mind of the investigator. In this he found himself in opposition to the Kantians. Nor did he hold to a materialism. Rather, he insisted that every part of the world is both material (acting in accord with the laws of matter) and spiritual (acting in accord with the laws of mind). The characteristic of mind, or spirit, is the ability to mirror the world. Thus Carus was also a realist in his account of knowing. In ethics he held that the worth of any part of the world depends on the degree to which it knows—that is, mirrors—the whole. This is achieved through greater and greater knowledge of the laws of nature. Hence, devotion to knowledge is the way to greater goodness. Prayer is recommended as a means of changing the will of the man who prays so that he can mirror the one law in his actions.

See also Dewey, John; Hilbert, David; Mach, Ernst; Monism and Pluralism; Peirce, Charles Sanders; Poincaré, Jules Henri; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bibliography

Hay, William H. "Paul Carus: A Case-Study of Philosophy on the Frontier." *Journal of the History of Ideas* 17 (1956): 498–510.

Meyer, Donald Harvey. "Paul Carus and the Religion of Science." *American Quarterly* 14 (1962): 597–607.

Sheridan, James Francis. *Paul Carus: A Study of the Thought and Work of the Editor of the Open Court Publishing Company.* Ann Arbor, MI, 1957.

William H. Hay (1967)

CASO, ANTONIO

(1883–1946)

Antonio Caso, a Mexican philosopher and diplomat, was born in Mexico City in 1883 and died there in 1946. He was a professor of philosophy at the National University of Mexico, rector of that institution, lecturer at the Colegio Nacional, and ambassador to several South American nations. He wrote voluminously over a period of three decades and had great influence as a teacher. For his sources he turned especially to Henri Bergson but also to Immanuel Kant, Arthur Schopenhauer, and Edmund Husserl.

The metaphysics of Caso emphasizes process, freedom, life, and spirit. He conceived of reality as a fluent dynamism whose operations and forms are unified organically. The subject-predicate bias of traditional logic distorts reality by its apparatus of static terms related as in a closed machine. Modern science has more insight with its realization that even the physical world eludes a rigorous determinism. The individual particle has a factor of spontaneity; law is only statistical, applying to groups by virtue of the mutual compensation of individual irregularities. By the same token, living process has a unique character that cannot be reduced to the terms of physics and chemistry but stimulates and directs the material vehicle. A conscious living being discovers its own freedom in the simple act of willing a bodily movement: freedom coincides with causation from within. Consciousness is not passively derived from more primitive conditions by laws of association and evolution. On the contrary, the pure ego projects its own structures upon the data of raw feeling, thus supplying the objects of mature experience and the principles underlying those of association and evolution.

The ethics of Caso is concerned with two triads: that of things, individuals, and persons, and that of economy, disinterest, and love. Things are merely physical, are deficient in unity, are divisible, and are not subjects of value. Individuals are living beings that are indivisible but can be substituted for each other. The value of the merely biological is economy, found in egocentricity and utility and illustrated in nutrition, growth, reproduction, tool making, and death. Beyond individuals are persons, which add the character of spirit to life. Persons are capable of both disinterest and love. Disinterest suspends the mechanisms of selfishness and usefulness in the act of contemplation; love identifies the self with another in sympathy and service and is at its noblest in self-sacrifice. Persons are unique; they play a role as creators of values in society, and in them freedom is most advanced and responsible. Their interplay defines human culture, the enemies of which are individualism and totalitarianism; both are forms of egoism and of economic value. The error of totalitarian philosophy is to transfer the notion of the absolute from a universal principle of existence, where it is justified, to the state, where it does not exist. This philosophy has its source in Thomas Hobbes; it should not be imputed to G. W. F. Hegel, who placed art, religion, and philosophy above the state.

Caso's aesthetics begins with the concept of a surplus of energy, or vital excess, that is the basis of play, art, and the spirit of sacrifice. Art is distinguished from play and from the spirit of sacrifice by disinterest. In addition to the suspension of selfishness and usefulness, disinterest implies abstraction from questions of reality and goodness of the object contemplated. Disinterest preserves art from any possibility of immorality, which requires an interested attitude. It is associated with the intuitive nature of the aesthetic experience, since absorption in the object as an end favors appreciation of its full individuality. The nonconceptual nature of the experience is reconciled with the claim of universality, after the manner of Kant. The experience, however, does not terminate with an image within the mind. The conative tendency of psychic states leads to empathy, or projection of the state upon the outer world. Aesthetic empathy differs from the projection mentioned earlier in that it is emotional and concrete rather than logical and formal, and from that empathy and religious empathy in that it is disinterested. But natural objects do not readily satisfy the aesthetic need. Aesthetic empathy therefore leads to expression, or the creation of works of art, in which are consummated

the empathic tendency and disinterested intuition. In his account of intuition and expression, Caso claimed to follow Benedetto Croce, but he did not do so without wavering.

See also Bergson, Henri; Croce, Benedetto; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; Kant, Immanuel; Schopenhauer, Arthur; Sympathy and Empathy.

Bibliography

MAIN WORKS BY CASO

La filosofia de la intuición (The philosophy of intuition). Mexico City, 1914.

La existencia como economía, como desinterés y como caridad (Existence as economy, as disinterest, and as love). Mexico City, 1919.

Discursos á la nación mexicana (Discourses to the Mexican Nation). Mexico City: Porrua huos., 1922.

El concepto de la historia universal (The concept of universal history). Mexico City: Ediciones Botas, 1923.

Principios de estética (Principles of aesthetics). Mexico City: Publicaciones de la Sria. de Edvcacion, 1925.

La Persona humana y el estado totalitario (The human person and the totalitarian state). Mexico City: Universidad Nacional Autónoma, 1941.

WORKS ON CASO

Berndtson, Arthur. "Mexican Philosophy: The Aesthetics of Antonio Caso." *Journal of Aesthetics and Art Criticism* 9 (4) (June 1951): 323–329.

Romanell, Patrick. *Making of the Mexican Mind.* Lincoln: University of Nebraska Press, 1952. Ch. 3.

Arthur Berndtson (1967)

CASSIRER, ERNST

(1874–1945)

Ernst Cassirer, the German neo-Kantian philosopher, was born in Breslau, Silesia. He studied at the universities of Berlin, Leipzig, Heidelberg, and Marburg and taught first at Berlin. From 1919 to 1933 he was professor of philosophy at Hamburg University; and he served as rector from 1930 to 1933. Cassirer, who was Jewish, resigned his post in 1933 and left Germany. He taught at Oxford from 1933 to 1935, at Göteborg, Sweden from 1935 to 1941, and at Yale from 1941 to 1944. He died in New York City while a visiting professor at Columbia University.

Cassirer was both a prolific historian of philosophy and an original philosopher. His philosophy is in many important respects a development and modification of Immanuel Kant's critical philosophy, idealistic in outlook and transcendental in method. Like Kant, he holds that the objective world results from the application of a priori principles to a manifold that can be apprehended only as differentiated and ordered by them. His method is transcendental in the sense that he investigates not so much the objects of knowledge and belief as the manner in which these objects come to be known or are constituted in consciousness. His work has to some extent also been influenced by G. W. F. Hegel and, of his own contemporaries, by his teacher Hermann Cohen and by Edmund Husserl.

Cassirer differs from Kant mainly in holding that the principles by which the manifold of experience receives its structure are not static, but developing; and that their field of application is wider than Kant supposed. Kant, according to Cassirer, assumed that the science and mathematics of his day admitted of no philosophically relevant alternatives, and therefore he conceived the synthetic a priori principles of the understanding to be unchangeable. He could not foresee the development of non-Euclidean geometry, of the modern axiomatic method, of the theory of relativity, or of quantum mechanics. Also, in Kant's day many areas of human culture had not yet been subjected to scientific investigation: There existed in particular no developed science of language and no scientific treatment of religion and myth. The idea of the humanities or moral sciences (*Geisteswissenschaften*) arose only in the nineteenth century. Cassirer's professed aim was to extend Kant's static critique of reason, that is, his critique of the organizing principles of natural science and morality, into a dynamic critique of culture, that is, of the organizing principles of the human mind in all its aspects. This aim is apparent in all his works, especially in his magnum opus, *Die Philosophie der symbolischen Formen*.

THE NATURE OF SYMBOLIC REPRESENTATION

A fundamental problem for the Kantian philosophy had been to understand the conceptualization of experience, in particular the relation between concepts and that to which they apply. For Cassirer, conceptualization, that is, the apprehension of the manifold of experience as instantiating general notions or as perceptual matter exhibiting a conceptual structure, is merely a special case of what he calls "symbolization," "symbolic representation," or simply "representation." Symbolic representation, according to Cassirer, is the essential function of human consciousness and is cardinal to our understanding not only of the structure of science, but also of myth and religion, of language, of art, and of history. Man is a symbolizing animal.

Symbolization creates, and exhibits within our consciousness, connections between perceptual signs and their significance or meaning. It is the nature of symbolic representation in general to constitute, or bring into being, a totality that both transcends the perceptual sign and provides a context for it. The unity of sign and signified allows for distinction in thought, but not in fact—just as color and extension are separable in thought but not in fact. The given always shows itself as a totality, one part of which functions as a representative of the rest. This basic self-differentiation of every content of consciousness is given a more enduring structure by the use of artificial signs that, as it were, articulate the stream of consciousness and impose patterns on it. The artificial signs or symbols, like the Kantian concepts and categories, do not mirror an objective world, but are constitutive of it. Scientific symbols constitute, or bring about, only one kind of objective world—the world of science. Mythical pictures constitute the reality of myths and religion; the words of ordinary language constitute the reality of common sense.

To the three symbolic systems that articulate three types of reality under different "symbolic forms" there correspond three modes of the one function of symbolic representation. The first and most primitive of these modes Cassirer calls the "expression function" (*Ausdrucksfunktion*). In the world it constitutes, the primitive world of myth, the sign and its significance merge into each other. The difference between them exists, but is not consciously noted. The thunder by which a primitive god shows his anger is not merely an external sign that the god is angry. It is the god's anger. In the same way, in ordinary perception we often not merely associate a smile with a kind intention, but also perceive a kindly smile.

The second mode of symbolic representation is "intuition function" (*Anschauungsfunktion*), which by the use of ordinary natural languages constitutes the world of common sense. The intuition function differentiates our perceptual world into spatially and temporally related material objects or substances that become the bearers of properties, the more permanent properties being apprehended as distinctive of the various kinds of substance, the less permanent being apprehended as accidental. Aristotle's philosophy represents, according to Cassirer, a prescientific stage of thinking about objects, based on the predominance of symbolic representation in the mode of the intuition function.

The third mode of symbolic representation, the "conceptual function" (*reine Bedeutungsfunktion*) constitutes the world of science, which is a system of relations as opposed to a system of substances with attributes. The particular, in this mode, is not subsumed under a universal but rather under a principle of ordering, which relates particulars to each other in ordered structures that, Cassirer seems to hold, are always serial in nature. He finds the prototype of this kind of symbolization in the works of Richard Dedekind, Giuseppe Peano, Gottlob Frege, and their successors.

The transcendental inquiry into the nature and function of symbolic representation is supported by a wealth of illustrations taken from the history of philosophy, the natural sciences, general linguistics, anthropology, and the humanities. Symbolic representation as a fundamental and logically primitive function must be seen at work in order to be understood. The philosophical analysis of symbolic representation can hardly do more than point out that in any symbolic representation two moments, the symbol and the symbolized, are united into an essential unity yet stand in polar relationship to each other. It has been objected that this analysis, by identifying a unity with an opposition of two different moments, results in a contradiction. Cassirer's answer to this objection, and to accusations that his professedly Kantian position is really Hegelian, is that his philosophy is not intended as a logic or a metaphysics, but as a phenomenology of consciousness.

PHILOSOPHY OF CULTURE

The highly general character of Cassirer's analysis of symbolic representation gives flexibility to a philosophy of culture. It does not force the variety of the ever-changing contents and structure of culture into rigid and artificial molds. But the very generality of Cassirer's conception makes it, perhaps, too easy to fit it to any situation and comparably difficult to test. It also makes it difficult to place the conclusions of Cassirer's special investigations in order of importance. The order here followed is in the main that of the summary given at the end of his *Essay on Man*, itself a synopsis of his *Philosophie der symbolischen Formen*.

Cassirer holds that the polarity that he finds in the relation between symbol and significance or meaning continually expresses itself in two opposing tendencies, a tendency toward stabilization and a tendency toward the breaking up of permanent symbolic patterns. In myth and the primitive religions the conservative tendency is stronger. Mythological explanation explains patterns of the present in terms of origins in a remote past—a type of explanation still regarded in the Platonic dialogues as containing important elements of truth. The more

advanced religions exhibit the opposing evolutionary tendency at work. This is mainly the result of conceiving forces in nature as individuals and persons, and of the consequent emergence of the notion of morality as being rooted in personal responsibility.

In natural languages, through which the commonsense world of substances in public space and time is constituted, the conservative tendency shows itself in the rules to which a language must conform if communication is to be possible. The evolutionary tendency, which is equally essential, works through phonetic and semantic change. The psychology of the processes by which children acquire their language shows important similarities to the development of a language through succeeding generations in a community.

In the arts, the tendency toward new patterns, which has its source in the originality of the individual artist, predominates over the tendency to preserve a tradition. Yet traditional forms can never be entirely discarded, since this would imply the breakdown of communication, making art, which is a cultural and social phenomenon, impossible. The polarity in artistic creation is mirrored in the history of aesthetic theories. Theories of art as based on imitation and as based on inspiration have in one way or another continuously arisen in opposition to each other. Cassirer's own view of the nature of art is largely influenced by Kant's *Critique of Judgment*, in which the essence of artistic creation and aesthetic experience is held to lie in the interplay of the understanding, which imposes rules, and of the free imagination, which can never be completely subsumed under determinate concepts.

In science the stabilizing and objective tendency predominates over that toward change and subjective innovation. Cassirer's philosophy of science is recognizably Kantian, although Kant's absolute a priori is replaced in it by a relative a priori. Scientific theories contain, apart from empirical concepts and propositions, concepts that are a priori and propositions that are synthetic a priori with respect to a given theoretical system. This idea has proved both fruitful and influential and has been further developed by, among others, Arthur Pap, at one time a pupil of Cassirer. Relative a priori concepts and propositions are hardly distinguishable from the theoretical concepts and propositions admitted by logical positivist philosophers of science when it appeared that their original positions were not wholly tenable.

Cassirer regards language, art, religion, and science as aspects in a continuous development that although it is not predictable in advance, does show an organic unity. Every aspect expresses the fundamental function of symbolic representation in human consciousness and the power of man to build an "ideal" or symbolic world of his own, which is human culture. Cassirer's work depends to a very great extent on the illustrative power of his detailed analyses. For this reason it is difficult to do it justice in a brief survey, especially since philosophical disagreement with his critical idealism is quite compatible with a deep appreciation of his informed scholarship and his sensitive judgment as to what is and what is not important in the various symbolic and conceptual systems that he has investigated.

See also Aesthetic Experience; A Priori and A Posteriori; Cohen, Hermann; Frege, Gottlob; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; History and Historiography of Philosophy; Kant, Immanuel; Neo-Kantianism; Peano, Giuseppe.

Bibliography

SELECTED WORKS BY CASSIRER

Theoretical

Substanzbegriff und Funktionsbegriff. Berlin: Cassirer, 1910.

Zur Einsteinschen Relativitätstheorie. Berlin: Cassirer, 1921. These two works have been translated in one volume as *Substance and Function and Einstein's Theory of Relativity*. Chicago: Open Court, 1923; reprinted New York, 1953.

Philosophie der Symbolischen Formen. 3 vols. Berlin: Cassirer, 1923, 1925, 1929. *Index*, Berlin, 1931. Translated by Ralph Manheim as *Philosophy of Symbolic Forms*. 3 vols. New Haven, CT: Yale University Press, 1953, 1955, 1957.

Symbol, Myth, and Culture: Essays and Lectures of Ernst Cassirer, 1935–1945, edited by Donald Phillip Verene. New Haven, CT: Yale University Press, 1979.

Zur Logik der Kulturwissenschaften. Göteborg: Göteborgs Högskolas Årsskrift 47, 1942. Translated by S. G. Lofts as *The Logic of the Humanities*, New Haven, CT: Yale, 2001.

An Essay on Man. New Haven, CT: Yale University Press, 1944.

Historical

Leibniz's System in seinen wissenschaftlichen Grundlagen. Marburg: Elwert, 1902.

Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit. 3 vols. Berlin: Cassirer, 1906, 1907, 1920. Vol. 4 translated by W. H. Woglom and C. W. Hendel as *The Problem of Knowledge*. New Haven, CT: Yale University Press, 1950.

Freiheit und Form, Studien zur deutschen Geistesgeschichte. Berlin: Cassirer, 1916.

Kants Leben und Lehre. Vol. XI of *Immanuel Kant's Werke*, edited by Ernst Cassirer and Hermann Cohen. Berlin: Cassirer, 1918.

Individuum und Kosmos in der Philosophie der Renaissance. Leipzig: Teubner, 1927. Translated by Mario Domandi as *The Individual and the Cosmos in Renaissance Philosophy*. New York: Harper, 1964.

Die Platonische Renaissance in England und die Schule von Cambridge. Leipzig: Teubner, 1932. Translated by F. C. A. Koelln and James P. Pettegrove as *The Platonic Renaissance in England*. Austin: University of Texas Press, 1953.

Die Philosophie der Aufklärung. Tübingen: Mohr, 1932. Translated by F. C. A. Koelln and James P. Pettegrove as *The Philosophy of the Enlightenment*. Princeton, NJ: Princeton University Press, 1951.

WORKS ON CASSIRER

Cassirer, Toni. *Mein Leben mit Ernst Cassirer*. Hildesheim: Gerstenberg, 1981.

Itzkoff, S. W. *Ernst Cassirer: Scientific Knowledge and the Concept of Man*. Notre Dame, IN: University of Notre Dame Press, 1971.

Krois, J. *Cassirer: Symbolic Forms and History*. New Haven, CT: Yale University Press, 1987.

Lipton, D. *Ernst Cassirer: The Dilemmas of a Liberal Intellectual in Germany, 1914–1933*. Toronto: Toronto University Press, 1978.

Paetzold, H. *Ernst Cassirer—Von Marburg nach New York: eine philosophische Biographie*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1995.

Schilpp, P. A., ed. *The Philosophy of Ernst Cassirer*. Evanston, IL: Library of Living Philosophers, 1949. Critical studies and full bibliography to 1949. For bibliography to 1964, see H. J. Paton and Raymond Klibansky, eds., *Philosophy and History, Essays Presented to Ernst Cassirer* (new ed., New York, 1964).

S. Körner (1967)
Bibliography updated by Thomas Nenon (2005)

CASTRO, ISAAC OROBIO DE

See *Orobio de Castro, Isaac*

CATEGORICAL IMPERATIVE

Immanuel Kant (1724–1804) introduced the term "categorical imperative" to characterize the fundamental principle of morality as it presents itself to beings. The principle is categorical, or unconditional, because it is valid for all humans, indeed, for all rational beings, independently of any particular desires or aims they may have. It presents itself as an imperative precisely because human beings have desires and aims that can be incompatible with the unconditional demands of the principle of morality and thus those demands often present themselves as obligations and constraints. Hence the propositional content of the fundamental principle of morality is identical for all rational beings, but its coloration as an imperative is distinctively human. For Kant, since there is a single fundamental principle of morality, there is, properly speaking, only a single categorical imperative, although more specific moral duties and obligations derivable from it are themselves unconditionally valid for any agent in the situation in which they arise. Kant contrasts the categorical imperative with "hypothetical imperatives," which express only the necessity of adopting certain means to achieve certain ends that are themselves merely conditional. Hypothetical imperatives can also present themselves to us as constraints, because we are not always sufficiently rational even to accept willingly the means to ends that we have willingly adopted, but in the case of hypothetical imperatives, we are not under any moral constraint to adopt the ends concerned.

Kant anticipated his mature distinction between categorical and hypothetical imperatives in his *Inquiry concerning the Distinctness of the Principles of Natural Theology and Morality* of 1764. There he wrote, "Every *ought* expresses a necessity of the action and is capable of two meanings. … Either I ought to do something (as a *means*) if I want something else (as an *end*), or I *ought immediately* to do something else (as an *end*) and make it actual." He argued that the former do not really express obligations at all; rather, they express only "recommendations to adopt a suitable procedure, if one wish[es] to attain a given end." Genuine obligations, by contrast, are "subordinated to an end which is necessary in itself." Kant's examples of ends that might be necessary in themselves were advancing the greatest total perfection and acting in accord with the will of God (Kant 1764; in Kant 1900, 2: 298; in Kant 1992, p. 272). The first of these is the ultimate end of morality according to Christian Wolff (1679–1754) and Alexander Baumgarten (1714–1762), and the latter the ultimate end of morality according to their Pietist opponent Christian August Crusius (1715–1775). In his *Anweisung, vernünftig zu leben* (Guide to living rationally; 1744/1964), Crusius himself anticipated the distinction that Kant made in the *Inquiry* by contrasting duties of prudence, which are grounded "only in certain ends already desired by us," with true obligations, which are grounded in "moral necessity" lying "in a law and in our owing fulfillment of it," and ultimately, in the case of "the obligation of virtue, or true obligation in a narrower sense," in divine law (§161). A widespread account of Kant's development of his mature conception of the categorical imperative is that he moved from the idea of an unconditional obligation grounded in a necessary end to the idea of an unconditional obligation that does not depend on any end whatever. Below, that will turn out to be misleading.

Kant first published his mature account of the categorical imperative in his *Groundwork for the Metaphysics of Morals* (1785). There Kant distinguished the categorical imperative from two kinds of hypothetical imperatives, namely, hypothetical imperatives of skill, which simply prescribe practically necessary means to realize entirely optional ends, and the hypothetical imperative of prudence, which prescribes means to an end that all human beings have as a matter of fact, namely happiness. Kant described the imperatives of skill as "problematic" (debatable, since the ends are optional) and the imperative of prudence as "assertoric" (impelled by the goal of happiness). Because the end of happiness is universal but not obligatory and because it is also indeterminate what will actually make anyone happy, the imperative of prudence can give rise only to "counsels of prudence." Finally, Kant stated, "There is one imperative that, without being based upon and having as its condition any other purpose to be attained by certain conduct, commands this conduct immediately. … It has to do not with the matter of the action and what is to result from it, but with the form and principle from which the action itself follows" (Kant 1785; in Kant 1900, 4:415–416; in Kant 1996, pp. 68–69). This is the categorical imperative, which is apodictic (certain).

In the *Groundwork*, Kant gave his first official formulation of the categorical imperative and the one to which he most frequently refers in subsequent works. This is that one "must act only in accordance with that maxim through which you can at the same time will that it become a universal law" (Kant 1785; in Kant 1990, 4:421; in Kant 1996, p. 73). He reached this formulation by different routes in the first and second sections of the book. In the first section, he began with the claim that only a good will is of unconditional value, and then argued that a good will is demonstrated in acting from the motive of duty, where "duty is the necessity of an action from respect for law" (Kant 1785; in Kant 1900, 4:400; in Kant 1996, p. 55), rather than in acting from any inclination toward a particular end or object. The good will having thereby been deprived of any inclination to realize it with action, nothing is left as its principle "but the conformity of actions as such with universal law" (Kant 1785; in Kant 1900, 4:402; in Kant 1996, p. 56).

In the second section, Kant argued that the formula of universal law follows from the very concept of the categorical imperative, since once it is stipulated that such an imperative "contains no condition to which it would be limited, nothing is left with which the maxim of action is to conform but the universality of a law as such" (Kant 1785; in Kant 1900, 4:420–421; in Kant 1996, p. 73). In his *Critique of Practical Reason* of 1788 (1996), Kant derives a similar formulation of the categorical imperative from the initial premises that any practical law must be necessary, but that any objective for action is empirical and contingent—a circumstance that leaves only the form of a law to furnish content for the categorical imperative (theorem III, Kant 1788, in Kant 1900, 5:27; in Kant 1996, p. 160).

In the *Groundwork*, Kant offers four further formulations of the categorical imperative. The first of these is "*Act as if your maxim were to become by your will a universal law of nature*" (Kant 1785; in Kant 1900, 4:421; in Kant 1996, p. 73), where a maxim is a proposal to perform a certain type of action for a certain end. H. J. Paton (1947) held that this introduces a teleological conception of nature into Kant's argument, and this is true in Kant's first illustration of how the imperative yields a prohibition of suicide. But since all that Kant explicitly meant by a law of nature is a law that is uniformly followed, this formulation, like the initial one, requires only that you consider whether you could act on your proposed maxim if in fact everyone else were also to act on it. In the second formulation, Kant said that "a possible categorical imperative" needs a ground in "something *the existence of which in itself* has an absolute worth, something which *as an end in itself* could be a ground of determinate laws," and stated that this ground is "the human being and in general every rational being" (Kant 1785; in Kant 1900, 4:428; in Kant 1997, p. 78). This leads Kant to reformulate the imperative as follows: "*So act that you use humanity, whether in your own person or in the person of any other, always as an end and never merely as a means*" (Kant 1785; in Kant 1900, 4:429; in Kant 1996, p. 80). By "humanity" Kant meant just the capacity to set and pursue ends (Kant 1785; in Kant 1900, 4:437; in Kant 1996, p. 86; Kant 1797; in Kant 1900, 6:387, 392; in Kant 1996, pp. 518, 522), so this requirement means that the human capacity to set and pursue ends should itself always be an end and never merely a means. Kant interpreted this requirement in turn to mean that the categorical imperative requires that you act only for ends that others can accept or even adopt for themselves. Third, Kant reformulated the imperative as "the *principle* of a human will *as a will giving universal laws through all its maxims*" (Kant 1785; in Kant 1900, 4:432; in Kant 1996, p. 82), which requires that any maxim be part of a universally acceptable *system* of maxims. Finally, he formulated the imperative as the requirement that "all maxims from one's own lawgiving are to harmonize into a possible kingdom of ends" (Kant 1785; in Kant 1900, 4:436; in Kant 1996, p. 86), which is "a

whole both of rational beings as ends in themselves and of the ends of his own that each may set himself" (Kant 1785; in Kant 1900, 4:433; in Kant 1996, p. 83).

This formulation makes explicit that to treat everyone as an end involves not only acting only on universally acceptable maxims but also allowing and promoting the individual ends of each insofar as doing so is consistent with treating all as ends in themselves. This sequence of formulations thus shows that the normative force of the categorical imperative is grounded on recognition of a necessary end, thus that the distance between Kant's mature formulation and his initial formulation of twenty years earlier is not as great as it initially seems, and that far from proscribing actions in behalf of particular, contingent ends, the categorical imperative prescribes such actions to the extent that such ends are freely chosen and are consistent with universal law. This is the foundation for Kant's doctrine of duties.

G. W. F. Hegel famously charged that Kant's categorical imperative is an "empty formalism," that is, that it either presupposes some already accepted particular end or else licenses any action that anyone is willing to universalize. This is clearly false, since the imperative requires consistency between any maxim on which you are proposing to act and the universalization of that maxim. Moreover, as the analysis above shows, universalization includes the requirement that your maxim be universally acceptable. This means that it is not enough that *you* be willing for your maxim to be universalized; everyone must be willing. More recent authors, including Marcus G. Singer (1971), Onora O'Neill (1975, 1989), and Allen Wood (1999), have considered cases in which clearly permissible maxims seem to fail the test of universalizability while clearly impermissible maxims seem to pass it. This shows that considerable care is needed in properly formulating maxims to be tested by the categorical imperative. John Rawls (2000) has interpreted the categorical imperative as yielding a "CI-procedure," which can be directly applied to individual maxims or proposals of action, while Barbara Herman (1993) has argued that it rather yields "rules of moral salience," that is, general factors of moral relevance that need to be considered in undertaking any particular action. The latter seems closer to Kant's own use of "categorical imperative" in the *Metaphysics of Morals* (1797) to yield general categories of duty, although Kant himself sometimes interpreted the requirement of being universalizable to apply to very specific types of action, as in his notorious argument of 1798 that lying is always wrong, no matter what the circumstances.

In addition to these questions about the interpretation and application of the categorical imperative, it has been criticized from a number of other points of view. Philippa Foot (1972/1978) has argued that categorical form is not sufficient to show that a requirement is moral, since rules of etiquette are also stated in categorical form. She concluded that both etiquette and morality, in spite of their categorical form, are really systems of hypothetical imperatives, to be adopted only if one wants to be regarded as polite or moral respectively. Bernard Williams (1985) accepted the categorical imperative as formulating the demands of morality, but raised questions about whether these demands are "overriding," that is, whether one's own personal projects and goals must always be sacrificed to the demands of morality in cases of conflict between them. R. M. Hare (1971) likewise accepted that moral principles have the form of categorical imperatives, or universal prescriptions, while raising the question of whether such prescriptions must always be accepted. These latter objections suggest that Kant was correct to use the concept of the categorical imperative to characterize the demands of morality, but that there is room to debate both whether he correctly identified the ground of any possible categorical imperative and whether morality itself is overriding.

See also Deontological Ethics; Duty; Kant, Immanuel; Kantian Ethics; Moral Rules and Principles.

Bibliography

PRIMARY SOURCES

Crusius, Christian August. *Anweisung, vernünftig zu leben* (1744). In *Die philosophische Hauptschriften*, edited by Giorgio Tonelli. Hildesheim: Olms, 1964. Selection translated in *Moral Philosophy from Montaigne to Kant*, edited by J. B. Schneewind. Vol. 2, 568–585. Cambridge, U.K.: Cambridge University Press, 1990.

Hegel, Georg Wilhelm Friedrich. *Elements of the Philosophy of Right* (1821). Translated by H. B. Nisbet. Cambridge, U.K.: Cambridge University Press, 1991.

Hegel, Georg Wilhelm Friedrich. *The Scientific Ways of Treating Natural Law, Its Place in Moral Philosophy, and Its Relation to the Positive Sciences of Law* (1802). Translated b y T. M. Knox. Philadelphia: University of Pennsylvania Press, 1975.

Kant, Immanuel. *Critique of Practical Reason* (1788). In his *Practical Philosophy*, edited and translated by Mary J. Gregor, 137–271. Cambridge, U.K.: Cambridge University Press, 1996.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals* (1785). In his *Practical Philosophy*, edited and translated by Mary J. Gregor, 37–108. Cambridge, U.K.: Cambridge University Press, 1996.

Kant, Immanuel. *Inquiry concerning the Distinctness of the Principles of Rational Theology and Morality* (1764). In his

Theoretical Philosophy, 1755–1770, edited and translated by David Walford, in collaboration with Ralf Meerbote, 243–275. Cambridge, U.K.: Cambridge University Press, 1992.

Kant, Immanuel. *Kants Gesammelte Schriften*, edited by the Königlich Preussischen (later Deutschen und Berlin-Brandenburgischen Akademie der Wissenschaften. 29 vols. Berlin: George Reimer (later Walter de Gruyter and Company), 1900–.

Kant, Immanuel. *Metaphysics of Morals* (1797). In his *Practical Philosophy*, edited and translated by Mary J. Gregor, 353–603. Cambridge, U.K.: Cambridge University Press, 1996.

Kant, Immanuel. "On a Supposed Right to Lie from Philanthropy" (1798). In his *Practical Philosophy*, edited and translated by Mary J. Gregor, 605–615. Cambridge, U.K.: Cambridge University Press, 1996.

SECONDARY SOURCES

Aune, Bruce. *Kant's Theory of Morals*. Princeton, NJ: Princeton University Press, 1979.

Ebbinghaus, Julius. "Interpretation and Misinterpretation of the Categorical Imperative." *Philosophical Quarterly* 4 (1954): 97–108.

Foot, Philippa. "Morality as a System of Hypothetical Imperatives." *Philosophical Review* 81 (1972): 305–316. Reprinted in her *Virtues and Vices*, 157–173. Berkeley: University of California Press, 1978.

Guyer, Paul. "The Possibility of the Categorical Imperative." *Philosophical Review* 104 (1995): 353–385. Reprinted in his *Kant on Freedom, Law, and Happiness*, 172–206. Cambridge, U.K.: Cambridge University Press, 2000.

Hare, R. M. *Practical Inferences*. Berkeley: University of California Press, 1971.

Herman, Barbara. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Höffe, Otfried. *Categorical Principles of Law: A Counterpoint to Modernity*. Translated by Mark Migotti. University Park: Pennsylvania State University Press, 2002.

Kerstein, Samuel J. *Kant's Search for the Supreme Principle of Morality*. Cambridge, U.K.: Cambridge University Press, 2002.

McDowell, John. "Are Moral Requirements Categorical Imperatives?" *Proceedings of the Aristotelian Society*, supp. vol. 52 (1978): 13–29. Reprinted in his *Mind, Value, and Reality*, 77–94. Cambridge, MA: Harvard University Press, 1998.

Nisters, Thomas. *Kants kategorischer Imperativ als Leitfaden humaner Praxis*. Freiburg, Germany: Verlag Karl Alber, 1989.

O'Neill, Onora. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge, U.K.: Cambridge University Press, 1989.

[O'Neill], Onora Nell. *Acting on Principle: An Essay on Kantian Ethics*. New York: Columbia University Press, 1975.

Paton, H. J. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. London: Hutchinson, 1947.

Patzig, Günter. *Ethik ohne Metaphysik*. Göttingen, Germany: Vandenhoeck und Ruprecht, 1971.

Pogge, Thomas W. "The Categorical Imperative." In *Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*, edited by Otfried Höffe, 172–193. Frankfurt am Main, Germany: Vittorio Klostermann, 1989. Reprinted in *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, edited by Paul Guyer, 189–214. Lanham, MD: Rowman and Littlefield, 1998.

Rawls, John. *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press, 2000.

Schnoor, Christian. *Kants kategorischer Imperativ als Kriterium der Richtigkeit des Handelns*. Tübingen, Germany: J. C. B. Mohr, 1989.

Schwaiger, Clemens. *Kategorische und andere Imperative: Zur Entwicklung von Kants praktischer Philosophie bis 1785*. Stuttgart, Germany: Fromann-Holzboog, 1999.

Schwemmer, Oswald. *Philosophie der Praxis: Verusch zur Grundlegung einer Lehre vom moralischen Argumentieren in Verbindung mit einer Interpretation der praktischen Philosophie Kants*. 2nd ed. Frankfurt am Main, Germany: Suhrkamp, 1980.

Singer, Marcus George. *Generalization in Ethics: An Essay in the Logic of Ethics, with the Rudiments of a System of Moral Philosophy*. New York: Atheneum, 1971.

Stratton-Lake, Philip. "Formulating Categorical Imperatives." *Kant-Studien* 83 (1993): 317–340.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Williams, T. C. *The Concept of the Categorical Imperative: A Study of the Place of the Categorical Imperative in Kant's Ethical Theory*. Oxford, U.K.: Clarendon Press, 1968.

Wood, Allen W. *Kant's Ethical Thought*. Cambridge, U.K.: Cambridge University Press, 1999.

Paul Guyer (2005)

CATEGORIES

Philosophical categories are classes, genera, or types supposed to mark necessary divisions within our conceptual scheme, divisions that we must recognize if we are to make literal sense in our discourse about the world. To say that two entities belong to different categories is to say that they have literally nothing in common, that we cannot apply the same descriptive terms to both unless we speak metaphorically or equivocally.

ARISTOTELIAN THEORY

The word *category* was first used as a technical term in philosophy by Aristotle. In his short treatise called *Categories,* he held that every uncombined expression signifies (denotes, refers to) one or more things falling in at least one of the following ten classes: substance, quantity, quality, relation, place, time, posture, state, action, and passion. By "uncombined expression" Aristotle meant an expression considered apart from its combination with other expressions in a sentence, and he intended his account to apply only to those expressions we now call

"descriptive" and "nonlogical." Logical expressions, such as "not," "or," "some," and "every," are excluded; these were called by medieval philosophers "syncategorematic," to distinguish them from the categorematic expressions covered by Aristotle's account of categories.

Each of the ten classes of entities signified constitutes a category, or genus, of entities, and each categorematic expression is said to be an expression in the category constituted by the class of entities it signifies. The nouns "plant" and "animal," for example, signify kinds of substances and are said to be expressions in the category of substance; the nouns "color" and "justice" signify kinds of qualities and are said to be expressions in the category of quality. On the other hand, the adjectives "colored" and "just" signify, respectively, colored and just things (substances) and also connote (consignify) the qualities color and justice. Aristotle labeled such expressions "derivative terms" or "paronyms" and held that instead of signifying substances simply, as expressions in the category of substance do, they signify substances derivatively by connoting accidents of substances.

Although Aristotle implied that his ten categories constitute the ten highest genera of entities and hence the only true genera—the only genera that cannot be taken as species of higher genera—he also implied that it is not essential to his theory that the categories be exactly ten in number or even that they be mutually exclusive and exhaustive. Categories are listed in various of Aristotle's writings, but the list usually stops short of ten without indication that categories have been omitted. He explicitly stated that no absurdity would result if the same items were included in both the category of quality and that of relation. He remarked that the expressions "rare," "dense," "rough," and "smooth" do not signify qualities, since they apply to a substance with reference to a quality it possesses, yet he did not specify in which category or categories these expressions are included. Despite these indications that his theory of categories is not entirely complete, medieval philosophers generally wrote as though Aristotle's list of ten provided a final, exhaustive enumeration of the highest genera of being.

What is essential to Aristotle's theory of categories is that substances be properly distinguished from accidents and essential predication from accidental predication. Any entity, regardless of the category in which it is included, can be an entity referred to by the subject term of an essential predication. "Man is an animal." "Red is a color." "Four is a number." "A year is twelve months." The subject terms denote entities that fall, respectively, in the categories of substance, quality, quantity, and time, and the predication in each case is essential. On the other hand, only entities in the category of substance can be entities referred to by subject terms of accidental predication. There is no such thing as an accident of an accident; accidents happen to substances and not to other accidents. "Red is darker than orange" does not assert something that happens to be, but need not be, true of red; it asserts what is essentially true of red, something that red must always be if it is to remain the color red. "Red is John's favorite color" does not assert anything that may happen to be true of red; rather, it asserts something that may happen to be true of John. To undergo change through time while remaining numerically one and the same thing is what principally distinguishes substances from entities in other categories. If John ceases to regard red as his favorite color, we say not that red has changed while remaining the same color but that John has changed while remaining the same person.

Categorematic expressions, for Aristotle, are technically "predicates," but they are not "predicates" in a sense that keeps them from serving as subject terms in essential predication. The minor term of an Aristotelian "scientific syllogism" occurs only as a subject, though Aristotle gave no examples in which it is a proper name. He regarded the ultimate subject terms in demonstration as common names marking species that are not further divided. Such expressions are still "predicates" in that like more generic terms they are applied to individuals in answer to the question What is it? But proper names are in a class by themselves; they are applied only in answer to the question Who? or Which? and are not "predicates" at all. Yet if proper names are thus not categorematic expressions, they are still fundamental to Aristotle's theory of categories. Without proper names there are no names for the subjects of accidental as distinct from essential predication. Man as such is an animal—"man" names every person indifferently if it names any, and the question of naming which one (or ones) does not arise. But only some man (or men) is (are) snub-nosed, and until the question Which? is answered by a proper name the subject of the accidental predication remains unnamed.

CATEGORY-MISTAKES. If we ask what, according to Aristotle's theory, would be the sort of thing often called today a "category-mistake," we must distinguish a mistake that violates what is essential to the theory from a mistake that violates a particular category-difference marked by the theory. Only a mistake of the first kind is strictly a category-mistake. Mistakes of the second kind form a subclass of equivocations. In his *Topics* (107a3–17), Aristotle listed as one example of equivocation the sentence

"The musical note and knife are sharp." That "sharp" is here used equivocally is shown by the fact that a musical note and a knife belong to different categories. A musical note is a kind of sound, and sounds are qualities. (Aristotle argued in *On the Soul,* 420a25–28, that we speak of the sound of a body as we speak of the color of a body.) A knife is a kind of substance, and one who believes that "sharp" applies in the same sense to musical notes and to knives may be said to have made the category-mistake of confusing a quality and a substance. Yet an appeal to category-differences is not necessary to expose the equivocation, and many equivocations cannot be exposed in this way because there is no violation of a category-difference. Aristotle claimed that the equivocal use of "sharp" in the example is also exposed by the fact (among others) that musical notes and knives are not compared with respect to their sharpness. Two notes may be equally sharp, or two knives, but not a note and a knife. Again, two flavors are equally sharp, but not a flavor and a note or a flavor and a knife. The equivocation in "The flavor and note are sharp" is exposed, although since flavors and sounds are both qualities there is no violation of a category-difference.

The appearance of absurdity produced by an equivocation can always be removed and literal meaning restored by distinction between the different senses of the crucial words. But with a genuine category-mistake there is no literal meaning to restore. In a passage in his *Posterior Analytics* (83a30–33), where he was discussing features of essential and accidental predication, Aristotle remarked that Plato's forms can be dismissed as mere sound without sense. The point is illustrated by a sentence like "The color white is white." The sentence may seem to make sense if one claims that since the color white is the standard by which we judge things to be white, it is itself white. But the sense is only apparent, because whatever is white remains numerically one and the same object even if its color changes. Such an object cannot be the quality, that is, the color white itself, as we then have the absurdity that the color white changes its color. Plato's theory of forms, as Aristotle interpreted it, makes the mistake of confusing accidental with essential predication. "The color white is the color white" is not an accidental but a trivially true essential predication; it is clearly not what is intended by the Platonic assertion that the color white is white. But the latter is just as absurd as the assertion that sitting sits.

Except in the passage in the *Posterior Analytics,* Aristotle did not refer to Plato's forms as mere sound without sense. Plato's theory has certain affinities with Aristotle's metaphysical account of substance as a composite of form and matter, and in his *Metaphysics,* Aristotle criticized Plato's forms, not as sound without sense, but as entities that fail to do the job they should, since they cannot be formal causes (991a11; 1033b26) and lead to an infinite regress (the third-man argument: 990b17). His criticism of the theory of forms receives attention in the history of philosophy mainly in this context of form, matter, and substance, and the passage in the *Posterior Analytics* that dismisses the forms as sound without sense is generally passed over or dismissed as a result of more than usual hostility toward Platonists. Yet apart from hostility, Aristotle was required by his theory to regard a sentence like "The color white is white" strictly as a category-mistake.

KANTIAN THEORY

Aristotle's theory dominated discussion of categories until the work of Immanuel Kant, where we find a radically new conception of a category. Kant professed in his theory of categories to have achieved what Aristotle had tried but failed to achieve in such a theory. Instead of beginning with uncombined expressions, Aristotle should have started with expressions of statements or judgments. Every statement is universal, particular, or singular in quantity; affirmative, negative, or infinite in quality; categorical, hypothetical, or disjunctive in the relation of its parts; and problematic, assertoric, or apodictic in modality (*Critique of Pure Reason,* "Transcendental Analytic," I, 2–3). Each of these twelve ways in which judgments are classified in logic corresponds to a function of the understanding indispensable to the formation of judgments, and each such function yields a category, or pure concept of the understanding, in one of the four major divisions of categories: quantity, quality, relation, and modality. The function, for example, of relating subject to predicate in a categorical judgment yields the relational category of substance and accident, and the function of relating antecedent to consequent in a hypothetical judgment yields the relational category of cause and effect.

Kant's conception of substance leads to important departures from Aristotle in the treatment of common names and paronyms. Whether an expression serves as a common name or as a paronym depends on its function in a given statement and not on its signification as an uncombined expression. "Stone," for example, serves as a common name of the substance in which a change occurs in "The stone grows warm," but it serves to specify a kind of change that occurs in a substance in "The sand

becomes stone." In the second case "stone" serves as a paronym; it connotes certain properties, such as hardness and solidity, and denotes any substance, such as a certain amount of sand, that acquires these properties. For Aristotle the change from sand to stone is substantial change, or coming to be, rather than alteration; for Kant substantial change is impossible because substance is related to accident as that which undergoes alteration is related to that which becomes and ceases to be. A substance is altered when one of its accidents ceases to be and is followed by another accident, so accidents, not substances, become and cease to be.

With Kant's theory there are no ordinary equivocations that can be exposed as category-mistakes, since categories are pure (formal), as opposed to empirical, concepts. "Substance" and "quality," in Aristotle's theory, are the highest generic terms that apply, respectively, to knives and sounds, so the equivocation in "The knife and musical note are sharp" can be exposed as a confusion of a substance and a quality. In Kant's theory, by contrast, generic terms represent empirical concepts, and an equivocation that confuses genera, as "The knife and musical note are sharp" confuses bodies and sounds, is not a category-mistake but a confusion of empirical concepts. One makes a category-mistake—violates what is essential to Kant's theory—by misapplying a category rather than by mistaking the category in which an entity belongs. The important point is that Kant's categories apply only to phenomena or appearances, not to entities or things in themselves. Every appearance can be judged according to every category and cannot be said to belong properly in one category rather than another. An appearance of red, for example, has extensive magnitude equal to a spatial area and is hence a quantity; it has intensive magnitude as a sensation with a certain degree of intensity and is hence a quality; it is related to further appearances as accident is to substance and effect to cause; and in relation to other appearances it is possible, actual, or necessary.

In Aristotle's theory, on the contrary, a redness is properly an accident in the category of quality; it exists in a substance from which it may be separated in thought but not in being. The extensive magnitude comprising a spatial area is a quantity of the substance and not of the redness; the intensity of the sensation of redness is a quality of the perceiving subject. Questions concerning the cause or the possibility, actuality, and necessity of the redness can be answered only by references to the substance that is said to be red. When the color is separated in thought from the substance the resulting abstract entity, the color red, can be characterized essentially (red, for example, is darker than orange), but to take it as an entity that itself has accidents is to make the category-mistake of confusing a quality with a substance.

To say that the color red is red is, for Kant, to misapply the relational category of substance and accident. Categories can be applied correctly only to phenomena, and in the case of a relational category both terms of the relation must be phenomena. The phrase "the color red" stands for the concept under which appearances of red are subsumed and not for an appearance that may be related to an appearance of red as substance to accident. This sort of category-mistake needs little attention since with Kant's theory there is no compelling tendency of the human mind to confuse a concept with its instances. But there is a natural tendency to make the mistake of applying categories to what are technically, for Kant, ideas and ideals; the former give rise to antinomies of pure reason and the latter to fallacious proofs of God's existence. Platonism in the form that gains a hold on men's minds is the mistake of applying the category of existence to ideals, not the mistake of confusing a concept with its instances. Along with antinomies and fallacious proofs of God, Kant argued for a third kind of category-mistake, a mistake that occurs when categories are misapplied in judgments about a thinking substance; the result is a set of equivocations giving rise to what Kant called "paralogisms of pure reason." These three kinds of category-mistakes are to be exposed not as sound without sense but as illusions to which the human mind is naturally prone.

POST-KANTIAN THEORIES

Although Kant's theory of categories marks the single most important development in the subject since Aristotle, his list of twelve categories never acquired anything like the dominant role once held by Aristotle's list of ten. Kant's influence has been to change the conception of how a list of categories should be formed, rather than to provide the list itself. Instead of looking for the highest genera of being, the most universal kinds of entities, one should look for the most universal forms of understanding presupposed in the formation of judgments. The strong influence of Kant is evident in the theories of categories of such philosophers as G. W. F. Hegel, Edmund Husserl, and Charles Sanders Peirce.

Peirce's theory is closely connected with his contributions to logic, but his conception of what constitutes a category is sufficiently Kantian to distinguish his theory radically from the theory usually associated with the development of modern logic.

THEORY OF TYPES. Bertrand Russell originally devised his theory of types as a means of avoiding a contradiction he had discovered in Gottlob Frege's logic, but the theory has profound implications for philosophy in general, and under its influence "category" has come to be used frequently as a synonym for "logical type."

As the theory of types is presented in *Principia Mathematica,* its cardinal principle (called by Russell the "vicious-circle principle") is that whatever involves all of a collection must not be one of the collection. The class of white objects, for example, includes (and hence involves) all white objects, and to say that this class is itself a white object is to violate the principle and to utter nonsense. The set of entities consisting of all white objects and the class of white objects is for Russell an "illegitimate totality," a set that "has no total" in the sense that no significant statement can be made about all its members. The purpose of the theory of types is to provide a theoretical basis for breaking up such a set into legitimate totalities. A totality is legitimate when and only when all its members belong to the same logical type, and two entities are of different logical types when and only when their inclusion in the same class yields an illegitimate totality. Whenever an entity involves all the members of a given class its logical type is said to be higher than the type of the members of this class. Logical types thus form an infinite hierarchy with individuals at the lowest level, or zero type, classes of individuals at the next level, then classes of classes, and so on. Since to every class there corresponds a defining property of that class, there is an equivalent hierarchy of logical types with individuals again at the lowest level, but with properties of individuals next, then properties of properties of individuals, and so on. "*X* is a member of the class of white objects" is equivalent to "*X* is white," and the two sentences "The class of white objects is a white object" and "The color white is white" are equally expressions of a type-mistake or category-mistake and are equally nonsensical.

The theory of types, if true, gets rid of the contradiction Russell wanted to avoid. This contradiction arises when the class of all classes that are not members of themselves is said to be or not to be a member of itself. According to the theory of types the attempt to make either assertion violates the vicious-circle principle and results in nonsense. But if this way of avoiding the contradiction is to be satisfactory, there must be reasons for accepting the theory of types other than the fact that if it is accepted the contradiction it was designed to avoid is avoided. Efforts to find such reasons have carried investigations concerning the theory of types from the sphere of technical issues in mathematical logic into the sphere of philosophical issues in a theory of categories. Developments in both spheres have often proceeded independently, and even though technical work in mathematical logic has developed alternatives to the theory of types (especially to the theory as first stated by Russell), the fact that the theory is not needed to avoid the original contradiction is not in itself conclusive evidence that the theory has nothing to be said for it as a theory of categories.

Russell offered in support of the theory of types the fact that it outlaws not only conditions giving rise to the paradox concerning class membership but also those giving rise to an indefinite number of other paradoxes of self-reference, including the ancient paradox of the liar. But alternative ways of avoiding these other paradoxes have been developed. More serious than its nonuniqueness as a consistent solution to the problems it was designed to avoid is a difficulty intrinsic to the theory itself. Even if the theory is true, there seems to be no way to state it without contradiction. The word *type* illustrates the point. In stating the theory one uses this word, which is itself a particular entity, with reference to all entities, so one entity is made to involve the collection of all entities. Russell tried to cope with the difficulty by proposing that a difference in logical type be taken as a difference in syntactical function rather than a difference in the totalities to which two entities may be legitimately assigned. Instead of saying that the color white and a table are of different logical types because the latter but not the former can be included in the class of all white objects without forming an illegitimate totality, we may say that the phrases "the color white" and "a table" belong to different logical types because the latter but not the former yields a significant statement when it replaces *X* in the sentence-form "*X* is white."

Reference to linguistic expressions rather than entities avoids a vicious-circle fallacy because the hierarchy of types asserted by the theory then includes only the totality of expressions within a given language, not the totality of all entities. But any given statement of the theory must be in a metalanguage whose expressions are not included in the totality of expressions covered by the statement. While the theory can thus never be applied to the language in which it is itself stated, it can always in principle be restated in a further language (a meta-metalanguage) so that it applies to the language in which it was originally stated as well as the language to which it originally applied. Universal application of the theory is thus possible in principle by proceeding up an infinite hierarchy of languages, while the application of the theory to each

particular language asserts the existence of an infinite hierarchy of types of syntactical functions within that language. But in neither case is there the simple assertion that the class of all entities comprises an infinite hierarchy of logical types.

The conception of logical type as syntactical function is much easier to maintain when the expressions typed are those of an artificial language, such as a logical calculus, rather than those of a natural language, such as English. Generalization about the totality of expressions in an artificial language is easy because this totality is generated by the rules one must lay down if one is to construct an artificial language in a clear and definite sense. But such relativity to the rules of an artificial language makes it impossible to maintain all that was originally claimed for the theory of types. Russell was originally understood as claiming to have *discovered* that what appears to be stated by sentences like "The color white is white" and "The class of white objects is a white object" is simply nonsense. But then it seems that the most one can say is that Russell *constructed* an artificial language (a calculus or formalism) in which the translations of these English sentences are not well-formed formulas. The mere construction of such a language is clearly not the same as the discovery that in point of logic certain apparent statements are really nonsense. The case against Russell's original claim is all the more damaging in view of the fact that formalisms have since been constructed in which translations of certain sentences that are nonsense according to the theory of types are well-formed formulas, and the contradiction the theory of types was designed to avoid does not appear. Enlarging the notion of logical type to include semantic as well as syntactical function does not change the picture. Semantic rules for an artificial language are necessary if one is to do certain things with the language, but these rules, like syntactical rules, are stipulated in the construction of the formalism; addition of such rules in no way furthers the claim to having discovered that certain sentences are nonsense rather than having constructed a language in which they become nonsense.

CATEGORIES AS DISCOVERED IN A NATURAL LANGUAGE

The claim to discovery is essential to a theory of categories, and the claim may still be made if types are found among the expressions of a natural language rather than imposed on the expressions of an artificial language. Instead of beginning with the vicious-circle principle as defining a condition we must impose on any language if we want to make sense, we may begin with expressions in the natural language we ordinarily use—expressions with which we assume we make sense, if we make sense at all—and try to determine what differences in type our making sense requires us to recognize in these expressions. This sort of approach is taken by Gilbert Ryle in *The Concept of Mind,* where he considers expressions we use in talking about mental powers and operations and argues that certain of these expressions cannot belong to the same type or category as others. Ryle's test for a category-difference is a case where one of two expressions cannot replace the other without turning the literal meaning of a sentence into an absurdity. To begin with an obvious case, when "the man" in "The man is in bed" is replaced by "Saturday" the result is clearly an absurd sentence if taken literally. Less obvious cases often go undetected by philosophers and remain a source of philosophical confusion. "He scanned the hedgerow carefully" becomes absurd when "saw" replaces "scanned," although the absurdity disappears when the adverb is omitted. Failure to note that "to see" belongs in the category of "achievement" verbs while "to scan" is a "task" or "search" verb has misled philosophers to posit a mental activity corresponding to seeing that is analogous to the genuine activity of scanning.

For Ryle categories are indefinitely numerous and unordered. The totality of categories is not in principle an infinite hierarchy of types; categories provide no architectonic such as Kant's fourfold division of triads; and there is no distinction setting off one category from all the others as basic regardless of their number, as Aristotle's distinction between substance and accident. There are thus no mistakes that are strictly category-mistakes rather than ordinary equivocations or absurdities. Ryle explains in his article "Categories" that he uses "absurdity" rather than "nonsense" because he wants to distinguish a category-mistake from mere sound without sense. According to Ryle, a category-mistake is not a meaningless noise but a remark that is somehow out of place when its literal meaning is taken seriously; many jokes, he observes, are in fact "type-pranks."

WHAT IS A THEORY OF CATEGORIES?

The above observations suggest that Ryle has no *theory* of categories at all—no principles by which categories can be determined and ordered. Yet he seems unwilling to give up all claims to a theory of categories. He is especially concerned with countering the impression that category-differences are on a par with differences created by a particular set of linguistic rules. In his article "Categories" he

considers briefly the question What are types of? He suggests that instead of saying absurdities result from an improper coupling of linguistic expressions, it is more correct to say that they result from an improper coupling of what the expressions signify. But one must be wary of saying that types are types of the *significata* of expressions. A phrase like "*significata* of expressions" can never be used univocally, because such use presupposes that all *significata* are of the same type. Ryle claims we can get along without an expression that purports to specify what types are types of, since the functions of such an expression are "purely stenographic"; if we want an expression performing these functions, he suggests "proposition-factor" but cautions that to ask what proposition-factors are like is ridiculous since the phrase "proposition-factor" has all possible type-ambiguities.

Ryle seems hardly to have advanced the question of the status of a theory of categories beyond the point where Russell left it. It appears to be just as difficult to establish category-differences by appeal solely to ordinary language as to establish them by appeal solely to an artificial language. J. J. C. Smart points out, in "A Note on Categories," that with Ryle's test of a category-difference we are led to make very implausible (if not absurd) claims about category-differences. When, for example, "table" replaces "chair" in "The seat of the chair is hard," the result seems clearly an absurd sentence. Yet if "table" and "chair" do not belong in the same category, what words do? If the phrase "category-difference" is to have anything like the force it has had from Aristotle to Russell, the claim to having discovered that "table" and "chair" are expressions in different categories is itself absurd. Though Ryle may not want to make the claim, he cannot avoid it and maintain his test of a category-difference.

Yet Ryle, whatever his intentions, may be said to have established the negative point that absurdity alone is never a sufficient test of a category-mistake. Aristotle, Kant, and Russell each began with metaphysical or logical principles that purport to set limits of literal sense; a violation of these principles results either in sound without sense or in intellectual illusion, and in both cases in more than simple absurdity. Ryle appears to want the advantages of a theory of categories and at the same time to avoid the embarrassment of having to defend its principles. Such a theory promises to rid philosophy of many fallacious arguments and contradictions, but the promise is worthless if the principles of the theory are no more tenable than the arguments and contradictions it sweeps away. Aristotle's metaphysics of substance and accident, Kant's transcendental logic, and Russell's elevation of the vicious-circle principle have proved as philosophically debatable as Platonic forms, proofs for the existence of God, and paradoxes of self-reference. It is comforting to believe that such debatable principles can be discarded and that the forms, proofs, and paradoxes can be exposed as category-mistakes by appeal to nothing more than what a man of common sense will recognize as an absurdity in his own ordinary language. But unfortunately our common use of "absurdity" covers too much. One can hardly hope to rid philosophy of Platonic forms with no more argument than the claim that saying the color white is white is like saying the seat of a table is hard.

Ryle also calls attention to another negative point about a theory of categories. The theory cannot have a subject matter in the usual sense. We cannot generalize about all proposition-factors, all entities, or all of whatever it is types are said to be types of as we generalize about, for example, all bodies or all biological organisms. We may say that every proposition-factor is of some type, but we cannot say what it is like regardless of its type as we can say what every body or biological organism is like regardless of its type. Since everything we can talk about is a proposition-factor, we have nothing with which they can be contrasted; we do, however, have things with which to contrast bodies and biological organisms. Ryle sees this point as forcing us to accept a phrase like "proposition-factor" as merely a kind of dummy expression we may use to preserve the ordinary grammar of "type" and "category," although the important thing is not to preserve the grammar but to avoid the error of thinking we can preserve it with other than a dummy expression. If we take "proposition-factor" as a metalinguistic expression applying to factors in a particular language, we succeed in preserving the grammar without a dummy expression, but only at the price of making categories relative to a particular set of linguistic rules. The use of a dummy expression is at least consistent with the claim (which Ryle seems to want to make) that a recognition of absurdity is not relative to the rules of a particular language. We may be said to recognize, regardless of our language, the absurdity of saying that the seat of a table is hard or that the color white is white, although we are unable to give criteria of absurdity.

Aristotle tried to cope with the subject-matter problem by holding that while we cannot generalize about all entities as we can about all bodies or all biological organisms ("being is not a genus," as he put it), we can have a science of being because there is one primary type of being—substance—and every other type exists, by being

an accident of substance. Although we have, then, nothing with which to contrast all beings, we can contrast substances with accidents, and the science of substance is the science of being *qua* being in that conditions for the being of substance are conditions for the being of everything else. A theory of categories may thus be founded on the principle that substances alone can have accidents and all categories other than substance are categories of accidents. For Kant categories do not distinguish beings or entities but a priori forms of understanding, and, unlike Aristotle's beings or Ryle's proposition-factors, these forms comprise not everything we can talk about but only necessary conditions for judgments about objects of experience. The forms stand in sharp contrast with other objects of discourse and constitute a single subject matter belonging to the science of transcendental logic.

Neither Aristotle's nor Kant's theory of categories seems immune to the objection that its subject matter is created rather than discovered. Aristotle's pronouncements about substance and accident and Kant's about forms of understanding each provide principles that yield a scheme of categories, but one may ask whether the pronouncements are anything more than rules for the construction of a certain kind of language—whether the construction of an Aristotelian metaphysics or that of a Kantian transcendental logic provides a theory of categories with anything more than an artificial language within which certain category-differences are established. An answer to this question is proposed by P. F. Strawson in his *Individuals.* Strawson suggests that theories of metaphysics have tended to be either descriptive or revisionary. A metaphysics is descriptive insofar as it yields a scheme of categories that describes the conceptual scheme we actually presuppose in ordinary language. A theory becomes revisionary to the extent that it leads to a departure from our ordinary scheme. Strawson cites the metaphysical theories of Aristotle and Kant as descriptive, those of René Descartes, Gottfried Wilhelm Leibniz, and George Berkeley as revisionary. While all five philosophers construct special languages, only Aristotle and Kant do so in a way that results in a scheme of categories that describes the conceptual scheme of our ordinary language.

But if in this sense Aristotle and Kant in their theories of categories describe rather than create a subject matter, what they describe is not what they claim as their subject matter. Strawson professes in his own theory of categories to describe the conceptual scheme of our ordinary language, but he does not profess to give principles of being *qua* being or a transcendental deduction of pure concepts of the understanding. If Aristotle and Kant to some extent describe the scheme Strawson sets out to describe, this achievement was certainly not their primary objective, and since they differ radically at crucial points, as in their views of alteration and substantial change, they can hardly be said in any case to describe the same scheme. One must say, rather, that each offers metaphysical or transcendental hypotheses that purport to account for and establish the necessity of the conceptual scheme underlying common sense. One may of course accept much of what they say in description of their schemes as true of what one takes to be our commonsense scheme and yet reject their hypotheses. With the rejection there is no need to defend the hypotheses' claims to a metaphysical or transcendental subject matter, but one then needs to explain how our commonsense scheme is subject matter for description. A description of common features in the grammars of Indo-European languages is not exactly what Strawson means by a description of the conceptual scheme of our ordinary language. But it can hardly be said that his efforts to distinguish the two descriptions are entirely successful. In some of his arguments he seems to appeal to metaphysical hypotheses of his own and hence to have a theory accounting for, and not simply a description of, the conceptual scheme he claims as his subject matter. In other arguments he seems, like Ryle, to make an ultimate appeal to our commonsense recognition of absurdity.

The construction of a theory of categories as descriptive metaphysics differs, according to Strawson, from what has come to be called philosophical, or logical, or conceptual analysis. But the difference is not "in kind of intention, but only in scope and generality." Strawson describes philosophical analysis as relying on "a close examination of the actual use of words," and while this is "the best, and indeed the only sure, way in philosophy," what it can yield is not of sufficient scope and generality "to meet the full metaphysical demand for understanding." But Strawson does not elaborate the demand and gives no criterion for deciding when philosophical analysis must give way to descriptive metaphysics. He sometimes implies that we may pass imperceptibly from one to the other, and this may be the case if to do descriptive metaphysics is simply to articulate what is presupposed in a given philosophical analysis. But it can hardly be the case if descriptive metaphysics, unlike philosophical analysis, has its own peculiar subject matter—being *qua* being, pure concepts of the understanding, our commonsense conceptual scheme, or whatever. Philosophical analysis is clarification of thought about a given subject matter, and to articulate the presuppositions of a given

analysis is not to analyze a new subject matter but only to push the original analysis as far as we can. In the end we may arrive at distinctions that agree with what philosophers from Aristotle to Strawson have called "category-differences," and there is no harm in using the label if we mean only that the distinctions are ultimate in the analysis we have given and not also that they have to be supported by a hypothesis about a special subject matter. We can hardly make the additional claim without passing beyond the point where we can hope for help from philosophical analysis.

HISTORICAL NOTES

STOICS AND NEOPLATONISTS. In place of Aristotle's ten categories the Greek Stoics substituted four "most generic" notions or concepts: substratum, or subject; quality, or essential attribute; state, or accidental condition; and relation. The Stoic view, as well as the Aristotlelian doctrine, was criticized by the Neoplatonist Plotinus. In his *Sixth Ennead* Plotinus argued that the ultimate categories are neither the Aristotelian ten nor the Stoic four but correspond to the five "kinds" listed in Plato's *Sophist*: being, rest, motion, identity, and difference. The central point for Plotinus was that different categories apply to the intelligible and sensible worlds, the ultimate categories applying only to the former. Plotinus's views on categories figured prominently in medieval discussions only as they were considerably modified by his pupil Porphyry. In Porphyry's short commentary on Aristotle's *Categories,* generally known as the *Isagoge* (*Εἰσαγωγὴ,* "Introduction"), he accepted Aristotle's list of ten but raised Plotinian questions about the way they exist. He noted that categories are genera and asked whether genera and species subsist (exist outside the understanding) or are in the naked understanding alone; whether, if they subsist, they are corporeal or incorporeal; and finally, whether they are separated from sensibles or reside in sensibles. He remarked that these questions are too deep for an introductory treatise, and we have no record of how he thought they should be answered.

BOETHIUS. Boethius translated the *Isagoge* into Latin, along with Aristotle's *Categories* and *On Interpretation.* He also wrote a commentary on the *Isagoge,* offering answers to Porphyry's unanswered questions, and thus began a tradition, which persisted throughout the medieval period, of accepting Porphyry's questions as presenting the fundamental issues for any account of categories. Since genera and species appear most prominently as genera and species of substances, the issues centered first of all in the signification of common nouns taken as names of kinds of substances. The medieval "problem of universals" thus arose from Porphyry's questions about Aristotle's categories, and prominent medieval philosophers, such as Peter Abelard, Thomas Aquinas, John Duns Scotus, and William of Ockham, are known as conceptualists, realists, or nominalists because of their answers to these questions. The important point for a history of theories of categories is that the discussion of the problem of universals by major figures in medieval philosophy occurred within an unquestioned framework provided by Aristotle's theory of categories—in particular, within a framework that presupposed the basic Aristotelian interrelation of substance and accident and essential and accidental predication.

LOCKE AND HUME. The Aristotelian framework broke down in modern pre-Kantian philosophy. Signs of the breakdown were evident in Thomas Hobbes and Descartes, but its full force appeared in John Locke and David Hume. With Locke's account of substance as an "unknown something" underlying appearances, essential predication in the category of substance becomes impossible, and the signification of common nouns supposed to name kinds of substances can be fixed only by "nominal essences," by conventional factors, rather than by Ockham's "natural signs in the soul." Essential predication, and hence necessary truth, remains possible only when the subjects are things of our own creation ("mixed modes") and not when they are substances in the real world.

The full consequences of Locke's departure from an Aristotelian framework were drawn by Hume. If it is impossible to know what something in the real world necessarily (essentially) is, it is also impossible to know that any one thing in the real world is necessarily connected with another or that any state of a thing at one time is necessarily connected with its state at another time. In other words, not only substance but also causality—an equally if not more fundamental notion (though not recognized as a category by Aristotle)—is made a matter of habit and custom. The stage was set for Kant to answer Hume with a radically new theory of categories.

HEGEL. Despite the radical differences between Kantian and Aristotelian categories, two basic points of similarity remain: (1) Categories provide form but not content for cognitive discourse about the world and thus serve to distinguish what we can meaningfully say in such discourse from what we may seem to say when we make category-mistakes or misapply categories. (2) Categories presup-

pose the substance-accident (subject-predicate) form basic to Aristotelian logic. Hegel's philosophy retains neither of these points of similarity, although he adopted the Kantian view that the clue to a system of categories is to be found in logic. But instead of turning to logic as a study of forms of reasoning without regard for content, Hegel turned to logic as a dialectical process in which form and content are inseparable. The essential nature of this process is seen not in the forms under which subject and predicate are brought together in the premises of reasoning to make affirmative, negative, disjunctive, hypothetical, and other types of judgment but in the basic stages through which the process itself repeatedly moves. These stages Hegel called "thesis," "antithesis," and "synthesis," and he took them as interrelating the basic ideas, notions, or principles of reason, which he also called "categories." This interrelation of categories constitutes both Hegel's system of philosophy and what he held to be the "system of reality." The categories, then, are many, and their exact number cannot be determined until the system of reality is fully articulated. Hegel thus marked the beginning of a tradition in modern philosophy, in which "category" means simply any basic notion, concept, or principle in a system of philosophy.

This use of "category" is standard not only among Hegel's progeny of absolute idealists but also among metaphysicians generally, who dissociate themselves from analytical philosophy. The use remains even when there is no vestige of Hegel's threefold pattern of thesis, antithesis, and synthesis as a means of ordering the principles of speculative philosophy. The categorial scheme in Alfred North Whitehead's *Process and Reality,* for example, is readily understood as dealing with the sort of notions Hegel called "categories" but hardly with categories in the Aristotelian-Kantian sense of setting limits of cognitive meaning, a sense that still survives in analytical philosophy.

PEIRCE. The collapse of Kant's theory of categories is inevitable, according to Peirce, as logic advances beyond the subject-predicate form recognized by Aristotle. So long as statements like "John gave the book to Mary" are not seen as possessing a logical form fundamentally different from and coordinate with the simple subject-predicate form of statements like "John is tall," categories are determined by what may be taken as different forms of this one-subject–one-predicate relation. Aristotle and Kant analyzed the forms differently, but the relation analyzed was the same. With the development of logic beyond Aristotle (a development to which Peirce made significant contributions), statements like "John gave the book to Mary" are recognized as statements with three-place predicates (x gave y to z) and are different in logical form from statements with one-place predicates (x is tall). Peirce claimed to have demonstrated in his "logic of relatives" that although one-place, two-place, and three-place predicates are basically different in logical form, predicates with more than three places have no features of logical form not already found in three-place predicates.

The demonstration remains one of the more questionable parts of his logic, but Peirce accepted it as proof that in formal logic there are but three fundamentally different types of predicates and hence that there are but three categories. He sometimes referred to his categories as the "monad," the "dyad," and the "polyad," but he preferred the more general expressions "firstness," "secondness," and "thirdness." As genera (or modes) of being, the categories are designated as "pure possibility," "actual existence," and "real generality." A pure possibility stands by itself, determined by nothing but conditions of internal consistency; what actually exists stands in relation to other existences and to some extent both determines and is determined by them; a true generalization is a representation related to other representations, to actually existing things, and to pure possibilities. In his philosophical cosmology Peirce had three universes corresponding to the three modes of being, and in his semeiotic theory, or theory of signs, he developed an extensive classification of signs, with the main divisions triadic, each triad comprising a firstness, a secondness, and a thirdness. Although Peirce's categories thus function architectonically somewhat as Hegel's thesis, antithesis, and synthesis, they serve, as Hegel's triad does not, to set limits of cognitive meaning. Though Peirce did not use the phrase "category-mistake," he said repeatedly in his later writings that nominalism, which he regarded as the great error in the history of philosophy, arises from the failure to recognize real generality as a mode of being distinct from actual existence. In arguing that universals have no actual existence, the nominalist has failed to see that to ask in the first place whether they have such existence is a category-mistake. In his final years Peirce labored to show that the pragmatic criterion of meaning, which he propounded early in his career, is not only consistent with but actually necessitated by his theory of categories.

HUSSERL. The role of categories in setting limits of cognitive meaning figures prominently in the philosophy of Husserl. To determine "primitive forms" or "pure categories" of meaning is the first task of a "pure philosophical grammar." The fundamental form is that of propositional meaning, and other primitive forms, such

as the nominal and adjectival, are forms of meaning that belong to constituents of a proposition. After determining these pure categories of meaning, pure logical grammar turns to primitive forms or categories of the composition and modification of meaning (forms such as those exhibited by propositional connectives and modal expressions). In addition to a pure logical grammar, Husserl held, there are a pure logic of consistency (noncontradiction) and a pure logic of truth. The picture is further complicated in that pure logic may be taken as giving rise to a formal ontology and, again, developed into a transcendental logic. A full account of categories requires the full development of logic in all its phases, and in this respect Husserl's view of categories seems reminiscent of Hegel. But at no point (even in formal ontology) did categories cease for Husserl to be purely formal and become inseparable from content. Husserl was careful to distinguish the kinds of nonsense precluded by his categories from nonsense of content (*inhaltlich Unsinn*). A phrase like "if-then is round" is nonsense because it violates a category-difference, a condition of meaningfulness established by logic alone; a phrase like "the seat of the table is hard" violates no such condition, and its nonsense arises from a material, not a formal (logical), incompatibility. While at times Husserl's language may suggest what Rudolf Carnap and others have since called "syntactical categories," it should be noted that Husserl had nothing like Carnap's technical distinction between syntax and semantics and that the "syntactical categories" of Husserl's pure logical grammar are in Carnap's sense neither purely syntactical nor semantical.

FREGE AND WITTGENSTEIN. In their philosophies of mathematics and logic both Peirce and Husserl remained close enough to Kant not to accord set theory the fundamental role it has come to play in logic and the foundations of mathematics. Frege, although he did not present any of his views under the heading "a theory of categories," did far more than Peirce or Husserl to shape the discussion of categories in the twentieth century. Frege analyzed sense and reference, concept and object (notions fundamental to Peirce's and Husserl's theories of categories) in a way that permitted him to take set theory as basic in mathematics and to define cardinal numbers as classes of classes. Russell's efforts to cope with the contradictory notion of the class of all classes not members of themselves (a notion one seems forced to admit with Frege's analysis) produced the theory of types.

The conclusion suggested by the difficulties encountered in the theory of types, that categories as setting limits of cognitive meaning are not proper subject matter for a theory, was first advanced by Ludwig Wittgenstein. In his early work, *Tractatus Logico-philosophicus,* Wittgenstein spoke of the limits of cognitive meaning as the ineffable, as what can be shown but not said. In his later writings he repudiated the suggestion that the limits constitute an ineffable subject matter, something to be unveiled but not articulated as a theory by philosophical analysis. Nevertheless, with the assumption of such subject matter philosophical clarity is to be achieved by the construction of an ideal language, a language is stripped of all superfluous symbolism and is hence unable to give the illusion of transcending the ineffable limits of cognitive meaning. But if this assumption is itself an illusion, as Wittgenstein later held, if we can no more show than we can state the limits of *all* language, then philosophical clarity can be achieved only piecemeal, context by context; there is no short cut via an ideal language. And a fortiori there is no universal scheme of categories to be unveiled, let alone to be established by a theory. Wittgenstein's influence may be seen in the hesitation of Ryle, Strawson, and other present-day analytical philosophers to claim that categories should (or can) have the absolute universality claimed in theories of categories from Aristotle's to the theory of types.

See also Aristotle; Berkeley, George; Boethius, Anicius Manlius Severinus; Descartes, René; Frege, Gottlob; Hegel, Georg Wilhelm Friedrich; Hume, David; Husserl, Edmund; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Peirce, Charles Sanders; Platonism and the Platonic Tradition; Plotinus; Porphyry; Russell, Bertrand Arthur William; Ryle, Gilbert; Smart, John Jamieson Carswell; Strawson, Peter Frederick; Type Theory; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

STANDARD HISTORIES

Ragnisco, P. *Storia critica delle categorie dai primordi della filosofia greca fino al Hegel,* 2 vols. Florence, 1871.

Trendelenburg, A. *Historische Beiträge zur Philosophie, Vol. I, Geschichte der Kategorienlehre.* Berlin: G. Bethge, 1846.

NEWER WORKS

Ackrill, J. L. *Aristotle's Categories and De Interpretatione.* Oxford: Clarendon Press, 1963. A new translation, written for the serious student of philosophy who does not read Greek. Copious notes, constituting over half the volume, provide an excellent scholarly commentary.

Anscombe, G. E. M., and P. T. Geach. *Three Philosophers.* Ithaca, NY: Cornell University Press, 1961. Contains penetrating but difficult discussion applying recent

techniques of analysis to problems of categories in Aristotle, Thomas Aquinas, and Frege.

Black, Max. "Russell's Philosophy of Language." In *The Philosophy of Bertrand Russell,* edited by P. A. Schilpp. Evanston, IL: Library of Living Philosophers, 1946, 229–255. Difficulties with the original theory of types as a theory of categories. See also Russell's reply on pp. 691–695.

Cross, R. C. "Category Differences." *PAS* (1958–1959): 255–270.

Hall, Everett W. "Ghosts and Categorical Mistakes." *Philosophical Studies* 7 (1956): 1–6.

Hall, Everett W. *Philosophical Systems: A Categorical Analysis.* Chicago: University of Chicago Press, 1960. Sketchy survey of different philosophical systems as different categorial schemes.

Harrison, Bernard. "Category Mistakes and Rules of Language." *Mind* 74 (1965): 309–325.

Hillman, D. J. "On Grammars and Category Mistakes." *Mind* 72 (1963): 224–234.

Pap, Arthur. "Types and Meaninglessness." *Mind* 69 (1960): 41–54.

Passmore, John. *Philosophical Reasoning.* New York: Scribners, 1961. Difficulties of maintaining a theory of categories are discussed in Ch. 7.

Popper, Karl R. *Conjectures and Refutations.* New York: Basic, 1962. Chs. 11–14 offer criticisms of the view that limits of meaningfulness can be set by discovery of types or categories.

Quine, W. V. *Word and Object.* Cambridge, MA: Technology Press of the Massachusetts Institute of Technology, 1960. Presents a skillful defense of the view that categories are relative to language.

Rorty, Richard. "Pragmatism, Categories, and Language." *Philosophical Review* 70 (1961): 197–223.

Ryle, Gilbert. "Categories." In *Logic and Language,* Second Series, edited by A. G. N. Flew, 65–81. Oxford: Blackwell, 1953.

Ryle, Gilbert. *The Concept of Mind.* London: Hutchinson's University Library, 1949.

Ryle, Gilbert. "Systematically Misleading Expressions." In *Logic and Language,* First Series, edited by A. G. N. Flew. Oxford, 1951. Pp. 11–36.

Shwayder, D. S. *Modes of Referring and the Problem of Universals.* Berkeley: University of California Press, 1961. Technical discussion of points crucial to a theory of categories, especially Strawson's.

Smart, J. J. C. "A Note on Categories." *British Journal for the Philosophy of Science* 4 (1953): 227–228.

Sommers, Fred. "Types and Ontology." *Philosophical Review* 72 (1963): 327–363.

Strawson, P. F. *Individuals.* London: Methuen, 1959.

Thompson, Manley. "On Category Differences." *Philosophical Review* 66 (1957): 468–508.

Warnock, G. J. "Categories and Dilemmas." In *English Philosophy since 1900.* London and New York: Oxford University Press, 1958. Ch. 7.

Manley Thompson (1967)

CATEGORIES, SYNTACTICAL AND SEMANTICAL

See *Syntactical and Semantical Categories*

CATHARSIS

See *Katharsis*

CATTANEO, CARLO

(1801–1869)

Carlo Cattaneo is possibly the most interesting Italian philosopher of the nineteenth century, and was a distinguished scholar in history, economics, linguistics, and geography. Born in Milan, he received a law degree from the University of Pavia, where for some years afterward he taught Latin and the humanities. In 1839 he founded the journal *Il Politecnico,* which he described as "a monthly repertory of studies applied to culture and social prosperity." Cattaneo led the 1848 Milanese insurrection against Austrian rule, the story of which he related in a masterly booklet, *L'insurrezione di Milano nel 1848* (in *Scritti storici e geografici,* Vol. IV, Florence, 1957; first published in French in Paris, 1848). When the first Italian war of independence ended in failure, in 1849, Cattaneo went into exile, first in Paris and then in Lugano, Switzerland, where for several years he taught philosophy in the local lyceum. Although he was appointed a deputy to the Italian parliament in 1860, he refused to enter the parliament house in order not to have to swear allegiance to the king. He continued to spend most of his time at Lugano, where he edited a new series of *Il Politecnico* from 1860 to 1863, the first series having been suspended in 1844.

The main influence on Cattaneo was the Lombard Enlightenment philosophy espoused by his teacher G. D. Romagnosi, which was interested in scientific inquiry as related to the well-being of society and concerned with progressive government—facets visible in the work of Alessandro Volta and Cesare Beccaria. Cattaneo blended this inheritance with reflection on his own research in fields other than philosophy but generally disregarded philosophical tradition. He developed an original though unsystematic body of ideas that can best be described as an empirical, scientifically minded phenomenology of history or a nonidealistic historicism. The contemporary reader may catch a Marxian ring or occasionally find a

resemblance to such thinkers as Wilhelm Dilthey, G. H. Mead, and John Dewey.

For Cattaneo the philosopher's task consists in clarifying objective current historical problems rather than subjective difficulties. There is no single problem to be made the center of systematic speculation, nor any logical or genetic "first truth" on which the chain of deductive reasoning may be hung. There is instead a plurality, itself subject to change, of well-determined and interrelated problems. There are no final solutions to problems, but only a body of perfectible solutions, which are discovered not by absolute reason but by general human reasonableness. Logic is the theory of scientific research; in philosophy, too, the experimental method, which unites men, must supersede metaphysics, whose continuous veerings divided men.

We know in order to act. The aim of all intellectual endeavor is to change the face of Earth for the good of humankind: Both nature and society must be "transformed" by man-invented techniques. Insofar as he brings about a knowledge that is public and beneficial, the philosopher is "a craftsman" who works "for the common people"—"we are all workmen if we supply something useful to mankind." To such philosophy Cattaneo contrasted "the philosophy of the schools," whose "ontological hammer" generated "a hidden, priestly wisdom scorning the common people," drawing on "fantastic hypotheses and imaginary intuitions," and "consuming itself in the repetition of empty formulae"—with the result of "throwing wide-open an immeasurable gap between doctrine and fact about man." In saying such things Cattaneo had in mind particularly Antonio Rosmini-Serbati, who was then trying to reconcile philosophical Catholicism with the subjectivism of modern philosophy.

For Cattaneo thought is social action, and it must be studied in the various human activities. There is no essence of thought to be reached directly. To become acquainted with his own nature, man must not recede into himself but rather must go out into the world to collect information. A complete science of thought amounts to knowledge of all that mankind has produced. By "mankind" Cattaneo meant empirical men in their finite world; while professing to be a follower of Giambattista Vico (who was at the time almost unknown), he was highly critical of Vico's oversimplified principles of interpretation, especially of the notion of historical cycles ("Su la *Scienza nova* del Vico," 1839; "Considerazioni sul principio della filosofia," 1844).

Cattaneo intended the phenomenology of history to overcome in a new way the traditional opposition of appearance and reality. What appears to us is what there is—all the reality we can or must cope with—and we cannot reach it outside the social development of humankind (see especially "Un invito alli amatori della filosofia," 1857). This must be construed methodologically, according to what Cattaneo labeled the "psychology of associated minds." The "solitude of the new-born in front of *things*" is a philosophical myth. "Even sensation is from the beginnings a social fact," and "whatever idea one comes to conceive is never the operation of a solitary mind but rather of several associated minds." (*Psicologia delle menti associate,* 1859–1863, unpublished; quotations taken from *Scritti filosofici,* Vol. II, p. 14; Vol. I, p. 448; Vol. II, p. 16). To help us understand the varieties of human history, a social psychology supported by scientific method must replace individual psychology as connected with that "lobby of theology" which was "[René] Descartes' solitude of consciousness."

See also Beccaria, Cesare Bonesana; Descartes, René; Dewey, John; Dilthey, Wilhelm; Historicism; Mead, George Herbert; Romagnosi, Gian Domenico; Vico, Giambattista.

Bibliography

ADDITIONAL WORKS BY CATTANEO

In addition to the essays mentioned in the article, Cattaneo's main philosophical work consists of studies on Romagnosi, Tommaso Campanella, Wilhelm von Humboldt, and others, the forewords to the annual volumes of *Il Politecnico,* and the lectures (*Lezioni*) delivered at the Lugano Lyceum on cosmology, psychology, ideology (see the pages on categories and language), logic, and law and morals. Students of Cattaneo's philosophy, however, should take into account also many of his writings in other fields.

There are several anthologies of Cattaneo's papers, especially the philosophical ones, the best being those edited by Gaetano Salvemini (Milan, 1922) and Franco Alessio (Florence, 1957). A complete edition of Cattaneo's works, divided into five sections according to subject matter, appeared in 1956 in Florence; the philosophical section, edited by N. Bobbio, comprises three volumes.

WORKS ON CATTANEO

On Cattaneo's life, see the Salvemini introduction. On his historical position, see Mario Fubini, *Il romanticismo italiano* (Bari: Laterza, 1953), passim. On his thought, see the introductions by Alessio and Bobbio, and also Bruno Brunello, *Cattaneo* (Turin, 1925); Alessandro Levi, *Il positivismo politico di Carlo Cattaneo* (Bari: G. Laterza & figli, 1928); and Luigi Ambrosoli, *La formazione di Carlo Cattaneo* (Milan: R. Ricciardi, 1959).

OTHER RECOMMENDED WORKS

Scritti filosofici. Edited by Afredo Saloni. Bari: Editori Laterza, 1965.

I problemi dello Stato italiano. Milano: Mondadori, 1966.

Tutte le opere. Edited by Luigi Ambrosoli. Milan: A. Mondadori, 1967.

Scritti letterari, artistici, linguistici e vari. Edited by Agostino Bertani. Firenze, Le Monnier, 1968.

Scritti scientifici e tecnini. Edited by Carlo G. Lacaita. Firenze: Giunti, G. Barbèra, 1969.

Carlo Cattaneo nel primo centenario della morte. Edited by Adriano Soldini. Bellinzona: Dipartimento della pubblica educazione, 1970.

Carlo Cattaneo. (Studio. Con un'appendice di scritti inediti o dimenticati di C. Cattaneo.) Edited by Biancamaria Frabotta. Lugano: Fondazione Ticino nostro, 1971.

Una filosofia militante. Studi su Carlo Cattaneo. Edited by Norberto Bobbio. Torino: G. Einaudi, 1971.

Opere scelte. Edited by Delia Castelnuovo Frigessi. Torino: G. Einaudi, 1972.

L'uomo e la storia. Storiografia, filosofia della storia, antropologia. Edited by Ferruccio Focher. Milano: Mursia, 1973.

L'opera e l'eredità di Carlo Cattaneo. Edited by Carlo G. Lacaita. Bologna: Il mulino, 1975–1976.

Introduzione a Cattaneo. Edited by Umberto Puccio. Torino: G. Einaudi, 1977.

Antologia degli scritti politici. Edited by Giuseppe Galasso. Bologna: Il mulino, 1978.

Elementi di teoria della propagazione ondosa. Edited by S. Pluchino. Bologna: Pitagora, 1981.

Indice dell'edizione nazionale delle opere di Carlo Cattaneo. Edited by Pietro Esposito. Firenze: Le Monnier, 1995.

Filosofia civile e federalismo nel pensiero di Carlo Cattaneo. Scandicci (Firenze): Nuova Italia, 1996.

La linea lombarda del federalismo. Roma: Gangemi: Fondazione nuovo millennio, 1999.

Del pensiero come principio d'economia publica = Intelligence as principle of public economy. Milano: Libri Scheiwiller, 2001.

Carteggi di Carlo Cattaneo. Serie I, Lettere di Cattaneo. Firenze: Felice Le Monnier; Bellinzona: Edizioni Casagrande, 2001.

Carteggi di Carlo Cattaneo. Serie II, Lettere dei corrispondenti. Firenze: Felice Le Monnier; Bellinzona: Edizioni Casagrande, 2001 .

La scienza nuova dell'umanità: scritti vichiani 1836–1861. Edited by Giuseppe Cospito. Genova, 2002.

Ferruccio Rossi-Landi (1967)
Bibliography updated by Michael J. Farmer (2005)

CAUSAL APPROACHES TO THE DIRECTION OF TIME

What account is to be given temporal priority and of the direction of time? One natural view is that no account is needed (Oaklander 2004), a position that can be defended by arguing, first, that one immediately perceives the succession of events (Bergson 1912), and second, that if one can immediately see that events stand in the relation of temporal priority, then the concept of that relation is primitive and unanalyzable.

There are, however, important objections to this view and to the supporting argument. As regards the latter, the question arises whether perception of change does not turn out, on closer scrutiny, to involve not only a momentary visual state but also short-term memories of immediately preceding visual states. If so, then the acquisition of a belief that something is moving or changing will involve inference, and succession will not be something immediately perceived.

As regards the view itself, one problem is that temporal priority is a relation with certain properties: It is impossible for an event to be earlier than itself; if A is earlier than B, B cannot be earlier than A; and if A is earlier than B, and B earlier than C, then A must be earlier than C. If the concept of the earlier than relation is analytically basic, then no account can be given of these necessary truths: they will have to be treated as synthetic *a priori*. By contrast, if the idea of temporal priority is analyzable, then it should be possible to show that these necessary truths are analytic.

One can assume, then, that the concept of temporal priority must be analyzable. What are the possibilities? The answer is that three main types of accounts have been offered. First, philosophers who favor a tensed account of the nature of time often maintain that the tensed properties of pastness, presentness, and future are basic properties and that the tenseless temporal relations of simultaneity and temporal priority are to be analyzed in terms of those tensed properties (e.g., Broad 1933, Sellars 1962, Prior 1967). According to this view, then, the direction of time logically supervenes on the tensed properties of events.

A second approach holds that if events stand in the relation of temporal priority, and if time has a direction, then such facts must be reducible to properties and relations recognized by physics. The idea, accordingly, is to analyze the relation of temporal priority and the direction of time in terms of such things as the direction of increase in entropy, the direction of the expansion of the universe, or the direction of irreversible processes (e.g., Popper 1956; Grünbaum 1971, 1973; Sklar 1974).

A third possibility is a causal approach. Here the idea is, first, that causal processes involve a direction, and, second, that causal facts are more basic than temporal facts,

with the result that the direction of time can be analyzed in terms of the direction of causation.

How do these three alternatives fare? As regards the first, there are two crucial objections. First, it is clear that the relation of temporal priority cannot be analyzed in terms of the tensed properties of pastness, presentness, and futurity alone, since one event may be earlier than another, though both have the same tensed property of pastness. One needs, then, to introduce additional tensed concepts, such as those of one event's being more past than, and more future than, another. These latter, however, are not plausible candidates for primitive concepts, since then one would be unable to explain, for example, why event A's being more future than event B entails that A is future and B is future. However, if one attempts to analyze those concepts, the natural way of doing so is in terms of the concept of the past, and the concept of the future, with the concept of temporal priority. Such analyses, however, will make the analysis of temporal priority in terms of tensed concepts implicitly circular.

Second, even the concept of futurity itself is not a plausible candidate for a basic concept, since it is plausible that it is concepts that pick out immediately given properties and relations that are analytically basic, and the concept of the future does not pick out a property of events that can be immediately perceived. However, if the concept of the future must be analyzed, how is this to be done except in terms of the idea of the present with the idea of temporal priority? So, once again, the attempt to analyze the relation of temporal priority in terms of tensed concepts can be seen to be circular.

In the case of the second approach—which involves analyzing temporal priority in terms of specialized scientific concepts, such as those of entropy and the expansion of the universe—there are also two main objections. First, most proposals for a scientific analysis of temporal priority entail that it is possible that the universe might undergo a temporal reversal. For the universe, rather than expanding forever, may stop expanding, and then begin contracting. Moreover, if this were to happen, entropy would at some point stop increasing and begin decreasing. The direction of time cannot be analyzed, therefore, in terms of the direction of increase in entropy or in terms of the direction of the expansion of the universe, since such analyses entail the unacceptable consequence that the resulting contraction of the universe would be earlier than the time at which the universe stopped expanding.

Second, there are logically possible worlds that contain temporally ordered events, but no increase in entropy or expansion of the universe. Consider, for example, two uncharged particles rotating endlessly about one another due to gravitational attraction. Accordingly, the concept of temporal priority cannot be analyzed in terms of such scientific concepts.

The conclusion, therefore, is that the first two approaches to the analysis of the concept of temporal priority appear unsatisfactory. If this is so, one is left with the third alternative—that of analyzing temporal priority in causal terms.

A CAUSAL THEORY OF THE DIRECTION OF TIME AND TEMPORAL PRIORITY

The idea of analyzing the concept of temporal priority in causal terms is not a recent development, since it dates back at least to Gottfried Wilhelm Leibniz (1715/1969) and Immanuel Kant (1781/1961). In more recent years it was advanced by the mathematician Alfred A. Robb (1914, 1921), and by philosophers such as Henryk Mehlberg (1935, 1937), Hans Reichenbach (1956), D. H. Mellor (1981, 1995, 1998), and Michael Tooley (1987, 1997), among others.

Before setting out a causal theory, it will be best to address an initial objection, the thrust of which is that it may well be, as many philosophers and scientists believe (e.g., Lewis 1976), that backward causation is logically possible, and, if this is so, how can the direction of time be defined in terms of the direction of causation?

One response, adopted by some advocates of a causal approach (Mellor 1981, 1995, 1998; Tooley 1987, 1997), is to argue that backward causation is not logically possible. However, a different response is available. For if one considers, for example, Dr. No traveling backward in time, then it is natural to say that the temporal ordering of events inside his time machine is opposite to the temporal ordering of events outside of it. If so, then in a world where there is backward causation, one needs the concept of the local direction of time, which can be defined in terms of the direction of causal processes in that region. One could then go on to introduce the idea of the overall direction of the universe, defined, as David Lewis (1976, 1979) suggests, in terms of the direction of most causal processes.

How can temporal priority be analyzed in causal terms? A natural starting point is the following postulate:

(*P*) If *A* causes *B*, then *A* is earlier than *B*.

This gives one a sufficient condition for one event's being earlier than another, but it does not provide a necessary condition. So how can one arrive at necessary and sufficient conditions for one event's being earlier than another?

To arrive at an answer, consider the following two plausible claims:

(*Q*) If *A* is earlier than *B*, and *B* is simultaneous with *C*, then *A* is earlier than *C*.

(*R*) If *A* is simultaneous with *B*, and *B* is earlier than *C*, then *A* is earlier than *C*.

These two postulates, with (*P*), then entail two further, more comprehensive propositions relating causation to temporal priority:

(*S*) If *A* causes *B*, and *B* is simultaneous with *C*, then *A* is earlier than *C*;

(*T*) If *A* is simultaneous with *B*, and *B* causes *C*, then *A* is earlier than *C*.

However, in addition, these two conditions, in conjunction with the fact that temporal priority is a transitive relation, entail another, much more encompassing condition:

(*U*) If $\{A_1, A_2, \ldots, A_i, \ldots, A_{n-1}, A_n\}$ is a set of *n* instantaneous events such that, for every $i < n$, either A_i causes A_{i+1}, or A_i is simultaneous with A_{i+1}, and if, in addition, there is some $i < n$ such that A_i causes A_{i+1}, then A_1 is earlier than A_n.

Principle *U*, entailing, as it does, principles *R*, *S*, and *T*, and more as well, is a comprehensive principle relating causation to temporal priority, and that it follows from the conjunction of the noncausal principles *Q* and *R* with the modest claim involved in *P* shows how powerful principle *P* is.

Principle *U*, of course, still gives one only a sufficient condition for one event's being earlier than another. The idea now, however, is that the sufficient condition that is given by *U* is also a necessary condition. If this is right, then the relation of temporal priority can be analyzed as follows:

A is earlier than *B*

means the same as

For some number *n*, there is a set of *n* instantaneous events $\{A_1, A_2, \ldots, A_i, \ldots, A_{n-1}, A_n\}$ such that, first, *A* is identical with A_1, and *B* is identical with A_n;

second, for every $i < n$, either A_i causes A_{i+1}, or A_i is simultaneous with A_{i+1}; and,

third, there is some $i < n$ such that A_i causes A_{i+1}.

This proposed analysis does, of course, involve a temporal notion—namely, that of simultaneity. However, that will be an objection to the analysis only if the concept of simultaneity itself has to be analyzed in terms of temporal priority. The latter, however, does not seem likely, since it would seem possible for there to be a world that consists of a single moment, containing states of affairs all of which are simultaneous with each other.

OBJECTIONS TO A CAUSAL ACCOUNT

Causal analyses of temporal priority are exposed to a number of objections, many of them advanced by J. J. C. Smart (1971). Among the most important are the following. First, given that the laws of physics do not, with one possible exception, involve any asymmetry, is it possible to explain causal priority without appealing to temporal priority? Second, it is surely logically possible for there to be events that have temporal location, but that have neither causes nor effects. However, this would seem to be ruled out by a causal analysis of temporal priority. Third, is it not also logically possible for there to be moments of time at which no events take place—perhaps because the world contains gappy causal laws? But then there would be no way of ordering that moment relative to other moments. Finally, and even more dramatically, is it not logically possible for there to be a spatiotemporal world that contains no events at all? But then there would be no causal relations, and so, according to a causal theory of temporal priority, no ordering of times in such a world.

With regard to the first objection, the answer is that most present-day analyses of causation offer accounts of the direction of causation that do not involve any appeal to temporal priority (Lewis 1973; Tooley 1987, 1997; Mellor 1995). As regards the second objection, it does not tell against the account set out earlier, since an event that does not itself enter into any causal relations may have temporal location by being simultaneous with an event that does enter into causal relations.

The third and fourth objections are more threatening. One way of responding to these objections is by appealing to possible events and causal relations. Here the idea is, in the case of the third objection, that if the world had been different at certain times, there would have been events when, as things stand, there are no events, and that it is those possible causal relations that make it the case that the time when no events occur has a temporal loca-

tion. Similarly, in the case of the totally empty spatiotemporal world, if there had been events at some times, these would have caused events at other times, and it is those possible causal relations that serve to order moments of time.

The problem with this sort of response is that if temporal order is to be analyzed causally, it seems clear, especially in the case of the totally empty world, that there are no truth makers for counterfactuals concerning such possible events. A different response, however, is available (Tooley 1987, 1997). The basic idea is that if one adopts a realist conception of space-time, then the continued existence of space-time is itself something that requires explanation if it is not to be a cosmic accident. However, what sort of explanation is possible, other than one according to which regions of space-time themselves causally give rise to other regions of space-time? If such immanent causal connections between spatiotemporal regions are possible, then the temporal ordering of different moments of time can, on a causal theory, be given by those causal relations, rather than only by causal relations between events in space-time.

See also Physics and the Direction of Time; Time; Time, Being, and Becoming.

Bibliography

Bergson, Henri. *Matter and Memory*. New York: Macmillan, 1912.

Broad, C. D. *Examination of McTaggart's Philosophy*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1933.

Davies, P. C. W. *The Physics of Time Asymmetry*. Berkeley: University of California Press, 1974.

Grünbaum, Adolf. "The Meaning of Time." In *Basic Issues in the Philosophy of Time*, edited by Eugene Freeman and Wilfrid Sellars, 195–228. LaSalle, IL: Open Court, 1971.

Grünbaum, Adolf. *Philosophical Problems of Space and Time*. 2nd ed. Dordrecht, Netherlands: D. Reidel, 1973.

Kant, Immanuel. *Kritik der reinen Vernunft* (1781). Translated as *Critique of Pure Reason*. London: Macmillan, 1961. See book 2, chapter 2, section 3 of the "Transcendental Analytic."

Leibniz, Gottfried Wilhelm. "The Metaphysical Foundations of Mathematics" (1715). In *Philosophical Papers and Letters*. 2nd ed. Translated and edited by Leroy E. Loemker, 666–674. Dordrecht, Netherlands: D. Reidel, 1969.

Lewis, David. "Causation." *Journal of Philosophy* 70 (1973): 556–567.

Lewis, David. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (1979): 455–476.

Lewis, David. "The Paradoxes of Time Travel." *American Philosophical Quarterly* 13 (1976): 145–152.

Lewis, David. *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986.

Malament, David. "Causal Theories of Time and the Conventionality of Simultaneity." *Noûs* 11 (1977): 293–300.

McCall, Storrs. *A Model of the Universe: Space-Time, Probability, and Decision*. Oxford, U.K.: Clarendon Press, 1994.

McCall, Storrs. "Objective Time Flow." *Philosophy of Science* 43 (1976): 337–362.

Mehlberg, Henryk. "Essai sur la théorie causale du temps." *Studia Philosophica* 1 (1935): 119–258; 2 (1937): 111–231.

Mehlberg, Henryk. "Physical Laws and Time's Arrow." In *Current Issues in the Philosophy of Science*, edited by Herbert Feigl and Grover Maxwell, 105–138. New York: Holt, Rinehart, and Winston, 1961.

Mellor, D. H. *The Facts of Causation*. New York: Routledge, 1995.

Mellor, D. H. *Real Time*. New York: Cambridge University Press, 1981.

Mellor, D. H. *Real Time II*. New York: Routledge, 1998.

Oaklander, L. Nathan. *The Ontology of Time*. Amherst, NY: Prometheus, 2004.

Oddie, Graham. "Backwards Causation and the Permanence of the Past." *Synthese* 85 (1990): 71–93.

Popper, Karl. "The Arrow of Time." *Nature* 177 (1956): 538.

Price, Huw. *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*. New York: Oxford University Press, 1996.

Prior, Arthur N. *Past, Present, and Future*. Oxford, U.K.: Clarendon Press, 1967.

Prior, Arthur N. *Time and Tense*. Oxford, U.K.: Clarendon Press, 1968.

Reichenbach, Hans. *The Direction of Time*. Berkeley: University of California Press, 1956.

Robb, Alfred A. *The Absolute Relations of Time and Space*. Cambridge, U.K.: Cambridge University Press, 1921.

Robb, Alfred A. *A Theory of Time and Space*. Cambridge, U.K.: Cambridge University Press, 1914.

Sellars, Wilfrid. "Time and the World Order." In *Minnesota Studies in the Philosophy of Science*. Vol. 3, edited by Herbert Feigl and Grover Maxwell, 527–616. Minneapolis: University of Minnesota Press, 1962.

Shoemaker, Sydney. "Time without Change." *Journal of Philosophy* 66 (1969): 363–381.

Sklar, Lawrence. *Space, Time, and Spacetime*. Berkeley: University of California Press, 1974.

Smart, J. J. C. "Causal Theories of Time." In *Basic Issues in the Philosophy of Time*, edited by Eugene Freeman and Wilfrid Sellars, 61–77. LaSalle, IL: Open Court, 1971.

Smith, Quentin. *Language and Time*. New York: Oxford University Press, 1993.

Tooley, Michael. *Causation: A Realist Approach*. New York: Oxford University Press, 1987.

Tooley, Michael. *Time, Tense, and Causation*. Oxford, U.K.: Clarendon Press, 1997.

Von Bretzel, Philip. "Concerning a Probabilistic Theory of Causation Adequate for the Causal Theory of Time." *Synthese* 35 (1977): 173–190.

Michael Tooley (2005)

CAUSAL CLOSURE OF THE PHYSICAL DOMAIN

The thesis that physics is causally closed asserts that:

> Every physical effect has a sufficient immediate physical cause, insofar as it has a sufficient immediate cause at all.

If this thesis is true, it distinguishes physics from all other subject domains. The biological realm is not causally closed, for example, because biological effects often have nonbiological causes, as when the impact of a meteorite precipitated the extinction of the dinosaurs. Again, meteorology is not causally closed: The burning of carbon fuels—a nonmeteorological event—is causing global warming. Nor, importantly, is the mental realm causally closed: A mental pain can be caused by sitting on a physical thumbtack, and a train of thought can be interrupted by a loud noise.

Physics, by contrast, does seem to be causally closed. If one considers any physical effect, then there will arguably always be some prior physical cause: People expect to be able to account for physical effects without leaving the physical realm itself. In particular, this seems to hold even for physical effects that take place within the bodies of conscious beings. When the muscle fibers in a person's arm contract, this is presumably due to electrochemical activity in the nerves, which is due to prior physical activity in the person's motor cortex, and so on. In principle, it would seem possible to account for this entire sequence solely in terms of the resources offered by physics itself, and without making any essential appeal to any other subject matter.

At first pass, the causal closure of physics is solely a claim about how things go within physics itself. It does not assert that everything is physical, only that everything physical that has a cause has a physical cause. As such, it does not rule out realms of reality that are distinct from the physical realm. It is entirely consistent with the causal closure of physics itself; there should be nonphysical realms that operate independently. The closure of the physical says only that *within the physical realm* every physical effect has a physical cause.

Even so, the causal closure of the physical does give rise to a powerful argument for reducing many prima facie nonphysical realms to physics: It suggests that *anything that has a causal impact on the physical realm* must itself be physical. The reason is that the causal closure of the physical seems to leave no room for anything nonphysical to make a causal difference to the physical realm, because it specifies that every physical effect already has a physical cause.

Intuitively, of course, people take it that many prima facie nonphysical events, such as biological, meteorological, and mental events, do have physical effects. A burrowing animal can dig a hole in the ground; a hurricane can destroy houses; one's current thoughts can give rise to patterns on a computer screen. However, the causal closure of the physical says that these effects already have physical causes. So it seems that the only way to respect the causal efficacy of realms such as the biological, meteorological, and mental is to conclude that they are not distinct from the physical after all. If one wants to maintain that the animal's burrowing, the hurricane, and a person's thoughts have physical effects, then apparently there is a need to identify these processes with the physical causes that their physical effects are already known to have.

Note that this line of reasoning only argues for a reduction to physics of those realms that *do* have physical effects. The causal closure of the physical provides no argument against the possibility of nonphysical realms that lack any physical effects. For example, it is arguable that mathematical, moral, and modal facts have no physical effects. If this is right, then the causal closure of the physical offers no reason to collapse these realms into the physical. (Of course, there may be other arguments against the possibility of such nonphysical realms of reality, such as their epistemological inaccessibility, but that is a different matter.)

The remainder of this entry contains three sections: First, a discussion of the evidence for the causal closure of physics from a historical perspective; second, a consideration of how the thesis can be made properly precise; and finally, an examination of the details of the argument that causal closure implies physicalism about the mental and similar realms.

A HISTORICAL PERSPECTIVE ON THE EVIDENCE FOR THE CAUSAL CLOSURE OF PHYSICS

Why should one believe the causal closure of physics (which for the moment shall be regarded as the simple claim that every physical effect has a sufficient physical cause)? If this thesis is true, it is not an a priori matter, but something that follows from the findings of science. But exactly which findings? What part of science, if any, argues that physical is causally closed?

At first sight it may seem that causal closure follows from the presence of conservation laws in physics: If there are laws specifying that important physical quantities stay constant over time, does not this show that the later values of physical quantities must be determined by earlier values? However, it depends what conservation laws one has. Not just any set of physical conservation laws rule out the possibility of nonphysical causes for physical effects.

Thus consider Descartes's early seventeenth-century physics. This was based on the conservation of *amount of motion*, which Descartes took to be the product of the masses of all bodies by their scalar *speeds*. (So amount of motion is different from momentum, which is the product of mass by vectorial *velocity*: A car going round a bend at a constant speed conserves amount of motion but not momentum.) As Leibniz observed, Descartes's conservation of amount of motion alone leaves plenty of room for nonphysical causes to intrude on the physical realm. In particular, if mental causes (operating in the pineal gland?) cause particles of matter to change their direction (but not their speed), this would not in any way violate the conservation of amount of motion.

Descartes's physics might allow an independent mind to affect the brain, but Descartes's physics is wrong, as Leibniz further observed. Leibniz himself replaced Descartes's law of the conservation of "motion" with the two modern laws of conservation of (vectorial) momentum and of (scalar) kinetic energy, and thereby arrived at what are now regarded as the correct laws governing impacts. Leibniz's physics, unlike Descartes's, did indeed imply that the later values of all physical quantities are determined by their earlier values, and therewith the causal closure of the physical. However, Leibniz did not draw the modern physicalist conclusion that the mind must therefore be identical to the brain. Because it seemed incontrovertible to him that mind and brain must be ontologically separate, he instead inferred from the causal closure of physical that the mind in fact has no causal impact on the physical world. (It only appears to do so because of the "preestablished harmony" with which God has arranged both the mental and physical worlds.)

Whereas Leibniz's physics implies the causal closure of the physical, this is not true of the Newtonian system of physics that replaced it at the end of the seventeenth century. The crucial difference is that, where Leibniz upheld the central principle of the "mechanical philosophy" and maintained that all changes of velocity are due to impacts between material particles, Newton allowed that accelerations can also be caused by disembodied forces, such as the force of gravity. Moreover, Newton's system was open-ended about the range of different forces that existed. In addition to gravity, Newton and his followers came to recognize magnetic forces, chemical forces, and forces of adhesion—and indeed vital and mental forces, which arose specifically in living bodies and sentient beings. If one counts vital and mental forces as nonphysical (and this point will be revisited in the next section), then the admission of such forces undermines the causal closure of the physical. For it means that physical effects, in the form of accelerations of particles of matter, will sometimes be due to the operation of nonphysical vital or mental causes.

Newtonian physics has its own conservation principles, but unlike Leibniz's they do not uphold the causal closure of the physical. Crucially, Newton's physics differs from Leibniz's in the way the conservation of energy must be understood. The existence of Newtonian forces means that Leibniz's conservation of kinetic energy no longer holds true; for example, two bodies receding from each other will slow down due to their mutual gravitational attraction, and so lose kinetic energy. Newtonian conservation additionally needs the notion of *potential* energy: the latent energy stored when bodies are "in tension" in force fields, as when two receding gravitating bodies cease to move apart and are about to accelerate together again. The notion of potential energy was not prominent in early Newtonian physics, but by the middle of the nineteenth century physicists concluded that all forces operated so as to conserve the *sum* of potential and kinetic energy—any loss of kinetic energy would mean a rise in potential energy, and vice versa.

This emergence of the modern version of the "conservation of energy" placed strong restrictions on what kinds of forces can exist, but it by no means ruled out vital and mental forces. Provided that the fields of these forces stored in latent form any losses of kinetic energy they occasioned (consider by way of comparison the notion of "nervous energy"), their presence would be perfectly consistent with the conservation of kinetic plus potential energy. True, the conservation of kinetic plus potential energy did apparently imply that all forces must be governed by deterministic force laws (otherwise what would ensure that they always paid back any kinetic energy they borrowed?), and this greatly exercised many Victorian thinkers, especially given that nothing in early Newtonian physics had ruled out spontaneously arising mental forces. But, even so, the Newtonian conservation

of energy did not stop *deterministic* vital and mental forces affecting the physical realm.

Nevertheless, during the late nineteenth and twentieth centuries an increasing number of scientists have come to doubt the existence of vital and mental forces. The most significant evidence seems to have come from physiology and molecular biology, rather than from physics itself. During this period a great deal has come to be known about the workings of biological systems (including brains), and there has been no indication that anything other than basic physical forces are needed to account for their operation. In particular the twentieth century has seen an explosion of knowledge about processes occurring within cells, and here too there is no evidence that this involves anything other than familiar physical chemistry. The result has been that the overwhelming majority of scientists now reject vital and mental forces, and accept the causal closure of the physical realm.

THE CAUSAL CLOSURE THESIS REFINED

Much recent discussion of the causal closure thesis has revolved around the question of exactly how "physical" should be understood in the claim that every physical effect has a physical cause. As Carl Hempel originally observed, advocates of the causal closure thesis seem to face a dilemma. On the one hand, they can equate "physical" with the category of phenomena recognized by *current* physical theory. But then it seems implausible that "physics" in this sense is closed; past form suggests strongly that physics will in time come to posit various new fundamental causal categories. Alternatively, advocates of causal closure might wish to equate "physical" with the ontology of some ideal *future* physics. But then it is hard to see how the causal closure of the "physical" could have any current philosophical significance, given that people are as yet ignorant of exactly what this "physical" includes.

However, this dilemma is by no means inescapable. True, neither current physics nor ideal future physics gives us a suitable notion of "physics" for framing the causal closure thesis. But this does not mean there are not other suitable notions of "physics." Indeed there are arguably a number of *different* ways of understanding "physics" that will yield a well-evidenced and contentful causal closure thesis.

For a start, one could simply define *physical* as "neither essentially mental nor biological." This understanding of "physical" was in effect assumed at the end of the last section, in the argument that the nonexistence of vital or mental forces establishes the causal closure of physics. Note that nothing in that argument assumed a definitive list of fundamental physical categories; rather the thought was simply that this list would *not* include any sui generis mental or vital entities. This is a relatively inclusive understanding of "physical"; it counts as a "physical" cause anything that is not mental or vital, and to this extent renders the causal closure of the physical a relatively weak thesis. But even so it remains a thesis of much philosophical interest, because it still argues that any mental or vital causes of physical effects must be identical to causes that can be identified without using mental or vital categories.

A rather stronger reading of "physical" would take it to cover any categories of the *same general kind as are recognized by current physical theory*. Now the list of fundamental "physical" categories will be taken to include not just anything nonvital or nonmental, but more specifically only items that display the same kind of spatiotemporal pervasiveness and simple mathematical characterizability as those assumed in contemporary physics. Again, there seems good reason to suppose that "physics" in this sense is casually closed, and therefore that anything that in this sense has "physical" effects must itself be "physical."

Finally, and even more specifically, there is the option of equating "physical" with *microscopic*. Modern physical theory characteristically operates at a level of microscopic spatiotemporal detail. Correspondingly, it is plausible that every microscopic effect can be accounted for by (a combination) of microscopic causes. This version of the causal closure thesis thus argues that anything that has microscopic effects must itself be identical to (a combination) of microscopic causes.

The remainer of this section deals with some complications in the formulation of the causal closure thesis. In the version of the thesis at the beginning of this article, every physical effect was referred to as having a "sufficient immediate physical cause," rather than just having "a physical cause." This was to ensure that the physical realm is genuinely causally closed. The specification that the physical cause be "sufficient" is needed to ensure that it causes the physical effect by itself, and not solely in virtue of its conjunction with some sui generis nonphysical cause—such a mixed cause would obviously violate the causal closure of physics. Again, the requirement that the physical cause be "immediate" is needed to ensure that it not produce the physical effect only via some intermedi-

ary nonphysical cause—such nonphysical intermediaries would again violate the causal closure of physics.

The earlier formulation of the causal closure thesis also specifies that every physical effect has a sufficient immediate physical cause "insofar as it has a sufficient immediate cause at all." The reason for this latter qualification is to accommodate the indeterminism of modern quantum mechanics, which states that certain physical effects are random, without any sufficient determining cause. It remains the case, however, that according to quantum mechanics these random physical effects still have their *probabilities* fixed by sufficient immediate physical causes. And this in itself will sustain the argument that anything that affects the physical realm must itself be physical. At first sight it may seem that quantum indeterminism creates room for nonphysical causes (determinations of the will, perhaps) to exert a downward influence on the physical realm, by influencing whether or not certain random physical events occur. But this in itself would violate the causal closure of the physical, understood now as including the claim that the *probabilities* of underdetermined quantum physical events are fixed by sufficient immediate physical causes. For, if a nonphysical cause influences whether or not random physical events occur, it must presumably make a difference to the probabilities of those events, and this itself will contradict the thesis that those probabilities are already fixed by sufficient physical causes.

THE ARGUMENT FROM CAUSAL CLOSURE TO PHYSICALISM

What follows now is a closer look at the argument that moves from the causal closure of the physical to the conclusion that anything with a physical effect must itself be physical. The focus will be on the case of *mental* causes of physical effects, but most of the points made will apply to items with physical effects generally.

Recall the point that the argument gets no grip on realms that have no physical effects. As mentioned earlier, Leibniz used this point to evade physicalism about the mental by holding that the mental and physical realms are causally insulated from each other, albeit unfolding in "preestablished harmony." Contemporary philosophers who share Leibniz's conviction that mental states cannot possibly be physical tend to adopt a somewhat different ploy. Instead of denying any causal contact between conscious mind and brain, they allow that brain processes cause conscious mental effects but deny that these conscious states then have any converse influence on the physical realm. The contemporary philosophers of mind Frank Jackson and David Chalmers have both argued in favor of this "epiphenomenalist" position (Jackson 1982, Chalmers 1996). By viewing conscious states as "causal danglers" that exert no independent influence on the physical realm, they avoid any conflict with the thought that the causal closure leaves no room for anything nonphysical to make a difference to physical effects.

Perhaps there is another loophole in the argument from causal closure. In effect, this argument holds that a nonphysical mind cannot have physical effects because then those effects would have too many sufficient causes—both a nonphysical mental cause and the physical cause guaranteed by causal closure. However, such overdetermination of effects by two sufficient causes is not unknown. Imagine a case of a man who is simultaneously shot and struck by lightning, where either cause would have sufficed for his death on its own. Why should the physical effects of mental causes not similarly be overdetermined by two independent causes?

However, it is not clear that this is a good comparison. Overdetermination by distinct causes occasionally occurs by chance. But if a nonphysical mind has physical effects, then causal closure means that overdetermination of those effects will be routine. This calls for some explanation of why the two independent causes—mental and physical—should always be found together. If the two causes really are distinct, then will not some mechanism be needed to ensure that a sufficient physical cause is in place whenever a mental cause has a physical effect? However, no plausible such mechanism suggests itself.

So the possibility of overdetermination by both physical and mental causes does not support a distinct mental realm in the face of the causal closure of physics. However, there is another sense in which the causal closure of physics does leave it open that all behavioral effects may have both a physical cause and a nonidentical mental cause.

It has been a common theme in much recent physicalist philosophy of mind that mental properties are not *type-identical* to physical properties: many physicalist philosophers of mind are persuaded that, because beings with different physical constitutions can share mental properties such as pain, mental properties must be functional properties that are variably (or multiply) realized by physical properties, or disjunctions of physical properties, or some other kind of property that is metaphysically fixed by (supervenes on) physical properties, but not strictly identifiable with them.

Now, to the extent that causes involve properties, this denial of type-identity for mental and physical properties means that the mental and physical causes of behavioral effects cannot be strictly identical. However, this kind of double causation does not amount to the kind of overdetermination by distinct causes that was argued to be unacceptable above. As long as mental causes supervene metaphysically on physical causes, they are not fully distinct from them, and there is already a built-in explanation for why there should always also be a physical cause (as required by the causal closure of the physical) whenever a mental cause produces a behavioral effect. The denial of type identity creates some space between mental and physical causes, but not so much as to render it mysterious that they are always found hand in hand.

See also Causation: Philosophy of Science; Physicalism.

Bibliography

Chalmers, D. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

Jackson, F. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Kim, J. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

Papineau, D. *Thinking about Consciousness*. Oxford: Oxford University Press, 1992.

David Papineau (2005)

CAUSAL OR CONDITIONAL OR EXPLANATORY-RELATION ACCOUNTS

Edmund Gettier attacked the traditional analysis of knowledge by showing that inferring a true belief from a false but justified belief produces a justified true belief that does not qualify as knowledge. Subsequent analyses of knowledge were motivated in large part by the wish to avoid examples of the type Gettier used. One way to do so is to insist that a belief must be connected in some proper way to the fact that makes it true in order for it to count as knowledge. In Gettier's examples beliefs are only accidentally true since there are no proper connections between them and the facts that make them true. Analyses that require such connections may either retain or drop the justification condition from the traditional analysis. Without it they are thoroughly externalist analyses since they require only that a belief be externally connected with the fact that makes it true, not that the subject be able to specify this connection.

One intuitive way to specify the proper connection is to say that it is "causal": The fact that makes a belief true must help cause the belief in the subject if the subject is to have knowledge. When this causal relation holds, the truth of the belief is nonaccidental. The causal analysis of knowledge therefore excludes standard Gettier-type cases, but it seems on reflection to be both too weak and too strong: too strong in that knowledge of universal propositions, mathematical truths, and logical connections seems to be ruled out if these cannot enter into causal relations; too weak in allowing knowledge when a subject cannot distinguish a fact that causes her belief from relevant alternatives. Suppose, for example, that a subject *S* cannot tell red expanses from green ones but believes that there is a red expanse before her whenever either a red or a green expanse is there. Then, on an occasion in which a red expanse is before *S* the usual sort of perceptual causal connection will hold, but knowledge that the expanse is red will be lacking.

A different way to specify the necessary connection that handles the sort of case just cited is provided by the "conditional" account. According to this account, *S* knows that *p* only if *S* would not believe that *p* if *p* were not true. In close possible worlds in which *p* is not true, it must be the case that *S* does not believe it. This rules out the case of the red and green expanses since, in a close world in which the expanse is not red but green, *S* continues to believe it is red. A further condition required by this account is that in close worlds in which *p* continues to be true but other things change, *S* continues to believe that *p*.

The conditional account handles both Gettier's cases and those that require the distinction of relevant alternatives. But once again there are examples that seem to show it both too weak and too strong. That the first condition is too strong can be shown by a variation on the color expanse example. Suppose that *S* cannot tell red from green but is very good at detecting blue. Then, on the basis of seeing a blue expanse *S* can come to know that there is not a green expanse before her. But if this proposition were false (if there were a green expanse before her), she would still believe it true (she would think she was seeing red). That the second condition is too strong seems clear from the case of a very old person whose mental capacities are still intact but soon will fail him. That there are close worlds in which he does not continue to believe as he does now by exercising those

capacities does not mean that he cannot know various facts now through their exercise.

That these conditions are too weak can perhaps be shown by cases in which someone intentionally induces a Gettier-type belief in *S*. In this case, if the belief were not true, it would not have been induced in *S*, and yet *S* does not know. Such a case might or might not be ruled out by the second condition, depending on how it is specified and on how the second condition is interpreted. But there are other cases that seem more certainly to indicate that the conditions are too weak. If *S* steadfastly believes every mathematical proposition that she entertains, then the conditions will be met, but she will not know all the true mathematical propositions that she entertains.

An analysis of knowledge should not only accommodate various intuitions regarding examples; it should also be useful to the normative epistemologist in reconstructing the structure of knowledge and addressing skeptical challenges. The conditional account, as interpreted by its main proponent, Robert Nozick (1981), has interesting implications regarding skepticism. According to it, I can know various ordinary perceptual truths, such as that I am seated before a fire, even though I cannot know that there is no Cartesian demon always deceiving me. This is because in the closest possible worlds in which I am not before the fire, I do not believe that I am (I am somewhere else with different perceptual evidence). But in the closest world in which there is a Cartesian demon, I do not believe there is one (since all my perceptual evidence remains the same). These implications are welcome to Nozick but are troubling to other philosophers. My knowledge of being before the fire depends on the demon world not being among the closest in which I am not before the fire. But, according to the conditional account, I cannot know that this last clause is true. Hence, I cannot show that my knowledge that I sit before the fire is actual, as opposed to merely being possible, and it seems that I ultimately lack grounds for being convinced that this is so. Furthermore, implications regarding more specific claims to knowledge and skeptical possibilities are counterintuitive as well. For example, according to this account I cannot know that my son is not a robot brilliantly constructed by aliens, although I can know that I do not have a brilliantly constructed robot son.

A third way of specifying the required connection that makes beliefs true is to describe it as "explanatory." If *S* knows that *p*, then the fact that *p* must help to explain *S*'s belief. To see whether this account handles the sorts of cases cited, we would need to define the notion of explanation being used here. One way to do so is in terms of a certain notion of probability: Roughly, *p* explains *q* if the probability of *q* given *p* is higher than the probability of *q* in the relevant reference class (reflecting relevant alternatives); put another way, if the ratio of (close) possible worlds in which *q* is true is higher in the worlds in which *p* obtains than in the relevant contrasting set of worlds. Given this interpretation, the analysis handles the perceptual discrimination case. In it *S* does not know there is a red expanse before him because its being red does not raise the probability of his belief that it is relative to those possible worlds in which this belief is based on its being green. The analysis also allows knowledge in the variation that defeats the conditional account. In it *S* knows that there is not a green expanse before her since the fact that the expanse is not green (i.e., it is blue) explains her belief that it is not green. Since the account must allow explanatory chains, it can be interpreted so as to include knowledge of mathematical propositions, which do not enter into causal relations. In the usual case in which *S* has mathematical knowledge that *p* her belief must be explanatorily linked to *p* via some proof. The truth of *p* makes a proof possible, and the ratio of close worlds in which *S* believes *p* must be higher in worlds in which there is a proof than in the overall set of worlds.

The explanatory account needs to be filled out further if it is to accommodate cases involving intentionally produced beliefs resembling Gettier's examples since in such cases the fact that *p* helps to explain why the belief that *p* is induced in *S*. As an externalist account, it would also need to provide defense for the claim that *S* can know that *p* even when, from his point of view, he has no good reasons for believing *p*. The analysis does suggest an approach to answering the skeptic different from that suggested by the conditional account. A proponent of this analysis would answer the skeptic by showing that nonskeptical theses provide better explanations of our ordinary beliefs than do skeptical theses.

See also Epistemology; Nozick, Robert.

Bibliography

Goldman, Alan H. *Empirical Knowledge*. Berkeley, CA.: University of California Press, 1988.

Goldman, Alvin I. "A Causal Theory of Knowing." *Journal of Philosophy* 64 (1967): 357–372.

Goldman, Alvin I. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73 (1976): 771–791.

Klein, Peter. "Knowledge, Causality, and Defeasibility." *Journal of Philosophy* 73 (1976): 792–812.

Nozick, Robert. *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.

Sainsbury, R. M. "Easy Possibilities." *Philosophy and Phenomenological Research* 57 (4) (1997): 907–919.

Alan H. Goldman (1996)
Bibliography updated by Benjamin Fiedor (2005)

CAUSATION: METAPHYSICAL ISSUES

Causal concepts have surely been present from the time that language began, since the vast majority of action verbs involve the idea of causally affecting something. Thus, in the case of transitive verbs of action, there is the idea of causally affecting something external to one—one finds food, builds a shelter, sows seed, catches fish, and so on—while in the case of intransitive verbs, or at least those describing physical actions, it is very plausible that they involve the idea of causally affecting one's own body—as one walks, runs, jumps, hunts, and so on.

It was not long after the very beginning of philosophy in ancient Greece that serious reflection concerning the nature of causation arose, with Aristotle's famous discussion of causation in Book 2 of his *Physics*. The result was Aristotle's doctrine of four types (or, perhaps, aspects) of causes—material, formal, efficient, and final—an account that was immensely influential for about two thousand years.

What was not realized at any point during this time, however—perhaps because of the sense of familiarity with the idea of causation occasioned by the almost ubiquitous presence of causal concepts in even the most rudimentary parts of language—is that the concept of causation gives rise to very serious, puzzling, and difficult philosophical questions. Thus it was only many centuries after Aristotle, with David Hume and his famous discussions of the relation of cause and effect (1739–1740 and 1748), that philosophers realized that the idea of causation was by no means simple and straightforward.

Why did Hume see what so many thoughtful philosophers before him had not? The reason, it would seem, was that Hume held—as did the other British empiricists, John Locke and Bishop (George) Berkeley—that while some concepts can be analyzed in terms of other concepts, in the end analysis must terminate in ideas that apply to things in virtue of objects' having properties and standing in relations that can be immediately given in experience. Hume therefore asked whether the relation of causation was one that could be given in immediate experience. His conclusion was that it could not. The question for Hume, accordingly, was how the concept of causation could be analyzed in terms of ideas that do pick out properties and relations that are given in experience, and once this question was in view, Hume was able to show that arriving at a satisfactory answer was a very difficult matter.

FUNDAMENTAL ISSUES AND ALTERNATIVE VIEWS

One of the central issues in the philosophy of causation concerns, then, this Humean problem: Is the concept of causation basic and unanalyzable, or, on the contrary, does it stand in need of analysis? If it does need to be analyzed, how can this be done?

Many different answers have been offered to these questions. But the various approaches can be divided up into four general types: direct realism, Humean reductionism, non-Humean reductionism, and indirect, or theoretical-term, realism.

This fourfold division, in turn, rests upon the following three distinctions: first, that between reductionism and realism; second, that between Humean and non-Humean states of affairs; and, third, that between states that are immediately observable and those that are not. Let us, then, consider each of these distinctions in turn, starting with that between reductionism and realism.

REALISM VERSUS REDUCTIONISM. The realism-versus-reductionism distinction in this area arises in connection with both causal laws, and causal relations between states of affairs, and gives rise to a number of related theses. In the case of causal relations between states of affairs, a thesis that is essential to reductionism is this:

Basic Reductionism with respect to causal relations. Any two worlds that agree both with respect to all of the non-causal properties of, and relations between, particulars, and with respect to all causal laws, must also agree with respect to all of the causal relations between states of affairs. Causal relations are, then, logically supervenient upon the totality of instances of non-causal properties and relations, together with causal laws.

But while this thesis is an essential part of a reductionist view of causation, it is not sufficient. The reason is that this thesis can be combined with a view of causal laws according to which they obtain in virtue of atomic, and therefore irreducible, facts. What is needed, then, is a reductionist thesis concerning causal laws, and here there are two important possibilities:

Strong Reductionism with respect to causal laws. Any two worlds that agree with respect to all of the non-causal properties of, and relations between, particulars, must also agree with respect to causal laws. Causal laws are, then, logically supervenient upon the totality of instances of non-causal properties and relations.

Moderate Reductionism with respect to causal laws. Any two worlds that agree both with respect to all of the non-causal properties of, and relations between, particulars, and with respect to all laws of nature, must also agree with respect to causal laws. Causal laws are, then, logically supervenient upon the totality of instances of non-causal properties and relations, together with laws of nature.

What lies behind this strong reductionism versus moderate reductionism distinction? The answer is that while most philosophers who are reductionists with regard to causation tend to identify laws of nature with certain cosmic regularities, it is possible to be a reductionist with regard to causation while holding that laws are more than certain cosmic regularities: One might hold, for example, that laws of nature are second-order relations between universals. Such a reductionist would reject Strong Reductionism with regard to causal laws, while accepting Moderate Reductionism.

Each of these two reductionist theses concerning causal laws then entails, in conjunction with the Basic Reductionist thesis concerning causal relations, a corresponding thesis concerning causal relations between states of affairs:

Strong Reductionism with respect to causal relations. Any two worlds that agree with respect to all of the non-causal properties of, and relations between, particulars, must also agree with respect to all of the causal relations between states of affairs. Causal relations are, in short, logically supervenient upon the totality of instances of non-causal properties and relations.

Moderate Reductionism with respect to causal relations. Any two worlds that agree both with respect to all of the non-causal properties of, and relations between, particulars, and with respect to all laws of nature, must also agree with respect to all of the causal relations between states of affairs. Causal relations are, then, logically supervenient upon the totality of instances of non-causal properties and relations, together with laws of nature.

To be a reductionist with regard to causation, then, is to accept the Basic Reductionist thesis with respect to causal relations, and either the Strong or the Moderate Reductionist thesis with respect to causal laws. This then commits one either to the Strong Reductionist thesis or the Moderate Reductionist thesis with respect to causal relations.

A realist with regard to causation, correspondingly, is one who rejects either the Basic Reductionist thesis concerning causal relations, or else both the Strong and the Moderate Reductionist theses with regard to causal laws, or all of these.

HUMEAN VERSUS NON-HUMEAN REDUCTIONISM. In addition to the gulf between reductionism and realism, there are also very important divides within both reductionism and realism. In the case of reductionism, the crucial division involves a distinction between what may be called Humean and non-Humean states of affairs. So let us now turn to that distinction.

Different authors offer different characterizations of what a Humean state of affairs is. The basic idea, however, is that Humean states of affairs are ones that consist of particulars having properties and standing in relations, where the properties and relations in question are, in some sense, immediately observable. The idea of being immediately observable can then be interpreted in different ways. A very restrictive interpretation would be one where immediate observation is equated with direct acquaintance, so that only properties and relations that are the objects of Hume's simple ideas—that is, properties and relations that can be immediately given in experience—are classified as immediately observable. Alternatively, one could construe the idea of immediate observation more broadly, so that any properties and relations that can be directly or noninferentially perceived would count as immediately observable.

What would be an example of a non-Humean state of affairs? One type would be any state of affairs that involves a dispositional property or power, since even if, for example, one sees something in the process of dissolving in water, an inference is involved if one is to arrive at the conclusion that the object is such that it is disposed to dissolve when it is in water, since its dissolving on the occasion in question could be a pure accident, or could be caused entirely by some external force, rather than being due to an intrinsic property of the object itself. So an inference is involved, and therefore the water-solubility of an object cannot be an object of direct perception.

Some twentieth-century approaches to causation attempt to analyze causation in terms of powers and propensities. Such approaches are reductionist, but not of a Humean sort.

DIRECT VERSUS INDIRECT REALISM WITH REGARD TO CAUSATION. Realists with regard to causation either reject the Basic Reductionist thesis concerning causal relations, or else both the Strong and the Moderate Reductionist theses concerning causal laws. But there is a crucial divide within realist approaches, and it concerns the question of whether causal states of affairs are immediately observable. According to direct realism, some causal states of affairs are immediately observable; according to indirect, or theoretical-term realism, no causal states of affairs are immediately observable.

What causal states of affairs are directly observable, according to a direct realist approach to causation? Since it is not at all plausible that one can be directly acquainted with causal laws, the relevant states of affairs must consist of causal relations between states of affairs. Thus direct realism can be defined as a version of realism that claims that the relation of causation is immediately given in experience.

Indirect, or theoretical-term realism rejects this claim, maintaining either that the relation of causation is itself an irreducible, theoretical relation, or, alternatively, that causal laws are irreducible, theoretical states of affairs, or both. Either way, then, the relation of causation is not directly observable.

DIRECT REALISM

We can now turn to a consideration of the four general types of approaches to causation, beginning with direct realism. This view of causation involves four main theses: first, that the relation of causation is directly observable; second, that that relation is not reducible to non-causal properties and/or relations; third, that the relation of causation is also not reducible to non-causal properties and/or relations together with causal laws—since such a reduction would entail that one could not be directly acquainted with the relation of causation; fourth, that the concept of the relation of causation is analytically basic.

A number of philosophers have claimed that the relation of causation is observable, including David Armstrong (1997), Elizabeth Anscombe (1971), and Evan Fales (1990). Thus Anscombe argues that one acquires observational knowledge of causal states of affairs when one sees, for example, a stone break a window, or a knife cut through butter, while Fales, who offers the most detailed argument in support of the view that causation is observable, appeals especially to the impression of pressure upon one's body, and to one's introspective awareness of willing, together with the accompanying perception of the event whose occurrence one willed.

Suppose that it is granted that in such cases one does, in some straightforward sense, observe that one event causes another. Does this provide one with a reason for thinking that direct realism is true? For it to do so, one would have to be able to move from the claim that the relation of causation is thus observable to the conclusion that it is not necessary to offer any analysis of the concept of causation, that the latter can be taken as analytically basic. But observational knowledge, in this broad, everyday sense, would not seem to provide adequate grounds for concluding that the relevant concepts are analytically basic. One can, for example, quite properly speak of physicists as seeing electrons when they look into cloud chambers, even though the concept of an electron is certainly not analytically basic. Similarly, the fact, for example, that sodium chloride is observable, and that one can tell by simply looking and tasting that a substance is sodium chloride does not mean that the expression 'sodium chloride' does not stand in need of analysis.

But might it not be argued in response, first, that, one can observe that two events are causally related in precisely the same sense in which one can observe that something is red; second, that the concept of being red is analytically basic, in virtue of the observability of redness; and therefore, third, that the concept of causation must, for parallel reasons, also be analytically basic?

This response is open, however, to the following reply. If a concept is analytically basic, then one can acquire the concept in question only by being in perceptual or introspective contact with *an instance* of the property or relation in question that is picked out by the concept. One could, however, acquire the concept of a physical object's being red in a world where there were no red physical objects: It would suffice if things sometimes looked red, or if one had hallucinations of seeing red things, or experienced red after-images. The concept of a physical object's being red must, therefore, be definable, and cannot be analytically basic.

What is required if a concept is to be analytically basic? The answer that is suggested by the case of the concept of redness is that for a concept to be analytically basic, the property or relation in virtue of which the concept applies to a given thing must be such that that property or relation is immediately given in experience, where a property or relation is immediately given in experience only if, for any two qualitatively indistinguishable experiences, the property must either be given in both or given in neither.

Is the relation of causation immediately given in experience? The answer is that it is not. For given any

experience *E* whatever—be it a perception of external events, an awareness of pressure upon one's body, or an introspective awareness of some mental occurrence, such as an act of willing, or a process of thinking—it is logically possible that appropriate, direct stimulation of the brain might produce an experience, *E**, that was qualitatively indistinguishable from *E*, but which did not involve any causally related elements. So, for example, it might seem to one that one was engaging in a process of deductive reasoning, when, in fact, there was not really any direct connection at all between the thoughts themselves—since all of them were in fact being caused instead by something outside of oneself. Causal relations cannot, therefore, be immediately given in experience in the sense that is required if the concept of causation is to be unanalyzable.

Let us now turn to objections to direct realism. The first has, in effect, just been set out. For if, for any experience in which one is in perceptual or introspective contact with the relation of causation, there could be a qualitatively indistinguishable, hallucinatory experience in which one was not in contact with the relation of causation, it would be possible to acquire the concept of causation without ever being in contact with an instance of that relation. But such experiences are logically possible. So the concept of causation must be analyzable, rather than being analytically basic.

Second, it seems plausible that there is a basic relation of causation that is necessarily irreflexive and asymmetric, even if this is not true of the ancestral of that relation. If either reductionism or theoretical-term realism is correct, one may very well be able to explain the necessary truths in question, since the fact that causal concepts are, on either of those views, analyzable means that those necessary truths may turn out to be analytic. Direct realism, by contrast, in holding that the concept of causation is analytically basic, is barred from offering such an explanation of the asymmetry and irreflexivity of the basic relation of causation. It therefore has to treat these as a matter of synthetic *a priori* truths.

Third, direct realism encounters epistemological problems. Thus, features such as the direction of increase in entropy, or the direction of the transmission of order in non-entropic, irreversible processes, or the direction of open forks, often provide evidence concerning how events are causally connected. In addition, causal beliefs are often established on the basis of statistical information—using methods that, especially within the social sciences, are often very sophisticated. Given an appropriate analysis of the relation of causation, one can show why such features are epistemologically relevant, and why the statistical methods in question can serve to establish causal hypotheses, whereas if causation is a basic, irreducible relation, it is not at all clear how either of these things can be the case.

HUMEAN REDUCTIONISM

Humean reductionist approaches to causation are of three main types: first, accounts that analyze causation in terms of conditions that in the circumstances are nomologically necessary, sufficient, or both; second, accounts in which counterfactual conditionals play the crucial role; and third, accounts based upon probabilistic relations of a Humean sort.

CAUSES AND NOMOLOGICAL CONDITIONS. This first Humean reductionist approach comes in different forms. According to perhaps the most common version, a cause is a condition that is necessary in the circumstances for its effect. To say that event *c* is necessary in the circumstances for event *e* is roughly to say that there is some law, *l*, and some circumstance, *s*, such that the nonoccurrence of *c*, in circumstance *s*, together with law *l*, logically entails the nonoccurrence of *e*.

It may be held instead that a cause is a condition that is sufficient in the circumstances for its effect. To say that event *c* is sufficient in the circumstances for event *e* is to say that there is some law, *l*, and some circumstance, *s*, such that the occurrence of *c*, in circumstance *s*, together with law *l*, logically entails the occurrence of *e*. Finally, it has also been suggested that for one event to cause another is for its occurrence to be both necessary and sufficient in the circumstances for the occurrence of the other event.

What problems do such approaches encounter? Perhaps the most serious difficulty concerns the direction of causation. Suppose, for example, that our world were a Newtonian one, and thus one where the basic laws were time-symmetric. Then the total state of the universe in 1950 would have been both necessary and sufficient not only for the total state in 2050 but also for the total state in 1850. It would therefore follow that events in 1950 had caused both events in 2050 and events in 1850.

Less general objections are also important. First, if a cause is necessary in the circumstances for its effect, this precludes cases of causal preemption, in which event *d* would have caused event *e* were it not for the presence of event *c*, which both caused *e* and prevented *d* from doing so. In such a case *c* is not necessary for *e* since, if *c* had not occurred, *e* would have been caused by *d*. Second, cases of

causal overdetermination are also ruled out. For if both *c* and *d* are causally sufficient to bring about *e*, and both do so, then neither *c* nor *d*was necessary in the circumstances for the occurrence of *e*.

These objections can be avoided if one holds instead that a cause is sufficient in the circumstances for its effect. But then other objections emerge. In particular, it follows that there can be no causal relations if all the laws of nature are probabilistic. This is a serious difficulty, especially given the indeterministic nature of quantum mechanics.

COUNTERFACTUAL CONDITIONAL APPROACHES. A second important reductionist approach attempts to analyze causation using subjunctive conditionals. One way of arriving at this approach is by analyzing causation in terms of necessary or sufficient conditions (or both) but then interpreting the latter, not as nomological connections, as above, but as subjunctive conditionals. Thus one can say that *c* is necessary in the circumstances for *e* if, and only if, had *c* not occurred *e* would not have occurred, and that *c* is sufficient in the circumstances for *e* if, and only if, had *e* not occurred *c* would not have occurred.

John L. Mackie (1965/1993, 1974) took this tack in developing a more sophisticated analysis of causation in terms of necessary and sufficient conditions. Thus, after defining an INUS condition of an event as an insufficient but necessary part of a condition which is itself unnecessary but exclusively sufficient for the event, and then arguing that *c*'s being a cause of *e* can then be analyzed as *c*'s being at least an INUS condition of *e*, Mackie asked how necessary and sufficient conditions should be understood. For general causal statements, Mackie favored a nomological account, but for singular causal statements he argued for an analysis in terms of subjunctive conditionals.

The most fully worked-out subjunctive conditional, or counterfactual approach, however, is that of David Lewis (1973/1986, 1979/1986, 2000). His basic strategy involves analyzing causation using a narrower notion of causal dependence and then analyzing causal dependence counterfactually: (1) an event *c* causes an event *e* if, and only if, there is a chain of causally dependent events linking *e* with *c*; (2) an event *g* is causally dependent upon an event *f* if, and only if, had *f* not occurred *g* would not have occurred.

Causes, so construed, need not be necessary for their effects because counterfactual dependence, and hence causal dependence, are not necessarily transitive. Nevertheless, Lewis's approach is closely related to necessary-condition analyses of causation since the more basic relation of causal dependence is a matter of one event's being counterfactually necessary in the circumstances for another event.

What problems arise for such approaches? One objection involves overdetermination, where two events, *c* and *d*, are followed by an event *e*, and where each of *c* and *d* would have been causally sufficient, on its own, to produce *e*. If it is true, in at least some actual or possible cases of this sort, both that *c* causes *e* and that *d* causes *e*, then one has a counterexample to Lewis's counterfactual analysis.

A second objection involves cases of preemption; that is, cases where there is some event *c* that causes *e*, but where there is also some event *d* that did not cause *e*, but that failed to do so only because the presence of *c* prevented it from doing so.

Until the late twentieth century, the discussion of preemption had focused on cases where one causal process preempts another by blocking the occurrence of some state of affairs in the other process, and a variety of closely related ways of attempting to handle this type of preemption have been advanced, involving such notions as fragility of events, quasi-dependence, continuous processes, minimal-counterfactual sufficiency, and minimal-dependence sets (Lewis 1986, Menzies 1989, McDermott 1995, Ramachandran 1997). But none of these approaches can handle the case of trumping preemption, advanced by Jonathan Schaffer (2000), where one causal process preempts another without preventing the occurrence of any of the states of affairs involved in the other causal process.

Third, there is once again the problem of explaining the direction of causation. One possibility is to define the direction of causation as the direction of time, but neither Mackie nor Lewis favors that approach: both think that backward causation is logically possible. Mackie's main proposal appeals to the direction of irreversible processes involving the transmission of order—such as with outgoing concentric waves produced by a stone hitting a pond—and Lewis advances a somewhat related proposal, in which the direction of counterfactual dependence, and hence causal dependence, is based upon the idea that events in this world have many more effects than they have causes. But the problem with both of these suggestions is that the relevant features are at best contingent ones, and it would seem that, even if the world had neither of these features, it could still contain causally related events.

A final objection, and the most fundamental of all, is concerned with the truth conditions of the counterfactuals that enter into the analysis. One familiar approach to counterfactuals maintains that the truthmakers for counterfactuals concerning events in time involve causal facts (Jackson 1977). Such analyses cannot of course be used in an analysis of causation, on pain of circularity. Accordingly, Lewis formulated his analysis of causation in terms of counterfactuals whose truth conditions are a matter of similarity relations across possible worlds (Stalnaker 1968, Lewis 1973). It can be shown, however, by a variant on an objection advanced by Bennett (1974) and Fine (1975), that this account of counterfactuals does not yield the correct truth-values in all cases (Tooley 2003). Moreover, the same type of counterexample also shows an analysis of causation based on such conditionals will generate the wrong truth-values in the cases in question.

PROBABILISTIC APPROACHES. Among the more significant developments in the philosophy of causation since the time of Hume is the idea, motivated in part by quantum mechanics, that causation is not restricted to deterministic processes. This has led several philosophers to propose that causation itself should be analyzed in probabilistic terms.

The central idea is that causes must make their effects more likely. This idea can, however, be expressed in two rather different ways. The traditional approach, developed by Hans Reichenbach (1956), I. J. Good (1961/1962), and Patrick Suppes (1970), focuses upon types of events and involves the notion of positive statistical relevance. According to this notion, an event of type *C* is positively relevant to an event of type *E* if and only if the conditional probability of an event of type *E*, given an event of type *C*, is greater than the unconditional probability of an event of type *E*. The basic idea, then, is that for events of type *C* to be direct causes of events of type *E*, a necessary condition is that the former be positively relevant to the latter.

But do causes necessarily make their effects more likely? Consider two types of diseases, *A* and *B*, governed by the following laws. First, disease *A* causes death with probability 0.1, while disease *B* causes death with probability 0.8. Second, contracting either disease produces complete immunity to the other. Third, in condition *C*, an individual must contract either disease *A* or disease *B*. (Condition *C* might be a weakening of the immune system.) Finally, assume that individual *m* is in condition *C* and contracts disease *A*, which causes his death. Given these conditions, what if *m*, though in condition *C*, had not contracted disease *A*? Then *m* would have contracted disease *B*. But if so, then *m*'s probability of dying had he not contracted disease *A* would have been 0.8—higher than his probability of dying given that he had contracted disease *A*. So the claim that lies at the heart of probabilistic approaches—that causes necessarily make their effects more likely—cannot be true.

NON-HUMEAN REDUCTIONISM

Traditional probabilistic approaches, in analyzing causation in terms of statistical relations, offered a Humean reductionist account of causation. In the late twentieth century, however, an alternative type of probabilistic approach to causation was suggested, one that involves analyzing causation in terms of propensities, or objective chances. Objective chances, however, do not logically supervene upon the totality of Humean states of affairs, as is shown by the fact, for example, that if atoms of a given type take a certain average time t to undergo radioactive decay, that fact is logically compatible with different objective chances of such atoms' undergoing decay within a given period of time. An analysis of causation that involves objective chances is therefore a reductionist account of a non-Humean sort.

OBJECTIVE CHANCE APPROACHES TO CAUSATION. A number of philosophers —such as Edward Madden and Rom Harré (1975), Nancy Cartwright (1989), and C. B. Martin (1993)—have both advocated an ontology in which irreducible dispositional properties, powers, propensities, chances, and the like, occupy a central place, and maintained that such an ontology is relevant to causation. Often, however, the details have been rather sparse. But a clear account of the basic idea of analyzing causation in terms of objective chances was set out in 1986 both by D. H. Mellor and by David Lewis and then, in the 1990s, Mellor offered a very detailed statement and defense of this general approach in his book *The Facts of Causation* (1995).

Mellor's approach, in brief, is roughly as follows. First, Mellor embraces an ontology involving objective chances, where the latter are ultimate properties of states of affairs, rather than being logically supervenient upon causal laws together with non-dispositional properties, plus relations. Second, Mellor proposes that chances can be defined as properties that satisfy three conditions: (1) The Necessity Condition: if the chance of *P*'s obtaining is equal to one, then *P* is the case; (2) The Evidence Condition: if one's total evidence concerning *P* is that the chance of *P* is equal to *k*, then one's subjective probability

that *P* is the case should be equal to *k*; (3) The Frequency Condition: the chance that *P* is the case is related to the corresponding relative frequency in the limit. Third, chances enter into basic laws of nature. Fourth, Mellor holds that even basic laws of nature need not have instances, thereby rejecting reductionist accounts in favor of a realist view. Fifth, any chance that *P* is the case must be a property of a state of affairs that temporally precedes the time at which *P* exists, or would exist. Finally, and as a very rough approximation, a state of affairs *c* causes a state of affairs *e* if and only if there are numbers *x* and *y* such that (1) the total state of affairs that exists at the time of *c*—including laws of nature—entails that the chance of *e* is *x*, (2) the total state of affairs that would exist at the time of *c*, if *c* did not exist, entails that the chance of *e* is *y*, and (3) *x* is greater than *y*.

This approach to causation is open to three main types of objections. First, this account necessarily involves the Stalnaker-Lewis style of counterfactuals, and, as was noted earlier, such a closest-worlds account of counterfactuals is unsound.

Second, there are a number of objections that can be directed against the view that objective chances are ontologically ultimate properties, one of which is as follows. Imagine that the world is deterministic, that every temporal interval is divisible, and that all causation involves continuous processes. Suppose that *x* at time *t* has an objective chance equal to 1 of being *C* at time $(t + \Delta t)$. Then there are an infinite number of moments between *t* and $(t + \Delta t)$, and for every such moment, *t*, it must be the case either that *x* at time *t* has an objective chance equal to 1 of being *C* at time *t*, or that *x* at time *t* has an objective chance equal to 1 of not being *C* at time *t*. But then, if objective chances are ontologically ultimate, intrinsic properties of things at a time, it follows that *x* at time *t* must have an infinite number of intrinsic properties—indeed, a non-denumerably infinite number of properties.

This view of the nature of objective chances involves, accordingly, a very expansive ontology indeed. By contrast, if objective chances, rather than being ontologically basic, supervene on categorical properties plus causal laws, this infinite set of intrinsic properties of *x*, at time *t* disappears, and all that one may have is a single, intrinsic, categorical property—or a small number of such properties— together with relevant laws of nature.

Third, there are objections to the effect that, even given this view of objective chances, the resulting account of causation is unsound. Here one of the most important is that, just as in the case of attempts to analyze causation in terms of relative frequencies, it can be shown that the crucial claim that a cause raises the probability of its effect remains unsound when one shifts from relative frequencies to objective chances.

INDIRECT, OR THEORETICAL-TERM, REALISM

Direct realism with regard to causation is, as we saw earlier, deeply problematic. There is, however, a very different form of causal realism, according to which causation is a theoretical relation between events. On this view, all knowledge of causal states of affairs is inferential knowledge, and the concept of causation stands in need of analysis. But unlike reductionist accounts, the relevant analysis does not imply that causal states of affairs are logically supervenient upon non-causal states of affairs.

A THEORETICAL-TERM REALIST ACCOUNT OF CAUSATION. This approach to causation involves finding postulates that serve to define implicitly the relation of causation. One suggestion here (Tooley 1990), for example, starts out with postulates for causal laws that say, very roughly, that the *a posteriori* probabilities of effects are a function of the *a priori* probabilities of their causes, whereas, by contrast, the *a posteriori* probabilities of causes are not a function of the *a priori* probabilities of their effects. Then, when one adds the further postulate that causal laws involve the relation of causation, the result is an implicit definition of the relation of causation. That implicit definition can then be converted into an explicit one by using one's preferred approach to the definition of theoretical terms. So, for example, if one adopts a Ramsey/Lewis approach, the relation of causation can be defined as that unique relation between states of affairs that satisfies the relevant open sentences corresponding to the postulates in question.

REALISM OR REDUCTIONISM?

Reductionist approaches to causation are, as we have seen, exposed to a variety of objections. In addition, however, there are general objections that appear to tell against any reductionist approach. Two especially important ones are, first, that the Basic Reductionist Thesis is unsound, and, second, that reductionism cannot provide a satisfactory account of the direction of causation.

SINGULARISM AND CAUSAL LAWS. According to the Basic Reductionist Thesis, causal relations are logically supervenient upon the totality of instances of non-causal properties and relations, together with causal laws. But

this thesis is exposed to a number of objections, such as the following. Assume that indeterministic laws are logically possible and that, in particular, it is a basic law both that an object's acquiring property *P* causes it to acquire either property *Q* or property *R*, but not both, and that an object's acquiring property *S* also causes it to acquire either property *Q* or property *R*, but not both. Suppose now that some object simultaneously acquires both property *P* and property *S* and then immediately acquires both property *Q* and property *R*. The problem now is that, given that the relevant laws are basic, there cannot be any non-causal facts that will determine which causal relations obtain. Did the acquisition of *P* cause the acquisition of *Q*, or did it cause the acquisition of *R*? On a reductionist approach, no answer is possible. Accordingly, causal relations between events cannot be logically supervenient upon causal laws plus non-causal states of affairs.

REDUCTIONISM AND THE DIRECTION OF CAUSATION. What determines the direction of causation? Reductionists have advanced various suggestions, but some arguments seem to show that no reductionist account can work. One such argument appeals to the idea of a very simple world—consisting, say, of a single particle, or of two particles rotating endlessly about one another. Such simple worlds would still involve causation since the identity over time of the particles, for example, requires causal relations between their temporal parts. But since such worlds are time-symmetric, the events in them will not exhibit any non-causal patterns that could provide the basis for a reductionist account of the direction of causation. Accordingly, no reductionist account of the direction of causation can generate the correct answer for all possible worlds. It would seem, then, that only a realist account of causation will do.

See also Anscombe, Gertrude Elizabeth Margaret; A Priori and A Posteriori; Aristotle; Armstrong, David M.; Bennett, Jonathan; Berkeley, George; Cartwright, Nancy; Hume, David; Lewis, David; Locke, John; Mackie, John Leslie; Philosophy of Statistical Mechanics; Reductionism in the Philosophy of Mind; Realism; Reichenbach, Hans; Suppes, Patrick.

Bibliography

Anscombe, G. E. M. *Causality and Determination*. Cambridge, U.K.: Cambridge University Press, 1971.

Aristotle. *Physics*. Book 2.

Armstrong, David M. *A World of States of Affairs*. Cambridge, U.K.: Cambridge University Press, 1997.

Armstrong, David M., and Adrian Heathcote. "Causes and Laws." *Noûs* 25 (1991): 63–73.

Bennett, Jonathan. "Counterfactuals and Possible Worlds." *Canadian Journal of Philosophy* 4 (1974): 381–402.

Cartwright, Nancy. *Nature's Capacities and their Measurement*. Oxford: Clarendon Press, 1989.

Davis, Wayne. "Probabilistic Theories of Causation." In *Probability and Causation: Essays in Honor of Wesley Salmon*, edited by James H. Fetzer, 133–160. Dordrecht: Reidel, 1988.

Dowe, Phil. *Physical Causation*. New York: Cambridge University Press, 2000.

Dowe, Phil, and Paul Noordhof, eds. *Cause and Chance*. London and New York: Routledge, 2004.

Eells, Ellery. *Probabilistic Causality*. Cambridge, U.K.: Cambridge University Press, 1991.

Ehring, Douglas. *Causation and Persistence*. New York: Oxford University Press, 1997.

Fair, David. "Causation and the Flow of Energy." *Erkenntnis* 14 (1979): 219–250.

Fales, Evan. *Causation and Universals*. London and New York: Routledge, 1990.

Fine, Kit. "Critical Notice—*Counterfactuals*." *Mind* 84 (1975): 451–458.

Good. I. J. "A Causal Calculus." Parts 1 and 2. *British Journal for the Philosophy of Science* 11 (1961): 305–318; 12 (1962): 43–51.

Hausman, Daniel M. *Causal Asymmetries*. Cambridge, U.K.: Cambridge University Press, 1998.

Hesslow, Germund. "Two Notes on the Probabilistic Approach to Causation." *Philosophy of Science* 43 (1976): 290–292.

Hume, David. *An Enquiry concerning Human Understanding*. London: 1748.

Hume, David. *A Treatise of Human Nature*. London: 1739–1740.

Jackson, Frank. "A Causal Theory of Counterfactuals." *Australasian Journal of Philosophy* 55 (1977): 3–21.

Lewis, David. "Causation." *Journal of Philosophy* 70 (1973): 556–567. Reprinted, with postscripts, in *Philosophical Papers*. Vol. 2. Oxford: Oxford University Press, 1986.

Lewis, David. "Causation as Influence." *Journal of Philosophy* 97 (4) (2000): 182–197.

Lewis, David. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (1979): 455–476. Reprinted, with postscripts, in *Philosophical Papers*. Vol. 2. Oxford: Oxford University Press, 1986.

Lewis, David. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

Lewis, David. "Postscripts to 'Causation.'" In his Philosophical Papers. Vol. 2, 172–213. Oxford: Oxford University Press, 1986.

Mackie, John L. "Causes and Conditions." *American Philosophical Quarterly* 2 (1965): 245–264. Reprinted in *Causation*, edited by Ernest Sosa and Michael Tooley, 33–55. Oxford: Oxford University Press, 1993.

Mackie, John L. *The Cement of the Universe*. Oxford: Oxford University Press, 1974.

Madden, Edward H., and Rom Harré. *Causal Powers*. Oxford: Blackwell, 1975.

Martin, C. B. "Power for Realists." In *Ontology, Causality and Mind*, edited by John Bacon, Keith Campbell, and Lloyd

Reinhardt. Cambridge, U.K.: Cambridge University Press, 1993.
McDermott, Michael. "Redundant Causation." *British Journal for the Philosophy of Science* 40 (1995): 523–544.
Mellor, D. H. *The Facts of Causation*. London: Routledge, 1995.
Mellor, D. H. "Fixed Past, Unfixed Future." In *Contributions to Philosophy: Michael Dummett*, edited by Barry Taylor, 166–186. The Hague: Nijhoff, 1986.
Menzies, Peter. "Probabilistic Causation and Causal Processes: A Critique of Lewis." *Philosophy of Science* 56 (1989): 642–663.
Ramachandran, Murali. "A Counterfactual Analysis of Causation." *Mind* 106 (1997): 263–277.
Reichenbach, Hans. *The Direction of Time*. Berkeley and Los Angeles: University of California Press, 1956.
Salmon, Wesley C. "Probabilistic Causality." *Pacific Philosophical Quarterly* 61 (1980): 50–74.
Salmon, Wesley C. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press, 1984.
Schaffer, Jonathan. "Trumping Preemption." *Journal of Philosophy* 97 (4) (2000): 165–181.
Sosa, Ernest, and Michael Tooley, eds. *Causation*. Oxford: Oxford University Press, 1993.
Stalnaker, Robert C. "A Theory of Conditionals." In *Studies in Logical Theory, American Philosophical Quarterly*, Monograph 2, edited by Nicholas Rescher, 98–112. Oxford: Blackwell, 1968.
Strawson, Galen. *The Secret Connexion: Causation, Realism, and David Hume*. Oxford: Oxford University Press, 1989.
Suppes, Patrick. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland, 1970.
Tooley, Michael. *Causation: A Realist Approach*. Oxford: Oxford University Press, 1987.
Tooley, Michael. "The Nature of Causation: A Singularist Account." *Canadian Philosophers, Canadian Journal of Philosophy* Suppl. 16 (1990): 271–322.
Tooley, Michael. "Probability and Causation." In *Cause and Chance*, edited by Phil Dowe and Paul Noordhof. London and New York: Routledge, 2004.
Tooley, Michael. "The Stalnaker-Lewis Approach to Counterfactuals." *Journal of Philosophy* 100 (7) (2003): 321–327.
Tooley, Michael. *Time, Tense, and Causation*. Oxford: Oxford University Press, 1997.
Von Bretzel, Philip. "Concerning a Probabilistic Theory of Causation Adequate for the Causal Theory of Time." *Synthèse* 35 (1977): 173–190.
Von Wright, Georg Henrik. *Explanation and Understanding*. Ithaca, NY: Cornell University Press, 1971.

Michael Tooley (1996, 2005)

CAUSATION: PHILOSOPHY OF SCIENCE

In *The Critique of Pure Reason* (first published in 1781), the German philosopher Immanuel Kant maintained that causation was one of the fundamental concepts that rendered the empirical world comprehensible to humans. By the beginning of the twenty-first century, psychology was beginning to show just how pervasive human reasoning concerning cause and effect is. Even young children seem to naturally organize their knowledge of the world according to relations of cause and effect.

It is hardly surprising, then, that causation has been a topic of great interest in philosophy, and that many philosophers have attempted to analyze the relationship between cause and effect. Among the more prominent proposals are the following: Causation consists in the instantiation of exceptionless regularities (Hume 1975, 1999; Mill 1856; Hempel 1965; Mackie 1974); causation is to be understood in terms of relations of probabilistic dependence (Reichenbach 1956, Suppes 1970, Cartwright 1983, Eells 1991); causation is the relation that holds between means and ends (Gasking 1955, von Wright 1975, Woodward 2003); causes are events but for which their effects would not have happened (Lewis 1986); causes and effects are connected by physical processes that are capable of transmitting certain types of properties (Salmon 1984, Dowe 2000).

It often happens, however, that advances in science force people to abandon aspects of their common sense picture of the world. For example, Einstein's theories of relativity have forced people to rethink their conceptions of time, space, matter, and energy. What lessons does science teach about the concept of causation?

RUSSELL'S CHALLENGE

In 1912, the eminent British philosopher Bertrand Russell delivered his paper "On the Notion of Cause" before the Aristotelian Society. In this paper, he claimed that the notion of cause had no place in a scientific worldview:

> All philosophers, of every school, imagine that causation is one of the fundamental axioms or postulates of science, yet, oddly enough, in advanced sciences such as gravitational astronomy, the word "cause" never appears … To me, it seems that … the reason why physics has ceased to look for causes is that, in fact, there are no such things. The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. (p. 1)

Russell was not alone in this view. Other writers of the period, such as Ernst Mach (the German physicist and philosopher of science), Karl Pearson (the father of mod-

ern statistics), and Pierre Duhem (French physicist, as well as historian and philosopher of science), also argued that causation did not belong in the world of science. This view was shared by the logical positivists, a group of philosophers working primarily in Austria and Germany between the World Wars whose ideas shaped much of philosophy of science in the twentieth century. A general suspicion of causal notions also pervaded a number of fields outside of philosophy, such as statistics and psychology.

CAUSATION IN SCIENCE

Despite Russell's remark, it is simply false that the word "cause" (and its cognates) does not appear in the advanced sciences. Russell's claim can be readily refuted by perusing any leading science journal. Admittedly, some uses of the word "cause" and its cognates have specific technical meanings—such as talk of "causal structure" in connection with the general theory of relativity—but frequently enough these words are used in their ordinary English sense. To cite just one example, an issue of *Physical Review Letters* from 2003 contains an article titled "Specific-Heat Anomaly Caused by Ferroelectric Nanoregions in Pb(Mg[sub 1/3]Nb[sub 2/3])O[sub 3] and Pb(Mg[sub 1/3]Ta[sub 2/3])O[sub 3] Relaxors." Moreover, it has become common in physics to classify a variety of phenomena as "effects": there is the "Hall effect," the "Kondo effect," the "Lamb-shift effect," the "Zeeman effect," and so on. But surely "cause and effect" are an inseparable pair: where there are causes, there are effects that are caused by them, and where there are effects, there are causes that cause them.

The person on the street is more likely to encounter causal claims from the medical sciences, such as: "Cholesterol in the bloodstream causes hardened arteries, which in turn causes heart attacks." While the medical sciences may not be as advanced as Russell's example of gravitational astronomy, it is implausible to think that these causal claims are the result of conceptual confusion, or are otherwise scientifically disreputable.

Despite the falsehood of its most provocative claim, however, Russell's paper does succeed in highlighting a number of important and interesting problems about the role of causation in science.

ANTI-FUNDAMENTALISM

Although the advanced sciences have hardly eschewed talk of causation, it is true that the deepest physical principles—such as Newton's three laws of motion, his law of universal gravitation, Maxwell's equations governing the electric and magnetic fields, Schrödinger's equation governing the evolution of quantum systems, and Einstein's field equations relating the distribution of mass-energy in the universe with the structure of space and time—make no mention of causation. All of these principles take the form of mathematical equations and act as constraints on possible states of physical systems (under suitable mathematical characterizations). A given sequence of states may be compatible with, for example, Newton's laws of motion, but nothing in those laws explicitly says that certain states (or aspects of those states) cause others. This suggests that the causal relation is not part of the constitution of the world at the deepest metaphysical level, a view that the historian and philosopher of science John Norton labels "anti-fundamentalism" (Norton 2003). Indeed, the world described by fundamental physics is in many ways at odds with the ordinary picture of a world regimented by cause and effect relationships.

ASYMMETRY

People normally think of causation as both *asymmetric* and *temporally biased*. It is asymmetric in the sense that if *C* is a cause of *E*, then (always? typically?) *E* is not a cause of *C*. This claim must be stated with some care. It may be, for instance, that anxiety is a cause of insomnia, which is in turn a cause of anxiety. But it is one's anxiety on Monday evening that causes insomnia on Monday night, which in turn causes anxiety on Tuesday morning. Monday night's insomnia is not both the cause and the effect of one and the same episode of anxiety. Causation is temporally biased in the sense that causes (always? typically?) occur before their effects in time.

By contrast, the fundamental laws of physics mentioned above are all *time-reversal invariant*. That is, if a particular sequence of states of a physical system is consistent with the laws of physics, then the temporally reversed sequence is also consistent with those laws. The laws of physics do not discriminate between the past and the future in the way that causation does, with two possible exceptions. The first exception involves the statistical laws governing the decay of certain mesons. While these laws exhibit a slight temporal asymmetry, the phenomena in question seem too esoteric to be of much help in understanding the asymmetry of causation.

The second exception is the second law of thermodynamics, which states that the entropy of a closed system can increase but never decrease. Thus a closed system whose entropy is increasing is consistent with the second law, while the temporal reverse of this system is not. The second law of thermodynamics is not, however, a funda-

mental law. The entropy of a physical system is determined by the physical state of the particles that make up the system, as characterized in terms of ordinary physical parameters such as position and momentum. These particles are in turn governed by the time-reversal invariant laws already mentioned. It is thus something of a mystery how the asymmetric second law of thermodynamics can arise from the underlying symmetric dynamics governing the constituents of thermodynamic systems. One prominent view is that the second law of thermodynamics is the result of de facto temporal asymmetries in the boundary conditions of the universe.

There have been a few attempts to ground the asymmetry of causation in the second law of thermodynamics. The basic idea is that the best characterization of our physical universe will include not only the fundamental laws of physics, but also the statement that in the past our universe was in a state of very low entropy—the so-called "past hypothesis."When entertaining various counterfactual suppositions, one conjoins those suppositions with the laws of physics and the past hypothesis to determine what the world would be like if those suppositions were true. Because people hold fixed features of the past, but not of the future, when entertaining contrary-to-fact suppositions, any changes from the actual world introduced in those suppositions will tend to entail significant changes in the future but only insignificant changes in the past. In this way, macroscopic features of the future will counterfactually depend upon what is true in the present, whereas macroscopic features of the past will not. This asymmetric relation of counterfactual dependence can then serve as the basis of an account of causation (such as that of David Lewis in "Causation" [1986]). If this account is correct, then the existence of an asymmetric causal relation is not guaranteed by the laws of physics but is rather the consequence of contingent asymmetries in the boundary conditions of the world.

The best-known attempt to account for causal asymmetry is the common cause principle, first formulated by the German-American Philosopher Hans Reichenbach and presented in his posthumously published book *The Direction of Time* (1956). For Reichenbach, temporal order and causal order are conceptually intertwined. Reichenbach defines causation in terms of probabilities and temporal order, but temporal order is itself defined in terms of asymmetries in probabilities. Let *A* and *B* be two events that are probabilistically *correlated*; in other words, the probability that *A* and *B* will occur together, *P(A & B)*, is greater than the product of the individual probabilities, *P(A)P(B)*. (If the two probabilities are equal, then *A* and *B* are said to be probabilistically *independent*.) An event *C* is said to "screen off" *A* from *B* if it renders them *conditionally* independent; that is, if $P(A \,\&\, B|C) = P(A|C)P(B|C)$. If there is an earlier event *C* that screens off *A* from *B*, but no later event that does so, then the trio *ABC* forms a *conjunctive fork open to the future*. If there is a later screener-off *E*, but no earlier one, then *ABE* is a conjunctive fork open to the *past*. Finally, if there is an earlier and a later screener-off, then that is a *closed* fork. According to Reichenbach, the overwhelming majority of open forks are open to the future, and this probabilistic asymmetry provides the basis for the distinction between the past and the future. Reichenbach further held that if two events *A* and *B* are correlated, and neither is a cause of the other, then there exists a common cause of *A* and *B* in their mutual past that screens *A* off from *B*.

Reichenbach believed that his common cause principle was related to the second law of thermodynamics. Think of *A & B* as one possible state of a physical system, the other possible states being *A & ~B*, *~A & B*, and *~A & ~B*. A probability distribution over these states in which *A* and *B* are correlated contains *information*, in a sense that is made precise within the mathematical field of information theory. From a formal perspective, information is inversely related to entropy. Thus a correlation between *A* and *B* is like a low entropy state of a physical system, and it is to be explained in terms of an earlier causal interaction between the system and its external environment.

There are a number of difficulties facing Reichenbach's common cause principle. The principle seems to fail for certain quantum phenomena involving distant correlations, such as the one featured in the famous thought experiment by the physicists Albert Einstein, Boris Podolski, and Nathan Rosen, in their 1935 paper "Can Quantum Mechanical Description of Reality Be Considered Complete?" In a simplified version of this setup, two particles form a single system in which the total spin is zero. If the particles are separated, and the spin of each particle is measured, they will always be found to have opposite spins. There is thus a correlation between the outcome of the two measurements. Neither measurement result can be a cause of the other, for the measurements can be conducted at such a great distance that not even a light signal could connect the two. Yet a series of mathematical and empirical results, beginning with the work of the physicist John Bell in 1964, show that there can be no earlier state of the two-particle system that screens off the measurement outcomes.

A further problem is that it is unclear why Reichenbach's fork asymmetry should hold within the physical

framework of classical statistical mechanics. Within this framework, a system possesses a microstate that evolves deterministically according to Newton's laws of motion. An "event" *A* is just a coarse-grained characterization of the state of the system at a particular time, consistent with many different microstates. A probability distribution is defined over the possible states of the system. Suppose that the events *A* and *B* are correlated according to this probability measure, and that there is an earlier event *C* that screens off *A* from *B*. It is possible to take the image of *C* under the deterministic dynamics of the system; that is, one can evolve each microstate in *C* to some point in time after the occurrence of *A* and *B* and collect the resulting set of microstates into a new event *C′*. By construction, *C′* will stand in the same probability relations with *A* and *B* that *C* did. Hence, *C′* will be a later event that screens off *A* from *B*, and *ABCC′* will form a closed fork. Because this procedure is fully general, it is not clear how there can be forks open to the future at all. One possible reply to this worry is that in such a closed fork, the later screener-off C' will just be a heterogeneous collection of microstates, and hence will not qualify as an "event" in the relevant sense. This reply raises two new questions: first, which sets of microstates constitute genuine events? Second, why should we expect that only earlier screeners off will be genuine events?

FURTHER CAUSAL ANOMALIES

There are a number of further respects in which the world described by fundamental physics seems not to be one ruled by relations of cause and effect. It is well known that certain quantum-mechanical phenomena such as radioactive decay appear to be *indeterministic*. For example, even a complete description of the present state of a carbon-14 atom cannot allow one to predict whether or not it will decay during a certain period of time, but will instead yield only a probability that decay will occur. If the atom does eventually decay, can anything be said to cause the decay event? This kind of indeterminism provides part of the motivation for attempts to analyze causation in terms of probabilities. But even probabilistic theories of causation have difficulties when indeterminism is coupled with the sorts of distant correlations described in the previous section.

Moreover, even classical Newtonian physics admits indeterminism. For example, John Norton, in "Causation as Folk Science," describes a system consisting of a point mass sitting at the apex of a bell-shaped dome. Newton's laws of motion permit the point mass to rest there indefinitely, but they also allow it to begin sliding down the side of the dome in an arbitrary direction after an arbitrary finite time. No force is necessary to dislodge the mass: the sudden motion of the mass down the side of the dome is fully consistent with the constraint that at every instant, the force acting on the mass (due to the pull of gravity, and the reactive push of the dome's wall) is proportional to its acceleration. Such a motion thus appears to be entirely uncaused.

Einstein's general theory of relativity also gives rise to causal anomalies. For example, the Austrian-American mathematician Kurt Gödel showed that Einstein's field equations permitted solutions in which there were *closed causal curves*. Thus it may be possible for a billiard ball to get knocked, continue rolling along its new trajectory, and then eventually bump into its earlier self, knocking it into that new trajectory in the first place. Such a scenario appears to be at odds with people's ordinary conception of causation as an asymmetric relation, for the collision between the older and younger billiard ball causes the trajectory of the younger ball, which in turn causes that collision.

CAUSAL INFERENCE

One of Russell's targets in "On the Notion of Cause" was the so-called "law of causality"; indeed, it is this law, rather than the "notion of cause" itself, whose utility is compared to that of the British monarchy. Russell cites a formulation of this principle from the nineteenth-century British philosopher John Stuart Mill: "The Law of Causation, the recognition of which is the main pillar of inductive science, is but the familiar truth, that invariability of succession is found by observation to obtain between every fact in nature and some other fact which has preceded it." (Mill 1856, p. 359.)

According to Mill, science discovers causal relationships by discovering invariable regularities in nature, and the success of science presupposes the pervasiveness of such regularities. Russell was certainly right to challenge the importance of this law to science—not because science is not in the business of discovering causal relationships, but because causal inference in science does not rest upon the discovery of perfect regularities.

Causal inference presents a prima facie difficulty, first articulated by the Scottish philosopher David Hume in 1739. Suppose that one billiard ball collides with a second, causing it to move. One can observe the motion of the first billiard ball; and one can observe the motion of the second billiard ball; but one cannot observe the causation that connects the two together. How, then, is a person to acquire knowledge of causal relationships?

Traditionally there have been two main lines of response to this problem. One line that has already been mentioned is to reject the notion of causation on the grounds that it is inaccessible to empirical investigation. The second line, adopted in different ways by Hume, Mill, and a number of twentieth and twenty-first century philosophers, is to try to spell out systematic connections between causation and observable phenomena such as empirical regularities in order to explain how the former can be inferred from the latter. The "law of causation" championed by Mill and attacked by Russell stems from this second line of response to the problem. (A third possibility, defended in the early part of the twentieth century by the French-American philosopher C. J. Ducasse, and in the middle of the twentieth century by the Belgian psychologist André Michotte, is to reject the claim that causation is not subject to direct perception. Even if this is possible in special cases such as billiard ball collisions, however, this hardly seems to be an adequate explanation for causal knowledge generally.) This problem concerning the empirical accessibility of causation has been a driving force behind attempts to banish causation, and also behind attempts to provide causation with a sound philosophical analysis.

In fact, however, causal inference is neither impossible nor a matter of reading causal relations off universal regularities or correlations. Causal inference, like other forms of scientific inference, is broadly "hypothetico-deductive" in character. A causal hypothesis is formulated, and in conjunction with various background assumptions (often involving causal relationships themselves), it is used to derive predictions about what types of correlations will be observed. These predictions are then compared with observations. In this way, causal hypotheses may be subjected to empirical test without the need for a direct reduction of causal claims to claims about regularities and the like.

EXPERIMENTATION

The most reliable causal knowledge comes not from passive observations, but from controlled experimentation. In the medical sciences, the experiments often take the form of randomized clinical trials. Consider the claim that a particular drug causes lowered blood pressure. How might one test this claim? One possibility would be to make the drug available on the open market and observe hypertension patients who choose to take the drug and those who do not. There is a problem with this methodology. Suppose that the drug is expensive; one might expect that patients who buy the drug will be wealthier on average then those who do not. Wealthier patients might enjoy any number of other benefits—such as access to better healthcare generally, better diets, and so on—that influence whether or not they experience a reduction in hypertension. If one finds that patients who take the drug do in fact experience greater reduction in blood pressure levels than those who do not, it can still not be known whether this reduction is due to the drug or due to one of the other advantages associated with wealth. In a randomized trial, it is determined randomly which patients will receive the drug and which will be given a placebo instead. Randomization helps to ensure that treatment is not correlated with any other causes that might influence recovery.

This example helps to show the importance of the distinction between genuine causal relationships, on the one hand, and mere regularities or correlations on the other. Suppose that the drug is available only to wealthy patients, and that patients who take the drug fare better, on average, than those who do not. If this correlation is due to the wealth of the patients who use the drug, rather than to any effect of the drug itself on hypertension, then one would not expect the correlation to persist under various policy interventions. For example, if the drug were to be covered by insurance, so that less wealthy patients could also afford to take the drug, then the correlation between use of the drug and lowered hypertension would disappear. As the philosopher Nancy Cartwright puts it in her paper "Causal Laws and Effective Strategies" (1983), causal relationships support "effective strategies," while mere correlations or regularities do not. It is for this reason, Cartwright argues, contrary to the opinion of Russell, that the notion of cause cannot be dispensed with. It is also for this reason that one often finds the most self-conscious attention to the specific concerns of causal inference in those branches of science that have a practical dimension, such as medicine and agronomy.

In many areas of science, randomized trials are not feasible. This may be due to the inability to produce the putative cause at will, or it may be due to the lack any analog of a control group that receives placebos. Nonetheless, in the experimental setting, it is often possible to isolate the influence of the cause under investigation by preventing other causes from operating. For example, an experiment might be conducted within a metallic container to eliminate external magnetic influences; or the experimental apparatus may be set afloat in a pool of mercury to prevent vibrations from being transmitted through the floor of the laboratory (as was done in the famous Michelson-Morley experiment of 1887, which failed to

detect any effect of the earth's motion on the speed at which light traveled). Sometimes the experimental preparations are more mundane, such as thoroughly dusting the apparatus to eliminate the effects of stray dust particles, or even removing pigeons found nesting in the apparatus (as was required by Arno Penzias and Robert Wilson, who discovered the cosmic microwave background in 1965).

CAUSAL MODELS

In some fields, such as macroeconomics, epidemiology, and sociology, experimental manipulation is simply not feasible, and causal relationships must be inferred from observed correlations. Beginning around 1990 has been an explosion of interest in developing causal modeling techniques to facilitate such nonexperimental causal inferences. Two important works that have garnered a substantial amount of attention from philosophers are *Causation, Prediction and Search* (2000), by the philosophers Peter Spirtes, Clark Glymour, and Richard Scheines, and *Causality: Models, Reasoning, and Inference* (2000) by the computer scientist Judea Pearl. Both frameworks employ graphs to represent causal relationships among sets of causal variables. The variables in a set V form the nodes of a graph, and certain pairs of variables are connected by *edges* in the graph. In a *directed graph*, the edges take the form of arrows, which point from one variable into another. If a graph over the variable set V contains an arrow from the variable X to the variable Y, that indicates that X is a *direct cause* of Y (also called a *parent* of Y): the value of X has an effect on the value of Y that is not mediated by any other variable in the set V.

The causal structure represented by a directed graph is connected to a probability distribution over the values of the variables by the *causal Markov condition*. This condition states that, conditional upon the values of its direct causes, the values of a variable are probabilistically independent of the values of all other variables, except for its effects. In other words, a variable's parents screen off that variable from all other variables, except for its effects. (The causal Markov condition is closely related to Reichenbach's common cause principle, discussed above.)

With the help of the causal Markov condition, as well as other conditions such as the *minimality* and the *faithfulness* conditions, a graph representing causal relationships among a set of variables will serve as a model that makes predictions about probabilistic relationships among the variables. In particular, it predicts that certain variables will be dependent or independent of others, either unconditionally, or conditional upon the values of other variables. These predictions can then be tested using normal statistical means.

The most obvious use of these methods is to test whether a postulated set of causal relationships among the variables in the set V is consistent with the statistical data about the values of those variables. But there are other types of problems where these methods can be applied. Even if one does not begin by hypothesizing a specific causal model, it is possible to determine which sets of causal relations among a variable set are consistent with the statistical data. Typically, the data will not single out one causal model, but will only pick out an equivalence class of statistically indistinguishable models. In this case, background knowledge may help to narrow the set of plausible models. In a different sort of problem, one begins with a qualitative causal model and uses it to make quantitative predictions about the effects of interventions that have not yet been performed.

It is important to note that the causal Markov condition is not an a priori constraint on the relationship between causal structure and probability. It can fail, for instance, if a variable set V omits a variable that is a common cause of two variables included in V. The causal Markov condition is at best an empirical assumption that holds for a wide variety of causal structures, and hence any application of techniques based on the causal Markov condition to infer causal relationships from probabilistic data carries substantive empirical presuppositions. A number of critics have charged that these presuppositions severely limit the utility of the new causal modeling techniques.

CONCLUSIONS

Contrary to Russell's claim, causal notions are as pervasive in science as they are in philosophy and everyday life. New scientific techniques continue to be developed for the discovery of causal relationships. Nonetheless, the world as it is described by the deepest physical principles bears little resemblance to a world that is regimented by asymmetrical causal relationships. Thus there remain a number of deep puzzles about how causal relationships can emerge from physical laws that themselves make no mention of causality.

See also Causation, Metaphysical Issues; Probability and Chance.

Bibliography

Albert, David. *Time and Chance*. Cambridge, MA: Harvard University Press, 2000.

Arntzenius, Frank. "Physics and Common Causes." *Synthese* 82 (1990): 77–96.
Bell, John. "On the Einstein-Podolsky-Rosen Paradox." *Physics* 1 (1964): 195–200.
Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford: Clarendon Press, 1983.
Corry, Richard, and Huw Price, eds. *Causation and the Constitution of Reality*. Oxford: Oxford University Press, 2005.
Dowe, Phil. *Physical Causation*. Cambridge, U.K.: Cambridge University Press, 2000.
Ducasse, Curt J. "On the Nature and Observability of the Causal Relation." *Journal of Philosophy* 23 (1926): 57–68.
Duhem, Pierre. *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press, 1991.
Earman, John. *A Primer on Determinism*. Dordrecht, Netherlands: Riedel, 1986.
Eells, Ellery. *Probabilistic Causality*. Cambridge, U.K.: Cambridge University Press, 1991.
Einstein, Albert, Boris Podolski, and Natha Rosen. "Can Quantum Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777–780.
Field, Hartry. "Causation in a Physical World." In *Oxford Handbook of Metaphysics*, edited by Michael Loux and Dean Zimmerman. Oxford: Oxford University Press, 2003.
Gasking, Douglas. "Causation and Recipes." *Mind* 64 (1955): 479–487.
Gopnik, Alison, and Laura Schulz. *Causal Learning: Psychology, Philosophy and Computation*. Oxford: Oxford University Press, 2005.
Hempel, Carl Gustav. "Aspects of Scientific Explanation." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965.
Hume, David. *An Enquiry concerning Human Understanding*. Oxford: Oxford University Press, 1999.
Hume, David. *A Treatise of Human Nature*, 2nd ed. Oxford: Clarendon Press, 1975.
Kant, Immanuel. *A Critique of Pure Reason*. Translated by Paul Guyer and Allen Wood. Cambridge, U.K.: Cambridge University Press, 1998.
Lewis, David. "Causation," with Postscripts. In *Philosophical Papers, Volume II*. Oxford: Oxford University Press, 1986.
Mackie, John. *The Cement of the Universe*. Oxford: Clarendon Press, 1974.
McKim, Vaughn, and Stephen Turner. *Causality in Crisis? Statistical Methods and the Search for Causal Knowledge in the Social Sciences*. Notre Dame, IN: University of Notre Dame Press, 1997.
Michotte, André. *The Perception of Causality*. New York: Basic Books, 1963.
Mill, John Stuart. *A System of Logic: Ratiocinative and Inductive*. 4th edition. London: Parker and Son, 1856.
Moriya, Yosuke, Hitoshi Kawaji, Takeo Tojo, and Tooru Atake. "Specific-Heat Anomaly Caused by Ferroelectric Nanoregions in Pb(Mg(sub 1/3]Nb(sub 2/3])O(sub 3] and Pb(Mg(sub 1/3]Ta(sub 2/3])O(sub 3] Relaxors." *Physical Review Letters* 90 (2003): 205901.
Norton, John. "Causation as Folk Science." *Philosopher's Imprint* 3 (4) (2003). Available from www.philosophersimprint.org/003004/.
Pearl, Judea. *Causality: Models, Reasoning, and Inference*. Cambridge, U.K.: Cambridge University Press, 2000.
Pearson, Karl. *The Grammar of Science*. 3rd ed. reprint. New York: Meridian, 1957.
Price, Huw. *Time's Arrow and Archimedes' Point*. Oxford: Oxford University Press, 1996.
Reichenbach, Hans. *The Direction of Time*. Berkeley: University of California Press, 1956.
Russell, Bertrand. "On the Notion of Cause." Proceedings of the Aristotelian Society 13 (1913): 1–26
Salmon, Wesley. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press, 1984.
Spirtes, Peter, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*, 2nd ed. Cambridge, MA: MIT Press, 2000.
Suppes, Patrick. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland, 1970.
van Fraassen, Bas. "The Charybdis of Realism: Epistemological Implications of Bell's Inequality." *Synthese* 52 (1982): 25–38.
von Wright, Georg. *Causality and Determinism.* New York: Columbia University Press, 1975.
Woodward, James. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press, 2003.

Christopher R. Hitchcock (2005)

CAUSATION IN INDIAN PHILOSOPHY

Indian philosophical theories, from their earliest speculative cosmologies and explorations of the nature of human existence—in the Vedas and Upanisads, whose compositions were completed by roughly the first half of the first millennium BCE—emphasized the plight of humans and their struggle towards a soteriological goal. An understanding of the evolution of the world and the place of human beings within it held out the hope of improving their lot, either in some other place after death or in the next life in the round of deaths and rebirths. Or even, as the Upanisads suggested, in the ultimate avoidance of rebirth itself—a theme adopted by much Indian philosophy thereafter.

As in Western metaphysical speculations about the nature of the cosmos and man's place within it, the Indian thinkers made central and vital use of the concept of a cause—*karana* in Sanskrit—and progressively developed a sophisticated understanding of this concept.

VEDAS AND UPANISADS

The earliest Vedic answers to the question of cosmological evolution suggested a god or gods, variously named and described, as creating and ruling over the human

world. Such views invoked probably the most obscure and difficult application of the concept of causation—that of creation—but had at least the merit of putting men and gods in a continuing relationship. Men could worship their gods, and indeed could wield a degree of control, through religious ceremonies that aimed to elicit benefits from them.

The Upanisads took a more subtle turn, concentrating on a deeper understanding of the nature of man himself. The "inner self," the *atman*, was distinguished from its physical embodiment and was taken to proceed through a series of rebirths according to a causal law of *karma*—whereby moral merit or demerit dictated the nature of the next rebirth. Ultimately it would hopefully achieve release from rebirths and acquire its final state of bliss (*moksa*).

The period from the fourth to the second century BCE was one of quite subtle developments, with new and deeper ideas of the causal operation of the law of *karma*, of the nature of human existence, and of the nature of and route to the soteriological end for man. The Hindu *Bhagavad Gītā* was composed—a part of the great epic the *Mahabharata* (the actual period of composition is still much disputed)—and two nonorthodox systems of ideas were introduced: Jainism and Buddhism. Interestingly, both Jainism and Buddhism have no place for deities in their systems, human existence and progression to the ultimate state of release from rebirth being said to depend on the efforts of the individual. We will look at these three systems, and at just some of the later developments through the classical period of Indian philosophy.

BHAGAVAD GĪTĀ

The *Bhagavad Gītā* (Song of the Lord) takes the form of a dialogue between the warrior-prince (*ksatriya*) Arjuna and Lord Krishna, who is a human manifestation (*avatara*) of the god Vishnu. Arjuna hesitates to lead his army into battle against his cousins who have usurped control of the state, suffering a confusion about which duty he should follow: fight to rectify the wrong they have done to society or refrain from fighting to protect his family and caste. Krishna argues that Arjuna should fight. The world is in a final epoch of the cycle of evolution and corruption, a process of dissolution that requires his coming to advise mankind on correct behavior. As Vishnu, he has designed the nature of human society with its hierarchy of castes and their associated socioreligious duties. By the law of *karma*, the *ātman* of each individual goes through the process of birth-death-rebirth (*sāmsara*), gaining merit according to good deeds and demerit according to bad. *Karma* in this context therefore has moral, religious, and soteriological dimensions. *Moksa*, final release, is achieved through individual effort. And the central theme of the *Bhagavad Gītā* is the doctrine of *karma-yoga*, a route to salvation that involves acting according to established socioreligious duties, for the sake of maintaining the social fabric and for pleasing god.

Quite apart from the question whether *karma-yoga* actually resolves a conflict of duties such as Arjuna's, there is a further question: whether the *Bhagavad Gītā* really leaves any room for freedom of action for Arjuna, or indeed mankind in general. The text ascribes such enormous powers to Vishnu that individual human effort seems futile. Nature—the world in which the *atman* becomes embodied—is a creation of Vishnu. It involves the interplay between three "strands" (*gunas*), called *sattva*, *rajas*, and *tamas*—which can be translated as "goodness," "passion," and "inertia," respectively. All nature is but the playing out of the interaction between these *gunas* in a mechanistic, deterministic way. The balance of the *gunas* in a particular individual also dictates his character and hence his actions. The *atman* cannot affect the *gunas*, and there seems no chance of choosing to follow the path of *karma-yoga*, much less any other activities.

The *Bhagavad Gītā* adds further worries for its *karma-yoga* theme, for Vishnu has foreknowledge of all that will happen, and retains a tight control over all actions—overt and psychological—of all human beings. "The Lord abides in the hearts of all beings, O Arjuna, causing them to turn round by His power as if they were mounted on a machine," declares Krishna in the final chapter. So the *Bhagavad Gītā* is a brave but flawed attempt to teach an ethics of engagement in traditional socioreligious duties. The law of *karma* was supposed to allow human beings to strive towards *moksa*, the law itself being a creation of Vishnu to ensure a just outcome for our efforts. The text's failure to sustain this account perhaps goes a long way toward explaining why a good deal of later philosophical speculation (if not common religious practices), including much of so-called orthodox or Hindu philosophy, found no room for a deity as originating and controlling human existence.

JAINISM

Jainism was founded in the sixth century BCE by Vardhamana—who became known as Mahavira (Great spiritual hero)—and is named after the Sanskrit word for *conqueror* (*jina*). (Though Vardhamana left no texts, a particularly important text was composed by Umasvati

some nine centuries later: *Tattvarthadhigama Sutra*, or *Discourse on the Nature of Things*.) It is system that supposedly commends itself to reason. Rejecting the authority of the Vedas, it nevertheless keeps the idea of a spiritual substance, a *jiva*. Entrapment in the round of births and deaths is seen as a consequence of fine polluting karmic dust that restricts the all-knowing ability of the *jiva*. The route to salvation involves the elimination of this pollution, to achieve the state of perfect knowledge (*kevala*). In contrast to the complex interpretation of the workings of the law of *karma* in the orthodox tradition, the Jain account might appear a straightforward theory of physical causation. Yet the process of karmic improvement nevertheless has a serious moral dimension, for it involves a commitment to five "vows of restraint": nonviolence (*ahimsa*), truthfulness (*satya*), no theft (*asteya*), sexual continence (*brahmacharya*) and nonattachment to worldly pleasures (*apigraha*). By the individual's own efforts, therefore, the desired end of perfect knowledge is achieved.

BUDDHISM

Buddhism was founded, also in the sixth century BCE, by Siddartha Gautama, who became known as the Buddha (the Enlightened One) and spent many years proclaiming his insights into the predicament of the cycle of births and deaths and the route to release into *nirvana*. He left no writings of his own, but his teachings are recorded in the collection known as the *Sutta Pitaka* of the Pali work the *Tripitaka* (Three baskets of tradition). The Buddha taught a system of ideas that was in stark contrast to the earlier orthodox Vedic tradition, rejecting any reliance on those texts, on the priestly caste (the Brahmins), and on the orthodox depiction of salvation. Nothing brings out this contrast more than claim that reality has these three marks: impermanence (*anitya*), no-soul (*anatman*) and suffering (*duhkha*). A standard depiction of reality (*brahman*) by the Hindu tradition is quite the opposite: being as a permanent (*sat*), consciousness (*cit*) and bliss (*ananda*).

The Buddha's system is supposedly based upon observation, both of the world outside him and of the inner workings of his mental world. Crucially, he could not observe an *atman*. Instead, he reports as his fundamental discovery that all the ingredients observed obey a general principle of "dependent origination" (*pratityasamutpada*). Whatever comes into existence is the causal consequence of previous existents. Causal generation has a complex form where a number of such previous existents produce together the new existent. And each and every existent is momentary. Applying this general principle to the specific case of a sentient being, he classified all its momentary causal ingredients into five groups (*skandhas*). These can be rendered as these (following their later interpretation in the work *Milindapanha*, or Questions of King Milinda): thoughts (*vijnana*), feelings (*vedana*), volitions (*samskara*), perceptions (*samjna*) and bodily ingredients (*rupa*). And, most crucially, there being a complex interplay between the ingredients both within and across the groups, he identifies as the *fundamental* causal factor driving them all—through this life and through into rebirths—the thought "I am a permanent entity."

This cognitive error, involved as it is in the Hindu idea of the *ātman*, is the root cause of all grasping—for fame, for power, and for all other worldly goods—and therefore the root cause of suffering and rebirths. Only the correction of this error can lead to salvation. Moreover, this correction leads to a general change in motivations for action, whereby selfish desires are replaced by altruistic ones such as compassion, and the adoption of such altruistic desires in its turn helps to achieve the cognitive correction.

Within this new account of the human predicament is clearly embedded a sophisticated theory of causation. Dependent origination, the momentariness of the ingredients of causal chains, and the necessity linking the steps in causal development, together offer an impressive analysis of *karana*. Later Buddhist thinkers further sophisticated these ideas and indeed developed the theme that each new causal product is *genuinely* new, for the effect is not already existent in the cause. Such is the doctrine of *asatkaryavada*, the nonexistence of the effect in the cause.

SANKHYA

Sankhya is an orthodox school that, in common with Jainism and Buddhism, finds no room for a deity. The earliest authoritative text of the school is the *Sankhyakarika* (Verses on discrimination) of Isvarakrishna. Though this was probably composed in the fifth century CE, it is thought that the system of ideas can be traced back into the Vedic period.

There are, in this system, two kinds of substance: the experiencer and the experienced. The former (comparable to the *atman* of the Upanisads) is *purusa*, an inactive "silent witness" of the latter, *prakrati* or nature. *Purusas* are eternal and numerous, whereas *prakrati* is eternal and singular. The account of *prakrati* in Isvarakrishna's text is

a complex story about its evolution out of an original state of equipose between the *gunas*.

Sattva is the strand of nature that is productive of consciousness or intelligence; *rajas* is the strand productive of activity; *tamas* the strand productive of resistance. The original state of equipose is *pradhana*, meaning "the inferred one" because its existence is claimed on the basis of inference by analogy from experience. The first evolute is *Mahat* (the Great One) or *buddhi* (the subtle material that forms the basis of consciousness). Next comes *ahamkara* (the basis of individuation or self-sense), and then evolution takes two directions where either *sattva* or *tamas* predominates. Through the *sattva* route evolve *manas* (mind, of perhaps better brain), the five organs of perception, and the five organs of action. Through the *tamas* route evolve the five subtle elements (essences of sound, touch, taste, smell, and sight), and the five gross elements (ether, air, light, water, and earth) that are the constituents of all gross matter.

At first sight the process seems to be a cosmic evolution, with at least some roots in the early Vedic tradition. Yet it clearly is also designed to explain the nature of *sāmsara* and *moksa* for individual *purusas*. But why does nature evolve in this way? There is no deity to start it and plan its process. A *purusa* becomes entrapped in *samsara* by becoming engrossed in the play of nature before it, and, losing its awareness of its distinction from *prakrati*, it conceives itself as an embodied self, as an actor within the natural world. To achieve *moksa* it needs to regain its awareness of its distinct status as the pure inactive witness of *prakrati*. The Sankhyans indeed identify the following two purposes behind the evolution of *prakrati*: it evolves to provide experience for *puruṣas* yet at the same time to provide the possibility of this ultimate release from *sāmsara*.

SANKARA'S CRITIQUE

Sankara, the eighth-century Hindu philosopher, criticizes the Sankhyan system's explanation of the evolution of *prakrati* as follows: neither *prakrati* nor *puruṣas* can provide the efficient cause (*nimitta karana*) of this evolution, for *prakrati* is insentient—it lacks *cit*, or intelligence—and *purusas* are inactive. Such evolution cannot be spontaneous, for no spontaneous activity is evident in experience. However, the Sānkhyans believe they can find such cases; but the important issue between them and Sankara seems to be more fundamental. The Sānkhyans are working with the idea of the purpose of evolution, as opposed to causation. The evolution of *prakrati* is a natural development that serves the purposes of *puruṣas*, and no intelligent designer is required contrary to Sankara's insistence. We might well compare the Sānkhyan approach to that of Aristotelian teleological explanation.

Sankara's criticism comes in his major text, the *Brahma-sutra-bhasya* (Commentary on the verses concerning reality). He is a major figure in the Vedanta school, which takes its inspiration from the ancient Upanishads. Unlike the Sankhyans, he is unwilling to engage in speculative reasoning beyond the words of those texts and claims to be merely restating their essential message. Other figures in the Vedanta tradition also wrote commentaries on the *Brahma sutra*, and we can judge Sankara's philosophical inventiveness from the quite striking differences in the contents of those commentaries.

Both Sankara and the Sānkhyans adopt a view of causation whereby the effect preexists in the material cause (*upadana karana*)—called *satkaryavada*. They differ, however, in the detail. For the Sankhyans the evolution of *prakrati* is a *real* process of natural unfolding out of the potentialities of the *gunas*—a position known as *parinamavada*. Sankara, however, finds difficulties with the notion of potentiality and argues instead for the more extreme position of the identity of the effect with the cause—there is only a *merely apparent* transformation from cause to effect. Applying this claim—known as *vivartavada*—to the case of the emergence of the experienced world out of the one real thing, *Brahman*, which is undifferentiated consciousness, the implication is that the experienced world is but an illusory appearance of *Brahman*. The route to *moksa* is the realisation of this difficult truth.

NYĀYA

Nyaya is another orthodox school, beginning with the third-century BCE text by Gautama, the *Nyāya Sutras* (Verses on argument). Important commentaries were written by Vatsyayana and Uddyotakara in about the third and sixth centuries CE, respectively, and substantial developments continued with the Navya-nyāya (or "new Nyāya") thinkers of the fourteenth to sixteenth centuries. Since argument or reasoning is often concerned with causal relations in the observed world, the Nyāya philosophers gave considerable attention to an analysis of such relations.

Causation, on their understanding, is the real production of new things out of the parts of matter (ultimately atoms). This is another version of *asatkaryavada*, for the effect is a *new* existent. From threads we can make a cloth, and from clay we can make a pot: The cloth and

the pot are new products of the causal process. They do, however, stand in a special relationship to the threads and clay, a relationship called *samavaya* (inherence). The cloth, for example, is said to inhere in the threads as *one in many*, one thing in many things; just as much as the threads are parts of the cloth as *many in one*. The idea of a material cause (*upadanakarana*) is given this new interpretation by this school—the matter or parts out of which something is made is called the "inherent cause"(*samvayikarana*).

Causation also involves an efficient cause (*nimittakarana*) or causes, such as the work of the weaver and the motions of the loom. Any case of causal production is likely to involve a multitude of factors—actions or material ingredients and all their individual qualities—and the Nyāya philosophers duly classify such factors further in terms of their efficacious or peripheral role in the process. A cause, in the final analysis, is the sum of the causal factors that are the invariant and unconditional antecedent of the effect.

The Nyāya account was criticized by both Buddhists and Sankara. For the Buddhists it is in stark contrast to the aggregate (*skandha*) theory, according to which the "new" product is merely the sum of the parts, and they try to fault the special relation of inherence that the Nyāya theory makes central to its account. Sankara, too, finds this relation logically flawed, since it leads to an infinite regress. If the parts and the new object are related by this *samavaya* relation, what relates *it* to the parts and the object? It seems it would have to be another case of *samavaya*, and then the same question arises again—without end.

See also Atomic Theory in Indian Philosophy; Brahman; God in Indian Philosophy; Negation in Indian Philosophy.

Bibliography

Gambhirananda, S., trans. *Brahma-sutra-bhasya of Sri Sankaracarya*. 2nd ed. Calcutta: Advaita Ashrama, 1972.

Hiriyanna, M. *Outlines of Indian Philosophy*. London: George Allen & Unwin, 1932. An established substantial introduction to Indian philosophy.

Jaini, J. L., trans. *Umasvati's Tattvarthadigamasutra*. Arrah, India: The Central Jaina Punlishing House, 1920.

Jha, G. *Nyayasutras of Gautama*. Poona: Oriental Book Agency, 1939.

O'Flaherty, W. D., ed. *The Rig Veda*. Harmondsworth: Penguin Books, 1981.

Radhakrishnan, S., ed. *The Principal Upanisads*. Oxford and New York: Oxford University Press, 1966.

Radhakrishnan, S., and C. A. Moore, eds. *A Sourcebook in Indian Philosophy*. Princeton, NJ: Princeton University Press, 1957. A useful resource containing extracts and sometimes complete translations (including the *Bhagavad Gītā*) of primary source materials.

Sastri, S. S. S., trans. *The Sankhyakarika of Isvara Krishna*. Madras: University of Madras, 1948

Warren, H. C., ed. *Buddhism in Translations*. Cambridge, MA: Harvard University Press, 1915.

Brian Carr (2005)

CAUSATION IN ISLAMIC PHILOSOPHY

According to the Qur'anic position, God is the voluntary creator of the universe. In causal theory, one finds an apparently necessary connection between cause and effect. Islamic philosophy experiences a profound tension between these two ideas—the Qur'anic legacy of God's will and the idea of independent causes leading to effects. From this perspective, one may observe four stages in the concept of causation in Islamic philosophy.

THE FIRST STAGE

The first stage, beginning with the rise of Islam in the seventh century and extending well into the tenth, is dominated by the Qur'anic understanding of cosmos, which assigns God as the fundamental cause of the universe and of the events taking place within it. A cause is thus conceived as a "means" or "way" conditioned or provided by God as a blessing to achieve something, as indicated in the following verses: "Do they not look at the camels how they are created? And at the sky how it is raised? And at the mountains how they are fixed firm? And at the earth how it is spread out?" (Qur'an 88:17–20); "it is God who causes the seed and the date-stone to split and sprout. He brings forth the living from the dead, and brings forth the dead from the living …" (Qur'an 6:95–104; also 67:3–4; 24:39; 2:118). Early philosophers of the Kalām Theology School attempted to express this Qur'anic understanding by their metaphysics of atoms and accidents. They argue that because each atom is created and annihilated at every instance, no being can subsist by itself and have an effect on another body except through the creation of an omnipotent God. In this scheme, causation is conceived as a creation at every instance, including human actions. Abū'l-Ḥasan ʿAlī ibn Ismaʿīl al-Ashʿarī (d. 935) argued that "God wills everything which can be willed" and that every instance of causation is to be conceived within the domain of this all-embracing divine will (1953, p. 33).

THE SECOND STAGE

In the second stage the Muslim Neoplatonic Aristotelians establish a philosophical theory claiming the necessary connection of cause and effect. Abū Yūsuf Yaʿqūb ibn Isḥāq al-Kindī (c. 801–866), Abu Naṣr Muḥammad al-Fārābī (870–950), and Abū ʿAlī al-Ḥusayn ibn Sīnā (980–1037) are the proponents of this school. Al-Kindī and al-Fārābī thus establish an emanationist system of universe that follows from God necessarily. This world system is decidedly necessitarian, neatly elaborated by Ibn Sīnā in a causally deterministic way. In his scheme, the universe is conceived as a hierarchical order of beings, which offers a cosmic pattern for causation in general and a model for all causal interactions. Each being is connected to the next in a necessarily ordered chain of causation beginning with God through the heavenly spheres down to the remote spheres of dark and primitive matter. The philosophers of the Kalām School vehemently objected to this theory claiming that, if accepted, the Qurʾanic understanding of God's absolute will and power becomes vacuous.

THE THIRD STAGE

Three prominent philosophers represent the third stage: Abū Ḥāmid Muḥammad al-Ghazālī (d. 1111), Abūʾl-Walīd Muḥammad ibn Rushd, known as Averroes (d. 1198), and Ṣadr al-Dīn Muḥammad ibn Ibrāhīm al-Shirāzī, known as Mullā Ṣadrā, or Ṣadrā (d. 1641). Against the philosophers of the Neoplatonic Aristotelian School, al-Ghazālī argues along Humean lines that people observe in existence not a necessity but two things that are contiguous. The connection, therefore, between a cause and its effect is due to the prior decree of God, who creates them side by side. What does not have a free will cannot enter into a temporal relation. When a piece of cotton burns, it is not the fire that is burning, for fire is inanimate and in itself has no action. What proof can be given that the fire is the agent? The only proof is that people observe an act of burning, not any other mediating factor. Therefore, existing contiguously with a thing does not prove causation between two things. Ghazali denies skepticism by arguing that the repeated occurrence of events fixes unshakably in our minds the belief in their occurrence according to past habit.

Ibn Rushd objected to this theory, arguing that in denying the necessity of a causal link, al-Ghazālī's motive was to defend the exclusive prerogative of God's sovereignty and efficient causal agency in all events. But the denial of this connection involves the rejection of an agent in an act, and hence, the logical ground for the idea of God as an efficient cause is destroyed. Moreover, logic implies the existence of causes and effects, and knowledge of these effects can only be obtained through knowledge of their causes. Hence, denial of causes implies the denial of knowledge, which, in turn, implies that nothing can be really known.

Mullā Ṣadrā developed an existential theory of causation based on the primacy of existence. An abstract notion of existence arises in the mind, but that notion cannot yield true reality. For, in each case, existence is a unique individual in an ongoing process of renewal. Essences arise in the mind as a result of this process when existence becomes further diversified into modes. It is existence that moves within this process; both the cause and the caused are existence; the essence is caused to arise in the mind in connection to particular beings. Causation must be considered within that existential process in which the problem of necessary connection does not arise. In each instance of causality there is a temporal emergence in which the temporal emergent, that is, the cause, is not the true cause but only a preparatory condition for it. The true cause in such an emergence is, therefore, the eternal creative act of God. In that case, this process is continuous, not discrete, involving change in the substance of everything that moves within the process.

THE FOURTH STAGE

In the fourth stage one finds primarily the idea of causal explanation on the basis of the Qurʾanic notion that God acts regularly and that there is no change in this regular course of action, called *sunnat Allah*. No thinker in this stage paid more attention to the problem of causation than the twentieth century thinker, Bediüzzaman Said Nursi of Turkey (d. 1960). Nursi uses two arguments to defend al-Ghazālī's theory of causation. The first is the argument from theodicy that establishes that "might and majesty require causes to be veiling occasions of God's omnipotence for the human mind" (Nursi 1996, p. 1278). God creates things for certain good ends. If causes are not seen as veils for God's acts, the human mind will directly infer God in all natural phenomena and attribute the seemingly evil results of these actions to him. This inference harms God's might and glory. Similarly, we may not be able to see good results immediately and thus blame God for evil. The second argument claims that "God's uniqueness and glory require causes to withdraw their interference from the actual efficacy" (Nursi 1996, p. 1278). The nature of an effect exhibits a perfection that is the result of a rational planning and omnipotence. These

qualities are not inherent in the causes producing their effect; hence, the true cause is outside the event, deduced by the mind and experienced by the awakened heart. There is thus only one true cause, God, who assures people of the causal nexus through the first argument by theodicy.

See also al-Fārābī; al-Ghazālī, Muhammad; al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotelianism; Averroes; Avicenna; Causation: Metaphysical Issues; Islamic Philosophy; Mullā Ṣadrā; Neoplatonism.

Bibliography

al-Ashʿarī, Abūʾl-Ḥasan ʿAlī ibn Ismaʿīl. *Al-Ibanah ʿan Usul al-Diyanah: The Elucidation of Islam's Foundation*. Translated with an introduction and notes by Walter Klein. New Haven, CT: American Oriental Society, 1940. Reprint, New York: Kraus Reprint Corporation, 1967.

al-Ashʿarī, Abūʾl-Ḥasan ʿAlī ibn Ismaʿīl. *Kitab al-Lumaʿ*. In *The Theology of Al-Ash'ari*. Translated by Richard J. Beyrouth: Imprimerie Catholique, 1953.

Colish, M. L. "Avicenna's Theory of Efficient Causation and its Influence on St. Thomas Aquinas." *Tommaso d'Aquino nel suo Settimo Centenario* 1 (1975): 296–306.

Druart, T-A. "Al-Fārābī's Causation of the Heavenly Bodies." *Islamic Philosophy and Mysticism*, edited by Parviz Morewedge. Delmar, NY: Caravan, 1981.

Fakhri, Majid. *Islamic Occasionalism and its Ctitique by Averroes and Aquinas*. London: George Allen & Unwin Ltd., 1958.

Ivry, A. L. "Averroes on Causation." In *Studies in Jewish Religious and Intellectual History, Presented to Alexander Altmann*, edited by Siegfried Stein and Raphael Loewe. Tuscaloosa: University of Alabama Press, 1979.

Goodman, Lenn Evan. "Did al-Ghazālī Deny Causality?" *Studia Islamica* 47 (1978): 83–120.

Gyekye, K. "Al-Ghazālī on Causation," *Second Order* 2 (1973): 31–39.

Kogan, Barry S. *Averroes and the Metaphysics of Causation*. Albany: State University of New York Press, 1985.

Nursi, Bediüzzaman Said. *Mesnevi-i Nuriye*. Printed in the *Risale-i Nur Külliyati*. 2 vols. Includes the complete works of Nursi. Istanbul: Nesil Publications, 1996.

Shanab, R. E. A. "Ghazālī and Aquinas on Causation." *Monist* 58 (1974): 140–150.

Alparslan Açikgenç (2005)

CAVELL, STANLEY

(1926–)

Stanley Cavell, American philosopher and long-time professor of philosophy at Harvard University, has written on epistemology, philosophy of language, moral philosophy, and aesthetics; on Shakespeare and Romanticism and Samuel Beckett; on modernism in the arts, classic Hollywood film comedies and melodramas of the 1930s and 1940s, and opera; on his most direct influences, J. L. Austin and Ludwig Wittgenstein, especially with reference to their attempts to draw words back to their everyday homes; on Friedrich Nietzsche and Martin Heidegger, who articulate our perhaps inevitable ambivalence toward what the latter calls "average everydayness"; on Kant, who in limiting knowledge to make room for faith makes the conditions and boundaries of human understanding and the recognition of our finitude dominant themes for subsequent thought; and also on the Kantian inheritance in the transcendentalism of Thoreau and Emerson, who conceptualize these issues in terms of lost contact with things themselves and the possibility of an intimacy regained that allows for acceptance of the world's independence from us. Cavell's circle of interests has its unity: his overarching concern is with philosophy's aspirations to self-knowledge and with obstacles the intellect erects to self-knowledge, particularly in the form of distortions of self-expression and loss of voice. Cavell links these threats to skepticism, conceived not just as a general doubt about the extent of our cognitive capacities, but as an expression of a tragic condition of withdrawal haunting the present age. Later, he finds acknowledgment of and response to this condition in images of recovery articulated in the dimension of the moral life he calls "Emersonian perfectionism."

Several essays in *Must We Mean What We Say* (1969) defend the salience of philosophical appeals to ordinary language. In doing so, they prepare for the comprehensive diagnosis of skepticism and the impulses behind it offered in Cavell's central work, *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy* (1979). Because appeals to "what we say when" draw on knowledge of native tongues, they do not directly refute the skeptic by convicting him of linguistic mistakes. The skeptic, after all, remains a master of language. On the other hand, because skeptical procedures do not fully fit ordinary ways of raising and responding to doubts about particular claims, Cavell interprets skepticism's negative conclusions about the limits of human knowledge not as failures of certainty, but as intellectualized disappointment with the sources of our capacities for making sense of the world.

Accordingly, part one of *The Claim of Reason* offers a reading of the later Wittgenstein's notion of criterion, on which criteria constitute not certainty, but the relevance and applicability of our concepts to worldly circumstances. On this view, our capacity to speak intelligibly is

based on nothing deeper (nor less deep) than our agreement in judgment, which agreement is not secured prior to particular judgments. Criteria are thus subject to repudiation, as our agreements may seem to run thin. The skeptic errs in implying that criteria should be grounded in something deeper, lest our whole conception of things be deemed irredeemably subjective. But because the skeptic reminds us of the repudiability of criteria, the skeptic's progress (or lack thereof) conveys an important moral: our sense of things is not a cognitive accomplishment.

Part two elaborates the external world skeptic's failure to live up to his own self-conception as a perfect knower. This skeptic faces a dilemma: either he fails to specify concrete claims about the external world for scrutiny, or his doubts about the claims he does single out do not generalize to all beliefs about external objects. Here Cavell discerns a truth behind the external world skeptic's efforts—that our relation to the external world as a whole is not a matter of knowledge about an, as it were, all-encompassing object, but rather one of acceptance. While such a conclusion may seem to exacerbate the skeptic's sense that we are cut off from the world, Cavell asks whether this discomfort, expressive of disappointment with ordinary modes of inquiry, criteria—even our manner of involvement with things—is self-imposed.

Part three of *The Claim of Reason* explores the nature of practical reasoning and the limits of both morality and traditional moral theorizing. Cavell sets himself against the "moralization of morality": the assumption that if morality is genuinely rational, it must rest on rules grounding its verdicts and rendering it competent to assess the value of every action. Much as the skeptic prescinds from actual practices of evaluating epistemic claims, so the moralist refuses the concept of morality by failing to locate its role in everyday life.

Part four, exploring symmetries and asymmetries between external world and other minds skepticism, argues that in the case of other minds, acknowledgment of others—not certainty about their inner lives—is in question. The tragic fate of the present age is that for the most part, we live our skepticism, tending, as a matter of historical fact, to shirk our responsibilities in knowing others and in making ourselves known to them. At stake is the voice—our expressiveness, and the barriers we erect to it.

Cavell's later writings explore his sense that responding to our tragically skeptical state, working through the issue of the voice, is a crucial task of modernity. Cavell reads romanticism (exemplified in Wordsworth and Coleridge as well as Emerson and Thoreau, thematized most explicitly in *In Quest of the Ordinary: Lines of Skepticism and Romanticism* [1998]) as registering both the success of and dissatisfaction with Kant's settlement with skepticism. Acknowledging that the quest for knowledge, at least as conceived by skepticism, blocks our access to the things themselves, romanticism seeks other routes to their recuperation. These lie in the particulars of our ability to make sense of them, despite the lack of philosophical grounding for our ways of doing so. At the same time, in reading the defining texts of moral perfectionism (especially in *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism* [1990] and *Cities of Words: Pedagogical Letters on a Register of the Moral Life* [2004]), Cavell finds in this openness the potential for the creation or discovery of a self capable of articulating its own identity, its own ideals and possibilities, again without need of a foundation from outside. In large part, recovery from the threat of skepticism lies in everyday uses of words, not because they express a set of commonly-held beliefs, but insofar as they manifest a responsiveness to ourselves and the world that enables us to find our conditions intelligible.

Bibliography

WORKS BY CAVELL

Must We Mean What We Say. New York: Scribners, 1969.

The World Viewed: Reflections on the Ontology of Film. New York: Viking, 1971.

The Senses of Walden. New York: Viking, 1972.

The Claim of Reason: Wittgenstein, Skepticism, Morality and Tragedy. Oxford: Oxford University Press, 1979.

Pursuits of Happiness: The Hollywood Comedy of Remarriage. Cambridge, MA: Harvard University Press, 1981.

Themes Out of School: Effects and Causes. San Francisco: North Point Press, 1984.

Disowning Knowledge: In Six Plays of Shakespeare. Cambridge, U.K.: Cambridge University Press, 1987.

In Quest of the Ordinary: Lines of Skepticism and Romanticism. Chicago: University of Chicago Press, 1988.

This New Yet Unapproachable America: Lectures after Emerson after Wittgenstein. Chicago: University of Chicago Press, 1988.

Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism. Chicago: University of Chicago Press, 1990.

A Pitch of Philosophy: Autobiographical Exercises. Cambridge, MA: Harvard University Press, 1994.

Philosophical Passages: Wittgenstein, Emerson, Austin, Derrida. Oxford: Basil Blackwell, 1995.

Contesting Tears: The Melodrama of the Unknown Woman. Chicago: University of Chicago Press, 1996.

Emerson's Transcendental Etudes. Stanford, CA: Stanford University Press, 2003.

Cities of Words: Pedagogical Letters on a Register of the Moral Life. Cambridge, MA: Harvard University Press, 2004.

SELECTED SECONDARY WORKS

Eldridge, Richard, ed. *Stanley Cavell*. Cambridge, U.K.: Cambridge University Press, 2003.

Goodman, Russell, ed. *Contending with Stanley Cavell*. Oxford: Oxford University Press, 2004.

Gould, Timothy. *Hearing Things: Voice and Method in the Writings of Stanley Cavell*. Chicago: University of Chicago Press, 1998.

Mulhall, Stephen. *Stanley Cavell: Philosophy's Recounting of the Ordinary*. Oxford: Clarendon Press, 1994.

Edward Minar (2005)

CAVENDISH, MARGARET

(1623?–1673)

Margaret Cavendish was born into the Lucasses, a family of English gentry. She does not seem to have had an education that was in any way remarkable for a young woman of her time. Indeed, she reports that while she had the usual tutors, her mother "cared not so much for our dancing and fidling, singing and prating of several languages" (Cavendish 1667), deeming honesty and civility more important. One consequence is that Cavendish was never able to speak or read any language but her native English. In 1643, when she was about twenty, she became a maid of honor to Charles I's wife, Queen Henrietta Maria, and the next year she followed the queen into exile in Paris.

While at the court in exile, she met and subsequently married William Cavendish, who eventually became the Duke of Newcastle and who was a widower some twenty years her senior. The marriage seems to have been a happy one, and indeed, it is Margaret, a second and childless wife, who lies buried next to William in Westminster Abbey. Margaret Cavendish found a husband who supported her ably in her intellectual endeavors. In marrying into the Cavendish family, she became a member of a family that had been in the forefront of the intellectual life of the time. Newcastle's cousins, the Devonshires, were patrons of Thomas Hobbes, and Newcastle and his brother, Sir Charles Cavendish, had as part of their circle a number of leading thinkers, including Marin Mersenne, Pierre Gassendi, and René Descartes. It is not clear to what extent this wider circle was available to Cavendish, but both her husband and her brother-in-law were prepared to encourage and to instruct her as she developed her intellectual interests. Cavendish published copiously, in a wide variety of genres, throughout her life, both while she and her husband lived in exile in Holland and after they returned to England in 1660, after the restoration of Charles II. The Newcastles lived on their return at the family estate, Welbeck, in Nottinghamshire, but made visits to London. During one of these visits, Cavendish made a ceremonial visit to the Royal Society, unusual in that they did not otherwise admit women to its meetings. Cavendish died in 1673, at the relatively young age of fifty, some three years before her husband.

Cavendish published over a dozen works, including poetry, plays, epistolary treatises, a life of her husband and a shorter one of herself, a novel, and some six works in natural philosophy. Cavendish reworked her ideas about natural philosophy throughout her life, improving them as she enlarged her reading and altered her vocabulary and her grasp on the issues about which she was writing. Among her works in natural philosophy, probably the best and most interesting are her last, *Grounds of Natural Philosophy* (1668/1996), where she lays out her material in its most organized form, and two slightly earlier works, *Philosophical Letters* (1664/1994) and *Observations upon Experimental Philosophy* (1666/2001). These last two are especially interesting because, in them, Cavendish situates her own views against a commentary on several leading thinkers of her day.

From *Grounds of Natural Philosophy* one learns the basic premises of Cavendish's approach to natural philosophy. She tells the reader there can be no substance but body, which exists in degrees of purity. While the less pure parts of matter are inert, the purer parts are self-moving and are endowed with self-knowledge. These come in two sorts, again distinguishable by their degree of purity: a sensitive part, which is living, and a rational part, which understands. Natural phenomena are to be explained in terms of the doings of matter, under the guidance of reason and as carried out by sense. Thus, Cavendish's account of nature is one of a number of accounts that try to explain natural phenomena in terms of the motions that lead to the division and composition of otherwise undifferentiated matter. Cavendish has absorbed and is working within one of the dominant explanatory paradigms of her day.

As *Philosophical Letters* and *Observations upon Experimental Philosophy*make clear, Cavendish developed her own version of this paradigm. *Philosophical Letters* consists of a series of letters to a fictional female correspondent discussing passages of Hobbes, Descartes, Henry More, and Francis Mercury van Helmont, with a final, less focused part answering a number of different questions and mentioning a number of different authors, including Galileo Galilei, Walter Charleton, and Robert Boyle. Unlike Hobbes and Descartes, Cavendish

rejects the idea that there can be purely mechanical explanations for such human functions as sensation, insisting on the self-moving, knowledgeable nature of sensation, which she says “patterns out” or imitates objects sensed. She rejects a mechanical or “transfer” theory of motion as unintelligible and provides an alternative, under which all motion is self-generated action on the basis of self-knowledge, rather than a passive reaction to impact. Thus, while a materialist, Cavendish is not a mechanist, but a vitalist. She energetically distinguishes herself, however, from other contemporary vitalists, like More, on the grounds that More’s immaterial plastic spirit of nature, as immaterial, is impotent to move matter. Cavendish’s vitalism is materialist and not dualist. Cavendish’s position can be seen as developing in conversation with a number of related theorists, with whom she shares a number of views, while carving out her own position.

In *Observations upon Experimental Philosophy* Cavendish takes on the experimenters of the Royal Society, in particular Robert Hooke. She criticizes Hooke for supposing that microscopes provide a unique view into the heart of things, on the grounds that adding a dubious instrument to a dubious sense organ does not improve matters. Her overall approach is to urge the claims of reason to give understanding over the deliverances of the senses. Although arguing for the special virtues of reason, Cavendish does not suppose that reason is a source of certainty in natural philosophy. Instead, her approach is probabilistic. Toward the end of *Philosophical Letters* she writes that

> the undoubted truth in Natural Philosophy is, in my opinion, like the Philosophers Stone in Chymistry, which has been sought by many learned and ingenuous Persons and which will be sought as long as the Art of Chymistry doth last; but although they cannot find the Philosophers Stone, yet by the help of this Art they have found out many rare things, both for use and knowledg. (1664/1994, p. 508)

While one cannot attain undoubted truth, to refuse to be guided by it would be like refusing to take medicine on the grounds that one will die eventually.

See also Boyle, Robert; Descartes, René; Galileo Galilei; Gassendi, Pierre; Hobbes, Thomas; Materialism; Mersenne, Marin; More, Henry; Vitalism; Women in the History of Philosophy.

Bibliography

WORKS BY CAVENDISH

Philosophical Fancies. London: N.p., 1653

Philosophical and Physical Opinions. London: Printed for J. Martin and J. Allestrye, 1655.

“Philosophical Letters: or, Modest Reflections upon some Opinions in Natural Philosophy” (1664). In *Women Philosophers of the Early Modern Period*, edited by Margaret Atherton. Indianapolis, IN: Hackett, 1994.

Observations upon Experimental Philosophy (1666). Edited by Eileen O’Neill. New York: Cambridge University Press, 2001.

A True Relation of My Birth, Breeding and Life, as printed in *The Life of William Cavendish, Duke of Newcastle*. London: 1667, 157–158.

Grounds of Natural Philosophy (1668). Introduction by Colette V. Michael. West Cornwall, CT: Locust Hill Press, 1996.

“The Blazing World.” In *The Description of a New World, called the Blazing World and Other Writings*, edited by Kate Lilley. New York: New York University Press, 1992.

Political Writings. Edited by Susan James. New York: Cambridge University Press, 2003.

WORKS ABOUT CAVENDISH

Bowerbank, Sylvia, and Sara Mendelson, eds. *Paper Bodies: A Margaret Cavendish Reader*. Orchard Park, NY: Broadview Press, 2000.

Broad, Jacqueline. “Margaret Cavendish.” In *Women Philosophers of the Seventeenth Century*, 35–64. New York: Cambridge University Press, 2002.

Hutton, Sarah. “Anne Conway, Margaret Cavendish, and Seventeenth-Century Scientific Thought.” In *Women, Science, and Medicine, 1500–1700: Mothers and Sisters of the Royal Society*, edited by Lynette Hunter and Sarah Hutton, 218–234. Stroud, U.K.: Sutton Publishing, 1997.

Hutton, Sarah. “In Dialogue with Thomas Hobbes: Margaret Cavendish’s Natural Philosophy.” *Women’s Writing* 4 (3) (1997): 421–432.

Hutton, Sarah. “Margaret Cavendish and Henry More.” In *A Princely Brave Woman: Essays on Margaret Cavendish, Duchess of Newcastle*, edited by Stephen Clucas, 185–198. Aldershot, U.K.: Ashgate, 2003.

Hutton, Sarah. “Science and Satire: The Lucianic Voice of Margaret Cavendish’s *Description of a New World called the Blazing World*.” In *Authorial Conquests: Essays on Genre in the Writings of Margaret Cavendish*, edited by Line Lottegnies and Nancy Weitz, 161–178. Madison, NJ: Fairleigh Dickenson University Press, 2003.

James, Susan. “The Innovations of Margaret Cavendish.” *British Journal for the History of Philosophy* 7 (2) (1999): 219–244.

Margaret Atherton (2005)

CELSUS

Celsus, a Middle Platonist (Origen wrongly called him an Epicurean) critic of Christianity, wrote the *Alethes Logos*

(True doctrine) about 178 CE. We know the work—whose title derives from a Platonic expression (*Meno* 81a)—only through quotations in Origen's reply, *Contra Celsum,* composed seventy years later. Celsus began his work by assuming the character of a Jew and attacking Christian views from this standpoint. Then he proceeded on his own to demonstrate their inadequacy in relation to the basic axioms of contemporary philosophical theology, especially with regard to the doctrines of God and providence and poetic-philosophical inspiration; as a Platonist he found the Christian idea of the Incarnation both impossible and immoral. At the end of his work he urged the Christians to abandon their irrational faith and join him in upholding the state and its religion. After Christianity was recognized by the Roman government, Celsus's work was destroyed.

The theology of Celsus is based, in his own view, on an ancient tradition handed down, especially among oriental wise men, from remote antiquity. This tradition, the "true doctrine," informed him of the existence of one god known by many names and worshiped by all pious men. Such a "polytheistic monotheism," he believed, had been perverted or misunderstood, first by the Jews and then by the Christians. If they were to return to the tradition, they would abandon their irrational exclusiveness and would recognize the divine right of the one emperor. His work thus culminates in a theology of politics.

Origen's reply is important not only because in it his philosophical theology, developed earlier, is clearly expressed in relation to Celsus's views, but also because it shows the extent to which he agreed with Celsus in opposing more literal religious conceptions. Each held, for example, that his own authoritative traditions are to be understood symbolically, whereas the other's traditions must be meant literally. But Origen finally took his stand on the particularity of the Hebrew-Christian tradition, which Celsus found totally unacceptable.

See also Origen; Platonism and the Platonic Tradition.

Bibliography

Bader, R. *Der Ἀληθὴς λόγος des Kelsos.* Stuttgart and Berlin, 1940. Critical edition of Greek text.

Chadwick, Henry. *Origen: Contra Celsum.* Cambridge, U.K. Cambridge University Press, 1953. Translation with introduction and very full notes.

Robert M. Grant (1967)

CENSORSHIP

"Censorship" is the suppression of speech or symbolic expression for reason of its message. Liberal Western constitutionalism has traditionally condemned censorship on both instrumental and intrinsic grounds, classically articulated by John Stuart Mill in *On Liberty.* In this traditional liberal view, freedom of speech instrumentally serves the ends of truth and self-government. Censorship, by entrenching orthodoxy and suppressing dissent, impedes the advancement of truth and the processes of democratic change. Freedom of speech is also intrinsically valuable, in this view, as an aspect of human autonomy. Censorship illegitimately interferes with that autonomy, because speech, unlike action, typically causes others no harm. The proper response to bad speech is more speech, not government regulation.

Late-twentieth-century and early-twenty-first-century critics have challenged both the instrumental and the intrinsic justifications for freeing speech from censorship. First, some suggest that the power to speak is so unequally distributed that free competition in the marketplace of ideas is unlikely to produce either truth or democracy. For example, advocates of regulating campaign advertisements argue that wealthy voices dominate and thus distort political debate, and advocates of hate-speech regulation argue that racial epithets and invective perpetuate a form of cultural white supremacy in which minority voices are effectively silenced. These critics would turn the traditional free-speech principle on its head. In their view freedom of speech helps to entrench the existing status quo while government regulation of the speech of powerful groups can level the playing field. Redistribution of speaking power would advance truth and political equality better than a regime of laissez-faire.

Second, some critics argue that the defense of free speech on autonomy grounds undervalues the harms that speech causes. On this view speech regulation ought to be more widely allowed to protect the countervailing autonomy interests of listeners or bystanders. Liberal constitutional democracies generally permit censorship only to avert a narrow range of material harms. For example, incitement to riot may be forbidden, as may publication of the movements of troops at war. But censorship is rarely permitted on the ground that speech will cause disapproval, anger, alarm, resentment, or offense on the part of the audience. American constitutional law categorically forbids such justifications. Legal systems that permit them do so only in exceptional contexts: For example, British law forbids expressions of racial hatred, and some

international human rights laws forbid advocacy of genocide.

Free-speech critics argue that such exceptions should be more the rule. First, some argue, government should be free to prevent injury, not only to bodies, but also to hearts and minds, including the injury caused by expressions of caustic opinion. Second, others argue, speech should be regulable for its social impact, even in the absence of immediate physical harm. On this view speech is not self-regarding but rather helps to structure social life. Thus, for example, pornography, hate speech, and graphic television violence inculcate attitudes that make society more immoral, sexist, racist, lawless, or violent than it would be if a different rhetoric prevailed. Speech helps construct society by socializing behavior, and reconstructing society, in this view, requires regulating speech.

At stake in these debates is whether speech will continue to be understood, like religious and reproductive practices, as presumptively a matter for private resolution, or instead will be subject to greater government regulation in the pursuit of social ends, including that of maximizing the quantity or diversity of speech itself.

See also Democracy; Liberty; Mill, John Stuart.

Bibliography

Bork, R. "Neutral Principles and Some First Amendment Problems." *Indiana Law Journal* 49 (1971): 1.

Coetzee, J. M. *Giving Offense: Essays on Censorship.* Chicago: University of Chicago Press, 1996.

Dworkin, R. "The Coming Battles over Free Speech." *New York Review of Books,* 11 June 1992, 55.

Gates, H. L., Jr. "Let Them Talk." *New Republic,* 20 September 1993, 37.

Haworth, Alan. *Free Speech.* New York: Routledge, 1998.

MacKinnon, C. A. *Only Words.* Cambridge, MA: Harvard University Press, 1993.

Matsuda, M. J., et al. *Words that Wound: Critical Race Theory, Assaultive Speech, and the First Amendment.* Boulder, CO: Westview Press, 1993.

Meiklejohn, A. *Free Speech and Its Relation to Self-Government.* New York: Harper, 1948.

Mill, J. S. *On Liberty.* London: J. W. Parker and Son, 1859.

Scanlon, T. "A Theory of Free Expression." *Philosophy and Public Affairs* 1 (1972): 204.

Strossen, N. *Defending Pornography: Free Speech, Sex, and the Fight for Women's Rights.* New York: Scribners, 1995.

Sunstein, C. R. *Democracy and the Problem of Free Speech.* New York: Free Press, 1993.

Kathleen M. Sullivan (1996)

Bibliography updated by Philip Reed (2005)

CHAADAEV, PËTR IAKOVLEVICH

(c. 1794–1856)

Pëtr Iakovlevich Chaadaev was a Russian thinker and writer. He was a member of the old nobility (his mother's father was the celebrated historian Mikhail Mikhailovich Shcherbatov [1733–1790]). He studied at Moscow University and participated in the great war of 1812 and in the subsequent campaign against Napoleon Bonaparte in Europe. In 1816–1817, while an officer in the Hussars, he met and became friends with Aleksandr Sergeevich Pushkin (1799–1837), who in his young years dedicated three letters in verse to Chaadaev. In 1821 Chaadaev resigned from military service, cutting short what had promised to be a brilliant career. From 1823 to 1826 he traveled in Europe (England, France, Italy, Switzerland, and Germany), where he became acquainted with Friedrich Wilhelm Joseph von Schelling and Hugues Félicité Robert de Lamennais, whose religious-philosophical ideas made a profound impression on him. At that time he also became friendly with a number of representatives of certain European religious sects, who were adherents of Catholic socialism. The acquaintance with European culture, social heritage, and ideas precipitated a spiritual crisis in Chaadaev: the transition from Enlightenment deistic beliefs about the universe to a modern version of Christianity, consisting in a syncretic union of religion, philosophy, history, sociology, natural science, art, and literature.

After his return Chaadaev wrote (from 1829 to 1831) his main work: *Lettres philosophiques*. It was written in French and consisted of eight treatises in the form of letters addressed to a lady. This work signified the start of an original Russian philosophy, as well as the formation of a new worldview for Chaadaev. Here, Chaadaev attempted to develop a religious justification for the social process. The establishment of a "perfect order on earth" is possible, in his opinion, only by means of the direct and constant action of "Christian truth," which, through the continuous intellectual interaction of many generations, forms the foundation of "the universal-historical tradition" in the movement of social history and facilitates "the education of the entire human race" (1991 Vol. 1, p. 644). In Chaadaev's view this social idea of Christianity evolved, first, in Catholicism. This idea defined, as Chaadaev points out in the first letter, "the sphere in which Europeans live and in which alone under the influence of religion the human race can fulfill its ultimate purpose" (p. 652).

From this premise Chaadaev infers that European successes in the domains of culture, science, law, and material progress were the fruits of Catholicism as a socially active, political religion; and therefore these successes could serve as the starting point of a higher synthesis. The interpretation of Christianity as a historically progressive social development became for Chaadaev the foundation of a critique of the contemporary Russian situation. In Russia Chaadaev found neither "elements" nor "embryonic indications" of European progress. In his opinion the reason for this was that, when it initially separated from the Catholic West, Russia "erred concerning the true spirit of religion": Russia did not recognize "the purely historical side," that is, the socially transformative principle, to be an inner property of Christianity (658). The consequence of this was that Russia lagged behind Europe and had not gathered "all the fruits" of science, of culture, of civilization, of a well-ordered life. Chaadaev believed that, for Russia to achieve the successes of European society, it was insufficient for it simply to adopt the European forms of development: It had to change everything from the beginning, by repeating, under the flag of the salvific Catholic idea, the entire history of western Europe.

The first "Philosophical Letter" was published in the Moscow journal *Teleskop* (1836). This publication produced in thinking Russia an impression similar to a "rifle shot resounding in a dark night" (in the words of Alexander Ivanovich Herzen, 1954–1965). After its publication the journal was prohibited by the government, and its editor-publisher, N. I. Nadezhdin (1804–1856) was arrested and expelled from Moscow, while Chaadaev himself was declared, "by imperial order," to be insane. This "Philosophical Letter" was the only work of Chaadaev's to be published during his lifetime. Chaadaev's conclusions in this letter provoked a serious critique and disputation in circles of the Russian intelligentsia. Despite the official prohibition of the polemic around the *Philosophical Letters*, there were serious responses to them from Pushkin, P. A. Viazemskii (1792–1878), Aleksandr Ivanovich Turgenev (1784?–1846), Filip Filipovich Vigel (1786–1856), D. P. Tatishchev (1974–), Schelling, and others. By and large, these commentators did not agree with Chaadaev, but they recognized that it was legitimate and timely to formulate philosophical problems connected with solving the riddle of "the sphinx of Russian life" (in Herzen's words). Chaadaev's publication also provoked a serious split in Russian social life, a split that acquired the character of a dispute that, in principle, could never be resolved.

Although Chaadaev was prohibited from publishing his ideas, he continued his philosophical search. To accusations that he was insufficiently patriotic, he responded with the article "L'apologie d'un fou" (The apology of a madman; written in 1837 but first published in Paris in 1862), in which, speaking about Russia, he affirms that "we are called to solve most the problems of the social order, to answer the most important questions which preoccupy mankind" (1991 Vol. 1, p. 675). Here, he admits that the traditions of Orthodox Christianity possess indisputable merits and have played a beneficial role in the formation of the Russian mind. He is prepared to see Russia's calling in the fact that "at the proper time [it] would offer a solution to all the questions provoking disputation in Europe." In the 1840s Chaadaev's house in Moscow became the center of an important literary and philosophical circle.

Following in Chaadaev's footsteps, many Russian writers and philosophers became sufficiently bold to pose and ask into fundamentally important but hitherto systematically unexplored problems of social development. This exploration made it possible to clarify conceptions regarding the historical evolution of Russia, and it had a significant influence on the formation of the two fundamental trends in Russian social thought: the Westernizing orientation (Timofei Nikolaevich Granovskii [1813–1855], Vissarion Grigor'evich Belinski, Herzen, and Konstantin Dmitrievich Kavelin) and the Slavophile orientation (Aleksei Stepanovich Khomiakov, Ivan Vasil'evich Kireevskii, Konstantin Sergeevich Aksakov [1817–1860], and Yu. F. Samarin [1819–1876]. Chaadaev himself found a common language with representatives of both camps, although he also critiqued both; at various times he was invited to contribute to journals that held diametrically opposed positions.

Chaadaev's ideas on the philosophy of history proved to be a stimulus for such different thinkers as Khomiakov, Herzen, Apollon Aleksanrovich Grigor'ev (1822–1864), Konstantin Nikolaevich Leont'ev, Nikolai Iakovlevich Danilevskii (1822–1865), and Vladimir Sergeevich Solov'ëv (Solovyov). In essence, these ideas marked the start of the development of an original Russian philosophy.

Chaadaev's esthetic judgments reflected the influence of his "one idea"; they are subordinate to the moral ideal worked out by him. For Chaadaev, beauty in art is inseparable from truth and goodness. The artist is a guide leading people toward endless perfection; in transient things the artist discerns the milestones on this path. Somewhat paradoxically, Chaadaev condemned the art of

antiquity, in which, he believed, "all the moral elements were chaotically confused" (1991 Vol. 1, p. 359). In contrast, Gothic art was, for Chaadaev, "something sacred and heavenly," serving as an expression of moral feelings and compelling man "to lift his gaze toward heaven" (p. 359). In contemporaneous letters Chaadaev valued Nikolai Vasil'evich Gogol's (1809–1852) *Selected Passages from a Correspondence with Friends* (1846), in which "among weak and even sinful pages there are pages of astonishing beauty, full of infinite truth" (1991 Vol. 2, p. 1991). Chaadaev's aesthetic judgment was defined by his moral creed: "[M]oderation, tolerance, and love for all that is good, whatever form it might take" (p. 200).

Chaadaev's legacy was most accurately assessed by Khomiakov, who wrote in 1860:

> An enlightened mind, an artistic feeling, a noble heart—those are the qualities that attracted everyone to him. But at a time when it appeared that Russian thought had become submerged in heavy and involuntary sleep, he was especially valuable to us because he was awake and awakened others, because in the thickening darkness of that time he did not allow the lamp of truth to go out.

See also Aesthetic Judgment; Belinskii, Vissarion Grigor'evich; Enlightenment; Herzen, Aleksandr Ivanovich; Kavelin, Konstantin Dmitrievich; Kireevskii, Ivan Vasil'evich; Lamennais, Hugues Félicité Robert de; Leont'ev, Konstantin Nikolaevich; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY CHAADAEV

Russian Philosophy, edited by James M. Edie, James P. Scanlan, and Mary-Barbara Zeldin, with George L. Kline, 101–154. Chicago: Quadrangle Books, 1965.

Philosophical Letters, and Apology of a Madman. Translated by Mary-Barbara Zeldin. Knoxville: University of Tennessee Press, 1969.

Polnoe sobranie sochinenii i izbrannye pis'ma (Collected works and selected letters). 2 vols. Moscow: Izd-vo "Nauka," 1991.

P. Ya. Chaadaev: Pro et contra. Antologiia (P. Ya. Chaadaev: Pro et Contra. An Anthology). St. Petersburg, Russia, 1998.

WORKS ABOUT CHAADAEV

Herzen, A. I. *Sobranie sochinenii v 30—ti tomakh* (Works in 30 volumes). Moscow: USSR Academy of Sciences, 1954–1965.

Khomiakov, A. S. *Polnoe sobranie sochinenii* (Complete works). Moscow: Universitetskaia tipografiia, 1900.

McNally, Raymond T. *Chaadayev and His Friends: An Intellectual History of Peter Chaadayev and His Russian Contemporaries*. Tallahassee, FL: Diplomatic Press, 1971.

Moskoff, Eugene A. *The Russian Philosopher Chaadayev: His Ideas and His Epoch*. New York: N.p., 1937.

Zenkovsky, V. V. *A History of Russian Philosophy*. 2 vols. Translated by George L. Kline. New York: Columbia University Press, 1953. Originally published in two volumes under the title *Istoriia russkoi filosofii* (Paris 1948–1950).

Viacheslav Koshelev (2005)
Translated by Boris Jakim

CHAIN OF BEING

See *Lovejoy, Arthur Oncken*

CHAMBERLAIN, HOUSTON STEWART
(1855–1927)

Houston Stewart Chamberlain, the Anglo-German race theorist and philosophical and historical writer, was born in Southsea, near Portsmouth, England. Despite his English birth and family, his early indifference toward England and all things English developed into a lifelong hatred. Chamberlain was brought up by relatives in France. After being forced to attend schools in England, he returned to England only briefly, in 1873 and 1893. A nervous breakdown determined the course of his physical and mental development. (Frequently ill, hypersensitive, neurotic, he was crippled during the last thirteen years of his life by an incurable paralysis.) He traveled in western and central Europe for nine years seeking a cure. A German tutor inspired him to turn his mind to German literature and philosophy, and eventually he chose Germany as his home. As early as 1876 he wrote, "My belief that the whole future of Europe—that is, of world civilization—is in Germany's hands has become a certainty" (*Lebenswege meines Denkens,* p. 59).

Chamberlain's intellectual development began with the study of botany and other natural sciences; this was soon completely supplanted by a preoccupation with philosophy, literature, theology, art, and history. The turning point of his life was his meeting his future father-in-law, Richard Wagner, "the sun of my life," whom Chamberlain considered the greatest poet and musician of all time. Johann Wolfgang von Goethe inspired the central concept of Chamberlain's picture of the world and his "theory of life," the concept of *Gestalt* (form) as the

expression of all that is timeless and unchangeable. The *Gestalt* is encountered as the primary concept in the intuition of everything living (*Anschauung*) and must be grasped and interpreted in thought. It is the key to metaphysics and art, two fields which Chamberlain passionately defended against rationalism and "the coarsely empirical theory of evolution."

RACE

Chamberlain's "Lebenslehre" (Theory of life), which he first drafted in 1896 (it was not published until 1928 and was then titled *Natur und Leben* [Nature and life]), presented the position of most of his later writings, a position to which he frequently sacrificed historical truth in *Die Grundlagen des 19. Jahrhunderts* (*Foundations of the Nineteenth Century*), his weakest but best-known work. Chamberlain upheld "Life," intuition, metaphysics, "holy art" in the Wagnerian sense, and antidemocratic thought against rationalism, biological materialism (of Jewish origin), the superficial belief in progress, and moral decadence. His *Weltanschauung*—a favorite word of Chamberlain's—is closely related to Wagner's theory of decadence and regeneration. It carries with it the urge to improve the world, and Chamberlain felt himself called into the battle for moral renewal not of humanity in general (he spoke derogatorily of "the ghost, humanity"), but of the Teutonic culture and people. To save culture from the threat of materialism was also the declared aim of his books on Immanuel Kant and Goethe.

In the *Grundlagen* Chamberlain represented history as a conflict of opposing philosophies of life, represented by the Jewish race on the one hand and by the Germanic-Aryan race on the other. The application of the biological idea of race to the study of cultural phenomena was widespread around the turn of the twentieth century. Under the influence of Charles Darwin, it was used by anthropologists, ethnologists, religious historians, and others. It could serve both as a basis for scholarly interpretation and as a vehicle for racism, following the example of Comte Joseph Arthur de Gobineau. It was natural for Chamberlain to take over the concept of race from his scientific studies, but the significance he gave to it went beyond what was tenable in the light of the scientific knowledge then available and even denied the relevance of scientific criticism: "Even if it were proved that there had never been an Aryan race in the past, we are determined that there shall be one in the future; this is the decisive point of view for men of action" (*Grundlagen,* 1st ed., Vol. I, p. 270). Intuition and instinct, an overwhelming irrationalism, the capacity to sweep away logical contradictions—these are the major characteristics of this "historical" work.

Without ever giving a precise definition of "race," Chamberlain considered it to be the "*Gestalt* in particular, transparent purity" (*Natur und Leben,* p. 152) "Only thoroughbred 'races,'" he held, "accomplish the extraordinary" (*Rasse und Persönlichkeit,* p. 75). In connection with his race theory, Chamberlain emphasized the significance of nations: "It is almost always the nation as a political entity that creates the conditions for the formation of a race, or at least for the highest expressions of the race" (*Grundlagen,* 1st ed., Vol. I, p. 290). The awareness of racial identity, not physical characteristics, determined a race. Thus Chamberlain could speak of the English or Japanese "races" and also employ the term in a very broad sense, as when he included the Slavs and Celts among the Teutonic peoples.

Race was always dominant in Chamberlain's thought, whether he was describing the "heritage of the old world" as Hellenic art and philosophy, Roman law, and the coming of Christ; the cultureless chaos of peoples which separated the ancient from the modern world; or the role of the Jews and the Teutonic peoples, who entered Western history as "pure" races and whose antagonism shaped the modern world. He recognized the existence of other historical forces, such as religion or the desire for power, but he placed them far below race in importance. He was thus led to the paradox of trying to prove that the historical Jesus, whose birth he regarded as "the most important date in the entire history of humanity," was not a Jew. Chamberlain denied that the Jewish people possessed any metaphysical inclinations or philosophical tendencies. Their outstanding characteristics in his view were materialism and rationalism. They were thus incapable of religion and could not have produced the man Jesus. The Jews served Chamberlain as a dark foil for the image of the Germanic peoples, whom he celebrated as the creators of "all present culture and civilization" and whose standard-bearers were the Germans. Paul Joachimsen, in a memorial article, described the aim of the *Grundlagen* as "to demonstrate the elements of Western cultural development in the light of an Aryan theodicy." But whereas Joachimsen considered Chamberlain's work as a document already belonging to the past, we know today what terrible consequences his ideas had when they were translated into reality after his death. The chief ideologist of National Socialism, Alfred Rosenberg, showed himself to be Chamberlain's disciple in his *Mythus des 20. Jahrhunderts* (Myth of the twentieth century).

GOETHE AND KANT

One must not interpret Chamberlain's personality exclusively by the *Grundlagen.* His philosophical books on Kant and Goethe provide a far more solid basis for judgment and are more representative of his inclination and his intellectual position. His *Goethe* (1912) is a milestone in studies of the poet. Chamberlain was concerned to present "a clear, enthusiastic, and at the same time a critically reflective, grasp of this great personality in its essence and effect." Chamberlain found in Goethe the same polarities which he found in himself: nature and freedom, intuition and concept, poet and scholar, Christian and pagan—in brief, "the juxtaposition of opposed vocations." Jean Réal rightly described *Goethe* as "full of originality, of depth, and of prejudice" ("Houston Stewart Chamberlain et Goethe").

Chamberlain interrupted his studies of Goethe, which he pursued for more than twenty years, in order to write his *Immanuel Kant* (1905). Through Kant's limitation of the possibility of metaphysics Chamberlain came to realize the place of religion in human life. This side of Kant's thought appealed to Chamberlain's antirationalistic, vitalistic tendencies.

During World War I, Chamberlain composed fanatical anti-English propaganda. He was an intimate of Kaiser Wilhelm II from 1901 until well into the kaiser's exile in the Netherlands. He was quite naturally unable to come to terms with the Weimar Republic and turned his sympathies to Adolf Hitler, whom he first met in 1923. *Mensch und Gott* (Man and God), written in Chamberlain's old age, is an impressive attempt at a philosophical synthesis but casts no light on his personality as a whole. One can agree with the judgment of Friedrich Heer in *Europa—Mutter der Revolutionen* (Stuttgart, 1964, p. 6): "H. S. Chamberlain presents himself as a highly significant symbol combining high culture and barbarism."

See also Darwin, Charles Robert; Gobineau, Comte Joseph Arthur de; Goethe, Johann Wolfgang Von; Kant, Immanuel; Racism.

Bibliography

PRIMARY WORKS

Richard Wagner. Translated by G. Ainslie Hight. Munich: Bruckmann, 1897.

Die Grundlagen des 19 (Foundations of the Nineteenth Century), 2 vols. Translated by John Lees. London: Lane, 1911.

Goethe. Munich: Bruckmann, 1912.

Immanuel Kant. Die Persönlichkeit als Einführung in das Werk, 2 vols. Translated by Lord Redesdale. London: Lane, 1914.

Deutsches Wesen. Ausgewählte Aufsätze. Munich: Bruckmann, 1916.

Lebenswege meines Denkens (My thought's path through life). Munich: Bruckmann, 1919.

Mensch und Gott. Betrachtungen über Religion und Christentum. Munich: Bruckmann, 1921.

Rasse und Persönlichkeit. Aufsätze. Munich: Bruckmann, 1925.

Briefe 1882–1924 und Briefwechsel mit Kaiser Wilhelm II. Edited by Paul Pretzsch, 2 vols. Munich: Bruckmann, 1928.

Natur und Leben. Edited by Jakob von Uexkiill. Munich: Bruckmann, 1928.

Foundations of the Nineteenth Century, 2 vols. Translated by John Lees. New York: Fertig, 1968.

SECONDARY WORKS

Baylen, Joseph O., and Rolf F. Munster. "Adolph Hitler as Seen by Houston Stewart Chamberlain: A Forgotten Letter." *Duquesne Review* 12 (2) (1967): 81–88.

Biddiss, Michael D. "History as Destiny: Gobineau, H. S. Chamberlain, and Spengler." *Transactions of the Royal Historical Society* 7 (1997): 73–100.

Chamberlain, Anna. *Meine Erinnerungen an Houston Stewart Chamberlain.* Munich: Beck, 1923.

Field, Geoffrey G. *Evangelist of Race: The Germanic Vision of Houston Stewart Chamberlain.* New York: Columbia University Press, 1981.

Joachimsen, Paul. "Houston Stewart Chamberlain." *Zeitwende* 3 (1927): 347–361, 430–439.

Lindner, Erik. "Houston Stewart Chamberlain: The Abwehrverein and the 'Praeceptor Germaniae,' 1914–1918." *Leo Baeck Institute Year Book* 37 (1992): 213–236.

Réal, Jean. "Houston Stewart Chamberlain et Goethe." *Études germaniques* 5 (1950): 154–166.

Réal, Jean. "La lettre à l'amiral Hollmann (1903) ou Guillaume II à l'école de Houston Stewart Chamberlain." *Études germaniques* 6 (1951): 303–312.

Réal, Jean. "The Religious Conception of Race: Houston Stewart Chamberlain and Germanic Christianity." In *The Third Reich; A Study under the Auspices of the International Council for Philosophy and Humanistic Studies with the assistance of UNESCO.* New York: Praeger, 1955.

Seillière, Ernest. *Houston Stewart Chamberlain, le plus récent philosophe du pangermanisme mystique.* Paris: La Renaissance du livre, 1917.

Stackelberg, Roderick."Houston Stewart Chamberlain and Idealized Racism." In *Idealism Debased.* Kent, OH: Kent State University Press, 1981.

Stolberg-Wernigerode, Otto Graf zu. "Houston Stewart Chamberlain." In *Neue deutsche Biographie,* Vol. 3. Edited by Historische Kommission bei der Bayerischen Akademie der Wissenschaften. Berlin: Duncker and Humblot, 1957.

Ziegenfuss, Werner. "Houston Stewart Chamberlain." In *Philosophen-Lexikon,* Vol. 1. Edited by Werner Ziegenfuss. Berlin: de Gruyter, 1949.

Antonia Ruth Schlette (1967)
Translated by Tessa Byck
Bibliography updated by Philip Reed (2005)

CHANCE

Much is asked of the concept of *chance*. It has been thought to play various roles, some in tension, or even incompatible, with others. Chance has been characterized negatively as the absence of causation; yet also positively—the ancient Greek "tyche" reifies it—as a cause of events not governed by laws of nature, or as a feature of laws of nature. Chance events have been understood epistemically as those whose causes are unknown; yet also objectively as a distinct ontological kind, sometimes called "pure" chance events. Chance gives rise to individual unpredictability and disorder; yet it yields collective predictability and order: stable long-run statistics and, in the limit, aggregate behavior susceptible to precise mathematical theorems. Some authors believe that to posit chances is to abjure explanation; yet others think that chances are themselves explanatory. During the Enlightenment, talk of chance was regarded as unscientific, unphilosophical, the stuff of superstition or ignorance; yet at the beginning of the twenty-first century it is often taken to be a fundamental notion of our most successful scientific theory, quantum mechanics, and a central concept of contemporary metaphysics.

Chance has both negative and positive associations in daily life. The old word in English for it, "hazard," which derives from French and originally from Arabic, still has unwelcome connotations of risk; "chance" evokes uncertainty, uncontrollability, and chaos. Yet chance is also allied with luck, fortune, freedom from constraint, and diversity. And it apparently has various practical uses and benefits. It forms the basis of randomized trials in statistics, and of mixed strategies in decision theory and game theory; it is appealed to in order to resolve problems of fair division and other ethical stalemates; and it is even thought to underpin biological and cultural adaptation. Throughout history, "chance" devices have been a source of entertainment, as well as of scorn.

A BRIEF HISTORY OF THEORIES OF CHANCE

The study of gambling games motivated the first serious mathematical study of chance by Blaise Pascal and Pierre de Fermat in the mid-seventeenth century, culminating in the *Port Royal Logic*. But inchoate ideas about chance date back to antiquity. Epicurus, and later Lucretius, believed that atoms occasionally underwent uncaused, indeterministic swerves—an early doctrine of pure chance. Aristotle, by contrast, believed that all events are necessary and regarded what we call coincidences (as in "We met at the market place by chance") as the intersections of independent deterministic causal chains—a view later shared by Thomas Aquinas, Antoine Augustin Cournot, and John Stuart Mill. Augustine believed that God's will controls everything, and thus that nothing happens by chance. In the middle ages, Averroes had a notion of "equipotency" that arguably resonated with Gottfried Wilhelm Leibniz's and later Pierre Simon de Laplace's ideas about "equipossibility," which undergirded their classical interpretation of probability: The probability of an event is the ratio of the number of equipossible cases in which it occurs to the total number of such cases. Girolamo Cardano, Galileo, Fermat, and Pascal also anticipated this interpretation.

Throughout the development of probability theory during the seventeenth through nineteenth centuries by authors such as Christian Huygens, Jakob Bernoulli, Thomas Bayes, Pierre Simon de Laplace, the Marquis de Condorcet, Abraham de Moivre, and John Venn, the fortunes of chance were at best mixed. De Moivre called chance "a mere word." David Hume captured the attitude of his time when he wrote, "'Tis commonly allowed by philosophers that what the vulgar call chance is nothing but a secret and conceal'd cause" (Hume 1975, p. 130). The triumphs of Newtonian mechanics engendered great confidence in determinism, personified by Laplace's image of an intelligent being (the so-called "Laplacean demon") for whom "nothing would be uncertain and the future, as the past, would be present to its eyes" (Laplace 1951, p. 4). Eliminativism about chance in nature had, moreover, good theological credentials: God's omniscience apparently made the world safe for determinism. But even the atheist Bertrand Russell insisted that a chance event is merely one whose cause is unknown. F. H. Bradley found the very notion of chance unintelligible.

Nonetheless, other intellectual developments set the stage for a revival of chance. With the burgeoning of social statistics in the nineteenth century came a realization that various social phenomena—births, deaths, crime rates, etc.—while unpredictable on an individual basis, conformed to large-scale statistical regularities. A somewhat analogous pattern of collective order from individual chaos appeared in statistical mechanics. The social sciences and then the physical sciences thus admitted statistical laws into their conceptual repertoire. This culminated in the early twentieth century with the advent of quantum mechanics, which appeared to show that chance was irreducible and ineradicable. Andrey Kolmogorov's axiomatization of probability came soon after

Werner Heisenberg and Erwin Schrödinger brought quantum mechanics to its apogee.

Meanwhile, chance was also making a comeback in philosophy. Charles Sanders Peirce defended pure chance on the basis of empirical evidence. William James saw the postulation of chance as a way to resolve the apparent conflict between determinism and free will. To be sure, philosophers such as John Stuart Mill, Moritz Schlick, and C. D. Broad thought that capricious chance could provide no ground for genuine freedom. Nevertheless, chance had regained its respectability. In the 1950s Hans Reichenbach's work on probabilistic causation placed chance in the limelight in the philosophy of science.

THE MATHEMATICS OF CHANCE

The mathematics of chance, unlike its philosophy, is relatively uncontroversial. That mathematics is widely taken to be probability theory. In Kolmogorov's theory (1933/1950), events are assigned numerical values between 0 and 1 inclusive:

$$P(X) \geq 0$$

$$P(\Omega) = 1$$

(Here Ω is the universal set of all possible outcomes.) The probability of one of two mutually exclusive events occurring is the sum of their probabilities:

$$P(X \cup Y) = P(X) + P(Y) \text{ if } X \cap Y = \emptyset$$

(This law has an infinite generalization.) And the conditional probability of *A* given *B* is as follows:

$$P(A|B) = P(A \cap B)/P(B) \text{ for } P(B) > 0$$

While Kolmogorov's theory remains the orthodoxy, some philosophers (e.g., James Fetzer, Paul Humphreys, Karl Popper) question its appropriateness for chance.

CHANCE IN SCIENCE

Probability was introduced into physics in the late nineteenth century, when James Clerk Maxwell and Ludwig Boltzmann grounded thermodynamics in statistical mechanics. The status of this probability was an important interpretive issue, but it was not universally regarded as objective chance. Statistical mechanics was based on Newtonian particle mechanics, which was apparently deterministic. There are profound and ongoing controversies over the existence and nature of chance in both statistical mechanics and quantum mechanics.

In nonrelativistic quantum mechanics, according to the canonical Copenhagen interpretation, there are two rules for the evolution of a physical system:

- *Schrödinger's equation* prescribes a deterministic evolution for the state of the system. Typically, the state is a superposition (combined state) of the various definite-property states that the system might possess (e.g., definite position, definite momentum, etc.). While the system is in a superposition, it has no single value for such quantities.
- The *collapse postulate* is where chance enters quantum mechanics. Upon measurement of such a superposition, the state instantaneously collapses to one of the quantity's eigenstates (definite-property states). Which one is a matter of chance, the probability for each being derivable by Born's rule.

Albert Einstein considered this intrusion of chance into microphysics an unacceptable violation of causality and hoped for an underlying deterministic theory, with hidden variables, that explains the apparently chancy behavior of quantum systems. In 1935, Einstein, Boris Podolsky, and Nathan Rosen (EPR) insisted that there must be such an underlying theory, arguing that the quantum-mechanical description of a certain two-particle system is incomplete. Neils Bohr and Werner Heisenberg effectively criticized the EPR argument, and since an experimental test of an EPR pair of particles appeared to be physically unrealizable, most physicists quickly forgot the debate.

In 1952 David Bohm proposed a variant of the EPR setup using two coupled particles with correlated spins. Bohm's variant was both immune to the criticisms of Bohr and Heisenberg and physically realizable. In 1965 John Bell proved a now-legendary theorem stating that no local hidden-variable theory, of the type desired by Einstein, could replicate the statistical predictions of quantum mechanics for the correlated spins. Contrary to what the EPR paper had assumed, an underlying hidden-variable theory that assigned definite local values of spin to individual particles was incompatible with the predictions of quantum mechanics. Physicists then realized that a decisive experimental test was possible, and numerous experiments were performed in the 1970s, culminating in Alain Aspect's 1982 experiments, widely regarded as decisive. Nature sided with Bohr and Heisenberg, not Einstein.

Ironically, however, this confirmation of the predictions of quantum mechanics did not definitively show

that God plays dice, to use Einstein's memorable phrase. In 1952 Bohm also formulated a hidden-variable variant of quantum mechanics that ascribes definite positions to all particles at all times, reproduces all the experimental predictions of standard quantum mechanics, and is perfectly deterministic. This is consistent with Bell's theorem. No *local* hidden-variable theory can match the predictions of quantum mechanics for coupled particles, but Bohm's version of quantum mechanics is *nonlocal*: A particle in one place may be affected, instantaneously, by distant events. Einstein would have approved of Bohm's theory for its deterministic microphysics and disapproved of it for violating the even more cherished precept of no nonlocal interactions.

There are other versions of quantum mechanics besides Bohm's that reject chancy collapses. It is thus unclear whether the success of quantum-mechanical theories implies a fundamental indeterminism in nature, and whether future experiments can resolve the issue.

Evolutionary biology is another area of science in which the existence and role of chance has been sharply debated. Evolutionary fitness is held by some philosophers and biologists to be fundamentally chancy, while others disagree.

PHILOSOPHICAL ACCOUNTS OF CHANCE

Now, at the beginning of the twenty-first century, "chance" is typically taken to be synonymous with "objective probability," as distinguished from epistemic or subjective probability. *Frequentists*, originating with Venn, identify chance with relative frequency. For example, the chance that a particular coin lands heads is the frequency of tosses on which it so lands, divided by the total number of tosses. If we restrict ourselves to actual outcomes, then such frequencies will presumably be finite. A concern is that the outcomes may ill-reflect the true chances; a fair coin may land heads nine times out of ten. At the extreme, the *problem of the single case*, various events are unrepeatable, yet arguably have nontrivial chances (e.g., the outcome of the next presidential election). In such cases, mismatch between chance and relative frequency is guaranteed. Sometimes we might include in the reference class for a given event various other events. For example, regarding your chance of getting cancer, the class might include various other people like you. But there may be competing classes that yield different relative frequencies. You may belong both to the class of smokers and the class of those with no family history of cancer. What, then, is the real chance? This is the *problem of the reference class*.

Some frequentists follow Richard von Mises in requiring the sequences of trials that ground chances to be *infinite*, and thus presumably hypothetical. Then the chance of an outcome type is identified with its *limiting* relative frequency. (Further randomness constraints might also be imposed on the sequences.) Counterintuitively, such "chances" are then sensitive to the ordering of the trials (a sequence with infinitely many heads and tails can be rearranged to give whatever limiting relative frequency we like). Moreover, the appeal to hypothetical trials, let alone infinitely many of them, may betray the empiricist and scientific scruples that made frequentism initially appear attractive, for such "chances" are not constrained by anything in our experience.

Historically associated with Peirce and Popper, *propensity* accounts of chance postulate primitive dispositions, or tendencies, possessed by various physical systems. Propensity theories fall into two broad categories. According to *single-case* propensity theories, propensities measure the tendencies of a system to produce given outcomes; according to *long-run* propensity theories, propensities are tendencies to produce long-run outcome frequencies over repeated trials. The former have been advocated by the later Popper, David Miller, and James Fetzer; the latter by the early Popper, Paul Humphreys, and Donald Gillies.

Adopting a long-run view answers a need for testability of propensity attributions, one arguably found wanting for single-case propensity attributions. A long-run attribution may be held falsified if the long-run statistics diverge too much from those expected. However, defining propensities in terms of long-run relative frequencies may render single-case chance attributions problematic. This poses a dilemma for the long-run propensity theorist. If propensities are linked too closely to long-run frequencies, the view risks collapsing into a variant of frequentism. But if the view is cast so as to make single-case chance attributions possible, it risks collapsing into a variant of the single-case propensity view.

Long-run propensity theories may be motivated by the worry that in a single case there can be factors present that are not part of the description of the chance setup but that affect the chances of various outcomes. If the long-run propensity theorist responds by, in effect, falling back on long-run frequentism, the single-case propensity theorist goes the other way, embracing all causally or physically relevant details as part of the chance setup, determining the single-case chance (though we cannot measure it) for any given trial. The chance of each outcome is determined by everything that might influence

the evolution of the setup. Propensity theories of this type respect some of our physical and causal intuitions, but pay a price epistemically. Since each single-case setup is presumably unique, we cannot use frequencies to estimate the chances or to falsify hypotheses about them.

A final problem, specifically for *conditional* propensities, is *Humphreys' paradox*. If $\Pr(A|B)$ is a propensity, it seems to have a built-in causal direction, from B to A; the "inverse" conditional probability $\Pr(B|A)$ can often be calculated, but it appears to get the causal direction wrong. Various authors argue that inverse probabilities cannot be considered propensities, earlier events not having propensities to arise from later events. Thus, not all conditional probabilities may be interpretable as propensities.

While frequentist and propensity theories have dominated philosophical accounts of chance, a recent recurring proposal is that "chance" be viewed as a theoretical term similar to others in the sciences, such as "mass" or "fitness." In this post-positivist era, philosophers mostly agree that such terms cannot be reduced to non-theoretical terms. Instead, we may view theoretical terms as implicitly defined by their roles in scientific and philosophical theories. This approach avoids many of the difficulties discussed above, but it may not satisfy philosophers who find something troubling about the very notion of chance (see below). It also renounces giving a philosophical account of chance with *normative* status—claiming, for example, that theorists *should* admit objective chances into quantum mechanics but not into economics.

Pioneering work by David Lewis on the connections between chance and *credence* (subjective probability) has inspired *Humean best-system* theories. They share these tenets:

- Chances are defined so that their distinctive connection with credences is rendered transparent (see "Chance and Credence" below).
- Chances supervene on (are determined by) the entire history of actual events, and not on anything modal that does not itself supervene on the actual.
- Chances are determined by the laws of nature: the regularities of a best system (theory) that optimizes the balance of simplicity, strength (covering as many phenomena as possible), and fit (how typical actual events are, given the chances posited by the system).

Humean best-system accounts aim to be as acceptable to empiricists as finite frequentism, while avoiding the defects of that account.

CHANCE AND CREDENCE

Perhaps the most crucial demand we make of chances is that they guide our bets, expectations, and predictions—that they be guides to life in the face of uncertainty. This role is captured by some chance-credence principle or other, the most common coinage recently being Lewis's Principal Principle (Lewis 1986, p. 83–132):

(PP) $Cr(A|ch(A) = x \;\&\; E) = x$

Here Cr is one's credence function, A is a proposition, $ch(A)$ is the chance of A (presumably time-indexed), and E is further evidence that one may have. For (PP) to be applicable, E cannot be relevant to whether A is true or false, other than by bearing on the chance of A. (PP) codifies something crucial about chance. A touchstone for any theory of chance is that it should underwrite (PP). There is considerable controversy over which theory (if any) can meet this challenge.

CHANCE AND DETERMINISM

Determinism is the thesis that any complete past or present state of the world, conjoined with the laws of nature, entails all future events. In a deterministic world, some insist, chance has no work left to do, the entire future being already determined by past events. Philosophers are divided over whether determinism rules out (nontrivial) chances. Since the definition of determinism says nothing about chance, more is needed to argue that determinism rules out chances.

D. H. Mellor, Popper, and others who view propensities as fundamental physical loci of indeterminism, see an immediate inference from determinism to the nonexistence of chances. Frequentists such as Venn and Reichenbach see no such inference: intermediate frequencies can exist in both deterministic and indeterministic worlds. The Humean best-system approach leaves open whether a deterministic system of laws can include chance laws (although Lewis rejects this possibility). And on the implicit-definition approach, intermediate chances and determinism coexist just in case our fundamental physical theories are deterministic but some scientific theory postulates objective probabilities. Statistical mechanics uses chances, but its underpinnings are deterministic, and typical uses of chance in biology and the social sciences involve no presumption for or against determinism, as Isaac Levi (1990) and others have argued.

Nor does indeterminism guarantee the existence of chances. Fundamental physical laws may fail to entail a unique future without being probabilistic. However, if these laws are probabilistic, as some interpretations of

quantum mechanics contend, then chances are apparently guaranteed on any but a skeptical/subjectivist view.

SUBJECTIVISM, SKEPTICISM ABOUT CHANCE, AND EXCHANGEABILITY

Chance is meant to play a certain theoretical role. It is a further matter what, if anything, actually *plays* this role. According to Bruno de Finetti, nothing does. "Probability does not exist," he said (1990, p. x), meaning that *chance* does not exist and that all probability is subjective. Skepticism about chance is easily assimilated to skepticism about kindred modal notions—possibility, counterfactuals, causation, laws of nature—that seem not to be straightforwardly reducible to nonmodal notions, in particular, notions congenial to an empiricist. And skepticism specifically about chance can be based on further arguments, for one can be skeptical not just about its modality, but also about its putative *degrees*. Subjectivists have also argued that chance is redundant, its alleged role being completely discharged by credences. Richard Jeffrey, Bas van Fraassen, Brian Skyrms, and others have developed subjectivist positions in the spirit of de Finetti.

Moreover, the mathematics of chance (unlike the other modal notions) permits a particular eliminativist gloss. A sequence of trials is said to be *exchangeable* with respect to a probability function if the probabilities of trial outcomes are invariant under finite permutations of trials; probabilities may be sensitive to the numbers of outcomes of each kind, but not to their ordering. De Finetti (1990) showed that when this condition is met, there is a unique representation of the probability distribution over the trials as an expectation of simpler probability distributions according to which the trials are independent and identically distributed. For example, if your credences over the results of repeated coin tossing are exchangeable, then it is as if you treat the trials as tosses of a coin of unknown bias, with credences over the possible biases. Subjectivists have argued that this delivers some of the supposed benefits of chance, without any questionable metaphysics.

CONCLUSION

Many of the perplexities about chance—its controversial metaphysics, its seeming resistance to reduction, its epistemological recalcitrance, etc.—are familiar from other modal notions. But chance has been handled in mathematics and philosophy with more precision than those other notions. In the process, still further perplexities have been born. For the foreseeable future, at least in the writings of philosophers and philosophically minded scientists, chance is probably here to stay.

See also Probability and Chance.

Bibliography

A BRIEF HISTORY OF THEORIES OF CHANCE

David, F. N. *Games, Gods, and Gambling: A History of Probability and Statistical Ideas*. Mineola, NY: Dover, 1962.

Franklin, James. *The Science of Conjecture: Evidence and Probability before Pascal*. Baltimore, MD: Johns Hopkins University Press, 2001.

Hacking, Ian. *The Emergence of Probability*. Cambridge, U.K.: Cambridge University Press, 1975.

Hacking, Ian. *The Taming of Chance*. Cambridge, U.K.: Cambridge University Press, 1990.

Hume, David. *A Treatise of Human Nature*, edited by L. A. Selby-Bigge, 2nd ed. Oxford: Clarendon Press, 1975.

Laplace, Pierre Simon. *A Philosophical Essay on Probabilities*. English edition 1951 (originally published 1814). New York: Dover Publications Inc.

THE MATHEMATICS OF CHANCE

Billingsley, Patrick. *Probability and Measure*. 3rd ed. New York: John Wiley and Sons, 1995.

Kolmogorov, Andrei N. *Grundbegriffe der Wahrscheinlichkeitrechnung*. Berlin: J. Springer, 1933. Translated as *Foundations of Probability*. New York: Chelsea, 1950.

Skyrms, Brian. *Choice and Chance: An Introduction to Inductive Logic*. 4th ed. Belmont, CA: Wadsworth, 1999.

CHANCE IN SCIENCE

Albert, David. *Time and Chance*. Cambridge, MA: Harvard University Press, 2000.

Bell, John S. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge, U.K.: Cambridge University Press, 1987.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777–780.

Sklar, Lawrence. *Physics and Chance*. Cambridge, U.K.: Cambridge University Press, 1993.

Sober, Elliott. *Philosophy of Biology*. 2nd ed. Boulder, CO: Westview Press, 2000.

PHILOSOPHICAL ACCOUNTS OF CHANCE

Fetzer, James H. *Scientific Knowledge: Causation, Explanation, and Corroboration*. Dordrecht, Netherlands: D. Reidel, 1981.

Fine, Terrence. *Theories of Probability*. New York: Academic Press, 1973.

Gillies, Donald. "Varieties of Propensity." *British Journal for the Philosophy of Science* 51 (2000): 807–835

Loewer, Barry. "David Lewis's Humean Theory of Objective Chance." *Philosophy of Science* 71 (2004): 1115–1125.

Mellor, D. H. *The Matter of Chance*. Cambridge, U.K.: Cambridge University Press, 1971.

Miller, David. *Critical Rationalism: A Restatement and Defence*. La Salle, IL: Open Court Press, 1994.

Sober, Elliott. "Evolutionary Theory and the Reality of Macro Probabilities." In *Probability in Science*, edited by Ellery Eells and James H. Fetzer. La Salle, IL: Open Court, 2005.

CHANCE AND CREDENCE

Lewis, David, "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*. Vol II., University of California Press, 263-293, 1980; reprinted with postscripts in *Philosophical Papers*, Vol. II. Oxford: Oxford University Press, 83-132, 1986. Lewis, David. "Humean Supervenience Debugged." *Mind* 103 (1994): 473–490.

Popper, Karl. "The Propensity Interpretation of Probability." *British Journal of the Philosophy of Science* 10 (1959): 25-42.

Vranas, Peter. "Who's Afraid of Undermining?" *Erkenntnis* 57 (2002): 151–174.

CHANCE AND DETERMINISM

Levi, Isaac. "Chance." *Philosophical Topics* 18 (2) (1990): 117–148.

Loewer, Barry. "Determinism and Chance." *Studies in History and Philosophy of Modern Physics* 32 (2001): 609–620.

Strevens, Michael. *Bigger than Chaos: Understanding Complexity Through Probability*. Cambridge, MA: Harvard University Press, 2003.

SUBJECTIVISM, SKEPTICISM ABOUT CHANCE, AND EXCHANGEABILITY

de Finetti, Bruno. *Theory of Probability*, Vol. 1. Chichester: Wiley Classics Library, John Wiley & Sons, 1990

Skyrms, Brian. "Bayesian Projectibility." In *Grue: The New Riddle of Induction*, edited by Douglas Stalker, 241–262. Chicago: Open Court, 1994.

WEB RESOURCES

"Chance." The Chance Project, Mathematics Department, Dartmouth College. Available from http://www.dartmouth.edu/~chance/.

Hájek, Alan. "Probability, Interpretations of." In *The Stanford Encyclopedia of Philosophy*, summer 2003 ed., edited by Edward N. Zalta. Available from http://plato.stanford.edu/archives/sum2003/entries/probability-interpret/.

Alan Hájek (2005)
Carl Hoefer (2005)

CHANNING, WILLIAM ELLERY

(1780–1842)

William Ellery Channing, America's most famous Unitarian minister, was described by Ralph Waldo Emerson as "one of those men who vindicate the power of the American race to produce greatness." Channing, born in Newport, Rhode Island, was graduated from Harvard in 1798. The following two years he spent as a tutor in Richmond, Virginia, and in private study. During this period he underwent a profound religious experience, and in 1801 he returned to Harvard for theological study. He was ordained the minister of Boston's Federal Street Congregational Church in 1803 and held this pastorate throughout his life. He died in Bennington, Vermont.

Channing was not an original or profound thinker, a systematic philosopher, or a great writer. His significance in the history of ideas lies in his representative influence, his achievement in expressing and synthesizing the diverse strands of thought that appeared in America at the end of the eighteenth and the beginning of the nineteenth centuries.

Although Channing was celebrated in his own lifetime as a man of letters (his critical essays on John Milton, Napoleon Bonaparte, and François Fénelon were widely read both here and abroad), his lasting reputation stands on his attempt to develop an "enlightened" religious faith for the Americans of his generation. Jonathan Edwards had responded to the spirit of the Enlightenment by employing the ideas of John Locke and Isaac Newton to revitalize Calvinist dogma. Channing employed the liberating spirit of eighteenth-century thought to free Christianity from an outmoded theology. "God has given us a rational nature," he said in his famous sermon "Unitarian Christianity" (1819), "and will call us to account for it." Without denying the authority of Scripture, Channing argued that men should "reason about the Bible precisely as civilians do about the Constitution under which we live." This rational approach to revelation led Channing to reject the "irrational and unscriptural doctrine of the Trinity." Substituting the moral perfection of God for the Calvinist conception of divine sovereignty, Channing also repudiated such doctrines as natural depravity and predestination. "It is not because his will is irresistible but because his will is the perfection of virtue that we pay him allegiance," Channing asserted. "We cannot bow before a being, however great and powerful, who governs tyrannically."

As a religious thinker Channing was liberal but not radical. Eighteenth-century skepticism had no place in his thinking. He was influenced considerably by Scottish "commonsense" philosophers, such as Adam Ferguson and Richard Price, and in his discourse "The Evidences of Revealed Religion" (1821) he relied heavily on the traditional arguments of William Paley in attempting to refute David Hume and assert the validity of miracles.

Channing is also important for his influence on the New England transcendentalists. Like Jean-Jacques Rousseau, whose writings he admired, he was partly an Enlightenment figure and partly a romantic. Channing's romanticism is most apparent in the sermon "Likeness to

God" (1828), in which he asserted that humankind discovers God not only through Scripture and rational inquiry but also through consciousness. Long before Emerson's famous essays were published, Channing was preaching that in all its higher actions the soul had "a character of infinity" and describing sin as "the ruin of God's noblest work." Despite the fact that Channing never professed enthusiasm for the "new views," the similarity between his conception of the divine potential in human nature and the later pronouncements of Emerson and Theodore Parker is unmistakable. The path to transcendentalism lay through Unitarianism, and it was Channing who helped to pave the way.

Finally, Channing is significant for his humanitarian influence. His belief in the parental character of God and the dignity of humanity provided an ideological base for humanitarian efforts, and he spoke out in favor of most of the reform causes of his day. His pamphlet against slavery, written in 1835, attracted wide attention. Although Channing always shied away from radical solutions to social disorder, no one was more influential in articulating the gospel of human dignity that nourished most American reformers before the Civil War.

See also Edwards, Jonathan; Emerson, Ralph Waldo; Enlightenment; Fénelon, François de Salignac de la Mothe; Ferguson, Adam; Hume, David; Locke, John; Milton, John; Newton, Isaac; Paley, William; Parker, Theodore; Price, Richard; Rousseau, Jean-Jacques.

Bibliography

The most usable edition of Channing's works is a one-volume edition (Boston, 1886). There are several full-length studies of Channing, including David Edgell's *William Ellery Channing: An Intellectual Portrait* (Boston: Beacon Press, 1955); Arthur Brown's *Always Young for Liberty* (Syracuse, NY: Syracuse University Press, 1956); and Robert L. Patterson's *The Philosophy of William Ellery Channing* (New York: Bookman Associates, 1952).

Irving H. Bartlett (1967)

CHAOS THEORY

A physical system has chaotic dynamics, according to the dictionary, if its behavior depends sensitively on its initial conditions, that is, if systems of the same type starting out with similar sets of initial conditions can end up in states that are, in some relevant sense, very different. But when science calls a system chaotic, it normally implies two additional claims: That the dynamics of the system is relatively simple, in the sense that it can be expressed in the form of a mathematical expression having relatively few variables, and that the geometry of the system's possible trajectories has a certain aspect, often characterized by a strange attractor.

Chaos theory proper, it should be noted, has its home in classical physics (and other kinds of dynamics that share the relevant properties of classical physics). The extent to which chaotic mathematics is fruitful in understanding the quantum realm is still a matter of debate.

SENSITIVE DEPENDENCE ON INITIAL CONDITIONS

In the popular imagination a chaotic system is one whose future state may be radically altered by the smallest of perturbations—as when the fluttering of a butterfly's wings creates a disturbance whose size is inflated to the point where it tips the meteorological balance on the other side of the globe, creating a tornado where there would otherwise have been none. Though the "butterfly effect" marvelously engages human fear and wonder at the unpredictability of things, it captures rather less completely what is interesting and distinctive about modern chaos theory.

The idea of an inherent unpredictability in human and other affairs due to the inflation of small disturbances is an old one. Swift wrote in *Thoughts on Various Subjects* (1711) that "A Wise man endeavors, by considering all Circumstances, to make Conjectures, and form Conclusions: But the smallest Accident intervening, (and in the Course of Affairs it is impossible to see all) doth often produce such Turns and Changes, that at last he is just as much in doubt of Events, as the most ignorant and unexperienced Person" (p. 415).

Modern mathematics is able to characterize the sensitivity of initial condition dependence in various ways that lie far beyond Swift's means. Notions such as the Liapunov exponent help to quantify the speed at which the trajectories of systems starting out with similar initial conditions will diverge. Measure theory quantifies something like the chance that a small initial difference will lead to a relatively large difference in outcome, in systems where not every small change makes such a difference. There is nothing here, though, that would have astounded Swift.

SIMPLICITY

The central insight of chaos theory is that systems governed by simple equations, that is, systems whose behav-

ior can be characterized by a small number of variables, called low dimensional systems, are often sensitive to initial conditions. At first blush this realization has a pessimistic cast. Most obviously it leads to the conclusion that even a simple dynamics may be unpredictable in the medium to long term, as which of two significantly different outcomes occurs may depend on such first details of the initial conditions as to lie beyond the resolving power any reasonable observational effort.

Somewhat less obviously certain kinds of sensitivity to initial conditions impede systematic dynamical understanding. A famous example closely connected to the origins of chaos theory is the three body problem, the task of elucidating all the properties of the dynamics of a three body system in Newtonian gravitational theory. In 1890 Henri Poincaré showed that three body systems can tend to chaos in the modern sense of the word, and concluded that a systematic treatment of three body dynamics would be difficult if not impossible.

Chaos can be an impediment to prediction and systematic understanding in low dimensional systems then. However, if low dimensional chaos is bad news for the study of systems known to have low-dimensional dynamics, it is good news for the study of systems known only to have chaotic dynamics. Traditionally such systems were modeled by complex equations, if at all; chaos theory introduces the serious possibility that these systems may be governed by equations with very few variables. Underlying the complex appearances may be a simple reality. The prospect of finding a hidden simplicity in such complex phenomena as turbulent flows, the weather, the movements of financial markets, and patterns of extinction is what most excites proponents of chaos theory. (Much the same prospect animates the advocates of catastrophe theory, the study of cellular automata, "complexity theory," and so on.)

To what extent can the nature of this hidden simplicity, if it exists, be divined? Given sensitive dependence on initial conditions, it is difficult to find the simple equation that best predicts the observed phenomena, since small errors in measuring initial conditions can make even the true model look like a bad predictor. More feasible is to infer some of the more interesting properties of the putative underlying law, such as the degree of sensitivity to initial conditions and certain geometrical aspects of the dynamics induced by the law (discussed below).

Under favorable conditions this information can be used to model accurately the behavior of chaotic systems to some extent—or at least that is the hope both of academic chaoticians and of those hoping to use the mathematics of chaos theory to beat the financial markets.

By far the boldest posit made in undertaking such work is the assumption that there is a simple dynamic law lying behind the subject system's complex behavior. For elaborate systems such as ecosystems and economies, the assumption of dynamic simplicity is often no more than a leap of faith; however, Strevens describes some circumstances in which ecosystems and some other complex systems have a low dimensional macrodynamics.

The Geometry of Chaos Trace the trajectory of a paradigmatically chaotic system through the space of possible states and the result is a complicated tangle of looping paths. It is the geometry of this tangle more than anything else—more even than sensitive dependence per se—that is distinctive of chaos (though there is disagreement as to which feature of the geometry is most important).

One especially striking feature of such trajectory tangles is their often-fractal structure: They cut out a shape in the space in which they are embedded so intricate that mathematicians ascribe it a fractional dimension. Such a shape is a strange attractor (strictly speaking an attractor only if it is a set of trajectories that systems starting from some points outside the attractor eventually join).

Many of the more interesting properties of chaotic systems can be understood as arising from the intricate geometry of the trajectory tangle. One well-known example is the appearance of "period-doubling cascades" in systems that are moving from a periodic to a chaotic regime of behavior: As some parameter affecting the system's dynamics is tweaked, the system first oscillates between two states, then between four states, then eight states, and so on, with shorter and shorter times between each successive doubling, until it goes chaotic. What is interesting about this behavior is that it turns up in many physically quite different kinds of systems, and that there are certain aspects of the period doubling, notably the rate at which the doubling increases, that are the same (in the limit) in these otherwise rather different systems. This universality in chaotic systems holds out the promise of understanding the behaviors of a considerable range of systems in terms of a single mathematical—in this case, it turns out, a geometrical—fact. So far however the wider significance of this understanding is unclear.

A more practical part of chaotic geometry is the use of limited data about the behavior of chaotic systems to reconstruct to a certain extent the geometry of the system's trajectory. Suppose that the behavior a chaotic sys-

tem is characterized by three variables, so that the system's "trajectory tangle" is a subset of three-dimensional space. Suppose also that only a single property of the system's dynamics can be observed, a function of the values of the three variables. In favorable conditions, this single set of observations can be used to recover the geometrical structure of the three-dimensional dynamics. Various predictions, quantitative and qualitative, can then be made from the recovered geometry.

This is a powerful technique, as it assumes no knowledge of the number or even the nature of the underlying variables. However its success does depend on, among other things, the simplicity assumption explained above: The technique supposes that there are no more than a small number of variables.

CHAOS AND PROBABILITY

The disorderly behavior of chaotic systems can be called "random" in a loose and popular sense. Might the behavior of at least some such systems be random in a stronger sense? The suggestion that chaos might provide a foundation for probabilistic theories such as statistical mechanics has been one of the more fruitful contributions of chaos theory to philosophy.

The best scientific theories of certain deterministic or near deterministic systems are probabilistic. Perhaps the most prominent examples are the systems characterized by statistical mechanics and population genetics; the simplest examples are various gambling setups such as a roulette wheel or a thrown die. The probabilistic characterization of these systems is apt because the various events that make up their behavior (die throws or deaths, for example) are patterned in characteristically statistical ways, that is, in ways that are captured directly by one or other of the canonical probability distributions in statistical theory.

The mathematics of chaos offers an explanation of the probabilistic aspect of these patterns, and so offers an explanation of the success of probabilistic theories applied to certain sorts of deterministic systems.

The explanation, or rather the family of explanations, is quite complex, but it can be loosely characterized in the following way. A paradigmatically probabilistic pattern has two aspects: A short term disorder, or randomness, familiar to every gambler, and a long term order that is quantified by the statistics characterizing a probability distribution, such as the one-half frequency of "heads" in a long series of coin tosses.

Chaotic systems are capable of producing probabilistic patterns because they are capable of producing both this short term disorder and the requisite kinds of long-term order. The short-term disorder is due to the sensitive dependence on initial conditions; the long-term order to other aspects of the "geometry of chaos," principally chaotic dynamics' resemblance to a "stretch-and-fold" process.

Nowhere near all chaotic systems, it should be noted, generate probabilistic patterns. Indeed this area of investigation is not, in a certain sense, mainstream chaos theory: There are no strange attractors or period-doubling cascades, though there is a characteristically chaotic geometry to the relevant trajectory tangles. As well as explaining the success of probabilistic theorizing in science, chaos has been put forward—for much the same reasons—as a foundation for the metaphysics of probability, on the principle that what explains the probabilistic pattern is deserving to a considerable extent of the name probability.

PHILOSOPHICAL SIGNIFICANCE

What is the philosophical significance of chaos? With respect to general philosophy of science, opinion is divided. Some philosophers, for example Stephen Kellert, have argued that chaos theory requires the abandoning of prediction as the touchstone of successful science, a new conception of the nature of scientific explanation, and the end of reductionism. Others, for example Peter Smith, have argued that these conclusions are too extreme, and that insofar as they are justified, chaos theory is not necessary for their justification, though it may well have brought to philosophy's attention problems previously wrongly ignored.

With respect to certain foundational questions about science, the significance of chaos is less controversial. The notion of determinism and (in the context of processes that are deterministic deep down) the notions of randomness and probability cannot be discussed without reference to work on dynamical systems since Poincaré that falls within the ambit—broadly conceived—of chaos theory.

See also Geometry; Philosophy of Physics; Poincaré, Jules Henri; Probability and Chance; Swift, Jonathan.

Bibliography

Bass, Thomas. *The Predictors*. New York: Henry Holt, 1999.

Belot, Gordon, and John Earman. "Chaos Out of Order: Quantum Mechanics, The Correspondence Principle and

Chaos." *Studies in History and Philosophy of Modern Physics* 28 (1997): 147–182.
Earman, John. *A Primer on Determinism*. Dordrecht: Reidel, 1986.
Kellert, Stephen H. *In the Wake of Chaos*. Chicago: Chicago University Press, 1993.
Ornstein, Donald, and Benjamin Weiss. "Statistical Properties of Chaotic Systems." *Bulletin of the American Mathematical Society* 24 (1991): 11–116.
Sklar, Lawrence. *Physics and Chance*. Cambridge, U.K.: Cambridge University Press, 1993.
Smith, Peter. *Explaining Chaos*. Cambridge, U.K.: Cambridge University Press, 1998.
Stewart, Ian. *Does God Play Dice? The Mathematics of Chaos*. Oxford, Blackwell, 1989.
Strevens, Michael. *Bigger than Chaos: Understanding Complexity through Probability*. Cambridge, MA: Harvard University Press, 2003.
Suppes, Patrick. "Propensity Representations of Probability." *Erkenntnis* 26 (1987): 335–358.

Michael Strevens (2005)

CHARDIN, PIERRE TEILHARD DE

See *Teilhard de Chardin, Pierre*

CHARRON, PIERRE
(1541–1603)

Pierre Charron, a skeptical philosopher and theologian, was born in Paris in a family of twenty-five children. He studied at the universities of Paris, Bourges, Orléans, and Montpellier and received a law degree from Montpellier in 1571. Sometime during his student years he became a priest. He was a successful preacher and theologian in southern France, serving as preacher in ordinary to Queen Margaret of Navarre and as a theological advisor and teacher in various dioceses in the Midi. In spite of his many worldly successes, he tried to retire to a monastic order in 1589 but was refused admittance because of his age.

During the 1580s Charron met Michel Eyquem de Montaigne in Bordeaux and became his close friend and disciple. Montaigne made Charron his intellectual heir, adopting Charron as his son. After Montaigne's death in 1592 Charron wrote his major works: *Les trois veritez* (Bordeaux, 1593), *Discours chrestiens* (Bordeaux, 1601; Paris, 1604), *De la sagesse* (Bordeaux, 1601), and *Petit traicté de sagesse* (written in 1603, published posthumously in Paris, 1606). These works were popular and were republished often in the seventeenth century, especially the skeptical *De la sagesse*, which was highly influential in disseminating skeptical views and arguments into philosophical and theological discussions and played an important role in the development of modern thought, libertinism, and fideism.

OPPOSITION TO CHARRON

Serious efforts to suppress and reject Charron's skeptical views were made by such figures as the Jesuit Father François Garasse, who in 1623 accused Charron of having supplied *le brèviare des libertins* and of having been a secret atheist trying to destroy religion. His work, which was first condemned in 1605, was seen as more dangerous than Montaigne's, partly because Charron was a professional theologian, partly because he wrote more didactically. Pierre Chanet, a Protestant medical doctor, published *Considerations sur la sagesse de Charon* (1643), an attempted Aristotelian refutation of Charron's skepticism about the possibility of knowledge.

Although Charron, like Montaigne, was attacked on many sides, his views were also defended and advanced by the so-called *libertins érudits*—Gabriel Naudé, Guy Patin, François de La Mothe Le Vayer, and Pierre Gassendi—and were supported in varying degrees as theologically orthodox by various French Counter-Reformation leaders. Pierre Bayle considered Charron an excellent and prime representative of fideistic Christian thought. Interest in and concern with Charron's views diminished in the eighteenth century, and he came to be considered a second-rate and derivative Montaigne whose style lacked the freshness and literary quality of his mentor's. In the light of more recent criticism suggesting that Montaigne was or might have been a sincere believer and that his skepticism was part of a theological movement of the period, Charron, too, has begun to be reexamined and reevaluated.

CHARRON'S VIEWS

The first statement of Charron's views was the *Trois veritez*, a tract against Calvinism and the views of its French leader, Philippe Duplessis-Mornay. The three truths Charron sought to establish were that God exists, that Christianity is the correct view of God, and that Catholicism is the true statement of Christianity. Most of this enormous work deals with the last claim. However, the work begins with a brief discourse on knowledge of God, developing skepticism about the possibility of human knowledge in this area, on the basis of both human rational limitations and the nature of God. One's

own capacities are so limited and unreliable that it is doubtful that one could really know anything in either the natural or the supernatural realm. God's nature is infinite and therefore surpasses all attempts to define or limit it. Hence, one cannot know, in rational terms, what God is. Thus, the greatest theologians and philosophers know as much or as little about God as do the humblest artisans. One's knowledge consists only of negative information, what God is not. In fact, Charron announced, "the true knowledge of God is a perfect ignorance about Him" (*Trois vritez*, 1595, p. 26).

Charron combined the skeptic's views about the inadequacy and unreliability of human knowledge with the mystic's and negative theologian's view that God is unknowable because he is infinite and then utilized this combination to attack atheism. The denial that God exists proceeds from some definition of God, from which absurd conclusions are then drawn. Such a definition can only be the result of human presumption, the attempt to measure divinity by human means, and, as such, is worthless, since atheists do not, and cannot, know what they are talking about.

Throughout the *Trois veritez* Charron argued principally in a negative way, trying to show that it is unreasonable not to believe in God, Christianity, and Catholicism and that the evidence adduced by opponents is unreliable or dubious. He often contended that opponents, usually Calvinists, had to base their case on the results obtained by the weak and miserable human capacities, employing these defective results as measures of divine truth.

DE LA SAGESSE

Charron's skeptical defense of the faith was made more explicit in *De la sagesse* and in his defense of that work, *Petit traicté de sagesse*. His major thesis was that since man cannot discover any truth except by revelation, morality should be based on following nature, except when guided by divine light. To support this thesis, Charron first put forth most of Montaigne's skeptical views in an organized fashion. One must first know oneself ("The true science and the true study of man is man," *De la sagesse*, book 1, chapter 1), and this involves knowing the limitations on what one can know. Charron presented the traditional skeptical critique of sense knowledge, questioning whether one possesses the requisite senses for gaining knowledge, whether one can distinguish illusions and dreams from veridical experience, and whether one can, in view of the enormous variability of sense experiences, determine which ones correspond to objective states of affairs. Next, he raised skeptical questions about one's rational abilities, contending that one possesses no adequate or certain criteria that enable one to distinguish truth from falsehood. He pointed out that in fact one believes things mainly as a result of passions and social pressures, not reasons and evidence. One actually functions a as beast and not as a rational being. Hence, one should accept Montaigne's contention that men possess no genuine principles unless God reveals them. Everything else is only dreams and smoke.

The second book of *De la sagesse* presents a discourse on the method for avoiding error and finding truth, in view of the human predicament. Charron's method closely resembles the one René Descartes set forth later: examine all questions freely and dispassionately, keep prejudice and emotions out of all decisions, develop a universality of mind, and reject any decisions that are in the slightest degree dubious. This skeptical method, Charron claimed, is of greater service to religion than any other there may be. It leads one to reject all dubious opinions until one's mind is "blank, naked and ready" to receive the divine revelation on faith alone. The complete skeptic will never be a heretic, since if he or she has no opinions, he or she cannot have the wrong ones. If God pleases to give him or her information, then the skeptic will have true knowledge. Until the skeptic receives the revelation, he or she should live by a *morale provisoire*, following nature. The last book of *De la sagesse* presents this theory of natural morality, showing how one ought to live as a skeptic and noble savage if one has no divine guidance.

De la sagesse was one of the first important philosophical works to be written in a modern language and to present a moral theory apart from religious considerations. Some considered the work a basic didactic statement of Pyrrhonian skepticism, challenging both traditional philosophical claims to knowledge and religious ones and thus preparing the ground for a thoroughly naturalistic view of human nature and conduct. Charron claimed that the argument in *De la sagesse* only represented part of his view, dealing with the human situation apart from divine guidance.

The overall theory stated in his various works, his ecclesiastical career, and the piety expressed in his *Discours chrestiens* suggest that he was a sincere fideist, who saw skepticism as a means of destroying the enemies of the true faith while preparing the soul for salvation.

The problem of interpreting Charron's views involves a larger issue, that of assessing the purport of the revival of skepticism in the Renaissance and the relation of this revival to Reformation and Counter-Reformation

thought. Skeptical thought, perhaps, played several different and possibly incompatible roles in the period. Both then and now, skeptics like Charron could provide the "rationale" both for antirational fideism and for irreligious naturalism.

See also Bayle, Pierre; Fideism; Gassendi, Pierre; La Mothe Le Vayer, François de; Montaigne, Michel Eyquem de; Naturalism; Reformation; Renaissance; Skepticism, History of.

Bibliography

WORKS BY CHARRON

Oeuvres (1635). Geneva, Switzerland: Slatkine Reprints, 1970.

WORKS ABOUT CHARRON

Adam, Michel. *Études sur Pierre Charron*. Talence, France: Presses univeritaires de Bordeaux, 1991.

Bayle, Pierre. "Charron (Pierre)." In *Dictionnaire historique et critique*. 2 vols. Rotterdam: Reinier Leers, 1697.

Belin, Christian. *L'oeuvre de Pierre Charron, 1541–1603: Littérature et théologie de Montaigne à Port-Royal*. Paris: Slatkine, 1995.

Bremond, Henri. "La folle 'sagesse' de Pierre Charron." *Le correspondant* 252 (1913): 357–364.

Busson, Henri. *La pensée religieuse francise de Charron à Pascal*. Paris: Librairie philosophique J. Vrin, 1933.

Charron, Jean. "Pierre Charron." In *A Critical Bibliography of French Literature. Vol. 3, The Seventeenth Century*, edited by Nathan Edelman, 476–478. Syracuse, NY: Syracuse University Press, 1961. Annotated bibliography of works about Charron.

Charron, Jean. *The "Wisdom" of Pierre Charron: An Original and Orthodox Code of Morality*. Chapel Hill: University of North Carolina Press, 1961.

Dini, Vittorio, and Domenico Taranto, eds. *La Saggezza moderna: Temi e problemi dell'opera di Pierre Charron*. Naples: Edizioni scientifiche italiane, 1987.

Gray, Floyd. "Reflexions on Charron's Debt to Montaigne." *French Review* 35 (1962): 377–382.

Gregory, Tulio. *Genèse de la raison classique de Charron à Descartes*. Paris: Presses Universitaires de France, 2000.

Horowitz, Maryanne Cline. "The Origin of Pierre Charron's Concept of Natural Law in Man." PhD diss., University of Wisconsin, 1970.

Julien-Eymard d'Angers. *Pascal et ses précurseurs: L'apologétique en France de 1580 à 1670*. Paris: Nouvelles Éditions latines, 1954.

Julien-Eymard d'Angers. "Le stoïcisme en France dans la première moitié du XVIIe siècle. Les origines (1575–1616)." *Études franciscaines* 2 (1951): 389–410.

Maia Neto, José R. "Charron's Construction of *Épochè* as Wisdom." Presented at the annual meeting of the Renaissance Society of America, April 1–3, New York: 2004.

Maia Neto, José R. "Charron's *Epoché* and Descartes' *Cogito*: The Sceptical Base of Descartes' Refutation of Scepticism." In *The Return of Scepticism: From Hobbes and Descartes to Bayle*, edited by Gianni Paganini, 81–113. Dordrecht, Netherlands: Kluwer Academic, 2003.

Paganini, Gianni. *Scepsi Moderna: Interpretazioni dello scetticismo da Charron a Hume*. Cosenza, Italy: Busento, 1991.

Popkin, Richard H. "Charron and Descartes: The Fruits of Systematic Doubt." *Journal of Philosophy* 51 (1954): 831–837.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Rice, Eugene F. *The Renaissance Idea of Wisdom*. Cambridge, MA: Harvard University Press, 1958.

Sabrié, Jean B. *De l'humanisme au rationalisme: Pierre Charron (1541–1603), l'homme, l'oeuvre, l'influence*. Geneva, Switzerland: Slatkine Reprints, 1970.

Richard H. Popkin (1967, 2005)

CHARTRES, SCHOOL OF

A cathedral school existed at Chartres as early as the sixth century but did not become famous until the eleventh and twelfth centuries. Under Bishop Fulbert (d. 1028), a pupil of Gerbert of Aurillac, students, among them Berengar of Tours, flocked to Chartres to study the *trivium* and *quadrivium*, medicine and theology. Later, Bishop Ivo brought renown in canon law. The high point was reached in the early twelfth century under Bernard of Chartres and his brother Theodoric (Thierry) and their pupils Gilbert of Poitiers (de la Porrée), William of Conches, and Clarembald of Arras. Also associated with the school in various ways were Bernard of Tours, Adelard of Bath, Alan of Lille, and John of Salisbury. The Chartrains of this period were humanists who loved the literature and philosophy of classical antiquity. The richness of their program of studies is evident in Theodoric's *Heptateuch*, a handbook of the seven liberal arts and a collection of the authors who were read. In the early twelfth century Chartres was the center of Latin Platonism. Plato himself was known only indirectly through a fragment of the *Timaeus* in the translation and commentary of Chalcidius and through Macrobius, Apuleius, Seneca, and Boethius, whose *Opuscula Sacra* and *Consolatio Philosophiae* were much commented on. Devotion to Platonism produced realist interpretations of the problem of universals, speculations about the Ideas, matter and form, cosmological thought, and discussions about the world soul. Aristotle was generally less highly esteemed. The Chartrains knew only his logical writings (the *Organon*), including the *logica nova* (the rediscovered *Prior Analytics, Topics,* and *Sophistic Refutations*), which makes an early appearance in Theodoric's *Heptateuch.* Under the inspiration of Boethius, attempts were

made to reconcile Aristotelianism and Platonism. Theology was presented largely in philosophical clothing. Confident of the harmony of faith and learning, the Chartrains attempted to establish the existence of God by numerical speculations, to synthesize Platonic cosmology and biblical revelation, and to compare the Platonic world soul with the Holy Spirit, as in William of Conches. God was considered to be the form of all being, a view that has been called pantheistic by some historians. Greek and Arabian writings on medicine, astronomy, and mathematics, including works by Hippocrates, Galen, Ptolemy, Euclid, al-Khwarizmi, Johannitius, and others were circulated and read in translation. In the early twelfth century Chartres was without a peer as a school of classical and humane learning and of Platonism, and it was rivaled in philosophy only by Paris. The bloom was fading fast by midcentury, but the influence of the school continued to be marked among the disciples of Gilbert of Poitiers, in thirteenth-century writings on natural philosophy, and still later in the works of Nicholas of Cusa.

See also Aristotle; Bernard of Chartres; Bernard of Tours; Boethius, Anicius Manlius Severinus; Galen; Gerbert of Aurillac; Gilbert of Poitiers; Hippocrates and the Hippocratic Corpus; John of Salisbury; Nicholas of Cusa; Plato; Platonism and the Platonic Tradition; Seneca, Lucius Annaeus; Theodoric of Chartres; William of Conches.

Bibliography

Clerval, A. *Les écoles de Chartres au moyen âge.* Paris: A. Picard, 1895.

Geyer, B. *Die patristische und scholastische Philosophie.* Basel, 1927, Pp. 226–252.

Gregory, T. *Anima Mundi. La filosofia di Guglielmo di Conches.* Florence: G.C. Sansoni, 1955.

Parent, J. *La doctrine de la création dans l'école de Chartres.* Paris: J. Vrin, 1938.

Wulf, M. de. *History of Mediaeval Philosophy.* Translated by E. C. Messenger. London: Nelson, 1952. Vol. I, pp. 173–188. Translation of *Histoire de la philosophie médiévale.* Louvain: Institut Supérieur de Philosophie, 1934.

David Luscombe (1967)

CHATEAUBRIAND, FRANÇOIS RENÉ DE
(1768–1848)

François René de Chateaubriand, the French author, was born at Saint-Malo in Brittany and educated at Dol-de-Bretagne and Rennes in preparation for studying for the priesthood at the Collège de Dinan. Finding that he had no vocation, he followed the tradition of his social class and became an army officer instead. In 1788 he joined the order of the Knights of Malta, went to Paris, and began to associate with men of letters. From then on literature was his chief interest in life, though his literary career was paralleled by a career in diplomacy and politics. In 1803 he was appointed an attaché at the French embassy in Rome, and upon the return of Louis XVIII to power he played a role in politics in the Ministry of the Interior. His main diplomatic post was that of French plenipotentiary at the Congress of Verona, an account of which he published in 1838.

Chateaubriand's political as well as his religious views were in a state of constant flux. As a young man he had been favorable to the revolution, but he was soon disillusioned and in 1792 went into voluntary exile in London. There he published his *Essai historique, politique et moral sur les révolutions,* which he later retracted. This work was clearly influenced by the Philosophes, especially Jean-Jacques Rousseau, and, though far from atheistic, was definitely favorable to deism and opposed to Christianity. As Charles-Augustin Sainte-Beuve showed a half-century later in his *Causeries du Lundi,* the printed version of Chateaubriand's views was much less extreme than what he really thought. Having undergone a personal crisis when he learned of the death of his mother, he returned from exile in 1800 and began the preparation of one of his most famous works, *Le génie du Christianisme.* The aim of the volume was to persuade the public that Christianity had as many themes worthy of artistic expression as paganism. It produced, said Sainte-Beuve, "a whole army of parlor Christians." This was precisely the goal of its author, to make Christianity fashionable.

In September 1816, Chateaubriand published his pamphlet *De la monarchie selon la charte,* which preached political liberalism in a constitutional monarchy. This brought on his temporary political ruin, but he soon recovered and was utilized by the government in various diplomatic posts. Toward the close of his life he developed an intimacy with Mme. Récamier and her circle but withdrew from politics and devoted himself to the preparation of his memoirs, the *Mémoires d'outretombe* (published posthumously in 1849).

Chateaubriand's contributions to French philosophy were indirect. The early *Essai sur les révolutions* made it clear that he considered any type of philosophy to be antireligious and religion to be a substitute for philosophy. In it he attempted to show that no philosophy could

ever hope to reach the truth, for truth was discovered not by reasoning but by some inner light, a kind of feeling (*sentiment*), perhaps what Blaise Pascal called the heart. It was this belief that appeared in such works as *Atala,* where the theme of the Noble Savage is developed. Though Atala is herself a Christian, she is a Christian by sentiment, not by reason, and her form of Christianity was believed by her inventor to be higher and nobler than that deduced by argument.

Similarly, Chateaubriand anticipated William Wordsworth in maintaining even as a young man that in the contemplation of nature, in the sense of the landscape, there is a spontaneous revelation of the truths of morality and religion. The famous passage "Night among the American Savages," which terminates the *Essai* and was reprinted in part in the *Génie du Christianisme,* is not only a description of a moonlight scene near Niagara Falls but also an evocation of the nobility of soul that belongs only to men who have lived in a state of cultural primitivism far from the contamination of society. Like Rousseau, Chateaubriand pitted nature and society against each other, and it is significant that in this passage the Indians are only two women, two small children at the breast, and two warriors. There is no mention of a tribe or village. The sole contact these people have with anything outside themselves is with the "ocean of trees." But it is to be noted that far from reinforcing the sense of individuality, this contact, on the contrary, induces an absence of all distinct thoughts and feelings, a kind of mystical union with that God who is nature itself.

This type of anti-intellectualism reappeared in the *Génie du Christianisme.* Chateaubriand said in the preface to this work that he turned away from eighteenth-century liberalism when he learned of his mother's death. He was in exile in London at the time. "I wept," he wrote, "and I believed." The evidence of tears was proof of the truths of Catholicism, as in the *Essai* the feelings aroused by natural scenery were proofs of the truth of deism. But Catholicism is hardly a religion spontaneously kindled in the hearts of all people. It is a religion initiated and developed in society. Hence, Chateaubriand found himself aligned with the Traditionalists, a group as far from Rousseauistic sentimentalism as can be imagined. For whereas Joseph Marie de Maistre and the Vicomte de Bonald believed reason was the faculty that united human beings, the sentimentalists believed it was what divided them into conflicting sects.

It was perhaps for this reason that Chateaubriand emphasized the gifts Christianity had made to European culture. He wrote at the height of the Neoclassical movement, when the masters were Jacques Delille in poetry, Antonio Canova in sculpture, and Jacques Louis David in painting. They, of course, found their inspiration in classical mythology and history. Chateaubriand tried to prove that there was more to be found in the Catholic tradition. However true this may have been, the point he was making was that to the extent that any set of beliefs increases the amount of beauty and goodness in the world, that set of beliefs is true. There is a concealed pragmatic test here that is of interest historically and would probably not be able to resist criticism. But at a time when men had lived through a period of horror brought on by the suppression of religion, it was understandable that they should attribute the horrors to the philosophy they believed had generated the antireligious practices. To Chateaubriand at this time the one alternative to philosophy was Catholicism, not that natural religion which he had lauded in the *Essai.* And this belief he never abandoned. He was not the type of writer to set down a body of premises from which he would deduce certain inferences. On the contrary, his hatred of philosophy was such that he simply stated his conclusions as his heart dictated; it remained for others to disentangle the form of his argument. He established a cultural atmosphere rather than a set of doctrines, and his works are more properly viewed as long poems of a purely lyrical nature than as doctrinal treatises.

Bibliography

WORKS BY CHATEAUBRIAND

Essai historique, politique et moral sur les révolutions. London, 1797.

Le génie du Christianisme. Paris, 1802. See especially the preface.

Les martyrs. Paris, 1809.

Oeuvres complètes, 20 vols. Paris, 1858–1861. Introduction by C. A. Sainte-Beuve.

WORKS ON CHATEAUBRIAND

Bertrin, Abbé Georges. *La sincérité religieuse de Chateaubriand.* Paris, 1899.

Blum, Christopher O., ed. and tr. *Critics of the Enlightenment: Readings in the French Counter-Revolutionary Tradition.* Wilmington, DE: ISI, 2004.

Chinard, G. *L'exotisme américain dans l'oeuvre de Chateaubriand.* Paris: Hachette, 1918.

Grimsley, Ronald. *Soren Kierkegaard and French Literature: Eight Comparative Studies.* Cardiff: Wales University Press, 1966.

Hilt, Douglas. *Ten against Napoleon.* Chicago: Nelson-Hall, 1975.

Law, Reed G., and Bobbie W. Law. *From Reason to Romanticism.* Denver: Big Mountain Press, 1966.

Sainte-Beuve, C. A. "Chateaubriand, anniversaire du Génie du Christianisme." In *Causeries du Lundi.* Vol. X, pp. 74–90. Paris: Garnier, 1855.

Sainte-Beuve, C. A. *Chateaubriand et son groupe littéraire sous l'Empire.* Paris, 1869.

George Boas (1967)
Bibliography updated by Tamra Frei (2005)

CHATTON, WALTER

(c. 1285–1343)

Walter Chatton was born in the village of Chatton in Northumbria. He entered the Order of Friars Minor at a young age and pursued the normal course of theological studies. His first lectures on the Sentences of Peter Lombard, called *Reportatio*, were held between 1321 and 1323. At the time Chatton, with William of Ockham and Adam Wodeham, was located in one of the Franciscan studia, probably London or Oxford, where Wodeham was the scribe or *reportator* of Chatton's lectures. A second commentary on the Sentences (incomplete), called *Lectura*, dates from 1328 to 1330. Besides these two Sentence commentaries, a single set of *Quodlibetal Questions* (incomplete) survives. Chatton became the fifty-third regent master for the Franciscans at Oxford in 1330. He went to Avignon in 1333 and was appointed by Popes Benedict XII (d. 1342) and Clement VI (c. 1291–1352) as one of the examiners of the writings of Thomas Waleys (d. 1349) and Durandus of Saint-Pourcain. He was appointed as bishop designate of the diocese of St. Asaph in Wales but died before the see had become vacant.

In virtually every distinction, question, and article of his lectures, Chatton attacks the views of Ockham, who in turn was appraised of these criticisms by Ockham's most noteworthy disciple, Wodeham. Chatton's other favored opponent was Peter Aureol, who had frequently criticized Chatton's favorite philosopher-theologian, John Duns Scotus. It is practically impossible to follow Chatton's train of thought without knowledge of the views of Ockham and Aureol.

One of Chatton's frequently invoked hermeneutical principles was designated as "my proposition" and can be called the antirazor as the foil of Ockham's principle of parsimony. If a situation cannot be adequately described by two propositions, then a third must be invoked, and if this is not adequate a fourth is required and so on.

In the domain of natural philosophy, Chatton was an indivisibilist, who viewed the continua, both permanent and successive, quantitative and temporal, as composed of indivisibles or instants. The argument being that whatever God by his absolute power can do successively, he could do instantaneously, and thus there would be, according to the divisibilists' view, an infinite multitude capable of accretion ad infinitum. Chatton is conscious that he is in the minority and is counter to the views of Aristotle and most philosopher-theologians.

Concerning the ten Categories of Aristotle, Ockham held that only substance and quality enjoyed extramental existence. In contrast, Chatton claimed that all the categories in one way or another were distinct realities and he took every opportunity to attack Ockham's claim that quantity was simply extended substance and not extramentally real.

According to Ockham relations as such are not some *tertia quid*. A white thing A and a white thing B both regarding their fundament whiteness and their distinct termini as things enjoy extramental reality, but this does not mean that their relation of similarity requires extramental status. Naturally, Chatton posits *res respectivae* and counters Ockham's views whenever possible.

Initially, Ockham held that concepts were nothing more than *esse obiectiva* (their being known) and not accidents or qualities of the mind. Because of Chatton's critique, Ockham modified his view and admitted that concepts were qualities of the mind. However, this did not mean that universals qua universals were things outside the mind, such that Ockham is best qualified as a conceptualist (nominalist in the medieval sense), where as Chatton and Scotus are best classified as moderate realists.

Chatton's other principal adversary was Aureol. The latter had criticized Scotus's opinion that a univocal concept of being was absolutely essential in any attempt to prove the existence of God. Aureol noted that the modes "finite" and "infinite" did not come under the purview of "being" as univocal. Chatton admits the objection while claiming that there is a concept of being that includes all its modes, including the ultimate individual difference or individual property (the word *haeceitas* occurs rarely and perhaps only once in Scotus's writings) and is a purely logical concept and not a metaphysical one.

Scotus's view of the principle of individuation as not being a double-negation (Henry of Ghent), a determinate quantity (St. Thomas Aquinas), or a collection of accidents (Porphyry and Anicius Manlius Severinus Boethius), but something positive that Scotus called the ultimate or individual difference or property, came under considerable attack from his successors. Ockham would

claim that no such principle was required because God created individuals and not species, genera, or universals. Chatton, however, strove to defend Scotus's view even while cognizant of its difficulties.

Just as Chatton regularly attacked Ockham, so Wodeham frequently criticized Chatton's views, particularly if Chatton was seen as misinterpreting or misunderstanding Ockham's positions.

In the realm of theology Chatton may be read as favoring positive theology, namely, as concerned with what the scriptures and the church fathers had to say. He is less concerned about what God might do or what he might have done by his absolute power (hypothetical theology).

Chatton is thus one of the earliest Scotises and his views attest to the intellectual ferment of his age. He is an interpreter of Scotus and offers alternative approaches in philosophical-theological discourse to his fellow Franciscans, Aureol, Ockham, and Wodeham.

See also Duns Scotus, John; Peter Aureol; William of Ockham; Wodeham, Adam.

Bibliography

WORKS BY CHATTON

Quaestio utrum quantum et continuum componantur ex indivisibilibus sicut ex partibus integralibus, edited by J. E. Murdoch and E. A. Synan. *Franciscan Studies* 26 (1966): 212–288.

Reportatio et Lectura super Sententias: Collatio ad Librum Primum et Prologus, edited by Joseph C. Wey. Toronto, Canada: Pontifical Institute of Mediaeval Studies, 1989, viii + 430 pp.

Reportatio super Sententias. 3 vols., edited by Joseph C. Wey and Girard J. Etzkorn. Toronto, Canada: Pontifical Institute of Mediaeval Studies, 2002–2004.

Reportatio super Sententias. Liber II, dist. 1–20, edited by J. C. Wey and G. J. Etzkorn. Toronto: Pontifical Institute of Mediaeval Studies, 2004, xii + 370 pp.

Reportatio super Sententias. Liber III, dist. 1–33; Liber IV, qq. 1–11, edited by J. C. Wey and G. J. Etzkorn. Toronto: Pontifical Institute of Mediaeval Studies 2005, xiii + 399 pp.

WORKS ABOUT CHATTON

Baudry, L. "Gauthier de Chatton et son Commentaire des Sentences." *Archives d'histoire doctrinale et littéraire du moyen âge* 14 (1943–1945): 337–369.

Brown, Stephen F. "Medieval Supposition Theory in Its Theological Context." *Medieval Philosophy and Theology* 3 (1993): 121–157.

Brown, Stephen F. "Walter Chatton's Lectura and William of Ockham's *Quaestiones in libros Physicorum Aristotelis*." In *Essays Honoring Allan B. Wolter*, edited by William A. Frank and Girard J. Etzkorn, 81–93. St. Bonaventure, NY: Franciscan Institute, 1985.

Fitzpatrick, Noel A. "Walter Chatton on the Univocity of Being: A Reaction to Peter Aureoli and William of Ockham." *Franciscan Studies* 31 (1971): 88–171.

Gál, G. "Gualteri de Chatton et Guillemi de Ockham controversia de natura conceptus universalis." *Franciscan Studies* 27 (1967): 199–212.

Maurer, A. "Ockham's Razor and Chatton's Anti-razor." *Medieval Studies* 46 (1984): 463–475.

Tachau, Katherine H. "The Early Reaction to Aureol and Ockham: The Views of Walter Chatton." In *Vision and Certitude in the Age of Ockham: Optics, Epistemology, and the Foundations of Semantics, 1250–1345*, 180–208. Leiden, Netherlands: E. J. Brill, 1998.

Girard J. Etzkorn (2005)

CHEMISTRY, PHILOSOPHY OF

Ideas about the diversity of matter in terms of elements and compound substances and their transformations have been pivotal to any scientific or prescientific approach to nature. From ancient natural philosophy and alchemy to modern nineteenth-century chemistry, these ideas were made the basis of philosophical systems and became the target of critical reflection. After a temporary interruption when modern philosophy of science focused on mathematical physics, philosophy of chemistry emerged anew in the 1980s and has become a flourishing field in which philosophers, chemists, and historians of chemistry are engaged. While many of the old philosophical issues have been rediscovered and discussed, new issues have also appeared as a result of shifts of general philosophical focus, alliances with historians and sociologists of science, the development of chemistry, and changes in its role in society.

ONTOLOGICAL ISSUES

The objects of chemistry are subject to many ontological debates beyond simple issues of definition, and these debates also have an impact on epistemological and methodological issues. Following the example of microphysics, many philosophers and chemists take atoms and molecules as the basic objects of chemistry. Yet despite the numerous techniques available to visualize molecules, the notion of a molecule is a theoretical concept with many model assumptions that do not apply to nonmolecular substances, such as water, metals, and salts. It is not so much the lack of optional microstructural descriptions for these substances, but the variety of models, which are continuously refined and adapted to certain contexts and

problems, that makes such models a weak basis for defining the basic objects of chemistry. Another option is to take material substances, either elementary or compound, as the basic objects. Yet, far from being phenomenologically given entities, pure substances are the final results of infinite purification operations; that is, they are ideal laboratory artifacts. This fact has in turn inspired operational definitions. Whether one takes microstructures or pure substances as basic is not an arbitrary decision, but rather has direct impact on chemical classification and all derived concepts, because there is no simple one-to-one relationship between the two kinds of entities. There are microstructures without corresponding pure substances, and there are substances with many different microstructures.

A second but related ontological issue is about natural kinds in chemistry. Microstructuralists, following Hilary Putnam, have claimed that water is a natural kind because it is determined by a microstructural essence. This claim faces the problems mentioned above. Yet the substance-based approach to natural kinds is confronted not only with a potentially infinite number of possibly essential properties (see below) but also with the artificiality of pure substances. Even if pure substances were stable kinds independent of our conceptualization, they are not independent of laboratory purification. Nonetheless, the experimental reproducibility of sufficiently pure substances provides, within limits, a successful operation to ensure relatively stable kinds.

A third ontological issue is about whether substances (or microstructures) or transformations are the basic objects of chemistry. This issue refers to the general debate between substance and process philosophy. If not closed in bottles, substances continuously undergo chemical reactions and are only intermediate states in an ongoing process. Quantum chemistry describes even these states as processes. Furthermore, traditional chemical characterization of substances goes by chemical properties, that is, by all the dispositions of substances to transform into other substances under certain conditions, including the presence of still other substances as reactants. Substance philosophers define a chemical reaction as the change of certain substances, whereas process philosophers define a substance by its characteristic chemical reactions. A third option, proposed by Joachim Schummer, combines substances and processes in a network of dynamic relations, as the proper object of chemical research. On this view, substances and reactivities mutually define each other. Answering the ontological question has direct consequences on whether chemists can best organize their knowledge in the form of substance databases, reaction databases, or combined substance-reaction databases.

Although all substances and transformations are usually considered objects of chemistry, the metaphysical distinction between natural and synthetic pervades both commonsense and chemical reasoning. Yet the notion of natural substances—substances that can be isolated from natural resources by purification—is questionable. Not only is purification a technical operation; also, most elements would have to count as synthetic when natural resources are lacking. On the other side of the ledger, all substances that can be isolated from natural resources can also be synthesized in the early twenty-first century, which undermines the distinction. Furthermore, we have little evidence to claim that a synthetic substance will never be isolable from natural resources in the entire universe.

EPISTEMOLOGICAL AND METHODOLOGICAL ISSUES

A central epistemological issue is whether chemical knowledge can be complete or not. Microstructural essentialists claim that a perfect microstructural description of any substance yields complete chemical knowledge. However, chemical properties are not manifest properties but dispositional relations (that is, relations of the form "*A* under certain conditions is disposed to react with *B* to form *C* and *D*"). This means that the structure of experimental chemical knowledge is relational, dispositional, and open-ended. Because new properties are defined by new conditions and new potential reactants (currently produced at 15.5 million new chemicals per year), experimental chemical knowledge can increase indefinitely without reaching a state of being complete. It is an open question to what extent theoretical approaches can compensate for the incompleteness on the experimental level.

Chemistry differs from other sciences in that its theoretical concepts need to serve different methodological goals. Besides the traditional goals of accurately describing, explaining, and predicting phenomena, theoretical concepts in chemistry also fulfill purposes of classification and synthesis. By 2004 the chemical classification system had distinguished 78.3 million different substances and ordered them by classes and subclasses. And beyond mere prediction of phenomena, theoretical concepts provide experimental guidelines for producing millions of new substances and reactions per year. For all three methodological goals, the main theoretical

approach has been chemical-structure theory, which emerged in mid-nineteenth century and has been influenced and diversified by many different developments since, including quantum chemistry and spectroscopic instrumentation. Apart from this theory, a multitude of other theoretical concepts and models have been developed for particular substance classes and phenomena and for various purposes.

The main methodological issue in current philosophy of chemistry is to bring order to this complex picture without imposing upon chemistry methodologies tailored to other disciplines. Several case studies have shown that received approaches, for instance, Karl Popper's view that science makes progress by falsifying theories, are rather useless in chemistry. There is some agreement that chemists favor methodological pluralism and pragmatic application of models, rather than methodological universalism and the ideal of a single axiomatic theory. A study on scientific realism has suggested that entity realism, rather than theory realism, is a more appropriate methodological ideal in chemistry. The received methodological focus on methods of justification has been widened to include methods for research, that is, for developing new knowledge. Many detailed studies on the different kinds and uses of models in chemistry, from theoretical chemistry to chemical engineering, have been undertaken. Besides the impact of quantum mechanics (see the next section), the impact of spectroscopic instrumentation on theoretical concepts since the mid-twentieth century has received considerable attention, in fact, so much attention that interest in the "instrumental revolution" has replaced the older focus on the eighteenth-century "chemical revolution" by Antoine-Laurent Lavoisier and others. The methodological integration of both chemical analysis and synthesis, which form the major experimental activity of chemists, has overcome received distinctions between science and technology. Studies on the formal sign-language system of chemistry, consisting of structural formulas and reaction mechanisms, have illuminated its multipurpose theoretical capacity, but further studies are required to understand changes stemming from various theoretical and experimental developments.

REDUCIBILITY TO PHYSICS

Whether chemistry is reducible to physics is a question that could come up only in the mid-nineteenth century, when modern physics emerged as its own discipline, because the former meaning of "physics" (natural science or natural philosophy) included chemistry as a branch. Before then, mechanical (physical) approaches were among several competing approaches within theoretical chemistry, though not very successful. The question became meaningful only with the development of quantum mechanics and its application to chemistry since the late 1920s. Following a speech by Paul Dirac in 1929, many quantum physicists and philosophers of physics have taken for granted that the whole of chemistry would be reducible to quantum mechanics, and so would be part of physics.

Wary of making such bold claims, philosophers have carefully distinguished between different meanings of "reduction." An ontological reduction claims that the supposed objects of chemistry are actually nothing other than the objects of quantum mechanics and that quantum-mechanical laws govern their relations. In its strong, eliminative version, an ontological reduction states that there are no chemical objects proper. Antireductionists argue that theoretical entities are determined by their corresponding theories, and that theoretical entities of different theories cannot be identified. For instance, from the different meanings of the term "electron" in quantum electrodynamics and in chemical-reaction mechanisms, they conclude that the term "electron" has different references, which rules out an ontological reduction. An epistemological or theoretical reduction claims that all theories, laws, and fundamental concepts of chemistry can be derived from first-principle quantum mechanics as a more basic and more comprehensive theory. This claim has prompted many detailed studies (see below). Methodological reductionism, while acknowledging the current failure of epistemological reduction, recommends applying quantum-mechanical methods to all chemical problems, because that would be the most successful approach in the long run (approximate reductionism). But the mere promise of future success is not convincing unless accompanied by a comparative assessment of different methods. By modifying the popular notion that the whole is nothing but the sum of its parts, philosophers have developed two further versions of reductionism. Emergentism acknowledges that new properties of wholes (say of water) emerge when the parts (say oxygen and hydrogen) are combined, but it does not deny that the properties of the whole can be explained or derived from the relations between the parts (epistemological reductionism). Supervenience, in a simple version, means that, although epistemological reductionism might be wrong, the properties of a whole asymmetrically depend on the properties of the parts, so that every change of the properties of the whole is based on changes of the properties of the parts or the relations between the parts, but not the other way round.

When these terms are applied to the reduction of chemistry to quantum mechanics, that is, to chemical entities as wholes and quantum-mechanical entities as parts, emergentism and supervenience presuppose elements of epistemological or ontological reductionism. Thus, criticism of these positions applies accordingly. For instance, if one denies that chemical electrons are the same as quantum-electrodynamic electrons or, more generally, that quantum-mechanical entities are proper parts of chemical wholes, one ends up rejecting supervenience altogether.

Recent criticism has focused on epistemological reductionism by pointing out the technical limits of quantum mechanics with regard to particular chemical concepts, laws, and problems. Two quantum chemists, Guy Woolley and Hans Primas, have shown that the concept of molecular structure, which is central to most chemical theories, cannot be derived from first-principle quantum mechanics, because molecular structures cannot be represented by quantum-mechanical observables. Eric Scerri has argued that current quantum-mechanical approaches cannot calculate the exact electronic configuration of atoms, which was formerly considereda successful reduction of the chemical law that underlies the periodic system of elements. Jaap van Brakel has pointed out that successful applications of quantum mechanics to chemical problems frequently include model assumptions and concepts taken from chemistry. Joachim Schummer has argued that quantum-mechanical approaches are nearly absent and useless in areas that chemists are mainly concerned with: chemical reactions, synthesis, and classification.

Criticism of the reduction of chemistry to quantum mechanics, as the lowest level in the standard hierarchy of reductions, also challenges microreductionism as a general position and thus contributes to general philosophy. In the most detailed philosophical study on various forms of reductionism, Jaap van Brakel has used the case of chemistry to argue for a kind of pragmatism in which the "manifest image" of common sense and the empirical sciences is epistemologically primary over the "scientific image" of microphysics. Nikos Psarros presupposes a rejection of reductionism in his extensive project of seeking the cultural foundation of chemical concepts, laws, and theories in prescientific cultural practices, norms, and values. For many others, including Joachim Schummer, rejecting reductionism supports a pragmatist and pluralist position that clearly distinguishes between fields of research where quantum-mechanical approaches are strong and even indispensable and those where they are poor or even useless compared to other approaches. Once reductionism has lost its function of securing the unity of the sciences, new relationships between chemistry and other disciplines could become subject to philosophical and historical investigations, including studies of such multidisciplinary fields as atmospheric science, biomedical science, materials science, and nanotechnology.

FURTHER TOPICS

Current philosophy of chemistry reaches far beyond ontological, epistemological, and methodological issues. On the one hand, there are strong trends in historical research. Pertinent classical works on chemistry by such philosophers as Aristotle, Immanuel Kant, Georg Wilhelm Friedrich Hegel, Pierre Duhem, Ernst Cassirer, and Gaston Bachelard have been rediscovered, and these have allowed reinterpretations of the history of philosophy of science. Philosophical works by chemists of the past, such as Benjamin C. Brodie, Wilhelm Ostwald, Frantisek Wald, Edward F. Caldin, Fritz Paneth, and Michael Polanyi, have also been rediscovered. Historians and philosophers of chemistry have explored the development of many fundamental concepts in chemistry, such as chemical substance, element, atom, the periodic system of elements, molecular structure, chemical bond, chemical reaction, affinity, and aromaticity. In addition, important historical developments in chemistry have been philosophically scrutinized, such as the transitions from alchemy to modern chemistry and from phlogistic to antiphlogistic chemistry; the emergence of physical chemistry, quantum chemistry, and biochemistry; and the development of molecular-model building and instrumentalization.

On the other hand, philosophers of chemistry have also applied theoretical insights to practical problems, discovered a wider spectrum of philosophical perspectives on chemistry, and engaged in contemporary issues. Epistemological and ontological studies have found useful applications in chemistry education and information management. Beyond the traditional scope of philosophy of science, perspectives on chemistry from philosophy of technology, language, culture, and literature, and from metaphysics, aesthetics, ethics, sociology, and public understanding of science have all been exploited. For instance, studies on the role of visualization and aesthetics in chemical research have been undertaken to understand the heuristics and dynamics of research in a broader cultural context beyond traditional epistemic and technological goals. Philosophers and historians have investigated the historical roots and the cultural value conflicts underlying the widespread chemophobic attitude of society and the peculiar opposition of natural

versus chemical. In addition to taking up general professional ethics, philosophers have challenged the legitimacy of chemical-weapon research, questioned the alleged moral neutrality of synthesizing new substances for scientific purposes, discussed the scope of moral responsibility of chemists for their synthetic products, and developed moral frameworks for assessing chemical-research practice. Finally, with the rise of nanotechnology, in which chemistry is particularly involved, philosophers of chemistry have taken a leading role in discussing the societal and ethical implications of this nanotechnology of the ultra-small.

Bibliography

Baird, Davis. *Thing Knowledge: A Philosophy of Scientific Instruments*. Berkeley, CA: University of California Press, 2003.

Bensaude-Vincent, Bernadette. *Eloge du mixte: Matériaux nouveaux et philosophie ancienne*. Paris: Hachette, 1998.

Bhushan, Nalini, and Stuart Rosenfeld, eds. *Of Minds and Molecules: New Philosophical Perspectives on Chemistry*. New York: Oxford University Press, 2000.

Burbidge, John W. *Real Process: How Logic and Chemistry Combine in Hegel's Philosophy of Nature*. Toronto: University of Toronto Press, 1996.

Earley, Joseph E., ed. *Chemical Explanation: Characteristics, Development, Autonomy*. New York: New York Academy of Science, 2003.

Foundations of Chemistry. Published since 1999. Includes special issues on "The Periodic System" (2001) and "Explanation and the Chemical Sciences" (2004).

Hoffmann, Roald. *The Same and Not the Same*. New York: Columbia University Press, 1995.

Hyle: International Journal for Philosophy of Chemistry. Published since 1995. Includes special issues on "Models in Chemistry" (1999–2000), "Ethics of Chemistry" (2001–2002), "Aesthetics and Visualization in Chemistry" (2003), and "Nanotech Challenges" (2004). Available from http://www.hyle.org/.

Janich, Peter, and Psarros, Nikos, eds. *The Autonomy of Chemistry*. Würzburg, Germany: Königshausen und Neumann, 1998.

Janich, Peter, and Nikos Psarros, eds. *Die Sprache der Chemie*. Würzburg, Germany: Königshausen und Neumann, 1996.

Klein, Ursula, ed. *Tools and Modes of Representation in the Laboratory Sciences*. Dordrecht, Netherlands: Kluwer, 2001.

Laszlo, Pierre. *La parole des choses, ou Le langage de la chimie*. Paris: Hermann, 1993.

Morris, Peter J. T., ed. *From Classical to Modern Chemistry: The Instrumental Revolution*. Cambridge, U.K.: Royal Society of Chemistry, 2002.

Primas, Hans. *Chemistry, Quantum Mechanics, and Reductionism: Perspectives in Theoretical Chemistry*. Berlin: Springer, 1981.

Psarros, Nikos. *Die Chemie und ihre Methoden: Ein philosophische Betrachtung*. Weinheim, Germany: Wiley-VCH, 1999.

Psarros, Nikos, and Kostas Gavroglu, eds. *Ars mutandi: Issues in Philosophy and History of Chemistry*. Leipzig, Germany: Leipziger Universitätsverlag, 1999.

Psarros, Nikos, Klaus Ruthenberg, and Joachim Schummer, eds. *Philosophie der Chemie*. Würzburg, Germany: Königshausen und Neumann, 1996.

Schummer, Joachim. *Realismus und Chemie: Philosophische Untersuchungen der Wissenschaft von den Stoffen*. Würzburg, Germany: Königshausen und Neumann, 1996.

Sobczynska, Danuta, Pawel Zeidler, and Ewa Zielonacka-Lis, eds. *Chemistry in the Philosophical Melting Pot*. Frankfurt, Germany: Peter Lang, 2004.

Van Brakel, Jaap. *Philosophy of Chemistry: Between the Manifest and the Scientific Image*. Louvain, Belgium: Leuven University Press, 2000.

Joachim Schummer (2005)

CHENG HAO

(1032–1085)

Cheng Hao, also called Cheng Mingdao, was cofounder, with his brother Cheng Yi, of the neo-Confucian school of Nature and Principle (*li*). He held some minor official posts but devoted most of his life to teaching.

By making principle the foundation of his philosophy and identifying it with the nature of man and things, Cheng Hao and his brother set the pattern for the neo-Confucian philosophical movement known since the eleventh century as the school of Nature and Principle. To Cheng Hao principle was the principle of nature (*tian li*), a concept that he evolved himself; it was the natural law. It had all the characteristics of principle as conceived by Cheng Yi, but as the principle of nature it was self-existent and unalterable. Whereas Cheng Yi stressed the doctrine that principle is one but its manifestations are many, Cheng Hao emphasized more strongly the principle of production and reproduction as the chief characteristic of nature. To him the spirit of life was in all things. This creative quality was *ren,* the highest good. In man, *ren* becomes humanity, or love, which makes him the moral being he is. It enables him to embrace all things and heaven and earth as one body.

Whatever is produced in man, that is, whatever is inborn in him, is his nature. In its original, tranquil state, human nature is neither good nor evil. The distinction arises when human nature is aroused and manifested in feelings and actions and when these feelings and actions abide by or deviate from the mean. The chief task of moral and spiritual cultivation is to calm one's nature through absolute impartiality and the identification of

internal and external life. To achieve this end Cheng Hao advocated sincerity and seriousness.

There can be no denying that Cheng Hao was the more idealistic and his brother the more rationalistic. Cheng Hao more or less concentrated on self-cultivation, whereas his brother advocated both seriousness and learning. Under the influence of Buddhism Cheng Hao also advocated quietism. The two brothers had vastly different temperaments and therefore showed divergent tendencies, but it is not true, as some scholars claim, that one was monistic and the other dualistic.

See also Buddhism; Cheng Yi; Chinese Philosophy: Overview; Chinese Philosophy: Buddhism.

Bibliography

The works of Cheng Hao and his brother Cheng Yi are in *Er Cheng quan shu* ("The Complete Works of the Two Ch'engs").

Chan, Wing-tsit. *A Source Book in Chinese Philosophy.* Princeton, NJ: Princeton University Press, 1963.

Fung Yu-lan. *A History of Chinese Philosophy.* Translated by Derk Bodde, Vol. II. Princeton, NJ: Princeton University Press, 1953.

Graham, Angus G. *Two Chinese Philosophers, Ch'eng Ming-tao and Ch'eng Yi-ch'uan.* London: Lund Humphries, 1958; LaSalle, IL: Open Court Press, 1992.

Graham, Angus C. "What Was New in the Ch'eng-Chu Theory of Human Nature?" In *Studies in Chinese Philosophy and Philosophical Literature.* Albany: SUNY Press, 1990.

Feng Youlan. "Cheng Hao and Cheng Yi." *Chinese Studies in Philosophy* 13 (Win–Spr 1981–82): 127–182.

Wing-tsit Chan (1967)

Bibliography updated by Huichieh Loy (2005)

CHENG I

See *Cheng Yi*

CH'ENG MING-TAO

See *Cheng Hao*

CHENG YI

(1033–1107)

Cheng Yi, or Cheng Yi-chuan, was the most outstanding Chinese teacher of his time, a lecturer to the emperor on Confucian classics, and cofounder, with his brother Cheng Hao, of the neo-Confucian school of principle (*li*) that dominated Chinese thought for many centuries.

The central concept of the school is principle. The concept, negligible in ancient Confucianism, had been developed by the neo-Daoists and Buddhists, but the Cheng brothers were the first to build their philosophy primarily on it. To them, principle is self-evident and self-sufficient, extending everywhere and governing all things. It is laid before our very eyes. It cannot be augmented or diminished. It is many, but it is essentially one, for "definite principles" are but principle. "Principle is one but its manifestations are many." It is universal truth, universal order, universal law. Most important of all, it is the universal principle of creation. It is dynamic and vital. Man and all things form one body because all of them share this principle. It is identical with the mind and with the nature of man and things. Since principle is principle of creation and since life-giving is good, principle is the source of goodness. To be good is to obey principle. Thus, principle is both natural and moral and both general and specific. It has meaning as an abstract reality, but more so as the moral law of man.

The relation between principle and material force, which actualizes things, is not a dualistic one. Although Cheng Yi said that "material force exists after physical form and is therefore with it whereas the Way [principle] exists before form and is therefore without it," he also said that "what makes yin and yang [material force] is the Way." Material force is the physical aspect of principle. In the process of creation each operation is new, for material force is perpetually generated by Origination. (Origination is comparable to creation, except that it is natural and self-caused and is not an act of any being.)

To understand principle one can study one thing intensively or many things extensively. One can also read books, study history, or handle human affairs, for all things and affairs, including blades of grass, possess principle. This intellectual approach makes Cheng's system strongly rationalistic. The approach, however, is balanced by the moral, for whereas "the pursuit of learning depends on the extension of knowledge," "self-cultivation requires seriousness." This dual emphasis reminds one of the Buddhist twofold formula of meditation (dhyana) and wisdom (prajna).

See also Buddhism; Cheng Hao; Chinese Philosophy: Overview; Chinese Philosophy: Buddhism.

Bibliography

The works of Cheng Hao and his brother Cheng Yi are in *Er Cheng quan shu* ("The Complete Works of the Two Ch'engs").

Chan, Wing-tsit. *A Source Book in Chinese Philosophy.* Princeton, NJ: Princeton University Press, 1963.

Fung Yu-lan. *A History of Chinese Philosophy.* Translated by Derk Bodde, Vol. II. Princeton, NJ: Princeton University Press, 1953.

Graham, Angus G. *Two Chinese Philosophers, Ch'eng Ming-tao and Ch'eng Yi-ch'uan.* London: Lund Humphries, 1958; LaSalle, IL: Open Court Press, 1992.

Graham, Angus C. "What Was New in the Ch'eng-Chu Theory of Human Nature?" In *Studies in Chinese Philosophy and Philosophical Literature.* Albany: SUNY Press, 1990.

Feng Youlan. "Cheng Hao and Cheng Yi." *Chinese Studies in Philosophy* 13 (Win–Spr 1981–82): 127–182.

Wing-tsit Chan (1967)
Bibliography updated by Huichieh Loy (2005)

CH'ENG YI-CH-UAN

See *Cheng Yi*

CHERNYSHEVSKI, NIKOLAI GAVRILOVICH

See *Chernyshevskii, Nikolai Gavrilovich*

CHERNYSHEVSKII, NIKOLAI GAVRILOVICH
(1828–1889)

Nikolai Gavrilovich Chernyshevskii, the Russian literary and social critic, was the guiding spirit of Russian nihilism and a major representative of positivistic materialism in nineteenth-century Russian philosophy.

Chernyshevskii was born in Saratov, Russia. The son of an Orthodox priest, he attended a theological seminary before entering the University of St. Petersburg in 1846. After his graduation in 1850, he taught secondary school in Saratov until 1853, when he returned to St. Petersburg, secured a master's degree in Russian literature, and began writing for leading reviews. He soon became a principal editor of *Sovremennik* (The contemporary), and by the early 1860s was the foremost spokesman of radical socialist thought in Russia. Arrested in 1862, he was banished to Siberia in 1864 and passed the remaining twenty-five years of his life in forced exile. He was permitted to return to Saratov, in failing health, a few months before his death.

In his student days Chernyshevskii was attracted to the writings of the French socialists and of G. W. F. Hegel and the left-wing Hegelians. In 1849 he read Ludwig Feuerbach's *Essence of Christianity* and by 1850 had formed an allegiance to Feuerbach that was decisive in his philosophical development. He was also influenced by the English utilitarians, notably John Stuart Mill, whose *Principles of Political Economy* he translated into Russian in 1860.

Chernyshevskii's master's dissertation and first philosophical work, *Esteticheskie otnosheniia iskusstva k deistvitel'nosti* (The aesthetic relation of art to reality; St. Petersburg, 1855), is a critique of Hegelian aesthetics "deduced" (as Chernyshevskii later expressed it) from Feuerbach's naturalistic principles. Chernyshevskii argued that art is an aesthetically inferior substitute for concrete reality. The essential purpose of art is to reproduce the phenomena of real life that are of interest to man, compensating for his lack of opportunity to experience the reality itself. The derivative purposes of art, which give it a moral dimension, are to explain this reality for the benefit of man and to pass judgment upon it. Chernyshevskii developed his aesthetic views further, emphasizing the social context of art, in his *Ohcerki gogolevskogo perioda russkoi literatury* (St. Petersburg, 1855–1856; translated as *Essays on the Gogol Period of Russian Literature*).

In his chief philosophical work, a long essay titled *Antropologicheskii printsip v filosofii* (The anthropological principle in philosophy; 1860), Chernyshevskii exhibited his acceptance of Feuerbach's anthropologism and adopted the materialistic position he retained throughout his life. By "the anthropological principle" Chernyshevskii meant the conception of man as a unitary organism whose nature is not bifurcated into "spiritual" and "material" elements. He argued that philosophical questions can be resolved only from this point of view and by the methods of the natural sciences. Indeed, in all their essentials such questions had already been resolved by the sciences, according to Chernyshevskii: Man is a complex chemical compound whose behavior is strictly subject to the law of causality, who in every action seeks his own pleasure, and whose character is determined by the features of the environment within which he is obliged to act.

On this basis Chernyshevskii advocated "rational egoism"—an ethical theory of enlightened egoistic utilitarianism—and maintained that radical reconstruction

of the social environment is needed to create happy and productive individuals. He portrayed these "new people" and the socialist order of the future in a novel, *Chto delat'?* (*What Is to Be Done?*, St. Petersburg, 1863), which was the principal literary tract of Russian nihilism and was for decades enormously influential in the radical movement. In his socioeconomic thought in general Chernyshevskii emphasized the peasant commune and the artel and is considered an important forerunner of Russian Populism.

Chernyshevskii was a severe critic of neo-Kantian phenomenalism. In a number of letters and in the essay *Kharakter Chelovecheskovo Znaniya* (The character of human knowledge; Moscow, 1885), written in exile, he espoused epistemological realism and condemned the skepticism and "illusionism" (as he called it) of such scientists as Rudolf Virchow and Emil Heinrich Du Bois-Reymond.

Bibliography

WORKS BY CHERNYSHEVSKII

Polnoye Sobraniye Sochineni (Complete works). 16 vols. Moscow, 1939–1953.

Sobranie sochinenii (Works). 5 vols. Moscow: Pravda, 1974.

Selected Philosophical Essays. Westport, CT: Hyperion, 1981.

What Is to Be Done?. Translated by M. R. Katz. Ithaca, NY: Cornell University Press, 1989.

WORKS ON CHERNYSHEVSKI

Paperno, I. *Chernyshevsky and the Age of Realism: A Study in the Semiotics of Behavior*. Stanford, CA: Stanford University Press, 1988.

Randall, F. B. *N. G. Chernyshevskii*. New York: Twayne Publishers, 1967.

Venturi, Franco, *Il populismo russo*. 2 vols. Turin, 1952. Translated by Francis Haskell as *Roots of Revolution*. New York: Knopf, 1960.

Woehrlin, W. F. *Chernyshevskii: The Man and the Journalist*. Cambridge, MA: Harvard University Press, 1971.

Zenkovskii, V. V. *A History of Russian Philosophy*. Translated by G. Kline. New York and London: Routledge, 2003.

James P. Scanlan (1967)

Bibliography updated by Vladimir Marchenkov (2005)

CHICHERIN, BORIS NIKOLAEVICH
(1828–1904)

Boris Nikolaevich Chicherin, a Russian philosopher, was educated at Moscow University, where he studied under both K. D. Kavelin and T. N. Granovskii. Until 1868 he was a professor at Moscow University; he also served briefly as tutor to the royal family and as mayor of Moscow (1881–1883). He was cautiously liberal in politics and, after an early period of agnosticism, devoutly Russian Orthodox in religion.

Chicherin wrote substantial critical studies of Vladimir Solov'ëv (1880) and Auguste Comte (1892), as well as several works on philosophy of law and on the state. His ethical individualism, like that of N. I. Kareev, was close to Immanuel Kant's, but, unlike Kareev, Chicherin was an orthodox Hegelian in logic, ontology, and philosophy of history. This eclecticism generated an unresolved tension in his thought. On the one hand Chicherin asserted that great men are merely "organs and instruments of a universal spirit" and that, under certain conditions, a nationality (*narodnost'*) "may become an individual person." On the other hand he insisted that man as a rational creature and "bearer of the Absolute" is an end in himself and must not be "treated as a mere instrument."

Chicherin asserted, with N. K. Mikhailovskii, that "not society, but individuals, think, feel, and desire"; he opposed the "monstrous notion" that society is a higher organism, an all-devouring Moloch, whose function is "to make mankind happy by putting it in chains." Chicherin was alert to encroachments by the social and political spheres on the private and personal realm; he saw the individual—the "foundation-stone of the entire social edifice"—as a single spiritual substance, possessed of reason and free will, and hence of a moral worth and dignity that demand respect.

Chicherin saw the dialectical movement of both thought and being as a passage from initial unity to final multiplicity, through the two intermediary stages of relation and combination. Thus, more explicitly than G. W. F. Hegel, he converted the dialectical triad into a tetrad.

See also Agnosticism; Comte, Auguste; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Kareev, Nikolai Ivanovich; Kavelin, Konstantin Dmitrievich; Philosophy of Law, History of; Mikhailovskii, Nikolai Konstantinovich; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

Two of Chicherin's works, *Polozhitel' naia filosofiia i edinstvo nauki* (Positive philosophy and the unity of science; Moscow: n.p., 1892) and *Osnovaniia logiki i metafiziki* (Foundations of logic and metaphysics; Moscow: n.p., 1894), have been translated as *Philosophische Forschungen* (Heidelberg, 1899). Chicherin's *Filosofiya prava* (Philosophy of law) was published in Moscow in 1900.

For discussion of Chicherin, see V. V. Zenkovsky, *Istoriia russkoi filosofii,* 2 vols. (Paris, 1948 and 1950), translated by G. L. Kline as *A History of Russian Philosophy,* 2 vols. (New York: Columbia University Press, 1953), pp. 606–620.

George L. Kline (1967)

CHICHERIN, BORIS NIKOLAEVICH [ADDENDUM]

In the last quarter of the twentieth century the reputation of Boris Nikolaevich Chicherin underwent a remarkable revival, both in Russia and the West. Already before the collapse of the Soviet Union Chicherin fascinated those Soviet philosophers of law who sought stealthily to combine civil liberties with state power. That fascination, masked by an accompanying critique of Chicherin's bourgeois liberalism, was expressed in an important 1975 book by Valerii Dimitrievich Zor'kin. The collapse of the Soviet regime in 1991, the unexpected elevation of Zor'kin to the post of chief justice of the Russian Constitutional Court, the broad search by intellectuals for new ways to combine freedom and authority in the post-Soviet era, and a general scholarly reconsideration of the Russian national tradition in philosophy—all these factors contributed indirectly to the new interest in Chicherin's political thinking.

At the beginning of the new millennium Chicherin has found new admirers not among Russian liberals but among moderate conservatives who approve of his doctrine of the state as absolute or undivided sovereign, who applaud his pragmatic recognition that individual liberty must be balanced against the general needs of society, who share his support for capitalism constrained only by the needs of the economically defenseless, and who find his Realpolitik in statecraft wiser than dogmatic nationalism or naive internationalism. In addition to studying his political philosophy, post-Soviet Russian scholars have examined anew Chicherin's philosophy of history, making a much more positive assessment than before of his advocacy of a modified Hegelian approach to understanding the laws or regularities of historical development. The tendency has been to regard him as an important innovator, one of the originators of the influential state school of historical writing.

In the West Chicherin has been interpreted as the most important theoretician of liberalism in Russia, the figure who between 1855 and 1866 systematized hostility toward serfdom and defense of civil rights into a coherent liberal political program favoring the gradual introduction into Russia of the rule of law. Chicherin's program sharply distinguished between civil rights (freedom of conscience and speech) and political rights (freedom of suffrage, constitutional guarantees, and representative government). He argued that Russian political culture at midcentury was not yet mature enough for political rights but that it could responsibly uphold civil rights. This view, based on Baron de Montesquieu's notion that liberty rests on a complex relationship among the geographical, cultural, social, political, and historical institutions prevailing in a given country, made Chicherin unpopular with the radical left and recalcitrant right.

In 1882–1883 Chicherin warned in his two-volume book *Sobstvennost' i gosudarstvo* (Property and the state) that individual liberty in Europe and Russia was being endangered by "a new monster rising above the state: it is called 'society'" (Chicherin 1882, p. xix). His apprehension that social pressure for equality would soon destroy liberty bears strong resemblance to Alexis Tocqueville's (1805–1859) fear of the "tyranny of the majority." Consequently, during the last two decades of his life Chicherin stood as Russia's strongest advocate of individual liberty against society and the state. His program came to approximate what Friedrich Augustus von Hayek would later call "classical liberalism" or what other scholars would name "the old liberalism" in distinction to the new, social liberalism that came to prevail in the West after John Stuart Mill. The philosophical foundations of that program, both Hegelian and Kantian, were elucidated in his remarkable *Filosofiia prava* (Philosophy of law; 1900). In it Chicherin made plain his antipathy to the collectivist idealism of Plato, Johann Gottlieb Fichte, and Karl Christian Friedrich Krause; to Benthamite utilitarianism; to Rudolf von Jhering's (1818–1892) command theory of law; to Russian socialism in all its variants; to Marxism; and to Vladimir Sergeevich Solov'ëv's (Solovyov) mystical fusion of law and morality.

That Chicherin's name has been appropriated both by Russian *étatist* conservatives and Western individualists may point back to the "unresolved tension in his thought" (George Louis Kline's phrase) between Hegelian determinism and Kantian individualism, but may also be an indication of Chicherin's life-long effort to find an appropriate balance between authority and liberty, duty and right, the needs of society and the requirements of the individual. His conviction that it is impossible in politics to realize simultaneously all values in their fullness and that some values (e.g., liberty and equality) are irrec-

oncilable in principle anticipated the value pluralism of Isaiah Berlin.

Bibliography

WORKS BY CHICHERIN

Liberty, Equality, and the Market: Essays. Translated and edited by G. M. Hamburg. New Haven, CT: Yale University Press, 1998.

Sobstvennost' i gosudarstvo, Chast' 1. Moscow: Tipografiia Martynova, 1882.

WORKS ABOUT CHICHERIN

Hamburg, G. M. *Boris Chicherin and Early Russian Liberalism, 1828–1866*. Stanford, CA: Stanford University Press, 1992.

Walicki, Andrzej. "Boris Chicherin: The 'Old Liberal' Philosophy of Law." In *Legal Philosophies of Russian Liberalism*, 105–164. Oxford, U.K.: Clarendon Press, 1987.

Zor'kin, Valerii Dimitrievich. *Is istorii burzhuazno-liberal'noi politicheskoi mysli Rossii vtoroi poloviny XIX–nachala XX v* [From the history of bourgeois liberal political thought in Russia in the late 19th and early 20th century]. Moscow: Izdatel'stvo Moskovskogo Universiteta, 1975.

G. M. Hamburg (2005)

CHINESE PHILOSOPHY

This composite entry is comprised of the following sub-entries:

OVERVIEW

In its twenty-five hundred years of evolution Chinese philosophy has passed through four periods: the ancient period (until 221 BCE), when the so-called Hundred Schools contended; the middle period (221 BCE–960 CE), when Confucianism emerged supreme in the social and political spheres, only to be overshadowed in philosophy first by Neo-Daoism and then by Buddhism; the modern period (960–1900), when Neo-Confucianism was the uncontested philosophy, although by no means without variety or conflicts of its own; and the contemporary period (from 1912), when Neo-Confucianism, having become decadent and being challenged by Western philosophy, first succumbed to it, then was revived and reconstructed, but at mid century was overwhelmed by Marxism.

ANCIENT PERIOD: HUNDRED SCHOOLS (UNTIL 221 BCE)

The Hundred Schools, which included individual agriculturalists, diplomatists, military strategists, and other independent thinkers, had one thing in common, their primary concern with man both as an individual and as a member of society. This humanistic note was dominant from the earliest times and characterized all schools. The most prominent of the schools were the Confucianists, the Daoists, the Mohists, the Logicians, the Yin Yang school, and the Legalists.

Chinese thought at the dawn of civilization was dominated by the fear of spiritual beings. During the Shang dynasty (1751–1112 BCE) the Chinese would do nothing important without first finding out, through divination, the pleasure of the spirits. But when the Zhou overthrew the Shang, in 1112 BCE, human talent was needed to consolidate the newly established kingdom and to fight the surrounding barbarians. Human skill in irrigation proved to be more effective than praying to the spirits for rain. And the tribal anthropomorphic Lord (*Di*), who controlled human destiny at his whim, was now replaced by impartial and universal Heaven (*Tian*). The Mandate of Heaven (divine election) for the House of Zhou to rule rested on the moral ground that rule belongs to the man of virtue. In the final analysis, it was man's ability and virtue that counted. Humanism had reached a high pitch.

CONFUCIAN SCHOOL. The person who elevated humanism to the highest degree was Confucius (551–479 BCE). His central concerns were the "superior man" and a well-ordered society. Up to his time the ideal man was the aristocrat, the *junzi* (literally, "son of a ruler") a perfectly natural concept in a feudal society. In a radical departure from the past, Confucius formulated an entirely new ideal, the superior man, one who is wise, humane, and courageous, who is motivated by righteousness instead of profit, and who "studies the Way [Dao] and loves men." This conception of the superior man has never changed in the Confucian tradition.

Nature of the individual. Confucius never explained how it is possible for one to become a superior man. He seemed to imply that man is good by nature, but he said only that "by nature men are alike but through practice

they have become far apart." It was necessary to explain how we know that man can be good. Mencius (c. 372–c. 298 BCE), one of his two major followers, supplied that explanation. From the facts that all children know how to love their parents and that a man seeing a child about to fall into a well will instinctively try to save him, Mencius concluded that man's nature is originally good, possessing the "Four Beginnings"—humanity (*ren*), righteousness (*yi*), propriety (*li*), and wisdom—and the innate knowledge of the good and the innate ability to do good. Evil is due not to one's nature but to bad environment, lack of education, and "casting oneself away." The superior man is one who "develops his mind to the utmost" and "nourishes his nature."

Xunzi (c. 295–c. 238 BCE), although holding essentially the same idea of the superior man, contended that the original nature of man is evil. He argued that by nature man seeks for gain and is envious. Because conflict and strife inevitably follow, rules of propriety and righteousness have been formulated to control evil and to train men to be good. Propriety and righteousness are not native moral characteristics of man but the artificial efforts of sages. Thus, Xunzi was directly opposed to Mencius. Nevertheless, both were truly Confucian because their central objective was the good man.

Nature of society. Confucius wanted a society governed by men of virtue who, through personal examples and moral persuasion rather than law or punishment, would bring about the people's welfare and social order. Mencius, applying his theory of original goodness, reasoned that if a ruler applies his originally humane mind to the administration of his government, he will have a humane government, and what Confucius desired will naturally ensue. Xunzi, on the other hand, felt that since man's nature is evil, he needs rulers to regulate him by law and teachers to guide him by rules of propriety and righteousness. Once more he and Mencius were opposed, but again they aimed at the same thing—namely, a well-ordered society.

Relation of the individual and society. The Confucian school, then, is devoted to the harmonious development of the individual and society. This theme is systematically presented in the little classic *The Great Learning,* traditionally ascribed to the Confucian pupil Zengzi (505–c. 436 BCE). It consists of eight successive steps: the investigation of things, the extension of knowledge, the sincerity of the will, the rectification of the mind, the cultivation of the personal life, the regulation of the family, national order, and world peace. The goal is a harmonious world in which man and society are well developed and adjusted.

The harmony of the individual and society rests on several basic ideas. Foremost of these is humanity (*ren*). Confucius discussed humanity more than any other subject, and throughout history it has remained one of the key concepts in Confucianism. Previously the term connoted particular virtues, such as kindness, benevolence, and affection. Confucius interpreted it to mean the general virtue, the foundation of all particular virtues. Humanity is the moral character, which enables man to attain true manhood. The moral character is developed in oneself and in one's relations with others. A man of *ren,* "wishing to establish his own character, also establishes the character of others." Thus, *ren* has two aspects, conscientiousness (*zhong*) and altruism (*shu*).

Following Confucius, Mencius stressed humanity. But he almost always mentioned humanity and righteousness (*yi*) together, the first in the Confucian school to do so. By this time a clear distinction between what is good, correct, or proper and what is evil, incorrect, or improper had to be made. He wanted the innate sense of correctness fully exercised. Xunzi felt the same necessity to define correctness, but he sought to achieve this end through the precision of and distinctions made in law, rules of propriety, and music.

Another idea behind the harmony of the individual and society is the rectification of names. For Confucius it meant verifying or implementing an exact correspondence between titles of rank and actual fulfillment of responsibilities. Mencius, however, took "rectification" to mean correcting errors in one's heart (moral errors). Xunzi gave it a logical interpretation. To him rectification was distinguishing the concepts of names and actualities, similarities and differences, and particularity and generality. In doing this he developed the only logical aspect, in the formal sense, of ancient Confucianism. Confucius, Mencius, and Xunzi all believed that when names are rectified the positions of the individual and society will be well adjusted.

The third concept basic to social harmony is the mean (*zhongyong*). By this Confucius chiefly meant moderation as a guide to human action, but he implicitly referred to the ideals of centrality and harmony as well. The reference to centrality and harmony was greatly elaborated in the classic *The Doctrine of the Mean,* traditionally ascribed to Confucius's grandson Tzu-ssu (492–431 BCE). Centrality (*zhong*) consists in not deviating from the mean, and harmony (*yong*) exists in the common, the ordinary, and the universal. Centrality in the individual is

the state of equilibrium in one's mind before the feelings are aroused, and harmony is the state after they are aroused. In society centrality and harmony together mean complete concord in human relations. Ultimately, through the moral principle, heaven and earth will attain their proper order and all things will flourish in a harmonious universal operation. At this point the doctrine of the mean assumed metaphysical significance, which made it a profound influence on Neo-Confucianism.

When the individual behaves correctly and society operates in the right manner, the Way is said to prevail. The Way (Dao) is the moral law, or moral order. It is the Way of Heaven. Heaven was no longer conceived of as the anthropomorphic Lord (*Di*), the greatest of all spiritual beings. To Confucius, Heaven was the origin of all things the Supreme Reality, whose purposive character is manifested in the Way. The Supreme Being only reigns, leaving the Way to operate by itself. But no one can be separated from this Way, and for the Way to be meaningful it must be demonstrated by man. "It is man that can make the Way great," Confucius said. The note of humanism was sounded again.

DAOIST SCHOOL. To the Confucian school Dao was a system of moral truth, the expression of Heaven. To the Daoist school, however, it was Nature itself. Laozi (c. sixth century BCE), the founder of the school, equated Dao with Heaven, the "self-so" (*ziran*), and the One. It is eternal, spontaneous, nameless, and indescribable, at once the beginning of all things and the way in which they pursue their course. It is nonbeing, not in the sense of nothingness but in the sense of not being any particular thing. It is absolute and mystical. When it is possessed by an individual thing, it becomes that thing's character or virtue (*de*). The ideal life of the individual, the ideal order of society, and the ideal type of government are all based on it and guided by it. As the way of life it denotes simplicity, spontaneity, tranquility, weakness, and, most important of all, nonaction (*wuwei*), or, rather, letting Nature take its own course. Laozi's concept of Dao was so radically different from those of other schools that his school alone eventually came to be known as the Daoist school (Daojia).

Zhuangzi (born c. 369 BCE), Laozi's chief follower, took a step forward and interpreted Dao as the Way of unceasing transformation. In so doing he gave Dao a dynamic character. In the universal process of constant flux all things are equalized from the point of view of Dao. At the same time, since everything transforms in its own way, its individual nature is to be respected. Thus, in the ideas of Zhuangzi there is a curious combination of universality and particularity, a point that had far-reaching effect on later Daoist developments.

Although the Daoist school was definitely more transcendental than the Confucian, its chief concern, like that of the Confucian school, was man. Laozi discoursed mainly on government, and Zhuangzi discussed at great length the way to find spiritual freedom and peace. There is no desertion of society or the individual in Daoism.

The dominant notes in the Daoist school were, however, oneness and naturalness. It is not surprising that the Daoists strongly attacked other schools, particularly the Confucian, for making distinctions of all kinds. But so far as interest in man and society was concerned, the school agreed with the Confucian and other schools.

MOHIST SCHOOL. The Daoist school in time became strong enough to compete with Confucianism, but in the ancient period it was the Mohist school, founded by Mozi (c. 470–c. 391 BCE), that rivaled Confucianism in prominence. In practically all its major doctrines it stood opposed to Confucianism. The most serious and irreconcilable issue was that between the Mohist doctrine of universal love and the Confucian doctrine of love with distinctions. Mozi wanted people to love other people's parents as they love their own, whereas the Confucianists, especially Mencius, insisted that although one should show love to all, one should show special affection to his own parents. Otherwise there would be no difference between other people's parents and one's own, and family relationships would collapse.

In further opposition Mozi condemned religious rites and musical festivals as economically wasteful; the Confucianists held that ceremonies and music are necessary to provide proper expression and restraint in social behavior. This conflict on the practical level stemmed from the fundamental opposition of utilitarianism and moralism. In this issue, as in the issue of universal versus graded love, Mozi justified his doctrines on the basis of "benefits to Heaven, to spiritual beings, and to all men."

Mozi also attacked the Confucianists' teaching of humanity (*ren*) and righteousness (*yi*), for advocating them but for failing to recognize that humanity and righteousness originated with Heaven. As he repeatedly said, it is the will of Heaven that man should practice humanity and righteousness, be economical, and practice universal love, and it is man's duty to obey the will of Heaven. Of all the ancient schools only the Mohist placed ethics on a religious basis.

LOGICIANS. The Mohist doctrine of universal love was subscribed to by the Logicians. Their main interest, however, lay in a discussion of names and actualities. The school was small and has left little imprint, if any, on subsequent Chinese intellectual history. But it was the only school devoted to such metaphysical problems as existence, relativity, space, time, quality, actuality, and causes. Its most outstanding scholars were Hui Shi (c. 380–c. 305 BCE) and Gongsun Long (born 380 BCE). To Hui Shi things were relative, but to Gongsun Long they were absolute. The former emphasized change, whereas the latter stressed universality and permanence. The Logicians employed metaphysical and epistemological concepts that were primitive and crude, but they were the only group in ancient China interested in these concepts for their own sake.

YIN YANG SCHOOL. While the schools mentioned above were thriving, the Yin Yang school prevailed and influenced all of them. We know nothing about its origin or early representatives, but its ideas are simple and clear. Basically, it conceived of two cosmic forces, one yin, which is negative, passive, weak, and disintegrative, and the other yang, which is positive, active, strong, and integrative. All things are produced through the interaction of the two. Associated with the theory of yin and yang is that of the five agents, or elements (*wuzing*)—metal, wood, water, fire, and earth. According to this theory things succeed one another as the five agents take their turns. Originally the two doctrines were separate. It is generally believed that Zou Yan (305–240 BCE), the representative thinker of the Yin Yang school, was the one who combined the interaction of yin and yang with the rotation of the five agents.

Yin and yang were at first conceived as opposed to each other, succeeding each other, or complementary to each other. The five agents, too, were conceived as overcoming one another or producing one another. Eventually all alternatives were synthesized so that harmony reigns over conflict and unity exists in multiplicity. Yin, yang, and the five agents are forces, powers, and agents rather than material elements. The whole focus is on process, order, and laws of operation. Existence is viewed as a dynamic process of change obeying definite laws, following definite patterns, and based on a preestablished harmony.

One implication of this doctrine is the correspondence and at the same time the unity of man and Nature, for both are governed by the same process. Another is that the universe is a systematic, structural one, determinate, describable, and even predictable. Still another implication is that the universe is a perpetual process of rotation. Just as the five agents rotate, so history proceeds in cycles, and just as yin and yang increase and decrease, so things rise and fall. The Yin Yang school, more than any other, put Chinese ethical and social teachings on a cosmological basis. Generally speaking, its ideas have affected every aspect of Chinese life, be it metaphysics, art, marriage, or even cooking. Wherever harmony is sought or change takes place, the forces of yin and yang are at work.

LEGALIST SCHOOL. Philosophically the Legalist school is the least important because it had no new concept to offer. In fact, it did not concern itself with ethical, metaphysical, or logical concepts, as other schools do. Its chief objective was the concentration of power in the ruler. Within the Legalist school there were three tendencies—the enforcement of law with heavy reward and punishment, the manipulation of statecraft, and the exercise of power. The school, called Fajia (meaning school of law) in Chinese, had many representatives, some of them prime ministers, but the most outstanding was Han Feizi (died 233 BCE), who combined the three tendencies of his school.

The Legalist school assumed the evil nature of man and rejected moral values in favor of concrete results. In insisting that laws be applicable to all, it unwittingly subscribed to the doctrine of the equality of all men, and in insisting that assignments be fulfilled with concrete results, it strengthened the doctrine of the correspondence of names and actualities. There is no doubt that compared to other schools, it looked to circumstances rather than principles and to the present rather than the past. It agreed with them in one respect, that life is in a process of constant change.

The Legalists helped the Qin to liquidate the feudal states and establish a new dynasty in 221 BCE. The Qin enforced the Legalists' totalitarian philosophy, suppressed other schools, and burned their books in 213 BCE. The contest of the Hundred Schools now came to an end.

MIDDLE PERIOD (221 BCE–960 CE)

The Legalists ruled the Qin with absolute power and tolerated no other schools, but other schools were by no means totally absent from the scene. When the dynasty was overthrown by the Han in 206 BCE, some of these schools reemerged, carrying with them a crosscurrent of thought. The result was a syncretic movement.

SYNCRETIC CONFUCIANISM. Confucianism became the state ideology in 136 BCE It was supreme in government, society, education, and literature and remained so until the twentieth century. But philosophically it was almost overwhelmed by the doctrine of yin and yang. This can readily be seen in the philosophies of the *Book of Changes* and Dong Zhongshu.

The *Book of Changes* (*Yijing*) is a Confucian classic, but the Daoists also made much use of it. (Tradition ascribes part of the work to Confucius, but it was most probably composed several centuries later, although portions may have been in existence in Confucius's lifetime.) It shows the strong impact of the Yin Yang school. According to the *Book of Changes* creation of the world begins with the Great Ultimate (*taiji*), which engenders yin and yang. Yin and yang, in their turn, give rise to the four forms of major and minor yin and yang. The four forms produce the eight elements (*bagua*), which, through interaction and multiplication, produce the universe. The cosmogony is naive and elementary, but it introduced into Confucianism the strong features of Daoist naturalism and the interaction of yin and yang. Since then the Confucianists have viewed the universe as a natural and well-coordinated system in which the process of change never ceases.

The syncretic spirit was also strong in Dong Zhongshu (179–104 BCE), the most outstanding Confucian philosopher of the period. He combined the Confucian doctrines of ethics and history with the ideas of yin and yang. Greed and humanity, the two foremost moral qualities, he correlated with yin and yang, respectively. Likewise, he equated human nature and feelings with yang and yin and thereby with good and evil. All things are grouped into pairs or into sets of five to correspond to yin and yang and the five agents. Ultimately they are reduced to numbers. In this arrangement historical periods parallel the succession of the five agents, and man, the microcosm, corresponds to Nature, the macrocosm. But Dong went beyond the idea of mere correspondence. To him, things of the same kind activate each other. There is the universal phenomenon of mutual activation and influence that makes the universe a dynamic, organic whole.

Unfortunately, this doctrine soon degenerated into superstition. Early in the Han dynasty (206 BCE–220 CE) there was a wide belief in prodigies, which were taken to be influences of Nature on man or vice versa. Wang Chong (27 CE–c. 100 CE), an independent thinker, revolted against this. He declared that Heaven (Nature) takes no action and that natural events, including prodigies, occur spontaneously. Man is an insignificant being in the vast universe, and he does not influence Nature or become a ghost at death to influence people. In addition, Wang Chong insisted that any theory must be tested by concrete evidence, and he supported his own theories with numerous facts. Thus, he raised rationalistic naturalism to a height never before reached in Chinese history and prepared for the advent of rationalistic and naturalistic Neo-Daoism, which was to replace Confucian philosophy.

NEO-DAOISM. Under the influence of the doctrine of the correspondence of man and Nature and the belief in prodigies, Han dynasty thinkers were chiefly concerned with phenomena. Thinkers of the Wei-Jin period (220–420), however, went beyond phenomena to find reality behind space and time. They were interested in what is profound and abstruse (*xuan*), and consequently their school is called Xuan Xuanxue ("profound studies") or the Metaphysical school. They developed their doctrines in their commentaries on the *Laozi,* the *Zhuangzi,* and the *Book of Changes,* the "three profound studies." To Wang Bi (226–249), the most brilliant Neo-Daoist, ultimate reality is original nonbeing (*benwu*). It is not nothingness but the pure being, original substance, which transcends all distinctions and descriptions. It is whole and strong. And it is always correct because it is in accord with principle (*li*), the universal rational principle that unites all particular concepts and events. The note of principle was a new one. It anticipated Neo-Confucianism, which is based entirely on it.

Guo Xiang (died 312), another famous Neo-Daoist, developed his theory in his comments on Zhuangzi's doctrine of self-transformation. To Guo Xiang, things transform themselves according to principle, but each and every thing has its own principle. Everything is therefore self-sufficient, and there is no need for an overall original reality to combine or govern them, as Wang Bi believed. Whereas Wang Bi emphasized nonbeing, the one, and transcendence, Guo Xiang emphasized being, the many, and immanence.

As a movement Neo-Daoism did not last long, but its effect on later philosophy was great. It raised the Daoist concepts of being and nonbeing to a higher level and thereby formed the bridge between Chinese and Buddhist philosophies.

BUDDHISM. In the first several centuries Buddhism existed in China as a popular religion rather than as a philosophy. When Buddhists came into contact with the Chinese literati, especially the Neo-Daoists, in the third

century, they matched Buddhist concepts with those of Daoism, identifying *Tathatā* (Thusness, *Nirvāṇa*) with the Daoist "original nonbeing," for example. Under Neo-Daoist influence, early Buddhist schools in China all engaged in discussions on being and nonbeing.

Middle Doctrine and Dharma Character. The problems of being and nonbeing largely characterize the two major Buddhist schools that developed in China in the sixth century, the Middle Doctrine (Zhonglun), or Three Treatise (San-lun), school and the Dharma Character (Faxiang), or Consciousness Only (Weishi), school. The Middle Doctrine school, systematized by Jizang (549–623), was based on three Indian scriptures—the *Mādhyamika Śāstra* (Treatise on the Middle Doctrine), by Nāgārjuna (c. 100–200), the *Dvādaśamikāya Śāstra* (Twelve gates treatise), also by Nāgārjuna, and the *Śata Śāstra* (One-hundred verses treatise), by Ārya-deva (exact dates unknown), a pupil of Nāgārjuna. This school regarded both being and nonbeing as extremes whose opposition must be resolved in a synthesis. The synthesis, itself a new extreme with its own antithesis, needs to be synthesized also. In the end all oppositions are dissolved in the True Middle or emptiness. The school was essentially nihilistic and is often called the school of Nonbeing.

In contrast, the Consciousness Only school, which was founded by Zuangzang (596–664), regarded all dharmas (elements of existence) and their characters—that is, the phenomenal world—as real, although only to a certain degree because they are illusory, apparent, and dependent. The school divides the mind into eight consciousnesses, the last of which contains "seeds" or effects of previous deeds and thoughts that affect future deeds and thoughts. Future deeds and thoughts are "transformations" of present ones, and present ones are "transformations" of past ones. When an individual attains perfect wisdom all transformations are transcended. In these transformations dharmas are produced. Some, the products of imagination, have only illusory existence. Others have dependent existence because they depend on causes for their production. But those of the "nature of perfect reality" have true existence. Since the school accepts dharmas and their character as real, it is often called the school of Being.

In spite of the fact that their basic problems of being and nonbeing are Chinese, the two schools were essentially no more than Indian schools transplanted to Chinese soil. They lacked the spirit of synthesis and were too extreme for the Chinese, and they declined after a few centuries, a relatively short time compared to other schools. In the meantime the Chinese spirit of synthesis asserted itself, notably in the Tiantai (Heavenly Terrace) and Huayan (Flower Splendor) schools.

Tiantai. According to the Tiantai school, which was founded by Zhiyi (538–597) in the Tiantai Mountains, dharmas are empty because they have no self-nature and depend on causes for production. This is the Truth of Emptiness. But since they are produced, they do possess temporary and dependent existence. This is the Truth of Temporary Truth. Thus, dharmas are both empty and temporary. This is the Truth of the Mean. Each truth involves the other two so that three are one and one is three. This mutual identification is the true state of all dharmas. In the realm of temporary truth—that is, the phenomenal world—all realms of existence, whether of Buddhas, men, or beasts, and all characters of being, such as cause, effect, and substance, involve one another, so that each element, even an instant of thought, involves the entire universe. This all-is-one-and-one-is-all philosophy is expressed in the famous saying "Every color or fragrance is none other than the Middle Path."

Huayan. In the same spirit of synthesis, the Huayan school, established by Fazang (596–664), propagated the doctrine of the universal causation of the realm of dharmas. This realm is fourfold. It contains the realm of facts, the realm of principle, the realm of principle and facts harmonized, and the realm of all facts interwoven and mutually identified. Principle is emptiness, static, the noumenon, whereas facts are specific characters, dynamic, constituting the phenomenal world. They interact and interpenetrate and in this way form a perfect harmony. This doctrine rests on the theory of the six characters, which states that each dharma possess the six characteristics of universality, speciality, similarity, difference, integration, and disintegration. Thus, each dharma is both one and all. The world is in reality a perfect harmony in all its flowery splendor.

Chan. Whereas Buddhist philosophy in the sixth and seventh centuries came to be more and more Chinese with the Tiantai and Huayan schools, Confucian philosophy remained dormant. In the eighth and ninth centuries its very life was threatened by the growth of Chan, or the Meditation school (Zen in Japan).

The Meditation doctrine, introduced from India by Bodhidharma (fl. 460–534), aimed at the realization of the Ultimate Reality through sitting in meditation. Its emphasis was on concentration to the point of absence of thought in order to get rid of attachments. As the Meditation school developed it conceived of the mind as split into the true mind, which does not have thought or attachments to the characters of dharmas, and the false

mind, which has them. Sitting in meditation was the effort to get rid of them.

Hui Neng (638–713), an aboriginal from the south, rose in revolt against the tradition. He and his followers refused to divide the mind but maintained that it is one and originally pure. Erroneous thoughts and erroneous attachments are similar to clouds hiding the sun. When they are removed the original nature will be revealed and great wisdom obtained. The way to discover the original nature is calmness and wisdom. Calmness does not mean not thinking or having nothing to do with the characters of dharmas. Rather, it means not being carried away by thought in the process of thought and being free from characters while in the midst of them. Sitting in meditation is useless, and external effort, such as reciting scriptures or worshiping Buddhas, is futile. When the mind is unperturbed by selfishness or deliberate effort and is left to take its own course, it will reveal its pure nature, and enlightenment will come suddenly. Instead of assuming a dualistic nature of the mind, ignoring the external world, and aiming at uniting with the Infinite, as Indian meditation did, Chinese meditation assumed the original goodness of nature, took place in the midst of daily affairs, and aimed at self-realization.

Chinese influences on Chan are obvious. Buddhism had become characteristically Chinese, with its interest in the here and now. It swept all over China. The Confucian Way was in imminent danger of disappearance. Han Yu (768–824), the greatest Confucianist of the Tang dynasty (618–907), had to defend the Confucian Way and demanded that Buddhist and Daoist books be burned. His contribution to Confucian philosophy is negligible, but he paved the way for Confucian awakening.

MODERN PERIOD: NEO-CONFUCIANISM (960–1912)

The combination of the wide spread of Chan and the attractiveness of the Huayan and Tiantai metaphysics, as well as the Chan psychology, woke the Confucianists from a long slumber. For centuries, within the Confucian school itself, efforts had been confined to textual studies and flowery compositions. Reaction, long overdue, now set in. Consequently in the early years of the Song dynasty (960–1279) Confucianists raised new problems and attempted to find solutions.

Since the *Book of Changes* had exerted tremendous influence throughout the ages, the Confucianists naturally turned to it for inspiration and support. But instead of using it for divination, as the Daoists did, they used it for a study of human nature and destiny on the basis of principle. This new movement eventually came to be known as the school of Nature and Principle (Xingli Xue or, in English, Neo-Confucianism).

The man who opened the vista and determined the direction of Neo-Confucianism was Zhou Dunyi (also called Zhou Lianxi, 1017–1073). Elaborating on the cosmogony of the *Book of Changes,* he held that in the evolution of the universe from the Great Ultimate through the two material forces of yin and yang and the five agents to the myriad things, the five agents are the basis of the differentiation of things, whereas yin and yang constitute their actuality. The two forces are fundamentally one. Consequently the many are ultimately one and the one is actually differentiated in the many. Both the one and the many have their own correct states of being. The nature and destiny of man and things will be correct in their differentiated state if they all follow the same universal principle. This was the central thesis of Neo-Confucianism for the next several centuries. The influence of the Buddhist one-in-all-and-all-in-one philosophy is unmistakable.

RATIONALISTIC NEO-CONFUCIANISM. Neo-Confucianism developed in two different directions, the rationalistic school of Principle and the idealistic school of Mind.

Cheng–Zhu philosophy. The central figures in the rationalistic movement were Cheng Yi (Cheng Yichuan, 1033–1107), who formulated the major concepts and provided the basic arguments, and Zhu Xi (1130–1200), who supplemented and refined them and brought Neo-Confucianism into a systemic, rationalistic whole. At the center of the school is its concept of principle (*li*); its other major concepts are the Great Ultimate, material force, the nature of man and things, the investigation of things, and the moral quality of humanity, or *ren.*

The idea of principle, virtually absent in ancient Confucianism, probably came from Neo-Daoism and Buddhism. If so, it was employed to oppose them. In the view of the Neo-Confucianists of the Song dynasty both Daoist nonbeing and Buddhist emptiness are too abstract, but their principle is concrete. Cheng Yi repeatedly said that for a thing to exist there must first be its principle, the law according to which it will exist. Principle is definite, correct, self-evident, and self-sufficient. It is in each and every thing. Put differently, the principle for each particular thing is a definite one.

Since the possible number of things in the world is infinite, the number of actual and potential principles is infinite. As new things appear, new principles are realized.

In the production and reproduction in the universe the process of daily renewal never ceases. This is a principle in itself, and there is always a new principle to make a new thing possible. But all principles are at bottom one, called the Great Ultimate. As substance the Great Ultimate is one, but as it functions it is manifested in the many, or the innumerable concrete things. The Great Ultimate is both the sum total of all principles and principle in its oneness.

The manifestations of the Great Ultimate depend on material force, which actualizes things. Operating as yin and yang, material force provides the stuff that makes a thing concrete. Things differ from one another because of their material endowments, and they resemble one another because of principle. Principle as the Great Ultimate exists before physical form (*xing er shang*), whereas material force exists after physical form (*xing er xia*). Logically speaking, principle is prior to material force, but as Zhu Xi emphasized, they are never separate. Without material force principle would be neither concrete nor definite, and without principle there would be no law by which material force could operate. In the universe there has never been any material force without principle or principle without material force.

When principle is endowed in man it becomes his nature. Man's nature is originally good because principle is good, and principle is good because it is the source of all goodness. Evil arises when feelings are aroused and deviate from principle. In this respect Neo-Confucianism retains the traditional Confucian doctrine that Nature is good whereas feelings are sources of evil. The Song Neo-Confucianists made a sharp distinction between the principle of Nature and selfish human desires.

Through moral cultivation selfish desires can be eliminated and the principle of Nature realized. To the rationalistic Neo-Confucianists the first step toward cultivation was the investigation of things (*gewu*). According to Cheng Yi every blade of grass and every tree possesses principle. Therefore, all things should be investigated. One can investigate by studying inductively or deductively, by reading books, or by handling human affairs. When things are investigated, as *The Great Learning* taught, one's knowledge will be extended, one's will sincere, one's feelings correct, and one's personal life cultivated. When this is done one will have fully developed one's nature and fulfilled one's destiny.

The development of human nature, according to the Cheng Yi–Zhu philosophy, does not stop with personal perfection but involves all things. This is where the concept of *ren* comes in. To Cheng Yi and Zhu Xi, as to previous Confucianists, *ren* is humanity, the moral quality that makes man a true man. But under the influence of the century-old Confucian doctrine of the unity of man and Nature and also the cosmological scale of Buddhist ethics, the Neo-Confucianists applied the concept of *ren* to all things and said that through it man can "form one body with heaven, earth, and all things." Furthermore, they added a new note to *ren* by interpreting the word in its other sense, that of seed or growth. *ren* was then understood to be the chief characteristic of heaven and earth, the production and reproduction of things. This life-giving character is the highest good. It is inherent in man's nature. Man's duty is to develop it and put it into practice. Neo-Confucianism returned to the chief topic and fundamental ethical concern of Confucius and gave it new meaning.

As has been indicated, Zhu Xi and Cheng Yi were the chief figures of rationalistic Neo-Confucianism. However, Cheng Yi's older brother Cheng Hao, their uncle Zhang Zai, and Shao Yong, who with Cheng Yi and Zhou Dunyi are called the Five Masters of early Song Neo-Confucianism, also contributed substantially to it.

Cheng Hao. Cheng Hao (Cheng Mingdao, 1032–1085) shared many ideas with his brother. The two were really the twin leaders of the school in its formative stage. Whereas Cheng Yi stressed the idea of principle as one and its manifestations many, Cheng Hao stressed principle as production and reproduction. He saw the spirit of life in everything, which impressed him much more than the rational character of things. Furthermore, to Cheng Hao the highest principle was the principle of Nature, a concept he evolved himself. He believed that principle is more than the rational basis of being. It is the principle of Nature, the self-evident universal truth that carries with it the dictate to distinguish right from wrong and the imperative to do good. Instead of focusing his attention on the investigation of things, he directed it to the calmness of mind. Only when the mind is calm—that is, free from selfishness, cunning, and deliberate effort—can it be peaceful. One can then respond to things as they come and naturally maintain a balance between the internal and the external. Cheng Hao considered understanding the nature of *ren* to be of the greatest importance. The man who has such an understanding will be free from all opposition between the self and the other and will be able to form one body with all things. It can easily be seen that although he differed from his brother on many points, Cheng Hao strengthened Neo-Confucianism by providing it with warmth and spirituality.

Zhang Zai. Unlike the Cheng brothers, Zhang Zai (Zhang Hengqu, 1020–1077) regarded principle not as

above or different from material force but as the law according to which material force operates. He identified material force with the Great Ultimate and considered yin and yang as merely the two aspects of material force. As substance, before consolidation takes place, material force is the Great Vacuity (*taixu*). As function, in its activity and tranquility, integration and disintegration, and so forth, it is the Great Harmony. But the two are the same as the Way (Dao). In its ultimate state material force is one, but in its contraction and expansion and the like it is manifested in the many. Similarly, in ethics *ren* is one, but in its application in the various human relations, as filial piety toward parents, brotherly respect toward brothers, and so on, it is many. Zhang Zai's advocacy of the concept of vacuity was too Daoistic to be attractive to his fellow Neo-Confucianists, but in making the doctrine of the one and the many the metaphysical foundation of Confucian ethics, he made "a great contribution to the Confucian school," in Zhu Xi's description.

Shao Yong. Shao Yong (1011–1077) agreed with his contemporaries that there are supreme principles governing the universe, but he added that they can be discerned in terms of numbers. In his cosmology change is due to spirit; spirit gives rise to number, number to form, and form to concrete things. Since the Great Ultimate engenders the four forms of major and minor yin and yang, Shao Yong used the number 4 to classify all phenomena. In his scheme there are the four seasons, the four heavenly bodies, the four kinds of rulers, the four periods of history, and so on. Since the structure of the universe is mathematical, elements of the universe can be calculated and objectively known. The best way to know is to "view things as things." All these are new notes in Neo-Confucianism that set Shao Yong apart from the rest. He was as much interested in the basic problems of principle, nature, and destiny as were other Neo-Confucianists. However, he hardly discussed social and moral problems, and his whole metaphysical outlook was too near Daoist occultism to be considered part of the main current of rationalistic Neo-Confucianism.

IDEALISTIC NEO-CONFUCIANISM. In spite of the fact that the rationalistic Neo-Confucianists tried to maintain a balance between principle and material force in metaphysics and between the investigation of things and moral cultivation in the way of life, they tended to be one-sided in their emphasis on principle and the investigation of things.

Lu Xiangshan. Opposition to these trends arose in Zhu Xi's own time, notably from his friend and chief opponent, Lu Xiangshan (Lu Jiuyuan, 1139–1193). Cheng Yi and Zhu Xi had regarded mind as the function of man's nature, which is identical with principle. To Lu mind *was* principle. It is originally good and endowed with the innate knowledge of the good and the innate ability to do good, as Mencius had taught long before. It is one and indissoluble. There is no such distinction as that between the moral mind, which is good, and the human mind, which is liable to evil, a distinction made by Zhu Xi. Both the principle of Nature and human desires are good, and they should not be contrasted, as they were by Zhu Xi. The mind fills the whole universe. Throughout all ages and in all directions there is the same mind. It is identical with all things, for there is nothing outside the Way and there is no Way outside things. In short, the mind is the universe. To investigate things, then, is to investigate the mind. Since all principles are inherent and complete in the mind, there is no need to look outside, as did Cheng Yi and Zhu Xi.

This thoroughgoing idealism shows not only the influence of Mencius but also the impact of Buddhism. However, Lu was no less a critic of Buddhism than were other Neo-Confucianists. Actually, he criticized Zhu Xi not to promote Buddhism but to uphold Confucianism. In his opinion the way of Zhu Xi led to a divided mind, aimless drifting, and devotion to isolated details that meant little to life. Lu advocated instead a simple, easy, and direct method of recovering one's originally good nature. It consisted in having a firm purpose, "establishing the nobler part of one's nature," and coming to grips with fundamentals. In short, Zhu's way was "following the path of study and inquiry," whereas Lu's way was "honoring the moral nature."

Lu's opposition did not have any immediate effect, for rationalistic Neo-Confucianism was too strong to be checked. It dominated the Chinese intellectual world for several hundred years. By the fifteenth century, however, it had degenerated into concern only with isolated details and had lost touch with the fundamentals of life. There was no longer any intellectual creativity or moral vigor in it.

Wang Yangming. Opposition rose again, this time from Wang Yangming (Wang Shouren, 1472–1529), who pushed the idealistic movement to its highest point in Chinese history. Wang reiterated most of Xiangshan's ideas but carried some of them to new heights. Like Lu, he said that the mind is principle and that things are in the mind, but he emphasized the direction of the mind—that is, the will. To him a thing (or affair) was nothing but the mind determined to realize it. There is no such thing

as filial piety, for example, unless one is determined to put it into practice and actually does so. Like Lu, Wang said that the investigation of things is the investigation of the mind; however, he added that since the most important aspect of the mind is the will, the sincerity of the will must precede the investigation of things, an idea diametrically opposed to Zhu Xi's contention that as things are investigated, one's will becomes sincere. Going beyond Mencius's doctrine of the innate knowledge of good, Wang held that because of one's innate ability to do good, one necessarily extends the innate knowledge into action. Knowledge and action are really identical; one is the beginning and the other the completion. Here are two original doctrines, the extension of the innate knowledge and the unity of knowledge and action, both of which represent new steps in Chinese thought.

Wang Fuzhi. For 150 years the idealistic philosophy of Wang Yangming dominated China, putting Zhu Xi's rationalism on the defensive. A number of philosophers attempted compromise, without much success. In the seventeenth century Wang's idealism declined, and Zhu Xi's rationalism reasserted itself. But rationalism enjoyed neither monopoly nor prominence, for revolts arose one after another. From the seventeenth century on, Confucianists regarded both Zhu and Wang as too speculative. The spirit of the time demanded the evident, the concrete, and the practical.

One of the first to rebel was Wang Fuzhi (Wang Chuanshan, 1619–1692). He rejected the central Neo-Confucian thesis that principle is a universal, transcending and prior to material force. Instead, he contended that principle is identical with material force. It is not a separate entity that can be grasped but the order and arrangement of things. The Great Ultimate and the principle of Nature are no transcendent abstractions. They, along with the mind and the nature of things, are all within material force. Wang Fuzhi boldly declared, "The world consists only of concrete things." He also refused to accept either the distinction between the principle of Nature and human desires or the subordination of human desires.

Dai Zhen. In the same spirit, Dai Zhen (Dai Dongyuan, 1723–1777) attacked the Neo-Confucianists, particularly those of the Song dynasty, for their conception of principle. He said that they looked upon principle "as if it were a thing." To him principle was nothing but the order of things, and by things he meant daily affairs, such as drinking and eating. The way to investigate principle, he thought, is not by intellectual speculation or by introspection of the mind but by critical, analytical, minutely detailed, and objective study of things based on concrete evidence. Dai Zhen's conception of principle led him to oppose vigorously the Neo-Confucianists' view of human feelings and desires, which he thought they had undermined. In his belief principle can never prevail when feelings are not satisfied, for principles are merely "feelings that do not err." Dai Zhen perpetuated the Neo-Confucian doctrine that the universe is an unceasing process of production and reproduction, except that to him Nature, like principle, was but an order.

Kang Youwei. By the end of the nineteenth century there was a swing back to the philosophy of Wang Yangming. The sad situation in China called for dynamic and purposive action that only an idealism like Wang's could provide. All of these factors conditioned the thought of Kang Youwei (1858–1927), the greatest Confucianist of the time. In an attempt to translate Confucian philosophy into action he enunciated the extraordinary theory that Confucius was first and last a reformer. Kang himself engineered the abortive political reform of 1898. Obviously influenced by the Christian concepts of utopia and progress, he envisaged the Age of Great Unity. In his theory of historical progress history proceeds from the Age of Chaos to the Small Peace and finally to the Great Unity, when nations, families, classes, and all kinds of distinctions will be totally abolished. The philosophical basis for this utopia is his interpretation of *ren.* He equates it with what Mencius called "the mind that cannot bear" to see the suffering of others. It is compassion. It is also the power of attraction that pulls all peoples together. As such it is ether and electricity, which permeate all things everywhere.

Kang was philosophically superficial but historically important. He showed that at the turn of the twentieth century China was at a philosophical crossroad.

CONTEMPORARY PERIOD (FROM 1912)

Philosophy in twentieth-century China was indeed confusing and chaotic, but certain tendencies could clearly be seen. There was first of all importation from the West. In the first three decades Charles Darwin, Ernst Heinrich Haeckel, Friedrich Nietzsche, Arthur Schopenhauer, Henri Bergson, Immanuel Kant, René Descartes, William James, John Dewey, Karl Marx, and others were introduced, each with his champion. Of these, James and Dewey were the most influential, since pragmatism was advocated by Hu Shih, leader of the intellectual revolution. Only Marxism, however, has remained strong, and it has become the established state philosophy.

Under the stimulation of Western philosophy both Confucianism and Buddhism resurged from a long

period of decadence. In the 1920s and early 1930s, Ouyang Jingwu (1871–1943), strongly impressed by Western idealism, sought to revive Buddhist idealism as it was centuries ago, and his opponent, Abbot Taixu (1889–1947), attempted to transform Buddhist idealism in the light of Western philosophy. Since neither knew Western philosophy or was really a philosopher, their movements, though extensive and vigorous, resulted more in religious reform than in intellectual advancement, and in the late 1930s their work quickly disappeared from the philosophical scene.

The renewal of Confucian philosophy, however, was different. Feng Youlan (1895–1990) developed his philosophy on the basis of rationalistic Neo-Confucianism, and Xiong Shili (1885–1968) built his on the foundation of idealistic Neo-Confucianism. Since the 1930s they became the two most prominent philosophical thinkers in China. While importation from the West and reconstruction of traditional philosophy were going on, certain philosophers tried to evolve their own systems out of Western thought. The most successful of these was Zhang Dongsun (1886–1962), who alone produced a comprehensive and mature philosophy.

Feng Youlan. Trained in philosophy at Columbia University, Feng Youlan derived his rationalism from the Neo-Confucianism of Cheng Yi and Zhu Xi and converted Neo-Confucian concepts into formal logical concepts. According to him, his "new rationalistic Confucianism" is based on four main metaphysical concepts—principle, material force, the substance of Dao, and the Great Whole. The concept of principle is derived from the Cheng-Zhu proposition "As there are things, there must be their specific principles." A thing must follow principle, but principle does not have to be actualized in a thing. It belongs to the realm of reality but not actuality and is purely a formal concept. The concept of material force is derived from the Cheng-Zhu proposition "If there is principle, there must be material force" by which a thing can exist. Material force is basic to the concept of existence but does not itself exist in the actual world. It is only a formal logical concept. The concept of Dao means a "universal operation," the universe of "daily renewal" and incessant change. Finally, the Great Whole, in which one is all and all is one, is also a formal concept, being the general name for all, not an assertion about the actual world. It corresponds to the Absolute in Western philosophy.

Basically, Feng's philosophy is a combination of Neo-Confucianism and Western realism and logic. Feng called his own system a "new tradition." It is new not only because it has interpreted Neo-Confucian ideas as formal concepts. In addition, Feng's system has replaced Neo-Confucianism, which is essentially a philosophy of immanence, with a philosophy of transcendence. To Feng the world of actuality is secondary.

In 1950, Feng repudiated his philosophy because it "neglects the concrete and the particular," but in 1957 he still maintained that Confucius was an idealist rather than a materialist. This suggests that he was not entirely Marxian in his interpretation of Chinese thought. He remained the most important Chinese philosopher of the last thirty years—the most original, the most productive, and the most criticized.

Xiong Shili. Xiong Shili called his philosophy the "new doctrine of consciousness-only." According to his main thesis reality is endless transformation of closing and opening, which constitute a process of unceasing production and reproduction. The original substance is in perpetual transition at every instant, continually arising anew and thus resulting in many manifestations. But reality and manifestations, or substance and function, are one. In its closing aspect original substance has the tendency to integrate, resulting in what may temporarily be called matter, whereas in its opening aspect it has the tendency to maintain its own nature and be its own master, resulting in what may temporarily be called mind. This mind itself is one part of the original mind, which in its various aspects is mind, will, and consciousness.

Xiong's terminology comes from the *Book of Changes* and the Buddhist Consciousness Only school, but his basic ideas—the unity of substance and function and the primacy of the original mind—come from Neo-Confucianism, especially that of Wang Yangming. He avoided Zhu Xi's bifurcation of principle and material force and Wang's subordination of material force to the mind and has provided the dynamic idea of change in Neo-Confucianism with a metaphysical foundation.

Zhang Dongsun. The theory of Zhang Dongsun (born 1886) has been variously called revised Kantianism, epistemological pluralism, and panstructuralism. Chiefly formulated between 1929 and 1947, it is derived from Kant but rejects Kant's bifurcation of reality into phenomena and noumena and Kant's division of the nature of knowledge into the a posteriori and the a priori. To Zhang knowledge is a synthesis of sense data, form, and methodological assumptions. Perception, conception, mind, and consciousness are all syntheses, or "constructs," and constructs are products of society and culture. He maintained that although he combined Western logic with modern psychology and sociology, his sys-

tem was his own. During World War II he shifted more and more from metaphysics to the sociology of knowledge and thus was drawn closer and closer to Marxism.

During the years since World War II neither Xiong's, Feng's nor Zhang's philosophy has become a movement, although Xiong has exercised considerable influence on a number of young philosophers. While Zhang is keeping silent, Xiong maintaining his position, and Feng still reconsidering his philosophy, Marxism has become the triumphant and official system of thought. It demands that philosophy be practical, scientific, democratic, and for the masses. Traditional philosophy is being studied and will survive, but it is being interpreted in a new light.

See also Buddhism; Communism; Mysticism, History of.

Bibliography

Chan, Wing-tsit, ed. *A Source Book in Chinese Philosophy.* Princeton, NJ: Princeton University Press, 1963.

Creel, H. G. *Chinese Thought from Confucius to Mao Tse-tung.* Chicago: University of Chicago Press, 1953.

De Bary, Wm. Theodore, Wing-tsit Chan, and Burton Watson, comps. *Sources of Chinese Tradition,* with contributions by Yi-pao Mei, Leon Hurvitz, T'ung-tsu Ch'u, and John Meskill. 2 vols. New York: Columbia University Press, 1960.

Fung, Yu-lan. *A Short History of Chinese Philosophy.* New York: Macmillan, 1948.

Fung, Yu-lan. *A History of Chinese Philosophy,* trans. Derk Bodde. Vol. 1, The Period of the Philosophers, and vol. 2, The Period of Classical Learning. Princeton, NJ: Princeton University Press, 1937, 1952, 1953.

Graham, A. C. *Disputers of the Tao.* La Salle, IL: Open Court, 1989.

Graham, A. C. *Studies in Chinese Philosophy and Philosophical Literature.* Albany: State University of New York Press, 1990.

Hall, David L., and Roger T. Ames. *Anticipating China: Thinking through the Narratives of Chinese and Western Culture.* Albany: State University of New York Press, 1995.

Hall, David L., and Roger T. Ames. *Thinking from the Han: Self, Truth and Transcendence in China and the West.* Albany: State University of New York Press, 1998.

Hansen, Chad. *A Daoist Theory of Chinese Thought.* New York: Oxford University Press, 1992.

Hsiao, Kung-chuan. *A History of Chinese Political Thought.* Vol. 1, From the Beginnings to the Sixth Century A.D. Translated by F. W. Mote. Princeton, NJ: Princeton University Press, 1979.

Liu, Shu-hsien. *Understanding Confucian Philosophy: Classical and Sung-Ming.* Westport, CT: Greenwood Press, 1998.

Loewe, Michael, ed. *Early Chinese Texts: A Bibliographic Guide.* Berkeley: University of California Press, 1993.

Mote, Frederick W. *Intellectual Foundations of China.* 2nd ed. New York: McGraw-Hill, 1971, 1989.

Munro, Donald J. *The Concept of Man in Early China.* Stanford, CA: Stanford University Press, 1969.

Needham, Joseph. *Science and Civilisation in China.* Vol. 2, History of Scientific Thought. Cambridge, MA: Cambridge University Press, 1956.

Nivison, David S. *The Ways of Confucianism.* Chicago and La Salle, IL: Open Court, 1996.

Roetz, Heiner. *Confucian Ethics of the Axial Age: A Reconstruction under the Aspect of the Breakthrough toward Postconventional Thinking.* Albany: State University of New York Press, 1993.

Schwartz, Benjamin. *The World of Thought in Ancient China.* Cambridge, MA: Belknap Press of Harvard University Press, 1985.

Shun, Kwong-loi. *Mencius and Early Chinese Thought.* Stanford, CA: Stanford University Press, 1997.

Van der Leeuw, Karel L. "The Study of Chinese Philosophy in the West: A Bibliographic Introduction." *China Review International,* 6, (2) (Fall 1999): 332–372.

Vittinghoff, Helmolt. "Recent Bibliography in Classical Chinese Philosophy," *Journal of Chinese Philosophy* 28: 1 & 2 (March/June 2001).

Waley, Arthur. *Three Ways of Thought in Ancient China.* London: Allen and Unwin, 1939.

Wing-Tsit Chan (1967)
Bibliography updated by Huichieh Loy (2005)

BUDDHISM

In India, Buddhism was a heterodox religious movement against the authority of the Vedas, the Bible of orthodox Hinduism. Gautama Buddha (c. 563–c. 483 BCE) dismissed the extreme ascetic way of life often adopted by Indian religious believers and taught the middle way. While Hindu philosophers asserted the existence of atman (I, self, ego, or soul) as the innermost essence of a human being and ontologically identified this essence with Brahma, the absolute reality of the universe, the Buddha repudiated the ideas of atman and Brahma, and proclaimed that everything is causally conditioned and nothing is absolute, permanent, and eternal.

All Buddhists have accepted the Buddha's teaching of anatman (nonself), but have apprehended his philosophical message differently. For the early, conservative Hinayana Buddhists, the Buddha's denial of ātman implies and even entails the existence of dharmas (divine laws), changing realities of the universe, and impermanent constituents of human beings. But later, progressive Mahayana Buddhists contended that the concept of dharma is as unintelligible as that of ātman. Both monistic absolutism and pluralistic realism are extreme views and should be eradicated. The true teaching of the Buddha is that all things are empty (*sunya*).

Both conservative and progressive Buddhist teachings had been introduced to China by the first century

CE. The Chinese preferred Mahayana and revered Nāgārjuna (c. 163–263) as the father of Mahayana Buddhism. The first Mahayana school founded by Nāgārjuna in India was named Mādhyamika, a name derived from the Sanskrit noun *madhyamā*, meaning middle or neutral. The Mahayana philosophy of emptiness as the middle way had laid a fine foundation for the development of Buddhism in China. The creation of new Chinese Buddhist schools—such as Tiantai, Huayan, Chan, and Pure Land—was directly or indirectly related to Nāgārjuna's philosophy.

THE SANLUN PHILOSOPHY OF EMPTINESS

In China, Indian Mādhyamika Buddhism is called the Sanlun (three-treatises) School. Nāgārjuna's *Madhyamakārikā* (Middle way treatise), *Dvādaśanikāyaśāstra* (*Twelve Gate Treatise*), and *Śataśāstra* (Hundred verse treatise), with the main verses by Āryadeva (third cent.), are devoted to the philosophy of emptiness and have been emphasized by Chinese Sanlun Buddhists. For Chinese Sanlun Buddhists, the notions of anatman, the middle way, and emptiness are synonymous in the Buddha's philosophy. Thus, the Sanlun school is also known as the middle-way school (*Zhongdao Zong*) and the emptiness school (*Kong Zong*).

More than any other Chinese philosophers, the Sanlun masters had a great interest in logical analysis and logical argument. They analyzed the dynamic and static worldviews, and they critically examined the nature and function of language and basic linguistic units such as subject, predicate, and predication. They questioned the essence and use of truth, knowledge, and logic, and they investigated various logical concepts and constructs such as right and wrong, negation and affirmation, and the meaning of thesis, antithesis, and synthesis in rational reasoning and conceptual disputes.

Usually people accept motion or change as an undeniable fact of experience. Even the Buddha, as well as the *Yijing* (Book of changes), seems to teach that all things are in a constant state of flux. Laozi's *Daodejing* (Way and power classic) also proclaims that "reversing" is the Dao (Way) of heaven. But under the influence of Indian Mahayana philosophy, Sengzhao (373–414), a brilliant Sanlun philosopher, wrote the famous essay *Wu buqian lun* (Things do not shift), arguing that motion is empty. He analyzed motion and pointed out that so-called motion consists of a part that has already passed (*yiqu*), a part that has yet to pass (*weiqu*), and a part that is passing (*qushi*). Change cannot be found in the part already passed, since it is already gone. Nor can it be found in the part yet to pass, since it is not yet. Nor can it be apprehended in the part that is passing, since passing makes sense if and only if there is an act of passing. But in examining whether there is an act of passing, we cannot use the act of passing to establish an act of passing without begging the question. So motion is impossible.

Zeno, a Greek philosopher, was well known for his argument that motion is impossible. Unlike Zeno, the Chinese Sanlun denial of motion does not entail the affirmation of rest. For Sanlun Mādhyamika, the concept of rest cannot be established either. Rest is the cessation of motion. If it is real, it must happen at some place and time. Does rest occur where something has already past, or where something has yet to pass, or where something is passing? None of these can be established. Therefore there can be no rest, or cessation of motion. For Sanlun masters, motion and rest are both empty, devoid of definite nature or essence, and hence not real. So one cannot maintain that reality is either permanent or impermanent. Therefore, any substantive or dynamic metaphysics must be repudiated.

According to Chinese Sanlun Mādhyamikas, philosophers appear to be very intelligent, but actually have often been fooled by language. Both Hindu and traditional Buddhist metaphysicians have failed to see the emptiness of words and names. Laozi understood the inadequacy of human language, as can be seen in the opening to his *Daodejing*, where he wrote, "The way that can be stated is not the real Way; Names that can be named are not real names." However, Laozi and later Daoists did not logically analyze language and did not present discursive arguments to substantiate their philosophy. Following Nāgārjuna's philosophy, Chinese Sanlun masters did logically analyze language, arguing that language is a conceptual game (*xilun*).

Sanlun masters critically examined the nature and the structure of conceptual and verbal statements, and argued that the relationship between two basic linguistic units, the subject (*kexiang*) and the predicate (*xiang*), cannot be rationally well formed, and that predication in our ordinary use of language is really not intelligible. They studied the precise relationship between the subject and the predicate, examining whether they are identical or different from one another. On the one hand, if two are identical, they are one, and it makes no sense to call one a subject and the other a predicate. Logically, the sentence is then a tautology and does not say anything about the world. Hence, in this case, predication is doing no real work. On the other hand, if the subject and the predicate

differ, predication is again unintelligible, since being (the similarity of subject and predicate) and not being (the difference between subject and predicate) cannot be at the same place at the same time. Hence it is absurd to unite what is different to form one sentence describing the same thing. Since every logical or conceptual statement consists of a subject and predicate, reality cannot be intelligently described. Therefore, so-called logic is in essence illogical.

In the view of the Sanlun masters, language and logic are empty. They are conventional and do not have a priori or absolute validity. Words have no definite meaning in themselves. The meaning of a term is not the object for which it stands, but depends on conditions and circumstances. If conditions change, the meaning of the word changes and might even be lost.

For the Sanlun masters, conceptualization, like a fish trap, has no intrinsic value and reality by itself, though it does have a practical use and can be employed to attract unenlightened persons to Buddhism. Yet the true message of the Buddha's teachings can be properly apprehended only if people comprehend the emptiness of words and discard conceptualization. Jizang (549–623), the most eminent Sanlun master, stated, "It is not that language is given in order to have Dharma [the Buddha's truth or teachings], but rather that Dharma is presented in order to eliminate language (Jizang 1854, p. 94c; Cheng 1984, p. 119).

According to Jizang, without practical benefits, truth and logic would lose their meaning. In ordinary life, humans have all sorts of emotional and intellectual attachments; they are attached to some view and stick to some law or principle. To free them from attachment, the Buddha preached a certain truth and followed a certain logic. To avoid the substantive or static view of the universe, he taught that everything is in flux, and to repudiate the dynamic view of the universe, he claimed that existence is real. Actually, terms such as "being" and "nonbeing," "permanent" and "impermanent," "to be" and "not to be," "real" and "unreal" are all empty. The Buddha's message can be regarded as the truth insofar as it helps dispel ignorance and illusion.

Ultimately, all conceptualizations should be discarded, and one should be silent. Such silence is not a form of absolutism or nihilism, but the manifestation of *prajñā* (wisdom). For ordinary people, to know is to know something; epistemology assumes objects to be known, acts of knowing, and a knower. In the ordinary way of thinking, an assertion of knowledge implies an ontological commitment. *Prajñā* is not to know something, but rather to apprehend that reality is empty, and so to be freed from attachments. In his essay *Boruo wuzhi* (*Prajñā* as nonknowing), Sengzhao (384–414) stated, "Real *prajñā* is as pure as empty space, without knowing, without seeing, without acting, and without objects. Thus knowledge is in itself without knowing, and does not depend on anything in order to be without knowing (Sengzhao 1858, p. 153; Cheng 1984a, p. 105).

To apprehend the empty logic of *prajñā*, one should understand, according to Jizang's *Sanlun xuanyi* (Profound meaning of the three treatises), that the refutation of erroneous views is the illumination of the right view (*poxie xianzheng*). In ordinary or even Aristotelian logic, negation and affirmation differ. Negation is usually asserted with the aim of affirming, establishing a thesis: Not *P* implies something other than *P*; the denial of a thesis entails the affirmation of an antithesis. For Chinese Sanlun masters, enlightened persons are empty-minded, free from affirmation and negation. Negation is used merely to repudiate erroneous views or to affirm negation itself. Not *P* means only the absence of *P*. *Prajñā* is the absence of any view, and is not a view in itself. The refutation of erroneous views and the affirmation of right views are not separate acts but the same. If a right view is held in place of an erroneous view, it becomes a new erroneous view and requires refutation. For Jizang, "Originally there was nothing to affirm and now there is nothing to negate (Jizang 1852, p. 6; Cheng 1984a, p. 47). An attachment to some view is a sickness (*bing*), and the logic of emptiness is the medicine (*yao*) to cure this intellectual sickness.

The Buddhist doctrine of emptiness, according to Chinese Sanlun masters, is not a metaphysical view. Rather, it is the doctrine that one should repudiate all metaphysical views, and to do so requires not the presentation of another metaphysical view, but simply the abolition of all metaphysics.

Emptiness (*śūnyatā*) is essentially a soteriological device. It is merely an instrument for eliminating extreme views. If there is no extreme to be removed, there need be no affirmation or negation. The so-called right view is really just as empty as the wrong view, and it is cited as right "only when there is neither affirmation nor negation." If possible, one should not use such terms as "right" and "wrong." For Jizang, "we are forced to use the word 'right' in order to put an end to wrong. Once wrong has been ended, then right no longer remains. Then the mind is attached to nothing." Even emptiness is empty. Jizang contended, "If one still clings to emptiness, then there is

no medicine that can eliminate the disease (Jizang 1852, p. 7).

The Sanlun philosophy was brilliant and authentically derived from Indian Mādhyamika thought. But the philosophy was too abstract and too Indian for the Chinese. Consequently, the Sanlun School declined in China after the death of Jizang in 623. However, its teachings inspired various Chinese Buddhists to develop new Buddhist movements in China.

THE ROUND TEACHING OF TIANTAI

Tiantai Buddhism, a sect of Mahayana Buddhism, had no Indian counterpart and was founded in China in the sixth century. It was initiated by Huiwen (550–577) and was well established by Zhiyi, also known as Zhikai (538–597), the greatest Tiantai master. Zhiyi lived and taught in the area of Mt. Tiantai in Zhejiang province, and hence the school came to be called the Tiantai School (Tendai in Japanese). Tiantai masters examined the Buddhist scriptures and held that the text *Saddharma-puṇḍarīka* (Lotus of the wonderful law) contains the best and most perfect doctrinal teaching of the Buddha, and consequently this school is also known as the Lotus (*huafang*) School.

Since ancient times Chinese have tended to think holistically or inclusively. Confucians and Daoists tended to observe things as they are and, with increasing ontological penetration, to see differences. The wonder of the universe, for Confucianism and Daoism, is a harmony among diversities and even opposites. According to the Yin-Yang School, the *Yijing* (Book of changes), and the *Daodejing* (Way and power classic), the universe is a united whole. It is composed of pairs of opposites: yin and yang, positive and negative, male and female, right and wrong. The interaction of yin and yang produces all things and all kinds of movement. Following this vein of thinking, Tiantai masters adopted the *yuanjiao* (round approach, doctrine, or teaching) and developed a philosophy of all in one and one in all.

Tiantai Buddhism disliked the analytic approach. For Tiantai masters, the analytic approach is a deductive and exclusive way of thinking that may reduce a complex world to one single reality, as seen in Hinduism, or a few simple fixed entities, as seen in Theravada Buddhism. Such thinking is one-sided and extreme, and hence should be eradicated. To avoid extremes, Tiantai Buddhists maintained that the Buddha's dharma is the direct observation, and pure and total description, of what is immediately given. Buddhism, for Tiantai masters, seeks to describe or to see things as they present themselves. "What the Buddha has accomplished is the teaching foremost, rare and inconceivable. Only the Buddhas can realize the true nature of all things; that is to say, all things are thus-formed, thus-natured, thus-substantiated, thus-caused, thus-forced, thus-activated, thus-circumstanced, thus-effected, thus-enumerated and thus-beginning-ending-completing (*Saddharma-pundarika* [The wonderful law of lotus], chapter 2).

Tiantai Buddhists held that Hindu and other Buddhist philosophers had distorted the original or true state of the things and polluted our comprehension of the universe. Tiantai philosophy sought to return to things themselves, that is, to penetrate to original, pure phenomena as they present themselves before any conceptualization or analytic judgment. "Thus-formed, thus-natured," in Tiantai teaching, indicates things as they present themselves. The Buddha's dharma, for Tiantai masters, seeks to penetrate to the fundamental or original data, to return to reality (*rushi*), as they appear to us in immediate experience. One should return to things themselves by means of direct awareness. For Tiantai masters, whoever sees things in this way sees what they called the original or true state of things (*zhufa shixiang*).

According to Tiantai masters, secular and even Buddhist philosophers have often ignored the richness of the universe and chopped complex, concrete, living facts into one absolute reality or a few simple elements. Tiantai Buddhists dismissed such philosophies as discriminative doctrines (*biejiao*). They did not divide the harmonious world into noumenon (*li*) and phenomena (*shi*). Nor did they reduce one concrete thing to another or give up any assertion; instead, they attempted to describe each fact in its fullness. They called their attitude and approach to the world the *yuanjiao* (round teaching, doctrine, or approach).

According to the round approach, all things, absolute and relative, are a united whole; noumenon is phenomena, and phenomena are noumenon. The relationship between the one and the many is like that of the ocean and waves. One ocean cannot be an ocean without many waves, and many waves cannot occur outside the one ocean. Thus, all is one, and one is all.

The Tiantai round approach is also used to apprehend Buddhist truth. Nāgārjuna is said to have taught, "Emptiness is called the middle way. For it is a provisionary name for causality (Nāgārjuna, "Zhong Lun" [The middle treatise], 18). For Huiwen and his followers, Nāgārjuna's statement taught that causality (*yinyuan*, dependent coarising) indicates lack of permanence and hence emptiness (*kong*), and thus it can serve as a substi-

tute name (*jiaming*) for the middle way (*zhongdao*). This awakened Huiwen to perceive the triple truth of emptiness, of temporariness, and of the mean. For Tiantai Buddhists, all things are empty because they are causally conditioned and hence are devoid of self-nature, but they do have temporary existence. Things by nature are empty and temporary; this principle constitutes the mean. These three—emptiness, temporariness, and the mean—penetrate one another and are found perfectly harmonized and united. A thing is empty but exists temporarily. It is temporary because it is empty. The fact that everything is empty and at the same time temporary constitutes the middle truth. One should consider the three truths not as separate but as a perfectly harmonious threefold truth.

In reality, the three truths, according to Tiantai Buddhism, are three in one and one in three. The principle is one, but its explanation is threefold, and each of the three truths has the value of all. From the perspective of emptiness, we may deny the existence of the temporary and the middle, for we consider emptiness as transcending all. The three principles would be empty. The same is the case from the perspective of temporariness or the mean. So when one principle is empty, all will be empty; when one is temporary, all will be temporary; when one is middle, all will be middle. These three principles are otherwise called identical emptiness, identical temporariness, and identical mean, and also the absolute threefold truth.

HUAYAN BUDDHISM AND THE MYRIAD MANIFESTATIONS

Huayan Buddhism, another sect of Mahayana Buddhism, was founded in China in the seventh century. It is named after the title of its chief scripture *Huayan jing* (*Avatamsaka sutra*, Flower-wreath sutra). According to this school, the Buddhist dharma is like the seed of a fine plant. It was planted by the Buddha in India; it grew and produced branches and leaves; eventually it blossomed, bearing beautiful flowers. Early Buddhism, various Hinayana and Mahayana schools, are the branches and leaves of the dharma. Huayan Buddhism is the flower of the dharma, the highest and the most splendid outcome of the Buddha's dharma.

The Huayan School was initiated by Dushun, also known as Fashun (557–640), but Fazang (643–712) is usually considered the real founder of the school because he was responsible for the final systematization of its teachings. Like Tiantai Buddhists, Huayan Buddhists developed a philosophy of one in all and all in one, and they also called their way of conceptualizing things *yuanjiao* (round approach, teaching, or doctrine). They wanted to observe and describe all things, phenomenal and noumenal, as purely and as fully as possible. They first rejected ordinary empiricism, which cuts up things into simple sense data, and they questioned Indian scholastic Buddhism, which reduced complex phenomena to simple dharmas. For Huayan masters, genuine phenomena are not the same as sensory phenomena. Alleged empirical facts or sensory appearances are really constituted phenomena and do not represent the true state of things.

The denial that sensory appearances represent the true state of the things does not, however, imply that Huayan masters denied sensory appearance in the world. In his famous *Jin shizi zhang* (Essay on the gold lion), Fazang used a gold lion to illustrate the case. We do perceive sensory phenomena such as a gold lion, but the appearance of the gold lion is not of a real lion. The proper understanding of things is that the true essence of such things is something other than physical form.

Sensory phenomena are empty; they do appear to exist, but their state of existence is not genuine. In the strict sense, sense experience is the manifestation of illusion (*huan*). To understand genuine being or genuine phenomena, one must contemplate things without qualities (*wuxiang*) by suspending one's natural belief in the existence of sense qualities or sense data. For Fazang, "To contemplate the qualityless is [to contemplate] the fact that the qualities of the tiniest part of matter arise out of the evolution of mind … , lacking any inherent nature of their own. This fact is called that of the qualityless" (Fazang, *Huayanjing yihaibaimen* [The hundred gates on the meaning of the flower splendor scripture] 1875, p. 627).

The contemplation of things without qualities, according to Huayan Buddhism, leads one to apprehend *li* (principle, noumenon) and to know the essence of the world. Such contemplation is a kind of empty-minded, disinterested observation of things, both objective and subjective, in their fullest breadth and depth. Thus seeing things as they are and as they are not is a round approach. For Huayan masters, it would empty or open up one's mind to see that being and nonbeing produce each other, to see that qualities and the qualityless complement each other, and thus to be aware of the essential relationship between phenomena and noumenon. Fazang wrote, "Noumenon does not interfere with phenomenon, what is pure is ever mixed. [Likewise] phenomena ever comprise noumenon in its totality, for what is mixed is ever so.… There is no barrier between what is pure and what is mixed" (Fazang, *Huayanjing yihaibaimen* [The hun-

dred gates on the meaning of the flower splendor scripture] 1875, p. 627).

Every event or fact is rich and complex. In describing the complex world, Huayan Buddhists claimed that a tiny particular thing involves and embraces all things in totality. Fazang wrote, "All things of the senses are revealed in their true essence and become merged into one great mass. Great functions arise, every one of which represents the Absolute. The myriad manifestations, despite their variety, harmonize and are not disparate. The all is the one, for all things equally have the nature of non-being. The one is the all, for cause and effect follow in an unbroken sequence. In their power and function, each implies the other and freely rolls up or spreads out. This is called the perfect teaching of the One Vehicle [the highest Buddhist truth]" (Fazang, *Jinshizi Zhang* [Essay on the gold lion], chapter 7).

In Huayan Buddhism, the universe is composed of an infinite number of possible differentiated worlds (*dharmadhātu*). As a whole, the universe is to be regarded as fourfold: the world of phenomena (*shifajie*, the realm of facts), the world of noumenon (*lifajie*, the realm of principle), the world of phenomena and noumenon united (*shiliwuaifajie*), and the world of phenomena united or interwoven with other phenomena (*shishiwuaifajie*). For Huayan masters, the Tiantai round approach is not inclusive or comprehensive enough. It merely touches on the first three realms of the universe but fails to see the world of phenomena united with other phenomena. According to the Huayan School, Huayan Buddhism better and more fully investigates and describes things themselves than other teachings. From its preeminent doctrine of *yuanjiao*, one can see that all things form a harmonious whole by mutually penetrating (*xiangru*) and mutually identifying (*xiangji*), and that phenomena are "the fact and the world of fact perfectly harmonized" (*shishiwuaifajie*) (Cheng 1984b, p. 222).

The distinct feature of Tiantai and Huayan Buddhism is their propagation of *yuanjiao* (round teaching, doctrine, or approach). In many ways, the Chinese round teaching in both Tiantai and Huayan philosophies is similar to Western phenomenology. Phenomenology can be seen as a purely descriptive study of any subject matter in which phenomena are described by means of direct awareness (*Anschauung*). In phenomenology, phenomena are not identified with sense experience or sense data, and the truth and falsity of phenomenological statements do not depend on sensory observation. For phenomenologists, sensory observation is instituted and categorized under certain general concepts, and hence in the strict sense, sensory experience is already constituted or polluted. The ideal of phenomenology is to return "to the things themselves" (*zu den Sachen selbst*). Actually, this is also the ideal of the round teaching that Tiantai and Huayan masters had tried to practice.

The object of phenomenological research includes whatever can conceivably be experienced, even what occurs in wild dreams. Phenomenologists do not neglect any aspect of our experience and seek to describe all things in their full possible concreteness. Tiantai and Huayan Buddhists had a similar objective in their round approach. Like phenomenologists, they aimed to investigate things, both subject and object, in their fullest breadth and depth. This is why Huayan masters taught the fourfold *dharmadhātu* as a way of exploring the infinite number of possible differentiated worlds, and claimed that their philosophy was "more round," "more complete," better, and higher than Tiantai and other Buddhist teachings. This is also why Huayan Buddhism is said to be the most splendid flower of the Buddha's dharma.

The phenomenological approach has negative and positive aspects, involving, as it does, turning away from something and turning toward something else. Negatively, it avoids preconceptions and brackets constituted phenomena. Positively, it turns to the things themselves and describes them as purely and as fully as possible. The negative aspect has a positive function: to facilitate genuine intuition of the given. In a similar way, the Buddhist round approach has a double character: *zhi* (cessation, stoppage, or stillness) and *guan* (observation, contemplation, awareness, or examination). *Zhi* is like Husserl's *epochē*, the suspension of all natural belief in the objects of experience. This is not to deny the world, but to become a disinterested spectator who can rediscover what has previously been lost. By means of *guan*, one can penetrate to the essence of things and obtain the unattached insight of true reality.

The phenomenological method is said to involve a change of attitude. One must look at the world with new eyes. The result is said to be a change in one's experience. The method of *zhiguan* in the Buddhist round approach also involves a change of attitude. One transforms from an attached way of life to an enlightened one, experiencing a sense of peace and transcendence: "How calm, still and pure! How deep, stable, and quiet! How pure and clear the inner silence! It functions without the character of functioning, and acts without the character of acting" (Fazang, *Dasheng Zhiguan Famen* [The Mahayana method of zhiquan], chapter 4.) Through this transformation, the true state of all things is apprehended, and

the universe is seen as the manifestation of an absolute mind, known as *zhenru* (true reality) or *rulaizang* (*tathāgatagarbha*).

It is interesting to see that the final outcome of the method of *zhiguan* is similar to that of Husserl's transcendental deduction, namely, the discovery of transcendental consciousness or mind. This is not subjective idealism, because subject and object, as well as the absolute and the relative, are seen to be interdependent, mutually penetrating, and even mutually identifying.

CHAN (ZEN)

Unlike Tiantai and Huayan Buddhism, the Chan School (Zen in Japanese), founded in China in the sixth century, does not aim to establish a round doctrine or to fully describe the universe. Chan Buddhism was claimed not to be a doctrine at all but a way of avoiding systematic views. Chan stories repeatedly teach that Chan Buddhism is not a body of fixed truths; instead, it is the abandonment of all views.

Although Chan masters did not develop theories, Chan Buddhism has some philosophical foundations found in Western philosophy: critical inquiry, autonomy, intellectual freedom, and creativity. Socrates is well known for saying that the unexamined life is not worth living. For him and many others, the philosophical enterprise consisted of inquiry rather than an accumulation of final truths. In the West, philosophy has often been regarded as the highest form of inquiry because, unlike other sciences, it alone does not involve presuppositions. True philosophers take nothing for granted. Similarly, Chan masters took nothing for granted.

Chan Buddhism has been critical of Buddhism viewed as a religion. Often a religion presupposes the authority or divinity of its founder and the infallibility of his words, but Chan Buddhism invoked no such presuppositions. Chan masters often rejected any special status for Gautama Buddha and repudiated the certainty of Buddhist scriptures. When the Buddha was born, he is alleged to have proclaimed, "Above the earth and below the heavens, I alone am the Honored One!" Chan master Yunmen (864–949) commented on this saying, "If I had been with him at the moment of his uttering this, I would surely have struck him dead with one blow and thrown the corpse into the maw of a hungry dog" (Suzuki 1964, p. 40). Chan masters would not subscribe to the views of a religious leader. One must enlighten oneself. Enlightenment (*wu*) must occur within and be done personally. In fact, according to Chan masters, any person who obtains enlightenment is a Buddha.

For Socrates, the autonomous activity of philosophy was integral to being genuinely human. For Chan Buddhism, to live genuinely is to live the life of enlightenment, and to live the life of enlightenment is to live autonomously. Simply following the Buddha faithfully and practicing the dharma diligently does not engender an enlightened outlook. Rather, one must be autonomous (*zizhu*). In the Chan lifestyle, a true Buddhist conducts his life freely and leisurely (*ziyou zizai*).

The main message of Chan Buddhism, believed to have been composed by Bodhidharma (470–543), is succinctly stated thus:

> A special transmission outside scriptures;
>
> No dependence upon words and letters;
>
> Direct pointing at human mind;
>
> Seeing into one's own nature to attain Buddhahood (Dumoulin 1988, p. 85).

Chan masters repudiate any blind acceptance of scriptures, for "the entire scriptures from beginning to end are nothing but deceitful words" (Chung-yuan 1971, p.143). The so-called holy scriptures of Buddhism have often been set aside, thrown away, and even burned by Chan masters.

The radical approach of Chan Buddhism created a refreshing Buddhist epistemology that emphasized opening up the mind to the serious issue of what truth is. For some Buddhists, truth is objective and can be spoken and written about. In this view, the Buddha and the patriarchs transmitted truth, and the scriptures contain their messages, often identified with the dharma. But for Chan Buddhism, truth is not something objective, nor can it be spoken and written about. The Buddhist dharma is not conveyed by ink marks on the pages of scripture. Huineng (638–713), the sixth patriarch, was said to be illiterate, and yet was a Chan master. When Fada, a devout monk, studied the *Lotus Sutra* three thousand times and still could not understand it, he came to ask Huineng for instruction. The master said, "The Dharma is quite clear; it is only your mind that is not clear. Whether Sutra-reciting can enlighten you or not all depends on yourself. … If the mind is deluded, the *Lotus* [*Sutra*] turns you around, if the mind is enlightened, you turn around the *Lotus* [*Sutra*]" (Huineng 1952, p. 24).

Truth, for Chan Buddhism, is something living and personal. To equate truth with a proposition is to objectify and conceptualize it, to make it static and dead. For Chan masters, "The real truth is nothing else but one's own mind. Thus … the real teaching must be transmitted

directly from one mind to another" (Chung-yuan 1971, p. 86). Genuine spiritual education occurs in personal communication between Chan master and disciple. This mind-to-mind transmission resembles what Martin Buber, the great twentieth-century thinker, described as an I-thou relationship, rather than an I-it relationship. To see the truth as an object and to conceptualize it is to shift from a personal I-thou point of view to an impersonal I-it understanding.

True meditation, a central Chan practice, does not refer to sitting in a certain posture with legs crossed, but to "the brightening up of the mind-works" (Suzuki 1956, p. 85). Mazu (709–788) used to sit diligently and frequently in meditation. Master Huairang (677–744) asked him, "Virtuous one, why are you sitting in meditation?" Mazu replied, "I want to become a Buddha." Thereupon the master picked up a tile and rubbed it repeatedly in front of the hermitage. Mazu asked, "What is the master doing?" Huairang answered, "I am polishing the tile to make a mirror." Mazu exclaimed, "How can you make a mirror by polishing a tile?" The master responded, "How can you make a Buddha by practicing sitting meditation?" (*Jingde Chuandeng Lu*, Vol. 5). The monk was said to be enlightened immediately, and later became a great Chan master.

The personal experience of the dharma, according to Chan masters, is not remote, abstract, or transcendent. It occurs in one's present daily life. Zhaozhou (778–897) asked Master Nanquan (748–834), "What is the Dao [the Way]?" The master replied, "Everyday-mindedness is Dao." In another instance, after attaining great enlightenment under Mazu, Pangyun stated, "I am an ordinary man who fulfills his daily tasks. How plain are the Buddhist teachings!" According to the Chan School, "In the carrying of water and the chopping of wood—therein lies Dao" (Chung-yuan 1971, p. 145).

An enlightened person does not live outside samsara, or this world of rebirth, and he should not ignore karma, or cause and effect. He should treasure this life and value the virtues of labor in daily affairs. Baizhang (720–814), who founded the Chan monastic order, was said to live by the principle "A day without work—a day without eating" (Dumoulin 1988, p. 103). When he was old and his disciples hid his tools, he refused to eat until he could work again. Chan practitioners do not adhere to rigid moral precepts, but practice a work ethic in daily life.

Chan philosophy has similarities with contemporary ordinary-language philosophy in that both favor the ordinary use of language. For Chan Buddhism, however, any concrete fact or lived experience is rich and complex. Things may appear to be simple and ordinary, yet are really quite complicated and extraordinary. Chan Buddhists have sometimes used metaphysical statements to convey their understanding, and have also expressed themselves through strange words and strange acts (*qiyan qixing*). Consequently, the Chan literature abounds with irrational statements and absurd actions.

The use of strange words and strange acts in Chan Buddhism actually accords with the Mādhyamika practice of revealing the truth of emptiness and the middle way. According to Nāgārjuna, the Buddha's dharma was given, and hence should be understood, by means of twofold truth, a convenient term for the perspectives of conventional and ultimate truth. The former sees things from a viewpoint deluded by attachment, while the latter sees things without attachment.

Following Nāgārjuna's philosophy, Chan masters often expressed themselves through twofold truth, and hence their teachings and practices may be apprehended from two standpoints. Ordinary sentient beings do not see the emptiness of all things. So, to comply with conventional usage, Chan masters may say, "I see" or "you should see" the objects of right knowledge. But from a higher, unattached standpoint, all things are empty, and so the same master may also state, "I do not see" or "one should not see" any right object; on the contrary, one should see the emptiness of all things. For instance, once Shenhui asked Huineng, "Do you see or not?" The master replied, "I both see and also do not see." The puzzled disciple asked, "How can you see and also not see?" The master instructed, "If your mind is attached, you do not see; if your mind is without attachment, you see." The seemingly inconsistent expressions of Chan Buddhism were delivered with twofold truth in mind. Understood in this light, they are not as illogical as they might appear.

Although Chan Buddhism is a Mahayana practice, in many ways it strongly reflects Chinese thinking and feeling. Such Chan ideas as *xing* (nature, essence, own nature), *xin* (mind, human mind), *foxing* (Buddha nature), *foxin* (Buddha mind), and the key message that everyone has a Buddha nature are more like Chinese Confucian thought than Indian Mahayana Buddhism.

The central message that Bodhidharma brought from India to China in the sixth century also differs from the Indian Mahayana philosophy of emptiness. For Nāgārjuna, all things are causally conditioned and hence empty of any nature (*xing*) of their own. So, whereas the Chan masters instruct one to see the nature of things (*jianxing*), the Indian Mahayana scriptures teach the believer to reject this idea. Chan practice has to be

regarded as a special transmission outside the scriptures. For Chan masters, one cannot and should not follow the ancient Indian scriptures literally; otherwise one will never be enlightened. Therefore, Chan Buddhism advises, "No dependence on words and letters."

The notion of the nature of things (*xing*) was important in the minds of Chinese thinkers long before Buddhism was introduced to China. Both orthodox and less orthodox Confucianists accepted the view that things had natures. Confucius and Mencius are well known as saying that human nature is good. In their teachings, the mind (*xin*) is the nature of a human. This human nature or mind is more important for its axiological value rather than for its ontological substance, in contrast with such notions as Hindu Brahma and atman, Theravada Buddhist *svabhava* (inborn nature), and Greek substratum. This notion of value makes humans valuable and endows them with a spiritual quality. Without this nature, a person would be merely a beast. With this nature, a person can become a sage. According to Mencius, Confucianism teaches that one should exhaust one's mind and know one's nature (*jinxin zhixing*). One who practices this will be a gentleman and a sage.

Chan masters skillfully assimilated the Confucian sense of nature, mind, and sagehood into Buddhist thought. The result of this skillful measure (*fangbian*, *upaya* in Sanskrit) was the doctrine of "direct pointing at the human mind; seeing into nature to attain Buddhahood" (Dumoulin 1988, p. 85). Huineng opened his famous *Platform Sutra* with the same message: "Virtuous ones! The Bodhi-nature is originally pure. Making use of this mind alone, one can directly become a Buddha" (Huineng 1952, opening statement).

Inspired by Confucian thought, Chan masters transformed the traditional Buddhist doctrine of gradual enlightenment into the teaching and practice of abrupt or sudden enlightenment (*dunwu*). In Indian Buddhist teachings, not everyone has a Buddha nature and can become a Buddha. But according to Confucius and Mencius, all human beings are alike in nature and become different owing to different external environments. In the original state, humans have innocent, fine minds that cannot bear to see the suffering of others. But this mind was lost. The aim of education is to recover what has been lost. Can we find the original mind? Mencius's answer was positive and optimistic. He wrote that the original mind is "all already complete in oneself," and that the truth "is not far to seek, but right by oneself." Following this positive, optimistic philosophy, Chan masters proclaimed that everyone has a Buddha nature, is able to become a Buddha, and can suddenly attain enlightenment.

PURE LAND'S MESSAGE OF HOPE

While most Buddhists took a positive view of human nature, Pure Land Buddhism (Jingtu, Jōdo in Japanese), also founded in China in the sixth century, acknowledged human weakness and was pessimistic about individual efforts to achieve nirvana. Reading scriptures, sitting in meditation, keeping moral precepts, understanding the dharma, and training for enlightenment are all fine, but really too much and too extreme for most. Pure Land Buddhism is a protest against, as well as a step away from, intellectual, scriptural, and disciplinary forms of Buddhism. The main message of Pure Land Buddhism is that one cannot and need not attain nirvana by effort, but may obtain it with the help and compassion of Amitabha Buddha.

The Pure Land message, according to this school, was the Buddha's original teaching, which was rediscovered by Nāgārjuna, later revered as the first patriarch of the Pure Land School. According to Pure Land masters, Nāgārjuna taught, "Although there are innumerable ways in the teachings of the Buddha, they can be classified roughly: the difficult way and the easy way." The difficult way is to approach *Avaivartike* (a state of no return to the world of delusion) by diligently following the eightfold path and practicing the six virtues of perfection (*pāramitās*); the easy way teaches faith in Amitabha Buddha. The Mahayana doctrine of emptiness is the teaching of the easy way, for it teaches the emptiness of all our views and efforts.

Although Indian Mahayana teachings appear to differ from Pure Land Buddhism, Tanluan (476–524), the real founder of Pure Land Buddhism in China, is said to have been inspired by Nāgārjuna's philosophy of emptiness. He drew on Nāgārjuna's *Dasabhumi-vibhastra* (A commentary by Nagarjuna on the ten stages in bodhisattva wisdom) to advocate that, because humans have little spiritual capacity, they should not pursue the difficult path. Traditional religious life represents the difficult way, which, more properly speaking, according to Tanluan, is the teaching of enlightenment through one's own power (*zili*). But the Mahayana way teaches salvation by relying on an external power (*tali*). By relying on the Buddha's help and compassion, one can empty oneself and be awakened and saved by the Buddha's help and compassion.

Pure Land Buddhism shifted the focus of the Buddha's dharma from the discipline of observing moral pre-

cepts (*vinaya*) and an emphasis on wisdom (*prajñā*) to the spread of compassion (*karuṇā*). Pure Land Buddhism is a religion of repentance, mercy, forgiveness, and grace. One obtains salvation by faith and devotion rather than by work or learning. Life, according to the Buddha's first noble truth, involves suffering. Yet formal religion has not made our lives more comfortable; on the contrary, it has frustrated and confused the minds of many clerics and laypeople because few can sustain the rigors of mastering Buddhist doctrine, either by practicing monastic discipline or by studying scriptures. An enlightened Buddha would see this state of suffering, have compassion, and be willing to help humans rise from the ocean of sufferings.

Amitabha, a compassionate bodhisattva according to Pure Land Buddhism, saw the miserable condition of sentient beings and determined to extend his great mercy to them, making forty-eight vows to save them. Failing his vows, he would not become a Buddha. Thus, while people may not be smart enough to digest Buddhist scriptures and may not have time to sit in meditation, they may yet hope for salvation by calling on the name of Amitabha Buddha (*Amituofo* in Chinese). The recitation or invocation of Amitabha's name became the trademark of Pure Land Buddhism. This simple act was said to help people enter into the western paradise, or the Pure Land. In fact, it became the most common Buddhist practice in China, Korea, and Japan, and the most popular means for salvation by which millions have sought release from suffering. So Pure Land Buddhism transformed Buddhism into a popular religion by preaching the simple gospel of hope.

Pure Land masters equated chanting "Amituofo" with Buddhism. Daochuo (562–645), the second patriarch of Pure Land Buddhism, was said to repeat the name of Amitabha Buddha seventy thousand times per day. Chanting "Amituofo" was believed to enable a person to be reborn in the Pure Land. Here, religious language does not describe the universe nor analyze truth. Rather, it is a calling for help, a therapy to relieve anxiety, frustration, despair, and other sufferings in life. The sound of "Amituofo" seems to have a power to comfort people and pacify the mind. The ultimate cause of the effectiveness of the invocation, according to Pure Land masters, is Amitabha Buddha himself, who aspired to save all beings. It is really through the power of Amitabha Buddha's vows that mortals, by reciting his name, can be released from the hell fires that a life of sin and evil bring on.

The power of chanting "Amituofo" is good news, because even persons who have committed the most egregious sins can be saved if they recite the name of Amitabha Buddha. According to Shandao (613–681), an eminent Pure Land master, Pure Land Buddhism not only offers salvation to known sinners, but also leads good people to repent and confess their sins. Those who sincerely acknowledge and believe that they are sinful, lowly persons continually involved in error and shut off from salvation are enlightened Buddhists. If one can repent of sin, no matter how small the sin, one will gain a deep sense of release from suffering and can aspire to birth in the Pure Land through Amitabha Buddha's vows. Confession, repentance, humility, and forgiveness, rather than punishment and condemnation, are the virtues promoted and practiced by the Pure Land community.

BUDDHISM IN CHINESE CULTURE

From the sixth century, Indian Buddhism became sinicized. Divergent Chinese Buddhist philosophies and practices were assimilated and fitted into the Chinese tradition, and exercised a lasting influence on almost every aspect of Chinese life. By the eighth century, Chinese Buddhism became firmly established and triumphantly spread throughout China. Chinese culture became an aggregation and synthesis of Confucianism, Daoism, and Buddhism. However, this syncretism did not go easily and smoothly. There were three major persecutions of Buddhists in Chinese history. The most devastating one occurred in 845. After this, most Buddhist schools declined in China. Then the Chan and Pure Land schools became predominant over other Buddhist schools and practices.

From the Song dynasty (960–1279) onward, chanting "Amituofo" has been the major religious practice among devout Buddhists. Chan philosophy was attractive to and popular among Chinese intellectuals, and was a vital cultural force, especially in literature and the arts. In fact, it led Confucian scholars to reexamine classical Confucian philosophy and develop neo-Confucianism, even though neo-Confucian scholars frequently attacked Buddhism when defending their orthodox teachings. Like Chan Buddhists, neo-Confucian scholars cultivated the mind, and even used Buddhist terms, some equating *li* (principle, reason) with the Dao, and others with the mind. Like Tiantai and Huayan Buddhists, many Confucianists adopted the round approach to develop an all-in-one and one-in-all worldview. In many ways what was new in neo-Confucianism was quite Buddhist in spirit.

The influence of Buddhism can also be seen in twentieth-century new Confucianism, as in Feng Youlan's (1895–1988) famous book *Xin lixue* (A new study of principle). Like metaphysically minded Buddhists, Feng

investigated the principles in and behind things with the aim of reaching the highest sphere of life, namely "forming one body with all things." Xiong Shili (1885–1968), the founder of twentieth-century new Confucianism, was obviously a Buddhist Confucian. He promoted the Mahayana philosophy of consciousness only (*weishi*) and reinterpreted the Confucian metaphysics found in the *Yijing* (Book of changes) in the light of this doctrine. His eminent disciples, among them Tang Junyi (1909–1978) and Mou Zongsan (1909–1995), examined the round approach (*yuanjiao*), and they debated whether Tiantai or Huayan philosophy represented the highest teaching. Mou Zongsan found the Tiantai School to be the best. To develop his moral metaphysics, he adopted Tiantai philosophy, especially the idea that phenomena are noumenon and noumenon is phenomena. Tang Junyi, Fang Dongmei (1899–1977), and many other twentieth-century Confucian scholars have contended that Huayan philosophy, rather than Tiantai philosophy, represented a fuller development of Buddhist thought.

Fang Dongmei, just before his death, made the following statement:

> From emptiness I came.
>
> To emptiness I return.
>
> Emptying the emptiness without possessing any being
>
> It is in nowhere that my heart will dwell (Shen 2003).

Thus, the latest approaches to Confucianism have been profoundly influenced by Buddhist thought. One cannot properly understand Chinese philosophy or the history of Chinese thought without knowing Buddhist philosophy.

See also Chinese Philosophy: Overview; Chinese Philosophy: Contemporary.

Bibliography

Chen, Kenneth. *Buddhism in China*. Princeton, NJ: Princeton University Press, 1964.

Cheng, Hsueh-li. *Empty Logic: Mādhyamika Buddhism from Chinese Sources*. New York: Philosophical Library, 1984a.

Cheng, Hsueh-li. *Exploring Zen*. New York: Peter Lang, 1996.

Cheng, Hsueh-li. "Phenomenology and T'ien-t'ai and Hua-yen Buddhism." *Analecta Husserliana* XVII (1984b): 222.

Chung-yuan, Chang. *Original Teaching of Ch'an Buddhism*. New York: Vintage Books, 1971.

Cook, Francis H. *Hua-yen Buddhism: The Jewel Net of Indra*. University Park: Pennsylvania State University Press, 1977.

De Bary, William Theodore. *The Buddhist Tradition in India, China, and Japan*. New York: Random House, 1972.

Dumoulin, Heinrich. *Zen Buddhism: A History*, Vol. 1: *India and China*. New York: Macmillan, 1988.

Fazang. *Dasheng Zhiguan Famen* (The Mahayana method of zhiquan). Chapter 4.

Fazang. "Huayanjing yihaibaimen" (The hundred gates on the meaning of the flower spendor scripture). *Taisho* 1875, Chapter I.

Fazang. *Jinshizi Zhang* (Essay on the gold lion). Chapter 7.

Gimello, Robert M., and Peter N. Gregory, eds. *Studies in Ch'an and Hua-yen*. Honolulu: University of Hawaii Press, 1983.

Huineng. *The Sutra of Huineng*. Hong Kong: H.K. Buddhist Distributor Press, 1952.

"Jingde Chuandeng Lu" (Jingde record on the transmission of light). *Taisho* 2076, Vol. 5.

Jizang. "Er Di Yi" (The meaning of twofold truth). *Taisho* 1854, p. 94c.

Jizang. "Sanlun Xuanyi" (The profound meaning of three treatises). *Taisho* 1852, p.6.

Nagarjuna. *Zhong Lun* (The middle treatise), XXIV: 18.

Saddharma-pundarika (The wonderful law of lotus). Chapter 2.

Sengzhao. "Zhao Lun," Part III. *Taisho* 1858, p. 153.

Shen, Vincent. "Fang Dongmei." *Encyclopedia of Chinese Philosophy*. New York: Routledge, 2003.

Suzuki, D. T. *An Introduction to Zen Buddhism*. New York: Grove Press, 1964.

Suzuki, D. T. *Zen Buddhism*. Garden City, New York: Doubleday Anchor Books, 1956.

Swanson, Paul. *Foundations of T'ien-t'ai Philosophy*. Berkeley, CA: Asian Humanities, 1989.

Takakusu, Junjirō. *The Essentials of Buddhist Philosophy*. Honolulu: University of Hawaii Press, 1947.

Wright, Arthur F. *Buddhism in Chinese History*. Stanford, CA: Stanford University Press, 1959.

Hsueh-li Cheng (2005)

CONFUCIANISM

CONFUCIANISM AS AN ETHICAL TRADITION

After the Zhou people conquered the Shang people in the middle of the eleventh century BCE, the early Zhou kings ruled by letting feudal lords govern vassal states. As their powers grew, feudal lords fought one another and resisted the Zhou king until the state of Qin conquered all other states in 221 BCE. A number of ethical and political thinkers lived in the period from the sixth to third century BCE, proposing different ways of restoring order as well as ideal ways of life for human beings. Among them, several thinkers, including Confucius (sixth century BCE), Mencius (fourth century BCE) and Xunzi (third century BCE), as well as their followers, were regarded as belonging to the same movement of thought. This move-

ment of thought was referred to retrospectively in the Han dynasty (206 BCE to 220 CE) as *rujia*, or the school of *ru*. The English term *Confucianism* is now often used as a translation of *rujia* to refer to this school of thought.

Unlike what the term *Confucianim* suggests, the expression *rujia*, or "the school of *ru*," does not bear any special relation to the name of the individual known as Confucius. Instead, *ru* referred to a social group that already existed before the time of Confucius. The group consisted of professional ritualists who performed rituals in such ceremonial contexts as funeral rites, sacrifices to ancestors, and marriage ceremonies. In addition, these ritualists were often professional teachers, not just of rituals but also of other disciplines such as music. Certain individuals who were members of this group in virtue of being professional ritualists and teachers (including Confucius, Mencius, and Xunzi) came to develop concerns that were no longer restricted to rituals or to their own economic sustenance. Instead, they directed their attention to finding a remedy for the chaotic social and political situation of the times and to establishing the ideal way of life for human beings. They believed that the remedy lay with the maintenance and restoration of certain traditional norms and values, including but going beyond rituals, and proposed that, ideally, people should follow a way of life that embodies such norms and values. Unlike what the term *Confucianism* might suggest, these norms and values did not originate with Confucius but date back to a much earlier time.

Still, in referring to this movment of thought as *rujia* or "the school of *ru*," the Chinese did regard Confucius as the first and most important thinker of the movement. Both Mencius and Xunzi, the two other major Confucian thinkers from that period, also regarded themselves as defending Confucius's teachings, and their different developments of Confucius's teachings competed for influence in the Han dynasty. In the Tang dynasty (618–907), the Confucian thinker Han Yu (768–824) regarded Mencius as the true transmitter of Confucius's teachings, and this view was endorsed by Zhu Xi (1130–1200) of the Song dynasty (960–1279). Zhu Xi included the *Analects* (*Lunyu*) of Confucius and the *Mencius* (*Mengzi*), along with the *Great Learning* (*Daxue*) and *Centrality and Commonality* (*Zhongyong*), the latter two texts dating probably to early Han, among the Four Books. These texts eventually became the canons of the Confucian school, and Mencius came to be regarded as second only to Confucius in importance. Different kinds of Confucian teachings continued to evolve after Zhu Xi's times, represented by major figures such as Wang Yangming (1472–1529) of the Ming dynasty (1368–1644) and Dai Zhen (1724–1777) of the Qing dynasty (1644–1912).

Suppose we characterize ethics in terms of a concern with the question how one should live, where the scope of "one" is supposed to extend considerably beyond the person raising the question. Confucian thinkers do share a concern of this kind. Furthermore, they are reflective not just in having a conception of how one should live, but also in being concerned with the proper spirit behind the observance of rituals and other traditional norms, and with the grounds for observing these traditional norms and values. This warrants describing them as ethical thinkers. Also, although there are substantive differences in the views of different Confucian thinkers, these thinkers also share a broad similarity, both in defending certain traditional norms and values and in the use of certain common key terms in elaborating on their thinking. They share the same allegiance to Confucius's teachings and, after the time of Zhu Xi, also share a conception of certain canonical texts that define the Confucian school. These similarities warrant regarding them as belonging to the same tradition of thought and describing Confucianism as an ethical tradition. The rest of the article will elaborate on some of the main characteristics of this ethical tradition.

CONCEPTION OF THE SELF

To start with, let us consider how the Confucians view the self and the human constitution. They use the term *ti*, often translated as "body," to talk about a person's body, and they also have ways of referring to parts of the body, such as the four limbs and the senses. These parts of the body are not regarded as inert; not only do they have certain capacities, such as the eye's capacity of sight, but they also exhibit certain characteristic tendencies. For example, the four limbs are drawn toward rest, while the senses are drawn toward such ideal objects as beautiful colors or pleasurable objects of taste. Such tendencies are referred to as *yu*, a term often translated as "desires" and paired with an opposite term often translated as "aversion." These terms have, respectively, the connotations of being drawn toward and being repelled by certain things. The terms can be used not just for parts of the body but also for the person as a whole to describe how the person is drawn toward things like life and honor and repelled by things like death and disgrace.

That human beings have such tendencies as part of their basic constitution is regarded as a fact about them that is pervasive and difficult to alter. Facts of this kind are referred to as the *qing* of human beings, where *qing*

means "facts" and, in this context, the connotation of certain facts about human beings that reveal what they are genuinely like. Later, *qing* comes to refer to what we would describe as emotions, including such things as joy, sorrow, and anger, these also being regarded as parts of the basic constitution of a person.

There is another feature of the Chinese view of the person for which it is difficult to find a Western equivalent. The body of a person is supposed to be filled with *qi*, a kind of energy or force that flows freely in and gives life to the person. *Qi* is responsible for the operation of the senses; for example, it is supposed to make possible speech in the mouth and sight in the eyes. Conversely, it can be affected by what happens to the senses; for example, *qi* can grow when the mouth takes in tastes and the ear takes in sounds. Also, *qi* is linked to the emotions, and what we would describe as a person's physical and psychological well-being is regarded as dependent on a proper balance of *qi*. For example, both illness and such emotional responses as fear are explained in terms of the condition of *qi*.

Among the different parts of the person, special significance is attached to *xin*, the organ of the heart that is viewed as the site of what we would describe as cognitive and affective activities. *Xin*, a term often translated as "heart" or "mind," can have desires (*yu*) and emotions (*qing*) and can take pleasure in or feel displeasure at certain things. It can also deliberate about a situation, direct attention to and ponder about certain things, and keep certain things in mind. One capacity of the heart/mind (*xin*) that is particularly important for Confucian thinkers is its ability to set directions that guide one's life and shape one's person as a whole. Such directions of the heart/mind are referred to as *zhi*, a term sometimes translated as "will."

Zhi can refer to specific intentions such as the intention to stay in or leave a certain place, or to general goals in life such as the goal of learning to be a sage. It is something that can be set up, nourished, and attained; it can also be altered by oneself or swayed under others' influence, and lost through insufficient persistence or preoccupation with other things. Early texts sometimes compare setting one's *zhi* in certain directions to aiming at a target in archery, and *zhi* is sometimes used interchangeably with another character that means "recording something" or "bearing something in mind." Probably, *zhi* has to do with the heart/mind's focusing itself on and constantly bearing in mind certain courses of action or goals in life, in such a way that *zhi* will guide one's action or one's life unless it is changed by oneself or under others' influence or unless one is led to deviate from it by other distractions. *Zhi* (directions of the heart/mind) differs from *yu* (desires) in this respect: although *zhi* pertains to the heart/mind, *yu* can pertain to the heart/mind or to parts of the body such as the senses or the four limbs. Furthermore, whereas *zhi* involves focusing the heart/mind in a way that guides one's actions or one's life in general, *yu* involves tendencies that one may choose to resist rather than act on.

With this survey of the different aspects of the person as background, let us consider the notion of self as it applies to Confucian thought. Now, besides the use of first-person pronouns, the Chinese language has two characters with the meaning of "oneself." *Zi* is used in reflexive binomials referring to one's doing something connected with oneself, such as one's examining oneself or bringing disgrace upon oneself. *Ji* is used to talk about not just one's doing something connected with oneself but also others doing something connected with oneself (such as others appreciating oneself), oneself doing something connected with others (such as oneself causing harm to others), or one's desiring or having something (such as a certain character) in oneself. The two characters differ in that the former emphasizes one's relation to oneself, whereas the latter emphasizes oneself as contrasted with others. In addition, the character *shen*, which is used to refer sometimes to the body and sometimes to the person as a whole, can also be used to refer to oneself or to one's own person when prefixed with the appropriate possessive pronoun.

These linguistic observations show that the Chinese have a conception of the way one relates to oneself. Furthermore, in connection with Confucian thought, the characters just mentioned are often used to talk about one's examining oneself and cultivating oneself on the basis of such self-examination. This further observation shows that Confucian thinkers also work with a conception of one's being related to oneself in a self-reflective manner, with the capacity to reflect on, examine, and bring about changes in oneself. So they have a conception of the self in the sense of a conception of how one relates to oneself in this self-reflective manner.

Confucian thinkers ascribe the capacity of self-reflection just described to the heart/mind, to which they also ascribe a guiding role. They emphasize the importance of self-cultivation—that is, the process of constantly reflecting on and examining oneself, setting one's heart/mind in the proper direction, and bringing about ethical improvements in oneself under the guidance of the heart/mind. There has been extensive disagree-

ment within the Confucian tradition about how the heart/mind can set itself in the proper direction. For example, Mencius and Xunzi disagree about whether a certain ethical direction is already built into the heart/mind and whether one should derive the proper direction by reflecting on the heart/mind or by learning from the outside. Later, Zhu Xi and Wang Yangming disagree in the different emphases they place on learning and on attending to the heart/mind in the process of self-cultivation. Despite such disagreements, they all regard the heart/mind as that which guides the process of self-cultivation.

Furthermore, they also agree on another distinctive feature of the heart/mind—not only can it set directions that guide the person's life and shape the person as a whole, but it is also independent of external control in having the capacity to hold on to the directions it sets without being swayed by external forces. For example, both the *Analects* and the *Mencius* emphasize its guiding role, comparing the directions (*zhi*) of the heart/mind to the commander of an army. In addition, the *Analects* notes one point of dissimilarity—although an army can be deprived of its commander, even a common person cannot be deprived of the directions set by the heart/mind. Such directions can, of course, be influenced by outside factors, but the point is that the heart/mind has the capacity to resist such influences and, for the Confucian thinkers, one should ideally cultivate oneself to attain such a steadfastness of purpose after having set the heart/mind in the proper directions. This independence of the heart/mind from external control is also emphasized by Xunzi, who compares the heart/mind to the position of the ruler and the senses to the offices of government; like the ruler, the heart/mind issues order but does not take order from anything.

Not only is the heart/mind independent of external control, but it also has the capacity to constantly step back to reflect on and improve its own operations. Three early Confucian texts—the *Xunzi*, *Great Learning*, and *Centrality and Commonality*—emphasize the idea that the heart/mind should cautiously watch over its own activities to ensure that all of them, however minute or subtle, are completely oriented in an ethical direction. This idea is presented in terms of watching over *du*, where *du* refers to the minute and subtle workings of the heart/mind that are not yet manifested outwardly and to which one alone has access. The idea is taken up by later Confucian thinkers, who in addition emphasize the importance of watching out for and eliminating what they call "selfish desires," that is, the distortive influences in the heart/mind that might lead one to deviate from the ethical direction. This aspect of Confucian thought shows that the Confucians ascribe to the heart/mind a self-reflexiveness; for any of its own activities, however minute and subtle, it has the capacity to reflect on and reshape such activities to ensure their orientation in an ethical direction. This self-reflexiveness is related to the independence of the heart/mind from external control—even though its activities can be influenced by external circumstances, the heart/mind has the capacity to constantly step back and reshape its own activities under the conception of what is proper, which it forms on the basis of its own reflections.

Given their emphasis on the distinctive role of the heart/mind, did Confucian thinkers believe in some kind of mind-body distinction? In a sense, they do emphasize a distinction between the heart/mind and other aspects of the person. The heart/mind has the distinctive capacity to reflect on these other aspects and on its own activities, to form a conception of what is proper, and to regulate and shape other aspects of the person and its own activities under such a conception. On the other hand, the distinction that the Confucian thinkers emphasize pertains to the distinctive capacities and modes of operation of the heart/mind rather than to the heart/mind as a distinctive kind of entity that occupies a "mental" as opposed to a "physical" realm. The character *xin*, translated here as "heart/mind," refers to the organ of the heart that is a part of the body just as the senses are. And just as the heart/mind can operate in the manner described earlier, the senses also have their own modes of operation, such as distinguishing between and being drawn toward certain sensory objects. What distinguishes the heart/mind from other parts of the body is not that it pertains to a "mental" as opposed to a "physical" realm but that its modes of operation are different from, and enable it to perform a guiding function in relation to, other parts of the body.

Furthermore, there is also a sense in which Confucian thinkers deemphasize the distinction between the heart/mind and other aspects of the person. Earlier, we considered the Confucian emphasis on one's cautiously watching over the minute and subtle activities of the heart/mind, activities that are not yet outwardly manifested. In elaborating on this idea, the relevant texts also emphasize the point that, though initially not discernible from the outside, these activities of the heart/mind will inevitably be manifested outwardly, and so one cannot conceal from others the way one truly is. Indeed, the different aspects of the person described earlier are all inter-

active. For example, the life forces (*qi*) that fill the body can be affected by what happens to the body, such as the tastes that the mouth takes in and the sounds that the ear hears; conversely, the life forces can generate speech in the mouth and sight in the eyes. Also, the directions (*zhi*) of the heart/mind can guide and shape the life forces while depending on the life forces for their execution; conversely, the directions of the heart/mind can be swayed if the life forces are not adequately nourished.

It follows from the intimate link between the heart/mind and the life forces, and between the life forces and the body, that the heart/mind is also intimately linked to the body. Various Confucian texts observe how the condition of the heart/mind makes a difference to one's bodily appearance. For example, Mencius observes how one's ethical qualities, while being rooted in one's heart/mind, are reflected in one's face, back, and the four limbs, while the *Great Learning* observes how virtue adorns the whole person just as riches adorn a house. Thus, while the heart/mind is distinguished from other aspects of the person by its modes of operation and its guiding role, it is at the same time intimately linked to other aspects of the person. It is not a kind of "private" or "inner" entity that eludes observation by others, but its condition is inevitably reflected in other parts of the person. In their emphasis on self-cultivation, the Confucians have in mind a transformation not just of the heart/mind but of the person as a whole. Accordingly, if the self is viewed as the object as well as the subject of self-reflection and self-cultivation, it would be more appropriate to describe the Confucian conception of the self as comprising not just the heart/mind but the whole person, including various parts of the body.

Indeed, not only does self-cultivation affect one's whole person, but it also has an attractive and transformative power on others, a power that many Confucians regard as the ideal basis for government. For them, the ideal goal of government is to transform people's character, and the way to accomplish this is to first cultivate oneself and to let the transformative power of one's cultivated character take effect. This does not mean that governmental policies are not important. However, proper policies are themselves a manifestation of the cultivated character of those in power, and properly carrying out policies transmitted from the past also requires a cultivated character. So the ultimate basis for order in society lies with cultivating oneself, and there is an intimate link between self-cultivation and transformation of others' character.

ETHICAL IDEAL

Having considered the Confucian conception of the self, let us consider the nature of the ethical ideal that the Confucians espouse. This ideal is presented through several key terms, three of the most important being *li*, *yi*, and *ren*.

Li originally referred to rites of sacrifice and subsequently broadened in scope to include rules governing ceremonial behavior in various social contexts, such as marriages and burials, as well as ways of presenting gifts, receiving guests, asking after the health of parents, or having audience with a prince. Subsequently, its scope broadened further to include rules governing behavior appropriate to one's social position, such as supporting one's parents in their old age. Though the term can be used to include social norms in general, *li* often retains the connotation of ceremonial behaviour. The *Xunzi*, for example, although sometimes using *li* interchangeably with *li yi* ("rites and propriety") to refer to various social norms, more often uses *li* in connection with ceremonial practices and their minute details. Whether it is the ceremonial or nonceremonial that is emphasized, *li* includes only rules that are part of a continuing cultural tradition and that pertain to the relations between people in different social positions or in recurring social contexts; behavior such as saving a drowning person, though proper, is not a matter of *li*. Also, Confucian thinkers emphasize the importance of the proper spirit behind the observance of *li*, which include attitudes such as respectfulness, attentiveness, and seriousness.

From a contemporary perspective, it may appear puzzling how rules as diverse as those ranging from details of rituals to rules governing conduct between people in different social positions could be placed together under one single concept. However, the rules of *li* do exhibit a unity both in the attitude that they are supposed to reflect and in the social functions they perform. A serious and reverential attitude toward others underlies both the observance of the responsibilities one has in virtue of one's social position and the observance of rules governing ceremonial behavior; a breach of *li*, even in ceremonial contexts—such as being dressed improperly when receiving a guest—demonstrates a lack of the proper attitude and constitutes a serious offense. And, just as the rules governing interaction between individuals in different social positions promote order and minimize conflict, the rules governing ceremonial behavior promote harmony and the proper channeling and beautification of one's feelings in those areas of life associated with strong emotions, such as funerals and mourning or marriage

ceremonies, during which individuals from different families become united as one family. The common spirit underlying the various rules of *li* and their common social functions show that their being grouped together is not based on a failure to distinguish between categorically different areas of life.

Another point worth noting is that the Confucian attitude toward *li* is not entirely conservative. Although the *Analects* contain only one passage that apparently endorses, on economic grounds, a deviation from an existing *li* practice, the *Mencius* is more explicit in asserting that *li* can be overridden by other considerations in exigencies. The *Xunzi* discusses the importance of adapting *li* to the changing circumstances of life, and later Confucian thinkers such as Wang Yang-ming also observe that what is of importance is to preserve the spirit behind *li* rather than to adhere to its minute details.

The Confucian readiness to deviate from or adopt *li* relates to another key term in Confucian thought, *yi*. *Yi* has the earlier meaning of a proper regard for oneself or a sense of honor, involving one's not brooking an insult, and lack of *yi* is often linked to disgrace in early texts. It is subsequently used to refer to what is proper or fitting to a situation, and is linked to *chi*, a character often translated as "shame."

Chi is a reaction to an occurrence or situation that one regards as beneath oneself and potentially lowering one's standing, and it is like shame in presupposing standards to which one is seriously committed. However, it is unlike shame in that it can be directed not just to past occurrences that fall below such standards but also to future prospects of such occurrences. Although *chi* can be directed to the manner in which one is treated in public, it is not typically associated with the thought of being seen or heard, and the typical reaction associated with it is not hiding or disappearing. Rather, it is associated with the thought of one's being tainted by a certain occurrence, and the typical reaction associated with it is to "wash off" what is tainting by distancing oneself from or remedying the situation. Even when directed to the past, it does not carry the connotation of dwelling on the past occurrence, but instead emphasizes a firm resolution to remedy the situation. It is more like the attitude of regarding something as contemptible or beneath oneself, and is linked to ideas such as disdain or a refusal to do certain things.

Yi, for Confucian thinkers, has to do with a firm commitment to certain ethical standards, involving one's disdaining and regarding as beneath oneself anything that falls below such standards. These standards include not being treated in a disgraceful manner as measured by certain public norms, and so one common example of *yi* behavior is a refusal to accept treatment in violation of *li*. However, they also include other measures that go beyond what is honourable or disgraceful by public standards; the *Xunzi* emphasizes a distinction between social honor and disgrace, as opposed to "propriety" (*yi*) honor, and disgrace. Accordingly, *yi* can also provide a basis for departing from a rule of *li*.

The firm commitment that *yi* involves is also related to a certain attitude toward external goods not within one's control. The Confucians advocate one's not being swayed in one's purpose by such external considerations and one's willingly accepting the consequences. In face of adversities to oneself or the prospect of great profits, one is supposed not just to conform to what is proper in one's behavior but also to be free from any distortive influences that might lead to a deviation from what is proper. One should not be subject to fear or uncertainty in face of adversities, and one should willingly accept such adversities, an attitude conveyed in the use of the the character *ming*.

Though often translated as "fate" or "destiny," *ming* does not refer to some opaque force operative in human events that cannot be thwarted. Instead, it serves primarily to express a certain attitude toward occurrences that go against one's wishes and to which one attaches importance, an attitude that follows upon one's recognition of certain constraints on one's activities. The constraints may be causal in that the occurrences are actually not within one's control, such as the failure of one's political endeavors or unexpected illness or death. The constraints may be normative such that the occurrences are something one could alter even though such alteration would involve improper conduct. Whichever is the case, having done what one could within the limits of what is proper, one should willingly accept the undesirable outcome by not engaging in improper conduct to alter things and not worrying about that outcome. Instead, one should resolve to redirect attention to other pursuits, such as Confucius's turning his attention to teaching after having accepted the failure of his political mission.

Finally, let us turn to the Confucian notion of *ren*. In its earlier use, *ren* refers either to kindness, especially from a ruler to his subjects, or to the qualities distinctive of members of certain aristocratic clans. It is used by Confucian thinkers sometimes in a broader sense to encompass all the ideal ethical attributes for human beings and sometimes to refer to a specific ethical attribute that emphasizes affective concern for others. Even for

early Confucians such as Mencius, such affective concern should extend not just to human beings, but also to certain kinds of animals. For later Confucians of the Sung-Ming period, it involves a concern for everything, including plants and what we would describe as inanimate objects. For both early and later Confucians, *ren* involves a gradation. One should have a special concern for parents and family members that one does not have for other people, not just in the sense of a more intense affection but also in the sense of observing certain special obligations to them as defined by *li*. One's relation to other human beings also differs from one's relation to other animals and objects; for example, in the case of animals bred for food, *ren* toward them is primarily a matter of one's being sparing in their use, not using them in excess, and not treating them in an abusive manner.

In later Confucian thought, *ren* is understood in terms of two ideas associated with Heaven (*tian*), which has the connotations of both a supreme diety and the underlying purpose or design of the natural order. In early texts, Heaven, the ideal ruler, and even oneself are often described as forming one body with other people and things. Later Confucian thinkers continue to advocate similar ideas and characterize *ren* in these terms. Heaven and Earth and the ten thousand things originally forming one body with myself, and *ren* involves attaining this state of unity with all things. Though one may have deviated from this state of existence, the task of self-cultivation is to enlarge one's heart/mind until one sees everything as connected to oneself. This idea is sometimes put in terms of a medical analogy. Just as medial texts refer to as a lack of *ren* numbness in one's limbs, an inability to feel for other people is also a lack of *ren*.

In early texts, Heaven is also regarded as what gives birth to things, and its operation is described in terms of a ceaseless life-giving force, an idea highlighted in the early text *Book of Change* (*Yijing*). In later Confucian thought, Heaven's giving birth to and nourishing the ten thousand things is described as its *ren*. The heart/mind of humans should be identical with the heart/mind of Heaven and Earth, which is to give life to things. This is *ren* in the human context, a quality compared to the life-giving power of a seed. This idea of a ceaseless life-giving force is related to the idea of forming one body with the ten thousand things—in giving life to all things, it is as if all things are part of one's own body.

With the above explication of the ethical ideal as background, let us consider the Confucian view of the relation between the self and the social order. As in the case of the relation between the heart/mind and other aspects of the person, there is a sense in which Confucian thinkers emphasize the independence of the self from the social order, and a sense in which they emphasize their intimate relation.

As we have seen, Confucian thinkers emphasize the capacity of the heart/mind to reflect on one's own life, including the activities of the heart/mind itself, as well as its capacity to reshape one's life and its own activities on the basis of such reflection. In virtue of such a capacity, one also has the capacity to step back from one's place in the social order and assess one's relation to it. In the *Analects*, for example, we find passages describing hermitlike individuals who shun the social and political order, at times ridiculing Confucius and his disciples for their persistent and (to these individuals) futile attempts to bring about social and political reform. The Confucian emphasis on the preparedness to deviate from or adapt traditional norms, less explicit in the *Analects* but more conspicuous in the *Mencius* and the *Xunzi*, also presupposes a capacity to step back and reflect on the existing social order.

At the same time, Confucian thinkers also view the self as intimately related to the social roles one occupies. In viewing human beings as a species distinct from other animals, they see the distinction as residing in the capacity of human beings to draw social distinctions and to abide by social norms associated with such distinctions. The point is found explicitly in the *Xunzi*, which states that what makes human beings human beings is not their biological or physiological constitution but their capability of social differentiation and distinction. It also accounts for Mencius's observation that someone who denies social distinctions or fails to make use of this social capacity is, or has become close to, a lower animal. Later Confucians such as Zhu Xi, although acknowledging that certain other animals exhibit something like social relations, also emphasize that human beings are different from other animals in their unique ability to bring to fruition such relations.

Also, as we have seen, Confucian thinkers advocate an ethical ideal that is informed by the traditional social setup that they advocate. The ideal involves a general observance of traditional norms that govern people's behavior either in virtue of the social positions they occupy (such as being a son or an official) or within other kinds of recurring social contexts (such as a host receiving a guest or sacrificial ceremonies); it alsoinvolves the embodiment of certain attitudes (such as reverence) appropriate to such behavior. In addition, it involves the cultivation of desirable qualities within various social

contexts, such as filial piety within the family or devotion when serving in government. Confucian thinkers do acknowledge the importance of a preparedness to deviate from or adapt traditional norms, and later Confucians such as Zhu Xi and Wang Yangming explicitly mention that the fine details of the *li* of ancient times are not all applicable to their times. Nevertheless, they see such deviation and adaptation as themselves based on a certain rationale underlying the social order that, although calling for changes in details in response to changing circumstances, is at the same time something that can be realized only in the evolving social order. It is through participating in this social order and letting oneself be shaped by it that one becomes fully human.

SELF-CULTIVATION

Let us now turn to self-cultivation, the process of, shaping one's own character out of a reflexive concern with the kind of person one is. Confucius stressed learning and reflection as part of the process. The former involves drawing moral lessons from the cultural heritage, which includes such elements as poetry, history, rites (*li*), music, and archery, and embodying such lessons in one's life. The latter involves reflecting on what one has learned so as to adapt it to one's present circumstances. Confucian thinkers after Confucius's times developed different views of human nature, different views of what the basic human constitution is like prior to learning and social influence. These different views have led to different conceptions of self-cultivation.

Some Confucian thinkers, such as Xunzi and Dai Zhen, emphasize the basic biological desires of human beings in elaborating on the basic human constitution. For them, living up to the Confucian ideal is instrumental in satisfying these basic human desires. According to Xunzi, when human beings act out of these desires without regulation, strife and disorder follow. The Confucian Way regulates and transforms such basic desires so that people can satisfy them in an orderly fashion. On this view, self-cultivation involves reshaping and transforming basic human desires, something that Xunzi at times compares to straightening a crooked piece of wood. Dai Zhen also emphasizes the basic biological desires and feelings. He sees the Confucian Way as a matter of one's using one's desires and feelings as a way of gauging others' desires and feelings, and one's satisfying others' desires as one would one's own.

Certain Confucian thinkers, such as Mencius, view human nature primarily in terms of ethical predispositions that human beings share. Mencius opposes the biological conceptions of human nature of his contemporaries and argues that the human heart/mind has a sense of propriety (*yi*) and that human beings do give precedence to propriety over biological desires. He believes that human beings already share certain ethical predispositions, such as the sense of commiseration upon suddenly seeing a child on the verge of falling into a well or the sense of shame when a beggar is given food in an abusive manner. For him, self-cultivation is a process of fully developing these ethical predispositions, a process that he compares to the development of a sprout into a full-grown plant. By directing attention to and nourishing these ethical predispositions, everyone is able to attain the ethical ideal.

Sung-Ming Confucians such as Zhu Xi and Wang Yangming, being self-professed Mencians, draw on the Mencian view but develop it in a different direction. Unlike Mencius, who regards human beings as having ethical predispositions that require nourishment to develop into the ideal ethical attributes, they regard these attributes as already present in the heart/mind in a full-blown form. Certain distortive influences, which they call selfish desires and sometimes selfish thoughts, can prevent the ethical attributes from fully manifesting themselves. *Si*, the character translated here as "selfish," has to do with focusing on oneself, or on people and things with which one forms close associations, in a way that inappropriately neglects other people and things. It involves a separation of the self from other people and things, preventing the life-giving force of *ren* from reaching all things and detracting from one's original unity with them. So, for both Zhu Xi and Wang Yangming, self-cultivation involves restoring the original state of the heart/mind, thereby allowing full manifestation of the ethical attributes. This process is illustrated with analogies such as the clear mirror obscured by dust or still water disturbed by sediments; the ethical task is to remove the dust or sediments to restore the original clarity of the mirror and of water. Zhu Xi and Wang Yangming differ on how to implement this task, the former emphasizing learning and the latter recommending focusing on the operations of the heart/mind.

The Confucian emphasis on self-cultivation, the process of one's doing something to shape one's own character out of a reflective concern with the kind of person one is, is arguably one of the more distinctive features of Confucian ethical thought. This emphasis, however, can lead to the worry that it involves a misdirection of one's ethical attention. This worry can take two different

forms—that this concern with one's own character is either too other-directed or too self-directed.

First, some may be concerned that this emphasis on self-cultivaiton may involve an excessive concern for others' opinion of oneself, especially if it involves one's thinking in terms of cultivating such attributes as *ren* and *yi*. The thought is that the terms that refer to ethical attributes are typically used in third-person descriptions rather than in the content of the ethical person's deliberations. So it appears that the first-person exercise of cultivating these attributes in oneself involves being concerned primarily with the way others would describe oneself. It seems that, in aiming at *ren* and *yi* or at becoming like the ancient sages, one's primary concern is with one's being describable by others in a certain way or with acquiring the kind of stature that the ancient sages have in others' eyes. If so, this kind of concern does seem other-directed in a disturbing way.

Part of the response to this worry is that, even if we grant that *ren* and *yi* are more often used by others as a third-person description of the ethical person, a concern with these attributes need not be a concern with one's being describable by others in a certain way. Instead, it can be a concern with one's becoming like the kind of person that one would oneself describe in this way. That is, the third-person description in terms of *ren* and *yi* can be a description of others by oneself rather than of oneself by others. Furthermore, in being concerned with becoming like the kind of person that one would oneself describe in this way, one's primary object of concern is not with the description but with having a certain character that can be described in this way. Likewise, a concern to be like the ancient sages can be a concern with one's character being like theirs, rather than with one's having the kind of stature that they have in others' eyes.

Still, even if a concern with *ren* and *yi* need not be a concern with how others view oneself, it is a concern with one's own character, and this can lead to the second worry that such a concern may be too self-directed. This concern can be too self-directed in two ways: One may be concerned with preserving or promoting one's own self-image as a certain kind of person, or one may be making one's own character the most important ethical consideration, more important than other-regarding considerations. These two forms of the worry are different. The first focuses on the way in which one is concerned with one's character, how it can take on a distortive form so that one's object of concern is one's image of oneself rather than one's character as such. The second focuses on the importance one attaches to one's own character, how one puts undue weight on one's character in comparison to other-regarding considerations.

In connection with the first form of the worry, we have seen that a concern with *ren* and *yi* is a concern with improving one's character. Just as such a concern need not be a concern with the way others' view oneself, it need not be a concern with preserving or promoting one's own self-image. However, the worry about a concern with self-image may arise with regard to the particular actions that one performs, in relation to both acts of *ren* and acts of *yi*. Let us therefore consider the two kinds of action in turn.

In the case of *ren*, let us take a helping action as example. Suppose one's thought in helping is that one should be doing what is *ren*. If so, it seems that what one is concerned with is that one gives expression to one's *ren* character, that one does what is *ren*, or that one preserves one's image of oneself as a *ren* person. In any case, it seems that there is indeed a misdirection of one's attention in acting.

It is unclear, though, that the Confucians would advocate performing such acts with thoughts about one's own *ren*. For example, in the case of the child on the verge of falling into a well, one's compassionate response is described as a direct response to the imminent death of the child, unmediated by thoughts about one's own character. It is true that, in cases in which one acts not out a sufficient concern for others but out of a concern that one should become the kind of person who would be so moved, one might act with the thought of doing what is *ren*. Even so, one's acting with such a thought is itself a way of transforming oneself so that one will act out of a more direct concern for others. Although, ideally, one should not need to act with such a thought, one's doing so is instrumental in the attainment of this ideal and so should not itself be problematic.

Yi involves a firm commitment to distancing oneself from certain things that one regards as below oneself. In acting out of such a commitment, it seems, one's primary concern is with avoiding smears on one's own character, which is a self-directed kind of concern. Now, even if this is correct, it seems that this kind of self-directed concern need not be problematic for actions that do not—at least directly—affect the well-being of others. For example, in the case of the beggar's rejecting food given with abuse, there does not seem to be anything problematic with the thought that to submit to such treatment to avoid starvation is beneath one's dignity. If there is something problematic about acting out of this kind of concern, it will

have to do with acts that also affect the well-being of others.

Let us therefore consider an act of this kind, such as King Wu's overthrowing the corrupt last king of the Shang dynasty. In the description of this occurrence in the *Mencius*, there is a reference to *chi*, or regarding something as below oneself. Now, although King Wu's attitude was that he regarded it as below him that he, who was in a position to remedy the situation, should allow the people's suffering to continue, there are two ways in which he was also acting out of a concern that is not self-directed. First, what he regarded as below him is also something he would view with aversion if done by someone else in a comparable position. That is, although he reacted with *chi* because of his special relation to the situation, underlying this reaction is the more general attitude of aversion directed to the act, whether by himself or by others, of allowing avoidable suffering to continue. So, in acting out of *chi*, he was in part also acting out of a more general concern that an act of this kind did not take place. Second, his acting out of *chi* is not exclusive of his acting out of a genuine concern for the people. Presumably, it was because he had such concern that he regarded it as below him that the situation be allowed to continue. As long as this other-regarding concern also played a role in his action, his action did not seem to suffer from a misdirection of attention.

This last point assumes that a concern to avoid what is below oneself and a concern for others converge; but what if the two should come into conflict? This takes us to the second form of the worry about an excessive concern with oneself: the worry that one may attach too much weight to one's own character by comparison to other-regarding considerations. Indeed, Mencius himself had been accused of precisely this kind of self-centeredness. The *Mencius* contains several examples of his refusing to see a ruler because he had not been summoned or treated in accordance with certain rules of *li* appropriate to his position. His critics made the point that, if only he had been willing to "bend" himself a little and have audience with the ruler, he might have been able to effect desirable political changes and thereby help the people. By insisting on an adherence to *li*, he was apparently putting more weight on preserving his own sense of honor than on the well-being of the people.

This is a serious charge, and Mencius's response was to draw on the early Confucian view about the transformative power of a cultivated character. The basis of order in society is the cultivated character of those in power, and what Mencius sought to accomplish in the political realm was to "straighten out" those in power. And straightening out others depends on one's being straight oneself; there has never been a case of one's bending oneself while succeeding in straightening others. So, according to Mencius, it is not possible to achieve the desired political changes by bending oneself. And, to the extent that the well-being of the people depends on a reform of the political order, which in turn depends on the transformative effect of a cultivated person, there cannot be a conflict between a concern for one's character and a concern for others. The same point applies to the relation between one's character and the character of others. Given the belief that the transformative effect on others' character is a natural outgrowth of one's cultivating one's own character, there cannot be a conflict between a concern for one's character and a concern for others' character.

See also Chinese Philosophy: Overview; Chinese Philosophy: Contemporary.

Bibliography

PRIMARY SOURCES (IN TRANSLATION)

Theodore de Bary, William, Wing-tsit Chan, and Burton Watson, eds. *Sources of Chinese Tradition*. 2 vols. New York: Columbia University Press, 1960.

Chan, Wing-tsit, trans. *Instructions for Practical Living by Wang Yang-ming*. New York: Columbia University Press, 1963.

Chan, Wing-tsit, trans. *Reflections on Things at Hand: The Neo-Confucian Anthology Compiled by Chu His and Lu Tsu-ch'ien*. New York: Columbia University Press, 1967.

Chan, Wing-tsit, ed. and trans. *A Source Book in Chinese Philosophy*, Princeton, NJ: Princeton University Press, 1963.

Chin, Ann-ping, and Mansfield Freeman, trans. *Tai Chen on Mencius*. New Haven: Yale University Press, 1990.

Lau, D. C., trans. *Confucius: The Analects*. London: Penguin, 1979.

Lau, D. C., trans. *Mencius*. London: Penguin, 1970.

Watson, Burton, trans. *Hsun Tzu: Basic Writings*. New York: Columbia University Press, 1963.

SECONDARY SOURCES

Creel, H. G. *Chinese Thought from Confucius to Mao Tse-tung*. Chicago: University of Chicago Press, 1953.

Fung, Yu–lan. *A History of Chinese Philosophy*. 2 vols. Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1937, 1952, 1953.

Graham, A. C. *Disputers of the Tao: Philosophical Argument in Ancient China*, La Salle, IL: Open Court, 1989.

Graham, A. C. *Two Chinese Philosophers*. 2nd. ed. La Salle, IL: Open Court, 1992.

Liu, Shu-hsien. *Understanding Confucian Philosophy: Classical and Sung-Ming*. Westport, CT: Greenwood Press, 1998.

Munro, Donald J. *The Concept of Man in Early China*. Palo Alto, CA: Stanford University Press, 1969.

Nivison, David S., *The Ways of Confucianism: Investigations in Chinese Philosophy*. La Salle, IL: Open Court, 1996.

Roetz, Heiner. *Confucian Ethics of the Axial Age: A Reconstruction under the Aspect of the Breakthrough toward Postconventional Thinking*. Albany: State University of New York Press, 1993.

Schwartz, Benjamin. *The World of Thought in Ancient China*. Cambridge, MA: Belknap Press of Harvard University Press, 1985.

Shun, Kwong-loi. *Mencius and Early Chinese Thought*. Palo Alto, CA: Stanford University Press, 1997.

Shun, Kwong-loi, and David Wong, eds. *Confucian Ethics: A Comparative Study of Self, Autonomy, and Community*. Cambridge, U.K.: Cambridge University Press, 2004.

Tu, Wei-ming. *Confucian Thought: Selfhood as Creative Transformation*. Albany: State University of New York Press, 1985.

Tu, Wei-ming. *Humanities and Self-Cultivation: Essays in Confucian Thought*. Berkeley CA: Asian Humanities Press, 1979.

Kwong-loi Shun (2005)

CONTEMPORARY

The People's Republic of China (PRC) was established in 1949 under the leadership of Chairman Mao Zedong (1893–1976). This date marks an important watershed in the development of Chinese philosophy. Since 1949 the official ideology of the communist regime on mainland China has without question been Marxism-Leninism-Maoism, while other thoughts were ruthlessly suppressed—especially during the period of Cultural Revolution from 1966 to 1977. The nationalist regime had been driven to the island of Taiwan, which in the early twenty-first century still carries the banner of the Republic of China (ROC, 1912–). The official ideology of the Republic of China was the Three People's Principles, formulated by Sun Yat-sen (1866–1925).

Yet apart from anticommunist and political struggles, other thoughts were more or less tolerated. In Hong Kong—a British colony not taken back by China until 1997—freedom of speech was protected. Furthermore, refugee scholars were allowed to develop and express their thoughts in borrowed space and time. Additionally, there was the Chinese diaspora overseas. Among the various elements in the development of Chinese philosophy, there are several mainstreams of thought that can be discerned. This entry will examine Maoism, Western liberalism, and contemporary neo-Confucianism.

MAOISM

The success of Maoism has been credited to Mao Zedong's talent to adapt Marxism-Leninism to the Chinese soil. Mao contributed two articles—"On Contradiction" and "On Practice"—to expound the philosophy of dialectical materialism. He believed he had succeeded the orthodox line of Marxism-Leninism after the death of Stalin, and as such battled against Western imperialism on the one hand and soviet revisionism on the other. According to Feng Youlan (Fung Yu-lan), Mao's thought went through three stages (Feng 1992).

In the first stage Mao advocated new democracy; in the second, he promoted socialism; and in the last stage he was obsessed with extreme leftist thought. The first stage was represented by Mao's essay "On New Democracy," published in 1940 (Mao 1967). According to his diagnosis at that time, China was not ready for a socialist revolution and had to go through a transitory stage of new democracy. This new democracy was led by the proletariat, the workers and farmers, in conjunction with the petite bourgeoisie. Together they formed a united front, and even the national capitalists were allowed to play a part. This united front joined forces to deal with the problems of a semicolonial, semifeudal society.

When the PRC was established in 1949, the government policy followed such a guideline. But the stage abruptly ended in 1954 when the constitution was drafted. Mao's thought entered the second stage, composed of big change: The goal for the next five years was to accomplish a socialist revolution. Thus the nature of the revolution determined what should be done at any particular stage in the revolution, regardless of actual societal conditions. The telos was a kind of utopian socialism, which surfaced after Mao took control to become the great helmsman of the new China.

In the final stage of his thought, Mao went against the bureaucracy and his own party organization and initiated the disastrous Cultural Revolution, putting his authority behind the Gang of Four. His intention was to do away with private property and to establish communes, which he believed would allow poor people to eat without pay. The release of the destructive powers of the Red Guards caused damages unprecedented in Chinese history.

Mao died in 1976, and in 1977 the Gang of Four was removed by Deng Xiaoping (1904–1997), who adopted an open policy to bring about the revival of China. Since then, capitalism has been seen as a necessary stage for China to go through before the socialist revolution can be implemented. As the doors of China opened to the outside world, many intellectuals were attracted by thought other than Marxism. One editorial in the *People's Daily News* in the 1980s said that Marx's ideas were the product of the nineteenth century, that they did not provide all

the answers, and hence it was desirable to further develop Marxism. Revisionism, therefore, seems to no longer be a crime, and a prevalent view shared by both mainland China and overseas scholars in the late-twentieth and early twenty-first centuries is that a healthy interaction between Marxism, Western liberalism, and Confucianism may find a direction for the future of China.

WESTERN LIBERALISM

Western liberalism was imported to China near the end of Qing dynasty (1644–1912), but somehow it failed to adapt well to the Chinese soil. The most famous liberal in the early Republic of China era was Hu Shih (1891–1962), a disciple of John Dewey, who during the New Culture Movement (c. 1919) promoted the ideals of democracy and science vigorously, urging for wholesale westernization or modernization without reservation (Chow 1960). Yet his approach by gradual reform quickly lost its appeal and radicalism became the vogue. After the communists took over mainland China in 1949, Hu left for the United States, then in 1958 went to Taiwan to serve as the head of Academia Sinica, where he remained until his untimely death in 1962. In his later years he disengaged from political activities and avoided making severe criticisms of the nationalist government under the leadership of Generalissimo Chiang Kai-shek (1887–1975). Instead, he put an emphasis on tolerance.

Among the liberals in Taiwan, one individual who stood out was Yin Haiguang (Yin Hai-kwong, 1919–1969), honored as a spokesman for the democratic movement in Taiwan under the authoritarian rule by the nationalist regime. In 1966 the government (the ministry of education) prevented Yin from teaching his classes at Taiwan University. He would be diagnosed with cancer in 1967 and die two years later.

A follower of logical positivism, Yin was not a deep thinker. His mentors included Bertrand Russell, A. N. Whitehead, Karl Popper, and F. A. Hayek. For Yin, only formal and empirical sciences are cognitively meaningful, yet he was willing to risk his life to fight for the implementation of democratic ideals. In his later years he returned to tradition and lauded Mencius's affirmation of moral courage to defend what is right under adverse environment (Yin 1966). Because he dared to stand up against the mighty powers of an authoritarian regime, eventually Yin was viewed as a martyr and gained respect because he was able to fulfill the duties of an intellectual as he saw it. At the turn of the twenty-first century certain aspects of Western liberalism have held a great attraction for some liberal-minded intellectuals on mainland China since it opened its doors to the outside world.

Both in Hong Kong and Taiwan, the influx of various trends of Western thought has not ceased. In recent years, these trends poured into mainland China with great speed. Although a majority of Chinese intellectuals in the twentieth century criticized Confucianism, that tradition never died. In fact, the most creative talents were found in the contemporary New Confucian movement, which sought to bring about a synthesis between East and West (Bresciani 2001). Despite the prediction of Joseph Levenson in the late 1960s that Confucianism would become something dead that could only be found in museums (Levenson 1968), it appears to be thriving at the present time like a phoenix reborn from the ashes.

CONTEMPORARY NEO-CONFUCIANISM

Confucianism may mean different things to different people, but it is possible to adopt the following threefold division (Liu 1998): (1) spiritual Confucianism, the tradition of the great Confucian thinkers; (2) politicized Confucianism, the tradition of Han Confucianism that served as the official ideology of the dynasties; and (3) popular (or vulgar) Confucianism, belief at the grassroots level that emphasizes family values, diligence, and education—and note that Confucianism in this last stage cannot be separated from beliefs in popular Daoism and Buddhism or various kinds of superstitions. The three forms of Confucianism must be kept distinct on the conceptual level, however, although in reality they are intricately related. Indeed, institutional Confucianism died when the last dynasty was overthrown in 1912, but the other forms of Confucianism survive. For example, some sociologists, such as Peter Berger, believe that vulgar Confucianism has contributed a great deal to the economic miracles accomplished since the 1970s by Japan and the so-called Four Mini-dragons: Taiwan, Hong Kong, Singapore, and Korea (Berger and Hsiao 1988). Politicized Confucianism has also attracted a large number of admirers. The cover of the June 14, 1993, issue of *Time* magazine was a portrait of Confucius; the issue reported that Francis Fukuyama, author of *The End of History*, was of the opinion that the kind of soft authoritarianism practiced in Singapore posed a greater challenge to Western liberalism than did Islam.

Spiritual Confucianism is a vigorous movement of thought. In 1986, mainland China designated Contemporary New Confucianism as a national research program for a period of ten years (Fang 1997). At first its scope was

not clearly defined; it included scholars with various backgrounds, such as scholar-thinker Liang Shuming (Liang Sou-ming, 1893–1988), scholar-statesman Zhang Junmai (Carsun Chang, 1887–1969), historian Qian Mu (Ch'ien Mu, 1895–1990), and intellectual historian and political commentator Xu Fuguan (Hsü Fu-kuan, 1903–1982). After broad consultations and extensive discussions under the guidance of Fang Ke-li and Li Jingquan, the directors of the program, ten case studies were completed.

In addition to the above-named scholars, the program also included six philosophers: Xiong Shili (Hsiung Shih-li, 1885–1968), Feng Youlan (1895–1990), He Lin (Ho Lin, 1902–1992), Fang Dongmei (Thomé H. Fang, 1899–1977), Tang Junyi (T'ang Chün-i, 1909–1978), and Mou Zongsan (Mou Tsung-san, 1909–1995). Later, four younger scholars were included: Yu Yingshi (Yü Ying-shih, 1930–), Liu Shuxian (Liu Shu-hsien, 1934–), Cheng Zhongying (Cheng Chung-ying, 1935–), and Du Weiming (Tu Wei-ming, 1940). The addition of Ma Yifu (1883–1967), a noted scholar in classics studies from the older generation, was also added. Eventually, fifteen names were chosen, and these fifteen may be assigned to four groups in three generations (Liu 2003):

The First Generation:

Group I: Liang, Xiong, Ma, and Zhang

Group II: Feng, He, Qian, and Fang

The Second Generation:

Group III: Tang, Mou, and Xu

The Third Generation:

Group IV: Yu, Liu, Cheng, and Du

Liang, Xiong, and Ma have been recognized as the three elders in the first generation, all of whom chose to remain in mainland China. Only Zhang, as the leader of a third force political party, fled overseas. Liang is seen as the person who initiated the movement, but it was Xiong, known only in a small scholarly circle, who became the spiritual leader of contemporary neo-Confucianism in the narrower sense. The three important representatives of the movement, Tang, Mou, and Xu, were disciples of Xiong.

The scholars in Group II were somewhat younger. Feng and He chose to remain in mainland China. Qian Mu fled to Hong Kong, where he and Tang became the cofounders of New Asia College, an undisputed center for contemporary neo-Confucianism. Fang, once a teacher of Tang, fled to Taiwan, where he taught at Taiwan University and had Liu and Cheng among his disciples. Xu and Mou also went to Taiwan, and from 1963 to 1969 made Tunghai University in Taichung a second center for contemporary neo-Confucianism.

THE IMPORTANCE OF THE THREE GENERATIONS. Without question, the mainstay of contemporary neo-Confucianism is represented by the second generation of scholars: Tang, Mou, and Xu. They signed the famous "Manifesto for a Reappraisal of Sinology and Reconstruction of Chinese Culture," drafted by Tang and issued on New Year's Day in 1958. The other signatory was Zhang, of the first generation (Liu 1996). The third generation scholars, disciples of Hong Kong and Taiwan New Confucians, received advanced academic training and had teaching careers in the United States, thereby acquiring an international dimension in their thought (Liu 2003).

For obvious reasons only refugee scholars outside of mainland China were able to make significant contributions to the further development of Confucian thought. Fang, of the older generation, received his academic training in the United States. He had a grand scheme of philosophy of culture with a comparative perspective of a fourfold division: ancient Greek, modern European, Chinese, and Indian. He also strongly criticized the dualism of modern European thought, which was believed to have a hidden nihilistic tendency (Fang 1957). Fang opted for messages of creative creativity and comprehensive harmony of primordial Confucianism (Fang 1981), as well as urging others to overcome the limitations of different cultures in order to bring about a synthesis of East and West.

Tang and Mou were also well versed in Western philosophy. While Tang had a Hegelian bent, Mou showed an unmistakable Kantian temperament. The famous manifesto drafted by Tang urged the sinologists in the West to study the Chinese culture not just through the eyes of the missionaries, the archaeologists, or the political strategists, but with a sense of reverence and sympathetic understanding of that culture. According to the manifesto, the wisdom of Chinese philosophy is crystallized in its philosophy of mind and human nature (*xin* and *xing*), an unmistakable reference to Sung-Ming neo-Confucianism. Although recognizing the need for the Chinese culture to learn from the West by absorbing its achievements in science and democracy, the manifesto claims that there is something invaluable in the Chinese tradition and suggests that the West may learn from Chinese thought in the following five items:

(1) The spirit to assert what is here and now and to let everything go (in order for nature to take its own course);

(2) All-round and all-embracing understanding or wisdom;

(3) A feeling of warmth and compassion;

(4) The wisdom of how to perpetuate the culture;

(5) The attitude that the whole world is like a family.

THE THREE TRADITIONS' DOCTRINE. In his later years Tang devoted himself to tracing the origins of insights in traditional Chinese philosophy. The last work he published was a comprehensive system of philosophy conceived in his lifetime. The book deals with the whole existence of humans and tries to understand the different activities of the mind, distinguishing the following nine worlds:

(1) the world of discrete things;

(2) the world of species and genus in terms of empirical generalization;

(3) the world of functional operation;

(4) the world of perceptions interpenetrating with one another;

(5) the world of contemplation of what is transcendent and vacuous;

(6) the world of moral practice;

(7) the world of aspiration toward God;

(8) the world of emptiness (*śūnyatā*) of both the self (*ātman*) and elements (*dharma*);

(9) the world of the embodiment of heavenly virtues.

Mou, however, was perhaps the most original thinker in his generation. Going further than the manifesto, he formulated the doctrine of three traditions:

(1) The assertion of Daotong (the tradition of the Way): We must assert the value of morality and religion, jealously guarding the fountainhead of the universe and human life as realized by Confucius and Mencius through a revitalization of the learning of the mind and the human nature.

(2) The development of *Xuetong* (the tradition of learning): We must expand our cultural life and further develop the learning subject as to absorb the Western tradition of formal sciences such as logic and mathematics on the one hand and empirical sciences on the other.

(3) The continuation and expansion of *Zhengtong* (the tradition of politics): We must recognize the necessity of adopting the democratic system of government as developed in the West in order to fulfill truly the political ideal of a government of humanity of the sages and worthies in the past.

After Mou dug deeply in the Chinese tradition and devoted himself to scholarly studies of Daoism, Confucianism, and Buddhism, in the latest stage of his thought he brought into focus a comparative perspective. In *Intellectual Intuition and Chinese Philosophy* he pointed out that the major difference between Chinese and Western philosophies lies in the fact that the three major Chinese traditions all believed in the possibility of intellectual intuition, whereas major Western traditions deny that there is such a possibility. Mou used Kant as his point of departure because Kant believed that all human knowledge must depend on sensible intuition, and only God has intellectual intuition. For Kant, freedom of the will, immortality of the soul, and the existence of God can only be postulates of the practical reason, hence Mou was of the opinion that Kant could only develop a metaphysics of morals, not a moral metaphysics. In *Phenomenon and the Thing-in-itself* Mou made a distinction between what he called "ontology with adherence" and "ontology without adherence." The former has been highly developed in the Western traditions, and the latter has been elaborately formulated in the Oriental traditions. When the infinite mind puts restrictions on itself, the knowing subject is formed; this is the result of a dialectical process. The adherence of the knowing mind and the realization of the infinite mind actually share the same origin. It is here that a foundation for the unity of the two perspectives can be found (Liu 1989).

Because the second generation neo-Confucians developed their ideas within a most adverse environment, they tended to stress what is positive in the Confucian tradition. But the third generation neo-Confucians face a very different context—they presuppose the pluralistic framework of the West. For example, Du Weiming feels that there is no need to prove that the Confucian tradition is better than other spiritual traditions, so long as it can be shown that it is one of the worthy spiritual traditions in the world; from modern to postmodern era, the flexible understanding of reason and the emphasis on harmony in the Confucian tradition are to its advantage. Liu Shuxian offers a new interpretation of *liyi fenshu* (one principle, many manifestations) which he inherited from

Sung-Ming neo-Confucian philosophy. He fully realizes that the Confucian tradition certainly does not have a monopoly of one principle, which would find different manifestations in different spiritual traditions. For example, the Golden Rule, credited to Confucius by Hans Küng, has been formulated in different ways in the East and in the West as well as in the ancient and modern times (1993). Thus, Liu vigorously supported the formulation of a global ethic in the awakening of a global consciousness. Because the world has turned into a global village with only limited resources, and because people need to live peacefully together, certainly the Confucians will have something significant to say in the future.

See also Chinese Philosophy: History of.

Bibliography

Berger, Peter L., and Hsin-huang Hsiao, eds. *Search for an East Asian Development Model*. New Brunswick, NJ: Transaction Publishers, 1988.

Bresciani, Umberto. *Reinventing Confucianism: The New Confucian Movement*. Taibei, Taiwan: Ricci Institute, 2001.

Chan, Wing-tsit. *A Source Book in Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1963.

Chow, Tse-tsung. *The May Fourth Movement: Intellectual Revolution in Modern China*. Cambridge, MA: Harvard University Press, 1960.

Fang, Ke-li. *Xiandai xinruxue yu zhongguo xiandaihua* [Contemporary New Confucianism and the modernization of China]. Tianjin, China: Tianjin renmin chubanshe, 1997.

Fang, Thomé H. *Chinese Philosophy: Its Spirit and Its Development*. Taibei, Taiwan: Linking Publishing, 1981.

Fang, Thomé H. *The Chinese View of Life*. Hong Kong: The Union Press, 1957.

Feng Youlan (Fung Yu-lan). *The Hall of Three Pines: An Account of My Life*. Translated by Denis C. Mair. Honolulu: University of Hawaii Press, 2000.

Feng Youlan. *Zhongguo xiandai zhexueshi* [A history of modern Chinese philosophy]. Hong Kong: Zhonghua shuju, 1992.

Fukuyama, Francis. *The End of History and the Last Man*. New York: Free Press, 1992.

Furth, Charlotte, ed. *The Limits of Change: Essays on Conservative Alternatives in Republican China*. Cambridge, MA: Harvard University Press, 1976.

Küng, Hans, and Karl-Josef Kuschel, eds. *A Global Ethic: The Declaration of the Parliament of the World's Religions*. London: SCM Press, 1993.

Levenson, Joseph R. *Confucian China and Its Modern Fate*. 3 vols. Berkeley: University of California Press, 1968.

Liu, Shu-hsien (Shuxian). "Postwar Neo-Confucian Philosophy: Its Development and Issues." In *Religious Issues and Interreligious Dialogues*, edited by Charles Wei-hsun Fu and Gerhard E. Spiegler. Westport, CT: Greenwood Press, 1989.

Liu, Shu-hsien. "Confucian Ideals and the Real World: A Critical Review of Contemporary Neo-Confucian Thought." In *Confucian Traditions in East Asian Modernity*, edited by Tu Wei-ming, 92–111. Cambridge, MA: Harvard University Press, 1996.

Liu, Shu-hsien. *Essentials of Contemporary Neo-Confucian Philosophy*. Westport, CT: Praeger Publishers, 2003.

Liu, Shu-hsien. *Understanding Confucian Philosophy: Classical and Sung-Ming*. Westport, CT: Greenwood Press, 1998.

Makeham, John, ed. *New Confucianism: A Critical Examination*. New York: Palgrave Macmillan, 2003.

Mao Zedong (Mao Tse-tung). *Selected Works of Mao Tse-tung*. Vol. 2. Beijing: Foreign Languages Press, 1967.

Mou Zongsan (Mou Tsung-san). *Xianxiang you wuzesheng* [Phenomenon and the thing-in-itself]. Taibei, Taiwan: Xuesheng shuju, 1975.

Tang Junyi (T'ang Chün-i). *Shengming cunzai yu zinlin jingjie* [Human existence and the worlds of the mind]. 2 vols. Taibei, Taiwan: xuesheng shuju, 1977.

Yin, Haiguang (Yin Hai-kwong). *Zhongguo wenhua te changwang* [Reappraisal of cultural change in modern China]. 2 vols. Taibei, China: Book World, 1966.

Shu-hsien Liu (2005)

DAOISM

Philosophical Daoism (also spelled Taoism) dates from the classical period (fifth through third century BC) and conventionally refers to the contents of the *Zhuangzi (Chuang Tzu)* and the *Laozi (Lao Tzu or Daode Jing/Tao-Te Ching)*. Some extend the term to cover less philosophical transitional texts of popularized Daoism of the Han (second century BC)—for example, the *Liezi* and the *Huainanzi*. Another movement, called Neo-Daoism, dates from the end of the Han (200-plus). The term "Daoism" is fundamentally misleading since no group, no leader, and no association linked those thinkers. The Han historians who coined the term centuries later viewed the philosophers as founders of their credulous religion, Huang-Lao, which flourished after classical philosophy was extinguished by Qin despotism (220 BCE). The main basis for the classification was thus: (1) their philosophical interest in the concept of *dao* (way or normative guide); and (2) relatively skeptical, anarchic, antisocial attitudes which contrasted with Confucianism.

PHILOSOPHICAL DAOISM: A QUICK TOUR

The concept of *dao* (tao) was central to ancient Chinese philosophizing. It is essentially a normative, practical concept—a way or guide to action. Almost all ancient Chinese thinkers philosophized about *dao*, about choosing, reforming, following *daos* as well as understanding

their relation to "constant" nature (*tian*$^{\text{nature:sky}}$), to human nature (*xing*$^{\text{nature}}$), and to society.

Those subsequently classed as daoist thinkers are distinguished by their more metaethical interest in *dao* in contrast to Confucians and Mohists who mainly advocated a variety of normative *daos*. Daoists discussed mainly three kinds of *dao*: human (or social) *dao*; *tian*$^{\text{natural}}$ *dao*; and "Great" *dao*. When I instruct you to cross the road on the green light, I am delivering a bit of human *dao*. Natural *dao* (often translated heavenly *dao*) is akin to what we would consider the constancies of science. Natural *dao* is the way things reliably (constantly) happen. The salient contrast is that Chinese thinkers do not elaborate this idea with the idea of law—universal, modally necessary propositions. Great *Dao* refers to the entire actual history of space and time—whatever has happened, is happening or will eventually happen in the universe is part of Great *Dao*. *Dao* is simply the counterpart of Ludwig Wittgenstein's "All that is the case." What amounts to determinism in Chinese thought is treating *tian dao* and Great *Dao* as identical—the constancies in nature make only one world history possible.

Daoist philosophers typically distance themselves from various human *dao* (paradigmatically Confucian or Mohist *dao*) by contrasting them to natural *dao* and/or Great *Dao* (the actual *dao*). Ancient Chinese moralists had tended to treat *tian*$^{\text{nature:sky}}$ as the authority for their account of the correct human *dao*. "Primitivist" daoism, usually regarded as an earlier form associated with Yang Zhu (sometimes called "Yangism") and the *Laozi* advocates being natural and rejects the social (historical or conventional *daos*). "Mature" Daoist analysis, typified by the inner chapters of the *Zhuangzi* centers on the insight that while human *daos* are normative, neither the natural nor the actual *dao* are, and that the guidance of any human *dao* depends on and presupposes natural *dao*.

Mature Daoism avers that nature does not authorize or endorse any *particular* social *dao*. This claim has two versions: pluralist and primitivist. Denying *one* is compatible with their being either *many* or *none*. The pluralist (relativist) reading of the claim takes it to entail that de facto rival practices are natural *daos* in virtue of their being actual practices. Thus, they are continuous with natural constancies. Great *Dao* versions suggest that all *actual* rival *daos* are part of Great *Dao* simply in virtue of actually being followed—*daos* are made by "walking" them. Pluralist Daoists end up vaguely associated with anarchism because they reject the Confucian-Mohist assumption that political authority exists to bring about a harmony of *daos*—making everyone follow a single *dao*. The social world survives as well (or better) when people follow different ways of life. Focus on either *tian*$^{\text{nature:sky}}$ *dao* or Great *Dao* thus undermines the sense that it is imperative to impose any particular first-order *dao* on all of society.

Primitivist versions of Daoism, however, typically take the form that nature does endorse a particular normative *dao*, though not a human one. That a single, constant, correct way of life cannot be expressed or presented in practices, rules, narratives, maps, examples, songs or any other human or social form of communication and advocacy. Though we usually think in terms of one natural *dao*, there could, in principle, be multiple, equally "primitive" *daos*.

The ambiguity between these two versions is neatly expressed in the opening stanza of the *Daode Jing*: "Any *dao* that can *dao* (guide) is not constant *dao*." Both primitive and mature versions underwrite a shared theme of harmony with nature—the pluralist seeing the point of such harmony as permissive and tolerant and the primitivist seeing it as more intolerant, as rejecting or prohibiting any conventional *dao*.

Metaphysically, Daoism is naturalistic though religious versions of Daoist primitivism evoke mystical and supernaturalist themes that remind interpreters of European, Middle Eastern, and South Asian mystical supernaturalism. The various mystical analyses, following the Indo-European model, are buttressed by an intuitive epistemology. Detaching from social *daos* means eschewing language, words, and norms of use that underwrite public discourse (reasoning) about what to do. The essential form language norms take is learned inclinations to distinguish or discriminate what is "this" from "not this." Intuitionism advocates "recovering" the simple, primitive, pre-social dispositions by forgetting names and the attitudes that linked them to action.

Translated to the language of Indo-European rationalism, this line of reasoning treats the pluralist insight as entailing an irrationalist absolutist conclusion. It endorses the tempting illicit inference from relativism (our distinctions are "socially constructed") to dogmatic, absolutist monism (there are no distinctions in [moral] reality). Religious and other nonphilosophical interpreters view this non-sequitur as the essence of Daoism.

HISTORICAL OUTLINE: THE RANGE OF DAOISM(S)

Moralizing schools proposed rival *dao* (social guiding discourse) for general order. The term "Daoism" was

applied to the general reflection on what it was to propose, accept, reject, a *dao*—the epistemology, semantics and metaphysics of *dao*. This meta-focus on *dao* inclined these thinkers to a variety of metaethical positions: skeptical, relativist, and mystical. Their doubt about the goal of unifying society's *dao* and their philosophical "distancing" from direct advocacy of a normative *dao* made them seem ethically amoral and politically liberal, libertarian and even anarchistic. Zhuangzi in particular emphasized his differences with the moralists—the Confucians and Mohists.

Religious Daoism, by contrast, is an extremely broad classification of popular and/or local religions that took many different forms at different times in China are distinguished mainly negatively—by their not identifying primarily with Confucianism or Buddhism. The earliest known example was the "Huang-Lao" sect that flourished at the beginning of China's philosophical dark age induced by Qin repression and Han Confucian orthodoxy. Religious Daoism's relation to philosophical Daoism is both controversial and obscure. Daoism acquired organizational religious trappings from its interactions with Buddhism after the latter arrived in China around the second century AD. Characterizing aspects of Daoism as "religious" prior to that time is simply to draw a distinction between relatively credulous, superstitious, popular readings and more reflective, skeptical, philosophical readings of the same texts.

It is common to trace a Daoist political "ethos" to hermits who lectured Confucius against social involvement. Another way to trace Daoism's origins runs through the shadowy figure of Yang Zhu. He seems to have drawn on something like the romantic conflict between what Yang Zhu thought of as our nature (our natural mode of development) and social-political structures. We know of him mainly through the Mencius's attack on him as an egoist. Mencius reports that Yang Zhu refused to risk a single body hair to "save the empire." The Yangist theme survives in Daoism—where it is also known as "primitivism."

In more orthodox Chinese moralizing, Confucian and Mohist, human nature is "shaped" via our being socialized through following some *dao*—a shared guiding discourse. Primitivists resist this social shaping and seek to "restore" natural and spontaneous patterns of action. Mencius's attack probably distorts primitivism, therefore, since it arguably rejects only social or conventional *daos*, not all guidance.

The picture is further blurred by subsequent history of Daoism. These interpretive tensions became part of the Chinese conception of Daoism in later periods. A philosophical "dark age" prior to the Han followed by the credulous Confucian orthodoxy that emerged as the "spirit" of the Han dynasty, effectively extinguished critical philosophical thought.

The autocratic rulers of the Qin and Han were highly superstitious and the courts tended toward the more authoritarian, dogmatic readings. Such readings sustain their claims of special or esoteric access to knowledge (for example, trance states induced by breathing exercises) that imperiously ignores any demand for deeper justification (for example, intuitionism or mysticism).

The Han eventually enshrined Confucianism as an official orthodoxy and the basis of the examinations that qualify one for political office. This inclined Confucian theorists to view their tradition as embracing and subsuming all other learning and treating Daoism as compatible with Confucianism. This task was made easier by emphasis on the more "religious" texts identified as Daoist. These Han texts borrowed stories, attitudes, and phrases from the two original daoist writings. Hence, the *Liezi*, the *Huainanzi*, and the *Baopuzi* were often included among the classics of Daoism.

The fall of the Han saw the emergence of a "mixed teaching"—Neo-Daoism. Its most influential writers were avowed Confucians motivated by the urge to "harmonize" the two traditions. The first, Wang Bi (c. 300) probably was familiar mainly with the Han religious echoes of Daoist thinking. He interpreted the *Laozi* alongside a cosmological divination manual, *The Book of Changes* (*I Ching* or *Yijing*). He treated *dao* as a term of creationist cosmology. *The Book of Changes* with its *yin-yang* account of change and its generational cosmology thus entered the list of Daoist texts.

Neo-Daoism, in turn, eventually facilitated and informed the assimilation of the newly imported Indian Buddhism which in turn inspired the development of a uniquely Chinese form of Buddhism—Chan (Zen). This blended outlook is one major vehicle by which Daoism survived in Chinese thinking.

One could say that Daoist philosophy, per se, was successfully extinguished by imperial suppression of thought that initiated China's philosophical dark age and the institution of the Han "official Confucian orthodoxy" which cemented it firmly in place. The result is that so many religious forms are regarded as Daoist, the term may have no meaningful value in identifying a unified philosophical trend or movement.

AN "INTERNAL" HISTORY OF PHILOSOPHICAL DAOISM

The *Zhuangzi* contains its own history of thought (Chapter 33) tracing the development of ideas leading up to Zhuangzi's doctrine, which we *could* call "mature" Daoism. The word "Daoism" however, had still not been coined and the *Zhuangzi* history takes itself as simply tracing a movement of thought motivated by progressive attempts to remove bias and replace narrow perspectives with more impartial ones.

This "internal" account of a "Daoist" dialectic starts from a traditional baseline and takes its first step with thinkers like the Mohists (90/33/17). They question the assumption that tradition is right and seek for a neutral standard for deciding which social *dao* to use in the public moral education project. They legitimized their chosen *dao* by appealing to a standard—the intentions of *tian*[nature:sky]. That was equivalent to a *natural* distinction between benefit and harm (the *natural* urge to benefit and aversion to harm). This standard motivates their utilitarianism, which formed such a demanding morality that they "wore out their heels" running from state to state stopping wars, opposing despots, relieving starvation and so forth.

The next phase of the urge to impartiality leads to a version of primitivism—attempting to identify and remove the biases that socialization has instilled in the *xin*[heart:mind]. It is these, they argue that divide men, induce competition, fuel disagreement, and make life miserable by creating desires which can be satisfied for only a few. They targeted the obviously social desires such as status, "cultivated" tastes, honor and so on (other texts attribute related slogans to them; for example, "to be insulted is no disgrace" and "farewell to narrowness.") The natural desires, they argued, are few and more easily satisfied. They are also universal in contrast to the cultivated desires which differ depending on tradition.

The first clear focus on the meta-nature of *daos* is then attributed to Shen Dao, Tian Ping, and Peng Meng. Shen Dao is also famous as a contributor to legalism. They started from the shared early classical assumption that we should follow a natural *dao*[guide]. And, like Mencius, thought of language as the paradigm of what is "unnatural." So impartiality comes from avoiding all language—all judgment about what is *shi-fei*[this-not this]. They motivate this by developing a concept of "Great *Dao*"—Great *Dao* is collection of all things and all events in a kind of everything concept. It is *dao*-like because it is a process, not an object. It is the history of all objects through all time, including the future. It "leaves nothing out."

The various competing normative *dao*[guides] imply that some possible future course of events (a *way* things might go) is the one we should "walk." To *dao*[guide] that *dao*[guide] is to recommend the future histories that result from the selected "walking." To learn one of those *dao*[guides] is to learn how to contribute to bringing about some future history. While there are many such possible future histories, there is only one actual history—one actual past and one actual future. He calls that actual history of the world, that actual course of events that all things will follow, Great *Dao*. The actual is natural so the Great *Dao*, the natural pattern of behaviors, events, and processes, requires no learning, no knowledge, no language or *shi-fei*[this-not this] distinctions.

This conception of the actual *dao* of the past and future has a deterministic flavor. Nothing we do can "miss" the Great *Dao*. Even a clod of earth cannot miss it (HY/92/33/50). From this conception of the world as all that is or will be, Shen Dao draws fatalistic sounding conclusions—"abandon knowledge and discard self" (HY/92/33/45). Flow with the inevitable and be indifferent, make not *shi-fei*[this-not this] judgments. He rejects all moral (and other) teaching and just … lives …

The account is critical of Shen Dao's theory: "Shen Dao's guide does not lead to the conduct of a living man but the tendency of a dead man. It is really very strange. … They made reversing what is human a constant value; didn't take the common view and couldn't avoid inconsistency. That which they called a guide was a non-guide and what they approved could not but be wrong. [They] did not know how to guide. …" (HY 92/33/51-4). Laozi, traditionally regarded as the founder of Daoism slots into the "internal history" at this point. He, like Shen Dao, is attracted to the conception of impartiality that underlies a recommendation that we "abandon knowledge." However, the *Laozi* does not seem to appeal to a Great *Dao* conception to justify it. Its appeal is more to "freedom" than to "determinism." The freedom, however, is relative to society, conventions and language.

The *Daode Jing* contains a classic expression of the ancient Chinese contrast theory of names. Words come in opposites and are learned together. In learning them, we learn to divide things in one of a range of possible ways and become "blind" to alternate ways of doing it. Along with the socially sanctioned distinctions, we learn to desire in socially prescribed ways thus acquiring the society's *dao*. Behavior motivated by this system of names, divisions and desires is called *wei*[deem:do] and the Laozi

advocates that we should avoid *wei*-ing (*Wu-wei* = lack deeming-actions).

Thus if we can "forget" what we learn from conventional society, we can return to natural spontaneous action—symbolized by the newborn child. The child does move, but the motions are not motivated by any conception of how to divide the world into socially sanctioned categories or conditioned, socialized desires. We recover a natural freedom that is also a much reduced level of simpler desires that will enable people to live in peace—not necessarily together because the "natural" structure of primitive desires may only support society at the level of Neolithic villages. This idyllic return is sometimes called primitivism and the *Daode Jing* contains a classic depiction of this peaceful world of agrarian villages whose peaceful, contented inhabitants lack any incentive even to visit the next door village—though they can "hear the cocks crow and the dogs bark" (Ch. 81).

This primitivism still countenances a "natural" *dao*—a prescriptive course of action that originates from *tian*[nature:sky] and contrasts with the artificial *daos* of moralizing society—the Confucians and the Mohists. The moralists like Mozi would also maintain that their conception of the moral *dao* was the *dao* of *tian*[nature]. Mencius, similarly, appeals to the normative authority of *tian*[nature] to justify action according to the "innate" (Confucian) moral tendencies that grow in the *xin*[heart-mind]. Mencius is critical of Yang Zhu, a figure often treated as a proto-Daoist who, thought he seems to have no metatheory of Dao, is reputed to espouse a version of primitivism (which some call Yangism). Implicitly that his argument also seems to appeal to the notion of *tian*[nature:sky] as an authority—where the command of *tian*[nature:sky] is a "simpler" dictate to care for the essentials of life and abandon the dangers of political and social involvement.

In fact, despite leaving behind the deterministic tone of Shen Dao's "Great *Dao*," the *Laozi* is caught in a similar paradox. Shen Dao's reasoning illustrates what is wrong with any *dao* that has blanket anti-language interpretations. His advice to "abandon knowledge" is self-refuting advice since that falls within the range of what it advocates abandoning—prescriptive doctrines. It is a prescriptive paradox—if we obey it, we disobey it. So we can continue to learn *daos* and still be natural. The deep point is that natural *daos* are irrelevant to the issues being debated by the moralists. In being natural they lose their capacity to guide. The could not warrant any particular *shi-fei*[this-not this] that was relevant to judging or choosing some human action. Everything that happens must be the same—either all *shi*[this:right]; all *fei*[not-this:wrong], all both or all neither. The crucial implication of his approach is that an injunction like "be natural" has no normative force.

So the trend of thought first recognized as forming the pattern we have come to call Daoism is one that reacts to Confucian conventionalism by trying to find a more universal, impartial point of view. It may include movements such as primitivism that seek to remove all social-conventional influence and those that analyze how conventions shape and induce our attitudes, desires and actions, and to a general interest in natural or transcendent standards or "ways" that undermine dogmatic conventionalism. The hint of paradox in the latter positions may be recognized and embraced or accepted as inescapable. The paradox, particularly of the anti-language implications of these developments led to the mature phase of philosophical daoism.

MATURE PHILOSOPHICAL DAOISM: THE ZHUANGZI

Classical thinkers found names and language relevant to not only Daoism. The *Analects* community of Confucianism became committed to "rectifying" names to make role-based behavior guides prescriptively reliable. Gongsun Long presented himself as defending Confucian practice with his one-name-one-thing rule. Confucians may not have wanted to claim him since he derived from this his notorious commitment to the assertability of "white horse not horse." Mohism produced the most sophisticated of these theories—an early version of semantic realism that may have derived from reflections on Mozi's three *fa*[standards] (standards) of *yan*[language] (language).

Mohists argued that name boundaries are determined by objective similarities and differences in things. So, against the name "rectifiers," they maintained that a "reality" can properly be called by several names—at times general or particular. But this position, they discovered, still left many puzzles. One was this: Which similarities count in correct naming or types? The others puzzles concern how to deal with compounding of names and strings. Mohists developed no syntactical theory of word-roles such as adjective and verb. All descriptive terms that picked out parts of reality were called "names." The way the parts combine when the words combine, however, struck them as irregular. Beyond noting and classifying some of the variation, however, Mohists did not seem to propose a systematic solution.

Zhuangzi, traditionally cited as the second daoist philosopher, engaged frequently in discussion with Hui Shi, cited as among the members of the "school of

names." Hui Shi however, seems to represents a third posture within the school of names—name relativism. His implicit criticism of the realists emerges when he draws attention to terms that are implicitly comparative; for example, 'large', before, and 'high', and indexicals, for example, 'today' and 'south.' These do not apply in virtue of objective similarity. More theoretically, he averred that any two things, no matter how different, are similar in *some* respect, and no matter how similar, are different in *some* respect. So any two things can be included in the scope of a term in virtue of some similarity or could be placed included in a different range and named by different terms in virtue of *some* difference. Thus placement of things in a named range, even if based on similarity and difference, is not a *constant* or *reliable dao*.

Despite the relativist ground of his linguistic reasoning, Hui Shi seems to have drawn an absolutist conclusion. Since all linguistic divisions are relative, the absolute or language-independent world must be devoid of any distinctions—a mystical, unnamable, one. In effect, distinctions and differences among things are socially constructed. Since naming is based on distinctions and since language is constructed of names, Hui Shi's ending position resembles that of Shen Dao and Laozi—the familiar, anti-language, Daoist, mystical, monism—all is an inexpressible one. From this, Hui Shi then drew a hybrid Mohist-Daoist prescriptive conclusion, "Love all things equally, the cosmos is one body."

The Later Mohists, however, diagnosed the problem in these defeatist, negative conclusions about language and distinction-based judgment. The problem may take different forms and the Mohist diagnosis is repeated in three forms. First, distinctions are manifest in language in *shi-fei*[this-not this] indexicals. With a name in view, it does or does not apply to some indexically accessed item—"this" or "not this." To have a distinction is for something to be "not-this." So, to oppose all distinctions is to oppose *fei*-ing. However to oppose *fei*-ing is to *fei fei*-ing. Anyone who does so confronts a pragmatic contradiction. To *fei fei*-ing is to *fei*. Similarly, to *deem* all language as *bei*[not-acceptable] (not-acceptable) is *bei*[not-acceptable]. And since, in the context of primitivism, such views amount to rejecting learning or education in language the third form is to teach that teaching is wrong is wrong. This explains the paradox that plagued Shen Dao and Laozi and any doctrine (for example, some interpretations of Mencius) that denigrate principles or linguistic guides in general in favor of following only "natural" *daos*.

Zhuangzi takes the paradox seriously and responds first by abandoning all such anti-language claims and appeals to *tian*[nature] as an authority for *dao* while, second, still sustaining his skeptical, relativist distance from Confucian convention. The trick is to note that all (actually existing) language is natural—as natural as such nonhuman sounds as the whistling of wind or twittering of birds. Humans, their societies, and their languages are products of nature as much as are ground squirrels and their high-pitched chirps.

The disagreeing human thinkers—the 10,000 distinctions and differences marked in language—are among the "pipes of *tian*[nature]." Thus he avoids taking an anti-language stance while still standing as an ironic "Daoist" distance from convention. From that stance he can continue to "poke fun" at the moralists—not for pretending to express "natural" *dao* (they do) but insisting that others, their opponents, *do not*. The sense in which theirs are natural (as indeed they are) is the sense in which their opponents *daos* are natural too.

All the warring discourse *daos* are, by hypothesis, natural. How, Zhuangzi asks, can a language exist without its being acceptable (in that community) to speak it? All *dao* that are actually walked (that generate behavior) are (in virtue of having emerged naturally in a natural world) *natural*. All the *daos* that anyone may *actually* appeal to in condemning rival *daos* must exist—and hence be natural. This does not entail that all *possible daos* are natural, but all existing and, no doubt, *many* that don't exist, are natural in the sense they have or might emerge in nature—for example, without any supernatural intervention.

This allows Zhuangzi to continue Daoism's trend of detachment and ironic neutrality in the fervent debates among the moralists. He also generalizes earlier themes (noted above) of how different *daos* shape our attitudes, our desires, and our descriptive language. This generalization emerges with a hint of mild skepticism. The anti-language position is an error position—all doctrines of morality/reality are false. The Mohist paradox undermines that. We cannot consistently conclude that no claims in language are correct. Neither can we claim that all linguistic utterance is correct since judgments according the standards of one natural language conflict with those of another. Each treats its own use as acceptable and others an unacceptable.

Each position in debate about *dao* depends on presuppositions—a presupposed *dao*. The *dao* is the *way* of speaking that language. An argument for one *dao* over another, also presupposes an implicit *dao*—a *way* of choosing which *dao* to follow in a situation. And even when we have selected which *dao* to follow, we may disagree about now to interpret our agreed *dao*. A way of

interpreting *this dao* is still another presupposed *dao*. When we perform any *dao*, we perform an entire hierarchy of *daos*.

Thus, winning an argument doesn't give one assurance of being right *simpliciter*—it presupposes some way of picking a winner. There is no completely *neutral* way of assigning a right and wrong. Zhuangzi's skepticism is not based on contrasting human cognitive weakness with some ideal perspective, but draws on the infinite regress of standards involved in guidance by a way, a *dao*. It is skepticism because it doesn't deny we might have acted rightly, only that we cannot know in ways that we could, correctly show others is correct (that is move as judged from their *daos*). Skepticism in the *Zhuangzi* rests on the observation that we cannot be sure we are using the right standard of "knowing."

Zhuangzi expresses another aspect of the same point. Consider the standard objection to "Ideal Observer" theories of morality (eg. Right = if what some God-like judge would judge to be right). The objection goes that what the Ideal Observer should do is irrelevant to my decision about what I, an ordinary observer, should do. The *Zhuangzi* frequently reflects on the theme of how a perfect perspective is neither useful nor comprehensible to us. What God should do is wildly irrelevant to the real practical questions that confront us. We could neither understand nor use the answer to "What would a perfect person do?"

This leads in two ways back to the naturalism characteristic of Daoism. Not only are *actual* conventions natural, but there is a natural *dao* that both guides the selection (evolution) of conventional *dao* and guides the interpretation of them. Thus, the *Zhuangzi* argues that there is no way to disentangle the realm of *tian*$^{\text{nature:sky}}$ and the realm of *ren*$^{\text{human}}$—no way to ground the claims of moralists or mystics to have found *the* single naturally correct way.

THE FATE OF DAOISM UNDER THE EMPIRE

Philosophy in China suffered a dark age initiated in the third century BC by the emergence of the imperial structure under a totalitarian ideology. Authoritarian misgivings about the tendency of philosophers to cast doubt on conventional ways of making distinctions and the assumed political goal of unifying the social world under one *dao* motivated this political authoritarism. Paradoxically, most would say the Qin favored Daoism since the superstitious rulers sought in it the "secret of long life." This, of course, served rather to replace Daoism's philosophical reflections with credulous religious dogma and the interpretation dominant in "legalist" commentary. The Qin dynastic family and its so-called "legalist" *dao* of governance lasted only one generation. However, the institutional structure survived in the Han which anointed Confucianism as the unifying *dao*.

This totalizing position led Confucianism to an eclecticism which sought to embrace everything from royal superstitions to naturalistic cosmic forces (yin and yang) and ground them in a Confucian *dao* that they could use to manipulate the ruler. Confucians assured us that this moralized cosmology "incorporated Daoism" harmoniously with Confucianism, implying that Daoism entailed the superstitious yin-yong cosmology they had worked out. Several new "Daoist" texts emerged which mixed quotations from the *Zhuangzi* and this moralized yin-yang cosmology. These further tended to shape the accepted interpretation of Daoism to better suit the bureaucracy's purpose.

The stasis of a naturalized Confucianism and a cosmologized Daoism survived with slight variation until the modern period. The main interruption was the importation of Buddhism, which continued the stable institutional model and the tendency to eclectic blending into a single harmonious officially recognized *dao*. Imperial rule tended always to patronize some "approved" eclectic ideology. In such contexts, the critical and skeptical quality of thought of the classical period never reemerged.

The fall of the Han brought with it a crisis in confidence in this cosmic Confucianism. This did little to disrupt the assumption of fundamental compatibility of Daoism and Confucianism, but did lead to greater focus on the classical Daoist philosophical texts. The result has been called Neo-Daoism, though the main thinkers identified themselves as Confucians. They inflated the cosmic interpretation of Daoism into something closer to pure metaphysics (via its explicit interest in the contrast of *being and non-being)*. The result was a holistic "round" metaphysics with *non-being* at the core (basic *non-being*) and "being" as the periphery (functionally oscillating fluctuations in the field of "*non-being*").

The main point of controversy within Neo-Daoism was whether the *non-being* was really nothing or really "something." From there the philosophical level of Daoist discourse declines more than advances. Wang Bi interpreted the *Laozi* in tandem with a cosmological divination manual, the *Yi-Jing*, into what he assumed was a single system. *Non-being* was the basis of everything, the great-ultimate that was also the non-ultimate. Wang Bi

explained the dynamic between *non-being* and *being* not as causation, but as the relation of "substance and function."

Guo Xiang interpreted the *Zhuangzi* giving us the received version and conforming to the outlines of Wang Bi's system except for his Parmenidean insistence that *non-being* simply was not. This he coupled with a radically un-Parmenidean view that *being* constantly changed "of itself." The Neo-Daoist systems were the originating models of the puzzling substance-function dualism that re-appeared regularly in most later philosophical systems, both Buddhist and Neo-Confucian, right up through modern times.

Neo-Daoist speculations on *being and non-being* helped facilitate early discussions of Buddhist philosophy—particularly the puzzle of the nature of Nirvana (and thus of the Buddha-Nature). Buddhism, however, brought with it the apparatus of monastic ecclesiastical authority that bequeathed more familiar religious structures to the existing fragments of "Daoist" superstition. The resulting religious movements are what survived the Buddhist period into modern times as what the West came to know as "Daoism."

KEY DAOIST CONCEPTS

This section explores two concepts that play a central role in Daoist philosophy—*dao*[guide] and *de*[virtuosity] (virtuosity). Together the terms have come to mean something like "ethics." We, however, take them to be the basic concepts of a broader notion of normativity. The normativity is broader because there can be a *dao* (and *de*) of language (correct way of use) of knowledge (ways to know) as well as to act.

DAO (WAY, GUIDE, ROAD). The main characteristic that justified Chinese historians in identifying a school to call Daoism is philosophical interest in the concept of *dao*. The almost universally accepted translation is a primitive of English—"way." So it subsumes "manner," "course," "technique," "system," "fashion," "custom," "style," "practice," "tradition," "discipline," "road," "direction," "path," and so forth.

A way is an answer to a "how" or "what-to-do" question. We typically use talk of ways in advising someone. Ways are thus *practical* (prescriptive or normative) concepts. A road, as a concrete (or asphalt!) example guides us and facilitates our arrival somewhere. Ways are prescriptive structures that have physical realizations. We can refer to the physical forms as ways or *daos* without thereby recommending them. The *Zhuangzi* reminds us that thievery has a *dao*. We can use both *dao* and "way" simply to describe—as when a Confucian undertakes to pursue his father's *dao* for three years after his death or we say, "I saw the way you did that."

There are interesting differences between *dao* and "way." Classical Chinese language lacks pluralization; for example, not simply has no plurals, but has no grammatical role for plurals. Most common nouns function like collective nouns, roughly analogous to plurals or mass nouns of English. So *dao* is more like "ways" or "way-stuff" or "the way-part or aspect of things" than it is like "*a* way." Like other common nouns, *dao* has a part-whole structure, that is additive—two parts simply yield a larger part of the same thing. What we describe as one way would function like one part or component of what in Chinese we call *dao*. Multiplicity in common nouns in ancient Chinese emerges via modification. So they might discuss, for example, my-*dao*, Sage-King's-*dao*, natural-*dao*, past-time's-*dao* and so forth. This feature explains why *dao* might appear more *metaphysical* than "way" and helps appreciate familiar Daoist spatial metaphors like "humans encounter each other in *dao* as fish do in water" (*Zhuangzi* Ch. 6). *Dao* is a little like the water—a feature of the realm in which humans live, work, and play. To be human is to be in a framework of ways to act, go, and speak. So-called Daoists are more likely to play with these metaphysical metaphors than Confucians or Mohists—who mainly point to (their favored part of) *dao*.

A second difference is that unlike "way," *dao* may be used as a verb. The best-known example is the famous first line of the *Dao*de Jing. Literally "*dao* can be *dao* not constant *dao*." For the middle *dao*, roughly one out of three translators uses "speak," another third use "tell" and the rest use near synonyms such as "expressed," "defined in words," or "stated." In a famous Confucian example of this use, Confucius criticizes *dao*-ing the people with laws rather than *dao*-ing them with ritual. (This verbal sense is often marked by a graphic variation *dao*[to direct].)

"Speak" is in some ways too narrow and in others too broad as a way translating this verbal use. It is too broad because in Western tradition, speaking is conventionally linked to describing, representing, picturing, expressing, defining, or "capturing" some reality. The Chinese verbal use resembles more what a European would express by "advocate" "acknowledge" or "recommend," for example, to "guide-speak." To *dao* is to put guidance into language. *Dao*-ing is giving advice.

"Speak" is, in other ways, too narrow. One can *dao* in written form or even by example—as when we *dao* with law or with ritual (texts or exemplars). Consider, again,

the concrete translation for *dao*: "road" or "path." A woodsman with an ax *daos* when he chops bark from the trees as he enters the forest; He is *dao*-ing when he is "blazing" the trail. As the *Zhuangzi* notes, "A *dao* is made by walking it."

What a road shares with a pattern of blazes in a forest is that both, like maps and verbal instructions, can serve as normative guides. What they also share, is that they require *reading* and *interpretation*. In following any kind of *dao*, we "interpret" it. This might be hard, as when we interpret blazes in bark, or piles of stones left by boy scouts or a Hansel's string of bread crumbs or the two-days-old tracks left by a deer. Or it may be relatively easy, as when I follow the asphalt ribbon between my house and the store. These examples should illustrate the symbolic guiding nature of all "roads." To interpret a road/path/*dao* is to extract guidance in the form of an actual "walking," not to develop a theory or belief. This use of "interpretation" is more familiar in artistic contexts—in music, dance, or drama. The interpretation of a score, line or character in a play consists in a performance of it.

The metaphysics of *dao* should mark this distinction between normative way *types* (treated as guides) and interpretive tokens (the result of practically interpreting a guide). The token is an actual history, a string of actions. The token may itself be *taken as a guide* (that is as an exemplary model), but in that case, it in turn requires interpretation. There are various *ways* to follow the example. We have to extrapolate from the exemplar's situation to our own. So the distinction between type and token ways can be relative; it is actually a distinction between normative and descriptive senses. When we treat a token as subject to evaluation relative to some normative *dao*-type, it is descriptive. When we treat it as a *model* guiding our own performance, it is normative.

This should help us understand the notion of natural *daos*. Roads and ways need not be human constructions. Nature's "engineers," deer or mountain goats, also make paths. Famously animals from ants to pack rats, swallows to dogs make or mark and read their own ways. Other species may read and use these as humans do when lost in the mountains. Other "ways" are pure natural possibilities of sequences that will result in attaining a goal; for example, their being a way through a forest or across a river. That way consists of their being a fallen log or several large stones in a fortuitous configuration—fortuitous, that is, from the point of view of human actors. We discover these structures in nature as we "feel our way" along. We may learn to read natural signs and exact guidance from natural clues better over time.

However, the concept of a naturally constant *dao* threatens to follow the Great *Dao* into losing its normative role. The *Zhuangzi* recognized this danger most clearly in pointing out that all recommendations (all prescribed *dao*) are *natural* in virtue of actually being advanced and promoted. Nature—the structure of natural constancy—does not select any of these. Any selection requires a *dao* interpreter and interpreters select using different standards—higher-level *daos* of selection and interpretation. Thus, while some *daos* are impossible, the appeal to nature does not adjudicate among any actual rival formulations, such as those of Confucians versus Mohists. Nature does not evaluate or prescribe its possibilities. Like the Great *dao*, they just are.

DE (VIRTUOSITY, VIRTUE, POWER). A Daoist formula for *de* is "*dao* within." Translators most commonly use "virtue" as a translation but hurry to remind us that it is "virtue" in the ancient Greek sense of an excellence. "Power" can work as an alternative translation because it reflects the link between *de* and successful action or achievement for its possessor. This author prefers "virtuosity" to capture both the sensitivity to context and *fit* and to remind us of the aesthetic features of these normative concepts. Virtuosity is the capacity of a performer to "interpret" a score-like *dao* into a *superb* performance (in that theater, for that audience, and so forth). Thus *de* is the capacity to perform *dao* correctly—successfully, beautifully, and well.

Daoist reflections on *de* sometimes point to "natural" or prelearned capacities to learn or perform some *dao* with skill. Think of Wittgenstein's talk of the unexplained human ability to catch on and continue the correct grammar of a human language. This stress sometimes suggests the hardwiring or the machine language translation required to implement or interpret other programming. While many Daoist comments may be taken to refer to such "natural" skills, the concept of *de* itself seems to include *de* that is acquired in the process of learning, internalizing, practicing, and fine-tuning our performance of some *dao*.

Natural *dao* is presupposed in learning social *daos* in a number of ways. There is a natural way humans learn to acquire and perform normative *dao* and there are different ways to perform in different natural contexts. There are both natural and social-practice ways to select which *dao*-type to execute and multiple ways to evaluate the performance-generated *dao*-tokens. So, as Zhuangzi

observes, we live so pervasively in such a "sea" of *dao* that, like fish in water, we forget that we forget *dao* as we swim around it.

See also Chinese Philosophy: Overview; Chinese Philosophy: Contemporary.

Bibliography

Alt, Wayne. "Logic and Language in the Chuang-tzu." *Asian Philosophy* 1 (1) (1991): 61–76.

Ames, Roger, ed. *Wandering at Ease in the Zhuangzi*. Buffalo, NY: SUNY Press, 1998.

Berling, Judith A. "Paths of Convergence: Interactions of Inner Alchemy, Taoism, and Neo-Confucianism." *Journal of Chinese Philosophy* 6 (1979): 123–.

Blofeld, John. *Taoism: The Quest for Immortality*. London: Mandala Books, 1979.

Carus, Paul. *The Canon of Reason and its Virtue*. Chicago: Open Court, 1913.

Chen, Ellen Marie. "Nothingness and the Mother Principle in Early Chinese Taoism." *International Philosophical Quarterly* 9 (September 1969): 391–405.

Ch'en, Ku-ying. *Lao Tzu: Text, Notes, and Comments*. Translated by Rhett Y. W. Young and Roger T. Ames. San Francisco, CA: Chinese Materials Center, 1977.

Cook, Scott. "Zhuang Zi and His Carving of the Confucian Ox." *Philosophy East and West* 47 (4) (1997): 521–554.

Creel, Hurlee G. *What Is Taoism?* Chicago: University of Chicago Press, 1970.

Cua, Antonio S. "Opposites As Complements: Reflections on the Significance of Tao." *Philosophy East and West* 31 (2) (April 1981): 123–140.

Duyvendak, J. J. L. *Tao Te Ching*. London: John Murray, 1954.

Fu, Charles Wei-hsun. "Creative Hermeneutics: Taoist Metaphysics and Heidegger." *Journal of Chinese Philosophy* 3 (1976): 115–143.

Fung, Yu-lan. *History of Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1952.

Girardot, Norman J. *Myth and Meaning in Early Taoism: The Theme of Chaos (hun-tun)*. Berkeley: University of California Press, 1983.

Graham, Angus, tr. *The Book of Lieh-tzu*. London: John Murray, 1960.

Graham, Angus. *Chuang tzu: The Inner Chapters*. London: Allen & Unwin, 1981.

Graham, Angus. "Chuang-tzu's Essay on Seeing Things as Equal." *History of Religions* 7 (1969): 137–159.

Graham, Angus. "Daoist Spontaneity and the Dichotomy of 'Is' and 'Ought.'" In *Experimental Essays on Chuang-tzu*, edited by Victor Mair, 3–23. Honolulu: University of Hawaii Press, 1983.

Graham, Angus. "The Date and Composition of the Lieh-tzu." *Asia Major* 8 (2) (1961): 139–198.

Graham, Angus. *Disputers of the Tao: Philosophical Argument in Ancient China*. La Salle, IL: Open Court, 1989.

Graham, Angus. "How Much of the Chuang-tzu Did Chuang-tzu Write?" *Journal of the American Academy of Religion* 47 (3) (1979).

Hall, David L. "Process and Anarchy: A Taoist Vision of Creativity." *Philosophy East and West* 28 (3) (July 1978): 271–285.

Hansen, Chad. *A Daoist Theory of Chinese Thought*. New York: Oxford University Press, 1992.

Hansen, Chad. "A Tao of Tao in Chuang Tzu." In *Experimental Essays on Chuang-tzu*, edited by Victor Mair, 24–55. Honolulu: University of Hawaii Press, 1983.

Hansen, Chad. "The Zhuangzi: A Historical Introduction." In *The Dao of Zhuangzi*, edited by Tsai Chih Chung, 9–22. Garden City, NY: Anchor Books, 1997.

Henricks, Robert G. *Lao-tzu: Te-Tao Ching: A New Translation Based on the Recently Discovered Ma-wang-tui Manuscripts*. New York: Ballantine, 1989.

Hsiao, Kung chuan. *A History of Chinese Political Thought*, Vol. 1: *From the Beginnings to the Six Dynasties*. Princeton, NJ: Princeton University Press, 1979.

Hu, Shih. "A Criticism of Some Recent Methods Used in Dating Lao Tzu." *Philosophy East and West* 40 (1) (1989): 17–33.

Kasulis, T. P. "The Absolute and the Relative in Taoist Philosophy." *Journal of Chinese Philosophy* 4 (1977): 383–394.

Kjellberg, Paul, and Philip J. Ivanhoe, ed. *Essays on Skepticism, Relativism, and Ethics in the Zhuangzi*. Buffalo, NY: SUNY Press, 1996.

Kohn, Livia, ed. *Taoist Meditation and Longevity Techniques*. Ann Arbor: Center for Chinese Studies Publications, University of Michigan, 1989.

Kupperman, Joel J. "Not in So Many Words: Chuang Tzu's Strategies of Communication." *Philosophy East and West* 39 (3) (July 1989): 311–317.

Lau, D. C., tr. *Lao Tzu: Tao Te Ching*. Baltimore: Penguin Books, 1963.

Lau, D. C., tr. "The Treatment of Opposites in Lao-tzu." *Bulletin of the Society for Oriental and African Studies* 21 (1958): 344–360.

Liu Xiaogan. *Classifying the Zhuangzi Chapters*. Ann Arbor: Center for Chinese Studies, University of Michigan, 1994.

Mair, Victor, ed. *Experimental Essays on Chuang-tzu*. Honolulu: University of Hawaii Press, 1983.

Mair, Victor, tr. *Tao Te Ching: The Classic Book of Integrity and the Way*. New York: Bantam Books, 1990.

Maspero, Henri, tr. *Taoism and Chinese Religion*. Amherst: University of Massachusetts Press, 1981.

Munro, Donald, J. *The Concept of Man in Early China*. Stanford, CA: Stanford University Press, 1969.

Needham, Joseph. *Science and Civilization in China*. Vol. 2: *History of Scientific Thought*. Cambridge, U.K.: Cambridge University Press, 1962.

Robinet, Isabelle. "Metamorphosis and Deliverance from the Corpse in Taoism." *History of Religions* 19 (1) (1979): 37–70.

Roth, Harold D. *The Textual History of the Huai Nanzi*. Ann Arbor, MI: Association of Asian Studies, 1992.

Roth, Harold D. "Who Compiled the Chuang Tzu?" *Chinese Texts and Philosophical Contexts*, edited by Henry Rosemont, 84–95. La Salle, IL: Open Court, 1991.

Saso, Michael. *Taoism and the Rite of Cosmic Renewal*. Pullman: Washington State University Press, 1972.

Schwartz, Benjamin. *The World of Thought in Ancient China*. Cambridge, MA: Harvard University Press, 1985.

Sivin, N. "On the Word 'Taoist' As a Source of Perplexity, with Special Reference to the Relations of Science and Religion in Traditional China." *History of Religions* 17 (3–4) (February/May 1978): 303–330.

Smullyan, Raymond. *The Tao Is Silent*. New York: Harper and Row, 1977.

T'ang, Chün-i. "Cosmologies in Ancient Chinese Philosophy." *Chinese Studies in Philosophy* 5 (1) (Fall 1973): 4–47.

Van Norden, Bryan V. "Competing Interpretations of the Inner Chapters of the 'Zhuangzi.'" *Philosophy East and West* 46 (2) (1996): 247–269.

Waley, Arthur, tr. *The Way and Its Power: A Study of the Tao Te Ching and its Place in Chinese Thought*. London: Allen & Unwin, 1934.

Watson, Burton. *The Complete Works of Chuang Tzu*. Records of Civilization: Sources and Studies No. LXXX, Columbia College Program of Translations from the Oriental Classics. New York: Columbia University Press, 1968.

Watts, Alan Wilson. "The Philosophy of the Tao." In *The Way of Zen*, 23–48. Harmondsworth: Penguin Books, 1962.

Watts, Alan Wilson. *Tao: The Watercourse Way*. Harmondsworth: Penguin Books, 1975.

Welch, Holmes. *Taoism: The Parting of the Way*. Boston: Beacon Press, 1966.

Welch, Holmes, and Anna Seidel, eds. *Facets of Taoism*. New Haven, CT: Yale University Press, 1979.

Wong, David. "Taoism and the Problem of Equal Respect." *Journal of Chinese Philosophy* 11 (1984): 165–183.

Wu, Kuang-ming. *Chuang Tzu: World Philosopher at Play*. New York: Crossroad Publishing Company, 1982.

Yearley, Lee. "The Perfected Person in the Radical Chuang-tzu." In *Experimental Essays on Chuang-tzu*, edited by Victor Mair. Honolulu: University of Hawaii Press, 1983.

Yearley, Lee. "Zhuangzi's Understanding of Skillfulness and the Ultimate Spiritual State." In *Essays on Skepticism, Relativism, and Ethics in the Zhuangzi*, edited by Paul Kjellberg and Philip J. Ivanhoe, 152–182. Buffalo, NY: SUNY Press, 1996.

Zhuangzi. "The Zhuangzi." Citations (HY) from the Harvard Yenching Institute Sinological Index Series, Supplement no. 20. Cambridge, MA: Harvard University Press, 1956.

Chad Hansen (2005)

ETHICS

The first recorded dynasty in Chinese history is the Shang (1766–1050 BCE). It came to an end when the Zhou family overthrew the Shang and justified its act on the grounds that the Shang kings had become corrupt and forfeited the right to rule conferred by the ruling force of the world, *tian* (which literally means "sky" and is usually translated as "Heaven"). Although the Zhou kings claimed validation by Heaven, their rule declined in the time of Kongzi (551–479 BCE; better known in the West by his Latinized name Confucius), entailing a breakdown of the social, political, and moral order. The dao, the way or path, that the Zhou claimed to possess was lost. As Angus C. Graham (1989) puts it, the primary question of the age was: Where is the dao? Whoever could rediscover it could regain the *de*, the human power and excellence, that enabled the early Zhou kings to create the golden age of harmony and flourishing that was lost. Each philosophy of the ancient period provides its version of the dao.

CONFUCIAN ETHICS

Benjamin Isadore Schwartz (1985) characterizes the Confucian dao as emphasizing respect for rightful authority, where the rightfulness of authority is based on the achievement of ethical excellence. Confucianism is a virtue ethic because of its central focus on three interrelated subjects: character traits identified as the virtues; the good and worthwhile life; and contextualist modes of ethical deliberation. The virtues are traits of character that are necessary for living a good life and that typically involve judging in the context at hand what must be done. The virtues belong to the *junzi* (the noble person, most often translated as "gentleman"), who is living in accord with the dao. Such a person can be said to have realized in a high degree the overall ethical excellence that befits human beings.

Consider the virtue of *ren*. In 12.22 Kongzi identifies *ren* with loving or caring for people. Understood in this way, *ren* is one particular virtue among a number possessed by the *junzi* such as wisdom and courage. Translators such as D. C. Lau (1979) focus on this meaning of *ren* and translate it as "benevolence." However, *ren* has a much broader meaning in the *Analects* (but perhaps not in the *Mengzi*, where it seems restricted to the meaning of a particular virtue). At a number of places, *ren* is associated with an array of different virtues: for example, the observance of ritual in 3.3, and sympathetic understanding of others in 6.30. In fact, *ren* seems so closely associated with the ideal of the *junzi*, or morally noble person, that it seems to stand for complete human excellence.

A virtue distinctive of the Confucian ethic is that of ritual observance or ritual propriety. In the *Analects* the rituals (*li*) include ceremonies of ancestor worship, the burial of parents, and the rules governing respectful and appropriate behavior between parents and children. In general, the *li* in the *Analects* are ceremonies or customary practices that express one or more of several ethically significant attitudes: reverence, respect, care, gratitude, and a feeling of indebtedness. Later, the word came to cover a broad range of customs and practices that spelled out courteous and respectful behavior between people occu-

pying specific social stations. Herbert Fingarette (1972) argues that this emphasis on ritual propriety conveys the profound insight that ceremonies, customs, and conventions constitute much of what is distinctive about human activity. A handshake means nothing unless understood against a background of conventions that establish the relevant physical movements as a way of greeting another person. So too, many of the ways of respecting and expressing care or gratitude toward others are possible only because they have been conventionally established as ways to express those attitudes. This implies that *ren* as complete human excellence cannot be understood as something separate and independent from *li*.

As Kwong-loi Shun (1993) points out, however, this does not mean that *ren* reduces to any given set of practices adopted by a community, since alternative practices in a different community might be devised to express the same ethical attitude or there might be nonconventional means of expressing that attitude. *Li* also are portrayed in the *Analects* as crucial for the project of ethical self-cultivation: dedicated observance of the rites, along with a sincere commitment to have the appropriate attitudes they are conventionally established to express, are crucial for developing and strengthening the dispositions to have those attitudes. Sincerely engaging in a ritual that is conventionally established to express reverence for parents or ancestors makes stronger the disposition to revere. Finally, there is an aesthetic dimension to the ethical importance of *li*. One is more or less graceful and elegant in the performance of *li*. One has made such observance more or less a second nature, from which it flows effortlessly and spontaneously. One who is so accomplished lives a life of beauty, and this is part of the *junzi* ideal.

The concept of *yi* refers both to that which is right or appropriate for the given situation and to the trait of character that consists in reliably identifying and acting on what is right. The *Analects* 4:10 says that the *junzi* is not predisposed to be for or against anything, but goes with what is *yi*. As Antonio Cua (1998) points out, traditional rules of ritual propriety provide one with a sense of what is courteous and respectful action given standard contexts, while the virtue of *yi* allows one to identify when those rules need to be set aside in exigent circumstances. In 4A17 Mengzi (371–289 BCE; better known in the West as Mencius) observes that to save the life of one's drowning sister-in-law one must suspend the customary rule of propriety prohibiting the touching of man and woman when they are giving and receiving. When his interlocutor wants to apply this idea of suspending the usual rules of propriety to save the entire country from drowning, Mengzi replies that one saves one's sister-in-law with one's hand but cannot save the country from drowning in chaos and corruption with one's hand. The country can only be pulled out by the dao.

This passage not only illustrates that one may have to set aside customary rules of propriety in exigent circumstances but also that analogy is a way to judge what is *yi* for the situation at hand. One starts with a case where the judgment seems right (touching the sister-in-law when she is drowning) and attempts to transfer a like judgment to a like situation. One can also criticize an analogy by pointing out a relevant unlikeness between the two cases. One cannot save the country through violations of ritual propriety, but only through setting it back on the dao, which itself may require one to observe propriety on many occasions.

The term *de* is used to refer to the power or excellence that a thing can achieve when it acts according to the dao for things of its kind. One who gets or attains the dao and achieves virtues such as *yi* achieves a power or excellence appropriate to things of one's kind. In the *Analects* the human *de* that is *ren* brings with it a power to influence and attract other people and even the surrounding environment. Human *de* is a kind of moral charisma that comes with the achievement of the virtues just discussed and can be possessed by any good human being, but when it appears in rulers, it allows them to command others without appealing to physical threats (as Edward Slingerland [2003] points out, such an ideal may manifest the theme of *wu-wei*, or effortless action, which is traditionally associated with the daoists). The Confucian prescription for bringing China back to the dao is partly based on this belief in *de* as possessed by rulers and in the strategy of Confucian scholars offering their moral advice to rulers.

Western interest in Confucianism rests on reasons that are similar to those underlying the enduring interest in Western virtue ethics. Those who are skeptical of the modernist ambition to construct ethical theories around general principles of action, seeing them as too vague and abstract to provide much guidance on the one hand, and too reductivist to capture the rich array of ethical considerations on the other hand, turn to the ancient ideals of good character and judgment that are sensitive to context. Virtue ethics also tend to embody the theme that the ethical life of right (and in the case of Chinese and contemporary Western virtue ethics) caring relationship to others is necessary for human flourishing. In the *Mengzi* this theme emerges in identification of the distinctively

human potentials with the incipient tendencies to develop the moral virtues.

At the same time, Confucian ethics is distinctive for the centrality it gives to family life in its conception of the good life. Part of the reason for this centrality lies in the Confucian appreciation for the family as the first arena in which care, respect, and deference to legitimate authority is learned (*Analects* 1.2). The way in which particularist reasoning is illustrated in historical stories such as those about Shun is also a distinctive feature of Confucian ethics. These stories present paradigms of good judgment and of good individuals, from which persons engaged in ethical self-cultivation should learn through analogy with relevantly similar situations in their own lives. Those who hold that much moral learning and reflection is accomplished through the telling of and listening to stories and other narratives have reason to study Confucianism. Another distinctive feature of Confucian ethics, as mentioned earlier, is the emphasis it gives to rituals as providing much of the distinctive substance of human life, as a necessary dimension of moral self-cultivation, and as contributing to the aesthetic dimension of the good life.

Mengzi and Xunzi (313–238 BCE) engaged in a vigorous, provocative debate over human nature and whether there are natural tendencies that form the basis for development of a good person. Mengzi holds that there are innate moral concepts that infuse intuitive judgments and feelings (people spontaneously feel compassion for a child about to fall into a well; a beggar knows intuitively to reject food that has been thrown on the ground and trampled on). Xunzi holds that human nature is dominated by the desire for gain and sensual gratification and that rather than having a natural basis, morality is invented to prevent the destructive conflict caused by people acting from their natures.

The contrasting ways in which Mengzi and Xunzi portray moral development raise important issues about the relation between reasoning, feeling, and moral judgment. Mengzi tends to portray moral perception, reasoning, and feeling as working in concert in ways that call into question any strict separation between perception and reasoning on the one hand and feelings such as shame and compassion on the other (Wong 1991, 2002). By contrast, Xunzi holds that the mind has the power to shape and retrain its desires and feelings, but his portrait of moral development seems to presuppose appropriate conative and affective elements that form the base for such reshaping (Van Norden 1992, Wong 1996). Taking all these distinctive features together, it is fair to say that Confucianism offers an especially rich moral psychology that Antonio Cua (1998, 2005), Philip J. Ivanhoe (2000), David S. Nivison (1996), and Shun (1997) illuminate.

Another respect in which Confucianism differs from modern Western moral theories bears especially on the cross-cultural comparison of values. Confucian morality lacks a focus comparable to that found in modern Western moralities on individual rights to liberty and to other goods, where the basis for attributing such rights to persons lies in a moral worth attributed to each individual independently of what conduces to individual's responsibilities to self and others. Confucianism rather assumes that the ethical life of responsibility to others and individual flourishing are inextricably intertwined, and in such a way, Craig Ihara (2004) argues, that the individual's dignity is honored without resort to the concept of rights. A frequent Western interpretation of Confucian ethics is that it subordinates the individual to the group. David L. Hall and Roger T. Ames (1998) respond that this interpretation erroneously presupposes that the individual and community are potentially at odds in ways the Chinese tradition does not conceive them to be. The nature of the individual is conceived relationally, they argue, so that it is just plain wrong to have the Chinese separating the individual from the group in the first place, much less subordinating the individual to the group.

Another frequent criticism from the Western side is that Confucianism fails to provide adequate protection to those legitimate interests an individual has that may conflict with community interests. On the other side, Henry Rosemont (1991) criticizes rights-focused moralities for ignoring the social nature of human beings and of portraying human life in an excessively atomistic fashion. Against those who argue that Confucianism does not sufficiently protect the individual, Rosemont (2004) replies that the Confucian framework of responsibilities to others can afford significant economic and social protections to the individual and arguably addresses the human need for community and belonging better than rights frameworks.

Moreover, it is arguable that rights in some sense can play a role in the Confucian tradition, even if such rights are not grounded in the idea of the independent moral worth of the autonomous individual. Within that tradition, Joseph Chan (1999) argues, rights might function to protect individuals' interests when the right relationships of care irretrievably break down. Furthermore, rights to be protected in one's speech can receive a Confucian justification as conducive to the health of the community. Mengzi advised kings to attach more weight to the opin-

ions of their people than to those of their ministers and officers in making certain crucial decisions. Xunzi recognized the need for subordinates to speak their views freely to their superiors. If one carries the reasoning in Mencius and Xunzi one step further, one sees the need to protect a space in which they may speak freely without fear of suppression, and hence derive a right in the "thin" sense of what one has whenever one has justifiable claims on others to assure one's possession of things or one's exercise of certain capacities (Wong 2004).

That there are developments of each tradition that bring each closer to the other may suggest that each could learn from the other. One might worry about the kind of individualism that prompts citizens in affluent nations such as the United States to tolerate gross inequality of opportunity. One therefore might look to a tradition that appreciates the way people thrive or falter within specific communities that nurture or shut them out. On the other side, a tradition that has tended to value the idea of social harmony at the cost of sufficiently protecting dissenters pointing out abuses of power or just plain bad judgment by authorities would do well to look at enduring traditions that do not value social harmony as highly.

MOHIST ETHICS

Mozi (470–391 BCE) is said to have begun as a student of Confucianism and eventually came to reject it in favor of a consequentialism that in important respects anticipates Western utilitarianism by two millennia. While Confucians saw the problem with China as loss of respect for authority and a related loss of moral basis for authority, Mozi saw the problem as partiality. Heads of families knew only to love their own families and mobilized their families to usurp others. Lords knew only to love their own states and consequently mobilize their own to attack others. Such partiality causes destructive conflict that harms everyone, so the proper conclusion is to override one's own tendencies to partiality and to practice *jianai*, sometimes translated as "universal love" but arguably better translated as "impartial concern."

Schwartz (1985) points out that *ai* in the *Mozi* means neither *Eros* nor *agape* but something closer to a concern for all that is justified on the basis that its practice advances one's own welfare and the welfare of those to whom one is partial. The doctrine of impartial concern, when combined with Mozi's emphasis on evaluating beliefs according to the benefits and harms that result from them, qualifies him as a kind of utilitarian. His is not, however, a hedonistic or welfare utilitarianism of the kind most commonly represented in the Western tradition. His conception of benefit and harm refers to no psychological goods and harms such as pleasure and pain but exclusively to material goods and harms such as enriching the poor, increasing the population, and bringing about order.

Because he advocated impartial concern, Mozi had no use for the Confucian doctrine of graded concern: that the degree of one's concern should depend on the nature of one's relationship to the person in question (one's family being owed the most concern). Because he relied on pragmatic appeals to people's existing interests to justify his moral position and because he took the rationalist position that people should have no trouble doing what they see to be in their interests, he saw no use for Confucian ritual as a mode of moral self-cultivation.

The traditional attitude toward Mozi is that he was a relatively minor philosopher, but that is changing. His criticism of Confucian-graded concern and his advocacy of impartial concern is of broad interest and raises the question of how to fit within a coherent moral framework the special concern parents and children ought to have toward one another with the universal and equal concern one ought to have toward all persons as persons (Wong 1989). Moreover, Chad Hansen (1992) argues persuasively that Mozi's vigorous argumentation against the Confucians constituted a pivotal point, after which subsequent Confucian thinkers such as Mengzi and Xunzi had to defend Confucianism with argument. Mozi was unique in developing explicit standards for argumentation, and his school developed a distinctive focus on questions of logic, argumentation, and philosophy of language.

DAOIST ETHICS

The two great daoist texts of the ancient period are the *Daodejing* (Book of the way and its power; traditionally but dubiously attributed to the historical figure Laozi, c. sixth century BCE), and the *Zhuangzi* (a good part of the first seven chapters, the so-called inner chapters, was probably written by the historical figure Zhuangzi, c. 360 BCE). It may seem paradoxical to write of daoist ethics, but daoism thrives on (apparent?) paradox. On the one hand daoism expresses strong skepticism about distinctions between good and bad, right and wrong. On the other hand it also makes recommendations that add up to putting forward a way of life. Joel J. Kupperman (1999) observes that the *Zhuangzi* commends a way of life that does not take oneself and one's ideas so seriously.

The way of life commended in the *Zhuangzi* also includes openness to what might escape one's current

conceptualizations and preconceptions. One is invited to see that one's conceptualizations of the world are inevitably incomplete and distorting. One attempts to order the world by sorting its features under pairs of opposites, but opposites in the real world never match neatly with one's conceptual opposites. Real "opposites" escape one's attempts to cleanly separate them. Despite one's best efforts, they switch places in one's conceptual maps, blur, and merge into one another. That is why chapter 2 of the *Zhuangzi* says that the sage recognizes a *this*, but a *this* that is also *that*, a *that* that is also *this*. In chapter 5, men who have had their feet amputated as criminal punishment are scorned by society, but not by their daoist masters, who see what is of worth in them. In fact, both the *Zhuangzi* and *Daodejing* express an underlying suspicion of the needs that evaluative judgments serve; it is precisely to dominate or to undermine others that one subsumes them under the disfavored halves of one's value dichotomies.

The *Zhuangzi* further emphasizes the need to accept the inevitable in human life, the need to manage one's desires to achieve tranquility in the face of the inevitable, and to identify with the world that makes acceptance and management of desires possible. Both the *Zhuangzi* and *Daodejing* commend *wu-wei*, literally translated as "non-action," but meaning something like unforced acting with the grain of things. It is a style of action that consists in being receptive rather than aggressive, following from behind rather than leading in front, accommodating rather than confrontational, and being flexible and ready to change with the situation rather than rigid and operating from general predetermined principles. Seeing what is of worth in people and getting attuned to the grain of things are themes that stand in tension with the skepticism expressed by both the *Zhuangzi* and *Daodejing*, and one of the central interpretative problems is how to reconcile them (Hansen 1992, Kjellberg and Ivanhoe 1996, Wong 2005).

The *Zhuangzi* addresses such recommendations largely to the private individual who has become disaffected with the popular striving after conventional success and with the earnest moral idealism of the Confucians. By contrast, the *Daodejing* often addresses its recommendations to rulers, and even when it does not it expresses a primitivist social philosophy that holds that humanity was at its best when its desires were the fewest and when it did not guide itself through self-conscious valuing. Chapter 19, for example, says, "Exterminate the sage, discard the wise, and the people will benefit a hundredfold; exterminate benevolence, discard rectitude, and the people again will be filial; exterminate ingenuity, discard profit, and there will be no more thieves and bandits" (Lau 1985, p. 23). The rejection of conventional success and earnest idealism is here paired with the promise that if one stops trying to impose one's will on others (along with the usual value dichotomies) one may actually result in the ends one originally hoped to achieve.

What is interesting about *wu-wei* as applied to political leadership, as Michael Lafargue (1992) points out, is that it implies an organic notion of social harmony that contrasts with the conception of harmony as imposed by a dominating person who stands out from the rest of the group. A leader in an organic social group models the kind of self-effacement and sparseness of desire that all members should have. One suspects that such a leader must do more than model to be effective, but the *Daodejing* does not dispense specific advice or strategies. It rather provides metaphors from nature about the strength to be found in water and in valleys, associated with the female, that can overmatch the strength to be found in rock and in mountains, associated with the male.

In the *Daodejing* both the skepticism about the adequacy of conceptual structures and the confidence in *wu-wei* have traditionally been thought to be rooted in a monistic vision of the universe that is centered on the notion of the dao. Consider chapter 4 of that text where the dao is described as being empty, as seeming something like the ancestor of the myriad of things, as appearing to precede the Lord (*di*). In chapter 1, the constant dao is characterized as nameless, and the nameless is the origin of Heaven and Earth. Insofar as it is named, one could call it the mother of all things. The dao of the *Daodejing* might be the indeterminate ground in which determinate things are incipient, as suggested by Robert Neville (1989). Chung-ying Cheng (1989) suggests that the embrace of an indeterminate ground of the determinate may reflect the decision to give the phenomenon of change a fundamental place in ontology, rather than an absolutely stable being as in Parmenidean ontology and as later reflected in Aristotelian and Cartesian notions of substance. One reason for a continuing Western interest in Chinese metaphysics has partly been fueled by the perception that contemporary physics has undermined the strategy of giving determinate being ontological primacy.

BUDDHIST ETHICS

In Chinese strains of Buddhism, especially Chan Buddhist texts such as the *Liuzu tanjing* (Platform sutra of the

sixth patriarch) by Huineng (638–713), there is also a sense that evaluative categories cannot reliably order the world and the confidence that one can become attuned to the world so as to move with its grain. This is not surprising since daoism profoundly influenced Buddhism on its importation into China. However, Buddhist ethics is distinguished by its special emphasis on the elimination of suffering and on the way it explains suffering by referring to the human attachment to the self as a fixed ego entity. The Buddhist scripture, *The Questions of King Milinda* (Conze 1959), articulates a view of the self as based on nothing more than a floating collection of various psychophysical reactions and responses. Contrary to the folk belief, there is no fixed center or relatively unchanging ego entity. One's bodily attributes, various feelings, perceptions, ideas, wishes, dreams, and in general a consciousness of the world display a constant interplay and interconnection that leads one to believe that there is some definite *I* that underlies and is independent of the ever-shifting series, but there is only the interacting and interconnected series.

In Buddhism, this view of the self has deep practical implications. It points toward the answer to human suffering, which ultimately stems from a concern for the existence and pleasures and pains of the kind of self that never existed in the first place. The recognition that none of the "things" of ordinary life are fixed and separate entities, anymore than the self is, leads to recognizing all of life as an interdependent whole and to the practical attitude of compassion for all of life. One can only be struck by the similarity between the Buddhist view of the self and David Hume's doubts in *The Treatise of Human Nature* about the existence of a unitary and stable self. Such a conception of the self may lay claim to one's renewed attention because it fits better with a naturalized conception of human beings as part of this world and not as Cartesian-thinking substances that somehow operate apart from the rest of nature. Consider also Derek Parfit's (1984) argument that acceptance of a Humean or a Buddhist view of the self can lead to sense that one is less separate from other selves and to a wider concern when one's projects seem not so absolutely different from other people's projects. Some might see Buddhist impersonal concern as unreasonably demanding of human beings who are so strongly partial to themselves and their own (a criticism made of utilitarianism also), but as Owen Flanagan (1991) argues, that Buddhism is a vibrant and long-lived tradition with many committed practitioners provides some support for the viability of impersonal concern as an ideal that is capable of claiming allegiance and influencing how people try to live their lives.

Another concern some have about Buddhist ethics is that it appears to advocate a dampening of desire and attachment to things and people. Attachment and clinging to the impermanent is deemed the root of suffering. There is a similar vein of thought in daoism, but combined with a more complex attitude that allows attachments to remain in a transformed state, allowing one to accept the death of a loved one as part of the process of change that one embraces and even celebrates. Chapter 18 of the *Zhuangzi* portrays its namesake as sobbing on the death of his wife, but stopping and even turning to drumming on a pot and singing after he realizes that his wife has gone to become a companion to spring, summer, autumn, and winter. This more complex attitude also surfaces in Buddhism, and not surprisingly in Chinese versions of it such as Chan (later becoming Zen in Japan), where it is stressed that enlightenment is to be found in the ordinary, in one's life here and now, not in a rejection of or escape from this life.

NEO-CONFUCIAN ETHICS

The neo-Confucian Zhuxi (1130–1200) reinterpreted ethical themes inherited from the classical thinkers and grounded them in a cosmology and metaphysics, partly as a response to the growing influence of daoism and Buddhism in his time. The dao or way of Heaven is expressed in principle (*li*, not to be confused with the *li* that means ritual propriety). It is embedded in something like the indeterminate ground of the daoists, but results in the myriad of determinate things when it is sheathed in *qi*, the material energy stuff of the universe. This sheathing, however, also results in base emotions and conflict. The task of human beings is to return to their own original goodness through purification of *qi* so that *li* can be expressed, an idea that is similar to the Buddhist theme that the Buddha nature is present in all things and that enlightenment is attained through purification of that nature. Another great neo-Confucian, Wang Yangming (1472–1529), seems more pragmatic than metaphysical. He taught of the sage who formed one body with Heaven and Earth and the myriad things, but he showed little of Zhu's interest in the *li* or principle of existent things, focusing rather on the rectification of the base thoughts of the mind.

THE LEGALIST CRITIQUE OF AN ETHICS-BASED APPROACH TO GOVERNMENT

Confucian, daoist, and Buddhist ethics recommend in one way or another the project of self-cultivation result-

ing in significant self-transformation, even if the basis of such transformation is present in human nature. In Confucian ethics and in some versions of daoist and Buddhist ethics, this transformation can result in the ethical transformation of a whole society. The legalist Hanfeizi (281–233 BCE) expresses skepticism about the ambitions of such projects, and in particular the Confucian project of bringing a society back to the dao through the ethical self-cultivation of the ruling elite. Hanfeizi argues that widespread good behavior, never mind the right motivations, is an achievement requiring fortuitous circumstances. He does not dispute the Confucian belief that the sage-kings of ancient times were virtuous and ruled over a harmonious and prosperous society. He does dispute that their virtue was the primary cause. What about, he asks, those kings in more recent times who were *ren* and *yi*, benevolent and righteous, and who got wiped out for their trouble? Virtue is not the explanation of success or failure. The explanation, argues Hanfeizi, has much more to do with the scarcity of goods in relation to the number of people.

Hanfeizi's subsequent emphasis on authority, on clear and consistent law, backed by severe punishment for its violation, is designed not to provide an alternative method of making the people follow the dao, but first and foremost to prevent the worst things from happening, the worst forms of chaos, bloodshed, and human misery. Legalism is commonly regarded as a philosophy of pure Realpolitik, but is perhaps better conceived as an ethic and political philosophy that is shaped by a severe pessimism about human nature and about the practicality of moral idealism.

SOME METHODOLOGICAL ISSUES

A common Western perception of Chinese ethical teaching is that it is "wisdom" literature, composed primarily of stories and sayings designed to move the audience to adopt a way of life or to confirm its adoption of that way of life. By contrast, Western ethical philosophy is systematic argumentation and theory. One reason to think there is such a difference is the fairly widespread wariness in Chinese philosophy of a discursive rationality that operates by deduction of conclusions about the particular from high-level generalizations. Confucians seem more willing than daoists to articulate their teachings in the form of principles, but in accordance with the conception of *yi* as action that is right for the circumstances at hand, such principles seem to function as designators of values or general considerations that ought to be given weight in judgments about what to do. Never lost is recognition of the necessity for the exercise of discretion in judgment according to the particular circumstances at hand. However, such contextualist themes appear in Western philosophical traditions, beginning with Aristotle. Perhaps it is fairest to say that the Chinese and Western traditions have differed over the emphasis and relative dominance accorded to particularism versus top-down normative theorizing.

Arne Naes and Alastair Hanay (1972) characterize Chinese philosophy as invitational in its method of persuasion, meaning that it portrays a way of life in a vivid fashion so as to invite the audience to consider its adoption. The *Analects*, for example, portrays the ideal of the *junzi* as realized by persons of genuine substance who are undisturbed by the failure of others to recognize their merits (1.1: "To be unrecognized by others yet not complain, is this not the mark of the junzi?"). In the *Mengzi* 2A2, such a person possesses a kind of equanimity or heart that is unperturbed by the prospects of fame and success. This unperturbed heart corresponds to the cultivation of one's *qi* (vital energies) by uprightness.

One might be able to see such passages as appealing to experiences the audience might have in its encounters with persons who do seem to possess special strength, substance, and tranquility through identification with and commitment to a cause they perceive to be far greater than themselves. One need not interpret such sayings as attempting to persuade by the pure emotive effect of certain words, as in propaganda. Rather, they may correspond to a way of doing philosophy that attempts to say something about values in life that can be supported by experience, even if not all testimony will agree (Kupperman 1999). The daoists recommend a way of life that they explicitly characterize as one that cannot be argued for, but their recommendation receives some support through commonly shared experience.

Consider again the notion of *wu-wei* and its illustrations in the *Zhuangzi* through stories of exemplary craft. Most famously, Zhungzi's Cook Ding cuts up an ox so smoothly and effortlessly that his knife never dulls, as if he is doing a dance with his knife as it zips through the spaces between the joints. He does this not through "perception and understanding" but through *qi*, the vital energies of the body. Suggested here is a portrait of acting in the world that consists of complete and full attention to present circumstances so that the agent can act with the grain of things (the Cook Ding passage refers to *tianli* or heavenly patterns). Such a portrait does resonate with the actual experience of craftspeople, artists, athletes, musicians, and dancers who have advanced beyond self-

conscious technique and rule-following, who become fully absorbed in the experience of working with the material, the instruments, or in the movement of their bodies and who experience their actions as an effortless flow and in fact perform at high levels. In such ways, Chinese thinkers draw a picture of the world that must in the end be evaluated by explanatory power in some broad sense. One must ask whether the picture helps make sense of one's experience of the world (again in a broad sense of experience not limited to quantifiable observations in replicable experiments) and whether it preserves features of that experience that one thinks are *prima facie* genuine.

The contrast between Chinese philosophy as invitational and Western philosophy as argumentative has some truth in it, but the difference is more a matter of degree than an absolute contrast. It was Aristotle in the *Nicomachean Ethics*, after all, who said that discussions about the good in human life cannot be properly assimilated by the young because they do not have enough experience of life. And Plato, despite his insistence on the centrality of argumentation to philosophy, dispatches the short analytical arguments presented in book 1 of the *Republic* in favor of long expository portraits of the ideal city-state and the harmonious soul for the rest of that work, often presenting little or no argument for some of his most crucial claims. Other of his claims, about the divisive effects of family loyalties and the ill effects of democracy, obviously appeal to experience, even if not all testimony will agree. Furthermore, as noted earlier, Mozi's criticism of Confucians required response in kind. Shun (1997) reveals the extensive argumentative context behind Mengzi's response to the Mohists. Methods of argumentation reach their most sophisticated state of development in Xunzi (see Cua 1985), who vigorously criticizes Mozi's, Zhuangzi's, and Mengzi's theory of human nature.

Differences in the way philosophy is done may reflect differences in the interests philosophy is meant to satisfy. Hansen (1992) argues that the classical Chinese thinkers did not conceive of the primary function of language to be descriptive and as attempting to match propositions with states of affairs, but as a pragmatic instrument for guiding behavior. Western interpreters have been unable to see this, argues Hansen, because they have imposed their own concerns with correspondence truth and metaphysics on the Chinese tradition. One result, in his view, is the wrong-headed interpretation of daoism as founded on the mystical doctrine of attunement to a metaphysically absolute dao. Hall and Ames (1987) criticize Fingarette's (1972) influential interpretation of Confucius's dao as an ideal normative order transcending the contingencies of time, place, history, and culture. Hall and Ames argue Confucius's dao was not conceived as a tradition and language-independent reality against which linguistically formulated beliefs were to be measured as reliable or unreliable, but in fact a cumulative creation of individuals working from within a context provided by a society's tradition, consisting of customs, conventions, conceptions of proper behavior and good manners, and conceptions of right conduct and of what is of ultimate value and of what lives are worth living. Such controversies indicate the continuing vibrancy of the Chinese philosophical tradition as it interacts with the West.

See also Chinese Philosophy: Religion; Chinese Philosophy: Social and Political Thought.

Bibliography

Chan, Joseph. "A Confucian Perspective on Human Rights for Contemporary China." In *The East Asian Challenge for Human Rights*, edited by Joanne R. Bauer and Daniel A. Bell. Cambridge, U.K.: Cambridge University Press, 1999.

Cheng, Chung-ying. "Chinese Metaphysics as Non-metaphysics: Confucian and Daoist Insights into the Nature of Reality." In *Understanding the Chinese Mind: The Philosophical Roots*, edited by Robert E. Allinson. Hong Kong: Oxford University Press, 1989.

Conze, Edward, tr. *Buddhist Scriptures*. London: Penguin, 1959.

Cua, Antonio. *Ethical Argumentation: A Study in Hsün Tzu's Moral Epistemology*. Honolulu: University of Hawaii Press, 1985.

Cua, Antonio. *Human Nature, Ritual, and History: Studies in Xunzi and Chinese Philosophy*. Washington, DC: Catholic University of America Press, 2005

Cua, Antonio. *Moral Vision and Tradition: Essays in Chinese Ethics*. Washington, DC: Catholic University of America Press, 1998.

Fingarette, Herbert. *Confucius: The Secular as Sacred*. New York: Harper Torchbooks, 1972.

Flanagan, Owen. *Varieties of Moral Personality*. Cambridge, MA: Harvard University Press, 1991.

Graham, Angus C., tr. *Chuang-Tzu: The Inner Chapters*. Indianapolis, IN: Hackett, 2001.

Graham, Angus C. *Disputers of the Tao: Philosophical Argument in Ancient China*. La Salle, IL: Open Court, 1989.

Hall, David L., and Roger T. Ames. *Thinking from the Han: Self, Truth, and Transcendence in Chinese and Western Culture*. Albany: SUNY Press, 1998.

Hall, David L., and Roger T. Ames. *Thinking through Confucius*. Albany: SUNY Press, 1987.

Hansen, Chad. *A Daoist Theory of Chinese Thought: A Philosophical Interpretation*. New York: Oxford University Press, 1992.

Ihara, Craig. "Are Claim Rights Necessary? A Confucian Perspective." In *Confucian Ethics: A Comparative Study of Self, Autonomy, and Community*, edited by Kwong-loi Shun

and David B. Wong. New York: Cambridge University Press, 2004.

Ivanhoe, Philip J. *Confucian Moral Self Cultivation*. 2nd ed. Indianapolis, IN: Hackett, 2000.

Kjellberg, Paul, and Philip J. Ivanhoe. *Essays on Skepticism, Relativism, and Ethics in the Zhuangzi*. Buffalo: SUNY Press, 1996.

Knoblock, John, tr. *Xunzi: A Translation and Study of the Complete Works*. 3 vols. Stanford, CA: Stanford University Press, 1988.

Kupperman, Joel J. *Learning from Asian Philosophy*. New York: Oxford University Press, 1999.

Lafargue, Michael. *The Tao of the Tao Te Ching*. Albany: SUNY Press, 1992.

Lau, D. C., tr. *Confucius: The Analects*. New York: Penguin, 1979.

Lau, D. C., tr. *Tao Te Ching*. New York: Penguin, 1985.

Naes, Arne and Alastair Hanay. *Invitation to Chinese Philosophy: Eight Studies*. Oslo: Universitetsforlaget, 1972.

Neville, Robert. "The Chinese Case in a Philosophy of World Religions." In *Understanding the Chinese Mind: The Philosophical Roots*, edited by Robert E. Allinson. Hong Kong: Oxford University Press, 1989.

Nivison, David S. *The Ways of Confucianism*, edited by Bryan Van Nordan. La Salle, IL: Open Court, 1996.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Clarendon Press, 1984.

Rosemont, Henry. *A Chinese Mirror: Moral Reflections on Political Economy and Society*. La Salle, IL: Open Court, 1991.

Rosemont, Henry. "Whose Democracy? Which Rights? A Confucian Critique of Modern Western Liberalism." In *Confucian Ethics: A Comparative Study of Self, Autonomy, and Community*, edited by Kwong-loi Shun and David B. Wong. New York: Cambridge University Press, 2004.

Schwartz, Benjamin Isadore. *The World of Thought in Ancient China*. Cambridge, MA: Belknap Press, 1985.

Shun, Kwong-loi. "Jen and Li in the Analects." *Philosophy East and West* 43 (1993): 457–479.

Shun, Kwong-loi. *Mencius and Early Chinese Thought*. Stanford, CA: Stanford University Press, 1997.

Slingerland, Edward. *Effortless Action: Wu-Wei as Conceptual Metaphor and Spiritual Ideal in Early China*. New York: Oxford University Press, 2003.

Van Norden, Bryan, "Mengzi and Xunzi: Two Views of Human Agency." *International Philosophical Quarterly* 32 (1992): 161–184.

Watson, Burton, tr. *Basic Writings of Mo Tzu, Hsün Tzu, and Han Fei Tzu*. New York: Columbia University Press, 1967.

Watson, Burton, tr. *The Complete Works of Chuang Tzu*. New York: Columbia University Press, 1968.

Wong, David B. "Is There a Distinction between Reason and Emotion in Mencius?" and a reply to a commentary by Craig Ihara. *Philosophy East and West* 41 (1991): 31–44, 55–58.

Wong, David B. "Reasons and Analogical Reasoning in Mengzi." In *Essays on the Moral Philosophy of Mengzi*, edited by Xiusheng Liu and Philip J. Ivanhoe. Indianapolis, IN: Hackett, 2002.

Wong, David B. "Rights and Community in Confucianism." In *Confucian Ethics: A Comparative Study of Self, Autonomy, and Community*, edited by Kwong-loi Shun and David B. Wong. New York: Cambridge University Press, 2004.

Wong, David B. "Universalism versus Love with Distinctions: An Ancient Debate Revived." *Journal of Chinese Philosophy* 16 (1989): 252–272.

Wong, David B. "Xunzi on Moral Motivation." In *Chinese Language, Thought, and Culture: Nivison and his Critics*, edited by Philip J. Ivanhoe. Chicago: Open Court, 1996.

Wong, David B. "Zhuangzi and the Obsession with Being Right." *History of Philosophy Quarterly* 22 (2005): 91–107.

David B. Wong (2005)

LANGUAGE AND LOGIC

This entry focuses on concepts, issues, and themes of Chinese philosophy that involve language in view of its relation to reality, thought, and logic; the discussion is thus arranged on three central concerns in this regard: the issue of the relation between language and reality; the issue of the relation between language and thought; and the issue of the relation between language and logic. This entry is neither a historical study nor a comprehensive survey of the relevant ideas of thinkers from different historical periods, although there will inevitably be references to them. It is known that the term "logic" has been ambiguously and vaguely used; in this entry on language and logic in Chinese philosophy, first, by "logic" is meant primarily two things: (1) logical *reasoning* as embedded or expressed in natural (Chinese) language; and (2) the syntactic-semantic structure of Chinese language that underlies the surface grammar of Chinese language. The author neither pretend nor plan to discuss them exhaustively but to the extent that the issues to be addressed bear on Chinese philosophy and/or that the issues to be addressed are philosophically interesting. In this sense, this entry is not a discussion of logic or logical thought on their own in the history of Chinese thought, no matter how the term "logic" is understood.

LANGUAGE AND REALITY

The issue of the relation between language and reality has been one classical concern in philosophical study of language concerning what language is about, in Chinese tradition as well as in Western tradition. The classical issue emerged in Chinese tradition in terms of the issue of *ming-shi* (*ming* means "name" while *shi* means "reality") in its broad sense. In this part is discussed how some representative approaches in Chinese tradition explore four aspects of the issue: first, the issue of the issue of *zheng-*

ming (name rectification), the issue of reference, the issue of whether language can capture reality, and the issue of the relation between truth concern and *dao* concern.

NAME AND ACTUALITY: NAME-RECTIFICATION APPROACH. In the pre-*Qin* period three figures put forward their doctrines of name rectification: Confucius (551–479 BCE), Gongsun Long (320–250 BCE), and Xun Zi (298–238 BCE). The points of Confucius' and Gongsun Long's accounts are here rendered more philosophically interesting, though Xun Zi suggested a much more systematical account of names. The focus is on the first two due to space.

It is known that Confucius' major concern is with moral and social issues. His doctrine of name rectification serves his major concern. Nevertheless, the focus here is on those interesting points suggested in this doctrine from the point of view of philosophy of language. The passages in the *Analects* that are directly related to the issue of name rectification are three: 13.3, 12.11, 12.17, of which I make full citation to give a complete account (my translations).

> 13.3: Zi-lu asked, "If the ruler of the *Wei* State has you in charge of the state administration to governing the state and the people, what would be the priority of your administration?" Confucius replied, "It would surely be the rectification of names." Zi-lu wondered, "Is it so? What a pedantic way! Why is there need to bring in the rectification of names?" Confucius said, "You, how unenlightened you are! When a *junzi* (an enlightened gentleman) is ignorant of something, he is not expected to offer any opinion on it. If names are not rectified, then what is said in speech would not be in accord with things as they are (supposed to be); if what is said in speech would not be in accord with things as they are (supposed to be), then what is [supposed] to be done by using words would not be accomplished; if what is [supposed] to be done by using words is not be accomplished, then the [adequate] socially established ritual rules as manifested via ceremonies and music will not implemented; if these [adequate] socially established ritual rules will not implemented, then punishment will not be just; if punishment would not be just, then the people will not know where to move forward. Therefore, a *junzi* should give names only to those that surely can be adequately delivered in speech and deliver in speech only what surely can be adequately carried out in practice."
>
> 12.17: Jikangzi asked Confucius about governing. Confucius replied, "To govern is to rectify. If you lead the people by rectifying yourself, who would dare not to be rectified?"
>
> 12.11: Duke Jing of the Qi State asked Confucius about governing. Confucius replied, "Let the ruler [those that bear the title 'ruler' in the society] be the ruler [become what is prescriptively symbolized by the name 'ruler'], the minister be the minister, the father be the father, and the son be the son." The Duke said, "Excellent! Surely, if a ruler is not the ruler, a minister not the minister, a father not the father, and a son not the son, then, even if there are all the grain, how could I get to eat it?"

Confucius' doctrine of name rectification might as well be another way of presenting his teachings on moral cultivation and adequate governing: the teaching delivered in 12.17 is to rectify yourself to fit what those terms that signify your ranks, duties, functions and moral attributes mean (12:17), which amounts to sageliness within, while the teaching delivered in 12.11 is to participate in rectifying others to fit what those terms that signify their ranks, duties, functions, and moral attributes mean (12:11), which amounts to kingliness without. However, what really interests us here is some, explicitly or implicitly, suggested general point concerning the relation between language and reality. Let us start with an apparent puzzle: There appears to be a tension between the suggested two kinds of rectification approaches. On the one hand, the trademark title of this doctrine is "name rectification," and, as highlighted in 13.3, Confucius emphasizes the significance of name rectification. Nevertheless, on the other hand, 12.17 and 12.11 indicate that what is rectified is actually the persons who bear the (social-title) name. Which one is the primary goal while which serves as means? What is the due relation between the two kinds of rectification? Why doesn't Confucius directly emphasize rectifying the moral agent?

The reason seems to be this. To rectify the person (self and others) for the sake of self cultivation and of social reform, there needs a standard or norm that per se needs language as means or even as medium for the sake of its being carried out, communicated and passed on. Actually this is a two-level rectification process with the goal of rectifying the agent into a certain prescriptively specified person. The first step is to take a semantic accent strategy: instead of directly talking about how to rectify

the agent, it is to first rectify her (social-title) name under examination through assigning it a certain due prescriptive content which specifies the standard or norm to be met by any eligible referents of the name and thus gives the primary identity condition of such referents. The second step is to rectify the agent based on the primary identity condition of the expected referents of the name that has been established in the preceding semantic accent strategy.

An interesting point concern the relation between name and actuality, which is implicitly suggested by Confucius' account of name rectification is this. The due identity condition of actuality of a thing (say, a ruler) is not simply its status-quo happening or current appearance (say, the ruler-title-bearing person); rather, it consists in realization of its due place without transgressing its due scope (say, the person who really possesses the moral character that is expected for the ruler); name rectification will play its important or even indispensable role through the name carrying out and delivering the norm which specifies such a due place of the thing that is normatively denoted by the name (say, through rectifying the name "ruler").

If Confucius' account only implicitly suggests the foregoing point concerning social-title names and their due referents, one of Gongsun Long's contributions in this regard lies in his explicitly making the point in more general terms concerning any name and its related actuality and in a more sophisticated way. In this essay "*Ming-Shi-Lun*" (On name and actuality), Gongsun Long explains, "What the heaven and earth produce are things. When a thing goes its own way without transgressing its limit, it achieves its actuality (*shi*); when its actuality goes its own way without being out of its track, it achieves its due place (*wei*). If a thing goes beyond its due place, it is in wrong place; if a thing is in its due place, it is in right place. One is expected to rectify a thing in wrong place into right place; one is not expected to challenge a thing in due place by virtue of it being in wrong place. The rectification of a thing is the rectification of its actuality; the rectification of its actuality is implemented through the rectification of its name. Once its name is rectified, the standards for 'that' and 'this' will be formed up and stabilized" (my translation). Gongsun Long here emphasizes that a thing needs to go its own way without transgressing its limit to achieve its actuality; he further stresses that, once a thing achieves its actuality, there remains an issue of how to keep its actuality in due place; he explicitly points out that the so-called name rectification lies in rectifying the actuality of a thing in its due place through rectifying the due content of its name which identifies such due place and thus gives due identity condition for the thing and its actuality.

THE ISSUE OF REFERENCE: PURPOSE-PERSPECTIVE-SENSITIVITY APPROACH. But the above Confucius' and Gongsun Long's views on the relation of name and actuality via their accounts of name rectification would raise one general question concerning the issue of reference: whether, and in which way, the subject would contribute to the identity of a thing when she refers to the thing. Though with their distinctive backgrounds and concerns, Gongsun Long, the Mohist, and Zhuang Zi are kindred in spirit on this issue, taking essentially the same approach to the effect that an referring agent's referring action, which involves her purpose and focus, assigns a certain identity to the thing referred to, or specifies some aspect(s) of the referent as its identity (or multiple identities) and that, sensitive to one's purpose and focus, one is entitled to make her perspective shift in one's referring practice to focus on some other aspect of the referent as its identity. This approach might as well be called the "purpose-perspective-sensitivity approach." This section will focus on Gongsun Long's account and then briefly present Zhuang Zi's view; the Mohist relevant point will be addressed when the Mohist view on reasoning is discussed in the "Language and Logic" part of this essay.

In his essay "*Zhi-Wu-Lun*" (On referring to things), Gongsun Long emphasizes "No things [that are identified or named as things] are not what are referred to [by linguistic names] … if there is no referring in the world, nothing can be called a 'thing.' If without referring [names], can anything in the world be called 'what is referring to'?" (my translation). Gongsun Long's point here is that the relevant contributing elements involved in the subject's act of referring via a name (such as what is the subject's purpose, which aspect of the referent the subject intends to seek or focus on) make their intrinsic contributions to identity of the referent of the name. This point is also explicitly and emphatically addressed in his essay, "*Bai-Ma-Lun*" (On the white horse), as indicated in the passage "What makes a white horse a horse is their same [common] aspect given that it is what is sought. If what is sought is the common aspect, a white horse would be not distinct from (*bu-yi*) a horse [in regard to the common aspect]. If what is sought is not some distinct but the same aspect, then why is it that yellow and black horses meet what is sought in one case but not in the other? It is evident that the two cases are distinct" (my translation).

This crucial passage gives the fundamental rationale behind a number of Gongsun Long's arguments for the thesis "[the] white horse [is] not [the] horse." The statement "The white horse is not the horse" is just another way to say in our ordinary discourse "The white horse has its distinct aspect which the horse does not [necessarily] have," while the statement "The white horse is the horse" is just another way to say in our ordinary discourse "The white horse has its common aspect which the horse [necessarily] does have." Each of the two can be right, depending on which aspect of the white horse the referring subject is seeking or focusing on and thus refers to concerning the identity of the white horse. In so doing, she alerts us to avoiding the danger of over-assimilating distinctions, especially when the distinctive aspects need to be emphatically focused on.

Zhuang Zi proceeds essentially in the same direction on the issue (*The Zhuang-Zi*, Inner Chapter 3 "*Yong-Sheng-Zhu*"). Given an ox as whole already there, now what is its identity? How should one refer to it in terms of language? How should one identify it? As something exclusively determined by its "essence" or as a pack of flesh and bones? It seems to Zhuang Zi that, based on one's specific purpose, one can legitimately refer an ox as a pack of flesh and bones. One can say that, from the Zhuang Zi style view of the philosophy of language, the relation between language and an object in the world is not one-to-one relation but many-to-one relation: There are multiple referring expressions that refer to various genuine aspects of the same object. Depending on one's purpose, one is entitled to take a certain perspective to focus on one aspect of the object and thus identify the object as what the referring expression capturing that aspect would tell. What is important is that these distinctive referring expressions refer to different aspects of the same object, the ox as a whole, which are metaphysically complementary to each other.

THE *DAO* CONCERN IN CHINESE PHILOSOPHY AND ITS LANGUAGE ENGAGEMENT. A classical issue in Chinese tradition is whether, through language engagement, we can capture and deliver the ultimate reality, which the Chinese term "*dao*" primarily means (that is, the so-called metaphysical *dao*). The term "language engagement (with an ultimate concern)" means any reflective endeavor to capture (reach or characterize) what is ultimately concerned through language. Let us have a case examination of the opening statement, *Dao-ke-dao-fei-chang-Dao*, of Chapter 1, of the daoist classic the *Dao-De-Jing* whose legendary author is Lao Zi. For one thing, this passage has been considered to give a representative or classical presentation of the daoist attitude toward the relation between language and the world; for another thing, many subsequent interpreters in Chinese tradition resorts to this passage to make their points in this regard.

One standard, and also most prevalent, interpretative translation (Creel 1983) of this passage is this: "The *dao* that can be told of [in language] is not the eternal *dao*" (Chan 1963, p.139). According to this interpretation, what the first statement reveals is a fundamental daoist insight that is strikingly similar to that of Wittgenstein's well-known idea about the spoken and the unspoken: Language expressions or formulations cannot really capture what those expressions or formulations aim to say; any language engagement is doomed to fail to capture the genuine *dao*; the genuine *dao* has to be captured in a way that is beyond language; contemplation of the *dao* in silence requires sharply distinguishing the eternal *dao* from what can be formulated or captured in (or by) language, for the two are simply opposed to each other. This standard interpretation is partially correct: The *dao* that has been characterized in terms of language does not exhaust, and is not identical to, the genuine *dao*.

Although this interpretative translation has been circulated for a long time and does deliver part of the daoist message, it has been challenged whether it completely captures and delivers Lao Zi's genuine point as a whole in the context of the *Dao-De-Jing*. It is not merely because this standard interpretative translation neither syntactically nor semantically captures the Chinese original but also because it seems to miss some important point of Daoism in this regard. Another interpretation (Mou 2000) gives the following interpretative translation of the opening statement: "The *Dao* can be reached in language [*Dao-ke-dao*], but the *Dao* that has been characterized in language is not identical with, or does not exhaust, the eternal *Dao* [*fei-chang-Dao*]." Though partially agreeing to the first interpretation, this interpretation differs from the first one in this significant aspect: The *dao* that have been captured in language is not bogus *dao* but still parts of the genuine *dao*; and this understanding of the partial *dao* in terms of language engagement would significantly contribute to our capturing the *dao* as a whole. This is based on one crucial characteristic of the metaphysical *dao*: The metaphysical *dao* as unifying force that runs through the whole universe is not something separate or beyond and above all those finite things in the world that are particular and concrete; particular things in the universe, *wan-wu* (ten-thousand things), which obtain the power from the *dao*, are considered as manifestations of

the metaphysical *dao* and individualized-particularized *daos*; the relation between the metaphysical *dao* and its manifestations in *wan-wu* is essentially *yin-yang* complementary; *dao* and *wan-wu* are interdependent, interpenetrating, interactive and correlative.

Epistemologically speaking, and from the point of view of language engagement, the metaphysical *dao* thus can be somehow captured through our language and our understanding of *wan-wu*. In this way, the point of the second interpretation is this: Instead of indiscriminately giving a negative claim against any language engagement with the ultimate concern, in the opening statement, Lao Zi reveals a two-sided transcendental insight which, on the one hand, positively affirms the role of the language-engaged finite point of view in capturing the ultimate concern and, on the other hand, alerts us to the limitation of the finite point of view and emphasizes the transcendental dimension of the *dao*.

It is noted that, in this regard, A. C. Graham's view seems to be much more moderate than the foregoing standard interpretation when he explains why there is the trouble with words: "The trouble with words is not that they do not fit at all but they always fit imperfectly; they can help us towards the Way, but only if each formulation in its inadequacy is balanced by the opposite which diverges in the other direction" (Graham 1989, p. 219). Nevertheless, the above second interpretation is more moderate than Graham's to this extent: It is not the case that the language engagement *always* fits *imperfectly*. That really depends on which part, dimension, or layer of the *dao* is set out to be captured in language engagement and on what kind of language function is at issue.

First, if a language engagement does not pretend to be exhaustive or conclusive regarding the *dao* but rather takes a finite point of view, it is reasonable to say that what has been captured in language in that case does fit adequately. When a language engagement takes a finite point of view, what is needed is not to reject such a finite point of view per se, but to hold the transcendental insight simultaneously, which would alert us to the limitation of the finite point of view and its due scope. Second, capturing something in language does not necessarily mean imposing a definition or formulation with a certain fixed format, meaning or usage. For instance, in contrast to *mere* description and descriptive designation, rigid designation via direct reference is one way to reach the genuine *dao* as a whole, as Lao Zi's own language-engagement practice illustrates (for example, Lao Zi did somehow successfully use the term *dao* to designate the *dao* as a whole).

TRUTH CONCERN AND *DAO* CONCERN. It seems that the truth concern is a dominant concern in Western tradition while the *dao* concern is a dominant concern in Chinese tradition. What is the relation between the truth concern and the *dao* concern? Are they dramatically and totally different reflective concerns in philosophy? (Given that the term "dao" primarily means the metaphysical *dao* concerning the way of the world as it is, especially in Daoism, and that any reflective concern, including the *dao* concern, that is open to criticism and self-criticism needs to be characterized in terms of language, and also given that one important aspect of the truth concern is about the relation between language and reality, this is a significant topic concerning the relation between language and reality at the meta-philosophical level.)

Although, as this author sees it, a silent majority of philosophers who are familiar with Chinese philosophy have considered both concerns essentially in accordance with each other, some scholars argue otherwise. There are two representative views. One takes it that, in contrast to what is called "Western sentential philosophy," the dominant portion of the classical Chinese philosophy is a non-sentential philosophy that is not essentially related to those concepts that are intrinsically connected with sentential philosophy like proposition (or semantic content), truth and belief (Hansen 1985/2003). This argument might as well be called the "no-sentential-concern argument." Another view takes it that the significant part and the primary concern of the classical Chinese philosophy have been considered be its moral concern and its ethical accounts; and the moral concern is not with how to understand impersonal material world but with the ethical constitution in the human society. In this way, it is not the by-default account of truth (the correspondence account) but a pragmatic account of truth that plays the role (Hall 1997, 2001). This argument might as well be called the "pragmatic truth argument."

In contrast to the silent majority's presupposed position, these views have been voiced prominently and loudly especially in West and thus have left on many who are not familiar with Chinese philosophy the impression to the effect that there is no truth concern in Chinese philosophy and that the truth concern in Western tradition and the *dao* concern in Chinese tradition are dramatically different from each other. This impression is incorrect at least to the following extent: First, it is highly controversial; second, to many experts, it is not so. But their views deserve careful examination, and the involved issue deserves a systematic discussion, instead of being silently dismissed. Though it is not a place to give such a detailed

discussion here, I intend to use the following strategy to assist the interested reader in examining the issue: this entry briefly addresses a number of basic things, to which one needs to pay due attention when one intends to explore the issue and give adequate evaluation of the competing views, but which might be ignored by some advocates of the above mentioned challenges.

Let us start with our pre-theoretic, or "folk," understanding of truth: A true (linguistic) sentence or statement (or the thought/belief it delivers) describes or characterizes (extra-linguistic) things as they are. When the term "our" is used here, its reference by no means includes only people in West but surely also includes people in Chinese speaking regions, now and in the past, no matter how such a pre-theoretic understanding has been indicated in their natural languages—whether it is expressed by a unified single term in a phonetic language (like "truth" in English), or it is expressed via various multiple-character phrases in the Chinese ideographic language (such as *shi-shi-qiu-shi*, meaning "seeking what things actually are," or *qiu-dao*, meaning "pursuing the *dao*/way of the world"). For convenience, this pre-theoretic understanding of truth is sometimes called our pre-theoretical "correspondence" understanding to highlight the accordance relation of our thought or our language with (the *dao*/way of) the world (including the human society) in which truth under such an understanding consists. Now the reader can think about this: Given that those approaches in Chinese philosophy as discussed in the preceding sections of this part are all distinctive illustrations of the *dao*-concern on several significant fronts and thus that all of them are thus intended to capture and deliver extra-linguistic things as they are, are those approaches dramatically separate and different from the reflective truth concern that is based on the foregoing pre-theoretic understanding of truth?

At this point three notes are due. First, the metaphysical commitment of our pre-theoretic understanding of truth per se as presented above is minimal: It does not commit to any ad hoc metaphysical criterion for what counts as reality, and it is compatible with a variety of ontological accounts of extra-linguistic things (say, snow's being white). For example, a *realist* pre-theoretic "correspondence" understanding of truth is actually a *combination* of our pre-theoretic understanding *of truth and a realist* ontological understanding or explanation of what counts as, say, snow's being white. (In this way, any argument that resorts either to the fact that some specific version of the truth concern in West is combined with some unfavorable metaphysical explanation or to the fact that a certain metaphysical account of the *dao* is so different from some representative metaphysical understanding of what counts as reality cannot automatically imply that the *dao* concern in Chinese tradition and the truth concern in Western tradition are dramatically different.)

Second, it is arguably right that our pre-theoretical "correspondence" understanding of truth plays its important and enormous explanatory role both in our daily lives and in our reflective lives (including philosophical inquiries). In most cases, whether for the sake of psychological satisfaction, intellectual enjoyment, scientific honesty, legal obligation or success of our actions or even for its own sake, we intend to understand what really happen(ed) around us rather than illusions, we hope that others tell us truths instead of lies or mere wishful thinking, we want to know those beliefs, thoughts or statements that are true. Moreover, in almost all of cases, we (even for those who advocate some understandings of truth that clearly revise or go against our pre-theoretic understanding of truth) seriously intend that the genuine contents of our own thoughts and claims to be delivered (or represented) to, and understood by, others "correspondently"—or without distorting or losing their original contents; we intend to behave in a way that does not go against the laws or *dao*/way of the world. In this sense and to this extent, it is not merely the case that, in many situations, with such a pre-theoretic understanding of truth, we consciously pursue truths; rather, it is a stronger case: whether consciously or unconsciously, we unavoidably presuppose our pre-theoretic "correspondence" understanding of truth both in our ordinary folk talks and in our reflective talks including philosophical discourses, either as one central explanatory norm to regulate and explain the purpose of our thoughts and actions or as one important explanatory basis to explain some other significant things in our folk and reflective lives.

Third, there is the distinction between truth nature and truth criterion: The former is examined by asking what truth is, what truth consist in or what it is for a statement (or belief) to be true, while the latter is examined by asking what is the criterion by which one can identify, judge and distinguish true statements from false ones. Our pre-theoretic understanding of truth is about the truth nature instead of truth criterion. Actually, the foregoing three notes indicate three significant respects, among others, in which one can critically examine the relation between the truth concern and the *dao* concern in Chinese tradition as well as the nature and due functions of major competing theoretic accounts of truth in

the Western tradition. (For further discussion of the issue, see Mou 2006.)

LANGUAGE AND THOUGHT

Besides the issue of the relation between language and reality, another important concern of philosophical reflection on language is the issue of the relation between language and thought. This entry will discuss this issue as explored in Chinese tradition in two fronts: (1) the issue of the relation between speech and ideas in mind in regard to whether and to what extent the former can capture and deliver the latter; (2) a reflective concern with how the structure of Chinese language bear on the orientation of philosophical thought in Chinese tradition.

SPEECH AND IDEAS: FOUR APPROACHES. The relation between speech and ideas in mind is one central concern in the so-called *yan-yi-zhi-bian*, that is, the debate on the relation between speech (*yan*) and meaning (*yi*, in the sense to be explained), which originated in the *Wei-Jin* period; but the following discussion will not limited to a number of representative approaches in this debate during that time but incorporates some other representative approach in Chinese tradition. (Note that, though using the ready-made translation “speech and meaning” of *yan-yi* here for the sake of convenience, and though *yi* in this debate also means *dao*-like principles in its metaphysical sense and the human understanding of them, by *yan-yi* is meant “speech and ideas in mind” in this context.) In the following, four representative approaches are focused on: (1) the “meaning-delivery-beyond-speech-capacity” approach; (2) “forgetting-speech-once-achieving-meaning” approach; (3) the “meaning-delivery-within-speech-capacity” approach; (4) the context-sensitivity approach. The first three approaches are three representative approaches in the *yan-yi-zhi-bian*during the *Wei-Jin* period, though the first two are actually kindred in spirit (see Chen 2004 for a recent discussion; my interpretation of the third approach is somewhat different than his), while the fourth one is my interpretative elaboration of the relevant points of Ji Zang’s Buddhist Middle-Way doctrine of double truth.

The “meaning-delivery-beyond-speech-capacity” approach was advocated by Ji Kan (223–262). This approach’s main arguments are these. First, some of our ideas in mind are so delicate and sophisticated that speech simply cannot capture them. Second, our ideas in mind are dynastic while speech is static, and therefore speech cannot fully capture ideas in mind. The “forgetting-speech-once-achieving-meaning” approach was advocated by Wang Bi (226–249). This approach acknowledges a certain important role played by speech as a means to achieve meaning. For example, when one intends to understand some other’s ideas or when one intends to have one’s own ideas to be understood by some other, one has to rely on speech to understand them or express them.

But this approach still takes it that eventually speech would hinder one’s understanding ideas per se and so that one should forget speech once achieving the ideas. This line of thought sounds like a Wittgenstein’s well-known metaphor to the effect that, once one climbs up on the building by means of a ladder, one needs to discard the ladder to keep oneself in the high position. It is noted that, though the first and second approaches have their differences in emphasis and focus, they share the basic positions concerning the relation between speech and ideas. Both think that ideas are primary while speech is only secondary, that ideas and speech can, and should, be separate and that at most speech serves merely as a means and makes no contribution to the constitution of thought and ideas.

The “meaning-delivery-within-speech-capacity” approach is suggested by Ouyang Jian (?–300). This view has been ignored for a long time and not a strong voice in the traditional Chinese philosophy in contrast to the mainstream approach on this issue; but some of the points of this approach deserve a close examination. Ouyang argues that:

> Surely one can achieve a principle in the form of ideas in one’s mind; however, without language [as media and as means], those ideas cannot exist in a smooth and coherent way. Given that a thing has been stabilized in a certain definite aspect, without language [in terms of name], one cannot identify and thus distinguish the thing [in view of the stabilized definite aspect] from the others. If one’s ideas cannot exist in a smooth and coherent way through the role of language, they cannot hold tight in connection with each other; if the thing cannot be identified and distinguished in terms of name, the distinctive ideas and insights cannot be shown evidently. But, as a matter of fact, the distinctive ideas can be shown evidently in terms of distinctive names, and speech holds ideas tight in connection and in a smooth and coherent way. Let us see why it is so. It is not because a thing has its ready-made fixed name; it is not because a principle has its fixed unchangeable language

> expression. The reason is this. When people intend to capture things as they are in a distinguishing way, they give them distinctive names. When people intend to declare their distinctive ideas, they employ distinctive language expressions that fit distinctive ideas. Names [and/or their meanings] change in accordance with the transformation of their named things; while speeches [and/or their meanings] change in accordance with the change of the contents of ideas. It is just like echo responds to sound, shadows attaches to body; they do not exist as two separate things. If they are not separate things, then speech can fully capture ideas; this is why I hold on my position.

(*YI-WEN-LEI-JU*, VOL.19)
(MY TRANSLATION)

There are two interesting points that seem to really engage with the two preceding views. First, speech is not merely a means but also a medium of ideas at least in regard to its contribution to their internal coherent construction. Second, as far as speech as means is concerned, though speech is relatively static and stable, that certainly does not mean that language is just as static as a dead thing; language itself also keep changing responding to the change of what it is to express. This is true as evidenced by the history of the development of natural languages. Although Ouyang's first point is still quite vaguely made and expressed, his position makes distinct contribution on the issue.

The fourth approach, the context-sensitivity approach, suggested by Ji Zang (540–623), a significant figure of Chinese Buddhism who elaborated and systematized *Mahayana* doctrine of Buddhism. Ji Zang's doctrine of double truth has interesting implications from the point of view of philosophy of language. First, a brief outline of the major ideas of his double-truth account. It seems to Ji Zang that there are two kinds of truth, truth in the common sense and truth in the higher sense, on each of three varying levels; what is the truth in the higher sense at a lower level becomes merely truth in the common sense at the higher level. At the first level, the common people take all things as really being and know nothing about their non-being, while the Buddhas have told them that actually all things are non-being and empty. At the second level, to say that all things are being is one-sided, but to say that all things are non-being is also one-sided; at this level, the Buddhas would say that what is being is simultaneously what is non-being. At the third level, saying that the middle truth consists in what is not one-sided means to make distinctions, and so this is merely a common sense truth; the higher truth consists in saying that all distinctions are themselves one-sided, and the middle path is neither one-sided nor not-one-sided. That amounts to denying the adequacy of any speech to capture the truth in the higher sense at this highest level, that is, the highest truth, which needs to be contemplated in silence.

Although Ji Zang as a Buddhist thinker still maintains that the highest truth cannot be captured and delivered via language but has to be contemplated in silence, but he emphasizes that all those truths, both in the common sense and in the higher sense and both at the first level and at the second level, can be captured and delivered in terms of language that involves relatively stabilized and fixed conceptual distinctions. With his explicitly distinguishing truths in distinct senses and at distinctive levels and acknowledging important role played by language at the first and second levels, Ji Zang's general point is philosophically interesting: we need to have it sensitive to the context whether speech can effectively capture and deliver the truths, that is, our understandings and comprehensions of the world.

There are two notes concerning evaluation of the foregoing views. First, to evaluate the ancient thinkers' views here, we indeed need to pay attention to those still valuable thoughts; on the other hand, we also need to note that one of the reasons why those ancient thinkers held that speech is not able to fully capture meaning is this: some conceptual and explanatory resources in contemporary philosophy that are available to us to capture and deliver some sophisticated ideas and thoughts were simply unavailable to those ancient thinkers; so there is no wonder why they felt the linguistic means then available to them were not sufficient to capture some complicated thoughts and ideas. Second, as emphasized at the outset, the term *yi* in the *yan-yi-zhi-bian* (the debate on the relation between speech and meaning) has its much wider coverage than what the term "thought" in the contemporary debate on the relation between language and thought is to cover: the latter primarily mean propositional thoughts while the former's coverage includes non-propositional ideas, emotions and some characteristic existential experience; a claim putting into doubt or denying the capacity of speech to capture such non-propositional mental things could be compatible with the positions by those whose primary concern is with the relation between language and propositional thoughts.

THE ISSUE OF THE STRUCTURE OF CHINESE LANGUAGE AND REFLECTIVE WAY OF THINKING IN CHINESE PHILOSOPHY. It seems that certain characteristic features of Chinese language influence or encourage some orientations in the Chinese (folk and reflective) way of thinking. Due to the topic, the focus will be on such influence on reflective inquiry in Chinese philosophy. Nevertheless, the reflective way of thinking is not separate from, but largely in accordance with the folk way of thinking via some reasonable pre-theoretic intuitive understanding on those issues that deserve further reflection.

We start with some known facts about certain characteristic features of Chinese ideographic language and Western phonetic language (say, English) in comparison that might, to some extent and in a certain scope, reflect some distinctive orientations or tendencies in the ways of thinking of the two linguistic communities. We know how we as English speakers give our names and addresses: We first give our given names (thus being called "first name") and last give our family names (thus being called "last name"). However, in Chinese, the family name goes first (thus the family name is really the first name in Chinese way) and then the given name (for example, the real order of my whole name in Chinese is "Mou Bo" instead of "Bo Mou"). For, in philosophical terms, the family (name) is both metaphysically and logically prior to the individual (name), and the former provides a necessary holistic background for understanding the latter. By the same token, in contrast to its way in English, a mailing address (taking mine as an example) should go this way when delivered in Chinese: "USA, California, San Jose, San Jose State University, Department of Philosophy, Mou Bo"; that is, the larger thing goes first while the smaller thing next. It is arguably right that the structure of Chinese language in this respect to some extent bears on the orientation of the way of thinking of the Chinese people as a whole.

(There are two notes. First, when the word "bear on" is used instead of "influence" alone, what is meant is that the relation between the former and the latter is bi-directional instead of one-directional. The actual situation might be this: When the way of the Chinese language originally formed up, it was influenced by the way of thinking of the people around that time; on the other hand, when such a way has become relatively stable and been followed and passed on generation by generation, it has conversely influences the way of thinking of the future Chinese language speakers to some extent. Second, the foregoing influence certainly implies neither that the people speaking in Chinese tend to put the family/the collective interest first nor that, say, English speakers tend to do otherwise. Even if such a distinctive order of which one is mentioned first, next, and last indeed influence which one would first go in mind at some level, surely one can say that, though saying things in a certain order, I actually think about all the involved things once for all simultaneously. Exactly how it would happen if any has yet to be carefully examined.)

Now, through a representative case analysis, we examine how the structure of Chinese language bears on the orientation of reflective inquiries in Chinese philosophy through one case analysis. The Platonic one-many problem has been a long-term issue in the Western philosophical tradition. The problem begins with the following observation: objects around us share features with other objects; and many particular individuals, say, horses bear the same name "horse." The Platonic one-many problem presupposes that there is one single universal entity which is common or strictly identical across all those particular concrete horses and by virtue of which many individual horses bear the same name "horse"; the single universal entity is labeled "horseness." The Platonic one-many problem is how to characterize the status of universals and the ways by which particulars share universals. However, there seems to be one puzzle: why the classical Platonic one-many problem in the Western philosophical tradition has not been consciously posed in the Chinese philosophical tradition and why, generally speaking, classical Chinese philosophers seem less interested in debating the relevant ontological issues. One suspects that the structures (the surface and deep ones together) and uses of different languages might play their roles in pushing philosophical theorization in different directions; the ways of speaking and writing of the Chinese language might somehow reveal and reflect Chinese folk ideology and then influence the ways in which certain philosophical questions are posed and certain ontological insights are formed.

The problem of relating Chinese thought to the structure and functions of the Chinese language has for generations tantalized sinologists and those philosophers who are concerned with the problem. Nevertheless, in the last two decade, some significant progress has been made in this regard. Chad Hansen (1983) advances a novel and provocative theory about the nature of the classical Chinese language. The central thesis of Hansen's theory is his mass-noun hypothesis. Its main ideas are these: (1) the (folk) semantics of Chinese nouns are like those of mass-nouns (i.e., those nouns referring to the so-called inter-

penetrating stuffs, like the nouns "water" and "snow"), and naming in Chinese is not grounded on the existence of, or roles for, abstract entities (either on the ontic level or on the conceptual level) but rather on finding "boundaries" between things; (2) influenced by the mass-noun semantics, the classical Chinese semantic theorists and ontological theorists ew words in ways that are natural to view mass nouns rather than count nouns, and Chinese theorists tend to organize the objects in the world in a mereological stuff-whole model of reality (the term "mereology," in its technical sense, means the (mathematical) theory of the relation of parts to whole).

In this way, according to Hansen, the language theory of classical Chinese philosophers differs fundamentally from the language theory of Western philosophy. This hypothesis has been challenged mainly in three ways. One way is to challenge the mass-stuff model from the perspective of a holographic process ontology (Cheng 1987, Hall and Ames, 1987).

Although some scholars also emphasize the implicit ontology of Chinese language, they focus on the case analysis of the typical philosophical nouns or terms, such as *tai-ji*, *wu*, *yin-yang*, *wu-xing*, which constitute the basic lexico (vocabulary) of Chinese metaphysical systems as found in the writings of the early Confucianists, the early daoist, and Neo-Confucianists. They argue that those nouns stand for interpenetrating wholes and parts in a quite different sense from Hansen's: the individual things behave in the on-going patterns and in the events or processes of interaction among them, and the universe behaves as an organic whole with parts exemplifying the structure of the whole; they claim that Chinese words in general share this ontological feature of combining universality and particularity, abstractness and concreteness, activity and the result of activity. In this way, some writers (Hall and Ames, 1987) prefer to consider the relations of "parts" and "wholes" in terms of the model of "focus" and "field" and take Chinese ontological views as holographic rather than mereological.

Another way is to directly challenge Hansen's mass noun hypothesis, arguing that there is a clear grammatical distinction in classical Chinese between count nouns and other nouns (Harbsmeier 1989, 1991). Claiming that there is a clear grammatical distinction in classical Chinese between count nouns and other nouns (generic nouns and mass nouns), Harbsmeier (1991) insists that the mass-noun hypothesis is "historically implausible and grammatically quite wrong-headed." However, as Hansen himself emphasizes (1992), his mass-noun hypothesis is not a syntactic claim that classical Chinese nouns have mass-noun grammar but a semantic interpretive hypothesis that the semantics of Chinese nouns may be like those of mass nouns, and classical Chinese theorists view words in ways that are natural to view mass nouns. So it seems to Hansen that Harbsmeier systematically confuses syntax and semantics and misinterprets his semantic hypothesis. Although one can agree with Hansen at this point, Harbsmeier's criticism is not irrelevant in the following sense. It seems that Harbsmeier insists that his alleged distinction between count nouns and other nouns is not merely grammatical but also semantic (or takes the grammatical difference in question to have semantic implications); Hansen thus needs to deal with the linguistic (semantic) evidence against his hypothesis that the semantics of classical Chinese nouns may be like those of mass-nouns.

The foregoing first challenge from the point of view of a holographic process ontology could be compatible with Hansen's approach; for the process ontology is essentially compatible with the ontological position, a kind of nominalism, presupposed or implied by Hansen's mereological mass-stuff hypothesis. Hansen's view is given in a semantic perspective that can be compatible with a pragmatic perspective with its focus-field orientation. I have responded to Hansen's view in a similar semantic perspective and within the same mereological-analysis track. But, disagreeing with Hansen's mass-noun hypothesis, I suggest and argue for a collective-noun hypothesis (Mou 1999). Its main ideas are these: (1) Chinese common nouns typically function, semantically and syntactically, in the way collective-nouns (that is, those nouns that denote collections of individual things, like the English nouns "people" or "cattle") function, and the folk semantics of Chinese nouns are like those of collective-nouns; (2) their implicit ontology is a mereological ontology of collection-of-individuals both with the part-whole structure and with the member-class structure, which does justice to the role of abstraction at the conceptual level; and (3) encouraged and shaped by the folk semantics of Chinese nouns, the classical Chinese theorists of language take this kind of mereological nominalism for granted; as a result, the classical Platonic one-many problem in the Western philosophical tradition has not been consciously posed in the Chinese philosophical tradition, and classical Chinese philosophers seem less interested in debating the relevant ontological issues. This mereological collection-of-individuals model of reality would provide a more reasonable interpretation of the semantics of classical Chinese nouns and the classical Chinese ontological theory. The collective-noun hypothesis makes a stronger claim that Chinese nouns do

not function as count nouns but typically function, both syntactically and semantically, as collective-nouns.

LANGUAGE AND LOGIC

As indicated at the outset, the term "logic" in this essay on language and logic in Chinese philosophy means two related things: first, logical reasoning as embedded or expressed in natural (Chinese) language; second, the syntactic-semantic structure of Chinese language that underlies the surface grammar of Chinese language. The two are related in this way: The reasoning as embedded in a language is intrinsically connected with the syntactic structure of such a language and makes sense in view of its semantic structure which per se is related to its syntactic structure; thus, to understand the reasoning as embedded in natural (Chinese) language, one needs to understand its syntactic-semantic structure. Actually, in the discussions of the previous two parts, the second issue has already been addressed in view of its relations with the central concerns there. In this part, we focus on the first issue. With space limitation, the strategy is this: we will start with an examination of some reasoning patterns in the Mohist discourse and then raise a general issue about the due relation between two modes of reasoning; that is, deductive reasoning versus evocative reasoning, in view of Chinese philosophical practice.

REASONING PATTERNS IN THE MOHIST DISCOURSE. The two trademark basic principles for deductive reasoning are the principle of non-contradiction and the law of identity, both of which are expected to be observed for the sake of good deductive reasoning. The principle of noncontradiction states that it is not the case that both that p and not p (where p is any proposition). The law of identity states that everything is identical with itself (for everything x, $x = x$). We begin with an example of reasoning via Aristotelian deductive logic:

Pr.1 If x is y, then to do something to x is to do it to y.

Pr.2 Robbers are people.

Therefore, killing robbers is killing people.

Pr.3 It is wrong to kill people.

Therefore, it is wrong to kill robbers.

However, the Mohist disagrees to this reasoning, arguing that killing robbers is not killing people. Their reason is this. In our ordinary language use, we often shift our attention from what is shared between them to what is distinct between them, depending on the nature of context and concrete situation. The Mohist distinguishes three sorts of contexts and considered the case of "killing robbers/killing people" as one case of the second kind (Graham 1989).

(1) The involved context would typically call our attention to what is shared between involved parties: In such a kind of contexts, for example, we say "Black horses are horses" or "Riding black horses is riding horses." (Typically, for the purpose of riding a horse, the color of the horse does not matter.) One example given in the Mohist text is this: "Huo is a person; to be concerned for Huo is to be concerned for persons." In the context of the Mohist text, Huo is a slave who is too humble for one to be concerned for anything about them except that he is a person; someone concerned for him is concerned for anyone as a person. Also note that the Mohist held the view of universal concern for anyone.

(2) The involved context would typically call our attention to what is distinct between involved parties. Consider three sentences in such a kind of contexts. First, "A carriage is wood; but riding a carriage is not riding wood": Typically, what is concerned with in the context of talking on riding something is whether or not the thing has the riding-function. Second, "Her younger brother is a handsome man; but loving her younger brother is not loving a handsome man": Typically, in this context, loving him is not for his looks. Third, "Robbers are people; abounding in robbers is not abounding in people, being without robbers is not being without people: Typically, in this context, what is called attention to is something distinct with robbers.

(3) The involved context would typically call our attention to both what is common and what is distinct between involved parties. One might say both "The white horse is not the horse" (in so saying, as analyzed before in view of Gongsun Long's approach, one pays attention to the distinct aspect of the white horse from the horse) and "Riding the white horse is riding the horse" (for the sake of riding a horse, the color of the horse does not matter).

What the Mohist calls our attention to is a variety of reasoning patterns embedded in our linguistic practice and the context in which reasoning utterances are made. In contrast, deductive reasoning focuses on logical necessity and logical entailment that seems to be concerned about only in the context (1) among the foregoing three kinds of contexts as the Mohist identifies. It is noted that such a focus-shift is not supposed to make at random but has its due metaphysical foundation: an object of study really possesses its multiple aspects/layers/dimensions. When saying "robbers are people," one focuses on the aspect of robbers, A, that makes them being people; nev-

ertheless, when saying "killing robbers is not killing people," one's focus shifts to some other aspect of robbers, A*, which is possessed by robbers rather than by the other people and which makes robbers deserve being killed (from the Mohist point of view): killing robbers for the sake of A*; that does not amount to killing people for the sake of A* because people generally speaking do not possess A*. Note that this challenge is rather to the indiscriminate applicability of deductive reasoning at the surface level than to its applicability to various extents in different linguistic contexts. That constitutes a deep reason why the Mohist view, as A. C. Graham points out, has its "Wittgensteinian look," which emphasizes the language use and claims that meaning consists in use.

LOGICAL VERSUS EVOCATIVE ARGUMENTATIONS IN CHINESE PHILOSOPHY. It is known that any philosophical inquiry needs to base its conclusion on justification or argumentation rather than simply dogmatically taking something for granted. There are two basic modes of argumentations in philosophy, one is logical (in its narrow sense) and the other evocative, though sometimes only the former is highlighted and celebrated. The two modes of argumentations are sometimes contrasted as "logical versus rhetoric," "inferential versus preferential" or "probative versus prohairetic." A logical argument is a set of statements in which one or more of the statements, the premises, purports to provide a reason or evidence for the truth of another statement, the conclusion, either in deductive way or in inductive way. When it does, we say that the premises entail or support the conclusion, or that the conclusion "follows from" the premises. We traditionally divide logical arguments into deductive and inductive arguments. The term "evocative" is used in contrast to the term "logical" used in the narrow sense; it means producing or suggesting or triggering (generally speaking, evoking) some subsequent thought or conclusion primarily in some non-"logical" way, neither deductively nor inductively as specified above. Among a variety of evocative argumentations, what have been often addresses especially in humanities are argument by (relevant) analogy [drawing its conclusion by evoking a similarity between some particular aspect of two things from, or on the basis of, their similarity in some other particular aspect(s) or in some other general aspect], argument by appealing to value [drawing its conclusion by appealing to one's value which is appreciated through one's life-experience and understanding of the world (and/or the human society)], and argument by appealing to (credible) authority [draws its conclusion by appealing to trustworthy and knowledgeable authority on the issue under examination].

Both modes of argumentations are widely used in the classical Chinese philosophy. Let us consider some examples in Confucius' *Analects* to illustrate the point. Contrary to some unjustified impression, this classical text is not lack of deductive reasoning; though some of the deductive-reasoning cases need one to be careful enough to identify between lines, some other are quite evident—for example, the reasoning given in the previous citation where Confucius' doctrine of name rectification is discussed. On the other hand, the argumentation implicit in Confucius' version of the Golden Rule as delivered in 6.28 of the *Analects* illustrate both argument by analogy and argument by appealing to value. Its conclusion is that one should treat others in a certain moral way; which way? One is expected to identity the way partially based on how one would like to be treated: Due to the common human-being identity among human beings that result in similarities in many relevant aspects between human moral beings, [Confucius' version of] the Golden Rule guides the moral agent to "draw the analogy from oneself [the way one would desire to be treated]" to how to treat others in a moral way (that is, to evoke the similarity in regard to what would be desired and what would be rendered moral, by both the moral agent and the moral recipient). Furthermore, the moral agent is not expected to start from nowhere but to be a moral agent with (a certain degree of) moral sensibility; that is, the virtue of *ren*; this initial moral sensibility serves as the internal starting point of how the moral agent is to adequately draw the analogy. This moral value, according to Confucius, is commonly, more or less, shared by all human moral agents; this moral value would thus contribute to what would be rendered moral by both the moral agent and the moral recipient (the similarity in this regard). In this way, Confucius' version of the Golden Rule appeals to the moral value to justify a reasonable version of the Golden Rule. Through this example, one can see how argument by analogy and argument by appealing to value interplay in the argumentation in Confucius' version of the Golden rule.

Indeed, when appealing to value and appealing to authority, one should be careful; otherwise, one might fall into fallacies. But, what is at issue is not whether people and philosophers have ever made adequate argument by appealing to value or to authority in their reflective practice. Philosophers do it, more or less, directly or indirectly, and explicitly or implicitly, in their argumentations and explanations, and both in the Chinese tradition and

in the Western tradition. What is really philosophically interesting is how to do it in some adequate way to avoid fallacies. In this connection, unlike deductive reasoning, there is no formal rule manual available but some general guidelines. Those general guidelines, largely, present themselves as explanations of what constitute fallacies in reasoning or argumentation. One case is the fallacy of dubious authority regarding current situation: An argument commits this fallacy when it mistakes some person as a trustworthy and knowledgeable authority about the current situation. Another case is the fallacy of relevance: mistaking relevant dissimilarities as irrelevant: An argument (typically, an argument by analogy) commits this fallacy when it mistakes relevant differences between two (kinds of) things as irrelevant to the issue under examination. For example, to illustrate the fallacy of relevance, let us consider how Mencius criticizes an argument by analogy made by Gao Zi, his contemporary in regard to the original human (moral) nature:

> Gao Zi said, "[Original] Human [moral] nature is like the willow tree, and righteousness is like making a drinking cup. To turn human nature into humanity and righteousness is like turning the willow tree into cups." Mencius responded, "Could you make the cups out of the willow tree without violating its nature, or do you have to violate the nature of the willow tree before you can make the cups? If you have to violate the nature of the willow tree in order to make cups, then [based on your analogy] do you have to also violate human nature in order to make it into humanity and righteousness? Your analogy would lead all people in the world to consider humanity and righteousness as the source of disaster [because they required the violation of human nature]!" (*Mencius* 6A:1. My modification of the translation in Chan 1963, p. 51)

Mencius here criticizes Gao Zi for his inadequately paying attention only to some superficial similarity between making a cup and building character but ignoring a crucial difference between making a cup out of the willow tree (injuring the willow) and building human moral character from the human original moral nature (without involving violence and injury); in this way, in our terminology here, Mencius actually criticizes Gao Zi for his mistaking one significantly relevant dissimilarity between both as irrelevant in his argument by analogy.

The relation of the two modes of argumentations together with their respective nature and status in philosophical inquiry has been under reflective examination. We can think about a number of questions in view of the cited cases above in the traditional Chinese philosophy and through examining our own reflective practice in argumentation: When carrying out deductive (or evocative) argumentation, could one's argumentation be totally immune from evocative (or deductive) argumentation? (Think about where premises in many deductive arguments come from; also think about whether one still needs to rely on a certain standard and resort to the two basic principles of deductive reasoning mentioned above in some way when carrying out evocative argumentation.) One strategic methodological point in regard to the relation between the two modes of argumentation is that they come into "mutually supportive overall harmonization" (Rescher 1994, p. 58). That is especially true in view of Chinese philosophical practice. For this orientation is kindred in spirit with the *yin-yang* way of thinking which emphasizes the complementary nature between seemingly competing approaches. Indeed, the *yin-yang* way of thinking has fundamentally influenced the orientation of mentality, and the way of carrying out reflective argumentation, of subsequent Chinese thinkers in various schools or movements.

See also Chinese Philosophy: Metaphysics and Epistemology.

Bibliography

Ames, Roger, and David Hall. *Thinking through Confucius*. Albany, NY: SUNY Press, 1987.

Analetcs. Translated by D. C. Lau as *Confucius: The Analects* (Harmondsworth, Middlesex: Penguin Classics, 1979).

Chan, Wing-tsit. *A Source Book in Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1963.

Chao, Y. R. "Notes on Chinese Grammar and Logic." *Philosophy East and West* 5 (1955): 31–41.

Chen, Bo. "Debate on Language-Meaning and the Skeptic Argument about Meaning," Presented at the ISCWP international conference "Davidson's Philosophy and Chinese Philosophy" (June 8 and 9, 2004; Beijing, China). Chinese abstract in *World Philosophy* 5 (2004): 26–28.

Cheng, Chung-ying. "Inquiries into Classical Chinese Logic." *Philosophy East and West* 15 (1965) (3–4): 195–216.

Cheng, Chung-ying. "Logic and Language in Chinese Philosophy." *Journal of Chinese Philosophy* 14 (1987): 285–307.

Chmielewski, Janusz. "Notes on Early Chinese Logic (I)." *Rocznik Orientalistyczny* 26 (1) (1962): 7–21.

Creel, Herrlee. "On the Opening Words of the Lao-Tzu." *Journal of Chinese Philosophy* 10 (1983): 299–329>

Dao-De-Jing (Lao-Zi) (in Chinese). *Lao Tzu, Tao te ching*. Translated by D. C. Lau. Harmondsworth, Middlesex: Penguin Classics, 1979.

Fung, Yiu-ming. *On Gong-Sun-Long-Zi*. Taipei: Dongda Publishing House, 2000 (in Chinese).

Fung, Yu-lan. *A History of Chinese Philosophy*, 2vols. Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1952–1953.

Gong-Sun-Long-Zi (in Chinese).

Graham, Angus C. *Disputers of the Tao*. La Salle, IL: Open Court, 1989.

Graham, Angus C. *Studies in Chinese Philosophy and Philosophical Literature*. Albany: SUNY Press, 1990.

Graham, Angus C. *Yin-Yang and the Nature of Correlative Thinking*. Singapore: Institute of East Asian Philosophies, 1986.

Hall, David. "The Import of Analysis in Classical China—A Pragmatic Appraisal." In *Two Roads to Wisdom?—Chinese and Analytic Philosophical Traditions*, edited by Bo Mou, 153–167. Chicago: Open Court, 2001.

Hall, David. "The Way and the Truth." In *A Companion to World Philosophies*, edited by in Eliot Deutsch and Ron Bontekoe, 214–224. Oxford: Blackwell, 1997.

Hansen, Chad. "Chinese Language, Chinese Philosophy, and 'Truth.'" *Journal of Asian Studies* 44 (3) (1985): 491–519.

Hansen, Chad. *A Daoist Theory of Chinese Thought*. New York: Oxford University Press, 1992.

Hansen, Chad. *Language and Thought in Ancient China*. Ann Arbor: University of Michigan Press, 1983.

Hansen, Chad. "The Metaphysics of *Dao*." In *Comparative Approaches to Chinese Philosophy*, edited by Bo Mou, 205–224. Aldershot, U.K.: Ashgate, 2003.

Harbsmeier, Christoph. "Marginalia Sino-logica." In *Understanding the Chinese Mind: The Philosophical Roots*, edited by Robert E. Allinson, 125–166. New York: Oxford University Press, 1989.

Harbsmeier, Christoph. "The Mass Noun Hypothesis and the Part-Whole Analysis of the White Horse Dialogue." In Chinese Texts and Philosophical Contexts, edited by Henry Rosemont, Jr., 49–66. La Salle, IL: Open Court, 1991.

Hu, Shih. *The Development of the Logical Method in Ancient China*. Shanghai: Commercial Press, 1922.

Mei, Tsu-lin. "Chinese Grammar and the Linguistic Movement in Philosophy." *Review of Metaphysics* 14 (1961): 463–492.

Mei, Tsu-lin. "Subject and Predicate: A Grammatical Preliminary." *Philosophical Review* 70 (1961): 153–175.

Mencius. Translated by D. C. Lau. Hammondsworth, Middlesex: Penguin, 1970.

Mou, Bo. "Truth Pursuit and Dao Pursuit." In *Davidson's Philosophy and Chinese Philosophy*, edited by Bo Mou. Leiden, The Netherlands: Brill Academic Publishers, 2006.

Mou, Bo. "The Structure of the Chinese Language and Ontological Insights: A Collective-Noun Hypothesis." *Philosophy East and West* 49 (1999): 45–62.

Mou, Bo. "Ultimate Concerns and Language Engagement: A Re-Examination of the Opening Message of the *Dao-De-Jing*." *Journal of Chinese Philosophy* 27 (4) (2000): 429–439.

Reding, Jean-Paul. "Analogical Reasoning in Early Chinese Philosophy." *Asiatische Studies* 40 (1986).

Rescher, Nicholas. *Metaphilosophical Inquiries*. Princeton, NJ: Princeton University Press, 1994,

Rosemont, Henry Jr. "On Representing Abstractions in Archaic Chinese." *Philosophy East and West* 24 (1974).

Schwartz, Benjamin I. *The World of Thought in Ancient China*. Cambridge, MA: Harvard University Press, 1985.

Wu, Kuang-ming. "Counterfactuals, Universals and Chinese Thinking." *Philosophy East and West* 37 (1987): 84–94.

Yi-Jing (I Ching). Translated by Cary F. Beynes as *The Book of Changes*. London: Routledge & Kegan Paul, 1968.

Zhuang-Zi. Translated by Burton Watson as *Chuang Tzu: Basic Writings*. New York: Columbia University Press, 1996.

Bo Mou (2005)

METAPHYSICS AND EPISTEMOLOGY

In traditional Chinese philosophy, epistemology was not an explicitly developed discipline, even if Chinese philosophers since ancient times were interested in problems related to human knowledge and developed some implicit theories of knowledge. Traditionally, there was no technical Chinese term equivalent to "epistemology" in Western philosophy, for which Chinese now use the terms "zhishilun" and "renshilun" as modern translations. In contrast, metaphysics has been a central interest of Chinese philosophy, traceable back to its origin in the *Yijing* or *Zhouyi* (The [Zhou] book of changes, c. 6th –5th century BCE). The discourse on the Way (*daolun*), in various forms, has always been an essential constituent of traditional Chinese philosophy. The term "xingershangxue," or simply "xingshangxue," now serving as the Chinese translation of the term "metaphysics" in Western philosophy, comes from the great appendix of the *Zhouyi*, where we read, "What is above forms [*xing er shang*] is the Way [*dao*]; what is under forms [*xing er xia*] is concrete things [*qi*]" (Kong, juan, p. 158). Knowledge of metaphysical reality, essential to Chinese philosophy, is also a fundamental concern of Chinese theory of knowledge.

The following discussion will first deal with Chinese theories of knowledge, ascending from ordinary knowledge to science to wisdom. All three moments have their metaphysical presuppositions, especially wisdom, which is in essence the knowledge of ultimate reality and thus leads to metaphysics properly speaking.

KNOWLEDGE

It is easy to identify some texts in which traditional Chinese philosophers discussed the subject-object structure of knowledge or the knower-known relation in the process of knowing. For example, Xunzi (298–238 BCE) said, "That by which one can know is human nature; that which can be known are the principles of things" (p. 523). Mo Di (fl. 400 BCE) said, "Wisdom (*zhi*) is the capacity

… by which, when one knows, one necessarily knows (as with eyesight)" (p. 212), and "Wisdom, by means of the capacity to know when in contact with things, enables one to describe it, like the seen" (p. 212). Unlike in Western epistemology, where the relation between subject and object or knower and known plays an essential role, in Chinese philosophy, this is only instrumental to a deeper dynamic process in which the individual attains knowledge of external things and cognitively appropriates objects in the world for building a meaningful life.

Chinese philosophers distinguished different types of knowledge, such as the Mohists' distinction between knowledge by hearsay, knowledge by explanation, and knowledge by personal experience, and Mencius's distinction between knowledge by the senses and knowledge by thinking. But more important is the Chinese concern with how to prepare the mind to know external things as they are, without bias. This can be seen in the Huanglao Daoist ideas of emptying (*xu*), unifying (*yi*), and quieting (*jing*) the mind. These notions were later developed by Xunzi as a way to attain the great clear enlightened state of mind (*da qingming*). Xunzi can be seen as the greatest theorist of knowledge in Chinese philosophy. The last master thinker in the Jixia Academy (374–221 BCE), Xunzi developed his epistemological thinking as theoretical support for scholarly argumentation in the academy, which consisted of different competing schools.

In the manner of an intellectualist, Xunzi emphasized humans' cognitive ability to discern right and wrong, which he termed "discernment" (*bian*). When expressed in discourse, this ability is displayed in what he called "discerning discourse" or "argumentation" (*bianshuo*). Xunzi conceived of the Way as the ultimate standard for discerning right and wrong, which included classes (*lei*), coherence (*tong*), and distinctions (*fen*) as subcriteria. Since things exist in different classes or categories (*lei*), their corresponding names should also be divided similarly or differently, as the case may be. The function of discourse is to make proper distinctions and classifications (*fen*) among things and names. Finally, all classifications and distinctions in discourse should be composed into a coherent system (*tong*).

To judge right from wrong well, one has to keep one's mind in a great clear enlightened state, attained by making one's mind empty, one, and still (*xu, yi er jing*), ideas that Xunzi received and developed from the Daoists, especially (370–290 BCE). Xunzi understood that when the mind is empty, "what has already been stored [in the mind] does not hinder the reception of new knowledge," that when the mind is one, "the knowledge of particular things does not hinder their unity," and that when the mind is still, "dreams and noisy fancies do not disorder one's knowing mind" (Xunzi, p. 510).

According to Xunzi, in the process of knowing and arguing for one's knowledge, one must, negatively, discard all obscuring factors and, positively, be alert to other, easily neglected aspects of an issue in dispute. Human knowledge is expressed by concepts, which, for Xunzi, are names (*ming*). Names can be analyzed according to the concepts of intentions and extensions of Western formal logic. With respect to intentions, Xunzi distinguished between names discerning superiority/inferiority and names discerning identity/difference, representing concepts respectively indicating values and facts, for him the former being higher than the latter. With respect to extensions, Xunzi made the distinction between generic names (*gongming*) and specific names (*bieming*), analyzable by reference to the relations of "inclusion" and "belonging to" between classes and subclasses. Classes can be seen as the basis of all deductive and inductive reasoning. Since the Way, as the ultimate standard for judging right from wrong, can be classified into different classes (*lei*), *lei* is imbued with both logical and ontological meanings.

SCIENCE

Before modern European science emerged in the sixteenth century, Chinese science was much more advanced than European science, as shown by Joseph Needham in *Science and Civilization in China*. Chinese philosophers were often enthusiastic about and full of scientific knowledge. For example, Zhu Xi (1130–1200), though living in the twelfth century, was well acquainted with different kinds of scientific knowledge, and is therefore a good example to showcase the philosophical import of Chinese science. Zhu Xi can be seen as the great synthesizer of medieval Chinese scientific knowledge and its philosophical foundation, even if he lived earlier than such Western medieval thinkers as Robert Grosseteste (1175–1253), Roger Bacon (1210–1292), Albert the Great (1200–1280), St. Bonaventura (1217–1274), and St. Thomas Aquinas (1225–1274). Zhu Xi's *Wenji* (Collected writings) and *Yulei* (Classified conversations) display his rich knowledge in the domains of calendrical astronomy, botany, music and harmonics, geomancy, medicine, etc. Also, he frequently discussed matters of science with his disciples, sometimes for the sake of scientific knowledge, sometimes to illustrate Chinese classical texts.

Some of Zhu Xi' observations on natural phenomena are quite interesting and true. For example, he said, "Mountains were formed by the elevation of sea bottom."

He then proceeded to prove it by pointing to the presence of seashells on top of mountains, saying, "On high mountains there are often seen shells of oyster and shellfish in the rocks. These rocks must have been earth in ancient times, and those shells from oysters and shellfish in the water. The lower becomes the higher, the soft becomes hard. This phenomenon is worthy of pondering upon, for these facts can be verified" (1999, bk. 5, p. 19). In this particular case, Joseph Needham admires Zhu Xi, writing, "Zhu Xi recognized the fact that the mountains had been elevated since the day when the shells of the living animals had been buried in the silt mud of the sea-bottom" (1959, p. 598). Note that Zhu Xi's remarks here concern mountain formation as well as fossils of sea animals. In other areas of science, Zhu Xi also correctly observed that the source of moonlight was the sun, and he correctly explained such phenomena as tides and eclipses of sun and moon.

Even if Zhu Xi was full of natural knowledge and was rational in attitude, he was not satisfied with the technical dimension of scientific knowledge and sought deeper understanding by exhausting the principles of all things and developing a holistic vision of reality. In fact, his interest in knowledge of nature should be understood in his philosophical concepts of *gewu zhizhi* (investigating things to extend knowledge). For Zhu Xi, *li*, meaning principle, reason, or order, could be found in everything and was worthy of investigation. He said, "As high as the Ultimate Infinite, the Great Ultimate, and as low as one herb, one tree, as tiny as one insect, each has its principle. … If we leave one thing uninvestigated, then we lack one principle" (1999, bk. 1, p. 295). The object of Zhu Xi's investigation was the order existing in other things, which presupposed that things and their principles possessed a certain otherness. The attainment of knowledge would include knowledge of other things and knowledge of self, or better said, a detour through the other that leads to a return to oneself, as when one finally achieves sudden penetration into the nature of things and attains transparent self-knowledge. Thus, the investigation of things is a detour in which one first goes outside one's self to the other and by knowing the other, one can finally come back to knowing one's own self.

So much for the Zhu Xi example. What is to be said about the epistemological specificity of Chinese science in comparison with European modern science? Generally speaking, Western modern science was historically grounded in the Greek heritage of *theōria*, the disinterested pursuit of truth and sheer intellectual curiosity. Aristotle said in his *Metaphysics* that science began in a way of life that included leisure (*rhaistōnē*) and recreation (*diagōgē*), such as the Egyptian priests enjoyed who discovered geometry. They did not need to care about daily necessities of life and could wonder about the causes of things and seek knowledge for its own sake. The result of their wonder was theories, whose meaning, according to Aristotle, was determined, in one sense, by practice, "not in virtue of being able to act, but of having the theory for themselves and knowing the causes" (*Metaphysics* 981b 6–7), and, in another sense, with respect to universal objects, seen by Aristotle as the first characteristic of science (*Metaphysics*, 982a 3–10, 20–23).

In contrast, Chinese science in general began as a concern leading not to universal theorization but to universal praxis. It was because of his concern with the destiny of the individual and society that Chinese began to philosophize. The great appendix of the *Zhouyi* asserts that the study of changes began with concern and anxiety over natural calamity, not in leisure and recreation. It also suggests that the practical intention of Chinese science was to serve as guidance for a universal praxis. Nevertheless, both modern European science and traditional Chinese science are concerned with the universal, or better, the universalizable, character of science, the one more with universal theories, the other more with universal praxis, yet both of them criticize and seek to transcend particular interests, with a view to attaining universality.

Because of its pragmatic concern, Chinese traditional science, in thinking about the secrets of nature, tends to use concrete images and construct concrete models for understanding natural phenomena. These images or models came directly from an intuitive or speculative vision of reality. Models in traditional Chinese science were based on analogies, that is, they were concrete models of images or small-scale models that combined the functions of explanation and pragmatic operation. For example, the construction of astronomic clepsydras (water clocks), very important in Chinese astronomy and hydraulics, expressed the genius of traditional Chinese science. Here were models that linked the movements of the heavens with the more visible movement of water or other fluid to create a visible image. In modern terms, the Chinese way of thinking in science is more analogous in form, giving birth to images and icons, which provide a more intuitive grasp of a situation in action. By contrast, the construction of models in modern European science is guided by theories presentable in mathematic form. Such models serve to mediate between mathematical theories and concrete empirical data. Modern European science, as exemplified by Newton's physics and Leibniz's

mathesis universalis (a universal science modeled on mathematics), is akin to the digital way of thinking and provides a more structured and lucidly conceptualizable understanding.

The special features of traditional Chinese science in comparison with modern European science concerning the epistemic structures involved in the process of constructing scientific knowledge are these: First, on the rational side, modern European science, in constructing theories, uses logically and mathematically structured languages to formulate theories of local validity, that is, with explanatory and predictive power in a particular domain of phenomena. In comparison, Chinese traditional science did not utilize logico-mathematical structures in its theory formation. Chinese never pondered about the structure of language to the point of elaborating a logic system for the formulation and control of scientific discourse. Mathematics, although highly developed, was used only for describing and organizing data, not for formulating theories. Chinese quasi-scientific theories, lacking logical and mathematical structure, were principally presented through intuition and speculative imagination. They might have the advantage of offering insight into the totality of life and environment and giving a reasonable interpretation of them, but these "theories" somehow lacked the rigor of structural organization and logical formulation.

Second, on the empirical side, modern European science is characterized by well-controlled systematic experimentation, which, by elaborating on the sensible data and our perception of them, keeps in touch with the real world, but in an artificial, technically controlled way. In contrast, the empirical data in traditional Chinese sciences were gathered through detailed but passive observations, with or without the assistance of instruments. Traditional Chinese science seldom tried any systematically organized experimentation to exercise active artificial control over human perception of natural objects.

Third, in modern European science, there is conscious checking of the correspondence between the rational side and the empirical side to combine them into a coherent whole so as to serve the objective of explaining and controlling the world. The rational side of science builds up a theoretical vision of the world, while the empirical side relates this vision to the scientist's sensible construction and controlled experience of the world. Philosophical reflection, in checking the correspondence between these two aspects, assures us of their coherence and unity. In contrast, traditional Chinese science did not conceive of any interactive relation such as deduction/falsification or induction/verification or tests/confirmation to relate empirical knowledge and its intelligible ground of unity. Although Chinese traditional science did have its visions of proper science and knowledge in general, it did not have modern European science's epistemological reflection and philosophy of science—disciplines that check the nature of and correspondence between empirical and rational constructs.

Still, there is unity in traditional Chinese science. Confucius (551–479 BCE) said that there is a unity binding, or a guiding thread penetrating, all his knowledge. Confucius thus seemed to affirm the complementary interaction between empirical data and thinking. He said, "He who learns without thought is confused. He who thinks without learning is in danger" (*Analects* 2.15). These words remind one of Kant's proposition that sensibility without concepts is blind, whereas concepts without sensibility are void. But we should be clear that the mode of unity in traditional Chinese science was a kind of mental integration with ultimate reality through ethical praxis. Here praxis or practical action was not the technical application of theories to control concrete natural or social phenomena. Rather, it was an active process of realizing what is proper in the life of the individual and society. Science and technology are not to be ignored, but must be reconsidered in the context of this ethical praxis.

From the analysis above, it becomes clear that traditional Chinese science should be characterized as reasonable, and not rational in the sense of modern Western science. To be rational, one has to control the gathering of empirical data through systematic experimentation, to formulate theories in logico-mathematical language, and to check the relation of empirical data and theories through philosophical reflection. By contrast, to be reasonable, one has to find meaning for human life with reference to the totality of existence. Chinese philosophy, in its quest for what is reasonable, was caught in the tension between reference to the totality of human existence and reference to the totality of all existence. Confucianism insists on referring to the totality of human existence, whereas Daoism seeks to escape from the all too human tendencies of humanist philosophy and to refer rather to the totality of all existence, as expressed by the concept of the Way (*dao*). Daoist philosophy, as a philosophy anchored in the Way and the totality of all existence, and Confucianism, as a philosophy anchored in the totality of human existence, exemplify two complementary aspects of Chinese reasonableness.

WISDOM

Wisdom is the common concern of Chinese epistemology and metaphysics. Ultimately speaking, in all Chinese philosophical traditions, wisdom is what one's knowledge should finally achieve, and wisdom in some sense always refers to what is really real, to ultimate reality. In ancient China, the same ideogram (*zhi*) was used for both knowledge and wisdom, but later a radical was added to the character to differentiate wisdom (*zhi*) from knowledge (also *zhi*). The modern term for wisdom is *zhihui*. Chinese Mahayana Buddhism, while using *zhihui*, prefers the term *banruo*, a Chinese phonetic translation of the Sanskrit *prajñā*. When Xuanzang (596–664) set up a system of regulations for his translation project, he showed a particular respect for the term *banruo* in his "five categories of terms not to be translated" (*wu bu fan*), while the Chinese term *zhihui* appeared for him to be superficial. Nevertheless, the term *zhihui* was also often used in Chinese Buddhism to express the idea of wisdom.

In Confucianism, wisdom means three things. First, wisdom means accumulating knowledge under a unifying thread or penetrating unity, as Confucius said. In this sense, knowledge comes from investigating the natures or principles of things so as to be able to unfold them according to their natures, instead of imposing theories upon them or exploiting their energy for human short-term interests. Second, wisdom means achieving total self-understanding. For Wang Yangming (1472–1529), this entails achieving one's inborn knowledge, completely developing one's true nature, and arriving at one's full potential of the moral knowledge proper to humans. Finally, wisdom means awareness of one's own destiny or heavenly mandate. Confucius took his understanding of his heavenly mandate, at age fifty, as a crucial point of his life. Also, the *Doctrine of the Mean* (c. 5th century BCE) says, "Wishing to know man, he must not fail to know Heaven" (Chan, p. 105).

In Daoism, *Laozi*, despite its critical and negative attitude toward instrumental knowledge and calculation, as shown in its negative use of the term "knowledge," nevertheless uses the term *ming*, defined as enlightened knowledge of the constant law of nature: "To know harmony is to accord with the constant; to know the constant is wisdom" (chap. 55). According to Daoism, to be wise, which is more than possessing mere intellectual knowledge, is to know the constant laws of nature, and from there, to be one with the Way and thereby to live a life of freedom, understood not as merely making free choices or arriving at autonomous decisions, but rather as complying with the spontaneous rhythms of nature.

In Chinese Mahayana Buddhism, the Chinese term *banruo* is taken to mean only perfect wisdom. This is a development of the Indian tradition, where the term *prajñā* means knowledge as well as wisdom, perfect wisdom as well as imperfect wisdom. In Chinese Buddhism, wisdom means attaining enlightenment, a state in which one understands that all is empty and thus seeks to rid oneself of original ignorance. The term *zhihui* (wisdom) was used to translate *prajñā*, especially in Weishi's (for example, in Xuanzang's Cheng Weishi Lun, 659 CE) concept of transforming consciousness into wisdom (*zhuan shi de zhi* or *zhuan shi cheng zhi*). For the Weishi School, more Indian than Chinese, wisdom arises from Alaya consciousness (Alya vijñāna). But for the Sanlun School, wisdom means realizing the ultimate emptiness of the world. In Chan (Zen) Buddhism, wisdom is the immediate self-realization of Buddhahood in the details of everyday life.

In his *Banruo wuzhi lun* (Wisdom as nonknowing), Sengzhao (383–414) distinguished wisdom from common knowledge. For him, knowledge is epistemologically structured by the relation of the knowing subject and known object, and therefore is relative and limited to a particular object. The content of knowledge is expressed in logical propositions that should be free of logical contradiction. In contrast, wisdom is all-knowing and comprehends all things, including itself. Therefore, it lacks subject-object structure and is not limited and relative to any particular object. Its self-awareness results from its own crystal-clear mirroring and not from any self-reflection or intuition. For Sengzhao, wisdom was a mysterious function of a mind characterized by emptiness, and emptiness he identified with ultimate reality, which belongs to the ontology and therefore is beyond all logical considerations, including the principle of noncontradiction. For Sengzhao, wisdom was absolutely pure and was beyond all sorts of delusions arising from relative knowledge.

Jizang (549–623) developed a typology of three types of wisdom. First was ultimate wisdom (*shixiang banruo*), which penetrates into ultimate reality, or the emptiness of all things. This is the ultimate ground of the other types of wisdom. Second was illumining wisdom (*zhengguan banruo*), which throws light upon the ultimate reality in all its different facets and manifestations. In this application of ultimate wisdom in meditating on the essence of each and every thing, one comes to see that each of them is empty. Third was linguistic wisdom (*wenzi banruo*), which enables one to give powerful linguistic expression to the perfect congruence between ultimate reality and its manifestations.

METAPHYSICS AS KNOWLEDGE OF ULTIMATE REALITY

Metaphysics concerns knowledge of ultimate reality. Even if all the schools of Chinese philosophy used "dao" (the Way) as a common term to refer to ultimate reality, there were other terms used in different schools, even different terms used by different philosophers within one school. For example, in Confucianism, different Confucians took the concepts of heaven, humanness, sincerity, and principle or reason as ultimate reality. In the following sections, we will see what different schools took as ultimate reality: heaven, humanness, sincerity, and principle in Confucianism, the Way in Daoism, and emptiness in Buddhism.

ULTIMATE REALITY IN CONFUCIANISM. Generally, the concept of ultimate reality in Confucianism moves from heaven (*tian*), a residue from ancient Chinese religious beliefs; to humanness (*ren*) in Confucius himself; then to sincerity (*cheng*) in Zisi, Confucius's grandson; and to mind (*xin*) or principle (*li*) in neo-Confucianism. In the prephilosophical tradition, the *Shijing* (Book of odes) and the *Shangshu* (Book of documents) used the concept of heaven, imbued with a religious sense, to represent God on High. A residue of this notion could still be found in Confucius when he said, "If heaven wished to destroy this legacy, we latecomers would not have access to it. Since heaven is not going to destroy this culture of ours, what can the people of Kuang do to me?" (*Analects* 9.7). Confucius also said that he prayed to heaven, yet heaven, though manifesting itself through regular cosmic movement, remained silent, thus maintaining a certain unfathomability. Confucius said, "Does heaven speak? And yet the four seasons turn, and the myriad things are born and grow within it" (*Analects* 17.19).

Confucius's proper contribution consisted in proposing the concept of humanness (*ren*) as a transcendental foundation for ritual (*li*). Humanness, a transcendental capacity in each person, had an ontological dimension in that it presupposed that all beings are interconnected, and this allows humans to be affected and respond to other people and things. Confucius considered this transcendental capacity of each person to affect and respond to others as the transcendental foundation of ritual. Sometimes humanness was combined with the Way to specify the way of humanness. With this metaphysical move, the concept of *ren* achieved metaphysical status in neo-Confucians such as Zhou Dunyi (1017–1073) and Zhang Zai (1020–1077), who extended humanness to the whole cosmos (a cosmic humanness), surely a metaphysical concept. Also, Zisi (493–406 BEC), Confucius's grandson, developed Confucius's idea of ultimate reality in *Zhongyong* (*Doctrine of the Mean*) with the concept of sincerity (*cheng*), which had two levels of meaning: On the psychological level, *cheng* meant being true to one's own self; on the metaphysical level, *cheng* meant the really real, truth, or reality itself.

Under the influence of Tiantai Buddhism and Chan Buddhism, idealist neo-Confucians such as Lu Xiangsan (1139–1193) and Wang Yangming took mind imbued with moral values to be ultimate reality. Such a mind was attainable through moral practice and moral effort. They thereby laid the foundation for a kind of moral metaphysics. In affirming that the Great Ultimate is principle or reason (*li*), the realist neo-Confucian Zhu Xi took principle or reason to be ultimate reality. For Zhu Xi, even if everything has its own principle, by way of metaphysical participation they share their reality with the cosmic principle that ultimately governs the whole world.

ULTIMATE REALITY IN DAOISM. Daoism coherently used "Way" (*dao*) as a metaphysical concept to denote various levels of metaphysical reality and ultimate reality itself. Etymologically, the ideogram for *dao* is composed of two elements, the head and the act of walking on a way. Together they mean a way on which one could find direction and a way to some point. Though *dao* was never limited to the idea of a physical way, this image of a way suggests the meaning of *dao*: The *dao* puts everything on its way. In common use, *dao* also means "to say," "to speak," or "to discourse," such as the second "dao" in the opening of *Laozi*, which says, "The way that can be spoken of is not the constant Way" (chap. 1). In Daoism, the function of discourse is always negative. Discourse, once said, must be hushed; words, once written, must be erased. One can never discourse about ultimate reality in any human language. This is quite different from Western philosophy, from the beginning of which emphasis has been on the function of language, of logos, to express reality. Apart from these two levels of meaning, *dao* in Daoism has three other uses that are more philosophical:

First, it refers to laws of becoming or laws of nature, especially in the term *tiandao* (the Way of heaven) or *tiandi zhi dao* (the Way of heaven and earth). In Daoism, the laws of nature have two aspects: (1) The structural law says that all things are structurally constituted of different yet complementary elements, such as being and nonbeing, yin and yang, movement and rest, weak and strong, and so on. (2) The dynamic law says that once a state of affairs has developed to the extreme in a process of

change, it will naturally move to its opposite state of affairs.

Second, it refers to the origin giving birth to all things. If all things are regulated by laws of nature, there must be an origin that gave birth to all things, there must be a cosmic law. Normally, the origin gives birth to all things in a process of differentiation and complexification, as indicated by these words in *Laozi*: "The Way gave birth to one. One gave birth to two. Two gave birth to three. Three gave birth to all things" (chap. 42).

Finally, it refers to ultimate reality. The Way ultimately represents the ever self-manifesting act of existence. If there is an origin giving birth to all things, then before the origin, there must be a self-manifesting act of existence, defined in relation to all things. The self-manifesting act of existence is reality itself, whereas everything we say about the Way is but a constructed reality, which can never be reality itself. One can mention the Way to express something about it, but what is said becomes a constructed reality and not reality itself. To keep one's mind open to reality itself, all human constructions stand in need of further deconstruction.

Most of the time in traditional Daoism, these three levels of the Way were closely related one another, so closely that they were often mixed up and seldom clearly distinguished in the texts. It is with philosophical effort that they can be analyzed into clearly distinguished aspects of a well-connected whole. This is to say, in traditional Daoism, ultimate reality and its multifaceted manifestations can be logically distinguished but are not ontologically distinct.

ULTIMATE REALITY IN BUDDHISM. Chinese Mahayana Buddhism, like Indian Buddhism, takes emptiness as ultimate reality. Although the Sanskrit term *śūnyatā* has many meanings in the Indian tradition, its Chinese equivalent *kong* has three major philosophical meanings, each with its own focus: First, on the ontological level, emptiness means that all things come and go through interdependent causation and therefore lack any self-nature or substance of their own. Second, on the spiritual level, it means that the spiritual achievement of the sage consists in total freedom, attaching himself neither to being nor to nonbeing, neither to dualism nor to nondualism, not even to any form of spiritual achievement, no matter how high or deep it is. To keep one's spirit totally free, one must even empty the emptiness. Third, on the linguistic level, emptiness means that all the words we use are artificially constructed, without any fixed correspondence or reference to reality.

Although Indian Buddhism put more emphasis on the ontological and the linguistic senses of emptiness, Chinese Mahayana Buddhism, generally speaking, emphasized mostly the spiritual sense of emptiness. For example, in the *Buzhen kong lun* (On the emptiness of the unreal), Sengzhao, appropriating Daoist philosophy, interpreted emptiness as the spiritual achievement of a sage (though he also gave other meanings to the term "emptiness"). For example, we read, "The sage moves within the thousand transformations but does not change, and travels on ten thousand paths of delusion but always goes through. This is so because he leaves the empty self-nature of things as it is and does not employ the term 'emptiness' to make things empty" (Chan 1963, p. 356, with corrections).

The spiritual achievement of a sage, who has no attachment to the realm of either being or nonbeing, not even any attachment to his own spiritual achievement, results from a mysterious function of his mind, which on the one hand is nonsubstantial and empty, yet on the other hand is mysterious in function and self-transcending. Because of this, the Way (emptiness as the ultimate reality) and sagehood are not far away from us and can be realized at the moment of enlightenment. "Things when touched become real. … Man when enlightened becomes mysterious" (Sengzhao, vol. 45, pp. 152–153). The idea of a mysteriously enlightened mind rendering real all things touched by it significantly influenced other Chinese Mahayana schools, especially Tiantai and Chan. In Tiantai and Chan Buddhism, the mind was taken to be ultimate reality.

INBORN KNOWLEDGE AND MORAL METAPHYSICS

Idealist neo-Confucianists, such as Lu Xiangsan and Wang Yangming, considered moral knowledge as inborn and the realization of moral knowledge to be the only access to the really real. They were idealists in the sense that they took mind as the ultimate reality, identifying the human mind and the cosmic mind, which they saw as the ontological source of all values and moral knowledge. For them, knowledge meant mainly moral knowledge and was therefore value-laden. Since moral knowledge comes from the mind, humans must be capable of knowing it before all empirical knowledge. As a kind of innate knowledge, it is to be realized through human moral effort and moral practice, called "realization of innate knowledge" by Wang Yangming. Innate moral knowledge is like a permanent light within everyone, arising before the emotions. The individual realizes it by overcoming

selfish tendencies, and thereby arrives at ultimate reality. Morality was thus considered a pragmatic way to access ultimate reality, and thus had metaphysical import.

Inheriting this line of thought, Mou Zongsan (1909–1995), a well-known figure in modern Confucianism, proposed the idea of moral metaphysics (*daode de xingshangxue*). He distinguished between moral metaphysics and the metaphysics of morals, the latter being a metaphysical study of morality and therefore moral philosophy rather than metaphysics. His idea of moral metaphysics represented an effort to emphasize the role of Confucianism and moral actions in Chinese metaphysical thinking. He also distinguished between the moral metaphysics of Confucianism and the liberation metaphysics (*jietuo de xingshangxue*) of Daoism and Buddhism. Even for Mou, these three traditions of Chinese philosophy saw the human mind as capable of intellectual intuition (*zhi de zhijue*), yet he preferred the Confucian way of attaining ultimate reality through moral practice and moral self-awareness. He thought that humans could achieve intellectual intuition through moral action and realize the noumenon of humanness (*ren*), which represented for him the ultimate reality or the thing in itself. Sometimes Mou named it "the free infinite mind/heart," or "the true self," that, as noumenon, possessed universality, infinity, and creativity, and through a process of self-negation similar to Johann Gottlieb Fichte's "I" positing a "non-I," it could unfold itself into a world of phenomena. In Mou's philosophy, intellectual intuition is an act of self-awareness of the free infinite mind, which replaced the concept of a personal God in Christianity, Islam, and Judaism.

Mou's moral metaphysics, by making Confucianism a kind of metaphysics, and thus making Confucian moral praxis an instrument for attaining ultimate reality, neglected the proper value and practical methods of Confucian moral praxis. Also, he considered morality a matter of finding one's true self, without relation to others, and thus without a proper ethical dimension. In this way, Confucianism lends its own weak points to a grand metaphysical system modeled after European philosophy. Also, by positing such an exclusively moral metaphysics, Mou neglected other metaphysical experiences, such as those in encounters with nature, in artistic creativity, in religious piety, and in historical encounters—all so rich in metaphysical implications in traditional Chinese culture. In his absolute idealism, Mou blurred and even confused the distinction between reality itself and constructed reality.

METAPHORICAL METAPHYSICS

Chinese philosophical traditions such as Confucianism, Daoism, and Buddhism all hold that ultimate reality, whatever its name, always has an unfathomable dimension and therefore is hidden from all human constructions and human languages. For this reason, the terms Chinese philosophers use to indicate ultimate reality—terms such as *tian* (heaven), *ren* (humanness), *cheng* (sincerity), *dao* (the Way), the mind, principles, emptiness, etc.—are used metaphorically rather than descriptively or ostensively. They express ideas about ultimate reality with a certain tangible image of it, which is to say that they are in some sense image-ideas, instead of pure ideas. Chinese philosophers, when grasping ultimate reality with enlightening insight, tend to form original image-ideas, something between a pure idea and an iconic image, thereby retaining the holistic character of the manifestation and the intuitive nature of the perception. This idea-image evokes the richness of ultimate reality without exhausting it, and therefore has the status of a metaphor.

This basic characteristic of Chinese metaphysics provides foundations for Chinese artistic, moral, and scientific practices and historical actions. Artistic creativity, by imagination and poetic transformation, renders this idea-image into a concrete iconic image and thereby materializes it. In moral and ethical reasoning, practical reason brings the idea-image to bear on an ethical situation, leading one to intervene and thereby take moral responsibility. In science, natural philosophers built models with reference to image-ideas, creating analogical images of reality so as to grasp natural processes in an organic and holistic way. In the historical arena, one can discern, by referring to idea-images, traces of notions of ultimate reality in the historical events and actions taken by historical agents. In this sense, Chinese art, ethics, science, and history are imbued with metaphysical significance.

Generally speaking, metaphor allows us to see one thing as something else. In other words, metaphor has an "as-structure," a term first used by Martin Heidegger to characterize interpretation. In the Chinese tradition, metaphysics or discourse on the Way is already a metaphorical interpretation of ultimate reality. Compared with the original manifestation of ultimate reality, various ways of realizing idea-images also possess an as-structure, in the sense that they allow us to see ultimate reality as idea-images, the later thereby serving a certain metaphorical function. In this sense, Chinese metaphysics can be characterized as a kind of metaphorical metaphysics. Viewing it in this way, one can achieve a true understanding of the spirit of Chinese philosophy.

See also Chinese Philosophy: Language and Logic; Chinese Philosophy: Religion.

Bibliography

Aristotle. *Metaphysics*. In *The Complete Works of Aristotle*, Vol. 2, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Chan, Wing-tsit, trans. and ed. *A Source Book in Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1963.

Fang, Thomé. *Chinese Philosophy: Its Spirit and Its Development*. Taipei: Linking, 1981.

I Ching, or Book of Changes. 2 vols. Translated by Cary F. Baynes from the German version of Richard Wilhelm. New York: Pantheon Books, 1950.

Kong Yingda. *Zhouyi zhengyi*. In *Shisanjing zhushu* (Commentary on the thirteen Confucian classics), Vol. 1. Taipei: Yiwen yinshuguan, 1985.

Lau, D. C., trans. *Confucius: the Analects*. London: Penguin, 1979.

Lau, D. C., trans. *Lao Tzu: Tao Te Ching* London: Penguin, 1963.

Mo Di. *Mozi*. Taipei: Guanwen shuju, 1965.

Mou Zongsan. *Zhi di zhixue yu Zhongguo zhexue* (Intellectual intuition and Chinese philosophy). Taipei: Taiwan shangwuyin shuguan, 1971.

Mou Zongsan. *Xinti yu xingti* (The substance of mind and the substance of nature). Taipei: Zhengzhong shuju, 1978.

Mou Zongsan. *Yuanshan lun* (On supreme good). Taipei: Taiwan xuesheng shuju, 1985.

Needham, Joseph. *Science and Civilization in China*, Vol. 3. Cambridge, U.K.: Cambridge University Press, 1959.

Robinson, H. Richard. *Early Mādhyamika in India and China*. Delhi: Motilal Banarsidass, 1976.

Shen, Vincent (Shen Tsingsong). "Laozi de xingshang sixiang" (Laozi's metaphysics). *Zhexue yu wenhua* 15 (12) (1988): 814–822.

Shen, Vincent (Shen Tsingsong). "Zhuangzi de daolun: Dui dangdai xingshang kunhuo de yige jieda" (Zhuangzi's discourse on the Way: Answers to contemporary metaphysical questions). *Guoli zhengzhi daxue zhexue xuebao* 1 (1994): 19–34.

Shen, Vincent (Shen Tsingsong). *Wuli zhihou: Xingshangxue de fazhan* (After physics: The development of metaphysics). Taipei: Niudun, 1987.

Sengzhao. *Zhaolun* (Treatises of Zhao). In *Taishō shinshū daizōkyō*, edited by Takakusu Junjirō and Watanabe Kaigyoku, Vol. 45. Tokyo: Taishō Issaikyō Kankōkai, 1924–1932.

Tang Junyi. *Zhongguo zhexue yuanlun: Yuandao pian* (Inquiry on Chinese Philosophy: Inquiry on the Dao). 3 vols. Hong Kong: New Asia Institute, 1973.

Xunzi. *Xunzi*. In *Er shi er zi* (Works of twenty-two masters), Vol. 4. Taipei: Prophet Press, 1976.

Zhang Dainian. *Key Concepts in Chinese Philosophy*. Translated by Edmund Ryden. New Haven, CT: Yale University Press, 2002.

Zhu Xi. *Zhuzi yulei* (Classified conversations of Master Zhu). Books 1 and 5. Beijing: Zhonghua shuju, 1999.

Vincent Shen (2005)

RELIGION

The subject of the religious dimensions of Chinese philosophy covers both a vast time period—at least two and a half millennia—and a vast array of religious traditions, including theistic religions like Islam and Christianity. This entry, however, will focus on only a few topics and on two indigenous traditions—Confucianism and Daoism—and those streams of Buddhism often, if controversially, said to be most characteristically Chinese, such as the Chan (Japanese Zen) tradition. These traditions not only share features but adherents of each, even fierce adherents, often adopted ideas and practices from the other traditions in ways that can seem disconcerting to people familiar with only Western religions. (Thus, the truth in the clichés that a person can be a Confucian at work and a Daoist at home or that Chinese Buddhists often employ Confucian ethical ideas.) This phenomenon raises interesting philosophical questions about the meaning, in China, of a religious tradition as well as about the character of an adherent's structure of beliefs, but this entry will treat them only obliquely in what follows.

Another, perhaps more vexing concern is the nature of the relationship between religion and philosophy in China. This subject has often been examined, in both China and the West, in a way that basically reflects the desire to guarantee that Chinese philosophy has none of the baleful qualities that characterize religion (religion in this context usually means folk religion, or put more baldly, superstition and magic). The relationship is, however, complicated, and one can best approach it by discussing the notion of religion as it affects this subject.

THE NOTIONS OF RELIGION AND RELIGIOUS THOUGHT

The attempt to define the phenomenon of religion and religious thought (and thereby also specify the forms of various religions and their processes of thinking) is a modern Western project—and one with many critics. It often combines attempts to map out a sphere of human life in a reasonably objective fashion with a desire to improve human life, which usually means to make human affairs more rational. The treatment of Chinese religion, a notoriously messy phenomenon or set of phenomena, exemplifies most that is bad and good in the project.

One is, however, interested in only one facet of that treatment, and one can begin this investigation by turning to Clifford Geertz's (1973) sophisticated and

immensely influential account (or definition) of religion. For him a religion is: "(1) a system of symbols which acts to (2) establish powerful, pervasive, and long-lasting moods and motivations in men [an ethos] by (3) formulating conceptions of a general order of existence [a worldview] and (4) clothing these conceptions with such an aura of factuality that (5) the moods and motivations seem uniquely realistic." Geertz comments at length on this definition, but of special importance is his claim that "a group's ethos is rendered intellectually reasonable by being … adapted to the … world view … while the world view is rendered emotionally convincing by being presented as an image of [the ethos]" (pp. 89–90).

This account seems to fit much Chinese religious thought, even to pinpoint crucial features of it. This is especially true if one focuses both on how Geertz exhibits the porous boundaries between elite and popular attitudes and, especially, on how he develops the ways his five elements interact with each other to create a closed perspective, a seamless web of reflection and reinforcement, in which ethos and worldview interact. Daoism is replete with stunning examples of these processes, but they also appear prominently in other traditions.

Moreover, and of special significance for this entry, the religious thought that results is detailed and can even count as rigorous, given its premises and notions of entailment. Furthermore, it often involves cosmological or cosmogonic subjects, many of which include science-like accounts of both natural and human phenomenon. (Correlative thinking, which finds homologies between the natural and human to control the latter through aligning it with the former, is only one of the most prominent and famous instances.) Finally, much of this thinking depicts a universe that is well, even fabulously, endowed with a great variety of beings (e.g., see Wang Bi 1994; Ko Hung 1966; Robinet, 1997, pp. 115–148, 195–256.).

Such thinking shares, however, few of the assumptions and notions that would make it credible to one's common contemporary experience, much less to one's understanding of philosophy. That is, this thinking cannot be formulated in a way that meets the conditions of plausibility found in an experience informed by, for instance, modern scientific explanation, historical consciousness, and ideas about the rights of all humans. Moreover, it makes difficult any understanding both of changes within a tradition, even if such changes were not always clearly recognized by participants, and of reasoned conversations among adherents of different traditions.

A focus on these kinds of Chinese religious thought, a common focus among scholars today, surely can be justified as a valuable and integral kind of historical inquiry. Nevertheless, this kind of focus also underlies the common perception that almost all Chinese philosophy lacks real analytical rigor, sophisticated modes of inquiry, and evident ways in which to attend to significant alternative views and therefore to reflect critically on given social forms.

The aim here, in contrast to the kind of historical inquiry just noted, is to show the ways in which many religious dimensions in Chinese thought are imperfectly captured by such criticisms and therefore have a claim on one's philosophical attention. Indeed, by examining three topics at length, and mentioning several others, this entry will illustrate how one can enrich an understanding of both philosophy and religion that, understandably, arises mainly from Western examples. One can, that is, be involved in the distinctive sort of intellectual exercise in which one tries to place oneself within a world that is much larger than the world one normally inhabits, a world the modern situation has, thankfully, forced on one. Before beginning that, however, one needs briefly to examine a different approach to the idea of religion.

The word *religious* refers of course to many phenomena, but most important in this case is the reference to an orientation that differs from and judges many features of the ordinary world, even if it also underlies other features of that world. (It is the latter feature that Geertz's [1973] examination usually emphasizes.) The religious, in this sense, rejects commonplace approaches to human fulfillment because normal life contains too many apparently insurmountable difficulties and because a marvelous spiritual actualization is possible.

The dialectic of great need and grand fulfillment means this orientation fits within what can be called a discontinuous or nonameliorative type of religion. This type of religion is, in many ways, fundamentally discontinuous with the activities and expectations of normal life; it seeks far more than just to ameliorate the problems ordinary life produces, and therefore it makes a substantial break with normal life rejecting attempts to build on what is already present. In contrast, in an continuous or ameliorative religion people work within the framework normal life provides. Their aim is to deepen and extend the best ideas and practices people have. The latter applies bandages to what are perceived as minor wounds, while the former calls for major surgery. The latter, of course, labels that kind of major surgery mutilation, while the

former labels the mere application of bandages malpractice.

The distinction between discontinuous and continuous religion is not a simple binary one. Rather, the two types define the ends of a continuum and specific features of a religious tradition will fit at places within the continuum. Significant features of Confucianism, for example, fit on the continuous end and significant features of early Daoism and Chan Buddhism on the discontinuous end. (Moreover, many of the most substantial debates among and within these traditions are illuminated by understanding where on the continuum a disputed feature, such as meditational activity, fits.) Nevertheless, all of them have pronounced discontinuous elements, and they appear prominently in their most able philosophical thinkers (e.g., see Mencius 1970, Book 2A; Chuang-Tzu 2001, pp. 76–93; Hui-neng 1967, #28–#47, #149–#174; Graham 1958, pp. 67–91).

These traditions may lack many of the discontinuous qualities that characterize theistic religious, but they also share other features, if often manifesting them in a distinctive way. Put schematically, three discontinuous religious elements are especially important in these traditions. One is a focus on a sacred realm that is related to but differs dramatically from the human realm. This realm provides thinkers with a perspective from which they can evaluate ordinary activities in ways that most people find perplexing at best and insulting at worst. Another is a belief that kinds of empowerment occur that exceed what appears in ordinary life, are crucial to people attaining any true flourishing, and can produce people who transcend the limits of ordinary understanding. Specifying exactly how this empowerment operates and how sacred and human realms interact, even how independent they are, is difficult enough, however, to demand special uses of language. (As it will be seen, despite its theoretical imprecision the needed language aids rather than impedes the fundamental spiritual discipline of self-cultivation.)

A final discontinuous element is the distinctive quality of members' adherence to the traditions of which they are a part. These thinkers recognize that traditions contain regenerative powers individuals alone could not produce and yet also can be a source of debilitating false fixities. This recognition leads them both to treasure their traditions and to be extremely sensitive to the dangers present in false teachings, misleading authorities, and the communities that gather around them.

TOPICS IN CHINESE RELIGIOUS THOUGHT

One can now turn to three subjects that illustrate the distinctive contribution to philosophy that religious elements in Chinese philosophy can provide. One is the subject of ritual, a second the differences and similarities between normal and religious excellences, and a third the need to employ various genres to present religious realities persuasively, a topic that helps one understand some distinctive, formal features of Chinese philosophy. Before beginning that examination, however, one should briefly note six other topics that illustrate the range of pertinent material that could be examined. This entry will, if telegraphically, describe each of the six; provide one paradigmatic, and accessible, instance from early Chinese thought; and then note resemblances to Western materials, thereby risking the embarrassment brief generalizations can produce.

(1) In each tradition one finds the belief that skillful people, and thus the idea of skillfulness, manifests a religious excellence that tells one much about the character of perfected action, thought, and selfhood. Little that resembles this focus, and the understanding that results from it, has ever appeared in Western philosophical discussions (Chuang-Tzu 2001, pp. 62–65, 135–142).

(2) Each tradition treats, if often with different results, the question of participation in or retreat from social or political involvement, seeing it as a choice that can be understood philosophically only if one grasps the religious dimensions of each alternative. A version of the question has, of course, been central in the West, but its religious dimensions—and what arises from considering them—has been far less central (Mencius 1970, Books 4A, 5B).

(3) Each tradition sees the purported religious excellence of aimless wandering (*you*), powerfully presented in early Daoism, as an ideal that raises the deepest philosophical questions about ordinary ideas of intention and responsibility. The absence of a similar ideal, and thus of the resulting questions, means certain religious challenges to basic ideas about human purposes and obligations are never fully engaged in the West (Chuang-Tzu 2001, pp. 43–47, 66–75).

(4) An agnostic posture toward many central religious notions is understood by many thinkers in these traditions to be a mark of true spiritual achievement. Neither the implications nor the gen-

eral importance of this posture are probed in the Western as they are in China, although the attitude surely is not absent (Xunzi 1994, pp. 3–32. 88–112).

(5) Humor is a crucial religious excellence in early Daoism (as well as in the Chan tradition) and the philosophical import of the perspective humor generates is illustrated constantly and occasionally even analyzed at length. With some notable exceptions, humor is rarely a central subject in especially Western theistic traditions (Chuang-Tzu 2001, pp. 122–125, 207–210).

(6) All these traditions provide myriad illustrations of the ways in which commentaries on texts thought to be religiously authoritative constitute a, perhaps the, major way in which philosophical thinking is both motivated and constrained. The resemblance to traditional practices in the West is close here, but both the elusive character of the Chinese texts commented on and the character of the constraints on inquiry provide illuminating insights (Confucius 2003, Books 4–7).

Any of these six topics could productively be examined at length, but now the focus of this entry will shift to three topics that are especially illuminating for one's inquiry: ritual, the relationship between normal and religious virtues, and the genres needed to present religious realities persuasively.

RITUAL

Ritual, probably the most adequate of the multitudinous translations of the character *li*, is surely among the most distinctive and complex of all Chinese notions. Put simply, the single notion covers two activities that most contemporary Westerners think are quite different. One activity is solemn, explicit religious activities such as marriage or internment services. The other activity, however, is what can be called etiquette or, more accurately, those reasonable and humane learned conventions that make up the ethos of a culture (e.g., see Xunzi 1994, pp. 49–73; Chuang-Tzu 2001, pp. 87–93; Robinet 1997, pp. 166–183; Gregory 1991, pp. 41–43, 274–285; Ching 2000, pp. 72–90).

Ritual covers, then, everything from the solemn performance of an elaborate ceremony to the "excuse me" after a sneeze. Explicit religious activities and social activities are, that is, part of one continuum, although there are, of course, notable differences. In specifically religious rituals, for example, the focus is on humans facing thresholds, situations where people move to a new state or respond to what lies beyond their ordinary routines.

The combining of these two senses of ritual is open to the criticism that it displays an unsophisticated kind of thinking that fails to differentiate what can and should be separated. The defense of the combination, one most evident in Confucianism but present in the other two traditions, rests on the notion that social rituals are more than just pedestrian social facts. Social and religious rituals resemble each other, that is, because both are sacred ceremonies that express and foster a spontaneous coordination that is rooted in reverence. Moreover, both exemplify learned, conventional behavior that manifests distinctly human rather than simply animal-like actions. Both therefore promote crucial human qualities and respond to central human needs.

Ritual is, then, a notion of overarching significance in Chinese religious thought and contests about its character and value are frequent. In fact, debates about ritual often served to focus debates among competing visions of life. Seeing, therefore, the various views of ritual (social and religious) that continually appear, if in somewhat different forms in different times and traditions, can help one understand the philosophical import of the idea.

Especially prominent in these debates are three kinds of attacks on ritual and three defenses of it. Put telegraphically, the different positions are as follows. One group attacks rituals as a wasteful, even unjust, use of scarce natural and human resources. Another group attacks them as a social artifice that distorts significant human capacities and reinforces destructive social organizations. A third group sees them as an inadequate form of social control that is best replaced by clear rewards and punishments.

One defense of ritual sees in it a process that activates transhuman forces and uses those forces to help humans. Another justifies ritual in terms of the innate human capacities for it or even inertial tendencies toward it; human beings, that is, need ritual if they are to be actualized. A last defense believes rituals are sanctified by tradition; they therefore need no real justification and must always be meticulously followed no matter what the apparent price.

Many, although surely not all, of the most sophisticated thinkers from these traditions think each of these approaches is flawed, and they therefore reinterpret religious rituals both to win outsiders' assent and to deepen their own and other adherents' assent. They usually proceed, to focus just on religious rituals, by distinguishing

among three different kinds. First are rituals that are useless or even harmful; sacrificing a pig to cure an illness falls in this class. Second are rituals that adorn life in important but not critical ways; rituals to produce rain fall in this class. Third are rituals that provide a crucial service to human life; death or internment rituals fall in this class.

Such reformulations manifest a set of common characteristics, and they are worth noting because of what they tell one about attitudes to religion. Most generally, the overarching goal of all these reformulations is the protection and encouragement of fully flourishing human activity. That goal provides the criterion both for dividing necessary from unnecessary religious ideas and actions and for reforming the meaning of the necessary ones. Second, these reformulations critically examine all simple anthropomorphic descriptions of the transhuman realm and replace them with designations that are symbolic or stress the mysterious. Third, if closely related to the second, they criticize depictions of activities that describe a manipulative relationship between the human and the transhuman. In fact, they often redescribe those activities in terms of how feelings are rearranged and spiritual attitudes are generated.

The grounds these thinkers use to defend ritual tell one much about the role of religion in Chinese philosophy. They usually, that is, presume that one is frail in ways that often are difficult for one to accept. Not only does one live between origins and terminations one cannot control but one also constantly faces the numinous. Moreover, one is prone to primordial reactions, and one must treat them in a fashion that both controls their destructive side and nurtures their constructive side. Stringent limits, then, define what people can do; they cannot immediately form themselves into what simple rational judgment might commend them to be. Ritual roles present, for instance, roles that people have no real choice but to assume, with the role of the mourner, however defined, being perhaps the clearest instance.

Underlying this perspective is a negative judgment about a philosophical approach that rejects ritual because it desires to produce a rigorous and coherent picture of the world that will provide simple, reasonable grounds for ideas and actions. Proponents of this approach reject internment rituals, for example, because to them the principle of noncontradiction is crucial; a person is either dead or not dead. They want to face life and death directly and they put everything into clear-cut categories.

Against such an approach, it is argued that when life is seen clearly, and in the death of others one sees it especially clearly, one can neither make it into a coherent understandable whole, nor respond adequately to it by focusing on simple practical expedients. Human life is too fragile and delicate, too complex and contradictory to capture in simple rational systems. One touches life as it actually is, then, only through the complex pathos, the human contradictions, the struggle to find peace, and the openness to the numinous that rituals exemplify. A related but different kind of judgment on some kinds of ordinary philosophical attitudes underlies the second subject, and to that one may now turn.

RELIGIOUS AND ORDINARY VIRTUES

To distinguish between religious and ordinary virtues (or excellences) in Chinese thought might seem to be problematic or even just wrong-headed, but investigating the subject can illustrate, among other things, the usefulness of examining apparently inapplicable Western ideas in the Chinese context. The relevant Chinese thinkers never, of course, make any formal distinction between normal and religious virtues, except when discussing those virtues that bind only those adepts who adopt monastic rules. Moreover, their general conceptual framework does not lead them (and probably literally could not allow them) to distinguish between what, say, Catholic Christianity calls natural and supernatural virtues. They surely would, that is, reject any distinction that rests on a clear-cut differentiation between what humans cause and what a deity, distinguished by the quality of aseity or being unmoved, causes.

Nevertheless, a crucial feature of much Chinese philosophy is the conviction that some virtues (or unnamed features of some virtues) have a special character. They produce actions and attitudes that both differ from normal virtues and change a range of normal actions in profoundly important ways. In fact, a number of traditional Western ways of theoretically distinguishing religious and normal virtues seem to be implied. Examples include sharp distinctions made among the kinds of objects pursued, among the goals of the intentions manifested, among the precise forms of behavior produced, and among the kinds of empowerment displayed. Moreover, and perhaps even more striking, many think humans are susceptible to transformations so total as to make some individuals fundamentally different from the rest, to make them, for example, the possessors of truly extraordinary abilities to affect the natural and human world (e.g., see Chuang-Tzu 2001, pp. 96–99, 143–150; Yearley 1990, pp. 144–168; Gregory 1991, pp. 255–274; Graham 1958, pp. 96–118).

A simpler example of such transformations concerns the role of distinctive kinds of belief in adherents' lives, a role that both resembles and differs from the role of faith in some theistic traditions. These beliefs go considerably beyond the evidence that would, and should, compel assent in normal affairs, and they often play a prominent role in guiding action. They include beliefs about the significance of certain books and historical figures, but most revealing may be beliefs about the role of some virtues and perspectives.

Two straightforward but illuminating examples of these latter beliefs come from what is probably the earliest part of the earliest (and arguably most important) book in Chinese thought, Confucius's *Analects* (Lunyu): "Virtue (*de*) never dwells in solitude; it will always bring neighbors" (4, 25), and "The Master said 'In the morning hear the Way (*Dao*); in the evening, die content'" (4, 8). Each passage represents a dramatic enough claim to be considered religious, as well as, of course, a claim that can be, and was, probed philosophically.

Indeed, beliefs like these are often at the center of debates with those people who lack them because they find them either unintelligible or unjustified. The religious perspectives that define each of these traditions, that is, are far from self evident to everyone. They include attitudes and confident judgments about many matters that seem problematic or even bizarre to many people outside the tradition. Moreover, adherents within a tradition also entertain questions about their own beliefs; they are not inoculated against the queries and doubts that other people manifest. One crucial spiritual dynamic in all these traditions, then, is to see obvious problems in their own position, if one uses either ordinary standards or another tradition's standards, and yet continually to reaffirm specific, central beliefs.

These Chinese ideas on religious beliefs, as well as on other virtues, reflect an ontological perspective in which the sacred realm and the ordinary realm are closely intertwined, in which an organismic, an interrelated and interdependent, cosmology operates. Indeed, Chinese religious thought manifests in its own fashion the ontological principle that guides the analysis of this topic in, say, Aristotelian Christianity: The sacred does not destroy but presupposes and perfects the normal. Unlike many traditions the ordinary is not, that is, eradicated by the religious and replaced by something fundamentally different. (This feature probably most clearly distinguishes discontinuous Chinese religious traditions from most other discontinuous traditions.) Rather, the ordinary provides the basis that is developed into a more actualized form. In fact, one can even argue that Chinese thinkers are able to develop this principle more fully than could Aristotelian Christians because they lack those theological ideas that impede a full development, notably the notion of a natural order created by a God distinguished from it by aseity.

These ideas lead Chinese philosophers to understand (perhaps even more clearly than do their counterparts in other traditions) that treating religious virtue well involves a kind of balancing of opposing demands. On the one hand, religious virtues are virtues where one cannot draw on too many normal presumptions and arguments to defend, or even to make plausible, the virtue else it ceases to be a religious virtue. On the other hand, however, one cannot simply disregard normal presumptions and arguments else the virtue ceases to be a plausible option for most people. This activity involves balancing on a tightrope, a posture that recalls Ludwig Wittgenstein's comment that an "honest religious thinker is like a tightrope walker … [who] almost looks as though he were walking on nothing but air … [because his] support is the slenderest imaginable … [and] yet it really is possible to walk on it" (1984, p. 73e).

The balancing act involves not falling into either of two dangerous alternatives. On the one hand, the religious virtue must not rely on notions that no reasonable person can really entertain seriously. The claim that only through sacrificial rituals is one able to appease a spirit's anger or a dead person's perturbation is an example of such a notion. The virtues cannot rest, then, on what seems to sensible people to be implausible ideas. On the other hand, if the virtue is truly a religious one, it must not rely on such common and sensible notions that most people would, with little thought, accept it. The idea that one should help others if the help causes neither pain nor dislocation would be an example. The virtue cannot, then, simply repeat the conventional wisdom of the day. The need for this kind of distinctive balancing when presenting religious materials leads one to the genres such an approach demands.

GENRE AND THE PERSUASIVE PRESENTATION OF RELIGIOUS REALITIES

The delicate kind of balancing we see in the presentation of religious virtues leads directly to the subject of the ways in which religious features affect the genres, the modes of presentation, manifest in much Chinese philosophy. (These choices about genre are, moreover, especially significant because a number of these philosophers

were capable of, and well trained in, more rigorous forms of theoretical analysis.) Indeed, the rationale for presenting religious features in different genres provides an excellent way to examine the widespread perception that Chinese philosophy often does not seem to operate as philosophy ought to operate (e.g., see Mencius 1970, Book 6A; Chuang-Tzu 2001, pp. 48–61, 106–107; Hui-neng 1967, #1–#10; Wang Yang-Ming 1963, # 139, #168–#171, #226–231).

It is said, that is, that Chinese thought (as noted earlier) lacks sustained formal argumentation, sophisticated forms of analytic inquiry, and evident ways in which to reflect critically on presuppositions. These criticisms can witness to the kind of disabling parochialism (and circularity) that allows for little discussion, but their more powerful forms focus well the subject of how presentation and persuasion operate in Chinese religious philosophy.

The best way to approach this subject is to look at responses to a simple, deceptively simple, question: How can one persuasively represent a world, a world the understanding of which is crucial to any true human fulfillment, that far exceeds one's normal understanding? Representing that world persuasively is critical because only through such representation can one keep before people realities central to any religious vision but discontinuous with ordinary understandings. Representing that world is exceedingly difficult because it involves presenting realities that differ from, and even challenge, people's ordinary perspectives. The needed kinds of language must therefore persuade people in ways that differ from the kinds of persuasion either logical argument or even ordinary language utilize.

One illuminating instance of such a mode, and one much favored in China, is the use of concise, compelling, and often elusive locutions, such as the two from Confucius noted earlier about virtue always bringing neighbors and about the Way and death. These expressions provide one with a great deal of textured material in a terse, striking form. Indeed, they both arrest and often stay with one because they give one something intriguing and rewarding to which one can return. These statements can, then, embed themselves in one's mind and lead one to mull them over, searching out their various implications and applications to one's own life. Such statements become meditational objects that work on one, as do all meditational objects, in both evident and mysterious ways. A specific literary device like this aims, then, to produce in the reader fascination, sympathetic identification, attentive perplexity, and other even more complex emotional states, such as pretending. All are states that can produce significant personal changes.

Put another way, two features of religious perspectives make necessary forms like these: First, simple rational arguments about such perspectives will only rarely affect those people who most need help, a group that includes most of everyone at different points in their lives. Second, those arguments, or even the appropriate principles they produce, will often not fundamentally affect most people in those situations where they most need help.

Especially important, therefore, is persuading people that ideas, actions, and perspectives they find odd, perplexing, or simply wrong are worth considering, even worth adopting. And that task's difficulty is heightened by the relative absence, in comparison to theistic traditions, either of limpid theological propositions about the sacred's character or of graphic accounts in authoritative texts of the actions of the highest sacred beings.

The problem, then, is how to employ language that is odd, often very odd, and yet still have it be persuasive. (It resembles, therefore, the balancing needed in presenting ordinary and religious virtues, but the scope of operations is much wider.) To attain the needed representation one must stretch language beyond its evident limits and recognizable shapes while one also understands that such stretching seems to violate those forms and expectations that allow language to convey meaning. How can one, that is, represent a world that transcends one's ordinary categories and even reference points in a way that is both realistic and persuasive? The Chinese responses to these interrelated questions, evident more in their practices than in their theories, rest on three ideas, or more accurately claims, that are implicit in their practice.

The first and most obvious involves distinguishing the persuasively presented from the well argued and emphasizing the former. Presentations may, then, fail to fit well the criteria for good argument, indeed may even be instances of reasoned attempts to doubt the value, when the subject is religious, of many kinds of reasoned arguments. (*Persuasively presented* is, of course, a considerably wider category than is *well argued*, and utilizing it necessarily involves one, as it will be seen, in the treatment of issues about the character of rhetorical presentation and those subjects that follow in its wake.)

Second is a general characteristic of much Chinese religious thought: the judgment that considerations about the deepest religious matters best manifest their distinctive subject matter when understood as treating

irresolvable but illuminating and productive tensions. These tensions arise from the presence of apparently conflicting ideas and experiences each of which is irreducible; any resolution, therefore, that even diminishes the tensions must be rejected. Indeed, a resolution need not even be sought because keeping the tension's irresolvability in mind both enables people to understand better the character of religious reflection and presentation and clarifies their relationship to religious realities.

The third notion or claim is a direct corollary of the preceding two: the idea that literalism is the most dangerous of all human deformations at least when religion is the subject, and probably even when life itself is the subject. Literalism can take different forms but at one end of the spectrum is an unwillingness (or perhaps inability) to read beyond a surface meaning—literalism in its most evident sense. At the other end, however, is an unwillingness (or perhaps inability) to do anything but read beyond the most evident sense; the wooden pursuit of allegorical readings displays another, more abstract kind of literalism. The first fails to grasp the import of the representation; the second's easy movement beyond the surface overlooks all the surface's rich texture.

Chinese religious thinkers, then, usually focus less on straightforward conceptual analysis or argument and more on persuasive presentations that work with irresolvable but revelatory and productive tensions, aim to change people's understanding and action, and nurture the avoidance of literalism. That focus helps to explain their use of genres that are, to employ Western categories, more often literary than theoretical.

The use of these genres, genres that aim to present realities that can be made evident or compelling in no other way, means rhetoric is crucial. And that means that one does not simply face passages that are the shadowgraphs of ideas, passages that can be put into propositional forms that leave no remainder. The language used is not the mere adornment of an idea; it is constitutive of the idea. The language used is not just a device to persuade the recalcitrant or intellectually inept. Rather, it is what makes possible any appropriation of the proper perspective.

Processes of persuasion like these can be thought to be problematic for many reasons. For example, the process seems to disregard too many significant, if ordinary, kinds of thinking; the process will often fail to provide warrants for adjudicating differences; and the process is not attentive enough to the need for the theoretical analysis and justification of at least many rhetorical statements.

Chinese thinkers are aware of these problems. They understand they must evaluate rhetoric and that such an evaluation involves both a detailed understanding of how rhetoric works and a grasp of the character and appropriate roles of logical argument. (In fact, they often display a remarkable grasp of different rhetorical modes and therefore also of the ways in which such modes may obfuscate.) Probably most important, however, is a recognition that rhetorical presentations are part of a more general process of self-cultivation that involves teachers, various spiritual disciplines, and participation in a tradition. This remains true despite the difficulties they often see in the ordinary understandings of self-cultivation and of traditions that dominate most communities.

Nevertheless, it remains true that they often gravitate to distinctive genres when presenting religious perspectives because only those genres can produce what must be produced. They accept, that is, a version of the "good person criterion," a criterion also evident in the Aristotelian tradition (Yearley 1990, pp. 62–72). They believe that a person's character determines what can be perceived and understood, and therefore that the ultimate measure of a person or an action's excellence, or even meaning, is the excellence of the person who makes the judgment. Most important here, they recognize that this criterion has dramatic implications for both presentation and persuasion, and they are more than willing to live with the consequences.

See also Buddhism; Chinese Philosophy: Daoism; Chinese Philosophy: Ethics; Chinese Philosophy: Metaphysics and Epistemology; Confucius.

Bibliography

Ching, Julia. *The Religious Thought of Chu Hsi*. New York: Oxford University Press, 2000.

Chuang-Tzu. *Chuang-tzu: The Inner Chapters*. Translated by Angus Graham. Indianapolis, IN: Hackett, 2001.

Confucius. *Confucius Analects, with Selections from Traditional Commentaries*. Translated by Edward Slingerland. Indianapolis, IN: Hackett, 2003.

Geertz, Clifford. *The Interpretation of Cultures: Selected Essays*. New York: Basic, 1973.

Graham, Angus. *Two Chinese Philosophers: Ch'eng Ming-tao and Ch'eng Yi-Yuan*. London: Lund Humphries, 1958.

Gregory, Peter. *Tsung-mi and the Sinification of Buddhism*. Princeton, NJ: Princeton University Press, 1991.

Hui-neng. *The Platform Sutra of the Sixth Patriarch: The Text of the Tun-Huang Manuscript*. Translated by Philip Yampolsky. New York: Columbia University Press, 1967.

Ko Hung. *Alchemy, Medicine, and Religion in the China of A.D. 320: The Nei P'ien of Ko Hung (Pao-p'u tzu)*. Translated by James R. Ware. Cambridge, MA: MIT Press, 1966.

Mencius. *Mencius*. Translated by D. C. Lau. Baltimore, MD: Penguin, 1970.

Robinet, Isabelle. *Taoism: Growth of a Religion*. Translated by Phyllis Brooks. Stanford, CA: Stanford University Press, 1997. [op. 1992]

Wang Bi. *The Classic of Changes: A New Translation of the I Ching as Interpreted by Wang Bi*. Translated by Richard Lynn. New York: Columbia University Press, 1994.

Wang Yang-Ming. *Instructions for Practical Living and Other Neo-Confucian Writings by Wang Yang-Ming*. Translated by Wing-tsit Chan. New York: Columbia University Press, 1963.

Wittgenstein, Ludwig. *Culture and Value*, edited by G. H. Von Wright. Chicago: University of Chicago Press, 1984.

Xunzi. *Xunzi: A Translation and Study of the Complete Works*. Vol. 3. Translated by John Knoblock. Stanford, CA: Stanford University Press, 1994.

Yearley, Lee H. *Mencius and Aquinas: Theories of Virtue and Conceptions of Courage*. Albany: SUNY Press, 1990.

Lee H. Yearley (2005)

SOCIAL AND POLITICAL THOUGHT

Chinese philosophy began in the sixth century BCE with social and political philosophy as a response to the collapse of traditional bronze-age feudal society (Shang and Zhou dynasties). As the loyalty of the nobility to the Zhou kings began to give way to the realpolitik of sheer military might, dozens of small kingdoms vied with one another for imperial domination in what became known appropriately as the Warring States period (475–221 BCE). This strife ended when the kingdom of Qin finally conquered the last of its competitors to unite, for the first time, the many warring states into a single military, imperial empire.

The problem was how to unify and rule such a heterogeneous collection of different ethnic groups. Into this breach came China's first philosophers. In the feudal period, social custom was maintained by etiquette as practiced and maintained by the aristocrats and by punishments applied to ordinary people. People did the morally right thing not out of an inner sense of obligation but simply because this was the prescribed behavior for anyone born into a particular class. As the feudal order broke up, Confucians attempted to convert the customary etiquette of the hereditary nobility into an internally felt and inwardly directed moral imperative for everyone. The Daoists advocated a back-to-nature simplification in which government does little and simply lets the people pursue their own affairs, as they traditionally did for centuries. The Mohists tried to break down ethnic and tribal boundaries through the practice of impartial universal love. And the legalists tried to extend the role of punishment more broadly to everyone: aristocrats and educated elite, as well as peasants. There were thus four recommended replacements for the dying feudal social order: develop a universal personal morality (Confucians); return to a nearly anarchic state of nature (Daoists); embrace a policy of universal impartial love (Mohists); establish and universally enforce applicable law (legalists).

CONFUCIUS

Confucius (551–479 BCE) was the first thinker to offer new methods for the postfeudal period. Philosophers in the early Warring States period tended to be either conservatives who wanted to preserve the old values of the dying feudal system or revolutionaries who wanted to start afresh with new ideas and a new set of values. Confucius was one of the conservatives. He sought to revive on a new foundation the values of the Zhou dynasty, the last of the feudal regimes. Confucius claimed not to be an original thinker; he always said he was just preserving the past. But since the feudal order had virtually disappeared by the time Confucius was born, he realized that these values could only be preserved by restructuring them so as to meet the new conditions. In this effort he was certainly original. In the past, aristocratic feudal values were informally handed down from parents to children in elite noble families. Confucius was the first Chinese thinker to advocate that these values be systematized, logically defended, universalized, and formally taught to everyone. If everyone would learn and practice the ancient virtues of loyalty to elders and rulers (*cheng*), moral righteousness (*yi*), and compassion for others (*ren*), the country would be well run, contented, and prosperous. Indeed, Confucius insisted that if people would simply fulfill the roles assigned to them, all would behave virtuously. For example, if the ruler would act like a ruler (that is, protect and care for his people), the country would be well taken care of. The ruler who takes advantage of his subjects to enrich himself is not a true ruler and should not be called a ruler (*wang*). This is what Confucius called the "rectification of names": Things should be called by their right names, and people should live up to the roles assigned them by their designated titles.

Beyond assuming a universal basis for such an extension of the old feudal values, Confucius did not develop a theory of universal human nature. This he left to his followers. When the Daoists and other competing schools of philosophy criticized Confucius for foolishly trying to revive the old values of the nobility of a then defunct feu-

dal era, Confucius's followers responded by arguing that all people have the same basic nature, that to be happy and successful, people must fulfill this nature, and that this required developing and following the traditional Chinese virtues of the ancient feudal nobility.

Soon after Confucius's death there arose many competing schools of philosophy in China and many competing varieties of Confucianism itself, reflecting serious disagreements among the followers of Confucius, especially Mencius and Xunzi. When the small warring states were finally united for the first time under the military dictatorship of Ying Zheng (259–210 BCE), China's first imperial emperor, the legalist philosophers (*fajia*) convinced the emperor that the only way to truly unite the country was to eliminate all the argumentative and therefore divisive schools of philosophy (except legalism). In 213 BCE all the philosophy books the government could find were burned, and some Confucians and other philosophers who refused to abandon their philosophical practice were buried alive. Shortly afterwards, during the Han dynasty (206 BCE–220 CE), the legalist position was reversed, an attempt was made to revive the ancient schools of philosophy, and after a prolonged debate, Confucianism was adopted over Daoism and Mohism as the official philosophy of the state. The only other philosophical school besides Confucianism to survive was Daoism, which then became a private philosophy of the educated elite, whose public lives were primarily Confucian. Confucianism thereby became China's official legitimizing discourse justifying imperial rule, although quietly and behind the scenes legalism continued to provide the practical basis for actual rule.

Confucius did not develop the systematic ethical theory we know today as Confucianism. This was left to his followers: Mencius, Xunzi, and Dong Zhongshu. Confucius himself sought to return to traditional feudal values (such as filial piety) of the Zhou dynasty. Confucius's critics, especially the Daoists, attacked this attempt to rehabilitate the traditional virtues. Just because such virtues may have worked in the past, what reason was there to suppose that these rules were applicable to the new circumstances? A theory of ethics should be based on something more permanent in nature and reality, not the historical conventions of a particular society at a particular time and place. In response to this criticism, the followers of Confucius developed the idea that the traditional Chinese virtues could be defended as based on an unchanging and eternal human nature shared by people of all classes, at all times, and at all places.

But *is* there a common human nature, and if so, what is that nature? Here the followers of Confucius could not reach agreement. Mencius held that human nature is essentially good, while Xunzi argued that it is essentially evil, to which Dong Zhongshu argued that it is both, that human nature comprised two opposing elements, one good and the other evil.

MENCIUS

In many ways Mencius (c. 376–c. 292 BCE) followed the example of Confucius. Like Confucius, he divided his time between offering (mostly unwanted) advice to the rulers of his day and teaching students in a private capacity. Also like Confucius, Mencius sought to rectify names. Things should be called by their correct names, and things (people) that did not live up to their names (titles) should not be called by those names. For example, when Mencius was asked whether it was morally permissible to kill a king, Mencius replied in effect that if this so-called king is a true king, then of course it is wrong to kill him, but if this so-called king is not a true king but only a tyrant, then killing him is not killing a king but only a tyrant, and so is morally permissible.

But in other ways Mencius deviated from the path followed by Confucius and most other Confucians before and after him. More than any other Confucian, including Confucius himself, Mencius emphasized the morality of following nature, human nature, as opposed to social convention. In this respect Mencius comes close to the Daoists. In one episode recorded in his book, Mencius tells the story of the man from Song who helped his rice plants grow taller by pulling on them, which of course caused them to wither and die. Mencius's point is that the farmer should have let nature take its course. Of course, the rice has to be planted and then transplanted and weeded and watered and protected from birds and other animals, but beyond that the rice plants should be left to grow and develop all by themselves without further human interference. By analogy, we should not impose alien social practices on children but should help them cultivate their innate human nature.

By the time of Mencius there were many competing schools of philosophy. Mencius often bemoaned the enormous popularity of what he regarded as two extremist schools of thought, the followers of Mozi (c. 468–c. 376 BCE) and Yang Zhu (440–360 BCE), Mozi arguing that we should love all people equally and Yang Zhu that we should not lift a finger to help anyone but ourselves. Like Confucius, Mencius followed a middle path between these two extremes, emphasizing the traditional Chinese

virtue of filial piety, that is, loyalty to members of one's own family and obligation to those immediately above and below in the social hierarchy of one's particular community.

According to Mencius, all people are born with the potential and tendency to be kind-hearted and virtuous, though this potential can either be nourished and developed, so that the individual becomes good, or else neglected, thwarted, and perverted, so that the individual becomes bad. Mencius was not saying that children are moral from birth. He realized that they must be trained and taught, and that they learn by practice and experience. He also realized that neglected or mistreated children will usually become bad. Nonetheless, his theory is that in either case there is an innate tendency or disposition to be good.

In his most famous example, Mencius asked what is the immediate and spontaneous response of any person upon seeing a child about to fall into a well. Mencius said that everyone naturally and spontaneously wants to rush to help the child. This does not mean that everyone is a morally good person. It only means that everyone is born with the germ of the Confucian virtue of compassion (*ren*), along with the germ of the other traditional Chinese virtues of righteousness (*yi*), propriety (*li*), and wisdom (*zhi*).

Mencius valued this distinctively human capacity for virtue above all other parts of a person and urged readers to honor, preserve, and develop that part of themselves above all else. "To know one's nature is to know heaven [*tian*]" (*The Mencius*). That is, humans alone have the capacity to realize what their nature is and to choose to follow it and in this way to consciously align themselves with heaven. Mencius saw this human capacity as a bodily part (the benevolent mind that cannot bear to see others suffer), though the highest part, coming from heaven, in contrast with the rest of the human body, which people share with the lower animals and which comes from earth.

Although Mencius did not stress rationality as what is distinctively human, as Western philosophers do, he did stress the capacity of the mind (*xin*) to think. In *Mencius* (the book), Mencius argued that some people become better than others (even though they all have the same human nature) because they realize the value of this small but superior part of themselves.

Finally, Mencius is often called the most democratic of Chinese philosophers. Although Mencius, like Confucius before him, defended the ancient feudal traditions, especially of the Zhou dynasty, he radically reinterpreted them to conform to his own political ideas of equality. For example, in feudal times, society was arranged for the benefit of the aristocracy and defended as being mandated by Heaven (*tian ming*). Of course, it is hard to tell who has the mandate of heaven, except in a circular way by who actually rules. The ruler in power can claim to enjoy the mandate of heaven, and the only argument against this claim by those who oppose him is their ability to oust him from power and take over themselves. Mencius offers an independent, noncircular criterion for who has and does not have the mandate of heaven: that the ordinary people support and are happy with the government. The only justification for government and economic policy is that it serve the people. To give another example, in feudal times the division of political and economic duties was hereditary, whereas for Mencius this division of labor can only be justified on grounds of merit. Let each person serve according to his innate ability, whether as farmers, teachers, or government officials.

XUNZI

In the third century BCE the most prominent Confucian and one of the most important philosophers in China was Xunzi (340–245 BCE). Xunzi argued, against Mencius, that human nature is essentially evil, that is, selfish and aggressively antisocial. It is only through education, training, discipline, and the threat of punishment, Xunzi argued, that people become socially cooperative. Xunzi speculated that originally men were free to follow their own selfish bent without fear of recrimination or punishment. But when they realized that they were as often the victims of aggressive abuse as its perpetrators, that they were getting robbed as often as they were robbing from others, they willingly accepted the authority of a ruler capable of maintaining order and punishing transgressions. Like the social-contract philosopher Thomas Hobbes, Xunzi argued from people's evil nature to the need for a strong central government to control human behavior by education and a system of rewards and punishments. If strong governmental authority were removed, Xunzi speculated, chaos would result as the strong rode rough shod over the weak, with no law enforcement to prevent or punish them.

One major difference between Mencius's and Xunzi's theories of human nature is that Mencius defined human nature as what is uniquely and distinctively human, whereas Xunzi defined human nature as what all people are born with, even if this allotment is also shared with lower animals. Relating his theory to ancient, prephilo-

sophical Chinese traditions, Xunzi said that human nature is the product of two factors. One, the contribution of heaven (*tian*), gives human beings the rational and intelligent capacity to be civilized, cultured, and virtuous, and the other, the contribution of earth (*di*), is our animal nature, which we are conscious of as feeling and emotion. Whereas Mencius said that we receive from heaven the germ of moral virtue, Xunzi maintained that we receive from heaven only the capacity or potential for virtue and civilized life.

Thus, for Xunzi, a person at birth is just like one of the lower animals except for possessing the capacity of becoming civilized and virtuous. If we define human nature as Xunzi did (as what all people are born with), then we will point to the tendencies people actually have to be greedy, selfish, and aggressive, but if we define human nature as Mencius did (as what is unique to people), then we will tend to discount greedy behavior, since it is shared with lower animals, and to emphasize instead the capacity of humans to develop virtuous behavior, to become moral creatures. As Mencius says, "Slight is the difference between man and the brutes. The ordinary person loses this distinguishing feature, while the true human being retains it" (*The Mencius*).

In a sense, the difference between Mencius and Xunzi is very small. Both acknowledged that we have natural desires for food and sex, and both acknowledged that we have the capacity to resist such desires when it is dangerous or inappropriate to indulge them. The difference is largely a matter of the relative weight placed on nature and nurture. Xunzi thought that because human beings are intelligent, when they realize the difficulties that uncontrolled indulgence in the desire for food and sex can lead to, they seek to set limits on those desires. Like Mencius, Xunzi acknowledged that the ordinary person can become a sage.

According to Xunzi, human goodness comes from development of human culture. Culture is uniquely human. "Heaven has its seasons, Earth has its resources, man has his culture" (*The Xunzi*). Humans should properly take what comes from heaven and earth and create a distinctly human culture. Just as Mencius held there were the four germs of human goodness, so Xunzi held there are the four germs of evil, all of which spring from the innate desire for profit and sensual pleasure. How, then, do humans become good? And what motivates them to become good if they are inherently evil? Xunzi developed two lines of argument.

First, humans need (and know they need) some kind of social organization, cooperation, and mutual support. To secure the required social organization, they need rules of conduct, ceremonial rites (*li*). (Ceremonial rites were of greater importance to Xunzi than to Confucius, who stressed compassion for others, *ren*.) We need rules of conduct to set limits on the satisfaction of desires.

Second, we need morality (*li*), culture, civilization to complete our humanity. The rules of conduct cultivate and refine our humanity. Unlike the Daoists, who rejected what comes from humans to return to nature, Xunzi advocated the way of humanity.

Xunzi further developed Confucius's sophisticated view of the utility of elaborate ceremonies, without the need for belief in conscious humanlike deities. For Xunzi, this involved a kind of aesthetic distance. We have both intellect and emotion. We intellectually know that death is the end of everything (and that gods cannot help improve the weather through prayers), but we emotionally need to hold on to some hope of something better to follow death (and the possibility of some help from a benevolent heaven). So we create in our rituals a kind of poetic imagination in which we believe and disbelieve all at the same time. Ordinary people can believe literally, while educated people can appreciate the same ceremonies aesthetically and symbolically. For civility, human emotions must have a physical embodiment, which distances the emotion from its natural expression. Thus, art and music become a way of inculcating proper social attitudes in the educational process and avoiding natural, animalistic expressions of such attitudes as aggression, for example.

Xunzi regarded dispute and argumentation as a sign of political disorder, and so encouraged the idea that a return of political order (at the end of the feudal period) would lead to the end of philosophical disputes and argumentation among the many different contending schools of philosophy. Unfortunately, through Xunzi's influence on the legalists Li Si and Han Fei, this contributed to the famous book burning of 213 BCE.

In some ways Xunzi resembled the Daoists, especially in his rationalist, scientific attitude toward nature or heaven (*tian*). In other ways, however, Xunzi followed Confucius in arguing that we ought nonetheless to keep up all the state ceremonies, even sacrifices to ancestors and gods, though they have no real causal effect. Why? Because these practices were socially beneficial; the emperor publicly praying for a good harvest did not make the crops grow any better, but it did help unify the people and organize their efforts toward a common goal.

DONG ZHONGSHU

Is there a human nature shared by all people, and if so what is it like? As discussed above, the followers of Confucius could not agree. Mencius held that it was essentially good, while Xunzi argued that it was essentially evil, to which Dong Zhongshu (179–104 BCE) argued that it was both, that human nature was composed of two opposing elements, one good and the other evil. Dong, in other words, found a middle ground between the views of Mencius and Xunzi (though probably closer to Xunzi). He agreed with Mencius that in a sense human nature contains the germ of goodness, but he disagreed with Mencius that this is enough to say that humans are by nature good. The germ of goodness is not actually good any more than a tomato seed is a tomato or an egg is a chicken. To become good, that germ must be nurtured and cultivated. He thus agreed more with Xunzi's emphasis on the need for government to educate and train people to become good citizens. Whereas Mencius said that goodness is a natural "tendency" of people, Dong claimed it is a mere "potential."

Dong also developed the theory that human nature must compete with people's innate tendencies toward greed and selfishness. In Dong's human psychology, the opposing forces of yin (emotion and feeling) and yang (our distinctively human nature) are in constant conflict with one another. If both these tendencies are innate, one may ask, are they both not parts of human nature? The answer here can be related to the idea that human nature is a normative concept. Like Mencius, Dong would like to say that human nature is the higher, better part of humans, the morally good part (derived from the positive yang aspect of heaven), which humans alone are capable of. The instinctive, emotional part (derived from the negative yin aspect of heaven), which all humans possess but also share with lower animals, is just as innate, but lacks the normative quality of the morally good potential of human nature.

The main difference between Mencius and Dong lies in their views of the role of government in fostering moral goodness. Mencius would have government take a far less intrusive role, merely encouraging and cultivating the germ of moral goodness already there. In contrast, Dong, like Xunzi, thinking of the enormous challenge of the Qin and Han dynasties in unifying the many previously warring states, held that government must mold and shape humans, who have the capacity for goodness but cannot become good without the intervention of the state. Lurking in the background of this Confucian debate lay the worry that moral cultivation, however noble an ideal, would not politically unify the vast military empire without strong state coercion.

MOZI

Mozi, or Mo Di as he is also known (c. 468–c. 376 BCE), was China's second philosopher, after Confucius. In his own lifetime and for two hundred years following his death, Mohism was at least as influential as Confucianism or any other early Chinese school of philosophy. But by the time of the Han dynasty (206 BCE–220 CE), Confucianism and Daoism had absorbed all the other schools of philosophy, and from then on, Mozi exercised little influence.

Philosophy arose in China at the end of the feudal period, and many scholars believe that the Confucians emerged from the ritual advisors (the *ru*) to the early feudal lords, while the Mohists emerged from the feudal warrior class. Certainly, Mozi's philosophy is much more down to earth, practical, and less elitist than Confucianism. Mozi opposed Confucius on several grounds, but four stand out as most important: that right action is determined by its practical results and consequences and not, as Confucius had urged, because duty required it, regardless of the consequences; that one should not privilege members of one's own family, especially one's parents, siblings, and children, but should love everyone equally; that morality should be based not on an unchanging human nature, which may or may not exist, but on our ability to transform people into morally better individuals through education and law; and that we should honor and obey a personal God, who rules heaven and earth, rewarding the faithful and punishing all others.

Mozi argued that the cause of the world's ills was the fact that people loved each other partially, that you love your mother and your clansmen more than you love my mother and my clansmen, for example, and that the cure for the world's ills is therefore to embrace universal, impartial love, in which everyone loves everyone else equally. Where there is competition, partial love leads "us" to hate and want to destroy "them." And so we have discrimination, ethnic cleansing, genocide, and warfare. How does one overcome these tendencies? According to Mozi, "Partial love should be replaced by universal, impartial love" (*The Mozi*.

Mozi realized (with the help of his Confucian critics) that impartial love is contrary to our ordinary feelings; you will tend to favor your relatives over mine. The Confucians were naturally appalled at Mozi's rejection of the traditional Chinese virtue of filial piety, that one's pri-

mary responsibility in life is to one's own parents and children. The Confucians therefore vigorously argued against Mozi's views on impartial love, arguing that since this is contrary to nature, no one could or would follow Mozi's advice (even if he were right). Nonetheless, Mozi argued that a system of rewards and punishments can induce and socially condition people to practice universal love (if not actually to *feel* love equally toward everyone).

Specifically, he argued that if the ruler urges people to love one another impartially, they would strive to do so; that since God created humans and loves them all impartially, God wants us to love each other impartially and rewards us when we do and punishes us when we do not; and that this too encourages people to embrace impartial love. Mozi did not think or argue that we are born with a sense of universal love of humanity in our hearts, only that we can be trained to adopt such an attitude. In this regard, Mozi argued that humans are infinitely pliable and can be molded into any form desired by the government (either to love partially or to love impartially).

Contrary to Confucius, Mozi argued that we should do the right thing to receive the rewards (*li*) we will receive by doing the right thing and to avoid the punishment we will suffer in this life and the next if we do the wrong thing. Sometimes Mozi argued that we should do what will produce the best results for everyone, not just for ourselves, and here he sounds like the nineteenth-century British utilitarians (Jeremy Bentham, James Mill, and John Stuart Mill), who argued that we should always do what will produce the greatest happiness for the greatest number of people.

Like Confucius, Mozi took his political theories to government leaders, offering his advice on how to improve government performance and social conditions, and like Confucius, his advice was largely ignored. Mozi was utilitarian in the sense that his standard for judging a philosophical position was whether it will benefit the people. Like John Stuart Mill, Mozi produced a theory that is more social and political than moral. That is, he was less interested in describing why individuals should love their neighbors as much as themselves than in telling government leaders how ordinary people can be motivated to practice universal, impartial love and how this will benefit the country as a whole.

Mozi explicitly criticized Confucius and the Confucians for preaching atheism (since this makes the gods angry, and the gods will then take it out on the people, making their lives miserable). He also criticized the Confucians for extravagance in spending on lavish state ceremonies (including musical ceremonies) and funerals (including three years of mourning), for proposing a complex educational system (it is simply too much to master all the old Zhou-dynasty ritual and history classics), and finally for relying too much on fate (*ming*). Like Xunzi, Mozi argued that without government there would be chaos and hardship, with constant disagreements over what should be done, and that the people thus decided that it is better to have an absolute dictator to decide disputes for all.

DAOISM

Like Confucians and Mohists, Daoists also tried to influence government, and very nearly succeeded in convincing the Han emperor Wu Di (r. 141–87 BCE) to choose Daoism over Confucianism as the official philosophy of the state. Only the extraordinary influence of Dong Zhongshu led the emperor to give the nod in the end to Confucianism.

The Daoists favored the natural over the artificial and mercilessly criticized the Confucians for their emphasis on the humanly created civilized culture of art and literature, ritual and custom, which children must learn through an elaborate process of socialization and acculturation. The Daoists were especially critical of the Confucians' attempts actively to foster and promote morality. Sometimes the Daoists expressed themselves by saying that one should practice "nonaction" (*wu wei*), which, the context makes clear, does not mean doing nothing, which is impossible, but rather not acting too deliberately, purposefully, or self-consciously, that is, not trying so hard, just letting events take their natural course.

Trying too hard to do anything, the Daoists thought, only proves how lacking one is in that regard. Also, generally speaking, the harder we try, the less we succeed, the Daoists argued. Morality, like humor and lightheartedness, cannot be learned by rote, by mechanically following some set of rules, the Daoists insisted, but must spring from the heart spontaneously. Since morality is generally pitted against natural impulses, the Daoists were firmly opposed to morality as it is generally understood, that is, as a set of socially approved guidelines or rules to which all are expected to conform.

The Daoists also found themselves at the opposite extreme from the Confucian moral theory of government. The Daoists advocated just letting events take their natural course, leaving well enough alone. According to the Daoists, events happen naturally, spontaneously (*ziran*), of their own accord. The principle that directs the

growth and development of creatures and other things in the world is not some cause from outside, but a guiding force stemming from within those creatures. This is the natural and therefore preferred order of things. The worst thing one can do, especially rulers, is to try to improve on this natural order by enacting and enforcing laws.

It seems perfectly natural for rulers to feel that affairs are not going as well as they might and therefore to try to figure out what would make them better and to enact laws to bring about those changes. But for the Daoists, it is better for governments to let the people alone. Ordinary people have been managing their affairs from time immemorial, not by following formally enacted laws, but simply by following time-honored traditions and customs, which generally work just fine. By trying to improve the situation, the ruler may upset these established customs, confuse people, and make the situation worse.

Before governments found it necessary to introduce harsh laws to regulate behavior and punishments to enforce those laws, people lived simply without the need for laws. The ruler should keep government at this simple, primitive level. It is better not to give the ordinary people fancy ideas or encourage them to improve their lot. Keep them ignorant and simple. The Confucians were wrong to encourage knowledge and virtue. By insisting on learning and moral training, they made people feel ignorant and immoral, sense a need to study and learn what they did not know, and want to reform their ordinary ways of behaving.

Thus, even moral education is bad, according to the Daoists, because it tries to force on people overly sophisticated and difficult culture that goes against their nature. In direct opposition to the Confucians, the Daoists therefore rejected indoctrination in the traditional virtues: compassion, righteousness, propriety, and wisdom. If you have to teach morality, that is a sure sign that the situation has been allowed seriously to deteriorate. When affairs are running smoothly, the people naturally and spontaneously know what to do and how to behave—without thinking about it and without the need of books and formal instruction. Just like children, people are happier this way, not feeling inadequate and unhappy because they are constantly told how ignorant or sinful or uneducated or uncultivated they are.

It is also a mistake, the Daoists argued, to encourage the acquisition of expensive goods and higher standards of living. This just makes people envious of their richer neighbors and leads them to lie and steal and even kill to enrich themselves. The wise ruler will keep the people ignorant of fancy, expensive goods. If they never see such goods, they will never want them and never be tempted to stray from their simple everyday lives to get goods they cannot afford. Once the ruler allows inflated desires and competition among the people, the ruler must promulgate and enforce laws to prevent people from stealing and taking advantage of one another. But the more laws are passed and enforced, the more people see the laws and the government as their enemy, and therefore the more they will try to break the laws and overthrow the government. And this requires still more laws and severer punishments, in a vicious upward spiral.

LEGALISM

The legalist theory was best expressed by Han Fei, also known as Han Feizi (c. 280–233 BCE), at the end of the Warring States period. The legalists thought that it was not enough just to leave the people to their traditional customs, as the Daoists recommended, but that it was too much to transform everyone into a moral agent, as the Confucians and Mohists proposed. Neither sort of advice really takes into account what rulers themselves want. Rulers are generally not interested in being morally good, nor are they satisfied in just keeping the people quiet and docile. They usually have their own agendas: to gain fame by conquering neighboring kingdoms, to enrich themselves and their families, or perhaps both. Since most of the early Chinese philosophers were trying to persuade rulers of the time how best to govern, the legalists thought it better to advise rulers on how to achieve what rulers themselves wanted than to try to get them to accept the moralistic goals of the philosophers (who had no experience in ruling). The problem that Confucius faced in training kings to be philosophers was that the kings did not want to be philosophers—they wanted lives of action, wealth, and power. The legalists (ever political realists) accordingly dropped the more ambitious normative project of formulating the ends that governments should strive for and opted instead for a more instrumental approach to how to achieve the goals rulers already had.

To accomplish these political goals, the legalists advised the rulers to adopt a law-and-order administration supported by a strict system of rewards and punishments. Like their Western counterparts, the legalists were realists, arguing that it is not necessary for the king to be morally virtuous or for the bulk of the population to practice moral behavior. All the king needs to do, the legalists maintained, is to decide what he wants and then to insure compliance by formulating clear laws with absolutely certain rewards (for obeying these laws) and

punishments (for disobeying them), and the people will do whatever the king wants. After all, he is the king. He does not have to follow the moral principles of someone else—certainly not those of a philosopher! The king can propose whatever he wants and call this "justice" and make others call it "justice" as well, however inherently unjust his proposals may in fact be. And since he has the army to back him up, he cannot be seriously challenged.

Most Chinese philosophers were conservatives, revering and urging a return to a halcyon past, as Confucius thought the Zhou dynasty had been. Han Fei, on the other hand, as a legal and historical realist, argued that different historical eras face different problems requiring different solutions, and that the solutions of the past are not necessarily appropriate for the present. The story he offered is of a farmer who, seeing a hare kill itself by running into a tree stump, abandons farming to sit and wait by the tree stump for another hare. In the new expansionist military dictatorships following the end of the feudal period, a strict system of rewards and punishments for clearly formulated and promulgated laws is a much surer way of ensuring compliance than moral education, the legalists felt. Even if the ruler enacts a system of universal moral education, how many people are actually going to become moral agents, always doing the right thing simply because it is the morally right thing to do?

The ruler also needs statecraft (*shu*). He need not do the work himself; he need simply hold people to their job descriptions (the rectification of names). As a pragmatist, the ruler is concerned not with the methods needed to achieve results but only with the results. If the minister lives up to his job description, he is rewarded; otherwise, punished. After a while, incompetents do not apply.

In a sense, the legalist ruler follows the Daoist injunction of nonaction: "doing nothing, yet there is nothing that is not done" (*Daodejing*). And all this rests securely on the simple but basic foundation of human self-interest. Like his teacher Xunzi, Han Fei thought that human nature was evil, but unlike Xunzi, he sought not change human nature through education and training but only to establish a workable system of government built on this self-interested human nature. The legalists were strangely like the Daoists: Do not fight human nature; work with it. Even the Daoist Zhuangzi (c. 369–c. 286 BCE), seems to agree with the legalist principles of management: "The superior must have no activity, so as thus to have control of events; but the subordinates must have activity, so as thus to be controlled by events. This is the invariable way" (*The Zhuangzi*). The tax collector, for example, knows that he must collect taxes. If at the end of the year he has collected his allotted quota, the ruler rewards him; if not, the ruler punishes (and replaces) him. He may fail because of drought and famine, in which case, but, whether fair or unfair, he will lose his job—if not his head. In this way, the job gets done. The subordinates are controlled by events, and yet the ruler has done nothing except employ the right statecraft.

In another way, however, the legalists advocated the complete opposite of what the Daoists advocated. The Daoists held that human beings were completely innocent; the legalists that they were completely self-interested. The Daoists upheld individual freedom; the legalists, absolute social control. The Daoists regarded the legalists as shallow pragmatists—they knew that certain methods worked, but they had no idea *why* they worked.

CONCLUSION

In a way, the Daoists agreed with Confucius and Mencius that the ruler needs to have fundamental knowledge of human nature. Like the Confucians but unlike the Daoists, the legalists developed a social and moral philosophy in tune with the breakdown of feudal class distinctions. The Confucians and Mohists were revolutionary and idealistic—they sought to transform human nature (or, in the case of the Mohists, at least human behavior) by developing an inner sense of right and wrong through education. In contrast, the more realistic and pragmatic legalists developed methods for controlling people with their self-interested natures.

For two millennia China's political philosophy has been a combination of openly espousing legitimizing Confucian discourse while silently employing the more pragmatic legalist methods to achieve the ruler's objectives, along with Daoist principles of not interfering in the day-to-day affairs of the vast majority of the peasant population where their affairs did not conflict with the ruler's personal objectives.

See also Chinese Philosophy: Ethics.

Bibliography

Chan, Wing-tsit. *A Source Book in Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1963.

Creel, Herrlee G. *Confucius and the Chinese Way*. New York: Harper Torchbooks, 1960.

Fung, Yu-lan. *A History of Chinese Philosophy*. 2 vols. Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1983.

Fung, Yu-lan. *A Short History of Chinese Philosophy*. New York: Macmillan, 1966.

Graham, Angus C. *Disputers of the Dao*. La Salle, IL: Open Court, 1989.
Graham, Angus C. *Studies in Chinese Philosophy and Philosophical Literature*. Albany, NY: SUNY Press, 1990.
Hanson, Chad. *A Daoist Theory of Chinese Thought*. Oxford, U.K.: Oxford University Press, 1992.
Ivanhoe, Philip J., and Bryan W. Van Norden, eds. *Readings in Classical Chinese Philosophy*. New York: Seven Bridges Press, 2000.
Knoblock, John. *Xunzi*. 3 vols. Stanford, CA: Stanford University Press, 1988–1994.
Lao-tzu (Laozi). *Tao Te Ching: The Classic Book of Integrity and the Way*. Translated by Victor Mair. New York: Bantam Books, 1990.
Lau, D. C. *Confucius: The Analects*. New York: Penguin Books, 1979.
Lau, D. C. *Mencius*. New York: Penguin Books, 1970.
Lowe, Scott. *Mo Tzu's Religious Blueprint for a Chinese Utopia*. Lewiston, U.K.: Edwin Mellen Press, 1992.
Rosemont, Henry, ed. *Chinese Texts and Philosophical Contexts*. La Salle, IL: Open Court, 1991.
Schwartz, Benjamin. *The World of Thought in Ancient China*. Cambridge, MA: Harvard University Press, 1985.

H. Gene Blocker (2005)

CHINESE ROOM ARGUMENT

In 1980 the philosopher John R. Searle published in the journal *Behavioral and Brain Sciences* a simple thought experiment that he called the "Chinese Room Argument" against "Strong Artificial Intelligence (AI)." The thesis of Strong AI has since come to be called "computationalism," according to which cognition is just computation, hence mental states are just computational states:

COMPUTATIONALISM

According to computationalism, to explain how the mind works, cognitive science needs to find out what the right computations are—the ones that the brain performs to generate the mind and its capacities. Once we know that, then every system that performs those computations will have those mental states: Every computer that runs the mind's program will have a mind, because computation is hardware-independent: Any hardware that is running the right program has the right computational states.

THE TURING TEST

How do we know which program is the right program? Although it is not strictly a tenet of computationalism, an answer that many computationalists will agree to is that the right program will be the one that can pass the Turing Test (TT), which is to be a system that is able to interact by e-mail with real people exactly the way real people do—so exactly that no person can ever tell that the computer program is not another real person. Alan M. Turing (1950) had suggested that once a computer can do everything a real person can do so well that we cannot even tell them apart, it would be arbitrary to deny that that computer has a mind, that it is intelligent, that it can understand just as a real person can.

This, then, is the thesis that Searle set out to show was wrong: (1) mental states are just computational states, (2) the right computational states are the ones that can pass the TT, and (3) any and every hardware on which you run those computational states will have those mental states too.

HARDWARE-INDEPENDENCE

Searle's thought experiment was extremely simple. Normally, there is no way I can tell whether anyone or anything other than myself has mental states. The only mental states we can be sure about are our own. We cannot be someone else, to check whether they have mental states too. But computationalism has an important vulnerability in this regard: hardware-independence. Because any and every dynamical system (i.e., any physical hardware) that is executing the right computer program would have to have the right mental states, Searle himself can execute the computer program, thereby himself becoming the hardware, and then check whether he has the right mental states. In particular, Searle asks whether the computer that passes the TT really understands the e-mails it is receiving and sending.

THE CHINESE ROOM

To test this, Searle obviously cannot conduct the TT in English, for he already understands English. So in his thought-experiment the TT is conducted in Chinese: The (hypothetical) computer program he is testing in his thought-experiment is able to pass the TT in Chinese. That means it is able to receive and send e-mail in Chinese in such a way that none of its (real) Chinese pen-pals would ever suspect that they were not communicating with a real Chinese-speaking and Chinese-understanding person. (We are to imagine the e-mail exchanges going on as frequently we like, with as many people as we like, as long as we like, even for an entire lifetime. The TT is not just a short-term trick.)

SYMBOL MANIPULATION

In the original version of Searle's Chinese Room Argument he imagined himself in the Chinese Room, receiving the Chinese e-mails (a long string of Chinese symbols, completely unintelligible to Searle). He would then consult the TT-passing computer program, in the form of rules written (in English) on the wall of the room, explaining to Searle exactly how he should manipulate the symbols, based on the incoming e-mail, to generate the outgoing e-mail. It is important to understand that computation is just rule-based symbol manipulation and that the manipulation and matching is done purely on the basis of the shape of the symbols, not on the basis of their meaning.

Now the gist of Searle's argument is very simple: In doing all that, he would be doing exactly the same thing any other piece of hardware executing that TT-passing program was doing: rule-fully manipulating the input symbols on the basis of their shapes and generating output symbols that make sense to a Chinese pen-pal—the kind of e-mail reply a real pen-pal would send, a pen-pal that had understood the e-mail received, as well as the e-mail sent.

UNDERSTANDING

But Searle goes on to point out that in executing the program he himself would not be understanding the e-mails at all! He would just be manipulating meaningless symbols, on the basis of their shapes, according to the rules on the wall. Therefore, because of the hardware-independence of computation, if Searle would not be understanding Chinese under those conditions, neither would any other piece of hardware executing that Chinese TT-passing program. So much for computationalism and the theory that cognition is just computation.

THE SYSTEM REPLY

Searle correctly anticipated that his computationalist critics would not be happy with the handwriting on the wall: Their "System Reply" would be that Searle was only *part* of the TT-passing system. That whereas Searle would not be understanding Chinese under those conditions, the *system as a whole* would be!

Searle rightly replied that he found it hard to believe that he plus the walls together could constitute a mental state, but, playing the game, he added: Then forget about the walls and the room. Imagine that I have memorized all the symbol manipulation rules and can conduct them from memory. Then the whole system is me: Where's the understanding?

Desperate computationalists were still ready to argue that somewhere in there, inside Searle, under those conditions, there would lurk a Chinese understanding of which Searle himself was unaware, as in multiple personality disorder—but this seems even more far-fetched than the idea that a person plus walls has a joint mental state of which the person is unaware.

BRAIN POWER

So the Chinese Room Argument is right, such as it is, and computationalism is wrong. But if cognition is not just computation, what is it then? Here, Searle is not much help, for he first overstates what his argument has shown, concluding that it has shown (1) that cognition is not computation at all—whereas all it has shown is that cognition is not all computation. Searle also concludes that his argument has shown (2) that the Turing Test is invalid, whereas all it has shown is that the TT would be invalid if it could be passed by a purely computational system. His only positive recommendation is to turn brain-ward, trying to understand the causal powers of the brain instead of the computational powers of computers.

But it is not yet apparent what the relevant causal powers of the brain are, nor how to discover them. The TT itself is a potential guide: Surely the relevant causal power of the brain is its power to pass the TT! We know now (thanks to the Chinese Room Argument) that if a system could pass the TT via computation alone, that would not be enough. What would be missing?

THE ROBOT REPLY

One of the attempted refutations of the Chinese Room Argument—the "Robot Reply"—contained the seeds of an answer, but they were sown in the wrong soil. A robot's sensors and effectors were invoked to strengthen the System Reply: It is not Searle plus the walls of the Chinese Room that constitutes the Chinese-understanding system, it is Searle plus a robot's sensors and effectors. Searle rightly points out that it would still be him doing all the computations, and it was the computations that were on trial in the Chinese Room. But perhaps the TT itself needs to be looked at more closely here:

BEHAVIORAL CAPACITY

Turing's original Test was indeed the e-mail version of the TT. But there is nothing in Turing's paper or his arguments on behalf of the TT to suggest that it should be

restricted to candidates that are just computers, or even that it should be restricted to e-mail! The power of the TT is the argument that if the candidate can do everything a real person can do—and do it indistinguishably from the way real people do it, as judged by real people—then it would be mere prejudice to conclude that it lacked mental states when we were told it was a machine. We don't even really know what a machine is, or isn't!

But we do know that real people can do a lot more than just send e-mail to one another. They can see, touch, name, manipulate, and describe most of the things they talk about in their e-mail. Indeed, it is hard to imagine how either a real pen-pal or any designer of a TT-passing computer program could deal intelligibly with all the symbols in an e-mail message without also being able to do at least some of the things we can all do with the objects and events in the world that those symbols stand for.

SENSORIMOTOR GROUNDING OF SYMBOLS

Computation, as noted, is symbol manipulation, by rules based on the symbols' shapes, not their meanings. Computation, like language itself, is universal, and perhaps all-powerful (in that it can encode just about anything). But surely if we want the ability to understand the symbols' meanings to be among the mental states of the TT-passing system, this calls for more than just the symbols and the ability to manipulate them. Some, at least, of those symbols must be grounded in something other than just more meaningless symbols and symbol manipulations—otherwise the system is in the same situation as someone trying to look up the meaning of a word in a language—let us say, Chinese—that he does not understand … in a Chinese-Chinese dictionary! E-mailing the definitions of the words would be intelligible enough to a pen-pal who already understood Chinese, but they would be of no use to anyone or anything that did not understand Chinese. Some of the symbols must be grounded in the capacity to recognize and manipulate the things in the world that the symbols refer to.

MIND READING

So the TT candidate must be a robot, able to interact with the world that the symbols are about—including us—directly, not just via e-mail. And it must be able to do so indistinguishably from the way any of the rest of us interact with the world or with one another. That is the gist of the TT. The reason Turing originally formulated his test in its pen-pal form was so that we would not be biased by the candidate's appearance. But in today's cinematic sci-fi world we have, if anything, been primed to be overcredulous about robots, so much more capable are our familiar fictional on-screen cyborgs than any TT candidate yet designed in a cog-sci lab. In real life our subtle and biologically based "mind reading" skills (Frith and Frith 1999) will be all we need once cognitive science starts to catch up with science fiction and we can begin T-Testing in earnest.

THE OTHER-MINDS PROBLEM

Could the Chinese Room Argument be resurrected to debunk a TT-passing robot? Certainly not. For Searle's argument depended crucially on the hardware-independence of computation. That was what allowed Searle to "become" the candidate and then report back to us (truthfully) that we were mistaken if we thought he understood Chinese. But we cannot become the TT-passing robot, to check whether it really understands, any more than we can become another person. It is this parity (between other people and other robots) that is at the heart of the TT. And anyone who thinks this is not an exacting enough test of having a mind need only remind himself that the Blind Watchmaker (Darwinian evolution), our "natural designer," is no more capable of mind reading than any of the rest of us is. That leaves only the robot to know for sure whether or not it really understands.

See also Artificial Intelligence; Computationalism; Functionalism; Machine Intelligence.

Bibliography

Anderson, David, and B. Jack Copeland. "Artificial Life and the Chinese Room Argument." *Artificial Life* 8 (4) (2002): 371–378

Brown, Steven. "Peirce, Searle, and the Chinese Room Argument." *Cybernetics and Human Knowing* 9 (1) (2002): 23–38

Dyer, Michael G. "Intentionality and Computationalism: Minds, Machines, Searle, and Harnad." *Journal of Experimental and Theoretical Artificial Intelligence* 2 (4) (1990): 303–319.

French, Robert. "The Turing Test: The First Fifty Years." *Trends in Cognitive Sciences* 4 (3) (2000): 115–121.

Frith, Christopher D., and Uta Frith. "Interacting Minds—A Biological Basis." *Science* 286 (1999): 1692–1695.

Harnad, Stevan. "The Annotation Game: On Turing (1950) on Computing, Machinery and Intelligence." In *The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, edited by Robert Epstein and Grace Peters. Amsterdam: Kluwer Academic, 2004. http://cogprints.org/3322.

Harnad, Stevan. "Can a Machine Be Conscious? How?" *Journal of Consciousness Studies* 10 (4–5) (2003): 69–75, http://cogprints.org/2460/.
Harnad, Stevan. "Minds, Machines, and Searle." *Journal of Theoretical and Experimental Artificial Intelligence* 1 (1989): 5–25, http://cogprints.org/1573/.
Harnad, Stevan. "On Searle on the Failures of Computationalism." *Psycoloquy* 12 (61) (2001), http://psycprints.ecs.soton.ac.uk/archive/00000190/.
Harnad, Stevan. "The Symbol Grounding Problem." *Physica D* 42 (1990): 335–346, http://cogprints.org/3106/.
Harnad, Stevan. "Symbol-Grounding Problem." In *Encyclopedia of Cognitive Science*. London: Nature Publishing Group, 2003, http://cogprints.org/3018/.
Harnad, Stevan. "What's Wrong and Right about Searle's Chinese Room Argument?" In *Essays on Searle's Chinese Room Argument*, edited by M. Bishop and J. Preston. New York: Oxford University Press, 2001, http://cogprints.org/4023/.
Overill, Richard E. "Views into the Chinese Room: New Essays on Searle and Artificial Intelligence." *Journal of Logic and Computation*14 (2) (2004): 325–326.
Pylyshyn, Zenon W. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press, 1984.
Searle, John R. "Explanatory Inversion and Cognitive Science." *Behavioral and Brain Sciences* 13 (1990): 585–595.
Searle, John R. "The Failures of Computationalism: I." *Psycoloquy* 12 (62) (2001), http://psycprints.ecs.soton.ac.uk/archive/00000189/.
Searle, John. R. "Is the Brain's Mind a Computer Program?" *Scientific American* 262 (1) (January 1990): 26–31.
Searle, John. R. "Minds and Brains without Programs." In *Mindwaves: Thoughts on Intelligence, Identity, and Consciousness*, edited by Colin Blakemore and Susan Greenfield. Oxford, U.K.: Blackwell, 1987.
Searle, John. R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (3) (1980): 417–457, http://www.bbsonline.org/documents/a/00/00/04/84/index.html.
Searle, John R. *Minds, Brains, and Science*. Cambridge, MA: Harvard University Press, 1984.
Souder, Lawrence. "What Are We to Think about Thought Experiments?" *Argumentation* 17 (2) (2003): 203–217.
Turing, Alan M. "Computing Machinery and Intelligence." *Mind* 59 (236) (1950): 433–460, http://cogprints.org/499/.
Wakefield, Jerome C. "The Chinese Room Argument Reconsidered: Essentialism, Indeterminacy, and Strong AI." *Minds and Machines* 13 (2) (2003): 285–319.

Stevan Harnad (2005)

CHISHOLM, RODERICK
(1916–1999)

Roderick Chisholm was a twentieth-century American philosopher who made major contributions in almost every area of philosophy, but most notably in epistemology and metaphysics. Chisolm was an undergraduate at Brown University from 1934 to 1938 and a graduate student at Harvard from 1938 to 1942. He served in the military from 1942 to 1946, and then, after briefly holding a teaching post with the Barnes Foundation and the University of Pennsylvania, he returned in 1947 to Brown University, where he remained until his death.

EPISTEMOLOGY

In epistemology Chisholm was a defender of foundationalism. He asserted that any proposition that it is justified for a person to believe gets at least part of its justification from basic propositions, which are themselves justified but not by anything else. Contingent propositions are basic insofar as they correspond to self-presenting states of the person, which for Chisholm are states such that whenever one is in the state and believes that one is in it, one's belief is maximally justified. There are two types of self-presenting states: intentional states (ways of thinking, hoping, fearing, desiring, wondering, intending, etc.) and sensory states (ways of being appeared to by the various senses). A noncontingent proposition is basic if understanding it is sufficient for understanding that it is true and also sufficient for making it justified. "2 + 3 = 5" and "If Jones is ill and Smith is away, then Jones is ill" are examples of such propositions, says Chisholm.

Self-presentation and understanding are among the basic sources of epistemic justification, but according to Chisholm there are other sources as well. The most important of these other sources are perception, memory, belief coupled with a lack of negative coherence (e.g., no inconsistencies among the propositions believed), and belief coupled with positive coherence (i.e., mutual support among the proposition believed). For each of these sources, Chisholm forwards an epistemic principle that describes the conditions under which the source generates justification.

Despite his thinking that there are many sources of epistemic justification, Chisholm is rightly regarded as a foundationalist because all the sources are such that they can produce justified beliefs only because some propositions are justified basically. For example, Chisholm's principles concerning perception and memory make reference to propositions that are justified because they correspond to self-presenting states. In the case of perception, the relevant states are sensings, and for memory the relevant states are beliefs, in particular, beliefs to the effect that one remembers something. In a similar spirit, Chisholm says that coherence relations among propositions are not capable of generating justification for

propositions that have nothing else to recommend them; their role instead is to increase the degree of justification that propositions have by virtue of being supported by basic propositions.

Chisholm is also a proponent of internalism in epistemology, in two senses of the term. First, he thinks that epistemic justification supervenes on human conscious states; thus, whether one's beliefs are justified is determined by one's own internal states rather than by conditions obtaining in one's external environment. Second, he thinks that the conditions, if any, that justify one's beliefs are accessible to one; thus, one is always able to determine if one reflects carefully enough, whether one's beliefs are justified.

Chisholm's epistemology is resolutely antiskeptical. Indeed, he says that the proper way to begin doing epistemology is by presupposing that some human beliefs are justified and that indeed some constitute knowledge. Epistemology, so conceived, becomes primarily a search for the conditions that account for these beliefs being justified. A second task is to define the conditions that turn a true belief into knowledge. Chisholm's approach to this latter task is to defend a nondefeasibility account of knowledge. One knows a proposition *p*, he says, whenever one believes *p*, *p* is true, and *p* is nondefectively evident for that person, where *p* is nondefectively evident that person (some details aside) just in case there is a set of basic propositions that justify *p* and that justify nothing false.

METAPHYSICS

Chisholm also had well-worked-out views on almost every major issue in metaphysics, but his most influential views were concerned with thought and language, ontology, action, and material bodies.

With respect to thought and language, Chisholm was a defender of the primacy of thought; the intentionality of language is to be understood in terms of the intentionality of thought, he says, rather than conversely. He develops this idea in his direct attribution theory of reference. At the heart of the theory is a proposal that people are able to refer to things other than themselves by directly attributing properties to them and that people indirectly attribute properties to things by directly attributing properties to themselves. For example, if John is the only person in a room with Sally and John is wearing a blue sweater, then by directly attributing to herself the property of being a person *x* such that the only other person in the room with *x* is wearing a blue sweater, Sally indirectly attributes to John the property of wearing a blue sweater and thereby refers to John. Using these notions of direct and indirect attribution, Chisholm provides an account of various semantic notions including sense and reference.

In ontology, Chisholm's view is that there are only two kinds of entities: attributes and the individual things that have these attributes. Everything else, including propositions, states of affairs, possible worlds, and sets, can be understood in terms of these two categories. Attributes are possible objects of thought—more specifically, what people are able to attribute, either directly (to themselves) or indirectly (both to themselves and other things). Thus in ontology, Chisholm once again is a defender of the primacy of thought in that he uses the phenomenon of intentionality to identify and understand what kinds of entities there are.

His theory of action is an indeterministic one. The fundamental notions are those of undertaking and causing, and with respect to the latter notion he carefully distinguishes among necessary causal conditions for an event, sufficient causal conditions, and causal contributions. With these notions in hand, he opposes compatibilist attempts to understand what it is for a person to be free to undertake something, insisting that one has undertaken to do something freely only if there was no sufficient causal condition for one to undertake it (although there may have been extensive causal contributions to the undertaking).

Much of Chisholm's work on bodies is concerned with puzzles about the persistence of physical bodies through time, and most of these puzzles, in turn, are concerned with apparent violations of Leibniz's principle of the indiscernibility of identicals. According to this principle, if *X* and *Y* are identical, then whatever is true of *X* is also true of *Y*. One famous puzzle, for example, is the ship of Theseus. Even if one plank of the ship is replaced at a time *t*, it is the same ship—namely Theseus's—that exists before *t* and after *t*, and yet the ship might appear to have different properties before *t* and after *t*. Chisholm attempts to solve this and other puzzles about the identity of physical bodies through time by using his fundamental ontological categories, attributes, and individual things, to make precise the seventeenth-century distinction between substances and their modes.

See also Classical Foundationalism; Internalism and Externalism in Ethics; Persistence; Reference.

Bibliography

Bogdan, R., ed. *Roderick M. Chisholm*. Dordrecht, Netherlands: D. Reidel, 1986. Critical essays on Chisholm and a helpful self-profile.

Chisholm, R. *The First Person: An Essay on Reference and Intentionality*. Minneapolis: University of Minnesota Press, 1981. A detailed defense of the direct attribution theory of reference.

Chisholm, R. *Human Freedom and the Self*. Lawrence, KS: University of Kansas Press, 1964. A defense of his indeterministic account of human freedom.

Chisholm, R. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957. His first major work on epistemology.

Chisholm, R. *Person and Object: A Metaphysical Study*. London: Open Court Publishing Company. 1976. Most of his metaphysical positions are defended in this volume.

Chisholm, R. *Theory of Knowledge*. 3rd edition. Englewood Cliffs, NJ: Prentice-Hall, 1987. His most influential work in epistemology; the later editions contain important modifications of his earlier views.

Lehrer, K., ed. *Analysis and Metaphysics: Essays in Honor of R. M. Chisholm*. Dordrecht, Netherlands: D. Reidel, 1975. Critical essays on Chisholm's metaphysics and epistemology.

Richard Foley (1996, 2005)

CHOICE, AXIOM OF

See *Set Theory*

CHOMSKY, NOAM

(1928–)

Noam Chomsky is the foremost linguistic theorist of the post–World War II era, an important contributor to philosophical debates, and a notable radical activist. His influence is felt in many other fields, however, most notably, perhaps, in the area of cognitive studies.

Chomsky's main achievement was to distinguish linguistic competence from its manifestations in performance and to characterize competence as a system of explicit rules for the construction and interpretation of sentences. Indeed, this achievement provided a model for investigations, in this and other cognitive domains, that replaced then-dominant models based on the notion of analogy and oriented to the causal explanation of behavior.

The competence of individuals to use their language is constituted, on Chomsky's account, by their (tacit) knowledge of a formal grammar (or system of rules); their linguistic performance, involving the deployment of such knowledge, may be influenced by a host of extraneous factors that need not be accounted for by the grammar itself but, instead and if possible, by subsidiary theories (e.g., of perceptual processing, etc.). Furthermore, knowledge of such a system of rules permits a kind of creativity in performance that exhibits itself in the novelty, in relation to speakers' prior linguistic experiences, of (many of) the sentences they actually produce. (Crudely put, they can understand and produce sentences they have never before encountered.)

The competence/performance distinction reflects Chomsky's preference for "Galilean" theorizing (i.e., for a "modular" approach), and its introduction was tremendously liberating. A direct attack on performance, under broadly behavioristic auspices, had proved barren, for reasons Chomsky identified with devastating clarity in his review of B. F. Skinner's *Verbal Behavior.* Also pertinent was Chomsky's analysis of linguistic creativity in a second, distinct sense: the appropriateness and yet stimulus-independence (and therefore causal inexplicability) of much of what a speaker says in concrete circumstances. Shifting the linguist's problematic from behavior to the system underlying behavior was probably Chomsky's most important contribution to the development of "scientific" studies of social phenomena. (Of course, the competence/performance distinction owes much to Ferdinand de Saussure's earlier distinction between *langue* and *parole.* But Saussure did not think of the system underlying behavior as primarily rule-based, and so his distinction proved less fertile than Chomsky's.)

In a series of works beginning with *Cartesian Linguistics,* Chomsky took up what he came to call "Plato's problem"—that of explaining how the gap is bridged between individuals' limited opportunities, as children, for acquiring knowledge of their (native) language(s) and the competence to make many subtle and complex discriminations that, as mature speakers, they do indeed possess. He solved this problem, siding with classical rationalists such as Gottfried Wilhelm Leibniz, by assuming the existence, as an innate species-wide attribute, of a "universal grammar." During the course of language acquisition, limited data fixes the values of free "parameters" associated with this grammar, thus providing a basis for full-blown knowledge of the language that far exceeds the ordinary "inductive" implications of these data.

Chomsky has also been a notable advocate, very significantly in a discipline previously marked by instrumentalist assumptions about theorizing, of a realist perspective on theoretical entities and processes. In early work deep structures were postulated as sources, via

transformations, of familiar superficial structures of sentences. So, for instance, a superficially passive sentence was said to be derived from the same deep structure as its active counterpart. And while it might have been more in line with then-contemporary practice to treat these so-called deep structures as pure postulates, useful in simplifying the description and taxonomization of the superficial sentences of our "experience," Chomsky advocated, instead, that they be treated as having psychological reality and thereby fostered many profound psycholinguistic studies intended to bear out or refute this contention. A topic of continuing importance is whether it is only structures or, instead, derivational processes as well that are to be treated as "real."

Less noticed by commentators is Chomsky's profoundly individualistic approach to linguistic phenomena. For him, language itself is a secondary phenomena; primacy is accorded to an individual's competence, a purely psychological phenomenon. Indeed, Chomsky explains the coordination of linguistic interaction, not by reference to any transpersonal system of conventions (as might be thought appropriate in relation to other social phenomena), but, instead, to a harmony—between the competence of the speaker and the marginally different competence of the hearer—that depends largely on the innate constraints on their (typically) quite separate episodes of language acquisition. Even if each learns in isolation from the other, and has quite (though not "too") different experiential bases for learning, each will acquire an "idiolect" that is accessible to the other: Otherwise rather different data-sets fix the free parameters of the universal grammar in sufficiently similar ways to permit mutual intelligibility.

Other philosophically important themes in Chomsky's work include: (1) his identification of the ideological interests that are served by certain allegedly "scientific" approaches to the study of human behavior; (2) his argument for treating the capacity for language as species-specific and thus as an aspect of the human "essence"; (3) his speculations about the possibility that there are innate limitations on the human capacity for knowledge of the world; and (4) his continued defense, in the face of broadly "postmodernist" opposition, of the role of reason in understanding and improving the human condition and of the viability of the notion of "progress" in relation to these projects.

See also Behaviorism; Cognitive Science; Leibniz, Gottfried Wilhelm; Modernism and Postmodernism; Postmodernism.

Bibliography

WORKS BY CHOMSKY

Syntactic Structures. The Hague: Mouton, 1957.

"A Review of B. F. Skinner's *Verbal Behavior.*" In *The Structure of Language,* edited by J. A. Fodor and J. J. Katz. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Cartesian Linguistics. New York: Harper, 1966.

Language and Mind. New York: Harcourt Brace, 1972.

For Reasons of State. London: Fontana, 1973.

Language and Problems of Knowledge. Cambridge, MA: MIT Press, 1988.

WORKS ON CHOMSKY

D'Agostino, F. *Chomsky's System of Ideas.* Oxford: Clarendon Press, 1986.

Harman, G., ed. *On Noam Chomsky.* New York: Anchor Press, 1974.

Kasher, A., ed. *The Chomskyan Turn.* Oxford: Blackwell, 1991.

Sampson, G. *Liberty and Language.* Oxford: Oxford University Press, 1979.

Fred D'Agostino (1996)

CHRISTIANITY

The present entry is restricted to Christian belief and scarcely touches on the origins of "Christianity" or its history and institutional forms. Among Christian beliefs only a few can be treated; certain others, such as the existence and attributes of God, are discussed in other entries.

CHRISTIAN BELIEF

Perhaps the first thing that should be said about Christian belief is that it does not constitute a philosophy. That is to say, it is not a metaphysical system comparable, for example, to Platonism or the systems of Aristotle and Benedict de Spinoza. Although the body of Christian doctrine does consist largely of metaphysical beliefs, in the sense that they are beliefs whose scope transcends the empirical world, it differs from what are usually identified as philosophical systems by its essential relation to and dependence on particular historical events and experiences. Such systems as Platonism begin with philosophical concepts and principles and seek by means of these to construct a comprehensive mental picture of the universe. Christianity, on the other hand, begins with particular, nonrecurrent historical events that are regarded as revelatory and on the basis of which Christian faith makes certain limited statements about the ultimate nature and structure of reality.

The relationship between experience and discursive reflection in Christianity can be brought out by distinguishing two orders of Christian belief. There is a primary level, consisting of direct reports of experience, secular and religious, and a secondary level, consisting of theological theories constructed on the basis of these reports.

At the primary level Christian literature affirms a number of both publicly verifiable historical facts and "religious facts," or "facts of faith." The latter consist of incidents in the history of Israel as understood and participated in by the prophets and in the life of Jesus as he was responded to by the apostles, these events being seen by faith as revelatory of God. The resulting testimonies of the prophets and apostles are not formulations of theological doctrine but direct expressions of moments of intense religious experience. The four New Testament gospels are writings on this primary level, recording events that occurred either within the purview of secular history or within the religious experience of the early Christian community.

Within this primary stratum of Christian belief certain facts of faith have always stood out as being preeminently important. By means of these Christianity has defined itself in distinction to other religions. Among the total body of those who have called themselves Christians there is no universally agreed-on list of these defining facts of faith, except insofar as such lists have been adopted, locally or more widely, by particular Christian communions, sects, or movements. However, it is safe to say that the main streams of contemporary Christianity, claiming continuity of faith with the first Christian generation, affirm at least the following: the reality of God and the propriety of speaking of him in a threefold manner, as Father, Son, and Spirit; the divine creation of the universe; human sinfulness; divine incarnation in the person of Jesus, the Christ; his reconciliation of man to God; his founding of the Christian church and the continuing operation of his Spirit within it; and an eventual end to human history and the fulfillment of God's purpose for his creation. Stated in this general form these are facts of faith that cumulatively define Christianity. Many further tenets are regarded as essential by different subgroups within Christianity, but the above probably constitute the permanent core that is acknowledged by virtually the whole of Christendom, past and present.

The second order of Christian belief consists in theological theories or doctrines that seek to explain these facts of faith and to relate them to one another and/or to human knowledge in general. The formulation of doctrines is essentially a discursive and speculative activity, differing from theory construction in secular philosophy only in that the theologian includes in his data, and indeed accords a central and determinative importance to, the special facts of Christian faith.

This distinction can now be illustrated by reference to some of the central Christian themes, noting both the relevant facts of faith and the theological theories that have been developed about them.

CREATION. The doctrine of creation (which Christianity holds in common with Judaism) stands somewhat apart from the other doctrines to be described below. The others have arisen out of reflection on specific historical phenomena, but belief in the divine creation of the universe, although connected with the religious experience of absolute dependence on God, has presumably been arrived at primarily as an implicate of the monotheistic understanding of God as the sole ultimate reality.

The doctrine of the divine creation of the universe out of nothing stands in contrast to other conceptions of its origin. This doctrine denies that the universe is eternal, although the denial does not entail the belief that it was created at some moment in time—Augustine, for example, taught that time is itself an aspect of the created world. The doctrine also excludes the Platonic notion of a Demiurge fashioning the world out of a formless matter and the Neoplatonic notion of the physical universe's coming to be by emanation from the Absolute. In distinction to these ideas the doctrine of *creatio ex nihilo* asserts that the universe has been summoned into existence out of nothing (that is, not out of anything) by the creative will and purpose of God.

INCARNATION. Jesus was born about 5 BCE in Palestine and was executed by crucifixion at Jerusalem probably in 29 or 30 CE. There immediately arose a conviction among his disciples, reflected in all the New Testament documents, that he had been raised by God from the dead, and under the compulsion of this conviction the Christian church came into existence, witnessing to both the divine status and the saving power of Jesus, now proclaimed as the Christ.

The beliefs of Jesus' disciples about him are reflected in the four memoirs, or gospels, which were produced in different centers of the apostolic church during the second half of the first century. On the one hand, these depict him as fully and authentically human, subject, like other men, to temptation, hunger, pain, fatigue, ignorance, and sorrow. But at the same time they affirm that

he is Lord, Messiah (*Christos*), the Son of God. This extremely exalted view reaches its highest expression in the Fourth Gospel, which claims in its prologue to Jesus' life that the Word (Logos), which was in the beginning with God, and was God, and through which all things were made, "became flesh, and dwelt among us, full of grace and truth; and we have beheld his glory, glory as of the only Son from the Father" (John 1:14; the conception of the Logos in the Fourth Gospel derives both from the Word and the wisdom of God in the Old Testament and from the Logos as the universal principle of reason in Greek philosophy). The faith that Jesus was the Christ apparently arose out of a practical acceptance of his status as one who had authority to forgive sins, to declare God's mind toward man, to reveal the true meaning of the divine Law, to heal diseases, and to assume that men's eternal destiny and welfare was bound up with their responses to him. This practical acknowledgment of his unique authority probably crystallized into conscious conviction as to his deity under the impact of the resurrection events.

In the gospels these two beliefs, identifying Jesus both as a son of man and as the Son of God, occur together without any attempt to theorize about the relationship between them. Thus, this primary stratum of Christian literature contains, as data for theological reflection, reports of (*a*) the publicly observable fact that Jesus was a man, and (*b*) the fact of faith that he was divine, in that "in him all the fullness of God was pleased to dwell" (Colossians 1:19).

During its first four centuries of life these data provided the church with its chief intellectual task. The eventual outcome of the Christological debates, formalized by the Council of Chalcedon (451), was not to propound any definitive theory concerning the relationship between Jesus' humanity and his divinity but simply to reaffirm, in the philosophical language of that day, the original facts of faith. The various views that were from time to time branded as heretical came under this condemnation because directly or by implication they denied one or the other of the two fixed points of Christian thought in this field, the human and divine natures of Christ.

The first of the Christological heresies, the Docetism of some of the Gnostics in the first and second centuries, denied the real humanity of Christ, suggesting that he was a human being in appearance only. The motive behind this theory was to exalt his divine status, but the effect was to deny one of the foundation facts of Christianity as historically based faith. The next great heresy, Arianism, in the fourth century, went to the opposite extreme, denying continuity of being or nature between the Godhead and Christ and regarding him as a created being, so that "there was a time when he was not" (*ἦν ὅτε οὐκ ἤν*). It was in the controversy with Arianism that the notion of substance (*οὐσία, substantia*) became a key category in the Christological debates. Arius declared that the Son was *ὁμοιούσιον τῷ πατρί* (of *like* substance with the Father), whereas the Council of Nicaea (325), excluding Arianism as a heresy, insisted that the Son was *ὁμοούσιον τῷ πατρί* (of the *same* substance as the Father). It was made clear by Athanasius, the champion of orthodoxy, that the iota's difference between these formulations involved an immense religious difference, for only a savior who came from the Godward side of creation could offer man an ultimate salvation. This Homousian Christology was reaffirmed by the Council of Chalcedon and has ever since been the position of the main streams of historic Christianity.

Since the mid-nineteenth century a number of theologians (for example, the Ritschlian school and H. R. Mackintosh) who accept the Nicene and Chalcedonian affirmations of the full humanity and real deity of Christ have questioned the adequacy of the category of substance in terms of which that affirmation was made. They have pointed out that it belongs to the thought-worlds of Plato and Aristotle and that it is a static notion, contrasting in this respect with such characteristically dynamic biblical categories as purpose and action. Accordingly there is now a fairly widespread tendency to describe the incarnation as a complex event constituting God's self-revealing action in man's history. In the New Testament records we see God at work in and through a human life, dealing with human beings in a way that makes plain the divine nature in its relation to man. The acts and attitudes of Jesus toward the men and women with whom he had to do were God's acts and attitudes in relation to those particular individuals, expressed in the finitude of a human life. Along these and other lines Christological discussion continues.

THE TRINITY. The Trinitarian doctrine is a second-order Christian belief. It was gradually developed within the church both to take account of certain data at the experiential level and to aid the development of the general system of Christian doctrine, some of the key points of which are related by the Trinitarian framework.

The New Testament basis for this doctrine was the Christian community's threefold awareness of God, first as the transcendent moral creator witnessed to in the prophetic tradition received from Judaism; second, as

having been at work among them on earth in the person of Christ; and third, as the Holy Spirit, which was referred to apparently indiscriminately as the Spirit of God and the Spirit of Christ, inspiring and guiding both individuals and the Christian community.

The doctrine of the Trinity developed in close conjunction with Christology and made possible the completion of the church's thought concerning the person of Christ. For it had never been the accepted Christian conception that God, simply as such and in his totality, became man in the incarnation. The belief that "God was in Christ" (2 Corinthians 5:19) was held in conjunction with the belief that God was also and at the same time sustaining and governing the universe. The God who was incarnate in Christ was the God who had created heaven and earth. This was expressed by the affirmation that God is both Father and Son; and the reality of the Spirit, operating in the world both before and after the thirty or so years of the incarnation, required the further expansion into a Trinitarian formulation. Thus, the doctrine of the Trinity (*a*) asserts the full deity of Christ as the second person of the Trinity; (*b*) prohibits a too simple conception of incarnation (as one branch of the theological tradition has put it, Christ is *totus deus,* wholly God, but not *totum dei,* the whole of God); and (*c*) recognizes the universal presence and activity of God in the world as the divine Spirit. This latter point is of great practical importance because it entails a Christian message not only about God's actions in the past but also about a divine activity in the present that can directly affect the individual today.

In the Trinitarian discussions that accompanied the Christological debates one of the main questions concerned the issue of equality versus subordination within the Trinity. Is the Son subordinate to the Father, or the Spirit to both? The answer that was eventually embodied in the *Quicunque vult,* or "Athanasian" Creed, of the sixth century was that the members of the Trinity are coeternal and have an equal divine status; the Son is eternally begotten by the Father, and the Spirit eternally proceeds from the Father and the Son. (The latter point was the occasion of the rift in the sixth century between the Eastern church, with its center at Constantinople, and the Western church, with its center at Rome. In its original form the Nicene Creed described the Spirit as proceeding (only) from the Father. Later the Western church added the famous *filioque*—"and the Son"—an insertion that Eastern Christianity rejected as an unwarrantable tampering with the creed.)

In the accepted Trinitarian language the Father, the Son, and the Holy Spirit are spoken of as three "Persons," the Latin *persona* having been used to translate the Greek *ὑπόστασις* (which had displaced *πρόσωπον*—literally, "face"—in this context). *Persona* is not, of course, the equivalent of "person" in the modern sense of an individual center of consciousness and purpose. Originally a persona was the mask worn by an actor, then his part in the play, and then by further extension any part a person might play in life. Thus, whereas *τρεῖς ὑπόστασεις* suggests three divine entities, *tres personae* suggests three roles or functions of the deity. These two different conceptions have each been developed in Christian thought, leading to what have been called respectively "immanent," or "ontological," and "economic" theories of the Trinity.

According to the ontological theories the doctrine of the Trinity is an affirmation about the transcendent metaphysical structure of the Godhead. It asserts that God in his inner being consists of three divine realities that are individually distinct and yet bound together in a mysterious unity—"three in one and one in three." The extreme form of this view is the "social" conception of the Trinity as comprising three consciousnesses. According to the economic theories, on the other hand, the doctrine is about God specifically in his relation to the world. It asserts that the one God has acted toward humankind in three distinguishable ways—in creation and providence, in redemption, and in inner guidance and sanctification. God must indeed, in his inner being, be such as to become related in these ways to his creation, but this does not necessarily require the postulation of three distinct and yet intimately related divine realities.

REDEMPTION. That human beings are sinful is a theological statement of the observable fact that men and women are persistently self-centered and that even their highest moral achievements are quickly corrupted by selfishness. Yet although we thus fail, exhibiting a chronic moral weakness and poverty, our failure is not inevitable; we are ourselves, at least in part, responsible for it. The biblical story of the fall of man depicts this situation by means of the myth that man was originally created perfect but fell by his own fault into his present state, in which he is divided both in himself and from his fellows and God.

At its primary level of belief Christianity claims that by responding to God's free forgiveness, offered by Christ, men are released from the guilt of their moral failure (justification) and are drawn into a realm of grace in which

they are gradually re-created in character (sanctification). The basis of this claim is the Christian experience of reconciliation with God and, as a consequence, with other human beings, with life's circumstances and demands, and with oneself. The "justification by faith" of which Paul spoke, and which represented the main religious emphasis of the Reformation of the sixteenth century, means that men are freely accepted by God's gracious love, which they have only to receive in faith. In Paul Tillich's contemporary restatement, a man has only to accept the fact that although unacceptable even to himself, he is accepted by God.

In this case, work at the secondary level of theological reflection did not begin seriously until the church had been preaching the fact of divine reconciliation and atonement for about a thousand years. Anselm, in the eleventh century, taught that the death of Christ constituted a satisfaction to the divine honor for the stain cast upon it by man's disobedience, and this remains the core of Catholic atonement doctrine. Martin Luther and John Calvin, in the sixteenth century, spoke of Christ's death as a substitutionary sacrifice by which Christ suffered in his own person the punishment that was justly due humankind, and this remains the core of official Protestant atonement doctrine. In the nineteenth century, however, the thought was developed (going back to Anselm's contemporary Peter Abelard) that God's forgiveness does not need to be purchased by Christ's death, but that this brings home to the human heart both man's need for divine forgiveness and the reality of that forgiveness. There were in the twentieth century and on into the twenty-first century continuing efforts to understand Christ's redeeming work in a way that would bring together the valid insights in these and other traditional views, each of which by itself has seemed one-sided.

HEAVEN, HELL, AND JUDGMENT. Jesus impressed upon his hearers in the strongest possible terms the absolute importance of decisions made and deeds performed in this present life. He regarded men and women as free and responsible persons on whose daily choices depended their own final good and happiness or irretrievable loss and failure. In doing this he used the traditional language of heaven and hell, which were understood until comparatively recently in terms of a prescientific cosmology, with heaven located in the sky above our heads and hell in the ground beneath our feet. Heaven is now generally conceived of as the enjoyment of the full consciousness of God's presence and participation in the divine "kingdom," which represents the final fulfillment of God's purpose for his creation; and hell is viewed as self-exclusion from this.

There are many perennially debated questions in this area. Are men divinely predestined, some to eternal salvation and others to eternal damnation ("double predestination"), as Augustine and Calvin taught? Does "hell" signify an eternal state, or is it a temporally bounded purgatorial experience that might lead to eventual salvation? (The adjective *αἰώνιος*, which is used in the New Testament, can mean either "eternal" or "for the aeon, or age"). Or does "hell" perhaps signify sheer annihilation? Can the final frustration of God's purpose by the loss of part of his human creation be reconciled with his ultimate sovereignty, and does the idea of never-ending torment, as a form of suffering out of which no good is finally brought, rule out the possibility of a Christian theodicy? Are all men to be finally saved ("universalism"), or only some?

In relation to such questions it is perhaps useful to distinguish between two standpoints from which eschatological statements may be made. There is the existential standpoint of "real life," in which we exercise a fateful responsibility in our moral choices and are confronted with the tremendous alternatives of spiritual life and death, symbolized by heaven and hell. There is also the detached standpoint of theological reflection, in which it seems possible to deduce from the two premises of the sovereignty and the love of God that although damnation is abstractly conceivable and is known in existential experience as a dread possibility, God's saving purpose in relation to his creatures will nevertheless in the end be triumphant, and eternal loss will remain an unrealized possibility.

THE CHURCH. Although Christianity as historically institutionalized lies outside the narrow scope of this entry, it must be added that Christian faith has always drawn people together into a community of faith, or church. The largest Christian institution, the Roman Catholic Church, holds that the authentic Christian community is defined by its visible continuity, manifested in a succession of bishops and popes, with the earliest church. Protestantism holds that the Christian community is defined by a different continuity, that of faith, and affirms that the external institutions associated with Christian faith are continually in need of reformation in the light of the original Christian data embodied in the scriptures.

See also Abelard, Peter; Anselm, St.; Arius and Arianism; Aristotle; Calvin, John; Gnosticism; God, Concepts of; Heaven and Hell, Doctrines of; Luther, Martin; Neopla-

tonism; Platonism and the Platonic Tradition; Ritschl, Albrecht Benjamin; Spinoza, Benedict (Baruch) de.

Bibliography

THE BIBLE

Old Testament. 7th ed. Edited by R. Kittel. Stuttgart, 1951. Hebrew text.

New Testament. 2nd ed. Edited by G. D. Kilpatrick. London, 1958. Greek text.

American Revised Standard Version. New York, 1952–1957. Old and New Testaments and Apocrypha in translation.

New English Bible. Vol. 1: New Testament. London, 1961. Translation.

COLLECTIONS OF CHRISTIAN LITERATURE

Baillie, John, John T. McNeill, and Henry P. Van Dusen, eds. *Library of Christian Classics,* 26 vols. London and Philadelphia, 1954–.

Migne, J. P., ed. *Patrologia Graeca.* 162 vols. Paris, 1857–1866.

Migne, J. P., ed. *Patrologia Latina.* 221 vols. Paris, 1844–1864.

Roberts, Alexander, and James Donaldson, eds. *The Ante-Nicene Fathers.* 10 vols. Grand Rapids, MI, 1951.

Schaff, Philip et al., eds. *Nicene and Post-Nicene Fathers of the Christian Church.* Grand Rapids, MI, 1952–1956. Two series, each in 14 vols.

ENCYCLOPEDIAS

Campenhausen, H. F. von et al., eds. *Die Religion in Geschichte und Gegenwart.* 3rd ed. 7 vols. Tübingen: Mohr, 1957–1965.

Hastings, James, ed. *Encyclopedia of Religion and Ethics,* 13 vols. Edinburgh, 1908–1926.

Vacant, A. et al., eds. *Dictionnaire de Théologie catholique.* 3rd ed. 15 vols. Paris, 1923–1950.

HISTORY OF CHRISTIANITY

Latourette, K. S. *A History of the Expansion of Christianity.* 7 vols. New York and London: Harper, 1937–1945.

HISTORY OF CHRISTIAN THOUGHT

Harnack, Adolf von. *Lehrbuch der Dogmengeschichte.* 4th ed. 3 vols. Tübingen, 1909–1910. Third ed. translated by Neil Buchanan as *History of Dogma.* 7 vols. London, 1894–1899; New York: Russell and Russell, 1958.

Kelly, J. N. D. *Early Christian Doctrines.* London: A. and C. Black, 1958.

Seeberg, Reinhold. *Lehrbuch der Dogmengeschichte.* 4th ed. 4 vols. Basel, 1953–1954. Translated by C. E. Hay as *Textbook of the History of Doctrines.* Grand Rapids, MI, 1952.

Wolfson, H. A. *The Philosophy of the Church Fathers.* Cambridge, MA: Harvard University Press, 1956–.

ROMAN CATHOLICISM

Denzinger, H. J. D. *Enchiridion Symbolorum Definitionum et Declarationum de Rebus Fidei et Morum,* 29th ed. Freiburg, Germany, 1952. Translated by R. J. Defarrari as *The Sources of Catholic Dogma.* St. Louis, MO: Herder, 1957.

Schmaus, Michael. *Katholische Dogmatik.* 5 vols. Munich, 1948–1958.

Smith, G. D., ed. *The Teaching of the Catholic Church.* 2 vols. London: Burns, Oates and Washburne, 1948; New York: Macmillan, 1949.

Thomas Aquinas. *Summa Contra Gentiles.* Editio Leonina Manualis. Rome, 1934. Translated by A. C. Pegis, J. F. Anderson, V. J. Bourke, and C. J. O'Neil as *On the Truth of the Catholic Faith.* 4 vols. New York, 1955–1957.

Thomas Aquinas. *Summa Theologiae.* 6 vols. Editio altera Romana. Rome, 1894. Translated by the English Dominican Fathers. London, 1911–1922; New York, 1947–1948.

PROTESTANTISM

Barth, Karl. *Kirchliche Dogmatik.* Munich: Kaiser, 1932–. Translated as *Church Dogmatics,* edited by G. W. Bromiley and T. F. Torrance. Edinburgh: Clark, 1936–.

Calvin, John. *Institutio Christianae Religionis.* Basel, 1536. Critical ed. by P. Barth and W. Niesel. Munich, 1926. Translated by F. L. Battles as *Institutes of the Christian Religion,* edited by J. T. McNeill. 2 vols. Philadelphia: Westminster Press, 1961.

Luther, Martin. *A Commentary on St. Paul's Epistle to the Galatians.* Edited and translated by Philip Watson. London, 1953.

Luther, Martin. *Lectures on Romans.* Translated by W. Pauck. London and Philadelphia, 1961.

Luther, Martin. *Werke.* 58 vols. Weimar, 1883–1948.

Schleiermacher, Friedrich. *Der Christliche Glaube.* Critical ed. by M. Redeker. Berlin, 1960. Translated as *The Christian Faith,* edited by H. R. Mackintosh and J. S. Stewart. Edinburgh, 1928; New York: Harper and Row, 1963.

Tillich, Paul. *Systematic Theology.* 3 vols. Chicago: University of Chicago Press, 1951–1964.

EASTERN ORTHODOXY

Bulgakov, Sergei N. *L'Orthodoxie.* Paris, 1932. Translated by E. S. Cram as *The Orthodox Church.* London: Centenary Press, 1935.

Tsankov, Stefan. *Das orthodoxe Christentum des Ostens.* Berlin, 1928. Translated by D. A. Lowrie as *The Eastern Orthodox Church.* London, 1929.

Zernov, Nicholas. *Eastern Christendom.* New York: Putnam, 1961.

CONTEMPORARY ECUMENICAL MOVEMENT

Bell, G. K. A., ed. *Documents on Christian Unity.* 4 vols. London, 1924–1958.

Neill, Stephen C., and R. Rouse, eds. *A History of the Ecumenical Movement, 1517–1948.* Philadelphia: Westminster Press, 1954.

CREATION

Gilkey, Langdon B. *Maker of Heaven and Earth.* Garden City, NY: Doubleday, 1959.

THE PERSON OF CHRIST

Baillie, D. M. *God Was in Christ.* London and New York: Scribners, 1948.

Hendry, G. S. *The Gospel of the Incarnation.* Philadelphia: Westminster Press, 1958.

Knox, John. *Jesus: Lord and Christ.* New York: Harper, 1958.

Mackintosh, H. R. *The Doctrine of the Person of Jesus Christ.* London and New York: Scribners, 1912.

Pittenger, W. Norman. *The Word Incarnate.* New York: Harper, 1959.

ATONEMENT

Anselm. *Cur Deus Homo?,* edited by F. S. Schmitt. Bonn, 1929. Also in *Opera Omnia,* edited by F. S. Schmitt, Vol. II. Anonymous translation in Ancient and Modern Library of Theological Literature. London, 1889. "Satisfaction" theory.

Aulen, Gustaf. *Christus Victor.* Translated by A. G. Hebert. London: Society for Promoting Christian Knowledge, 1931. "Christus Victor" theory.

Denney, James. *The Christian Doctrine of Reconciliation.* London: George H. Doran, 1918. "Penal-substitutionary" theory.

Mozley, J. K. *The Doctrine of the Atonement.* New York: Scribners, 1916. Historical.

Rashdall, Hastings. *The Idea of Atonement in Christian Theology.* London: Macmillan, 1919. "Moral influence" theory.

TRINITY

Franks, R. S. *The Doctrine of the Trinity.* London: Duckworth, 1953. Historical.

Hodgson, Leonard. *The Doctrine of the Trinity.* London: Nisbet, 1943.

Welch, Claude. *In His Name.* New York, 1952. Welch and Hodgson represent the two types of theory described in this entry.

ESCHATOLOGY

Althaus, Paul. *Die letzten Dinge.* 4th ed. Gütersloh, 1933.

Cullmann, Oscar. *Christus und die Zeit.* Zürich, 1946. Translated by F. V. Filson as *Christ and Time.* Philadelphia, 1950; rev. ed., London, 1962.

Hügel, F. von. *Eternal Life.* Edinburgh: T. and T. Clark, 1912.

CHRISTIANITY AND OTHER RELIGIONS

Farmer, H. H. *Revelation and Religion.* London: Nisbet, 1954.

Kraemer, H. *Religion and the Christian Faith.* London: Lutterworth Press, 1956.

OTHER RECOMMENDED TITLES

Beaty, Michael, ed. *Christian Theism and the Problems of Philosophy.* Notre Dame, IN: University of Notre Dame Press, 1990.

Feenstra, Ronald, and Cornelius Plantinga Jr., eds. *Trinity, Incarnation, and Atonement.* Notre Dame: University of Notre Dame Press, 1989.

Flint, Thomas, ed. *Christian Philosophy.* Notre Dame, IN: University of Notre Dame Press, 1990.

Morris, Thomas. *The Logic of God Incarnate.* Ithaca, NY: Cornell University Press, 1986.

Murray, Michael, ed. *Reason for the Hope Within.* Grand Rapids, MI: Eerdmans, 1999.

Padgett, Alan, ed. *Reason and the Christian Religion: Essays in Honour of Richard Swinburne.* Oxford: Oxford University Press, 1994.

Plantinga, Alvin. *Warranted Christian Belief.* Oxford: Oxford University Press, 2000.

Swinburne, Richard. *The Christian God.* Oxford: Clarendon Press, 1994.

Swinburne, Richard. *Responsibility and Atonement.* Oxford: Clarendon Press, 1989.

Swinburne, Richard. *The Resurrection of God Incarnate.* Oxford: Clarendon Press, 2003.

Van Inwagen, Peter. *God, Knowledge and Mystery: Essays in Philosophical Theology.* Ithaca, NY: Cornell University Press, 1995.

Vesey, Godfrey, ed. *The Philosophy in Christianity.* Cambridge, U.K.: Cambridge University Press, 1990.

White, Vernon. *Atonement and Incarnation: An Essay in Universalism and Particularity.* Cambridge, U.K.: Cambridge University Press, 1991.

John Hick (1967)
Bibliography updated by Christian B. Miller (2005)

CHRYSIPPUS

(c. 279–206 BCE)

Chrysippus, the Stoic philosopher born at Soli, in Cilicia, became the third leader of the Stoa at Athens upon the death of Cleanthes, in 232 BCE. This post he held until his own death. Because of his defense of the Stoa against the attacks of Arcesilaus and the skeptical Academy, and undoubtedly also on the basis of his voluminous writings, it was said in antiquity "if there had been no Chrysippus, there would be no Stoa." He wrote 705 books, about half of which, judging from the catalog preserved by Diogenes Laertius, dealt with logic and language. None of his works is extant, though quotations from his books and assessments of some of his views have survived in the works of other ancient authors.

Chrysippus's epistemology is empirical. Presentations of objects are produced in the ruling part of the soul by movements engendered in the sense organs of the percipient. Illusory presentations can be distinguished from those that are veridical by deliberation, which consists in checking any given presentation against a fund of common notions, that is, families of remembered similar presentations; if the presentation is found to be sufficiently like some common notion, one may assent to it, thus acknowledging its veridical character.

Propositions are either simple or nonsimple. The truth condition of a simple proposition is the occurrence of the fact it conveys. The truth conditions of nonsimple propositions are functions of the truth-values of their ingredient propositions.

Chrysippus formulated five undemonstrated argument forms whose variables are to be specified by propositions. Among them are forms of the *modus ponens* and the *modus tollens* arguments. Arguments of varying complexity can be constructed by combining two or more of these basic forms. Chrysippus enjoyed a particular renown for his competence as a dialectician.

The moral philosophy of Chrysippus is concerned primarily with a statement of the final end of life and the relation of other things to it and with a consideration of the emotions and therapy for those enslaved by them. The final good is "to live in accordance with one's experience of the things which come about by nature." This is equivalent to living in accordance with reason, which in man supervenes upon instinct as a guide in life. The excellence of reason is wisdom, or knowledge of what is really good and what is really bad. Chrysippus's view in regard to the source of this knowledge is ambivalent. On the one hand—and this is obviously the doctrine that coheres best with his epistemology—it derives from generalizations made upon particular experiences. On the other hand, there are fragments implying that his knowledge is innate.

Emotions are great obstacles to happiness and are to be totally eradicated. In keeping with his monistic psychology, which rejects the Platonic doctrine of a tripartite soul, Chrysippus conceived of an emotion as a recently formed false judgment about the goodness or badness of something; such a judgment causes "a forceful and excessive impulse." Therapy for the emotions consists in persuading their victims that the judgments constituting the emotions are false.

The dominant motifs of the natural philosophy of Chrysippus are monism and determinism. The one substance that converts periodically into an elaborately structured universe has two constant aspects, a passive one and an active one. The passive is matter; the active is identified variously as reason, pneuma (spirit or breath), and God. Chrysippus regards so-called individual substances not as discrete units of matter but rather as "parts" of one primary substance. Everything that occurs is controlled unexceptionably by fate, which is "the continuous causal chain of the things that exist." Nothing comes about except in accordance with antecedent causes. Even in the case of states of affairs that might seem to be of a spontaneous or uncaused nature, "obscure causes are working under the surface." Chrysippus believed that humans were responsible for their conduct, and he sought in several ways to show that such a belief was not undermined by the rigorously deterministic view he espoused.

See also Arcesilaus; Cleanthes; Determinism, A Historical Survey; Dialectic; Diogenes Laertius; Greek Academy; Platonism and the Platonic Tradition; Stoicism.

Bibliography

Arnim, Hans V. *Stoicorum Veterum Fragmenta*. Leipzig: Teubner, 1903.

Bréhier, E. *Chrysippe et l'ancienne Stoïcisme*. Paris, 1951.

Mates, B. *Stoic Logic*. Berkeley: University of California Press, 1953.

Pohlenz, M. *Die Stoa*. Göttingen: Vandenhoeck and Ruprecht, 1948.

Josiah B. Gould Jr. (2005)

CHUANG TZU

See *Zhuangzi*

CHUBB, THOMAS

(1679–1747)

Thomas Chubb, the English Arian and deist, was born at East Harnham, near Salisbury, the son of a maltster. Receiving little formal education, he read widely in geography, mathematics, and theology while working as apprentice to a glovemaker and, later, as a tallow chandler. At one time he lived in the house of Sir Joseph Jekyll, master of the rolls, in the capacity, it is alleged, of a sort of superior servant. Through the kindness of friends (one of whom was the celebrated surgeon William Cheselden) and the sales of his candles, his last years, spent at Salisbury, were largely devoted to study and to the presidency of a debating society. Chubb's importance, frequently overlooked, lies in the fact that a self-educated and humble artisan developed a good style of writing and mastered the prevalent rationalistic thinking sufficiently well to compete on equal terms with highly educated upper-class scholars and divines. He was the first, and one of the few, leading English deists of poor circumstances (only Peter Annet and Thomas Morgan shared this humble background). With Chubb it was apparent that deism had filtered down to the level of the common people and had become widespread.

Chubb's first publication was an Arian tract, *The Supremacy of the Father Asserted,* inspired by William Whiston's *Primitive Christianity Revived* of 1711 and published in 1715 upon the recommendation of Whiston.

Although Chubb went through an early phase of Arminianism and was always hard pressed to reconcile Jehovah with the rationalistic concept of a Supreme Being, he nevertheless became and remained a "Christian deist." Skeptical of the Jewish revelation, he was less so of

the Islamic and openly accepted the Christian, at least as he understood it. In *The True Gospel of Jesus Christ asserted* (1732) and *The True Gospel of Jesus Christ Vindicated* (1739) he identified the essence of Christianity with the few simple principles of natural religion as found, for example, in Lord Herbert of Cherbury. He openly compared the propagation of primitive Christianity with the then current spread of Methodism and thereby rejected the claims of supernatural power associated with the early church. He defended his sort of rationalistic Christianity against some of the aspersions of that formidable deist Matthew Tindal. Although Voltaire had some kind words to say about Chubb, it is unlikely that he had read many of Chubb's tracts and certainly did not accept the concept of "Christian deism."

Chubb, like the general run of deists, found reason sufficient to guide humankind to God's favor and the happiness of another world; he was suspicious of mystery and of miracles and critical of some passages in the Scriptures; he regarded revelation not as divine but as the work of honest men who gave a fair and faithful account of matters of fact; he was dubious about a particular providence and, therefore, of prayer; he argued against prophecy and miracle and believed in the dignity of human nature and in free will. Among the multitudinous answers to Chubb from the more orthodox, the foremost came in 1754 from Jonathan Edwards of Massachusetts. *A Careful and Strict Enquiry into The modern prevailing Notions of the Freedom of Will, Which is supposed to be essential To Moral Agency, Vertue and Vice, Reward and Punishment, Praise and Blame,* Edwards's chief claim to philosophical fame, devotes no fewer than nineteen pages to the refutation of Chubb on free will. Chubb, it may reasonably be inferred, was widely read in America.

In fine, though adding little constructive thought to the deistic movement, this humble and least formally educated of the English deists was definitely one of its most valuable and popular spokesmen. In the nonpejorative sense of the term he was a candid freethinker.

See also Deism.

Bibliography

Chubb was prolific in publication, and his ardent deism was expressed in the titles of a few of his chief works: *The Comparative Excellence and Obligation of Moral and Positive Duties* (1730); *A Discourse concerning Reason, With regard to Religion and Divine Revelation* (1731); *The Sufficiency of Reason in Matters of Religion Farther Considered* (1732); *The Equity and Reasonableness of the Divine Conduct, In Pardoning Sinners upon Their Repentence Exemplified* (1737), which was directed against Bishop Butler's famous *Analogy of Religion* of the previous year; *An Enquiry into the Ground and Foundation of Religion. Wherein Is shewn, that Religion Is founded in Nature* (1740); and *A Discourse on Miracles, Considered as Evidence to Prove the Divine Original of a Revelation* (1741).

Other works by Chubb include *Four Tracts* (1734) and *Some Observations Offered to Publick Consideration…. In which the Credit of the History of the Old Testament Is Particularly Considered* (1735). The posthumous *Works of Mr. Thomas Chubb,* 2 vols. (London, 1748) contains the valuable "Author's Farewell to his readers."

See also Sir Leslie Stephen's *History of English Thought in the Eighteenth Century* (London: Smith Elder, 1876; the paperback, 2 vols., New York: Harcourt Brace, 1963, follows the revised edition of 1902) and the general bibliography under the "Deism" entry.

OTHER RECOMMENDED TITLE

Bushell, Thomas L. *The Sage of Salisbury: Thomas Chubb, 1679–1747.* New York: Philosophical Library, 1967.

Ernest Campbell Mossner (1967)
Bibliography updated by Tamra Frei (2005)

CHU HSI

See *Zhu Xi*

CHURCH, ALONZO

(1903–1995)

Alonzo Church, an American logician and philosopher, was born in Washington, D.C. He received his PhD from Princeton in 1927, having written his dissertation under Oswald Veblen on alternatives to the axiom of choice. He spent a year at Harvard and then a year in Europe, studying first at Göttingen and then at Amsterdam with L.E.J. Brouwer. He returned to Princeton where he was professor of mathematics from 1929 to 1967, after which he moved to UCLA to become professor of mathematics and philosophy. He retired from teaching at UCLA in 1990. Church's most important contributions to logic were his analysis of the concept of effective computability and his proof of the undecidability of first-order logic (Church's theorem).

A function of natural number is effectively computable if there is an algorithm—a surefire method requiring no ingenuity to follow—that will yield the value of the function for any given natural number as input. Church devised a formal system, the lambda calculus (which subsequently became an important tool in computer science), and proposed that a function of natural numbers be taken to be computable if it is *lambda definable*—definable by way of a

formula in the calculus. The analysis has little to recommend it initially, but experience with intuitively computable functions led Church to conjecture that every such function is lambda definable—a conjecture now known as *Church's thesis.* Alan Turing gave a more compelling analysis of computability in terms of abstract computing machines (Turing machines) and it was subsequently shown that lambda definability is equivalent to this notion of *Turing computability*. Various other analyses have been proposed and all have turned out to be equivalent to Church's definition. This is often regarded by logicians as evidence for the correctness of the conjecture. Church's thesis is now almost universally accepted.

Say, for instance, that a property of an expression is (effectively) decidable if there is an algorithm for deciding whether or not any given expression has the property. This notion can be identified with a certain sort of effective computability by supposing that all expressions have been assigned numbers (in some effectively determinate way) and then saying that a property of an expression is *effectively decidable* if there is an algorithm that will yield 0 (*no*) when applied to the number for the expression if the expression does not have the property and will yield 1 (*yes*) if the expression does have the property. If one then identifies the existence of such an algorithm with the lambda definability (or Turing computability) of that function, as Church's (or the Church-Turing) thesis proposes, one has a precise definition of effective decidability. Church's theorem shows that the property of being a valid formula of first-order predicate logic is not decidable in this sense. Thus, unlike the propositional calculus for which truth tables yield an effective procedure for deciding tautologousness, the validity of a first-order formula can not be decided, yea or nay, by any uniform algorithmic procedure.

Church's most important philosophical contributions involve the realism-nominalism controversy in the philosophy of mathematics and logic and problems and theories about meaning. He was a realist or Platonist about abstract entities and provided powerful arguments against various attempts to explain away such entities.

Rudolf Carnap and others associated with logical positivism displayed a general animosity toward such abstracta as numbers, functions, properties, and propositions. Carnap attempted to analyze sentences ostensibly ascribing belief in a proposition to someone in terms of sentences and a relation of "intensional isomorphism" between sentences. Roughly, the relation holds when the sentences in question are made up of necessarily equivalent parts, arranged in the same order. Church objected that a sentence ascribing a belief to someone does not mention a sentence of a particular language. He goes on to give a detailed and compelling refutation of Carnap's specific proposal. The method used, what is now called the "translation argument," appears to be of general applicability and makes it seem implausible that any replacement of propositions by more concrete things such as sentences will be successful. Church also raised powerful objections to nominalist maneuvers by A. J. Ayer and Israel Scheffler. Problems about the notion of synonymy were raised by Nelson Goodman and Benson Mates. Church answered these decisively.

Church's work on the logic of sense and denotation, a formal intensional logic incorporating some of Gottlob Frege's ideas about meaning, was one of his most important projects for philosophy, but it remains unfinished. The basic new idea is the "delta-relation"—the relation that holds between the sense of an expression and the denotation of that expression in some possible (N.B.) language. This is taken to be a logical relation and it is said that the sense is a *concept of* the denotation. It is postulated that a concept (the sense of some expression in some possible language) is a concept of at most one thing. And if *F* is a concept of a function *f* and *X* is a concept of an object *x*, then *F[X]* is a concept of *f(x)*. Church assumes that one can construe concepts of functions as certain functions on concepts, so that *F[X]*, plausibly taken to be a certain complex entity, is just construed as application of the function *F* to an argument *X*.

Various difficulties were encountered in working out this last idea, as well as in developing an axiomatic treatment of a *criterion of identity* for concepts that would render them suitable for the analysis and logic of the propositional attitudes—belief, knowledge, and the like. Modifying Carnap's notion of intensional isomorphism, Church proposed that two sentences (or other complex expressions) express the same proposition (or concept) if they are *synonymously isomorphic*—roughly, that they consist of synonymous expressions arranged in the same order. The development of axioms for the logic of sense and denotation that this idea suggests Church calls "Alternative (0)." Church was unable to complete an adequate formalization of this important conception.

See also Ayer, Alfred Jules; Brouwer, Luitzen Egbertus Jan; Carnap, Rudolf; Computability Theory; First-Order Logic; Frege, Gottlob; Goodman, Nelson; Logic, History of; Mathematics, Foundations of; Meaning; Realism; Turing, Alan M.

Bibliography

PRIMARY WORKS

"A Formulation of the Logic of Sense and Denotation." In *Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer*, edited by Paul Henle. New York: Liberal Arts Press, 1951.

"Logic and Analysis." *Atti del XII Congresso Internazionale de Filosofia* Venezia, 12–18 settembre, 1958.

"On Carnap's Analysis of Statements of Assertion and Belief." *Analysis* 10 (1950): 97–99.

"Outline of a Revised Formulation of the Logic of Sense and Denotation," pt. 1. *Noûs* 7 (1973): 24–33.

"Outline of a Revised Formulation of the Logic of Sense and Denotation," pt 2. *Noûs* 8 (1974): 135–156.

"Propositions and Sentences." In *The Problem of Universals—A Symposium*, 1–12. Notre Dame, IN: University of Notre Dame Press 1956.

SECONDARY WORKS

Anderson, C. Anthony. "Alonzo Church's Contributions to Philosophy and Intensional Logic." *The Bulletin of Symbolic Logic* 4 (1998): 129–171.

C. Anthony Anderson (2005)

CHURCH FATHERS

See *Patristic Philosophy*

CHURCH'S THESIS

See *Computability Theory*

CHWISTEK, LEON

(1884–1944)

Leon Chwistek, a Polish mathematical logician, philosopher, aesthetician, essayist, and painter, was a lecturer at the University of Kraków and from 1930 a professor of mathematical logic at the University of Lvov.

THEORY OF REALITIES

The central problem of Chwistek's philosophy was a criticism of the idea of a uniform reality. It had been shown by Bertrand Russell that in logic admission of the totality of all functions of x produces contradictions; Chwistek claimed that in philosophy, likewise, many obscure and misleading thoughts result from the assumption of a single all-inclusive reality.

The results of this criticism led Chwistek to the thesis of a plurality of realities. Out of many possible realities four are particularly important to philosophy. The first, the reality of natural objects, is assumed by common sense; natural objects are of a given form regardless of our perception. Chwistek's defense of natural reality and our knowledge of it is reminiscent of the British common-sense philosophy of the nineteenth century. The objects studied in physics are not natural; the telescopic and microscopic worlds, matter, and the particles upon which the forces are supposed to act form a second reality. They are constructions, not something naturally given. The third reality, that of impressions, the elements of sensation, as studied by David Hume or Ernst Mach, forms the world of appearances. The fourth reality is that of images, produced by us and dependent on our will, fantasy, and creative processes.

All four of these realities are necessary to account for our knowledge. In addition, when we reflect that we speak about a reality, we cannot include ourselves or our reflection in this reality. Such a reflection must be a part of a higher reality. Otherwise confusions and contradictions arise. The act of discourse cannot be a part of the universe of discourse.

AESTHETICS

Chwistek applied the doctrine of plurality of realities to investigations in many areas—aesthetics, for example. Natural reality is dealt with by primitive art. In primitive art each object is given one color only, and perspective is not obeyed. The primitivist paints not as he sees but as things are supposed to be by themselves. He uses his vision, but mainly he uses his knowledge about the world. Realism in art depicts the physical reality as it is conceived at a given time. Impressionism is the art of the reality of impressions; it flourished in a society that had developed psychological research and made psychologism its fundamental scientific method. Futurism is the art of free images, of an actively created reality of fantasy and mental constructions.

In each style of art the artist tries to give a perfect form to his creation independent of the kind of reality he is working with. The form is the common feature of all works of art. Thus, Chwistek justified all styles by relating them to different realities, and he advocated formism: evaluation of form, not of reality, is the proper aesthetic evaluation.

MATHEMATICS AND SEMANTICS

Chwistek extended his pluralism to mathematics. There is no one system of mathematics, but there are many mutually exclusive systems. Various geometries coincide only

in part. When we build analysis based on logic, we can accept, reject, or accept the negations of some extralogical existence axioms, such as the axiom of choice, the axiom of infinity, and the assumption of the existence of transfinite cardinal numbers. Logic itself should not decide any existence problem.

This restrained program for logic was paired with the requirement that logic be understandable in a nominalistic manner and deal with expressions in a constructive, mechanically computable way. Among principles often accepted as logical are some propositions questionable from the constructivist point of view—for example, the axiom of reducibility and the axiom of extensionality. The axiom of reducibility has to do with the distinction between predicative and impredicative concepts. An impredicative concept is a concept definable only by a definiens containing a quantifier that accepts as one of its values the very concept being defined. Russell and Chwistek ruled out such definitions as involving a vicious circle.

As was incisively pointed out by Kurt Gödel (in *The Philosophy of Bertrand Russell*, P. A. Schilpp, ed. [Evanston, IL, 1946], pp. 135–138), impredicative definitions involve a vicious circle only if one takes, as Chwistek did and Russell did not, a nominalistic attitude toward logic. Only if the quantifier is understood as a summary reference (infinite conjunction) to all of its values that are expressions and if one of the values of a quantifier that occurs in the definiens is the expression that is the definiendum do we presuppose what we want to define. Russell was not a nominalist. His exclusion of impredicative definitions was a way of avoiding antinomies. By differentiating between ranges of values of variables according to the way the quantifier binding a variable occurs, Russell constructed the ramified theory of logical types. This is a somewhat awkward theory. In analysis we want to speak about, for example, the real number that is the least upper bound of a set of real numbers that has a bound. To introduce this concept we must quantify over real numbers greater than all real numbers of a class that includes the least of them. Russell's theory avoids this impredicativeness by setting the least upper bound in a different logical type from the starting real numbers. But then the least upper bound and the real numbers involved cannot be values of the same variables, and several statements about particular sets of real numbers (for example, that a given function is continuous) are impossible.

To overcome this difficulty Russell accepted the axiom of reducibility, which says that every propositional function is coextensive with a predicative one. In many cases we cannot construct such a predicative function, and therefore constructivists, such as Chwistek, cannot accept this axiom. Moreover, for a nominalist, that two propositional functions are coextensive is not a sufficient guarantee of their identity. Thus, Chwistek attempted the task, which Russell called "heroic," of forming a purely constructivist system of the foundations of mathematics without impredicative definitions, the axiom of reducibility, or the axiom of extensionality. He observed, as F. P. Ramsey did, that results similar to Russell's can be obtained by the simple theory of types (where one distinguishes only between variables ranging over individuals, properties of individuals, properties of such properties, etc.) instead of the more complicated ramified theory. But simple type theory is inconsistent with the axiom of intensionality, which Chwistek wanted to be free to accept and which asserts the nonidentity of the concepts defined by two different propositional functions (even if they are coextensive).

The systems Chwistek constructed for the foundations of mathematics were such that they answered the philosophical needs of their author. They were admittedly more complicated than Russell's. "But it may be erroneous to think that clear ideas are never complicated; while we must agree that many simple ideas are, as a matter of fact, very obscure." Chwistek presented several formulations of his attempts at a constructivist theory, all of them too sketchy to be judged definitive. The relation to other constructivist systems is hard to establish. The last few versions were called "rational metamathematics." This theory deals with expressions, some of which are theorems.

A principal part of rational metamathematics, the fundamental system of semantics, uses two specific primitive signs, *c* and *, about which we stipulate that *c* is an expression and that if *E* and *F* are expressions, then **EF* is an expression. These formation rules assign a definite tree (or grouping) structure to each finite expression as well as to any two expressions written one after the other. Some of the allowed combinations of *c* and * may have no meaning—in this Chwistek was a formalist. To some other expressions we assign meaning, and in accordance with this assignment we accept proper axioms. We take 0 to be an abbreviation of **cc*. The fundamental substitution pattern (*EFGH*)—which is read "*H* is the result of substituting *G* for every occurrence of *F* in *E*"—is taken to be an abbreviation of *****EE***FF***GG***HH*. The Sheffer stroke function, |*EF*, is regarded as an abbreviation of ****EE****EE***EE*****FF****FE***FF*. Identity = *EF* stands for (*EOOF*).

See also Aesthetics, History of; Aesthetics, Problems of; Gödel, Kurt; Russell, Bertrand Arthur William; Semantics.

Bibliography

SELECTED WORKS BY CHWISTEK

Wielość Rzeczywistości (Plurality of realities). Kraków, 1921.

"The Theory of Constructive Types." *Annales de la Société Polonaise de Mathématique* 2 (1924, for 1923): 9–48, and 3 (1925, for 1924): 91–141.

"Fondements de la métamathématique rationnelle." *Bulletin de l'Académie Polonaise des Sciences et des Lettres,* Series A (1933): 253–264. Written with W. Hetper and J. Herzberg.

"Remarques sur la méthode de la construction des notions fondamentales de la métamathématique rationnelle." *Bulletin de l'Académie Polonaise des Sciences et des Lettres,* Series A (1933): 265–275. Written with W. Hetper and J. Herzberg.

Granice Nauki. Lvov and Warsaw, 1935. Rev. ed. translated by H. C. Brodie and A. P. Coleman as *The Limits of Science.* London: K. Paul, Trench, Trubner, 1948.

Wielość Rzeczywistości w Sztuce (Plurality of realities in art). Edited by K. Estreicher. Warsaw, 1960.

Pisma Filozoficzne i Logiczne (Philosophical and logical writings). Edited by K. Pasenkiewicz, 2 vols. Warsaw: Naukwoe, 1961–1963.

A reformulation of Chwistek's system of semantics by John Myhill appears in the *Journal of Symbolic Logic* 14 (1949): 119–125, and 16 (1951): 35–42.

Palace Boga. Próba rekonstrukcji. Warsaw: Panstwowy Instytut Wydawniczy, 1968.

Niejedna rzeczywistosc: racjonalizm krytyczny Leon Chwistka. Sens i rzeczywistosc. Crakow: Inter esse, 2004.

H. Hiż (1967)

CICERO, MARCUS TULLIUS

(106–43 BCE)

The Roman orator and statesman Marcus Tullius Cicero, of Arpinum, had a lifelong interest in philosophy and wrote a number of philosophical works during periods of forced retirement from public life. He was well acquainted with the four main Greek schools of his time and counted among his friends and teachers the Epicureans Phaedrus and Zeno, the Stoic Posidonius, the Peripatetic Staseas, the Academics Philo and Antiochus, and many others. He identified himself primarily with the Academy, though he found much to admire also in the Stoa and Lyceum. He rejected Epicureanism.

In a famous passage in a letter to Atticus (xii, 52, May 21, 45 BCE), with reference to some of his books on philosophy, Cicero calls them copies ("apographa"), written with little effort; he supplied only the words ("Verba tantum adfero, quibus abundo"). A week earlier he had written: "It is incredible how much I write, even at night; for I cannot sleep" (*Ad Atticum* xiii, 26). Modern scholars have found in such passages support for the view that these writings are chiefly valuable for the reconstruction of lost Greek originals, which Cicero in his haste sometimes misunderstood or jumbled together. The search for sources has been a major preoccupation of Ciceronian scholars for almost a century.

A more generous view is that in spite of his own statements Cicero's philosophical writings are more than hasty copies of Greek originals; they present a fairly coherent and modestly original system of thought. At a minimum Cicero took from the Academy a framework for his views. The Platonism of the New Academy had abandoned the search for truth and was occupied, rather, with the confrontation of conflicting opinions. Carneades, its leading spokesman, had even devised criteria for preferring one opinion to another. Within such a framework Cicero examines alternative views and makes his selection (though not necessarily in terms of Carneades' criteria). The views examined extend to all three commonly accepted branches of philosophy—logic, physics, ethics—and the presentation follows an orderly plan. Within this broad coverage, however, are many unresolved conflicts; clearly, Cicero's primary purpose was to offer to his Roman readers a wide range of philosophical opinions rather than to construct a well-integrated system.

PHILOSOPHY AND RHETORIC

Whatever originality Cicero's views possess is not in their components (he believed that the Greeks had already exhausted the varieties of possible opinions) but in their combination. The most conspicuous feature of his thought is the union of philosophy with rhetoric. This union carries with it some criticism of Socrates, who was blamed for their separation (see *De Oratore* iii, 61), and appears to align Cicero with Isocrates rather than Plato; yet he does not consider the union incompatible with Platonism. Carneades had prepared the way for a reconciliation between rhetoric and the Academy when he made philosophy a contest between opinions, and Greek theoretical rhetoricians had long since sought to implement Plato's prescription in the *Phaedrus* for a scientific rhetoric. Cicero could also point to the literary excellence of the dialogues as evidence that Plato was a master of the rhetorical art (ibid. i, 47).

The union of rhetoric and philosophy gave Cicero the materials for construction of his humanistic ideal.

The highest human achievement lies in the effective use of knowledge for the guidance of human affairs. Philosophy and the specialized disciplines supply the knowledge, and rhetorical persuasion makes it effective. Each is useless without the other, and the great man is master of both. Cicero associates this ideal with a free society—that is, a constitutional republic in which persuasion rather than violence is the instrument of political power. He believes that Rome has the essential features of such a state but that unless a great man is found to guide it, its freedom is in jeopardy.

Commitment to the union of eloquence and knowledge led Cicero to the view that if the statesman-philosopher is to speak persuasively on all subjects, he must have knowledge of all subjects. But recognizing the impossibility of such a requirement, Cicero advocated liberal education as the best approximation. An important part of liberal education is the study of philosophy, and Cicero's philosophical works provided materials for this study. Thus, in his philosophical writings no less than in his great public orations, he was combining wisdom and eloquence in the service of the Roman people.

PHILOSOPHICAL WORKS

The literary form that Cicero used emphasizes his didactic intent. Most of the philosophical works are dialogues, preceded by an introduction in defense of philosophical studies. The speakers are distinguished Romans, including Cicero himself, and frequently the listeners are young men just beginning their political careers. Conflicting views are presented in long speeches, with few interruptions. Sometimes the clash of opinions leads to insult and denunciation, especially when Epicureans are involved, but personal abuse of one speaker by another is avoided. There is hardly a vestige of dramatic conflict in such dialogues as *Tusculanae Disputationes,* where the conversation is between a young man and his preceptor. In two late works, *De Officiis* (*On Duties,* addressed to Cicero's son) and *Topica* (addressed to a young lawyer, Trebatius), the dialogue form is discarded.

In logic Cicero wrote *Academica,* in two versions (45 BCE), on the dispute between dogmatists and Academic skeptics about the criterion of truth; only portions of these are extant. *Topica* (44 BCE), though usually grouped with the rhetorical works, is also on logic. The title is from Aristotle, but the treatment is not. Cicero compiles a single exhaustive list of kinds of argument without distinction between the philosophical and the rhetorical.

There are three works, planned as a unit, on physics: (1) *De Natura Deorum,* (2) *De Divinatione,* and (3) *De Fato* (45–44 BCE). They present Epicurean, Stoic, and Academic arguments and counterarguments about religion and cosmology. Cicero himself was inclined to accept the Stoic arguments for a divine providence, but he rejected the Stoic doctrine of fate.

The major ethical writings are *De Finibus Bonorum et Malorum* (45 BCE), in which Epicurean, Stoic, and Peripatetic ethical views are examined; *Tusculanae Disputationes* (45 BCE), on fear of death, on pain, on distress of mind, and on other matters; and *De Officiis* (44 BCE), a practical ethics based on Stoic principles.

On political theory Cicero wrote two dialogues with titles taken from Plato. There is *De Re Publica* (51 BCE), from which the famous "Dream of Scipio" is an excerpt. The subject matter of the "Dream" ensured its preservation; it portrays the virtuous soul enjoying a more perfect existence after death in the region above the moon. The rest of the work is fragmentary. The other political dialogue, *De Legibus* (date uncertain), depicts Roman law as a very nearly perfect realization of Greek (chiefly Stoic) theory.

Some of the rhetorical works, especially the first book of *De Oratore* (55 BCE), discuss the relation of philosophy to rhetoric and present the ideal of the great man in whom both are united.

Minor works on philosophical themes include *Paradoxa Stoicorum, De Senectute, De Amicitia,* and the lost *Consolatio* and *Hortensius.* Cicero also translated two Platonic dialogues, *Protagoras* (lost) and *Timaeus* (W. Ax, ed., Leipzig, 1938).

See also Ancient Skepticism; Antiochus of Ascalon; Carneades; Greek Academy; Hellenistic Thought; Philo of Larissa; Stoicism.

Bibliography

TEXTS AND TRANSLATIONS

The standard Latin editions for most of Cicero's works are published by Teubner (various editors). New critical editions of the philosophical works are being published by Oxford University Press; two have appeared (more are imminent): *De Finibus,* edited by L. D. Reynolds (1998), and *De Officiis,* edited by Michael Winterbottom (1994). Texts and facing translations of the entire corpus, including all philosophical works except a partial Latin translation of Plato's *Timaeus,* are published by the *Loeb Classical Library,* various editor-translators (Cambridge, MA: Harvard University Press, 1914–1949). Superior translations of philosophical works include:

Academica: Brittain, Charles. *Cicero: On Academic Scepticism.* Indianapolis, IN: Hackett, 2005.

De Amicitia: Powell, Jonathan G. F. *Cicero: On Friendship and The Dream of Scipio.* Warminster: Aris and Phillips, 1990.

De Fato: Sharples, Robert W. *Cicero: On Fate, and Boethius: The Consolation of Philosophy IV.5–7, V.* Warminster: Aris and Phillips, 1991.

De Finibus Bonorum et Malorum: Annas, Julia, and Raphael Woolf. *Cicero: On Moral Ends.* Cambridge, U.K.: Cambridge University Press, 2001.

De Legibus: Zetzel, James E. G. *Cicero: On the Commonwealth and On the Laws.* Cambridge, U.K.: Cambridge University Press, 1999.

De Natura Deorum: Walsh, Patrick G. *Cicero: The Nature of the Gods.* Oxford: Oxford University Press, 1997.

De Officiis: Atkins, E. Margaret, and Miriam T. Griffin. *Cicero: On Duties.* Cambridge, U.K.: Cambridge University Press, 1991.

De Oratore: May, James M., and Jakob Wisse. *Cicero: On the Ideal Orator.* New York: Oxford University Press, 2001.

De Re Publica: Zetzel, James E. G. *Cicero: On the Commonwealth and On the Laws.* Cambridge, U.K.: Cambridge University Press, 1999.

Topica: Reinhardt, Tobias. *Cicero: Topica.* Oxford: Oxford University Press, 2003.

Tusculanae Disputationes: Graver, Margaret. *Cicero on the Emotions: Tusculan Disputations 3 and 4.* Chicago: University of Chicago Press, 2002.

Dyck, Andrew R. *A Commentary on Cicero, De Legibus.* Ann Arbor: University of Michigan Press, 2004.

Dyck, Andrew R. *A Commentary on Cicero, De Officiis.* Ann Arbor: University of Michigan Press, 1996.

Gawlick, Gunther, and Woldemar Görler. "Cicero." In *Die Philosophie der Antike,* Vol. 4, edited by Helmut Flashar. Basel: Schwabe, 1994. A comprehensive survey of his philosophical work and thought, with extensive, analytical bibliographies.

Inwood, Brad, and Jaap Mansfeld, eds. *Assent and Argument: Studies in Cicero's Academic Books.* Leiden: Brill, 1997.

MacKendrick, Paul. *The Philosophical Books of Cicero.* London: Duckworth, 1989. Useful synopses of each work.

Powell, Jonathan G. F., ed. *Cicero the Philosopher.* Oxford: Oxford University Press, 1995. Articles on diverse aspects of his thought.

Powell, Jonathan G. F., and John A.North, eds. *Cicero's Republic = Bulletin of the Institute of Classical Studies,* suppl. 76. London: University of London, 2001. Articles on his political thought.

Rawson, Elizabeth. *Cicero: A Portrait.* Ithaca, NY: Cornell University Press, 1983. A thoughtful and engaging biography emphasizing his intellectual life.

P. H. DeLacy (1967)

Bibliography updated by Stephen A. White (2005)

CIRCULARITY IN EPISTEMOLOGY

See *Epistemology, Circularity in*

CIVIL DISOBEDIENCE

The idea of civil disobedience comes out of the tradition of social and political protest whose best known advocates are the nineteenth-century American transcendentalist Henry David Thoreau, the Indian reformer Mohandas Gandhi, and the American civil rights leader, Martin Luther King, Jr. While the idea of civil disobedience has diverse roots, the views of these activist/thinkers set the stage for academic and popular discussion.

Philosophical discussions of civil disobedience generally focus on two questions. First, what is civil disobedience? Second, can acts of civil disobedience be morally justified?

DEFINING "CIVIL DISOBEDIENCE"

The definition of civil disobedience that best accords with the tradition of Thoreau, Gandhi, and King categorizes acts as civil disobedience if they have four features. They must be: (1) illegal; (2) nonviolent; (3) public; and (4) done to protest a governmental law or policy.

Thoreau's refusal to pay his taxes has all these features. It was illegal, nonviolent, and public. (Unlike a tax evader, Thoreau did not hide his not paying.) And, it was done to protest policies of the United States government that Thoreau thought were seriously unjust—support of slavery and an aggressive war against Mexico.

Actions such as Thoreau's are sometimes described as "conscientious refusal," refusing to obey a law that requires one to act immorally. While conscientious refusal is not identical with publicly protesting a policy, the two usually go together. Generally, people who refuse to obey unjust laws hope that their act will stimulate others to see that the law is wrong and to work for change. Thoreau spoke publicly about the reasons for his act, and his lecture became the classic essay "Civil Disobedience."

Gandhi and King went beyond individual conscientious refusal and organized large numbers of people to disobey the law as a means of protest. These illegal acts were intended to publicize serious injustices and to rally support for change. If enough people were to disobey an unjust law, it might be impossible for a government to enforce it.

Acts of civil disobedience cover a spectrum ranging from: (a) conscientious refusal by individuals; to (b) symbolic disobedience that is meant to convey a message about the wrongness of government policy; to (c) large-scale acts of disobedience that aim to render a government unable to carry out its policies.

Not everyone would accept the definition given above. Some argue that civil disobedients must accept the punishment, but this does not seem necessary. For example, someone who publicly burns a draft card might flee the country if the punishment were extremely severe; yet the original act would still be civil disobedience, even if the act of fleeing is not. John Rawls (1999) has argued that civil disobedience addresses a community's sense of justice, but this overlooks the fact that a community can have mistaken or conflicting conceptions of justice. Finally, some argue that civil disobedience can be violent, but this overlooks the connotations of the word "civil" and violates the tradition of Gandhi and King, who were explicitly committed to nonviolent strategies of resistance. Moreover, because violent acts require stronger types of justification, including them in the definition complicates the evaluation of civil disobedience. Violent acts will have to be distinguished from nonviolent ones when people try to see if civil disobedience can be morally justified. In the end, the test of definitions is that they help to clarify matters, and lumping together violent and nonviolent acts in this case does not seem helpful.

Using the definition above, the question "Is civil disobedience ever morally justified?" can be understood to mean "Is it ever morally permissible to engage in nonviolent, public violations of the law in order to protest a governmental law or policy?"

THE DUTY TO OBEY THE LAW

Asking whether civil disobedience can be morally justified presupposes that there is a moral duty to obey the law. If there were no such moral duty, then breaking the law would not need a special justification. In addition, people who think that civil disobedience can never be justified must believe that the moral duty to obey the law is absolute and can never be overridden by other moral concerns.

Socrates' arguments in the *Crito* are often taken as a source of the view that people must always obey the law. Socrates appears to argue that people must always obey the law because the state is like a parent and one must obey one's parents, that the state has benefited him and therefore should be obeyed, and that he has made a tacit agreement to obey the laws by living in Athens all his life. In the *Apology*, however, Socrates states that he will disobey the law if it requires him to violate the commands of the gods. Socrates, then, is a source of both the individualist tradition that approves civil disobedience and the authoritarian, statist tradition that condemns it.

In his *Leviathan*, Thomas Hobbes provides a famous argument for the duty of obedience to law. He argued that recognition of government's authority is justified because it is the only way for people to avoid a state of nature in which everyone is a threat to everyone else. If all people followed their own judgment and recognized no legal authority, this would lead to a situation of unlimited conflict in which life is "nasty, brutish, and short" (*Leviathan*, Ch. XIII). Hobbes thought that peace could be achieved only if people agree to obey a sovereign who enforces the laws. If everyone claims a right to act according to their own judgment and to disregard the law, then government would be undermined, and there would be a return to anarchy and a state of war by "every man, against every man" (*Leviathan*, Ch. XIII). In short, individuals must trade away their personal autonomy if peace and security are to be possible.

In a much discussed argument from the 1960s, Robert Paul Wolff turned Hobbes's argument on its head in order to defend a version of philosophical anarchism. Wolff agrees with Hobbes that governments claim authority over what citizens should do and thus take away personal autonomy. But, Wolff claimed, personal autonomy—deciding what is right and wrong for oneself and acting on those decisions—can never legitimately be traded away (Wolff 1976). Therefore, governmental authority can never be morally legitimate. From Wolff's anarchist perspective, it is obedience to law rather than disobedience that is morally questionable.

There is also a cynical tradition that sees laws as devices for protecting the interests of the rich and powerful. Thrasymachus, a character in Plato's *Republic*, defined justice as whatever is in the interests of the stronger. This idea is echoed in the Marxist view that the legal system is a device whose real purpose is to protect the property and power of the wealthy. This cynical perspective suggests that it is foolish to believe in a moral obligation to obey the law.

JUSTIFYING CIVIL DISOBEDIENCE

Debates about civil disobedience are often conducted in all-or-nothing terms. They presuppose either (a) that there is an absolute obligation to obey the law no matter what, or (b) that there is no obligation to obey the law at all. From this perspective, support for civil disobedience

leads to anarchism, whereas opposition to it requires mindless conformity to governmental authority.

A different tradition emerges from John Locke's *Second Treatise on Civil Government*. While Locke argued that governments and laws could be legitimate and should be taken seriously, he also defended a right of revolution in cases where the government violates the rights that it is supposed to defend. According to Locke, the duty to obey is conditional on the nature of the government. There is no duty to obey a tyrannical government. This Lockean view acknowledges a general moral duty to obey the law while recognizing that there are circumstances in which disobedience—and even revolution—might be justified. Locke's view is echoed in the American *Declaration of Independence*, which affirms a right to "alter or abolish" a government that violates its people's rights.

Defenders of civil disobedience, then, need not be anarchists. They can recognize the moral force of the law while at the same time believe that the moral force of the law is conditional. When the right conditions do not exist, various forms of disobedience—including civil disobedience—may be justified. If the conditions that warrant obedience to law do exist, then people who violate the law are acting wrongly. Just as obedience to law can be morally required in some cases and morally forbidden in others, so likewise civil disobedience can be justified in some cases and not in others.

The argument for civil disobedience is strongest when a specific law requires people to act immorally. A broader justification for disobedience arises when a government lacks legitimacy. Gandhi's campaign for Indian independence, for example, challenged the legitimacy of British colonial rule. If British rule was illegitimate, then there was no moral duty to obey British laws. Still, for both moral and tactical reasons, Gandhi used civil disobedience selectively.

KING'S DEFENSE OF SELECTIVE OBEDIENCE

While there are plausible justifications for disobedience to some laws and some governments, a serious problem faces people who engage in civil disobedience but nonetheless appeal to others to obey the law. Martin Luther King Jr.'s classic "Letter from a Birmingham Jail" discusses just this problem. Critics charged that King was inconsistent because he urged segregationists to obey laws that enforced racial equality at the same time that King and his followers stated their willingness to violate other laws. If selective obedience was permissible for King, why was it not permissible for his opponents?

King defended himself by providing criteria for justified disobedience. He argued that it is morally permissible to disobey the law: (a) when the law itself is unjust because it "degrades human personality" rather than respecting people; (b) when the laws are binding on a minority group but do not bind the majority that imposes it; (c) when those who are mistreated are deprived of rights of democratic participation in the process of enacting the law; or (d) when a proper law is unjustly applied so as to deprive people of their rights of protest. These conditions, he argued, were met by those campaigning for racial equality but not by those who supported segregation.

King's argument shows how one can consistently defend the right to disobey the law and also take obedience to law seriously. He recognizes a strong presumption in favor of obedience but argues that the presumption is overridden in the kinds of circumstances he describes.

UNJUSTIFIED CIVIL DISOBEDIENCE

Acts of civil disobedience are not as difficult to justify as forms of protest that use violence. Nonetheless, acts of civil disobedience can be morally wrong. For example, they can be committed on behalf of an unjust cause. Thoreau, Gandhi, and King all protested serious evils, but if a person mistakenly believes that a law or policy is unjust, then an act of disobedience against it will not be morally justified. Moreover, even if a law or policy is bad, its defects may not be serious enough to justify violating the law. If obedience to law is something people expect of others when they disagree with a law, then those same people are not justified if they disobey laws simply because they disagree with them. Disobedience must be reserved for serious cases, and even then, it may not be justified if legal means are available for effectively promoting change. It is only when effective, legal means are unavailable that civil disobedience is permissible. Finally, such acts can be wrong if they undermine just and valuable institutions.

A strong case, then, can be made for the view that civil disobedience can be morally justified under certain conditions. Whether specific acts of civil disobedience are justified, however, is often controversial. This is because people often disagree about the seriousness of the evils being opposed, the availability of other effective means of protest, and the long-term effects on valuable institutions and practices. People who agree that civil disobedience can be justified in theory can still disagree about whether it is justified in practice.

See also Hobbes, Thomas; King, Martin Luther; Locke, John; Thoreau, Henry David.

Bibliography

Bedau, Hugo Adam, ed. *Civil Disobedience in Focus*. London: Routledge, 1991.

Gandhi, M. K. *Non-violent Resistance*. New York: Schocken Books, 1961.

Hobbes, Thomas. *Leviathan*. Many editions.

King, Martin Luther, Jr. "Letter from a Birmingham Jail." In *Why We Can't Wait*. New York: Signet Books, 1964.

Locke, John. *Second Treatise on Civil Government*. Many editions.

Nathanson, Stephen. *Should We Consent to Be Governed?*. 2nd ed. Belmont, CA: Wadsworth, 2001.

Plato, *Apology*, *Crito*, and the *Republic*. Many editions.

Rawls, John. "The Justification of Civil Disobedience." In *John Rawls: Collected Papers*, edited by Samuel Freedman. Cambridge, MA: Harvard University Press, 1999.

Thoreau, Henry David. "Civil Disobedience" and "A Plea for Captain John Brown." In *Walden and Other Writings of Henry David Thoreau*, edited by Brooks Atkinson. New York: Modern Library, 1950.

Wolff, Robert Paul. *In Defense of Anarchism*. 2nd ed. New York: Harper and Row, 1976.

Zinn, Howard. *Disobedience and Democracy: Nine Fallacies on Law and Order*. New York: Vintage Books, 1968.

Stephen Nathanson (2005)

CIXOUS, HÉLÈNE

(1937–)

Hélène Cixous was born in Oran, Algeria, on June 5, 1937. Her father was of French-colonial and Jewish descent and her mother was Austro-German. Cixous grew up in Algeria, although she studied in France and began her academic career there. Her first text, *Le prénom de Dieu* (God's first name), was published in 1967. Since 1968, she has been a professor of English literature at Université de Paris VIII–Vincennes, a university considered "revolutionary" for its opposition to traditional institutional structures, which she helped found. Cixous also established the first women's studies center in Europe at Université de Paris VIII.

Cixous has been consistently concerned with the repressive and exclusionary consequences of institutional and systemic forms of power. She has been interested in both individual and collective liberation struggles, such as the liberation of the self from the impact of psychoanalysis, the liberation of women, and Third World struggles. She has published approximately fifty novels, plays, and theoretical essays. Within the United States, the best known of her writings have been "The Laugh of the Medusa" (1976) and *The Newly Born Woman* (1986). Much of her work has been originally published in French and has not been translated into English.

Cixous is well known for her notion of *écriture feminine*. In "The Laugh of the Medusa," Cixous maintains that to define a feminine practice of writing, or écriture feminine, is not possible since "it will always surpass the discourse that regulates the phallocentric system" that aims to theorize or enclose it (1976, p. 883). Cixous discusses her wariness of reductive language that would simplify or capture her practice of écriture feminine. Nonetheless, her basic attempt is to free language and to offer new ways of writing and speaking. To do so, she emphasizes the fictional and poetic elements in her writing. In questioning structures of power, Cixous advocates the freeing of self through writing. In turn, freeing the self (or the subject) means rethinking traditionally repressed categories; for example, woman, the body, and writing. Cixous argues against the association of the phallic subject with narcissism and death, which simultaneously equates women with death. In contrast to an emphasis on narcissism and death, Cixous suggests an economy of the gift—an economy that is based on giving and receiving. The exchange represented by an economy of the gift would mark a new mode of exchange, for Cixous, and would arise through linguistic changes. In turn, in Cixous's view, it is only through linguistic changes that social changes are possible. Thus, Cixous encourages women to "write themselves"; that is, women should write their bodies and their desires, which have always and only been written and discussed by men.

The transformation of the relationship between self and other is central to Cixous's writing and constitutes its political dimension. While Cixous wrote her dissertation on Irish author James Joyce, her emphasis on life over death separated her from him. Although Cixous recognized Joyce for his emphasis on transforming linguistic structures as a means of changing mental structures, Joyce ultimately maintained that one must lose (kill the other) in order to have (live). Despite Cixous's recognition of loss and death as inevitable for life, her aim is to emphasize life over death (thereby reversing the emphasis of many male authors). One way in which Cixous highlights life, and the economy of the gift, is through a focus on the mother and child relationship; specifically, the mother and daughter relationship. Cixous suggests that the woman/mother gives insofar as she nourishes the child. Woman is both the container and the contained. Woman's relationship to the Other, or to otherness, thus

differs from the relationship between man and the Other since things happen to him from the outside. Cixous uses the metaphor of "white ink," or of writing in breast milk, to convey the idea of reuniting with the maternal body. She also argues for a bisexuality that would extend subjectivity beyond dualisms to configure a multiple, rather than a fixed and static, subject.

In addition to Joyce, Cixous's work has been informed by several German and French philosophers, including Martin Heidegger and Jacques Derrida. In *La Venue à l'écriture* (1977), a strongly Derridean work, Cixous advances the position that écriture feminine is not necessarily writing by a woman; instead, it is writing likewise practiced by certain male authors (such as Joyce and Jean Genet). Cixous has furthered the work of psychoanalyst Jacques Lacan, though amidst controversy, by pointing out that women and men enter into the symbolic order (the structure of language) differently. She critiques Lacan's naming of the phallus as the center of the symbolic and suggests that this view marks language as "phallocentric" (the idea that the structure of language is centered by the phallus). In this regard, she both echoes and presses Derrida's insight that the Western privileging of spoken words over written words renders the structure of language as "logocentric." Like Derrida, she interrogates the binary structure of language in the West and exposes its role in maintaining oppressive structures of thought.

Often Cixous is placed alongside Luce Irigaray, Julia Kristeva, and Catherine Clément as being one of the French or continental feminists. However, the use of the phrase, "the French feminists," is problematic here in that it tends to conceal from consideration other feminists who are French. Moreover, the phrase overlooks the more complicated backgrounds of the so-called "French" feminists themselves. Not unlike these other thinkers however, and most notably Irigaray, Cixous has been charged with essentialism. That is, she has been criticized for engaging with an essential, identifiable, and named femininity within the texts she examines. Cixous's response to such accusations, not unlike Irigaray's, would be to claim that she does not intend to engage with a biological category "woman"; rather, she aims to interrogate the cultural position held by such categories within discourse and systems of language.

See also Feminism and Continental Philosophy.

Bibliography

WORKS BY HÉLÈNE CIXOUS

"The Laugh of the Medusa." Translated by Keith Cohen and Paula Cohen. *Signs* 1 (Summer 1976): 875–893.

The Newly Born Woman. With Catherine Clément. Edited by Betsy Wing. Minneapolis: University of Minnesota Press, 1986.

"Coming to Writing" and Other Essays. Edited by Deborah Jenson. Translated by Sarah Cornell, Deborah Jensen, Ann Liddle, and Susan Sellers. Cambridge, MA: Harvard University Press, 1992.

Hélène Cixous, Rootprints: Memory and Life Writing. With Mireille Calle-Gruber. New York: Routledge, 1997.

WORKS ABOUT HÉLÈNE CIXOUS

Conley, Verena Andermatt. *Hélène Cixous: Writing the Feminine*. Lincoln: University of Nebraska Press, 1991.

Sellers, Susan, ed. *The Hélène Cixous Reader*. New York: Routledge, 1994.

Wilcox, Helen, ed. *The Body and the Text: Hélène Cixous Reading and Teaching*. New York: Palgrave Macmillan, 1991.

Mary K. Bloodsworth-Lugo (2005)

CLANDESTINE PHILOSOPHICAL LITERATURE IN FRANCE

The body of clandestine literature in France that deals with philosophy, religion, ethics, and social problems is impressive. It can be traced back to the sixteenth century, and the diffusion, particularly wide between 1714 and 1740, of the allegedly atheistic treatise *La béatitude des Chrétiens ou le fléau de la foy,* published by Geoffroy Vallée in 1572, and of other tracts of early date bears witness to the continuity and vitality of the tradition of free thought in France. The term "Clandestine philosophical literature" usually refers to works known to have circulated in manuscript form during the first half of the eighteenth century and the importance of the subject lies in the fact that the circulation of these works provided one of the sources of the French encyclopedic movement and a solid foundation for liberalism. For the period between 1700 and 1750, I. O. Wade has listed 392 extant manuscripts of 102 different treatises, including 15 translations from other languages. Many more are known to have been in circulation.

The technique of the clandestine manuscript essay was used to circumvent the severe censorship and was most common between 1710 and 1740, when the activities of copyists, colporteurs, and the police were particularly vigorous. Works that found their way into print were often impounded, but they were copied and distributed until the French Revolution. Occasionally authors whose identities could be established were incarcerated. This happened to de Bonaventure de Fourcroy in 1698 for his

Doutes sur la religion proposées à Mss. les Docteurs de Sorbonne, but he soon secured his release from the Bastille. Most often the police found it futile to make arrests and concentrated on preventing the diffusion of the tracts. Public burning, usually in effigy, of works condemned by the Parlement of Paris did not prevent reprints and manuscript copies from being made in the Low Countries, one of the centers of the clandestine trade. After 1750, however, covert circulation became increasingly unnecessary, owing to the breakdown of the censorship, and a number of the more important treatises were printed, many with the indication of a false place of publication.

Voltaire, Henri-Joseph Dulaurens, Baron d'Holbach, and Jacques-André Naigeon, in their desire to foster deism or atheism, prolonged the life of the anonymous tracts by including them in collective volumes, such as *Nouvelles Libertés de penser* (Amsterdam, 1743, 1770), *Recueil nécessaire* (by Voltaire; Geneva, 1765, 1766, 1768, 1776), *L'évangile de la raison* (by Voltaire; Geneva, 1764, 1765, 1767, 1768), *Recueil philosophique* (by Naigeon; "Londres," 1770), and *Bibliothèque du bon sens portatif* (by Holbach; "Londres," 1773). The treatises constituted one of the main sources from which the *philosophes* drew their polemics.

Through the records preserved in the Archives de la Bastille and from statements appearing in manuscripts and letters by Dubuisson, Nicolas Fréret, G. de Bure, and Charles-Marie de la Condamine, we know something of the organization and diffusion of these manuscripts. Le Coulteux, Charles Bonnet, Lépiné, and a certain Mathieu or Morléon (who was incarcerated in 1729) are known to have specialized in the works of Henri de Boulainvilliers and his friends. These works were distributed often in the vicinity of the Procope and other cafés to listed patrons and initiates, including members of the clergy and the Parlement. Copies such as those of Jean Meslier's *Testament* were made by professionals, occasionally the personal secretaries of men like the Comte de Boulainvilliers, the Comte d'Argenson, and Chrétien-Guillaume de Lamoignon de Malesherbes, and the practice of employing copyists was continued throughout the century. The price of such copies varied greatly. A sum as prohibitively high as twenty pistoles is known to have been asked for Jean-Baptiste de Mirabaud's *Examen critique du Nouveau Testament.*

The clandestine movement, fed by new discoveries in science, reflected the climate of world opinion, an attitude to life and society, man and his welfare, God and the universe which, although not new, was reinforced by new arguments and gained an ever-increasing audience. Although the tracts appeared sporadically and were mostly anonymous, they share a few common characteristics and must be judged as a stage in the history of free thought, which goes back to the Renaissance in France and has its deepest roots in the works of Epicurus and Lucretius.

THEOPHRASTUS REDIVIVUS

The *Theophrastus Redivivus* (1659) is significant in that it establishes a link between the atheism of men of the Renaissance and that of men of the seventeenth century (it refers, for example, to Lucilio Vanini and Cyrano de Bergerac) and also of the eighteenth century, when it was secretly circulated. The author, possibly a regent in one of the Parisian colleges, wrote in Latin a 2,000-folio-page compendium of historical references. He developed the arguments that if God exists he is the Sun and that the world is eternal. For the author all religions are false, and miracles, oracles, prophecies, and revelations are man-made. The resurrection of the dead and the immortality of the soul are absurdities; happiness is to be found only in living according to nature, which is revealed to us through experience; there is no absolute good or evil, as we may deduce from the multiplicity of customs and laws; man is a species of animal endowed with speech and reason. Animals, however, are not totally devoid of these faculties. The author referred neither to Pierre Gassendi nor to René Descartes, but he did mention the treatise *De Tribus Impostoribus,* attributing to Frederick II the proposition that Moses, Christ, and Muḥammad were three remarkable impostors.

BACKGROUND

Throughout the seventeenth century the *libertins* in the wake of François Rabelais and Michel Eyquem de Montaigne became erudite skeptics, radical naturalists, associating freedom of morals and freedom of belief. As freethinkers they were prompted more by a feeling of revolt against asceticism and scholasticism than by any convincing argument. Gassendi contributed to the rehabilitation of Epicurus and Lucretius. Emmanuel Maignan, too, in his *Cursus,* evolved a philosophy that bridged Aristotle and Epicurus, linking matter and thought, sensationism and the spiritual world, and developing the idea of a scale of being. But it was from Descartes that the movement of free thought gained its greatest impetus. Cartesian rationalism and mechanism provided freethinkers with a new certainty and their systems with a new coherence. Long after his philosophy had been adopted by the Jesuits and had consequently grown

unpopular, Descartes continued to exercise a determining influence on free thought through the method he advocated. His philosophy, however, was commonly misunderstood by freethinkers and with Julien Offray de La Mettrie it culminated in an extreme mechanistic materialism that Descartes would have decried.

Benedict de Spinoza's influence on the clandestine literature was considerable but rather indirect. His work was largely known through the writings of other thinkers, like Pierre Bayle and Boulainvilliers, and his philosophy was commonly distorted by Cartesian misrepresentation. The *Ethics* was little known, and frequently its views were reconstituted through refutations. The *Tractatus Theologico-Politicus* was of interest on account of its biblical criticism, and in Holland, Jean Le Clerc, professor of philosophy and Hebrew at the University of Amsterdam, was allowed to carry on this critical work. In France, however, the uncompromising attitude of Jacques Bénigne Bossuet stifled biblical criticism. Richard Simon, a well-known teacher at the Oratorian school at Juilly who had admitted in his *Histoire critique du Vieux Testament* (1678) the truth of much of Spinoza's exegesis while recognizing the authority of the Bible, succeeded in offending both Catholics and Protestants and was expelled from the Oratorian congregation in 1678. He retired to continue his rational critique in two *instructions pastorales* (1702, 1703), *Histoire critique du texte du Nouveau Testament* (1683), *Histoire critique des versions du Nouveau Testament* (1690), and *Histoire critique des principaux commentateurs du Nouveau Testament* (1692).

Disputes that reached the general public—such as those over the authorship of the Pentateuch, in which Isaac La Peyrère, Thomas Hobbes, Spinoza, Simon, Le Clerc, and others held different views—led to much perplexity. The body of anonymous treatises that continued such discussions and in many cases rejected revelation is naturally large. These include the *Examen de la religion,* the *Analyse de la religion* (written after 1739), and the *Militaire philosophe* (composed between 1706 and 1711).

Hobbes's *Leviathan* (1651) seems to have been little known in France. Bayle's *Dictionnaire,* however, enjoyed great authority, and his *Lettre sur la comète de 1680* popularized the ideas that the conception of Providence did not rest on rational premises and that atheists could be good men. Bayle's views were those of a protestant, but his argument was such that his articles could easily be used to develop anti-Christian ideas. The anonymous writers also read Bernard Le Bovier de Fontenelle and knew something of the English deists whose thought developed along parallel lines. There were translations of works by Bernard Mandeville, Lord Bolingbroke, John Toland, Anthony Collins, and Thomas Woolston, but it was only after the publication of Voltaire's *Lettres anglaises* (1734), which discussed Newtonian physics and philosophy and the ideas of John Locke, that the English influence became significant. Gottfried Wilhelm Leibniz's influence, too, was felt only at a late stage, partly because he was known primarily through Bayle and also through Pierre-Louis Moreau de Maupertuis, whose ideas served to link the *Monadology* with Denis Diderot and materialism.

THE COTERIE OF BOULAINVILLIERS

The only group of writers known to have been involved in concerted action was that centered in the Comte de Boulainvilliers and closely linked with d'Argenson, the duc de Noailles, and the Académie des Inscriptions. This coterie included Nicolas Fréret, Mirabaud, César Dumarsais, and J.-B. Le Mascrier. Voltaire, in his *Dîner du comte de Boulainvilliers* (1767), attested to the important influence of this group, which was especially responsible for the diffusion of Boulainvilliers's *Esprit de Spinoza* (known to have existed by 1706 and first published in 1719 in Holland).

FRÉRET. Nicolas Fréret (1688–1761), a student of law, joined the coterie of Boulainvilliers at the age of nineteen. Fréret appended to copies of the *Histoire ancienne* an account of Boulainvilliers's life and works. In 1714 he was admitted to the Académie des Inscriptions; in 1715 he was imprisoned for some months in the Bastille, where he read Bayle's *Dictionnaire* and wrote a Chinese grammar. From 1720 to 1721 he was preceptor of the duc de Noailles.

The *Lettre de Thrasibule à Leucippe* (written c. 1722 and published in London, probably in 1768; also published in *Oeuvres de Fréret,* Vol. IV, London, 1775) is generally attributed to him. Systematic and Cartesian in its presentation, this treatise combines sensationist psychology and naturalist ethics. Thrasibule, a Roman, describes the early Christians as combining Jewish beliefs with Stoicism and as influenced by both monotheist and polytheist currents. He argues that knowledge is acquired through our senses and has only relative validity. Only the truths of mathematics and reason are universal. Religious beliefs however, do not spring from reason; it is reason alone that should guide man in regulating his life, establishing society and laws, and achieving happiness. This work can be seen as an early essay in comparative religion, and it sharply reflects the growing interest in the sci-

ence of law and social philosophy. It perhaps influenced Baron de Montesquieu, and Jean-Jacques Rousseau annotated it while engaged in writing the *Discours sur l'inégalité.*

Fréret is also reputed to be the author of an *Examen critique des apologistes de la religion chrétienne* (composed after 1733), which introduces the historical method adopted by Voltaire in, for example, the *Essai sur les moeurs* and the *Dictionnaire philosophique,* in which Voltaire acknowledged his debt. Fréret was held in high esteem as a savant. He was a chronologist, a geographer, an orientalist, and a philologist as well as a philosopher, and he delivered papers on a wide variety of subjects to the Académie des Inscriptions, becoming its permanent secretary in 1743. These *Mémoires de l'Académie* outline new methods for the study of prehistory and geography as well as history. Fréret specialized in mythology, opposing the *évhéméristes,* who believed that all myths had a basis in historical fact. A pioneer in comparative philology, he made known the Chinese linguistic system. His *Oeuvres complètes* were published by Leclerc de Septchênes in Paris, 1796–1799, but about half his works were omitted (many of his manuscripts bequeathed to the Académie des Inscriptions have never been published), and a few of the treatises included cannot be attributed to him.

MIRABAUD. Jean-Baptiste de Mirabaud (1675–1760) was educated by the Oratorian congregation and entered a military career. He then became secretary to the duchess of Orléans and preceptor of her two youngest daughters. In 1724 he translated *Gerusalemme liberata* by Torquato Tasso. He was elected to the Académie Française in 1726, becoming its secretary in 1742. Mirabaud read his manuscripts to select groups of friends. He was probably the author of four essays (described below), often to be found together, that threw new doubts on biblical chronology and promoted Fontenelle's method of oblique attack on miracles. Many of Mirabaud's notes recall ideas expressed in *La religion chrétienne analysée* (a popular post-1742 tract attributed by Voltaire and Claude François Nonnotte to Dumarsais). The *Opinion des anciens sur le monde* (c. 1706–1722) challenges the story of Genesis. In the *Opinions des anciens sur la nature de l'âme* (composed before 1728, published in *Nouvelles Libertés de penser*) Mirabaud pointed out that the Jews, the Greeks, and the Romans envisaged the soul as material and that the Egyptians introduced the belief in the immortality of the soul as a restraining influence on public morals. The *Opinion des anciens sur les Juifs* (c. 1706–1722), based on Jacques Basnage's *Histoire des Juifs* (1706), tries to prove that the Jews had no right to claim to be a "chosen" people. The *Examen critique du Nouveau Testament* (c. 1706–1722), which deals with the canonical and the noncanonical gospels, stresses that neither Philo nor Josephus mentioned Christ and that Christian morality conflicts with natural morality. Much of our information on Mirabaud is derived from the *Notice sur Jean-Baptiste de Mirabaud* (Paris, 1895), by Paul Mirabaud.

DUMARSAIS. César Chesneau Dumarsais (1676–1756), a grammarian, was personally known to Fontenelle and Voltaire and was associated with the *Encyclopédie* until his death. For a time he was preceptor in the family of John Law. Dumarsais edited, with Le Mascrier, some of the deistic works of Mirabaud and wrote a defense of Fontenelle's *Histoire des oracles* and probably the deterministic essay *Le philosophe* (written before 1728); edited by Herbert Dieckmann in 1948). He was probably responsible for *La religion chrétienne analysée* (also known as *Examen de la religion* and *Doutes,* in which inconsistencies in the Bible are shown up, the doctrine of original sin is attacked, and the doctrine of the Trinity is stated to be contrary to reason. It is argued that God should be worshiped without ceremony and that man must follow his reason, which is his *lumière naturelle,* and adopt a social morality incompatible with Christian dogma.

MESLIER. The most interesting of the clandestine authors was no doubt Jean Meslier (1664–1729), a priest who was directly or indirectly influenced by Spinoza. (Reading François de Salignad de La Mothe Fénelon's *Démonstration de l'existence de Dieu* and R.-J. de Tournemine's *Réflexions sur l'athéisme* helped Meslier clarify his ideas.) He identified nature with matter, which he saw as eternal and as endowed with movement. He favored a mechanical interpretation of nature, rejecting the arguments of those who believed in chance and in a divine design. In his 1,200-page *Testament,* Meslier listed the errors, illusions, and impostures of Christianity. His attack on Christianity is one of the most detailed and comprehensive ever written, and his materialistic system is particularly interesting in that it foreshadows many aspects of Diderot's thought.

Voltaire is known to have acquired a copy of the *Testament* and to have made extracts, which he dated 1742 and published in 1761 or 1762 under the title *Extrait.* The first edition sold out immediately and was followed in the same year by an edition of 5,000 copies. In 1772 Holbach published extracts under the title *Le bon sens du curé*

Meslier, and in 1789 Sylvain Maréchal published *Le Cathéchisme du curé Meslier.*

Meslier's social ideas were remarkable for the time. He claimed in very general terms that all men are equal and have the right to live, to be free, and to share in the fruits of the earth. He divided humankind into workers and parasites and saw in revolt the best hope of better conditions. He dreamed of a class struggle, not reconciliation.

OTHER WORKS

Among other anonymous works that cast doubts on the proofs of the truth of Christianity and allege contradictions in the Bible are five manuscript volumes of the *Examen de la Genèse* and the *Examen du Nouveau Testament* (probably written in the late 1730s or early 1740s), which are attributed to Mme. du Châtelet, Voltaire's mistress. She purports to have proved that the stories of the Bible relate barbarous and cruel events and cannot have been inspired by God. No doubt she received some help from Voltaire, but she relied chiefly on the work of Meslier and Woolston and especially on the *Commentaire littéral sur tous les livres de l'Ancien et du Nouveau Testament* (23 vols., Paris, 1707–1716) by Augustin Dom Calmet.

Among other manuscripts whose authorship is now known is *Le ciel ouvert à tous les hommes* (also titled *Le paradis ouvert* and *Nouveau Système de la religion chrétienne*), by the priest Pierre Cuppé, which must have been in draft in 1716. The tract never assails orthodoxy, but Cuppé submitted the Scriptures to scrutiny and preached toleration and brotherly love, concluding that all men are saved by God's love. Cuppé's stress on his respect for reason, as well as his deistic beliefs, led to his being considered a forerunner of French deism.

The author of *Le militaire philosophe* (1706–1711; published in London by Naigeon in 1768) is unknown. It is first a commentary of Nicolas Malebranche's views on religion. It gives a frank exposition of deism, which won Voltaire's commendation. After a strongly worded criticism of the Old and the New Testaments, the work rejects Christianity and develops the doctrine of natural religion, stressing the roles of reason and instinct. Man, who is both body and soul, is free and immortal, and his behavior should be governed by reason and by conscience. Man must worship God and abide by the golden rule. The author foreshadowed Montesquieu in his insistence on the absolute character of justice and the relative nature of civil laws and in his treatment of chance, which he rejected as an explanation of events. He anticipated Voltaire in his use of the figure of a watchmaker to explain the function of God. His idea that truth is to be found in the individual soul was later developed by Rousseau.

A widely disseminated treatise was *Israël vengé,* by Isaac Orobio, a Spanish Jew who escaped from the Inquisition to France and then to Holland and died in 1687 or 1688. His originally unpublished critical attack on the Christian religion was translated by A. Henriquez and published in London in 1770. It was circulated by Jean Lévesque de Burigny.

The *Jordanus Brunus Redivivus* is a materialistic compilation. The author believed in the Copernican system (and the existence of other solar systems with living beings) and in the eternity of matter. There are no innate ideas, no objective good or evil. Man is motivated by pain and by pleasure. Experience can deceive us. Reason alone is valid but must not be thought infallible. The laws of nature are eternal, but everything is in a state of flux. Certain passages of this work bring to mind Diderot's *Rêve de d'Alembert.* Other manuscripts whose authorship is uncertain include *Lettre d'Hypocrate à Damagette* (1700 at latest), *Recherches curieuses de philosophie* (1713), *Suite des Purrhoniens: qu'on peut douter si les religions viennent immédiatement de Dieu ou de l'invention des politiques pour faire craindre et garder les préceptes de l'homme* (c. 1723), *Traité de la liberté* (a determinist and materialist tract, probably by Fontenelle, c. 1700), *Essai sur la recherche de la vérité,* and *Dissertation sur la formation du monde* (1738), which was inspired by Lucretius and formulates transformist theories while upholding the conception of fixed species.

INFLUENCE

It will be seen that the clandestine tracts fall into two main categories, those written from the standpoint of critical deism and those that are atheistic, deterministic and materialistic. The outstanding eighteenth-century literary works based on this movement can be similarly characterized. Montesquieu's adoption of the letter form for *Les lettres persanes* (1731) may owe something to the *Lettre à Damagette,* and the views expressed in *Lettre persane 46* reflect those expressed in *La religion chrétienne analysée.* Voltaire, who adopted the form for his *Lettres philosophiques,* published anonymously in 1734, wrote in the same year a *Traité de métaphysique* (which Mme. du Châtelet kept under lock and key), which embodied his own deism as well as many of the ideas expressed in the clandestine literature.

Toward the middle of the century atheism gained ground, no doubt encouraged by such treatises as *Lettre*

sur la religion, sur l'âme et sur l'existence de Dieu. Diderot's *Pensées philosophiques,* published anonymously in 1746, allegedly at the Hague but actually in Paris, and condemned to be burned by the Parlement of Paris, is characteristic of this tendency. Although based on a translation of the Earl of Shaftesbury, the work succeeds in presenting an original and vividly expressed atheism side by side with more commonplace arguments in favor of natural religion. In particular it challenges Christian belief in miracles, outlining the principles of the new biblical criticism. In the eighteenth century alone the *Pensées philosophiques* ran to twenty editions (some with crude interpolations) and reprints. It was translated into German, Italian, and English and was the subject of long and heated controversy. Twelve signed or anonymous refutations by Protestants, Catholics, parliamentarians, and others were published, some of them, together with Diderot's text, being circulated in manuscript form.

As government policy wavered and censorship grew slack, an increasing number of the manuscripts of earlier date were published, and anonymity became a thin veil, if not a mere convention. The main current of what has become known as clandestine literature, which many have identified with the tradition of free thought, came to an end with the advent of Montesquieu, Voltaire, Rousseau, and Diderot. In their works it found its finest literary expression, and thanks to them it became integrated into coherent patterns that have won it a place in the history of ideas.

See also Boulainvilliers, Henri, Comte de; Meslier, Jean.

Bibliography

The most important reference works in the field are I. O. Wade, *Philosophical Ideas in France from 1700 to 1750* (Princeton, NJ, 1938), and J. S. Spink, *French Free-Thought from Gassendi to Voltaire* (London: University of London, Athlone Press, 1960). The following works are also very important: W. H. Barber, *Leibniz in France* (Oxford: Clarendon Press, 1955); A. Vartanian, *Diderot and Descartes* (Princeton, NJ: Princeton University Press, 1953); and P. Vernière, *Spinoza et la pensée française avant la Révolution* (Paris, 1954).

The reader should also consult E. R. Briggs, "L'Incrédulité et la pensée anglaise en France au début de XVIII[e] siècle," in *Revue d'histoire littéraire de la France* 41 (1934): 497–538; E. R. Briggs, "Pierre Cuppé's Debts to England and Holland," in *Studies on Voltaire and the XVIII Century,* edited by Theodore Besterman, Vol. VI (Geneva: Institut et Musée Voltaire, 1958), pp. 37–66; Pierre Brochon, *Le livre de colportage en France depuis le XVI[e] siècle, sa littérature, ses lecteurs* (Paris: Librairie Gründ, 1954); Herbert Dieckmann, ed., *Le Philosophe: Texts and Interpretation,* Washington University Studies, New Series, Language and Literature, No. 18 (St. Louis, MO, 1948); Herbert Dieckmann, "The Abbé Jean Meslier and Diderot's *Eleuthéromanes,*" in *Harvard Library Bulletin* 7 (Spring 1953): 231–235; J. P. Free, *Rousseau's Use of the Examen de la religion and of La Lettre de Thrasibule à Leucippe,* unpublished doctoral dissertation (Princeton, NJ, 1935); A. D. Hole, *Mirabaud's Contribution to the Deistic Movement and His Relation to Voltaire,* Princeton doctoral dissertation *Abstracts* (Princeton, NJ, 1952); G. Lanson, "Questions diverses sur l'histoire de l'esprit philosophique en France avant 1750," in *Revue d'histoire littéraire de la France* 19 (1912): 1–29, 293–317; A. R. Morehouse, *Voltaire and Jean Meslier* (New Haven, CT: Yale University Press, 1936); R. R. Palmer, *Catholics and Unbelievers in Eighteenth Century France* (Princeton, NJ: Princeton University Press, 1939); Renée Simon, "Nicolas Fréret, académicien," in *Studies on Voltaire and the XVIII Century,* edited by Theodore Besterman, Vol. XVII (Geneva: Institut et Musée Voltaire, 1961); J. S. Spink, "La Diffusion des idées matérialistes et anti-religieuses au début de XVIII[e] siècle: Le *Theophrastus Redivivus,*" in *Revue d'histoire littéraire de la France* 44 (1937): 248–255. See also Diderot's *Pensées philosophiques,* edited by Robert Niklaus (Geneva: Droz, 1950).

OTHER RECOMMENDED TITLES

Darnton, Robert. *The Corpus of Clandestine Literature in France, 1769–1789.* New York: Norton, 1995.

Sutcliffe, Adam. "Judaism in the Anti-Religious Thought of the Clandestine French Early Enlightenment." *Journal of the History of Ideas* 64 (1) (2003): 97–117.

Robert Niklaus (1967)
Bibliography updated by Tamra Frei (2005)

CLARKE, SAMUEL
(1675–1729)

Samuel Clarke, the most important British philosopher and theologian of his generation, was born in Norwich, England, on October 11, 1675. He took his BA degree at Cambridge in 1695, defending Isaac Newton's views. In 1697 he provided a new annotated Latin translation of Jacques Rohault's *Treatise of Physics*, and in his notes criticized René Descartes's physics in favor of Newton's. In that same year he was introduced into the Newtonian circle, probably by William Whiston (1667–1752), whom he had befriended. In 1704 he delivered his first set of Boyle Lectures, *A Demonstration of the Being and Attributes of God: More Particularly in Answer to Mr. Hobbes, Spinoza, and Their Followers*. They were so successful that he was asked to deliver the 1705 lectures as well under the title *A Discourse concerning the Unchangeable Obligations of Natural Religion and the Truth and Certainty of Christian Revelation*. His connection with Newton became official in 1706, when he translated the *Opticks* into Latin. In the same year Anthony Collins, a materialist follower of John

Locke's, engaged Clarke in a long and famous exchange on whether matter can think.

After becoming one of Queen Anne's (1665–1714) chaplains, Clarke was elevated in 1709 to the rectory of St. James's, Westminster. In 1712 Clarke published *The Scripture Doctrine of the Trinity*, which was accused of Arianism, the view that Christ is divine but created. The ensuing theological controversy culminated two years later in his humiliating promise to the Upper House of Convocation not to preach or write on the trinity any longer. However, suspicions of crypto-Arianism remained. François-Marie Arouet de Voltaire reports that Bishop Edmund Gibson (1669–1748) effectively prevented Clarke's elevation to the see of Canterbury by pointing out that Clarke was indeed the most learned and honest man in the kingdom, but had one defect: He was not a Christian.

After the Hanoverian accession Clarke developed a close relationship with Caroline of Anspach (1683–1737), the Princess of Wales, and through her mediation he engaged Gottfried Wilhelm Leibniz in the most famous philosophical correspondence of the eighteenth century. The exchange dealt with many of the issues that had occupied Clarke in his Boyle Lectures, such as divine immensity and eternity, the relation of God to the world, the soul and its relation to the body, free will, space and time, and the nature of miracles. It also discussed more strictly scientific topics, such as the nature of matter, the existence of atoms and the void, the size of the universe, and the nature of motive force, which were then often given both a philosophical and a scientific treatment. In 1717 Clarke published his translation of the correspondence with Leibniz together with an attack on a work by Collins denying the existence of free will. This was his last significant philosophical work, although in 1728 he wrote a short essay for the *Philosophical Transactions* trying to show, against the Leibnizians, that the proper measure of force is not mv^2 but mv. He died in 1729 after a short illness and was survived by his wife, Katherine, and five of his seven children.

Clarke was a polite and courtly man who, however, was vivacious with his friends and seems to have been fond of playing cards. He was also a classicist of repute, and seems to have held Marcus Tullius Cicero's views in high esteem. Voltaire, who met him, was impressed by his piety and admired his logical skills so much that he called him "a veritable thinking machine." Indeed, his reputation was such that in 1710 George Berkeley sent him the first edition of his *A Treatise concerning the Principles of Human Knowledge* (Clarke declined to comment on it).

THE ATTACK AGAINST NATURALISM AND THE DEFENSE OF NATURAL RELIGION

Clarke's primary philosophical interests lay in theology, metaphysics, and, to a lesser degree, ethics. His philosophical vocabulary and some of his metaphysical ideas were influenced by Descartes, whom he followed in holding that the world contains two types of substance, mind and matter, the combination of which constitutes humans. However, he sided with Nicolas Malebranche and Locke in denying that introspection lets one reach the substance of the soul. Indeed, like Locke and Newton, he held that one just does not know the substance of things. Furthermore, Clarke's overall judgment of Descartes was critical. He shared the view expressed by Henry More, Blaise Pascal, Pierre Bayle, and Leibniz that Descartes's system could be, and had been, used to further irreligion and had naturally developed into Spinozism. In particular, he believed that Descartes's identification of matter with extension, and therefore space, entails making it eternal and infinite. He defended natural religion from naturalism (the view that nature constitutes a self-sufficient system of which humans are but a part) and revealed religion from deism.

Clarke's attack against naturalism revolved around five connected points. First, God is a necessarily existent omnipotent, omniscient, eternal, omnipresent, and supremely benevolent person. Second, nature and its laws are radically contingent. God, endowed with a libertarian will, chose to create the world and to operate in it by a reasonable but uncaused fiat. Third, although space and time are infinite, matter is spatiotemporally finite, and being endowed only with *vis inertiae* it has no power of self-motion. Fourth, God is substantially present in nature (or better, nature is literally in God, since space and time are divine attributes) and constantly exercises his power by applying attractive and repulsive forces to bodies. Except for the law of inertia, which describes the essentially passive nature of matter, strictly speaking, the laws of nature do not describe the behavior of matter, which is just dead mass constantly pushed around, but the modalities of the ordinary operation of the divine power. As for occasionalism, natural laws prescribe the actions of the divine will rather than describe those of bodies.

Fifth, although the soul is extended and interacts with the body, it is necessarily immaterial because matter, being constituted of merely juxtaposed parts, cannot possibly think even by divine intervention; moreover, the soul has been endowed by God with a libertarian will.

The first four points guarantee that nature is not a self-sufficient system, so much so that without direct and constant divine physical intervention planets would fly away from their orbits, atoms would break into their components, and the machinery of the world would literally grind to a halt; the fifth guarantees that the soul is not a part of nature. In the remainder of this entry, it will be seen that these points emerge from a consideration of Clarke's views on God, free will, matter and the laws of nature, space and time, and the soul.

GOD

The proof of the necessary existence and attributes of God occupies most of *A Demonstration of the Being and Attributes of God*. The main lines of Clarke's argument are as follows. Since something exists now, something has always existed because nothing comes from nothing. What has existed from eternity is either an independent being (one having in itself the reason of its existence), or an infinite series of dependent beings. However, such a series cannot be the being that has existed from eternity because by hypothesis it can have no external cause, and no internal cause (no dependent being in it) can cause the whole series. Hence, an independent being exists. As a separate argument, Clarke also reasoned that since space and time cannot be thought of as nonexistent and they are obviously not self-subsistent, the substance on which they depend, God, must exist necessarily as well. Finally, teleological considerations show that God is necessarily endowed with intelligence and wisdom. In addition, God has, though not with metaphysical necessity, all the moral perfections, whose nature is the same in the divine being as in humans.

Clarke rejected the view of God as substantially removed from space and time. Divine eternity involves both necessary existence and infinite duration. Rather traditionally, the former consists in the fact that God contains the reason (but not the cause) of his own existence. The latter, however, cannot be identified with the traditional view that God exists in an unchanging permanent present without any succession since, like Newton, he considered such a position unintelligible. Consequently, Clarke attributed distinct and successive thoughts to God, as he perceived these as preconditions of the will. Hence, God is immutable with respect to his will and his general and particular decrees only in the sense that the divine being does not change his mind. However, as Clarke also made clear in his exchanges with Joseph Butler, God is not in space and time.

Clarke's criticism of the Scholastic view of divine immensity or omnipresence was analogous to that concerning eternity: the claim that the immensity of God is a point, as his eternity is an instant is, he held, unintelligible. However, while for Clarke God's temporal presence is analogous to humans' at least in involving temporal succession, his views about God's spatial presence were somewhat less clear because he did not explicitly state whether he adopted holenmerism (the view that the divine substance is whole in the whole of space and whole in each and every part) or the view that God is dimensionally extended. Nevertheless, there is evidence that he held the latter view. For Clarke vigorously denied Leibniz's charge that extension is incompatible with divine simplicity, because it introduces parts in God, without making any reference to holenmerism, and in addition he did not defend holenmerism from More's famous critique. Finally, Collins mentions him with Thomas Turner (1645–1714) and More as supporters of the dimensional extension of God.

For Clarke, divine eternity and immensity are to be identified with space and time. Usually, he held that space and time are just divine properties. However, in his fourth letter he also told Leibniz that, in addition, they are necessary effects of God's existence and necessary requirements for divine eternity and ubiquity, without supplying any argument to show that these different accounts are equivalent or even compatible. At other times, as in the letter to Daniel Waterland (1683–1740) and in the *Avertissement* to Pierre Des Maizeaux (1673–1745), in the latter of which Newton had more than a hand, he held that they are not, strictly speaking, properties.

As Leibniz and an anonymous correspondent (almost certainly Waterland) readily noted, echoing Bayle's critique of Newton and Malebranche, the identification of divine immensity with space endangers the simplicity of the divine being because space has parts, albeit not separable ones. Clarke's solution was to claim parity between spatial and temporal extendedness: Since the former is compatible with the simplicity of what "stretches" temporally, the latter is compatible with the simplicity of what stretches spatially. In addition, from the fact that the divine consciousness is extended, one should not infer that it is proper to talk about it in terms of spatial parts any more than it is to talk of the spatial parts of an instant of time although, as Newton had noted in the General Scholium to book 3 of *Principia*, an instant is the same everywhere.

FREE WILL

Clarke attached great importance to the issue of free will. He held that the highest form of freedom involves willing as one should, namely, having one's will in step with one's right values. He also believed that freedom of the will, or liberty, entails a libertarian power of self-determination (a point he emphasized against Leibniz's compatibilist views) and that it is a necessary condition both for that higher form of freedom and for religion. Thomas Hobbes's and Benedict (Baruch) de Spinoza's views—which in Clarke's mind Leibniz had de facto adopted—that everything happens deterministically or necessarily destroys liberty. Against them he held that the causal version of the principle of sufficient reason in the cosmological argument shows that the necessary being on which the contingent world depends must have a libertarian will. For the notion of a necessary agent is contradictory, as agency involves the libertarian capacity of suspending action. Moreover, if God operated necessarily, things could not be different from how they are. But the number of planets, their orbits, indeed, the law of gravitation itself could have been different, as any reasonable person (but not Spinoza) could plainly see. Furthermore, the obvious presence of final causes indicates that divine activity follows not necessary but architectonic patterns.

Besides attacking necessitarianism and determinism with arguments drawn from general metaphysical considerations, Clarke criticized the Hobbesian view that volition is caused by one's last evaluative judgment and the Spinozistic position that the two are identical. He was ready to grant that the understanding is fully determined to assent to a proposition perceived as true in the same way in which an open eye is fully determined to see objects. In this sense the assent is necessary. However, he held, the necessity of the last evaluative judgment is totally immaterial to the issue of freedom. In his judgment, his opponents were guilty of basic philosophical errors. On the one hand, if they maintained that the content of the evaluation, the evaluative proposition, is identical with the volition or causes it, they were confusing reasons with causes. As he explained to Collins, the proposition "doing X is better than doing Y" can provide a reason for action but cannot cause anything because it is an abstract entity. On the other hand, if Clarke's opponents maintained that not the evaluative proposition but one's perceiving or believing it is identical with, or a partial cause of, volition, then they were falling foul of a basic causal principle. Against Descartes, Clarke insisted that judging (assenting to what appears true and dissenting from what appears false) is not an action but a passion. But what is passive cannot cause anything active. So, there is no causal link between evaluation and volition. What causes the volition is the principle of action itself, which Clarke identified with the agent, that is, the spiritual substance.

Having shown that God is endowed with liberty, Clarke argued that humans are as well. Not involving qualities such as complete causal independence and self-existence, liberty is a power God can transfer to one. Furthermore, experience assures one that one has been granted liberty, since one's actions seem to one to be free, exactly as they would do on the supposition that one is really a free agent. Of course, he conceded, this does not amount to a strict demonstration; but denying that one has free will is on a par with denying the existence of the external world, a coherent but unreasonable option. The burden of proof, he felt, is not on the supporter of liberty, but on its denier.

MATTER AND THE LAWS OF NATURE

Clarke's views on matter are best seen in connection with his ideas about miracles. Like Joseph Glanville (1636–1680), Thomas Sprat (1635–1713), Robert Boyle, and Locke, he belonged to that group of English intellectuals associated with the Royal Society, who thought that miracles could be used as evidence for the claim that Christianity is the true religion. According to Clarke a miracle is a work effected in an unusual manner (by which he seems to have meant in a way not subsumable under the laws of nature) by God himself or some intelligent agent superior to man for the proof or evidence of some doctrine, or the attestation of the authority of some person. However, he claimed, "modern deists," noticing that nature is regular, have concluded that there are in matter certain absolutely inalterable laws or powers that render the course of nature unchangeable and therefore miracles impossible.

The deistic view, Clarke argued, is completely wrong. Everything that is done in the world is done either immediately by God himself or by sentient beings; matter is not capable of any laws or powers whatsoever, except for the negative power of inertia. Consequently, the apparent effects of the natural powers of matter—the laws of motion, gravitation, or attraction—are but the effects of God's acting on matter continually, either directly or through intelligent creatures. The course of nature, then, is just the divine will operating continuously and uniformly. This mode of operation is perfectly free and as easily altered as preserved at any time. Of course, Clarke admitted, the divine will infallibly follows necessarily cor-

rect judgments, and consequently God always acts on the basis of rules of "uniformity and proportion." However, given that the will, in God as in humans, is not causally determined by the understanding, the rules governing the ordinary power of God, a subset of which are the laws of nature, are freely self-imposed, and not the unavoidable result of the necessarily correct divine understanding. They are a manifestation of God's moral, and therefore free, attributes, not God's metaphysical, and therefore necessary, ones.

Clarke steadfastly maintained that matter has neither an essential nor an accidental power of self-motion. The first claim was common among early modern philosophers and held not only by the occasionalists but also by thinkers of different persuasions like Descartes, Locke, and Boyle. In fact, even Pierre Gassendi, who had upheld the notion of an active matter by claiming that atoms have an internal corporeal principle of action, had fallen short of claiming that they possess it essentially. Clarke's second claim, however, was more controversial. For although mechanists programmatically tried to substitute a nature made of inert particles for the living nature of Renaissance philosophy, the attempt soon ran into great difficulties. Strict mechanism proved inadequate to explain phenomena like exothermic reactions or the spring of the air, which causes a deflated closed balloon in a vacuum tube to expand. Consequently, mechanism was altered to include particles variously endowed with powers of motion, attraction, and repulsion.

Clarke's position on the activity of matter was radical: The various nonmechanical powers of particles are the result of direct divine or spiritual activity. He could not bring himself to accept active matter because he thought it a prelude to atheism. For, as noted earlier, he believed that denying divine continuous, direct intervention in nature in effect amounts to eliminating God, as John Toland had by endowing matter with essential self-motive powers. Clarke's views, however, had serious drawbacks. A God who is actually extended and constantly operates physically on matter looks suspiciously like the soul of the world, as Leibniz charged using Newton's identification in the *Opticks* of space as the *sensorium* of God. Similarly, the placement of gravitational forces within the purview of ordinary divine activity drew from Leibniz the accusation of obscurantism, a throwback to the quaint idea of angels causing the rotation of the spheres.

SPACE AND TIME

According to Clarke the ideas of space and time are the two first and most obvious simple ideas that exist. Like many of the philosophers who investigated the nature of space and time, he tended to produce arguments with regard to space, presumably leaving the reader to infer that parallel arguments could be drawn with respect to time. With Newton, he argued that while matter can be thought of as nonexisting, space exists necessarily because to suppose any part or the whole of space removed is to suppose it removed from and out of itself, namely taken away while it still remains, which is contradictory. Although space is not sensible, it is not nothingness, mere absence of matter, as it has properties such as quantity and dimensions. One might add other properties Clarke attributed to it, such as homogeneity, immutability, continuity, and, probably, impenetrability since bodies do not penetrate space but space penetrates them. For Clarke, space is also not an aggregate of its parts but presumably an essential whole preceding all it parts, a position motivated at least in part by the view that space is a divine property. As for Newton, space is necessarily infinite because limiting it is supposing it is bounded by something that itself takes up space or supposing it is bounded by other space, and both suppositions are contradictory.

Since absolute space has an essential and invariable structure independent of the bodies in it and is not altered by their presence, any possible world must conform to it, as creatures must be in space and God, whose power is limited to the metaphysically possible, cannot alter the essence of things. The same is true of time, which flows equably independently of anything in it. In short, in contrast to God all creatures occupy an absolute position in space and time that one may or may not be able to determine.

The introduction of absolute space, allegedly demanded by Newtonian physics, offered Clarke an immediate philosophical advantage in the fight against Spinoza. For it showed that the Cartesian identification of extension with matter, which had made possible Spinoza's excesses, was wrong—a consequence that was not lost on Bayle and was insisted on by Colin Maclaurin (1698–1746). Of course, the existence of absolute space introduced a new difficulty, that of its relation to God, but Clarke thought he had solved it by claiming that space and time are attributes of God or the result of divine existence.

THE SOUL

In 1706 Henry Dodwell (1641–1711) published a book in which he defended conditional immortality: One's soul is naturally mortal and following the death of the body can be kept in existence only by divine supernatural intervention. Clarke wrote an open letter to Dodwell complaining that he had opened the floodgates to libertinism by providing an excuse for the wicked not to fear eternal punishment. He then argued that the soul, being immaterial, is naturally immortal and gave his own version of the traditional argument for the immateriality of the soul from the alleged unity of consciousness, insisting that not even God could make matter conscious. Clarke's argument failed to convince Collins, who made no bones about his materialist leanings and intervened in defense of Dodwell. Clarke told Collins that if thinking in humans were a mode of matter, it would be natural to conceive that it may be the same in other beings. Then, Clarke continued, every thinking being, including God, would be ruled by the same absolute necessity governing the motion of a clock. The result would be the destruction of every possibility of self-determination and the undermining of the very foundations of religion.

Clarke's argument for the immateriality of the soul revolved around three basic claims. First, necessarily consciousness is an individual power, that is, each consciousness is one undivided entity, not a multitude of distinct consciousnesses added together. Second, an individual power cannot result from, or inhere in, a divisible substance; or, alternatively, an individual power can only be produced by, or inhere in, an individual being. Third, matter is not, and cannot possibly be, an individual being. The conclusion is that consciousness cannot possibly be the product of, or inhere in, matter.

For Clarke, although the soul is necessarily immaterial, it can causally affect the body because material qualities such as figure and mobility are deficiencies or imperfections that can be brought about by consciousness, which is a positive quality; moreover, one experiences the causal power by which one moves one's body. However, his position on whether the body causally affects the soul was less than clear. At times he leaned toward the view that it does, and at other times that it does not.

According to Clarke the soul is in space and is extended. As he eventually told Leibniz, the soul is in a particular place, the *sensorium*, which a part of the brain occupies. Clarke inferred the presence of the soul in the *sensorium* through an argument employing two independent premises: first, that something can act only where it is substantially, and second, that the soul interacts with the body. The conclusion is that the soul is substantially present where (at least) a part of the body is.

Saying that the soul must be substantially present where a part of the brain is does not fully determine how the soul is present. It rules out mere Cartesian operational presence, but it fails to determine whether the soul's presence is to be understood in terms of holenmerism or in terms of dimensional extension. However, there is cumulative evidence that for Clarke the soul is merely coextended with a part of the brain. Clarke used an analogy with space, which he took to be both extended and indivisible, to explain how the soul could be extended and indivisible; but holenmerism does not apply to space. He did not address Leibniz's accusation that the extension of the soul destroys its unity by appealing to holenmerism; rather, he defended the claim that the soul "fills" the *sensorium*. In sum, Clarke's views on freedom, with their ties to morality and religion, together with his views on causality, pushed him toward the thesis that the soul is extended.

ETHICS AND REVEALED RELIGION

Although some of his sermons contain interesting analyses of individual Christian virtues, the most sustained exposition of Clarke's ethics is contained in *A Discourse concerning the Unalterable Obligations of Natural Religion and the Truth and Certainty of the Christian Revelation*, his second set of Boyle Lectures. Clarke started by stating that clearly there are different relations among persons and that from these relations there arises a "fitness" or "unfitness" of behavior among persons. So, for example, given the relation of infinite disproportion between humans and God, it is fit that one honors, worships, and imitates the Lord. In other words, from certain eternal and immutable factual relations among persons there arise certain eternal and immutable obligations, which in their broad features can be rationally apprehended by anyone with a sound mind, although in some particularly complex cases one may be at a loss in clearly demarcating right from wrong. For Clarke, being grounded in necessary relations, morality, like geometry, is universal and necessary. As such, it is independent of any will, be it divine or human, and of any consideration of punishment or reward as anyone, but not Hobbes, can plainly see. So, Clarke's view thus far can be characterized as a variety of rationalist deontology.

For Clarke, morality has three main branches: dealing with duties toward God, other humans, and oneself—all grounded in the notion of fitness. Duties toward

others are governed by equity, which demands that one deals with other persons as one can reasonably expect others to deal with oneself, and by love, which demands that one furthers the happiness of all persons. Duties toward oneself demand that one preserves one's life and spiritual well-being so as to be able to perform one's duties. Suicide, then, is wrong.

Since God's will is uncorrupted by self-interest or passion, divine volitions and moral commands are extensionally equivalent. Hence, God wants one to follow morality, and such a desire is manifested in laws God has set up. But since laws require sanctions, and since such sanctions are not uniformly present in this life, moral laws are associated with reward and punishment in the next life. Moreover, human depravity makes the prospect of future sanctions a necessary incentive for proper behavior.

However, Clarke seemed prepared to go further, claiming against the Stoics and his beloved Cicero that in one's present state virtue is not the highest good (this being happiness) and that consequently it would be unreasonable, not just psychologically difficult, to lay down one's life for the sake of duty. Virtue, Clarke claimed, is not happiness but only a means to it, as in a race running is not itself the prize but the way to obtain it. The present sorry state of humankind, beset by ignorance, prejudice, and corrupt passions, renders divine revelation necessary, contrary to what deists think, and therefore the remaining lectures are mainly devoted to establishing the reliability of the Gospels.

Clarke's theory was criticized on several grounds. He never quite explained the nature of the relations among persons that ground morality, leaving both his followers and detractors to argue inconclusively about it. Hume famously charged Clarke's theory with motivational impotence because the intellectual perception of fitness cannot, alone, move the will. Matthew Tindal, who devoted chapter fourteen of his *Christianity as Old as the Creation* to an analysis of Clarke's ethics, noted that Clarke's rationalist strand hardly fits with his insistence on the need for Christian revelation, since his arguments establishing the reliability of scripture seem to require much more intellectual effort than the apprehension of one's moral duties. Even more pointedly, Tindal, who approved Leibniz's claim that the Chinese should send missionaries in natural theology and its subsequent morality to Europe, noted that revelation is neither necessary nor sufficient for proper moral behavior even for common people.

See also Arius and Arianism; Bayle, Pierre; Berkeley, George; Boyle, Robert; Butler, Joseph; Cicero, Marcus Tullius; Collins, Anthony; Deism; Descartes, René; Determinism and Freedom; Gassendi, Pierre; Hume, David; Laws of Nature; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Matter; Miracles; More, Henry; Newton, Isaac; Pascal, Blaise; Renaissance; Rohault, Jacques; Space; Spinoza, Benedict (Baruch) de; Stoicism; Time; Tindal, Matthew; Toland, John; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY CLARKE

Works of Samuel Clarke (1738). 4 vols. New York: Garland, 1978. This contains all of Clarke's published works.

A Demonstration of the Being and Attributes of God and Other Writings, edited by Ezio Vailati. New York: Cambridge University Press, 1998. This offers a selection of passages relevant to Clarke's first Boyle Lecture besides the lecture itself.

WORKS ABOUT CLARKE

Ariew, Roger, ed. *Correspondence: G. W. Leibniz and Samuel Clarke*. Indianapolis, IN: Hackett, 2000. This contains Clarke's famous exchange with Leibniz.

Attfield, Robin. "Clarke, Collins, and Compounds." *Journal of the History of Philosophy* 15 (1977): 45–54.

Attfield, Robin. "Clarke, Independence, and Necessity." *British Journal for the History of Philosophy* 1 (2) (1993): 67–82.

Ducharme, Howard. "Personal Identity in Samuel Clarke." *Journal of the History of Philosophy* 24 (1986): 359–383.

Ferguson, James P. *Dr. Samuel Clarke: An Eighteenth Century Heretic*. Kineton, U.K.: Roundwood Press, 1976.

Ferguson, James P. *The Philosophy of Dr. Samuel Clarke and Its Critics*. New York: Vantage Press, 1974.

Force, James. "Samuel Clarke's Four Categories of Deism, Isaac Newton, and the Bible." In *Scepticism in the History of Philosophy: A Pan-American Dialogue*, edited by Richard H. Popkin, 53–74. Dordrecht, Netherlands: Kluwer Academic, 1996.

Gay, John. "Matter and Freedom in the Thought of Samuel Clarke." *Journal of the History of Ideas* 24 (1963): 85–105.

Hoskin, Michael "Causality and Free Will in the Controversy between Collins and Clarke." *Journal of the History of Philosophy* 25 (1987): 51–67.

Hoskin, Michael "Mining All Within: Clarke's Notes to Rohault's *Traite de Physique*." *The Thomist* 24 (1961): 253–263.

Le Rossignol, James E. *The Ethical Philosophy of Samuel Clarke*. Leipzig, Germany: n.p., 1892.

Rowe, William L. *The Cosmological Argument*. Princeton, NJ: Princeton University Press, 1975.

Stewart, Larry. "Samuel Clarke, Newtonianism, and the Factions of Post-revolutionary England." *Journal of the History of Ideas* 42 (1981): 53–71.

Vailati, Ezio. *Leibniz and Clarke: A Study of Their Correspondence*. New York: Oxford University Press, 1997.

Whiston, William. *Historical Memoirs of the Life of Dr. Samuel Clarke*. London: N.p., 1730.

Yolton, John W. *Thinking Matter: Materialism in Eighteenth-Century Britain*. Minneapolis: University of Minnesota Press, 1983.

Ezio Vailati (2005)

CLASSICAL FOUNDATIONALISM

Classical foundationalism maintains that all knowledge and justified belief rest ultimately on a *foundation* of knowledge and justified belief that has not been inferred from other knowledge or belief. Because the classical foundationalist typically assumes an account of knowledge in terms of justified or rational true belief, it might be best to focus on the distinction invoked between inferentially and noninferentially justified beliefs. What is written in this entry will apply mutatis mutandis to the distinction between inferential and noninferential knowledge.

THE PRINCIPLE OF INFERENTIAL JUSTIFICATION

If one thinks about most of the beliefs one takes to be justified and asks what justifies them, it seems natural to answer in terms of other justified beliefs. A person's justification for believing that it will rain, for example, may consist in part of that person's justifiably believing that the barometer is dropping rapidly. But under what conditions can one justifiably infer the truth of one proposition *P* from another *E*? The classic foundationalist typically insists that to be justified in believing *P* on the basis of *E* one must be justified in believing *E*. So, for example, one cannot be justified in believing that the world will end tomorrow by basing that belief on an unsupported hunch that the earth will be hit by a giant meteor. More controversially, many classic foundationalists—at least implicitly—also seemed to presuppose that to be justified in believing *P* by inferring it from *E* one must also be justified in believing that *E* confirms (makes probable) *P* (where *E*'s entailing *P* is the upper limit of *E*'s making probable *P*). Thus, one cannot justifiably infer the arrival of Armageddon from a fortune-teller's prediction that the world will end tomorrow unless one has some good reason to believe that the fortune-teller's predictions make probable the occurrence of the events predicted. Call the principle stating both of the above requirements for justification the principle of inferential justification (PIJ):

> To be justified in believing *P* on the basis of *E* one must be: (1) justified in believing *E*; and (2) justified in believing that *E* makes probable *P*.

The principle of inferential justification is a crucial premise in the famous regress argument for foundationalism. If the principle is correct, then to be justified in believing some proposition *P* on the basis of some other evidence, *E1*, one would need to be justified in believing *E1*. But if all justification were inferential, then to be justified in believing *E1* one would need to infer it from something else *E2*, which one justifiably believes, and so on ad infinitum. This first regress is generated invoking only clause (1) of the principle of inferential justification. If the second clause is correct, the potential regresses proliferate endlessly. To be justified in inferring *P* from *E1* one must justifiably believe not only *E1* but also that *E1* makes likely *P*, and one must infer this from something else *F1*, which one must justifiably infer from some other proposition *F2*, which one justifiably infers …. And so on.

But one must also justifiably believe that *F1* makes likely that *E1* makes likely *P*, so one must justifiably infer *that* from some other proposition *G1*, which one justifiably infers …. And so on. If all justification were inferential then to justifiably believe any proposition *P* a person would need to complete not just one but an infinite number of infinitely long chains of reasoning. However, the human mind is finite and cannot complete infinitely long chains of reasoning. To avoid the absurd conclusion that people cannot ever be justified in believing anything whatsoever, we must suppose that some beliefs are justified without inference and that these noninferentially justified beliefs ground the justification of all other justified beliefs.

The principle of inferential justification is also often a critical assumption of classic skeptical arguments, most of which presuppose a strong form of foundationalism. So, for example, Hume seemed to conclude that we have no reason for believing any description of an external world when ultimately all we have to rely on as evidence is our knowledge of fleeting and subjective experience. The problem, Hume argued, is that we have no way of establishing sensations as reliable indicators of the existence of external objects that they take to be their cause. Indeed, the difficulty of avoiding a fairly radical skepticism within the constraints of classical foundationalism is one reason so many philosophers became disillusioned with the view.

NONINFERENTIAL JUSTIFICATION

Classical foundationalists refer to the foundations of knowledge and justified belief in a variety of ways—for example: noninferentially justified beliefs, self-evident truths, directly evident truths, incorrigible beliefs, infallible beliefs, and so on—but there is no consensus on what confers foundational status on a belief. Some, following Descartes, seek foundations in beliefs that do not admit of the possibility of error. As will be seen, the possibility in question may be interpreted in a number of different ways, but classical foundationalists usually invoked a very strong concept of possibility: If a belief is foundational it must be inconceivable that the belief be false. The having of the belief must somehow entail its truth. Thus Descartes famously purported to find an ideal foundation for knowledge in one's belief that one existed. It seems trivially true that if someone *S* really does believe that he or she exists, that belief couldn't possibly be false. *S* has to exist in order to believe that *S* exists (or to be in any other conscious state).

Still other foundationalists sought to identify noninferential justification with whatever fact is the truth-maker for the alleged noninferentially justified belief. So, for example, some foundationalists would claim that my justification for believing I am in pain—when I am—is the pain itself. Of course, such a view hardly qualifies as a philosophical theory until its proponent gives a principled account of how some truth-makers justify us in believing the claims they make true, whereas others do not.

Although it was not always spelled out, many other classical foundationalists sought the source of foundational knowledge in some relation (other than belief) obtaining between a believer and the truth conditions of what is believed. One metaphor often invoked is the concept of acquaintance. When one believes that one is in pain when one is in pain, for example, one is directly acquainted or confronted with the pain itself (the very state that makes true the proposition believed). It is the knower's direct confrontation with the relevant aspect of reality to which the truth in question corresponds that obviates the need for any inference. Another variation on the view might insist that noninferential justification consists not just in acquaintance with the fact that is the truth-maker for one's belief but also acquaintance with the correspondence between the truth bearer (sometimes taken to be a thought or "picture" of reality) and the truth-maker.

In addition to direct acquaintance with contingent facts that can yield noninferentially justified beliefs in empirical propositions, there may also be direct acquaintance with logical relations holding between propositions, states of affairs, or properties that yields direct knowledge of necessary truths. So, for example, one might claim that one's noninferential justification for believing that squares have four sides is constituted in part by one's acquaintance with the properties of being a square and having four sides and the way in which the former contains the latter. Or one might hold that one's noninferential justification for believing that nothing can be both red all over and blue all over at the same time is constituted in part by one's acquaintance with the way in which being red excludes being blue.

On the above view, one might locate the source of both a priori and a posteriori foundational knowledge in the same relation of acquaintance. Traditionally, philosophers have made a great deal of the distinction between a priori knowledge (knowledge of necessary truths that is in some sense independent of sense experience) and a posteriori knowledge (knowledge of contingent truth that somehow relies on sense experience). But it is hard to see in what sense knowledge of one's own beliefs, for example, fits neatly into this traditional way of making the distinction. That one believes that it will rain tomorrow is a contingent truth that one knows, but it doesn't seem that one's knowledge of that truth depends on sense experience. On the acquaintance theory, the difference between a priori and a posteriori knowledge might better be thought of as lying more on the side of the relata of the acquaintance relation than on the source of the knowledge.

CRITICISMS

Classical foundationalism has come under considerable attack from many different directions. The second clause of the principle of inferential justification is particularly controversial. A worry exists that it is far too strong a requirement for inferential justification and may simply invite a vicious regress. In assessing the claim that inferential justification requires access to a probability connection between one's premises and one's conclusion, it is important to make sure that the arguments one considers are not enthymematic. as we ordinarily talk, it is natural to describe the dark clouds overhead as evidence of an approaching storm. But it is doubtful that the the presence of the clouds by itself constitutes the entire body of evidence from which people predict the storm; it is the dark clouds together with one's knowledge of a past association between dark clouds and storms. One might argue that when one considers genuinely non-

enthymematic reasoning it is less plausible to suppose that one needs knowledge of connections between premises and conclusion in order to legitimately infer one's conclusion.

Still, even in the case of deductively valid arguments there is a great deal of plausibility to the claim that one cannot get justification for believing the conclusion of the argument unless one not only has reason to believe the premises but also sees the connection between premises and conclusion. To avoid regress, people need noninferential knowledge of connections between premises and conclusions; and whereas it may not be that hard to convince oneself that one can discover without inference that one proposition entails another, it is much harder to convince oneself that one can just "see" probability connections (connections that are lower than entailment).

Without noninferential awareness of probability, however, skepticism looms on the horizon. Of course, in deciding what one can or cannot be noninferentially justified in believing, the question of just what might constitute noninferential justification needs to be addressed.

WHAT CONSTITUTES NONINFERENTIAL JUSTIFICATION? Some would argue that the search for infallible beliefs as the foundations of knowledge is both fruitless and misguided—at least if infallibility is understood in terms of a belief's entailing the truth of what is believed. As has been shown, there are trivial examples of beliefs that do entail the truth of what is believed. My beliefs that "I exist; that I am conscious; that I have beliefs" are all trivially infallible in the sense defined. Critics have pointed out, however, that if one believes a necessary truth, one's belief will also trivially entail the truth of what is believed. If one says that *P* entails *Q* when it is impossible for *P* to be true while *Q* is false, then every proposition will entail a necessary truth—necessary truths cannot be false.

But surely belief in a necessary truth does not constitute knowledge if the person holds the belief as a matter of pure whimsy. If one becomes irrationally convinced that every third sentence of a book expresses a truth, and by employing this decision procedure for belief ends up, by a remarkable coincidence, believing an extraordinarily complex necessary truth (far too complex for one to even recognize as a necessary truth) it hardly seems right to suppose that one would have any justification whatsoever for believing that truth. Once one sees that the entailment relation between belief and the truth of what one believes is not sufficient for knowledge or justified belief, one might begin to wonder whether it is ever getting at the heart of any interesting *epistemic* concept.

Still other philosophers have pointed out that beliefs that entail their truth are few and far between, and that if knowledge rests on a foundation of these, then that foundation is precarious indeed. Consider a favorite example of a foundational belief offered by classical foundationalists: the belief one has that one is in pain. Believing that one is in pain seems to be a state logically distinct from the pain. As such it seems always at least conceivable that the belief could occur—perhaps produced by some evil demon—without the pain. for all we know, the brain state causally responsible for one's believing that one is in pain is a distinct brain state from the one causally responsible for the pain. If so, then one could presumably induce belief in pain without producing the pain. Yet if one cannot get foundational justification or knowledge for accepting descriptions of one's own psychological states, then an impoverished foundation indeed exists upon which to attempt to build an edifice of knowledge.

EXTERNALIST APPROACHES TO INFALLIBILITY. Some contemporary philosophers are sympathetic to the idea of direct knowledge understood in terms of beliefs that cannot be false, but have understood the relevant possibility in causal or nomological terms. Thus the circumstances that produce the belief that *P* may be *causally* sufficient for the truth of *P*. It is not easy to spell out in an interesting way how one might specify the relevant circumstances causally responsible for a belief, but this approach does succeed in raising an alternative to the classical foundationalists' emphasis on conceivability or logical possibility as the relevant concept to employ in defining epistemically interesting concepts of infallibility. It also raises the prospects of a much richer array of propositions being contained in the foundations. Such externalist approaches to understanding infallibility, however, are probably anathema to classical foundationalists, who typically wanted the conditions that constitute foundational knowledge or justification to be conditions to which people have a kind of unproblematic direct access (a desire that might itself raise again the specter of vicious regress).

It was pointed out earlier that one cannot very well identify truth-makers as the source of noninferential justification without giving a plausible account of just what gives some truth-makers a critical epistemic role that most others fail to have. But reliance on the concept of acquaintance to define the concept of foundational knowledge has not fared much better when it comes to

contemporary philosophical fashion. The standard line most often taken is that there is no such relation and, even if there were, it would be of no epistemic interest. Foundational knowledge must be knowledge of *propositions* if it is to yield the premises from which people can infer the rest of what they justifiably believe. But acquaintance with a fact seems to be a relation that has nothing to do with anything that has a truth value. Facts are not the kinds of things that can be true or false. How does acquaintance with a fact yield access to truth? Indeed, can one even make sense of reference to facts independently of truth? Some philosophers would argue that to refer to a fact is just another way of referring to a proposition's being true. If facts are reducible to truths, it would clearly be uninformative to locate the source of noninferential knowledge of truths in terms of acquaintance with facts to which truths correspond.

The claim that acquaintance with facts is not by itself constitutive of noninferential knowledge of truths is one that an acquaintance theorist might grant, however. As was noted earlier, one might introduce a critical role for truth-bearers to play in one's acquaintance theory. To be noniferentially justified in believing some proposition *P*, one might argue, one must be acquainted not only with the fact that *P* but the truth-making relation of correspondence between the thought that *P* and the fact that *P*. When one has acquaintance with the truth-bearer, the truth-maker, and the truth-making relation that holds between them, one is in a complex state that does just constitute the most fundamental kind of propositional knowledge.

NONINFERENTIAL JUSTIFICATION. To attack various versions of foundationalism is not, of course, to respond to the regress argument for foundationalism. It has already been noted that some contemporary foundationalists accept the fundamental idea that there are foundations to knowledge but reject classical accounts of what those foundations consist in. As was seen in considering alternative conceptions of infallibility, many externalists identify justificatory conditions for belief with the circumstances producing the belief. Reliabilists, for example, count a belief as justified if it is reliably produced and they allow that a belief might be reliably produced even if the input producing the belief involves no other beliefs. Such reliable "belief-independent" processes can end a regress of beliefs justified by reference to other beliefs. Reliabilist conceptions of noninferential justification also divorce noninferential justification from infallible justification. According to some reliabilists, a justified belief might result from nondoxastic input and be just barely more likely to be true than not.

As was true of those who seek foundations in the causal impossibility of mistake, reliabilists offer the prospect of a greatly expanded class of propositions that might be noninferentially justified. Like other versions of externalism, however, it is not clear that reliabilism succeeds in capturing a concept of justification that would interest the classical foundationalist. The classical foundationalist sought justification that would provide a kind of assurance of truth, and it is not clear how the causal origin of a belief by itself (when one has no access to that origin) could satisfy one's intellectual curiosity.

THE COHERENCE THEORY OF JUSTIFICATION

Historically, the other main alternative to classical foundationalism was the coherence theory of justification. The coherentist rejects the classical foundationalist's assumption that justification is linear in structure. According to the coherentist, there is no escape from the circle of one's beliefs—nothing can justify a belief but other beliefs. But one doesn't justify a belief by reference to other prior justified beliefs. Rather, each belief is justified by reference to its fit in an entire system of beliefs. When each belief does its part in contributing to a clear, coherent picture of the world, each belief is justified. The coherentist, however, faces a serious dilemma. The coherentist must choose between the view that coherence by itself confers positive epistemic status on the beliefs that cohere, and the view that it is one's awareness of the coherence between one's beliefs that confers such status. If the coherentists embrace the first horn of the dilemma, they are left with a view that seems vulnerable to counterexample. Does one really want to allow that if one consults one's astrologer and comes to believe a set of complex propositions that coincidentally cohere beautifully—even though due to their complexity one could never discover that coherence—the beliefs in question are all justified?

However, if one requires that one must be aware of the coherence among one's beliefs in order to acquire justification for those beliefs, one faces once again the regress that drove so many to foundationalism. To be aware of coherence one must be aware of the fact—that is, have a justified belief—that one has the beliefs one has and that they stand in various logical and probabilistic connections. But how does one come to know what one believes? If one answers in terms of coherence to which one has access the problem just arises again. If one gives

oneself unproblematic direct access to one's beliefs and the connections that hold between them, one has simply returned to classical foundationalism.

See also Analysis, Philosophical; A Priori and A Posteriori; Coherentism; Correspondence Theory of Truth; Descartes, René; Epistemology; Evidentialism; Hume, David; Knowledge, A Priori; Propositions.

Bibliography

Armstrong, David. "Is Introspective Knowledge Incorrigible?" Philosophical Review 72, (1963): 427–32.

Audi, Robert. *The Structure of Justification*. Cambridge, U.K.: Cambridge University Press, 1993.

Ayer, A. J. *The Foundations of Empirical Knowledge*. New York: St. Martin's Press. 1955.

BonJour, Laurence, and Sosa, Ernest. *Epistemic Justification*. Oxford: Blackwell, 2003.

Butchvarov, Panayot. *The Concept of Knowledge*. Evanston, IL: Northwestern University Press, 1970.

Chisholm, Roderick. *The Foundations of Knowing*. Minneapolois, MN: University of Minnestoa Press, 1982.

Conee, Earl, and Feldman, Richard. *Evidentialism*. Oxford: Oxford University Press, 2004.

De Paul, Michael, ed. *Resurrecting Old-Fashioned Foundationalism*. Totowa, NJ: Rowman and Littlefield, 2001.

Descartes, Rene. *Discourse on Method and Meditations*, trans. Laurence LaFleur. Indianapolis: Bobbs-Merrill, 1960.

Fales, Evan. *A Defense of the Given*. Lanham, MD.: Rowman and Littlefield, 1986.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Fumerton, Richard. *Metaepisptemology and Skepticism*. Lanham, MD: Rowman and Littlefield, 1996.

Klein, Peter. "Human Knowledge and the Infinite Regress of Reasons." *Philosophical Perspectives* 13 (1999): 297–325.

McGrew, Timothy. *The Foundations of Knowledge*. Lanham, MD: Littlefield Adams. 1995.

Plantinga, Alvin. "Justification in the 20th Century." In *Philosophical Issues 2: Rationality in Epistemology*, edited by Enrique Villanueva, 43–78. Atascadero, CA: Ridgeview Publishing Co., 1992.

Price, H. H. *Perception*. London: Methuen, 1950.

Russell, Bertrand. "Knowledge by Acquaintance and Knowledge by Description." *Proceedings of the Aristotelian Society*, 11, 108–128. Reprinted in Russell, Bertrand, *Mysticism and Logic*, London: Allen and Unwin, 1963, 152–167.

Sellars, Wilfred. *Science, Perception and Reality*. London: Routledge and Kegan Paul, 1963.

Richard Fumerton (1996, 2005)

CLASSICAL MECHANICS, PHILOSOPHY OF

Classical physics is the research tradition beginning with Isaac Newton's *Mathematical Principles of Natural Philosophy* (often called simply the *Principia*) of 1687, which was overtaken by relativity theory and quantum mechanics in the early twentieth century and is still undergoing lively development in such areas as chaos and catastrophe theory. The "Newtonian" physics canonized in textbooks includes many elements added long after Newton, such as vector notation, the analytical mechanics that Joseph-Louis Lagrange and William Hamilton developed in the late eighteenth and early nineteenth centuries, and the laws of energy conservation and of electromagnetic phenomena formulated in the mid-nineteenth century. Indeed, Leonhard Euler in 1749 was the first to express Newton's second law as the familiar relation between a body's instantaneous acceleration and the momentary force that the body experiences; Newton's own version of the law set the body's change in momentum during a finite period of time equal to the impulse on the body (the force times the period's length, for constant force over the period). However, the subject of this entry is the anachronistic classical mechanics found in textbooks.

Though classical mechanics is false, as relativity and quantum mechanics reveal, there are many reasons for philosophers to continue investigating its proper interpretation (i.e., what the world would be like if classical mechanics were true). Many of the difficulties encountered in trying to interpret modern physics also arise in connection with classical physics, but in a simpler context. Moreover, many of the venerable metaphysical and epistemological ideas vigorously developed by modern philosophers such as George Berkeley, David Hume, and Immanuel Kant are best understood in connection with the classical physics that originally prompted them. Furthermore, one should not wait to deploy one's interpretive faculties only after physics has secured the final theory of everything; if one did, progress in both physics and philosophy would suffer. Finally, although classical physics lacks some of the provocative features exhibited by relativity and quantum mechanics, it has long served philosophers as the exemplar of what genuine scientific understanding would be. By studying it, one can learn about the concepts, logic, and limits of science. This entry touches on only a few of the metaphysical and epistemological questions that classical physics provokes.

BASIC ONTOLOGY: MASS AND MATTER

Philosophers have worked to identify the ontologically fundamental objects, properties, and relations that classical mechanics posits. Among the candidates proposed have been distance, time interval, velocity, force, matter, mass, electric charge, and inertial reference frame. All raise difficult questions.

Mass is the single parameter relating a body's motion to the force on the body. Remarkably, that relation in classical mechanics is the same for macroscopic bodies as for their constituents; classical mechanics "scales up." Newton defined mass as measuring the amount of matter composing a body, though he did not define *matter* itself. In contrast, the nineteenth-century Scots physicist James Clerk Maxwell (1952), who formulated the laws of electromagnetism, defined mass in terms of momentum and energy, which he believed more fundamental. An alternative approach later pursued by Ernst Mach (1960), the Moravian physicist and philosopher, characterizes mass operationally: The masses of two bodies are related as the inverse ratio of their mutually induced accelerations when isolated from other bodies. If mass is not defined operationally, but instead is an intrinsic property responsible for resistance to force, then according to many philosophers, one knows the effects for which mass is responsible, but one cannot know what mass is in itself. Similar considerations apply to electric charge.

Classical mechanics is sometimes interpreted as deeming a macroscopic body to be a swarm of point bodies in a void. In an alternative interpretation classical mechanics takes bulk matter to be continuous space-filling stuff, or instead to be composed of many bodies of a small but finite dimension made of continuous media, having no internal structure, and separated by empty space. Newton's laws and Charles-Augustin de Coulomb's electrostatic force law are often codified in terms of pointlike bodies, whereas the basic equations of hydrodynamics and the theory of elastic solids are typically expressed in terms of continua. Mass points present obvious difficulties; when two collide, they must be inside each other and the gravitational force between them becomes infinite. Continuous media avoid the latter problem, since at a point, there is no finite quantity of mass; there is only mass density, defined as the limit of mass per volume as the volume becomes arbitrarily small. But collisions still present a problem: When two bodies collide, do they occupy a common point? Or is there simply no finite volume between them? (If a point separates them, then how are they in contact?)

BASIC ONTOLOGY: MOTION AND FORCE

A body's velocity is its position's instantaneous rate of change, and its acceleration is its velocity's instantaneous rate of change. As ordinarily defined, a quantity's rate of change at an instant is its average rate of change during a finite interval around that instant, in the limit of an arbitrarily short interval. Hence, a body's velocity at time t is just a mathematical property of the body's trajectory in a neighborhood of t, which includes some of the body's trajectory after t. But its velocity at t is supposed to be an initial condition in the causal explanation of its subsequent trajectory. That would apparently require points in that subsequent trajectory to help causally explain themselves. This is puzzling. Furthermore, consider a body moving uniformly across the surface of a smooth horizontal table and then falling off the edge. At the final moment that the body is on the table (assuming that the table includes its edge), its trajectory's second derivative is undefined; taken from the left it is zero, but taken from the right it is equal to the gravitational acceleration. Presumably, though, a body has a well-defined acceleration at all times.

Force is characterized by William Thomson and Peter Guthrie Tait in their canonical mid-nineteenth-century physics text as "a direct object of sense" (1895–1896, p. 220). However, other natural philosophers regarded forces as redundant in classical mechanics once fields are admitted as local causes or remote charges and masses are acknowledged as acting at a distance. The late nineteenth-century German physicist Heinrich Hertz regarded forces as mere calculational devices between cause and effect, "simply sleeping partners, which keep out of the business altogether when actual facts have to be represented" (1956, p. 11).

FUNDAMENTAL LAWS

Newton's three laws of motion, with his inverse-square law of gravity, are commonly regarded as the fundamental laws of classical physics. In the nineteenth century the laws of electromagnetism were added to them. The status of the conservation laws and the variational principles of classical physics remains more controversial, as will be seen.

Newton's second law is sometimes taken to be "The net force on a body, divided by its mass, equals its acceleration in any inertial frame of reference." But how is *inertial frame* defined? Mach suggested that inertial frames are frames where the universe's average matter is not accelerating. But this definition leads to predictions

that depart from those made by Newtonian mechanics regarding, say, a body in otherwise empty space or in a universe where (according to Newton) all other matter is accelerating. Sometimes *inertial frame* is taken to be defined by Newton's first law: a frame is inertial exactly when a body feeling no forces remains at rest or in uniform rectilinear motion in that frame. But then Newton's first law is true by definition.

The chief alternative is to presuppose points of absolute space, as Newton did, and to define an *inertial frame* as a rigid Euclidean frame at rest or in uniform rectilinear motion with respect to those points. However, even disregarding objections to absolute space as either empirically inaccessible or in contravention of metaphysical scruples, Newton's approach contains surplus ontological structure. A particular frame need not be privileged as at rest. Rectilinear uniform motion need only be distinguished from other paths; the frames pursuing such trajectories are inertial. *Inertial frame* is thereby defined independent of Newton's first law, which is not tautologous but just a consequence of Newton's second law. (Newton's first law is never instantiated, since every body feels some component gravitational forces.)

In this "neo-Newtonian" space-time, there is no fact of the matter regarding a body's velocity. (But there is a fact regarding its acceleration and its velocity relative to another body.) There is also no fact regarding the distance between two nonsimultaneous events, unlike in Newton's absolute space and time. All inertial frames are equal in neo-Newtonian space-time.

However, absolute velocity figures in the classical laws of electromagnetism. Absolute space is then no longer superfluous. This fact opened one of Albert Einstein's paths to relativity theory.

HOW MUCH DOES CLASSICAL PHYSICS SAY?

Thomson and Tait interpret Newton's second law as requiring every acceleration to be caused by some force (1895–1896, p. 223). But simply as an equation, Newton's laws make no explicit mention of causes and effects; they merely relate a perpetually isolated system's past and future states to its current state. Accordingly, Bertrand Arthur William Russell concludes that the notion of a causal relation (insofar as it goes beyond a correlation demanded by the laws) has no place in physics, but is "a relic of a bygone age" (1929, p. 247). David Lewis (1983–1986) draws a different moral, arguing instead that since classical physics reveals causal relations, those relations must supervene on the laws and the actual course of events.

There are many similar questions about how richly or austerely classical physics describes the world. For instance, the law of energy conservation might be interpreted as specifying that the universe's total quantity of energy is fixed. But it might instead be taken as saying more: That for any volume over any temporal interval, the change in energy within must equal the energy that has flowed across its boundary. Alternatively, the conservation laws (of mass, energy, linear momentum, and angular momentum) might not be understood as laws of classical physics at all. They do not follow immediately from Newton's laws of motion and gravity.

TIME-REVERSAL INVARIANCE

Newton's laws are time-reversal invariant. Roughly speaking, if a sequence of events is permitted by the laws, then the laws also permit those events to occur in reverse order. The laws recognize no difference between past and future just as they fail to discriminate among spatial directions. However, certain macroscopic processes are never observed to occur in reverse. For example, when two bodies of unequal temperature touch, heat flows from the warmer to the cooler body. Although there are configurations of the bodies' molecules that would lead by Newton's laws to the warmer body's becoming still warmer, many more configurations would produce the result one sees. So irreversibility can be reconciled with Newton's laws if, roughly speaking, all the possible microrealizations of a system's macrostate are equally likely.

But this equiprobability is not required by Newton's laws. Its origin remains puzzling. Furthermore, even if a closed system far from equilibrium (e.g., with an unequal distribution of heat) were much more likely to head toward equilibrium (i.e., to increase its entropy) than away from equilibrium, entropy's increase in the space-time region one observes would remain unexplained. It would still be mysterious why one's space-time region is so far from equilibrium in the first place.

MECHANISM AND DETERMINISM

Classical mechanics suggests that the universe is like a majestic clockwork, the laws fully determining the universe's past and future states given its present state, and a body changing its motion only because another body touches it. But gravity and electromagnetism apparently operate by action at a distance. Newton famously offered no hypotheses (*hypotheses non fingo*) regarding the means by which gravity operates. Accordingly, some natural

philosophers ceased to seek local causes for all effects. In contrast, Michael Faraday and Maxwell regarded fields of force as existing on a par with bodies. The field at a given location would cause a body there to feel a force. The field picture avoids positing action at a distance but departs significantly from the picture of material particles in the void: Fields occupy all locations, even where there is no ordinary matter.

In 1814 Pierre Simon de Laplace invoked his famous "demon" to explain the determinism of the clockwork universe:

> Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (1951, p. 4)

Twentieth-century research revealed that Laplace may have overstated the determinism of a universe governed by classical physics (although there is no obvious way in which the indeterminism of classical physics supports the freedom of the will). When two point bodies collide, their mutual gravitational interaction becomes infinite, yet the laws of energy and momentum conservation nevertheless allow an analytic solution to the classical equations of motion to be extended uniquely through the collision singularity. However, this extension is generally impossible when three bodies collide. Furthermore, Newton's laws enable a closed system of point bodies to undergo an infinite number of triple near collisions in a finite time (as the sequence of encounter times converges to some particular moment). By the slingshot effect resulting from these close approaches, certain bodies attain infinite acceleration in finite time and so afterward are absent from any finite region of the universe. They are literally nowhere to be found. Since Newton's laws are time-reversal invariant, they permit this sequence of events to proceed in reverse, so that "space invaders" suddenly appear in the system from nowhere. Determinism is thereby violated without a collision occurring. Of course, the invaders' unanticipated appearance violates mass, energy, and momentum conservation, illustrating that these principles fail to follow from Newton's laws alone.

ANALYTICAL MECHANICS

In 1661 Pierre de Fermat derived the law of refraction from the postulate that in traveling from one location to another, a ray of light takes the path that minimizes the travel time. To some (such as the eighteenth-century French mathematician Pierre-Louis Moreau de Maupertuis), Fermat's principle suggested that nature produces effects with the greatest economy, efficiency, or ease—demonstrating God's wisdom. However, this metaphysical moral was undermined somewhat by the discovery that light may also take the path of greatest travel time. For example, consider a point light source at the center of an ellipsoidal mirror. The points around the mirror's margin that can reflect light back to the center are exactly the two points along the mirror's minor axis (i.e., where the edge is closest to the center) and the two points along the mirror's major axis (i.e., where the edge is farthest from the center).

Fermat's principle was generalized by Euler, Lagrange, and Hamilton into the variational principles of analytical mechanics. Given the system's initial configuration (the initial positions and velocities of its particles) and final configuration, there are various paths (through configuration space) by which the system may get from one to the other. These paths may differ, for instance, in the time it takes the system to arrive at its final configuration and in the configurations through which the system passes along the way. Roughly speaking, the Euler-Lagrange "principle of least action" states that the time integral of the system's total kinetic energy is "stationary" along the actual path as compared to all sufficiently close possible paths. That is, roughly speaking, the sum of the kinetic energies at all the points along the path actually taken is a minimum, maximum, or saddle point as compared to the sums for similar paths that are not taken. (So "the principle of least action" does not demand that the action be "least.")

Similarly, Hamilton's principle states roughly that of all the possible paths by which the system may proceed from one specified configuration to another in a specified time, the actual path as compared to other possible, slightly different paths makes stationary the time integral of the system's Lagrangian (i.e., the difference between the system's total kinetic and potential energies). A possible path may violate energy conservation and other laws; Hamilton's principle picks out the path demanded by the laws. So to apply Hamilton's principle, scientists must contemplate counterlegals: what would have been the case, had the system violated natural laws in certain ways. But a possible path must respect the constraints on the

system, which may include a body's having to remain rigid or in contact with a certain surface.

These constraints may be plugged into the variational principles without the forces that constrain the system having to be specified. This gives variational principles a practical advantage over Newton's laws, since the forces of constraint may be unknown, and emphasizes the style of explanation that variational principles supply. Newton's laws are differential equations; they determine the instantaneous rates of change of the system's properties from the system's conditions at that moment, such as the forces on it. The system's trajectory over a finite time interval is then built up, point by point, and the forces are efficient causes of the system's acceleration. In contrast, variational principles make no mention of forces; instead, they invoke the system's energy. The explanations they supply specify no efficient causes. Variational principles involve integral equations; they determine the system's trajectory as a whole, rather than point by point.

TELEOLOGY

Explanations that use variational principles sound teleological; the system appears to aim at making a certain integral stationary. But then the system's final configuration apparently helps to explain the path that the system takes to that destination; later events help to explain earlier ones. That is puzzling. How does a light ray "know," at the start of its journey, which path will take less time? How can the light adjust the earlier part of its route to minimize its later path through optically dense regions (where it cannot travel as fast) unless it knows about those distant regions before it sets off?

Some natural philosophers (such as Max Planck) suggested that variational principles are more basic laws than Newtonian differential equations, especially considering that unlike Newton's laws, variational equations of the same form apply to any set of variables sufficient to specify the system's configuration. Other natural philosophers (notably Gottfried Wilhelm Leibniz) embraced both mechanical and teleological explanations as equally fundamental. Leibniz declared that there are

> two kingdoms even in corporeal nature, which interpenetrate without confusing or interfering with each other—the realm of power, according to which everything can be explained mechanically by efficient causes when we have sufficiently penetrated into its interior, and the realm of wisdom, according to which everything can be explained architectonically, so to speak, or by final causes when we understand its ways sufficiently. (1969, pp. 478–479)

Other natural philosophers (such as Mach) denied final causes but also denied efficient causes as well (allowing only the relations specified by natural laws). The most common view, however, has been to reject teleological explanations as a relic of anthropomorphic characterizations of nature and to regard variational principles as logical consequences of more fundamental, mechanical laws. The variational principles follow from the Newtonian differential equations roughly because the entire path can minimize the integral only if each infinitesimal part does (since otherwise, by replacing that part with another, one would create a new path with a smaller integral), and the minimum for each infinitesimal part reflects the gradient of the potential there, which is the force. The variational principle thus arises as a byproduct of the relation between the force and an infinitesimal section of the path.

See also Berkeley, George; Chaos Theory; Determinism, A Historical Survey; Faraday, Michael; Hamilton, William; Hertz, Heinrich Rudolf; Hume, David; Kant, Immanuel; Laplace, Pierre Simon de; Laws, Scientific; Leibniz, Gottfried Wilhelm; Lewis, David; Logic, History of: Precursors of Modern Logic; Mach, Ernst; Maxwell, James Clerk; Newton, Isaac; Philosophy of Physics; Planck, Max; Quantum Mechanics; Relativity Theory; Russell, Bertrand Arthur William.

Bibliography

Diacu, Florin, and Philip Holmes. *Celestial Encounters: The Origins of Chaos and Stability*. Princeton, NJ: Princeton University Press, 1996.

Earman, John. *A Primer on Determinism*. Dordrecht, Netherlands: Reidel, 1986.

Earman, John. *World Enough and Space-Time: Absolute versus Relational Theories of Space and Time*. Cambridge, MA: MIT Press, 1989.

Feynman, Richard, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*. 3 vols. Reading, MA: Addison-Wesley, 1963–1965.

Hertz, Heinrich. *The Principles of Mechanics*. Translated by D. E. Jones and J. T. Walley. New York: Dover, 1956.

Lange, Marc. *An Introduction to the Philosophy of Physics: Locality, Fields, Energy, and Mass*. Oxford, U.K.: Blackwell, 2002.

Laplace, Pierre Simon de. *A Philosophical Essay on Probabilities*. Translated by Frederick Wilson Truscott and Frederick Lincoln Emory. New York: Dover, 1951.

Leibniz, Gottfried Wilhelm. *Philosophical Papers and Letters*. 2nd ed. Translated and edited by Leroy E. Loemker. Dordrecht, Netherlands: Reidel, 1969.

Lewis, David. *Philosophical Papers*. 2 vols. New York: Cambridge University Press, 1983–1986.

Lindsay, Robert Bruce, and Henry Margenau. *Foundations of Physics*. New York: Wiley, 1936.

Mach, Ernst. *The Science of Mechanics: A Critical and Historical Account of Its Development*. 6th ed. Translated by Thomas J. McCormack. LaSalle, IL: Open Court, 1960.

Maxwell, James Clerk. *Matter and Motion*. New York: Dover, 1952.

Newton, Isaac. *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World*. 2 vols. Translated by Andrew Motte and revised by Florian Cajori. Berkeley: University of California Press, 1962.

Russell, Bertrand. "On the Notion of Cause." In *Our Knowledge of the External World*. London: Norton, 1929.

Sklar, Lawrence. *Philosophy of Physics*. Boulder, CO: Westview Press, 1992.

Thomson, William, and Peter Guthrie Tait. *Treatise on Natural Philosophy*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1895–1896.

Yourgrau, Wolfgang, and Stanley Mandelstam. *Variational Principles in Dynamics and Quantum Theory*. 3rd ed. London: Pitman, 1968.

Marc Lange (2005)

CLASSIFICATION

See *Definition; Logical Terms, Glossary of*

CLAUBERG, JOHANNES

(1622–1665)

Johannes Clauberg, a German Cartesian philosopher, was born in Soligen, February 24, 1622, and died in Duisburg, January 31, 1665. Though he lived a short life, his philosophical output was considerable; his name became almost synonymous with that of René Descartes in Germany. Clauberg studied in Cologne and Bremen, where he came under the influence of reformed scholasticism and the pedagogical and methodological ideals of Jan Amos Comenius. At Bremen he also met Tobias Andreae, whom he later joined in Groningen in 1644 after Andreae was appointed professor of History and Greek. He disputed some theses in 1646 and published his first independent treatise, *Ontosophia*, in 1647. Clauberg's initial works, including *Ontosophia*, do not display the influence of Descartes's philosophy, though Clauberg rewrote the book along Cartesian lines in later editions. After travels to France, to the Protestant Academy in Saumur and Paris (where he seems to have met some early Cartesians), and to England, Clauberg attended the lectures of the Cartesian Johannes de Raey in Leiden in 1648. It is clear that by 1648 Clauberg had become interested in Descartes's philosophy. Clauberg made his official entrance into the Cartesian world as a result of his participation in what is sometimes called the "Conversation with Burman." The latter is a manuscript of the University of Göttingen reporting a lengthy discussion between Descartes and (presumably) Frans Burman, a young theology student at Leiden. The discussion, conducted in Latin, apparently occurred on April 16, 1648, at Descartes's retreat in Egmont. According to the manuscript, Burman dictated his impressions of the meeting to Clauberg on April 20. Clauberg evidently kept a copy and had a second one made by some unknown scribe some months later; this is the surviving copy.

In that period Clauberg was approached about becoming a professor of theology in Herborn; he began his duties the following year, in 1649, as professor of philosophy instead, with occasional teaching in theology. However, he was not happy with his position; his teaching load was heavy and he probably resented the combination of theology and philosophy, protesting as well that the professor of theology had some teaching duties in philosophy. A conflict with his more conservative colleagues developed. On November 1, 1651, Clauberg's employer, the Count of Nassau, officially decreed that the only philosophy allowed in Herborn was Aristotelico-Ramist philosophy, either separately or jointly. As a result, Clauberg and his friend and fellow Cartesian Christoph Wittich, who had been appointed professor of mathematics, left Herborn in December 1651 and accepted posts in Duisburg, a town that fell under the jurisdiction of the Elector of Brandenburg. In Duisburg, Clauberg's position was initially Rector of the town's Gymnasium; when the Academy of Duisburg was opened in 1655 he and Wittich became doctors of theology. Clauberg married Catharina Mercator in 1652; they had one son and five daughters. For the rest of his life, the now-settled Clauberg lived the life of a professor in a small German town; he was even rector of the Academy in 1655 and 1659. He attracted many students to Duisburg, several of whom became professors themselves.

WORKS

Clauberg must have already started on his second book, *Defensio cartesiana*, when still in Leiden, though it was published only in 1652. It is primarily a reply to *Consideratio theologica* (1648), a detailed commentary on Descartes's *Discourse on Method* from an orthodox theological point of view, by the Leiden Professor Jacobus Revius. Clauberg also added materials attacking his erst-

while colleague Cyrianus Lentulus (or Lentz), Professor of Practical Philosophy at Hernborn. The *Defensio Cartesiana* provoked a reply from Revius, which Clauberg answered with *Initiatio Philosophi sive dubitatio cartesiana* (1655). The conflict also involved Andreae, who published a two-volume response to Revius in 1653–1654, triggering yet another treatise from Revius in 1654. In his defense of Cartesianism, Clauberg distinguished between Descartes's popular and his esoteric works; according to Clauberg, the *Discourse on Method* belongs to the first category, whereas the *Meditations* and *Principles of Philosophy* belong to the second.

The promulgation of Cartesianism required Clauberg to write a number of other works explaining Descartes's physics and metaphysics, such as *Paraphrasis in Renati Descartes Meditationes*, *Differentia inter Cartesianam et alias*, and *Physica*. Clauberg also published some volumes of disputations. But doubtless his most influential books were *Logica vetus et nova*, first published by Elzevier in 1654, and the smaller *Logica contracta*. After Clauberg's death, the Amsterdam professor of philosophy Johann Theodor Schallbruch provided an edition of his works, *Opera omnia philosophica*, partly based on unpublished material in the possession of Clauberg's son, Johann Christopher; the added material included Clauberg's notes on Descartes's *Principles of Philosophy*, his correspondence with Andreae, a biography of Clauberg, and a general index to all of Clauberg's treatises.

CARTESIANISM

Clauberg's work is a paradigm of what first-generation Cartesian scholastics needed to accomplish. Clauberg made progress elaborating Cartesian themes, such as espousing occasionalism for the relation between mind and body, and created texts to fill the gaps in the collegiate curriculum as it would be taught by a Cartesian. With the *Principles of Philosophy*, Descartes began the process of producing textbooks from which to teach Cartesian philosophy. However, scholastic textbooks usually had quadripartite arrangements mirroring the structure of the collegiate curriculum: logic, ethics, physics, and metaphysics. And Descartes produced at best only a partial physics and what could be called a general metaphysics; he did not finish his physics—he did not produce the expected final two parts of the *Principles of Philosophy* on animals and on man—and did not write a particular metaphysics. He did not produce a logic or ethics for his followers to use or to teach from. These things must have been perceived as glaring deficiencies in the Cartesian program and in the aspiration to replace Aristotelian philosophy in the schools.

So the Cartesians rushed in to fill the voids. One can understand Louis de la Forge's additions to the *Traité de l'homme*, for example, as an attempt to complete the physics, and Clauberg's later editions of *Ontosophia* or Baruch Spinoza's *Cogitata metaphysica*, for instance, as endeavors to produce a more conventional-looking metaphysics. Descartes, of course, saw himself as presenting Cartesian metaphysics as well as physics, both the roots and trunk of his tree of philosophy. But from the point of view of schools texts, the metaphysical elements of physics (general metaphysics) that needed to be discussed by Descartes—such as the principles of bodies: matter, form, and privation; causation; motion: generation and corruption, growth, and diminution; place, void, infinity, and time—were expected to be taught in a course on physics. The scholastic course on metaphysics (particular metaphysics) dealt with other topics, not discussed directly in the *Principles of Philosophy*, such as being, existence, and essence; unity, quantity, and individuation; truth and falsity; good and evil. Such courses usually also ended up with questions about knowledge of God, names or attributes of God, God's will and power, and God's goodness. The *Principles of Philosophy* by itself was not sufficient as a text for the standard course in metaphysics.

Clauberg's *Ontosophia*, however, discussed being in general, dividing it into its general and primary sense of "intelligible" being, a secondary and lesser sense of "something" to be distinguished from "nothing," and a third, particular sense of "real" being, being outside the intellect, or substance, contrasting it with accident and mode. Clauberg went on to talk about essence, existence, and duration. His remaining chapters concerned pairs of concepts such as one and many; true and false; good and evil; perfect and imperfect; distinct and opposite; the same and another; exemplar and image.

CARTESIAN LOGIC

Beyond completing Cartesian physics and metaphysics, there were even attempts at producing Cartesian ethics; a Latin-language manual called *Ethica*, printed in 1685, was said to have been authored by Descartes. Descartes never wrote such a work, but a translator was able to put together a tripartite treatise out of Descartes's own words: (1) on the greatest good, happiness, and free will; (2) on passions; and (3) on love. There were numerous stabs at creating Cartesian-style logic texts as well, Clauberg's *Logica vetus et nova* being first of its kind, together with the *Logique* of Jacques Du Roure. The attempt to publish

a Cartesian textbook that would mirror what was taught in the schools culminated in the famous multivolume works of Pierre-Sylvain Régis and of Antoine Le Grand, which included expanded versions of Cartesian physics and metaphysics, together with treatises on ethics and logic.

Scholastic logic, as taught in the seventeenth century, typically followed an order of topics dictated by the various books of Aristotle's *Organon: Categories, On Interpretation, Prior Analytics, Posterior Analytics, Topics, and Sophistical Refutations*. For example, after some preliminary questions on the usefulness of logic, whether logic can be called science or art, and the definition and divisions of logic, Scipion Dupleix wrote a six-part logic, corresponding to Aristotle's six logical works: (1) categories—that is, substance, quantity, quality, relations, and so forth; (2) nouns, verbs, and statements; (3) syllogism; (4) science and demonstration; (5) topics; and (6) paralogisms. One can say similar things about the logic textbooks of other early seventeenth-century scholastics, such as Eustachius a Sancto Paulo and Pierre du Moulin. Clauberg's *Logica contracta* followed a similar pattern, starting with the categories and continuing with attribute and accident, cause and effect, subject and adjunct, relation, whole and part, the same and other, universal and singular, definition, and division. His second part of logic began with the grades of judgment, qualitative statement, truth and falsity, opposition, conversion and equivalence, and composite statement, and continued with argument and syllogism, both perfect and imperfect, and true and false. Clauberg's third part of logic dealt with the grades of memory and his fourth part concerned teaching and dialectics, order, and fallacy. Very little of this was Cartesian.

A major problem to resolve in producing a Cartesian logic was that Descartes, in keeping with a standard Renaissance view, was extremely negative about the subject. According to Descartes in the *Discourse on Method* (repeating views he had previously elaborated in the unfinished *Rules*), syllogisms are useless: they serve to explain things one already knows, or even to speak without judgment on matters of which one is ignorant, rather than to learn them; although logic might contain true and good precepts, nevertheless there are so many other precepts mixed up with them, that are either harmful or superfluous, that it is practically impossible to separate them from one another. Descartes proposed instead his four rules of method—the rules of evidence, of the division of difficulties, of the order of inquiry, and of the completeness of enumerations—as a method of discovery exempt from the faults of formal logic.

However, Descartes also called his rules of method the principal rules of logic. According to Descartes, before applying himself to true philosophy a person who has only common and imperfect knowledge should study "logic," but not the logic of the Schools: Such logic corrupts good sense rather than increasing it. Descartes's logic instead teaches people to direct their reason with a view to discovering the truths of which they are ignorant. The more moderate late Cartesian views about logic were reinforced in a text familiar to Clauberg. Commenting to Burman on the *Discourse* passage about the harmful role of logic, Descartes supposedly asserted that his statements did not apply so much to logic, which provides demonstrative proof on all subjects, but to dialectic, which teaches how to hold forth on all subjects. Descartes's subtle shift in position allowed Clauberg to reinterpret Descartes's rules of method as part of logic, now integrated into a legitimate branch of learning that even included syllogisms.

Clauberg's *Logica vetus et nova* begins with a Prolegomena arguing, along Descartes's line from the end *of Principles of Philosophy*, Part 1, that the principal origin of error is to be found in the prejudices of childhood. Logic is the corrective for these mental imperfections; thus, in the first book of his logic, Clauberg devises a scheme that involves Descartes's rules of method and traditional logic, following the pattern of his *logica contracta*, as three "grades" or levels of logic. The first level has to do with accepting clear and distinct perceptions; it includes the rule of evidence and ends up with the rule about the division of difficulties, but it also discusses traditional topics such as: substance, attribute, and mode; essence and existence; universal and singular; definition; and division. The second level concerns right judgment and involves the rule about the order of inquiry, ending with the rule of the completeness of enumerations; it also discusses induction and syllogism. Clauberg's third level concerns memory.

Clauberg provided the initial pattern for Cartesian logic, though other Cartesians found it more expedient to follow more closely the scholastic order in logic, grafting on a section about method at the end of their treatises. Later Cartesian logics, such as the Port-Royal Logic and Le Grand's *Logick* are divided into four parts: (1) Ideas, including Aristotle's categories, universals, and names; (2) Propositions (or Judgments), truth and falsehood; (3) Reasoning (or Discourse), including syllogisms, topics, and sophisms; and (4) Method. By method, however,

these writers meant analysis and synthesis, which does not have to be anything particularly Cartesian, though we do find Descartes's rules of method enumerated in the chapters on analysis. The Port-Royal Logic supplanted Clauberg's logic and was ultimately adopted and abbreviated by Régis as his logic in his *General System of Cartesian Philosophy*. One can legitimately think, however, that Clauberg understood Descartes's views on logic better than subsequent Cartesians.

See also Cartesianism.

Bibliography

WORKS BY CLAUBERG

Tessarakas thesium philosophicarum, de logicae ab aliis disciplinis quibus cum vulgo confundi assolet distinctione … moderatore Tobias Andreae. Groningen, Netherlands: J. Nicolai, 1646.

Disputatio theologico-practica de conscientia … sub praesidio Matthiae Pasoris.… Groningen, Netherlands: Eissens, 1646.

Elementa philosophiae sive Ontosophia. Scientia prima, de iis quae Deo creaturisque suo modo communiter attribuuntur. Groningen, Netherlands: Nicolai, 1647. With editions in 1660 and 1664.

Defensio cartesiana adversus Iacobum Revium Theologum Leidensen, et Cyriacum Lentulum, professorum Herbornensem: pars prior exoterica, in qua Renati Cartesii Dissertatio de Methodo vindicatur, simul illustria Cartesianae logicae et philosophiae specimina exhibentur. Amsterdam: Elsevier, 1652.

Logica vetus et nova, modum inveniendae ac tradendae veritatis, in Genesi simul et analysi facili methodo exhibens. Amsterdam: Elzevier, 1654. With editions in 1656, 1658, 1683, 1685, and 1692.

Initiatio philosophi, sive dubitatio Cartesiana, ad metaphysicam certitudinem viam aperiens. Leiden, Netherlands: Wyngaerden, 1655.

De Cognitione Dei et nostri, quatenus naturali rationis lumine, secundum veram philosophiam, potest comparari, exercitationes centum. Duisburg, Germany: Wyngaerden, 1656, with another edition in 1685.

Logica contracta. 3rd ed. Duisburg, Germany: Cunradi, 1670. With editions in 1683, 1701, 1702, 1711, 1737, and 1721. There might have been a first edition plus a second edition in 1659. The *Logica contracta* is also printed with adaptations of Clauberg's works, such as *Ontosophia, quae vulgo metaphysica vocatur, notis perpetuis in philosophiae et theologiae studiosorum usum illustrata, a Joh. Henrico Suicero. In calce annexa est Claubergii logica contracta* (Zurich: Gessner, 1694).

Unterschied zwischen der cartesianischen und der sonst in den Schulen gebraüchlichen Philosophie. Duisburg, Germany: Wyngaerden, 1658. Dutch translation, Nijmegen: Van Hervelt, 1658. The work is best known in a Latin version, *Differentia inter Cartesianam et alias in Scholis usitatam Philosophiam conscripta.…* (Groningen, Netherlands: Edzardi, 1680) with another edition in 1681.

Paraphrasis in Renati Descartes meditationes de prima Philosophia. In quibus Dei existentia, et Animae humanae à Corpore distinctio demonstrantur. Duisburg, Germany: Wyngaerden, 1658, with another edition in 1668. Dutch translation, 1683.

Ars etymologica Teutonem e philosophiae fontibus derivata; id est, via Germanicarum vocum & origines & praestantiam detegendi, cum plurium, tum harum Vernunft, Suchen, Außspruch exemplis. Duisburg, Germany: Asendorf, 1663.

Physica, quibus rerum corporearum vis & natura … explicantur. Amsterdam: Elzevier, 1664.

Theologorum academiae Duisbergensis Johannis Claubergii et Martini Hundii disputationes selectae, quibus controversiae fidei adversus omnis generis adversarios. Duisburg, Germany: Haius, 1665.

Chilias thesium ad philosophiam naturalem pertinentium … disputanda in Academia Duisburgensi. Groningen, Netherlands: Cöllen, 1668.

Dictata physica privata, id est physica contracta seu theses physicae, commentario perpetuo explicatae. Frankfurt: Ilsner, 1681. With another edition in 1689.

Opera omnia philosophica. Edited by Johannes Theodor Schalbruch. 2 vols. Amsterdam: Blaeu, 1691. With another edition in 1710. Reprint, Hildesheim: Olms, 1968. (Contains *Physica Contracta*; *Disputationes Physicae*; *Theoria Corporum Viventium*; *Metaphysica de Ente, cum Notis*; *Paraphrasis in Meditationes Cartesii*; *Notae breves in Cartesii Principia Philosophiae: nunc primum editae*; *Exercitationes Centum de Cognitione Dei et Nostri*; *Logica vetus et nova*; *Logica contracta*; *Defensio Cartesiana*; *Dubitatio Cartesiana*; *Differentia Cartesianam inter & Vulgarem Philosophiam*; *Exercitationes & Epistolae Joh. Claubergii & Tob. Andreae varii argumentis, nunc primum editae*; plus an *Index locupletissimus* and a *Vita* of Clauberg by Henricus Christianus Henninus.)

WORKS ABOUT CLAUBERG

Balz, Albert G. A. *Cartesian Studies*. New York: Columbia University Press, 1951.

Bohatec, Joseph. *Die Cartesianische Scholastik in der Philosophie und Theologie der reformierten Dogmatik des 17. Jahrhunderts*. Leipzig, Germany: A. Deichert, 1912.

Brosch, Pius. *Die Ontologie des Johannes Clauberg. Eine historische Würdigung und eine Analyse Ihrer Probleme*. Greifswald, Germany: E. Hartmann, 1926.

Dibon, Paul. *L'enseignement philosophique dans les universités néerlandaises à l'époque pré-cartésienne*. Leiden, Netherlands: s.n., 1954.

Dibon, Paul. *Regards sur la Hollande du Siècle d'Or*. Naples: Vivarium, 1990.

Müller, Hermann. *Johannes Clauberg und seine Stellung im Cartesianismus*. Jena, Germany: H. Pohle, 1891.

Thijssen-Schoute, C. Louise. *Nederlands Cartesianisme*. Amsterdam: Noord-Hollandse Uitgeversmaatschappij, 1954.

Verbeek, Theo. *Descartes and the Dutch: Early Reactions to Cartesian Philosophy, 1637–1650*. Carbondale: Southern Illinois University Press, 1992.

Verbeek, Theo, ed. *Johannes Clauberg (1622–1665) and Cartesian Philosophy in the Seventeenth Century*. Dordrecht, Netherlands: Kluwer, 1999.

Weier, Winfried. *Die Stellung des Johannes Clauberg in der Philosophie*. Mainz, Germany: Johannes Gutenberg-Universität, 1960.

Wundt, Max. *Die deutsche Schulmetaphysik des 17. Jahrhunderts*. Tübingen, Germany: Mohr (Siebeck), 1939.

Roger Ariew (2005)

CLEANTHES

(c. 331–330 BCE–c. 230–229 BCE)

Cleanthes (sometimes referred to as Cleanthes of Assos) was the second head of the Stoic school. Ancient biographical information is found in Diogenes Laërtius's *Lives of the Philosophers* (7. 168-176) and in Philodemus's history of Stoicism (columns 18-19). Born in 331–330 BCE, in Asia Minor, he came to Athens in 281–280. He took over leadership of the school on the death of its founder, Zeno of Citium, in 262–261 and held that position until his own death in 230–229. The most important contemporary Stoic was Ariston of Chios, against whom Cleanthes defended the version of Zeno's legacy that became standard, insisting on the vital importance of logic and physics as well as ethics. His own student and successor, Chrysippus, maintained this integrated system. Cleanthes also defended Stoic epistemology against the skepticism of the academic Arcesilaus.

Cleanthes was a prolific author in every area of philosophy. He wrote four books of interpretation of Heraclitus, a defense of Zeno's natural philosophy, and works on the interpretation of poetry and myth, which probably aimed to show that ancient wisdom supported Stoicism. He is now better known for the surviving portions of his philosophical poetry, which includes thirty-nine lines of his *Hymn to Zeus* and four lines on the topic of fate. In physics he wrote on the basic principles (active and passive, God and matter), on cosmogony (a cyclical recreation of the cosmos punctuated by recurring conflagrations of all matter), and on cosmology (with a particular emphasis on the role of the sun as the organizing principle). In theology he is important for his theory about the origins of the conception of God reported in book two of Cicero's *On the Nature of the Gods* and other more dialectical arguments reported by Sextus Empiricus.

Cleanthes' response to the Master Argument of Diodorus Cronus was to hold (1) that there are possibilities which neither are nor will be true and (2) that the impossible does not follow from the possible; but (3) to deny that every past truth is necessary, thereby perhaps avoiding an excessively necessitarian version of his determinism. He was a materialist, holding that anything that causes or is caused must be material, but followed Zeno in invoking incorporeal predicates as necessary features of a causal account of material interaction. He may have been the first to use the term *lekton* ("sayable") for such items. He wrote several works on dialectic, logic, and epistemology, but ultimately his contribution in this area was eclipsed by that of his brilliant successor Chrysippus.

Like all Stoics he held that the soul is a material stuff, a warm, breathy substance capable of perception and intelligence; he invoked the authority of Heraclitus particularly for his psychology. One argument for the physical nature of the soul was the heritability of psychological traits as well as corporeal characteristics. Cleanthes held that the soul survived the death of the person but only until the next conflagration; postmortem survival of personal traits seems not to have been envisaged.

In ethics he held that living according to nature is equivalent to living virtuously and took a particularly strong anti-hedonistic stance, using parables and images to dramatize the starkness of the choices one must make in planning one's life. He held the controversial view that in planning one's life one must look only to cosmic nature rather than to any more limited nature (such as that of the species or the individual), a position that coheres with his theological and cosmological views. He wrote extensively on practical ethics, but held that the norms applicable specifically to individuals in their social roles must be based on general philosophical principles. He held a strong version of the unity of virtues, maintaining that it is a single disposition (called "strength" and "power" rather than "health") manifested as different virtues (such as courage and justice) according to the circumstances where it is applied. Virtue is a cognitive state consisting of the secure and irreversible knowledge of doctrines and factors relevant to decision making. Hence it is a permanent trait once achieved.

His psychology is often thought to have had a dualistic character because Galen, perhaps following Posidonius, exploited Cleanthes' writings when arguing against the monistic views of Chrysippus. It would, however, be a mistake to infer dualistic psychology from the fact that Cleanthes dramatized a debate between reason and emotion in a poem, no doubt for protreptic purposes. In all areas of his philosophy Cleanthes was committed to the main lines of Stoic orthodoxy as set down by his master Zeno.

See also Arcesilaus; Aristo of Chios; Chrysippus; Stoicism; Zeno of Citium.

Bibliography

H. Dörrie. *"Kleanthes" Real-Enzyklopedie der Altertumswissenschaft*. Suppl.-Bd. 1970 XII : Nachtr. 1705–1709.

Dorandi, T., ed. *Filodemo Storia dei filosofi: La stoà da Zenone a Panezio*. Leiden: Brill, 1994.

Guérard, C. "Cléanthe d'Assos." In *Dictionnaire des Philosophes Antiques*. Vol. 2, ed. R. Goulet. Paris: Éditions du Centre national de la recherche scientifique, 1989–2003.

Brad Inwood (2005)

CLEMENT OF ALEXANDRIA

(c. 150–c. 213)

Clement of Alexandria (full Latin name, Titus Flavius Clemens), the Christian theologian of the Alexandrian school, was born of pagan parents, probably in Athens. Clement learned from several teachers in the Mediterranean world before he came to Alexandria, where he studied under the Christian philosopher Pantaenus, a converted Stoic who was the head of the catechetical school. Clement remained in Alexandria from 175 to 202, writing and teaching, until he fled during the persecution of the emperor Septimius Severus. He died in Palestine.

Alexandria's heritage of learning, culture, syncretism, and religious mystery may be seen in his writing. His three major works form a trilogy that leads from paganism to mature Christianity. In the *Protrepticus* (Exhortation) he attacks the absurdities of pagan deities and exhorts his readers to turn to Christianity. In the *Paedagogus* (Tutor) he instructs Christians in the good life. In his chief work, the unfinished *Stromateis* (Patchwork), he sets down his philosophical opinions in unsystematic notes—"Gnostic notes concerning the true philosophy." This work, which represents the final stage of instruction, includes much material that he had learned from his teachers but hesitated to write about because of its difficult and sacred nature. He regards obscurity, compression of style, and haphazard arrangement as safeguards against the abuse of sophistry. Clement used the word *gnostic* because he wanted to show that there was a true Christian *gnosis,* or knowledge, which developed out of faith and which was better than the boasted knowledge of the heretical Gnostics. Gnosticism was especially strong in Alexandria. Clement put forward an attractive alternative to it and attacked what he considered to be its peculiar tenets of esoteric knowledge, dualism, and ethical determinism. Knowledge, he said, grows out of faith and is not distinct from it. There is one God who made all things. Men are free to choose the way they will go.

Clement wrote against the background of Middle Platonism, of Antiochus of Ascalon, Maximus of Tyre, Albinus, and Numenius, whose thought was governed by the problem of defining the relation between the one and the one-many, and of deriving the latter from the former. The difference between a one and a one-many, or between simple and complex unity, is like the difference between the unity of a pinpoint and the unity of a spider's web. In Middle Platonism these two unities were developed into divine entities. Simple unity is divine and transcendent, while complex unity is divine and immanent. Clement was influenced by the Alexandrian Jewish Platonist Philo, for whom God is a simple, bare unity and the Logos an all-embracing cosmic whole. Clement's thought is governed by the pattern of simple and complex unity; and his accounts of God, goodness, and truth are expressed in these terms.

God is the transcendent one, a simple unity, the ultimate first principle and cause of all things. The categories of logic cannot be applied to him. "Nor are any parts to be ascribed to him, for the one is indivisible." God cannot properly be named. The good names we give him are supports to our minds to stop us from erring. Taken separately, these names do not say what God is like, but together they show his power. While God cannot be known, the Son, or the Logos, is wisdom, knowledge, and truth. He unites in himself the world of Platonic forms, or "powers," as they are also called in later Platonism. "The Son is not simply one thing as one thing nor many things as parts, but one thing as all things. All things come from him. For he is the circle of all the powers rolled into one and united." Within this unity of the Son the individual believer is saved. Faith is union in him, while disbelief is separation, estrangement, and division. Paganism is wrong because it multiplies the nature of divinity, and Marcion, the Christian heretic, is wrong because he divides the supreme God from the Creator of the world, making two Gods instead of one.

God's goodness is perfect and unique. God does not prevent evil and suffering from taking place, but when they do, he turns them to good account. He may use suffering as a form of correction for sinners. After death, imperfect souls may be sanctified by an intelligent nonmaterial fire. The complex goodness of men is always assimilation to God—growing like him by participation in his goodness. Clement constantly refers to Plato's statement in the *Theaetetus* concerning assimilation to God. All men, says Clement, receive the image of God at their birth and all may then, as they choose, become assimilated to him and receive his likeness. In the *Paedagogus,*

Clement gives detailed instruction for Christian behavior. From Plato came the emphasis on self-knowledge, and the conception of evil as ignorance and virtue as knowledge. Virtue comes through discipline and the pursuit of goodness, without thought of ulterior gain. The harmony of the soul is aided by the harmony of the body. From Aristotle, Clement draws the notion of virtue as the fulfillment of man's function and the achievement of his end. This fulfillment is found in pursuing the mean between extremes and in possessing right reason. Clement draws heavily on Stoic ethics, commending what is in accord with nature and in harmony with reason. There is a class of things intermediate between good and evil. One should recognize the things that are in one's power and the things which are not, and avoid being dominated by one's irrational passions.

Clement speaks of truth in two ways. The simple elements of Christianity are true, and heresy is to be rejected as false. Truth is one and unique, powerful and strong in delivering men from error. It comes from God and is preserved within the tradition of the church. Second, Clement speaks of truth as including all that is consistent with basic Christian truth. This truth is a whole composed of many parts. It is one body from which each of the philosophical sects has torn a limb, or part, falsely imagining it to be the whole of truth. The many parts must be brought together, so that the perfect Logos, the truth, may be known. The truth of philosophy was partial, but real. It was for the Greeks, as the Law was for the Jews, a schoolmaster to bring them to Christ. Clement shared with others the quaint notion that the Greeks stole their ideas from the Hebrews.

Faith is an act not a process. Faith is the acceptance from God of an indemonstrable first principle from which all other truth may be deduced. It is a judgment of the soul, an Epicurean preconception, and a Stoic assent. Knowledge (*gnosis*) is both logical and spiritual, joining things together either by logical reasoning or by spiritual vision. The eighth book of the *Stromateis* is a notebook of logic composed of materials from various sources. It deals with demonstration and definition in an Aristotelian way, gives a Stoic refutation of the skeptical suspension of judgment (that is, if one must suspend judgment, then one should suspend judgment concerning suspense of judgment), and treats of cause, using both Stoic and Aristotelian terms. Causes may be original, sufficient, cooperating, and necessary. Spiritual knowledge is growth in Christ, awareness of God's universal presence, and union with him in love. Symbolism reveals hidden connections and points to unity. Knowledge is always a complex unity, while faith is a simple unity.

Clement achieved the first real synthesis of classical philosophy and Christianity. The Apologists had used particular ideas to bridge the gap between philosophy and Christianity. In Justin's writings, for example, God is described in terms of the Platonic ineffable being, and the divine reason implanted in men is expounded along Stoic lines; but there is no comprehensive conceptual framework that enables these and other ideas to modify one another. Clement's synthesis was developed by Origen, and the result was the theology of the fourth-century Greek Fathers and of Augustine.

See also Patristic Philosophy; Platonism and the Platonic Tradition.

Bibliography

TEXTS AND TRANSLATIONS

Stählin, Otto, ed. *Clemens Alexandrinus. Die griechischen christlichen Schriftsteller der ersten drei Jahrhunderte.* Vol. 39.2. Leipzig: Hinrichs, 1936. Vols. 12, 15, 17, 39.1, revised by Ludwig Früchtel and Ursula Treu. Berlin: Akademie-Verlag, 1970–1985.

By various hands, with parallel French translation: *Sources Chrétiennes,* Vols. 2b (*Protrepticus*), 23 (extracts from Theodotus), 30, 38 (*Stromata* 1–2), 70, 108, 158 (*Pedagogue*), 278–279, 428, 446, 463 (*Stromata* 4–7). Paris: Éditions du Cerf, 1965–2004.

Wilson, William. In *The Ante-Nicene Fathers: Translations of the Writings of the Fathers down to A.D. 325,* Vol. 2, edited by Alexander Roberts and James Donaldson, revised by A. Cleveland Coxe. Grand Rapids, MI: Eerdmans, 1951.

FURTHER READING

Behr, John. *Asceticism and Anthropology in Irenaeus and Clement of Alexandria.* Oxford: Oxford University Press, 2000.

Clark, Elizabeth A. *Clement's Use of Aristotle. The Aristotelian Contribution to Clement of Alexandria's Refutation of Gnosticism = Texts and Studies in Religion* 1. New York and Toronto: Edwin Mellen, 1977.

Lilla, Salvatore R. C. *Clement of Alexandria: A Study in Christian Platonism and Gnosticism.* London: Oxford University Press, 1971.

Osborn, Eric F. *The Philosophy of Clement of Alexandria.* Cambridge, U.K.: Cambridge University Press, 1957.

E. F. Osborn (1967)

Bibliography updated by G. R. Boys-Stones (2005)

CLIFFORD, WILLIAM KINGDON

(1845–1879)

A English mathematician and philosopher, William Kingdon Clifford was born in Exeter, the son of a justice of the peace. At the age of fifteen he went to Kings College, London. There he gained a minor scholarship to Trinity College, Cambridge, to which he went in 1863. He began to exhibit powers of originality in mathematics, publishing a number of mathematical papers during the year in which he first entered Cambridge.

At the university Clifford distinguished himself not only by his intellect but also by his singular character. As one of the most prominent undergraduates, he was soon invited to join the Apostles, an exclusive Cambridge club made up of the twelve most distinguished undergraduates of the time. Here he exhibited some of that breadth of learning and clarity of mind for which he was to be noted all his life. It appears that he was highly concerned about religious questions because he studied Thomas Aquinas and learnedly supported the Catholic position. Later, however, he became an agnostic and turned against religion; Herbert Spencer and Charles Darwin became the most important influences upon his thinking in many areas.

Clifford was elected a fellow of Trinity in 1868. In that year he began the practice of giving public lectures, a source from which most of his published work stems. He participated in a scientific expedition, which was wrecked off the coast of Catania, Sicily. In 1870 he was appointed professor of applied mathematics at University College, London. Soon after, he became a member of the most distinguished intellectual society of the day, the Metaphysical Society, as well as of the London Mathematical Society. Tragically, his life was drawing to a close, for he had contracted tuberculosis. His condition worsened, until by 1878 it was evident that the disease was far advanced. In 1879 he traveled south to try to counteract the disease, but he died on March 3 of that year.

During Clifford's lifetime he published only a textbook on dynamics and some scattered technical and nontechnical papers based on his lectures. It remained for a number of his friends to gather together his work. H. J. S. Smith edited the mathematical papers, F. Pollock the philosophical ones. The young Karl Pearson edited and completed his popular work on science, *The Common Sense of the Exact Sciences.*

SCIENTIFIC EPISTEMOLOGY

Clifford's philosophical views must be placed within the context of several major influences upon his thought: the Kantian frame in epistemology, the Riemannian frame in geometry, and the Darwinian frame in biology. On the basis of these and other influences, Clifford constructed a scientific epistemology and attempted to construct a scientific metaphysic. A discussion of his epistemology is first in order, since out of it grew his metaphysics. Clifford conceived of knowledge as a biological response to the world. Its structure, therefore, is determined by that adjustment. Nevertheless, any analysis of knowledge as such reveals that within it the form and the content of knowledge are distinguishable from each other. Immanuel Kant believed that he had determined a method to make this distinction in all cases. Clifford, taking his cue from Kant, believed that he too could make this distinction, but in a way that took into account the ultimately biological character of knowledge. He thought that an analysis of the foundations of science, and in particular of the axioms of geometry, would reveal that these axioms are forms of experience in the life of any particular individual. Thus, since the biological adaptation of the race has crystallized three-dimensional Euclidean space, this spatial framework has become the one in which individuals see spatially locatable objects. Clifford went even further in this direction by claiming that such a construction is ultimately a growth of experience which has been transformed into neural capacities. Thus, Clifford conceived of the form-content distinction of knowledge as one relative to the biological development of the race. What is at one time the content of experience is later, through a biological process, transformed into a form of experience.

The principles of geometry and arithmetic serve, for individuals, to structure their experience. They are or correspond to ways in which our sense data are "spatially" or "numerically" organized. Their logical status is therefore closely akin to the one that Kant assigned to them. They are a priori, for no experience is capable of verifying or falsifying them, whereas at the same time they are synthetic, since the predicate term is not contained in the subject term.

Within this framework of thought it is intelligible to discuss Clifford's concrete epistemological ideas. He offered analyses of what might be called (1) perceptual statements, (2) geometric, arithmetical, and even physical principles, and (3) belief statements in general.

PERCEPTUAL STATEMENTS. In various essays Clifford offered an analysis of perceptual statements concerning objects, persons, and the spatial aspects of objects and persons. In general, he refused to admit a phenomenalist analysis of such statements. In all cases some ideal conception, be it of "an eject" (a technical term that will be explained later) or of "a form of experience"—in other words, a conception which is not itself definable in terms of a set of sense experiences—enters into the meaning of the statement, either explicitly or implicitly. This is true with the qualification that Clifford sometimes suggested that statements about physical objects are reducible to statements about sense experiences.

GEOMETRIC, ARITHMETICAL, AND PHYSICAL PRINCIPLES. The analysis that Clifford provided of the several kinds of statements differed somewhat from one another, and it would be wise to examine them in sequence. As has already been indicated, the statements of geometry and arithmetic state universal and therefore formal characteristics of experience. In the case of geometric statements, Clifford asserted that they are universally true about the objects of our perceptions, in the sense that all perceptions of spatial relationships must conform to them. Furthermore, they are necessary, since the perceptions compatible with the negations of such statements are impossible. Clifford contended that Kant had established the necessary properties of space by a subjective method, a method of introspection, whereas Clifford attempted to demonstrate such properties by a consideration of the neurological bases of perception. The limits of what is perceptible, given man's neurological structure, were, for Clifford, what is known a priori to the individual, while those perceptions whose contradictions are not imperceivable, again given man's neurological structure, are known a posteriori. Clifford proceeded to demonstrate, to his satisfaction, that at this level of analysis both Euclidean and non-Euclidean space are compatible with the neurological structure of perception, and that it is a matter of the general explicatory value of a geometric theory as to which of the various geometries is to be accepted. Of course, man's neurological structure evolves over time, so that what is necessary at one time is not necessary at another—this indicates that Clifford used the term *necessary,* in this context, in a relative sense.

Clifford's analysis of arithmetical statements differs somewhat from his analysis of geometric statements. He thought that their validity depended upon several factors: (1) the tautological character of certain parts of language, (2) the acceptance of a general principle of the uniformity of nature of the kind that J. S. Mill suggested, and (3) the acceptance of an analysis of arithmetical operations in terms of the physical operation of counting. Numerals are assigned in a one-to-one correspondence with standard sets of objects, each set containing one member more than the preceding set. The operations of addition, multiplication, and, by implication, subtraction and division are next defined in terms of the physical operations of juxtaposition of sets of objects. Clifford then claimed that if the meaning of "distinct objects" were granted, along with the assumption that all objects maintain their identity through space and time (the uniformity of nature), then the laws of arithmetic can be seen to hold for all objects. On the basis of the natural numbers, he sketched the development of the more complex number systems.

Clifford did not have much to say about the status of physical laws and theories, except to suggest that there are some principles of physics that are, like the principles of geometry and arithmetic, rules for the ordering of sense impressions.

BELIEF STATEMENTS. Clifford's examination of the basis of belief in the natural sciences led him to a more general analysis of belief. Indeed, it was this general analysis of belief and the agnostic and antireligious conclusion to which it led that occasioned great opposition on the part of William James and others. Clifford claimed that no statement is worthy of belief unless all the possible evidence points to the truth of the statement. He recognized that in practice it is impossible to have available all the possible evidence about the truth or falsity of a proposition. Failure of memory, the expenses of collecting information, and a host of other factors contribute to this impossibility. But he claimed that an acceptance of the principle that similar causes have similar effects (another version of the principle of the uniformity of nature) permits our acceptance of many beliefs in cases where the standard of all possible evidence is not met. Such a principle permits an inductive inference from known facts to unknown ones, and thus permits us to make up for evidence we do not possess. These ideas are contained in his essay "The Ethics of Belief," to which James's famous essay "The Will to Believe" is a reply. In that essay James claimed that a belief is worthy of acceptance in some cases where there is no empirical evidence either for or against the content of the belief. And this criterion permitted James to believe in the existence of God.

SCIENTIFIC METAPHYSICS

Clifford's epistemological views were the occasion for his speculative metaphysical ideas. He had been wrestling with the problem of whether the existent world is wholly phenomenal in character or whether there are entities of a nonphenomenal character which go to make it up. In earlier essays—for example, "The Philosophy of Pure Sciences"—he inclined toward a purely phenomenalist view, but in his more mature and well-known essay "On the Nature of Things-in-Themselves" he reversed his former stand. Not all existence is phenomenal in character. He was clear, for example, that the ego cannot be analyzed in purely phenomenal terms. Clifford thus postulated the existence of what he termed "ejects" as well as of phenomenal "objects." An eject is distinguished from an object in the following way: An object can be an object of *my* consciousness, an eject is something *outside* my consciousness. Thus, another's ego (and this holds for all persons) is an eject; it is never in my consciousness. Clifford postulated that there are nonpersonal as well as personal entities that are ejects. The elements of ejects are themselves what Clifford called feelings. They are constituents of everything, he claimed, since the fact that there is a continuity of forms in nature gives assurance that, at least to some degree, any entity in nature possesses the same qualities that all others have. Since feelings are elements of consciousness, all entities therefore have this aspect of consciousness to a certain extent, although it is only to more complex entities that we ascribe consciousness. The elementary entities that are called "feelings" were considered by Clifford to be absolute existents and therefore things-in-themselves. Clifford then named these elementary entities mind-stuff, since they participate somehow in the character of the mental. Their necessarily incomplete representation in the mind of man is what is known as the material world. There exists a complex mirroring relation—indeed, Clifford used the image of two reflecting mirrors—between the external world and its representation in knowledge. Thus, Clifford's speculative metaphysic ultimately postulated a Spinozistic world in which the mental and physical are really two different ways of looking at the same world. Another possible interpretation of his thought is that all existence is ultimately infused with a psychic aspect, that is, that panpsychism is the most correct view of reality.

In conclusion, it is worthwhile mentioning several areas of thought in which Clifford was ahead of his time:

(1) Clifford recognized the fact that scientific laws are always "practically inexact." By this notion he wished to point out that a scientific law is never exactly confirmed by the evidence for it but rather is confirmed within the limits of experimental error. A law is accepted on the basis of experimental evidence even if that experimental evidence does not exactly coincide with what, on the basis of deductions from the law, one might expect to be confirming evidence. This is so simply because all measurement of evidence in modern scientific practice involves taking into account errors of measurement, and such errors of measurement must be "factored out" before a definite conclusion is reached as to the relevance of the evidence.

(2) Clifford, in the brief note "On the Space-Theory of Matter," declared himself to be in the geometric tradition that holds that the determination of the truth or falsity of geometrical axioms is empirical. Clifford saw that through a change in the basic assumptions of microgeometry (geometry of the infinitesimally small) he could work out a system of geometry and physics that would clear up the anomalies in physical theory that existed in his day. He saw that a reformulation of microgeometry in non-Euclidean terms could achieve this result, and in this respect he anticipated, at least in part, Albert Einstein's program. He never, however, carried through this program on his own; he merely suggested that such a program was feasible.

(3) Clifford showed the possibility, at least in principle, of constructing a wholly empirical geometry in the following special sense: Geometry is considered to be a set of statements about the relations between geometrical objects such as points, lines, planes, and volumes. These geometrical objects and relations, however, are themselves characterized in a completely empirical way, not as ideal objects, as they are usually characterized in most treatments of geometry. That is, they are identified with the physical objects or aspects of physical objects. The principles of geometry are then empirical statements whose truth or falsity is a matter of observation. This point of view is close to a geometric operationalism. Clifford's account of it is found in his book *The Common Sense of the Exact Sciences.*

See also Darwin, Charles Robert; Einstein, Albert; Epistemology; Geometry; James, William; Kant, Immanuel; Mill, John Stuart; Pearson, Karl; Perception; Thomas Aquinas, St.

Bibliography

For works by Clifford, see *The Common Sense of the Exact Sciences,* edited by Karl Pearson (New York: Appleton, 1885), reissued with a preface by Bertrand Russell and a new introduction by James R. Newman (New York: Dover, 1955);

Lectures and Essays, edited by F. Pollock, 2 vols. (London: Macmillan, 1879); and *Mathematical Papers,* edited by H. J. Smith (London: Macmillan, 1882).

For literature on Clifford, see George S. Fullerton, *A System of Metaphysics* (New York: Macmillan, 1904), which contains a criticism of Clifford's theory of mind-stuff; and A. S. Eddington, *The Nature of the Physical World* (Cambridge, U.K.: Cambridge University Press, 1929).

Howard E. Smokler (1967)

COCKBURN, CATHARINE TROTTER
(1679?–1749)

Catharine Trotter, according to her editor and biographer, was born on August 16, 1679, the younger of two daughters of David Trotter, a captain in the Royal Navy and his wife, Sarah Ballenden, of a well-connected Scottish family. Trotter's father died of the plague while on a voyage that was to have made his fortune. Instead, his family was forced to survive on an irregularly disbursed pension from the reigning monarch. Trotter was educated at home, and perhaps largely self-educated, although she seems to have taught herself French and Latin. She was a precocious writer, publishing a novella at a young age, followed by poems, and ultimately five plays, four appearing between 1695 and 1701 and the last in 1706, all of which achieved a certain renown.

In 1701 Trotter began to live with her married sister in Salisbury, where she remained until her own marriage in 1708. In Salisbury Trotter joined the circle surrounding Gilbert Burnet (1643–1715), the bishop of Salisbury, which included his wife, the devotional writer Elizabeth Burnet (1661–1709), and his cousin, Thomas Burnet of Kemnay (1656–1729), a lively correspondent of Trotter's. It was during her time in Salisbury that Trotter's theological and philosophical interests began to manifest themselves. In 1702 she published *A Defense of Mr. Locke's Essay*, a reply to criticisms of John Locke by yet another Thomas Burnet (c. 1635–1715), and in 1707 *A Discourse concerning a Guide in Controversies*, the fruits of her struggles with Roman Catholicism, justifying her decision to return to the Church of England.

In 1708, she married the clergyman Patrick Cockburn (1678–1749), and her scholarly interests were for some time suspended while she struggled to raise a family of four in somewhat reduced circumstances, brought on when her husband lost his curacy, on finding himself unable to swear the oath of abjuration on the ascension of George I (1660–1727) to the throne. In 1726 Cockburn was able to reconcile himself with this oath and became first the rector at Aberdeen and then the vicar of Long Worsley in Northumberland, where the family was still living at the time of Catherine's death in 1749.

With the restoration of the family fortunes, Cockburn's philosophical interests also revived, and in 1726 she published *A Letter to Dr. Holdsworth*. While she did not resume publishing until close to the end of her life, first bringing out *Remarks upon Some Writers in the Controversy concerning the Foundation of Moral Virtue and Moral Obligation* in 1743 and then *Remarks upon the Principles and Reasonings of Dr. Rutherforth's Essay* in 1747, it is clear from letters written throughout this period, particularly those to a niece, Anne Arbuthnot, that Cockburn maintained a lively reading program and developed her intellectual interests in correspondence.

Cockburn's works were collected by Thomas Birch and published after her death, and include, in addition to her published philosophical work, several hitherto unpublished pieces, a play, and a fascinating collection of letters. Some doubts have been raised about the dates of Cockburn's life supplied by Birch, stimulated by a letter to G. Burnet written in 1707, in which she reports the marriage of a son and the birth of a grandchild. Since, according to Birch's reckoning, this would make Trotter a mere twenty-seven, it has been suggested that her birth date should be pushed back to accommodate the birth of a son and grandson. There are some limits, however, on the extent to which Trotter's age in 1706 can be adjusted, since according to Birch's account, she was seventy-one at the time or her death and was publishing close to that time. An alternative possibility is that Trotter was not in fact the birthmother of the son she mentions casually to Birch.

Each of Cockburn's works takes roughly the same form, that of a loosely organized commentary on some other work, often itself critical in nature. Her earlier work defends Locke against various attacks, and her later work is written in defense of Samuel Clarke.

Her presentation then can appear somewhat diffuse and disorganized. In her early defense of Locke against Thomas Burnet, for example, she considers three different criticisms: that Locke's rejection of innate moral principles leaves him with no resources on which to ground one's knowledge of moral principles, that Locke provides no way in which he can establish God's veracity, and, finally, that an account of personal identity based like Locke's on consciousness instead of substance does not provide grounds for personal immortality.

There are, however, some common threads that tend to reappear in much of her work. In particular, Cockburn is very much embroiled in eighteenth-century attempts to walk a middle ground between deism and voluntarism. Her concern is to argue that human beings can, by means of their intellectual resources, derive an understanding of moral concepts based on their nature as sensitive, rational, and social beings. It is through this complex understanding of ourselves that we are able to work out what is suitable or fit for us. Cockburn argues that the complexity of our nature does not limit our grasp of what is fit for us simply to what is pleasant or what is in our self-interest, but that we can derive a full sense of our moral obligation from our nature as rational, social beings. Therefore, there is no need to turn to an otherwise unmotivated appeal to God's decrees to account for the full range of our moral obligations. Cockburn also wants to maintain that, while our understanding of the nature of these obligations rests on our understanding of ourselves, it is nevertheless God's decrees that give these principles the force of law. But since we know that God is good, and we understand, from our own case, what it is to be good, we also know that God would not require of us actions that are, as we understand it, irrational. Her position is designed to guard against both the view that human morality is entirely independent of religion and against the view that our obligations have only a religious and no rational support.

See also Locke, John.

Bibliography

WORKS BY COCKBURN

The Works of Mrs. Catharine Cockburn, Theological, Moral, Dramatic and Poetical. Several of Them Now First Printed. Revised and Published, with an Account of the Life of the Author. 2 vols. Compiled by Thomas Birch. London: J. and P. Knapton, 1751.

WORKS ABOUT COCKBURN

Atherton, Margaret. "Women Philosophers in Early Modern England." In *A Companion to Early Modern Philosophy*, edited by Steven Nadler, 404–422. Malden, MA: Blackwell, 2002

Bolton, Martha Brandt. "Some Aspects of the Philosophical Work of Catharine Trotter." *Journal of the History of Philosophy* 31 (4) (October 1993): 565–578.

Broad, Jacqueline. "Catharine Trotter Cockburn." In *Women Philosophers of the Seventeenth Century*, 141–165. New York: Cambridge University Press, 2002.

Nuovo, Victor. "Catharine Cockburn." In *The Dictionary of Seventeenth-Century British Philosophers*, edited by Andrew Pyle. Bristol, U.K.: Thoemmes Press, 2000.

Waite, Mary Ellen. "Catharine Trotter Cockburn." In *A History of Women Philosophers. Vol. 3, Modern Women Philosophers, 1600–1900*, edited by Mary Ellen Waite, 101–125. Dordrecht, Netherlands: Kluwer Academic, 1991.

Margaret Atherton (2005)

CODE, LORRAINE
(1937–)

Lorraine Code is a Canadian philosopher with interests in epistemology, feminist epistemology, and the politics of knowledge. She is Distinguished Research Professor of Philosophy at York University, where she is also appointed to the Graduate Programs in Women's Studies and Social and Political Thought. Code has authored five books and numerous articles, and has edited five collections. From 1999 to 2001, she served as a Canadian Council Research Fellow.

Code describes her work as an interrogation of local and global politics of knowledge that harm people and nature. She argues, for example, that traditional philosophical epistemologies foster the exploitation of people and nature by Western sciences and institutions because they include tenets that obscure the role of social and political relations in the formation of knowledge. Recently Code has undertaken the constructive project of developing an alternative, "ecologically modeled" epistemology that, she maintains, avoids the failings of traditional epistemologies.

Code's critical and constructive projects consistently focus on the ethical dimensions of knowledge making and epistemological accounts of it. In *Epistemic Responsibility* (1987) she argues that epistemic responsibility is not exhausted by "purely epistemological" standards. Code contends that an emphasis in epistemology on virtue and responsibility would result in attention to the social contexts of knowing, including the relevance of social relations and social roles to what is recognized as knowledge. Such analyses would, in turn lead, to more robust notions of epistemic responsibility.

A concern with the ethical implications of epistemology is also central in *What Can She Know?* (1991). Here Code focuses on the "alignments" in "mainstream epistemology": on one hand, characteristics its values (for example, objectivity and rationality) and, on the other hand, shifting conceptions of masculinity. Code argues that these alignments contribute to institutional knowledge (for example, in the sciences and law) and to social institutions that undermine women's abilities to act as

knowers while rendering invisible the politics of gender at work. She uses these alignments in a more general argument that subjective factors inform all knowledge claims and epistemic ideals. From this perspective, theories of knowledge that obscure the role of such factors are not just factually flawed, but they are also ethically flawed because they underwrite the continuation of a form of subjectivity that, although changing overtime, has consistently put women at a disadvantage. Not surprisingly, when Code poses the question in this work of whether a distinctly feminist epistemology is desirable, she is not enthusiastic. She holds that efforts to achieve universality, which she here attributes to epistemology in general, are at odds with the attention to particularity, context, and other aspects of subjectivity that her arguments call for.

In *Rhetorical Spaces: Essays on Gendered Locations* (1995), Code undertakes the kind of fine-grained studies she recommends. Her essays explore cases in which specific and rhetorically and socially constructed locations—including those of marginalization and power—have an impact on who is deemed credible and what counts as knowledge. In one, a victim of sexual harassment seeks to reconcile her memories with conflicting accounts and to understand how trusting herself relies in part on her credibility in the eyes of others. Other essays, focusing on institutionalized knowledge such as health care, explore ways in which everyday knowledge practices are sites of social interactions that contribute to or deny credibility to various subjects and groups.

In *Ecological Thinking* (2005) and elsewhere, Code builds from her earlier work to advance a sustained argument for what she calls "an ecologically modeled" theory of knowledge. Code maintains that explanatory models in ecology are promising for a theory of knowledge precisely because they assume a mutual dependency of organisms, an interrelatedness between their well-being and features of their environment, including features that are cruel. Code argues that such models contrast sharply with the individualism and instrumentalist conceptions of rationality that characterize traditional epistemology and obscure the ethics and politics of knowledge-making practices. Incorporating these noninstrumentalist ideas into epistemology, she maintains, would result in a model that could accommodate the insights of feminist, multicultural, and postcolonial studies into precisely those dimensions of knowledge making that have been obscured by traditional epistemology.

See also Epistemology; Feminist Philosophy; Feminist Epistemology.

Bibliography

BOOKS BY CODE

Epistemic Responsibility, Dartmouth: University Press of New England, 1987.

What Can She Know: Feminist Theory and the Construction of Knowledge. Ithaca, NY: Cornell University Press, 1991.

Rhetorical Spaces: Essays on Gendered Locations. New York and London: Routledge, 1995.

Ecological Thinking. Oxford: Oxford University Press, 2005.

Feminist Philosophy: Routledge Contemporary Introductions to Philosophy. New York and London, Routledge, 2005.

EDITED COLLECTIONS

Feminist Perspectives: Philosophical Essays on Method and Morals (with Sheila Mullett et al). Toronto: University of Toronto Press, 1988.

Changing Methods: Feminists Transforming Practices (with Sandra D. Burt). Orchard Park: Broadview Press, 1995.

Encyclopedia of Feminist Theories. New York and London: Routledge, 2000.

Feminist Interpretations of Hans-Georg Gadamer. University Park: Penn State University Press, 2002.

The Sex of Knowing (with Michelle DeDoeuff et al). New York and London: Routledge, 2003.

OTHER WORKS BY CODE

"What Is Natural about Epistemology Naturalized?" *American Philosophical Quarterly* 1 (1996): 1–22.

"Feminists and Pragmatists: A Radical Future?" *Radical Philosophy* 87 (1998): 22–30.

"How to Think Globally: Stretching the Limits of Imagination." *Hypatia: A Journal of Feminist Philosophy* 2 (1998): 73–85.

"Flourishing." *Ethics and the Environment* 1 (1999): 63–72.

Lynn Hankinson Nelson (2005)

COGNITIVE SCIENCE

Cognitive science is the interdisciplinary study of mind, in which the concepts and methods of artificial intelligence (AI) are central (Boden forthcoming). The most prominent disciplines within the field are AI, artificial life (A-life), psychology, linguistics, computational neuroscience, and philosophy—especially the philosophy of mind and language. Cognitive anthropology is included too, though often goes unseen under the label of evolutionary psychology.

The many relevant subfields include robotics, whether classical, situated, or evolutionary; studies of enactive vision, where the organism's own movements (of eyes and/or body) provide crucial information for acting in the world; the psychology of human-computer interaction, including various aspects of virtual reality such as avatars; and computational theories of literature, art,

music, and scientific discovery. Nonhuman minds are studied by computational ethology and neuroethology, and by A-life.

WHO IS A COGNITIVE SCIENTIST?

Not everyone working in the key disciplines is a cognitive scientist. Only those taking a computational approach to questions about mind are considered cognitive scientists.

Some AI workers, for example, are not cognitive scientists because they have no theoretical interest in human thought. Their aim is to challenge their ingenuity as computer engineers by getting a program or robot to do a task that people either cannot do or do not want to do. If hunches, or experimental evidence, about human psychology can help them achieve that goal, that is fine. But if nonhuman tricks are available, such as looking ahead in a chess game to consider all the legal possibilities for several moves, they will use them. These computer scientists are engaged in technological AI, not psychological AI. Only the latter is a proper part of cognitive science.

Even someone who does have a professional interest in human minds need not be a cognitive scientist. For instance, many social psychologists study patterns of interpersonal behavior without asking about the information processes that underlie them and make them possible. Even some cognitive psychologists insist that they are not cognitive scientists, because they follow James Gibson's (1979) affordance theory of perception—which allows for information pickup but not for information processing. (Their self-description is based on an overly narrow view of what cognitive science covers: Gibsonian insights have become prominent in various areas of cognitive science, such as enactive vision.)

Similarly, many linguists—sociolinguists and historical philologists, for instance—are not primarily concerned with just how language is generated and/or understood. But even those who do focus on these computational matters do not all agree. Chomskian linguistics, for example, was crucial in the rise of cognitive science and has deeply influenced the philosophy of mind; but non-Chomskian accounts of syntax have been developed since. In addition, theories of pragmatics have become at least as prominent as theories of syntax—and pragmatics is an aspect of situatedness, a concept of growing importance within cognitive science as a whole. As for anthropology, most anthropologists see their field as a hermeneutic enterprise, not a scientific one. They reject psychological explanations of culture in general, and computational accounts in particular.

COGNITIVE SCIENCE IS ABOUT MORE THAN COGNITION

It includes cognitive psychology, of course: the study of language, memory, perception, problem solving, and creative thinking. What is more, most research has focused on individual human adult cognition. However, other aspects of mind are studied too: motivation, emotion, choice (Dennett 1984), development, psychopathology, interpersonal phenomena, motor control, and animal psychology.

Consider emotion, for example. The role of emotion in problem solving, attitude formation, and neurosis were topics of research in AI and computational psychology in the early 1960s. But the problems were too difficult, and were largely dropped. Interest revived later, partly because of neuroscientific work on emotional intelligence and partly because of advances in the computational theory of scheduling in multigoal systems (Sloman 1993). Interdisciplinary conferences on the psychology, neuroscience, computer modeling, and philosophy of emotion blossomed at the turn of the century, when the topic became a prominent aspect of research.

Whether the focus of attention is on development or psychopathology, emotion or motor control, the prime interest for cognitive science is in the abstractly defined computational functions that generate the behavior concerned. But the neural mechanisms that implement them are often studied too. Despite the functionalist doctrine of multiple realizability, many cognitive scientists want to know how psychological functions are actually implemented in the brain. When functionalism began in the 1960s, little attention was paid to the nervous system by philosophers or AI scientists. Since the 1980s, that has been less true.

Indeed, neuroscience as such has become increasingly concerned with computational questions. On the one hand, there are theories (and computer models) of specific neural circuits doing closely specified things. For instance, cells in the retina and/or visual cortex that compute particular visual features, such as light gradients or surface textures; or cells in the female cricket's brain that enable her to discriminate the song of male crickets of the same species, and to move accordingly. On the other hand, there are broad-brush theories about the computational functions carried out by large areas of the brain, where the focus is less on specific individual cells than on general neuroanatomy: the different cell types, locations, and connections of the neurons.

DEVELOPMENTAL ISSUES

Most cognitive scientists study already established phenomena, although many include learning in their subject matter. Some, however, study—and model—mental development. And some do this because they believe that adult psychology cannot be properly understood without knowing how it developed. In short, they see the mind as an epigenetic system, deeply informed by its developmental history.

Epigenesis was stressed long ago by Conrad Waddington in biology and Jean Piaget in psychology. It is self-organized development, grounded in innate predispositions in continual dialectic interaction with the (internal and external) environment. For example, there are inborn dispositions to attend to broadly facelike stimuli, or to human speech-sounds. Once the attention is caught, learning can help develop the infant's pattern recognition and discriminatory powers. In some cases, such as face recognition, the neural mechanisms relevant at different stages have been largely identified.

An epigenetic view is not strictly environmentalist, nor strictly nativist either. Rather, it stresses the dialectical interplay between the two. Late twentieth-century work in developmental neuroscience and developmental psychology has therefore led to a radical reconceptualization of nativism (Elman et al. 1996). Some philosophers of biology have defined new accounts of self-organization and dynamical development accordingly (Oyama 1985).

WHAT IT MEANS TO SAY THAT COGNITIVE SCIENCE IS COMPUTATIONAL

Cognitive science employs computational models of mind in two senses.

First, the substantive concepts in its theories are computational. The mind is seen as some sort of computational system (just what sort is hotly disputed), and mental structure and mental processes are described accordingly (Haugeland 1997). So whereas many psychologists (and other scientists) use computers to express/clarify their theories, and especially to manipulate their experimental data, only cognitive scientists import computational ideas into their theories.

Second, computer modeling is often used to clarify and test computational theories of mind. Often, but not always, some work in cognitive science (in AI and psychology, not just in philosophy) employs computational concepts and insights, but with insufficient detail to allow programs to be written. When programming is possible, it provides several advantages. Even program failures can be scientifically illuminating, pointing out lacunae or mistakes in the theory, or fundamental limitations of the methodology being used. However, successes may be even more instructive. For if a program—or a robot—produces a given performance, one knows that it suffices to do so.

Whether real minds (or brains) use similar processes to produce equivalent performance is another matter: just because a program does something in a certain way, it does not follow that people do too. This question can be answered only by empirical evidence. Sometimes, a programmed theory models not only psychological phenomena at various levels, but also the details of their underlying neural base. In such cases, validation requires both psychological and neuroscientific evidence.

The references to computational ideas in the previous two paragraphs cover concepts rooted in two different intellectual traditions, namely, cybernetics and Turing computation. These were closely linked in the years when cognitive science began.

A seminal paper by Warren McCulloch and Walter Pitts (1943) prompted early work both in neural nets and in what is sometimes called GOFAI, or "Good Old-Fashioned AI." (It also influenced the design of the von Neumann computer.) McCulloch and Pitts integrated three key ideas of the early twentieth century: propositional logic, neuron theory, and Turing computation. They proved that anything expressible in the propositional calculus is computable, and can be mapped onto some specifiable neural net. In addition, they suggested that a fourth key idea—feedback, the core concept of cybernetics—could be defined in terms of these networks, in which case purpose and learning could be embodied in them too.

A few years later they published another paper, in which they argued that probabilistic networks are more like brains and can do things naturally that logic-based systems cannot (Pitts and McCulloch 1947). They still insisted, nevertheless, that their original, logical, account was correct in principle.

In short, the concept of computational systems is normally used within the field to cover both GOFAI and connectionism. (Some philosophers, however, restrict it to the former.) Cognitive science includes both.

Sometimes, the reason why a computational theory is not actually modeled is that suitable computer technology does not yet exist. By the same token, many advances in cognitive science have depended partly upon advances in computing technology. These include both increases in

size (computing power) and new types of virtual machine, embodying forms of computation that were not possible previously.

In some cases, the core ideas had already been defined long before the technology was available to test/explore them. Parallel distributed processing, for instance, was envisaged over twenty years before computers became powerful enough for it to be implemented in interesting ways. Similarly form-generating interactive diffusion equations and cellular automata were both first defined in the 1950s, but not extensively studied until the advent of large machines and computer graphics in the late 1980s. And genetic algorithms, glimpsed in the 1950s and defined in the late 1960s, were first implemented in the 1980s. Once the technology was available, further questions arose that had not been posed before.

SOME PHILOSOPHICAL PROBLEMS

Many philosophical disputes arise within cognitive science. One dispute concerns the relative merits of the two AI approaches mentioned above: classical (symbolic) AI and connectionism, or neural networks. The latter is broadly inspired by the basic structure of the brain. (Some recent work in artificial neural networks tries to take more account of the subtleties of real neurons; even so, these models are hugely oversimplified in comparison with the real thing.) There are several types of neural networks, but the one most widely used within cognitive science—and the one of greatest interest to philosophers—is parallel distributed processing, or PDP.

Some researchers champion only one of these AI approaches, whereas others admit both because of their complementary strengths and weaknesses. Symbolic AI, or GOFAI, is better for modeling behaviors that involve hierarchical structure, advance planning, deliberation, and/or strict sequential order. The conscious, deliberative aspects of the mind are best suited to this approach. Connectionism, by contrast, is better for modeling the tacit learning and knowledge involved in pattern recognition, including the fuzzy family resemblances between instances of one and the same concept.

It does not follow that all unconscious mental processes are best modeled by PDP systems. Some psychoneural theories of action errors, including various clinical syndromes, employ hybrid (mixed) models in which the hierarchical aspects represent both conscious and unconscious processing.

INTERNAL REPRESENTATION. Another debate concerns the nature and importance of various kinds of internal representation. Connectionist representations are different from GOFAI ones, and several philosophers have argued that they are closer to the neural representations that embody concepts (Churchland 1989; Clark 1989, 1993; Cussins 1990). Computational neuroscience has described further types of representation. One example is emulator systems, which are neural mechanisms whose physical dynamics mimic the temporal changes being represented. Another, based on the anatomy of the cerebellum, is a way of representing motor behavior that is based neither on logic (GOFAI) nor on statistics (PDP), but on noneuclidean tensor geometry.

Some philosophers follow the AI community and/or the neuroscientists, in accepting that representations may take many different guises, depending on the role they have evolved to play. Others, however, argue that only formal-symbolic structures, expressed in a *language of thought*, are properly termed representations, and that only these can generate human conceptual/linguistic thought (Fodor 2000).

NATURE OF COMPUTATION. The nature of computation is a third topic of controversy (Scheutz 2002). Most philosophers define it as Alan Turing did in the 1930s—and his is still the only really clear definition. However, practicing AI scientists think of computation in a number of different ways, based on virtual machines whose properties are different from those of a Turing machine. Moreover, some people are trying to go beyond Turing computation by defining new forms of computers (hypercomputers), some—but not all—of which involve quantum computing. Some of these may turn out to be relevant to human brains, but others will not.

MEANING IN THE REAL WORLD. A fourth area of philosophical discussion focuses on whether—and if so, how—meaning (intentionality) can be grounded in the real world—and whether it can properly be attributed to programs and/or robots. Evolutionary theories of intentionality rule out GOFAI programs (as do many philosophers), but—arguably—allow meaning to be ascribed to some evolved robots. The grounding problem, on this view, is solved by the way in which the relevant mechanism has evolved in situated, embodied systems.

Empirical work that is closely related to the problem of intentionality includes research on *theory of mind.* Very young children are unable to grasp that each person is an agent with their own set of beliefs and interests, which may differ from those of the child. So although the child realizes that adults (and even other children) know many

things that they do not, the child does not appreciate that someone else may believe something to be true that the child knows to be false. (This is why infants do not lie: they cannot conceive of doing so.) Normally, theory of mind develops spontaneously at around ages four or five, although in autistic children it apparently does not. In other words, inbuilt predispositions have evolved that lead the young child first to engage with other humans (maintaining eye contact, pointing to direct attention, turn-taking in communication, etc.), and eventually to attribute intentional states to them. Philosophers have asked (for instance) whether they do this by theorizing about other people's minds or by simulating, or empathizing with, them (Davies and Stone 1995).

CONSCIOUSNESS IN COMPUTATIONAL TERMS. A fifth philosophical puzzle concerns whether consciousness could be explained in computational terms—or in any other scientific, naturalistic, manner (Heil 2004, Newell 1980, Searle 1993). Research in various disciplines within cognitive science has shown that there is no such thing as the problem of consciousness; rather, there are many problems of consciousness, because the term is used to make many different distinctions. Some of these are much better understood than they were twenty years ago, thanks to computational work in AI, psychology, and neuroscience. Reflective self-consciousness, for example, and the bizarre dissociations of consciousness typical of multiple personality, are intelligible in terms of recursive processing, guiding procedures, and access limitations within complex hierarchical structures for perception, memory, and action.

Considerable controversy, however, still attends the problem of qualia. Some cognitive scientists argue that *qualia* can be analyzed in terms of complex dispositions for making discriminatory computations (Dennett 1991). Others see them as aspects of an irreducible informational feature of the universe, applying not only to human brains but to atoms as well (Chalmers 1996). Still others make further suggestions, including several based on quantum physics. In short, there are many theories of *qualia*, and no agreement about what a successful theory might look like.

OPPOSITION TO ORTHODOX COGNITIVE SCIENCE. A sixth controversy—or rather, batch of controversies—arises from recent work that opposes orthodox (neo-Cartesian) cognitive science (Cliff, Harvey, and Husbands 1993; Port and van Gelder 1995; Wheeler 2005). This involves both empirical theory/modeling and philosophical discussion. In general, it draws on the traditions of phenomenological philosophy and/or autopoietic biology, rather than Cartesianism. It rejects both symbolic and connectionist AI, and the concept of representation. It highlights embodied systems (not abstract simulations), embedded in their environment and responding directly to it. Examples include situated robotics in AI, dynamical systems theory, ecological psychology, and A-life studies of evolution and coevolution.

Philosophies inspired by these empirical researches include the theory of extended mind (Clark 1997). This starts from the position that minds must necessarily be embodied and that memory storage lies largely outside the skull (ideas familiar within phenomenology and GOFAI, respectively) and goes on to argue that an individual person's mind is extended over the surrounding cultural artifacts: language, customs, and material objects—from palaces to pencils. The claim is that mind is not merely deeply influenced by these things, but it is largely constituted by them.

Philosophical questions associated with A-life include whether evolution is a necessary characteristic of life, and whether the concept of autopoiesis captures the essence of life (Bedau 1996; Maturana and Varela 1987). If living things are defined as autopoietic systems—whose physical unity, boundaries, and self-maintenance are attained by self-organized metabolic processes—then questions about the origins of life take on a different color, as do questions about the possibility of strong A-life (life in computer memory)—so called by analogy to strong AI.

Philosophers of A-life consider not only the nature of life as such, but how and why it is related to mind. Must all minds be evolved, for example? Autopoietic theorists define all life as involving cognition, while insisting that only linguistic life (i.e., adult humans) involves representations. But questions remain about whether, and if so why, life really is essential for mind. By the same token, questions remain about whether the study of A-life is essentially unrelated to cognitive science or fundamental to it.

CULTURE AND COGNITIVE SCIENCE. Finally, culture-directed research in cognitive science raises philosophical questions too. One concerns the nature of group mind, or as it is more commonly called, distributed cognition (Hutchins 1995). Can one identify aspects of cognition that cannot be attributed to any single individual, but only to a team of enculturated persons acting in concert—and if so, can one model such phenomena in computers? Two more such questions concern the evolution

of information-processing mechanisms that underlie important cultural phenomena—religion or aesthetic appreciation, for example—and the evolution of culture as such.

See also Computationalism; Neuroscience; Psychology.

Bibliography

Boden, Margaret A. *The Creative Mind: Myths and Mechanisms.* 2nd ed. London: Routledge, 2004.

Boden, Margaret A. *Mind as Machine: A History of Cognitive Science*. Oxford: Oxford University Press, forthcoming.

Bedau, Mark A. "The Nature of Life." In *The Philosophy of Artificial Life*, edited by Margaret A. Boden. Oxford: Oxford University Press, 1996.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press, 1996.

Churchland, Paul M. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press, 1989.

Clark, Andy J. *Associative Engines: Connectionism, Concepts, and Representational Change*. Cambridge, MA: MIT Press, 1993.

Clark, Andy J. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press, 1997.

Clark, Andy J. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1989.

Cliff, David; Harvey, Inman; and Husbands, Philip. "Explorations in Evolutionary Robotics." *Adaptive Behavior* 2 (1993), 73–110.

Cussins, Adrian. "The Connectionist Construction of Concepts." In *The Philosophy of Artificial Intelligence*, edited by Margaret A. Boden. Oxford: Oxford University Press, 1990.

Davies, Martin, and Tony Stone, eds. *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell, 1995.

Dennett, Daniel C. *Consciousness Explained*. Boston: Little Brown, 1991.

Dennett, Daniel C. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press, 1984.

Dennett, Daniel C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, et al. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press, 1996.

Fodor, Jerry A. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press, 2000.

Gibson, James J. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979.

Haugeland, John, ed. *Mind Design II: Philosophy, Psychology, Artificical Intelligence.* 2nd ed. Cambridge, MA.: MIT Press, 1997.

Heil, John, ed. *Philosophy of Mind: A Guide and Anthology*. Oxford: Oxford University Press, 2004. See especially 3–6, and 9–10.

Hutchins, Edwin L. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.

McCulloch, Warren S., and Walter H. Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." In *The Philosophy of Artificial Intelligence,* edited by Margaret A. Boden. Oxford: Oxford University Press, 1990. Essay first published in 1943.

Maturana, Humberto R., and Francisco J. Varela. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: New Science Library, 1987.

Newell, Allen. "Physical Symbol Systems." Cognitive Science 4 (1980): 135–183.

Oyama, Susan. *The Ontogeny of Information: Developmental Systems and Evolution*. Cambridge, U.K.: Cambridge University Press, 1985.

Pitts, Walter H., and Warren S. McCulloch. "How We Know Universals: The Perception of Auditory and Visual Forms." In *Embodiments of Mind*, edited by Seymour Papert. Cambridge, MA: MIT Press, 1965. Essay first published in 1947.

Port, Robert F., and Timothy J. van Gelder, eds. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, 1995.

Scheutz, Matthias, ed. *Computationalism: New Directions*. Cambridge, MA: MIT Press, 2002.

Searle, John R. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press, 1993.

Sloman, Aaron. "The Mind as a Control System." In *Philosophy and the Cognitive Sciences*, edited by Christopher Hookway and Donald Peterson. Cambridge, U.K.: Cambridge University Press, 1993.

Wheeler, Michael W. *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press, 2005.

Margaret A. Boden (1996, 2005)

COHEN, HERMANN

(1842–1918)

Hermann Cohen, a neo-Kantian philosopher, was born at Coswig, Anhalt, Germany. His father, Gerson Cohen, was a teacher and precentor at the synagogue; his mother was Friederike née Salomon. In 1878 Hermann married Martha Lewandowski, the daughter of Professor Louis Lewandowski, who was also a precentor at the synagogue and a composer of Jewish ritual songs. In 1853 Hermann went to the gymnasium of Dessau, which he attended for some years. He left there prematurely and went to the Jewish Theological Seminary at Breslau. Later, as a student at the University of Breslau, he wrote the essay "Über die Psychologie des Platon und Aristoteles," which won the prize of the philosophical faculty in August 1863. On August 5, 1864, he took the bachelor's examination as an extramural pupil at the Breslau Matthias Gymnasium. In the fall of the same year he went for further university studies to Berlin. He wrote an essay, "Philosophorum de Antinomia Necessitatis et Contigentiae Doctrinae" and

entered it for a university prize. Since the prize was not awarded to him, he submitted the work (somewhat altered) to the philosophical faculty at Halle. On the basis of this work he was awarded the doctorate of philosophy by this faculty on October 27, 1865.

On his return to Berlin he published several studies, some of them in *Zeitschrift für Völkerpsychologie und Sprachwissenschaft.* Heymann Steinthal, the coeditor of this periodical, who was warmly interested in the very gifted young man, had stimulated his interest in social psychology. It was not until 1870 that his publications disclosed a special interest of their author in Kantian philosophy. In that year Cohen intervened in the Homeric struggle that had broken out between Adolf Trendelenburg and Kuno Fischer over Trendelenburg's criticism of the Kantian transcendental aesthetic. Trendelenburg agreed with Immanuel Kant that the concepts of space and time are a priori, but he denied their exclusion from things-in-themselves, which was, in Kant's opinion, an unavoidable consequence of their intuitive apriority. According to Trendelenburg, a third possibility was left, namely the validity of space and time with regard to all existing objects *in spite of* the apriority of their *concepts.* Fischer, defending Kant against the charge of leaving this "gap," insisted that Kant's assignment of both space and time to human sensibility, in the transcendental aesthetic, was irrefutable. Cohen, a pupil of Trendelenburg, but not a favorite one, in an essay published in the above periodical (7 [3]: 239–296) gave the Solomonic judgment. Trendelenburg was right in criticizing Fischer, but wrong in criticizing Kant.

PHILOSOPHICAL TEACHING

This judgment already contained in germ the whole of Cohen's future philosophical achievement. In the following year his first philosophical book, *Kants Theorie der Erfahrung* (Berlin, 1871) made it clear why, in his opinion, both Trendelenburg and Fischer were wrong. The teaching of the transcendental aesthetic, which showed space and time to be forms of our sensibility, had to be complemented by the teaching of the transcendental logic, where these forms are shown to be a priori conditions of possible experience. Possible experience, as Kant said throughout the *Critique,* is the only object of a priori knowledge. Therefore, the exclusive subjectivity of space and time, assumed by both parties to be Kant's complete view, disappears entirely if one takes into account the methodological difference between a psychological classification of space and time among native ideas and the Kantian transcendental theory of their being the a priori conditions of the possibility of experience.

By thus extending the matter in question to the whole of Kant's theory of a priori knowledge, Cohen gave evidence of the philosophical turn of his gifts. In 1873 he presented to the philosophical faculty of Marburg a treatise titled *Die systematischen Begriffe in Kants vorkritischen Schriften* (Berlin, 1873) with an application for the *venia legendi* (lectureship). On the recommendation of F. A. Lange, Cohen's application was accepted. Lange died two years later, and in January 1876 Cohen, proposed by the faculty, was appointed to the vacant chair. He devoted his work to the fortification and extension of his new interpretation of Kant, which from the beginning had aroused admiration for the author's energy and devotion, though many doubted the compatibility of Cohen's interpretation with Kant's real opinion.

In any case, Cohen found himself confronted with a serious problem. If the objectivity of space and time consisted in their being a priori conditions of the possibility of experience, the question remained from what principle experience itself derived its validity. There was no identity between the conditions of experience and the conditions of things-in-themselves. This was unquestionably Kant's teaching. But, as Cohen observed, Kant had a new concept of experience. Actually, the innovation—if there was one—was David Hume's, not Kant's. Experience, according to Hume, is a statement on matters of fact presupposing some connection of these matters by general rules. The difference between Kant and Hume is not in the concept of experience but in the question of whether it is possible to justify the universality of that intellectual presupposition with regard to the objects of sense perception. Hume claimed it is not possible; those a priori assumptions are not a matter of intelligence at all. Man is driven to them by the laws of nature, which make him believe automatically in the possibility of experience.

This might not be a satisfactory answer. But Cohen's solution to the question—to derive the objectivity of those presuppositions (including space and time) from their being a priori conditions of experience—was not only not satisfactory—it was no answer at all. It was an answer that answered by what was the subject of the question. If, therefore, Cohen wished neither to accept the unconditioned subjectivity of Kant's possibility of experience nor to fall back on Hume's skepticism—which way was left to him?

It was the way of a cryptopositivism. The objectivity of doubtful a priori assumptions, such as space, time, and the categories, was demonstrable, according to Cohen, by

means of the "fact of science" (*das Faktum der Wissenschaft*). Surely it was a historical fact that Isaac Newton had used these assumptions as principles in establishing his mathematical theory of the phenomena of nature. It was also a fact that Newton was far from justifying the assumption of these principles by deriving them from experience. But this by no means made the fact of their use as nonempirical principles of natural science equivalent to the fact of an existing a priori knowledge of nature. It was, on the contrary, evident that none of Newton's mathematical laws of natural phenomena, formulated in differential equations, could be called a knowledge of those phenomena if it was not verifiable by experience. How, then, could those principles presupposed by Newton's physics assume the character of a priori requirements for the cognition of nature by the mere fact of being presupposed by Newton, if the cognitive character of these presuppositions with regard to natural phenomena was demonstrable only by experience?

Despite this unanswerable question, Cohen boldly proclaimed that Newtonian science demonstrated by its own historical facticity the possibility of an a priori knowledge of nature by means of the concepts of space, time, and the Kantian categories. He called the manner of this demonstration the "transcendental method." It proved to be an enormous success. Cohen's pupils vied with each other in showing that modern science would not have been possible if its promoters had not presupposed what they actually had—that is, space, time, and the principles assigned by Kant to pure understanding. This, if it was meant to be a legitimation of a priori knowledge of natural phenomena by means of those principles, was clearly a vicious circle.

The desire to escape this consequence determined Cohen's philosophical development and the fate of neo-Kantianism in general. Cohen realized eventually that his transcendental method, if it were to prove effective with regard to a priori knowledge of nature, required the tearing down of the insurmountable barrier Kant had fixed between a priori and empirical knowledge by means of his distinction between sensibility as receptivity and understanding as spontaneity. Therefore Cohen posited a kind of thinking that originated by its own act the whole field of principles of our knowledge ("Denken des Ursprungs"). Thus, all human knowledge must be in principle a priori knowledge.

In *Die Logik der reinen Erkenntnis* (Berlin, 1902) Cohen elaborated this puzzling idea. He explained by abundant historical comments that the real task of metaphysics was the thinking of the origin. If this is to be regarded as more than an utter triviality, it testifies that the author, in order to escape the deadly embrace of Hume, fled into the arms of Johann Gottlieb Fichte and G. W. F. Hegel. Once more he fell victim to the ancient illusion of being able to understand Kant better than Kant himself by dropping the conditions essential to the very problem of transcendental philosophy. Thus Cohen, however unintentionally, encouraged a new movement from Kant to Hegel in German neo-Kantianism. Even Heideggerian existentialism claimed some kinship with the critique of pure reason by proclaiming the search for the "common root" of sensibility and intelligibility, necessarily problematical with Kant, as a way of salvation from all possible transcendental problems.

PRACTICAL PHILOSOPHY

Cohen similarly interpreted Kant's moral philosophy according to the maxim that to interpret Kant one must go beyond him in his *Kants Begründung der Ethik* (Berlin, 1877). He inherited from Trendelenburg's Aristotelianism the idea of virtue as the supreme problem of moral philosophy. Combined with the Kantian assumption of an a priori principle of morals, this idea generated the problem of ethics as the problem of an a priori science of virtue. Here again Aristotle intervened by his teaching that all other virtues are implied in justice. Thus, the problem of morals presented itself to Cohen as the problem of an a priori knowledge of justice. All a priori knowledge, according to Cohen's transcendental method, required some factual science to justify it. Kant did not presuppose any such factual science in his *Critique of Practical Reason.* In this Cohen believed Kant to be mistaken. According to him, morals does have a basic science, jurisprudence, because the idea of justice is the constitutional law of this science. If there were no a priori law of justice, the sort of systematic knowledge of the laws that the Romans assigned to *iurisprudentia* would not be possible. In identifying *iurisprudentia* with *scientia iusti,* Cohen found that the a priori character of Kant's categorical imperative was justified by the factual existence of jurisprudence.

POLITICS

It is easy to observe that autonomy as conditioned by the categorical imperative is by no means the principle of a society that, like the state, is realizable under the conditions of experience. And it is no less easy to see that the positive laws of a given state, the objects of jurisprudence, in spite of the possibility of their being systematically treated by jurisprudence, do not necessarily agree with

some a priori idea of justice. Nevertheless, the idea of a human society constituted by the law of autonomy meant a quite personal engagement to Cohen, above and beyond all philosophical subtleties concerning its meaning or its justification. This engagement drove him from the field of transcendental deductions into politics. It made him a public champion of those whose personal dignity granted by the law of autonomy was infringed upon by society. He eventually found himself among them. Some years after he settled at Marburg, anti-Semitism appeared on the German political stage. The famous historian Heinrich von Treitschke published in his *Preussische Jahrbücher* (Vol. 1879, No. 11) an article in which he called attention to an attitude allegedly adopted by a good many Jewish writers, whom he accused of being antinational and anti-Christian. He held that they should respect the feelings of the majority. The weak point in Treitschke's pleas was the authority that he assigned to what in his romanticism he called Christian German culture.

Cohen in his *Eine Bekenntnis in der Judenfrage* (Berlin, 1880), without attacking Treitschke's romantic idea of a law given by Germano-Christian feeling, boldly announced that the Jews already belonged to the German nation—not *in spite of* their being Jews, but *because* they were Jews. This, of course, was too much for both parties. But to Cohen the philosopher and learned Jewish theologian it seemed quite simple to demonstrate. The Germans, he argued, are the nation of Kant. The Jews are a nation whose creed has been purified by the prophets. The teachings of the prophets, as Cohen's learnedness interpreted them, were identical with Kant's ethical idealism. Therefore, whoever tells a Jew that he can belong to the German nation only at the cost of his religion denounces him as having no true morality of his own. From that time on, Cohen continued as a collaborator in the interpretation of Jewish tradition by adapting it to his philosophy. His writings in this field were edited by Bruno Strauss and published with an introduction by Cohen's admirer Franz Rosenzweig as *Hermann Cohens jüdische Schriften* (3 vols., Berlin, 1924).

Besides the startling historical and ideological identifications of his Germano-Jewish patriotism, there was yet another reason for Cohen's reputation as a political outsider. It was not unusual to support the workingman's longing for a decent living according to the law of humanity. All the so-called *Katheder-Sozialisten,* among them some of the most influential professors of the German Empire, did it. But the mixture of philanthropy and justice that Cohen considered the supreme principle of his moral philosophy made him believe in a basic accordance between the doctrine of Karl Marx and his own. Thus, he became responsible for the legend of a kinship between Kant and Marx. This was enough to color the politician Cohen with a red tinge—and if his true patriotic German feeling separated him from Jewish orthodoxy and Zionism, his rather innocent socialism did not make him a favorite with either his government or his faculty.

Hence, his retirement in 1912 brought a great disappointment with it. The faculty, not very fond of intricate transcendental deductions that were admired by students but doubted by philosophers, refused to give his chair to the man of his choice, Ernst Cassirer. The choice of his colleagues, Paul Natorp dissenting, was a young experimental psychologist.

LATER RELIGIOUS VIEWS

Deeply hurt, Cohen left Marburg and retired to Berlin. There he devoted himself to a lectureship at the Lehranstalt für Wissenschaft des Judentums, of which he was already a member of the board of trustees. Thus, he was again a theologian. Meanwhile, his philosophy had dissolved theology into a transcendental deduction of the eternity of cultural progress governed by the "social ideal"; namely, the community of autonomous beings. But in actual fact there was no solid deduction even of this eternity. The question of whether religion had any meaning at all arose again. Cohen answered it in two books, *Der Begriff der Religion im System der Philosophie* (Giessen, 1915) and *Die Religion aus den Quellen des Judentums* (Leipzig, 1919). In both of these works the point of departure lies in the observation that the belief in the eternity of cultural progress is of little comfort to the individual in his personal sufferings. Therefore, an empty space has been left by philosophy. This space may be filled by God as a savior bringing personal consolation to all people. Cohen found this idea of the Divine Being splendidly expressed by the prophets and the Psalmist. But the mere idea of a powerful personal Helper does not cause that Helper to exist; and since this idea, according to Cohen himself, could not be justified by his philosophical system, the question of a savior's existence was left entirely to personal conviction. To the great satisfaction of his religious friends, Cohen, when he died, seemed to be in full possession of this conviction.

AESTHETICS

The manner in which Cohen addressed religious problems in his last writings was prepared by his aesthetics. Aesthetics had been treated by Kant within the frame of

what he called the critique of judgment. Cohen's comment, published under the title *Kants Begründung der Ästhetik* (Berlin, 1889), once again disclosed the author's difficulty in harmonizing his own ideas in this field with the peculiar but at bottom simple Kantian theory of aesthetic pleasure.

In spite of the stock of questions left unanswered by Cohen's principles, he continues to live in the memory of philosophers as a Kantian who dominated to a great extent the philosophical discussions of his time. But if Cohen's own interpretation was attractive, it did not make Kant attractive; and his school of neo-Kantianism eventually expired. The unbearable viciousness of the famous gnosiological circle, wrongly imputed to Kant himself but inextricably woven into Cohen's own omnipresent transcendental method, drove the younger generation to the worship of new gods. But even so, Cohen has left a stimulus to study "that Kant" whom, as one of his pupils is reputed to have said, "nobody ever knew." The feeling expressed by these words was precisely Cohen's own feeling when he began his work.

See also Aesthetics, History of; Aesthetics, Problems of; Aristotle; Ethics, History of; Fichte, Johann Gottlieb; Fischer, Kuno; Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Lange, Friedrich Albert; Marx, Karl; Neo-Kantianism; Newton, Isaac; Rosenzweig, Franz.

Bibliography

WORKS ON COHEN

Cassirer, Ernst. "Hermann Cohen." *Social Research* 10 (2) (1943): 219–232.

Kinkel, Walter. *Hermann Cohen, Einführung in sein Werk.* Stuttgart, Strecker and Schröder, 1924.

Natorp, Paul. *Hermann Cohen als Mensch, Lehrer, und Forscher.* Marburg, Germany, 1918.

Rosmarin, T. W. *Religion of Reason, Hermann Cohen's System of Religious Philosophy.* New York: Bloch, 1936.

Vuillemin, J. *L'héritage kantien et la révolution copernicienne.* Paris: Presses universitaires de France, 1954.

Julius Ebbinghaus (1967)

COHEN, HERMANN [ADDENDUM]

Philosophical research between 1960 and 2004 looked at Cohen's thought from both a historical and a theoretical viewpoint. In the age of the integration of German Jews into German society, he was the foremost advocate of the need for a meeting between the Enlightenment and Judaism.

Cohen had an important influence on various philosophical fields. Ernst Cassirer's neo-Kantian approach to human culture (1943) and J. B. Soloveitchik's neo-Kantian attitude to religion, particularly Judaism (1986), owe their method to his work. Both Husserl's and Heidegger's interpretations of Kant's transcendental philosophy, and therefore the phenomenological or existentialist concept of the *self*, derive from Cohen's theory of knowledge (Dussort 1963, Vuillemin 1954). Hans Kelsen's juridical positivism was inspired by Cohen's idea of "natural right" (Winter 1980). Franz Rosenzweig's philosophy of divine revelation—as a bond between a human being and God through "religious love"—stems from Cohen (Altmann 1970). Lastly, Cohen's essays on the history of philosophy influenced Leo Strauss's interpretations of Spinoza, Maimonides, and the relation between "Jerusalem" and "Athens" (Kajon 2002).

Cohen's logic has inspired examinations into the fundamental principles of mathematics and physics (Holzhey 1986). Unlike Hegelianism on the one hand and postmodernism on the other, Cohen's ethics sought the relation between reason and the facts of law, state, and history (Gigliotti 1989). His aesthetics invites a criticism of *art for art's sake* (Poma 1997). His philosophy of religion expresses the need for a "religion of reason" which keeps the profundity of religious life (Zac 1984).

Cohen's thought stems from both Jewish tradition and European idealism, hence its fertile, albeit problematic, character.

Bibliography

COHEN'S WORKS

Werke "Hermann-Cohen-Archiv am Philosophischem Seminar Zürich" under the direction of Helmut Holzhey. Hildesheim, Germany: Olms, 1977–.

WORKS ON COHEN

Altmann, Alexander. "Hermann Cohens Begriff der Korrelation." In *Von der mittelalterlichen zur modernen Aufklärung: Studien zur jüdischen Geistesgeschichte*, 300–317. Tübingen, Germany: Mohr, 1970.

Cassirer, Ernst. "Hermann Cohen." *Social Research* 10 (2) (1943): 219–232.

Dussort, Henry. *L'école de Marbourg*. Paris: PUF, 1963.

Gigliotti, Gianna. *"Avventure e disavventure del trascendentale." Studio su Cohen e Natorp*. Napoli, Italy: Guida, 1989.

Holzhey, Helmut. *Cohen und Natorp*. Basel-Stuttgart, Germany: Schwabe, 1986.

Kajon, Irene. *Il pensiero ebraico del Novecento: Una introduzione*. Roma, Italy: Donzelli, 2002. (English edition: *Contemporary Jewish Philosophy: An Introduction*. London: RoutledgeCurzon, 2005.)

Kinkel, Walter. *Hermann Cohen: Einführung in sein Werk*. Stuttgart, Germany: Strecken und Schröder, 1924.

Natorp, Paul. *Hermann Cohen als Mensch, Lehrer, und Forscher*. Marburg, Elwert: 1918.

Poma, Andrea. *The Critical Philosophy of Hermann Cohen*. Translated by John Denton. Albany: State University of New York Press, 1997.

Rosmarin, T. W. *Religion of Reason: Hermann Cohen's System of Religious Philosophy*. New York: Bloch, 1936.

Soloveitchik J. B. *The Halakhic Mind*. New York: Seth Press, 1986.

Vuillemin, J. *L'héritage kantien et la révolution copernicienne*. Paris: Presses Universitaires de France, 1954.

Winter E. *Ethik und Rechtswissenschaft: Eine historisch-systematische Untersuchung zur Ethik-Konzeption des Marburger Neukantianismus im Werke Hermann Cohens*. Berlin: Duncker & Humblot, 1980.

Zac, Sylvain. *La philosophie religieuse de Hermann Cohen*. Paris: Vrin, 1984.

Irene Kajon (2005)

COHEN, MORRIS RAPHAEL

(1880–1947)

Morris Raphael Cohen, the American naturalistic philosopher, was born in Minsk, Russia. When twelve years old, he was brought to New York City by his parents, who immigrated to America in search of greater opportunity and freedom. In his early youth he came under the influence of the Scottish freelance scholar Thomas Davidson. Cohen was graduated from the College of the City of New York (City College) in 1900 and received his PhD in philosophy from Harvard University in 1906. At Harvard he studied under Josiah Royce, William James, and Hugo Münsterberg.

From 1912 to 1938, Cohen taught philosophy at City College. He was an outstanding teacher, and some of his students became eminent teachers, philosophers, and lawyers. He was a visiting lecturer in philosophy at Johns Hopkins, Yale, Stanford, and Harvard and from 1938 through 1941 was a professor at the University of Chicago. For years he gave courses at the New School for Social Research. He was also a lecturer at the law schools of St. John's University, Columbia, Yale, Harvard, Cornell, the University of Buffalo, and New York University. Although an agnostic, he had been a dedicated Jew. His wit, his critical spirit, his erudition, and his interest in a wide range of friends made him a colorful and animating person.

Cohen's philosophic interests included the philosophy of science, metaphysics, logic, social philosophy, legal philosophy, and the philosophy of history. His contribution to legal philosophy has been especially widely recognized.

METAPHYSICAL AND LOGICAL PRINCIPLES

Cohen's general philosophic outlook is naturalistic. There is no place in his philosophy for the extranatural and no place for extrascientific methods to attain knowledge. His outlook is also rationalistic, for he assumed that rationality is inherent in nature. His philosophy is based on three principles: rationality, invariance, and polarity. These three principles, coherently interwoven, provide his view of reality.

RATIONALITY. In its long history the concept of rationality has acquired a variety of meanings. It has meant logical order, inductive generalization, and wisdom. Each of these meanings has been significant. Cohen did not offer an inclusive definition of rationality, but in his philosophy of nature the first meaning is dominant and in his ethical and legal philosophies the third meaning is central.

Rationality as logical order may be considered methodologically or ontologically. Methodologically, it is a procedure to order our objects of thought in a logical way. Most philosophers, except for mystics and irrationalists, feel the necessity of such a procedure. Yet Cohen went beyond the methodological use of rationality and insisted on its ontological status. The rules of logic and pure mathematics "may be viewed not only as the principle of inference applicable to all systems but also as descriptive of certain abstract invariant relations which constitute an objective order characteristic of any subject matter" (*Reason and Nature*, p. 142).

For Cohen, as a logical realist, the formal aspects of logic apply to everything. As against idealists, positivists, and pragmatists, he was firm in insisting that the rational order is independent of human or superhuman mind. Idealists, according to him, deny the objectivity of logical order by giving it only a psychological status, but the psychological description of reasoning as a mental event cannot determine, according to him, whether a given logical argument is valid. Positivists, his arch philosophic enemies, fall short in a similar way. As sensations are considered the only deliverance of the external world, for

positivists logical connections are mere fictions. Pragmatists, he argued, similarly depreciate the status of rational order. In their attempt to interpret the truth of judgment in terms of practical consequences, they consider logical relations as merely practical tools of thought without any ontological standing.

However, Cohen admitted an element of contingency in nature. "By no amount of reasoning," he wrote, "can we altogether eliminate all contingency from our world" (ibid., p. 82). The universe is ultimately what it is, and contingency cannot be eliminated. And by contingency Cohen meant that the world contains an irrational element in the sense that "all form is the form of something which cannot be reduced to form alone" (*Studies in Philosophy and Science*, p. 11).

INVARIANCE. Science is not, as Cohen rightly pointed out, a mere observation of particular facts; it is never satisfied with stating only what has occurred. The aim of science is to determine the universal, invariant relations of particular events. To say that sulfur has melted at 125°C. is a mere statement of fact similar to the statement that Russians for generations have used the Cyrillic alphabet, but to say that sulfur always melts at 125°C. means that if ever anything conforms to the category of sulfur, it melts at this temperature. The second statement expresses not only a historical event but also an invariant relation that belongs to "the eternal present."

Although the essence of particular things is their invariant relations, our knowledge of these is only probable. Only in logic or in mathematics can we attain certainty; in the world of facts our knowledge is only probable, for we cannot prove that the opposite of a given factual statement is absolutely impossible.

POLARITY. According to the principle of polarity, opposites involve each other. As Cohen expressed it in *Reason and Nature*, "Opposites such as immediacy and mediation, unity and plurality, the fixed and the flux, substance and function, ideal and real, actual and possible, and so on, like the north (positive) and the south (negative) poles of a magnet, all involve each other when applied to any significant entity" (p. 165).

In addition to its methodological value as a guide to the clarification of ideas, the principle of polarity, like the principle of rationality, has ontological status. Empirical facts, such as the existence of the north and south poles, are said to be resultants of opposing tendencies. Cohen generalized this alleged fact as the principle of "the necessary copresence and mutual dependence of opposite determinations."

ETHICS

Historically, there have been two major opposing theories of morality—the absolutist and the relativist. Cohen examined both of these theories and found them unsatisfactory. The absolutist is too rigid and uncritical; the relativist is too chaotic, without guiding principles. Cohen thought the principle of polarity could reconcile the two opposing views. Actually, these two views provide a vantage point for arriving at the truth. Concretely, every issue of life involves choice. The absolutist is right "in insisting that every such choice logically involves a principle of decision," and the relativist is right "in insisting on the primacy of the feeling or perception of the demands in the actual case before us" (ibid., p. 438). We may thus have an ethical system that is rigorously logical and at the same time richly empirical. Such an ethics must be grounded in what human beings desire and believe, and yet its primary condition must be the logical analysis of judgment as to what constitutes right and wrong, good and evil—an ethic that is the rational formulation of our ends.

LAW

Cohen was a pioneer in introducing legal philosophy as a significant study to universities and law schools. As Leonora Cohen Rosenfield wrote, "His philosophical treatment of the law in relation to man and the social order may prove in time to be his foremost influence."

For Cohen law is essentially a system for the orderly regulation of social action. Jurisprudence must avoid the extremes of positivism and formalism. "Law without concepts or rational ideas, law that is not logical is like prescientific medicine—a hodge-podge of sense and superstition," yet law without reference to the actual facts of human conduct would be empty. A law is both stable and dynamic; it is a balance between prevailing customs and the emerging demands of society. Cohen was especially critical of what he called the "phonograph theory of law," the theory that the judge arrives at his decision in a mechanical way, according to unchanging laws. Cohen effectively argued that the judge's opinions on social and economic questions deeply influence his decisions. One of the chief merits of his analysis of law is his insistence on the interdependence of the factual and the normative. As he maintained, "Justice and the law, the ideal and the actual are inseparable, yet identifiable."

See also Idealism; James, William; Philosophy of Law, History of; Positivism; Pragmatism; Rationalism; Realism; Royce, Josiah.

Bibliography

WORKS BY COHEN

Reason and Nature: An Essay on the Meaning of Scientific Method. New York: Harcourt Brace, 1931. Cohen's major work.

Law and the Social Order: Essays in Legal Philosophy. New York: Harcourt Brace, 1933. Important work.

An Introduction to Logic and Scientific Method. New York: Harcourt Brace, 1934. Written with Ernest Nagel.

A Preface to Logic. New York: Holt, 1945. Important work.

The Meaning of Human History. La Salle, IL: Open Court, 1947. Important work.

A Dreamer's Journey. Boston: Beacon Press, 1949. Autobiography.

Studies in Philosophy and Science. New York: Holt, 1949.

Reason and Law: Studies in Juristic Philosophy. Glencoe, IL: Free Press, 1950.

The Faith of a Liberal: Selected Essays Freeport, NY: Books for Libraries Press, 1970.

WORKS ON COHEN

Baron, Salo W., Ernest Nagel, and Koppel S. Pinson, eds. *Freedom and Reason: Studies in Philosophy and Jewish Culture in Memory of Morris Raphael Cohen*. Glencoe, IL: Free Press, 1951. See Part I.

Cairns, Huntington. "The Legal Philosophy of Morris R. Cohen." *Vanderbilt Law Review* 14 (1960).

Kuhn, M. A. *Morris Raphael Cohen: A Bibliography*. New York: City College of New York Library, 1957. Available in the City College library.

Rosenfield, Leonora Cohen. *A Portrait of a Philosopher: Morris R. Cohen*. New York: Harcourt Brace, 1962.

Yervant H. Krikorian (1967)
Bibliography updated by Michael Farmer (2005)

COHERENCE THEORY OF TRUTH

The coherence theory is one of the two traditional theories of truth, the other being the correspondence theory. The coherence theory is characteristic of the great rationalist system-building metaphysicians Gottfried Wilhelm Leibniz, Benedict (Baruch) de Spinoza, G. W. F. Hegel, and Francis Herbert Bradley; but it has also had a vogue with several members of the logical positivist school, notably Otto Neurath and Carl Gustav Hempel, who were much influenced by the systems of pure mathematics and theoretical physics. According to the coherence theory, to say that a statement (usually called a judgment) is true or false is to say that it coheres or fails to cohere with a system of other statements; that it is a member of a system whose elements are related to each other by ties of logical implication as the elements in a system of pure mathematics are related. Many proponents of the theory hold, indeed, that each member of the system implies every other member. To test whether a statement is true is to test it for coherence with a system of statements. The system with which all true statements must cohere is said by its logical positivist supporters to be that accepted by the scientists of the contemporary culture. The metaphysical supporters of coherence, on the other hand, insist that a statement cannot properly be called true unless it fits into the one comprehensive account of the universe or reality, which itself forms a coherent system. In either case, no statement can be known to be true until it is known to cohere with every other statement of the system; where the system consists of all true statements, such knowledge is unattainable.

It is not altogether possible to give a plausible exposition of the theory independently of its close historical links with rationalist and idealist metaphysics, but the account might go something like this.

In practice, we sometimes reject as false an ordinary person's assertions—for instance, that he saw a ghost—or even a scientist's results—for instance, in experiments on extrasensory perception—on the ground that they do not cohere with the other commonsense or scientific views that we also hold as true.

MEANING OF TRUTH

In the exact and reputable science of pure mathematics, the logical test for the truth or acceptability of any proposition is whether it coheres with some of the other propositions, and ultimately with the axioms, of its system. In this test, which is not merely a practical one, for a proposition to cohere with other propositions is for it to be logically deducible from them. Further, this coherence is what we mean by calling such a proposition true.

INTERNAL RELATIONS. It is characteristic of the parts of a logical system like that of pure mathematics that no part would be what it is if its relations to the other parts were different from what they are. Thus, 2 would not be the number we associate with the numeral 2 if it were the third of 4 instead of the half of 4 or the cube root of 27 instead of the cube root of 8. Hence, it is said, the meaning and the truth of, for instance, "2 + 2 = 4" are bound up with the meaning and the truth of all the other statements in the arithmetical system; and our knowledge of

its meaning and its truth is bound up with our knowledge of their meaning and their truth. This principle that nothing would be what it is if its relations to other things were different—which is called the doctrine of internal relations—holds, say the metaphysical supporters of coherence, for every element, whether in thought or in reality. For example, they argue that we would not even understand, much less know the truth or falsity of, a statement about something blue if blue were "divorced in our thought from all the colours in the spectrum to which it is related by likeness and difference, all the shades within its own range, and all the definition it possesses in virtue of being thought as a quality rather than as a substance or a relation" (Brand Blanshard, *The Nature of Thought*, Vol. II, p. 316). Further, not only would we not know the meaning or truth of such a statement, but it also cannot properly be said to have its meaning or truth-value independently of its relations to other statements. The statement "Caesar crossed the Rubicon in 49 BCE" is said to be pregnant with a meaning "owing to the concrete political situation within which it took place" that it would not otherwise have.

DEGREES OF TRUTH. A corollary of the principle of internal relations and of the coherence theory in general is the doctrine of degrees of truth. If the truth of any given statement is bound up with, and can only be seen with, the truth of all the statements of the system and thus is bound up with the whole system, it is argued that individual statements as such are only partly true—and, therefore, partly false—while only the whole system is wholly true. "Truth," said Bradley, "must exhibit the mark of expansion and all-inclusiveness."

CRITERION OF TRUTH

Coherence theorists might admit that their arguments hitherto have been drawn from the nature of the a priori reasoning typical of mathematics and metaphysics; but some have also claimed that an examination of the a posteriori reasoning of the empirical sciences and ordinary life also supports the theory, not only as giving the meaning of "truth" but also as giving the test of truth (ibid., pp. 226–237). In testing for truth it is obvious, runs the claim, that coherence is our only criterion when dealing with statements about the past. No one can now compare the statement that the battle of Hastings was fought in 1066 with anything else than other statements, such as those that occur in documents, history books, or works of art. However, we can contrast with this a statement about something present, such as "There is a cat on the mat." If asked how you would test this, your reply might be "I would look and see. If what I saw corresponded to what was asserted, I would call the judgment true." However, you are assuming that "there is some solid chunk of fact, directly presented to sense and beyond all question, to which thought must adjust itself" (ibid., p. 228). What you take and use as a fact is really "another judgement or set of judgements, and what provides the verification is the coherence between the initial judgement and these" (ibid.). Consider how much of your previous experience and education, how great an exercise of your powers of conceptualization, has gone into your perception of the cat on the mat; how much, in a word, your supposed perception of fact is really a judgment, since, without a stock of judgments, what is seen could never be identified as a cat and a mat, respectively. Your test of the truth of the judgment that there is a cat on the mat or your comparison with what was there turns out to be a comparison of the original judgment with another judgment. This example, in addition, shows not only that coherence is the test or criterion of truth, but also that it gives the meaning of "truth," for it shows that the truth of the tested judgment consists in its coherence with other judgments and not with something other than a judgment.

ASSUMPTIONS OF THE THEORY

The arguments used by supporters of the coherence theory rest on various assumptions about meaning, fact, thought, and judgment that are linked partly with the impression made on them by the a priori reasoning of mathematics and logic and partly with their theory of knowledge.

A PRIORI AS PARADIGM OF TRUTH. Metaphysics is traditionally nonempirical; its conclusions are a priori deductions from certain tenets, such as George Berkeley's "To be is to be perceived" or Zeno's analysis of infinity. The conceptual statements typical of philosophy—such as that no one can know what is false, that no one can know what has not yet been proved, or that no one can know what is going to be—are true or false because of logical relations between such concepts as knowledge, truth, proof, and the future. Further, ever since Plato, mathematics has been the metaphysician's ideal; Leibniz's system was based on certain principles that he held to characterize logic and mathematics, and Spinoza's famous book on ethics is subtitled "proved in geometrical order." Some of the logical positivists, because of their training in mathematics and theoretical physics, sought to establish all knowledge as a vast system of logically interrelated statements expressed in the language of physics. In such systems, the criterion of truth is indeed

the coherence of the statement under consideration with at least some other members of the system.

Criticism. Coherence of a statement with other members of the system is not sufficient to prove the coherence theory of truth. First, the a priori statements typical of pure mathematics, unlike the empirical statements of science and everyday life, serve not to give information about characteristics of objects in the world but to show the various conclusions that can be derived from a given set of axioms and a given set of rules for operating on them. It is no objection to the truth of a given mathematical statement that there are or may be other systems with whose members it does not cohere or that it is a member of a system with no application to the world.

However, it is an objection to coherence as the meaning of "truth" or as the only criterion of truth that it is logically possible to have two different but equally comprehensive sets of coherent statements between which there would be, in the coherence theory, no way to decide which was the set of true statements. To reject a particular empirical statement like "He saw a ghost" because it conflicts with the body of our beliefs is not to assimilate the judgments of everyday life to those of mathematics, since this rejection, unlike the analogous one in mathematics, is made only because we think the body of our everyday beliefs has already been shown to be true of the world. Coherence of one judgment with another is accepted as a practical test of truth only because the second judgment is independently accepted as true.

Metaphysical supporters of the coherence theory distinguish their comprehensive system from particular systems such as those of mathematics by linking it to experience by means of their theory of knowledge, which assimilates what is thought, what is experienced, and what is. This appeal to experience and reality is indeed an inconsistency in the metaphysical version of the coherence theory, but it is more sensible than the position of the logical positivist supporters of the theory, who, in the name of consistency, allow that mutually incompatible but internally coherent systems of statements differ not in truth but only in the historical fact that our contemporaries have adopted one of the systems.

Second, there is in the a priori statements typical of mathematics and philosophy a close connection between meaning and truth. Such statements as "Twice two is half of eight" or "What is known cannot be false" are true in virtue of the meanings of the words that express them; it is because the meanings of the words are internally related as they are that these statements are true. It is not because of the relations between the meanings of "knowledge" and "breakfast," however, that it is true that no one knows what Pompey had for breakfast on the day he was murdered, nor is it because of the relations between the meanings of "two" and "four" that it is true that I made two mistakes on page four of my typescript.

Third, even within mathematics coherence gives the criterion, not the meaning, of truth. Mathematical statements are true in virtue of the criterion of coherence with each other, whereas it would seem that empirical statements are true in virtue of the criterion of correspondence with the nature of the world. However, to say that either kind of statement is true is to say that what it asserts is a fact. Whether "*X* is *Y*" is a mathematical or an empirical statement, if "*X* is *Y*" is true, then it is a fact that *X* is *Y*.

Fourth, even when confined to mathematics, the coherence doctrine of degrees of truth does not seem tenable. The fact that a given statement in mathematics is not true unless it coheres with some (or even all) other statements in the system does not imply that it is not itself wholly true; it could at most imply that it does not give the whole truth.

Ambiguities in degrees of truth. It is worth pointing out here how the theory of degrees of truth depends for its plausibility and its air of paradox on various ambiguities. There are at least three different ways in which we may qualify truth. First, we commonly ask how true something is, meaning how much truth is there in it, and commonly reply that it is partly, entirely, or perfectly true. For example, [in 1967] the report that African-Americans in the southern U.S. have been deprived of their right to vote might be said to be not quite true, either on the supposed grounds that they have been denied the opportunity to exercise their right rather than been deprived of it or that, although there has been a deprivation of the right, it is women who have been deprived.

Second, instead of asking how much truth there is in something, we may quite differently ask how much of *the* truth there is in it. To ask how much truth there is in something is to ask how much of what is not true is included; to ask how much of *the* truth there is in something is to ask how much of what is true is not included. A particular statement could be perfectly true without containing more than a minute proportion of the whole truth. Being wholly true is not the same as being the whole truth, nor is being partly true the same as being part of the truth. What is only partly true is necessarily partly false, but what is part of the truth may be entirely true.

Third, we can, in the case of general statements like "Water boils at 100° C," ask how far or under what conditions is it true. It may, for example, be true of water at sea level but not at high altitudes.

When coherence theorists say that every statement is only partly true, they usually seem to mean that every statement is only part of the truth, since nothing but the whole system of statements can give the whole of the truth. What they mean, therefore, is quite correct but wrongly expressed, because they have confused the first and the second of the above qualifications of truth. A typically ambiguous assertion is Blanshard's remark that "the trueness of a proposition is indistinguishable from the amount of truth it contains." At other times, as in their discussion of mathematical statements, by "degrees of truth" they mean "true in certain conditions." Thus, the statement "2 + 2 = 4" is said to be only partly true, as it is true in pure mathematics but not necessarily in all applied fields. Here again, what is meant is correct enough—not that such statements are not perfectly true, but that they are not universally true. The main reason, however, for the coherence theorists' belief in degrees of truth is based on a mistaken deduction from their doctrine of internal relations. Because each statement is, according to this doctrine, logically connected with other statements, it follows both that the truth of each statement is dependent on the truth of other statements and that our knowledge of its truth depends on our knowledge of the truth of these other statements. What appears to be true might turn out to be false when its further connections become known. Hence, it is said, "a given judgement is true in the degree to which its content could maintain itself in the light of a completed system of knowledge." This conclusion, however, is mistaken. A statement can be perfectly true in itself even though it would not have been true unless it had been connected in certain ways with other true statements; and it can be perfectly true whether we know this or not.

EPISTEMOLOGICAL ASSUMPTIONS. The second main influence in the usual defense of the coherence theory—that of a particular theory of knowledge—can be seen most prominently in the argument for transforming the commonsense belief that a statement (or judgment) is true if and only if it corresponds to facts into the doctrine that the judgment is true if and only if it coheres with another judgment or set of judgments. The first move in this transformation is from (*a*) "'There is a cat on the mat' is true if and only if it corresponds to the fact that there is a cat on the mat" to (*b*) "'There is a cat on the mat' is true if and only if it corresponds to the situation described as 'There is a cat on the mat.'" This is an illegitimate move, however, since a fact is not a situation, an event, or an object; otherwise we would have to postulate negative and conditional situations, events, and objects, to be described by such statements as "It is a fact that no one has yet succeeded in doing this" and "It is a fact that anyone who did succeed would be munificently rewarded." Hence, even if the moves designed to show that the situations, events, and objects we discover are not independent of our method of discovering them were valid, they would not show that facts are not independent of our methods of discovering them.

The second move in the transformation is from (*b*) "'There is a cat on the mat' is true if and only if it corresponds to the situation, event, or object describable as 'There is a cat on the mat'" to (*c*) "'There is a cat on the mat' is true if and only if it corresponds to what is verified to be a cat on the mat." This is illegitimate, however, since (*b*) is an explanation, although a false one, of the meaning of "true," whereas (*c*) contains the reason why someone might hold that there is a cat on the mat. Something can be true without anyone's knowing it to be true, although, of course, no one would sincerely say it was true unless he thought he knew it was. Idealist supporters of the coherence theory, like Bradley, move easily from (*b*) to (*c*) because they tend to identify reality with experience and knowledge, what is with what is experienced or with what is known. Further, they move distractingly to and fro between assertions about truth and assertions about *the* truth (the whole truth, the ultimate truth, a part of the truth), from assertions about the notion of truth to assertions about that which actually happens to be true. Thus, they speak of the identity of reality and truth when they mean the identity of reality and *the* truth, that is, what is true.

The third move in the transformation is from (*c*) "'There is a cat on the mat' is true if and only if it corresponds to what is verified to be a cat on the mat" to (*d*) "'There is a cat on the mat' is true if and only if it corresponds to a verification, or an experience, that would be expressed in the judgment (or, in logical positivist language, "the observation statement") 'I see (or there is) a cat on the mat.'" Because of this move they rule out the correspondence theory as a test of the truth of statements about the past, since there can be no verifying experience about what happened in the past. This move, too, is illegitimate because it assimilates what is verified, or experienced, to the verification, or experience, of it—the cat on the mat that I perceive to my perception of the cat on the mat. Such an assimilation is a standard part of the theory

of knowledge of the Idealist metaphysicians, but an analogous assimilation is made by some logical positivists who, in their talk about observation statements, do not carefully distinguish between the report of what is discovered and that of which it is a report. Having reached (*d*), the coherence theorist then emphasizes how much our previously acquired powers of judgment are exercised in this experience. He concludes that the second term with which our original judgment that there is a cat on the mat corresponds is not, as we thought, a fact; it is really another judgment or set of judgments.

Whether the whole argument is designed to show that correspondence is really coherence when the correspondence is put forward as giving the nature of truth or only when it is put forward as giving the criterion of truth, it seems equally invalid.

What the coherence theory really does is to give the criteria for the truth and falsity of a priori, or analytic, statements. Any attempt to change the meaning of "coherence" from coherence with other statements to coherence with fact (or reality of experience) is to abandon the theory. A merit of the theory is that it sees that the reasons for calling an analytic statement true or false are not those which some correspondence theorists, primarily thinking of empirical statements, try to fasten on all statements. When it sets itself up as the theory of truth, its mistake is twofold. First, it suggests that the criteria appropriate to a priori, or analytic, statements apply to every kind of statement; what the metaphysicians really did was to suppose all statements to be a priori.

Second, it confuses the reasons, or criteria, for calling a statement true or false with the meaning of "truth" or "falsity." As far as the criteria of truth are concerned, we can say only of a priori, or analytic, statements that they are true because they cohere with each other, and only of empirical statements that they are true because of what the world is like; however, as far as the meaning of truth is concerned, we can say of any kind of statement that it is true if it corresponds to the facts. Thus, as well as saying that a true a priori statement coheres with other statements in the system, we can also say that it corresponds to the a priori facts. It may be a fact that the sum of the angles of a Lobachevskian triangle is less than two right angles and also that the field of Waterloo is a mile square. What we must remember is that although both sorts of statements, if true, state the facts—tell us how things are—this amounts to something different in the two cases; the size of the angles of a Lobachevskian triangle is not something in the world in the way that the size of the field of Waterloo is.

See also Analytic and Synthetic Statements; Blanshard, Brand; Bradley, Francis Herbert; Coherentism; Correspondence Theory of Truth; Fallacies; Hegel, Georg Wilhelm Friedrich; Hempel, Carl Gustav; Idealism; Leibniz, Gottfried Wilhelm; Logical Positivism; Neurath, Otto; Rationalism; Spinoza, Benedict (Baruch) de; Truth.

Bibliography

Ayer, A. J. "The Criterion of Truth." *Analysis* 3 (1935–1936): 28–32. Reprinted in *Philosophy and Analysis*, edited by M. Macdonald. Oxford: Blackwell, 1954. Attacks the coherence theory as put forward by the logical positivists.

Bender, John W., ed. *The Current State of the Coherence Theory*. Boston: Kluwer, 1989.

Blanshard, Brand. *The Nature of Thought*. London: Allen and Unwin, 1939. Vol. II, Chs. 25–27.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Bradley, F. H. *Appearance and Reality*. Oxford, 1893. Chs. 15 and 24.

Bradley, F. H. *Essays on Truth and Reality*. Oxford: Clarendon Press, 1914. Chs. 7 and 11.

Dancy, Jonathan. "On Coherence Theories of Justification: Can an Empiricist Be a Coherentist?" *American Philosophical Quarterly* 21 (1984): 359–365.

Davidson, Donald. "A Coherence Theory of Truth and Knowledge." In *Reading Rorty*, edited by Alan R. Malachowski. Cambridge, MA: Blackwell, 1991.

Devitt, Michael. *Realism and Truth*. Oxford: Blackwell, 1984.

Ewing, A. C. *Idealism: A Critical Survey*. London: Methuen, 1934. Ch. 5.

Hempel, C. G. "On the Logical Positivists' Theory of Truth." *Analysis* 2 (4) (1935): 49–59.

Joachim, H. H. *The Nature of Truth*. Oxford: Clarendon Press, 1906.

Khatchadourian, Haig. *The Coherence Theory of Truth: A Critical Evaluation*. Beirut: American University, 1961.

Kirkham, R. L. *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press, 1992.

Putnam, Hilary. *Reason, Truth, and History*. Cambridge, U.K.: Cambridge University Press, 1981.

Quine, W. V., and J. S. Ullian. *The Web of Belief*, 2nd ed. New York: Random House, 1978.

Rescher, Nicholas. *The Coherence Theory of Truth*. Oxford: Clarendon Press, 1973.

Russell, Bertrand. *An Inquiry into Meaning and Truth*. London, 1940. Ch. 10.

Schlick, Moritz. "Facts and Propositions." *Analysis* 2 (5) (1935), 65–70. Reprinted in M. Macdonald, ed. *Philosophy and Analysis*. Oxford, 1954. Attacks the coherence theory as put forward by the logical positivists.

Sellars, Wilfrid. *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Press, 1991.

Sosa, Ernest. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Sosa, Ernest. "Reflective Knowledge in the Best Circles." *Journal of Philosophy* 94 (1997).

Sylvan, Richard. "On Making a Coherence Theory of Truth True." *Philosophica* (1990): 77–105.

Walker, Ralph. *The Coherence Theory of Truth: Realism, Anti-Realism, Idealism*. London: Routledge, 1989.

Woozley, A. D. *Theory of Knowledge*. London: Hutchinson, 1949. Pp. 146–169.

Alan R. White (1967)
Bibliography updated by Benjamin Fiedor (2005)

COHERENTISM

One of the three major views of the nature of epistemic justification, the coherence theory (or "coherentism") experienced a revival during the 1970s and 1980s after its near total eclipse earlier in the twentieth century. Although its origins can be traced to idealists, including Francis Bradley, Bernard Bosanquet, and Brand Blanshard, the coherence theory has more recently been espoused by empiricist-minded contemporary philosophers such as Wilfrid Sellars, Nicholas Rescher, Keith Lehrer, Gilbert Harman, and Laurence Bonjour. The coherence theory of justification stands as an alternative to both the more traditional foundations theory and the view called reliabilism. It should not be confused with a coherence theory of truth. A coherence theorist about justification can acknowledge a fact that cripples the coherence theory of truth, namely, that there are instances of coherent, hence justified, beliefs in falsehoods.

Although the details of different versions of the coherence theory vary widely, all versions share a positive thesis and a resulting negative claim. The coherence theory's positive thesis is that a belief is justified or warranted for a person to the degree that that belief coheres with the rest of that person's belief system. As a fabric derives its strength from the reciprocal ties and interconnections among its constitutive threads, so, for the coherentist, beliefs derive their justification from their interconnectedness with one's other beliefs. The negative claim endorsed by all coherentists is that foundationalism is in error when it asserts that some of our justified beliefs are privileged or basic—that is, their justification is at least partly independent of their connectedness with other held beliefs.

The coherentist's picture of mutual support or fit among our beliefs departs (to varying degrees) from the strictly linear image of justification that classical foundationalism endorses. For the foundationalist epistemic justification is transmitted to nonbasic beliefs, from those that are basic or foundational, along lines of inference and explanation. Inferred beliefs are justified by those from which they are inferred. For the coherentist the belief's justificatory status has less to do with the grounds on which a belief is based and more to do with the whole cluster of relations (of consistency, implication, probability, explanation, and the like) that more or less strongly fix that belief within the network of other held beliefs.

The exact nature of epistemic coherence, however, is very difficult to clarify, and disagreements occur even among coherentists. Some have argued that coherence is always and ultimately explanatory coherence, a question of whether a belief is a member of the best overall explanatory account accessible to an individual. Others claim that there are justificatory relations of comparative reasonableness of competing beliefs that reflect concerns wider than explanation alone, including measures of subjective probability and the relative informativeness of the proposition believed. Logical consistency seems to be a minimal necessary condition for maximal coherence, but some have argued that at least certain inconsistencies are unavoidable but do not so undermine coherence as to prevent beliefs from being justified. Speaking generally, coherence is a property of a belief system that is determined by the (various) connections of intelligibility among the elements of the system. Most agree that these include deductive, inductive, and abductive relations, as well as other explanatory and probabilistic connections. Some writers, especially pragmatists, are prepared to add relations such as the relative simplicity or the power of the explanations contained in one's belief system as contributors to overall coherence.

Motivation for the coherence theory comes most directly from finding foundationalism unworkable and believing as a consequence that some version of coherence must be correct. Another motivation comes from the observation that it seems apt and possible to ask about any belief what a person's reasons are for holding it. The theory also appears particularly compatible with the realization that all instances of epistemic justification are defeasible—that is, the justification of a given belief is always liable to undermining by other held beliefs, no matter how strong the initial grounds or evidential basis of the belief might be. Since undermining can come from any element of one's system that might be negatively relevant to a specific belief, it appears that complete epistemic justification, the kind necessary to support claims of knowledge, is sensitive to all of the connections among our beliefs, precisely as the coherence theorist urges. This

argument for the coherence theory is not decisive, however, since foundationalists can freely admit that warrant is undermined by a lack of coherence while still rejecting the coherentist's positive claim that coherence is the source of all epistemic justification.

In addition to the unclarities surrounding measuring degrees of coherence, numerous objections have been offered to coherentism. Four have been particularly prominent.

THE CIRCULARITY OBJECTION

If there are no foundational beliefs that act as the ultimate source of epistemic justification, and if the lines of justification transmission are not infinitely long (which appears absurd given the finitude of our mental capacities), then the coherence theory seems forced to claim that justification can be ultimately but not viciously circular. It is not immediately clear how circularity of this sort is anything but vicious, no matter how wide the circle may be, even though some have argued that wideness of a justificatory circle immunizes against viciousness. But if *A* is the source of justification for *B,* how can *B* be the source of justification for *A*? The coherentist can reply that the "source" of justification is the entire belief system. The linear model of justification on which the circularity objection is based may not be forceful against a more holistic construal of the relation. Taken as a holistic and higher-order relation constituted by lower-order reciprocal relations (at least some of which are asymmetric, such as "explaining" and "being explained by"), coherence might be able to avoid the problem of vicious circularity.

THE PROBLEM OF PERCEPTUAL BELIEFS

Certain simple and apparently immediate perceptual beliefs seem to be justified for us on the basis of the perceptual experience we currently are having rather than on any considerations about how that belief coheres with the rest of our belief system. Experience often seems to warrant beliefs that are anomalous—that is, do not cohere with already-held beliefs. In such cases we do not think that we are justified in rejecting the new belief on grounds of incoherence but often concede that revision of some previously held beliefs is appropriate. Coherentists have replied to this objection by arguing that the justification of even the most immediate perceptual belief requires that that belief cohere with our metabeliefs regarding how reliable or trustworthy we take our perceptual processes to be in the particular conditions. It is such metabeliefs that make it more reasonable to accept the anomalous perceptual experience than it is for us to conclude that we are hallucinating or have been deceived in some fashion. The introduction of metabeliefs into the explanation why immediate perceptual beliefs are often justified for us has struck many, however, as overintellectualizing our epistemic situation, as well as possibly reintroducing foundational principles into the theory of justification.

THE ISOLATION OBJECTION

This objection, closely related to the problem of perceptual beliefs, begins with the observation that coherence is a cognitively internal relation, relating belief to belief. But might not a thoroughly coherent system of beliefs nonetheless fail to be justified because they are not properly linked to the external perceptual circumstances? Would acceptance of a coherent fiction be justified if it were entirely the product of wishful thinking? The continual perceptual input we receive from the world must be assimilated into our belief system or else the justification for those beliefs will often suffer from undermining. The coherence theory seems too internalist to be a complete theory of epistemic justification, the objection concludes. Since coherence does not necessarily serve the epistemic goals of pursuing truth and avoiding error in our belief system, further constraints seem necessary if our notion of justification is to relate appropriately to knowledge. Coherentists respond in a number of ways to the isolation objection.

One alternative is to admit the objection's force and add a requirement that all justified systems include the belief that certain kinds of spontaneously occurring beliefs such as perceptual and memory beliefs are reliable or likely to be true. Demonstrating that this constraint is not an ad hoc amendment to coherentism is a difficult matter. A similar requirement applied to acceptances based on spontaneous wishful thinking would be obviously ad hoc and unacceptable. Some have suggested that metabeliefs about the trustworthiness of our perceptual beliefs in certain circumstances are not ad hoc and are important and legitimate members of our belief system, justified, as all beliefs are, through their coherence with our other beliefs. Whether such beliefs can be noncircularly defended, whether they constitute a sort of foundational belief, and whether they are realistically necessary for epistemic justification are each open matters.

THE INFERENTIAL-STRUCTURE OBJECTION

The foundationalist's traditional view—that whether one is epistemically justified in believing some proposition depends crucially upon the actual course of inference taken in arriving at a belief—is not easily relinquished. Coherence, however, is a relation determined only by the contents of beliefs and not by the order in which they have been inferred. Consequently, it appears possible that a series of beliefs inferred one from the other in a wholly fallacious manner might nevertheless cohere maximally with a background system of beliefs as long as there is another valid (but unused) course of inference that does connect them. This leads to the conclusion that, even if the coherence theory adequately captures the concept of epistemically justifiable beliefs relative to a system, it fails to explicate the notion of being justified in believing a proposition. Coherentists have responded to this challenge by relying once more on metabeliefs, claiming that when we infer *A* from *B* and *B* from *C* we also accept or believe that *A* follows from *B,* and not, for example, that *C* follows from *A*. Incorrect metabeliefs will, on some versions of coherentism, cause incoherence and loss of justification, keeping blatantly fallacious reasoning from ending in justified beliefs. This response, however, may generate an infinite regress of metabeliefs. Not all uses of inference schemes contain premises stating that the scheme is valid. One can infer *B* from *A* without first having to infer that *B* follows from *A*. Some coherentists answer this and other objections by admitting that their proposed conditions for coherence constitute ideals to which human knowers should aspire but seldom in actuality achieve. Debate over the merits of the coherence theory promises to continue unabated.

See also Blanshard, Brand; Bosanquet, Bernard; Bradley, Francis Herbert; Classical Foundationalism; Coherence Theory of Truth; Epistemology; Epistemology, History of; Harman, Gilbert; Lehrer, Keith; Reliabilism; Rescher, Nicholas; Sellars, Wilfrid.

Bibliography

Audi, R. "Foundationalism, Coherentism, and Epistemological Dogmatism." In *Philosophical Perspectives,* edited by J. Tomberlin, Vol. 2. Atascadero, CA: Ridgeview, 1988.

Bender, J. W., ed. *The Current State of the Coherence Theory.* Philosophical Studies, Vol. 44. Dordrecht: Kluwer Academic, 1989.

Bonjour, L. *The Structure of Empirical Knowledge.* Cambridge, MA: Harvard University Press, 1985.

Lehrer, K. *Theory of Knowledge.* Boulder, CO: Westview Press, 1990.

Plantinga, A. *Warrant: The Current Debate.* Oxford: Oxford University Press, 1993.

Pollock, J. *Contemporary Theories of Knowledge.* Totowa, NJ: Rowman and Littlefield, 1986.

John W. Bender (1996)

COLERIDGE, SAMUEL TAYLOR
(1772–1834)

Samuel Taylor Coleridge, the critic, romantic poet and philosopher, was born four years before the publication of Jeremy Bentham's *Fragment on Government,* and died only two years before the death of Bentham's most influential disciple, James Mill, at a time when the young John Stuart Mill was making a brilliant success in political journalism. The striking fact about Coleridge's place in English intellectual history, however, is that he developed a form of idealism in virtual isolation from the mainstream of empirical philosophy. In developing his own philosophical insights, Coleridge turned to Immanuel Kant. He had two reasons for doing this. First, he was deeply dissatisfied with the mechanistic theory of mind still flourishing in English philosophy, since he was unable to formulate within its terms certain views about poetic imagination; while Kant's *Critique of Judgment* (1790) had, however, set out with great rigor, and within a much more tractable conceptual framework, views essentially similar to Coleridge's own.

Second, Coleridge thought he saw in Kant's Transcendental Dialectic a way of combating the chronic latitudinarianism in English theology that had predominated throughout the eighteenth century and continued until the time of the Oxford Movement. But it must be remembered that although Coleridge was a serious student of Kant and one of Kant's earliest and ablest English interpreters, he was not a systematic or academic philosopher. His philosophical writings are always disorganized, eclectic, aphoristic. Philosophy became for him what poetry had always been: a necessary means for self-analysis, for the objectification of his personal engagement with life.

PHILOSOPHICAL DEVELOPMENT

What can be very schematically called the first stage in Coleridge's philosophical development was a highly enthusiastic acceptance in 1794 of David Hartley's theory of association and the "necessitarianism" which that doctrine seemed to imply. Also at this time, after an intense

study of John Locke and of William Godwin's *Inquiry concerning Political Justice* (1793), Coleridge became strongly inspired by the Enlightenment ideal of social perfectibility. So inspired was he that in December of that year, having had these enthusiasms reciprocated by Robert Southey, he left Cambridge without taking his degree. In January 1795 he lectured at Bristol on religion and politics and became preoccupied with Southey on the project of a pantisocracy, an ideal socialist community consisting of twelve young men and their wives, which was to be established on the banks of the Susquehanna. This project never really got under way; but its rather serious practical outcome for Coleridge was his marriage on October 4, 1795, to the uncomplicated Sara Fricker, sister of Southey's pantisocratic fiancée. Coleridge's early marriage was unfortunate because it prevented his developing what would have been in every way a more compatible relationship with Sara Hutchinson, whom he met through the Wordsworths in 1799 and whose inaccessibility he spent the greater part of his life lamenting. (Thus the celebrated *Dejection: An Ode,* written in 1802, should be considered more as a crescendo in this lament than as a statement of any alleged conflict between imagination and metaphysics.)

Despite his temporary acquiescence in Hartley's psychology, it was in fact Hartley's theology that most of all appealed to Coleridge. In particular, Hartley's idea of an ascending scale of affections, from primary sensations of pleasure and pain through new complexes of association to self-interest and eventually to sympathy, moral sense, and theophany (*Religious Musing,* 1794–1796) made a lasting impression on him. To this idea, conceived of mechanistically by Hartley, Coleridge later found an organically conceived analogue in Friedrich von Schelling's *Naturphilosophie.* Possibly in 1795, and certainly in 1796, Coleridge read George Berkeley. The next important stage of his philosophical development consisted in the replacement of Hartley's passive concept of mind by Berkeley's never consistently expressed notion of finite mind being actually creative in perception and imagination when it is considered as participating in the infinite, all-productive mind of God. Once more it was the place of God in the philosophy of Berkeley that most concerned Coleridge; and Berkeley's view of nature as purposive, as divine language, found expression in a number of poems written between 1796 and 1800 (for instance, *Destiny of Nations,* ll. 18–20; *Frost at Midnight,* ll. 59–62; *Apologia pro Vita Sua*).

By 1797 the Godwin-Hartley-necessity phase was over. It is probably significant that Coleridge emancipated himself from the mechanical theory of mind at the same time that he lost his once firmly held belief in the ideals of the French Revolution (*France: An Ode,* 1798). In September 1798, Coleridge accompanied the Wordsworths to Germany. After a short meeting in Hamburg with the poet F. G. Klopstock, Coleridge left the Wordsworths to see the countryside and settled himself at the University of Göttingen in order to improve his German and to collect material for a biography of Gotthold Ephraim Lessing. At Göttingen he attended the biological lectures of J. F. Blumenbach and had theological arguments with disciples of the rationalist J. G. Eichhorn. He returned to England in July 1799, transporting £30 worth of German philosophy books "with a view to the one work, to which I hope to dedicate in silence the prime of my life." This work was his never-completed *Opus Maximum.* Thus, the third period of Coleridge's philosophical development was a long assimilation of Kant and the German romantic philosophers, particularly Schelling, which he began in earnest in 1801 and continued well beyond 1816, when he was settled in the London house of James Gillman and able to write his most important philosophical works.

PHILOSOPHY AND FAITH

That "seminal" quality of mind that J. S. Mill detected in Coleridge and praised so highly needs, as we shall see, slight reevaluation. Mill was perhaps right in claiming that the "Germano-Coleridgean" school had done more for the philosophy of human culture than any of their predecessors could have done. Yet, in stressing the great contributions made to social theory by a series of Continental thinkers from Johann Gottfried Herder to Jules Michelet and in attributing to Coleridge simply a share in those contributions, Mill tended to ignore the less philanthropic and more personalistic aspects of European romanticism. For Coleridge was a post-Kantian "philosopher of life" in the tradition of Heinrich Heine's *Die romantische Schule.* For example, the closeness in particular doctrines and virtual identity in general philosophical orientation between Coleridge and Friedrich von Schlegel is remarkable. Both thinkers are essentially religious critics of the Enlightenment's secular anthropology. That man is a "fallen creature … diseased in his will" is a principle as axiomatic to Coleridge and Schlegel as it is self-dramatizing and even morally pernicious to the philosophical radicals.

Where Bentham and his followers write primarily as social reformers seeking, in the manner of David Hume and Claude-Adrien Helvétius, a means of harmonizing

individual egoism with the general good of society, the "Germano-Coleridgeans" take man's tragic alienation from God to be the fundamental datum not only of religion but also of philosophy. For the Benthamites the area of moral significance is in socioeconomic relationships, the external actions of everyday public association. For Coleridge, on the other hand, almost as much as for Søren Kierkegaard, the locus of reality is in the individual's experience of God. Thus, with thinkers like Coleridge philosophy inevitably becomes a form of theosophy. Religion is the highest exercise of the human spirit, and philosophy is a kind of rational prolegomenon that prepares the way for man's fuller appreciation of his relationship with God. Philosophy does this by trying to ascertain "the origin and primary laws (or efficient causes) either of the world man included (which is Natural Philosophy)—or of Human Nature exclusively, and as far as it is human (which is Moral Philosophy)." The remaining branch of philosophy, according to Coleridge, is epistemology, which deals with "the question concerning the sufficiency of the human reason to arrive at the solution of both or either of the two former problems.

REASON AND UNDERSTANDING

The core of Coleridge's epistemology is contained in his distinction between Reason and Understanding and his insistence that these differ not in degree but in kind. Although the terminology Coleridge uses here is decidedly Kantian, Kant's distinction between understanding (*Verstand*) and reason in the narrow sense (*Vernunft*) is only superficially similar to Coleridge's. Like his parallel distinctions between Imagination and Fancy, Genius and Talent, Symbol and Allegory, Coleridge's contrast between Reason and Understanding is more evaluative than descriptive and well illustrates his characteristic attempt to keep empiricist and associationist concepts in a subordinate position within a larger idealist framework. Understanding is "the faculty of judging according to sense … the faculty by which we reflect and generalize," which roughly corresponds to Locke's definition of it as "the power of perception." In other words, it is what Coleridge takes to be the pragmatic reasoning faculty of the empiricists.

The Coleridgean Reason, however, is a higher and more esoteric faculty that has at least three not very clearly differentiated functions. In its "speculative" aspect, Reason (1) provides us with basic logical rules of discourse, the so-called laws of thought; (2) is the origin of synthetic a priori truths in mathematics and science; and, in its most important "practical" aspect (3) is "the source of ideas, which … in their conversion to the responsible will, become ultimate ends." Reason produces Ideas or ideals that, although not capable of demonstration, are nevertheless not self-contradictory and may have a clear and distinct form. But they can also, says Coleridge, be more like an instinct or longing: "a vague appetency towards something which the Mind incessantly hunts for … or the impulse which fills the young Poet's eye with tears, he knows not why."

What Coleridge's distinction amounts to is this: "Understanding" is a pejorative blanket term for the negative aspects of eighteenth-century logic and science, while "Reason" is an approbatory label for those personal ideals and religious beliefs that are psychologically foreign to, or at least not logically entailed by, scientific empiricism. "Reason" thus is clearly allied with Christian faith. Coleridge is not, then, doing a piece of straight conceptual analysis in making this distinction, even though he often writes as if he thinks he is. Instead, he is persuasively psychologizing in an attempt to reorient contemporary philosophical attitudes into unison with contemporary Christian ideals. The barely disguised function of Coleridge's distinction is to give metaphysical respectability to those Ideas of God, freedom, and immortality that Kant had rightly regarded as merely regulative rather than constitutive elements of knowledge.

MIND AND NATURE

Philosophy must begin, says Coleridge, with a primary intuition that can be neither merely speculative nor merely practical, but both in one. Here Coleridge significantly modifies the views of Schelling. If the existence of external nature is taken to be the primary intuition, as in natural philosophy, then it becomes necessary to explain how mind or consciousness can be related to it. If, conversely, mind is taken to be primary, as in the Cartesian *Cogito,* we must account for the existence and significance of nature. The only satisfactory way to do either of these things is to suppose that there is in fact no dualism between nature and mind. Nature appears as extrinsic, alien, and in antithesis to mind. The difference is not absolute, however, but merely one of degree of consciousness and, consequently, of freedom.

Nature is mind or spirit slumbering, unconscious of itself. It is representable under the forms of space and time, subject to the relations of cause and effect, and requires an antecedent explanation. Mind, however, originates in its own (that is, God's) acts and exists in a realm of freedom. But if in its turn this qualitative difference between nature and mind is to be accounted for, a first

cause must be postulated that is itself neither exclusively mind nor exclusively nature, subject or object, but the identity of both. Such a first cause or unconditional principle could not be a natural thing or object because each thing is what it is in consequence of some other thing. Nor can this principle be mind as such, because mind exists only in antithesis to nature. (Rather than indulging in tautology here, Coleridge seems to be making the phenomenologist's point that consciousness is always intentional; i.e., is consciousness *of* something.) The unconditioned must be conceived, apparently, as a primeval synthesis of subject and object, consciousness and nature, in the self-consciousness of God. In God or Spirit lies the identity of the two, of being and knowing in the "absolute I AM."

Thus nature and mind seem to be conceived by Coleridge as two dialectical opposites resulting from God's free act of self-alienation in becoming self-conscious. On this last point, however, he is in his published works particularly (and perhaps necessarily) obscure. Unlike Johann Gottlieb Fichte and Schelling, Coleridge wishes to combine the dialectics of the Identity-Philosophy with the traditional Christian concept of dualism between creature and creator. In the unpublished *Opus Maximum* and other manuscripts, he elaborates this point of divergence from the Germans by distinguishing the "personeity" of God from the "personality" of man and goes to great lengths in accounting for the problem of evil. What is important and seminal in Coleridge's metaphysics, however, is not its details or conclusions, but the rich suggestiveness of its basic categories applied to certain problems in aesthetics and social theory.

IMAGINATION AND FANCY

From the formal dialectics of his idealism Coleridge drew a living description of how the artist's mind works. Since conscious life exists only through contradiction, or doubleness, the whole of nature out of which conscious life develops must exhibit opposing forces in the reconciling and recurrence of which "consists the process and mystery of production." Art is produced through that same dialectical struggle for the reconciliation of opposites that takes place between mind and nature. Art is not, then, merely imitative, but symbolic of reality. Like all symbols (as Coleridge defines them), it is consequently an inherent part of the process it represents; and the artist as creator, his consciousness being the focus of nature and Idea, matter and form, becomes symbolic of God. So, like God, the artist or Genius must suffer alienation in order to create. He needs to be in a special sense disinterested, emotionally aloof for a while from his subject matter and from himself. For in the joy of creation "individuality is lost." He must first "eloign himself from nature in order to return to her with full effect." Just as in the cosmic struggle for synthesis, so in the microcosm of art and the individual artist's mind, there is an attempted fusion of conscious and unconscious forces.

The artist (Coleridge usually considers the case of the poet) achieves such fusions in virtue of his special psychological makeup; that is, through his having the power of Imagination. Coleridge's theory of Imagination, however, does not neatly reflect any of the everyday uses of "imagination" distinguished by modern linguistic analysts. His poet does not create through merely imaginary (unreal) fantasy, nor does he imagine in the sense of making to himself or his reader a kind of supposal, veridical or false. And although it is of course true that the poet is imaginative in being creative or inventive, it is not the case, according to Coleridge, that it is in this fact alone that the poet's Imagination consists.

Nor is Imagination "invention" in the sense that it adds to the real, as common usage might suggest. Instead, as we have seen, Coleridge's view is that the poem and the poet are microcosmic analogues, indeed symbolic parts, of reality. His theory is not concerned, then, with an elucidation of ordinary senses of "in imagination" or even with ordinary senses of "with imagination." It is, typically, a piece of speculative (though not therefore unempirical) psychology that is the rather overweighted vehicle for a value judgment. In this and certain other respects, Coleridge's theory of Imagination has interesting affinities with Jean-Paul Sartre's theory in which imagination is related to the notion of nihilation of consciousness. Needless to say, Sartre is borrowing from a later development of the same German tradition to which Coleridge was indebted.

Coleridge considers three things: primary Imagination, secondary Imagination, and Fancy. The power of primary Imagination is not peculiar to poets, but is standard psychological equipment for all men. It is Coleridge's term for what he considers to be finite mind's repetition in perception of God's creative act. His view seems to be that by synthetically perceiving and categorizing things that are not me, I become conscious of myself, and that this state of human self-consciousness is analogous to God's own creative schizophrenia. Secondary Imagination is the specialized poetic faculty. Differing only in degree and in its mode of operation from primary Imagination, it is the poet's power of unifying

chaotic experience into the significant form of art. Thus, secondary or poetic Imagination "dissolves, diffuses, dissipates, in order to recreate … it struggles to idealize and to unify. It is essentially *vital,* even as all objects (*as* objects) are essentially fixed and dead."

Fancy, on the other hand, differs in kind from Imagination. While poetic Imagination is organic in its operation, producing true analogues of God's creation, Fancy is merely mechanical, aggregative; it is at best imitative rather than symbolic and the instrument of Talent, as opposed to Genius. Fancy is in fact that lower-grade imagination that Locke and Hume set beside sense and memory as a third, nonreferential, source of ideas. Thus Fancy is allied to Understanding, while Imagination, in its ability to transcend and transform the phenomenal, is allied to Reason. It embodies in works of art that inner struggle between nature and mind within which art and Genius are temporary points of resolution.

Despite Coleridge's unhelpful talk about Imagination and Fancy being mental faculties, there is no doubt that the concrete application of these essentially evaluative concepts leads to a highly practical literary criticism. To mention only one instance, Coleridge's conception of the work of art as in some degree analogous to a biological organism and his distinction between mechanical regularity and organic form in poetry has had the greatest possible influence on modern criticism. Largely through the far-reaching implications of his distinction between Imagination and Fancy, Coleridge became the first English writer on poetry since the Renaissance to embody the highest powers of critical response within a framework of philosophical concepts that seemed to explain and reinforce that response rather than to inhibit or destroy it.

MORALS AND POLITICS

Although Coleridge was in his ethical theory a follower and acute critic of Kant, he is interesting today not so much for his own positive views as for his attack upon utilitarianism. Coleridge launches this attack in two ways. First, he tries to demonstrate the logical absurdity of the greatest happiness principle by reductio ad absurdum techniques; second, he "postulates the Will," which involves the claim that the utilitarian notion of personality is psychologically inadequate. On the logical side, Coleridge opens fire with the surprisingly modern assertion that the whole of moral philosophy is contained in one question: "Is *Good* a superfluous word … for the pleasurable and its causes—at most a mere modification to express degree and comparative duration of pleasure?" His reply is that the meaning of *good* can be decided only by an appeal to universal usage, for the distinction between *good* and *pleasurable,* which, he holds, is common to all languages of the civilized world, must "be the consequent of a common consciousness of man as man."

Then, avoiding the error J. S. Mill was soon to make, Coleridge distinguishes between things that are good because they are desired, and things that are or ought to be desired because they are good. This leads him to conclude that *good* cannot be defined simply in terms of pleasure or happiness. Against the Benthamite view that the agent's motive has nothing to do with the morality of his action, Coleridge makes two points, partly logical and partly psychological. The utilitarian position cannot generally hold, he says, because it follows from it that I could do a morally right act by sheer chance. But such complete lack of inward, conscious participation on my part could never be a sufficient criterion for my acting morally. The utilitarian principle therefore confounds morality with law. Moreover, it is no defense here to say that the principle was put forward as a criterion for judging the morality of the action and not that of the agent, because this last distinction is "merely logical, not real and vital." Acts cannot be dissociated from an agent any more than ideas from a mind.

In his social philosophy, Coleridge writes in the tradition of Edmund Burke. His mature views are contained in *On the Constitution of the Church and State,* which was begun as an attempt to formulate objections to various bills for Catholic emancipation and finished as an idealist treatise containing the whole logomachy of organism and the reconciliation of opposites. In any society there are always two antithetical forces at work. Since, dialectically speaking, "opposite powers are always of the same kind, and tend to union," Coleridge's idea of a well-functioning society is the nonrevolutionary reconciliation of forces working for permanence with forces working for progression. These he identifies with, respectively, the aristocratic, landed interest and the bourgeois, commercial interest of early Victorian England; a monarch also being required to maintain cohesion.

Coleridge's habit of generalizing from the history and the contemporary pattern of British political institutions rather than, as he alleges, drawing a description of the idea of a state, should at least make suspect his application of these largely a priori principles. This habit leaves Coleridge, like G. W. F. Hegel, wide open to the charge of surrounding the constitution of his own country with an aura of metaphysical sanctity to which it has no claim. Despite such ruinous methodology, however, what Coleridge has to say about the intelligentsia and the part

it has to play in the dissemination of culture has been influential.

Coleridge contrasts *cultivation* with *civilization.* Civilization he takes to denote external, material social progress, while cultivation is more inward and personal: the "harmonious development of these qualities and faculties that characterize our humanity." So that cultivation can take place, Coleridge proposes the formation of a state-endowed class, the "clerisy" or "national church," which would effectively consist of professors of liberal arts officially established throughout the country. The national church would, however, be in no sense identical with the Church of England or with any purely religious organization. Its purpose would be to preserve the results of learning, to "bind the present with the past" and to give every member of the community an understanding of his social rights and duties. The almost limitless possibilities for authoritarianism in such an arrangement are, again, obvious. Nevertheless, in Coleridge's *Church and State* the idea of culture as something independent of material progress was first systematically introduced into English thinking, and was from then onward available in various forms, not merely to influence society but also to judge it.

CONCLUSION

Though it is no doubt true that Coleridge was, with Bentham, one of the great seminal minds of England in his age, it is not true without qualification that the cultural powers wielded by Bentham and Coleridge were "opposite poles of one great force of progression." Here Mill was surely indulging in public-spirited wish fulfillment rather than relating the facts. Coleridge and his German contemporaries undoubtedly brought to social consciousness those deeper insights into the nature of the individual and the organic complexities of human associations that were classically synthesized by Hegel in the *Philosophy of Right* (1821). Yet the inherent ambiguity of these insights has today become a disturbing commonplace. Mill inevitably overlooked the darker side of romanticism. For once the romantic artist or philosopher ceases to believe in God, he tends either to find a new object of veneration in history or hero worship or, more recently, to relinquish his very inwardness and imagination in solipsistic nausea. It was Coleridge's curious fortune that he never lost his belief in God.

See also Aesthetics, History of; Bentham, Jeremy; Berkeley, George; Burke, Edmund; Cartesianism; Enlightenment; Fichte, Johann Gottlieb; Godwin, William; Hartley, David; Hegel, Georg Wilhelm Friedrich; Helvétius, Claude-Adrien; Herder, Johann Gottfried; Hume, David; Idealism; Imagination; Kant, Immanuel; Kierkegaard, Søren Aabye; Lessing, Gotthold Ephraim; Locke, John; Mill, James; Mill, John Stuart; Romanticism; Schelling, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von.

Bibliography

The Complete Works of S. T. Coleridge, edited by W. G. T. Shedd, 7 vols. (New York: Harper, 1853 and 1884), was very far from complete. Fortunately, Professor Kathleen Coburn is now in the process of preparing the first scholarly and really comprehensive edition of practically everything Coleridge wrote, apart from the *Collected Letters,* edited by E. L. Griggs, 6 vols. (Oxford: Clarendon Press, 1956–1971). Coburn has already brought out two volumes of Coleridge's *Notebooks* (London and New York, 1957 and 1962) and his *Philosophical Lectures* (London: Pilot Press, 1949). Coburn's *Inquiring Spirit: A New Presentation of Coleridge from His Published and Unpublished Prose Writings* (New York: Pantheon, 1951) gives an exciting foretaste of what is to come.

Among the prose works essential for a study of Coleridge as a thinker, note particularly the following: *Biographia Literaria* (New York: Kirk and Merein, 1817), edited by John Shawcross, 2 vols. (Oxford: Clarendon Press, 1907); *The Friend,* 3 vols. (London: Fenner, 1818); *Aids to Reflection* (London: Taylor and Hessey, 1825); *On the Constitution of the Church and State, According to the Idea of Each* (London, 1830); *Coleridge on Logic and Learning,* edited by A. D. Snyder (London, 1929); *S. T. Coleridge's Treatise on Method,* edited by A. D. Snyder (London: Constable, 1934); *Coleridge's Shakespearean Criticism,* edited by T. M. Raynor (London: Constable, 1930); *Specimens of the Table Talk of the late S. T. Coleridge,* edited by H. N. Coleridge, 2 vols. (London: Murray, 1835).

On Coleridge's general philosophical position, see especially John H. Muirhead, *Coleridge as Philosopher* (London: Allen and Unwin, 1930) and Elizabeth Winkelmann, *Coleridge und die Kantische Philosophie* (Leipzig: Mayer and Müller, 1933). Both these works are reliable but philosophically very old-fashioned. Among the very large number of articles on Coleridge's thought, A. O. Lovejoy, "Coleridge and Kant's Two Worlds," reprinted in *Essays in the History of Ideas* (Baltimore: Johns Hopkins Press, 1948) deserves special attention. James D. Boulger, *Coleridge as Religious Thinker* (New Haven, CT: Yale University Press, 1961) gives a good account of the later Coleridge.

On the theory of imagination, see especially James V. Baker, *The Sacred River* (Baton Rouge: Louisiana State University Press, 1957); I. A. Richards, *Coleridge on Imagination* (2nd ed., New York: Norton, 1950); René Wellek, *A History of Modern Criticism* (London, 1965), Vol. 2, Ch. 6; J. M. Cameron, "Poetic Imagination," in *Proceedings of the Aristotelian Society,* n.s. 62 (1961–1962): 219–240. The best short work in this field is Gordon McKenzie, *Organic Unity in Coleridge,* University of California Publications in English, Vol. 7, No. 1 (Berkeley: University of California Press, 1939), 1–108. R. H. Fogle, *The Idea of Coleridge's*

Criticism (Berkeley: University of California Press, 1962) is suggestive if not always epistemologically acute.

A reliable account of Coleridge's political thought is given in John Colmer, *Coleridge: Critic of Society* (Oxford: Clarendon Press, 1959). J. S. Mill, "Bentham" and "Coleridge," in *London and Westminster Review* (1838 and 1840) will always remain great classics. Also see F. R. Leavis, *Mill on Bentham and Coleridge* (London: Chatto and Windus, 1950) and Raymond Williams, *Culture and Society, 1780–1950* (New York: Columbia University Press, 1958), Part II, Ch. 3. Justus Buchler, *The Concept of Method* (New York: Columbia University Press, 1961) contains an interesting analysis of the *Treatise on Method* and compares Coleridge's views with those of other thinkers, including Bentham.

OTHER RECOMMENDED WORKS

Appleyard, J. A. *Coleridge's Philosophy of Literature: The Development of a Concept of Poetry, 1791–1819*. Cambridge, MA: Harvard University Press, 1965.

Barfield, Owen. *What Coleridge Thought*. Middletown: Wesleyan University Press, 1971.

Coleridge, Samuel Taylor. *The Friend: A Series of Essays to Aid in the Formation of Fixed Principles in Politics, Morals, and Religion with Literary Amusements Interspersed*. London: W. Pickering, 1844.

Coleridge, Samuel Taylor. *The Relation of Philosophy to Theology, and of Theology to Religion*. London: Ward, 1851.

Gallant, Christine, ed. *Coleridge's Theory of Imagination Today*. New York: AMS Press, 1989.

Hamilton, Paul, D. Phil. *Coleridge's Poetics*. Oxford: Basil Blackwell, 1983.

Haney, David P. *The Challenge of Coleridge: Ethics and Interpretation in Romanticism and Modern Philosophy*. University Park: Pennsylvania State University Press, 2001.

Haney, John Louis. *The German Influence on Samuel Taylor Coleridge*. New York: Haskell House, 1966.

Hedley, Douglas. *Coleridge, Philosophy and Religion: Aids to Reflection and the Mirror of the Spirit*. Cambridge, U.K.: Cambridge University Press, 2000.

Hodgson, John A. *Coleridge, Shelley, and Transcendental Inquiry: Rhetoric, Argument, Metapsychology*. Lincoln: University of Nebraska Press, 1989.

Lockridge, Laurence S. *Coleridge the Moralist*. Ithaca: Cornell University Press, 1977.

Marks, Emerson R. *Coleridge on the Language of Verse*. Princeton, NJ: Princeton University Press, 1981.

McKusick, James C. *Coleridge's Philosophy of Language*. New Haven, CT: Yale University Press, 1986.

McNiece, Gerald. *The Knowledge that Endures: Coleridge, German Philosophy and the Logic of Romantic Thought*. London: Macmillan, 1992.

Modiano, Raimonda. *Coleridge and the Concept of Nature*. Tallahassee: Florida State University Press, 1985.

Newlyn, Lucy. *The Cambridge Companion to Coleridge*. Cambridge, U.K.; New York: Cambridge University Press, 2002.

Orsini, Gian Napoleone Giordano. *Coleridge and German Idealism: A Study in the History of Philosophy with Unpublished Materials from Coleridge's Manuscripts*. Carbondale: Southern Illinois University Press, 1969.

Perkins, Mary Anne. *Coleridge's Philosophy: The Logos as Unifying Principle*. Oxford: Clarendon Press; New York: Oxford University Press, 1994.

Perry, Seamus. *Coleridge and the Uses of Division*. Oxford: Clarendon Press; New York: Oxford University Press, 1999.

Swinden, Patrick. "Coleridge and Kant." In *Literature and the Philosophy of Intention*. Basingstoke, Hampshire: Macmillan; New York: St. Martin's Press, 1999.

Taylor, Anya. *Coleridge's Defense of the Human*. Columbus: Ohio State University Press, 1986.

Wheeler, Kathleen M. "Coleridge's Attack on Dualism." In *Romanticism, Pragmatism, and Deconstruction*. Oxford, U.K.; Cambridge, MA: Blackwell, 1993.

Wylie, Ian. *Young Coleridge and the Philosophers of Nature*. Oxford; New York: Oxford University Press, 1989.

Michael Moran (1967)
Bibliography updated by Desiree Matherly Martin

COLET, JOHN

(1466–1519)

John Colet, the Christian humanist and English educator, was the founder of St. Paul's School for Boys, leader of the "Oxford Reformers" Sir Thomas More and Desiderius Erasmus, and chief transmitter of Florentine Platonism from Italy to such English Renaissance figures as Edmund Spenser, John Donne, and John Milton. The son of a London lord mayor, Colet took a master's degree from Oxford (1490) and then explored Plato, Plotinus, and Origen in Latin translation. From 1493 to 1496, he traveled in France and Italy. The appealing tradition that he studied in Florence under Marsilio Ficino was shattered in 1958 when Sears Jayne discovered correspondence between Colet and Ficino in a copy of Ficino's *Epistolae* (1495) at All Soul's College, Oxford. This correspondence shows that Colet never visited Florence or met Ficino.

Upon his return to Oxford in 1496, Colet delivered Latin lectures on St. Paul's Epistles to the Romans and Corinthians. The visiting Erasmus and others applauded as Colet, frequently quoting the Florentine Platonists, propounded a new "historical approach" to the study of Scripture. In 1504 Colet was appointed dean of St. Paul's Cathedral, where, contrary to custom he preached frequently and in English. His congregation included the young lawyer Thomas More.

Colet's penchant for controversy is illustrated by his *Convocation Sermon* (1512), in which he wrathfully condemned his own bishops for their moral laxness. Charges of heresy provoked by this sermon were dismissed by his friend Archbishop Warham, but Colet was soon again involved in controversy. He attacked the war policy of

Henry VIII and was summoned to court; but Henry, after hearing Colet's arguments, was so dazzled that he made the dean a royal chaplain.

Colet's chief contribution to philosophy was his remarkably successful attempt to blend pagan and Christian thought. In practice Colet followed the approach of St. Augustine, who argued that pagan philosophy, when properly controlled, is a useful handmaiden for Christianity. By pagan philosophy, Colet understood especially Florentine Platonism, a weird conglomeration of original Platonism, later Neoplatonism, and private Florentine speculation on man, love, beauty, and mystical union. Much of this speculation came to Colet through Ficino's *Theologia Platonica* (1482) and Giovanni Pico della Mirandola's *Heptaplus* (1489), both of which he admiringly quoted or paraphrased in his scriptural treatises.

Despite his debt to the Florentines, Colet avoided the heretical Florentine approach which proclaimed that pagan philosophy and Christianity are equal and even identical. Instead, Colet was careful, as was his model Augustine, to purge pagan views of heretical "errors" before merging them with Christian doctrine. For example, Colet favored the Platonic soul-body terminology over Paul's spirit-flesh, but rejected Plato's dictum that the soul alone comprises the total personality. Again, Colet accepted the Neoplatonic view that Creation was a merging of form and matter, yet he was careful to emphasize that this form is not an emanationist overflow from God's essence, but rather an entity created by God outside himself. In the realm of redemption, Colet accepted Plato's position that only a harmonized soul can govern the body, but he deviated from Plato in insisting that such harmonization can come only from the Holy Spirit's infusion of sanctifying grace. Even in the delicate area of mysticism, Colet borrowed from the *Symposium* the view that love transforms the lover into the object loved.

Whether Colet was as successful in adhering to Catholic as to generally Christian doctrine is a controversial issue. A doctrinal cleavage between Colet and More would seem to be reflected in the *Dialogue on Tyndale* (1529), where More strongly rebuts a form of religion (described in words almost identical to Colet's *Exposition of Romans*), which condemns, as mere shadows, all types of external religion such as sacraments, vestments, and ritual. A comparative study of Colet and More suggests that Colet might have found himself in grave difficulty with Catholic authorities had he lived until the doctrinal reformation of 1534.

Bibliography

During the period 1867–1876, Joseph Hirst Lupton issued five volumes in which he edited Colet's *Treatise on Sacraments* and translated the remainder of Colet's extant Latin works: two treatises on the Hierarchies of the pseudo-Dionysius (c. 490–550); the lectures on St. Paul's First Epistle to the Corinthians; a similar series on the Epistle to the Romans; a Genesis commentary in the form of *Letters to Radulphus*; an exposition (never given as lectures) on the first five chapters of Romans; and *On Christ's Mystical Body, the Church.* Lupton's translations have been out of print since 1893, but Bernard O'Kelly began modern translations with *John Colet's Enarratio Primum S. Pauli Epistolam ad Corinthios* (Oxford, 1963). Colet's only English works, the *Convocation Sermon* and the *Right Fruitful Monition* pamphlet, are printed in the appendix to Lupton's *A Life of John Colet* (London: Bell, 1887), still the indispensable biography. Another biographical classic is Frederic Seebohm's *The Oxford Reformers* (London: Longmans Green, 1867). Colet's influence as educator is detailed in M. F. McDonnell's *The History of St. Paul's School for Boys* (London, 1909).

Since World War II there has been a revival of interest in Colet's thought, as evidenced by Eugene Rice Jr., "John Colet and the Annihilation of the Natural," in *Harvard Theological Review* 45 (July 1952): 141–163; Albert Duhamel, "The Oxford Lectures of John Colet," in *Journal of the History of Ideas* 14 (October 1953): 493–510; Ernest William Hunt, *Dean Colet and His Theology* (London: Church Historical Society, 1956); Leland Miles, *John Colet and the Platonic Tradition* (La Salle, IL: Open Court, 1961; London, 1962), and Sears Jayne, *John Colet and Marsilio Ficino* (London: Oxford University Press, 1963).

OTHER RECOMMENDED TITLES

Gleason, John B. *John Colet.* Berkeley: University of California Press, 1989.

Kaufman, Peter I. *Augustinian Piety and Catholic Reform: Augustine, Colet, and Erasmus.* Macon, GA: Mercer University Press, 1982.

Miles, Leland. "Platonism and Christian Doctrine: The Revival of Interest in John Colet." *Philosophical Forum* 21 (1963–1964): 87–103.

Leland Miles (1967)
Bibliography updated by Tamra Frei (2005)

COLLIER, ARTHUR

(1680–1732)

Arthur Collier, an English idealist philosopher, was born at Langford Magna, Wiltshire, where his father was rector. In 1697 he entered Pembroke College, Oxford, but transferred in 1698 to Balliol. He took orders and in 1704 succeeded to the family living at Langford Magna. Such events as mark his life were of a private character. He was in constant financial difficulties, arising, it is said, from his own impracticality and the extravagance of his wife; his writings did nothing to bring him into contact with a

wider world since scarcely anybody read them. He was buried at Langford on September 9, 1732.

Collier makes no mention of John Locke. He read George Berkeley (with whose views his own partly coincide), but only after the publication of Collier's major work, *Clavis Universalis* (1713). René Descartes, Nicolas Malebranche, and Collier's neighbor John Norris were the philosophers who particularly interested Collier, although he was also considerably influenced by Francisco Suárez and other late scholastic philosophers.

Malebranche and John Norris had argued that perception provides us with no direct evidence for the existence of an external world. They did not deny, however, the existence of such a world, even though it is an embarrassment to their metaphysics. They retain it for theological reasons. Collier agreed with them in rejecting the view that perception reveals an external world to us but went on to argue that the very conception of an external world is self-contradictory.

PHILOSOPHICAL VIEWS

In the Introduction to *Clavis Universalis* Collier begins by explaining just what he wishes to assert and what to deny. His starting point is that what we perceive is "in the mind"; the objects of perception, that is, depend upon the mind for their existence. In denying their externality Collier is denying their independence or self-subsistence; he is not at all denying that they exist. "It is with me a first principle," he writes, "that whatsoever is seen, is." Indeed, even what is imagined must exist, since it is an actual object of mind. Collier does not deny, either, that what we perceive seems to us to be independent of our minds. But, he suggests, this "quasi-externeity" also characterizes what we imagine as much as what we see. The difference between types of objects of perception lies only in the degree of vividness with which they are perceived.

Collier is not, of course, alleging that our mind causes the ideas which it has. Ideas, he says, exist in the mind qua perceiver, not qua voluntary agent. Nor is he asserting that the ideas which other people perceive are internal to my mind. "The world which John sees is external to Peter, and the world which Peter sees is external to John." Peter's world and John's world may be similar, but they are numerically different. The crucial point for Collier is that every object must be "in-existent" to some mind; every object has existence, but no object has "extra-existence."

To establish his main conclusions, Collier makes use of two main lines of argument, to each of which a book of the *Clavis Universalis* is devoted. In the first book he sets out to show that we have no good reason for believing that objects exist externally to mind. It is generally supposed that we directly perceive them to be external, but the "quasi-externeity" of objects is no proof, he argues, that they are really external. Everybody admits that in hallucinations, for example, we can suppose objects to be external which are not in fact external. As for the Cartesian argument that there must be an external world because otherwise God would have deceived us when he implanted in us so strong an inclination to believe that there is, Collier points out that according to Descartes himself we are constantly mistaken about what is and what is not a property of the external world. If we can be mistaken about the externality of colors, for example, without God's veracity being impugned, why not about the existence of objects?

Thus far, Collier's argument has been in some measure an *argumentum ad hominem*; he has supposed it to be an intelligible hypothesis that there is an external world and has argued only that there is no good reason for accepting that hypothesis. In the second book he goes further. The concept of an external world is, he says, riddled with contradictions. To establish this point, he calls upon the commonplace skeptical arguments of his time, which had ordinarily been used, however, to demonstrate that the concept of the physical world is as full of mysteries and obscurities as are the concepts of theology rather than to show that it does not exist. Philosophers have demonstrated, Collier argues, that an external world must be finite and that it must be infinite, that it must be infinitely divisible and that it cannot be infinitely divisible, that it is capable of motion and that it cannot be capable of motion. Faced with this situation, we have no alternative but to declare that the very concept of an external world is self-contradictory. Finally, he argues, no intelligible account can be given of the relation between an external world and God. Stress its dependence on God's will, and its externality vanishes; stress its externality, and it takes on the attributes of God.

In a letter to the publisher Nathaniel Mist, Collier pushes his argument slightly further. The subtitle of *Clavis Universalis* was, he now says, misleading insofar as in it he professed to provide "a demonstration of the non-existence or impossibility of an external world." This suggests that the existence of an external world is a possibly true, even if in fact a false, hypothesis. The correct account of the matter is that the doctrine that an external world exists is "neither true nor false"; it is "all-over non-

sense and contradiction in terms," the very concept of an external world being self-contradictory.

RELIGIOUS VIEWS

Collier's other publications consist of *A Specimen of True Philosophy, in a Discourse on Genesis* (1730), which is designed as a preliminary essay to a complete commentary on the Bible, and a series of seven sermons published as *Logology* (1732). These works are primarily theological. Collier's metaphysical views are more clearly formulated in the brief "Confession" he wrote in 1709 but did not publish. There is, he says, one substance, God, which is "being itself, all being, universal being." The existence of everything else is dependent upon the existence of God not only causally but also in the sense that particular things have no substance of their own. However, although everything but God is ultimately dependent on him, everything except Christ is also relatively dependent on something else; qualities "in-exist" in objects, objects in the mind, and the mind in Christ, through whom God made the transition from universality to particularity. Not unnaturally, Collier was accused of Arianism. He thought of himself, however, as reconciling the Arians and the orthodox by admitting Christ's dependence on God but asserting his priority to all created things and even to time, Christ's begetting being "the first pulse of time."

In Great Britain attention was first drawn to Collier's work by Thomas Reid and Dugald Stewart, but he has never exerted any real influence, being overshadowed by Berkeley. In Germany he attracted some attention as a result of an abstract of the *Clavis Universalis* published in the *Acta Eruditorum* (1717) and a German translation by John Christopher Eschenbach in 1756. He is quoted by Christian Wolff, and it is sometimes supposed, without any real evidence, that the Kantian antinomies derive from his work.

See also Arius and Arianism; Berkeley, George; Descartes, René; Idealism; Locke, John; Malebranche, Nicolas; Norris, John; Reid, Thomas; Stewart, Dugald; Suárez, Francisco; Wolff, Christian.

Bibliography

For *Clavis Universalis* see Ethel Bowman's edition (Chicago: Open Court, 1909) or Samuel Parr, *Metaphysical Tracts of the Eighteenth Century* (London: E. Lumley, 1837), which also includes *A Specimen* and a brief précis of *Logology.* Antonio Casiglio has translated *Clavis Universalis* into Italian, with notes (Padua, 1953), and has included as an appendix to an article on Collier's *True System* an Italian translation of that work in *Sophia* 23 (3–4) (1955): 302–321. Collier's "Confession" is in Robert Benson, *Memoirs of the Life and Writings of the Rev. Arthur Collier* (London: E. Lumley, 1837).

See also John Henry Muirhead, *The Platonic Tradition in Anglo-Saxon Philosophy* (London and New York: n.p., 1931); Georges Lyon, *L'idéalisme en Angleterre au XVIIIe siècle* (Paris: F. Alcan, 1888). On Collier and Berkeley see George Alexander Johnston, *The Development of Berkeley's Philosophy* (London: Macmillan, 1923), Appendix I; for Collier and Immanuel Kant see A. O. Lovejoy, "Kant and the English Platonists," in *Essays Philosophical and Psychological in Honor of William James … by His Colleagues at Columbia University* (New York: Longmans, Green, 1908), and H. J. de Vleeschauwer, "Les antinomies kantiennes et la *Clavis Universalis* d'Arthur Collier," in *Mind* 47 (187) (1938): 303–320.

John Passmore (1967)

COLLINGWOOD, ROBIN GEORGE

(1889–1943)

Robin George Collingwood, the English philosopher and historian, was born in Coniston, Lancashire. His father, W. G. Collingwood, friend and biographer of John Ruskin, educated him at home until he was old enough to enter Rugby and imbued him with a Ruskinian devotion to craftsmanship and art and an adult attitude toward scholarship. Although Collingwood later wrote contemptuously of most of his teachers at Rugby and praised Oxford chiefly for leaving him to himself, his undergraduate work in Greek and Latin was excellent and in *literae humaniores* (philosophy and history from Greek and Latin texts), brilliant. He was elected to a fellowship at Pembroke College in 1912, and to the Waynflete professorship in 1934. Except for a period of service with the admiralty intelligence during World War I, he remained at Oxford throughout his career, until in 1941 illness compelled him to retire.

Although he always considered philosophy his chief vocation, Collingwood was a pupil of the great Romano-British archaeologist F. J. Haverfield. Since he alone of Haverfield's pupils both survived the war and remained at Oxford, Collingwood considered it his duty to transmit Haverfield's teachings to others. Although he was a competent excavator, most of Collingwood's work was theoretical. Both in suggesting questions that excavation might answer and in drawing together and interpreting the results of others' excavations, he was brilliant. The final monuments to his historical labors are his sections

on Roman Britain in the first volume of the *Oxford History of England* (1936; 2nd ed., 1937) and in Tenney Frank's *An Economic Survey of Ancient Rome* (5 vols., New York, 1933–1940). To these must be added his extensive contributions to the revised edition of the British section of Theodor Mommsen's *Corpus Inscriptionum Latinarum*, begun by Haverfield, for which Collingwood drew each inscription from his own accurate rubbings.

The consensus of present-day archaeologists appears to be that Collingwood's "imperishably accurate" work on inscriptions will prove more valuable than his works of synthesis and interpretation. Collingwood himself expected that his interpretations would be superseded, but he was convinced that first-rate thinking in history, as in natural science, remains valuable even if further evidence requires that its conclusions be revised. In most of his work his willingness to propose hypotheses was fruitful. He knew something that cautious historians often forget—that nothing is evidence except for or against some hypothesis.

Collingwood's philosophical work falls roughly into three periods: (1) 1912–1927, his acceptance of idealism; (2) 1927–1937, his mature philosophy of the special sciences, conceived as resting on an idealist foundation; and (3) 1937–1943, his rejection of idealism. His ethical and political views will be discussed separately.

ACCEPTANCE OF IDEALISM

In his first book, *Religion and Philosophy* (London, 1916), Collingwood maintained three doctrines familiar to readers of his later work: (1) that creations of the human mind, no matter how primitive, must be studied historically, not psychologically; (2) that historical knowledge is attainable; and (3) that history and philosophy are identical. What he meant by this third doctrine depends on what he meant by "history" and by "philosophy"; in subsequent years he changed his mind about both.

In his *Autobiography* (London, 1939) Collingwood related that in 1917 a publisher rejected a manuscript, *Truth and Contradiction*, in which he had reached conclusions about truth and about the relation between history and philosophy that are characteristic of his thought at a much later period. Those conclusions are that truth or falsity does not belong to propositions but to complexes of questions and answers; that all such complexes rest on "absolute presuppositions" that are neither true nor false; and that since the business of philosophy is to elicit the absolute presuppositions held by different people at different times, philosophy is really a branch of history.

Since Collingwood destroyed the manuscript of *Truth and Contradiction* after writing his *Autobiography*, it is impossible to ascertain how closely the earlier work anticipated the later. However, in *Ruskin's Philosophy* (London, 1920), a lecture delivered in 1919, he asserted that a man's philosophy is "the [set of] principles which … he assumes in all his thinking and acting"; and he went on to maintain that since most men do not know what their philosophy is, "it is the attempt to discover what people's philosophy is that marks the philosopher." At least until 1919, therefore, Collingwood conceived of philosophy as a historical investigation of humankind's ultimate and largely unacknowledged principles, but it may be doubted whether Collingwood at that time denied that ultimate principles are either true or false. In *Ruskin's Philosophy* he sympathized with G. W. F. Hegel's refusal to accept as ultimate any dualism, whether of reason and understanding or of theory and practice. And two years later, in an essay, "Croce's Philosophy of History" (*Hibbert Journal* 19 [1921]: 263–278), he attacked Benedetto Croce for holding that philosophy was being "absorbed" into history, so that it is "cancelled out entirely as already provided for" by history. Collingwood did not then think that either history or philosophy in the ordinary sense could absorb the other but rather that each, if seriously pursued, leads to the other. He agreed with the "idealistic" Giovanni Gentile that they are "poised in equilibrium."

PUBLICATION OF *SPECULUM MENTIS*. *Speculum Mentis* (Oxford, 1924) was Collingwood's first attempt to construct a philosophical system. In it he critically reviewed five "forms of experience," ordered according to the degree of truth each attains.

Art. Art, the lowest form of experience, Collingwood defined after Croce as pure imagination, which he distinguished from sensation, on the one hand, assertion, on the other. Unlike sensation, imagination is active and has its own guiding principle, Beauty. "Beauty," however, must be defined in terms of imagination and not vice versa. As a form of experience, the deficiency of art is that while in itself a work of art is neither true nor false, it inevitably suggests assertions: It is expressive. Despite Croce's definition, then, imagination in art is in conflict with expression in art, and their conflict shows that art alone cannot satisfy the human spirit.

Religion. Art gives rise to religion, in which something imagined is affirmed as real. Like art, religion has its own guiding principle, holiness. The artistic consciousness does not affirm that what it imagines is real; but religion, even Christianity, which Collingwood considered its

highest form, affirms something imagined—a Father in heaven, the Real Presence in the sacrament, the resurrection of the dead—as real. These affirmations, Collingwood held, symbolize something true; but religion requires that they be affirmed in their symbolic form: "A philosopher would not be regarded as a Christian for subscribing to a statement which he declared to be a mere paraphrase of the Apostles' Creed in philosophic terms."

Christianity, by affirming the incarnation and atoning death of God, symbolizes the overcoming of the opposition between man and God. This unity of man with God symbolizes man's capacity to attain nonsymbolic, direct knowledge.

Science. Theoretical science is the first form of experience in which man tries by reason to grasp truth. But theoretical science, whether a priori as in mathematics or empirical as in natural science, is abstract. Natural science is the application of mathematics to the empirical world, conceived as subject to laws (mechanism) and composed of an ultimate undifferentiated stuff (materialism). But the world, as we experience it, is not merely mathematical, mechanical, and material. Theoretical science is therefore only supposition: Its truths are hypothetical. It can say truly, "If there were an *S*, there would be *P*," where *S* and *P* are events in a material world specified in mechanistic terms; but mechanistic terms are not unconditionally applicable to the world of experience. They are abstract; and to abstract is to falsify.

History. History appears to offer a way of escape from the abstractness of theoretical science; for it treats of the world of experience as a concrete temporal process. In their highest development, all theoretical sciences—physics and biology no less than the social sciences—assume a historical form. But history, too, has its characteristic deficiency. At bottom it is an extension of the historian's perception; and a perceived world is alien to its perceiver: a spectacle. Perception can never be knowledge because it can never grasp the whole historical process, and what is beyond the perceiver's ken may have implications for what is within it. Every specialist in a period is ignorant of a large part of what came before it, and his ignorance "introduces a coefficient of error into his work of whose magnitude he can never be aware." Even if this were not true, he could not escape the limitation of all attempts at knowledge in which subject and object are distinct. Since what is merely object is alien, it is falsified by the very process of appropriating it.

Philosophy. But one form of experience, philosophy, yields truth. Philosophy is self-knowledge. In it the distinction between knowing subject and known object vanishes. The self that is known is that which has attained all the subordinate forms of experience—art, religion, science, and history—and corrected their distortions. Philosophy has no positive content of its own: It is the awareness of what is true in those subordinate forms. In knowing their limitations it transcends them. Hence the absolute mind exists in the life of each individual mind to the extent that the individual mind raises and solves problems in any form of experience; as long as this process goes on, each mind is infinite. "The truth is not some perfect system of philosophy: It is simply the way in which all systems, however perfect, collapse into nothingness on the discovery that they are only systems."

MATURE PHILOSOPHY

PHILOSOPHY OF RELIGION. From 1924 to 1930, Collingwood further explored the positions of *Speculum Mentis*, especially those in aesthetics and religion. For the most part he remained content with his earlier theory of art, but in an essay, "Reason Is Faith Cultivating Itself" (*Hibbert Journal* 26 [1927]: 3–14), and a pamphlet, *Faith and Reason* (London, 1928), he abandoned the doctrine of *Speculum Mentis* that religion is essentially symbolic. Religion, he argued, can rid itself of superstition. Christianity correctly insists that there is a sphere of faith that transcends reason and is its basis. Neither the belief that the universe is rational nor that life is worth living can be established by scientific or ethical inquiry, yet they underlie natural science and rational ethics. Popular Christianity expresses those beliefs symbolically; but symbolizations are not essential to it. The ignorant believer who denounces philosophical or scientific paraphrases of Christian dogmas has no right to speak for Christianity.

PHILOSOPHY OF HISTORY. In his *Autobiography* Collingwood recorded that during the summer of 1928 he finally perceived the flaw that had vitiated his philosophy of history in *Speculum Mentis*. He presented his revised views in a pamphlet, *The Philosophy of History* (London, 1930). In 1936 he wrote the lectures that are the fullest statement of these views and that make up the greater part of his *Idea of History* (Oxford, 1946). The error he detected in *Speculum Mentis* was that the historical past is a spectacle, an object alien to the historian's mind. It has two roots: the realist error that knowing is fundamentally like perceiving; and the idealist error that the same thought cannot exist in different contexts. Against the realists, Collingwood maintained that every thought is an act that may be performed at different times and by different minds. A historian can know that Caesar

enacted a certain thought if he can reconstruct that thought in his own mind (so reenacting it) and demonstrate by evidence that his reconstruction is true of Caesar. Against the idealists, he maintained that, while some contexts change the character of a thought, others do not. The fact that, with my knowledge of modern geometry, I rethink one of Euclid's thoughts, for instance, the forty-fifth proposition of his first book, does not entail that my thought is different from Euclid's.

The key to Collingwood's conception of historical verification is his repeated declaration that historical method is "Baconian," a matter of putting evidence to the question. Given any piece of evidence, more than one reconstruction can be made of the action of which it is a relic. But each reconstruction, taken together with other knowledge, will entail consequences different from those of its fellows. A given reconstruction is established if no consequence that can be drawn from it conflicts with the evidence and if every other reconstruction has some consequence that does conflict with it. If a historian cannot show that one reconstruction, and only one, can be reconciled with the evidence, he must suspend judgment.

Historians must not only show what happened but also explain it. Collingwood proved that the two tasks are accomplished together. The past happenings that historians are concerned to discover are acts; and an act is a physical event that expresses a thought. To discover that an act took place includes discovering the thought expressed in it; and discovering that thought explains the act.

NATURAL SCIENCE. Just as in *The Idea of History* and in the writings that preceded it Collingwood had demolished the historical skepticism of *Speculum Mentis*, so in a set of lectures written in 1933–1934, which became *The Idea of Nature* (Oxford, 1945), he renounced his earlier skepticism about natural science and confessed that since "the knowledge acquired for mankind by Galileo and Newton and their successors … is genuine knowledge," philosophy must ask "not whether this quantitative material world can be known but why it can be known." His answer to that question, however, was equivocal. Collingwood named three constructive periods in European cosmological thought: the Greek, the Renaissance, and the modern, each with its characteristic view of nature. But he said curiously little about the question, "Why is one view of nature replaced by another?" In his introduction to *The Idea of Nature* he declared that "natural science must come first in order that philosophy may have something to reflect on," which suggests that views of nature change only as scientific thought changes; but in his exposition of the change from the Renaissance to the modern view of nature and in his criticisms of modern views, he often wrote as though philosophy might decide what is or is not a tenable view of nature without referring to natural science at all.

METAPHYSICS. Abandoning his earlier view that philosophy is no more than awareness of the limitations of subordinate forms of experience, Collingwood, in his *Essay on Philosophical Method* (Oxford, 1933), assigned philosophy the task of "thinking out the idea of an object that shall completely satisfy the demands of reason." He no longer rejected natural science and history as offering false accounts of such an object. Instead, he described each as limited in its aims. Natural science attempts to find true universal hypothetical propositions; history seeks true categorical propositions, but only about individuals in the world. The propositions of philosophy must be both categorical (about something existent) and universal (about everything existent). Hence, its object can only be the *ens realissimum*, the being that comprehends all being, of which all finite beings are appearances.

Although distinct from history, philosophy is nevertheless closely allied to it. Just as the various definitions that have been proposed for any philosophical concept constitute a scale of forms, of which the lower are appearances of the higher, so do the various metaphysical systems that purport to give an account of the *ens realissimum*. The way to knowledge in metaphysics is through critical reflection on its history.

REJECTION OF IDEALISM

AESTHETICS AS THEORY OF LANGUAGE. In 1937 Collingwood was invited to revise or to replace his *Outlines of Philosophy of Art* (London, 1925), in which he had largely followed the theory of art in *Speculum Mentis*. He chose to replace it; and his new book, *The Principles of Art* (Oxford, 1938), moved closer to Croce, whose article "Aesthetic" Collingwood had translated for the 1929 edition of the *Encyclopaedia Britannica*. Collingwood began by assuming that an aesthetic usage of the word *art* has been established in the modern European critical tradition and that it is the business of aesthetics to define what *art* so used means. The classical definition of art as representation (*mimesis*), in all its varieties, confounded art with craft (*techne, ars*), that is, with the production of something preconceived. Analysis shows that none of the classical definitions state either a necessary or a sufficient condition of art. Works of art may be, and commonly are,

also works of craft. But what makes something a work of art and determines whether it is a good or a bad one is not what makes it a work of craft.

A work of art is an imaginative creation; the function of imagination is to raise what is preconscious (for instance, mere feeling) to consciousness by giving it definite form. Since this activity is expression, Collingwood repudiated his earlier stand and accepted Croce's doctrine that imagination and expression are identical. He also accepted Croce's view that all expression, in any medium, is linguistic; for any form by which the preconscious is raised to consciousness is linguistic. Language thus begins in the cradle. Children speak before they learn their mother tongues.

The primitive language of the cradle is too narrow in range to serve the purposes of any but infants; it must be enriched by "intellectualizing" it so that it can express thoughts as well as feelings. An intellectualized language is one containing "conceptual" terms, and all conceptual thinking is abstract.

An intellectualized language does not cease to be expressive; rather its range of expressiveness is increased. Art is, therefore, not an activity cut off from, say, science. Every fresh linguistic utterance is imaginative and can be considered a work of art. Hence Croce was right when he said that there is poetry without prose, but no prose without poetry. And since it is the nature of art to be expressive, good art is successful expression. Bad art is the malperformance of the act of bringing preconscious thoughts and feelings to consciousness, a malperformance that misrepresents what is thought and felt. It can arise only in a corrupt consciousness. Critics can detect bad art, works of corrupt consciousness, by comparing them with successful works.

PHILOSOPHY OF MIND. In his last book, *The New Leviathan* (Oxford, 1942), Collingwood amplified and corrected the philosophy of mind he had outlined in *The Principles of Art*. Mind is consciousness, and while every act of consciousness has an object, no act of consciousness involves consciousness of itself. The various functions of consciousness are stratified into orders. The most primitive of them is consciousness of feeling. An act involving consciousness of a primitive act belongs to a higher order. Collingwood distinguished five such orders: primitive consciousness, appetite, desire, free choice, and reason. In principle, there is no upper limit to the orders of consciousness; for in reasoning about an act of reason a higher-order act is brought into being.

Holding that feeling (that is, sensation with its emotional charge) is not an act of consciousness, Collingwood denied that one can become conscious of an act of consciousness by introspection or inner sense. All acts of consciousness are linguistic; mind is the child of language. In analyzing the various forms of language, Collingwood reiterated his conclusion in *The Principles of Art* that conceptual thinking is abstract, and he expressly repudiated the idealist doctrine that to abstract is to falsify.

All theories of the relation between body and mind betray a philosophical misconception. Body and mind are not two related substances: They are man as investigated in two different ways, physiologically and historically. There is no conflict between physiology and history. To hold that Brutus's movement in stabbing Caesar can be investigated and explained physiologically does not imply that Brutus's act cannot be investigated historically nor does it detract from the value of a historical explanation of that act. Here Collingwood strikingly anticipated Gilbert Ryle's view as expressed in *The Concept of Mind* (New York, 1950).

LATER METAPHYSICS. In his *Autobiography* Collingwood reaffirmed his adherence to the conception of metaphysics as a historical science of absolute presuppositions which he claimed to have reached in *Truth and Contradiction*. In the *Essay on Metaphysics* (Oxford, 1940) he amplified this position. Every science, whether theoretical or practical, consists in asking and answering questions; and every sequence of questions rests ultimately on absolute presuppositions that are not answers to questions. Since truth or falsity belongs only to answers to questions, absolute presuppositions are neither true nor false. The task of metaphysics is to ascertain what is absolutely presupposed in a given society and how one set of absolute presuppositions has come to be replaced by another. Metaphysicians, however, must not criticize the absolute presuppositions they discover; for criticism presupposes that they are either true or false. A society does not consciously change its absolute presuppositions. Since most men are quite unconscious of their absolute presuppositions, any change in them is unconscious too and comes about because of internal strains.

Collingwood did not acknowledge what must have been obvious to his readers, that in the *Autobiography* and in the *Essay on Metaphysics* he had jettisoned the metaphysics of the *Essay on Philosophical Method*. His views in the *Essay on Metaphysics* are so incoherent that some sympathetic critics have ascribed his change of mind to

illness. (Both the *Autobiography* and the *Essay on Metaphysics* were written while he was recovering from a series of strokes.) However, his conception of metaphysics in the *Essay on Philosophical Method*, no less than his earlier conception in *Speculum Mentis*, rested on idealist doctrines from which he had been gradually freeing himself. He still believed that philosophical concepts are not abstract. The doctrine that philosophical propositions are both categorical and universal cannot be detached from the idealist theory of the concrete universal. But both in *The Principles of Art* (written before his illness) and in *The New Leviathan* Collingwood explicitly declared that all concepts are abstract.

Although in his *Autobiography* Collingwood repudiated his earlier idealist conception of philosophy, his views about religion, natural science, and history remained virtually intact. Nor were his views on art altered by his later historicism in metaphysics. This suggests that his change of mind in 1938 may be less fundamental than has been thought. After 1924 the main direction of Collingwood's thought was opposed to skepticism in the special sciences. His earlier skepticism had sprung from his idealistic rejection of abstract thinking and his conviction that philosophical thought is not abstract. By 1938 his work on the philosophy of art and the special sciences had overthrown both these errors, and it became clear that he could no longer hold the idealistic metaphysics of the *Essay on Philosophical Method*. It is natural that in seeking something to put in its place he reverted to his youthful historicism, and that it in turn proved inadequate. His inability to find a substitute for idealism does not show that he was mistaken in rejecting it; nor does it prejudice his achievements in aesthetics, philosophy of history, and philosophy of mind.

ETHICS AND POLITICS

In *Speculum Mentis* Collingwood recognized three forms of ethics: utilitarian, in which action is conceived as a means to an end; duty or concrete ethics, in which action is conceived as determined by the will to act in accordance with the moral order of the objective world; and absolute ethics, in which the distinction between the individual and society, and with it the sense of abstract law, disappears. The first form was held to be characteristic of science, the second of history, and the third of philosophy. Both in *Speculum Mentis* and in the *Essay on Philosophical Method* he represented the forms of ethics on a scale in which the higher forms complete and correct the lower.

Collingwood never renounced this triadic scheme, although in *The New Leviathan* he proposed a new view of the connection between morality and theoretical science, namely, that theoretical science reflects moral practice. Teleological science reflects utilitarian morality; "regularian" science reflects a morality of law; and history reflects the concrete morality of "duty."

In *The New Leviathan* Collingwood set out to bring the "classical politics" of Thomas Hobbes and John Locke up to date. He accepted the classical conception of politics as bringing men out of a state of nature into a state of civil society. Essentially, political life is a process in which a nonsocial community (i.e., the state of nature) is transformed into a social one. This cannot happen unless the rulers understand that social life is a life in which people freely engage in joint enterprises. Civilization is "a process whereby a community undergoes a change from a condition of relative *barbarity* to one of *civility*." Barbarism is hostility to civilization; but although barbarous communities always strive to destroy civilized ones, in the long run the defeat of barbarism is certain.

See also Aesthetics, History of; Art, Expression in; Croce, Benedetto; Determinism in History; Galileo Galilei; Gentile, Giovanni; Hegel, Georg Wilhelm Friedrich; Historicism; Hobbes, Thomas; Idealism; Imagination; Locke, John; Newton, Isaac; Philosophy of History; Presupposing; Renaissance; Ruskin, John; Utilitarianism.

Bibliography

The obituary essays by R. B. McCallum, T. M. Knox, and I. A. Richmond in *Proceedings of the British Academy* 29 (1943): 463–480, together with T. M. Knox's editorial preface to Collingwood's *The Idea of History*, are indispensable and contain full bibliographies. E. W. F. Tomlin, *R. G. Collingwood* (London: Longmans Green, 1953) is useful but elementary. Alan Donagan, *The Later Philosophy of R. G. Collingwood* (Oxford: Clarendon Press, 1962) treats of Collingwood's work after 1933.

OTHER RECOMMEDED WORKS

Collingwood, R. G. *Roman Britain*. Rev. ed. Oxford: Clarendon, 1966.

Collingwood, R. G. *Essays in the Philosophy of Art*. Bloomington, Indiana University Press, 1966.

Collingwood, R. G. *Faith & Reason; Essays in the Philosophy of Religion*. Edited by Lionel Rubinoff. Chicago, Quadrangle Books, 1968.

Collingwood, R. G. *Essays in Political Philosophy*. Edited by David Boucher. Oxford: Clarendon, 1989.

Alan Donagan (1967)
Bibliography updated by Michael J. Farmer (2005)

COLLINS, ANTHONY

(1676–1729)

Anthony Collins, the English deist, freethinker, theologian, and philosopher, was born at Hounslow, near London, the son of Henry Collins, a well-to-do gentleman. Anthony Collins was educated at Eton and at King's College, Cambridge, and for a while was a student in the Temple. This training in the law later enabled him to maintain an excellent reputation for many years as justice of the peace and deputy lieutenant in Middlesex and in Essex. He was married twice to daughters of the landed gentry. A devoted admirer of John Locke both as philosopher and as writer on religion, Collins, aged twenty-seven, made the pilgrimage to Oates early in 1703 to meet the master, then aged seventy. They were strongly attracted to one another. Later that year Locke wrote poignantly to Collins: "You complain of a great many defects [in yourself] and that complaint is the highest recommendation I could desire to make me love and esteem you and desire your friendship. And if I were now setting out in the world, I should think it my great happiness to have such a companion as you, who had a true relish of truth … and, if I mistake not you have as much of it as I ever met with in anybody." In his will Locke left Collins a legacy of £110 and some books and maps, and named him one of three trustees of his estate. Collins arranged tributes to the master that appeared in 1708 as *Some Familiar Letters between Mr. Locke and several of his friends* and in 1720 as *A Collection of Several Pieces of Mr John Locke, published by M. Des Maizeaux under the direction of Mr Anthony Collins*.

By that time Collins had made a lasting, if at the time a notorious, name for himself through a series of outspoken yet restrained publications, all of which were anonymous (although most sophisticated readers were aware of the author's identity). The more important include *An Essay concerning the Use of Reason in Propositions, the Evidence wherof depends upon Human Testimony* (1707); *Priestcraft in Perfection: Or, A Detection of the Fraud of Inserting and Continuing this Clause (The Church hath Power to Decree Rites and Ceremonys, and Authority in Controversys in Faith) In the Twentieth Article of the Articles of the Church of England* (1710); *A Discourse of Free-Thinking, Occasion'd by the Rise and Growth of a Sect call'd Free-Thinkers* (1713; actually published late in 1712); *A Philosophical Inquiry concerning Human Liberty* (1715).

In 1711 Collins made the first of many visits to Holland, where he met numerous men of intellect. Soon after the appearance of the *Discourse of Free-Thinking*, with accompanying public uproar, Collins visited Holland briefly, possibly for reasons of prudence. His later major works include *A Discourse of the Grounds and Reasons of the Christian Religion* (1724), which elicited thirty-five replies within two years and which Bishop Warburton later named one of the most plausible books ever written against Christianity, admitting that the replies might have been left to confute one another; *The Scheme of Literal Prophecy Considered* (The Hague, 1725; London, 1726), a sequel to the *Discourse; A Discourse concerning Ridicule and Irony in Writing* (1727); and the *Dissertation on Liberty and Necessity* (1729). This last, together with the earlier *Philosophical Inquiry concerning Human Liberty*, constitutes a powerful statement of the doctrine of necessitarianism. By and large, it is to be noted, the English deists upheld the freedom of the will.

During all this time Collins carried on vigorous, frequently witty, controversies with—to name but a few—Henry Dodwell the elder, famous nonjurist; and such clerical antagonists as Richard Bentley, the classical scholar; Samuel Clarke, the rationalist; and William Whiston, the biblical literalist. His health weakened by repeated attacks of the stone, Collins died late in 1729 and was buried in Oxford chapel. It is said that despite a lifetime of controversy, he was never attacked on the basis of his character. Collins represents the philosophical skeptic in the true sense of the word.

FREETHINKING

The right and the necessity to inquire freely and fearlessly into all subjects, especially religion, was Collins's constant and fundamental thesis. Its master statement is the *Discourse of Free-Thinking*, but it was adumbrated in two earlier works. The *Essay concerning the Use of Reason* makes the point that reason is "that faculty of the Mind whereby it perceives the Truth, Falsehood, Probability or Improbability of Propositions." Truth and falsehood are known rationalistically and are certain. Probability may take the form of opinion when discovered by reason or of faith when perceived by testimony. Testimony is the foundation of much of our knowledge but can never impugn the natural (rationalistic) notions implanted in the mind of man. The Bible, consequently, is not to be taken seriously when it portrays God in human terms; certain parts of the Bible are to be accepted, while others are to be rejected. Thus, Collins combined Locke's arguments for the reasonableness of Christianity and morality and religious principles with the rationalistic Common Notions of Lord Herbert of Cherbury. *Priestcraft in Perfection* carried the attack, common to most deists of the eighteenth

century, against the dogmas of established churches. Such dogmas, Collins argued, must be viewed as fraudulent when contrary to reason. The appeal to mystery and to things above reason simply will not do.

The title page of the *Discourse of Free-Thinking* is embellished with several quotations: one from the Old Testament, one from the New Testament, one from Cicero, and one from the earl of Shaftesbury. The influence of Shaftesbury is apparent throughout, but Collins was less hesitant to employ the method of ridicule (as is fully attested in the *Discourse concerning Ridicule and Irony in Writing*). The general definition of the right to think freely was applied mainly to religion. Collins pointed out that the new science and the new philosophy had exposed many errors of the past; the Reformation was the result of fearless thinking on the part of a few leaders; the abundant literature of travel exposed the superstitions of peoples throughout the world and also the infinite numbers of pretenders to divine revelation. Freedom had exorcised the witches that so plagued James I and Charles I: "great numbers of witches have been almost annually executed in England, from the remotest antiquity to the late Revolution; when the liberty given and taken to think freely, the Devil's power visibly declin'd, and England as well as the United Provinces ceas'd to be any part of his Christian territories." (The "Witches Act" of 1603 was to be repealed in 1736.)

With tongue in cheek, Collins suggested that the Society for Propagating the Gospel in Foreign Parts was really a freethinking organization because infidels must be asked to examine and to reject their native traditional religions in order to accept true religion. He further suggested that such zealous divines as Francis Atterbury, George Smalridge, and Jonathan Swift be drafted annually for this enterprise in the same manner as "military missionarys." The argument then turned against the priests of all ages who are responsible for quibbling about biblical interpretations and end up calling one another atheists. The Bible, Collins continued, is clearly replete with corrupted texts—30,000 in the New Testament alone, according to one authority. Its text, therefore, is to be examined in the same scholarly and critical manner as the texts of all ancient books. The *Discourse* concluded with a refutation of the standard objections to freethinking. Atheism is not, after all, the worst of all evils; enthusiasm and superstition hold that title, according to Francis Bacon. Cicero was quoted to confute the claim that some false ideas are necessary for the good of society (an early version of the Marxian notion of religion as the opiate of the people). A long list of freethinkers was given, including Socrates, Plato, Aristotle, Epicurus, Plutarch, Cicero, and Seneca among the ancient pagans; Solomon and the prophets of the Old Testament; Josephus, the Pharisee; Origen, the Church Father; and, among the moderns, Bacon, Thomas Hobbes, and Archbishop Tillotson ("whom all English free-thinkers own as their head"). Collins then asserted that he might well have added other names, such as Michel Eyquem de Montaigne, René Descartes, Hugo Grotius, Richard Hooker, Lord Falkland, Lord Herbert of Cherbury, John Milton, Ralph Cudworth, Sir William Temple, and the master, Locke. All enemies of freethinking were branded crackbrained and enthusiastical, malicious, ambitious, inhumane, ignorant, or brutal—or courters of priests, women, and the mob.

BIBLICAL CRITICISM

The Discourse of the Grounds and Reasons of the Christian Religion and the *Scheme of Literal Prophecy Considered* follow the rational, scholarly methods for biblical criticism described earlier, but concentrate on the question of the fulfillment of Old Testament prophecies in the New Testament. The most cogent attacks are on the virgin prophecy in the book of Isaiah and the unusually specific prophecies in the book of Daniel. In both works Collins pursued the theme of the necessity of thinking freely and went out of his way to defend the right of Whiston, one of his chief adversaries, to think freely—although wrongly, as he saw it—about prophecy. Whiston was a literalist, and Collins had no great difficulty and no little sport in pointing out the absurdities to which Whiston was driven. Collins himself had promised to investigate the miracles of the New Testament but was unable to do so before his final illness and death, and the task fell to Thomas Woolston.

Like John Toland before them, Collins and Woolston forced the issue of the scriptural canon upon the orthodox and opened the way in England for historical criticism.

See also Deism; Woolston, Thomas.

Bibliography

Attfield, Robin. "Clarke, Collins and Compounds." *Journal of the History of Philosophy* 15 (1977): 45–54.

Berman, David. "Anthony Collins: His Thought and Writings." *Hermathena* (1975): 49–70.

Bibliotheca Anthony Collins. London, 1732.

Broome, J. H. "Une Collaboration: Anthony Collins et Desmaizeaux." *Revue de littérature comparée* 30 (1956): 161–179.

Cranston, Maurice. *John Locke, A Biography*, 460–477. London: Longmans Green, 1957. The Collins-Locke friendship.

Disraeli, Isaac. *Curiosities of Literature*. Vol. III, 333–343. Boston: Veazie, 1860.

Dybikowski, James. "Anthony Collins (1676–1729)." In *British Philosophers, 1500–1799 (Dictionary of Literary Biography, Vol. 252)*, edited by Peter Fosl. Farmington Hills, MI: Thomson Gale, 2002.

Dybikowski, James. "Anthony Collins' Defense of Free-Thinking." In *Scepticism, Clandestinity and Free-Thinking*, edited by Gianni Paganini. Paris: Champion, 2002.

Hahn, Joseph. *Voltaires Stellung sur Frage der menschlichen Freiheit in ihrem Verhältnis sur Locke und Collins*. Borna and Leipzig, 1905.

Nichols, John. *Illustrations of the Literary History of the Eighteenth Century*. Vol. II, 148–150. London: Nichols and Bentley, 1817.

Snobelen, Stephen. "The Argument over Prophecy: An Eighteenth-Century Debate between William Whiston and Anthony Collins." *Lumen* 15 (1996): 195–213.

Thorschmid, Urban G. *Critische Lebensgeschichte Anton Collins*. Dresden and Leipzig, 1755.

Torrey, Norman L. *Voltaire and the English Deists*. New Haven, CT: Yale University Press, 1930. Ch. 3, "Voltaire and Anthony Collins: Metaphysics and Prophecies."

See also the general bibliography to the *Deism* entry.

Ernest Campbell Mossner (1967)
Bibliography updated by Tamra Frei (2005)

COLORS

The phenomena of "color" pose a special puzzle to philosophers characterizing the mind, the world, and the interaction of the two. In various ways, both subjective and objective, both appearance and reality, color has been the subject of wide disagreement. Besides the extreme view that colors are literally sensations—which would imply that they are not in the category of properties and that they last precisely as long as sensations—the main views are these.

PHYSICALISM

D. M. Armstrong, J. J. C. Smart, and others have suggested that red (or being red), for example, is a physical property—perhaps a surface physical characteristic (like Robert Boyle's "textures")—or a propensity to reflect some kinds of light more than others. The threat that physical science might be unable to find a predicate coextensive with "red" seems small; but there are challenges to the idea that any such property can be identified with red. First, will a physical property have the same higher-level properties as red does? Red is a "unique" color—there is a "pure" shade of red with no hint of any other color (unlike orange, every shade of which evidently contains red and yellow); however, it seems nonsense to say that some reflectance characteristic is "unique." The physicalist may perhaps reply: A reflectance characteristic can indeed have the property of "uniqueness"—if that is understood as the property of *suggesting to a normal observer no hint of any other color.* (That higher-level property will no doubt be the subject of a later reduction.) A second challenge is this: Ordinary people surely know, for example, that red is more similar to orange than to blue, but if colors are properties whose true nature is revealed only in science, then (until they know more science) they should be in no position to know this. The physicalist may have a reply: This kind of knowledge is of phenomenal similarity, not physical similarity—and on that ordinary perceivers are authoritative. Both challenges suggest an important point, however—that physicalism can at best be a theory about properties that we think of initially without any thought of physical science.

DISPOSITIONAL VIEWS

The view that colors are dispositions to produce experiences has long been nearly an orthodoxy in the field. Proposed by Boyle and John Locke it has seemed a perfect way to capture the connections between color concepts and color experience. You cannot, it seems, grasp the idea of red unless things sometimes look red to you. And you cannot have a full grasp of the idea unless you realize that your color judgments will be defeasible if it turns out that either you or the conditions are abnormal. The proposal may be strengthened by adding an actuality operator: To be red, an object needs to look red to such observers and in such conditions as *actually* count as normal. This last phrase shall be abbreviated as "to look red [etc.]."

A preliminary worry can perhaps be met. Are there any such things as "normal conditions" and "normal observers"? Normal conditions vary hugely with the nature of the object and with our interests; in some cases (e.g., bioluminescent fish) there may be no clear answer to the question what normal conditions are. "Normal observers" pose a further problem: Even when we rule out "color-blind" people, there is surprising disagreement among the remainder (e.g., over which shade of green is "unique"). These problems may not be fatal. If there is indeterminacy in the truth of "*x* is disposed to look red to normal observers under normal conditions," there may be an exactly corresponding indeterminacy in the truth of "*x* is red"—the moral may be that some things have no determinate color, not that color is mischaracterized by the dispositional thesis.

Dispositional views vary according to whether they take the experience of a thing's looking red to be a sensation or a representation. The sensationalist version faces the suspicion that the required "sensations of red" (or the "red' regions of the visual field" in C. Peacocke's language) are mythical creatures of a modern-day sense-datum theory. The view also implies that when an object looks red, it looks disposed to produce red' regions in the visual field. And that seems excessively sophisticated.

The representational version has a related problem: if "red" literally means "disposed to look red [etc.]," then "looks red" will have to mean "looks disposed to look red [etc.]"—which is surely false. This—like related objections about circularity—shows that "red" cannot mean "disposed to look red [etc.]"; but it may not rule out a nonobvious identity of redness and the disposition to look red [etc.], or an a priori necessary coextensiveness.

A final challenge—for both versions of the dispositional view—is more serious. Imagine a yellow object that also emits death rays, so that anyone who looks at it is killed before he can see its color. The object will be yellow but have no disposition to produce experiences as of yellow in normal observers. (The example is due to Saul Kripke.) One can indeed insist that the object would look yellow to normal observers if only we masked the death rays. But we need to mask the death rays without masking or changing the color. And there is no knowledge of what that amounts to, independent of a substantial conception of what color is. We may believe, for example, that the color of a surface is a matter of the way it reflects incident light; so we can change and mask anything that leaves intact the object's *way of changing incident light.* But if we have that belief, it is no thanks to the definition of yellow as simply "the disposition to look yellow [etc.]." Our prime conception of color must have a different source.

VIEWS AVAILABLE

If the physicalist and dispositional views can at best be true with respect to properties first identified by some other route, then we need a new account of our thought about color and of the object of that thought. If color thinking contains an error, the options are projectivism and eliminativism; more easily overlooked is the possibility that color thinking may contain no error and a nonreductive simple realism be the appropriate view.

PROJECTIVISM

Galileo Galilei and (at times) René Descartes and Locke are the first of many to treat colors as properties of experiences, which we wrongly "project" onto external objects. Attractive though the view is, it faces two tasks. It must establish its right to a sensational conception of color vision; and it must clarify what exactly is meant by "projecting" a sensation. The difficulty is to find a precise account of projection that does not make the process so absurd that humans could not commit it or so innocent that it is not actually a mistake.

REPRESENTATIONAL ERROR THEORY AND ELIMINATIVISM

Some have suggested that color vision is representational—color vision involves the apparent representation of properties of external objects, but there is in fact no suitable external referent. C. L. Hardin has a related view: colors are properties neither of external objects nor of experiences. Colors are to be "eliminated," though there remain "chromatic perceptual states," which are to be reduced to neural states.

The strengths of these views must lie in the careful analysis of what is involved in naive thought about color. If naive thought makes fundamental assumptions that are false, then error theory must be the right conclusion. But an everyday commitment to the notions of normal observers and normal conditions may (as we have seen) not be disastrous. Incoherence in everyday color thought may have to be sought elsewhere.

AUTONOMY VIEWS

If color experience apparently represents features of physical objects, what is to prevent us from saying that (in ordinary cases) it correctly represents features of those things, namely colors? These colors would need to be supervenient upon physical properties, though they might or might not be reducible to them. (The model might be Davidson's or Jerry Fodor's view of mental properties.) Colors would have their place in a scheme of explanation that was autonomous with respect to physics, in that its legitimacy was not dependent upon ratification by physics. And that explanatory scheme would no doubt make connections between the colors we see and the contingencies of our perceptual system—thus making those colors not only genuine features of external objects but also in a certain way subjective and relative.

The view needs to overcome the suspicions that the only genuine properties are those recognized in physics, or those intelligible from an "absolute" point of view. A defense is needed of the idea that the world (and not just the mind) can contain subjective items, and an account of

the mind's thought about such items. Until these tasks are achieved the autonomy view will at best be programmatic. If they cannot be achieved, the option seems to be an error theory. They are large tasks, central in the philosophy of mind and metaphysics, and it is a measure of the difficulty of the topic that they have taken so long to come clearly to light.

See also Armstrong, David M.; Boyle, Robert; Davidson, Donald; Descartes, René; Eliminative Materialism, Eliminativism; Galileo Galilei; Kripke, Saul; Locke, John; Metaphysics; Philosophy of Mind; Physicalism; Projectivism; Qualia; Smart, John Jamieson Carswell.

Bibliography

Armstrong, D. M. *A Materialist Theory of the Mind.* London: Routledge and Kegan Paul, 1968. Chap. 12.

Boghossian, P., and J. D. Velleman. "Colour as a Secondary Quality." *Mind* (1989): 81–103.

Hardin, C. L. *Color for Philosophers: Unweaving the Rainbow.* Indianapolis: Hackett, 1988.

Hilbert, D. R. *Color and Color Perception: A Study in Anthropocentric Realism.* Stanford, CA: Center for the Study of Language and Information, 1987.

Peacocke, C. *Sense and Content.* Oxford: Clarendon Press, 1983. Chap. 2.

Westphal, J. *Colour.* Oxford: Blackwell, 1987; 2nd ed., 1991.

Justin Broackes (1996)

COMBINATORY LOGIC

Combinatory logic is a branch of mathematical logic that analyzes certain processes, such as substitution, which are associated with variables. These processes are taken for granted in most formulations of logic, but they are complex, and since a fundamental part of the resulting theory is recursively undecidable the analysis is not trivial. Combinatory logic contributes to simplifying the ultimate foundations of mathematical logic and to explaining the paradoxes; it contains an arithmetic in which exactly those numerical functions that are partial recursive are representable; and it has potential applications to the deeper study of such areas as logical calculuses of higher order, computer programming, and linguistics.

Before one can define combinatory logic precisely, it is necessary to explain some notions concerning formal systems. This will be done in the next section. In the following section the definition will be given and a plan presented according to which the later sections of this article will develop the subject. Each technical term is defined by the context in which it appears in italics.

FORMAL SYSTEMS

Consider a formal system of the following type: There is a class of formal objects, or *obs,* constructed from certain primitive obs, or *atoms,* by certain *operations*; each ob has a unique such construction. Among these operations a binary one, called *application,* is singled out. If this is the only operation, the system is called *applicative*; otherwise it is *quasi-applicative.* There is a unique unary predicate, symbolized by the sign "$\vdash$" used as prefix; the *elementary statements* are then of the form

(1) $$\vdash X,$$

where X is an ob. The *elementary theorems* form an inductive subclass of the elementary statements; they are generated from certain initial ones, the *axioms,* by *deductive rules.* The atoms, obs, elementary statements, and axioms are definite classes—that is, it can be effectively ascertained whether a proposed member of one of them is actually a member—but concerning the elementary theorems it is required only that the correctness of a derivation by the deductive rules can be effectively checked. Combinatory logic takes such a system as basis. Other sorts of system exist, but all those ordinarily used in mathematical logic can be reduced to the above type.

Assuming such a system, we observe the following conventions: The application of X to Y is symbolized as (XY). Parentheses are omitted according to the rule of association to the left and also to the rule that outside parentheses are superfluous, so that $XY_1Y_2Y_3$ is the same ob as $(((XY_1)Y_2)Y_3)$. A *combination* of given obs is an ob formed from some or all of them by application alone. The sign "$\equiv$" stands for definitional identity; "$\rightarrow$" and "$\rightleftarrows$" for metatheoretic implication and equivalence, respectively. Finally, "$=$" is defined, say, by

(2) $$X = Y \rightleftarrows \vdash \mathbf{Q}XY,$$

where $\mathbf{Q}$ is a specific ob, the axioms and rules being such that equality has the appropriate properties.

With such a formal system one associates two sorts of ontology. On the one hand, some persons insist on describing more definitely what the obs are; on the other hand, one may give a description of the meaning one intends for the elementary statements. The first description will be called a *representation*; the second description will be called an *interpretation.*

For a representation it is customary to state that the obs are words in an object language. We will not do that here—all symbols belong to the *U-language* (metalanguage)—but it can be done quite easily for any given

object language with two or more symbols. This permits a certain freedom in regard to use and mention.

An interpretation for combinatory logic may be described as follows (this is for motivation only and does not imply a commitment to any special type of metaphysics): One associates with certain obs *contensive* (known from prior experience) notions called *interpretants.* The fact that Y is the interpretant of X will be expressed simply as X means Y. Then if X means a function and Y means a possible value for the first assignment of that function, XY will mean the result of assigning the intepretant of Y as value to the first argument of X. Thus, if X means the addition function of natural numbers and Y means the number 1, then XY will mean a form of the successor function, and if Z means the number 2, then XYZ will mean the number 3. This device reduces many-place functions to unary ones without postulating ordered pairs. An elementary statement (1) will mean that X means an asserted statement; the interpretation is a valid one when every asserted statement is true.

DEFINITION AND DIVISIONS OF COMBINATORY LOGIC

The usual formal systems contain a special class of obs, usually atoms, called (*formal*) *variables.* These are so named in the formalization and play a special role, such that arbitrary obs can be substituted for them (perhaps under restrictions). Variables do not have interpretants; obs containing them mean functions in which they stand for arguments. Thus, the elementary statements of *Principia Mathematica,* Sec. 1A, are not about *p, q, r* but about negation (~) and alternation ∨); the interpretants of its elementary theorems state rather complex relationships, indicated by the variables, between these functions.

Let $\mathscr{H}$ be a system as defined earlier, and let $\mathscr{H}(x_1, \cdots, x_m)$ be the system formed by adjoining $x_1, \cdots, x_m$ as variables—that is, as new atoms—without further changes. As stated above, the natural interpretant of an ob M of $\mathscr{H}(x_1, \cdots, x_m)$ is that function over $\mathscr{H}$ whose value for arguments $a_1, \cdots, a_m$) is the result of substituting $a_1, \cdots, a_m$ for $x_1, \cdots, x_m$, respectively, in M. Let us say that an ob X of $\mathscr{H}$ designates M when and only when

$$Xx_1x_2 \cdots x_m = M \qquad (3)$$

is derivable in $\mathscr{H}(x_1, \cdots, x_m)$. The system $\mathscr{H}$ is called *combinatorially complete* when and only when such an X exists for every M. A *constant* (that is, an ob of an $\mathscr{H}$ containing no variables) X is called a *proper combinator* when and only when it designates a combination of variables alone; a *combinator* is any combination of proper combinators.

Combinatory logic may now be defined as that branch of logic which studies combinators. This is tantamount, at least for applicative systems, to studying combinatorial completeness.

There are two methods of achieving combinatorial completeness. The first is to postulate a designator for every M. This idea leads to the *theory of λ-conversion,* which is discussed in the next section. It is a quasi-applicative system with bound variables. The other method is to exhibit all combinators as combinations of certain atomic ones, after which we can get along with an applicative system. This leads to *synthetic combinatory logic,* to which the rest of the article is devoted. The two approaches have been shown to be equivalent.

The subject of combinatory logic divides itself into two parts in another way. In the first part, called *pure combinatory logic,* one introduces no constant atoms except combinators and those atoms necessary to define equality and pays no attention to whether the obs have interpretants. In the second part, called *illative combinatory logic,* one introduces atoms meaning other logical notions, such as implication, quantification, and semantical categories. The question whether an ob has an interpretant, and if so, what sort of interpretant, belongs to the illative theory.

THEORY OF λ-CONVERSION. In the theory of λ-conversion we postulate that given $M, x_1, \cdots, x_m$, there is an X in $\mathscr{H}$ such that (3) holds. This X, in Alonzo Church's notation, is $\lambda x_1 \cdots x_m M$. It suffices to postulate this for $m = 1$, for we can define

$$\lambda x_1 \cdots x_m x_{m+1} M \equiv \lambda x_1 \cdots x_m (\lambda x_{m+1} M).$$

Thus, we need only a binary operation forming λxM from x and M. This operation is the only primitive operation besides application. Thus, x is a variable and is bound (in a natural extension of the usual sense) in λxM. One must distinguish, just as in predicate calculus, free and bound occurrences of variables. One understands "ob of $\mathscr{H}$" to include any ob formed from atoms of $\mathscr{H}$ and variables without free occurrences of variables not in $\mathscr{H}$. Further, given an ob M, a variable x, and an ob N, we define $[N/x]M$ (subject to restrictions to prevent confusion of bound variables) as the ob obtained by the substitution of N for x in M.

In view of the intended interpretation, the following are acceptable (subject to the stated restrictions) as axiom schemes:

$$(\alpha) \qquad \lambda xM = \lambda y[y/x]M,$$

$$(\beta) \qquad (\lambda xM)N = [N/x]M.$$

Along with this one has the rules for equality, which give as a special case

$$(\xi) \qquad M = N \rightarrow \lambda xM = \lambda xN.$$

The equality relation is called *convertibility,* and "cnv" is often used instead of "=." We call (β) (as well as η and δ below) a *replacement scheme.* The definition is equivalent to saying that X cnv Y when and only when X can be converted to Y by zero or more successive applications of replacement schemes in either direction. There is also defined a relation of *reducibility,* indicated by "red," in which the replacement schemes can be used from left to right only. An ob is said to be in *normal form* when and only when no replacement scheme can be so applied to it.

There are various modifications of this system. In λI-*conversion* (the original λ-conversion), λxM is defined only when M contains a free occurrence of x; in λK-*conversion* this restriction is dropped. Again the additional axiom scheme (for x not free in U)

$$(\eta) \qquad \lambda x(Ux) = U$$

is acceptable from interpretations which maintain a strong extensionality principle. If it is adopted, the theory is here called *$\lambda\eta$-conversion,* in contrast to the original *$\lambda\beta$-conversion.* Finally one may introduce axiom schemes (δ) which single out special constants $\delta_1, \delta_2, \cdots$ and allow constants of the form $\delta_k U_1 U_2 \cdots U_{nk}$, where n_k is fixed by δ_k and the U_j are in normal form, to be replaced by other constants determined in some uniform manner. Note that (δ) is, in principle, illative.

The various forms of λ-conversion have differences in interpretation. In λI-reduction no component is dropped; hence, if X has a normal form, so does every part of X. This is not true for λK-reduction, a disadvantage if one identifies possession of an interpretant with having a normal form. Again, if one accepts (η), every ob means a function (in some sense), and sometimes this is unacceptable. However, one may prefer to make such distinctions in the illative theory.

The principal result concerning λ-conversion is the Church–Rosser theorem. This states that if X cnv Y, then one can find effectively a Z such that X red Z and Y red Z. Thus, two different combinations of variables are never interconvertible; this establishes consistency. In 1936, Alonzo Church and J. B. Rosser ("Some Properties of Conversion," in *Transactions of the American Mathematical Society* 39, pp. 472–482) proved the theorem for $\lambda I\beta$-conversion; it has since been extended to all forms of λ-conversion.

The decision problem for all equations $X = Y$ was shown by Church in 1936 ("An Unsolvable Problem of Elementary Number Theory," in *American Journal of Mathematics* 58, pp. 345–363) to be recursively unsolvable, as was the problem of determining whether X has a normal form. This result was the basis of Church's later proof of the recursive unsolvability of the decision problem for predicate calculus.

Since every kind of λ-conversion is equivalent to a synthetic theory and vice versa, the results described below for the synthetic theory are also results of λ-conversion and in some cases were first so obtained.

FOUNDATIONS OF PURE SYNTHETIC THEORY. Table 1 contains a list of special combinators. The names assigned to the combinators are in the first (X) column and the values of m and M to be used in equation (3) are in the second and third columns. The other columns will be explained later. In the various formulations certain of the combinators will be atomic; the corresponding equations (3) will then be assumed as axiom schemes in which 'x_1', $\cdots$, 'x_m' stand for arbitrary obs.

We seek to define, for arbitrary M, $x_1, \cdots, x_m$, an X such that (3) holds. The X so defined will be $[x_1, \cdots, x_m]M$; this means the same thing as $\lambda x_1 \cdots x_m M$ but is a defined, not a postulated, concept. One way of defining it is to use an induction on m, as above, and then, for $m = 1$, an induction on the structure of M. The latter can be obtained, for instance, by defining X to be **K**M when M does not contain x, to be **I** when M is x, and to be **S**X_1X_2 when $M \equiv M_1M_2$ and we have already defined $X_1 \equiv [x]M_1$, $X_2 \equiv [x]M_2$. Such an algorithm defines all combinators in terms of **I**, **K**, **S** as atoms; the definitions are very long, but suitable modifications improve matters. The fourth column of the table gives some definitions obtained by suitably modified algorithms. Other modifications give definitions in terms of **I**, **B**, **C**, **S** for all cases where M actually contains x; these are suitable for an analogue of λI-conversion.

Thus, we get a definition for $[x]M$ compatible with any of the forms of λ-conversion if we postulate schemes (3) as stated, together with the properties of equality. The analogues (with $[x]M$ in place of λxM) of (α) and (β) will then hold. But we do not have the analogue of (ξ), nor do we have an extensionality principle

$$(\zeta) \qquad U_1x = U_2x \rightarrow U_1 = U_2$$

even under the restrictions that are appropriate for $\lambda\beta$-conversion. One can obtain these properties by adjoining a finite number of *combinatory axioms.* Examples of these axioms are

(4) $\mathbf{SK} = \mathbf{KI}$,

(5) $\Psi\mathbf{SK} = \mathbf{BK}$.

Given a form of λ-conversion, we can choose these axioms so that there is a many–one mapping of the resulting system into the λ-conversion and another one vice versa, such that an equation in either system is a theorem exactly when its image is in the other. Thus, λ-conversion and the synthetic theory are equivalent. Bruce Lercher, in 1963, extended these considerations to include (δ).

It is possible to define, in several ways, a combinator **Y** such that for any *X*, $\mathbf{Y}X = X(\mathbf{Y}X)$. If Γ means negation, then **Y**Γ means the same as its own negation. For $\mathbf{Y} \equiv \mathbf{WS}(\mathbf{BWB})$, this is the notion at the root of the Russell paradox. Thus, in a combinatorially complete system one cannot exclude the paradoxes; one must explain them in the illative theory.

In the foregoing, equality can be taken as primitive. Then the axioms consist of the combinatory axioms, all instances of the reflexive law, and (3) (for atomic combinators); the rules are the usual rules for equality. When we press the analysis deeper so as to define equality by (2), the schemes (3) become rules; for example, that for **S** gives the pair of rules (one in each sense)

(6) $\vdash U(\mathbf{S}XYZ) \rightleftarrows \vdash U(XZ(YZ))$,

whereas reflexivity can come from an axiom. The result is a system with a finite number of axioms (no axiom schemes) and about a dozen rules, each with one or two premises and otherwise no more complex than (6) and such that the premises uniquely determine the conclusion. There are also only a finite number of atoms—variables are used only in the metatheory—and the single operation of application. The structure is therefore very simple. But all functions of variables can be performed therein, and with suitable illative additions it can form a basis for almost any logical system.

COMBINATORY ARITHMETIC. From the formal standpoint the natural numbers are constructions from a single atom, 0, by a single unary operation, σ. On this basis one can develop the usual recursive arithmetic, and one can explain how to count. Assume that such a system is given, represented, say, in the words in some alphabet

TABLE 1

X	*m*	*M*	*Definition*	*FX*
B	3	$x_1(x_2x_3)$	**S(KS)K**	$\mathbf{F}(\mathbf{F}\beta\gamma)\mathbf{F}(\mathbf{F}\alpha\beta)(\mathbf{F}\alpha\gamma))$
C	3	$x_1x_3x_2$	**S(BBS)(KK)**	$\mathbf{F}(\mathbf{F}\beta(\mathbf{F}\alpha\gamma))(\mathbf{F}\alpha(\mathbf{F}\beta\gamma))$
I	1		———	$\mathbf{F}\alpha\alpha$
K	2		———	$\mathbf{F}\alpha(\mathbf{F}\beta\alpha)$
S	3	$x_1x_3(x_2x_3)$	———	$\mathbf{F}(\mathbf{F}\alpha(\mathbf{F}\beta\gamma))(\mathbf{F}(\mathbf{F}\alpha\beta)(\mathbf{F}\alpha\gamma))$
W	2	$x_1x_2x_2$	**SS(KI)**	$\mathbf{F}(\mathbf{F}\alpha(\mathbf{F}\alpha\beta))(\mathbf{F}\alpha\beta)$
ϕ	4	$x_1(x_2x_4)(x_3x_4)$	**B(BS)B**	
ψ	4	$x_1(x_2x_3)(x_2x_4)$	———	

with only one letter. These words we shall call the *natural numbers.* Further, let σ be the successor function, δ the predecessor function, τ the ordered-pair function, and μ the operation such that for any numerical function *f*, μf is the least *n* for which $f(n) = 0$ and is undefined if there is no such *n*.

One can find a representation of the natural numbers as combinators; indeed, there are many choices. For any such choice let angle brackets "⟨ ⟩" symbolize the combinatory analogues of the arithmetic notions indicated within them. Thus, $\langle n\rangle$ is, for any numeral *n*, the *combinatory numeral*which represents it, $\langle +\rangle$ the analogue of addition, etc. The analogues are often not uniquely determined.

The first representation, by Church in 1933, chose $\langle n\rangle$ so that $\langle n\rangle f$ is the *n*th iterate of *f* (the first iterate being *f* itself). If one has **K**, then $\langle 0\rangle$ is **KI** and $\langle\sigma\rangle$ is **SB**. Then $\langle n\rangle$ is the $\mathbf{Z}_n$ of H. B. Curry and Robert Feys (*Combinatory Logic*). Further, $\langle +\rangle$, $\langle\cdot\rangle$, and $\langle e\rangle$>, where $e(x,y) = x^y$, have simple definitions (for example, ϕ**B**, **B**, and **CI**, respectively) from which their arithmetical properties follow. There are other proposals for combinatory numerals; one, made by Dana Scott in 1963, has a simple $\langle\delta\rangle$. For the sake of generality, $\langle n\rangle$ is here unspecified, but a **Z** is postulated such that $\mathbf{Z}\langle n\rangle = \mathbf{Z}_n$. If $\langle n\rangle \equiv \mathbf{Z}_n$, then $\mathbf{Z} \equiv \mathbf{I}$.

Next one can define combinators $\mathbf{D}$ ($\equiv \langle\tau\rangle$), $\mathbf{D}_1$, $\mathbf{D}_2$ such that

$$\mathbf{D}_1(\mathbf{D}xy) = x,\ \mathbf{D}_2(\mathbf{D}xy) = y.$$

For instance (as Paul Bernays suggested in 1936),

$$\mathbf{D} = [x,y,z].\mathbf{Z}z(\mathbf{K}y)x.$$

For this

(7) $\mathbf{D}xy\langle 0\rangle = x,\ \mathbf{D}xy\langle\sigma n\rangle = y$,

so that $\mathbf{D}_1$ and $\mathbf{D}_2$ can be $[x]x\langle 0\rangle$ and $[x]x\langle 1\langle$, respectively. One can also define **D** in terms of $\langle\delta\rangle$ rather than **Z**.

Next a combinator **R** can be defined such that

(**8**) $$\mathbf{R}xy\langle 0\rangle = x,\ \mathbf{R}xy\langle\sigma\rangle = y\langle n\rangle(\mathbf{R}xy\langle n\rangle).$$

If $x = \langle g\rangle$ and $y = \langle h\rangle$, where g and h are, respectively, k-place and $(k + 2)$-place numerical functions, $\mathbf{R}xy$ can be taken as $\langle f\rangle$, where f is the $(k + 1)$-place numerical function defined by the "primitive recursion scheme" from g and h. Since the other processes of forming primitive recursive functions have combinatory analogues, definition of **R** will ensure that $\langle f\rangle$ is defined for any primitive recursive f.

Several definitions of **R** exist. The first (given by Bernays in 1936) depends on the fact that $f(n)$ can be obtained (for $k = 0$, as an example) by iterating n times, starting with $\tau(0,g)$, the function ϕ such that $\phi(\tau(x,y)) = t(\sigma x,h(x,y))$ and taking the second member. This leads to the definition (in two stages)

$$\boldsymbol{Y} \equiv [u]\mathbf{D}(\langle\sigma\rangle(\mathbf{D}_2u))\ (y(\mathbf{D}_1u)\ (\mathbf{D}_2u)),$$

$$\mathbf{R} \equiv [x,y,z](\mathbf{D}_2(\mathbf{Z}z\boldsymbol{Y}(\mathbf{D}\langle 0\rangle x))).$$

Another possibility is to define a combinator $\boldsymbol{\Omega}$ such that for given obs p, q, r, the ob $t = \Omega pqr$ satisfies the conditions

(**9**) $$t\langle n\rangle = \begin{cases} p\langle n\rangle & \text{if } r\langle n\rangle = \langle 0\rangle, \\ qf\langle n\rangle & \text{if } r\langle n\rangle = \langle 1\rangle, \end{cases}$$

thus:

$$\boldsymbol{Y} \equiv \mathbf{D}(\mathbf{K}p)([u](q([z](u(rz)uz)))),$$

$$\boldsymbol{\Omega} \equiv [p,q,r,x](\boldsymbol{Y}(rx)\boldsymbol{Y}x).$$

For $p \equiv \mathbf{K}x$, $q \equiv [u,z](y(\langle\delta\rangle z)(u(\langle\delta\rangle z)))$, $r = \mathbf{I}$, the ob $[x,y]\Omega pqr$ is an **R**, different from the foregoing, satisfying (8). There are still other ways of defining **R**. Since $\langle\delta\rangle$ can be defined as $\mathbf{R}\langle 0\rangle\mathbf{K}$ and **Z** as **R** (**KI**)(**K**(**SB**)), we have any primitive recursive function as soon as we have either **Z** or $\langle\delta\rangle$ and a discrimination for $\langle 0\rangle$.

We can go further. If we take $p \equiv \mathbf{I}$, $q \equiv [u,z]\ (u(\langle\sigma\rangle z))$, while r is a given function $\langle g\rangle$, then Ωpqr is an $\langle f\rangle$ such that $f(n) = n$ if $g(n) = 0$ and otherwise $f(n) = f(\sigma n)$. This shows that we can define $\langle\mu\rangle$ in terms of the above q as $[x](\Omega\mathrm{I}qx\langle 0\rangle)$. Consequently, every partial recursive numerical function is definable by combinators. The converse of this thesis follows by the usual arguments involving Gödel numeration.

These conclusions are not greatly affected if one restricts the system to correspond with restricted forms of λ-conversion. The passage from η-conversion to β-conversion hardly makes any difference. The omission of **K** complicates things somewhat—one needs ordered triples instead of ordered pairs. But the main conclusion, that every partial recursive function is definable by combinators and vice versa, stands.

Some generalizations are known. One can define by combinators certain transformations between obs and their Gödel numbers. An extension to recursive functionals of certain types can be obtained by using an analogue of (δ). There is also an extension to certain transfinite ordinal numbers.

ILLATIVE THEORY. By definition illative combinatory logic includes all considerations where there are atoms which neither are combinators nor are necessary to express equality. We consider here those cases in which the new atoms mean ordinary logical notions—for example, II (absolute universality), **P** (implication), Ξ (relative universality or formal implication), **F** (functionality—**F**XYZ means that Z is a function from X into Y), Σ (instantiality), Λ (conjunction), Γ (negation), Θ (descriptive quantifier), etc. In addition, we need obs meaning semantical categories, such as **E** (the category of all obs—**E** is definable, for example, as **WQ**), **H** (propositions), **J** (individuals), **M** (sets), **N** (numbers), etc.

The meaning of these obs is expressed more precisely by the rules associated with them. For the first four obs these are

RULE Π: $\vdash \Pi X\ \&\ \vdash \mathrm{E}U \to\ \vdash XU.$

RULE **P**: $\vdash \mathbf{P}XY\ \&\ \vdash X \to\ \vdash Y.$

RULE **Ξ**: $\vdash \boldsymbol{\Xi}XY\ \&\ \vdash XU \to\ \vdash YU.$

RULE **F**: $\vdash \mathbf{F}XYZ\ \&\ \vdash XU \to\ \vdash Y(ZU).$

These rules, when relevant, are to be postulated in addition to the combinatory rules given earlier; the latter can be summarized as

RULE Eq: $\vdash X\ \&\ X = Y \to\ \vdash Y.$

These notions are, of course, interdefinable; in fact, one can take as atoms either **F**, Ξ, or Π and **P** and define the others as follows (there are two possible definitions for Ξ in terms of **F**):

$$\mathbf{F} \equiv [x,y,z](\Xi x(\mathbf{B}yz)),$$

$$\Xi \equiv [x,y](\Pi([z](\mathbf{P}(xz)(yz)))),$$

$$\Pi \equiv \Xi\mathbf{E},$$

$$\mathbf{P} \equiv \Psi\Xi\mathbf{K} = [x,y](\Xi(\mathbf{K}x)(\mathbf{K}y)),$$

$$\Xi'' \equiv [x,y](\mathbf{F}xy\mathbf{I}),\ \Xi'' \equiv [x,y](\mathbf{F}x\mathbf{I}y).$$

The system based on **F** as primitive is called $\mathscr{F}_1$, or the theory of *functionality*; that on Ξ as $\mathscr{F}_2$, or the theory of *restricted generality*; and that on Π and **P** as $\mathscr{F}_3$. With reasonable axioms, these are listed in order of increasing strength.

Although the Church–Rosser theorem shows that pure combinatory logic is consistent, in the illative theory one easily runs into contradictions. Thus, if one were to assume

(10) $\qquad \vdash \mathbf{P}(\mathbf{P}\alpha(\mathbf{P}\alpha\beta))(\mathbf{P}\alpha\beta)$

as an axiom scheme, with the Greek letters standing for arbitrary obs, the theory would be inconsistent in the sense that (1) would hold for any *X*. But (10) is a thesis of the *absolute* (that is, positive intuitionistic) *propositional algebra.* Thus, it is necessary, if the theory is to contain this algebra, that (10) be a theorem scheme with a restricted range for the Greek letters. In its early stages illative combinatory logic will have axiom schemes with such restrictions. Later, perhaps, these schemes will be reduced to axioms by quantifying over a suitable category.

This requires some sort of machinery of categories or types. Such machinery is taken for granted in the usual systems of mathematical logic. It consists of four items: (*a*) a list of primitive categories (such as those listed above), (*b*) devices for forming derived categories, (*c*) assignments of the primitive notions to categories, (*d*) means for determining the categories of composite notions. Of these items (*a*) and (*c*) are special to the theory considered, but (*b*) and (*d*) are general processes which are appropriate for study in combinatory logic. Since composite obs are formed by application alone, one would expect a means of assigning a category to *XY* when those for *X* and *Y* are known; the general principle is that if *X* is a function from α to β and *Y* belongs to α, then *XY* belongs to β. This principle is expressed by Rule **F**, so the basis for this generalized theory of types is $\mathscr{F}_1$.

From the illative standpoint one would expect that each combinator would be assigned a category depending on parameters expressing that it is a function transforming from certain sorts of categories to categories of certain other sorts. Such *functional characters* for some basic combinators of the table are listed there in the fifth column. Assignments of these characters to the atomic combinators would then be axiom (or at least theorem) schemes of $\mathscr{F}_1$. However, these schemes cannot be accepted with the Greek letters standing for arbitrary obs, for if one so accepts **FW**, the theory is again inconsistent. Even $\mathscr{F}_1$ has to be formulated with restrictions on the Greek letters.

The most radical restriction is the requirement that the Greek letters range over an inductive class of F-obs generated from certain otherwise unspecified atoms θ_1, $\theta_2, \cdots$ by the operation of forming $\mathbf{F}\alpha\beta$ from α and β. One further restricts Rule Eq thus:

RULE **Eq'**: $\qquad \vdash \xi X \ \&\ X = Y \rightarrow \vdash \xi Y.$

The resulting theory is called the *basic theory of functionality.* In this theory every elementary statement will be of the form

(11) $\qquad \vdash \xi X,$

where ξ is an F-ob and *X* is a combinator. The theory is demonstrably consistent. If *X* is a *stratified* combinator—that is, if *X* satisfies (3) and one can derive $\vdash \eta \boldsymbol{M}$ by Rule **F** alone from the axiom schemes and assignments of categories to the variables—then one can derive a statement of the form (11) stating that *X* has the appropriate functional character. There is a converse to this which is somewhat difficult to state, but it shows that the *X*'s for which (11) can be derived are greatly restricted; in particular, they have a normal form.

There are several "stronger" theories of functionality with less drastic restrictions. A theory in which one uses only combinators that do not repeat variables can be constructively proved consistent without restrictions on the Greek letters. Constructive consistency proofs have been obtained for some other theories of $\mathscr{F}_1$.

All these systems of $\mathscr{F}_1$ are extremely weak. To obtain stronger and natural theories one proceeds to $\mathscr{F}_2$ (or adds assumptions to $\mathscr{F}_1$ which are equivalent to this). In $\mathscr{F}_2$ reasonable schemes **ΞI**, **ΞK**, etc., with Greek letters restricted to a class of "canonical obs," can be formulated from which the corresponding **FI**, **FK**, $\cdots$ can be derived. Thus, $\mathscr{F}_2$ has a deduction theorem; it also includes the absolute propositional calculus of pure implication. There is a Gentzen-like theory of "verifiability" from which it follows that certain weak forms of the illative theory are consistent.

The study of illative combinatory logic is still in its preliminary stages. Little is known, for example, about $\mathscr{F}_3$. It is clear that ordinary logical systems can be founded on a combinatory basis, but little has been pub-

lished along this line. On such a basis E. J. Cogan, in 1955, analyzed the foundations of Gödel's set theory and also the predicate calculus and some other calculuses; owing to an unfortunate oversight in the definition of "class," the system was inconsistent, as Rainer Titgemeyer showed in 1961, but the necessary changes are rather minor. Other investigations of this sort are in the process of development or publication. Some authors, such as F. B. Fitch, go in a somewhat different direction.

In illative combinatory logic we are dealing with concepts of such generality that we have little intuition in regard to them. This explains why proposals by competent logicians beginning with Gottlob Frege (not all combinatory, but the principle applies) have later proved inconsistent. We must, indeed, proceed by trial and error. No doubt we shall continue to find both inconsistencies in weaker systems and consistency proofs of stronger systems. In due course nonfinitary methods will be used, and much is to be expected of them. But the possibility remains that we may always be interested in systems for which neither consistency nor inconsistency is known.

Combinatory logic was inaugurated by Moses Schönfinkel in 1924. He introduced the notion of application, the combinators **B**, **C**, **I**, **K**, **S** (his *Z, T, I, C, S*), and an illative notion *U.* He showed how statements of logic could be expressed in terms of these notions, but he gave no deductive theory of them. He became ill shortly after writing the paper and was unable to do anything further in the subject. Curry, beginning in 1929, produced the first deductive synthetic theory and introduced the terminology used here. The theory of λ-conversion was developed by Church from 1932. Subsequent improvements were made by these authors and by Rosser, S. C. Kleene, Bernays, Fitch, and Paul Rosenbloom. The present state of the subject is the result of an interaction of the work of these authors and their students.

See also Church, Alonzo; Frege, Gottlob; Logical Paradoxes; Logic, History of; Quantifiers in Formal Logic.

Bibliography

Items listed here either are of broad scope or treat aspects of the subject not included in the text. For sources, details, etc., see the bibliographies in the works cited.

Church, Alonzo. *The Calculi of Lambda-Conversion.* Princeton, NJ: Princeton University Press, 1941; 2nd ed., 1951. The standard work on λ-conversion, with references.

Church, Alonzo. "The Richard Paradox." *American Mathematical Monthly* 41 (1934): 356–361. Expounds a system modeling illative notions on the pure theory.

Cogan, E. J. "A Formalization of the Theory of Sets from the Point of View of Combinatory Logic." *Zeitschrift für mathematische Logik und Grundlagen der Mathematatik* 1 (1955): 198–240. Contains a readable introduction and some information on illative theory going beyond Curry and Feys (below). For an inconsistency, see text, section on illative theory.

Curry, H. B. "The Elimination of Variables by Regular Combinators." In *The Critical Approach to Science and Philosophy,* in honor of Karl R. Popper, edited by Mario Bunge, pp. 127–143. New York: Free Press of Glencoe, 1964. Contains, on pp. 128–131, a discussion of the "ontology" of combinatory logic, partly in answer to Quine (below).

Curry, H. B., and Robert Feys. *Combinatory Logic,* Vol. I. Amsterdam: North-Holland, 1958. A monograph, with full bibliography, historical sketch, etc.

Fitch, F. B. "The System CΔ of Combinatory Logic." *Journal of Symbolic Logic* 28 (1963): 87–97. A nonfinitary system with consistency proof. Refers to Fitch's other papers.

Quine, W. V. "Variables Explained Away." *Proceedings of the American Philosophical Society* 104 (1960): 343–347. An alternative formulation of combinatory logic, to be used in combination with first-order predicate calculus, which, Quine claims, involves fewer ontological presuppositions.

Rosenbloom, Paul. *The Elements of Mathematical Logic.* New York: Dover, 1950. Very compact. Departures from standard notations hinder reading.

Rosser, J. B. *Deux Esquisses de logique.* Paris: Gauthier-Villars, 1955. A masterly exposition of certain parts of the pure theory.

Schönfinkel, Moses. "Über die Bausteine der mathematischen Logik." *Mathematische Annalen* 92 (1924): 305–316. Still readable and interesting for motivation. The statement in the postscript (by Heinrich Behmann) to the effect that parentheses can be eliminated by **Z** (that is, **B**) alone is erroneous.

Haskell B. Curry (1967)

COMEDY

See *Humor*

COMENIUS, JOHN AMOS

(1592–1670)

John Amos Comenius, also called Komensky, the Czech philosopher of education and theologian, was born in Uhersky Brod. Comenius was a member of the Community of the Moravian Brethren (*Unitas Fratrum*) and studied Protestant theology at the universities of Herborn and Heidelberg. Shortly after his return to Moravia, the Thirty Years' War broke out. The Protestant Czechs were defeated by the Catholic Hapsburg monarchy, and Comenius became a permanent exile. Elected bishop of the *Unitas* in 1632, he considered it his main mission as a pas-

tor and as a theological writer to preserve the faith and unity of the dispersed Moravian brethren.

In his writings, which range from such topics as theology, politics, philosophy, and science (as he understood science) to linguistics and education, as well as in his personal life, he combined such contradictory strands of thought as world immanence and world transcendence, interest in science and dependence on false prophets, progressivism, and apocalyptic expectations. In order to understand this mingling of ideas, we must project ourselves into the baroque age, when so many illustrious minds were wandering from one extreme to another. Thus, despite scholastic and Calvinist influences during his years of study, Comenius's concept of the divine regime contained a notable admixture of Neoplatonic, evolutionary, mystical, and pantheistic ideas. God was for him the God of Nature as well as the God of Heaven. However, all these pantheistic leanings did not shake the foundations of Comenius's faith, and throughout his life he clung to the fundamentals of the Christian dogma. Nevertheless, it was the cosmic curiosity in Comenius's religion that opened his mind to the unfolding of the natural and humanistic sciences. Yet Comenius lacked any real understanding of science in the Newtonian sense. The generic concept under which he subsumed the new scientific pursuits was that of "Light," to be understood as both the "Light of God" and the light of reason that God has kindled in man in order to guide him on his way toward eternal truth.

No doubt a certain utopian chiliasm inspired Comenius, but he also shared with the greatest minds of his time the enthusiasm about a new discovery, the discovery of "method," understood as a form of systematic and empirical inquiry which would guarantee the harmonization between man's reason and the natural—and perhaps even the supernatural—universe. The man who impressed Comenius most of all was Francis Bacon. Through Bacon, he became convinced that the new inductive method would shed light not only on the *arcana naturae* but also on the mysteries of the human mind and of human learning. The long title of Comenius's *Great Didactic* (*Didactica Magna*) tells the reader that the author believes he has found a system to teach "all things to all men." Comenius was one of the first to grasp the significance of a methodical procedure in schooling, to project a plan of universal education, and to see the significance of education as an agency of international understanding. Often quoted are the eight principles of teaching that Comenius expounds in Chapter 9 of the *Great Didactic,* in strange analogy to what he supposes to be the economy and order of the sun's functioning in the universe. Still valid in these principles is the emphasis on the interrelation between mental maturity and learning, on the participation of the student, and on the logical interconnection of the subjects in the curriculum.

Education—to be extended to both sexes, all men, and all peoples—should be crowned by a *pansophia* (encyclopedic synthesis of universal knowledge), with the aim of a *dilucidatio* (systematic interpretation) of the order of all things within the cosmic order. For the promotion of the great and worldwide mission of education, Comenius recommended a "Universal College" of the great and wise men of the whole world, and an easily constructed international language for the peace and "for the reform of the whole world" and as an "antidote to the confusion of thought."

In 1668 he dedicated a treatise, *The Way of Light* (*Via Lucis*), "to the torch bearers of this enlightened age, members of the Royal Society of London, now bringing real philosophy to a happy birth." He expressed the "confident hope" that through their endeavors "philosophy brought to perfection" would "exhibit the true and distinctive qualities of things … for the constantly progressive increase of all that makes for good to mind, body, and estate."

Bibliography

WORKS BY COMENIUS

The Great Didactic. Translated by M. W. Keatinge. London: A. and C. Black, 1923.

The Way of Light. Translated by E. T. Campagnac. Liverpool: Liverpool University Press, 1938.

The Labyrinth of the World and the Paradise of the Heart. Translated by Matthew Spinka. Chicago: National Union of Czechoslovak Protestants in America, 1942.

John Amos Comenius, Selections, edited by Jean Piaget. New York, 1957. Contains references to more-recent Czech literature.

WORKS ON COMENIUS

Bušek, Vratislav, ed. *Comenius.* Translated by Káca Polácková. New York: Czechoslovak Society of Arts and Sciences in America, 1972.

Kozik, František. *The Sorrowful and Heroic Life of John Amos Comenius.* Prague: State Educational Publishing House, 1958.

Kvacala, J. *J. A. Comenius.* Berlin, 1914. Contains a large bibliography.

Laurie, S. S. *John Amos Comenius.* Boston, 1885.

Petru, Eduard. "The Harmonizing Influence of God in the Understanding of J. A. Comenius." *Ultimate Reality and Meaning* 20 (2 and 3) (1997): 99–106.

Sadler, John E. *J. A. Comenius and the Concept of Universal Education.* London: Allen and Unwin, 1966.
Ulich, Robert. *History of Educational Thought.* New York: American Book, 1950.
Ulich, Robert. *Three Thousand Years of Educational Wisdom.* 2nd ed. Cambridge, MA: Harvard University Press, 1954.

Robert Ulich (1967)
Bibliography updated by Tamra Frei (2005)

COMMON CAUSE PRINCIPLE

No correlation without causation. This is the most compact formulation of Reichenbach's Common Cause Principle (RCCP). More explicitly RCCP is the claim that if two events A and B are correlated, then either A and B stand in a causal relation, $R_{cause}(A, B)$, or, if A and B are causally independent, $R_{ind}(A, B)$, then there is a third event C, a so-called Reichenbachian common cause that brings about the correlation by being related to A and B in a specific manner spelled out in the following definition, first given by Reichenbach (1956): Event C is called a (Reichenbachian) common cause of the correlation

$$p(A \wedge B) - p(A)p(B) > 0 \tag{1}$$

if the following conditions hold:

$$p(A \wedge B|C) = p(A|C)p(B|C) \tag{2}$$

$$p(A \wedge B|C^{\perp}) = p(A|C^{\perp})p(B|C^{\perp}) \tag{3}$$

$$p(A|C) > p(A|C^{\perp}) \tag{4}$$

$$p(B|C) > p(B|C^{\perp}) \tag{5}$$

Here A, B, and C are assumed to be elements in a Boolean algebra $\mathcal{S}$ and they are to be interpreted as representatives of random events. $p(A|C) = p(A \wedge C)/p(C)$ and so on denote the conditional probability of A on condition C, $C^{\perp}$ denotes the complement of C, and it is assumed that none of the probabilities involved is equal to zero.

RCCP is a metaphysical claim about the causal structure of the world, and it has been debated extensively in the philosophical literature whether RCCP is a valid principle. How could RCCP fail? The first step in any attempt to falsify RCCP is to display common cause incomplete probability spaces, that is, probability spaces that contain at least one correlation that does not have a common cause in the given probability space. Common cause incomplete probability spaces exist; however, the mere existence of such probability spaces does not entail that RCCP is not valid because RCCP is not the claim that given a correlated pair (A, B) of events in $\mathcal{S}$ there has to exist a common cause C that belongs to $\mathcal{S}$: RCCP is a pure existence claim, not requiring the common cause to belong to the specific set of events $\mathcal{S}$. If, however, one wishes to maintain the validity of RCCP against the threat posed by the existence of common cause incomplete probability spaces, one has to be able to claim that the probability space $(\mathcal{S}, p)$ is consistently extendable into a larger probability space $(\mathcal{S}', p')$ that does contain a common cause of the given correlation. If this can be done, one calls $(\mathcal{S}, p)$ "common cause completable" (with respect to the given correlation). It can be shown that every common cause incomplete probability space is common cause completable with respect to any finite set of correlations in it. (It is an open problem whether common cause extendability with respect to an infinite number of correlated events also holds.)

In view of common cause completability of probability spaces, one can always defend RCCP against attempts of falsification by referring to "hidden" common causes—"hidden" in the sense of not being accounted for in the set of events $\mathcal{S}$. Thus any successful falsification of RCCP must require some properties of the common cause in addition to those required by Reichenbach. One such possible requirement is that different correlations have a *common* common cause. One can show that different correlations cannot in general have a common common cause—not even in case of two correlations.

Assuming that RCCP is valid, one is led to the question of whether our theories predicting probabilistic correlations can be causally rich enough to contain also the causes of the correlations. According to RCCP, causal richness of a theory $(\mathcal{S}, p)$ would manifest in the theory's being causally closed: $(\mathcal{S}, p)$ is called *common cause closed with respect to* R_{ind}, if for every pair (A, B) of correlated events such that $R_{ind}(A, B)$ holds, there exists a common cause C in $\mathcal{S}$ of the correlation.

Whether a probabilistic theory is common cause closed with respect to the causal independence relation R_{ind} depends on how R_{ind} is specified. The weaker R_{ind} (i.e., the more pairs of random events are causally independent) the stronger the notion of common cause closedness with respect to R_{ind} and the more difficult it is for $\mathcal{S}$ to be common cause closed with respect to R_{ind}. For instance no probability space with a finite set of random events can be common cause closed with respect to the weakest R_{ind} (i.e., if $R_{ind}(A, B)$ holds for all A and B). However if $R_{ind}(A, B)$ is strong enough to imply that the presence of A implies neither the presence of B nor the presence of $B^{\perp}$

(and conversely, replacing *A* with *B*) then finite probability spaces can be common cause closed (with respect to R_{ind})—though they are not necessarily so. EPR correlations predicted by quantum mechanics are generally viewed as ones that might not admit a common cause type explanation—if the common causes are required also to conform to relativistic causality (such common causes are called "local").

Proving the impossibility of local common causes of EPR correlations involves two difficulties: First one has to link RCCP to quantum mechanics, which is non-trivial task since Reichenbach's notion of common cause was defined in terms of classical probability theory, not in terms of quantum mechanics. Second one has to formulate "locality" of common causes. One can approach the first problem in two ways: (i) reformulating Reichenbach's notion of common cause in terms of non-classical (quantum) probability spaces; (ii) representing quantum probabilities and quantum correlations in terms of classical probability theory.

Reichenbach's notion of common cause can be reformulated in terms of non-classical probability theory, where $\mathcal{S}$ is replaced by the lattice of projections of a von Neumann algebra and *p* by a state on the von Neumann algebra. The notions of common cause and of common cause completability can be adapted to the non-commutative case, and it can be shown that every non-commutative probability space also is common cause completable. Relativistic causality can also be formulated in terms of non-commutative probability spaces—the resulting theory is known as local algebraic quantum field theory. Locality of common causes of EPR correlations predicted by local quantum field theory can be defined by requiring the common causes to belong to a spacetime region located within the intersection of backward light cones of the spacelike separated regions containing the correlated observables. Whether such localized common causes exist in quantum field theory is an open problem, only partial results are known.

One can also take approach (ii) and formulate locality conditions for the hypothetical common causes of EPR correlations predicted by non-local, non-relativistic quantum mechanics—now represented in classical probability theory. These locality conditions express two sorts of independence: (i) the statistical independence of the random events of choosing measurements in the two wings of a typical correlation experiment and (ii) the statistical independence between choosing measurements in any wing and presence of any combination of the hypothetical common causes of spin correlations in different directions. Again it is an open question whether common causes satisfying these locality conditions can exist. It is known however that the EPR correlations between outcomes of spin measurements in different directions cannot have a common common cause because the assumption of common common causes of EPR correlations in different directions implies Bell's inequality.

See also Bell, John, and Bell's Theorem; Causation: Philosophy of Science; Neumann, John von; Reichenbach, Hans.

Bibliography

Arntzenius, Frank. "The Common Cause Principle." *PSA* 2 (1992): 227–237.

Cartwright, Nancy. "How to Tell a Common Cause: Generalization of the Conjunctive Fork Criterion." In *Probability and Causality*, edited by James H. Fetzer, 181–188. Boston: Reidel, 1988.

Fetzer, James H. Fetzer *Probability and Causality*. Boston: Reidel, 1988.

Hofer-Szabó, Gabor, Miklos Rédei, and Laszlo Szabó. "Common-Causes Are Not Common Common-Causes." *Philosophy of Science* 69 (2002): 623–636.

Hofer-Szabó, Gabor, Miklos Rédei, and Laszlo Szabó. "On Reichenbach's Common Cause Principle and Reichenbach's Notion of Common Cause." *The British Journal for the Philosophy of Science* 50 (1999): 377–398.

Hofer-Szabó, Gabor, Miklos Rédei, and Laszlo Szabó. "Reichenbach's Common Cause Principle: Recent Results and Open Questions." *Reports on Philosophy* 20 (2000): 85–107.

Placek, Tomasz. *Is Nature Deterministic?* Cracow: Jagellonian University Press, 2000.

Rédei, Miklos, and Stephen Summers. "Local Primitive Causality and the Common Cause Principle in Quantum Field Theory." *Foundations of Physics* 32 (2002): 335–355.

Rédei, Miklos. "Reichenbach's Common Cause Principle and Quantum Correlations." In *Modality, Probability and Bell's Theorems*, NATO Science Series, II. Vol. 64, edited by Tomasz Placek and Jeremy Butterfield, 259–270. Boston: Kluwer Academic, 2002.

Rédei, Miklos. "Reichenbach's Common Cause Principle and Quantum Field Theory." *Foundations of Physics*, 27 (1997): 1309–1321.

Reichenbach, Hans. *The Direction of Time*. Los Angeles: University of California Press, 1956.

Salmon, Wesley. "Why Ask 'Why?'?" In *Proceedings and Addresses of the American Philosophical Association* 51 (1978): 683–705.

Salmon, Wesley. "Probabilistic Causality." *Pacific Philosophical Quarterly* 61 (1980): 50–74.

Salmon, Wesley. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press, 1984.

Sober, Elliott. "Common Cause Explanation." *Philosophy of Science* 51 (1984): 212–241.

Sober, Elliott. "The Principle of the Common Cause." In *Probability and Causality*, edited by James H. Fetzer, 211–228. Boston: Reidel, 1988.
Sober, Elliott. "Venetian Sea Levels, British Bread Prices, and the Principle of Common Cause." *The British Journal for the Philosophy of Science* 52 (2001): 331–346.
Spohn, Wolfgang. "On Reichenbach's Principle of the Common Cause." In *Logic, Language and the Structure of Scientific Theories*, edited by Wesley Salmon and Gereon Wolters, 211–235. Pittsburgh: University of Pittsburgh Press, 1991.
Suppes, Patrick. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland, 1970.
Suppes, Patrick, and Mario Zanotti. "When Are Probabilistic Explanations Possible?" *Synthese* 48 (1981): 191–199.
Uffink, Jos. "The Principle of the Common Cause Faces the Bernstein Paradox." *Philosophy of Science*, Supplement, 66 (1999): 512–525.
Van Fraassen, Bas. "The Charybdis of Realism: Epistemological Implications of Bell's Inequality." In *Philosophical Consequences of Quantum Theory*, edited by Jamers Cushing and Ernan McMullin, 97–113. Notre Dame: University of Notre Dame Press, 1989.
Van Fraassen, Bas. "The Pragmatics of Explanation." *American Philosophical Quarterly* 14 (1977): 143–150.
Van Frassen, Bas. "Rational Belief and the Common Cause Principle." In *What? Where? When? Why?*, edited by Robert McLaughlin, 193–209. Boston: D. Reidel, 1982.

Miklós Rédei (2005)

COMMON CONSENT ARGUMENTS FOR THE EXISTENCE OF GOD

Numerous philosophers and theologians have appealed to the "common consent" of humankind (the *consensus gentium*) as support for certain doctrines. Richard Hooker, for example, in his *Treatise on the Laws of Ecclesiastical Polity* appeals to this common agreement of humankind in justifying his view that the obligatory character of certain moral principles is immediately evident. Most frequently the conclusions supported in this way were those asserting the existence of God and the immortality of the human soul. In the present entry we shall confine ourselves to common consent arguments for the existence of God.

Among those who favored arguments of this kind were Cicero, Seneca, Clement of Alexandria, Herbert of Cherbury, the Cambridge Platonists, Pierre Gassendi, and Hugo Grotius. In more recent times these arguments were supported by numerous distinguished Protestant and Catholic theologians. G. W. F. Hegel did not accept the argument, but he thought that it contained a kernel of truth. Rudolf Eisler, in his *Wörterbuch der philosophischen Begriffe,* ranks the argument fifth in importance among so-called proofs of the existence of God, and this seems an accurate estimate of its place in the history of philosophy. At the same time, J. S. Mill was probably right when he observed that, as far as the "bulk of mankind" is concerned, the argument has exercised greater influence than others that are logically less vulnerable. Although there are hardly any professional philosophers at the present time who attribute any logical force to reasoning of this kind, it is still widely employed by popular apologists for religion.

Some supporters claim relatively little. "In no form," wrote the nineteenth-century theologian Robert Flint, "ought the argument from general consent to be regarded as a primary argument. It is evidence that there are direct evidences—and when kept in its proper place it has no inconsiderable value—but it cannot be urged as a direct and independent argument" (*Theism,* p. 349). Cardinal Mercier similarly regarded the argument as "indirect or extrinsic." It does not by itself prove the existence of God, but it is a "morally certain indication that there are proofs warranting the assertion that God exists" (*A Manual of Modern Scholastic Philosophy,* Vol. II, p. 55). Father Bernard Boedder and G. H. Joyce claim a great deal more. Boedder (*Natural Theology,* p. 63) regards it as an "argument of absolute value in itself." The universal consent "of nations in the recognition of God must be deemed the voice of universal reason yielding to the compelling evidence of truth." Later, however, he admits that it is not "absolutely conclusive, except when taken in conjunction with the argument of the First Cause" (ibid., p. 75). Joyce, a twentieth-century writer to whom we owe one of the fullest and clearest statements of one version of the argument, is far more sanguine. He calls it without any qualification a "valid proof of the existence of God" and seems to regard the conclusion as established with "perfect certainty."

The argument has rarely been stated by any philosopher in the form of a simple appeal to the universality of belief in God. In this form it is patently invalid and invites Pierre Bayle's comment that "neither general tradition nor the unanimous consent of all men can place any injunction upon truth." There is, on the face of it, no reason why the whole of humankind should not have been as wrong on a speculative topic as it has been on some more empirical questions on which, history teaches, it has been mistaken. The actual versions of the argument advanced by philosophers are more complicated and can be conveniently grouped into two classes. In the first we

have arguments in which the universality of belief, for reasons peculiar to this particular case, is taken as evidence either that the belief itself is instinctive or that it is due to longings or needs which are instinctive. It is then concluded, for a variety of reasons, that the belief must be true. In the second group we have arguments according to which the universality of the belief, in conjunction with the claim that believers used reason in arriving at their position, is treated as evidence for the existence of God. We shall refer to arguments of the first kind as "biological" versions and to those of the second kind as the "antiskeptical dilemma." Whatever the shortcomings of these arguments may be, they cannot be dismissed simply on the ground that the whole of humankind may well be mistaken.

Although no doubt some of the disputes in which philosophers and others have engaged in this connection are antiquated and sometimes have a slightly preposterous ring to modern ears, other related issues are still very much with us. For example, it is still maintained by a number of influential philosophers and psychologists that people are "by nature" religious, so that the spread of skepticism and atheism is likely to lead to highly undesirable results. "It is safe to say," writes Carl Jung about his patients, "that every one of them fell ill because he had lost that which the living religions of every age have given their followers" (*Modern Man in Search of a Soul,* p. 254). Nor are attempts lacking even in our own day to show that everybody "really" believes in God, no matter what they may say or think. In the course of evaluating various forms of the Argument from Common Consent, we shall have occasion to say something about these more contemporary issues as well.

BIOLOGICAL FORMS OF THE ARGUMENT

INSTINCTIVE BELIEF IN GOD. A familiar version of the biological form of the Argument from Common Consent is found in Seneca's *Epistulae Morales* (Letter 117):

> We are accustomed to attach great importance to the universal belief of mankind. It is accepted by us as a convincing argument. That there are gods we infer from the sentiment engrafted in the human mind; nor has any nation ever been found, so far beyond the pale of law and civilization as to deny their existence.

Seneca did not elaborate on the nature of the "sentiment" that is "engrafted in the human mind," but later writers did, especially when replying to criticisms such as John Locke's. In the course of his polemic against the theory of innate ideas, Locke had rejected the initial premise of the argument as plainly false. His reasons were twofold. First, he noted with regret that there were atheists among the ancients, and also, more recently, "navigation discovered whole nations among whom there was to be found no notion of a God" (*Essay concerning Human Understanding,* Book I, Sec. IV). Aside from questioning the prevalence or even the existence of unbelief, the usual reply to this kind of criticism has been to make a distinction between two senses in which an idea or a belief may be said to be innate or instinctive. Such an assertion may mean that the idea or the belief is present in the human mind at birth as an image or some other actual "content," or it may amount to the much milder claim that it is present as a disposition to arrive at the belief when noticing certain things in the world or in oneself (usually this is stated very strongly to the effect that, when noticing the things in question, the person cannot help coming to believe in God). It is then explained that belief in God is instinctive in the latter or dispositional sense only. To avoid the charge of triviality, the defenders of the argument usually insist that because of this disposition, teaching or indoctrination is not required. Thus Charles Hodge, who makes it clear that he advocates a doctrine of the innateness of belief in God in the dispositional sense, adds that "men no more need to be taught that there is a God, than they need to be taught that there is such a thing as sin" (*Systematic Theology,* Vol. I, p. 199). "Adam," he also writes, "believed in God the moment he was created, for the same reason that he believed in the external world. His religious nature, unclouded and undefiled, apprehended the one with the same confidence that his senses apprehended the other" (ibid., pp. 200–201).

Several comments are in order here. To begin with, the theory that belief in God is innate does not become vacuous when it is stated in this way, so long as we are told what the facts are in the presence of which a human being cannot help coming to believe in the existence of God. However, when these facts are specified as the adjustments of organisms to their environment or as our experiences of duty and obligation (and these are the ones most frequently mentioned), Locke's objection seems to be fundamentally intact. For, apart from the question of primitive tribes, a great many of the unbelievers in Western culture appear to have been fully exposed to these facts. But this does not usually move the proponents of the argument. Aside from certain rejoinders that will be discussed later, their formulations usually contain highly elastic words that make possible a speedy disposition of apparent negative instances. The

unbeliever may have been exposed to the relevant facts but not "adequately"; or he may have been exposed to them adequately, but his religious nature may have been "clouded" or "defiled"; or, contrary to outward appearances, the unbeliever may really believe, but the belief may be so faint as to be barely perceptible. This last method was adopted by Hodge when faced with the negative evidence drawn from the observations of blind deaf-mutes. Unbelievers like Ludwig Büchner had pointed to several famous cases, including that of Laura Bridgman, who either could not be brought to form an idea of God at all or who reported that, prior to instruction, no such idea had entered their minds. As far as is known, Hodge never made any empirical studies of blind deaf-mutes, but this did not prevent him from replying with full confidence. "The knowledge obtained by Christian instruction so much surpasses that given by intuition," he assures us, that the purely intuitive knowledge of the blind deaf-mute "seems as nothing" (ibid., p. 197).

At this stage one must raise the following questions: Under what circumstances would a human being *not* possess an innate belief in God? More specifically, let us suppose that a person observes the facts of organic adjustment and experiences a sense of duty and obligation but nevertheless maintains, with all appearance of sincerity, that he does not believe in God. Under what circumstances would it be true to say that he had observed the facts adequately, that his religious nature was not clouded or defiled, but that he nevertheless had no belief in God? Unless these questions are satisfactorily answered, the argument does not really get off the ground. For it is meant to be based on an empirical premise, and the premise will not be empirical if it is retained no matter how human beings may respond to the stimuli that are supposed to activate the innate disposition to believe in God.

Waiving this difficulty, and granting that the distinction between the two senses in which a belief may be instinctive circumvents the first of Locke's objections, the argument would still be open to his second criticism, namely, that the universality of an idea or a belief does not establish its innateness. It may well be possible, Locke argued, to account in other ways for the universal occurrence of an idea or the general agreement on a topic. The ideas of the sun and heat, he wrote, are also universal without being "natural impressions on the mind" (op. cit., Book I, Sec. 2). Locke, who was primarily concerned with the origin of the *idea* of God rather than with any question of the universality of belief in God, claimed that he could give an adequate account of how this idea arose in the human mind without an appeal to innate ideas, and John Stuart Mill later offered a detailed account of how belief in God might be universal without being instinctive. Reasons for rejecting such accounts would have to be offered before one could infer the innateness of belief in God from its universality.

Mill, one of the few great philosophers of recent times to discuss this argument in detail, objected to it on several other grounds as well. Assuming a belief to be innate or instinctive, he asked why this should be any reason whatsoever for regarding it as true. The only justification for this transition that Mill could think of he dismissed as begging the question. This is "the belief that the human mind was made by a God, who would not deceive his creatures" (*Three Essays on Religion,* p. 156), which of course presupposes what is to be proved. Whether this is in fact the only possible justification of the inference from the innateness of a belief to its truth, Mill's observation that the former does not by itself afford evidence for the latter seems to be very well taken. The force of his point, however, may be obscured because instinctive beliefs are frequently referred to as a priori and because this and related expressions are ambiguous. In this present context, calling a proposition a priori simply means that it was not affirmed as the result of instruction. In other contexts, and more commonly, to say that a proposition is a priori logically implies that it is a necessary truth and hence requires no empirical confirmation. It should be clear that if a proposition is a priori in the former sense, it does *not* automatically follow that it is a necessary truth or a truth at all. If an empirical or, more generally, a nonnecessary proposition were instinctively entertained, it would stand just as much in need of proof or confirmation as any other; and, except for a few defenders of the Ontological Argument, believers and unbelievers alike are satisfied that "God exists" does not express a necessary proposition.

Flint and others complained that Mill was unfair because there are versions of the argument that cannot be accused of circular reasoning. It will become clear in the next section that the antiskeptical form of the argument is in fact immune from such a criticism (and Mill was probably not familiar with it). However, it is difficult to see how the version of the biological argument which we have been discussing can bridge the transition from the instinctiveness of belief to its truth without introducing God as guarantor of the instinct's trustworthiness.

INNATE YEARNING FOR GOD. There is, however, another version of the biological argument which can

perhaps be stated in such a way as to avoid the charge of circular reasoning. This version, moreover, has certain additional advantages over the one considered previously. "All the faculties and feelings of our minds and bodies," writes Hodge, "have their appropriate objects; and the possession of the faculties supposes the existence of those objects." Thus the eye, "in its very structure, supposes" that there is light to be seen, and the ear would be "unaccountable and inconceivable" without the existence of sound. "In like manner" our religious feelings and aspirations "necessitate" the existence of God (op. cit., p. 200). "The yearning for some kind of God," in the words of Chad Walsh, a contemporary defender of the argument, "does point toward an in-built hunger in each of us—a hunger for something greater than we are." But every other hunger has its normal gratification. This is true of physical hunger, of love and sex, and of our craving for beauty. If, similarly, our religious hunger did not have its proper gratification, it would be difficult to see "how it got built into our natures in the first place. What is it doing there?" (*Atheism Doesn't Make Sense,* p. 10).

This version of the argument escapes one of the difficulties of the version considered earlier. It can very plausibly be argued that absence of belief in God does not prove absence of a yearning for God; and in fact there are undoubtedly unbelievers who wish they could believe. But, granting that the existence of unbelievers does not prove that the wish for God's reality is not universal, this version of the argument nevertheless appears to be open to a number of fatal or near-fatal objections. To begin with, there seem to be exceptions here also. There seem to be people who not only do not believe in God but who are also devoid of any hunger for God. Furthermore, even if this hunger were universal, it might, as before, be possible to explain it on some basis other than that it is innate; or, putting the point differently, one would have to be satisfied that all such explanations are inadequate before one could conclude that it is innate. More seriously and waiving such objections as that the analogy between "religious hunger" and either physical hunger or having organs like eyes and ears is more than dubious, statements to the effect that we have eyes because there is light are objectionable on several grounds. Neither the observed facts nor contemporary biological theory warrants any such assertion. We are entitled to say that we have eyes and that there is light and that the eyes are useful because there is light, so that, other things being equal, organisms with eyes are likely to win out in the struggle for survival against organisms without eyes. Many kinds of biological variations are not similarly useful, but these are rarely noticed by proponents of the argument. When reading the teleological formulations of these writers—Walsh's question "How did the longing get *built into* [italics added] our nature in the first place?" or Hodge's remark that "possession of the faculties *supposes* [italics added] the existence of the appropriate objects"—one can hardly avoid the suspicion that although God is not explicitly brought into the premise of the argument, these authors surreptitiously introduce a designer who supplied organisms with their native equipment in order to fit them to their environment. It might indeed be possible to establish the existence of a designer on other grounds, but in the present context the defender of the argument is guilty of circular reasoning and thus would not escape Mill's stricture. No such circularity is involved if the instinctive desire is made the basis of an argument for *immortality* after the existence of a beneficent deity has been independently established. Dugald Stewart, in his *Philosophy of the Active and Moral Powers of Man* (Edinburgh, 1828, Book III, Ch. 4), offered such an argument, observing, "whatever desires are evidently implanted in our minds by nature, … we may reasonably conclude, will in due time be gratified under the government of a Being infinite both in power and goodness." Stewart was not guilty of circular reasoning, since he thought that he had previously proved the existence of God by means of the Design Argument.

THE ANTISKEPTICAL DILEMMA

JOYCE'S ARGUMENT. One of the most carefully developed statements of the second main form of the Common Consent Argument is found in G. H. Joyce's *The Principles of Natural Theology.* There are three stages to this form of the argument. (1) As in the biological versions, it is contended that practically all human beings, past and present, can be counted as believers in God. Here, however, it is not maintained that there are innate tendencies in human beings to believe in God. If anything, the opposite is true: people crave liberty of action and resent any being with superior authority. If, nevertheless, nearly all human beings are "perfectly certain" of the existence of their "absolute Master," this can be so only because "the voice of reason" is so clear and emphatic: "All races, civilized and uncivilized alike, are at one in holding that the facts of nature and the voice of conscience compel us to affirm this [the existence of God] as certain truth" (op. cit., p. 179). (2) If the whole of humankind were mistaken in a conclusion of this kind, it would follow that something is amiss with man's intellect, that "it is idle for man to search for truth." In that event, pure skepticism would be the only alternative. (3) How-

ever, all of us, unless we wish to be perverse, realize that "man's intellect is fundamentally trustworthy—that, though frequently misled in this or that particular case through accidental causes, yet the instrument itself is sound" (ibid.). Since reason is fundamentally trustworthy, universal skepticism is not a serious alternative to the acceptance of humankind's conclusion that God exists.

Some writers, though not Joyce, are concerned to add that on this topic great men are at one with the masses of believers. "Even for the independent thinker," writes John Haynes Holmes, "there is such a thing as a consensus of best opinions which cannot be defied without the weightiest of reasons" ("Ten Reasons for Believing in Immortality," in *A Modern Introduction to Philosophy,* edited by Paul Edwards and Arthur Pap, New York, 1965, p. 241). If there were no God and no afterlife, the deceived would include, in the words of James Martineau, "the great and holy whom all men revere." Whom are we to reverence, he goes on, "if the inspirations of the highest nature are but cunningly-devised fables?" (ibid.).

Joyce is aware that the "common consent" of the human race on this subject has been challenged from two sides. To the criticism that there are unbelievers at the present time and that the history of Western countries records instances of other unbelievers, he replies that these are so few in comparison with the number of believers that they do not affect the "moral unanimity of the race," and he adds that he never meant to claim that literally everybody who ever lived has affirmed the existence of God. To the criticism that there are primitive peoples without a belief in God or at least in one God, Joyce replies that there is in fact no race without religion and that even where there is belief in a plurality of gods, it is invariably found that "the religion recognizes a supreme deity, the ruler of gods and men" (p. 182). Joyce concedes that the supreme deity of primitive religions often lacks some of the characteristics attributed to God by Christian and Jewish monotheists. But this does not affect the argument, since "an idea of God does not cease to deserve that name because it is inadequate" (p. 181). A person may be said to believe in God if he believes in a "Supreme Being, personal and intelligent, to whom man owes honor and reverence" (ibid.), regardless of what else he also believes or fails to believe.

OBJECTIONS TO JOYCE'S ARGUMENT. The claim that belief in God is practically, if not indeed strictly, universal in the human race is shared by defenders of both forms of the Common Consent Argument. We shall discuss the difficulties of such a position in some detail in the next section. Meanwhile, it should be pointed out that even if the moral unanimity of humankind on this subject is not questioned, the argument, as presented by Joyce, appears to be open to two powerful objections.

To begin with, it presupposes that all or most believers in God arrive at their belief by means of reason or the intellect. If this is not the case, then the argument clearly fails, since nothing derogatory about reason would follow if it was not the source of the mistaken conclusion. In actual fact, it seems more than doubtful that the majority of men use reason in any significant sense in arriving at belief in God or even in fortifying their belief after their original acceptance of it. In making this observation, "reason" is not used in any specially narrow sense. A person may, in a perfectly familiar and proper sense, be said to have arrived at a conclusion by means of reason without having set out any formal arguments. However, there seems to be a good deal of evidence that the majority of human beings came to their belief in God by traditional indoctrination. Nor is it particularly plausible to maintain that originally this belief was the product of reason. If reason had anything to do with it, its role, in the opinion of most contemporary psychologists, was probably quite subsidiary. Joyce's view that man's natural inclinations would lead to denial rather than to belief in God seems highly doubtful. There is a good deal of disagreement about the exact psychological mechanisms involved, but the majority of psychologists seem to think that man's loneliness and helplessness, as well as his animistic propensities, incline him to belief in protective (and also hostile) cosmic powers. This does not, of course, mean that such beliefs cannot also be adequately supported by rational considerations, but it does undermine Joyce's argument.

It should be emphasized that the view just outlined is by no means confined to antireligious psychologists. Fideistic theists would most certainly endorse these observations, as would many believers who have stressed the evil and suffering in the observable world. Indeed, most of the defenders of the biological form of the Common Consent Argument would be opposed to Joyce's account. "Our own consciousness," in the words of Charles Hodge (op. cit., pp. 199–200), "teaches us that this is not the ground of our own faith. We do not reason ourselves into the belief that there is a God; and it is very obvious that it is not by … a process of ratiocination, that the mass of the people are brought to this conclusion."

Even if this difficulty could be overcome, however, and if it were granted that human beings arrive at their belief in God by reason, Joyce's argument would still be in

trouble. If "universal skepticism" stands for the view that human beings can never find the true answer to *any* question, then it is not implied by the rejection of the universal belief of humankind in God. All kinds of other explanations of the "universal error," short of "the radical untrustworthiness" of human reason, seem possible and cannot be ruled out without further ado. It has, for example, been widely held by Kantians, nineteenth-century positivists, and fideists that human reason, while trustworthy as long as it deals with empirical and purely formal issues, is not fit to handle questions transcending experience.

As for the observations of Martineau and Holmes, several points are in order. To begin with, "appeals to the best opinion" are of logical force only in areas in which there are experts, as there are in physics or dentistry, for example. In this sense there is no such thing, either for the independent thinker or for anybody else, as a "consensus of best opinion" when we come to such questions as the existence of God or the immortality of the soul. Furthermore, just as there have been great men and great philosophers who believed in God, so there have also been great men and great philosophers who did not. Since presumably both groups cannot be right, we will be left with the conclusion that men who deserve to be "reverenced" are occasionally mistaken—no matter which view is taken on this subject. Finally, there is nothing about the loftiness of an "inspiration" that guarantees its truth. People whose loftiness makes them believe the best about their neighbors are probably as often mistaken as those whose lack of loftiness makes them believe the worst.

IS BELIEF IN GOD UNIVERSAL?

Let us now turn to a discussion of the detailed objections to the premise which all forms of the Common Consent Argument share, namely, that all or practically all human beings are believers in God.

ANTHROPOLOGICAL OBJECTIONS. To begin with, there is a series of objections based on what is known or allegedly known about primitive tribes and about religions which are not monotheistic. We have already seen that Locke believed, on the basis of the reports of travelers, that there were whole nations without the notion of God. This view was widely advocated by anthropologists and sociologists in the nineteenth century, many of whom did not rely on the reports of others but spent long periods studying the beliefs and habits of primitive peoples at first hand. It was developed in considerable detail by Sir John Lubbock in his pioneering work, *Prehistoric Times,* and it had the unqualified endorsement of Charles Darwin, who, in *The Descent of Man* (Ch. 3), reported confirmations in his own experience with the Fuegians. The denial that belief in God is universal was an essential part of the position of the so-called evolutionary anthropologists. They maintained that there was a gradual transition from animism, via fetishism, to a belief, first, in many gods and then, finally, in a single God. Several of these writers, however, used the word *religion* very broadly to include belief in any unseen spiritual agencies, and in this sense both E. B. Tylor (the eminent evolutionary anthropologist) and Darwin were ready to admit that *religious* belief appeared to be universal among the less civilized tribes. The philosopher Fritz Mauthner, who followed this tradition, expressed himself very strongly on the subject. In *Der Atheismus und seine Geschichte im Abendlande* (Vol. IV, Ch. 10) he accused Christian missionaries of "translation impertinence" in dragging out of aborigines the confession that they believed in a heavenly Father, when more careful investigation revealed that they did not mean anything of the kind. He also protested against the trick, as he called it, by advocates of the *consensus gentium,* of using the word *religion* ambiguously—at first in the broad sense of Tylor and Darwin, in which it may be plausible to maintain the universality of religion, and then shifting to the narrower sense, required by their argument, in which it implies belief in God or gods.

Critics of the argument have also pointed out that there are numerous tribes believing in polytheism without having in their theology one supreme deity. Hence, even if the argument were otherwise sound, it could not prove the existence of a single Supreme Being.

Finally, it has been maintained that there are religions, of which Buddhism is the most notable instance, which have no belief in God at all.

To the last of these criticisms, the customary answer is that, while the founder of such a religion may indeed not have believed in God or gods, once these religions spread, they acquired theologies—and sometimes exceedingly extravagant ones at that. Joyce, who was familiar with this objection, regarded the example of Buddhism as highly favorable to his argument. It was his contention that no religion or philosophical system which rules out belief in God "has ever succeeded in maintaining a prominent hold on any people" (p. 197). In China, Buddhism flourished, but there it became a polytheistic religion. In India, on the other hand, where the original agnostic teachings were not substantially

changed, the Buddhist creed could not hold its own and had to give way to modern Hinduism.

The existence of polytheistic religions is not, of course, questioned by defenders of the Common Consent Argument. Some, indeed, like Flint and Mercier, are willing to concede that the argument, by itself, does not favor a stronger conclusion than that God *or* gods exist. This, however, is not the usual reaction. Recent advocates of the argument have commonly challenged the entire scheme of the evolutionary anthropologists. Basing their argument largely on the work of the Austrian anthropologist Father Wilhelm Schmidt (1868–1954) and others belonging to the "theological school," they deny that polytheism antedates monotheism and insist, furthermore, that in every polytheistic religion there is one supreme deity. According to Schmidt, the simplest peoples are also the oldest, and they are believers in a very pure monotheism. Their God possesses all the main attributes of the God of Christianity and Judaism: he is the creator of reality, he supplies the foundation of morality, and he is also omnipotent, omniscient, and supremely good. As societies became more complex, this monotheism became transformed into various kinds of animism, polytheism, and ancestor worship. Even among these later cultures, Schmidt finds "a clear acknowledgment and worship of a supreme being," while all other "supernormal beings" are regarded as far inferior and subject to him.

It would be idle to get involved here in the controversies between Schmidt's school and other schools of anthropology, especially since there are objections to the Argument from Common Consent which can be evaluated without taking sides on anthropological issues. Perhaps the only comment worth making is that while contemporary anthropologists are willing to credit Schmidt and other members of the theological school with some sound criticisms of the evolutionary anthropologists and with a good deal of impressive field work, the great majority of them regard his basic theories as quite unsupported by the available evidence.

UNBELIEVERS IN THE WESTERN WORLD. The other main challenge to the claim that belief in God is universal consists in pointing to the unbelievers in Western culture. It is admitted that unbelievers are a minority, but it is argued that they are and have for some time been too significant a minority not to affect the "moral unanimity of mankind" on this subject. This challenge and the various attempted rebuttals deserve, but have very rarely received, extended discussion.

"Belief" redefined. One way in which the significance of individual unbelievers may be discounted is apparent in the tendency of some Protestant writers to define "belief in God" or "religion" or both so broadly as to make it virtually impossible for a human being not to be a believer or to be religious. In our own day such writers frequently follow Paul Tillich's definition of an atheist as someone who believes that "life is shallow" and of an irreligious person as someone who has "no object of ultimate concern." However, the use of such definitions to do away with unbelievers achieves a victory which is purely illusory. It will now indeed be possible to call a man like Denis Diderot a believer and religious. But in the sense in which there was a dispute about the existence of unbelievers, namely, whether there are people who do not believe in the existence of what is usually understood by "God," Diderot and countless other people will still have to be classified as unbelievers. Moreover, if the premise of the Common Consent Argument is now a true proposition, with "believer" used in the *new* sense, the conclusion established, if any, would not be the one originally aimed at. It would not show that God exists but rather, using Tillich's redefinitions, that life is not shallow and that there are objects of ultimate concern.

Unbelievers discounted as abnormal. One of the favorite devices used to defend the *consensus gentium* against irritating exceptions has been the charge that unbelievers are in effect too morally or mentally defective to count as representative of human opinion. Strangely enough, this tactic was used by Pierre Gassendi, who was highly critical of Herbert of Cherbury's argument and from whom, in view of his own independence of thought, one might have expected something better. In the course of expounding his version of the argument, Gassendi minimized the number and importance of atheists, declaring that they are either "intellectual monstrosities" or "freaks of nature." More recently this defense was adopted by some eminent nineteenth-century Protestant theologians. Thus A. H. Strong, in a text that was widely used in Protestant seminaries, observed that just as the oak must not be judged by the "stunted, the flowerless specimens on the edge of the Arctic Circle," so we must not take account of unbelievers in judging the nature of man (*Systematic Theology,* Vol. I, p. 56). One of the rivals of Strong's book was Hodge's *Systematic Theology.* Hodge was not to be outdone. A man's hand, he reminds us, may be so hardened as to lose the sense of touch, but this does not prove that the hand is not "normally the great organ of touch." Similarly, it is possible that "the moral nature of a man may be so disorganized by vice or by a false philosophy as to have its testimony for the existence of God

effectually silenced" (op. cit., p. 198). Human beings cannot abandon belief in God "without derationalizing and demoralizing their whole being" (p. 201); and the belief, or rather lack of belief, of such a "derationalized" and "demoralized" individual does not count.

Perhaps two brief comments will be sufficient here. First, Hodge at least is begging the question when he refers to the "false philosophy" that silences the testimony for the existence of God. The question is precisely whether this is a false philosophy. If this were already known, there would be no need for the Argument from Common Consent. Second, and more important, anybody having the slightest familiarity with the history of unbelief must surely protest that many of the outstanding thinkers of the last two centuries were avowed unbelievers. Like other mortals, they may have been frequently in error, but to dismiss them as freaks, to compare them to stunted, flowerless oaks, or to regard their moral nature as disorganized by vice is surely outrageous nonsense.

Unbelief discounted as an illusion. Some of those who regard the unbeliever as "unnatural" or "monstrous" do not, perhaps, wish to refer to any actual human being. This may be so because some of them also maintain that *really* everybody is a believer in God even though he may say the opposite and believe that he believes the opposite. (The strategy here is rather different from the redefinitional maneuver described above.) Hodge, for example, offers two reasons in support of such a position. First, unbelief is such an unnatural state that it cannot last. "Whatever rouses the moral nature, whether it be danger, or suffering, or the approach of death, banishes unbelief in a moment (ibid., p. 198). There seems to be an obvious answer to this. It is true that unbelievers have become converted or reconverted on occasions, but it is equally true that others have remained unbelievers right to the end of their lives. Furthermore, those who became converted must *really* have been unbelievers before their change of position, or else there would have been no conversion. To this it must be added that there are also shifts in the opposite direction, and if a person does not count as an unbeliever *at all* because he ultimately becomes a believer, then those who change from belief to unbelief will have to be counted as unbelievers exclusively.

Hodge's second reason would probably have a much wider appeal. "It is hardly conceivable," he writes, "that a human soul should exist in any state of development, without a sense of responsibility, and this involves the idea of God. For the responsibility is felt to be not to self, nor to men, but to an invisible Being, higher than self and higher than man" (ibid., p. 197). Hodge is certainly not alone in taking the line that if a person is a moral creature and not lacking in sensibility, then he must be a believer in God. Even at the present time there are many people who seem to rule out a priori the possibility that a good person can be an unbeliever. To give just one illustration, Justice William O. Douglas wrote a highly laudatory preface to a recent collection of the court pleas of Clarence Darrow (*Attorney for the Damned*). Darrow had repeatedly stated and defended his agnosticism, and he never once retracted this position. Nevertheless, seeing that Darrow was such a kind and compassionate man, Douglas remarks: "Darrow met religious bigotry head-on … but he obviously believed in an infinite God who was the Maker of all humanity."

There are several confusions in reasoning of this kind. To begin with, the criteria which all of us employ to determine that a man is kind, that he does not lack sensibility, that he shows responsibility in his relations with other human beings—that, in short, he is a "moral person" or a good man—are quite distinct from those which we employ when determining that he is a believer in God. This at any rate must be so if the statement that all believers in God and only believers in God are good is to be, as it is usually taken to be (both by those who accept it and by those who deny it), a factual claim and not a tautology.

Second, the claim that responsibility is invariably felt not to oneself or to other men but to an invisible Being is unwarranted. Assuming that some people do on occasions feel responsibility to an invisible Being, this is certainly not true of all. If people who assure us that they feel responsible, but not to an invisible Being, are to be discounted or disbelieved, why should we count and accept the assurances of those who say that they feel responsible to the invisible Being? Moreover, it appears that the attitude, even of religious believers, is not generally in accord with Hodge's account. If a believer borrows money and considers himself obligated to return it, he surely, like an unbeliever, regards himself as obligated to the person who lent him the money and not to anybody else. Suppose believers were asked the following question in such a situation: "If an atheistic philosopher persuaded you that God does not exist, but if otherwise the situation remained exactly the same—you needed the money badly, your friend helped you without hesitation, you promised to repay him as soon as possible, and so on—would you still consider yourself obligated to repay the loan?" It seems very doubtful that more than a handful of believers would reply that they no longer regarded themselves as obligated.

Questions about whether a person who says that he believes or disbelieves a proposition and who is apparently not lying, does really believe or disbelieve it, are complicated by the fact that "belief" is an ambiguous word. Without entering into any subtleties or attempting an elaborate analysis, it may be granted that there is nothing absurd in the suggestion that a person may sincerely regard himself as an unbeliever when in fact he is a believer, or vice versa. It is helpful in this connection to distinguish belief in terms of verbal responses and positions adopted in purely theoretical contexts from belief insofar as it is exhibited in actions and in involuntary responses, especially to critical situations.

Bertrand Russell discusses this question in a little-known essay titled "Stoicism and Mental Health." He points out that people who say, with all appearance of sincerity, that they believe in an afterlife seem to fear their own death or regret the death of their friends as much as those who say that they do not believe in an afterlife. He explains this "apparent inconsistency" by remarking that the belief in the afterlife is in most people "only in the region of conscious thought and has not succeeded in modifying unconscious mechanisms" (*In Praise of Idleness,* paperback ed., London, 1960, pp. 133–134). Many of us, like Russell, are inclined to regard the latter, the sense in which belief is expressed in involuntary responses and not merely in theoretical contexts, as the "deeper" sense. We say that a man has reached and avows a certain conclusion, but "deep down" he really believes the opposite. It must be conceded to the defender of the Argument from Common Consent that there are people who are unbelievers in the verbal and theoretical sense but who in a deeper sense do believe in God. This is notoriously true of some who are brought up in a religious home and much later come under the influence of skeptical thinkers.

Nevertheless, the Common Consent Argument is not really helped by this admission. For, in the first place, there can be no reasonable doubt that a good many people are unbelievers in both senses; and second, not a few cases are known of believers, that is, people who sincerely believe in God in terms of their verbal and theoretical responses whose actions show them to be unbelievers "deep down." This fact has been repeatedly stressed by religious writers when castigating some of the members of their own groups as "practical atheists."

Unbelief seen as a negligible influence. Some defenders of the argument are quite ready to admit the existence of highly educated unbelievers. In other words, they question neither the genuineness of the lack of belief nor the intellectual standing of unbelievers. However, they add to this the fact that unbelievers have failed and are bound to fail to make any major inroads on humankind at large. "We find a disposition on the part of some few philosophers to dispute the validity of the belief," writes Boedder (op. cit., p. 68), "but nevertheless the belief has proved to be persistent and indestructible in the mass of humankind. It is this persistency among the mass of men, retained even in the teeth of skeptical opposition, on which our argument is based."

Sometimes a comparison is made between the unbelievers and the philosophers who deny the existence of an external world or the reality of space and time but are rightly laughed off by ordinary people whose common sense is intact. Granting that the ordinary person is in some sense right as against the philosopher who denies the reality of time, to confine ourselves to one such case, the comparison seems to be very weak in more ways than one. For one thing, unbelief in matters of religion is not at all confined to professional philosophers or to people who are naturally referred to as intellectuals. Furthermore, as G. E. Moore has pointed out, the philosophers who *say* such things as "time is unreal" and who presumably in some sense also believe this, also say things and cannot help saying things which indicate that *they also do not believe it.* The very philosophers who say that time is unreal nevertheless use clocks, complain when their students are late, plan for the future, and engage in the same activities that the ordinary man regards as presupposing the reality of time. Nothing even remotely comparable can be found in the case of unbelievers as a class.

However, returning to the original question, it is not at all certain that unbelieving philosophers and other critics of belief in God have not significantly affected the masses. There seems to be a good deal of evidence to the contrary; but even if it were true and the impact has in fact been negligible, this could be explained quite plausibly without supposing either that belief in God is inherent or, as Boedder claims, that reason, properly used, is certain to lead to a theological conclusion.

ARE MEN BY NATURE "GOD-SEEKERS"?

There are philosophers and psychologists of influence who either do not believe in God at all or who, at any rate, do not favor the enterprise of buttressing belief in God by means of "proofs" but are nevertheless concerned to maintain that human beings are by nature religious—that they are, in Max Scheler's phrase, "God-seekers." They would point out that it is this question of "philosophical anthropology," and not any question about the

validity of the Common Consent Argument, which is of real interest and human importance. Though perhaps invalid as a proof of the *existence of God,* the Common Consent Argument does embody an important insight about the *nature of man.*

These writers are a great deal more sophisticated than most of the traditional defenders of the argument, whose views we considered in preceding sections. They do not at all deny that, in the most obvious sense, the world is full of unbelievers, but they would add that a great many of these unbelievers feel a strong urge to worship something or somebody and therefore invent all kinds of surrogate deities. Man's "gods and demons," writes Jung, "have not disappeared at all; they have merely got new names." Those, in the words of Miguel de Unamuno, "who do not believe in God or who believe that they do not believe in Him, believe nevertheless in some little pocket god or even devil of their own." "Religious agnosticism," writes Scheler, "is not a psychological fact, but a self-deception … it is an essential law [*ein Wesensgesetz*] that every finite spirit believes either in God or in an idol. These idols may vary greatly. So-called unbelievers may treat the state or a woman or art or knowledge or any number of other things as if they were God" (*Gesammelte Werke,* Vol. V, pp. 261–262). Scheler adds that what needs explanation is not belief in God, which is original and natural, but unbelief or, rather, belief in an idol. The situation is not infrequently compared with the sexual instinct and what we know about the consequences of its suppression. If the sexual instinct does not find natural gratification, it does not cease to be operative but becomes diverted into other and less wholesome channels. The worship of institutions and human deities is said to be a similarly pathological phenomenon.

An evaluation of this position, which amounts in effect to an endorsement of the theory of the religious instinct without inferring the existence of God from it, is not possible here because it would involve elaborate discussions of child psychology and the causation of neurosis and "alienation." Here we can only observe that in the opinion of many contemporary thinkers there is no reason whatever to suppose that human beings are "by nature" religious. In their opinion the "hunger for God," in its orthodox no less than in its newer "substitute" expressions, is invariably the result of certain deprivations and traumatic experiences. People who suffer from insufficient contact with other human beings and who do not find the natural world satisfying will tend to experience longings for something supernatural or feel a need to endow human beings with supernatural attributes. Some of these writers would go further and maintain that traditional religion, through its life-denying morality and irrational taboos, is itself in no small measure responsible for the existence of the type of personality that displays the hunger for God. Sigmund Freud, who took this position, conceded that those in whom the "sweet—or bittersweet—poison," as he called religion, had been instilled early in life were unable to dispense with it later on. The same, he added, is not true of others who have been brought up more soberly. "Not suffering from neurosis," they will "need no intoxicant to deaden it."

See also Cambridge Platonists; Cicero, Marcus Tullius; Clement of Alexandria; Cosmological Argument for the Existence of God; Darwin, Charles Robert; Degrees of Perfection, Argument for the Existence of God; Freud, Sigmund; Gassendi, Pierre; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Herbert of Cherbury; Hooker, Richard; Locke, John; Martineau, James; Mercier, Désiré Joseph; Mill, John Stuart; Moore, George Edward; Moral Arguments for the Existence of God; Ontological Argument for the Existence of God; Russell, Bertrand Arthur William; Scheler, Max; Seneca, Lucius Annaeus; Stewart, Dugald; Teleological Argument for the Existence of God; Tillich, Paul; Unamuno y Jugo, Miguel de.

Bibliography

There is no full-length study in any language of the different forms of the Common Consent Argument. The major reference works contain either no entries or else very brief and unhelpful ones. Even Rudolf Eisler's "Consensus Gentium," in *Wörterbuch der philosophischen Begriffe,* 3 vols., 4th ed. (Berlin: E.S. Mittler, 1930), devotes less than a page to this subject.

The fullest defenses of the argument are found in Charles Hodge, *Systematic Theology,* 3 vols. (New York: Scribners, 1871–1873), Vol. I; Bernard Boedder, *Natural Theology* (London, 1896); and G. H. Joyce, *The Principles of Natural Theology* (London: Longmans, Green, 1923). Briefer discussions also favoring the argument are contained in A. H. Strong, *Systematic Theology,* 3 vols. (Philadelphia: Griffith and Rowland, 1907), Vol. I; Robert Flint, *Theism* (London: Blackwood, 1877); Hermann Ulrici, *Gott und die Natur,* 3 vols., 3rd ed. (Leipzig: T.O. Weigel, 1875), Vol. I; and Cardinal Mercier, *A Manual of Modern Scholastic Philosophy,* translated by T. L. and S. A. Parker, 2 vols., 3rd ed. (London: K. Paul, Trench, Trubner, 1928), Vol. II. The famous nineteenth-century biologist G. J. Romanes, in "The Influence of Science upon Religion," which forms Part I of his *Thoughts on Religion* (Chicago: Open Court, 1895), defends the biological form of the argument as proving not that there is a God but that *if* "the general order of nature is due to Mind," then the character of that Mind is "such as it

is conceived to be by the most highly developed form of religion."

A popular contemporary statement of the biological version of the argument is advanced in Chad Walsh, *Atheism Doesn't Make Sense* (Cincinnati, n.d.). Among earlier writers, Cicero defended the argument in *De Natura Deorum,* Book II, Sec. II, translated by C. D. Yonge as *The Nature of the Gods* (London, 1892); by Herbert of Cherbury in *De Veritate,* translated by M. H. Carré (Bristol, U.K.: University of Bristol, 1937); and by Pierre Gassendi in *Syntagma Philosophicum,* in his *Opera Omnia,* edited by H. L. H. de Montmorency and F. Henri, Vol. I (Lyons: Lavrentii Anisson and Ioan, 1658).

One of the earliest criticisms of the argument was by Locke, in *Essay concerning Human Understanding* (London: Thomas Bassett, 1690), Book I. There are brief and unsystematic critical discussions in several of the works of the freethinkers of the seventeenth, eighteenth, and nineteenth centuries, including Bayle, Paul-Henri Holbach, and Büchner, but the first detailed and systematic critique is found in J. S. Mill, *Three Essays on Religion* (New York: Henry Holt, 1874). More recently, the argument has been attacked in John Caird, *Introduction to the Philosophy of Religion* (Glasgow: J. Madehose, 1880); Fritz Mauthner, *Der Atheismus und seine Geschichte in Abendlande,* Vol. IV (Stuttgart, 1923); and in two books by Josef Popper-Lynkeus: *Das Individuum und die Bewertung menschlicher Existenz* (Dresden: C. Reissner, 1910) and *Über Religion* (Vienna: R. Löwit, 1924). Popper-Lynkeus' criticisms are, for the most part, an elaboration of David Hume's remark that "the conviction of the religionists, in all ages, is more affected than real." Hume's discussion of this topic occurs in Sec. XII of *The Natural History of Religion* (London, 1757; critical ed. with introduction by H. E. Root, London: A. and C. Black, 1956), a work which also anticipates many of the conclusions of the evolutionary anthropologists of the nineteenth century. There is a discussion, at once critical and sympathetic, in Miguel de Unamuno, *The Tragic Sense of Life in Men and in Peoples,* translated by J. E. Crawford Flitch (New York, 1921).

Two more recent works surveying the evidence concerning the religious beliefs of primitive tribes are Guy E. Swanson, *The Birth of the Gods* (Ann Arbor: University of Michigan Press, 1960), and W. J. Goode, *Religion among the Primitives* (Glencoe, IL: Free Press, 1951). Wilhelm Schmidt's theory is stated in his *The Origin and Growth of Religion,* translated by H. J. Rose (London; Methuen, 1931). A view similar to Schmidt's was expressed by Andrew Lang in various works, including *The Making of Religion* (London: Longmans, Green, 1898) and *Magic and Religion* (London: Longmans, Green, 1901). The *Anthropological Review* 2 (1864): 217–222, contains an interesting summary of an address by the Reverend F. W. Farrar, "On the Universality of Belief in God and in the Future State," in which a great deal of evidence is presented to the effect that neither belief in God nor belief in an afterlife is universal. The discussion following Farrar's address is also reported, and most of the participants, including W. R. Wallace, fully supported Farrar's negative conclusion.

J.-H. Leuba, *The Belief in God and Immortality* (Boston: Sherman, French, 1916), presents evidence concerning belief and unbelief among academic groups in the United States in the early years of the twentieth century. Unfortunately, there has been virtually no study in depth of religious belief and unbelief in the general population of any country.

Jung's views on the natural religious needs of human beings and the sickness of modern men who have lost their religion are stated in *Psychology and Religion* (New Haven, CT: Yale University Press, 1938) and *Modern Man in Search of a Soul,* translated by W. S. Dell and C. F. Baynes (New York: Harcourt Brace, 1933). Scheler's similar views are found in his *Vom Ewigen im Menschen,* in *Gesammelte Werke,* 4th rev. ed., Vol. V (Bern: Francke, 1954). The opposite position is defended by Sigmund Freud in *The Future of an Illusion,* translated by W. D. Robson-Scott (New York: H. Liveright, 1928), and Wilhelm Reich, in *The Mass Psychology of Fascism,* translated by T. P. Wolfe (New York: Orgone Institute Press, 1946). The views of Freud and Reich are foreshadowed in Ludwig Feuerbach, *The Essence of Christianity,* translated by George Eliot (London, 1853).

Paul Edwards (1967)

COMMON SENSE

Several things can be learned about common sense from Dr. Johnson's attempt to refute George Berkeley by kicking the stone. Its philosophical incompetence is not one of them. Dr. Johnson of course misunderstood Berkeley, and his misunderstanding was not a collapse of common sense. He thought that if stones had, as Berkeley said, no "material substance" and were collections of "ideas," a boot ought to go through them without resistance. And if Berkeley had been maintaining that solid objects were only apparently solid and were really collections of what we would ordinarily call ideas, the refutation would have been an appropriate reaction of common sense.

THE NOTION OF COMMON SENSE

Whatever other aspects of meaning the word *sense* may retain in the compound "common sense," it has prominently the force of sense as opposed to nonsense. In what is contrary to common sense there is always something more or less—but obviously—nonsensical. It produces the feeling, varying in strength according to circumstances, that argument is only precariously in place in dealing with it. For to deploy arguments at all *directly* against the manifestly absurd is to invest it with some intellectual dignity and to muffle its self-annihilating character. It is, moreover, to invite the suspicion that one has failed to recognize absurdity, and such failure has a very foolish look. As a man of redoubtable common sense, Dr. Johnson kept dialectic for the right occasion. He did not kick the stone formally in the name of common sense, but his action has traditionally been praised

and condemned as a piece of commonsense behavior. Yet he was demonstrating against a philosopher who was also determined to be on the side of common sense.

BERKELEY. Berkeley's notebooks contain the reminder to himself: "To be eternally banishing Metaphisics &c & recalling Men to Common Sense" (*Philosophical Commentaries*, No. 751). Confident that he could always secure the neutrality of common sense when he could not have its assistance, Berkeley went about his own metaphysical enterprise, which was to exhibit the dependence of physical objects on their being perceived. His *Three Dialogues* (1713) is studded with references to common sense: to opinions that are "repugnant" or "shocking" to it, to its "dictates," to the judgment of men of "plain common sense." The objections that have to be most carefully answered are those which appear to proceed from common sense. Since the issues concern mainly the world of perception, the man of common sense in the *Dialogues* is eminently the man who "trusts his senses," who will not tolerate the suggestion that the things he sees and handles are not real things but their mere representations.

The eighteenth century also brought into existence, in France and Scotland, philosophies of common sense—philosophies, to a greater or lesser degree, centered on this notion. They safeguarded what they held to be the beliefs (or "truths") of common sense by defending its authority and—in the Scottish philosophy—by exposing contraries of these beliefs to its blunt rejection.

COMMON-SENSE BELIEFS. It may be asked whether common sense had beliefs until philosophers engaged in its defense ascribed them to it. The *Oxford English Dictionary* lists a variety of meanings for the expression. Three of these, referring to a mental endowment, might be taken together: ordinary understanding—without which a man is out of his mind, or feeble-minded (an early meaning); ordinary, practical, good sense in everyday affairs; and the "faculty of primary truths." Ordinary understanding is not obviously, and practical, good sense is obviously not, the sort of thing that could stamp a set of beliefs with a special character. The third of these meanings is marked "philosophical." A further meaning must be noticed: "the general sense, feeling, or judgement of mankind." Here common sense seems to be a cluster of beliefs or persuasions, somehow "felt" to be true by most people. An argument drawn from common sense, in this case, would amount to an appeal to an ancient tribunal of opinion, common consent. (The most absolute modern proponent of this tribunal has probably been Lamennais, in his *Essai sur l'indifférence*, Paris, 1817–1823.)

Philosophers have frequently meant by common sense an intuitively based common consent. And the philosophers, during and after the eighteenth century, who have argued from common sense and for its beliefs have often thought of common sense in this way. They have, however, as often thought of it in a more ordinary way, as the common sense that is opposed—always at first sight, sometimes irreconcilably—to high and obvious paradox.

Can the common sense that is opposed to gross paradox properly be thought of as having beliefs, however strong? If there is some artificiality in saying that common sense has beliefs, there is none in speaking of its rejection of an opinion; the reason—it might be suggested—is that common sense does not declare itself in advance of attack upon it. The man of plain, ordinary common sense cannot readily be said, for instance, to believe that the things around him continue to exist in his absence—the idea of their not doing so does not cross his mind. But when he encounters the contrary opinion, his common sense asserts itself. On the supposition that the declarations of common sense are essentially reactive, to ascribe to it beliefs specified by what it rejects—and this the philosophers who have maintained its beliefs seem often to have intended—would be a minor linguistic innovation, justified in that it makes its commitments explicit. The supposition would have to be modified in some cases. It does not come naturally to us to speak of a *belief* in our personal identity through time, because this identity is something of which we are aware. Nevertheless, it can be argued that here also common sense has commitments which are not apparent before its reaction to various assertions.

REACTION TO SKEPTICISM. A philosophy of common sense is a natural reaction to the fact, or to the threat, of philosophical paradox or skepticism. The French Jesuit Claude Buffier (1661–1737) saw us as threatened, since René Descartes, with skepticism about all matters of fact beyond the range of our consciousness, the states of which cannot be doubted. What we need is unimpeachable authority for the fundamental convictions shared by all normal men about matters of fact with respect to which consciousness can give no guarantees. Common sense supplies it. It puts us into assured possession of such "first truths" as that there is an external world, that our minds are incorporeal, that we are capable of free agency. First truths have characteristic marks: No attack upon them, and no attempt to prove them, can operate from premises that surpass them in clarity or evidence. They are, and always have been, acknowledged by the vast

majority of humankind. Those who imagine they reject them act like others in conformity with them.

HUME. David Hume's work *A Treatise of Human Nature* (1739–1740) produced by reaction a more important philosophy of common sense than Buffier's. In parts of the *Treatise*—to isolate what gave the book its most generally "shocking" aspect—things were reduced to the contents of the mind and the mind to its contents. While many of Hume's conclusions are capable of a milder interpretation than they were given by his readers, Hume himself did not pretend that a number of them were anything but profoundly disturbing to our natural beliefs. At the same time he thought these beliefs had us too tightly in their grip for reasoning to be able to pry us loose. In the *Treatise* "nature" has the last word, but its meaning is left uncertain. We must submit, but whether in submitting to nature we are also submitting to truth is quite another matter. In Hume's later *An Enquiry concerning Human Understanding*, "common sense and reflection" are mentioned as correcting, in some degree, the indiscriminate doubt of an extreme skepticism, but nature and reasoning are still seen as coming into conflict. However, it should be remarked that there is another side to Hume in which these skeptical tendencies are in abeyance.

THE "SCOTTISH SCHOOL"

REID. A central purpose of Thomas Reid's *An Inquiry into the Human Mind on the Principles of Common Sense* (1764), and of his two later books, was, with Hume kept steady in view, to defend common sense against philosophical paradox and skepticism. It was for Reid a doubly difficult undertaking; if, as he held, the truths of common sense were self-evident, how could they be denied? And again, if they were self-evident, how could they be made evident when denied?

The great source of paradoxical or skeptical repudiations of common sense, Reid thought, was an innocent-looking theory that he believed philosophers had very generally adopted in order to explain the possibility of our awareness of anything beyond the present contents of our minds. According to this theory, such awareness is secondhand, necessarily mediated by "ideas" within our minds that are representative substitutes for external things. As its implications were drawn out, the theory, Reid maintained, committed philosophers to a steadily increasing range of conflict with common sense, with no stopping before "ideas," losing their representative character, monopolize existence. The "theory of ideas" is to be found in John Locke, needing only, Reid believed, an unsparing logic such as Hume's to produce Hume's world. (Locke's *An Essay concerning Human Understanding* [1690], has a deceptively commonsense air; its tone is down-to-earth, and experience is set up as the source of knowledge. Locke wanted no paradoxes, and when they were approached by what he said, he was not very efficient at drawing conclusions.)

The truths of common sense cannot be made evident by deductive proofs, but, Reid maintained, there is always absurdity in opinions contrary to its dictates. His most general procedure in defending common sense was to remind us of its command over us. Common sense has so fundamentally determined the scaffolding of ordinary language that the philosopher, in trying to word an opinion which is against common sense, is liable to need another language; and his utterance is continually threatened with incoherence between its structure and its content. The beliefs of common sense govern the behavior even of those who repudiate them in opinion, and they are only fitfully repudiated even in opinion; the paradoxical or skeptical philosopher is no sooner off his guard than he is believing with, as well as acting like, other men. Reid stressed a truism about matters of common sense: They lie within "the reach of common understanding." If this were not so, the judgment of the great bulk of humankind would carry no weight against a philosopher's superior competence. But in "a matter of common sense, every man is no less a competent judge than a mathematician is in a mathematical demonstration" (*Intellectual Powers*, Essay VI, Ch. 4). Whether or not something is a matter of common sense may well have to be investigated—prejudices shamming common sense must be exposed; what Reid denied is that the philosopher is in a better position than anyone else to pronounce on the truth of what really comes from common sense.

Many of the opinions that Reid rejected as contrary to common sense do not appear to be in conflict with the necessities of action he held common sense to impose. Thus, he attacked Berkeley as having denied the existence of a material world, but Berkeley denied that the truth of his opinion would make any changes in our experience; stones, for instance, would remain the solid objects we find them to be. Reid's limited success in vindicating the beliefs of common sense by pointing to inconsistencies between the profession and the practice of dissenters was connected with his interpretation of many of these beliefs; they presented themselves to him as containing an element that lies beyond verification by experience and that might therefore be called metaphysical. He construed, for example, our belief in the existence of a mate-

rial world as disallowing any phenomenalistic account of the nature of material things, our belief in personal identity as involving a reference of all our experience to its (immaterial) subject, our belief in the freedom of our will as involving indeterminacy of choice.

REID'S FOLLOWERS. The notion of an appeal to common sense in great matters of philosophical dispute was crudely taken up by two of Reid's contemporaries, James Beattie (the poet) and James Oswald. When they were regarded as its representatives, the school that became associated with Reid's name could easily be spoken of as appealing to "the judgment of the crowd." Dugald Stewart (1753–1828), teaching and writing with Reid's moderation, though without his penetrating simplicity, consolidated the school's position in Scotland, and his books helped to make the influence of the ideas he shared with Reid strongly felt in France and America.

SIR WILLIAM HAMILTON. Sir William Hamilton (1788–1856) produced a philosophy in which doctrines of Reid and Immanuel Kant were fused into an unstable compound. It proclaimed the sovereignty of common sense and compromised its deliverances, which for Reid were necessarily objective, with an ambiguous assertion of the "relativity" of knowledge. According to Hamilton, the convictions of common sense come to us with the backing of our entire cognitive nature. They are tests of other truth; their own must be presumed, for they are too elementary to have antecedents from which they could be derived. The only possible falsification of common sense would be demonstrated inconsistency in its deliverances, and this would bring in epistemological chaos. J. S. Mill's *An Examination of Sir William Hamilton's Philosophy* (1865) gave a reactionary, obscurantist look to the authority that Reid and Hamilton claimed for common sense. The "psychological" method, which Mill opposed to their "introspective method," was damagingly designed to show how a belief—such as everyone's belief in an external world—had grown up, taking on in the process the appearance of obviousness; the psychological method would undermine the doctrine that a belief is a dictate of nature by exhibiting its natural history.

CRITICAL COMMON SENSE

Reid and Hamilton both thought that criticism is or may be necessary in order to determine whether a belief is in fact a belief of common sense. They also held, however, that once this fact is established, it follows that the belief is true. The label "critical common sense" might be used, not too misleadingly, to distinguish from this position those philosophical views which combine the greatest respect for common sense with the insistence or admission that at least some of its beliefs are open to critical revision.

ARISTOTLE. If common sense is identified with what is commonly believed and its criticism is thought of as designed to elicit and defend the truth in common beliefs, then Aristotle may be called the first common-sense philosopher. "We must," Aristotle said, "as in all other cases, set the observed facts before us and, after first discussing the difficulties, go on to prove, if possible, the truth of all the common opinions about these affections of the mind, or, failing this, of the greater number and the most authoritative; for if we both refute the objections and leave the common opinions undisturbed, we shall have proved the case sufficiently" (*Nicomachean Ethics*, 1145b2–7; cf. 1172b35–1173a2, *Eudemian Ethics* [attributed to Aristotle], 1216b26–35).

C. S. PEIRCE. The "Critical Common-sensism" argued for by the American philosopher Charles Sanders Peirce (1839–1914) was largely defined in relation to the views held by the Scottish school. It saw the beliefs of common sense, Peirce said, as changeless, the same for all men at all times. It rightly thought of them as having a kind of instinctive character—but instincts can undergo modification. Peirce was sure that these beliefs show some modification as people become civilized and civilization develops. They are not, as ordinarily held, beliefs that have been up for acceptance or rejection; they exist as lifelong "belief-habits." And they possess a logical feature in virtue of which they are doubt-resistant when criticized: They have an essential vagueness. Peirce illustrated this with "our belief in the Order of Nature." Let an attempt be made to give this belief precision, and what results will be found disputable. "But who can think that there is *no* order in nature?" (*Collected Papers*, Vol. V, p. 359).

The "Critical Common-sensist," Peirce said, tries to "bring all his very general first premisses to recognition" and to develop "every suspicion of doubt of their truth" (ibid., p. 363). But the doubt he is looking for must be the real thing, not "paper" doubt; we can no more induce genuine doubt by an act of will than we can give ourselves a surprise by deciding to. "Strong thinkers" are "apt to be great breath-holders," but holding one's breath against belief is not doubting. In claiming "indubitability" for a belief of common sense, Peirce was not declaring its truth—"propositions that really are indubitable, for the time being" may "nevertheless be false" (ibid., p. 347). The future holds possibilities of surprise for all our beliefs. Yet

Peirce seems to have held that though any one of our indubitable beliefs might turn out to be false, they could not all do so.

HENRY SIDGWICK. "Common sense organised into Science," Henry Sidgwick (1838–1900) remarked, "continually at once corrects and confirms crude Common Sense" (*Lectures on the Philosophy of Kant*, p. 425). Sidgwick saw common sense as a great mass of ore, rich in valuable metals, that needs philosophical smelting. It must have removed "inadvertencies, confusions, and contradictions" (ibid., p. 428). However, the procedures by which this is done—rigorous reflection, the adjustment of its beliefs to the assured results of science—are not alien to it. Sidgwick's *Methods of Ethics* (1874) contains a detailed examination of the "morality of common sense," directed toward showing its frequent vagueness, its areas of indecision, its compromises between conflicting ideas, and also toward showing how its fundamental convictions can be taken up into a form of utilitarianism that can reasonably claim the acquiescence of common sense.

G. F. STOUT. For G. F. Stout (1860–1944), common sense has been self-correcting in its evolution and it is still to some extent modifiable. The man in the street is not to be taken as its representative; the common sense of philosophical importance resides in the consensus of ignorant and educated belief. This unanimity is the result of a long development, during which idiosyncrasies of opinion have been worn down by mutual attrition, and mistakes—which common sense itself can see to be such—have been corrected. Common sense is less a matter of particular beliefs than "the persistence of plastic tendencies to certain most general and comprehensive views" (*Mind and Matter*, p. 11). These include such strongly metaphysical dispositions as "the tendency to find Mind in Nature generally" (ibid., p. 14). When a conflict arises between common sense and some scientific or philosophical opinion, the final decision, Stout maintained, rests with common sense, "however indirectly"; for common sense must either be provided with reconciliatory explanations or be brought to see that the considerations in favor of the opinion more than cancel the presumption against it.

RUSSELL AND BROAD. It is convenient to mention here two contemporary philosophers who have thought that there are philosophical opinions which can be described as common sense but who have thought that some of these opinions are quite radically mistaken. Science takes common sense as its starting point, Bertrand Russell says; it has arrived at results with regard to the nature of physical things and their relation to perception that are incompatible with parts of the "metaphysic" of common sense. One does what one can for common sense, but, according to C. D. Broad, sometimes not much is possible; nor should a philosopher feel disturbed at a break with common sense that results from seeing together facts that average people notice only separately and from taking into account other facts of which they are altogether ignorant.

COMMON SENSE AND ORDINARY LANGUAGE

G. E. MOORE. G. E. Moore (1873–1958) did not think that common sense never errs. He seems often to have treated universal, or very general, acceptance as the identifying mark of a commonsense belief, and, as he mentions, things that everybody once believed have turned out to be false. He was prepared to allow that, for all he knew to the contrary, there might be many false propositions included within the vague boundaries of "the Common Sense view of the world." Moore had no special interest in critically sifting the beliefs of common sense for truth and falsity. He was primarily interested in its massive certainties.

Moore's paper "A Defence of Common Sense" (1925) lists sets of propositions that are as obviously true as almost any imaginable: for instance (with considerable paraphrase for the sake of brevity), propositions stating that the earth has existed for many years; that its inhabitants have been variously in contact with, or at different distances from, one another and other things; and that these facts are matters of common knowledge. According to Moore, these "truisms," taken together, imply the truth of the commonsense view of the world in certain of its "fundamental features," for they imply that there are material things, space, time, and other minds besides one's own—in a clear meaning of each of the expressions "material thing," "space," "time," and so on. The abstract words contain ambiguities that are absent from, for example, "The earth has existed for many years," but Moore thought that some philosophers who have denied the existence of material things, of space, of time, or of other minds besides their own are to be understood as having expressed views incompatible with such banally obvious truths. He thus regarded them as paradoxically uttering opinions inconsistent with what they themselves know to be true. They constantly reveal this knowledge in its incompatibility with their opinions; a solipsistic

philosopher, for example, sets himself to persuade others that he alone exists.

There is very great doubt, Moore thought, about the correct "analysis," in some important respects, of propositions of common sense that are quite certainly true. (Roughly, for Moore, the analysis of a concept or a proposition lays bare its structure by indicating the concepts it implicitly contains and the way they are combined.) Moore did not think that a phenomenalistic analysis of the concept of a material thing could be ruled out as absolutely impossible. It follows that, in his judgment, a philosopher who was using the sentence "Material things do not exist" simply to word a phenomenalistic doctrine and to repudiate its alternatives would not be repudiating a conviction of common sense that is manifestly true. And if this is so, it is hard to see what a philosopher could have in mind in using the words that would constitute such repudiation. By contrast, denials of the "reality" of space and time on the ground that their concepts are self-contradictory do appear to be in irreconcilable conflict with the most commonplace facts about position and distance, and about past, present, and future.

NORMAN MALCOLM. The philosophical paradoxes that Moore attacked on many different occasions are construed in Norman Malcolm's essay "Moore and Ordinary Language" as disguised, variously motivated rejections of common language, and Moore's defense of common sense is construed as its vindication. A philosopher declares, for instance, "We can never know for certain the truth of any empirical statement." As interpreted by Malcolm, he is saying that it is never right to say "I know for certain" when it is *logically* possible that one is mistaken, that the words are always improperly used in this situation. Moore's reply, characteristically translating from the abstract to the concrete, pointed out the absurdity of anyone's suggesting, when he is sitting on a chair, that he believes he is, that he very probably is, but that he does not know it for certain. What Moore's reply did, on Malcolm's interpretation, was "to appeal to our language-sense," "to make us feel how queer and wrong" it would be to speak here in the way the philosopher proposes and substitute "believe" for "know for certain" or to turn to such words as "probable" ("Moore and Ordinary Language," p. 354).

"A philosophical paradox," Malcolm says (pp. 359–360), "asserts that, whenever a person uses a certain expression, what he says is false." However, from the fact that the expression has a use in ordinary language, it follows, Malcolm argues, that it is free from self-contradiction (since a self-contradictory expression necessarily has no use) and therefore that it *can* be employed to make true statements. And this is enough to refute the paradox. Whether or not people always say something false when using these expressions becomes a matter to be settled by matter-of-fact evidence, and the paradoxical philosopher does not deal in evidence of this sort.

In Malcolm's essay a stronger claim is made in effect for Moore's refutations: They produce indisputably true statements employing the expressions that the paradoxes reject, for they present *paradigms* of the correct application of these expressions. And it is maintained that we could not learn the meaning of some expressions without such paradigms or standard cases; that we could not learn, for example, the meaning of "material thing" without being shown examples of material things, or the meaning of spatial and temporal expressions without acquaintance with spatial and temporal relations, or the meaning of "certain," "probable," "doubtful" without being introduced to the contrasted situations to which they apply. Thus, a statement denying that there is anything answering to one of these expressions must be false. Scrutiny of "the argument from paradigm cases" has been an incident in the recent shift of philosophical interest from common sense (at least under that name) to ordinary language.

LUDWIG WITTGENSTEIN. The way to philosophical paradox is opened, according to Ludwig Wittgenstein (1889–1951), when some feature of ordinary language is misconstrued as only philosophers are likely to misconstrue it. This disorder, along with such other characteristic philosophical aberrations as directionless bafflement, is to be got rid of by bringing words back from their alienation in metaphysical discourse to the familiar surroundings from which they have been abstracted and watching them at work there. Philosophers have not carelessly misunderstood ordinary language; it is waiting for them with "bewitchment" and "illusion." In the emancipation that is achieved when one is able to "command a clear view" of the functioning of language, everything is left, but seen to be, "as it is." Wittgenstein rarely mentioned common sense. He referred in *The Blue Book* (Oxford, 1958, p. 48) to the "common-sense philosopher" (such as Moore or Reid) who, "*n.b.*, is not the common-sense man." The commonsense man, Wittgenstein may be taken to suggest, is man before the philosophical Fall.

See also Paradigm-Case Argument.

Bibliography

THE NOTION OF COMMON SENSE

Berkeley, George. *Three Dialogues between Hylas and Philonous*. London, 1713.

Buffier, Claude. *First Truths*. Translated anonymously. London: J. Johnston, 1780.

Isaacs, Nathan. *The Foundations of Common Sense*. London: Routledge and Paul, 1949.

Lewis, C. S. "Sense." In *Studies in Words*. Cambridge, U.K.: Cambridge University Press, 1960.

Also see the works listed below by Grave, Reid, Stewart, Peirce, Stout (*Mind and Matter*), Malcolm, Moore (*Philosophical Papers*, Ch. 2), and Woozley.

THE "SCOTTISH SCHOOL"

Beattie, James. *An Essay on the Nature and Immutability of Truth*. Edinburgh: A. Kincaid and J. Bell, 1770.

Chastaing, Maxime. "Reid, la philosophie du sens commun." *Revue philosophique de la France et de l'étranger* 144 (1954): 352–399.

Grave, S. A. *The Scottish Philosophy of Common Sense*. Oxford: Clarendon Press, 1960.

Mill, John Stuart. *An Examination of Sir William Hamilton's Philosophy*. London: Longman, Green, Longman, Roberts and Green, 1865. Chs. 3, 8–14.

Reid, Thomas. *Works.* 2 vols, edited by Sir William Hamilton. Edinburgh, 1846–1863. Hamilton's appendices include a long historical and expository dissertation on the notion of common sense.

Stewart, Dugald. *Works.* 2 vols, edited by Sir William Hamilton. Edinburgh, 1854–1860. Vol. III, Ch. 1.

"CRITICAL COMMON SENSE"

Broad, C. D. *The Mind and Its Place in Nature*. New York: Harcourt, Brace, 1925. Chs. 2–4.

Broad, C. D. "A Reply to My Critics." In *The Philosophy of C. D. Broad*, edited by P. A. Schilpp, 803–805. New York: Tudor, 1959.

Peirce, C. S. *Collected Papers.* 6 vols, edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: Belknap Press of Harvard University Press, 1931–1935. Vol. V, pp. 293–313, 346–375.

Russell, Bertrand. "Reply to Criticisms." In *The Philosophy of Bertrand Russell.* 2nd ed., edited by P. A. Schilpp, 700–705. Evanston, IL: Library of Living Philosophers, 1946.

Russell, Bertrand. *Human Knowledge*. London, 1948. Part III; Part IV, Ch. 10.

Sidgwick, Henry. *The Methods of Ethics.* 6th ed. London: Macmillan, 1901.

Sidgwick, Henry. "The Philosophy of Common Sense." In *Lectures on the Philosophy of Kant*, edited by J. Ward, 406–429. London: Macmillan, 1905.

Stout, G. F. *Studies in Philosophy and Psychology*. London: Macmillan, 1930. Ch. 6.

Stout, G. F. *Mind and Matter*. Cambridge, U.K.: Cambridge University Press, 1931.

COMMON SENSE AND ORDINARY LANGUAGE

Campbell, C. A. "Common-sense Propositions and Philosophical Paradoxes." *PAS* 45 (1944–1945): 1–25.

Chappell, V. C. "Malcolm on Moore." *Mind* 70 (279) (July 1961): 417–425.

Chisholm, R. M. "Philosophers and Ordinary Language." *Philosophical Review* 60 (1951): 317–328. Malcolm's reply follows immediately.

Duncan-Jones, A. E., and A. J. Ayer. "Does Philosophy Analyse Common Sense?" *PAS*, Supp. 16 (1937): 139–176. Symposium.

Flew, A. G. N., ed. *Essays in Conceptual Analysis*. London: Macmillan, 1956. Chs. 1 and 6. Discusses paradigm cases.

Malcolm, Norman. "Moore and Ordinary Language." In *The Philosophy of G. E. Moore*, edited by P. A. Schilpp, 345–368. Evanston and Chicago: Northwestern University, 1942.

Malcolm, Norman. "Defending Common Sense." *Philosophical Review* 58 (1949): 201–220.

Moore, G. E. "A Reply to My Critics." In *The Philosophy of G. E. Moore*, edited by P. A. Schilpp, 660–675. Evanston and Chicago: Northwestern University, 1942.

Moore, G. E. *Some Main Problems of Philosophy*. London: Allen and Unwin, 1953. Chs. 1, 5–7, 11.

Moore, G. E. *Philosophical Papers*. London: Allen and Unwin, 1959. Ch. 2, "A Defence of Common Sense," and Ch. 7, "Proof of an External World."

Passmore, J. A. *Philosophical Reasoning*. London: Duckworth, 1961. Ch. 6. Discusses paradigm cases.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Blackwell, 1953.

Woozley, A. D. "Ordinary Language and Common Sense." *Mind* 62 (247) (July 1953): 301–312.

OTHER RECOMMENDED TITLES

Audi, Robert. *The Architecture of Reason: The Structure and Substance of Rationality*. New York: Oxford University Press, 2001.

Barker, S. F., and T. C. Beauchamp, eds. *Thomas Reid: Critical Interpretations*. Philadelphia: Monograph Series, 1976.

Chisholm, Roderick. *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice-Hall, 1966; 2nd ed., 1977; 3rd ed., 1989.

Cornman, James W. *Perception, Common Sense, and Science*. New Haven, CT: Yale University Press, 1975.

Foley, Richard. *Intellectual Trust in Oneself and Others*. New York: Cambridge University Press, 2001.

Goldman, Alan H. "Epistemic Foundationalism and the Replaceability of Ordinary Language." *Journal of Philosophy* 79 (1982): 136–154.

Guyer, Paul. "Kant on Common Sense and Scepticism." *Kantian Review* 7 (2003): 1–37.

Lehrer, Keith. *Thomas Reid*. London: Routledge, 1989.

Lemos, Noah. "Common Sense and 'A Priori' Epistemology." *Monist* 81 (3) (1998): 473–487.

Lycan, William. "Moore against the New Skeptics." *Philosophical Studies* 103 (2001): 35–53.

Moore, G. E. *Philosophical Studies*. London: Routledge and Kegan Paul, 1922.

Moser, Paul K. "A Defense of Epistemic Intuitionism." *Metaphilosophy* 15 (1984): 196–209.

Moser, Paul K. "Epistemology (1900–Present)." In *Routledge History of Philosophy*, Vol. 10: *Philosophy of the English Speaking World in the 20th Century*, edited by John Canfield. London: Routledge, 1996.

Reid, Thomas. *Essays on the Active Powers of the Human Mind,* edited by B. Brody. Cambridge, MA: MIT Press, 1969.

Reid, Thomas. *Essays on the Intellectual Powers of Man,* edited by B. Brody. Cambridge, MA: MIT Press, 1969.

Schneewind, J. B. "Scottish Common Sense Philosophy." In *The Cambridge Dictionary of Philosophy*, edited by Robert Audi. Cambridge, U.K.: Cambridge University Press, 1995; 2nd ed., 1999.

Wolterstorff, Nicholas. "Reid on Common Sense." In *The Cambridge Companion to Thomas Reid*, edited by Terence Cuneo. Cambridge, U.K.: Cambridge University Press, 2004.

S. A. Grave (1967)
Bibliography updated by Benjamin Fiedor (2005)

COMMUNISM

The voluntary disbanding of the communist state of the Soviet Union in 1991 was the practical defeat of a certain theory of communism as the economic, social, and political antithesis and opponent of the liberal democratic capitalist state that first emerged in the developed Western societies. According to Francis Fukuyama (1992), citing Georg Wilhelm Friedrich Hegel's theory of history, the "death of communism" marks the triumph of liberal democratic states as the paramount achievement of human history. Any further opposition to the extension of the liberal democratic model could only come in the form of regressive social movements seeking to avoid the trauma of inevitable change by clinging to ancient dogmas.

Still, as capitalism becomes the unrivaled global economic system, spilling over the bounds of the nation-state, the social and political achievements and perspectives of the liberal democracies are increasingly being jeopardized by the economic logic of capitalism itself. That the economic power of global corporations imposes demands that most nation-states ignore at their peril necessitates a reappraisal of a complacent triumphalism. In historical retrospect and freed from much of the ideological partisanship of the cold war period, it becomes clear that the challenge of communist claims of social egalitarianism and economic efficiency did much to stimulate progressive social and democratic changes in Western societies throughout the nineteenth and twentieth centuries (Hobsbawn 1994). Rather than a choice between two distinct models, it appears that the thesis of capitalism and the antithesis of communism produced in the West an evolving mixture of elements from both ideal models (Lawler 2001).

MARX'S CONCEPTION OF THE STAGES OF COMMUNISM

Indeed, the perspective of communism as an aspect or dimension of the internal evolution of Western society was the view recommended by the foremost exponent of communism, Karl Marx, who argues that the working people "have no ideals to realize, but to set free elements of the new society with which old collapsing bourgeois society itself is pregnant" (1987, p. 355). In criticizing the conception of communism as an ideal to be realized in the future by contrast to the existing and undesirable state of affairs of the present, Marx distinguishes his "dialectical" understanding of communism from that of rival "nihilistic" theories of communism (Lawler 1994).

NIHILISTIC AND DIALECTICAL COMMUNISMS. The most prominent exponent of the nihilistic conception of communism, and Marx's opponent at the time of the *Communist Manifesto* (1848), was the Russian communist Nikolai Bakunin (1814–1876). In his "Appeal to the Slavs" written in 1848 while he was fleeing arrest in Germany, Bakunin writes:

> Look! The Revolution is all around. It alone is powerful. The new spirit with its ability to dissolve has irrevocably penetrated humanity; it is burrowing into and overturning the deepest and darkest layers of European society. And the Revolution will not rest until it has completely destroyed the old dislocated world and created in its place a new and better world. Thus all the vigour and strength, all the certainty of triumph is in it and only in it. In it alone is life; outside it is death. (Pirumova, Itenberg, and Antonov 1990, pp. 85–86)

Marx and his partner, Friedrich Engels, rejected this utopian and nihilistic vision of creating an alternative society, however egalitarian and committed to social justice, out of the destruction of the old world. Evoking the realist historical perspective of Hegel that "[w]hat is rational is actual and what is actual is rational" (1991, p. 20), Marx argues that only by studying the real world and its actual movement is it possible to discern the internal forces and trends that bring about change, development, and transformation. Communism, he then argues, is a real movement that is actually taking place within the present capitalist society.

TEN HOURS BILL. For example, one of the major social events of the first half of the nineteenth century in England was the passage of a series of factory acts, including

the Ten Hours Bill, which limited the workday for women and children to ten hours. Marx describes this modest achievement as of historic significance, "It was the first time that in broad daylight the political economy of the middle class succumbed to the political economy of the working class" (1987, "Inaugural Address of the Working Men's International Association, September 28, 1864"). In Marx's conception the political economy of the middle class, or capitalism, is the pure, unfettered rule of private property and production for the market. Therefore, in limiting the operation of the free market for the sake of the well-being of working people, the factory acts evinced the partial triumph of communism over capitalism taking place within capitalism itself. Other such elements of communism that emerged in the industrial capitalism of the West during the nineteenth and especially the twentieth centuries included free public education, national health care plans (such as, in the United States, Medicare and Medicaid), and national pension or social security plans, as well as laws further limiting the time of the working day and establishing legal conditions for the self-organization of labor through trade unions.

From Marx's perspective the history of Western capitalism presents evidence for the growing emergence within the evolution of capitalism of embryonic elements of an alternative society whose basic characteristics are already discernable, not from the constructions of ideal theory, but from the requirements of actual historical development. A detailed study of Marx's thought on the nature of communism reveals six stages or phases of communist development, beginning with the factory acts and similar infusions of social consciousness into the operation of the capitalist market economy: two phases of communism within capitalism, two phases of the transition between capitalism and communism, and two phases of communism *per se* (Lawler 1998).

DEFINITION OF COMMUNISM. In the *Communist Manifesto*, when Marx projects the final outcome of this evolution, he formally defines communism as "an association, in which the free development of each is the condition for the free development of all" (Marx and Engels 1976, p. 506). The core idea of communism is the all-round freedom of the individual to develop latent abilities without the age-old restrictions that come from the necessities of mere physical survival. Such free development of each is the foundation of an integrally free society. When he further elaborates on this definition in his *Critique of the Gotha Program* (1875), he writes of the highest stage of the evolutionary process, the second phase of communism *per se*:

> In a higher phase of communist society, after the enslaving subordination of the individual to the division of labor, and therewith also the antithesis between mental and physical labor, has vanished; after labor has become not only a means of life but life's prime want; after the productive forces have also increased with the all-round development of the individual, and all the springs of cooperative wealth flow more abundantly—only then can the narrow horizon of bourgeois right be crossed in its entirety and society inscribe on its banners: from each according to his ability, to each according to his needs! (1989, p. 87)

The dramatic final maxim of communism, "from each according to his ability, to each according to his needs," cited out of context as the sum and substance of Marx's conception, appears as an unrealizable, utopian ideal. However, this definition must be comprehended as the outcome of previous stages of historical development. Distribution according to need is only possible at a certain stage or phase of historical evolution when "all the springs of cooperative wealth flow more abundantly." And this abundance of social wealth presupposes both the alienation of labor and the alienation of this alienation—that is, the progressive emergence of creative human labor, labor that has become "not only a means of life but life's prime want" (1989, p. 87). These conditions of a fully developed communism emerge within the previous history of market-oriented society.

LOWER STAGE OF COMMUNISM OR SOCIALISM. For Marx production for the market, although further limited by laws aimed at individual and social well-being, continues well past the communist revolution (initiating the transition between capitalism and communism) and into the lower phase of communism *per se*. In the lower phase of communism, often called socialism, distribution or the individual's income is geared to the quantity and quality of the work that the individual performs. This is the principle of "bourgeois right" that arises out of the requirements of market exchange in which qualitatively different products are equalized by their economic value. Because individuals differ in terms of their needs—for example, one person is single, the other has children to support—the principle of fairness, right, or law according to which each is paid according to work performed results in inequality in real conditions of life.

In this lower phase of communism, however, bourgeois "principle and practice are no longer at logger-

heads" (1989, p. 86), as is the case in capitalism. In capitalism the principle of justice calling for "an honest day's pay for an honest day's work" is systematically violated by the fact that individuals do not receive according to their actual labor, but according to the value of their labor power or ability to work. While it is asserted that workers are generally paid according to the work they perform, their wages in fact tend to reflect merely the value of goods and services needed to reproduce them as workers. The difference between the wage thus determined and the value of the goods actually produced is surplus value, the basis of capitalist profit. Paying workers according to the work they actually perform, the principle of the first phase of communism, overcomes the contradiction in capitalism between abstract principle and real practice. But if bourgeois right is finally realized only in this lower phase of communism, both practical inequity and the alienation of labor nevertheless continue.

ALIENATION OF LABOR. The alienation of labor, first described by Marx in his early *Economic and Philosophic Manuscripts of 1844*, consists in the individual's having to work to live, to subsist (1975). When people work only for the sake physical survival, they are subverting their essential human powers. Labor, for Marx, is the defining feature of human beings, distinguishing "the worst architect from the best of bees" (1996, p. 189)—the ability to creatively transform and channel the forces of nature to achieve distinctively human goals. The capacity for creative activity or labor arises out of the nature of the human being as a species being, that is, as a being who is directly concerned with the species as a whole. It is this connection of the individual with the human species, as epitomized in the use of language, that raises consciousness beyond the animal level of concern for (mostly) individual needs to the level of universality that constitutes reflective thought itself. Thus, while the animal is satisfied when its present hunger (and that of its immediate family) is appeased, the human individual is not content until the threat of hunger is banished in general, in terms of the future of the group and ultimately of the species as a whole.

Hence, when people survive only by selling their labor, working not to express their creative ability but to prolong their biological existence, they are alienating this distinctive feature of their humanity. Creative, essentially human, activity is barely possible where the necessities of survival force individuals to engage in repetitive physical work for up to sixteen hours per day. The distinctive human gift is squandered when children are forced into mindless labor from an early age. So when Marx examines the Ten Hours Bill, he recognizes the essential core of communist humanism: restricting the amount of time individuals are forced to work to survive and thereby freeing them, however minimally, to develop their own creative powers. Hence, one of the essential demands of the *Communist Manifesto* is free education for children and the elimination of child labor.

That much of the political platform set forth in the *Communist Manifesto* has in fact been realized in the course of the later evolution of the Western capitalist societies is therefore evidence, from Marx's point of view, not of the triumph of capitalism, but of the incipient emergence, taking place already within capitalism itself, of what he projects as the outcome of this process, the free development of each of communist society. Only when the prime need of the majority of people is to engage in creative activity is the alienation of labor fully overcome. But the seeds of this development and its embryonic growth begin within capitalism. The historic advances of social democracies face new challenges in the early twenty-first century as an unrivaled capitalism emerges on a global scale beyond the controls of the nation-state. Capitalist economic logic implicitly pits workers of advanced countries against those of newly developing nations without centuries of struggle for the rights of the free development of each. Marx's ringing conclusion to the *Communist Manifesto* has therefore become even more relevant: Working people of all countries, unite! (Marx and Engels 1976, p. 519)

COMMUNISM IN THE HISTORY OF WESTERN PHILOSOPHY

ASIATIC COMMUNISM. Viewed in this way, communism is not an alien social theory inserted abruptly at one juncture into Western philosophy by Marx, and then given a more hospitable reception in non-Western states such as Russia and China. The communism of the Soviet Union and China reflects an altogether different historical dynamic rising out of what Marx called, in his characterization of the socioeconomic structure of this part of the world, the "Asiatic mode of production" (1989, p. 263). In this mode of production the ruler, the tsar of Russia or the emperor of China, centralizes both political and economic power in his own hands. The dynamics of Western capitalism involves, on the contrary, the relative separation of political power from economic evolution—a separation that continues, for Marx, until, with the full development of communism, "the public power will lose its political [i.e., repressive] character" (Marx and Engels 1976, p. 505). From this point of view the "cult of per-

sonality" of Joseph Stalin (1879–1953) in Russia and Mao Zedong (1893–1976) in China, with state centralization and command of the economy, reflects a kind of Asiatic communism, or a communism developing within the Asiatic mode of production, rather than the communism that Marx discerned as emerging within the womb of Western capitalism. Marx's conception that communism and the market coexist and interpenetrate well after the communist revolution, allowing for a distinct phase of "market socialism," diverges sharply from this "Eastern" approach to communism (Lawler 1998).

POSSESSIVE INDIVIDUALISM. If in terms of content Marx's theory of communism is based on a study of Western society, in terms of philosophical form it is the outcome primarily of one of two major streams in early modern Western philosophy (Lawler 2006). One stream regards the individual as a self-interested being, urged on deterministically by desires arising out of nature, environment, and upbringing, and using reason as a means to achieve maximum individual satisfactions and advantages. The "possessive individualism" (Macpherson 1975) of the modern world sets it apart from the ancient Greco-Roman and medieval views of the individual as constituted by birth or nature for various relatively fixed social functions regarded as necessary for the good of the hierarchically ordered social whole. For the self-interested individual of modern times, the good of the social whole is a means to the individual's own well-being. The classical expression of this trend is the *Leviathan* of Thomas Hobbes, for whom the equal restrictions imposed by the laws of the state (i.e., bourgeois right) establish the civil liberties of capitalist society, including "the liberty to buy, and sell, and otherwise contract with one another; to choose their own abode, their own diet, their own trade of life, and institute [instruct] their children as they themselves think fit; and the like" (1952, p. 113).

Adam Smith's *Wealth of Nations* propounds an economic justification of this perspective, in which the social good or wealth of nations is the largely unintended outcome of individualistic endeavors of production for the market. But in contrast to Hobbes's emphasis on the laws of the state, for Smith the economy is the base of the social order and the state and its laws of formally equal freedoms constitute a secondary framework. In his conception of the primacy of the economic base in relation to the political superstructure, Marx continues such economic materialism (Marx writes: "The totality of these relations of production constitutes the economic structure of society, the real foundation, on which arises a legal and political superstructure and to which correspond definite forms of social consciousness" [1989, p. 263]). However, as seen earlier, the "political economy of the working class" or communism enters this picture when the free operation of buying and selling, and production for the market, are restricted by laws directly aimed at promoting the social good. A radically different understanding of the relation between the individual and the community is implied in the emergence of such communist laws.

PLATO'S HIERARCHICAL COMMUNISM. The second line of thought is continuous with the traditional ancient and medieval view that sees the deliberate promotion of the social good as the highest aim of individual flourishing. In the classical formulation of Plato the public good demands communist or communal ownership of property on the part of the ruling guardians of society to prevent them from using their positions of power for private gain. Such public good also requires the perpetuation of what Plato calls a shameful lie, that is, that the souls of individuals are composed of finer or baser metals, from gold and silver for the rulers and their children to brass and iron for the farmers and artisans and their offspring. This is a lie for Plato, because the souls of human beings are not material, and their destinies, evolving over many lifetimes, are ultimately subject to their own choices (1952, book 10, pp. 437–441). Nevertheless, for the peace and order of society it is necessary that

> none of [the rulers] should have any property of his own beyond what is absolutely necessary.… Gold and silver we will tell them that they have from God; the diviner metal is within them, and they have therefore no need of the dross which is current among men, and ought not to pollute the divine by any such earthly admixture. (1952, book 3, 341)

The modern proponents of the social good are not ashamed to propagate openly their Platonic spiritualism. However, like their counterparts in the stream of possessive individualism, they reject the fixed hierarchies of the past and adopt the standpoint of free and equal individuals. But such equality they ground on the freedom of consciousness or spirit.

DESCARTES'S EGALITARIAN COMMUNISM. Modern egalitarian communism replaces ancient hierarchical communism by stressing the primacy of the free, self-conscious individual, whose awareness that "I think" is, for René Descartes, the foundation of modern scientific method. The self-conscious individual in the modern idealist or spiritualist tradition achieves full self-

development only by working directly for the good of others, the good of society as a whole, in such a way that each associates with the other cooperatively in rewarding activities of mutual and common endeavor. Thus, for Descartes the highest activity for the individual is the pursuit of scientific truth, and the motive of this activity is the practical application of scientific knowledge for the well-being of all. Descartes continues the medieval view of the primacy of the social whole, but abandons its aristocratic foundations in a hierarchy of social classes. Nothing is so equally distributed as reason or good sense (Descartes 1952, p. 69), and this common reason is the foundation of all science and the quest for those truths that will liberate humankind from the immense suffering that is due to ignorance and error. Each individual is capable of joining in the step-by-step ascent to truth that science elaborates as it progressively gains access to the laws of the natural and human orders. Thus, recognizing the limitations of his own individual accomplishments, Descartes begs

> all well-inclined persons to proceed further by contributing, each one according to his own inclination and ability, to the experiments which must be made, and then to communicate to the public all the things which they might discover, in order that the last should commence where the preceding had left off; and thus, by joining together the lives and labours of many, we should collectively proceed much further than any one in particular could succeed in doing. (p. 69)

METAPHYSICAL BASIS OF COMMUNISM: MATERIALISM OR SPIRITUALISM? Descartes's metaphysical conception of the human individual as a spiritual being occupying a physical body contrasts with Hobbes's materialist view of the human being as primarily a physical being capable somehow of mental *phantasmata*. Both founders of modern philosophy appeal to the requirements of modern science. But whereas Hobbes regards the new physics of Galileo Galilei as demanding a starting point in the inertial straight-line motion of externally moved matter, Descartes sees the ultimate foundation of science in thinking itself, in the self-conscious "I" that is free to depart from the illusions of sensory perception so as to reconstruct a true picture of the world according to a step-by-step method of thought. While the possessive individualism of Adam Smith's justification of the free market correlates with this first metaphysical option, a fundamentally social orientation, anticipating Marx's theory of species being, follows from the second. It may seem paradoxical to locate Marx's philosophical ancestry in the spiritualism of Plato and Descartes rather than the materialism of Hobbes, but Marx's materialism is a dialectical materialism that is opposed to the mechanistic materialism of Hobbes that was also influential for Smith.

Descartes's metaphysical hierarchy of spirit or consciousness over matter and the body is expressed in practical, ethical, and social requirements. The pursuit of objects that diminish when they are shared with others should be subordinated to the pursuit of objects that are not so diminished. External material wealth diminishes when shared with others, and so one tends to separate oneself from others when one pursues them. However, because one recognizes the good in others, one should freely focus one's mind on the pursuit of those goods that do not diminish when shared, such as knowledge, health, and virtue. So, in a manner reminiscent of Plato's communism, Descartes establishes the community of shared goods as taking precedence over the pursuit of material wealth:

> But I distinguish between those of our goods which can be lessened through others possessing the like, and those which cannot be so lessened. … But virtue, knowledge, health, and in general all other goods considered in themselves without regard to glory are not in any way lessened in us through being found in many others; and so we have no grounds for being distressed because they are shared by others. (1991, pp. 321–322)

When one shares one's ideas with others, one loses nothing, but enriches both oneself and others. It is a win-win situation. When, however, one pursues limited material goods, then what one person gains the other loses. Therefore, scientific philosophy prescribes a social ethics in which the pursuit of goods of the first type has precedence over pursuit of goods of the second type. So Descartes prescribes the basic maxim of a reasonable and good society: from each according to ability, in cooperation with others, and for the good of all. Therefore, a good society is one in which the creative development of each individual is freely associated with that of other individuals, and working together they promote the full development of society as a whole. But that is just Marx's definition of communism in the *Communist Manifesto*.

LEIBNIZ'S REPUBLIC OF SPIRITS. Similarly, Gottfried Wilhelm Leibniz's monadic human "spirits" achieve their highest development when they are aware of their harmony with one another and actively promote that

harmony. So Leibniz writes that "[s]pirits are of all substances the most capable of perfection and their perfections are different in this that they interfere with one another the least, or rather they aid one another the most, for only the most virtuous can be the most perfect friends" (1951a, p. 342). The outcome of such universal social friendship Leibniz calls, variously, the moral world, the city of God, the republic of spirits (p. 343), and "the kingdom of final causes" (1951b, p. 132).

ROUSSEAU'S SOCIAL CONTRACT OF THE POOR. Reflecting this latter term of Leibniz in his *Groundwork of the Metaphysics of Morals*, Immanuel Kant calls the pursuit of the "kingdom of ends" the culminating formulation of the categorical imperative. It was not primarily to Leibniz that Kant turned for his moral theory, however, but to Jean-Jacques Rousseau, whom Kant regarded as the Isaac Newton of moral science (Meld Shell 1996, 81–82). Rousseau heightens the critique of the philosophy of individual self-interest with his analysis of the Hobbsean social contract as a deceptive strategy on the part of the rich to mobilize the poor in defense of their property, for the meager concession of gaining formal political rights. He describes with sarcasm the real essence of this social contract of the rich, "You need me, for I am rich and you are poor. Let us come to an agreement between ourselves. I will permit you to have the honor of serving me, provided you give me what little you have for the trouble I will be taking to command you" (1976, p. 186). Rousseau would have appreciated the sardonic remark of the communist writer Anatole France that "[t]he law, in its majestic equality, forbids the rich as well as the poor to sleep under bridges, to beg in the streets, and to steal bread" (1894/1992, p. 550).

If the materialist philosophy of self-interest underlies this deception, Rousseau finds an alternative basis of community in the heart-felt promptings of the human soul, which the wise educator of *Émile* must nourish by turning the sympathies of youth toward the common human being, those poor and oppressed victims of the fraudulent social contract. Thus, the tutor of Émile counsels:

> To excite and nourish this nascent sensitivity, to guide it or follow it in its natural inclination, what is there to do other than to offer the young man objects on which the expansive force of his heart can act—objects which swell the heart, which extend it to other beings, which make it find itself everywhere outside of itself—and carefully to keep away those which contract and concentrate the heart and tighten the spring of the human *I*? (Rousseau 1979, pp. 222–223)

On such a basis an authentic social contract can be established in which what is emphasized is not the equality of formal legal and political rights (bourgeois right) but relative equality of the conditions of existence. Against the contracted "I" of the philosophy of self-interest, Rousseau emphasizes the expansive "I" that identifies with "the general will." What distinguishes the general will from the particular will is

> not so much the number of votes as the common interest that unites [the citizens], for in this institution each person necessarily submits himself to the conditions he imposes on others. … And asking how far the respective rights of the sovereign and the citizens extend is asking how far the latter can commit themselves to one another, each to all and all to each. (1976, p. 34)

What is crucial is the prevention of the extremes of wealth and poverty, not a mathematical or formal equality, and the means for doing this involve manifold rectifications of the existent inequalities of conditions of life, involving the use of a progressive income tax and universal public education. A society based on the principle of "each to all and all to each" is just Marx's definition of communism in the *Communist Manifesto*.

KANT'S KINGDOM OF ENDS. Kant takes up Rousseau's general will in his formulations of the categorical imperative, culminating in the conception of a "kingdom of ends," according to which one can "abstract from the personal differences between rational beings, and also from the content of their private ends—to conceive a whole of all ends in systematic conjunction" (1956, pp. 100–101). Making it clear that he does not primarily have in mind the establishment of formal legal and political rights, Kant stresses economic relations of production and exchange of goods as an integral part of such systematic conjunction of ends or goals united under the moral consciousness. The kingdom of ends formulation of the categorical imperative asserts a systematic hierarchy of ends as follows:

> What is relative to universal human inclinations and needs has a *market price*; what, even without presupposing a need, accords with a certain taste—that is, with satisfaction in the mere purposeless play of our mental powers—has a *fancy price*; but that which constitutes the sole condition under which anything can be an end in itself has not merely a relative value—that is, a

price—but has an intrinsic value—that is, dignity (p. 102).

Universal respect for the dignity of the human being establishes a community based on common humanity that economic goals must not violate and to which they should be subordinated. Kant repudiates Adam Smith's idea that if everyone pursues their individual interests, the good of all, defined in terms of quantity of goods, will take care of itself. Smith is also far from Descartes's cooperative search for truth, Leibniz's republic of spirits who "aid one another the most," or Rousseau's heart-based community whose maxim is "each to all and all to each," when he writes of the principle of the modern economy:

> It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages. Nobody but a beggar chooses to depend chiefly upon the benevolence of his fellow-citizens (1952, p. 7).

In another formulation of the moral society, Kant, referring with Leibniz to the Gospels, calls the goal toward which all morality ultimately points "the highest good (the Kingdom of God)" (1993, p. 135). Leibniz and Kant interpret the Gospels as promoting a *this-worldly* kingdom based on spiritual truth, as Jesus said, "The Kingdom of God is within you" (Luke 17:21). The highest good is a unity of virtue and happiness in which happiness is "in exact proportion to morality" (Kant 1993, p. 117). That is, it is a society in which people who perform their moral duties are happy—meaning, that they have their legitimate needs and wants satisfied. Marx merely reformulates this principle for the highest stage of communism: from each according to ability; to each according to need. That is, with the realization of a society whose governing principle is the highest good, people will perform their duties as creative individuals, working in accord with the good of all, and their needs and wants will be satisfied, from the goods and services provided by society, independently of any strict measurement of their contributions. People who contribute less, materially speaking, but still perform their duty according to their ability, are able to satisfy their particular needs just as freely as those who contribute more. They do not, however, work for the sake of satisfying their needs—which for Kant constitutes heteronomy and for Marx is the general characteristic of the alienation of labor.

IDEAL OF THE HIGHEST GOOD: A FANTASY OR AN EMERGING REALITY? The problem with this ultimate goal of morality, Kant says, is that it seems unrealizable in the real world that one observes around oneself, that is, the world that is governed by the laws enunciated by Adam Smith and that Marx calls the "political economy of the middle class." In this empirical reality the satisfaction of needs is not based on the performance of moral duty, but on market-based laws of supply and demand that can bring misery and death to whole portions of the population as a result of changes in fashion and fad. Writing about the same time as Adam Smith, and well before the Ten Hours Bill of the next century, Kant sees no clear expressions of a countervailing "communist" tendency in the real world capable of counteracting the actual operation of the economy based on self-interest. But unless the moral principle is capable of being realized, he says, it must be "fantastic, directed to empty imaginary ends, and consequently inherently false" (Kant 1993, p. 120). The whole of Kant's moral theory as he understands it thus hangs on the empirical possibility of its being realizable. The apparent contradiction between moral ideal and empirical reality constitutes what Kant calls "the antinomy of practical reason" (pp. 199–126).

Marx again reformulates Kant when he rejects the pursuit of communism as an abstract ideal raised against the real world, as well as Hegel, who opposes the "empty ideal" of a better society and insists that "[w]hat is rational is actual and what is actual is rational" (Hegel 1991, p. 20). Kant's own solution to the problem hinges primarily on the recognition that history does in fact move in the direction of a society based on the moral ideal (Van der Linden 1988). But to justify this conception he must establish the validity of a teleological view of history. His third *Critique of Judgment*, as well as many of his historical essays, argues for this perspective. In this way Kant paves the way for the historical approach of Hegel, who sees all of human history as the expression of the dynamic of "spirit," which he defines as "'I' that is 'We,' and 'We' that is 'I'" (Hegel 1977, p. 110). Therefore, what is both actual and rational in the course of history, according to Hegel, is what Marx later calls communism (MacGregor 1984).

See also Bakunin, Mikhail Aleksandrovich; Descartes, René; Dialectical Materialism; Engels, Friedrich; Galileo Galilei; Hegel, Georg Wilhelm Friedrich; Historical Materialism; Hobbes, Thomas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Marxist Philosophy; Marx, Karl; Materialism; Newton, Isaac; Nihilism; Plato;

Rousseau, Jean-Jacques; Smith, Adam; Social Contract; Socialism.

Bibliography

Descartes, René. *Great Books of the Western World. Vol. 31, Discourse on Method*, edited by Robert Maynard Hutchins. Chicago: Encyclopaedia Britannica, 1952.

Descartes, René. "To Chanut, 6 June 1647." In *The Philosophical Writings of Descartes. Vol. 3, The Correspondence*. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. New York: Cambridge University Press, 1991.

France, Anatole. *The Red Lily* (1894). In *Bartlett's Familiar Quotations*, edited by John Bartlett. 16th ed. Boston: Little, Brown, 1992.

Fukuyama, Francis. *The End of History and the Last Man*. New York: Free Press, 1992.

Hegel, Georg Wilhelm Friedrich. *Elements of the Philosophy of Right*. Translated by H. B. Nisbet; edited by Allen W. Wood. New York: Cambridge University Press, 1991.

Hegel, Georg Wilhelm Friedrich. *The Phenomenology of Spirit*. Translated by A. V. Miller. Oxford, U.K.: Clarendon Press, 1977.

Hobbes, Thomas. *Great Books of the Western World. Vol. 23, Leviathan*, Part 2. Edited by Robert Maynard Hutchins. Chicago: Encyclopaedia Britannica, 1952.

Hobsbawm, Eric. *Age of Extremes: The Short Twentieth Century, 1914–1991*. New York: Viking Penguin, 1994.

Kant, Immanuel. *Critique of Practical Reason*. 3rd ed., edited and translated by Lewis White Beck. New York: Macmillan, 1993.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. New York: Harper and Row, 1956.

Lawler, James. Forward to *Socially Mixed Economies: How Social Gains Develop in Opposed Systems*, by John Weber. Lanham, MD: Lexington Books, 2001.

Lawler, James. "Marx as Market Socialist." In *Market Socialism: The Debate among Socialists*, edited by Bertell Ollman, 23–52. New York: Routledge, 1998.

Lawler, James. "Marx's Theory of Socialisms: Nihilistic and Dialectical." In *Debating Marx*, edited by Louis Patsouras, 173–209. Lewiston, NY: EmText, 1994.

Lawler, James. *Matter and Spirit: The Battle of Metaphysics in Early Modern Philosophy before Kant*. Rochester, NY: University of Rochester Press, 2006.

Leibniz, Gottfried Wilhelm. "Essay on Dynamics." In *Selections*, edited by Philip P. Wiener. New York: Scribner, 1951b

Leibniz, Gottfried Wilhelm. *Selections. Vol. 36, Discourse on Metaphysics*, edited by Philip P. Wiener. New York: Scribner, 1951a.

MacGregor, David. *The Communist Ideal in Hegel and Marx*. Toronto, Canada: University of Toronto Press, 1984.

Macpherson, C. B. *The Political Theory of Possessive Individualism: Hobbes to Locke*. New York: Oxford University Press, 1975.

Marx, Karl. *Karl Marx, Frederick Engels: Collected Works. Vol. 3, Economic and Philosophic Manuscripts of 1844*. Translated by Richard Dixon et al. New York: International Publishers, 1975.

Marx, Karl, and Frederick Engels. *Karl Marx, Frederick Engels: Collected Works. Vol. 6, The Manifesto of the Communist Party*. Translated by Richard Dixon et al. New York: International Publishers, 1976.

Marx, Karl. *Karl Marx, Frederick Engels: Collected Works. Vol. 22, July 1870–October 1871*. Translated by Richard Dixon et al. New York: International Publishers, 1987.

Marx, Karl. *Karl Marx, Frederick Engels: Collected Works. Vol. 24, Critique of the Gotha Program*. Translated by Richard Dixon et al. New York: International Publishers, 1989.

Marx, Karl. *Karl Marx, Frederick Engels: Collected Works. Vol. 29, Contribution to the Critique of Political Economy*. Translated by Richard Dixon et al. New York: International Publishers, 1989.

Marx, Karl. *Karl Marx, Frederick Engels: Collected Works. Vol. 35, Capital*, Vol. 1. Translated by Richard Dixon et al. New York: International Publishers, 1996.

Meld Shell, Susan. *The Embodiment of Reason: Kant on Spirit, Generation, and Community*. Chicago: University of Chicago Press, 1996.

Pirumova, N., B. Itenberg, and V. Antonov. *Russia and the West: Nineteenth Century*. Translated by Vitaly Baskakov, Patricia Beryozkina, and Maureen Ryley. Moscow: Progress Publishers, 1990.

Plato. *Great Books of the Western World. Vol. 7, Republic*, edited by Robert Maynard Hutchins. Chicago: Encyclopaedia Britannica, 1952.

Rousseau, Jean-Jacques. *Discourse on the Origin of Inequality*. In *On the Social Contract, Discourse on the Origin of Inequality, Discourse on Political Economy*. Translated and edited by Donald A. Cress. Indianapolis, IN: Hackett, 1983.

Rousseau, Jean-Jacques. *Émile: or, On Education*, edited by Allan Bloom. New York: Basic Books, 1979.

Smith, Adam. *Great Books of the Western World. Vol. 39, The Wealth of Nations*, edited by Robert Maynard Hutchins. Chicago: Encyclopaedia Britannica, 1952.

Van der Linden, Harry. *Kantian Ethics and Socialism*. Indianapolis, IN: Hackett, 1988.

James M. Lawler (2005)

COMMUNITARIANISM

In the 1980s communitarians displaced Marxists as the most prominent critics of liberal political theory. Communitarians share a belief that liberalism is excessively individualistic or atomistic, ignoring people's dependence on communal relationships. They differ in where they locate this flaw. Some criticize the liberal ideal of freedom of choice, arguing that people's ends in life are defined by their communal ties, not freely chosen (Sandel 1984). Others accept the ideal of freedom of choice, but criticize liberalism for ignoring its social and cultural preconditions (Taylor 1989). Still, others argue that moral reasoning is dependent on communal traditions, so that

liberal claims to universal validity are illegitimate (Walzer 1983, MacIntyre 1981).

Commentators sometimes distinguish between backward-looking and forward-looking versions of communitarianism (Phillips 1993). The former asserts that healthy communal bonds existed in the past, lament the decline of community as a result of the increasing emphasis on individual choice and diverse ways of life (the "permissive society"), and seek to retrieve a conception of the common good. This sort of communitarianism is difficult to distinguish from traditional conservatism and is widely criticized for ignoring the ways that most communities historically excluded women, gays, or racial and religious minorities (Frazer 1999). By contrast, forward-looking communitarians disavow nostalgia for the past, accept that individual choice and cultural diversity are now permanent features of modern life, and acknowledge that earlier forms of community were too narrow and exclusive to be retrievable today. Hence, they seek to build new bonds of community that integrate diverse groups and lifestyles, for example, by promoting forms of patriotism, democratic citizenship, or civil society that encourage people from different backgrounds to work together. A more complex version of communitarianism is backward-looking at the local level, allowing ethnic or religious communities to uphold a traditional way of life even if it requires restricting individual freedom, while adopting a forward-looking model at the national level, where the multiplicity of different groups in society must cooperate.

In response to the communitarian critique many liberals attempt to show that they, too, are sensitive to the importance of community and culture and that they can accommodate at least the forward-looking dimensions of communitarianism. Hence, a proliferation of theories of liberal republicanism, liberal patriotism, liberal multiculturalism, and liberal civil society have been witnessed. All these are intended to show that a liberal society is not exclusively individualistic and can accommodate and support a rich array of collective identities and associations, without compromising the basic liberal commitment to the protection of individual civil and political rights.

Given these developments, the original liberal-communitarian debate of the 1980s has given way to a number of new, more differentiated positions and issues. Instead of a stark choice between individualism and communitarianism, one now faces a range of debates about how to sustain bonds of moral solidarity and political community in an era of individual rights and cultural diversity: How to build a common national identity without suppressing ethnic and religious diversity? How to nurture feelings of trust and solidarity in mass societies where people share little in common? How to foster a vibrant public sphere that encourages civic participation and democratic dialogue? How to support family life without imposing traditional gender roles? How to educate children to be public-spirited citizens without inculcating a narrow chauvinism? Communitarianism does not provide a single perspective or framework for answering these questions, and there is a growing sense that the communitarian label obscures as much as it reveals about someone's position on them. Indeed, virtually all the major writers associated with the original communitarian critique express reservations about the label. Nonetheless, these are all questions that have been put on the agenda of political philosophy by the communitarian critique of liberalism. Communitarianism may be fading as a recognizable school of political philosophy, but communitarian concerns have come to dominate political philosophy at the start of the twenty-first century.

See also Liberalism.

Bibliography

Avineri, Shlomo, and Avner de-Shalit, ed. *Communitarianism and Individualism*. New York: Oxford University Press, 1992.

Frazer, Elizabeth. *The Problems of Communitarian Politics: Unity and Conflict*. New York: Oxford University Press, 1999.

Lehman, Edward W., ed. *Autonomy and Order: A Communitarian Anthology*. Lanham, MD: Rowman & Littlefield, 2000.

MacIntyre, Alisdair. *After Virtue: A Study in Moral Theory*. London: Duckwork, 1981.

Mulhall, Stephen, and Adam Swift. *Liberals and Communitarians*. 2nd ed. Oxford, U.K.: Blackwell, 1996.

Phillips, Derek L. *Looking Backward: A Critical Appraisal of Communitarian Thought*. Princeton, NJ: Princeton University Press, 1993.

Sandel, Michael J., ed. *Liberalism and Its Critics*. New York: New York University Press, 1984.

Taylor, Charles. "Cross-Purposes: The Liberal-Communitarian Debate." In *Liberalism and the Moral Life*, edited by Nancy L. Rosenblum, 64–87. Cambridge, MA: Harvard University Press, 1989.

Walzer, Michael. *Spheres of Justice: A Defence of Pluralism and Equality*. Oxford: Blackwell, 1983.

Will Kymlicka (1996, 2005)

COMPOSITIONALITY

The principle of compositionality is the claim that the meaning of a complex expression is determined by its structure and the meanings of its constituents. Normally the thesis is taken to be about some particular language; questions of structure and constituency are then settled by the syntax of that language. By extension, we can talk about compositionality in other representational systems—thoughts, traffic signs, musical notation, and so on—as long as they have their own syntax.

VARIETIES OF COMPOSITIONALITY

The principle is not committed to a specific conception of syntax and semantics, which is why it can be employed in debates between proponents of different conceptions (see, by way of comparison, Partee 1984). Still, if we reject all constraints on either structure or meaning, compositionality becomes trivial. As T. M. V. Janssen (1986) has shown, we can turn any meaning function on a recursively enumerable set of expressions into a compositional one, as long as we can replace the syntactic operations with different ones. And as W. Zadrozny (1994) has shown, we can turn an arbitrary meaning function into a compositional one, as long as we replace the old meanings with new ones from which they are uniformly recoverable. But because the task of semantics is to identify a meaning assignment that respects both what our best syntax tells us about structure and what our best intuitions tell us about synonymy, these results do not show compositionality to be empirically empty (compare Kazmi and Pelletier 1998; Westerståhl 1998; Dever 1999).

Although hardly trivial, the principle as stated is rather weak. For example, consider a view, according to which the meaning of a declarative sentence *s* is the set of possible worlds where *s* is true. According to such a view, tautologies are synonymous, even though (because Rudolf presumably has some tautological beliefs and lacks others) sentences resulting from embedding tautologies in the context of "Rudolf believes that … " are not. Intuitively, this is a violation of compositionality (compare Carnap 1947, sec. 14). Still, the semantics is *not* in conflict with compositionality as stated, because tautologies might differ structurally or in the meaning of their constituents, which could explain how embedding them may yield nonsynonymous sentences. The strengthening of the principle that *is* incompatible with this view requires that the meaning of a complex expression be determined by the meanings of its *immediate* constituents and the syntactic way these constituents are combined. (*e* is an immediate constituent of *e*' if *e* is a constituent of *e*' and *e*' has no constituent of which *e* is a constituent.) Call the strengthened principle *local compositionality* and the original one *global compositionality*.

Compositionality rules out the existence of a pair of nonsynonymous complex expressions built up from synonymous constituents through identical syntactic operations within the same language. As the principle is usually construed, it says nothing about the possibility of such pair of complex expressions existing in *distinct* languages. Still, intuitively, if the Estonian sentence s_1 and the Aramaic sentence s_2 mean different things despite having identical syntactic structure and pairwise synonymous constituents, we should conclude that either Estonian or Aramaic is not compositional. (The same structure and the same meanings of constituents cannot *determine* more than one meaning.) If we want our principle of compositionality to yield this result, we need to strengthen it: we could demand, for example, that there be a single function for all possible human languages that maps the structure of a complex expression and the meanings of its constituents to the meaning of that complex expression (compare Szabó 2000, p. 500). Call this principle *cross-linguistic compositionality* and the original one *language-bound compositionality*.

So, there are at least four versions of the principle of compositionality: language-bound global, language-bound local, cross-linguistic global, and cross-linguistic local. The first is the weakest and it corresponds to how the principle is officially announced; the last is the strongest and it better captures what is typically taken for granted.

There are three well-known claims that are also occasionally referred to as compositionality principles. The first is the *building principle*, which states that the meaning of a complex expression is built up from the meanings of its constituents. This is a fairly strong claim, at least if we take the building metaphor seriously. For then the meanings of complex expressions must themselves be complex entities whose structure mirrors that of the sentence (compare Frege 1984 [1892], p. 193; Frege 1979 [1919], p. 255). The second is the *rule-to-rule principle*, according to which every syntactic rule corresponds to a semantic one that assigns meanings to the output of the syntactic rule on the basis of the meanings of its inputs. If we assume that an arbitrary function deserves to be called a rule, this is equivalent to language-bound local compositionality. The third is the *principle of substitutivity*, according to which if two expressions have the same meaning, then substitution of one for the other in a third

expression does not change the meaning of the third expression. Assuming that the semantics is *Husserlian*—that substitution of synonyms at a single syntactic position within a larger expression never changes the meaningfulness of the larger expression (compare Husserl 1913, p. 318)—this is also equivalent to language-bound local compositionality. (For the equivalence results, see Hodges 2001, theorem 4. If we want to insist—plausibly—that semantic rules must be at least computable, the rule-to-rule principle is stronger than language-bound local compositionality. The assumption that the semantics is Husserlian is far from trivial—it entails, for example, that because "Jacques is likely to leave" is meaningful and "Jacques is probable to leave" is not; "likely" and "probable" are not synonyms [compare Gazdar et al. 1985, p. 32].)

FORMAL EXPRESSION

Since R. Montague (1974), it has been customary to capture compositionality formally as the existence of a homomorphism between a syntactic and a semantic algebra. Let the syntactic algebra be a partial algebra $\mathbf{E}=\langle E, (F\gamma)_{\gamma\in\Gamma}\rangle$, where E is the set of (simple and complex) expressions and every $F\gamma$ is a syntactic operation on E with a fixed arity. Let m be a meaning assignment function from E to M, the set of meanings. Let F be a k-ary syntactic operation on E; then m is F-compositional if there is a k-ary partial function G on M such that whenever $F(e_1, \dots, e_k)$ is defined,

$$m(F(e_1, \dots, e_k))=G(m(e_1), \dots, m(e_k)).$$

Finally, let m be compositional just in case m is F compositional for every operation of the syntactic algebra. Whenever m is compositional, it induces the semantic algebra $\mathbf{M}=\langle M, (G\gamma)_{\gamma\in\Gamma}\rangle$ on M and it is a homomorphism between **E** and **M** (compare Westerståhl 1998). (For details and formal results, see Janssen 1986, 1997; Hodges 2001.) As stated, this captures language-bound local compositionality.

ARGUMENTS FOR COMPOSITIONALITY

The argument most frequently used to support the compositionality of natural languages is the *argument from productivity*. It goes back (at least) to Frege, who claimed that "the possibility of our understanding sentences which we have never heard before rests evidently on this, that we can construct the sense of a sentence out of parts that correspond to words" (Frege 1980 [1914?], p. 79). The argument is an inference to the best explanation, which can be expanded and rephrased without assuming that meanings are Fregean senses as follows. Because speakers of a language can understand a complex expression *e* that they have not previously encountered, it must be that they (perhaps tacitly) know something on the basis of which they can figure out, without any additional information, what *e* means. If this is so, something they already know must determine what *e* means. But this knowledge cannot plausibly be in general anything but knowledge of the structure of *e* and knowledge of the meanings of the primitive constituents of *e*.

If successful, the argument from productivity establishes global language-bound compositionality. To show that a language is locally and/or cross-linguistically compositional requires detailed empirical investigation. As an argument for global language-bound compositionality, it can be criticized on the ground that although we clearly do understand some complex expressions we have never heard before, it is not self-evident that we could *in principle* understand *all* complex expressions in this manner. In fact, it is hard to see how the sort of general considerations mentioned by the argument from productivity could rule out the existence of isolated *exceptions* to compositionality. (Isolated putative exceptions are often declared to be idioms. Criteria for being an idiom are controversial [compare Nunberg, Sag, and Wasow 1994].)

Besides productivity, two other features of our language comprehension are cited in support of compositionality. One is *unboundedness*: although we are finite beings, we have the capacity to understand each of an infinitely large set of complex expressions. (An example from Platts 1979, p. 47: "The horse behind Pegasus is bald," "The horse behind the horse behind Pegasus is bald," "The horse behind the horse behind the horse behind Pegasus is bald," and so on.) 47. From unboundedeness, productivity follows (assuming that finite beings cannot encounter infinitely many expressions), and thus this is not really an independent consideration. The other feature of language comprehension that supports compositionality is *systematicity*: that there are definite and predictable patterns among the sentences we understand. (For example, anyone who understands "The rug is under the chair" can understand "The chair is under the rug" and vice versa.) Because productivity does not follow from systematicity the argument from systematicity provides independent support for compositionality.

In fact, systematicity supports a stronger principle. The standard explanation for why understanding "black dog" and "white cat" is sufficient for understanding "black cat" and "white dog" is that we can *decompose* the

meanings of complex expressions into the meanings of their constituents and then *compose* these into meanings of other complex expressions. The best explanation for the possibility of our ability to compose the meanings of complex expressions from the meanings of their constituents is supposed to be compositionality. By parity of reasoning, the best explanation for the possibility to decompose the meanings of complex expressions into the meanings of their constituents is *inverse compositionality*: that the meaning of any complex expression determines the meanings of its lexical constituents (as well as its syntactic structure) (compare Fodor and Lepore 2002, p. 59; Pagin 2003, p. 292). Compositionality and its inverse yield the view that the meanings of complex expressions can be viewed as having a structure isomorphic to the syntactic structures of those expressions, which in turn, may capture the idea behind the metaphor of the building principle.

See also Meaning; Syntax.

Bibliography

Carnap, R. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press, 1947.

Dever, J. "Compositionality as Methodology." *Linguistics and Philosophy* 22 (1999): 311–326.

Evans, G. *The Varieties of Reference*. Oxford, U.K.: Clarendon Press, 1982.

Fodor, J., and E. Lepore. "Why Compositionality Won't Go Away: Reflections on Horwich's 'Deflationary' Theory." In *The Compositionality Papers*. Oxford, U.K.: Clarendon Press, 2002.

Frege, G. "Letter to Jourdain." 1914? In *Philosophical and Mathematical Correspondence*, edited by G. Gabriel et al., 78–80. Chicago: University of Chicago Press, 1980.

Frege, G. "Letter to Ludwig Darmstaester." 1919. In *Posthumous Writings*, edited by H. Hermes et al., 253–257. Chicago: University of Chicago Press, 1979.

Frege, G. "On Concept and Object." 1892. In *Collected Papers on Mathematics, Logic, and Philosophy*, edited by B. McGuinness, 182–194. Oxford, U.K.: Blackwell, 1984.

Gazdar, G., E. Klein, G. Pullum, and I. Sag. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press, 1985.

Hodges, W. "Formal Features of Compositionality." *Journal of Logic, Language, and Information* 10 (2001): 7–28.

Husserl, E. *Logische Untersuchungen II/1*. Tübingen, Germany: Max Niemeyer, 1913.

Janssen, T. M. V. "Compositionality." In *Handbook of Logic and Language*, edited by J. van Benthem and A. ter Meulen, 417–473. Amsterdam: Elsevier, 1997.

Janssen, T. M. V. *Foundations and Applications of Montague Grammar*. Amsterdam: Mathematisch Centrum, 1986.

Kazmi, A., and F. J. Pelletier. "Is Compositionality Formally Vacuous?" *Linguistics and Philosophy* 21 (1998): 629–633.

Montague, R. "Universal Grammar." In *Formal Philosophy: Selected Papers of Richard Montague*, edited by R. Thomason, 222–246. New Haven, CT: Yale University Press, 1974.

Nunberg G., I. Sag, and T. Wasow. "Idioms." *Language* 70 (1994): 491–538.

Pagin, P. "Communication and Strong Compositionality." *Journal of Philosophical Logic* 32 (2003): 287–322.

Partee, B. "Compositionality." In *Varieties of Formal Semantics: Proceedings of the Fourth Amsterdam Colloquium, September 1982*, edited by F. Landman and F. Veltman, 281–312. Dordrecht, Netherlands: Foris, 1984.

Platts, M. *Ways of Meaning: An Introduction to the Philosophy of Language*. London: Routledge, 1979.

Szabó, Z. G. "Compositionality as Supervenience." *Linguistics and Philosophy* 23 (2000): 475–505.

Westerståhl, D. "On Mathematical Proofs of the Vacuity of Compositionality." *Linguistics and Philosophy* 21 (1998): 635–643.

Zadrozny, W. "From Compositional to Systematic Semantics." *Linguistics and Philosophy* 17 (1994): 329–342.

Zoltán Gendler Szabó (2005)

COMPUTABILITY THEORY

0. THE INFORMAL CONCEPT

Computability theory is the area of mathematics dealing with the concept of an effective procedure—a procedure that can be carried out by following specific rules. For example, one might ask whether there is some effective procedure—some algorithm—that, given a sentence about the positive integers, will decide whether that sentence is true or false. In other words, is the set of true sentences about the positive integers decidable? Or for a much simpler example, the set of prime numbers is certainly a decidable set. That is, there are mechanical procedures, that are taught in the schools, for deciding of any given positive integer whether or not it is a prime number.

More generally, consider a set S, which can be either a set of natural numbers (the natural numbers are 0, 1, 2, …), or a set of strings of letters from a finite alphabet. (These two situations are entirely interchangeable. A set of natural numbers is much like a set of base-10 numerals, which are strings of digits. And in the other direction, a string of letters can be coded by a natural number in a variety of ways. The best way is, where the alphabet has k symbols, to utilize k-adic notation, which is like base-k numerals except that the k digits represent 1, 2, … , k, without a 0 digit.) One can say that S is a decidable set if there exists an effective procedure that, given any natural number (in the first case) or string of letters (in the sec-

ond case), will eventually end by supplying the answer: "Yes" if the given object is a member of *S* and "No" if it is not a member of *S*.

And by an effective procedure here is meant a procedure for which one can give exact instructions—a program—for carrying out the procedure. Following these instructions should not demand brilliant insights on the part of the agent (human or machine) following them. It must be possible, at least in principle, to make the instructions so explicit that they can be executed by a diligent clerk (who is good at following directions but is not too clever) or even a machine (which does not think at all).

Although these instructions must of course be finite in length, no upper bound on their possible length is imposed. It is not ruled out that the instructions might even be absurdly long. Similarly, to obtain the most comprehensive concepts, no bounds are imposed on the time that the procedure might consume before it supplies the answer. Nor is a bound imposed on the amount of storage space (scratch paper) that the procedure might need to use. One merely insists that the procedure give an answer eventually, in some finite length of time.

Later, in section 7, more restrictive concepts will be considered, where the amount of time is limited in some way, so as to exclude the possibility of ridiculously long execution times. Initially, however, one wants to avoid such restrictions, to obtain the limiting case where practical limitations on execution time or memory space are removed.

This description of effective procedures, vague as it is, already shows how limiting the concept of decidability is. It is not hard to see that there are only countably many possible instructions of finite length that one can write out (using a standard keyboard, say). There are, however, uncountably many sets of natural numbers (by Cantor's diagonal argument). It follows that almost all sets, in a sense, are undecidable.

The following section will look at how the foregoing vague description of effective procedures can be made more precise—how it can be made into a mathematical concept. Nonetheless, the informal idea of what can be done by effective procedure, that is, what is calculable, can be useful.

For another example, consider what is required for a string of symbols to constitute an acceptable mathematical proof. Before one accepts a proof and adds the result being proved to the storehouse of mathematical knowledge, one insists that the proof be verifiable. That is, it should be possible for another mathematician, such as the referee of the paper containing the proof, to check, step by step, the correctness of the proof. Eventually, the referee concludes either that the proof is indeed correct or that the proof contains a gap or an error and is not yet acceptable. That is, the set of acceptable mathematical proofs should be decidable. This fact will be seen (in section 4) to have significant consequences for what can and cannot be proved. The conclusion follows that computability theory is relevant to the foundations of mathematics.

Before going on, one should broaden the canvas from considering decidable and undecidable sets to considering the more general situation of partial functions. Let U be either the set $\mathbb{N} = \{0,1,2, \dots\}$ of natural numbers or the set Σ^* of all strings of letters—all words—from a finite alphabet Σ. Then a k-place partial function on U is a function whose domain is included in $U^k = U \times U \times \dots \times U$ and whose range is included in U. And one can say that such a function is total if its domain is all of U^k.

For a k-place partial function f, one can say that f is an effectively calculable partial function if there exists an effective procedure with the following property:

- Given a k-tuple x in the domain of f, the procedure eventually halts and returns the correct value for $f(x)$
- Given a k-tuple x *not* in the domain of f, the procedure does not halt and return a value

(Strictly speaking, when U is $\mathbb{N}$, the procedure cannot be given numbers, it must be given numerals. Numerals are bits of language, which can be communicated. Numbers are not. Thus, the difference between $U = \mathbb{N}$ and $U = \Sigma^*$ is even less than previously indicated.)

For example, the partial function for subtraction

$$f(m,n) = \begin{cases} m - n & \text{if } m \geq n \\ \uparrow & \text{otherwise} \end{cases}$$

(where $\uparrow$ indicates that the function is undefined) is effectively calculable, and procedures for calculating it, using base-10 numerals, are taught in the elementary schools.

The concept of decidability can then be described in terms of functions: For a subset S of U^k, one can say that S is decidable if its characteristic function

$$C_S(x) = \begin{cases} \text{Yes} & \text{if } x \in S \\ \text{No} & \text{if } x \notin S \end{cases}$$

(which is always total) is effectively calculable. Here, "Yes" and "No" are some fixed members of U, such as 1 and 0 in the case of $\mathbb{N}$.

Here, if $k = 1$, then S is a set of numbers or a set of words. If $k = 2$, then one has the concept of a decidable binary relation on numbers or words, and so forth.

And it is natural to extend this concept to the situation where one has half of decidability: Say that S is semidecidable if its partial characteristic function

$$c_S(x) = \begin{cases} \text{Yes} & \text{if } x \in S \\ \uparrow & \text{if } x \notin S \end{cases}$$

is an effectively calculable partial function. Thus, a set S of words—a language—is semidecidable if there is an effective procedure for recognizing members of S. One can think of S as the language that the procedure accepts.

The following is another example of a calculable partial function:

$F(n)$ = the smallest $p > n$ such that both p and $p + 2$ are prime

Here, it is to be understood that $F(n)$ is undefined if there is no number p as described; thus F might not be total. For example, $F(9) = 11$. It is not known whether or not F is total. Nonetheless, one can be certain that F is effectively calculable. One procedure for calculating $F(n)$ proceeds as follows. "Given n, first put $p = n + 1$. Then check whether or not p and $p + 2$ are both prime. If they are, then stop and give output p. If not, increment p and continue." What if $n = 10^{1000}$? On the one hand, if there is a larger prime pair, then this procedure will find the first one, and halt with the correct output. On the other hand, if there is no larger prime pair, then the procedure never halts, so it never gives an answer. That is all right, because $F(n)$ is undefined—the procedure should not give any answer. (Of course, F is total if and only if (iff) the twin prime conjecture is true.)

Now suppose one modifies this example. Consider the total function:

$$G(n) = \begin{cases} F(n) & \text{if } F(n)\downarrow \\ 0 & \text{otherwise} \end{cases}$$

Here, $F(n)\downarrow$ means that $F(n)$ is defined so that n belongs to the domain of F. Then the function G is also effectively calculable. That is, there exists a program that calculates G correctly. That is not the same as saying that one knows that program. This example indicates the difference between knowing that a certain effective procedure exists and having the effective procedure in one's hands.

One person's program is another person's data. This is the principle behind operating systems (and behind the idea of a stored-program computer). One's favorite program is, to the operating system, another piece of data to be received as input and processed. The operating system is calculating the values of a two-place "universal" function, as in the following example.

Suppose one adopts a fixed method of encoding any set of instructions by a single natural number. (First, one converts the instructions to a string of 0s and 1s—one always does this with computer programs—and then one regards that string as naming a natural number under a suitable base-2 notation.) Then, the universal function

$\Phi(x, y)$ = the result of applying the instructions coded by y to the input x

is an effectively calculable partial function (where it is understood that $\Phi(x, y)$ is undefined whenever applying the instructions coded by y to the input x fails to halt and return an output). Here are the instructions for Φ: "Given x and y, decode y to see what it says to do with x, and then do it." Of course, the function Φ is not total.

Using this universal partial function, one can construct an undecidable binary relation, the halting relation H:

$$\begin{aligned} (x, y) \in H &\Leftrightarrow \Phi(x, y)\downarrow \\ &\Leftrightarrow \text{applying the instructions coded by } y \text{ to input } x \text{ halts} \end{aligned}$$

To see that H is undecidable, one can argue as follows. Suppose that, to the contrary, H is decidable. Then the following function would be effectively calculable:

$$f(x) = \begin{cases} \text{Yes} & \text{if } \Phi(x, x)\uparrow \\ \uparrow & \text{if } \Phi(x, x)\downarrow \end{cases}$$

(Notice the use of the classical diagonal construction.) (To compute $f(x)$, one first would decide if $(x, x) \in H$. If not, then $f(x)$ = Yes. If $(x, x) \in H$, however, then the procedure for finding $f(x)$ should throw itself into an infinite loop, because $f(x)$ is undefined.) The function f cannot possibly be effectively calculable, however. Consider any set of instructions that might compute f. Those instructions have some code number k, but f has been constructed in such a way that $f(k)$ differs from the output from the result of applying instructions coded by k to the

input k. (They differ because one is defined and one is not.) So these instructions cannot correctly compute f; they produce the wrong result at the input k. And so one has a contradiction. That the previous relation H is undecidable is usually expressed by saying that "the halting problem is unsolvable"; that is, one cannot effectively determine, given x and y, whether applying the instructions coded by y to the input x will eventually terminate or will go on forever.

While the concept of effective calculability has been described in somewhat vague terms here, the following section will give a precise (mathematical) concept of a computable partial function. And then it will be argued that the mathematical concept of a computable partial function is the correct formalization of the informal concept of an effectively calculable partial function. This claim is known as Church's thesis or the Church-Turing thesis. Church's thesis, which relates an informal idea to a formal idea, is not itself a mathematical statement, capable of being given a proof, but one can look for evidence for or against Church's thesis; it all turns out to be evidence in favor.

One piece of evidence is the absence of counterexamples. That is, any function examined thus far that mathematicians have felt was effectively calculable, has been found to be computable.

Stronger evidence stems from the various attempts that different people made independently, trying to formalize the idea of effective calculability. Alonzo Church used λ-calculus, Alan M. Turing used an idealized computing agent (later called a Turing machine), and Emil Post developed a similar approach. Remarkably, all these attempts turned out to be equivalent, in that they all defined exactly the same class of functions, namely, the computable partial functions!

The study of effective calculability originated in the 1930s with work in mathematical logic. As noted earlier, the subject is related to the concept on an acceptable proof. Since the development of modern computers the study of effective calculability has formed an essential part of theoretical computer science. A prudent computer scientist would surely want to know that, apart from the difficulties the real world presents, there is a purely theoretical limit to calculability.

1. FORMALIZATIONS

In the preceding section, the concept of effective calculability was described only informally. Now, these ideas will be made more precise (i.e., will be made part of mathematics). In fact, several approaches to doing this will be described: idealized computing devices, generative definitions (i.e., the least class containing certain initial functions and closed under certain constructions), programming languages, and definability in formal languages. It is a significant fact that these different approaches all yield exactly equivalent concepts.

TURING MACHINES. In early 1935 Alan M. Turing was a twenty-two-year-old graduate student at King's College in Cambridge. Under the guidance of Max Newman, he was working on the problem of formalizing the concept of effective calculability. In 1936 he learned of the work of Alonzo Church at Princeton University. Church had also been working on this problem, and in his 1936 paper "An Unsolvable Problem of Elementary Number Theory" he presented a definite conclusion: that the class of effectively calculable functions should be identified with the class of functions definable in the λ-calculus, a formal language for specifying the construction of functions. Moreover, he showed that exactly the same class of functions could be characterized in terms of formal derivability from equations.

Turing then promptly completed writing his paper, in which he presented a different approach to characterizing the effectively calculable functions, but one that—as he proved—yielded once again the same class of functions as Church had proposed. With Newman's encouragement, Turing then went to Princeton for two years, where he wrote a doctoral dissertation under Church.

Turing's paper remains a readable introduction to his ideas. How might a diligent clerk carry out a calculation, following instructions? He might organize his work in a notebook. At any given moment his attention is focused on a particular page. Following his instructions, he might alter that page, and then he might turn to another page. And the notebook is large enough that he never comes to the last page.

The alphabet of symbols available to the clerk must be finite; if there were infinitely many symbols, then there would be two that were arbitrarily similar and so might be confused. One can then without loss of generality regard what can be written on one page of notebook as a single symbol. And one can envision the notebook pages as being placed side by side, forming a paper tape, consisting of squares, each square being either blank or printed with a symbol. At each stage of his work, the clerk—or the mechanical machine—can alter the square under examination, can turn attention to the next square or the previous one, and can look to the instructions to

see what part of them to follow next. Turing described the latter part as a "change of state of mind."

Turing wrote, "We may now construct a machine to do the work" (1936–1937, p. 251). Of course, such a machine is now called a Turing machine, a phrase first used by Church in his review of Turing's paper in *The Journal of Symbolic Logic*. The machine has a potentially infinite tape, marked into squares. Initially, the given input numeral or word is written on the tape, but it is otherwise blank. The machine is capable of being in any one of finitely many states (the phrase "of mind" being inappropriate for a machine). At each step of calculation, depending on its state at the time, the machine can change the symbol in the square under examination at that time, can turn its attention to the square to the left or to the right, and can then change its state to another state.

The program for this Turing machine can be given by a table. Where the possible states of the machine are q_1, …, q_r, each line of the table is a quintuple $\langle q_i, S_j, S_k, D, q_m\rangle$, which is to be interpreted as directing that whenever the machine is in state q_i and the square under examination contains the symbol S_j, then that symbol should be altered to S_k and the machine should shift its attention to the square on the left (if $D = L$) or on the right (if $D = R$), and should change its state to q_m. For the program to be unambiguous, it should have no two different quintuples with the same first two components. (By relaxing this requirement regarding absence of ambiguity, one obtains the concept of a nondeterministic Turing machine, which will be useful later, in the discussion of feasible computability.) One of the states, say q_1, is designated as the initial state—the state in which the machine begins its calculation. If one starts the machine running in this state and examining the first square of its input, it might (or might not), after some number of steps, reach a state and a symbol for which its table lacks a quintuple having that state and symbol for its first two components. At that point the machine halts, and one can look at the tape (starting with the square then under examination) to see what the output numeral or word is.

Now suppose that Σ is a finite alphabet and that f is a k-place partial function on the set Σ^* of words. One says that f is Turing computable if there exists a Turing machine M that, when started in its initial state scanning the first symbol of a k-tuple $\vec{w}$ of words (written on the tape, with a blank square between words, and with everything to the right of $\vec{w}$ blank), behaves as follows:

- If $f(\vec{w})\downarrow$ (i.e., if $\vec{w} \in \text{dom } f$), then M eventually halts, and at that time it is scanning the leftmost symbol of the word $f(\vec{w})$ (which is followed by a blank square).
- If $f(\vec{w})\uparrow$ (i.e., if $\vec{w} \notin \text{dom } f$), then M never halts.

This definition can be readily adapted to apply to k-place partial functions on $\mathbb{N}$.

Then Church's thesis, also called—particularly in the context of Turing machines—the Church-Turing thesis, is the claim that this concept of Turing computability is the correct formalization of the informal concept of effective calculability. Certainly, the definition reflects the ideas of following predetermined instructions, without limitation of the amount of time that might be required. (The name *Church-Turing thesis* obscures the fact that Church and Turing followed different paths in reaching equivalent conclusions.)

As will be explained shortly, Church's thesis has by now achieved universal acceptance. Kurt Gödel, writing in 1964 about the concept of a formal system in logic, involving the idea that the set of correct deductions must be a decidable set, said that "due to A. M. Turing's work, a precise and unquestionably adequate definition of the general concept of formal system can now be given" (Davis 1965, p. 71).

The robustness of the concept of Turing computability is evidenced by the fact that it is insensitive to certain modifications to the definition of a Turing machine. For example, one can impose limitations on the size of the alphabet, or one can insist that the machine never move to the left of its initial starting point. None of this will affect that class of Turing computable partial functions.

Turing developed these ideas before the introduction of modern digital computers. After World War II he played an active role in the development of early computers and in the emerging field of artificial intelligence. (During the war, he worked on deciphering the German battlefield code Enigma, work that remained classified until after Turing's death.) One can speculate whether Turing might have formulated his ideas somewhat differently, if his work had come after the introduction of digital computers.

PRIMITIVE RECURSIVENESS AND MINIMALIZATION. For a second formalization of the calculability concept, a certain class of partial functions on $\mathbb{N}$ will now be defined as the smallest class that contains certain initial function and is closed under certain constructions.

For the initial functions, one can take the following simple total functions:

- The zero functions, that is, the constant functions f defined by the equation:

$$f(x_1, \dots , x_k) = 0$$

- The successor function S, defined by the equation:

$$S(x) = x + 1$$

- The projection functions I^k_n from k-dimensions onto the nth coordinate,

$$I^k_n(x_1,\dots,x_k) = x_n$$

where $1 \leq n \leq k$.

One can form the closure of the class of initial functions under three constructions: composition, primitive recursion, and minimalization.

A k-place function h is said to be obtained by composition from the n-place function f and the k-place functions $g_1, \dots , g_n$ if the equation

$$h(\vec{x}) = f(g_1(\vec{x}), \dots , g_n(\vec{x}))$$

holds for all $\vec{x}$. In the case of partial functions, it is to be understood here that $h(\vec{x})$ is undefined unless $g_1(\vec{x}), \dots , g_n(\vec{x})$ are all defined and $\langle g_1(\vec{x}), \dots , g_n(\vec{x})\rangle$ belongs to the domain of f.

A $(k + 1)$-place function h is said to be obtained by primitive recursion from the k-place function f and the $(k + 2)$-place function g (where $k > 0$) if the pair of equations

$$\begin{aligned} h(\vec{x},0) &= f(\vec{x}) \\ h(\vec{x},t+1) &= g(t,h(\vec{x},t),\vec{x}) \end{aligned}$$

holds for all $\vec{x}$ and t.

Again, in the case of partial functions, it is to be understood that $h(\vec{x}, t + 1)$ is undefined unless $h(\vec{x}, t)$ is defined and $\langle t, h(\vec{x}, t), \vec{x} \rangle$ is in the domain of g.

For the $k = 0$ case, the one-place function h is obtained by primitive recursion from the two-place function g with the number m if the pair of equations

$$h(0) = m$$

$$h(t + 1) = g(t, h(t))$$

holds for all t.

Postponing the matter of minimalization, one can define a function to be primitive recursive if it can be built up from zero, successor, and projection functions by use of composition and primitive recursion. In other words, the class of primitive recursive functions is the smallest class that includes the initial functions and is closed under composition and primitive recursion.

Clearly, all the primitive recursive functions are total. One can say that a k-ary relation R on $\mathbb{N}$ is primitive recursive if its characteristic function is primitive recursive.

One can then show that a great many of the common functions on $\mathbb{N}$ are primitive recursive: addition, multiplication, … , the function whose value at m is the $(m + 1)$st prime, …

On the one hand, it is clear that every primitive recursive function should be regarded as being effectively calculable. On the other hand, the class of primitive recursive functions cannot possibly comprehend all total calculable functions, because one can easily "diagonalize out" of the class. That is, by suitably indexing the "family tree" of the primitive recursive functions, one can make a list $f_0, f_1, f_2, \dots$ of all the one-place primitive recursive functions. One can then consider the diagonal function $d(x) = f_x(x) + 1$. Then d cannot be primitive recursive; it differs from each f_x at x. Nonetheless, if one makes the list tidely, the function d is effectively calculable. The conclusion is the class of primitive recursive functions is an extensive but proper subset of the total calculable functions.

Next, one can say that a k-place function h is obtained from the $k + 1$-place function g by minimalization and one writes

$$h(\vec{x}) = \mu y\,[g(\vec{x}, y) = 0]$$

if for each $\vec{x}$, the value $h(\vec{x})$ either is the number y such that $g(\vec{x}, y) = 0$ and $g(\vec{x}, s)$ is defined and is nonzero for every $s < y$, if such a number y exists, or else is undefined, if no such number y exists. The idea behind this μ-operator is the idea of searching for the least number y that is the solution to an equation, by testing successively $y = 0, 1, \dots$

One can obtain the general recursive functions by adding minimalization to the closure methods. That is, a partial function is general recursive if it can be built up from the initial zero, successor, and projection functions by use of composition, primitive recursion, and minimalization.

The class of general recursive functions is (as Turing proved) exactly the same as the class of Turing computable functions. And Church's thesis therefore has the

equivalent formulation that the concept of a general recursive function is the correct formalization of the informal concept of effective calculability.

What if one tries to diagonalize out of the class of general recursive functions, as one did for the primitive recursive functions? As will be argued later, one can again make a tidy list $\varphi_0, \varphi_1, \varphi_2, \ldots$ of all the one-place general recursive partial functions. And one can define the diagonal function $d(x) = \varphi_x(x) + 1$. In this equation, $d(x)$ is undefined unless $\varphi_x(x)$ is defined. The diagonal function d is indeed among the general recursive partial functions, and hence is φ_k for some k, but $d(k)$ must be undefined. No contradiction results.

The class of primitive recursive functions was defined by Gödel, in his 1931 paper on the incompleteness theorems. Of course, the idea of defining functions on $\mathbb{N}$ by recursion is much older and reflects the idea that the natural numbers are built up from the number 0 by repeated application of the successor function. The theory of the general recursive functions was worked out primarily by Stephen Cole Kleene, a student of Church.

The use of the word *recursive* in the context of the primitive recursive functions is entirely reasonable. Gödel, writing in German, had used simply *rekursiv* for the primitive recursive functions. Retaining the word *recursive* for the general recursive functions was a, however, historical accident. The class of general recursive functions—as this section shows—has several characterizations in which *recursion* (i.e., defining a function in terms of its other values, or using routines that call themselves) plays no role at all.

Nonetheless, the terminology became standard. What are here called the computable partial functions were until the late 1990s standardly called the partial recursive functions. And for that matter, computability theory was called recursive function theory for many years, and then recursion theory. And relations on $\mathbb{N}$ were said to be recursive if their characteristic functions were general recursive functions.

An effort is now being made, however, to change what had been the standard terminology. Accordingly, this entry speaks of computable partial functions. And it will call a relation computable if its characteristic function is a computable function. Thus, the concept of a computable relation corresponds to the informal notion of a decidable relation. In any case, there is definitely a need to have separate adjectives for the informal concept (here, *calculable* is used for functions, and *decidable* for relations) and the formally defined concept (here, *computable*).

LOOP AND WHILE PROGRAMS. The idea behind the concept of effective calculable functions is that one should be able to give explicit instructions—a program—for calculating such a function. What programming language would be adequate here? Actually, any of the commonly used programming languages would suffice, if freed from certain practical limitations, such as the size of the number denoted by a variable. One can give here a simple programming language with the property that the programmable functions are exactly the computable partial functions on $\mathbb{N}$.

The variables of the language are $X_0, X_1, X_2, \ldots$ Although there are infinitely many variables in the language, any one program, being a finite string of commands, can have only finitely many of these variables. If one wants the language to consist of words over a finite alphabet, one can replace X_3, say, by X'''.

In running a program, each variable in the program gets assigned a natural number. There is no limit on how large this number can be. Initially, some of the variables will contain the input to the function; the language has no input commands. Similarly, the language has no output commands; when (and if) the program halts, the value of X_0 is to be the function value.

The commands of the language come in five kinds:

(1) $X_n \leftarrow 0$. This is the clear command; its effect is to assign the value 0 to X_n.

(2) $X_n \leftarrow X_n + 1$. This is the increment command; its effect is to increase the value assigned to X_n by one.

(3) $X_n \leftarrow X_m$. This is the copy command; its effect is just what the name suggests; in particular it leaves the value of X_m unchanged.

(4) Loop X_n and endloop X_n. These are the loop commands, and they must be used in pairs. That is, if $\mathcal{P}$ is a program—a syntactically correct string of commands—then so is the string:

loop X_n

$\mathcal{P}$

endloop X_n

What this program means is that $\mathcal{P}$ is to be executed a certain number k of times. And that number k is the initial value of X_n, the value assigned to X_n before

one starts executing $\mathcal{P}$. Possibly, $\mathcal{P}$ will change the value of X_n; this has no effect at all on k.

(5) While $X_n \neq 0$ and endwhile $X_n \neq 0$. These are the while commands; again, they must be used in pairs, like the loop commands, but there is a difference. The program

$$\text{while } X_n \neq 0$$

$$\mathcal{P}$$

$$\text{endwhile } X_n \neq 0$$

also executes the program $\mathcal{P}$ some number k of times. Now, however, k is not determined in advance; it matters very much how $\mathcal{P}$ changes the value of X_n. The number k is the least number (if any) such that executing $\mathcal{P}$ that many times causes X_n to be assigned the value 0. The program will run forever if there is no such k.

And those are the only commands. A while program is a sequence of commands, subject only to the requirement that the loop and while commands are used in pairs, as illustrated. Clearly, this programming language is simple enough to be simulated by any of the common programming language, if one ignores overflow problems.

A loop program is a while program with no while commands; that is, it has only clear, increment, copy, and loop commands. Note the important property: A loop program always halts, no matter what. It is easy, however, to make a while program that never halts.

One can say that a k-place partial function f on $\mathbb{N}$ is while-computable if there exists a while program $\mathcal{P}$ that, whenever started with a k-tuple $\vec{x}$ assigned to the variables $X_1, \ldots, X_k$ and 0 assigned to the other variables, behaves as follows:

- If $f(\vec{x})$ is defined, then the program eventually halts, with X_0 assigned the value $f(\vec{x})$.
- If $f(\vec{x})$ is undefined, then the program never halts.

The loop-computable functions are defined in the analogous way. There is the difference, however, that any loop-computable function is total.

Theorem. (a) A function on $\mathbb{N}$ is loop-computable iff it is primitive recursive.

(b) A partial function on $\mathbb{N}$ is while-computable iff it is general recursive.

The proof in one direction, to show that every primitive recursive functions is loop-computable, involves a series of programming exercises. The proof in the other direction involves coding the status of a program $\mathcal{P}$ on input $\vec{x}$ after t steps, and showing that there are primitive recursive functions enabling one to determine the status after $t + 1$ steps, and the terminal status. Because the class of general recursive partial functions coincides with the class of Turing computable partial functions, one can conclude from the previous theorem that while-computability coincides with Turing computability.

DEFINABILITY IN FORMAL LANGUAGES. In his 1936 paper in which he presented what is now known as Church's thesis, Church utilized a formal system, the λ-calculus. Church had developed this system as part of his study of the foundations of logic. In particular, for each natural number n there is a formula $\boldsymbol{n}$ of the system denoting n, that is, a numeral for n. More important, formulas could be used to represent the construction of functions. He defined a two-place function F to be λ-definable if there existed a formula $\boldsymbol{F}$ of the λ-calculus such wherever $F(m, n) = r$ then the formula $\{\boldsymbol{F}\}(\boldsymbol{m}, \boldsymbol{n})$ was convertible, following the rules of the system, to the formula $\boldsymbol{r}$, and only then. An analogous definition applied to k-place functions.

Church's student, Stephen Cole Kleene, showed that a function was λ-definable iff it was general recursive. (Church and his student, J. B. Rosser, were also involved in the development of this result.) Church wrote in his paper, "The fact … that two such widely different and (in the opinion of the author) equally natural definitions of effective calculability turn out to be equivalent adds to the strength of reasons … for believing that they constitute as general a characterization of this notion as is consistent with the usual intuitive understanding of it" (Alonzo 1936, p. 346).

Earlier, in 1934, Gödel, in his lectures at Princeton, formulated a concept now referred to as Gödel-Herbrand computability. He did not, however, at the time propose the concept as a formalization of the concept of effective calculability. The concept involved a formal calculus of equations between terms built up from variables and function symbols. The calculus permitted the passage from an equation $A = B$ to another equation obtained by substituting for a part C of A or B another term D where the equation $C = D$ had been derived. If a set $\mathscr{E}$ of equations allowed the derivation, in a suitable sense, of exactly the right values for a function f on $\mathbb{N}$, then $\mathscr{E}$ was said to be a set of recursion equations for f. Once again, it turned out that a set of recursion equations existed for f iff f was a general recursive function.

A rather different approach to characterizing the effectively calculable functions involved definability in first-order logic over the structure of the natural numbers with addition and multiplication. Say that a k-place partial function f on $\mathbb{N}$ is a Σ_1-function if the graph of f (i.e., the $(k + 1)$-ary relation $\{\langle x_1, \ldots, x_k, y\rangle \mid f(x_1, \ldots, x_k) = y\}$) is definable in the structure with universe $\mathbb{N}$ and with the operations of addition, multiplication, and exponentiation, by an existential formula (i.e., a formula consisting of a string of existential quantifiers, followed by a quantifier-free part). Then the class of partial Σ_1-functions coincides exactly with the class of partial functions given by the other formalizations of calculability described here. Moreover, Yuri Matijasevič showed in 1970 that the operation of exponentiation was not needed here.

Finally, say that a k-place partial function f on $\mathbb{N}$ is representable if there exists some finitely axiomatizable theory T in a language having a suitable numerals $\boldsymbol{n}$ for each natural number n, and there exists a formula φ of that language such that (for any natural numbers) $f(x_1, \ldots, x_k) = y$ iff $\varphi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{y})$ is a sentence deducible in the theory T. Then once again the class of representable partial functions coincides exactly with the class of partial functions given by the other formalizations of calculability described here.

2. BASIC RESULTS

First, one has the remarkable fact that all the formalizations of the preceding section yield exactly the same class of partial functions on $\mathbb{N}$. And this fact is not only remarkable, it is also reassuring, indicating that the concept captured by the formalizations—the concept of a computable partial function—is natural and significant. Moreover, it gives evidence that the concept captured by the formalizations is actually the correct formalization of the informal concept of effective calculability. That is, it gives evidence for Church's thesis (or the Church-Turing thesis). This thesis was first set forth by Church in a 1935 abstract, and then published in full in his 1936 paper. (At the time, Church was unaware of Turing's approach, but he knew of several of the other formalizations described in the preceding section.) This assertion, that computability is the precise counterpart to effective calculability, is not really a mathematical statement susceptible of proof or disproof; rather, it is a judgment that one has found the correct formalization of the one's informal concept.

The situation can be compared to one encountered in calculus. An intuitively continuous function (defined on an interval) is one whose graph one can draw without lifting the pencil off the paper. To prove theorems, however, some formal counterpart of this notion is needed. And so one gives the usual definition of ε-δ-continuity. One should ask if the precise notion of ε-δ-continuity is an accurate formalization of intuitive continuity. If anything, the class of ε-δ-continuous functions is too broad. It includes nowhere differentiable functions, whose graphs cannot be drawn without lifting the pencil—there is no way to impart a velocity vector to the pencil. Nonetheless, the class of ε-δ-continuous functions has been found to be a natural and important class in mathematical analysis.

In a similar spirit, one can ask how accurately the formal concept of computability captures the informal concept of effective calculability. As with continuous functions, the precisely defined class (of computable functions) appears to be, if anything, too broad. It includes functions for which any procedure will, for large inputs, require so much computing time and memory (scratch paper) space as to make implementation absurd. Computability corresponds to calculability in an idealized world, where length of computation and amount of memory are disregarded. (This will be discussed further in section 7.) In any case, however, the class of computable partial functions has been found to be a natural and important class in mathematical logic.

Empirical evidence that the class of computable functions is not too narrow is provided both by the fact that the attempted formalizations (as described in section 1) have all yielded the equivalent concepts, and by the fact that no counterexample have arisen—the functions considered thus far that mathematicians have felt were effectively calculable have turned out to be computable. In the decades since 1936, Church's thesis has gained universal acceptance.

NORMAL FORM. In each of the formalizations described in the preceding section, one can in a straightforward way code the instructions for any computable partial function by a natural number e. In the case of Turing machines, e encodes the set of quintuples that determine the machine's operation. In the case of a function built up from the zero, successor, and projection functions by primitive recursion and minimalization, e encodes the ancestral tree describing exactly how the function is built up. In the case of while programs, e encodes the program.

Normal form theorem. There is a ternary computable relation T and a total computable function U with the following property: For each 1-place computable

partial function f on $\mathbb{N}$, there is a natural number e such that

$$f(x) = U(\mu y T(e, x, y))$$

for every number x.

Here (as elsewhere), equality has the natural meaning: Either both sides of the equation are defined and are the same, or else both sides are undefined.

One can construct the relation T (called the Kleene T-predicate) so that $T(e, x, y)$ expresses the idea that e encodes the instructions for f, and y encodes the entire history of the step-by-step computation of f with input x, from the beginning through the final step at which the computational procedure comes to a halt. Then the function U (the upshot function) extracts from y what the answer or output is.

The normal form theorem can be extended to k-place functions. One can make a $(k + 2)$-ary computable relation T_k such that for each k-place computable partial function f, there is a number e such that

$$f(x_1, \dots, x_k) = U(\mu y T_k(e, x_1, \dots, x_k, y))$$

for every $x_1, \dots, x_k$. Moreover, one can construct T_k and U so that they are even primitive recursive.

The significance of the normal form theorem is that it allows one to form a universal partial computable function. One can define

$$\varphi_e(x) = U(\mu y T(e, x, y))$$

(where, of course, $\varphi_e(x)\uparrow$ if the right side of the equation is undefined, which happens if there does not exist a y such that $T(e, x, y)$). Then on the one hand, $\varphi_e(x)$ is a computable partial 2-place function of x and e. And on the other hand, each 1-place computable partial function equals φ_e for some e. That is,

$$\varphi_0, \varphi_1, \varphi_2, \dots$$

is a complete list of all the computable partial 1-place functions.

Similarly, one can extend these ideas to k-place partial functions:

$$\varphi_e^k(x_1, \dots, x_k) = U(\mu y T_k(e, x_1, \dots, x_k, y))$$

Then

$$\varphi_0^k, \varphi_1^k, \varphi_2^k, \dots$$

is a complete list of all the computable partial k-place functions.

Whenever one has such a list, one can diagonalize out of it. One can define the set K by the condition

$$x \in K \Leftrightarrow \varphi_x(x)\downarrow$$

so that a number x (thought of as encoding a program for computing a partial function) belongs to K iff that program, given x itself as input, halts and returns a value.

Then the diagonal function

$$d(x) = \begin{cases} \varphi_x(x) + 1 & \text{if } x \in K \\ 0 & \text{otherwise} \end{cases}$$

is a total function, but it cannot equal φ_e for any e, because it differs from φ_e at e. So d is not a computable function. If K were computable, however, then d would be computable, because the partial function $\varphi_x(x) + 1$ is computable.

One can conclude that K is not a computable set; its characteristic function C_K is not a computable function. But the partial characteristic function

$$c_k(x) = \begin{cases} 1 & \text{if } x \in K \\ \uparrow & \text{otherwise} \end{cases}$$

is a computable partial function; $c_k(x) = 1 + 0 \cdot \varphi_x(x)$.

Theorem. For a set A of numbers, the following are equivalent:

(1) The partial characteristic function of A is a computable partial function

(2) A is the domain of some computable partial function

(3) For some computable binary relation R,

$$x \in A \Leftrightarrow R(x, y) \text{ for some } y$$

(Here (2) $\Rightarrow$ (3) because $x \in \text{dom } \varphi_e \Leftrightarrow T(e, x, y)$ for some y. And (3) $\Rightarrow$ (1) because one can use the function $1 + 0 \cdot \mu y R(x, y)$.)

A set A with the properties of this theorem is said to be computably enumerable (c.e.). The concept of a c.e. set is the formalization of the informal concept of a semidecidable set, discussed in section 0. And Church's thesis assures one that it is the correct formalization.

In the previously standard terminology mentioned earlier, a set A with the properties of the theorem was said to be recursively enumerable (r.e.). In fact, this terminol-

ogy—especially the abbreviation—has become so well established that the prospects for reform are uncertain.

The theorem extends to the case where A is a k-ary relation on $\mathbb{N}$; now in part (3) the relation R is $k + 1$-ary. Thus, one may speak of c.e. (or r.e.) relations on $\mathbb{N}$.

Unsolvability of the halting problem. The binary relation $\{\langle x, y\rangle \mid \varphi_y(x)\downarrow\}$ is c.e. but not computable.

This relation—the halting relation—cannot be computable lest the previous diagonal function d be computable. It is c.e., because $\varphi_y(x)\downarrow \Leftrightarrow \exists z T(y, x, z)$.

If one defines $W_e = \mathrm{dom}\varphi_e$, then as a consequence of the normal form theorem, one has a complete list

$$W_0, W_1, W_2, \ldots$$

of all the c.e. sets. The set K can be described simply as $\{x \mid x \in W_x\}$.

The following is not hard to see:

Kleene's theorem. A set is computable iff both it and its complement are c.e.

For example, the complement $\overline{K}$ is not only noncomputable, it is not even c.e.

For another example of an undecidable set, take the set of programs that compute total functions:

$$\mathrm{Tot} = \{e \mid \varphi_e \text{ is total}\}$$

The same argument used for K shows that Tot is not computable. Moreover, Tot is not c.e. In fact, a slightly stronger statement holds: There is no c.e. set P such that $\{\varphi_e \mid e \in P\}$ coincides with the class of total computable functions. Thus, if P is a c.e. set of programs that compute only total functions, then there must be some total computable function with no program in P.

Rice's theorem. Suppose that C is a collection of computable partial 1-place functions, and let I be $\{e \mid \varphi_e \in C\}$. Then I is computable only in two trivial cases: when C is empty and when C is the collection of all computable partial functions.

For example, suppose one focuses attention on a particular computable partial function f. Rice's theorem asserts that one cannot always decide of a given program whether or not that program correctly computes f.

The name *computably enumerable* corresponds to yet another characterization: A set is c.e. iff there is a Turing machine (augmented with a suitable output tape) that can generate, in some order, the members of that set, one after another. More formally, a set S of natural numbers is c.e. iff it is either empty or is the range of some total computable function f, that is, $S = \{f(0), f(1), \ldots\}$. In fact, one can even insist that f be primitive recursive. In general the function f will not enumerate the members of S in numerical order, however (i.e., f will not in general be an increasing function). The range of an increasing function (or even of a nondecreasing function) will always be a computable set.

It is easy to see that if f is a two-place computable partial function, then the result of holding one variable fixed (as a parameter)

$$g(x) = f(36, x)$$

is a one-place computable partial function g. Often, one needs the more subtle fact: A program for g can be effectively found from the program for f and the value of the parameter.

Parameter theorem. There is a total computable function ρ such that

$$\varphi_e(t, x) = \varphi_{\rho(e,\ t)}(x)$$

for all e, t, and x.

The analogous statement holds for more variables, that is, for an m-tuple $\vec{t}$ and an n-tuple $\vec{x}$ in place of t and x. The parameter theorem commonly goes by the cryptic name of the S-m-n theorem.

A deeper result is the following theorem, which is due to Kleene.

Recursion theorem. For any computable partial function g, one can find an e such that

$$\varphi_e(x) = g(e, x)$$

for all x.

Again, x can be replaced by an n-tuple $\vec{x}$. The proof of the recursion theorem is similar to the argument used to produce self-referential sentences in number theory, such as those used in proving Gödel's incompleteness theorem.

3. AXIOMATIZABLE THEORIES

The connection between computability theory and logic hinges on the fact that proofs must be effectively recognizable.

The concept of a proof is basic to logic. What exactly is a proof? As indicated in section 0, for a proof to be acceptable, it must be possible—in principle—to fill in enough steps that a hard-working graduate student (or a referee) can verify its correctness. One cannot demand that this student have the same brilliant insight that the proof's discoverer had. Nor can one demand that the stu-

dent spend an infinite amount of time checking an infinite number of cases. What one can insist is that, given a correct proof (with all the steps filled in), the student will eventually complete the verification and stand up and say, "Yes, this proof is correct."

This is just to say, however, that the set of correct proofs must be at least semidecidable. And typically one expects that the set will even be decidable, lest the student work forever attempting to verify an incorrect proof. In an axiomatic theory, one expects to be able to tell (effectively) an axiom from a nonaxiom, and one expects to be able to determine (effectively) whether or not a rule of inference is being correctly applied.

Even with the weaker property of semidecidability, it follows that the set of theorems—the set of sentences that have proofs—is semidecidable. (Given a sentence, one could employ the brute-force procedure of going through all strings of symbols in a systematic way, spending more and more time on each, attempting to verify that it is a proof of that sentence.) That is, the set of theorems must be c.e.

More formally, assume one has a first-order language, such as the language for set theory. Formulas are (or can be made to be) strings over a finite alphabet, so the concepts of computability theory are applicable. (It is being assumed here that the language has a reasonably simple array of nonlogical symbols.) One can define a theory to be a set of sentences closed under logical consequence. In particular, for a set A of formulas adopted as axioms, one can obtain the theory T_A consisting of all sentences that are logical consequences of A.

Theorem. (a) If A is a computable set or a c.e. set of axioms, then the set T_A of logical consequences of A is c.e.

(b) (Craig's theorem.) Conversely, if a theory T is c.e., then there is a computable set A of axioms such that T is the set of logical consequences of A.

Part (a) follows from the Gödel completeness theorem for first-order logic. The set of logical consequences of A is the same as the set of sentences derivable from A in the predicate calculus. If one has a machine that can generate the axioms, then one can organize a machine to generate the theorems.

Part (b) utilizes the simple fact that if one can generate the members of T in some order,

$$T = \{\tau_0, \tau_1, \tau_2, \dots\}$$

then one can generate a suitable set of axioms in increasing order:

$$A = \{\tau_0, \tau_0 \wedge \tau_1, \tau_0 \wedge \tau_1 \wedge \tau_2, \dots\}$$

So A is computable.

If one defines a theory T to be axiomatizable if there exists a computable set of axioms for it (or equivalently, if there exists a c.e. set of axioms for it), then there is the conclusion: A theory is axiomatizable iff it is c.e.

For example, the usual ZFC axioms for set theory form a computable set of axioms, so the set of theorems of ZFC is a c.e. set. At the other extreme, taking the set of axioms to be empty, one can conclude that the set of valid sentences is c.e. The set of valid sentences is, however, undecidable:

Church's theorem. Assume the language has at least one two-place predicate symbol. Then the set of valid sentences is not computable.

4. GÖDEL INCOMPLETENESS THEOREM

This section examines Gödel's first incompleteness theorem, from the point of view of computability theory. As the context, first-order theories of arithmetic, that is, theories dealing with the natural numbers with the operations of addition and multiplication, will be considered. Certainly, the study of the natural numbers with addition and multiplication is a basic part of mathematics, in the real sense that it is the topic in mathematics that school children study first.

The structure that is focused on here

$$\mathfrak{N} = (\mathbb{N}; 0, S, +, \times)$$

consists of the set $\mathbb{N}$ of natural numbers with the distinguished element 0 and the operations of successor (S), addition (+), and multiplication (×). The first-order language corresponding to this structure has quantifiers ∀ and ∃ ranging over $\mathbb{N}$, a constant symbol **0** for the number 0, and function symbols **S**, +, and × for successor, addition, and multiplication.

The set of all sentences of this language that are true in standard structure $\mathfrak{N}$ will be called the theory of true arithmetic. Although this theory deals with basic topics, it is by no means trivial. For example, it is not hard to see that the set of prime numbers is definable in $\mathfrak{N}$, that is, one can write down a formula $\pi(x)$ of the language that is satisfied in $\mathfrak{N}$ when the number n is assigned to x iff n is prime:

$$\begin{aligned} n \text{ is prime} &\Leftrightarrow \models_{\mathfrak{N}} \pi[n] \\ &\Leftrightarrow \models_{\mathfrak{N}} \pi(\boldsymbol{n}) \end{aligned}$$

where one substitutes for x the numeral $\boldsymbol{n}$ for n, that is the numeral **SS···S0**. Using this formula π one can then write

down a sentence in the language that expresses the twin prime conjecture, or a sentence that expresses Goldbach's conjecture. But the truth or falsity of these conjectures remains unknown.

What can one say quantitatively about the complexity of the theory of true arithmetic? It will be seen in this section that the theory is not c.e. and hence is not an axiomatizable theory. One connection between $\mathfrak{N}$ and computability is expressed by the result:

Theorem. Every computable relation over $\mathbb{N}$ is definable in the structure $\mathfrak{N}$. That is, for each computable k-ary relation $R \subseteq \mathbb{N}^k$ there is a formula $_\rho$ defining R in $\mathfrak{N}$:

$$\langle n_1, \ldots, n_k \rangle \in R \Leftrightarrow \vDash_{\mathfrak{N}} \rho[n_1, \ldots, n_k]$$

As an immediate consequence of this theorem, one can conclude that c.e. relations are also definable in $\mathfrak{N}$. This is because any c.e. relation Q is the domain of some computable relation R:

$$m \in Q \Leftrightarrow \langle m, n \rangle \in R, \text{ for some } n$$

Thus, if $\rho(x, y)$ defines R in $\mathfrak{N}$, then $\exists y \rho(x, y)$ defines Q. Moreover, $\neg\exists y \rho(x, y)$ defines the complement $\overline{Q}$ of Q. And by repeating the previous argument, the domain of $\overline{Q}$ is definable.

The conclusion is that any relation over $\mathbb{N}$ that is obtainable from the computable relations by the operations of forming the domain (i.e., projection) and forming the complement, iterated any number of times, will be definable in the structure $\mathfrak{N}$. (The converse is also true; these are exactly the definable relations; see section 6.)

In particular, the set $\overline{K}$ is definable in $\mathfrak{N}$, where K is the c.e. but noncomputable set constructed earlier. That is, there is some formula $\kappa(x)$ that defines K, so that $\neg\kappa(x)$ defines $\overline{K}$ and

$$n \in \overline{K} \Leftrightarrow \neg\kappa(\boldsymbol{n}) \text{ is true in } \mathfrak{N}.$$

It follows from this, however, that the set of sentences (of the language) true in $\mathfrak{N}$ cannot be semidecidable, lest equivalence yield an effective procedure for recognizing membership in this. Thus, one comes to the conclusion that truth in arithmetic is not a c.e. concept:

Theorem. The set of sentences true in $\mathfrak{N}$ is not c.e.

(An elaboration of this argument would give Tarski's theorem: The set of sentences true in $\mathfrak{N}$, when converted to a set of natural numbers, is not definable in $\mathfrak{N}$.)

This theorem, with the previous section, asserts that the theory of true arithmetic is not axiomatizable. So any axiomatizable subtheory fails to give all of true arithmetic.

Gödel's incompleteness theorem. For any axiomatizable subtheory T of true arithmetic, one can find a true sentence that is not derivable in T.

In fact, here is how one can find that true, underivable sentence. Let

$$J = \{n \mid T \vdash \neg\kappa(\boldsymbol{n})\},$$

the set of numbers that T "knows" are in $\overline{K}$. Because T is axiomatizable (and hence c.e.), the set J is c.e., and so $J = W_j$ for some number j. Moreover, J is a subset of $\overline{K}$ so it cannot be all of that set; there is a number in $\overline{K}$ that is not in J. In fact, j is such a number.

That is, the sentence $\neg\kappa(\boldsymbol{j})$ is a true sentence (j is really in $\overline{K}$) that T does not prove (T does not know that $j \in \overline{K}$). Thus, the sentence $\neg\kappa(\boldsymbol{j})$ is a specific witness to T's incompleteness.

And what might this sentence $\neg\kappa(\boldsymbol{j})$ say? Interpreted in $\mathfrak{N}$ it speaks of numbers and their sums and products. One can give it a more interesting translation, however:

$$\begin{aligned} \neg\kappa(j) \ &\text{says } j \in \overline{K} \\ &\text{i.e., } j \notin W_j \\ &\text{i.e., } j \notin J \\ &\text{i.e., } T \nvdash \neg\kappa(j) \end{aligned}$$

That is, the witness (the true unprovable sentence) asserts, in a sense, its own unprovability!

The conclusion is that the computability theory approach to Gödel's incompleteness theorem, based on c.e. sets, is not so different from the more traditional approach, which uses a diagonal construction to produce a sentence asserting, in a sense, its own unprovability.

5. DEGREES OF UNSOLVABILITY

Some unsolvable problems are more unsolvable than others. To make sense of this idea, one can employ the concept of relative computability.

Consider a fixed set B of natural numbers. Then a partial function f should be considered effectively calculable relative to B if there is a procedure that computes f and is effective except that it is allowed to appeal to an "oracle" for B. An oracle for B can be thought of as a device that, given a number x, responds by saying whether or not x is in B.

Any of the formalizations of calculability given in section 1 can be augmented to incorporate such an oracle. For example, in the case of primitive recursive functions, one can simply add the characteristic function of B

as a new initial function. As before, the various formalizations give exactly the same class of partial functions. Thus, one may speak unambiguously of computability relative to a set B.

Of course, if B is a computable set, however, the computability relative to B is simply equivalent to computability. For a noncomputable set, however, some noncomputable functions will become computable relative to B (the characteristic function of B, for one).

The concept of relative computability was introduced by Turing in a 1939 paper. At first glance, it seems an odd concept, combining as it does the most constructive approach to functions (that of computability) with the least constructive approach (that of a magical oracle). It is to Turing's credit that he perceived the value of the concept.

For sets A and B of natural numbers, one can say that A is computable *in* B, or that A is Turing reducible to B (written $A \leq_T B$) if the characteristic function of A is computable relative to B. That is, saying that $A \leq_T B$ implies that membership in A is no harder to decide than is membership in B. The $\leq_T$ relation is transitive and is reflexive on $\mathcal{P}\mathbb{N}$ (i.e., it is a preordering). Informally, transitivity of $\leq_T$ corresponds to connecting machines in series. Consequently, the symmetric version

$$A \equiv_T B \Leftrightarrow A \leq_T B \text{ and } B \leq_T A$$

is an equivalence relation on $\mathcal{P}\mathbb{N}$, and $\leq_T$ gives a partial ordering of the equivalence classes. These equivalence classes are called degrees of unsolvability, or simply degrees.

There is a least degree **0**, the class of the computable sets. Each degree must be a countable collection of sets (because there are only countably many programs), and so there are $2^{\aleph_0}$ equivalence classes altogether. Any two degrees have a least upper bound. The earlier construction of a noncomputable set K can be relativized:

$$x \in B' \Leftrightarrow \varphi_x^B(x)\downarrow$$

(where φ_x^B is the partial function computed, relative to B, by the program e). Then the degree of B' is strictly larger (under $\leq_T$) than the degree of B; thus, there is no largest degree.

The set B' is called the jump of B. Thus, the jump operation can be applied to a set to obtain a set of higher degree, and this operation can be iterated:

$$B <_T B' <_T B'' <_T B''' <_T \ldots$$

The degrees are not linearly ordered. It is possible to construct simultaneously sets A and B in such a way as to sabotage each machine that might reduce one set to the other. In fact, much more is true; one can construct $2^{\aleph_0}$ degrees that are all incomparable to each other under the ordering.

One can define a degree to be c.e. if it contains a c.e. set. These degrees are of particular interest because they are the degrees of axiomatizable theories. The least degree **0** is the degree of the decidable theories. Earlier, a noncomputable c.e. set $K = \{x \mid \varphi_x(x)\downarrow\}$ was constructed. So the degree of K, denoted **0'**, is a c.e. degree greater than **0**. The halting problem for Turing machines (regarded as a set of integers) also has degree **0'**. It is not hard to show that **0'** is the largest c.e. degree: for every c.e. degree ***a*** one has ***a*** $\leq_T$ **0'**. (Thus, any c.e. set of degree **0'** is $\leq_T$-complete for c.e. sets, in the sense that every other c.e. set is computable in it.)

A number of undecidable axiomatizable theories turn out to have degree **0'**: the validities of predicate calculus (with at least a binary predicate symbol), first-order Peano arithmetic, ZF set theory (if consistent), and others.

In 1944 Emil Post raised the question whether there were any c.e. degrees other than **0** and **0'**. This question, which became known as Post's problem, was finally answered in 1956 (two years after Post's death), independently by Richard Friedberg (in his Harvard senior thesis) and by A. A. Mučnik (in the Soviet Union). They showed that intermediate c.e. degrees do indeed exist, and in great profusion. Gerald Sacks later showed that any countable partial ordering can be embedded—as a partial ordering—in the partial ordering of c.e. degrees.

Although the natural axiomatizable theories have turned out to have either degree **0** or degree **0'**, Solomon Feferman showed that every c.e. degree contains some axiomatizable theory.

There is a simpler way in which questions about membership in one set might be effectively reducible to questions about another set. One can define A to be many-one reducible to B (written $A \leq_m B$) if there is a total computable function f such that

$$x \in A \Leftrightarrow f(x) \in B$$

for all natural numbers x. The idea is that each question "$x \in A$?" about A is reduced by f to one question about B. Moreover, if there is such a reduction function f that is one to one, then one can say that A is one-one reducible to B (written $A \leq_1 B$). Clearly,

$$A \leq_1 B \Rightarrow A \leq_m B \Rightarrow A \leq_T B$$

and in general neither arrow can be reversed. Again, both $\leq_1$ and $\leq_m$ are preorders, so the corresponding symmetric relations

$$A \equiv_1 B \Leftrightarrow A \leq_1 B \;\&\; B \leq_1 A \text{ and } A \equiv_m B \Leftrightarrow A \leq_m B \;\&\; B \leq_m A$$

are equivalence relations on $\mathcal{P}\mathbb{N}$, and the equivalence classes (the one-one degrees and the many-one degrees) are partially ordered. John Myhill showed that if $A \equiv_1 B$, then there is a total computable permutation of $\mathbb{N}$ taking A onto B.

It is not hard to make a c.e. set that is $\leq_1$-complete for c.e. sets, that is, every c.e. set is one-one reducible to it. In fact, K is such a set.

6. DEFINABILITY IN ARITHMETIC

As in section 4, let

$$\mathfrak{N} = (\mathbb{N}; 0, S, +, \times)$$

be the standard structure for arithmetic, consisting of the set $\mathbb{N}$ of natural numbers with the distinguished element 0 and the operations of successor, addition, and multiplication. In this structure, what sets (or relations or functions) are definable by first-order formulas? In section 4 it was noted that every computable relation is definable in arithmetic, and section 5 used the fact that some non-computable sets (such as $\overline{K}$) are also definable. Now, one can approach the matter more systematically.

Say that a relation (on $\mathbb{N}$) is arithmetical if it is definable in $\mathfrak{N}$. Of course, only countably many relations can be arithmetical, because there are only countably many formulas. One wants to classify these relations, according to the quantifier depth of the defining formulas.

From section 2 it is known that a relation A is c.e. iff it is the domain of some computable relation R:

$$\vec{m} \in A \Leftrightarrow R(\vec{m}, n) \text{ for some } n.$$

If $\rho(x_1, \dots, x_k, y)$ defines R, then the formula $\exists y \rho(x_1, \dots, x_k, y)$ defines A, so that A is "one quantifier away" from being computable. One can say that such relations A are Σ_1. (Yuri Matijacevič showed in 1970 that in fact every c.e. relation is definable by an existential formula, that is, one of the form

$$\exists y_1 \cdots \exists y_1\, \rho(x_1, \dots, x_k, y_1, \dots, y_1)$$

where ρ is quantifier-free, but that fact is not needed here.)

Next, call a relation Σ_2 if it is definable by a formula

$$\exists y_1 \forall y_2\, \rho(x_1, \dots, x_k, y_1, y_2)$$

where ρ defines a computable relation. Call a reaction Σ_3 if it is definable by a formula

$$\exists y_1 \forall y_2 \exists y_3\, \rho(x_1, \dots, x_k, y_1, y_2, y_3)$$

where ρ defines a computable relation), and so forth.

The dual concept, where one reverses existential and universal quantifiers, gives the Π_k relations. That is, call a relation Π_1 if it is definable by a formula

$$\forall y\, \rho(x_1, \dots, x_k, y)$$

call it Π_2 if it is definable by a formula

$$\forall y_1 \exists y_2\, \rho(x_1, \dots, x_k, y_1, y_2),$$

call it Π_3 if it is definable by a formula

$$\forall y_1 \exists y_2 \forall y_3\, \rho(x_1, \dots, x_k, y_1, y_2, y_3),$$

and so forth, where in each case ρ defines a computable relation.

Then the Π_k relations are exactly the complements of the Σ_k relations. In effect, one is measuring how far away a relation is from decidability.

By adding vacuous quantifiers, one sees that any Σ_k relation is both Σ_{k+1} and Π_{k+1}. And every definable relation appears somewhere in this hierarchy, because it will be definable by a prenex formula, the quantifier-free part of which always defines a computable (in fact primitive recursive) relation.

For example, the set $\{e \mid \varphi_e \text{ is total}\}$ of programs of total functions is Π_2, because φ_e is total iff $\forall m\, \exists n\, T(e, m, n)$. By Kleene's theorem, a relation is computable iff it is both Σ_1 and Π_1. The set K is Σ_1 but not Π_1. And in analogy to this fact,one can, for each k, construct a set that is Σ_k but not Π_k. Thus, letting the noun Σ_k denote the collection of all Σ_k relations, one has proper inclusion in both of the chains

$$\Sigma_1 \subset \Sigma_2 \subset \Sigma_3 \subset \cdots$$

$$\Pi_1 \subset \Pi_2 \subset \Pi_3 \subset \cdots$$

and in both cases the union of the chains is exactly the class of arithmetical relations. One can say that these chains define the arithmetical hierarchy.

From the point of view of the arithmetical hierarchy, one can obtain Tarski's theorem that the theory of true arithmetic is not arithmetical:

Tarski's theorem. The set of sentences true in $\mathfrak{N}$, regarded as a set of numbers, is not definable in $\mathfrak{N}$.

Let T be the set of true sentences. It suffices to show, for each k, that T cannot be Σ_k. Let A be a set that is arithmetical but not Σ_k (as indicated, there is such a set, and one can even make it Π_k). Then A is definable by some formula $\alpha(x)$ and

$$n \in A \Leftrightarrow \alpha(\boldsymbol{n}) \in T$$

which shows that $A \leq_m T$ (where T has been identified with the corresponding set of numbers). That is, for some total computable function f (which substitutes numerals into α),

$$n \in A \Leftrightarrow f(n) \in T.$$

If, contrary to one's hopes, T were Σ_k, then the previous line would let one conclude that A is also Σ_k, which it is not.

There is also a connection between the arithmetical hierarchy and relative computability, as in section 5. The following result extends the fact that a relation is Σ_1 iff it is c.e.

Post's theorem. (a) A relation is Σ_2 iff it is c.e. in $\emptyset'$, the jump of the empty set.

(b) More generally, a relation is Σ_{k+1} iff it is c.e. in $\emptyset^{(k)}$, the kth jump of the empty set.

7. FEASIBLE COMPUTABILITY

Up to now, this entry has approached computability from the point of view that there should be no constraints on the time required for a particular computation, or on the amount of memory space that might be required. The result is that some total computable functions will take a long time to compute. If a function f grows rapidly, then for large x it will take a long time simply to generate the output $f(x)$. There are also, however, bounded functions that require a large amount of time.

To be more precise, suppose one adopts one of the formalizations from section 1 (any one will do), and one defines in a reasonable way the "number of steps" or the "time required" in a computation. (Manuel Blum converted the term *reasonable* into axioms for what a *complexity measure* should be.) Then Michael Rabin showed that for any total computable function h, no matter how fast it grows, one can find another total computable function f with range $\{0, 1\}$ such that for any program e for f (i.e., $f = \varphi_e$), the time required for e to compute $f(x)$ exceeds $h(x)$ for all but finitely many values of x. (The function f is constructed in stages, in such a way as to sabotage any fast program that might try to compute f.)

Is there a more restricted concept of "feasibly computable function" where the amount of time required does not grow beyond all reason, where the amount of time required is an amount that might actually be practical, at least when the input to the function is not absurdly large? To this vague question, an exact answer has been proposed.

Once can call a function f polynomial-time computable (or for short, P-time computable) if there is a program e for f and a polynomial p such that for every x, the program e computes $f(x)$ in no more than $p(|x|)$ steps, where $|x|$ is the length of x.

This definition requires some explanation and support. If f is a function over Σ^*, the set of words over a finite alphabet Σ, then of course $|x|$ is just the number of symbols in the word x. If f is a function over $\mathbb{N}$, then $|x|$ is the length of the numeral for x. (Here, one comes again to the fact that effective procedures work with numerals, not numbers.) So if one uses base-2 numerals for $\mathbb{N}$, then $|x|$ is about $\log_2 x$.

Moreover, there was vagueness about exactly how the number of steps in a computation was to be determined. Here the situation is encouraging: The class of P-time computable functions is the same, under the different reasonable ways of counting steps.

Back in sections 0 and 1 there was the encouraging fact that many different ways of formalizing the concept of effective calculability yielded exactly the same class of functions. As remarkable as that fact is, even more is true. The number of steps required by one formalization is bounded by a polynomial in the number of steps required by another. For example, there exists a polynomial p (of moderate degree) such that a computation by a Turing machine that requires n steps can be simulated by a loop-while program that requires no more than $p(n)$ steps. Consequently, the concept of a P-time computable function is robust: One can get the same class of functions, regardless of which formalization from section 1 is employed. To be sure, the degrees of the polynomials will vary somewhat, but the class of P-time functions is unchanged.

Encouraged by this result, and inspired in particular by 1971 work of Stephen Cook, people since the 1970s have come to regard the class of P-time functions as the correct formalization of the idea of functions for which computations are feasible, without totally impractical running times.

By analogy to Church's thesis, the statement that P-time computability corresponds to feasibly practical computability has come to be known as Cook's thesis or the Cook-Karp thesis. (The concept of P-time computability appeared as early as 1964 in work of Alan Cobham. Jack Edmunds in 1965 pointed out the good features of P-time algorithms. Richard Karp in 1972 extended Cook's work.)

So what are the P-time computable functions? They form a subclass of the primitive recursive functions. All the polynomial functions are P-time computable, as are some functions that grow faster than any polynomial. There is, however, a limit to the growth rate of P-time computable functions, imposed by the fact that printing an output symbol takes a step. That is, there is the following growth limitation property: If f is computable in time bounded by the polynomial p, then $|f(x)| \leq |x| + p(|x|)$. This prevents exponential functions from being P-time computable; there is not enough time to write down the result.

Often, P-time computability is presented in terms of acceptance of languages (i.e., sets of words). Where Σ is the finite alphabet in question, consider a language $L \subseteq \Sigma^*$. One can say that $L \in$ P if there is a program and a polynomial p such that whenever a word w is in L, then the program halts on input w (i.e., it "accepts" w) in no more than $p(|w|)$ steps, and whenever a word w is not in L, then the program never halts on input w (i.e., the program does not accept w). This is equivalent to saying that the characteristic function of L is P-time computable, because one can add to the program an alarm clock that rings after time $p(|w|)$. For example, it is now known that the set of prime numbers (as a set of words written in the usual base-10 notation) belongs to P.

Of course, if the characteristic function of L is P-time computable, then so is the characteristic function of its complement, $\overline{L}$. That is, P = co-P, where co-P is the collection of complements of languages in P.

Informally, L is in P if L is not only a decidable set of words, but moreover there is a fast decision procedure for P—one that can actually be implemented in a practical way. For example, finite graphs can be coded by words over a suitable finite alphabet. The set of two-colorable graphs (i.e., the set of graphs that can be properly colored with two colors) is in P, because it is fast to check that the graph has no cycles of odd length. The set of graphs with an Euler cycle is in P, because it is fast to check that the graph is connected and that every vertex has even degree.

What about three-colorable graphs or graphs with Hamiltonian cycles? Here, there are no known fast decision procedures, but there are weaker facts: Given a proper coloring with three colors, it is fast to verify that it is indeed a proper coloring. Given a Hamiltonian cycle, it is fast to verify that it is indeed Hamiltonian. Both three-colorable graphs and Hamiltonian graphs are examples of languages that belong to a class known as NP.

One way to define NP is to use nondeterministic Turing machines. (The acronym NP stands for *nondeterministic polynomial time.*) Back in section 1 the definition of a Turing machine demanded that a machine's table of quintuples be unambiguous, that is, that no two different quintuples have the same first two components. By simply omitting that demand, one can obtain the concept of a nondeterministic Turing machine. A computation of such a machine, at each step, is allowed to execute any quintuple that begins with its present state and the symbol being scanned. Then, one can say that $L \in$ NP if there is a nondeterministic Turing machine M and a polynomial p such that whenever a word w is in L, then *some* computation of M starting from input w halts in no more than $p(|w|)$ steps, and whenever a word w is not in L, then no computation of M starting from input w *ever* halts. (An accepting computation can be thought of as having made a number of lucky guesses.)

There is an equivalent, and somewhat more workable, characterization along the lines of Σ_1 definability: $L \in$ NP iff there is binary relation $R \in$ P and a polynomial p such that for every word w,

$$w \in L \Leftrightarrow \exists y[|y| \leq p(|w|) \text{ and } R(w, y)].$$

Another example of a language in NP is SAT, the set of satisfiable formulas of sentential logic. The truth-table method for determining whether a formula with n sentence symbols is satisfiable involves forming all 2^n lines of the formula's truth table and looking to see if there is a line making the formula true. This is not, however, a feasible algorithm, because 2^{80} microseconds greatly exceeds the age of the universe. If one (nondeterministically) guesses the correct line of the table, however, then one can quickly verify that the formula is true under that line.

There is a clear analogy between computable and c.e. sets on the one hand, and P and NP on the other hand. The computable sets are decidable; the sets in P are decidable by fast algorithms. And c.e. sets are one existential quantifier away from being computable; sets in NP are one existential quantifier away from being in P. Moreover, there are c.e. sets that are complete with respect to $\leq_m$; there are NP sets with a similar property. One can say that

L_1 is P-time reducible to L_2 if there is a P-time computable (total) function f that many-one reduces L_1 to L_2. The following result was proved independently by Cook (1971) and Leonid Levin (1973):

Cook-Levin theorem. SAT is in NP, and every NP language is P-time reducible to SAT.

In other words, SAT is NP-complete. Karp showed that many other NP languages (three-colorable graphs, Hamiltonian graphs, and others) are NP-complete.

P VERSUS NP. How far does the analogy between NP and c.e. go? It is known that there are noncomputable c.e. sets, and a set is computable iff both it and its complement are c.e. While it is clear that $P \subseteq NP \cap$ co-NP (i.e., every language in P is in NP, as is its complement), it is not known whether P = NP, or if NP is closed under complement.

The diagonalization that produces a noncomputable c.e. set can be relativized in a straightforward way to show that for any fixed oracle B, there is a set B' that is c.e. in B but not computable in B. Might some diagonal argument produce a set in NP that was not in P? Would that argument then relativize? The definitions of P and NP extend easily to P^B and NP^B, where the computations can query the oracle B (in one step).

In a 1975 paper, Theodore Baker, John Gill, and Robert Solovay showed that there are oracles B and C such that on the one hand $P^B = NP^B$ and on the other hand $P^C \neq NP^C$. This result suggests that the P versus NP question is difficult, because whatever argument might settle the question cannot relativize in a straightforward way. It has also been shown that if one chooses the oracle B at random (with respect to the natural probability measure on $\mathcal{P}\mathbb{N}$), then $P^B \neq NP^B$ with probability 1.

The P versus NP question remains the outstanding problem in theoretical computer science. In recognition of this fact, the Clay Mathematics Institute is offering a million-dollar prize for its solution.

8. ANALYTICAL HIERARCHY

The ideas in section 5 can be utilized to consider partial functions that take as input not only numbers (or words over a finite alphabet), but sets of numbers or, more generally, functions from $\mathbb{N}$ to $\mathbb{N}$. One can think of the computational procedure as being given a set or a function if it is given an oracle for it.

Let $\mathbb{N}^\mathbb{N}$ be the set of all total functions from $\mathbb{N}$ to $\mathbb{N}$. For a function α in $\mathbb{N}^\mathbb{N}$, a calculation can query an oracle for α by giving it a number n. The oracle then supplies (in one step) the number $\alpha(n)$. For example, the partial function whose value at α is the least n, if any, for which $\alpha(n) = 0$ is a computable partial function on $\mathbb{N}^\mathbb{N}$. One can broaden the concept of computability to include partial functions that take as inputs k numbers and l members of $\mathbb{N}^\mathbb{N}$, and produce numbers as outputs.

For definiteness, suppose that f is a partial function on $\mathbb{N} \times \mathbb{N}^\mathbb{N}$. Informally, f is effectively calculable if there exists an effective procedure that, when given a number x and an oracle for an α in $\mathbb{N}^\mathbb{N}$, eventually halts and returns the correct value $f(x, \alpha)$ if this is defined, and never halts if $f(x, \alpha)$ is undefined. As before, the various formalizations in section 1 all can be adapted to incorporate inputs from $\mathbb{N}^\mathbb{N}$. One can thereby obtain the concept of a computable partial function on $\mathbb{N} \times \mathbb{N}^\mathbb{N}$.

The basic results of section 2 can be adapted to this broader situation. An essential point is that any one computation takes finitely many steps before producing its output and so can make use of only finitely many values from the given oracles. To obtain a normal form theorem, one again needs to adopt a way of encoding an entire step-by-step history of a computation when a program e is given an input number x and an oracle α. It is natural to do this in such a way that, where y is the number encoding the history, the oracle is asked for values $\alpha(t)$ only for $t < y$. Let $\overline{\alpha}(y)$ be a number encoding the finite sequence $\alpha(0), \alpha(1), \ldots, \alpha(y-1)$ consisting of the first y values of α.

For the Kleene T-predicate, one now needs for $T(e, x, s, y)$ to say that e encodes a program, and y encodes the step-by-step history of the computation that program produces on input x, where the oracle supplies values according to the sequences coded by s. This is a decidable property of e, x, s, and y.

Normal form theorem. There is a 4-ary computable relation T on $\mathbb{N}$ and a total computable function U with the following property: For each computable partial function on $\mathbb{N} \times \mathbb{N}^\mathbb{N}$, there is a natural number e such that

$$f(x, \alpha) = U(\mu y T(e, x, \overline{\alpha}(y), y))$$

for every x and α.

Analogous results hold for partial computable functions with more arguments. It is interesting to note that because U and T have only natural numbers as arguments, computability on $\mathbb{N} \times \mathbb{N}^\mathbb{N}$ can be characterized in terms of computability on $\mathbb{N}$.

As before, one can define a subset of $\mathbb{N} \times \mathbb{N}^\mathbb{N}$ to be computable if its characteristic function is computable. Informally, this means that one has an effective decision procedure that, given $\langle x, \alpha \rangle$, decides whether or not it is in

the set. Of course, the decision procedure will be able to utilize only finitely much information about α before rendering a verdict. Because of this fact, any computable set will be both open and closed in the natural topology on $\mathbb{N} \times \mathbb{N}^{\mathbb{N}}$ (where $\mathbb{N}$ has the discrete topology and $\mathbb{N}^{\mathbb{N}}$ has the product topology).

Moreover, one can define the c.e. sets to be the ones whose partial characteristic function is a computable partial function. From the normal form theorem, it follows that if Q is c.e., then there is a computable ternary relation R on $\mathbb{N}$ such that

$$Q(x, \alpha) \Leftrightarrow \exists y R(x, \overline{\alpha}(y), y)$$

for every x and α. Any such set Q will be open in the natural topology.

In section 6 the connection between computability and definability in arithmetic was examined. The definable relations formed a hierarchy, where the place of relation in the hierarchy (Σ_k or Π_k) depended, roughly, on how many quantifiers away from being computable it was.

Now, one can extend those ideas to second-order definability in arithmetic, where besides quantifiers over $\mathbb{N}$, quantifiers over $\mathbb{N}^{\mathbb{N}}$ can be used. One can start with

$$\Sigma_0^1 = \Pi_0^1 = \text{the class of arithemetical relations.}$$

Furthermore, one can define $\sum_{k+1}^1$ to consist of relations definable from Π_k^1 relations by prefixing existential quantifiers over $\mathbb{N}^{\mathbb{N}}$. Similarly, one can define Π_{k+1}^1 to consist of relations definable from $\sum_k^1$ relations by prefixing universal quantifiers over $\mathbb{N}^{\mathbb{N}}$. Finally, one can define Δ_k^1 to be $\sum_k^1 \cap \Pi_k^1$.

As with the arithmetical hierarchy, one has proper inclusion in both of the chains

$$\textstyle\sum_1^1 \subset \sum_2^1 \subset \sum_3^1 \subset \cdots$$

$$\Pi_1^1 \subset \Pi_2^1 \subset \Pi_3^1 \subset \cdots$$

and in both cases the union of the chains is exactly the class of relations that are second-order definable in $\mathfrak{N}$. One can say that these chains define the analytical hierarchy.

See also Cantor, Georg; Church, Alonzo; Computationalism; Computing Machines; First-Order Logic; Gödel, Kurt; Gödel's Theorem; Logic, History of; Machine Intelligence; Mathematics, Foundations of; Modern Logic; Peano, Giuseppe; Tarski, Alfred; Turing, Alan M.

Bibliography

Barwise, Jon, ed. *Handbook of Mathematical Logic*. Amsterdam, Netherlands: North-Holland, 1977. Part C, "Recursion Theory," comprises eight papers on computability theory.

Church, Alonzo. "An Unsolvable Problem of Elementary Number Theory." *American Journal of Mathematics* 58 (1936): 345–363.

Davis, Martin, ed. *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems, and Computable Functions*. Hewlett, NY: Raven Press, 1965. A collection of fundamental papers by Kurt Gödel, Alonzo Church, Alan M. Turing, J. B. Rosser, Stephen Cole Kleene, and Emil Post.

Griffor, Edward R., ed. *Handbook of Computability Theory*. Amsterdam, Netherlands: Elsevier, 1999. A collection of eighteen papers exploring topics in computability theory.

Herken, Rolf, ed. *The Universal Turing Machine: A Half-Century Survey*. New York: Oxford University Press, 1988.

Kleene, Stephen Cole. *Introduction to Metamathematics*. Amsterdam, Netherlands: North-Holland, 1952.

Rogers, Hartley. *Theory of Recursive Functions and Effective Computability*. New York: McGraw-Hill, 1967.

Soare, Robert I. *Recursively Enumerable Sets and Degrees: A Study of Computable Functions and Computably Generated Sets*. Berlin: Springer-Verlag, 1987.

Turning, A.M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42 (1936–1937): 230–265. A correction, 43 (1937): 544–546.

Herbert B. Enderton (2005)

COMPUTATIONALISM

Computer science has been notably successful in building devices capable of performing sophisticated intellectual tasks. Impressed by these successes, many philosophers of mind have embraced a computational account of the mind. *Computationalism*, as this view is called, is committed to the literal truth of the claim that the mind is a computer: Mental states, processes, and events are computational states, processes, and events.

THE BASIC IDEA

What exactly are computational states, processes, and events? Most generally, a physical system, such as the human brain, implements a computation if the causal structure of the system—at a suitable level of description—mirrors the formal structure of the computation. This requires a one to one mapping of formal states of the computation to physical state-types of the system. The mapping from formal state-types to physical state types can be called an *interpretation function I*. *I* allows a

sequence of physical state-transitions to be seen as a computation.

An example should make the central idea clear. A physical system computes the addition function if there exists a one-to-one mapping from numbers to physical state types of the system such that any numbers *n*, *m*, and *n*+*m* related as addends and sums are mapped to physical state types related by a causal state-transition relation. In other words, whenever the system goes into the physical state specified under the mapping as *n*, and then goes into the physical state specified under the mapping as *m*, it is caused to go into the physical state specified under the mapping as *n*+*m*.

Traditionally, computational processes have been understood as rule-governed manipulations of internal symbols or representations—what computer scientists call *data structures*. Though these representations typically have meaning or semantic content, the rules apply to them solely in virtue of their structural properties, in the same way that the truth-preserving rules of formal logic apply to the syntax or formal character of natural language sentences, irrespective of their semantic content. Computationalism thus construes thinking as a type of mechanical theorem-proving.

Computationalism has been the predominant paradigm in cognitive psychology since the demise of behaviorism in the early 1960s. The failure of behaviorism can be traced in no small part to its refusal to consider the inner causes of behavior—in particular, the capacity of intelligent organisms to represent their environment and use their representations in controlling and modulating their interactions with the environment. Computationalism avoids this failing, explaining intelligent behavior as the product of internal computational processes that manipulate (construct, store, retrieve, and so on) symbolic representations of the organism's environment.

Many philosophers of mind find computationalism attractive for two reasons. First, it promises a physicalistic account of mind; specifically, it promises to explain mental phenomena without positing any mysterious nonphysical substances, properties, or events. Computational states are physically realized in the computer; they are just the physical states specified by the mapping *I*. Computational operations, as noted, are purely mechanical, applying to the objects in their domain—typically, symbols—in virtue of their structural properties. Moreover, computationalism, if true, would show how it is possible for mental states to have both causal and representational properties—to function as the causes of behavior, and to be about things other than themselves. Mental states, on this view, are relations to internal symbols, and symbols have a dual character: They are both physically constituted, hence causally efficacious, and bearers of meaning.

A second reason why philosophers of mind find computationalism attractive is that it promises a nonreductive account of the mind. A serious problem with reductive physicalist programs—such as central state identity theory—is that they are overly chauvinistic; they seek to identify mental state-types, such as pain, with their specific physical (i.e. neural) realization in humans, thus denying mentality to systems that lack human physiology. Functionalists, by contrast, take it to be a contingent fact about mental states and processes that they are realized in neural matter; these same mental processes may, in other creatures or devices, be realized in other ways (e.g., in a silicon-based circuitry). According to functionalism, it is the causal organization of a system—rather than its intrinsic physical makeup—that is crucial to its mentality. Computationalism endorses—and affords a precise specification of—the basic idea of functionalism. The computational characterization given by *I* provides an abstract characterization of that causal organization for a given system. Computational explanation is itself a species of functional explanation; it provides an analysis of a cognitive capacity in terms of the organized interaction of distinct components of the system, which are themselves functionally characterized—that is, described abstractly in terms of what they do rather than what they are made of.

A commitment to computationalism by philosophers of mind has frequently taken the form of a commitment to a computational construal of the *Representational Theory of Mind* (hereafter, RTM-C), which is an account of propositional attitudes, such states as beliefs, desires, hopes, fears, and so on. According to RTM-C, propositional attitudes are relations to internal representations—for example, to believe that P is to bear a certain relation to a token of an internal representation that means that P. Each attitude type is construed as a distinct computationally characterizable relation to an internal representation; thus, believing is one type of computational relation, and desiring another. The RTM-C has been advertised, by, for example, Jerry Fodor in *Psychosemantics* (1987), as a scientific vindication of the commonsense practice of explaining a subject's behavior by appealing to his or her propositional attitudes. If true, it would underwrite the practice of individuating propositional attitudes along two distinct dimensions, by *attitude* and by *content*. Subjects can hold various attitudes toward a single proposition; they may believe, doubt, or

fear that the conflict in the Middle East will never by resolved. And subjects bear the same relation—belief, say—to many different propositions. On the RTM-C, the various attitudes correspond to distinct computational operations, and distinct data structure-types over which these operations are defined have distinct contents. The transparency of the relation between the commonsense explanatory scheme and the underlying computational realization of human psychology is an attractive feature of the view. However, it may also seem rather surprising that the two explanatory structures are virtually isomorphic. (Imagine if commonsense physics had anticipated the basic explanatory structure of quantum physics.)

The RTM-C is a heavily committed empirical hypothesis about the nature of the mind. Unlike computationalism, which claims simply that the mind is a computer, RTM-C purports to specify in broad outline the computational architecture of the mental processes that produce behavior. Computationalism is therefore compatible with the falsity of RTM-C. Is there any reason to believe RTM-C? It has been claimed that existing work in computational cognitive science provides empirical support for the RTM-C. Fodor (1987) points out that computational models of human cognitive capacities construe such capacities as involving the manipulation of internal representations. Psycholinguistic theories, for example, explain linguistic processing as the construction and transformation of structural descriptions, or *parse trees*, of the public language sentence being processed. It should be noted, however, that in order for a psychological theory to provide support for the RTM-C—for the claim that to have an attitude A toward a proposition P is to bear a computational relation R to a internal structure that means that P—it is not sufficient that the theory posits computational operations defined over internal representations. The posited representations must have appropriate contents—in particular, they must be interpreted in the theory as the contents of attitudes that one is prepared to ascribe to subjects independently of any commitment to the RTM.

For example, consider a psycholinguistic theory that explains a subject's understanding of the sentence "the dog bit the boy" as involving the construction of a parse tree exhibiting the constituent structure of the sentence. The theory supports the RTM only if there are independent grounds for attributing to the subject the content ascribed to the parse tree. There may be grounds for attributing to the subject a belief in a certain distal state of affairs—that a specific dog bit a specific boy—but this is not the content ascribed to the parse tree by the psycholinguistic theory. The parse tree's content is not even of the right sort. It does not represent a distal state of affairs; it represents the constituent structure of the sentence comprehended. The psycholinguistic theory supports the RTM only if the subject has propositional attitudes about the grammatical constituents of the sentence, such things as noun phrases and determiners. Such attitudes may be attributed to subjects as a consequence of the acceptance of the RTM-C, but these attitudes would not provide independent empirical support for the view.

SOME GENERAL ISSUES

Developments in computer science in the 1980s, in particular, the construction of *connectionist* machines—devices capable of performing cognitive tasks but without fixed symbols over which their operations are defined—have necessitated a broadened understanding of computation. Connectionist processes are not naturally interpretable as manipulations of internal symbols or data structures. Rather, connectionist networks consist of units or nodes whose activation increases or decreases the activation of other units to which they are connected until the ensemble settles into a stable configuration. Because connectionist networks lack symbols, they lack the convenient "hooks" to which, in the more traditional *classical* models, semantic interpretations or meanings are attached. Semantic interpretations, in connectionist models, are assigned either to individual units (in localist networks), or, more commonly, to patterns of activation over an ensemble of units (in distributed networks). Therefore, representation in connectionist devices is, in one respect, not as straightforward as it is in classical devices, because it is not as transparent which states or structures of the device count as *representations*. But issues concerning how the interpreted internal states or structures acquire their meaning—in other words, how a given semantic interpretation is justified—are fundamentally the same for the two kinds of machines.

Connectionist networks have had some success modeling various learning tasks—most notably, pattern-recognition tasks. There is continuing discussion within cognitive science about whether connectionist models will succeed in providing adequate explanations of more complex human cognitive capacities without simply implementing a classical or symbol-based architecture. One issue, originally raised by Jerry Fodor and Zenon Pylyshyn (1988), has turned on whether connectionist networks have the resources to explain a putative property of thought—that cognitive capacities are systemati-

cally related. For example, subjects can think the thought *the dog bit the boy* only if they can think the thought *the boy bit the dog*. A classical explanation of *systematicity* would appeal to the constituent structure of representations over which the operations involved in these capacities are defined—these representations contain the same constituents, just differently arranged. Connectionists, of course, cannot provide this sort of explanation—their models do not contain structured representations of the sort that the explanation requires.

Whether systematicity constitutes a decisive reason to prefer classical over connectionist cognitive models depends on several unresolved issues: (1) how pervasive the phenomena really is. It is certainly not true generally that if one can entertain a proposition of the form aRb, then one can entertain bRa. One can think the thought *the boy parsed the sentence* but not *the sentence parsed the boy*; (2) whether classical cognitive models are in fact able to provide real explanations of the systematic relations holding among cognitive capacities, rather than simply a sketch of the form such explanations would take in classical models. A real explanation of the phenomena would require, at least, the specification of a compositional syntax for the internal system of representation, something that classicists have so far been unable to provide; and (3) whether connectionist models are in fact unable to explain the systematic relations that do hold among cognitive capacities.

While strong claims have been made on both sides of this dispute, the question remains open. If it turns out that the mind has a connectionist architecture, then it would be expected that a perspicuous account of this architecture would reveal many cognitive capacities to be systematically related. For example, a characterization of the state of the network that consists in an English speaker's understanding of the sentence "the dog bit the boy" would cite the activation levels of various nodes of the network. The subject's understanding of the sentence "the boy bit the dog" would, presumably, activate many of the same nodes, and the explanation for the systematic relation between these two states would appeal to a dynamical account of the network's state transitions. These nodes would not be *constituents* of the subject's thought(s), in the sense required by classical models. And yet the relation between the two thoughts would not be merely accidental but instead would be a lawful consequence of general features of the network's architecture.

Questions such as whether connectionist devices will prove capable of modeling a wide range of complex cognitive capacities, and whether the best explanation of human cognitive capacities will advert to connectionist or classical computational processes, are properly understood as issues within computationalism. It should be noted, however, that the majority of philosophers committed to computationalism tend to interpret computation in classical terms, claiming that mental processes are manipulations of symbols in an internal code or language of thought. (See Jerry Fodor's *The Language of Thought* [1975] for the most explicit account of this view.) For this reason, and for ease of exposition, this entry will continue to refer to computational processes as manipulations of internal representations.

Computationalism requires a *psychosemantics*—that is, an account of how the postulated internal representations (or, in connectionist devices, states of the network) acquire their meaning. In virtue of what fact does a particular data structure mean *snow is white* rather than *2+2=4*? The meanings of natural language sentences are fixed by public agreement, but internal symbols must acquire their meanings in some other way. Philosophers committed to computationalism (and, hence, typically to physicalism) have assumed that an appropriate semantics for the language of thought must respect a "naturalistic constraint," the requirement that the conditions for a mental representation's having a particular meaning must be specifiable in nonintentional and nonsemantic terms. There have been various proposals for a naturalistic semantics. *Information-based* theories identify the meaning of a mental representation with the cause of its tokening in certain specifiable circumstances. *Teleological* theories hold that the meaning of a mental representation is determined by its biological function, what it was selected for.

No proposal is without serious problems, and the difficulty of accounting for the possibility that thoughts can *misrepresent* is the most widely discussed. But the difficulty of specifying naturalistic conditions for mental representation does not undermine computationalism itself. Cognitive scientists engaged in the business of developing computational models of cognitive capacities seem little concerned with the naturalistic constraint, and their specifications of semantic interpretations for these models do not obviously respect it. (See Frances Egan [1995] for argument.) There is no reason to think that the physicalistic bona fides of computational models are thereby impugned.

SUCCESSES AND OBSTACLES

As a hypothesis about the nature of mind, computationalism is not uncontentious. Important aspects of the

mental have so far resisted computational analysis, and computational theorists have had little to say about the nature of conscious experience. While computers perform many intellectual tasks impressively, no one has succeeded in building a computer that can plausibly be said to feel pain or experience joy. It is possible that consciousness requires an explanation in terms of the biochemistry of the brain. In other words, the computational strategy of prescinding from the neural details of mental processes may mean that conscious phenomena will escape its explanatory net.

If conscious mental phenomena resist computational analysis, then the computational model of mind cannot be said to provide a general account of the human mind; however, the model may still provide the basis for a theory of those cognitive capacities that do not involve consciousness in any essential way. Cognitive psychologists have applied the computational model to the study of language processing, memory, vision, and motor activity, often with impressive results. Domain-specific processes—such as syntactic processing and early vision—have proved most amenable to computational analysis. It is likely that the information available to these processes is tightly constrained. So-called *modular* processes lend themselves to computational treatment precisely because they can be studied independently of the rest of the cognitive system. One does not need to know how the whole mind works to characterize the relatively simple interactions involved in these processes. The idea that perceptual processes are modular, at least up to a certain point, is well supported. Modular processes tend to be more reliable—they take account of information in the input before being influenced by the system's beliefs and expectations. This is especially important for the perception of novel input. And modular processes are faster—the process does not have to find and retrieve relevant information from memory for the processing to proceed. Ultimately, of course, perceptual processes will have to be integrated with the rest of the system if they are to provide a basis for reasoning, belief formation, and action.

Perceptual mechanisms, as characterized by computational accounts, typically rely on physical constraints—that is, on general information true of the subject's environment—to aid the recovery of perceptible properties of that environment. This information is assumed to be available only to the process in question—not stored in memory, and hence not available to the system for reasoning tasks. For example, the *structure from motion* visual mechanism, characterized by Shimon Ullman in *The Interpretation of Visual Motion* (1979), computes the structure of objects in the scene from information obtained from relative motion. The mechanism computes the unique rigid structure compatible with relatively minimal input data (three distinct views of four non-coplanar points in the object), in effect making use of the fact that objects are rigid in translation. Without the assumption of rigidity, more data is required to compute an object's shape. Whether or not Ullman's model accurately describes the human visual system, the general strategy of positing innate assumptions about the environment that simplify the processing is methodologically sound, given that perceptual mechanisms may be assumed to be adaptations to that environment.

Domain-general processes—such as decision making and rational revision of belief in response to new information—have so far resisted computational treatment. Their intractability is due in part to the fact that general constraints on the information that may be relevant to solutions are difficult, if not impossible, to specify. A system capable of passing the *Turing test*—the requirement that it convince an interlocutor that it is a person for a short period of time—must have access to a vast store of information about the world, about how agents typically interact with that world, and about the conventions governing conversation among agents in the world. All this information must be stored in the system's memory. At any point in the conversation, the system must be capable of bringing that information to bear on the selection of an appropriate response from the vast number of meaningful responses available to it. Human agents, of course, have no trouble doing this. The relevant information is somehow just there when it is needed. The task for the computational theorist is to characterize how this vast store of information is represented in the system's memory in such a way that relevant information can be accessed efficiently when needed. This formidable technical problem is known in the field of Artificial Intelligence (AI) as the *knowledge representation problem*.

A related problem, known in AI circles as the *frame problem*, concerns how a system is able to continuously update its knowledge store as the world around it changes. Every change has a large number of consequences. For example, the typing of the previous sentence on this author's computer requires the provision of a plausible example of the generalization just typed. It also changes the arrangement of subatomic particles in the room, yet it doesn't affect the Dow Jones industrial average or the price of crude oil. The author needs to keep track of some of these consequences, but most can and

should be ignored. How, then, does the author update her knowledge store to take account of just those changes that are relevant (for her) while ignoring the vast number that are not? Unless the frame problem can be solved, or otherwise sidestepped, computationalism has a slim chance of providing a general account of human cognitive capacities.

GENERAL OBJECTIONS

Opponents of computationalism have offered arguments purporting to show that the human mind cannot be a computer. One class of objection, typified by Roger Penrose's *The Emperor's New Mind* (1989), takes as its starting point Kurt Gödel's result that any formal system powerful enough to do arithmetic can yield a sentence that is undecidable—that is, a sentence such that neither it nor its negation is provable within the system. A human observer, the argument continues, can see that the undecidable sentence is true; therefore, the human's cognitive abilities outstrip that of the machine. For the argument to establish that human minds are not machines it would have to demonstrate that human cognitive abilities simultaneously transcend the limits of all machines. No version of the argument has succeeded in establishing this strong claim.

A second objection claims that any physical system, including a rock or a piece of cheese, may be described as computing any function, thus computationalism's claim that the human mind is a computer is utterly trivial. If everything is a computer, then computationalism reveals nothing interesting about the nature of mind. The following is John Searle's version of the argument in *The Rediscovery of the Mind* (1992). Recall that to characterize a physical system as a computer is to specify a mapping from formal states of a computation to physical state-types of the system. Take some arbitrary function, say the addition function, and some physical system, say a particular wall. Though the wall appears to be in a constant state, it is known that the wall is made up of atoms in continuous motion. Its physical state is constantly changing. The microphysical state of the wall at time t1 can be interpreted as *two* and its microphysical state at time t2 as *three*, and its microphysical state at time t3 as *five*. And similarly for other combinations of addends and sums. Under this interpretation the physical state transitions of the wall implement the addition function. The wall is an adder!

It is possible, in the above sense, to describe any physical system as computing any function. This does nothing, however, to damage computationalism's claim that the mind is a computer. There are significant differences between the interpretation function under which the wall is an adder, and the interpretation function under which a hand calculator is an adder. One important difference is that in order to know how to interpret the wall's states as sums one has to compute the addition function oneself. The triviality argument does point to an important task for theorists concerned with the foundations of computational cognitive science—namely, specifying the adequacy conditions on interpretation that allow a computational characterization of a physical system to be predictive and explanatory of the systems's behavioral capacities.

A third objection to computationalism has been made by John Searle in his 1980 article "Minds, Brains, and Programs." According to Searle's *Chinese Room* argument, genuine understanding cannot be a computational process. The manipulation of symbols according to rules that operate only on their structural properties is, according to Searle, a fundamentally unintelligent process. The argument, which many have found unconvincing, is formulated explicitly for classical computational models—yet if Searle is right it would apply to any mechanical model of the mind, and hence to connectionist models as well.

It is unlikely that a philosophical argument of the sort discussed in this section will prove computationalism false. Computationalism is a bold empirical hypothesis about the nature of mind that will be evaluated by the explanatory fruit it bears. There is reason for cautious optimism, though substantial progress needs to be made on some formidable technical issues before theorists of mind can be confident that computationalism is true.

See also Artificial Intelligence; Chinese Room Argument; Cognitive Science; Machine Intelligence; Psychology.

Bibliography

Copeland, Jack. *Artificial Intelligence: A Philosophical Introduction*. Oxford: Blackwells, 1993.

Cummins, Robert. *Meaning and Mental Representation*. Cambridge, MA: MIT Press, 1989.

Egan, Frances. "Computation and Content." *The Philosophical Review* 104 (1995): 443–459.

Fodor, Jerry. *The Language of Thought*. New York: Thomas Y. Crowell, 1975.

Fodor, Jerry. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press, 1987.

Fodor, Jerry, and Zenon Pylyshyn. "Connectionism and Cognitive Architecture: A Critical Analysis." *Cognition* 28 (1988): 3–71.

Penrose, Roger. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press, 1989.

Searle, John. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (1980): 417–424.

Searle, John. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press, 1992.

Sterelny, Kim. *The Representational Theory of Mind: An Introduction*. Oxford: Blackwells, 1990.

Ullman, Shimon. *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press, 1979.

Frances Egan (2005)

COMPUTATIONAL MODEL OF MIND

See *Computationalism*

COMPUTER ETHICS

Computer ethics is a branch of applied ethics that considers ethical issues raised or significantly amplified by computer technology. The field is sometimes referred to by other terms such as "cyberethics," "information ethics," "information communications technology ethics," "global information ethics," and "Internet ethics." But, whatever the field is called, the computer remains the essential technological feature. Although some computing technology, for example the abacus, is centuries old, computer ethics has developed as a philosophical field with the advent of modern, digital, electronic computing. Modern computing technology, which includes hardware, software, and networks, is highly flexible and powerful. Computers can be programmed and in some cases trained to perform a wide range of functions. Because of this logical malleability computers carry out numerous and diverse applications in society. Computer chips are ubiquitous. They are embedded in everyday items such as cars and clothing, toys and tools, and pets and people.

Communication that depends upon computer technology has grown dramatically through widespread use of the cell phones, global positioning systems, and the Internet. In the early twenty-first century, people in developed countries live in computationally revolutionized and informationally enriched environments. Because computing has become so integrated in society, computer ethics has expanded dramatically to issues involving most activities within society including education, law, business, government, and the military. Through its extensive growth computer ethics is a field of applied ethics that intersects and affects virtually all other branches of applied ethics.

Computer ethics is interesting philosophically, not merely because computing technology is widely used, but because the application of computing technology raises intriguing conceptual issues and serious ethical problems for society. This happens frequently because computers are logically malleable and can be configured to perform old tasks in new ways and to accomplish strikingly new tasks. When computing technology is deployed in novel ways, ethical guidelines for its use are frequently unclear or nonexistent. This creates policy vacuums that may be accompanied by conceptual confusions about how to understand the computerized situation adequately. Hence, computer ethics typically demands doing more than routinely applying ethical principles to ethical issues in computing. Rather computer ethics requires an analysis of the nature and impact of the computing technology and the corresponding formulation and justification of policies for the ethical use of such technology. Listing all of the subject matter of computer ethics would be difficult as the field continues to expand as the application of computing grows, but broadly speaking traditional areas of investigation and analysis include privacy, property, power, security, and professionalism.

Because computers rapidly store and search vast amounts of information, privacy has been an ongoing concern of computer ethics. Personal information in medical documents, criminal records, and credit histories is easily retrieved and transmitted to others electronically, and as a result individuals are vulnerable to the improper disclosure of sensitive information and to the introduction of unknown errors into their records. The threat to privacy has been increasing in part because computing technology enables an enormous amount of information gathering to occur in subtle and undetectable ways. Internet stores track purchases of individuals and place cookies on personal computers inconspicuously. Computerized cameras in satellites, public places, private establishments, and personal cell phones record without notice. Computers utilizing global-positioning satellites routinely track locations of vehicles. Spyware installed on computers surreptitiously surveils the computing activities of unsuspecting users. In general, personal information can be collected from many sources and potentially assembled in databases that can be further merged, matched, and mined to construct profiles of the lives of individuals. Many fear that the widespread use of computers to collect information is creating a panopticon

society in which too many details of individual lives are known by others, leaving people with dramatically reduced levels of privacy. Philosophical analyses of the nature of privacy, the policies to protect privacy, and the justifications for privacy are more important than ever.

Property is also a major issue within computer ethics. This has become increasingly important because of the significant growth in hardware and software and the computerization of many popular products including art, photos, music, movies, and games that are produced, transmitted, and portrayed using a digital format. Because digital information can be copied so easily and accurately, the extent to which digital products should be owned and protected is heavily debated. Some libertarians on this issue argue that "information wants to be free" and that traditional intellectual property restrictions should not apply. For instance, those in the open source software movement advocate licensing that permits the free redistribution of software and requires accessibility to a program's source code so that it can be tested and improved by others. Those who advocate the ownership of intellectual digital property argue that with ownership comes pride and profit incentive that will generate digital products that otherwise would never be produced.

Debates over the rights of ownership raise many difficult philosophical issues. What is it that is owned and how should it be protected? A computer disk itself does not have much value; it is the information on the disk that matters. Information seems to be nothing more than an idea and ideas are not normally given intellectual property protection. As an example, consider again computer programs. Computer programs are algorithmic and hence mathematical in nature. This suggests that computer programs, like the Pythagorean theorem, should not be owned at all. However, computer programs generally are fixed in a tangible medium and are lengthy, original human expressions. As such they are appropriately covered by copyright protection. Yet, in their operation on machines computer programs are often novel, useful, nonobvious processes and hence are properly patentable. How, or even whether, computer programs should be protected depends largely on one's philosophical analysis of the nature of computer programs and on a justification of protecting intellectual property.

The basic philosophical issues of computerized property extend well beyond computer programs to every product in digital form. A movie that costs millions of dollars to make can be copied at no significant cost. If a movie is copied illegally using the Internet, to what extent should various contributors be held accountable—the person downloading the copy, the person who maintains a directory on the Internet informing people where copies are located, the person who makes a digital version available for others to copy, the company that makes the software specifically designed to copy movies easily over the Internet, or the Internet service provider?

Computers can create and shift relationships of power. Because computers allow individuals to perform tasks more easily and to accomplish some activities that they could never do without them, those who have access to computers have access to power. As a consequence, an obvious social concern is the disparity in advantage of those who have access to computing, for example in school, over those who do not. Unequal distribution of power may require ethical countermeasures to ensure fairness. To what extent, for example, should disabled citizens be assured of equal access to computing technology? To some degree the Internet has helped to correct this imbalance of power and even shift power toward the individual. For a modest fee individuals can advertise personal items for sale on the Web to a large audience. Politicians who are not well connected to an established political group can run an Internet campaign to express their ideas and to solicit funds. Independent hotel operators can unite through an Internet reservation service to compete with the larger hotel chains.

But the Internet's ability to shift power to the individual allows one person to solicit children to arrange illicit sexual encounters, to send spam e-mails to millions of people, and to spread viruses and worms. Moreover, Internet power shifts can sometimes result in making the strong even stronger. Large corporations can outsource jobs to cheaper labor markets and dominant militaries can enhance their capabilities with computerized communication and weapons. These power shifts raise philosophical questions about what the new relationships should be One of the most important power questions is who should govern the Internet itself.

The issue of rights and responsibilities of individuals on the Internet is complex because the Internet that supports the Web is worldwide. Different countries have different laws and customs and therefore have different concerns about the Web. Any given country may have great difficulty enforcing its concerns with information coming and going beyond its borders. Consider differences with regard to free speech as just one example. France and Germany have been concerned about prohibiting hate speech. China has targeted political speech. In the United States the focus has been largely on controlling pornography over the Internet. Even within a coun-

try's borders free speech often raises perplexing conceptual issues. For instance, should pornography that utilizes virtual children be regulated differently than pornography displaying actual children? But, even assuming agreement on the law, how does a country stop or punish a violator of free speech on the Internet who is located in some remote location in the world? Should the law be change to accommodate the realities of the Internet?

Not surprisingly security is as a fundamental problem on the Internet. Computer users can act from a distance over networks and thereby can accomplish goals without being observed. Hackers can break into computers and remove or alter data without being detected. Ordinary citizens can use tools on the Web to gather information from public documents in order to steal the identities of others. Terrorists can disrupt entire networks that control vital resources such as the electric power grid. The lack of security on the Internet is reminiscent of Plato's story of the ring of Gyges that allowed a shepherd to act invisibly. Plato posed the question, Why should someone be just if he can get away with being unjust? Plato's question is not just an abstract theoretical issue given the availability of current computer technology. If an Internet user can act unjustly and get away with it, why should he or she not do it?

Many people who design and operate computing systems regard themselves as computing professionals. But, given that anyone, regardless of educational background, can be hired to do computing, what does it mean to claim that someone is a computing professional? To what standards, including ethical standards, should computing professionals adhere? Although several codes of ethics have been offered to clarify what duties and responsibilities computer professionals have, professional responsibility has been difficult to establish for at least two reasons. First, unlike medicine and law, the field does not have a tradition of professional qualifying examinations and licensing, and therefore enforcement of any code of ethics is difficult. Second, the nature of computing itself makes the assessment of responsibility difficult. Computer programs are often enormously complex, written by dozens of people, and incomprehensible to any one person. Moreover, such large computer programs are brittle in that a tiny, obscure error can shatter the performance of the entire system under certain conditions. To what extent should computing professionals be regarded as liable when such difficult to predict errors lead to major failures or even catastrophic results?

Although traditionally computer ethics has focused on the ethics of computing situations, a philosophically rich part of the field is computational ethics that considers the impact computing has or theoretically may have on ethics itself. Philosophical issues in this area include questions such as: In what ways can ethical decision making be properly assisted by computational methods? In principle, could a computer ever make appropriate ethical decisions? Could computer implants in humans enhance and possibly alter human values? And, could a computer, or perhaps a robot, ever have rights or moral responsibilities?

See also Applied Ethics; Power; Property; Rights.

Bibliography

Bynum, T. W., and S. Rogerson, eds. *Computer Ethics and Professional Responsibility*. Oxford: Basil Blackwell, 2004.

Johnson, D. G. *Computer Ethics*. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2000.

Johnson, D. G., and H. Nissenbaum, eds. *Computers, Ethics, and Social Value*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

Moor, J. H. "What Is Computer Ethics?" *Metaphilosophy* 16 (4) (1985): 266–275.

Spinello, R. A. *CyberEthics: Morality and Law in Cyberspace*. 2nd ed. Sudbury, MA: Jones and Bartlett, 2003.

Spinello, R. A., and H. T. Tavani, eds. *Readings in Cyberethics*. 2nd ed. Sudbury, MA: Jones and Bartlett, 2004.

Tavani, H. T. *Ethics and Technology: Ethical Issues in an Age of Information and Communication Technology*. Hoboken, NJ: John Wiley, 2004.

James H. Moor (1996, 2005)

COMPUTING MACHINES

Any thorough discussion of computing machines requires the examination of rigorous concepts of computation and is facilitated by the distinction between mathematical, symbolical, and physical computations. The delicate connection between the three kinds of computations and the underlying questions "What are machines?" and "When are they computing?" motivate an extensive theoretical and historical discussion. The relevant outcome of this discussion is formulated at the beginning of section 3.

The paradigm of the first kind of computation is given when a human calculator determines, by finitely many and mathematically meaningful steps, the values of number-theoretic functions for particular arguments. The informal concept of such effectively calculable functions is thought to be captured by Kurt Gödel's concept of general recursive functions. The latter notion was introduced in 1934 and arose in an intellectual context that

includes the contemporaneous development of David Hilbert's program as well as earlier steps toward modern logic and abstract mathematics.

Alan M. Turing and Emil Post initiated in 1936 a shift from mathematically meaningful steps to basic, not further analyzable ones that underlie mathematical computations. They investigated symbolic processes carried out by human calculators and proposed essentially the same model of symbolic computation that is mathematically presented now by a Turing machine. Turing took, however, an additional, most significant step: He devised a universal machine that can execute the program of any Turing machine, and he had it carry out the necessary symbolic operations. This construction allowed him to prove the effective unsolvability of the halting and decision problems.

The physical details of how a universal machine could actually be constructed did not matter for Turing's theoretical investigations in 1936, but obviously did when he was involved in designing and building an Automatic Computing Engine (ACE). In modern digital computers controlled physical processes are used to realize, efficiently, the stepwise operations of a universal machine. That seems to be true even for quantum computing. In analog computing physical processes are used in a different way. However, independently of the mode of the physical computation, the question can be raised, whether there are physical processes that are carried out by a computing machine, but do not fall under Turing computations.

1. MATHEMATICAL COMPUTATIONS

Human calculating provides a rich prehistory for the development of machines that can take over routine computational tasks. This prehistory points to the pervasive impact computing machines will develop: from the broadly intellectual and socioeconomic to the highly focused scientific. Before coming to the technological challenge of building machines that mimic processes on symbolic configurations, one has to address the problem of determining the nature of such processes and those aspects that are crucial for their machine implementation. After all, physical representations of the symbolic configurations are needed, and machines have to perform on them physical operations that correspond to the symbolic ones.

1.1. PREHISTORY. In the sixteenth and seventeenth centuries Wilhelm Schickard, Blaise Pascal, Gottfried Wilhelm Leibniz, and others constructed mechanical calculators to carry out basic arithmetical operations. The calculatory roots go back, however, not just to ancient Greece but also to Egypt and Mesopotamia; important developments took place also in China, India, and in many different parts of the world under Arab influence. It is no accident for the evolution of computing that algebra and algorithm etymologically come from the same Arab source: the title of a widely used book and the agnomen of its author (Muhammed ibn Musa al-Kwarizmi).

The construction of mechanical calculators is prima facie narrower than the development of other scientific tools and yet it was pursued as having, potentially, a much broader impact through the intimate connection of computing with mathematics and logic. That was clearly sensed and expressed with great expectations by Leibniz. Of course, there had been aids to computation in the form of neatly arranged configurations of pebbles, for example. Another efficient aid had been the Chinese abacus that allows, via a good representation of natural numbers, the human calculator to add, subtract, multiply, and divide through strictly local manipulations of beads. The configurations of the abacus serve as the representation of input, intermediate results, and output of the calculation; they are essentially aids to memory.

The difference between abacus-like devices and mechanical calculators (as developed by Schickard, Pascal, and Leibniz) is formulated in an illuminating way by Charles Babbage:

> Calculating machines comprise various pieces of mechanism for assisting the human mind in executing the operations of arithmetic. Some few of these perform the whole operation without any mental attention when once the given numbers have been put into the machine.
>
> Others require a moderate portion of mental attention: these latter are generally of much simpler construction than the former, and it may also be added, are less useful. (1864/1994, p. 30)

The abacus certainly requires a moderate portion of mental attention, whereas Babbage's difference engine is perfectly in line with the development of automatic computing machines. The difference engine was intended to determine the values of polynomials for given arguments by the method of finite differences; the results were to be printed by the machine and to create reliable tables useful for astronomy and navigation. The evolution of the difference engine brought to light the economic importance of computing and the consequent governmental

support of related research. (The British government sponsored Babbage's work; the Swedish government supported the work of Georg Edward Scheutz, who was inspired by a description of Babbage's machine and constructed a difference engine in 1834; an improved version was built between 1851 and 1853 with funds from the Swedish Academy.)

Babbage took later another important conceptual step when developing his analytical engine. He followed the lead of Joseph Marie Jacquard, who had used "cards with holes" as a means of programming a loom to weave intricate patterns. Babbage devised, but never fully constructed, a programmable computing machine with a rather modern organization. In chapter 8 of *Passages from the Life of a Philosopher*, Babbage writes after having described the process of the Jacquard loom:

> The analogy of the Analytical Engine with this well-known process is nearly perfect. The Analytical Engine consists of two parts:
>
> 1. The store in which all the variables to be operated upon, as well as all those quantities which have arisen from the result of other operations, are placed.
>
> 2. The mill into which the quantities about to be operated upon are always brought.Every formula which the Analytical Engine can be required to compute consists of certain algebraical operations to be performed upon given letters, and of certain other modifications depending on the numerical value assigned to those letters. (1864/1994, p. 89)

Evidently, *store* corresponds to the memory and *mill* to the central processing unit of a contemporary computer. The programming constructs in Babbage's design are of such a general character that, Robin Gandy (1980) asserts, the number theoretic functions that are Babbage calculable are precisely those that are Turing computable.

The generality of computational issues, beyond their connection with arithmetic and analysis, is emphasized through the algebraic treatment of logic in the hands of George Boole, Augustus De Morgan, Charles Sanders Peirce, and Ernst Schröder, among others. In this line of research the decision problem was formulated and considered as a central issue. Even in the traditional Aristotelian presentation of logic computational features were considered to be significant by Raymundus Lullus and, importantly, by Leibniz in his project of constructing a universal language and an appropriate calculus ratiocinator. A logical machine in that tradition was built by William Stanley Jevons and described in the *Proceedings of the Royal Society* for January 20, 1870. Finally, it should be mentioned that Gottlob Frege claimed in *Grundgesetze der Arithmetik* (1893) that in his logical system "inference is conducted like a calculation," but continued, "I do not mean this in a narrow sense, as if it were subject to an algorithm the same as … ordinary addition and multiplication, but only in the sense that there is an algorithm at all, i.e., a totality of rules which governs the transition from one sentence or from two sentences to a new one in such a way that nothing happens except in conformity with these rules."

Within mathematics at that time, Leopold Kronecker insisted on the decidability of mathematical notions and the calculability of functions. These logical and mathematical developments were joined in formal mathematics, when Hilbert exploited the effective metamathematical description of formal theories (in his consistency program) and shifted effectiveness requirements from mathematics to metamathematics. It is here that modern computability theory found its ultimate motivation through the emphasis of the decision problem (*Entscheidungsproblem*) in the Hilbert School and the systematic articulation of the significance of Gödel's incompleteness theorems; both issues required a rigorous mathematical concept of effective method or mechanical procedure. Though these issues could have been addressed in their formulation for symbolic configurations, it took a detour through the calculability of number theoretic functions to arrive at sharp mathematical notions.

1.2. UNIFORM CALCULABILITY. Richard Dedekind formulated in his 1888 essay "Was sind und was sollen die Zahlen?" the general concept of a primitive recursive function and proved that all these calculable functions can be made explicit in his logicist framework. Dedekind's idea for the proof was very abstract, namely, to show the existence of unique solutions for functional equations of the form

$$\psi(0) = \omega,$$

$$\psi(\varphi(n)) = \theta(\psi(n)),$$

where ω ?is an element of N, φ is the successor function, and θ an arbitrary given function from N to N. This general point recurs in the early 1920s, for example, in the work of Hilbert, Thoralf Skolem, and Jacques Herbrand. However, the existence of solutions is no longer to be guaranteed by abstract logicist or set theoretic considerations, but by the availability of suitable calculation procedures. Implicit in these discussions is the specification of the class PR of primitive recursive functions. Hilbert's 1925 essay, "On the Infinite," defines this class inductively,

in almost the standard contemporary form, by specifying initial functions and closing under the definitional schemas of composition and primitive recursion. One shows by an easy inductive argument that the values of primitive recursive functions can be determined by an effective procedure for any given argument. All primitive recursive functions are in this sense calculable, but there are calculable functions that are not primitive recursive. Hilbert discussed an example due to Wilhelm Ackermann prominently already in 1925.

Herbrand viewed the Ackermann function in 1931 as finitistically calculable. In his systems of arithmetic he considered different classes F of finitist functions for which recursion equations were available. The defining axioms for the elements in F had to satisfy in particular this calculability condition, which had to be proved by finitist means, "We must be able to show, by means of intuitionistic [i.e., finitist] proofs, that with these [defining] axioms it is possible to compute the value of the functions univocally for each specified system of values of their arguments" (letter to Gödel in *Gödel's Collected Works V*, p. 15). The issue of characterizing classes of finitistically calculable functions was crucial for Herbrand's reflections on Gödel's second incompleteness theorem and its impact on Hilbert's program. Inspired by Herbrand's formulation, Gödel defined in his Princeton lectures of 1934 the class of general recursive functions; its definition no longer depends on the problematic concept of finitist provability.

Gödel's class of functions includes all primitive recursive functions and those of the Ackermann type. Assume, Gödel suggests, you are given a finite sequence $\psi_l, \dots, \psi_k$ of known functions and a symbol ϕ for an unknown one. Then substitute these symbols "in one another in the most general fashions" *Gödel's Collected Works I*, p. 368) and equate certain pairs of the resulting expressions. If the selected set of functional equations has exactly one solution, consider ϕ as denoting a "recursive" function; the definition of general recursive functions is obtained by insisting on two restrictive conditions. The first stipulates a standard form of certain terms, whereas the second condition demands that for every 1-tuple $k_1, \dots, k_l$ there is exactly one m such that $\phi(k_1, \dots, k_l) = m$ is a derived equation. The set of derived equations is specified inductively. The basic clauses guarantee that all numerical instances of a given equation as well as all true equalities $\psi_{ij}(x_1, \dots, x_n) = m$ are derived equations. The rules that allow steps from already obtained equations to additional ones are formulated as follows:

(**R.1**) Replace occurrences of $\psi_{ij}(x_1, \dots, x_n)$ by m, if $\psi_{ij}(x_1, \dots, x_n) = m$ is a derived equation;

(**R.2**) Replace occurrences of $\phi(x_1, \dots, x_1)$ on the right-hand side of a derived equation by m, if $\phi(x_1, \dots, xl) =$ m is a derived equation.

Gödel emphasized two central features in his definition when comparing it to Herbrand's: first, the precise specification of mechanical rules for carrying out numerical computations in a uniform way; second, the formulation of the regularity condition requiring calculable functions to be total, but without insisting on a finitist proof of that fact.

1.3. NORMAL FORM AND THE μ-OPERATOR. Using Gödel's arithmetization technique to describe provability in the equational calculus Stephen Cole Kleene analyzed the class of general recursive functions in 1936. The uniform and effective generation of the derived equations allowed him to establish an important theorem that is called now Kleene's normal form theorem: *for every recursive function* φ *there are primitive recursive functions* ψ *and* ρ *such that* $\varphi(x_1, \dots, x_n)$ *equals* $\psi(\varepsilon y.\rho(x_1, \dots, x_n, y) = 0)$, *where for every n-tuple* $x_1, \dots, x_n$ *there is a y such that* $\rho(x_1, \dots, x_n, y) = 0$. The latter equation expresses that y is (the code of) a computation from the equations that define φ for the arguments $x_1, \dots, x_n$; $\varepsilon y.\rho(x_1, \dots, x_n, y) = 0$ provides the smallest y, such that $\rho(x_1, \dots, x_n, y) = 0$, if there is a y for the given arguments (it yields 0 otherwise). Finally, the function ψ considers the last equation in the given computation and determines the numerical value of the term on the r.h.s of that equation, which is a numeral and represents the value of φ for the given arguments $x_1, \dots, x_n$. This theorem, or rather its proof, is remarkable: it allows to establish equivalences of different formulations with great ease; what is needed for the proof is only that the inference or computation steps are all primitive recursive.

Hilbert and Paul Bernays had introduced in the first volume of their *Grundlagen der Mathematik* (1934) a μ-operator that functioned in just the way the ε-operator did for Kleene. The μ notation was adopted later by Kleene and is still being used in computability theory. Indeed, the μ-operator is at the heart of the definition of a new class of number theoretic functions, the so-called μ-*recursive functions*, and the normal form theorem is the crucial stepping stone in proving that this class of functions is co-extensional with that of Gödel's general recursive ones. The μ-recursive functions are specified inductively in the same way as the primitive recursive ones, except that a third closure condition is formulated:

if $\rho(x_1, \ldots , x_n, y)$ is μ-recursive and for every n-tuple $x_1, \ldots , x_n$ there is a y such that $\rho(x_1, \ldots , x_n, y) = 0$, then the function $\theta(x_1, \ldots , x_n)$ given by $\mu y.\rho(x_1, \ldots , x_n, y) = 0$ is also μ-recursive.

Gödel's concept characterized a class of calculable functions that contained all known effectively calculable functions. Footnote 3 of the Princeton lectures seems to express a form of Alonzo Church's thesis. In a letter to Martin Davis of February 15, 1965, Gödel rejected that interpretation, "The conjecture stated there only refers to the equivalence of 'finite (computation) procedure' and 'recursive procedure.' However, I was, at the time of these lectures, not at all convinced that my concept of recursion comprises all possible recursions; and in fact the equivalence between my definition and Kleene's … is not quite trivial" (Davis 1982, p. 8). At that time in early 1934 Gödel was equally unconvinced by Church's proposal that effective calculability should be identified with λ-definability; he called the proposal "thoroughly unsatisfactory." That was reported by Church to Kleene on November 29, 1935. In the following year Gödel observed the absoluteness of general recursive functions: If the value of a general recursive function can be computed in a finite or even transfinite type extension of arithmetic, then it can be computed already in arithmetic. Gödel added this important observation to his 1936 paper *On the Length of Proofs* and viewed it as providing evidence that an important and stable class of functions had been isolated. The next section presents considerations of some of the pioneers, obviously including Gödel, as to their reasons why the mathematically rigorous notion of machine computation introduced by Turing, and not general recursiveness, was ultimately viewed as providing the correct concept of mechanical procedure.

2. CONCEPTUAL ANALYSIS

Returning to the beginning, one notices a shift from effective mathematical calculations to mechanical operations of a machine. Church maintained in 1935 that the former are properly captured by the calculations involving general recursive functions. Clearly, if one has appropriate machines that allow the calculation of the base functions and mimic composition, recursion, and minimization, then all recursive functions and thus all effectively calculable ones are seen to be machine computable. Gödel argued in exactly that way in his 193? paper and drew broader conclusions. He asserted that the characteristics of his equational calculus "are exactly those that give the correct definition of a computable function." He expanded that assertion by, "That this really is the correct definition of mechanical computability was established beyond any doubt by Turing" (*Gödel's Collected Works III*, p. 168). The equivalence between general recursiveness and Turing computability is taken to support this claim.

2.1. CHURCH'S THESIS.

Almost a year after his conversation with Gödel, Church came back to his proposal in a letter to Bernays dated January 23, 1935; he conjectured that the λ-calculus may be a system that allows the representability of all constructively defined functions. When Church wrote this letter, he knew that all general recursive functions are λ-definable; the converse was established in collaboration with Kleene by March 1935. This mathematical equivalence and the quasi-empirical adequacy of λ-definability provided the background for the public articulation of Church's thesis. Church announced it in a talk contributed to the meeting of the American Mathematical Society in New York City on April 19, 1935, but formulated it with general recursiveness, not λ-definability as the mathematically rigorous notion.

In his 1936 paper Church restated his proposal for identifying the class of effectively calculable functions with a precisely defined class. To give a deeper analysis Church discussed, in section 7 of his paper, two methods of characterizing the effective calculability of number-theoretic functions. The first of these methods uses the notion of algorithm, and the second employs the notion of calculability in a logic. He argues that neither method leads to a definition that is more general than recursiveness. These arguments have a parallel structure, and this entry discusses only the one pertaining to the second method. Church considers a logic **L**, that is a system of symbolic logic whose language contains the equality symbol =, a symbol { }( ) for the application of a unary function symbol to its argument, and numerals for the positive integers. He defines, "F is *effectively calculable* if and only if there is an expression f in the logic **L** such that: $\{f\}(\mu) = \nu$ is a theorem of **L** iff F(m) = n; here, μ and ν are expressions that stand for the positive integers m and n." Church claims that F is recursive, assuming that **L** satisfies a certain step condition that amounts to requiring the theorem predicate of **L** to be recursively enumerable. The claim follows immediately by an application of the μ-operator; the argument parallels that for Kleene's normal form theorem.

The general concept of calculability is thus explicated by that of derivability in a logic, and Church uses the step condition to sharpen the idea that within such a logical formalism one operates with an effective notion of immediate consequence. The thesis is thus appealed to

only in a special case. Given the crucial role this condition plays, it is appropriate to view it as a normative requirement: The steps of any effective procedure (governing derivations of a symbolic logic) must be recursive. If this requirement is accepted and a function is defined to be effectively calculable as above, then Church's step-by-step argument proves that all effectively calculable functions are recursive.

Church gave two reasons for the thesis, namely, (1) the quasi-empirical observation that all known calculable functions are general recursive and (2) the mathematical fact of the equivalence of two differently motivated notions. A third reason comes directly from the 1936 paper, the step-by-step argument from a core conception. However, Church and Gödel found in the end Turing's machine model of computation much more convincing. Church's 1937 review of Turing's paper for the *Journal of Symbolic Logic* asserts that Turing computability has the advantage over general recursiveness and λ-definability of "making the identification with effectiveness in the ordinary (not explicitly defined) sense evident immediately" (pp. 42–43)

2.2. **FINITE MACHINES**. Church's more detailed argument for the immediate evidence starts out as follows:

> The author [Turing] proposes as a criterion that an infinite sequence of digits 0 and 1 be "computable" that it shall be possible to devise a computing machine, occupying a finite space and with working parts of finite size, which will write down the sequence to any desired number of terms if allowed to run for a sufficiently long time. As a matter of convenience, certain further restrictions are imposed on the character of the machine, but these are of such a nature as obviously to cause no loss of generality—in particular, a human calculator, provided with pencil and paper and explicit instructions, can be regarded as a kind of Turing machine.

He then draws the conclusion, "It is thus immediately clear that computability, so defined, can be identified with … the notion of effectiveness as it appears in certain mathematical problems" (pp. 42–43). Why Turing's notion should convey this immediate conviction Church does not explain; the step from a computing machine "occupying a finite space and with working parts of finite size" to Turing machines is not deepened.

Gödel commented on Turing's notion in his 1951 Gibbs lecture publicly for the first time and made remarks similar to Church's. He explores there the implications of the incompleteness theorems, not in their original formulation, but in a "much more satisfactory form" that is "due to the work of various mathematicians." He stresses, "The greatest improvement was made possible through the precise definition of the concept of finite procedure, which plays such a decisive role in these results" (*Gödel's Collected Works III*, p. 304). There are, Gödel points out, different ways of arriving at a precise definition of finite procedure, which all lead to exactly the same concept.

However, and here is Gödel's substantive remark on Turing, "The most satisfactory way … [of arriving at such a definition] is that of reducing the concept of finite procedure to that of a machine with a finite number of parts, as has been done by the British mathematician Turing" (*Gödel's Collected Works*, pp. 304–305). Gödel does not expand on this brief remark. In particular, he gives no hint of how reduction is to be understood or why the concept of such a restricted machine is equivalent to that of a Turing machine. At this point, it seems, the ultimate justification lies in the pure and perhaps rather crude fact that finite procedures can be reduced to computations of finite machines.

In a deep sense, neither Church nor Gödel seem to have recognized the distinctive character of Turing's analysis, that is, the move from arithmetically motivated calculations to general symbolic processes that underlie them. Most importantly in the given intellectual context, these processes have to be carried out programmatically by human beings: the *Entscheidungsproblem* had to be solved by humans in a mechanical way; it was the normative demand of radical intersubjectivity between humans that motivated the step from axiomatic to formal systems. For this reason Turing brings in human computers and exploits the limitations of their processing capacities, when proceeding mechanically. The Turing machine is in the end nothing but, as Gandy (1980) puts it, a codification of the human computer.

2.3. **COMPUTORS**. One can call a human computing agent who proceeds mechanically a *computor*; such a computor operates on finite configurations of symbols and, for Turing, deterministically so. At issue is then, how does one step from calculations of computors to computations of Turing machines? Turing explores, as he put it, the extent of the computable numbers (or, equivalently, of the effectively calculable functions) by considering two-dimensional calculations in a child's arithmetic book. Such calculations are reduced to symbolic steps on linear configurations of such a simple character that a

Turing machine operating on strings (instead of letters) can carry them out. Turing's argument concludes, "We may now construct a machine to do the work of the computer. … The machines just described [string machines] do not differ very essentially from computing machines as defined in §2 [letter machines], and corresponding to any machine of this type a computing machine can be constructed to compute the same sequence, that is to say the sequence computed by the computer" (*The Undecidable*, p. 138).

It is important to recall Turing's goal of isolating the basic steps of computations, that is, steps that need not be further subdivided. This leads to the demand that the configurations, which are operated on, must be immediately recognizable by the computor. Combined with the evident limitation of the computor's sensory apparatus, this demand motivates convincingly two restrictive conditions:

(**B**) (*Boundedeness*) There is a fixed finite bound on the number of configurations a computor can immediately recognize

(**L**) (*Locality*) A computor can change only immediately recognizable (sub-) configurations

Turing's considerations, sketched earlier, lead rigorously from operations of a computor on linear configurations to operations of a letter machine and can be generalized to other syntactic or graphic configurations. It should be noted that these constraints apply to Turing machines, but are violated by Gödel's equational calculus, as the replacement operations naturally involve terms of arbitrary complexity.

Turing's analysis secures the generality of mathematical results (e.g., of the incompleteness theorems) and their conclusiveness (e.g., of the undecidability of predicate logic) by respecting the intellectual context that appealed to effective operations carried out by humans without invoking higher mental capacities. It was after all the decision problem, the *Entscheidungsproblem* in the title of Turing's 1936 article, that motivated Turing's work. Its positive solution required "a procedure … that permits—for a given logical expression—to decide the validity, respectively satisfiability, by finitely many operations." Hilbert and Ackermann gave that formulation (pp. 72–73) in their book *Grundzüge der theoretischen Logik* (1928) and considered the decision problem as the main problem of mathematical logic. Why that problem should be considered as the main problem of mathematical logic is stated clearly in their remark, "The solution of this general decision problem would allow us to decide, at least in principle, the provability or unprovability of an arbitrary mathematical statement" (p. 86). Taking for granted the finite axiomatizability of set theory or some other fundamental theory in first-order logic, the general decision problem is solved when that for first-order logic has been solved.

A negative solution of the decision problem required, however, a rigorous characterization of finite procedures and a proof that none of them answers Hilbert and Ackermann's demand. Turing did both, as he gave a convincing conceptual analysis, established the effective unsolvability of the halting problem (or rather of the equivalent printing problem), and showed how to reduce it to the decision problem. Thus, if the latter were effectively solvable, then the halting problem would be; but as it is not, one has a contradiction. The proof of the unsolvability of the halting problem makes crucial use of a particular Turing machine, the universal machine U that, when presented on its tape with the program of a Turing machine M and an input, executes M's program for that input. This particular machine will play a special role in the next section.

3. PHYSICAL REALIZATION

For the further considerations, the most significant outcome of the previous historical and conceptual examination can be restated sharply as follows: Turing's notion of machine computation is obtained by an analysis of symbolic calculations carried out by computors. To put it negatively, Turing's notion is not obtained by an independent analysis of physical devices with the goal of, first, defining a general notion of machine and, second, introducing an appropriate concept of computation for such machines. It was only in 1980 that Gandy gave an analysis of machines and the deterministic computations they can carry out. This is presented in the second subsection below and will be followed, in the last subsection, by a description of the special features of quantum computers. However, what amounts to the physical implementation of Turing's universal machine U is discussed first. That is an absolutely central step in the development of modern computing machines and was taken in intricately intertwined ways, it seems, by Turing and John von Neumann; their work shaped the architecture of modern computers.

3.1. IMPLEMENTING U.

In the years following World War II Turing worked on various aspects of the design and the actual building of a practical version of his universal machine U. During the last three months of 1945

he wrote a remarkable document titled *Proposal for Development … of an Automatic Computing Engine (ACE)* and connected, in a lecture to the London Mathematical Society of February 20, 1947, the work on the ACE explicitly with his early theoretical work:

> Some years ago I was researching on what might now be described as an investigation of the theoretical possibilities and limitations of digital computing machines. I considered a type of machine which had a central mechanism, and an infinite memory which was contained on an infinite tape. … It was essential in these theoretical arguments that the memory should be infinite. It can easily be shown that otherwise the machine can only execute periodic operations. Machines such as the ACE may be regarded as practical versions of this same type of machine. (Turing 1947, pp. 106–107)

Turing characterized the ACE in his lecture as a typical large-scale electronic digital computing machine. From a mathematical perspective, Turing viewed being digital as the most relevant property of the ACE, since digital machines can work to any desired degree of accuracy and are not restricted, as analog machines are, to a particular type of computational problem.

From a practical point of view, the property of the ACE to be an electronic machine Turing considered as extremely important: it was to guarantee high speed and thus make it possible to execute complex procedures. The latter possibility requires, beyond the speed of basic operations, an appropriate organization of the machine, so that it can proceed fully automatically—without having to interact with a human operator—while executing a procedure. Turing emphasized, alluding to his universal machine:

> It is intended that the setting up of the machine for new problems shall be virtually only a matter of paper work. Besides the paper work nothing will have to be done except to prepare a pack of Hollerith cards in accordance with this paper work, and to pass them through a card reader connected with the machine. There will positively be no internal alterations to be made even if we wish suddenly to switch from calculating the energy levels of the neon atom to the enumeration of groups of order 720. It may appear somewhat puzzling that this can be done. How can one expect a machine to do all this multitudinous variety of things? The answer is that we should consider the machine as doing something quite simple, namely carrying out orders given to it in a standard form which it is able to understand. (Turing 1946, p. 21)

In the 1947 lecture he made the connection to the universal machine explicit; after discussing memory extensively, he claims that digital computing machines such as the ACE are just "practical versions of the universal machine." He continues, "There is a certain pool of electronic equipment, and a large memory. When any particular problem has to be handled the appropriate instructions for the computing process involved are stored in the memory of the ACE and it is then 'set up' for carrying out that process."

The requirements for building a universal machine can in the end only be satisfied, if the machine is not only digital and electronic but also large scale, as it involves demands for "storage of information or mechanical memory." Indeed, Turing pointed out already in the ACE Report that "the memory needs to be very large indeed." The principled as well as the practical issues of implementation overlapped at this point with developments in the United States. Indeed, Turing recommended reading his report "in conjunction with J. von Neumann's *Report on the EDVAC*." (Herman H. Goldstine [1972] and Andrew Hodges [1983] present complementary views on the tenuous connection between the two projects; a balanced perspective is given by Hodges [1983, pp. 555–556, note 5.26.)

von Neumann completed a first draft of his report on June 30, 1945; the report emerged out of work with the group of J. Presper Eckert and John Mauchly at the Moore School of Electrical Engineering (University of Pennsylvania, Philadelphia). The group had built one of the first electronic calculators, the Electronic Numerical Integrator and Computer (ENIAC) and was evaluating a new memory system for a second, more sophisticated calculator, the Electronic Discrete Variable Calculator (EDVAC). The demand for a large, readily accessible memory emerged out of computational practice, namely, the need to have fast access to instructions, but also to fixed constant parameters and statistical data. That was to be achieved by storing them in the machine; von Neumann writes, "The device requires a considerable memory. While it appeared, that various parts of this memory have to perform functions which differ somewhat in their nature and considerably in their purpose, it is nevertheless tempting to treat the entire memory as one organ" (Goldstine 1972, p. 194).

von Neumann shifted the attention from the technological problems of having a larger memory to logical ones concerning the basic structure of machines with a central control mechanism and extensive memory. This structure is discussed in detail by Goldstine (1972, pp. 204–210).

A higher level of generality was attained in the Electronic Computer Project at the Institute for Advanced Study at Princeton University; this project was begun in March 1946 and directed by von Neumann. The resulting IAS Computer can be viewed as a prototype of all modern computers (the "von Neumann machine"). Its basic architecture, however, is similar to that of the ACE; it is the balance between arithmetical and fundamental logical operations that is distinctive. Goldstine describes the issue as follows:

> The work of Post and Turing made it very clear that from the point of view of formal logics there was no problem to devise codes which were "in abstracto adequate to control and cause the execution of any sequence of operations which are individually available in the machine and which are, in their entirety, conceivable by the problem planner." The problem is of a practical nature and is closely allied to that connected with the choice of elementary operations in the arithmetic organ. (1972, p. 258)

Turing and von Neumann made different compromises between simplicity of basic machine operations and complexity of programs needed to execute mathematical or symbolic procedures. These choices were obviously informed not only by their different computational experience and goals but also by their broader philosophical outlook. (That is movingly described by Hodges [1983, pp. 320–333].)

3.2. **DISCRETE MACHINES**. Turing's U can be realized within practical limits by physical devices, and one can raise the question whether these devices are just doing things faster than humans can do, or whether they are in a principled way computationally more powerful. Church, as recalled earlier, asserted in 1937 that finite machines are essentially Turing machines; in Gödel's remarks (from 193? and 1951) that assertion is taken for granted. The claim seems to be plausible, but it does require an argument. On the one hand, there may be physical systems that do not obey the same restrictions as computors and consequently may be able to carry out computations not possible for a computor. On the other hand, there may be physically grounded limits for machines in the same way that there are psychologically based constraints for computors.

The character of individual computational steps was at the heart of the conceptual analysis. Because of physical constraints, such steps cannot be accelerated unboundedly or be made arbitrarily complex (Mundici and Sieg 1995, §3). However, there seems to be the possibility of sidestepping these constraints by usingwith massively parallel operations. Cellular automata, introduced by Stanislaw Ulam and von Neumann, operate in parallel; they do not satisfy the boundedness condition (**B**), as the configurations affected in a single computation step are potentially unbounded. They can simulate universal Turing machines and yield discrete simulations of complex physical processes. Konrad Zuse, for example, reflected on digital formulations of physics in his essay *Rechnender Raum* (1967). Edward Fredkin advocated the use of (reversible) cellular automata in physics and conjectured in his *Digital Mechanics* "that there will be found a single cellular automaton rule that models all of microscopic physics; and models it exactly" (1990, p. 254). The interested reader should consult Rolf Herken (1988), Tommaso Toffoli and Norman Margoulis (1987), and, of course, Stephen Wolfram (2002).

Gandy addresses the issue of parallel machine computations in his essay "Church's Thesis and Principles for Mechanisms" (1980), where he proposes a particular mathematical description of discrete mechanical devices and their computations. He then follows Turing's three steps of pertinent analysis, articulation of constraints, and proof of a reduction theorem. The central and novel aspect of Gandy's formulation lies in the fact that it incorporates parallelism in complete generality. Cellular automata fall directly under Gandy's formulation. And yet, the reduction theorem shows that everything calculable by a device satisfying the constraints, a Gandy machine, is already computable by a Turing machine. Here is a sketch of the main considerations.

Gandy (1980) introduces the term *discrete mechanical device* to make it vivid that his analysis is not at all concerned with analog devices, but rather with machines that are discrete and proceed step-by-step by step from one state to the next. Gandy considers two physical constraints as fundamental for such devices: (1) a lower bound on the size of atomic components and (2) an upper bound on the speed of signal propagation. Together, these constraints guarantee what the sensory limitations guarantee for computors, namely that in a given unit of time there is a bound on the number of different observable configurations and of possible actions

on them. However, the incorporation of massive parallelism into the mathematical description takes in Gandy's essay a substantial amount of complex mathematical work. In Wilfried Sieg's "Calculations by Man and Machine: Conceptual Analysis" (2002), Gandy machines are axiomatized as special discrete dynamical systems, and this presentation makes clear that they are radical generalizations of Turing machines: the latter modify one bounded part of a state, whereas the former operate in parallel on arbitrarily many bounded parts to arrive at the next state of the system.

Discrete computing machines in the broadest sense, when only constrained by physically motivated boundedness and locality conditions, do not reach beyond the computational power of Turing machines; that is the general moral. Every mathematical model of physical processes faces at least two questions: How accurately does the model capture physical reality, and how efficiently can the model be used to make predictions? It is distinctive for modern developments that, on the one hand, computer simulations have led to an emphasis of algorithmic aspects of scientific laws and that, on the other hand, many physical systems are being considered as computational devices, but under what conditions can a physical system really be viewed in that way? To have one important data point for reflections on this question, this entry will now look at the case of particular quantum systems.

3.3. **QUANTUM COMPUTERS**. Suppose one has a photon that impinges on a beam splitter and then propagates via two different paths. Quantum theory describes the photon as going partly into each of these two components. The state of the photon is given by the superposition of the two states associated with the two components of the original beam. Any observation of the photon, however, results in either the whole photon or nothing at all. This implies that after a measurement (1) the photon changes its state from being partly in one beam and partly in the other to being entirely in one of the beams, and (2) any interference effect is lost since one of the beams no longer enters into the description of the photon. If a second beam splitter combines the two beams, then the photon will be observed with probability one in a single beam. This certainty is because of quantum interference. Quantum computation arises from the possibility of exploiting a multiplicity of parallel computational paths in superposition as well as quantum interference to amplify the probability of correct outcomes of computations.

As the photon can be in a coherent superposition of being in two beams, the basic unit of quantum information, a *qubit* (from quantum bit), is a two-state system that can be prepared in a superposition of the two logical states 0 and 1. If a computational state can be reached through several alternative paths, then its probability is the squared modulus of the sum of all the probability amplitudes for the constituent paths. (Probability amplitudes determine probabilities and these have to add up to one for any quantum computational state.) Since the probability amplitudes are complex numbers, they may cancel each other and produce destructive interference or enhance each other and produce constructive interference.

Imagine a computation that starts in the input state 0 and reaches the output state in two steps. Suppose a computational step can mimic the action of a beam splitter and generate a superposition of two intermediate output states, 0 and 1 with probability amplitudes $c_0 = i/\sqrt{2}$ and $c_1 = 1/\sqrt{2}$. Then the probability of each output is the same: $|c_0|^2 = |c_1|^2 = 1/2$. However, if the output state is measured after two computational steps, then the probability of the output 1 is one: The action of a beam splitter can be perfectly simulated by quantum computing operations that have no classical analogs. One of these is the $\sqrt{NOT}$, which when applied twice results in the logical operation *NOT*.

Since quantum mechanics describes a state transformation by means of a unitary operator, any quantum computing operation is a unitary transformation on qubits. The description of a quantum Turing machine (*QTM*) is derived from a Turing machine, but using quantum theory to define the operations carried out by the computer, which is now a physical system. Quantum interference allows a *QTM* to act on coherent superpositions of a given state and evolves them via unitary operators into other superpositions, from which the next state results with a certain probability. Any unitary operation on n qubits can be decomposed into simple operations on one or two qubits.

A collection of n qubits constitutes a quantum register of size n (the analogue of a Turing machine tape). A quantum register of two qubits can store all four numbers $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ in superposition. Adding qubits increases the storage capacity of the register exponentially: given a quantum register of size L, a *QTM* can in one computational step perform the same mathematical operation on 2^L numbers; a classical machine has to repeat the same computation 2^L times or has to use 2^L different processors working in parallel. However, if one

tries to read a number out of a superposition of the 2^L output states, then one sees just one randomly chosen number. Only after an appropriate number of computational steps can one obtain a single final result that depends—in constructive ways—on all 2^L intermediate results.

This is how quantum algorithms work. Grover's algorithm, as an example, can determine an element from an unsorted list of *N* items in approximately $\sqrt{N}$ steps. A classical algorithm that scans the entries one by one requires on average *N*/2 steps. Another quantum algorithm, due to Peter Shor, can factorize large integers efficiently. Here, the difference in performance between the quantum and classical algorithms seems exponential. Quantum algorithms solve some important problems more efficiently than classical ones, but they do not increase the class of computable functions.

If, using Ludwig Wittgenstein words, Turing machines are humans who calculate, then quantum Turing machines are physical systems that calculate. What made this shift possible was Deutsch's analysis leading to the assertion, "Every finitely realizable physical system can be perfectly simulated by a universal Turing computing machine operating by finite means" (Deutsch 1985, p. 99). Following David Deutsch (1985), a computing machine operates by finite means if: (1) only a finite subsystem is in motion during anyone step; (2) the motion depends only on the state of a finite subsystem; and (3) the rules that specify the motion can be given finitely in the mathematical sense (e.g., by an integer). "Turing machines," Deutsch asserts, "satisfy these conditions, and so does the universal quantum computer" (p. 100). Thus, boundedness conditions also play a significant role in characterizing the computation of a quantum system.

4. CONCLUDING REMARKS

Computing machines have taken over the tasks of computors and transcend in important ways (e.g., of power and efficiency) human computational capacities. The takeover has two bases: (1) aspects of physical or intellectual reality have a finite symbolic representation, and (2) machines can take on (part of) the effective manipulation of the physical tokens involved in a representation. The latter may consist of just simulating the mechanical steps in human operations, as Turing machines do, or it may involve complex physical processes that are used in a different way, as in the case of quantum computers, when a suitable theoretical description allows them to perform a massively parallel calculation, so to speak, in a single step.

The concrete technological and scientific challenges of building a quantum computer seem enormous. Broad issues surrounding computing machines in general are multifarious and reach from the mathematically fundamental to the methodologically problematic. Can representations, for example, contain infinite components? Are there physical processes that can be viewed as computations, but do not fall within the Turing limits? What is the conceptual nature of analogue computations? Do they have to have a mathematical description that allows a calculable determination? What are the critical physical issues concerning measurement?

The ultimate challenge, articulated by Turing, is to have machines exhibit intelligence. Implementing the universal machine U meant for Turing to build a machine with discipline; producing intelligence required in addition initiative. Here, then, is the core of Turing's challenge, "Our task is to discover the nature of this residue as it occurs in man, and to try and copy it in machines" (*Gödel's Collected Works*, p. 125). Computing machines have become in their modern form scientific tools to explore, in particular, one's own intellectual nature.

See also Boole, George; Church, Alonzo; Computability Theory; Computationalism; De Morgan, Augustus; First-Order Logic; Frege, Gottlob; Gödel, Kurt; Gödel's Theorem; Hilbert, David; Jevons, William Stanley; Leibniz, Gottfried Wilhelm; Logic, History of: Modern Logic; Logic Machines; Machine Intelligence; Mathematics, Foundations of; Neumann, John Von; Pascal, Blaise; Peirce, Charles Sanders; Quantum Mechanics; Turing, Alan M.

Bibliography

The classical 1936 papers by Church, Kleene, Post, and Turing are all reprinted in *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems, and Computable Functions*, edited by Martin Davis (Hewlett, NY: Raven Press, 1965). Gödel's papers are all available in his *Collected Works*. 5 vols. (New York: Oxford University Press, 1986–2003. Turing's papers from the late 1940s are collected in *Mechanical Intelligence*, edited by D. C. Ince (Amsterdam, Netherlands: North-Holland, 1992). The classical textbook presenting the Turing-Post development of computability theory is Martin Davis's *Computability and Unsolvability* (New York: McGraw-Hill, 1958). Finally, the references to von Neumann's papers and reports are all in Goldstine (1972).

Babbage, Charles. *Passages from the Life of a Philosopher* (1864). New Brunswick, NJ: Rutgers University Press, 1994.

Davis, Martin. "Why Gödel Didn't Have Church's Thesis." *Information and Control*, 54 (1/2) (1982): 3–24.

Deutsch, David. "Quantum Theory: The Church-Turing Principle and the Universal Quantum Computer." *Proceedings of the Royal Society*, A 400 (1985).

Gandy, Robin. "Church's Thesis and Principles for Mechanisms." In *The Kleene Symposium*, edited by Jon Barwise, H. Jerome Keisler, and Kenneth Kunen. Amsterdam, Netherlands: North-Holland, 1980.

Goldstine, Herman H. *The Computer from Pascal to von Neumann*. Princeton, NJ: Princeton University Press, 1972.

Herken, Rolf, ed. *The Universal Turing Machine. A Half-Century Survey*. New York: Oxford University Press, 1988.

Hodges, Andrew. *Alan Turing: The Enigma*. New York: Simon and Schuster, 1983.

Mundici, Daniele, and Wilfried Sieg. "Paper Machines." *Philosophia Mathematica* 3 (1995).

Sieg, Wilfried. "Calculations by Man and Machine: Conceptual Analysis." In *Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman*, edited by Wilfried Sieg, Richard Sommer, and Carolyn Talcott. Urbana, IL: Association for Symbolic Logic, 2002.

Toffoli, Tommaso, and Norman Margoulis. *Cellular Automata Machines: A New Environment for Modeling*. Cambridge, MA: MIT Press, 1987.

Wolfram, Stephen. *A New Kind of Science*. Champaign, IL: Wolfram Media, 2002.

Wilfried Sieg (2005)
Rossella Lupacchini (2005)

COMTE, AUGUSTE

(1798–1857)

Auguste Comte was a French positivist philosopher. Positivism may be viewed as either a philosophical system and method or as a philosophy of history. In the latter aspect, Comte's work was almost an early history of science. He has a good claim to having originated the new science of sociology; certainly, he coined the term. His political philosophy, elaborated on the basis of his positive sociology, was a noteworthy attempt to reconcile science with religion, and the ideals of the Revolution of 1789 with the doctrine of the counterrevolution of his own time. His influence on nineteenth-century thought was strong, he had numerous disciples, such as Émile Littré, and sympathetic supporters, such as John Stuart Mill. His ideas still have important meaning and interest.

LIFE

Comte was born in Montpellier, France. Although his family were ardent Catholics, he announced at the age of fourteen that he had "naturally ceased believing in God." At this time he also seems to have abandoned his family's royalism and to have become a republican.

Comte's relations with his family were strained throughout his life. His mother, twelve years older than her husband, clutched at the son. She once wrote asking for word from him "the way a beggar asks for bread to sustain life" threatening that he would know what he had lost only when she was dead. His father and sister constantly complained of ill health; the latter appears to have suffered from hysteria. Comte portrayed them all as covetous and hypocritical and accused them of keeping him in financial distress. The facts, however, suggest that they did what they could for the son and brother whom they loved and admired but found so strange. It is necessary, in order to understand Comte's philosophy and polity, to comprehend his family's compelling influence on him. Although he rejected the ties to his parents and sister (he also had a brother), with their Catholic royalism and their strong emotional demands, these ties reasserted themselves in altered form in his later life and thought. These same family bonds also become important in understanding his nervous breakdown.

EDUCATION. Two events are outstanding in Comte's early life: his attendance at the École Polytechnique and his service as secretary to Claude-Henri de Rouvroy, Comte de Saint-Simon. The École Polytechnique, founded in 1794 to train military engineers and rapidly transformed into a general school for advanced sciences, was the product of both the French Revolution and the rise of modern science and technology, and it became the model for Comte's conception of a society ordered by a new elite. Although he was there for only a short period, from 1814 to 1816, he immersed himself in the scientific work and thought of such men as Lazare Carnot, Joseph Lagrange, and Pierre Simon de Laplace. Indeed, it was Lagrange's *Analytical Mechanics* that inspired Comte to expound, by means of a historical account, the principles animating each of the sciences.

Expelled from the École at the time of its royalist reorganization, Comte remained in Paris instead of returning home, as his parents desired. He came under the influence of the *idéologues* (Comte de Volney, Pierre-Jean Georges Cabanis, and Comte Destutt de Tracy) and, through his wide reading, of the political economists Adam Smith and J. B. Say, as well as of such historians as David Hume and William Robertson. Of major importance was the Marquis de Condorcet, whom Comte called "my immediate predecessor," and whose *Sketch for a Historical Picture of the Progress of the Human Mind* provided an outline of history in which developments in science and technology played a prominent role in humankind's rise through various stages to a period of enlightened

social and political order. Then, in August 1817, he became secretary to Saint-Simon. This crucial relationship lasted seven years, until it dissolved in acrimony.

COMTE AND SAINT-SIMON. The question of what Comte owed his patron, and what he added to the latter's ideas, is vexed. Both men were responding to the same intertwined challenges of the French, scientific, and industrial revolutions. Both sought a science of human behavior, called social physiology by Saint-Simon, and both wished to use this new science in the effort to reconstruct society. Saint-Simon, the older man, had priority in some of the ideas: He was first to announce the law of the three stages, talked of organic and critical periods, and called for a new industrial-scientific elite. Moreover, Comte's early work, including the fundamental opuscule, "Prospectus des travaux scientifiques nécessaires pour réorganiser la société," appeared as the last part of a work that also included two of Saint-Simon's writings.

However, Comte's development of the ideas—for example, the encyclopedic range of data with which he supported the idea of the three stages—went far beyond Saint-Simon and ultimately established a qualitative difference in their systems. Further, where Saint-Simon hoped to deduce his new social science from existing knowledge, such as the law of gravitation, Comte saw each science as having to develop its own method. Comte also perceived that such a development came historically; that is, only in the course of the progress of the human mind. And whereas Saint-Simonianism evolved toward a vague socialism, Comte's thought emerged as a philosophical or scientific position.

LATER LIFE. After the angry break with Saint-Simon, Comte, who could never obtain a satisfactory university post, supported himself primarily by tutoring in mathematics. Gradually, beginning in 1826, he also lectured on his new philosophy to a private audience composed of many of the outstanding thinkers of his time: Henri Marie de Blainville the physiologist, Jean Étienne Esquirol the psychologist, Jean Baptiste Joseph Fourier the mathematician, and others. From these lectures came Comte's major work, the six-volume *Cours de philosophie positive* (1830–1842).

Meanwhile, Comte entered into connubial arrangements, which were only later formalized in a macabre religious ceremony (Comte was then in the midst of his nervous breakdown) insisted upon by his mother. Although Comte was nursed back to health by his wife, the marriage was unhappy and was finally dissolved in 1842. Two years later, Comte met Mme. Clothilde de Vaux and fell deeply in love, and from this love may have come his new emphasis on a universal religion of humanity. In any case, after the *Cours,* which forms the core of Comte's positivism—the part that had the most influence on subsequent philosophers—came such various attempts to set up the religion of humanity as the *Système de politique positive* (1851–1854), and the *Catéchisme positiviste* (1852). In 1857, worn out from his labors, Comte died in wretchedness and isolation. Behind him he left only his monumental attempts at synthesis of many of the most important intellectual strands of his period.

POSITIVE PHILOSOPHY

Comte's positive philosophy emerged from his historical study of the progress of the human mind—the western European mind. India and China, he claimed, had not contributed to the development of the human mind. Indeed, by mind he really meant the sciences: astronomy, physics, chemistry, and physiology (biology). Mathematics, for Comte, was a logical tool and not a science.

THE THREE STAGES. The history of the sciences shows that each goes through three stages: the theological, the metaphysical, and the positive. The progress of each field through the three stages is not only inevitable but also irreversible; it is, in addition, asymptotic—that is, we always approach, but never obtain, perfect positive knowledge.

Briefly, Comte's view of each of the three stages is as follows: In the theological stage, man views everything as animated by a will and a life similar to his own. This general view itself goes through three phases; animism, or fetishism, which views each object as having its own will; polytheism, which believes that many divine wills impose themselves on objects; and monotheism, which conceives the will of one God as imposing itself on objects. Metaphysical thought substitutes abstractions for a personal will: Causes and forces replace desires, and one great entity, Nature, prevails. Only in the positive stage is the vain search for absolute knowledge—a knowledge of a final will or first cause—abandoned and the study of laws "of relations of succession and resemblance" seen as the correct object of man's research.

Each stage not only exhibits a particular form of mental development, but also has a corresponding material development. In the theological state, military life predominates; in the metaphysical state, legal forms achieve dominance; and the positive stage is the stage of industrial society. Thus, Comte held, as did G. W. F.

Hegel, that historical development shows a matching movement of ideas and institutions.

According to Comte, the first science to have gone through the triadic movement was astronomy, whose phenomena are most general and simple, and that affects all other sciences without itself being affected. (For instance, chemical changes on Earth, while they affect physiological phenomena, do not affect astronomical or physical phenomena.)

METHODOLOGY. In the *Cours,* Comte attempted to demonstrate, by a mass of detail, that each science is dependent on the previous science. Thus, there can be no effective physics before astronomy, or biology before chemistry. Further, the history of the sciences reveals the law that as the phenomena become more complex (as biological phenomena are more complex than astronomical), so do the available methods by which those phenomena may be treated—for example, the use of comparative anatomy in contrast to simple observation of planetary movement.

In this part of his work, Comte demonstrated the real power and flexibility of his approach. In contrast to René Descartes, who saw only one right method of conducting the reason—the geometrical method—Comte believed that each science develops by a logic proper to itself, a logic that is revealed only by the historical study of that science. He explicitly named Descartes as his predecessor and claimed to have fulfilled Descartes's work by studying the mind historically instead of merely abstractly. In Comte's view, the logic of the mind cannot be explained in a priori fashion, but only in terms of what it has actually done in the past. In this respect, Comte's position implies a fundamental revolution in philosophy.

Himself a mathematician, Comte objected to the overextended use of mathematics. In his view, mathematics was simply one tool among many. He admitted that while in principle all phenomena might be subject to mathematical treatment, in practice those phenomena far up the scale in complexity, such as biology or his hoped-for new science of sociology, were not amenable to such an approach. On the other hand, Comte sharply dissociated the positive method from the inquiry into first causes; as we have seen, this would be metaphysical, not positive, knowledge.

Observation. The first means of scientific investigation, according to Comte, is observation. We observe facts, and Comte would agree with the logical positivists of our day that a sentence that is not either a tautology or an assertion of empirical facts can have no intelligible sense. However, by the observation of a fact, Comte—perhaps more sophisticated than many of his latter-day followers—did not mean having a Humean sensation or a complex of such sensations. He meant an act of sensing that was connected, at least hypothetically, with some scientific law. Comte admitted that the simultaneous creation of observations and laws was a "sort of vicious circle" and warned against the perverting of observations in order to suit a preconceived theory. However, he insisted that the task of the scientist was to set up hypotheses about invariable relations of phenomena, concomitantly with their verification by observation.

Experimentation. After observation, understood in this sense, experimentation is the next available method. Since it can be resorted to only when the regular course of a phenomenon can be interfered with in an artificial and determinate manner, the method is best suited to physics and chemistry. In biology, interestingly enough, Comte suggested that disease—the pathological case—while not determined beforehand, could serve as a substitute for experimentation.

Comparison. For the more complex phenomena of biology and sociology, the best available means of investigation is comparison. In biology this might be comparative anatomy. In social science, the method might take the form of comparing either coexisting states or consecutive states: The first method anticipated anthropology; the latter comprised historical sociology.

SOCIOLOGY. Comte described the study of consecutive social states as a "new department of the comparative method." This "new department" was the final science to be developed by man, and the only one that had not yet entered the positive stage: sociology. As the last phenomena to be considered as falling under invariant laws, social phenomena were the ones that would give meaning to all the rest. Only by perceiving through the new science of sociology that man is a developing creature who moves through the three stages in each of his sciences could we understand the true logic of his mind.

Comte acknowledged both Baron de Montesquieu and Condorcet as his predecessors in the science of sociology, for they, too, had perceived that social phenomena appear to obey laws when correctly considered. However, the task of bringing sociology into the positive stage, or at least up to its threshold, was performed by Comte alone. He officially announced the advent of the new science in the fourth volume of the *Cours,* 47th lesson, when he proposed the word *sociologie* for what Lambert Adolphe Jacques Quételet had named *physique sociale.*

Statics and dynamics. Comte divided sociology into two parts: statics and dynamics. Social statics is the study of political-social systems relative to their existing level of civilization; that is, as functioning cultural wholes. Social dynamics is the study of the changing levels of civilization; that is, the three stages. The division into statics and dynamics is merely for analytic purposes: The distinction is one between two different ways of organizing the same set of social facts (just as, for example, in biology students of comparative anatomy and of evolution classify the same facts in different ways).

Order and progress. Statics and dynamics, then, are branches of the science of sociology. To this classification, Comte added a division between order and progress, which he conceived as abstractions about the nature of the society studied by sociology. (He further complicated the matter by using the terms *organic* and *critical* or *negative* to describe various periods.) Thus, order exists in society when there is stability in fundamental principles and when almost all members of the society hold similar opinions. Such a situation prevailed, Comte believed, in the Catholic feudal period, and he devoted numerous pages to analyzing the ideas and institutions of medieval social structure.

In contrast to the concept of order, and using images that remind one of the Hegelian dialectic, Comte posited what he called the idea of progress. He identified this progress with the period bounded by the rise of Protestantism and the French Revolution. What was now needed, Comte told his readers, was the reconciliation or synthesis of order and progress in a scientific form. Once a science of society had been developed, opinions would once again be shared and society would be stable. According to Comte, people did not argue over astronomical knowledge, and, once there was true social knowledge, they would not fight over religious or political views. Liberty of conscience, Comte declared, is as out of place in social thought as in physics, and true freedom in both areas lies in the rational submission to scientific laws.

The gradual becoming aware of and understanding of these invariable laws was what Comte meant by progress. (One of these invariable laws, incidentally, was that society must develop in a positive direction.) Thus, in the Middle Ages, when society found its order in terms of shared religious ideas, sociology was in the theological stage, and the French Revolutionary period witnessed the emergence of the metaphysical stage. As has been explained, Comte denigrated the period of progress, from the rise of Protestantism to the French Revolution, while from the point of view of social dynamics, he had to praise the progressive movement toward positivism that took place during this "negative" period. Comte's classification was neither always clear nor consistent.

POLITICAL PHILOSOPHY. Comte's sociology was overly intertwined with his conception of the right polity. In Comte's view, society had broken down with the French Revolution. The Revolution had been necessary because the old order, based on outdated "theological"—Catholic—knowledge, no longer served as a respectable basis for shared opinions; it had been undermined by the progress of the sciences. The Revolution itself offered no grounds for the reorganization of society because it was "negative" and metaphysical in its assumptions. The task, therefore, was to provide a new religion, and a new clergy, that could once again unify society. Comte's solution was a science on which all could agree. In place of the Catholic priesthood, Comte proposed a scientific-industrial elite that would announce the "invariable laws" to society. It was a bold effort to synthesize the old regime (as conceived by Comte) and the Revolution, and to meet the problems of a modern industrial society with the insights about the need for order and shared certainty that were revealed in the theological-feudal period. These insights, religious in nature and intuitive in form, were now to be reformulated by Comte and his followers in terms of positive science.

POSITIVE RELIGION. Comte, in responding to the actual problems of his time, was also working out a synthesis of two bodies of thought. Montesquieu and Condorcet have already been mentioned as the sources of Comte's conception of social statics and social dynamics. Comte's views on organic and critical periods, and his dislike of Protestantism as negative and productive only of intellectual anarchy, were undoubtedly derived from the Catholic counterrevolutionary thinkers Vicomte de Bonald and Comte Joseph de Maistre, whom he began to read around 1821. It was Bonald, in fact, who first announced that one did not argue over social truths any more than one argued over the fact that 2 plus 2 equals 4, and de Maistre stated that Protestantism is a negative ideology. Comte rewarded de Maistre by putting his name in the calendar of positivist saints.

Now the positivist calendar was a product of Comte's increasing turn from his earlier mainly philosophical and scientific interests to a form of mysticism. Comte appointed himself the high priest of a new religion of humanity. The new "religion"—based on Comte's positive science—had its holy days, its calendar of saints (which included de Maistre, Adam Smith, Frederick the

Great, Dante Alighieri, and William Shakespeare), and its positive catechism. It was nontheistic, for Comte never reverted to a belief in God or in Catholic dogma. As an effort to replace the Catholic religion with a new version of the cult of reason of 1793, it is of great interest, but it was not this aspect of Comte's work that influenced such important figures as Littré and J. S. Mill and it is not what is generally meant when one speaks of Comte's Positivism.

EDUCATIONAL THEORY. It was on the basis of the earlier, rather than the later, parts of his work that Comte sought to regenerate education. To know a given science, Comte believed, one must know the sciences anterior to it. According to this scheme, the sociologist must first be trained in all the natural sciences, whose knowledge has already gone through the three stages and become positive. (A by-product of this approach to education was Comte's conviction that the proposed method of studying would aid each science by suggesting answers to its problems from other fields.) Positive education was a necessary foundation for the positive polity, as well as for the positive sociology.

COMTE AND SOCIALISM. To round off this presentation of Comte's thought, a brief word is in order on the relationship of his views to the emerging proletarian movement. The goal of Comte's polity was never the affluent society, although he believed that every social measure ought to be judged in terms of its effect on the poorest and most numerous class. He sought, instead, a moral order, with the positive religion enjoining everyone "to live for others." The two classes from which Comte expected the greatest moral influence were women and proletarians, and he relied on their respective charms and numbers to soften the selfish character of the capitalists. In this way, class conflict would be abolished, and the owners of industry would be moralized instead of eliminated. Comte was against the abolition of private property; on the other hand, he joined Karl Marx in attacking the individualist attitudes and behavior of the property-owning classes. In this context, it is interesting to note that Marx, who claimed not to have read Comte until 1866, when he judged his work "trashy," had as a friend the Comtian Professor E. S. Beesly, who chaired the 1864 meeting establishing the International Workingmen's Association.

CRITICISM AND ASSESSMENT

Against Comte's entire system, various criticisms may be lodged. J. S. Mill took Comte to task for not giving a place in his series of sciences to psychology (instead, Comte concentrated on phrenology) and commented that this was "not a mere hiatus in M. Comte's system, but the parent of serious errors in his attempt to create a Social Science."

Perhaps there is a connection between Comte's disregard of introspective psychology and his unquestioned faith in the possibility of an ultimate positive stage of society and knowledge. For example, Comte did not even consider the question of how we can be sure that the positive stage is the last one. Since the human mind and its logical procedures, in Comte's own view, can be known only in terms of experience, it is at least theoretically possible that another stage might be reached. And how can we be sure that, although the positive method has been extended to all natural phenomena, it can be extended to human phenomena? Even if we grant this—and admittedly it is an appealing and useful assumption—does the discovery of laws regulating human phenomena put us in possession of a final science of humanity? At this point, are we not still without a science of ethics, a science that will tell us with complete positive certainty what end to pursue? Comte considered none of these questions, nor, with his neglect of introspective psychology, the further question of whether man's moral disposition is necessarily improved by the pursuit of science.

On another level, both Comte's sociology and his political philosophy can be criticized as embodying a wrong view of scientific procedure. In his best moments, he knew that science proceeds by free inquiry and constant redefinition of its "laws." However, in setting up a scientific elite, who were to announce fixed and stable laws to society, he betrayed his own insight. The polemic needs of his polity—ordered, organic, and positive—triumphed over the philosophic and scientific method he had so painstakingly elaborated in the *Cours.*

Along this same line of criticism, Comte can be charged with serious errors of fact. His anti-Protestant, pro-Catholic feelings led him to make sweeping and unexamined statements, such as that Protestantism was "anti-scientific" (a conclusion supported, perhaps, by Martin Luther's views, but undermined, for example, by the Puritan involvement in the Royal Society) and that Catholicism was a nonaggressive religion. Thus, speaking of the Crusades, Comte asserted, as a matter of fact: "All great expeditions common to the Catholic nations were in fact of a defensive character." Throughout his work, especially in the last three volumes of the *Cours,* which are devoted to his sociology rather than to the natural sciences, similar remarks are to be found.

Yet, with all the criticisms of either a conceptual or factual nature that can be leveled against Comte's position, one must not lose sight of his essential contributions. He did grasp the notion that knowledge in the various sciences is unified and related. His law of the three stages, while too rigid and schematized, did point to the different ways of viewing the world and to the fact that men at different stages of history have emphasized one way of ordering society more than another. And, most important, Comte did prepare the way for a new science, sociology, that would help study the interrelations of men in society and how these interrelations change in the course of history.

See also Bonald, Louis Gabriel Ambroise, Vicomte de; Cabanis, Pierre-Jean Georges; Condorcet, Marquis de; Dante Alighieri; Descartes, René; Destutt de Tracy, Antoine Louis Claude, Comte; Hegel, Georg Wilhelm Friedrich; Hume, David; Laplace, Pierre Simon de; Littré, Émile; Maistre, Comte Joseph de; Marx, Karl; Mill, John Stuart; Montesquieu, Baron de; Positivism; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Smith, Adam; Volney, Constantin-François de Chasseboeuf, Comte de.

Bibliography

There is no critical edition of Comte's works. His most important writings, all published in Paris unless otherwise stated, are as follows: *Opuscules de philosophie sociale 1819–1828* (1883), which includes the 1822 "Plan des travaux scientifiques nécessaires pour réorganiser la société"; *Cours de philosophie positive,* 6 vols. (1830–1842), and Harriet Martineau's English condensation of *Cours, The Positive Philosophy of Auguste Comte,* 2 vols. (London: John Chapman, 1853), which was personally approved by Comte; *Discours sur l'esprit positif,* prefixed to the *Traité philosophique d'astronomie populaire* (1844); *Discours sur l'ensemble du positivisme* (1848); *Calendrier positiviste* (1849); *Système de politique positive,* 4 vols. (1851–1854), translated by J. H. Bridges, Frederic Harrison, et al. as *The System of Positive Polity,* 4 vols. (London: Longmans, 1875–1877); *Catéchisme positiviste* (1852), translated by Richard Congreve as *The Catechism of Positive Religion* (London: John Chapman, 1858); *Appel aux conservateurs* (1855); and *La synthèse subjective* (1856).

In addition, see P. Valat, ed., *Lettres d'Auguste Comte à M. Valat* (1870); *Lettres d'Auguste Comte à John Stuart Mill, 1841–1846* (1877); *Testament d'Auguste Comte* (1884); *Lettres à des positivistes anglais* (London: Church of Humanity, 1889); *Correspondance inédite d'Auguste Comte,* 4 vols. (1903–1904); and *Nouvelles Lettres inédites. Textes présentés par Paulo E. de Berredo-Carneiro* (1939).

The most important work on Comte, essential to a study of his intellectual development, is Henri Gouhier, *La jeunesse d'Auguste Comte et la formation du positivisme,* 3 vols. (Paris: Vrin, 1933–1941). The same author's *La vie d'Auguste Comte* (Paris: Gallimard, 1931) is the best biography. For an analysis of Comte's philosophical ideas, J. S. Mill, *Auguste Comte and Positivism* (London: Trubner, 1865) is still obligatory. See also Thomas Whittaker, *Comte and Mill* (London: A. Constable, 1908); Lucien Lévy-Bruhl, *La philosophie d'Auguste Comte* (Paris: Alcan, 1900), translated by Kathleen de Beaumont-Klein as *The Philosophy of Auguste Comte* (New York: Putnam, 1903); Émile Littré, *Auguste Comte et la philosophie positive* (2nd ed., Paris: Hachette, 1864); J. Delvolvé, *Reflexions sur la pensée comtienne* (Paris: Alcan, 1908); and Pierre Ducassé, *Méthode et intuition chez Auguste Comte* (Paris: Alcan, 1939).

On Comte's religious attitudes, see George Dumas, *Psychologie de deux Messies positivistes: Saint Simon et Auguste Comte* (Paris: Alcan, 1905).

Jean Lacroix, *La sociologie d'Auguste Comte* (Paris: Presses Universitaires de France, 1956) is one of the most interesting books on Comte. For a critical view of Comte's sociology in relation to morality, H. B. Acton, "Comte's Positivism and the Science of Society," in *Philosophy* 26 (October 1951) is valuable.

Treating Comte as a historian of science are Paul Tannery, "Comte et l'histoire des sciences," in *Revue générale des sciences* 16 (1905), and George Sarton, "Auguste Comte, Historian of Science," in *Osiris* 10 (1952). In this connection, John C. Greene, "Biology and Social Theory in the Nineteenth Century: Auguste Comte and Herbert Spencer," in *Critical Problems in the History of Science,* edited by Marshall Clagett (Madison: University of Wisconsin Press, 1959) is interesting. Frank Manuel, *The Prophets of Paris* (Cambridge, MA: Harvard University Press, 1962), Ch. 6, and F. A. Hayek, "Comte and Hegel," in *Measure* 2 (1951), are rewarding; the Hayek article treats Comte as a historicist.

Bruce Mazlish (1967)

CONCEPTS

Concepts are customarily regarded as intermediaries between mind and world. They are the basic elements of thoughts and the tools by which one classifies things. Concepts are central to the philosophy of mind, and they are often implicated in theories of meaning. There are also some who think that philosophical method is largely a matter of conceptual analysis. There is considerable consensus on the importance of concepts, and, to a lesser extent, on the roles that concepts play, but beyond that there is rampant disagreement. For example, philosophers disagree about the ontology of concepts, the acquisition of concepts, and the content of concepts. In the twentieth century, psychologists began to weigh into these debates, and since the 1970's, much theorizing about concepts has been informed by interdisciplinary dialogue. This entry surveys dominant theories.

WHAT ARE CONCEPTS FOR?

Within philosophy concepts are most often defined as the elements or ingredients of thoughts. Concepts are for thinking. When one ascribes a thought, such as "aardvarks are nocturnal," one typically assumes that the bearer of that thought has a concept of aardvarks and a concept corresponding to the property of being nocturnal. It is sometimes said that a concept is to a thought as a word is to a sentence, but this formula can mislead, because some philosophers do not believe that thought is language-like. However, even those who resist the view that thought is language-like may be attracted to one crucial point of analogy: concepts are believed to be combinable. Those who possess a concept of aardvarks can form the thought that aardvarks are nocturnal, the thought that aardvarks are quadrupeds, or the thought that aardvarks are insectivores, assuming possession of these other concepts. Gareth Evans (1982) suggests that it is a condition on concept possession ("the generality constraint") that, if a person can have the thought that a is F, then that person should also be able to form every other thought of the form a is X, where X ranges over the concepts in that person's conceptual repertoire. Some philosophers think there may be restrictions (e.g., of intelligibility) on combination, but most agree that thought formation through conceptual combination is a central function of concepts.

A second function of concepts is categorization. Many philosophers think that concepts are the primary tools by which one determines that something falls into a category. One knows that two things are both turtles in virtue of having a turtle concept. Historically, some philosophers have reserved the word *concept* and closely related words for general kinds. On this usage there can be a concept of turtles, in general, but not a concept of a particular turtle, say Yertle. Other philosophers tend to say that there can be concepts of individuals and that concepts can be singular as well as general. When one identifies an individual, one can think of that as an act of categorization, broadly construed: One categorizes that individual as belonging to a class with one member. Concepts are implicated in the categorization of kinds and individuals.

Concepts are sometimes said to have a function in inference. This third function often works in concert with the second. One uses concepts to draw inferences about the things that one categorizes. If one encounters a shovel, one can infer that it is used for digging. The knowledge that shovels are used for digging is said, by many, to be contained in one's concept of shovels. Thus, when one applies the concept to some thing, one can use the concept to infer facts about that object.

Concepts are also widely presumed to play a role in linguistic meaning. For some, concepts simply are the meanings of words or components of meanings. On this view concepts are expressed when one uses words. Some philosophers' (especially those who favor reference-based semantic theories) concepts are not meanings. But these authors usually concede that concepts play a central role in the epistemology of language. One comes to understand a word by associating it with a concept. On either approach concepts and language will be closely related.

A fifth function of concepts is related to the other three, but is potentially dissociable. Concepts are said to be representations; they refer to things. Some theories of concepts encompass theories of reference. In this sense, concepts are intermediaries between mind and world.

There is controversy about what concepts are for, but the items on the preceding list are widely accepted. Concepts are usually postulated to play all or some of the preceding roles.

SOME ISSUES OF CONTROVERSY

In describing some of the functions of concepts, a few places of controversy have already been indicated. There are a number of other controversies that deserve special mention.

One issue concerns ontology. It is widely agreed that concepts are intermediaries between mind and world, but where do they reside? One possibility is that they are timeless *abstracta*. This view has been especially popular among those who identify concepts with word meanings. Many semantic theorists believe that meaning enjoy some autonomy from psychology. On this view the meaning of a word does not depend on the images or ideas any individual happens to possess. Others are attracted to this view because they regard concepts as a specification of the essential properties of the things to which they refer.

The concept of a triangle, on this view, might be a geometric definition. Triangles had that definition before anyone discovered it. In contrast, there are philosophers who locate concepts inside the head. On this approach a concept is a mental representation, which plays a causal role in information processing. Others regard concepts as psychological, but eschew talk of mental representations. For example, one might say that a concept is a skill or ability or an operation on mental representations. Immanuel Kant (1997) says concepts are rules for constructing or organizing images. In between those who say

that concepts are abstract and those who say they are psychological, there are social theories of concepts, according to which concepts supervene on human practices. What matters is not the contents of any individual's mind, but socially distributed patterns of deference, normative demands, and reason-giving behavior. There is room for uniting all these ontological perspectives into a single theory. For example, one could say that individuals have mental representations (psychological concepts) of community-enforced rules (social concepts) that dictate which timeless, essential properties their thoughts denote (abstract concepts).

Among those who think that concepts are mental representations, there are significant disagreements about representational format. Some think concepts are words in language-like mental code ("the language of thought hypothesis"), others claim they are mental images ("imagism"), and still others say they are weighted connections or patterns of activation in neural networks ("connectionism").

Those who think that concepts are mental representations also disagree about how concepts are attained. Some think that many concepts are innate, and some think that few or none are innate. There are controversies about how learned concepts are acquired. Concepts might be copied from experience, they might be abstracted, they might be learned by strengthening associations, or they might be acquired using a more deliberative procedure, such as the formation and testing of hypotheses. The innateness question is sometimes posed as a question of which concepts are primitive. Many philosophers believe that some concepts are primitive and others are assemblies or inferential networks built up from these. (When two concepts are combined to form a third, they are said to be "features" of that third concept.) Primitive concepts are often thought to be innate, so debates about this issue can sometimes be characterized as debates about how many primitives one has. Historically, however, some philosophers have assumed that many complex concepts are innate as well (such as the concept of God or of identity).

Those who think that concepts are abstract or otherwise external to individual minds sometimes talk about concepts using a definite article, "The concept of X." Those who think that concepts are mental representations are less likely to talk this way, leaving open room that different people may have different concepts of the same thing. There may be exceptions to this rule. It is natural to speak of technical concepts with a definite article ("The concept of natural selection") because there is sometimes just one correct formulation. In addition, some philosophers think concepts are individuated by their referents. On this view any concepts of the same thing will count as being the same concept. Hence, it would always make sense to talk about concepts using the definite article.

Another controversy surrounds the relationship between concepts and language. Besides the question of whether concepts are meanings (hence, whether language depends on concepts), there is a question of whether concepts depend on language. This conclusion has been defended by Ludwig Josef Johann Wittgenstein, Michael Anthony Eardley Dummett, and Donald Davidson. The arguments often turn on the claim that having concepts requires recognizing that thoughts and inferences can be mistaken, which depends in turn on belonging to a language community whose members give and demand reasons for utterances. In contrast, many think that concepts can be possessed without language, and, indeed, Jerry A. Fodor (1975) argues that language learning would be impossible without prior possession of concepts.

All these controversies are significant, but the main issue dividing competing theories of concepts has to do with content. Philosophers disagree about what information one knows in virtue of possessing concepts. One knows a great deal about many categories, but many philosophers believe that only some of this knowledge is conceptually constitutive. Some of this knowledge belongs to one's concepts, and the rest merely belongs to one's conceptions, where conceptions are thought to be more ephemeral and idiosyncratic than concepts. Theories of concepts can be distinguished by where they draw the concept-conception divide.

THE CLASSICAL THEORY

One theory of concepts has been so dominant in the history of philosophy that it has been dubbed "the classical theory." The name is apt, because the theory is championed by Plato. In classical theory, concepts are definitions: They specify conditions that are individually necessary and jointly sufficient for the categories they designate. In his dialogues Plato tries to uncover definitions of concepts such as justice, knowledge, piety, and love. On this approach specifying a concept of justice is a matter of specifying what it is to be just. It is unclear whether Plato thinks concepts are abstract entities or mental entities. He claims that people categorize things by recalling a life in a world of ideal forms, which they inhabited before life in the terrestrial world. Possessing a

concept is a matter of intuiting, through memory and reflection, the essence of these ideal forms.

Many philosophers have assumed that some version of the classical theory is correct. Kant (1997) says that concepts are rules that determine the conditions of category membership. He also suggests that many concepts contain other concepts, like houses made from bricks, and in the *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science, with Selections from the Critique of Pure Reason* (1997) he introduces the term *analytic* to refer to judgments whose predicate concepts are contained in their subject concepts. These judgments are, in effect, true by definition, as opposed to synthetic judgments, which are not true by definition, but must be discovered.

Gottlob Frege (1960) uses the term *concept* (*Begriff*) to refer, narrowly, to the concepts expressed by predicates, but he uses a more encompassing term *sense* (*Sinn*) to refer to the components of thoughts, and each of these, he suggests, can be identified with a descriptive content that determines reference. Frege insists that senses are abstract entities; if they were in the head, he thought they could not serve as the shared meanings of words. Inspired by Frege, Christopher Peacocke (1992) claims that concept possession involves the mastery of inferences, which play a central role in determining reference.

Rudolf Carnap (1956) claims that the concepts used in ordinary thought and talk are riddled with imprecision and that they need to be replaced by concepts that are explicitly defined. Analytic truths are stipulated, and hence immune from empirical refutation.

Defenders of the classical theory disagree about how concepts are attained. Plato obviously thinks concepts are innate, and Carnap thinks explicated concepts must be learned. For many classical theorists, some are innate and others are learned.

The classical theory has been criticized in various ways. Willard Van Orman Quine (1981) argues that the distinction between analytic and synthetic truths is unprincipled, because any putatively analytic claim could be revised under empirical pressure, if, for example, such a revision would be the most conservative way to alter a prevailing theory to accommodate new evidence. Hilary Putnam (1975) argues that definitions are not essential for reference; one can think about natural kinds (e.g., tigers, gold, and water) even if no one grasps the conditions that are necessary and sufficient for falling in the categories. Ludwig Wittgenstein (1953) argues that concepts often lack catchall definitions; instead, concepts group things together on the basis of family resemblances (games are his famous example). Psychologists support Putnam and Wittgenstein by showing that people rarely know the defining features of a category. Georges Rey (1983) counters that the psychological objections presuppose that concepts are in the head and readily available to consciousness—some classical theorists are willing to deny both assumptions.

CONCEPT EMPIRICISM

Plato does not say much about how concepts are mentally represented. Aristotle has more to say. He says that every concept is accompanied by an image. This idea inspired subsequent empiricist philosophers to propose that concepts are perceptual in nature. This basic claim is the essence of concept empiricism. Scholastic philosophers say that nothing is in the intellect that is not first in the senses. The British empiricists, such as John Locke (1979) and David Hume (1978), say that concepts are derived from percepts. Hume says concepts are simply copies of percepts or combinations of copied percepts, and Locke proposes that many concepts are acquired through abstraction from percepts (though there is some disagreement about what he and his scholastic predecessors meant by abstraction).

Concept empiricists differ in several ways from typical classical theorists. First, many concept empiricists are imagists, whereas many classical theorists are not. Concept empiricists usually say that concepts are mental representations (the British empiricists use the term *ideas*), whereas classical theorists often say they are abstracta. Concept empiricists emphasize learning, whereas traditional classical theorists assume that many concepts are innate. Concept empiricists claim that concepts refer either by resemblance or by causal relations to their referents, whereas classical theorists usually assume that concepts refer by satisfying lists of defining conditions.

Nevertheless, there are theories that straddle the border between the classical theory and concept empiricism. The verificationist theories of concepts advanced by Carnap (1956) and other logical positivists are a case in point. For a verificationist, concepts consist in conditions that are necessary and sufficient for reference, but these conditions are specified in observational vocabulary; a concept refers to that which satisfies perceivable conditions of verification.

Concept empiricism is widely believed to face serious objections. One has to do with the concepts of abstract categories. There seem to be concepts of virtue, truth, substance, cause, and being, yet none of these things has

any characteristic appearances. If concepts were all derived from perception, it would be hard to explain how people think about these things. Concept empiricists reply by either arguing that people do not have concepts of these things, or by reducing these concepts to perceptual features. Both strategies are hard to pull off.

Another objection is put forward by Kant (1997). He argues that one's capacity to perceive presupposes the possession of certain concepts (including concepts of time and space), which could not be derived from experience. Contemporary psychologists also argue that there is empirical evidence for innate concepts, which are evidently in place before experience, such as the concept of physical object or of number.

PROTOTYPE THEORY

When the classical theory came under attack in the middle of the twentieth century, new alternatives were sought. One alternative, already mentioned, was Wittgenstein's (1953) family resemblance account, according to which one comprehends categories by means overlapping features rather than a catchall definition. This suggestion spawned the emergence of the cluster theory, which identified concepts with features that are not individually necessary for category membership but sufficient when a sufficient number are in place. No one feature may suffice for being a game, and no one feature is necessary, but bring a few features together and one has a game. In effect, the cluster theory is a similarity theory of concepts; it says that one categorizes by looking for similarities with familiar instances.

In psychology, dissatisfaction with the classical theory and inspiration from Wittgenstein (1953) gave rise to the prototype theory. On this approach categorization is also a matter of assessing similarity to a set of features that are not individually necessary for category membership. Prototype theorists do not construe concepts as unwieldy clusters, but as summary representations capturing just those features that are most typical of the category. A prototype is a representation of features that are highly frequent, salient, and diagnostic for category membership. The prototype for the category bird might include features such as flies, has feathers, has a beak, and sings. Following Putnam (1975), philosophers sometimes use the term *stereotypes* for much the same thing. Psychologists, notably Eleanor Rosch and Carolyn Mervis (1975), support the postulation of prototypes by showing that people categorize prototypical category members faster, learn to recognize them earlier, and list prototypical features first when asked to describe a category.

Prototype theorists usually assume that concepts are mental representations, but they diverge on the format of these mental representations. Some say they are made up of images, some say they are patterns in connectionist networks, and some say they are lists of features coded in a language of thought. Like some concept empiricists, prototype theorists argue that concepts are often learned by abstracting from particular category instances, but prototype theorists do not always assume that concepts are grounded in perceptual experience. There can be prototypes for categories that are difficult to discern perceptually, such as a prototypical analytic philosophy paper, a prototypical democracy, or a prototypical lie.

Prototypes are often used in categorization, but some psychologists and philosophers argue that they should not be equated with concepts. One objection is that similarity to a prototype is not necessary for categorization and reference; a shaved, mute, tailless, three-legged dog is completely unlike the dog prototype but still falls under the category. Similarity to a prototype is also not sufficient for categorization and reference: a duck decoy is no duck. Another objection is that prototypes do not combine together compositionally: the prototype for a compound concept is often unlike the prototype for its parts. Pet fish prototypically live in bowls, but neither pets nor fish prototypically live in bowls. Fodor (1998) argues that concepts must combine compositionally to explain that one can generate an unbounded number of novel thoughts from a finite stock of concepts. For similar reasons, prototypes may violate Evans's Generality Constraint (1982), which implies that concepts can be freely recombined; someone might know the prototypes for red fruit and long hair without knowing the prototypes for red hair and long fruit.

THE THEORY THEORY AND HOLISM

Unconvinced by prototype theory, some psychologists developed an alternative, which is associated with the following basic tenets. First, not all concepts are alike; one must distinguish animal concepts, artifact concepts, psychological concepts, mathematical concepts, concepts of physical objects, and so on. Each of these classes is governed by different "folk theories" that comprise small collections of basic principle; for example, folk biology explains that animals have hidden genetic essences, and folk physics explains that solid objects cannot pass through each other. Second, some folk theories lead one to postulate defining essences (as in the case of folk biology), but, unlike classical theorists, psychologists do not assume that these essences are known to those who pos-

tulate them; this is called psychological essentialism. Third, each of the concepts within one of these classes may contain causal and explanatory features besides prototypical features; for example, a concept of birds may contain the belief that wings enable flight. Together, these tenets suggest that concepts are like scientific theories: they divide into domains, they postulate hidden features, and they play a role in explanation. The approach has been dubbed the *theory theory*.

Most theory theorists assume that some rudimentary folk theories are innate, but they disagree about which ones. They also disagree about whether one's innate theories remain intact over development, or whether they undergo significant transformations, akin to conceptual revolutions in science. On the latter view adult concepts may be incommensurable with the concepts of children and infants.

The theory theory has been primarily developed by psychologists, but related ideas can be found in philosophy. Quine's (1981) critique of the analytic-synthetic distinction has led some to believe that the basic units for understanding any given category is an entire theory. Quine differs from most psychologists in three respects: He does not claim that theories are insulated from each other (perturbations in one may have ripple effects); he does not claim that theories are mentally represented in the head (Quine is a behaviorist); and he assumes that theories are learned (one's initial sorting behavior is driven by superficial similarities). Still, one might appropriate Quine's ideas into a psychological theory by proposing that each concept is a mental representation individuated by its place in a complete network of mental representations. This would be a holistic theory of concepts.

Critics of the theory theory and holism worry that these approaches entail that concepts are rarely shared. If two people have different theories, then they have different concepts, and their ability to communicate and to obey the same psychological laws becomes difficult to explain. It is also unclear whether these approaches can explain how concepts are combined to form thoughts, because theories are too cumbersome to easily combine together.

INFORMATIONAL ATOMISM

The theory theory and holism pack a lot of information into concepts. Some philosophers prefer the opposite strategy. Fodor (1998) argues that just about every lexical concept (a concept expressed by a single word) is primitive: a primitive concept is one that is not individuated by its relation to any other. This is called atomism. Instead, concepts are individuated by their referents, and concepts refer by falling under the nomic control of properties; roughly, a cow concept refers to cows because it is a law that encounters with cows and causes cow concepts to be tokened. This is called informational semantics.

Informational atomism is unlike all the theories considered so far, because all the others assume that most lexical concepts are complex. A primary advantage of informational atomism is that it can explain how concepts are recombined compositionally. If concepts are primitive symbols, then they can retain their identity when combined, just as words retain their shape when placed into sentences. Concepts can also be easily shared on this view: Two people have the same concept if they have symbols that are under nomic control of the same properties regardless of any difference in their beliefs.

These advantages come at cost. If lexical concepts are primitive, then they cannot be used to explain the inferences one draws or the way one categorizes. For a thoroughgoing atomist, someone could possess a concept of bachelors without knowing that they are male or unmarried. Atomism has also been associated with radical concept nativism. Many philosophers assume that primitive concepts are innate and that complex concepts are learned; if all lexical concepts are primitive, then all are innate. Fodor (1981) used to embrace this consequence, and Fodor (1998) now argues that primitive concepts can be learned.

PHILOSOPHY AS CONCEPTUAL ANALYSIS

Beginning with Plato, one of the dominant methods for doing philosophy has been philosophical analysis. Practitioners begin with a specific concept and reflect on its content. In so doing, they hope to reveal not only how one thinks about the referent of that concept but also what the essence of that referent is. By reflecting on the concept of virtue, for example, one might reveal what it is to be virtuous. The viability of this method depends on which theory of concept is correct. If concepts are definitions, conceptual analysis can reveal the essence of things. But if concepts are merely assemblies of typical features, incomplete and revisable theories, or semantically primitive symbols, then conceptual analysis cannot reveal the essence of things. There is as yet no consensus on which theory of concepts is right, but at stake is the methodology of philosophy itself.

See also Content, Mental.

Bibliography

Carey, Susan. *Conceptual Change in Childhood*. Cambridge, MA: MIT Press, 1985.

Carnap, Rudolf. "Empiricism, Semantics, and Ontology." In *Meaning and Necessity: A Study in Semantics and Modal Logic*. 2nd ed. Chicago: University of Chicago Press, 1956.

Davidson, Donald. "Thought and Talk." In *Mind and Language*, edited by Samuel Guttenplan, 7–23. Oxford, U.K.: Clarendon Press, 1975.

Dummett, Michael. *The Interpretation of Frege's Philosophy*. Cambridge, MA: Harvard University Press, 1981.

Evans, Gareth. *Varieties of Reference* edited by John McDowell. New York: Oxford University Press, 1982.

Fodor, Jerry A. *Concepts: Where Cognitive Science Went Wrong*. New York: Oxford University Press, 1998.

Fodor, Jerry A. *The Language of Thought*. Cambridge, MA: Harvard University Press, 1975.

Fodor, Jerry A. "The Present Status of the Innateness Controversy." In *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 1981.

Frege, Gottlob. "On Sense and Reference." In *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black.bOxford, U.K.: Blackwell, 1960.

Hume, David. *A Treatise of Human Nature*, edited by Peter H. Nidditch. New York: Oxford University Press, 1978.

Kant, Immanuel. *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science, with Selections from the Critique of Pure Reason*, edited and translated by Gary Hatfield. New York: Cambridge University Press, 1997.

Keil, Frank C. *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press, 1989.

Locke, John. *An Essay concerning Human Understanding*, edited by Peter H. Nidditch. New York: Oxford University Press, 1979.

Margolis, Eric, and Stephen Laurence, eds. *Concepts: Core Readings*. Cambridge, MA: MIT Press, 1999.

Medin, Douglas, and Anthony Ortony. "Psychological Essentialism." In *Similarity and Analogical Reasoning*, edited by Stella Vosniadou and Anthony Ortony. New York: Cambridge University Press, 1989.

Murphy, Gregory, and Douglas Medin. "The Role of Theories in Conceptual Coherence." *Psychological Review* 92 (1985): 289–316.

Peacocke, Christopher. *A Study of Concepts*. Cambridge, MA: MIT Press, 1992.

Prinz, Jesse J. *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press, 2002.

Putnam, Hilary. "The Meaning of 'Meaning.'" In *Language, Mind, and Knowledge*, edited by Keith Gunderson. Minneapolis: University of Minnesota Press, 1975.

Quine, Willard Van Orman. "Two Dogmas of Empiricism." In *From a Logical Point of View: Nine Logico-philosophical Essays*. 2nd ed. Cambridge, MA: Harvard University Press, 1981.

Rey, Georges. "Concepts and Stereotypes." *Cognition* 15 (1983): 237–262.

Rosch, Eleanor, and Carolyn Mervis. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7 (1975): 573–605.

Smith, Edward E., and Douglas L. Medin. *Categories and Concepts*. Cambridge, MA: Harvard University Press, 1981.

Spelke, Elizabeth. "Initial Knowledge: Six Suggestions." *Cognition* 50 (1994): 431–445.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Basil Blackwell, 1953.

Jesse Prinz (2005)

CONDILLAC, ÉTIENNE BONNOT DE

(1714–1780)

Étienne Bonnot de Condillac was one of the French *philosophes*, known primarily for his development of the doctrine of "sensationism." According to this doctrine, not only all of one's thoughts but even the basic operations on these thoughts derive from sensation.

Condillac was born on September 30, 1714, in Grenoble, one of five children of Gabriel Bonnot, vicomte de Mably, and Catherine de la Coste. He took the name of Condillac after his father purchased an estate of that same name in 1720. Condillac was born with poor eyesight that prevented him from reading before the age of twelve, and he was considered in his childhood to possess only limited intellectual abilities. However, in 1730 he took up residence with his brother, the abbé de Mably, in Lyon to attend the Jesuit college there, and in 1733 he went to Paris to study at the Sorbonne, where he later became a seminarian at Saint-Suplice. Condillac defended his thesis in theology in 1739, and he took holy orders around 1741, though he subsequently devoted himself more to study than to pastoral work. Indeed, he was said to have celebrated Mass only once in his life. While in Paris Condillac frequented the salons and was exposed to the views of John Locke and Isaac Newton. He was influenced in particular by Locke's critique of innatism and Newton's method of explaining phenomena in terms of simple general principles drawn from experience.

Condillac was well connected in French Enlightenment circles. His cousin was Jean Le Rond d'Alembert, coauthor of the *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers* (1751–1765), and he was a friend of the other coauthor, Denis Diderot, as well as of Jean-Jacques Rousseau. The entries of the *Encyclopédie* on "Mémoire (*Métaphysiq*)," "Réflexion (*Logique*)," and "Signe (*Métaphysiq*)" reflect the influence of Condillac's views on these topics. The first of his philosophical writings was the *Essai sur l'origine des connaissances humaines* (Essay on the Origin of Human Knowledge), which was

published anonymously in 1746, after Diderot had helped him find a publisher. Around this time Condillac corresponded with the French scientist Pierre-Louis Moreau de Maupertuis, who was then the president of the Royal Prussian Academy in Berlin.

In 1746 Condillac submitted an essay on Gottfried Wilhelm Leibniz's theory of monads to a competition sponsored by the academy (it was not selected for the prize), and he was elected to this organization in 1749. Also in 1749 Condillac published his *Traité des systèmes* (Treatise on Systems), a critique of the metaphysics and methodologies of philosophers such as René Descartes, Nicolas Malebranche, Benedict (Baruch) de Spinoza, and Leibniz. He published his second main philosophical work, *Traité des sensations* (Treatise on Sensations), in 1754. The following year he published *Traité des animaux* (Treatise on Animals), a work that emphasizes the differences between human and animal souls, and "Extrait raisonné" of the *Traité des sensations*. In 1755 he also produced a "Dissertation sur la liberté" (Dissertation on Liberty), appended to the *Traité des sensations*, that addresses the issue of human freedom.

In 1758 Condillac became tutor to the young Prince of Parma, grandson of Louis XV. He spent nine years in Parma, during which time he wrote with the help of his brother the multivolume *Cours d'Etudes* (Course of Study), which was published in 1775. He returned to Paris in 1768, when he became a member of the Académie française, but left Paris again in 1773 to take up residence at the chateau de Flux, near Beaugency, which his niece had purchased for him. After that time he published a work on commerce in 1776 and a textbook on logic, which the comte Stanislas Félix Potocki had requested for his Polish schools in 1780. On August 3, 1780, Condillac died at his chateau after a return from a trip to Paris. He left behind an unfinished manuscript, *La langue des calculs* (The Language of Calculation), which was first published in 1798.

MIND AND SENSATIONS

RELATION TO LOCKE. In his *Essai* Condillac acknowledged his great debt to Locke's *Essay concerning Human Understanding*, and in particular to the attack there on innate ideas and to Locke's empirical investigation of the origin of human thought. This debt is reflected in the subtitle of the English translation of the *Essai*: "A Supplement to Mr. Locke's Essay." Even so, Condillac argued explicitly against Locke that one can know with certainty that the mind that is the subject of thought is an indivisible and immaterial substance wholly distinct from body (2001, I.i.§6, pp. 12f). In later years Condillac was especially concerned to distance himself from the materialism of more radical French Enlightenment figures such as Julien Offray de La Mettrie and Paul-Henri Thiry, Baron d'Holbach. In the *Essai* Condillac also distinguished himself from Locke by emphasizing the possibility that when separated from the body one's mind can derive knowledge independently of the senses. However, he noted that in one's present condition, that is, after the fall from the biblical state of innocence that the first humans experienced, the human mind is wholly dependent on the body, to the extent that one can have no thought that does not have a sensory origin. In Condillac's terms, all of one's thoughts are simply "*sensations transformées*."

Condillac's dualism informs his conclusion that sensations are modification of an immaterial mind. In his *Essai* he also claimed to follow Locke in holding that there are no sensory impressions in one's mind of which one is not conscious. Indeed, at one point he used this same point against the account in Locke's *Essay* of shape perception. This account addresses the speculation of Locke's friend, William Molyneux, that a man born blind would on recovering sight not be able to immediately distinguish a cube from a sphere by vision alone, without the aid of touch. Locke accepted this conclusion and claimed on the basis of this hypothetical case that one's perception of three-dimensional shapes involves not only sensations of light and color but also judgments that alter these sensations "without our taking notice of it." Condillac objected that the phenomenology of shape perception belies this account. One's sensations of light and colors render one immediately conscious of a three-dimensional world. Condillac did mention the 1729 report to the Royal Society in London by the English surgeon William Chesselden that subjects who had blinding cataracts removed could not recognize shapes. But he proposed that this result was due simply to the fact that the subjects were overwhelmed by the new sensory information and thus were unable to focus properly on the shapes (2001, I.vi.§16, p. 110).

RELATION TO BERKELEY. In a 1749 *Lettre sur les aveugles* (Letter on the Blind) Diderot charged that Condillac's *Essai* had failed to respond adequately to an idealism in George Berkeley that precludes any awareness of an external material world. Condillac in effect responded to this charge by attempting in his *Traité des sensations* to give an account of one's perception of the extended world that does not simply assume from the start that such a world exists. He introduced the example of a slowly animated statue to illustrate the manner in which one comes to per-

ceive the external world. This statue is supposed to possess initially only the sense of smell and to perceive this smell merely as an aspect of itself, and not as part of an external world (*Traité des sensations* I.i.2). Even when the statue comes to sense colors, the colors themselves are not considered as constituting distinct shapes. It is only with the sense of touch that the statue acquires an awareness of objects in space and attributes various sensible qualities to such objects (III.iii.§2). Here, Condillac abandoned his view in the *Essai* that one senses shapes by means of the sensations of light and color alone. He also granted in the *Traité*, in effect, that one is not immediately aware of everything in one's sensations. Even though sensations of color are shaped, one cannot recognize the shapes until one comes to associate colors with various tactile sensations.

In a supplement to his 1756 *Lettres à un Américan*, Joseph Adrien Lelarge de Lignac objected that, by allowing in the *Traité* that one has color sensations that are themselves extended, Condillac illicitly attributed to spirits a quality that pertains to bodies alone. In his "Lettre de M. l'abbé de Condillace à l'auteur des Lettres à un Américan," first published the same year, Condillac responded that colors are considered as manners of being of the mind only with respect to their chromatic features, and not with respect to their extension or shape. On the view in the *Traité*, one can recognize the colors as marking out shapes only when one associates them with tactile sensations and, on that basis, attributes the shapes to external objects. But there is still the question whether the color sensations themselves are extended, however one might consider them. Here, Condillac could perhaps draw on Berkeley's view in his *Principles of Human Knowledge* (1710) that extension exists in the mind not "by way of mode or attribute" but "by way of idea." There is still Berkeley's challenge that the extension that exists by way of idea can in no way "resemble" any purported extramental extension. But it is not clear that Condillac was too concerned to respond to this sort of challenge given the skeptical suggestion in his writings that one cannot know for certain whether any object exists external to mind and, if any does, what the nature of such an object is (*Traité des sensations* IV.v).

MENTAL OPERATIONS AND SIGNS

In his introductory remarks in the *Essai* Condillac claimed to have found a "fundamental fact of experience" that explains all operations involved in human knowledge, a fact that consists in "the connection of ideas, either with signs or among themselves" (2001, p. 5). An important part of Condillac's sensationism is his claim that not only the ideas but even their connections with signs or among themselves derive from sensation. He focused in particular on the initial connections forged through imagination, memory, and reminiscence. Imagination occurs when a perception is recalled at the sight of an object. This operation is possible because of an association between the object and perception set up by attention to their conjunction in experience. The attention is itself developed by associations of perceptions with sensations of pleasure and pain. Memory is a more developed operation that involves the recall not of the perception itself, but only of certain signs or circumstances associated with the object. Thus, memory is an imagination of these signs. Finally, reminiscence is the most developed of the operations, which involves not merely the formation of previously experienced perceptions, as in imagination, or previously experienced signs, as in memory, but also the recognition that the recalled perceptions or signs were experienced in the past. The ability so to recognize itself depends on the previous exercise of the imagination and memory.

In the *Essai* Condillac distinguished among three kinds of signs involved in the development of memory and reminiscence. The first two, accidental and natural signs, are not initially recognized as signs. Accidental signs are simply objects that have been experienced with certain circumstances, whereas natural signs are merely one's instinctual reactions to certain experiences. These two become signs only when they are actually associated with the circumstances or experiences. Instituted signs are those that one has chosen to induce thoughts. Though not required for the exercise of imagination and memory, the use of instituted signs allows one to have control over these operations. Such control in turn allows for the development of further rational operations such as abstraction and judgment that according to Condillac are not present in brute animals but are unique to humans.

In a 1752 letter to Maupertuis Condillac wrote that though he had tried to show in the *Essai* how the progress of the mind depends on language, "I was mistaken and gave too much to signs" (1947–1951, vol. 2, p. 536). The mistake here is indicated by Condillac's comment in a 1747 letter to Gabriel Cramer that his work was "not clear enough" on the point that natural and arbitrary signs "are the first principles of the development and progress of the operations of mind" (1953, p. 86). Condillac had of course indicated the importance of these kinds of signs in the *Essai*, but his mistake seems to have consisted in distinguishing them too greatly from instituted signs

involved in language. This would explain why he chose to focus in his *Traité* on the nature of sensation and mental operations before the start of language. There, even a statue without language is held to be capable of constructing a rich awareness of a spatially extended world on the basis of primitive sensory experience.

LANGUAGE AND ACTION

In the *Essai* Condillac criticized Locke for addressing the topic of words only after he had provided an account of ideas and mental operations. He insisted that the use of words is in fact "the principle that develops the seed of all our ideas" (2001, p. 8). Though the discussion in the *Traité* indicates that Condillac came to have a greater appreciation of one's prelinguistic abilities, he never relinquished the view that language is crucial for the development of mind. Whereas Descartes and Locke both suggested that thoughts or ideas are prior to and condition the use of language, Condillac insisted that it is the use of language that makes higher-order thoughts and mental operations possible. Here, one has a historical precedent for the "linguistic turn" in twentieth-century analytic philosophy.

Among the higher-order operations that require the use of language, Condillac singled out in particular a reflection that allows the mind to detach itself from current perceptions and apply itself to different objects. The *Essai* introduces the objection that the claim that this operation depends on language seems to be circular, since the use of instituted signs itself requires the abilities involved in reflection. Condillac responded to this objection that the nonlinguistic use of signs prepares the way for the mental operations required for the use of language and that these operations in turn make possible the development of reflection. He compared this relation between reflection and language to the discovery of algebraic signs by means of mental operations that had sufficient exercise to prepare the way for this discovery, but that were more primitive than the sort of mathematical thought that could not have occurred without this discovery (2001, II.i.§4, p. 115).

In the *Essai* Condillac claimed that spoken language derives from a "language of action" that involves voluntary control over nonlinguistic signs. He took the fact that such control develops over time to show that even the will derives from sensation. Still, he also seems to have indicated in the "Dissertation sur la liberté" that the freedom to direct attention is an original mental ability (1947–1951, vol. 1, p. 316). His sensationism thus appears to entail not that the will itself as a capacity derives from sensation, but that the employment of the capacity so derives. The employment of the will is made possible in particular by the habits that the instinctual use of natural and artificial signs produces.

Noam Chomsky claims to find in Antoine Arnauld and Claude Lancelot's *Grammaire générale et raisonnée, ou La grammaire de Port-Royal* a doctrine that posits innate "universal grammar" responsible for language (compare Arnauld and Lancelot 1966 and Chomsky 1966). The historical accuracy of this characterization is a matter of dispute (e.g., see the critical review of Chomsky in Aarsleff 1982, pp. 101–119), but what is undeniable is that Condillac offered an alternative to this sort of linguistics that attempts to explain language in terms of prelinguistic instincts and habits. This alternative was a particularly important influence for one of the classic texts in the field, Johann Gottfried Herder's *Abhandlung Über den Ursprung der Sprache* (1772).

See also Alembert, Jean Le Rond d'; Animal Mind; Arnauld, Antoine; Berkeley, George; Chomsky, Noam; Descartes, René; Diderot, Denis; Encyclopédie; Enlightenment; Experience; Holbach, Paul-Henri Thiry, Baron d'; La Mettrie, Julien Offray de; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Maupertuis, Pierre-Louis Moreau de; Newton, Isaac; Rousseau, Jean-Jacques; Sensationalism; Spinoza, Benedict (Baruch) de; Touch.

Bibliography

WORKS BY CONDILLAC

Oeuvres philosophiques. 3 vols., edited by Georges Le Roy. Paris: Presses Universitaires de France, 1947–1951.

Lettres inédites à Gabriel Cramer, edited by Georges Le Roy. Paris: Presses Universitaires de France, 1953.

Condillacou la joie de vivre, edited by Roger Lefévre. Paris: Seghers, 1966.

La logique/Logic, edited by W. R. Albury. New York: Abaris, 1980.

Les monades, edited by Laurence L. Bongie. Oxford, U.K.: Voltaire Foundation at the Taylor Institution, 1980.

Philosophical Writings of Étienne Bonnot, Abbé de Condillac. 2 vols. Translated by Franklin Philip. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982.

Essay on the Origin of Human Knowledge. Translated by Hans Aarsleff. New York: Cambridge University Press, 2001.

HISTORICAL WORKS

Arnauld, Antoine, and Claude Lancelot. *Grammaire générale et raisonnée, ou La grammaire de Port-Royal*, edited by Herbert E. Brekle. Studtgart-Bad: Frommann, 1966.

Chesselden, W. "An Account of Some Observations Made by a Young Gentleman Who Was Born Blind … and Was Couch'd between 13 and 14 Years of Age." In Royal Society

of London, *Philosophical Transactions* 35 (402) (April 1728): 447–450.
Diderot, Denis. "Letter on the Blind for the Use of Those Who See." In *Diderot's Early Philosophical Works*. Translated by Margaret Jourdain. New York: AMS Press, 1973.
Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers. 17 vols. Paris and Neufchâtel, Switzerland: Briasson et al, 1751–1765.
Lignac, Joseph Adrien Lelarge de. *Suite des "Lettres à un Américan" sur les IV et V Volumes de l'Histoire Naturelle" de M. de Buffon et sur le "Traité des Animaux" de M. l'abbé de Condillac*. 4 vols. Hamburg, Germany, 1756.

CONTEMPORARY WORKS
Aarsleff, Hans. *From Locke to Saussure: Essays on the Study of Language and Intellectual History*. Minneapolis: University of Minnesota Press, 1982.
Chomsky, Noam. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper and Row, 1966.
Derrida, Jacques. *The Archeology of the Frivolous: Reading Condillac*. Translated by John P. Leavy Jr. Lincoln: University of Nebraska Press, 1987.
Falkenstein, Lorne. "Étienne Bonnot de Condillac." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 1995–.
Knight, Isabel F. *The Geometric Spirit: The Abbé de Condillac and the French Enlightenment*. New Haven, CT: Yale University Press, 1968.
Lenois, Raymond. *Condillac*. Paris: F. Alcan, 1924.
Rousseau, Nicolas. *Connaissance et langage chez Condillac*. Geneva: Droz, 1986.
Sgard, Jean, ed. *Condillac et les problèmes du langage*. Geneva: Slatkine, 1982.
Sgard, Jean, ed. *Corpus Condillac, 1714–1780*. Geneva: Slatkine, 1981.

Tad M. Schmaltz (2005)

CONDITIONALS

Conditionals are sentences like the following:

(1) If Oswald did not kill Kennedy, then someone else did

(2) We will not go on the trip if it rains tomorrow

(3) If Oswald had not killed Kennedy, then someone else would have

(4) We would be playing tennis if it were not raining

Conditionals are often believed to be analyzable into a two-place sentence connective and two constituent sentences, the antecedent (the sentence introduced by "if") and the consequent. (Thus, [3] may be analyzed into a binary connective ["If it had been the case that … , then it would have been the case that …"] and the constituent sentences "Oswald did not kill Kennedy" and "someone else did [kill Kennedy].")

Many philosophers believe that there is an important difference between conditionals like (1) and (2) (which are commonly called "indicative conditionals"), and those like (3) and (4) (called "subjunctive" or "counterfactual"). Following Ernest W. Adams (1970), one can motivate this idea by considering (1) and (3). Suppose that you think that Oswald killed Kennedy, acting alone, and that no one else ever thought of committing this crime. You reject (3). But you accept (1): If you are wrong in thinking that Oswald did it, then someone else must be the culprit. Thus, it can be perfectly reasonable to assign different truth-values to the two conditionals. This indicates that an indicative conditional cannot in general have the same meaning as the corresponding counterfactual. Now suppose that this observation is combined with the suggested analysis of conditionals into two constituent sentences and a binary connective. Since (1) and (3) have the same constituent sentences, it is natural to conclude that their difference in meaning must result from a difference in meaning between the conditional connectives contained in the two sentences. The connective occurring in indicative conditionals, it seems, must differ semantically from the one found in counterfactuals.

This line of reasoning can be resisted in a number of ways. In particular, even if (1) and (3) belong to semantically different kinds of conditional, it is not obvious that the line between the two kinds coincides with that between indicative conditionals and counterfactuals. For all the argument shows, some indicative conditionals might have to be classified with (3) or some counterfactuals with (1), and under the influence of Vic H. Dudman (1984) some philosophers argue that indicative conditionals like (2) belong to the same class as (3).

Nonetheless, the standard view has it that conditionals are to be classified into indicatives and counterfactuals, and this entry will focus on theories that rest on this classification. The symbol "$\rightarrow$" will be used for the indicative and "$\Box\rightarrow$" for the counterfactual conditional connective.

INDICATIVE CONDITIONALS

Two of the main approaches to indicative conditionals will be considered.

THE EQUIVALENCE THESIS. Consider the mode of inference

(5) either B or not-A; therefore, if A, then B,

which is instantiated by the argument "Either the butler is guilty, or Fred lied about the ice pick. Therefore, if Fred said the truth, then the butler is guilty." This form of inference might appear to be valid. If it is, then an indicative conditional must be true whenever its antecedent is false and whenever its consequent is true. Moreover, it seems plausible that these are the only cases in which the conditional is true. It cannot be true if it has a true antecedent but a false consequent. (If someone says, "If it rains, she won't come," and it rains but she does come, then the utterance is not true.) This suggests that "$A \rightarrow C$" is true if and only if either A is false or C is true. In other words, "$A \rightarrow C$" has the same truth-conditions as the so-called material conditional, "$A \supset C$." This claim is sometimes called the "equivalence thesis."

It is well known that the equivalence thesis yields many seemingly absurd consequences. For instance, it makes (6) come out true, since (6) has a false antecedent and a true consequent:

(6) If Kennedy survived Oswald's assassination attempt, then he died in the assassination attempt.

Yet (6) does not seem to be assertable.

One strategy for dealing with such apparent counterexamples originates in work by Paul Grice (1991): According to the equivalence thesis, knowledge that A is false or that C is true is sufficient for knowing that "$A \rightarrow C$" is true. But if one's belief in the truth of a conditional rests solely on one's knowledge of the truth-values of its constituents (as in the case of [6]), then there is little point in asserting the conditional. For one could convey more information with fewer words by simply uttering the consequent, or the negation of the antecedent (as the case may be). If one utters the conditional anyway, then the audience, trusting the speaker not to do something pointless, will conclude that the speaker has reasons for believing the conditional that go beyond knowledge of the truth-values of its constituents. The utterance of the conditional would therefore be misleading, and the conditional, although true, is unassertable. When confronted with (6), one notes that it would be a mistake to assert it. This accounts for the feeling that there is something wrong with uttering the conditional. This impression can thus be explained without denying that the conditional is true.

The Gricean account has come in for criticism, but even if it is correct and apparent counterexamples to the equivalence thesis can be explained away, one may wonder whether the thesis is sufficiently well motivated. The previous argument for it rests on the assumption that the inference schema (5) is valid. But this premise has been questioned, because of apparent counterexamples to (5), such as "You will meet nobody, or at least not many people. Therefore, if you meet many people, then you will meet nobody."

The equivalence thesis can be supported in other ways, however: It is the simplest of all candidate truth-conditional theories of indicative conditionals. And Frank Jackson (1987, chapter 2) argues that, although the equivalence theorist must concede that an indicative conditional's degree of assertability can differ from its probability of truth, the equivalence thesis can be used to explain the assertability-conditions and can be supported by appeal to this explanatory power.

THE RAMSEY TEST. Another approach to the semantics of indicative conditionals originates in a footnote in a paper by Frank P. Ramsey (1990, p. 155, n. 1) and has been developed in detail by Adams (1975). It starts from the idea, which is sometimes called the "Ramsey test," that the degree to which a speaker ought to accept "$A \rightarrow C$" equals the person's subjective conditional probability $P(C|A)$ (i.e., $P(A \text{ and } C) / P(A)$), provided that $P(A)$ is not zero so that $P(C|A)$ is defined. (On other versions of this account, $P[C|A]$ measures the degree to which the speaker should regard the conditional as assertable. The discussion below will focus on the acceptability-conditional version of the thesis.) This hypothesis is strongly supported by its ability to predict pre-theoretical intuitions about individual conditionals. Suppose that I am about to cast a fair die. My probability that I will throw a six given that I will throw an even number is one-third, and this is also the degree to which I accept, "I will throw a six if I throw an even number."

One might be tempted to try to explain why the degree of acceptability of "$A \rightarrow C$" equals $P(C|A)$ by the assumption that

(7) a conditional "$A \rightarrow C$" expresses a proposition, and the probability that this proposition is true equals $P(C|A)$ in all probability distributions for which $P(C|A)$ is defined.

However, David K. Lewis shows that (7) is false (he proves this and some stronger results in his 1991a and 1991b). Instead of stating the proof, this entry will point in a nonrigorous and informal way in the direction of the reason why (7) is false (this seems more intuitively helpful than a formal proof):

Let each point of the rectangle in Figure 1 stand for a possible world, and let the rectangle as a whole represent

FIGURE 1

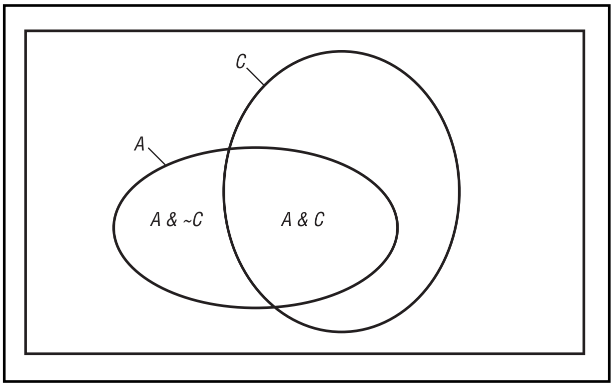

the totality of possible worlds. Propositions can be represented by the regions containing all and only the points that stand for worlds in which these propositions are true. One can model a belief system by distributing one kilogram of mud over the rectangle: If P(*X*) equals *p* in the belief system one intends to represent, then one places *p* kilograms of mud on the region representing the proposition *X*. Every possible way of distributing the mud corresponds to some probability distribution. If (7) were true, then there would have to be some region, namely the one representing the proposition expressed by "*A* → *C*," that bears an amount of mud equal to P(*C*|*A*), that is, to the ratio of the amount on the *A* & *C* region and that on the *A* region, whenever P(*A*) is not zero. However, it is easy to make it plausible that there is no such region. Assume that P(*C*|*A*) equals one-half, which is to say that there is the same (nonzero) amount of mud on the *A* & *C* region as on the *A* & ~*C* region. This is information about the relative amounts of mud on the two regions, and as such it tells one next to nothing about the absolute amount on any specific region. In particular, it seems intuitively plausible that, contrary to (7), there is no region that must be loaded with exactly half a kilogram of mud whenever the *A* & *C* region and the *A* & ~*C* region bear the same (nonzero) amount of mud.

These considerations suggest that there is no one region whose amount of mud equals the ratio of the amount of *A* & *C* mud and the amount of *A* mud whenever this ratio is defined. However, it might be that, whenever the mud is distributed in such a way that the ratio is defined, there is some region whose amount of mud equals the ratio, though it is a different region in different cases. (Note the difference in the scopes of the quantifiers.) Hence, as Bas van Fraassen (1976) points out, for all the argument of the last paragraph shows, it could be that an indicative conditional "*A* → *C*" expresses a proposition and that its probability of truth equals P(*C*|*A*) whenever this conditional probability is defined, but that the proposition expressed by the conditional varies systematically with the speaker's belief system. Philosophers have attempted to extend Lewis's proof so as to rule out this possibility.

As an alternative to finding truth-conditions that fit the Ramsey test, one might give up the idea that indicative conditionals express propositions and make the Ramsey test itself the centerpiece of one's semantic account. Such a theory raises two questions: (1) What account can be given of the meanings of compound sentences that embed indicative conditionals, such as "Either Fred will give you the money if you ask him, or he is more avaricious than Susie"? If indicative conditionals lack truth-values, then one cannot assign a meaning to this sentence using the usual truth-functional construal of the disjunction operator. However, as Allan Gibbard (1981, pp. 234–236) argues, that a nonpropositional account of indicative conditionals does not assign meanings to all compounds of conditionals might be a good thing. For many such compounds are so hard to understand that one may doubt that they have any clear meanings. (Consider "If Fred arrived yesterday if it rains tomorrow, then Susie was in Paris last week.") The thesis that indicative conditionals lack truth-conditions may explain such difficulties of interpretation. (2) The usual criterion for the acceptability of an inference form relates to whether it preserves truth, that is, to whether the conclusion of an instance of it must be true if the premises are true. If indicative conditionals cannot be true or false, then this criterion cannot be applied to inferences involving such conditionals. Adams (1975, chapter 2) tackles this problem by defining a new and independently motivated criterion of acceptability that is more widely applicable. According to this criterion, an inference must preserve probability, in a sense that Adams makes precise as follows: Call 1–P(*A*) the "uncertainty" of the proposition *A*; the uncertainty of a conditional "*A* → *C*" equals 1–P(*C*|*A*). An inference preserves probability just in case there is no probability distribution in which the uncertainty of the conclusion exceeds the sum of the uncertainties of the premises. Classically valid arguments satisfy this condition, as do intuitively acceptable inferences involving indicative conditionals.

COUNTERFACTUALS

Counterfactuals are used to analyze a wide range of philosophically important concepts, such as dispositions, causation, laws of nature, knowledge, practical rationality

(counterfactuals are used in decision theory), and freedom of action ("She would have acted differently if she had chosen to do so"). Theories of counterfactuals are of interest in part because they may make it easier to understand and evaluate counterfactual accounts of other notions.

GOODMAN'S ACCOUNT. In the seminal paper "The Problem of the Counterfactual Conditional" (1991) Nelson Goodman proposes an account of roughly the following form for a certain important class of counterfactuals:

> (8) "$A \Box\!\!\rightarrow C$" is true just in case C follows from A, the laws of nature, and suitable true supplementary premises.

This account fits the ordinary-life practice of evaluating counterfactuals well. In determining what would have happened to a certain match if it had been struck on a specific occasion, one needs to draw on knowledge of the particular circumstances, such as the knowledge that (D) the match was dry and (O) oxygen was present, and of the law that (L) dry matches start to burn when struck in the presence of oxygen. These items of knowledge, when combined with the assumption that the match was struck, entail that (B) it burned. This justifies the conclusion that the match would have burned if it had been struck.

Which truths count as "suitable supplementary premises" in the sense of (8)? Clearly, not every truth does. When evaluating the counterfactual "If the match had been struck … ," one cannot regard the truth that it was never struck as a suitable ancillary premise. More generally, if the antecedent is both self-consistent and consistent with the laws, then the suitable auxiliary premises must be consistent with the antecedent plus laws. Otherwise, the antecedent combined with the laws and the supplementary premises would entail everything, so that every counterfactual with the relevant antecedent would come out true—an unwelcome result if the antecedent is consistent.

This condition of consistency does not suffice as a criterion for the suitability of a truth as ancillary premise. For there are different sets of truths that meet the consistency constraint, and depending on which of them one regards as the set of suitable auxiliary premises, different counterfactuals come out true. If one uses (D) and (O) as supplementary premises in evaluating the conditional about the match, one can draw on one's knowledge that (L) is a law to conclude that the match would have burned if it had been struck. Availing oneself instead of (O) and (~B) as auxiliary premises, one can (again using [L]) establish that the match would not have been dry if it had been struck.

The task of stating conditions for a truth's suitability as supplementary premise is central to Goodman's project. After discussing the issue at length, he ends up proposing that a truth *P* is suitable only if *P* is cotenable with the antecedent of the conditional, where this means: It is not the case that *P* would have been false if the antecedent had been true. (For instance, [~B] is not cotenable with "the match was struck," since if the latter sentence had been true, (~B) would have been false. But [D] and [O] would still have been true and are therefore cotenable.) Since this criterion is formulated in counterfactual terms, it renders Goodman's theory circular—a problem of which Goodman is vividly aware. As will become clear below, more recent work on counterfactuals promises to deliver a solution.

THE POSSIBLE-WORLD ACCOUNT. In the late 1960s and early 1970s another account of counterfactual conditionals was developed by Robert C. Stalnaker (1991b) and Lewis (1973). Lewis neatly expresses the core idea: "'*If kangaroos had no tails, they would topple over*' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over" (1973, p. 1).

More formally, the theory is formulated in terms of possible worlds. A possible world in which the antecedent of a counterfactual is true is called an "antecedent-world." One can state the theory (in a somewhat simplified form) by saying that a counterfactual is true just in case its consequent is true in those antecedent-worlds that are most similar to the actual world.

Stalnaker's and Lewis's accounts differ in a number of ways. First, Stalnaker intends his theory to cover both indicative conditionals and counterfactuals, whereas the scope of Lewis's account is restricted to counterfactuals. Second, according to Stalnaker's truth-conditions, but not according to Lewis's, there is always one most similar possible antecedent-world. In consequence, Stalnaker's theory validates the principle of conditional excluded middle, ($A \Box\!\!\rightarrow C$) or ($A \Box\!\!\rightarrow {\sim}C$), whereas Lewis's account does not.

It is an advantage of the possible-world account that it can explain some noteworthy logical features of counterfactuals, namely the failure of a number of inference

FIGURE 2

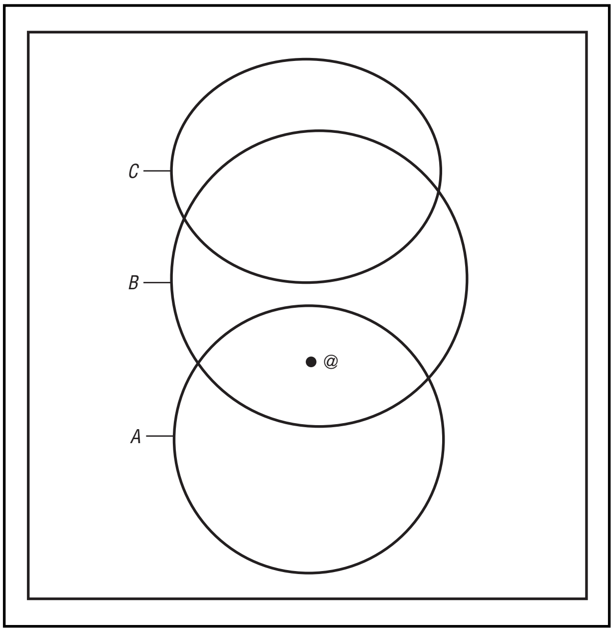

schemata that are valid for the material and strict conditionals, such as the following:

$A \Box\!\!\rightarrow B \therefore \sim B \Box\!\!\rightarrow \sim A$	(Contraposition)
$C \Box\!\!\rightarrow B, B \Box\!\!\rightarrow A \therefore C \Box\!\!\rightarrow A$	(Hypothetical syllogism)
$B \Box\!\!\rightarrow A \therefore (B \,\&\, C) \Box\!\!\rightarrow A$	(Strengthening the antecedent)

To see that these modes of inference are invalid, consider the following counterexamples:

- (Even) if Mary had qualified for the tournament, she (still) would not have won it. Therefore, if she had won the tournament, she (still) would not have qualified for it.
- If Hoover had been born a Russian, he would have been a communist. If Hoover had been a communist, he would have been a traitor. Therefore, if Hoover had been born a Russian, he would have been a traitor (Stalnaker 1991b).
- If we had told them about our plan, they would have been delighted. Therefore, if we had told them about our plan and its likely result, they would have been delighted.

The possible-world account explains the failure of these inference rules, as Figure 2 will render clear. Let the dot labeled "@" stand for the actual world, let the other points in the rectangle represent the other possible worlds, and let the smaller and greater spatial distances between the points represent smaller and greater degrees of similarity between the corresponding worlds. As before, propositions can be represented by regions within the rectangle. In the situation depicted, *B* is true in the possible *A*-worlds most similar to the actual world; but ~*A* is not true in the most similar possible ~*B*-worlds. Hence, while "$A \Box\!\!\rightarrow B$" is true, "$\sim B \Box\!\!\rightarrow \sim A$" is false. This shows that contraposition is invalid. Moreover, "$C \Box\!\!\rightarrow B$" and "$B \Box\!\!\rightarrow A$" are true while "$C \Box\!\!\rightarrow A$" is false, and "$B \Box\!\!\rightarrow A$" is true while "$(B \,\&\, C) \Box\!\!\rightarrow A$" is false, so that the diagram also represents counterexamples to hypothetical syllogism and strengthening the antecedent.

If the antecedent of a counterfactual is impossible, then there are no possible antecedent-worlds, so that it is vacuously true that the consequent is true in all the most similar possible antecedent-worlds. The possible-world account therefore entails that all counterfactuals with impossible antecedents are true. But that is implausible: Most philosophers would agree that Willard Van Orman Quine could not have been a hippopotamus, but it does not seem right to say that, if Quine had been a hippopotamus, he would have been a reptile. According to Daniel Nolan (1997) and others, this problem can be remedied if impossible worlds are allowed to figure in the account alongside possible worlds. On this view, impossible worlds are ordered by their comparative similarity to the actual world, just as possible worlds are. A counterfactual "$A \Box\!\!\rightarrow C$" with impossible antecedent is true just in case *C* is true in the most similar impossible *A*-worlds. Such an account, however, requires an ontology of impossible worlds, which not all philosophers are happy to accept.

SIMILARITY BETWEEN WORLDS. The notion of similarity between worlds that is used in the analysis of counterfactuals cannot be the one that governs ordinary offhand judgments about overall similarity. This was shown by Kit Fine (1975) among others. Fine used (9) as his example:

(9) If Nixon had pressed the button, there would have been a nuclear catastrophe.

(9) sounds correct. But offhand it may seem that an antecedent-world devastated by a nuclear explosion is much less similar to the actual world than an antecedent-world in which the signal disappears in the wire after the button-pressing, so that no harm is done. If the notion of offhand similarity were used in analyzing counterfactuals, the account would yield the incorrect verdict that (9) is false and that everything would have been fine if Nixon had pressed the button.

What are the standards of similarity that govern counterfactuals? Many philosophers who address this question assume that different standards are relevant in different contexts of utterance. This assumption is motivated by examples like the following (which is taken from Jackson 1977, p. 9): Frank is in a room on the tenth floor of a building. There is nothing that could break the fall of someone jumping out of the window. It seems safe to say that Frank would get badly hurt if he were to jump. But suppose that Frank says: "I would never jump from a tenth-floor window, unless I had made sure that there was a safety net. So, if I were to jump, a net would be in place, and I would be fine." Frank's reasoning might convince his audience that his counterfactual is true. And yet his conditional seems to be incompatible with the one stated before. The most obvious diagnosis is that the truth-conditions of counterfactuals are context-dependent. In some contexts worlds in which Frank jumps despite the absence of a net count as more similar than those in which he places a net below the window before jumping. In other contexts it is the other way around.

Some of those who take the truth-conditions of counterfactuals to be context-dependent (notably Lewis 1979), believe that there is such a thing as a default assignment of truth-conditions to them, an assignment that hearers choose when interpreting the utterance of a counterfactual unless their presumption in favor of it is removed by distinctive features of the context. That seems plausible enough in the example of the last paragraph: If presented with the case out of the blue and asked for a judgment, one would say that Frank would get badly hurt if he were to jump. It requires some stage-setting (like that provided by Frank's utterance) to create a context in which it seems right to say that he would be fine.

Attempts to describe the default truth-conditions of counterfactuals often start from a special case: counterfactuals whose antecedents are false and describe nomically possible matters of particular local fact. (9) can serve as an example. Pre-theoretical intuitions about this conditional seem to furnish two data points:

(1) *Counterfactual dependence is temporally asymmetrical*. If Nixon had pressed the button, then later on things would have been different from what they were actually like; but matters until shortly before the button-pressing would have been just as they actually were. The most similar antecedent-worlds must therefore be just like the actual world until a short time before the button-pressing, but might be different afterward.

(2) *Laws support counterfactuals*. If Nixon had pressed the button, then events would still have conformed to the actual laws of nature. The most similar antecedent-worlds must therefore be ones that evolve in accordance with the laws of the actual world. In particular, if the missile system is set up in such a way that the actual laws guarantee that button-pressing leads to a nuclear explosion, then there is a nuclear catastrophe in the most similar antecedent-worlds.

Suppose that determinism is true. In that case at least one of the principles (1) and (2) stands in need of some qualification. For determinism entails that every initial segment of the history of the actual world, together with the laws, determines the entire rest of history, and thus determines that Nixon does not press the button. This implies that no antecedent-world can both perfectly conform to the actual laws and be like the actual world throughout some initial segment of its history. Some philosophers (e.g., Lewis 1979) choose to solve this problem by allowing that the most similar antecedent-worlds contain violations of the actual laws, while others allow for backward counterfactual dependence over arbitrarily long periods of time (e.g. Bennett 1984; but see Bennett 2003, §80).

Note that Goodman's problem of specifying which truths are suitable supplementary premises resurfaces on the possible-world theory, in the shape of the question: Which of the actual matters of particular fact must obtain in an antecedent-world for it to count among the most similar? An account of the similarity relation will address this question and, if successful, will at last provide a non-circular solution to Goodman's problem.

See also Entailment, Presupposition, Implicature; Modal Logic; Paraconsistent Logics; Relevance (Relevant) Logics.

Bibliography

Adams, Ernest W. *The Logic of Conditionals: An Application of Probability to Deductive Logic*. Dordrecht, Netherlands: D. Reidel, 1975.

Adams, Ernest W. "Subjunctive and Indicative Conditionals." *Foundations of Language* 6 (2) (1970): 89–94.

Bennett, Jonathan. "Classifying Conditionals: The Traditional Way Is Right." *Mind* 104 (414) (1995): 331–354.

Bennett, Jonathan. "Counterfactuals and Temporal Direction." *Philosophical Review* 93 (1) (1984): 57–91.

Bennett, Jonathan. "Farewell to the Phlogiston Theory of Conditionals." *Mind* 97 (388) (1988): 509–527.

Bennett, Jonathan. *A Philosophical Guide to Conditionals*. New York: Oxford University Press, 2003.

Dudman, Vic H. "Parsing 'If'-Sentences." *Analysis* 44 (1984): 145–153.

Edgington, Dorothy. "On Conditionals." *Mind* 104 (414) (1995): 235–329.

Fine, Kit. "Review of Lewis's *Counterfactuals*." *Mind* 84 (335) (1975): 451–458.

Gibbard, Allan. "Two Recent Theories of Conditionals." In *Ifs: Conditionals, Belief, Decision, Chance, and Time*, edited by William L. Harper, Robert C. Stalnaker, and Glenn Pearce, 211–247. Dordrecht, Netherlands: D. Reidel, 1981.

Goodman, Nelson. "The Problem of the Counterfactual Conditional." In *Conditionals*, edited by Frank Jackson, 9–27. New York: Oxford University Press, 1991.

Grice, Herbert Paul. "Logic and Conversation." In *Conditionals*, edited by Frank Jackson, 155–176. New York: Oxford University Press, 1991.

Harper, William L., Robert C. Stalnaker, and Glenn Pearce, eds. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Dordrecht, Netherlands: D. Reidel, 1981.

Jackson, Frank. "A Causal Theory of Counterfactuals." *Australasian Journal of Philosophy* 55 (1977): 3–21.

Jackson, Frank. "Classifying Conditionals." *Analysis* 50 (1990): 134–147.

Jackson, Frank. *Conditionals*. Oxford, U.K.: Basil Blackwell, 1987.

Jackson, Frank, ed. *Conditionals*. New York: Oxford University Press, 1991.

Lewis, David K. "Counterfactual Dependence and Time's Arrow." *Nous* 13 (1979): 32–66.

Lewis, David K. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

Lewis, David K. "Probabilities of Conditionals and Conditional Probabilities." In *Conditionals*, edited by Frank Jackson, 76–101. New York: Oxford University Press, 1991a.

Lewis, David K. "Probabilities of Conditionals and Conditional Probabilities II." In *Conditionals*, edited by Frank Jackson, 102–110. New York: Oxford University Press, 1991b.

Nolan, Daniel. "Impossible Worlds: A Modest Approach." *Notre Dame Journal of Formal Logic* 38 (4) (1997): 535–572.

Ramsey, Frank P. "Law and Causality" (1929). In his *Philosophical Papers*, edited by D. H. Mellor. Cambridge, U.K.: Cambridge University Press, 1990.

Stalnaker, Robert C. "A Defense of Conditional Excluded Middle." In *Ifs: Conditionals, Belief, Decision, Chance, and Time*, edited by William L. Harper, Robert C. Stalnaker, and Glenn Pearce, 87–104. Dordrecht, Netherlands: D. Reidel, 1981.

Stalnaker, Robert C. "Indicative Conditionals." In *Conditionals*, edited by Frank Jackson, 136–154. New York: Oxford University Press, 1991a.

Stalnaker, Robert C. "Letter to van Fraassen." In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Vol. 1, *Foundations and Philosophy of Epistemic Applications of Probability Theory*, edited by William L. Harper and Clifford A. Hooker. Dordrecht, Netherlands: D. Reidel, 1976.

Stalnaker, Robert C. "A Theory of Conditionals." In *Conditionals*, edited by Frank Jackson, 28–45. New York: Oxford University Press, 1991b.

Van Fraassen, Bas. "Probabilities of Conditionals." In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Vol. 1, *Foundations and Philosophy of Epistemic Applications of Probability Theory*, edited by William L. Harper and Clifford A. Hooker. Dordrecht, Netherlands: D. Reidel, 1976.

Boris C. A. Kment (2005)

CONDORCET, MARQUIS DE

(1743–1794)

Marie-Jean-Antoine-Nicolas Caritat, Marquis de Condorcet, the French mathematician, historian of the sciences, political theorist, and social reformer, was one of the youngest of the Encyclopedists and the only prominent one to participate actively in the French Revolution. He was born in Ribemont in Picardy and was educated by the Jesuits at the Collège de Navarre. Admitted to the Académie des Sciences in 1769 on the basis of his early mathematical writings, he was elected its perpetual secretary in 1776 and ably depicted the progress of the sciences to a wide public in the customary eulogies (*Éloges*) of deceased academicians, which he presented in this position.

A protégé of Jean Le Rond d'Alembert, for whom Condorcet's election to the Académie Française in 1782 was regarded as a personal triumph, and of Baron de l'Aulne Turgot, who called him to the directorship of the mint during his abortive reforming ministry, Condorcet was active in the prerevolutionary campaigns for economic freedom, religious toleration, legal reform and the abolition of slavery. After his marriage to Sophie de Grouchy in 1786 their salon became one of the most brilliant and influential of the prerevolutionary period. He took part in the opening debates of the French Revolution as a member of the municipal council of Paris and was a convinced republican by the time he was elected to the Legislative Assembly in 1791. Prominent in this assembly, he directed his most sustained efforts toward the elaboration of a project for public education that had great influence on the eventual establishment of the French educational system.

In the National Convention, Condorcet's opposition to the death penalty led him to cast his vote against the execution of Louis XVI (he voted for the supreme penalty short of death). He then undertook the task of drawing up a draft constitution for the new republic, but although accepted by the committee on the constitution, his liberal

constitutional scheme—commonly known as the Girondin constitution of 1793—shared the unfortunate fate of the group with which it was associated. In July 1793, Condorcet's indignant defense of his constitution against that prepared by the Jacobins led to his denunciation and flight into hiding. He spent his remaining months of life secluded in Paris, working on the *Sketch for a Historical Picture of the Progress of the Human Mind* (*Esquisse d'un tableau historique des progrès de l'esprit humain*), published posthumously in 1795. He left his asylum in March 1794 and was arrested and imprisoned at Bourg-la-Reine, near Paris. He died during the first night of his imprisonment, either from exhaustion or from a self-administered poison.

PROBABILITY AND SOCIAL SCIENCE

It has often been assumed that Condorcet's increasing preoccupation with social and political affairs, if not the result of a sense of frustration with his mathematical investigations, was at least accompanied by a waning interest in them. Quite the reverse is true. Condorcet's experience at the Académie des Sciences fostered a sense of the power of science to elucidate even the realm of social behavior. His mathematical endeavors were intimately bound up with his fundamental intellectual concern. He aimed to bring to social questions the attitudes and methods of the physical sciences, thereby welding the broken elements of the moral and political sciences into a new social science, which he regarded as the necessary condition of a rational political and social order.

Condorcet seized upon the calculus of probabilities as the essential epistemological connection between the physical sciences and the science of man. All the truths of experience are merely probable, he argued. In the social sciences the observation of facts may be more difficult and their order less constant. The results of the social sciences may therefore be less probable than those of the physical sciences. But Condorcet maintained that the probability of all statements of experience can be expressed and evaluated mathematically within probability theory. Thus, while the statements attained by the social sciences may on occasions be less probable than those of the physical sciences, in Condorcet's view the mathematical estimate of their respective probabilities is equally certain. The meteorologist cannot be certain that it will rain tomorrow, for example, but if on the basis of his observations he can estimate the probability of its doing so as x:1, then he can be certain that there is a probability of x:1 that it will rain tomorrow. Similarly, the economist, who cannot be certain that the standard of living will continue to rise, can in theory arrive at a certain mathematical estimate of the probability of its doing so.

The significance of this argument can be best assessed in terms of the earlier epistemological claims to certainty made by René Descartes on behalf of the mathematical and physical sciences. Condorcet accepted the skeptic's evaluation of the physical sciences as being merely probable. But in arguing that probabilities in the physical sciences (like those in the social sciences) can be evaluated with mathematical certainty, he remained in a sense fundamentally Cartesian. Not only did he hold to the idea of certainty as the criterion of acceptable knowledge, but he also accepted mathematics as the paradigm of certain knowledge (although even this certainty is based in the last analysis, he was occasionally prepared to argue, on the observed constancy of the operation of the human mind). Condorcet's argument in this respect ranks with that of Giambattista Vico as one of the major eighteenth-century attempts to establish the validity of social science. But whereas Vico turned away from the mathematical and physical sciences in search of a historical and organic conception of his new science, Condorcet's probabilistic evaluation of the physical sciences served to integrate them with the science of man in an essentially mathematical conception of science. For Condorcet, the mathematician was able, by using the calculus of probabilities, to subject to the certain evaluation of mathematics even those areas of knowledge condemned by Descartes as untrustworthy. The calculus of probabilities provided a sure means of estimating the validity of our opinions and the probability of our expectations; it bound the moral and physical sciences together on a sliding scale of probabilities which could at all stages be evaluated with mathematical certainty.

Condorcet developed this conception in two very different works. In the first, the *Essay on the Application of Analysis to the Probability of Majority Decisions* (*Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, 1785), he set out to discover by means of the calculus of probabilities under what conditions there will be an adequate guarantee that the majority decision of an assembly or tribunal is true. In one of its applications he envisaged such an analysis as the means of solving a perennial problem of liberal thought, that of reconciling the claims of an elite to exercise special responsibilities in the process of decision making with the general principle of universal or majority consent. But the obscure mathematics of the essay and its inevitable reliance on unverifiable assumptions as to the probable

truth or error of the opinions of individuals composing social bodies have left it largely ignored by those interested in Condorcet's political theory. More recently, social mathematicians interested in elucidating the relationship between individual and collective choice (whether political or economic) have been able to disengage from the probabilistic framework of this work a theoretical model of collective decision making that is remarkably modern in its implications and approach. (See Black [1958] and Granger [1954]).

The *Essai sur l'application de l'analyse* was intended to convince academicians of the validity of Condorcet's contention that the moral and political sciences can be treated mathematically. The unfinished "Tableau général de la science, qui a pour objet l'application du calcul aux sciences morales et politiques" (General View of the Science Comprising the Mathematical Treatment of the Moral and Political Sciences) was meant for a different audience. It appeared in 1793 in a popular journal that sought to initiate citizens of the new French republic into the social science, or the art of the rational conduct of politics. Condorcet saw the new social mathematics (*mathématique sociale*) as a common, everyday science of conduct ("une science usuelle et commune," *Oeuvres*, Vol. I, p. 550) that would provide the essential foundation of a democratic, but rational, politics. He viewed man in all his conduct as a gambler. Each individual automatically and instinctively balances the probability of one opinion against that of another, the desired goal of a proposed action against its probable results. The mathematical science of man was intended not only as an objective description of social behavior but also as a scientific basis for individual conduct that would enable people to substitute for habitual and instinctive modes of thought and action the precise evaluation of reason and calculation. Social mathematics, coupled with an exact language based on precise philosophical analysis of our ideas, would free human beings from instinct and passion and restore the empire of reason in social affairs. It formed the essential link between scientific advance and moral progress, for evil, as Condorcet remarked, was far more often the result of an erroneous calculation of interest than the product of violent passion.

IDEA OF PROGRESS

In the *Sketch for a Historical Picture of the Progress of the Human Mind*, Condorcet turned to history for a demonstration of the power of reason and calculation in social affairs. The *Sketch* was only the hastily written introduction to a larger work on the history of science and its impact upon society which Condorcet had been contemplating for many years. Some of the fragments of this unfinished work are of considerable philosophical interest. One outlined a project for a universal, symbolic language of the sciences; another elaborated a decimal system of classification addressed to the much-debated problem of scientific classification. But it is with the *Sketch* itself that Condorcet's name and influence have been chiefly associated, and it is with that work—often regarded as the philosophical testament of the eighteenth century—that Condorcet bequeathed to the nineteenth century the fundamental idiom of its social thought, the idea of progress.

The aim of the *Sketch* was to demonstrate man's progressive emancipation, first from the arbitrary domination of his physical environment and then from the historical bondage of his own making. Condorcet shared with other eighteenth-century theorists a view of progress that depended ultimately upon man's cumulative ability to combine sensations and ideas (in the manner revealed by sensationalist psychology) to his own satisfaction or advantage. This Promethean psychological capacity functioned in the same manner in the human race as in the individual; it proceeded by way of a natural, self-revealing logic or "method," from the fundamental data of sense experience to the most general principles of the moral and physical sciences. Condorcet's main concern, therefore, was less to explain the growth of reason in itself—this growth was posited as natural—than to point to the destruction of the obstacles that had inhibited that growth or diverted the historical development of the mind from the natural logic of ideas.

Condorcet's hopes for future progress rested on two conclusions. First, he was convinced that the obstacles which had in the past threatened the advance and dissemination of reason—elitism and tyranny on the one hand; popular prejudice, ignorance, and social and political subjection, on the other—were finally being destroyed under the joint impact of scientific, technological, and political revolution. Second, he believed that the discoveries of sensationalist psychology had made it possible to articulate the natural and fundamental principles of the social art, or science, and he drew from the doctrine of the rights of man—grounded upon the "facts" of man's sensate nature—a comprehensive outline of the principles of liberal democracy that it would be the purpose of the social art to implement.

Although this belief in indefinite future progress was based on the general assertion that observation of past events warrants extrapolation as to the probable future,

Condorcet was not a strict historical determinist. Humans are subject to the general laws of physical nature, he maintained in an unpublished introduction to the *Sketch*, but they have the power to modify these laws and turn them to their own advantage. Although this power is feeble in the individual, when exercised by humankind collectively and over a long period, it can balance the forces of nature and can even be regarded as the work of nature itself. For if nature has endowed humankind collectively with the capacity to learn from experience, to understand its laws, and to modify their effects, the progressive emancipation of humans from nature is itself natural, and the growth of freedom is a natural law. The *Sketch* not only demonstrated the power of the social art but also made clear that it could succeed only as a communal and democratic art. It is this emphasis upon the collective experience and achievements of humankind, this concern with the "most obscure and neglected chapter of the history of the human race" (*Sketch*, Barraclough translation, p. 171)—namely, the progress of the mass of the people in society—that links Condorcet's view of history with his conception of social science.

See also Alembert, Jean Le Rond d'; Descartes, René; Encyclopédie; Mathematics, Foundations of; Philosophy of History; Progress, The Idea of; Turgot, Anne Robert Jacques, Baron de L'Aulne; Vico, Giambattista.

Bibliography

The standard edition of Condorcet's works is A. Condorcet-O'Connor and F. Arago, eds., *Oeuvres de Condorcet*. 12 vols. (Paris, 1847–1849). The *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Paris: De l'Imprimerie Royale, 1785) and Condorcet's many other mathematical writings are not included in this edition. For a bibliography of these mathematical works see Charles Henry, "Sur la Vie et les écrits mathématiques de Jean-Antoine-Nicolas Caritat marquis de Condorcet." In *Bulletino di bibliografica e di storia delle scienze matematiche e fisiche* 16 (1883): 271–324. See also Charles Henry, ed., *Correspondance inédite de Condorcet et de Turgot (1770–79)* (Paris: Charavay Frères, 1883); Léon Cahen, "Condorcet inédit: Notes pour le Tableau des progrès de l'esprit humain," in *La révolution française* 75 (1922): 193–212; Gilles-Gaston Granger, "Langue universelle et formalisation des sciences: Un Fragment inédit de Condorcet," in *Revue d'histoire des sciences* 7 (1954): 197–219; Alberto Cento, "Dei manoscritti del 'Tableau' di Condorcet," in *Rendiconto del Istituto Lombardo di Scienze e Lettere: Classe di Lettere e Scienze Morali e Storiche* 88 (1955): 311–324; K. M. Baker, "An Unpublished Essay of Condorcet on Technical Methods of Classification," in *Annals of Science* 18 (1962): 99–123. For a modern translation of the *Esquisse*, see *Sketch for a Historical Picture of the Progress of the Human Mind*, translated by June Barraclough, with an introduction by Stuart Hampshire (London: Weidenfeld and Nicolson, 1955).

For reference see Franck Alengry, *Condorcet, guide de la Révolution française, théoricien du droit constitutionnel et précurseur de la science sociale* (Paris: V. Giard and E. Brière, 1903); Duncan Black, *The Theory of Committees and Elections* (Cambridge, U.K.: Cambridge University Press, 1958); Janine Bouissounouse, *Condorcet, le Philosophe dans la Révolution* (Paris: Hachette, 1962); Léon Cahen, *Condorcet et la Révolution française* (Paris: Alcan, 1904); Alberto Cento, *Condorcet e l'idea di progresso* (Florence: Parenti, 1956); Gilles-Gaston Granger, *La mathématique sociale du marquis de Condorcet* (Paris, 1956); Frank E. Manuel, *The Prophets of Paris* (Cambridge, MA: Harvard University Press, 1962); and J. Salwyn Schapiro, *Condorcet and the Rise of Liberalism* (New York: Harcourt Brace, 1934).

See also Keith M. Baker, *Condorcet, From Natural Philosophy to Social Mathematics* (Chicago: University of Chicago Press, 1975); and Edward Goodell, *The Noble Philosopher: Condorcet and the Enlightenment* (Buffalo, NY: Prometheus, 1994).

Keith Michael Baker (1967)
Bibliography updated by Tamra Frei (2005)

CONFIRMATION THEORY

Predictions about the future and unrestricted universal generalizations are never logically implied by our observational evidence, which is limited to particular facts in the present and past. Nevertheless propositions of these and other kinds are often said to be confirmed by observational evidence. A natural place to begin the study of confirmation theory is to consider what it means to say that some evidence *E* confirms a hypothesis *H*.

INCREMENTAL AND ABSOLUTE CONFIRMATION

Let us say that *E* raises the probability of *H* if the probability of *H* given *E* is higher than the probability of *H* not given *E*. According to many confirmation theorists, "*E* confirms *H*" means that *E* raises the probability of *H*. This conception of confirmation will be called incremental confirmation.

Let us say that *H* is probable given *E* if the probability of *H* given *E* is above some threshold. (This threshold remains to be specified but is assumed to be at least one half.) According to some confirmation theorists, "*E* confirms *H*" means that *H* is probable given *E*. This conception of confirmation will be called absolute confirmation.

Confirmation theorists have sometimes failed to distinguish these two concepts. For example, Carl Hempel

(1945/1965) in his classic "Studies in the Logic of Confirmation" endorsed the following principles:

(1) A generalization of the form "All *F* are *G*" is confirmed by the evidence that there is an individual that is both *F* and *G*.

(2) A generalization of that form is also confirmed by the evidence that there is an individual that is neither *F* nor *G*.

(3) The hypotheses confirmed by a piece of evidence are consistent with one another.

(4) If *E* confirms *H* then *E* confirms every logical consequence of *H*.

Principles (1) and (2) are not true of absolute confirmation. Observation of a single thing that is *F* and *G* cannot in general make it probable that all *F* are G; likewise for an individual that is neither *F* nor *G*. On the other hand there is some plausibility to the idea that an observation of something that is both *F* and *G* would raise the probability that all *F* are *G*. Hempel argued that the same is true of an individual that is neither *F* nor *G*. Thus Hempel apparently had incremental confirmation in mind when he endorsed (1) and (2).

Principle (3) is true of absolute confirmation but not of incremental confirmation. It is true of absolute confirmation because if one hypothesis has a probability greater than ½ then any hypothesis inconsistent with it has a probability less than ½. To see that (3) is not true of incremental confirmation, suppose that a fair coin will be tossed twice, let H_1 be that the first toss lands heads and the second toss lands tails, and let H_2 be that both tosses land heads. Then H_1 and H_2 each have an initial probability of ¼. If *E* is the evidence that the first toss landed heads, the probability of both H_1 and H_2 given *E* is ½, and so both hypotheses are incrementally confirmed, though they are inconsistent with each other.

Principle (4) is also true of absolute confirmation but not of incremental confirmation. It is true of absolute confirmation because any logical consequence of *H* is at least as probable as *H* itself. One way to see that (4) is not true of incremental confirmation is to note that any tautology is a logical consequence of any *H* but a tautology cannot be incrementally confirmed by any evidence, since the probability of a tautology is always one. Thus Hempel was apparently thinking of absolute confirmation, not incremental confirmation, when he endorsed (3) and (4).

Since even eminent confirmation theorists like Hempel have failed to distinguish these two concepts of confirmation, we need to make a conscious effort not to make the same mistake.

CONFIRMATION IN ORDINARY LANGUAGE

When we say in ordinary language that some evidence confirms a hypothesis, does the word "confirms" mean incremental or absolute confirmation?

Since the probability of a tautology is always one, a tautology is absolutely confirmed by any evidence whatever. For example, evidence that it is raining absolutely confirms that all triangles have three sides. Since we would ordinarily say that there is no confirmation in this case, the concept of confirmation in ordinary language is not absolute confirmation.

If *E* reduces the probability of *H* then we would ordinarily say that *E* does not confirm *H*. However, in such a case it is possible for *H* to still be probable given *E* and hence for *E* to absolutely confirm *H*. This shows again that the concept of confirmation in ordinary language is not absolute confirmation.

A hypothesis *H* that is incrementally confirmed by evidence *E* may still be probably false; for example, the hypothesis that a fair coin will land "heads" every time in 1000 tosses is incrementally confirmed by the evidence that it landed "heads" on the first toss, but the hypothesis is still extremely improbable given this evidence. In a case like this nobody would ordinarily say that the hypothesis was confirmed. Thus it appears that the concept of confirmation in ordinary language is not incremental confirmation either.

A few confirmation theorists have attempted to formulate concepts of confirmation that would agree better with the ordinary concept. One such theorist is Nelson Goodman. He noted that if *E* incrementally confirms *H*, and *X* is an irrelevant proposition, then *E* incrementally confirms the conjunction of *H* and *X*. Goodman (1979) thought that in a case like this we would not say that *E* confirms the conjunction. He proposed that "*E* confirms *H*" means that *E* increases the probability of every component of *H*. One difficulty with this is to say what counts as a component of a hypothesis; if any logical consequence of *H* counts as a component of *H* then no hypothesis can ever be confirmed in Goodman's sense. In addition Goodman's proposal is open to the same objection as incremental confirmation: It allows that a hypothesis *H* can be confirmed by evidence *E* and yet *H* be probably false given *E*, which is not what people would ordinarily say.

Peter Achinstein (2001) speaks of "evidence" rather than "confirmation" but he can be regarded as proposing an account of the ordinary concept of confirmation. His account is complex but the leading idea is roughly that "*E* confirms *H*" means that (i) *H* is probable given *E* and (ii) it is probable that there is an explanatory connection between *H* and *E*, given that *H* and *E* are true. The explanatory connection may be that *H* explains *E*, *E* explains *H*, or *H* and *E* have a common explanation. Achinstein's proposal is open to one of the same objections as absolute confirmation: It allows evidence *E* to confirm *H* in cases in which *E* reduces the probability of *H*. Achinstein has argued that this implication is in agreement with the ordinary concept, but his reasoning has been criticized, for example, by Sherrilyn Roush (2004).

It appears that none of the concepts of confirmation discussed by confirmation theorists is the same as the ordinary concept of evidence confirming a hypothesis. Nevertheless, some of these concepts are worthy of study in their own right. In particular, the concepts of incremental and absolute confirmation are simple concepts that are of obvious importance and they are probably components in the more complex ordinary language concept of confirmation.

PROBABILITY

All the concepts of confirmation that we have discussed involve probability. However, the word "probability" is ambiguous. For example, suppose you have been told that a coin either has heads on both sides or else has tails on both sides and that it is about to be tossed. What is the probability that it will land heads? There are two natural answers: (i) ½; (ii) either 0 or 1 but I do not know which. These answers correspond to different meanings of the word "probability." The sense of the word "probability" in which (i) is the natural answer will here be called inductive probability. The sense in which (ii) is the natural answer will be called physical probability.

Physical probability depends on empirical facts in a way that inductive probability does not. We can see this from the preceding example; here the physical probability is unknown because it depends on the nature of the coin, which is unknown; by contrast the inductive probability is known even though the nature of the coin is unknown, showing that the inductive probability does not depend on the nature of the coin.

There are two main theories about the nature of physical probability. One is the frequency theory, according to which the physical probability of an event is the relative frequency with which the event happens in the long run. The other is the propensity theory, according to which the physical probability of an event is the propensity of the circumstances or experimental arrangement to produce that event.

It is widely agreed that the concept of probability involved in confirmation is not physical probability. One reason is that physical probabilities seem not to exist in many contexts in which we talk about confirmation. For example, we often take evidence as confirming a scientific theory but it does not seem that there is a physical probability of a particular scientific theory being true. (The theory is either true or false; there is no long run frequency with which it is true, nor does the evidence have a propensity to make the theory true.) Another reason is that physical probabilities depend on the facts in a way that confirmation relations do not. Inductive probability does not have either of these shortcomings and so it is natural to identify the concept of probability involved in confirmation with inductive probability. Therefore we will now discuss inductive probability in more detail.

Some contemporary writers appear to believe that the inductive probability of a proposition is some person's degree of belief in the proposition. Degree of belief is also called subjective probability, so on this view, inductive probability is the same as subjective probability. However, this is not correct. Suppose, for example, that I claim that scientific theory *H* is probable in view of the available evidence. This is a statement of inductive probability. If my claim is challenged, it would not be a relevant response for me to prove that I have a high degree of belief in *H*, though this would be relevant if inductive probability were subjective probability. To give a relevant defense of my claim I need to cite features of the available evidence that support *H*.

In saying that inductive probabilities are not subjective probabilities, we are not denying that when people make assertions about inductive probabilities they are expressing their degrees of belief. Every sincere and intentional assertion expresses the speaker's beliefs but not every assertion is about the speaker's beliefs.

We will now consider the concept of logical probability and, in particular, whether inductive probability is a kind of logical probability. This depends on what is meant by "logical probability."

Many writers define the "logical probability" of *H* given *E* as the degree of belief in *H* that would be rational for a person whose total evidence is *E*. However, the term "rational degree of belief" is far from clear. On some natural ways of understanding it, the degree of belief in *H*

that is rational for a person could be high even when *H* has a low inductive probability given the person's evidence. This might happen because belief in *H* helps the person succeed in some task, or makes the person feel happy, or will be rewarded by someone who can read the person's mind. Even if it is specified that we are talking about rationality with respect to epistemic goals, the rational degree of belief can differ from the inductive probability given the person's evidence, since the rewards just mentioned may be epistemic. Alternatively, one might take "the rational degree of belief in *H* for a person whose total evidence is *E*" to be just another name for the inductive probability of *H* given *E*, in which case these concepts are trivially equivalent. Thus if one takes "logical probability" to be rational degree of belief then, depending on what one means by "rational degree of belief," it is either wrong or trivial to say that inductive probability is logical.

A more useful conception of logical probability can be defined as follows. Let an "elementary probability sentence" be a sentence that asserts that a specific hypothesis has a specific probability. Let a "logically determinate sentence" be a sentence whose truth or falsity is determined by meanings alone, independently of empirical facts. Let us say that a probability concept is "logical in Carnap's sense" if all elementary probability sentences for it are logically determinate. (This terminology is motivated by some of the characterizations of logical probability in Carnap's *Logical Foundations of Probability*.) Since inductive probability is not subjective probability, the truth of an elementary statement of inductive probability does not depend on some person's psychological state. It also does not depend on facts about the world in the way that statements of physical probability do. It thus appears the truth of an elementary statement of inductive probability does not depend on empirical facts at all and hence that inductive probability is logical in Carnap's sense.

It has often been said that logical probabilities do not exist. If this were right then it would follow that inductive probabilities are either not logical or else do not exist. So we will now consider arguments against the existence of logical probabilities.

John Maynard Keynes in 1921 published a theory of what we call inductive probability and he claimed that these are logical. Frank Ramsey (1926/1980) criticizing Keynes's theory, claimed that "there really do not seem to be any such things as the probability relations he describes." The main consideration that Ramsey offered in support of this was that there is little agreement on the values of probabilities in the simplest cases and these are just the cases where logical relations should be most clear. Ramsey's argument has been cited approvingly by several later authors.

However, Ramsey's claim that there is little agreement on the values of probabilities in the simplest cases seems not to be true. For example, almost everyone agrees with the following:

> (5) The probability that a ball is white, given only that it is either white or black, is ½.

Ramsey cited examples such as the probability of one thing being red given that another thing is red; he noted that nobody can state a precise numerical value for this probability. But that is an example of *agreement* about the value of an inductive probability, since *nobody* pretends to know a precise numerical value for the probability. What examples like this show is merely that inductive probabilities do not always have numerically precise values.

Furthermore, if inductive probabilities are logical (i.e., non-descriptive), it does not follow that their values should be clearest in the simplest cases, as Ramsey claimed. Like other concepts of ordinary language, the concept of inductive probability is learned largely from examples of its application in ordinary life and many of these examples will be complex. Hence, like other concepts of ordinary language, its application may sometimes be clearer in realistic complex situations than in simple situations that never arise in ordinary life.

So much for Ramsey's argument. Another popular argument against the existence of logical probabilities is based on the "paradoxes of indifference." The argument is this: Judgments of logical probability are said to presuppose a general principle, called the Principle of Indifference, which says that if evidence does not favor one hypothesis over another then those hypotheses are equally probable on this evidence. This principle can lead to different values for a probability, depending on what one takes the alternative hypotheses to be. In some cases the different choices seem equally natural. These "paradoxes of indifference," as they are called, are taken by many authors to be fatal to logical probability.

But even if we agree (as Keynes did) that quantitative inductive probabilities can only be determined via the Principle of Indifference, we can also hold (as Keynes did) that inductive probabilities do not always have quantitative values. Thus if there are cases where contradictory applications of the principle are equally natural, we may take this to show that these are cases where inductive probabilities lack quantitative values. It does not follow

that quantitative inductive probabilities never exist, or that qualitative inductive probabilities do not exist. The paradoxes of indifference are thus consistent with the view that inductive probabilities exist and are logical.

How can we have knowledge of inductive probabilities, if this does not come from an exceptionless general principle? The answer is that the concept of inductive probability, like most concepts of ordinary language, is learned from examples, not by general principles. Hence we can have knowledge about particular inductive probabilities (and hence logical probabilities) without being able to state a general principle that covers these cases.

A positive argument for the existence of inductive probabilities is the following: We have seen reason to believe that a statement of inductive probability, such as (5), is either logically true or logically false. Which of these it is will be determined by the concepts involved, which are concepts of ordinary language. So, since competent speakers of a language normally use the language correctly, the wide endorsement of (5) is good reason to believe that (5) is a true sentence of English. And it follows from (5) that at least one inductive probability exists. Parallel arguments would establish the existence of many other inductive probabilities.

The concept of probability that is involved in confirmation can appropriately be taken to be inductive probability. Unlike physical probability, the concept of inductive probability applies to scientific theories. And unlike both physical and subjective probability, the concept of inductive probability agrees with the fact that confirmation relations are not discovered empirically but by examination of the relation between the hypothesis and the evidence.

EXPLICATION OF INDUCTIVE PROBABILITY

Inductive probability is a concept of ordinary language and, like many such concepts, it is vague. This is reflected in the fact that inductive probabilities often have no precise numerical value.

A useful way to theorize about vague concepts is to define a precise concept that is similar to the vague concept. This methodology is called explication, the vague concept is called the explicandum, and the precise concept that is meant to be similar to it is called the explicatum. Although the explicatum is intended to be similar to the explicandum, there must be differences, since the explicatum is precise and the explicandum is vague. Other desiderata for an explicatum, besides similarity with the explicandum, are theoretical fruitfulness and simplicity.

Inductive probability can be explicated by defining, for selected pairs of sentences E and H, a number that will be the explicatum for the inductive probability of H given E; let us denote this number by "$p(H|E)$." The set of sentences for which $p(H|E)$ is defined will depend on our purposes.

Quantitative inductive probabilities, where they exist, satisfy the mathematical laws of probability. Since a good explicatum is similar to the explicandum, theoretically fruitful, and simple, the numbers $p(H|E)$ will also be required to satisfy these laws.

In works written from the 1940s to his death in 1970, Carnap proposed a series of increasingly sophisticated explications of this kind, culminating in his *Basic System of Inductive Logic* published posthumously in 1971 and 1980. Other authors have proposed other explicata, some of which will be mentioned below.

Since the value of $p(H|E)$ is specified by definition, a statement of the form "$p(H|E) = r$" is either true by definition or false by definition, and hence is logically determinate. Since we require these values to satisfy the laws of probability, the function p is also a probability function. So we may say that the function p is a logical probability in Carnap's sense.

Thus there are two different kinds of probability, both of which are logical in Carnap's sense: Inductive probability and functions that are proposed as explicata for inductive probability. Since the values of the explicata are specified by definition, it is undeniable that logical probabilities of this second kind exist.

EXPLICATION OF INCREMENTAL CONFIRMATION

Since inductive probability is vague, and E incrementally confirms H if and only if E raises the inductive probability of H, the concept of incremental confirmation is also vague. We will now consider how to explicate incremental confirmation.

First, we note that the judgment that E confirms H is often made on the assumption that some other information D is given; this information is called background evidence. So we will take the form of a fully explicit judgment of incremental confirmation to be "E incrementally confirms H given D." For example, a coin landing heads on the first toss incrementally confirms that the coin has heads on both sides, given that both sides of the coin are the same; there would be no confirmation if the back-

ground evidence was that the coin is normal with heads on one side only.

The judgment that E incrementally confirms H given D means that the inductive probability of H given both E and D is greater than the inductive probability of H given only D. Suppose we have a function p that is an explicatum for inductive probability and is defined for the relevant statements. Let "$E.D$" represent the conjunction of E and D (so the dot here functions like "and"). Then the explicatum for "E incrementally confirms H given D" will be $p(H|E.D) > p(H|D)$. We will use the notation "$C(H, E, D)$" as an abbreviation for this explicatum.

The concept of incremental confirmation, like all the concepts of confirmation discussed so far, is a qualitative concept. For each of these qualitative concepts there is a corresponding comparative concept, which compares the amount of confirmation in different cases. We will focus here on the judgment that E_1 incrementally confirms H more than E_2 does, given D. The corresponding statement in terms of our explicata is that the increase from $p(H|D)$ to $p(H|E_1.D)$ is larger than the increase from $p(H|D)$ to $p(H|E_2.D)$. This is true if and only if $p(H|E_1.D) > p(H|E_2D)$, so the explicatum for "E_1 confirms H more than E_2 does, given D" will be $p(H|E_1.D) > p(H|E_2.D)$. We will use the notation "$M(H,E_1,E_2,D)$" as an abbreviation for this explicatum.

Confirmation theorists have also discussed quantitative concepts of confirmation, which involve assigning numerical "degrees of confirmation" to hypotheses. In earlier literature the term "degree of confirmation" usually meant degree of absolute confirmation. The degree to which E absolutely confirms H is the same as the inductive probability of H given E and hence is explicated by $p(H|E)$.

In later literature, the term "degree of confirmation" is more likely to mean degree of incremental confirmation. An explicatum for the degree to which E incrementally confirms H given D is a measure of how much $p(H|E.D)$ is greater than $p(H|D)$. Many different explicata of this kind have been proposed; they include the following. (Here "$\sim H$" means the negation of H.)

Difference measure: $p(H|E.D) - p(H|D)$

Ratio measure: $p(H|E.D) / p(H|D)$

Likelihood ratio: $p(E|H.D) / p(E|\sim H.D)$

Confirmation theorists continue to debate the merits of these and other measures of degree of incremental confirmation.

VERIFIED CONSEQUENCES

The remainder of this entry will consider various properties of incremental confirmation and how well these are captured by the explicata C and M that were defined above. We begin with the idea that hypotheses are confirmed by verifying their logical consequences.

If H logically implies E given background evidence D, we usually suppose that observation of E would incrementally confirm H given D. For example, Einstein's general theory of relativity, together with other known facts, implied that the orbit of Mercury precesses at a certain rate; hence the observation that it did precess at this rate incrementally confirmed Einstein's theory, given the other known facts.

The corresponding explicatum statement is: If $H.D$ implies E then $C(H,E,D)$. Assuming that p satisfies the laws of mathematical probability, this explicatum statement can be proved true provided that $0 > p(H|D) > 1$ and $p(E|D) < 1$.

We can see intuitively why the provisos are needed. If $p(H|D) = 1$ then H is certainly true given D and so no evidence can incrementally confirm it. If $p(H|D) = 0$ then H is certainly false given D and the observation that one of its consequences is true need not alter this situation. If $p(E|D) = 1$ then E was certainly true given D and so the observation that it is true cannot provide new evidence for H.

If H and D imply both E_1 and E_2, and if E_1 is less probable than E_2 given D, then we usually suppose that H would be better confirmed by E_1 than by E_2, given D. The corresponding explicatum statement is: If $H.D$ implies E_1 and E_2, and $p(E_1|D) < p(E_2|D)$, then $M(H, E_1, E_2, D)$. Assuming that p satisfies the laws of probability, this can be proved true provided that $0 < p(H|D) < 1$. The proviso makes sense intuitively for the same reasons as before.

If H and D imply both E_1 and E_2 then we usually suppose that E_1 and E_2 together would confirm H more than E_1 alone, given D. The corresponding explicatum statement is that if $H.D$ implies E_1 and E_2 then $M(H, E_1.E_2, E_1, D)$. It follows from the result in the previous paragraph that this is true, provided that $p(E_1.E_2|D) < p(E_1|D)$ and $0 < p(H|D) < 1$. The provisos are needed for the same reasons as before.

These results show that, if we require p to satisfy the laws of probability, then C and M will be similar to their explicanda with respect to verified consequences and, to that extent at least, C and M will be good explicata. In addition these results illustrate in a small way the value of explication. Although the provisos that we added make

sense when one thinks about them, the need for them is likely to be overlooked if one thinks only in terms of the vague explicanda and does not attempt to prove a precise corresponding result in terms of the explicata. Thus explication can give a deeper and more accurate understanding of the explicandum. We will see more examples of this.

REASONING BY ANALOGY

If two individuals are known to be alike in certain respects, and one is found to have a particular property, we often infer that, since the individuals are similar, the other individual probably also has that property. This is a simple example of reasoning by analogy, and it is a kind of reasoning that we use every day.

In order to explicate this kind of reasoning, we will use "*a*" and "*b*" to stand for individual things and "*F*" and "*G*" for logically independent properties that an individual may have (for example, being tall and blond). We will use "*Fa*" to mean that the individual *a* has the property *F*; similarly for other properties and individuals.

It is generally accepted that reasoning by analogy is stronger the more properties that the individuals are known to have in common. So for *C* to be a good explicatum it must satisfy the following condition:

(6) *C (Gb, Fa.Fb, Ga)*.

Here we are considering the situation in which the background evidence is that *a* has *G*. The probability that *b* also has *G* is increased by finding that *a* and *b* also share the property *F*.

In the case just considered, *a* and *b* are not known to differ in any way. When we reason by analogy in real life we normally do know some respects in which the individuals differ, but this does not alter the fact that the reasoning is stronger the more alike *a* and *b* are known to be. So for *C* to be a good explicatum it must also satisfy the following condition. (Here *F'* is a property that is logically independent of both *F* and *G*.)

(7) *C (Gb, Fa.Fb, Ga.F'a.~F'b)*.

Here the background evidence is that *a* has *G* and that *a* and *b* differ in regard to *F'*. The probability that *b* has *G* is increased by finding that *a* and *b* are alike in having *F*.

Another condition that *C* should satisfy is:

(8) *C (Gb, Ga, F'a. ~F'b)*.

Here the background evidence is merely that *a* and *b* differ regarding *F'*. For all we know, whether or not something has *F'* might be unrelated to whether it has *G*, so the fact that *a* has *G* is still some reason to think that *b* has *G*.

In *Logical Foundations of Probability* Carnap proposed a particular explicatum for inductive probability that he called *c**. In *The Continuum of Inductive Methods* he described an infinite class of possible explicata. The function *c**, and all the functions in Carnap's continuum, satisfy (6) but not (7) or (8). Hence none of these functions provides a fully satisfactory explicatum for situations that involve more than one logically independent property.

Carnap recognized this failure early in the 1950s and worked to find explicata that would handle reasoning by analogy more adequately. He first found a class of possible explicata for the case where there are two logically independent properties; the functions in this class satisfy (6) and (8). Subsequently, with the help of John Kemeny, Carnap generalized his proposal to the case where there are any finite number of logically independent properties, though he never published this. A simpler and less adequate generalization was published by Mary Hesse in 1964. Both these generalizations satisfy all of (6)-(8).

Carnap had no justification for the functions he proposed except that they seemed to agree with intuitive principles of reasoning by analogy. Later he found that they actually violate one of the principles he had taken to be intuitive. In his last work Carnap expressed indecision about how to proceed.

For the case where there are just two properties, Maher (2000) has shown that certain foundational assumptions pick out a class of probability functions, called P_I, that includes the functions that Carnap proposed for this case. Maher argued that the probability functions in P_I handle reasoning by analogy adequately and Carnap's doubts were misplaced.

For the case where there are more than two properties, Maher (2001) has shown that the proposals of Hesse, and Carnap and Kemeny, correspond to implausible foundational assumptions and violate intuitive principles of reasoning by analogy. Further research is needed to find an explicatum for inductive probability that is adequate for situations involving more than two properties.

NICOD'S CONDITION

We are often interested in universal generalizations of the form "All *F* are *G*," for example, "All ravens are black," or "All metals conduct electricity." Nicod's condition, named after the French philosopher Jean Nicod, says that generalizations of this form are confirmed by finding an indi-

vidual that is both *F* and *G*. (Here and in the remainder of this entry, "confirmed" means incrementally confirmed.)

Nicod (1970) did not mention background evidence. It is now well known that Nicod's condition is not true when there is background evidence of certain kinds. For example, suppose the background evidence is that, if there are any ravens, then there is a non-black raven. Relative to this background evidence, observation of a black raven would refute, not confirm, that all ravens are black.

Hempel claimed that Nicod's condition is true when there is no background evidence but I. J. Good argued that this is also wrong. Good's argument was essentially this: Given no evidence whatever, it is improbable that there are any ravens, and if there are no ravens then, according to standard logic, "All ravens are black" is true. Hence, given no evidence, "All ravens are black" is probably true. However, if ravens do exist, they are probably a variety of colors, so finding a black raven would increase the probability that there is a non-black raven and hence disconfirm that all ravens are black, contrary to Nicod's condition.

Hempel was relying on intuition, and Good's counterargument is intuitive rather than rigorous. A different way to investigate the question is to use precise explicata. The situation of "no background evidence" can be explicated by taking the background evidence to be any logically true sentence; let *T* be such a sentence. Letting A be "all *F* are *G*," the claim that Nicod's condition holds when there is no background evidence may be expressed in explicatum terms as

(9) $C\ (A,\ Fa.Ga,\ T)$.

Maher has shown that this can fail when the explicatum p is a function in P_I and that the reason for the failure is the one identified in Good's argument. This confirms that Nicod's condition is false even when there is no background evidence.

Why then has Nicod's condition seemed plausible? One reason may be that people sometimes do not clearly distinguish between Nicod's condition and the following statement: Given that an object is F, the evidence that it is *G* confirms that all *F* are *G*. The latter statement may be expressed in explicatum terms as:

(10) $C\ (A,\ Ga,\ Fa)$.

This is true provided only that *p* satisfies the laws of probability, $0 < p(A|Fa) < 1$, and $p(Ga|Fa) < 1$. (This follows from the first of the results stated earlier for verified consequences.) If people do not clearly distinguish between the ordinary language statements that correspond to (9) and (10), the truth of the latter could make it seem that Nicod's condition is true.

THE RAVENS PARADOX

The following three principles about confirmation have seemed plausible to many people.

(11) Nicod's condition holds when there is no background evidence.

(12) Confirmation relations are unchanged by substitution of logically equivalent sentences.

(13) In the absence of background evidence, the evidence that some individual is a non-black non-raven does not confirm that all ravens are black.

However, these three principles are inconsistent. That is because (11) implies that a non-black non-raven confirms "all non-black things are non-ravens," and the latter is logically equivalent to "all ravens are black," so by (12) a non-black non-raven confirms "all ravens are black," contrary to (13).

Hempel was the first to discuss this paradox. His initial statement of the paradox did not explicitly include the condition of no background evidence but he stated later in his article that this was to be understood. The subsequent literature on this paradox is enormous but most discussions have not respected the condition of no background evidence. Here we will follow Hempel in respecting that condition.

The contradiction shows that at least one of (11)-(13) is false. Hempel claimed that (11) and (12) are true and (13) is false but his judgments were based on informal intuitions, not on any precise explicatum or use of probability theory.

Our preceding discussion of Nicod's condition shows that (11) is false, contrary to what Hempel thought. On the other hand, our explicata support Hempel's view that (12) is true and (13) is false, as we will now show.

In explicatum terms, what (12) says is: If H', E', and $D'$are logically equivalent to H, E, and D respectively, then $C(H, E, D)$ if and only if $C(H', E', D')$. The truth of this follows from the assumption that *p* satisfies the laws of probability.

Now let "*F*" mean "raven" and "*G*" mean "black." Then (13), expressed in explicatum terms, is the claim $\sim C\ (A,\ \sim Fa.\sim Ga,\ T)$. Maher has shown that this need not

be true when p is a function in P_I; we can instead have *C (A, ~Fa. ~Ga, T)*. This happens for two reasons:

(a) The evidence *~Fa.~Ga* reduces the probability of *Fb.~Gb*, where *b* is any individual other than *a*. Thus *~Fa.~Ga* reduces the probability that another individual *b* is a counterexample to *A*.

(b) The evidence *~Fa.~Ga* tells us that *a* is not a counterexample to *A*, which a priori it could have been.

Both of these reasons make sense intuitively.

We conclude that, of the three principles (11)-(13), only (12) is true.

PROJECTABILITY

A predicate is said to be "projectable" if the evidence that the predicate applies to some objects confirms that it also applies to other objects. The standard example of a predicate that is not projectable is "grue," which was introduced by Goodman (1979). According to Goodman's defnition, something is grue if either (i) it is observed before time *t* and is green or (ii) it is not observed before time *t* and is blue. The usual argument that "grue" is not projectable goes something like this: A grue emerald observed before *t* is green, and observation of such an emerald confirms that emeralds not observed before *t* are also green. Since a green emerald not observed before *t* is not grue, it follows that a grue emerald observed before *t* confirms that emeralds not observed before *t* are not grue; hence "grue" is not projectable.

The preceding account of the meaning of "projectable" was the usual one but it is imprecise because it fails to specify background evidence. Let us say that a predicate ϕ is absolutely projectable if *C* $(\phi b, \phi a, T)$ for any distinct individuals *a* and *b* and logical truth *T*. This concept of absolute projectability is one possible explicatum for the usual imprecise concept of projectability. Let "*Fa*" mean that *a* is observed before *t* and let "*Ga*" mean that *a* is green. Let "*G'a*" mean that either *Fa.Ga* or *~Fa.~ Ga*. Thus "*G*'" has a meaning similar to "grue." (The difference is just that *G* uses "not green" instead of "blue" and so avoids introducing a third property.) Maher has proved that if p is any function in P_I then "*F*", "*G*", and "*G'* " are all absolutely projectable. It may seem unintuitive that "*G'* " is absolutely projectable. However, this result corresponds to the following statement of ordinary language: The probability that *b* is grue is higher given that *a* is grue than if one was not given any evidence whatever. If we keep in mind that we do not know whether *a* or *b* was observed before *t*, this should be intuitively acceptable. So philosophers who say that "grue" is not projectable are wrong if, by "projectable," they mean absolute projectability.

Let us say that a predicate ϕ is projectable across another predicate ψ if *C* $(\phi b, \phi a, \psi a.\sim\psi b)$ for any distinct individuals *a* and *b*. This concept of projectability across another predicate is a second possible explicatum for the usual imprecise concept of projectability.

It can be shown that if p is any function in P_I then "*G*" is, and "*G'* " is not, projectable across "*F*." So philosophers who say that "grue" is not projectable are right if, by "projectable," they mean projectability across the predicate "observed before *t*."

Now suppose we change the definition of "*Ga*" to be that *a* is (i) observed before *t* and green or (ii) not observed before *t* and not green. Thus "*G*" now means what "*G'* " used to mean. Keeping the definitions of "*F*" and "*G'* " unchanged, "*G'a*" now means that *a* is green. The results reported in the preceding paragraph will still hold but now they are the opposite of the usual views about what is projectable. This shows that, when we are constructing explicata for inductive probability and confirmation, the meanings assigned to the basic predicates (here "*F*" and "*G*") need to be intuitively simple ones rather than intuitively complex concepts like "grue."

See also Carnap, Rudolf; Einstein, Albert; Goodman, Nelson; Hempel, Carl Gustav; Induction; Keynes, John Maynard; Probability and Chance; Ramsey, Frank Plumpton; Relativity Theory.

Bibliography

Achinstein, Peter. *The Book of Evidence*. New York: Oxford University Press, 2001.

Carnap, Rudolf. "A Basic System of Inductive Logic, Part I." In *Studies in Inductive Logic and Probability*. Vol. 1, edited by Rudolf Carnap and Richard C. Jeffrey. Berkeley: University of California Press, 1971.

Carnap, Rudolf. "A Basic System of Inductive Logic, Part II." In *Studies in Inductive Logic and Probability*. Vol. 2, edited by Richard C. Jeffrey. Berkeley: University of California Press, 1980.

Carnap, Rudolf. *The Continuum of Inductive Methods*. Chicago: University of Chicago Press, 1952.

Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 1950. Second edition 1962.

Earman, John. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press, 1992.

Festa, Roberto. "Bayesian Confirmation." In *Experience, Reality, and Scientific Explanation*, edited by Maria Carla Galavotti and Alessandro Pagnini. Dordrecht: Kluwer, 1999.

Fitelson, Branden. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (1999): S362–S378.

Gillies, Donald. *Philosophical Theories of Probability*. London: Routledge, 2000.

Good, I. J. "The White Shoe *qua* Herring Is Pink." *British Journal for the Philosophy of Science* 19 (1968): 156–157.

Goodman, Nelson. *Fact, Fiction, and Forecast*. 3rd ed. Indianapolis, IN: Hackett, 1979.

Hempel, Carl G. "Studies in the Logic of Confirmation." *Mind* 54 (1945): 1–26 and 97–121. Reprinted with some changes in Carl G. Hempel. *Aspects of Scientific Explanation*. New York: The Free Press, 1965.

Hesse, Mary. "Analogy and Confirmation Theory." *Philosophy of Science* 31 (1964): 319–327.

Howson, Colin, and Peter Urbach. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court, 1993.

Keynes, John Maynard. *A Treatise on Probability*. London: Macmillan, 1921. Reprinted with corrections, 1948.

Maher, Patrick. "Probabilities for Two Properties." *Erkenntnis* 52 (2000): 63–91.

Maher, Patrick. "Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny." *Erkenntnis* 55 (2001): 183–216.

Maher, Patrick. "Probability Captures the Logic of Scientific Confirmation." In *Contemporary Debates in Philosophy of Science*, edited by Christopher R. Hitchcock. Oxford: Blackwell, 2004.

Nicod, Jean. *Geometry and Induction*. Berkeley and Los Angeles: University of California Press, 1970. English translation of works originally published in French in 1923 and 1924.

Ramsey, Frank P. "Truth and Probability." Article written in 1926 and published in many places, including *Studies in Subjective Probability*, 2nd ed., edited by Henry E. Kyburg, Jr. and Howard E. Smokler. Huntington, New York: Krieger, 1980.

Roush, Sherrilyn. "Positive Relevance Defended." *Philosophy of Science* 71 (2004):110–116.

Salmon, Wesley C. "Confirmation and Relevance." In *Minnesota Studies in the Philosophy of Science*. Vol. VI: *Induction, Probability, and Confirmation*, ed. Grover Maxwell and Robert M. Anderson Jr. Minneapolis: University of Minnesota Press, 1975.

Skyrms, Brian. *Choice and Chance*. 4th ed. Belmont, CA: Wadsworth, 2000.

Stalker, Douglas, ed. *Grue: Essays on the New Riddle of Induction*. Chicago: Open Court, 1994.

Patrick Maher (2005)

CONFUCIUS

(551–479 BCE)

Confucius (Kong Qiu) is one of the early Chinese philosophers and the founder of the ethical teaching known as Confucianism. He was born in a time of political, social, and spiritual crisis that had shattered the traditional way of life as well as the view of a world based on the conventions of ritual propriety (*li*) and the religion of Heaven (*tian*). The hierarchies of the patriarchal feudal system of the Zhou had fallen into decay, giving way to a new social mobility, and because of this, a small but influential middle class emerged. Its members became the clients of private teachers who imparted the knowledge needed in a society that ascribed increasing importance to individual capability instead of descent.

Confucius (a transcription of Kong fuzi—teacher Kong) was one of these teachers. He probably taught the practical "six arts" (writing, mathematics, ritual propriety, music, charioteering, and archery) and dealt with the texts handed down from the past that he is said to have edited and that constitute the core of the later Confucian classics. However, as documented by his "Collected Words" (*Lunyu*, a later compilation), the main focus of his teaching is morality. Confucius dedicates himself to an ideal of education that transcends the social boundaries and roles the disciples would possibly play in their present and later life—the ideal of becoming a gentleman (*junzi*), a truly moral person in solidarity with the community and rooted in self-respect. This endeavor is again embedded in the quest for a still higher goal: To rescue "this culture" from the flood in which it was drowning, and to "change the world" that had lost the dao, the right way (*LY* 9.5, 18.6).

To find a solution for the world in a time when tradition was in crisis enforced a reflection on the established norms in order to reconstruct and rescue their true meaning (*zheng ming*). This gives a philosophical ring to Confucius's ethics. He finds one of the paradigmatic answers to the challenge China's intellectuals were facing: How to redefine humankind's position in a world that had lost its foundation, without the possibility for reiterating the past. His answer is the internalization of ethics as a new basis for the ethical life, which entails both constant self-reference of the individual and norm reflexivity. It has always remained a Confucian conviction that there must be ethical rather than legal or organizational solutions to the basic problems of human existence; that morality must have primacy over all other concerns, also over politics; and that the human being as its agent is capable of moral cultivation. This makes the Confucian position distinct both from the Daoist return to nature and the legalist social engineering.

The general structure of ethics of the *Lunyu* may be described as comprising three steps: (1) In view of the sobering conditions of the time, the "gentleman" *turns away* from society; he no longer trusts public reputation

(*LY* 12.20) and the opinions of the majority (13.24, 15.28); and he is constantly prepared to be misjudged and not acknowledged by others (1.1, 1.16, 14.30, 15.19). He then (2) *turns into his inner self* where in private seclusion he develops self-respect (5.16, 13.19) as the basis for autonomous action and, given the absence of a strong religious backing of ethics, the ultimate reason for being moral. Through regular self-reflection and critical self-examination (1.4, 5.27, 12.4, 15.21) he safeguards the purity of his intentions, which, if necessary, will enable him to, as Confucius is quoted in the *Mengzi*, "withstand thousands or tens of thousands" (2a:2). However, in a final step, the moral actor (3) consciously "overcomes himself" and *returns* to society (*ke ji fu li*) (*LY* 12.1). He thus accepts his responsibility as a moral authority in the interest of the common good, rather than simply trying to stay "clean" in a world where "the *dao* does not prevail" (18.7), far away from the ideal of the "great community" (*da tong*, attributed to Confucius in *Liji* 9).

Return to society implies the critical acknowledgement of the given ethos of a hierarchical world dominated by the principle of male seniority. Without the handed-down rules of propriety (*li*) the human being would be without a firm "standing" (*lì*) (*LY* 8.8, 16.13, 20.3). However, the traditional canon of normative orientations is reconsidered and realigned to a new organizing center—humaneness (*ren*). Humaneness has "to start from oneself" rather than from external guidance (12.1); it is ideally followed for its own sake rather than for reasons of utility (4.2); and it is universally valid, even when one is among barbarian tribes (13.19).

Humaneness is explicated differently, however, in the *Lunyu*, the most conspicuous variants being its affective reading as love (12.22) and its cognitive reading as the golden rule (5.12, 6.30, 12.2, 15.23), the maxim that "consists of one word and can be practiced through all one's life" (15.24) and the "one that goes through all" (4.15). By humaneness in terms of the golden rule, the direct reciprocal relationship with the generalized "other" becomes one of the two complementary dimensions of ethics along with the concrete role orientation.

Confucius's ethics thus promises a "mean" comprising personal integrity and social integration, allowing one to keep faith with the conventional ethos while not surrendering to it. The "gentleman" as its protagonist will fulfill the duties owed to family and society and at the same time, "harmonious, but not conformist" (13.23), maintain a moral watchfulness and inner independence.

It was possible, however, to adopt this ethics with different accents, also because of the vagueness of many *Lunyu* passages and the opacity of its structure. The conflict of opinions about the true teaching of the master, the attempts to regain the original spirit lost in the course of its effective history, as well as Confucius's critique as a rebel, a ritualist, and a moralist out of touch with reality, apparently started shortly after his death. The debate still continues in the twenty-first century, with deontological, pragmatist, aestheticist, communitarian, and religious interpretations competing with each other.

See also Chinese Philosophy: Confucianism.

Bibliography

Creel, Herrlee G. *Confucius and the Chinese Way*. New York: Harper, 1960.

Lau, D. C. *Confucius: The Analects (Lun Yü)*. 2nd ed. Hong Kong: Chinese University Press, 1993.

Roetz, Heiner. *Confucian Ethics of the Axial Age*. Albany: SUNY Press, 1993.

Roetz, Heiner. *Konfuzius*. Munich: Beck, 1998.

Schaberg, David. "Sell it! Sell it! Recent Translations of Lunyu." *Chinese Literature: Essays, Articles, Reviews* 23 (2001): 115–139.

Van Norden, Bryan W., ed. *Confucius and the Analects: New Essays*. New York: Oxford University Press, 2002. Contains a comprehensive bibliography.

Heiner Roetz (2005)

CONNECTIONISM

"Connectionism" is an approach within cognitive science that employs neural networks, rather than computer programs, as the basis for modeling mentality. A connectionist system, or neural network, is a structure of simple neuronlike processors called nodes or units. Each node has directed connections to other nodes, so that the nodes send and receive excitatory and inhibitory signals to and from one another. The total input to a node determines its state of activation. When a node is on, it sends out signals to the nodes to which it has output connections, with the intensity of a signal depending upon both (1) the activation level of the sending node and (2) the strength or "weight" of the connection between it and the receiving node. Typically at each moment during processing, many nodes are simultaneously sending signals to others.

When neural networks are employed for information processing, certain nodes are designated "input" units and others as "output" units, and potential patterns of activation across them are assigned interpretations. (The remaining nodes are called "hidden units.") Typically a "problem" is posed to a network by activating a

pattern in the input nodes; then the various nodes in the system simultaneously send and receive signals repeatedly until the system settles into a stable configuration; the semantic interpretation of the resulting pattern in the output nodes is what the system currently represents, hence its "answer" to the problem. Connectionist systems are capable of "learning" from "experience" by having their weights changed systematically in a way that depends upon how well the network has performed in generating solutions to problems posed to it as a training regimen. (Typically the device employed is not an actual neural network but a simulation of one on a standard digital computer.)

The most striking difference between such networks and conventional computers is the lack of an executive component. In a conventional computer the behavior of the whole system is controlled at the central processing unit (CPU) by a stored program. A connectionist system lacks both a CPU and a stored program. Nevertheless, often in a connectionist system certain activation patterns over sets of hidden units can be interpreted as internal representations with interesting content, and often the system also can be interpreted as embodying, in its weighted connections, information that gets automatically accommodated during processing without getting explicitly represented via activation patterns.

Connectionist models have yielded particularly encouraging results for cognitive processes such as learning, pattern recognition, and so-called multiple-soft-constraint satisfaction (i.e., solving a problem governed by several constraints, where an optimal solution may require violating some constraints in order to satisfy others). For example, Terry Sejnowski and Charles Rosenberg trained a network they called NETtalk to convert inputs that represent a sequence of letters, spaces, and punctuation constituting written English into outputs that represent the audible sounds constituting the corresponding spoken English. (The phonetic output code then can be fed into a speech synthesizer, a device that actually produces the sounds.)

Philosophical discussion of connectionism has largely centered on whether connectionism yields or suggests a conception of mentality that is importantly different from the conception of mind-as-computer at the core of classical cognitive science. Several different nonclassical alternatives have been suggested; each has been alleged to fit well with connectionism, and each has been a locus of debate between fans and foes of connectionism. Three proposed interpretations of connectionism deserve specific mention.

On one view, the key difference between classical models of mental processes and connectionist models is that the former assume the existence of languagelike mental representations that constitute a so-called language of thought (LOT), whereas the latter supposedly favor representations that are alleged to be inherently non-languagelike in structure: namely, activation patterns distributed over several nodes of a network, so-called activation vectors. On this interpretation connectionism shares with classicism the assumption that cognition is computation over mental representations—that cognitive transitions conform to rules for transforming representations on the basis of their formal structure, rules that could be formulated as an explicit computer program. (In connectionist systems the rules are wired into the weights and connections rather than being explicitly represented. In classical systems some rules must be hard wired; and there may be—but need not be—other rules that are explicitly represented as stored data structures.) The key difference allegedly turns on the languagelike or non-languagelike structure of mental representations.

This construal of connectionism fits naturally with the idea that human cognition involves state transitions that are all essentially associative—in the sense that they reflect statistical correlations among items the system can represent and can be analyzed as the drawing of statistical inferences. Many fans of connectionism, including Patricia Churchland and Paul Churchland, evidently see things this way and tend to regard connectionism as breathing new life into associationism. Prominent foes of connectionism, notably Jerry Fodor and Zenon Pylyshyn, also see things this way; but they regard the link with associationism as grounds for maintaining that connectionism is bound to founder on the same general problem that plagued traditional associationism in psychology: namely, inability to account for the rich semantic coherence of much human thought. To overcome this problem, Fodor and Pylyshyn maintain, cognitive science must continue to posit both (1) mental representations that encode propositional information via languagelike syntactic structure and (2) modes of processing that are suitably sensitive to syntactic structure and are thereby sensitive to propositional content.

A second interpretation of connectionism claims that connectionist models do not really employ internal representations at all in their hidden units (and, a fortiori, do not employ internal representations with languagelike structure). This view has been defended—by Rodney Brooks, for example—on the grounds that puta-

tive representations in connectionist systems play no genuine explanatory role. It has also been defended—for instance, by Hubert Dreyfus and Stuart Dreyfus—on the basis of a Heideggerian critique of the notion of mental representation itself. The approach goes contrary to the views of most (but not all) practicing connectionists, who typically posit internal representations in connectionist models and assign them a central explanatory role.

A third interpretation assumes the existence of internal mental representations; and it does not deny—indeed, the version defended by Terence Horgan and John Tienson resolutely affirms—that mental representations often have languagelike structure. It focuses instead on the classical assumption that cognition is computation (see above). This third approach maintains (1) that much of human cognition is too rich and too subtle to conform to programmable rules and (2) that connectionism has theoretical resources for potentially explaining such nonalgorithmic cognitive processing. The approach stresses that there is a powerful branch of mathematics that applies naturally to neural networks: dynamical systems theory. According to this anticomputational construal of connectionism, there can be cognitive systems—subservable mathematically by dynamical systems, which in turn are subservable physically by neural networks—whose cognitive state transitions are not tractably computable. In other words, mental activity in these systems is too refined and too supple to conform to programmable rules. Humans are alleged to be such cognitive systems, and connectionism (so interpreted) is held to yield a more adequate picture of the mind than the classical computational picture.

One objection to this third interpretation of connectionism alleges that cognitive state transitions in a connectionist system must inevitably conform to programmable rules, especially since neural networks are simulable on standard computers. Another objection, directed specifically at the version that retains languagelike representations, alleges that the LOT hypothesis is intelligible only on the assumption that cognition is computation.

In much of the early philosophical debate between proponents and opponents of connectionism, the first interpretation was largely taken for granted. But as competing interpretations get articulated, defended, and acknowledged, philosophical discussion of connectionism and its potential implications becomes richer.

See also Cognitive Science; Fodor, Jerry A.; Language of Thought; Philosophy of Mind.

Bibliography

Aizawa, K. "Explaining Systematicity." *Mind and Language* 12 (1997): 115–136.

Bechtel, W., and A. Abrahamsen. *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks.* 2nd ed. Oxford: Blackwell, 2001.

Churchland, P. M. "Conceptual Similarity across Sensory and Neural Diversity: The Fodor/Lepore Challenge Answered." *Journal of Philosophy* 95 (1998): 5–32.

Clark, A., and P. Millican, eds. *Connectionism, Concepts, and Folk Psychology.* Oxford: Oxford University Press, 1996.

Cummins, R. "Systematicity." *Journal of Philosophy* 93 (1996): 561–614

Dawson, M. *Minds and Machines: Connectionism and Psychological Modeling.* Oxford: Blackwell, 2004.

Horgan, T., and J. Tienson. *Connectionism and the Philosophy of Psychology.* Cambridge, MA.: MIT Press, 1996.

Macdonald, C., ed. *Connectionism: Debates on Psychological Explanation.* Oxford: Blackwell, 1995.

Marcus, G. *The Algebraic Mind.* Cambridge, MA: MIT Press, 2001.

McLeod, P., K. Plunkett, and E. T. Rolls. *Introduction to Connectionist Modelling of Cognitive Processes.* Oxford: Oxford University Press, 1998.

Tomberlin, J., ed. *Philosophical Perspectives 9: AI, Connectionism and Philosophical Psychology.* Atascadero: Ridgeview Press, 1995.

Tuescher, C. *Turing's Connectionism. An Investigation of Neural Network Architectures.* London: Springer-Verlag, 2002.

Terence E. Horgan (1996)
Bibliography updated by Alyssa Ney (2005)

CONSCIENCE

Doubtless from the earliest times in which groups established social customs, or mores, and enforced them, members of such groups who were tempted to violate these mores could almost feel the disapproval of their fellows and hear in their own minds a protesting outcry, perhaps some primitive equivalent of "No!" or "Don't!" In the early eighteenth century such inner voices or feelings were described as edicts of one's moral sense or of one's "conscience." This kind of account of these restraining influences became explicit with the development of faculty psychology, which involved the view that there are different faculties of the human mind responsible for different capacities or abilities which the mind seems to exhibit. Reason was thought of as the rational faculty, emotion as a passional one, and volition as a faculty that enables us to reach decisions and make choices. The moral faculty was thought by some, the earl of Shaftesbury and Francis Hutcheson, for example, to operate through feelings. For instance, a feeling of repugnance would tend to be aroused by the thought of doing any-

thing immoral—anything in violation of the mores—and a feeling of approval by the thought of acting virtuously.

In contrast with this moral-sense type of theory, Samuel Clarke and Richard Price, among others, thought that it must be something akin to reason or the understanding which enabled us to distinguish right from wrong. Joseph Butler termed this faculty of the mind "conscience," and in more recent times this term has become the common one.

Modern behaviorists, to be sure, would not write of conscience as a mental faculty; they refer instead to "learned modes of reaction to stimuli." When one has been conditioned to respond in certain standard ways which are widely and strongly approved, one tends to find that one can break with such approved behavioral norms only after a genuine struggle and a stiff volitional conflict. In any case, whether we speak of the voice of conscience or of the voice of our group or of learned blockage and interference patterns, we often find that there are inhibitions to be overcome before we can break with the mores of our peers.

It has been suggested that a policeman, upholding the law, functions as a kind of government-supported externalized conscience. His mere presence in uniform suffices to warn us not to break, for example, the speed law that we are already bending a bit. Even animals below the human level can be trained to feel the force of such an externalized conscience. Cats, for example, can be trained not to sleep on the couch when humans are in the room. But it is difficult, to say the least, to teach them not to do so when no human observer is present to their senses. With human children and adults, by contrast, it is possible to develop an internalized conscience, which, even in the absence of all enforcers, will remind them, and even stimulate them strongly, not to do certain prohibited actions and to do certain required ones. The driver who stops his car at red traffic lights only when he sees or suspects that an officer is nearby has, like the cat, only an externalized conscience about this type of act, whereas one who habitually stops is, as we say, acting conscientiously—obeying, perhaps unconsciously, his internalized conscience.

That "the voice of conscience" is often effective seems clear, but it is also clear that it can and often does lose its effectiveness. A dutiful son may well adopt many of the mores of his father for a time and then gradually abandon them. If a person persists in violating his conscience, it will grow decrepit, bother him less and less effectively, and it may soon cease to deter him at all.

CONSCIENCE AS A RELIABLE GUIDE

As children many of us were taught that the voice of conscience is the voice of God and, hence, completely reliable. Some would claim, in more sophisticated terms, that although God gave us free will and does not infringe upon our freedom of choice, he nevertheless continues to lend us moral support. He gives us, through conscience, a means for distinguishing right from wrong. If we follow the guidance of conscience, we shall do our duty and act rightly. If we act contrary to its deliverances, we shall surely act wrongly.

There are, however, many difficulties with this kind of account and, indeed, with any other which claims that conscience is a sufficient guide to moral conduct.

First, the consciences of different people, whether members of the same or of different societies, often differ radically. Conscientious objectors to war and volunteers for wartime service usually disagree strongly as to the rightness of a given war. Cannibals do not share the conscientious objections to eating human flesh that vegetarians do, and both these groups differ from those who feel it is morally permissible to eat animal but not human flesh.

Second, there seem to be exceptions to all the edicts of conscience. Even within groups whose members share, say, a conscientious prescription against deliberately taking a human life, the exceptions that the various consciences allow to individuals vary greatly from one person to another. Lev Tolstoy, and presumably some Quakers, would insist that his conscience forbids the taking of a human life under any conditions. By contrast, although many of us verbally would fully accept the commandment not to kill, we would be likely in practice to find ourselves approving some acts of killing, for example in self-defense or in defense of others, and disapproving of some avoidances of killing, for example in a very deserving mercy case.

Third, conscience fails to provide guidance for many important and even some crucial moral questions. Many problems that we confront are so complex that we frankly have very little idea, and certainly no confirmed judgment or deliverance of conscience, as to which alternative is most worthy of being chosen. In many such cases, where getting adequate knowledge in the time available before a decision must be reached is impossible, we know in advance that we would be only too happy to do what is right if we could identify, with some reasonable degree of probability, the right alternative. A situation of this sort must frequently arise for people who cannot pass the

decision on to someone else. The president of the United States, for example, cannot avoid the responsibility for important decisions that must be made—very often on vastly less evidence than he would like to have. Similarly, there are many difficult problems to be decided by those of us who are less highly placed, problems where the decision will not indeed be world-shaking but where it will affect a number of lives in important ways. We often sweat with the desire to solve a difficult problem in the right way but are unable, in the time available before a decision must be taken, to find out which way is the right way. In complicated cases the relatively simple prescriptions of conscience tend to prove quite inadequate.

It is not that the prescriptions of our conscience are worthless; they are often of value in reminding us of the moral views which have been taken by other members of our peer group. Awareness also of the edicts that spring from the consciences of others with different backgrounds not infrequently throws light on our own problem. But in complicated and novel cases, the edict of another's conscience cannot provide us with certain knowledge as to what ought to be done.

SOURCES OF DELIVERANCES OF CONSCIENCE

Psychologists, anthropologists, and other social scientists have gathered empirical evidence as to various sources of the deliverances of conscience. Many of the edicts of our conscience seem to have come to us while we still rested at our mother's knee. These were usually simple in form but quite effective for many years. Others came from our fathers, from teachers, from preachers, from lecturers and writers, from friends whom we respected. This wide variety of the sources of the edicts that now emanate apparently from our own consciences explains many things about them: their vagueness, their variability, their changing authority over us. As suggested by behaviorists, at least some of them rest on conditioned responses instilled in us at an early age by repetitions we no longer remember.

Examination of a particular edict of conscience throws significant light on "our inner voices." Suppose we warn our sons, ages four and six, to stay off a railway trestle near our home. We say with great emphasis, "*Never* go out on that trestle, no matter what." One day the younger boy pursues his gay red ball onto the trestle. The older boy rushes to him and pulls him off the trestle just before a train crosses it. Will we punish him for breaking our "absolute" rule? Obviously not. Our consciously instilled rule, now a command of conscience, has its values, positive and negative. It needs supplementation as soon as increasing maturity permits rational consideration. And to this phase anyone who has attained knowledge and discretion should surely move on.

UNIVERSALIZABILITY OF MORAL PRESCRIPTIONS

Since the edicts of conscience have pedestrian empirical sources and are subject to exceptions, it was natural for Immanuel Kant to insist, through his categorical imperative, that every valid moral principle must hold universally: "So act that you can will the maxim or principle of your action to be a universal law, binding on the will of every rational being." This requirement has two facets. First, for an act to be moral it must be done not on whim or impulse or as a mere reflex response to stimuli, but in accordance with some moral principle or maxim. Second, this principle must be one that the agent is willing to have universally adopted. This requirement that a person should act only on a principle that he is willing to have universally adopted seems to introduce undesirable psychological factors that might tend to vary radically from one person to the next. Thus, a pessimist like Arthur Schopenhauer might approve of universal suicide and be willing to have everyone else do so, whereas an optimist might be willing to have everyone work toward increasing the population. Such a formulation of the universalizability principle would thus lead to incompatible moral edicts.

To eliminate such psychological factors and to state the principle in a way closer to Kant's intent, Richard M. Hare urges that a moral principle, to be applicable to a person *A,* must also be applicable in like circumstances to any similar person *B.* Although Hare's intent seems clear, he does not specify the degree of similarity required. Complete identity would make the principle useless. On the other hand, it seems clear that Hare was not suggesting, for example, that because it is right for *A* to make love to his wife, it is also proper that *B,* who is like *A* in various respects, should also make love to Mrs. *A.* Perhaps the universalizability thesis is best stated as follows: If it is right for *A* to do an act of kind *X* in a set of circumstances *C,* then it is right for any *B* who is like *A* in all relevant respects to do an act of kind *X* in circumstances like *C* in all relevant respects. So stated, the principle is analytically and thus necessarily true. But whether we can ever know in practice that both sets of circumstances and both agents are alike in all relevant respects is highly doubtful. It would be difficult, if not impossible, even to specify these respects. But we do know what is meant by this pre-

scription, and we sometimes know with a fair degree of probability that the required likenesses are present.

Because the universalizability principle is analytic, it is necessarily true. But it is an "If … then …" statement: If *A* should do *X* in *C,* then *B* should do *Y* in *D,* where the similarities between *A* and *B, C* and *D,* and *X* and *Y* meet the requirements previously mentioned. Quite aside from the difficulties of knowing whether or not these requirements are met, the statement tells us only that if its antecedent is true, its consequent is also true. But to know the antecedent to be true—that *A* ought to do *X* in *C*—we must turn to experience for an answer. To know anything to be good on the whole, we must know if its existence (or occurrence) is preferable to its nonexistence. To know any act to be right, we must know that no possible alternative is preferable to it. Such preferability presupposes empirical knowledge of values. The possibility of such knowledge is a matter of controversy, but many, including the present writer, believe it to be attainable.

See also Behaviorism; Butler, Joseph; Clarke, Samuel; Emotion; Hare, Richard M.; Hutcheson, Francis; Kant, Immanuel; Moral Motivation; Moral Rules and Principles; Price, Richard; Reason; Schopenhauer, Arthur; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Tolstoy, Lev (Leo) Nikolaevich; Volition.

Bibliography

WORKS OF HISTORICAL INTEREST

Aquinas, Thomas, Saint. "Conscience, Question 17, Articles 1–5," and "Synderesis, Question 16, Articles 1–3." In *Truth (Quaestiones Disputate de Veritate)*, translated by J. V. McGlynn. Chicago: H. Regnery, 1952–1954.

Balguy, John. *The Foundation of Moral Goodness.* London: John Pemberton, Part I, 1728; Part II, 1729.

Baylor, Michael G. *Action and Person: Conscience in Late Scholasticism and the Young Luther*. (Studies in Medieval and Reformation Thought, 20) Leiden: E. J. Brill, 1977.

Butler, Joseph. *Five Sermons Preached at the Rolls Chapel and a Dissertation upon the Nature of Virtue*. Indianapolis: Hackett, 1983.

Cicero. *De Finibus*. Translated by H. Rackham. Cambridge, MA: Harvard University Press, 1914.

Cicero. *De Natura Deorum*. Translated by H. McGregor. Hammondsworth, UK: Penguin, 1972.

Clarke, Samuel. *A Discourse concerning the Unchangeable Obligations of Natural Religion and the Truth and Certainty of the Christian Revelation,* 2nd corrected ed. London: Knapton, 1706.

Hutcheson, Francis. *An Inquiry into the Original of Our Ideas of Beauty and Virtue.* London, 1725.

Jonsen, Albert, and Stephen Toulmin. *The Abuse of Casuistry: A History of Moral Reasoning*. Berkeley: University of California Press, 1988.

Kant, Immanuel. *The Moral Law, or Kant's Groundwork of the Metaphysics of Morals.* Translated with analysis and notes by H. J. Paton. London: Hutchinson, 1948.

Potts, Timothy C. *Conscience in Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1980.

Price, Richard. *A Review of the Principal Questions and Difficulties in Morals.* London, 1758.

Shaftesbury, 3d Earl of (Anthony Ashley Cooper). *An Inquiry concerning Virtue or Merit.* London: Bell, Castle, and Buckley, 1699.

MORE RECENT WORKS

Andrew, Edward G. *Conscience and Its Critics. Protestant Conscience, Enlightenment Reason, and Modern Subjectivity*. Toronto: University of Toronto Press, 2001.

Broad, C. D. "Some Reflections on Moral-Sense Theories in Ethics." *PAS* 45 (1944–1945): 131–186.

Fingarette, Herbert. *Self-Deception*. London: Routledge & Keegan Paul, 1969.

Hoose, Jayne, editor. *Conscience in World Religions*. Notre Dame IN: University of Notre Dame Press, 1999.

Hursthouse, Rosalindk Gavin Lawrence, and Warren Quinn, editors. *Virtues and Reasons: Philippa Foot and Moral Theory: Essays in Honour of Philippa Foot*. Oxford: Clarendon Press, 1995.

Jacobs, Jonathan. *Choosing Character. Responsibility for Vice and Virtue*. Ithaca: Cornell University Press, 2001.

Kohlberg, Lawrence. "Conscience as Principled Responsibility: On the Philosophy of Stage Six." In *Conscience: An Interdisciplinary View*, edited by Gerhard Zecha and Paul Weingartner. Dordrecht: D. Reidel, 1987.

Kohlberg, Lawrence. *Essays on Moral Development*, Vol. 1: *The Philosophy of Moral Development*. San Francisco: Harper & Row, 1981.

Langston, Douglas C. *Conscience and Other Virtues*. University Park: Pennsylvania State University Press, 2001.

Nussbaum, Martha. *The Fragility of Goodness*. Cambridge, U.K.: Cambridge University Press, 1986.

Raphael, D. Daiches. *The Moral Sense.* London: Oxford University Press, 1947.

Ross, W. David. *Foundations of Ethics.* Oxford: Clarendon Press, 1939.

Ross, W. David. *The Right and the Good.* Oxford: Clarendon Press, 1930.

Ryle, Gilbert. "Conscience and Moral Convictions." In *Conscience*, edited by John Donnelly and Leonard Lyons. Staten Island NY: Alba House, 1973.

Slote, Michael. *From Morality to Virtue*. New York: Oxford University Press, 1992.

Smith, T. V. *Beyond Conscience. A Critical Examination of Various Doctrines of Conscience.* New York: McGraw-Hill, 1934.

Charles A. Baylis (1967)
Bibliography updated by William O'Neill (2005)

CONSCIOUSNESS

The term *consciousness* refers to several distinct, but related phenomena that figure in the mental functioning of people and other creatures.

KINDS OF CONSCIOUSNESS

One of these phenomena is closely tied to simply being awake. An individual is conscious if it is awake and responsive to sensory stimulation; a person or other creature that is asleep, in a coma, or knocked out is not conscious.

There are also other phenomena we refer to as consciousness. One is conscious *of* something if one senses or perceives the thing or has some suitable thought about it; being conscious of something is being aware of that thing. Because we use a grammatical object to specify what somebody is conscious of, it is convenient to call this phenomenon *transitive consciousness*, as against an individual's being awake and responsive to sensory input, which we can call *creature consciousness* (Rosenthal 1990).

We sometimes describe the states one is aware of as constituting one's current mental life as a stream of consciousness. But there are, in addition, thoughts, desires, feelings, and perceptions that occur outside that stream of consciousness, of which one is wholly unaware. Even though one is unaware of these states, they are nonetheless part of one's mental functioning. We call the states that occur in somebody's stream of consciousness *conscious*, in contrast with those of which that individual is wholly unaware. This is a third use of the term *conscious*. Because consciousness of this sort is a property of mental states, such as thoughts, desires, feelings, and perceptions, we can call it *state consciousness*.

Sometimes we focus deliberately and attentively on some feeling or perception we have; such focused awareness of our mental states is called *reflective*, or *introspective consciousness*. And we call the explicit consciousness of the self to which these states belong *self-consciousness*.

There is disagreement about what connections hold among these several kinds of consciousness. Some theorists hold that an individual cannot be creature conscious—that is, awake and responsive to sensory stimulation—unless at least some of its mental states are conscious. Doubtless that is true for ordinary humans; people are never conscious without being in some conscious states. But, if perceptions and feelings can occur without being conscious, there is in principle no reason why some creatures might be awake and responsive to sensory stimulation even though none of their feelings, perceptions, or other mental states are conscious states.

Not all theorists, however, accept that feelings, perceptions, thoughts, and desires can occur without being conscious. Even Sigmund Freud (1961), who championed the idea of unconscious desires and thoughts, drew the line at qualitative states, such as sensations and feelings. All feelings, he held, are conscious, though we can loosely characterize feelings as unconscious when one is unclear or mistaken about what they are about.

Others, such as Thomas Nagel (1974) and John R. Searle (1990, 1992), accept that nonconscious states occur that function in ways similar to conscious feelings, perceptions, and thoughts, but deny that those nonconscious states are full-fledged feelings, perceptions, or thoughts.

It is also a matter of some controversy whether mental states are conscious in virtue of one's being conscious *of* those states. Earlier writers, such as René Descartes (1984–1991) and John Locke (1975), always described the states we now call conscious as states that one is conscious of. But, because they also held that we are conscious of all our mental states, they saw no need to use the term *conscious* to mark a distinction between those mental states we are conscious of and those we aren't. When, in the late nineteenth century, it became widely accepted that individuals are in some mental states of which they are wholly unaware, the term *conscious* came to mark the contrast between those mental states one is conscious of and those one is not.

Though not all theorists agree that a mental state's being conscious involves one's being conscious of that state, that has long been the dominant view.

QUALITATIVE CONSCIOUSNESS

Mental states fall into three broad groups. Some, such as beliefs, thoughts, desires, hopes, and expectations, have intentional content that can be described by a sentential clause. Thinking is always thinking *that* something is the case; the clause beginning with *that* specifies what it is that one thinks. Similarly with one's desires, hopes, and expectations. The intentional content, described by a *that* clause, specifies what it is that one desires, hopes, or expects.

Pains and other sensations, by contrast, have no intentional content, but instead exhibit some qualitative character, such as the quality of painfulness or the color qualities of visual sensations. A third group of states includes emotions and perceptions, which exhibit both qualitative character and intentional content.

Many theorists hold that the consciousness of qualitative states is something different from the consciousness of other mental states, and that it demands special treatment. Giving an informative theoretical account of qualitative consciousness, on their view, faces special difficulties.

Thus Ned Block (1995) has urged that qualitative states exhibit a special kind of consciousness, which he calls *phenomenal consciousness*, or *phenomenality*. A state has phenomenal consciousness, according to Block, if there is something it is like to be in that state, which happens only when the state has some qualitative character. There is nothing it is like simply to think or believe something, even when one's thought or belief is conscious. Phenomenal consciousness occurs only with states that have qualitative character.

Block (1995) distinguishes a state's having phenomenal consciousness from its having what he calls *access consciousness*. A state is access conscious if its content is poised to figure in reasoning and in the rational control of action and speech. Some qualitative states, which exhibit phenomenal consciousness, also exhibit access consciousness; intuitively, they are the qualitative states one is conscious of. By contrast, when a state is phenomenally conscious but not access conscious, one is wholly unaware of the state. And there is often compelling empirical or theoretical reason to think that qualitative states of which one is unaware do occur (see the next section in this entry).

Block's notion of access consciousness echoes Daniel C. Dennett's (1993) idea that a state's being conscious consists in its having "cerebral celebrity," that is, if it has a widespread effect on memory and on the control of behavior. It also accords with the cognitive theory of consciousness advanced by the psychologist Bernard J. Baars (1988), on which a state is conscious if it occurs in a global workspace that maximizes its connections with other states and behavior.

PROBLEMS ABOUT QUALITATIVE CONSCIOUSNESS

A state's being access conscious consists in its having suitable connections with other states and with behavior. So the notion of access consciousness invites a functionalist account (Lewis 1972, Putnam 1975), on which a state's mental properties are a matter of such connections. Many theorists, however, deny that any such an account can work for qualitative consciousness. They insist that, because conscious qualitative character is an intrinsic property of sensations, it cannot be understood in terms of connections that sensations have with other things.

EXPLAINING CONSCIOUS QUALITIES. The new physics pioneered in the seventeenth century by Galileo Galilei, Descartes, and Isaac Newton holds that we can explain the nature and behavior of physical objects only insofar as we can describe them in mathematical terms. Since commonsense physical qualities, such as color and sound, seem to resist mathematical description, some have followed Locke in construing such properties as powers to cause the corresponding mental qualities. But conscious mental qualities also resist mathematical description, and it may seem that no parallel move is possible for them. Many conclude that conscious mental qualities lie outside the reach of physical explanation, and possibly, therefore, any informative explanation. Thus, Locke argued that sensations either are nonphysical states or, if they are states of material bodies, they must be "superadded" to those bodies by God.

In a somewhat similar spirit, Joseph Levine (2000) has argued that there is an "explanatory gap" that blocks any intelligible explanation of conscious qualitative states in terms of physical processes. Similarly, Nagel (1974) has argued that none of the available naturalist theories of mind can explain what it's like for one to be in a mental state. And David Chalmers (1996) has described as the "hard problem" of consciousness the question why relevant brain processes are accompanied by qualitative consciousness at all, and why particular brain events are accompanied by specific types of mental quality.

Levine has urged that, though we cannot explain qualitative consciousness in physical terms, qualitative consciousness might nonetheless be physical in nature. Others conclude instead that qualitative consciousness cannot have a physical nature, arguing that any such physical nature would make possible a physical explanation of qualitative consciousness. And Colin McGinn (1991) has argued that, though consciousness is physical in nature, we lack the cognitive capacity to understand how that is possible.

It is unclear, however, that the considerations used to support these views are compelling. We rely on well-developed theories to draw systematic connections among things in nature. Even the most ordinary connections among natural processes seem surprising and unintelligible without any commonsense theory to provide context, and come to seem rational only when subsumed by some suitable theory. So it may be that the ties between conscious mental qualities and brain processes now

seem unintelligible only because we still have no well-developed theory that links them. But by itself, that current lack gives us no reason to doubt that we will some day have such a theory. And coming to have one would likely overcome any prior intuitive concerns about explaining qualitative consciousness, just as physics and chemistry have made intuitively acceptable various explanations of our commonsense world that would previously have seemed outlandish.

Introspection may also seem to support intuitive doubts about whether rational explanation of conscious qualitative character is possible, since introspection provides no clue about how such an explanation might proceed. But we have no reason to think that introspection would help here; introspection can at most tell us about the qualities themselves, not how they connect with other things.

THE KNOWLEDGE ARGUMENT. Frank Jackson (1982, 1986) has argued that conscious qualitative states are not physical because, even if one knew everything physical there is to know about our psychological and neural functioning, one would not thereby know what it's like for one to be in particular qualitative states. Jackson imagines a neuroscientist who knows everything physical about visual functioning but has never seen anything except in black, white, and gray. Still, Jackson argues, this neuroscientist, on consciously seeing red for the first time, would learn something new, namely, what it's like for one to see red. Since the neuroscientist already knew everything physical about seeing red, the new knowledge of it what it's like for one to see red cannot be knowledge of something physical.

Jackson (2003) has since repudiated this argument, maintaining now that what it's like for one to see red is purely a matter of intentional content. This view is a version of representationalism, which is discussed later in this entry. Others have responded differently to Jackson's original argument, urging that what one would learn on first consciously seeing red is not factual knowledge, but only a kind of acquaintance (Churchland 1985) or an ability to recognize the quality in question (Lewis 1990; cf. Loar 1997).

QUALITIES AND CONSCIOUSNESS. Some theorists contend that qualitative states of which we are in no way conscious cannot occur. Indeed, the very term *qualia* (singular *quale*) is often applied to mental qualities with the implication that such qualities cannot occur without being conscious. But there is compelling reason to holdthat mental qualities do occur outside our stream of consciousness.

Individuals sometimes perceive things without being at all aware that they are doing so. In so-called masked priming experiments (Marcel 1983), subjects briefly exposed first to one visual stimulus and then to another may consciously see only the second. Nonetheless, it is plain that subjects do see the first stimulus, since it affects subsequent behavior in ways characteristic of seeing those stimuli. Thus, subjects who report seeing only the second stimulus can nonetheless make strikingly accurate guesses about the first.

There are other such cases. Individuals with lesions in the cortical area primarily responsible for vision may be wholly unaware of seeing a stimulus and yet guess about its visible character, again with great accuracy, exhibiting what Lawrence Weiskrantz (1997) has called blindsight. These individuals see stimuli, but are not in any way conscious of seeing them. And subliminal perceiving, of which one is wholly unaware, sometimes occurs even in everyday situations.

Not only do qualitative states sometimes occur without being conscious; there are circumstances in which we are conscious of ourselves as being in qualitative states that are different from those we are actually in. John Grimes (1996) reported that subjects will continue to see a highly salient object as unchanged in color or other respects if the relevant changes occur during a saccade, since no visual input reaches the brain during saccades. A subject may thus attentively look at something red but be conscious instead of seeing green. Such a subject would presumably have a sensation of red, despite being conscious of having a sensation of green. Our consciousness of our qualitative states can sometimes be strikingly inaccurate.

According to Block (1995), cases of qualitative states that occur outside our stream of consciousness are phenomenally conscious states that lack access consciousness. Access consciousness makes the difference, he urges, between qualitative states of which we are intuitively aware and those of which we are not. But it's likely that even qualitative states that intuitively occur in one's stream of consciousness sometimes lack access consciousness, on Block's official definition. Visual states near the periphery of our visual field are conscious but are not, without some shift in attention, poised to figure in any general way in reasoning and the rational control of action and speech. Similarly with other perceptual states that lie outside our focus of attention but are nonetheless part of our stream of consciousness. It is

likely that access consciousness has more to do with attention than with consciousness.

QUALITY INVERSION Qualitative states figure in perceiving. There is a distinctive mental quality that occurs when one sees something red and a different quality when one sees something green; similarly for perceptible properties accessible by modalities other than vision. That raises the question whether particular mental qualities might play different perceptual roles from one individual to another, or even different roles in the same individual at different times. The question is not about the slight variations in the way people see things, which are detectable in standard ways, but about whether particular mental qualities could play different perceptual roles in ways undetectable by others.

Locke held that such inversion of mental qualities is at least conceivable, and many contemporary theorists share that view. This idea very likely reflects a conviction that mental qualities are individuated solely by the way one is conscious of them, that is, by how they appear to consciousness. If any other factors do figure in the individuation of mental qualities, those factors would enable the detection in others of inversion in the perceptual roles of their mental qualities.

But if mental qualities were individuated only by how we are conscious of them, they would differ only in the way they appear to consciousness. And then mental qualities could not occur without being conscious. Indeed, the evidence that mental qualities do occur without being conscious provides ways of determining their occurrence independent of consciousness. So, that evidence also suggests that any conceivable quality inversion would have to be detectable. It is therefore likely that any satisfactory way of individuating mental qualities will rely on their role in perceiving, independent of whether that perceiving occurs consciously (Rosenthal 2005).

CONSCIOUSNESS AND INTENTIONALITY

As noted earlier, intentional states, such as thoughts, desires, doubts, and expectations, occur both consciously and not consciously.

Freud posited intentional states that are not conscious as the best explanation of various otherwise inexplicable conscious thoughts and desires and various bits of behavior. Thus, a person may do just those things and have just those conscious thoughts and desires that we would expect if the person also had certain other thoughts and desires. And, if the person is unaware of being in those other thoughts and desires, we can best explain the behavior and conscious states by supposing that the person has those thoughts and desires, but they are simply not conscious. Such reasoning again invites a functionalist account of intentional states, on which the intentional properties of a state is a matter of its connections with other states, behavior, and sensory stimulation. But even apart from a functionalist account, it is widely accepted that intentional states with particular contents have characteristic causal connections with other intentional states and with behavior, and that is all Freud's argument requires. Such reasoning is compelling, moreover, independent of the special kinds of case that interested Freud.

Experimental work in social psychology shows that subjects sometimes report having beliefs or desires that would make sense of a situation or conform to social expectations, despite compelling evidence that these subjects do not actually have those beliefs and desires (Nisbett and Wilson 1977). Not only are we sometimes unaware of our thoughts and desires; in such confabulatory cases we are conscious of ourselves as having thoughts and desires that we do not have.

Searle (1990) has argued that intentional states, properly so called, cannot occur without being conscious. As he notes, one's thoughts and desires always represent things in terms of some aspects and not others; Oedipus had a desire to marry a particular woman, but his desire did not represent that woman as his mother. And, as Oedipus's case illustrates, how one's intentional states represent the things they are about makes a difference both to one's mental life and one's behavior.

According to Searle, the way one's intentional states represent things cannot make a difference to one unless those states are conscious. He concludes that genuine thoughts and desires cannot occur without being conscious. But the way one's thoughts and desires represent things can make a difference to one even if those states are not conscious. A thought or desire need not be conscious to affect one's other intentional states and one's behavior, and it will affect those things differently depending on the way it represents things. Genuine intentional states can occur without being conscious.

THEORIES OF CONSCIOUSNESS I

Theories of consciousness often rely on the traditional idea that a state's being conscious involves one's being conscious in some way *of* that state. States of which one is in no way conscious are not conscious states. When we are conscious of something, moreover, we can tell others

about it. So a standard test for whether somebody is in a conscious state is whether that person can report being in the state. If somebody can report having a particular thought, feeling, or perception, that state is conscious; if the person cannot report being in the state, it is not. This rule of thumb underlies typical methodology in experimental psychology no less than everyday practice.

But the commonsense observation that mental states are states of which we are conscious goes only so far. A theory of consciousness must also specify how it is that we are conscious of those states. One important feature of our consciousness of those states was highlighted by Descartes, who insisted that we are immediately conscious of our mental states. When we are conscious of a mental state, it seems that nothing mediates between that state and our awareness of it. A theory of consciousness must explain this sense of immediacy in the way we are aware of our conscious states.

INNER SENSE. We also seem to be immediately conscious of things when we perceive them; nothing seems to mediate between the things we perceive and our perceptions of them. This encourages the hypothesis, advanced by Locke, Immanuel Kant (1998), and others, that mental states are conscious because we sense or perceive them. A thought, feeling, or perception is conscious because one is aware of that state by way of some faculty of inner sense, or some internal monitoring mechanism that involves the higher-order perceiving of that state. This theory has traditionally been the most widely held explanation of consciousness; contemporary advocates include David M. Armstrong (1978) and William G. Lycan (1996).

But there are difficulties with this theory. Sensations and perceptions always exhibit some qualitative character; sensing a red object, for example, involves a sensation's having a mental quality of red, as against a mental quality of blue, green, and so forth. Our consciousness of our mental states, however, does not involve any qualitative character. This is obvious when the state we are are conscious of is a thought or desire, which itself has no qualitative character; plainly, no mental quality figures in the way we are conscious of those intentional states.

Qualitative character does figure when we are conscious of sensations and perceptions. But these mental qualities are just the qualities we are conscious of our sensations and perceptions as having. As Aristotle (1993) noted, there are no higher-order qualities in virtue of which we are conscious of our qualitative mental states, in the way the mental quality of red enables us to see red objects. Our higher-order awareness of our conscious states may resemble perceiving in other ways, but qualitative character is so central to perceiving that no form of awareness that fails to involve mental qualities can count as perceiving.

Inner-sense theorists often urge that the higher-order sensing or perceiving they posit serves the function of monitoring our mental states, much as perceiving monitors external objects and bodily conditions. But perceiving is not the only way that the mind might monitor itself. And cases of confabulatory awareness, which Richard E. Nisbett and Timothy DeCamp Wilson have demonstrated, do not in any case fit neatly with a model based on monitoring.

INTRINSIC THEORIES. But we need not appeal to inner sense to capture the apparent immediacy of our consciousness of many mental states. If our awareness of our conscious states were internal to those states themselves, nothing could mediate between a state and one's awareness of it; such consciousness would be intrinsic to each conscious state. This theory, advanced by Franz Brentano (1973) and possibly Aristotle, also has a number of contemporary advocates.

But the intrinsic theory also faces difficulties. There are thoughts and desires that we sometimes have consciously and other times not. A sensation that results from a particular stimulus may be conscious if that stimulus occurs alone, but not conscious if the very same stimulus is followed in a suitable way by a second, masking stimulus (Marcel 1983). It is unclear how we can explain such variation if consciousness is literally built into our mental states.

The problem is particularly pressing when one particular state passes between being conscious and not being conscious. Some perceptual or bodily sensations that are not very intense may be conscious or not depending on where one focuses one's attention. But, since shifts in attention are extrinsic to particular sensations, such shifts should leave consciousness unaffected if consciousness is indeed an intrinsic aspect of mental states.

Brentano held that consciously hearing something makes us conscious of two things: the sound one hears and the hearing itself. And he maintained that we are conscious of the hearing in the way that having a thought about something makes us conscious of that thing, even when we don't perceive it. The intentional content of hearing something, according to Brentano, makes us conscious both of the thing heard and the hearing itself.

Perhaps hearing can have two intentional objects in this way. But there are other cases for which Brentano's

model does not work. Doubting something does not make one conscious of the thing one's doubt is about. Consider, then, a case of doubting that it is raining. Even if one's doubt is about both the rain and the state of doubting itself, that will not make one conscious of the doubting. A mental affirmation that one has that doubt would make one conscious of the doubting. But that mental affirmation could not be intrinsic to the doubting, since no mental state involves more than one mental attitude. Similar considerations apply to wondering about something, and many other mental attitudes.

Inner sense and Brentano's intrinsic theory both sought to explain the way our awareness of our conscious states is immediate. But all we really need to explain is why such awareness appears to be immediate, since we do not know that it actually is. Indeed, perceiving is also subjectively unmediated, but we know that there is much that actually mediates between our perceptions and the things we perceive. So the same may well be so with the way we are aware of our conscious states. Despite the subjective impression of immediacy, there may well be mediation we are not subjectively aware of. All we need to explain is the subjective sense of immediacy, and neither the analogy with perceiving nor the intrinsic theory is required for that.

HIGHER-ORDER THOUGHTS. On Brentano's intrinsic theory, every conscious state makes us conscious of itself, in much the way that having a thought about something makes one conscious of that thing. This theory cannot work, at least for cases like doubting and wondering. Inner sense, by contrast, faces the difficulty that, because the awareness of our conscious states does not involve higher-order mental qualities, that awareness cannot be sensing or perceiving. This suggests combining features of the two theories so as to avoid the difficulties of each. Perhaps we are aware of our conscious states by having thoughts about them, as Brentano urged, but those thoughts are distinct from the states we are conscious of, as inner sense maintains about the higher-order perceptions it posits.

This appeal to higher-order thoughts that are distinct from the mental states they make us conscious of avoids the foregoing difficulties that face inner sense and Brentano's theory. The higher-order-thought theory, advanced by David M. Rosenthal (1986, 1990, 2005) and others, also allows for an explanation of the subjective immediacy of our awareness of our conscious states; the theory can require that these higher-order thoughts are independent of any conscious inference. If we are unaware of any inference on which a higher-order thought relies, we will be unaware of any mediation between the states we are aware of and our awareness of them. So, such awareness will be subjectively unmediated. Indeed, we would seldom be aware of these higher-order thoughts, since a third-order thought would be needed for any second-order thought to be conscious. And our typically being unaware of our higher-order thoughts would enhance the subjective sense that our consciousness of our mental states is immediate.

Critics have urged two major difficulties for this theory. One involves the possibility that higher-order thoughts will sometimes misrepresent what mental states we are in. But it is arguable that consciousness does sometimes misrepresent things, as in the confabulatory cases noted earlier. And there are very likely psychological pressures that prevent such misrepresentation from becoming too extreme.

Nonetheless, some have insisted that such misrepresentation cannot occur, in effect relying on the traditional idea, advanced by Descartes, Locke, and others, that the mind is transparent to itself. But there is compelling reason to reject that transparency claim. Confabulatory consciousness and other phenomena show that consciousness does occasionally mislead us; moreover, others sometimes know what we are thinking and feeling better than we ourselves do. Consciousness is neither infallible nor exhaustive.

Another challenge to the higher-order-thought theory pertains specifically to conscious qualitative states. How can higher-order thoughts, which themselves lack qualitative character, result in qualitative states' being conscious? How can simply having a thought result in there being something it is like for one to be in a qualitative state?

This challenge echoes the concern that an explanatory gap may make it impossible to understand how conscious mental qualities could arise as a result of particular neural events. Put most generally, the concern is how conscious qualities can result from anything else. But as noted earlier, connections among things in nature seem intelligible only when we have a well-established theory that subsumes those connections. Since higher-order thoughts are seldom conscious, introspection cannot tell us whether they result in conscious qualities. But it may be that the connections higher-order thoughts have with the thoughts we have about the things we perceive result in our being conscious of the mental qualities that figure in such perceiving.

In any case, no alternative theory has a response to this challenge that is at all satisfactory. By itself, simply positing that our awareness of qualitative states is intrinsic to those states does nothing to explain why there is something it is like for one to be in those states. And inner sense faces a regress, since it could help only if the higher-order perceptions themselves had conscious mental qualities, and we would then need to explain what gives rise to those higher-order conscious qualities.

THEORIES OF CONSCIOUSNESS II

The foregoing theories differ about whether our awareness of our conscious states is intrinsic to those states or external to them, and about whether that awareness is due to our perceiving those states or to our having thoughts about them. But there are other issues about which theories differ as well.

DISPOSITIONAL THEORIES. Peter Carruthers (2000) has argued that a state's being conscious does not require the actual occurrence of a higher-order thought, but only a disposition for such a higher-order thought to occur. Carruthers urges that having an actual higher-order thought for each conscious state would result in cognitive overload, unlike one's merely being disposed to have higher-order thoughts.

But there is no reason to think that our cortical resources cannot accommodate actual higher-order thoughts, and dispositions would themselves make substantial cortical demands. Nor is it obvious that dispositions will do. Since being disposed to have a thought about something does not make one conscious of anything, merely being disposed to have a higher-order thought would not make one aware of one's mental states. Carruthers seeks to meet this difficulty by endorsing the view that the intentional content a state has is partly a matter of what other states it is disposed to cause. So, when a state is disposed to cause a higher-order thought, that very state has higher-order content, which makes one conscious of that state. But, since the state itself has the higher-order content, this view faces the same difficulties that tell against Brentano's theory.

REPRESENTATIONALISM. Some have sought to meet the challenge about conscious mental qualities by denying that there are any. According to representationalism, we are never conscious of any mental qualities, but only the perceptible properties of physical objects, and the states in virtue of which we are conscious of them are purely intentional states. When we see something red, on this view, the only quality we are aware of is the redness of the thing seen; we are not in addition aware of some mental red. Advocates of this view, such as Gilbert Harman (1990), Dennett (1991), Fred Dretske (2000), Armstrong, and Lycan, point out that we never seem to be conscious of two distinct qualities of red, nor to switch from being conscious of the redness of physical objects to being conscious of a mental quality of the seeing itself. Descartes also espoused a form of representationalism, since he regarded all mental phenomena as having only intentional properties, and construed sensations either as purely intentional states or as nonmental bodily states.

But, as Wilfrid Sellars (1963), Sydney Shoemaker (1996), and Rosenthal (2005) have argued, perceptual sensations resemble and differ in ways that reflect the similarities and differences among perceived physical properties. And it is natural to construe the properties in virtue of which those sensations resemble and differ as mental qualities. When we introspectively attend to our qualitative states, moreover, we sometimes become conscious of the relevant qualities *as* qualities of our experiences. So it may well be that we are, after all, often aware of mental qualities that our qualitative states exhibit.

Some theorists, such as Dretske (1993) and Searle (1992), reject the idea that a mental state's being conscious is a matter of one's being conscious of that state. A state's being conscious, on their view, does not involve some higher-order awareness of that state. Rather, according to Dretske (1993), a state is conscious if, in virtue of one's being in that state, one is conscious of something. This is sometimes called a first-order theory of consciousness, in contrast to theories that posit some higher-order awareness.

This account faces a difficulty, however. Perceptions sometimes occur without being conscious. But it is arguable that even those perceptions make us conscious of things. If perceiving something primes one for some conscious state or some behavior, then one was conscious of the thing one perceived even if it did not seem to one that one perceived it. On Dretske's (1993, 2000) view, however, any state in virtue of which one is conscious of something is conscious. And that has the unwelcome result that even the perceptions we seem subjectively not to have are conscious.

Searle (1992) holds that we can subjectively draw no distinction between a conscious state and one's consciousness of it. He concludes that no higher-order awareness figures in a state's being conscious. But when we focus introspectively on our conscious states, we are often aware both of the state thus scrutinized and of the

scrutinizing itself. And even if we could not draw that distinction subjectively, we might still have sound theoretical reasons to insist on it.

DENNETT'S THEORY. Dennett (1991) has developed an important theory of consciousness, which emphasizes cases in which consciousness misrepresents what mental states we are in. Visual information that is not central to our focus of attention can be highly degraded, but still we seem subjectively to see things in sharp detail throughout our field of vision. Dennett argues that consciousness extrapolates from available visual information to create a full picture of the environment, in effect filling in missing visual information and providing missing details.

In thus distinguishing the way consciousness represents things from our actual visual states, Dennett's (1991) view resembles higher-order theories, on which our higher-order awareness of mental states is distinct from the states themselves, and so can misrepresent them. But Dennett rejects such higher-order views, arguing that there is no distinction between the way things appear and our awareness of how they appear. So, he construes the divergence between consciousness and visual states not in terms of two mental levels—perceiving and our consciousness of perceiving—but rather as the difference between the way things consciously appear and the subpersonal neural events in virtue of which things appear that way.

Dennett shares with the higher-order-thought theory a view of consciousness as a kind of self-interpretation. On both views we interpret ourselves as being in various commonsense psychological states. But, unlike higher-order-thought theorists, Dennett denies that we are actually in any of the commonsense psychological states we interpret ourselves as being in. The only states that figure in psychological functioning are subpersonal neural events of content fixation, complex patterns of which subserve such functioning.

Searle (1992) and Dennett (1991, 1993) both reject any distinction between the mental states one is in and one's awareness of those states, but they do so for different reasons. Searle rejects that distinction because he holds that we cannot draw it subjectively. Dennett, by contrast, maintains that the psychological states we are conscious of ourselves as being in do not actually occur. But it is arguable that suitable patterns of the subpersonal events of content fixation Dennett countenances actually constitute the mental states of commonsense psychology. If so, we can distinguish between those commonsense mental states and our higher-order awareness of them.

NEURAL CORRELATES, FUNCTION, AND THE SELF

Whether or not conscious qualitative states are physical in nature, few doubt that something specific in brain functioning correlates with qualitative consciousness. This has led to speculation about what that neural correlate of consciousness (NCC) is.

According to Francis Crick and Christoph Koch (1990) the NCC involves the occurrence of synchronized neural oscillation of 35 to 75 hertz in sensory cortex, a synchrony sufficient for a vigorous coalition of neurons firing together. One thing that favors this hypothesis is that such synchronized neural oscillation seems to figure in the way different qualitative properties are bound together in conscious experience. As Anne Treisman (1986) has shown, visual qualities pertaining to color, shape, motion, and orientation occur independently in the early stages of visual processing; so there is a *binding problem* of explaining how they come together in conscious experience. But these qualities are bound together even when qualitative states are not conscious. So, the neural factors operative in such binding may not be the same as those responsible for qualitative consciousness, and synchronized oscillation may subserve only binding, independent of consciousness.

Mental functioning plays a variety of roles, allowing animals to negotiate their way in the world and to satisfy various needs and desires. What, then, is the specific function of consciousness? The answer depends on which kind of consciousness is at issue. Creature consciousness, which consists in an animal's being awake and responsive to sensory stimulation, plainly functions to enable an animal to satisfy needs and avoid danger. Similarly with transitive consciousness, which consists in a creature's being conscious *of* various things.

State consciousness, by contrast, consists in a state's occurring in a creature's stream of consciousness, and it is somewhat less clear what function this has. A mental state's function depends on its causal connections with other states and with behavior and sensory stimuli, and the causal properties of thoughts and desires are mainly a matter of their intentional content, whether or not they are conscious. Similarly, the causal properties of qualitative states depend on their mental qualities; visual sensations of red interact causally in ways suitably different from visual sensations of green, again whether or not they are conscious. So it is unclear what additional function might result from these states' being conscious.

One standard answer is that such consciousness functions to enhance reasoning and planning; perhaps decisions and thinking will be more rational if one is conscious of one's thoughts and desires. This idea underlies Block's (1995) claim that a state is access conscious if its content is poised to figure in the rational control of action and speech, Baars's (1988) related suggestion that conscious states occur in a global workspace, and Dennett's (1993) that consciousness is cerebral celebrity.

But many thoughts and desires have global effects on other states and on behavior even when they are not conscious. And even when planning and thinking is not conscious, it often is rational, as when we solve problems by sleeping on them. Indeed, this is just what we should expect if the causal potential of thoughts and desires depends mainly on their intentional content. There is, moreover, compelling evidence that we are conscious of our decisions only after those decisions have been made (Libet 1985), so that being conscious of those decisions cannot affect whether we make them.

According to Dretske (1993), any state in virtue of which one is transitively conscious of something is a conscious state. So on that view, the function of state consciousness coincides with that of transitive consciousness. But if, instead, a state is conscious just in case one is in some suitable way conscious of that state, the function of state consciousness will rather be whatever function is added by one's being thus conscious of the state. And that may be relatively marginal.

When we introspect our mental states by deliberately and attentively focusing on them, we are conscious of the states we introspect as states of ourselves, and we are in that way conscious of ourselves as centers of consciousness. There are several questions about the nature of such self-consciousness. David Hume (1978) urged that, though we are aware of many of our mental states, we are not aware of anything in addition to those states which we might call a self. Hume was operating with a perceptual model of awareness, and it is plain that we do not perceive such a self. But we are sometimes conscious of things not only by perceiving them, but also by having thoughts about them as being present. So having higher-order thoughts to the effect that one is in various mental states will make one conscious of oneself as being in those states, and hence conscious of the self to which those states belong.

As Descartes and Kant stressed, our mental states are conscious in a way that seems to involve an important unity; we are conscious of them as all being states of a single unitary self or center of consciousness. It is not obvious whether some actual unity underlies this appearance of conscious unity (see Marcel 1993, Rosenthal 2005). But even explaining that subjective appearance requires more than simply explaining the consciousness of the relevant mental states.

See also Knowledge Argument; Qualia; Subjectivity.

Bibliography

Aristotle. *De Anima: Books II, III* (with passages from Book I). 2nd ed. Translated by D. W. Hamlyn, with added material by Christopher Shields. Oxford, U.K.: Clarendon Press, 1993.

Armstrong, D. M. "What Is Consciousness?" *Proceedings of the Russellian Society* 3 (1978): 65–76. Reprinted in expanded form in *The Nature of Mind*, by D. M. Armstrong (St. Lucia, Queensland: University of Queensland Press, 1980, pp. 55–67).

Baars, Bernard J. *A Cognitive Theory of Consciousness*. New York: Cambridge University Press, 1988.

Baars, Bernard J., William P. Banks, and James B. Newman, eds. *Essential Sources in the Scientific Study of Consciousness*. Cambridge, MA: MIT Press/Bradford Books, 2003.

Block, Ned. "On a Confusion about a Function of Consciousness." *The Behavioral and Brain Sciences* 18 (2) (June 1995): 227–247.

Block, Ned, Owen Flanagan, and Güven Güzeldere, eds. *The Nature of Consciousness: Philosophical Debates*. Cambridge, MA: MIT Press/Bradford Books, 1997.

Brentano, Franz. *Psychology from an Empirical Standpoint* [1874]. Translated by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister; edited by Linda L. McAlister. London: Routledge & Kegan Paul, 1973.

Carruthers, Peter. *Phenomenal Consciousness: A Naturalistic Theory*. New York: Cambridge University Press, 2000.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.

Churchland, Paul M. "Reduction, Qualia, and the Direct Introspection of Brain States." *The Journal of Philosophy* 82 (1) (January 1985): 8–28.

Clark, Austen. *Sensory Qualities*. Oxford, U.K.: Clarendon Press, 1993.

Crick, Francis, and Christoph Koch. "Towards a Neurobiological Theory of Consciousness." *Seminars in the Neurosciences* 2 (1990): 263–275.

Dennett, Daniel C. *Consciousness Explained*. Boston: Little, Brown, 1991.

Dennett, Daniel C. "The Message Is: There Is No *Me*dium." *Philosophy and Phenomenological Research* 52 (4) (December 1993): 919–931.

Descartes, René. *The Philosophical Writings of Descartes*, 3 vols. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch (vol. III with Anthony Kenny). New York: Cambridge University Press, 1984–1991.

Dretske, Fred. "Conscious Experience." *Mind* 102 (406) (April 1993): 263–283. Reprinted in *Perception, Knowledge, and Belief*, by Fred Dretske, pp. 113–137.

Dretske, Fred. *Perception, Knowledge, and Belief: Selected Essays*. New York: Cambridge University Press, 2000.

Freud, Sigmund. *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. 19, The Ego and the Id*. Translated and edited by James Strachey. London: Hogarth Press, 1961.

Gennaro, Rocco J., ed. *Higher-Order Theories of Consciousness: An Anthology*. Philadelphia, PA: John Benjamins, 2004.

Grimes, John. "On the Failure to Detect Changes in Scenes across Saccades." In *Perception*, edited by Kathleen Akins, 89–110. New York: Oxford University Press, 1996, pp. 89–110.

Harman, Gilbert. "The Intrinsic Quality of Experience." In *Philosophical Perspectives. Vol. 4, Action Theory and Philosophy of Mind*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1990, pp. 31–52.

Hume, David. *A Treatise of Human Nature* [1739]. 2nd ed., edited by L. A. Selby-Bigge. Text revisions and variant readings by Peter H. Nidditch. Oxford, U.K.: Clarendon Press, 1978.

Jackson, Frank. "Epiphenomenal Qualia." *Philosohical Quarterly*. XXXII, 127 (April 1982): 127–136.

Jackson, Frank. "Mind and Illusion." In *Minds and Persons*, edited by Anthony O'Hear. New York: Cambridge University Press, 2003, pp. 251–71. Reprinted in *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, edited by Peter Ludlow, Yujin Nagasawa, and Daniel Stoljar (Cambridge, MA: MIT Press/Bradford Books, 2004, pp. 421–442).

Jackson, Frank. "What Mary Didn't Know." *The Journal of Philosophy* 83 (5) (May 1986): 291–295.

Kant, Immanuel. *Critique of Pure Reason* [1787]. Translated and edited by Paul Guyer and Allen W. Wood. New York: Cambridge University Press, 1998.

Levine, Joseph. *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press, 2000.

Lewis, David. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (3) (December 1972): 249–258.

Lewis, David. "What Experience Teaches." In *Mind and Cognition: A Reader*, edited by William G. Lycan. Cambridge, MA: Basil Blackwell, 1990.

Libet, Benjamin. "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action." *The Behavioral and Brain Sciences* 8 (4) (December 1985): 529–539.

Loar, Brian. "Phenomenal States." In *The Nature of Consciousness: Philosophical Debates*, edited by Ned Block, Owen Flanagan, and Güven Güzeldere. Cambridge, MA: MIT Press/Bradford Books, 1997, pp. 597–616.

Locke, John. *An Essay concerning Human Understanding*. Edited from the fourth (1700) edition by Peter H. Nidditch. Oxford, U.K.: Clarendon Press, 1975.

Lycan, William G. *Consciousness and Experience*. Cambridge, MA: MIT Press/Bradford Books, 1996.

Marcel, Anthony J. "Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition." *Cognitive Psychology* 15 (1983): 197–237.

Marcel, Anthony J. "Slippage in the Unity of Consciousness." In *Experimental and Theoretical Studies of Consciousness*, edited by Gregory R. Bock and Joan Marsh. Chichester, NY: Wiley, 1993, pp. 168–186.

McGinn, Colin. *The Problem of Consciousness: Essays towards a Resolution*. Oxford, U.K.: Basil Blackwell, 1991.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Nagel, Thomas. "What Is It Like to Be a Bat?" *The Philosophical Review* 83 (4) (October 1974): pp. 435–450. Reprinted in *Mortal Questions*, by Thomas Nagel (Cambridge, U.K.: Cambridge University Press, 1979, 165–179).

Nisbett, Richard E., and Timothy DeCamp Wilson. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* LXXXIV (3) (May 1977): 231–259.

Putnam, Hilary. "The Nature of Mental States." In *Mind, Language, and Reality: Philosophical Papers, Volume 2*. New York: Cambridge University Press, 1975, pp. 429–440. Originally published as "Psychological Predicates," in *Art, Mind, and Religion*, edited by W. H. Capitan and D. D. Merrill (Pittsburgh: University of Pittsburgh Press, 1967, pp. 37–48).

Rosenthal, David M. *Consciousness and Mind*. Oxford, U.K.: Clarendon Press, 2005.

Rosenthal, David M. "A Theory of Consciousness." Report 40/1990, Center for Interdisciplinary Research (ZiF), University of Bielefeld. Reprinted in *The Nature of Consciousness: Philosophical Debates*, edited by Ned Block, Owen Flanagan, and Güven Güzeldere (Cambridge, MA: MIT Press/Bradford Books, 1997, pp. 729–753).

Rosenthal, David M. "Two Concepts of Consciousness." *Philosophical Studies* 49 (3) (May 1986): 329–359. Reprinted in *Consciousness and Mind*, by David M. Rosenthal.

Seager, William. *Theories of Consciousness: An Introduction and Assessment*. New York: Routledge, 1999.

Searle, John R. "Consciousness, Explanatory Inversion, and Cognitive Science." *The Behavioral and Brain Sciences* 13 (4) (December 1990): 585–696.

Searle, John R. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press/Bradford Books, 1992.

Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." In *Science, Perception, and Reality*. New York: Humanities Press, 1963; Atascadero, CA: Ridgeview Publishing Company, 1991.

Shoemaker, Sydney. *The First-Person Perspective and Other Essays*. New York: Cambridge University Press, 1996.

Siewert, Charles P. *The Significance of Consciousness*. Princeton, NJ: Princeton University Press, 1998.

Simons, Daniel J., and Ronald A. Rensink, "Change Blindness: Past, Present, and Future." *Trends in Cognitive Sciences* 9 (1) (January 2005): 16–20.

Treisman, Anne. "Features and Objects in Visual Processing." *Scientific American* 255 (5) (November 1986): 114–125.

Weiskrantz, Lawrence. *Consciousness Lost and Found: A Neuropsychological Exploration*. New York: Oxford University Press, 1997.

David M. Rosenthal (2005)

CONSCIOUSNESS IN PHENOMENOLOGY

For Edmund Husserl, the two basic features of consciousness are intentionality and temporality. Intentionality means that all consciousness is directed to some object. The thesis that consciousness is temporal means not only that all conscious states have a temporal location but that each of them has within itself a temporal structure and that the temporal structure of consciousness is the basis for all other determinations of consciousness and its objects.

Husserl's philosophical method proceeds through an analysis of conscious life. However, because all consciousness is intentional, the analysis of the forms and structures of various kinds of consciousness (including volitional, emotional, and evaluative, as well as theoretical) is also the appropriate way to analyze the essential forms and structures of various kinds of objects. Because Husserl also believes that consciousness involves at least implicit self-consciousness of one's own mental states, the focus on consciousness shifts the analysis to a sphere that is immediately and directly given in reflection and is therefore the source of apodictic certainty, the transcendental ego. In later works Husserl qualifies this assertion by pointing out that self-givenness even for ideal objects never necessarily involves absolute certainty, so that all purported givenness requires reconfirmation. He also turns his attention to the sphere of passive synthesis, whose results may be directly given to us, while the operations that originally generate them are not, so that a phenomenological reconstruction or intentional analysis is necessary to reveal sedimented or initially hidden and prepredicative elements of consciousness.

Jean-Paul Sartre considered himself a philosopher of consciousness during the first half of his career. He subscribed to the Cartesian ideal of the cogito as the starting point of philosophy and placed a premium on the apodictic evidence it yielded. But he valued consciousness as much for its freedom and spontaneity as for its epistemological translucency. In fact, it was the relevance of translucency to moral responsibility that led him to deny both a transcendental ego and the Freudian unconscious and to posit a "prereflective *Cogito*."

In his *The Imaginary* Sartre describes imaging consciousness as the locus of "negativity, possibility, and lack." Because we are able to "hold the world at bay" and "derealize" perceptual objects imagistically, he argues, we are free. Imaging consciousness becomes paradigmatic of consciousness in general (being-for-itself) in *Being and Nothingness*. Adopting Husserl's thesis that all consciousness is intentional, he insists that this intentionality is primarily practical, articulating a fundamental project that gives meaning/direction (*sens*) to our existence.

Sartre makes much of the prereflective self-awareness that accompanies our explicit awareness of any object, including our egos as reflective objects. Because we are always implicitly self-aware, it is unnecessary to seek self-consciousness in an endless infinity of reflections on reflections or to chase after a subject that cannot be an object (the transcendental ego). The unblinking eye of prereflective consciousness makes possible both bad faith and its overcoming through what he calls "purifying reflection," the authentic "choice" to live at a creative distance from one's ego.

Husserl's students such as Aron Gurwitsch and Ludwig Landgrebe and most of the subsequent figures within the phenomenological tradition such as Martin Heidegger and Maurice Merleau-Ponty built upon Husserl's and Sartre's insights into the importance of self-awareness, intentionality, and temporality—often under other names—but they also stress the prepredicative and the practical nature of this awareness as well as its limitations. Hence they avoid the term "consciousness" for the most part because of its association with Cartesian aspirations to complete self-transparency and absolute autonomy in human knowledge and action that they reject.

See also Authenticity; Descartes, René; Time, Consciousness of.

Bibliography

Husserl, E. *Cartesianische Meditationen*. Husserliana 1, edited by S. Strasser. The Hauge: M. Nijhoff, 1950. Translated by D. Cairns as *Cartesian Meditations: An Introduction to Phenomenology*. The Hauge: M. Nijhoff, 1960. See especially secs. 6–22 regarding intentionality and reflection. See secs. 30–41 regarding the temporal nature of consciousness and passive genesis.

Husserl, E. *Formale und transzendentale Logik*. Husserliana 17 edited by P. Janssen. The Hauge: M. Nijhoff, 1974. Translated by D. Cairns as *Formal and Transcendental Logic*. The Hauge: M. Nijhoff, 1969. Sec. 58 distinguishes between self-givenness and infallibility.

Husserl, E. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. Husserliana 3. The Hauge: M. Nijhoff, 1976. Translated by F. Kersten as *Ideas Pertaining to a Pure Phenomenology and Phenomenological Philosophy. First Book*. The Hauge: M. Nijhoff, 1982. See especially secs. 34–62 regarding the intentionality of consciousness and its accessibility to pure reflection.

Nenon, Thomas. "Freedom, Responsibility, and Self-Awareness in Husserl." *New Yearbook for Phenomenology and Phenomenological Philosophy* 2 (2002): 1–21.

Sartre, J.-P. *L'etre et le néant*. Paris: Gallimard, 1943. Translated by H. E. Barnes as *Being and Nothingness*. New York: Philosophical Library, 1956.

Sartre, J.-P. *L'imaginaire*. Paris, 1940. Translated by Jonathan Webber as *The Imaginary*. London: Routledge, 2004.

Sartre, J.-P. *La transcendence de l'ego*. Paris, 1937. Translated by R. Kirkpatrick and F. Williams as *The Transcendence of the Ego: An Existentialist Theory of Consciousness*. New York: Noonday, 1957.

Thomas Flynn (2005)
Thomas Nenon (2005)

CONSENSUS GENTIUM

See *Common Consent Arguments for the Existence of God*

CONSEQUENTIALISM

As a name for any ethical theory or for the class of ethical theories that evaluate actions by the value of the consequences of the actions, "consequentialism" thus refers to classical utilitarianism and other theories that share this characteristic.

Classical utilitarianism, in the philosophies of Jeremy Bentham, John Stuart Mill, and Henry Sidgwick, was consequentialist, judging actions right in proportion as they tended to produce happiness, wrong as they tended to produce pain. In the nineteenth and early twentieth centuries much of the criticism of utilitarianism was directed at the hedonistic value theory on which the ethical theory was founded. Some philosophers, such as G. E. Moore, agreed with the claim of utilitarianism that acts are right insofar as they produce good consequences, wrong as they produce bad consequences, but put forward a richer theory of value, claiming that other things besides pleasure and pain are of intrinsic value and disvalue. Such theories were sometimes labeled ideal utilitarianism. The term *consequentialism* is now used in a generic sense to include both hedonistic and nonhedonistic theories.

The term was probably introduced into usage by Elizabeth Anscombe in "Modern Moral Philosophy" (1958), an essay in which she claims that there is little difference between strictly consequentialist theories and other moral theories from Sidgwick on that permit forbidden acts to be overridden by consequentialist considerations. For example, W. D. Ross, who was an intuitionist in opposition to utilitarianism, even "ideal" utilitarianism, believed that a prima facie wrong action, such as the deliberately unjust punishment of an innocent person, could be outweighed by some consequentialist consideration such as the national interest. One contrast with consequentialism, then, is absolutism, the claim that there are some actions that are never right, whatever the consequences.

In the most usual usage of *consequentialism* as a term for ethical theories, however, it is contrasted, not only to absolutism, but to any theory, such as Kantianism, intuitionism, virtue ethics, rights theories, and so on, that does not in some way make consequences the determinant of right and wrong. The consequences may be considered indirectly. Distinctions have been made between act utilitarianism, which judges acts right or wrong according to the consequences of the particular act, case by case, rule utilitarianism, which judges acts right or wrong according to whether the acts are in accord with or in violation of a useful rule—that is, a rule whose general practice would have good consequences (or better consequences than any feasible alternative rule)—and motive utilitarianism, which judges acts right or wrong if stemming from a motive that, as a motive for action, generally has good consequences. These distinctions carry over to consequentialism as a generic category of ethical theories, and one can speak of act consequentialism, rule consequentialism, and so on. Consequentialist theories can also have a place for virtues and for rights, if the inculcation of certain virtues or the respect for certain rights has good consequences. But for the consequentialist the virtues or rights are not ultimate. Their value is dependent upon their contribution to good consequences.

Abstracting from the alternative theories of value, there are still important controversies regarding consequentialist theories. Some are problems of measuring consequences or making interpersonal comparisons, whatever the theory of value, but these cannot be addressed in the abstract. Another is the theory of responsibility. One prominent criticism of consequentialism, stated, for example, by Bernard Williams (1973), is that it does not adequately distinguish between positive and negative responsibility. The claim is that consequentialism is indifferent between states of affairs that are produced by what an agent does and those that occur because of what someone else would do that the agent could prevent. It becomes an agent's responsibility to prevent someone else from doing harm as well as not to do harm oneself. Related to this is the claim that consequentialism undermines agent integrity. For example, someone opposed to research in chemical and biological

warfare might be required to engage in such research to prevent someone else from doing it more zealously. Another criticism is that if it is formulated as a "maximizing" theory, requiring the maximization of best consequences, consequentialism goes beyond the limits of obligation. For example, one would be morally obligated to spend one's wealth and income on others as long as there is anyone who could benefit more than oneself.

There are four basic kinds of responses that the consequentialist can make to these criticisms. One is to stick to the theory, saying that these things are morally demanded, even if not generally recognized in our selfish and self-centered society, as Peter Singer (1971–1972) argues concerning famine relief. Another is to challenge the implications of the examples, claiming that for moral agents to focus energy and attention on their own lives with integrity to their own principles has better consequences than doing otherwise. A third strategy for a nonhedonist is to attempt to avoid some of these objections by enriching the theory of value, such as to claim that integrity is something that is intrinsically valuable. A fourth strategy is to modify the structure of the theory. Michael Slote (1984) has argued in favor of a "satisficing" rather than a maximizing theory. Samuel Scheffler (1982) has proposed a "hybrid" theory that permits an agent either to maximize best consequences or to pursue the "agent-centered prerogative" of not always doing so.

See also Anscombe, Gertrude Elizabeth Margaret; Bentham, Jeremy; Deontological Ethics; Hedonism; Metaethics; Mill, John Stuart; Moore, George Edward; Rights; Ross, William David; Sidgwick, Henry; Utilitarianism; Virtue Ethics.

Bibliography

Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33 (1958): 1–19.

Bennett, J. "Whatever the Consequences." *Analysis* 26 (1965–1966): 83–102.

Brandt, R. B. "Utilitarianism and Moral Rights." *Canadian Journal of Philosophy* 14 (1984): 1–19.

Crisp, R. "Utilitarianism and the Life of Virtue." *Philosophical Quarterly* 42 (1992): 139–160.

Darwall, Stephen, ed. *Consequentialism*. Oxford: Blackwell, 2003.

Feldman, Fred. "On the Consistency of Act- and Motive-Utilitarianism: A Reply to Robert Adams." *Philosophical Studies* 70 (1993): 201–211.

Feldman Fred. *Pleasure and The Good Life*. Oxford: Oxford University Press, 2004.

Feldman, Fred. *Utilitarianism, Hedonism, and Desert*. New York: Cambridge University Press, 1997.

Hooker, B. *Ideal Code, Real World*. Oxford: Clarendon Press, 2000.

Hooker, B., E. Mason, and D. E. Miller. *Morality, Rules, and Consequences*. Edinburgh: Edinburgh University Press, 2000.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101 (1991): 461–482.

Kagan, S. "Does Consequentialism Demand Too Much? Recent Work on the Limits of Obligation." *Philosophy and Public Affairs* 13 (1983–1984): 239–254.

Mulgan, Tim. *The Demands of Consequentialism*. Oxford: Clarendon Press, 2001.

Norcross, Alasdair. "Consequentialism and Commitment." *The Pacific Philosophical Quarterly* 78 (1997): 380–403.

Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Pettit, Philip, ed. *Consequentialism*. Aldershot: Dartmouth, 1993.

Pettit, Philip. "The Consequentialist Perspective." In *Three Methods of Ethics*, by M. Baron, P. Pettit, and M. Slote, Oxford: Blackwell, 1997.

Railton, Peter. "Alienation, Consequentialism and the Demands of Morality." *Philosophy and Public Affairs* 13 (1984): 134–171, reprinted in Scheffler, ed. *Consequentialism and Its Critics*, 93–133.

Scarre, Geoffrey. *Utilitarianism*. London: Routledge, 1996.

Scheffler, Samuel, ed. *Consequentialism and Its Critics.* Oxford: Oxford University Press, 1988.

Scheffler, Samuel. *The Rejection of Consequentialism*. Oxford: Clarendon Press, 1982.

Shaw, William. *Contemporary Ethics: Taking Account of Utilitarianism*. Cambridge, MA: Blackwell, 1999.

Singer, P. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1971–1972): 229–243.

Slote, M. "Satisficing Consequentialism." *Aristotelian Society* suppl. 43 (1984): 139–163.

Smart, J. J. C. "An Outline of a System of Utilitarian Ethics." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and B. Williams. Cambridge, U.K.: Cambridge University Press, 1973.

Williams, B. "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and B. Williams. Cambridge, U.K.: Cambridge University Press, 1973.

Henry R. West (1996)

Bibliography updated by Michael J. Farmer (2005)

CONSERVATION PRINCIPLE

Conservation principles tell us that some quantity, quality, or aspect remains constant through change. Such principles already appear in ancient and medieval natural philosophy. In one important strand of Greek cosmology, the rotation of the celestial orbs is eternal and immutable. In optics at least from the time of Euclid (fl. c. 300 BCE), when a ray of light is reflected, the angle of reflection is

equal to the angle of incidence. According to some versions of the medieval impetus theory of motion, impetus permanently remains in a projected body (and the associated motion persists) unless the body is subject to outside interference. Such examples abound.

In the seventeenth century, conservation principles began to play a central role in scientific theories. Galileo Galilei, René Descartes, Christian Huygens, Gottfried Leibniz, and Isaac Newton founded their approaches to physics on the principle of inertia—the principle that a body will undergo uniform rectilinear motion unless interfered with. A multitude of other conservation principles gained currency during the seventeenth century—some still with us, some long ago left behind.

Descartes is an interesting example of an author who attempted to derive all of his physical principles from conservation laws (1991 [1644], esp. pt. 2, secs. 36–42). Descartes believed that the principles of his physics could be derived from the God's immutability, supplemented only by very weak assumptions about the existence of change in the world. He claimed, in fact, that we ought to postulate the strongest conservation laws consistent with such change. These laws were that God preserves at all times the total quantity of motion in the world (the quantity of motion of a body being the product of its volume and its speed), that each thing remains in the same state in every respect unless interfered with, and that in collisions the quantity of motion gained by one body is balanced by the quantity of motion lost by the other. The rest of his physics was supposed to follow from these principles alone.

The most remarkable of seventeenth-century analyses of conservation principles is contained in Huygens's essay on elastic collisions (1977). Huygens began by assuming that if two collinearly moving bodies of equal size move toward one another with equal speeds, in the resulting collision they simply exchange velocities. He then showed that it follows from the principle of Galileian relativity—that an experiment has the same outcome whether performed in a laboratory at rest or in a laboratory in uniform rectilinear motion—that whatever the initial velocity of such bodies, the result of a collision is always that velocities are simply exchanged. Huygens went on to analyze collisions between bodies of unequal size, again relying heavily on Galileian relativity. Among the consequences of his analysis were a number of conservation laws for systems of particles interacting only via elastic collisions: that the center of gravity of such a system undergoes uniform rectilinear motion, that the total kinetic energy of such a system is constant in time, and that the relative velocities (*mv*) of a pair of colliding particles is unchanged by a collision.

In a sense that will be spelled out below, the principle of Galileian relativity is a symmetry principle. So one of the things that Huygens accomplished was to show that from a symmetry principle one could deduce conservation principles. For an extensive class of physical theories—essentially, all of classical (or nonquantum) physics—it is now possible to establish a deep connection between symmetry principles and conservation laws. The balance of the discussion here provides an elementary introduction to the ideas relevant to understanding this connection.

SYMMETRY

At the most abstract level, a structure is a set of objects instantiating some set of properties and relations. A symmetry of a given structure is a permutation of the set of objects of the structure that leaves invariant all the properties and relations involved in the structure. For any structure, the identity map on its set of objects is, trivially, a symmetry.

For example, suppose that three points have relative pair-wise distances of three, four, and five units. Then there is no nontrivial symmetry that preserves these distances. But if the points instead form the vertices of an equilateral triangle, there will be several nontrivial symmetries, such as any transformation that interchanges two vertices while leaving the third fixed. We will be interested here in dynamic symmetries. As an intuitively plausible example, ordinary translation and rotation in space should be symmetries of any decent physical theory set in Euclidean space. Note that this does not mean that every translation or rotation will be a symmetry of the states allowed by the theory: The theory might treat the behavior of a finite number of point masses, in which case no configuration of the material points could be invariant under any nontrivial translation or under more than finitely many rotations. Rather, in such a case the invariance of the theory amounts to this: The dynamics of the theory is indifferent to the location or orientation of the system in Euclidean space, in that a translation or rotation of any state allowed by the theory will not change the dynamic evolution predicted by the theory, so long as the evolution of the new state is described relative to coordinate axes that have also been translated or rotated.

Its necessary to make all of this a bit more precise. Specifying a physical theory typically involves specifying a set of physically possible states and a dynamics defined on this space of states. Most often, the states involved will

be possible instantaneous states of the system, such as the instantaneous positions and momenta of a set of particles, or the values of some field and of its time derivative at each point of space. These states will be collected together to form a space with some interesting mathematical structure (there is no need to be very specific about this structure at this stage). For convenience, a strict form of determinism will be assumed, under which the dynamics is given by the rule that if the state of the system at a given time is *a*, then its state *t* units of time later will be *b*, which we write as $a \underset{t}{\rightarrow} b$. A symmetry of this dynamics, *S*, will be a one-to-one mapping from the state space onto itself that leaves invariant all of the structure defined on this space, including the arrow relation. So $a \underset{t}{\rightarrow} b$ if and only if $S(a) \underset{t}{\rightarrow} S(b)$.

THE HAMILTONIAN APPROACH

Remarkably, almost all the equations of motion that arise in classical physics can be derived within the mathematical framework of Hamiltonian mechanics.

Consider the Newtonian *n*-body problem (*n* point masses interacting according to Newton's law of gravitation). We construct the phase space for this problem, the space of dynamically possible states of the particles. Choosing a point in this 6*n*-dimensional abstract space amounts to specifying the position and momentum of each of the *n* particles (collision states with two or more particles coinciding in position are ruled out a priori, since the expression for the force of gravitational attraction between coincident particles is ill defined). Now, by the nature of the Newtonian equation of motion ($F = ma$), specifying the positions and momenta of the particles at some initial time suffices to determine their positions and momenta at other times (indeed, at all other times, unless a collision or other singularity occurs). So the dynamic content of the theory takes this form: Specifying a point in the phase space determines a curve in the phase space through that point—the idea being that if the given point represents the state of the system of *n* particles at time $t = 0$, then the curve tells us which points of the phase space represent states of the system at earlier and later times. These curves have the following nice feature: They partition the phase space, in the sense that exactly one curve passes through each point of the phase space (that at least one curve passes through each point follows from the dynamic content of the theory; that no more than one does so is a reflection, roughly speaking, of the determinism of this theory).

At the heart of the Hamiltonian approach lie three insights: (1) The phase space of the system, just in virtue of being a space of possible positions and momenta, carries a natural mathematical structure called a "symplectic form" (a closed nondegenerate two-form). (2) This structure allows the association to each nice real-valued function on the phase space of a family of curves that partitions the phase space. (3) The curves encoding the dynamics of the theory are thus associated with the Hamiltonian for the theory—the function that assigns to each point in phase space the total energy of the corresponding physical state (here the total energy is the sum of the kinetic energy and the gravitational potential energy).

These insights carry over to underwrite a Hamiltonian treatment of a vast assortment of classical (or nonquantum) physical theories. To develop a Hamiltonian treatment, consider the space of initial data for the equations of motion, and take this as the phase space of one's theory, showing that it comes equipped with a natural symplectic form (or generalization thereof) that allows one, in general, to pass from a function on the phase space to a set of curves partitioning the phase space—and in particular to pass from the Hamiltonian function assigning to each state its total energy to the curves on the phase space encoding the dynamic content of the equations of motion of the theory. This strategy works for rigid bodies, systems of moving particles subject to many sorts of constraints, many field theories, and some theories of material continua such as fluids and elastic bodies.

SYMMETRIES IN THE HAMILTONIAN APPROACH

So under the Hamiltonian approach, a theory consists of a phase space (representing the possible dynamic states of the theory) equipped with a symplectic form (or generalization) and a Hamiltonian function. Below, this symplectic form will be referred to as the geometrical structure of the phase space, although it is important to keep in mind that this structure is different in kind from the sort of metric structure that is normally treated in geometry.

A symmetry of a Hamiltonian theory is a one-to-one mapping from the phase space onto itself that preserves the geometric structure of the phase space and the Hamiltonian. Because these latter two objects are smooth, it follows that symmetries are continuous and differentiable to all orders.

In the *n*-body problem, for instance, all symmetries are smooth maps from the phase space onto itself that correspond to some combination of the following actions: (a) shifting by some fixed amount the positions

of all particles in the Euclidean space in which they move; (b) rotating the orientation of the system of particles in Euclidean space by some fixed amount; (c) shifting the temporal origin by some fixed amount (that is, associating each state with the state that normally precedes or follows it by the given amount of time); (d) applying related discrete symmetries, such as a mirror reflection of the positions of the particles or an interchange of negative and positive senses of time.

Because each symmetry of a Hamiltonian theory leaves invariant all of the structure on the phase space that was used to define the dynamics, it also leaves invariant the curves that encode the dynamics—as we should expect from our general account of dynamic symmetries above. (The operation of a Galileian boost is not a symmetry in the present sense. Boosting a system does not leave its Hamiltonian invariant. A boost changes the kinetic energy of each particle and in general alters the total kinetic energy of the system, while leaving the potential energy of the system unchanged. But Galileian boosts do leave invariant the set of dynamic curves of the *n*-body problem.)

NOETHER'S THEOREM

A remarkable consequence of the geometric structure of the phase space is that in any Hamiltonian system, the Hamiltonian is constant along each curve encoding the dynamics. That is, the total energy of the system is a conserved quantity of the dynamics: If one state evolves into another, each has the same total energy.

Are there additional conserved quantities—functions on our phase space that are constant along the curves encoding the dynamics? To find them, consider any one-parameter continuous family of symmetries of a Hamiltonian system closed under composition—such as the family of spatial translations by a varying amount in a given direction in the *n*-body problem. What happens if we allow such a family to act on a point in the phase space of a Hamiltonian system? To find out, we construct a curve in the phase space that describes how each symmetry in our one-parameter family acts on the initial state. By performing this operation for each point in the phase space, we construct partitions the phase space. For the sort of theories that arise in practice, we can then find a nice function on our phase space that the geometric structure associates with this family of curves. From the geometric structure and from the fact that the family of curves in question arises via the action of a family of symmetries that preserve the Hamiltonian, it follows that this function will itself be a constant of motion of the physical theory under consideration—that this new function, associated with our one-parameter family of symmetries, is constant along each of the dynamic curves associated with the Hamiltonian of our theory.

In this way, for any well-behaved Hamiltonian theory, we can construct a conserved quantity corresponding to each one-parameter family of symmetries of the theory. Indeed, we can find as many functionally independent conserved quantities as there are dimensions in the complete family of symmetries of the theory. In the case of the *n*-body problem, we find seven conserved quantities: the Hamiltonian, the components of the total linear momentum of the system, and the components of the total angular momentum of the system. These conserved quantities correspond, respectively, to the invariance of the system under time translations, to its invariance under spatial translations in each direction, and to its invariance under rotation.

These insights derive from work of Emmy Noether in 1918 (1971), though they assume a somewhat different form in her work, since she worked in the Lagrangian framework rather than the Hamiltonian framework.

See also Classical Mechanics, Philosophy of; Philosophy of Physics.

Bibliography

Brading, Katherine, and Elena Castellani, eds. *Symmetries in Physics: Philosophical Reflections*. Cambridge, U.K.: Cambridge University Press, 2003.

Descartes, René. *Principles of Philosophy* (1644). Translated by Valentine Miller and Reese Miller. Dordrecht, Netherlands: Kluwer, 1991.

Huygens, Christian. "The Motion of Colliding Bodies." Translated by Richard J. Blackwell. *Isis* 68 (1977): 574–597.

Noether, Emmy. "Invariant Variation Problems" (1918). Translated by M. Tavel. *Transport Theory and Statistical Physics* 1 (1971): 184–207.

Singer, Stephanie Frank. *Symmetry in Mechanics: A Gentle, Modern Introduction*. Basel, Switzerland: Birkhäuser, 2001.

Van Fraassen, Bas. *Laws and Symmetry*. Oxford: Oxford University Press, 1989.

Zuckerman, Gregg. "Action Principles and Global Geometry." In *Mathematical Aspects of String Theory*, edited by Shing-Tung Yau. Singapore: World Scientific, 1987.

Gordon Belot (2005)

CONSERVATISM

Conservatives hold that the aim of political arrangements is to make a society good, that a society is good if those

living in it can make good lives for themselves, and that good lives are satisfying for oneself and beneficial for others. The political arrangements that make such lives possible are discovered by historical reflection, which discloses various enduring traditions. People participate in them because they conceive of good lives in terms of the beliefs, values, and practices these traditions provide. Conservatism is not an uncritical defense of all traditions but only ones that have endured because people have found participation in them satisfying and beneficial. The justification or criticism of traditions, therefore, is based on their success or failure in fostering good lives. Conservatism has different versions depending on the views held about history, values, the relation between individuals and their society, and between human nature and evil. Conservatives agree that these are the pivotal political topics and that political arrangements should be based on the views held about them, but they disagree about what these views should be.

HISTORY

Conservatives believe that the starting points of political thought should be the prevailing political arrangements, rather than a hypothetical contract, a theory about an ideal society, a conception of the common good for all of humanity, or some basic value that always overrides any other value that may conflict with it. Some conservatives, however, believe that it is not a coincidence that certain political arrangements have been historically conducive or detrimental to good lives. They hold that there is a metaphysical explanation of their success or failure: the existence of a moral order in reality. Lives are good if they conform to this moral order and bad if they do not, and the same is true of political arrangements. These conservatives recognize that there are serious disputes about whether the order is (a) hierarchical—having The Good at its pinnacle, as supposed by the ancient Greek philosopher Plato, (b) providential, embodied in natural law, as held in the thirteenth century by Saint Thomas Aquinas and his many followers since then, or (c) an unfolding of the dialectic of clashing forces culminating in the final unity of reason and action, as claimed by Georg W. F. Hegel in nineteenth-century Germany. They nevertheless all assume that there is a moral order. Their task is to find out what it is, or, if it has already been revealed in some canonical text, to find out how it should be interpreted. Disputes about these matters are taken to show only the infirmity of human understanding or motivation, not that the existence of the moral order is doubtful.

The historical record of societies whose political arrangements have been based on a supposed moral order, however, is most alarming. They have tended to indoctrinate unwilling or uninformed people, leaving them no opportunity for choice. Such societies have not been freer of misery than less dogmatic ones. But they have added the misery peculiar to themselves of recognizing authorities who have claimed privileged access to the true and the good and thought that only human shortcomings stand in the way of good lives. This was taken by them to justify coercion, silencing dissenters, and indoctrinating the rest. Many conservatives, such as Edmund Burke in the eighteenth century and Michael Oakeshott in the twentieth, have rejected this approach to politics because of its grave dangers.

Other conservatives, such as the Frenchman Joseph de Maistre (1753–1821), deny that the right political arrangements can be justified by reason. It makes no difference to them whether the proffered reasons are metaphysical, scientific, or historical. They believe that all reasons are ultimately based on assumptions accepted on faith. Their rejection of reason as a guide, however, leaves them with the problem of deciding what political arrangements they ought to favor. The solution they have offered is to be guided by their faith and perpetuate existing arrangements because familiarity makes them safer than untried alternatives.

The problems of this approach are as evident from the historical record as those of the preceding one. Faith breeds dogmatism, persecution of those who hold other faiths or none at all, and it provides no justification for regarding political arrangements based on one faith as better than contrary ones based on other faiths. Moreover, the perpetuation of political arrangements on account of their familiarity makes their improvement virtually impossible because even flawed familiar ones will be judged preferable to dangerous unfamiliar possibilities.

An alternative to relying on a moral order or on faith is the fallibilism of the Frenchman Michel de Montaigne in the sixteenth century, the Scottish David Hume in the eighteenth century, and, closer to our times, George Santayana (1863–1952) and Oakeshott. They do not deny that there may be a moral order, only that reliable knowledge of it can be had. They find the historical evidence testifying to humans' fallibility much more convincing than the success of efforts to overcome it. They think that the claims of revelation, canonical texts, or putative eternal verities stand in need of persuasive evidence. But the available evidence is as questionable as the claims based

on it. Fallibilists believe that it is far more prudent to look to the historical record of various political arrangements than to endeavor to justify or criticize them by relying on metaphysical speculations or faith, either of which is more dubious than the historical record.

Fallibilism, however, does not commit conservatives to the denial that it is possible to adduce reasons for or against political arrangements. What they deny is that good reasons must be universal and timeless. They reject the fideistic repudiation of reason, accept the importance and desirability of being as reasonable as possible, and claim that political arrangements should be based on the historical evidence available for them. Fallibilists want political arrangements to be firmly rooted in the experiences of the people who are subject to them. Since these experiences are inevitably historical, it is to their history that these conservatives look for evidence. Thus they avoid basing political arrangements on metaphysical speculation about what lies beyond experience and suspecting reasonable evaluations because of a global distrust of reason.

It seems, then, that the most defensible conservative belief about history is the fallibilist one. There is a presumption in favor of the enduring political arrangements of a society because their endurance is prima facie reason for supposing that they foster satisfying and beneficial lives. In the absence of contrary reasons, the enduring political arrangements ought to be maintained. It is possible, of course, that the arrangements have endured because of coercion or manipulation. If the case for change is based on cogent evidence that the arrangements have endured for reasons other than fostering good lives, then it should be seriously considered. But if the case is inspired merely by metaphysical, contractarian, fideistic, or utopian speculations, then much stronger reasons are needed before they could mount a reasonable challenge to political arrangements that have stood the test of time and attracted the allegiance of many people.

VALUES

Commitment to political arrangements that make good lives possible requires a view about what makes lives good. But there are countless valued activities, obligations, virtues, and satisfactions, countless ways of combining them and evaluating their respective importance, and so there seem to be countless ways in which lives can be good. Conservatives, therefore, must have a view about the diversity of values because the arguments for or against particular political arrangements depend on it. The problem is that there are three incompatible views about the diversity of values: absolutist, relativist, and pluralist.

Absolutists believe that the diversity of values is apparent, not real. They concede that there are many values, but they think that there is a universal standard that can be used to rank them. This standard may be a highest value, such as Plato's Good; the tranquility of ancient Greek and Roman Stoics; the love of God postulated by Saint Augustine, the fifth-century Bishop of Hippo; or the idea of general happiness advanced by nineteenth-century English Utilitarians. Other values, then, can be ranked on the basis of their contribution to the realization of the highest value. Or the standard may be a principle, like the Ten Commandments of the Old Testament, the Golden Rule of the New Testament, or the categorical imperative formulated by the German philosopher Immanuel Kant in the eighteenth century. If a choice needs to be made among different values, the principle will determine which ought to take precedence. Contemporary absolutists—for instance, the English John Finnis, the American Germain Grisez, and the German-American Eric Voegelin—argue that some political arrangements are preferable to others because they conform more closely to the universal standard than the alternatives. However, the candidates for a universal standard are also numerous and face the same problems as the values whose diversity is supposed to be diminished by them. Absolutists acknowledge this problem and explain it in terms of human shortcomings that prevent people from recognizing the true standard. The history of religious wars, persecutions, and tyrannies, aiming to rectify human shortcomings, testifies to the dangers inherent in this explanation.

Relativism is opposed to absolutism. Relativists regard the diversity of values as real: There are many values; many ways of combining and ranking them; and there is no universal standard that could be appealed to in resolving disputes about them. A good society, however, requires some consensus about values, and political arrangements should reflect this consensus. If the consensus changes, the arrangements should change as well. According to relativists, then, what counts as a value and how important it is depends on the consensus of a society. A value is what is valued in a particular context; all values, therefore, are context-dependent.

Values and the political arrangements that reflect them can be reasonably justified or criticized, but, relativists believe, the reasons that can be given for or against them count as reasons only within the context of a society. Since reasons ultimately rest on the prevailing con-

sensus, they will not and are not meant to persuade outsiders. The ultimate appeal of relativists is to point at their arrangements and say: This is what we do here. If relativism takes a conservative form, as in Burke, or in the nineteenth-century Germans Johann Gottfried Herder and Wilhelm Dilthey, it often results in the romantic celebration of national identity, of the spirit of a people and an age, of the shared landscape, historical milestones, ceremonies, stylistic conventions, manners, and rituals that unite a society.

Relativism may seem to avoid the dangers of dogmatism and repression that threaten absolutism, but it is, in fact, equally prone to them. That the values of a society are not thought to be binding outside of it does not mean that the values of other societies will be regarded tolerantly. Because the world is full of people and societies whose values are inimical to the relativist's, there is good reason to guard jealously the relativist's values. Furthermore, if the justification of the values of a society is the prevailing consensus, then any political arrangement becomes justifiable, provided a sufficiently large consensus favors it. Slavery, female circumcision, the maltreatment of minorities, child prostitution, the mutilation of criminals, blood feuds, bribery, and any other noxious political arrangement may be exempted from censure on the grounds that it happens to be valued in its context.

These pitfalls of absolutism and relativism make them unreliable guides to the evaluation of political arrangements. It is with relief, then, that some conservatives in the last and present century—for instance, Oakeshott, Gordon Graham, and John Kekes—have turned to pluralism as an alternative to these flawed views. Pluralists are in partial agreement and disagreement with both absolutists and relativists. According to pluralists, there is a universal standard, but it applies only to values that must be protected by all political arrangements if they are to foster good lives. This standard is universal and minimal. It is possible to establish with reference to it some values required by all good lives, but not all the values that may make lives good. This leads to recognizing some political arrangements as necessary and to allowing a generous plurality of possible political arrangements beyond the necessary minimum. The standard accommodates part of the universalism of absolutism and part of the contextualism of relativism.

The source of pluralism's standard is human nature. To understand human nature sufficiently for the purposes of this standard does not require scientific research or commitment to some metaphysical belief or to natural law. It is enough for it to concentrate on normal people in a commonsensical way. It will then become obvious that good lives depend on the satisfaction of basic physiological, psychological, and social needs such as nutrition, shelter, rest, companionship, self-respect, the hope for a good or better life, the division of labor, justice, predictability in human affairs, and so forth. The satisfaction of these basic needs is a universal requirement of all good lives, whatever may be their social context. If the political arrangements of a society foster their satisfaction, it is reasonable to support them; if they hinder their satisfaction, it is reasonable to reform them.

If absolutists merely asserted and relativists merely denied this requirement, then the former would be right and the latter wrong. But both go beyond this point: absolutists hold that all the values that make lives good are to be evaluated by a universal standard and relativists deny that there is any universal standard at all. Pluralists think that the minimum requirements of human nature set a universal standard, but beyond it there is a plurality of values, of ways of ranking them, and of good lives that embody these values and rankings. According to pluralists, then, the political arrangements of a society ought to protect minimum requirements of good lives and foster a plurality of values beyond the minimum.

The combination of pluralism and conservatism provides two important possibilities. The first is the justification of political arrangements that protect the minimum requirements and the criticism of political arrangements that violate them. This possibility sets the goal of political action and makes possible reasonable comparisons among different societies on the basis of how well they protect the conditions on which all good lives depend. This version of conservatism avoids the objection to relativism that it sanctions any political arrangement so long as a large consensus supports it. The second is that the best guide to the political arrangements a society ought to have beyond the minimum is the history of the society because it is most likely to provide the evidence for or against the prevailing political arrangements. This second possibility avoids the dangers of dogmatism and repression that beset absolutism.

The political arrangements favored by this version of conservatism are based on a familiar list of values: justice, freedom, the rule of law, order, legal and political equality, prosperity, peace, civility, happiness, and so forth. There is likely to be a significant overlap among the lists conservatives, liberals, and socialists may draw up. But there will also be a significant difference: conservatives are genuine pluralists, whereas the liberal and socialists are not. Liberals and socialists are committed to regarding

some of these values as more important than the others. What makes them liberals or socialists is their claim that when the values they favor conflict with others on the list, then the favored ones should prevail. If they did not claim this, they would cease to be liberals or socialists. Conservatives reject this approach. Their concern is to protect the whole system of these values. This sometimes requires favoring a particular value over another, sometimes the reverse. Conservative pluralists hold this to be true of all values. They differ from liberals and socialists in refusing to make an a priori commitment to resolving conflicts in favor of any particular value in the prevailing system of values.

INDIVIDUALS AND SOCIETY

Good lives must be satisfying and beneficial, but these requirements often conflict because satisfying lives may not be beneficial and beneficial lives may not be satisfying. This raises the question of which requirement should be dominant, and it has far-reaching political consequences how it is answered. Some twentieth-century conservatives—for instance, Friedrich von Hayek, Shirley Letwin, and Robert Nozick—favor individual autonomy over social authority. Their position is virtually indistinguishable from classical liberalism or libertarianism. Other conservatives—such as James Fitzjames Stephen in nineteenth-century England, Voegelin, and the twentieth-century English thinker Roger Scruton—think that social authority should prevail over individual autonomy, if they conflict. As before, there is an intermediate view between these two extremes, namely, that of twentieth-century traditionalist conservatives, represented, among others, by Oakeshott, the American Edward Shils, and Kekes.

Conservatives who stress autonomy at the expense of authority face two serious problems. First, they assume that good lives must be autonomous and cannot involve the recognition of some form of authority over oneself. If this were so, no military or devoutly religious life, no life in static, traditional, hierarchical societies, no life that involves the subordination of individual will and judgment to what is regarded as a higher purpose could be good. This would require regarding of the vast majority of lives outside of modern Western societies as bad. The mistake involved is to slide from the reasonable view that autonomous lives may be good to the unreasonable view that a life cannot be good unless it is autonomous. Second, if a good society is one that fosters good lives, then the precedence of autonomy over authority cannot be right, since autonomous lives may be frustrated or harmful. It is obvious that social authority must prevail over the autonomy of fanatics, criminals, fools, and so on.

The problems of authoritarianism are no less serious. There is no guarantee that if social authority prevails over individual autonomy, the resulting lives will be satisfying. Lives cannot be pronounced satisfying by some authority. Whether they actually are satisfying must ultimately be judged by the individuals themselves. Their judgments, of course, may be influenced by social authority. But no matter how strong that influence is, it cannot override the judgment of individuals in finding what satisfies them. As the lamentable historical record shows, however, this has not prevented countless religious and ideological authorities from stigmatizing individuals who reject their prescriptions as heretics, pagans, maladjusted, or sinful. The result is a repressive society whose dogmatism is reinforced by specious moralizing.

How, then, is the question to be answered? There is no need to insist that either individual autonomy or social authority should systematically prevail over the other. Both are necessary for good lives, but neither is sufficient. Instead of engaging in futile arguments about their comparative importance, it is far more illuminating to understand that they are interdependent aspects of the same underlying activity of individuals trying to make good lives for themselves. The connecting link between them is tradition.

A tradition is a set of customary beliefs and practices that have endured from the past to the present and attracted the allegiance of people so that they wish to perpetuate it. Traditions may be religious, horticultural, scientific, political, stylistic, moral, aesthetic, commercial, medical, legal, military, educational, architectural, and so on. They permeate human lives. When individuals aim at a good life, part of what they are doing is deciding which traditions they should participate in. They may make conscious, deliberate, clear-cut yes-or-no choices, or they may unconsciously, unreflectively fall in with familiar patterns, or they may be at various points in between. The bulk of the activities of individuals concerned with living in ways that strike them as good is composed of participation in the various traditions of their society.

Participation involves the exercise of autonomy. Individuals choose and judge; their wills are engaged; they learn from the past and plan for the future. But they do so in the frameworks of various traditions that authoritatively provide them with a range of choices, with matters that are left to their judgments, and with standards that within a tradition determine what choices and judgments are reasonable or unreasonable. Their exercise

of autonomy is the individual aspect of their conformity to their tradition's authority, which is the social aspect of what they doing. They act autonomously by following the authoritative patterns of the traditions to which they feel allegiance. When a Catholic goes to confession, a violinist gives a concert, or a football player scores a touchdown, then the individual and the social, the autonomous and the authoritative, the traditional pattern of doing it and an individual doing it are inextricably mixed. To understand what is going on in terms of individual autonomy is as one-sided as it is to do so in terms of social authority. Both play an essential role, and understanding what is going on requires understanding both the roles they play. Traditionalism rests on this understanding, and it is the political response to it to maintain political arrangements that foster participation in the various traditions that have endured in a society.

Traditions may be vicious, destructive, stultifying, nay-saying, and thus detrimental to good lives. Part of the purpose of political arrangements is to draw distinctions among traditions that are unacceptable (like slavery), suspect but tolerable (like pornography), and worthy of encouragement (like university education). Traditions that violate minimum requirements of good lives should be prohibited. Traditions that have shown themselves to make questionable contributions to good lives should be tolerated but not encouraged. Traditions whose record testifies to their importance for good lives should be cherished.

The obvious question is who should decide which tradition is which and how that decision should be made. Traditionalist conservatives say that the decision should be made by those who are legitimately empowered to do so through the political arrangements of their society and they should make the decisions by reflecting on the historical record of the tradition in question. From this three corollaries follow. First, those who are empowered to make the decisions ought to be able to view the prevailing political arrangements from a historical perspective. The process works well if it empowers people who are not ill-educated, preoccupied with some single issue, inexperienced, or have qualifications that lie in some other field of endeavor. Traditionalist conservatives are clearly not populists. Second, a society that proceeds in this manner is pluralistic because it fosters a plurality of traditions. It does so because it sees as the justification of its political arrangements that they foster good lives, and fostering them depends on fostering the traditions participation in which may make lives good. Third, the society is tolerant because it is committed to having as many traditions as possible. Its political arrangements place the burden of proof on those who wish to proscribe a tradition. If a tradition has endured, if it has the allegiance of enough people to perpetuate it, then there is a prima facie case for it. That case may be, and often is, defeated, but the initial presumption is in its favor.

This outlook leads traditionalists to favor limited government. They do not think that the purpose of its political arrangements is to impose a conception of a good life. The political arrangements of a limited government interfere as little as possible with the traditions that flourish among people subject to it. The purpose of its arrangements is to enable people to live as they please, not to force them to live in a particular way. One of the most important ways of accomplishing this is to have a wide plurality of traditions as a bulwark between individuals and the government that has power over them. Traditionalist conservatives thus believe that a good society should have political arrangements that balance the claims of individual autonomy and social authority. This balance is maintained by the mediation of the traditions of a society that make autonomy possible and provide many of the forms it might take.

HUMAN NATURE AND EVIL

Conservatives tend to take a dim view of progress. They are not so foolish as to deny that great advances in pure and applied science have changed human lives for the better. But they have also changed them for the worse. Advances have been both beneficial and harmful. They have certainly enlarged the stock of human possibilities, but the possibilities are for both good and evil, and new possibilities are seldom without new evils. Evil is an obstacle to the betterment of the human condition. Unjust war, genocide, tyranny, torture, terrorism, the drug trade, concentration camps, racism, the murder of religious and political opponents, easily avoidable epidemics and starvation are some familiar and widespread evils. If evil is understood as serious unjustified harm caused by human beings, then the conservative view is that the prevalence of evil is a permanent condition that cannot be significantly altered.

The prevalence of evil reflects not just a human propensity but also a contingency that influences what propensities people have and develop. The propensity for evil is itself a manifestation of deeper, more pervasive influences, which operate through genetic inheritance, environmental factors, the crimes, accidents, pieces of good or bad fortune, the historical period, society, family, and so on. The same contingency also affects people

because others, whom they love, depend on, and with whom their lives are intertwined in other ways, are as subject to it as they are themselves.

Pessimistic conservatives, such as Thomas Hobbes in seventeenth-century England and Niccolo Machiavelli in sixteenth-century Florence, think that the prevalence of evil reveals that human nature is basically evil. Optimistic conservatives, such as Hume and Oakeshott, reject pessimism because they think that the right sort of political arrangements will make evil less prevalent. Opposed to both are realistic conservatives—for example, Montaigne, Stephen, and Santayana—who hold that whether the balance of good and evil propensities and their realization by people tilts one way or another is a contingent matter over which individuals and their political arrangements have insufficient control.

Realistic conservatives do not think that the human condition is devoid of hope, but they have no illusions about the limited control a society has over its future. Their view is not that evil propensities are uncontrollable. Rather, human beings have both good and evil propensities and neither they nor societies can exercise sufficient control to make good propensities reliably prevail over evil ones. The right political arrangements help of course, just as the wrong ones make matters worse. But even under the best political arrangements a great deal of contingency remains and it places beyond human control much good and evil. The chief reason for this is that the efforts to control contingency are themselves subject to the very contingency they aim to control. And that, of course, is the fundamental reason why realistic conservatives doubt the possibility of significant improvement of the human condition.

Realistic conservatives do not believe that it is a matter of indifference what political arrangements are made. Political arrangements cannot guarantee the victory of good over evil, but they can influence how things go. Whether that is sufficient at a certain time and place is itself a contingent matter insufficiently within human control. The attitude that results from realizing this combines the acceptance of the fact that not even the best political arrangements guarantee good lives with the motivation to make political arrangements as good as possible.

This view accounts for another significant difference between conservative and liberal or socialist politics: the insistence of conservatives on the importance of political arrangements that hinder evil. This difference is a result of the conservative rejection of the optimistic belief, shared by liberals and socialists, that the prevalence of evil is the result merely of bad political arrangements, which tend to corrupt people, and if political arrangements were good, evil would be much less prevalent. Realistic conservatives reject this optimism. They do not think that evil is prevalent merely because of bad political arrangements. They ask why political arrangements are bad. And the answer must be that political arrangements are made by people, and they are bound to reflect the propensities of their makers. Bad political arrangements are ultimately traceable to evil human propensities. Since the propensities are subject to contingencies over which human control is insufficient, there is no guarantee that political arrangements can be made good. Nor that, if they were made good, they would be sufficient to hinder evil.

Conservatives insist, therefore, on the necessity and importance of political arrangements that hinder evil: moral education, the enforcement of morality, the treatment of people according to what they deserve, the importance of swift and severe punishment for serious crimes, and so on. They oppose the prevailing attitudes that lead to agonizing over the criminal and forgetting the crime, the absurd fiction of a fundamental moral equality between habitual evildoers and their victims, guaranteeing the same freedom and welfare-rights to good and evil people, and so forth. Conservatives think that the aim of justice is to uphold the rule of law that assures that people get what they deserve.

Political arrangements that are meant to hinder evil are liable to abuse. Conservatives know and care about the historical record that testifies to the dreadful things that have been done on the many occasions when such arrangements have gone wrong. The remedy, however, cannot be to refuse to make the arrangements; it must be to learn from history, and try hard to avoid their abuse. Conservatives know that in this respect, as in all others, contingency will be a permanent obstacle to success. But this is precisely the reason why political arrangements are necessary for hindering evil. Realistic conservatives face the worst and try to deny scope to it, rather than endeavor to build a Utopia on optimistic illusions.

OVERVIEW

The most reasonable version of conservatism is fallibilist, pluralist, traditionalist, and realist. It avoids metaphysical and fideistic dogmatism. It denies that the content of good lives is given by a system of absolute values; accepts that good lives have some universal albeit minimal, requirements; and holds that some, but not all, values are context-dependent. It recognizes that both individual

autonomy and social authority are necessary for good lives, and resolves their conflicts by balancing their claims. It rejects both optimism based on utopian illusions and pessimism that registers only human corruption. It sees human nature as having both good and evil propensities and strives for political arrangements that foster the first and curb the second. Conservatism is a view of politics guided by history and aiming at the betterment of society within the limits set by the contingency of life and human imperfection.

See also Augustine, St.; Burke, Edmund; Dilthey, Wilhelm; Evil; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hobbes, Thomas; Human Nature; Hume, David; Kant, Immanuel; Liberalism; Machiavelli, Niccolò; Maistre, Comte Joseph de; Montaigne, Michel Eyquem de; Nozick, Robert; Oakeshott, Michael; Plato; Pluralism; Santayana, George; Social and Political Philosophy; Thomas Aquinas, St.; Value and Valuation.

Bibliography

Aquinas, Saint Thomas. *Summa Theologiae*. London: Eyre and Spottiswoode, 1963–1975.

Burke, Edmund. *Reflections on the Revolution in France*. Harmondsworth: Penguin, 1968.

Finnis, John. *Natural Law and Natural Rights*. Oxford: Clarendon Press, 1980.

Graham, Gordon. "True Conservatism—the Rejection of Ideology." In *Politics in Its Place*. Oxford: Clarendon Press, 1986.

Hayek, Friedrich von. *The Constitution of Liberty*. Chicago: Regnery, 1960.

Hegel, Georg W. F. *The Philosophy of Right*. Translated by T.M. Knox. Oxford: Oxford University Press, 1952.

Hobbes, Thomas. *Leviathan*. London: Dent, 1962.

Hume, David. *An Enquiry concerning the Principles of Morals*. Oxford: Clarendon Press, 1961.

Hume, David. *Essays Moral, Political, and Literary*. Indianapolis: Liberty Press, 1985.

Kekes, John. *A Case for Conservatism*. Ithaca: Cornell University Press, 1998.

Kirk, Russell. *The Conservative Mind*. South Bend, IN: Gateway, 1978.

Letwin, Shirley. *The Pursuit of Certainty*. Cambridge, U.K.: Cambridge University Press, 1965.

Letwin, Shirley. *The Gentleman in Trollope*. Cambridge, MA: Harvard University Press, 1982.

Machiavelli, Niccolo. *The Prince*. In *Selected Political Writings*. Translated and edited by David Wootten. Indianapolis: Hackett, 1994.

Montaigne, Michel de. *Essays*. In *The Complete Works of Montaigne*. Translated by Donald M. Frame. Stanford: Stanford University Press, 1958.

Muller, Jerry Z. ed. *Conservatism*. Princeton, NJ: Princeton University Press, 1997.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Oakeshott, Michael. *The Politics of Faith and the Politics of Scepticism*, edited by Timothy Fuller. New Haven: Yale University Press, 1996.

Oakeshott, Michael. *Rationalism in Politics*, edited by Timothy Fuller. Indianapolis: Liberty Press, 1991.

O'Sullivan, Noel. *Conservatism*. New York: St. Martin's Press, 1976.

Plato. *The Republic*. Translated by Robin Waterfield. Oxford: Oxford University Press, 1993.

Santayana, George. *Dominations and Powers*. New York: Scribner's, 1951.

Scruton, Roger. *The Meaning of Conservatism*. Harmondsworth: Penguin, 1980.

Shils, Edward. *Tradition*. Chicago: University of Chicago Press, 1981.

Shils, Edward. *The Virtue of Civility*, edited by Steven Grosby. Indianapolis: Liberty Press, 1997.

Stephen, James Fitzjames. *Liberty, Equality, Fraternity*. Cambridge, U.K.: Cambridge University Press, 1967.

Voegelin, Eric. *Order in History*. Baton Rouge: Louisiana State University Press, 1954–1987.

John Kekes (2005)

CONSTRUCTIVISM, MORAL

Moral constructivism is a metaethical view about the nature of moral truth and moral facts (and properties), so called because the intuitive idea behind the view is that such truths and facts are human constructs rather than objects of discovery. More precisely, constructivism involves both a semantic thesis about moral sentences and a two-part metaphysical thesis about the existence and nature of moral facts and properties. According to the semantic thesis, ordinary moral sentences purport to be fact-stating sentences and thus purport to be genuinely true or false. And, according to the metaphysical thesis, there are moral facts whose existence and nature are in some sense dependent upon human attitudes, agreements, conventions, and the like. Thus, constructivism represents a metaethical view in partial agreement with versions of moral realism. Like the realist, the constructivist is a so-called cognitivist (descriptivist)—moral sentences have descriptive content and thus purport to be genuinely fact stating. Again, like the realist, the constructivist accepts the view that there are moral facts that serve as the truth makers of true moral sentences. But unlike the realist, the constructivist rejects the idea that there are moral facts (and properties) that are independent of human attitudes, conventions, and the like.

It is useful to distinguish between simple and sophisticated versions of constructivism as well as between nonrelativist and relativist versions. Simple versions of constructivism are represented by certain views that would construe moral truth in terms of the actual attitudes of individuals or actual agreements within cultures about matters of moral concern. More sophisticated versions of constructivism construe moral truths (and associated moral facts and properties) in terms of the hypothetical attitudes of individuals or perhaps hypothetical agreements among members of a group reached under suitably constrained circumstances. Nonrelativist versions of constructivism maintain that all individuals and groups whose attitudes, agreements, and so forth provide the basis for moral truths and facts do or would accept the same set of basic moral norms with the result that there is a single set of moral truths and facts. Usually, such views are wedded to some version or other of sophisticated constructivism.

Thus, a version of the ideal-observer view of moral truth—according to which basic moral truths are represented by the moral norms that would be accepted by an ideal observer, where the notion of an ideal observer is so characterized that all ideal observers will agree on the same set of basic moral norms—is a version of sophisticated nonrelativist constructivism. Relativist versions of constructivism allow that there may be more than one individual or group with differing attitudes and agreements that serve as the basis for different (and conflicting) sets of basic moral norms. Versions of moral relativism, according to which moral truths and facts are a matter of what basic moral norms a culture in fact accepts, represent versions of simple, relativistic constructivism; versions of relativism, according to which moral truths and facts are a matter of what would be accepted under conditions that are ideal for choosing such norms, represent sophisticated relativistic versions of constructivism. Versions of the ideal-observer view are relativistic if they allow that there can be ideal observers who would accept different (and conflicting) sets of moral norms. So-called Kantian constructivism of the sort elaborated and defended by John Rawls, which appeals to choices made by hypothetical individuals behind a veil of ignorance (a version of contractarianism), is yet another sophisticated and apparently nonrelativistic constructivist view.

Constructivism, at least in its sophisticated versions, is supposed to capture what is plausible about moral realism, leaving behind what is problematic about realist views. Thus, constructivism can accommodate quite well certain "objective pretensions" of commonsense moral thinking. Some of these pretensions have to do with the form and content of moral discourse. A good many moral sentences are in the declarative mood (e.g., "Abortion, except in cases of rape and incest, is wrong") and are thus naturally interpreted as genuinely fact-stating sentences. Moreover, some such sentences appear to make references to (putative) moral facts and properties (e.g., "The evil of American slavery was partly responsible for its demise as an institution"). Other objective pretensions have to do with such activities as moral deliberation, debate, and argument. These critical practices are seemingly aimed at arriving nonarbitrarily at true or correct moral views, ones that would ideally resolve intrapersonal and interpersonal conflict and uncertainty about moral issues. Like realism, constructivism is attractive in apparently being able to accommodate such objective pretensions of ordinary moral discourse. Moreover, it attempts to accommodate these features without endorsing the sorts of metaphysical commitments to independently existing moral properties and facts countenanced by the realist. In short, at least certain versions of constructivism boast a robust notion of moral objectivity without problematic metaphysical commitments.

One serious challenge to constructivism is represented by the argument from moral error. According to constructivism, moral truths and associated facts are to be understood in terms of the attitudes and agreements of individuals and groups. However, if we take ordinary moral discourse and argument seriously, then since such discourse and argument presuppose that there are right answers to moral questions whose correctness outstrips any actual or even ideal set of attitudes or agreements, the constructivist view cannot be correct. To understand this objection more clearly, it will be useful to distinguish between basic and nonbasic moral truths and facts. Basic moral truths and facts are of a quite general sort, properly expressed by moral principles, and are the direct objects of choice by those under ideal conditions of moral thought and deliberation. Nonbasic moral truths and facts are those truths and facts that, in some sense, follow from the basic ones together with nonmoral information.

Now the constructivist can allow for certain sorts of errors in moral judgment. For instance, simple moral relativism can allow that individuals and groups can be mistaken about particular moral judgments owing to misinformation about particular cases or perhaps to faulty reasoning from basic moral principles to concrete cases. This kind of moral relativism, however, cannot allow for error at the level of actual agreements, since

such agreements constitute basic moral truths. The sophisticated constructivist can allow for error at the level of communal agreement, since it is possible on such views that the actual agreements of actual groups are at odds with those hypothetical choices constitutive of moral truth on this sort of view. However, the sophisticated constructivist cannot allow for error at the level of choice made under ideal conditions—call this "deep moral error." After all, the constructivist construes such choice as constitutive and not just evidence of basic moral truths and facts. But, so the objection goes, given our critical practices, we can sensibly raise questions about the truth of those moral principles and norms that are chosen under ideal circumstances. This indicates that moral truth is one thing and the norms and principles chosen even under the most ideal of circumstances is another. Hence, constructivism, in both its simple and sophisticated versions, is not acceptable.

In response, the constructivist can perhaps block the argument from moral error in the following way. First, the constructivist can note that it is dubious that our critical practices presuppose that deep moral error—error at the level of choice under ideal conditions—is possible. After all, our commonsense critical practices are not finely tuned to subtle differences in metaethical positions, and, in particular, common sense does not (so the constructivist might plead) make any distinction between the sort of realist objectivity that presupposes the possibility of deep moral error and a kind of constructivist objectivity that denies this possibility. Can we, for instance, really make sense of the idea that we might be mistaken about such basic moral principles as one that prohibits torture for fun? Furthermore, the constructivist can question the basic move featured in the argument from moral error—that is, the move from (1) it is quite sensible to raise questions about choices that purport to be made under ideal conditions to (2) an explanation of this phenomenon requires moral realism. Granted, the supposed gap between the truth of moral principles on the one hand and choice of such principles under ideal conditions on the other is one way to explain how we can sensibly raise questions about the truth of moral judgments made under ideal conditions, but this is not the only way to make sense of such critical stances.

The constructivist can note that in the context of everyday discussion where we have to judge whether or not to accept the moral judgments of others, one can sensibly raise questions about some judgment by raising questions about the judger herself. After all, whatever is involved according to the constructivist in being ideally well situated for choosing basic moral principles, it is not likely to involve features of the judger and her situation that are easy to detect. For example, part of being ideally well situated would seem to require having all sorts of factual information, being free from certain forms of bias, and properly weighing the interests of parties affected by the choice of principles. But it is difficult to determine that someone has satisfied these and other relevant desiderata for being well situated. So, even if it is not possible for someone who really is well situated to be mistaken in moral judgment, it is possible for critics who acknowledge that such error is not possible to raise sensible questions about the truth of a person's moral judgment. Hence, although the constructivist cannot allow for the possibility of deep moral error, she can plausibly argue that our commonsense critical practices do not presuppose that deep moral error is possible. Moreover, she can go on to accommodate the idea that it makes sense to criticize those who are ideally situated. The constructivist, it would appear, can plausibly respond to the argument from moral error.

See also Metaethics; Moral Realism; Rawls, John.

Bibliography

Brink, D. O. *Moral Realism and the Foundations of Ethics.* Cambridge, U.K.: Cambridge University Press, 1989. In chapter 2 Brink uses the argument from moral error against constructivism. Appendix 4 is a critical discussion of Rawlsian constructivism.

Firth, R. "Ethical Absolutism and the Ideal Observer." *Philosophy and Phenomenological Research* 12 (1952): 317–345. A classic statement of the ideal-observer version of constructivism.

Milo, R. "Skepticism and Moral Justification." *Monist* 76 (1993): 379–393. Milo defends a contractarian version of constructivism.

O'Neill, O. "Constructivisms in Ethics." In *Constructions of Reason.* Cambridge, U.K.: Cambridge University Press, 1990. O'Neill criticizes Rawls's version of constructivism and sketches what she takes to be a more plausible version of the view inspired by Immanuel Kant's writings.

Rawls, J. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77 (1980): 515–572. An elaboration of a constructivist view that centrally involves a Kantian conception of persons.

Timmons, M. "Irrealism and Error in Ethics." *Philosophia* 22 (1993): 373–406. A critical discussion of the argument from moral error.

Wong, D. *Moral Relativity.* Berkeley: University of California Press, 1984. Wong both criticizes various versions of moral relativism and defends his own version.

Mark Timmons (1996)

CONSTRUCTIVISM AND CONVENTIONALISM

"Conventionalism" and "constructivism" are kindred, often overlapping positions, asserting that the subject matter of some area of inquiry is not fully mind-independent. Conventionalism and constructivism are not well-defined names of positions but labels adopted—as often by critics as by advocates—to emphasize one positive aspect of positions in a wide range of areas; consequently, these terms group together a variety of positions with varying motivations. In general, the label "conventionalism" is applied to positions that claim the truths in some area are so in virtue of the conventions of a linguistic or conceptual scheme, while "constructivism" emphasizes that a position assigns to the cognitive faculties of humans some role in "making" the objects or facts in the area in question.

CONVENTIONALISM

Conventionalists claim either that the truths of some subject matter—such as mathematics or logic, or of a certain sort, such as necessary truths, or some dispute, such as whether Euclid's parallel postulate holds of our physical world—are matters of convention rather than of how the world is independent of mind. Some extreme versions of conventionalism take the fact that it is a matter of convention what our words mean (we could have used *cat* to designate Napoleon Bonaparte) to show that all truth is conventional. However, its being a convention that "Napoleon" names Napoleon hardly makes it conventional that Napoleon was defeated at Waterloo. An interesting conventionalism must assert something more than the conventionality of word meaning and must rest on something more than wild inference from it.

One area in which conventionalism is familiar, though controversial, is necessary truth. This was one cornerstone of logical positivism; from the seeming a priori nature of necessary truths, the positivists argued (some would say claimed) that since a priori knowledge cannot be of (mind-independent) facts, necessary truths must be analytic, which they understood as true by definition. Given that mathematics is also a priori, their argument was applied there as well. This sort of epistemological argument is typical of conventionalist views: Arguing that our methods for ascertaining what is so in some area could not give us knowledge of a mind-independent world, they claim that this knowledge would not be problematic on the assumption that what is fundamentally under investigation are our conventions. Some sorts of conventionalism are also supported by metaphysical considerations such as naturalistic concerns about what, in the mind-independent world, could make for the relevant sort of truth. This sort of argument is common to necessary truths, mathematics, ethics, and other areas with normative import; plainly, such arguments need to be supplemented with an account of how it is that conventions can provide the relevant features.

Saul Kripke's arguments that there are necessary truths that are a posteriori—and, so, not analytic—seemed to some to undermine conventionalism about necessity (Kripke 1980). It has, however, been argued that conventions could explain the necessity of these truths without the truths themselves being analytic—that is, true by convention (Sidelle 1989). This may indicate that in general conventionalism, with respect to a subject matter, does not require that all target truths themselves be analytic but only that conventions be responsible for the features that purportedly cannot be adequately handled within a realistic interpretation.

Aside from the claim that certain truths are so by convention, another common conventionalist position is that some dispute is a conventional rather than factual matter. Jules Henri Poincaré's famous conventionalism about geometry is of this sort. He claims that the choice among systems of geometry, for describing the physical world, is not an issue of which is true but of which is most convenient or useful. By adopting any of them, we could modify our physics so as to have equally full and correct descriptions of the world; indeed, this last claim is the basis for his view that the issue is conventional rather than factual. Rudolf Carnap offers a similar view about ontological disputes between, for instance, phenomenalists and materialists. Both of these views illustrate that "conventional" does not as such imply "arbitrary," as pragmatic differences may be quite genuine; we can also see that the plausibility of conventionalism in some area depends largely on how implausible it is to claim that the issue, or truth in question, is a matter of mind-independent fact.

On a more local level, some disputes can appear "purely verbal," as perhaps whether some politician is conservative. When this is plausible, the issue may be said to be a matter of convention or choice rather than fact. The conventionalism of Poincaré, Carnap, and others is akin to this, only in a wider application. In book 3, chapters 7 and 11, of his *Essay,* John Locke speculates that many of the "great disputes" are of this sort.

As applied to areas in which the truths are well established (mathematics or logic, for instance), conventional-

ism is fundamentally a deflationary interpretive position, urging that we not mistake the metaphysical status of these truths. Applied to areas of controversy—ontological or essentialist claims, or whether whales are fish—conventionalism claims that disputes here can only be over what our conventions in fact are, or what they should be, either pragmatically or perhaps morally. In either case, if conventionalism is right, our focus and methods of investigation—and certainly our understanding of what is at stake—for the questions at hand would probably require alteration.

CONSTRUCTIVISM

Thomas Kuhn, by virtue of his *The Structure of Scientific Revolutions,* may be considered constructivism's leading protagonist of the mid-to-late twentieth century, despite not adopting the label himself and expressing unease at having it assigned to him. He writes of scientists within different paradigms—roughly, methodological and theoretical traditions or frameworks—as studying different worlds and of their paradigms as in "a sense … constitutive of nature" (Kuhn 1970, p. 110, chaps. 10, 13), at least suggesting a constructivism about the world studied by science. Kuhn's major concerns are epistemological; he argues that scientific procedure is deeply theory laden and encodes ontological and theoretical commitments that it is incapable of testing. How, then, can such a method give us knowledge of the world? Those who see a constructivist in Kuhn have him answer that the world under investigation is itself partly a product of the investigating paradigm. This puts Kuhn in the tradition of Immanuel Kant, except that the features we "impose" upon the phenomenal world are not (as for Kant) necessary for the possibility of experience, but, rather, contingent features of current science. It is important to note that, even as interpreted, this constructivism does not have scientists making the world out of whole cloth with their paradigms; rather, there is something mind-independent that "filters through" the conceptual apparatus of the paradigm. This is a central difference between constructivism and idealism. The object of scientific study is, however, not this mind-independent world, but rather that which results "through the filter."

Other philosophers, as well as historians and sociologists of science, have taken the supposedly arational or nonobjective features guiding scientific judgment to establish that scientific truth is relative to one's background theory or paradigm. This is sometimes then articulated as the view that these theories or paradigms in part "make" the objects of study—that is, as constructivism. Indeed, many positions that formerly would have simply been called relativist came, in the late twentieth century, to be called constructivist by their protagonists; arguments in their support tend to be of the familiar relativist sort and thus have the same strengths and problems. It should be noted that neither constructivism nor conventionalism need take a relativistic form.

Another philosopher associated with constructivism is Nelson Goodman, due largely to his *Ways of Worldmaking.* Goodman argues that no sense can be made of the notion "the (one) way the world is"; rather, there are lots of ways the world is, depending on the conceptual apparatus one brings. This sort of position is found in many philosophers since Kant, often argued on the trivial ground that one cannot describe or investigate the world without using a system of representation, therefore (*sic*) the world investigated is not mind-independent but partly constructed by our conceptual scheme. This is sometimes added to, or confused with, the relativistic considerations mentioned above. What needs to be explained is how we are supposed to get this substantive conclusion from the banal premise. Why can't the objects represented by the elements of a system of representation—by the name "Tabby," say—be wholly and utterly mind-independent? And even if we add the fact that there can be different schemes of representation, why can it not simply be that they pick out different features of a mind-independent reality? What gives Goodman his special place is that he supplements this argument with the claim that different schemes may be such that their claims conflict with each other, but there can be no grounds for maintaining that one is correct and the other not. Goodman uses as examples the claims that the planets revolve around the sun and that the sun and other planets revolve around Earth. Both, he claims, must be judged as correct (within the appropriately formulated total systems), but they cannot simply be seen as two notationally different descriptions of a single world (thus differing from Poincaré's conventionalism). The success of this argument depends on whether one can simultaneously make out that these claims genuinely do conflict with each other and that, so understood, neither of them can be judged to be true while the other is false.

PROSPECTS

While both Kuhn and Goodman offer relatively global constructivist positions, there are constructivists about essences, moral and aesthetic properties, mathematical objects, and in principle anything. The same is true of conventionalism. Both conventionalism and construc-

tivism are motivated primarily by negative considerations against a realistic understanding of the subject matter in question; this is sometimes supplemented with positive arguments that by understanding the matter as concerning our conventions or choices we can get a better explanation of the phenomena at hand. Often, the negative arguments are very quick and fail to fully consider the range of options available to realists (Scheffler [1966] presents good discussion), and sometimes they fail to consider whether their positive proposals actually fare any better. Plainly, the plausibility of these positions depends on how well these arguments can be made out, and this may vary drastically across the different subject matters for which conventionalist and constructivist proposals have been offered. Additionally, if these positions are even to be candidates for serious consideration, defenders must be prepared to offer further proof. Conventionalists must specify some sense in which that which is purportedly so by convention would have been otherwise had our conventions been different, and constructivists must describe some sense in which the purportedly constructed objects would not have existed without our input.

See also A Priori and A Posteriori; Carnap, Rudolf; Conventionalism; Geometry; Goodman, Nelson; Kant, Immanuel; Kripke, Saul; Kuhn, Thomas; Locke, John; Logical Positivism; Poincaré, Jules Henri.

Bibliography

Carnap, R. "Empiricism, Semantics, and Ontology." In *Meaning and Necessity.* Chicago: University of Chicago Press, 1947. A classic exposition of the conventionalist's treatment of apparently ontological questions as fundamentally linguistic.

Dummett, M. *Frege: Philosophy of Language.* 2nd ed. Cambridge, MA: Harvard University Press, 1981. Chap. 16, esp. pp. 560–583. An interesting presentation of what may be seen as a constructivist position on objects or a conventionalist position on identity criteria.

Goodman, N. *Ways of Worldmaking.* Indianapolis: Hackett, 1978. For useful discussion, see C. G. Hempel, "Comments on Goodman's *Ways of Worldmaking,*" and I. Scheffler, "The Wonderful Worlds of Goodman," both in *Synthese* 45 (1980): 193–209.

Horwich, P. "A Defence of Conventionalism." In *Fact, Science and Morality,* edited by G. MacDonald and C. Wright. Oxford, 1986.

Knorr-Cetina, K. "The Ethnographic Study of Scientific Work: Towards a Constructivist Interpretation of Science." In *Science Observed: Perspectives on the Social Study of Science,* edited by K. Knorr-Cetina and M. Mulkay. London, 1983.

Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.

Kuhn, T. S. *The Structure of Scientific Revolutions.* 2nd ed. Chicago: University of Chicago Press, 1970.

Kyburg, H. E. "A Defense of Conventionalism." *Noûs* 11 (1977): 75–95.

Latour, B., and S. Woolger. *Laboratory Life: The Social Construction of Scientific Facts.* 2nd ed. Princeton, NJ: Princeton University Press, 1986.

Lewis, D. *Convention.* Cambridge, MA: Harvard University Press, 1969. Argues that sense can be made of the idea of nonexplicit conventions, *contra* Quine (1953).

Locke, J. *An Essay concerning Human Understanding* (1689). Locke's discussion of substance in book 3, chap. 7, classically expounds his conventionalism about essence and kinds.

Niiniluoto, I. "Realism, Relativism, and Constructivism." *Synthese* 89 (1991): 135–162. Critical discussion of constructivism in the social sciences.

Pap, A. *Semantics and Necessary Truth.* New Haven, CT: Yale University Press, 1958. An extended argument against conventionalism about necessity.

Poincaré, H. *Science and Hypothesis.* London: Walter Scott, 1905.

Putnam, H. "The Refutation of Conventionalism." In *Mind, Language, and Reality.* Cambridge, U.K.: Cambridge University Press, 1975.

Quine, W. V. O. "Truth by Convention." In *The Ways of Paradox and Other Essays.* New York: Random House, 1966. A classic argument against the possibility of truth by convention alone.

Quine, W. V. O. "Two Dogmas of Empiricism." In *From a Logical Point of View.* Cambridge, MA: Harvard University Press, 1953. Classical argument against the analytic-synthetic distinction and so against truth by convention.

Reichenbach, H. *Experience and Predication.* Chicago, 1938.

Scheffler, I. *Science and Subjectivity.* Indianapolis, IN: Bobbs-Merrill, 1967. Responds to arguments based on the history of science against objectivity in science, particularly Kuhn's, which are often used to bolster constructivist positions about the objects of scientific study.

Sidelle, A. *Necessity, Essence, and Individuation: A Defense of Conventionalism.* Ithaca, NY: Cornell University Press, 1989.

Alan Sidelle (1996)

CONTENT, MENTAL

Beliefs, desires, perceptions, and other mental states and events are said to possess content. We attribute such states and events with sentences such as

(1) Arabella believes that the cat is crying.

(1) contains a propositional attitude verb ("believes") and a sentence complement ("the cat is crying"). The verb specifies a type of mental state (belief), and the complement sentence indicates the content of the state. On most accounts this content is the proposition expressed by that sentence. Propositions have been variously conceived as abstract entities composed of modes of presen-

tation, sets of possible worlds, sets of synonymous sentences, and structured entities containing individuals and properties. All these accounts agree that propositions determine truth conditions. Some mental states and events (e.g., desiring to visit Paris) seem to have contents that are not propositions. However, for most of the current discussion, contents will be identified with propositions and contentful mental states with mental states that possess truth conditions.

Both natural-language sentences and mental states possess contents. The relation between content properties of the two items is controversial. Some philosophers think that natural-language expressions derive their contents from mental states, while others hold that, at least in some cases, the dependency goes the other way. In any case, it is plausible that there are mental states whose contents cannot be expressed or cannot be completely expressed by sentences of English (or other natural languages). For example, the full propositional content of a person watching the sun set is only partially captured by an attribution such as "*A* sees that the sun is setting." Also, some of the states posited by cognitive psychology and the mental states of animals plausibly have contents that fail to correspond to any contents expressible in English.

Content apparently endows mental states with a number of remarkable features. First, they or their constituents refer to extramental reality. When a person perceives that the sun is setting her perception refers to and thus puts her into contact with the sun. Second, they seem to be essentially normative. For example, a person ought to believe that the sun is setting only if the sun is setting, and if she believes that the sun is setting she ought not believe that the sun is not setting. Third, they apparently cause other mental states and behavior in virtue of their contents. For example, Arabella's belief that the cat is crying causes her to feed it. Fourth, a person can apparently know the contents of her own thought a priori and with an authority available only to her.

It is difficult to see how anything can exemplify all these features. The problem is especially difficult for philosophers who endorse naturalism, the view that all genuine properties are constituted by or realized by properties that are mentioned in true theories of the natural sciences. Content properties are prima facie so different from physical and biological properties as to raise the question of whether they are natural properties.

Hilary Putnam (1975) and Tyler Burge (1979) described thought experiments that have been taken to have important consequences for the nature of mental contents. Putnam imagined two thinkers, Oscar and twin-Oscar, who are identical with respect to their intrinsic neurophysiological properties but whose environments differ. Specifically. Oscar shares our environment, but twin-Oscar lives on twin-earth where the abundant substance that quenches thirst, fills the twin-earth oceans, and so forth is not H_2O but XYZ. H_2O's and XYZ's superficial properties are identical, and the two substances are indistinguishable without chemical analysis. Putnam claims that, while Oscar's sentence "Water is wet" and the thought he expresses with it are about H_2O, the same sentence in twin-Oscar's language and the thought he expresses with it are about XYZ. The two thoughts differ in their propositional contents, since one is true if and only if (iff) H_2O is wet and the other iff XYZ is wet. Putnam supports these conclusions with the intuition that, were Oscar and twin-Oscar to learn that the substances each refers to with the word *water* differ in their chemical natures, they would agree that their utterances of "Water is wet" possessed different truth conditions.

Putnam's thought experiment has been taken to establish the truth of content externalism, the thesis that the individuation conditions of mental content are partially external to the thinker. The point generalizes to other mental states whose contents are the same as the contents of sentences containing natural-kind terms such as *water*. Burge described further thought experiments that he thinks show that practically all thoughts expressible in natural language are externally individuated, and others have argued that all mental states that express extramental truth conditions are externally individuated (LePore and Loewer 1986).

Some philosophers (Fodor 1987, Loar 1987) react to content externalism by granting that mental states possess externally individuated contents but adding that they also possess narrow contents that are not externally individuated. Oscar's and twin-Oscar's beliefs possess the same narrow content. Philosophers sympathetic to narrow contents raise a number of considerations. One is the Cartesian intuition that thinkers in the same intrinsic state have the same mental lives. It seems essential to our conception of a mental life that it possess content, so there must be some kind of content that such thinkers share. The other consideration is that the causal powers of Oscar's and twin-Oscar's mental states seem to be, in an important way, the same. Jerry Fodor (1987) claims that if these causal powers involve the states' contents, then that content must be narrow.

Whether or not these considerations are persuasive, it has proved difficult to formulate a satisfactory notion of narrow content. If natural-language sentences express

only externally individuated contents, then we do not attribute narrow content with sentences such as (1). While identity of intrinsic neurophysiological states is sufficient for identity of narrow content, it is not a plausible necessary condition. To adopt it as such would make it enormously unlikely that two people have ever shared the same narrow content state and impossible for a state to maintain its content in the course of reasoning. While some proposals for necessary and sufficient conditions for identity of narrow content have been forthcoming (Fodor 1987), there is little agreement concerning whether they are correct or, for that matter, whether a notion of narrow content is even needed.

Externalism seems to be in tension with our having a priori knowledge of the contents of our thoughts (Boghossian 1989). If the content of the thought (e.g., that water is wet) is individuated in part by external factors, then it seems that a person could know that she is thinking this thought only if she knows that those external factors obtain, and thus it is implausible that such knowledge is a priori. One response to this is to grant that we have a priori knowledge only of narrow contents. But a number of philosophers (Burge 1988, Warfield 1994) have responded that the tension is only apparent. Burge claims that judgments of the form "I am now thinking that water is wet" are self-verifying, since one cannot make the judgment without thinking the thought that the judgment is about. If this is correct, then externalism and a priori knowledge of content are not always incompatible. But such self-verifying thoughts seem to be a very special case of the thoughts whose contents we seem able to know a priori. It is likely that little progress concerning the epistemology of content can be made without an account of the nature of contentful mental states.

The dominant view in the philosophy of mind is that contentful mental states are functionally individuated internal states. Some philosophers (Dretske 1981, Fodor 1987) posit that these states are partially constituted by mental representations that are the bearers of propositional content. Mental representations are conceived of as picturelike (mental images), maps, or linguistic expressions. One view (Fodor 1979) is that mental representations are expressions in a language of thought, Mentalese. On this account thinking that the cat is crying involves tokening a Mentalese sentence with the content that the cat is crying. The thought inherits its content from the semantic properties of its constituent sentence, which in turn obtains its content from the semantic properties of its constituent expressions. Fodor identifies concepts with Mentalese expressions. So, for example, possessing the concept cat is being able to token a Mentalese expression that refers to cats. Some philosophers (Peacocke 1986) have argued that the contents of perceptual states are nonconceptual. If so, then the contents of these states are not borne by Mentalese expressions.

The nature of the bearers of mental content is best seen as an empirical issue. Fodor (1987) cites the fact that thought is productive and systematic as support for the language-of-thought hypothesis. Productivity is the capacity to produce complicated thoughts by combining simpler thoughts, and systematicity involves being able to think thoughts that are systematically related to each other, as are the thoughts that Bill loves Newt and that Newt loves Bill. Fodor argues that the language-of-thought hypothesis provides the best explanation of these phenomena, since languages are productive and systematic. Further, cognitive scientists have constructed theories of cognitive processes, language comprehension (Pinker 1994), perception (Marr 1982), and so forth that involve subpersonal contentful mental representations. For example, on one such theory understanding a natural-language sentence involves tokening a representation of its grammatical structure. These representations are not accessible to consciousness and have contents that are not usually available as the contents of a person's beliefs.

There have been various attempts to specify conditions in virtue of which mental states or mental representations possess their contents. Some of these are attempts to naturalize content properties. Following are brief descriptions of the main proposals.

According to interpretationist theories (ITs; Davidson 1984, Lewis 1974) our practices of interpreting one another partially constitute the contents of mental states. On Donald Davidson's approach interpretation is constrained by principles of rationality and charity. These principles say, roughly, that a person's mental states are generally rational and her beliefs are generally true. According to Davidson the evidential base for an assignment of contentful mental states to a person consists of her dispositions to hold true sentences under various conditions. She believes that *p* (desires that *p*, etc.) iff assignments of content to her sentences and to her mental states that systematize these holding true dispositions and that conform to the principles of charity and rationality assign to her the belief that *p* (desire that *p*, etc.).

On ITs, content properties are holistic, since whether or not a person exemplifies a particular contentful mental state depends on what other mental states she exemplifies and on their relations to each other and to environmental conditions. Davidson's IT is externalist,

since a state's content is partially determined by relations to environmental conditions. But his account does not provide a naturalistic account of content, since it explains content in terms that presuppose content: holding true, rationality, truth. The primary difficulty with extant ITs is their vagueness. No one has formulated the principles of rationality and charity with sufficient clarity to permit an evaluation of proposed ITs.

According to conceptual role semantics (CRS), the content of a mental representation (or mental state) is determined by the inferential relations among representations and causal relations between representations and extramental events (Block 1986, Loar 1981, Sellars 1963). In this respect CRS is similar to IT. The difference is that, whereas ITs employ holistic principles of interpretations (rationality and charity), CRS attempts to spell out inferential patterns associated with particular concepts. CRS seems plausible for the logical connectives. For example, if a thinker is disposed to infer the representation *A#B* from *A* and *B* and vice versa, then # is the thinker's conjunction concept. Some philosophers (Peacocke 1992) have attempted to formulate conditions that are necessary and/or sufficient for possessing certain predicate concepts. It appears that any such account is committed to a substantial analytic-synthetic distinction, since it will hold that certain inferences involving a concept are necessary to having the concept (Fodor and LePore 1992). Willard Van Orman Quine's arguments (1960) that there are no analytic inferences poses an important problem for CRS.

Another approach is informational semantics (Dretske 1981, Stalnaker 1984). These theories are supposed to provide naturalizations of content; that is, they specify naturalistic properties that are claimed to be sufficient for possessing content. Informational theories claim that the content of a belief is constituted by the information the belief state carries under certain conditions. A state *S* carries the information that a property *P* is instantiated just in case the occurrence of *S* is caused by and nomically implies the instantiation of *P*. Informational theories have difficulty accounting for the possibility of error, since if a belief state has the content that *p* it carries the information that *p*. To solve this problem Fred Dretske proposed that the content of a belief is the information that it carries during what he calls "the learning period." A different suggestion (Stalnaker 1984) is to identify belief content with the information the belief state carries under epistemically optimal conditions. Barry Loewer (1987) has argued that these accounts are not successful as naturalizations, since they appeal to notions—learning, epistemic optimality—that themselves presuppose semantic notions.

Fodor has developed a sophisticated variant of informational theories that applies to the reference of Mentalese predicates. On this account, asymmetric dependency theory (ADT), a Mentalese predicate *C* refers to, for example, the property of being a cow if it is a law that cows cause *C*s, and any other causal relation between something other than cows and *C*s depends on this law but not vice versa. That is, if the other causal relations were to fail, it would still be a law that cows cause cows, but if the law were to fail, so would the other causal relation.

ADT is an atomistic account of content in that, contrary to CRS and ITs, it implies that the property of possessing a particular reference is metaphysically independent of inferential connections among thoughts and, indeed, independent of the existence of any other items with content. Whether or not one sees this as an advantage will depend on how one views the analytic-synthetic distinction. Obviously, ADT makes heavy use of metaphysical notions that are less than perspicacious, so one may wonder about its naturalistic credentials. It has also been argued (Boghossian 1991) that it is equivalent to an optimal-conditions account and is subject to the objections that show that account not to be a naturalization.

Teleological theories of content ground the contents of mental states in biological functions. The biological functions of a system in an organism are those of its features that increased the organism's fitness. Teleological accounts are quite elaborate, but the basic idea (Millikan 1984, Papineau 1992) is that there are desire-producing and belief-producing biological systems with certain biological functions. The desire-producing system has the function of producing states that tend to bring about certain effects. The effect associated with a particular desire is its content. The belief-producing systems have the function of producing states that tend to be tokened when certain states of affairs obtain. The state of affairs thus associated with a belief is its content.

Teleological accounts are appealing, since they are naturalistic, assign biological significance to contentful states, and seem to supply them with a kind of normativity. But various serious objections have been raised to teleological theories of content (Fodor 1992). The most serious is that it is doubtful that teleological considerations are sufficient to assign determinate contents to mental states. A desire state will typically tend to bring about a number of different advantageous effects. Natural

selection does not select any one of these effects as the content of the desire. Similarly, natural selection will not single out one of the states of affairs a belief state will typically be associated with as its unique content.

Whether or not content properties can be naturalized is an open question. Some consider it a very important question, since they think that if content properties cannot be naturalized then they are unsuitable to appear in scientific theories or, even worse, that they do not exist or are uninstantiated (Stich 1983). The unsuitability of content properties for science would be a blow to the emerging cognitive sciences. But the nonexistence of content properties would be devastating to the way we think about ourselves and others, since these ways are permeated with attributions of contentful states. In fact, it has been argued (Boghossian 1990) that the thesis that there are no content properties is incoherent. Fortunately, no dire consequences strictly follow from the failure of naturalization. It may be that content properties are natural but not naturalizable (McGinn 1991). It is possible that, while content properties are natural, connections between them and properties that occur in the natural sciences are too unsystematic or too complicated for us to discern. But whether or not this is so is also an open question.

Following Gareth Evans's discussion (1982), there has been growing interest in the proposals that there is a distinction between *conceptual* and *nonconceptual* content and that the latter plays a significant role in perception (and perhaps imagination) and in subdoxastic (and so unavailable to consciousness) cognitive processing. Exactly what this distinction amounts to, whether there is nonconceptual content, and what its explanatory and epistemological roles may be are all controversial matters.

Beliefs and other propositional attitudes involve relations to thoughts (or propositions), and concepts are constituents of thoughts. It follows that for someone to have the belief, for example, that the Supreme Court is about to convene, he must have the concepts *supreme court*, *about to convene*, and so on. A widely held necessary condition for concept possession is that one has the concept *C* only if one can think an array of thoughts involving *C*. This is similar to the systematicity that Jerry Fodor (1987/1990) appealed to support his claim that mental representations involved in thought are languagelike. In fact, he and others (who do not necessarily share his views about meaning) think of concepts as words in the mental language deployed in thinking and in propositional attitudes. Evans observed that it is plausible that there are mental states whose contents are not conceptually articulated in this languagelike way. Visual perception seems to involve such states and processes. When one is looking at, for example, a sunset over a distant mountain range, one's perception seems much richer than what can be expressed in thought. There are particular colors and shapes represented in the perception that one is not able to represent in thought. Further, there do not seem to be components of visual representations that one can combine in the systematic ways in which concepts can be combined. In addition to perceptual states, the mental states of animals and the subdoxastic mental representations of humans posited by cognitive scientists are also said to have nonconceptual contents.

On some accounts of mental content, it is not clear that there can be nonconceptual content. For example, accounts like Donald Davidson's (1984), in which there is an intimate relationship between mental contents and the contents of public-language expressions and in which rationality constraints play a role in content determination, seem to preclude there being contents that cannot be expressed in public language. John McDowell (1994), who advocates such an account, has argued against the existence of nonconceptual content. Specifically, he thinks it is essential to mental states with content that they enter into rational and justifying relations with one another and claims that this requires that their contents be articulated conceptually. Famously, views like these resist attributing contentful mental states to animals and to subdoxastic mental processes, since animals and subdoxastic mental processes cannot harbor concepts.

Philosophers who think of mental content in terms of information (examples are Evans, Fodor, and Fred Dretske) can allow for states with nonconceptual content since non-conceptual representations can possess informational content. Dretske thinks that nonconceptual content is more basic than and prior to conceptual content, and that the latter is in some way derived from the former. Since information makes a division of possible worlds into those in which the information is correct and those in which it is not an informational state can stand in semantic relations of entailment and incompatibility with conceptual representations. On the other hand, it is not obvious how nonconceptualized information states can be involved in inference and reasoning.

There are a number of issues that advocates of nonconceptual content need to address. One is whether the distinction between conceptual and nonconceptual content is really a distinction between *kinds* of content or a distinction between different *ways* of representing content. Of course, this depends on what one takes content to be. As noted, states with nonconceptual content, like

those with conceptual content divide possible worlds into those that are, and those that are not, in conformity with the content. However, conceptual content is often thought of as involving structure that reflects conceptual composition. The question is whether this structure is better understood as a species of content or belongs to the representation that represents the content. Those who think of contents solely in possible-world terms, such as Robert Stalnaker (1998), will see structure as belonging to the representation.

Another issue is that the alleged nonconceptual content of a perception seems to be informationally much more rich than the content of a thought. At the same time, it seems to be finer-grained in that there are distinctions that can be made in perception that we do not and perhaps cannot represent conceptually. It is not clear how these two features fit together. One idea is to think of nonconceptual content as analogous to pictorial or maplike content (Peacocke 2001). If this is correct, it raises the question of whether the pictorial structure belongs to the content or to the representation. Also, as mentioned above, there are issues concerning the epistemological role of nonconceptual content. Can a nonconceptual perceptual state justify a perceptual belief that it causes? Finally, it is not clear whether the contents involved in perception, the mental states of animals, and subdoxastic states and processes are all the same kind of nonconceptual contents. Indeed, theorists who appeal to subdoxastic states and processes often posit sentencelike representations as involved in mental computations. Their contents are nonconceptual only in that they are not available to thought and propositional attitudes.

See also Belief; Concept; Davidson, Donald; Internalism versus Externalism; Knowledge, A Priori; Language of Thought; Naturalism; Philosophy of Mind; Propositional Attitudes: Issues in the Philosophy of Mind and Psychology; Propositional Attitudes: Issues in Semantics; Putnam, Hilary; Quine, Willard Van Orman.

Bibliography

Block, N. "Advertisement for a Semantics for Psychology." In *Midwest Studies in Philosophy*. Vol. 10. Minneapolis: University of Minnesota Press, 1986.

Boghossian, P. "Content and Self-Knowledge." *Philosophical Topics* 17 (1989): 5–26.

Boghossian, P. "Naturalizing Content." In *Meaning and Mind*, edited by B. Loewer and G. Rey. Oxford, 1991.

Boghossian, P. "The Status of Content." *Philosophical Review* 99 (1990): 157–184.

Burge, T. "Individualism and Self-knowledge." *Journal of Philosophy* 85 (1988): 644–658.

Burge, T. "Individualism and the Mental." In *Midwest Studies in Philosophy*. Vol. 4. Minneapolis: University of Minnesota Press, 1979.

Davidson, D. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 1984.

Dretske, F. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

Evans, Gareth. *The Varieties of Reference*. Oxford, U.K.: Oxford University Press, 1982.

Fodor, J. *The Language of Thought*. Cambridge, MA, 1979.

Fodor, J. *Psychosemantics*. Cambridge, MA: MIT Press, 1987.

Fodor, J. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 1990.

Fodor, J., and E. LePore. *Holism: A Shopper's Guide*. Oxford: Blackwell, 1992.

Grice, P. "Meaning." *Philosophical Review* 66 (1957): 377–388.

LePore, E., and B. Loewer. "Solipsistic Semantics." *Midwest Studies in Philosophy* (1986).

Lewis, D. *Convention*. Cambridge, MA: Harvard University Press, 1969.

Lewis, D. "Radical Interpretation." *Synthese* 23 (1974): 331–344.

Loar, B. *Mind and Meaning*. Cambridge, U.K.: Cambridge University Press, 1981.

Loar, B. "Social Content and Psychological Content." In *Contents of Thought*, edited by R. Grimm and D. Merrill. Tucson: University of Arizona Press, 1986.

Loewer, B. "From Information to Intentionality." *Synthese* 70 (1987): 287–317.

Marr, D. *Vision*. San Francisco: Freeman, 1982.

McDowell, John. *Mind and World*. Cambridge, MA: Harvard University Press, 1994.

McGinn, C. *The Limits of Philosophy*. Oxford, 1991.

Millikan, R. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press, 1984.

Papineau, D. *Naturalism*. Oxford, 1992.

Peacocke, C. "Does Perception Have a Non-conceptual Content." *Journal of Philosophy* 98 (2001): 239–264.

Peacocke, C. *A Study of Concepts*. Cambridge, MA: MIT Press, 1992.

Peacocke, C. *Thoughts: An Essay on Content*. New York: Blackwell, 1986.

Pinker, S. *The Language Instinct*. New York: Morrow, 1994.

Putnam, H. "The Meaning of Meaning." In his *Mind, Language, and Reality*. Cambridge, U.K.: Cambridge University Press, 1975.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Schiffer, S. *Meaning*. Oxford: Clarendon Press, 1972.

Sellars, W. *Science, Perception, and Reality*. New York: Humanities Press, 1963.

Stalnaker, R. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Stalnaker, Robert. "What Might Non-conceptual Content Be?" In *Concepts: Philosophical Issues*, edited by Enrique Villanueva. Atascadero, CA: Ridgeview, 1998.

Stich, S. *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press, 1983.

Warfield, F. "Externalism and Self-Knowledge." *Analysis* (1994).

Barry Loewer (1996, 2005)

CONTEXTUALISM

The term "contextualism" has been used to denote many different philosophical theories. Within epistemology alone, there are two broad categories of theories that have been called "contextualist": subject contextualism and attributor contextualism.

SUBJECT CONTEXTUALISM

A few basic concepts are needed to explain subject contextualism. Let *S* be an epistemic subject, a being whose cognitive attitudes are proper targets of epistemic evaluation. Let *C* be a cognitive attitude that *S* has. *C* may be a belief, a judgment, a high degree of confidence, an affirmation or endorsement of some kind—any attitude that is a proper target of epistemic evaluation. *C* has a propositional content *p*. Finally, let *x* be the situation in which *S Cs* that *p*. We will hereafter specify the target of epistemic evaluation as "*S*'s *C*ing that *p* in *x*."

According to subject contextualism, whether *S C*s that *p* in *x* constitutively depends on features of *x* that are metaphysically independent of *S*'s cognitive attitudes and of the truth-values of the propositional contents of *S*'s cognitive attitudes. As these features of *x* vary, so too does the epistemic status of (the degree of truth attached to) *S*'s *C*ing that *p*, even if *S*'s cognitive attitudes and the truth-values of their propositional contents all remain fixed. One or another version of such a view has been suggested in various passages in the writings of C. S. Peirce, Ludwig Wittgenstein, John Dewey, Karl Popper, W. V. Quine, and J. L. Austin. But only since the mid-1970s has subject contextualism been developed with any precision and generality.

The different versions of subject contextualism differ from each other in at least two ways. First, these different versions specify different features of *x* as relevant to constitutively determining whether *S C*s that *p* in *x*. Second, these different versions of subject contextualism specify different ways in which the relevant feature of *x* can determine whether *S C*s that *p*. By differences of the first kind, we can distinguish the various theories of subject contextualism that have been propounded into three broad groups.

According to one group of subject contextualist theories (Stine 1976, Goldman 1976, Dretske 1981), the epistemic status of *S C*ing that *p* in *x* constitutively depends on the objective probability of *p*'s being true in *x*. Other things being equal, the higher the objective probability of *p*'s being true in *x*, the higher the epistemic status of *S C*ing that *p* in *x*.

The most prominent argument in favor of this first variety of subject contextualism proceeds from consideration of case pairs such as the following (from Goldman 1976): Suppose that, in normal daylight, Henry, who has normal visual powers, has an unobstructed view of a barn right in front of him. Henry sees the barn, has a normal visual experience as of a barn, and believes that there is a barn in front of him. If there is nothing unusual about the case, then Henry knows that there is a barn in front of him. But now suppose that Henry's environment is full of barn facades that look exactly like barns from the angle and distance at which Henry is currently viewing the real barn in front of him. In this second case, just as in the first case, Henry has a true belief that there is a barn in front of him. And in both cases, this belief is based on Henry's seeing the barn, on his normal visual experience as of a barn. But in the second case, unlike the first, Henry does not know that there is a barn in front of him. What difference in the two cases could account for this difference in whether or not Henry knows? Subject contextualists of the first variety say that the difference in Henry's extrapsychological environment—in particular, the frequency of barn facades in his environment—is responsible for the difference in whether he knows.

Opponents of this first variety of subject contextualism typically respond to the preceding argument by offering an alternative explanation of why Henry knows, in the first case but not in the second case, that there is a barn in front of him. In each case, they contend, Henry believes that there is a barn in front of him partly because he believes that, in his environment, things that look like barns from his vantage point typically are barns. So his belief that there is a barn in front of him is partly based on the latter epistemic belief. And the epistemic belief is true in the first case and false in the second case. It is the difference in the truth-value of the epistemic belief that explains why Henry knows in the first case but does not know in the second case—or so say the opponents of this first variety of subject contextualism.

According to a second group of subject-contextualist theories (Annis 1978, Williams 1992, Henderson 1994, Klein 1999), whether *S C*s that *p* in *x* constitutively depends on the inquiry that takes place in *x*. The inquiry fixes which considerations confer positive epistemic status upon *S*'s *C*ing that *p*. The main argument for this second variety of subject contextualism, propounded in different ways by Annis, Williams, and Klein, has to do with the regress of reasons. According to this argument, neither foundationalism nor coherentism offers a correct account of structure of epistemic reasons, or justifica-

tions. Foundationalism cannot offer a correct account, because it is committed to the unsustainable claim that some cognitive attitudes—the foundational beliefs—are intrinsically justified. And coherentism cannot offer a correct account, because it is committed to claiming either that circular reasoning provides justification or that each belief in a coherent set is a foundational belief. But if there are no foundational beliefs and if circular reasoning does not provide justification, then how can positive epistemic status accrue to *S*'s *C*ing that *p*? According to this second variety of subject contextualism, this question is best answered as follows: *S*, in a particular context of inquiry, makes certain presuppositions. These presuppositions can provide epistemic reasons, or justifications, for *S*'s other cognitive attitudes in that same inquiry. But when *S* moves into a different context of inquiry, some presuppositions in the earlier inquiry may be put into question in the new inquiry, and other propositions that were in question in the earlier inquiry may simply be presupposed in the new inquiry.

To the preceding argument, Henderson adds that because our cognitive competence is limited, in ways that can be empirically ascertained, we are incapable of forming beliefs about our environment without taking a great deal for granted. What we need to take for granted to form needed beliefs will vary from task to task. Since we are incapable of forming the beliefs that we need to form without taking a great deal for granted, we cannot be epistemically obligated to do otherwise. Our epistemic obligations cannot exceed our cognitive potential. Since our belief-forming processes require us to take more or less for granted, depending on the cognitive task at hand, our epistemic obligations must allow us to take more or less for granted, depending on the task.

Opponents of this second variety of subject contextualism typically respond to the preceding arguments by defending foundationalism or coherentism. While foundationalist and coherentist theories of justification might hold us to normative standards that we do not commonly meet, this is not a problem for those theories if, in the epistemological realm, "ought" does not imply "can."

According to a third group of subject-contextualist theories (Fantl and McGrath 2002, Hawthorne 2003, Stanley 2004), whether *S C*s that *p* in *x* constitutively depends on how the truth-value of *p* affects *S*'s interests, or perceived interests, in *x*. Other things being equal, the higher the cost, or perceived cost, of *S*'s being wrong that *p*, the less likely that *S C*s that *p*. The most prominent argument in favor of this third variety of subject contextualism considers pairs of cases such as the following (adapted from Fantl and McGrath 2002): Suppose that you are at the train station waiting for the train to New York. You would like to get on the express train so that you can be in New York by dinnertime, but it does not matter all that much to you whether you get there by dinnertime or not. You ask someone else on train platform, "Is the next train an express train, or a local?" and your honest and knowledgeable interlocutor sincerely tells you that it is an express. You believe her, and you have no reason to distrust her. In this situation, it seems that you know, and are justified and warranted in believing, that the next train is an express. But now suppose that the situation is exactly the same except that your life depends on your being in New York by dinnertime. In this case, the testimony of your honest and knowledgeable interlocutor does not justify or warrant—let alone give you knowledge—that the next train is an express. When so much depends on your being right, knowledge, justification, warrant, etc., all require more than they otherwise would require.

Opponents of this third variety of subject contextualism typically respond to the preceding argument by claiming that what depends on a subject's actual or perceived interests is not the epistemic status of the subject's *C*ing that *p*, but rather the rationality of the subject's acting as if *p* were true. For these opponents, whether *S C*s that *p* is fixed independently of *S*'s actual or perceived interests.

These are the main varieties of subject contextualism, and the arguments concerning them. Some subject contextualists are also attributor contextualists, and almost all attributor contextualists are also subject contextualists.

ATTRIBUTOR CONTEXTUALISM

A few more basic concepts are needed to explain what attributor contextualism is. Let *A* be an epistemic attributor, someone who epistemically evaluates *S*'s *C*ing that *p* in *x*. *A*'s evaluation of *S*'s *C*ing that *p* in *x* will also be a cognitive attitude of some sort, either expressed (as in the case of an assertion) or not (as in the case of silent thought). *A*'s evaluation may concern whether *S* knows that *p* in *x*, or it may concern whether *S* is justified, reasonable, rational, or warranted in *C*ing that *p* in *x*, or it may concern whether *S* has adequate grounds, evidence, or reasons to *C* that *p* in *x*. More generally, an epistemic evaluation or appraisal of *S*'s *C*ing that *p* in *x* is a determination of whether *S C*s that *p* in *x*. Let *y* be the situation that *A* is in when *A* evaluates whether *S C*s that *p* in

x. We will hereafter specify the act of epistemic evaluation as "*A*'s evaluation, in *y*, of *S*'s *C*ing that *p* in *x*."

According to attributor contextualism, the semantic value (the truth) of *A*'s evaluation (of *S*'s *C*ing that *p* in *x*) constitutively depends on features of *y*. As these features of *y* vary, so too does the semantic value of *A*'s evaluation, even with everything else held fixed. The earliest prominent statements of such a view appear in Lewis (1979) and Dretske (1981). The view gained widespread notice following the publication of Cohen (1988), DeRose (1995), and Lewis (1996).

Different versions of attributor contextualism specify different features of *y* as relevant to constitutively determining the semantic value of *A*'s evaluation of *S*'s *C*ing that *p* in *x*. And even if two attributor contextualists agree about which features of *y* are relevant to constitutively determining the semantic value of *A*'s evaluation, they might still disagree about precisely *how* those features of *y* are relevant. By differences of the first kind, we can distinguish the various attributor-contextualist theories on offer into several groups. Although there is thus some diversity among attributor-contextualist theories, a single line of argument has generally been used to support attributor contextualism.

The argument in question proceeds from consideration of cases similar to those commonly adduced to support the third variety of subject contextualism. Suppose that Jones and Smith are at the train station trying to catch a train to New York. They want to know whether the next train is an express. They ask a bystander if he knows whether the next train is an express. The bystander looks at a schedule and replies, "Yes, I know. It is an express." It turns out that Jones and Smith have to be in New York as soon as possible, and cannot afford the mistake of getting on a local train. Jones says, "That schedule could easily have been outdated. That guy does not really know that the next train is an express." So the bystander claims to know that the next train is an express, but Jones claims that the bystander does not know that the next train is an express. Who is right?

If the bystander is right that he knows, then is Jones making a false assertion when he says that the bystander does not know. Suppose that the bystander's warrant for thinking that the next train is an express is precisely what Jones takes it to be. If such warrant is strong enough to give the bystander knowledge, then, it seems, it is also strong enough to give Jones and Smith knowledge. But if Jones and Smith have warrant strong enough to give them knowledge that the next train is an express, they have no reason to check further whether the next train is an express. Since they clearly do have a reason to check further whether or not the next train is an express, their warrant cannot be strong enough to give them knowledge of whether or not it is. But if they do not have enough warrant for knowledge, then it seems that the bystander does not know, since he does not have any more warrant than they do.

Suppose that the bystander is wrong, that he does not know that the next train is an express. In this case, it seems that most of the knowledge attributions that we make in ordinary life are wrong as well, since our warrant for most of what we claim to know is no greater than is the bystander's warrant for the claim that next train is an express. So if the bystander does not know that the next train is an express, then most of us know very little of what we ordinarily claim to know.

How can we avoid simply granting that Jones and Smith are right to deny that the bystander knows, or that the bystander is right to claim to know? The attributor contextualist avoids granting this by claiming that the truth-values of knowledge attributions are relative to the context in which the attribution is made. Relative to the context in which the bystander claims to know, her claim is true. But relative to the context in which Jones claims that the bystander does not know, his claim is true. So both claims are true, and they do not contradict each other. These assertions only appear to contradict each other because we fail to notice that "knows" requires a higher standard (or signifies a more stringent epistemic relation) in one context of attribution than in the other context of attribution than in the other context of attribution. Analogous arguments may support the conclusion that ascriptions of other epistemic properties, not just ascriptions of knowledge, are semantically sensitive to context.

Opponents of attributor contextualism will typically reply to an argument like the preceding in one or both of the following two ways. First, like Bach (2005), they may claim that, although the bystander knows that the next train is an express, Jones and Smith do not know, and that is because knowledge requires sufficient confidence. Although the bystander is sufficiently confident that the next train is an express, Jones and Smith are not, and so they do not satisfy one of the necessary conditions of knowledge. If, without any further investigation, Jones and Smith claimed to be sufficiently confident of the bystander's claim, then, given how much is at stake in their being right, their degree of confidence would be irrational.

This suggests a second line of response to the attributor-contextualist argument above: Even if Jones and Smith, prior to doing any further investigation, are sufficiently confident that the next train is an express, their confidence is unreasonable, given how much is at stake. Knowing that *p* requires not simply that one be sufficiently confident that *p*, but moreover that one's level of confidence be reasonable. Since, without doing any further investigation, Jones and Smith cannot be reasonably confident that the next train is an express, they also cannot know that the next train is an express, even if they are sufficiently confident of it (by the bystander's standards), even if they share all of the bystander's evidence for it, and even if the bystander himself knows. On this second line of response, we resolve the problem set out by the attributor contextualist's argument by appealing to subject contextualism of the third variety distinguished above.

To bolster the argument for attributor contextualism in the face of these objections, attributor contextualists must now attempt to run their thought experiments while controlling for variation in *S*'s level of confidence and also in *S*'s level of reasonable confidence. To do this, they must focus on a particular epistemic subject *S* in a particular context *x*, and then find variation in the truth conditions for asserting that *S C*s that *p* in *x*. To make the case for attributor contextualism, they must make sure that the variation they discover is variation in the truth conditions of an ascription of knowledge, and not simply in the conditions under which we are inclined to make, or are warranted in making, the ascription. DeRose (2004) undertook to do all this.

Three other sorts of argument that have commonly been used to support attributor contextualism. The first is an argument to the effect that attributor contextualism provides the best response to a skeptical argument like the following:

> Premise 1: I cannot possibly know that I am not a brain in a vat being electrochemically stimulated to have realistic experiences.
>
> Premise 2: If I knew that I have hands, then I could deduce, and thereby come to know, that I am not a brain in a vat being electrochemically stimulated to have realistic experiences.
>
> Conclusion: I do not know that I have hands.

While some philosophers would simply to deny one of the premises, attributor contextualists typically take such denial to be implausible. So how can attributor contextualists avoid accepting the skeptical conclusion of such an argument? They can do so by claiming that the skeptical conclusion is true *only relative* to contexts of attribution that we enter into by thinking (in some way or other) about premise 1. Relative to other, more commonplace, contexts of attribution, premise 1 is false, as is the skeptical conclusion of the argument. Attributor contextualists typically take this response to the skeptical argument above to be more plausible than any alternative response, and they take this to count as a point in favor of attributor contextualism concerning knowledge attributions. Analogous arguments have been adduced in favor of attributor contextualism concerning attributions of other epistemic properties.

A second style of argument in favor of attributor contextualism proceeds from premises concerning the epistemic properties in question. For instance, Dretske (1981) and Lewis (1996) both claim that knowledge that *p* involves having infallible grounds for one's belief that *p*. For them, one's grounds are infallible just in case all alternatives to *p* are ruled out. But "all," like other quantifiers, involves a contextually restricted domain of quantification. Ruling out "all" alternatives means ruling out all those that fall within the contextually restricted domain of quantification. So on this account, some form of attributor contextualism is true of knowledge attributions. Analogous arguments show that attributor contextualism is true of attributions of other epistemic properties as well.

Finally, a third style of argument in favor of attributor contextualism proceeds from premises concerning the conversational function of epistemic-property attributions and epistemic appraisal. Such arguments (e.g., Neta 2002, Schaffer 2004) claim that for attributions of epistemic properties to serve the function that they are supposed to serve, they must be semantically sensitive to contexts. For instance, if a knowledge attribution of the form "*S* knows why *p*" functions to signal to one's interlocutors that they can trust *S* on the topic of why it is that *p*, then whether it is appropriate to make such an attribution depends on whether one should signal to one's interlocutors that they can trust *S*, and this in turn depends on features of the conversational context. If this appropriateness depends on conversational context because the truth conditions depend on conversational context, then attributor contextualism is true of knowledge-why attributions. Analogous arguments may lead to attributor contextualism for attributions of other epistemic properties as well.

These are some of the main lines of argument that have been adduced in favor of one or another variety of

attributor contextualism. Here is a review of those varieties:

According to one group of attributor-contextualist theories (Cohen 1986, 1988, 1999), what varies with the context of attribution is the threshold of evidential support for *p* that must be exceeded for *S* to *C* that *p* (for *S* to know, to be justified, to be warranted, etc.). In a particular context *x*, *S*'s evidence confers a certain level of epistemic support on the proposition that *p*. Does that level of support suffice to warrant asserting that *S Cs* that *p*? This first kind of attributor contextualism takes the answer to this question to depend on features of the context of attribution *y*.

According to a second group of attributor-contextualist theories (Dretske 1981, Lewis 1996, Schaffer 2004), what varies with the context of attribution is the range of relevant alternatives to *p* that *S*'s evidence must rule out for *S* to *C* that *p*. To say that *S*'s evidence must rule out these alternatives to *p* is not to say anything about what *S* does or does not know, or about what *S* does or does not believe. Rather, it is to say that *S* has adequate evidence only if these alternatives to *p* do not obtain. Epistemologists standardly assume that *S*'s evidence cannot rule out all alternatives to *p*—in particular, it cannot rule out the alternative that *p* is false but *S* is being deceived by a deceiving spirit into believing that *p*. But *S*'s evidence does not need to rule out all alternatives to *p* for *S* to *C* that *p*. Rather, for *S* to *C* that *p*, *S*'s evidence must rule out only the relevant alternatives to *p*. Which alternatives are relevant? That depends upon the context of attribution *y*.

According to a third group of attributor-contextualist theories (DeRose 1992, 1995; Heller 1995, 1999), what varies with the context of attribution is the range of possible situations throughout which, in order for *S* to count as knowing that *p*, *p* must be true if and only if *S Cs* that *p*. In a particular context *x* in which *S Cs* that *p*, *S* has the disposition to *C* that *p* just in case certain conditions obtain. *S Cs* that *p* if and only if there is an adequate range of conditions under which *S* is disposed to *C* that *p*. But what range of conditions is "adequate"? That is relative to a context of attribution *y*.

According to a third group of attributor-contextualist theories (DeRose 1992, 1995; Heller 1995, 1999), what varies with the context of attribution is the range of possible situations in which, for *S* to count as knowing that *p*, it is required that *p* is true if and only if *S Cs* that *p*. In a particular *x*, *S Cs* that *p*just in case certain conditions obtain (such as, for instance, *S*'s being an authority). What is the range of conditions throughout which this biconditional must hold for *S* to know that *p*? That is determined by the context of attribution *y*.

A view that combines features of the last two varieties of attributor contextualism is defended by Rieber (1998). According to Rieber, *S* knows that *p* if and only if the fact that *p* explains *S*'s belief that *p*. An explanation answers the question "Why?" "Why" questions are contrastive: To ask "Why is it that *p*?" is always, at least implicitly, to ask "Why is it that *p* rather than that *q*?" For Rieber, ascriptions of knowledge inherit the contrasts of explanation statements. Thus, for Rieber, for *S* to know that *p* is for there to be some contrast proposition *q* such that *S* knows that *p* rather than that *q*. And for the latter to hold true, on Rieber's account, the fact that *p*, rather than the fact that *q*, must explain *S*'s belief that *p*. On Rieber's view, then, *S* knows that *p* if and only if the fact that *p* (rather than the contrast proposition *q*) explains the fact that *S* believes that *p*. On Rieber's view, the context of attribution *y* determines the contrast proposition.

Finally, according to the most recently espoused version of attributor contextualism (Neta 2002, 2003a, 2004), what varies with the context of attribution is the range of propositions, or of psychological states, that count as part of *S*'s evidence set. Relative to some contexts of attribution, *S*'s evidence set may include nothing more than *S*'s own current states of consciousness. But relative to more ordinary contexts of attribution, *S*'s evidence set may include various propositions about, say, widely known results of experiments that took place completely independently of *S*. More generally, according to attributor contextualism concerning evidence, whether *S Cs* that *p* depends on *S*'s evidence for *p* for other epistemic properties *C* as well (e.g., knowing, being justified, having warrant).

THE RECENT CONTROVERSY OVER ATTRIBUTOR CONTEXTUALISM

Since the late 1990s, attributor contextualism has been subject to two sorts of objections. According to the first sort of objection, the problem with attributor contextualism is that it implausibly attributes to native speakers a significant level of semantic self-ignorance. We can see this either by thinking about attributor-contextualist responses to skeptical arguments, as Schiffer (1996) and Rysiew (2001) do, or by thinking about the consequences of attributor contextualism for our practices concerning disquotation of knowledge attributions, as Hawthorne (2003) and LePore and Cappellen (2004) do.

According to attributor-contextualist responses to skeptical arguments, such arguments gain their plausibil-

ity because, when going through these arguments, we confuse the propositions that our epistemic-property attributions express with the propositions that these attributions would express in certain other contexts. But, according to the proponents of this objection, it is implausible to claim that native speakers do indeed suffer from this confusion. Again, according to attributor contextualists, attributor *A* can, in some contexts, truthfully assert something of the following form: "*S* does not know that *p*, even though *S* speaks truthfully when *S* says, 'I know that *p*,' " or, more generally, "*S* is not justified or warranted in *C*ing that *p*, even though *S* speaks truthfully when *S* says, 'I am justified or warranted in *C*ing that *p*.' " But such assertions appear self-contradictory to native speakers. Thus, the attributor contextualist is committed to claiming that native speakers are wrong to think that such assertions are self-contradictory. Once again, the attributor contextualist is committed to attributing a significant level of semantic self-ignorance to native speakers.

In response to this first line of objection, Neta (2003b) raises the question of whether the level of self-ignorance that attributor contextualism posits is any greater than the level of self-ignorance about their own language that native speakers routinely display at other levels of linguistic analysis (e.g., pragmatics, syntax, phonology). If native speakers generally do not realize that there is a difference between the "t" sound in "butter" and the "t" sound in "putter," then why should they realize that there is a difference in the meaning of terms of epistemic appraisal, "*C*," as they occur in different contexts? Of course, naive speakers can be brought to notice the difference between the "t" sound in "butter" and the "t" sound in "putter." Can the attributor contextualist bring naive speakers (even if not theoretically invested philosophers) to discern a difference in the meaning of terms of epistemic appraisal as they occur in different contexts? This remains an open empirical question.

Stanley (2000, 2004), pursuing the second line of objection, has argued that attributor contextualism is empirically implausible because there is no well-established precedent for the particular kind of semantic context-sensitivity that attributor contextualists posit in our epistemic-property attributions. These empirical arguments have been rebutted most recently by Ludlow (2005) and DeRose (2005), but this issue, like many other empirical issues in semantics, remains unsettled. Indeed, Unger (1984, 1986) has argued that there is no empirically ascertainable fact of the matter as to whether attributor contextualism is true.

See also Austin, John Langshaw; Coherentism; Dewey, John; Dretske, Fred; Epistemology; Lewis, David; Peirce, Charles Sanders; Popper, Karl Raimund; Quine, Willard Van Orman; Skepticism, History of; Wittgenstein, Ludwig Josef Johann.

Bibliography

Annis, David. "A Contextualist Theory of Epistemic Justification." *American Philosophical Quarterly* 15 (1978): 213–219.

Austin, J. L. "Other Minds." *Proceedings of the Aristotelian Society*, supp. vol. 20 (1946): 149–187.

Bach, Kent. "The Emperor's New 'Knows.' " In *Contextualism in Philosophy*, edited by Gerhard Preyer and Georg Peter. Oxford, U.K.: Oxford University Press, 2005.

Cappelen, Herman, and Ernest LePore. *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Oxford, U.K.: Blackwell, 2004.

Cohen, Stewart. "Contextualism, Skepticism, and the Structure of Reasons." In *Epistemology, 1999*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1999.

Cohen, Stewart. "Knowledge and Context." *Journal of Philosophy* 83 (1986): 574–583.

Cohen, Stewart. "Knowledge, Context, and Social Standards." *Synthese* 73 (1987): 3–26.

Cohen, Stewart. "How to Be a Fallibilist." In *Epistemology, 1988*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1988.

DeRose, Keith. "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research* 52 (1992): 913–929.

DeRose, Keith. "The Ordinary Language Basis for Contextualism and the New Invariantism." *Philosophical Quarterly* 55 (2005): 172–198.

DeRose, Keith. "The Problem with Subject Sensitive Invariantism." *Philosophy and Phenomenological Research* 68 (2004): 346–350.

DeRose, Keith. "Solving the Skeptical Problem." *Philosophical Review* 104 (1995): 1–51.

Dretske, Fred. "The Pragmatic Dimension of Knowledge." *Philosophical Studies* 40 (1981): 363–378.

Fantl, Jeremy, and Matthew McGrath. "Evidence, Pragmatics, and Justification." *Philosophical Review* 111 (2002): 67–94.

Gendler, Tamar, and John Hawthorne, eds. *Oxford Studies in Epistemology*. Vol. 1. Oxford, U.K.: Oxford University Press, 2005.

Goldman, Alvin. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 73 (1976): 771–791.

Hawthorne, John. *Knowledge and Lotteries*. Oxford, U.K.: Oxford University Press, 2003.

Heller, Mark. "Contextualism and Anti-luck Epistemology." In *Epistemology, 1999*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1999.

Heller, Mark. "The Simple Solution to the Problem of Generality." *Noûs* 29 (1995): 501–515.

Henderson, David. "Epistemic Competence and Contextualist Epistemology: Why Contextualism Is Not Just the Poor

Man's Coherentism." *Journal of Philosophy* 91 (1994): 627–649.

Klein, Peter. "Human Knowledge and the Infinite Regress of Reasons." In *Epistemology, 1999*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1999.

Lewis, David. "Elusive Knowledge." *Australasian Journal of Philosophy* 74 (1996): 549–567.

Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.

Ludlow, Peter. "Contextualism and the New Linguistic Turn in Epistemology." In *Contextualism in Philosophy*, edited by Gerhard Preyer and Georg Peter. Oxford, U.K.: Oxford University Press, 2005.

Neta, Ram. "Contextualism and the Problem of the External World." *Philosophy and Phenomenological Research* 66 (2003a): 1–31.

Neta, Ram. "Perceptual Evidence and the New Dogmatism." *Philosophical Studies* 119 (2004): 199–214.

Neta, Ram. "Skepticism, Contextualism, and Semantic Self-Knowledge." *Philosophy and Phenomenological Research* 67 (2003b): 396–411.

Neta, Ram. "*S* Knows That *p*." *Noûs* 36 (2002): 663–681.

Preyer, Gerhard, and Georg Peter, eds. *Contextualism in Philosophy*. Oxford, U.K.: Oxford University Press, 2005.

Rieber, Steven. "Skepticism and Contrastive Explanation." *Noûs* 32 (1998): 189–204.

Rysiew, Patrick. "Contextualist Solutions to Scepticism." *Noûs* 35 (2001): 477–514.

Schaffer, Jonathan. "Contrastive Knowledge." In *Oxford Studies in Epistemology*. Vol. 1, edited by Tamar Gendler and John Hawthorne. Oxford, U.K.: Oxford University Press, 2005.

Schaffer, Jonathan. "From Contextualism to Contrastivism." *Philosophical Studies* 119 (2004): 73–103.

Schiffer, Stephen. "Contextualist Solutions to Scepticism." *Proceedings of the Aristotelian Society* 96 (1996): 317–333.

Stanley, Jason. "Context and Logical Form." *Linguistics and Philosophy* 23 (2000): 391–434.

Stanley, Jason. "On the Linguistic Basis for Contextualism." *Philosophical Studies* 119 (2004): 119–146.

Stine, Gail C. "Skepticism, Relevant Alternatives, and Deductive Closure." *Philosophical Studies* 29 (1976): 249–261.

Tomberlin, James E., ed. *Epistemology, 1988*. Atascadero, CA: Ridgeview, 1988.

Tomberlin, James E., ed. *Epistemology, 1999*. Atascadero, CA: Ridgeview, 1999.

Unger, Peter. "The Cone Model of Knowledge." *Philosophical Topics* 14 (1986): 125–178.

Unger, Peter. *Philosophical Relativity*. Minneapolis: University of Minnesota Press, 1984.

Williams, Michael. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. Oxford, U.K.: Basil Blackwell, 1992.

Ram Neta (2005)

CONTINENTAL PHILOSOPHY

Continental philosophy is a term that arose after the Second World War in English-speaking countries as a name for philosophical approaches that take as their point of departure the work of certain nineteenth and twentieth centuries figures from Continental Europe, especially Germany and France, whose themes and methods were different from those of the analytical philosophy common at most leading British and American philosophy departments at that time. As a general term it includes movements such as phenomenology, existentialism, critical theory, hermeneutics, psychoanalytically oriented philosophy, structuralism, deconstruction, and postmodernism as well as feminist theory, race theories, and other critical social theories to the extent that they draw on one or more of these other movements. Its themes can range across all of the traditional philosophical areas—from epistemology, metaphysics, and ethics to aesthetics, philosophy of language, philosophy of religion, philosophy of science, and studies in the history of philosophy, to name just a few.

General comparisons between the intellectual life in Britain and Continental Europe go back at least to the nineteenth century. However, the current opposition between analytical and continental philosophy can be traced back above all to polemical attempts on the part of leading English analytical philosophers, in particular those at Oxford, to justify their own approach to philosophy and distinguish it from the predominant philosophical movements in France and Germany during the 1950s such as phenomenology and existentialism. Gilbert Ryle (1971) and R. M. Hare (1960), for instance, were outspoken in contrasting the way they and their colleagues approached philosophy from philosophy as practiced in the rest of Europe, with a decidedly negative assessment of continental philosophy's Germanic roots. The primary targets at the time were Husserl, Heidegger, Sartre, and Merleau-Ponty, although it is not clear that their attackers had read many of their works closely.

The term became fairly common, however, only during the late 1960s and early 1970s when it was embraced, first in the United States and then later in Great Britain and other English-speaking countries, as a positive term by philosophers who used it as a name for their own work Continental philosophy was still defined in opposition to analytical philosophy, but without necessarily carrying the negative connotations it had for British and American analytical philosophers. Though discussions about the

differences between analytic versus continental philosophy and the relative merits of each are still common, the term has increasingly become more of a commonly accepted, though still somewhat vague descriptive term used by both proponents and opponents of the different ways of doing philosophy that have been grouped together under the general heading of "continental" (often with a capital "C") in contrast to "analytical" philosophy.

From the outset, even though it suggests a geographical reference, the term "continental philosophy" has referred only to those figures from continental Europe whose approaches were not consistent with those of the project of earlier analytical philosophy. It specifically excluded thinkers such as Frege, Carnap, and Wittgenstein, whose work was viewed in a positive light by the English opponents of continental philosophy, because they were seen rather as precursors to or representatives of analytical philosophy. Moreover, even though the original point of difference goes back primarily to differences in philosophical work being conducted in Britain as opposed to continental European countries after the Second World War, these differences were projected backwards into the history of philosophy. Most observers agree that there were no clearly identifiable differences along geographic lines in philosophy as practiced in England and on the continent before the twentieth century. However, because analytic philosophy as practiced after World War II excluded idealistic philosophy and some other nineteenth-century movements that were originally from Germany or France from its lineage, the term continental philosophy soon came to include not only postwar and earlier twentieth-century philosophical movements from the continent, but also philosophers such as Kant, Hegel, Marx, Kierkegaard, Nietzsche, and others from eighteenth- and nineteenth-century European philosophy whose approaches were not consistent with the historical development of analytical philosophy.

Toward the end of the twentieth century it became common to speak of a "continental tradition" instead of simply referring to "continental philosophy." One reason for this is that many of the adherents and leading practitioners of these directions in philosophy come from countries all around the world instead of just Europe. Another reason is that, from the outset, the fate of this term has been tied to that of its opposite, namely "analytical philosophy," and critical developments within the latter movement led many of its adherents to refer to a "postanalytical" phase that is still part of an analytical tradition of philosophy. As philosophers from both traditions became increasingly familiar with each others' work and many of the original claims in both traditions have had to be modified, it became increasingly difficult to provide a simple characterization of what continental philosophy is and hence also of how it is and is not different from analytical philosophy. Increasingly, some scholars, especially those outside of English-speaking countries, have begun to draw on ideas and resources from both traditions to address philosophical problems across the differences that had previously divided them, so that it makes good sense to think of both continental and analytical philosophy as competing traditions of philosophy and less as two clearly delineated camps situated in different geographic locations.

See also Critical Theory; Deconstruction; Existentialism; Feminist Philosophy; Marxist Philosophy; Phenomenology; Postcolonialism.

Bibliography

Critchley, Simon. "Introduction: What Is Continental Philosophy?" In *A Companion to Continental Philosophy*, edited by Simon Critchley and William R. Schroeder, 1–17. Oxford, U.K.: Blackwell, 1998.

Glendinning, Simon. "Introduction: What is Continental Philosophy?" In *The Edinburgh Encyclopedia of Continental Philosophy*, edited by Simon Glendinning, 3–20. Edinburgh, U.K.: Edinburgh University Press, 1999.

Hare, R.M. "A School for Philosophers." *Ratio* 2 (1960): 107–120.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Ryle, Gilbert. "Phenomenology versus *The Concept of Mind*." In *Collected Papers*. London: Hutchinson, 1971.

Sallis, John, Hugh Silverman, and Thomas M. Seebohm, eds. *Continental Philosophy in America*. Pittsburgh: Duquesne University Press, 1983.

West, David. *An Introduction to Continental Philosophy*. Cambridge, U.K.: Polity, 1966.

Thomas Nenon (2005)

CONTINUITY

In the decades bracketing the turn of the twentieth century, the real number system was dubbed the *arithmetic continuum* because it was held that this number system is completely adequate for the analytic representation of all types of continuous phenomena. In accordance with this view, the *geometric linear continuum* is assumed to be isomorphic with the arithmetic continuum, the axioms of geometry being so selected to ensure this would be the case. In honor of Georg Cantor (1845–1918) and Richard

Dedekind (1831–1916), who first proposed this mathematico-philosophical thesis, the presumed correspondence between the two structures is sometimes called the *Cantor-Dedekind axiom*.

Since their appearance, the late nineteenth-century constructions of the real numbers have undergone set-theoretical and logical refinement, and the systems of rational and integer numbers on which they are based have themselves been given a set-theoretic foundation. During this period the Cantor-Dedekind philosophy of the continuum also emerged as a pillar of standard mathematical philosophy that underlies the standard formulation of analysis, the standard analytic and synthetic theories of the geometrical linear continuum, and the standard axiomatic theories of continuous magnitude more generally.

Since its inception, however, there has never been a time at which the Cantor-Dedekind philosophy has either met with universal acceptance or has been without competitors. The period that has transpired since its emergence as the standard philosophy has been especially fruitful in this regard, having witnessed the rise of a variety of constructivist and predicativist theories of real numbers and corresponding theories of analysis as well as the emergence of a number of alternative theories that make use of infinitesimals. Whereas the constructivist and predicativist theories have their roots in the early twentieth-century debates on the foundations of mathematics and were born from critiques of the Cantor-Dedekind theory, the infinitesimalist theories were intended to either provide intuitively satisfying (and, in some cases, historically rooted) alternatives to the Cantor-Dedekind conception that have the power to meet the needs of analysis or differential geometry, or to situate the Cantor-Dedekind system of real numbers in a grander conception of an arithmetic continuum.

Speculation regarding the nature and structure of continua and of continuous phenomena more generally therefore naturally falls into three periods: the period of the emergence and eventual domination of the Cantor-Dedekind philosophy, and the periods before and after. These three periods are considered in this entry in historical turn.

THE ARISTOTELIAN CONCEPTION

Before the Cantor-Dedekind philosophy the idea of the continuum stood in direct opposition to the discrete and was generally thought to be grounded in our intuition of extensive magnitude, in particular of spatial or temporal magnitude, and of the motion of bodies through space. Some of the essential characteristics of what emerged as the standard ancient conception were already described by Anaxagoras of Clazomenae (c. 500–c. 428 BCE) when he observed that "Neither is there a smallest part of what is small, but there is always a smaller (for it is impossible that what is should cease be [no matter how far it is being subdivided])" (Kirk, Raven, and Schofield 1983, p. 360). Thus, not only is the continuum infinitely divisible, but through the process of division it cannot be reduced to discrete indivisible elements that are, as Anaxagoras picturesquely put it, "separated from one another as if cut off with an axe" (ibid. p. 371). However, while ingredients of the standard ancient conception are already found in the writings of some of the pre-Socratics, it was Aristotle (384–322 BCE), inspired by the writings of the geometers of his day, who provided its earliest systematic philosophical treatment.

Central to Aristotle's analysis is the distinction between discrete and continuous quantity; whereas the former lack, the latter have, a common boundary at which the parts join to form a unity. For Aristotle, number—by which he meant the positive integers greater than or equal to two—is discrete, whereas measurable magnitude—lines, surfaces, bodies, time, and place—are continuous. Lines, in particular, are continuous because "it is possible to find a common boundary at which its parts join together, a point" (*Categories* 6, 5^a1–2, in Aristotle 1984, p. 8); in the cases of surfaces and bodies, the common boundaries are lines, and lines or surfaces, respectively, and in the case of time they are moments.

Motion for Aristotle is also continuous, its continuity being a reflection of spatial and temporal continuity (*Physics* IV.11, 219^a10–13, in Aristotle 1984, p. 371). It is this reflection or isomorphism, for Aristotle, that endows continuous motion with its familiar characteristic properties such as the absence of spatial jumps and the absence of temporal pauses.

Aristotle's characterization of the continuous emerges as the culmination of the following series of definitions he offers in the third chapter of Book V of his *Physics*:

> Things are said to be *together* in place when they are in one primary place. … Things are said to be in *contact* when their extremities are together. … A thing is in *succession* when it is after the beginning in position or in form or in some other respect in which it is definitely so regarded, and when further there is nothing of the same kind as itself between it and that to which it is in succession. … A thing that is in succession and touches [i.e., is in contact] is *con-*

> *tiguous*. … The *continuous* is a subdivision of the contiguous: things are called continuous when the touching limits become one and the same … (*Physics* V.3, 226b22–227a12, in Aristotle 1984, pp. 383–384)

Aristotle maintains that the previous definition implies that

> nothing that is continuous can be composed of indivisibles: e.g. a line cannot be composed of points, the line being continuous and the point indivisible. For the extremities of two points can neither be *one* (since of an indivisible there can be no extremity as distinct from some other part) nor *together* (since that which has no parts can have no extremity, the extremity and the thing of which it is the extremity being distinct). … Moreover, it is plain that everything continuous is divisible into divisibles that are always divisible; for if it were divisible into indivisibles, we should have an indivisible in contact with an indivisible, since the extremities of things that are continuous with one another are one and in contact. (*Physics* VI.1, 231a24–29, 231b15–19, in Aristotle 1984, pp. 390–391)

For Aristotle, the infinite divisibility of the continuous—a property, which on occasion, he appears to take to define the continuous—is a *potential* infinite. Indeed, for Aristotle, the infinite, which is a property of a process rather than of a collection or of a substance, is always potential as opposed to *actual* or completed; that is, no matter which finite stage of the process has been completed, in principle another such stage can be completed. Processes may be infinite with respect to addition or division. Moreover, in the case of spatial continua, in particular, it is the very process of division from which points arise. Thus, while a line segment contains an infinite number of points and an infinite number of parts, for Aristotle it does so only potentially. It is the infinite divisibility of the continuum in this sense that Aristotle appeals to in his treatment of the various paradoxes of Zeno of Elea that are intended to challenge the coherence of the continuity of space, time, matter, and motion. It was also this conception of the continuum that was the dominate conception among philosophers, scientists, and mathematicians alike until the time of Cantor and Dedekind.

NONSTANDARD ANCIENT CONCEPTIONS

However, while Aristotle's theory was the dominate theory until well into the nineteenth century, it never achieved hegemony. Among the ancients, in particular, there were a number of alternative conceptions of continua, including a variety of atomistic conceptions (Furley 1967, Sorabji 1983, White 1992) and the nonatomic conception of the *Stoics* (Sambursky 1959, White 1992). While atomic theories tended to apply solely to the physical realm, there appear to have been atomistic conceptions of geometrical continua as well. Democritus, for example, apparently held that a cone was made up of an infinite number of parallel sections, each of the same indivisible thickness; some who sought to square the circle, including Antiphon, also appear to have embraced atomic theories of geometrical objects. The Stoics, on the other hand, while continuing to adhere to the Aristotelian conception in the mathematical realm, and even to infinite divisibility in the physical realm, may well have distanced themselves from the standard conception in an important respect. According to the interpretation introduced by Shmuel Sambursky (1959) and championed by Michael J. White (1992), the Stoics maintained that there are no points, edges, and surfaces serving as sharp boundaries in physical continua, but rather regions of indeterminacy in which the parts of bodies and adjacent bodies blend. Sambursky (1959, p. 98) likens the physical continuum of the Stoics to the fluid intuitionistic conception of L. E. J. Brouwer, and White proposes instead that "[p]erhaps the best place to look for contemporary elucidation of the Stoic idea is the nonstandard mathematics based on L. A. Zadeh's fuzzy-set theory" (1992, p. 288).

Unlike the Stoics, Aristotle maintains that the physical continuum is a reflection of the geometrical continuum. Indeed, according to Aristotle, "geometry investigates natural [ie. physical] lines, but not qua natural" (*Physics* II.2, 194a9–10, in Aristotle 1984, p. 331). It was this widely held ancient view that the physical mirrors the geometrical that bequeathed to the geometers and their ideas regarding continua an influence far beyond the mathematical domain.

ANCIENT GEOMETRICAL CONCEPTIONS

Aspects of the distinction between discrete number and continuous magnitude are conspicuous in Euclid's *Elements*. Whether Euclid (flourished c. 300 BCE) was directly influenced by the Aristotelian corpus or by earlier geometric practice, however, is the subject of dispute. In any case, reminiscent of Aristotle's characterization, Euclid regards number as a multitude composed of units and believing that one is not itself a number but the unit of number, he appears to identify the numbers with the

positive integers greater than one. Geometrical magnitude, on the contrary, for Euclid, is infinitely divisible. Line segments, in particular, can not only be bisected (Book 1, Proposition 10) *ad infinitum*, they can be divided into *n* congruent segments for each positive integer *n* (Book 6, Proposition 9).

Other ingredients of the Euclidean synthesis that shed important light on the nature of the classical conception of the geometrical continuum are the theories of proportions and incommensurable magnitudes presented in Books 5 and 10, respectively, and the so-called method of exhaustion that is developed in Book 12.

Though arguably the result of an evolutionary process (Knorr 1975, 1978), the theory of proportions developed in Book 5 is usually attributed in its entirety to Eudoxus (c. 400–c. 350 BCE), who, like his contemporaries Plato and Aristotle, lived about half a century before Euclid's *Elements* were compiled. Central to the theory is the concept of a *ratio*:

> A ratio [says Euclid] is a sort of relation in respect of size between two magnitudes of the same kind. … Magnitudes are said to have a ratio to one another which are capable, when multiplied [by a positive integer], of exceeding one another. … Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are equal to, or alike fall short of the latter equimultiples respectively taken in corresponding order. (Book 5, Definitions 3–5 in Heath 1956, Volume 2, p. 114)

While Euclid never contends that two magnitudes of the same kind necessarily have a ratio to one another, his geometry (with the one possible exception of his treatment of horn angles [Book 3, Proposition 16]) is limited to systems of magnitudes for which this is the case. Following Otto Stolz (1842–1905), such systems are said to satisfy the axiom of Archimedes (although it is Eudoxus to whom Archimedes (c. 287–212 BCE) attributes the proposition). In contemporary parlance, if *A* and *B* are members of a given system of magnitudes, *A* is said to be *infinitesimal relative* to *B* if *A* and *B* do not have a ratio to one another and *A* is smaller than *B*. Collaterally, if *A* is infinitesimal relative to *B*, *B* is said to be *infinite relative* to *A*. Thus, in Euclid's geometry, no line segment is either infinitesimal or infinite relative to another segment, and analogous results hold for planer and solid figures as well. Moreover, as in the case of line segments, where there is a well-defined means of subtracting the smaller of two magnitudes of the same kind from the larger, the absence of infinitesimal magnitudes of a given kind precludes the existence, more generally, of magnitudes of a given kind that differ by an infinitesimal amount.

Among the virtues of the theory of proportions of Book 5 is that, unlike the older Pythagorean theory that was based on ratios of integers, it is applicable to both *commensurable* and *incommensurable* magnitudes. Following Euclid, "Those magnitudes are said to be *commensurable* which are measured by the same measure, and those *incommensurable* which cannot have any common measure" (Book 10, Definition 1 in Heath 1956, Volume 3, p. 10). The commensurable-incommensurable dichotomy is as close as the ancients came to the modern dichotomy of rational and irrational numbers, a dichotomy that is central to the Cantor-Dedekind conception of the continuum. The discovery of the existence of incommensurable magnitudes, which is usually attributed to the fifth-century Pythagoreans, was significant because it showed that not every pair of magnitudes of the same kind (straight line segments, rectilinear plane figures bounded by such segments, and so on) has a common measure that divides each an exact integral number of times. Thus, for example, since the side and diagonal of a square are incommensurable, it was not possible (given the ancients' conception of measure) to measure all the sides of even so simple a figure as an isosceles right triangle employing a common unit of measure. These and related discoveries, coupled with their conception of number as a multitude of units, convinced the ancients that it was impossible to bridge the gap between the discrete domain of number and the continuous domain of geometry.

Guided by their intuitions about geometrical continua, the Greeks assumed that simple curvilinear planer figures such as circles, ellipses, and segments of parabolas have areas and perimeters of the same kinds as those of polygons, and they made analogous assumptions about the surface areas and volumes of solids such as spheres, cylinders, and cones. The misleadingly term *exhaustion* was introduced by Gregory of St. Vincent (1584–1667) to describe the method devised by Eudoxus, incorporated into the *Elements* by Euclid, and later extended by Archimedes to measure these and other lengths, areas, and volumes in a rigorous fashion without appealing to either the infinitesimal techniques of the Newtonian and Leibnizian calculi or the passage to the limit concept that has been characteristic of the standard approach to calcu-

lus since its arithmetization during the latter part of the nineteenth century.

As early as 430 BCE Hippocrates of Chios established that the area of a lune (that is, a curvilinear area bounded by circular arcs) of a particular kind is equal to the area of a square. Soon thereafter, Antiphon contended it was possible to obtain a rectilinear figure having the same area as a circle by beginning with an inscribed regular polygon, say a square, and constructing successively more inclusive inscribed regular polygons until the area of the circle was exhausted. Surprisingly, however, he held that the area of the circle would be exhausted after a sufficiently large finite number of steps (perhaps believing that the side of the polygon would coincide with a small arc of the circumference of the circle). Bryson (c. 420 BCE) later developed an alternative account where he considered a circle C sandwiched between a finite series of successively more inclusive inscribed regular polygons, on the one hand, and a finite series of successively less inclusive circumscribed regular polygons, on the other. He maintained that for some positive integer n there is an n-sided regular polygon P, whose area equals the area of C, that properly contains and is properly contained within the aforementioned inscribed and the circumscribed polygons, respectively. To reach his conclusion he appears to have invoked a continuity principle to the effect that a magnitude passes from a smaller to a greater value solely through values of magnitudes of the same kind. The reliance on this principle, which was criticized by Proclus, John Philoponus, and others, was later obviated by the Eudoxean approach.

Central to the exhaustion approach is an alternative continuity principle—the so-called *bisection principle*—that, following Euclid, may be stated as follows:

> Two unequal magnitudes being set out, if from the greater there be subtracted a magnitude greater than or equal to its half, and from that which is left a magnitude greater than or equal to its half, and if this process is repeated continuously, there will be left some magnitude which will be less than the lesser magnitude set out. (Book 10, Proposition 1 in Heath 1956, Volume 3, p. 14)

Using the Archimedean axiom, Euclid proves the bisection principle for the case where the magnitudes subtracted are greater than half the given magnitude, and immediately thereafter observes that "the theorem can be similarly proved even if the parts subtracted be half" (Book 10, Proposition 1 in Heath 1956, Volume 3, p. 15).

The exhaustion method essentially consists of showing that the magnitude (or, more often, ratio of magnitudes) M in question is equal to another magnitude (or ratio of magnitudes) M' one already knows how to determine by showing—using a pair of *reductio ad absurdum* arguments, each of which employs the principle of bisection—that neither $M < M'$ nor $M' < M$. To draw the conclusion that $M' = M$, a tacit appeal is made the presupposition, alluded to earlier, that either $M < M'$, $M' < M$, or $M' = M$. In the version developed by Archimedes, by evoking the geometrical properties of the geometrical object whose perimeter, area, or volume is to be determined, two sequences $I_1, I_2, \dots, I_n$, and $C_1, C_2, \dots, C_n$, consisting respectively of inscribed and circumscribed lines, polygons, or polyhedrons are constructed whose corresponding perimeters, areas, or volumes are such that $I_n < M < C_n$ and $I_n < M' < C_n$ for all n. Using the principle of bisection, it is then either shown that, given $\varepsilon > 0$, $C_n - I_n < \varepsilon$ for n sufficiently large or that, given $\alpha > 1$, $C_n/I_n < \alpha$ for n sufficiently large.

Archimedes's version of the method of exhaustion resembles and, to some extent, inspired the technique of integration later employed in the calculus. Before the development of the calculus, however, a variety of the concepts and techniques inherited from the ancient geometers would undergo marked change. Of these perhaps none has had a more profound or lasting impact on theories of continua than the rethinking of the number concept and its relation to the geometrical continuum.

EARLY MODERN THEORY OF REAL NUMBERS

The early modern theory of real numbers began to emerge when mathematicians such as Simon Stevin (1548–1620) argued that not only is 1 also a number but that there is also a complete correspondence between (positive) number and continuous magnitude, as well as a parallelism between certain geometrical constructions and the now familiar arithmetic operations on numbers. In a number of his works, including his influential *L'Arithmétique* (1585), Stevin expresses the matter thus:

> I consider the relation between number and magnitude to be such that what can be done by the one can be done by the other. …
>
> … *these two quantities cannot be distinguished by continuity or discontinuity*.
>
> To a continuous magnitude corresponds the continuous number to which it is attributed. (Strong 1976, p. 105; Klein 1968, p. 195; Stevin 1585/1958, p. 501)

This viewpoint soon led to, and was implicit in, the analytic geometry of René Descartes (1596–1650), and was made explicit by John Wallis (1616–1703) and Isaac Newton (1643–1727) in their arithmetizations thereof. Influenced by work of Wallis, Isaac Barrow (1630–1677), and others, the (positive) numbers came to be associated with Eudoxian-Euclidean ratios that were assumed to exist between the magnitudes of a given kind and a selected unit magnitude of the same kind (compare Klein 1968, Pycior 1997). In accordance with the Eudoxian-Euclidean framework, no two such magnitudes of the same kind could differ by an infinitesimal amount. In his popular and influential *Arithmetica Universalis* (1707/1728), Newton extends the correspondence between numbers and ratios to include negative numbers and zero, but whereas some of his predecessors had identified the positive numbers with the symbolic representations of the ratios, Newton identifies numbers with the abstracted ratios themselves. Emphasizing his sharp break with the ancient conception of number, says Newton:

> By *Number* we understand not so much a Multitude of Unities, as the abstracted Ratio of any Quantity, to another Quantity of the same Kind, which we take for Unity. And this is threefold; integer, fracted, and surd: An *Integer* is what is measured by Unity, a *Fraction*, that which a submultiple Part or Unity measures, and a Surd, to which Unity is incommensurable. (Newton 1728/1967, p. 2)

That zero could not be a number in accordance with this definition did not preclude Newton from asserting it was, and the careful treatment that late nineteenth-century mathematicians would recognize to be required to handle ratios involving negative quantities is nowhere to be found.

THE CALCULUS OF NEWTON AND LEIBNIZ

During the sixteenth century the works of Archimedes were widely studied by Western mathematicians and served as the chief source of inspiration for the seventeenth-century development of the infinitesimal calculus, the branch of mathematics erected by Newton and Gottfried Wilhelm Leibniz (1646–1716) for the study of continuously varying magnitudes or quantities. The conception of the continuum embraced by most mathematicians of the period was geometrical or kinematical by nature and grounded in intuition. It was commonplace to consider a curve as a path of a moving point, the curve being continuous insofar as motion itself was presumed to be continuous. Moreover, perhaps as a result of the medieval speculations on the infinite and the continuum, the mathematicians of the day, unlike their mainstream Greek counterparts, were not adverse to employing infinitesimal techniques and appeals to the actual infinite in these and related works. Some authors, such as Galilei Galileo (1564–1642), following in the footsteps of such fourteenth-century thinkers as Henry of Harclay, Nicholas Bonet, Gerard of Odo, Nicholas of Autrecourt, and John Wyclif (Murdoch 1982), maintained that line segments, surfaces, and solids are made up of an actual infinite number of indivisible or infinitesimal elements. And similar ideas were employed by Johannes Kepler (1571–1630), Bonaventura Cavalieri (1598–1647), and others in their determinations of areas and volumes and by Barrow in his determinations of tangents to curves, determinations that would be the focus of the unifying algorithmic frameworks that would come to be called the calculus.

Following in the footsteps of their just-cited forerunners, infinitesimal techniques were employed by Newton and Leibniz in their treatments of the calculus, but unlike some of their predecessors neither of them attributed ontological status to either the actual infinite or the actual infinitesimal. Both regarded infinitesimals—or incomparables as Leibniz sometimes called them—as varying quantities in a state of approaching zero that serve as useful fictions to abbreviate their mathematical proofs. The abbreviated proofs in turn, they contended, could be replaced by limit-based proofs the latter of which not only constitute the rigorous formulation of calculus but are a direct version of the indirect method of exhaustion due to Archimedes. Newton and Leibniz also agreed that the justification for the limit-based proofs lay in the concept of continuity but they differed on the justification itself. Whereas Newton sought it in terms of one's intuition of continuous flow, Leibniz sought it in terms of his *law of continuity*, a philosophical principle to the effect that "[n]o transition is made through a leap" or that "nothing takes place suddenly." The natures of their respective attempts at justification, however, only begin to intimate that the limit-based proofs envisioned by Newton and Leibniz, while akin to, are by no means identical with the limit-based proofs that emerged during the nineteenth century.

Unlike the calculus of today, the calculi of Newton and Leibniz were not concerned with functions but with variable quantities, their rates of change, and so on. However, whereas Newton regarded these quantities as varying at finite rates with respect to time, Leibniz envisioned

them as ranging over discrete sequences of values that successively differ by infinitesimal amounts. Underlying this difference was a difference in fundamental concepts: For Newton it was the *fluxion* or finite rate of change of the variable with respect to time, and for Leibniz it was the just-cited infinitely small differences or the differential. As a result, in Leibniz's treatment of the calculus the limit concept was suppressed or at least disguised, whereas it was explicit in Newton's formulation. In the case of differentiation, for example, since for Leibniz it is the distinct differentials dx and dy that are fundamental, their ratio dy/dx is of principal significance, whereas for Newton, especially in his later treatment, it is the derivative itself—as a ratio of fluxions or an ultimate ratio of evanescent quantities—that is of central importance.

That there were foundational difficulties with the science of continuously varying quantities was well known among seventeenth- and eighteenth-century mathematicians including Newton and Leibniz themselves. These difficulties were brought into sharp focus by George Berkeley (1685–1753) in his stinging critique of the logical and ontological foundations of the calculus titled *The Analyst* (1734/1992). According to Berkeley, there is no justification for attributing existence to either limits or infinitesimals: A limit of a ratio is either a limit of finite quantities, and therefore not an ultimate ratio, as Newton contended, or it is a mysterious indeterminate ratio 0/0; and if the infinitesimal quantities dx and dy are not equal to zero, one has the problem of explaining how it is possible that $x + dx = x$, and if they are equal to 0, once again one has the problem of explaining the meaning of the indeterminate 0/0. It was with these and related quandaries that mathematicians concerned with the study of continuously varying magnitude grappled well into the nineteenth century. Some eighteenth-century mathematicians, such as Colin Maclaurin (1698–1746) and Jean Le Rond d'Alembert (1717–1783), attempted to address the foundational worries with refinements of the limit approach of Newton, but it was the infinitesimal approach of Leibniz that emerged as the dominant approach of the day. Moreover, the remaining puzzlement over infinitesimals no longer applied solely to the fictional infinitesimals of Leibniz, but to the actual infinitely large and the actual infinitely small numbers and magnitudes employed to great effect throughout the eighteenth and early nineteenth century by a host of distinguished analysts working in the Leibnizian tradition including Jakob Bernoulli (1654–1705), Johann Bernoulli (1667–1748), Daniel Bernoulli (1700–1782), and Leonhard Euler (1707–1783), to name only a few.

THE ARITHMETIZATION OF ANALYSIS

During the nineteenth century, building on the work begun in 1821 by Augustin-Louis Cauchy (1789–1857) in his *Cours d'analyse*, the calculus was given a rigorous foundation that is still accepted today. By the middle of the century, developments in subject persuaded many mathematicians that the traditional concepts of the calculus were too imprecise, unreliable, and ineffective to provide such a basis. It was held that the traditional relation between real quantities and intuitively given continuous magnitudes such as straight lines was more of a hindrance than an aid in achieving that end as was the then familiar reliance on infinitesimals. In response, the modern *arithmetico-set-theoretic* conception of a real number emerged when a number of mathematicians including Cantor (1872/1939) and Dedekind (1872/1996) introduced systems of real numbers that were designed to dispense with the former and provide a basis for the calculus that made superfluous the latter.

Cantor's system is based on *Cauchy sequences* of rational numbers. A sequence $\{r_n\}$ of rational numbers (indexed over the natural numbers) is said to be a Cauchy sequence (or *fundamental sequence*, as Cantor called it) if for every rational number $\varepsilon > 0$ there is a natural number k such that $|r_m - r_n| < \varepsilon$ for all m, $n > k$. Two such sequences $\{r_n\}$ and $\{s_n\}$ are said to be equivalent if for every rational number $\varepsilon > 0$ there is a natural number k such that $|r_n - s_n| < \varepsilon$ for all $n > k$. In modern parlance, Cantor's construction amounts to identifying the set $\mathbb{R}$ of real numbers with the set of all equivalence classes of Cauchy sequences of rational numbers. If an equivalence class contains the Cauchy sequence $\{r_n\}$ where $r_n = r$ for all n, the equivalence class corresponds to the rational number r, otherwise it corresponds to an irrational number. Each irrational number is associated with the equivalence class containing the Cauchy sequence consisting of the initial segments of its unique nonperiodic decimal representation. For example, $\sqrt{2}$ is associated with the equivalence class containing the Cauchy sequence

$$r_0 = 1; r_1 = 1.4; r_2 = 1.41; r_3\ 1.414; \ldots$$

If α and β are real numbers represented by the Cauchy sequences $\{r_n\}$ and $\{s_n\}$, then the sum and product of α and β are represented by the Cauchy sequences $\{r_n + s_n\}$ and $\{r_n s_n\}$, respectively; and it is said that $\alpha > \beta$ if there is an $a > 0$ such that $r_n \geq s_n + a$ for sufficiently large n.

Dedekind's system, by contrast, is based on *cuts* of the ordered set $\mathbb{Q}$ of rational numbers. By a cut (A_1, A_2) of $\mathbb{Q}$ Dedekind means a partition of $\mathbb{Q}$ into two nonempty sets A_1 and A_2 in which every member of A_1 pre-

cedes every member of A_2. Dedekind identities ℝ with the set of all cuts of ℚ; if A_1 has a greatest member or A_2 has a least member, say, r, the cut (A_1, A_2) is associated with the rational number r; otherwise it defines an irrational number. $\sqrt{2}$, in particular, is defined by the cut

$$(\{a \in \mathbb{Q}: a < 0 \vee a^2 \leq 2\}, \{a \in \mathbb{Q}: a^2 > 2\}).$$

Given two real numbers $\alpha = (A_1, A_2)$ and $\beta = (B_1, B_2)$, Dedekind stipulates that $\alpha < \beta$ if A_1 is a proper subset of A_2. He also defines $\alpha + \beta = (C_1, C_2)$ where $C_2 = \mathbb{Q} - C_1$ and C_1 is the set of all $c \in \mathbb{Q}$ such that $c \leq a_1 + b_1$ for some $a_1 \in A_1$ and some $b_1 \in B_1$, and further observes that the remaining familiar arithmetic operations on real numbers can likewise be defined.

On the basis of Cantor's and Dedekind's systems of real numbers, whose equivalence (modulo the then tacitly emerging underlying set-theoretic assumptions) would soon be established, the classical concepts of the calculus, including Cauchy's and Bernhard Riemann's (1826–1866) modern definitions of the derivative and the integral, were reformulated in a rigorous fashion using the now familiar δ, ε-techniques, as were the concepts of convergence, sum of an infinite series, and continuity to name only a few. Central to this development was Karl Weierstrass's (1815–1897) replacement of Cauchy's dynamic conception of the limit concept, together with its Newtonian connotations of continuous motion, with a static purely arithmetical formulation. Instead of setting $\lim_{x\to\xi} f(x) = L$ provided that $f(x)$ approaches L as x approaches ξ, Weierstrass set $\lim_{x\to\xi} f(x) = L$ provided that, given a positive real number ε, there is a positive real number δ such that $|f(x) - L| < \varepsilon$ if $0 < |x - \xi| < \delta$. With the host of limit dependent concepts so reformulated, the calculus assumed the form that one still finds in the standard textbooks of the early twenty-first century.

CONTINUOUS FUNCTIONS

As was already noted, the calculus of Newton and Leibniz was not a calculus of functions. It was Euler in the middle of the eighteenth century who placed the concept of function and, in particular, the concept of continuous function at the center of analysis, and it was Cauchy in 1821, and independently Bernard Bolzano (1781–1848) in 1817, who gave the concept its now standard meaning. Following these authors, a function $f(x)$, defined in a neighborhood of a point ξ, is said to be *continuous at* ξ if $\lim_{x\to\xi} f(x) = f(\xi)$; and $f(x)$ is said to be *continuous in a closed interval* $[a, b]$ if it is continuous at each point of the interval, it being understood that the limits corresponding to the endpoints a and b are right-sided and left-sided limits, respectively. Thus, within the Weierstrassian framework, $f(x)$ is continuous in a closed interval iff for each member ξ of the interval and for each positive real number ε there is a positive real number δ such that $|f(x) - f(\xi)| < \varepsilon$ whenever $|x - \xi| < \delta$.

The Cauchy-Bolzano conception of continuity accords nicely with the intuition that the values of a continuous function f differ slightly when its arguments differ slightly and, hence, with its geometric analog that the graph of f does not have a *break* or *jump* in the interval in question. Indeed, using the Cauchy-Bolzano definition nineteenth-century mathematicians were able to show that formal replacements of a number of the familiar intuitions about continuous functions and curves could be established as theorems including the following two:

> The Intermediate Value Theorem. If f is a continuous function on a closed interval $[a, b]$ of ℝ and $f(a) < \xi < f(b)$ for some ξ, then there is a $c \in (a, b)$ with $f(c) = \xi$;
>
> Extreme Value Theorem. If f is a continuous function on a closed interval $[a, b]$ of ℝ, then f has a maximum value at some $c \in [a, b]$ and f has a minimum value at some $d \in [a, b]$.

The Cauchy-Bolzano conception of continuity is local by nature, referring to the behavior of the function in the neighborhood of a point. Even the notion of continuity in a closed interval is defined in terms of continuity at every point of the interval. A more global conception of continuity that gradually emerged during the process of rigorization that implies but is not implied by the Cauchy-Bolzano conception is that of *uniform continuity*. Following Heinrich Heine (1821–1881), who carefully separated the two notions that had apparently been conflated by Cauchy, a function f is said to be uniformly continuous on a set A of real numbers if, for each positive real number ε, there is a positive real number δ such that for each pair of members ξ and ξ' of A, $|f(\xi) - f(\xi')| < \varepsilon$ whenever $|\xi - \xi'| < \delta$. Essentially, this asserts that for a given ε, the same δ for the continuity condition works for all members of A. The following result, which is of central importance in both standard analysis and a number of the nonstandard alternatives that will be discussed later on, is also due to Heine: f is uniformly continuous on a closed interval $[a, b]$ of ℝ, whenever f is continuous on $[a, b]$.

It is important to emphasize that the class of functions that are deemed to be continuous by standard analysts are more inclusive than those envisioned implicitly or explicitly by their seventeenth-, eighteenth-, and even

early nineteenth-century predecessors. Basically, all the curves treated by seventeenth-century analysts were expressed everywhere by one and the same algebraic or transcendental equation and were, accordingly, continuous in the now standard sense. In the mid-eighteenth century so-called discontinuous functions were introduced into analysis by Euler, though they would not be recognized as such today. According to Euler's distinction, which was used up to the time it was replaced by that of Bolzano and Cauchy, a function is continuous if it is characterized by a single analytic expression, and it is discontinuous if it lacks any analytic expression, as in the case of freehand curves, or, if it is defined by different analytic expressions in a finite number of different intervals, the points at which the analytical expressions change being the points of discontinuity.

Euler's points of discontinuity correspond to points of the curve having no well-defined derivative. Accordingly, if one thinks of a continuous curve as the path of a moving point—an intuition that played an important heuristic role in the development of the calculus—Euler's points of discontinuity correspond to points at which the moving point has no well-defined direction. With this in mind it is not difficult to understand why during much of the nineteenth century it was widely believed that functions that are continuous in the Cauchy-Bolzano sense may fail to have derivatives at no more than a finite number of points. In fact, a number of mathematicians including Bolzano himself attempted to prove just this. Mathematicians were therefore surprised when in 1861 Weierstrass provided an example of a continuous function that is nowhere differentiable. A similar such function was constructed by Bolzano in 1834, but like the remainder of his work, it did not become known to the mathematical community till after the work of Weierstrass.

The Cauchy-Bolzano conception of discontinuity, by contrast with Euler's, is closer to that of *discontiguity*, an extreme case being P. G. Lejeune Dirichlet's (1805–1859) so-called monster function—the nowhere continuous function defined on the real line by the condition

$$f(x) = \begin{cases} 0, \text{ if } x \text{ is rational} \\ 1, \text{ if } x \text{ is irrational.} \end{cases}$$

Euler apparently was aware of the existence of discontinuous functions in the modern sense (Youschkevitch 1976, pp. 64–65), but they did not play a fundamental role in the calculus of his time. With the work of Riemann on the convergence of Fourier series during the middle of the nineteenth century, however, this all changed, and they have come to enjoy widespread application not only in analysis, but also in empirical science from where they came. Indeed, referring to their early nineteenth-century roots in the work of Joseph Fourier, the great philosopher-mathematician Jules Henri Poincaré (1854–1912) musingly observes:

> Fourier's series is a precious instrument of which analysis makes continual use, it is by this means that it has been able to represent discontinuous functions; Fourier invented it to solve a problem of physics relative to the propagation of heat. If this problem had not come up naturally, we should never have dared to give discontinuity its rights; we should still long have regarded continuous functions as the only true functions. (Poincaré 1913, p. 286)

THE CANTOR-DEDEKIND CONTINUUM

Central to Cantor's and Dedekind's constructions of the real number system was the underlying belief that only after providing a precise definition of a continuum based on the science of number would it be possible to lend precision to the idea of the continuity of the Euclidean straight line and of continuous magnitude more generally. For this purpose they turned to properties of $\mathbb{R}$ whose continuity they assumed as a *mathematico-philosophical thesis*. According to Cantor, the continuity of $\mathbb{R}$ consists in its being both *connected* and *Cauchy complete*; it is connected because whenever a and a' are elements of the system such that $a < a'$, then for any positive element ε of the system, there is a finite sequence $a_1, \dots, a_n$ of elements of the system where $a < a_1 < \dots < a_n < a'$ such that $a_1 - a, \dots, a' - a_n < \varepsilon$; and it is Cauchy complete since every convergent sequence of elements of the system has a limit in the system. Dedekind, by contrast, identified the continuity of $\mathbb{R}$ with its being a totally ordered system having what is today called the Dedekind continuity property; that is, whenever the system is partitioned into two nonempty subsets X and Y such that every member of X precedes every member of Y, then either X has a greatest member or Y has a least member, but not both.

Connectivity, in Cantor's sense, was soon recognized to be equivalent the Archimedean axiom for a large class of structures including ordered fields, the latter of whose roots lie in analytic geometry. Indeed, since the time that Wallis and Newton incorporated directed segments into Cartesian geometry, it had been loosely understood that given a unit segment AB of a line L of a classical Euclid-

ean space, the collection of directed segments of L emanating from A including the degenerate segment AA itself constitutes an Archimedean ordered field with AA and AB the additive and multiplicative identities of the field and addition and multiplication of segments suitably defined. These ideas were made precise by Giuseppe Veronese (1854–1917) in his *Fondamenti di geometria* (1891) and by David Hilbert (1862–1943) in his *Grundlagen der Geometrie* (1899), works on the foundations of geometry from which the modern conceptions of Archimedean and non-Archimedean ordered fields emerged. It was also these and related works on the foundations of geometry that gave rise to the following familiar characterizations of the arithmetic continuum whose continuity properties are associated with Cantor, Dedekind, Bolzano, and Hilbert, respectively:

(1) $\mathbb{R}$ is (up to isomorphism) the unique Archimedean ordered field that is Cauchy complete

(2) $\mathbb{R}$ is (up to isomorphism) the unique Dedekind continuous ordered field

(3) $\mathbb{R}$ is (up to isomorphism) the unique ordered field having the least upper bound property (that is, every subset of the system that is bounded above has a least upper bound)

(4) $\mathbb{R}$ is (up to isomorphism) the unique Archimedean ordered field that admits no proper extension to an Archimedean ordered field

Each of these characterizations of $\mathbb{R}$ makes use of metrical conceptions. However, in 1895 Cantor demonstrated that it is possible to provide a categorical characterization of the ordered set of real numbers and, hence, of a Cantor-Dedekind linear continuum, using order-theoretic concepts alone. Another such characterization that emerged soon thereafter is the following one given by Edward V. Huntington (1874–1952) that indicates what besides simple density—the set-theoretic analog of infinite divisibility—is required to characterize the order type of $\mathbb{R}$ (1917):

(1) $\langle K, < \rangle$ is a totally ordered set having neither a first element nor a last element;

(2) $\langle K, < \rangle$ is dense, that is, if a and b are elements of K such that $a < b$ there is an element x in K such that $a < x < b$;

(3) $\langle K, < \rangle$ is Dedekind complete, that is, if K_1 and K_2 are any two nonempty subsets of K, such that every element belongs either to K_1 or K_2 and every element of K_1 precedes every element of K_2, then there is at least one element x in K such that any element that precedes x belongs to K_1, and every element that follows x belongs to K_2;

(4) the class K contains a denumerable subset K' in such a way that between any two elements of K there is an element of K'.

Reflecting on the then newly developed order-theoretic conception of the mathematical continuum, the great philosopher-mathematician Jules Henri Poincaré perceptively remarks that:

> [t]he continuum so conceived is only a collection of individuals ranged in a certain order, infinite in number, it is true, but *exterior* to one another. This is not the ordinary conception, wherein is supposed between the elements of a continuum a sort of intimate bond which makes of them a whole, where the point does not exist before the line, but the line before the point. Of the celebrated formula, "the continuum is unity in multiplicity," only the multiplicity remains, the unity has disappeared. The analysts are none the less right in defining their continuum as they do, for they always reason on just this as soon as they pique themselves on rigor. But this is enough to apprise us that the veritable mathematical continuum is a very different thing from that of the physicist and that of the metaphysician. (1913, p. 43)

To some extent these views are a reflection of those of Cantor and Dedekind themselves. For example, distancing himself from a long line of metaphysicians, Cantor writes:

> The concept of the "continuum" has … always evoked the greatest differences of opinion and even vehement quarrels. This lies perhaps in the fact that, because the exact and complete definition of the concept has not been bequeathed to the dissentients, the underlying idea has taken on different meanings; but it may also be (and this seems to me the most probable) that the idea of the continuum had not been thought out by the Greeks (who may have been the first to conceive it) with the clarity and completeness which would have been required to exclude the possibility of different opinions among their posterity. Thus we see Leucippus, Democritus, and Aristotle consider the continuum as a composite which consists from parts divisible without end, but Epicurus and Lucretius construct it out of their atoms considered as finite things. Out of this a great quarrel arose among the

> philosophers, of whom some followed Aristotle, others Epicurus; still others, in order to remain aloof from this quarrel, declared with Thomas Aquinas that the continuum consisted neither of infinitely many nor of a finite Anzahl [number] of parts, but of *absolutely no* parts. … Here we see the *medieval-scholastic origin* of a point of view which we still find represented today, in which the continuum is thought to be an unanalysable concept, or as others express themselves, a pure a priori intuition which is scarcely susceptible to a determination through concepts. Every arithmetical attempt at determination of this *mysterium* is looked on as a forbidden encroachment and repulsed with due vigor. Timid natures thereby get the impression that with the "continuum" it is not a matter of a *mathematically logical concept* but rather of religious dogma. (1883/1996, p. 903)

Moreover, as to the necessity of even conceiving space as continuous, Dedekind remarks, "If space has at all a real existence it is not necessary for it to be continuous" (1872/1996, p. 772). Indeed,

> If anyone should say that we cannot conceive of space as anything else than continuous, I should venture to doubt it and call attention to the fact that a far advanced, refined scientific training is demanded in order to perceive clearly the essence of continuity and to understand that besides rational quantitative relations, also irrational, and besides algebraic, also transcendental quantitative relations are conceivable. (1888/1996, p. 793)

Bertrand Russell, who played the leading role in introducing the ideas of Cantor and Dedekind to the English-speaking philosophical world, goes even further when he remarks, "Whether the axiom of continuity be true as regards our actual space, is a question I see no means of deciding. For any such question must be empirical, and it would be quite impossible to distinguish empirically what may be called a rational space from what might be called a continuous space" (1903, p. 440).

However, despite these and other such pronouncements made in the years bracketing the turn of the twentieth century, and despite the ongoing speculation about quantized space and time that emerged soon thereafter with the advent of the quantum theory (compare Cepek 1961, pp. 223–243; Sorabji 1983, pp. 381–383, 447) and that was redirected toward overcoming the difficulties of harmonizing the quantum theory with the relativistic theory of gravity and space-time (compare Markopoulou 2004; Smolin 2001, 2004), it became and remains commonplace among philosophers and physicists to assume not only that space and time are continuous, as most of their modern predecessors had supposed, but also that they are continuous in the sense of Cantor and Dedekind. Whether this assumption should be construed instrumentally or realistically there is a multiplicity of views (compare Earman 1989, chapters 8 and 9; Maddy 1997, chapter 6; Hellman 1998), as are the views regarding the testable status of the thesis itself (compare Forrest 1995, Markopoulou 2004).

MODERN EUCLIDEAN GEOMETRY AND THE CONTINUUM

At least as far back as the seventeenth century there were thinkers who observed that there are places in the *Elements* where Euclid tacitly employs continuity assumptions that are not warranted by the axioms and common notions he assumes. For example, in his proof that given any segment, there is an equilateral triangle having the given segment as one of its sides (Book 1, Proposition 1), Euclid assumes

> *The Circular Continuity Principle*: If a circle C has one point inside and one point outside another circle C', then the two circles intersect in two points.

And in his proof that through a point outside a given line there is a line perpendicular to the given line (Book 1, Proposition 12), he assumes

> *The Line-Circle Continuity Principle*: If one endpoint of a segment is inside a circle and the other outside, then the segment intersects the circle at one point.

Among the thinkers who thus criticized Euclid was Leibniz in his *Specimen geometriae luciferae* (c. 1695/1962, p. 284). Such criticisms were significant not only because they drew attention to gaps in Euclid's reasoning but also because they intimated that, contrary to the ancient and the then still standard view, infinite divisibility is not sufficient for continuity. In fact, in the just-cited paper as Ernst Cassirer (1902, p. 183) importantly observes, Leibniz departed from his usual acceptance of the standard view and explicitly stated just this.

During the late nineteenth century geometers showed that by supplementing the then newly emerging refinements of Euclid's system of axioms with continuity axioms that ensured the satisfaction of the Cantor-Dedekind axiom, one could establish the circular continuity principle and the line-circle continuity principle

and with these all of Euclid's continuity needs (compare Heath 1956, pp. 1:234–240; Greenberg 1993, pp. 93–101). However, as Cantor and Dedekind were aware, whereas the Cantor-Dedekind axiom suffices for the continuity needs of Euclid, it goes beyond those needs. Cantor makes this point, albeit only implicitly, when (following his proof-sketch that to each point of the Euclidean line there corresponds a real number) he maintains:

> In order to complete the connection of numerical quantity with the geometry of the straight line, one must only add an *axiom* which simply says that conversely every numerical quantity also has a determined point on the straight line, whose coordinate is equal to that quantity. … I call this proposition an *axiom* because by its nature it cannot be generally proved. (Cantor 1872/1932, p. 96)

And Dedekind makes the point more forcefully and explicitly when he revealingly observes:

> If we select three non-collinear points *A*, *B*, and *C* at pleasure, with the single limitation that the ratios of the distances *AB*, *AC*, *BC* are algebraic numbers, and regard as existing in space only those points *M*, for which the ratios of *AM*, *BM*, *CM* to *AB* are likewise algebraic numbers, then it is easy to see that the space made up of the points *M* is everywhere discontinuous. But in spite of this discontinuity, and despite the existence of gaps in this space, all constructions that occur in Euclid's *Elements*, can, so far as I can see, be just as accurately effected here as in a perfectly continuous space; the discontinuity of this space would thus not be noticed in Euclid's science, would not be felt at all. (1888/1996, p. 793)

The ordered field of real algebraic numbers to which Dedekind is referring is an instance of a *Euclidean ordered field*, that is, an ordered field in which every positive element is the square of some element of the field. Besides being Euclidean, the ordered field of real algebraic numbers is both countable and Archimedean. During the twentieth century it emerged that a model of all the axioms of (a modern refined version of) Euclidean geometry less the Archimedean axiom and the Cantor-Dedekind axiom satisfies the circular continuity principle iff it satisfies the line-circle continuity principle iff a line of the model is modeled by a Euclidean ordered field (compare Tarski 1959/1986; Hartshorne 2000, pp. 104–112, 144–145; Greenberg 1993, pp. 143–144). If, following Euclid, the Archimedean axiom is also assumed, the Euclidean ordered fields must be Archimedean as well. It is essentially for this reason that Euclidean ordered fields have been so named. For historically important examples of modern refined versions of Euclidean geometry, see David Hilbert (1899/1971) and Alfred Tarski (1959/1986); and for examples of non-Archimedean Euclidean ordered fields and their corresponding non-Archimedean Euclidean geometries, see Philip Ehrlich (1997a).

SET THEORY AND THE CONTINUUM

The Cantor-Dedekind theory of the continuum was originally formulated in a *naïve* set-theoretic framework, grounded in intuitions about sets that included the then radical assumption that infinitely many entities could be collected together in a set. Within this framework Cantor established the existence of an exhaustive hierarchy $\aleph_0$, $\aleph_1$, $\aleph_2$, … of increasingly large infinite cardinals, proved that the cardinality of the set of rational numbers is $\aleph_0$, that the cardinality of $\mathbb{R}$ is $2^{\aleph_0}$, and that $2^{\aleph_0}$ is greater than $\aleph_0$. In the early decades of the twentieth century Cantorian set theory was placed on an axiomatic basis that sidestepped a medley of paradoxes that had befallen the naïve theory. In honor of two of its principal architects, Ernst Zermelo (1871–1953) and Abraham A. Fraenkel (1891–1965), the theory has come to be designated ZFC, where ZF indicates the body of axioms outside the axiom of choice. Of the open problems inherited from the naïve theory none was regarded more important than that of deciding the veracity of the *continuum hypothesis* (CH)—Cantor's conjecture that the cardinality of the continuum is $\aleph_1$. However, in 1938 Kurt Gödel showed that ZFC + CH is consistent if ZFC is, and twenty-five years later Paul Cohen demonstrated if ZFC is consistent so is ZFC + ¬CH, thereby establishing the independence of CH from standard set theory. Since the work of Cohen there has been a good deal of speculation on the part of philosophers and logicians whether or not the axioms of ZFC should be supplemented with one or more additional axioms that would settle the matter one way or another (Kanamori 2003). Gödel (1947/1983), who held a platonist view of sets, maintained that CH is either objectively true or objectively false and, accordingly, promoted the search for additional axioms to settle the matter. Many set theorists, including Cohen (1990), however, believe that there is nothing in the intuitive concept of set that would recommend the adoption of an additional axiom that would conclusively settle CH one way or another. The views of Gödel and Cohen, however, only begin to indicate the range of opinion on the matter as is evident from the debate between Soloman Feferman et al. (2000), the writings of Donald Martin and H. G. Dales (Dales and

Oliveri 1998), and the intriguing though controversial views of W. Hugh Woodin (2001, 2002, 2004). Both models of ZFC + CH and ZFC + ¬CH are being explored by set theorists with perhaps a bit more attention being devoted to the latter. Mathematicians who are not set theorists, however, tend to use CH freely for the purpose of theorem proving, their reliance on CH being indicated in the statement of the theorem.

NONSTANDARD THEORIES OF CONTINUA

Although the Cantor-Dedekind theory of real numbers and philosophy of the continuum have occupied privileged positions in standard mathematical philosophy since the decades following the turn of the twentieth century, it has never enjoyed the complete allegiance of either philosophers or mathematicians. Early opponents such as Hermann Hankel (1839–1873) and Paul du Bois-Reymond (1831–1889) were critical of the attempts by Cantor, Dedekind, and others to treat irrational numbers formally and without the concept of continuous magnitude, and others such as Leopold Kronecker (1823–1891) complained, on the contrary, that the arithmetization had not gone far enough. Still others, including Emil Borel (1871–1956) and a young L. E. J. Brouwer (1881–1966), continued to regard the continuum as a primitive concept given to one directly by geometric intuition that was not amenable to analysis (compare Troelstra 1982); and others, including Hermann Weyl (1885–1955) (1918) as well as a more mature Brouwer (compare 1918, 1924, 1952), while embracing an analytical approach questioned one or another aspect of the logico-set-theoretic underpinnings of the Cantor-Dedekind theory. Another complaint that was, and to some extent still is, a stumbling block to the acceptance of the Cantor-Dedekind theory is the contention that the Cantor-Dedekind philosophy of the continuum is committed to the reduction of the continuous to the discrete, a program whose philosophical cogency, and even logical consistency, had been called into question over the centuries. Inspired by arguments originating with Aristotle, and reiterated and further developed by Immanuel Kant (1724–1804), Franz Brentano (1838–1917), Edmund Husserl (1859–1938), and others, a string of late nineteenth- and twentieth-century mathematicians and philosophers beginning with du Bois-Reymond (1882) maintained that unlike the unextended points that, by their lights, compose the Cantor-Dedekind continuum, the elements of a genuine continuum must themselves have extension if the continuum itself is to have extension. This view led Charles Sanders Peirce (1939–1914) to sketch a nonarithmetic theory of the continuum based on infinitesimals (1898/1992, 1900), and it played important contributing roles in the development of Brouwer's and Weyl's aforementioned intuitionist and predicativist alternatives to the standard conception as well.

Veronese (1889, 1891, 1894), however, while agreeing that the parts of a continuum must be intervals as opposed to points, held that for the sake of geometry the geometer may treat the line as an ordered collection of unextended points; moreover, holding that one's intuitive conception of the continuum is independent of the Archimedean axiom, he developed a general axiomatic theory of continua that was not only satisfied by the standard arithmetic continuum but by certain structures with infinitesimals as well, and he illustrated the latter by means of a synthetic construction of a non-Archimedean ordered field that is continuous in his sense. Veronese's non-Archimedean continuum was placed on a logically sound arithmetic foundation by Tullio Levi-Civita (1873–1941), who therewith provided the first analytic constructions of non-Archimedean ordered fields (1892–1993/1954, 1898/1954).

Building on the work of Levi-Civita, Hans Hahn (1879–1934) constructed non-Archimedean number systems (1907) having properties that generalize the aforementioned continuity properties of Dedekind and Hilbert, and he demonstrated that his number systems collectively provide a panorama of the finite, infinite, and infinitesimal numbers that can enter into a non-Archimedean theory of continua based on the concept of an ordered field (compare Ehrlich 1995, 1997, 1997a). Throughout the remaining first half of the twentieth century there continued to be important contributions to the theory of the continuum including the algebraic (Artin and Schreier 1926/1965) and logical (Tarski 1939/1986, 1948/1986) versions of the theory of elementary continuity. During the 1950s, under the influence of A. A. Markov (1856–1922), a Russian school of constructive analysis was developed based on a continuum consisting of real numbers with a *recursive* Cauchy sequence (compare Kushner 1984), and during the 1960s Errett Bishop (1928–1983) introduced an alternative constructive approach to analysis (1967) based on a model of the continuum whose theorems, unlike those of Brouwer and the Russian school, are all provable in classical mathematics.

Also during roughly the same period, interest in Weyl's predicative theory was rekindled by Feferman (1998), who developed his own predicative approach to analysis. In addition, since the late 1950s a number of

nonstandard theories of continua have appeared that make use of infinitesimals including those arising from Abraham Robinson's (1918–1974) nonstandard approach to analysis (1961/1969, 1966/1974), those arising from F. W. Lawvere (1979, 1980) and Anders Kock's (1981) ideas on smooth infinitesimal analysis and synthetic differential geometry, the partially ordered continuum of Curt Schmieden and Detlef Laugwitz (1958), and Ehrlich's so-called absolute arithmetic continuum (1987, 1989, 1992) based on J. H. Conway's (1976) theory of *surreal numbers*. Still another theory that arose during this period is the theory of fuzzy real numbers based on Lofti Asker Zadeh's (1987) theory of fuzzy sets.

Of the nonstandard theories mentioned earlier, the intuitionist, Bishop-style constructivist, predicativist, and the Robinsonian and post-Robinsonian infinitesimalist theories have been given the most attention by philosophers. In the subsequent three sections these will be considered in turn with some attention paid to the Russian constructivist theory as well.

CONSTRUCTIVIST THEORIES

Constructivism is a rubric that has come to designate a family of approaches to the foundations of mathematics that are loosely united by their opposition to certain forms of mathematical reasoning employed in the mainstream mathematical community. However, as the term loosely suggests, there are significant differences between the various schools and substantial differences in attitude can be found even among the representatives of a given school or a single representative over time. However, whether the result of their rejection of actual infinities or the universal applicability of certain principles of classical logic, or their insistence on the use of algorithmic constructions of one form or another, constructivists have always found themselves at odds with the Cantor-Dedekind theory of the continuum and have sought to provide alternatives that are constructively sanctioned by their own particular lights. To distinguish the resulting intuitionist, Russian constructivist, and Bishop-style constructivist arithmetic continua from $\mathbb{R}$, from now on they will be denoted as $\mathbb{R}_I$, $\mathbb{R}_R$, and $\mathbb{R}_C$, respectively.

Before the late 1960s the constructivist theory of the continuum that attracted the greatest attention is the intuitionistic theory of Brouwer. Until 1914, Brouwer, like Borel before him, embraced a holistic view in which the continuum is regarded as a primitive notion given directly by intuition that cannot be understood as the totality of its individually definable elements. Thereafter, while still clinging to certain aspects of the irreducible conception, Brouwer adopted a more analytic view in which the continuum, while not a completed nondeumerable totality, can be more and more completely specified in a never-ending fashion with one's increasing knowledge as a medium of free development.

The basis of Brouwer's conception of the continuum is the concept of a *choice sequence*, a concept not acceptable to classical mathematics. According to Brouwer, the construction of such a potentially infinite sequence is always incomplete in the sense that at any given instant of its construction the sequence is limited to a finite number of terms. A choice sequence α is given by a fixed initial segment $\alpha_1, \alpha_2, \ldots, \alpha_n$ of mathematical objects along with a corresponding set of restrictions $R_1, R_2, \ldots, R_n$, where R_n restricts the range of possible choices for α_{n+1}. In particular, real numbers are introduced by Brouwer as choice sequences that are Cauchy sequences of rational numbers. Of course, being a choice sequence, the notion of a Cauchy sequence of rational numbers must be appropriately understood. More specifically, according to the intuitionist one can assert that a potentially infinite sequence $r_1, r_2, \ldots$ of rational numbers is a Cauchy sequence only if one knows there is a procedure that, given any positive integer k, effectively produces a natural number N along with a proof that N has the specified Cauchy property, for example, $|r_m - r_n| < 1/k$ for all $m, n < N$.

For Brouwer, such a choice sequence is not a technique for approximating some preexisting real number—it is the choice sequence itself, growing in time, that is the real number. Some such real numbers are introduced by letting the choices be prescribed by a fixed algorithm. These so-called *lawlike* sequences lead to the existence of real numbers such as e, π, and √2. Other real numbers arise from forbidding any restriction on the rational numbers one chooses outside of assuring that the choice sequence is Cauchy. Between these two extremes, however, there is a wide range of possibilities for introducing real numbers. Until 1927 Brouwer did not place any restrictions on choice sequences—having regarded them as sufficiently explained by the freedom of a supposed ideal mathematician generating them—but thereafter he became more specific and continued to revise his conception of permissible choice sequences until the early 1950s (compare Troelstra 1982, pp. 472–474).

The adoption of choice sequences forces a nonclassical logic on intuitionists that rejects the universal validity of the *law of excluded middle*. For example, given the incomplete nature of choice sequences the intuitionist has no right to assume for an arbitrary pair of choice sequences α and β having identical initial segments that α

$= \beta \vee \alpha \neq \beta$. The logic that has come to be embraced by intuitionists is a subtheory of classical logic called *intuitionistic logic* (compare Heyting 1971, chapter 7). For the intuitionist, logic does not provide a foundation for mathematics but emerges from one's mathematical practice. The adoption of intuitionistic logic leads intuitionists to a theory of real numbers and corresponding theory of continuity that differs markedly from their classical counterparts. For example, besides the just-cited deviation, it is not possible for intuitionists to prove precise analogs of the following classical results for their own system $\mathbb{R}_I$ of real numbers: $\forall x, y \in \mathbb{R}(xy = 0 \rightarrow (x = 0 \vee y = 0))$; $\forall x, y \in \mathbb{R}(x \leq y \vee x \geq y)$; $\forall x, y \in \mathbb{R}(x > y \vee x = y \vee x > y)$; every subset of $\mathbb{R}$ that is bounded above has a least upper bound; $\forall x \in \mathbb{R}(x \text{ is rational} \vee x \text{ is irrational})$. Nor can they prove the intermediate value theorem or the extreme value theorem. From Brouwer's perspective, this inability is not a limitation since each of the previous assertions implies an instance of the law of excluded middle that is not intuitionistically sanctioned (compare Bridges and Reeves 1999, pp. 72–73). This attitude, and the embrace of intuitionistic logic more generally (with the aforementioned implications for their own theories of real numbers and corresponding theories of continuity) is a common thread that binds constructive mathematicians. From the constructivist point of view, accepting the law of excluded middle as a universal principle would mean the existence of a universal procedure for generating for each proposition P, either a proof of P or a proof of $\neg P$, where a procedure for generating a proof of $\neg P$ is understood to be a method for generating a contradiction from a hypothetical proof of P. However, such a procedure is not available; if it were one could decide a proposition P—such as Goldbach's conjecture—the truth of which has not been decided.

Brouwer's concept of a choice sequence that is Cauchy corresponds, as Heyting notes, "to the intuitive concept of the continuum as a possibility of a gradual determination of points" (1971, p. 34). To develop an adequate theory of continuity and of analysis more generally, however, the mathematician must be able to talk about classes of such real numbers and functions. From the standpoint of the intuitionist, however, one cannot collect them together into a Cantorian set—there are simply too many of them. Rather, for the intuitionist, they are held together in a *spread*—roughly speaking, a growing tree, whose emerging paths through the tree correspond to the various ways an initial segment of a choice sequence can be continued (compare Heyting 1971, chapter 3).

Moreover, to obtain the central continuity theorems concerning such classes of real numbers and functions, Brouwer introduced two fundamental ideas governing the mathematical treatment of choice sequences: the *weak continuity principle for numbers* (compare Veldman 2001, Atten and Dalen 2002, Troelstra and Dalen 1988, chapter 5) and the *principle of bar induction* (compare Kushner 2001). The continuity principle makes choice sequences serviceable by contending that a total function from choice sequences to natural numbers never requires more input than an initial segment (of a choice sequence) to generate its output; and the induction principle ensures, among other things, that the entire intuitionistic continuum can be treated in a constructively manageable fashion (compare Atten 2004, chapters 3–4).

The adoption of the continuity principle and the induction principle gives the intuitionistic theory of continuity its own distinctive constructivist flavor and leads to even more striking deviations from the classical theory than those listed earlier. For example, in virtue of the continuity principle, the analogs of the aforementioned classical results that cannot be established as a result of the adoption of intuitionistic logic are now provably false (compare Bridges and Richman 1987, pp. 4–5; Troelstra and Dalen 1988, pp. 257–258). What is perhaps the most notorious such deviation, however, is that all functions from $\mathbb{R}_I$ to $\mathbb{R}_I$ are continuous, and uniformly continuous at that. This apparent absurdity arises in part from the fact that the contention "f is a function defined on all of $\mathbb{R}_I$" is substantially stronger construed intuitionistically than is the contention "f is a function defined on all of $\mathbb{R}$" construed classically. Consider, for example, the classical discontinuous function f defined by

$$f(x) = \begin{cases} 0, \text{ if } x < 0 \\ 1, \text{ if } x \geq 0. \end{cases}$$

From Brouwer's perspective, f is not a function at all since one cannot prove $\forall x \exists y(f(x) = y)$ by intuitionistically sanctioned means. In particular, one cannot prove $\forall x(x < 0 \vee x \geq 0)$ insofar as the definition of f does not tell one how to compute $f(x)$ if x is a number for which one cannot assert either $x < 0$ or $x \geq 0$. Closely related to this is still another striking deviation, the so-called *unsplittability* of the intuitionistic continuum; that is, unlike the Cantor-Dedekind arithmetic continuum, there do not exist two disjoint nonempty subsets of an interval of the intuitionistic continuum whose union is the given interval, nor are there such partitions of the intuitionistic continuum whose union is the continuum itself. Accordingly,

for the intuitionist, as for Anaxagoras and Aristotle before them, it is not possible to separate out a point from their continuum or from an interval thereof.

While Brouwer's theory has attracted a good deal of attention from philosophers and logicians, it has received comparatively little attention from standard mathematicians. Whether this is because of the philosophical precepts underlying it, the highly nonclassical nature of the mathematical arguments it employs, the belief that the resulting mathematics is too impoverished, or simply the absence of a perceived need for it, is difficult to say. In 1967, however, Brouwer's theory was given an especially stinging critique, not by standard mathematicians, but by Bishop, whose treatise *Foundations of Constructive Analysis* is widely credited with having breathed new life into constructive mathematics. In the treatise's polemical opening chapter Bishop describes the construction and motivation underlying Brouwer's theory of the continuum in the following biting terms:

> Brouwer became involved in metaphysical speculation by his desire to improve the theory of the continuum. A bugaboo of both Brouwer and the logicians has been compulsive speculation about the nature of the continuum. In the case of the logicians this leads to contortions in which various formal systems, all detached from reality, are interpreted within one another in the hope that the nature of the continuum will somehow emerge. In Brouwer's case there seems to have been a nagging suspicion that unless he personally intervened to prevent it, the continuum would turn out to be discrete. He therefore introduced the method of free-choice sequences for constructing the continuum, a consequence of which the continuum cannot be discrete because it is not well enough defined. This makes mathematics so bizarre it becomes unpalatable to mathematicians, and foredooms the whole of Brouwer's program. This is a pity, because Brouwer had a remarkable insight into the defects of classical mathematics, and he made a heroic effort to set things right. (1967, p. 6)

Bishop sought to place analysis on a constructivist foundation that is free of the perceived difficulties mentioned earlier, a project that has been extended by a number of other mathematicians including Douglas Bridges and Fred Richman. In Bishop's arithmetic continuum $\mathbb{R}_C$ a real number is simply defined as a sequence $\{x_n\}$ of rational numbers that satisfies the condition $|x_m - x_n| \leq m^{-1} + n^{-1}$ (for all integers $m, n \geq 1$); though some authors, following Troelstra and Dirk van Dalen (1988, pp. 253–254), prefer to use equivalence classes of certain sequences of rational numbers in their place. Thus, for Bishop, as for Markov, every real number is a lawlike Cauchy sequence. Using a system of axioms for constructive ordered fields with a formulation of the Archimedean axiom and a constructive formulation of the least upper-bound principle, Bridges (1999) shows that $\mathbb{R}_C$ can be characterized in a manner that closely resembles one of the aforementioned standard axiomatizations of $\mathbb{R}$. Working independently, Alberto Ciaffaglione and Pietro Di Gianantonio (2002) and Herman Geuvers and Milad Niqui (2002) provide equivalent axiomatizations of $\mathbb{R}_C$ that employ constructive versions of Cauchy completeness in place of Bridges's least upper-bound condition. Assuming the axiom of countable choice (compare Troelstra and Dalen 1988, pp. 189–190)—an axiom that is frequently adopted by constructive mathematicians—Geuvers and Niqui (2002) further establish the categoricity of the axiomatizations. Absent the choice axiom (or some equivalent thereof), there are models of the axioms that are not isomorphic to $\mathbb{R}_C$—in particular, Troelstra and Dalen's version of the constructive continuum based on Dedekind cuts (1988, pp. 270–274).

Bishop tended to distinguish his theory of analysis from the classical theory by emphasizing the former's demand for algorithmic constructions. Following in the footsteps of Brouwer, Bishop took the concept of an algorithm as a primitive, undefined notion and was led to reject the universal validity of the law of excluded middle by interpreting mathematical existence strictly in terms of computability or constructivity. Bridges and, especially, Richman speculate that the theory of analysis that emerges from Bishop's approach may be regarded as the subtheory of the classical theory that is obtainable employing intuitionistic logic as opposed to classical logic as the underlying logic (Bridges and Richman 1987; Richman 1990, 1996; Bridges 1999). Since (in accordance with Heyting's axiomatization of intuitionistic logic [1971]) one passes from intuitionistic logic to classical logic by embracing the universal validity of the law of excluded middle, on their view Bishop's theory may be regarded as the subtheory of the classical theory that can be obtained without appealing to the instances of this classical law that are not intuitionistically sanctioned (Bridges and Richman 1987, p. 120).

Besides being a subtheory of the classical theory, Bishop's theory may be regarded as a subtheory of the intuitionist and Russian constructivist theories as well

(Bridges and Richman 1987, chapter 6). Whereas, in accordance with Bridges and Richman's view, one moves from Bishop's theory to the classical theory by embracing the universal validity of the law of excluded middle, to move from Bishop's theory to the intuitionistic theory one introduces Brouwer's weak continuity principle along with a seminal consequence of Brouwer's principle of bar induction called the *fan theorem*, and to pass from Bishop's theory to the Russian constructivist theory one adds a consequence of Church's thesis that all computable sequences of natural numbers are recursive (Bridges and Richman 1987, chapter 5). Talk of such passages, however, applies solely to theories in the logician's narrow sense of the term; it ignores the divergent philosophical motivations and mathematical trappings of the four theories. With respect to the Russian constructivist theory, for example, it ignores that every real number is a recursive real number, that algorithms are Markov algorithms, that functions are Gödel numbers of algorithms that compute them, and so on.

The differences between the classical, Russian constructivist, and intuitionist theories, however, go beyond their respective philosophical motivations and mathematical trappings; they have different theories of continuity, as is evident from the following theorems of the three respective theories:

> (Classical): Some functions f:[0, 1]→$\mathbb{R}$ are not continuous and, hence, not uniformly continuous
>
> (Russian Constructivist): Whereas all functions f:[0, 1]→$\mathbb{R}_R$ are continuous, some are not uniformly continuous
>
> (Intuitionist): All functions f:[0, 1]→$\mathbb{R}_I$ are not only continuous, they are uniformly continuous

Despite these differences, Bishop's theory manages to lie comfortably within the common core of the three theories in part because in Bishop's theory attention is restricted from the outset to functions that are assumed to be uniformly continuous on each closed interval of $\mathbb{R}_C$.

Restricting attention to the just-said functions, Bishop managed to obtain a surprisingly robust theory of analysis that includes among a wide range of other theorems constructive analogs of the intermediate value theorem and the extreme value theorem. Like many of the theorems of constructive analysis, the latter two theorems differ from their classical counterparts by having weakened conclusions or strengthened hypotheses. One constructive version of the intermediate value theorem asserts that a uniformly continuous function f:[a, b]→$\mathbb{R}_C$ takes on a value as close to the given intermediate value as one pleases, and the constructive version of the extreme value theorem asserts that a uniformly continuous function f:[a, b]→$\mathbb{R}_C$ does have a maximum (minimum), though the maximum (minimum) is not necessarily assumed. There is also a constructive version of the intermediate value theorem that asserts that in a particular class of cases (which includes all real-analytic functions), the intermediate value in question is in fact realized (compare Bishop and Bridges 1985, pp. 40–41; Troelstra and Dalen 1988, pp. 292–295).

However, despite the strength of Bishop's analysis and its compatibility with classical mathematics, Bishop's theory, like its intuitionist and Russian constructivist forerunner's, has not attracted the kind of attention from classical mathematicians its practitioners had hoped for (compare Bridges and Reeves 1999, p. 67). Moreover, while praising the significance of Bishop's accomplishment, some devotees of Brouwer's theory have questioned the adequacy of the analysis of the continuum that emerges from Bishop's approach. In particular, they are concerned that in Bishop's theory, like Markov's before it, the continuum of real numbers is restricted to those real numbers introduced by lawlike Cauchy sequences. For example, in his monograph devoted to Brouwer's theory, Mark van Atten remarks:

> One may, like Markov and Bishop, settle for just the lawlike sequences … but while practical, that also amounts to ducking the issue of how to model the full continuum. Brouwer's achievement is to have found a way to analyze the continuum that does not let it fall apart into discrete elements … and it is constructive to boot. (2004, p. 33)

PREDICATIVE THEORIES

Between classical mathematics, in which arbitrary sets are embraced, and Bishop's constructive mathematics, in which only algorithmically constructed objects are permissible, there is an intermediate approach called *predicativism*, in which only definable sets are considered, and in which quantifiers over sets are interpreted as ranging only over sets that have previously been defined. Although having its roots in Russell's and Poincaré's attempts to lay the blame for the paradoxes of set theory on definitions that define entities in terms of classes to which they belong—so-called *impredicative* definitions—it was Weyl, in his monograph *Das Kontinuum* (1918/1994), who first undertook the development of a theory of the arithmetic continuum and of analysis on it in a predicatively acceptable fashion.

Central to Weyl's critique of the classical theory is its dependence on the proposition that every nonempty set of real numbers that is bounded above has a least upper bound, the definition of the least upper bound of a set being inextricably impredicative. Weyl proposed overcoming this by employing the predicatively sanctioned proposition: Every nonempty sequence of reals having an upper bound has a least upper bound. Using this idea, Weyl constructed a restricted set of real numbers containing all reals that are expressible as Dedekind cuts definable in his system. Although the set of standard real numbers not definable in Weyl's system is everywhere dense, Weyl showed that on the basis of his continuum most, if not all, of the nineteenth-century analysis of piece-wise continuous functions can be carried out predicatively. On the other hand, as Weyl conceded, substantial and significant portions of modern analysis are not obtainable in his system including "the more far-reaching integration- and measure theories of Riemann, Darboux, Cantor, Jordan, Lebesque and Carathéodory" (Weyl 1918/1994, p. 86).

In the years following the publication of *Das Kontinuum* Weyl abandoned his own approach in favor of Brouwer's intuitionistic framework. Soon thereafter, however, he returned to the standard mathematical fold and distanced himself from Brouwer's school and from foundations work more generally. In the ensuing years, Weyl's predicative theory lay largely dormant until the 1960s, when a number of authors including, most prominently, Feferman undertook a formalization and systematic analysis of Weyl's system as well as the development of a variety of predicative extensions thereof (compare Feferman 1964, 1988/1998, 1993/1998). Unlike Weyl, who worried about the purported *vicious circles* associated with impredicative definitions, Feferman was motivated in part by the concern that the unbridled use of such definitions presupposed a strong form of platonic realism regarding sets, a view he found philosophically objectionable; he was also interested in providing an analysis of predicativity itself, as well as with the purely logical question of the extent to which analysis can be carried out by predicative means. One of Feferman's extensions, called W in honor of Weyl, has been proven to be sufficiently strong to permit the reconstruction of almost all of classical analysis as well as important portions of modern analysis that are not obtainable in Weyl's original system. Feferman maintains that:

> While there are clearly parts of theoretical analysis that cannot be carried out in W because they make essential use of the 1.u.b. axiom applied to sets rather than sequences, or because they make essential use of transfinite ordinals or cardinals, or because they deal with nonseparable spaces, *the working hypothesis that all of scientifically applicable analysis can be developed in W has been verified in its core parts*. What remains to be done is to examine results closer to the margin to see whether this hypothesis indeed holds in full generality. (1993/1998, p. 294)

INFINITESIMALIST APPROACHES

Following Emil Artin (1898–1962) and Otto Schreier (1901–1929), an ordered field *K* may be said to be *real-closed* if it admits no extension to a more inclusive ordered field that results from supplementing *K* with solutions to polynomial equations with coefficients in *K* (1926/1965). Intuitively speaking, real-closed ordered fields are precisely those ordered fields having no holes that can be filled by algebraic means alone. Tarski demonstrated that real-closed ordered fields are precisely the ordered fields that are first-order indistinguishable from ℝ or, to put this another way, they are precisely the ordered fields that satisfy the elementary (i.e., first-order) content of the Dedekind continuity axiom (1939/1986, 1948/1986). For this reason they are called *elementary continua*. While ℝ is the best-known elementary continuum, it is hardly the only one.

Some elementary continua, like ℝ, are Archimedean, though most are non-Archimedean; and among the latter many are extensions of ℝ. In the early 1960s Robinson (1961, 1966) made the momentous discovery that among the real-closed extensions of the reals there are number systems that can provide the basis for a consistent and entirely satisfactory nonstandard approach to analysis based on infinitesimals. Robinson motivated his work with the following words:

> It is our main purpose to show that these models [i.e. number systems] provide a natural approach to the age old problem of producing a calculus involving infinitesimal (infinitely small) and infinitely large quantities. As is well known, the use of infinitesimals, strongly advocated by Leibnitz and unhesitatingly accepted by Euler fell into disrepute after the advent of Cauchy's methods which put Mathematical Analysis on a firm foundation. Accepting Cauchy's standards of rigor, later figures in the domain of nonarchimedean quantities concerned themselves only with small fragments of the edifice of Mathematical Analysis. We men-

> tion only du Bois-Reymond's Calculus of Infinites [1875] and Hahn's work on nonarchimedean fields [1907] which in turn were followed by the theories of Artin-Schreier [1926] and, returning to analysis, of Hewitt [1948] and Erdös, Gillman and Henriksen [1955]. Finally, a recent and rather successful effort at developing a calculus of infinitesimals is due to Schmieden and Laugwitz [1958] whose number system consists of infinite sequences of rational numbers. The drawback of this system is that it includes zero-divisors and that it is only partially ordered. In consequence, many classical results of the Differential and Integral calculus have to be modified to meet the changed circumstance. (1961/1979, p. 4)

Being elementary continua, Robinson's number systems do not have the just-cited drawbacks of the number system of Schmieden and Laugwitz. By analogy with Thoralf Skolem's (1934) *nonstandard model of arithmetic*, a number system from which Robinson drew inspiration, Robinson called his totally ordered number systems *nonstandard models of analysis*. These number systems, which are often called *hyperreal number systems* (Keisler 1976, 1994), may be characterized as follows: Let $\langle \mathbb{R}, S: S \in \mathfrak{F} \rangle$ be a relational structure where $\mathfrak{F}$ is the set of all finitary relations defined on $\mathbb{R}$ (including all functions). Furthermore, let $^*\mathbb{R}$ be a proper extension of $\mathbb{R}$ and for each *n*-ary relation $S \in \mathfrak{F}$ let $^*S$ be an *n*-ary relation on $^*\mathbb{R}$ that is an extension of *S*. The structure $\langle ^*\mathbb{R}, \mathbb{R}, ^*S{:}S \in \mathfrak{F} \rangle$ is said to be a hyperreal number system if it satisfies the *Transfer Principle*: Every *n*-tuple of real numbers satisfies the same first-order formulas in $\langle \mathbb{R}, S{:}S \in \mathfrak{F} \rangle$ as it satisfies in $\langle ^*\mathbb{R}, \mathbb{R}, ^*S{:}S \in \mathfrak{F} \rangle$.

The existence of hyperreal number systems is a consequence of the compactness theorem of first-order logic and there are a number of algebraic techniques that can be employed to construct such a system. One commonly used technique is the *ultapower* construction (Keisler 1976, pp. 48–57; Goldblatt 1998, chapter 3), though not all hyperreal number systems can be obtained this way. By results of H. Jerome Keisler (1963; 1976, pp. 58–59), however, every hyperreal number system must be (isomorphic to) a *limit ultapower*.

Using the transfer principle, one can develop satisfactory nonstandard conceptions and treatments of all of the basic concepts and theorems of the calculus including those from the theories of integration, differentiation, and continuity to name only a few (compare Keisler 1986, Goldblatt 1998, Loeb 2000). For example, it follows from the transfer principle that a real-valued function *f* is continuous at $a \in \mathbb{R}$ iff $^*f(x)$ is infinitesimally close to $^*f(a)$ whenever *x* is infinitesimally close to *a*, for all $x \in {}^*\mathbb{R}$. On the basis of this result one may prove various classical properties governing the continuity of real-valued functions including the intermediate and extreme value theorems (Goldblatt 1998, pp. 79–80). It should be emphasized, however, that Robinson's discoveries do not provide vindication of the Leibnizian formalism or of the seventeenth- and eighteenth-century preanalytic formalisms more generally. For example, whereas Leibniz conceived of differentiation and integration in terms of ratios of and infinite sums of infinitesimals, respectively, for Robinson they are real numbers that are infinitesimally close to such ratios and sums. On the other hand, nonstandard analysis not only demonstrates that the branch of mathematics erected for the study of continuously varying magnitude can be fully developed using infinitely large and infinitely small numbers as Leibniz and his followers had envisioned but it also provides one with an intuitively satisfying alternative to the standard picture of a continuum and of continuous phenomena more generally that is mathematically adequate and logically sound relative to classical mathematics.

Modern analysis, however, goes far beyond the traditional province of the calculus, dealing with arbitrary sets of reals, sets of sets of reals, sets of functions from sets of reals to sets of reals, and the like. Importantly, nonstandard analysis is entirely applicable to this expanded arena as well. However, the methods of superstructures (Robinson 1966) and internal set theory (Nelson 1977) that are most usually employed for this purpose are of little relevance here (compare Chang and Keisler 1990, §4.4; Robert 1988).

Unlike $\mathbb{R}$, the structures that may play the role of $^*\mathbb{R}$ in a hyperreal number system are far from being unique up to isomorphism. From a purely mathematical point of view this causes no difficulty and from the standpoint of varying applications can even be advantageous (compare Keisler 1994, p. 229). On the other hand, if one takes $^*\mathbb{R}$ to be a model of *the* continuous straight line of geometry—something practitioners of nonstandard analysis tend not to do—the absence of uniqueness is a bit disconcerting. Still, as several nonstandard analysts including Tom Lindstrøm (1988, p. 82) and Keisler (1994, p. 229) emphasize, even $\mathbb{R}$ is not as unique as one would like to think since its uniqueness up to isomorphism is in fact relative to the underlying set theory. In particular, by retaining the construction of $\mathbb{R}$ and supplementing the set theory with additional axioms, one can change the

second-order theory of the real line. This leads Keisler (1994) to suggest that not only is ZFC not the appropriate underlying set theory for the hyperreal number system but also that set theory might have developed differently had it been developed with the hyperreal numbers rather than the real numbers in mind. According to Keisler, an appropriate set theory "should have the power set operation to insure the unique existence of the real number system, and another operation which insures the unique existence of the pair consisting of the real and hyperreal number systems" (p. 230).

Consistent with the previous observation, one type of axiom that is used to secure categoricity is a saturation axiom (Keisler 1976, pp. 57–60). As the name suggests, saturation axioms ensure that the line is extremely rich. A hyperreal number system $\langle {}^*\mathbb{R}, \mathbb{R}, {}^*S: S \in \mathfrak{F}\rangle$ is said to be *κ-saturated* if any set of formulas with constants from $^*\mathbb{R}$ of power less than κ is satisfiable whenever it is finitely satisfiable. If κ is the power of $^*\mathbb{R}$, the hyperreal number system is said to be *saturated*. Although there is a wide range of hyperreal number systems in ZFC that are saturated to varying degrees of power less than the power of $^*\mathbb{R}$, saturated hyperreal number systems do not exist in ZFC. In virtue of classical results from the theory of saturated models, however, there is (up to isomorphism) a unique saturated hyperreal number system of power κ whenever $\kappa > 2^{\aleph_0}$ and either κ is (strongly) inaccessible or the generalized continuum hypothesis (GCH) holds at κ (i.e., $\kappa = \aleph_{\alpha+1} = 2^{\aleph_\alpha}$ for some α). So, for example, by supplementing ZFC with the assumption of the existence of an uncountable inaccessible cardinal, one can obtain uniqueness (up to isomorphism) by limiting attention to saturated hyperreal lines having the least such power (Keisler 1976, p. 60).

However, as Ehrlich (2002, 2004) emphasizes, perhaps the most remarkable of all elementary continua that may play the role of $^*\mathbb{R}$ in a hyperreal number system (and bring categoricity to the hyperreal line to boot) is Conway's ordered field of *surreal numbers* (1976/2001), a system that was not created with nonstandard analysis in mind. This would correspond (to within isomorphism) of adopting a hyperreal number system that is the union of an elementary chain of ω_α-saturated hyperreal number systems where α ranges over the class *On* of all ordinals. Though such models do not exist in ZFC, they can be suitably characterized and shown to exist (up to isomorphism) in von Neumann–Bernays–Gödel (NBG) set theory with the axiom of global choice (Ehrlich 1989). Since NBG is a conservative extension of ZFC, its sets have the same properties as those of standard set theory (compare Fraenkel, Bar-Hillel, and Levy 1973). The idea of employing such a hyperreal number system to establish the categoricity of the hyperreal line appears to be due (at least implicitly) to Keisler (1976, p. 59; 1994, p. 233; theorem 3 of addendum to Ehrlich 1989), but guided by reasons of simplicity and convenience he chooses the least uncountable inaccessible cardinal approach instead.

The ordered field of surreal numbers, which Conway calls *No*, is so remarkably inclusive that, subject to the proviso that numbers—construed here as members of ordered (number) fields—be individually definable in terms of sets of NBG, it may be said to contain "All Numbers Great and Small." In this respect, *No* bears much the same relation to ordered fields as the system of real numbers bears to Archimedean ordered fields. Ehrlich (1987, 1989a, 1992, 2002) suggests that whereas the real number system may be regarded as an *arithmetic continuum modulo the Archimedean axiom*, the system of surreal numbers may be regarded as a sort of *absolute arithmetic continuum modulo NBG*. To lend credence to this thesis, Ehrlich provides a variety of categorical axiomatizations of $\mathbb{R}$ making use of novel continuity axioms (that are equivalent to any of the familiar continuity axioms) with axioms for Archimedean ordered fields (or Archimedean real-closed ordered fields) and shows that by simply deleting the Archimedean axiom one obtains categorical axiomatizations of *No* (Ehrlich 1992, theorems 1, 4, and 6). Ehrlich also introduces a natural generalization of Dedekind's conception of a gap, called a *set-gap*, and provides further evidence for the thesis by showing that whereas $\mathbb{R}$ is (up to isomorphism) the unique elementary continuum having no set-gaps that satisfies the Archimedean condition, *No* is (up to isomorphism) the unique elementary continuum having no set-gaps that satisfies the *On*-Archimedean condition, the latter being a natural generalization of the Archimedean condition that is appropriate for *No* (Ehrlich 1992, Lemma 1, Theorem 7; 2001, pp. 1255–1256). Critical to the proof of the latter result is Ehrlich's (1988, 1989, 2001) further characterization of *No* (up to isomorphism) as the unique elementary continuum such that for all subsets *X* and *Y* of the field where every member of *X* precedes every member of *Y* there is a member of the field lying strictly between those of *X* and those of *Y*. Intuitively, this characterizes *No* (up to isomorphism) as the unique ordered field having neither algebraic limitations nor order-theoretic limitations that are definable in terms of sets of standard set theory.

Besides its distinguished structure as an ordered field, *No* has a rich hierarchical structure that emerges from the recursive clauses in terms of which it is defined. This algebraico-tree-theoretic structure, or *simplicity hierarchy*, as Ehrlich (1994, 2001) calls it, depends on *No*'s structure as a lexicographically ordered binary tree and arises from the fact that the sums and products of any two members of the tree are the simplest possible elements of the tree consistent with *No*'s structure as an ordered group and an ordered field, respectively, it being understood that x is *simpler than* y just in case x is a predecessor of y in the tree. Among the remarkable consequences of this algebraico-tree-theoretic structure is that much as the surreal numbers emerge from the empty set of surreal numbers by a transfinite recursion that yields chains of increasingly less and less simpler numbers, the recursive process of defining *No*'s arithmetic in turn gives rise to chains of increasingly richer and richer numbers systems with the result that an isomorphic copy of every elementary continuum emerges in *No* as the union of a chain of elementary continua each of which is an initial subtree of *No* (Ehrlich 2001).

Conway (1976/2001) shows that besides the reals *No* contains a natural isomorphic copy of Cantor's ordinals, and hence, by virtue of the axiom of choice, the cardinals as well (Ehrlich 2001, pp. 1253–1256). Ehrlich (1988, 2001, 2002, 2004) notes that *No* also provides a natural setting for the non-Cantorian theories of the infinite (and infinitesimal) pioneered by Veronese, Levi-Civita, Hilbert, and Hahn in connection with their work on non-Archimedean-ordered algebraic and geometric systems and by du Bois-Reymond, Stolz, G. H. Hardy (1877–1947), and Felix Hausdorff (1868–1942) in connection with their work on the rate of growth of real functions (compare Ehrlich 1994, 1995, 2005; Fisher 1981). This, together with the observation about the relationship between *No* and hyperreal number systems, leads Ehrlich (2002, 2004) to observe that over and above providing a panorama of the entire set-theoretic spectrum of numbers great and small (modulo NBG), the purported absolute arithmetic continuum provides a unifying framework for many of the most important totally ordered systems of finite, infinite, and infinitesimal numbers that have played and continue to play prominent roles in mathematics since the days of Cantor and Dedekind.

Within a decade of the development of nonstandard analysis, Lawvere proposed a profound and novel approach to differential geometry based on infinitesimals. Unlike Robinson, who was stimulated by Leibniz's idea that the properties of infinitesimals should reflect the properties of the reals, Lawvere's ideas more closely mirror the heuristic ideas of geometers who envisioned a vector tangent to a surface at a point as a tiny arc of a curve having the vector tangent to it. Building on Lawvere's ideas, Kock (1981) presents a systematic treatment of the theory under the rubric *synthetic differential geometry* (SDG).

Unlike the nonzero infinitesimals employed in nonstandard analysis, the nonzero infinitesimal elements of SDG are *nilpotent*, that is, each such infinitesimal d satisfies the condition $d^2 = 0$. Nilpotent infinitesimals are not *invertible* (in the sense that they have no multiplicative inverses) and as such a line in SDG in not modeled by a field or a portion thereof. Rather, in SDG a line is modeled by a ring $\Re$ containing a subset $D = \{d \in \Re : d^2 = 0\}$ which satisfies the

Kock-Lawvere axiom: For every mapping $f: D \to \Re$, there is precisely one $b \in \Re$, such that for all $d \in D$, $f(d) = f(0) + d \cdot b$.

Geometrically speaking, the Kock-Lawvere axiom asserts that the graph of every function $f: D \to \Re$ is a piece of the unique straight line through $(0, f(0))$ with slope b. It is a consequence of this assumption that in SDG a tangent vector to a curve C at a point p is a nondegenerate infinitesimal line segment around p coincident with C.

Another consequence of the Kock-Lawvere axiom is that in SDG, unlike in Euclidean geometry, there are pairs of points in the plane that are connected by more than one straight line. In this regard, SDG resembles Johannes Hjelmslev's (1873–1950) *natural geometry* (compare Kock 2003), a geometry that was designed to mirror real (as opposed to ideal) sense experience and that also employs nilpotent infinitesimals. However, unlike in natural geometry, in SDG there are pairs of points in a plane that are not connected by any line at all. This arises in part from the fact that whereas the nilpotent infinitesimals in natural geometry have "a quantitative (linear ordered) character," those employed in SDG do not (Kock 2003, pp. 226–228). For an axiomatization of "Euclidean Geometry with Infinitesimals" inspired by SDG, see Succi Cruciani (1989).

A space X in SDG is said to be *indecomposable* if no proper nonempty part U of X is *detachable* in the sense that there is a part V such that $U \cup V = X$ where $U \cap V = \emptyset$. There are models of SDG in which a classical space $\mathbb{R}^n$ has a counterpart X that is indecomposable if X is *connected*. John Bell takes this to imply that "the connected continua of SDG are *true* continua in something like the Anaxagoran sense" (1995, p. 56). In this respect, they are

also reminiscent of the unsplittable continuum of Brouwer; however, the similarity is not perfect and varies depending on the axioms adopted for SDG (Bell 2001).

Another respect in which SDG is similar to Brouwer's theory is the failure of the intermediate value theorem in its underlying theory of analysis. In fact, in SDG, unlike in Brouwer's system, the theorem even fails for some polynomials (Moerdijk and Reyes 1991, pp. 317–318), a failure that runs contrary to the thinking of Leibniz and Euler let alone Bolzano, Cauchy, and Weierstrass. Accordingly, while SDG may provide a viable alternative for differential geometry, its underlying analysis may not be as well suited to provide a natural alternative for classical analysis, at least not if it hopes to mirror the latter's most central ideas regarding continuity.

Unlike nonstandard analysis, which is developed in a set-theoretic setting, SDG is developed in a category-theoretic framework. Moreover, whereas the underlying logic employed in nonstandard analysis is classical logic, in SDG the underlying logic is intuitionistic logic. In SDG every function $f:\Re\rightarrow\Re$ is differentiable and, hence, infinitely differentiable (i.e. *smooth*) as well as continuous in the sense that it sends neighboring points to neighboring points. It is sometimes maintained (compare Bell 1995, p. 56) that it is the ubiquitous nature of continuity within SDG that forces the change from classical to intuitionistic logic. This, however, is apt to be misleading since it is possible to develop a theory of continua in which the continuity of functions from the continuum to the continuum is likewise ubiquitous though the underlying logic is classical (compare the so-called Cauchy continuum due to Schmieden and Laugwitz [1958; Laugwitz 2001, p. 134]). Rather, it is the Kock-Lawvere axiom that underlies the incompatibility of SDG with classical logic (compare Lavendhomme 1996, pp. 2–5). It is therefore interesting to note that Paolo Giordano (2001), by suitably modifying the axiom, presents a variation of SDG based on nilpotent infinitesimals in which the underlying logic is entirely classical, and he observes that the nilpotent infinitesimals could be supplemented with invertible infinitesimals as well. Earlier, Ieke Moerdijk and Gonzalo E. Reyes (1991), while retaining the underlying intuitionistic logic, also introduced an alternative approach in which invertible as well as nilpotent infinitesimals are employed. The work of Moerdijk, Reyes, and Giordano, much like the pioneering work of Lawvere and Kock, provides still other models of mathematical continua.

CONCLUDING REMARKS

"*Bridging the gap between the domains of discreteness and of continuity*, or between arithmetic and geometry, is a central, presumably even *the* central, problem of the foundations of mathematics." So write Fraenkel, Yehoshua Bar-Hillel, and Azriel Levy in their mathematico-philosophical classic *Foundations of Set Theory* (1973, p. 212). Cantor and Dedekind of course believed they had bridged the gap with the creation of their arithmetico-set theoretic continuum of real numbers, and it remains a well-entrenched tenet of standard mathematical philosophy that indeed they had. At the same time, Cantor was overly sanguine when in 1883 he seemed to suggest, or at least implied, that his theory of the continuum, unlike that of the ancients, had "been thought out … with the clarity and completeness … required to exclude the possibility of different opinions among [its] posterity" (Cantor 1883/1996, p. 903). Indeed, while Cantor and Dedekind had succeeded in replacing the vague ancient conception with a clear and complete arithmetico-set-theoretic conception, a conception that was adequate for the needs of analysis, differential geometry, and the empirical sciences, they did not, nor could not, free their theory of its logical, theoretical, and philosophical presuppositions, nor could they preclude the possibility that other adequate conceptual schemes, each self-consistent, could be devised offering alternative visions of the continuum.

However, it was critiques of the former that gave rise to some of its competitors and the realization of the logical possibility of the latter that gave rise to others. To some extent, the architects of each of its competitors were motivated by the belief, or at least the hope, that their respective theories are or with time would be adequate for the needs of analysis (or differential geometry), though in the cases of the constructivist and predicativist architects *analysis* was equated with *legitimate analysis* constructively and predicatively construed. Outside of the overarching question of the historical needs of analysis, the question of whether legitimate analysis thus understood is adequate for the needs of the empirical sciences and the physical sciences, in particular, is the subject of dispute (compare Fletcher 2002; Hellman 1993, 1993a, 1997, 1998; Bridges 1995; Billinge 2000; Bridges and Ishihara 2001; Feferman 1988/1998, postscript). Nonstandard analysis has bypassed all of these questions since from the standpoint of the standard domain it is as strong as or even stronger than standard analysis depending on what one assumes (Henson and Keisler 1986). Moreover, like their late nineteenth- and early twentieth-century non-Archimedean geometric forerunners, non-

standard analysis and the infinitesimalist approaches more generally have drawn attention to the possibility of physical continua whose logical cogency let alone physical possibility had long been in doubt. Whether empirical science will require such a theory, as some already contend (compare Fenstad 1987, 1988) and others, following Veronese (compare 1909/1994, p. 180), will not rule out, only time will tell. Nevertheless, while showing little sign of displacing the standard theory, the constructivist, predicativist, and infinitesimalist alternatives have performed, and continue to perform, important logical and philosophical service. Nonstandard analysis has also had real success in shedding important light on and establishing significant new results in various areas of analysis, theoretical physics, and economics (compare Albeverio, Luxemburg, and Wolff 1995; Arkeryd, Cutland, and Henson 1997; Loeb and Wolff 2000). However, whether nonstandard analysis or any of the other nonstandard theories canvassed earlier, together with its corresponding theory of the continuum, will eventually assume the status of the standard theory (or even stand alongside the standard theory as a co-standard theory) remains to be seen.

See also Infinity in Mathematics and Logic.

Bibliography

ANCIENT AND MEDIEVAL PHILOSOPHICAL AND COSMOLOGICAL THEORIES

Aristotle. *The Complete Works of Aristotle: The Revised Oxford Translation*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Furley, David. J. *Two Studies in the Greek Atomists*. Princeton, NJ: Princeton University Press, 1967.

Kirk, Geoffrey Stephen; John Earle Raven, and Malcolm Schofield. *The Presocratic Philosophers: A Critical History with a Selection of Texts*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983.

Kretzmann, Norman, ed. *Infinity and Continuity in Ancient and Medieval Thought*. Ithaca, NY: Cornell University Press, 1982.

Murdoch, John. "Infinity and Continuity." In *The Cambridge History of Later Medieval Philosophy*, edited by Norman Kretzmann, Anthony Kenny, and Jan Pinborg. New York: Cambridge University Press, 1982.

Pyle, Andrew. *Atomism and Its Critics: From Democritus to Newton*. Bristol, U.K.: Thoemme Press, 1995.

Sambursky, Shmuel. *Physics of the Stoics*. London: Routledge and Kegan Pau1, 1959.

Sorabji, Richard. *Time, Creation, and the Continuum: Theories in Antiquity and the Early Middle Ages*. Ithaca, NY: Cornell University Press, 1983.

White, Michael J. *The Continuous and the Discrete: Ancient Physical Theories from a Contemporary Perspective*. Oxford, U.K.: Clarendon Press, 1992.

ANCIENT MATHEMATICAL THEORIES

Fowler, David. *The Mathematics of Plato's Academy*. 2nd ed. Oxford, U.K.: Clarendon Press, 1999.

Heath, Sir Thomas L., ed. *The Thirteen Books of Euclid's Elements: Translated from the Text of Heiberg*, Vols. 1–3. New York: Dover, 1956.

Knorr, Wilbur R. *The Ancient Tradition of Geometric Problems*. Boston: Birkhäuser, 1986.

Knorr, Wilbur R. "Archimedes and the Pre-Euclidean Proportion Theory." *Archives Internationales d'Histore des Sciences* 28 (1978): 183–244.

Knorr, Wilbur R. *The Evolution of the Euclidean Elements: A Study of the Theory of Incommensurable Magnitudes and Its Significance for Early Greek Geometry*. Dordrecht, Netherlands: D. Reidel, 1975.

Thiele, Rüdiger. "Antiquity." In *A History of Analysis*, edited by Hans Jahnke. Providence, RI: American Mathematical Society, 2003.

THE EARLY MODERN THEORY OF REAL NUMBERS

Klein, Jacob. *Greek Mathematical Thought and the Origin of Algebra*. Translated by Eva Brann. Cambridge, MA: MIT Press, 1968.

Newton, Isaac. "Universal Arithmetik." In *The Mathematical Works of Isaac Newton*. Vol. 2, edited by Derrk T. Whiteside. New York: Johnson Reprint Corporation, 1967.

Pycior, Helena. *Symbols, Impossible Numbers, and Geometric Entanglements*. New York: Cambridge University Press,1997.

Stevin, Simon. "Arithmetique." In *The Principal Works of Simon Stevin*, Volume 2B, edited by D. J. Struik. Amsterdam, Netherlands: C. V. Swets and Zeitlinger, 1958.

Strong, Edward W. *Procedures and Metaphysics: A Study in the Philosophy of Mathematical-Physical Science in the Sixteenth and Seventeenth Centuries*. Merrick, NY: Richwood Publishing Co., 1976.

HISTORY OF THE CALCULUS

Baron, Margret. *The Origin of the Infinitesimal Calculus*. New York: Oxford University Press, 1969.

Berkeley, George. *The Analyst* (1734). *De Motu and The Analyst: A Modern Edition with Introductions and Commentary by Douglas M. Josseph*. Dordrecht, Netherlands: Kluwer Academic, 1992. *The Analyst* was originally published in 1734.

Boyer, Carl. *The Concepts of the Calculus and Its Conceptual Development*. New York: Dover, 1959.

Edwards, C. H., Jr. *The Historical Development of the Calculus*. New York: Springer-Verlag, 1979.

Jahnke, Hans Niels, ed. *A History of Analysis*. Providence, RI: American Mathematical Society, 2003.

LEIBNIZ'S VIEWS

Bertoloni-Meli, Domenico. *Equivalence and Priority: Newton versus Leibniz*. Oxford, U.K.: Clarendon Press, 1993.

Bos, Henk J. M. "Differentials, Higher-Order Differentials, and the Derivative in the Leibnizian Calculus." *Archive for History of Exact Sciences* 14 (1974): 1–90.

Cassirer, Ernst. *Leibniz' Systeme in seinen wissenschaftlichen Grundlagen*. Marburg, Germany: N. G. Elwert, 1902.

Leibniz, Gottfried Wilhelm. *The Labyrinth of the Continuum: Writings on the Continuum Problem, 1672–1686*. Translated

by Richard T. W. Arthur. New Haven, CT: Yale University Press, 2001.

Leibniz, Gottfried Wilhelm. "Specimen geometriae luciferae." (c. 1695). In *Leibniz Mathematische Schriften*. Vol. 7, edited by C. I. Gerhardt, 260–298. Hildesheim, Germany: G. Olms, 1962.

NEWTON VIEWS

Bertoloni-Meli, Domenico. *Equivalence and Priority: Newton versus Leibniz*. Oxford, U.K.: Clarendon Press, 1993.

Guicciardini, Niccolò. "Newton's Method and Leibniz's Calculus." In *A History of Analysis*, edited by Hans Niels Jahnke. Providence: American Mathematical Society, 2003.

Kitcher, Philip. "Fluxions, Limits, and Infinite Littleness." *Isis* 64 (221) (1973): 33–49.

GENERAL SET THEORY

Fraenkel, Abraham A., Yehoshua Bar-Hillel, and Azriel Levy. *Foundations of Set Theory*. 2nd ed. Amsterdam, Netherlands: North-Holland, 1973.

Levy, Azriel. *Basic Set Theory*. Berlin: Springer-Verlag, 1979.

THE CANTOR-DEDEKIND THEORY OF THE CONTINUUM

Cantor, Georg. "Grundlagen einer allgemeinen Mannigfaltigkeitslehre: Ein mathematischphilosophischer Versuch in der Lehre des Unendlichen" (1883). Translated by William B. Ewald as "Foundations of a General Theory of Manifolds: A Mathematico-Philosophical Investigation into the Theory of the Infinite." In *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Vol. 2, edited by William B. Ewald. Oxford, U.K.: Clarendon Press, 1996.

Cantor, Georg. "Über die Ausdehnung eines Satzes der Theorie der trigonometrischen Reihen" (1872). In *Georg Cantor Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, edited by Ernest Zermelo. Berlin: J. Springer, 1932.

Dedekind, Richard. "Stetigkeit und irrationale Zahlen" (1872). Translated by W. W. Beman as "Continuity and Irrational Numbers." In *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Vol. 2, edited by William B. Ewald (Oxford, U.K.: Clarendon Press, 1996).

Epple, Moritz. "The End of the Science of Quantity: Foundations of Analysis, 1860–1910." In *A History of Analysis*, edited by Hans Niels Jahnke. Providence, RI: American Mathematical Society, 2003.

Ferreirós, José. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*. Berlin: Birkhäuser Verlag, 1999.

Hilbert, David. *Foundations of Geometry* (1899). 10th ed. Translated by Leo Unger. LaSalle, IL: Open Court, 1971.

Hilbert, David. *Grundlagen der Geometrie*. Leipzig, Germany: Teubner, 1899.

Huntington, Edward V. *The Continuum and Other Types of Serial Order: With an Introduction to Cantor's Transfinite Numbers*. 2nd ed. Cambridge, MA: Harvard University Press, 1917.

Poincaré, Henri. *Science and Hypothesis*. In *The Foundations of Science*. Translated by George Bruce Halsted. New York: Science Press, 1913. "Mathematical Magnitude and Experience" is a translation of a reprinting in modified form of "Le Continu Mathématique," *Revue de Métaphysique et de Morale* (1893).

Russell, Bertrand. *The Principles of Mathematics*. Cambridge, U.K.: Cambridge University Press, 1903.

THE CONTINUUM HYPOTHESIS

Cohen, Paul. "Conversation with Paul Cohen." In *More Mathematical People: Contemporary Conversations*, edited by Donald J. Albers, Gerald L. Alexanderson, and Constance Reid, 43–58. Boston: Harcourt, Brace, Jovanovich, 1990.

Cohen, Paul. "Independence of the Continuum Hypothesis." *Proceedings of the National Academy of Sciences* 50 (1963): 1143–1148; 51 (1964): 105–110.

Dales, H. G., and G. Oliveri. *Truth in Mathematics*. Oxford, U.K.: Clarendon Press, 1998.

Feferman, Solomon et al. "Does Mathematics Need New Axioms?" *Bulletin of Symbolic Logic* 6 (4) (2000): 401–446.

Gödel, Kurt. *The Consistency of the Continuum Hypothesis*. Princeton, NJ: Princeton University Press, 1940.

Gödel, Kurt. "What Is Cantor's Continuum Problem?" (1947). In *Philosophy of Mathematics: Selected Readings*. 2nd ed., edited by Paul Benacerraf and Hilary Putnam. New York: Cambridge University Press, 1983.

Kanamori, Akihiro. *The Higher Infinite*. 2nd ed. Berlin: Springer-Verlag, 2003.

Maddy, Penelope. *Naturalism in Mathematics*. Oxford, U.K.: Clarendon Press, 1997.

Woodin, W. Hugh. "The Continuum Hypothesis." *Notices of the American Mathematical Society* 48 (6) (2001): 567–576; (7) (2001): 681–690; 49 (1) (2002): 46.

Woodin, W. Hugh. "Set Theory after Russell: The Journey Back to Eden." In *De Gruyter Series in Logic and Its Applications (Number 6)*, edited by Godehard Link. Berlin: Walter de Gruyter, 2004.

CONSTRUCTIVIST APPROACHES: GENERAL BACKGROUND

Beeson, Michael J. *Foundations of Constructive Mathematics: Metamathematical Studies*. Berlin: Springer-Verlag, 1985.

Bridges, Douglas, and Fred Richman. *Varieties of Constructive Mathematics*. New York: Cambridge University Press, 1987.

Troelstra, A. S., and D. van Dalen. *Constructivism in Mathematics: An Introduction*. Amsterdam, Netherlands: North-Holland, 1988.

BROUWER'S INTUITIONISTIC APPROACH

Atten, Mark van. *On Brouwer*. Belmont, CA: Wadsworth, 2004.

Atten, Mark van, and Dirk van Dalen. "Arguments for the Continuity Principle." *Bulletin of Symbolic Logic* 8 (3) (2002): 329–347.

Brouwer, L. E. J., and Walter P. van Stigt. "Part I: L. E. J. Brouwer." In *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*, edited by Paolo Mancosu. New York: Oxford University Press, 1998.

Dalen, Dirk van, ed. *Brower's Cambridge Lectures on Intuitionism*. New York: Cambridge University Press, 1981.

Dummett, Michael. *Elements of Intuitionism*. 2nd ed. Oxford, U.K.: Clarendon Press, 2000.

Fraenkel, Abraham A., Yehoshua Bar-Hillel, and Azriel Levy. *Foundations of Set Theory*. 2nd ed. Amsterdam, Netherlands: North-Holland, 1973.

Heyting, Arend. *Intuitionism: An Introduction*. 3rd rev. ed. Amsterdam: North-Holland, 1971.

Kushner, Boris. "On Brouwerian Bar Induction." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

McCarthy, David. "Intuitionism in Mathematics." In *The Oxford Handbook of Philosophy of Mathematics and Logic*, edited by Stewart Shapiro. New York: Oxford University Press, 2005.

Posy, Carl. "Intuitionism and Philosophy." In *The Oxford Handbook of Philosophy of Mathematics and Logic*, edited by Stewart Shapiro. New York: Oxford University Press, 2005.

Stigt, Walter P. van. *Brouwer's Intuitionism*. Amsterdam, Netherlands: North-Holland, 1990.

Troelstra, A. S. "On the Origin and Development of Brouwer's Concept of Choice Sequence." In *The L.E.J. Brouwer Centenary Symposium*, edited by A. S. Troelstra and Dirk van Dalen. Amsterdam, Netherlands: North-Holland, 1982.

Veldman, Wim. "Understanding and Using Brouwer's Continuity Principle." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

BISHOP'S CONSTRUCTIVE APPROACH

Bishop, Errett. *Foundations of Constructive Analysis*. New York: McGraw-Hill, 1967.

Bishop, Errett, and Douglas Bridges. *Constructive Analysis*. New York: Springer-Verlag, 1985.

Bridges, Douglas. "A Constructive Look at the Real Number Line." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Bridges, Douglas. "Constructive Mathematics: A Foundation for Computable Analysis." *Theoretical Computer Science* 219 (1–2) (1999): 95–109.

Bridges, Douglas, and Steve Reeves. "Constructive Mathematics in Theory and Programming Practice." *Philosophia Mathematica* (3) 7 (1999): 65–104.

Bridges, Douglas, and Fred Richman. *Varieties of Constructive Mathematics*. New York: Cambridge University Press, 1987.

Ciaffaglione, Alberto, and Pietro Di Gianantonio. "A Tour with Constructive Real Numbers." In *Types for Proofs and Programs (Durham, 2000)*, edited by Paul Callaghan et al., 41–52. Berlin: Springer-Verlag, 2002.

Geuvers, Herman, and Milad Niqui. "Constructive Reals in Coq: Axioms and Categoricity." In *Types for Proofs and Programs (Durham, 2000)*, edited by Paul Callaghan et al., 79–95. Berlin: Springer-Verlag, 2002.

Richman, Fred. "Interview with a Constructive Mathematician." *Modern Logic* (3) 6 (1996): 247–271.

Richman, Fred. "Intuitionism as Generalization." *Philosophia Mathematica* 5 (1990): 124–128.

RUSSIAN CONSTRUCTIVE APPROACH

Beeson, Michael J. *Foundations of Constructive Mathematics: Metamathematical Studies*. Berlin: Springer-Verlag, 1985.

Kushner, Boris A. *Lectures on Constructive Mathematical Analysis*. Translated by Elliott Mendelson. Providence, RI: American Mathematical Society, 1984.

CONTINUITY AND MATHEMATICAL PHYSICS

Billinge, Helen. "Applied Constructive Mathematics: On G. Hellman's 'Mathematical Constructivism in Spacetime.'" *British Journal for the Philosophy Science* 51 (2) (2000): 299–318.

Bridges, Douglas. "Constructive Mathematics and Unbounded Operators. Reply to: 'Constructive Mathematics and Quantum Mechanics: Unbounded Operators and the Spectral Theorem.'" *Journal of Philosophical Logic* 24 (5) (1995): 549–561.

Bridges, Douglas, and Hajime Ishihara. "Constructive Unbounded Operators." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

Capek, Milic. *The Philosophical Impact of Contemporary Physics*. Princeton, NJ: D. Van Nostrand Company, Inc., 1961.

Earman, John. *World Enough and Space-Time*. Cambridge, MA: MIT Press, 1989.

Fletcher, Peter. "A Constructivist Perspective on Physics." *Philosophia Mathematica* 10 (1) (2002): 26–42.

Forrest, Peter. "Is Space-Time Discrete or Continuous? An Empirical Question." *Synthese* 103 (3) (1995): 327–354.

Hellman, Geoffrey. "Constructive Mathematics and Quantum Mechanics: Unbounded Operators and the Spectral Theorem." *Journal of Philosophical Logic* 22 (3) (1993): 221–248.

Hellman, Geoffrey. "Gleason's Theorem Is Not Constructively Provable." *Journal of Philosophical Logic* 22 (2) (1993a): 193–203.

Hellman, Geoffrey. "Mathematical Constructivism in Spacetime." *British Journal for the Philosophy Science* 49 (3) (1998): 425–450.

Hellman, Geoffrey. "Quantum Mechanical Unbounded Operators and Constructive Mathematics—A Rejoinder to Bridges." *Journal of Philosophical Logic* 26 (2) (1997): 121–127.

Markopoulou, Fotini. "Planck-scale Models of the Universe." In *Science and Ultimate Reality. Quantum Theory, Cosmology, and Complexity. In Honor of the 90th Birthday of John Archibald Wheeler*, edited by John D. Barrow, Paul C. W. Davies, and Charles L. Harper, Jr. Cambridge, U.K.: Cambridge University Press, 2004.

Poincaré, Henri. *The Value of Science. In The Foundations of Science*. Translated by George Bruce Halsted. New York and Garrison, NY: The Science Press, 1913. Reprinted in 1921. "Analysis and Physics," which is Chapter 5 of this book, is a translation of a reprinting in modified form of "Sur les Rapports de l'analyse pure et de la physique Mathématique," which was first published in *Acta Mathematica* 21 (1897): 331–341.

Smolin, Lee. *Three Roads to Quantum Gravity*. New York: Basic Books, 2001.

Smolin, Lee. "Quantum Theories of Gravity: Results and Prospects." In *Science and Ultimate Reality. Quantum Theory, Cosmology, and Complexity. In Honor of the 90th Birthday of John Archibald Wheeler*, edited by John D. Barrow, Paul C. W. Davies, and Charles L. Harper, Jr. Cambridge, U.K.: Cambridge University Press, 2004.

WEYL'S PREDICATIVE APPROACH AND ITS AFTERMATH

Bell, John L. "Hermann Weyl on Intuition and the Continuum." *Philosophia Mathematica* (3) 8 (2000): 259–273.

Feferman, Solomon. "Predicativity." In *The Oxford Handbook of Philosophy of Mathematics and Logic*, edited by Stewart Shapiro. New York: Oxford University Press, 2005.

Feferman, Solomon. "Systems of Predicative Analysis." *Journal of Symbolic Logic* 29 (1964): 1–30.

Feferman, Solomon. "Weyl Vindicated: Das Kontinuum 70 Years Later (1988)." In *In the Light of Logic*, edited by Solomon Feferman. New York: Oxford University Press, 1998.

Feferman, Solomon. "Why a Little Bit Goes a Long Way: Logical Foundations of Scientifically Applicable Mathematics (1993)." In *In the Light of Logic*, edited by Solomon Feferman. New York: Oxford University Press, 1998.

Parsons, Charles. "Realism and the Debate on Impredicativity, 1917—1944." In *Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman*, edited by Wilfried Sieg, Richard Sommer, and Carolyn Talcott. Natick, MA: A. K. Peters, 2002.

Scholz, Erhard. "Herman Weyl on the Concept of Continuum." In *Proof Theory: History and Philosophical Significance*, edited by Vincent F. Hendricks, Stig Andur Pedersen, and Klaus Frovin Jorgensen. Dordrecht, Netherlands: Kluwer Academic, 2000.

Weyl, Hermann. *Das Kontinuum* (1918). Translated by Stephen Pollard and Thomas Bole as *The Continuum: a Critical Examination of the Foundation of Analysis*. New York: Dover Publications Inc., 1994.

Weyl, Hermann, et al. "Part II: H. Weyl." In *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*, edited by Paolo Mancosu. New York: Oxford University Press, 1998.

ELEMENTARY CONTINUITY

Artin, Emil, and Otto Schreier. "Algebraische Konstruktion reeller Körper" (1926). In *The Collected Papers of Emil Artin*, edited by Serge Lang and John T. Tate. Reading, MA: Addison-Wesley, 1965.

Tarski, Alfred. *Completeness of Elementary Algebra and Geometry* (1939). In *Collected Papers of Alfred Tarski, Vol. 2, 1935–1944*, edited by Steven R. Givant and Ralph N. McKenzie. Basel, Switzerland: Birkhäuser Verlag, 1986.

Tarski, Alfred. *A Decision Method for Elementary Algebra and Geometry* (1948). In *Collected Papers of Alfred Tarski, Vol. 3, 1945–1957*, edited by Steven R. Givant and Ralph N. McKenzie. Basel, Switzerland: Birkhäuser Verlag, 1986.

Tarski, Alfred. "What Is Elementary Geometry?" (1959). In *Collected Papers of Alfred Tarski, Vol. 4, 1958–1979*, edited by Steven R. Givant and Ralph N. McKenzie. Basel, Switzerland: Birkhäuser Verlag, 1986.

Tarski, Alfred, and Steven Givant. "Tarski's System of Geometry." *Bulletin of Symbolic Logic* 5 (2) (1999): 175–214.

Sinaceur, Hourya. "Calculation, Order, and Continuity." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Sinaceur, Hourya Benis. *Fields and Models: From Sturm to Tarski and Robinson*. New York: Springer-Verlag, 2006.

CONTINUITY AND EUCLIDEAN GEOMETRY

Ehrlich, Philip. "From Completeness to Archimedean Completeness: An Essay in the Foundations of Euclidean Geometry." In *A Symposium on David Hilbert*, edited by Alfred Tauber and Akihiro Kanamori. *Synthese* 110(1) (1997a): 57–76.

Greenberg, Marvin J. *Euclidean and Non-Euclidean Geometries: Development and History*. 3rd ed. New York: Freeman, 1993.

Hartshorne, Robin. *Geometry: Euclid and Beyond*. New York: Springer-Verlag, 2000.

Heath, Sir Thomas L. ed. *The Thirteen Books of Euclid's Elements: Translated from the Text of Heiberg*, Volumes 1–3. New York: Dover Publications, Inc., 1956.

Tarski, Alfred. "What is Elementary Geometry?" (1959). In *Collected Papers of Alfred Tarski*, *Vol. 4, 1958–1979,* edited by Steven R. Givant and Ralph N. McKenzie. Basel: Birkhäuser Verlag, 1986.

Tarski, Alfred, and Givant, Steven. "Tarski's System of Geometry." *Bulletin of Symbolic Logic* 5(2) (1999): 175–214.

INFINITESIMALIST APPROACHES: HISTORICAL BACKGROUND

Ehrlich, Philip. "Hahn's *Über die nichtarchimedischen Grössensysteme* and the Development of the Modern Theory of Magnitudes and Numbers to Measure Them." In *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, edited by Jaakko Hintikka. Dordrecht, Netherlands: Kluwer Academic, 1995.

Ehrlich, Philip. "The Rise of Non-Archimedean Mathematics and the Roots of a Misconception I: The Emergence of Non-Archimedean Systems of Magnitudes." *Archive for History of Exact Sciences*, (Springer-Verlag Online) (2005).

Fisher, Gordon. "The Infinite and Infinitesimal Quantities of du Bois-Reymond and Their Reception." *Archive for History of Exact Sciences* 24 (2) (1981): 101–164.

DU BOIS-REYMOND'S CONCEPTION

Du Bois-Reymond, Paul. *Die allgemine Functionentheorie*. Tübingen, Germany: Verlag der H. Laupp'schen Buchhandlung, 1882.

Du Bois-Reymond, Paul. "Sur la grandeur relative des infinis des fonctions." *Annali di matematica pura ed applicata* 4 (1870–1871): 338–353.

Du Bois-Reymond, Paul. "Ueber die Paradoxen des Infinitärcalcüls." *Mathematische Annalen* 11 (1877): 149–167.

VERONESE'S NON-ARCHIMEDEAN CONTINUUM

Ehrlich, Philip. "Dedekind Cuts of Archimedean Complete Ordered Abelian Groups." *Algebra Universalis* 37 (2) (1997): 223–234.

Ehrlich, Philip. "General Introduction." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Fisher, Gordon. "Veronese's Non-Archimedean Linear Continuum." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Veronese, Giuseppe. *Fondamenti di geometria a più dimensioni e a più specie di unità rettilinee esposti in forma elementare*. Padua, Italy: Tipografia del Seminario, 1891.

Veronese, Giuseppe. *Grundzüge der Geometrie von mehreren Dimensionen und mehreren Arten gradliniger Einheiten in elementarer Form entwickelt. Mit Genehmigung des Verfassers nach einer neuen Bearbeitung des Originals übersetzt von Adolf Schepp*. Leipzig: Teubner, 1894.

Veronese, Giuseppe. "La geometria non-Archimedea" (1909). Translated by Mathieu Marion as "On Non-Archimedean Geometry." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Veronese, Giuseppe. "Il continuo rettilineo e l'assioma V di Archimede." *Memorie della Reale Accademia dei Lincei, Atti della Classe di scienze naturali, fisiche e matematiche* (4) 6 (1889): 603–624.

THE NON-ARCHIMEDEAN CONTRIBUTIONS OF LEVI-CIVITA AND HAHN

Ehrlich, Philip. "From Completeness to Archimedean Completeness: An Essay in the Foundations of Euclidean Geometry." In *A Symposium on David Hilbert*, edited by Alfred Tauber and Akihiro Kanamori. *Synthese* 110 (1) (1997a): 57–76.

Ehrlich, Philip. "Hahn's *Über die nichtarchimedischen Grössensysteme* and the Development of the Modern Theory of Magnitudes and Numbers to Measure Them." In *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, edited by Jaakko Hintikka. Dordrecht, Netherlands: Kluwer Academic, 1995.

Hahn, Hans. "Über die nichtarchimedischen Grössensysteme." *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Wien, Mathematisch - Naturwissenschaftliche* Klasse 116 (Abteilung IIa) (1907): 601–655.

Laugwitz, Detlef. "Tullio Levi-Civita's Work on Non-Archimedean Structures (with an Appendix: Properties of Levi-Civita Fields)." *Tullio Levi-Civita Convegno Internazionale Celebrativo Del Centenario Della Nascita, Accademia Nazionale Dei Lincei Atti Dei Convegni Lincei, Rome* 8 (1975): 297–312.

Levi-Civita, Tullio. "Sugli infiniti ed infinitesimi attuali quali elementi analitici" (1892–1893). In *Tullio Levi-Civita, Opere Matematiche, Memorie e Note, Volume primo 1893–1900*. Bologna, Italy: Nicola Zanichelli, 1954.

Levi-Civita, Tullio. "Sui Numeri Transfiniti" (1898). In *Tullio Levi-Civita, Opere Matematiche, Memorie e Note, Volume primo 1893–1900*. Bologna, Italy: Nicola Zanichelli, 1954.

NONSTANDARD ANALYSIS

Albeverio, Sergio A., Wilhelm A. J. Luxemburg, and Manfred P. H. Wolff, eds. *Advances in Analysis, Probability, and Mathematical Physics: Contributions of Nonstandard Analysis*. Dordrecht, Netherlands: Kluwer Academic, 1995.

Arkeryd, Leif O., Nigel J. Cutland, and C. Ward Henson, eds. *Nonstandard Analysis: Theory and Applications*. Dordrecht, Netherlands: Kluwer Academic, 1997.

Chang, C. C., and H. Jerome Keisler. *Model Theory*. 3rd ed. Amsterdam, Netherlands: North-Holland, 1990.

Di Nasso, Mauro, and Marco Forti. "On the Ordering of the Nonstandard Real Line." In *Logic and Algebra*, edited by Yi Zhang. Providence, RI: American Mathematical Society, 2002.

Fenstad, Jens Erik. "The Discrete and the Continuous in Mathematics and the Natural Sciences." In *L'Infinito Nella Scienza*, edited by Giuliano Toraldo di Francia. Rome: Istituto Della Enciclopedia Italiana, Fondata Di G. Treccani, 1987.

Fenstad, Jens Erik. "Infinities in Mathematics and the Natural Sciences." In *Methods and Applications of Mathematical Logic: Contemporary Mathematics*. Vol. 69, edited by Walter A. Carnielli and Luiz Paulo de Alcantara. Providence, RI: American Mathematical Society, 1988.

Goldblatt, Robert. *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*. New York: Springer-Verlag, 1998.

Henson, C. W., and H. Jerome Keisler. "On the Strength of Nonstandard Analysis." *Journal of Symbolic Logic* 51 (2) (1986): 377–1386.

Keisler, H. Jerome. *Elementary Calculus*. 2nd ed. Boston: Prindle, Weber and Schmidt, 1986.

Keisler, H. Jerome. *Foundations of Infinitesimal Calculus*. Boston: Prindle, Weber and Schmidt, 1976.

Keisler, H. Jerome. "The Hyperreal Line." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Keisler, H. Jerome. "Limit Ultrapowers." *Transactions of the American Mathematical Society* 107 (1963): 383–408.

Lindstrøm, Tom. "An Invitation to Nonstandard Analysis." In *Nonstandard Analysis and Its Applications*, edited by Nigel Cutland. New York: Cambridge University Press, 1988.

Loeb, Peter A. "An Introduction to Nonstandard Analysis." In *Nonstandard Analysis for the Working Mathematician*, edited by Peter A. Loeb and Manfred Wolff. Dordrecht, Netherlands: Kluwer Academic, 2000.

Loeb, Peter A., and Manfred Wolff, eds. *Nonstandard Analysis for the Working Mathematician*. Dordrecht, Netherlands: Kluwer Academic, 2000.

Nelson, Edward. "Internal Set Theory: A New Approach to Nonstandard Analysis." *Bulletin of the American Mathematical Society* 83 (1977): 1165–1198.

Robert, Alain. *Nonstandard Analysis*. New York: Wiley, 1988.

Robinson, Abraham. "The Metaphysics of the Calculus" (1967). In *Abraham Robinson Selected Papers. Vol. 2, Nonstandard Analysis and Philosophy*, edited by W. A. J. Luxemburg and Stephan Körner. New Haven, CT: Yale University Press, 1979.

Robinson, Abraham. "Non-standard Analysis" (1961). In *Abraham Robinson Selected Papers. Vol. 2, Nonstandard Analysis and Philosophy*, edited by W. A. J. Luxemburg and Stephan Körner. New Haven, CT: Yale University Press, 1979.

Robinson, Abraham. *Non-standard Analysis*. Amsterdam, Netherlands: North-Holland, 1966.

Robinson, Abraham. *Non-standard Analysis*. Rev. ed. Amsterdam: North-Holland Publishing Company, 1974.

Skolem, Thoralf. "Über die Nichtcharakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen." *Fundamenta Mathematica* 23 (1934): 150–161.

Zakon, Elias. "Remarks on the Nonstandard Real Axis." In *Applications of Model Theory to Algebra, Analysis, and*

Probability, edited by W. A. J. Luxemburg. New York: Holt, Rinehart and Winston, 1969.

SMOOTH INFINITESIMAL ANALYSIS

Bell, John L. "The Continuum in Smooth Infinitesimal Analysis." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

Bell, John L. "Infinitesimals and the Continuum." *The Mathematical Intelligencer* 17 (2) (1995): 55–57.

Bell, John L. *A Primer of Infinitesimal Analysis*. New York: Cambridge University Press, 1998.

Giordano, Paolo. "Nilpotent Infinitesimals and Synthetic Differential Geometry in Classical Logic." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

Kock, Anders. "Differential Calculus and Nilpotent Real Numbers." *Bulletin of Symbolic Logic* 9 (2) (2003): 225–230.

Kock, Anders. *Synthetic Differential Geometry*. New York: Cambridge University Press, 1981.

Lavendhomme, René. *Basic Concepts of Synthetic Differential Geometry*. Dordrecht, Netherlands: Kluwer Academic, 1996.

Lawvere, F. William. "Categorical Dynamics." In *Topos Theoretic Methods in Geometry*, edited by Anders Kock. Aarhus Matematisk Institut: Various Publication Series, 1979.

Lawvere, F. William. "Toward the Description in a Smooth Topos of the Dynamically Possible Motions and Deformations of a Continuous Body." *Cahiers de Topologie et Géométrie Différentielle Catégoriques* 21(4) (1980): 377–392.

McLarty, Colin. "Defining Sets as Sets of Points." *Journal of Philosophical Logic* 17 (1988): 75–90.

Moerdijk, Ieke, and Gonzalo E. Reyes. *Models for Smooth Infinitesimal Analysis*. New York: Springer-Verlag, 1991.

Succi Cruciani, Rosanna. "Euclidean Geometry with Infinitesimals." *Rendiconti di Matematica e delle sue Applicazioni, Serie* (Serie VII) 8 (4) (1989): 557–578.

THE WORK OF SCHMIEDEN AND LAUGWITZ

Laugwitz, Detlef. "Kurt Schmieden's Approach to Infinitesimals: An Eye-Opener to the Historiography of Analysis." In *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, edited by Peter Schuster, Ulrich Berger, and Horst Osswald. Dordrecht, Netherlands: Kluwer Academic, 2001.

Laugwitz, Detlef. "Leibniz's Principle and Omega Calculus." In *Le Continu Mathematique, Colloque de Cerisy*, edited by Hourya Sinaceur and Jean-Michel Salanskis. Paris: Springer-Verlag, 1992.

Laugwitz, Detlef. "Ω-Calculus as a Generalization of Field Extension—An Alternative Approach to Nonstandard Analysis." In *Nonstandard Analysis: Recent Developments*, edited by Albert Emerson Hurd. Berlin: Springer-Verlag, 1983.

Schmieden, Curt, and Detlef Laugwitz. "Eine Erweiterung der Infinitesimalrechnung." *Mathematische Zeitschrift* 69 (1958): 1–39.

SURREAL NUMBERS

Conway, J. H. *On Numbers and Games*. London: Academic Press, 1976.

Conway, J. H. *On Numbers and Games*. 2nd ed. Natick, MA: A.K. Peters, 2001.

Conway, J. H. "The Surreals and Reals." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Ehrlich, Philip. "The Absolute Arithmetic Continuum and the Unification of All Numbers Great and Small." In *Philosophical Insights into Logic and Mathematics (Abstracts).* Nancy, France: Université de Nancy Laboratoire de Philosophie et d'Histoire des Sciences, Archive Henri Poincaré (2002): 41–43.

Ehrlich, Philip. "The Absolute Arithmetic and Geometric Continua." In *PSA 1986*. Vol. 2, edited by Arthur Fine and Peter Machamer. Lansing, MI: Philosophy of Science Association, 1987.

Ehrlich, Philip. "Absolutely Saturated Models." *Fundamenta Mathematicae* 133 (1) (1989): 39–46.

Ehrlich, Philip. "All Number Great and Small." In *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by Philip Ehrlich. Dordrecht, Netherlands: Kluwer Academic, 1994.

Ehrlich, Philip. "An Alternative Construction of Conway's Ordered Field No." *Algebra Universalis* 25 (1988): 7–16. Errata, *Algebra Universalis* 25 (1988): 233.

Ehrlich, Philip. "Number Systems with Simplicity Hierarchies: A Generalization of Conway's Theory of Surreal Numbers." *Journal of Symbolic Logic* 66 (3) (2001): 1231–1258. Corrigendum, 70 (3) (2005): 1022.

Ehrlich, Philip. "Surreal Numbers: An Alternative Construction (Abstract)." *Bulletin of Symbolic Logic* 8 (3) (2002a): 448.

Ehrlich, Philip. "Surreal Numbers and the Unification of All Numbers Great and Small." *Bulletin of Symbolic Logic* 10 (2) (2004): 253.

Ehrlich, Philip. "Universally Extending Arithmetic Continua." In *Le Continu Mathematique, Colloque de Cerisy*, edited by Hourya Sinaceur and Jean-Michel Salanskis. Paris: Springer-Verlag, 1992.

Ehrlich, Philip. "Universally Extending Continua." *Abstracts of Papers Presented to the American Mathematical Society* 10 (1989a): 15.

CAUCHY'S CONTINUUM

Cleave, J. P. "Cauchy, Convergence and Conitinuity." *British Journal for the Philosophy of Science* 22 (1971): 27–37.

Grabiner, Judith V. *The Origins of Cauchy's Rigorous Calculus*. Cambridge, MA.: MIT Press, 1981.

Lakatos, Imre. "Cauchy and the Continuum." *The Mathematical Intelligencer* (1) (1978): 151–161.

Laugwitz, Detlef. "Definite Values of Infinite Sums: Aspects of the Foundations of Infinitesimal Analysis around 1820." *Archive for History of Exact Sciences* 39 (1989): 195–245.

Laugwitz, Detlef. "Infinitely Small Quantities in Cauchy's Textbook." *Historia Mathematica* (14) (1987): 258–274.

Lützen, Jesper. "The Foundations of Analysis in the 19th Century." In *A History of Analysis*, edited by Hans Niels Jahnke. Providence, RI: American Mathematical Society, 2003.

Robinson, Abraham. "The Metaphysics of the Calculus" (1967). In *Abraham Robinson Selected Papers. Vol. 2, Nonstandard Analysis and Philosophy*, edited by W. A. J. Luxemburg and Stephan Körner. New Haven, CT: Yale University Press, 1979.

Robinson, Abraham. *Non-standard Analysis*. Rev. ed. Amsterdam, Netherlands: North-Holland, 1974.

Spalt, Detlef D. "Cauchys Kontinuum. Eine historiografische Annäherung via Cauchys Summensatz." *Archive for History of Exact Sciences* 56 (4) (2002): 285–338.

BERNARD BOLZANO

Rusnock, Paul. *Bolzano's Philosophy and the Emergence of Modern Mathematics*. Amsterdam, Netherlands: Editions Rodophi B. V., 2000.

Russ, Steve, ed. *The Mathematical Works of Bernard Bolzano*. New York: Oxford University Press, 2004.

CONTINUOUS FUNCTIONS

Youschkevitch, A. P. "The Concept of Function up to the Middle of the 19th Century." *Archive for History of Exact Sciences* 16 (1) (1976): 37–85.

CHARLES SANDERS PEIRCE

Ehrlich, Philip. "The Peircean Linear Continuum: A Surreal Model (Abstract)." *The Bulletin of Symbolic Logic* 11(4) (2005).

Herron, Timothy. "Charles Sanders Peirce's Theories of Infinitesimals." *Transactions of the Charles S. Peirce Society* 33 (3) (1997): 590–645.

Myrvold, Wayne. "Peirce on Cantor's Paradox and the Continuum." *Transactions of the Charles S. Peirce Society* 31 (3) (1995): 508–541.

Noble, N. A. Brian "Peirce's Definitions of Continuity and the Concept of Possibility." *Transactions of the Charles S. Peirce Society* 25 (2) (1989): 149–174.

Peirce, Charles Sanders. "Infinitesimals" (1900). In *Collected Papers of Charles Sanders Peirce*. Vol. 3, edited by Charles Hartshone and Paul Weiss. Cambridge, MA: Harvard University Press, 1935.

Peirce, Charles Sanders. "The Logic of Continuity." In *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, edited by Kenneth Laine Ketner. Cambridge, MA: Harvard University Press, 1992.

Potter, Vincent, and Paul Schields. "Peirce's Definitions of Continuity." *Transactions of the Charles S. Peirce Society* 13 (1) (1977): 20–34.

Putnam, Hilary. "Peirce's Continuum." In *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, edited by Kenneth Laine Ketner, with an Introduction by Kenneth Laine Ketner and Hilary Putnam. Cambridge, MA.: Harvard University Press, 1992: 37–54.

FUZZY REAL NUMBERS

Lowen, Robert. *Fuzzy Set Theory: Basic Concepts, Techniques, and Bibliography*. Dordrecht, Netherlands: Kluwer Academic, 1996.

Zadeh, Lotfi Asker. *Fuzzy Sets and Applications: Selected Papers*. Edited by R. R. Yager et al. New York: Wiley, 1987.

Philip Ehrlich (2005)

CONTINUUM PROBLEM

See *Set Theory*

CONTRACTUALISM

Contractualism, as a distinctive account of moral reasoning, was originally advanced by T. M. Scanlon in his widely admired paper "Contractualism and Utilitarianism" (1982) and was later elaborated on in detail in his book *What We Owe to Each Other* (1998). Drawing on an understanding of the significance of the social-contract metaphor that has its roots in Jean-Jacques Rousseau, rather than Thomas Hobbes, contractualism offers distinctive and interrelated answers to two central questions of philosophical theorizing about moral reasoning. First, what explains the importance of morality for people motivated to comply with the requirements of morality? Second, what kinds of reasons support judgments that particular acts or types of acts are right or wrong? Consequentialism provides what is undoubtedly the most familiar answer to this question. Contractualism seeks to provide a plausible alternative.

The contractualist account of why those who seek to comply with the requirements of morality care about being so guided presupposes a general approach to understanding the nature of value. The central idea of the presupposition is that to take something to be of value is to have reasons to regard it positively and reasons to act in certain ways with regard to it, some of which are required by the value of the thing in question. For instance, one's appreciation of the value of *The Last Supper* might take the form of planning trips to go and admire it, watching documentaries about it, reading scholarly works that deepen one's appreciation of it, worrying about its deterioration due to age, and debating the merits of various proposals to restore it with others who share one's passion.

Recognition of this value need not express itself in one's attitudes in these ways, though they are certainly rationally appropriate ways of responding to the value. But not all ways of engaging with something of value are optional. Some reasons for engaging with something of value in particular ways are reasons that all persons capable of making evaluative judgments are required to take account of in their practical deliberations. A person's indifference to *The Last Supper* does not alter the fact that he has reason not to ridicule or disparage it (even in his thoughts), not to urinate on it, not to attack it with a can of spray paint. These reasons, which are demanded by

respect for the value of the *The Last Supper*, apply to all individuals irrespective of their particular tastes and inclinations. Such reasons can be usefully characterized as *categorical reasons*.

Just as there are categorical reasons that flow from the value of the *The Last Supper*, so there are categorical reasons that flow from the value of human life. The distinctive value of human life, on the contractualist account, lies in the human capacity to assess reasons and justifications, to select among various reasons for wanting one's life to go a certain way, and thus to actively live and govern one's life (Scanlon 1998, p. 105). We have reason, then, to have certain attitudes toward, and give consideration to, the interests of others in our practical deliberations, namely, out of respect for the value of others as rationally self-governing beings. Failure to do so is a rational mistake, a failure to respond appropriately to all the relevant reasons for our behavior. This conclusion follows from the theory of value presupposed by contractualism and a specific characterization of the value of human life, neither of which are distinctively contractualist.

In answering why complying with morality matters to people who are morally motivated, contractualism holds that there are more than just rational reasons for respecting the value of another human being. Intuitively, there is a significant difference between failing to respect the value of a human and the kind of failure of respect exhibited by, for instance, proposing to film a rock video in the Sistine Chapel or building a McDonald's on the Great Wall of China. What accounts for the difference, according to contractualism, is the value of *mutual recognition*. Rational creatures living their lives in ways respectful of one another's value as rational creatures creates a special relation between them, a moral community of the kind that Immanuel Kant called the "Kingdom of Ends" and "a systematic union of various rational beings through common laws" (1902–, 4: 433). It is the kind of moral community that John Stuart Mill had in mind when he spoke of "unity with our fellow creatures." Respecting the value of others as persons, then, has a special importance for the morally motivated because they value the kind of relationship with others created by so living. This ideal of a moral community is at the heart of the contractualist characterization of moral reasoning.

Standards must guide individuals in their deliberations if they are to live on terms of respect for one another's value as persons. Contractualism characterizes these standards as principles for the general regulation of how individuals ought to deliberate in various situations. It asserts that those who care about the justifiability of their conduct toward similarly motivated others cannot reasonably reject these standards as a basis for informed, unforced, general agreement. Thus, principles that the morally inclined cannot reasonably reject play an important interpersonal role in regulating how individuals should relate to one another. They do so by fixing the attitudes and treatment that individuals are entitled to legitimately demand, and have demanded, out of respect for each other's value as rational creatures. In other words, these principles fix legitimate expectations concerning how individuals should deliberate in various situations. On this account, one person *wrongs* another when he fails to regulate his deliberations as the other is legitimately entitled to expect.

Whether or not a principle cannot be reasonably rejected is assessed according to the implications (broadly construed) of licensing individuals to reason as required and permitted by the proposed principle. Contractualism is both more restrictive and more permissive than consequentialism concerning what counts as a relevant implication of a proposed principle. It is more restrictive in that it does not regard as relevant facts about the aggregate value of the outcome likely to result from general compliance with the principle. The only relevant considerations are those that have to do with the implications of a principle for the life of an *individual* with a particular point of view. Different relevant implications can emerge from consideration of a principle from different points of view. This restriction on relevant implications rules out appeals to the aggregate value of an outcome as relevant for assessing a principle. One outcome may be worse than another with respect to aggregate value without being worse from the point of view of each individual. Contractualism and consequentialism are thus diametrically opposed on the relevance of considerations having to do with the aggregate value of potential outcomes.

Contractualism is more permissive than consequentialism in counting, as relevant, considerations that have nothing do with what is likely to happen as a consequence of individuals being licensed to treat one another in certain ways. Consider, for example, a principle that licenses a designated authority periodically to force randomly chosen individuals to serve as test subjects for dangerous medical experiments. In addition to the consequences for the lives of some unlucky individuals, contractualism will also allow as relevant consideration of the fact that such a principle would turn the bodies of individuals into a form of public property. That is, it would undermine the exclusive authority of individuals concerning decisions

about how their bodies are to be used, a prerogative that plays a fundamental role in an agent's understanding of his life as *his own*.

Assessing the validity of a principle requires both identifying the relevant considerations that ought to be taken into account in its assessment and combining them in a judgment about whether it is reasonable to reject the principle. Consequentialists claim that the right way to combine relevant implications of a principle is to aggregate their value. This sum is then compared to the aggregate value of the implications of possible alternatives. The valid principle is the one whose implications sum to the greatest aggregate value.

Contractualism adopts a different approach to this problem. Contractualism starts from the position that what the morally motivated person cares about is that his comportment toward another person be justifiable *to that person* as respectful of that person's value as a person. Justification *to another* requires that one's comportment toward the other be justified in light of what *that person cares about*. A principle is *justifiable to* a person, then, if he has reason to judge it to be justified (even if he himself does not recognize that reason) in light of the values that structure his particular point of view.

The central contractualist insight is that respect for the value of another as a person requires not merely that one take the implications of one's actions for that person's well-being into account, but that one be guided, in one's thinking about one's comportment toward that person, by a principle *justifiable to* that person. The impact of a possible principle on any person's well-being may be relevant to assessing the principle, but it will be so derivatively, as a consideration picked out as relevant by the master consideration of what is justifiable to that person. One's conduct may have negative implications for another, but if one has been guided by a principle justifiable to him, he has no grounds for complaint on the grounds that one has failed to give his interests the kind of consideration in one's deliberations that he is owed out of respect for his value as a person. If, in how one relates to another, one is guided by a principle that is justifiable to him, that principle can rightly be characterized as *authorized* by him.

Principles that no one can reasonably reject, then, enable individuals to relate to one another on terms of mutual respect for the value of one another as persons. They do so because a principle that no one can reasonably reject is justifiable to *any* individual from his point of view, provided at least that he values living with others on terms of mutual respect.

To arrive at a valid principle, we have to combine the implications of a proposed principle to arrive at an all-things-considered judgment about whether it is reasonable to reject the principle. At the heart of the contractualist approach to doing this is the requirement that a valid principle be justifiable to anyone from his own point of view. Assessing a proposed principle requires that one consider the point of view of the individual who stands to be most seriously burdened by it. Can such an individual reasonably reject the principle? On the contractualist account, that depends on the implications of a plausible alternative principle for those with other points of view. If every alternative principle to one that seriously burdens you will more seriously burden someone else, then you cannot reasonably reject the principle, as another individual's having to bear a burden that could be avoided by your bearing a lesser burden is justifiable *to you*. A valid principle is justifiable to the person who has the strongest reason for wanting to reject that principle.

This approach to how all the relevant implications of a principle are to be taken account of in an all-things-considered judgment of its validity stands in sharp contrast to that favored by consequentialist accounts. According to consequentialism, a principle that seriously burdens an individual can be justified by appeal to the aggregate value of the benefits secured under that principle for those with other points of view. Contractualism does not permit trade-offs of this kind among persons. A principle that seriously burdens you may secure benefits for others whose aggregate value outweighs the burden it places on you. But that fact has no bearing on whether the principle is justifiable *to you*, as it does not point to the viewpoint of another to whom any other principle, which does not so seriously burden you, would be justifiable. Under contractualism, our motivation for morality rules out aggregative considerations as relevant for the assessment of principles.

Consequentialism has a hard time making sense of commonsense prohibitions against treating others in certain abominable ways in circumstances where the consequences of doing so have great positive aggregate value. Nonconsequentialists argue that there is no problem in understanding the rationale for these prohibitions if one locates the basis for claims of wrongdoing in the very character of the prohibited way of treating others. One way of trying to articulate more clearly what the nonconsequentialist has in mind is the Kantian injunction to treat others, never as mere means, but always as ends in themselves. Contractualism, in locating the basis of a per-

son's claim to have been wronged in his having been treated in a way not justifiable *to him*, powerfully illuminates the compelling insight to which Kant's injunction draws our attention.

See also Constructivism, Moral; Discourse Ethics; Hobbes, Thomas; Locke, John; Rawls, John; Rousseau, Jean-Jacques; Social Contract.

Bibliography

Kant, Immanuel. *Gesammelte Schriften*. Vols. 1–29. Berlin: G. Reimer, 1902–. Page numbers of this edition are printed in the margin of translations in the *Cambridge Edition of the Works of Immanuel Kant*.

Scanlon, T. M. "Contractualism and Utilitarianism." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams, 103–128. Cambridge, U.K.: Cambridge University Press, 1982.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Rahul Kumar (2005)

CONVENTIONALISM

In the physical sciences, some very basic facts or principles appear to have a status that is difficult to categorize: not simply empirically discovered; not purely analytic (true by virtue of already established meanings); fundamental, but without quite being ordinary physical laws. Incompatible-looking alternative principles are conceivable; sometimes we can even see how an alternative physical framework could be built on them. Such principles are held, by some philosophers, to be true by *convention*. They are parts of our physical theories that had to be conventionally chosen by us over other incompatible postulates, whether or not we were overtly aware of this element of choice.

The most famous examples of putative conventional truths are to be found in our theories of space and time, and will be discussed below. But some are not directly related to space and/or time: in classical physics, Isaac Newton's famous 2nd law, F = ma, is an example. This law at first sight looks like it cannot be a convention, for surely, as the center of his mechanics, the 2nd law is far from true by stipulation. But in the Newtonian paradigm, the 2nd law served as ultimate arbiter of the questions (a) whether an external force is, in fact, acting on a given object; and (b) if so, what its magnitude was. While not necessarily immune to rejection or revision in the long run, this principle was nevertheless a postulate that helped constitute the meaning of other terms such as "force," and functioned as something *akin to* a definition, with a warrant *akin to* a prioricity.

Conventionalism of the Duhem-Quine sort holds that one can *always* maintain the truth of the 2nd law (or any other conventional truth), come what may. The backing for this claim comes from the *holism* of scientific theory testing (Pierre Duhem) or more generally, of conceptual frameworks (W. V. O. Quine): Since the things we hold true form an interconnected web, any one belief or postulate that faces apparent disconfirmation may be preserved as "true," as long as we are willing to make enough compensatory adjustments among other beliefs. (For example, the 2nd law may be held true in the face of motions that appear not to conform to it, as long as we are willing to postulate the existence of hitherto-undiscovered forces acting on the relevant bodies.)

However, one can usually imagine circumstances in which unbearable tensions arise in our conceptual frameworks from the insistence on retention of the putative conventional principle, and one is effectively forced to give it up. If this is right, then the original claim of conventionality—that the principle in question is a mere stipulation or definition—looks like something of an exaggeration. Are there in fact *any* choices in the creation of adequate physical theory that are genuinely free, conventional choices (as, for example, choice of units is), without being completely trivial (as, again, choice of units is)? Many philosophers have thought that space-time structures give us true examples of such conventionality. The debates over conventionalism form a significant fraction of twentieth-century philosophy of space and time, and work continues in a wave of recent books and papers (Friedman 1999, Ryckman 2004). We cannot hope to do justice to the depth and complexity of the arguments of the major thinkers here, but will limit ourselves to introducing the main themes and key arguments, directing the reader to further resources in the bibliography.

CONVENTIONALISM ABOUT SPACE

Before the eighteenth century all philosophers of nature assumed the Euclidean structure of space; it was thought that Euclid's axioms were true a priori. The work of Nikolai Lobachevsky, Bernhard Riemann, and Carl Friedrich Gauss destroyed this belief; they demonstrated, first, that consistent non-Euclidean constant-curvature geometries were possible, and later that even variably curved space was consistent and analytically describable. But what, exactly, does it mean to say that space is Euclidean or Riemannian? A naïve-realist interpretation can of course be

given: there exists a *thing*, space, it has an intrinsic structure, and that structure conforms to Euclid's (or Riemann's) axioms. But space, so described, is not observable in itself; only the material phenomena governed by physical laws are. When philosophers gave attention to this fact, they realized that our physical theories always contain assumptions or postulates that *coordinate* physical phenomena with spatial and temporal structures. Light rays in empty space travel in straight lines, for example; rigid bodies moved through space without stresses do not change their length; and so on. So-called *axioms of coordination* are needed to give meaning and testability to claims about the geometry of space.

The need for axioms of coordination seems to make space for conventionalism. For suppose that, under our old axioms of coordination, evidence starts to accumulate that points toward a non-Euclidean space (triangles made by light rays having angles summing to less than 180 degrees, for example). We could change our view of the geometry of space; but equally well, say conventionalists, we could change the axioms of coordination. By eliminating the postulate that light rays in empty space travel in straight lines (perhaps positing some "universal force" that affects such rays), we could continue to hold that the structure of space itself is Euclidean. According to the strongest sorts of conventionalism, this preservation of a conventionally chosen geometry can always be done, come what may. Henri Poincaré (1902/1952) defended the conventionality of Euclidean geometry; but he also conjectured that it would always be *simpler* to construct mechanics on assumption of Euclidean geometry. (Poincaré argued, on the basis of the work of Hermann von Helmholtz and Sophus Lie on free-mobility mechanics, that the possible geometries among which we must make a conventional choice are just three: Euclidean, Riemannian, and Lobachevskian. He thus did not consider the possibility of the variably curved space-time introduced by Albert Einstein in 1912).

Poincaré did not defend a wide-ranging, Duhem-Quine style of conventionalism; rather his view might better be thought of as neo-Kantian (as was also true of Hans Reichenbach in his first book on space and time [1920[). The Euclidean status of space is a convention that plays the role of a constitutive, a priori axiom with respect to mechanics and the rest of physics. By contrast, although he held it to be synthetic, arithmetic was *not* conventional for Poincaré: we have no choice but to regard it as true. But when it comes to the choice between Euclidean and (say) Lobachevsky geometry for real space, Poincaré's defense of the tenability of the Euclidean convention becomes basically an instance of the Duhem-Quine thesis: by making compensatory adjustments in our physics (specifically, introducing "universal forces" of the right sort), we can continue to hold that space is Euclidean even if direct measurements with rods and light-rays do not conform to that geometry. We will explore this idea further via Reichenbach, its most vigorous proponent in the twentieth century.

Reichenbach introduces the basic argument for conventionalism with an example that has become a classic:

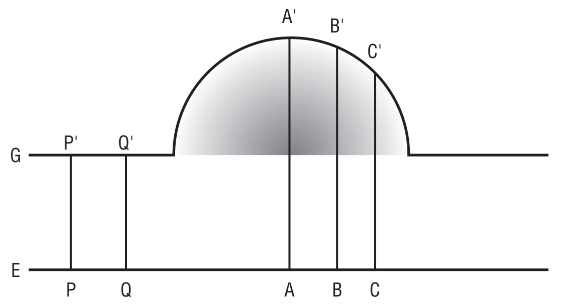

G represents a planar surface on which some 2-D beings live; we suppose that the surface is "really" flat almost everywhere, except for a non-flat hump centered around the point A'. The people on the plane can "know" that they have this hump in their space because of the way triangles' angles measure inside the hump. Using their measuring rods, they regard the segments A'B' and B'C' as equal in length.

Now we suppose that G is actually made of glass, and light shining from above casts shadows of everything on G onto the plane E below. People on E have their own measuring rods and so on. Let's suppose that, as it happens, the measuring rods on E behave exactly like the shadows of the E-rods: declaring AB congruent to BC, for example. Reichenbach has us suppose that there is a heating source under E that causes measuring rods to expand as they approach A, with no heat beyond the limits of the shadow of the hump. If the beings on E knew nothing of heat and how it expands measuring rods, what will they conclude? Like the G-people, they will conclude that their space has a non-Euclidean hump in it, centered on A.

The example brings to light the apparent impossibility—at least under the described circumstances—of determining whether one "really" lives in a curved space, or a flat Euclidean space with certain "universal" forces affecting things like measuring rods, light rays, and so on. (Reichenbach's "universal" forces affect every object in exactly the same way, and cannot be shielded out; they are clearly modeled on the force of gravity, for reasons we

will see below.) Now Reichenbach makes the key move in arguing for his conventionalism about physical geometry, and it is based on his verificationist empiricism: given that there is no way in principle to determine which of these is the case, we should reject the question itself as a mistake based on false presuppositions. There is no fact of the matter, in the case discussed, about whether G is "really" flat or rather "really" has a hump. A conventional choice must be made concerning whether to keep the geometry flat or not; after that, one can determine the presence (or absence) of universal forces as required.

We should note the irony here: in order to introduce his conventionalism, Reichenbach had to present us with hypothetical cases in which there *is* a nonconventional fact of the matter about the intrinsic geometry of space, then argue that we should disbelieve in these facts after all. Realists about space (or space-time) respond to Reichenbach precisely on this point: The fact that we cannot determine the geometry of space beyond any possibility of doubt, due to the logical possibility of other physical theories postulating a different geometry, does not entail that there *is no* fact of the matter—unless, of course, one subscribes to the most far-reaching of verificationist views of meaning and truth.

But Reichenbach is not quite so easily dismissed. The "intrinsic" geometries of G and E were introduced as a crutch for the imagination, to get us ready to see how the combination of a geometry G and a set of physical postulates about forces, F, are only testable (hence meaningful) together. Once the point is understood, we realize that our "intrinsic" geometries by themselves had no significance. The combination of G and F together, by contrast, is both meaningful and testable: the E-residents can certainly tell that their world is such that *if held to have a Euclidean space, then there are universal forces acting in the A-region; or if held to have no universal forces, then space is non-Euclidean in the region around A*. Which they decide to adopt is up to them, and the decision is a conventional one (perhaps based on simplicity or convenience).

ANTI-CONVENTIONALIST RESPONSES. Roberto Torretti (1983) and others have criticized Reichenbach's notion of a universal force, arguing that (a) gravity does not meet the criteria established by Reichenbach; and (b) physicists would never take such a stipulated, truly-unverifiable concept seriously. Recalling the analogy used by Reichenbach (and Poincaré before him) of the deformation of a measuring rod by heat, notice that real material objects respond differently to temperature changes: steel expands while a ceramic contracts when heated, for example. The differential response of some materials to the physical "force" of heat is crucial to its playing a significant role in physics. Gravity, too, is a force that affects different objects differently: a meter-stick made of steel with ball-shaped ends will change its length little, if at all, when held vertical in a gravity field like the Earth's; but a meter-stick made of foam rubber with steel ball-shaped ends will change significantly in the same gravity field. The force of gravity is indeed universal in the sense of (1) affecting all massive bodies equally *per unit of mass*, and in (2) being un-shieldable. But the universal forces Reichenbach discusses would appear to be rather different, affecting all bodies equally on a *per unit volume* basis, so as to change their sizes by exactly the same amount, regardless of internal constitution. Torretti argues that there are no forces in real physics that act in such a way.

Interestingly, though it was not known to Reichenbach, there are illustrations of potentially conventional elements that may be discerned in classical Newtonian gravitation theory, though not involving Reichenbachean universal forces. The arguably conventional choices are in fact two-fold (see Friedman 1983). First, one may add an arbitrary (constant) universal acceleration to every body in space: This acceleration changes no observable phenomena, and in fact is implemented simply by adding a term to the gravitational potential Φ. This extra term in the potential generates a universal gravitational force that accelerates every object at the same rate—seemingly a real-science example of Reichenbach's universal forces, but in fact different: The force does not deform any body's shape or behavior relative to other bodies, hence does not change the Euclidean geometry of space. Physicists customarily chose Φ so as to make such overall-acceleration equal to zero.

Secondly, mathematicians in the early twentieth century discovered how to transform Isaac Newton's gravity theory into a curved-space-time theory analogous to Einstein's General Theory of Relativity (GTR, about which more below): in this formulation, there are no gravity forces and instead the local geometry of space-time is curved, non-Euclidean. (Note however that this is only true of space-*time*, not space on its own—that remains flat, i.e., Euclidean.) Still, the example illustrates the conventionalist point: we might have had to choose whether to consider space-time flat/Euclidean, and let gravity be a universal force explaining why things do not always move on straight-line paths (geodesics); or instead, eliminate the "force" of gravity, allow that our space-time is curved, and hold that all bodies follow geodesics of the curved geometry of space-time. If we imagine that physics had

really turned out to show our world perfectly Newtonian, it is easy to see that we might find the conventionalist viewpoint attractive, compared to a realism that denies us the possibility of knowing what sort of space-time we live in, whether we are moving uniformly or instead with a frightful acceleration, and so forth.

Conventionalism and GTR. Discussions of conventionalism took a dramatic turn because of the work of Einstein. With its variably curved space-time, GTR obviously posed new challenges and opportunities for both sides on the conventionality of geometry. In the first half of the twentieth century GTR was widely viewed as vindicating a significant conventionalism or neo-Kantian "constitutive a priori" element in physics. In addition to Reichenbach, Ernst Cassirer, Moritz Schlick, and Adolf Grünbaum are some notable figures of twentieth-century philosophy who argued for the conventionality of space-time's geometry in the context of GTR (see Ryckman [2004] for an extensive and nuanced discussion of this early interpretive wave). Recent scholars have tended to be skeptical that any nontrivial conventionalist thesis is tenable in GTR; Friedman, Torretti and Hilary Putnam are prominent examples here.

Grünbaum, a student of Reichenbach's, recast the arguments for conventionalism in a non-epistemological form, more suited to the post-positivist climate of the 1960s and 1970s. He also brought forth a novel argument for the conventionality of geometry, based on the intrinsic *metrical amorphousness* of a continuous space. If space were composed of discrete atoms or chunks, it would thereby have a built-in metric. The distance between the ends of a meter stick would be determined by the number of space-atoms traversed by the line of its center, for example. But if, as most physical theories postulate, space(-time) is a continuum, then it cannot have any such built-in metric. (Grünbaum seems to be thinking of space as, intrinsically, just a topological manifold.) The metrical properties must be imposed extrinsically, by phenomena and bodies existing in the space (or space-time). And again, we must adopt conventions about which processes, bodies etc. are taken as constitutive for the geometry of space. See Grünbaum (1973) and Friedman (1983) for extensive discussion.

Einstein's GTR gave impetus to conventionalism in several ways; here we will mention just one. Consider Einstein's Equivalence Principle (EP), which says that a body that is uniformly accelerating (e.g., a rocketship) may consider itself as "at rest," but in the presence of a gravitational field that pulls everything downward. Conversely, according to the EP, a body freely "falling" under a gravitational force may equally well consider itself as "at rest" in a space without any gravitational forces. (The EP was, we see, implicit in our discussion of the two conventional elements in Newtonian gravity theory above.) A strong reading of this principle leads to the view that the existence or non-existence of a gravitational field is not a fact "out there" in the world, but rather something which we must arbitrarily decide. However, since gravitational fields (i.e., regions of local curvature of space-time) caused by bodies like planets and stars *can* be empirically distinguished from gravity-free regions—the EP is only true "locally," and to first approximation in small regions—the apparent freedom to choose turns out to be illusory.

Conventionality of Simultaneity. But it was in 1905, rather than 1915, that Einstein gave the greatest boost to conventionalism. In the astounding first few pages of "On the Electrodynamics of Moving Bodies," the paper that introduced the Special Theory of Relativity (STR) Einstein overthrew the Newtonian view of space-time structure—and, in passing, noted that *part* of the structure with which he intended to replace it had to be chosen by convention. That part was *simultaneity*. Einstein investigated the operational significance of a claim that two events at different locations happen simultaneously, and realized that it must be defined in terms of some clock synchronization procedure. The obvious choice for such a procedure was to use light-signals: Send a signal from event A for observer 1, have it be received and reflected back by observer 2 (at rest relative to 1), event B, and then received by 1 again at event C. The event B is then simultaneous with an event E, temporally midway between A and C.

Or is it? To suppose that it is, is to assume that the velocity of light on the trip from A to B is the same as its velocity from B to C (or, more generally, that light has the same velocity in a given frame, in all directions). This seems like a very good thing to assume. But can it be verified? Einstein thought not. All ways of directly measuring the one-way velocity of light seemed to require *first* having synchronized clocks at separated locations. But if this is right, we are going in circles: we need to know light's one-way velocity to properly synchronize distant clocks, but to know that velocity we need antecedently synchronized clocks.

To break the circle, Einstein thought we needed to make a conventional choice: We *stipulate* that event E is simultaneous with B (i.e., that light's velocity is uniform and direction-independent). Other choices are clearly possible, at least for the purposes of developing the

dynamics and kinematics of STR. Following Reichenbach, these are synchronizations with ε ≠ ½ (ε being the proportion of the round-trip time taken on the outbound leg only, freely specifiable between 0 and 1 exclusive). Adopting one of these ε ≠ ½ choices is equivalent to stipulating that the velocity of light is different in different spatial directions, without offering any physical reason for the difference, which some philosophers and physicists would find objectionable. It is also a recipe for calculational misery of a very pointless kind. But the Einstein of 1905, and many philosophers since then, thought that such a choice cannot be criticized as objectively *wrong*. Ultimately, they say, distant simultaneity is not only frame-relative, but partly conventional. It is important to see how different the situation is from Newtonian physics, in which there is no upper limit to the velocity of causal signals. In Newtonian physics, as long as we prohibit objects from moving " backward" in time, the existence of arbitrarily high velocities means that, given a specific event *here*, only one instant of time *there* can be chosen as simultaneous; that is, there is no scope for conventionality of simultaneity at all.

Many philosophers have been skeptical of the conventionality of simultaneity in STR. In 1967 Brian Ellis and Peter Bowman argued that slow clock transport offers a means of synchronizing distant clocks that is independent of the velocity of light. Their idea was this: in STR, of course, when a clock is accelerated from rest in a given frame up to some constant velocity, then decelerated to rest again at a distant location, there are the notorious time-dilation effects that prevent us from regarding the clock as having remained in synch with clocks at its starting point (the accelerated clock will have fallen behind the rest-clock—though this can, again, only be *directly* verified if it is brought back to its starting place for comparison with the rest clock). And calculation of the size of the effect depends on having established a distant-simultaneity convention (i.e., a choice of ε). So it looks as though carrying a clock from observer 1 to observer 2 will not let us break the circle.

But Ellis and Bowman noted that the time dilation effect tends to zero as clock velocity goes to zero, and this is independent of ε-synchronization. Therefore, an "infinitely slowly" transported clock allows us to establish distant synchrony, and measure light's one-way velocity. Infinitely slow transport is not, of course, a practical method for synchronizing clocks. The point is rather this: Since we can prove mathematically that the time dilation effect goes to zero as velocity of transport approaches zero, we can establish the *conceptual* point that the one-way velocity of light is non-conventional. Conventionalists were not persuaded, and the outcome of the fierce debate provoked by Ellis and Bowman's paper was not clear (Norton 1986).

In 1977 David Malament took up the conventionalist challenge from a different perspective. One way of interpreting the claim of conventionalists such as Grünbaum is this: The observable *causal structure* of events in an STR-world does not suffice to determine a unique frame-dependent simultaneity choice. By "causal structure" we mean the network of causal connections between events; loosely speaking, any two events are causally connectable if they *could* be connected by a material process or light-signal. In STR, the "conformal structure" or light-cone structure at all points is the idealization of this causal structure. It determines, from a given event, what events could be causally connected to it (toward the past or toward the future). Grünbaum and others believed that the causal structure of space-time by no means singles out any preferred way of cutting up space-time into "simultaneity slices."

Malament showed that, in an important sense, they were wrong. The causal/conformal structure of Minkowski space-time *does* pick out a unique frame-relative foliation of events into simultaneity slices. Or rather, more precisely, the conformal structure suffices to determine a unique relation of *orthogonality*. If we think of an ε-choice as the choice of how to make simultaneity slices relative to an observer in a given frame, then Malament showed that the conformal structure is sufficient to define a unique, orthogonal foliation, which corresponds to Einstein's ε = ½ choice. But most conventionalists do not view Malament's result as a refutation of their view (Janis 1983), in part because Malament's proof starts from assumptions that are arguably already in violation of the spirit of the conventionalist's view.

In recent years philosophers have begun to consider whether quantum theories may shed light on the debates concerning simultaneity; see Gunn and Vetharaniam (1995) and Karakostas (1997) for arguments for and against the idea that quantum field theory refutes the conventionalist claim. Bain (2000) shows that while it is true that one can formulate quantum field physics in coordinate systems corresponding to ε ≠ ½ simultaneity, by choosing the *generally* covariant formulation of the theory (i.e., a formulation that is valid, roughly speaking, in any coordinate system whatsoever), doing so requires the introduction of a new mathematical object or "field," whose role is basically to represent the standard orthogonal simultaneity slices. That is, while one can nominally

"choose" a simultaneity standard different from ε ≠ ½, the compensatory adjustments one is forced to introduce in order to make the theory work are such that one can see the "true" temporal structure lurking just under the surface. Conventionalists and anti-conventionalists disagree, of course, over whether the scare-quotes may be removed from the "true" in this verdict.

See also Philosophy of Physics; Space; Time.

Bibliography

Bain, J. "The Coordinate-Independent 2-Component Spinor Formalism and the Conventionality of Simultaneity." *Studies in History and Philosophy of Modern Physics 31B* (2000): 201–226.

DiSalle, R. "Spacetime Theory as Physical Geometry." *Erkenntnis* 42 (1995): 317–337.

Duhem, P. *The Aim and Structure of Physical Theory*. Translated by P. Wiener. Princeton, NJ: Princeton University Press, 1954.

Ellis, B., and P. Bowman. "Conventionality in Distant Simultaneity." *Philosophy of Science* 34 (1967): 116–136.

Friedman, M. *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton, NJ: Princeton University Press, 1983.

Friedman, M. *Reconsidering Logical Positivism*. Cambridge, U.K.: Cambridge University Press, 1999.

Gunn, D., and I. Vetharaniam. "Relativistic Quantum Mechanics and the Conventionality of Simultaneity." *Philosophy of Science* 62 (1995): 599–608.

Janis, A. "Simultaneity and Conventionality." In *Physics, Philosophy and Psychoanalysis*, edited by R, Cohen, 101–110. Reidel: Dordrecht, 1983.

Karakostas, V. "The Conventionality of Simultaneity in the Light of the Spinor Representation of the Lorentz Group." *Studies in the History and Philosophy of Modern Physics* 28 (1997): 249–267.

Malament, D. "Causal Theories of Time and the Conventionality of Simultaneity." *Noûs* 11 (1977): 293–300.

Norton, J. "The Quest for the One Way Speed of Light." *British Journal for the Philosophy of Science* 37 (1986): 118–120.

Poincaré, H. *Science and Hypothesis* (1902). New York: Dover Books, 1952.

Putnam, H. "An Examination of Grünbaum's Philosophy of Geometry." In *Philosophy of Science: The Delaware Seminar*, Vol. 2, edited by B. Baumrin, 205–255. New York: Interscience, 1963.

Redhead, M. "The Conventionality of Simultaneity." In *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, edited by J. Earman, A. Janis, G. Massey, and N. Rescher, 103–128. Pitsburgh, PA: University of Pitsburgh Press, 1993.

Reichenbach, H. *Philosophie der Raum-Zeit-Lehre.*Berlin and Leipzig: Walter de Gruyter, 1928. Translated by Maria Reichenbach and John Freund as *The Philosophy of Space and Time* (New York: Dover, 1957).

Reichenbach, H. *Relativitätstheorie und Erkenntnis apriori*. Berlin: Julius Springer1920. Translated by Maria Reichenbach as *The Theory of Relativity and A Priori Knowledge* (Berkeley and Los Angeles: University of California Press, 1965).

Ryckman, T. *The Reign of Relativity: Philosophy in Physics 1915–1925*. New York: Oxford University Press, 2004.

Torretti, R. *Relativity and Geometry*. Oxford, U.K.: Pergamon Press, 1983.

Carl Hoefer (2005)

CONVERSATIONAL IMPLICATURE

The concept of conversational implicature is due to the work of Paul Grice, and in particular to his paper "Logic and Conversation," which was delivered in 1967 and instantly became highly influential, although it was not published until 1975. A key goal of this paper was to defend the traditional logical understanding of connectives like *and* against what he saw as the excesses of ordinary language philosophy. He did this by drawing a sharp distinction between what is strictly speaking *said* and what is *conversationally implicated*. Consider sentence (1), below.

(1) Amanda and Beau fell in love and got married.

An utterance of (1) will typically suggest that the falling in love preceded the marriage. However, if *and* has its bare logical meaning, (1) may be true even if the marriage was initially loveless. According to Grice, (1) might indeed be true under these circumstances—because, strictly speaking, *and* contributes no more than its logical meaning to what is said. Grice claimed that the extra suggestion of temporal order was a conversational implicature. Conversational implicatures are an important part of communication, but (according to Grice) they have no effect on truth value. This is because they are not a part of what is strictly speaking said.

Grice argued that conversational implicatures arise from our adherence to (and presumption that others will adhere to) what he called the Cooperative Principle: "[m]ake your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (1989b, p. 26). (He took this principle to govern conversation, but he also took it that the principle would have correlates in other cooperative endeavors.) In its broadest outline Grice's idea was that we presume that others are being cooperative—and we will generally make whatever supplementary assumptions are required to maintain this presumption. This presumption is what

allows for the communication of conversational implicatures.

Grice takes it that we generally follow the Cooperative Principle by following four more specific maxims of cooperation (which, like the Cooperative Principle itself, he takes to have correlates in other endeavors):

> Quantity: Make your contribution as informative as (neither more nor less informative than) is required (1989b, p. 26).
>
> Quality: "Try to make your contribution one that is true": "do not say what you believe to be false" and "do not say that for which you lack adequate evidence" (1989b, p. 27).
>
> Relation: "Be relevant" (1989b, p. 27).
>
> Manner: "Be perspicuous": "1. Avoid obscurity of expression; 2. Avoid ambiguity; 3. Be brief (avoid unnecessary prolixity); 4. Be orderly" (1989b, p. 27).

Any of these maxims may play a role in generating conversational implicatures.

Grice characterizes conversational implicature as follows, and most scholars have followed him in this characterization.

> A man who, by (in, when) saying (or making as if to say) that *p* has implicated that *q* may be said to have conversationally implicated that *q*, provided that:
>
> (1) he is to be presumed to be following the conversational maxims, or at least the Cooperative Principle;
>
> (2) the supposition that he is aware that, or thinks that, *q* is required to make his saying or making as if to say *p* (or doing so in *those* terms) consistent with this presumption; and
>
> (3) the speaker thinks (and would expect the hearer to think that the speaker thinks) that it is within the competence of the hearer to work out, or grasp intuitively, that the supposition mentioned in (2) is required.
>
> (GRICE 1989, PP. 30–31)

To see how all of this machinery works in the generation of a conversational implicature, one must return to an utterance of (1).

(1) Amanda and Beau fell in love and got married.

Typically, the audience will assume that the speaker is being cooperative (so condition 1 is met). A cooperative speaker will follow the maxim of Manner, and the maxim of Manner dictates orderly presentation. If one is being orderly, one will generally present events in the order in which they occurred, so the audience must assume that the speaker thinks that Amanda and Beau's love predated their marriage. (Condition 2 is met). The speaker surely realizes that the audience is capable of working this out, so condition 3 is met. A typical utterance of (1), then, will implicate something like (1*).

(1*) Amanda and Beau fell in love and *then* got married.

There are a variety of ways that conversational implicatures may be generated. The above mechanism does not rely on the speaker's utterance being one whose explicit content would be uncooperative, but this latter sort of implicature (created by violating or flouting a maxim) is also an important one. It is crucial, for example, to irony, hyperbole, and understatement.

KINDS OF CONVERSATIONAL IMPLICATURE

Grice distinguished two main kinds of conversational implicature, *generalized* and *particularized*. Generalized conversational implicatures are ones that are *usually* carried by a certain word or phrase, while particularized conversational implicatures depend far more heavily on context. (Grice also allowed for nonconversational implicatures, such as *conventional* implicatures, with no role whatsoever for context. These are not of concern here.) The example discussed so far, involving (1), is that of a generalized conversational implicature: utterances of sentences involving *and* will usually carry the suggestion of temporal order that it is present in this example. Particularized implicatures depend far more heavily on context. Suppose you are hiring for a philosophy job, you ask me what I think of my student Charla, and I reply with nothing but (2).

(2) Charla reads a lot.

(2) is clearly not adequately informative, given your question—philosophy jobs require a great deal more than reading. By violating the maxim of Quantity in this way, I conversationally implicate that Charla is not a good philosopher. But in a different context, there would be no such implicature (consider, for example, a context in which I was asked for names of people who might like to join a book group). Because the implicature depends so heavily on context, it is particularized rather than generalized.

A great deal of work has been done, especially in linguistics, on generalized conversational implicatures, and

various mechanisms for their generation. Some of the most important work on this topic is by Laurence Horn (1972) and Stephen Levinson (2000). Generalized conversational implicatures have also played a particularly important role in philosophy. For Grice, these implicatures were especially important for their role in explaining certain problematic intuitions—like the intuition that (1) is false if the marriage preceded the love. He argued that generalized conversational implicatures are especially difficult to distinguish from what is strictly speaking said and that they may therefore give rise to mistaken intuitions. Philosophers since Grice have followed up on this thought by using generalized conversational implicatures in explanations of recalcitrant intuitions in a wide variety of areas.

It is worth noting that many cases that Grice took to be ones of generalized conversational implicature are very much disputed. For example, Robyn Carston (1991) has argued for a notion of saying (or, in her preferred terminology, *explicating*) on which the meaning of *and* is just what Grice would have taken it to be, yet nonetheless an utterance of a sentence like (1) *says* (rather than implicates) that Amanda and Beau's love preceded their marriage.

(1) Amanda and Beau fell in love and got married.

Carston's work on such examples is a part of a broader debate on the notion of what is said, which is not addressed here. But for some other approaches that also result in examples like the above counting as said, see Jeffrey King and Jason Stanley (2005), François Recanati (1989), and Dan Sperber and Dierdre Wilson (1986). For objections to reconstruing such generalized conversational implicatures as a part of what is said, see Kent Bach (2001), Laurence Horn (1992), Stephen Levinson (2000), Michael O'Rourke (2003), and Jennifer Saul (2002b).

TESTING FOR IMPLICATURE

Grice does not offer necessary and sufficient conditions that would allow one to test conclusively whether a given claim is a conversational implicature. However, he does offer certain necessary conditions for conversational implicature that can provide partial tests, and these have been widely accepted. Two especially important ones are cancelability and calculability. For more on testing for conversational implicatures, see Jerrold Sadock (1978).

CANCELABILITY. Because all conversational implicatures depend at least to some extent on context, it is always possible to cancel a conversational implicature by indicating either explicitly or implicitly that the implicatures should not be taken as present. For example, one might utter (1C):

(1C) Amanda and Beau fell in love and got married, but not in that order. Because (1C) contains an explicit cancellation of the conversational implicature standardly carried by (1), that implicature will not be carried by an utterance of (1C).

This contrasts with the case of saying. An attempt to "cancel" something that is said results only in a contradiction. To see this, consider an utterance of (1C*):

(1C*) Amanda and Beau fell in love and got married, but they didn't get married.

Applying this test shows us that the claim that Amanda and Beau got married is definitely not a conversational implicature, while the claim that their love preceded their marriage may well be.

While failure of the cancelability test does indeed indicate that one is not dealing with a conversational implicature, passing the cancelability test cannot be taken to decisively established that one is dealing with a conversational implicature. There are at least two reasons for this. First, a case of disambiguation may resemble one of cancellation, as in (3):

(3) He is in the grip of a vice, but not the mechanical kind.

Second, speaking loosely may result in an appearance of cancellation. Grice's own example (1989a, p. 44) concerns the fact that one may acceptably say, "Macbeth saw Banquo, even though Banquo was not there to be seen," even if it is known by all that Banquo merely hallucinated. Because one might be using the verb "see" in a loose way, this apparent cancellation does not indicate that utterances involving "see" merely implicate that what is seen exists.

CALCULABILITY. According to Grice, a putative conversational implicature is not a conversational implicature unless it is possible for audience to work out that the presence of the implicature is required in order to understand the speaker as cooperative. This calculation is meant to draw on knowledge of the linguistic meaning of the sentence uttered, the maxims of conversation, relevant background information, and the specific context. If no explanation can be given of how an audience would perform a calculation like this, a hypothesis that a particular conversational implicature is present must be rejected.

This necessary condition is also widely accepted. But its exact interpretation is a matter of some controversy. In

particular, there is disagreement over what it requires psychologically on the part of the hearer: must the hearer actually have distinct conscious representations of what is said and what is conversationally implicated (as argued by François Recanati in his work (1989)? Or are the requirements much more minimal (as argued by Kent Bach (2001), Manuel Garcia-Carpintero (2001), and Kenneth Taylor (2001)? The calculability requirement has proved very important: it has been used, for example, to argue for a more expansive conception of what is said (as in Recanati's work, as well as Robyn Carston's [1991]); to argue for and to object to particular invocations of conversational implicature; and (as in Wayne Davis's [1998]) to raise quite general concerns about the viability of Grice's theory of implicature.

GRICE'S TAXONOMY. It is very common to maintain that speaker meaning must divide exhaustively into what is said and what is implicated. Thus, any claim that the speaker means but does not say must be an implicature. (It need not, however, be a conversational implicature, since it could be a conventional implicature.) It is not entirely clear, however, that this is the right way to understand relationship between speaker meaning and implicature. For objections to this view of the relationship, see Kent Bach (1994) and Jennifer Saul (2002b).

See also Meaning; Pragmatics; Presupposition; Semantics.

Bibliography

Bach, Kent. "Conversational Impliciture." *Mind and Language* 9 (2) (1994): 124–161.

Bach, Kent. "You Don't Say?" *Synthese* 128 (2001): 15–44.

Carston, Robyn. "Implicature, Explicature, and Truth-Theoretic Semantics." In *Pragmatics: A Reader*, edited by Steven Davis. New York: Oxford University Press, 1991.

Davis, Wayne. *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge, U.K.: Cambridge University Press, 1998.

Garcia-Carpintero, Manuel. "Gricean Rational Reconstructions and the Semantics/Pragmatics Distinction." *Synthese* 128 (2001): 93–131.

Grice, Paul. "Further Notes on Logic and Conversation." In *Studies in the Way of Words*, 41–57. Cambridge, MA: Harvard University Press, 1989a.

Grice, Paul. "Logic and Conversation." In *Studies in the Way of Words*, 3–21. Cambridge, MA: Harvard University Press, 1989b.

Horn, Laurence. "On the Semantic Properties of Logical Operators in English." Diss. University of California at Los Angeles, 1972.

Horn, Laurence. "The Said and the Unsaid." In *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory*, 163–192. Columbus, OH: Ohio State University Department of Linguistics, 1992.

King, Jeffrey, and Jason Stanley. "Semantics, Pragmatics, and the Role of Semantic Content." In *Semantics versus Pragmatics*, edited by Zoltan Szabo, 111–164. New York: Oxford University Press, 2005.

Levinson, Stephen C. *Pragmatics*. Cambridge, U.K.: Cambridge University Press, 1983.

Levinson, Stephen C. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press, 2000.

O'Rourke, Michael. "The Scope Argument." *Journal of Philosophy* 100 (2003): 264–288.

Recanati, François. "The Pragmatics of What Is Said." *Mind and Language* 4 (1989): 295–329.

Sadock, Jerrold. "On Testing for Conversational Implicature." In *Syntax and Semantics*. Vol. 9: *Pragmatics*, edited by Peter Cole and Jerry Morgan. New York: Academic Press, 1978.

Saul, Jennifer. "Speaker Meaning, What Is Said, and What Is Implicated." *Noûs* 36 (2) (2002a): 228–248.

Saul, Jennifer. "What Is Said and Psychological Reality: Grice's Project and Relevance Theorists' Criticisms." *Linguistics and Philosophy* 25 (2002b): 347–372.

Sperber, Dan, and Dierdre Wilson. *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Taylor, Kenneth. "Sex, Breakfast, and Descriptus Interruptus." *Synthese* 128 (2001): 45–61.

Jennifer Saul (2005)

CONWAY, ANNE

(1631–1679)

Anne Conway (Anne Finch, Viscountess Conway), the English philosopher, was born in London. Her education was primarily informal and self-directed. Her associates included Henry More, Ralph Cudworth, Francis Mercury Van Helmont, William Harvey, and Robert Boyle, the latter two as physicians for her serious headaches. Later in life she scandalized More by becoming a Quaker.

WORK AND INFLUENCE

Conway's sole published work, *The Principles of the Most Ancient and Modern Philosophy,* published posthumously in 1690, shows the influence of the Cambridge Platonists, Kabbalism, and Neoplatonism. It criticized Thomas Hobbes, Benedict de Spinoza, and René Descartes, and influenced Gottfried Wilhelm Leibniz, who, during the year he was introduced to her work by Van Helmont in 1696, adopted her term *monad* and used it in a quite similar way (Merchant 1979). A notable difference between their uses of the term is that, while Leibniz's monads are purely spiritual, Conway's are both physical and spiritual.

Leibniz refers directly to Conway in his *New Essays* (book 1, chap. 1) as one of the better advocates of vitalism.

METAPHYSICS

Conway begins the *Principles* by asserting without proof the existence of a perfect God, the description of which is influenced by Neoplatonism and Kabbalism. Conway's God is one of three kinds of substance, each with its own essence. God is a complete, self-sufficient fountain that necessarily emanates Christ, the second kind of substance, and through the mediation of Christ, who shares some attributes with God, others with creatures, necessarily emanates creatures—the third kind of substance. Because emanative creation is creation "out of" God rather than "out of" nothing, creatures have a share of the divine attribute of life. Since all creatures are of the same kind of substance, they have a single essence, differing only modally from one another. Thus, spirit or mind and body are not "really distinct." There are many degrees of corporeity, and thus "a Thing may more or less approach to, or recede from the State and Condition of a body or a Spirit" (Conway 1982, p. 192). Conway draws the further conclusion that creatures are interconvertible: A horse, for example, can turn into a bird and spirits can turn into bodies (p. 177).

Not only God's creative act, but all of God's actions flow automatically from God's nature. Thus, God does whatever does not involve a contradiction. Conway's deity, like Leibniz', is timeless. Both Conway and Leibniz consider time to be relative to succession and motion; they consider succession and motion to be inferior analogues of eternity and the divine will, respectively, and thus to belong only to creatures (Conway 1982, p. 161).

Conway employs the concept of mediation, introduced in her account of creation, to explain action at a distance as well as causation between bodies and spirits. All created substances, in addition to sharing an essence, are interconnected by means of "Subtiler Parts," which are the "Emanation of one Creature into another." These mediated connections facilitate action at a distance and form "the Foundation of all Sympathy and Antipathy which happens in Creatures" (Conway 1982, p. 164). Conway offers, by contrast to the mechanical philosophy, a fairly direct account of the intelligibility of causation based on the concepts of similarity (or sympathy) and mediation. Similarity between cause and effect, as in the case of causation among bodies, renders causation directly intelligible, "because Things of one, or alike Nature, can easily affect each other." Mediation is required in the case of mind-body causation, because a soul is a "Spiritual Body" (pp. 214–215).

Since Conway regards interconnection as primitive, she requires no detailed explanations of causal interactions. Here she contrasts markedly with mechanistic philosophers' demands for explanations using motion and passive matter as primitives. Conway nonetheless incorporates causation by motion into her overall account of causation: Motion, especially vital motion, and divine emanation do not differ intrinsically from one another but are analogically related.

See also Boyle, Robert; Cambridge Platonists; Causation: Metaphysical Issues; Causation: Philosophy of Science; Cudworth, Ralph; Descartes, René; Harvey, William; Hobbes, Thomas; Kabbalah; Leibniz, Gottfried Wilhelm; More, Henry; Neoplatonism; Spinoza, Benedict (Baruch) de; Vitalism.

Bibliography

Conway, A. *The Principles of the Most Ancient and Modern Philosophy.* Edited by P. Lopston. The Hague: Nijhoff, 1982.

Duran, J. "Anne Viscountess Conway: A Seventeenth-Century Rationalist." *Hypatia* 4 (1989): 64–79.

Frankel, L. "Anne Finch, Viscountess Conway." In *A History of Women Philosophers,* Vol. 3, edited by M. E. Waithe. Dordrecht: Kluwer Academic, 1991.

Frankel, L. "The Value of Harmony." In *Causation in Early Modern Philosophy,* edited by S. Nadler. University Park: Pennsylvania State University Press, 1993.

Merchant, C. "The Vitalism of Anne Conway: Its Impact on Leibniz's Concept of the Monad." *Journal of the History of Philosophy* 17 (1979): 255–269.

Nicolson, M. H. *Conway Letters.* New Haven, CT: Yale University Press, 1930.

Lois Frankel (1996)

COPENHAGEN INTERPRETATION

The Copenhagen interpretation is the standard textbook interpretation of quantum mechanics. The term covers a range of divergent views, loosely related to Bohr's complementarity interpretation. The consensus of the physics community is that Einstein lost the debate to Bohr about the "completeness" of quantum mechanics at the Solvay conference of October 1927, and that Bohr's analysis of the experimental situation in quantum mechanics in terms of the notion of complementarity allows one to make sense of a universe that is indeterministic 'all the way down,' so that quantum states (that in general assign

probabilities between 0 and 1 to the outcomes of experiments) are as complete as they can be.

It is difficult to pin down the Copenhagen interpretation. Heisenberg—who seems to have coined the term "Copenhagen interpretation" (see Howard's "Who Invented the Copenhagen Interpretation" for a discussion)—concedes differences between his own position and Bohr's, but concludes that "we really meant the same." The term is generally taken to cover such radical views as Wigner's, that "the quantum description of objects is influenced by impressions entering my consciousness" and John Wheeler's notion of a "participatory universe":

> The dependence of what is observed upon the choice of experimental arrangement made Einstein unhappy. It conflicts with the view that the universe exists 'out there' independent of all acts of observation. In contrast Bohr stressed that we confront here an inescapable new feature of nature, to be welcomed because of the understanding it gives us. In struggling to make clear to Einstein the central point as he saw it, Bohr found himself forced to introduce the word 'phenomenon.' In todays words Bohr's point—the central point of quantum theory—can be put into a single, simple sentence. 'No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon.' It is wrong to speak of the 'route' of the photon in the experiment of the beam splitter. It is wrong to attribute a tangibility to the photon in all its travel from the point of entry to its last instant of flight. A phenomenon is not yet a phenomenon until it has been brought to a close by an irreversible act of amplification such as the blackening of a grain of silver bromide emulsion or the triggering of a photodetector. In broader terms, we find that nature at the quantum level is not a machine that goes its inexorable way. Instead what answer we get depends on the question we put, the experiment we arrange, the registration device we choose. We are inescapably involved in bringing about that which appears to be happening.
>
> (WHEELER 1983, PP. 184–185)

It is doubtful that Bohr would have endorsed Wheeler's formulation as a friendly amendment to complementarity. In a cautionary remark about misleading terminology, he writes:

> In this connection I warned especially against phrases, often found in the literature, such as "disturbing of phenomena by observation" or "creating physical attributes to atomic objects by measurements." Such phrases, which may serve to remind of the apparent paradoxes in quantum theory, are at the same time apt to cause confusion, since words like "phenomena" and "observations," just as "attributes" and "measurements," are used in a way hardly compatible with common language and practical definition.
>
> (BOHR 1948, P. 237)

THE REJECTION OF EINSTEIN'S REALISM

The common strand linking these different positions is the rejection of Einstein's realism—the "ideal of the detached observer," as Pauli put it somewhat pejoratively in a letter to Max Born (dated March 30, 1954). Einstein's position can be characterized by two informal independence principles: A separability principle and a locality principle. The separability principle is the principle that if two physical systems are spatially separated (or, in a relativistic setting, space-like separated), then each system can be characterized by its own properties, independently of the properties of the other system. That is, each system separately has its own "being-thus," as Einstein put it: A characterization in terms of certain properties intrinsic to the system, insofar as it is a separable system. The locality principle is the requirement that no influence on a system can directly affect another system that is spatially separated from it. In particular, a measurement performed on a system cannot alter any properties of another system that is spatially separated from it. The Copenhagen idea is that, in some sense (notwithstanding Bohr's discomfort with the terminology), the dynamical variables of quantum mechanics—the so-called "observables" of the theory—"only have values when you look," where the notion of "looking" is understood in a certain way (depending on the version: As involving the specification of a classically describable experimental set-up, or an interaction with a macroscopic measuring instrument that does not involve an ultimate conscious observer, or a measurement process that does involve the activity of a conscious observer, etc.). This claim is justified by citing examples of quantum interference characterized by Heisenberg's uncertainty relations, such as the double-slit experiment, or beam splitter experiments, or by appealing to the irreducible disturbance of a measured system in a quantum mechanical measurement interaction.

MEASUREMENT AND INTERFERENCE

Now it is generally recognized that the mere fact that measurements disturb what we measure does not preclude the possibility that observables have determinate values, or even that measurements might be exploited to reveal these values in suitably designed measurement contexts. (The "disturbance" terminology itself suggests the existence of determinate values for observables, prior to measurement, that are "disturbed" or undergo dynamical change in physical interactions.) And there is no warrant in the theory for interpreting the Heisenberg uncertainty relations for observables like position and momentum as anything more than a constraint on the possibility of preparing ensembles of systems in which these observables are simultaneously "sharp"–that is, as anything more than a constraint on the reciprocal distribution of the determinate values of these observables in quantum measurements.

Even interference phenomena, by themselves, say nothing about whether or not observables have determinate values in the absence of measurements, unless some interpretative principle is introduced. The usual story, in the case of a double-slit photon interference experiment, for example, is that you get the wrong distribution of hits on the screen behind the slits if you calculate the distribution on the assumption that each individual photon goes through one or the other of the two slits, when the photon is prepared in a quantum state that is represented algebraically in the theory as a linear sum (superposition) of a state in which the photon goes through slit 1 and a state in which the photon goes through slit 2. The photon is supposed to exhibit "wave-particle duality" and "go through both slits at once" to produce the characteristic interference pattern on the screen, where the photon finally manifests its presence as a particle. In passing through the slits, the photon behaves like a wave, a physical influence spread out over both slits, but in hitting the screen, it behaves like a particle, something localized at a point.

The loophole in the argument is the assumption of a specific link between attributing a determinate value to a quantum observable (like position, in the case of a photon going through one of two slits), and attributing a specific quantum state to the photon. This depends on an interpretative principle, the so-called "eigenvalue-eigenstate link," that a quantum system has a determinate value (an "eigenvalue") for an observable if and only if the quantum state is in a specific state called the "eigenstate" of the observable associated with the specific eigenvalue. If we reject this principle, then we can attribute a determinate value (an eigenvalue) to the observable associated with the photon going through slit 1 or slit 2, exclusively, without assigning the associated state (the eigenstate) to the photon. This is precisely what observer-free hidden variable interpretations like Bohm's theory accomplish.

Interference *per se* represents no obstacle to the simultaneous determinateness of noncommuting observables. The justification for assuming constraints on the simultaneous determinateness of quantum mechanical observables comes, rather, from the hidden variable 'no go' theorems of Kochen and Specker (1967) and Bell (1964), which severely limit the assignment of values to observables.

IS THE COPENHAGEN INTERPRETATION INSTRUMENTALIST?

For Bohr, a quantum "phenomenon" is an individual process that occurs under conditions defined by a specific, classically describable experimental arrangement, and an observable can be said to have a determinate value only in the context of an experiment suitable for measuring the observable. The experimental arrangements suitable for locating an atomic object in space and time, and for a determination of momentum-energy values, are mutually exclusive. We can choose to investigate either of these "complementary" phenomena at the expense of the other, so there is no unique description of the object in terms of determinate properties.

Summing up a discussion on causality and complementarity, Bohr writes:

> Recapitulating, the impossibility of subdividing the individual quantum effects and of separating a behaviour of the objects from their interaction with the measuring instruments serving to define the conditions under which the phenomena appear implies an ambiguity in assigning conventional attributes to atomic objects which calls for a reconsideration of our attitude towards the problem of physical explanation. In this novel situation, even the old question of an ultimate determinacy of natural phenomena has lost its conceptual basis, and it is against this background that the viewpoint of complemenarity presents itself as a rational generalization of the very ideal of causality.

(1949, p. 31)

Pauli characterizes Bohr's position this way:

> While the means of observation (experimental arrangements and apparatus, records such as spots on photographic plates) have still to be described in the usual 'common language supplemented with the terminology of classical physics,' the atomic 'objects' used in the theoretical interpretation of the 'phenomena' cannot any longer be described 'in a unique way by conventional physical attributes.' Those 'ambiguous' objects used in the description of nature have an obviously symbolic character.
>
> (1948, pp. 307–308)

The complementarity interpretation can be understood as the proposal to take the classically describable experimental arrangement (suitable for either a space-time or a momentum-energy determination) as defining what Bohr calls a quantum "phenomenon." A current approach is to refer to the macroscopic character of our measuring instruments, and to show that the nature of the interaction between such systems and the environment is of a specific sort that results in a physical process called "decoherence" that ensures the "classical" character of the instrument. Some version of this idea is incorporated into the Copenhagen interpretation, sometimes extended by claims such as Wheeler's. According to this view, then, the properties we attribute to a quantum object after a measurement depend partly on what we choose to measure, not solely on objective features of the system itself. To echo Pauli, the properties are "ambiguous" or merely "symbolic."

At first blush it would seem that the Copenhagen interpretation is thoroughly anti-realist, and in some contemporary versions straightforwardly instrumentalist. However, Don Howard in "Who Invented the Copenhagen Interpretation" has argued persuasively that Bohr's complementarity interpretation, as distinct from the Copenhgen interpretation, should be construed as a realist interpretation of quantum mechanics. For the contemporary philosophical debate on the Copenhagen interpretation, see Cushing (1994) and Beller (1999).

See also Bohr, Neils; Quantum Mechanics.

Bibliography

Bell, John S. "On the Einstein-Podolsky-Rosen Paradox." *Physics* 1 (1964): 195–200. Reprinted in *Speakable and Unspeakable in Quantum Mechanics*. Cambridge, U.K.: Cambridge University Press, 1987.

Beller, Mara. *Quantum Dialogue: The Making of a Revolution*. Chicago: University of Chicago Press, 1999.

Bohm, David, and Basil J. Hiley. *The Undivided Universe: An Ontological Interpretation of Quantum Mechanics*. London: Routledge, 1993.

Bohr, Niels. "Discussion with Einstein on Epistemological Problems in Modern Physics." In *Albert Einstein: Philosopher-Scientist*, edited by Paul A. Schilpp, 201–241. La Salle, IL: Open Court, 1949.

Bohr, Niels. "On the Notions of Causality and Complementarity." *Dialectica* 2 (1948): 312–319.

Born, Max, ed. *The Born-Einstein Letters*. London: Walker and Co: 1971.

Cushing, James T. *Quantum Mechanics, Historical Contingency and the Copenhagen Hegemony*. Chicago: University of Chicago Press, 1994.

Einstein, Albert. "Quantenmechanik und Wirklichkeit." *Dialectica* 2 (1948): 320–324. Translated as "Quantum Mechanics and Reality." In *The Born-Einstein Letters*, edited by Max Born, 170. See the accompanying letter from Einstein to Born dated April 5, 1948, and the letter dated March 18, 1948, on p. 164.

Fuchs, Christopher A., and Asher Peres. "Quantum Theory Needs No Interpretation." *Physics Today*, 70: 2000.

Heisenberg, Werner. "The Development of the Interpretation of Quantum Theory." In *Niels Bohr and the Development of Physics*, edited by Wolfgang Pauli, 12–29. London: Pergamon Press: 1955.

Heisenberg, Werner. "Quantum Theory and its Interpretation." In *Niels Bohr: His Life and Works as Seen by his Friends and Colleagues*, edited by S. Rozental, 94–108. New York: Wiley Interscience, 1967.

Howard, Don. "Who Invented the Copenhagen Interpretation?" Available at http://www.nd.edu/dhoward1/Papers.html.

Kochen, Simon, and E. P. Specker. "The Problem of Hidden Variables in Quantum Mechanics." *Journal of Mathematics and Mechanics* 17 (1967): 59–87.

Pauli, Wolfgang H. "Editorial on the Concept of Complementarity." *Dialectica* 2 (1948): 307–311.

Pauli, Wolfgang H. Letter to Max Born, dated March 30, 1954. In *The Born-Einstein Letters*, edited by Max Born, 218.

Wheeler, John A. "Law without Law." In *Quantum Theory and Measurement*, edited by John H. Wheeler and Wojciech H. Zurek, 183–213. Princeton, NJ: Princeton University Press, 1983.

Wigner, Eugene P. "Remarks on the Mind-Body Question." In *The Scientist Speculates*, edited by I. J. Good, 284–301. London: Heinemann, 1962.

Jeffrey Bub (2005)

COPERNICUS, NICOLAS
(1473–1543)

Nicolas Copernicus, or Mikolaj Kopernick, was a Polish clergyman, physician, and astronomer, and the propounder of a heliocentric theory of the universe. He was born at Torun (Thorn) on the Vistula. He studied liberal arts, canon law, and medicine at the universities of

Kraków (1491–1494), Bologna (1496–1500), and Padua (1501–1503) and received a doctorate in canon law from the University of Ferrara in 1503. Through the influence of his uncle, the bishop of Ermland, Copernicus was elected *in absentia* as a canon of the cathedral of Frauenburg in 1497. By 1506 he had returned to Poland, serving as physician to his uncle until 1512, when he took up his duties as canon. Copernicus's duties as canon involved him in the complex diplomatic maneuverings of the time and in the administration of the cathedral's large estates. In his own day he was more widely known as a physician than as an astronomer. He was one of the few persons in northeastern Europe to have a knowledge of the Greek language, and the one book he published without the urging of colleagues was a Latin translation of the poems of Theophylactus Simocatta, a seventh-century Byzantine poet. Copernicus's competence in economics was shown in some reports on money, presented to the Prussian diet, in which he anticipated a form of Gresham's law.

Copernicus's interest in astronomy was probably aroused at Kraków by the mathematician Wojciech Brudzewski and spurred on at Bologna by the astronomer Domenico Maria da Novara. Copernicus's first documented astronomical observation was made in Bologna in 1497. Twenty-seven such observations were used in his major treatise; others he recorded in the margins of books in his library. By 1514 he was so well known as an astronomer that he was asked by Pope Leo X to assist in the reform of the calendar, a task he declined because the motions of the sun and the moon had not yet been sufficiently determined.

Although Copernicus's major work, *De Revolutionibus Orbium Coelestium Libri IV,* was not published until 1543, the year of his death, he had been developing his theories at least from about 1512, the approximate date of his *Commentariolus* (a short outline of his system which he gave in manuscript copies to a few trusted friends). The first published account of his system was the *Narratio Prima* of his disciple and biographer (the biography is no longer extant), Georg Joachim Rheticus, in 1540. It was Rheticus who finally induced Copernicus to allow the publication of his major work.

LATE MEDIEVAL ASTRONOMY

The difference between Copernicus's theory and the then prevailing Ptolemaic system of astronomy can be stated briefly. The Copernican system was heliocentric rather than geocentric and geostatic; it placed the sun close to the center of the universe and Earth in orbit around the center, rather than postulating an immobile Earth at the center of the universe. But the full significance of this statement can be understood only via an examination of the *ad hoc* character of late medieval astronomy. Such late scholastic thinkers as Robert Grosseteste, Thomas Bradwardine, Jean Buridan, Nicholas Oresme, and Nicholas of Cusa had perceived the theoretical virtues and explanatory power of the heliocentric principle, as had Ptolemy himself long before. They understood the imperfections of the Ptolemaic techniques; yet they conceded that observational evidence did not clearly favor either theory—as was the case until the late sixteenth century. On scriptural grounds these thinkers accepted orthodox geocentrism; but they aired, more fully and deliberately than any of their predecessors, the arguments in support of terrestrial movement. They played *advocatus diaboli* with precision and imagination.

But prior to Copernicus astronomy was a piecemeal undertaking. Such problems as the prediction of a stationary point, or of an occultation, were dealt with one at a time, planet by planet. There was no conception that one planet's current stationary point might be related to another planet's later occultation. Techniques were employed as needed, and problem solving was not systematically integrated. Copernicus's theory changed this piecemeal approach forever. He effected a Kantian revolution in astronomy perhaps even more than Immanuel Kant effected a Copernican revolution in philosophy. Copernicus relocated the primary observational problem, that of explaining the apparent retrograde motions of the planets, by construing the motions not as something the planets "really" did "out there," but as the result of our own motion. Earth's flight around the sun makes other circling objects sometimes appear to move backward in relation to the fixed stars. Although either the Ptolemaic or the Copernican theory could be reconciled with sixteenth-century observations, Copernicus's view did not require investing those planets with queer dynamical properties, such as retrogradations-in-fact; a planet that actually halted, went into reverse, halted again, and then proceeded "forward" would be a strange physical object indeed. Rather, in Copernicus's view, all planets, including Earth, had the same kind of motion—a simple motion that explained the observed retrogradations.

It had been clear even to the ancients that the view that Earth was in the exact center of the universal system and that all celestial bodies moved about Earth in perfect circles could not generate predictions and descriptions even remotely close to the observed facts. In order to generate the right predictional numbers as well as tractable

orbital shapes, the Ptolemaic astronomers made a number of ad hoc assumptions. They moved Earth from the exact center of the planetary array; they used the geometrical center of the system as a reference point from which to calculate planetary distances; and they invented a third point, the *punctum aequans* (a mere computational device without physical significance, a device that Copernicus described as "monstrous"), around which the centers of the planetary epicycles described equal angles in equal times. No mechanisms known in nature or in art, however, have one center from which distances are determined, another from which velocities are determined, and a third from which observations are made. Moreover, the location of all these points and the choice of angular velocities around them were fixed arbitrarily and ex post facto simply to cope with each new observation as it turned up.

Even had Ptolemaic astronomy achieved perfection in predicting and describing, it was still powerless to explain planetary motion. One might ask how a theory that could describe and predict perfectly could in any way lack explanatory power; but Copernicus would have distinguished between the mere capacity of a theory to generate accurate numbers, and its further ability to provide an intelligible foundation for comprehending the phenomena studied. Even had the Ptolemaic system been able to predict accurately any future position of each moving point of light, Copernicus would still have asked what these points of light were, and what systematic mechanical interconnections existed between them.

An imaginative scholar, aware of the many difficulties posed by the Ptolemaic system as it had been developed over the years, and knowing (as Copernicus did) the accounts of ancient heliocentric theorists, might have only been expected to continue to seek improvements within the Ptolemaic system by incorporating promising heliocentric devices from his Scholastic predecessors (if he knew them) and from the ancients. Any gifted astronomer of Copernicus's day bent on improving astronomy "from the inside" would thus have had to take heliocentrism seriously.

In fact, Copernicus's books and Rheticus's summary might be viewed as an articulate and systematic expression of much late medieval planetary thinking. The ties with fifteenth-century Scholastic thought are everywhere apparent. But the primary insight of *De Revolutionibus,* although not novel, was boldly carried out and very much sharpened in detail. It was a comprehensive attempt to make the science of that day work better; it was not explicitly a plan for a new science of tomorrow. The dramatic consequences, largely unanticipated by Copernicus, are a tribute to his thoroughness as a student of nature and not to any self-conscious desire to level the orthodoxy around him.

THE COPERNICAN ALTERNATIVE

Copernicus was led to conclude that, in view of the plethora of epicycles required by the Ptolemaic system to account for the observed motion of the heavenly bodies, it must contain some basic error. He found that the assumption of a moving Earth, however absurd and counterintuitive it appeared, led to a much simpler and aesthetically superior system. Imagine yourself on the outer edge of a merry-go-round, sitting in a swivel chair. The constant rotation of the chair, when compounded with the revolution of the chair around the center of the merry-go-round, would generate—to say the least—complex visual impressions. Those impressions are compatible either with the motion as just described or with the supposition that it is the chair which is absolutely fixed and that all of the visual impressions stem from the motion of the merry-go-round about the chair-as-center and of a like motion of the walls of the building in which it is housed. The actual observations could be accounted for by either hypothesis. But what is easy to visualize in this example was extraordinarily difficult to comprehend in astronomical terms. That it was Earth that rotated and twisted, and revolved around the sun, seemed contrary to experience, common sense, and Scripture. Yet it was this simple alternative hypothesis that, for reasons demanded by astronomy, Copernicus espoused.

COPERNICUS'S REVOLUTION

Fundamentally, then, Copernicus argued that the observational intricacies of planetary motion were not real, but merely apparent. This argument made planetary motion simpler to comprehend but our own motion more intricate and therefore harder to believe. That was the fundamental objection to Copernicus's innovation.

But one must be quite clear about the nature of the theory. It was not a celestial dynamics, even in the sense that Johannes Kepler's theory of the causes of planetary motion (in terms of primitive spokes of force radiating from the sun) was a celestial dynamics. Copernicus, like his predecessors, was no astrophysicist; he was concerned with positional astronomy—the kinematics of planetary appearances, the motions of stellar lights against the black bowl of the sky and the underlying geometry that would, with a minimum of ad hoc assumptions, make those motions intelligible. So, both the *Almagest* and the *De Rev-*

olutionibus were concerned with plantetary kinematics exclusively—the latter in a systematic way, the former in the manner of a recipe collection. And even as a kinematic theory, Copernicus's theory was less adequate than those of Tycho Brahe and Kepler. He believed that the planets moved in perfect circles, an assumption shattered by Kepler's discovery of elliptical orbits. There is nothing in Copernicus to compare with Kepler's second law—that planets sweep out equal circumsolar areas in equal times. Nor is there anything to compare to Kepler's third law, correlating the time a planet requires to circle the sun with its distance from the sun. (And only when Kepler's three laws were added to the Galileo-Descartes law of inertia, and Isaac Newton's law of universal gravitation, was there developed a genuine celestial mechanics.) Copernicus's contributions consisted in a redeployment of the established elements of Ptolemaic positional astronomy. It is in this sense that he has been, and should be, viewed as the last great medieval astronomer.

SIMPLICITY OF COPERNICUS'S THEORY

Copernicus's theory was not psychologically simpler than competing systems. A moving Earth, and a sun and stars that do not "rise" and "circle" us, seemed contrary to experience. Also, a theoretical apparatus that linked all astronomical problems instead of leaving them to be faced one at a time could not constitute an easier system of calculation. Indeed, in the sixteenth century, heliocentrism was psychologically far more complex than the theories men were accustomed to.

Was Copernicus's conception perhaps simpler in that, as a formal theoretical system, it did not require primitive new ideas for each new problem or for the times when old problems led to difficulties? It invoked nothing like a *punctum aequans*; that is, it invoked fewer independent conceptual elements (primitive terms) merely to explain aberrant calculations than did other astronomies. But this point is insufficient to explain the sense in which Copernicus's system manifests "simplicity." Computational schemes had been proposed by Caelio Calcagnini and Geronimo Fracastoro that were simpler in that they were built on smaller sets of primitive notions. But they were so inadequate to the observational tasks of astronomy that it would have been as idle to stress their simplicity as it would be today to press for the theoretical adoption of John Dalton's atom because of its simplicity; the issue of simplicity does not arise except between theories that are comparable in explanatory and predictional power.

It has been urged that Copernicus's theory was numerically simpler, in that it required only 17 epicycles to the Ptolemaic 83. But the Ptolemaist, because he addressed his problems singly and without regard for the configurational complexities of taking all planets at once, never had to invoke 83 epicycles simultaneously. The number was usually no more than 4 or 5 per individual calculation.

This error is analogous to that involved in referring to a Ptolemaic "system" at all. Such a system results only from taking all individual calculating charts for the separate planets, superimposing them, running a pin through the centric "Earthpoint," and then scaling the orbits up or down so they do not collide. This scaling is determined by a principle of order wholly unconnected with any part of the Ptolemaic epicycle-on-deferent technique. In contrast, Copernicus's system locates the planets in a circumsolar order such that their relative distances from, and their angular velocities around, the sun are in themselves sufficient in principle to describe and predict all stationary points, retrograde arcs, occultations, and the brightening and dimming of the planetary lights. Thus, since Copernicus linked all planets, and invented systematic astronomy, he had to invoke all the epicycles his theory needed en bloc. The number of epicycles in any calculation would tend to be greater, not less, than that required in a corresponding Ptolemaic problem.

Copernicus's scheme is systematically simpler. It required more independent concepts than some others, but these were deductively interlocked. Copernicus was astronomy's Euclid. He constructed out of the disconnected parts of astronomy as he found it a systematic monument of scientific theory. The *De Revolutionibus* is psychologically and quantitatively more complex than anything that had gone before, but it was deductively simpler. What Euclid had done for geometry, and what Newton was later to do for physics, Copernicus did for positional astronomy.

IMPORTANCE OF THE THEORY

It has been argued that, as formalizations, the Copernican and Ptolemaic theories were strictly equivalent (D. J. de S. Price 1959), geometrically equivalent (A. R. Hall), even "absolutely identical" (J. L. E. Dreyer). But characterizing the theory as no more than "an alternative frame of reference plus some anti-Aristotelian philosophy" obscures the sense in which the heliocentric system and the geocentric systems of the sixteenth century were really equivalent. They were not equivalent in the sense that every consequence of the one was also derivable from the other.

Even when construed as mere geometrical calculations on paper, what the Ptolemaist would generate within his theory as corresponding to a stationary point in Mars's orbit is not congruent with what the Copernican would generate. The orbits were accorded different shapes in both theories, so points on those shapes, although viewed at the same angle from Earth, will not be superimposable. Nonetheless, every line-of-sight observation inferable within the one theory is completely inferable in the other. As positional astronomy, the two theories were observationally equivalent; no astronomer then could distinguish the two by comparing them with known facts. (Even today the Nautical Almanac is virtually a textbook of geocentric observation-points.) But the theories were neither formally equivalent nor physically equivalent, and certainly not absolutely identical. This is a difference that should make a difference to a philosopher.

With Sigmund Freud, man lost his Godlike mind; with Charles Darwin his exalted place among the creatures on Earth; with Copernicus man had lost his privileged position in the universe. The general intellectual repercussions of this fact are more dramatic than any consequences within technical astronomy, where one can speak of the Keplerian "revolution" but of not more than a Copernican "disturbance."

For the broad history of ideas, however, the implications of Copernicanism can hardly be exaggerated. Even religious revolutionaries such as Martin Luther and Philipp Melanchthon came to view Copernicus's position with abhorrence. His views challenged the literal interpretation of Scripture, the philosophical and metaphysical foundations of moral theory, and even common sense itself. The result was a massive opposition, learned and lay, to the reported ideas of Copernicus. It was the slow, sure acceptance of the technical *De Revolutionibus* by natural philosophers that ultimately quieted the general clamor against heliocentrism. Without the riotous reaction against it, Copernicus's book might have been but a calm contribution to scholarship somewhat like Pierre Simon de Laplace's *Mécanique céleste.* In the sixteenth and seventeenth centuries, however, the name Copernicus became a battle cry against the establishment in religion, in philosophy, and in natural science. It was a cry amplified in the world of wider scholarship and theology—far beyond Copernicus's original pronouncements. For Copernicus epitomized the well-trained, thorough, and rigorous sixteenth-century natural philosopher. He sought to make the theories he had inherited work better than when he found them. The history of ideas is charged with such figures. The difference is that Copernicus was presented with a theory that was incapable of further internal revision and improvement. The only recourse was fundamental overhaul—the consequences of which we still feel today.

See also Bradwardine, Thomas; Buridan, Jean; Darwin, Charles Robert; Freud, Sigmund; Grosseteste, Robert; Kant, Immanuel; Kepler, Johannes; Laplace, Pierre Simon de; Luther, Martin; Melanchthon, Philipp; Nicholas of Cusa; Oresme, Nicholas.

Bibliography

WORKS BY COPERNICUS

Edward Rosen is the editor and translator of *Three Copernican Treatises* (New York: Columbia University Press, 1939), which contains a translation of the *Commentariolus,* a "Letter against Werner," and Rheticus's *Narratio Prima.*

De Revolutionibus Orbium Coelestium Libri IV is available in a facsimile text and a Latin printed text edited with notes by S. Kubach (1944). The translation by C. G. Wallis in *Great Books of the Western World,* Vol. 16 (Chicago: Encyclopaedia Britannica, 1952) is to be used with caution.

WORKS ON COPERNICUS

Armitage, Angus. *Copernicus, the Founder of Modern Astronomy.* London: Allen and Unwin, 1938.

Hanson, N. R. "Contra-Equivalence." *Isis* 55 (1964).

Hanson, N. R. "The Copernican Disturbance and the Keplerian Revolution." *Journal of the History of Ideas* 22 (1961): 169–184.

Kuhn, Thomas. *The Copernican Revolution.* Cambridge, MA: Harvard University Press, 1957.

Mizwa, S. P. *Nicholas Copernicus, 1543–1943.* New York: Kosciuszko Foundation, 1943.

Price, D. J. de S. "Contra-Copernicus." In *Critical Problems in the History of Science,* edited by Marshall Claggett. Madison: University of Wisconsin Press, 1959.

Prowe, Ludwig. *Nicolaus Coppernicus,* 3 vols. Berlin, 1883–1885.

Rudnicki, J. *Nicholas Copernicus.* Translated by B. W. A. Massey. London: Copernicus Quatercentenary Celebration Committee, 1943.

Norwood Russell Hanson (1967)

CORBIN, HENRY
(1903–1978)

The French Islamicist and philosopher Henry Corbin was born in Paris on April 14, 1903. He studied with such French scholars as Étienne Henry Gilson, Emile Brehier, and Louis Massignon. To expand the scope of his studies, he learned over a dozen classic and modern languages. His interest in philosophy took him to Germany where he made an acquaintance with Ritter, Karl Löwith, Alexan-

der Kojve, and Martin Heidegger. He translated several works of Heidegger into French, including *What Is Metaphysics*.

Corbin's main philosophical interest during the 1930s was the relationship between philosophy and mysticism. This was a major factor in his decision to study Islamic philosophy. Louis Massignon, then the head of Islamic studies at Sorbonne, introduced him to the works of the twelfth century Muslim philosopher Shihāb al-Dīn Yahyā al-Suhrawardī (d. 1191). Suhrawardī founded a philosophical school called the School of Illumination and sought to combine philosophical analysis with mystical experience—a theme that runs through Corbin's works.

In 1940, during World War II, Corbin went to Istanbul, Turkey, to study the manuscripts of Suhrawardī's works. He stayed in Turkey for the next five years. Then in 1945 he went to Tehran, Iran, where he founded an institute of Iranian studies under the French-Iran Institute. This is the beginning of Corbin's lifelong involvement with what he came to call "Persian Islam" (*islam iranien*). Iran became a spiritual birthplace for him.

Corbin was a prolific writer. Even though his scholarly works are mostly devoted to the philosophical exposition of Islamic or "Oriental" thought, they are permeated by his lifelong concern to resuscitate the mystico-philosophical outlook of such mystical philosophers as Suhrawardī, Ibn al-ʿArabi, Mullā Ṣadrā, and Emmanuel Swedenborg. One key term in Corbin's thought is *mundus imaginalis* (the *ʿalam al-khayal* of the Muslim philosophers). Not to be confused with "imaginary" world, *mundus imaginalis* refers to an intermediary stage between the purely intellectual and empirical worlds. For Corbin, this is the realm of angels and spiritual visions where sensible forms become immaterial and intelligible forms take on an "imaginal" character and dimension. This is where heaven and earth meet in the metaphysical sense of the term. Corbin believed that the European intellectual tradition has lost sight of this crucial concept, severing its relation with the "angelic world" and lending religious justification to the Cartesian dualism of body and soul.

In his readings, Corbin followed the tradition of spiritual hermeneutics (*ta'wil*), "returning" words to their original meanings and thus going back to the "beginning." He called himself a phenomenologist in the sense of "removing the veils of ignorance," (*kashf al-mahjub*). In his philosophical quest, Corbin gave some of the best examples of what is sometimes called comparative philosophy, and his immense knowledge of European and Asian philosophies allowed him to do much more than simply compare or juxtapose different ideas and concepts.

See also Illuminationism; Nasr, Seyyed Hossein.

Bibliography

Jambet, Christian, ed., *Cahiers de l'Herne Henry Corbin*. Paris: Herne, 1981.

Nasr, Seyyed Hossein, ed. *Mélanges offerts à Henry Corbin*. Tehran, Iran: Institute of Islamic Studies, McGill University, Tehran Branch, 1977.

Shayegan, Daryush. *Henry Corbin: La topographie spirituelle de l'islam iranien*. Paris: Editions de la Difference, 1990.

Ibrahim Kalin (2005)

CORDEMOY, GÉRAUD DE

(1626–1684)

Géraud de Cordemoy, a French lawyer, historian, and philosopher, was born in Paris. From about 1657 on, he frequented Cartesian circles and soon developed some distinctive (but seemingly un-Cartesian) theses, such as atomism and occasionalism. He met Bishop Jacques Bénigne Bossuet in October 1668 and through him became *lecteur ordinaire* to the dauphin in 1673. During this period he was engaged, at Bossuet's order, on a biography of Charlemagne, which was understood as involving a complete history of the French monarchy. He did not finish the work, though it was completed and published posthumously in two volumes by his oldest son, Louis-Géraud, along with other unpublished manuscripts and a three-volume collection of his works. Cordemoy was elected to the Académie Française in 1675. His fame rests on his attempts to extend Cartesian philosophy to the fields of language and communication, in *Discours physique de la parole* (1668), and on his advocacy of Cartesian orthodoxy, such as the doctrine of animal-machines and the consistency of Cartesianism with Genesis in *Lettre écrite à un scavant religieux* (1668), but above all, Cordemoy is known for the views he defended in *Le discernement du corps et de l'ame* (1666).

For René Descartes, the principal attribute of body is extension and that of mind is thought. The first half of this tenet was thought to entail the impossibility of both atoms and the void: atoms, because extension was considered indefinitely divisible, and the void, because space was identified with the extension of bodies. Cordemoy offered a variation of Cartesian mechanical philosophy—everything in the physical world is explained in terms of

the size, shape, and motion of particles—but one that required atoms and the void. He rejected the indefinite division of body and the Cartesian identification of space with extension. He distinguished body and matter, matter being an assemblage of bodies, and claimed that bodies as such were impenetrable and could not be divided or destroyed.

He even criticized the Cartesian principle of individuation of bodies as shared motion, pointing out that a body at rest between other bodies would have to constitute a single body with the other bodies, even though one has a clear and natural idea of a body at rest between other bodies. Cordemoy proposed that shape, rather than motion, distinguishes the indivisible atoms. Cordemoy further reduced all forms of motion to local motion, arguing that no body has motion by itself. The prime mover—ultimately God—is necessarily a mind, though one's mind is not capable to begin, stop, or accelerate a motion; it can only change its direction. According to Cordemoy, this change of direction or "determination" of motion is not a change in the quantity of motion. The application of this analysis to the problem of the union of the soul and body led Cordemoy to occasionalism: Changes in the soul occur on the occasion of motions in its body, and vice versa.

Cartesians, such as Robert Desgabets and Nicolas Malebranche, criticized Cordemoy's atomism, though others, such as Gottfried Wilhelm Leibniz, praised it. Leibniz rejected Cordemoy's physical atoms, but spoke of Cordemoy as recognizing something of the truth when he tried to save the substantial unity of bodies; according to Leibniz, something lacking extension is required for the substance of bodies, otherwise there would be no source for the reality of phenomena. Other aspects of Cordemoy's philosophy were options generally discussed at the time; Cartesians, such as Johannes Clauberg and Louis de La Forge, had already proposed versions of occasionalism. But Cordemoy produced a more systematic treatment, starting from the physics of motion, which anticipated a number of Malebranche's theses, including the soul knowing itself directly by consciousness, but without a clear idea, and the existence of the body being known only indirectly, as the object of faith.

See also Cartesianism.

Bibliography

WORKS BY CORDEMOY

Le monde de Mr Descartes ou le traité de la lumière et autres principaux objets des sens, avec un discours de l'action des corps et un autre des fièvres, composés selon les principes du même auteur. Paris: Jacques le Gras, 1664. The *Discours de l'action des corps* is the second *discours* of *Le discernement du corps et de l'ame*.

Le discernement du corps et de l'ame, en six discours. Paris: F. Lambert, 1666, with editions in 1670, 1671, 1679, 1680, and 1683.

Discours physique de la parole. Paris: F. Lambert, 1668, with editions in 1671, 1677, and 1704. English translation: *A Philosophical Discourse concerning Speech*. London: John Martin, 1668.

Copie d'une lettre écrite à un scavant Religieux de la Compagnie de Jesus: pour montrer, I. que le Systeme de Monsieur Descartes & son opinion touchant les bestes n'ont rien de dangereux, II. et que tout ce qu'il en a écrit semble estre tiré du premier Chapitre de la Genese. Paris: N.p., 1668. English translation: *A Discourse Written to a Learned Friar*. London: Moses Pitt, 1670.

Tractatus physici duo: I, de corporis et mentis distinctione; II, de loquela. Geneva: J. Pictetum, 1679. This is a Latin translation of *Le discernement du corps et de l'ame* and *Discours physique de la parole*.

Histoire de France. 2 vols. Published by Louis-Géraud de Cordemoy. Paris: J.-B. Coignard, 1685–1689.

Dissertations physiques sur le discernement du corps et de l'ame. 2 vols. Paris: Veuve D. Dion, 1689–90. This contains *Le discernement du corps et de l'ame*, *Discours physique de la parole*, and *Lettre écrite à un scavant religieux*.

Divers traitez de métaphysique, d'histoire et de politique. Published by Louis-Géraud de Cordemoy. Paris: Veuve de J.-B. Coignard, 1691.

Les oeuvres de feu monsieur de Cordemoy. 3 vols. Published by Louis-Géraud de Cordemoy. Paris: C. Remy, 1704.

Oeuvres philosophiques, edited by Pierre Clair and François Girbal. Paris: Presses Universitaires de France, 1968.

WORKS ABOUT CORDEMOY

Balz, Albert G. A. *Cartesian Studies*. New York: Columbia University Press, 1951.

Battail, Jean-François. *L'avocat philosophe Géraud de Cordemoy (1626–1684)*. The Hague: Nijhoff, 1973.

Nicolosi, Salvatore. *Il dualismo da Cartesio a Leibniz: Cartesio, Cordemoy, La Forge, Malebranche, Leibniz*. Venice, Italy: Marsilio, 1987.

Prost, Joseph. *Essai sur l'atomisme et l'occasionalisme dans la philosophie cartésienne*. Paris: Henry Paulin, 1907.

Scheib, Andreas. *Zur Theorie individueller Substanzen bei Géraud de Cordemoy*. New York: Peter Lang, 1997.

Roger Ariew (2005)

CORDOVERO, MOSES BEN JACOB

(1522–1570)

Moses ben Jacob Cordovero, the Jewish legalist and mystic, was the outstanding systematizer of the kabbalah. The place of his birth is not known; his father probably was

among the Jews expelled from Cordova, Spain. Cordovero's career centers in Safad, the little town in Palestine that had a period of glory in the sixteenth century. Here, after studying with three distinguished rabbinical teachers—Joseph Caro, Jacob Berab, and Moses di Trani—he was ordained at an early age and became one of the leading figures of the community. His kabbalistic studies were begun at the age of twenty, under the direction of his brother-in-law, Solomon Alkabez, and became the major concern of the remainder of his life. Isaac Luria, who was to become the key figure in a new, more theosophic version of kabbalistic teachings, was originally a pupil of Cordovero.

Cordovero wrote at least ten important kabbalistic books, of varying lengths, during his lifetime. From the philosophic point of view, the greatest of these was *Pardes Rimmonim* (A Garden of pomegranates; first printed at Kraków, 1591). This large book attempted to present a systematic exposition of kabbalistic ideas and to justify them by deductive rational argumentation instead of the usual methods of kabbalistic exegesis. The word *Pardes* (*PRDS*) in the title acrostically represents the four modes of interpretation of Scripture: *peshat,* literal interpretation; *remez,* allegorical, or hinting, interpretation; *derash,* homiletical interpretation; and *sod,* mystical interpretation. Among the subjects emphasized by Cordovero in his treatment are God's unity, God's will, God's knowledge and thought, God's wisdom and goodness, God's many names, and God's relation to creation; the emanations (*sephirot*), both individually and collectively, the reason for there being precisely ten emanations, and the mystery of their multiplicity in unity; the Shekinah; angels; soul; being; prophecy; the relation of correspondence between the upper and lower worlds and the necessity of each to the other; the Law and the commandments; the mysteries of the Law; the secrets of the Hebrew alphabet; man and Israel; righteousness; time; freedom and bondage; the service of God. Cordovero was one of the first writers to stress the idea of *zimzum,* the voluntary self-shrinkage of God to make room for the material world.

Because of his rational discussion of all these subjects and his successful philosophic justification of them, in terms of his own presuppositions, Cordovero may well be regarded as the climactic figure of the earlier period of kabbalistic speculation. To what extent he was also intrigued by the more practical or "magical" aspects of kabbalah, we cannot tell.

See also Kabbalah.

Bibliography

WORKS BY CORDOVERO

The Palm Tree of Deborah. Translated by Louis Jacobs. New York: Hermon Press, 1974.

WORKS ON CORDOVERO

Idel, M. *Language, Torah, and Hermeneutics in Abraham Abulafia.* Albany: State University of New York Press, 1989.

Robinson, Ira. "Moses Cordovero and Kabbalistic Education in the Sixteenth Century." *Judaism* 39 (1990): 155–162.

Robinson, Ira. *Moses Cordovero's Introduction to Kabbala.* New York: Yeshiva University Press, 1994.

Sack, Bracha. "The Influence of Cordovero on Seventeenth-Century Jewish Thought." In *Jewish Thought in the Seventeenth Century,* edited by Isadore Twersky and Bernard Septimus. Cambridge, MA: Harvard University Press, 1987.

Sack, Bracha. "Some Remarks on Rabbi Moses Cordovero's Shemu'ah be-'Inyan ha Gilgul." In *Perspectives on Jewish Thought and Mysticism,* edited by A. Ivry. E. Wolfson and A. Arkush. Amsterdam: Harwood Academic, 1998.

Scholem, G. *On the Mystical Shape of the Godhead: Basic Concepts in the Kabbalah.* New York: Schocken, 1991.

Tirosh-Samuelson, H. "Philosophy and Kabbalah." In *Cambridge Companion to Medieval Jewish Philosophy,* edited by D. Frank and O. Leaman, 218–257. Cambridge, U.K.: Cambridge University Press, 2003.

Wolfson, E. "Jewish Mysticism: A Philosophical Overview." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman, 450–498. London: Routledge, 1997.

J. L. Blau (1967)
Bibliography updated by Oliver Leaman (2005)

CORRESPONDENCE THEORY OF TRUTH

The term "correspondence theory of truth" has circulated among modern philosophical writers largely through the influence of Bertrand Russell, who sets the view (which he himself adopts) that "truth consists in some form of correspondence between belief and fact" against the theory of the absolute idealists that "truth consists in *coherence*," that is, that the more our beliefs hang together in a system, the truer they are.

ANCIENT AND SCHOLASTIC VERSIONS OF THE THEORY

The origins of the word *correspondence*, used to denote the relation between thought and reality in which the truth of thought consists, appear to be medieval. Thomas Aquinas used *correspondentia* in this way at least once, but much more often he used other expressions and preferred most of all the definition of truth that he attributed to the ninth-century Jewish Neoplatonist Isaac Israeli:

Veritas est adaequatio rei et intellectus (truth is the adequation of things and the intellect). At one point he expanded this to *adaequatio intellectus et rei, secundum quod intellectus dicit esse, quod est, vel non esse, quod non est*. This is an echo of Aristotle's "To say of what is that it is not, or of what is not that it is, is false; while to say of what is that it is, and of what is not that it is not, is true." Other Scholastics sometimes said that a proposition is true when and only when *ita est sicut significat* ("the thing is as signified"); this too is in line with the Aristotelian account, in which "is" is not restricted to the meaning "exists"—the definition also covers the point that to say of what is so that it is not so, or of what is not so that it is so, is false; while to say of what is so that it is so, and of what is not so that it is not so, is true. This simple statement is the nerve of the correspondence theory; we shall continually return to it.

PLATO. Aristotle did not originate the correspondence theory but took it over from Plato's *Sophist*. There it was developed with an eye on a rejected alternative—not the coherence theory, which is a comparatively late invention (G. E. Moore is probably correct, in his "Truth" article in Baldwin's *Dictionary*, in tracing its vogue to Immanuel Kant), but one that we may call the existence theory, which also crops up in the *Theaetetus*. In this latter dialogue Socrates tries to find what differentiates true from erroneous belief, and the first suggestion he considers is that whereas true belief is directed toward what is, false belief is directed toward what is not. This view is rejected on the ground that just as to see or hear what is not is to see or hear nothing, and to see or hear nothing is just not to see or hear at all, so to "think what is not" is to think nothing, and that is just not to think at all, so that erroneous thought, on this view, would just not be thinking at all.

The same theory is considered in the *Sophist*, but here an alternative is put forward. Thought is compared with speech (it is the soul's dialogue with itself), and the important thing about speech is that in order to be true or false it must be complex—only complete statements are true or false, and these must consist of both nouns and verbs. (These points are also stressed by Aristotle.) As simple examples of complete statements, Plato gives "Theaetetus is-sitting-down" and "Theaetetus is-flying." The first of these is true because Theaetetus is sitting down, and the second is false because he is not flying. This escapes the difficulties of the existence theory because it abandons the suggestion that thinking is a simple direction of the mind toward an object—if it were that, its verbal expression would not have to be a complete sentence but could be just a name—and so opens up the possibility for thinking to be erroneous even though what is thought about, such as Theaetetus, is perfectly real.

The existence theory, however, dies hard and has continued to maintain itself, not merely as a rival to the correspondence theory but even more as something that theory is in constant danger of becoming. (The two views continually oscillate, for example, in the early work of Russell and Moore.) It is easy to equate the complexity of thinking with its having a complex object—for instance, Theaetetus's-sitting-down or Theaetetus's-flying—which exists if the thought is true and does not if it is not.

ARISTOTLE. There is no trace of the above slide or degeneration in Aristotle, nor even of a conscious resistance to it, but he has passages that have some bearing on it and that in any case develop a little further the correspondence theory itself. For example, having said that the distinguishing mark of a substance or individual thing is that it may have opposite qualities at different times, he resists a suggestion that statements and opinions would count as things by this criterion, since they may be at one time true and at another time false—for example, the statement or opinion that a person is sitting down will be true while he is doing so but false when he stands up. This, Aristotle suggests, is unfair, because what is in question here is not any genuine alteration in the statement or opinion itself, but rather in the facts outside it by which its truth or falsehood is measured. "For it is by the facts of the case, by their being or not being so, that a statement is called true or false."

Sometimes Aristotle represents the verification of statements by facts as a kind of causation. Causation, he says, differs from implication because even where implication is reciprocal we can distinguish the cause from the effect:

> The existence of a man, for instance, implies the truth of the statement in which we assert his existence. The converse is also the case. For if he exists, then the statement in which we assert his existence is true, and conversely, if the statement in which we assert his existence is true, he exists. But the truth of the statement is in no way the cause of his existence, though his existence is in a way the cause of the truth of the statement. For we call the statement true or false according as he exists or not.

(*CATEGORIES* 14B15–20)

(What Aristotle calls a cause here is perhaps something more like a criterion.)

MEGARIAN "LIAR" PARADOX. The Platonic–Aristotelian correspondence theory was not long formulated when a distressing consequence, or apparent consequence, of it was pointed out by Eubulides, a member of the school of Megara, which seems to have conducted constant warfare against various basic Platonic–Aristotelian positions. Eubulides invited his hearers to consider a man who says "I am lying" or "What I am now saying is false." According to the Platonic–Aristotelian view, this is true if what the man is saying *is* false—it is true if it is itself false—and false if what he is saying is *not* false—false if it is true. Therefore, in at least this one case, that view leads to the position that whatever we say about the truth or falsehood of an utterance entails its own opposite. We may note, too, that in this instance the Aristotelian one-sided dependence of the truth or falsehood of a proposition on the related matter of fact does not hold, since the related matter of fact in this instance *is* precisely the truth or falsehood of the proposition. This "paradox of the liar" was much discussed by both ancient and medieval writers and still presents a serious problem to anyone attempting to give a satisfactory general account of truth and falsehood.

STOICS. What is substantially the Platonic–Aristotelian account of truth is also found among the Stoics, but with modifications. The Stoics held that truth in the primary sense is a property of statements or *axiomata*, not in the sense of sentences but in the sense of what the sentences state or mean. These *axiomata* exist independently of their being expressed by sentences, and the "meanings" of false sentences exist just as much as the meanings of true ones—that is to say, *axiomata* include objective falsehoods as well as objective truths. (This is not, therefore, the existence theory.) Describing the Stoics' account of truth from this point on, Diogenes Laertius says that the *axioma* expressed by "It is day" is true if it is day and false if it is not. This is an example rather than a general theory; Sextus Empiricus says that the kind of *axioma* called simple and definite—the kind that would be expressed by a sentence of the form "This *X Y*'s" (for instance, "This bat flies")—is true when the predicate belongs to the object denoted by the demonstrative. This, however, only defines "true" for the simplest type of proposition. For other types we know that the Stoics laid down such rules as that an *axioma* of the form "Some *X Y*'s" is true if and only if there is some true *axioma* of the form "This *X Y*'s," and one of the form "*p* and *q*" is true if and only if both of its components are; but we do not know whether they regarded such rules as actually defining "true" for these forms. It is scarcely likely that they saw them as parts of a single "recursive" definition of truth, such as is found in Alfred Tarski, but they laid the foundations for such a development.

MEDIEVAL LOGICIANS. Such statements of truth conditions, as we now call them, were also laid down and discussed by the logicians of the later Middle Ages, although they generally treated truth as a property not of abstract *axiomata* but of spoken and written sentences. Besides the truth conditions of sentences containing "not," "or," and "some," they considered those of sentences containing such expressions as "possibly" and verbs in the past and future tenses. They observed, for example, that while in general a past-tense statement *is* true if and only if the corresponding present-tense statement *was* true, and a statement that something *could have been* so *is* true if and only if the statement that it *is* so *could have been* true, there are exceptions to such rules. For example, "Something white *was* black" *is* true, but "Something white *is* black" was never true. The rule here is that a past-tense predication *is* true if the corresponding predication *was* true of the individuals to which the subject term now applies; for instance, "Something white was black" is true if "It is black" could have been truly said in the past of a thing that is now white. "It could have been that no proposition is negative" is true, since God might have annihilated all negative propositions; but "No proposition is negative" could in no circumstances have been true, since the mere existence of this sentence (which is itself negative) falsifies it. The rule is rather that a sentence *de possibili* is true if and only if things could have been as the corresponding unqualified sentence says they are.

The later medieval logicians also implicitly modified the Platonic–Aristotelian theory in order to cope with the "liar" and similar paradoxes. John Buridan, for example, although he preserved the formula that a sentence is true when *ita est sicut significat*, gave a somewhat un-Aristotelian twist to the meaning of *significat*. According to Buridan, the man who says "I am saying something false," and says nothing else, really is saying something false, not because things are otherwise than as his sentence *significat formaliter* but because they are otherwise than as his sentence *significat virtualiter*. A sentence "virtually" signifies whatever follows from itself together with the circumstances of its utterance, and what follows from this particular sentence together with the circumstances of its

utterance is that it is both true and false; since this is never the case, things are not as it "virtually" says they are, and it is false.

MOORE'S CORRESPONDENCE THEORY

In the twentieth century a particularly extended and fruitful discussion of the correspondence theory is found in a series of lectures given by G. E. Moore in 1910–1911. Here truth and falsehood first appear as properties of what are called propositions. Moore uses the term *proposition* to mean not an indicative sentence but what such a sentence *means*, an *axioma* in the Stoic sense. When we both hear and understand a spoken sentence, and both see and understand a written one, there is something apprehended by us over and above the sentence, and while this apprehension or understanding is the same kind of act in all cases, what is apprehended is in general different when different sentences (such as "Twice two are four" and "Twice four are eight") are involved and therefore is distinguishable from the act of apprehending. We also "constantly think of and believe or disbelieve, or merely consider, propositions, at moments when we are neither hearing nor seeing any words which express them"; for example, when we "apprehend a proposition, which we desire to express, before we are able to think of any sentence which would express it." In this lecture Moore is quite confident that "there certainly are in the Universe such things as propositions," and that it is propositions rather than sentences or acts of belief that are true or false in the primary sense. We often say that beliefs are true or false, but this is only because the word *belief* is often used not for an act of believing but for what is believed; for instance, if we say that two different people have the same belief, we mean to identify what they believe rather than their respective acts of believing, and what is believed is simply a proposition in Moore's sense. Acts of believing and sentences could, however, be said to be true or false in a secondary sense, when what is believed or expressed is a true or false proposition.

MOORE'S LATER POSITION. Moore returns to the subjects of true and false beliefs and the nature of propositions in later lectures in the series, but now he seems to move somewhat away from the position outlined above. He leads up to them with a problem he states as follows: "Suppose a man believes that God exists; … then to say that his belief is true seems to be exactly equivalent to saying that it *is a fact* that God exists or that God's existence is a fact" (*Some Main Problems in Philosophy*, p. 250). Quite generally it seems that "the difference between true and false beliefs is … that where a belief is true, there what is believed is a fact; whereas where a belief is false, there what is believed is not a fact" (ibid.). Even where a belief is false, however, there does seem to be something that is believed.

> A man believes in God's existence and it seems quite plain that he is believing in something—that there is such a thing as what he believes in, and that this something is God's existence. It seems quite plain, therefore, that there *is* such a thing as God's existence, *whether* his belief is true or false. But we have just seen that if his belief is false, then God's existence is *not* a fact. And what is the difference between saying that there is such a thing as God's existence and (saying) that God's existence is a fact? (ibid.)

This is the problem of the *Theaetetus* all over again—if a false belief has no real object, how can it be a belief at all?

DENIAL THAT PROPOSITIONS EXIST. Moore raises the above question with regard to a more certainly false proposition, namely, that the hearers of his lecture were at that time hearing the noise of a brass band; and he then restates, but no longer with conviction, his earlier theory. We *could* say that there was indeed such a thing as their hearing a brass band then but that this was a proposition, not a fact. But, Moore argues, this theory admits in the case of the phrase "the fact that they are hearing a brass band" that what looks like the name of a real object of a possible belief is not one, so why should we not say this also of the phrase "the proposition that they are hearing a brass band"? Moore is thus led to the view that the subject–verb–object form of assertions about beliefs is misleading. His new theory, he says, "may be expressed by saying that there simply are no such things as *propositions*. That belief does *not* consist … in a relation between the believer, on the one hand, and another thing which may be called the proposition believed" (ibid., p. 265). He cannot give any satisfactory alternative analysis of belief statements to supplant the one he has abandoned, but he thinks he can give an account of the truth and falsehood of beliefs without one.

FALSE BELIEF AND FACTS. In developing the account of truth and falsehood of beliefs, Moore considers the case of a friend believing that he has gone away for his holidays, and begins in a thoroughly Aristotelian vein. "If this belief of his is *true* then I *must* have gone away … and, conversely, … *if* I *have* gone away, then this belief of his certainly *is* true" (ibid., p. 274). And similarly, "if this belief is *false*, then I *can't* have gone away … and con-

versely, *if* I have not gone away, then the belief that I *have* gone away certainly *must* be false" (ibid., p. 275). However, this statement of necessary and sufficient conditions does not constitute a definition of truth and falsehood, for "when we assert: 'The belief that I have gone away is true,' we mean to assert that this belief has some property, which it shares with other true beliefs," whereas "in merely asserting 'I have gone away,' we are not attributing any property at all to this belief" (ibid., p. 276). For "Plainly I might have gone away without my friend believing that I had; and if so, his belief would not be true, simply because it would not exist." This objection, however, suggests that Moore's having gone away would not after all be a sufficient condition, but only a necessary one, of his friend's belief being true; and it could be met by defining the truth of his friend's belief, not simply as Moore's having in fact gone away but as this together with his friend's believing it.

The problem remains, however, of generalizing this to cover all cases, which Moore goes about solving as follows: "We can see quite plainly," he says, "that this belief, if true, has to the fact that I have gone away a certain relation which that belief has to no other fact," a relation which cannot be defined in the sense of being analyzed, but with which we are all perfectly familiar and which "is expressed by the circumstance that the name of the belief is 'the belief that I have gone away,' while the name of the fact is 'that I have gone away'" (ibid.). Moore proposes to call this relation correspondence, and "To say that this belief is true is to say that there is in the Universe *a* fact to which it corresponds, and to say that it is false is to say that there is *not* in the Universe any fact to which it corresponds" (ibid., p. 277).

FACTS RATHER THAN PROPOSITIONS. It is essential to Moore's final account that although there are no propositions, there are facts. A belief, even if true, does not consist in a relation between a person and a fact, but the truth of a belief does. He is also at pains to insist that facts "are" or exist in the very sense in which, say, chairs and tables do. He concedes that as a matter of usage we find it natural to say "It is a fact that bears exist," while we do not find it at all natural to say "That bears exist, is" (or "That bears exist, exists", or even "The existence of bears exists"), but he thinks this simply reflects our acute sense of the difference in kind between facts and other things—they are real objects but objects of a very special sort. We also express their character by calling them truths, or by prefixing "It is true that" to them as an alternative to "It is a fact that." This property of *being* a truth or fact is to be carefully distinguished from the "truth" which is possessed by some beliefs and which consists, as previously explained, in correspondence to a truth or fact.

RUSSELL'S CORRESPONDENCE THEORY

In Moore's account of truth and falsehood, it will be seen, there are two elements that are a little mysterious and that he is reluctantly compelled to leave in that condition—the correct analysis of belief statements and the nature of the correspondence that entitles us to use the same form of words in describing the content of a belief and in asserting the fact to which, if true, it corresponds.

Shortly before Moore gave these lectures, Russell had made an attempt to elucidate just these points. In the concluding section of a paper he gave before the Aristotelian Society in 1906, there is a hint of this explanation, which is more fully developed in various writings of the period 1910–1912. He suggests in the 1906 paper that a belief may differ from an idea or presentation in consisting of several interrelated ideas, whose objects will be united in the real world into a single complex or fact if the belief is true, but not otherwise, so that a false belief is indeed "belief in nothing, though it is not 'thinking of nothing,' because it is thinking of the objects of the ideas which constitute the belief." In the later versions this is expanded to the view that a belief consists in a many-termed relation, the number of terms always being two more than that occurring in the fact to which, if true, the belief corresponds. For example, if it is a fact that Desdemona loves Cassio, then in this fact the two terms Desdemona and Cassio are "knit together" by the relation of loving, while if it is a fact that Othello believes that Desdemona loves Cassio, then the four terms Othello, Desdemona, the relation of loving, and Cassio are "knit together" in this fact by the relation of believing. The correspondence between the belief and the fact, when the fact exists and the belief is therefore true, consists in a certain characteristic semiparallelism between the ordering of the last terms of the belief relation and the ordering of the terms by their ordering relation in the fact. Knowing and perceiving, on the other hand, really are relations between the knower or perceiver and the fact known or perceived (which of course must *be* a fact for knowledge or perception to occur).

CRITICISMS. The above theory is open to a number of objections, some of which have been particularly well stated by P. T. Geach, and one of which, due in essence to Ludwig Wittgenstein, had already led Russell to abandon

the theory, in a course of lectures on logical atomism delivered in 1918.

Belief and what is the case. Russell's 1906–1912 theory—and indeed even Moore's more vague theory, of which it is a possible filling out—makes it altogether too mysterious that the very same words should be used to express what is believed and what is actually the case if the belief is true. (At most, there is in some languages a slight but regular formal alteration when the latter is put into *oratio obliqua* to give the former.) As Wittgenstein puts it (*Philosophical Investigations*, Para. 444), "One may have the feeling that in the sentence 'I expect he is coming' one is using the words 'he is coming' in a different sense from the one they have in the assertion 'He is coming.' But if it were so how could I say that my expectation had been fulfilled?"—that *the very thing* I expected had come to pass?

This severance of the senses of the *oratio obliqua* and *oratio recta* forms of the same sentence is exacerbated in Russell's account, as Geach points out, by its consequence that believing is not one relation but several, since the number of terms it requires differs with the number of terms required by the relation that occurs among its objects (for instance, while Othello's believing that Desdemona loves Cassio is a 4-termed relation, his believing that Desdemona gave Cassio a certain ring would be a 5-termed one). This difference arises even when we are only considering beliefs of which the apparent objects are simple relational propositions; still more radical differences would have to be admitted with believings apparently directed toward compound and general propositions. This point was, indeed, stressed by Russell himself from the outset and seems never to have been regarded by him as a serious objection to the theory, since in his 1918 lectures, even when he had abandoned the view that necessitated it, we find him saying that "belief will really have to have different logical forms according to the nature of what is believed" (*Logic and Knowledge*, p. 226), so that "the apparent sameness of believing in different cases is more or less illusory."

There is here, it seems, a remnant of the ramified theory of types that Russell at first thought necessary to deal with such paradoxes as that of the "liar." According to this theory, propositions are not only of different logical forms but also of different logical types, and "truth" and "falsehood" must be differently defined for each type; indeed, even such ordinary logical functions as negation and conjunction must be understood differently according to the types of propositions to which they are attached. Even by the time he was exercising the influence acknowledged in Russell's 1918 lectures, Wittgenstein had definitely abandoned this theory: "*Any* proposition can be negated. And this shews that 'true' and 'false' mean the same for all propositions (in contrast to Russell)" (*Notebooks 1914–1916*, p. 21).

Verbs in judgments. What Russell did successfully assimilate from Wittgenstein at this period was that in such judgments as that Othello believes that Desdemona loves Cassio, "both verbs have got to occur as verbs, because if a thing is a verb it cannot occur otherwise than as a verb." He also says:

> There are really two main things that one wants to notice in this matter that I am treating of just now. The *first* is the impossibility of treating the proposition believed as an independent entity, entering as a unit into the occurrences of the belief, and the *other* is the impossibility of putting the subordinate verb on a level with its terms as an object term in the belief. That is a point in which I think that the theory of judgment which I set forth once … was a little unduly simple, because I did then treat the object verb as if one could put it as just an object like the terms. (*Logic and Knowledge*, p. 226)

"Every right theory of judgment," as Wittgenstein puts it, "must make it impossible for me to judge that 'this table penholders the book' (Russell's theory does not satisfy this requirement)" (*Notebooks 1914–1916*, p. 96).

Propositions in judgments. Russell's objection ties up in two ways with Wittgenstein's that "a proposition itself must occur in the statement to the effect that it is judged." In the first place, it is by inserting the "proposition itself" into the "statement to the effect that it is judged" that we enable the subordinate verb to occur as a verb and not disguised as an abstract noun. (It looks, in fact, as if these "two main things" that Russell says we must notice cannot be observed together.) We might put the two objections together thus: Because the use of abstract nouns is *always* something to be explained, it is more illuminating to say that "Othello ascribes unfaithfulness to Desdemona" (where "ascribes" is apparently a 3-termed relation with "unfaithfulness" as one of its terms) means exactly what is meant by "Othello believes that Desdemona is unfaithful" than it is to say that the second means exactly what is meant by the first.

Facts as objects. The second way in which the two objections come together is more complicated, and it can be gathered from an extended discussion of what may at first seem another point: that Russell's 1906–1912 theory,

like Moore's of 1910, still takes "facts" seriously as a special sort of object. On this point Russell's 1918 view is a little obscure. He seems not to have changed at all on this subject, and describes it as one of those truisms that "are so obvious that it is almost laughable to mention them," that "the world contains *facts*, … and that there are also *beliefs*, which have reference to facts, and by reference to facts are either true or false" (*Logic and Knowledge*, p. 182). He sharply contrasts facts with propositions in this respect. "If we were making an inventory of the world, propositions would not come in. Facts would, beliefs, wishes, wills would, but propositions would not" (ibid., p. 214). This last remark occurs in a criticism of an attempt by Raphael Demos to eliminate the *negative* fact that a certain piece of chalk is not red from the "inventory of the world" by equating it with the fact that the chalk has some other positive but incompatible color. "Even if incompatibility is to be taken as a sort of fundamental expression of fact," Russell says to this, "incompatibility is not between facts but between propositions. … It is clear that no two *facts* are incompatible" (ibid.). And since propositions do not have being independently, this "incompatibility of propositions taken as an ultimate fact of the real world will want a good deal of treatment, a lot of dressing up before it will do." However, Russell's own alternative, that there are irreducibly negative facts—for instance, the fact that it is not the case that this piece of chalk is red—equally involves the consequence that there are facts that contain real falsehoods as constituents. This Russell himself pointed out in his 1906 paper, and it led him then to be more hesitant than he was later about dismissing the notion of objective falsehoods. Even if, he says in this paper, we can remove the suggestion that false beliefs have objective falsehoods for their objects:

> There is … another argument in favour of objective falsehood, derived from the case of true propositions which contain false ones as constituent parts. Take, e.g., "Either the earth goes round the sun, or it does not." This is certainly true, and therefore, on the theory we are considering, it represents a *fact*, i.e. an objective complex. But it is, at least apparently, compounded of two (unasserted) constituents, … of which one must be false. Thus our fact seems to be composed of two parts, of which one is a fact, while the other is an objective falsehood. ("On the Nature of Truth," pp. 47–48)

The real moral of all this is surely that if propositions must go, facts must go, too; but Russell seems to shrink from this step.

Elsewhere in the 1918 lectures, however, he says that facts, although apparently real in a way in which propositions are not, have the extraordinary property that they cannot be named. In the first place, they are not named by propositions (sentences). For this he has a rather strange argument, taken from Wittgenstein. Whereas Moore thought of a false belief as one that corresponds to no fact at all, Wittgenstein held that a false statement does correspond to a fact, but in the wrong way. Hence, to quote Russell's exposition of the theory:

> There are *two* propositions corresponding to each fact. Suppose it is a fact that Socrates is dead. You have two propositions: "Socrates is dead" and "Socrates is not dead." And those two propositions corresponding to the same fact, there is one fact in the world, that which makes one true and one false. … There are two different relations … that a proposition can have to a fact: the one the relation that you may call being true to the fact, and the other being false to the fact. (*Logic and Knowledge*, p. 187)

This means that a proposition does not name a fact, since in the case of a name, there is only one relation that it can have to what it names. Further,

> You must not run away with the idea that you can name facts in any other way; you cannot. … You cannot properly name a fact. The only thing you can do is to assert it, or deny it, or desire it, or will it, or wish it, or question it. … You can never put the sort of thing that makes a proposition to be true or false in the position of a logical subject. (ibid., p. 188)

RAMSEY AND THE LATER WITTGENSTEIN

Russell's whole position, as it stands, is difficult to maintain. If there are really individual objects to which the common noun "fact" applies, and we can sometimes actually perceive them (Russell continued to hold this in the 1918 lectures), then if at the time of our perceiving one our language has no name for it, why can we not invent one and christen the thing on the spot? However, there is not just superstition, but something true and important, behind the statement of Russell and Wittgenstein that facts cannot be named, and they both identify it in the end. "When I say 'facts cannot be named,'" Russell admitted in 1924, "this is, strictly speaking, nonsense. What can be said without falling into nonsense is: 'The symbol for a fact is not a name.'" Or better, perhaps: to state a fact is not to name an object. Whatever may be the

case with "that" clauses, *sentences* aren't names of anything; just as, whatever may be the case with abstract nouns, verbs are not names of anything—they are not names at all, but have other functions; naming is one thing, saying or stating another. Even Plato saw that this distinction was important.

But can we not *name what a sentence says*, for instance, by the corresponding "that" clause? Not really—"*what* a sentence says," although a good sense can be given to it, is a misleading expression; when it means anything, it means "*how* a sentence says things are" or, better, "how *we* say things are" when we use the sentence in question. To *name* what we are saying is to *say* what we are saying, and to name what we are thinking or wishing is similarly to say what we are thinking or wishing. "I think that bears exist" is, therefore, not to be analyzed as "I think (that bears exist)," which suggests that "that bears exist" is one term of the relation expressed by "think" but rather as "I think that (bears exist)," where "bears exist" does not even look like a name (it looks like, and is, a sentence) and "think that" does not look like the expression of a relation. If Othello thinks that Desdemona loves Cassio, there is indeed a 3-termed relation between Othello, Desdemona, and Cassio (not, as Russell thought, a 4-termed one between Othello, Desdemona, Cassio, and loving), but this relation consists in his *thinking that she loves him*, that is, the relation is expressed by the whole complex verb "——thinks that——loves——," not by the simple "——thinks that——," which does not express any relation at all, since its second gap is not filled by the name of an object but by a sentence, which does not name but *says* what he thinks (how he thinks things are). The plain "thinks," without the "that," means nothing at all. I may, indeed, use forms of expression like "I think something that Jones doesn't think" or "Something that Jones thinks is not true," but the "thing" in this "something" is no more to be taken seriously than the "what" in "what I say"—these sentences, respectively, mean simply "For some *p*, I *think that p* and Jones does not think that *p*" and "For some *p*, Jones thinks that *p* but it is not the case that *p*." The correspondence theory can now assume the simple form: "*X* says (believes) truly that *p*" means "*X* says (believes) that *p*, and *p*"; and "*X* says (believes) falsely that *p*" means "*X* says (believes) that *p*, and not *p*."

RAMSEY. The above position was very lightly sketched in 1927 by Frank Plumpton Ramsey, who says in effect that the words *fact* and *true* in their primary use are inseparable parts of the adverbial phrases "truly," "in fact," "it is a fact that," and "it is true that"; and these, attached to some sentence, say no more than this sentence says on its own. "It is false that *p*" or "That *p*, is contrary to fact" similarly says no more than the simple "Not *p*." Thus there are not only no falsehoods but no facts or truths either, any more than there is an entity called "the case" involved in the synonymous phrase "It is the case that." This part of Ramsey's view has led some writers to set it in opposition to the correspondence theory as a "no truth" theory, but Ramsey also discusses more complex uses of "true" in which there is something more like a juxtaposition of what a man says and what is so. In particular he considers the statement "He is always right"—"Whatever he says is true"—and renders this as "For all *p*, if he says that *p*, it is true that *p*," and this in turn as "For all *p*, if he says that *p*, then *p*." This may seem to require a further verb in its second clause, but there is already a "variable verb" implicit in the variable *p*.

WITTGENSTEIN. We may expand Ramsey's discussion of the more complex uses of "true" by taking up a suggestion of the later Wittgenstein (which, indeed, we have already used a bit). In the *Tractatus*, Wittgenstein says that "the general form of propositions is: this is how things are." In the *Investigations*, criticizing this identification, he reminds us that "This is how things are" is itself a proposition, an English sentence applied in everyday language, as in "He explained his position to me, said that this was how things were, and that therefore he needed an advance." "This is how things are" can be said to stand for any statement and can be employed as a propositional *schema*, but only because it already has the construction of an English sentence. Wittgenstein continues, "It would also be possible here simply to use a letter, a variable, as in symbolic logic. But no one is going to call the letter 'p' the general form of propositions."

"This is how things are," although a genuine proposition, is nevertheless being employed only as a propositional variable. "To say that this proposition agrees or does not agree with reality would be obvious nonsense." "This is how things are" is a propositional variable in ordinary speech in much the same way that a pronoun is a name variable in ordinary speech. In Wittgenstein's example, the "value" of this "variable" is given by a specific sentence uttered earlier, much as the denotation of a pronoun may be fixed by a name occurring earlier. "I'm desperate—that's how things are" is like "There's Jones—he's wearing that hat again." "This (that) is how things are" is a pro-sentence. But we may also obtain a specific statement by "binding" this variable, as in "However he says things are, that's how they are," that is, Ramsey's "For all *p*, if he says that *p*, then *p*." We speak truly whenever things are as we say they are, and falsely when they are

not. There was a hint of this way of putting things when the later Scholastics equated *est vera*, said of a sentence, with *ita est sicut significat* or *qualitercumque significat, ita est*—"*however* the sentence signifies (that the case is), *thus* it is"—avoiding the possibly misleading "*What* the sentence says is so."

These "misleading" forms, however, *need* not mislead us, once the whole picture has been spread out, and we can soften our earlier skepticism by agreeing that after all there *are* facts, and that there are falsehoods, if all that is meant by "There are facts" is "For some *p, p*" and by "There are falsehoods" "For some *p*, not *p*." We can say, too, that there are both facts and falsehoods that have never been either thought or asserted, that is, we can insist on the objective or mind-independent character of propositions, if by this we mean that for some *p*, both *p* and it has never been thought or said that *p*. (We cannot, of course, give examples of such facts or falsehoods, for to do so would be to state them, and then they would not be unstated; but this is no more strange than that there should be people—as there certainly are—whose names we do not know, although we cannot in the nature of the case name any specific examples.) It is significant that Moore in his last years contrived to assimilate a Ramsey-like account of truth without losing any of his earlier sense of the mind-independent and speech-independent character of what is so. Propositions about propositions, he said in effect, are *not* propositions about sentences precisely because the words *proposition, true*, and *false* are eliminable—just because "The proposition that the sun is shining is true" is equivalent to and perhaps identical with the plain "The sun is shining," it neither says anything about sentences nor entails that there are such things, since the sun could obviously *be* shining even if no one ever said so.

TARSKI'S SEMANTIC THEORY

In the theories of Ramsey and the later Moore, truth is a quasi property of a quasi object. What is really defined in them is not a property of anything, but rather what it is to say with truth that something is so; it is an account of the adverbial phrase "with truth" rather than of the adjective "true." The late medieval treatment of "true" as a straightforward adjective applying to straightforward objects—sentences—was revived in the twentieth century, and developed with extraordinary precision, elegance, and thoroughness, in a paper by Alfred Tarski that is one of the classics of modern logic.

"TRUE" AS A METALINGUISTIC ADJECTIVE. A sentence, Tarski points out, is true or false only as part of some particular language. The Schoolmen were sensitive to this point also; Buridan, for example, observed that if we neglect it, we will be trapped by such arguments as the following: "A man is a donkey" is a true sentence if and only if a man is a donkey; but "A man is a donkey" could have been a true sentence (since we could have used it to mean what we now mean by "White is a color"); ergo it could have been that a man is a donkey. Moore was fond of making similar points.

Further, Tarski argues, a sentence asserting that some sentence *S* is a true sentence of some language *L*, cannot itself be a sentence of the language *L*, but must belong to a metalanguage in which the sentences of *L* are not used but are mentioned and discussed. He is led to this view by the paradox of the "liar" which he presents, after Jan Łukasiewicz, as follows: He uses the letter *c* as an abbreviation for the expression "the sentence printed on page 158 [of his paper], line 5 from the top," and the sentence printed there is "*c* is not a true sentence." By the ordinary Aristotelian criterion for the truth of sentences, we may say "'*c* is not a true sentence' is a true sentence if and only if *c* is not a true sentence." But "*c* is not a true sentence" is precisely the sentence *c*, so we may equate the preceding with "*c* is a true sentence if and only if *c* is not a true sentence," which is self-contradictory. The contradiction is eliminated if we put "of *L*" after "true sentence" throughout and deny the principle "'*c* is not a true sentence of *L*' is a true sentence of *L* if and only if *c* is not a true sentence of *L*" on the ground that *c* is not a true sentence of *L* under any conditions whatever, because it is not a sentence of the language *L* at all but of its metalanguage *M*.

Similar paradoxes lead to similar conclusions about such terms as "is a name of" or "signifies"—in fact, all terms that concern the relations between the expressions of a language and the objects which this language is used to describe or talk about. All such semantic terms must occur, not in the language that they concern, but in the associated metalanguage. This metalanguage must contain names for expressions in the object language and may also contain descriptions of the structure of such expressions; for instance, we might be able to say in it that one sentence is the negation of another, meaning by this that it is formed from that other by prefixing the expression "It is not the case that" to it. Tarski is attempting to state the conditions under which, for a given language *L*, we can define the term "true sentence" (and perhaps other semantic expressions) in terms of this basic metalinguistic apparatus, and in such a way as to entail all

sentences, in the metalanguage *M*, of the form "*x* is a true sentence if and only if *p*," where *x* is a *name* of some sentence of *L* (we need not write "sentence of *L*" in the formula, since in *M* this is what "sentence" means), and *p* is the *translation* into *M* of this same sentence. (*M* could include *L* as a part of itself, in which case the sentences of *L* would be their own translations into *M*.) Note that this criterion of a satisfactory definition of truth, which Tarski calls the Convention *T*, is not itself such a definition in *M* of truth in *L*, since it talks *about* expressions of *M*, and about their relation to what they mean (they "name" sentences of *L*), and so is itself not in the metalanguage *M* but in the meta-metalanguage.

Since in many (meta)languages we form the name of an expression by putting that expression in quotation marks, the following might seem to meet Tarski's criterion: "For all *p*, '*p*' is a true sentence if and only if *p*." This, one might think, would immediately yield such individual cases as "'Snow is white' is a true sentence if and only if snow is white" (given, of course, that "Snow is white" is a sentence of *L*). This will not do, however, for by enclosing the fourteenth letter of the alphabet in quotation marks (however we use that letter elsewhere) we simply form the name of the fourteenth letter of the alphabet. Hence, what we get by instantiation of the proposed definition are, for example, the sentences "The fourteenth letter of the alphabet is a true sentence if and only if snow is white" and "The fourteenth letter of the alphabet is a true sentence if and only if snow is not white," which together entail that snow is white if and only if snow is not white, a contradiction.

"RECURSIVE" DEFINITION OF TRUTH. If the language *L* contained only the two simple sentences "Snow is white" and "Grass is green," plus such compounds as could be formed by prefixing "It is not the case that" to a sentence and by joining two sentences by "or," we might offer the following "recursive" definition of "true sentence":

(1) "Snow is white" is a true sentence if and only if snow is white, and "Grass is green" if and only if grass is green.

(2) The sentence formed by prefixing "It is not the case that" to a given sentence *S* is true if and only if *S* is not true.

(3) The sentence formed by placing "or" between the two sentences S_1 and S_2 is true if and only if either S_1 or S_2 is true.

There is a mathematical device for turning such "recursive" definitions into ordinary ones, so this feature of the above need not worry us; but we are clearly not very far along if we have to begin by listing all elementary sentences and defining "true" for each of them.

Suppose we enrich *L* by adding "Snow is green" and "Grass is white" to the elementary sentences, and enrich *M* by calling "snow" and "grass" names and "is green" and "is white" predicates, and defining an elementary sentence as a name followed by a predicate. We may then alter (1) above to "For any name *X* and predicate *Y*, the sentence *XY* is true if and only if the predicate *Y* applies to the object named by *X*." This, however, assumes that the metalanguage already contains the semantic expressions "names" and "applies to"; if it does not, we can only "define" them by saying that "snow" names snow, "grass" names grass, "is white" applies to *X* if and only if *X* is white, and "is green" applies to *X* if and only if *X* is green.

This is still not very satisfactory, but it is Tarski's basic procedure, except that for his simplest *L* he takes a language in which there is only one predicate, the relative or two-place predicate "is included in," and no names at all, but only variables standing for names of classes; sentences are formed from "sentential functions" by prefixing a sufficient number of universal quantifiers ("for all *x*," "for all *y*") to "bind" all the variables in the function. That is, the sentences in this language are ones like "For all *x, x* is included in *x*," and ones in which "not" and "or" are used either inside or outside the quantifiers or both—"For all *x* and *y*, either *x* is included in *y* or *y* is included in *x*" or "It is not the case that for all *x*, it is not the case that *x* is included in *x*" (this last can be abbreviated to "For some *x, x* is included in *x*"). Tarski so defines "sentential function" as to cover sentences as special cases (they are simply those sentential functions in which all the variables are bound by quantifiers) and defines the "satisfaction" of a sentential function by a class or classes (a notion very like that of a predicate's "applying to" an object) in such a way that the truth of a sentence becomes the satisfaction by *all* classes and groups of classes of the function which "is" the sentence in question.

To develop this in a little more detail: Tarski defines "sentential function" recursively, by saying that a variable followed by "is included in" followed by a variable is a sentential function, and so are expressions formed by joining sentential functions by "or" or by prefixing "It is not the case that" or a universal quantifier to a sentential function. "Satisfaction" is more complicated, for Tarski wishes to run together such cases as that the function "*x* is included in *x*" is satisfied by the class *A* if and only if *A*

is included in *A*; "For all *y, x* is included in *y*" is satisfied by *A* if and only if for all *y, A* is included in *y*; "*x* is included in *y*" is satisfied by the pair of classes *A* and *B* if and only if *A* is included in *B*; "*x* is included in *y* or *y* in *z*" by the trio of classes *A, B*, and *C* if and only if *A* is included in *B* or *B* in *C*; and so on. To cover all such cases he introduces the notion of an infinite numbered sequence of classes, numbers his variables, and says that the sequence *f* satisfies the function "v_m is included in v_n" if and only if the *m*th member of *f* is included in the *n*th; the rest is done recursively—*f* satisfies the negation of a function Φ if and only if it does not satisfy Φ itself, the disjunction of Φ and Ψ if and only if it satisfies either Φ or Ψ, and the universal quantification of Φ with respect to the *n*th variable if and only if Φ is satisfied both by *f* itself and by all sequences that are like *f* except in having a different *n*th term.

This last part of the definition is difficult but crucial. How it works can best be seen by considering a simple example. The function "v_1 is included in v_2" is satisfied by all sequences such that the first member is included in the second. The function "For any v_2, v_1 is included in v_2" is satisfied by a sequence *f* if and only if *f* is one of the sequences satisfying the preceding function and the preceding function is still satisfied if we replace *f* by any sequence otherwise like it but with a different second term. This means, in view of what sequences satisfy the first function, that a sequence will satisfy the second function if and only if its first member is included in its second, whatever class that second member may be. Finally, consider the function "For all v_1, for all v_2, v_1 is included in v_2." This is satisfied by a sequence *f* if and only if *f* satisfies the preceding function (the second) and if the preceding function is still satisfied if we replace *f* by any sequence otherwise like it but with a different *first* term. This means, in view of what sequences satisfy the second function, that a sequence will satisfy the third only if its first member is included in its second, whatever class either of them may be, that is, if and only if every class is included in every class. It is clear that if this function were satisfied by any sequence at all, it would be satisfied by every sequence whatever. (In fact, of course, it is not satisfied by all, and therefore not by any.) In some cases a sentential function will be satisfied by any sequence whatever, even though it contains free variables—as is the case with "v_1 is included in v_1"—but if it is thus satisfied and has all its variables bound—that is, is not merely a sentential function but a sentence—it will be, in Tarski's sense, "true."

TRUTH AND CORRESPONDENCE. Tarski goes on to consider a more complicated language in which there are variables of two logical types, and an ingenious extension of the notion of a sequence enables him to define "true sentence" for this language also; but when he comes to consider "languages of infinite order," in which there are variables of an infinity of logical types, he has a proof (very similar to Gödel's proof of the incompletability of arithmetic) that any definition of either "truth" or "satisfaction" in terms of the basic material he allows himself would result in the provability of some sentence contravening his Convention *T*, that is, of the negation of some sentence of the form "*x* is a true sentence if and only if *p*," in which *x* is a name in the metalanguage of a sentence in the language studied and *p* is the translation into the metalanguage of the same sentence. Even with such a language, however, it is possible to introduce into the metalanguage the undefined semantic expression "true sentence" and so to axiomatize the metalanguage, thus enriched, that all sentences of the form indicated in the Convention *T* will be provable in it, and also desirable general theorems about truth, such as that "For any sentence *x*, either *x* is a true sentence or the negation of *x* is a true sentence." "Truth," introduced in this way, has something of the mysteriousness of the "correspondence" introduced without analysis by Moore, but Tarski has not merely a suspicion but a proof that, where "truth is understood as a property of sentences of the language in question, such acceptance of a semantic term without definition is inevitable.

See also Modal Logic; Negation.

Bibliography

Aristotle's definition of truth is in his *Metaphysics* 1011b26 ff., and there are further discussions in *Categories* 4a10–4b19 and 14b12–23, and *De Interpretatione* 16a10–19. As to Aristotle's Platonic sources, Plato states in the *Theaetetus* 188E–189A the problem solved in the *Sophist* 240D and 260C–263D.

Thomas Aquinas's echoes of Aristotle are in his *De Veritate*, Q. 1, A. 1, and *Summa contra Gentiles*, Book I, Ch. 59. Buridan's *Sophismata* unfortunately has not been reprinted in modern times, but there are accounts of his treatment of the "liar" and similar paradoxes in E. A. Moody, *Truth and Consequence in Mediaeval Logic* (Amsterdam: North-Holland, 1953), and A. N. Prior, "Some Problems of Self-Reference in John Buridan," in *Proceedings of the British Academy* 48 (1962): 115–126.

G. E. Moore's 1910–1911 lectures were published under the title *Some Main Problems of Philosophy* (London: Allen and Unwin, 1953); see especially Chs. 6 and 13–16. Bertrand Russell's multiple relation theory of truth is developed in his "On the Nature of Truth," in *PAS* (1906–1907): 28–49;

Philosophical Essays (London: Longmans, Green, 1910), Ch. 7; the introduction to the first edition of *Principia Mathematica* (Cambridge, U.K.: Cambridge University Press, 1910), pp. 42–45; and *Problems of Philosophy* (New York: Henry Holt, 1912). P. T. Geach's criticisms of the multiple relation theory are in his *Mental Acts* (London: Routledge and Paul, 1957), Sec. 13; Wittgenstein's, in his *Notebooks 1914–1916* (Oxford: Blackwell, 1961), pp. 21 and 96–97. Russell's recantation is in his 1918 lectures on *The Philosophy of Logical Atomism* (edited by R. C. Marsh), which is included in the collection *Logic and Knowledge*, also edited by R. C. Marsh (London: Allen and Unwin, 1956); see especially pp. 182, 187–188, 214, 218, 225–226. His "double correspondence" theory is in *The Analysis of Mind* (London: Allen and Unwin, 1921), Lecture 13. The more subtle views of the later Russell and Wittgenstein on propositions and facts appear in Russell's 1924 paper "Logical Atomism," included in his *Logic and Knowledge* and in Wittgenstein's *Blue and Brown Books* (Oxford: Blackwell, 1958), pp. 30–38, and *Philosophical Investigations* (Oxford: Blackwell, 1953), Paras. 134 and 444.

The key passage from F. P. Ramsey is in his paper "Facts and Propositions," included in *The Foundations of Mathematics*, edited by R. B. Braithwaite (London: K. Paul, Trench, Trubner, 1931), pp. 142–143. For later views of G. E. Moore, exhibiting some influence from Ramsey, see R. B. Braithwaite, ed., *The Commonplace Books of G. E. Moore* (London, 1962), pp. 228–231, 319, and especially 374–377. For the use of a Ramsey-type definition in the handling of paradoxes, see A. N. Prior, "On a Family of Paradoxes," in *Notre Dame Journal of Formal Logic* 2 (1961): 16–32, and J. L. Mackie, "Self-Refutation—A Formal Analysis," in *Philosophical Quarterly* (July 1964): 1–12.

Alfred Tarski's "The Concept of Truth in Formalized Languages" was first printed as a paper in 1931, then modified and enlarged in a German translation in 1935, of which there is an English version in his *Logic, Semantics, Metamathematics* (Oxford: Clarendon Press, 1956), Paper 7. Tarski states its main ideas less formally in "The Semantic Conception of Truth and the Foundations of Semantics," in *Philosophy and Phenomenological Research* 4 (1944): 341–376. Notable comments on it include Max Black, "The Semantic Definition of Truth," in *Analysis* 9 (March 1948): 49–63, and J. F. Thomson, "A Note on Truth," in *Analysis* 9 (April 1949): 67–72.

OTHER RECOMMENDED TITLES

Alston, William. *A Realist Conception of Truth*. Ithaca, NY: Cornell University Press, 1996.

Armstrong, D. M. *A World of States of Affairs*. Cambridge, U.K.: Cambridge University Press, 1997.

Austin, J. L. "Truth." In *Philosophical Papers*, edited by J. O. Urmson and G. J. Warnock. Oxford: Oxford University Press, 1950.

Bealer, George. *Quality and Concept*. Oxford: Clarendon Press, 1982.

Blackburn, Simon, and K. Simmons, eds. *Truth*. Oxford: Oxford University Press, 1999.

Coffa, Alberto. *The Semantic Tradition from Kant to Carnap*. Cambridge, U.K.: Cambridge University Press, 1991.

David, Marian. *Correspondence and Disquotation: An Essay on the Nature of Truth*. Oxford: Oxford University Press, 1994.

Davidson, Donald. "A Coherence Theory of Truth and Knowledge." In *Reading Rorty*, edited by Alan R. Malachowski. Cambridge, MA: Blackwell, 1990.

Davidson, Donald. "The Structure and Content of Truth." *Journal of Philosophy* 87 (1990): 279–328.

Devitt, M. *Realism and Truth*, 2nd ed. Oxford: Blackwell, 1991.

Forbes, Graham. "Truth, Correspondence, and Redundancy." In *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth & Logic*, edited by G. Macdonald and C. Wright. Oxford: Blackwell, 1986.

Kirkham, R. L. *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press, 1992.

Moser, Paul K. *Philosophy after Objectivity*. New York: Oxford University Press, 1993.

Pitcher, G., ed. *Truth*. Englewood Cliffs, NJ: Prentice Hall, 1964.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Schmitt, F. F. *Truth: A Primer*. Boulder, CO: Westview Press, 1995.

Soames, Scott. *Understanding Truth*. Oxford: Oxford University Press, 1999.

Strawson, P. F. "Truth." In *Truth*, edited by G. Pitcher. Englewood Cliffs, NJ: Prentice Hall, 1964.

A. N. Prior (1967)
Bibliography updated by Benjamin Fiedor (2005)

COSMOGONY

See *Cosmology*

COSMOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD

The cosmological argument is actually a family of arguments that seek to demonstrate the existence of a sufficient reason or first cause of the existence of the cosmos. Among the proponents of the cosmological argument stand many of the most prominent figures in the history of western philosophy: Plato, Aristotle, Ibn Sīnā, al-Ghazālī, Maimonides, Anselm, Aquinas, Scotus, Descartes, Spinoza, Leibniz, and Locke, to name but some. The arguments offered by these thinkers can be grouped into three basic types: (1) what may be called the *kalam* cosmological argument for a first cause of the beginning of the universe; (2) the Thomist cosmological argument for a sustaining ground of being of the world; and (3) the Leibnizian cosmological argument for a sufficient reason why anything at all exists.

The *kalam* cosmological argument derives its name from the Arabic word designating medieval Islamic scholasticism, the intellectual movement largely responsible for developing this version of the cosmological argument. It originated in the efforts of Christian philosophers such as John Philoponus who, out of their commitment to the biblical teaching of *creatio ex nihilo*, sought to rebut the Aristotelian doctrine of the eternity of the universe. The argument aims to show that the universe had a beginning at some moment in the finite past and, because something cannot come out of nothing, must therefore have a transcendent cause that brought the universe into being.

By contrast the Thomist cosmological argument, named for the medieval philosophical theologian Thomas Aquinas, seeks a cause that is first—not in the temporal sense, but in the sense of rank. On Aquinas's Aristotelian-inspired metaphysic, every existing finite thing is composed of essence and existence and is therefore radically contingent. A thing's essence is a set of properties that serve to define what that thing is. Now if an essence is to be instantiated, there must be conjoined with that essence an act of being. This act of being involves a continual bestowal of being, or the thing would be annihilated. Essence is in potentiality to the act of being, and therefore without the bestowal of being the essence would not be instantiated.

For the same reason no substance can actualize itself; for in order to bestow being upon itself it would have to be already actual; a pure potentiality cannot actualize itself but instead requires some external cause. Although Aquinas argued that there cannot be an infinite regress of causes of being (because in such a series all the causes would be merely instrumental and so no being would be produced, just as no motion would be produced in a watch without a spring even if it had an infinite number of gears) and that therefore there must exist a first uncaused cause of being, his actual view was that there can be no intermediate causes of being at all, that any finite substance is sustained in existence immediately by the ground of being. This must be a being that is not composed of essence and existence and, hence, requires no sustaining cause. One cannot say that this being's essence includes existence as one of its properties, for existence is not a property, but an act, the instantiating of an essence. Therefore, one must conclude that this being's essence just *is* existence. In a sense, this being has no essence; rather it is the pure act of being, unconstrained by any essence. It is, as Thomas says, *ipsum esse subsistens*, the act of being itself subsisting. Thomas identifies this being with the God whose name was revealed to Moses as "I am" (Exod. 3.15).

The German polymath Gottfried Wilhelm Leibniz, for whom the third form of the argument is named, sought to develop a version of the cosmological argument from contingency without the Aristotelian metaphysical underpinnings of the Thomist argument. "The first question which should rightly be asked," he wrote, "will be, *Why is there something rather than nothing?*" ("The Principles of Nature and of Grace, Based on Reason," §7, p. 527). Leibniz meant this question to be truly universal, not merely to apply to finite things. On the basis of his principle of sufficient reason (PSR) that "no fact can be real or existent, no statement true, unless there be a sufficient reason why it is so and not otherwise" ("The Monadology," §32, p. 539), Leibniz held that this question must have an answer. It will not do to say that the universe (or even God) just exists as a brute fact. There must be an explanation why it exists. He went on to argue that the sufficient reason cannot be found in any individual thing in the universe, nor in the collection of such things which is the universe, nor in earlier states of the universe, even if these regress infinitely. Therefore, there must exist an ultramundane being that is metaphysically necessary in its existence, that is to say, its nonexistence is impossible. It is the sufficient reason for its own existence as well as for the existence of every contingent thing.

THE LEIBNIZIAN COSMOLOGICAL ARGUMENT

Undoubtedly, the most controversial premise in the Leibnizian cosmological argument is the PSR. The principle as stated in "The Monadology" has seemed, to many, evidently false. Not every fact can have an explanation, for there cannot be an explanation of what has been called the big conjunctive contingent fact (BCCF) that is itself the conjunction of all the contingent facts there are; for if such an explanation is contingent, then it, too, must have a further explanation; whereas if it is necessary, then the fact explained by it must also be necessary. But the explanation cannot have a further explanation, because the BCCF includes all the contingent facts there are; and the fact explained by it cannot be necessary, because the BCCF is contingent.

Some proponents of the cosmological argument have responded to this objection by abandoning the PSR and agreeing that one must ultimately come to some explanatory stopping point that is simply a brute fact, a being whose existence is unexplained. For example, Richard Swinburne (1991) argues that God, as the brute

ultimate, is the best explanation of why everything else exists, because as a unique and infinite being God is simpler than the variegated and finite universe.

But other theists have sought to defend the Leibnizian argument without retreating to the dubious position that God is a contingent being. They have either challenged the assumption that there is a BCCF or sought to provide an acceptable explanation of it. It may well be that the existence of a BCCF is inherently paradoxical (compare the set of all truths), so that its existence cannot just be assumed. But if there is such a fact, then the claim that its explanation cannot be found in a necessary truth presupposes that explanations must entail the facts they serve to explain. If some fact is materially implied by a necessary truth, then it may be explained by that truth without itself being necessary.

Some theists have suggested that the BCCF may be explained by the necessary truth that God has weighed the reasons for creating each world and has freely chosen which world to create. Moreover, the claim that the BCCF cannot be explained by some contingent truth assumes, even more controversially, that no contingent truth can be self-explained. The reason why the BCCF is true may be simply because each of its conjuncts is true; nothing more is needed to explain why the BCCF is true than the truth of its atomic constituents, each of which has an explanation for its truth. Or again, it may be supposed that the explanation for the BCCF is that God freely wills the BCCF. Because that explanation is itself a contingent fact, it is also a constituent of the BCCF willed by God. It may then be regarded as self-explained or its explanation may be that God wills that he wills the BCCF, which fact will also be a constituent of the BCCF to be similarly explained in terms of yet another conjunct. This regress seems to be as innocuous as a series of entailments such as its being true that it is true that *p*. The entire regress is contained in the BCCF and so is willed by God.

This debate is, in any case, somewhat academic because the cosmological argument does not depend for its success on anything so strong as Leibniz's own version of the PSR. For example, in their discussion of Hartry Field's anti-Platonist claim that it is an inexplicable contingency whether mathematical objects exist, Crispin Wright and Bob Hale (1992), while rejecting the demand for an explanation of something such as the BCCF, nevertheless maintain that explicability is the default position and that exceptions to this rule have to be *explicable* exceptions—some explanation is needed for why no explanation is possible. For example, they claim that if physical existence is at issue, Leibniz's question, "Why is there something rather than nothing?" is an unanswerable question if a satisfactory explanation of why a physical state of affairs obtains has to advert to a causally prior situation in which it does not obtain, because a physically empty world would not cause anything. Wright and Hale believe that the demand for an explanation of the contingency of physical existence is preempted by the restrictive principle that "the explanation of the obtaining of a (physical) state of affairs must advert to a causally prior state of affairs in which it does not obtain" (Wright and Hale 1992, p. 128). Such a principle will be seen by the theist, however, as not at all restrictive, because the explanation of why the physical world exists can and should be provided in terms of a causally prior nonphysical state of affairs involving God's existence and will.

The proponent of the Leibnizian cosmological argument could generate an argument by holding, in conjunction with the above principle, that the obtaining of any physical state of affairs has an explanation. Or the proponent could claim that for any contingently existing thing, there is an explanation why that thing exists; or assert that everything that exists has an explanation of its existence, either in the necessity of its own nature or in an external cause; or, more broadly, maintain that in the case of any contingent state of affairs, there is either an explanation for why that state of affairs obtains or else an explanation of why no explanation is needed. All of these are more modest, nonparadoxical, and seemingly plausible versions of the PSR.

A simple statement of a Leibnizian cosmological argument might run as follows:

(1) Anything that exists has an explanation of its existence, either in the necessity of its own nature or in an external cause.

(2) If the universe has an explanation of its existence, that explanation is God.

(3) The universe exists.

(4) Therefore the explanation of the existence of the universe is God.

The version of the PSR in premise (1) is compatible with there being brute facts or states of affairs about the world. But there are two kinds of being: necessary beings, which exist of their own nature and so have no external cause of their existence, and contingent beings, whose existence is accounted for by causal factors outside themselves. Numbers might be prime candidates for the first sort of being, whereas familiar physical objects fall under the second kind of being.

Premise (2) is, in effect, the contrapositive of the typical atheist response to Leibniz that on the atheistic worldview the universe simply exists as a brute contingent thing. Atheists typically assert that, there being no God, the universe just exists inexplicably. Moreover, (2) seems plausible in its own right, for if the universe, by definition, includes all of physical reality, then the cause of the universe must (at least causally prior to the universe's existence) transcend space and time and therefore cannot be physical or material. But there are only two kinds of things that could fall under such a description: either an abstract object or else a mind. But abstract objects do not stand in causal relations. Therefore it follows that the explanation of the existence of the universe is an external, transcendent, personal cause—which is one meaning of "God."

Finally, premise (3) states the obvious—that there is a universe. Because the universe exists, it follows that God exists.

It is open to the nontheist to retort that whereas the universe has an explanation of its existence, that explanation lies not in an external ground but in the necessity of its own nature; in other words, (2) is false. The universe is a metaphysically necessary being. This is an extremely bold suggestion. One may safely say that there is a strong intuition of the universe's contingency. A possible world in which no concrete objects exist certainly seems conceivable. People generally trust their modal intuitions on other matters with which they are familiar; if they are to do otherwise with respect to the universe's contingency, then the nontheist needs to provide some reason for such skepticism other than the desire to avoid theism.

THE THOMIST COSMOLOGICAL ARGUMENT

Still, it would be desirable to have some stronger argument for the universe's contingency than modal intuitions alone. Could the Thomist cosmological argument help out here? If successful, it would show that the universe is a contingent being causally dependent upon a necessary being for its continued existence. The difficulty with appeal to the Thomist argument, however, is that it is difficult to show that things are, in fact, contingent in the special sense required by the argument. Certainly things are naturally contingent in that their continued existence is dependent upon myriad factors including particle masses and fundamental forces, temperature, pressure, entropy level, and so forth, but this natural contingency does not suffice to establish things' metaphysical contingency in the sense that being must continually be added to their essences lest they be spontaneously annihilated. Indeed, if Thomas's argument does ultimately lead to an absolutely simple being whose essence is existence, then one may well be led to deny that beings are metaphysically composed of essence and existence if the idea of such an absolutely simple being proves to be unintelligible.

THE *KALAM* COSMOLOGICAL ARGUMENT

But perhaps the *kalam* cosmological argument can reinforce the Leibnizian argument. For an essential property of a metaphysically necessary being is eternality, that is to say, existing without beginning or end. If the universe is not eternal, then, it could not be a metaphysically necessary being.

It may be countered that a being with a temporal beginning or end could be metaphysically necessary in that it exists in all possible worlds. But the notion of metaphysical necessity that underlies this suggestion fails to take tense seriously and may therefore seem inadequate. Metaphysicians have in recent years begun to appreciate the metaphysical importance of whether time is tensed or tenseless; that is to say, whether items in the temporal series are ordered objectively as past, present, or future, or whether, alternatively, they are ordered merely by tenseless relations of *earlier than*, *simultaneous with*, and *later than*. Possible worlds semantics is a tenseless semantics and so is incapable of expressing the significance of one's view of time and tense. In particular, it is evident that a truly necessary being, one whose nonexistence is impossible, must exist at every moment in every world. It is not enough for it to exist at only some moment or moments in every possible world; for the fact that there exist moments in various worlds at which it fails to exist shows that its nonexistence is not impossible. Furthermore, a truly metaphysically necessary being must exist either timelessly or sempiternally in any tensed world in which it exists, for otherwise its coming into being or ceasing to be would again make it evident that its existence is not necessary, even if it existed at every moment in worlds in which time had a beginning or end.

But it is precisely the aim of the *kalam* cosmological argument to show that the universe is not sempiternal but had a beginning. It would follow that the universe must therefore be contingent in its existence. Indeed, the *kalam* argument shows the universe to be contingent in a special way: It came into existence out of nothing. The nontheist who would answer Leibniz by holding that the existence of the universe is a brute fact, an exception to the PSR, is

thus thrust into the awkward position of maintaining not simply that the universe exists eternally without explanation, but rather that for no reason at all it magically popped into being out of nothing, a position that might make theism look like a welcome alternative.

The *kalam* cosmological argument may be formulated as follows:

(1) Whatever begins to exist has a cause.

(2) The universe began to exist.

(3) Therefore, the universe has a cause.

Conceptual analysis of what it means to be a cause of the universe then helps to establish some of the theologically significant properties of this being.

Premise (1) seems obviously true—at the least, more so than its negation. It is rooted in the metaphysical intuition that something cannot come into being from nothing. Moreover, this premise is constantly confirmed in human experience. The conviction that an origin of the universe requires a causal explanation seems reasonable, for on the atheistic view, if the universe began at the big bang, there was not even the *potentiality* of the universe's existence prior to the big bang, because nothing is prior to the big bang. But then how could the universe become actual if there was not even the potentiality of its existence? It makes much more sense to say that the potentiality of the universe lay in the power of God to create it.

Often it is said that quantum physics furnishes an exception to premise (1), because on the subatomic level events are said to be uncaused. This objection, however, is based on misunderstandings. In the first place, not all scientists agree that subatomic events are uncaused. A great many physicists today are dissatisfied with this view (the so-called Copenhagen Interpretation) of quantum physics and are exploring deterministic theories like that of David Bohm (Cushing, et al, 1996). Thus, quantum physics is not a proven exception to premise (1). Second, even on the traditional, indeterministic interpretation, particles do not come into being out of nothing. They arise as spontaneous fluctuations of the energy contained in the subatomic vacuum, which constitutes an indeterministic cause of their origination. Thus, there is no basis for the claim that quantum physics proves that things can begin to exist without a cause, much less that the universe could have sprung into being uncaused from literally nothing.

Premise (2), the more controversial premise, may be supported by both deductive, philosophical arguments and inductive, scientific arguments. Classical proponents of the *kalam* argument contended that an infinite temporal regress of events cannot exist, because the existence of an actually infinite number of things leads to intolerable absurdities.

It is usually alleged that this sort of argument has been invalidated by Georg Cantor's work on the actual infinite and by subsequent developments in set theory (e.g., Sobel). But this allegation gratuitously presupposes a Platonistic view of mathematical objects that the argument's defender is at liberty to reject. Cantor's system and set theory may be taken to be simply a universe of discourse, a mathematical system based on certain adopted axioms and conventions. The argument's defender may hold that whereas the actual infinite may be a fruitful and consistent concept within the postulated universe of discourse, it cannot be transposed into the spatio-temporal world, for this would involve counterintuitive absurdities. A fictionalist understanding of abstract objects or a divine conceptualism combined with the simplicity of God's cognition is at least a tenable alternative to Platonism.

A second argument for the beginning of the universe offered by classical proponents of *kalam* is that the temporal series of past events cannot be an actual infinite because a collection formed by successive addition cannot be actually infinite, an argument that eventually became enshrined in the thesis of Kant's first antinomy concerning time. Sometimes the problem is described as the impossibility of traversing the infinite. In order for one to have "arrived" at today, temporal existence has, so to speak, traversed an infinite number of prior events. But before the present event could arrive, the event immediately prior to it would have to arrive; and before that event could arrive, the event immediately prior to it would have to arrive; and so on ad infinitum. No event could ever arrive, because before it could elapse there will always be one more event that has had to have happened first. Thus, if the series of past events were beginningless, the present event could not have arrived, which is absurd.

It is frequently objected that this sort of argument illicitly presupposes an infinitely distant starting point in the past and then pronounces it impossible to travel from that point to today. But if the past is infinite, then there would be no starting point whatever, not even an infinitely distant one. Nevertheless, from any given point in the past, there is only a finite distance to the present, which is easily traversed.

But proponents of the *kalam* argument have not in fact assumed that there was an infinitely distant starting point in the past. The fact that there is *no beginning* at all,

not even an infinitely distant one, seems only to make the problem worse, not better. To say that the infinite past could have been formed by successive addition is like saying that someone has just succeeded in writing down all the negative numbers, ending at -1. And, one may ask, how is the claim that from any given moment in the past there is only a finite distance to the present even relevant to the issue? For the question is how the *whole* series can be formed, not a finite portion of it. To think that because every *finite* segment of the series can be formed by successive addition the whole *infinite* series can be so formed is to commit the fallacy of composition.

A third argument for the universe's beginning is an inductive argument based on contemporary evidence for the expansion of the universe. According to the standard big bang model, as time proceeds, the distances separating galactic masses become greater. It is important to understand that the model does not describe the expansion of the material content of the universe into a preexisting, empty space, but rather the expansion of space itself. This has the astonishing implication that as one reverses the expansion and extrapolates back in time, the universe becomes progressively denser until one arrives at a so-called singularity at which space-time curvature, along with temperature, pressure, and density, becomes infinite. It therefore constitutes an edge or boundary to space-time itself.

The history of twentieth-century cosmology has, in one sense, been a series of failed attempts to craft acceptable nonstandard models of the expanding universe in order to avert the absolute beginning predicted by the standard model. Whereas such theories are possible, it has been the overwhelming verdict of the scientific community than none of them is more probable than the big bang theory. There is no mathematically consistent model that has been so successful in its predictions or as corroborated by the evidence as the traditional big bang theory. For example, some theories, such as the oscillating universe (which expands and recontracts forever) or the chaotic inflationary universe (which continually spawns new universes), do have a potentially infinite future but turn out to have only a finite past. Vacuum fluctuation universe theories (which postulate an eternal vacuum out of which this universe is born) cannot explain why, if the vacuum was eternal, one does not observe an infinitely old universe. The no-boundary universe proposal of James Hartle and Stephen Hawking, if interpreted realistically, still involves an absolute origin of the universe even if the universe does not begin in a singularity, as it does in the standard big bang theory. More recently proposed ekpyrotic cyclic universe scenarios based on string theory or M-theory have also been shown not only to be riddled with problems, but, most significantly, to imply the origin of the universe that its proponents sought to avoid. Of course, scientific results are always provisional, but there is no doubt that the defender of the *kalam* argument rests comfortably within the scientific mainstream.

A fourth argument for the finitude of the past is also an inductive argument, appealing to implications of physical eschatology. According to the second law of thermodynamics, processes taking place in a closed system always tend toward a state of equilibrium. The universe is, on a naturalistic view, a gigantic closed system, because it is everything there is and there is nothing outside it. What this seems to imply is that, given the probability that the universe will expand forever, the universe will in the finite future degenerate into a cold, dark, lifeless, highly dilute condition, as it asymptotically approaches equilibrium. Now if, given enough time, the universe will reach such a condition, then why is it not in such a condition now, if it has existed forever, from eternity? Because it is not in such a state, the universe must have begun to exist.

Some have tried to escape this conclusion by proposing an oscillating model of the universe that restores an appearance of youth to an infinitely old cosmos. But even apart from the physical and observational problems plaguing such models, the thermodynamic properties of such a universe imply the very beginning that its proponents sought to avoid—for entropy increases from cycle to cycle in such a model, which has the effect of generating larger and longer oscillations with each successive cycle. Hence, the oscillating model has an infinite future, but only a finite past.

Even if this difficulty were avoided, a universe oscillating from eternity past would require an infinitely precise tuning of initial conditions in order to last through an infinite number of successive bounces. A universe rebounding from a single, infinitely long contraction is, if entropy increases during the contracting phase, thermodynamically untenable and incompatible with the initial low entropy condition of the expanding phase. Postulating an entropy decrease during the contracting phase in order to escape this problem would require one to postulate inexplicably special low entropy conditions at the time of the bounce in the life of an infinitely evolving universe. Such a low entropy condition at the beginning of the expansion is more plausibly accounted for by the presence of a singularity or some sort of quantum creation event.

From the two premises it follows logically that the universe has a cause. Protagonists of *kalam* maintain that a conceptual analysis of what properties must be possessed by such an ultramundane cause enables one to recover a striking number of the traditional divine attributes, revealing that if the universe has a cause, then an uncaused, personal creator of the universe exists who sans the universe is beginningless, changeless, immaterial, timeless, spaceless, and enormously powerful. This creator will be, as Leibniz maintained, the sufficient reason why anything at all exists.

See also Causation; Cosmology; Cosmos; God, Concepts of; Leibniz, Gottfried Wilhelm; Thomas Aquinas, St.

Bibliography

Adams, Fred C., and Gregory Laughlin. "A Dying Universe: The Long-Term Fate and Evolution of Astrophysical Objects." *Reviews of Modern Physics* 69 (2) (1997): 337–372.

Beck, W. David. "The Cosmological Argument: A Current Bibliographical Appraisal." *Philosophia Christi* 2 (2000): 283–304.

Burrill, Donald R. *The Cosmological Arguments*. Garden City, NY: Doubleday, 1967.

Craig, William Lane. *The Cosmological Argument from Plato to Leibniz*. New York: Barnes & Noble, 1980.

Craig, William Lane. *The* Kalam *Cosmological Argument*. London: Macmillan, 1979. Reprint, Eugene, OR: Wipf and Stock, 2000.

Craig, William Lane. "Naturalism and Cosmology." In *Analytic Philosophy without Naturalism*, edited by A. Corradini, S. Galvan, and J. Lowe. London: Routledge, 2005.

Craig, William Lane, and Quentin Smith. *Theism, Atheism, and Big Bang Cosmology*. Oxford: Clarendon Press, 1993.

Cushing, James T., Arthur Fine, and Sheldon Goldstein. *Bohmian Mechanics and Quantum Theory: An Appraisal*. Boston Studies in the Philosophy of Science 184. Dordrecht: Kluwer Academic Publishers, 1996.

Davis, Stephen T. *God, Reason, and Theistic Proofs*. Grand Rapids, MI: Eerdmans, 1997.

Gale, Richard M. *On the Existence and Nature of God*. New York: Cambridge University Press, 1991.

Harrison, Jonathan. *God, Freedom, and Immortality*. Avebury Series in Philosophy. Burlington, VT: Ashgate, 1999.

Hick, John. *Arguments for the Existence of God*. London: Macmillan, 1971.

Leibniz, G. W. "The Principles of Nature and of Grace, Based on Reason." In *Leibniz Selections*, edited by Philip P. Wiener, 522–533. New York: Charles Scribner's Sons, 1951.

Leibniz, G. W. "The Monadology." In *Leibniz Selections*, edited by Philip P. Wiener, 533–552. New York: Charles Scribner's Sons, 1951.

Mackie, John L. *The Miracle of Theism*. Oxford: Clarendon Press, 1982.

Martin, Michael. *Atheism: A Philosophical Justification*. Philadelphia: Temple University Press, 1990.

Pruss, Alexander. Ex Nihilo Nihil Fit: *A Study of the Principle of Sufficient Reason*. Cambridge, U.K.: Cambridge University Press, 2005.

Rowe, William L. "Circular Explanations, Cosmological Arguments, and Sufficient Reasons." *Midwest Studies in Philosophy* 21 (1997): 188–199.

Rowe, William L. *The Cosmological Argument*. Princeton, NJ: Princeton University Press, 1975.

Sobel, Jordan Howard. *Logic and Theism: Arguments for and against Beliefs in God*. Cambridge, U.K.: Cambridge University Press, 2004.

Swinburne, Richard. *The Existence of God*. Rev. ed. Oxford: Clarendon Press, 1991.

Vallicella, William. "On an Insufficient Argument against Sufficient Reason." *Ratio*10 (1997): 76–81.

Wright, Crispin, and Bob Hale. "Nominalism and the Contingency of Abstract Objects." *Journal of Philosophy* 89 (1992): 111–135.

William Lane Craig (2005)

COSMOLOGY

The term *cosmology* stands for a family of related inquiries, all in some sense concerned with the world at large. Two main subgroups of uses may be distinguished: those belonging to philosophy and those belonging to science.

"Cosmology" has received wide currency as a name for a branch of metaphysics, ever since Christian von Wolff, in his *Discursus Praeliminaris de Philosophia in Genere* (1728), gave cosmology a prominent place in his classificatory scheme of the main forms of philosophical knowledge and distinguished this branch from ontology, theology, and psychology. (See *Discourse on Philosophy in General*, translated by R. J. Blackwell, Indianapolis, 1963, Para. 77). Despite the severe strictures that Immanuel Kant leveled against the pursuit of rational cosmology in his *Critique of Pure Reason*, the term has continued to enjoy a standard use among many philosophers. For example, it occupies a central place in the manuals of scholastic philosophy; these adhere, for the most part, to the Wolffian scheme of classification of the branches of metaphysics. The term has been used, too, by many philosophers not in the scholastic tradition; for example, A. E. Taylor in his *Elements of Metaphysics* (London, 1903) assigns to cosmology the task of considering "the meaning and validity of the most universal conceptions of which we seek to understand the nature of the individual objects which make up the experienced physical world, 'extension,' 'succession,' 'space,' 'time,' 'number,' 'magnitude,' 'motion,' 'change,' 'quality,' and the more complex categories of 'matter,' 'force,' 'causality,' 'interac-

tion,' 'thinghood,' and so forth" (p. 43). Cosmology is sometimes understood even more broadly, as being synonymous with speculative philosophy in its most comprehensive sense. Thus in Alfred North Whitehead's *Process and Reality* (New York, 1929), whose subtitle is "An Essay in Cosmology," the attempt is made to construct a categorial scheme of general ideas "in terms of which every element of our experience can be interpreted" (p. 4).

In its second major use, the term *cosmology* designates a science in which the joint efforts of the observational astronomer and the theoretical physicist are devoted to giving an account of the large-scale properties of the astronomical or physical universe as a whole. The task of constructing models of the universe that are suggested by and tested by appeals to the observational findings of the astronomer distinguishes the enterprise of scientific cosmology from the a priori investigations of rational cosmology (as a branch of metaphysics) and the purely conceptual and categorial analyses of the speculative philosopher. Nevertheless, even scientific cosmology poses a number of philosophical questions. The sum of these—and they are principally methodological and epistemological in character—constitutes the philosophy of scientific cosmology. The present entry is concerned with the philosophy of cosmology in this sense. Attention will be focused on a central theme in this area: the question whether cosmology must employ a method different from that employed in other empirical sciences because of its distinctive subject matter, namely, the universe as a unique system.

DESCRIPTION OR EXPLANATION?

Is the familiar distinction between description and explanation (or the corresponding one drawn between sciences still in the early stages that are primarily descriptive, and those that have progressed to the predominating use of the explanatory aspects of theory) a distinction that can be profitably applied in giving an account of the logic of cosmology? No simple and unqualified answer can be given. For, on the one hand, cosmology, in attempting to gain knowledge of the universe as a whole, certainly is not content to rest with the observational reports of the astronomer, and therefore cannot be classed with the descriptive sciences. On the other hand, in advancing to the level of theory, as cosmology in a qualified sense certainly does, it is not primarily concerned with the explanation of laws—as is the case with other explanatory sciences.

If by description is meant giving an account of some single event or object in observational terms, or (in an extended sense of "description") formulating a generalization (law) in observational terms which refers to the observable or measurable properties and relations of a class of events, then cosmology, which is interested in giving an account of the universe as a whole, is not engaged in description. Even if we recognize, as we must, the descriptive activities of observational cosmology as a branch of observational astronomy, these fall short of giving us an adequate account of the universe as a whole. All that astronomy can give us is a description of the domain of objects and events within the range of its most powerful instruments. At the present time, however, these instruments, have not reached the limit, if there is a limit, to what is in principle observable. Moreover, even if the universe were in some sense finite and wholly explorable by actually or theoretically available instruments, the statement that what is thus observationally explored is in fact the universe as a whole would not be warranted by observational evidence alone. Such a statement could not, therefore, be part of the description of the universe, insofar as this description is a report of what is found. The claim that the universe is open to complete inspection requires the support of theory. It is a statement which is not included in the description, but is a rider *to* the description—to the effect that the description as given is of the universe as complete; considerations other than purely observational ones are needed to support this claim.

If cosmology is not content with description, does it then aim at giving explanations? Here our answer must be qualified. In the case of ordinary empirical generalizations, where there are multiple instances of some phenomenon of which we have examined a limited number, we say that the law supported by this evidence may be used as a reliable rule of inference. Since the law applies to a kind of subject matter, or a type of phenomenon, it can be upheld as a useful means for predicting and explaining those instances that can be brought within its scope. But in cosmology, the primary goal is not to establish laws. The universe, by definition, is a unique object or system. Cosmology does not undertake to establish laws about universes; at best one can establish laws about the constituents of the universe. The relation that the observable portion of the universe bears to the universe is that of part to presumed whole, rather than that of instance to law. Hence, if to explain means to bring an instance under a law, this mode of explanation, which is a characteristic concern of other branches of physical science, does not characterize cosmology.

Can it be said, then, cosmology aims at giving explanations in the sense in which theories are employed to explain laws? Here our answer, once more, cannot be a straightforward "yes" or "no." The characteristic device employed by theoretical cosmology is a model of the universe, and a model in many respects functions precisely as a theory does. It is a conceptual construction that cannot be said to be a mere report of what is already found in observation, nor even an anticipated description of what might be found in future observations. Rather, it is a means for making the observational data themselves intelligible. However, the facts that the cosmologist wants to explain are not laws in the ordinary sense of the term, and so in this respect the purpose of a model of the universe is not identical with that of a theory in the ordinary sense.

Consider, for example, the question "Why is the apparent magnitude of galaxies correlated with their red shift?" This question asks for an explanation of an important datum of observation. The observed fact is sometimes called Hubble's "law," but it is a law only in the peculiar sense in which we refer to Kepler's laws as "laws." That is, Hubble's law tells us something about a particular distribution or process of a unique set of objects, namely the system of galaxies, just as Kepler's laws tell us something about the orbits of the planets in our solar system, not in any solar system. In general, however, laws of science are characterized by their universal form. They are unrestricted in scope and are not ostensibly tied to objects or events specifically located in some particular space-time region. Thus Newton's law of gravitation, for example, says that for *any* two bodies, the gravitational force that holds between them is inversely as the square of their distance apart and proportional to their masses. Now when we deal with the system of bodies and processes that constitute the unique configuration we call the universe, we are not dealing with *any* configuration of events and objects; we are dealing with the configuration actually observed and given.

An interesting and important question that can be raised here, however, is whether the unrestricted laws of ordinary physics are not themselves, in a more profound sense, relatively restricted, since they apply to bodies or phenomena within the ultimately unique configuration that constitutes the physical universe. From this point of view, the study of cosmology sets the environment and limiting framework for all other branches of physical science. Hence it is not unreasonable to expect—as E. A. Milne, D. W. Sciama, and others have pointed out—that one may hope to understand the laws of physics themselves in terms of the unique background making up the universe studied in cosmology. Such a claim, however, is associated only with certain specific models, namely, the kinematic model as worked out by Milne and the steady state model as sketched by Hermann Bondi and Sciama, and therefore this idea of explaining all laws by a cosmological model cannot be held up as a working goal for all cosmological models. In fact, the majority of models developed within the framework of general relativity theory have not been designed to embody this feature.

OBSERVATION AND THEORY

The study of cosmology has two lines of attack, that of the observational astronomer and that of the theoretical physicist. One might say that both the observational astronomer and the theoretical cosmologist are studying the universe, though from different vantage points, or that one supplies observational data about the universe that the other undertakes to interpret; but this is, at best, only a sketch of the situation and is in some ways seriously misleading. For it will not do to say that both the astronomer and the theoretical cosmologist are studying the universe, as if the universe is laid out for identification before them and the only difference between them is in approach and method. If we look more closely at the study of cosmology, the situation is rather different.

The observational astronomer is not confronted with the universe as an observationally complete whole. Instead, he obtains observational clues from various instruments about a large population of identifiable subsystems—namely, individual galaxies and clusters of galaxies. This population of observable entities is sometimes referred to as "the observable universe." However, this phrase is not to be understood in the sense that we have independent means for identifying the universe and that we wish to refer to it insofar as it is being observed. "The observable universe" is not the same as "the universe observed." What the astronomer reports on of relevance to cosmology is an observable population of galaxies and clusters of galaxies. These observational reports have to do with such matters as the spatial distribution of galaxies, their systematic motions, density, spectroscopic patterns, individual shapes, and stellar composition.

The population of subsystems that makes up the observable universe is now, as in a sense it must always be, a finite population. With the advance in the power and sensitivity of instruments, knowledge of the extent of this population and the refinement in the details of the reports about this population are improved. Although it is regarded as likely that further advances in observa-

tional resources will disclose a wider population of subsystems similar to those already observed, it must be remembered that it is always possible for further observations to disclose as basic constituents of the universe astronomical units of a higher degree of inclusiveness than galaxies or clusters of galaxies, or even entities of an altogether different type from those heretofore disclosed. Whatever may be the case in the future, it certainly is the case at present that what comes within the observational reaches of the astronomer is definitely not the universe as an absolute whole, if there is in fact such a whole.

When we say that the theoretical cosmologist studies the universe in order to understand it or make it intelligible, what is it that he studies? He does not study the universe in any direct way, if that means having before him a readily identified object which he tries to comprehend, for example by subsuming it under some law. Nor, as we have just seen, does the universe he studies consist of a complete population of entities about which the observational astronomer furnishes him detailed reports. The theoretical cosmologist is not given information about the universe as a whole, nor even about what lies beyond the immediate range of the astronomer's instruments.

What then does he study? A brief and simple answer is to say that he constructs a model of the universe and that he studies the way in which this model may be used to interpret the observational data already available. The cosmologist will use his model to interpret the data assembled by the observational astronomer and to guide the astronomer in the search for further data. Insofar as the use of theoretical models proves satisfactory, we may say that cosmology has helped us to understand the universe and to make it intelligible. This is not to be understood, however, as meaning that even at the end of a relatively successful course of inquiry, the cosmologist has been able to confront the universe directly as some kind of readily identifiable object, system, or class of objects. What is to be understood by "the universe," in short, can only be approached and identified through the use of models, not independently of them.

THE MODEL AND ITS CONSTRUCTION

The kind of model that the cosmologist constructs is wholly conceptual rather than material. It consists of different sorts of symbols including ordinary language, mathematical language, diagrams, and charts, all of which will normally be employed in presenting a given model. A model of the universe is not something that can be directly visualized or completely represented in a pictorial diagram. Consider, for example, a typical model in which use is made of a geometric mode of representation according to which the galaxies are treated as a set of mathematical points that trace out a set of geodesic curves in space-time. In this case, the metric of this set of points is given by the general Robertson-Walker expression for the space-time interval (ds):

$$ds^2 = dt^2 - \left(\frac{R^2(t)}{c^2}\right)\left\{\frac{dr^2 + r^2\,d\theta^2 + r^2\sin^2\theta\,d\phi^2}{(1 + kr^2/4)^2}\right\}$$

where $R(t)$ is the expansion factor, k is a constant whose value determines whether space is Euclidean or non-Euclidean, c is the velocity of light, and r, θ, and ϕ are spatial coordinates. In addition to the specification of purely geometric or kinematic features, which are specified by introducing appropriate values for the curvature constant (k) and the expansion factor ($R(t)$), a model will also require some assignment of specific dynamic or gravitational properties to the entities thus represented. Additional formulas will then be required, and these will normally involve relativity theory or some equivalent branch of physics. It is clear that however much a simple diagram making use of dots and lines may serve to give us a visual representation of what we are talking about, this hardly suffices to encompass all those additional features of the model not included in the diagram.

Although the cosmologist cannot inspect the original, the universe itself, he nevertheless undertakes to make a model of it. How is this done? The answer is to be found by noting the various clues and sources to which the cosmologist appeals in determining the properties to be assigned to his model. These are of two principal types: observational clues provided by the astronomer, and theoretical principles thought to be of relevance to the cosmologic problem.

OBSERVATIONAL CLUES. In general, the observational data the astronomer gathers aid the cosmologist by suggesting ways of assigning certain idealized properties to the model, by providing empirically ascertained values for the constants and variables in the model, and by offering tests for the adequacy of the model as a tool for predicting observable matters of fact.

Idealized properties. The kinds of entities and their properties that the astronomer observes suggest to the cosmologist the lines to follow in developing a simplified and idealized conception of the universe. Let us take some examples. The galaxies, though of enormous physical bulk, may be considered for purposes of the model as

particles making up a continuous and perfect fluid. The advantage of treating the galaxies in this fashion is that it permits a great simplification of the problem, to which readily available mathematical tools of representation and calculation may be applied. Here, of course, the cosmologist adopts a technique that is universally adopted in other branches of physical science and with similar justification. If necessary, suitable corrections to this idealization can always be introduced when application is made of the model to "describe" the actual universe.

An important feature of the domain of galaxies already observed is their spatial distribution. The actual spatial distribution of the galaxies is roughly homogeneous and isotropic when fairly large volumes of space are considered. On a smaller scale, departures from homogeneity become more noticeable, in the clustering of the galaxies, for example. When still smaller volumes of space are investigated, homogeneity breaks down altogether. In general, then, the claim to the uniformity of distribution of galaxies can be upheld only if one takes a sufficiently large unit of volume, say 3.5×10^8 parsecs in diameter. Yet in constructing his model, and as a first approximation, the cosmologist will assign a complete homogeneity to his model of the universe. The expression "cosmological principle" is commonly used to designate this feature of spatial homogeneity. Models that satisfy this cosmological principle, and thus possess the feature of spatial homogeneity, are known as uniform model universes.

When put into mathematical language, a uniform model universe is one possessing a constant curvature at a given moment of time. In the language of general relativity theory, since the density and pressure of material that make up the model are the same in all volumes of space at a given time, whatever their size, a geometric representation of this fact will involve the use of one or another of the spaces of constant curvature. All segments of space of the entire universe will have the same curvature. Such a model clearly requires a process of idealizing and simplifying the spatial distribution of bodies actually observed. For if we were to use the language of geometry to describe the actually observed spatial distribution, we should have to note the actual local departures from homogeneity or constancy of curvature.

In constructing a model of the universe that embodies the feature of spatial homogeneity or constancy of curvature, it is not enough to specify what that curvature is at the present moment of cosmic time. A fully determined model requires (in addition to other features) that the spatial properties of the universe be specified for any point in its past or future. Here there are, broadly speaking, two possibilities. According to one, the spatial properties of the universe remain the same at all times; this view is upheld by those who adhere to the "perfect cosmological principle" and use it to define the properties of the steady state model. A second alternative is to adopt the cosmological principle in its more restricted form as designating merely spatial uniformity, as is the case with the orthodox cosmological models of general relativity. For such models, the entire history of the universe, from a spatial point of view, could be specified if one knew just one thing—the rate at which the distance between any two galaxies changes with time. In a universe characterized simply by the cosmological principle, since an observer would always find a spatially isotropic distribution of particles about him, the only basic feature subject to change is a temporally noticeable feature, namely, changes in the density of the distribution of particles. Such changes in the density might then serve to define a cosmic "clock."

Empirically obtained values. A second important function that the appeal to observational data serves in the construction of cosmological models is that of yielding empirically obtained values for some of the constants and variables of theory. For example, in relativistic uniform models of the expanding universe, the defining characteristics of a particular model need to be specified by assigning values for the following quantities: the cosmological constant (λ), the temporal pattern of the universe as determined by evaluating the function $R(t)$, the values for the velocity factor and the acceleration factor in the velocity-distance relation that specify how the galaxies are moving, the density (ρ) and the pressure (p) of the material and energetic content that fills the universe, and the curvature constant (k). Observational evidence is, at present, either not available at all or not accurate enough to give sufficiently precise determinations for all of these terms. The cosmologist must, therefore, use whatever data is available to eliminate those models that are incompatible with present observations and to suggest lines of inquiry that will help to narrow the field down to those models that can be further tested by observation.

One overall condition for the acceptance of a model is, of course, the consistency of the empirically obtained values it proposes. In a particular model, a particular combination of empirically assigned curvature and density values, for example, may lead to a calculated "age of the universe" that will be inconsistent with an independently obtained estimate for the lower bound of such an "age"; the estimated time scale of the universe will be too

short. In general what is sought is a model all of whose empirically ascertainable values are mutually consistent within the available limits of accuracy.

Empirical tests. Finally, as a natural extension of the point just made, we see how the data obtained by the astronomer serve to test the calculated numerical values for quantities appearing in the cosmologist's equations or other qualitative predictions made on the basis of a given model. Thus whether the extremely remote galaxies at the horizon of the now observable population of galaxies have roughly the same characteristics as those that are nearer is an important question much discussed at the present time as a means for evaluating the rival claims of the steady state and evolutionary theories. The steady state theory claims that galaxies that are at the outer limits of observability should have roughly the same characteristics as those at lesser distances. According to various "evolutionary" models, those same remote galaxies, from which we receive light and other forms of radiation emitted billions of years ago, could, in effect, tell us about the earlier stages of the evolution of the universe. Since conditions at the time of emission were presumably different from what they are now, these very remote galaxies should display differences from those that are nearer to us in at least some of their properties, and these differences should give us valuable clues about the course of development of the universe as a whole. In this regard a number of delicate questions that are the subject of much controversy have arisen in current research.

THEORETICAL IDEAS. A second major source of ideas in the construction of cosmological models is to be found in the conceptual resources of mathematical physics. Here there are two broad possibilities that confront the cosmologist.

Use of established principles. As a first possibility, the cosmologist may turn to some already established body of physical theory as expressed in fundamental principles and derived laws. Such theory will normally have already been found to be successful in dealing with a variety of physical problems of lesser scope than, and wide differences from, the purely cosmological problem. The cosmologist will nevertheless propose to see to what extent the same general body of ideas may be used when applied to the distinctive subject matter of cosmology. He will investigate to what extent the universe as a physical system has a detailed structure that may be articulated and specified by means of the selected physical theory. For example, he may use Newtonian mechanics to construct a model of the universe. Isaac Newton himself drew, in a general qualitative way, the cosmological consequences of using the inverse square law of gravitation as a guide. He argued that the universe, throughout its infinite space, must be filled by a more or less evenly distributed matter. For if all the matter that exists were to be confined to a finite "island" in an infinite "ocean" of space, it would have a center of mass toward which, in time, all matter would move by gravitational attraction. The fact is that no such motion is found, and Newton concluded, therefore, that matter is distributed uniformly throughout an infinite space.

At the present time, the primary and predominant source to which the cosmologist turns is the general theory of relativity as expressed in Albert Einstein's general field equations. These equations specify the relations between the space-time metric of any physical domain and its material or energetic content. The discovery of solutions to those field equations that are of special relevance to the cosmologic situation has led to the construction of several varieties of relativistic models. The other major use to which the field equations of general relativity have been put takes the form of the Schwarzschild solution. It was this solution that first afforded the opportunity for testing the predictive and explanatory powers of the theory as a whole. Karl Schwarzschild's solution is particularly applicable to a physical system such as we encounter in the solar system, namely, a single massive particle (the sun) in whose neighborhood we may study the behavior of much smaller masses (the planets) and light rays. The success of its predictions and explanations has been the primary basis for the confidence placed in the general theory.

To return then to cosmology: Within the broad class of homogeneous, or uniform, model universes we may distinguish the nonstatic models and the static models. Among the static models is Einstein's original model of 1917, which pictured the universe as finite and unbounded; in the light of the subsequent discovery of the mutual recession of the galaxies, it is no longer considered adequate. The nonstatic models include the ever-expanding-universe models that originate from zero or in some finite volume, and oscillating models that undergo alternate contractions and expansions. Within each of these groups, individuating characteristics for a particular model are to be found in the choice of values for the curvature, age, density, and cosmological constant. No single model has as yet been universally adopted.

Creation of new principles. The other broad possibility for furnishing theoretical ideas for cosmologic models is one in which the cosmologist, instead of appealing to

already established principles or laws, for example those of relativistic mechanics, undertakes to create afresh basic principles thought to be of special relevance to the cosmologic problem. By way of illustration, there is the conflict of the 1930s and 1940s between the way in which E. A. Milne sought to establish his kinematic model and the more orthodox procedures of relativistic cosmology. Although Milne did use the formulas of special relativity, he did not take these over directly from Einstein's own presentation; Milne attempted instead to derive them from what he thought of as more basic and primitive postulates. These postulates, he claimed, state the conditions for the measurement of time and for the communication of results by different timekeepers and observers.

A more recent example of the same sort of procedure is the steady state model of the universe proposed by Bondi and Thomas Gold in 1948. In support of this model, it is argued that since the universe is unique, there is no reason to believe that the laws which apply to smaller-scale physical phenomena, for example in laboratory terrestrial physics, or even in the domain of gravitational phenomena in the solar system, need be expected to apply to the universe as a whole. Therefore, instead of taking such laws as the point of departure in investigating the physical properties of the universe as a whole, it is suggested that the cosmologist can actually enjoy a far greater freedom than is believed possible in orthodox relativistic cosmologies. Let the cosmologist adopt any "laws" or principles which he believes are appropriate to the study of the universe as a whole, even though these may not have been established or confirmed in other (smaller-scale) areas of physical phenomena. The important thing is to see whether using these laws and principles leads to confirmable empirical results and whether they help to increase and deepen our understanding of the universe.

Those who favor this view (Bondi and Gold among others) determine some of the major features of the steady state model by appeal to the specially introduced postulate known as the perfect cosmological principle. This principle was not in prior use in other branches of physics but was introduced because of its special relevance to cosmology. (Fred Hoyle's model of the steady state universe proceeds along more conventional lines, at least in this respect. Although it differs from the expanding universe models of general relativity in abandoning the principle of the conservation of matter—in order to make possible the idea of the continuous creation of matter—it appeals for its basic physical principles, although in modified form, to the field equations of general relativity.)

One general motive that seems to inspire the setting up of specially devised principles for cosmologic models is the desire to show that the science of cosmology is basic to all other physical sciences. Instead of appealing to other branches of physics for principles to be used in describing the features of the universe as a whole, it is thought desirable that one should be able, eventually, to show that the laws of ordinary physics can be linked with the properties of the universe as a whole. The universe would then disclose itself to be a unitary physical system within which it would be possible in principle to deduce ordinary physical laws from the principles of cosmology. Milne undertook to show how, for example, the inverse square law of gravitation, among other things, could be deduced from such more fundamental cosmological ideas. Similarly, within the framework of a steady state model, Sciama attempts to show how the local inertial properties of matter can be linked (as Ernst Mach originally proposed) with the distribution of masses in the universe at large.

From a logical point of view, there is no reason to discourage such efforts. On the contrary, the realization of such a goal would be of immeasurable significance for all of science, and one should in logic suspend judgment until such a program can be carried through with some fair degree of success.

Meanwhile, it is necessary to point out that some of the writers who favor this approach put methodological interpretations on the use and warrant for the specially devised principles that are not acceptable, whatever the eventual success or promise of the program as a whole. Thus Milne and Bondi, who support different models, are each concerned to stress what they take to be a special method for cosmology—as contrasted with other branches of physics. Milne, for example, thought of ordinary physics as employing an inductive method, whereas cosmology, he believed, should be based on a deductive method. Cosmology, he argued, should not employ the laws of ordinary physics to the extent that these are inductively warranted. This was his major complaint against what he took to be the faulty procedure of relativistic cosmologies founded on the "inductively established" principles of general relativity. In making this claim, Milne was in error, since the principles of general relativity theory are not, as he thought, ordinary inductive generalizations.

In fact, Milne's own appeal to "self-evidence" as the warrant for introducing his preferred cosmological prin-

ciples must be rejected, for the appeal is groundless and fails to support the certainty and uniqueness which he claimed for his principles. In constructing a model of the universe, the cosmologist is engaged in setting up a theoretical tool for dealing with the facts of observation. Whether he gets his theoretical principles by "borrowing" them from some other branch of physics or whether he creates them especially for the problem at hand is of secondary importance to what he does with these principles once he has them and how he evaluates the results he achieves. There is a common method that characterizes cosmology regardless of the particular model being proposed or favored, and it is precisely the same method which is employed in other branches of physics. Moreover, the same criteria of evaluation need to be brought to bear in the appraisal of results in cosmology as in other areas of science. Far from including any appeal to self-evidence or to similar rationalistic demands, a satisfactory model requires the constant support provided by observational evidence.

COGNITIVE WORTH OF MODELS

Consideration of the goals set by scientific cosmology gives rise to a central philosophical question, that of determining the cognitive worth of any cosmological model. This is an epistemological question and may be put in terms of the traditional issues separating the realist and the conceptualist (or the instrumentalist). Should we say with the realist that cosmological models offer us an account of the structure of the independently existing universe, or, rather, should we say with the conceptualist that these models are simply useful means of presenting and interpreting observational data?

As a basis for clarifying the issue at hand, it will be helpful to point out a fundamental ambiguity in the use of the term *universe* itself. Employed without the qualifying adjectives "observed" or "observable," it may have at least two quite distinct senses. One meaning of *universe* is "that to which the observed universe belongs"; another is "that which is characterized by a cosmological model." So far as the realist is concerned, the two meanings are equivalent; in his view the universe defined by a cosmological model will be the same universe as the one described by the expression "that to which the observed universe belongs." But the realist's view of cosmological models cannot be assumed in advance to be the only tenable one. Thus the distinction suggested here has the merit of permitting us to keep this question open. If later a realist philosophy is accepted, the appropriate modifications can be made. Clearly, we do not need to commit ourselves to the position that everything properly said of "the universe" in the sense of "that which is characterized by a cosmological model" can also be said of "the universe" in the sense of "that to which the observed universe belongs." For example, we might want to attribute the property of being "a whole" or "an absolute totality" to the universe as characterized by a particular model, but not to the universe in the sense of that to which the observed universe belongs.

Cosmology aims at articulating the character of the universe *as* a whole. To that extent, then, it rests on the methodological postulate that the universe *is* a whole. The specific character of the whole will, of course, be variously described by various models. What remains fixed, however, is the assumption that the goal of cosmology is to characterize the universe *as* a whole. Therefore the statement "The universe is a whole" is in this context an analytic statement, a matter of definition. But note that it is a definition in which "the universe" is used to signify "that which is characterized by a cosmological model." Not only does cosmology require that, as a matter of definition, the universe be thought of as a whole (in the sense of being intelligible in the way that mathematical classes, geometrical relations, or physical systems are); it also postulates that the universe as a whole is unique or absolute. This means that there is just one such class, pattern, or system, and that all other physical processes or systems of lesser duration or spatial extent are to be taken only as parts of this all-embracing whole. Since each model will so define the universe, it would be a misuse of language to speak of a plurality of universes. Again, of course, the precise structure of this unique or absolute whole will, at least in some respects, vary from model to model.

But what if "the universe" means "that to which the observed universe belongs"? Is the statement "The universe is an absolute, unique whole" still analytic? To this the answer must be no. For when we use "the universe" in this sense, we move from methodology to ontology. In contrast to the case of the universe as defined by a cosmological model, we are no longer committed by the basic methodological postulate of cosmology to saying that the universe *is* a whole. True, in setting up a science, it may be necessary to presuppose the existence of some pervasive structure as the object of study. Yet such a presupposition need not be binding on what the universe is existentially. So long as "the universe" means simply "that to which the observed universe belongs," nothing in this meaning contains analytically the notion of its being a "whole" or an "absolute whole." Indeed, even if we grant

that the observed universe is structured in some manner, this does not entail that the wider universe of which it is a part is also pervasively structured. Nor does the fact that we describe the observed universe as "part of" or "that which belongs to" something else require us to say that the universe to which it belongs is a unique or absolute whole. For our reliance on such terms as *part, whole*, and *belong* reveals merely that the mind, in reaching into the unfamiliar, must use analogies in order to relate the unfamiliar to what it already knows.

The universe as the "something more" than the observed universe may well be a complete, unique and intelligibly structured whole. But the claim that we are able to say so is something to which we need not commit ourselves. It is better left as an open question, since, strictly speaking, it is one on which we neither have nor can have any knowledge. Stipulating an affirmative answer by definition does not, of course, establish such knowledge.

See also Creation and Conservation, Religious Doctrine of; Einstein, Albert; Mach, Ernst; Newton, Isaac; Relativity Theory; Taylor, Alfred Edward; Time; Whitehead, Alfred North; Wolff, Christian.

Bibliography

HISTORICAL SURVEYS

De Sitter, W. *Kosmos*. Cambridge, MA: Harvard University Press, 1932.

Koyré, A. *From the Closed World to the Infinite Universe*. Baltimore: Johns Hopkins Press, 1957.

Kuhn, T. S. *The Copernican Revolution*. Cambridge, MA: Harvard University Press, 1957.

Munitz, M. K., ed. *Theories of the Universe from Babylonian Myth to Modern Science*. Glencoe, IL: Free Press, 1957.

SYSTEMATIC SURVEYS

Bondi, H. *Cosmology*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1961.

Bondi, H. et al. *Rival Theories of Cosmology*. London: Oxford University Press, 1960.

Bonnor, W. B. *The Mystery of the Expanding Universe*. New York: Macmillan, 1964.

Couderc, P. *The Expansion of the Universe*. London: Faber and Faber, 1952.

Eddington, A. *The Expanding Universe*. Cambridge, U.K.: Cambridge University Press, 1933.

Hubble, E. P. *Observational Approach to Cosmology*. Oxford, 1937.

Hubble, E. P. *The Realm of the Nebulae*. New Haven, CT: Yale University Press, 1936.

Lemaitre, G. *The Primeval Atom*. New York: Van Nostrand, 1950.

McVittie, G. C. *Fact and Theory in Cosmology*. London: Eyre and Spottiswoode, 1961.

Milne, E. A. *Modern Cosmology and the Christian Idea of God*. Oxford: Clarendon Press, 1952.

Milne, E. A. *Relativity, Gravitation, and World-Structure*. Oxford: Clarendon Press, 1935.

Sciama, D. *The Unity of the Universe*. Garden City, NY: Doubleday, 1959.

Tolman, R. C. *Relativity, Thermodynamics and Cosmology*. Oxford: Clarendon Press, 1934.

Whitrow, G. J. *The Structure and Evolution of the Universe*. London: Hutchinson, 1959.

PHILOSOPHICAL PROBLEMS

Bondi, H. "Fact and Inference in Theory and in Observation." In *Vistas in Astronomy*, edited by A. Beer. London: Pergamon Press, 1955. Vol. I.

Bondi, H. "Philosophical Problems of Cosmology." In *British Philosophy in the Mid-Century*, edited by C. A. Mace. London: Allen and Unwin, 1957.

Davidson, W. "Philosophical Aspects of Cosmology." *British Journal of the Philosophy of Science* 13 (50) (August 1962).

Dingle, H. "The Philosophical Aspects of Cosmology." In *Vistas in Astronomy*, edited by A. Beer. London: Pergamon Press, 1955. Vol. I.

Finlay-Freundlich, E. "Cosmology." In *International Encyclopedia of Unified Science* Vol. I, No. 8. Chicago: University of Chicago Press, 1951.

Harré, R. "Philosophical Aspects of Cosmology." *British Journal of the Philosophy of Science* 13 (50) (August 1962).

Johnson, M. *Time, Knowledge and the Nebulae*. London: Faber and Faber, 1945.

McVittie, G. C. "Rationalism and Empiricism in Cosmology." *Science* 133 (21) (April 1961).

Monist 47 (1) (Fall 1962). Entire issue.

Munitz, M. K. *The Mystery of Existence*. New York: Appleton-Century-Crofts, 1965.

Munitz, M. K. *Space, Time, and Creation*. Glencoe, IL: Free Press, 1957.

Toulmin, S. "Contemporary Scientific Mythology." In *Metaphysical Beliefs*, edited by S. Toulmin, R. W. Hepburn, and A. MacIntyre. London: SCM Press, 1957.

Milton K. Munitz (1967)

COSMOLOGY [ADDENDUM]

The key issue identified by Milton K. Munitz, whether cosmology should center on description or explanation, remains a central philosophical theme in the early twenty-first century. However, the context has changed dramatically since he wrote his entry, and that has changed the implications.

First, massive new data sets are available because of the extraordinary improvement of telescopes, detectors, and computer technology in the past three decades. Not only are optical, ultraviolet, and infrared observations of

galaxies possible in determining luminosities and spectra with unprecedented sensitivity but radio, X-ray, and gamma ray sky surveys are also possible. Galaxies have been detected up to a redshift of 6 and many quasar-stellar objects as well as multiple images of distant gravitationally lensed galaxies have been identified. Besides large-scale number-count and redshift surveys, the background radiation spectrum at all wavelengths have been measured. A key feature has been identification of cosmic blackbody radiation (CBR) with a perfect 3 degrees Kelvin (K) blackbody spectrum, which is isotropic to one part in 104 after allowing for Earth's motion relative to the rest frame of the radiation. Detailed observations have mapped its anisotropies over the whole sky at a sensitivity of better than one part in 105. Cosmology has changed from a data-poor to a data-rich subject.

Second, theory has developed dramatically, largely in a symbiotic relation with the observations, being used both to interpret them and to suggest new observations. This has happened in relation to gravitational theory, as regards astrophysics, and in relation to various branches of high-energy physics. This has changed the texture of cosmology from being essentially an exercise in geometry with an admixture of philosophy, to being a rich theoretical subject with relations to many branches of physics.

The cosmological application of Albert Einstein's theory of gravitation has been developed simultaneously. This theory predicted there must have been a start to the universe, but it was not clear if this was simply because of the special geometry of the standard (Friedmann-Lemaître) models of the universe, which are exactly isotropic and spatially homogeneous. More realistic models might show the prediction was a mathematical artifact. The singularity theorems developed by Roger Penrose and Stephen Hawking showed this was not the case: even for realistic geometries, classical gravitational theory predicts a beginning to the universe at a space-time singularity. Sophisticated perturbation theory was developed to underlie the theory of structure formation in the expanding universe, and dynamical systems studies related the behavior of whole classes of anisotropic models in suitable phase spaces, enabling identification of generic patterns of behavior.

The cosmological constant is a possible repulsive force term in the Einstein field equations. After Einstein discarded it in the 1930s, it was usually assumed to be zero until a decade ago, when observations of supernovae in distant galaxies showed the expansion of the universe is speeding up. The gravitational equations then imply presence of a cosmological constant or some equivalent form of dark energy. There is no known physics reason why this force should exist at this level where it is just detectable—quantum field theory suggests it should be enormously larger. Perhaps the problem is that the wrong theory of gravity is being used.

Astrophysics studies the physics and evolution of galaxies, clusters of galaxies, and any intergalactic medium there may be. It has led to three important deductions. First, both the universe itself and the matter in it are evolving: Radio source counts preclude the steady state model proposed by Hermann Bondi, Thomas Gold, and Sir Fred Hoyle. Second, the dynamics of galaxies and clusters of galaxies is only compatible with observations if there is a large amount of dark matter present, felt by its gravitational effects but not emitting detectable radiation. In terms of the dimensionless density parameter, $\Omega_{dm} = 0.3$. This is less than the critical value = 1 separating universes that recollapse in the future ($\Omega < 1$) from those that expand forever ($\Omega > 1$) when the cosmological constant vanishes, but much greater than that for visible matter ($\Omega_{vm} = 0.02$). Third, the CBR is relic radiation from a hot early state of the universe, with matter and radiation held in tight equilibrium when the universe is ionized at redshifts greater than z_{dec} 1100, but separately evolving for lower redshifts when electrons and nuclei are combined into atoms. The matter at $z = z_{dec}$ formed the last scattering surface (LSS) in the early universe and emitted the CBR. The universe is opaque for $z > z_{dec}$ but transparent for $z < z_{dec}$. Thus, the LSS delineates the visual horizon: One is unable to see to earlier times than its occurrence (because the early universe was opaque) or to detect any matter at larger distances than that one sees on the LSS (because of the speed of light limit on propagation of information).

The physics of the early universe (before decoupling) can be thought of in three stages. The hot big bang stage is the last one, when matter and radiation cooled from a high temperature ($> 10^{12}$K) to 4000K on the LSS. Nuclear physics processes with pair production and weak interactions lead to a well-understood physical evolution. At about 109K nucleosynthesis took place: the creation of the light elements (deuterium, helium, and lithium) from protons and neutrons (heavy elements such as carbon only formed later in the interior of stars). Theory and observation are in excellent agreement provided the density of baryons is low: $\Omega^{bar} = 0.02$. This leads to an important conclusion: the much more abundant dark matter detected astrophysically cannot be ordinary matter made up of baryons. Many attempts have been made to identify its nature, but this is still unknown.

Particle physics processes dominated the preceding era when exotic processes took place such as combination of a quark-gluon plasma to produce baryons. Quantum field theory effects were significant then, and this leads to an important possibility: Scalar fields producing repulsive gravitational effects could have dominated the dynamics of the universe at those times. This leads to the theory of the inflationary universe, proposed by Alan Guth: an extremely short period of accelerating expansion preceding the hot big bang era, leading to a cold and smooth vacuum-dominated state and ending in reheating: conversion of the scalar field to radiation, initiating the hot big bang epoch. This inflationary process is claimed to explain some philosophical puzzles: why the universe is so special (with spatially homogeneous and isotropic geometry and a uniform distribution of matter) and why the space sections are so close to being flat (the sign of the spatial curvature is still unknown). This theory led to a major bonus: a proposal that initial tiny quantum fluctuations were expanded to such a large scale by inflation that they provided seeds initiating growth by gravitational attraction of large-scale structures such as clusters of galaxies. This theory makes clear observational predictions for the spectrum of CMB anisotropies, which have since been spectacularly verified.

Quantum gravity processes are presumed to have dominated the earliest times, preceding inflation. There are many theories of the quantum origin of the universe, but none has attained dominance. The problem is that there is not a good theory of quantum gravity, so all these attempts are essentially different proposals for extrapolating known physics into the unknown. A key issue is whether quantum effects can remove the initial singularity and make possible universes without a beginning. Preliminary results suggest this may be so.

Thus, the present dominant cosmological paradigm is a quantum gravity era followed by inflation, leading on to the hot big bang epoch and finally the observable universe domain. A wealth of observations supports this dominant theory, but some theoretical proposals are being made that have no observational support; sometimes it will be impossible to ever obtain such support. This happens both as regards physics and geometry.

The limits of physics testing are reached because accelerators on Earth attaining the energies relevant in the early universe cannot be constructed. Consequently, when considering physical processes at the time of inflation and earlier, the extrapolation of known physics into the unknown is relied on. Some things are assumed to be unchanged (e.g., the use of standard variational principles) and others as indefinitely mutable (any potential that is convenient may be used). Thus, the claimed link of cosmology to high-energy physics is potential rather than real. In particular, no specific particle or field has been identified as underlying inflation, and no experimental test is possible for the various mechanisms proposed for creation of the universe. There is only one universe, and what happens during a rerun with the same or different initial conditions cannot be observed.

The limits of what is observable is given by the visual horizon, as discussed by Munitz. However, most cosmological models make predictions of what lies beyond, and these predictions can never be observationally verified. This is particularly important in the case of the chaotic inflation theory proposed by Andrei Linde, which claims that the expanding universe domain is imbedded in a still-exponentially expanding inflationary universe in which there are embedded countless other expanding universe domains similar to this one, the whole forming a fractal-like structure. However those other domains cannot be seen, so this model is observationally unverifiable; furthermore, the underlying physics is experimentally untested. Adherence to this model implies the victory of theory over experimental tests and observational verification.

There is one case where this kind of spatial observational limit does not obtain. This is when a small universe occurs, that is, a universe that closes up on itself spatially for topological reasons and does so on such a small scale that one can see right round the universe since the time of decoupling. Then one can see all the matter that exists, with multiple images of many objects occurring. This possibility is observationally testable, and indeed there are weak hints in the CMB anisotropies that this could actually be the case. Checking if this is possible or not is an important task; the nature of the observational relationship to the universe is fundamentally different if it is true.

A key issue is whether the universe is special in geometrical and physical terms. The assumption that the universe is geometrically special was encoded in Edward Arthur Milne's cosmological principle, taken as a founding principle in cosmology until the 1960s. Then Charles W. Misner introduced the idea of a universe with generic initial conditions that was isotropized at later times by physical processes, giving a physical cause for its special geometry. This concept became central to the inflationary universe paradigm, but is only partially successful because universes that are anisotropic enough may never inflate and may not lead to ordinary thermodynamics. Some degree of geometric specialty must have occurred.

As regards physics, the key point is that only a restricted set of physical laws and boundary conditions will allow life to exist. This has led to the anthropic question: Why does the universe have that special character that is favorable to life? The only physically based answer is that there physically exists a multiverse: an ensemble of numerous universes with varying properties. If there is a large enough ensemble with enough variation, it becomes virtually certain that some of them will just happen to get things right so that life can exist. However, despite various ingenious suggestions, this proposal is observationally and experimentally untestable. Adherence to the multiverse idea to gain explanatory power is another triumph of theory and explanation over observational testing and description. There is no way to determine the properties of any other universe in the multiverse if they do indeed exist, for they are forever outside observational reach.

Some proposals claim there may be an infinite number of universes in a multiverse and many cosmological models have spatial sections that are infinite, implying an infinite number of particles, stars, and galaxies. This proposal involves, however, an idea that some have argued is incoherent. Thus, David Hilbert (1964), for example, argues that infinity is quite different from a very large number: The word *infinity* denotes a quantity or number that can never be attained, and so will never occur in physical reality. If so, then this last proposal is not only unverifiable, but such as cannot be true. On the contrary, many other scholars—such as José A. Benardete (1964)—argue that an actual infinity of things is possible.

See also Einstein, Albert; Hilbert, David; Infinity in Mathematics and Logic; Quantum Mechanics; Philosophy of Physics.

Bibliography

Barrow, John D., and Frank J. Tipler. *The Anthropic Cosmological Principle*. New York: Oxford University Press, 1986.

Benardete, José A. *Infinity: An Essay in Metaphysics*. Oxford, U.K.: Clarendon Press, 1964.

Harrison, Edward. *Cosmology: The Science of the Universe*. 2nd ed. New York: Cambridge University Press, 2000.

Hawking, Stephen W., and George F. R. Ellis. *The Large Scale Structure of Space Time*. Cambridge, U.K.: Cambridge University Press, 1973.

Hilbert, David. "On the Infinite." In *Philosophy of Mathematics*, edited by Paul Benacerraf and Hilary Putnam, 134–151. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Leslie, John. "Anthropic Principle, World Ensemble, and Design." *American Philosophical Quarterly* 19 (1982): 141–152.

George F. R. Ellis (2005)

COSMOPOLITANISM

When the Cynic philosopher Diogenes of Sinope (c. 412–c. 323 BCE) was asked where he came from, he said "I am a citizen of the world" (Diogenes Laertius, *Lives of Eminent Philosophers*, bk. 6, chap. 63). The Greek term is "kosmopolitēs," the source of the English word "cosmopolitan." Cosmopolitanism is actually a range of views—moral, political, and cultural—affirming the importance and value of the community of all human beings. Against particular and local allegiance to the polis, city-state, or modern nation-state, the cosmopolitan would emphasize a general and far-reaching concern for humanity.

It remains unclear whether Diogenes' own view was meant to affirm a positive duty to humanity or only to deny the conventional obligations of citizenship associated with the polis. But the Greek Stoics, such as Zeno of Citium and Chrysippus in the third century BCE, developed the tradition by identifying the law of the cosmos with divine reason and extending world citizenship to everyone who lives in accordance with it. Roman Stoicism—especially as developed by Cicero, Seneca, Epictetus, and Marcus Aurelius—strongly influences modern cosmopolitanism by counting the possession of reason as a sufficient condition of membership in this foremost ethical community. Marcus Aurelius developed the idea of natural law as the common law of the polis of which all human beings are fellow citizens (*Meditations*, bk. 4). Nonetheless, the Roman Stoics readily acknowledged duties to one's country along with duties to humanity as a whole.

With advances of natural-law theory in the seventeenth century, international law, or the law of nations, got its first explicit modern statement in the theories of Hugo Grotius and Samuel Von Pufendorf. In the eighteenth century, Immanuel Kant, partly inspired by Stoicism, viewed all persons as members of a single community of rational agents, each of whom is free, equal, and independent. On these grounds he strongly criticized European colonialism and imperialism. In *Perpetual Peace* (1795), Kant argued for a federation of republics, each recognizing the human rights of all persons. (See Heater 1996 for a history of cosmopolitan thought.)

Cosmopolitanisms, as sets of moral, political, and cultural views, have developed significantly in the late twentieth century. Below are some of the most important arguments and distinctions made in recent debates, with particular emphasis on the core moral claims.

MORAL COSMOPOLITANISM

Moral cosmopolitanism is characterized by three basic commitments. First, it is a species of moral individualism, maintaining that the basic units of moral concern are human individuals rather than groups or other collectivities. Second, it is egalitarian, holding that each individual counts equally from a moral perspective, that is, that no person is worth more than any other and that every person is entitled to equal consideration. Finally, cosmopolitans are moral universalists, who believe that the proper scope of moral concern encompasses all persons, regardless of their ethnic, racial, cultural, religious, and national affiliations. In short, moral cosmopolitanism affirms the equal worth of every human individual, quite apart from any subgroup to which they might belong, along with a commitment to impartial concern.

The great interest in these ideas is for their possible implications for an account of the basic moral and political obligations of persons. A dominant puzzle is the apparent contradiction between (1) widely recognized special obligations and associative duties, for example, ties to one's family members, friends, fellow citizens, and compatriots, and (2) general duties to individual human beings, regardless of membership in any of these communities. How are special duties compatible with the requirement of equal concern?

To address this question, it will be useful to flag two truths about cosmopolitanism. First, it can be defended by a deeper moral theory, including utilitarianism, a theory of human rights, contractarianism, and a Kantian account of fundamental obligations explained ultimately by the categorical imperative. Such defenses are exemplified by some of the most notable recent thinkers in this tradition: the utilitarian Peter Singer, the human-rights theorists Henry Shue and Thomas Pogge, the contractarian Brian Barry, and the Kantian Onora O'Neill. Charles Beitz is less easily classified: His moral cosmopolitanism at times has drawn on contractarian thought and lately has issued in a sustained focus on human rights as the appropriate language of international justice.

The second truth about cosmopolitanism is that it can come in strong and moderate varieties, both sharing a commitment to helping other human beings regardless of citizenship, nationality, ethnicity, race, religious affiliation, and geographical location (Scheffler 2001). Strong moral cosmopolitans believe that universalist, egalitarian individualism entails that the basic moral claims of all human beings are the same, and that any special regard for some persons over others must be justified by the role such regard plays in promoting the good of the human community as a whole. As the prominent cosmopolitan Martha Nussbaum has said, the reason a cosmopolitan should show additional concern for the locals or fellow nationals "is not that the local is better per se, but rather that this is the only sensible way to do good" (1996, pp. 135–136). Moderate moral cosmopolitans, on the other hand, believe both that there are basic obligations toward all other human beings that each of us must recognize, and that particular affiliations—to family, nation, state, and so on—give rise to special duties justified independently of any instrumental value for promoting the good of humanity. On this view, associated duties do not derive from our universal duties to human beings in general.

David Miller's defense of nationality (1995, 2002) is a good example of a view strongly opposed to the idea that we have the same duties to each person in the world. For Miller, a nation is a community of belief, extended in history, active in character, tied to a particular homeland, and associated with a distinctive public culture. Compatriots share a common national identity and possess special reasons for recognizing duties to one another beyond those to persons generally. For one thing, nation-states involve institutionalized reciprocity, in which members contribute their efforts and wealth to the community for the benefit of fellow members. For another thing, nation-states pursue collective cultural projects involving distinctive choices about work, religion, and culture more generally, and these projects give rise to nationally different mixtures of burdens and benefits. The upshot of these two points is that fellow nationals owe to one another a range of duties that they do not owe to nonmembers, but these duties are compatible with the view that each person is due equal concern in virtue of their being human.

This dispute about the nature of cosmopolitan morality becomes especially acute in matters concerning just distribution. Here a range of views seem to deny the force or extent of cosmopolitan justice. Theorists such as Michael Walzer (1983) claim that these duties of distributive justice make sense only within the context of a community, such as a nation, within which the goods to be distributed are produced and shared. David Miller (1998) has defended the related view that some principles of distributive justice are comparative and some are noncomparative, and the comparative principles can apply only within communities and not globally. John Rawls (1999) has argued that the scope of distributive justice should be limited to the basic structure of a particular society conceived of as a self-contained cooperative venture for mutual advantage. On his view, justice beyond the nation-state is concerned with interstate rules aimed at

promoting toleration and peace worldwide, but questions about the distributive entitlements of particular individuals considered as such can gain no footing.

Such views seem to go beyond merely rejecting strong moral cosmopolitanism; they offer positive views that seem to minimize the substance of global duties of distributive justice. But it is precisely on the basic justice-related claims of individuals that something like strong cosmopolitanism appears most plausible. If one believes that all human beings possess the same rights to be free from torture, persecution, hunger, and homelessness, it seems natural to infer that our duties as human beings include aiming to bring about a world in which these rights are protected and promoted to the same extent for each and every person in the world. No amount of reciprocity between fellow nationals in one country can generate special duties to each other when there are countless foreign nationals suffering from deprivations of their basic interests. Moral cosmopolitanism and the idea of justice itself seem to share a fundamental commitment to impartial concern for all persons affected by an institutional framework. In a worldwide network of social, political, and economic institutions, distributive justice demands that each human being on the planet be entitled to concern from a perspective that includes their interests on a par with everyone else's. Special treatment for insiders is legitimate only if it can be justified to those excluded from it (Barry 1998, p. 145).

These sorts of considerations have led cosmopolitans to argue for strong obligations to alleviate the continuing dire suffering and death of millions of our fellow human beings. Peter Singer (1972) has defended the utilitarian view that we are morally required to stop such suffering where we can do so without sacrificing anything of comparable moral importance. The argument emphasizes the moral irrelevance of distance. A dying child on another continent obligates us just as much as a dying child next door. This cosmopolitan aspect of the case has been more readily accepted than the specifically utilitarian aspect of maximizing benefits, the demand for which has seemed difficult to square with commonsense views about the limits of moral obligation.

One influential line of argument proposed by Brian Barry (1973), Thomas Pogge (1989), and Charles Beitz (1999) has suggested that a consistent application of John Rawls's justly famous original-position argument for principles of distributive justice would lead in the direction of strong moral cosmopolitanism. If the Rawlsian veil of ignorance rules out knowledge of facts about oneself that unfairly skew one's choice of principles, then—along with sex, race, class, and conception of the good—citizenship too ought to be obscured from the contractors' considerations. If principles of distributive justice should not privilege or disadvantage people on the basis of characteristics they possess for which they are not responsible, then their citizenship should not affect their life prospects. If Rawls is correct that inequalities should be allowed only when they maximally benefit the worst-off group, the scope of principles of justice should encompass the least advantaged in the world.

POLITICAL COSMOPOLITANISM

There is a long tradition of favoring political institutions beyond the local or national—a view often allied with the need to promote global peace. While the positions of moral and political cosmopolitanism are distinct, political and legal proposals tend nonetheless to be linked to underlying moral views that emphasize the universal scope of concern for the interests of persons. Options for political cosmopolitanism take various forms, each an instance of the general institutional view that authority should be shifted from individual states to supranational political institutions (Beitz 1994, p. 124).

One option would be a single state encompassing the whole world. Immanuel Kant's rejection of a world state has been followed by later theorists, including John Rawls, who concurs with Kant's judgment that such a state would be either a global despotism or the backdrop for unending civil wars. But if duties to other persons have global scope, it seems reasonable to think that global institutions of some sort will be necessary to make sure that those duties are fairly distributed and that they achieve the goal of protecting human beings from avoidable harm.

Another approach is David Held's model of "cosmopolitan democracy" (2004), which envisages not a single world government but a range of reforms of international political and economic institutions in the name of democratic accountability, consent, and inclusiveness. Held's approach is distinctive in its appeal to democracy as the core value of global political legitimacy, but this is questioned by those who rank justice, rather than democracy, as the highest-ranking value underpinning any assessment of global political institutions.

CULTURAL COSMOPOLITANISM

Cultural cosmopolitanism is a view about the conditions under which individuals can generate an identity and live a good life. It emphasizes that cultures are constantly changing and that individuals can benefit from mixing

elements from different cultural traditions. Strong cultural cosmopolitans believe that individuals can live well only by drawing on a range of cultural traditions and practices, while moderate cultural cosmopolitans hold that a range of good lives can be grounded in both this sort of openness to cultural variation and a more traditional, inward-looking existence with its settled cultural commitments (Scheffler 2001, Waldron 1992).

While the strong position is more contentious, both views deny that lives can be good only when lived within the confines of a particular cultural or national tradition. Consequently, this form of cosmopolitanism is relevant to evaluating cultural nationalism and its attendant claim to political self-determination (Caney 2005)

CONCLUSION

The recent flourishing of cosmopolitan thought signals a recognition that any plausible account of politics, morality, distributive justice, and the good life for human beings should take seriously the idea that humanity is a community whose claims on us are both fundamental and far-reaching. A significant project for the future is developing a comprehensive account of the basis and implications of cosmopolitan political morality.

See also Chrysippus; Cicero, Marcus Tullius; Diogenes of Sinope; Epictetus; Grotius, Hugo; Kant, Immanuel; Marcus Aurelius Antoninus; Multiculturalism; Nussbaum, Martha; Postcolonialism; Pufendorf, Samuel von; Rawls, John; Republicanism; Seneca, Lucius Annaeus; Stoicism; Zeno of Citium.

Bibliography

Barry, Brian. "International Society from a Cosmopolitan Perspective." In *International Society: Diverse Ethical Perspectives*, edited by David Mapel and Terry Nardin, 144–163. Princeton, NJ: Princeton University Press, 1998.

Barry, Brian. *The Liberal Theory of Justice*. Oxford: Clarendon Press, 1973.

Beitz, Charles. "Cosmopolitan Liberalism and the States System." In *Political Restructuring in Europe: Ethical Perspectives*, edited by Chris Brown, 123–136. London: Routledge, 1994.

Beitz, Charles. *Political Theory and International Relations*. 2nd ed. Princeton, NJ: Princeton University Press, 1999.

Caney, Simon. *Justice beyond Borders: A Global Political Theory*. Oxford, U.K.: Oxford University Press, 2005.

Heater, Derek. *World Citizenship and Government: Cosmopolitan Ideas in the History of Western Political Thought*. Basingstoke, U.K.: Macmillan, 1996.

Held, David. *Global Covenant*. Cambridge, U.K.: Polity, 2004.

Miller, David. "Cosmopolitanism: A Critique." *Critical Review of International Social and Political Philosophy* 5 (2002): 80–85.

Miller, David. "The Limits of Cosmopolitan Justice." In *International Society: Diverse Ethical Perspectives*, edited by David Mapel and Terry Nardin, 164–181. Princeton, NJ: Princeton University Press, 1998.

Miller, David. *On Nationality*. Oxford, U.K.: Clarendon Press, 1995.

Nussbaum, Martha. "Patriotism and Cosmopolitanism" and "Reply." In *For Love of Country*, edited by Joshua Cohen, 3–17, 131–144. Boston: Beacon Press, 1996.

O'Neill, Onora. *Bounds of Justice*. Cambridge, U.K.: Cambridge University Press, 2000.

Pogge, Thomas. *Realizing Rawls*. Ithaca, NY and London: Cornell University Press, 1989.

Pogge, Thomas. *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms*. Cambridge, U.K.: Polity, 2002.

Rawls, John. *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.

Scheffler, Samuel. "Conceptions of Cosmopolitanism." In his *Boundaries and Allegiances*, 111–130. Oxford, U.K.: Oxford University Press, 2001.

Shue, Henry. *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy*. 2nd ed. Princeton, NJ: Princeton University Press, 1996.

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1972): 229–243.

Waldron, Jeremy. "Minority Cultures and the Cosmopolitan Alternative." *University of Michigan Journal of Law Reform* 25 (1992): 751–792.

Walzer, Michael. *Spheres of Justice*. Oxford, U.K.: Basil Blackwell, 1983.

Charles Jones (2005)

COSMOS

From Anaximander on, early Greek philosophers regarded the structure and regular processes of the world as central to their accounts of nature. However, their understanding of this order differed considerably. These processes might be viewed as harmony or balance and as the result of growth or conflict or an intelligence, or they might be considered the result of random collisions of particles. The order might involve cycles or it might be a single continuous development from a primal state. In some philosophers, order itself exemplifies the goodness of the world. Many of these elements can be found in nonphilosophical cosmologies as well, such as the emergence of the world from waters in the Babylonian *Enuma Elish*, or the Genesis creation story. What distinguishes Greek philosophers is the variety of their attempts to describe the world as ordered, their reflection on what such an account must consist in, their consideration of

the role of divinity in their accounts, and the depth of their attempts to provide rich, unified, explanatory accounts of the world. Scholars do not know who first used the word "kosmos" to describe arrangements of the world or parts of it, but it came to be a common word in denoting this central concept.

Kosmos normally means "fine or beautiful arrangement or order" and can refer to an array of warriors, a hairdo, or a government; by extension, it can apply to cosmetic accessories or even to each of the ten officials in a Cretan senate (that is, the components of an arrangement). In Aristotle's *Poetics*, it is a technical term for the spectacle of a play and also for ornamental diction. Early philosophers used the word to describe an order or arrangement in the world, but later "the kosmos" could refer to the world itself or at least to the most organized part of it, the heaven: "The kosmos is a system consisting of heaven and earth and the natures enclosed in these" (ps. Aristotle, *De mundo* 2).

The oldest extant, philosophical use of the word, to describe the balance of changes, occurs in Heraclitus (fr. 30), although a late doxographer, Aetius, says that Pythagoras was the first to name the enclosure of all things "kosmos." Even if Aetius is right (the claim is rejected by most modern scholars), it is indeterminate whether in each case the philosopher meant to speak of all changing elements, or even all things, as an order (the general use) or instead of the arrangement of all things, the world (the privileged use). The latter seems unlikely in early authors, but becomes probable when Empedocles (fr. 134) speaks of intellect darting through the kosmos, and almost certain in Democritus, Diogenes of Apollonia, and Philolaus at the end of the fifth century. Hence, Xenophon could say around 385 BCE that Socrates did not discuss the nature of "the kosmos as the wise call it." Similarly, Plato could have Socrates say that the wise call "kosmos" the whole of heaven, earth, gods, and man, as sharing community and friendship (*Gorgias* 507E–8A). A fundamental presupposition of the privileged use of "kosmos" is that the world is orderly and well arranged.

There is, however, a fundamental ambiguity even in the privileged use of "kosmos," which also reflects philosophical debates about the nature of the world. In the fourth century BCE, "the kosmos" can be used to refer to the entire world or just the system of stars, planets, sun, and moon. For Plato and Aristotle, the sublunary world is unorderly in comparison with the heaven and one task of the philosopher is to find the order it. The heaven is a better kosmos. But, depending on the interests of the text, the kosmos in some discussions might signify the entire world. Thus Aristotle can also speak of "the kosmos encompassing the earth," the region between the earth and the heaven (*Meteor*. A 2–8).

For Aristotle, as for most scientists until the sixteenth century, the world was spherical, consisting of three concentric layers, an outermost spherical shell for the fixed stars, then, contained within it, a spherical shell with the planets, sun, and moon and the apparatus by which they move (for Aristotle, an elaborate system of concentric spheres), and the sublunary sphere which has the earth as its center. Hence he distinguishes three senses of heaven (*De caelo* 9, 278b9–21), the limit of the periphery of the heaven or the spherical shell of the fixed stars (the first heaven), the spherical shell for the planets, sun, and moon are (the lower heavens), the sphere contained by the first heaven or the universe (all three layers). To these one may add the obvious first and lower heavens. It is plausible that in its privileged use, "the kosmos" could refer to any of these.

However, the universe need not be a kosmos, as is clear from ancient discussions of those philosophers who believed in many worlds, Democritus, Diogenes of Apollonia, and Epicurus. The many different systems of stars and earth are all "kosmoi," but neither the disordered universe composed out of all of them nor what lies in between them is itself a kosmos. So too, for the Stoics, the universe is not a kosmos since the universe includes not just the finite world sphere, but also an infinite void outside. Only what they call "the whole," the finite sphere encompassing the heaven and earth, is a kosmos (though within this "whole" there are three different arrangements they describe as a "kosmos": god or the divine moving principle; the ordering produced by this god; and the unity of the two).

In Greek mathematical astronomy, the kosmos is just a mathematical object, so that the connotations of orderliness are irrelevant. The primary goal of Greek astronomers from Eudoxus (fourth century BCE) on include mapping the heavens, determining the sizes and distances of all the bodies of the world, and constructing geometrical models that explain the apparent motions and phases of the heavenly bodies. With few exceptions, such as heliocentric theories (Aristarchus, c. 270 BCE), the kosmos will be a rotating sphere with the earth as center and whose poles determine the daily rotation of the stars.

See also Anaximander; Aristotle; Cosmology; Diogenes of Apollonia; Empedocles; Epicurus; Heraclitus of Ephesus; Leucippus and Democritus; Philolaus of Cro-

ton; Plato; Pre-Socratic Philosophy; Pythagoras and Pythagoreanism; Socrates; Stoicism; Xenophon.

Bibliography

Finkelberg, Aryeh, "On the History of the Greek Kosmos," *Harvard Studies in Classical Philology* 98 (1998): 103–136.

Kahn, Charles H. *Anaximander and the Origins of Greek Cosmology*. New York: Columbia University Press, 1960.

Kahn, Charles H. *The Art and Thought of Heraclitus*. Cambridge, U.K.: Cambridge University Press, 1979.

Kerschensteiner, Jula. *Kosmos; quellenkritische Untersuchungen zu den Vorsokratikern*. Munich: Beck, 1962.

Henry Mendell (2005)

COSTA, URIEL DA

(1585–1640)

Uriel da Costa, or Gabriel Acosta, an opponent of traditional religion, was born in Portugal to a New Christian family, that is, one forced to convert to Catholicism from Judaism. After completing studies at Coimbra, he held a minor church office. According to his autobiography, biblical studies led him back to Judaism, which he then expounded to his family as he deduced it from the Bible. The family fled to Amsterdam to escape the Inquisition and to practice their religion freely. Da Costa soon found that his biblical Judaism was in conflict with actual practices, which he claimed were too rigid and ritualistic. He attacked "the Pharisees of Amsterdam" and wrote a book arguing that the doctrine of the immortality of the soul was doubtful and unbiblical. The next year da Costa completed his *Examen dos tradiçoens Phariseas conferidas con a ley escrita* (Examination of the Traditions of the Pharisees Compared with the Written Law; 1624), a work considered so dangerous that the author was excommunicated by the Jews and arrested by the Dutch authorities as a public enemy of religion. He was fined, and the book was publicly burned. (Its contents can be reconstructed from a reply by Samuel da Silva.) In 1633 he sought readmission to the Jewish community. Though he had not changed his views, he needed the communal life, and so, he said, he would "become an ape among apes," and submit to the synagogue. However, he soon found himself doubting whether the Mosaic law was really God's law, and asking whether all religions were not human creations. He transgressed all sorts of Jewish regulations and observances, and finally was condemned for discouraging two Christians from becoming Jews. He was again excommunicated. In 1640 he submitted once more and underwent the most severe penance, first recanting before the whole synagogue, then receiving thirty-nine lashes, and finally lying prostrate while the congregation walked over him. He then went home, wrote his autobiography (*Exemplar Humanae Vitae*), and shot himself.

Da Costa's tragic career has made him a symbol of the dangers of religious intolerance, as well as a precursor of modern naturalism and higher criticism. One romantic painting shows him as a kindly scholar, holding young Benedict de Spinoza on his knee, teaching him.

Almost all our information about da Costa comes from his autobiography, published in 1687 from a Latin manuscript. It is not known if it is the original text or an altered version. Very little other data have turned up concerning his actual relations with Amsterdam Jewry or Spinoza. I. S. Révah's 1962 study, based on Portuguese Inquisition records, indicates that da Costa's initial conversion was not, in fact, from Catholicism to biblical Judaism, but rather to a peculiar Iberian form of crypto-Judaism. Then, Révah suggests, in Amsterdam da Costa developed first a biblical Judaism, and later a variety of deism or natural religion.

Da Costa's influence, from the eighteenth century onward, has been mainly on religious liberals opposing traditional orthodoxies. It is his martyrdom, rather than his doctrines (which we hardly know), that has affected people. Considering the many intellectuals gruesomely killed by Protestants and Catholics, it is odd that da Costa has stood out as *the* example of a freethinker destroyed by religious bigotry. Possibly Enlightenment and romantic thinkers could better accept a hero victimized by Judaism than one victimized by their own previous Christian traditions.

Bibliography

WORKS BY DA COSTA

Die Schriften des Uriel da Costa. With introduction, translation, and index, edited by Carl Gebhardt. Amsterdam: Hertzberger, 1922. Bibliotheca Spinozana Tomus II.

Une vie humaine. Translated, with a study of the author, by A.-B. Duff and Pierre Kann. Paris, 1926.

Examination of Pharisaic Traditions. Translated by H. Salomon and I. Sassoon. Leiden: Brill, 1993.

WORKS ON DA COSTA

"Acosta, Uriel." In *Jewish Encyclopedia,* Vol. I, 167–168.

Bayle, Pierre. "Acosta." In *Dictionnaire philosophique et critique.* Rotterdam, 1695–1697.

Israel, J., and D. Katz, eds. *Sceptics, Millenarians and Jews.* Leiden: Brill, 1990.

Popkin, R. "The Jewish Community of Amsterdam." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman, 600–611. London: Routledge, 1997.

Révah, I. S. "La Religion d'Uriel da Costa." *Revue de l'histoire des religions* 161 (1962): 45–76.

Révah, I. S. *Uriel da Costa et Les Marranes de Porto.* Paris: Centre Culturel Calouste Gulbenkian, 2004.

Salomon, H. "A Copy of Uriel da Costa's *Examen des tradicoes phariseas* Located in the Royal Library of Copenhagen." *Studia Rosenthaliana* 24 (1990): 153–168.

Richard H. Popkin (1967)
Bibliography updated by Oliver Leaman (2005)

COUNTERFACTUALS

See Appendix, Vol. 10

COUNTERFACTUALS IN SCIENCE

The term *counterfactual* is short for "counter-to-fact conditional," a statement about what would have been true, had certain facts been different—for example, "Had the specimen been heated, it would have melted." On the face of it, claims about what would or could have happened appear speculative or even scientifically suspect because science is an investigation of reality grounded in experimental evidence, and by definition people have experimental access only to the actual universe. Yet, despite their implicit reference to alternative possibilities, many counterfactuals are scientifically respectable because the criteria determining whether they are true depend wholly on facts about the actual universe. Counterfactuals are often important in science because they appear implicitly in the definitions of certain specific concepts such as "solubility" and "biological fitness," and because they are closely related to general scientific notions such as "law of nature" and "causation."

The exact definition of *counterfactual* is controversial. In philosophy, a counterfactual is a statement that can be paraphrased in the form, "If *A* were true, then *C* would be true." They are distinguished from indicative conditionals, which take the form, "If *A* is true, then *C* is true." The difference in meaning consists roughly in the kind of facts one keeps fixed when considering the hypothetical situation *A*. To evaluate "If Napoleon Bonaparte had been born in Spain, France would have been ruled by democrats," one imagines Napoleon for some reason being born in Spain instead of Corsica and then speculates about alternative histories for France, ignoring what is known about the specifics of Napoleon's actual reign of power. But when evaluating the indicative, "If Napoleon was born in Spain, France was ruled by democrats," one can imagine that somehow historians have made a mistake on this one issue of Napoleon's birth and retain other things known about Napoleon, such as his undemocratic rule of France.

Despite the clear difference in meaning between these two particular sentences, there is significant controversy about whether the distinction between indicatives and counterfactuals makes sense in general and whether it is the best way to categorize conditionals. Associated with such debates are subtleties regarding how truth applies to counterfactuals. For example, the name "counterfactual" is misleading in that one can use counterfactuals for situations that are known to be true. Believing "If the fish had mutated, it would have survived," is consistent with believing the fish did mutate. So, counterfactuals are not only about counter-to-fact possibilities, but sometimes about actual situations as well.

RELATION TO LAWS

What makes counterfactuals especially suitable for science is that the truth of counterfactuals depends largely on the general patterns that science aims to describe. One can reasonably say that a particular sample of salt is soluble in water even when the salt has never been dissolved and never will, on the grounds that because of its chemical structure, had it been placed in a sufficient amount of pure water, it would have dissolved. One is justified in making claims about what the salt counterfactually would have done in virtue of what other similar samples of salt have actually done and that person's knowledge of nature's regularities. In this way, the laws of nature can be understood as governing not only actual happenings but also what may have happened.

In one early philosophical treatment, Nelson Goodman tried to explain counterfactuals as a kind of elliptical expression. He thought counterfactuals such as, "Had I struck this match, it would have lit," should be understood as, "I struck this match and the laws of nature are true and … logically entails that the match lit," where the ellipses represent some unstated but true facts. In a typical situation one may complete the sentence as, "I struck this match, and the laws of nature are true, and the match was dry, and there was sufficient oxygen in the air, and the match had the proper chemicals in the tip, entails that the match lit." The value of Goodman's account is that it captures the idea that counterfactuals in science express consequences of actual or hypothetical facts following from the laws of nature.

A major problem with this account, as Goodman himself recognized, is that it fails to give any constructive advice about how to pick out the right facts to insert into the ellipses. Why should one insert the fact that the match was dry and infer that the match would have lit, rather than insert that the match did not light and infer that the match would have been wet? Because there is no principled way of answering this question, Goodman's theory is of limited value as a guide to determining the truth of counterfactual statements. Also, because many counterfactuals have nothing to do with laws ("If the circumference were only half as large, the radius would be …") and some require the actual laws to be abandoned ("If there were no friction …"), the elliptical account is not a general account of counterfactuals, and it was not immediately obvious how it would fit into a larger account.

SIMILARITY APPROACHES

The dominant approach to elucidating the meaning of counterfactuals is to think of them as having truth conditions given by similarity relations among possible worlds—that is, hypothetical universes—and more controversially that some more or less tractable notion of similarity tells how to evaluate specific counterfactuals. The justification for this is primarily formal. Robert Stalnaker and David Lewis developed a compelling family of logic systems describing counterfactual conditionals that do a remarkable job of justifying a wide range of intuitively plausible reasoning patterns. It is a feature of the logic that it can be interpreted using a notion of similarity among possible worlds. The way it works roughly is that the counterfactual "If *A* were the case, then *C* would be the case," is true when the worlds most similar to actuality among those where *A* is true are also worlds where *C* is true. Consider, "If this bird had three legs, it would have more legs than wings." The worlds where *A* is true are all the worlds where the bird has three legs, including worlds where it has three legs and three wings, worlds where it has three legs and four wings, and worlds where it has three legs, two wings, transparent feathers and a metallic beak. Intuitively, the minimal departure from actuality is for it to have one extra leg without any change to its wings, and so using common sense, one would say this counterfactual is true.

This illustration of how to determine whether a counterfactual is true involves an appeal to one's offhand, pretheoretical judgements of similarity, an appeal not mandated by the role similarity plays in the formal logic. It is a significant speculative leap to suppose that which counterfactuals are true depends on what human beings find similar. Nevertheless, inspired by David Lewis's work, there has been a serious philosophical research program dedicated to finding a plausible refinement of people's ordinary similarity concept to justify the usage of counterfactuals and more important to use counterfactuals in elucidating other concepts, such as causation.

CAUSATION

A large part of science is figuring out what causes what. The role of counterfactuals in this project is to express dependencies among logically independent elements of reality, dependencies that are often causal. In the vast literature on causation, counterfactuals appear in different roles, not all of them central. One tradition concerning causation is to take causal connections between facts or events to be primitive elements of reality holding together the pattern of various particular facts. In this tradition, counterfactuals are not crucial to the formulation or definition of causation, although they are useful for expressing consequences of causal relations.

Where counterfactuals become most important are in theories where causes are understood as the byproduct of physical processes that are themselves not fundamentally causal in nature. This tradition is compelling because fundamental physics uses equations establishing mathematical relationships between physical entities in a way that does not obviously indicate what causes what. Here a theory about counterfactual relationships between events can be constructed as part of a story that tells how the mathematical relations in physics could possibly account for truths such as "Lightning causes thunder."

Some theories of causation are literally counterfactual accounts of causation. They argue that the causal connection is really due to a counterfactual dependence relation. An event *E* counterfactually depends on the event *C* whenever if *C* had not happened, *E* would not have happened. In one famous version of the theory—by David Lewis—causation is identified with having a chain of events that are counterfactually dependent on one another, but other variations on the connection between causation and counterfactual dependence are possible. While counterfactual accounts need to successfully explain many aspects of causation, for them to be even superficially plausible, they need to explain the causal asymmetry—why in ordinary circumstances causes precede their effects. In counterfactual accounts, the difficult part of that explanation is to say why in an ordinary case of causation such as lightning causing thunder, one does not also have the lightning counterfactually depending

on the thunder, wrongly entailing that thunder causes lightning.

Explaining why thunder does not cause lightning is difficult if one followsthe orthodoxy of using anthropocentric ideas of similarity as a guide to counterfactual truth. It is plausible that lightning counterfactually does depend on thunder because a possible world with a bolt of perfectly silent lightning is intuitively stranger than a world with just one less bolt of lightning. This shortcoming for the counterfactual account of causation may be corrected by giving up on using naive judgements of similarity and instead concocting a suitable theory of similarity that fits the needs of causation. David Lewis's theory in "Counterfactual Dependence and Time's Arrow" (1979) has been a popular model for developing such an account. In following this strategy, the attempt is to defend the more general hypothesis that counterfactuals ordinarily exhibit a temporal asymmetry that in turn explains the difference between cause and effect.

COUNTERFACTUAL ASYMMETRY

When one considers how things may have been had only *X* not happened, typically one envisions alternate histories with an identical past where for some reason *X* did not happen. one then speculates how these alternate histories may have played out, leaving the future as open as the laws and circumstances allow. The practice of evaluating counterfactuals this way is asymmetric, treating the future but not the past as counterfactually dependent on the present.

Because there are counterfactuals having nothing to do with time, such as "If the variable x had been equal to three, then $x+1$ would have been equal to four," it is known that time asymmetry is not a part of the logic or meaning of counterfactuals per se. In a sense, it is wholly up to a person to choose whether he or she evaluate a given counterfactual symmetrically or asymmetrically. Nevertheless, it is an objective fact that nature tends to reward people for using the asymmetric ones. For example, it is sometimes useful to think, "If I were to shield myself now, I would avoid the next volley of arrows," and not so useful to think, "If I were to shield myself now, I would have avoided the previous volley." In this sense, counterfactual asymmetry is a natural fact perhaps amenable to scientific explanation. The project is to determine which physical structures vindicate the practice of evaluating counterfactuals asymmetrically. This includes determining to what extent the asymmetry is an aspect of people's particular perspective on nature, and to what extent the asymmetry is a feature of broader physical conditions and laws.

One idea is that there are fundamentally random processes that make the future chancy in a way that the past is not. This is problematic because although chances seem to imply a sense of openness for the future, it is not clear how chances imply a fixed past. Perhaps the intuition about chance in this case presupposes a theory where the past is given special fundamental significance as being in some sense more real than the future, or real in a different way. Spelling out such a deep metaphysical difference between past and future has proven difficult in itself, and clarification of its connection to chances has been problematic because the application of chances in science does not seem to require any such distinction.

Another group of proposed candidates for the explanation do not take counterfactual asymmetry to be a fundamental fact about reality or time itself, but as a contingent feature of the particular environment. A suggestion by Lewis is that typical processes exhibit a pattern where future facts "overdetermine" past facts in the sense that they give redundant evidence of the past. For example, after an explosion, there are many fragments around, each of which individually suggests an explosion, but there are often only a few facts beforehand that imply an explosion will occur—for example, a burning fuse.

Another idea is that counterfactual asymmetry is explained by cosmological facts, such as the universe is expanding from a smooth distribution of matter just after the big bang. This idea draws some plausibility from nature's two classes of time asymmetries. The first kind is a local asymmetry—a fact that applies directly to the physical process taking place. Examples of this first kind include chancy transitions in the physical state, and time-asymmetric evolutions such as one sees in certain high energy particle experiments involving weak decays. The second kind is an asymmetry in the boundary conditions. Irreversible phenomena such as mixing gasses or a hot and cool object settling to a single temperature are explained only when one posits special boundary conditions. Specifically, the explanation for why people regularly see thermodynamic asymmetries comes by way of the physics of the distant past being constrained in a way that the future is not. The connection to counterfactuals is that one's reason for thinking that causation is asymmetric comes from one's experience with asymmetric macroscopic phenomena, exactly the kind of phenomena whose asymmetries are explained by boundary conditions and not by local asymmetries. Hence, it is plausible to think that special facts about the beginning of the uni-

verse are a critical component of why counterfactuals and causation treat the past as more fixed—that is, why nature rewards people for evaluating counterfactuals in a way that treats the past as fixed.

These strategies that attempt to explain counterfactual asymmetry by way of contingent physical circumstances are interesting in that they allow for at least some counterfactual dependence of the past on the future. This seems reasonable because one wants to allow for counterfactual differences that arise from ordinary processes. If the population were greater right now than it actually is, this would have been because people would have had more children, not because people would have magically popped into existence. While having some counterfactual dependence of the past on the present is good for matching up the theory of counterfactual evaluation with pretheoretical intuitions about counterfactuals, it highlights a difficulty with the desired uses of counterfactual asymmetry. If the past counterfactually depends on the present, and the difference between cause and effect is purely given by the counterfactual asymmetry, then one would seem to have backwards causation, such as thunder causing lightning. So either counterfactual asymmetry can't do the job of grounding the cause-effect asymmetry or causation is a less robust notion than is ordinarily thought, defined with respect to a temporal asymmetry that at best is justified only in certain special cases—for example, human decisions—where there is little or no significant backwards dependence.

See also Causation: Philosophy of Science; Conditionals; Counterfactuals; Philosophy of Statistical Mechanics.

Bibliography

Goodman, Nelson. "The Problem of Counterfactual Conditionals." *The Journal of Philosophy* 44 (1947): 113–128.

Lewis, David. "Causation," *The Journal of Philosophy* 70 (1973): 556–567.

Lewis, David. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (1979): 455–476.

Lewis, David. *Counterfactuals*. Oxford: Blackwell, 1973.

Stalnaker, Robert. "A Theory of Conditionals." In *Studies in Logical Theory: American Philosophical Quarterly Monograph Series*, no. 2, edited by Nicholas Rescher. Oxford: Basil Blackwell, 1968.

Douglas Kutach (2005)

COUNTERPART THEORY

See *Modality, Philosophy and Metaphysics of*

COURNOT, ANTOINE AUGUSTIN

(1801–1877)

Antoine Augustin Cournot, the French mathematician, economist, philosopher, and educator was born in Gray, Haute Saône. He was educated at *collèges* at Gray (which now bears his name) and Besançon, and at the École Normale Supérieur in Paris. In addition to teaching at the universities of Lyon and Paris, he was head of the *Académie* at Grenoble and rector of the *Académie* at Dijon and succeeded André Marie Ampère as inspector general of studies. An able student of mechanics (including astronomy) and of mathematics, he applied probability theory to problems in both the physical and the social sciences. His work in economics early secured his reputation in that field, and he is now generally regarded as a founder of econometrics; as a philosopher he remains much less known.

Cournot is identified by Jean de la Harpe as a critical realist. This designation would be peculiarly appropriate were it not for the fact that this name has been taken by a group of American philosophers whose position is notably unlike that of Cournot in important respects. Since the term *critical realist* is equivocal, it may be advisable to refer to Cournot as a critical rationalist. Cournot is a realist of sorts in his metaphysics and more rationalist (albeit critically so) than empiricist or positivist in his epistemology. For him knowledge is a function of reason. The senses furnish neither the basis nor the criteria of knowledge, which not only can but does extend beyond their limits. Yet the senses do make important contributions to knowledge, especially by restraining its claims by challenging overextended speculations by confronting them with what William James aptly called "brute facts." Cournot rejects all dogmatic philosophies, whether rationalist or empiricist. Knowledge requires a continuing appraisal of all principles to determine both their grounds and the range of their legitimate applications. Specifically, he examines the established sciences to see whether they have any basic concepts in common. He discovers three such concepts—order, chance, probability. These three concepts lie at the heart of Cournot's philosophy and suffice to account for his rejection of many earlier and contemporary alternative positions. He rejects the idealistic basis and implications of Immanuel Kant's philosophy, but he accepts the critical intent of the Kantian program.

For Cournot, order is a basic category which, as "objective reason," relates to the nature of things and, as

"subjective reason," to the means through which we apprehend that nature. The major function of philosophy is to examine and criticize the efforts of subjective reason to know objective reason, making sure among other things that such closely related and often confused principles as "reason" and "cause," "rational order" and "logical order," are clearly differentiated in both their meaning and their function. We have knowledge when we apprehend the objective reason of things, but such knowledge is rarely complete and certain. Therefore, our knowledge is relative and probable, not absolute and apodictic, but it nonetheless rests on objective grounds, not on forms or categories native to the mind itself.

Cournot's unusual and cogent use of probability draws attention to a fundamental moderating element in his philosophy. His treatment of probability is developed most extensively in his *Exposition* (1843) and is used ingeniously and productively in his *Essai* (1851), *Traité* (1861), and *Matérialisme, vitalisme, rationalisme.* Long before putting these views to philosophic use, Cournot had applied them to problems in astronomy and in various fields of social studies, notably in economics, where he applied them with lastingly important results.

The calculus of probabilities is related to both order and chance. Both order and probability have plural meanings. Order as a category of the objective reason of things must not be confused either with logical order—that is, with the order essential to a formal system of ideas—or with causal order, by which Cournot means essentially what Aristotle called "efficient cause." The reason for a phenomenon must be distinguished from its cause, from the conditions or circumstances which give rise to it. *Cause* is related to the particular and unique; *reason* is related to the universal and abstract aspects of phenomena that are the ground for laws of general and fundamental relations among them, relations that are necessary, but not in themselves sufficient, conditions for the production of specific phenomena. Probability is of two sorts, mathematical and philosophical. Mathematical probability applies to those relatively rare situations in which the number and relative frequency of various possibilities can be numerically determined. Philosophical probability—which may attain practical, but never demonstrable, certainty—applies to the vastly more numerous cases in which such numerical determination is not possible. It involves an appraisal of evidence in terms of rational cogency where probabilities persuade and win the acquiescence of reasonable persons even though the relevant evidence is neither quantifiably manipulatable nor conclusive. We live continuously and inescapably with such probabilities; philosophical criticism is also largely concerned with them. In either case probability is a function of objective factors and conditions and not solely of our ignorance or other subjective factors, although these do contribute to our need to deal with probabilities of both types.

Of Cournot's three basic ideas, that of chance is least adequately developed. It is unfortunate that there is no specific and clear definition of this concept in its theoretical function, yet what the concept refers to is not at all unclear. Numerous examples leave no doubt about the meaning of the term as Cournot uses it. A chance occurrence is one in which there is an unpredictable conjunction of independent series of events, each series being internally related and having a determinable nature. However complete our knowledge of each independent series, events resulting from unpredictable conjunctions among them are contingent, unpredictable, fortuitous. Such events have causes, but they are not reducible to laws. The absence of reasons for such events is irreducible, chance, like order, being an objective feature of the nature of things. This doctrine is one source of the pluralism in Cournot's philosophy. In it he anticipates Émile Boutroux and suggests certain aspects of the philosophies of C. S. Peirce (for example, his "tychism") and M. R. Cohen (whose general philosophical position is not unlike Cournot's critical rationalism).

Another pluralistic aspect of Cournot's thought is indicated in the title of his last philosophic work, *Matérialisme, vitalism, rationalisme.* Countering the principles of Darwinian evolution, Cournot holds to the principle that living beings are distinguished from nonliving things by a unity and form suggestive of finality and by a vital principle inexplicable in physical and chemical terms. Here Cournot anticipates both Henri Bergson and the emergent evolutionists, notably Samuel Alexander and C. Lloyd Morgan.

In his consideration of such concepts as form, unity, simplicity, and symmetry, Cournot moves toward a transrationalism—that is, toward a view in which ideas that go beyond normal rational analysis and use, such as finality, purpose, and God, find a place. This development is consistent with, indeed perhaps it is a consequence of, his pluralism and his implied doctrine of levels and with his rejection of any reductionist view as these are evidenced by his assertion that the phenomena of life involve something not present in nonliving phenomena. Such ideas as simplicity and symmetry are relevant to rational investigation, to the discovery of the order and reason of things, as in the probabilistic assessment and choice between

otherwise equally adequate alternative hypotheses. In this sense such concepts are regulative ideas of reason. But Cournot argues that they are more than this, and in his treatment of these concepts he moves from a logic of reason toward an aesthetic of reason, in which the concept of order has a connotation more extensive than reason can explore. What effect does such a transrationalism have on the claimed objective existence of chance, the second concept so fundamental to Cournot's philosophy as a whole? None. Why this is the case is not adequately developed in Cournot's works, although a hint is found in *Exposition*: God lays out the laws or rational elements of reality and leaves to objective and inexpugnable chance the details of fortuitous occurrences. Therefore, even such a superior intelligence would, like man, be unable to foresee contingent events, although unlike man its assessment of what is contingent would not be complicated by subjective factors of the sort which inescapably limit and affect human judgment.

In developing his philosophy, Cournot deals with the nature of language, ethics, and aesthetics and with various social institutions and factors which contribute to civilization. He also discusses the nature of science, history, and philosophy and considers at some length the irreducible distinctions between them. His *Considérations* is a peculiarly interesting account of his handling of various historical matters.

See also Alexander, Samuel; Ampère, André Marie; Aristotle; Cohen, Morris Raphael; Critical Realism; Kant, Immanuel; Morgan, C. Lloyd; Peirce, Charles Sanders.

Bibliography

WORKS BY COURNOT

Mémoire sur le mouvement d'un corps rigide soutenu par un plan fixe. Paris, 1829.

Recherches sur les principes mathématiques de la théorie des richesses. Paris: Hachette, 1838. Translated into English by N. I. Bacon as *Researches into the Mathematical Principles of the Theory of Wealth.* Economic Classics series. London: Macmillan, 1897.

Traité élémentaire de la théorie des fonctions et du calcul infinitesimal. 2 vols. Paris: Hachette, 1841; 2nd ed., revised and corrected, 1857.

Exposition de la théorie des chances et des probabilités. Paris, 1843.

De l'origine et des limites de la correspondance entre l'algèbre et la géometrie. Paris, 1847.

Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique. 2 vols. Paris: Hachette, 1851; 2nd ed., 1912; 3rd ed., 1 vol., 1922. Translated into English by Merritt H. Moore as *An Essay on the Foundations of Our Knowledge.* New York: Liberal Arts Press, 1956.

Traité de l'enchaînement des idées fondamentales dans les sciences et dans l'histoire, 2 vols. Paris: Hachette, 1861; 2nd ed., 1 vol., 1911; 3rd ed., 1 vol., 1922. The 1911 and 1922 editions include a foreword by L. Lévy-Bruhl.

Principes de la théorie des richesses. Paris: Hachette, 1863.

Des institutions d'instruction publique en France. Paris: Hachette, 1864.

Considérations sur la marche des idées et des événements dans les temps modernes. 2 vols. Paris: n.p., 1872; republished with introduction by M. Mentré, 2 vols., Paris: Boivin, 1934.

Matérialisme, vitalisme, rationalisme: Études sur l'emploi des données de la science en philosophie. Paris, 1875; Paris: Hachette, 1923.

Revue sommaire des doctrines économiques. Paris: Hachette, 1877.

Souvenirs: 1760 à 1860. Paris: Hachette, 1913. Written in 1859; published with introduction and notes by E. P. Bottinelli.

WORKS ON COURNOT

Bottinelli, E. P. *A. Cournot métaphysicien de la connaissance.* Paris: Hachette, 1913.

Darbon, A. *Le concept du hasard dans la philosophie de Cournot.* Paris: Alcan, 1911.

Harpe, Jean de la. *De l'ordre et du hasard: Le Réalisme critique D'Antoine Augustin Cournot.* Memoirs of the University of Neuchâtel, Vol. IX. Neuchâtel: Secrétariat de l'Université, 1936.

Lévèque, R. *L' "élément historique" dans la connaissance humaine d'après Cournot.* Publications of the Faculty of Letters of the University of Strasbourg, No. 82. Paris: La Société d'édition Les Belles lettres, 1938.

Mentré, F. *Cournot et la renaissance du probabilisme au XIX[e] siècle.* Paris: Rivière, 1908.

Milhaud, G. *Études sur Cournot.* Paris: J. Vrin, 1927.

Revue de métaphysique et de morale 13 (3) (1905): 293–543. A special number dedicated to A. Cournot, with portrait. Articles by H. Poincaré, G. Milhaud, G. Tarde, C. Bouglé, A. Anpetit, F. Faure, A. Darlu, F. Vial, D. Parodi, R. Audierne, H.-L. Moore.

Ruyer, R. *L'humanité de l'avenir d'après Cournot.* Paris: F. Alcan, 1930.

Segond, J. *Cournot et la psychologie vitaliste.* Paris: Alcan, 1911.

Merritt Hadden Moore (1967)

COUSIN, VICTOR

(1792–1867)

Victor Cousin, the French philosopher and historian, was born in Paris and educated at the Lycée Charlemagne and the École Normale, where he studied under Pierre Laromiguière. He began his teaching career in 1815, assisting Pierre Paul Royer-Collard in his course on the history of philosophy at the University of Paris. Cousin studied German and read Immanuel Kant and F. H. Jacobi; but he was especially attracted to the works of Friedrich Wilhelm Joseph von Schelling, whose thought

had a permanent influence upon him. A trip to Germany in 1817 brought him into personal contact with both Schelling and G. W. F. Hegel, a fact which was later responsible for the accusation that he had rejected French philosophy in favor of Germany's. In 1821 Cousin was removed from his position because of his supposed antigovernmental views, and he used his freedom to make another trip to Germany. While there he was imprisoned, on charges that have never been entirely clear, but was freed after six months. Returning to France, he spent his time writing his philosophical and historical works and editing the works of other philosophers, including Proclus (6 vols., 1820–1827) and René Descartes (1826, 11 vols.), and beginning his translation of Plato (13 vols., 1822–1840). In 1828 he was restored to his post and from then on had an influential career as lecturer. He became a spokesman for the *juste milieu,* as he called it, which in philosophy meant eclecticism. Cousin's power increased when in 1840 he became minister of public instruction, director of the École Normale, and a member of the Institut de France. He was not only the most famous French philosopher of his time but also supreme dictator of who should teach philosophy and what should be taught. He had become, moreover, a power in the whole educational system of France when he published a report on Prussian education (1833). (This report was later translated into English in 1834 and distributed to the schools of Massachusetts by an act of the legislature.) At the advent to power of Louis Napoleon in 1848 Cousin retired from active teaching and spent his time in literary studies.

ECLECTICISM

Though Cousin started his career as a pupil of Laromiguière, it was the commonsense philosophy of Thomas Reid, as interpreted by Royer-Collard, that was the source of his own doctrine. To Cousin common sense was a fusion of the best that had been done in philosophy, combining the empiricism of sensationalism in epistemology with the spiritualism of religion. The epistemology of Étienne Bonnot de Condillac and his school, Cousin felt, because it made the spirit of man a simple passive victim of external forces, had led them to atheism and materialism, both of which were to be condemned. Atheism and materialism could not give men those permanent principles that would guide their moral life. Such principles were to be found only if men realized that their minds were active as well as passive, their activity consisting in their use of their a priori categories of substance and causality.

Though it is likely that Cousin got the idea of the complementary active and passive aspects of mentality from Schelling, he himself attributed it to Maine de Biran's self-scrutiny. This gave him a French origin for doctrines which were to guide French professors. Maine de Biran's active will, Cousin maintained, was balanced by sensibility, which "implies" the existence of an external world. Sensibility and active will were accompanied by reason, and thus Cousin revived the traditional threefold analysis of the mind. Corresponding to the three faculties was a threefold division of philosophical problems into that of the good, the beautiful, and the true. In his book *Du vrai, du beau et du bien* (1853) Cousin argued that these problems were united in a whole which absorbed what was valid in sensation (John Locke), reason (Plato), and the heart (for which he named no sponsor). These three parts of the soul are not independent of one another. Reason requires both sensation and the heart, sensation requires reason and the heart, and the heart requires both reason and sensation. By analogy epistemology, ethics, and aesthetics are all intertwined and inseparable except for purposes of exposition.

POLITICAL PHILOSOPHY

The political philosophy of Cousin was expressed in *Justice et charité,* a brief tract that he wrote as one of a series published by members of the Académie des Sciences Morales et Politiques in 1848. This tract is based on the same metaphor of the interdependence of separate things. The purpose of all the tracts in this series was to substantiate the right to property, the well-being of family life, popular freedom, and progress. Cousin opposed the idea of equality, the right to work, and governmental aid. Justice is the protection of natural rights, but every right implies a complementary duty. Men are all free, but their freedom resides only in the search for truth, in religious beliefs and practices, and in property. Justice demands that these rights be respected and protected by the state. On the other hand, charity demands that we abuse none of these rights, that we individually seek the truth and not perpetuate error, that we give others the religious freedom that we demand for ourselves, that we respect others' property as we would have them respect ours. In short, law is futile if it is not obeyed, and we cannot obey a law that is not enforced. Respect for the law is like charity in that it has no limits; for charity extends to all men and to liberty in all its forms.

AESTHETICS

Cousin was a strong believer in absolute beauty. His ideal work of art was the Apollo Belvedere. Art, he believed, is neither an imitation of nature (sensationalism) nor edification (moralism), but rather a vision of "the infinite." Though all arts utilize matter, they communicate to it "a mysterious character which speaks to the imagination and to the soul, liberates them from the real, and bears them aloft either gently or violently to unknown regions." These unknown regions are the country of God, the world of the ideal. Though this passage might seem to ally Cousin with the Romantic school, in fact it led him to give highest praise to the classicists of the seventeenth century. He was clearly under the influence of J. J. Winckelmann, who also admired the Apollo Belvedere as the *summum* of all ideal beauty and believed that all praiseworthy artists put into their works of art the ideal beauty of Plotinus. Cousin saw in beauty, as did Hegel, a sensuous manifestation of the Absolute, though he expressed it in different language.

At the same time Cousin admitted that one must not exaggerate the idealism of a work of art. All works of art speak to the senses as well as to the heart. The ideal must be presented to us in sensible form and it must also be agreeable to our feelings. A work of art that is beautiful was for Cousin a concrete presentation of the unity he found in eclecticism. Consequently, art that did not realize the potentialities of the sensuous, the rational, and the sentimental would not be of as high a rank as art that did. The conclusion was that poetry was the highest of all the arts. Its power of words is such that it can stimulate images, feelings (affections), and thoughts at one and the same time. It is thus a synthesis of all human powers.

HISTORIOGRAPHY

The pioneering editorial work of Cousin, mentioned above, made accessible to the public manuscripts that had been previously hidden in libraries. His eclecticism served him well in this field, for with the exception of the sensationalists of the eighteenth century, there were few philosophers of the past in whom he could not find some truth.

Cousin's *Philosophie sensualiste au XVIII^e siècle* (1819), a course of lectures, is the most biased of his historical studies, but still treats of Locke, Condillac, Claude-Adrien Helvétius, Saint-Lambert, and Thomas Hobbes in an interesting manner. His criticism of Locke, that Locke was unable by the very nature of his epistemology to account for universal and necessary ideas, Cousin's analysis of Condillac's notion that deduction is always tautological, and even Cousin's attacks on Helvétius are carefully based on the texts and are far from superficial. Fundamentally his objection to these thinkers was the pragmatic, moral, and religious consequences of their premises, an objection which obviously sprang from his own moral and religious convictions. His *Cours de l'histoire de la philosophie* (1829) was considered a work serious enough to be analyzed and commented upon by Sir William Hamilton in the *Edinburgh Review,* and, indeed, its exposition of the technique of historiography was thorough and based on a perception of genuine historical problems.

Cousin made the mistake of dividing all possible philosophies into four kinds—sensualism, idealism, skepticism, and mysticism—and thus helped to influence his successors in this area toward thinking of philosophies as always productive of systems. This division led Cousin to look for a unitary idea pervading each system, though the idea in question might be a simple metaphor or a theory of the origin of ideas which exfoliated into an ethics, aesthetics, theology, or other theoretical construct. Like Hegel, Cousin was given to envisioning philosophical systems as "expressive" of ages and peoples, as if an age or a people were homogeneous. Yet at the same time he admitted the heterogeneity of what he called populations as distinguished from peoples, the latter being unified in their beliefs and outlooks on the world's problems, the former being diversified or, as he would put it, not yet unified. Where there was diversity, there was nevertheless a predominant idea in every epoch, but alongside of it existed other ideas "playing a secondary but real role."

Each people, Cousin maintained, was given, presumably by God or by the inevitable course of history, an idea to represent, and its history was the realization of this idea. This idea expresses itself in all human concerns—in philosophy, religion, science, art, and morals. It is almost certain that Hegel was the source of this theory, though Cousin made no mention of his influence. He was willing, however, to give great credit to J. J. Brucker, Dietrich Tiedemann, and W. G. Tennemann; these last two, he believed, expressed a history of philosophy associated with Locke and Kant, respectively. As for the nineteenth century, Cousin held that it would not have its own history of philosophy until it had a representative philosophy. That philosophy would be a union of the two traditions referred to by Cousin as the nucleus of a "vast and powerful eclecticism."

It is customary to treat Cousin with patronizing disdain, and it is true that he was always ready to compromise with political power and adjust his conclusions and,

indeed, his methods of research to what he believed to be expedient. He succeeded in excluding from his "regiment," as he called it, philosophers whose views were not harmonious with his own. Thus neither Auguste Comte nor J. G. F. Ravaisson-Mollien nor Charles Renouvier, to cite but three names, were able to become members of the teaching staff of the University of Paris. On the other hand, Cousin did stimulate research into the classics of philosophy, and his very chauvinism turned men's attention to such neglected figures as Maine de Biran. His eclecticism was not real, for he rejected any philosophy whose supposed religious and ethical effects he thought were undesirable. Yet his notion that every philosophy contained some truth induced his pupils to look into them all and gave them a catholicity of interest that was unusual and almost unique.

See also Absolute, The; Art, Interpretation of; Comte, Auguste; Condillac, Étienne Bonnot de; Descartes, René; Hamilton, William; Hegel, Georg Wilhelm Friedrich; Helvétius, Claude-Adrien; Hobbes, Thomas; Idealism; Jacobi, Friedrich Heinrich; Kant, Immanuel; Laromiguière, Pierre; Locke, John; Maine de Biran; Mysticism, Nature and Assessment of; Plato; Ravaisson-Mollien, Jean Gaspard Félix; Reid, Thomas; Renouvier, Charles Bernard; Royer-Collard, Pierre Paul; Schelling, Friedrich Wilhelm Joseph von; Winckelmann, Johann Joachim.

Bibliography

With the omission of his editions of classical authors and literary studies, Cousin's works include *Fragments philosophiques* (Paris: A Sautelet, 1826); *Cours de l'histoire de la philosophie,* 3 vols. (Paris: Pichon et Didier, 1829), which together with *Cours de l'histoire de la philosophie moderne,* 5 vols. (Paris: Ladrange, 1841) contains many titles also published separately; *Justice et charité* (Paris: Pagnerre, 1848); and *Du Vrai, du beau et du bien* (Paris: Didier, 1853).

For literature on Cousin, see P. F. Dubois, *Cousin, Jouffroy, Damiron, souvenirs publiés avec un introduction par Adolphe Lair* (Paris: Perrin and Cie, 1902); Paul Janet, *Victor Cousin et son oeuvre* (Paris: Calmann Lévy, 1885); and Jules Simon, *Victor Cousin* (Paris: Hachette et Sie, 1887), which has been translated under the same title by M. B. Anderson and E. P. Anderson (Chicago: A.C. McClurg, 1888). For a generally hostile approach, from a positivistic point of view, see Lucien Lévy-Bruhl, *History of Modern Philosophy in France,* translated by G. Coblence (Chicago and London: Open Court, 1924), Ch. 12; for a favorable account see John Veitch and X, "Cousin, Victor," in *Encyclopaedia Britannica,* 11th ed. (Chicago, 1910).

George Boas (1967)

COUTURAT, LOUIS
(1868–1914)

Louis Couturat, the French philosopher and logician, studied at the École Normale Supérieure and earned an *agrégé* in philosophy and a licentiate in mathematics. He taught philosophy at the universities of Toulouse and Caen but soon gave up teaching in order to devote all of his time to his own researches.

Couturat first attracted attention with his important doctoral thesis, *L'infini mathématique* (Paris, 1896). At a time when the mathematicians were still questioning the validity of Georg Cantor's theories and when the majority of French philosophers, led by Charles Renouvier, were resolute advocates of finitism, Couturat presented a vigorous case in behalf of an actual infinite. In opposition to the formalist theories of number of Julius Dedekind, Leopold Kronecker, and Hermann Helmholtz, he bases number on magnitude—not on a strictly spatial intuition but on magnitude considered as the object of a "rational intuition." This is why, of the various generalizations of number—the arithmetical, the algebraic, the geometrical—he regards the geometrical as the most rational. His reasoning consisted of offering the actual infinite as a new generalization of number, analogous to those that resulted in signed numbers, fractions, irrationals, and imaginaries. All of these numbers at first seemed to be arithmetical nonsense, but they took on meaning once they were recognized as suitable for representing new magnitudes and for allowing various operations on them that were hitherto impossible. The justification for infinite numbers is that they are indispensable for maintaining the continuity of magnitudes.

From this point on, Couturat's studies proceeded in three areas closely associated in his mind—the history of philosophy, logic and the philosophy of mathematics, and the development of a universal language.

After writing an essay (his Latin complementary thesis) on the myths of Plato, he devoted himself to Gottfried Wilhelm Leibniz, the great infinitist, whose reinterpretation he undertook independently of Bertrand Russell but at the same time and in the same sense. As indicated by the title of his book *La logique de Leibniz* (Paris, 1901), Couturat had at first intended simply to study the precursor of modern logistic. He soon perceived, however, that Leibniz's "logic was not only the heart and soul of his system, but the center of his intellectual activity, the source of all his discoveries, … the obscure or at least concealed hearth from which sprang so many *fulgurations.*" The manuscripts he discovered at

Hanover, a copious collection of which he published in *Opuscules et fragments inédits de Leibniz* (Paris, 1903), further strengthened Couturat in this conviction. Considering only Leibniz's known, celebrated works, if we wish to find the real root of his system, we must look not to the *Monadology* or the *Theodicy* but to the *Discourse on Metaphysics,* together with the *Correspondence with Arnauld,* which is, as it were, a commentary on the *Discourse.* Taking the old formula *praedicatum inest subjecto* in all its rigor, Leibniz held that every true proposition can be resolved into identities provided one pursues its analysis to the end. Contingent or factual truths differ from the necessary truths of reason only in respect to the infinite length of the analysis, an analysis which God alone is able to complete. Couturat showed, with supporting texts, that all the theses of the Leibnizian metaphysics are obtained from this position and derive their unity from it. The system thus appears as a panlogism.

It is likewise to his interest in Leibniz that we may ascribe, indirectly, Couturat's important study "La philosophie des mathématiques de Kant," published in the *Revue de métaphysique* (1904) on the centennial of Immanuel Kant's death. In *L'infini mathématique* Couturat had already criticized the Kantian antinomies that claim to establish the impossibility of an actual infinite. He now concluded that "the progress of logic and mathematics in the nineteenth century has invalidated the Kantian theory and decided the issue in favor of Leibniz" and his ideal of a completely "intellectualized" mathematics. The majestic edifice of the three *Critiques* lacks the indispensable basement of a logic on a level with science. "The brass colossus has feet of clay."

Deploring the fact that C. I. Gerhardt, in editing Leibniz, had separated the mathematical writings from the philosophical, Couturat could not but associate himself with the task assumed by the newly founded *Revue de métaphysiqu?* of working for a *rapprochement,* unfortunately broken off in the nineteenth century, between philosophers and scientists. After the establishment of the *Revue* in 1893, scarcely a year passed when he did not publish one or more articles in this spirit (some thirty at the time of his death, plus three that appeared posthumously). Rather than present original views, he dedicated himself with great disinterestedness to making known the views of others, mainly foreigners. He explained to French philosophers the mathematical logic of Guiseppe Peano, the universal algebra of Alfred North Whitehead, and the foundations of geometry and the principles of mathematics according to Russell. He vigorously defended both the new logic (to whose diffusion he contributed with his *L'algèbre de la logique,* Paris, 1905) and the Russellian logistic. This involved him in a celebrated controversy with his former teacher Jules Henri Poincaré. Although at the time Poincaré was often able to score against his opponent, subsequent developments in logic and mathematics have been more favorable to Couturat on many points.

Couturat's admiration for Leibniz, who dreamed of a universal language; his adherence to logistic that he saw as the source of an algorithm disengaged from the contingencies and irregularities of the natural languages; his participation in the organization of the first International Congress of Philosophy (Paris, 1900); his active collaboration with André Lalande in the preparation of the *Vocabulaire technique et critique de la philosophie* (Paris, 1926); and his rationalism, which one may characterize as militant in the sense that his purpose was less to rediscover reason in things than to work to make it rule among men—all these converging concerns led him to devote himself more exclusively to a task which became a veritable apostolate for him—the creation and adoption of an international auxiliary language by the rationalization of Esperanto and Ido. He prepared himself for this mission first by studying and then by publishing, in collaboration with Léopold Léau, the *Histoire de la langue universelle* (Paris, 1903). After 1900, Couturat was the moving spirit of the Délégation pour l'Adoption d'une Langue Auxiliaire Internationale, initiated by Léau, and later of the Akademie di la Lingue Internaciona Ido. In 1908 he founded and directed until his death the monthly review *Progreso,* written in the reformed language and designed to propagate it. The opposition of many Esperantists and the death of Couturat, which happened to come at the very moment when a war that exacerbated national particularisms was breaking out, caused the abandonment of the project. His friends and admirers have often regretted that Couturat should have expended so much effort in vain and sacrificed his wide talent to a noble dream.

See also Cantor, Georg; Helmholtz, Hermann Ludwig von; History and Historiography of Philosophy; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Plato; Poincaré, Jules Henri; Renouvier, Charles Bernard; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bibliography

WORKS BY COUTURAT

L'algèbre de la logique. Paris: Gauthier-Vill, 1905. The 1914 edition (Paris) was republished at Hildesheim, Germany, in 1965, Georg Olms edition.

Les principes des mathématiques. Paris: Alcan, 1905. Republished under the same title with the addition of the article "La philosophie des mathématiques de Kant" in Hildesheim, Germany, in 1965.

La logique de Leibniz. Georg Olms edition. Hildesheim, Germany, 1961.

WORKS ON COUTURAT

Cassirer, Ernst. "Kant und die moderne Mathematik. Mit Bezug auf Russells und Couturats Werke über die Prinzipien der Mathematik." *Kantstudien* (1907).

Lalande, André. "L'Oeuvre de Louis Couturat." *Revue de métaphysique* 22 (1914): 644–688. Includes a detailed bibliography.

Robert Blanché (1967)
Translated by Albert E. Blumberg

CRAIG'S THEOREM

In mathematical logic, Craig's Theorem—not to be confused with Craig's Interpolation Theorem—states that any recursively enumerable theory is recursively axiomatizable. Its epistemological interest lies in its possible use as a method of eliminating "theoretical content" from scientific theories.

PROOF OF CRAIG'S THEOREM

Assume that S is a deductively closed set of sentences, the elements of which may be recursively enumerated thus $F(0)$, $F(1)$, …, $F(n)$, … where F is a recursive function from natural numbers to sentences (we assume that expressions, sentences, etc., have been Gödel-coded in some manner). The set of theorems of an axiomatic theory is automatically recursively enumerable. But in general a recursively enumerable set is not automatically recursive. An example of a recursively enumerable set that is non-recursive is the set of logical truths in a first-order language with a single dyadic predicate. This follows from Church's Theorem, which states that the general notion of first-order validity is undecidable. However, through a trick devised by Craig, we can define a recursive set Craig(S) whose deductive closure is S. Let A be a sentence and n a natural number. Let A^n be the $(n+1)$-fold conjunction $A \wedge \ldots \wedge A$. The sentence A^n is logically interdeducible with A. Next consider sentences of the form $F(n)^n$. Define Craig(S) to be $\{F(n)^n: n \in N\}$. The deductive closure of Craig(S) must be S, since each element of Craig(S) is equivalent to an element of S. Next we give an informal decision procedure for membership in Craig(S). Given a sentence A, to decide whether $A \in$ Craig(S), first check if A has the form B^n, for some sentence B and number n. Through unique readability this is checkable, and if A is not of this form then $A \notin$ Craig(S). So suppose that A is of the form B^n. We calculate $F(n)$, and if B is indeed $F(n)$ then $A \in$ Craig(S). And otherwise $A \notin$ Craig(S). The existence of a decision procedure for membership in Craig(S) implies that Craig(S) is recursive. The set Craig(S) is therefore a recursive axiomatization of the theory S.

CRAIGIAN ELIMINATION. The logical positivists held that, under a logical reconstruction, a scientific theory is an axiom system formulated in a language $L(O, T)$, in which extra-logical predicates and function symbols are classified as either O-terms, for observational properties, or T-terms, for theoretical properties. Statements in $L(O, T)$ can be classified as observational, theoretical, or mixed, depending upon the presence or absence of O-terms or T-terms. Deleting theoretical terms yields a sublanguage $L(O)$ whose sentences express observational or empirical claims about the world. Assume that the property of being an $L(O)$-sentence is recursive. Consider a recursively enumerable theory S in $L(O, T)$. The empirical content of S is the set of $L(O)$-theorems of S. This is a subtheory of S obtained by a restriction on a recursive property. So it is recursively enumerable too. According to Craig's Theorem there is a recursive set of $L(O)$-sentences whose deductive closure is the empirical content of S. According to these assumptions we can therefore recursively axiomatize the empirical content of any given scientific theory S, obtaining a recursive axiom system Craig(S), known as the Craigian reaxiomatization of S's empirical content.

PHILOSOPHICAL SIGNIFICANCE OF CRAIGIAN ELIMINATION

Instrumentalism or positivism about science involves a scepticism towards the non-observational content of a scientific theory. Lacking such content the Craigian reaxiomatization Craig(S) provides an object of rational belief compatible with instrumentalist or positivist scruples. Note that this elimination method need not be based on an observation/theory distinction. With obvious modification it can be used as a way of eliminating, for example, the mathematical content from a scientific theory formulated using mathematical predicates and quantification over sets, functions, and so forth, or as a way of eliminating theoretical content from a psychological theory that refers to mental states, and so on. Craigian reaxiomatization offers a possible elimination strategy for a variety of instrumentalist positions.

CRITICISMS OF CRAIGIAN ELIMINATION

Of the aforementioned there are two methodological criticisms. First, even if the original theory *S* is presented in a simple manner, the reaxiomatization Craig(*S*) will be complex and thus will violate the canon of *simplicity* which we might impose on admissible theories. Second Craig(*S*) is *parasitic* upon the original theory *S* and so does not really stand alone from the original theory. Indeed Craig(*S*) is a bizarre theory, having infinitely many axioms of the form A^n, where *A* is an empirical consequence of *S*. Hartry Field refers to Craigian reaxiomatization as "bizarre trickery" and complains that Craig(*S*) is "obviously uninteresting, since [it] does nothing towards explaining the phenomenon in question in terms of a small number of basic principles" (Field 1980, p. 8). A third criticism is that Craigian elimination rests on a mistaken conception of scientific theories, namely a *syntactic view* of theories. This criticism has been urged by Bas van Fraassen, who writes "empirical import cannot be isolated syntactically … the reduced theory [Craig(*S*)] is not a description of the observable part of the world of *S*; rather it is a hobbled and hamstrung version of *S*'s description of everything" (van Fraassen 1976, pp. 87–88). A final criticism attacks the tenability of the observation/theory distinction required. A simple example of this is that although "red" seems a paradigmatic observational term, we can nonetheless speak of red blood cells, which are too small to be visible to the naked eye (see Putnam 1962).

With respect to certain assumptions discussed above concerning the notion of "empirical content," Craig's Theorem tells us that we can reaxiomatize the empirical content of a scientific theory, thereby eliminating apparent reference to unobservable objects and properties. However this elimination procedure has not found many adherents, and it seems safe to say that the significance of Craigian elimination is primarily pedagogical.

See also Field, Hartry.

Bibliography

Craig, William. "On Axiomatizability Within a System." *Journal of Symbolic Logic* 18 (1953): 30–32.

Field, Hartry. *Science without Numbers*. Princeton, NJ: Princeton University Press, 1980.

Putnam, Hilary "What Theories Are Not." In *Logic, Methodology and Philosophy of Science*, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski. Stanford University Press, 1962. Reprinted in Hilary Putnam *Mathematics, Matter and Method: Philosophical Papers*, Vol. 1, Cambridge University Press, 1979.

van Fraassen, B. C. "To Save the Phenomena. " *Journal of Philosophy* 73 (1976): 623–632 (page references to the reprint in David Papineau, ed. *Philosophy of Science*. New York: Oxford University Press, 1996).

Jeffrey Ketland (2005)

CRATYLUS

Cratylus, an Athenian, was contemporary with Socrates but was probably considerably younger. He was, according to Aristotle, a follower of the doctrines of Heraclitus, and Plato, in his youth, was closely associated with him. Aristotle implies that this was before he came under Socrates' influence, although later sources put the influence of Cratylus upon Plato after the death of Socrates.

Cratylus took as his starting point the doctrine of the flux of phenomena (here assumed to have been a genuine doctrine of Heraclitus, despite G. S. Kirk's objections), and he capped Heraclitus's saying that one cannot step twice into the same river by adding "nor once either." His reason clearly was his contention that the river is changing even as you step into it. He ended by coming to the view that one ought not to say anything, but only move the finger, since no true statement can be made about a thing that is always changing. According to Aristotle, upon whose evidence the above account rests, Plato took from Cratylus the belief, which he maintained even in later years, that all sensible things are always in a state of flux and that there is no knowledge about them.

Plato in the *Cratylus* attributes to him the doctrine that everything has a right name of its own, fixed by nature, and somehow or other this one right name will point to the nature of the thing named.

At an early stage it became clear to modern critics that the contention that there is a right name that indicates the true nature of a thing is apparently inconsistent with the doctrine of a Heraclitean flux in phenomena, since this flux would prevent a thing from having an abiding nature. Attempts to explain this contradiction in Cratylus's position have been numerous. Frequently it has been supposed that Cratylus either did not have a doctrine of words at all or else did not believe in the flux doctrine.

All such explanations seem misguided. Aristotle makes it clear that the final step—the refusal to use words—came after a previous period when Cratylus was already a Heraclitean. The implications of Plato's account are also clear; Cratylus at the time of the dialogue had long been interested in the doctrines of Heraclitus, and he

also held the theory of words attributed to him. It might be that he failed to realize the inconsistency at the stage represented by the dialogue, and, when the inconsistency became clear, subsequently proposed to abandon speech. More probably, at the time of the dialogue he inclined to the view to which he is clearly attracted when Socrates mentions it, namely, that words themselves in some sense flow, and so point to the flowing nature of the objects to which they refer (*Cratylus,* 437D).

See also Aristotle; Heraclitus of Ephesus; Plato; Socrates.

Bibliography

Testimonia in Diels-Kranz, *Fragmente der Vorsokratiker,* Vol. II, 10th ed. (Berlin: Weidmann, 1961). See also V. Goldschmidt, *Essai sur le Cratyle* (Paris: H. Champion, 1940); G. S. Kirk, "The Problem of Cratylus," in *American Journal of Philology* 72 (1951): 225–253; D. J. Allan, "The Problem of Cratylus," ibid., 75 (1954): 271–287; H. Cherniss, "Aristotle, Metaphysics 987a32–b7," ibid., 76 (1955): 184–186.

G. B. Kerferd (1967)

CREATION AND CONSERVATION

Many religions view the universe as the creation of a divine being or beings. The value of such a view is manifold. Among other things, it implies that the world is at least partly a product of divine planning and governance. To the extent that this is so, the world can be expected to be an orderly place, made for a purpose and ruled by providence. Thus, humans can anticipate that what befalls them in earthly life will not have occurred by accident, that their fortunes and destiny are, at least in some measure, divinely ordained. The idea of divine governance of the world also offers a possible basis for grounding principles of moral conduct; finally, if human destiny is in the hands of a higher power, there is at least hope of a life beyond the grave.

The doctrine of creation characteristic of the Judeo-Christian-Islamic tradition is a very strong one, for at least three reasons. First, these religions understand the universe to be entirely the product of one God. Second, creation is understood to be *ex nihilo*—that is, the universe is not fashioned out of any preexisting thing. Third, the world is not just created by God "in the beginning," but is also sustained in being by him for its entire existence; thus, the fact that the world is able to persist beyond the present moment is as much owing to the creative action of God as is the fact that it exists at all. Taken together, these claims indicate that divine governance of the world is unified, close, and thorough, with implications about both the nature of the world and God's relation to it. These implications are related in turn to a number of problems in philosophical theology and to certain aspects of contemporary scientific cosmology.

THE COSMOLOGICAL ARGUMENT

No single account of creation could ever portray the richness of the Western philosophical tradition on the matter. There are, however, a number of common themes. Typically, treatments of God as creator begin with an argument for the existence of God—usually some version of either the teleological argument, which is based on premises about order or purpose in the universe, or the so-called cosmological argument, which invokes God as an uncaused first cause of all things. The latter argument shall be examined first.

The cosmological argument is traditionally presented as a deductive argument. Put very roughly, it asserts the following:

> (1) The universe of our experience need not have existed—that is, that its existence is contingent.
>
> (2) The existence of contingent things must have an explanation.
>
> (3) The only or best explanation for the existence of the universe is the creative activity of a necessarily existing "first cause."
>
> (4) Therefore, there is such an entity or being.

Further, it is held, this being is in fact the personal God of traditional Western monotheism.

Of the premises of this argument, the first seems plausible. The second is a version of the principle of sufficient reason. It is needed if the argument is to be deductively valid, but in the end it is probably damaging to the argument. The principle of sufficient reason is not a necessary truth, and so cannot be known a priori; any effort to establish its truth a posteriori is apt to be inconclusive, in part because the truth of the principle depends precisely on whether it holds with regard to cases such as the existence of the universe, which is precisely the point at issue.

Suppose, then, that premise number two is dropped. What remains is best interpreted as a kind of inductive argument, an inference to the best explanation, according to which the existence of the universe is a result of the

causal activity of a necessary being. In such an argument, it is best to separate two claims that are implicit in the third premise above:

(3a) The creative activity of a necessarily existing first cause is sufficient to explain the existence of the universe.

(3b) No alternative hypothesis is sufficient to do the job.

With this clarification, and premise number two now out of the way, an appropriate conclusion might now be something like:

(5) Relative to the evidence of experience, we have better reason for thinking there is a necessarily existing first cause of the universe than for thinking otherwise.

Because this is an inductive argument, the conclusion no longer follows necessarily from the premises. In this, however, it is no different from any inductive argument, including many that we find quite persuasive—for example, arguments for the existence of subatomic particles or even for exotic and unexpected phenomena such as quantum entanglement. No doubt, a skeptic may treat this feature as a reason for denying the conclusion, but that is not a point of interest. A skeptic can find a reason to deny any conclusion. A second important point is that the God postulated in the argument is described as a necessary being—that is, a being whose existence is necessary or a being that exists by its very nature. Some philosophers have questioned whether such a being is possible, and that is an issue worthy of consideration. It would, however, be logically misguided to greet this argument with the question, "What caused God?" By definition, a necessary being is existentially self-sufficient: It neither has nor requires a cause. Admittedly, it is not obvious that a necessarily existing first cause of the universe ought to be identified with the personal God of traditional belief. Proponents of cosmological arguments have, however, been well aware of this point; medieval demonstrations of a creator—that of Thomas Aquinas, for example—were often followed by lengthy consideration of what characteristics might appropriately be attributed to such a being.

Perhaps the most interesting issues about the cosmological argument concern premises 3a and 3b. Whether, as 3a asserts, the activity of a creator God explains the existence of the word will depend on how we understand that activity. Presumably, it does not consist in a sequence wherein God first commands that the universe exist, and the command then causes its existence. For if, as is usually supposed, causal relations are themselves contingent, then God would first have to create the causal mechanism by which his commands gain efficacy. This would require another command, and a vicious regress would ensue. How, then, should the activity of creation be understood? One attractive possibility is an analogy with human creation: for example, a writer envisioning a drama, a composer inventing a melody, or a scientist coming up with a hypothesis. On this kind of view, the first cause would indeed have to be conceived as personal, since the world would owe its existence to a knowing will, of which it would be the *content* as well as the product. That is to say, unlike the products of human creation, the universe would have its existence in God both as a concrete reality and as something known.

Premise 3b must itself be established inductively: That is, we must canvas the known alternatives to the hypothesis of a creator and show that they do not work. Of course, even if we succeed, it may be that some as yet undiscovered explanation for the existence of the world will be found superior to any invoking a creator God. Still, the cosmological argument is greatly strengthened if alternative hypotheses can be eliminated. Historically, the most favored alternative by far has been the hypothesis that the world had no temporal beginning but rather is infinite in duration, its existence at each moment being a causal consequence of the immediate past, from which it is generated in accordance with scientific law. In fact, however, this alternative is all but indefensible. Scientific laws, classically at least, are not diachronic: That is, they do not speak of causes that occur at one moment and effects at another. Action and reaction are simultaneous in Newton's scheme. The application of net force produces acceleration at the instant of application; if at t an object is not acted upon by a net force, then it is at rest or in uniform rectilinear motion at t.

We can, of course, deduce the state of a closed system at a later time if we know its state at t, but only if we *assume* as a premise that the mass/energy of which it is constituted will continue to exist. Nor will it do at this point to invoke conservation principles. For, again, the law that mass/energy is conserved holds only of closed systems—that is, systems in which, *ex hypothesi*, mass/energy is neither gained nor lost. Conservation laws are not, however, mere tautologies. They tell us something very important: that physics is not about things just *being*, but about how they change; and that although we may learn a great deal from science about the development of the universe over time, and the ways in which the items of our experience combine and separate and

change, physics has next to nothing to tell us about the existence of things.

A second difficulty with the alternative hypothesis is that we have no idea what it would be like for the universe to be able to bootstrap itself into the future by sustaining its own existence. No one has ever described a mechanism by which this could occur; and if one should deny the need to do so—invoking, say, a principle of "existential inertia" by which, once in existence, things "naturally" tend to continue existing— the explanation becomes empirically vacuous, a mere redescription of the phenomenon to be explained. Finally, even if one temporal stage of the universe could give rise to a succeeding one, the question that drives the cosmological argument would go unanswered.

For the problem is not why the universe exists at this moment, but why it exists at all. If I ask you why the bordelaise sauce in the upper container of a double boiler is hot, you may fairly explain that it is heated by the boiling water in the lower container. If I ask you why the water is boiling, you could in principle reply that this is in fact a triple boiler: that the water is heated by still further water, boiling merrily another level below. But if I ask you how heat gets into the system at all, it will not do to postulate an infinity boiler, for water is never anything but contingently hot. You have to come to something essentially hot—fire, perhaps—to answer my question. Similarly, if the problem is to explain the existence of contingent things, an infinite sequence of them is of no avail. Only if we postulate a necessary being will an explanation be possible.

CONSERVATION

If these points are correct, then the traditional cosmological argument is in a considerably stronger position than is often supposed. Its plausibility is the same, moreover, whether the past duration of the universe is finite or infinite. If there is never a natural accounting for the existence of contingent things, then at any moment of their existence, only the activity of a necessary being can explain them. This is the essence of the religious doctrine of conservation, according to which God is as much responsible for the persistence of the universe as for its being there at all. Many have thought, however, that the doctrine of conservation leads to a serious difficulty. If God is, at each moment, the cause of all that exists, what place can there be for natural causes? It is unreasonable to think God is causally responsible for the present existence of my chair unless he is also responsible for its properties. Indeed, the seventeenth-century philosopher Nicolas Malebranche argued that nothing else is possible, that it is self-contradictory to suppose God could create a chair that is neither at rest nor in motion, and has no color or mass or any other property characteristic of a chair. But if God is creatively responsible for all the properties of the chair at each moment of existence, are not natural causes simply redundant? God's will as creator must, after all, be presumed completely efficacious. But if it is, what efficacy is left to natural causes?

Malebranche's answer was blunt: none. He held a doctrine known as *occasionalism*, according to which the events we ordinarily take to be causes (for example, my pushing on the chair) are only occasions for God to exercise his own causal power (to create the chair in motion), the only causal power that is legitimately efficacious. But occasionalism has uncomfortable consequences. Perhaps the worst is that if it is true, we do not perceive the world in the way we think we do—that is, by the action of the things in it on our senses. Rather, our sensations are caused in us by God so as to match what is going on in the world. And then we are only a step away from the idealism of George Berkeley: that is, from moving to the conclusion that the so-called physical universe must be superfluous to God's plan, then denying its existence and attempting to reduce its contents to nothing but sets of ideas.

It is possible, however, to avoid occasionalism if the suggestion of the previous section is correct: that natural causation is not a matter of conferring existence to begin with. Much more plausible accounts are possible. In the realm of physical action in particular, causal interaction may be viewed as a matter of energy transfer, wherein quantities presumed to be conserved—motion, momentum, charge, and the like—are transmitted from one entity to another. As such, what we normally take to be causal processes (one billiard ball imparting motion to another, to cite the classic example) count as genuine exercises of causal power yet do not carry the suggestion of one event conferring existence on another. The case of perception, where what is caused is something mental, is notoriously more difficult. But if a similar solution can be found there, we have every reason to think both natural and divine causation can be accepted without setting up a false competition between them.

CREATION AND SCIENTIFIC COSMOLOGY

Philosophical arguments for the existence of God are widely understood to be reinforced by two recent developments in scientific cosmology. The first is the so-called

big bang theory, according to which the universe of our observation constitutes a kind of continuous explosion that commenced some 15 billion years ago from an initial singularity in which the entirety of space-time was compressed to a state so dense as to be indescribable by any known principles of physics. This theory is now well confirmed, and it implies that the universe did indeed have a temporal beginning—so radical a beginning, in fact, that it would be scientifically meaningless to speak of a time prior to the big bang.

That the universe had a temporal beginning is in line with many religious creation narratives. Theists have therefore tended to treat the big bang theory as confirmation of their views, holding that it is far more plausible to postulate a creator to explain the world's beginning than to claim it "just happened." Opponents have found the appeal to such a cause unscientific. Some have argued that, in fact, a divine act of creation is not possible, because even a divine cause must precede its effect, and the big bang allows for no time prior to itself. Others have posed mathematical models for the universe's origin that would avoid the claim of an initial singularity and the attendant implication of a temporal beginning.

The suggestion that a proper science ought not to be postulating a creator has much to be said for it; science is fairly taken to be solely concerned with natural phenomena and natural explanations of them. Theists who take comfort from big bang theories need not, however, be deterred by such a delimitation of scientific purview, given that they seek an explanation for the existence of the entire natural order—something a science thus delimited cannot in principle provide. As for the claim that even a divine cause must precede its effect, that seems mistaken. On the contrary: if, as is argued above, the laws that govern natural processes are synchronic rather than diachronic, then even natural causation must be understood in a way that makes cause and effect simultaneous—in which case any support for the claim that a divine cause must precede its effect vanishes. Humans may eventually come to understand the genesis of the cosmos according to some model other than the standard big bang. But the alternatives presently available face problems of internal coherence and of testability, and so have yet to offer strong competition.

The second development in cosmology that is often taken to support claims of a creator is the realization that living beings of the sort with which we are familiar owe their existence to a wondrously exacting fine-tuning of various physical parameters. For example, if the ratio of the universe's rate of expansion to its total mass were increased or decreased by only one part in a million, there would be no stars and planets to support life. If the strong nuclear force were increased by just 1 percent, it would have been impossible for carbon to form; an increase of 2 percent would have ruled out the formation of protons from quarks. On the opposite side, a decrease of 4 percent would have allowed no atoms other than hydrogen to form.

Examples like this can be multiplied at considerable length, and the likelihood that all the requirements for life that they embody should be satisfied in one universe is exceedingly remote. The fact that our universe does exactly that has therefore been held to justify a teleological argument for the existence of God. That is, it is argued that the only way to explain the fine-tuning of our universe for life is to postulate an intelligent creator who designed it to be so. Opponents have countered that the universe visible to us may be only one of a great many worlds, perhaps even an infinite number, in which many or even all possible combinations of basic parameters are displayed. If so, then the fine-tuning of our world might be "explained," at least in the weak sense that the appearance of such a universe would be made more likely, or even certain. In addition to being plainly ad hoc in character, these speculations too present problems of testability—not to mention the difficulty of truly explaining the existence of such an ensemble of worlds by specifying a mechanism that could cause it to appear, and guarantee its exhaustiveness. Like the big bang, however, the issue of fine-tuning is a subject of intense interest, and there is doubtless a good deal more to be said about both.

See also Berkeley, George; Cosmological Argument for the Existence of God; Cosmology; Laws, Scientific; Malebranche, Nicolas; Teleological Argument for the Existence of God; Thomas Aquinas, St.

Bibliography

Barrow, John, and Frank Tipler. *The Anthropic Cosmological Principle*. Oxford: Clarendon Press, 1986.

Craig, William L. "Professor Grünbaum on Creation." *Erkenntnis* 40 (1994): 325–341.

Craig, William L., and Quentin Smith. *Theism, Atheism, and Big Bang Cosmology*. Oxford: Clarendon Press, 1993.

Dowe, Phil. *Physical Causation*. Cambridge, U.K.: Cambridge University Press, 2000.

Freddoso, Alfred J. "Medieval Aristotelianism and the Case Against Secondary Causation in Nature." In *Divine and Human Action: Essays in the Metaphysics of Theism*, edited by Thomas V. Morris. Ithaca, NY: Cornell University Press, 1988.

Grünbaum, Adolf. "Creation as a Pseudo-Explanation in Current Physical Cosmology." *Erkenntnis* 35 (1991): 233–254.

Hawking, Stephen. *A Brief History of Time*. New York: Bantam Books, 1988.

Kvanvig, Jonathan L., and Hugh J. McCann. "Divine Conservation and the Persistence of the World." In *Divine and Human Action: Essays in the Metaphysics of Theism*, edited by Thomas V. Morris. Ithaca, NY: Cornell University Press, 1988.

Mackie, J. L. *The Miracle of Theism*. Oxford: Clarendon Press, 1982.

Malebranche, Nicolas. *Dialogues on Metaphysics*. Translated by Willis Doney. New York: Abarus Books, 1980.

Quinn, Philip L. "Divine Conservation, Secondary Causation, and Occasionalism." In *Divine and Human Action: Essays in the Metaphysics of Theism*, edited by Thomas V. Morris. Ithaca, NY: Cornell University Press, 1988.

Rowe, William L. *The Cosmological Argument*. Princeton, NJ: Princeton University Press, 1975.

Thomas Aquinas. *Summa Contra Gentiles*, Bk. III, Chs. 64–70. Translated by Vernon J. Bourke. Notre Dame, IN: University of Notre Dame Press, 1975.

Thomas Aquinas. *Summa Theologiae* I, Ques. 1–11. London: Blackfriars, 1964.

Hugh J. McCann (2005)

CREATIVITY

It has often been claimed that genuine creativity is largely if not entirely a matter of inspiration—the sudden, involuntary, and inexplicable outpouring of innovative ideas and actions. In many expressions of this thesis, including Plato's, the source of this outpouring is a sacred instance—a spirit or muse—while in other versions it is the unconscious mind. The antithesis to such inspirationist theses is the rationalist doctrine that all creativity is ultimately reducible to a form of calculation or more or less deliberate problem-solving.

Although both the extreme inspirationist thesis and the rationalist antithesis have adherents, many investigators find an intermediary position more tempting. Inspirationist theses are hard to square with basic, naturalist inclinations and with a commitment to scientific research. That creative behavior is complex and hard to explain does not mean that it is essentially mysterious or could never be modeled with some modest measure of accuracy. Inspirationism is further challenged by evidence that most if not all significant episodes of innovation require industry, rational thought, and action. Extreme rationalist accounts, on the other hand, confront testimony regarding the involuntary and sudden onslaught of important new ideas, such as vivid accounts of inspiration's "sudden illumination" offered by Pyotr Il'ich Tchaikovsky (1970 [1878]), Henri Poincaré (1952 [1908]), Albert Einstein, and scores of other impressive informants. Yet some of these same sources identify rational elements of creativity, as when Tchaikovsky goes on to underscore the necessity of daily, strenuous efforts. How to balance such divergent insights and data in a synthetic account of creativity remains a matter of great controversy.

UNDERSTANDING *CREATIVITY* AND *CREATION*

A nontrivial problem is that of specifying how the terms *creativity* and *creation* are to be understood in the first place. While there is widespread agreement that creative acts and their products must be new or innovative, there is disagreement as to the sense in which this is true, as well as with regard to other conditions on creativity. In what way are all creative acts novel or original? Although every particular action is new in the sense that this particular event never happened before, genuine creativity involves something far less common. Saying that creative acts must never have been performed before is not only vague, but overly restrictive. Is not someone's wholly independent repetition of a great discovery creative? With this question in view, Margaret A. Boden distinguishes between historical and psychological creativity. An act is *historically* creative, she proposes, only if it has never been had before by anyone else in all of human history. In contrast, *psychologically* creative acts may replicate previous inventions.

Even if one agrees to focus on what Boden calls psychological creativity—a move disputed by some theorists, including Mihály Csikszentmihályi—there remains the problem of coming up with a nontrivial elucidation of the novelty clause. To that end, Boden attempts to characterize radical psychological novelty in terms of the creative act's transformation of "a conceptual space," by which she means the principles that unify and structure a given domain of thinking and action. She contrasts the relatively uncreative writer who produces a new and interesting novel while conforming entirely to the rules of some established genre, to a genuinely creative writer who creates a strikingly new work that transforms generic patterns and establishes a new literary category. Boden further contends that what sets off the genuinely creative transformations of conceptual spaces is that their results could not have been thought before by the person working within that space. Given that a conceptual space is governed by a system of constitutive rules, its transfor-

mation entails that at least one of these rules is dropped or violated in a genuinely creative act. Jon Elster (2000) also explores the relation between creativity and rules or constraints of various sorts, arguing that originality—which may be either sterile or of genuine value—involves not merely a rebellious violation, but the revolutionary replacement of constraints.

David Novitz (2003) challenges Boden's proposal and identifies counterexamples. Some inventions, such as Thomas Edison's creation of the phonograph and Henri Matisse's use of color in his paintings, are not plausibly described as having arisen within a rule-governed conceptual space. Matisse, after all, was playing around with color combinations he found pleasing, hardly a pursuit organized by a system of rules. And some actions that do arise within a rule-governed practice or "space," such as a chess player's invention of a new opening, may nonetheless be genuinely creative. Thus if Boden's discussion offers insight into some forms of creativity, it does not adequately cover all of them.

Novitz defends an alternative, "recombination" theory of creativity. He proposes that creative acts are novel in the sense that they are not predicted by, and are surprising to a given population. Alternately, creative acts are those which would have been surprising had the members of the population become aware of them. Novitz does not specify how the population in question is to be identified, but does remark that the members of the population must have some familiarity with some of the ideas or objects that get recombined in the creative act. Another alternative is to say that it is the invention's creator or creators who must be surprised by the discovery.

Novitz argues that it is a mistake to associate creativity with the making of art. Many creative acts and inventions have nothing to do with the fine arts, and much art-making, or creation, is routine and devoid of creativity. Novitz joins a long tradition in specifying that genuine creativity must, in addition to manifesting a novel or surprising recombination of ideas or objects, bring forth a result having some real, positive value: "Creative acts are valued positively because they are intended to, and have the potential to, satisfy actual human needs and desires" (2003, p. 186). Novitz also allows that creative acts may also display a form of intrinsic value in addition to such instrumental value. These points are not, however, uncontroversial, as some authors are willing to allow that a fiendish or malicious invention, or intentionally immoral act, could be creative. Some forms of creativity may be useless. Nor is it clear that a viable conception of creativity need entail strong, realist commitments in the theory of value.

THE NATURE OF THE CREATIVE PROCESS

Additional controversy surrounds proposals concerning the nature or basic structure of the creative process or processes. One key issue has to do with the question whether the expression "the creative process" really refers to a single type of process or activity. Francis Sparshott (1981), John Hospers (1984–1985), and others state that there is no such thing as a single, determinate process involved in all creative acts, but instead, different sorts of processes having little or nothing in common. Another controversy concerns the extent to which creative activity can be adequately described as a species of problem-solving or means-end rationality. Vincent Tomas rejects the idea that artistic creation is "a paradigm of purposive activity" (1958, p. 2).David Ecker's description of the creation of art as "qualitative problem-solving" (1963) is critiqued by Monroe C. Beardsley (1965), who deems it a mistake to think that creative thinking, in the arts at least, is characteristically a matter of means-end calculations. Even if the sought-for aesthetic and artistic effects do depend on the artist's manipulation of some medium or media, creative work is not throughout guided by the effort to realize some preconceived goal or end: it would be unusual if the precise quality of the final painting were in the painter's mind from the start.

Various investigators have contended that there is a characteristic creative process having a hybrid nature. In an account popularized by Graham Wallas in *The Art of Thought* (1926), this process breaks down into four discrete stages. Creativity requires, first of all, apprenticeship and preparation: even the most brilliant innovator must learn his or her craft.Second comes "incubation," a stage in which the creative person stops working consciously on some problem, allowing unconscious processes to predominate. The result, when circumstances are favorable, is illumination or inspiration, the moment when some unexpected and innovative idea "pops" into mind. In the final stage of "verification," the creator assesses and revises what inspiration has yielded. As Beardsley observes, it would be wiser to replace talk of four linear stages with the idea of an interplay between two alternating phases, namely, preconscious invention and conscious criticism and selection of the latter phase's results. This is similar to Paul Valéry's contention that creative art-making is always a matter of both the spontaneous emergence of ideas, and conscious, means-end adjust-

ments and rearrangements of the latter (1957–1960 [1938]). Only their relative proportion varies, he adds.

Psychologists and cognitive scientists continue to attempt to provide models of complex creative processes in various domains, including musical composition, the formation of scientific hypotheses, the visual arts, and storytelling (for surveys, see Albert and Runco 1999; Boden 2004). The greatest challenge is perhaps that of providing detailed explanations and effective simulations of the processes that underlie and generate moments of inspiration, or "popping." Psychologists working in a range of traditions, including Gestalt theory, psychoanalysis, associationism, cognitive psychology, interactionism, systems theory, and so on, have devised elaborate labels for the mind's unconscious generation and selection of novel ideas. Some of these traditional insights have been revived in the development of computer simulations using connectionist and other approaches (Martindale 1995). And in a philosophical vein, Berys Gaut (2003) explores the Kantian connection between creativity, genius, and the imagination, taking metaphor's linking of diverse domains as a paradigm.

See also Imagination.

Bibliography

Albert, Robert S., and Mark A. Runco. "A History of Research on Creativity." In *Handbook of Creativity*, edited by Robert J. Sternberg. Cambridge, U.K.: Cambridge University Press, 1999.

Beardsley, Monroe C. "On the Creation of Art." *Journal of Aesthetics and Art Criticism* 23 (1965): 291–304.

Boden, Margaret A. *The Creative Mind: Myths and Mechanisms*. London: Weidenfeld and Nicholson, 1990. 2nd rev. ed., London: Routledge, 2004.

Boden, Margaret A. "What is Creativity?" In *Dimensions of Creativity*, edited by Margaret A. Boden. Cambridge, MA: MIT, 1994.

Ecker, David. "The Artistic Process as Qualitative Problem Solving." *Journal of Aesthetics and Art Criticism* 21 (1963): 283–90.

Elster, Jon. *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints*. Cambridge, U.K.: Cambridge University Press, 2000.

Feyerabend, Paul. "Creativity—A Dangerous Myth." *Critical Inquiry* 13 (1986–1987): 700–711.

Gaut, Berys. "Creativity and Imagination." In *The Creation of Art: New Essays in Philosophical Aesthetics*, edited by Berys Gaut and Paisley Livingston. Cambridge, U.K.: Cambridge University Press, 2003.

Ghiselin, Brewster, ed. The *Creative Process: A Symposium*. New York: Mentor, 1952.

Halper, Edward. "Is Creativity Good?" *British Journal of Aesthetics* 29 (1989): 47–56.

Hospers, John. "Artistic Creativity." *Journal of Aesthetics and Art Criticism* 43 (1984–1985): 243–255.

Kant, Immanuel. *Critique of Judgment* [1790]. Translated by Werner S. Pluhar. Indianapolis: Hackett, 1987.

Martindale, Colin. "Creativity and Connectionism." In *The Creative Cognition Approach*, edited by Steven M. Smith, Thomas B. Ward, and Ronald A. Finke. Cambridge, MA: MIT, 1995.

Novitz, David. "Creativity and Constraint." *Australasian Journal of Philosophy* 77 (1999): 67–82.

Novitz, David. "Explanations of Creativity." In *The Creation of Art: New Essays in Philosophical Aesthetics*, edited by Berys Gaut and Paisley Livingston. Cambridge, U.K.: Cambridge University Press, 2003.

Poincaré, Henri. *Science et méthode*. Paris: Flammarion, 1908. Translated by Francis Maitland as *Science and Method.* New York: Dover, 1952.

Sparshott, Francis. "Every Horse has a Mouth: A Personal Poetics." In *The Concept of Creativity in Science and Art*, edited by Denis Dutton and Michael Krausz. The Hague: Martinus Nijhoff, 1981.

Tchaikovsky, Pyotr Il'lich. "Letters" [1878]. In *Creativity: Selected Readings*, edited by P. E. Vernon. Harmondsworth, U.K.: Penquin, 1970.

Tomas, Vincent. "Creativity in Art." *The Philosophical Review* 67 (1958): 1–15.

Valéry, Paul. "L'invention esthétique" [1938]. In *Oeuvres.* 2 vols. Edited by Jean Hytier. Paris: Gallimard, 1957–1960.

Wallas, Graham. *The Art of Thought*. London: Jonathan Cape, 1926.

Paisley Livingston (2005)

CREIGHTON, JAMES EDWIN

(1861–1924)

James Edwin Creighton, an American idealist philosopher, was born in Pictou, Nova Scotia. Creighton was educated at Dalhousie College, Halifax (A.B., 1887), where one of his teachers was Jacob Gould Schurman, whom he later followed to Cornell University. He was appointed fellow in philosophy there in 1888 and studied in Leipzig and Berlin, returning to Cornell in 1889 as an instructor. He received his Ph.D. in 1892 with the thesis "The Will; Its Structure and Mode of Operation," and became associate professor. In 1895 he was elected Sage professor of logic and metaphysics, succeeding Schurman, and held that chair until his death. He received LL.D. degrees from Queens University (1903) and from Dalhousie (1914). While Creighton was dean of the graduate school at Cornell from 1914 to 1923, his flexible policies stimulated student initiative, but the administrative demands on his time limited his literary output. He

was coeditor of the *Philosophical Review* from 1892 until 1902, when he became sole editor, and he was American editor of *Kantstudien* from 1896 until his death.

Convinced that the intellectual life is a social venture, Creighton was a cofounder of the American Philosophical Association and in 1902 became its first president. His vigorous instruction influenced the development of philosophy in American education through the efforts of his students, twenty-two of whom honored him with a volume of articles, *Philosophical Essays* (New York, 1917), commemorating twenty-five years of his teaching.

Creighton's "speculative idealism" grew out of his view that philosophical inquiry must occur in the context of the history of ideas and must begin with "the standpoint of experience." But experience is not a simple, isolated particular which can be understood by analysis. Finite individuality has system implicit in it, and can be understood as a part of the order of the universe. It is unity in plurality and identity in difference. It is permeated with meaning. In short, Creighton identified it as the "concrete universal."

Thus, with Bernard Bosanquet, Creighton held that philosophical judgments are ways in which experience progresses toward its goal of intelligibility, and the task of such judgments is to disclose the implications of the dynamic coordinates of experience: mind, nature, and other selves. Reality cannot be identified with mind, will, or personality but must be comprehended as a system in which each entity plays a part as an individual and as a significant function of the purposeful whole. Epistemological problems traceable to Immanuel Kant's emphasis on the centrality of the knowing subject are artificial because mind by its very nature is already in touch with reality. Subject and object cannot be viewed as ontologically discrete but are correlative. Accordingly, Creighton dissociated himself from neorealism, which regards truth as a quality of single propositions; from pragmatism, which fails to see that thought modifies the internal structure of experience itself; and from Berkeleianism and other "mentalistic" idealisms, which interpret nature as a phase of mind, thereby transforming experience unnecessarily into an order of ideas instead of accepting objective reality as a direct intuition. Such idealisms, even Josiah Royce's absolutism, issue in subjectivism and thus deny the objective world. Creighton maintained that this conclusion would render all thought chaotic because the objective order is the presupposition of all rationality.

Appointed to the Carus lectureship in 1924, Creighton planned to develop his views on historical method in philosophy, but death intervened. He wrote virtually nothing on ethics, aesthetics, or religion, unlike his idealist contemporaries, but certain details of his system can be inferred from his excellent critical discussions of competing movements.

See also Bosanquet, Bernard; Idealism; Kant, Immanuel; Pragmatism; Royce, Josiah.

Bibliography

WORKS BY CREIGHTON

Translator with E. B. Titchener, of W. Wundt, *Lectures on Human and Animal Psychology.* London: S. Sonnenschein, 1894.

Translator with A. Lefevre, of F. Paulsen, *Immanuel Kant.* New York: Scribners, 1902.

An Introductory Logic. New York: Macmillan, 1909. The fifth edition (New York: Macmillan, 1932) was revised by one of Creighton's colleagues, H. R. Smart.

Studies in Speculative Philosophy, edited by H. R. Smart. New York: Macmillan, 1925. Fourteen of his 38 major articles were posthumously published in this volume, which is the best single source for his views, containing a select bibliography and his most representative essay, "Two Types of Idealism."

WORKS ON CREIGHTON

Blau, Joseph L. *Men and Movements in American Philosophy.* New York: Prentice-Hall, 1952.

Cunningham, G. Watts. *The Idealist Argument in Recent British and American Philosophy.* New York: Century, 1933.

Townsend, Harvey G. *Philosophical Ideas in the United States.* New York: American, 1934.

Critical discussions of Creighton's work can be found in the three books listed above.

Warren E. Steinkraus (1967)

CRESCAS, HASDAI
(1340–1410)

The Spanish rabbi and philosopher Hasdai Crescas was born in Barcelona, the scion of a distinguished family. He exercised considerable influence both in the Jewish community and at the Aragonese court. After the 1391 persecution of the Jews, in which his only son perished, Crescas moved to Saragossa, where he engaged in literary activity until his death.

Crescas's purpose was to defend Judaism from both internal and external subversion. To this end he composed his Spanish "Refutation of the Principles of Christianity" (extant only in Hebrew translation), a rational critique of Christian dogmatic theology, and his masterwork, *The Light of the Lord* (*Or Adonai*), conceived as an introduction to a legal code that was never composed. Crescas

wrote in the tradition of those thinkers, such as Judah Halevi and Nahmanides, who rejected the rationalistic compromising of Judaism with the teachings of Aristotle, but he differed from them in that he chose to combat the philosophers on their own ground. In this respect his position may be compared with that of Muhammad al-Ghazālī in Islamic philosophy. The *Light* is arranged as a dogmatic treatise, beginning with an exposition of the primary concept of God's existence and unity and followed by expositions of certain fundamental and subordinate doctrines. The first section, in which Crescas presents and criticizes the twenty-six basic propositions of physics which Maimonides (*Guide,* Part II, Introduction) culled from Aristotle, is concerned less with advancing a new system than with indicating the inadequacy of those of his forerunners. Crescas conceived of time as duration independent of motion and insisted on the possibility of a vacuum based on a conception of space as extension independent of body. These two notions enabled him to establish the existence of infinite time and space, thereby destroying the concept of the Aristotelian prime mover. Furthermore, the debate over creation ex nihilo is dismissed as futile since, in any event, all is derivative from God, who is the only necessary existent.

Crescas maintained both the literalness of the Biblical attributes and God's unity by advancing the Kalam-like theory of essential attributes compatible with God's absolute simplicity. These attributes are related to the subject as light rays are to the source of luminescence, one being inconceivable without the other, and are bound together by the unifying principle of the divine goodness. It is this goodness or perfection which characterizes the Divinity, rather than the Aristotelian concept of self-thinking thought.

The return of Crescas to the biblical conception of God is best exemplified in his treatment of the problem of the conflict between divine foreknowledge and human free will. Maimonides had taken refuge in the notion that God's knowledge has nothing in common with humanity's while Gersonides sacrificed divine knowledge of the future and the particulars to unconditional human free will. Rejecting both points of view, Crescas felt it unnecessary to reconcile divine knowledge (which he considered absolute) with free will but rather free will with causality. Definitely inclined toward determinism, he maintained that an act is contingent when considered in relation to itself but necessary in relation to its causes and to God's knowledge. Human consciousness of free will consists in the pleasure or disapproval felt when an act is committed.

Divine providence, prophecy, and immortality are not dependent on intellectual perfection, as in Maimonides and Gersonides, but rather on love and reverence for God, which is the purpose of the Divine Law and the universe. It is the substance of the soul itself, rather than the acquired intellect, which survives death.

Crescas's independence of Aristotle helped pave the way for such Renaissance thinkers as the younger Pico della Mirandola and Giordano Bruno. Of particular interest is Crescas's influence on the thought of Benedict (Baruch) de Spinoza, who knew his work well.

See also al-Ghazālī, Muhammad; Aristotle; Bruno, Giordano; Gersonides; Halevi, Yehuda; Islamic Philosophy; Jewish Philosophy; Maimonides; Spinoza, Benedict (Baruch) de.

Bibliography

The *editio princeps* of the *Or Adonai* (*Light of the Lord*) appeared at Ferrara in 1556 and has since been published several times. The first section, which deals with the 26 propositions, has been edited and translated into English by H. A. Wolfson in *Crescas' Critique of Aristotle* (Cambridge, MA: Harvard University Press, 1929). The section on free will was rendered into German by Philipp Bloch in *Die Willensfreiheit von Chasdai Crescas* (Munich, 1879).

Crescas's thought has been surveyed in M. Waxman, *The Philosophy of Don Hasdai Crescas* (New York: Columbia University Press, 1920) and M. Joel, *Don Hasdae Crescas' religionsphilosophische Lehren in ihrem geschichtlichen Einflusse dargestellt* (Breslau: Schletter, 1866). His influence on Spinoza has been discussed by Joel in *Zur Genesis der Lehre Spinozas* (Breslau: Schletter, 1871) and by D. Neumark in "Crescas and Spinoza," in *Essays in Jewish Philosophy* (Cincinnati, OH, 1929). For a full bibliography, see *Encyclopedia Judaica* (Berlin: Eschkol, 1928), Vol. V, pp. 698ff., 708.

Frank Talmage (1967)

CRESCAS, HASDAI [ADDENDUM]

Hasdai Crescas takes a radical anti-Aristotelian position, and yet he himself presents conclusions that are threatening to the traditional understanding of religion. For example, in attacking the views on the creation of the world by Maimonides and Gersonides he ends up presenting a theory that allows for an eternal world. The world could be eternal in the sense that it would be eternally dependent on God. According to Crescas there is no difficulty about the existence of a vacuum before the creation of the world, and so it is no good objecting to Aristotle that something

could not come from nothing. He even contemplates the possibility of this world being only one of a number of worlds, each existing along with the others.

Crescas is unusual also in accepting the existence of the infinite, a concept that many Aristotelians think suggests an absurdity, and the discovery of which is taken by Aristotelians to indicate an impossibility in the argument. The concept of infinity allows Crescas to envisage an infinite space in which a vast variety of worlds could exist.

Still, the qualms about infinity that his predecessors held had allowed them to argue for the necessity of a first cause, since otherwise the series of causes and effects would continue infinitely. Crescas's attack on Aristotle led him to propose a range of ideas and arguments that were to play a major part in the acceptance of new ways of thinking not only in philosophy but also in science.

Bibliography

Frank, Daniel H., and Oliver Leaman. *History of Jewish Philosophy*. London: Routledge, 1997.

Kellner, Menachem. *Dogma in Medieval Jewish Thought: From Maimonides to Abravanel*. New York: Oxford University Press, 1986.

WORKS BY CRESCAS

"Or Adonai (Light of the Lord)." In *With Perfect Faith: The Foundations of Jewish Belief*, edited by J. David Bleich. New York: Ktav, 1983.

The Refutation of Christian Principles by Hasdai Crescas. Translated by Daniel J. Lasker. Albany: SUNY Press, 1992.

WORKS ON CRESCAS

Robinson, James T. "Hasdai Crescas and Anti-Aristotelianism." In *Cambridge Companion to Medieval Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman. New York: Cambridge University Press, 2003.

Oliver Leaman (2005)

CRITERIOLOGY

"Science of criteria" or "criteriology" is a term, originally neoscholastic, for a theory of knowledge in which judgments are warranted or justified simply by conforming to certain criteria for correct judgment. These criteria are general principles that specify what sorts of considerations ultimately confer warrant on some judgments and that tend (tacitly) to guide self-reflective persons in checking and correcting their judgments. The epistemologist's task is to formulate these principles by reflecting on the considerations present and absent in various judgments we intuitively think of as warranted and unwarranted.

Different criteria may deal with different subject matters, degrees, and sources of warrant (e.g., in perception, memory, inference). Ultimately, there must be warranting considerations other than inferability from other warranted judgments. These must be internally accessible through introspection or reflection without relying on further warranted judgments. They will not be considerations such as whether nature designed us to be reliable judges but ones such as whether we ostensibly see or recall something or intuitively grasp or clearly and distinctly conceive something.

Many epistemologists argue that critical considerations need not guarantee truth or confer certainty, and whatever warrant they confer may be defeated. For instance, if one ostensibly sees something red, one is prima facie or defeasibly warranted in judging that one actually sees something red. The judgment might not be warranted when, despite ostensibly seeing something red, one has evidence that the illumination makes everything look red. We need additional principles specifying what considerations defeat warrant.

However, if criterial considerations do not guarantee truth, what makes a set of principles genuinely warranting? Putative common contingent features such as their overall reliability rest warrant on something beyond mere conformity to these principles and may allow for alternative principles. Criteriologists (e.g., Pollock 1974, 1986) often appeal to controversial, nonscholastic, views about concepts and truth influenced by Ludwig Wittgenstein. Criteria are internalized norms (rules) about when to make and correct judgments ascribing a concept. They characterize what persons must, in order to have a particular concept, tacitly know how to do in their judging and reasoning and be tacitly guided by. Criteria individuate our concepts and thus are necessarily correct. Although warranted judgments need not be true, we have no idea of their truth completely divorced from what undefeated criterial considerations warrant. Critics often respond: Surely this norm conformity must have a purpose beyond itself, like accurately representing the world?

See also Epistemology; Wittgenstein, Ludwig Josef Johann.

Bibliography

CRITERIOLOGIES

Coffey, P. *Epistemology or the Theory of Knowledge: An Introduction to General Metaphysics.* 2 vols. London: Longmans Green, 1917.

Pollock, John. *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman & Littlefield, 1986; 2nd ed. (with J. Cruz), 1999.

Pollock, John. *Knowledge and Justification*. Princeton, NJ: Princeton University Press, 1974.

CRITICAL DISCUSSIONS

Lycan, W. G. "Non-inductive Evidence: Recent Work on Wittgenstein's 'Criteria.'" *American Philosophical Quarterly* 8 (1971): 109–125.

Millikan, R. "Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox." *Philosophical Review* 99 (1990): 323–353.

Plantinga, A. *Warrant: The Current Debate*. New York: Oxford University Press, 1993.

Wright, C. J. G. "Second Thoughts about Criteria." *Synthese* 58 (1984): 383–405.

Bruce Hunter (1996)
Bibliography updated by Benjamin Fiedor (2005)

CRITICAL PHILOSOPHY

See *Kant, Immanuel; Neo-Kantiansim*

CRITICAL REALISM

Critical Realism is the title of a book by Roy W. Sellars published in 1916. The name was adopted by a group of philosophers who shared many of his views on the theory of knowledge. *Essays in Critical Realism: A Cooperative Study of the Problem of Knowledge* by Durant Drake, A. O. Lovejoy, J. B. Pratt, A. K. Rogers, George Santayana, C. A. Strong, and Sellars was published in 1920.

BACKGROUND

Much of the epistemological debate since the seventeenth century stems from the matter-mind dualism of René Descartes, who argued that what we know first and most surely is not a physical world but the existence of our own minds, and of John Locke, who argued that we are immediately acquainted only with our own ideas. Starting from these assumptions, how can one know a physical world external to the mind, if, indeed, such a world exists at all? Critical Realism is a chapter in this long debate. Some philosophers, finding it impossible to bridge the gap from a mental world to a material reality that transcends it, turned to some form of subjectivism or idealism; at the beginning of the twentieth century the dominant philosophy in Britain was the Neo-Hegelian idealism of F. H. Bradley and Bernard Bosanquet, and in America it was the voluntarism of Josiah Royce, the personalism of George H. Howison and Borden Parker Bowne, or the pragmatism of William James. But idealism, uncongenial to common sense and to ordinary interpretations of physical science, was followed by a reaction. Scientific knowledge seemed to support philosophical realism rather than idealism.

Shortly before the emergence of Critical Realism a group of philosophers, calling their view the New Realism, argued that even if it is true that whenever something is being perceived, it is an object for a mind, it does not follow that it has no existence except by being perceived. Hence, the idealist commits a fallacy if he concludes that the whole world is nothing but ideas from the truism that when something is known, it is an object for a mind. The American new realists, then—and here they could claim the support of such important British thinkers as G. E. Moore and Bertrand Russell—maintained that elements in perception can at the same time be elements in the physical world. Things do not cause ideas in us, as Locke would have said, so that we first know only ideas and then try to infer from them the nature of the real world which is never directly perceived. Rather, knowing is more akin to selecting, or throwing a light upon, aspects or parts of a world already there to be selected or illuminated by the light of consciousness.

THE CRITICAL REALIST POSITION

The critical realist agrees with the new realist in holding that there is an objective physical world; their disagreement is chiefly on the question of the relation of the datum of knowledge to its object. Physical things, or parts of them, cannot be directly presented to us in perception. Considering the great variety of what is perceived—the double image, the partially submerged bent stick, the toe that is felt after the leg has been amputated—under various conditions by both normal persons and those who are, for example, inebriated or color-blind, are we to say that the real world actually contains all that is disclosed in all these circumstances? And is there no such thing as error? The trouble is that the "direct" realist, by identifying the immediate data of knowledge with elements of the physical world, is trying to account for the universe with an insufficient number of categories or kinds of entities. The knower, whether he is conceived as an organism and a part of nature or as a mind, does not "take in" the physical world. According to Santayana, the datum is an essence, a Platonic universal, which has an identity by being just the character it is, whether it characterizes one or many things in nature or characterizes no existent whatsoever. The datum, the immediately intuited evidence of reality, cannot be numerically identical with any part of that reality.

It is on this epistemological point that the critical realist opposes both the idealist and the direct realist. Whatever exists and whatever its character may be, no datum, or essence, given in experience exists, at least not in the sense in which we say that the objects of perception exist. As Santayana says in *Scepticism and Animal Faith*, "Existence is never given." When the astronomer talks about the moon, he does not mean by "the moon" the yellow disk image that may come to your mind; no doubt, a different image will come to the mind of your companion. If both of you understand the astronomer, "the moon" will mean to both of you, and to the astronomer, the same object to which your thoughts or perceptions are referred—namely, the distant satellite of Earth to which you ascribe certain physical properties. The words or images are the symbols of a meaning, but the essence of the word or image is not, in general, the essence of its meaning. The essence of the meaning is intended to be, but in cases of error will not be, the essence of the actual moon in the sky. This distinction, perhaps difference, between the nature of an image or sense datum and the nature of the object known by means of it is still more obvious when we consider feelings instead of visual images. When sympathizing with a person who has a toothache, we do not say, "I feel the way you feel"; we say, "I know how you feel." Knowing about another person's toothache is not having a toothache.

In perception, as distinguished from thought or conception, there is a tendency to identify image with meaning, so that an effort of analysis is required to separate image, meaning, and object. Paradoxically, the meaning is often psychologically prior to the image. For example, we may perceive a penny as round and as "out there" before noting that in the given perspective it presents an elliptical image. We can then analyze the situation into the image (elliptical), the meaning (round), and the belief that a round object was out there. Error is possible because there may be no object having the same essence as that contemplated in the meaning we have given to the elliptical image that was presented to us. A resolute skeptic who doubts all existence cannot be proved to be mistaken, but if he is consistent, he should be as skeptical of the existence of other minds and even of his own living self as he is of a physical world. Since the idea of change is no guarantee of actual change or process, he should arrive at an inarticulate solipsism of the present moment.

Yet there is no doubt that philosophers as well as laypeople normally believe in the existence of themselves and other minds and ascribe at least some of the characters they intuit to things that exist in space at present or past times. In memory and in the belief in history, the referent of present thoughts is a world of things and events believed to be existentially real and independent of any intuitions, present or past. An actual past or future is not given in any datum, but when one speaks, as David Hume did, of having or of being a succession of perceptions, one posits the existence of a temporal series of events and thereby instantiates in existence one or more essences. To ascribe existence to an essence as such would be a logical or categorial error; it would equally be a logical error to assert that an essence had been intuited by some mind or that some event or perception had occurred and at the same time deny that there is any factual temporal existence. The ontological status proper to essences is timeless subsistence. Actual intuitions come to exist on particular occasions, but knowledge of what they mean, says Santayana, "involves a leap of faith and action from the symbol actually given in sense or thought, to some ulterior existing object."

In *Essays in Critical Realism* Santayana argued that a child reaching for the moon is in quest of an object deployed in a physical world along with the outstretched arm and other bodies. If the moon did not transcend experience, if what is experienced were itself the object striven for, it would already be attained, and there would be no biological need to employ the presently intuited essence as a symbol for an existence still to be reached. There would be no knowledge about anything nor any need for it. If there is any validity in our scientific and commonsense beliefs, our intuitions are engendered in a biological organism by a natural environment. Matter in flux embodies now one essence, now another, and the set of propositions that describes all that exists at all times constitutes the realm of truth. Truth is therefore that part of the realm of essence that happens to characterize existence, and to have knowledge is to believe what is true.

But believing a proposition does not guarantee its truth, beyond the truth of the fact that it is believed. The terms of our beliefs are, in general, symbolic rather than literal representations of nature. Does it make any difference, then, if we clothe nature with intuited essences that are more fanciful than true as long as they are signals for successful action? If a pragmatist at this point suggests that truth means no more than the verification in later experience of the anticipated result of action guided by the earlier experience, the realist cannot agree. The pragmatist does agree with the realist that in knowing there is a reference beyond the immediate having of perceptions, but for the pragmatist the consummation of knowing, the successful working of an idea, does not go beyond

experience; the referent of an idea is another experience. This avoids the problems of a mind-matter dualism and avoids the unanswerable question: How can we know when our ideas correctly represent external things? But the realist sees the pragmatist position as a reversion to idealism and subjectivism and will have none of it. If the pragmatist, to escape idealism, speaks in naturalistic terms, he admits all that the realist asks for. Lovejoy quotes William James: "Practically our minds meet in a world of objects which they share in common" (*Essays in Radical Empiricism*, New York, 1912, p. 79) to show that the practical man, going about his business of solving problems, must assume the existence of an external world; it is important that he discover what its properties were antecedent to and independent of the inquiry in which that discovery is made. Phenomenalism and positivism, sharing with pragmatism the view that the referent of all that can be meaningfully said about real existence must be, in principle, capable of being found in direct experience, are likewise rejected by the critical realist.

How, then, is knowledge of an external reality possible? The critical realist maintains that Locke erred in taking his own ideas to be the objects of knowledge. Knowledge, Locke said, is nothing but the perception of the agreement or disagreement of our ideas. When he comes to a discussion of "knowledge of real existence," however, he is forced to abandon his own definition, and true knowledge becomes the correspondence between ideas and external things. The critical realist argues that Locke should have recognized that when ideas are used in knowing, as distinguished from being merely entertained or had as an experience, there is always reference to an object other than themselves. But merely insisting that data have a referent beyond themselves does not tell us why we should believe one interpretation of them to be a truer description of the facts than any other interpretation. In his more skeptical mood Santayana tells us that knowledge is only faith mediated by symbols, yet in *The Realm of Matter* he sets forth what he takes to be the "indispensable properties of substance." Presumably, he means literally true properties. Substance has parts external to one another and, being in flux and unequally distributed, constitutes a spatial and temporal field of action. These are very nearly those primary qualities that Locke had said resemble the ideas the mind has of them, and if the critical realist seems to have a better case for his position than Locke had, it is chiefly because the sciences have supplied us with a detailed account of the mechanism of perception.

The scientist finds by actual experiment that the date of emission of light from the star, the distorting intervening media between the star and the observer, and the physiological peculiarities of the observer's body all condition what turns up at what time in the experience of the observer. But this scientific account cannot be used by the critical realist to support his position without begging the question. What is proved is that whenever something is found in our world, we can also find something else related to it; scientific knowledge consists of finding what is related to what. This supports the critical realist's thesis that experience depends upon a reality outside all possible experience only if it is assumed from the outset that the experimental data used by the astronomer and by the physiologist are experienced effects of a physical star and a physical organism. The scientist could interpret his explanatory theories on idealist or pragmatist, instead of on critical realist, assumptions. Hence, it is not what the scientist finds, but the epistemology he happens to assume, that supports critical realism. The best that can be said for this realist assumption is that it may be the most economical way to predict and control our experiences and that it may even be the truth about reality.

DIFFERENCES AMONG CRITICAL REALISTS

Some of the critical realists, including Sellars, believe that their position is not best interpreted, even by some within their own camp, when a curtain of essences, ideas, or sense data is drawn between the perceiver and the objects he wants to know by means of such data. For in that case, as in Locke's representative perception theory, the essences or ideas are themselves the only possible objects of knowledge. Sellars would escape this difficulty by what he believes to be a more adequate account of perceiving. When a biological organism has sensations—that is, is affected by an object in the environment with which it must come to terms—the sensation functions as information about the object that caused it. Perception is a response; it is an act of taking the sensation as the appearance of the external object. It is not the sensation or the sense datum that appears; it is by means of the sensation that the object appears. A sophisticated analyst might make the qualities of the sensation the object of his study, but then he is no longer using them to decipher things.

Sellars finds an ally in the English philosopher Gilbert Ryle, who follows common sense in the belief that we perceive trees and hands, not sense data. Here it would seem that Sellars has left the critical realists to join the direct realists, but he would insist that he is not taking

either a direct presentational view or a Lockean representational view. The mediating role of sensation, which determines how the object will look, is not to be ignored. We look *with* our sensations but not *at* them.

In addition to some differences about the role of essences and of sense data, the critical realists are not all in accord on questions of metaphysics. Sellars and Santayana could be called metaphysical monists because for them only one kind of substance—matter—exists. The psyche of which Santayana speaks is the conscious material organism. Sellars thinks of the so-called mental functions not as being carried on by a substantial mind but as ways in which biological organisms, after a long evolutionary development, have learned to respond to stimuli.

Lovejoy, on the other hand, maintains that only a psychophysical dualism is a tenable corollary of an epistemological dualism; only a mind could have sensations and thoughts and intend or mean objects by them.

There has, then, been considerable divergence in the views of thinkers who were, and many who still are, called critical realists. Some have drawn closer to the positions of the direct realists in America or in Britain, and it may be that the label will cease to characterize a definite epistemology.

Bibliography

Boyd, Richard. "Scientific Realism and Naturalistic Epistemology." In *PSA 1980: Proceedings of the 1980 Biennial Meeting of the Philosophy of Science Association.* Vol. 2, edited by Peter D. Asquith and Ronald N. Giere, 613–662. East Lansing, MI: Philosophy of Science Association, 1982.

Chisholm, Roderick. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.

Churchland, Paul M. *Scientific Realism and the Plasticity of Mind*. Cambridge, U.K.: Cambridge University Press, 1979.

Drake, Durant. *Mind and Its Place in Nature*. New York: Macmillan, 1925.

Drake, Durant et al. *Essays in Critical Realism: A Cooperative Study of the Problem of Knowledge*. London, 1920.

Dretske, Fred. *Seeing and Knowing*. London: Routledge and Kegan Paul, 1969.

Goldman, Alan. *Empirical Knowledge*. Berkeley: University of California Press, 1989.

Haack, Susan. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Oxford: Blackwell, 1993.

Hanley, Richard. "Much Ado about Nothing: Critical Realism Examined." *Philosophical Studies* 115 (2) (2003): 123–147.

Kulp, Christopher, ed. *Realism/Antirealism and Epistemology*. Lanham, MD: Rowman & Littlefield, 1997.

Lovejoy, A. O. *The Revolt against Dualism*. New York: Norton, 1930.

Moser, Paul K. *Knowledge and Evidence*. New York: Cambridge Unversity Press, 1989.

Murphy, A. E. "The Fruits of Critical Realism." *Journal of Philosophy* 34 (1937): 281–292.

Pratt, J. B. *Personal Realism*. New York: Macmillan, 1937.

Quinton, Anthony. *The Nature of Things*. London: Routledge and Kegan Paul, 1977.

Santayana, George. *Realms of Being.* 4 vols. New York: Scribners, 1927–1940.

Santayana, George. *Scepticism and Animal Faith*. New York: Scribners, 1923.

Schilpp, Paul A., ed. *The Philosophy of George Santayana*. Evanston, IL: Northwestern University, 1940.

Sellars, Roy W. "American Critical Realism and British Theories of Sense Perception." *Methodos* 14 (55–56) (1962): 61–108.

Sellars, Roy W. *Critical Realism*. Chicago: Rand, McNally 1916.

Sellars, Roy W. *The Philosophy of Physical Realism*. New York: Macmillan, 1932.

Sellars, Wilfrid. *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Press, 1991.

A. G. Ramsperger (1967)
Bibliography updated by Benjamin Fiedor (2005)

CRITICAL THEORY

"Critical theory" is used to refer to the diverse body of work produced by members and associates of the Frankfurt Institute for Social Research after Max Horkheimer became its director in 1930. The first generation of what came to be called the Frankfurt school included, in addition to Horkheimer, such prominent figures as Theodor Adorno, Herbert Marcuse, Walter Benjamin, Erich Fromm, Leo Löwenthal, Franz Neumann, Otto Kirchheimer, and Frederick Pollock. The most influential members of the second generation are Jürgen Habermas, Karl-Otto Apel, and Albrecht Wellmer. As the variety of backgrounds and interests might suggest, critical social theory was conceived as a multidisciplinary program linking philosophy to history and the human sciences in a kind of "philosophically oriented social inquiry," as Horkheimer put it. Though very strongly influenced by Immanuel Kant and neo-Kantianism, Georg Wilhelm Friedrich Hegel and German idealism, Max Weber and Sigmund Freud, it was understood as a renewal of Marxism inspired in part by the earlier work of Georg Lukács and Karl Korsch. This updated Marxism would take account of the altered historical realities of advanced capitalism and integrate areas of inquiry neglected by traditional Marxism, such as philosophy and political theory, cultural studies (including studies of mass culture), and social psychology (appropriating psychoanalysis for social theory). With the rise of National Socialism, the institute moved briefly to Geneva and Paris in 1933 and

then in 1934 to Columbia University in New York, where its journal, the *Zeitschrift für Sozialforschung,* continued to be published until 1941, the last volume in English. Early in the 1950s, Horkheimer and Adorno reestablished the institute in Frankfurt. Habermas became an assistant there in 1955.

The original project of a critical social theory advanced by Horkheimer was a version of Karl Marx's *Aufhebung* of philosophy in social theory and practice. Philosophy was to become a sociohistorical, practically oriented critique of reason and its claimed realizations. While the dominant forms of reason were often distorted in the interests of dominant classes, the aim of critical theory was, not simply to negate them, but, by examining their genesis and functions, to transform them and enlist them in the struggle for a better world. The insistence on the "truth content" of the "bourgeois ideals" of freedom, truth, and justice, the refusal to abandon them as mere ideology, was severely tested by the horrors of World War II. Early in the 1940s, in their collaborative reflections on the "dialectic of enlightenment," Horkheimer and Adorno offered a much more pessimistic view of the history of reason. Keying on a tendency that Weber had emphasized, the relentless spread of "instrumental" rationality, they revered Marx's positive evaluation of scientific-technological progress. It was now seen as the core of a domination that had spread to all spheres of life and, in the process, had immobilized the potential agents of social change. In this "totally administered society" with what Marcuse later called its "one-dimensional man," critical theory could at best reveal the unreason at the heart of what passed for reason, without offering any positive account of its own.

Habermas's work since the 1960s might be viewed as an attempt to avoid this impasse by introducing into critical theory a fundamental shift in paradigms from the philosophy of the subject to the theory of communication and from means-ends rationality to communicative rationality. This serves as the basis for an altered diagnosis of the ills of modernity—as rooted, not in rationalization as such, but in a one-sided rationalization driven by economic and administrative forces—and an altered prescription for their cure, the democratization of public opinion and will formation in an effectively functioning public sphere, where issues of general concern are submitted to rational, critical public debate.

See also Adorno, Theodor; Apel, Karl-Otto; Benjamin, Walter; Freud, Sigmund; Habermas, Jürgen; Hegel, Georg Wilhelm Friedrich; Horkheimer, Max; Kant, Immanuel; Lukács, Georg; Marxist Philosophy; Marx, Karl; Neo-Kantianism; Weber, Max.

Bibliography

Arato, A., and E. Gebhardt, eds. *The Essential Frankfurt School Reader.* New York: Urizen Books, 1978.

Bronner, S., and D. Kellner, eds. *Critical Theory and Society: A Reader.* New York: Routledge, 1989.

Ingram, D., and J. Simon-Ingram, eds. *Critical Theory: The Essential Readings.* New York: Paragon House, 1992.

Jay, M. *The Dialectical Imagination.* Boston: Little Brown, 1973.

Wiggershaus, R. *The Frankfurt School.* Translated by M. Robertson. Cambridge, MA: MIT Press, 1994.

Thomas McCarthy (1996)

CROCE, BENEDETTO
(1866–1952)

Benedetto Croce was the best-known Italian philosopher of the twentieth century. His universally and justly celebrated book on aesthetics, *Estetica come scienza dell'espressione e linguistica generale* (1902), which became the first volume of his systematic "philosophy of the spirit," was a foundation stone in the great revival of historical idealism in Italy between 1900 and 1920. In a long and diligent life devoted almost entirely to scholarly studies, Croce gained an international reputation in the fields of aesthetics, literary criticism, cultural history, and historical methodology; and he exercised in these areas an influence so pervasive that it cannot yet be definitively estimated.

LIFE AND WORKS

Born at Pescasseroli, in the Abruzzi, of a family of wealthy landowners, Croce never needed to earn a living. He displayed an early bent for literary and historical research but never seriously entered on an academic career, preferring to be master of his own course of study. From 1883—when his parents were killed, and he himself buried and injured, in an earthquake—until 1886 he lived with his uncle Silvio Spaventa (brother of the philosopher Bertrando) in Rome, and for a time he attended the university there. At the university he came under the influence of Antonio Labriola, who led him to the study of Johann Friedrich Herbart and, later, of Karl Marx. These studies left a lasting mark on his philosophy. After 1886 he lived permanently in Naples.

In 1893 Croce published his first philosophical essay, "La storia ridotto sotto il concetto generale dell'arte"

(History brought under the general concept of art), a title that foreshadowed the main concerns of his mature thought. In 1898, while working on a book on Marx (*Materialismo storico ed economia marxista*), he entered into correspondence with his younger contemporary, Giovanni Gentile, who was similarly occupied. Thus began a friendly collaboration that lasted twenty-five years. In 1900 came the first sketch of Croce's *Aesthetic.* In 1903 he founded the journal *La critica,* and in 1904 he became an editorial adviser to the publishers Laterza of Bari. For the rest of his life he exercised an ever-increasing influence on the literary and academic world through these two channels.

Even as the volumes of Croce's philosophy of spirit were being published, his association with Gentile was leading him to a reexamination of G. W. F. Hegel. He published his results in 1907 (*Ciò che è vivo e ciò che è morto nella filosofia di Hegel*) and made appropriate revisions in his *Estetica* and in his *Logica come scienza del concetto puro* (1905). *Filosofia della pratica, economia ed etica* appeared at Bari in 1909. In 1911 he published *La filosofia di Giambattista Vico*—Giambattista Vico was the other major influence on his thought—and in the succeeding years he wrote the essays that appeared at Bari in 1917 as the culminating volume of his system, *Teoria e storia della storiografia.*

In 1910 Croce was made a life member of the Italian senate, but he was not then actively involved in politics. He was a neutralist prior to Italy's entry into World War I in 1915; and in the postwar crisis, he became minister of public instruction in Giovanni Giolitti's last cabinet (1920–1921). With Gentile's help, Croce drafted a reform of the school system, rejected at the time but later incorporated in the fascist *Riforma Gentile* of 1923–1924. Naturally, therefore, he regarded the first fascist administration with some benevolence. His breach with fascism (and with Gentile) came with the establishment of an overt dictatorship in January 1925. He drafted a celebrated "Protest" against Gentile's "Manifesto of Fascist Intellectuals" and thus became identified as the chief antifascist intellectual, a role he worthily maintained through more than fifteen years of almost complete political isolation and retirement. He emerged briefly in 1929 to speak in the senate against the concordat with the Vatican. After the fall of fascism he became a leader of the revived Liberal Party and served once more as a cabinet minister for a short period in 1944.

During his years of isolation, Croce wrote voluminously and his thought developed significantly. His aesthetics reached its final form only in *La poesia* (1936). His opposition to fascism is often apparent in his literary criticism, but it expressed itself more naturally in his historical writing and in theoretical reflection on politics and history, where it led to vital developments in his thought.

Croce celebrated his eightieth birthday by founding and endowing the Institute for Historical Studies, which is still located in his former home. In spite of a serious stroke in 1950, he went on working right up to his death.

AESTHETICS

When Croce's philosophical interests were first aroused in 1893, he was a historical and literary scholar who accepted most of the assumptions of the French positivism then dominant in the circles in which he moved. But controversy led him to ask himself whether history was an art or a science, and he made a decisive choice in favor of the idealist view of the great Hegelian philosopher of art and literary historian Francesco De Sanctis (1817–1883). Initially, his idealist aesthetics was set in a context of a realistic metaphysics, of which there were still some signs in the *Aesthetic* of 1902; but the attempt to expound his view systematically, combined with his discovery of Vico and rediscovery of Hegel, led to the development of his full-fledged idealism. Thus, his aesthetic theory was the original foundation of his philosophy of spirit, although it might fairly be argued that the theory of moral judgment became more fundamental in the final form of his system. Croce himself distinguished four phases in his reflections on aesthetics. Some critics have held one or more of the later phases to be inconsistent with his system as a whole, but they will here be viewed as part of a continuous and essentially consistent evolution.

AESTHETIC INTUITION. It is characteristic of idealist aesthetics to regard aesthetic experience as a kind of cognition. Following Vico and De Sanctis, Croce regarded it as the primitive form of cognitive experience. Intuition is a nonconceptual form of knowledge; it is the awareness of a particular image either of outward sense (a person or a thing) or of inner sense (an emotion or a mood). Intuitions possess a kind of ideal being and validity that is independent of and ontologically prior to any question of existence or nonexistence. Croce's use of the term *intuition* derived directly from Immanuel Kant's use of *Anschauung,* and he originally thought of the external world as a Kantian manifold of sensation, which we organize into distinct perceptions through the intuitive faculty of imagination. Thus, history was initially "subsumed under the general concept of Art," as the subform of art that is concerned with the ordering of intuitions of

actual existence. He soon abandoned this position, but if Kant's theory of space and time as the "forms of intuition" through which the sensible manifold is organized is recalled, it can be seen how Croce's view applies to the plastic arts, which he often seemed to ignore. His own background and interests were predominantly literary, and his theory frequently seems specifically devised to meet the needs of literary critics who have to deal with poems, which are uniquely individual entities created in the conceptual (or logically universal) medium of language. Croce himself fostered this illusion by insisting that aesthetics was "the general science of language." This is a very Pickwickian contention on his part, since the conceptual function of words and symbols in factual communication—which must surely be regarded as fundamental in a general theory of *language*—is specifically excluded from his "science of expression"; and all forms of nonconceptual communication—even nonverbal ones—are included in it.

LYRICAL INTUITION. If it had not been for his overriding concern with poetry, Croce might never have advanced to the second phase of his aesthetics, the theory that all intuition is "lyrical" in character. The problem he faced was essentially one of defining what it is that is nonconceptually communicated in poetry by way of language. His answer was that poetry communicates emotions and moods, it expresses for cognitive contemplation different aspects of the practical personality of man. Here the "circle of the spirit," the doctrine that man's theoretical activity has his practical reality as its one and only object, comes into view. By means of this doctrine, Croce was able to dispense with the last residues of naive realism present in his basically Kantian epistemology. Some doctrine of this sort was certainly needed if the view that art is nonconceptual cognition was to be maintained. As Croce said in 1908, in his lecture announcing the doctrine, "An image that does not express a state of mind has no theoretical value." But the need might well have appeared less pressing, and the solution less natural and obvious, if he had not always thought primarily about poetry.

It is easy—especially if one reads only the *Breviary of Aesthetics* (1912), as many English-speaking students do—to misinterpret Croce's theory that all art is lyrical as a type of romanticism, which he was, in fact, absolutely opposed to. His doctrine was that art is the expression of emotion, not just for its own sake but as a special kind of cognitive awareness. He was seeking a middle way between the intellectualist errors of classical theorists, with their artificial canons, rules, and genres (all of which he categorically rejected), and the emotional excesses of the romantics, with their glorification of immediate feeling. His critique of classical intellectualism is easily grasped; but it is a mistake to think, as some critics have, that his "lyricism" is radically inconsistent with his own systematic rationalism. Unlike Gentile, Croce always refused to identify intuitions as "feelings" or to formulate his theory in terms of "feeling" at all, because he held that "feeling" was an ambiguous concept which when clarified referred to the practical impulse that is the content of intuition.

COSMIC INTUITIONS. How can the expression of emotion produce cognitive awareness? This was the problem that Croce faced in the third phase of his thought—his theory that all intuition has a "cosmic" aspect. Again, some doctrine of the sort was required by his basic thesis that intuition is cognitive of particulars without reference to their existential status. Simply as images they provide experience of the universal human spirit. This self-validating character, this reference to universal humanity (not as an abstract nature or essence but as the activity of the spirit revealing itself in personal experience and in history as a whole) is what Croce called the cosmic aspect of genuine intuition. Some intuitions, however, are more directly cosmic than others and are hard to characterize in terms of specific emotions; this was the classical counterweight against lyrical romanticism in Croce's thought. It was apparently suggested to him by an essay of Wilhelm von Humboldt on Johann Wolfgang von Goethe, and he applied it in critical studies of such masterpieces as *Faust* and the *Divine Comedy.* Oddly enough, however, it was neither Dante Alighieri nor Goethe, but Ariosto, who served as Croce's paradigm of the cosmic poet. Croce earnestly desired to avoid confusion between the proper lyrical unity of a poem and the logical coherence of a philosophical system. His own critical practice even provides some justification for the view that the whole cosmic phase of his theory was an aberration. The truth is rather that it was an inescapably necessary complement of his general view and that his critical practice suffered from an antiphilosophical bias.

LITERATURE AND ART. The final phase of Croce's aesthetic theory is the theory of literature in *La poesia,* which forms the negative corollary of his theory of intuition. Much that is ordinarily classified as art was, in Croce's view, not properly art at all because in it the purity of intuitive cognition is subordinated to various practical ends, such as entertainment or intellectual and moral instruction. For instance, he declared the *De Rerum*

Natura to be a work of literature, not of art; and although this is an extreme case of his critical bias, it is easy to see what led him to it, since the passionate conviction and practical aim of Lucretius are evident in every line of the poem.

LOGIC OF HISTORY AND THE SCIENCES

As aesthetics is the science of pure intuition, so logic is the science of pure concept; and as pure intuition is the form in which we imaginatively express some particular aspect of the human spirit, so pure conception is the form in which we rationally evaluate these particular manifestations and relate them to one another and to the spirit as a systematic unity. Thus on the one hand, conceptual cognition presupposes intuition because it requires intuitions as its material; and on the other hand, aesthetics, the science of intuitive cognition, is only a subdivision of logic because beauty is a form of the pure concept. Concepts presuppose intuitions but are not derivable from them; and any evaluation or correlation of intuitions—even the categorizing of them as intuitions—presupposes concepts. This is the "dialectic of the distincts," which Croce insisted was more ultimate and fundamental than the Hegelian dialectic of opposites. His model here was Kant, rather than the often-cited Vico. For Vico, as for Hegel, poetic cognition was already an immature form of reason, or, in other words, reason develops out of it; whereas for Croce, as for Kant, the two functions were quite distinct and interdependent, although not equally primitive. Croce's aesthetics was a new transcendental analytic, and his logic was a new deduction of the categories.

KNOWLEDGE. For Croce, however, the words *reason* and *knowledge* meant something very different from what they meant to Kant. Croce's work was a "critique of historical reason," and the knowledge that he regarded as genuine was historical knowledge. It is only to historical judgments that the predicates "true" and "false" are properly applicable. According to Croce, the scientific knowledge of Kant's *Critique* was a myth, and belief in this myth was one type of logical error. (Croce offered an exhaustive analysis of the types of logical error as a sort of negative proof of his own deduction.) Science and scientific investigation are forms of practical activity, not of cognition. They cannot be genuinely cognitive because they are founded on pseudo concepts, not on the genuine forms of the pure concept.

Thus, for example, if a child reports that "the cat is on the mat," this is a statement of historical fact and its truth or falsity can be established. But if a scientist says, "The cat is a mammal with such and such properties," the words *cat* and *mammal,* together with all the property-terms, are abstract universals, artificial summaries of actual aesthetic and historical experience. These abstractions are enormously useful in practical experience—indeed, they are vital to the intelligent planning of our lives—but they could only be the basis of genuine knowledge if we were endowed with the kind of rational intuition into the "real essences" of things that is described in Plato's myths.

The forms of the pure concept are the distinct forms of the spirit itself, since only a proof that some form of the spirit is "distinct" in Croce's sense could establish the a priori validity of a proposed category or standard of judgment. There are four such forms and, hence, four ways in which our experience can be cognitively categorized and evaluated. Any proper element of experience can be considered from two theoretical and two practical points of view; it can be evaluated intuitively, rationally, economically, or morally.

ERROR. In his theory of error, Croce followed René Descartes and Antonio Rosmini. He regarded all genuine error as caused by the intrusion of practical motives into theoretical contexts. He was primarily concerned with philosophical errors, such as the belief that science is knowledge or the belief in myth (a historical narrative possessing absolute significance), which he took to be the origin of religion. About mistakes in historical interpretation, his view appears to have been that (if the historian advances his hypotheses in a properly tentative spirit) they are not really errors but stages in the development of truth.

PHILOSOPHY AND HISTORY. Under the influence of Gentile, Croce accepted the Hegelian identification of philosophy with the history of philosophy and reduced even the a priori judgments of his own logic (for example, that there are four forms of the spirit) to the status of historical judgments. He did this because he held that no one could "close the gates of truth" against further progress. Yet he never accepted Gentile's view that this formal concession to the future meant that all deductions of "the forms of the spirit" were mistaken; he remained convinced that his logic possessed an eternal validity. In his view, the unity of philosophy and history was a unity of distincts.

ECONOMICS AND LAW

The most fundamental of all distinctions in Croce's philosophy is the distinction between theory and practice. Goaded by the actual idealists who sought to unify theory and practice in the "pure act," Croce tried to justify this distinction by arguments that were largely wasted, because his opponents did not really deny the distinction any more than he denied the unity. The only point at issue was the more general question of whether the unity arose from a dialectic of opposed moments or of distinct forms.

ECONOMIC UTILITY AND VITALITY. It has already been shown how the circle of the spirit first appeared when Croce recognized practical impulse as the presupposed content of intuition. It would seem to follow that the practical manifestation of the spirit is somehow more primitive than the spirit's theoretical functions; but the implication is, at best, only a partial truth, for Croce claimed also that the primitive form of practical activity—economic volition—presupposes both forms of theoretical activity. He had learned from his long study of Marx and of the English classical economists that the calculation of economic utility is a rational process and that economic action involves historical judgment. The practical impulse that intuition presupposes, considered in itself, is not yet the conscious action of the spirit; it is only the blind urge of organic life out of which the spirit emerges. But the origin of volition in vitality is what accounts for the independence that Croce always ascribed to economic utility as a distinct spiritual form. Critics objected from the beginning that there was a paradox involved in treating utility as an autonomous form of value. There is no such thing as simple usefulness; there is only usefulness for some purpose. It is really life or vitality that is the primitive category of action. In later writings Croce recognized this, but he continued to hold that economic action is the first form of action in the true sense.

In spite of Croce's insistence that the "utility" of the economists is a fundamental philosophical category, his logic does not allow the admission of economics itself as a genuine philosophical science. The work of economists, like that of all other scientists, belongs to the category of utility itself, not to that of truth. "Economic man" is a paradigm case of a pseudo concept.

LAW AND UTILITY. It is more surprising, perhaps, to find the concept of law subsumed under that of utility in Croce's system. The Kantian model, which we have appealed to several times, might lead us to expect moral law as the universal form of practical consciousness. Law in fact functions as a transitional notion in Croce's system because it may be obeyed either from motives of duty or from motives of expediency. Croce held, however, that in the making and execution of law we should be guided strictly by considerations of social utility, since no one can make a genuinely moral judgment about what is right for a whole class of cases defined abstractly. Laws are of necessity framed in terms of the pseudo concepts of economics and social science; even the moral habits and rules we adopt as our own guides are similarly abstract. They belong among the instruments of life, not among its purposes. Because so much of the work of government is also instrumental, Croce tried at first to formulate a purely economic theory of political action in general. This view he subsequently abandoned.

ETHICS AND POLITICS

Moral action and moral judgment are the distinct universal forms of practical consciousness corresponding to economic action and economic rationality. The dialectic of the distinction is closely analogous to that of the two theoretical forms. There can be economic acts that are not moral (for example, historical explanation of an *im*moral act is bound to be at the economic level); but there cannot be moral acts that have no vital utility (asceticism or abstract moralism is a moral error). On the other hand, practical activity cannot concretely achieve rationality at the economic level without superseding that level. There can be no theory of economic life except from an independent ethical point of view. This is shown by the inconsistency of utilitarian ethics, which attempts to justify individual self-sacrifice by smuggling in moral principles that have not themselves been accounted for. Confined strictly to the economic level, rational people would live in the Hobbesian state of nature, and all the consequences of Hobbesian philosophy would follow.

Moral, as distinct from economic, consciousness is the awareness of some definite act as a duty overriding private inclinations. Moral judgment declares the act to be a duty because it embodies some universal spiritual value (which may fall under the category of beauty, truth, or social utility or be a distinctively moral good). Whatever category the value belongs to, if the act is a moral duty, there is always a sense of "harmony with the Universe." The moral point of view is the final all-embracing awareness of the spirit as a whole, in its wholeness; hence this is the point of view from which true history can be written.

FREEDOM. Because he held that all true judgment is historical, Croce could do little except offer historical illustrations of his view. Reflection on the nature of history itself and on the reason for rejecting scientific concepts as pseudo concepts, however, throws further light on goodness as a distinct category of the spirit. Science fails to be genuine knowledge because the spirit in all its forms always exhibits spontaneity and individual uniqueness. At the moral level, this spontaneity becomes conscious freedom and self-possession. History is "the story of liberty," and freedom is another name for the good as a distinct form of value.

ETHICO-POLITICAL HISTORY. Gentile buttressed an ethical theory similar to Croce's with the Hegelian conception of the national state as an ethical organism and as the bearer of the spirit in history. Croce admitted that if one interpreted the concept "state" broadly enough, this was a legitimate way of viewing it. But he was initially more inclined to think of politics as an economic art or technique of directing selfish passions into orderly channels (as if there were no conflicting moral ideals in political life). The advent of fascism taught him that both of these extreme views were mistaken. Politics does involve moral consciousness, but the absorption of all morality into the "ethical state" is a "governmental concept of morality" unacceptable in a society of free men. The true bearer of the spirit is the individual moral agent, and the state contains the dialectic of practical life as a whole (economics and ethics). The ethical universal is only fully revealed in the history of the state so conceived. Political life, as the unity in which all spiritual activities (even poetry) have a place, is raised to the ethical level in the consciousness of the historian who writes ethico-political history. This is the complete expression of the spirit in which philosophy and history are unified. Croce's work as a historian, particularly in *La storia del regno di Napoli* (History of the Kingdom of Naples; 1925), illustrated how this concept applies to periods of decadence as well as periods of progress.

The "circularity of the spirit" might seem to require that this form of historical consciousness become the content of poetic intuition. But Croce never made this point, and he does not seem to have held this view. The circle of the spirit, as he describes it, closes by returning from vitality to poetry. Ethico-political history transcends the circle altogether because it is the perfected consciousness of the spirit in its circularity.

See also Aesthetics, History of; Aesthetics, Problems of; Dante Alighieri; De Sanctis, Francesco; Descartes, René; Error; Gentile, Giovanni; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Herbart, Johann Friedrich; Humboldt, Wilhelm von; Idealism; Intuition; Kant, Immanuel; Marx, Karl; Rosmini-Serbati, Antonio; Spaventa, Bertrando; Vico, Giambattista.

Bibliography

WORKS BY CROCE

Croce's *Opere complete* as of 1965 consists of 67 volumes.

La storia ridotto sotto il concetto generale dell'arte. Naples, 1893.

Materialismo storico ed economta marxista. Palermo, 1900. Translated by C. M. Meredith as *Historical Materialism and the Economics of Karl Marx.* New York: Macmillan, 1914.

Estetica come scienza dell'espressione e linguistica generale. Bari: Laterza, 1902. Translated by Douglas Ainslie as *Aesthetic.* New York, 1909; 2nd complete ed., 1922.

Logica come scienza del concetto puro. Naples, 1905. Translated by Douglas Ainslie as *Logic.* New York, 1917.

Ciò che è vivo e ciò che è morto nella filosofia di Hegel. Bari, 1907. Translated by Douglas Ainslie as *What Is Living and What Is Dead in the Philosophy of Hegel.* London: Macmillan, 1915.

Filosofia della pratica, economia ed etica. Bari: Laterza, 1909. Translated by Douglas Ainslie as *Philosophy of the Practical.* New York, 1913.

La filosofia di Giambattista Vico. Bari: Laterza, 1911. Translated by R. G. Collingwood as *The Philosophy of Giambattista Vico.* New York: Macmillan, 1913.

Breviario di estetica. Bari: Laterza, 1913. Translated by Douglas Ainslie as *The Essence of Aesthetic.* London: Heinemann, 1921. First published in *The Book of the Opening of the Rice Institute.* Houston, TX, 1912.

Teoria e storia della storiografia. Bari: Laterza, 1917. Translated by Douglas Ainslie as *History—Its Theory and Practice.* New York: Harcourt Brace, 1921.

Contributo alla critica di me stesso. Naples, 1918. Translated by R. G. Collingwood as *Autobiography.* Oxford: Clarendon Press, 1927.

Frammenti di etica. Bari: Laterza, 1918. Translated by Arthur Livingston as *The Conduct of Life.* New York: Harcourt Brace, 1924.

Goethe. Bari: Laterza, 1919. Translated by E. Anderson. London: Methuen, 1923.

Ariosto, Shakespeare, Corneille. Bari: Laterza, 1920. Translated by Douglas Ainslie. New York, 1920.

La poesia di Dante. Bari: Laterza, 1921. Translated by Douglas Ainslie as *The Poetry of Dante.* New York: Holt, 1922.

Elementi di politica. Bari: Laterza, 1922. Translated by Salvatore Castiglione as *Politics and Morals.* London: Allen and Unwin, 1946.

La storia del regno di Napoli. Bari: Laterza, 1925.

Aesthetica in nuce. Bari, 1928. Translated by R. G. Collingwood as "Aesthetics," in *Encyclopaedia Britannica,* 14th ed. New York and London, 1929. Vol. VII.

Storia d'Italia dal 1871 al 1915. Bari: Laterza, 1928. Translated by Cecilia M. Ady as *History of Italy 1871–1915.* Oxford: Clarendon Press, 1929.

Storia d'Europa nel secolo XIX. Bari: Laterza, 1932. Translated by Henry Furst as *History of Europe in the Nineteenth Century.* New York: Harcourt Brace, 1933.

La poesia. Bari: Laterza, 1936.

La storia come pensiero e come azione. Bari: Laterza, 1938. Translated by Sylvia Sprigge as *History as the Story of Liberty.* London: Allen and Unwin, 1941.

My Philosophy. London: Allen and Unwin, 1949. Essays selected by R. Klibansky and translated by E. F. Carritt.

WORKS ON CROCE

Borsari, Silvano. *L'opera di Benedetto Croce.* Naples: Istituto Italiano per gli Studi Storici, 1964.

Caponigri, A. R. *History and Liberty.* London: Routledge and Paul, 1955. A study of the historical works.

Nicolini, Fausto. *Benedetto Croce.* Turin, 1962. A definitive biography.

Orsini, G. N. G. *Benedetto Croce, Philosopher of Art and Literary Critic.* Carbondale: Southern Illinois University Press, 1961. A masterly survey of Croce's aesthetic writings.

Piccoli, R. *Benedetto Croce.* London: J. Cape, 1922. Still the best general introduction to Croce, but it covers only the period from 1893 to 1920.

H. S. Harris (1967)

CRUSIUS, CHRISTIAN AUGUST

(1715?–1775)

Christian August Crusius, the German Pietist philosopher and theologian, was born at Leuna, Saxony. Educated in Leipzig, he was appointed extraordinary professor of philosophy there in 1744, and professor of theology in 1750. Crusius initiated the third wave of Pietist attacks on Wolffianism by a series of dissertations (1739–1745), and continued it in his four main philosophical works (1744–1749). He later turned to theological studies, lost interest in philosophy, and founded a new theological school, the Biblicoprophetic school, which partially diverged from Pietism. He later became *canonicus* at the Meissen Theological Seminary.

Crusius's reputation in his own time and his influence on his contemporaries was second among Pietist philosophers only to Christian Thomasius. The collaboration of his close follower, A. F. Reinhard, with Pierre-Louis Moreau de Maupertuis and the Berlin Academy in their polemics against Wolffianism established a link between Christian Wolff's Pietist and academic opponents. Several later philosophers acknowledged Crusius as their teacher, although they combined a Crusian background with more advanced trends of French and English origin. These thinkers contributed considerably to the renewal of German philosophy after the dissolution of the Wolffian school. In theology Crusius's influence was even stronger.

Crusius's importance was forgotten or suppressed soon after his death, especially among theologians, and has not yet been fully reestablished because of the hostility of the subsequently dominant rationalist and philological schools to the trend of his theology. As a philosopher, Crusius was nearly voted into oblivion, along with most other minor eighteenth-century philosophers, by idealistic historiographers. He was rediscovered by the new philological historiographers, chiefly in connection with his influence on Immanuel Kant.

ORIGIN OF CRUSIUS'S THOUGHT

After 1730, Wolff and his school began to recover from his expulsion from Halle University in 1723, and from the loss by most of his pupils of their professorships, an attack launched for personal and political reasons by his Pietist opponents. The Pietists were gradually deprived of official support and were more and more restricted to theoretical controversy with Wolff. However, Wolff's system of philosophy was a much more modern, comprehensive, and technically refined body of doctrines than those in the obsolete and clumsy treatises of Thomasius, Franz Budde, and Andreas Rüdiger. A far-reaching reform in the doctrine and quality of Pietist philosophy was needed for it to face the Wolffian doctrine and counteract it successfully. Crusius's teacher, A. F. Hoffman (1703–1741), developed the logical doctrines of Thomasius and Rüdiger, taking into account Wolff's new philosophical techniques and achievements and accepting some of his doctrines, in his own *Vernunft-Lehre* (Leipzig, 1737). Crusius's own logic was inspired by Hoffman's refined and comprehensive handbook, whose quality and thoroughness substantially met the most modern requirements. Hoffman's early death prevented him from publishing the treatises on the other branches of philosophy that he had announced in 1734, but Crusius proceeded along Hoffman's lines, both improving and completing his lifework. Crusius provided the Pietist school with a renewed, efficient, and modern theoretical platform that temporarily assured its philosophical survival, outlived orthodox Wolffianism, and led to a far-reaching change in German philosophy.

METHODOLOGY AND LOGIC

Crusius's methodology, the foundation of his philosophical attitude, was based on two central ideas, both originating in the Pietist tradition. Philosophy is not, as it was

for Wolff, a pure "science of possible things insofar as they are possible," but is based on existing things. Second, human understanding has very narrow limits; theoretical certainty is impossible concerning many fundamental points whose only foundation is moral certitude or revelation. The mysteries of religion are not only beyond human reason, as Wolff claimed, but also contradict it. Something may be unthinkable for human reason that is not so for God or in itself.

Crusius held that the most general principle of human knowledge is neither the principle of identity nor the principle of contradiction, but a principle concerning what we can and cannot think: What cannot be thought at all is false, and what cannot be thought of as false is true. Our notions of identity and contradiction are based on this principle, which he called the principle of *cogitabilitas*. It is an inner criterion, depending on the nature of the human understanding.

Crusius further held that human reason cannot reach ultimate truth. Knowledge begins with experience, both inner and outer, and in many cases is stopped in its analysis of an order of facts by certain notions that, although they are not simple in themselves, cannot be further analyzed by man. Even if an analysis is completed and man does reach some simple basic notion, this notion cannot be demonstrated or deduced from a unique source. Each notion must be intuited singly by connecting it with concrete examples.

It is therefore impossible, according to Crusius, to assume that the method of philosophy is identical with the method of mathematics. Mathematics deals with very simple properties of things and its objects are exhaustively defined, whereas many notions relating to objects of philosophical thought can neither be known with intuitive distinctness nor analyzed by man. Again, mathematics proceeds only by demonstration and solely on the basis of the principle of identity. Philosophy, on the other hand, frequently must revert to moral certainty and is based on several different principles and on the knowledge of fact.

The main characteristics of Crusian logic, as expounded in his *Weg zur Gewissheit und Zuverlässigkeit der menschlichen Erkenntnis* (Way to certainty and reliability of human knowledge; Leipzig, 1747), follow from these views. Crusius connected logic with methodology. His logic contained much empirical psychology and many informal concrete and practical rules for obtaining or verifying knowledge, including rules for experimentation. Because Crusius so limited the field of theoretical demonstration, he presented a highly developed logic of probability (which he called moral certitude), covering, among other topics, induction, hypothesis, and the reliability of testimony. The last was essential in the justification of revelation.

Both for Crusius and for Wolff, knowledge derived only from the senses, but the main characteristics of Crusius's methodology allowed his successors to be much more receptive to English and French empiricism, sensationalism, and commonsense philosophy than were orthodox Wolffian rationalists. This receptivity was partially due to John Locke's strong influence on Thomasius, but the ethical and mystical sources of these Pietistic attitudes was most important.

METAPHYSICS

Crusius, in his *Entwurf der nothwendigen Vernunftwahrheiten* (Sketch of necessary rational truths; Leipzig, 1745), divided metaphysics into ontology, theology, cosmology, and pneumatology, in explicit opposition to Wolff's ordering of the metaphysical sciences.

ONTOLOGY. Ontology begins, not with first principles, but with the notion of a thing in general, directly connected with the notion of a "really given thing." Only after introducing these notions did Crusius discuss essence, existence, and causality. Crusius regarded existence as indefinable and as a primary notion arising from sensation.

In his discussion of causality, Crusius expounded a principle of determining reason, his version of Gottfried Wilhelm Leibniz's principle of sufficient reason. Crusius held, against Wolff, that a sufficient reason suffices only for free actions insofar as they are free. Rational truths and natural events not depending on free causes need a more cogent foundation, a determining reason. This principle does not derive from the principle of identity, but rather from what we must conceive or what we cannot conceive as united or separate, and thus from a new case of the principle of *cogitabilitas*. Crusius, aiming at a sharper distinction between mechanism and free actions, held that the real nature of causality is unknown and that our knowledge of causal connections is based on the constant conjunction of two events in experience. This, of course, cleared the path for the members of his school to accept the Humean critique of the causal connection.

Crusius's ontology reveals a general characteristic of his metaphysics. His was not a monolithic system beginning with a single principle and deducing from it all subsequent notions and propositions, as was Wolff's. Rather, it was founded both on several independent principles

and on a multitude of elementary notions that could be defined only by an appeal to reality (by their concrete representation)—notions such as existence, space, time, and force; or, in psychology, the particular powers of the soul, some mental faculties, and pleasure and pain. Through Hoffman Crusius derived this view from Locke's doctrine of simple ideas, but he supposed that the number of elementary notions (which he once called categories) could be infinite.

THEOLOGY. Rational theology followed immediately after ontology, instead of being—as for Wolff—the final section of metaphysics, because Crusius held that God's existence is a necessary foundation for cosmology and pneumatology. Crusius denied the Ontological Argument: God's existence can be proved by moral evidence only, and his attributes cannot, properly speaking, be understood by humankind—among other things, positive infinity is beyond human reason. The human notion of God is partially relative and partially negative; nevertheless, it is certain. God is different from created beings both in degree and in essence. Among the attributes of God, Crusius stressed his free will, which is limited by the principle of contradiction and by his goodness. In God and God alone, intellect and will are a single power.

COSMOLOGY. Crusius held that matter is composed of a multitude of simple substances. Simple material substances are extended, and the infinite divisibility of matter is impossible. Simple substances have an essential, though not absolutely necessary, force. They act upon each other only by motion and contact. Physical space and time are real, but they are neither independent beings (substances), nor properties, nor relations (all of these concern the metaphysical essence of things). Space and time are intimately connected with existence; they are conditions of things insofar as such things exist. There is no real space or time without substance to fill it; outside the real world there is only possible (not sensible) space or time, which is infinite and filled by God. There are empty spaces in the world (otherwise movement would be impossible), but they are only physically—and not metaphysically—empty, because they are filled with God's presence. Mathematical space and time are distinct from physical space and time, and are abstracted from the relations of things.

Crusius was trying to offer a new set of solutions to the difficulties of the traditional doctrines of substance, of space and time and their limits, and of the void, while avoiding the concepts of *res extensa*, Leibnizian monads, and atoms, as well as the contradictions presented by the real space and time of René Descartes and Isaac Newton and the ideal space and time of Leibniz and Wolff. His doctrine resembled that of Locke, but it was a mixture of well-chosen elements of the traditional views connected by doubtful subtleties.

PNEUMATOLOGY. In his pneumatology, or rational psychology, Crusius rejected Thomasius's spiritual materialism but retained some of its characteristics. He held that finite spirits are simple unextended substances, but that they fill a space and share with material substances the power of motion. Thus, a real interaction between spiritual and material substances is possible, and the doctrines of preestablished harmony and occasionalism are unnecessary. The human soul is an independent substance with two fundamental powers, thinking and willing, both of which are a complex of several independent lesser powers.

Crusius was, in general, very cautious in his pneumatology, and frequently appealed to the limitations of human reasoning. For instance, he held that the immortality of the soul could be proved only if God's existence were presupposed—that is, by an appeal to moral certitude.

NATURAL PHILOSOPHY

Crusius's treatise on natural philosophy, *Anleitung, über natürliche Begebenheiten ordentlich und vorsichtig nachzudenken* (Introduction to regular and prudent reflections on natural events; 2 vols., Leipzig, 1749), was by far the least original of his works. Nevertheless, he was the first important Pietist philosopher to accept mechanism. In this work, Pietist philosophy finally renounced animism and adopted the more modern Cartesian and Leibnizian views, although it was still opposed to Newton's theory of gravitation. Crusius stressed the difficulties of physics and the purely hypothetical character of much of our knowledge of the particular laws of nature.

ETHICS

Crusius's first major work was a treatise on ethics, *Anweisung, vernunftig zu leben* (Instructions for a Reasonable Life; Leipzig, 1744). Hoffman's influence on Crusius is clear. Ethics, for Crusius, is not based on reason alone, but also on revelation. Natural duties have been imposed on humanity through God's free choice.

THE WILL. Crusius split Wolff's empirical psychology into two parts. He incorporated the first part, concerned with the cognitive power, into logic. The second, con-

cerned with the will, he placed in ethics. Moral goodness consists in the conformity of the human will with God's will. The human will is a power to act on the understanding, on the body, and on the will itself, but its connection with the understanding is not altogether clear. We are immediately conscious of freedom, which is the main property of the human will. The will is moved by sufficient reason, which does not necessitate, and therefore the will is free.

DUTY. The second section of the *Anweisung*, on ethics proper, discusses human duties. An action is moral if it is done out of obligation only, and not in quest of happiness. Virtue is formally conditioned by a coincidence of human will and divine law, and is materially conditioned by love for God. Divine law is known through conscience, which is an immediate power of moral judgment founded on a sort of common sense called moral taste. Evil originates in a wrong use of free will, which, when it submits to unreasonable impulses, corrupts human understanding and the true representation of goodness.

A third section of the *Anweisung* was devoted to moral theology; a fourth, to natural law; and a fifth, to prudence, which was closely studied in Thomasius's school and partially corresponded to Kant's technical imperatives.

REVEALED THEOLOGY

In his revealed theology Crusius united orthodox Pietist doctrines with those of a dissident Pietist, J. A. Bengel (1687–1752). Bengel and Crusius carried to an extreme the Pietist belief that the Bible is an organic whole inspired by God and historically true throughout. The Pietists held that Scripture is the only source of theological truth, and rejected all exegetical developments, even those of Protestant divines. No rational criticism of the Bible was permitted; its meaning could be penetrated only by a kind of empathy or inner light. Crusius stressed a theology of history, founded on biblical prophecies, that tried to explain the whole history of Christianity and to reveal its future aim in a second coming of Christ.

CRUSIUS'S INFLUENCE ON KANT

Recent historical scholarship has stressed Crusius's importance in Kant's development, and the view that Kant's philosophy was rooted in Wolff's system has been more and more questioned. Recent research has shown that Kant, educated in the Pietistic, eclectic, and anti-Wolffian milieu of Königsberg University, was mainly trying in his precritical development (1745–1768)—despite the nonorthodox Wolffian influence of his teacher, Martin Knutzen—to counteract Wolffian philosophy in an increasingly original way. He therefore appealed both to recent anti-Wolffian trends—to Maupertuis and his Berlin circle and through Maupertuis to Newton—and to Crusius, the new leader of Pietist philosophy and only nine years his senior, whose reputation grew tremendously from 1744 on.

Crusius's influence on Kant consists in six main points, some of which were also held by other Pietist philosophers or by Maupertuis. Crusius stressed the limits of human understanding, a theme that recurs in Kant's writings under different forms from 1755 on. He rejected the Ontological Argument, as did Kant after 1755, and he later rejected all theoretical proofs of God's existence. He assumed a multiplicity of independent first principles; Kant did so after 1755. He denied the importance of formal logic, and simplified it. He rejected the possibility of defining existence, and accepted a multiplicity of simple notions. He rejected the mathematical method as applied to philosophy. Kant adopted these last three positions in 1762.

Kant's Crusianism reached its climax in his *Untersuchung über die Deutlichkeit der Grundsätze der naturlichen Theologie und der Moral* (Investigations concerning the Distinctness of the Fundamental Principles of Natural Theology and Morals; Berlin, 1764), written in 1762. By 1763 Kant's enthusiasm for Crusius's philosophy was waning, but he did not reject the six tenets above and was still influenced by Crusius on individual points as late as the 1770s. J. Bohatec has claimed that Crusius's doctrines in revealed theology exerted some influence on Kant's late works in religion.

See also Descartes, René; Kant, Immanuel; Knutzen, Martin; Leibniz, Gottfried Wilhelm; Locke, John; Maupertuis, Pierre-Louis Moreau de; Newton, Isaac; Ontological Argument for the Existence of God; Pietism; Probability and Chance; Revelation; Rüdiger, Andreas; Sensa; Thomasius, Christian; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY CRUSIUS

Die philosophischen Hauptschriften, edited by Giorgio Tonelli. Hildesheim, 1964–.

Dissertations

De Corruptelis Intellectus a Voluntate Pendentibus. Leipzig, 1740.

De Appetitibus Insitis Voluntatis Humanae. Leipzig, 1742.

De Usu et Limitibus Principii Rationis Determinantis, Vulgo Sufficientis. Leipzig, 1743.

Opuscula Philosophico-Theologica. Leipzig, 1751. Reprint of above three dissertations.

Theological Work

Hypomnemata ad Theologiam Propheticam. Leipzig, 1764–1778.

WORKS ON CRUSIUS

Adickes, Erich. *Kantstudien*. Kiel and Leipzig, 1895.

Bohatec, J. *Die Religionsphilosophie Kants*. Hamburg, 1938.

Burkhardt, Hans. "Modalities in Language, Thought and Reality in Leibniz, Descartes and Crusius." *Synthese* 75 (1988): 183–215.

Campo, Mariano. *La genesi del criticismo kantiano*. Varese, 1953.

Delitzsh, Fr. *Die biblisch-prophetische Theologie, ihre Fortbildung durch Christian August Crusius, und ihre neueste Entwicklung seit der Christologie Hengstenbergs*. Leipzig, 1845.

Festner, C. *Christian August Crusius als Metaphysiker*. Halle, 1892.

Heimsoeth, Heinz. "Metaphysik und Kritik bei Chr. Aug. Crusius" (1926). Reprinted in his *Studien zur Philosophie Immanuel Kants*. Cologne, 1956.

Marquardt, A. *Kant und Crusius*. Kiel, 1885.

Schmalenbach, H. *Leibniz*, 553–560. Munich, 1921.

Schmucker, J. *Die Ursprünge der Ethik Kants*. Meisenheim an der Glan, 1951.

Seitz, A. von. *Die Willensfreiheit in der Philosophie des Christian August Crusius*. Würzburg, 1899.

Tonelli, Giorgio. *Elementi in Kant precritico*. Turin, 1959. Vol. I.

Tonelli, Giorgio. "Kant, dall'estetica metafisica all'estetica psicoempirica." *Memorie della Academia delle scienze di Torino*, Series 3, Vol. 3, Part 2 (1955).

Tonelli, Giorgio. "La question des bornes de l'entendement humain au XVIIIe siècle." *Revue de metaphysique et de morale* (1959): 396–427.

Tonelli, Giorgio. "Der Streit über mathematische Methode in der Philosophie in der ersten Hälfte des XVIII Jahrhunderts." *Archiv für Philosophie* 9 (1959).

Watkins, Eric. "The Development of Physical Influx in Early Eighteenth Century Germany: Gottsched, Knutzen, and Crusius." *Review of Metaphysics* 49(2) (1995): 295–339.

Watkins, Eric. "Forces and Causes in Kant's Early Pre-Critical Writings." *Studies in History and Philosophy of Science* 34a(1) (2003): 5–27.

Wundt, Max. *Die deutsche Schulphilosophie im Zeitalter der Aufklärung*, 254–264. Tübingen, 1945.

Wundt, Max. *Kant als Metaphysiker*, 60–81. Stuttgart: Enke, 1924.

Giorgio Tonelli (1967)
Bibliography updated by Tamra Frei (2005)

CUDWORTH, RALPH

(1617–1688)

Ralph Cudworth was one of the leading figures among the Cambridge Platonists, a group of seventeenth-century philosopher theologians. He was born in Aller, Somerset, to a minister who had been a fellow of Emmanuel College, Cambridge. Educated at home by his stepfather, John Stoughton, until 1632, he then entered Emmanuel College. There he was influenced by Benjamin Whichcote, founder of the Cambridge Platonist school. In 1639 he was elected a fellow of Emmanuel, and received the bachelor of divinity degree in 1645, defending for his examination Whichcote's thesis that good and evil are eternal and immutable. This examination, with its opposition to any system that makes morality contingent on will, whether human or divine, already betrays Cudworth's distance from the rigorous Calvinism with which Emmanuel College had always been associated. Nevertheless, Cudworth did have some sympathy with political aspects of the Puritan cause. He was appointed by Parliament master of Clare College and Regius professor of Hebrew in 1645, and served as advisor to Oliver Cromwell's secretary of state on several government appointments. In 1647 he was invited to preach to a sharply divided House of Commons.

In his sermon of March 31, 1647, Cudworth urged parliamentarians not to legislate on doctrinal matters, arguing that salvation depends not on speculative details but on living a life of Christlike love and forbearance. This emphasis on morality over doctrine was characteristic of the Cambridge Platonists and influential for the later Latitudinarian divines. Cudworth was appointed master of Christ's College in 1654, and succeeded in retaining his appointment at the time of the Restoration. He remained in the post until his death in 1688. In 1654 Cudworth was married. None of his sons survived him, but his daughter, Damaris, later Lady Masham (1658–1708), took custody of her father's writings and became a philosopher in her own right. Intimate friend and correspondent of John Locke, she published *A Discourse concerning the Love of God* in 1696.

METAPHYSICS

Cudworth's massive *True Intellectual System of the Universe* (1678) was the only one of Cudworth's principal writings to be published during his lifetime. Lengthy as it is, the published volume represents only the first installment of a three-part project originally sketched by Cudworth, with the parts devoted respectively to attacking

mechanistic or atomistic determinism, theological determinism (Calvinism), and Stoic determinism. The *True Intellectual System of the Universe* in its published form constitutes a defense of theistic atomism and an attack on "Hylopathic" atheism. Hylopathic atheism, which claims that all things can be explained by reference to material atoms, with no need to invoke spirit or incorporeal substance, was an important target because it was represented in Cudworth's own day by Thomas Hobbes. A secondary target of the book is "Hylozoic" atheism, differing from Hylopathic in attributing life to matter, but still materialistic, and worthy of Cudworth's notice because of its recent revival by Benedict Spinoza. Rather than engage directly with Hobbes or Spinoza, Cudworth's argument is directed against ancient schools of philosophy, and much of it consists in a *consensus gentium* argument; atheism is an anomaly or aberration from an original truth that has been acknowledged from the beginning. This original true system accepted atomism, but only as an account of matter or corporeal reality. Properly understood, atomism reveals matter to be essentially passive or inert, thus making clear that only the existence of incorporeal substance can explain the origin, motion, and organization of matter.

While interested in the *ancient theology* argument that Plato's insights (particularly what Cudworth regarded as a concept of the Trinity in Plato) derived originally from divine revelation through Moses, Cudworth was finally content to claim that there is a natural *prolepsis* or anticipation of the idea of God existing throughout all times and places. Atheism is thus a willful destruction of this prolepsis, and Hobbes and Spinoza are not new threats, but reincarnations of old foes. The other leading philosophical thinker among the Cambridge Platonists, Henry More (1614–1687), devoted more direct attention to Spinoza's thought than did Cudworth, and differed from Cudworth as well in seeking, along with Joseph Glanvill (1636–1680), empirical evidence for the existence of incorporeal substances in cases of witchcraft and demonic possession.

In embracing atomism, Cudworth was making common cause with Cartesian dualism and rejecting scholastic accounts of substantial forms. Mind cannot be simply a property of material objects. For Cudworth, though, it is passivity, not extension, that essentially constitutes matter, and it is activity, rather than self-consciousness, that essentially constitutes incorporeal substance. The key challenge facing Cartesian dualism was to account for the interaction between body and soul, corporeal and incorporeal substance. Cudworth's solution was to appeal to active incorporeal powers that mediate between wholly passive matter and self-conscious soul, creating a vital union between them. Each finite soul has a finite field of action—its own body.

An analogous solution allows Cudworth to articulate the relationship between God and the world. While God is not bound to physical creation as a finite soul is bound to its body, Plastic Nature does serve as an intermediary between God and world that, like the lower powers of the soul, allows for a vital connection between the two. Critical of Descartes's suggestion that the existing ordered universe could have originated from a single initiating divine act, Cudworth argued that an ongoing divine influence was necessary if the material universe was to maintain an ordered motion. At the same time, God is not required, as in *occasionalism*, to attend directly to each and every detail of order in the universe. *Plastic Nature*, an unconscious power that pursues not its own but divine purposes, imposes order and finality on the material world. Nothing works according to mere chance, but according to final causes, divine intentions mediated by Plastic Nature.

Cudworth's Plastic Nature is similar to More's *Hylarchic Principle*, but Cudworth did not follow More's contentions that both material and immaterial substance are extended and that space is infinite. The concept of Plastic Nature was influential for biologist John Ray and for philosophical biology generally up through Darwin. Pierre Bayle attacked Cudworth's plastic powers as atheistic in tendency in eliminating the need for direct divine action to account for purpose displayed in the physical world. Denis Diderot and Jean d'Alembert's *Encyclopédie* (1751–1772) included a detailed account of Cudworth's theory, and the theory may, via Paul Janet, have influenced modern doctrines of the unconscious.

On the one hand, Cudworth cannot be finally understood as opposing the new philosophies of Hobbes and Spinoza merely on the basis of allegiance to an outmoded neoplatonism; his *consensus gentium* argument advances at the same time a contemporary position. On the other hand, the baroque erudition displayed in the *True Intellectual System of the Universe* was out of step with the leaner philosophical style of his contemporaries. This did not prevent the work from achieving significant influence in its own day, and in fact the text served for several generations as a key resource on Greek philosophy. But it did in the nineteenth and twentieth centuries often mean that Cudworth's importance as an interlocutor of Hobbes, Spinoza, and Descartes was not fully appreciated.

EPISTEMOLOGY

Cudworth's Platonist epistemology is developed at length in *A Treatise of Eternal and Immutable Morality*. The *Treatise* was published in 1731 at the behest of Cudworth's grandson, Lady Masham's son. Cudworth argues against empiricism that knowledge is more than a mirroring or representation of reality, and that the mind is more than a blank sheet of paper upon which the objects of sense are inscribed. Knowledge can never arise solely out of sense experience. When we sense, we sense particulars, but when we know, we know by means of universals. Cudworth insists that universals must precede the empirical particulars that they organize and make sense of; they are not abstracted from particulars, for this act of abstraction would be unmotivated and undirected unless one already knew the universal at which one was aiming. Sense allows the soul to perceive the appearances of things, but not clearly to comprehend them.

The universals of which we have knowledge are eternal and immutable. But the fragmented nature of human nature points beyond itself to God. Given the eternal and immutable nature of intelligible ideas, they cannot solely be modifications of limited and finite intellects, which only come to know them in time, if at all. It is God's infinite and eternal mind that, in perceiving itself, eternally perceives these ideas. Rather than innate ideas, human souls possess innate activities or tendencies, a capacity to exert themselves so as to participate in a limited way in divine self-knowledge. Human persons do not, though, arrive at knowledge by comparing their ideas with ideas in the mind of God. Pointing out the impossibility of such a comparison, Cudworth insists simply on clear intelligibility as the criterion of truth. Descartes, Cudworth argues, fell into circularity in seeking further to defend the criteria of clarity and distinctness by proving that God is not a deceiver. The criterion of clear intelligibility is self-evident and depends on no external support.

ETHICS

As the title of *A Treatise* suggests, among the eternal and immutable ideas that may be known as clearly intelligible are moral principles. In fact, Cudworth's epistemological discussion is occasioned and motivated by his concern to defeat voluntarist and relativist accounts of morality. This concern reached back to Whichcote, but Cudworth was both much more learned than Whichcote about ancient and more recent Platonism and much more connected to contemporary philosophical discourse. As in the *True Intellectual System of the Universe*, one of Cudworth's key targets is Hobbes, who argues that right and wrong are relative concepts, based solely in convention. Cudworth also attacks "diverse modern theologers" who argue that morality is created by divine fiat, naming among them William of Ockham and Pierre d'Ailly and one contemporary theologian, the Polish Jan Szydlowski. The Calvinistic theology of Cudworth's Puritan contemporaries is a looming unnamed target. Descartes's argument that the natures and essences of things, including moral good and evil, must depend on the contingent will of God in order not to be independent of God, receives particular criticism. Following Plato's *Euthyphro*, Cudworth argues that things are not good because they are willed by God; rather, God wills things because they are good. It is either eternally true or eternally false that something is good; no act of will can change this. The good is not, though, an external constraint on divine freedom, but God's essential nature.

If the *True Intellectual System of the Universe* was originally intended as a comprehensive critique of all forms of determinism, Cudworth's many manuscripts defending "freewill" represent his efforts to articulate a positive account of free human action and a moral psychology to accompany that. None of these manuscript treatises were published during Cudworth's lifetime, and it is unclear how widely they may have circulated. Lady Masham may well have shared them with Locke and Shaftesbury. The shortest of the manuscripts was published in 1838 as *A Treatise of Freewill*, testifying to ongoing interest in Cudworth's thought. In an innovative move, Cudworth rejects traditional faculty psychology; the will and understanding are not distinct faculties in the soul, but activities of the soul. Drawing on Stoic terminology, Cudworth argues that the soul's *hegemonikon* or ruling power "is the soul as comprehending itself, all its concerns and interests, its abilities and capacities, and holding itself, as it were in its own hand, as it were redoubled upon itself" (p. 178). It is through this reflexive capacity that the soul is able to adjudicate among conflicting passions, dictates of conscience, and inferences of reason, and act as a unified self. Cudworth considers that in identifying this capacity for reflexive deliberation he has successfully shown that persons are not determined by any "antecedent necessary causes" (p. 179). It is far from clear that this is so, although Cudworth's account of the *hegemonikon* does make it possible to speak intelligibly of the soul's self-determination and moral responsibility.

For Cudworth, moral agency and thus moral responsibility rest on the capacity to survey in a comprehensive way possible courses of action and to pass judgment on

which is best. Only action that issues from such reflection can properly be regarded as one's own. God neither has nor needs a *hegemonikon*, being simple and unified. God's freedom consists in unfailingly acting according to God's own perfect nature, and God's self-determination in the fact that nothing outside of God determines divine action. Insofar as human self-determination takes the form of an active pursuit of the good, human persons come to participate increasingly fully in God and God's goodness. Thus human "freewill" can be employed in order to arrive at a more perfect, more godlike, freedom.

See also Cambridge Platonists; Epistemology; Ethics; Metaphysics.

Bibliography

WORKS BY CUDWORTH

The True Intellectual System of the Universe. Stuttgart-Bad Cannstatt, Germany: Friedrich Fromann Verlag, 1964. Facsimile of 1678 edition.

"A Sermon Preached Before the House of Commons." In *The Cambridge Platonists*, edited by C. A. Patrides. Cambridge, MA: Harvard University Press, 1970.

A Treatise concerning Eternal and Immutable Morality, with *A Treatise of Freewill*. Edited by Sarah Hutton. Cambridge, U.K.: Cambridge University Press, 1996.

WORKS ABOUT CUDWORTH

Darwall, Stephen. "Cudworth: Obligation and Self-Determining Moral Agency." In *The British Moralists and the Internal Ought, 1640–1740*. Cambridge, U.K.: Cambridge University Press, 1992.

Hutton, Sarah. "Ralph Cudworth, God, Mind and Nature." In *Religion, Reason and Nature in Early Modern Europe*, edited by Robert Crocker. Dordrecht, Netherlands: Kluwer, 2001.

Passmore, J. A. *Ralph Cudworth: An Introduction*. Cambridge, U.K.: Cambridge University Press, 1951.

Scott, D. "Platonic Recollection and Cambridge Platonism." *Hermathena* 149 (1990): 73–97.

Thiel, Udo. "Cudworth and Seventeenth-Century Theories of Consciousness." In *The Uses of Antiquity: The Scientific Revolution and the Classical Tradition*, edited by Stephen Gaukroger. Dordrecht, Netherlands: Kluwer, 1991.

Jennifer A. Herdt (2005)

CULVERWEL, NATHANAEL
(1618?–1651?)

Nathanael Culverwel, the religious and moral philosopher commonly if rather misleadingly described as a Cambridge Platonist, was probably a son of Richard Culverwel, rector of St. Margaret's, in London, although neither his parentage nor the date of his birth is certain. He certainly grew up in a Calvinist atmosphere. In 1633 he was admitted to Emmanuel College, Cambridge, where he encountered the teachings of Benjamin Whichcote, the spiritual leader of Cambridge Platonism. Ralph Cudworth was slightly junior to him as an undergraduate at Emmanuel but was elected to a fellowship three years before Culverwel's election in 1642. John Smith was of the same generation. Culverwel's contemporaries refer in somewhat obscure terms to troubles that beset him in later life; these may have included some sort of mental breakdown. He died not later than 1651.

Culverwel published nothing during his lifetime. Shortly after his death, however, William Dillingham prepared for publication a discourse titled, in Culverwel's typically metaphorical style, *Spiritual Opticks: or a Glasse discovering the weaknesse and imperfection of a Christians knowledge in this life* (1651). This was sufficiently successful to encourage Dillingham to proceed to the publication of a manuscript by Culverwel, composed, Dillingham says, about 1646, which was obviously intended, although incomplete, to be a book—*An Elegant and Learned Discourse of the Light of Nature*. In the same volume Dillingham included a number of Culverwel's sermons. Prefixed to the *Discourse* is an essay by Culverwel's brother Richard that asserts that in its present form the *Discourse* is somewhat misleading, since the praise of reason which it contains was to have been followed by another section in which the limitations of reason would have been more strongly insisted upon. That judgment is borne out by the tone of Culverwel's sermons, which are severely Calvinist.

The *Discourse*, as it stands, is an elaboration of Whichcote's favorite quotation (from Proverbs 20:27), which Culverwel translates as "The understanding of a man is the candle of the Lord." Insofar as it is critical of those who "blaspheme reason," the *Discourse* is written in Whichcote's spirit. However, its philosophical tone is in many respects Aristotelian rather than Platonic; Culverwel sharply criticizes the "fanciful ideas" of "the Platonists," under which heading he almost certainly includes his Emmanuel colleagues. (None of them had yet published, so that although—unusually for his time—Culverwel makes precise references to such near-contemporaries as Lord Herbert of Cherbury, Lord Brooke, and Sir Kenelm Digby, he could not refer to the Cambridge Platonists in similarly definite terms.) When Culverwel speaks with enthusiasm of Plato, it is of the *Laws* or the *Republic* rather than of John Smith's favorite, the *Phaedo*; quite unlike Smith or Cudworth he rarely pays any attention to the Neoplatonists. On the other hand, he writes

with great approval not only of Aristotle but also of the Scholastics, especially Thomas Aquinas and Francisco Suárez, and even of Francis Bacon, to whom the Platonists were generally strongly opposed.

He differs from the Platonists on four crucial points. The first is epistemology; he disagrees with them, as he puts it, about "the time at which the candle of the Lord is lighted." It is true that at an early stage in the *Discourse* (Ch. 7) he writes: "There are stamped and printed upon the being of man some clear and indelible principles, some first and alphabetical notions, by putting together of which it can spell out the law of nature," a passage which it is natural to read as a defense of innate ideas. Later, however, in Chapter 11, he argues quite explicitly against the doctrine of innateness, even in the modified form in which Platonists like Cudworth held it. First principles—which he describes as having "so much of certainty in them, that they are near to a tautology and identity"—arise, he argues, "from the observing and comparing of objects"; these principles are not inherent in our minds. He strongly criticizes Plato and René Descartes in Chapter 14 for "too much scorning and slighting" of sensations. Sensation, he admits, is no more than "the gate of certainty," but only through this gate can certainty enter the soul. Otherwise, the soul would remain "a blank sheet."

Second, he criticizes the Platonist tendency to diminish the gap between human and divine by treating the human soul as having a degree of divinity, as being, insofar as it is rational, an ingredient in divine reason. The candle of the Lord, he argues, is lit by God but is no part of God's light. God's light is like the sun; a candle is but a wavering, imperfect light even when it is at its brightest. Men cannot hope to be godlike, the ideal the Platonists set before themselves.

This is connected with the third point of difference. Culverwel continued to be a Calvinist; he continued to believe, therefore, that no human being is worthy of salvation. In a sermon titled "The Act of Oblivion," addressed to a congregation presumed to belong to the elect, he says that God "might have written thy name in his Black Book, with fatal and bloody characters, and made his justice glorious in thy misery and damnation"; God had chosen otherwise because he so chose, not because any members were deserving of a higher destiny. If God has chosen to save Socrates, he argues, this can only be because God gave a private revelation to him, not because Socrates was a worthy man. God may well have chosen to save Aristophanes rather than Socrates. God's decrees, Culverwel insists, are absolute; it is ridiculous to suppose that a man can save himself from the damnation decreed for him merely by exercising an act of choice, by choosing to be good. Nothing could be further from the spirit of Cambridge Platonism than Culverwel's unmitigated Calvinism.

Finally, and this again is connected with his Calvinism, Culverwel's emphasis as a moral philosopher is on law rather than on reason. He agrees with the Platonists, it is true, that some acts are good in their own nature and that some relationships are peculiarly just and rational; however, the performance of such acts, he argues, does not constitute a moral good. Essentially, he says, morality is a matter of obedience to rule, and there can be rules only when there is a lawgiver. The obligatoriness of moral laws depends upon the fact that they are commanded by God. Even though the lawgiving is itself a rational act, even though moral laws are based upon the lawgiver's apprehension of "the eternal relations of things," even though it is by our reason that we discover their nature, command, not reason, is still the foundation of morality. A capacity for obeying rules, he suggests, is the distinguishing mark of a rational being; moral rules apply to men, not to animals, just because men are capable of following rules. But human rationality does not in any way constitute the obligatoriness of the rules.

Following Hugo Grotius, Culverwel devotes a great deal of attention to the concept of a natural law and its relation to the laws of nations. In the *Discourse*, as his argument proceeds, the importance of law comes more and more to the fore, and the importance of reason recedes, although Culverwel takes the two to be intimately connected. For Culverwel, as for so many of his antirationalist successors, Abraham's sacrifice of Isaac is the crucial case. This was decreed, and the decree had to be obeyed, he argues, even though it goes against all our concepts of a rational morality; "the candle durst not oppose the sun."

One can discern a tension in Culverwel's work between his Calvinism and the Platonism he had learned from Whichcote. A very similar tension between empiricism and rationalism, between the concept of law and the concept of reason, is manifest in John Locke, and it is more than likely that Locke was strongly influenced by Culverwel's *Discourse*, most obviously, but by no means exclusively, in the *Essays on the Law of Nature*, which he wrote in 1660.

See also Aristotelianism; Aristotle; Bacon, Francis; Cambridge Platonists; Cudworth, Ralph; Descartes, René; Grotius, Hugo; Herbert of Cherbury; Locke, John;

Neoplatonism; Plato; Platonism and the Platonic Tradition; Socrates; Suárez, Francisco; Thomas Aquinas, St.; Whichcote, Benjamin.

Bibliography

The best edition of the *Elegant and Learned Discourse of the Light of Nature* is edited by John Brown, with a critical essay by John Cairns (Edinburgh: Constable, 1857). The sermons, including *Spiritual Opticks*, can be read in the 1652 edition of the *Discourse* (London: John Rothwell; reprinted, 1654). There are extracts in John Wesley, *Christian Library*, Vols. IX–X (London, 1819–1827); and Ernest Trafford Campagnac, *The Cambridge Platonists* (Oxford: Clarendon Press, 1901).

See also William Cecil de Pauley, *The Candle of the Lord, Studies in the Cambridge Platonists* (London: Society for Promoting Christian Knowledge, 1937); A. C. Scupholme, "Nathanael Culverwel—a Cambridge Platonist," in *Theology* 38 (225) (1939); Wolfgang von Leyden, *John Locke: Essays on the Law of Nature* (Oxford: Clarendon Press, 1954); Jay G. Williams, "Nathaniel Culverwel (1619?–1651?)," in *British Philosophers, 1500–1799 (Dictionary of Literary Biography, Vol. 252)*, edited by Peter Fosl (Farmington Hills, MI: Thomson Gale, 2002); and the bibliography to the *Cambridge Platonists* entry.

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

CUMBERLAND, RICHARD
(1631–1718)

Richard Cumberland, the bishop and moral philosopher, was born in London, the son of a London citizen. Educated at St. Paul's School, in 1648 he entered Magdalene College, Cambridge, where, distinguishing himself both by his scholarship and by his capacity for friendship, he was elected a fellow in 1656. He first studied medicine, but he finally decided to enter the church, accepting preferment in 1658 to the rectory of Brampton, Northamptonshire, and in 1667 to the rectory of All Hallows at Stamford, Lincolnshire. In 1661, Cambridge appointed him one of its twelve official preachers, and he kept in close touch with Cambridge intellectual life. Cumberland earned the reputation of being an exceptionally staunch Protestant. Report has it that the attempt of James II to reintroduce Roman Catholicism into England produced in him a dangerous fever. Such zeal did not go unrewarded under William III, and although quite without personal ambition, Cumberland was consecrated as bishop of Peterborough on July 5, 1691. He performed his episcopal duties with diligence until his death in 1718.

Jewish history was Cumberland's main interest. In 1686 he published *An Essay towards the Recovery of the Jewish Measures and Weights*. His domestic chaplain and son-in-law Squier Payne published in 1720 *Sanchoniatho's Phoenician History*, translated with a commentary by Cumberland. This monument of misplaced scholarly ingenuity derived its immediate inspiration from Hugo Grotius. With no qualms about the authenticity as history of Sanchoniatho's cosmogony, Cumberland devoted himself to identifying its personages with characters in the Old Testament. A sequel, *Origines Gentium Antiquissimae; or Attempts for Discovering the Times of the First Planting of Nations*, appeared in 1724.

Cumberland's sole philosophical work, *De Legibus Naturae* (1672), was designed, as the subtitle explains, as a refutation of Thomas Hobbes—the first full-length philosophical reply to Hobbes to be published. Written in an inelegant Latin, badly printed, ill-organized, intolerably diffuse, Cumberland's treatise did not attract much contemporary attention. In 1692, with Cumberland's approval, James Tyrrell prepared an abridgment and translation as *A Brief Disquisition of the Law of Nature*, hoping to draw attention to Cumberland's main ideas. But the abridgment was a poor one (in addition, Tyrrell's own views were mingled with Cumberland's) and failed in its main purpose. Eighteenth-century philosophers were more interested in Cumberland's work than his contemporaries had been; he anticipated their ambitions and preoccupations. A complete English translation was prepared by John Maxwell in 1727, and what has become the standard translation was published, with copious annotations by John Towers, in 1750. A French translation by Jean Barbeyrac (1744) ran into two editions.

Cumberland's point of departure is Grotius's *De Iure Belli et Pacis* (1625). Grotius, or so Cumberland interprets him, had based his demonstration of the existence and binding force of natural laws upon the consensus of civilized opinion. Very conscious of Hobbes, Cumberland sets out to supplement Grotius by demonstrating that natural laws are founded on "the nature of things," as distinct from the commands of sovereign rulers. To that extent Cumberland's aims coincide with Ralph Cudworth's, but unlike Cudworth he does not base his argument on Platonic metaphysics. Nor does he criticize, as did the Cambridge Platonists, the mechanical worldview; indeed, he wholeheartedly accepts it. He thinks of his approach as scientific and nonmetaphysical. He sets out to construct an ethics that, although Christian, is independent of revelation and, although demonstrating that morality is eternal and immutable, is based on "the evi-

dence of sense and experience." These were to be the typical eighteenth-century specifications for a satisfactory moral theory.

Cumberland begins by arguing that there is a single natural law from which all moral laws can be derived—the law, namely, that an agent secures his own good by the promotion of the good of the whole to which he belongs. If this single law is based on "the nature of things," if its truth can be demonstrated from experience, then, he thinks, morality rests secure. And, he argues, experience reveals to us—he draws upon his medical training to illustrate the point—that the parts of a whole secure their own welfare only when they work for the good of the whole to which they belong. A bodily organ, for example, is at its healthiest when it is most effectively securing the health of the body. This truth men could recognize, so Cumberland argues against Hobbes, even in a state of nature. Thus, the foundation of moral laws is not the will of the sovereign.

Benevolence, Cumberland further maintains, is natural to humankind. Even brute animals, indeed, devote themselves to the welfare of their fellow brutes. A state of nature, therefore, would not be, as Hobbes suggested, a war of all against all; their human instincts, not the pressure of a sovereign will, lead men to cooperate with their fellow men in society. Certainly, Cumberland admits, men sometimes act in opposition to the good of the whole, just as an organ of the body will sometimes infect, rather than work toward the health of, the organism of which it forms a part. The fact remains, however, that the "natural impetus of man" is toward securing the common good, just as the general tendency of a bodily organ is to make the body healthier. The legislator's rewards and punishments, like medicine, are directed toward correcting abnormalities; they are not the original springs of moral action.

All moral concepts, Cumberland tries to show, are definable in terms of the single natural law that men secure their own welfare by pursuing the common good. An act is "naturally good" if by virtue of its own nature it tends toward the common good; it is "right" if it is the shortest way to that end; it is "morally good" if it conforms to the natural law. Particular virtues are similarly deducible from the obligation of pursuing the common good; to show that the common good ought to be our objective is at the same time to show that we ought to be law-abiding, just, temperate, and obedient to God.

Most of what were to be the leading eighteenth-century moral theories can be found somewhere suggested, if nowhere fully worked out, in *De Legibus Naturae*. Cumberland argues in detail that moral principles are analogous to the propositions of mathematics, and Samuel Clarke learned much from him on this point. Cumberland also sketches a moral calculus of the sort Francis Hutcheson was to employ; there are many resemblances between his moral philosophy and the third earl of Shaftesbury's; he has been described as the first systematic utilitarian; the organic theory of morality and of the state is conspicuous in his work; resemblances between Cumberland and Benedict de Spinoza are easy to detect.

Accounts of his moral philosophy differ widely, depending on which of the manifold tendencies in his thinking commentators stress. In Cumberland's own eyes, however, the crucial points are (1) there is a law of nature, defined as a proposition of "unchangeable truth and certainty … which lays firm obligations upon all outward acts of behaving, even in a state of nature"; (2) this law enjoins upon us the pursuit of the common good and assures us that by pursuing the common good we achieve happiness and personal perfection; (3) observation of the world, including man's nature, demonstrates the truth of this law; (4) all other moral precepts are applications of the law of nature to particular forms of human action.

See also Cambridge Platonists; Clarke, Samuel; Cudworth, Ralph; Ethics, History of; Grotius, Hugo; Hobbes, Thomas; Hutcheson, Francis; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Spinoza, Benedict (Baruch) de.

Bibliography

The best edition of Cumberland's *De Legibus Naturae* is John Tower's 1750 translation (Dublin), which also includes, as Appendix IV, Squier Payne 's *Life of Cumberland*, the main biographical source. For commentaries, see Frank Elsworth Spaulding, *Richard Cumberland als Begründer der englischen Ethik* (Leipzig: Fock, 1894); Ernest Albee, *A History of English Utilitarianism* (London: Allen and Unwin, 1902); Frank Chapman Sharp, "The Ethical System of Richard Cumberland, and Its Place in the History of British Ethics," in *Mind* 21 (83) (1912): 371–398; James Clarke, *Richard Cumberland and Natural Law: Secularisation of Thought in Seventeenth-Century England* (Cambridge, U.K.: James Clarke, 1987); William Ewald, "The Biological Naturalism of Richard Cumberland." *Jahrbuch fur Recht und Ethik* (8 [2000]: 125–141); Murray Forsyth, "The Place of Richard Cumberland in the History of Natural Law." *Journal of the History of Philosophy* (20 [1982]: 23–42); M. A. Stewart, M. A., ed. *English Philosophy in the Age of Locke* (New York: Oxford University Press, 2000).

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

CUSANUS

See *Nicholas of Cusa*

CYNICS

"Cynics," the "dog philosophers" of the Greek and Roman world, so called almost certainly from the nickname of Diogenes of Sinope, were not a continuous school of theoretical philosophy but an erratic succession of individuals who from the fourth century BCE to the sixth century CE preached, through ascetic practice and mordant denunciation of established convention, a more or less similar way of life designed to lead to the happiness of the individual. Consequently there is no established doctrinal canon by which to define an "orthodox" Cynic, and the ancient but still lively debate as to whether Antisthenes or Diogenes was the founder of Cynicism is an unreal one. Nevertheless, despite marked variations of stress and tone in individual exponents, Diogenes was always regarded as the arch-Cynic, and a sufficient number of characteristic attitudes recur to identify the movement.

The nature of the existing evidence of Cynicism is highly unsatisfactory. The written works with which Diogenes was credited have not survived, and doxographies are few and of uncertain origin (for example, Diogenes Laërtius, Bk. 6, 70–73). Since Diogenes's life was his main testament, the largest class of evidence is anecdotal, with all the uncertainties and elaborations of an oral tradition. Information of his pupils and of Cynics of the third century BCE is tantalizingly fragmentary. Even the comparatively abundant material on contemporary Cynicism from the first century CE comes from outside the movement, from sympathizers of such diverse interests as Epictetus, Dio Chrysostom, and Julian, or from satirists such as Lucian.

TEACHING

The Cynics believed that happiness was found in "virtuous" action, which was the practical expression of self-realization (*arete* and "know thyself"). This state was in turn produced by a rational awareness of the distinction between natural and artificial values. External and physical goods such as wealth, reputation, pleasure, conventional duties arising from family, property, or state, and all traditional inhibitions, whether social or religious, were condemned as unnatural tyrannies that fettered a man through desire, indulgence, and the ignorance of a confused and corrupt society—the three causes of human misery. Freedom was secured by "following nature" by means of self-discipline whose end was self-sufficiency (*autarkeia*); since man was vulnerable and perverted through his emotions and desires, happiness could be guaranteed only by the understanding and strength of mind to want nothing, lack nothing. And since the artificial currency of human standards was thought to be, not an indifferent factor, but an active corruption to be eradicated, Cynics wished not merely to devalue the coin (like Socrates and the Stoics), but to deface it (*paracharattein*); hence, the most characteristic feature of Cynicism was an asceticism that sought to reduce physical wants to a minimum, as in the case of the animals after which Cynics were named, and to achieve spiritual independence like gods.

Independence was not to be achieved, however, by the withdrawal of a hermit; the Cynic engaged in an active crusade that required a continual training (*askesis*) to harden the body and temper the spirit in the very face of temptation, and thus to free the natural "perceptions" and capacities for virtuous actions. The toiling, painful effort of this moral struggle (*ponos*) was categorized as a good, the steep short cut to virtue, which evoked the only natural pleasure; and the legend of Herakles's life of service spent in successfully overcoming labors was sanctified as an ideal of freedom and self-fulfillment. He and the Cynic, whether slave or oppressed, ruled himself as his own master and, therefore, was the ideal king among men. Essentially individualistic and largely antisocial in advocating independence from any community, Cynicism was the most radical philosophy of spiritual security offered to fill the social and moral vacuum created in the fourth century BCE by the dissolution of the city-state political organism. Yet there was a strong philanthropic impulse in the movement in the sense that the gospel of Herakles, the ideal king, was a spiritual evangel for all men, to be preached by personal example. The Cynic saw himself as "scout and herald of God," dedicating his own labors as a reconnaissance for others to follow; he was the "watchdog of mankind" to bark at illusion, the "surgeon" whose knife sliced the cancer of cant from the minds of others. Cynics deliberately adopted shamelessly shocking extremes of speech and action to jolt the attention and illustrate their attack on convention.

Fearlessness in criticism was a virtue, useful to further Cynic ideals, but it was also open to abuse, as was the license of affected shamelessness. There was always a real danger that the negative, denunciatory side of Cynicism would predominate, the more so since happiness was most often described as freedom from misery, and virtue,

practical wisdom, and right reason remained somewhat nebulous terms. The Cynics did not offer arguments to intellectuals, whose theories they despised as useless. Rather, they offered the ideal practical example of autonomy of will through their own actions, bringing by the very vilification of luxury and sensual indulgence and by the justification of poverty, spiritual hope to the poor, disenchanted and oppressed. Thus the more formal types of philosophical instruction were abandoned and three new literary genres fostered: the *chreia*, or short anecdotal quip with a pungent moral tang; the diatribe, or popular sermon in conversational style; and Menippean satire.

HISTORY OF THE MOVEMENT

The most influential of Diogenes's converts was Crates of Thebes. Joined by his wife in a life devoted to Cynic ideals, he earned by his humanity and good works the affectionate name of "Door-Opener." He wrote philosophical tragedies and poetry about a Cynic paradise named the island of Pera. In the third century BCE. Bion of Borysthenes, a wandering preacher, was "the first to tart up philosophy" by popularizing the diatribe; Menippus of Gadara initiated a new type of satire mingling seriocomic themes in prose and verse (his works are lost); Cercidas of Megalopolis applied Cynic ideas to practical politics by proposing reforms attacking social inequalities in the refounding of his city; the fragments of Teles, a dull Megarian schoolmaster, throw some light on Bion and earlier Cynics. After a quieter, although not dormant, period Cynicism revived in the first century CE with some encouragement from Stoicism: Demetrius was prominent in the Stoic-flavored opposition to the emperor in the seventh decade; Dio Chrysostom found solace for his exile in an amalgam of Cynic and Stoic practice; Epictetus, the Stoic, admired Diogenes.

The second century records the apogee of Cynic influence and extravagance. The leading figures differed sharply. The philanthropy and popularity of Demonax of Cyprus contrasted with the brutal scorn of Oenomaus of Gadara. Peregrinus Proteus, a convert from Christianity, was an irrepressible radical with a touch of the mystic; he burned himself to death before huge crowds at the Olympic festival. These were men of ideals; but Lucian and Julian also record with disgust a riffraff of confidence tricksters and professional beggar-preachers masquerading under the Cynic uniform of cloak, knapsack, and stick. The peculiar animal-divine polarity of Cynicism attracted both saints and rogues. In the history of Greek thought Cynicism was most influential on the development of Stoicism, first through Zeno and then much later with Epictetus, who gave noble expression (3.22) to the most uncompromisingly radical ethic that anyone attempted to put into practice in the ancient world.

It is tempting to recognize Cynic traits in other civilizations, as Onesicratus, the admiral and historian of Alexander, did on encountering the gymnosophist Indian fakirs. In medieval times, the mendicant friars are more apposite than anchorites, especially when one considers the complementary virtues of Franciscans and Dominicans (*Domini canes*).

See also Antisthenes; Asceticism; Diogenes of Sinope; Epictetus; Lucian of Samosata; Megarians; Socrates; Stoicism; Zeno of Citium.

Bibliography

SOURCES, COMMENTARIES, AND TRANSLATIONS

Giannantoni, G. *Socratis et Socraticorum Reliquiae*. 4 vols. Naples: Bibliopolis, 1990.

For Crates: Diehl, E. *Anthologia Lyrica Graeca* I. Leipzig, 1958.

For Teles: Hense, O. *Teletis Reliquiae*. Freiburg, 1889.

Malherbe, A. J. *The Cynic Epistles*. Missoula: Scholars Press for the Society of Biblical Literature, 1977.

Müseler, E. *Die Kynikerbriefe*. 2 vols. Paderborn: F. Schöningh, 1994.

Caizzi, F. D. *Antisthenis Fragmenta*. Milan, 1966.

Von Fritz, K. *Quellenuntersuchungen zur Leben und Philosophie des Diogenes von Sinope*. Philologus Suppl. 18, Leipzig, 1926.

STUDIES

Billerbeck, M. *Die Kyniker in der modernen Forschung*. Amsterdam: B. R. Grüner, 1991.

Caizzi, F. D. "Antistene." *Studi Urbinati* 1 (1964) 25–76.

Downing, F. G. *Cynics and Christian Origins*. Edinburgh: T & T Clark, 1992.

Dudley, D. R. *A History of Cynicism from Diogenes to the Sixth Century A.D*. London: Methuen, 1937.

Goulet-Cazé, M.-O. "Le cynisme à l'époque impériale." *Aufstieg und Niedergang der Römischen Welt* 2.36.4. Berlin: W. de Gruyter, 1990.

Goulet-Cazé, M.-O. *Sophiês Maiêtores*. In *Chercheurs de Sagesse: Hommage à Jean Pepin*. Collection des Études Augustiniennes, Serie Antiquité 131 (1992) 5–36.

Goulet-Cazé, M.-O., and R. Goulet, eds. *Le cynisme ancien et ses prolongements: Actes du Colloque intérnational du C.N.R.S*. Paris: Presses universitaires de France, 1993.

Hahm, D. "Diogenes Laertius VII: On the Stoics." *Aufstieg und Niedergang der Römischen Welt* 2.36.6. Berlin: W. de Gruyter, 1992.

Mansfeld, J. "Diogenes Laertius on Stoic Philosophy." *Elenchos* 7 (1986) 296–382.

I. G. Kidd (1967)
Bibliography updated by Scott Carson (2005)

CYRANO DE BERGERAC, SAVINIEN DE

(1619–1655)

Savinien de Cyrano de Bergerac, a soldier, man of letters, and freethinker, was born in Paris, where he died thirty-six years later; he resembled only superficially the hero of Edmond Rostand's romanticized drama (1897). Hostile to the formal authoritarian education to which he had been subjected at the Collège de Beauvais, he was persuaded to serve in the army, where he gained a considerable reputation as a duelist and writer of verses. His military career came to an end when he was wounded at the siege of Arras in 1640. Between 1642 and 1651 he studied philosophy assiduously, with special stress on Pierre Gassendi and René Descartes, and was, according to some, a pupil of Gassendi himself. Descartes's principle of methodical doubt, Gassendi's rehabilitation of Epicurus, and the attendant influence of a newly translated Lucretius were all forces providing a common philosophical denominator which drew Cyrano closer to his fellow *libertins*—Gabriel Naudé, François La Mothe Le Vayer, and Molière, among others. At the same time he was emerging as a burlesque poet of consequence and a redoubtable political writer who first attacked and then defended the Machiavellian statecraft of Cardinal Mazarin. In 1652 he entered the service of the Duc d'Arpajon under whose protection he brought out in 1654 his *Oeuvres diverses,* which included the boldly rational *Lettre contre les sorciers* and a farcical comedy, *Le pédant joué,* from which Molière borrowed two passages for *Les fourberies de Scapin.* In 1654 Cyrano also published an intellectually challenging and ideologically daring tragedy, *La mort d'agrippine.* A falling beam, dislodged by accident—or perhaps intentionally—brought death a year later.

Cyrano's reputation as an intellectual libertine, propagator of subversive ideas, satirist of man and his foibles, and as a figure in the vanguard of scientific thought—already firmly established before 1655—received increased notoriety with the posthumous appearance of *L'autre monde, ou les états et empires de la lune et du soleil,* which described imaginary voyages to the moon and the sun, respectively. The first of the two parts of this work was made public in truncated form by the author's friend Le Bret in 1657. The second part, either unfinished or censored (the original manuscript has vanished), was published in 1762.

Despite borrowings and suggestions from a variety of sources, Cyrano's work, particularly when compared with that of many of his contemporaries, is strikingly original. Subscribing to the still little known and highly controversial Copernican theory, he adhered to the principle that all is relative in the universe and attacked religious and philosophical anthropocentrism. In fact he was the first to link closely together a criticism of the religion of Moses and the philosophy of Aristotle. In the man-machine–beast-machine debate, he stressed the idea of continuity among all living creatures. A forerunner of Denis Diderot's materialism, he outlined a calculation of probability according to which atoms, by means of chance and infinite time alone, could, in their innumerable combinations, create the organized world known to man. Furthermore he demonstrated an awareness of the forces of gravitation, the laws of which Isaac Newton was to discover and define several decades later. But he did not have Gassendi's gift for observation and experimentation or Descartes's aptitude for mathematics. He was more a popularizer of science than a true scientist. Indeed, he was the originator of science fiction.

The chief significance of Cyrano lies in the fact that he epitomized the general mental attitudes among the freethinkers of his period: enmity toward tradition, interest in ethical and scientific progress, and fondness for philosophical abstractions. As such he was eminently representative of those engaged in a protracted intellectual struggle which revealed the great trend of the French critical spirit—a spirit that was to gain increased momentum in the eighteenth century and to approach fulfillment with the publication of Diderot's encyclopedia.

See also Aristotle; Descartes, René Gassendi, Pierre; La Mothe Le Vayer, Francois de; Newton, Isaac.

Bibliography

WORKS BY CYRANO

Modern editions of de Bergerac's works include F. Lachèvre, ed., *Les oeuvres libertines de Cyrano de Bergerac* (Paris: E. Champion, 1921); H. Weber, ed., *Cyrano de Bergerac. L'autre monde* (Paris: Editions Sociales, 1959); and Richard Aldington, ed. and tr., *Voyages to the Moon and Sun* (New York: Orion, 1962).

WORKS ON CYRANO

See P.-A. Brun, *Savinien de Cyrano Bergerac. Sa vie et ses oeuvres* (Paris: A. Colin, 1893); René Pintard, *Le libertinage érudit dans la premiere moitié de XVIII[e] siècle* (Paris: Boivin, 1943); J.-J. Bridenne, "A la recherche du vrai Cyrano de Bergerac," *Information littéraire* (November–December 1953).

Otis Fellows (1967)

CYRENAICS

The Cyrenaics were a school of philosophy founded by Aristippus of Cyrene in the first quarter of the fourth century BCE. Although he had two sons, Aristippus designated his daughter Arete as his intellectual heiress. She in turn bestowed the succession on her son Aristippus, called "the Mother-taught." Apparently it was mainly he, a contemporary of Aristotle, who developed the more technical aspects of Cyrenaic doctrines. Cyrenaics were always included in lists of philosophical schools drawn up by the historians even though they had no fixed headquarters (unlike the Academy, the Lyceum, the Garden, etc.). There were several subschools referred to by the names of individuals, as Hegesiacs, Annicerians, and so on. They seem to have carried on the tradition of the Sophists of Socrates' time, being loosely associated itinerant teachers offering, for fees, instruction in general culture and on particular philosophical doctrines. Their pupils were supposed to learn from them how to live the good life, specifically, how to get along with anybody in any circumstances, as their founder put it.

The Cyrenaics were hedonists. They regarded it as self-evident that pleasure is the goal of life, for pleasure and avoidance of pain are what all living creatures seek by nature. The sage best knows how to attain a life of as many pleasures, interspersed with as few pains, as feasible and how to bear the pains when they come, as come they must. Unlike the Epicureans, the Cyrenaics regarded pleasure not negatively as mere absence of pain—dead people are in that condition—but as positive feeling, notably what is experienced in eating, drinking, and sex.

The younger Aristippus formulated a physiological analysis. There are three kinds of internal bodily motions: rough, smooth, and intermediate, which he compared to a tempestuous sea, a gently undulating sea, and doldrums respectively. Pleasure is the perception, by "internal feeling," of smooth motion; pain, of rough motion. Pleasures thus are particular present-moment happenings in individuals. These motions and their perceptions include satisfactions and dissatisfactions not so obviously internal to the body, such as gratitude and the pleasure one takes in the prosperity of one's country. The Cyrenaics noted that thought, not simply perception, enters into pleasure/pain distinctions. For example, watching a man really dying is painful, but to see an actor "die" on stage may be pleasurable. Nevertheless, plainly corporeal pleasures and pains are, in general, more intense, which is why they are prescribed as rewards and punishments.

Like his teacher Socrates, the elder Aristippus did not concern himself with natural science, which he deemed useless for furthering the good life. His grandson justified this rejection by advancing a skeptical theory of knowledge, of greater present-day interest perhaps than Cyrenaic ethics, for it is the closest ancient forerunner of modern phenomenalism and subjectivism. The only things one knows infallibly and certainly, he held, are one's feelings. These are internal states of the body. Things outside us produce the feelings—the Cyrenaics never doubted the external world—but one cannot know what those things are in themselves and how they operate. Something not yellow in itself may produce the sensation of yellow in a person with jaundice, and so on through the usual litany. Strictly, then, when in the presence of snow, one ought to say not "I see something white" but rather "I am being whitened" or, even better, "I am being affected whitely." Statements of these forms are the only ones knowable as absolutely true or false. Furthermore, if someone else in the same situation says—sincerely, let us assume—that he too is being affected whitely, then he speaks the truth, but from this it cannot be inferred that his feeling is identical to one's own. We apply the word "white" conventionally in the context of snow, but we have no way of knowing that the feeling it refers to is identical in everyone. Thus although the Cyrenaics did not explicitly raise the problem of other minds, in maintaining this possibility of an inverted spectrum they came close.

Cyrenaic skepticism helped also to justify Cyrenaic hedonism. Choices, as Socratics insisted, should be based on knowledge, not opinion or conjecture. But the scope of knowledge is limited to feelings, including pleasure and pain. Therefore, it is not only natural but rational to base our choices on pleasure and pain.

The most notable later Cyrenaics were Hegesias, Anniceris, and Theodorus, all active at the turn of the fourth to third century BCE. Hegesias, called "the Death Persuader," was an ancient Schopenhauer. From hedonism, surprisingly but straightforwardly, he deduced an unmitigated pessimism. The only good is pleasure; the only evil is pain. But as things are and must be, pains so predominate over pleasures that a life adding up to a pleasurable net balance is impossible. Therefore, suicide is eminently rational. Hegesias wrote a book, *The Man Starving Himself to Death*, in which the title character describes in detail the unavoidable ills of life. It was said that he lectured on this theme with such eloquence that some of his auditors killed themselves, whereupon the Greek king of Egypt, Ptolemy I Soter ("the Savior"), for-

bade him to deliver any more such addresses. Thus Hegesias perhaps had the dubious honor of being the first professor to have had his academic freedom curtailed by the government. He did not kill himself. There is a further similarity to Arthur Schopenhauer in his counsel "We should not hate people, but educate them."

Anniceris altered Cyrenaic hedonism by putting mental pleasures on a par with bodily ones, or even preferring them. Moreover, he softened the Cyrenaic egoism, declaring that the sage might forgo particular pleasures for the sake of friendship (as Epicurus maintained). He was credited with having ransomed Plato when that philosopher was for sale in the slave market of Aegina, though there are chronological and other difficulties with the story.

Theodorus, Anniceris's pupil, took free speech and the flouting of conventional pieties to an extreme even for the Greeks. He said that the sage would not fight for his country, for why should he put his wisdom at risk of extinction for the sake of the stupid masses? (Theodorus, unlike Hegesias, did not say no to life.) Aristippus had said that if all the laws were abrogated, the sage would continue to behave as before. Theodorus turned this proto-Kantian ethic all the way around, declaring that the sage might steal, commit adultery, even pillage temples if the occasion demanded—such acts being evil not by nature but only supposedly so to restrain the stupid. Extending his teacher's view on precedence among pleasures, for pleasure/pain he substituted joy/sorrow (primarily mental feelings) as the basic ethical contrast. He even went so far as to hold, as the Cynics did, that matters of the body are "indifferent." Threatened with crucifixion after he insulted Lysimachus, king of Macedon and in consequence ruler of Athens at the time, Theodorus contemptuously replied that it did not matter to him whether he rotted in the ground or in the air. (The threat apparently was not carried out.) But in place of Anniceris's amiability he reinstated hard-boiled egoism, claiming that the sage, being self-sufficient, has no need of friends.

With Diagoras of Melos and Euhemerus of Tegea, also a Cyrenaic, Theodorus was one of only three Greek thinkers who unequivocally proclaimed that there are no gods or demons at all, thereby earning the sobriquet "the Atheist." At a party in Athens, Hipparchia, wife of Crates the Cynic and a philosopher in her own right, chopped logic with him, saying, "What would not be wrong when done by Theodorus would not be wrong when done by Hipparchia. Now, it would not be wrong for Theodorus to strike himself. Therefore it would not be wrong for Hipparchia to strike Theodorus." Theodorus made no answer but instead pulled up her dress. (Doing so was not, or not merely, a display of classical male chauvinism, for it would not have been wrong for Hipparchia to pull up her own dress.)

The modifications by Anniceris and Theodorus brought the Cyrenaics so close to Epicurean views that it is not surprising that we hear no more of them as a distinct school after the first half of the third century BCE, when it was displaced by the Epicurean school.

See also Aristippus of Cyrene; Epicurus; Hedonism; Schopenhauer, Arthur.

Bibliography

Giannantoni, Gabriele. *I Cirenaici*. Florence, Italy: G. C. Sansoni, 1958.

Mannebach, Erich, ed. *Aristippi et Cyrenaicorum fragmenta*. Leiden, Netherlands: E. J. Brill, 1961.

Tsouna, Voula. *The Epistemology of the Cyrenaic School*. Cambridge, U.K.: Cambridge University Press, 1998.

Wallace Matson (2005)

D

DAI ZHEN

(1723–1777)

Dai Zhen, styled Dongyuan, was a critical-minded neo-Confucian in the Qing period. He is noted as a critic of neo-Confucianism of the Song (960–1279) and Ming (1368–1644) periods and made original contributions to the critical exposition of the philosophy of Mencius (c. 371–c. 289 BCE). Even though not known as a philosopher in his own time, his work in neo-Confucian criticism and exposition received more attention after Hu Shi wrote about his philosophy in the 1930s.

Dai Zhen was born in the Huizhou area of Anhui Province at a place known as Longfu (Tunxi) of Xiuling County. Although Huizhou was prosperous and produced outstanding academic talents, Dai Zhen, coming from a poor family with no academic traditions, received no good formal schooling. His success as a scholar derived from his own dedication to self-study.

Dai Zhen wrote *Yuan shan* (*Inquiry into Goodness*), his first philosophical treatise, from age 33 to age 41. With this work as a foundation, Dai Zhen then introduced quotations from the classics to support his philosophical points, and this later work, together with *Yuan shan*, became *Xuyan* (Prefatory words). Consolidating and concentrating on *Mengzi* (The book of Mencius), he sorted out and expanded *Xuyan* into a philosophical commentary on the key notions of *Mengzi*, producing *"Mengzi" ziyi shuzheng* (Commentary on the meanings of terms in *Mengzi*). This was his last work, which he completed at age 44.

Although Dai Zhen strongly objected to the abstract use of principles (*li*), he did not deny the importance of reason when applied correctly to concrete matters. But how does one acquire an understanding of reason and principle? The answer is twofold: by correctly reading the classical texts on which doctrines of moral reasoning are based and by clearly reflecting on what reason and principle concretely signify. One must first authenticate the classical texts and semantically and philologically determine their meaning. Only then can one correctly read and interpret them. In this sense textual criticism is highly relevant to understanding the principles and moral reasoning embodied in the texts of the classical philosophers.

The usefulness of textual criticism for understanding reasons and principles is, of course, no explanation of the rise of textual criticism in the Qing period. Most well-known textual critics were not interested in discovering or rediscovering the principles and moral reasoning of

the classical texts. But for Dai Zhen, textual criticism is essential for such discovery. He wrote, "The ultimate idea of the classics is the Way. We use words [*ci*] to understand the Way. We use the linguistic study of the text [*xiaoxue wenzi*] to understand the words. From such study we come to understand the discourse, from the discourse we come to understand the mind and intent of the ancient sages" (1995, p. 378). Hence, for Dai Zhen, the purpose of textual criticism of a classic is to retrieve the original meaning of the text. On this basis one can then come to understand the moral reasoning and principles behind the texts.

Dai Zhen was perhaps the first modern Chinese scholar to formulate a textual hermeneutics that combines historical linguistics with philosophical reflection for reading classical texts. He was also one of the earliest pioneers in philosophical hermeneutics in the whole world.

Another important philosophical contribution of Dai Zhen's was his objection to separating reason from feelings and desires in Song and Ming neo-Confucianism and hence his stress on understanding in terms of human feelings and desires. This position came from his deep appreciation of the naturalistic cosmology of the *Yijing* (Book of changes), where he found sources of human nature and human reason. He took the productivity of life (*shengsheng*) as the most basic fact of reality. The purpose of this productivity gives purpose to the interplay of yin and yang and is called the Way, he explained. This interplay results in the unceasing transformation of life and the ordering of things (*tiaoli*) in heaven and on earth. From the productivity of life and the ordering of things Dai Zhen derives the virtues of humanity (*ren*) and moral reason (*yi*), which he regards as inherent in these two processes.

See also Chinese Philosophy. Confucius; Cosmology; Human Nature; Hu Shi; Mencius; Reason.

Bibliography

WORKS BY DAI ZHEN

Tai Chên's "Inquiry into Goodness." Translated by Chung-ying Cheng. Honolulu, HI: East-West Center Press, 1968.

Dai Zhen wenji. Beijing: Zhonghua shuju, 1980.

Dai Dongyuan quanji. Beijing: Qinghua daxue chubanshe, 1993.

"Gujing jie Gou Chen xu." In *Dai Zhen quanshu*. Vol. 6. Hefei, China: Huangshan shushe, 1995.

WORKS ON DAI ZHEN

Hu Shi. *Dai Dongyuan de zhexue*. Taipei: Yuanliu chuban gongsi, 1968.

Torbjoern Loden. "Dai Zhen's Evidential Commentary on the Meaning of the Words of Mencius." *Bulletin of the Museum of Far Eastern Antiquities* 60 (1988): 165–313.

Yu Yingshi. *On Dai Dongyuan and Zhang Xuecheng*. Hong Kong, 1976.

Chung-ying Cheng (2005)

D'ALEMBERT, JEAN LE ROND

See *Alembert, Jean Le Rond d'*

DAMASCIUS

(c. 462–c. 538)

Damascius was a neoplatonic philosopher and the last head of the Academy in Athens. He was born around 462 CE in Damascus and studied in Alexandria and Athens. In 515 he became head of the Academy, which, through his reforms and teaching, would see a final flourishing. After the closing down of the Academy by Emperor Justianus in 529, Damascius and six colleagues went into exile at the court of King Chosroes in Persia. They returned in 532, having been granted the freedom to continue their philosophical work. Damascius died in Syria sometime after 538.

His writings include the "Life of Isidore" (Isidore was his teacher and predecessor), in which he offers a privileged insight in the history of the pagan Platonic school in the fifth century CE; and commentaries on Plato (preserved are those on the *Parmenides*, the *Philebus*, and the *Phedo*). He is, however, mainly known for his treatise "On the First Principles" (*De principiis*), an ingenious philosophical speculation about the first causes of all things.

Damascius had no ambition to develop a better metaphysical system than his predecessors. His own thought is primarily aporetic: He raises critical questions in the margin of the doctrine of the principles, as it had been developed in the neoplatonic tradition, and confronts the doctrine with all sorts of difficulties. When he ventures a solution—and on many issues he can be original (for instance, his doctrine on time)—he again puts that solution into a question with new aporias (or doubts). Damascius's work is in many ways a critical analysis of the position of Proclus, who, in his view, was too preoccupied with logical coherence and system building. He raises questions about the One and multiplicity, about procession and return, about triads of principles,

and about concepts such as power—not in order to discredit all philosophical discourse skeptically, but to clarify what is inadequate in the formulations of his predecessors.

The most fundamental aporia is discussed at the beginning of the treatise: Is the first principle itself a part of the whole of which it is the principle? But if it is a part, how could it still have the status of a principle? If it is outside the whole, how can we understand that the whole originates from it? The first principle, it seems, is neither principle nor cause, nor does it fit in any other category used to explain relations between beings: It is an ineffable "nothing" we have to postulate beyond the one whole. This "ineffable" is even beyond the "One" that is the first principle of all things.

More than any other Platonic philosopher, Damascius is aware of the precarious nature of all rational discourse when people deal with questions that go beyond the limits of what they can experience. More than any other, he explored the boundaries of rationality; he tried, by all means, to say what could not be said, because about the first principles one can only speak using analogies and "indications" (*endeixeis*), which are as such unfitting to indicate divine realities.

Damascius's sharp critical mind does not, however, bring him to skepticism. If philosophical systems remain tentative and fragile, there is also the mythological tradition and religious practice, to which Damascius remains devoted. Damascius is, together with Proclus, our main source on Chaldean and Orphic theologies. In many ways his work is a wonderful swan song of pagan Hellenism.

See also Greek Academy; Neoplatonism; Plato; Proclus.

Bibliography

WORKS BY DAMASCIUS

Damascius: Traité des premiers principes. 3 vols, edited by Leendert Westerink and Joseph Combès. Paris: Les Belles Lettres, 1986–1991.

Damascius: Commentaire du Parménide de Platon. 4 vols, edited by Leendert Westerink and Joseph Combès. Paris: Les Belles Lettres, 1997–2003.

WORKS ABOUT DAMASCIUS

Athanassiadi, Polymnia, ed. *Damascius: The Philosophical History*. Athens: Apamea Cultural Assocation, 1999.

Hoffmann, Philippe. "Damascius." In *Dictionnaire des philosophes antiques*. Vol. 2, 541–593. Paris: CNRS, 1994.

Carlos Steel (2005)

DANTE ALIGHIERI

(1265–1321)

Dante Alighieri, the author of the *Divine Comedy,* was born in Florence of a middle-class family with some pretensions to nobility. It is likely that he frequented the church schools, and he probably spent a year at the University of Bologna. He fought in the battle of Campaldino (1289) and a few years later married Gemma Donati, by whom he had at least three children. He took part in the government of his native city, serving on various city councils (1295–1297, 1301), as prior (1300), and as ambassador to San Gimignano (May 1300) and later to Rome (October 1301), where his mission was to negotiate with the pope to bring about a just peace between the warring factions of White Guelphs and Black Guelphs. Aided by the intervention of Charles of Valois, the Blacks took over the city and Dante, a White, went into exile. He wandered from court to court of medieval Italy, with especially long sojourns at Verona and at Ravenna, where he spent the last three years of his life. He seems to have served his patrons as adviser and on occasion specifically as ambassador; it was after an embassy to Venice on behalf of Guido da Polenta, Lord of Ravenna, that the poet died.

By choice Dante might well have devoted himself to political life: circumstance deprived him of this opportunity and constrained him to put his great gifts to the service of letters; his masterpiece, the *Divine Comedy,* is generally regarded as the supreme poetic achievement of the Western tradition and has assured his fame. His *Vita Nuova* is the story of his idealistic love for Beatrice, presumably of the Portinari family, who married Simone de' Bardi and died in 1290. The *Convivio*, composed after the author went into exile, is a didactic work; the *De vulgari eloquentia* is a milestone in the history of linguistics, being the first serious study of a vernacular tongue; and the *De monarchia* is the vehicle for Dante's expression of his political theory. Mention should also be made of his *Rhymes,* a collection of verses of varying kinds—some purely lyrical, some moralistic, and some, one might say, philosophical.

To what extent Dante may properly be considered a philosopher depends on one's definition of the term. Richard McKeon does not consider him such "by the crucial test that, despite the philosophic doctrines that crowd his poems, scholars have been unable to agree concerning what his attitude toward the philosophers he uses is." But this is to make a very special category of philosophers. The best statement of Dante's attitude is found at the

beginning of the *Convivio* (Banquet), where he represents himself not as one of the great (scholars and philosophers) who actually sit at the banquet table but rather as one who, sitting at their feet, passes on to others the crumbs that he is able to pick up. This would make him on the one hand at least an eager student of philosophy and on the other what we should now call a popularizer, if the term may be used without disparagement. And within the great area of philosophy his major interest was in ethics and politics. Let us concede that in the field of pure speculation his mind was alert and curious rather than original. Like his contemporaries he was for the most part content to follow Aristotle as interpreted by Thomas Aquinas, with recourse to what he thought of as "Platonic" where it suited him. His use of his authority, his stature as a poet, and his influence, which still endures, make it worthwhile to study his philosophical posture in some detail.

THE *VITA NUOVA* AND *CONVIVIO*

If a drive to seek eternal truth, permanent universals, and order in things is the proper attribute of a philosopher, as it would seem to be, then Dante's claim to the cherished title is reasonable. Perhaps his first work, the *Vita Nuova,* is the most dramatic example of this precisely because, paradoxically, it is not a philosophical work at all. It is a love story of intimate and personal nature, grounded, it would seem, in historical fact but taking on the air of a spiritual parable; its immediate sources are not in works of philosophy but in the love cult of the Middle Ages. Yet the construction and the apparatus betray a disciplined intent; the prose and poetry are mingled in a strict architectural pattern; and each of the poems is followed by an analysis composed in the tradition of Scholasticism. Digressions on the nature of personification and the meaning of certain terms are evidence of what one might fairly call the philosophical manner. Beatrice herself becomes in the course of the confessional narrative something very close to a theological and thus a quasi-philosophical concept.

It is, however, the *Convivio* that is the most purposefully "philosophical" of Dante's canon. It was inspired, the author tells us, by the reading of Cicero and Boethius, and Dante in fact seems to see himself as having much in common with the latter, also a victim of political injustice, and as turning to the same source for consolation. It is noteworthy, too, that Dante, like Boethius, attempts—consciously, one suspects—to set philosophy free from its entanglement with Christian theology. His definition of philosophy in the third tractate goes back to Pythagoras, and in Book IV, in the course of enumerating the virtues appropriate to the successive ages of man, he turns to the pagans (such as Aeneas and, very strikingly, Cato), to exemplify such virtues. All but startling is his eulogy: "And what earthly man was more worthy than Cato to signify God? Truly none." Such an attitude toward the "ideal pagan" dramatizes the author's celebrated exposition of the two beatitudes (II, 4): one in speculation and contemplation, the other in proper conduct of the active life; the former is "higher" than the latter, which, however, clearly is not "subordinate": "It is typical of Dante," says Étienne Gilson, "to base the autonomy of an inferior order on its very inferiority."

In this connection the plan of the *Convivio* (if it may be called a plan, for, unlike most of Dante's works, the book seems to have grown of itself) is very revealing of the author's concept of the uses, if not the nature, of philosophy. The first tractate is highly personal, stating that the genesis of his interest was his need for consolation in his exile and his feeling that his "image" in Italy had suffered somewhat from the youthful and impassioned portrait that emerged from the pages of the *Vita Nuova.* In the second tractate he avows that in effect philosophy, "the fairest and noblest daughter of the universe," is the new lady who has replaced Beatrice in his heart. In the third tractate he discusses the meaning of philosophy, which he finds to signify "love of and zeal for wisdom," adding that philosophy has "as its subject understanding and as its form an almost divine love of the thing understood." Presumably "understanding" can be applied to the various fields of study Dante had enumerated in the second tractate, composing an ingenious correlation between the sciences and the heavens of the Ptolemaic system. Of these branches the highest for any medieval theologian (theology itself is in the empyrean, beyond the physical cosmos) would be metaphysics, but it is significant that Dante brackets it with physics in the starry heaven and puts ethics in the loftiest physical sphere, the primum mobile, morality being "the science that disposes us rightly for the other sciences" even as the crystalline heaven sets in motion all the other spheres. In fact, the largest part of the work, the fourth treatise, is given over to a study of true nobility, its source and its effects.

Dante finds this human excellence to be not the Aristotelian "inherited wealth and good manners" but rather a God-given grace, the nature of which is evident in its fruits. The fruits, which are enumerated in chronological order, are all of such a nature as to be properly called social virtues. Dante's ideal is not a mystic or a visionary but, in the best sense of the term, a man of the world, liv-

ing in a community and serving it to the best of his ability—certainly an Aristotelian concept. Only in the stage of "decrepitude" does Dante say that the good man's thoughts should turn to God and the afterlife, and even this passage, beautiful as it is, has about it a tone more pagan than Christian. It is noteworthy that all the men chosen to exemplify the appropriate virtues are men of action, in many cases pagans but also including such ambiguous characters as Lancelot and Guido da Montefeltro, the *condottiere.* Thus the *Convivio*, dedicated to the glorification of philosophy, ends by being a rule of good living, high-minded, to be sure, but practical as well. Noteworthy too is the rather lengthy excursus of Book IV (Chs. 4–5) that is inserted to justify the Roman Empire. Dante finds historical correspondences between the empire and the church, affirms that Christ chose to come to Earth at the time the world was best governed and at peace (that is, under Augustus), and concludes with a panegyric to Rome. This is the more interesting because some of his data are traceable to St. Augustine, whose view of imperial Rome was quite opposite.

THE *DE MONARCHIA*

The *De monarchia,* developing the latent and the tentative attitudes of the *Convivio,* may well contain Dante's most original contribution to philosophical thought. Written, it seems likely, either during or shortly after Henry VII of Luxembourg's descent into Italy (c. 1313), it is an eloquent defense of the imperial cause or, more accurately, principle. The work is divided into three parts: in the first Dante shows the necessity for the rule of one monarch in temporal affairs; in the second he argues that for historical reasons such a monarch should be the Roman emperor; and in the third he defends the thesis that the emperor, although he owes deference to the pope, should not be subordinate to the pontiff in temporal matters.

It is the first book that is the most fascinating to the student of Dante the philosopher. Briefly, the main argument is that peace is a necessity if humanity is to actualize its potential intellect in the highest degree; and there can be no assurance of peace, national rivalries being what they are and greed being as strong as it is, unless the world is governed by one prince, supreme above all nations and beyond the temptations of *cupiditas.* In the course of defining the collective potential intellect, Dante invokes the name of Averroes, thus laying himself open to a charge of heresy (and indeed the *De monarchia* was solemnly burned and remained on the Index for many years).

Gilson, however, has well made the point that the collective potential intellect of humanity as conceived by Dante was not a "being," as was the "possible intellect" (or kind of oversoul) of Averroes, but rather a "community." Indeed, in the course of his arguments in the first book Dante follows Thomistic reasoning, but unlike Thomas, who never so much as mentioned the word *emperor,* he applies it to secular purposes. Conceding the superiority of contemplation over action and, by inference, of the spiritual over the temporal, he nevertheless stresses the importance of the machinery necessary to perfect the fulfillment of man's proper endowment in the active life and his happiness in this world. So too at the end he readily concedes that the emperor owes the pope the respect of a younger brother, but while thus indicating that the spiritual life is superior, he seems also to imply that it is separate and independent; both pope and emperor would, in his theory, derive their authority directly from God. The result is in fact a kind of political facet of the Averroistic double truth, as contemporary critics were quick to point out. Gilson, for whom Dante is no Averroist, nevertheless commends him for seeing clearly "that one cannot entirely withdraw the temporal world from the jurisdiction of the spiritual world without entirely withdrawing philosophy from the jurisdiction of theology" and adds that Dante's perception of this fact gives him "a cardinal position in the history of mediaeval political philosophy." In this sense and with a practical intent characteristic of Dante, the *De monarchia* reaffirms the underlying thesis of the *Convivio.*

THE *DIVINE COMEDY*

It has been argued by some critics that the *Divine Comedy* is in essence a repudiation of the secular and independent *Convivio* and *De monarchia* and is evidence of a kind of "Conversion" of the poet, resulting either from some inner crisis or from his despair at the defeat of Henry VII. Perhaps if we say that in the *Comedy* the substance of the earlier works is utilized as a preparation for the vision, a basis for the mystic superstructure rather than as a finality in itself, we may speak of "conversion," but not, in the opinion of this writer, if the word carries any suggestion of rejection. It is true that the devotional element is novel and important: the intercession of the Virgin Mary makes it possible for the poet to undertake the supernatural journey and to enjoy the vision that crowns it. The vision itself is of a mystical nature, adumbrated perhaps in the *Vita nuova* but totally absent from the "philosophical" works. Concern with purely theological matters—the Incarnation, predestination, divine justice, and the like—

bulk large in the *Comedy,* which also contains (in *Paradiso* XI) a very interesting example of the *contemptus mundi* posture, otherwise quite uncharacteristic of Dante. The poet is also very careful to point out the error of the belief in Averroistic oversoul (*Purgatorio* XXV). Such elements have led to discussion of Dante's Augustinianism as opposed to his Thomism. (T. K. Swing has argued that in his manipulation of these doctrines "Dante is the first to accomplish a consistent elucidation of the teleological destiny of the Christian soul through a metaphysical scheme.") It is true that the presence of St. Bernard as Dante's last guide and, as it were, sponsor for his ultimate vision, gives dramatic emphasis to the Neoplatonic or Augustinian strain. But if the substitution of rapture for reason represents the victory of Augustine over Thomas, it also carries us beyond the limits of philosophy and perhaps out of the area of our proper concern here.

We may yet affirm, in the face of all such elements as noted, that the *Comedy* is, in the author's intent, primarily an exposition of ethics; the letter to Can Grande specifically defines it as having for its subject "man, liable to the reward or punishment of Justice, according as through the freedom of the will he is deserving or undeserving." And in this area the frame of reference is, as it was in the *Convivio,* Aristotelian and Thomistic—not without some original sallies of Dante's own. The presence in the *Paradiso* of the Latin Averroist Siger of Brabant, for example, may be interpreted as an affirmation of the autonomy and dignity of the "contemporary profane science" (Pierre Mandonnet) of Aristotelian philosophy. But from the point of view of ethical investigation, the *Inferno* is the most interesting part of the work, for here, dealing not with the way of salvation, which is no longer possible to the damned, nor with the ultimate doctrines, interesting only to saints, Dante is in a sense free to formulate his own code of morality. Clearly his inclusion of pagans and other non-Christians in hell indicates his intent to establish a code of behavior for all men; his hell is nonsectarian, broadly speaking. His main divisions of incontinence, violence, and fraud are ingeniously worked out from a combination of Aristotle, Cicero, and Thomas; interesting too is his creation of the "vestibule" for the lukewarm spirits and his peopling of the limbo with the souls of the virtuous pagans. Nor does the "converted" Dante abandon his appreciation of the second beatitude; not only do the pagans in limbo enjoy quite a comfortable immortality but Cato, so much revered in the *Convivio,* reappears as the guardian of purgatory, where he symbolizes free will; and, most startling of all, in the heaven of Jupiter the Trojan Ripheus is shown as an example of the "baptism of desire" that would make it possible for a good man, totally ignorant of the Mosaic or Christian message, to win salvation. To be sure, this is rare and does not avail to save Vergil or Aristotle, but on the other hand salvation in Christian terms also is ultimately a matter of predestined grace: without being unorthodox, Dante, in the example of Ripheus, has revealed his deep concern for ultimate justice. Indeed, the analysis of sin in the *Inferno,* as Kenelm Foster has pointed out, has its genesis in a conception of justice and presupposes society. The souls in the *Inferno* have "injured" others, have broken the social fabric in one way or another; even the heretics seem to be there because they have misled their followers rather than because of their own arrogant pride (a sin not specifically classified in the *Inferno*). We may also remark that Dante's concern for the good life on Earth does not desert him: The theory of the two "suns" necessary for the proper illumination of humankind reappears in the *Purgatory*; the emperor is glorified (a reserved seat awaits Henry VII in the celestial rose); and certain cabalistic prophecies indicate Dante's hope for a *dux* who will lead the temporal world back to order and sanity. "The *Divine Comedy* is as much a political as it is a religious poem," says A. Passerin d'Entrèves, and surely in that climactic work both politics and religion are seen *sub specie philosophiae.* If Dante is not a true philosopher, he is certainly a magnificent amateur.

See also Aristotelianism; Aristotle; Augustine, St.; Augustinianism; Averroes; Boethius, Anicius Manlius Severinus; Cicero, Marcus Tullius; Continental Philosophy; Gilson, Étienne Henry; Love; Neoplatonism; Pythagoras and Pythagoreanism; Siger of Brabant; Thomas Aquinas, St.; Thomism.

Bibliography

WORKS BY DANTE

Convivio. Translated by Philip Wicksteed. London, 1903.

Epistolae. With English translation by Paget Toynbee. Oxford, 1920.

Le opere di Dante. Florence: R. Bemporad, 1921; 2nd ed., 1960.

The Portable Dante. Edited by Paolo Milano. New York: Viking, 1947. Selected translations.

Divine Comedy. Translated by H. R. Huse. New York: Rinehart, 1954. A useful, prose, line-for-line version.

Monarchy. Translated by Donald Nicholl. London: Weidenfeld and Nicolson, 1954.

WORKS ON DANTE

Davis, C. T. *Dante and the Idea of Rome.* Oxford: Clarendon Press, 1957.

Entrèves, A. P. d'. *Dante as a Political Thinker.* Oxford: Clarendon Press, 1952.

Foster, Kenelm. "The Theology of the 'Inferno.'" In *God's Tree.* London: Blackfriars, 1957.

Gilson, Étienne. *Dante et la philosophie.* Paris: J. Vrin, 1939. Translated by David Moore as *Dante the Philosopher.* New York: Sheed and Ward, 1949.

Mandonnet, Pierre. *Dante le théologien.* Paris, 1935.

Mazzeo, J. A. *Medieval Cultural Tradition in Dante's Comedy.* Ithaca, NY: Cornell University Press, 1960.

Mazzeo, J. A. *Structure and Thought in the Paradiso.* Ithaca, NY: Cornell University Press, 1958.

McKeon, Richard. "Poetry and Philosophy in the Twelfth Century." In *Critics and Criticism: Ancient and Modern,* edited by R. S. Crane et al. Chicago: University of Chicago Press, 1952.

Nardi, Bruno. *Dante e la cultura medievale.* Bari, Italy, 1949.

Santayana, George. *Three Philosophical Poets.* Cambridge, MA: Harvard University Press, 1910; New York, 1953. Latter edition is paperback.

Swing, T. K. *The Fragile Leaves of the Sibyl: Dante's Master Plan.* Westminster, MD: Newman Press, 1962.

Thomas Goddard Bergin (1967)

DANTO, ARTHUR

(1924–)

Arthur Danto's contributions to the philosophy of art have been shaped by his experiences as art maker, art critic, and art lover. He earned a bachelor's degree in fine arts from Wayne State University in 1948. For the next decade, his woodcuts were shown in such important venues as the Art Institute of Chicago, the Detroit Institute of Arts, Los Angeles County Museum, the Museum of Fine Arts in Boston, and the National Gallery of Art in Washington, D.C. During this period of active art-making, he completed a doctorate in philosophy at Columbia University in 1952 and began his half-century-long appointment in the Columbia philosophy department.

One of Danto's central aims for the first thirty years of his career was to render the ideas of nineteenth- and twentieth-century continental philosophers such as Hegel, Nietzsche, and Sartre accessible and useful to analytical philosophers. Danto's writing about these figures is clear and often critical. He has also published penetrating overviews of core fields such as the philosophy of science, the philosophy of action, the theory of knowledge, the philosophy of history, and philosophical psychology.

Nevertheless, this philosopher remained an artist and passionate art lover. He had come to New York to study philosophy just when that city emerged as the center of innovative achievement in the art world. The art Danto encountered in the museums and galleries he frequented was conceptually challenging.

Andy Warhol's 1964 work *Brillo Box* provoked a key question: What makes *Brillo Box*—a replica of the box used to ship packages of Brillo pads—a work of art, suitable for display in a museum or gallery, when perceptually indiscernible objects—the actual Brillo boxes created en masse by the manufacturer—are relegated to grocery displays or storerooms? This is a philosophical query, but also an integral part of experiencing *Brillo Box* as art, for the art lover encountering *Brillo Box* is initially transfixed by questions about its status.

Danto's famous essay "The Art World" (1964) initiated an answer that he refined and elaborated over the ensuing fifteen years. Danto asked how commonplace objects that never could have been art in earlier times not only had gained the possibility of being art by 1964 but also appeared to be the art necessary for that time. Danto presumes that philosophy should accept, not correct, the phenomena of art-world practice and discourse. Therefore, the traditional questions of philosophy of art and philosophical aesthetics must be transformed to fit the art world's realities.

Danto's more fully elaborated position, first presented in full in *The Transfiguration of the Commonplace* (1981), is that art history and art theory contribute experiential (albeit not sensuous) properties to certain objects. These properties make the difference in experiencing objects as art. Absent being experienced at the appropriate art-historical moment, and through the lens of compelling art-theoretical understandings that offer illuminating interpretive hypotheses, objects do not rise to the status of art.

Seeing affinities between Danto's focus on art-world practice and his own view that it is artists, critics, and curators who decree which objects should be treated as art, George Dickie heralded the advent of the institutional theory of art. Danto's view differs from Dickie's in many ways, however. For example, a key idea in Danto's, but not in Dickie's, thought is that art distinctively embodies meaning, or at least embodies questioning.

Danto takes modern art's history to be a quest for answers about the general (transhistorical) nature and identity of art. Art in our time has achieved a philosophical self-consciousness that acknowledges rather than veils ontological questions about its own nature. But in pursuing its own ontology, art transcends its limits and is transfigured into philosophy. Persisting in this transgressive aim, art subsequently executes its own end, turning its back on philosophical anxiety about what art must be. Art thereby is liberated to place itself freely in the service of a multiplicity of values rather than to embrace a single

value that is uniquely aesthetic. Danto's theory of the end of art is empirical, not prescriptive. He explains where art has arrived, and why, rather than directing where art should go. In such a pluralistic age as our own, when everything is possible, what principles should guide the art critic? This question, traditionally a concern of philosophical aesthetics, is of special interest to Danto because of another artworld role he fills, that of art critic. In 1984 Danto became the art critic for *The Nation* magazine. Much of his writing since that time has been criticism of works of art or reflections on art criticism. His *Encounters and Reflections: Art in the Historical Present*, a collection of art criticism, won the National Book Critics Circle Prize for Criticism in 1990.

In general, Danto's art criticism is about understanding artistic processes, not assessing aesthetic outcomes. Some philosophers fault him for stamping his philosophy of art with his style of art criticism and thereby giving artistic considerations priority over aesthetic values. Others praise him for developing a philosophical theory of art into which enlightening art criticism is tightly woven. Danto seeks to explain rather than steer the direction of art. Art criticism, as Danto understands the practice, deploys artistic judgment to detect an object's content and explain how the object embodies or presents what it is about. Yet Danto himself offers no developed philosophical analysis of artistic embodiment, neither of the process nor of the criteria of success. His signature stance is to observe from the intersect of philosophy and criticism. His strategy is to gently and genially compel art criticism to confront its own implicit abstractions and generalizations, while persuasively propelling philosophy to engage with the puzzling particulars of the world of art.

See also Art, Expression in; Art, Style and Genre in.

Bibliography

Danto, Arthur C. *After the End of Art*. Princeton, NJ: Princeton University Press, 1997.

Danto, Arthur C. "The Art World." *Journal of Philosophy* 61 (1964): 571–584

Danto, Arthur C. *Encounters and Reflections: Art in the Historical Present*. New York: Farrar, Straus & Giroux, 1990.

Danto, Arthur C. *Narration and Knowledge, Including the Integral Text of Analytical Philosophy of History*. New York: Columbia University Press, 1990.

Danto, Arthur C. *The Transfiguration of the Commonplace*. Cambridge, MA: Harvard University Press, 1981.

"Danto and His Critics: Art History, Historiography, and After the End of Art." *History and Theory* 37 (1998): 1–143.

Dickie, George. *Art and the Aesthetic: An Institutional Analysis*. Ithaca, NY: Cornell University Press, 1974.

Rollins, Mark, ed. *Danto and His Critics*. Cambridge, MA: Basil Blackwell, 1993.

Anita Silvers (2005)

DAOISM

See *Chinese Philosophy: Daoism; Laozi; Mysticism, History of*

DARWIN, CHARLES ROBERT

(1809–1882)

Charles Robert Darwin, the British biologist whose theory of organic evolution revolutionized science, philosophy, and theology, was born at Shrewsbury. He attended the universities of Edinburgh and Cambridge but was not attracted by his medical studies at the first or by his theological studies at the second. Near the end of his undergraduate days he formed a friendship with J. T. Henslow, professor of botany at Cambridge, "a man who knew every branch of science" (*Autobiography of Charles Darwin*). This association, together with an enthusiasm for collecting beetles and a reading of works by Wilhelm von Humboldt and John Herschel, generated in him "a burning zeal to contribute to the noble structure of Natural Science." The opportunity to do so on a large scale arose when Henslow secured for him the post of naturalist "without pay" aboard the H.M.S. *Beagle,* then about to begin a long voyage in the Southern Hemisphere. Thus, between 1831 and 1836 Darwin was able to make extensive observations of the flora, fauna, and geological formations at widely separated points on the globe. This experience determined the course of his life thereafter and laid the foundation for many of his fundamental ideas. On his return he lived in London for six years, where he became acquainted with leading scientists of the day. Sir Charles Lyell, Sir Joseph Hooker, and T. H. Huxley were among his most intimate friends. In 1842 he took up residence at Down, a secluded village in Kent. Here, during the forty years until his death, he conducted the researches and wrote the works that made him famous. He was buried in Westminster Abbey close to the grave of Sir Isaac Newton.

Darwin's productivity, despite recurrent bouts of illness, was prodigious. His publications ranged over such diverse subjects as volcanic islands, coral reefs, barnacles, plant movement, the fertilization of orchids, the action of

earthworms on the soil, the variation of domesticated animals and plants, and the theory of evolution. Even if he had never written *The Origin of Species* (1859) and *The Descent of Man* (1871), he would still be regarded as one of the great biologists of the nineteenth century. Of course, it was these two books that made him the initiator of a revolution in thought more far-reaching than that ushered in by Nicolas Copernicus. He established beyond reasonable doubt that all living things, including man, have developed from a few extremely simple forms, perhaps from one form, by a gradual process of descent with modification. Furthermore, he formulated a theory (natural selection), supporting it with a large body of evidence, to account for this process and particularly to explain the "transmutation of Species" and the origin of adaptations. As a result, the biological sciences were given a set of unifying principles, and man was given a new and challenging conception of his place in nature.

It was characteristic of Darwin that he came to these conclusions by his own observations and reflections. When he embarked on the *Beagle,* his outlook was "quite orthodox." He accepted without question the fixity of species and their special creation as depicted in Genesis. Doubts began to arise in his mind during the ship's visit to the Galápagos Archipelago in 1835, when he noticed that very small differences were present in the so-called species inhabiting separate islands. The doubts were reinforced by his observation of fossils on the Pampas and the distribution of organisms on the South American continent as a whole. He was "haunted" by the idea that such facts "could be explained on the supposition that species gradually became modified." In July 1837 he "opened his first notebook" to record additional facts bearing on the question, but it was not until he happened to read Thomas Robert Malthus's *Essay on Population* in October 1838 that he found an explanatory theory from which the above "supposition" followed. He then proceeded to formulate the principle of natural selection, which is simply "the doctrine of Malthus applied with manifold force to the whole animal and vegetable kingdoms." Darwin never professed to have invented the idea of organic evolution, of the mutability of species, or even of natural selection. What he did profess was to have produced the first scientific proof that these ideas apply to the living world.

Unlike some lesser men of science, Darwin was not inclined to rush into print in order to establish a proprietary right to his theory. His modesty and single-minded desire to find out the truth forbade any such action. Accordingly, the theory underwent several preliminary formulations. It was first set down in a short abstract in 1842 and two years later was expanded into an essay that both Lyell and Hooker read. Early in 1856 Lyell advised Darwin to write a full-length account of his views. It was when this manuscript, which would have been "three or four times as extensive" as *The Origin of Species,* was about half finished that Alfred Russel Wallace's paper, which contained virtually the same ideas that Darwin was working out, arrived at Down from the Malay Archipelago. The resulting crisis was resolved by having a joint communication from the two men read at a meeting of the Linnaean Society on July 1, 1858. Between September of that year and November 1859, Darwin "abstracted" the large manuscript and produced his classic. *The Origin of Species* appeared on November 24 in an edition of 1,250 copies, all of which were sold on the first day. Ultimately, six editions containing many revisions were published.

Despite the interest that *The Origin of Species* excited, it was by no means universally approved at first. In the scientific world support for it came from Darwin's friends, but others expressed opposition that often took the form of objections to the modes of explanation and proof employed in the work. Darwin's use of historical or genetic explanations, his implicit adoption of statistical conceptions ("population thinking," as it is now called), and his practice of introducing conjectures or "imaginary illustrations" to buttress his argument were repugnant to biologists who held that scientific explanation must consist in bringing directly observed phenomena under general laws. Believers in this oversimplified model also disliked his notion of "chance" variations and his repudiation of "any law of necessary development." Before long, however, the cumulative force of Darwin's arguments, augmented by the case put forward in *The Descent of Man,* convinced the great majority of biologists, so that opposition from this quarter had disappeared by 1880.

The popular reaction to Darwin's theory focused on its religious and ideological implications. These were recognized to be hostile to the Establishment. Hence, Darwin found himself enthusiastically supported by radicals, rationalists, and anticlericals and vehemently attacked by reactionaries, fundamentalists, and priests. He shrank from entering into this controversy, which was altogether distasteful to him, but T. H. Huxley, who enjoyed crossing swords with theologians, took a different stand. Appointing himself "Darwin's bulldog," he relentlessly pursued such antievolutionists as Bishop Wilberforce and W. E. Gladstone. His efforts had a good deal to do with creating the image of Darwin as an enemy of the Bible, the church, and Christianity.

This image was, in fact, fairly close to the truth. Darwin's religious beliefs, as he relates in his *Autobiography,* underwent a change from naive acceptance of Christian doctrines to reluctant agnosticism. In the two years following his return from the voyage of the *Beagle* he was "led to think much about religion." Doubts were engendered in his mind about the historical veracity of the Gospels, the occurrence of miracles, and the dogma of everlasting damnation of unbelievers (which he calls "a damnable doctrine"). By reflection on such matters he "gradually came to disbelieve in Christianity" and wondered how anybody could wish it to be true.

A similar erosion occurred in connection with his belief in the existence of a personal God. When he wrote *The Origin of Species,* Darwin accepted a vague theism or deism. In the last chapter he speaks of laws having been "impressed on matter by the Creator" and of life's powers "having been breathed by the Creator into a few forms or into one." He was thus able at the time to deny that it was his intention "to write atheistically." Yet it was also clear to him that the theory of natural selection exploded the old argument for theism based on the presence of design in the organic world. The vast amount of suffering and misery that exists seemed to him a strong argument against any belief in a beneficent First Cause. He had moods in which it seemed difficult or impossible to conceive that "this immense and wonderful universe, with our conscious selves, arose through chance." In the end, however, he concluded "that the whole subject is beyond the scope of man's intellect. … The mystery of the beginning of all things is insoluble by us; and I for one must be content to remain an Agnostic."

Darwin's reflections on religion, although not systematic, provide a good example of his intellectual integrity. "I have steadily endeavored," he wrote in his *Autobiography,* "to keep my mind free, so as to give up any hypothesis, however much beloved (and I cannot resist forming one on every subject), as soon as facts are shown to be opposed to it." That statement might well serve as his epitaph.

See also Copernicus, Nicolas; Darwin, Erasmus; Darwinism; Evolutionary Ethics; Evolutionary Theory; Herschel, John; Humboldt, Wilhelm von; Huxley, Thomas Henry; Malthus, Thomas Robert; Newton, Isaac; Philosophy of Biology; Wallace, Alfred Russel.

Bibliography

WORKS BY DARWIN

The Autobiography of Charles Darwin (London, 1958) was edited by his granddaughter Nora Barlow, who restored the material omitted from the original. The original *Autobiography* was first published in 1887 as part of the *Life and Letters of Charles Darwin,* but many passages of the manuscript were omitted because they contained candid and caustic judgments of persons and of the Christian religion. These omitted passages, amounting to nearly six thousand words, were restored in the 1958 edition.

The *Life and Letters of Charles Darwin* (2 vols., New York: Basic Books, 1959) was edited by Francis Darwin. The 1959 edition has an introduction by George Gaylord Simpson.

Among the many editions of *On the Origin of Species* are a facsimile of the first edition, with an illuminating introduction by Ernst Mayr (Cambridge, MA: Harvard University Press, 1964), and a variorum text of the six editions, edited by Morse Peckham (Philadelphia: Univ. of Pennsylvania Press, 1959). The *Origin* is also available in paperback with a foreword by George Gaylord Simpson (New York, 1962).

For the voyage of the *Beagle* see *Charles Darwin's Diary of the Voyage of H.M.S. Beagle* (Cambridge, U.K.: Cambridge University Press, 1933), which was edited from the manuscript by Nora Barlow, and Nora Barlow's edition of *Charles Darwin and the Voyage of the Beagle* (London: Pilot Press, 1945), for which she has written an introduction.

Darwin and Wallace's *Evolution by Natural Selection* (London: Cambridge University Press, 1958) has been edited, with an introduction, by Gavin De Beer. This volume contains the Linnaean Society papers.

WORKS ON DARWIN

For works on Darwin see Alvar Ellegård's *Darwin and the General Reader* (Göteborg, Sweden, 1958) and Gavin De Beer's excellent *Charles Darwin: Evolution by Natural Selection* (London: T. Nelson, 1963).

T. A. Goudge (1967)

DARWIN, ERASMUS
(1731–1802)

Erasmus Darwin, an English physician, man of science, and poet, was the grandfather of Charles Darwin, whose evolutionary views he partly anticipated, and of Francis Galton. Like Charles he was educated at Cambridge, where he took the M.B. degree in 1755. For more than forty years he practiced medicine at Lichfield and Derby and gained a wide reputation for his skill, intellectual vigor, and originality of character. Among his friends were Jean-Jacques Rousseau, whom he met in 1766, and Joseph Priestley. He corresponded with both men. In 1784 he founded the Philosophical Society at Derby to stimulate interest in the sciences. He wrote copiously,

with varying degrees of success. His chief prose works are *Zoonomia or the Laws of Organic Life* (2 vols., London, 1794–1796) and *Phytologia or the Philosophy of Agriculture and Gardening* (London, 1799). Two long poems embodying his views about the origin and development of life, *The Botanic Garden* (London, 1789) and *The Temple of Nature* (London, 1803), were not taken seriously by his contemporaries, although Darwin himself was rather proud of them. Samuel Taylor Coleridge likened the poems to "mists that occasionally arise at the foot of Parnassus" and coined the word *darwinizing* to describe their biological speculations. After his death Erasmus Darwin was forgotten until interest in his ideas revived as a result of the fame of his grandson Charles.

An important feature of Erasmus Darwin's work is the relation it establishes between early evolutionary theory and the embryological controversy of the preformationists and the epigenesists. In "Of Generation," Chapter 39 of *Zoonomia,* Darwin argues against the doctrine that each new individual is already "preformed" on a minute scale in the reproductive cell from which it is developed. He defends an epigenetic position according to which new individuals develop by utilizing material from the environment to generate new parts. Hence, there is a transformation of a relatively undifferentiated egg into a complex organism. From this position it is only a short step to the view that life in general has evolved by a similar transformation.

Darwin actually took this step but did not provide a systematic justification of it. His writings are a curious mixture of observed facts, sober scientific judgments, and extravagant speculations, all designed to support the conclusion that living things, different from one another as they now are, originated from one "primal filament" that existed long ago. Through the ages organisms have altered to meet altered conditions of life. The result has been a continuous perfecting of their capacities. "This idea of the gradual formation and improvement of the animal world accords with the observations of some modern philosophers" (*Zoonomia,* Vol. I). An evolution of life has undoubtedly occurred.

Among the items of evidence adduced to support this contention are some that anticipate matters later embodied in *The Origin of Species.* Thus, Erasmus Darwin calls attention to such phenomena as the metamorphosis of tadpoles into frogs, the changes produced by the domestic breeding of animals, the specialized adaptations to climatic conditions, and, above all, "the essential unity of plan in all warm-blooded animals." These things oblige us to believe that all organisms have been derived from "a single living filament."

Embedded in Darwin's work are the rudiments of a theory about the causes of evolution. What he says foreshadows the more finished theory of the Chevalier de Lamarck. Environmental stimuli act on organisms that are endowed with the unique power of "irritability or sensibility." The organisms respond in accordance with their wants, desires, and dislikes. Thus, the bodily characteristics required to satisfy the organisms' demands are produced. These characteristics are inherited by some members of succeeding generations and favor them in the struggle for existence, which is depicted in lurid terms by Darwin in *The Temple of Nature.*

The facts that man's body bears traces of his evolution from lower forms of life and that Earth itself appears to have come into being gradually by the operation of natural processes in no way led Darwin to doubt the existence of "the Great Architect" of the cosmos. His solid and complacent deism enabled him to regard God as simply "the Great First Cause," who infused spirit and life into the primal filament and gave it the potentiality to evolve. "The whole of nature may be supposed to consist of two essences or substances, one of which may be termed spirit and the other matter" (*Zoonomia,* Vol. I, Section 1). The "whole of nature" was designed by the Great Architect. Indeed, God "has infinitely diversified the works of His hands, but has at the same time stamped a certain similitude on the features of nature, that demonstrates to us, that the whole is one family of one parent."

Darwin's views mark the close of the era of romantic speculation about natural history and the advance into an era of systematic observation and generalization. He did not, however, succeed in formulating any enduring principles. Perhaps his major achievement was acquiring the characteristics of scientific curiosity, independence of mind, and intellectual power that were transmitted to his descendants.

See also Coleridge, Samuel Taylor; Darwin, Charles Robert; Evolutionary Theory; Lamarck, Chevalier de; Priestley, Joseph; Rousseau, Jean-Jacques.

Bibliography

For material bearing on the once notorious controversy between Charles Darwin and Samuel Butler, in which the assessment of Erasmus Darwin's ideas played a part, see Charles Darwin, *Life of Erasmus Darwin: An Introduction to an Essay on His Works by Ernst Krause* (London, 1879), and Samuel Butler, *Evolution, Old and New* (London, 1879), Chs.

12–14. The complex story of the controversy is given in the complete edition of *The Autobiography of Charles Darwin*, edited by Nora Barlow (London, 1958), Appendix, Part 2, pp. 167–219.

See also Hesketh Pearson, *Doctor Darwin: A Biography* (London: Dent, 1930), and Desmond King-Hele, *Erasmus Darwin* (New York: Scribners, 1964).

T. A. Goudge (1967)

DARWINISM

The term *Darwinism* has both a narrow and a broad meaning. In the narrow sense, it refers to a theory of organic evolution presented by Charles Darwin (1809–1882) and by other scientists who developed various aspects of his views; in the broad sense, it refers to a complex of scientific, social, theological, and philosophical thought that was historically stimulated and supported by Darwin's theory of evolution. Biological Darwinism—the first sense—was the outstanding scientific achievement of the nineteenth century and is now the foundation of large regions of biological theory. Darwinism in the second sense was the major philosophical problem of the later nineteenth century. Today, Darwinism no longer provides the focus of philosophical investigation, largely because so much of it forms an unquestioned background to contemporary thought.

Darwin's theory is an example of scientific innovation that has had reverberations into the farthest reaches of human thought. It is fair to say that every philosophical problem appears in a new light after the Darwinian revolution. In order to outline the connections between biological and philosophical Darwinism, it will first be necessary to describe Darwin's own views and to discuss various criticisms that were directed against them. It will then be possible to describe Darwinism in the broader sense, and to distinguish the various ways in which the scientific theory has afforded material for philosophical inquiry.

DARWIN'S THEORY

The theory of the origin of species by means of natural selection was the discovery of Darwin and Alfred Russel Wallace (1823–1913). Both Darwin and Wallace had stated the theory in a series of papers delivered before the Linnaean Society on July 1, 1858. The members of the Linnaean Society listened without enthusiasm and apparently without much understanding, but in fairness to them, it should be observed that Wallace and Darwin did not present their theory forcefully on this occasion. Some of the shattering implications of the theory were not drawn in detail, and the evidence in its support, which Darwin in particular had amassed, was barely hinted at. Wallace's paper "On the Tendency of Varieties to Depart Indefinitely from the Original Type" was a discussion of a widely accepted argument in favor of the "original and permanent distinctness of species," namely, that the varieties that are produced by artificial selection in domesticated species never vary beyond the limits of the original wild species, and that whenever artificial selection is relaxed, the domesticated varieties revert to the ancestral form. These facts were interpreted by naturalists as evidence for an innate conservative tendency in nature that kept all variation within the bounds defined by the unbridgeable gaps between species.

But, Wallace argued, the view that artificial selection can produce only new varieties, never new species, rests on the false assumption that naturalists possess a criterion for distinguishing the species from the variety. Moreover he stated, "This argument rests entirely on the assumption that varieties occurring in a state of nature are in all respects analogous to … those of domestic animals, and are governed by the same laws as regards their permanence or further variation. But it is the object of the present paper to show that this assumption is altogether false." Overproduction, together with heritable variations, some of which are better adapted to the circumstances of life, will tend to make varieties depart indefinitely from the ancestral type, bringing about changes that will eventually amount to the origin of a new species. Wallace accounted for the reversion of domestic varieties by pointing out that the ancestral type is better adapted to life "in a state of nature," and consequently the very same principles that bring about progress in nature also bring about the reversion of domestic varieties.

Wallace aimed his argument precisely at the philosophical presupposition that for so long had stood in the way of a proper interpretation of natural selection, namely, that the species—being the exemplar of a divine archetype—is as well adapted as it could be and, consequently, that variation away from the type will automatically be selected against. Natural selection, according to this interpretation, is an agency of permanence, not change. One of Wallace's, as well as Darwin's, most original contributions consisted in breaking the hold of this idea.

Wallace's argument is implicit in Darwin's Linnaean Society papers, but the focus is different. Instead of challenging accepted opinion, Darwin added up well-known

facts. With great eloquence he described the prevalent overproduction of animals and plants: "Nature may be compared to a surface on which rest ten thousand sharp wedges touching each other and driven inwards by incessant blows." The wedges are held back by large numbers of "checks" that bring about the death, or prevent the mating, of individuals. "Lighten any check in the least degree, and the geometrical powers of increase in every organism will almost instantly increase the average number of the favored species." He called attention to the extreme heritable variability of animals under domestication. In nature there is also variation, although no doubt not as much. Some variants will be better adapted to their environments than others and will tend to survive and propagate. "Let this work of selection on the one hand, and death on the other, go on for a thousand generations, who will pretend to affirm that it would produce no effect?" To the effects of this natural selection, Darwin added the effect of "the struggle of males for females."

Both Wallace and Darwin had stated the essence of the theory of the *Origin of Species* (1859). The *Origin* itself is mainly a sober, scrupulously fair, and thoroughly documented elaboration and defense of the doctrine of natural selection presented in the Linnaean Society papers. Darwin set out to accomplish three things: (*a*) to show that evolution has in fact occurred; (*b*) to describe the mechanism of evolution; and (*c*) to account for the major facts of morphology, embryology, biogeography, paleontology, and taxonomy on the evolutionary hypothesis.

THE FACT OF EVOLUTION. Darwin freely admitted that we do not directly observe the process of evolution. The time needed even for the origin in nature of a new variety is far too long. Consequently, the case for the occurrence of evolution is simply the same as the case for its scope and mechanism, and Darwin did not have access to direct evidence for the efficacy of natural selection—a gap that was not filled until the twentieth century. Darwin argued that life is too short for direct evidence but that certain facts force the conclusion upon us that there must be evolution; and if we adopt the hypothesis, a wide range of hitherto unconnected facts may be given a uniform explanation.

THE MECHANISMS OF EVOLUTION. Darwin described three mechanisms that tend to effect the evolution of populations. These are natural selection, sexual selection, and the inheritance of characteristics acquired during the lifetime of the individual organism.

Natural selection. In the *Origin* Darwin placed the greatest weight on evolution by natural selection. It operates in conjunction with sexual selection and the inheritance of acquired characters and, Darwin argued, there are some features of organisms that could have developed only by natural selection. Indeed, it seems that the theory of natural selection was partially inspired by his observations on the *Beagle* voyage (1831–1836) of local variations, particularly in the islands of the Galápagos Archipelago, that could not be accounted for on Lamarckian grounds.

The theory of natural selection as Darwin presented it may be summarized as follows: (1) Populations of animals and plants exhibit variations. (2) Some variations provide the organism with an advantage over the rest of the population in the struggle for life. (3) Favorable variants will transmit their advantageous characters to their progeny. (4) Since populations tend to produce more progeny than the environment will support, the proportion of favorable variants that survive and produce progeny will be larger than the proportion of unfavorable variants. (5) Thus, a population may undergo continuous evolutionary change that can result in the origin of new varieties, species, genera, or indeed new populations at any taxonomic level. Darwinian natural selection may accordingly be defined as a differential death rate between two variant subclasses of a population, the lesser death rate characterizing the better-adapted subclass.

Darwin was careful to present evidence for every hypothesis in his account of natural selection. It was especially necessary to argue that natural populations do exhibit the requisite amount of variation and that the variation is heritable. He cited, among other things, the extreme variability of domestic plants and animals and the well-known fact that new varieties can be propagated. He admitted that the causes of variation were unknown; but he argued that changing environmental conditions greatly increase variability by action on the reproductive system, thereby providing material for natural selection when it is most needed. This is "indefinite variability." In addition, there is "definite variability," due to the direct action of the environment on the body of the organism. "Definite variations" are heritable; they provide material for natural selection and, being responsive to the environment, are more likely than chance variations to be adaptive.

"The laws governing inheritance," he remarked, "are for the most part unknown." This lack of knowledge turned out to be the most serious obstacle to the further development of the theoretical foundations of selection

theory in the nineteenth century; but, as Darwin noted, although the laws of inheritance were unknown, a number of the phenomena of inheritance were known, and those were probably all that were required for the theory of natural selection. Most important is the obvious fact that progeny bear an overwhelming resemblance to their parents, although they differ in some degree. In addition, Darwin was familiar with the intermittent appearance of hereditary characters, with sex-linked and sex-influenced characters, and with the tendency for a character to appear in the progeny at the same developmental stage that it appears in the parents.

For natural selection to be an agency of change rather than an agency of permanence, it is necessary that some variations from the ancestral type represent better adaptations. Darwin pointed out that, in fact, every organism could be better adapted to its ordinary environment; and that, moreover, environments change.

Pre-Darwinian taxonomy ascribed a very special significance to the species, as against varieties, genera, and the higher taxonomic groups. The species was regarded by the pious as the unalterable work of God; the limits laid down by the diagnostic features of any species established the limits of possible variation within the species. Thus, although any biologist would be willing to countenance the origin of new varieties or subspecies, brought about by the operation of biological laws, most were unwilling to admit the possibility of the origin of new species by natural processes. The title of Darwin's book was aimed precisely at this conception. Like Wallace, he argued that there is no difference in principle between the diagnostic characters of varieties and species; therefore, to admit the origin of new varieties amounts to admitting the possibility of new species—and if new species appear, so may new genera, families, and so on. He cited the existence of "doubtful species"—groups that cannot be definitely placed at either the variety or species level—and the general inconsistency of taxonomists in the identification of species.

Sexual selection. In the Linnaean Society papers Darwin described the second mechanism of evolution as the "struggle of males for females." The theory was developed further in the *Origin,* and it occupied some two-thirds of the pages of his *Descent of Man and Selection in Relation to Sex* (1871). In these later statements of the theory, the struggle of males for females is a special case of a more general phenomenon. Suppose that a population is divided in some proportion between males and females and suppose for the sake of simplification that all of the males and females are equally well endowed for the struggle for survival. Now, Darwin argued, it may happen that either the males or females are unequally endowed with some characteristic that will increase their propensity to leave progeny. There will then be selection in favor of that characteristic, even though it will not be favored by natural selection. All such cases Darwin calls "sexual selection." It is clear that different sorts of characteristics can influence the probability of having offspring. Some individuals, for example, may possess behavior patterns that lead to the fertilization of a larger percentage of eggs or have more efficient organs of copulation. Or they may have some advantage in the competition for mates—migratory male birds may arrive early at the breeding grounds and be ready to receive the more vigorous females, leaving the culls for their tardy brothers; or the females may for some reason prefer plumage or displays of a certain character; or some males may aggressively drive away other males; and so on. Finally, some characteristics that are also useful in the struggle for survival might also be useful in the competition for mates; for example, the antlers of male deer may do double duty against both rivals and predators.

Darwin appeals to sexual selection in order to account for the evolution of such things as mating rituals and secondary sexual characteristics, such as breeding plumage in birds. He regards it as especially significant in the evolution of man. The loss of body hair, for example, is attributed to systematic choice among man's ancestors of mates that exhibited large regions of bare skin.

The inheritance of acquired characters. Darwin's work was plagued by ignorance and misinformation concerning the laws of heredity. The principles of segregation and independent assortment, which form a cornerstone of contemporary evolution theories, were discovered by Gregor Mendel in 1864; but his paper remained unnoticed until 1900. Moreover, although "sports" were well known to biologists, the concept of mutation had not been clearly formulated. Consequently, the modern theory of the origin of genetic variation in populations was not available to Darwin; instead, he suggested that some variations are due to the action of the environment on the germplasm and that others are due to the effects of use and disuse. For example, if an animal's skin is tanned by sunlight, this may induce changes in its germplasm that will result in its progeny possessing pretanned skin; or if a wolf develops his muscles by chasing rabbits, his pups may inherit larger muscles. These mechanisms, if they exist, would account for some variability. But they would also account for some evolutionary change even in the absence of natural or sexual selection. Since, accordingly,

there seemed to be no sound reason for rejecting the inheritance of acquired characters and since the doctrine would aid in explaining both variability and evolutionary change, Darwin was led to adopt it and to give it increasing weight in his later years. This aspect of Darwin's views is often labeled Lamarckism, but the Chevalier de Lamarck himself, although he did accept the inheritance of the effects of use and disuse, did not accept the doctrine of the direct action of environmental factors on the germplasm.

THE SCOPE OF EVOLUTIONARY THEORY. It is clear that Darwin regarded his theory as revolutionary. He believed that all the traditional branches of biology would be transformed and deepened; familiar phenomena would take on a new significance; apparently unconnected facts could be regarded as mutually related. Even the vocabulary of the older biology would acquire new meanings: "The terms used by naturalists, of affinity, relationship, community of type, paternity, morphology, adoptive characters, rudimentary and aborted organs, etc., will cease to be metaphorical, and will have a plain signification." Natural history would acquire the fascination, not of a catalog of *curiosae,* but of a labyrinth that may be charted.

> When we no longer look at an organic being as a savage looks at a ship, as something wholly beyond his comprehension; when we regard every production of nature as one which has had a long history; when we contemplate every complex structure and instinct as the summing up of many contrivances, each useful to the possessor, … when we thus view each organic being, how far more interesting—I speak from experience—does the study of natural history become!

And not only would the old biology be put on a new foundation; whole new fields of research would become possible. For example, "Psychology will be securely based on the foundation … of the necessary acquirement of each mental power and capacity by gradation. Much light will be thrown on the origin of man and his history."

The major part of the *Origin* is devoted to the detailed application of the theory of natural selection to a range of biological phenomena. It is impossible to give more than a general impression of the thoroughness, detail, and diversity of Darwin's evidence. The modern reader cannot fail to be impressed not only by Darwin's immense learning but also by his subtlety of insight—his ability to locate those phenomena that lend his theory the most striking support.

The *Origin* as a whole provides, on the one hand, a sweeping portrait of the history and biology of living things, a portrait whose internal balance and consistency are easily discernible. On the other hand, Darwin fills selected regions of his portrait with careful detail, exhibiting the applicability of his theory to a variety of phenomena. These two aspects of his work constitute both the argument for the fact of evolution and the argument for the truth of his account of its mechanisms.

In the broad portrait Darwin shows how the main facts of known fossil successions, the relation of living fauna and flora to recent fossil forms, the geographical distribution of species, the connection between morphology and function, and the major features of embryological development are explicable by his theory. He applies it in detail to such phenomena—to mention only a few—as rudimentary organs, insect metamorphosis, the divergence of island and mainland forms, and sexual dimorphism. He provides us with a discussion of taxonomy that is philosophically superior to many contemporary accounts, arguing, among other things, in favor of the special significance for the taxonomist of embryological and phylogenetic studies.

Darwin was always sensitive to the effect that his views might have on the general public. In composing the *Origin* he decided to avoid the whole topic of man's evolution; the book would be a sufficiently bitter pill without explicitly treating a subject that was "so surrounded with prejudices." His only explicit reference to man was the remark quoted above, that "light will be thrown on the origin of man and his history." Darwin's successors, however, were not so cautious. Sir Charles Lyell (1797–1875) discussed the question in 1863. Shortly thereafter, Wallace published his paper "The Origin of Human Races and the Antiquity of Man Deduced from the Theory of Natural Selection." T. H. Huxley (1825–1895) and a number of Continental morphologists, particularly Ernst Haeckel (1834–1919), produced a series of studies aimed at showing the similarity of man and the anthropoid apes and giving speculative reconstructions of man's ancestry. Thus, by the date of Darwin's *Descent of Man* (1871), the controversy over man was in full swing, and there were already a number of alternative theories that Darwin had to consider, such as whether the races of men are distinct species.

Darwin showed a wise unwillingness to acknowledge any known nonhuman species, living or extinct, as ancestral to man. We have so far examined, he argued, only animals that have diverged from the prehuman stock. For instance, the anthropoid apes and man have a common

ancestor, but its remains have not been found. Nor did he identify species that are ancestral to the primates, the mammals, or even the vertebrates. He did trace a general line of descent: Old World ape, a lemurlike animal, some "forms standing very low in the mammalian series," marsupials, and monotremes. No true reptile is an ancestor of man. All the classes of vertebrates may have been derived from a remote ancestor similar to the larvae of the tunicates. With a flash of romanticism, Darwin wrote: "In the lunar or weekly recurrent periods of some of our functions we apparently still retain traces of our primordial birthplace, a shore washed by the tides."

In the *Descent of Man* evolution by the inheritance of acquired characters and by sexual selection plays a larger role than in the *Origin.* Darwin admitted that he had been accused of overrating the importance of natural selection, but added, "whether with justice the future will decide." His relative retreat from natural selection was probably occasioned by two factors: first, his doubts as to whether Earth is old enough for evolution by natural selection without substantial help from faster mechanisms; second, his belief that man is in many ways less the child of violent nature than his ancestors, a belief that requires considerable appeal to sexual selection and to the development of moral and spiritual qualities through social usage.

CRITICISMS OF DARWIN'S THEORY

In spite of the resistance that Darwin's theory aroused on other than scientific grounds, the weight of his arguments was largely—but with many notable exceptions—sufficient for the younger generation of biologists. In 1872, in the sixth edition of the *Origin,* Darwin was in a position to write, "At the present day almost all naturalists admit evolution under some form." It was, like any novel and important theory, carefully scrutinized for empirical weaknesses. We shall describe the major ones and indicate how they were dealt with.

The most damaging scientific objections were the following:

(1) Darwin had no direct evidence for the effectiveness of natural selection, let alone for the origin of new species.

(2) Darwin could not show a single species that was transitional between two known species.

(3) Complex organs, such as the vertebrate eye, could not have evolved by stages, since they would have been useless at any preliminary stage and hence would have given their possessor no selective advantage.

(4) If evolution has taken place, then some evolutionary trends must have continued past the point of usefulness to the organism. Such trends could not be accounted for by Darwinian selection.

(5) Earth is not old enough for evolution to have taken place.

(6) Evolution by natural selection is incompatible with the laws of inheritance.

(7) There is no inheritance of acquired characters.

The first two objections were commonly raised in the nineteenth century; they are genuine questions that require some sort of answer. Darwin, however, was not in a position to answer them in a way that would satisfy everybody, since the weight that one assigns to them depends in part upon personal preference.

INDIRECT EVIDENCE. With regard to the first objection Darwin pointed out that natural selection cannot be directly observed; we can only present indirect evidence in its favor. On this point he was mistaken. Natural selection has been directly studied in the twentieth century, both experimentally (in fruit fly populations, for example) and in nature (for instance, the development of so-called industrial melanism). But even today Darwin's and Wallace's contention that evolution by natural selection can pass the species limit has no direct support. Darwin recognized, however, that it is no fatal objection to a theory if some of its components are not subject to direct verification.

TRANSITIONAL SPECIES. On the second criticism—the absence of forms intermediate between species—Darwin had a double-barreled answer. He admitted that, for instance, we know of no forms intermediate between man and the apes. But we have innumerable examples of species that are in process of giving rise to new species, namely, those that have varieties or subspecies. These polytypic species (as they are now called) are intermediate between other species which, to be sure, have not yet evolved, but which are in process of evolving.

When it was further objected that we ought to have better examples of demonstrable ancestors of existing species, Darwin appealed to the incompleteness of the fossil record. This is the correct answer, but one that is hardly satisfying to a skeptic. Again, the weight that one would assign to the objection depends upon personal preference.

DEVELOPMENT OF COMPLEX ORGANS. Darwin was well aware of the difficulty in accounting for the origin of structures that would be useless, even deleterious, until they were essentially complete. The eye, he wrote, gave him "a cold shudder." In such cases as the eye, however, he had no alternative but to appeal to natural selection. Therefore, he was compelled to argue that in point of fact all the earlier stages in the evolution of the eye were useful in the struggle for survival. Darwin himself provided us with the standard textbook example: he constructed a plausible sequence of stages that could have led to the human eye. Each stage is a functional eye; and something similar to each stage does exist in one or another living species. The criticism has the form, "Such and such *could not* have happened." It can be countered piecemeal, by showing in a variety of cases how it *could* have happened.

ORTHOGENETIC TRENDS. A great many of Darwin's critics accepted the fact of evolution but entered reservations concerning his account of the mechanisms of the process. The reservations were of several types. Some rejected "Lamarckism," by which they meant simply the inheritance of acquired characters; they were known as the Neo-Darwinians. Others doubted that there was such a process as sexual selection. Still others, however, believed that there must be an evolutionary process that Darwin had not identified at all. The evidence consisted in the existence of apparently nonfunctional evolutionary trends. Trends that continue over long periods and that are relatively straight-lined—for example, increasing size in horses and increasing length of sabers in the saber-toothed cat—came to be called orthogenetic trends. The question was whether orthogenetic trends could be accounted for on Darwinian principles.

Wallace argued (in "Geological Climates and the Origin of Species," *Quarterly Review,* 1869) that the development of man's brain could not be so accounted for. Man's apelike ancestors, he argued, had reached a certain stage of evolution and then, over a period of some ten million years, remained largely unchanged except for a steady orthogenetic increase in the size and complexity of the brain. This was an unprecedented episode in the history of life, for it freed man from those ordinary pressures of natural selection that so often led to close specialization and ultimate extinction. Moreover, the brain acquired abilities that could not have been exercised in a primitive environment, such as the power to construct speculative systems of ideas or the insight into spiritual reality. These are present in modern man, but would have been useless in man's primitive ancestors. Natural selection operates only on abilities that are actually so exercised as to give an advantage in the struggle for life. "An instrument," Wallace concluded about the brain, "has been developed in advance of the needs of its possessor." Later he wrote: "A superior intelligence has guided the development of man in a definite direction, and for a special purpose, just as man guides the development of many animal and vegetable forms." Thus we avoid the "hopeless and soul-deadening belief" that man is the product of "blind eternal forces of the universe."

Darwin looked upon this as a failure of nerve, a hankering after miraculous origins for man. "I can see no necessity for calling in an additional and proximate cause in regard to man," he wrote in a letter to Wallace. Nevertheless, Wallace's position, fitting as it did the efforts of many theologians to come to grips with Darwinism, gained a number of adherents, and although the main line of evolutionary theory has bypassed it, even now versions of Wallace's position turn up from time to time.

Wallace had argued that the evolution of the brain was an orthogenetic trend that outstripped its usefulness. Others argued that trends sometimes continued even after they had become positively deleterious. A favorite example was the teeth of the saber-toothed cat, which, it was alleged, were valuable as weapons up to a certain length, but which finally became detrimental by interfering with feeding. There would be selection against increased tooth length under these conditions; consequently, it was argued, some cause other than natural selection must have operated. A variety of theories were proposed—for example, those of Karl Nägeli (1817–1891) and E. D. Cope (1840–1897). These theories posited an otherwise unknown internal principle of change, which was compared to the laws of embryological development, to the principle of inertia, or, as with Henri Bergson, to creative spiritual activity. Since the theories accounted for nothing other than the alleged orthogenetic trends, they have always had a peripheral position in the history of evolutionary thought. Moreover, subsequent analysis of orthogenesis has shown that in most cases the trends are in fact adaptive; and in those cases where they are not adaptive, contemporary theory provides various possible sorts of explanation compatible with the doctrine of natural selection, such as the explanation that if a trend affects only adults past the breeding age, it will not be selected against.

AGE OF EARTH. In 1865 William Thomson, Lord Kelvin, published a paper titled "The Doctrine of Uniformity in Geology Briefly Refuted." Its argument was aimed at Lyell and his followers, who had maintained that Earth as we

now find it is not the result of a series of catastrophes, but is the outcome of the ages-long operation of geological processes that we can still observe. This viewpoint, known as uniformitarianism, was widely accepted among geologists even before the publication of the *Origin,* having been impressively established in Lyell's *Principles of Geology* (1834). It was in fact an earlier application of the idea of evolution. But uniformitarianism required vast reaches of time; consequently, Kelvin was prodding its weakest point when he argued that Earth could not be as old as the geologists supposed. Grant, Kelvin argued, that Earth was once a molten sphere; then it could not have solidified much over twenty million years ago, or it would now be cooler, through dissipation of its heat, than we actually find it. The biological consequences were clear: there was not enough time for evolution to have produced the forms we now see.

Darwin was deeply concerned by this reasoning. As far as he could tell, it was perfectly sound; on the other hand, he was perfectly convinced that Earth had supported life for a much longer time. His later emphasis on Lamarckism was probably an attempt to provide an evolutionary process that was swifter than natural selection. But this was a half measure; in fact, Darwin simply swallowed what he believed to be a contradiction—a not uncommon occurrence in the history of science. It turned out that Kelvin's argument was mistaken, since he was unaware of an additional source of heat within Earth, namely radioactive decay.

LAWS OF INHERITANCE. As noted above, the evolutionists of the nineteenth century worked in ignorance of the principles of genetics discovered by Mendel; this lack was by far the most serious theoretical gap in the Darwinians' arguments. It now appears that no fundamental innovation in evolutionary theory was possible until the gap was filled. Biologists of the nineteenth century accepted a rough theory of blending inheritance, that is, the view that the characteristics of the progeny of sexual crosses were intermediate between the characteristics of the parents. This theory was seldom explicitly defended, since everyone was familiar with a variety of phenomena that were incompatible with it, for example, blue-eyed children of brown-eyed parents. Nevertheless, when biologists theorized at all on the subject, the theory produced was ordinarily a vague and suitably guarded version of the theory of blending inheritance.

In 1867 Fleeming Jenkin ("The Origin of Species," *North British Review*) pointed out that the blending theory was incompatible with the theory of natural selection as ordinarily presented by the Darwinians. He argued that if favorable variations appeared in a population, their characteristics, even if favored by natural selection, would soon be lost in the vast population pool by crossing with individuals of the normal type. Assume, for instance (as Jenkin did), that a white man is greatly superior to a black man and that a white man is shipwrecked on a black-populated island. "He would kill a great many blacks in the struggle for existence; he would have a great many wives and children.… But can anyone believe that the whole island will gradually acquire a white, or even a yellow population?" Jenkin's argument in essence is this: the white man's children will be darker than their father; and it is impossible on the blending theory that their descendants could become lighter, whatever the effects of natural selection might be.

Again, Darwin was forced to admit the strength of a powerful objection that he was unable to counter directly. At best, he could only argue that natural selection would be effective if adaptive variations were sufficiently common; the black island could become white, for example, if there were a steady influx of shipwrecked sailors. He actually had no evidence that adaptive variations were sufficiently common; instead, he retreated more and more to the Lamarckian theory that variation is due to the effects of activity in the environment and would accordingly be largely adaptive.

Unlike the answer to Kelvin's objection, which could not have been offered in the nineteenth century, the answer to Jenkin was available but remained unknown except to a few, who did not see its significance. Mendel's paper on plant hybridization established an alternative to the blending theory of inheritance. Mendel showed that there were discrete genetic factors that pass unchanged from generation to generation and are hence not subject to Jenkin's swamping effect. Mendel had established that the character of these factors (genes) is not changed by other factors in the germplasm and that the factors segregate independently of one another in gamete formation. (He was unaware of the phenomenon of linkage.) Researchers of the literature on heredity recovered Mendel's work in 1900; and in 1904 William Bateson (1861–1926), in *Genetics and Evolution,* applied Mendel's laws to the theory of natural selection, thus answering Jenkin's objection.

The new genetics turned out to be far more significant for the theory of evolution than merely answering Jenkin's objection. The history of scientific Darwinism in the twentieth century was mainly the story of a series of advances in genetics, and the working out of their conse-

quences for evolution. Mendel's laws were correlated with the behavior of the chromosomes in meiosis; the concepts of chromosome and gene mutation were introduced; linkage was discovered and understood; and statistical methods were employed in the analysis of the dynamics of genetic change in natural populations. One major gain of these developments was a systematic understanding of the origin and maintenance of genetic variability—the question that was so troublesome for Darwin. Another was the final decline of the Lamarckian aspect of Darwinism.

ACQUIRED CHARACTERS. The Neo-Darwinians had already denied the inheritance of acquired characters, but their evidence against it, like the Neo-Lamarckians' evidence in its favor, was largely anecdotal. August Weismann (1834–1914) had presented the theory that life is essentially a continuous stream of germplasm that from time to time gives rise to whole organisms; the organisms die but the germplasm is immortal. The stream can divide (gamete formation) and merge (fertilization), thus accounting for variability. This view was employed by Weismann and others as a theoretical argument against the inheritance of acquired characters, for it is an easy step from the continuity of the germplasm to its independence of somatic influences. The emergence of Mendelism shed a new light on Weismann's theory. The mechanism of "immortality"—self-replication of chromosomes—was elucidated, and evidence accumulated that the chromosomes were indeed uninfluenced, or influenced only randomly, by somatic factors.

PHILOSOPHICAL DARWINISM

We have considered Darwinism as a biological theory; we may now consider its wider intellectual connections. These are many and complex, so it will be necessary to select the most important—those which now seem to be enduring ingredients of speculative thought or those which struck the people of the later nineteenth century with the greatest force. The differences between the climate of opinion—the ordinary presuppositions, ideas about the proper pattern of argument, assumptions as to proper method, in short, the worldview—of the mid-nineteenth and twenty-first centuries is large, comparable in degree to the differences between the Middle Ages and the Renaissance. Of course the change had many causes, but the advent and absorption of Darwinism, while in part an effect of other currents, was also one major cause.

We shall consider the connections of Darwin's theory in three major regions: scientific cosmology, theology, and social doctrine.

SCIENTIFIC COSMOLOGY. Scientists have general views about the way things are. The scientists of any historical period are likely to share a common set of views, with, of course, individuals differing over one or another point to some degree. These general views, insofar as they concern a subject matter of professional scientific interest and insofar as they are capable of influencing method, methodology, or empirical formulations, may be called cosmological. They differ from the ordinary statements of a science (for example, "organisms overproduce," "acquired characters are not inherited") in degree of determinateness. They are so formulated that they are exempt from immediate verification and falsification but subject to specification, by means of a series of semantical decisions, into determinate, verifiable propositions. A good example of such a cosmological proposition is "Nature makes no jumps," or "Nature has no gaps." Darwin, unlike many of his contemporaries, was fond of making this remark (in Latin); he employs it in the Linnaean Society papers and subsequently quotes it again and again. It constitutes part of Darwin's cosmology and is a point on which the nineteenth century was deeply divided. It is clear that the sense of the proposition is not sufficiently determinate, as it stands, for verification. But it can be construed to mean, for instance, that evolution is gradual or that the apparent gaps between living species can be filled if we consider a sufficient stretch of history.

These properties of cosmological belief have important implications. First, it is possible to arrive at a cosmology by a process akin to generalization—an empirical statement can be construed as the determinate form of an indeterminate proposition, which in turn can be applied to new subject matters. This is the formal pattern of the influence of science on cosmology. Second, the precise verbal formulation of a cosmological belief is relatively unimportant; indeed, it can affect thought without being explicitly formulated at all. For cosmological beliefs do not function as premises of empirical arguments; rather, they impart color to empirical argument, affecting its form and conceptual materials.

Darwin's biological theory was itself supported by prior developments in cosmological belief. The theory of evolution by natural selection did not occur to Darwin in an intellectual vacuum. Most important of these cosmological beliefs was uniformitarianism, the belief that nature operates everywhere and always by the same sorts

of law. This view Darwin had imbibed from Lyell's *Principles of Geology*; it became cosmological by construing the geological theory as exhibiting a general truth about the way things, including livings things, are. This particular belief is already a powerful stimulus to look at organic nature as the outcome of a historical process, although, to be sure, the belief does not entail this conclusion.

A second belief, which Darwin inherited and was seen to support, was the necessity of taking time seriously. This meant, among other things, that the past is long. By the date of the *Origin* there was little actual evidence on the age of Earth, let alone the age of the universe. Outside scientific circles, the prevailing view was that Earth and universe were the same age, something on the order of thousands of years. As long as this is accepted, evolution is evidently most improbable. Some geologists, in particular James Hutton (1726–1797), had, on the other hand, argued that Earth is infinitely old—an important argument, since it helped to accustom scientists to the possibility of vast stretches of time and change. Geologists after Hutton were willing to help themselves to as much time as they needed, and Darwin gladly followed suit.

Taking time seriously, however, gained a deeper meaning after the publication of the *Origin,* namely, that change is a fundamental feature of nature. This constituted part of the cosmology of every Darwinian. It meant that the process of change is not merely the reshuffling of preexisting materials in accordance with physical law but that the materials themselves are subject to alteration. For instance, as applied to biology it meant that the fundamental form, the species, did not merely exhibit eternal law but changed in such a way that new regularities of behavior replaced the old. In the favored terminology of the nineteenth century, we may say that taking time seriously meant that the laws of nature are subject to change.

Structures and patterns of behavior, then, have to be regarded as historically conditioned. This is the cosmological aspect of the most characteristic post-Darwin view of method, the insistence upon the investigation of origins, together with the view that such investigation can be scientific. Thus, we find the development of the idea of a human prehistory, the application of elaborate schemes concerning, as they were called, stages of development—spiritual, social, political, moral—and the belief that, at least in outline, the future of man may be successfully charted.

Pre-Darwinian biological theory was strongly influenced by the view that all living things are patterned after an eternal idea or archetype. This was held not only for the species but also for other taxonomic categories and for anatomical structures as well. Taxonomists were fond of describing, for example, the ideal vertebrate or mollusk; and morphologists described the ideal organ. One of the achievements of Darwinism was to break the hold of this notion on taxonomic and anatomical theory. Darwin was finally able to write, in *Descent of Man,* "A discussion of the beau ideal of the liver, lungs, kidneys, etc., as of the human face divine, sounds strange to our ears."

THEOLOGY. The expressed doctrines of theology are related to empirical propositions as cosmological doctrines are related to the natural sciences. The role of Darwin's theory as a generator of such indeterminate beliefs naturally is well exemplified in theology. On the one hand it was immediately taken to be in prima facie opposition to a number of theological doctrines, especially the following: the uniqueness of man as God's supreme creation; the importance of natural theology; and the dominant theory, in Protestant circles, that the Bible is an authoritative source of beliefs about the natural world.

The first theological reaction to Darwinism can only be described as one of outrage; but by the close of the century, theologians having decided that since they must live with Darwinism, they ought to love it, the outlines of a reconciliation had been sketched. Even further, Darwinism was allowed to guide the formation of a new brand of theology. We shall consider first the reaction.

As we have seen, Darwin's readers were quick to grasp the consequences of the *Origin* for man himself. These consequences immediately aroused the most intense feelings. These feelings were quite justified, for Christian theology demands that man be considered unique; and his uniqueness was universally interpreted as ontological separateness from the rest of creation. The geologist Adam Sedgwick (1785–1873), for example, spoke no more than common opinion when he wrote in 1850 that man is a barrier to "any supposition of zoological continuity—and utterly unaccounted for by what we have any right to call the laws of nature." The Darwinians not only argued that man is continuous with the animal kingdom and subject to the laws of nature; they also asserted that his mental, moral, and spiritual qualities evolved by precisely the same processes that gave the eagle its claws and the tapeworm its hooks. Such opinions were a threat to the deepest level of Christian doctrine, and were bound to be, until man's uniqueness could be given a new theological interpretation.

Moreover, the furor over the animal nature of man was heightened, especially in Britain, by local circum-

stances. T. H. Huxley compared man and the ape with endless zest, knowing how the comparison annoyed his opponents. For apes and monkeys were thought to be oversexed and obscene; in addition, the British took very seriously the principle that a man's standing in the world is dependent on the standing of his ancestors. Thus the literature of the period is enlivened by comic remarks, such as, "Are you descended from an ape, Mr. Huxley, from your mother's or your father's side?" (Bishop Wilberforce) and "You can't wash the slugs out of a lettuce without disrespect to your ancestors" (John Ruskin). But the symbol of the ape squatting in one's family tree was no more than an expression of dismay at being swallowed up in the infinite forms of nature. The twentieth century did not fully regain its equanimity on this point. Pius XII wrote that a Catholic may accept a doctrine of evolution, but should beware of doubting that there was a first man and woman. And consider this passage from the speech of William Jennings Bryan at the Scopes trial (1925): "We are told just how many species there are, 518,900. … and then we have mammals, 3,500, and there is a little circle and man is in the circle, find him, find man."

The edifice of traditional theology was touched at other points. Early-nineteenth-century theologians placed heavy weight on the cooperation of science and religion. The clergyman-naturalist was a familiar figure. It was thought that the intricacy and systematic interconnections of nature exhibited the handiwork of God; to study them was an act of piety. More specifically, natural teleology was the mainstay of natural theology. William Paley's *Natural Theology* (1802) is a good example. He holds that God's creation is totally good, that the organs of living things are almost perfect, that all animals have their just share of happiness, and that all this demonstrates with thousandfold certainty the existence and beneficence of God. An older natural theology tended to see evidences of God's design throughout nature; but Paley, and others after him, such as Thomas Chalmers in the *Bridgewater Treatises* (1834), rest their case on the structure of living things: consider, they suggest, the hand, the heart, the eye (especially the eye); they are complex and adapted for their functions to a degree that transcends all possibility of chance correlation.

By hindsight this attitude appears curiously self-defeating as well as vulnerable. The religiously inspired examination of organic adaptation was precisely one factor that led to Darwin's account of the origin of adaptation. His theory made the last citadel of divine teleology in nature untenable except, of course, for a few holdouts; but it was also widely interpreted as refuting all natural teleology, especially by the German materialists. "Chance" had been defined by Paley as "the operation of causes without design," and on this definition Darwinism leaves the origin of species to chance.

Theology in the middle half of the nineteenth century was especially vulnerable to Darwinism on a second point, namely, its extreme Biblicism and, even further, its literalism in biblical interpretation. It hardly needs saying that Darwinism is incompatible with any literal construction put upon either the Old Testament or the New Testament. The laity and most of the clergy, however, insisted upon such constructions. Matthew Arnold quotes the following as prevailing opinion in England: "Every verse of the Bible, every word of it, every syllable of it, every letter of it, is the direct utterance of the Most High"—a view Coleridge describes as "Divine ventriloquism." The matter was not so extreme outside of Britain, but the fact remains that Protestant education and practice relied heavily on the study and interpretation of the Bible.

The intellectual compromise that gradually emerged seems obvious today; the problem was not to think of it but to accept it. It consists in admitting that man is part of nature and that he is indeed, even in his spiritual aspects, the outcome of an evolutionary process. But lowly origins do not detract from a unique present. And the process of evolution is either guided, as Wallace suggested, or is itself the mode and manner of God's creation. Indeed, it was sometimes argued that Darwinism provides us with an elevated conception of God. Canon Charles Kingsley, for example, wrote to Darwin as follows: "I have gradually learnt to see that it is just as noble a conception of the Deity to believe that he created primal forms capable of self-development …, as to believe that He required a fresh act of intervention to supply the *lacunas* which He Himself had made." This passage is quoted by Darwin with some changes in later editions of the *Origin.* As Kingsley also put it, Darwin allows us to get "rid of an interfering God—a master-magician, as I call it" in favor of a "living, immanent, ever-working God."

The final step in this direction was to give God an even more intimate metaphysical connection with natural process. This step had been taken by previous philosophers—Benedict (Baruch) de Spinoza and G. W. F. Hegel, for example; but it was repeated under the aegis of Darwinism by Bergson, Alfred North Whitehead, and a number of Protestant thinkers. The problem of a divine nature that is both perfect and yet incomplete is one contemporary heritage of Darwinism.

SOCIAL DOCTRINE. The social thought of the later nineteenth century drew so heavily from the theories of evolution that its major ideas became known as social Darwinism. The 1850s were a period of revolutionary fervor in the streets as well as the academies, and political ideologists seized on Darwin as their major intellectual spokesman. His views, or rather selected aspects of them, presented ideal material for application to ethical, economic, and political problems.

It is convenient to divide social Darwinism into a political right and left, using these terms in their rough, contemporary editorial-page sense. In adopting Darwinism to social questions, it must be admitted that the right wing had the best of the bargain. In Europe these were the men whose interests were vested in hereditary privilege and in the factories and institutions of the industrial revolution. On the grounds of these interests they defended themselves against any attempt to justify social revolution, governmental control, unionism, or socialism in any of its many nineteenth-century forms. The ideology that was developed, with the help of Darwinism, in order to facilitate this defense also committed them, in various combinations, against such things as child-labor legislation, poor laws, compulsory safety regulations, and public education. A similar ideology provided the United States with its justification for the undisturbed economic expansion, speculation, and competition that we associate with the robber barons.

On the other hand, Darwinism was employed by the social reformers. Karl Marx wanted to dedicate the first volume of *Das Kapital* to Darwin. George Bernard Shaw, although he criticized the theory of natural selection, defended his socialism with the help of his version of Bergson's creative evolutionism. The reformers saw Darwinism as the final demonstration that no particular economic or political institution—however hallowed by tradition or supported by existing theories—need be regarded as unalterable. The forms of society, like the forms of life, are local, temporary, and functional and may accordingly be changed (for the better) without shaking the foundations of the cosmos.

In short, the biology and cosmology of Darwinism was capable of being all things to all men. It enjoyed this status by virtue of its ability to inspire and lend a measure of apparent scientific support to the following major ideas:

(1) The vision of a science that was historical, and at the same time a rigorous application of natural law, inspired a new vision of a science of society. Herbert Spencer (1820–1903), whose evolutionism antedated the *Origin,* became the symbol of this ideal wedding of history and sociology. He drew elaborate comparisons between social structures and the forms of living organisms and saw societies as undergoing a progressive evolution in which egoism would be gradually replaced by altruism through a mechanism analogous to the inheritance of acquired characters. Sociology stood in relation to society as evolutionary biology stood to the phenomena of organic nature.

(2) The process of natural selection, interpreted as the survival of the fittest, provided a means for explaining social process. The American political economist William Graham Sumner (1840–1910), for example, saw society as the outcome of a social struggle in which each man, in pursuing his own good, can succeed only at the expense of others. The fittest in this social struggle are the ruthless, the imaginative, the industrious, the frugal. They climb to the top, and it is right that they should do so. The idle, infirm, and extravagant are losers, not adapted to the realities of their world, and thus legitimately subject to elimination by "social selection." Sumner presents society with an alternative: either "liberty, inequality, survival of the fittest," or "not-liberty, equality, survival of the unfittest." Self-made millionaires are the paradigm of the fittest. They are "a product of natural selection, acting on the whole body of men to pick out those who can meet the requirement of certain work to be done."

This doctrine of the financially successful as the cream of the universe naturally had a sympathetic audience. John D. Rockefeller, Andrew Carnegie, and Theodore Roosevelt were supporters, although Roosevelt believed that the unfit were entitled to some protection.

(3) Darwinism provided a rationale for Adam Smith's doctrine of the "Invisible Hand." Smith had supposed that while each man follows his innate tendency to "truck, barter, and trade," men's efforts would automatically dovetail in such a way that the economic good of society as a whole would be served. And Darwin had shown that the net result of each organism's engaging in a struggle for its own welfare was continuous evolution of the species as a whole in the direction of better adaptation to its environment. The political implications of this viewpoint are clear.

The central ethical question raised by the social Darwinists is this: granted that man is subject to natural law, and even granted further that he is subject to some form of natural or social selection, can one legitimately derive from this such policies as laissez-faire? Alfred Russel Wallace had argued that with the advent, under divine guidance, of man's brain, the evolution of man was no longer

controlled by natural selection, so that inference from the doctrine of natural selection to ethical policy would be illegitimate. Huxley provided a similar argument: Man represents an island of cultural evolution in a sea of Darwinian change. These issues have largely passed into history, however, due to the philosophical point that whether or not to support a law of nature is not a question for decision.

The fate of Darwinism since the twentieth century has been mixed. Social Darwinism is of no more than historical interest. It is rightly regarded as philosophically naive and, moreover, as concerned with social questions that are not of contemporary interest. The same is largely true of the theological battles over the significance of evolution. Current theology exhibits a sublime indifference to the questions that agitated Huxley and Bishop Wilberforce. It must be pointed out, however, that modern theology is free to pursue other problems because of the clarification of the status of man and of the relation of science to theology that emerged from the Darwinian debate.

In biological theory proper, Darwin's theory remains secure. His Lamarckism is no longer accepted, if we discount some periodic revivals in the former Soviet Union; and the doctrine of sexual selection is still a matter of some debate. But the major theory of the *Origin,* evolution by natural selection, is the framework of modern evolutionary theory. This modern account—sometimes called the synthetic theory and sometimes, rather confusingly, Neo-Darwinism—accepts in toto the doctrine of natural selection as described above but develops it in a manner that Darwin himself could not have envisaged. The synthetic theory may fairly be described as Darwin's theory of natural selection, deepened by the absorption of twentieth-century genetics and systematically applied to the whole range of biological phenomena.

The absorption of genetics accounts for the novel developments in the doctrine of natural selection itself. Darwin thought of natural selection fundamentally as differential survival, and he regarded the organism as the natural unit that is subjected to selective pressures. With the advent of Mendelian genetics, and especially of the statistical study of the genetics of populations, these two Darwinian conceptions underwent a significant change. From the geneticist's point of view, differential survival is subordinate to differential reproduction of genetic materials; evolution is simply temporal change in the genetic constitution of a population. The simplest model of evolutionary change would be the following: Suppose that we have in a population two alleles, a_1 and a_2, of a gene a, and that a_1 is present in the proportion p, and a_2 in the proportion $1-p$. Then any temporal change in the value of p would be a case of reproductive differential between a_1 and a_2; and it would be an evolutionary change in the population. Some biologists simply identify such differential reproduction with natural selection, in which case sexual selection is a special case of natural selection. The natural unit of selection becomes the gene rather than the whole organism.

This conception of natural selection is not incompatible with Darwin's. Differential survival is still the major cause of differential reproduction of genes; and there is still a clear and obvious sense in which the organism is the fundamental unit of natural selection. But the new conception of natural selection facilitates the discussion of a large range of questions, for example, the roles of isolation and migration in evolution; the effectiveness of very small selective advantages; the roles of gene mutations, sex-linkage, and dominance; and so on. The modern theory has much to say on these topics that could not have been foreseen by Darwin, but nothing that he could not readily endorse.

See also Arnold, Matthew; Bergson, Henri; Darwin, Charles Robert; Darwin, Erasmus; Ethics, History of; Evolutionary Ethics; Evolutionary Theory; Good, The; Haeckel, Ernst Heinrich; Hegel, Georg Wilhelm Friedrich; Huxley, Thomas Henry; Lamarck, Chevalier de; Laws of Nature; Marx, Karl; Paley, William; Racism; Ruskin, John; Smith, Adam; Spinoza, Benedict (Baruch) de; Sumner, William Graham; Teleology; Wallace, Alfred Russel; Whitehead, Alfred North.

Bibliography

DARWIN AND WALLACE

Darwin, Charles Robert. *On the Origin of Species by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* London, 1859. A variorum text, edited by Morse Peckham (Philadelphia, 1959), was published in paperback with an introduction by G. G. Simpson (New York, 1962). There is also a Modern Library edition (New York, 1949).

Darwin, Charles Robert. *The Descent of Man, and Selection in Relation to Sex.* London: J. Murray, 1871.

Darwin, Charles Robert. *The Voyage of the Beagle.* London and New York, n.d. A reissue by J. M. Dent and E. P. Dutton of their 1906 edition of the *Journal of Researches into the Geology and Natural History of the Various Countries Visited by H. M. S. Beagle, 1832–1836.*

Darwin, Charles Robert, and Alfred Russel Wallace. *Evolution by Natural Selection.* Cambridge, U.K.: Cambridge University Press, 1958. This contains Darwin's sketch of

1842, his essay of 1844, and the Darwin and Wallace papers read before the Linnaean Society in 1848.

Wallace, Alfred Russel. *Contributions to the Theory of Natural Selection.* London: Macmillan, 1870.

Wallace, Alfred Russel. *Darwinism.* London: Macmillan, 1889.

Wallace, Alfred Russel. "The Origin of Human Races and the Antiquity of Man Deduced from the Theory of Natural Selection." *Journal of the Anthropological Society of London* (1864).

CULTURAL ASPECTS OF DARWINISM

Dewey, John. *The Influence of Darwinism on Philosophy.* New York, 1910.

Eiseley, Loren. *Darwin's Century; Evolution and the Men Who Discovered It.* Garden City, NY: Doubleday, 1958. Contains many illuminating discussions of the interplay of philosophical and scientific theories.

Fothergill, Philip. *Historical Aspects of Organic Evolution.* London: Hollis and Carter, 1952. A history of evolutionary theories.

Goudge, T. A. *Ascent of Life; a Philosophical Study of Evolution.* Toronto: University of Toronto Press, 1961.

Gray, Asa. *Natural Science and Religion.* New York: Scribners, 1880. A topical discussion of the way theism looked to an evolutionist.

Himmelfarb, Gertrude. *Darwin and the Darwinian Revolution.* Garden City, NY: Doubleday, 1959.

Hofstadter, Richard. *Social Darwinism in American Thought, 1860–1915.* Philadelphia, 1944.

Huxley, Thomas Henry. *Evolution and Ethics, and Other Essays.* New York, 1898.

Schneider, Herbert. "The Influence of Darwin and Spencer on American Philosophical Theology." *Journal of the History of Ideas* 6 (1945).

MODERN EVOLUTIONARY THEORY

Darlington, C. D. *The Evolution of Genetic Systems.* Cambridge, U.K.: Cambridge University Press, 1939.

De Beer, Gavin. *Embryology and Evolution.* Oxford: Clarendon Press, 1930.

Dobzhansky, Theodosius. *Genetics and the Origin of Species,* 3rd ed. New York: Columbia University Press, 1951.

Fisher, R. A. *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press, 1930. This is the classic application of statistical methods to the dynamics of evolving populations; together with the Darlington and Dobzhansky works, it affords a good introduction to the crucial relations between evolution and population genetics.

Mayr, Ernst. *Systematics and the Origin of Species.* New York: Columbia University Press, 1942.

Ross, H. H. *A Synthesis of Evolutionary Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

Simpson, G. G. *The Meaning of Evolution.* New York: New American Library, 1951. The best general introduction to the modern synthetic theory, this is a revised and abridged paperback edition of *The Meaning of Evolution: A Study of the History of Life and Its Significance for Man.* New Haven, CT, 1949.

Simpson, G. G. *The Major Features of Evolution.* New York: Columbia University Press, 1953. This and the above work by Simpson are the best nontechnical introductions for the general reader.

Morton O. Beckner (1967)

DAVID BEN MERWAN AL-MUKKAMMAS

See *Muqammis, David ben Merwan al-*

DAVID OF DINANT

The materialistic pantheist of the Middle Ages David of Dinant taught at Paris near the beginning of the thirteenth century. Apart from this fact, almost nothing is known of his life. It is uncertain whether he derived his name from Dinant in Belgium or Dinan in Brittany. His major work, *De Tomis, Hoc Est de Divisionibus,* is probably identical with the *Quaternuli* condemned at a provincial council in Paris in 1210, and his writings were among those banned at the University of Paris in 1215 by the papal legate, Robert de Courçon. Our knowledge of his ideas is largely derived from Albert the Great, Thomas Aquinas, and Nicholas of Cusa.

David developed his philosophy at a time when Latin Christian thought was facing an almost unprecedented challenge from rival world views. Neoplatonism, introduced into the medieval West by John Scotus Erigena and popularized in the twelfth century by numerous translations of Arabic works, was the first great non-Christian system to impress the medieval mind, but by the early thirteenth century Aristotelianism loomed large, and other Greek philosophies were not unknown. Attempts were made to blend the Christian doctrine of creation with these doctrines, notably with the Neoplatonic theory of emanation, with the result that the distinctive character of the biblical conception of the relation between the world and God was at least occasionally obscured.

The title of David's *De Tomis* suggests some indebtedness to Erigena's *De Divisione Naturae,* and David's pantheism may well have been inspired to some extent by his reading of Erigena's work. His thought seems, however, to have been more strongly influenced by ancient Greek materialism, as described in Aristotle's *Physics* and *Metaphysics,* and by certain Aristotelian ideas dialectically manipulated in the manner of the early medieval logicians.

David's interpretation of reality was essentially monistic. He first divided the objects of knowledge into three classes and then presented individual objects within each class as mere modes of a primary reality. Thus, bodies are modes of matter (*hyle*), souls are modes of mind (*nous*), and eternal substances or separated forms are modes of God. Furthermore, these three primary realities are themselves essentially one being or substance.

David supported this doctrine by a dialectical argument based on the logical notion of a "difference" (*differentia*) that, when added to a genus, forms a species. Such *differentiae,* he argued, can be predicated only of composite beings. God, mind, and prime matter, however, are all simple realities, and can therefore include no *differentiae.* Consequently, they must be substantially identical.

David's monism may be further characterized as materialistic. In his view, neither God nor matter possesses form, since beings determined by form are individual, composite substances. God and matter, therefore, cannot be known by an assimilation of their forms through abstraction. If in fact the intellect knows both God and matter, the explanation must be that it is already identical with them. Furthermore, if both God and matter are unformed, they are nothing but being in potentiality. Being in potentiality, however, is the definition of prime matter. Properly speaking, then, the ultimate reality, which is at once God, mind, and matter, is best described as matter.

See also Albert the Great; Aristotelianism; Aristotle; Erigena, John Scotus; Medieval Philosophy; Neoplatonism; Nicholas of Cusa; Pantheism; Thomas Aquinas, St.

Bibliography

Gabriel Théry, *David de Dinant. Étude sur son panthéisme matérialiste* (Paris, 1925), is the only book on David of Dinant. See also A. Birkenmajer, "Découverte de fragments manuscrits de David de Dinant," in *Revue néo-scolastique de philosophie* 35 (1933): 220–229, which does not, however, affect the accepted picture.

Eugene R. Fairweather (1967)

DAVIDSON, DONALD

(1917–2003)

Donald Davidson was born in 1917 in Springfield, Massachusetts, and graduated from Harvard in 1939. After serving in the United States Navy, Davidson returned to Harvard, where he wrote his doctoral dissertation on Plato's *Philebus* (1990a). After he received his PhD in 1949, Davidson went on to do extensive work in decision theory, in collaboration with Patrick Suppes and others. After many years at Stanford, and somewhat shorter stays at Princeton, Rockefeller, and Chicago, Davidson in 1981moved to the University of California at Berkeley, where he was appointed Willis and Marion Slusser Professor of Philosophy. Davidson lived in Berkeley for the rest of his life, continuing to produce important work until his death in 2003.

The early confrontation with the methodological challenges of giving empirical application to rational-choice theory had a lasting influence on Davidson. It is apparent in his later formulation of philosophical questions regarding action, the mental and linguistic meaning. Davidson's views on these matters have gradually come to articulation through a series of papers presenting detailed arguments pertaining to specific problems. In each of three areas of philosophy, Davidson elaborated a set of closely interconnected and highly influential doctrines. This entry looks briefly at each area in turn, emphasizing certain general features characteristic of Davidson's philosophical approach. The entry concludes with a glance at key themes of Davidson's later work, in which he elaborates an anti-representationalist conception of mind and of philosophy.

THE CAUSAL THEORY OF ACTION

Davidson's view, first set out in print in *Actions, Reasons, and Causes* (1963), is that individuals must consider the reasons for their actions—combinations of propositional attitudes, paradigmatically belief-desire pairs—to be also their causes. In this and related papers Davidson granted a main premise of the anti-causalist view prevailing at the time, that the teleological form of action-explanation makes such explanation irreducibly different from the nomological form characteristic of explanation in the natural sciences. What is distinctive about action-explanation is that it identifies the events involved (the action and its explanatory antecedent) in terms that reveal them to be part of a rational pattern. Davidson proceeded to challenge anti-causal orthodoxy, however, by arguing that it does not follow from this irreducible difference that action explanation is not a species of causal explanation.

A striking aspect of Davidson's view is the claim that the appeal to reason on which action explanations turn will be genuinely explanatory only insofar as the particular events thus rationally related are also related as cause and effect, and hence, for Davidson, may also be charac-

terized by nomologically related descriptions. But Davidson insisted that the explanatory efficacy of a reason-explanation in no way depends on one possessing the law-evincing descriptions of the particular events in question. Indeed, in the typical case one enjoys the full benefits of effective action explanation without the slightest idea of what those descriptions might be. Davidson's work thus reconciles the following three fundamental claims.

First, when an event is cited in successful explanation of another event, the former is a cause of the latter. Second, causal relations between events entail nomological relations between them. Third, *explanans* and *explanandum* in action explanations are captured in terms that cannot be subsumed under strict law. This reconciliation trades on a particular conception of the relation between cause and law. If two events are causally related, they are so related no matter how described. The nomological relation, however, obtains between kinds of events; laws, as Davidson said, are linguistic, and so while causal relations are extensional, and causally related events necessarily fall under what he terms strict laws (that is, laws that are "free from caveats and *ceteris paribus* clauses … treating the universe as a closed system" (1993 [2005a], p. 191), they instantiate such law only under some appropriate description. Hence, the descriptions under which two causally related events appear in successful action explanation may be such that no amount of knowledge of strict causal law would allow one to infer the action from knowledge of the conditions cited in the explanation of it (1963a; 1993; 1995).

ANOMALOUS MONISM

A crucial element in Davidson's account of action is the distinction between a particular event and the descriptions that sort particular events under kinds. This same distinction is central also to his claims about the nature of the mental and its relation to the physical. In "Mental Events" (1970) and subsequent papers (1991b, 1993), Davidson argued that what it is to be a mental event is to be an event that falls under a mental predicate; that is to say, for Davidson, an event is mental just in case it falls under a description that ineliminably involves an intentional term. Correspondingly, what it is to be a physical event is to fall under a physical predicate. Physical predicates are of diverse kinds; a subset of physical terms are the predicates of developed physics. They form an ideal vocabulary the constitutive purpose of which is to track the causal structure of the world by displaying all events as they fall under maximally strict laws.

Since Davidson conceived of events as extensionally identified spatio-temporal particulars constituting nodes in the causal network, W. v. O. Quine's basic ontological dictum expresses also for Davidson an important truth; it is the unique business of physics to aim for full coverage. What this means for Davidson is that all events, qua nodes in the causal network, must be describable in the terms of physics. Yet some events are also mental, and Davidson argued that his physicalism supports no reductivist or eliminativist conclusions. For while all particular mental events are also particular physical events, no particular *kind* of mental event is a particular *kind* of physical event. The reason for this is that intentionality, which for Davidson is the mark of the mental, is constituted, in his view, by one's efforts to characterize fellow creatures as rational according to an inter-subjective standard. One is able to view fellow language-users thus because an individual has at his or her disposal two kinds of conceptual resources. One keeps track of other people by keeping tabs on objective environmental relations in which human beings are all embroiled in various and changing ways. At the same time, one is able to deploy a set of concepts—of belief, desire, and so on—which allows one to construct accounts not just of objective environmental relations, but also of how these relations appear to someone to be.

This system of double bookkeeping allows an individual to absorb a great deal of variation and irregularity in human behavior by accounting for objective anomalies in terms of subjective variables. But this strategy remains informative and useful only insofar as the essential discrepancies between subjective perspective and objective reality that interpretation exploits are prevented from becoming arbitrary or chaotic—were that to happen, the subjective would lose its explanatory purpose, it would simply mark the place where explanation ends.

This is why the interpretive construction of the subjective perspective must be tightly constrained; as Davidson stated, making sense of others "we will try for a theory that finds [them] consistent, believer[s] of truths, and lover[s] of the good (all by our own lights, it goes without saying)" (1970 [1980a], p. 222). This constraint on the application of intentional terms is often referred to as the "principle of charity." It reflects the fact that only the attitudes (though not only the rational attitudes) of a recognizably rational subject may be invoked in a genuinely explanatory way to account for the subject's behavior. Moreover, as Davidson later emphasized, because rationality considerations govern an individual's application of propositional-attitude concepts, these concepts

are irreducibly causal, "identified in part by the sorts of action they are prone to cause, given the right conditions" (1991b [2001], p. 216).

As he further pointed out, "the right conditions" are themselves not independently characterizable. The phrase, marking the interdependence of the application conditions of mental predicates, remains an ineliminable qualification of the sort of platitudinous generalizations that express the content of our psychological terms. By contrast, Davidson argued, the application of the predicates of physics—aimed at the formulation of strict law—cannot itself depend on causal concepts (1991b). The application conditions of terms related by strict empirical law must be independently specifiable. The real difference, then, between the mental and the physical, and the reason for the irreducibility of one to the other, stems from the fact that the vocabulary of physics and the vocabulary of psychology have evolved under the pressure of distinctively different interests. What one wants from the former are modes of description that allow people to interact with each other as persons. What one wants from the latter are laws "as complete and precise as we can make them; a different aim" (1991b [2001], p. 217).

TRUTH AND MEANING

In the philosophy of language, Davidson is associated with the view that an individual may account for linguistic competence by appropriately characterizing the evidence available to and resources required by an idealized interpreter (1973). There are two fundamental aspects to this position. What, Davidson asked, might one know such that by knowing it one would be able to say what a speaker of a given language meant by some arbitrary utterance? His answer is a theory of truth for that language, an account of the logical structure of a language of the sort that Alfred Tarski demonstrated how to construct (1967a; 1990b; 2005b). The condition of adequacy for such a theory is an adaptation of what Tarski called "convention T." One has, Davidson proposed, a theory of meaning for a given language L provided one has a theory that entails for each sentence of L an instance of the schema, "s is true in L if and only if p." In this schema, s would be replaced by an expression that mentions a sentence of L (for example by means of quotation marks), and p replaced with any sentence of the language in which the theory is stated that is true if and only if the sentences mentioned by s is true. Such a theory provides, based on finite resources, a recursive characterization of the truth conditions for any sentence of L. While all that is demanded by convention T is that the theorems of the theory—known as T-sentences—capture co-extension of truth-values, "the hope," as Davidson said, "is that by putting appropriate formal and empirical restrictions on the theory as a whole, individual T-sentences will in fact serve to yield interpretations." (1973, [1984], p 124).

This proposal has spawned a great deal of work in formal semantics, guided by the aim of accounting for natural-language idioms in terms of their deep structure, or logical form, which makes explicit their truth-theoretical composition. For Davidson, the notion of logical form is extremely powerful; constrained on the one hand by one's intuitions concerning entailment relations, and on the other by the logical resources of Tarskian truth-theory construction, the uncovering of logical form functions as a crucible within which crystallize the ontological categories human language commits a person to. So for example, support for an ontology of events takes the form of an argument that one cannot account for the entailment relations intuitively characteristic of action sentences within the logical confines of a Tarskian theory of truth for a language unless one is willing to see such sentences as quantifying over events (1967b).

If a theory of truth is to serve as a theory of meaning for a language, tone needs to know how an interpreter may arrive at such a theory for a language she does not know. What is required for a recursive truth-theory to have empirical application? This question points to the other main aspect of Davidson's conception of linguistic understanding, an aspect where Quine's influence is most apparent. Observing the utterances of a speaker but knowing neither what the speaker means nor what the speaker believes, the interpreter will face endless alternative explanations of any observed piece of behavior. However, she can narrow the range of possibilities dramatically, by assuming that the speaker's behavior, including the speaker's linguistic behavior, embodies a rational response to salient features of her environment. This assumption of rationality is defeasible with respect to any particular attribution within the context of the construction of a theory of the meanings of someone's words and the contents of the person's thoughts. Davidson's point is emphatically not that one is never irrational, or that the irrational cannot be interpreted. Rather, the lesson is that irrationality is conceptually parasitic, diagnosable only against a background of reason (1982b; 1985). Thus, in what Davidson called radical interpretation, the interpreter may inductively construct a theory of truth for the speaker based on observations of behavior only by assuming that the speaker's mental

life—her thoughts, actions and utterances—constitutes a largely rational whole (1973; 1980b). This assumption, compulsory for the radical interpreter, is often referred to as the principle of charity.

Even while minimizing the irrationality of her subject in accordance with charity, the radical interpreter will be able to produce for a speaker alternative theories of belief and meaning, theories that comport equally well with the empirical evidence (i.e., with the speaker's utterances and their contexts). This indeterminacy Davidson regards as innocuous; the salient facts about meaning and mind are what such differing theories have in common. If alternative theories are empirically equivalent, this means that there is more than one way of stating the facts that interpretive theories are designed to capture. This is no threat to the viability of interpretation (1974a; 1979; 1990b; 1991b).

CHALLENGES

With respect to Davidson's view of action, the most serious objection holds that Davidson's theories cannot indicate how action explanation actually can be explanatory at all. The point of the objection is that one cannot reconcile the three fundamental claims regarding explanation, cause, and law to which Davidson's work is committed (see aforementioned text). One claim—advanced, for example, by Jerry Fodor—is that informative action explanation must somehow draw on the explanatory power of nomic relations, in which case Davidson's irreducibility-claim would be threatened. An alternative view—defended by anti-causalists like George Wilson—is that the explanatory force of reason-explanation is *sui generis*, and does not depend on reasons being causes. This would jeopardize Davidson's conception of event monism.

With regard specifically to anomalous monism, Jaegwon Kim and others have argued that Davidson's view renders the mental causally inert. So reason explanations cannot really be explanations at all, since Davidson believed that any genuine explanation of an event, including an action, must pick out actual causal relations. Partly because of their different views on the individuation of events, this conflict is difficult to assess. However, if one grants Davidson his fundamental claims—that is, that the difference between the mental and the physical is a matter of vocabulary of description, and that events should be extensionally conceived—then his concept of supervenience ensures that a change in the truth value of the relevant kind of mental predicate ascription will entail some difference or other in causal relations. Naturally, alternatives to Davidson's Humean conception of causality and of the relations between causality and law are frequently at play in criticisms of Davidson's account. Davidson relied on this conception both in arguing for anomalous monism and in reconciling the irreducibility of action-explanation with a causal view of action.

As for Davidson's philosophy of language, there have been objections at various levels to the idea that a theory of meaning for a language must take the form of a Tarskian truth-theory. Even while accepting the proposal that a theory of meaning should take the form of a theory of truth, one may ask, for example, why theorists should restrict themselves, in producing a formal semantics for a language, to the resources of first-order predicate calculus. A great majority of scholars now doubt the prospects of an account of natural language semantics couched in purely extensional terms.

ANTI-REPRESENTATIONALISM

Davidson's contention that theories of truth as Tarski defined them give the structure of theories of meaning is best viewed as a pragmatic methodological commitment. What supports Davidson's most innovative philosophical conclusions is the more general point that one must understand meaning in terms of truth, in conjunction with his insistence—following Quine—on a third-person perspective to meaning and mind. This view makes the conditions of interpretation constitutive of content. Together, these commitments yield an account of the concept of truth constrained by the methodological requirements of interpretation. This account contrasts both with traditional correspondence theories and with epistemic accounts of the sort advanced, for example, by Hilary Putnam. It is also distinct from disquotationalist or deflationist theories such as that of Paul Horwich (1990b; 2005b).

The significance of these core commitments is readily apparent in Davidson's argument aimed to discredit the duality of representational scheme and empirical content on the grounds that it presupposes the notion of an untranslatable language (1974b). If truth and meaning are interlocking concepts whose features are illuminated by an account of the methodology of an ideal interpreter, the idea of alternative representations of reality that are mutually semantically impenetrable is not coherent. This argument also marks a dividing line between Davidson and Quine. For the metaphysical opposition between what is given to the mind on the one hand, and the processes brought to bear on that given, on the other hand, is the very duality in terms of which empiricism

faces its defining challenge, namely to articulate a coherent notion of sensory evidence (1982a). On this fundamental score, Quine has remained within the bounds of empiricism (1990c). Davidson, on the other hand, has gone on explicitly to reject the basic metaphor of mind as inner space on which empiricism rests. For Davidson the hold of this metaphor reveals itself in the persistence of the interdependent notions of mental states as representational and of truth as correspondence, which, in turn, inextricably entangle philosophy in the problems of relativism and skepticism (1986a; 1987; 1988; 1990b; 2005b).

Davidson's alternative to the representational view of mind is most succinctly expressed in the thesis that there is thought only when there is actual communication (1989a; 1989b; 1991a; 1991b; 1992). On this controversial view, knowledge of one's own mental state, knowledge of the so-called external world, and knowledge of the mental states of others appear mutually interdependent (1991b). This blocks the very possibility of a skeptical or relativist challenge from arising, insofar as these are typically constructed around arguments that purport to show the impossibility of deriving any one of the three kinds of knowledge from either or both of the other two. This impossibility is something Davidson's work accepts—indeed insists on. Against the skeptic or relativist his claim is simply that the three forms of knowledge stand or fall together; denying one is to deny all, and to deny all is just to deprive our intentional concepts of any application.

This position rests on two key claims. One is that shared linguistic understanding is a prerequisite for any standard of objectivity (1991b). Such a standard gives content to the very distinction exploited by the propositional-attitude verbs between what is and what seems from some perspective to be, and hence, on Davidson's conception, is a prerequisite of thought. The other is the claim that the idea of shared linguistic understanding presupposes actual communication (1986b). The mental is thus what one reveals when one subjects a certain vaguely delimited range of causal relations to a particular kind of description, the terms of which presuppose the mutual recognition of subjects interacting in a shared world.

This view carries with it the commitments to event monism, to the constitutive role of rationality for content, and to a view of human agents as an integral part of the natural world, that have always been evident in Davidson's work. The distinction between extensionally conceived particulars and their descriptions remains pivotal. But the upshot is fundamentally at odds with the governing metaphors of modern epistemology-centered philosophy: "A community of minds," Davidson concluded, "is the basis of all knowledge; it provides the measure of all things." And he added: "It makes no sense to question the adequacy of this measure, or to seek a more ultimate standard" (1991b [2001], p. 218). However one assesses the plausibility of the considerations Davidson offers in support of this position, cognizance of the thorough-going externalism on which it is based should lead one to see it not as a species of antirealism or idealism, but as a profound rejection of foundationalist aspirations. Systematically linking the content of one's concepts to one's communicative practices as agents in the world, Davidson's work articulates a recognizably pragmatist view of mind, nature, and philosophy.

See also Action; Anomalous Monism; Philosophy of Language; Philosophy of Mind; Semantics; Supervenience.

Bibliography

WORKS BY DAVIDSON

Essays

The following essays have been reprinted in the works following this list, as indicated by year of publication:
"Actions, Reasons and Causes" (1963). Reprinted in 1980a.
"Truth and Meaning" (1967a). Reprinted in 1984.
"The Logical Form of Action Sentences" (1967b). Reprinted in 1980a.
"Mental Events" (1971). Reprinted in 1980a.
"Radical Interpretation" (1973). Reprinted in 1984.
"Belief and the Basis of Meaning" (1974a). Reprinted in 1984.
"On the Very Idea of a Conceptual Scheme" (1974b). Reprinted in 1984.
"The Inscrutability of Reference" (1979). Reprinted in 1984.
"A Unified Theory of Thought, Meaning and Action" (1980b). Reprinted in 2004.
"Empirical Content" (1982a). Reprinted in 2001.
"Paradoxes of Irrationality" (1982b). Reprinted in 2004.
"Incoherence and Irrationality" (1985). Reprinted in 2004.
"A Coherence Theory of Truth and Knowledge" (1986a). Reprinted with "Afterthoughts, 1987" in 2001.
"A Nice Derangement of Epitaphs" (1986b). Reprinted in 2005a.
"Knowing One's Own Mind" (1987). Reprinted in 2001.
"The Myth of the Subjective" (1988). Reprinted in 2001.
"What is Present to the Mind?" (1989a). Reprinted in 2001.
"The Conditions of Thought" (1989b). Reprinted in 2001.
"The Structure and Content of Truth" (Journal of Philosophy, Vol. 97. 1990b.
"Meaning, Truth and Evidence" (1990c). Reprinted in 2005a.
"Epistemology Externalized" (1991a). Reprinted in 2001.
"Three Varieties of Knowledge" (1991b). Reprinted in 2001.
"The Second Person" (1992). Reprinted in 2001.
"Thinking Causes" (1993). Reprinted in 2005a.
"Laws and Cause" (1995). Reprinted in 2005a.

Works

Essays on Action and Events. Oxford: Oxford University Press, 1980a.

Inquiries into Truth and Interpretation. Oxford: Oxford University Press, 1984.

Plato's Philebus. New York: Garland, 1990a.

Subjective, Intersubjective, Objective. Oxford: Oxford University Press, 2001.

Problems of Rationality. Oxford: Oxford University Press, 2004.

Truth, Language, and History. Oxford: Oxford University Press, 2005a.

Truth and Predication. Cambridge, MA: Harvard University Press, 2005b.

WORKS ABOUT DAVIDSON

Brandl, J., and W. Gombocz, eds. *The Mind of Donald Davidson*. Amsterdam: Rodopi, 1989.

Dazenbrock, W. R., ed. *Literary Theory after Davidson*. University Park, PA: Penn State Press, 1989.

Evnine, Simon. *Donald Davidson*. Palo Alto, CA: Stanford University Press: 1991.

Hahn, L. ed. *The Library of Living Philosophers*. Vol. 27: *The Philosophy of Donald Davidson*. Lasalle, IL: Open Court, 1999.

Joseph, M. *Donald Davidson*. Kingston, ON: McGill-Queen's University Press, 2004.

Kotatko, P., P. Pagin, and G. Segal, eds. *Interpreting Davidson*. Palo Alto, CA: CSLI Publications, 2001.

LePore, E., ed. *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, 1986.

LePore, E., and K. Ludwig. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press, 2005.

LePore, E., and B. McLaughlin, eds. *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, 1985.

Ludwig, K., ed. *Donald Davidson*. Cambridge, U.K.: Cambridge University Press.

Malpas, J. E. *Donald Davidson and the Mirror of Meaning*. Cambridge, U.K.: Cambridge University Press, 1992.

Preyer, G., F. Siebelt, and A. Ulfig, eds. *Language, Mind and Epistemology: On Donald Davidson's Philosophy*. Dordrecht: Kluwer Academic Publishers, 1994.

Ramberg, B. *Donald Davidson's Philosophy of Language: An Introduction*. Oxford: Blackwell, 1989.

Stoecker, R., ed. *Reflecting Davidson*. Berlin: De Gruyter, 1993.

Vermazen, B., and M. Hintikka, eds. *Essays on Davidson: Actions and Events*. Oxford: Oxford University Press, 1985.

Bjørn T. Ramberg (1996, 2005)

DA VINCI, LEONARDO

See *Leonardo da Vinci*

DEATH

Although most of the great philosophers have touched on the problem of death, few have dealt with it systematically or in detail. Frequently, as in the case of Benedict (Baruch) de Spinoza, an author's views on the subject are known to us from a single sentence; and at almost all stages in Western history we are likely to discover more about the topic in the writings of men of letters than in those of technical philosophers. Whether this relative reticence on the part of philosophers should be attributed to a general lack of interest or to other causes is a moot point. Arthur Schopenhauer, who was the first of the major philosophers to deal extensively with the subject, declared that death is the muse of philosophy, notwithstanding that the muse is seldom avowed. And the existentialist philosophers from Søren Kierkegaard to the present have more or less consistently endorsed Schopenhauer's contention; Albert Camus's declaration in *The Myth of Sisyphus* (1942) that suicide is the only genuine philosophical issue is an extreme but notable case in point. On the other hand, most contemporary Anglo-American analytic philosophers probably regard the paucity of materials on death as evidence of the subject's resistance to serious philosophical inquiry. In general, they wish to exclude the subject of death from the area of legitimate philosophical speculation, either as a part of their campaign against metaphysics or on the grounds that the subject can be more adequately dealt with by psychologists and social scientists. The psychologists and social scientists have, in fact, recently given signs of a willingness to explore the question. One such indication was a symposium on the psychology of death at the 1956 American Psychological Association Convention, which resulted in the publication in 1959 of an anthology including contributions from scholars in a wide variety of fields. Unfortunately, as several of the contributors to this volume lamented, the number of experimental studies actually undertaken has been disappointingly small.

THE KNOWLEDGE OF DEATH

The primary concern of most philosophers who have dealt with the question of death has been to discover ways in which we may mitigate or overcome the fear it tends to inspire. There are, however, several other loosely related problems that have also tended to excite interest or controversy and that it will be advisable to discuss first. How does man learn of death? Is death a natural phenomenon, or does it require explanation in nonnatural terms? What specific psychological or social conditions tend to heighten the awareness and fear of death?

AWARENESS OF DEATH. The clearest and simplest answer to the first of these questions was given by Voltaire, who stated: "The human species is the only one which knows it will die, and it knows this through experience" (*Dictionnaire philosophique*). Although some persons have questioned whether man is the only animal who knows he will die, arguing that certain of the lower animals appear to show some vague presentiment of approaching extinction, it appears to be unquestioned that man alone has a clear awareness of death and that man alone regards death as a universal and inevitable phenomenon. The interesting question is how man knows he will die. The view that experience alone gives knowledge of death derives support from the ignorance of death displayed by many children and from anthropological data indicating that many primitive peoples refuse even as adults to regard death as necessary or universal. However, a number of twentieth-century philosophers contested this view, especially Max Scheler and Martin Heidegger, who argue that the awareness of death is an immanent, a priori structure of human consciousness. Although neither of these authors offers anything in the nature of scientific evidence for his position, it is not easily refuted; for, if one grants current notions about levels of consciousness, apparent ignorance of death may be interpreted as merely superficial and attributed to some form of repression. Moreover, the imperfect knowledge of death among primitive peoples is a fact that could be used against those who argue that the knowledge of death comes from experience, since the hazards of their lives expose primitive peoples to an earlier and greater experience of death than is common among civilized men. At the very least it must be granted that the knowledge of death depends not only upon experience but also upon a level of mental culture that makes it possible to interpret experience accurately.

Ironically, Sigmund Freud, who more than anybody else has habituated us to think in terms of levels of consciousness and has thereby rendered credible the idea that knowledge of death may exist despite apparent ignorance, stated that the consciousness, not the apparent ignorance, of death is merely superficial, the unconscious being firmly convinced of its immortality. How Freud could reconcile this belief, which dates from the period of World War I, with his later belief in the unconscious death wish is not clear.

DEATH: A NATURAL PHENOMENON? Is death a natural phenomenon? Most persons today tend to find this question a bit foolish. It is noteworthy, however, that most primitive peoples attribute death to the agency of gods or demons who are jealous of human achievements. Equally significant is the Christian explanation of death as punishment for the sins of Adam. It should also be observed that if by a "natural" phenomenon one means a fact that can be fully understood and explained by empirical inquiry, death is not a natural phenomenon for Heidegger or Scheler. This reluctance to explain death in terms of natural causes has an interesting parallel in the reluctance to explain life itself naturalistically, and the religious or metaphysical perspectives that give rise to nonnaturalistic interpretations of life tend also to occasion nonnaturalistic interpretations of death.

VARIATIONS IN CONSCIOUSNESS OF DEATH. Are there great variations in the awareness or fear of death from person to person, from epoch to epoch, from culture to culture? If so, how are these variations to be explained? Surprisingly, very little attention has been given to these questions. The most interesting and almost the only hypothesis on this topic is that of Johan Huizinga and Paul-Louis Landsberg, who, each in his own way, link the consciousness of death to individualism. According to these authors, the consciousness of death has been most acute in periods of social disorganization, when individual choice tends to replace automatic conformity to social values; they point especially to classical society after the disintegration of the city-states; to the early Renaissance, after the breakdown of feudalism; and to the twentieth century. This hypothesis has yet to be fully confirmed or disconfirmed by careful historical and anthropological study. However, it is true that late antiquity, the early Renaissance, and the twentieth century made unusually great contributions to the literature on death.

THE FEAR OF DEATH

With respect to the fear of death, the great divide is between those who argue that only the hope of personal immortality will ever reconcile men to death and those who argue that the fear of death may be mitigated or overcome even when death is accepted as the ultimate extinction of the individual person. The second group, which is remarkably heterogeneous, may be subdivided according to the techniques recommended for allaying fears.

THE EPICUREANS. One of the oldest of the "solutions" to the fear of death was that of Epicurus and his followers. According to Epicurus, the fear of death is based upon the beliefs that death is painful and that the soul may survive to experience pain or torture in an afterlife.

Since both of these beliefs are mistaken, it suffices to expose them as such. Although death may be precipitated by painful disease, death itself is a perfectly painless loss of consciousness, no more to be feared than falling asleep. And since the soul is merely a special organization of material atoms, it cannot survive physical destruction. "Death," Epicurus said, "is nothing to us. … It does not concern either the living or the dead, since for the former it is not, and the latter are no more" (*Letter to Menoeceus*). It is hardly necessary to point out that many persons have questioned Epicurus's conception of the soul and consequently have rejected his views with respect to its immortality. The principal criticism, however, is that the Epicureans have falsely diagnosed the cause of humankind's fear of death. Death terrorizes us, not because we fear it as painful, but because we are unwilling to lose consciousness permanently. The twentieth-century Spanish existentialist philosopher Miguel de Unamuno reports that "as a youth and even as a child, I remained unmoved when shown the most moving pictures of hell, for even then nothing appeared to me quite so horrible as nothingness itself."

THE STOICS. The later Stoics, especially Seneca, Epictetus, and Marcus Aurelius, offered a more complicated and elusive view of death. Seneca said that to overcome the fear of death we must think of it constantly. The important thing, however, is to think of it in the proper manner, reminding ourselves that we are but parts of nature and must reconcile ourselves to our allotted roles. He recurrently compared life to a banquet from which it is our obligation to retire graciously at the appointed time, or to a role in a play whose limits ought to satisfy us, since they satisfy the author. The fear of death displays a baseness wholly incompatible with the dignity and calm of the true philosopher, who has learned to emancipate himself from finite concerns. Essential to the Stoic outlook was the Platonic view that philosophizing means learning to die; that is, learning to commune with the eternal through the act of philosophic contemplation.

Although much of Stoic thinking on death crept into later Christianity, the contemporary Christians saw in this thinking a sinful element of pride. Death, Augustine said, is a punishment for human sin, and the fear of death cannot be overcome except through divine grace. Others find it highly questionable whether one can reasonably accept the metaphysical underpinnings of the Stoic view, most especially the belief in a providential order of Nature.

SPINOZA. A third solution is that of Spinoza. He wrote: "A free man thinks of nothing less than of death, and his wisdom is not a meditation upon death but upon life" (*Ethics,* Prop. LXVII). Since Spinoza did not elaborate, it is possible to argue almost endlessly about the precise import of this famous remark. Most often, however, it is interpreted to mean that men can and should allay the fear of death simply by diverting their attention from it, and some persons have argued that by his nature man tends to—perhaps must—follow this advice. François de La Rochefoucauld, for instance, averred that man can no more look directly at death than he can look directly at the sun. One fundamental criticism of this position comes from the Stoics and the existentialists, both of whom maintain that the fear of death can be allayed only by facing it directly. A second criticism consists in pointing out that the fear of death is frequently an involuntary sentiment that cannot be conquered by a merely conscious decision or a bare act of will. It is not enough to tell people not to think of death; one must explain how they can avoid thinking of it.

DEATH AND THE GOOD LIFE. This brings us to a fourth view on death, a view that was felicitously put by Leonardo da Vinci. Just as a day well spent brings happy sleep, so, he said, a life well spent brings happy death. Painful preoccupation with death has its source in human misery; the cure is to foster human well-being. A happy man is not seriously pained by the thought of death, nor does he dwell on the subject. This view was held by many Enlightenment thinkers, most notably the Marquis de Condorcet. It also appears to be the view of most pragmatists and of Bertrand Russell.

There are two counterarguments. The first is the theme prevalent in several branches of Christianity concerning the total impossibility of attaining happiness on Earth. The second is the even more familiar and prevalent Christian theme that in order to achieve happiness in this life, one must first conquer the fear of death. Happiness, therefore, is not a cure; it is a consequence of the cure.

DEATH WITHOUT CONSOLATION. In sharp contrast to this last position is that of a long line of nineteenth-century and more recent philosophers, from Schopenhauer to contemporary existentialists. For them human well-being or happiness, at least as traditionally conceived, is totally impossible to achieve; and if the individual is to experience such rewarding values as life does permit, he must uncompromisingly embrace the tragedy of the human condition, clearheadedly acknowledging such evils as death. Like the Stoics, these authors would

have us think constantly of death. Unlike the Stoics, however, they do not offer us the consolation of belief in a providential order of nature. From the standpoint of Being or Nature, the death of the individual is totally meaningless or absurd.

For Schopenhauer the finite, empirical self is a manifestation of a cosmic will that has destined man to live out his life in suffering or painful striving. The only remedy is to achieve a state of indifference or pure will-lessness—a state best known in moments of pure aesthetic contemplation but to which the awareness of death substantially contributes.

According to Friedrich Nietzsche, the superior man will not permit death to seek him out in ambush, to strike him down unawares. The superior man will live constantly in the awareness of death, joyfully and proudly assuming death as the natural and proper terminus of life.

Heidegger and Jean-Paul Sartre, like most existentialists, urge us to cultivate the awareness of death chiefly as a means of heightening our sense of life. The knowledge of death gives to life a sense of urgency that it would otherwise lack. The same point has been made by Freud, who compared life without the consciousness of death to a Platonic romance or to a game played without stakes.

Heidegger makes the additional claim, although here Sartre parts company with him, that the awareness of death confers upon man a sense of his own individuality. Dying, he says, is the one thing no one can do for you; each of us must die alone. To shut out the consciousness of death is, therefore, to refuse one's individuality and to live inauthentically.

See also Augustine, St.; Camus, Albert; Epictetus; Epicurus; Euthanasia; Existentialism; Freud, Sigmund; Heidegger, Martin; Immortality; Kierkegaard, Søren Aabye; La Rochefoucauld, Duc François de; Leonardo da Vinci; Life, Meaning and Value of; Marcus Aurelius Antoninus; Nietzsche, Friedrich; Platonism and the Platonic Tradition; Reincarnation; Sartre, Jean-Paul; Scheler, Max; Schopenhauer, Arthur; Seneca, Lucius Annaeus; Spinoza, Benedict (Baruch) de; Stoicism; Unamuno y Jugo, Miguel de; Voltaire, François-Marie Arouet de.

Bibliography

Jacques Choron, *Death and Western Thought* (New York: Collier, 1963), is a fairly comprehensive review of what Western philosophers have had to say on the subject of death. It is especially recommended for its wealth of quotations. Herman Feifel, ed., *The Meaning of Death* (New York: McGraw-Hill, 1959), is an anthology containing many contemporary essays by psychologists, sociologists, and workers in allied fields.

Russell's views on death may be found in "The Art of Growing Old," in his *Portraits from Memory and Other Essays* (London: Allen and Unwin, 1956) and in "What I Believe," in his *Why I Am Not a Christian* (London: Allen and Unwin, 1957).

Heidegger's views will be found in his *Being and Time* (New York: Harper, 1962), Part II, Ch. 1; Sartre's in his *Being and Nothingness* (New York: Philosophical Library, 1956), pp. 531–553.

An interesting work by a Roman Catholic existentialist is Paul-Louis Landsberg, *Essai sur l'experience de la mort* (Paris, 1951).

Freud's views are expressed in "Thoughts for the Times on War and Death," in *Standard Edition of the Complete Psychological Works of Sigmund Freud,* edited by James Strachey and Anna Freud (London: Hogarth Press, 1957), Vol. XIV, 288–317.

Johan Huizinga, *The Waning of the Middle Ages* (London, 1952), provides the best account of that author's reflections.

For an analysis of primitive attitudes toward death, see Lucien Lévy-Bruhl, *Primitive Mentality,* translated by Lilian A. Clare (London: Allen and Unwin, 1923).

Robert G. Olson (1967)

DEATH [ADDENDUM]

In recent decades death has garnered considerable philosophical attention in three principal areas: medical ethics, value theory, and metaphysics.

In medical ethics, interest has centered on determining the criterion of death. The most common criterion for the death of human beings is irreversible loss of consciousness, but this postulate remains controversial for those who see humans principally as animals rather than as conscious beings. A human animal may be said to be alive even if he or she is not conscious. The meaning of "irreversibility" is also controversial. Would Jane Doe be dead if she lost consciousness but her consciousness could be recovered if she were injected with a serum that will not be invented for another thousand years? Medical ethicists have also debated the relevance of the principle of double effect regarding situations in which a doctor is the cause of death. Followers of the principle of double effect will allow morphine to be administered to a patient in order to stop his suffering, even allowing for the possibility that the morphine will kill the patient. But this outlook does not countenance such an injection if the patient's death is the main objective of the act.

In the theory of values, the argument of Epicurus (341–270 BCE) still engenders controversy. Epicurus argued that if death is the annihilation of a person, then the death cannot be bad for that person. For something to be bad, there must be some thing or person for whom it is bad, but if the person who has died has ceased to be, then (for that person) there is no bad or evil. Epicurus employed this argument to alleviate the fear of death. Epicurus also argued that just as one does not fear or bemoan one's nonexistence before birth, one ought not to lament one's nonexistence after death.

Epicurus's reasoning has elicited a number of counterarguments. Some have argued that death is bad for the person insofar as it involves missed opportunities or possibilities. The death of a person involves his absence, and such an absence can be a bad state of affairs even if it is not bad for the person or for others. It has also been suggested that the Epicurean bid for a symmetry of neither regretting past nonexistence or future nonexistence is not plausible because a person may well wish she had been alive in some earlier era.

Philosophical reflection on death has implications for an overall theory of values, involving questions about rights (if the dead have ceased to be, do they have any rights whatsoever?), promise-keeping (do you have a duty to keep a promise to someone who has ceased to be?), and the environment (should environmental harms be understood principally in terms of the death of individual animals and plants or of species)?

Still other thinkers have wrestled with the issue of whether or not the death of humans is in fact a case of annihilation. This debate has unfolded mainly among specialists in the metaphysics of personal identity, philosophy of mind, and philosophy of religion. Dualist conceptions of human nature—according to which there is a distinction between a person or soul or mind on the one hand and his or her body on the other—allow for the possibility that the physical body may be annihilated without destroying the person or soul or mind. The case for dualism and the coherence of disembodiment has been bolstered by reports of out-of-the-body experiences (OBE). Even if the empirical testing of such OBEs is inherently problematic, some dualists have appealed to the apparent coherence of such reports in arguing that disembodiment is a bona fide possibility. Philosophical speculation on the afterlife contains a great deal of debate on the coherence of competing thought experiments. It may appear that materialist accounts of persons, according to which a person *is* his or her body, are ill-suited to lending any credibility to an afterlife scenario, but late-twentieth-century philosophers have challenged this conclusion. Some have argued that a physical object (such as a human body) can cease to exist and then be re-created. This process entails what has been called a "gap inclusive" or "intermittent" identity, for it posits an interval when a person temporarily ceases to exist.

Philosophers have also advanced a constitutional argument, according to which human beings are material objects because they are constituted exclusively by material objects. The constitutional relationship is not one of strict identity, so it is possible for the same thing (a person) to survive the change of his or her constitutive parts. Some argue that a person (understood as a material body) can survive the loss of his or her present body or perhaps come to have a new body or even be reconstituted by nonphysical individuals. The philosophical investigation into the possibility of an afterlife is often located in the broader philosophical enterprise of weighing the merits of naturalism, theism, and various nontheistic religious traditions.

Bibliography

Almeder, Robert. *Beyond Death: The Evidence for Life After Death.* Springfield, IL: Charles C. Thomas, 1987.

Baker, Lynne Rudder. *Persons and Bodies.* Cambridge, U.K.: Cambridge University Press, 2000.

Feldman, Fred. *Confrontations with the Reaper: A Philosophical Study of the Nature and Value of Death.* New York: Oxford University Press, 1992.

Hick, John. *Death and Eternal Life*. Louisville, KY: Westminster John Knox, 1994).

Moody, Raymond. *Life After Life: The Investigation of a Phenomenon—Survival of Bodily Death.* Bantam Doubleday Dell, 1976.

Nagel, Thomas. *Mortal Questions.* Cambridge, U.K.: Cambridge University Press, 1979.

Taliaferro, Charles. *Consciousness and the Mind of God.* Cambridge, U.K.: Cambridge University Press, 1994.

Taliaferro, Charles. "Sensibility and Possibilia: A Defense of Thought Experiments." *Philosophia Christi* 3:2, 201 (2001): 403–420.

Rick Fairbanks (2005)
Charles Taliaferro (2005)

DECISION THEORY

Decision theory provides a general, mathematically rigorous account of decision making under uncertainty. The subject includes rational choice theory, which seeks to formulate and justify the normative principles that govern optimal decision making, and descriptive choice theory, which aims to explain how human beings actually

make decisions. Within both these areas one may distinguish individual decision theory, which concerns the choices of a single agent with specific goals and knowledge, and game theory, which deals with interactions among individuals. This entry will focus on rational choice theory for the single agent, but some descriptive results will be mentioned in passing.

DECISION PROBLEMS

It is standard to portray decision makers as facing choices among acts that cause desirable or undesirable consequences when performed in various states of the world. Acts characterize those aspects of the world that an agent can directly control. States specify contingencies beyond her control that might influence the consequences of acts. Each combination of an act A and state S fixes a unique consequence $A(S)$ that describes the result of doing A in S. When there are only finitely many acts and states the decision situation can be represented as a matrix:

	S_1	$S_2 \ldots$	S_n
A_1	$A_1(S_1)$	$A_1(S_2) \ldots$	$A_1(S_n)$
A_2	$A_2(S_1)$	$A_2(S_2) \ldots$	$A_2(S_n)$
A_3	$A_3(S_1)$	$A_3(S_2) \ldots$	$A_3(S_n)$
:	:	:	:
A_m	$A_m(S_1)$	$A_m(S_2) \ldots$	$A_m(S_n)$

The agent decides the row, the world decides the column, and these together determine the consequence.

In any well-formed decision problem (1) the value of each consequence is independent of the act and state that bring it about, (2) each consequence is sufficiently detailed to settle every matter about which the agent intrinsically cares, (3) neither acts nor states have any value except as a means for producing consequences, and (4) the agent will not believe that she has the ability to causally influence which state obtains. When these conditions are met, the agent's goals and values affect her decision only via her desires for consequences, and her beliefs influence her choice via her uncertainty about which state obtains. The agent will use her beliefs about states to select an act that provides the best means for securing a desirable consequence.

For theoretical purposes, it is useful to idealize the decision setting by assuming that the repertoire of actions is rich. Specifically, for each consequence c there is a constant act $[c]$ that produces c in every state of the world, and, for any acts A and B, and any disjunction of states E, there is a mixed act $A_E \cup B_{\sim E}$ that produces A's consequence when E holds and B's consequence when $\sim E$ holds. While real agents will typically be unable to realize such recherché prospects as these, imagining that decision makers have attitudes toward them often helps one determine which realistic acts should be performed.

This model applies to one-choice decisions made at a specific time. Early decision theorists believed that sequences of decisions made over time could be reduced to one-shot decisions among contingency plans, or strategies, but this view now has few adherents. The topic of dynamic decision making lies beyond the scope of this entry. For relevant discussions, see Peter Hammond (1988), Edward McClennen (1990), and James M. Joyce (1995).

SUBJECTIVE EXPECTED UTILITY

The central goal of rational choice theory is to identify the conditions under which a decision maker's beliefs and desires rationalize the choice of an action. According to the standard model of decision-theoretic rationality, an action is rational just in case, relative to the agent's beliefs and desires, it has the highest subjective expected utility of any available option. This subjective expected utility (SEU) theory has its roots in the work of Blaise Pascal, Daniel Bernoulli, Vilfredo Pareto, and Frank P. Ramsey, and finds its fullest expression in Leonard J. Savage's *Foundations of Statistics* (1972). According to SEU a rational agent's basic desires can be represented by a utility function $\boldsymbol{u}$ that assigns a real number $\boldsymbol{u}(c)$ to each consequence c. The value of $\boldsymbol{u}(c)$ measures the degree to which c would satisfy the agent's desires and promote his or her aims.

Likewise, the agent's beliefs can be characterized by a subjective probability function $\boldsymbol{P}$ whose values express the agent's subjective degrees of confidence, or credences, in the states of the world. $\boldsymbol{P}$ is assumed to be unique, and $\boldsymbol{u}$ is unique once the choice of a unit and a zero for measuring utilities are fixed. Given $\boldsymbol{P}$ and $\boldsymbol{u}$, the expected utility of each act A is a weighted average of the utilities of its consequences, so that $\boldsymbol{Exp_{P,u}}(A) = \sum_{i=1} \boldsymbol{P}(S_i)\boldsymbol{u}(A(S_i))$. According to the core doctrine of SEU, the choice of an act is rational only if it maximizes the chooser's subjective expected utility, so that $\boldsymbol{Exp_{P,u}}(A) \geq \boldsymbol{Exp_{P,u}}(B)$ for all acts B. This should not be taken to suggest that the agent sees herself as maximizing expected utility, or even that she has the concept of expected utility. SEU does not propose expected utility maximization as a decision procedure, but as a way of assessing the results of such procedures. Rational decision makers merely act as if they maximize subjective expected utility; they need not explicitly do so.

REPRESENTATION OF RATIONAL PREFERENCE

A central challenge for SEU is to find a principled way of characterizing credences and utilities. Following the lead of Ramsey (1931), the standard solution involves proving a representation theorem that shows how an agent's beliefs about states and desires for outcomes are related to her all-things-considered preferences for acts. The agent is assumed to make three sorts of comparative evaluations between acts: She might strictly prefer A to B, written $A > B$, weakly prefer A to B, $A \gtrsim B$, or be indifferent between them, $A \approx B$. These relations hold, respectively, just in case the agent judges that, on balance, A will do more than, at least as much as, or exactly as much as, B will to satisfy her desires and promote her aims. The totality of such evaluations is the agent's preference ranking.

Early decision theorists, motivated by a misguided scientific methodology, thought of preferences as operationally defined in terms of overt choices, so that, by definition, an agent prefers A to B if and only if (iff) she will incur a cost to choose A over B. Even though this sort of behaviorism remains firmly ensconced in some areas of economics, it has been widely and effectively criticized (Sen 1977, Joyce 1999). In the end, preferences are best thought of as subjective judgments of the comparative merits of actions as promoters of desirable outcomes. While such judgments are closely tied to overt choice behavior, the relationship between the two is nowhere near as direct and unsophisticated as behaviorism suggests.

The representation theorem approach seeks to justify SEU by (1) imposing a system of axiomatic constraints on preference rankings, (2) arguing that these express requirements of rationality, and then (3) proving that any preference ranking that satisfies the axioms can be associated with a probability P and a utility u such that each of $A >, \gtrsim, \approx B$ hold iff, respectively, $\boldsymbol{Exp}_{P,u}(A) >, \geq, = \boldsymbol{Exp}_{P,u}(B)$. An agent whose preferences can be represented in this way evaluates acts as if she were aiming to maximize expected utility relative to $\boldsymbol{P}$ and $\boldsymbol{u}$.

FRAME INVARIANCE

All versions of SEU share a common set of core principles. The first says that logically equivalent redescriptions of prospects should not alter preferences.

> *SEU_1 Frame Invariance.* The evaluation of an act should not depend on how its consequences happen to be described.

People often violate this constraint. Consider the following two decision framings due to E. Shafir and A. Tversky (1995):

- You receive \$300, and are then given a choice between getting another \$100 for sure or getting \$200 or \$0 depending on the toss of a fair coin.
- You receive \$500, but are then forced to choose between returning \$100 for sure or returning \$200 or \$0 depending on the toss of a fair coin.

Since both decisions offer a sure \$400 or a fifty-fifty chance of \$300 or \$500, SEU_1 requires agents to make the same choice in each case (though it does not tell them which choice to make). As it turns out, most people make the safe choice in the first case and take the sure \$400, but they make risky choice in the second case by taking the fifty-fifty gamble. Cognitive psychologists attribute this violation of SEU_1 to the following two irrational tendencies of human decision makers:

> Divergence from Status Quo. People are more concerned with incremental gains and losses, seen as changes in the status quo, than with total well-being or overall happiness.

> Asymmetrical Risk Aversion. People eschew risk when pursuing gains, but to seek risk when avoiding losses.

Under the first description, where the status quo is \$300, people see themselves as trying to secure an additional gain, and so opt for the safe alternative. Under the second description, where the status quo is \$500, people see themselves avoiding losses, and so incline toward the risky choice. These divergent attitudes are irrational given that the options are effectively identical.

VALUE INDEPENDENCE

The second principle requires each act to have a value that depends only on the values and probabilities of the outcomes it might cause.

> SEU_2 *Value Independence*. If the agent prefers A to B in a decision where C is not an option, then she should still prefer A to B even if C is an option, provided that C's inclusion does not provide any information about state probabilities.

Apparent counterexamples to SEU_2 as a requirement of rationality always involve violations of the proviso. For example, R. Duncan Luce and Howard Raiffa (1957) discuss a diner who, thinking he is in a greasy spoon, prefers salmon to steak, but then orders steak when told that snails are on the menu. SEU_2 is vindicated by the obser-

vation that the availability of snails provides the diner with evidence for thinking that he is in fine restaurant, and this alters his views about the comparative merits of the salmon and steak. Other common violations of SEU_2 are clearly irrational. For example, D. Redelmeier and E. Shafir (1995) show that physicians are less likely to prescribe ibuprofen to patients in pain when they have the option of prescribing the inferior drug piroxicam than when piroxicam is unavailable. While this sort of behavior does not discredit SEU_2 as a normative principle, it does show that it is inaccurate as a description of human behavior.

ORDERING

The third principle rules out preference cycles in which $A \gtrsim B$, $B \gtrsim C$, but $A > C$, and it requires that the preference ranking be complete in the sense that exactly one of $A > B$, $A \approx B$ or $B > A$ always hold.

> SEU_3 *Ordering*. Preference rankings completely order the set of acts.

Though some dispute anticyclicality, and Peter C. Fishburn (1991) even develops an acyclic decision theory, the prohibition against cycles remains among the most widely accepted principles of rational preference. On views that equate preferences and choices, preference cycles are irrational because they leave the agent open to exploitation as a "money pump": she will freely trade *C* for *B* and *B* for *A*, and then pay a fee to exchange *C* for *A*, thereby getting nothing for something. Even if choice is not equated with preference, cycles are still problematic. Many seemingly rational cycles treat preferences as partial, rather than all-things-considered evaluations. For instance, one might prefer an expensive shirt to a moderately priced one on the basis of style, and prefer the moderately priced shirt to a cheap shirt on the basis of durability, but prefer the cheap shirt to the expensive one on the basis of price. Here what seems to be a rational preference cycle is really a failure to integrate considerations of style, durability, and price into an all-things-considered value judgment.

Failures of evaluative discrimination can also seem to generate rational preference cycles. Suppose a vinophile, who cares only about how his wine tastes, cannot taste any difference between wine *A* and wine *B*, or between wine *B* and wine *C*, but can taste that *C* is better than *A*. It is tempting to think that the vinophile should be indifferent between *A* and *B* and between *B* and *C*, but should prefer *C* to *A*. A clearer understanding of the situation shows that this is incorrect. A person should only be indifferent between prospects when he lacks any reason, on balance, for preferring one to the other. The vinophile, however, has reason to favor *B* over *A* since *B* is indistinguishable in taste from a wine superior to *A*. He also has reason to favor *C* over *B* since *B* is indistinguishable from a wine inferior to *C*. Properly speaking, then, the vinophile is not indifferent between *A* and *B* or between *B* and *C*: his preferences run $B \gtrsim A$, $C \gtrsim B$, and $C > A$, but neither $A \approx B$ nor $B \approx C$ is true.

One might worry that the vinophile's reasons seem insufficient to justify strict preferences. It would, for example, be silly for him to pay anything to trade a bottle of *A* for a bottle of *B* (unless he could convert the latter into a bottle of *C* for a small enough fee). While this is a legitimate concern, it tells against completeness rather than anticyclicality. When an agent cannot precisely discriminate the qualities of prospects on which his evaluations depend, or when these qualities are themselves vague or indeterminate, his preference ranking will be incomplete: for certain options, all three of $A > B$, $A \approx B$, and $B > A$ will fail. Sometimes both $A \gtrsim B$ and $B \gtrsim A$ will fail as well, in which case the agent has no views about the comparative merits of *A* and *B*. Alternatively, as in the vinophile example, the agent might determinately weakly prefer *B* to *A* even though he neither strictly prefers *B* to *A* nor is indifferent between them. So, while $A \approx B$ and $B > A$ each entail $B \gtrsim A$, the latter is consistent with the falsity of both $A \approx B$ and $B > A$. Besides indeterminacy or vagueness in values, incompleteness in preferences can arise via an imprecision in credences. In both sorts of cases it can be perfectly rational to have an incomplete preference ranking.

One response to these considerations, which is advocated in Isaac Levi (1980), Richard Jeffrey (1983), and Mark Kaplan (1983), is to construe SEU_3's completeness clause as a requirement of coherent extendibility. Instead of asking an agent to completely order acts, one demands merely that there be at least one complete preference ranking (usually there will be many) that satisfies all other requirements of rationality, and that agrees with the agent's preferences whenever she has definite preferences. One then represents vague or indeterminate preferences by giving up the idea that the agent's attitudes can be modeled by a single probability/utility pair (given a unit and zero for utility). Rather, there will be a representing set $\boldsymbol{R}$ of $(\boldsymbol{P}, \boldsymbol{u})$ pairs that agree with the agent's preferences in the sense that, for any options *A* and *B*, each of $A >, \gtrsim, \approx B$ hold iff, respectively, $\boldsymbol{Exp}_{P,u}(A) >, \geq, = \boldsymbol{Exp}_{P,u}(B)$ holds for every $(\boldsymbol{P}, \boldsymbol{u})$ pair in $\boldsymbol{R}$. Act *A* is unambiguously choice worthy only if maximizes expected utility relative to every $(\boldsymbol{P}, \boldsymbol{u})$ pair in $\boldsymbol{R}$. It is admissible when it maxi-

mizes expected utility relative to some such pair. There is no generally accepted procedure for handling situations where no admissible act is unambiguously choice worthy. Some theorists would say that the agent's beliefs and desires are too indefinite to justify any choice as rational. Others, most notably Levi (1980), maintain that principles of decision making that outrun expected utility maximization come into play in this situation. For example, Levi allows agents to decide among admissible options using *maximin*, that is, by selecting the act whose worst consequence is at least as good as the worst consequence of any alternative.

COMPARATIVE PROBABILITY

The next principle of SEU forges a link between rational preference and rational belief. A wager on event E is an act of the form $[c]_E \cup [d]_{\text{-}E}$ where $[c] > [d]$. Such a wager produces the desirable consequence c in every state consistent with E and the undesirable consequence d in every state consistent with ˜E. Intuitively, a person should prefer such a wager more strongly the more likely she takes E to be. More precisely, given any events E and F, $[c]_E \cup [d]_{\text{-}E}$ should be preferred to $[c]_F \cup. [d]_{\text{-}F}$ exactly if E as more probable than F. The following axiom is meant to ensure that this is so.

> SEU_4 *Comparative Probability*. Assuming $[c] > [d]$, if the agent prefers $[c]_E \cup [d]_{\text{-}E}$ to $[c]_F \cup [d]_{\text{-}F}$, she must also prefer $[c^*]_E \cup [d^*]_{\text{-}E}$ to $[c^*]_F \cup [d^*]_{\text{-}F}$ for any consequences such that $[c^*] > [d^*]$.

SEU_4 can seem implausible when the values of consequences vary with the world's state. Suppose, for example, that c and d are monetary fortunes that one might have in ten years, say c = \$500,000 and d = \$400,000. Let E and F be hypotheses about the cumulative rate of inflation over the decade: E puts the figure at 60 percent, while F puts it at 10 percent. Even if one regards E as the more probable hypothesis, one might still prefer to wager on F since one's fortune will be worth more if F is true.

There are two standard responses to this problem. Savage (1972) maintains that decision problems of this sort, in which the values of consequences depend on states, are ill formed. He argues any such problem could be transformed into a well-formed decision by a suitable subdivision of consequences. In the previous example, c would be split into c_1 = "\$500,000 after cumulative inflation of 60 percent," and c_2 = "\$500,000 after cumulative inflation of 10 percent." Alternatively, one might opt for a state-dependent utility theory, which replaces SEU_4 by a weaker condition and allows the values of consequences of vary with states (for details, see Karni 1993; Schervish, Seidenfeld, and Kadane 1990).

INDEPENDENCE AND THE SURE-THING PRINCIPLE

The most controversial tenet of SEU is the independence axiom:

> SEU_5 *Independence*. Preference among acts that have exactly the same consequences when E is false should depend exclusively on whathappens when E is true. If $A_E \cup C_{\text{-}E}$ is preferred to $B_E \cup C_{\text{-}E}$ forsome act C, then $A_E \cup D_{\text{-}E}$ is preferred to $B_E \cup D_{\text{-}E}$ for all acts D.

To illustrate, consider the following act types, where c, d, c^* and d^* are known consequences, and x ranges over possible consequences.

	S_1	S_2	S_3
A_x	c	d	x
B_x	c^*	d^*	x

SEU_5 says that an agent's preference between A_x and B_x should not depend on x's value. More generally, it requires agents to have well-defined conditional preferences: A is preferred to B in the event of E just in case $A_E \cup C_{\text{-}E} > B_E \cup C_{\text{-}E}$ for some (hence any) C.

SEU_5 has the following intuitive consequence:

> *Sure-Thing Principle*: Let $E_1, E_2, \dots, E_n$ be mutually exclusive, collectively exhaustive events. If A is weakly preferred to B conditional on each E_i, then A is weakly preferred to B simpliciter. Moreover, if A is strictly preferred to B conditional on some event that is not judged certainly false, then A is strictly preferred to B.

Independence and the sure-thing principle have been quite controversial. Some apparent failures of SEU_5 arise in ill-formed decision problems whose states are not independent of acts. For example, imagine a man who has to drive home from a party where alcohol is being served. He likes to drink, but worries about getting home safely. Suppose he frames his decision like this:

	Car accident	No accident
Drink	–100	1
Teetotal	–101	0

Since the consequences of drinking are better that those of refraining both in the event of an accident and otherwise, it looks as if the sure-thing principle advocates

drinking, which is clearly bad advice given that drinking increases the probability of an accident. Problems of this sort led Jeffrey (1983) to develop an evidential version of decision theory in which independence is only valid for decisions in which acts provide no evidence about the occurrence of any state. Reflections on Newcomb problems, in which acts and states are causally independent but evidentially correlated, led causal decision theorists like Robert Stalnaker (1981), Allan Gibbard and William Harper (1978), and Brian Skyrms (1980) to insist that the two principles be restricted to decisions in which the choice of an act has no causal influence over states.

The most famous objections to SEU_5 are the paradoxes of Maurice Allais (1953) and Daniel Ellsberg (1961), which seem to show that SEU rules out certain rational attitudes toward risk and uncertainty. An act involves risk when the agent knows the objective probabilities with which its consequences will obtain. It involves uncertainty when the agent's information allows a range of possible risk profiles for consequences. SEU_5 entails that, insofar as decision making is concerned, all legitimate considerations of risk and uncertainty are fully captured in expected utilities. The Allais and Ellsberg paradoxes suggest, to the contrary, that risk and uncertainty are nonseparable quantities: one cannot express them as weighted averages of their values conditional on disjoint events. If this is correct, then an agent need not have any fixed preference between the act types A_x and B_x because *x*'s value might provide information about the relative risk or uncertainty of the two options, and this information might justifiably influence the agent's preferences.

The Allais paradox envisions an agent who chooses between *A* and *A** and then between *B* and *B** (with the know probabilities listed).

	0.10	0.01	0.89
A	$1,000,000	$1,000,000	$1,000,000
*A**	$5,000,000	$0	$1,000,000
B	$1,000,000	$1,000,000	$0
*B**	$5,000,000	$0	$0

Empirical studies show that people systematically violate independence when presented with such choices. They "play it safe" and select *A* over *A** in the first choice, but favor the riskier option *B** over *B* in the second. The standard rationale for these choices assumes (1) that there is more risk involved in choosing *A** over *A* than there is in choosing *B** over *B*, and (2) that it is rational minimize this risk even when doing so violates independence.

Ellsberg's paradox shows something similar with respect to judgments of uncertainty. Suppose a ball will be drawn at random from an urn that holds thirty red balls and sixty white or blue balls in an unknown proportion. One chooses between *C* and *C** and then between *D* and *D**.

	Red	White	Blue
C	$100	$0	$0
*C**	$0	$100	$0
D	$100	$0	$100
*D**	$0	$100	$100

Here most people prefer *C* to *C** and *D** to *D*. Interestingly, when gains are replaced by losses, people still violate independence, but both choices are reversed. People thus seem to prefer risk to uncertainty when they have something to gain, but prefer uncertainty to risk when they have something to lose. Those who regard Ellsberg's paradox as a counterexample to SEU maintain that such nonseparable preferences for risk over uncertainty or uncertainty over risk are entirely rational.

Some proponents of SEU (see Broome 1991) respond by arguing that the consequences in the Allais and Ellsberg paradoxes are underdescribed. For example, the standard pattern of preferences in Allais can be rationalized by noting that, when the 0.01 event occurs, agents who choose *A** over *A* may feel regret (because they passed up a sure thing), while those who choose *B** over *B* will feel no regret (because they probably would have ended up with nothing anyhow). For such agents, the decision matrix really looks like this:

	0.10	0.01	0.89
A	$1,000,000	$1,000,000	$1,000,000
*A**	$5,000,000	$0 with regret	$1,000,000
B	$1,000,000	$1,000,000	$0
*B**	$5,000,000	$0 without regret	$0

Likewise, if an agent feels uneasy when gains ride on uncertain prospects (or losses ride on risky prospects), then the correct description of the Ellsberg problem is this:

	Red	White	Blue
C	$100	$0	$0
*C**	$0 with uneasiness	$100 with uneasiness	$0 with uneasiness
D	$100 with uneasiness	$0 with uneasiness	$100 with uneasiness
*D**	$0	$100	$100

If these matrices accurately describe the decisions, then neither the Allais or Ellsberg paradoxes provide a genuine counterexample to SEU_5.

These sorts of rationalizing responses are weakened by their dependence on substantive assumptions about the psychology of risk, uncertainty, and regret that are not universally accepted (see Loomes and Sudgen 1982, Weber 1998). An alternative is to argue that the usual preferences in the Allais and Ellsberg paradoxes are simply irrational. In Allais, for example, agents assume that the disparity in risk between *A* and *A** exceeds the disparity in risk between *B* and *B**. This may be a mistake. One way to determine differences in risk is to consider the costs of insuring against the incremental risk one incurs by trading one option for another. Someone who switches from *A** to *A* in Allais can offset this risk by purchasing an insurance policy that pays out $1,000,000 contingent on the 0.01 event. Notice, however, that the risk incurred by switching from *B** to *B* can be offset by the same policy. Since a single policy eliminates both risks there is reason to think that the actual change in risk is the same in each case. Similar things can be said about the Ellsberg choosers, who implicitly assume that they decrease their uncertainty more by switching from *C** to *C* than they do by switching from *D** to *D*. So, if one measures disparities in risk or uncertainty by the costs of insuring against it, then SEU is safe from the Allais and Ellsberg examples.

Opponents of SEU will, of course, deny that risks should be measured by the costs of insuring against them. Ultimately, the issue will be resolved by the development of a convincing measure of risk. While there is a well-known theory of risk aversion within SEU, there is no universally accepted method for quantifying risk itself. The best work in this area, which builds on M. Rothschild and J. E. Stiglitz (1970), suggests that risk is indeed separable.

ALTERNATIVES TO SEU

While subjective expected utility theory remains firmly ensconced as the standard model of rational decision making for individuals, a number of alternatives have been developed. One kind of approach seeks to relax independence while preserving most other aspects of SEU. Especially noteworthy here is the "generalized expected utility analysis" of Mark Machina (1982), and the "weighted utility model" of Soo-Hong Chew and Kenneth R. MacCrimmon (1979). Alternatively, one can reject maximizing conceptions of rationality altogether and see decision making as matter of *satisficing* relative to fixed constraints. For example, G. Gigerenzer et al. (1999) seek to replace the single all-purpose prescription to maximize expected utility by an ecological model of rationality in which decision makers employ a set of simple, highly localized decision heuristics. These heuristics efficiently generate choices that produce desirable consequences in the contexts where they tend to be employed, but they can go badly awry when used in out of context. For discussion of further nonstandard decision theories, see Robert Sugden (2004).

Interesting though these alternatives are, none has seriously challenged the normative status of SEU. Though highly idealized, and far from adequate as a description of human behavior, SEU remains the best overall account of rational decision making.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Game Theory; Pareto, Vilfredo; Pascal, Blaise; Probability and Chance; Ramsey, Frank Plumpton; Savage, Leonard; Sen, Amartya; Statistics, Foundations of.

Bibliography

Allais, Maurice. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'École Americaine." *Econometrica* (1953).

Broome, John. *Weighing Goods: Equality, Uncertainty, and Time*. Cambridge, MA: Basil Blackwell, 1991.

Chew, Soo-Hong, and Kenneth R. MacCrimmon. "Alpha-nu Theory: A Generalization of Expected Utility Thoery." *Working Paper* 669 (1979). University of British Columbia, Vancouver.

Ellsberg, Daniel. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics* 75 (1961): 643–669.

Fishburn, Peter C. 1991. "Nontransitive Preferences in Decision Theory." *Journal of Risk and Uncertainty* 4 (1991): 113–134.

Gibbard, Allan, and William Harper. "Counterfactuals and Two Kinds of Expected Utility." In *Foundations and Applications of Decision Theory*, edited by C. Hooker, J. Leach, and E. McClennen. Dordrecht, Netherlands: D. Reidel, 1978.

Gigerenzer, G, et al. *Simple Heuristics that Make Us Smart*. New York: Oxford University Press, 1999.

Hammond, Peter. "Consequentialist Foundations for Expected Utility." *Theory and Decision* 25 (1988): 25–78.

Jeffrey, Richard. *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press, 1983.

Joyce, James M. *The Foundations of Causal Decision Theory*. New York: Cambridge University Press, 1999.

Kaplan, Mark. "Decision Theory as Philosophy." *Philosophy of Science* 50 (1983): 549–577.

Karni, E. "Subjective Expected Utility Theory with State-Dependent Preferences." *Journal of Economic Theory* 60 (1993): 428–438.

Levi, Isaac. *The Enterprise of Knowledge*. Cambridge, MA: MIT Press, 1980.

Loomes, Graham, and Robert Sugden. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal* 92 (1982): 805–824.

Luce, R. Duncan, and Howard Raiffa. *Games and Decisions*. New York: Wiley, 1957.

Machina, Mark. "'Expected Utility' Analysis without the Independence Axiom." *Econometrica* 50 (1982): 277–323.

McClennen, Edward. *Rationality and Dynamic Choice: Foundational Explorations*. New York: Cambridge University Press, 1990.

Ramsey, Frank P. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*, edited by Richard Braithwaite. London: Kegan Paul, 1931.

Redelmeier, D., and E. Shafir. "Medical Decision Making in Situations that Offer Multiple Alternatives." *Journal of the American Medical Association* 273 (4) (1995): 302–305.

Rothschild, M., and J. E. Stiglitz. "Increasing Risk: I. A Definition." *Journal of Economic Theory* 2 (1970): 225–243.

Savage, Leonard J. *The Foundations of Statistics*. 2nd ed. New York: Dover, 1972.

Schervish, M., T. Seidenfeld, and J. Kadane. "State-Dependent Utilities." *Journal of the American Statistical Association* 85 (1990): 840–847.

Sen, Amartya. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy and Public Affairs* 6 (1977): 317–344.

Shafir, E., and A. Tversky. "Decision Making." In *An Invitation to Cognitive Science*. Vol. 3, *Thinking*. 2nd ed., edited by Edward Smith and Daniel Osherson, 77–100. Cambridge, MA: MIT Press, 1995.

Skyrms, Brian. *Causal Necessity*. New Haven, CT: Yale University Press, 1980.

Stalnaker, Robert. "Letter to David Lewis, May 21, 1972." In *IFs: Conditionals, Belief, Decision, Chance, and Time*, edited by Robert Stalnaker, William Harper, and Glen Pearce. Dordrecht, Netherlands: D. Reidel, 1981.

Sugden, Robert. "Alternatives to Expected Utility: Formal Theories." In *Handbook of Utility Theory*. Vol. 2, edited by Peter Hammond, Salvador Barberà, and Christian Seidl. Dordrecht, Netherlands: Kluwer Academic, 2004.

Weber, Michael. "The Resilience of the Allais Paradox." *Ethics* 109 (1998): 94–118.

James M. Joyce (2005)

DECONSTRUCTION

Deconstruction is a philosophical-critical approach to textual analysis that is most closely associated with the work of Jacques Derrida in philosophy and the Yale School (Paul DeMan, J. Hillis Miller, Geoffrey Hartman) in literary theory and criticism. Derrida draws the term *déconstruction* from his interpretation of Martin Heidegger as a way to translate two Heideggerian terms: *Destruktion*, which means not destruction but a destructuring that dismantles the structural layers in a system; and *Abbau*, which means to take apart an edifice in order to see how it is constituted or deconstituted. For Derrida, then, deconstruction, in the context of philosophy, refers to a way to think the structured genealogy of philosophy's concepts while exposing what the history of these concepts has been able to obscure or exclude. By displaying those concepts that the philosophical tradition both authorizes and excludes, a deconstructive reading seeks to work within the closed field of metaphysical discourse without at the same time confirming that field. Instead, it allows a text to dismantle itself by exposing the internal inconsistencies and implicit significations that lie concealed within the language of the text.

One way to understand deconstruction is in terms of a critique of the binary, oppositional thinking that, for Derrida, is central to the history of philosophy. This is to say, each term in the Western philosophical/cultural lexicon is accompanied by its binary opposite: intelligible/sensible, truth/error, speech/writing, reality/appearance, mind/body, culture/nature, good/evil, male/female, and so on. These oppositions do not peacefully coexist, however: one side of each binary opposition has been privileged and the other side devalued. A hierarchy has been established within these oppositions, as the intelligible has come to be valued over the sensible, mind has come to be valued over body, and so on. The task of deconstruction is to dismantle or deconstruct these binary oppositions: to expose the foundational choices of the philosophical tradition and to bring into view what the tradition has repressed, excluded, or—to use the Derridean terminology—marginalized.

As a critical practice, the deconstruction of these oppositions involves a double movement of overturning and displacement. The first phase initiates an overturning of the hierarchy that valorizes the term traditionally subordinated by the history of philosophy: for example, privileging writing over speech, signifier over signified, or the figurative over the literal. But this privileging is temporary and strategic, for in overturning a metaphysical hierarchy, deconstruction seeks to avoid reappropriating the same hierarchical structure; it is the hierarchical oppositional structure itself that underwrites the metaphysical tradition, and to remain within the binary logic of metaphysical thinking will only reestablish and affirm these oppositions. The second phase of deconstruction destabilizes the inversion by showing the arbitrary nature of the process of hierarchical valorization itself and displaces the hierarchy altogether by intervening with a new "undecidable" term—for example, *difference*, *trace*, *pharmakon*, *supplement*—that resists the formal structure imposed by the binary logic of philosophical opposition. Much of Derrida's early work involves elucidating the play of these undecidables: the play of *différence*, which both differs

and defers; the play of the trace, which is both present and absent; the play of the *pharmakon*, which is both poison and cure; the play of the supplement, which is both surplus and lack. By displaying the choices by means of which the philosophical tradition constitutes itself *as* a tradition, Derridean deconstruction opens the possibility to think difference other than as opposition and hierarchy.

Within literary criticism, the deconstructive method is used to show that the meaning of a literary text is not fixed and stable. Instead, by exploring the dynamic tension within a text's language, literary deconstruction reveals the literary work to be not a determinate object with a single correct meaning but an expanding semantic field that is open to multiple, sometimes conflicting interpretations.

See also Structuralism and Post-structuralism.

Bibliography

Bloom, Harold, Paul DeMan, Jacques Derrida, Geoffrey Hartman, and J. Hillis Miller. *Deconstruction and Criticism*. London: Continuum, 1979.

Culler, Jonathan D. *On Deconstruction: Theory and Criticism in the 1970s*. Ithaca, NY: Cornell University Press, 1982.

DeMan, Paul. *Allegories of Reading: Figural Language in Rousseau, Nietzsche, Rilke, and Proust*. New Haven, CT: Yale University Press, 1979.

DeMan, Paul. *Blindness and Insight: Essays in the Rhetoric of Contemporary Criticism*. 2nd ed. Minneapolis: University of Minnesota Press, 1983.

Derrida, Jacques. *Dissemination*. Translated by Barbara Johnson. Chicago: University of Chicago Press, 1981. Originally published as *La Dissémination* (Paris: Du Seuil, 1972).

Derrida, Jacques. *Margins of Philosophy*. Translated by Alan Bass. Chicago: University of Chicago Press, 1982. Originally published as *Marges de la Philosophie* (Paris: Éditions de Minuit, 1972).

Derrida, Jacques. *Of Grammatology*. Translated by Gayatri Chakravorty Spivak. Baltimore, MD: Johns Hopkins University Press, 1976. Originally published as *De la Grammatologie* (Paris: Éditions de Minuit, 1967).

Derrida, Jacques. *Positions*. Translated by Alan Bass. Chicago: University of Chicago Press, 1981. Originally published as *Positions: Entretiens avec Henri Ronse, Julia Kristeva, Jean-Louis Houdebine, Guy Scarpetta* (Paris: Éditions de Minuit, 1972).

Derrida, Jacques. *Speech and Phenomena, and Other Essays on Husserl's Theory of Signs*. Translated by David B. Allison. Evanston, IL: Northwestern University Press, 1973. Originally published as *La voix et le phénomène: Introduction au problème du signe dans la phénoménologie de Husserl* (Paris: Presses Universitaires de France, 1967).

Derrida, Jacques. *Writing and Difference*. Translated by Alan Bass. Chicago: University of Chicago Press, 1978. Originally published as *L'écrituré et la difference* (Paris: Éditions de Seuil, 1967).

Gasche, Rodolphe. *The Tain of the Mirror: Derrida and the Philosophy of Reflection*. Cambridge, MA: Harvard University Press, 1988.

Hartman, Geoffrey. *Criticism in the Wilderness*. New Haven, CT: Yale University Press, 1982.

Miller, J. Hillis. *The Ethics of Reading: Kant, De Man, Eliot, Trollope, James, and Benjamin*. New York: Columbia University Press, 1989.

Miller, J. Hillis. *The Linguistic Moment: From Wordsworth to Stevens*. Princeton, NJ: Princeton University Press, 1985.

Norris, Christopher. *Deconstruction: Theory and Practice*. London: Methuen, 1982.

Alan D. Schrift (2005)

DEDUCTION

See *Logic, History of: Modern Logic: From Frege to Gödel*

DE FINETTI, BRUNO

(1906–1985)

Bruno de Finetti, an Italian mathematician, was born in Innsbruck, Austria. On the death of his father, the six-year-old de Finetti and his mother moved to Trento (then in Austrian possession). At thirteen he suffered severe osteomyelitis in the left leg; surgery left him permanently lame. In 1923 he entered the Politecnico di Milano to study engineering, his father's and grandfather's profession. In his third year he transferred to the new University of Milan, from which he graduated in 1927 with a degree in applied mathematics. While still an undergraduate he published the first of a series of articles on Mendelian population genetics, developing the first mathematical model with overlapping generations.

From graduation until 1931 de Finetti worked at Rome's Istituto Centrale di Statistica. This was a period of intense and productive research, resulting in publication of a series of mathematical and foundational works on probability. The mathematical works made his name internationally known. The foundational works set out the subjectivist interpretation of probability that he was to advocate all his life. Two stand out: "Sul significato soggettiva della probabilità" (1931) and the remarkable "Probabilismo" (1931), remarkable not least, but certainly not only, for its fascist peroration.

Between 1931 and 1946 de Finetti worked in the actuarial office of the Assicurazioni Generali insurance

company in Trieste. At the same time he taught at the Universities of Padua and Trieste. In this period de Finetti's range widened to include actuarial and financial mathematics, economics, the automation of actuarial procedures (an interest reflected in the postwar years in his advocacy of computing and the use of simulation methods in statistics), and mathematics education. From the early 1950s his works became better known in the English-speaking world, thanks to the advocacy of the American statistician Leonard Savage. In 1947 de Finetti was appointed to the chair of financial mathematics in Trieste. In 1954 he moved to the Faculty of Economics at the University of Rome "La Sapienza"; in 1961 he transferred to the Faculty of Sciences in which he was a professor of the theory of probability until his retirement in 1976. De Finetti died in 1985.

In the 1970s de Finetti was active in Italian politics, standing as a parliamentary candidate for the Radical Party; for a while he edited the party's *Notizie Radicali*. On one occasion a judge ordered his arrest for antimilitarist campaigning.

What de Finetti's life exhibits is a concern for the tying of ideas to applications. The cornerstone of the radical subjectivist interpretation of probability, summed up in de Finetti's claim (in the preface to the English translation of his *Teoria delle probabilità* [1974]), "PROBABILITY DOES NOT EXIST" is that only concepts that can be given an operational, practical significance are meaningful. The radical subjectivist denies the meaningfulness of talk of objective, unknown probabilities. Probability is degree of belief/credence/conviction. De Finetti, as Frank Plumpton Ramsey before him (in work unknown to de Finetti), gave a Dutch book argument to show that a rational person's degrees of belief satisfy the axioms of the probability calculus: degrees of belief are revealed in the betting odds the person considers fair; a rational person does not bet so as to lose money with certainty; fair betting quotients avoid certain loss just if they satisfy the axioms of the probability calculus. Conditional probabilities are handled by conditional bets, bets that are canceled if a given event does not occur. (This led de Finetti to a logic of conditional events: $B|A$ is true if A and B are both true, false if A is true and B is false, and neither if A is false, corresponding to the cases when the bet on B conditional on A is won, lost, and canceled. The idea has resurfaced from time to time in work on the indicative conditional of natural language and on production rules in computer science.)

One axiom is the subject of dispute. In Andrei Nikolajevich Kolmogorov's (1903–1987) *Foundations of the Theory of Probability* (1933) the axiom that probabilities add across a countably infinite partition is adopted as mathematically expedient. De Finetti urged its rejection. Much is known of the consequences of giving up this axiom, but de Finetti's line has not won general acceptance.

Not a philosopher by training, de Finetti found parallels to his thought in the Italian pragmatists Mario Calderoni and Giovanni Vailati (a mathematician), and the man-of-letters Giovanni Papini. Later he saw Humean connections in his influential work on exchangeable and partially exchangeable sequences of events and random variables. A sequence of events of N types is partially exchangeable if the probability that n_1 events of the first type, n_2 events of the second type, … , and n_N events of the Nth type all occur depends only on the numbers $n_1, n_2, \dots, n_N$. For exchangeability, $N = 1$. De Finetti sees this notion as the subjective analogue of (and correction to) talk of independent trials with unknown probability and as making mathematically precise David Hume's account of induction and causation. This comes about through representation theorems. Take the case of an infinite sequence of exchangeable events. From the probability, for various n, that n events all yield favorable outcomes, one can infer the probabilities of r favorable outcomes in n trials, $0 \leq r \leq n$. The distributions of these relative frequencies for different n tend, as n increases, to a limit distribution that functions exactly as a distribution over an unknown probability, so that the probability of any definable event is the expectation with respect to this distribution of the probability it would have were one dealing with a sequence of independent events of constant probability. Exchangeability is preserved as one conditionalizes on the outcomes of any finite number of trials, so, provided the initial limit distribution assigns a nonzero probability to an interval containing it, one obtains a sequence of limit distributions increasingly weighted toward the observed relative frequency as the number of observed instances increases. This encapsulates de Finetti's account of learning from experience and inductive inference, his "translation into logic-mathematical terms of Hume's ideas" (1938, p. 194).

With the acceptance by today's philosophers of science of semantic realism and, increasingly, pluralism in the philosophy of probability, de Finetti's eliminativist reading of what is now called the de Finetti representation theorem is little in favor. But there has been a huge increase in the application both to scientific reasoning generally and to statistics in particular of the subjectivist

interpretation of probability, usually under the name Bayesianism.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Calderoni, Mario; Hume, David; Mathematics, Foundations of; Papini, Giovanni; Probability and Chance; Ramsey, Frank Plumpton; Savage, Leonard; Vailati, Giovanni.

Bibliography

WORKS BY DE FINETTI

"Probabilismo: Saggio critico sulla teoria della probabilità e sul valore della scienza." *Logos* (Naples) 14 (1931): 163–219. Translated by R. C. Jeffrey, M. C. DiMaio, and M. C. Galavotti as "Probabilism: A Critical Essay on the Theory of Probability and on the Value of Science." *Erkenntnis* 31 (2–3) (1989): 169–223.

"Sul significato soggetiva della probabilità." *Fundamenta Mathematicae* 17 (1931): 298–329. Translated as "On the Subjective Meaning of Probability." In *Probabilità e induzione/Induction and Probability*, edited by Paolo Monari and Daniela Cocchi, 291–321 (Bologna, Italy: Editrice Clueb, 1993).

"La logique de la probabilité." In *Actes du congrès international de philosophie scientifique*. Vol. 4, 31–39. Paris: Hermann, 1936. Translated by R. B. Angell as "The Logic of Probability." *Philosophical Studies* 77 (1995): 181–190.

"La prévision: ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincaré* 7 (1937): 1–68. Translated by Henry E. Kyburg Jr. as "Foresight: Its Logical Laws, Its Subjective Sources." In *Studies in Subjective Probability*. 2nd ed., edited by Henry E. Kyburg Jr. and Howard E. Smokler, 57–118. Huntington, NY: Robert E. Krieger, 1980.

"Sur la condition d'équivalence partielle." *Actualités scientifiques et industrieles No. 739* (*Coloque Genève d'Octobre 1937 sur la Théorie des Probabilités, 6ième partie*), 5–18. Paris: Hermann, 1938. Translated by P. Benacerraf and R. C. Jeffrey as "On the Condition of Partial Exchangeability." In *Studies in Inductive Logic and Probability*, Vol. 2., edited by Richard C. Jeffrey, 193–205. Berkeley: University of California Press, 1980.

Teoria delle probabilità: Sintesi introdutiva con appendice critica. 2 vols. Turin, Italy: Einaudi, 1970. Translated by Antonio Machí and Adrian Smith as *Theory of Probability: A Critical Introductory Treatment*. 2 vols. New York: Wiley, 1974–1975.

Probability, Induction, and Statistics: The Art of Guessing. New York: Wiley, 1972.

"Probability and My Life." In *The Making of Statisticians*, edited by J. Gani, 3–12. New York: Springer-Verlag, 1982.

Probabilità e induzione/Induction and Probability, edited by P. Monari and D. Cocchi. Bologna, Italy: Editrice CLUEB, 1993.

Filosofia dela probabilità, edited by A. Mura. Milan, Italy: Il Saggiatore, 1995.

WORKS ABOUT DE FINETTI

Cifarelli, Donato Michele, and Eugenio Regazzini. "De Finetti's Contribution to Probability and Statistics." *Statistical Science* 11 (4) (1996): 253–282.

Diaconis, Persi, and David Freedman. "De Finetti's Generalizations of Exchangeability." In *Studies in Inductive Logic and Probability*. Vol. 2, edited by Richard C. Jeffrey, 233–249. Berkeley: University of California Press, 1980.

Hintikka, Jaakko. "Unknown Probabilities, Bayesianism, and de Finetti's Representation Theorem." In PSA 1970 *In Memory of Rudolf Carnap*, edited by Roger C. Buck and Robert S. Cohen, 325–341. Dordrecht, Netherlands: D. Reidel, 1971.

Howson, Colin, and Peter Urbach. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court, 1993.

Jeffrey, Richard. *Subjective Probability: The Real Thing*. New York: Cambridge University Press, 2004.

Jeffrey, Richard C., and Maria Carla Galavotti, eds. "Bruno de Finetti's Philosophy of Probability." *Erkenntnis* 31 (2–3) (1989).

Milne, Peter. "Bruno de Finetti and the Logic of Conditional Events." *British Journal for the Philosophy of Science* 48 (1997): 195–232.

Peter Milne (2005)

DEFINITION

The problems of definition are constantly recurring in philosophical discussion, although there is a widespread tendency to assume that they have been solved. Practically every book on logic has a section on definition in which rules are set down and exercises prescribed for applying the rules, as if the problems were all settled. And yet, paradoxically, no problems of knowledge are less settled than those of definition, and no subject is more in need of a fresh approach. Definition plays a crucial role in every field of inquiry, yet there are few if any philosophical questions about definition (what sort of thing it is, what standards it should satisfy, what kind of knowledge, if any, it conveys) on which logicians and philosophers agree. In view of the importance of the topic and the scope of the disagreement concerning it, an extensive reexamination is justified. In carrying out this conceptual reexamination, this article will summarize the main views of definition that have been advanced, indicate why none of these views does full justice to its subject, and then attempt to show how the partial insights of each might be combined in a new approach.

All the views of definition that have been proposed can be subsumed under three general types of positions, with, needless to say, many different varieties within each type. These three general positions will be called "essentialist," "prescriptive," and "linguistic" types, abbreviated as "E-type," "P-type," and "L-type," respectively. This clas-

sification is not intended as a precise historical summary, but merely as a useful schema for stating some of the problems and disputes. Thus, some outstanding philosophers may very clearly belong to one of these types. Others who, for the purposes of this article, are placed in a certain class hold positions varying considerably from the presentation to be given. It must therefore be borne in mind that not all the criticisms that will be made apply to all philosophers included in the class being criticized. Writers whose accounts of definition fall largely under the E-type include Plato, Aristotle, Immanuel Kant, and Edmund Husserl. Those who support P-type views include Blaise Pascal, Thomas Hobbes, Bertrand Russell, W. V. Quine, Nelson Goodman, Rudolf Carnap, C. G. Hempel, and most contemporary logicians. Supporters of L-type views include John Stuart Mill (in part), G. E. Moore (in part), Richard Robinson, and most members of the school of linguistic analysis.

According to essentialist views, definitions convey more exact and certain information than is conveyed by descriptive statements. Such information is acquired by an infallible mode of cognition variously called "intellectual vision," "intuition," "reflection," or "conceptual analysis." Prescriptive views agree with essentialism that definitions are incorrigible, but account for their infallibility by denying that they communicate information and by explaining them as symbolic conventions. Although linguistic views agree with essentialism that definitions communicate information, they also agree with prescriptivism in that they reject claims that definitions communicate information that is indubitable. The linguistic position is that definitions are empirical (and therefore corrigible) reports of linguistic behavior.

ESSENTIALISM

An essentialist account was first proposed by Socrates and Plato. Socrates is renowned for having brought attention to the importance of definitions. His favorite type of question, "What does (virtue, justice, etc.) mean?," became the characteristic starting point of philosophical inquiry. But Socrates did not make clear what kind of answer he was looking for. In Plato's *Euthyphro* Socrates is reported to have said that the kind of answer he expected to his question "What is piety?" was one giving an explanation of "the general idea which makes all pious things to be pious" and "a standard to which I may look and by which I may measure actions." He did not explain, however, what he meant by "idea" and "standard" nor how one produces an "idea" or a "standard" when one is defining a term. Richard Robinson, in his book *Plato's Earlier Dialectic* (p. 62), has suggested that the question "What is *X?*" is more ambiguous than Socrates realized and that it may be answered in all sorts of ways, depending on the context in which it is asked.

PLATO

PLATO. Plato's attempts in his later dialogues to explain the meaning of the Socratic question "What is *X?*" constitute the celebrated Theory of Forms, the trademark of Platonic metaphysics and epistemology. In a passage of central importance (*Republic* VI), Plato distinguished two kinds of objects of knowledge (sensible things and forms) and two modes of knowledge (sense perception and intellectual vision). Sensible things are objects of opinion, while abstract forms are objects of philosophical knowledge. Physical objects, shadows, and images are imperfect and ephemeral copies of forms; our perceptual knowledge of them is an inaccurate approximation to our knowledge of their abstract archetypes. Definitions describe forms, and since forms are perfect and unchanging, definitions, when arrived at by the proper procedure, are precise and rigorously certain truths. Empirical statements describe objects of perception and are therefore only more or less reliable approximations to truth.

Models and copies. Plato's analogy between definitions and empirical descriptions—an analogy upon which all E-type theories of definition rest—is supplemented by a second analogy between the relation of a model to a copy and the relation of a definition to an individual predication. This analogy was suggested by Socrates when he asked for "a standard to which I may look and by which I may measure actions." Plato describes the process of coming to know as if it were like the procedure of a craftsman producing a piece of sculpture or a house. The sculpture is a "copy" of the subject who models for it; the house is in one sense a "copy" of the architect's blueprint, in a somewhat different sense a "copy" of a small-scale model, and in still a third sense a "copy" of the idea in the mind of the builder. Plato's frequent references to the arts and crafts in his exploration of conceptual problems indicate that the analogy of the model-copy relation plays a central role in his theory of knowledge.

Thus, Platonic essentialism provides two sets of answers (both of which rest on metaphors) to the questions "What kind of statements are definitions?" "What purpose do they serve?" and "How are they to be judged as good or bad?" It suggests primarily that definitions are descriptions of objects that are somehow analogous to tables, chairs, and other familiar things; that these definitions serve the purpose of providing descriptive informa-

tion about their objects; and that they are confirmed by a mode of cognition somehow analogous to sense perception, yet independent of the sensory organs. Secondarily, Platonic essentialism specifies the relation between the objects of definitions and those of empirical descriptions by characterizing the former as models of which the latter are "copies."

Adequacy of the model metaphor. Metaphors are apt or inapt, illuminating or misleading, according to two criteria: (1) the number and importance of the known points of resemblance between the things compared; and (2) the number and importance of previously unnoted facts suggested by the metaphor. To what extent does Plato's metaphor of the unseen model satisfy these criteria?

The primary term of comparison in Plato's metaphor is the abstract form or universal that a definition allegedly describes. The secondary term is the model for a painting or, alternatively, a tailor's pattern. As the painter looks to his model and the tailor to his pattern, the philosopher can look to the forms for the specifications that identify things as instances of one class rather than another, as well as for exact information about the properties of that class.

What are the known points of resemblance between forms and models, on which this metaphor is grounded? Merely to ask this question is already to see that the metaphor is defective from the start, since there cannot possibly be *any* literal points of comparison. The Platonic forms, unlike models and patterns, have no observable properties by virtue of which they can be said to "resemble" anything at all. Thus, if the model metaphor has any value, it must lie entirely in what the metaphor suggests, rather than in its literal grounds.

Primarily, the model metaphor suggests that definitions and their corollaries constitute all there is to knowledge. Whenever a question of fact or of judgment is raised in the Platonic dialogues, it is treated as a problem of definition. For example, when, in the *Euthyphro,* Socrates and Euthyphro argue about the propriety of a son's prosecuting his father for murder, Socrates proceeds as though the issue could be settled by arriving at a clear definition of piety—as though one could then look at the definition, look at the action, and decide whether they coincide. We can identify a portrait or a garment by comparing it with its model or its pattern, but we cannot classify and judge an action in the same way. Description and evaluation are seldom matters of identification by comparison with a pattern. In this respect Plato's essentialism is misleading rather than illuminating.

The metaphor of the unseen model also suggests that definitions provide us with precise and rigorous knowledge in the way that blueprints make possible a high degree of uniformity and precision in productive arts such as architecture. But definitions increase precision only when they change the original meanings of words for technical purposes. Generally speaking, a definition can be no more precise than the concept it defines, at the risk of shifting to a different concept. Our concept of what constitutes an adult is vague; if we try to make it precise by specifying an exact age at which childhood is divided from adulthood, we merely lose sight of what we started out to talk about by replacing the concept of maturity with that of having passed a certain birthday.

The model metaphor is not entirely misleading; it suggests at least one genuine resemblance between the terms it compares. The relation between definitions and empirical descriptions is, in one respect, rather similar to the relation between portraits and their models. We judge a portrait (to some extent) by noting whether the portrait looks like the model; we verify the empirical description "This table is round" by looking at the table to see whether it has the properties definitive of tables and of roundness. But if we are asked, "Is that person a good model?" or "Is that definition a good definition?" we cannot look toward anything of which the model is himself a portrait, and we cannot look at a definitional form of which the particular definition is itself an instance. Definitions are not evaluated in the same way as empirical descriptions, just as models are not judged in the same way as their portraits. Thus the analogy between definitions and empirical descriptions from which Platonic essentialism starts eventually contradicts itself.

ARISTOTLE. One can find in Aristotle's works anticipations of every later theory of definition, but he gave high priority to his own brand of essentialism, whereby he explained the nature of "real" as distinguished from "nominal" (that is, prescriptive or linguistic) definition. Like Plato, Aristotle stressed the similarity between definitions and statements of fact, and he assumed that definitions convey precise and certain information. But Aristotle employed a different supporting metaphor to explain the special nature of definitions. The most noteworthy feature of his many discussions of definition is his insistence that a real definition should provide a causal explanation of the thing defined. In the *Physics,* Aristotle distinguished four types of causes—formal, material, final, and efficient. He characterized the first three types as "internal," while efficient causes are (usually) "external" to their effects. Internal causes are not available to

public inspection, but must be discovered in abstract intuition. The causal explanation provided by a real definition is in terms of one or more of these three internal types of cause.

Definition and causality. It is not easy to explain just what Aristotle meant by "internal cause." Part of what he seems to have meant is that, unlike "incidental" causes, internal causes are necessary for their effects. But it is by no means clear what sense of necessity is involved in this instance. To explain this necessity as causal would be a case of circular reasoning. On the other hand, to say that the necessity is logical seems only another way of saying that the effect is definable in terms of its cause, which is again circular reasoning. As an example of a causal definition, Aristotle defined a lunar eclipse as the privation of the moon's light because of the interposition of the earth between the moon and the sun (*Posterior Analytics* 90a). This example confirms the suggestion that for something to be an internal causal is for it to be part of a definition. But the difficulty then arises that definition has been explained by internal causality, internal causality by necessity, and necessity by definition. Thus, Aristotle's eclipse leaves us in the dark about definition.

Classification and explanation. The trouble is that the idea of internal causality is a metaphor. An essential cause is not "internal" to the thing defined as a kernel is inside a nut, but only metaphorically "inside."

This metaphor suggests two important but dubious principles: that scientific knowledge consists entirely of definitions and their corollaries and that systematic classification is identical with theoretical explanation. If to define a term is, at the same time, to provide a causal explanation of what it denotes and if the classification of a thing in terms of its species and differentia is sufficient for deducing the laws of its behavior, then the work of scientific inquiry is completed when a comprehensive system of classification has been constructed. Thus, Aristotle wrote in the *Posterior Analytics* (90b) that "Scientific knowledge is judgment about things that are universal and necessary, that the conclusions of demonstration *and all scientific knowledge* follow from the first principles" and that "the first principles of demonstration are definitions" (italics added).

That scientific knowledge is not entirely derivable from a set of definitions and that systematic classification is only one small aspect of scientific procedure need hardly be argued. Aristotelian concepts of causality and explanation have been almost completely expunged from modern science, and causes are conceived of in quite different ways. But it is not the archaic character of Aristotle's use of "cause" and "explanation" that concerns us here. It is largely a matter of terminological convenience whether we continue to use these words in the Aristotelian manner or confine them to the procedures of modern physical science. In regard to the problem of clarifying the functions and criteria of definitions, however, Aristotle's claim that definitions reveal the internal causes of their definienda must be criticized not as a false, but as a misleading, metaphor, for it dissolves the very distinctions which it is intended to explain—namely, the distinction between definitions and empirical statements of fact, that between the method of evaluating definitions and the method of confirming factual hypotheses, and that between the distinctive functions of definition and the general aims of scientific inquiry.

IDEAS AND CONCEPTS. A third metaphor that has been employed in the support of E-type views of definition originated in Cartesian dualism. René Descartes himself leaned toward a prescriptive account of definition, which will be considered later. But John Locke, Kant, Husserl, and other philosophers who accepted the Cartesian division between the "inner world" of the mind and the "outer world" of physical events took the essentialist position that philosophical inquiry should provide information about a special set of objects ("ideas" for Locke, David Hume, and Husserl; "concepts" for Kant, Heinrich Rickert, and G. E. Moore) discoverable by an infallible mode of cognition ("reflection" for Locke and Husserl; "analysis" for Hume, Kant, Rickert, and Moore).

According to Locke, the outer world of material objects and their motions is describable by the laws of physics, while the inner world of ideas is describable by the laws of psychology that are discovered by reflection on the contents of the mind. These contents are simple and complex ideas; the task of philosophy is to analyze complex ideas into their simple elements and to describe their mode of combination.

Kant distinguished between "analytical" and "synthetic" definitions, regarding the former as the identification of the simple elements (predicates) out of which concepts are formed by the understanding and the latter as the formation of rules of serial order that provide the synthetic a priori postulates of mathematics and physics.

The philosophers under consideration, like their predecessors, assumed that definitions convey knowledge of objects (ideas, images, essences, concepts, or meanings) whose special nature guarantees precision and certainty and that this remarkable kind of knowledge is acquired through a special mode of cognition (reflection,

introspection, intuition, or conceptual analysis). The literal content of the private-world metaphor thus seems to be identical with that of the essentialist metaphors already considered. The differences between the private-world and essentialist metaphors (other than terminological ones) must be sought in the suggestive implications of the metaphor. But there is an important difference between philosophers such as Locke, Hume, and Husserl, who reserve the word *definition* for conventions of word usage and do not consider their introspective analyses of ideas to be definitions, and those such as Kant, Rickert, and C. I. Lewis, who regard philosophical definitions as products of conceptual analysis.

Both groups employ the Aristotelian distinction between real and nominal definitions, except that members of the first group avoid calling the results of their introspective studies "definitions" because they think of them as descriptions of the workings of the mind analogous to descriptions of a clock that has been taken apart for inspection. They think of the special mode of cognition by means of which they discover how simple ideas are organized into complex ideas as inner vision or grasp, which is analogous to sight and touch. But members of the more abstractly minded group compare the special faculty by which real (or analytic or explicative) definitions are discovered to the experience (familiar to logicians and mathematicians) of recognizing logical relation, rather than comparing it to any type of sense perception. They speak of "understanding the meanings of words," of "logical analysis," of "understanding what is contained in a concept," rather than of seeing or grasping the "contents of the mind." There are, then, two kinds of world imagined by these theorists: a world of privately visible or tangible ideas, sense data, secondary qualities, and so forth and a world of abstract concepts or meanings. Some, like Kant, Husserl, and, most systematically, C. I. Lewis, posit both kinds of worlds.

What then do these two metaphors suggest, and how illuminating are their implications? The metaphor of the private world of sense data that is allegedly described by definitions of complex ideas suggests that such definitions, like reports of hallucinations, dreams, and other private experiences, must be taken at face value (provided that they are sincerely and consistently expressed), since they cannot be checked by public observation. This would account for the unchallengeable character of definitions and their analytic corollaries, in contrast to the corrigibility of empirical statements. But this view deprives definitions of any claim to objective validity and entails that every person has a right to his own definitions, in the same way that everyone has a right to his own dreams.

The metaphor of the world of concepts and meanings also attributes a self-certifying character to definitions but fares better with respect to the commonsense fact that we balk at some definitions and accept others—for the recognition of logical relations, no matter how intuitive, is a socially shared experience. We immediately and privately understand, see, or grasp that a statement of the form $P \cdot Q$ implies a statement of the form Q, but we can also argue the fact and summon evidence (in the form of postulates of a logical system) to prove it. But this metaphor, which of all those we have considered comes closest to not being a metaphor at all and blends imperceptibly into a prescriptive concept of definition, suggests both too much and too little. It suggests that definitions are logical truths and possess logical certainty. But although some definitions are worse than others, all logical truths are normatively equal. Moreover, the metaphor fails to indicate how definitions can be evaluated other than by their formal consistency (the standard by which we confirm a system of logical truths). Yet a definition of a cow as a three-legged animal would be universally rejected on grounds having nothing to do with inconsistency. The denial of a logical truth can be shown to involve a contradiction, but the denial of a definition leads to contradiction only if one has already accepted the definition. Although consistency is a sufficient condition for a system of logical truths, it is merely a necessary condition for sound definitions; yet no additional conditions are provided by logistic phenomenalism.

PRESCRIPTIVISM

E-type views claim that definitions are statements and that they make assertions that can be pronounced true or false. Essentialists, however, have difficulty explaining how and why definitions differ from ordinary statements of fact, and hence they fall back on metaphors. P-type theories avoid this trouble by denying that definitions are statements of any kind. The prescriptivist assimilates definitions to imperative sentences rather than to declarative sentences and endows them with the function of syntactic or semantic rules for prescribing linguistic operations.

There are two main varieties of prescriptivism. The nominalist variety explains definitions as semantic rules for assigning names to objects, while the formalist variety regards definitions as syntactic rules for abbreviating strings of symbols. P-type views of definition can be traced back to the Greek Sophists and Skeptics, but this article will concentrate on the modern sources of these

views. The rebirth of science in the seventeenth century was accompanied by a sweeping rejection of medieval thought, in particular the medieval concept of definition as the penetration by metaphysical intuition into a realm of changeless forms. The nominalist theories of language employed by Sophist and Cynic contemporaries of Plato to undermine belief in the objectivity of knowledge, and again by the more radical medieval Scholastics to subvert the control of theology over science, became, in the seventeenth century, a cornerstone of the reconstruction of knowledge on a new scientific foundation.

Seventeenth-century writings on definition are not entirely free of the influence of classical essentialism. Seventeenth-century prescriptive theories of definition try to avoid the obscurities of essentialism by repudiating the informative role of definitions, but they cannot provide adequate criteria for distinguishing good definitions from bad without presupposing some sort of informative role for them.

NOMINALISM. For Francis Bacon and Hobbes, definitions possessed a therapeutic function, as a means of clearing up or avoiding ambiguous, vague, and obscure language. Regarding semantic confusion as the main source of intellectual trouble, they proposed to clear the way for a new system of knowledge by subjecting existing concepts to the test of definitional reduction to observable and measurable properties. Definition was thus a surgical knife for cutting away metaphysical encrustations, as described by Bacon in paragraph 59 of the *Novum Organum*:

> But the idols of the market-place are the most troublesome of all: idols which have crept into the understanding through the alliances of word and names, and this it is that has rendered philosophy and the sciences sophistical and inactive. Whence it comes to pass that the high and formal discussions of learned men end oftentimes in disputes about words and names: with which it would be more prudent to begin, and so by means of definitions reduce them to order.

Thomas Hobbes also stressed the clarifying role of definitions, taking geometry as his model. In the *Leviathan* he wrote:

> Seeing then that truth consists in the right ordering of names in our affirmations, a man that seeketh precise truth had need to remember what every name he useth stands for … or else he will find himself entangled in ffords as a bird in lime twigs. And therefore in geometry, which is the only science which it hath pleased God hitherto to bestow on mankind, men begin at settling the significations of their words: which settling of significations they call *definitions,* and place them in the beginning of their reckoning.

Definitions thus clear up ambiguities and "settle significations," rather than communicate information about a realm of essences. They are introduced at the beginning of inquiry, as in geometry, rather than at the culmination of inquiry, as in metaphysics and Aristotelian natural science.

According to Hobbes, all knowledge consists in the "right ordering of names in affirmation." A proposition connects one name to another, and an inference adds or subtracts one proposition from another. The structure of scientific thought thus maps the structure of the physical world. It would seem then that, for Hobbes, all scientific knowledge is derivable from definitions. Yet Hobbes also stressed the role of perception in knowledge. The solution to this paradox lies in Hobbes's conception of naming. All inquiry is deductive except for the assignment of names to things, and it is to the assignment of names that we must look for the empirical sources of knowledge. But it follows that definitions as assignments of names must be as informative for Hobbes as they are for Plato or Aristotle. This conclusion leads to a further paradox, for, according to Hobbes, definitions provide no information at all; they express conventional decisions to use particular signs as names of particular objects.

There is an ambiguity in Hobbes's account of definitions that must hamper any attempt to reduce definitions to assignments of names. In order to make definitions entail all the propositions of scientific knowledge, Hobbes had to include, in the notion of naming, all the cognitive functions that we ordinarily distinguish from naming. He first compared the highly abstract and sophisticated definitions of concepts in mathematics and natural science to simple naming procedures such as baptism. Then, in order to account for the conspicuous differences between the two kinds of procedures, he was compelled to reinject into the notion of naming the very distinctions he set out to eliminate. The reduction of definitions to assignments of names only *appears* to solve the problem of whether definitions are informative: It first suggests that definitions are as arbitrary as acts of naming and then suggests that naming is, after all, not always arbitrary.

EARLY FORMALISM. Although the language used by the Cartesians of the seventeenth century in discussing defi-

nitions was similar to that of Bacon and Hobbes, their emphasis and direction of interest was different. Bacon and Hobbes were primarily concerned with the role of definitions in achieving semantic clarity, the Cartesians were more interested in the role of definitions in deductive inference. They developed a conception of definitions as theoretically dispensable abbreviations whose value lies solely in the notational economy they make possible. Cartesian references to "names" are rather misleading since, unlike Hobbes, the Cartesians did not regard assignment of names as the initial and fundamental process of inquiry from which the rest of knowledge is derived. This role was taken over by axioms and postulates that relate "simple" (i.e., indefinable) terms to each other, definitions then being introduced as rules for substituting brief expressions for logical complexes of simple terms.

Descartes did not give much attention to the subject of definition. In rejecting classical syllogistic logic as the framework of scientific inference, he abandoned the emphasis on terms or classes as the basic units of inference in favor of propositional units. The simplest inference became, for Descartes, the intuitive recognition of the implication of one proposition by another. Consequently, postulates replaced definitions as the foundation of deductive science, and essential definitions ceased to represent the highest goal of knowledge.

Pascal's analysis of the nature and function of definitions made explicit the view of definition implicit in Descartes's theory of knowledge. The main elements of Pascal's discussion are formalistic. However, it is not free of ambiguity with respect to the purely notational role of definitions as against the informative role ascribed to them by essentialists.

Pascal's theory of definition is expounded in a brief essay, *De l'esprit géométrique* (*Oeuvres,* 14 vols., Léon Brunschvicg and E. Boutroux, eds., Paris, 1904–1914). He began by distinguishing two types of definition, *définitions de nom,* which he claimed to be the only type appropriate in science, and an unnamed type that seems to be what Aristotle called "real," the type favored by essentialists, about which he thereafter says nothing more.

Définitions de nom are said to be "mere impositions of names upon things that have been clearly indicated in perfectly intelligible terms," as, for example, the definition of "even number" as "number that can be divided by two without remainder." Such definitions, Pascal claimed, are conventional labels that need have nothing in common with the things they name. They communicate no information about their *nominata,* expressing merely the decision of the writer to use them in the prescribed manner. The sole limitation on *définitions de nom* is that they be internally and mutually consistent.

When he discussed the methodology of definition, Pascal no longer regarded the relation between language and reality as purely conventional. We must make sure "not to define things that are clear and are understood by everyone." Geometry provides the model for definitional procedure. "It does not define such things as *space, time, motion, number, equality* … because these terms so naturally designate the things to which they refer, for those who understand the language, that the intended clarification would be more likely to obscure them than to instruct." One might think that, in saying "space naturally designates" its referent, Pascal meant that the word *space* is so familiar that everyone understands what it signifies. But why, then, should he interdict any definition of "space"? If definitions are notational conventions, there could be no objection to stipulating a new use of the word. Indeed, the ordinary use of "space" is quite different from its technical use in mathematics. Why, then, is it improper to define either the ordinary or the mathematical use? Surely, Pascal was not thinking of the word *space,* but of space itself as an irreducible entity that cannot be analyzed into simpler components, and if so, then he was thinking of definition not as a notational convenience, but as an informative mode of analysis.

The Cartesian theory of knowledge by which Pascal was guided conceives of the world as a system of elements combined according to mathematical laws to form complex objects and events. While Descartes stressed the analytical reduction of complex propositions to simple ones (i.e., axioms), Pascal joined definitions to axioms as the basis from which the deductive reconstruction of science should start. But common to all the Cartesians is the assumption that knowledge is a mathematical mapping of the structure of nature. In the light of this epistemological atomism, the conventional character attributed to definitions contrasts sharply with the requirement that they correspond to an antecedent natural order—a requirement that leads back to essentialism.

MODERN FORMALISM. The formalistic conception of definitions as rules of notational abbreviation was only vaguely anticipated by seventeenth-century philosophers, who failed to separate this purely syntactic procedure from epistemological considerations such as mapping the order of nature. Only in recent times have formalistic discussions of definition been purified of epistemological assumptions, by (among others) Russell, Alfred North

Whitehead, W. V. Quine, Rudolf Carnap, C. G. Hempel, and Nelson Goodman. But it remains doubtful whether this purely formalistic view either is or can be consistently maintained.

Russell and Whitehead, in *Principia Mathematica* (Vol. I, p. 11), define a definition as follows:

> A definition is a declaration that a certain newly introduced symbol or combination of symbols is to mean the same as a certain other combination of symbols of which the meaning is already known. … It is to be observed that a definition is, strictly speaking, no part of the subject in which it occurs. For a definition is concerned wholly with the symbols, not with what they symbolize. Moreover, it is not true or false, being the expression of a volition, not of a proposition.

This characterization of definition is not consistently syntactical. It defines *definition* in terms of sameness of meaning, while claiming that a definition "is concerned wholly with the symbols, not with what they symbolize." Later in the same passage, Russell and Whitehead declare:

> In spite of the fact that definitions are theoretically superfluous, it is nevertheless true that they often convey more important information than is contained in the propositions in which they are used. This arises from two causes. First, a definition usually implies that the *definiens* is worthy of careful consideration. … Secondly, when what is defined is … something already familiar …, the definition contains an analysis of a common idea. (p. 12)

The first and last sentence in the passage above express a nonsyntactical attitude toward definitions. Definitions turn out to be highly informative, and we seem to have returned to an essentialist view of the matter. But a further qualification has been attached, namely, "when what is defined is … something already familiar." In fact, two types of definition are being considered, one being a rule of notational abbreviation and the other an "analysis of an idea." But if some definitions are "analyses of ideas" and are highly informative, then these are the important kinds of definitions, and the formalist view proclaimed at the outset loses its force.

Similar difficulties attend the efforts of other modern logicians to deal with the problem of definition from a purely formal point of view. Thus, W. V. Quine, after asserting that "a definition is a convention of notational abbreviation," qualified his statement as follows:

> Although signs introduced by definition are formally arbitrary, more than such arbitrary notational convention is involved in questions of definability; otherwise any expression might be said to be definable on the basis of any expressions whatever. … To be satisfactory … a definition … not only must fulfill the formal requirement of unambiguous eliminability, but must also conform to the traditional usage in question. ("Truth by Convention," in *Readings in Philosophical Analysis,* edited by H. Feigl and W. Sellars, New York, 1949, p. 252)

Nelson Goodman took the same position and fell into the same difficulties:

> In a constructional system … most of the definitions are introduced for explanatory purposes. … In a formal system considered apart from its interpretation, any such definitional formula has the formal status of a convention of notational interchangeability once it is adopted; but the terms employed are ordinarily selected according to their usage, and the correctness of the interpreted definition is legitimately testable by examination of that usage. (*The Structure of Appearance,* p. 3)

In common with many other logicians, Quine and Goodman distinguish between the function of definitions "in a formal system" and their function when the system is interpreted—that is, when definite meanings are assigned to the symbols of the system. But this distinction overlooks the fact that from a purely formal standpoint, there is no such thing as a definition at all. Before it is interpreted, the formula that we interpret as a definition is just a string of marks. From a "purely formal standpoint," not only is there no difference between a definition and a notational abbreviation, but there is no difference between a definition and *any* other kind of formula. There are only various strings of marks, some permitted by the rules of formation of the system, others excluded by these rules. Consequently, the distinction made by Quine and Goodman between definitions in a formal system and those in an interpreted system is seriously misleading.

Rudolf Carnap and C. G. Hempel have tried to clarify the difference between informative definitions and mere notational abbreviations by distinguishing between "old" and "new" concepts. Definitions of old concepts are called "explications" by Carnap and "rational reconstructions" by Hempel, while both call definitions of new concepts "notational conventions." When we are

"explicating" or "reconstructing" a concept, our definitions are subject to evaluation by the criteria of conformity to usage and increase of precision (Rudolf Carnap, *The Logical Syntax of Language,* p. 23). When definitions are introduced solely for the purpose of abbreviation, only the criterion of consistency applies. One must therefore wonder why Carnap and Hempel should bother to call notational abbreviations "definitions," since they have nothing whatever in common with explications.

Perhaps the answer to this question lies in the logical difficulties lurking within the notion of explication. What does it mean to "reconstruct" or "explicate" a concept, and what precisely is the difference between "old" and "new" concepts? If definitions of old concepts must conform to established usage, are they not true or false statements about language usage, in which case the distinction between definitions and empirical statements disappears? These problems lead naturally into the linguistic theory of definition.

LINGUISTIC THEORIES

Anticipations of a linguistic view of definition may be found in classical writings (for example, in Aristotle's discussion of "nominal definition") and in the nominalist and formalist positions previously considered. But while early nominalism attempted to reduce all the varied functions of words to that of proper names and thus to reduce meaning to the arbitrary assignment of a name to an object, formalism added linguistic considerations as an inessential afterthought. The first step from nominalism to an L-type view proper was taken by John Stuart Mill, although his formulations are permeated with elements of both nominalism and essentialism. A further step was taken by G. E. Moore, but Moore's discussion also contains a heavy strain of essentialism. The clearest formulation of the linguistic view was provided by Richard Robinson in his book *Definition,* which has the distinction of being the only book in the English language devoted to this subject.

In his *System of Logic,* J. S. Mill defined "definition" as follows: "The simplest and most correct notion of a Definition is, a proposition declaratory of the meaning of a word: namely, either the meaning which it bears in common acceptation, or that which the speaker or writer … intends to annex to it" (10th ed., p. 86).

Mill then explained that a definition is a "verbal proposition" that "adds no information to that which was already possessed by all who understood the name (defined)"—a tautology that Mill mistook for an important observation. But, unlike the thoroughgoing prescriptivist, Mill did not regard definitions as purely conventional stipulations, at least insofar as terms in general use are concerned:

> It would, however, be a complete misunderstanding of the proper office of the logician in dealing with terms already in use, if we were to think that because a name has not at present an ascertained connotation, it is competent to anyone to give it such a connotation at his own choice. The meaning of a term actually in use is not an arbitrary quantity to be fixed, but an unknown quantity to be sought. (p. 91)

At this point, Mill conceded that some definitions are not mere "declarations" but convey some kind of information about "unknown quantities to be sought." Mill gave two reasons for this departure from prescriptivism. The first consideration involves him in a tug of war between nominalist and linguistic theories. "Since names and their significations are entirely arbitrary, such (verbal) propositions are not, strictly speaking, susceptible of truth or falsity, but only of conformity or disconformity to usage or convention; and all the proof they are capable of is proof of usage" (p. 92).

In this instance, Mill first denied and then asserted that definitions are informative. If "all the proof they are capable of is proof of usage," then they *are* capable of proof after all, despite his initial disclaimer of this possibility.

Mill's second reason for ascribing at least a quasi-informative function to some definitions resembles, to some extent, the phenomenalist conception of definition as analysis of complex ideas into simple constituents. Mill wrote:

> A name, whether concrete or abstract, admits of definition, provided we are able to analyze, that is, to distinguish into parts, the attribute or set of attributes which constitutes the meaning both of the name and of the corresponding abstract. … We thus see that to frame a good definition of a name in use is not a matter of choice but of discussion … not merely respecting the usage of language, but respecting the properties of things, and even the origin of these properties. (p. 91)

The source of Mill's shifts of emphasis and inconsistencies lies in the ambiguity of his notion of meaning. At times he identified the meaning of a term with the object it "names," at other times with the customary usage of the word, and at still other times with an abstract object or "idea" capable of being divided into simpler parts. Thus,

depending on which conception of meaning he had in mind, he thought of a definition as the stipulation of a name, a report of linguistic usage, or the analysis of a complex idea into its constituent parts.

G. E. MOORE. The extent to which G. E. Moore's approach to definitions can properly be called "linguistic" is debatable. Moore placed less stress on the linguistic aspect of definition than later philosophers such as Gilbert Ryle, Peter Frederick Strawson, and Robinson, who were influenced by Moore's analytical method. For Moore, as for Socrates, the clarification of language was only a means toward the discovery of deeper philosophical truths. But there can be no doubt that Moore inspired others to concern themselves with language and that his painstaking attention to the nuances of words was the most distinctive feature of his work.

In his *Principia Ethica,* Moore characterized "analytical" definitions (the kind produced by philosophical analysis) as follows: "Definitions of the kind that I was asking for, definitions which describe the real nature of the object or notion denoted by a word and which do not merely tell us what the word is used to mean, are only possible when the object or notion is complex" (p. 7).

In order to indicate the kind of descriptive information that he expected philosophical definitions to provide, Moore offered an example that is as misleading as it is famous: "When we say … 'The definition of horse is "a hoofed quadruped of the genus Equus"' … we may mean that a certain object, which we all of us know, is composed in a certain manner: that it has four legs, a head, a heart, a liver, etc., all of them arranged in definite relations to one another" (p. 8).

This passage is curious; it suggests that an analytical definition lists the physical parts of the thing defined. The example, however, gives the species and differentia of the class of horses but does *not* mention any physical parts. In commenting on this passage in his *Reunion in Philosophy* (p. 184), Morton White has observed that Moore shifted inadvertently from logical to physical complexity.

In later writings, Moore maintained that concepts are the proper subject matter of definition. "To define a concept," he wrote, "is the same thing as to give an analysis of it" ("Reply to My Critics," in *The Philosophy of G. E. Moore,* pp. 664–665). It is not easy to tell just what Moore meant by "concept analysis." For the analysis of a concept, he offered three criteria that add up to the relation of synonymity of expressions. Thus, despite his explicit effort to find an informative function for definitions that goes beyond the explanation of how words are used, it is not unreasonable to conclude that all that his obscure notion of "analyzing a concept" finally comes to is linguistic clarification. In denying that analytic definitions "merely tell us what the word is used to mean," Moore was rejecting the view that definitions are generalizations about common usage and suggesting that they have a more explanatory function. But he never made clear what that function is.

In the only full-length volume in English devoted to the study of definition, Richard Robinson formulated a purely linguistic account of definitions as reports of word usage. But he thought it necessary to supplement his main view with a "stipulative," or prescriptive, account. The reasons for his vacillation are that reports of usage are empirical generalizations, while definitions are, if acceptable at all, necessary truths, and that stipulations are uninformative, while definitions are highly informative. Thus, neither the linguistic nor the prescriptive interpretation accounts for all features of definitions. But the mere juxtaposition of the two can hardly overcome the defects of each taken separately.

A PRAGMATIC-CONTEXTUAL APPROACH

Linguistic theories of definition brought needed attention to the close relation between definitions and the meanings of words, but they erred in identifying meanings either with objects or concepts allegedly denoted by words or with linguistic usage. A correct theory of definition would unite the partial insights of E-type, P-type, and L-type views without relying on misleading metaphors, denying the obvious informative value of definitions, or reducing definitions to historical reports of linguistic behavior.

Why should essentialists and linguistic philosophers claim that definitions convey knowledge, while prescriptivists deny that they do? In some sense of the word *knowledge,* anyone would agree that definitions communicate knowledge. The problem is to identify a special sense of "knowledge" that is appropriate to definitions but does not require us to postulate obscure essences or to reduce definitions to historical reports. This special kind of knowledge may be knowledge of how to use words effectively. Use, unlike usage, is functional. As Gilbert Ryle has observed, there are misuses and ineffective uses, but there is no such thing as a misusage or ineffective usage ("Ordinary Language," in *Philosophical Review* 42 [1953]). Usage is what people *happen* to do with words and is determined by habits, while use is what *should* be done with words and is governed by rules. To

explain the right use of a word, as distinct from merely reporting its usage, a definition must give the rules that guide us in using it. In this respect definitions are rules, rather than descriptions or reports.

All three traditional theories of definition assume, mistakenly, that if definitions convey knowledge, then the knowledge they convey is of the same type as that conveyed by ordinary statements of fact. Essentialists conclude that the knowledge conveyed by definitions is descriptive knowledge of essences, linguistic philosophers conclude that it is descriptive knowledge of language usage, while prescriptivists maintain that definitions do not convey knowledge of any kind. There has been a strikingly similar three-way dispute over the status of value judgments: nonnaturalists hold that value judgments convey knowledge of an abstract realm of "values"; naturalists maintain that they convey knowledge of observable causal relations; and emotivists assert that they convey no knowledge whatsoever. Arguments about whether definitions and value judgments convey true or false information mistakenly presuppose that all information must be of the descriptive type, thus overlooking the fact that cookbooks, military manuals, Sunday sermons, and do-it-yourself instruction sheets all convey, in various ways, the kind of normative information that Ryle has called "knowledge-how" in *The Concept of Mind* (Ch. 2). Practical or ethical advice may be regarded as stating rules that inform us how to act effectively, while definitions provide rules that inform us how to speak or write effectively. In either case it may be said that the information conveyed is subject to being evaluated as good or bad, but not to being verified as true or false.

APPLICATIONS OF A CONTEXTUALIST VIEW. The three views of definition distinguished above fail to provide adequate criteria for distinguishing good definitions from bad ones. They assume that the criteria of a good definition can be stated independently of the specific context in which the definition is offered and the purpose it is intended to serve. But no brief list of criteria can be given that would enable us to judge at sight whether a definition is adequate. The most we can do on a general level is to classify the kinds of rules of use that definitions provide and the kinds of discursive purposes they serve, and to say generally that definitions are good if and only if they serve the purpose for which they are intended.

Thus, an evaluation of a definition must begin with the identification of the point or purpose of the definition, and this requires knowledge of the discursive situation in which the need for the definition arises. We use words to incite ourselves and others to action, to express and share emotions, to draw attention to things, to memorize, to make inferences, to evoke and enjoy images, to perform ceremonies, to teach, to exercise, and to show off. It is when we are unsure of the most effective use of an expression for one of these purposes that we seek a definition.

LINGUISTIC RULES. Rules governing the uses of words can be sorted into three main types: (1) referring rules, which aid us in identifying the things or situations to which a word may be applied; (2) syntactical rules, which govern the ways in which a word may be combined with other words to form phrases and sentences; and (3) discursive rules (the most difficult to formulate), which indicate when we may use language metaphorically (as in poetry) and when we must use it literally (as in science), as well as indicating differences of category or logical type (for example, the rule that one cannot predicate human qualities such as intelligence of inanimate things such as machines) and indicating when a word should be used in one sense rather than another (for example, *space* in mathematics as distinguished from physics). Discursive rules are the genuinely philosophical rules.

Rules for defining. The practical value of any account of the nature of definition is to be found in the clarity of the standards it provides for judging when a definition is good or bad. How does the pragmatic-contextualist account fare in this respect?

A number of rules of thumb for evaluating definitions have become canonical in the literature on the subject despite the fact that they make no clear sense in terms of any of the traditional views. The following rules can be found in practically every textbook on logic. They were first suggested by Aristotle in his *Topica* and have survived without change by sheer weight of tradition:

(1) A definition should give the essence or nature of the thing defined, rather than its accidental properties.

(2) A definition should give the genus and differentia of the thing defined.

(3) One should not define by synonyms.

(4) A definition should be concise.

(5) One should not define by metaphors.

(6) One should not define by negative terms or by correlative terms (thus, one should not define north as the opposite of south, or parent as a person with one or more children).

Significance of the rules. Rule 1, which makes sense only according to the essentialist theory, is nevertheless accepted by many writers who hold a prescriptive or linguistic view of definition, although these writers usually mean that a definition should indicate the properties that *define* the meaning of the term in question rather than those that just happen to hold true of the objects to which the term applies. But in such a case, the rule is vacuous; it asserts only that a definition should define rather than describe.

Rule 2 deserves its high status only if one accepts Aristotle's extension of biological classification to metaphysics, but it retains a limited value when it is reinterpreted in linguistic terms. We may understand "genus" to mean what Ryle has called the logical grammar of a term. The term defined need not be the name of any natural species or, for that matter, any object whatsoever. In defining words such as *function,* we do not identify a class of objects. We define a function as a certain type of relation, thus indicating that whatever can be said about relations in general can also be said about functions in particular. We thus provide a rule of syntax governing the word *function,* indicating with what other words it may be combined. The differentia of function—namely, that the relation is many-one between two variables—is a referring rule (criterion of identification) that helps us to identify the situations or formulas to which the term *function* may be applied. But it is wrong to think that the genus and differentia are necessary for a good definition. What must be stated in a definition varies with the definition's purpose. The genus may already be known and only the differentia needed or vice versa. Moreover, there are types of definition, such as contextual and recursive definition, that cannot be expressed in genus-differentia form. Contextual and recursive definitions provide rules for substituting a simpler expression for each of an infinite number of complex expressions of a given type.

Synonyms. The rule that forbids defining by a synonym makes sense only on the contextualist view of definitions as rules of use, although it has long been cited by supporters of the traditional views. The same books that cite this rule also insist that the definiendum must be logically equivalent to the definiens. But a synonym is just an expression that is logically equivalent to a given expression. The trouble seems to be that the term *synonym* is employed in a vaguely restricted sense to signify not just any logically equivalent expression, but a very brief one. Thus, we often find the injunction, "Do not define a word by a single other word." But this formulation, while sufficiently clear, is misleading. Is a two-word definition, such as "phonograph disc" for "record," a case of defining by a synonym or not? Just how many words may the definiens contain if it is not to violate this rule?

To make matters worse, the prohibition of synonyms is inconsistent with rule 4, which demands that a definition be concise; indeed, the more concise the definiens, the more it looks like a synonym. However, we can understand a rule only if we know what specific purpose the rule is intended to serve. A contextualist view of definitions provides the following solution to the conflict between conciseness and nonsynonymity:

Single-word definitions are seldom useful because a person who does not know the rules governing the definiendum, is not likely to know the rules governing the definiens. The more words there are in the definiens, the more likely it is that those for whom the definition is offered are familiar with some of the words and thus understand some of their rules of use. Everyone has experienced the frustration of looking up a word in a dictionary and being confounded by some equally unfamiliar synonym.

But why should definitions be concise if the greater the number of words, the greater are our chances of at least partial comprehension? One obvious answer is that brief explanations are easier to remember. A second answer is that a lengthy definiens is more likely to suggest some rules of use that are inessential to the definiendum. But the most important consideration has to do with the kind of discursive context in which the definition is employed. In mathematics and in other formal contexts such as jurisprudence and contractual language, the purpose of most definitional equations is to abbreviate discourse or notation. In such cases it is a virtue rather than a defect for the definiens to be long and complicated, since it is precisely this fact that makes the definiendum worth introducing as an abbreviation. Moreover, the complexity of the definiens is less likely to produce confusion in technical contexts because of the great pains taken to preserve consistency and precision of language. In contrast, the rule of conciseness is more appropriate to informal discourse, in which definitions are intended to translate or otherwise clarify an expression unfamiliar to some of the participants. In informal discourse, the definiens should be brief, while in formal contexts, the longer and more complicated the definiens, the more useful the definition. Clearly, one can make little sense of criteria of good definitions without specifying the context in which and the purpose for which a definition is needed.

Figurative language. Why should a definition avoid figurative language? This traditional injunction is proba-

bly a result of the concentration of classical philosophy on formal discursive contexts such as mathematics and natural science, in which figures of speech are usually out of place. But in informal contexts such as conversation, literature, public debate, and even the less technical discussions of scientists, figurative language may well be the most effective way of getting a point across, and it is certainly the only way to define expressions whose meaning is essentially figurative (for example, *fathead* may be defined as "a fool puffed up with vanity"). No literal definiens can do justice to the nuances of natural discourse, as every translator knows from bitter experience.

Negative and correlative terms. Why not define by the use of negative or correlative terms? This injunction, in contrast to rule 5, holds for informal discourse and becomes senseless when applied to formal discourse. It is perfectly proper in mathematics or logic to define "$-p$" as "the negation of p" or to define "$F^{-1}(x)$" as "the inverse of the function $F(x)$." The reason for prohibiting negative and correlative definitions in informal contexts is that a person who is unclear about the rules of use of the definiendum would be just as puzzled about the rules of use of a negative or correlative definiens.

Meaning equations. In light of the preceding discussion, it is advisable to look again at the problem of synonymity. It has already been noted that every meaning equation—that is, every definition of the form "*E*" means (or means the same as) "*x, y, z*"—provides a definiens that is synonymous with its definiendum. The very point of the definition is to assert this synonymity and thus to transfer the rules of use already known to govern the definiens to the presumably less familiar definiendum. In order to make sense of the traditional injunction against synonymous definitions, we found it necessary to interpret the synonymity in question as a special and restricted subtype of synonymity, measured by the number of words in the definiens. But although it is absurd to require that a meaning equation must not offer synonyms (in the general sense of "synonym"), it is quite sensible to cast doubt on the usefulness of meaning equations. Meaning equations provide a kind of definition misleadingly called "explicit," in contrast to axioms and postulates, which are frequently regarded as "implicit" or "partial" definitions.

It is unfortunate that meaning equations have come to be called "explicit" definitions, because their function, as we have seen, is to transfer rules of use from definiens to definiendum without articulating the rules in question, so that the rules remain implicit. The most explicit kind of definition, the kind that actually states the rules governing the use of an expression, is a very complicated matter. Outside of technical contexts, it is doubtful whether complete definitions of this kind can ever be provided. On the other hand, it is just as doubtful whether a complete articulation of all the rules of use of the definiendum need be given. We seldom, if ever, require more than one or a few rules of reference, logical grammar, or relevant discourse that happen to be obscure to us in a particular context. Thus, meaning equations are frequently neither the most valuable nor the most appropriate kind of definition. In technical discourse, contextual, recursive, and operational definitions play a far more important role than mere notational abbreviations. And in nontechnical contexts, such as teaching a child or a foreigner the use of a word, definitions by illustration, by enumeration of instances or enumeration of subclasses, and by an indefinite number of other devices (depending on the ingenuity and linguistic sensitivity of the parties concerned) are usually more appropriate and effective than meaning equations. The evaluation of specific definitional procedures remains an important task for philosophically minded experts in each field of discourse and inquiry.

See also Aristotle; Art, Definitions of; Bacon, Francis; Brunschvicg, Léon; Carnap, Rudolf; Descartes, René; Essence and Existence; Geometry; Goodman, Nelson; Hempel, Carl Gustav; Hobbes, Thomas; Hume, David; Husserl, Edmund; Intuition; Kant, Immanuel; Language, Philosophy of; Lewis, Clarence Irving; Locke, John; Logical Terms, Glossary of; Medieval Philosophy; Mill, John Stuart; Moore, George Edward; Pascal, Blaise; Plato; Proper Names and Descriptions; Quine, Willard Van Orman; Rickert, Heinrich; Russell, Bertrand Arthur William; Ryle, Gilbert; Semantics; Socrates; Strawson, Peter Frederick; Universals, A Historical Survey; Whitehead, Alfred North.

Bibliography

Aristotle. *Works.* Edited by W. D. Ross. Oxford, 1928. See especially *Physics* 192–195, *Metaphysics* 982–984, *Posterior Analytics* 90, *Topics.*

Ayer, A. J. *Language, Truth, and Logic,* 2nd ed. London: Gollancz, 1946. Chs. 3 and 4.

Bacon, Francis. *Novum Organum.* In *Works,* edited by J. Spedding. London, 1901. Vol. IV.

Black, Max. *Problems of Analysis.* Ithaca, NY: Cornell University Press, 1954. Ch. 2.

Carnap, Rudolf. *Introduction to Semantics.* Cambridge, MA: Harvard University Press, 1942. Secs. 6, 24.

Carnap, Rudolf. "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–471 and 4 (1937): 1–40.

Carnap, Rudolf. "The Two Concepts of Probability." *Philosophy and Phenomenological Research* 5 (1945): 513–532.

Carré, M. H. *Realists and Nominalists.* Oxford: Oxford University Press, 1946.

Church, Alonzo. "Definition." In *Dictionary of Philosophy,* edited by D. D. Runes. New York: Philosophical Library, 1942.

Copi, I. M. *Introduction to Logic.* New York, 1948. Ch. 4.

Descartes, René. *Rules for the Direction of the Understanding.* In *Works.* Translated by E. S. Haldane and G. R. T. Ross. Cambridge, U.K., 1911. Vol. I.

Dewey, John, and A. F. Bentley. "Definition." *Journal of Philosophy* 44 (1947): 281–306.

Dubs, Homer H. *Rational Induction.* Chicago: University of Chicago Press, 1930.

Frege, Gottlob. *Translations from the Writings of Gottlob Frege,* edited by P. T. Geach and Max Black. Oxford, 1952.

Goodman, Nelson. *The Structure of Appearance.* Cambridge, MA: Harvard University Press, 1951. Ch. 1.

Hempel, C. G. *Fundamentals of Concept Formation in Empirical Science.* Chicago: University of Chicago Press, 1952. Ch. 1, Parts 2–4.

Hobbes, Thomas. *Leviathan.* Book I, Secs. 3–5.

Husserl, Edmund. *Erfahrung und Urteil.* Hamburg: Claassen and Goverts, 1948. Pp. 410 ff.

Husserl, Edmund. *Ideas.* Translated by W. R. Boyce Gibson. New York: Macmillan, 1931. Vol. I.

Kant, Immanuel. "Introduction." In *Critique of Pure Reason,* translated by Norman Kemp Smith. London, 1929.

Kaplan, Abraham. "Definition and Specification of Meaning." *Journal of Philosophy* 43 (1946): 281–288.

Lenzen, Victor. "Successive Definition." In *Procedures of Empirical Science.* Chicago, 1938.

Lewis, C. I. *An Analysis of Knowledge and Valuation.* La Salle, IL: Open Court, 1946. Pp. 105 ff.

Locke, John. *Essay concerning Human Understanding.* Edited by A. C. Fraser. Oxford, 1894. Vol. I.

Maritain, Jacques. *Philosophy of Nature.* Translated by I. Byrne. New York: Philosophical Library, 1951.

Mill, John Stuart. *A System of Logic.* London, 1879. Pp. 72ff., 436ff.

Moore, G. E. *Principia Ethica.* Cambridge, U.K., 1913. Ch. 1.

Moore, G. E. "Reply to My Critics." In *The Philosophy of G. E. Moore,* edited by P. A. Schilpp, 660–667. Evanston, IL: Northwestern University Press, 1942.

Pepper, S. C. "The Descriptive Definition." *Journal of Philosophy* 43 (1946): 29–36.

Plato. *The Dialogues of Plato.* Translated by Benjamin Jowett. See especially *Charmides, Euthyphro, Meno, Republic, Sophist,* and *Theaetetus.*

Quine, W. V. "Truth by Convention." In *Philosophical Essays for A. N. Whitehead.* London and New York: Longmans, Green, 1936.

Rickert, Heinrich. *Zur Lehre von der Definition.* Tübingen, 1929.

Robinson, Richard. *Definition.* Oxford, 1954.

Robinson, Richard. *Plato's Earlier Dialectic.* Ithaca, NY: Cornell University Press, 1941.

Russell, Bertrand, and A. N. Whitehead. *Principia Mathematica.* Cambridge, U.K.: Cambridge University Press, 1910. Vol. I.

Scriven, Michael. "Definitions, Explanations, and Theories." In *Minnesota Studies in the Philosophy of Science II,* edited by H. Feigl, G. Maxwell, and M. Scriven, 99–195. Minneapolis: University of Minnesota Press, 1958.

Scriven, Michael. "Definitions in Analytical Philosophy." *Philosophical Studies* 5 (1954).

Wittgenstein, Ludwig. *Philosophical Investigations.* Translated by G. E. M. Anscombe. New York: Macmillan, 1953.

Raziel Abelson (1967)

DEGREES OF PERFECTION, ARGUMENT FOR THE EXISTENCE OF GOD

The proof for the existence of God from degrees of perfection, sometimes called the Henological Argument, finds its best-known expression as the fourth of Thomas Aquinas's "Five Ways" in his *Summa Theologiae* Ia, 2, 3. It is here quoted in full:

> The fourth way is based on the gradation observed in things. Some things are found to be more good, more true, more noble, and so on, and other things less. But comparative terms describe varying degrees of approximation to a superlative; for example, things are hotter and hotter the nearer they approach what is hottest. Something therefore is the truest and best and most noble of things, and hence the most fully in being; for Aristotle says that the truest things are the things most fully in being. Now *when many things possess some property in common, the one most fully possessing it causes it in the others: fire,* to use Aristotle's example, *the hottest of all things, causes all other things to be hot.* There is something therefore which causes in all other things their being, their goodness, and whatever other perfections they have. And this we call God.

COMPARATIVES AND SUPERLATIVES

A distinctive feature of the Fourth Way is the principle that "comparative terms describe varying degrees of approximation to a superlative"; for example, suppose "whiter than" is such a comparative term. The judgment that bond paper is whiter than newsprint would then be more adequately expressed as "The color of bond paper is closer to *pure white* than is the color of newsprint." How-

ever, the new comparative term "closer to" (that is, "more closely resembles," "more similar to") is used in exactly the same sense when none of the things compared is a superlative, for example, in "The color of bond paper is closer to the color of newsprint than the color of newsprint is close to the color of lemons," and here "closer to" obviously does not describe a degree of approximation to pure white. If "closer to," used to compare colors, does describe degrees of approximation to a superlative, the superlative must be the greatest possible similarity between colors, that is, qualitative identity of colors. Perhaps the initial judgment should then be expressed "The similarity between the color of bond paper and pure white is closer to the *greatest possible similarity* than is the similarity between newsprint and pure white." But here there is still a comparative term, "closer to," used to compare similarities between colors. It seems impossible to define a comparative term by means of a superlative without using another comparative term, and we are on our way to an infinite regress. If all comparative terms describe degrees of approximation to a superlative, then any comparative judgment implicitly refers to infinitely many superlatives.

But perhaps not all comparative terms describe degrees of approximation to a superlative. Suppose "closer to" (as used to compare colors) does not, and therefore the infinite regress can be cut short. Then "closer to" can be used to define "whiter than," and the definition need not refer to pure white, or to any other superlative. This is a reason for denying that "whiter than" describes a degree of approximation to a superlative. The definition runs as follows:

First it must be given, perhaps simply by fiat, that color *B* is whiter than color *A*. *B* need not be pure white, or superlatively white. Then any color *X* is between *A* and *B* if and only if both *X* is closer to *A* than *A* is close to *B* and *X* is closer to *B* than *B* is close to *A*. If *X* is between *A* and *B*, then *X* is whiter than *A*, and *B* is whiter than *X*. If *X* is different from both *A* and *B* and is not between *A* and *B*, then (1) *X* is whiter than *B* if and only if *X* is closer to *B* than *X* is close to *A* and (2) *A* is whiter than *X* if and only if *X* is closer to *A* than *X* is close to *B*. Two colors, *X* and *Y*, can be compared with respect to whiteness by (1) comparing *X* with the initially given pair in the manner just described and (2) similarly comparing *Y* with either the pair *A* and *X* or the pair *B* and *X*.

Superlative terms can be defined by means of comparatives more easily than comparative terms can be defined by means of superlatives. For example, "Brand *X* is the whitest bond paper if and only if Brand *X* is whiter than any other bond paper." Or "Brand *X* is the whitest bond paper if and only if no other bond paper is whiter than Brand *X*." On the second definition there can be more than one whitest bond paper. On the first definition there can be only one; and it is therefore possible that nothing satisfies the first definition. Such nonequivalent forms of definition are possible whatever the kind of superlative term defined; either form may be used if it is not confused with the other. Both definitions above define a relative superlative term. *Whitest* is defined with respect to a certain class, the class of bond papers. Since not only bond paper is white, neither definition rules out the possibility that something other than bond paper is whiter than the whitest bond paper. A universal superlative term is defined with respect to the class of everything of which the corresponding comparative term is predicable. For example, "*X* is the whitest thing if and only if nothing is whiter than *X*." Both relative and universal superlative terms can be absolute superlative terms. An absolute superlative term is defined by means of a modal term such as *possible* or *can*. "*X* is pure white if and only if it is not *possible* for anything to be whiter than *X*." There are as many senses of an absolute superlative term as there are relevant senses of *possible*.

Any comparative term can be used to define some superlative term. For example, "greater than" can be used to define "greatest prime number": *n* is the greatest prime number if and only if *n* is a prime number and there is no prime number greater than *n*. But it has been proved that there is no greatest prime number—that the predicate "greatest prime number" cannot be truly predicated of any number. This raises a general question: How can we know whether a particular superlative term could possibly be truly predicated of something? One can define "pure white," but this gives no assurance that there might possibly be something that is pure white. Perhaps we do not know what we are talking about when we talk about "pure white"; for perhaps there can be nothing to talk about, just as there can be nothing to talk about when we talk about "the greatest prime number." A superlative term should be suspected of not being truly predicable of anything possible unless there is a reason to think otherwise, and such a reason is not provided by the fact that the superlative term can be defined by a perfectly understandable comparative term.

Such a reason is sometimes provided when the superlative term can be defined without using any corresponding comparative or superlative terms. Definitions of this sort will usually, perhaps always, employ a universal quantifier. For example, "An object is (absolutely) *pure*

gold if and only if *all* its atoms are atoms of gold. A *perfect* reflector is one that reflects *all* the light falling on it." Definitions of the form "Something is pure ______ if and only if it contains no impurities" or "something is a perfect ______ if and only if it has *no* imperfections" will not do by themselves. The terms "contains *no* impurities" and "has *no* imperfections" are as problematic as the particular superlative terms they define and should be used without qualms only if they can be characterized independently. "Absolutely pure minestrone soup" can be defined as "minestrone soup completely free of impurities," but this is no help until we have a complete list of possible impurities. Aniline dyes are definitely impurities in soups. Some batches of minestrone soup are therefore definitely purer than others. But starting from an incomplete list of possible impurities, there is no obvious way, other than by arbitrary stipulation, of making a complete list. It seems that "absolutely pure minestrone soup" can therefore be given a clear sense only by stipulation. We do not need to give it a clear sense in order to talk sensibly about some batches of soup being purer than others.

A comparative term is often much clearer than the corresponding superlative term; one can often know how to use a comparative term without at all knowing how to use the corresponding superlative term. It seems reasonable to deny that such comparative terms describe degrees of approximation to a superlative.

PERFECTIONS

Thomas stated his principle quite generally, but presumably he would have been willing to qualify it. He argued himself that there can be nothing that is unlimited in size (*Summa Theologiae* Ia, 7, 3) and he would deny, reasonably, that the comparative term "longer than," for example, describes degrees of approximation to a superlative. The argument from degrees of perfection does not lead to the heretical conclusion that God is pure white or pure red. Still less does it lead to the impossible conclusion that God is both pure white and pure red or that God is both perfectly circular and perfectly triangular. The argument is concerned only with perfections whose predication does not imply any sort of imperfection. If a thing is white, it must be extended; if extended, it must be divisible; and if divisible, it must be perishable. Perishability is an imperfection, and therefore whiteness, like all other properties that exist only in something extended, can exist only in things less than completely perfect. Perfections that involve absolutely no imperfection are sometimes called "transcendental perfections." The traditional list includes being, unity, truth, goodness, nobility, and sometimes beauty and intelligence. Thomas thought that anything, a member of any genus, and God, who is not a member of any genus, could have these perfections. For Thomas's argument the principle about comparison need be true only of the transcendental perfections.

The principle about comparison is generally dubious, and it is particularly dubious with the transcendental perfections. Goodness, for example, is sensibly predicated of something only when it is understood as being of some kind. One who asserts of something "It is good" should be prepared always to answer the question "A good *what*?" Things of a certain kind are good in virtue of having certain characteristics; things of another kind in virtue of having others. Thus, if comparisons of goodness describe degrees of approximation to a superlative, then comparison with respect to any of the different characteristics admitting of degrees in virtue of which different kinds of things are good must also describe degrees of approximation to a superlative. The restriction of the comparative principle to transcendental perfections is not much of a restriction.

Those who do not subscribe to a Thomistic metaphysics, or to one like it, will not find any reason to accept the principle that comparisons of perfections describe degrees of approximation to a superlative. It is not surprising that Thomas's philosophy contains enough material to construct more arguments for God's existence than he formulated explicitly. Some of these back up the Fourth Way. For example, Thomas's philosophical theology makes great use of the Aristotelian distinction between act and potency: "Each thing is perfect according as it is in act, and imperfect as it is in potency" (*Summa contra Gentiles* I, 28, 6). Furthermore, something whose actuality is less than complete must be caused by something else with at least as much actuality (I, 28, 7). Bearing these two principles in mind, the argument from degrees of perfection can be reformulated as follows:

> Some things are found to be more perfect than others. Thus, some things have less than the superlative degree of perfection. Since a thing's perfection is its actuality, these things have less than the superlative degree of actuality. Something whose actuality is less than complete must be caused by something else with at least as much actuality. The resulting hierarchy of causes cannot be infinite, so there must be a first cause whose actuality is complete, who is pure act, and who therefore has all perfections in a superlative degree. And this we call God.

Thus reformulated, the Fourth Way resembles the First Way, the argument from efficient causality, and the Second Way, the argument from change. And it is susceptible to the same sorts of familiar objections raised against them. These objections, however, may seem less forceful against the Fourth Way than against the other arguments. A modern reader who is untroubled by the idea of an infinite hierarchy of efficient causes may well balk at the idea of an infinite hierarchy of increasing perfection. And one who claims that a proof of a first cause does not prove *God's* existence may admit that a proof of an absolutely perfect being does. However, this does not make the argument from degrees of perfection more convincing than the other proofs. The argument is now generally neglected, and a modern nonbeliever is not likely to be much influenced by it. For its premises will seem plausible only to one who accepts metaphysical principles, which in turn will seem plausible only to one who has a prior belief in the existence of God.

The reformulation of the Fourth Way given earlier brings out the relevance of the relation between comparative and superlative to other parts of Thomas's system. A central doctrine of Thomas's philosophical theology is that God is *pure act,* that there neither is nor could be any potency in him. Even if it is granted that we can learn, from Aristotle's and Thomas's examples, how to compare some things as being more or less in act, this gives us no reason to suppose that the superlative term "pure act" is intelligible or that it could possibly apply to something.

See also Aristotle; Perfection; Popular Arguments for the Existence of God; Thomas Aquinas, St.

Bibliography

The Fourth Way seems the most Platonic of Thomas Aquinas's Five Ways. See, for example, Plato, *Phaedo,* 100–101. Two books in English on Thomas's relation to Plato are Arthur Little, *The Platonic Heritage of Thomism* (Dublin, 1950), and R. J. Henle, *Saint Thomas and Platonism* (The Hague: Nijhoff, 1956). The Fourth Way is the major topic of Little's book. Henle's book has a very complete bibliography.

The principle that comparative terms describe varying degrees of approximation to a superlative appears several times in Thomas's *Summa contra Gentiles* (I, 28, 8, I, 42, 19; I, 62, 5). It does not, however, appear in the very compressed argument for God's existence (I, 13, 34), where an argument from degrees of truth is attributed to Aristotle. Aristotle's *Metaphysics* II, 993b25–30, a passage mentioned in both the *Summa Theologiae* and the *Summa contra Gentiles* versions of the argument, does seem adaptable to Thomas's purposes. *Metaphysics* IV, 1008b31–1009a5, mentioned only in the *Summa contra Gentiles* version, does not. Aristotle probably should not be counted among the philosophers who employed or would be willing to employ the argument from degrees of perfection to prove the existence of a perfect being.

Several arguments from degrees of perfection appear in the writings of Augustine; see, for example, Bk. V, Sec. 11, and Bk. VIII, Sees. 4 and 5, of *De Trinitate,* the work of Augustine's referred to in the preface of Anselm's *Monologion.* Anselm's arguments from degrees of perfection appear in the first four chapters of the *Monologion.*

The Blackfriars edition of Thomas's *Summa theologiae* (New York: McGraw-Hill, 1964–; only a few of the projected 60 vols. have been published so far) has the Latin text, along with a new English translation. The second volume, Ia. 2–11 (New York McGraw-Hill, 1964), translated by Timothy McDermott, contains appendices by Thomas Gilby, "The Fourth Way" and "Perfection and Goodness."

Thomas's doctrine of "analogical" predication is usually invoked to explain the notion of a transcendental perfection. This, as well as the act-potency distinction, is discussed in Knut Tranøy, "Thomas Aquinas," in *A Critical History of Western Philosophy,* edited by D. J. O'Connor (New York: Free Press of Glencoe, 1964), and further references are given.

The Fourth Way is the least widely accepted of Thomas's proofs for the existence of God. References to the disputes are given in Étienne Gilson, *The Christian Philosophy of St. Thomas Aquinas* (New York: Random House, 1956). See also Gilson's *Elements of Christian Philosophy* (New York: Doubleday, 1960).

René Descartes presented a proof for the existence of God from the degrees of perfection found in ideas. See his *Third Meditation,* his *Principles of Philosophy,* I. xviii., Objection II of the second set of *Objections to the Meditations,* and his *Reply* to this objection.

Bobik, Joseph. "Aquinas's Fourth Way and the Approximating Relation." *The Thomist* 51 (1987): 17–36.

Brady, Jules. "Note on the Fourth Way." *New Scholasticism* 48 (1974): 219–232.

Catan, John, ed. *St. Thomas Aquinas on the Existence of God: Collected Papers of Joseph Owens.* Albany: State University of New York Press, 1980.

Dewan, Lawrence. "St. Thomas's Fourth Way and Creation." *The Thomist* 59 (1995): 371–378.

Kenny, Anthony. *The Five Ways: St. Thomas Aquinas' Proofs of God's Existence.* London: Routledge and Kegan Paul, 1969.

Morreall, John. "Aquinas' Fourth Way." *Sophia* 18 (1979): 20–28.

Urban, Linwood. "Understanding St. Thomas's Fourth Way." *History of Philosophy Quarterly* 1 (1984): 281–296.

Wippel, John. *The Metaphysical Thought of Thomas Aquinas: From Finite Being to Uncreated Being.* Washington, DC: Catholic University of America Press, 2000.

David Sanford (1967)
Bibliography updated by Christian B. Miller (2005)

DEISM

Deism (Lat. *deus,* god) is etymologically cognate to theism (Gr. *theos,* god), both words denoting belief in the

existence of a god or gods and, therefore, the antithesis of atheism. However, as is customary in the case of synonyms, the words drifted apart in meaning; theism retained an air of religious orthodoxy, while *deism* acquired a connotation of religious unorthodoxy and ultimately reached the pejorative. Curiously, however, the earliest known use of the term *deist* (1564) already had this latter intent, although it was by no means consistently retained thereafter. The situation is complicated by a late eighteenth- and nineteenth-century technical metaphysical interpretation of deism, in which the meaning is restricted to belief in a God, or First Cause, who created the world and instituted immutable and universal laws that preclude any alteration as well as divine immanence—in short, the concept of an "absentee God." A further complication has been the acceptance of natural religion (religion universally achievable by human reason) by many eminent Christian theologians throughout the course of many centuries. Such theologians also believed in revelation and in personal divine intervention in the life of man, a position that had been made clear and authoritative by St. Thomas Aquinas. No sharp line can be drawn between the doctrines of such rationalistic theologians and those of deists, especially those who termed themselves "Christian deists." Nor is it accurate to maintain that the historical deists (mainly of the seventeenth and eighteenth centuries), like the philosophical deists, altogether denied the immanence of God, even though they did tend to become more and more critical of the necessity of any revelation and of the Hebraic-Christian revelation in particular. It is therefore necessary to distinguish between the two types of deists. The remainder of this entry will be devoted to a survey of historical deism.

EARLY HISTORY OF DEISM

To attempt to disentangle the antecedents of historical deism—intertwined as they are with rationalistic natural religion on the one hand, and with skepticism on the other—would indeed be foolhardy. Skepticism itself might end in Pyrrhonism or atheism or fideism. It is safe to generalize, however, that any tendency away from religious dogmatism, implicit faith and the mysterious, and in the direction of freedom of thought on religious matters, was in some measure a premonitory symptom of deism.

The earliest known use of the word *deist* was by Pierre Viret, a disciple of John Calvin, in his *Instruction chrétienne* (Geneva, 1564), Vol. II, "Epistre" (signed, Lyons, December 12, 1563). Viret regarded it as an entirely new word that (he claimed) the deists wished to oppose to *atheist:* According to him, the deist professes belief in God as the creator of heaven and earth but rejects Jesus Christ and his doctrines. Although those unidentified deists were learned men of letters and philosophy, they were bitterly attacked by Viret as monsters and atheists. This definition and commentary was given wide circulation through Pierre Bayle's citation in his article on Viret in the *Dictionnaire historique et critique* (1697; English translation, 1710). The word *deist* remained unknown in England until 1621, when it appeared in Robert Burton's *Anatomy of Melancholy* (III. iv. II. i). After discussing atheists and near-atheists, Burton continues: "Cousin-germans to these men are many of our great Philosophers and Deists," who, although good and moral, are yet themselves atheists. These "great Philosophers and Deists" likewise remain unidentified. A century and a half later, David Hume (1711–1776), in his *History of England,* ventured to name James Harrington, Algernon Sidney, and Sir John Wildman, among others, as the reputed leaders of the deists under the Commonwealth. The first interpretation of *deist* in both French and English as a euphemism for *atheist* was not followed by Dr. Samuel Johnson, who, in his *Dictionary* (1755), defined *deist* as "a man who follows no particular religion but only acknowledges the existence of God, without any other article of faith."

The first appearance of *deism* seems to have been in John Dryden's preface to his poem *Religio Laici* of 1682, where he equated it with natural religion. Dr. Johnson agreed: "The opinion of those that only acknowledge one God, without the reception of any revealed religion." Neither Dryden nor Johnson, evidently, regarded deism as disguised atheism. The notion of deism, however, if not the word itself, is to be found in one form or another throughout the Renaissance until, in the late seventeenth century, the Englishman Charles Blount openly acknowledged that he was a deist.

Beginning in the early sixteenth century, general contributions to the development of deism include such broad movements as anti-Trinitarianism, Unitarianism, secularism, anticlericalism, Erastianism, Arminianism, and Socinianism, the rise of the sects, and the general revolt against authority. It may be argued that all of these currents and undercurrents were united in the increasing trend away from religious persecution and toward religious toleration, the glorification of the natural powers of man, and the endorsement of the right to think and to publish freely on all religious and political subjects.

DEISM IN BRITAIN

The British deists constituted no conspiracy and formed no school of thought; they were highly individualistic, frequently unknown to one another, and sometimes at odds with one another. They were less systematic philosophers than thoughtful writers on practical moral, religious, and political issues. In 1704 the rationalist Anglican theologian Samuel Clarke distinguished four varieties of deists: those who denied providence; those who acknowledged providence in natural religion but not in morality; those who, while denying a future life, admitted the moral role of the deity; and finally, those who acknowledged a future life and the other doctrines of natural religion. The following summary of the leading deists will testify to the general truth of Clarke's subtle distinctions.

LEADING BRITISH DEISTS AND THE RISE OF DEISM.

Lord Herbert of Cherbury. Lord Herbert (1583–1648) never called himself a deist and had but a single acknowledged disciple, Charles Blount; nevertheless, he exerted considerable influence and deserves the title of "the father of English deism" bestowed on him in 1714 by Thomas Halyburton in *Natural Religion Insufficient.* Lord Herbert's *De Veritate, Prout Distinguitur a Revelatione, a Verisimili, a Possibili, et a Falso* was published in Paris in 1624, in London in 1633, and again in 1645. The first edition, therefore, postdated Burton's avowal of the existence of many deists by three years. In the expanded London edition of 1645, Herbert laid down the religious Common Notions that constitute the rationalistic basis of deism and that were to be assumed, if not always acknowledged, by virtually all succeeding deists. These principles are (1) that there is one supreme God; (2) that he ought to be worshiped; (3) that virtue and piety are the chief parts of divine worship; (4) that man ought to be sorry for his sins and repent of them; (5) that divine goodness dispenses rewards and punishments both in this life and after it. These truths, he argued, are universal, and may be apprehended by reason. Revelation is not openly repudiated, but by implication is rendered supererogatory. (Somewhat incongruously, however, Herbert prayed for a sign from Heaven that would grant permission to publish *De Veritate,* and was satisfied that he had received it.) Herbert treated Scripture as ordinary history, ridiculed bibliolatry, and overtly attacked priestcraft, and disavowed faith as a basis for religion. His *De Religione Gentilium* (1663) is one of the earliest studies of comparative religion.

Propagation of Deism. Although precise documentation is not available, deism was ripening between the time of Herbert and Blount, through such various and overlapping influences as humanism in general, the philosophy of Thomas Hobbes, the idealism of Harrington, the naturalistic biblical exegesis of Benedict de Spinoza and others, the corruption of the clergy, the widespread religious rationalism of the Cambridge Platonists and other Latitudinarians, the "sweet reasonableness" of John Locke, and the scientific approach of Isaac Newton—all of which were contributing to religious and political toleration. By the close of the seventeenth century, a new and memorable influence was added—the pervasive presence of the skepticism of Bayle. The first direct attack on British deism, Bishop Stillingfleet's *Letter to a Deist* (1677), acknowledges that owning to the being and providence of God but expressing "a mean esteem" of the Scriptures and the Christian religion had become a common theme.

Charles Blount. Beginning in 1679, Blount (1654–1693) was an indefatigable propagandist who, in the battle for freedom on all fronts, learned to resort to indirect methods in order to keep clear of the law. His *Summary Account of the Deist's Religion* (1693), which appeared posthumously during the same year in which he committed suicide, is his most outspoken work.

The year 1610 marks the last burning of heretics in England. Yet the matter of legal suppression of heterodox works is of vital importance in understanding and assessing the writings of the deists. The strict Press Licensing Act of 1662 was allowed to drop by 1695, but the blasphemy laws were still in effect. The ecclesiastical courts had the power to imprison heretics for a period of six months; in 1676 Lord Chief Justice Hale ruled that through common law the Court of King's Bench had jurisdiction over blasphemy, because Christianity is "parcel of the laws of England"; and finally, in 1698 a vicious statute was enacted under which any acknowledged Christian who made any accusation whatsoever against the Christian religion could be rendered incapable of holding office, of taking legal action, of purchasing land, and, if the blasphemy was repeated, would be made to suffer three years' imprisonment without bail. Such repressive measures drove the heterodox into various evasive techniques. Irony, innuendo, ridicule, raillery, allegorical interpretation of the Scriptures, fictitious analogies, frequent use of the dialogue and epistolary forms, the claim to be "Christian deists," pseudonymity, and anonymity not only successfully hampered legal

prosecution but also made it difficult for modern historians to ascertain the genuine beliefs of the writers.

After Herbert and Blount, the foremost British deists were John Toland, Anthony Collins, and Matthew Tindal, and of somewhat less consequence, William Wollaston, Thomas Woolston, Thomas Chubb, Thomas Morgan, Henry St. John Bolingbroke, and Peter Annet. Others, such as the earl of Shaftesbury and Bernard Mandeville, have been labeled deists with some justification, and many others without justification, even including orthodox clergymen who emphasized natural religion, expressed scruples about specific biblical passages or voiced doubts about specific biblical miracles.

John Toland. Toland (1670–1722) produced in 1696 his most famous deistical work, the very title of which spells out its major thesis: *Christianity not Mysterious: Or a treatise Shewing That there is nothing in the Gospel Contrary to Reason, Nor above it: And that no Christian Doctrine can be properly call'd a Mystery.* The treatise is basically rationalistic and is reminiscent of Herbert's *De Veritate.* It opposes not only biblical mysteries, but also challenges the validity of the biblical canon and points out corruptions in biblical texts. It mocks the implicit faith of the Puritans and their bibliolatry, and severely censures the vested interests of priests of all denominations. Philosophically, Toland was in the tradition of Giordano Bruno, René Descartes, Spinoza, Gottfried Wilhelm Leibniz, and, to a lesser extent, of Locke. Eclectic and somewhat inconsistent in his opinions, he was a freethinker and a deist, a materialist and a pantheist (the first use of the word *pantheist* is found in 1705 in his *Socinianism truly stated*). With his great learning, Toland became a figure of international renown, for the first time bringing deism to a wide reading public through a profusion of bold controversial publications.

Anthony Collins. Collins (1676–1729) was a well-to-do and well-educated gentleman and magistrate. At the age of twenty-seven he earned the respect and friendship of Locke. Two early works, *An Essay concerning the Use of Reason* (1707) and *Priestcraft in Perfection* (1709), prepared the way for the more famous *Discourse of Free-Thinking* (1713), in which the right to think and publish freely is examined chiefly as it pertains to religion. Enthusiasm and superstition are considered more evil than atheism; modern science and the Protestant Reformation are presented as examples of courageous freethinking that have relieved many from age-old errors, including witchcraft; and priests are blamed for trivial quarreling among themselves over biblical interpretations and are held responsible for many corrupt texts. An impressive list of freethinkers is furnished from the ancient Greeks, Romans, and Hebrews; from the Church Fathers; and from the moderns, ranging from Michel Eyquem de Montaigne to John Tillotson and Locke.

Collins defended his style of writing in *A Discourse concerning Ridicule and Irony in Writing* (1727); his philosophical doctrine of necessitarianism (wherein he differs from the doctrine of free will espoused by most deists) is developed in a *Philosophical Inquiry concerning Human Liberty* (1715) and a *Dissertation on Liberty and Necessity* (1729); and his biblical criticism, mainly of the supposed fulfillment of Old Testament prophecies in the New Testament, in the *Discourse of the Grounds and Reasons of the Christian Religion* (1724) and the *Scheme of Literal Prophecy Considered* (1725). Collins is unquestionably the most readable and urbane of the British deists.

Matthew Tindal. A law fellow at All Souls College, Oxford, and advocate at Doctors' Commons, Tindal (1657?–1733) was the most learned of the British deists, as well as the most significant historically. His *Christianity as Old as the Creation: Or, The Gospel A Republication of the Religion of Nature* (1730), composed in dialogue form, was at once recognized as "The Deist's Bible," and elicited over 150 replies, the most famous of which is Bishop Joseph Butler's *The Analogy of Religion* (1736). Although a declared admirer of Locke, Tindal deduces the being and attributes of God by a priori reason. As man reasons downward from the knowledge of the attributes of God to knowledge of himself, the religion of nature, including all the moral precepts requisite for leading the life of virtue and achieving ultimate salvation, then follows. Scripture, replete with ambiguities, is not only unnecessary but is actually confusing to men of reason; and according to Tindal, all men of whatever education or status in life are capable of Right Reason. Some Old Testament heroes are inspected in detail and are found wanting in virtue; even some New Testament parables are subjected to critical comment. Tradition is repudiated as a basis for Christianity, since it can be used equally as the basis for any and all religions. The customary deistical castigation of priestcraft is combined with this repudiation of tradition. Tindal, a rationalist, always maintained the title of "Christian deist."

LESSER ENGLISH DEISTS. The remaining British deists, already named, each made some personal contribution to the movement, however small.

William Wollaston. A graduate of Sidney Sussex College, Cambridge, Wollaston (1660–1724) took holy orders, but through the unexpected inheritance of a large

fortune he was able to devote himself to moral philosophy and general learning. His *The Religion of Nature Delineated* (1724) was well received by Queen Caroline the Illustrious, as well as by the public at large. It was attacked, however, by the American deist Benjamin Franklin and was subjected to ridicule by Lord Bolingbroke, the British deist. Unlike most deistical treatises, it contains no biblical criticism of any sort. Almost purely rationalistic, it has obvious affinities, in a simplified form, with Herbert of Cherbury's religious Common Notions. Man knows truth (that is, things as they are) by means of reason; vice, or the denial of things as they are, is a lie. To seek happiness is man's duty, because happiness, or the excess of pleasure over pain, is part of man's approach to truth. Man is by nature not fundamentally selfish; his search for truth must take into account the happiness of others. It is altogether likely that Bishop Butler, in *The Analogy of Religion,* had Wollaston at least partly in mind when he reproved extreme religious rationalism as "that idle and not very innocent employment of forming imaginary models of a world, and schemes of governing it."

Thomas Woolston. Woolston (1670–1731), fellow of Sidney Sussex College, Cambridge, and Christian divine, was a deist of another stamp. A disciple of Anthony Collins, who had spearheaded the assault on biblical prophecies, Woolston extended the assault to biblical miracles. Influenced by the writings of the Greek Church Father Origen, he interpreted Scripture allegorically, was subsequently deemed out of his mind by his adversaries and, as a result, in 1720 was deprived of his fellowship. In 1705 he first employed the allegorical method in *The Old Apology for the Truth of the Christian Religion against the Jews and Gentiles Revived,* and later published a series of anticlerical tracts against those who spurned it. But it was a series of six *Discourses On the Miracles of our Saviour, In View of the Present Contest between Infidels and Apostates* (1727–1729) that brought prosecution by the government, ending in 1729 with a conviction of blasphemy. Sentenced to a fine of £100, imprisonment for one year, and security for good behavior during life, he died in jail in 1731, unable to pay the fine. A fighter for freedom of thought and publication for all, Woolston ironically fell the victim of his own principles. The six *Discourses* take a colloquial and frequently witty dialogue form, with a fictitious learned Jewish rabbi presenting Woolston's queries concerning fifteen New Testament miracles. Woolston's madness may possibly have been real (in which case his sentence was truly infamous), but his tracts read more like the strong convictions of a strong mind. He was one of two of the leading British deists (the other being Annet) to suffer punishment by the government.

Thomas Chubb. An Arian and "Christian deist," Chubb (1679–1746) was a self-educated and humble artisan. Writing for the common people, Chubb was also able to hold his own with the educated upper classes, divines, and scholars. He mastered the widespread rationalism of the early eighteenth century and propagated its basic ideas through prolific publication, as is observable in such works as *A Discourse concerning Reason, With Regard to Religion and Divine Revelation* (1731) and *An Enquiry Into the Ground and Foundation of Religion. Wherein is shewn, that Religion is founded in Nature* (1740). Another approach is taken in *A Discourse on Miracles, Considered as evidence to prove the Divine Original of a Revelation* (1741), a work influenced by Toland and Woolston. Although he is skeptical of the Hebrew revelation, Chubb is never skeptical of the Christian, as is manifested in *The True Gospel of Jesus Christ asserted* (1732) and *The True Gospel of Jesus Christ Vindicated* (1739). In these two tracts, Chubb employs natural religion as proof of Christian religion. He defends the miraculous propagation of primitive Christianity against the aspersions of the deist Tindal. A believer in free will, Chubb was answered at considerable length by the eighteenth-century American theologian Jonathan Edwards in *A Careful and Strict Enquiry into The modern prevailing Notions of the Freedom of Will* (1754).

Thomas Morgan. A Welsh "Christian deist," divine, and medical doctor, Morgan (d. 1743) came from a poor family (as did Chubb and Annet). Morgan combined the religious Common Notions of Lord Herbert with some of the principles of historical biblical criticism found in the writings of Toland and Chubb. He opposed Chubb, however, on the question of free will. Morgan's chief contributions to the deistical controversy are to be found in *The Moral Philosopher, in a Dialogue between Philalethes, a Christian Deist, and Theophanes, a Christian Jew* (1737), and its two sequels. His general historical criticism of Scripture stresses the many ambiguities that permit many different interpretations of biblical texts by believers who truly attempt to understand their significance. All history, therefore, is simply probability, and infallibility is fostered by priestcraft for selfish purposes. Toleration, reasonableness, and freedom are necessary to combat superstition and persecution.

Henry St. John, Viscount Bolingbroke. Tory statesman, historian, deist, and wit, Bolingbroke (1678–1751) left his philosophical and religious compositions to be published posthumously in 1754 by David Mallet. Regarded by Dr. Johnson as a "blunderbuss" against religion and morality, Bolingbroke's *Works* were regarded by

Hume as unoriginal and feeble. In the twentieth century, Voltaire's long-alleged great indebtedness to Bolingbroke has been discredited, and the claim that Alexander Pope's *Essay on Man* was founded on Bolingbroke's *Fragments or Minutes of Essays* has been vigorously challenged. As a philosopher Bolingbroke is a rationalist, but a curiously inconsistent one. In one passage he states that only Right Reason can demonstrate the Being of Deity, yet in another, that only empiricism can prove the Being of Deity. Paradoxes abound: No universal revelation has ever been made, but modern religion can benefit by the study of primitive religions—for example, of China and Egypt. Like all the deists, Bolingbroke regarded the baneful influence of priestcraft as a major cause of the corruption of religious texts and religious traditions. With Bolingbroke, the course of British rationalistic deism, stemming from that of Lord Herbert in the middle of the seventeenth century, up to the middle of the eighteenth century, had been pretty well played out, but there was always opportunity for remorseless repetition and intensified publicity.

Peter Annet. Schoolmaster Annet (1693–1769) may be regarded as the last of the old-line deists. An outspoken freethinker, Annet advocated the freedom to divorce and, in a long series of tracts, attacked the Resurrection of Jesus and the character and conversion of St. Paul. His truculent assault on the credibility of all miracles in general, and those of the Old Testament in particular, carried on in *The Free Enquirer* of 1761, brought a governmental charge of blasphemous libel to which Annet pleaded guilty. The inhumane sentence against a man aged seventy included imprisonment for a month, two pilloryings, hard labor for a year, a fine, and bonds of security for good behavior during life. Annet survived this flagrant miscarriage of justice with its attendant humiliation and returned to schoolmastering until his death. The ascription to him of the authorship of the notorious *History of the Man after God's own Heart* (1761) has been disproved by modern scholarship. Although he contributed little fresh to the deistical movement, Annet, like Chubb, wrote directly to the people in their own language.

THE RATIONALISTIC CLIMATE OF OPINION. Little has been said so far about the rationalistic "orthodox" of the seventeenth and eighteenth centuries, those Latitudinarians, who were closely akin to the deists, except on the one crucial point of raising objections against Christian revelation. Nevertheless, both groups were united in a contemptuous rejection of Tertullian's dictum, *credo quia impossibile est*; in this respect, there was no warfare between reason and religion. In a 1670 defense of the orthodox rationalists, a Latitudinarian was succinctly defined as "a gentleman of a wide swallow."

Ralph Cudworth. Cudworth (1617–1688) may be taken as representative of the small but important band of Cambridge Platonists who sought to synthesize the spirit of Christianity with that of Greek philosophy by affirming that reason is spiritual as well as intellectual. Cudworth distinguishes between fundamental and nonfundamental religious doctrines: "I perswade myself, that no man shall ever be kept out of heaven, for not comprehending mysteries that were beyond the reach of his shallow understanding; if he had but an honest and good heart, that was ready to comply with Christ's commandments" (*A Sermon before the House of Commons, March 31, 1647.*) In *The True Intellectual System of the Universe* (1678), Cudworth argues cogently against fatalism. His posthumous *Treatise concerning Eternal and Immutable Morality* (1731) derives morality from natural law rather than from the positive precepts of revelation. Another member of the group, Benjamin Whichcote, states their position admirably: "If you would be religious, be rational in your religion." In short, the Cambridge Platonists stood for reason and moderation.

John Tillotson. Tillotson (1630–1694), archbishop of Canterbury and great champion of Anglicanism, employed rationalistic arguments against the Catholic use of tradition and authority. Observing that these same arguments could be turned against Christianity itself, the deists frequently seized upon Tillotson's authority and quoted his arguments in this new context. Collins went so far as to name him the man "whom all English freethinkers own as their head."

THE NEW SCIENCE. It might be expected that the New Science, which had made such great strides from Nicolas Copernicus to Newton, would have precipitated warfare between science and religion as it did in the nineteenth century, following Charles Darwin's *Origin of Species* (1859). But insofar as Britain was concerned, such was not the case, for Francis Bacon had enunciated the principle of a rigid dichotomy between science and religion that, on the whole, was adhered to during the seventeenth century. Indeed, science was more generally used as a bulwark for Christianity than the reverse—notably, in the case of the Latitudinarians. Newton himself was a student of Old Testament prophecies and believed in the Scriptures as inerrant guides.

The "skeptical chemist" Robert Boyle wrote orthodox religious tracts, one of which had the ancillary pur-

pose of proving that by being "addicted" to experimental philosophy, a man is assisted rather than indisposed to being a good Christian. In 1691 Boyle endowed a lectureship for the proof of the Christian religion against the attacks of infidels. Great efforts were made to replace a priori reasoning with the argument from design. Richard Bentley, the first Boyle Lecturer, corresponded with Newton in preparing *The Folly of Atheism and what is now called Deism* (1692). William Derham's two lectures, *Physico-Theology* (1713) and *Astro-Theology* (1715), continued the effort. Nevertheless, the bulk of the Boyle Lectures, from the beginning to 1732, are almost purely rationalistic, as, for example, Clarke's *Demonstration of the Being and Attributes of God* (1704) and *Discourse concerning the Unchangeable Obligations of Natural Religion, and the Truth and Certainty of the Christian Religion* (1705). Collins gibed that until Clarke's "demonstration" of the existence of God, nobody had doubted the fact; and Franklin, in his autobiography, acknowledged that he became a deist after reading some of the Boyle Lectures. The New Science, in effect, had relatively little influence on the course of the deistical controversy, since neither side squarely faced the problem of the relationship of science to religion.

THE DECLINE AND FALL OF REASON. Rationalistic refutations of deism were prolific and formidable but achieved relatively little because they had so much in common with those of deism. Tindal had forced upon the apologists acceptance of the natural sufficiency of reason in theology. Thus, if deism was to be defeated, it had to be from a citadel other than that of an infallible and universal reason. One of the infrequent replies to Tindal's direct challenge, "Dare any say that God is an Arbitrary Being, and His laws not founded on the eternal reason of things?" (*Christianity as Old as the Creation*) was *The Case of Reason, Or Natural Religion Fairly and Fully Stated* (1731). Its pietistic author, William Law (1686–1761), better remembered for his *A Serious Call to a Devout and Holy Life* (1729) and as a forerunner of John Wesley (1703–1791), totally disavowed Right Reason in the areas of morality and religion, and argued for historical evidence and implicit faith.

Bishop Joseph Butler (1692–1752) offered in the *Rolls Sermons* (1726) an important revaluation of the authority of conscience and in the *Analogy of Religion* (1736) a matter-of-fact defense of Christianity; he sought to prove by analogy that all deistical objections against revelation were equally applicable to natural religion. The danger of this argument (which employed some of the methods of science and of Lockean empiricism) was that it might conceivably drive readers to become skeptical of both kinds of religion, to espouse atheism, or to retreat into implicit faith.

Bishop George Berkeley's (1685–1753) *Alciphron, or the Minute Philosopher* (1732), with its subtitle "Containing an Apology for the Christian religion against those who are called Freethinkers," is a brilliant series of polemical dialogues, but it contains little of his highly controversial and much misunderstood philosophical denial of abstract ideas and of "matter," for which Berkeley was frequently accused of being a skeptic. His *The Analyst* (1734), addressed to an "infidel mathematician" (presumably Edmund Halley), adopts the hazardous method of defending orthodoxy by asserting that the axioms of mathematics are as irrational and incomprehensible as the mysteries of Christianity.

Law and Butler had paved the way for antirationalistic assaults on deism, the former through faith, the latter through matter of fact. The argument for faith was implemented in *Christianity Not Founded on Argument* (1742) by Henry Dodwell ("the younger"), who had as little use for historical proofs as for intellectual proofs. According to Dodwell, the Boyle Lectures, like all rationalistic efforts, had only succeeded in spreading infidelity; external proofs have no real evidential value; probability reigns; so in the final analysis, there is no other way to approach religion, than to believe because you wish to believe. With Dodwell's appeal to emotionalism, the "enthusiasm" of Wesley was just around the corner.

Conyers Middleton (1683–1750), Anglican clergyman, and equally antirationalistic, pressed the historical argument against external proof of the validity of religious claims in his *Free Inquiry into the Miraculous Powers which are supposed to have subsisted in the Christian Church from the Earliest Ages through several successive Centuries* (1749). Professedly denying the supernatural powers associated with the growth of Catholicism, Middleton could scarcely have been unaware that the same arguments could also be used to attack Gospel miracles, and that there is in actuality no breach between sacred and profane history.

Fatal blows to the Age of Reason (as differentiated from the Age of Enlightenment) came simultaneously on two levels—intellectually, from Hume and emotionally, from Wesley. What might be termed the deistical side of Hume can most readily be seen in "Of Miracles" and "Of a Particular Providence and of a Future State" (1748), "The Natural History of Religion (1757), and *Dialogues concerning Natural Religion* (1779), the last of which comes to the purposefully lame conclusion "that the

cause or causes of order in the universe probably bear some remote analogy to human intelligence." Natural religion, whether of the rationalistic or matter-of-fact variety, can lead only to doubt, uncertainty, and suspension of judgment. In reality, of course, Hume was no deist, but rather an antideist, a skeptic who destroyed the vulnerable a priori basis of deism.

At about the same time, Wesley attacked deism through "enthusiasm," the doctrine of continuous personal inspiration and inner conversion of the soul: "By grace are ye saved through faith." The fatal blows had been delivered; the Age of Reason had fallen and deism was dead. Or was it? The question will be taken up after brief considerations of deism in France, Germany, and America.

DEISM ON THE CONTINENT

The term *Enlightenment* was unknown in Britain during the eighteenth century, although its spirit was plainly manifest. When it did appear in the nineteenth century, it was employed in the derogatory sense of shallow and pretentious intellectualism coupled with unreasonable contempt for tradition and authority. In eighteenth-century France and Germany, on the contrary, full-fledged movements of *Éclaircissement* and *Aufklärung* were under way and were winning important intellectual and political victories. The present section will confine itself, insofar as possible, to religion and will deal with only a few predominant thinkers.

VOLTAIRE. Without stopping to investigate such sixteenth-century precursors as Jean Bodin, Rabelais, Pierre Charron, and Montaigne, or such seventeenth-century precursors as Descartes, Pierre Gassendi, Bernard Le Bovier de Fontenelle, and Bayle, it is well to proceed directly to François-Marie Arouet, universally known as Voltaire (1694–1778), the greatest of the French deists. Banishment to England (1726–1729) by order of the ancien régime put the already widely known poet, playwright, *philosophe* (and later, historian and novelist) into the scientific atmosphere of Newton, the philosophical and religious atmosphere of Locke and some of the earlier deists (Voltaire had already known Bolingbroke in France), and the literary neoclassical atmosphere of Jonathan Swift and Pope. Much impressed by the relatively tolerant attitudes of the English as compared to the rigid censorship of the ancien régime, Voltaire published in London in 1733 *Letters concerning the English Nation.* A surreptitiously arranged French version of 1734, *Lettres philosophiques,* speedily burned by the common hangman, was Voltaire's first bombshell against governmental and church tyranny. Thereafter, his remorseless battle cry of *Écrasez l'infâme!* was to be heard throughout a long life of polemic.

Although he consistently used the word *theist* in reference to himself, Voltaire was a deist in the tradition of the British deists, never attacking the existence of Deity but always the corruptions of church and priestcraft. As late as 1770, in a letter to Frederick the Great voicing strong disapproval of the avowed atheism of many of the *philosophes,* Voltaire repeated his conviction that if God did not exist, it would be necessary to invent him. The *Lettres philosophiques* eulogizes the Quakers as ideal deists for their freedom of thought and their freedom from dogmatism and clericism; attacks Blaise Pascal's Pyrrhonism, which leaves man only the alternative of implicit faith; praises the philosophical empiricism and religious reasonableness of Locke; and seeks to convert the scientists of France to the Newtonian system. Other writings on religion and morality, *Poème sur la loi naturelle* and *Poème sur le désastre de Lisbonne,* both of 1756, as well as the famous novel *Candide* (1759), assail the doctrine of philosophical optimism and, indeed, of divine benevolence. Believing as he did in a natural religion based on reason, Voltaire's chief onslaughts were upon dogmatism, superstition, fanaticism, and tyranny. His *Traité sur la tolérance* (1763), a classic denunciation of oppression, occasioned by the infamous Calas *affaire* of 1762, was followed in 1764 by the witty and effective *Dictionnaire philosophique.* Like most of the so-called deists, Voltaire was fundamentally a humanist seeking to better the condition of humankind.

JEAN-JACQUES ROUSSEAU. Novelist, political writer, deist, *philosophe* and anti-*philosophe,* Rousseau (1712–1778) remains one of the most inscrutable literary and philosophical geniuses of all time—a supreme individualist doting upon his own uniqueness. Born a Protestant, he became a Catholic, and finally a deist. His *Confessions* reveals that it was the reading of Voltaire's *Lettres philosophiques* that first incited him to study, to think, and to become a dedicated man of letters.

In touching solely upon Rousseau's role as a deist, it is fitting to examine the "Profession of Faith of a Savoyard Vicar," part of the fourth book of *Émile, ou de l'éducation* (1762). The first book had opened with the affirmation that everything is good as it comes from the Author of all things, but that everything degenerates in the hands of man. The fourth book seeks to develop and clarify this thesis, using, for prudential purposes, a vicar as

spokesman. Jettisoning metaphysical proofs of God and subscribing to no strict system, the vicar simply feels God within himself, as a world governor of will, intelligence, power, and goodness. This beneficent deity is to be worshiped from the heart, and not through artificial forms. Yet it is paradoxically evident that while mere animals are happy, superior man is miserable. Why? asks the vicar. He replies to his own question that far from being a simple uncompounded creature, man is actually a being of contradictions. Self-love is natural to him, but a sense of justice or conscience or inner light is innate; he has the power to will things, but does not always exert this power to enforce his will.

Man, therefore, is the author of evil: Born good, he acquires vice. God, infinitely powerful, is infinitely good and supremely just. To emulate God in seeking justice is man's only source of happiness. In this respect, natural religion, learned through conscience, is sufficient. Christian revelation, on the one hand, is fraught with difficulty, mystery, obscurity, and dogma. Its majesty, sublimity, and beauty, on the other hand, bear witness to its divinity: It is not a manmade invention; indeed, it remained Rousseau's "pillow-book" throughout life. Rousseau, in brief, is a sentimental and primitivistic, rather than a "hard," rationalistic deist. Yet, in substance, his "soft" sentimental deism is actually not far removed from the religious Common Notions of Lord Herbert or even from Spinoza's Doctrines of Universal Faith.

Rousseau's device of using the Savoyard vicar as spokesman for his own deism was unsuccessful; *Émile* was publicly burned and an order was issued for the arrest of the author, who was forced to flee the country. Except for his much later autobiographical writings, *Émile* was Rousseau's last major work.

ATHEISM. Aside from Voltaire, who subscribed to "hard" deism, and Rousseau, who dispensed the "soft" variety, the *philosophes* were not deists at all. To them, deism was but the starting point on the road to atheism. Their militant atheism, as well as their dogmatic belief in constant and inevitable progress and the perfectibility of man, shocked Gibbon and Hume, and greatly disturbed both Voltaire and Rousseau. The names of Jean Le Rond d'Alembert and Denis Diderot (editors of the *Encyclopédie*), Baron d'Holbach (and his "atheistical club"), Claude-Adrien Helvétius, F. M. Grimm, Julien Offray de La Mettrie, …tienne Bonnot de Condillac, and Marquis de Condorcet can hardly be excluded from the list of atheistical *philosophes* or, at least, those well on the road to atheism. Deism in France, although considerably influenced by deism in England, was much more extreme both religiously and politically, simply because England had already made considerable social progress. In France, deism was part and parcel of the general move toward materialism, freedom of thought and publication, freedom from the tyranny of the ancien régime in the affairs of state and church, that ultimately exploded in the Revolution.

DEISM IN GERMANY. The course of the *Aufklärung* differed in major respects from the analogous movements in Britain and France, and developed later. Under the domination of the earlier Leibniz-Wolff philosophy, rational supernaturalism generally prevailed. After 1740 (the year of the accession of Frederick the Great, the first modern freethinking king), numerous translations of the British deists and of their orthodox refuters (as indicated in G. W. Alberti's *Briefe betreffend den allerneusten Zustand der Religion und der Wissenschaften in Gross-Brittannien* of 1752–1754, J. A. Trinius's *Freydenker-Lexicon* of 1759, and U. G. Thorschmid's *Freidenker-Bibliothek* of 1765–1767) introduced a new influence. Although the German *philosophes* were widely read, there was little of French radicalism in either their religious or political thinking. Among out-and-out deists (called *Freidenkers,* or Freethinkers), the names of Karl Bahrdt, Johann Eberhard, Johann Edelmann, and Hermann Samuel Reimarus must be mentioned.

Hermann Samuel Reimarus. The apology of Reimarus (1694–1768) for natural religion as opposed to atheism and materialism, written in 1755, was Englished in 1766 as *The Principal Truths of Natural Religion Defended and Illustrated.* His direct attacks on Christianity, through a painstaking study of New Testament texts, included "On the Object of Jesus and His Apostles" and "On the Story of the Resurrection," and were published posthumously (1774–1778) by Gotthold Ephraim Lessing as *Fragments of an Anonymous Work found at Wolfenbüttel.*

Gotthold Ephraim Lessing. Lessing (1729–1781), distinguished man of letters and author of the *Laokoon* (1766) and *Nathan the Wise* (1779), was a freethinker in the nonabusive sense of the term. He should probably not be classified as a typical deist, since he professed belief in natural revelation in his last publication, *The Education of the Human Race* (1780), and at the close of his life he is said to have privately acknowledged pantheistic beliefs. Lessing's lifelong friend Moses Mendelssohn (1729–1786), a Jewish freethinker, is customarily classified as a deist in the loose usage of the term.

Immanuel Kant. The case of Immanuel Kant (1724–1804), the greatest of the German philosophers, is highly instructive. Born and educated as a religious Pietist, he came under the influence of Newtonian physics and always remained interested in science. In theology his three most famous critiques, stimulated by the "mitigated scepticism" of Hume, agree with Hume in principle. The *Critique of Pure Reason* (1781) presses beyond Hume in criticizing proofs of the existence of God; the *Critique of Practical Reason* (1788) is concerned with moral experience in natural religion; and the *Critique of Judgement* (1790), in a sense, mediates between the first two. Kant's position as a "Christian deist," however, is best expressed in *Religion within the Limits of Reason Alone* (1792–1794). The limits of religion, basically naturalistic, are set in conscience or practical religion. Christianity is stripped of mystery and tradition and is treated as a purely moral religion—in fact, the only purely moral one; God is the moral Creator of the world, and it is the duty of the good man to worship him. Kant's transcendental philosophy is beyond the scope of this entry, but it is relevant to say that Kant was the leader of the *Aufklärung,* which he defined as the freeing of man from the self-imposed bondage of the mind, and proclaimed as its motto *sapere aude* ("dare to know").

DEISM IN THE UNITED STATES

The works of the British deists, as well as those of the defenders of the faith, were well known in American intellectual circles, commencing with the second quarter of the eighteenth century. In the latter half of the century, Voltaire's "hard" deism and, especially, Rousseau's "soft" deism were widely disseminated; but the atheism of the *philosophes* made little headway. The Great Awakening, triggered by the preaching of Edwards in 1734 and bolstered by the preaching of the English Methodist George Whitefield, militated against orthodox Puritanism and in favor of republicanism both in religion and politics, but the atmosphere of rationalism still prevailed. Before the Revolution, however, deism made relatively little progress. Among the intelligentsia at Harvard, nevertheless, the Dudleian Lectures were established in 1755 for the purpose of explicating natural religion. Alarms sounded by the orthodox that deism was sweeping the country were unjustified. However, the Treaty of Paris in 1763 and the French alliance at the time of the Revolution undeniably quickened the spread of radical Gallic ideas.

MAJOR AMERICAN DEISTS.

Benjamin Franklin. Franklin (1706–1790), man of letters, scientist, and diplomatist, as early as 1723 acknowledged himself a deist to intimate friends but circumspectly continued church attendance throughout life, thereby setting the conservative pattern followed by most of the leaders of the colonial and Revolutionary periods. In London in 1725 Franklin published his *Dissertation on Liberty and Necessity, Pleasure and Pain* in opposition to the free-will doctrine of the British deist Wollaston. However, Franklin shortly repudiated and suppressed this juvenile work. When he was about twenty-two, he drafted "Articles of Belief and Acts of Religion," a creed not unlike Lord Herbert of Cherbury's religious Common Notions and one that sustained him for life. Prudence and practicality characterize all of Franklin's publications and actions. *Poor Richard's Almanack* (1732–1757) is the essence of common sense, or how to get along in the world without unduly disturbing society; his list of virtues by no means coincides with the Christian virtues.

Thomas Jefferson. Framer of the Declaration of Independence, diplomatist, vice president and twice president of the United States, and member of the Episcopal Church, Jefferson (1743–1826) was in reality a deist, rationalist, and, above all, a humanitarian. He compiled but never published what later came to be known as *The Jefferson Bible, being The Life and Morals of Jesus Christ of Nazareth.* This little work, a cento of clippings from the Gospels of Matthew, Mark, Luke, and John pasted in a blankbook, extols Jesus as a man for his moral teachings, omits ambiguous and controversial passages, and, while rejecting many of the supernatural elements, presents the core of Christian morality and is genuinely religious in tone. Religion, for Jefferson as well as for Franklin, was essentially a utilitarian moral code.

George Washington. Washington (1732–1799), general and first president of the United States, was a deist of a similar stripe. Although he always maintained a church pew, he was one of the leading statesmen who advocated total separation of state and church and who saw to it that no reference to Christianity or even to Deity was made in the Constitution. In answer to a direct question from a Muslim potentate in Tripoli, Washington acquiesced in the declaration of Joel Barlow, then American consul in Algiers, that "the Government of the United States of America is not in any sense founded on the Christian religion."

Thomas Paine. Born in England, Paine (1737–1809) arrived in America in 1774, bearing a letter of introduction from Franklin. A political theorist, diplomatist, and man of letters, Paine was a deist, but not overtly until the

publication in Paris of his *The Age of Reason: Being an Investigation of True and Fabulous Theology* (1794–1796). The first of its two books, intended to rescue deism from the reigning French atheism, is a more or less scientific assault upon revealed religion in general as being supererogatory to natural religion. The second book carries the attack directly to both the Old and New Testaments, arguing that the Bible is not the word of God and depicting Christianity as a species of atheism. Paine wrote vigorously and extensively and was outspoken in carrying his message to the common people, whose battles he had fought on the political, social, and economic fronts as well. In *The Age of Reason* the battleground was not new but was considerably enlarged from that of any earlier British deist. The work offended readers in France and shocked many in England and America who were laboring under the delusion that the deistical controversy was over and that orthodoxy had triumphed. Paine was rewarded for his efforts by banishment from England and by social obloquy in America. The patriot who throughout a long and turbulent career had accomplished so much for the new country, the man who had so vigorously combated atheism, was held to be an atheist, infidel, radical, and drunkard.

LESSER AMERICAN DEISTS. Paine was not the first acknowledged American deist, for the year 1784 produced *Reason the Only Oracle of Man, or a Compendious System of Natural Religion.* Its author, Ethan Allen (1738–1789), Revolutionary hero and leader of the Green Mountain Boys, had acquired his deism through early reading of the British deists. His book is flagrantly anticlerical and anti-Christian; he argues that a rationalistic universal religion of nature that provides the fundamentals of morality is all-sufficient and needs no supplementation. Both the Hebraic and the Christian testaments are subjected to ridicule. Like Paine, Allen was not so much an original thinker as a fearless propagandist.

Beginning in 1793, the blind ex-Baptist preacher Elihu Palmer (1764–1806) led a fiery deistical campaign from the lecture platform and by publication against the divine authority of the Bible. In 1794 he rushed to the defense of Paine's *Age of Reason* and in 1801–1802 published *Principles of Nature; or, a Development of the Moral Causes of Happiness and Misery among the Human Species.* From 1803 to 1805 he edited a weekly deistical paper, *Prospect; or, View of the Moral World.* Palmer also organized the Deistical Society of New York. With his many speeches and tracts designed to disseminate deism among the lower classes, Palmer was a most unusual deist, in that he was deliberately leading a popular crusade.

Philip Freneau (1752–1832), writer of patriotic verse, was also the American poet of the religion of nature and humanity, and his ideas were close to those of Paine. The very titles of such poems as "Belief and Unbelief: Humbly recommended to the serious consideration of creed makers," "On the Uniformity and Perfection of Nature," "On the Religion of Nature," tell their own story, without need of commentary.

DECLINE OF DEISM. During the eighteenth century, Puritanism in America had begun to crumble under the combined attacks of the Great Awakening, Methodism, and deism. "The Triumph of Infidelity" (1788), the poem by Timothy Dwight, orthodox president of Yale University, bears weak witness to the strength of deism. Shortly after 1800 deism became submerged in a revival of enthusiastic evangelism, particularly in the frontier areas, where intellectual attainments were hardly predominant. In New England, Unitarianism began making headway under the influence of Joseph Priestley, who in 1794 had immigrated from England. But elsewhere emotionalism, conservatism, reaction, and fideism were triumphant.

THE LEGACY OF DEISM

Historical deism, a term of many connotations, was essentially rationalism applied to religion, and as such was the counterpart to literary neoclassicism. Deism and neoclassicism flourished at approximately the same time, both stressing universality and shying away from particularity. In deism, this cardinal point meant that from the very beginning the Hebraic and Christian revelations were suspect, if not invariably attacked. Deism primarily put forth the view that the aim of religion is morality and that anything traditionally taught beyond morality is superfluous. The widely accepted distinction between constructive deism and critical deism, or, as it has also been put, deism before Locke and deism after Locke, or humanistic deism as opposed to scientific deism, will not survive the careful scrutiny and evaluation of leading deistical texts. Yet the prime position of Right Reason in deism did not prevent empiricism, in the form of scholarly examination of Scriptural texts and historicism, from assuming increasingly important roles. Edward Gibbon's purely naturalistic investigation into the early progress and establishment of the Christian religion in the famous (or infamous) fifteenth and sixteenth chapters of his *Decline and Fall of the Roman Empire* (Vol. I, 1776) was manifestly influenced, not only by the philosophical

skepticism of Hume, but also by the somewhat crude historical investigations of a number of the deists themselves.

One general development of the deistical movement, therefore, was the rise of "the higher criticism": The Bible was no longer deemed sacrosanct, and its verbal inspiration no longer dogmatically assumed. A second development was the greatly intensified study of comparative religion. A third development was the rise of "the philosophy of religion," spurred on by Hume's demonstration that no matter of fact, including the existence of God, can be proved a priori.

In actuality, deism did not die; it did not even fade away, and it still exists in fact, though perhaps not in name, for those who say (with Voltaire) that there must be a God and those who say (with Rousseau) that they know there is a God. Nor was deism vanquished, as has so often been asserted, by the superior talents of its orthodox opponents, by the exhaustion of the subject, or by the incapacities of its protagonists: Certainly, among the English, at least, Toland, Collins, and Tindal were the intellectual equals of most of their adversaries. By and large, both orthodox and heterodox alike were rational theists of a somewhat naive variety. Charles Leslie's *Short and Easy Method with the Deists* of 1696 proved, in actuality, neither short nor easy. The deists were long subjected to the *odium theologicum,* and the historians of the movement have almost without exception downgraded or slandered them socially as well as intellectually since the time of John Leland in the mid-eighteenth century. Even the foremost rationalists of the nineteenth century, Mark Pattison and Leslie Stephen (the latter produced the most complete and erudite history to date) are condescending. Rarely have the achievements of deism been acknowledged and appreciated, and then only in passing, in brief comments from specialized monographs, articles, and encyclopedia entries. No really satisfactory, complete, impartial, and scholarly account of the significance of the movement has as yet appeared.

Deism had somewhat different effects in different countries, depending on the different national cultural situations. By the close of the eighteenth century in England, it seemed, superficially at least, to have disappeared or gone underground. Yet in 1790, when Burke triumphantly asked, "Who born within the last forty years has read one word of Collins and Toland, and Tindal, and Chubb, and Morgan, and that whole race who called themselves Freethinkers? Who now reads Bolingbroke? Who ever read him through?" he was historically mistaken and premature in his inference. For in the nineteenth century, radical publishers such as William Benbow, William Hone, and, most notably, Richard Carlile (1790–1843), all of whom were political as well as religious reformers, flooded the popular market with periodicals (for example, *The Deist; or Moral Philosopher,* 1819–1820), pamphlets, and cheap reprints and excerpts from freethinkers of all ages, including the whole range of the British deists, the skeptical Hume, Voltaire and Rousseau of France, and Paine and Palmer of America. The campaign was continued by others throughout the nineteenth century and survives in the present century on a higher intellectual level by affiliations with Unitarianism, Fabian socialism, and rationalist and humanistic societies, among others.

In France, the true deism of Voltaire and Rousseau was overwhelmed by the atheism of most of the *philosophes,* a doctrine which inevitably contributed to the upheaval of the French Revolution. The course of these eighteenth-century developments may be said to be paralleled today, on the one hand, by widespread atheism and, on the other, by the militant anticlericism of even many of the devout. In Germany, early intellectual deism was followed by both the fideism of Friedrich Heinrich Jacobi and a new post-Humean variety of rationalism which began with Kant and the romanticists of the following century.

In America, deism was long submerged by evangelism among the semiliterate masses and by Unitarianism among the well educated. An aggressive antireligionism resurged in the 1870s with Robert Ingersoll, "the great agnostic," and a host of followers, such as William Brann in Texas in the 1890s with his world-famous newspaper *Brann's Iconoclast.* Today, rationalist and humanistic societies and Unitarianism are omnipresent.

With few exceptions, deists in all countries have been interested in political and social reform, and with the passage of time it has become virtually impossible to isolate the purely religious aspects. Deism remains a symptom of revolt against orthodoxy and dogmatism.

By way of summary and possible oversimplification, deism is the individual's affirmation of his right to think for himself on all subjects and to communicate his thoughts to others for the general welfare. It is the affirmation of the principle of the oneness of humanity. It marks the rise of secularism and the beginning of modernity in theology. In this sense it is still viable, and although freethinking today claims a philosophical substratum different from the simple rationalism of the seventeenth and eighteenth centuries, it is akin in spirit to historical deism. The early rise of deism in all countries

was strongly abetted by the growth of the spirit of toleration, and deism, in its turn, has strongly contributed to the continued growth and acceptance of toleration of other views. Perhaps, in the most universal sense, this is the major legacy of historical deism to the modern world.

Bibliography

USEFUL ANTHOLOGIES

Berlin, Isaiah, ed. *The Age of Enlightenment.* New York: New American Library, 1956.

Brinton, Crane, ed. *The Portable Age of Reason Reader.* New York: Viking Press, 1956.

Creed, John M., and John S. Boys Smith, eds. *Religious Thought in the Eighteenth Century.* Cambridge, U.K., 1934.

Fellows, Otis E., and Norman L. Torrey, eds. *The Age of Enlightenment.* New York: Crofts, 1942.

Gay, Peter, ed. *Deism: An Anthology.* Princeton, NJ: Van Nostrand, 1968.

Hampshire, Stuart, ed. *The Age of Reason: The 17th Century Philosophers.* Boston: Houghton Mifflin, 1956.

Torrey, Norman L., ed. *Les philosophes.* New York: Capricorn, 1960.

GENERAL HISTORIES

Betts, C. J. *Early Deism in France: From the So-Called "Déistes" of Lyon (1564) to Voltaire's "Lettres Philosophiques."* (1734). Boston: Martinus Nijhoff Publishers, 1984.

Byrne, Peter. *Natural Religion and the Nature of Religion: The Legacy of Deism.* London: Routledge, 1989.

Herrick, James. *The Radical Rhetoric of the English Deists: The Discourse of Skepticism, 1680–1750.* Columbia: University of South Carolina Press, 1997.

Lechler, Gotthard V. *Geschichte des englischen Deism.* Stuttgart, 1841.

Leland, John. *A View of the Principal Deistical Writers that have appeared in England in the last and present Century.* 3rd ed. 3 vols. London, 1754–1756. Early, voluminous, and vituperative apologetics.

Orr, John. *English Deism: Its Roots and Its Fruits.* Grand Rapids, MI: Eerdmans, 1934. Discursive, not too sound.

Sayous, Edouard. *Les déistes anglais et le christianisme, 1696–1738.* Paris, 1882.

Stephen, Leslie. *English Thought in the Eighteenth Century.* 2 vols. London, 1876; revised, 1902; paperback, New York, 1963. Vol. I is most scholarly study available but with a curious personal animus.

Walters, Kerry. *The American Deists: Voices of Reason and Dissent in the Early Republic.* Lawrence: University of Kansas Press, 1992.

GENERAL HISTORICAL BACKGROUND

Abbey, C. J., and J. H. Overton. *The English Church in the Eighteenth Century.* 2 vols. London, 1878.

Becker, Carl L. *The Heavenly City of the Eighteenth-Century Philosophers.* New Haven, CT: Yale University Press, 1932. Becker's brilliant paradox has dimmed over the years.

Bury, J. B. *A History of Freedom of Thought.* New York: Holt, 1913.

Cassirer, Ernst. *The Philosophy of the Enlightenment.* Princeton, NJ: Princeton University Press, 1951, and Boston, 1955.

Colie, R. L. *Light and Enlightenment: A Study of the Cambridge Platonists and the Dutch Arminians.* Cambridge, U.K.: Cambridge University Press, 1957.

Cragg, G. R. *From Puritanism to the Age of Reason.* Cambridge, U.K.: Cambridge University Press, 1950.

Farrar, A. S. *Critical History of Free Thought.* London: Murray, 1862.

Hall, Thomas C. *The Religious Background of American Culture.* Boston: Little Brown, 1930.

Havens, George R. *The Age of Ideas.* New York: Holt, 1955.

Hazard, Paul. *La crise de la conscience européenne.* 3 vols. Paris: Boivin, 1935.

Hazard, Paul. *La pensée européenne au XVIII^e siècle.* 3 vols. Paris, 1946. Both of Hazard's works are brilliantly comprehensive.

Humphreys, A. R. *The Augustan World.* London: Methuen, 1954. Useful survey.

Hunt, John. *Religious Thought in England from the Reformation to the End of the Last Century.* 3 vols. London: Strahan, 1870–1873.

Jordan, Wilbur K. *Development of Religious Toleration in England,* 4 vols. Cambridge, MA: Harvard University Press, 1932–1940. Highly important.

Lecky, W. E. M. *History of the Rise and Influence of the Spirit of Rationalism in Europe.* 2 vols. in one. London, 1910. Valuable.

Martin, Kingsley. *French Liberal Thought in the Eighteenth Century.* Boston, 1929.

M'Giffert, Arthur C. *Protestant Thought before Kant.* London, 1919.

Parrington, Vernon L. *Main Currents in American Thought.* 3 vols. New York: Harcourt Brace, 1927–1930.

Popkin, Richard, and Argo Vanderjagt, eds. *Scepticism and Irreligion in the Seventeenth and Eighteenth Centuries.* Leiden, Netherlands: Brill, 1993.

Robertson, J. M. *A History of Freethought in the Nineteenth Century.* 2 vols. London: Watts, 1929.

Robertson, J. M. *A Short History of Freethought Ancient and Modern.* 3rd ed. 2 vols. London, 1915. Both works are indispensable.

Smith, Preserved. *A History of Modern Culture.* 2 vols. New York, 1930–1934. Wide coverage.

Tulloch, John. *Rational Theology and Christian Philosophy in England in the Seventeenth Century.* 2 vols. Edinburgh: Blackwood, 1872.

Urwin, Kenneth. *A Century for Freedom.* London: Watts, 1946.

Willey, Basil. *The Eighteenth Century Background.* London: Chatto and Windus, 1940. Somewhat superficial.

SPECIALIZED STUDIES

Aldridge, Alfred O. "Shaftesbury and the Deist Manifesto." *Transactions of the American Philosophical Society,* n.s. 41 (2) (1951): 297–385.

Biddle, John. "Locke's Critique of Innate Principles and Toland's Deism." *Journal of the History of Ideas* 37 (1976): 411–422.

Boller, Paul F., Jr. *George Washington & Religion.* Dallas, TX: Southern Methodist University Press, 1963.

Bushell, Thomas. *The Sage of Salisbury: Thomas Chubb, 1679–1747.* New York: Philosophical Library, 1967.

Cole, G. D. H. *Richard Carlile.* London: Gollancz, 1943.

Courtines, Leo P. *Bayle's Relations with England and the English.* New York: Columbia University Press, 1938.

Cunliffe, Christopher, ed. *Joseph Butler's Moral and Religious Thought: Tercentenary Essays.* Oxford: Clarendon Press, 1992.

Daniel, Stephen. *John Toland: His Methods, Manners and Mind.* Montreal: McGill-Queen's University Press, 1984.

Fruchtman, Jack. *Thomas Paine and the Religion of Nature.* Baltimore: Johns Hopkins University Press, 1993.

Harkness, Douglas. *Bolingbroke: The Man and his Career.* London: Staples Press, 1957.

Hefelbower, S. G. *The Relation of John Locke to English Deism.* Chicago: University of Chicago Press, 1918.

Koch, G. Adolf. *Republican Religion: The American Revolution and the Cult of Reason.* New York: Holt, 1933. Indispensable.

Lovejoy, A. O. "The Parallel of Deism and Classicism." *Modern Philology* 29 (1932): 281–299. Important.

Luke, Hugh J., Jr. *Drams for the Vulgar: A Study of Some Radical Publishers and Publications of Early Nineteenth-Century London.* Unpublished PhD dissertation. University of Texas, 1963.

Merrill, Walter. *From Statesman to Philosopher: A Study in Bolingbroke's Deism.* New York: Philosophical Library, 1949.

Morais, Herbert M. *Deism in Eighteenth Century America.* New York: Columbia University Press, 1934.

Mossner, Ernest C. *Bishop Butler and the Age of Reason.* New York: Macmillan, 1936.

Noack, Ludwig. *Die Freidenker in der Religion.* 3 vols. Bern, 1853–1855.

Pattison, Mark. "Tendencies of Religious Thought in England." In *Essays and Reviews.* London, 1860.

Popkin, Richard H. "Scepticism in the Enlightenment." *Studies on Voltaire and the Eighteenth Century,* Geneva, xxiv/xxvii (1963): 1321–1345.

Redwood, J. "Charles Blount (1654–93), Deism, and English Free Thought." *Journal of the History of Ideas* 35 (1974): 490–498.

Russell, Bertrand. *Why I Am Not a Christian and Other Essays on Religion and Related Subjects,* edited by Paul Edwards. London: Allen and Unwin, 1957.

Salvatorelli, Luigi. *From Locke to Reitzenstein: The Historical Investigation of the Origins of Christianity.* Cambridge, MA, 1930.

Stromberg, Roland N. *Religious Liberalism in Eighteenth-Century England.* London: Oxford University Press, 1954. Sound survey.

Sullivan, Robert E. *John Toland and the Deist Controversy: A Study in Adaptations.* Cambridge, MA: Harvard University Press, 1982.

Tennant, F. R. *Miracle & Its Philosophical Presuppositions.* Cambridge, U.K.: Cambridge University Press, 1925.

Torrey, Norman L. *Voltaire and the English Deists.* New Haven, CT: Yale University Press, 1930. Indispensable.

Webb, Clement C. J. *Studies in the History of Natural Theology.* Oxford: Clarendon Press, 1915.

Winnett, A. R. "Were the Deists 'Deists'?" *Church Quarterly Review* 161 (1960): 70–77. Makes distinction between philosophical and historical deists.

Yolton, John W. *John Locke and the Way of Ideas.* London: Oxford University Press, 1956.

Ernest Campbell Mossner (1967)
Bibliography updated by Christian B. Miller (2005)

DELEUZE, GILLES
(1925–1995)

Gilles Deleuze, one of the most influential and prolific French philosophers of the postwar period, was born in Paris, and lived there, with a few exceptions, for the rest of his life. The son of a conservative, middle-class engineer, a veteran of World War I, Deleuze received his early elementary education in the French public school system. When the Germans invaded France, Deleuze's family was on vacation in Normandy, and he spent a year being schooled there. Deleuze traced his own initiation into literature and philosophy to his encounter with a teacher at Deauville named Pierre Halwachs (son of the sociologist Maurice Halwachs), who introduced him to writers such as André Gide and Charles Baudelaire. Early on, he later recalled, philosophical concepts struck him with the same force as literary characters, having their own autonomy and style, and he soon began to read philosophical works with the same animation and engagement as literary texts. During the occupation, Deleuze's older brother was arrested by the Nazis for resistance activities and deported; he died on the train to Auschwitz.

After the Liberation, Deleuze returned to Paris and undertook his khâgne (an intensive year of preparatory studies) at the prestigious Lycée Henri IV, and then studied the history of philosophy at the Sorbonne. He was taught by Jean Hippolyte and Ferdinand Alquié ("two professors I loved and admired enormously" [Deleuze, 1977, p. 12]), as well as Georges Canguilheim and Maurice de Gandillac, though like many of his peers he was as influenced by the writings of Jean-Paul Sartre as by the work of his academic mentors. He published his first book, *Empiricism and Subjectivity*, on David Hume, in 1953, when he was twenty-eight. In an era dominated by phenomenology and "the three Hs" (Hegel, Husserl, Heidegger), Deleuze's decision to write on empiricism and Hume was already a provocation, early evidence of the heterodox tendencies of his thought.

During the decade between 1953 and 1962—which he later referred to as "a hole in my life" (Deleuze 1990, p. 138)—Deleuze published little, moved among various

teaching positions in Paris and the provinces, and contracted a recurring respiratory ailment that would plague him for the rest of his life. In 1956 he married Fanny (Denise Paul) Grandjouan, a French translator of D. H. Lawrence, with whom he would have two children. In 1962 his groundbreaking study *Nietzsche and Philosophy* was published to considerable acclaim, cementing Deleuze's reputation in academic circles. In the decade that followed, Deleuze more or less published a book per year, most of them devoted to the work of a particular philosopher or writer: Kant (1963), Proust (1964), Nietzsche (1965), Bergson (1966), Sade and Masoch (1967), Spinoza (1968), and later Kafka (1975), Francis Bacon (1981), Michel Foucault (1986), and Leibniz (1988). *Difference and Repetition*, his magnum opus, appeared in 1968, followed by *Logic of Sense* in 1969.

In the same year, he met Félix Guattari, a militant psychoanalyst, with whom he wrote a number of influential books, including the two volumes of *Capitalism and Schizophrenia* (1972, 1980), which were overtly political texts written in the wake of the ferment of May 1968. The first volume, *Anti-Oedipus*, was a best-seller in France, and thrust Deleuze into the limelight as a public intellectual. In 1969 Deleuze took up a teaching post at the experimental campus of the University of Paris VII (at Vincennes and, later, St. Denis), where he gave weekly seminars until his retirement in 1987. Like Kant, he traveled little, and devoted his time to teaching and writing: Paris was his Konigsberg, France was his Prussia. He shunned academic conferences and colloquia, insisting that the activity of thought took place primarily in writing, and not in dialogue and discussion. By 1993 his pulmonary illness had confined him severely, making it increasingly difficult to read or write; he took his own life on November 4, 1995.

Deleuze's writings were strongly grounded in the history of philosophy, but he read widely in contemporary science and mathematics, and was well known for his interactions with the various arts. His early work was in part a reaction against Hegel, and more generally against the then-dominant post-Kantian tradition in philosophy. Kant's genius, for Deleuze, was to have conceived of a purely *immanent* critique of reason—a critique that did not seek, within reason, "errors" produced by external causes, but rather "illusions" that arise internally from within reason itself by the illegitimate (transcendent) uses of the syntheses of consciousness. Deleuze characterized his own work as a philosophy of immanence, but argued that Kant himself had failed to fully realize the immanent ambitions of his critique, for at least two reasons.

First, Kant made the immanent field immanent *to* a transcendental subject, thereby reintroducing an element of transcendence, and reserving all power of synthesis to the activity of the subject. In his first book, *Empiricism and Subjectivity* (1953), on Hume, Deleuze pointed to an empiricist reversal of this relation: whereas Kant's question had been "How can the given be given to a subject?" Hume's question had been "How is the subject (human nature) constituted within the given?" Deleuze would later characterize his own position as a "transcendental empiricism": the determination of an impersonal and pre-individual transcendental field in which the subject is itself the result or product of *passive syntheses* (of the body, habit, desire, the unconscious). Just as there is no universal reason but only historically variable processes of "rationalization" (Max Weber), so there is no universal or transcendental subject, but only diverse and historically variable processes of "subjectivation." Deleuze summarized his empiricism in terms of two characteristics: the abstract does not explain, but must itself be explained; the aim of philosophy is not to rediscover the eternal or the universal, but to find the singular conditions under which something new is produced (creativity).

Second, Kant had simply presumed the existence of certain "facts" (knowledge, morality) and then sought their conditions of possibility in the transcendental. But already in 1789, Salomon Maimon, whose early critiques of Kant helped generate the post-Kantian tradition, had argued that Kant's critical project required a method of *genesis*—and not merely a method of conditioning—that would account for the production of knowledge, morality, and indeed reason itself—a method, in other words, that would be able to reach the conditions of *real* and not merely *possible* experience. Maimon found a solution to this problem in a principle of difference: Whereas identity is the condition of possibility of thought in general, it is *difference* that constitutes the genetic and productive principle of real thought.

These two Maimonian exigencies—the search for *the genetic conditions of real experience* and *the positing of a principle of difference*—reappear like a leitmotif in almost every one of Deleuze's early monographs. *Nietzsche and Philosophy* (1962), for instance, suggests that Nietzsche completed and inverted Kantianism by bringing critique to bear, not simply on false claims to knowledge or morality, but on true knowledge and true morality, and indeed on truth itself: "genealogy" constituted Nietzsche's genetic method, and the will to power was his principle of difference. *Bergsonism* (1966) argues that Bergson's con-

cepts of duration, memory, and *élan vital* constitute the dimensions of the multiplicities of the real. Against the "major" post-Kantian tradition of Fichte, Schelling, and Hegel, Deleuze in effect posited his own "minor" post-Kantian trio of Maimon, Nietzsche, and Bergson. In rethinking the post-Kantian heritage, Deleuze would also retrieve the work of a well-known trio of pre-Kantian philosophers—Hume, Spinoza, and Leibniz—although from a decidedly post-Kantian viewpoint.

Deleuze's historical monographs were, in this sense, preliminary sketches for the great canvas of *Difference and Repetition* (1968), which marshaled these resources from the history of philosophy in an ambitious project to construct a metaphysics of difference. Normally, difference is conceived of as an empirical relation between two terms each of which has a prior identity of its own ("x is different from y"). In Deleuze, this primacy is inverted: identity persists, but it is now a secondary principle produced by a prior relation between differentials (*dx* rather than not-x). Difference is no longer an empirical relation but becomes a *transcendental* principle that constitutes the sufficient reason of empirical diversity as such (for example, it is the difference of potential in a cloud that constitutes the sufficient reason of the phenomenon of lightning). In Deleuze's ontology, the different is related to the different through difference itself, without any mediation. Although he was indebted to metaphysical thinkers such as Spinoza, Leibniz, and Bergson, Deleuze appropriated their respective systems of thought only by pushing them to their "differential" limit, purging them of the three great terminal points of traditional metaphysics (God, World, Self).

Deleuze's subsequent work was, to some degree, a working out of the metaphysics developed in *Difference and Repetition*. Deleuze considered himself a classical philosopher and conceived of his philosophy as a system—albeit an open and *heterogenetic* (non-totalizing) system—which might be summarized in terms of the following traditional rubrics, derived largely from Kant.

DIALECTICS (THEORY OF THE IDEA)

Difference and Repetition attempts to formulate a theory of Ideas (dialectics) based neither on an essential model of identity (Plato), nor a regulative model of unity (Kant), nor a dialectical model of contradiction (Hegel), but rather on a problematic and genetic model of difference. Ideas are what define the "essence" of a thing, but one cannot attain an Idea through the Socratic question "What is … ?" (which posits Ideas as transcendent and eternal), but rather through "minor" questions such as "Which one?" "Where?" "When?" "How?" "How many?" "In which case?" "From which viewpoint?"—all of which allow one to define the spatio-temporal coordinates of Ideas that are purely immanent and differential. The formal criteria Deleuze uses to define Ideas are largely derived from Leibniz and the model of the differential calculus, which provides a mathematical symbolism for the exploration of the real: things or beings are virtual and problematic multiplicities composed of singularities-events, which are prolonged in converging and diverging series, forming zones of indiscernibility where the multiplicities entering into perpetual becomings.

AESTHETICS (THEORY OF SENSATION)

What are the implications of a principle of difference for aesthetics? Kant had dissociated aesthetics into two halves: the theory of sensibility as the form of possible experience (the "Transcendental Aesthetic"), and the theory of art as a reflection on real experience (the "Critique of Aesthetic Judgment"). In Deleuze's work, these two halves of aesthetics are reunited: If the most general aim of art is to "produce a sensation," then the genetic principles of sensation are at the same time the principles of composition for works of art; conversely, it is works of art that are best capable of revealing these conditions of sensibility. Deleuze's writings on the various arts—including the cinema (*Cinema* I and II), literature (*Essays Critical and Clinical*), and painting (*Francis Bacon: The Logic of Sensation*)—must be read, not as works of criticism, but rather as philosophical explorations of this transcendental domain of sensibility. Deleuze locates the conditions of sensibility in an *intensive* conception of space and a *virtual* conception of time, which are necessarily actualized in a plurality of spaces and a complex rhythm of times (for instance, in the nonextended spaces and nonlinear times of modern mathematics and physics).

ETHICS (THEORY OF AFFECTIVITY)

Deleuze has similarly developed a purely immanent conception of ethics, an "ethics without morality." If morality implies an appeal to transcendent values as criteria of judgment (as in Kant's moral law), ethics evaluates actions and intentions according to the immanent mode of existence they imply. One says or does this, thinks or feels that: What mode of existence does it imply? This is the link Deleuze establishes between Spinoza and Nietzsche, his two great precursors as philosophers of immanence: each of them argued, in his own manner, that there are things one cannot do or think except on the condition of being base or enslaved, unless one harbors a

ressentiment against life (Nietzsche), unless one remains the slave of passive affections (Spinoza); and there are other things one cannot do or say except on the condition of being noble or free, unless one affirms life or attains active affections. The transcendent moral opposition (Good/Evil) is in this way replaced by an immanent ethical difference (good/bad). A bad or sickly life is an exhausted and degenerating mode of existence, one that judges life from the perspective of its sickness, which devalues life in the name of higher values. A good or healthy life, by contrast, is an overflowing or ascending mode of existence, capable of transforming itself depending on the forces it encounters, always opening up new possibilities of life, new becomings.

POLITICS (SOCIO-POLITICAL THEORY)

This immanent conception of ethics leads directly into Deleuze's political philosophy, which he developed most fully in the two volumes of *Capitalism and Schizophrenia*, with Félix Guattari. *Anti-Oedipus* (1972), under the guise of a critique of psychoanalysis, is in effect an immanent reworking of Kant's theory of desire in the *Critique of Practical Reason*. Since the capacities and affectivity (desire) of individuals is always effectuated within concrete socio-political "assemblages"—one of Deleuze's fundamental political concepts—the political philosophy presented in *A Thousand Plateaus* (1980) takes the form of a typology of social assemblages (primitive societies, the State, nomadic war machines, capitalism) that provide conceptual tools for analyzing the complex dimension of the actual situation: How are its mechanisms of power organized? What are the "lines of flight" that escape its integration? What new modes of existence does it make possible? What relations does it sustain between desire and power?

ANALYTICS (THEORY OF THE CONCEPT)

Finally, Deleuze's dialectic (the constitution of problems) leads directly into his analytic (concepts as cases of solution), which he presented in his late book *What Is Philosophy?* (1991, co-authored with Guattari). Deleuze defines philosophy as the art of creating concepts, as knowledge through pure concepts. But for Deleuze, the highest concepts are not a priori universals applicable to objects of possible experience (categories), but singularities that correspond to the structures of real experience. Concepts are self-referential—they posit their object in being posited—and are defined in terms of their consistency of their components (endo-consistency) and their relation to other concepts (exo-consistency). Deleuze's analytic should be evaluated critically in relation to competing theories of the concept (Frege, Russell), which often make use of scientific functions or logical propositions as their model. His analysis of the concepts of "sadism" and "masochism" in his 1967 book *Coldness and Cruelty* (and his concomitant critique of the notion of "sadomasochism") provides an excellent case study of his differential approach to concepts.

See also Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Nietzsche, Friedrich.

Bibliography

WORKS BY GILLES DELEUZE

Empiricism and Subjectivity: An Essay on Hume's Theory of Human Nature (1953). Translated by Constantin V. Boundas. New York: Columbia University Press, 1991.

Nietzsche and Philosophy (1962). Translated by Hugh Tomlinson. New York: Columbia University Press, 1983.

Kant's Critical Philosophy (1963). Translated by Hugh Tomlinson and Barbara Habberjam. Minneapolis: University of Minnesota Press, 1984.

Proust and Signs (1964). Translated by Richard Howard. London: Athlone Press, 2000.

Bergsonism (1966). Translated by Hugh Tomlinson and Barbara Habberjam. Minneapolis: University of Minnesota Press, 1988.

Masochism: Coldness and Cruelty (1967). Translated by Jean McNeil. New York: Zone Books, 1989.

Difference and Repetition (1968). Translated by Paul Patton. New York: Columbia University Press, 1994.

Expressionism in Philosophy: Spinoza (1968). Translated by Martin Joughin. New York: Zone Books, 1990.

The Logic of Sense (1969). Translated by Mark Lester, with Charles Stivale; edited by Constantin V. Boundas. New York: Columbia University Press, 1990.

Capitalism and Schizophrenia. With Félix Guattari. Vol. 1: *Anti-Oedipus* (1972). Translated by Robert Hurley, Mark Seem, and Helen Lane. Vol. 2: *A Thousand Plateaus* (1980). Translated by Brian Massumi. Minneapolis: University of Minnesota Press, 1983, 1987.

Kafka: Toward a Minor Literature (1975). With Félix Guattari. Translated by Sean Hand. Minneapolis: University of Minnesota Press, 1986.

Dialogues (1977). With Claire Parnet. Translated by Hugh Tomlinson and Barbara Habberjam. New York: Columbia University Press, 1987.

Francis Bacon: The Logic of Sensation (1981). Translated by Daniel W. Smith. London: Athlone Press, 2003.

Cinema. Vol. 1: *The Movement-Image* (1983). Vol. 2: *The Time-Image* (1985). Translated by Hugh Tomlinson, Barbara Habberjam, and Robert Galeta. Minneapolis: University of Minnesota Press, 1986, 1989.

Foucault (1986). Translated by Sean Hand. Minneapolis: University of Minnesota Press, 1986.

The Fold: Leibniz and the Baroque (1988). Translated by Tom Conley. Minneapolis: University of Minnesota Press, 1988.

Negotiations, 1972–1990 (1990). Translated by Martin Joughin. New York: Columbia University Press, 1995.

What Is Philosophy? (1991). With Félix Guattari. Translated by Hugh Tomlinson and Graham Burchell. New York: Columbia University Press, 1991.

Essays Critical and Clinical. (1993). Translated by Daniel W. Smith and Michael A. Greco. Minneapolis: University of Minnesota Press, 1997.

Daniel W. Smith (2005)

DEL VECCHIO, GIORGIO

(1878–1970)

Giorgio Del Vecchio, the Italian legal philosopher, was born in Bologna, the son of the economist Giulio Salvatore Del Vecchio. He studied in Italy and Germany and taught in Ferrara, Sassari, Messina, Bologna, and at Rome, where he was a professor from 1920, rector of the university from 1925 to 1927, and dean of the faculty of law from 1930 to 1938. He was dismissed by the fascists in 1938 because of his Jewish background. He resumed teaching in 1944 but was dismissed again in 1945, this time as a former fascist; he taught again from 1947 to 1953. He was named professor emeritus in 1955. Del Vecchio founded the *Rivista internazionale di filosofia del diritto* in 1921 and was its editor; he founded the Istituto di Filosofia del Diritto of the University of Rome in 1933 and the Società Italiana di Filosofia del Diritto in 1936.

Del Vecchio was influential in turning Italian legal thought from nineteenth-century positivism. His own position has been described as neo-Kantian idealism and as humanist ethical idealism. According to Del Vecchio, the thinking subject is necessarily conscious of the *other,* not merely as object, but as itself a subject. Hence, mutual recognition and respect are necessary, and it is possible to deduce for the mutual relations of subjects not merely a logical form but also an ideal content of justice based on respect for personality. Law is the objective coordination of possible actions between subjects according to an ethical principle, which in its highest expression is the principle of justice. Psychologically, the idea of justice is a necessary aspect of consciousness, found in rudimentary form even among animals. Historically, the idea has been realized with varying degrees of positivity in human societies, and continual effort is needed to realize it in the changing specific conditions of life. There are instances of "involution" (regression), but history on the whole shows a progressive evolution toward the understanding and realization of justice. These main ideas, stated in Del Vecchio's early writings, were developed with a wealth of historical learning in his *Lezioni di filosofia del diritto* and *La giustizia*; in other writings he applied them to particular problems of legal and political philosophy.

Del Vecchio, like other veterans of World War I, joined the fascist movement when it arose because he saw it as a defense against Bolshevism, and it is unjust to consider him a representative of fascist philosophy. For a time he did hope, mistakenly, that the fascist "strong state" might realize the "ethical state" that, by harmonizing individual freedoms, would enhance individual personality. Throughout the fascist period, however, Del Vecchio's fundamental teaching was unchanged; and he continued to assert the validity of natural law and to defend individual freedom against the statolatry of official fascist doctrine.

See also Continental Philosophy; Idealism; Justice; Legal Positivism; Political Philosophy, History of; Positivism.

Bibliography

PRIMARY WORKS

Presupposti, concetto e principio del diritto (The formal bases of law). Translated by John Lisle. New York: Macmillan, 1921.

La giustizia, 3rd ed. Rome: Editrice Studium, 1946.

Lezioni di filosofia del diritto, 6th ed. Milan: Giuffrè, 1948.

Sui principi generali del diritto (General principles of law). Translated by Felix Forte. Boston: Boston University Press, 1956.

Studi sullo stato. Milan, 1958.

Studi su la guerra e la pace. Milan, 1959.

Contributi alla storia del pensiero giuridico e filosofico. Milan: Giuffrè, 1963.

Humanité et unité de droit. Paris, 1963.

Man and Nature; Selected Essays. Edited by Ralph A. Newman; translated by A. H. Campbell. Notre Dame, IN: University of Notre Dame Press, 1969.

SECONDARY WORKS

Orecchia, Rinaldo. *Bibliografia di Giorgio Del Vecchio,* 2nd ed. Bologna, 1949.

Vela, Luis. *El Derecho Natural en Giorgio Del Vecchio.* Rome: Libreria editrice dell'Università Gregoriana, 1965.

Vidal, Enrico. *La filosofia giuridica di Giorgio Del Vecchio.* Milan: Giuffrè, 1951.

A. H. Campbell (1967)
Bibliography updated by Philip Reed (2005)

DEMIURGE

Demiurge, an anglicized form of *δημιουργός,* the ordinary Greek word for a workman, craftsman, or artificer, is

commonly used in Greek literature from Homer onward. In Homer it is applied to heralds, soothsayers, and physicians as well as to manual workers; but in later Greek it primarily means a craftsman or maker, such as a carpenter or a smith. Its importance in the history of philosophy derives almost entirely from Plato's *Timaeus*, in which a Demiurge, or Craftsman, is represented as ordering and arranging the physical world and bringing it as far as possible into conformity with the best and most rational pattern. In two other places (*Republic* 530A and *Sophist* 265C) Plato uses the word *δημιουργος*, or the corresponding verb, in connection with divine creation; and it occurs in one passage in Xenophon's Socratic discourses (*Memorabilia* 1.4.9), but these are all casual and isolated references. For our understanding of Plato's conception of creation we must rely almost exclusively on the *Timaeus*.

The *Timaeus* is, in fact, Plato's only substantial essay in physical theory and cosmology. There is disagreement about the date of the dialogue and about its place in the chronological order of Plato's writings; but it is generally agreed to be later than the great group of middle dialogues, from the *Phaedo* and *Symposium* to the *Republic* and *Phaedrus,* in which Plato expounds his most characteristic metaphysical and ontological doctrines. The substance of these doctrines is repeated and underlined in the *Timaeus* itself, which makes a sharp division between the eternal, transcendent, intelligible, unchanging world of true being or reality and the temporal, phenomenal, sensible, unstable world of mere becoming. It was this very contrast between the world of Forms and the world of sense that had led Plato to neglect physical research and speculation; and when he does turn to this subject in the *Timaeus,* he repeatedly insists that even his own best efforts in this field cannot produce more than an *εἰκὼς μυθός*—a "likely tale"—falling far short of the certainty and exactness that can be sought in mathematics and pure philosophy. He speaks of the whole doctrine of the *Timaeus* in the provisional, tentative manner in which he presents the eschatological myths of the *Gorgias, Phaedo, Republic,* and *Phaedrus.*

Against this background it may appear surprising that Plato ventured on these topics at all. His motives become plainer if we remember his own comments in the *Phaedo* (97C–99D) on the cosmology of Anaxagoras. Socrates first praises Anaxagoras for holding that *νοῦς*—Intelligence or Reason—ordered and arranged the world, imposing a rational plan on a preexisting chaos. He then complains that Anaxagoras did not pursue this line of thought to its proper conclusion: He uses Reason as a mere deus ex machina to explain the origin of the cosmic process as a whole but does not give detailed teleological explanations of particular things and events, showing that everything is arranged for the best. Anaxagoras resorts instead to the purely physical explanations that had been used by his Ionian predecessors, which is like trying to explain why Socrates does not escape from prison wholly in terms of bones and sinews, without reference to intelligence, intention, motive, and morality. Aristotle makes a similar comment in *Metaphysics* I,3: Anaxagoras stands out among his contemporaries and predecessors "like a sober man among drunkards," but he does not make proper use of his concept of cosmic *νοῦς*.

The *Timaeus* is Plato's attempt to carry out the program of rationalist cosmology that Anaxagoras had promised but had failed to fulfill. The Demiurge is portrayed as the agent who turns the initial chaos into a cosmos. Like a human craftsman, he arranges existing materials and does not create them. The conception of creation ex nihilo is foreign to the whole tradition of Greek thought. The Demiurge shapes his materials to conform as much as possible to the eternal intelligible model of the Forms. First, he makes other gods, the world soul that the cosmos requires as its motive principle, and the immortal part of the human soul. The created gods then complete the work by making physical things, including human bodies. The Demiurge's success is necessarily limited: the Reason that constitutes his pattern is opposed by a recalcitrant Necessity (*ἀνάγκη*) that hinders his work in something like the way in which a human craftsman may be frustrated by intractable materials—and no material is perfectly tractable. This obstacle to a faultless achievement by the Demiurge is also the main reason why Plato cannot hope to give more than a "likely tale" of the Demiurge's work.

It has been widely believed, from ancient times to the present day, that the Demiurge is a mythical figure and that Plato did not believe in the literal existence of such a creator-god. He is a personification of the Reason whose requirements he is represented as trying to embody in the nature of the cosmos. Even if he is literally meant, he must still be sharply contrasted with the creator-god of the Judeo-Christian tradition, not only because he is not in that sense a creator, but also because he is in no sense an object of worship.

It is more difficult to decide whether the process of creation is also mythical; whether Plato believed that the imposition of order on the physical world was a definite event that took place at some time in the past, or whether the narrative of the *Timaeus* is a presentation in chronological form of Plato's views about the relative value and

ontological priority of the various elements in the universe. According to this latter view, the story that bodies were created after souls would be a pictorial way of marking the inferiority of the body to the soul. Aristotle reports (*De Caelo* 279b33) that this was the tradition in Plato's Academy. The chronological picture is said to be used only for purposes of exposition, like a figure in geometry. Aristotle himself took the chronology literally, and he was followed in this by Plutarch; but the ancient authorities were nearly all on the other side.

Most modern scholars have disagreed with Aristotle, but he has had some notable supporters; and the question is still being debated. In support of the usual interpretation one may quote the parallel case of the *Republic,* where the building and dissolution of the ideal community is a pictorial means of presenting a logical analysis in chronological terms. Defenders of the opposite view point out that the word γέγουευ ("it came into being") gives an emphatic answer to the crucial question "Has the cosmos always been, or has it come to be, starting from some beginning?" (28B). However, the imagery of the *Republic* is equally emphatic. Once a man has chosen to represent one thing by painting a picture of another, the fact that he uses firm brush strokes and bright colors does not destroy its claims to be a picture.

The concept of the Demiurge was taken over by the Neoplatonists and by some Gnostic writers. To the Gnostics he was the evil lord of the lower powers, creator of the despised material world, and entirely separate from the supreme God. Their parody of the Demiurge as a clumsy imitator is blended with hostile satire of the Old Testament creator-God. Plotinus protested against their conception of the Demiurge as a source of positive evil in the world.

There is no clear case of any notable modern thinker whose teaching has been closely or directly influenced by the concept of the Demiurge, although there are hints of a similar idea in J. S. Mill's essay "Theism," where the word *Demiurgos* is applied to a God whose creative power is limited by the nature of his materials.

See also Anaxagoras of Clazomenae; Aristotle; Gnosticism; Greek Academy; Homer; Mill, John Stuart; Neoplatonism; Plato; Xenophon.

Bibliography

Archer-Hind, R. D. *The Timaeus of Plato.* London and New York, 1888. Text, translation, introduction, and notes.

Bury, R. G. *Plato, Timaeus, Critias, Clitopho, Menexenus, Epistulae.* Loeb Classical Library. London and New York, 1929. Text and translation.

Cornford, Francis M. *Plato's Cosmology.* London: K. Paul, Trench, Trubner, 1937. Translation of the *Timaeus,* with a running commentary.

Crombie, I. M. *An Examination of Plato's Doctrines.* London and New York: Humanities Press, 1963. Vol. II, *Plato on Knowledge and Reality,* Ch. 2.

Grube, G. M. A. *Plato's Thought.* London: Methuen, 1935. Ch. 5.

Hackforth, R. "Plato's Cosmogony." *Classical Quarterly,* n.s., 9 (1959): 17–22.

Taylor, A. E. *A Commentary on Plato's Timaeus.* Oxford: Clarendon Press, 1928. Prolegomena and notes.

Renford Bambrough (1967)

DEMOCRACY

Democracy is difficult to define, not only because it is vague, like so many political terms, but more importantly, because what one person would regard as a paradigm case another would deny was a democracy at all. The word has acquired a high emotive charge in the last hundred years; it has become good tactics to apply it to one's own favored type of regime and to deny it to rivals. The most diverse systems have been claimed as democracies of one sort or another, and the word has been competitively redefined, to match changes in extension by appropriate changes in intention. However, there is still this much agreement: Democracy consists in "government by the people" or "popular self-government." As such, it would still be universally distinguished from, say, a despotism that made no pretense of popular participation—the despotism of Genghis Khan or of Louis XIV, for instance—or from a theocracy, like the Vatican. There remains plenty of room for disagreement, however, about the conditions under which the people can properly be said to rule itself.

In the first place, what is "the people"? In ancient Greece, the *demos* was the poorer people; democracy meant rule of the poor over the rich. This is still the usage of those who identify the people with the proletariat and democracy with the rule of the working class. The word *people,* however, is often used to differentiate the subject mass from the ruling elite, as, for instance, when John Locke speaks of a tyrannical government putting itself into a state of war with the people. In this sense, "the people" necessarily means the ruled. Can the people, however, be said *to rule itself* in the same sense as it is said to be ruled by monarchs, oligarchs, and priests? To rule is,

generally, to prescribe conduct for someone else. There is a sense, it is true, in which moralists speak of ruling oneself, when by a kind of metaphor they speak of reason governing the passions. Again, a former colony becomes self-governing when its people is no longer ruled by outsiders; but this is not inconsistent with its still being ruled by *native* masters.

The usual paradigm of a people governing itself is the direct democracy of ancient Athens. Admittedly, citizenship was a hereditary privilege, excluding slaves and metics, and it is very doubtful whether, without this limitation, the citizen body would have been small enough for it to have operated as it did. Aside from this, however, the Athenian people governed itself in the sense that every individual could participate personally in policy decisions by discussion and voting, in a face-to-face situation. Athenian procedures are held to have been democratic in the sense that everyone was supposed to have an equal opportunity to state a case and influence decisions, even if, in some cases, individuals had ultimately to accept decisions that they had previously resisted. So today, in a similar sense, if a school or a department is said to be democratically run, we should expect its head to consult his staff on important issues and to concur in decisions to which he himself is opposed when the weight of opinion is against him. Self-government for a small group consists in general participation in the deliberative process, in which each person's voice carries a weight appropriate not to his status but to the merits, in the judgment of others, of what he has to say. If, despite continuing disagreement, a decision is essential, then it must be arrived at by majority vote. For it is not consistent with equal participation in decision making for any one individual to be privileged to say in advance that regardless of the distribution of opinions, his own or that of his group must prevail. That privilege excluded, decisions may be reached by lot or by vote; and if by vote, the opinion of either the lesser or the greater number may prevail. Deciding by lot was in fact used in Athens to fill certain public offices; it is a way of giving everyone an equal chance where advantages or privileges cannot be equally and simultaneously enjoyed; but to decide policy by lot would make nonsense of the procedure of public discussion, which is as integral to the democratic process as the idea of equality. The same would apply to a rule whereby whatever opinion received the fewest votes would prevail; for what point would there be in persuasion if it had no effect on the outcome or, still worse, if it actually reduced the chance of one's view being implemented? If a democratic decision is thought of, then, as the result of a fair confrontation of opinions, it must, at best, be generally agreed upon, and at worst, agreed upon by the majority.

CONDITIONS OF POLITICAL DEMOCRACY

Obviously, the conditions of face-to-face democracy, with direct participation, cannot be fulfilled within the political structure of modern states, both because of the size of their populations and because of the specialized knowledge needed to govern them. So although everyone may agree on what makes a small group democratic, when it comes to applying the concept to mass organizations, there is plenty of room for different interpretations of the principles to be applied and of the way to realize them under these very different conditions. Democracy now becomes representative government, that is, government by persons whom the people elect and thereby authorize to govern them.

Election and *representation* are themselves complex notions, however. In one sense, to be representative of a group may mean no more than to possess salient characteristics common to and distinctive of most of its members. In another, quasi-legal sense, one person may be said to represent another if, according to some code of rules, the consequences attached to an act of the representative are precisely those that would be attached to the act had it been performed by the principal himself; the representative can, in this case, *commit* the represented. In yet a third sense, one may represent another by looking after his interests, with or without his authorization (for example, the representation of infants in law). Now, democratic representation need not imply representation in the first sense, that of resemblance. Since an elected member of a legislature is taken to represent those who voted against as much as those who voted for him, he need not resemble those he represents, even in his opinions. Nor does he commit them as if they themselves had acted; the fact of their having legal duties does not depend on the fiction that, if their representative votes for a law, they have personally agreed to it. Their legal duties remain even if their representative voted against it. Nor must we necessarily accept moral responsibility for what is done by those who politically represent us, for in voting against them, we may have done the only thing open to us to disavow them.

Political representation is closer to the third sense of the term—the representation of interests; a democratic representative is usually thought to have the duty to watch over either the interests of his constituents or, as a member of an assembly representing the whole people,

the interests of the people at large. Nevertheless, he could still represent the interests of a group of people without their having had any part in choosing him. Some members of colonial legislatures in Africa used to be nominated by the governor to represent the interests of the unenfranchised native population. Precisely analogous, from the standpoint of the liberal democrat, is the case of a single-party system, where the ruling party invites the electors to endorse the candidate it has chosen to represent them. No matter how zealously the representative watched their interests, this would not count as democratic representation, precisely because the electors had had no part in selecting him. This view of democracy, therefore, is not compatible with tutelage; it implies the possibility not only of rejecting but also of freely proposing candidates, if none put forward by others is acceptable. Choosing and rejecting representatives is, indeed, the central act of participation by the citizens of a mass democracy, from which any effectiveness that they might have in other respects derives.

Closely related to election is the notion of the *responsibility* of the democratic representative. This means, in practice, that representatives must submit themselves periodically for reelection and, as a corollary, that they must be prepared to justify their actions and to attend to the experience and needs of their constituents, whose good will they must retain so long as they wish to remain in office.

DEMOCRACY AND POPULAR SOVEREIGNTY

It is often said that in a democracy the people's will is sovereign. But can the people be said to have a will? Opinions are divided on most things; there may be ignorance and apathy; on many questions only sectionally interested groups may have any clear opinions at all. Small groups, like committees, may reach agreed policies to which everyone feels committed; or in time of grave national danger, whole nations may discover a collective devotion to a single objective, overriding all conflicts of interests. However, although it might be intelligible to speak of a collective will in such cases, they are too limited or too rare to provide a framework for a general theory of democratic government. Such cases apart, one may speak of action, will, or decision in relation to collectivities only if their collective acts can be identified by some more or less formal procedure or if there are rules authorizing some identifiable individual to act *in the name of* the whole group. Thus, "Parliament has decided …" presupposes rules determining who are members of Parliament, defining their roles, and giving their several actions a collective significance and validity as "legislation." Are there analogous procedures, by virtue of which the people can be said to act or to express a will? Only by voting and by applying the majority principle in elections and referenda. And of course, applied to any particular collection of individual votes, different systems of voting or different arrangements of constituency boundaries can yield quite different results, each in its own rule context expressing "the people's will." Nevertheless, some people consider a system democratic to the extent that it approximates to government by referendum, though they would agree that this could not work as a day-to-day procedure. The doctrine that a government ought not to initiate policy changes without putting them to a vote in a general election (or, in a stronger form, that having done so, it is entitled—or obliged—to implement them forthwith) is a practical application of the popular-sovereignty view of democracy. A possible corollary sometimes derived from this last view is that it is undemocratic to oppose or impede any government acting with the people's mandate. Moreover, since the people is sovereign, the traditionally important safeguards against the abuse of power become otiose; for, in Jean-Jacques Rousseau's words, "the sovereign, being formed wholly of the individuals who compose it, neither has nor can have any interest contrary to theirs." Popular-sovereignty theory is always, therefore, on the brink of totalitarianism, since—as the French Jacobin party showed—it is only a short step from proclaiming the sovereignty of the people to claiming the unlimited authority of its elected representatives, to proscribing opposition, and to denying individuals any rights other than those which the government with majority support deems fit.

There is, of course, another view, closer to the tradition of liberal individualism, which sees democracy as a way of safeguarding and reconciling individual and group interests. For James Madison, the virtue of the new Constitution of the United States was that it permitted no faction, not even a majority, to deprive minorities of their natural rights, since it demanded the concurrence in action of independent authorities. The constitution was designed to balance diverse interests against one another, so that none might ever become a dominant and entrenched majority. More recent pluralistic accounts of democracy (or of what R. A. Dahl calls "polyarchy"), while more sophisticated, follow a similar approach. To be democratic, policy-making agencies must be sensitive to a wide range of pressures, so that no interest significantly affected by a decision will be left out of account. Popular participation consists not merely in voting, but

also in wide consultation with interest groups and in the whole process of public criticism and governmental self-justification. Democracy, according to this view, requires the dispersal, not the concentration, of power; every voter has his quantum, making him worth the attention of those who want to govern. The people is not homogeneous, but a highly diversified complex of interest groups with crisscrossing memberships. It rarely makes sense to talk of the majority, except with reference to the result of a particular election or referendum, to describe how the votes were cast. A sectional majority, if there were one, would have no intrinsic claim to rule. To govern, a party would have to piece together an electoral majority; but every elector would have his own reasons for voting as he did, and no party could say in advance that, since it had no potential supporters among the members of some particular group, that group could, therefore, be safely neglected. Admittedly, wherever group divisions coincide over a wide range of interests (as, for instance, in many polyethnic societies), these conditions might not be fulfilled, and there might be a built-in majority and minority. In such a case, no party aiming at majority support could afford to uphold a minority interest, and democracy would tend to give way to majority tyranny. Thus, where popular-sovereignty theorists see the majority as the expression of the supreme will of the people, writers such as Madison, Alexis de Tocqueville, J. S. Mill, and, more recently, Walter Lippmann and the pluralists have seen it as either a myth or a potential tyrant.

THE POSSIBILITY OF DEMOCRACY

According to elitist sociologists like Vilfredo Pareto, Gaetano Mosca, and Robert Michels, there is always, behind the democratic facade, an oligarchy, even though its members take turns at playing the key governing roles. Now obviously, in every organization leaders initiate action and followers concur, but the power relations between leader and led are not on that account always the same. Precisely because democracy is a form of political organization, it *must* also be a pattern of leadership; nevertheless, the way leaders gain and retain their authority; the extent to which their initiatives respond to the interests of those they lead; their need to listen to and answer criticism—these things distinguish a democracy in important ways from what we usually mean by an oligarchy.

For the Marxist, bourgeois democracy is a sham because equal political rights cannot equalize political power where economic power is unequal. This does not amount to saying that democracy is *necessarily* impossible, only that economic equality and a classless society are necessary conditions for it.

According to other critics, popular self-government is delusory because government calls for expertise that few voters possess. Most accept the directions of some party, to whose image they are irrationally committed, and are incapable of a rational choice of policy. However, except in the popular-sovereignty variant, democracy does not require the electors to choose policies. Their role is merely to choose governors whom they trust to deal fairly and efficiently with problems as they emerge, and to look for new governors when they are disillusioned. A party's public image need not be an irrational construct; it may accurately epitomize deep-rooted tendencies and traditional preferences and be a reliable guide to the spirit in which the party would govern.

JUSTIFICATION OF DEMOCRACY

Democracy, it is sometimes said, asks too much of ordinary men, who would never be prepared to maintain the lively and informed interest in politics that ideally it demands. This, however, presupposes a particular view of the purpose and justification of democratic government. For some writers, as J. S. Mill, men and women cannot be fully responsible, adult, moral persons unless they are "self-determining," that is, concerned about the ways in which their lives are to be controlled. This view is a development from an older natural-rights theory of democracy, according to which (in the words of Colonel Rainborough, the Leveller), "Every man that is to live under a government ought first by his own consent put himself under that government," this being a condition for preserving his natural autonomy as a rational being. Or again, for democrats in the tradition of Rousseau, men achieve moral fulfillment only as participants in the collective self-governing process, helping to give expression to the "General Will" for the "Common Good"; failure in this constitutes failure in one's moral duty as a citizen.

There is, however, a more strictly utilitarian theory, sketched by Jeremy Bentham and James Mill and implicit in a good deal of the work of democratic political scientists today. According to this view, the test of the adequacy of a political system is whether it tends to provide for the interests of the governed and protect them against the abuse of power. Democracy, they maintain, is likely to do this better than other systems. Active participation has no intrinsic virtue. Mill would have limited the franchise to men over forty, on the grounds that the interests of women and younger men would be adequately safe-

guarded by their husbands and fathers, and therefore universal suffrage would be an unnecessary expense. For many modern writers, politics is a second-order activity: if things are going well, there is really no reason for people who prefer to spend their time on other things to devote it to politics. Political activity, indeed, is often most vigorous, as in Germany before 1933, when passions are high and democracy is in imminent danger of collapse. Apathy may be a sign of political health, indicating that there are no irreconcilable conflicts nor serious complaints. If there is ground for disquiet, it is only that apathy may become so habitual that democracy's defenses may be found unmanned in the face of some future attack.

This is a prudential model of democracy, in which satisfaction is maximized and conflicts reconciled by pressures bringing countervailing pressures into operation. It leaves out of account, perhaps, the sense in which democracy moralizes politics. Because decisions have to be publicly justified, political debate is conducted in moral terms, reviewing the impact of decisions on all interests affected, not just on this or that pressure group. Moreover, the quantum of power one has as a citizen can be represented not simply as a lever for personal or sectional protection or advantage, but also as a public responsibility; for even when one's own interests are not affected, one is still a member of a court of appeal. The bystanders in a democracy are, in a sense, the guarantors that a political decision shall not simply register the strongest pressure but shall be a reasoned response to diverse claims, each of which has to be shown to be *reasonable,* in the light of whatever standards are widely accepted in the community.

See also Authority; Bentham, Jeremy; Civil Disobedience; Communism; Locke, John; Marxist Philosophy; Mill, James; Mill, John Stuart; Mosca, Gaetano; Pareto, Vilfredo; Political Philosophy, History of; Republicanism; Rousseau, Jean-Jacques; Social and Political Philosophy; Socialism; Sovereignty.

Bibliography

GENERAL

Benn, S. I., and R. S. Peters. *Social Principles and the Democratic State.* London: Allen and Unwin, 1959. Reissued as *Principles of Political Thought.* New York: Collier, 1964.

Dahl, Robert Alan. *Preface to Democratic Theory.* Chicago: University of Chicago Press, 1956. A formal analysis of types of democratic theory.

De Grazia, Alfred. *Public and Republic.* New York: Knopf, 1951. On theories of representation, with useful annotated bibliography.

Mayo, Henry B. *An Introduction to Democratic Theory.* New York: Oxford University Press, 1960.

Pennock, J. Roland. *Liberal Democracy: Its Merits and Prospects.* New York: Rinehart, 1950. Includes an extensive bibliography.

Sartori, G. *Democrazia e definizione,* 2nd ed. Bologna, 1958. Translated by the author as *Democratic Theory.* Detroit, 1962. Notes include extensive bibliographical references.

Wollheim, R. "Democracy." *Journal of the History of Ideas* 19 (1958): 225–242. Includes brief history and bibliographical references.

ON THE SEMANTICS OF "DEMOCRACY"

McKeon, R., ed. *Democracy in a World of Tensions.* Chicago: University of Chicago Press, 1951. Edited for UNESCO.

Naess, A., J. A. Christopherson, and K. Kvalø. *Democracy, Ideology, and Objectivity.* Oslo: Norwegian Research Council for Science and the Humanities by University Press, 1956. A semantic analysis of the McKeon volume.

ON THE HISTORY OF DEMOCRATIC IDEAS

Crosa, E. *La sovranità popolare dal medioevo alla rivoluzione francese.* Turin: Bocca, 1915.

Glover, T. R. *Democracy in the Ancient World.* Cambridge, U.K.: Cambridge University Press, 1927.

Gooch, G. P. *English Democratic Ideas in the Seventeenth Century* edited by H. J. Laski. Cambridge: Cambridge University Press, 1927.

Mill, James. *Essay on Government.* London, 1821; Cambridge, U.K.: Cambridge University Press, 1937.

Mill, John Stuart. *Considerations on Representative Government.* London: Parker, 1861. More recent edition by R. B. McCallum. Oxford: Blackwell, 1946.

Rousseau, J. J. *Political Writings of Jean-Jacques Rousseau.* Edited by C. E. Vaughan. Oxford, 1962. In French.

Rousseau, J. J. *Rousseau: Political Writings.* Translated and edited by F. Watkins. New York: B. Franklin, 1971.

Spitz, D. *Patterns of Antidemocratic Thought.* New York, 1949.

Talmon, J. L. *Origins of Totalitarian Democracy.* London: Secker and Warburg, 1952.

Tocqueville, Alexis de. *De la démocratie en Amérique,* 2 vols. Brussels: Meline, Cans, 1835–1840. Translated by H. Reeve as *Democracy in America,* edited by P. Bradley, 2 vols. New York: Knopf, 1945.

Woodhouse, A. S. P. *Puritanism and Liberty.* London: Dent, 1938.

Stanley I. Benn (1967)

DEMOCRACY [ADDENDUM]

It is widely agreed that democracy is a system of government in which the people rule. Since the term "democracy" is often also used to describe nonpolitical

communities—such as religious congregations, clubs, and athletic teams—democracy may be understood more generally to be a system in which the community is governed through the participation of its members. Thoroughly understanding democracy therefore requires answers to five key questions about this process of participation: What does democracy *presuppose*? What are the principal *instruments* that democracies must employ? What *conditions* are critical for its success? How can democracies be *appraised*? How can democracy be *theoretically defended*?

PRESUPPOSITIONS OF DEMOCRACY

For a democracy to be realized in any context, the community to be governed must self-consciously recognize itself as such, those entitled to participate in its government must be identified as citizens or members, and the extent of the community's concerns, both geographical and theoretical, must be at least roughly agreed upon. In short, democracy presupposes *community*.

Democratic government is possible only if members of the community can participate in decision making. They must be able to communicate effectively and to grasp the relations of means to ends. Participating citizens will not always be rational, of course, but at a minimum, democracy presupposes that members of the community have a *capacity for rationality*.

INSTRUMENTS OF DEMOCRACY

Democracies with many members must devise systems by which those members can reasonably effect their wills. All cannot speak; therefore some must speak and act for others. There are many kinds of representation (geographic, institutional, proportional, etc.). Whatever the manner of representation, every democracy of substantial size must have some *system of representation*.

Since there will be continuing and often serious disagreements among the participating members, every democracy must employ some *decision-making rules*. The rule of the majority, of the greatest number, is foremost among these, but other rules (qualified majorities, the rule of consensus, etc.) are also used in some contexts. A *system of representation* and *decision-making rules* within that system are essential instruments of every democracy.

CONDITIONS FOR THE SUCCESS OF DEMOCRACY

Self-government will not be sustainable unless some conditions are widely (but not necessarily universally) met within the community that aims to be democratic.

Some *material* conditions must be realized, chiefly a degree of economic well-being sufficient to permit its members to devote the needed time and energy to self-government. Ideally, the economic system should support, and even encourage, general participation. Yet which economic systems best do this is a topic of unending controversy.

Some *constitutional* conditions are essential for democratic participation in decision making. Among these the most essential are universal (or almost universal) protection of the *right to vote* and the *right to speak freely*. Thus a central condition of democratic success is a widely respected and legally protected liberty of all to publish with little or no restraint, to engage in robust political debate, to criticize intensely and vigorously those presently in authority.

Some *intellectual* conditions must be realized if general participation is to be reasonably successful. *Information* needed for decision making must be widely available; secrecy must be minimized. Citizens must be able to use this information. Thus an inevitable goal of every democracy will be *education*, broad and deep.

Some *psychological* conditions must also be realized. Citizens must permit intense opposition and debate while retaining good will. Citizens must be generally disposed to keep their expectations reasonable, to invite experimentation, and above all to encourage and accept compromise. A democracy is not likely to succeed unless its members, by and large, have the flexibility and resilience to bear defeat with patience when all does not go their way.

In a perilous world, political democracies require some *protective* conditions. A democratic state must be able to ward off international enemies, and therefore must sometimes rely on (while carefully controlling) military forces that themselves are not democratically organized. There must be security against attacks from without and subversion from within, but to achieve such security without sacrificing constitutional liberties is the greatest modern challenge for political communities that hope to become or remain democratic.

APPRAISING DEMOCRACIES

Self-government is a method of achieving the objectives of the members of a community. What they will seek cannot be known in advance. So appraising a democracy cannot depend on its goals, which we may despise. Democracies often make bad laws and sometimes behave immorally. The extent to which a community has

achieved a democracy may be appraised by estimating *the degree to which the will of its members is genuinely realized by its government*.

This yardstick of achievement has two principle dimensions. First, how *broadly* within the community is participation realized in fact? Excluding from participation any substantial fraction of the community directly undermines the self-government of the community as a whole. Breadth of participation is fundamental; universal participation is the ideal, never perfectly realized. Democracy is crippled when breadth is restricted by law, as when women could not vote or when serfs or slaves had no voice in community affairs. But democracy is also undermined when segments of the community are excluded in fact, even if not by law, as when ethnic minorities are oppressed in nations that profess democracy but informally limit participation. And when apathetic citizens ignore or abandon the process of participation, democracy is wounded.

Second, how *deep* is community participation? To gauge a democracy, one must estimate not only the number of citizens who vote, but also the quality of their interaction and discussion for the eventual vote. In a healthy democracy, elections are not the only manifestation of participation; member participation unceasingly goes on in the informal workings of its representative system.

Great breadth combined with substantial depth is exceedingly difficult to achieve, especially when the political community is very large, as most nations are. But that combination is the ideal against which every democracy (whatever its particular objectives) must be appraised.

THE THEORETICAL DEFENSE OF DEMOCRACY

Even when the instruments of democracy are well devised and the conditions of its success are realized in good measure, there remains the question, Why should we want democracy? Answers of two kinds may be given: We may *vindicate* the process by showing that the outcomes of self-government tend to be more beneficial than those of its alternatives. We may *justify* the process by showing that democracy is the form of government most nearly in accord with our most fundamental moral convictions, that is, by showing that universal participation in government is morally right.

VINDICATION. Democracy, its proponents contend, is the one system most likely to achieve the objectives we seek through government. Among the alternatives, it is the most likely to enact just laws, because all (or most) members are represented in the law-making process. It is the most likely to reach wise decisions (though, of course, it does not always or universally do so), because it provides maximal opportunities for all to contribute. Of all forms of government, it is the most likely to avoid violence and disorder, because all have opportunity to speak. And it is most likely to safeguard the freedoms of speech and conduct, just because those freedoms are so central to the democratic process itself. Democracies behave stupidly and badly at times, but all things considered and over the long run, they are likely to produce better outcomes for community members than those of any alternative system of government.

JUSTIFICATION. Democracy can be justified by showing that it is the one form of government that most fully gives to community members what they deserve, what is most fair to them. Democracy presupposes that community members are roughly equal—equal not in skill or strength, but in being persons with lives to live, and therefore equally entitled to a voice in community affairs. Only democracy can give them that voice. The autonomy that individuals prize in their lives is prized in the larger social sphere as well. In that larger sphere, autonomy can realized only when the members of a community, through participation in common affairs, govern themselves.

See also Civil Disobedience; Cosmopolitanism; Equality, Moral and Social; Multiculturalism; Postcolonialism; Republicanism.

Bibliography

Andersen, Kristi. *After Suffrage: Women in Partisan and Electoral Politics before the New Deal*. Chicago: University of Chicago, 1996.

Anderson, Christopher J., A. Blais, S. Bowler, et al., eds. *Losers' Consent: Elections and Democratic Legitimacy*. Oxford, U.K.: Oxford University Press, 2005.

Avritzer, Leonardo. *Democracy and the Public Space in Latin America*. Princeton, NJ: Princeton University Press, 2002.

Bird, Colin. *The Myth of Liberal Individualism*. Cambridge, U.K.: Cambridge University Press, 1999.

Calhoun, John C. *A Disquisition on Government* (1851). In his *"A Disquisition on Government," and Selections from the "Discourse"*. New York: Liberal Arts Press, 1953.

Carter, April, and Geoffrey Stokes, eds. *Democratic Theory Today: Challenges for the 21st Century*. Cambridge, U.K.: Polity Press, 2002.

Cohen, Carl. *Democracy*. New York: Free Press, 1973.

Elkin, Stephen, and Karol Soltan, eds. *Citizen Competence and Democratic Institutions*. University Park: Pennsylvania State University Press, 1999.

Friedman, Milton. *Capitalism and Freedom*. Chicago: University of Chicago Press, 1962.

Green, Thomas Hill. "Liberal Legislation and Freedom of Contract." In *Works of T. H. Green*, edited by R. L. Nettleship. Vol. 3. London: Longmans Green, 1888.

Gutmann, Amy, and Dennis Thompson. *Democracy and Disagreement*. Cambridge, MA: Harvard University Press, 1996.

Hayek, Friedrich. *The Road to Serfdom*. Chicago: University of Chicago Press, 1944.

Klosko, George. *Democratic Procedures and Liberal Consensus*. Oxford, U.K.: Oxford University Press, 2000.

Laski, Harold. *A Grammar of Politics*. London: G. Allen and Unwin, 1925.

MacDonald, J. Ramsay. *Parliament and Democracy*. London: National Labour Press, 1920.

Madison, James. "The Union as a Safeguard against Domestic Faction and Insurrection." *The Federalist*, no. 10 (1787). Available from http://www.constitution.org/fed/federa10.htm.

Meiklejohn, Alexander. *Political Freedom* (1948). New York: Harper, 1960.

Mill, John Stuart. *On Liberty* (1859). New Haven, CT: Yale University Press, 2003.

Paine, Thomas. *The Rights of Man* (1792). In his *"Rights of Man"; "Common Sense"; and Other Political Writings*. Oxford, U.K.: Oxford University Press, 1995.

Perry, Ralph Barton. *Puritanism and Democracy*. New York: Vanguard Press, 1944.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Richardson, Henry. *Democratic Autonomy: Public Reasoning about the Ends of Policy*. Oxford, U.K.: Oxford University Press, 2003.

Shapiro, Ian. *The State of Democratic Theory*. Princeton, NJ: Princeton University Press, 2003.

Sharansky, Natan. *The Case for Democracy: The Power of Freedom to Overcome Tyranny and Terror*. New York: Public Affairs Press, 2004.

Thomas, Norman. *Democratic Socialism*. New York: League for Industrial Democracy, 1953.

U.S. Supreme Court. *Schenck v. United States*, 249 U.S. 47 (1919). *Abrams v. United States*, 250 U.S. 616 (1919). *Gitlow v. New York*, 268 U.S. 652 (1925). *Whitney v. California*, 274 U.S. 357 (1927). *W. Va. State Board of Education v. Barnette*, 319 U.S. 624 (1943). *United States v. Dennis*, 341 U.S. 494 (1951). *Brown v. Board of Education*, 347 U.S. 483 (1954). *Yates v. United States*, 354 U.S. 298 (1957). *Barenblatt v. United States*, 360 U.S. 109 (1959). *Baker v. Carr*, 369 U.S. 186 (1962). "U.S." is an abbreviation for *United States Supreme Court Reports*.

Carl Cohen (2005)

DEMOCRITUS

See *Leucippus and Democritus*

DEMONSTRATIVES

Demonstratives are one type of *indexical*. Like other indexicals, demonstratives can be used to refer to different objects on different occasions. Some examples of demonstratives are *that*, *this*, *you*, *he*, *she*, *there*, *then*, *this dog* and *that yellow house*.

INDEXICALS AND DEMONSTRATIVES

Philosophers of language commonly distinguish between the *meaning* of a linguistic expression and its *referent*. For example, the definite descriptions *the president of the United States in 2003* and *the husband of Laura Bush in 2003* refer to the same individual (namely, George W. Bush), but differ in meaning. Indexicals (also known as *context-sensitive expressions*) lead many philosophers to distinguish between two different sorts of meaning. Consider the paradigm indexical *I* and suppose that Al and Bob both utter the sentence *I live in Chicago*. Their utterances of *I* have the same meaning, in one sense of *meaning*. Let us call the type of meaning that their utterances share *linguistic meaning*. But there are reasons to think that their utterances also differ in some other type of meaning. Al's utterance of *I* refers to Al, whereas Bob's utterance refers to Bob. Al and Bob also say different things: Al says that Al lives in Chicago, whereas Bob says that Bob does. Finally, one of their utterances may be true while the other is false. Let us call the type of meaning that their utterances do not share *content*. All utterances of *I* have the same linguistic meaning, but utterances that are produced by different speakers have different contents. Similarly for the sentence *I live in Chicago*.

A speaker's utterance of *I* refers to that speaker no matter what object he might want to refer to, even if he intends to refer to Napoleon Bonaparte as he utters *I* (because he mistakenly thinks that he is Napoleon) and even if he points at someone else as he produces his utterance. By contrast, the referent of a speaker's utterance of *he* depends on the speaker's intentions or pointing gestures. If Al intends to refer to George W. Bush, and points at Bush as he utters *He is a Republican*, then his utterance of *he* refers to Bush. If Al instead intends to refer to Bill Clinton, and points at Clinton, then his utterance of *he* refers to Clinton. Indexicals, such as *he*, whose reference and content depend on the actions or intentions of the speaker are commonly called *demonstratives*. Those that do not, such as *I*, are often called *pure indexicals*. David Kaplan (1989a, 1989b), whose work on indexicals is highly influential, claims that *he*, *she*, *this*, and *that* are demonstratives, whereas *I*, *today*, *tomorrow*, *yesterday*, *now*, and *here* are pure indexicals.

The pronouns *he* and *she* are often used as demonstratives, but they also have nondemonstrative uses. For instance, the pronoun *he* is used in roughly the same way as a bound variable of formal logic in the sentence *Every man thinks that he is handsome* (on the interpretation "for every man *x*, *x* thinks that *x* is handsome"). The pronoun *she* is used as an anaphor in *Fiona sat down. She picked up a pencil.* It is controversial how these uses of *he* and *she* are related to their demonstrative uses.

REFERENCE-FIXING FOR DEMONSTRATIVES

We previously observed that the reference of a demonstrative utterance depends on the speaker's intentions or pointing gestures. Kaplan's early work (1989b) tends to emphasize the importance of pointing gestures in determining reference. In this early work Kaplan says that an utterance of a demonstrative is typically accompanied by a demonstra*tion*, which is a public presentation of an object that is typically, though not always, a pointing gesture. The demonstration determines a demonstrated object (a demonstra*tum*) in a context, and the demonstratum is the referent of the demonstrative, in the context. Kaplan's later work (1989b) tends to emphasize the role of speakers' intentions. According to it demonstrations are directed towards objects by *directing intentions* and it is directing intentions that determine the referents of demonstrative utterances.

One difficulty for the view that pointing gestures determine reference is that not all utterances of demonstratives are accompanied by pointing gestures. Kaplan's early theory allows there to be demonstrations that are not pointing gestures, but unfortunately leaves unclear what demonstrations (in general) are. A problem for the view that directing intentions determine reference is that it is not clear what directing intentions are. Speakers typically have many different intentions when they utter demonstratives. When Gail utters *he*, she may simultaneously (1) intend to refer to Hal, (2) intend to refer to the man she sees, and (3) intend to refers to the man to whom others are referring with *he*. But these intentions may conflict, and it is unclear which of them (if any) is a directing intention.

KAPLAN'S SEMANTICS FOR PURE INDEXICALS

Kaplan (1989b) presents a semantics for indexicals that attempts to describe their various meanings. We shall first consider how his theory works with pure indexicals, and then consider how to extend it to simple demonstratives. (The following text concentrates on Kaplan's informal remarks about the semantics of indexicals. His formal logical system uses the apparatus of possible-worlds semantics.)

Kaplan's theory begins with the idea that a linguistic expression has a *content with respect to*, or *in*, a *context* of utterance. For instance, the word *I* has a content in every context, depending on who the agent of the context is. For every context *C*, there is an agent of *C*, a location of *C*, a time (or day) of *C*, and a possible world of *C*. The content of the word *I* in any context *C* is the agent of *C*, the content of *here* at *C* is the location of *C*, and the content of *now* and *today* at *C* is the time (day) of *C*. The linguistic meaning, or *character*, of an expression is a function whose value at any context *C* is its content in *C*. For instance, the character of *I* is a function on contexts whose value at any context *C* is the agent of *C*.

The content of a declarative sentence in a context is a *proposition*, which is an entity that can be an object of attitudes such as belief, doubt, and assertion. When a speaker assertively utters a sentence, she asserts the proposition that is the content of her sentence in her context, and if she is sincere, then she believes that proposition. On Kaplan's (informal) semantics, propositions have constituent structures that resemble the constituent structures of sentences: Just as sentences have words as parts or constituents, so propositions have individuals, properties, and relations as parts or constituents. If the content of sentence *S* in context *C* is proposition *P*, then the constituents of *P* are (roughly) the contents, in *C*, of the words in *S*. For example, suppose that *C** is a context in which Inga is the agent. Then the content of *I laugh* with respect to *C** is a proposition whose constituents are Inga and the property of laughing. If Inga laughs in the possible world of context *C**, then *I laugh* is true in context *C**.

The sentence *I am speaking* is false in some contexts, according to Kaplan, because there are contexts in which the agent is not speaking. However, every agent of every context exists in the possible world of that context. Therefore, *I exist* is true in every context. Thus, Kaplan (1989b) claims that *I exist* is a logical truth. But the content of *I exist* in a context is (usually) not a necessary truth. For example, the content of *I exist* with respect to context *C** above is the proposition that Inga exists. This is not a necessary proposition. Therefore on Kaplan's theory, the sentence *I exist* is a logical truth that fails to express a necessary truth in many contexts.

EXTENDING KAPLAN'S THEORY TO SIMPLE DEMONSTRATIVES

One way to extend Kaplan's (1989a, 1989b) theory to simple demonstratives involves adding more items to contexts. (A second way, which will not be discussed here, involves Kaplan's dthat-terms.) For instance, one can suppose that every context has a sequence of demonstrata, and that the content of $that_1$ with respect to context *C* is the first demonstratum of *C*, the content of $that_2$ is the second demonstratum of *C*, and so on. One can then say that the character of $that_1$ is a function whose value at every context *C* is the first demonstratum of *C*, the character of $that_2$ is a function whose value at every context *C* is the second demonstratum of *C*, and so on. Similarly, one can suppose that every context has a sequence of addressees and that the content of you_1 is the first addressee, the content of you_2 is the second addressee, and so on.

There are two difficulties with this extension. First, We saw earlier that the referent of a demonstrative in a context is determined, somehow, by the pointing gestures and intentions of the speaker. The preceding theory assumes that every context has a sequence of demonstrata. But as Kaplan (1989b) points out, the agents of many contexts are not pointing at any objects and do not have any intentions that are relevant to determining a referent for a demonstrative. So it is highly artificial to suppose that every context contains a sequence of demonstrata. Second, the English word *that* is a single linguistic expression with a single linguistic meaning. But the previous theory instead provides an infinite number of distinct subscripted demonstratives ($that_1$, $that_2$, …), each with its own character (Braun 1994).

BELIEF AND DEMONSTRATIVES

Propositions that have individuals as constituents are known as *singular propositions*. Kaplan's theory says that singular propositions can be asserted and believed. This claim is problematic. Suppose that John is wearing a shirt with a large stain on its back. Suppose that he sees the back of his shirt in a mirror, but does not realize that he is viewing his own shirt. Then he may sincerely say *I am wearing a clean shirt and he is not wearing a clean shirt*, as he points at the person reflected in the mirror. On Kaplan's theory, John asserts and believes the contradictory singular proposition that John is wearing a clean shirt and John is not wearing a clean shirt. But surely he does not believe a contradictory proposition.

Gottlob Frege gives similar reasons for thinking that the content of a proper name is not its referent, but is instead a sense that determines a referent. Kaplan (1989b) and John Perry (2000) respond to Fregean criticisms of Kaplan's theory and criticize Frege's theory of demonstratives.

COMPLEX DEMONSTRATIVES

Complex demonstratives are expressions of the form *that CN*, where *CN* is a common noun phrase. Examples are *that car*, *that man who is wearing a baseball hat*, and *that yellow house*. Kaplan's theory does not mention complex demonstratives, and it is not entirely clear how they should be integrated into a theory of demonstratives. The major issue concerns the property expressed by the common noun phrase inside a complex demonstrative. Is this property a part of the content of the entire complex demonstrative? For instance, does the content of an utterance of *that yellow house* include the property of being yellow? Or, alternatively, is the content of the complex demonstrative simply the object to which the utterance refers?

On the one hand, it seems that a speaker's intentions and demonstrations are relevant to determining the referent of *that yellow house*. In this respect, *that yellow house* resembles the simple demonstrative *that*. We concluded earlier that the content of a simple demonstrative is just its referent. Therefore, We have some reason to think that the content of a complex demonstrative is also its referent and to think that the property of being yellow is not a part of the content of *that yellow house*. David Braun (1994) and Nathan Salmon (2002) argue for this view of complex demonstratives. On the other hand, the complex demonstrative *that yellow house* has a syntactic form much like the syntactic forms of the definite description *the yellow house* and the quantifier phrase *some yellow house*. Most philosophers think that the contents of *the yellow house* and *some yellow house* include the property of being yellow. That is some reason to think that the content of *that yellow house* also includes the property of being yellow. Jeffrey C. King (2001) argues for this latter view.

See also Anaphora; Frege, Gottlob; Indexicals; Logical Form; Meaning; Modality and Quantification; Proper Names and Descriptions; Propositional Attitudes; Propositions; Sense.

Bibliography

Braun, David. "Demonstratives and Their Linguistic Meanings." *Noûs* 30 (1996): 145–173.

Braun, David. "Indexicals." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2001.

Braun, David. "Structured Characters and Complex Demonstratives." *Philosophical Studies* 74 (1994): 193–219.

Kaplan, David. "Afterthoughts." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein, 565–614 . New York: Oxford University Press, 1989a.

Kaplan, David. "Demonstratives." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein, 481–563 . New York: Oxford University Press, 1989b.

King, Jeffrey C. *Complex Demonstratives: A Quantificational Account*. Cambridge: MIT Press, 2001.

Perry, John. *The Problem of the Essential Indexical and Other Essays*. Expanded ed. Stanford, CA: CSLI Publications, 2000.

Salmon, Nathan. "Demonstrating and Necessity." *Philosophical Review* 111 (3) (2002): 497–537.

David Braun (2005)

DE MORGAN, AUGUSTUS

(1806–1871)

A British mathematician and logician, Augustus De Morgan was born at Madura, India, where his father was an army officer. After early education in the west of England, he entered Trinity College, Cambridge, in 1823 and graduated fourth wrangler in 1827. His refusal to subscribe to the religious tests then in force precluded him from further advancement at Cambridge, but he was fortunate enough to be appointed first professor of mathematics at the newly opened University of London. Because of his habit of resigning on matters of principle, he twice vacated this chair, once at the beginning and once at the end of his career; but he enjoyed, in the interval, the highest repute and affection as a teacher and had many pupils who later achieved distinction.

In addition to numerous important papers on the foundations of algebra and the philosophy of mathematical method, De Morgan was the author of several excellent elementary textbooks; a standard bibliography, *Arithmetical Books* (London, 1847); a large treatise on the calculus (London, 1842); and an enormous quantity of learned journalism, mostly in the shape of review articles in the London *Athenaeum* and contributions on mathematical and astronomical subjects to the *Companion to the Almanac* (1831–1857) and to the *Penny* (later *English*) *Cyclopaedia.* His best-known work in this line is the posthumously assembled *Budget of Paradoxes* (London, 1872), a still-diverting miscellany from the lunatic fringes of science and mathematics, originally serialized in the *Athenaeum.* Despite many years' service as secretary of the Royal Astronomical Society, De Morgan was in general suspicious of official bodies and distinctions, never sought membership in the Royal Society, and declined an Edinburgh LL.D. Indifferent to politics and society—and professedly hostile to the animal and vegetable kingdoms as well—he nonetheless maintained an extensive scientific correspondence with such friends as William Whewell, George Boole, Sir John Herschel, Sir William Rowan Hamilton (the mathematician), and John Stuart Mill. His crotchets did little to disguise his exceptional benevolence and firmness of character or to inhibit his talents as a humorist and a wit.

De Morgan's outlook was that of a philosophical mathematician and historian of science; he did not claim to be a philosopher in any narrow sense of the term. He admired Berkeley and followed him to the extent of holding the existence of minds to be more certain, as a fact of experience, than that of a material world. But his general attitude to such questions may be gathered from his remark that, while he would not dissuade a student from metaphysics, he would warn him, "when he tries to look down his own throat with a candle in his hand, to take care that he does not set his head on fire."

In common with other mathematicians of his time, De Morgan realized that algebra could be conceived as a system of symbols whose laws could be codified independently of any arithmetical or other interpretation that might be given to them. His logic had a similar aim. Deeply versed in the history of logic, he was able to freshen and illuminate the subject by generalizing its traditional principles along mathematical lines. In this respect he ranks as the chief precursor of Boole; but his views attained notice chiefly through the controversy that arose when Sir William Hamilton (of Edinburgh) accused him of plagiarizing the doctrine of a quantified predicate.

De Morgan's *Formal Logic* (London, 1847) represents the best-known, though by no means the most mature, statement of his logical views. Among its many excellences, the chapter on fallacies is worthy of mention. De Morgan's later work is dispersed in pamphlets and periodicals, most notably in five memoirs contributed to the *Cambridge Philosophical Transactions* (Vols. 8–10, 1847–1863) and in his *Syllabus of a Proposed System of Logic* (London, 1860, reprinted in *On the Syllogism* (London, 1964). Though too largely concerned with polemics against Hamilton, and hampered by a notation that found no acceptance, these writings display much originality in the handling of negative terms, compound propositions, and numerous unorthodox varieties of syl-

logistic reasoning. Apart from the well-known "De Morgan laws" for the negation of conjunctions and disjunctions (or logical sums and products), the most important development was the recognition that the copula performs its function in the syllogism solely by virtue of its character as a transitive and convertible relation. De Morgan was led by this to examine the logic of relations in general and so paved the way not only for Peirce's "logic of relatives" but for all that has since been done in this branch of the subject.

As a skilled actuary, who was often in demand as a consultant to insurance companies, De Morgan was not unnaturally interested in the mathematical theory of probability and the problems of applying it to the hazards of mortality and other types of experience. His treatise "Theory of Probabilities," in the *Encyclopaedia Metropolitana* (London, 1837) and the more popular *Essay on Probabilities* (London, 1838) were among the earlier discussions of this topic in English (see further relevant chapters of *Formal Logic* and the papers on the evaluation of argument and testimony attached to the first two Cambridge memoirs above). De Morgan's conception of probability was largely derived from Pierre Simon de Laplace, whose ideas (and errors) he was thus instrumental in propagating among his nineteenth-century successors. His method of approach was to construe the theory as an extension of formal logic, that is, as an investigation of the rules whereby propositions not absolutely certain affect the certainty of other propositions with which they are connected. He also employed the "inverse" procedures founded on Bayes's theorem, whereby, from known factual premises, it is sought to conjecture the probabilities of their likely or possible antecedents. In attempting to quantify the degree of uncertainty involved, De Morgan identified it with the amount of belief that is, or rather, that ought to be attached to it by a rational person, and proceeded on this basis to discuss the compounding and derivation of partial beliefs in accordance with the mathematical rules of the calculus of chances. His view of the matter was thus both a priori and subjective, though not in the objectionably psychological sense that has sometimes been ascribed to him. There are better reasons for censuring the technical errors he fell into through uncritical reliance on the Laplacean "rule of succession" and "principle of indifference"; even here, however, his confidence in the mathematical apparatus was often less blindly trusting than that of the writers who preceded him.

De Morgan's conception of scientific method may be gathered primarily from a review of Francis Bacon's works inserted in the *Budget of Paradoxes.* He there embraced what is essentially the modern "hypothetico-deductive" view of the subject; but one has to go to William Whewell before him or to W. S. Jevons after him to see it worked out in full.

See also Bacon, Francis; Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Berkeley, George; Boole, George; Hamilton, William; Herschel, John; Jevons, William Stanley; Laplace, Pierre Simon de; Logic, History of; Logic, Traditional; Logical Terms, Glossary of; Mill, John Stuart; Scientific Method; Whewell, William.

Bibliography

Apart from the works mentioned under the History of Logic entry there is not much literature on De Morgan. The best general accounts are in A. Macfarlane, *Ten British Mathematicians* (New York, 1916) and J. A. Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1957). The *Memoir* by his wife, Sophia Elizabeth De Morgan (London: Longmans, Green, 1882), contains an excellent account of him. The details of the quarrel with Hamilton may be found in Hamilton's *Letter to Augustus De Morgan* (Edinburgh, 1847) and De Morgan's *Statement in Answer* (London, 1847) or more readily in the appendices to De Morgan's *Formal Logic.*

P. L. Heath (1967)

DENNETT, DANIEL CLEMENT

(1942–)

Daniel Clement Dennett obtained his first degree at Harvard, where, as he tells us in *Brainchildren* (1998), he vigorously resisted the most influential American philosopher of the twentieth century, Willard van Orman Quine. He then did a D. Phil. in Oxford in a brief two years under Gilbert Ryle, the most influential Oxford philosopher of his time, finishing in 1966.

His first book was *Content and Consciousness* (1969). These two words, *content* and *consciousness*, encapsulate much of Dennett's mission. *Content* refers to the contents of the mind—all the beliefs, desires, values, emotions, hopes, expectations, memories, and so forth that make up the mind. *Consciousness* refers, of course, to consciousness. In Dennett's view, the correct order in which to examine these topics is content first and then consciousness. Dennett's central project is already clear in this book, the project of "naturalizing the mind." This is the

project of showing that mind and consciousness are simply aspects of brain and behavior, just as much open to investigation by cognitive psychology and neuroscience as other aspects of cognition. He has never waivered in this commitment.

CONTENT

Dennett's next book was a collection of essays, *Brainstorms*, written during the 1970s. This work helped launch a unique publishing enterprise, Bradford Books. Founded by Harry and Betty Stanton and subsequently absorbed by MIT Press, the Bradford insignia has become one of the most important collections of books in philosophy of mind and cognitive science in the English language.

Brainstorms begins with the first full articulation of Dennett's distinctive approach to mental content, the approach that he calls the intentional stance. According to Dennett, we can approach something in order to explain it from three stances: the physical stance, the design stance, and the intentional stance. Each has its own advantages and costs, but none is describing reality from the one correct perspective.

After editing, with Douglas Hofstadter, a charming collection of works by others on the mind, *The Mind's I* (1981), Dennett next turned to decision making and responsibility in an idiosyncratic little book called *Elbow Room: The Varieties of Free Will Worth Wanting* (1984). The book began life as John Locke Lectures in Oxford and espouses a brisk compatibilism between decisions being causally determined and decisions being free in any way that is "worth wanting." Interestingly, he returned to the topic of free will nearly twenty years later in *Freedom Evolves*(2003).

The year 1987 saw his second major collection of papers on content, *The Intentional Stance*. The papers in this collection are probably the most influential papers that Dennett has written. Near the end of the collection are two papers on evolutionary theory, a topic that was to loom large in his thinking in the 1990s.

Dennett's work on mental content has led him to questions about such topics as artificial content (AI [artificial intelligence]), the evolution of content, the relationship of content to the environment and brain (neuroscience), content in nonhumans (cognitive ethology), the nature of explanation in psychology and science generally, how content is represented and the different styles of mental representation, the relationship of representations to the brain, and how we ascribe mental content to ourselves and others.

CONSCIOUSNESS

At this point Dennett turned to consciousness, and a large book, *Consciousness Explained* (1991), ensued. For the first time, Dennett wrote a book deliberately aimed at a wide audience (it was not the last). Dennett laid out methods for studying consciousness, built a model of consciousness as a cognitive system, and discussed the nature of introspection (the consciousness we have of ourselves and our own mental states). He considered how consciousness evolved, pathologies of consciousness such as dissociative identity disorder (formerly multiple personality disorder), whether there is any real difference between how a mental state functions in us and how it feels to us (what philosophers call *qualia* or felt quality), what selves might be, the neural implementation of consciousness, and so on—just about every issue pertaining to consciousness.

This book has two main targets. One is the picture of conscious states that the tradition received from Descartes. This is the idea that there is something to a conscious state, some felt quality, that is unmistakably clear and clearly different from anything else in the world. The other is what Dennett calls the Cartesian theater, the idea that the conscious system is a kind of screen on which conscious states play before a little homunculus sitting in the middle of the theater. To replace the Cartesian picture in both its parts, Dennett proposed what he calls a Multiple Drafts Model (MDM) of consciousness. MDM treats consciousness as a kind of mental content, almost a matter of programming.

Dennett next wrote a shorter book pulling the two sides of his work together: *Kinds of Minds* (1996). Then he turned to a task that had been awaiting him for a long time: evolutionary theory. *Darwin's Dangerous Idea* (1995) was also published as a trade book and also enjoyed phenomenal success. Here Dennett argues for two main claims: (1) Darwin's theory of evolution is a "universal acid" that dissolves all manner of intellectual "skyhooks" and other pseudoscientific props that philosophers (and not just philosophers) have dreamed up to try to patch up hopeless theories; (2) yet contrary to those who see Darwin as the destroyer of all morality, the theory of evolution leaves one perfectly satisfactory approach to morality and political philosophy untouched: traditional western liberalism. Among the most important claims introduced in this book is that it is language that makes it possible for us to have our kind

of mind, a kind of mind that, by being able to cooperate with other minds and record the results of cooperation for others to build on, can figure out the physics of the universe, find cures for serious diseases, build Hubble telescopes and the Channel tunnel, and so on.

The book set off a stormy debate with Steven Jay Gould and others in the *New York Review of Books* in 1997. Gould insisted that Dennett had espoused an ultra-adaptionist position, assigning change in species to natural selection (selection on the basis of survival and reproductive fitness) over almost all other sources of change over time, such as cataclysmic changes in weather, exhaustion of habitats. Despite the heat that the debate generated (and some astonishingly uncollegial language), with the passage of time it now seems clear that the elements of agreement between the two of them are far greater than the elements of disagreement.

In the late 1990s, Dennett published another collection of essays, *Brainchildren* (1998), a remarkably diverse array of pieces mostly on consciousness and artificial intelligence. His most recent book is *Freedom Evolves* (2003). He is working on a book on religion. There are many sides to Dennett's contribution, but one of the most important is the way he challenges orthodoxies. He is a master at showing what is wrong with points of view with which he disagrees. One of his most characteristic techniques is to go after comfortable ideas with what he calls intuition pumps. Following is an example, the case of Mr. Chase and Mr. Sanborn:

> Mr. Chase and Mr. Sanborn both used to like a certain coffee. More recently it has lost its appeal. The reasons they give seem to differ markedly. Chase: "The flavor of the coffee hasn't changed but I just don't like that flavor very much now." Sanborn: "No, no, you are quite wrong. I would still like *that* flavor as much as ever. The problem is that the coffee *doesn't* taste that way anymore." (reconstructed from Dennett 1988, p. 50)

Dennett's target is the idea that there is always a clear distinction between a conscious state, in this case how something tastes to us, and how we react to it. When we read about Mr. Chase and Mr. Sanborn, we are meant to say to ourselves, "Hmmm, maybe the distinction is not so clear after all." One is then meant to see that similar doubts arise all over the place.

An expert high-seas sailor and an accomplished pianist and choral singer, Dennett is far from retirement. In addition to his prolific authorship of books, Dennett has written an average of ten papers per year for thirty-five years. He has taught at Tufts University for more than thirty years.

See also Cognitive Science; Consciousness.

Bibliography

WORKS BY DENNETT

Content and Consciousness. London: Routledge, 1969.

Brainstorms. Montgomery, VT: Bradford Books, 1978.

The Mind's I, edited with Douglas Hofstadter. New York: Basic Books, 1981.

Elbow Room: The Varieties of Free Will Worth Wanting. Cambridge, MA: MIT Press/A Bradford Book, 1984.

The Intentional Stance. Cambridge, MA: MIT Press/A Bradford Book, 1987.

"Quining qualia." In *Consciousness in Contemporary Society*, edited by A. Marcel and E. Bisiach, 42–47 Oxford: Oxford University Press: 1988.

Consciousness Explained. Boston: Little Brown, 1991.

Darwin's Dangerous Idea. New York: Simon and Schuster, 1995.

Kinds of Minds. New York: Basic Books, 1996.

Brainchildren. Cambridge, MA: MIT Press/A Bradford Book, 1998.

Freedom Evolves. New York: Penguin Books, 2003.

WORKS ABOUT DENNETT

Brook, Andrew, and Donald Ross. *Daniel Dennett: Contemporary Philosophy in Focus*. New York and Cambridge, U.K.: Cambridge University Press, 2002.

Dahlbom, Dagmar, ed. *Dennett and His Critics: Demystifying Mind*. Oxford: Blackwell, 1995.

Quine, Willard v. O. *Word and Object (Studies in Communication)*. Cambridge, MA: MIT, 1964.

Ross, Don, Andrew Brook, and David Thompson, eds. *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, MA: MIT Press/A Bradford Book, 2000.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949.

Andrew Brook (2005)

DEONTIC LOGIC

See *Modal Logic*

DEONTOLOGICAL ETHICS

Deontology is the view that because there are moral constraints on promoting overall best consequences, sometimes the right action is not the one whose consequences are best. The constraints that deontological theories emphasize are familiar from our everyday experience of morality: One ought to keep one's promises and be loyal to one's friends; one ought not to inflict unnecessary suf-

fering or to ignore one's debts of gratitude, and so on. Some deontological theorists see a unified basis for all such duties; others are frankly pluralist.

THE MEANING OF "DEONTOLOGY"

Apparently coined by Jeremy Bentham in the nineteenth century, the term "deontology" initially was used to refer, quite widely, to the "science of duty." This wide usage reflects the word's two Greek roots: *deon*, meaning "needful" or "fitting," and *logos*, meaning "science" or "discourse." Within a century, the term gained its narrower meaning; even this wide use of the term, however, carries some definite commitments. These stem from each of its two roots: (1) It is fitting to be virtuous; but it is not part of the concept of virtue that virtue is always needful or morally necessary. In focusing on the needful, deontology may thus either leave the moral virtues to one side or demote them to a derivative status. In its concern with duty, deontology also either ignores, or treats as peripheral, the nature of moral success—described by some as happiness or *eudaimonia*, by others as perfected moral worth. Deontology's principal terms of assessment are, instead, deontic: they concern what ought to be the case and, more specifically, what people ought (morally) to do. (2) It is possible to speak or discourse about almost anything; but holding that there is a *logos* of duty suggests that moral duties may be correctly described in general terms. This suggestion runs contrary to the views of at least some contemporary moral particularists, who deny that there are any general truths about what people ought to do.

By the middle of the twentieth century, "deontology" acquired its more specific meaning, which refers to a particular conception or theory of our moral duties. To say that something is one's duty is to represent a type of action as necessary in some way, but how? C. D. Broad noted that there are at least two ways. One way is to represent the action as a means best suited to our attaining some good end that we ought, unconditionally, to pursue: This he called "teleological necessity" (from the Greek word *telos*, meaning "end"). Another is to represent the action as one that we ought, unconditionally, to take, irrespective of the consequences: This he called "deontological necessity." Of course, it is perfectly possible to embrace both of these types of moral necessity—as did, for instance, Immanuel Kant, who recognized both obligatory ends and strictly prohibited actions. Yet this contrast between two types of moral necessity may also be used to divide moral theories into two groups.

"Teleological" theories, as it has become commonplace to say, hold that the rightness and wrongness of actions is *wholly* determined by their tendencies to generate good. "Deontological" theories deny this, holding that the right action at least *sometimes* expresses deontological necessity, which stands independent of teleology—even a teleology that tots up *overall* goodness. As we have seen, the initial, wide use of "deontology" suggests that there are general ethical truths. In keeping with that suggestion, the more specific idea of deontological theory, as it is usually presented, invokes the idea of a general type of duty. One or more generally statable moral *constraints* prohibits certain ways of pursuing good results. In this way, we arrive at the conception of deontology stated at the outset: Deontology is the ethical theory, or family of ethical theories, according to which there are constraints on promoting overall best consequences that imply that sometimes the right act is not the one whose consequences are best.

This distinction between teleological and deontological theories does not cover all possible ethical theories—not even all of the non-particularist ones that focus on duty rather than virtue or happiness. On the one hand, there are other ways of resisting the idea that the right act is the one whose consequences are best. Philippa Foot (1985) and others have questioned the coherence of this apparently all-purpose notion of overall goodness: Does it really make sense to ask, for any two states of affairs or any two alternative actions, which is "better?" Another possibility (emphasized by Samuel Scheffler [1982]) is to hold that the basis for deviating from what is for the best is not a set of constraints or duties, but rather a set of prerogatives or permissions: Perhaps we sometimes have moral latitude to act in some merely acceptable ways. On the other hand, there are ways of developing consequentialism that drop any reference to teleological necessity: Perhaps we simply rank (some) alternative actions as better or worse, on the basis of whatever considerations apply, interpreting "consequentialism" simply as holding that we ought to take the best available alternative. If this abstract understanding of consequentialism is taken to the limit, the contrast between deontology and consequentialism will blur. To see why, we must shift from the meaning of deontology to the merits of the view.

THE MERITS OF DEONTOLOGY

Deontological constraints are often called "agent-centered." The negative ones, for instance, direct people not to *do* certain things while not directing them to minimize the extent to which certain kinds of action *are done*.

Although these constraints are typically conceived as applying to everyone, that does not mean that they apply in the latter, impersonal, way, but only that they apply to each agent in the former, personal way. This distinction is implicit in St. Paul's principle from Romans 3:8, central in Alan Donagan's (1977) deontological theory: "Thou shalt not do evil in order that good may come of it." The principal difficulty in justifying deontology is to explain why agent-centered restrictions make sense. If breaking a promise or unnecessarily harming someone is a bad thing, then would not rationality dictate minimizing this type of bad, other things equal? Niccolò Machiavelli infamously wrote that princes need to learn how not to be good. They must be ready to use cruelty well in order to minimize cruelty in the long run; but perhaps the advice applies not only to princes. Should one not suffocate the crying child so as to prevent the evil soldiers from finding the refugees in the attic and killing them all? Such cases present what are known as "paradoxes of deontology."

Some deontological theorists simply deny that the paradoxes pose any problem: The deontological constraints stand on their own—as self-evident, a priori, or resting on divine authority—and entail nothing about minimizing bads. Others do attempt to defend deontology from the challenge posed by the paradoxes, in three ways. One is to defend the moral significance of the distinction between doing ill and allowing ill effects to happen. The doctrine of doing and allowing holds that, across a wide range of cases, there is a morally significant difference between the two. As Warren Quinn's (1993) sympathetic discussion reveals, it is not easy to explain why the bare difference between doing something and allowing something to happen should make a fundamental moral difference; a first step is to concede that the distinction matters only in certain contexts.

A second way to attempt to defend deontology is to develop Thomas Nagel's (1986) idea of "agent-centered" (or "agent-relative") *reasons*, which explain the point of deontological constraints. We can understand how certain moral reasons may not apply to everyone, but only to some people. This may be because of special relationships in which some people stand to others (friend, physician) or it may be because of the moral leeway we have to pursue what we care about. Perhaps the reasons that underwrite deontological constraints are similarly agent-relative. W. D. Ross (1988) suggested that each important deontological constraint reflects a different special relationship in which we can stand to others: as past benefactors, promisors, and so on. Alternatively, agent-relative reasons can be given a systematic place in moral theory. For example, T. M. Scanlon's (1998) contractualist theory holds that the rightness of actions is determined by principles that could not be reasonably rejected by anyone motivated to reach reasonable agreement on principles. The reasons that individuals might reject proposed principles, he suggests, will naturally include some agent-relative ones. In either sort of deployment, however, there are grounds for worrying that the agent-relative reasons presuppose deontological constraints rather than really justifying them.

A third way to defend deontological constraints is to deny that all goods call for one to promote them. Some goods—some valuable things—may instead call upon us to respect or honor them. Kant's seminal contribution to deontological thinking was his insistence that rational persons are to be respected, as having a dignity that is beyond all price. Having said that, one *might* turn around and argue that human rational dignity, as an agent-neutral value, is to be promoted. That would be to turn away from deontology. In contrast, one might hold that the appropriate attitude to human dignity is, in turn, to respect *it*. Frances Kamm (1992) has argued that human dignity is best respected by ensuring the inviolability of persons' basic rights.

THE PRIORITY OF RIGHT

As Kamm herself points out, resting deontological constraints on the value of human dignity begins to efface the distinction between deontology and consequentialist views. A fully abstract consequentialism can look to *any* relevant basis for holding that one alternative action is better or worse than another. Jamie Dreier (1993) has argued that the strictness of deontological constraints can easily be recast in a consequentialist mode by stipulating that some actions be ranked *lexically* better than others. Such an abstract consequentialism gives up the title to being a teleological view, as it does not develop its content on the basis of observed teleological necessities; but it remains recognizably consequentialist. Maintaining a firm contrast between deontological theories and consequentialist theories, therefore, depends upon resisting such an abstract recasting of consequentialism. Friends of deontology may want to understand "goods" or "goodness" somewhat more narrowly, as referring only to features of states of affairs or to values towards which the correct stance is promotion rather than respect, such as human well-being. The values that do not fall within the good—or at least some of them—may then be thought of as belonging to "the right"—the domain of rightness.

Accepting that there are values proper to the right, whose role is not to be promoted but rather to be honored and respected by the structure of duty, opens up many additional possible ways of defending deontology. Such an approach helps ground deontological constraints by relating them to some value, but in a non-teleological way; and this helps explain why we should care about acting morally. Barbara Herman (1993) has emphasized this layer in Kant's moral theory. Kant (1998) held that the a priori concepts of morality determine the content of the one unconditionally valuable thing, namely the good will, the will that acts from respect for the moral law. Although this value cannot, in Kant's view, be directly promoted, it helps characterize the value of acting morally. Our capacity to achieve this value also underwrites our dignity: We are worthy of respect because we are capable of acting with a good will. A contemporary example of this approach is Scanlon's contractualism. As noted above, Scanlon (1998) holds that the rightness of actions depends on whether they accord with principles that no one duly motivated to find agreement on principles could reasonably reject. What motivates us to act morally, on this interpretation? Scanlon's answer in *What We Owe to Each Other* is that it is "the positive value of living with others on terms that they could not reasonably reject" (1998, p. 162). This value seems to belong to the right, not to the good insofar as it is distinct from the right. Yet it is nonetheless something we might intelligibly care about. According to deontologists such as Kant and Scanlon, then, these considerations of rightness have a kind of structural priority over other types of value.

See also Categorical Imperative; Contractualism; Divine Command Theories of Ethics; Duty; Ethics; Kant, Immanuel; Kantian Ethics; Moral Rules and Principles; Ross, William David; Teleological Ethics.

Bibliography

Bentham, Jeremy. *Deontology, Together with a Table of the Springs of Action and the Article on Utilitarianism*, edited by A. Goldworth. Oxford: Clarendon Press, 1983.

Broad, C. D. *Five Types of Ethical Theory*. London: Routledge & Kegan Paul, 1930.

Darwall, Stephen, ed. *Deontology*. Malden, MA: Blackwell, 2003.

Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1977.

Dreier, Jamie. "Structures of Normative Theories." *Monist* 76 (1993): 22–40.

Foot, Philippa. "Utilitarianism and the Virtues." *Mind* 94 (1985): 196–209.

Herman, Barbara. "Leaving Deontology Behind." In *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Kamm, Frances. "Non-Consequentialism, the Person as an End-in-Itself, and the Significance of Status." *Philosophy and Public Affairs* 21 (1992): 354–389.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, edited and translated by M. Gregor. Cambridge, U.K.: Cambridge University Press, 1998.

Louden, Robert B. "Toward a Genealogy of 'Deontology.'" *Journal of the History of Philosophy* 34 (1996): 571–592.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Quinn, Warren. *Morality and Action*. Cambridge, U.K.: Cambridge University Press. 1993.

Ross, W. D. *The Right and the Good* (1930). Indianapolis: Hackett, 1988.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Scheffler, Samuel. *The Rejection of Consequentialism*. Oxford: Clarendon Press, 1982.

Henry S. Richardson (2005)

DERRIDA, JACQUES

(1930–2004)

Although he was not altogether happy with the fact, Jacques Derrida's name has become synonymous with deconstruction. Derrida was born in El-Biar, near Algiers, in 1930. In 1949 he left for Paris and in 1952 began to study at the École Normale Supérieure, where he taught from 1964 to 1984. Beginning in 1975, Derrida spent a few weeks each year teaching in the United States. While at Yale University Derrida collaborated with Paul de Man (1919–1983), leading to the extraordinary impact that deconstruction has had on the study of literature in the United States, an impact that quickly spread to other disciplines and countries.

Derrida's record of publications is remarkable. In 1962 he wrote an introduction to a translation of Husserl's *Origin of Geometry* that in many respects anticipates the later works. In 1967 he published a further study of Edmund Husserl, *Speech and Phenomena*; a collection of essays, *Writing and Difference*; and a reading of Ferdinand de Saussure, Claude Lévi-Strauss, and Jean-Jacques Rousseau, *Of Grammatology*. A rapid succesion of publications ensued, among the most important of which are *Dissemination* (1972), *Glas* (1974), *The Post Card* (1980), *Psyché* (1987), *Given Time* (1991), and *The Politics of Friendship* (1994). Derrida also published extensively on an increasingly broad range of subjects from literature and politics to art and architecture.

STYLES OF DECONSTRUCTION

Deconstruction is neither a method nor a negative critique. It is better understood as a strategy for reading texts under the influence of Husserl, Martin Heidegger, Friedrich Nietzsche, Sigmund Freud, Emmanuel Levinas, and Saussure. In the early years of deconstruction many of the most important readings were devoted to these thinkers, all of whom, except for Husserl, were treated in Derrida's 1968 lecture "Différance." Derrida justified this cross-fertilization of disparate authors by saying that their names served to define contemporary thought. This practice came to be generalized as *intertextuality* and came to be further enhanced as Derrida, in each new text, drew heavily on his previous readings. Because Derrida's language is both cumulative and parasitic on the texts that he is reading, attempts to formulate Derridean doctrines are often misleading. Hence it is more appropriate to focus on his strategies.

Most of Derrida's writings operate by close reading, and their impact depends on the capacity of this reading to account for details that more conventional readings either ignore or explain away. In clear contrast, not only with most modern trends in philosophy, but also with a widespread image of him, Derrida was immersed in the history of philosophy. For Derrida this was the only way to avoid unwittingly repeating the most classic gestures of philosophy, a danger that threatens every attempt to ignore that history and begin philosophy anew. Deconstruction locates itself within traditional conceptuality in order to find the radical fissures that it believes can be traced in every work of philosophy. Derrida was drawn to the apparent contradictions of the tradition and made them the starting point of his readings, whereas a more conventional treatment tends to stop short as soon as a contradiction is identified. Much that is strange, and to some even offensive, about Derrida's analyses arises because he attempted to uncover the structures that organize and so transcend or exceed conventional reason.

Particularly in his early writings, Derrida presented his deconstructive readings of individual works in the history of Western philosophy as directed against a certain understanding of that history as one in which presence had been privileged. The priority of presence was reflected throughout the binary oppositions that structured Western thought: presence over absence, speech over writing, inside over outside, and so on. Derrida's strategy was to show that those texts that were supposed to have exhibited this privilege of presence also reflected a certain counter-tendency. So, for example, texts that on the surface appear to privilege speech over writing also have moments in which the hierarchy was reversed. Following this reversal, Derrida sought to pass beyond the opposition to that which exceeds it: Hence, in the example given, he identifies what he calls a proto-writing, which is neither speech nor writing in the conventional sense, but that which is the condition of all forms of language.

Derrida drew his account of history of Western metaphysics in terms of presence from Heidegger, but in so far as Heidegger's account was directed toward the overcoming of metaphysics, Derrida located within that project an opposition between what was inside and outside metaphysics. He thus identified within the project of leaving metaphysics behind, a hierarchical opposition that was itself still metaphysical. By contrast, Derrida's own position was that one cannot stand unequivocally either within or outside metaphysics. This was reflected in his strategy of double reading. To any text that was conventionally conceived of as belonging to Western metaphysics, he added a new reading that showed how that same text could be understood as exceeding Western metaphysics, and to texts, such as those by Heidegger and Levinas, that presented themselves as passing beyond Western metaphysics, Derrida added a reading that drew them back into the conceptuality of Western metaphysics. The deconstruction lay not in the new reading alone, but in its juxtaposition with previous readings, which were not thereby supplanted so much as understood as belonging to the history of the text. This means that Derrida does not so much oppose the dominant reading, as that he adds another reading to it, so that the so-called double reading combats any attempt to locate the text in question either within or outside Western metaphysics.

There is, however, another style of deconstruction that has become increasingly widespread in Derrida's thought. It proceeds by the exploration of *aporias*, as will be illustrated later in a discussion of Derrida's conception of the gift. Because Derrida sometimes seems to give the aporias he investigates a universal status, deconstruction in this sense is no longer as attached to the conception of the history of philosophy as the history Western metaphysics, as was the case with his textual readings of philosophy. However, Derrida did not consider these two styles of deconstruction as independent of each other, so that it would be a mistake to suppose that he had abandoned the genealogical component of his work.

THE SUPPLEMENT

Supplement is one of the key terms of Derrida's challenge to Western metaphysics, understood as a unified body of

thought that privileges presence. He used this term to problematize the philosophical quest for a simple origin as a self-sufficient source. He identified a "logic of supplementarity," which is said to be "inconceivable to reason," according to which the supplement, by delayed reaction, produces that onto which it is said to be added (*Of Grammatology*, pp. 179 and 259). The force of Derrida's analysis relies heavily on the close readings of philosophical texts in which he uncovered this logic, most notably his reading of Jean-Jacques Rousseau in *Of Grammatology*.

The logic of supplementarity uncovers the rules that structure some of the apparent contradictions found in the texts of metaphysics. In the case of an author who courts paradox as readily as Rousseau, the task is particularly demanding. Derrida's diagnosis is that Rousseau wants to resist the conclusions he nevertheless cannot avoid. As a result, Rousseau's descriptions do not match with the declarations that reveal what he wants those descriptions to say. For example, Rousseau wants to identify the origin of language with speech and thereby make writing a "mere" supplement, but speech is itself a substitute for gesture, which is thereby, in a phrase whose apparent incoherence Derrida underlines, the primordial supplement. Derrida argued that instead of distinguishing Rousseau's use of "supplement" as addition from its use as "substitute," one should see the two senses as operating together (pp. 144–165). So, to continue with the example, much of what appeared contradictory in Rousseau's account of the origin of languages is found to arise because Rousseau wanted to locate the origin of language in the languages of the south but found himself having to draw constantly on the supplementary principles that he had associated with the languages of the north. The languages of the north were, therefore, not simply an external addition, but an alterity that must have been lodged within the system from the outset.

Derrida has exhibited the logic of supplementarity in other metaphysical texts. For example, in *Speech and Phenomena* (1973), Derrida located this operation in Edmund Husserl's account of language. Derrida identified a double tendency in Husserl, like that found in Rousseau. On the one hand Husserl wants to separate indication from solitary life, the strata of expression. On the other hand there are suggestions in Husserl's text that indication is constitutive of expression. The deconstruction of Husserl performed by this double reading is not a critique any more than the reading of Rousseau is. Neither thinker is criticized for failing to recognize the logic of supplementarity as such, not least because this logic has to be understood in terms of what metaphysics represses. The effacement of the primordial supplement is the condition of metaphysics, which thus can no longer be seen as a unity, as it was for Heidegger.

THE TRACE

That *trace* is another notion that Derrida employed as part of his contestation of the tradition of Western metaphysics understood in terms of the priority of presence is clear from his use of the phrase "a past that has never been present" to explicate it. The phrase itself is already found in Maurice Merleau-Ponty's *Phenomenology of Perception*, where it describes the unreflective fund of experience on which reflection draws. Derrida adopted the phrase in "Violence and Metaphysics" to explicate the notion of trace in the work of Levinas, who immediately introduced it into his own account.

Derrida employed various strategies to show that the trace challenges conventional thought. For example, in *Of Grammatology* (1976), when he introduced the concept of an originary trace or arche-trace, he underlined that it represents a contradiction because a trace, which is ordinarily possible only as an effect, is here posited as an origin. The point is to problematize the language and procedures of transcendental philosophy, especially transcendental phenomenology, on which the thought of the trace nevertheless depends for its articulation. This was already Levinas's aim when he appealed to the trace in his account of the possibility of ethics in terms of the face of the Other. The trace is more than a sign of remoteness; it is an irrecuperable absence. Levinas was serving notice that the face surpasses the limits of phenomenology and yet can be approached only through phenomenology. Similarly, even though Derrida makes Freud's failure to apply the effaceability of the trace to all traces a critical element of his reading, at the same time he explicitly recognizes Freud's unconscious as transcending transcendental phenomenology, just as the structure of delay in the sense of deferred effect (*Nachträglichkeit*) cannot be construed as a variation on the present.

These examples show how in the 1960s Derrida developed his account of the trace by gathering together the thought of such thinkers as Levinas and Freud, but he subsequently moved away from this largely parasitic approach. Most notably in *Cinders*, Derrida took the impossible thought of the trace to a different level by explicating it as ashes, with clear reference to the Holocaust. In this way the trace comes to define our epoch even more definitively, than when he drew on the

thinkers who, as he had put it earlier, had helped to define our epoch.

CRITICISM AND RESPONSES

If deconstruction's initial impact within the United States has been strongest in literature departments, this is in part because Derrida's conviction that absolute univocity is impossible is more readily welcomed by literary critics, who have always celebrated the multiplicity or meaning, than by philosophers, whose discipline has tended to encourage a reduction on controlling of equivocity. Whereas the dominant tendency in philosophy has been to mark different uses of a term in an effort to control the ambiguity, the deconstructive approach is to question the basis of any attempt to limit the associations of language. This approach has sometimes been confused with an invitation to so-called free play, in the sense of arbitrariness in interpretation, although Derrida has often rejected this way of reading his work. In exploring equivocity, Derrida is acknowledging and not ignoring the ambiguity of words. In the literary context the constraints of deconstruction are sometimes neglected for the freedom of literacy experimentation. This is less common in Derrida than in some of his followers, but it has given ammunition to the critics of deconstruction.

The most persistent criticism of Derrida arises from his claim in *Of Grammatology* that "there is nothing outside the text" (p. 158). This has sometimes been understood to mean that all reference to the social and historical context is ruled out, and even that the text has no referent. It is easy to show that Derrida has never practiced such an extreme aestheticization of the text. What he did mean is explained in "Living On," in which he sets out the concept of a text as a differential network that overruns all the limits assigned to it (p. 84). This, the so-called general text, is not conceived as a totality. It does not have an outside, anymore than it has an inside. As Derrida explained in the 1988 afterword to *Limited Inc.*, there is nothing outside context, which is almost the opposite of what he is often accused of saying by many who do not share his philosophical background in phenomenology, psychoanalysis, or structural linguistics and yet fail to make allowance for that fact.

One of the most persistent criticisms raised against Derrida in the 1970s and 1980s was that his thought was ill-placed to address ethical and political issues. The understanding, widespread at that time, that Derrida's 1964 essay "Violence and Metaphysics," subsequently reprinted with revisions in the collection *Writing and Difference*, was critical of Levinas for his evocation of ethics after Nietzsche and after Heidegger's *Letter on Humanism* seemed to block him from making any such contribution. This interpretation has now been abandoned in the face of Derrida's repeated invocations of Levinas in the course of his own efforts to contribute to an understanding of ethics. It is here that a form of deconstruction as the exploration of certain aporias has come into his own. Derrida takes up the idea of a duty to go beyond one's duty. So, for example, in *Given Time* Derrida introduced the aporia of the gift whereby a gift is only a gift and not a form of exchange if there is no return on the gift. Derrida pursued these conditions to the point where even being aware that one is making a gift of something would constitute a form of return, thereby making the gift impossible. Parallel studies of hospitality and forgiveness followed. However, it should always be remembered that, for Derrida, the impossibility of the gift or of hospitality, for example, does not mean that giving and hospitality do not happen. It means rather that they are singular events that exceed, and so cannot be explained in terms of what precedes them.

See also Deconstruction; Freud, Sigmund; Heidegger, Martin; Husserl, Edmund; Language; Levinas, Emmanuel; Merleau-Ponty, Maurice; Metaphysics; Nietzsche, Friedrich; Phenomenology; Rousseau, Jean-Jacques.

Bibliography

A complete list of works by Derrida in French and English up to 1992, established by Albert Leventure and Thomas Keenan, is available in *Derrida: A Critical Reader*, edited by D. Wood (Oxford: Blackwell, 1992).

DERRIDA'S MAJOR PHILOSOPHICAL WORKS

De la grammatologie. Paris: Édtions de Minuit, 1967. Translated by G. Spivak as *Of Grammatology* Baltimore: The John Hopkins University Press, 1976.

La carte postale. Paris, 1980. Translated by A. Bass as *The Post Card* Chicago: Chicago University Press, 1987.

La dissémination. Paris, 1972. Translated by B. Johnson as *Dissemination* Chicago: Chicago University Press, 1981.

La voix et le phénomène. Paris, 1967. Translated by D. Allison as *Speech and Phenomena* Evanston, IL: Northwestern University Press, 1973.

L'écriture et la différence. Paris, 1967. Translated by A. Bass as *Writing and Difference* Chicago: Chicago University Press, 1978.

Glas. Paris, 1974. Translated by J. Leavey, Jr., and R. Rand as *Glas* Lincoln, NE: University of Nebraska Press, 1986.

Introduction to *L'origine de la géométrie* by Edmund Husserl. Translated by J. Leavey, Jr., as Edmund Husserl's *Origin of Geometry*. Lincoln, NE: Nebraska University Press, 1989.

Limited Inc. Translated by Samuel Weber and Jeffrey Mehlman. Evanston: Northwestern University Press, 1988.

"Living On." In *Deconstruction and Criticism*. New York: Seabury Press, 1979, pp. 75–176.

Marges de la philosophie. Paris, 1972. Translated by A. Bass as *Margins of Philosophy* Chicago: Chicago University Press, 1982.

Psyché: Inventions de l'autre. Paris: Galilée, 1987.

WORKS ON DERRIDA

Bennington, G., and J. Derrida. *Jacques Derrida*. Paris: Seuil, 1991. Translated by Geoffrey Bennington as *Jacques Derrida* Chicago: Chicago University Press, 1993.

Direk, Z., and L. Lawlor, eds. *Jacques Derrida: Critical Assessments of Leading Philosophers*. 3 vols. London: Routledge, 2002.

Gasché, R. *The Tain of the Mirror*. Cambridge, MA: Harvard University Press, 1986.

Lawlor, L. *Derrida and Husserl: The Basic Problem of Phenomenology*. Bloomington, IN: Indiana University Press, 2002.

Llewelyn, J. *Derrida on the Threshold of Sense*. London: Macmillan, 1986.

Robert Bernasconi (1996, 2005)

DE SANCTIS, FRANCESCO

(1817–1883)

Francesco De Sanctis, the Italian liberal politician and political and literary critic, was born near Naples. Although trained for the law, he turned to the study of Italian culture. He taught at the Military School of Naples, but his participation in the Revolution of 1848 led to his dismissal, a three-year prison sentence, and banishment. He taught and lectured in Turin and Zürich, and returned to Naples in 1860 as governor of the province of Avellino. As director of the Ministry of Public Instruction he brought scholars of great repute to the University of Naples and fought for the secularization of the public schools. After becoming editor of the newspaper *Italia* in 1863, he continued to champion reforms and helped to establish the modern Italian tradition of combining philosophy and worldly affairs. In 1868 De Sanctis returned to literary criticism. Several years later he completed his *Storia della letteratura italiana* (History of Italian literature). He accepted the chair of comparative literature at the University of Naples in 1872, but in 1877 he resumed his political career as organizer of a liberal opposition party, vice-president of the Chamber of Deputies, and minister of public instruction.

De Sanctis developed no systematic aesthetics or political philosophy. His principles of criticism are implicit in his essays. Literary truth, for De Sanctis, is realized in form, but literature's connection with political and social life is the substance of its meaning and the true source of formal beauty. Form transforms an idea into art and is the instrument by which artistic truth is achieved; it is art itself. Content and ideas are, for artistic purposes, without truth. Form provides truth, artistic integrity, the capacity to project an experience or idea so as to bring it subjectively alive for an observer. It does so successfully when it is naturally wedded to the content and seems fused with it. Successful form is derived from the concrete vision of the poet as he reflects on a living experience of the language and forms of his age. This tie between the artist and his immediate image is the deepest source of true art. The language and ideas of art spring from and are shaped by the social and historical events against which they act in the mind of the artist. De Sanctis sought, by grasping history and language, to grasp the work of art as conceived by the artist. History, and specifically political history, provides the framework in which ideas are tested against each other and find concrete representation in artistic form.

Traditional criticism saw technical skill as the essence of poetry, but poetry is involved with the values of the moral, historical, and social orders it expresses and reflects. The philosophical commitments of the poet, his moods and personal objectives, are the stimuli, the raw materials from which an ordered piece of art is shaped. The essence of art is form, but form into which content has passed and fulfilled itself.

De Sanctis believed that the poet must be immersed in the life of his national community. The subject and object of art is the human being. The artist must study man, exhort him, laugh at him, understand him. The artist's manner of picturing human life gives art its truth; this truth is gained by mastery of the language of the age and absorption of its combinations and formal possibilities.

Although art is measured by aesthetic criteria, as a historical phenomenon it is subject to social and moral considerations. Therefore, De Sanctis was led from literary criticism to literary history to the history of Italian culture and ultimately to the relation and debt of Italian culture to Italian politics.

Politics, De Sanctis believed, is a reflection of the moral fiber of a nation. Political activities reflect a wider cultural context and have a special responsibility for that culture, through the power to stimulate or repress it. Politics is a national dialogue between the various sectors of the population. The capacity of the popular classes to participate in and guard a national political organism, to preserve its morality in the face of the tasks of national destiny, to absorb the style and content of past national leaders imprints the national style and goal on political behavior.

Many of De Sanctis's political essays are exhortations, expressions of concern over apathy and loss of morality in political life, as well as attempts to express the inner urgings of Italy. For De Sanctis morality and culture were intimately connected. Moral political activity carried out Italy's destiny, which its previous culture had marked for restored greatness. The politics of a great nation reflects its culture and is perpetually open to self-renewal through the participation of the bearers of that culture. If they cease to participate in the nation's political activity, the culture breaks down, and politics becomes immoral, politicians self-aggrandizing, and the people apathetic.

See also Aesthetics, History of; Aesthetics, Problems of; Art, Truth in; Political Philosophy, History of.

Bibliography

PRIMARY WORKS

Scritti rarii inediti o rari, 2 vols. Naples, Italy: Morano, 1898.

Scritti politici, 3rd ed. Naples, Italy: Morano, 1900.

Saggi critici, 3 vols. Bari, Italy: Laterza, 1952.

La scuola liberale e la scuola democratica. Bari, Italy: Laterza, 1953.

Storia della letteratura italiana (History of Italian literature), 2 vols. Translated by Joan Redfern. New York: Barnes and Noble, 1968.

SECONDARY WORKS

Cione, Edmondo. *Francesco de Sanctis,* 2nd ed. Milan: Casoni, 1944.

Croce, Benedetto. *Estetica (Aesthetics),* 2nd ed. Translated by Douglas Ainslie. London: Macmillan, 1922.

Davis, John A., and Paul Ginsborg, eds. *Society and Politics in the Age of the Risorgimento: Essays in Honour of Denis Mack Smith.* Cambridge, U.K.: Cambridge University Press, 1991.

Holliger, Max. *Francesco De Sanctis: Sein Weltbild und seine Ästhetik.* Freiburg: Paulusdruckerei, 1949.

Landucci, Sergio. *Cultura e ideologia in Francesco De Sanctis.* Milan: Feltrinelli Economica, 1964.

Muscetta, Carlo. *Francesco De Sanctis.* Rome: Laterza, 1975.

Russo, Luigi. *Francesco De Sanctis e la cultura napoletana.* Venice: La Nuova Italia, 1928.

Irving Louis Horowitz (1967)
Bibliography updated by Philip Reed (2005)

DESCARTES, RENÉ
(1596–1650)

In *Discourse on Method* (1637), his first published work, French philosopher and scientist René Descartes combined an intellectual autobiography with a popular presentation of the system he was to develop more rigorously in his *Meditations* (1641) and *Principles of Philosophy* (1644). *Meditations* begins with a radical attempt to doubt all past beliefs, but finds a proposition that resists doubt in the existence of the self as a thinking thing. It then uses this initial certainty as a basis for arguing that God exists, that mind and body are distinct, and that we can achieve certainty in the sciences if we assent only to clear and distinct ideas, provided we have first shown that God would not deceive us about those ideas. The *Principles* uses the metaphysics and epistemology laid out in the *Meditations* as the foundation for an ambitious attempt to provide a scientific account of the entire world.

CHILDHOOD AND FORMAL EDUCATION (1596–1618)

Descartes was born on March 31, 1596, in the village of La Haye, in Touraine, at the home of his maternal grandmother, with whom he lived after his mother's death in 1597. His father, Joachim Descartes, was a member of the gentry and a councilor in the parliament of Brittany whose duties required him to spend several months each year in Rennes. When René was four, his father remarried and moved to Rennes; René and two older siblings remained with his grandmother. We do not know much about his earliest years, but it appears that he was never close to his father, either as a child or as an adult. His grandmother died when he was about fourteen.

At ten, Descartes entered the Royal College at La Flèche, founded two years earlier by Henry IV and run by the Jesuit order. The first five years of the program were guided by the ideals of Renaissance humanism, and thus were devoted to studying Latin, Greek, and classical literature (especially Latin and the works of Cicero). The last three years were dedicated to instruction in numerous subjects, including:

- Thomistic-Aristotelian philosophy, including dialectic (Aristotle's *Organon*), natural philosophy (Aristotle's *Physics*; *On the Heavens* [*De caelo*]; and *On Generation and Corruption* [*De generatione et corruptione*], book I);
- Mathematics (arithmetic, geometry, and topics in applied mathematics, such as astronomy);
- Metaphysics (Aristotle's *On Generation and Corruption*, book II; *On the Soul* [*De anima*]; and *Metaphysics*);
- Moral philosophy (Aristotle's *Nicomachean Ethics* and the work of Jesuit casuists).

Study of Aristotle and Aquinas made extensive use of late-sixteenth- and early-seventeenth-century commentaries, especially those by the Jesuits at the University of

Coimbra. Sometimes the curriculum ignored aspects of Aristotle's thought difficult to reconcile with Christian doctrine. Instead of reading the theological portions of Aristotle's *Metaphysics*, whose remote God is not the creator of the universe, students read a creationist treatise from the first century CE, *On the World* (*De mundo*), which was mistakenly ascribed to Aristotle. In other cases the problems were faced. Thomas Aquinas had interpreted Aristotle's teaching about the soul as consistent with the Christian doctrine of personal immortality. In the sixteenth century Pietro Pomponazzi had argued persuasively that Aquinas misread Aristotle and that Aristotle in fact held the soul to be mortal. Pomponazzi himself did not deny immortality; his official position was that philosophy could neither prove nor disprove the immortality of the soul, but that revelation made it certain. However, since he tended to identify Aristotle's views with those views that human reason would naturally reach unaided by revelation, he came perilously close to the doctrine of double truth associated with the Averroist tradition—the idea that philosophical truth and theological truth may conflict irreconcilably. The Jesuit curriculum called for the teachers in its colleges to attack the authority of such commentators.

Descartes may have drawn an unintended conclusion from these disputes. When he first began to work out his own theory of knowledge, he wrote, "Whenever two people make contrary judgments about the same thing, it is certain that at least one of them is wrong, and it seems that neither of them has knowledge. For if one had a certain and evident argument, he would be able to propose it to the other in a way which would in the end convince his intellect" (Adam and Tannery X, p. 363). The persistence of the dispute about whether immortality was consistent with an Aristotelian theory of the soul probably encouraged Descartes to develop his own anti-Aristotelian theory. His studies in mathematics may also have encouraged skepticism about Aristotelian natural philosophy. At La Flèche the teachers used the texts of the Jesuit mathematician Christopher Clavius (1538–1612), who argued that mathematics was superior to the other supposed sciences, because it succeeded in eliminating all doubt, and that the other so-called sciences demonstrated their uncertainty by their inability to elicit consensus.

At the time, the Jesuit colleges dominated secondary education in France and had an immense influence on the formation of a generation of leaders in politics, philosophy, and religion. Their primary mission was to combat the Protestant heresy, but they also reciprocated the King's support, defending his claims to absolute power. The colleges required total immersion in the Jesuit educational program, permitting little contact with the outside world (parents included). Descartes had mixed feelings about this education. In his *Discourse on Method* he was quite critical of it, claiming it had not provided him with what he had hoped for—clear and certain knowledge of everything useful in life—but had instead left him embarrassed by many doubts and errors. It seemed to him that he gained nothing from his studies but an increasing awareness of his own ignorance. Still, he was careful to say he thought his school had given him as good an education as was then available, and he later recommended it to friends for their children.

Descartes completed the program at La Flèche in 1614. In *Discourse* he wrote that as soon as he was old enough to leave the control of his teachers, he completely quit the study of letters and sought what knowledge he could find either in himself or in the great book of the world. This is not entirely true. We do not know much about what he did in the years between 1614 and 1618, but we do know that he completed a degree in law at the University of Poitiers in 1616. His earliest biographer, Adrien Baillet, reports that Descartes spent the first year after leaving La Flèche in St. Germain-en-Lay, a village outside Paris. The Royal Gardens there contained remarkable statues, designed to move, play music, and even speak. Seeing these machines mimic the behavior of living, intelligent creatures may have helped make plausible to Descartes his later doctrine that animals are nothing but machines.

INFORMAL EDUCATION: ENCOUNTERING BEECKMAN (1618–1619)

In the summer of 1618 Descartes left France to join the army of Maurice of Nassau in the Netherlands. His legal studies would more naturally have led to a career in the law or government, possibilities he considered at various times in the next several years. But in 1618 military life was more enticing, with its opportunities to travel and to learn about the practical application of scientific theories. Maurice encouraged scientific research and employed one of the leading scientists in the Netherlands, Simon Stevin, to oversee his army's education in military technology. Among Stevin's scientific accomplishments were an experimental refutation (anticipating Galileo Galilei) of the Aristotelian theory that heavy bodies fall faster than light ones and his discovery of the "hydrostatic paradox": that the downward pressure of a liquid depends only on the height and base of its vessel and is independ-

ent of its shape. As a military engineer, Stevin developed a system of sluices to flood the fields, giving the Dutch a critical (if Pyrrhic) means of defense against invading armies. In 1618 the Dutch were still enjoying an extended truce in their war for independence from Spain, so Descartes saw no combat. In a letter from this period he wrote that he spent his time learning drawing, military architecture, and Dutch. But he also mentioned plans to write books.

The stimulus for this last ambition was his encounter with Isaac Beeckman, a Dutch scientist, several years his elder, whom he met in November 1618. In a letter to Beeckman, Descartes credits Beeckman with having roused him from his laziness, recalled to his mind the learning he had almost forgotten, and brought him back to serious pursuits. By the spring of 1619 Descartes was contemplating two works, one in mechanics, the other in geometry. In his enthusiasm, Descartes wrote to Beeckman and promised to embrace him as "the promoter and first author of my studies" (Adam and Tannery X, p. 162). Their later relations were not so cordial. In 1630 Descartes was to write Beeckman complaining that the older man was claiming too much credit for having been Descartes's teacher. Though they later reconciled, Descartes makes no mention of Beeckman in the *Discourse*.

Why was Descartes's encounter with Beeckman so important? First, as Beeckman put it, they shared a desire to "combine physics with mathematics in an exact way" (Gaukroger 1995, p. 69). Beeckman had been working on problems in this manner for some years and was delighted to find a like-minded colleague. One result was a short treatise that Descartes wrote on music, attempting to work out a mathematical relationship between the various sounds that appear pleasing to us in combination. He may have begun this treatise before he met Beeckman, but he finished it in December and presented it to Beeckman as a gift on New Year's Day, 1619.

Their more normal pattern of interaction was that Beeckman would set up Descartes with a problem in mechanics or some related area, and invite him to solve it. For example, Beeckman recognized that bodies falling freely in a void would accelerate uniformly. So he posed the following problem: Suppose that a body moving in a void will move eternally (in opposition to Aristotelian physics, Beeckman assumed a version of the principle of inertia). Suppose further (again in opposition to Aristotle) that there is a void between a falling stone and the Earth, and that the stone covers a given distance in a given time. How far will it fall in half that time? As early as 1604 Galileo had worked out the correct law governing the free fall of bodies, which implied that the velocity of the falling body is proportional to the duration of time that it falls, but he did not publish this result until 1638. Descartes concluded that the velocity was proportional to the distance covered, a mistake that Galileo had also made in his first attempt to solve the problem. What is important here is that Beeckman was encouraging Descartes to engage in the Galilean project of discovering laws governing the motion of bodies, expressible in mathematical formulas.

Another area where Beeckman influenced Descartes involved his program of explaining macroscopic physical phenomena in terms of the mechanical properties of the microscopic particles composing them. This program—generally now called "the corpuscularian hypothesis" or "the mechanical philosophy"—had connections with ancient atomism, but differed from atomism in important respects. It did not assume that the component particles were indivisible, and as Descartes was to develop it, it did not assume the existence of a void. Moreover, whereas ancient atomism had regarded the size and shape of the atoms as the primary explanatory factors, the corpuscularians emphasized the speed of the particles and direction of motion.

One problem that Beeckman set for Descartes was to explain in corpuscular terms Stevin's hydrostatic paradox. Descartes postulated that the weight of the column of water can be reduced to the force exerted by its particles in their tendency to downward motion and that each particle of water on the bottom of the container is connected with a particle on the surface by a unique line of particles along which the force (tendency to motion) is transmitted. It is surprising that Descartes thought this explanation worked. It seems plausible where the area at the surface and the area at the base are equal in size, but not in the cases that most require explanation, cases where the area at the surface is smaller (or larger) than the area at the base.

These examples give us an idea of the sort of study that Descartes might have included in the work on mechanics he was contemplating. His work in mathematics seems to foreshadow his discovery of analytic geometry, his most enduring contribution to the sciences. In a letter to Beeckman in March 1619, Descartes excitedly wrote that he hoped to discover "a completely new science," one that would "provide a general solution of all possible equations, involving any sort of quantity, whether continuous or discrete" (Adam and Tannery X, pp. 156–157). The path to solving these geometric or

arithmetic problems seems to have involved using complex instruments, "proportional compasses." Descartes devised compasses that not only solved the problem of dividing an angle into any number of equal smaller angles, but also solved cubic equations of varying degrees of complexity. He had not yet formulated the program, which he was to develop in his *Geometry*, of using algebraic means to solve geometric problems. But he had taken a step toward unifying arithmetic and geometry.

FINDING A VOCATION IN GERMANY (1619–1620)

In the spring of 1619, as the Thirty Years War was just beginning, Descartes set out for Germany to join the army of Maximilian of Bavaria, the leader of the Catholic League. In the *Discourse* he tells us he attended the coronation of Ferdinand II as Holy Roman Emperor. As he was returning to the army, the onset of winter detained him in a place where he had no one to talk to and no cares to trouble him. Shut up all day in a stove-heated room, he was alone with his thoughts. This isolation produced the first of two major turning points in his life. In a document now lost that Baillet saw and preserved (apparently in a mixture of paraphrase and quotation), Descartes wrote that on November 10, 1619, while "full of enthusiasm," he discovered "the foundations of a wonderful science." Descartes left behind conflicting indications of what this discovery was.

The account in the *Discourse*, written seventeen years later, implies that Descartes's discovery involved a decision that to make firm judgments in the sciences, he would have to rid himself of all his previous opinions and reconstruct his system of beliefs on new foundations, accepting nothing he had previously believed until he had squared it with reason. Too much of what he believed was based on uncritical acceptance in his youth of the opinions of others. The document Baillet saw—an account of three dreams that Descartes reported having had on the night of November 10, probably written not long after the event—doesn't suggest a project of ridding himself of all his past opinions. Nor does Descartes's earliest methodological work, *Rules for the Direction of the Mind*, which he apparently began around this time.

Of the three dreams, the most important was the last. In that dream Descartes found two books on a table. One he describes as a dictionary; the other was an anthology of poetry. When he opened the anthology, he found a poem by Decimus Magnus Ausonius (c. 310–395) that opens with the sentence *Quod vitae sectabor iter?* ("What path shall I follow in life?") Descartes said that while he was still asleep, he recognized he had been dreaming and began to interpret his dreams. He construed his discovery of the poem by Ausonius as indicating that he must choose the proper direction for his life. And the dictionary, which we should probably think of as more like an encyclopedia, he interpreted as representing a collection of all the sciences. The dream as a whole he took to indicate that the path he should choose in life was to pursue the sciences and demonstrate their fundamental unity. A fragment preserved from this period expresses that idea vividly: "The sciences now are masked, but if the masks were taken off, they would appear most beautiful. Someone who sees how the sciences are linked together will find them no harder to retain in his mind than the series of numbers" (Adam and Tannery X, p. 215). Descartes's vocation was to unmask the sciences. The ambition to construct a unified system of all scientific knowledge was to guide him for the rest of his life.

RULES FOR THE DIRECTION OF THE MIND (EARLIEST STAGES, 1619–1620)

Among the unfinished works Descartes left behind at his death was a treatise on methodology, which he apparently worked at, off and on, between 1619 and 1628: *Rules for the Direction of the Mind* (*Regulae*, for short). He intended the *Regulae* to be a three part work, each part of which would consist of twelve rules. Although he completed only the first part and about half of the second part, this is the most substantial work we have from the period before 1629. Its parts often seem inconsistent with one another, apparently reflecting different stages of the work's composition and the lack of any unifying revision. Nevertheless, the work sheds light on Descartes's development and later thought. Our best current theory of its composition—resulting from the analyses of Jean-Paul Weber (1964) and John Schuster (1980)—holds that Descartes wrote a part of Rule Four first, perhaps before the night of the three dreams, that he completed most of Part One sometime during the period from 1619 to 1620, and that he then set the work aside for several years, returning to it in the period between 1626 and 1628, when he added Rules Twelve through Twenty-One. After that, he abandoned the work, for reasons we can only guess.

The second half of Rule Four ("Rule IV-B," as it is now called, beginning at the bottom of Adam and Tannery X, p. 374) recounts Descartes's first investigations of mathematics and his disappointment with the ancient mathematicians. He found in them many propositions about numbers that he recognized as true after doing his

own calculations and many conclusions about figures that his authors reached by logical arguments. But they did not explain why these things should be true or how they had discovered them. Descartes conjectured that the ancient mathematicians possessed an algebraic method of discovery they concealed because it made the discovery of mathematical truths too easy. They feared that revealing their method would diminish people's respect for their accomplishments. Rule IV-B is entirely concerned with the project of developing a general method for discovering mathematical truth.

The second stage of the *Regulae* is more ambitious, aiming to formulate a methodology that applies to *all* the sciences. They are all, he says, "nothing but human wisdom, which always remains one and the same, however much its objects may differ" (Adam and Tannery X, p. 360). The sciences are so interconnected and interdependent that it is easier to learn them together than separately. What someone seeking truth in the sciences must first do is to consider how to increase his "natural light of reason," the cognitive abilities he is naturally endowed with. ("Ingenium," the term traditionally translated as "mind" in the title, *Regulae ad directionem ingenii*, might more happily be translated as "native cognitive powers.") Negatively, this means that we should rely only on intuition and deduction—that is, on propositions whose truth we can see distinctly, with certainty, when we attend to them carefully, without being confused by what our senses and imagination tell us, and on propositions that can be inferred from propositions of the first kind by a process of inference equally clear and certain.

Descartes claimed that the only genuine sciences discovered thus far were arithmetic and geometry. But he denied that these were the only areas where we could achieve absolute certainty by intuition and deduction. There are more intuitively certain truths than most people suspect. He gave as examples the propositions that he exists, that he thinks, and that a triangle is bounded by just three lines, among others. If we make proper use of such truths, not mixing them up with probable assumptions, we will be able to extend the certainty of mathematics to other areas. This was an attack, not only on reliance on the senses or imagination, but also on the scholastic use of "probable syllogisms," whose premises needed only the support of a majority of the wise to be acceptable. Descartes thought that in difficult matters the minority is more likely to be right than the majority.

Positively, Descartes's central message is that we must conduct our investigations in an orderly way, gradually reducing complex and obscure propositions to simpler ones, until we reach propositions simple enough that we can know them intuitively, that we can *see* their truth without the aid of other propositions. Once we have completed that reduction, we can work our way back, step by step, to the proposition whose truth or falsity we originally wished to determine. Suppose that the problem is to find the three mean proportionals between 3 and 48. We might not have any intuitions about the answer. But if we look for the single mean proportional between those numbers, 12, then we will have reduced the original problem to something more manageable, finding the mean proportionals between 3 and 12 and between 12 and 48. We can see easily enough that 3 is to 6 as 6 is to 12, and that 12 is to 24 as 24 is to 48. Seeing this enables us to see that 6, 12, and 24 are the numbers sought.

The visual metaphor here is deliberate. "Intuitus," the Latin noun translated as "intuition" is derived from a verb, "intueri," whose basic reference was to visual perception, though it was commonly extended to mental acts of consideration and contemplation in classical Latin. Descartes thought that we can learn how to better use our mental power of intuition by comparing it to vision. If we try to look at many physical objects at once, we see none of them distinctly. Likewise if we try to attend to many propositions in a single act of thought. We can improve our vision, both physical and mental, by focusing our attention on one simple object at a time.

Descartes emphasized that intuition is required not only for our knowledge of the premises of our inferences, but also in the inferential process itself. To have scientific knowledge of the conclusion of an inference, we must intuit not only the premises of the inference, but also the connection between the premises and the conclusion. To know by deduction that $2 + 2 = 3 + 1$, we must see not only that $2 + 2 = 4$ and that $3 + 1 = 4$, but also that our conclusion follows necessarily from these premises. We cannot avoid relying on intuition by insisting, with the Scholastics, that our arguments possess formal validity. Descartes accepted the classical skeptical critique of syllogistic reasoning: that it is useless as a means of acquiring knowledge, because the formalization of the argument—the addition of a suppressed conditional or universal premise to transform an enthymeme into a formally valid syllogism—accomplishes nothing. If the suppressed premise is evident, it is unnecessary for the argument's validity. If the suppressed premise is not evident, then all that the formalization of the inference accomplishes is to increase the number of assumptions requiring proof.

WANDERER YEARS (1621–1625)

After the initial burst of energy that produced the earliest stages of the *Regulae*, Descartes appears to have set the project aside for a while and to have produced no significant work. He traveled here and there, returning to France, visiting Italy (perhaps more than once), and finally returning to France for an extended stay in Paris. He sold the property that he had inherited from his mother, using the proceeds to secure a modest but regular income. This freed him from the need to earn a living. He probably made a pilgrimage to Loreto, Italy, fulfilling a promise made after the night of the three dreams. Apparently, he did not visit Galileo when he passed through Florence. But during this period he seems to have made the acquaintance of Marin Mersenne (a member of the Order of Minims, residing at a convent in Paris), who shared his interest in mathematics and the new mechanical philosophy.

In these years Mersenne was preoccupied with arguing against the radical, Pyrrhonian skepticism that Michel de Montaigne had popularized, which he regarded as a serious threat to religion and society. His "refutation," summed up in *La vérité des sciences contre les sceptiques ou pyrrhoniens* (*The Truth of the Sciences*; [1625]), conceded to the skeptic that we cannot have certain knowledge of the essences of physical things, but insisted that we can have certain knowledge in mathematics (including such applied mathematical disciplines as geometrical optics). He also argued that sense experience provided the basis for knowledge of the physical world, so long as it claimed to be no more than knowledge of appearances, not of the essences of things. If Descartes was not familiar with the Pyrrhonian skeptical challenge before his association with Mersenne in the 1620s, he must have been aware of it by then.

LIFE IN LIBERTINE PARIS (1625–1628)

When Descartes returned to Paris in 1625, he encountered a contentious intellectual scene. Not only were men like Mersenne concerned about the threat of Pyrrhonism, but Paris had just seen the trial of Théophile de Viau, a protestant poet whose writings contained suggestions of Lucretian atomism, a celebration of sensuality, and an advocacy of free thought. Only a few years earlier Giulio Cesare Vanini had been burned in Toulouse for spreading doctrines alleged to be materialistic. The "libertines," as these and other freethinkers were called, were accused of holding scandalous religious opinions and of leading a debauched, hedonistic way of life—the natural consequence of their denial of (or skepticism about) the afterlife. One of Mersenne's projects in this period was a lengthy attack on religious unorthodoxy, *L'impiété des déistes, athées et libertins de ce temps* (The Impiety of the Deists, Atheists, and Libertines of Our Time [1624]). Using, no doubt, a very generous criterion for atheism, Mersenne estimated that in Paris alone there were then at least 50,000 atheists (the population of the whole city at the time was only about 300,000).

Whether or not the threat was as grave as Mersenne claimed, it provoked a response that sought to repress any kind of unorthodoxy. In 1624 three men attempted to hold a public debate in which they would have challenged various theses in Aristotelian natural philosophy. There was apparently considerable public interest in the proposed debate, for it is said to have attracted a crowd of eight or nine hundred people. But the Theology Faculty at the Sorbonne prevented it from occurring. The men were banished from Paris on pain of death; the *parliament* prohibited anyone from holding or teaching theses "contrary to the ancient approved authors, and from holding any public debate other than those approved by the doctors of the Theology Faculty" (Gaukroger 1995, p. 136). The penalty for violating this edict would be death.

Descartes does not seem, at this stage of his life, to have engaged in these culture wars. His main preoccupations, apparently, were with solving problems in geometrical optics and resuming work on the *Regulae*. Sometime during these Paris years he evidently discovered the law of refraction known as Snell's law: When light passes from one medium to another, the sine of the angle of incidence is proportional to the sine of the angle of refraction. (Though Willebrord Snell discovered this law before Descartes, Descartes's discovery was independent of Snell's.) Knowledge of this law is required to solve a practical problem in optics, that of finding the anaclastic curve, the shape that the surface of a lens must have to collect parallel rays of light into a single focus. This knowledge was necessary to design a telescope that would provide a clearer image than existing telescopes did. When Descartes was doing this work, the telescope was a recent invention, dating back only to his years at La Flèche. Descartes was quite excited about the potential of the new scientific instrument for extending our knowledge of nature.

In his treatise on *Optics*, published in 1637 with *Discourse on Method*, Descartes tried to explain Snell's law micromechanically, in terms of the tendencies to motion of the particles involved in the transmission of light rays and the laws of motion, which he held applied to tenden-

cies to motion as well as to motions. He also proposed a solution to the problem of the anaclastic curve: that the lenses should have a hyperbolic shape. His new studies in optics also had an impact on his revision of the *Regulae*, providing him with an example of the use of the method that extended its range from mathematics to physics.

THE LATER *REGULAE* (1626–1628)

According to our best current theory of the composition of the *Regulae*, one change that Descartes made when he returned to this work was to add two examples of the method to Rule Eight in a passage now generally designated as "Rule Eight-C" (Adam and Tannery X, p. 393–396).

The first example deals with the problem of determining the anaclastic curve. Earlier rules prescribed that we gradually reduce complex and obscure propositions to simpler ones, till we reach an intuition of the simplest propositions, from which we can then retrace our steps till we achieve intuitive knowledge of the proposition we originally wished to know. Someone who follows this advice is supposed to see, first, that determining the anaclastic line depends on determining the relation between the angles of refraction and the angles of incidence. He will not be able to determine that relation by conjecture or by appeal to experience or by learning it from the philosophers. But he will make progress if he realizes that the relation depends on the change made in the angles by the difference of the media, and that this change depends on how a ray of light penetrates a transparent medium. Knowing how it penetrates that medium requires knowing the nature of light. Understanding light requires, finally, knowledge of what a natural power is in general.

In this passage Descartes did not say what a natural power is. Here he limited himself to claiming that because this is "the most absolute thing in the whole series," the most basic item in the investigation, which does not depend on anything more fundamental, it is something we will be able to grasp intuitively. Once we have done that, we will be able to retrace our steps—using our understanding of natural powers to understand the nature of light, our understanding of light to understand how it penetrates transparent media, and our understanding of light penetration to understand how a change in the medium changes the path of a light ray—until we are finally able to answer the question about the anaclastic curve. Elsewhere in the *Regulae* (Adam and Tannery X, p. 402), Descartes suggests that understanding the concept of a natural power requires reflection on the local motion of bodies—the idea apparently being that bodies always act on one another by transmitting motion from one body to another through contact.

The frequent talk of intuition and deduction in the *Regulae* is apt to suggest—what many people are inclined to believe on other grounds anyway—that Descartes's scientific methodology is wholly a priori. But this is not true, even in the *Regulae*. Although Descartes says in Rule Eight-C that we cannot determine by experience the relation of the angle of refraction to the angle of incidence, the reason seems to have been that the question is too complex to be resolved by an appeal to experience. "We can have certain experiential knowledge," he said, "only of things completely simple and absolute." We might infer that we derive from experience our intuition of what in this case is most simple—of what a natural power is, of how bodies naturally act on one another in cases where this action is immediately intelligible.

When Descartes was developing his theory of light in the *Optics*, he frequently used analogies from experience. The transmission of light from a luminous body to the object of illumination is like the transmission of resistance through a blind man's cane from the object in his path to his hand. The reflection of light from a shiny surface is like the motion of a tennis ball when it bounces off an impermeable surface. Refraction is like the motion of a tennis ball when it encounters a permeable surface. Not everything we observe is as immediately intelligible as these analogical cases are. The movement of iron filings subjected to the power of a magnet is mysterious. But it's part of Descartes's scientific program to try to understand such phenomena by reducing them to others that are readily intelligible. Insofar as Descartes appeals to common experience in support of these intuitions, his method in the sciences is not wholly a priori.

The other example that Descartes added when he returned to Rule Eight he describes as "the most notable of all." Everyone who loves truth in the slightest degree, he said, should, set himself, once in his life (*semel in vita*), the task of determining what truths human reason is capable of knowing and what questions are beyond our cognitive powers. This will prevent him from always being uncertain about what the mind can do and save him from wasting time on matters our faculties are not capable of dealing with. Someone who undertakes this task will discover that nothing can be known prior to the intellect, since knowledge of all other things depends on it, and it does not depend on knowledge of them. But though Descartes asserted the priority of the pure intellect, he also acknowledged that we have other instruments of knowledge in addition to the intellect. In Rule

Eight-C Descartes said that there are two such instruments: the imagination and the senses. In the text that follows (from Adam and Tannery X, p. 396, 26, to the end of Rule Eight, commonly dubbed "Rule Eight-D" and apparently a rewrite of Eight-C), he added a third instrument: memory.

What is most striking about this passage is its difference from well-known later texts to which it is in other respects quite similar. Three times in Rule Eight (once in Eight-C, twice in Eight-D) Descartes used the phrase "semel in vita" to refer to a project that everyone who wants to use his cognitive powers well should undertake once in his life. Each time he describes the project, with minor variations, in Lockean terms, as one which requires us to determine the limits of the human understanding. Descartes used the same phrase, "semel in vita," in the opening sentence of the First Meditation, also to refer to a project one should undertake once in one's life. But in the *Meditations* the project involves, not determining the limits of the human understanding, but *overthrowing all the opinions he has haphazardly acquired over the years, so he can start anew, from firm foundations*. Three years later, in *Principles of Philosophy*, he used the same phrase, "semel in vita," to again call all seekers after truth to the same overthrow of any past opinion in which they find even the smallest suspicion of uncertainty. As Descartes develops this project in the *Meditations*, it becomes clear that among the opinions to be rejected, at least provisionally, are those involving even the simplest truths of mathematics, which in the *Regulae* were paradigms of certainty. The call to radical doubt, undertaken in the hope of achieving absolute certainty—the most characteristic feature of Descartes's published works—is nowhere present in the *Regulae*.

SEEKING SOLITUDE IN THE NETHERLANDS (1629–1633)

Toward the end of the Paris years Descartes attended a meeting that was to become the second major turning point in his life. The meeting, attended by many of Paris's leading intellectuals, occurred at the palace of the papal nuncio. The speaker was a chemist/alchemist named Chandoux, who attacked Aristotelian natural philosophy as an inadequate basis for chemistry and apparently proposed a mechanistic approach in its place. Contrary to what we might have expected from the ban on criticisms of Aristotle in 1624, most of those present, who included Mersenne and Cardinal Pierre de Bérulle, received Chandoux's speech well. Descartes did not.

Bérulle noticed that Descartes did not share the group's enthusiasm and asked why. After politely trying to excuse himself from saying what he thought, Descartes gave an extended critique of Chandoux, praising his desire to rescue philosophy from "the perplexity of the schoolmen" (Baillet, p. 69), but faulting him for replacing it with something merely probable. If merely probable arguments are allowed, he contended, it is easy to make the false appear true and the true false. He then challenged the company to give him an example of an incontestable truth. When they did, he produced a dozen probable arguments designed to prove its falsity. Then he asked them for an example of an evident falsehood. When they provided one, he showed the falsehood to be credible by another dozen probable arguments. The group then asked him whether there was any infallible way of avoiding sophisms. He replied that he knew of no more certain way than to use the method he commonly followed, which was derived from mathematics, and which he thought sufficient to provide a clear demonstration of all truths.

Descartes's dialectical ingenuity made a deep impression on his audience, especially Cardinal Bérulle, who asked to see him privately. When they met later, Descartes claimed that if he continued his inquiries, the benefits to the public would be considerable. He could achieve results in medicine that would greatly improve people's health and results in mechanics that would greatly lessen people's labor. Bérulle replied that since God had given Descartes this extraordinary talent, he owed it to God and his fellow men to make full use of it. Descartes had been thinking about leaving Paris for some time. This conversation tipped the balance. He resolved to take up residence in the Netherlands, where he would find a more congenial climate and, more important, the solitude that would allow him to meditate without constant interruptions by his friends.

There is nothing up to this point to indicate that Descartes had entertained radical skepticism as a serious possibility. He must have been aware of the debates about Pyrrhonism and of some of the Pyrrhonist literature. But Baillet's account of the Chandoux episode and Descartes's own description of the event in 1631 both suggest a Descartes whose epistemology is very like that of the *Regulae*: dismissive of scholastic philosophy and of any reliance on probable arguments, confident of the certainty of mathematics and of the method he had modeled on his discoveries in mathematics. In his first years in the Netherlands, Descartes's attitude on these issues changed, it seems.

We know from a letter to Mersenne in November 1630 that soon after Descartes moved there (probably during the winter of 1628–1629), he began what he described as "a little treatise on metaphysics," in which he set out to prove "the existence of God and of our souls when they are separate from the body, from which their immortality follows" (Adam and Tannery I, p. 182). This "little treatise" sounds like an early version of the *Meditations*. It was evidently not complete at that point, and we do not know much about its specific content, but from an earlier letter to Mersenne (April 15, 1630) it appears that Descartes worked on this treatise for the first nine months that he was in the Netherlands.

In that letter of April 15, Descartes also described what seems to be a different "little treatise," begun more recently. To provide himself with an extra incentive to finish it as soon as possible, he promises to send it to Mersenne by the beginning of 1633. Mersenne will be amazed, he says, that he is taking such a long time to write a treatise so short that it requires only an afternoon to read. This treatise was apparently focused on physics. But Descartes said that he would not have been able to discover the foundations of his physics if he had not approached them by first trying to know God and himself, and that he had discovered how to prove metaphysical truths in a way more evident than the proofs of geometry. This suggests, for the first time, that Descartes was trying to ground physics on metaphysics, specifically on a metaphysics that focuses on a knowledge of God and the self. It also suggests, for the first time, that there may be something defective about geometrical proofs, that considered apart from a metaphysical foundation, they may be less certain than the metaphysical proofs Descartes discovered.

Why did Descartes then think geometrical proofs might need a metaphysical foundation? The letter of April 15 contains a possible clue. For the first time Descartes stated a doctrine for which he was to become notorious: the creation of the eternal truths. He wrote to Mersenne that in his treatise on physics he would discuss several metaphysical topics; in particular, he would defend the view that the eternal truths of mathematics have been established by God just as a king establishes laws in his kingdom, that they depend entirely on him, no less than does the rest of his creation.

There has been much debate about what this doctrine means and why Descartes held it. But our present concern is its relation to the certainty of mathematics. When Descartes argued, in the First Meditation, that even mathematical truths are subject to doubt, he did not invoke this doctrine that God created them. He simply appealed to the idea of God as an omnipotent being who created him, and who *could*, if God had chosen, have created him so imperfectly that his cognitive apparatus might lead him astray, even in the things that seem most evident to him. In the correspondence and elsewhere, we can see that Descartes thinks a proper understanding of omnipotence would conclude that it requires the ability to determine what the eternal truths are. But Descartes knows this was an unusual, controversial conception of omnipotence, and he did not deploy it in the First Meditation. In the *Discourse*, as he was about to justify his skepticism about mathematics, Descartes would write, "I don't know whether I should tell you of the first meditations that I had [in the Netherlands], for they are perhaps too metaphysical and uncommon for everyone's taste."

Since Aquinas, the dominant view among Scholastics was that God has the power to determine, by his will, what contingent truths are true, but that his will does not determine what necessary truths are true. Those truths were supposed to be grounded in God's intellect, in the ideas he has, not in his will. The meditator of the First Meditation approaches the question of what things are subject to doubt from the perspective of someone just beginning to philosophize, who presumably holds conventional views about what God's omnipotence implies. Moreover, the ability to interfere with his creatures' cognitive faculties would not seem to require God to have the power to create eternal truths. Surely, anyone with enough power to create the world of contingent beings must have the power to make one species of contingent being defective in its perception of necessary truths.

Skeptics nevertheless suggested that God's power might extend to eternal truths and used that thought to ground a doubt about mathematics and any other truths we might think necessary. In his *Apology for Raymond Sebond*, one of Montaigne's arguments for Pyrrhonian skepticism involves the claim that because God's power is incomprehensibly great, we speak irreverently if we say that there is something he cannot do. Among the things that Montaigne suggested it would be impious to say God cannot do are to make two times ten not equal twenty, to go back on his word, to cause a man who has lived not to have lived. Montaigne did not say that God *created* the eternal truths. But he did say it is arrogance on man's part to claim that God *cannot render false* the truths commonly classed as eternal. This category includes not only simple mathematical truths, but also metaphysical truths (such as "The past is immutable" and "Nothing is made of nothing") and moral truths (such as "A perfect being

would not go back on his word"). And he uses these accusations of irreverence and arrogance to justify his claim that we ought to suspend judgment about everything, including these supposedly incontestable truths.

Descartes agreed with Montaigne that we do not do justice to God's power if we say that there is something God cannot do. "We can assert that God can do everything within our grasp, but not that he cannot do what is beyond our grasp" (Adam and Tannery I, p. 146). To say that God's power is limited to what we can comprehend would be rash and disrespectful. Descartes did not want to say that God can render the eternal truths false. They are immutable. But they are immutable because God's will is immutable. So by the spring of 1630 Descartes thought there was a need to ground physics (including the truths of mathematics) in metaphysics, and his perception of this need was connected with his view that God's power is incomprehensibly great. He thought that he had a way to accomplish this grounding, and he began, but did not complete, a draft of a treatise on metaphysics like the *Meditations* that would accomplish that.

THE WORLD (1629–1633)

Descartes never did publish the treatise on physics he referred to in his correspondence as "*The World*" (or "my *World*"), but he used material from it in his *Optics* (1637) and *Meteorology* (1637) and *Principles of Philosophy* (1644). And portions of the work appeared after his death: one under the title *Le monde, ou Traité de la lumière* (The World, or Treatise on Light [1664/1979]), which reproduced the beginning of the treatise, another titled *Traité de Homme* (Treatise of Man [1662/1972]), which reproduced a later part of it. Here the entire physical treatise, as projected in the early 1630s, will be referred to as *The World*, and the title *Treatise on Light* will refer to the opening portion published in 1664.

The World originated in Descartes's concern with the problem of explaining parhelia, the bright spots that sometimes appear in a solar halo, caused by the refraction of sunlight through ice crystals in the atmosphere. In the summer of 1629 a friend had shown him a description of this phenomenon and asked him what he thought of it. Descartes set aside work on his treatise on metaphysics to see what he could make of it:

> My mind is not so strong that I can devote it to many tasks at once, and as I never make any discoveries except through a long train of diverse considerations, I must devote myself wholly to a subject when I wish to investigate some particular aspect of it" (Adam and Tannery I, p. 22–23).

His curiosity about parhelia led him first to inquire into meteorological phenomena in general (including the rainbow, another effect of refraction in the atmosphere) and then to the incredibly ambitious project of explaining "all the phenomena of nature, that is, the whole of physics."

The Treatise on Light begins by arguing that there can be a difference between the visual sensation we have of something and what there is in the object that produces this sensation. The common view, he claimed, is that our ideas are completely like the objects from which they proceed. But there are many experiences which should cast doubt on this. For example, in the case of sound, most philosophers think that the cause of our auditory sensations is a vibration in the air, which does not resemble those auditory sensations at all. Descartes proposed that something similar is true of light.

The first step in discovering the nature of light is to identify the bodies that we know produce light. These seem to be the stars and fire. Because the stars are too remote to be easily observable, Descartes concentrated on fire. When we watch fire burning a piece of wood, we observe that it sets the minute parts of the wood in motion and separates them from one another. It transforms the smallest particles into fire, air, and smoke, and leaves the grosser particles as ash. Scholastic philosophers might have supposed that in addition to these mechanical processes, there is a form of fire or a quality of heat involved. But Descartes (applying Ockham's razor) limited himself to what he saw as necessarily part of the process and did not postulate anything unnecessary to explain the phenomena. Since it is inconceivable that one body should move another except through its own motion, Descartes inferred that the body of the flame is composed of particles so small as not to be observable but moving so rapidly that, in spite of their small size, the force with which they act on the wood is great enough to disperse its particles. The motions of the flame particles cause both the light the flame produces and its heat, depending on which sense organs they encounter. But in neither case is there any resemblance between the cause and the idea.

Descartes observed that there is nothing anywhere in nature that is not changing. He suggested that the changes involved in combustion are not unusual: In many cases the cause of observable changes will lie in the motions of unobservable particles. The principle at the core of the mechanical philosophy is that all physical change is reducible to change of place in bodies, if not in bodies large enough to be observable, then in bodies so

small as to be unobservable. The fundamental differences among bodies are in the size, shape, and motion of their constituent particles.

Descartes distinguished three elements, which he called "fire," "air," and "earth." These are not to be identified with the elements traditionally so-called or with the familiar substances commonly so called. Fire is a very subtle liquid, made up of the smallest, fastest moving particles, which have no determinate size or shape. This permits them to fill the gaps between the particles of the other elements and makes it unnecessary to allow for the existence of a void. Air too is a subtle liquid. Its particles are small and fast-moving compared with those of earth, but large and slow-moving compared with those of fire. They have a determinate size and shape; almost all of them are round. The particles of earth are the largest and have little or no movement. Descartes emphasized that he did not attribute the traditional qualities of heat and cold, moisture and dryness, to his elements. He took those qualities to require explanation themselves, by appeal to the size, shape, and motion of bodies or their constituent particles.

Having explained the basic elements of his physics, Descartes then asked us to imagine that God has created, somewhere in space, a new world, made only of such matter, and has distinguished the different bodies in this world from one another only by the different motions he has given to the different parts of its matter. There is no order in the initial state of this imaginary world. Descartes's project was to show how a world like ours could emerge from this chaos, given only the laws of nature that God established when he created this world. (Genesis, of course, tells us how the world we are familiar with *was* created. Descartes professed to be interested only in exploring other ways God could have done it.)

The language Descartes used here—that God *established* the laws of nature—is reminiscent of the language he used in his letter to Mersenne, where he said that God established the eternal truths (referred to as "laws in nature") as a king establishes laws in his kingdom. The comparison to a king suggests that there is something arbitrary about this act and may also suggest that the laws are subject to change. Descartes wanted the suggestion of arbitrariness, but he did not want the suggestion of mutability. Writing to Mersenne in 1630, he anticipated an objection that Mersenne might encounter when he publicized Descartes's view (as Descartes encouraged him to do):

> They will tell you that if God had established these truths, he could change them, as a king makes his laws. To which one must reply: Yes, if his will can change.—'But I understand them as eternal and immutable.'—And I make the same judgment about God" (Adam and Tannery I, p. 145–146).

In the correspondence of 1630, Descartes invoked the immutability of God's will to explain the immutability of the eternal truths. In *The World* he used it to give content to the laws of nature.

God, "as everyone must know," is immutable. (Presumably, this could have been established by an argument from God's perfection in the 1629 treatise on metaphysics.) This entails that he always acts in the same way. This, in turn, entails that he continues to preserve the objects he created in the same way he created them. That does not mean that things do not change. On the contrary, since God endowed some of these things with motion when he created them, it means that he preserves that motion. This fact, combined with the absence of a void, entails that when beginning to move, bodies also begin to change and diversify their movements through their encounters with other bodies.

Descartes claimed to derive three principal laws of nature from God's immutability: (1) Each part of matter always continues to exist in the same state, so long as encounters with other bodies do not cause it to change. (2) When bodies push against one another, the total quantity of motion is preserved (one body cannot increase the motion of another body without losing as much of its own motion as it transfers to the other body). (3) Although the motion of a body is usually in a curved line, it always *tends to move* in a straight line. Because Descartes denied the existence of a void, he insisted that all motion must be "in some way circular." Since there is no empty space for a body to move into, it can move only by displacing other bodies. Ultimately, each moving body must be part of a chain of moving bodies that forms a closed curve of some sort. Though Descartes deduced these laws from God's immutability, which he presumably knew a priori, he also insisted that they agree well with what we find in experience. Aristotelian physics, he pointed out, assumes that motion will continue only as long as force continues to be applied to the moving object; so it has difficulty accounting for projectile motion.

From these assumptions about the nature of matter and the laws according to which matter moves, Descartes developed a theory about how a world like ours might have evolved from the chaos that he supposed God originally created in the hypothetical new world. Since that

world is a plenum, whose parts must displace other parts to move at all, and all motion must be in a closed curve, matter will naturally organize itself into vortices, masses of matter swirling, whirlpool-like, around a center. The element of fire will tend to concentrate at the centers of the vortices, composing the Sun and the stars; the element of earth will tend to form into large clusters, rotating around the centers of their vortices and carried along by particles of the second element. The planets are formed from such clusters, as are the comets. But the planets remain in one vortex, whereas the comets have motions that carry them from one vortex to another. This clearly heliocentric system implies not only that Earth rotates around the Sun, but also that our solar system is only one of many in the universe, each forming around the various stars, which are no longer embedded in a single sphere, as in the Aristotelian-Ptolemaic cosmology.

Descartes went on to offer explanations for the planets' rotations around their axes, the motions of satellites (moons) around their planets, the movement of the tides, weight, and light. The last he interpreted as resulting from the rotation of the Sun and the matter around it. This generates a radial pressure, which spreads outward from the Sun along straight lines from its center. He enumerated a dozen properties of light that he claimed this theory can account for: that it is propagated from all sides of the luminous body, to any distance, instantaneously, normally in a straight line, but subject to reflection when it encounters a body it cannot penetrate, and subject to refraction when it encounters a medium it can penetrate, and so on. The treatise as it has come down to us does not explain reflection and refraction but instead refers us to his *Optics*, published in 1637. In another work, *Meteorology* (1637), Descartes offered explanations of rainbows and of parhelia.

The work commonly known as the *Treatise of Man* was to have been part of *The World*, though it appears that it would not have come immediately after the *Treatise on Light*. We do not know how many intervening chapters are missing, but we can have a fair idea of their intended contents from the description of *The World* in Part Five of the *Discourse*. Among the topics covered would have been the formation of mountains, seas, springs, and rivers on Earth; the formation of metals in the earth and plants in the fields; and the nature of fire and its various properties, such as its ability to form glass from the ashes of the material it has burned.

The *Treatise of Man* was to have included accounts of both the body and the soul, though only the chapters dealing with the body survive. These chapters begin by asking us to imagine that God created a statue or machine made of earth (the element), which he intended to make as much as possible like us. Just as the *Treatise on Light* tries to show that God could have produced a world that would look just like ours, using only the materials and mechanisms Descartes described, so the *Treatise of Man* tries to show that God could have produced machines that would have looked and behaved just like the human body, using only such means. Though an exception would be the functions that Descartes thinks need to be attributed to the rational soul, notably the intelligent use of language, he claimed to give a mechanistic explanation of all the animal functions that Aristotelian philosophy attributed to vegetative and sensitive souls: the digestion of food; the beating of the heart; the nourishment and growth of members; waking and sleeping; the reception of light, sounds, smells, heat, and so on, by the sense organs; the transmission of ideas of these qualities to the brain; the retention of these ideas in memory; the internal movements of the appetites and passions; and the external movements of the limbs (insofar as their explanation does not depend on the actions of the soul).

We cannot go into these explanations here, but two points about them deserve notice. First, we know from his correspondence that Descartes spent a lot of time during these years dissecting animals to learn anatomy. However much works like the *Regulae* encourage the picture of Descartes as a purely a priori scientist, and however much justice there may be in that picture, it is clear that in practice Descartes believed it was necessary for scientists to do a great deal of data collection. When Mersenne wrote saying that he knew people so dedicated to advancing the sciences that they were willing to make all kinds of experiments at their own expense, Descartes replied,

> It would be very useful if some such person were to write the history of celestial phenomena, according to the Baconian method, describing exactly for us the present appearance of the heavens, without any explanations or hypotheses, reporting the positions of each fixed star in relation to its neighbors, and their differences in size, color, visibility, brilliance, and so on. He should tell us how far this accords with what ancient astronomers have written about it and what differences there are; for I have no doubt that the stars are constantly changing their relative position somewhat, in spite of being called fixed. (Adam and Tannery I, p. 251–252)

Descartes himself regularly offered hypotheses to explain phenomena. But he knew that a satisfactory explanation

of phenomena required good descriptions of the phenomena to be explained and that such descriptions required empirical inquiry. He might also have acknowledged that the necessary empirical inquiries would be guided by intuitions about what the ultimate explanation of the phenomena was likely to be. In this case, his call for a Baconian history seems to be guided by his conviction that the so-called fixed stars are not really fixed.

Second, although it may not be obvious that Descartes's scientific procedure in *The World* exemplifies his method, we can regard it as an extension of the method described in Rule Eight-C of the *Regulae*. Reflection on the problem of determining the anaclastic curve had persuaded Descartes that solving that problem would require understanding refraction, which would require understanding the nature of light and its transmission, and ultimately, understanding what a natural power is. In the *Regulae* Descartes was vague about what a natural power might be, suggesting only that it had something to do with local motion. When he began to write *The World*, his starting point was a similar problem in optics, explaining parhelia, which he must have realized would also involve understanding refraction. He already believed that understanding refraction required understanding the nature of light, and that understanding light required understanding what a natural power is. He then saw that understanding the concept of a natural power requires a full-fledged theory of the nature of bodies and the laws governing their motion, and that getting this right should enable him to explain all kinds of phenomena, both in the heavens and on Earth.

THE GALILEO AFFAIR AND ITS AFTERMATH (1633–1637)

Descartes never published *The World*, despite having worked hard on it for years and having achieved results he was very proud of. In 1616 the Church condemned as false and contrary to Scripture the Copernican doctrines that Earth moves and that the Sun is motionless. It prohibited a book by Paolo Foscarini that taught this doctrine, and suspended Nicolaus Copernicus's *On the Revolutions of the Heavenly Spheres* "until it should be corrected." Also, Cardinal Robert Bellarmine, in a private meeting with Galileo Galilei, ordered him to abandon the Copernican view. Descartes knew that the Copernican view had been censured. He probably did not know about Galileo's meeting with Bellarmine. He seems to have heard rumors that, in spite of the censure, some continued to teach the Copernican view "publicly, even in Rome."

In 1623 a Florentine cardinal friendly to Galileo became Pope Urban VIII. After discussions with the new pope, Galileo got permission to write a treatise on the Copernican system, provided he treat it as a mathematical hypothesis, no more than a convenient predictive device. Galileo apparently decided to test the limits of this permission. He wrote a dialogue in which one participant defended the Copernican theory, another defended the Ptolemaic theory, and a third played the role of uncommitted inquirer. His spokesman in the dialogue does not claim certainty for the Copernican theory. Moreover, he permitted his representative of orthodoxy to have the last word, proclaiming that, however plausible the pro-Copernican arguments might be, we can never know with certainty what the true explanation of the phenomena is. God, in his infinite power and wisdom, might have produced the phenomena in any number of ways. Finally, in the preface he claimed that his work treated the heliocentric theory as "a pure mathematical hypothesis," adopted for astronomical convenience, and that he was writing only to demonstrate that Italians were well aware of the scientific case *for* Copernicanism, and that the prohibition of 1616 had not been issued in ignorance. After long negotiations with the censors, he secured permission to publish his *Dialogue Concerning the Two Chief World Systems* in 1632.

But his precautions proved insufficient. It was evident to careful readers that he had crossed the line between hypothetical consideration and advocacy. In the spring of 1633 he was tried by the Inquisition on the charge of "vehement suspicion of heresy." What this language meant, in this case, was that he had presented views contrary to Scripture as if they were probable, but that there was some doubt as to whether he had the evil intention necessary for conviction of formal heresy. Found guilty in June, he was sentenced to house arrest for the rest of his life and required to abjure his errors.

In November 1633 Descartes wrote to Mersenne that he had tried to buy a copy of Galileo's *Dialogue*, which he heard had been published the year before. But when he looked for it, he learned that it had been confiscated and burned. He was so astonished, he said, "that I almost decided to burn all my papers, or at least, to let no one see them. For I could not imagine that he, who is an Italian, and even (as I hear) favored by the Pope … could have been made a criminal simply because he wanted (as doubtless he did) to establish the movement of the earth" (Adam and Tannery I, p. 270–271). The doctrine of the Earth's movement, Descartes writes, is so connected with the other parts of his own *World* that he could not detach

it without making the remainder very unsatisfactory. If that doctrine is false, "all the foundations of my philosophy are, too, for they demonstrate it quite evidently." When he learned several months later that Galileo had been condemned "even though he pretended that he proposed [the Copernican system] only hypothetically" (Adam and Tannery I, p. 271) he was especially concerned, since he had adopted a similar device himself, representing himself as merely telling a story about how God *could have* created a world like ours, while conceding that Genesis tells us how he *did* create it.

It is not clear what stage *The World* was in at this time. Descartes said that he had been on the point of sending it to Mersenne as a New Year's present, if it could be copied in time. Perhaps it was nearer completion than the surviving parts would suggest. For a while Descartes held out hope that there may be a way to publish it. Perhaps this action of the Inquisition had not yet been ratified by the pope or by a Church council. If so, it may not have the full authority of the Church behind it. But eventually Descartes decided to abandon his treatise for the time being and adopt a different plan. He decided to publish a semiautobiographical treatise on method, to be supplemented by three short treatises demonstrating the power of his method: *Optics*, *Meteorology*, and *Geometry*.

THE *DISCOURSE ON METHOD* AND ITS ESSAYS (1637)

At age forty-one, with only thirteen more years to live, Descartes published his first works. Thus began the public career that would earn him a reputation as the father of modern philosophy. As we have noted, Descartes's *Discourse on the Method of Conducting One's Reason Well and Searching for Truth in the Sciences* is partly autobiographical, but it is not very reliable in this regard, for it omits important events (such as his relation with Beeckman, his three dreams, and his encounter with Chandoux); it projects into the past ideas that Descartes probably had only at a later date (such as the idea of overturning all his past opinions to reconstruct his beliefs on firmer foundations); and it is silent on ideas that Descartes feared might cause his readers to raise objections he did not want to deal with (such as the creation of the eternal truths). Descartes himself warned us not to take his work too seriously as autobiography when he wrote that he was presenting it "only as a history, or, if you prefer, a fable." Descartes wanted us to read his work for its moral, for examples of conduct to imitate or avoid. But he also cautioned us that both fables and history have their dangers: Fables may make us think that something is possible when it is not, and even the most accurate histories, because of their selectivity, may make us conceive plans beyond our powers.

Examples of conduct to imitate would be examples of how to conduct our reason when we seek truth in the sciences. Descartes offered four rules that he said he found sufficient in this search:

> 1) never accept anything as true which he did not know evidently to be true, including nothing in his judgments except what has presented itself so clearly and distinctly to his mind that he had no reason to doubt it;
>
> 2) divide the difficulties he was examining into as many parts as possible;
>
> 3) conduct his thoughts in an orderly way, beginning with the simplest objects, and ascending gradually to the most complex; and
>
> 4) make enumerations so complete and reviews so comprehensive that he was sure he had not left anything out. (Adam and Tannery VI, p. 18–19)

Presented thus baldly, these rules probably do not give enough direction to be very useful. And in a letter to Mersenne in February 1637, Descartes disclaimed any intention to *teach* his method in the *Discourse*. His purpose there was only to *talk* about it, and his purpose in the scientific essays that accompanied the *Discourse* was to show what could be accomplished through its use. Even the essays do not, for the most part, purport to show the method at work. As Descartes explained in a letter to Antoine Vatier in February 1638, "I could not show the use of this method in [the three scientific treatises] because it prescribes an order for investigating things which is rather different from the one I thought I had to use to explain them" (Adam and Tannery I, p. 559). But Descartes made one exception to this generalization. He told Vatier he had given a sample of the method in his discussion of rainbows in the eighth chapter of *Meteorology*.

In that chapter Descartes began by noting that rainbows occur not only in the sky but also in the air near us, whenever there are many drops of water in the air illuminated by the Sun. We know this from our experience with fountains. He inferred from this that the rainbow arises only from the way light rays interact with drops of water, and from there move toward our eyes. Previously in the *Meteorology* he had shown that these drops are round; he also knew, presumably from experience, that the occurrence of a rainbow is independent of the size of the drops.

These reflections suggested an experiment that enabled him to examine the phenomenon close up, in circumstances he could control.

He filled a large, round flask with water and positioned it so that the Sun was coming from behind him as he faced it. Then he situated himself in relation to the flask so that he observed a bright red spot at its bottom. He discovered that a line drawn from his eye to the bottom of the flask made about a 42-degree angle with a line drawn from the Sun to the flask's bottom. Furthermore, no matter how he moved—nearer to the flask or further away, to the right or to the left, even if he made the flask revolve around his head—he always saw a red spot at the bottom, so long as the angle between his line of vision and the line of the Sun's rays remained about 42 degrees. If he increased the angle, the red disappeared. If he decreased it slightly, the spot did not cease to be colored, but divided into two less brilliant parts of different colors (yellow, blue, etc.). From this he inferred that if all the air in that direction were filled with such round drops of water, a red spot would appear in each drop where the angle between the Sun's rays and the line of vision was about 42 degrees, producing a continuous circle of red spots. Similar circles of other colors would be generated in drops that were at slightly more acute angles.

Through further experiments with the flask, Descartes discovered that the red spot did not disappear when the light source was blocked, so long as light was permitted to enter at the top of the flask and leave at the bottom, and so long as certain paths within the flask were not blocked. He inferred that the appearance of red at the bottom was caused by refraction of the Sun's rays as they enter at the top of the flask, their reflection from a point at the back of the flask, and their refraction again at the bottom of the flask as they leave it to move toward the eye. He proposed a similar explanation for the production of the secondary bow, which appeared at an angle of about 52 degrees and had its colors arranged in reverse order. This, he inferred, arises from a combination of two refractions and two reflections.

So far the phenomena Descartes was trying to account for depended essentially on the refractive index of water in relation to air, a figure that he could calculate accurately. And so far his explanation of the phenomena was basically right. But he still had not explained what he called the principal difficulty: Why do only those rays refracted at a certain angle cause certain colors to appear? To resolve this difficulty, he undertook a series of experiments with a prism, a similar object also known to produce a spectrum of colors. The prism differed in various ways from his flask, and these differences enabled him to eliminate as irrelevant certain features of the flask. To produce a spectrum of colors, it is not necessary that the medium through which the light passes have a curved surface, or that the light strike that medium at a particular angle, or that it be reflected, or that it be refracted more than once. But it is necessary that the light be refracted at least once.

At this point Descartes invoked his theory of the nature of light, that it is the action or movement of particles of air (the element), which must be imagined as little balls. These balls have two motions, one in the direction of their propagation, the other rotational. Different degrees of rotational motion produce different color sensations when they strike the eye. The differences in the colors produced when light is refracted arise from the fact that the refractive process imparts different degrees of rotational motion to the light particles. (For further details, see Gaukroger 1995, chap. 6.)

This part of Descartes's explanation has not fared well. But the example remains interesting in a number of respects. First, it illustrates Descartes's second and third rules: dividing a complex problem into as many parts as possible till you reach something simple and easy to understand, and then retracing your steps back to the complex phenomenon you were originally interested in. The complex phenomenon is the rainbow. The simple object is the individual drop of water seen to have one of the colors of the rainbow. By using a model of the simple object, which we can observe close up and manipulate, we can determine the conditions for its being seen as having the color it has, and we can determine how changing those conditions might produce different colors (or no colors at all). We then reconstruct the complex phenomenon from the simple model by recognizing that if we observed a mass of such simple objects in the sky, the ones observed at the right angle for producing a particular color would form a continuous circle of that color, and that other concentric circles of different colors would also be produced at different angles.

But this example also warns us that if we use the concepts of intuition and deduction to analyze our solution of this problem, we need to understand those concepts very broadly. Our understanding of how the simple objects behave involves a priori elements, insofar as we make use of geometry to deal with certain aspects of the problem (such as the shape of the bow). But it also involves numerous appeals to experience. It is by careful experiment that we determine that the same color is produced so long as the same angle is preserved, or that a

double refraction, combined with reflection, is involved in producing the primary bow, or that the refractive index of water in relation to air has the particular value it has. Our ordinary experience of fountains initially suggests a way of breaking the complex phenomenon down into simple elements. Experience is also involved, no doubt, in the theory of matter that Descartes's theory of light invokes. There is no a priori reason why there must be exactly three elements, having the properties that Descartes assumed they have. A priori considerations of simplicity and intelligibility speak for this theory when it is compared with the scholastic forms and qualities. But those considerations would not be sufficient to warrant acceptance of the theory if it were not capable of explaining a wide range of phenomena, as Descartes clearly thought it is.

The rules of the *Discourse*, then, are quite similar to the rules of the *Regulae*, provided that we interpret the concepts of the *Regulae* freely. But one notable feature of the *Discourse* is the absence of any explicit discussion of intuition and deduction. The ghosts of these concepts are present in the first rule, insofar as Descartes advises us to never accept as true anything we don't know to be evidently true, making no judgments except those that present themselves so clearly and distinctly to our minds that we have no reason to doubt them. This excludes reliance on merely probable assumptions. But it does not explicitly mention intuition or deduction. And it suggests a problem we have so far not considered, because so far it has not seemed to arise in the writings we have considered.

Throughout his work Descartes was clearly a foundationalist, at least in the minimal sense that he thought some of our beliefs are based on other beliefs we have, whereas some are not based on others. We can call the ones not based on others *basic beliefs*. Our basic beliefs provide the foundations for our system of beliefs; our derivative beliefs, the superstructure. This metaphor of our system of beliefs as like a building, which has foundations and a superstructure and might collapse if the foundations are not solid, is prominent in the *Discourse* and in the *Meditations*, but is only implicit in the *Regulae*, where Descartes presents arithmetic and geometry as the only genuine sciences yet discovered, superior to all other alleged sciences because of the certainty of their initial assumptions and the care with which mathematicians derive from those assumptions only conclusions clearly seen to follow from them.

But the *Regulae* does not have a criterion for distinguishing the absolutely certain from the merely probable. It assumes that mathematics is more certain than the other sciences because it is concerned with objects so pure and simple that it need make no assumptions that experience has rendered uncertain. At that point, that is all that Descartes thought it necessary to say to justify reliance on the assumptions of mathematics. But in the *Discourse* (and the *Meditations*), he is concerned with a problem his earlier work had not considered. It is not the problem of the creation of the eternal truths, but a different skeptical problem.

We are not born with fully mature cognitive faculties. Rather our faculties develop gradually as we grow to adulthood. While they are developing, we accept, uncritically, many propositions from parents, teachers, and others whose authority we have come to respect. Then we learn, sadly, that these are not perfectly reliable sources. The propositions we accepted in this way can seem quite obvious. Nevertheless, they lack a firm foundation, and we can be mistaken about them, even when they seem most obvious. This reflection gives us a reason, not only for doubting the specific propositions we have accepted from others and everything based on them, but also for wondering whether our cognitive faculties, our basic capacities for distinguishing truth from falsity, are as reliable as we thought they were.

The *Discourse* not only identifies this problem; it offers a solution for it. Descartes was not content, in this work, simply to say that the basic beliefs we acquire through intuition are indubitable or evident. He wanted to show that we perceive some propositions so clearly and distinctly that there is no reason to doubt them, even on a generous conception of what might constitute a reason for doubt. So in Part IV of the *Discourse* he embarked on a project of rejecting as false anything in which he could "imagine the least doubt." This is what is called the method of doubt. He was very permissive in what he counted as a ground of doubt. He was prepared to allow that even "the most extravagant suppositions of the skeptics" provide some ground for doubt. If a belief can survive that permissive a test, we cannot reasonably demand anything more certain as a foundation for our beliefs. If we are to doubt, we must have some reason to doubt. But if we want what survives our attempt to doubt to be absolutely certain, we must be thorough about the attempt; we must allow even the most improbable possibilities to count as reasons for doubt.

This might seem to be a quixotic quest if Descartes had not apparently discovered something that resists his attempts to doubt it: that he, who is engaged in this methodical doubt, and thus is thinking, exists. So we

encounter what is commonly referred to as "Descartes's *cogito*," a label deriving from the Latin version ("*Cogito, ergo, sum*") of an inference that appears in the *Discourse* in French: "*Je pense, donc, je suis*" ("I think; therefore, I exist").

Though there is something very compelling about that inference, it is not clear exactly what Descartes was claiming as his initial certainty. In the *Regulae* he had cited both "I think" and "I exist" as truths known certainly by intuition; if that's their status, then either proposition might be a suitable foundation for demonstrations. In the *Discourse* he seems to be inferring his existence from his thought, as if he can be certain of his existence because he can be certain that he thinks—and, moreover, certain that to think, it is necessary to exist (Adam and Tannery VI, p. 32–33). This suggests that his affirmation of his existence is the conclusion of the following demonstration:

(1) To think, it is necessary to exist.

(2) I think.

(3) Therefore, I exist.

This way of thinking about the *cogito* naturally raises the question of how Descartes can be certain of the premises of this demonstration. The *Discourse* does not explicitly ask that question, but it does have what looks like an answer to it, as far as the first premise is concerned, where Descartes says that he sees very clearly that (1) is true. Though the *Discourse* has not offered any *theory* of intuition, this looks like an appeal to intuition, a faculty whose reliability we might have thought was put in doubt when Descartes questioned the certainty of simple mathematical truths.

The *Discourse* does not even *seem* to answer the question as it concerns the certainty of the second premise. But in a letter that Descartes wrote to Henri Reneri in 1638, the answer seems to be that when we are thinking, we cannot doubt that we are thinking (Adam and Tannery II, p. 38). This may suggest the following argument for the certainty of (2):

(4) When we think, we cannot doubt that we think.

(5) I am thinking.

(6) Therefore, I cannot doubt that I am thinking.

But though Descartes often seems to accept (4), or propositions equivalent to it, there are times when he seems to reject such claims. Earlier in the *Discourse* he had written that "many people don't themselves know what they believe; for the act of thought by which one believes a thing being different from the act by which one knows that one believes it, the one often occurs without the other" (Adam and Tannery VI, p. 23). Moreover, the argument consisting of propositions (4) to (6), if offered as a demonstration of the certainty of (2), looks hopelessly question-begging: it assumes the truth (and certainty) of the proposition whose certainty it claims to prove. So the *cogito* argument of the *Discourse*, in spite of its fame and wide appeal, is problematic. Fortunately, the argument takes a different, and more attractive, form in the *Meditations*, as we shall see below.

The remainder of part IV gives a quick sketch of the argument Descartes would develop more fully and accurately in the *Meditations*. Having found one proposition that he knew to be true and certain, he provisionally formed a general rule: Whatever we conceive very clearly and very distinctly is true. Reflecting on his nature as a doubter, and hence as imperfect, he asked how he could have acquired his ideas of things other than himself. Most of them, he thought, he could have generated himself. But the idea of God is an exception. An imperfect being cannot cause itself to have the idea of a perfect being. So God must be the cause of his idea of God. God, therefore, must exist. To this causal argument he added a version of the ontological argument: If God is a perfect being, as we conceive him to be, then he cannot lack the perfection of existence. Having established the existence of God, he proceeded to argue that because everything real and true in us comes from a perfect being, the general rule he had provisionally adopted is correct: All our clear and distinct ideas must be true. And even those ideas that are not clear and distinct must have *some* foundation in truth. This account of Descartes's metaphysics raises issues that are best pursued in the discussion below of the *Meditations*.

Parts V and VI of the *Discourse* are primarily concerned with Descartes's *World*, which, he wrote, "certain considerations" prevent him from publishing. He tantalizes us with a summary of its contents, omitting any explicit mention of its Copernicanism, but strongly hinting that the Church's condemnation of Galileo is the reason that he could not publish at that time. He did not mention either the Church or Galileo by name, but what he did say must have left little doubt in the minds of informed readers: "People to whom I defer and who have no less authority over my actions than my reason has over my thoughts have disapproved an opinion in physics, published not long ago by someone else" (Adam and Tannery VI, p. 60). Descartes did not say whether he accepted this opinion, but he did say that before the authorities' censure he had not noticed anything in the work "preju-

dicial either to religion or to the state," so nothing would have prevented him from publishing this opinion himself if his reason had convinced him it was true. This censure, he said, made him fear that he might have made some mistake in his own theories. And that, combined with a fear of getting involved in time wasting controversies, made him decide not to publish, at least at that time.

Descartes had clearly not given up all hope of publishing *The World* during his lifetime. He even suggested that he had a duty to publish it: If, as he thought, he was on the path to developing a correct and comprehensive physics, giving an account not only of the heavens, but of all the principal kinds of bodies here on Earth, the potential benefits would be enormous. Such a science would enable us to become "the masters and possessors of nature." It would offer the hope of discovering new ways to maintain our health and prolong our lives. He saw only two obstacles to his achieving this goal: the brevity of life and the lack of observations. Though he presented the foundations of his physics as a priori ("To discover in general the principles or first causes of everything that exists or can exist in the world … , I considered nothing but God alone, who created the world"), he reported that as he proceeded from the first causes, through the first and most ordinary effects deducible from them, to more particular things, he found that the only way he could discover the causes of the particular effects was to construct what Bacon called crucial experiments.

Descartes's principles were so general that there were many ways he could deduce the effects from them. To determine which, among many possible ways to produce the effects, was the one God had chosen, he needed to set up situations where the alternative theories would have different observable consequences. To do that he would need money for research. Part 6 of the *Discourse* was, among other things, an appeal for money from public-spirited citizens who saw the value of his work and wished to aid him. But the whole project of the *Discourse* and its essays was also intended to generate such interest in his project that the Church would feel obliged to permit him to publish his *World* during his lifetime. Failing that, he would publish posthumously.

Other matters in Parts V and VI of the *Discourse* merit more discussion than they can receive here: Descartes's discoveries regarding the circulation of the blood, which he made independently of William Harvey, and his affirmation that the fundamental laws of nature are necessary truths that must be observed in any world that God might have created. Here we must limit ourselves to noting his provocative doctrine that animals are nothing but machines. In the portions of *The World* dealing with humankind, Descartes had tried to show that God could have produced machines that would have looked and behaved just as the human body does, using only matter of the kind Cartesian physics allows and the laws that follow from God's nature. Descartes aimed to give a mechanistic explanation of many different animal functions, all the functions, in fact, that humans share with the lower animals. He did not think mechanism could explain all human activities. Some, notably the intelligent use of language, could be explained only by the presence of a rational soul embedded in the machine. We can be certain from their language use, Descartes thought, that the human-looking bodies around us are inhabited by rational souls. (He was not troubled by the problem of other minds.) But nonhuman animals, which do not display intelligent language use, lack a rational soul; they are nothing but complex machines, lacking even sensations of the kind we have.

This doctrine had a strong impact, most of it in ways that Descartes would not have welcomed. Some thought it absurd to draw such a sharp distinction between humans and the rest of the animal kingdom. Some accused the Cartesians of being cruel to animals, or at least of having no good reason not to be. And some argued that Descartes was right about lower animals, but wrong to think that humans are fundamentally different. They too are nothing but very complex machines.

THE START OF CONTROVERSY (1637–1641)

After publishing the *Discourse* and its essays in June 1637, Descartes spent the next few years responding to criticisms of his work and, toward the end of the period, preparing to publish his *Meditations*. The criticism of the 1637 publications tended to focus, not on metaphysics or epistemology, but on his commitment to mechanistic explanations in science: of light, of the circulation of the blood, of animal behavior. In the early part of this period he tried to reassure friends in the Jesuit order that his work does not contain dangerous innovations. He boasted to Vatier (Adam and Tannery I, p. 564) that the faith had never been so strongly supported by human reasons as it was by his, and that transubstantiation, "which the Calvinists criticize as impossible to explain by the ordinary philosophy, is very easily explained by mine." But by 1640, the Jesuit priest Pierre Bourdin's criticism of his *Optics* had persuaded Descartes that he had to 'go to war with the Jesuits' (Adam and Tannery III, p. 752).

By that point Descartes had already begun revising his "little treatise" on metaphysics (the future *Meditations on First Philosophy*) and planned to circulate it privately among twenty or thirty theologians before making it public so that he could learn from their criticisms what needed to be corrected or added before publication (Adam and Tannery II, p. 622). Descartes told Mersenne that his book on metaphysics was to contain "all the foundations of my physics," but cautioned him not to tell people that, "for those who favor Aristotle might make more difficulty about approving them. I hope that readers will gradually get used to my principles, and recognize their truth, before they notice that they destroy Aristotle's principles" (January 28, 1641; Adam and Tannery III, p. 298).

Descartes was particularly keen to have the Sorbonne's approval of his work. This may seem out of character, for in the *Discourse* he said that since God has given each of us some capacity for distinguishing truth from falsity, he felt obliged not to be content with accepting the opinions of others (Adam and Tannery VI, p. 27). Presumably this is an obligation we all have. But experience had persuaded him that he needed the support of the authorities to get people to read his work carefully and to free himself from having to reply to quibbling, malicious critics (Adam and Tannery III, p. 184, 237–238).

When the *Meditations* first appeared in August 1641, the original plan had changed. Instead of circulating his work first among twenty or thirty theologians to get objections that might lead to changes, Descartes delegated most of the preliminary circulation of the work to Mersenne, who selected a smaller number of critics, not all theologians. Instead of modifying the text in the light of this criticism, Descartes left the text largely untouched, publishing the objections he received and his replies after the main text. Each critic could see the preceding objections and replies in composing his own.

The author of the first set of objections was a Dutch Catholic theologian named Johan van Kater (Johannes Caterus). Mersenne himself is generally credited with having written some or all of the anonymous second and sixth sets of objections. The third, fourth, and fifth sets of objections were written by Thomas Hobbes, Antoine Arnauld, and Pierre Gassendi, respectively. Those were the only objections included in the first edition. When the second edition appeared in 1642, there was an additional set of objections, by Father Bourdin, accompanied by Descartes's irate replies. Descartes was not a man to suffer fools gladly, and he found it easy to believe that his critics were fools. Sometimes he was right.

MEDITATIONS ON FIRST PHILOSOPHY (1641)

The title page of the first edition claims that Descartes was publishing it "with the approval of the learned" and that in his work he would demonstrate both the existence of God and the immortality of the soul. Neither of these claims was true. Though he and Mersenne tried, they were not able to get the approval of the Theology Faculty at the Sorbonne. While Descartes did offer several arguments for the existence of God, he did not even attempt to prove the immortality of the soul. Both these mistakes were corrected on the title page of the second edition, which appeared in the following year. But it is puzzling that they were made in the first place. Some have blamed them on Mersenne, who saw the work through the press. He is supposed to have hastily inferred from the Dedicatory Letter to the Theology Faculty that Descartes intended to prove the immortality of the soul. But in December 1640 Descartes warned Mersenne not to expect a proof of immortality in the *Meditations*. Descartes thought the most he could prove was that the mind is distinct from the body, not subject to die when the body does. Since God is omnipotent, he can always annihilate the mind (Adam and Tannery III, p. 265–266). The title page of the second edition claimed only a proof that mind and body are really distinct, and it dropped any claim to be approved by the learned.

The *Meditations* is a work with multiple agendas. No reasonable interpreter doubts that Descartes wanted to establish the religious conclusions announced on the title page of the second edition. But the First Meditation emphasizes a different aim: establishing something firm and lasting in the sciences. It is that project that has preoccupied most English-language students of Descartes and made the *Meditations* one of the most commonly used texts in modern universities. The project involves more than just validating our reliance on clear and distinct ideas. As Descartes said in a letter to Mersenne (January 28, 1641), he also wanted to accustom people to the foundations of his physics and destroy Aristotelian natural philosophy.

The First Meditation begins by recalling the project of the *Discourse*: ridding ourselves of all past beliefs. Descartes assumed that if a belief survives a thorough attempt to doubt it, and is permissive in what it counts as a valid ground of doubt, it will qualify as indubitable and provide a proper foundation for reconstructing our system of beliefs. If the fact that a belief is indubitable is to make it a proper foundation for a new system of beliefs, that indubitability cannot be a merely psychological mat-

ter. But facts about what we can and cannot believe are relevant to determining what is indubitable. We cannot doubt a belief at will. We must have *some reason* for doubt. That reason need not be probable enough to make the belief improbable. But if, after a thorough search for some reason, we cannot find even a slight reason for doubt, our inability to doubt the proposition is more than just a psychological fact about us.

How are we to proceed? If we had to question each of our beliefs individually, it would be an endless task to doubt them all. Fortunately, many of our beliefs are based on other beliefs. If we shake the foundation, we shake everything that rests on it. Most, if not all, of our beliefs are based on trust in the senses. Descartes actually said, early in the First Meditation, that *all* his past beliefs were. But when Frans Burman questioned him about this, he explained that the "I" who speaks to us in the *Meditations* is a man who is first beginning to philosophize, someone who holds the opinions anyone might hold, if he has not reflected critically on his beliefs. Call this fictional person "the meditator." Descartes does not endorse all the opinions the meditator expresses, any more than the author of a dialogue endorses all the opinions his characters express. Before the First Meditation is over, reflection will lead the meditator to drop this empiricist assumption. but in the beginning, empiricism rules.

The meditator briefly considers common cases of sense deception as a ground of doubt, but dismisses them because they support doubts only about small or distant objects, not a more general doubt about all material objects. More serious, he thinks, are the skeptical implications of dreaming. Each night, when he falls asleep, he dreams. In those dreams he has experiences just as vivid as his most vivid waking experiences. Or at least if there is a difference between his dreams and his most vivid waking experiences, it is not discernible during the dream. Only afterward, when he wakes up, does he realize that he was dreaming. So it is possible, for all he knows, that he is dreaming *now*, no matter how convincing his present experience seems to be. If this doubt can be raised about *any* sense experience, no matter how vivid, then no belief based on sense experience can be certain. And if all justified beliefs are based on sense experience, then no belief is certain.

That seems to be the conclusion the meditator reaches during the first stage of his reflections. But the *Meditations* are a dialogue within the meditator's mind, a dialogue between his skeptical side and his dogmatic side. After reflection it occurs to the meditator that perhaps arithmetic and geometry, those sciences that deal with the simplest and most general objects and care little whether their objects exist in nature, might not be affected by the dream argument. Sense experience is our primary means of knowing what is in nature. But if the mathematical sciences do not require objects actually existing in nature, they may not depend on sense experience. If they do not, they will not be impugned by an argument that shows sense experience to be unreliable. Moreover, it seems impossible that truths so clear should be suspected of falsity.

The meditator then reflects on the implications of a belief he has long held: that there is a God, who can do all things, and who has made him what he is. If there is such a being, it seems he might have created him (the meditator), not only with deceptive perceptions of everything around him (so that he seems to see Earth, a sky, and other extended objects even though there are no such things), but also with mistaken beliefs about even the simplest truths of mathematics—so that it seems evident that two added to three makes five, though this proposition is false. Of course, the meditator also believes that God is supremely good, and that such a being would not want him to be deceived. But the meditator does, after all, make mistakes. Evidently, if the meditator was created by a good God, it is consistent with God's goodness to permit him sometimes to be deceived. Couldn't it be consistent with God's goodness to make him always be deceived? Moreover, dropping the assumption that God created the meditator does not help. The less perfect his cause, the less reason he has to think that his cognitive faculties are not flawed.

The meditator has no answer to these arguments. He concludes that a legitimate doubt can be raised about *all* his former beliefs and that he has powerful (*validas*) and carefully considered (*meditatas*) reasons for these doubts. The reasons are powerful not because of their probability, but because of their scope, because they cast doubt on all kinds of beliefs, sense-based or not. The meditator insists that his former beliefs remain highly probable, more reasonable to believe than to deny. Later he will characterize the doubt based on the possibility of a deceiving God as "slight (*tenuis*) and … metaphysical." Many critics have asked how Descartes *knows* that the premises of his skeptical arguments are true. The answer is that he does not, and *need not*, claim to know that. Since the meditator is seeking absolute certainty, the only epistemic requirement for a legitimate ground of doubt is that the doubt not be one that he has compelling reasons to reject.

In the Second Meditation, having resolved to set aside as false anything that admits even the slightest

doubt, Descartes claims to finds his Archimedean point, a proposition that resists all attempts to doubt it, on which he can build his revised system of beliefs. His initial certainty is the existence of the self. But the argument for the certainty of his existence takes a different form than it had in the *Discourse*. The famous inference—"I think; therefore, I exist"—does not appear. Instead, the *cogito* paragraph concludes with the words "This proposition, *I am, I exist*, is necessarily true as often as I utter it or conceive it in my mind" (Adam and Tannery VII, p. 25).

This formulation, combined with the absence of any explicit inference and some obscure remarks Descartes makes in the second set of replies, has led some readers to think Descartes is claiming intuitive certainty for the proposition "I exist." But we must remember that in the *Discourse* and the *Meditations* Descartes was writing for readers who had not read the *Regulae*. In neither the *Discourse* nor the *Meditations* does he introduce intuition as a central concept in his epistemology. Moreover, like the *Discourse*, but unlike the *Regulae*, the *Meditations* has raised as yet unresolved doubts about those paradigms of intuitive knowledge, the simplest truths of mathematics.

There is an alternative to seeing the existence of the self as something which, if known at all, must be known either by intuition or by inference from intuitions. As the *cogito* paragraph opens, the meditator is reviewing his situation. He has rejected the existence of all bodies, but perhaps there is something incorporeal whose existence he cannot doubt. God, perhaps? But God does not yet qualify as an indubitable being; at this stage the meditator thinks he himself might be the cause of his thoughts about God. What about himself? Is his existence so bound up with the existence of his body that he cannot exist without it? No. *If he has convinced himself of something* (say that there are no bodies), *then he must exist*, whether bodies exist or not. Perhaps a supremely powerful deceiver is deceiving him about everything (including his own existence). But *if the deceiver is deceiving him, then he exists.*

The italics here emphasize two *cogito*-like conditionals that each have an antecedent hypothesizing some thought process that the meditator *may* be involved in (convincing himself, being deceived by the deceiver) and a consequent affirming his existence. The meditator does not commit to either of the antecedents. The point is that whatever skeptical hypothesis he entertains, and whether he is responsible for his beliefs or the deceiver is, it follows from that hypothesis that he exists. Descartes hit on a way to justify accepting something as a first principle without incurring reasonable accusations of dogmatism: if the truth of a proposition follows from any skeptical hypothesis that could validly be invoked to cast doubt on it, then it's permissible to accept that proposition as certain without other argument, specifically, without having to deduce it from some prior certainty and without having to appeal to an infallible faculty of intuition.

Any valid ground of doubt must entail the existence of the doubter. Although valid doubts need to satisfy only a weak epistemic requirement (that we not have compelling reasons to reject them), there is another condition they must also satisfy: They must explain, conjecturally at least, how the person engaged in the search for truth could be mistaken. But if they do that, they must say something of the form "Perhaps, but you could be mistaken because God is deceiving you, or you are dreaming, or you are yourself the source of this thought, etc." The skeptic, if he is rationally, and not dogmatically, skeptical in his attempt to cast doubt on our beliefs, must argue that there is some reason why things seem to us as they do, even though things are not as they seem. As soon as he does that, he concedes that we are thinking, and hence that we exist.

Descartes used the same procedure when he took up the next problem in the Second Meditation: What is this self whose existence the meditator is now certain of? The meditator starts from the beliefs that he assumes a beginner in philosophy would have and asks which of them, if any, can survive radical doubt. The meditator thinks that he is something that has both a body (something with shape and location, occupying space so as to exclude other bodies, perceptible to the senses, and movable by other bodies that come in contact with it) and a soul (a fine substance, like air or fire, infused throughout the body and responsible for nutrition, motion, sensation, and thinking).

Not many of these prereflective beliefs can survive the hypothesis that some supremely powerful malicious being is deceiving him. The meditator has already rejected, until it can be reestablished on firmer ground, the belief that there are bodies. So the self whose existence he is certain of is apparently not something corporeal, nor can it engage in functions requiring the existence of a body. Nutrition and movement must go. At first it seems that sensation too must go, since sensation apparently presupposes the existence of sense organs. Only thought remains. Just as the existence of the self follows from any hypothesis entertained to cast doubt on it, so (trivially) does its thinking. If Descartes's procedure for identifying indubitable first principles is sound, he could have taken "I think" as a first principle and demonstrated "I exist"

from that principle. Perhaps that is why he sometimes gives the appearance of doing that.

To say that the meditator is a thinking being is to attribute a number of different activities to him: that he understands many propositions, affirms some, denies others, and suspends judgment about still others. All this is implicit in the dialogue between his skeptical side and his dogmatic side. And on reflection, even sensation is something whose occurrence he cannot deny. Not having a body, he may not have sense organs, but he cannot deny that it sometimes *seems* to him *as if* he were perceiving something through some organ of the body he thought he had. And such seeming is a purely mental occurrence, immune to skeptical doubt. The skeptic assumes it in his attempt to explain why we had the ill-founded beliefs we had about bodies.

Toward the end of the Second Meditation, Descartes indulges in what looks like a digression. Though the meditator has not yet resolved his doubts about the existence of bodies, he says that he will give in to his natural inclination to believe that he knows bodies (which he can imagine and sense) more distinctly than he knows this mysterious self (which he can neither imagine nor sense). So he decides to examine one particular body, a piece of wax, to see what he knows distinctly in that object. He describes its properties: size, shape, color, hardness, temperature, taste, fragrance, etc. Then he takes the wax near a fire and notes the changes it undergoes in these changed circumstances. All its sensible properties change. What was cold becomes warm; what was hard becomes soft; and so on. But the wax, he says, remains (numerically) the same, in spite of its qualitative changes. No one doubts this. He concludes that the wax is not to be identified with any of its changing sensible properties. What he imagines distinctly in the wax is nothing but an extended something, capable of changing its shape, and capable of change in general.

Descartes draws a number of conclusions from this experiment. First, the wax, and bodies in general, are known, not by the senses or the imagination, but by the mind alone. The wax is capable of changing in many more ways than either the meditator's senses or imagination can encompass. Only the mind can grasp the wax. Second, the mind is better known than the body. Whenever the meditator judges, on the basis of sense evidence, that the wax exists, those sensations do not establish the existence of the wax. But they do establish the existence of the thinking being that judges that the wax exists.

What appears here to be a digression, not necessary to establish the main announced conclusions of the *Meditations*, does serve Descartes's unannounced purpose of insinuating the foundations of his physics. Just as the middle section of the Second Meditation clarified our prereflective concept of the soul, or mind, paring away the inessential to lay bare the essential property of thought, so the concluding section clarifies our prereflective concept of body. After the wax passage we know not to think of the sensible properties of bodies as essential to them. The only first-order property essential to any body is that it is extended. We also know not to think of bodies as inherently perceptible by the senses.

The wax passage serves another nonobvious purpose. It is characteristic of Descartes's method in the *Meditations* that he does not formally define his central concepts, but lets them emerge in informal ways. One of Descartes's central concepts is that of a clear and distinct idea, which he first mentioned prominently at the beginning of the Third Meditation, where he proposed his criterion of truth: Whatever he perceives clearly and distinctly is true. He did not define "clarity" and "distinctness" until he wrote his *Principles of Philosophy* (and even then the definitions are not very helpful). But the wax passage gives us a paradigm of what it is to acquire a clear and distinct idea. When the meditator begins to reflect on the wax, his idea of it is imperfect and confused. After he considers more attentively what the wax consists in and eliminates the inessential, his idea is clear and distinct.

The Third Meditation illustrates another way in which the process of acquiring clear and distinct ideas can work. When the meditator introduced the idea of God in the First Meditation, he explained the content of that idea by enumerating several attributes that he took God to have, among them that he created the meditator, that he can do all things, that he is supremely good, and that he is a source of truth. The problem the meditator faced was that he was not sure that all these attributes are united in one being. Perhaps he was created by an omnipotent being who is not supremely good and, far from being a source of truth, is a deceiver.

The idea of God is central to both arguments for the existence of God in the Third Meditation. At the heart of those arguments is the contention that the only possible explanation for the meditator's possessing an idea of God is that God does exist and has implanted an idea of himself in the meditator, much as a craftsman might stamp his mark on his work. But what exactly is the content of that idea? Descartes offers three answers to that question in the Third Meditation. The first two involve lists of divine attributes: God is supreme, eternal, infinite, omniscient,

omnipotent, and the creator of all things apart from himself (Adam and Tannery VII, p. 40); God is an infinite substance, independent, supremely intelligent, supremely powerful, and has created the meditator and everything else that exists, if there is anything else (Adam and Tannery VII, p. 45). The French translation of the *Meditations*, which appeared in 1647, adds immutability to both lists.

These varying lists have several notable features: All three include the idea that God is the creator and that he is omnipotent. The two lists in the Third Meditation both omit the attributes that gave trouble in the First Meditation, that God is supremely good and the source of truth. And the Third Meditation lists both include infinity, an attribute that will play an important part in the arguments for God's existence. But no two lists are identical. This highlights a problem to which the Third Meditation will suggest a solution. We cannot adequately explain the content of the idea of God by listing his attributes. We may know where to begin: with his being the creator and being omnipotent. But we do not know where to stop. If God is absolutely infinite, not only are his individual attributes infinite in themselves, he must have infinitely many of them. No finite mind will be able to list them all. And as we learned in the First Meditation, there may be disagreement about some candidates. If God created the meditator and is omnipotent but the meditator makes mistakes and is imperfect in other ways, is God supremely good and a source of truth?

The solution that the Third Meditation proposes is that God is best understood as a supremely perfect and infinite being (Adam and Tannery VI, p. 46), where this implies that he must have *all* perfections and *only* perfections. This formula is a generalization from the various lists of attributes, each of which is a perfection. It is a useful way of summing up those lists, since it covers attributes that may have been omitted, either inadvertently or because of the limitations of the mind compiling the list. But most important, it provides a criterion for deciding what should be on the list and what should not. If an attribute is a perfection, it should be; if it is not, it shouldn't be.

Is there such a perfect being? Descartes first addressed this question in the Third Meditation, mounting two arguments, each starting from the assumption that we have an idea of God of the kind described. In the third set of objections Thomas Hobbes challenged the claim that we have an idea of God. But Descartes replied that Hobbes's challenge depends on confusing ideas with images. Since God is an infinite being, we can, of course, have no image of God. But that does not mean that we cannot have an idea of him. "Whenever I express something in words and understand what I am saying, it is certain, from this very fact, that there is in me an idea of the thing the words signify" (Adam and Tannery VII, p. 160). If a theist affirms, and an atheist denies, the existence of God, and if they both understand what they are saying, they both have an idea of God.

But how can the mere fact that we have an idea of God lead to a proof of his existence? In the Third Meditation the arguments are causal. They depend first on the general causal maxim that there must be at least as much reality in the total efficient cause as there is in the effect (Adam and Tannery VII, p. 40). It was an axiom of ancient philosophy, which Descartes endorsed, that something cannot come from nothing. A stone that previously did not exist cannot now begin to exist unless it is produced by something in which there exists, "either formally or eminently," whatever is in the stone. Descartes never really explained what the quoted qualification means. It's clear that he did not think that the cause needs to have the same properties as the effect. If it did, then God, who is incorporeal, would not be able to create extended objects. It is also clear that if the cause does not have the same properties, it must have properties "of at least equally great perfection." There cannot be heat in an object not previously hot except from something that is "of an order at least as perfect as is the heat" (Adam and Tannery VII, p. 41). That language clearly does not mean that the cause needs to have heat in it. But it is unclear what restriction the language does place on possible causes.

From this general causal maxim the meditator infers a causal principle applying specifically to ideas: The cause of an idea must contain at least as much formal reality as the idea contains objective reality. If we understood what formal reality is, we would understand what objective reality is, since objective reality can be defined in terms of formal reality. Objective reality is a property of ideas as representative entities that is correlated with the formal reality of their objects. An idea that represents its object as possessing a very high degree of formal reality will have more objective reality than one that represents its object as possessing a lower degree of formal reality. To say that an idea has objective reality is *not* to say that its object exists. All ideas have some degree of objective reality, even though some ideas have non-existent objects. Similarly, all objects have some degree of formal reality, even though some objects do not exist.

Descartes's point is that all ideas have some content, and their content requires causal explanation. In the first

set of replies, he illustrates this with the example of someone who has the idea of a machine with a highly intricate design. The person might have acquired the idea of that machine by observing a real machine with that design. But perhaps there is no such machine. If not, we must seek some other cause for his conception of that object, perhaps in his extensive knowledge of mechanics. If he derived his idea of the machine neither from having observed such a machine nor from his knowledge of mechanics, he may have derived it from someone who had seen such a machine or had the requisite knowledge of mechanics. But whatever the cause, there must be a cause sufficient to produce that effect. The idea of God, as the idea of an infinite being possessing all the perfections that God is supposed to possess, has more objective reality than the idea of a finite substance does. Indeed, it has as much objective reality as it is possible for an idea to have, since its object has as much perfection as it is possible for an object to have.

Stripped to its essentials, the argument is as follows: 1) Each idea must have a cause possessing at least as much formal reality as the idea represents its object as having. 2) The idea of God represents its object as having the maximum possible formal reality. 3) Therefore, the only possible cause of our idea of God is a being that has the maximum possible formal reality (that is, equals all possible perfections). 4) Therefore, the idea of God must have God as its cause. 5) Therefore, God exists. This argument has generally not been well received by Descartes's readers, partly because of the obscurity of the causal principles involved, and partly because Descartes seems to have precluded himself from ever using such an argument.

The argument appeals to causal principles that Descartes said are known by natural light, a cognitive faculty whose deliverances cannot be doubted in any way. (Adam and Tannery VII, p. 38). As an example of one of the things so known, he gave the proposition: "From the fact that I doubt, it follows that I am." But just before he entered on this argument, he said that until he knew whether God exists and can be a deceiver, he could not be certain of anything (Adam and Tannery VII, p. 36). And he seemed there to regard the possibility of God's deception as a reason for doubting, not only simple truths of mathematics, but also the proposition "If I think that I am something, I am something"—a proposition that would presumably be known by that natural light whose deliverances are beyond doubt. It looks as though, to prove the reliability of the natural light, Descartes needs to construct a proof of the existence of a nondeceiving God. And to construct that proof, he needs to deploy premises known by the natural light, which he cannot be sure of until he is sure of his conclusion. The reasoning looks circular. The difficulty is known, therefore, as the Cartesian circle.

We will not have the materials to respond to this objection until we have considered the Fourth Meditation. But first we must note briefly that Descartes offers a second causal argument in the Third Meditation, beginning at the top of Adam and Tannery VII, p. 48. The focus of this argument is not on explaining the existence of the meditator's idea of God, but rather the meditator's own existence as a being possessing this idea. This argument has not persuaded many readers either, partly because it involves some of the same conceptual difficulties as the first argument. But it does introduce another restriction on causality, which had interesting consequences.

At one point in the argument the meditator considers the possibility that his existence as a being possessing the idea of God that he has, might be explained by saying that he has always existed, as he does now. This might not seem a plausible hypothesis, since few people are likely to think they have always existed. But Descartes's reason for rejecting it is curious. He replied that each person's life can be divided into countless parts, each completely independent of the others. From the fact that the meditator exists at one moment, it does not follow that he will exist at the next moment. Apparently he will not exist then, unless some cause creates him again at that time. The meditator thus requires a cause to sustain him in existence from one moment to the next, much as he requires a cause to bring him into existence, if he has not always existed. And that cause, of course, must be God.

What is interesting about this position is the assumption that for a cause to explain an effect, the existence of the effect must follow logically from the existence of the cause. The will of an omnipotent being can satisfy this requirement on causality. It is part of the notion of omnipotence that if an omnipotent being wills something, what it wills must occur. But no finite being appears able to satisfy the condition. For any supposed finite cause, it will always be possible for that being to exist without having the effect we suppose it to have. This restriction on causality looks like it will lead quickly to the occasionalist doctrine that no finite being is ever truly a cause, that God is the only real cause of anything that happens, apparent finite causes being merely occasions for his willing things to happen as they do. It is unclear whether Descartes saw that his argument might have these consequences.

At the end of the Third Meditation, having devoted most of his longest meditation to elaborating two complex arguments for the existence of God, Descartes makes a quick argument that God cannot be a deceiver. The God whose existence he has proven is a supremely perfect being, possessing all perfections and no defects. It is manifest by the natural light that all deception involves some defect. So God cannot be a deceiver. Of course, there is the awkward fact, noted in the First Meditation, that God's creatures do sometimes make mistakes. Not until the next meditation will Descartes attempt to reconcile his awareness of that with his conviction that a supremely perfect being created him.

The main line of response to this difficulty in the Fourth Meditation is a variation of a standard approach to the problem of evil: Though God created the meditator as he is, God is not responsible for the meditator's errors, because they arise from the meditator's misusing the free will God has given him. Free will is a good great enough to compensate for whatever evil is involved in the meditator's mistakes. If the meditator exercised his free will properly, he would not make mistakes.

In the Third Meditation, Descartes classified his thoughts into three kinds: ideas, which, though *not* images, are *like* images insofar as they represent their objects as possessing certain properties; volitions or emotions, which involve having an idea of an object and also having some affective attitude toward it (wanting it, disliking it, fearing it, etc.); and judgments, which involve having an idea of an object and affirming or denying something about that object. Only judgments can be true or false. The most common mistake the meditator makes is to judge that things outside him are as his ideas represent them. When they are not, as is often the case, he errs. But error, like any judgment, always involves an act of the will, either affirming something or denying it. The meditator makes judgments he does freely. If he makes a mistake, it is his fault, not God's.

The notion of freedom used here requires some examination. Within one sentence, Descartes suggests two very different conceptions of freedom. The sentence reads as follows:

> The will, or freedom of choice … , consists only in this, that we can do something or not do it (that is, affirm or deny, pursue or flee, the same thing), *or rather*, only in this, that when the intellect proposes something to us for affirmation or denial, pursuit or avoidance, we are so inclined that we do not feel we are determined to it by any external force. (Adam and Tannery VII, p. 57; emphasis added)

This puzzling sentence presents difficulties both of translation and of interpretation. But what Descartes seems to mean by it is that the first clause (before "or rather") describes one (indeterminist) way we can be free, and the second clause (after "or rather") describes another way we can be free (without assuming indeterminism).

Descartes's view seems to be this: *Much of the time*, when we affirm something, we could have denied it, and when we deny it, we could have affirmed it (or neither affirmed it nor denied it). This will be true under a variety of circumstances: We might have no evidence one way or the other; we might have evidence each way, but the evidence might not favor one way over the other; or the evidence for the proposition might outweigh the evidence against it, perhaps quite strongly, without being conclusive. In all these cases we will have the power to decide either way, and will be free under the first clause of Descartes's definition. This is often called a *liberty of indifference*, though that term has misleading connotations. It may suggest either that we have no evidence one way or the other, or that our evidence one way is no stronger than our evidence the other way. As Descartes conceived this liberty, that will not always be true. In cases where our evidence for a proposition is strong but inconclusive, as is our sense evidence for the existence of material objects, denial or suspension of judgment will be difficult, but not impossible.

But *sometimes*, Descartes thought, we find that we cannot help judging as we do. In the Second Meditation, when the meditator was examining whether there was anything in the world and noticed that it followed from the fact that he was examining this that he existed, he could not but judge that what he understood so clearly was true. He was not aware of any external force compelling him to judge thus. Rather, a great inclination of his will followed from a great light in his intellect. He seemed to be all the more free the less indifferent he was. This is what is sometimes called a *liberty of spontaneity*, a notion that suggests that the absence of external constraint is sufficient for freedom. It is not necessary for our freedom that we have the power to act differently than we do.

Descartes wanted to allow both a liberty of indifference and a liberty of spontaneity. When we do not have clear and distinct ideas, we possess a liberty of indifference. We can judge either way. When we do have clear and distinct ideas, we cannot judge otherwise, but we are still

judging spontaneously, and not under any kind of coercion. The absence of external coercion does not imply that there is no external causation of our judgments. Descartes explicitly allowed that God might be disposing the meditator's innermost thoughts to judge the way he did. That will not diminish his liberty of spontaneity, though it will mean that he no longer has a liberty of indifference.

Some critics have found Descartes's theory of judgment highly implausible. Benedict de Spinoza argued that Descartes was confusing judgments with utterances when he supposed that we might have a liberty of indifference in some of the cases where he claimed it. It is one thing to *say* that one's experiences of the external world might have no more basis than a dream, and quite another to actually *believe* it. The first is easy; the second may well be impossible. Again, is the liberty of indifference that Descartes requires to relieve God of responsibility for our errors compatible with his doctrine that God is continuously creating us at each moment of time? The doctrine of continuous creation seems to make us completely dependent on God; a liberty of indifference seems to make us at least partially independent of God.

For all the time that Descartes spent arguing that we have a liberty of indifference with respect to some ideas, in the final analysis he seems not to have relied on that liberty to reconcile God's goodness with the occurrence of error. At the end of the Fourth Meditation he conceded that God could easily have brought it about that the meditator would never make a mistake without losing his freedom. All God would have to do is to give the meditator clear and distinct ideas about everything he would ever have to make judgments about, or to implant in him a firm resolution to make judgments only about things he perceived clearly and distinctly. In the closing paragraphs of the Fourth Meditation it looks like Descartes's solution to the problem of error does not depend on free will at all, but on the thought that, although the *meditator* might be better if he never made mistakes, it is possible that the *world as a whole* is better for having in it beings who make mistakes. Variety is the spice of the universe.

Nevertheless, the doctrine of judgment in the Fourth Meditation has considerable systematic importance. The method of doubt requires that we suspend judgment about everything we have the slightest reason to doubt, that we withhold our assent from things we do not perceive clearly and distinctly. Moreover, it is arguable that Descartes's vindication of reason depends on our *in*ability to refrain from assenting to things we *do* perceive clearly and distinctly. As noted above, when Descartes was arguing in the Third Meditation that God exists and is not a deceiver, he frequently justified the assumptions of those arguments as things "manifest by the natural light." And it's not clear how, given the arguments of the First Meditation, he can repose confidence in that, or any other, cognitive faculty until he has first determined whether God exists and is a deceiver.

Since the mid-twentieth century, at least, commentators have been reluctant to accuse Descartes of blatant circularity. But there is no consensus about how he escapes the accusation. Here is one try. It is not controversial that Descartes thought that our clear and distinct ideas compel assent when we are attending to them. We may be able to doubt simple propositions of mathematics when we consider them under some general rubric, like "the things which seem most evident to me." But when we are actually focusing on a particular simple proposition of mathematics, we cannot in fact doubt it. It compels our assent. The same is true, Descartes thought, of some metaphysical propositions, such as "So long as I think I am something, I am something," and "If I exist now, then it will not be true at some later time that I never existed."

The arguments for God's existence and nondeception in the Third Meditation are constructed from two kinds of propositions. One kind reports the contents of the meditator's consciousness, specifically, the fact that he has an idea of God. This is a presupposition of the dialogue with the skeptic and amenable to the defense offered above for the propositions "I exist" and "I think." The other kind are general propositions, such as "A cause must have at least as much reality or perfection as its effect," and "Deception is a defect." If we perceive these things clearly and distinctly, we will not be able to doubt them when we attend to them. Descartes may not have thought that they are self-evident, in the sense that they command assent as soon as we understand the terms. But if they do not command assent, then we have not yet perceived them clearly and distinctly. We are confused in some way, perhaps by badly understood experiences that seem to refute the principles.

Suppose that we are able to construct an argument that God exists and is not a deceiver, relying entirely on propositions about contents of our consciousness that we cannot doubt and on general propositions that we perceive clearly and distinctly, which we also cannot doubt when we attend to them. If we perceive all these premises clearly and distinctly, and see equally clearly their connection with the conclusion, we cannot doubt the conclusion.

A skeptic might now say, "I understand that you *cannot* doubt that God exists and is not a deceiver. But that's just a fact about you. It doesn't mean the proposition is not worthy of doubt. Perhaps your creator is an omnipotent demon and this conviction of yours is just another of his tricks." On the interpretation offered here, Descartes would say that once he has a compelling argument to the conclusion that he has been created by a God who is not a deceiver, it is no longer enough to offer the mere supposition that a demon might be deceiving us when we assent to ideas that we cannot in fact doubt. In the First Meditation the hypothesis that an omnipotent creator might deceive us, even about matters most evident to us, constituted a valid ground of doubt, because we had no compelling argument against it. By the end of the Fourth Meditation we do have a compelling argument against it. So it no longer constitutes a valid ground of doubt. The validity of a ground of doubt is situational. What constitutes a valid ground of doubt at one stage of the argument, when we have no compelling argument against it, will no longer be valid when we do have such an argument. Descartes makes this clear in his reply to the seventh set of objections (Adam and Tannery VII, p. 473–474).

It may help to consider the Pyrrhonian skepticism that we find in Montaigne's *Apology for Raymond Sebond.* The Pyrrhonist advocates what Montaigne calls the principle of equipollence: For every argument in favor of a proposition, there is an equally strong argument against it. Montaigne's criterion for the strength of an argument is psychological persuasiveness. When someone who holds the principle of equipollence is confronted with a compelling argument that we have been created by a non-deceiving god, he can no longer cast doubt on that conclusion simply by hypothesizing the possibility of deception by an omnipotent being. He must produce an equally strong and compelling argument for the opposite conclusion. Absent such an argument, Descartes is entitled to his conclusion.

There is one other respect in which the situation at the end of the Fourth Meditation is different from the situation at the beginning of the Third Meditation. Now we have a clear and distinct idea of God. At the beginning of the Third Meditation we conceived of God simply as an omnipotent creator who was supposed to be supremely good and is the source of all truth. But we didn't see any necessary connection between these attributes, and we worried that we might have been created by a being who possessed some of these attributes, but not all of them. By the end of the Fourth Meditation we understand that what God is, essentially, is a supremely perfect being, who has all the perfections and no defects. Once we have seen this, we see that the hypothesis of an omnipotent deceiver is incoherent. It is not even a hypothesis that we can consider as a possibility.

In the Fifth Meditation, Descartes had two items on his agenda: considering the nature of material things and arguing once more for the existence of God. His most urgent task if he is to recover from his doubts, he said, is to determine whether he can have any certainty about material things. Before he could decide whether such things really exist, he needed to consider what distinct ideas he had of them. He prepared the ground for this consideration in the Second Meditation, where he identified extension as the one first-order property that remains constant in the wax as it changes. There his focus was on a particular body. Here it is on what he calls "continuous quantity … or the extension of this quantity—or rather, of the thing quantified—in length, breadth and depth" (Adam and Tannery VII, p. 63). So we are to think of geometrical space (continuous quantity) as a material substance extended in three dimensions, of which particular bodies are parts, each possessing its own size, shape, and position, and distinguished from the other parts by their varying motions. (Here again, Descartes is insinuating fundamental propositions of his physics.)

When Descartes reflected on his ideas of extended objects, he realized that he had countless ideas of geometrical objects, objects that may not exist anywhere outside his mind but that nevertheless have a definite nature, "a true and immutable nature," independent of his mind. He could demonstrate properties of these shapes, even though he might never have observed any shapes of the kind whose properties he was demonstrating. He may have observed triangles; it's unlikely that he ever observed a chiliagon (a thousand-sided polygon). But he could determine what its properties are, even if there are no chiliagons to observe. Whenever he saw clearly and distinctly that some property belongs to the true and immutable nature of some thing, that property really does belong to that thing. He had a clear and distinct idea of God as a supremely perfect being. He understood that to be supremely perfect, a being must possess all perfections, and that existence is a perfection. He inferred, then, that God must possess the perfection of existence.

This version of the ontological argument depends on a Platonic philosophy of mathematics, which Pierre Gassendi criticized in the fifth set of objections. Gassendi complained first that it seemed to him that it is hard to maintain that there are true and immutable natures apart

from God. He imagined Descartes replying that he was only saying what they say in the schools, that the essences of things are eternal, and that there can be true propositions about them. But Gassendi did not understand how there can be an essence of something—his example is man—if there are no things of that kind. At one point he seemed willing to concede that there is a sense in which "Man is an animal" can be true even if no men exist. But he said that the statement is true only if it is understood conditionally: "If something is a man, it is an animal." And he gave an analysis of that conditional that makes its truth apparently require the existence of some men:

> When man is said to be of such a nature that he cannot exist without being an animal, it is not on that account to be imagined that such a nature is something or is somewhere outside the intellect. The meaning is only that for something to be a man, he must be like the rest of those things to which we give the same name, man, on account of their mutual similarity. (Adam and Tannery VII, p. 320)

Gassendi also questioned whether existence is a perfection: "Existence is not a perfection either in God or in anything else; it is that without which no perfections can be present. … What does not exist has no perfections or imperfections. … If a thing lacks existence, we do not say it is imperfect … but say instead that it is nothing at all" (Adam and Tannery VII, p. 323). Though Gassendi focused on the idea that existence might be a perfection, his reasoning would seem to exclude its being a property of any kind. He treated existence not as something which is predicated of a thing, but as a precondition of any predication.

When Descartes replied, he was puzzled about what category Gassendi wanted to put existence in. Existence seemed to him analogous to omnipotence, something that can be predicated of a thing, and that therefore is a property. But then he rejected Gassendi's conditional analysis of essential predications. Gassendi's example, man, was one of the "universals of the dialecticians," that is, the Scholastic philosophers. Descartes preferred to focus on essences that we understand clearly and distinctly, like those of geometric figures. We cannot understand the latter essences the way the Scholastics and Gassendi did, as based on concepts abstracted from experience of instances of the concept, because there are no instances for us to experience. This is true not only for such unfamiliar figures as chiliagons, but also for such apparently common figures as triangles. The problem is that nothing in our experience strictly satisfies the definition of a triangle, which requires, among other things, that it be composed of straight lines. The lines we experience turn out, when examined closely, not to be perfectly straight. But we can recognize the figures we experience as approximations of the ideal geometric figures because we have ideas of the ideal figures from another source.

Descartes's objections to Gassendi's analysis of essential predication probably go deeper than his opposition to the scholastic theory that our concepts are formed by abstraction from experience. It seems likely that he would reject any conditional analysis of essential predications, even if it was not spelled out in abstractionist terms. Gassendi had complained that the essences Descartes was talking about could not have an immutable and eternal nature apart from God. Descartes replied that he did not claim that the essences of things exist independently of God. He conceived of them rather as depending on the will of God, and as being immutable only because God's will is immutable. Although Descartes did not explicitly invoke his doctrine of the creation of the eternal truths in the body of the *Meditations* itself –(he *may* have presupposed it in the Third Meditation), he did make it explicit in his replies to objections. (It comes up again in the sixth set of replies.) Reflection on the reasons that may have led Descartes to his doctrine of the creation of the eternal truths may also suggest a reason why he would reject Gassendi's conditional analysis of essential predications.

One problem that scholastic philosophers faced when they thought about essential predications was that, according to the orthodox theory of universal propositions, they have existential import. "All men are animals" entails that there are men. But if "All men are animals" is a necessary truth, so are its entailments. However, "There are men" is supposed to be a contingent truth, made true at the creation by God's will. Descartes may have been moved to compare the eternal truths with the laws that a king establishes in his kingdom because a king's laws depend for their validity not on the existence of violators of those laws, but only on the authority of their source. The king's prohibition on dueling does not depend on there being any duelists. Descartes may have felt that a conditional analysis of essential predications avoided one problem only to raise another equally difficult problem. On the hypothesis that there are no men, "If anything is a man, it is an animal" is a conditional whose antecedent is false. If this is a material conditional, it is true in such circumstances, as is the conditional "If anything is a man, it is a plant." If it is a modal conditional, it is unclear what the truth conditions for such conditionals are (if they have any).

Gottfried Wilhelm Leibniz raised an equally serious problem when he argued that Descartes needed to supplement his ontological argument with a proof that the concept of God is consistent. Arguably, that is what Descartes was trying to do, in a limited way, in the Fourth Meditation. That meditation tried to resolve an inconsistency that had threatened his concept of God since the First Meditation: that God was a perfect being who was nonetheless supposed to have created a very imperfect being. But emphasizing human freedom as a solution to that problem, even if it is not Descartes's final solution, only raised the question of whether human freedom is compatible with God's omnipotence, a problem Descartes would address in his *Principles of Philosophy*.

In the Sixth Meditation there are two announced items on the agenda: establishing the existence of bodies and proving that mind and body are distinct. The first step in approaching the latter problem is to recognize (1):

> (1) Whatever I clearly and distinctly understand can be made by God as I understand it.

The thought here seems to be that if I understand something clearly and distinctly, it must be free of contradiction, and that God, being omnipotent, can create anything that does not involve a contradiction. From (1) it follows that (2):

> (2) If I clearly and distinctly conceive myself as a thinking, non-extended thing, then God can create me as a thinking, nonextended thing.

Similarly, (3) also follows from (1):

> (3) If I clearly and distinctly conceive of body as an extended, non-thinking thing, God can create it as an extended, nonthinking thing.

In the Second Meditation Descartes' meditator, in his reasoning, achieved a state in which he satisfied the antecedents of (2) and (3). He had a clear and distinct idea of the wax as an extended thing, to which he did not ascribe any thought, and a clear and distinct idea of himself as a thinking thing, to which he did not ascribe anything corporeal. So he infers (4) and (5):

> (4) God can create me as a thinking thing, apart from my or any other body.
>
> (5) God can create my or any other body as an extended thing, apart from me or any other thinking thing.

To show that two things are really distinct, it does not matter what power is required to create them as separate substances. *According to the definition of a real distinction, two things are really distinct if they are substances and it's possible for each to exist without the other*. So this is sufficient to prove that:

> (6) I and my body are really distinct substances.

It is not obvious what is wrong with this argument, though it certainly has not lacked critics.

In the fourth set of objections Antoine Arnauld proposed the following counterexample. An individual might clearly and distinctly perceive that a triangle inscribed in a semicircle is right-angled, but not be aware of the Pythagorean theorem, according to which the square on the hypotenuse of a right-angled triangle must equal the sum of the squares on the other two sides. So he might doubt or deny that a particular triangle inscribed in a semicircle has the Pythagorean property. From Descartes's first assumption (1), he might infer (2'):

> (2') If I clearly and distinctly conceive the triangle inscribed in a semicircle as right-angled, but doubt or deny that this triangle has the Pythagorean property, then God can create a triangle inscribed in a semicircle that does not have that property.

The antecedent of this conditional might well be true, it seems, but the consequent attributes to God a power he cannot have, even if we accept Descartes's doctrine of the creation of the eternal truths. Even if God *could have* created a different nature for triangles, the immutability of his will entails that he *cannot now* create a triangle with a different nature (Adam and Tannery V, p. 160). Descartes replied at length to Arnauld's objection without ever seeming to meet the point. It is not obvious what he should have said to defend himself.

Though Descartes regarded mind and body as substances capable of existing apart from one another, he was also anxious to insist that he is very closely united to his body, "as it were, intermingled with it" (Adam and Tannery VII, p. 81), so that he composes one thing with it. His bodily sensations taught him this: He feels pain when this body is damaged, hunger when it needs food, thirst when it needs drink. He does not feel these sensations when similar things happen to other bodies. So, he said, nature taught him that he is not merely present to his body as a sailor is present to his ship. He thereby rejected what the medievals regarded as the excessive dualism of Plato. Bodily sensations are nothing but confused modes of thinking arising from the union of mind and body. It is not clear that this doctrine of mind-body union is compatible with the doctrine that mind and body are distinct. This was to become a major topic of debate after the publication of the *Meditations*, as we shall see.

Descartes's attempt to prove the existence of bodies has generated not so much debate as dismissal. God, he said, had given him a great propensity to believe that his sensations are caused by material things and no faculty for recognizing any alternative source for them. So Descartes did not see how God could be cleared of the charge of being a deceiver if his sensations were caused by something incorporeal. There must be corporeal things. They may not have all the properties he grasped by sensation, since there is much in those properties that is obscure and confused. But they must have all the properties he understands clearly and distinctly, that is, all the properties that are the subject of pure mathematics.

It is indeed hard to see how a perfectly good God could permit such a massive delusion. But Descartes here has weakened the conditions under which God can be judged to be a deceiver. In the Fourth Meditation, God would have been a deceiver if we had false beliefs that we could not help having. Now God is a deceiver even if we have false beliefs that we can help having, provided we are strongly inclined to believe them and have no way of telling that they are false. Perhaps the proper conclusion from this argument is that it is improbable that our belief in material objects is false. When we are dealing with beliefs that we *could* help having, we probably need to know something about God's purposes before we can decide whether or not he would be a deceiver if we held false beliefs under those circumstances. Descartes did not think that we can know what God's ends are. But probably the main reason the argument has not found much favor is that it does not seem that we should have to accept this complex theistic argument to see the existence of bodies as certain.

We should note an important negative conclusion that Descartes reaches in the Sixth Meditation: Even if we have been created by a nondeceiving God, we have no justification for believing that the things we perceive by the senses have *all* the properties we are inclined to ascribe to them. The properties of which we have confused and obscure ideas—the heat we attribute to hot bodies, the color we attribute to green bodies, and in general what later philosophers were to call "the secondary qualities of things"—these properties, insofar as we think of them as properties of external objects, need not resemble in any way the ideas we have of them. There must be some differences in the things themselves, between a hot object and a cold one, or between a red object and a green one. But so long as there is a systematic correlation between the differences in external objects and the differences in our sensations, we needn't suppose that there is anything in the objects themselves resembling color or heat. This was one of the fundamental principles of Descartes's physics that he slipped into the *Meditations*.

CONTINUED CONTROVERSY (1641–1644)

Even before the publication of the *Meditations* in August 1641, Descartes had begun work on his next major publication, his *Principles of Philosophy* (1644), which he sometimes referred to in the correspondence as his "*summa* of philosophy" or as his "philosophy" or as his "physics." His aim was to produce "a complete textbook" of his philosophy, combining metaphysics, physics, and biology, in the form of theses, "where, without any excess words, I will just present all my conclusions, with the true premises from which I derive them" (Adam and Tannery III, p. 233). There would be none of the false starts that gave the *Meditations* their dialectical character. When he first began planning this work, he thought of publishing it with a standard textbook of scholastic philosophy on which he would comment. He had selected Eustachius of St. Paul's *Summa philosophiae* for this purpose, but gave up that aspect of the project after Eustachius's death in December 1640. Clearly, he had decided that he could be more open about his anti-Aristotelianism, and could present his cosmology in a way that would escape condemnation.

When the second edition of the *Meditations* appeared in May 1642, it added not only Father Bourdin's objections and Descartes's replies, but also a letter from Descartes to Father Jacques Dinet, a former teacher at La Flèche and now the head of the Jesuit order in France, complaining about his treatment by Bourdin. Descartes had reason to be upset by Bourdin: He was long-winded, sarcastic, and unsympathetic in his interpretation of Descartes's views. Descartes said he showed the acumen of a bricklayer, not a Jesuit priest. And he wrote bad Latin. Though Descartes seems to have had a genuine affection for some members of the Jesuit order and respect for the quality of education the society provided in its schools, he was prone to see conspiracy in its members' actions. He worried that Bourdin's critique was not the opinion of one Jesuit priest, but represented a consensus among the Jesuits. He urged Dinet to read the *Meditations* himself—or if he did not have the time for that, to assign the task to members of the society more competent than Bourdin—and to let him know if they saw problems in his project. Dinet delegated the task to Father Etienne Charlet, formerly the rector at La Flèche and later Dinet's successor as head of the Jesuits in France, who was to

write to Descartes about his works. It appears that Charlet eventually demonstrated his personal good will toward Descartes and his work, but left him uncertain about the attitude of the society as a whole.

Descartes also included in his letter to Dinet an account of a controversy he was embroiled in at the University of Utrecht. In 1641 a follower of his, Henricus Regius, professor of medicine at the University, had engaged in a disputation there in which he presented his version of Cartesian natural philosophy, going further than Descartes judged it wise to go. Regius said that the union of mind and body was an accidental one, rather than substantial, and also denied the existence of substantial forms, those formal aspects of things that in scholastic natural philosophy were supposed to make them the kinds of things they are and explain their characteristic behavior. For these positions Regius came under attack from the rector of the university, Gisbert Voetius, who took the opportunity to hurl a few barbs in Descartes's direction as well.

Regius felt he needed to reply publicly, and Descartes advised him on what to say. Though Descartes thought Regius meant something acceptable when he declared that the union of soul and body was accidental—namely that mind and body are really distinct from one another, each capable of existing without the other—he warned Regius that the Scholastics would interpret this language differently, and that the best thing would be to claim ignorance of scholastic terminology and to say that the disagreement between them was only verbal. Regarding substantial forms, Descartes thought Regius should say that he did not wish to reject them absolutely, and that he meant merely that he had no need to invoke them in his scientific explanations. Saying that fire possesses the form of fire does not help us in any way to understand its ability to burn wood. This was the stance Descartes had taken, leaving it to his readers to draw the conclusion that if substantial forms were explanatorily useless, there was no reason to postulate them.

When Regius published his reply, he only made matters worse. The university condemned the new philosophy and forbade Regius to teach his course on physical problems. "Utrecht University, the first in the world to allow one of its professors to teach Cartesianism, was also the first that forbade its teaching" (Verbeek 1992, p. 19). When Descartes criticized Voetius in his letter to Father Dinet, Voetius responded by arranging for Martin Schoock, a professor at Groningen and a disciple of his, to write a book that accused Descartes, among other things, of atheism and of fathering numerous illegitimate children. (Descartes did, in fact, have one illegitimate child, a daughter whom he was quite fond of but who died in 1640, at the age of five.) The full story of the Utrecht affair—which ultimately involved lawsuits for libel, charges of perjury, and a prohibition on any discussion of Descartes, pro or con—is too complex to tell here (for further details, see Verbeek 1992).

PRINCIPLES OF PHILOSOPHY (1644)

The *Principles* was Descartes's most systematic work and the one his contemporaries went to for a definitive statement of his philosophy. It consists of four parts, the first dealing with metaphysics and epistemology, the second with general principles of physics, the third with celestial phenomena, and the fourth with terrestrial phenomena. Since Descartes himself preferred the exposition of his metaphysics and epistemology in the *Meditations* to the one he gave in Part I of the *Principles*, and since the science that dominates the remainder of the work is primitive by modern standards, most recent students of Descartes have neglected the *Principles*. Here we must limit ourselves to noting only a few of the many things it adds to what we know from our survey of Descartes's other works.

Among the additions is a metaphor that Descartes used in the Preface to the French translation of 1647: "The whole of philosophy is like a tree, whose roots are metaphysics, whose trunk is physics, and whose branches are all the other sciences, which reduce to three principal sciences, medicine, mechanics and morals" (Adam and Tannery IX-B, p. 14). This passage illustrates Descartes's conception of the close connection among disciplines that we now regard as quite separate, his ambition to found the sciences in metaphysics, and his hope that his foundational work would have practical consequences. Particularly intriguing is his ambition to derive a moral philosophy from his metaphysics and physics. We will see what that led to when we come to his last major work, *The Passions of the Soul*.

One delicate issue the *Principles* raises is the question of the extent of the universe. Copernicus had not claimed that the world was infinite, but later Copernicans, such as Giordano Bruno, did. Though we do not know the specific grounds for the Church's execution of Bruno in 1600, it seems likely that this was one of them. Since Descartes identified matter with (Euclidean) space, it might seem that he too would be committed to the infinity of the physical universe. But in the *Principles* (pt. I, secs. 26–27), he said that he was not. He reserved the term "infinite" for God alone and designated things in which

he could discover no limits—such as the extension of the world and the divisibility of matter—indefinite. Nevertheless, later in the *Principles* (pt. II, sec. 21), he passed from denying knowledge that there are any limits to the extent of the world to affirming knowledge that there are no limits to its extent. When he later began to develop his moral philosophy, he listed, as one of the truths most useful to us, the proposition that we must beware of supposing that there are limits to the extent of the world God created (*Principles*, pt. III, sec. 1). Descartes supposed that an appreciation of the vastness of God's creation would aid us in detaching ourselves from the things of this world (Adam and Tannery IV, p. 292).

Another theologically sensitive issue that Descartes dealt with in Part 1 is the problem of reconciling human freedom with God's omnipotence (secs. 37–41). Descartes's conception of freedom here seems to be more single-mindedly indeterminist than it was in the *Meditations*. If we are to deserve praise for our actions, we must be in some special way the author of those actions, and not have been determined to so act by our maker. We must have been able to do otherwise. That we have the power to assent or not to assent in many cases is as evident as any first principle, though this is not innate knowledge, but something we learn from what we experience within ourselves. On the other hand, now that we know God, we see that his power is so immense that it would be impious to think we could ever do something he had not foreordained.

Recent discussion of the problem of reconciling human freedom with God's attributes has tended to focus on showing that human freedom is consistent with divine foreknowledge. Descartes was more worried about showing it to be consistent with God's omnipotence. Perhaps Descartes thought that his identification of God's will with his intellect ruled out the possibility that he might foreknow without foreordaining. In any case, the solution that Descartes proposed is that we should maintain both our freedom and God's foreordination, even though we do not see how they could be compatible. God's power is infinite; our intellects are finite. So we should not expect to understand how they can be compatible, and we cannot give up two such certain truths merely because of a defect in our understanding. Had Descartes continued to allow the liberty of spontaneity that he recognized in the Fourth Meditation, it seems that he would not have needed to take this position.

In part II of the *Principles*, Descartes laid the groundwork for a version of Copernicanism that was supposed to avoid the censured claim that Earth moves. In sections 13 and 24 he gave a relativistic account of what we ordinarily mean when we say that a body moves: It changes its place, which is defined as its position relative to other bodies taken to be at rest. We will get different answers to the question of whether something is moving, depending on which other bodies we take as our frame of reference. Suppose that a man is sitting on the stern of a ship headed down river to the sea. We say that he is at rest if we consider his constant relation to the part of the ship where he is sitting. We say that he is moving if we consider his relation to the shore, since he is continually moving away from some parts and toward others. If we think that Earth is rotating on its axis and moving just as much from west to east as the ship is moving from east to west, we say that he is not moving—our frame of reference now being certain bodies in the heavens that we suppose to be motionless. But if we think that there are no such motionless points anywhere in the universe, we will conclude that nothing has a permanent place, except insofar as it is determined by our thought. In part II, section 13, Descartes foreshadowed an argument that, he said, makes it probable that there are no genuinely fixed points in the universe. We get that argument in part III, section 29, where he contended that if we follow ordinary usage, there is no reason to say it is the stars that are at rest rather than Earth.

Descartes seems to reject ordinary usage. In part 2, section 25, he said that if we want to understand motion "according to the truth of the matter," we ought to define it as "a transfer of one part of matter, or of one body, *from the neighborhood of those bodies immediately touching it*, considered as resting, *to the neighborhood of other bodies*" (emphasis added). So he treats the immediately surrounding bodies as a privileged reference frame. On this definition, Earth, strictly speaking, is at rest, even though there is admittedly a sense in which it is moving round the Sun. In Descartes's cosmology, it is at rest in relation to the fluid matter immediately surrounding it, which carries it round the Sun, just as a ship, neither driven by the wind nor hindered by an anchor, might be at rest in relation to the water around it, though it is imperceptibly carried out to sea by the tide (pt. III, sec. 26–28). Of course, as Descartes noted, the same thing can be said of all the other planets.

Cartesian scholars have often suspected Descartes of adopting this strict definition of motion simply because he could then claim that in his cosmology the Earth did not move, permitting him to adopt a basically Copernican astronomy without suffering the fate of Galileo. Descartes anticipated that his denial that the Earth moves might be

judged to be "merely verbal," intended to avoid censure. But he said that a careful reading of his work should remove that suspicion (Adam and Tannery V, p. 550). In any event, it is arguable that he had serious reasons, internal to his philosophy, for wanting to define motion in a way that would escape the relativism he saw in the common conception of motion. Motion is supposed to do a great deal of work in his mechanistic physics. As he said in the *Principles*, "All the variety in matter, all the diversity of its forms, depends on motion" (pt. II, sec. 23). To make that kind of explanatory use of motion, he needed it to be something that really exists in bodies, not something that is in them or not, depending on how you look at them. But his solution to the problem is highly problematic, and not only because it did not in the end protect him from condemnation by the Church. (For more on this complex issue, see Garber 1992, chap. 6.)

We cannot leave this all-too-brief discussion of the *Principles* without noting that at the end Descartes commented on the epistemological status he took his scientific theories to have. He claimed that they are *at least morally certain*, that is, certain enough that it would be reasonable to act on them (or perhaps unreasonable not to act on them), even if they are not absolutely or metaphysically certain (pt. IV, sec. 205). His principles explain so many phenomena that it hardly seems possible that they could be false. And some of his principles, he thought, are absolutely certain, because they are grounded in his certainty that God is supremely good and is not a deceiver (pt. IV, sec. 206). He mentioned mathematical demonstrations, the existence of material things, and "all evident reasonings about material things." He clearly hoped that his readers would find even more of his conclusions metaphysically certain.

ROYAL ADMIRERS, CONTINUING CONFLICTS (1644–1648)

Descartes dedicated his *Principles of Philosophy* to Princess Elisabeth, the daughter of Frederick V (formerly the Elector Palatine and briefly King of Bohemia) and Elisabeth Stuart (sister of Charles I of England). They had begun to correspond in 1643, after Descartes learned that the princess, who was living in exile in the Hague, had read his *Meditations* with approval. She pressed him with acute questions about the relation between mind and body, eliciting some surprising answers. Later their correspondence turned to questions of ethics and psychology, which prompted Descartes to write his last major work, *The Passions of the Soul* (1649), also dedicated to her. Though the extravagant mutual flattery that pervades their correspondence may be mere courtly etiquette, readers have sometimes wondered if Descartes did not harbor an affection for this sad, lovely, intelligent young woman that might have led to a romance, had not the difference in their ages, social station, and religion made that impossible. In any event, she proved to be a stimulating student.

Elisabeth began their correspondence by raising an issue that was to become central in the subsequent development of Cartesianism: How, in voluntary motion, can the mind, as a nonextended thing, cause its body, an extended thing, to move (Adam and Tannery III, p. 661)? Her paradigm for an intelligible causation of motion—and Descartes's paradigm too, we might have thought—involves the impact of one body on another, with the cause transmitting some of its motion to the body that begins to move. Impact requires contact, which requires extension in both cause and effect. A nonextended thing cannot have an impact on an extended one.

Descartes replied by saying that what explains the mind's power to move the body is its union with the body (Adam and Tannery III, p. 664). The notion of the union of mind and body is a primitive one, like extension and thought, which cannot be explained in terms of anything more fundamental. But Descartes thought that we demonstrate our possession of this notion when we attribute to so-called "real qualities," like weight, a force that moves bodies toward the center of Earth. Although we have no knowledge of weight, except as a force of a sort that has this effect, we find no difficulty in thinking of it as moving a body, even though we do not think that it does so by actually touching one surface against another. We find this easy to conceive because we experience in ourselves a power to move the body, and we infer that bodies possess qualities that have analogous powers. We call these qualities "real," meaning thereby that we conceive of them as being really distinct from the body that has them, and hence as a kind of substance. (In fact, as Descartes explained elsewhere, we think of them as a kind of spiritual substance, since we attribute goal-oriented behavior to them.)

Unsatisfied with this explanation, Elisabeth pointed out that real qualities are a disreputable part of scholastic natural philosophy that Cartesian physics aims to replace (Adam and Tannery III, p. 684). Descartes promised to give a proper mechanical explanation of such phenomena as the fall of heavy bodies to Earth, so that it will not be necessary to explain the mind's power to move the body in terms of occult qualities, powers known only by their effects. Since Elisabeth did not really understand weight,

she could not use its supposed causal powers to help her understand how the soul might act on the body. It would be easier for her, she confessed, to grant extension to the soul than to suppose that an immaterial being has the ability to move and be moved by a material one. In his reply (Adam and Tannery III, p. 694), Descartes gave her permission to do just that: to think of the soul as an extended being! Thinking of the soul as extended is just thinking of it as united to the body. Elisabeth was not satisfied with this reply either, which hardly seems consistent with saying that we have a clear and distinct idea of the mind as a thinking, nonextended substance. But she got no more from Descartes on this subject.

Later their correspondence turned to ethical questions, and Descartes recommended that they discuss Seneca's "De Vita Beata" (On the happy life). Elisabeth's life as a princess in exile was not a happy one. Descartes hoped that reading Seneca would help her overcome her depression. Evidently, he had not read Seneca, or had not read him recently, when he made that suggestion. When he did, he did not find much useful there. But when he made his own recommendations for achieving happiness, they had a distinctly Stoic flavor: We should use our reason to consider without passion the value of all the perfections, both of body and of soul, so that we can always choose the better. We should cultivate a firm and constant resolution to carry out what reason recommends as best without being diverted by our passions. Virtue consists in sticking to this resolution, and virtue, Descartes thought, is the path to contentment. But before long he decided that he needed to examine the passions in more detail, so that he could define them. This led to the first draft of his *Passions of the Soul*, written in the winter of 1645–1646.

While these positive developments were occurring, the controversy with Voetius continued and spread to the University of Leiden, where Jacob Revius, the dean of the Staten college at the University, attacked Descartes, and Adriaan Heereboord, Revius's subdean in the college, defended him. This time the principal issues were not so much Descartes's rejection of key ideas in scholastic philosophy as the positive doctrines of his own philosophy:

- Whether the method of doubt leads to skepticism—a reasonable concern, considering the problems Descartes faced in getting beyond the *cogito*
- Whether Descartes was guilty of blasphemy even to suggest the possibility that God might be a deceiver—not so reasonable, it seems, since Descartes had shown sensitivity to the issue by substituting the demon for God at the end of the First Meditation and had gone on to argue that the hypothesis of a deceiving God involves a contradiction
- Whether Descartes was guilty of atheism in rejecting the Thomistic versions of the cosmological argument for God's existence and replacing them with less satisfactory arguments of his own—a possibly reasonable concern, though the details of the critic's arguments show a poor understanding of Descartes's conception of an idea of God
- Whether Descartes was guilty of Pelagianism for excessively exalting free will

The principal basis for this last accusation was Descartes's claim, in the Fourth Meditation, that he experienced within himself a freedom of choice so great that he could not conceive of the idea of a greater freedom (Adam and Tannery VII, p. 57). It is above all in virtue of his freedom that he understood how he might have been made in the image of God.

The accusation of Pelagianism had come up in the correspondence with Mersenne as early as 1637 (Adam and Tannery I, p. 366). Descartes was always puzzled by it, since he understood the Pelagian heresy to involve the claim that an individual, using only his own natural powers, without a special act of divine grace, can achieve salvation. He knew that he had never made this claim, and he was happy to reject it when the situation required (Adam and Tannery III, p. 544). Nevertheless, when the curators at the University of Leiden forbade any discussion of Descartes's views, pro or con, and Descartes appealed to them, complaining that he must be permitted to defend himself against misrepresentation, the rector of the University, who was well disposed to Descartes, advised him to drop the appeal. The matter might be brought before an ecclesiastical council, where his opponents would surely win, not because of what he had said about freedom of the will, but "because they believe he is a Jesuit in disguise" (Verbeek 1992, p. 47). This was ironic, in view of the trouble Descartes was having with the Jesuits in France, but it was not the last of the ironies arising from Descartes's ambiguous position on free will, as we shall see later.

These were busy years for Descartes. One matter that occupied him was seeing that his principal Latin works were translated into French, so that they could be read by a broader audience. Various friends did the translations: Louis Charles d'Albert, Duke of Luynes, did the *Meditations*; Claude Clerselier, *Objections and Replies*; and the abbot Claude Picot, the *Principles*. The translations were published in 1647. In each case Descartes is supposed to

have reviewed them, presumably correcting anything he found faulty and occasionally adding text to explain his views more clearly. In principle, this means we might prefer the French translations of his works to the Latin originals. But it is not clear how much weight we can put on the French variations. We cannot know how carefully he reviewed the translations. Substantial variations almost certainly come from his hand. Smaller ones are doubtful. Older translations of Descartes into English blended the Latin and French texts. The now standard translation listed in the bibliography (Cottingham, Stoothoff, Murdoch, and Kenny) properly takes the Latin as the primary text, noting variations in the French.

In 1647 or 1648 Descartes initiated a quarrel with his former follower, Henricus Regius, who had developed positions at odds with Cartesian philosophy. Descartes first criticized Regius in the preface he wrote for the French translation of the *Principles* and later in the short work *Notae in programma quoddam* (Notes on a Program; also known as Comments on a Certain Broadsheet), published in 1648 and notable for its clarification of Descartes's views on innate ideas. Also in 1648, Descartes sat down for a long interview with a young Dutch theology student named Frans Burman. Burman prepared well for the interview, carefully reading Descartes's published works and asking probing questions about them. His record of Descartes's answers is a valuable source of information about Descartes's views, though sometimes it is not clear that Burman accurately transcribed what Descartes said.

Toward the end of this period, Descartes entered into a correspondence with Queen Christina of Sweden, who was making her court in Stockholm a center for learning. Most of their correspondence was conducted through Pierre-Hector Chanut, the French ambassador, and there is none of the give and take that makes his correspondence with Elisabeth so interesting. But Descartes's relationship with Christina was momentous in other ways, as we shall see.

THE PASSIONS OF THE SOUL (1649)

The Passions of the Soul is Descartes's most serious attempt to provide the moral philosophy promised in the preface to the French edition of the *Principles*. In a prefatory letter, Descartes said that he will treat the passions "only as a natural philosopher," not "as a rhetorician, or even as a moral philosopher." But this is somewhat misleading. Although the work begins with a quick course in Cartesian physiology (secs. 1–16), and broader and narrower definitions of the passions that emphasize their close connection with the body (secs. 17, 25, 27–29), it ends by making a moral evaluation of the passions that smacks more of Aristotelian moderation than Stoic rigor: The passions are all good in their nature; all we need do is to avoid their excess and misuse (sec. 211). Indeed, all the good and evil of this life depend only on the passions (sec. 212).

In the broad sense, the passions of the soul are perceptions the soul receives from the things they represent (sec. 17). Sometimes the things these perceptions represent are in the soul itself, as when we perceive our volitions, imaginations, etc. (sec. 19). Sometimes the things they represent are either in our body or in some external object that acts on our body. This category includes bodily sensations, sensations of external objects, and passions in the narrow sense. These last are defined as excitations of the soul that, though in fact proximately caused by some movement of the animal spirits, are not perceived as having that proximate cause, but are referred to the soul itself (sec. 27).

Descartes maintained that there are six "simple and primitive" passions: wonder, love, hatred, desire, joy, and sadness (sec. 69). All other passions are either combinations of the primitive passions or particular species of them. Like sensations, the passions help to preserve the mind-body union: Their use "consists in this alone, they dispose the soul to will the things nature tells us are useful and to persist in this volition" (sec. 52). They are nature's way of telling us what is useful to us and what is harmful, motivating us to pursue what is useful and avoid what is harmful. The sensation of fear incites the will to flight; the sensation of boldness incites the will to fight.

The connection between the movements of the animal spirits and the excitations of the soul they cause and sustain is no more perspicuous here than it was in the correspondence with Elisabeth. Descartes identified the locus of interaction as the pineal gland, selected for this role, it seems, because it is the only part of the brain that is not double, and because a slight movement of this gland can greatly alter the movements of the animal spirits and, conversely, a slight movement of the animal spirits can greatly alter the movement of this gland (secs. 31–32). But how a particular movement of the pineal gland can affect the soul and how an action of the soul can move the pineal gland are mysteries shrouded in silence. The connections, apparently, are established "by nature" (secs. 44, 50), that is, we assume, the will of God.

Descartes, it seems, thought that for the most part the regularities God has put in place work well for us. But just as in the Sixth Meditations, our bodily sensations can

sometimes lead us astray, causing us to want drink, say, when drink would be harmful to us, so can our passions sometimes lead us astray. "When we feel the blood stirred up, we should be warned and remember that everything which presents itself to the imagination tends to deceive the soul, and to make the reasons favoring the object of its passion appear much stronger than they are, and the opposing reasons much weaker" (sec. 211). If the passion favors some object that does not require immediate action, we should refrain from making any immediate judgment and distract ourselves with other thoughts, until our blood has cooled. If it incites us to an action requiring immediate action, we should reflect on the reasons that oppose that action, and follow them even if they seem weaker. This, Descartes said, is "the general remedy for all the excesses of the passions, and the easiest to put into practice" (sec. 211). Descartes is not at his best when he is doing moral philosophy.

DEATH AND CONDEMNATION (1649–1663)

In July 1649, in response to an invitation from Queen Christina, Descartes embarked for Stockholm, where he was to enhance the reputation of her court as an intellectual center and provide the queen with lessons in philosophy. This Swedish adventure did not end happily. When Descartes first arrived in October, his duties were minimal. But by mid-January he was required to give Christina five-hour lessons in philosophy, three mornings a week, beginning at five in the morning. Within two weeks he came down with pneumonia. By February 11, 1650, he had died.

Thirteen years later Descartes's works were placed on the Catholic Church's Index of Prohibited Books. For a long time it was unclear what the grounds for this condemnation were, but recently the Archives of the Congregation for the Doctrine of the Faith have been opened, permitting a clearer view of the Church's reasons and procedures. The Holy Office assigned two outside consultants to read Descartes's works and report on them: Joannes Tartaglia, to read the *Discourse on Method* (and its essays) and the *Meditations* (with the *Objections and Replies*); and Stephanus Spinula, to read *The Principles of Philosophy* and *The Passions of the Soul*. On the whole, the censors (especially Tartaglia) seem to have done their work carefully, attributing to Descartes only doctrines he actually held, or at least doctrines that might reasonably be inferred from what he wrote.

The censors found much to object to. Some were propositions in Cartesian physics where Descartes knew he was pushing the bounds of orthodoxy: the denial of substantial forms and real qualities; the doctrine that Earth moves, while the Sun is immobile; the doctrine that the physical universe has no limits. Others were fundamental doctrines of Cartesian epistemology: that the existence of the self as a thinking thing is the first evident truth, from which all other evident truths derive; that we cannot clearly understand what is true unless we know clearly that God exists and cannot deceive us; and that the standard Thomistic versions of the cosmological argument are unsatisfactory ways of proving God's existence.

Particularly interesting are the objections to two doctrines relating to human freedom: that the soul can easily acquire an absolute power over all its passions; and that freedom of the will does not require freedom from necessity, but only freedom from constraint. The first of these was a proposition that Spinoza also sharply criticized, in the Preface to part V of his *Ethics*. The second was one of five Jansenist propositions that the Church had censured in 1653. So while the Dutch Protestants accused Descartes of Pelagianism, the Catholic Church condemned him for Jansenism, that is, for siding with those within the Church who thought that in their reaction against Lutheran/Calvinist denial of free will the Jesuit theologians had succumbed to Pelagianism. The gate to doctrinal orthodoxy is narrow indeed.

See also Anselm, St.; Aristotelianism; Aristotle; Arnauld, Antoine; Augustine, St.; Berkeley, George; Cartesianism; Cosmological Argument for the Existence of God; Degrees of Perfection, Argument for the Existence of God; Galileo Galilei; Gassendi, Pierre; Hintikka, Jaakko; Hobbes, Thomas; Husserl, Edmund; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Matter; Mind-Body Problem; More, Henry; Newton, Isaac; Nicholas of Cusa; Ontological Argument for the Existence of God; Pascal, Blaise; Plato; Reid, Thomas; Ryle, Gilbert; Skepticism; Spinoza, Benedict (Baruch) de; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY DESCARTES

Discourse on Method, Optics, Geometry, and Meteorology. Translated by Paul Olscamp. Indianapolis, IN: Library of Liberal Arts, 1965. The only complete English translation of the scientific essays published with the *Discourse*.

Oeuvres. 11 vols., edited by Charles Adam and Paul Tannery. Paris: J. Vrin, 1974–1986. This is the standard edition of the texts in their original languages. All references to Descartes's

works in this entry are to volume and page of this edition, cited as "Adam and Tannery."

Treatise of Man. Cambridge, MA: Harvard University Press, 1972. The French text with translation and commentary by Thomas Steele Hall.

Descartes' Conversation with Burman. Translated by John Cottingham. Oxford, U.K.: Clarendon Press, 1976.

Descartes: His Moral Philosophy and Psychology. Translated by John Blom. New York: New York University Press, 1978. The only English language edition of Descartes's correspondence with Elisabeth that gives her letters to Descartes as well as his replies to her.

Le monde, ou Traité de la lumière (*The World*). French text with English translation by Michael Mahoney. New York: Abaris, 1979.

Principles of Philosophy. Translated by Valentine Rodger Miller and Reese P. Miller. Dordrecht, Netherlands: D. Reidel, 1983. The only complete English translation of this work.

Philosophical Writings. 3 vols. Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch, and (in vol. 3), Anthony Kenny. Cambridge, U.K.: Cambridge University Press, 1985–1991. The best comprehensive English language edition. It gives the volume and page numbers of the Adam and Tannery edition in the margins.

The Passions of the Soul. Translated by Stephen Voss. Indianapolis, IN: Hackett, 1989.

BIOGRAPHY

Baillet, Adrien. *The Life of Monsieur Descartes*. Translated by S. R. London: 1653. Available through Early English Books Online (http://eebo.chadwyck.com/home).

Gaukroger, Stephen. *Descartes: An Intellectual Biography*, Oxford, U.K.: Clarendon Press, 1995.

WORKS ON DESCARTES

Alanen, Lilli. *Descartes' Concept of Mind*. Cambridge, U.K.: Harvard University Press, 2003.

Ariew, Roger. *Descartes and the Last Scholastics*. Ithaca, NY: Cornell University Press, 1999.

Ariew, Roger, John Cottingham, and Tom Sorell. *Descartes' "Meditations": Background Source Materials*. Cambridge, U.K.: Cambridge University Press, 1998.

Ariew, Roger, Dennis Des Chene, Douglas Jesseph, Tad Schmaltz, and Theo Verbeek. *Historical Dictionary of Descartes and Cartesian Philosophy*. Lanham, MD: Scarecrow Press, 2003.

Ariew, Roger, and Marjorie Grene. *Descartes and His Contemporaries, "Meditations," Objections and Replies*. Chicago: University of Chicago Press, 1995.

Armogathe, Jean-Robert, and Vincent Carraud. "The First Condemnation of Descartes' *Oeuvres*: Some Unpublished Documents from the Vatican Archives." *Oxford Studies in Early Modern Philosophy* 1 (2003): 67–109.

Baker, Gordon, and Katharine Morris. *Descartes' Dualism*. London: Routledge, 1996.

Broughton, Janet. *Descartes' Method of Doubt*. Princeton, NJ: Princeton University Press, 2002.

Clarke, Desmond. *Descartes' Philosophy of Science*. Manchester, U.K.: Manchester University Press, 1982.

Cottingham, John, ed. *The Cambridge Companion to Descartes*. Cambridge, U.K.: Cambridge University Press, 1992.

Cottingham, John, ed. *Reason, Will, and Sensation*. Oxford, U.K.: Clarendon Press, 1994.

Curley, Edwin. *Descartes against the Skeptics*. Cambridge, MA: Harvard University Press, 1978.

Frankfurt, Harry. *Demons, Dreamers, and Madmen*. Indianapolis, IN: Bobbs-Merrill, 1970.

Garber, Daniel. *Descartes' Metaphysical Physics*. Chicago: University of Chicago Press, 1992.

Garber, Daniel. *Descartes Embodied: Reading Cartesian Philosophy through Cartesian Science*. Cambridge, U.K.: Cambridge University Press, 2001.

Gaukroger, Stephen, ed. *Descartes: Philosophy, Mathematics, and Physics*. Sussex, U.K.: Harvester Press, 1980.

Hatfield, Gary. *Routledge Philosophy Guidebook to Descartes and the "Meditations."* London: Routledge, 2003.

Kenny, Anthony. *The Anatomy of the Soul: Historical Essays in the Philosophy of Mind*. Oxford, U.K.: Basil Blackwell, 1973.

Kenny, Anthony. *Descartes: A Study of His Philosophy*. New York: Random House, 1968.

Koyré, Alexandre. *From the Closed World to the Infinite Universe*. Baltimore: Johns Hopkins University Press, 1957.

Matthews, Gareth. *Thought's Ego in Augustine and Descartes*. Ithaca, NY: Cornell University Press, 1992.

Menn, Stephen. *Descartes and Augustine*. Cambridge, U.K.: Cambridge University Press, 1998.

Rorty, Amélie Oksenberg, ed. *Essays on Descartes' "Meditations."* Berkeley: University of California Press, 1986.

Schuster, John. "Descartes' *mathesis universalis*, 1619–1628." In *Descartes: Philosophy, Mathematics, and Physics*, edited by Stephen Gaukroger. Sussex, U.K.: Harvester Press, 1980.

Verbeek, Theo. *Descartes and the Dutch: Early Reactions to Cartesian Philosophy, 1637–1650*. Carbondale: Southern Illinois University Press, 1992.

Voss, Stephen, ed. *Essays on the Philosophy and Science of Descartes*. Oxford, U.K.: Oxford University Press, 1993.

Weber, Jean-Paul. *La constitution du texte des "Regulae."* Paris: Société d'édition d'enseignement supérieur, 1964.

Williams, Bernard. *Descartes: The Project of Pure Enquiry*. Harmondsworth, U.K.: Penguin, 1978.

Williston, Byron, and André Gombay, eds. *Passion and Virtue in Descartes*. Amherst, NY: Humanity Books, 2003.

Wilson, Margaret Dauler. *Descartes*. London: Routledge and Kegan Paul, 1978.

Edwin Curley (2005)

DESCRIPTIONS

See *Proper Names and Descriptions*

DESCRIPTIONS, THEORY OF

See *Analysis, Philosophical; Existence*

DESGABETS, ROBERT

(1610–1678)

Robert Desgabets was a French Benedictine who offered a form of Cartesianism that departs from René Descartes's own account of the nature of substance and of one's knowledge of the self and of the external world. These departures are indicated in the two book-length texts from Desgabets published during his lifetime, but they are explicated most fully in manuscripts published only during the mid-1980s, in a definitive edition of his philosophical writings sponsored by *Studia Cartesiana*.

Desgabets was born in Ancemont in Verdun, a region annexed by France in 1552, to Jean des Gabets and Barbe Richard. He entered the Benedictine order in 1636 and taught philosophy and theology for over a decade at Saint-Evre in Toul. In 1648 he was named the Benedictine procurer general in Paris, and the following year he took up the position of professor of philosophy at Saint-Arnold in Metz. From 1653 to 1657 he served in administrative posts in various Lorraine abbeys. It was during this time that Claude Clerselier attempted to draw him into a defense of Descartes by sending him copies of Descartes's discussion in unpublished correspondence of the Catholic doctrine that the Eucharist involves the "transubstantiation" of bread and wine into Christ's body and blood. Desgabets endorsed Descartes's proposal that such transubstantiation occurs by means of the union of Christ's soul with the matter of the Eucharistic elements. What he added to this proposal was an argument against the possibility of the annihilation of this matter that appeals to the result in his 1654 manuscript "Traité de l'indéfectibilité des creatures," that material substance has an existence that is "indefectible," that is, indestructible and immutable.

In 1658 Desgabets spent a brief time in Paris on official business, and while there he participated in public discussions of Cartesian natural philosophy. He also offered for consideration a brief *Discours* on the transfusion of blood, which the French physician Jean-Baptiste Denis included in his 1668 *Lettre à Sorbière* in part to draw the attention of the English Royal Society to French research in this area.

Following his return to the Lorraine provinces in 1659, Desgabets worked to spread the teaching of Cartesianism in local Benedictine abbeys. In the mid-1660s he also became involved in the controversies in France over Jansenist theology associated most prominently with the convent of Port-Royal. Desgabets took the risky step of siding with the Jansenists and Port-Royal against the religious and political establishment.

Even so, Desgabets later split with the Port-Royalists on the issue of the Eucharist. One occasion for the rupture was the publication in 1671 of his *Considérations sur l'état présent de la controverse*. Jean Ferrier, the royal confessor, promptly condemned the work to Louis XIV as heretical, and Louis ordered François de Harlay de Champvallon, the archbishop of Paris, to censure it. When Harlay questioned the Port-Royalist solitaires Antoine Arnauld and Pierre Nicole about this text, they denounced it in no uncertain terms. In an audience with Clerselier Harlay also insinuated that Desgabets's tract was responsible for a 1671 decree from Louis to the University of Paris that marked the start of the official campaign against the teaching of Cartesianism in France. Pressure from above led Desgabets's Benedictine superiors to interrogate him the following year and to prohibit him from speaking on the issue of the Eucharist. The effects of the censure were felt even into the mid-eighteenth century, when the Benedictine authorities refused the request of some admirers of Desgabets to publish an edition of his writings.

Despite the 1672 censure, Desgabets subsequently became underprior and then prior of the provincial abbey of Breuil. Moreover, the censure did not bring about the end of his philosophical activity. In 1674 he engaged in correspondence with Nicolas Malebranche after the latter sent him a copy of the first volume of his *Recherche de la verité*. When Simon Foucher wrote *Critique de la recherche de la verité* that cast doubt on claims in Malebranche's text that mind and body are distinct substances and that ideas represent external objects, Desgabets composed a Cartesian refutation of Foucher's skeptical position. Desgabets's *Critique de la critique de la recherche de la verité* appeared in 1675, and like his 1671 *Considérations*, it was published anonymously. Desgabets further defended the fundamental tenets of his *Critique* in a manuscript commentary on the *Meditations*, the 1675 "Supplément à la philosophie de Monsieur Descartes."

In 1677 there was a series of conferences concerning Desgabets's distinctive version of Cartesianism that took place at the chateau of the Cardinal de Retz (Jean-François-Paul de Gondi) in Commercy. Around this same time Retz's secretary, Jean Corbinelli, led a discussion of the results of the conferences at a special meeting of Cartesians in Paris that included Malebranche. Shortly after these discussions, in March 1678, Desgabets died at his home abbey of Breuil, near Commercy.

MATTER, SUBSTANCE, AND THE COGITO

In commentary published with his 1840 edition of the Commercy conferences, Victor Cousin noted that "if dom Robert, in metaphysics, is a disciple of Descartes revolting against all the principles of his master, he is not so in physics. There he is a faithful Cartesian" (cited in Retz 1887, p. 345). The fidelity to Descartes in physics is indicated in a 1666 letter to Clerselier, in which Desgabets criticized as schismatic the attempt of the French Cartesian Géraud de Cordemoy to introduce a version of Cartesian physics that posits indivisible atoms and the void. Desgabets also argued against the atomist admission of vacua by appealing to Descartes's claim in the *Principles* that matter by its nature occupies all imaginable space.

However, Desgabets went further than Descartes in connecting the claim that matter fills all space to the conclusion that this matter is "indefectible" since not even God can annihilate any part of it. There may seem to be a similarity here to Descartes's view in the Synopsis to the *Meditations* that "body taken in general" is incorruptible since it cannot be destroyed by natural means. Still, Descartes argued in the *Meditations* that since creatures have a duration divisible into distinct parts, God can reduce them to nothing at any moment by refraining from conserving them. Desgabets explicitly rejected this line of argument when he charged Descartes with confusing the modes of a substance with the substance itself. In the case of the material world Desgabets allowed that particular bodies can and do go out of existence. However, he claimed that these bodies are merely modes of extended substance that exist only "secundum quid" as particular temporal determinations of that substance. Desgabets insisted that substance itself exists "simpliciter" in a manner that is wholly indivisible, and so not subject to temporal change (1675, p. 77f).

Desgabets's opposition to Descartes's view that substance has a divisible temporal duration is a clear case of his revolt "against the principles of the master." Another such case is provided by his charge that it is a "principal fault" in Descartes that he took the certainty of the cogito argument to show that the existence of the self is better known than and independent of the existence of body. In Descartes this conclusion is supported by the possibility of a hyperbolic doubt of the existence of the material world. In the "Supplément," however, Desgabets objected to the possibility of such doubt. In the first part of this text he urged that the cogito itself undermines this sort of doubt since it reveals that one's thoughts bear an essential connection to bodily motion (1983–1985, p. 5:183f). His argument stresses that reflection on the cogito occurs in a continuous time that is not intrinsic to thought as such but derives from the union of one's thought with motion. Desgabets relied explicitly here on the traditional Aristotelian definition of time as "the measure of motion." He also held, with other Cartesians, that the only motion is local motion, and further claimed, in orthodox Cartesian fashion, that local motion itself presupposes the existence of the particular bodies that are in motion.

These various premises help to explain his conclusion that the temporality revealed by reflection on the cogito could not exist if there were no bodies external to mind. This argument is somewhat reminiscent of the later appeal in Immanuel Kant to the temporality of consciousness in his "refutation" of a "problematic idealism" in Descartes that takes consciousness to reveal with certainty only the existence of the self. Whereas Kant emphasized that the existence of "outer things" is required for the determination of the temporal succession of inner experience, Desgabets held that the existence of bodies in motion is required for the presence of the temporal duration of one's thoughts.

Desgabets's "Cartesian refutation of idealism," as one might call it, is connected to his endorsement in his *Critique* of Foucher's rejection of Malebranche's orthodox Cartesian claim that one has a "pure intellect" that operates independently of the body. For Desgabets, that all one's thoughts are temporal reveals that they all involve a union with motion. Since he adopted the traditional view that the soul is united to the body through the senses, he accepted the scholastic maxim *Nihil est in intellectu quod prius non fuerit in sensu* (Nothing is in the intellect that was not first in the senses). Pierre Gassendi also had appealed to this maxim in response to Descartes's claim that one has a pure intellect, and this resemblance has led some commentators to label Desgabets as a Gassendist. Unlike Gassendi, however, Desgabets was firmly committed to a Cartesian dualism that distinguishes mind as thinking substance from body as extended substance.

IDEAS, EXTERNAL OBJECTS, AND ETERNAL TRUTHS

In the second part of the "Supplément" Desgabets argued that skepticism concerning the existence of extended substance is overturned by "the most simple, the best known, and the most evident of all principles," namely, that all simple ideas or conceptions correspond to real objects (1983–1985, p. 6:223). Desgabets took this principle to be linked to the claim that to perceive nothing is not to per-

ceive. He admitted that one can make false judgments about what one perceives. Indeed, he pointed to the scholastic claim that sensible qualities exist in bodies as a paradigmatic example of such a judgment, one that is to be corrected by "the great discovery of M. Descartes" that these qualities exist only in us (1983–1985, p. 5:164f). However, Desgabets's "intentionality principle," as commentators have called it, requires that ideas that succeed in representing extramental objects, such as the idea of body, presuppose that their objects actually exist in some sense. The qualification is required by Desgabets's distinction, which informs his discussion of the indefectibility of matter, between modes and the substances they modify. Desgabets allowed that one can conceive of modes that do not actually exist insofar as one can conceive of them as only possibly modifying an existing substance. In this way the nonexistent modes have a "true possibility" conferred on them by substance. Since substance cannot be conceived through any other feature of created reality, however, the possibility of its existence also cannot be conceived through anything else. Desgabets concluded that one cannot even conceive of a substance that is "purely possible" and does not actually exist. For him, then, the mere fact that one has an idea of extended substance, and so can conceive of it, suffices to show that this substance exists external to mind.

Desgabets admitted "an extreme difference between the thoughts of M. Descartes and mine" concerning the issue of the existence of the external material world, since Descartes allowed for the possibility that extended substance exist not in extramental reality but only "objectively" in one's mind (1983–1985, 6:223). However, one reason for Desgabets's extreme opposition derives from his development of Descartes's doctrine of the creation of the eternal truths. Descartes had introduced this doctrine in 1630 in correspondence with Marin Mersenne, in which he insisted that God's free and indifferent will is the efficient cause of the eternal truths. Desgabets took this position to indicate that there are no preexisting truths concerning creatures that constrain divine creation. But he also insisted that if eternal truths concerning bodies were grounded in a mental objective reality, then those truths would seem to be as contingent and mutable as one's mind. In Desgabets's version of the Cartesian doctrine the truths are grounded rather in an extended substance with an atemporal existence that is completely indefectible. Thus, the necessity and immutability of the relevant truths are assured, even given that God has freely created the indefectible substance that provides the foundation for these truths.

The juxtaposition in Desgabets of a strong voluntarism and a firm commitment to substantial indefectibility is found also in the work of the French Cartesian Pierre-Sylvain Regis, who called Desgabets "one of the greatest metaphysicians of our age." Regis endorsed Desgabets's arguments both for the claim that one's idea of extended substance reveals immediately the extramental existence of that substance and for the conclusion that temporal human thought requires a union with and thus presupposes the existence of bodily motion. In both Desgabets and Regis, then, radical doctrines concerning the indefectibility of substance, the intentionality of ideas, and the union of all human thought with motion constitute an unusual but philosophically sophisticated version of Cartesianism.

See also Cartesianism.

Bibliography

WORKS BY DESGABETS

Considérations sur l'état présent de la controverse touchant le Très Saint-Sacrement de l'autel, où il est traité en peu mots de l'opinion qui enseigne que la matière du pain est changé en celle du corps de Jésus-Christ par son union substantielle à son âme et à sa personne divine. Hollande: Sphère, 1671.

Critique de la critique de la recherche de la verité, ou l'on découvre le chemin qui conduit aux connoissances solides. Pour servir de reponse à la lettre d'un academicien. Paris: du Puis, 1675.

Dom Robert Desgabets: Oeuvres philosophiques inédites. 7 vols, edited by Joseph Beaude. Amsterdam: Quadratures, 1983–1985.

WORKS ABOUT DESGABETS

Armogathe, Jean-Robert. *Theologia cartesiana: l'explication physique de l'Eucharistie chez Descartes et dom Desgabets*. The Hague: Nijhoff, 1977.

Beaude, Joseph. "Cartésianisme et anticartésianisme de Desgabets." *Studia cartesiana* 1 (1979): 1–24.

Cook, Monte. "Desgabets's Representation Principle." *Journal of the History of Philosophy* 40 (2) (2002): 189–200.

Denis, Jean-Baptiste. *Lettre escrite à M. Sobière, with an "Extrait d'une lettre de Dom Robert des Gabets et Discours de la communication ou transfusion du sang pronouncé à Paris chez Montmor … en juillet 1658."* Paris: Cusson, 1668.

Easton, Patricia. "Desgabets, dom Robert." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 1995–. (http://plato.stanford.edu/).

"Journée D. Robert Desgabets du CNRS." *Revue de synthèse* 74 (1974).

Retz, Jean-François-Paul de Gondi, cardinal de. "Dissertations sur le cartésianisme, par le cardinal de Retz et le bénédictin dom Robert Des Gabets." In *Oeuvres de cardinal de Retz*, vol. 9. Edited by R. de Chantelauze. Paris: Hachette, 1887.

Rodis-Lewis, Geneviève. *L'anthropologie cartésienne*. Paris: Presses Universitaires de France, 1990.

Rodis-Lewis, Geneviève. "Robert Desgabets." In *Grundriss der Geschichte der Philosophie, begründet von Friedrich Ueberweg, völlig neubearbeitete Ausgabe. Die Philosophie des 17. Jahrhunderts. Vol. 2, Frankreich, und Niederlande, vol. 1*, edited by Jean-Pierre Schobinger. Basil: Schwabe, 1993.

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Tad M. Schmaltz (2005)

DESIGN

See *Philosophy of Technology*

DESTUTT DE TRACY, ANTOINE LOUIS CLAUDE, COMTE

(1754–1836)

Comte Antoine Louis Claude Destutt de Tracy, the French philosopher and propounder of the doctrine of Ideology, was born in Paris. Educated at the University of Strasbourg, he entered the army and served later as deputy of the Bourbonnais nobility to the States-General. Despite his noble rank he was a fervent partisan of reform in monarchical government, but by 1792 he had become disgusted with the extremists among the revolutionaries and retired from politics to Auteuil, where he joined the celebrated group of philosopher-scientists that found its center at the home of Madame Helvétius. Among his intimates were Pierre-Jean-Georges Cabanis and Marquis de Condorcet, Comte de Volney and Dominique Joseph Garat. Imprisoned for a year under the Terror, he began to study the works of Étienne Bonnot de Condillac and John Locke, the result of which was his elaboration of the discipline he called Ideology. The group associated with Destutt de Tracy took the name *Idéologues* from his doctrine. They became influential in 1795 in two new institutions, the École Normale and the Institut National, especially in the Second Class of the Institut National.

Ideology, according to Destutt de Tracy, is the analysis of ideas into the sensory elements of which he believed them to be composed. Training in this new science would replace classical logic, and, he maintained, if a man learned how to analyze his ideas, he would then discover which of them were founded in experience and which were groundless. Destutt de Tracy held that Ideology was a branch of zoology; all ideas had a physiological determinant. The child, with its weak sense organs, has nothing but sensation and memory; the adult, whose sense organs have become strengthened through use, has the powers of judgment and intelligence. It was therefore to be asked what the effect of habit would be on judgment. This question was put to the Second Class of the Institut National on 15 Vendémiaire, An VIII (October 6, 1799). The winning *mémoire* was that of Maine de Biran, at that time a young disciple of the *Idéologues,* and his *Mémoire sur l'habitude* (1802) formed the link between the French epistemological tradition of the eighteenth century and that of the nineteenth-century "spiritualists."

The word *thinking* in the works of Destutt de Tracy means, as it did for René Descartes, all conscious processes. Any immediate apprehension is called "feeling," whether it be sensory, emotional, or intellectual. Even memory and the perception of relations were "felt." But the feelings were not images; they were merely the awareness of whatever content might be before one. Destutt de Tracy called these contents ideas, following Locke. They were of four kinds: sensations, memories, judgments, and desires.

The question that puzzled Destutt de Tracy and, for that matter, most of the philosophers of this period in France was whether all consciousness is passive or whether some is active. If all were passive, then we should have no reason to believe in the existence of an external world. There is, however, according to Destutt de Tracy, one idea that gives us an intimation of a reality beyond ourselves, the idea of touch. When we put pressure upon an object, it resists. We cannot, at the same time, desire both a feeling and its annihilation. The feeling of resistance annihilates the desire to penetrate. Therefore, when we feel resistance, we are forced to conclude that there is a resisting object. In this way an element of activity was introduced into Destutt de Tracy's epistemology, an element that was to form the logical nucleus of the theories of his successors, Maine de Biran and Pierre Laromiguière.

Destutt de Tracy thought that the analysis of general ideas into elementary feelings would destroy the analyzer's faith in many of the teachings of religion. For if an idea could not be found to be either an elementary feeling or to be composed of such, it must be discarded. But many religious ideas cannot be so analyzed and therefore must be discarded.

Although the *Idéologues* had favored Napoleon Bonaparte's coup d'état of 1799, they soon opposed him, and in 1803 Napoleon suppressed the Second Class of the Institut. Destutt de Tracy's antireligious views, which

directly clashed with Napoleon's reestablishment of religion, were a major factor in Napoleon's act of suppression. The soon-to-be emperor, moreover, could not tolerate Destutt de Tracy's view that every man has the power to determine the truth and falsity of his ideas without recourse to authority and that among those ideas are those of right and wrong, both moral and political.

See also Cabanis, Pierre-Jean Georges; Condillac, Étienne Bonnot de; Condorcet, Marquis de; Continental Philosophy; Descartes, René; Ideology; Laromiguière, Pierre; Locke, John; Maine de Biran; Volney, Constantin-François de Chasseboeuf, Comte de.

Bibliography

PRIMARY WORKS

Quels sont les moyens de fonder la morale chez un peuple? Paris, 1798.

Observations sur le système d'instruction publique. Paris, 1801.

Eléments d'idéologie, 4 vols. Paris, 1801–1815.

Grammaire générale. Paris, 1803.

Logique. Paris, 1805.

Traité de la volonté et de ses effets. Paris, 1805.

Commentaire sur l'esprit des lois de Montesquieu (*A Commentary and Review of Montesquieu's Spirit of Laws*). Translated by Thomas Jefferson. Philadelphia, 1811.

De l'Amour. Edited by Gilbert Chinard. Paris, 1926.

A Treatise on Political Economy. Translated by Thomas Jefferson. Detroit: Center for Health Education, 1973.

SECONDARY WORKS

Boas, George. *French Philosophies of the Romantic Period.* Baltimore: Johns Hopkins Press, 1925.

Chinard, J. *Jefferson et les idéologues.* Baltimore: Johns Hopkins Press, 1925.

Head, Brian William. *Ideology and Social Science: Destutt de Tracy and French Liberalism.* Dordrecht: Nijhoff, 1985.

Kennedy, Emmet. *A* Philosophe *in the Age of Revolution, Destutt de Tracy and the Origins of* "Ideology." Philadelphia: American Philosophical Society, 1978.

Picavet, François. *Les idéologues.* Paris: Alcan, 1891.

Van Duzen, C. *The Contributions of the Idéologues to French Revolutionary Thought.* Baltimore, 1935.

Welch, Cheryl B. *Liberty and Utility: The French Ideologues and the Transformation of Liberalism.* New York: Columbia University Press, 1984.

George Boas (1967)
Bibliography updated by Philip Reed (2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DETERMINABLES – FUZZY LOGIC

3 volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT

Editor in Chief

MACMILLAN REFERENCE USA

An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

ENCYCLOPEDIA OF PHILOSOPHY
2nd edition

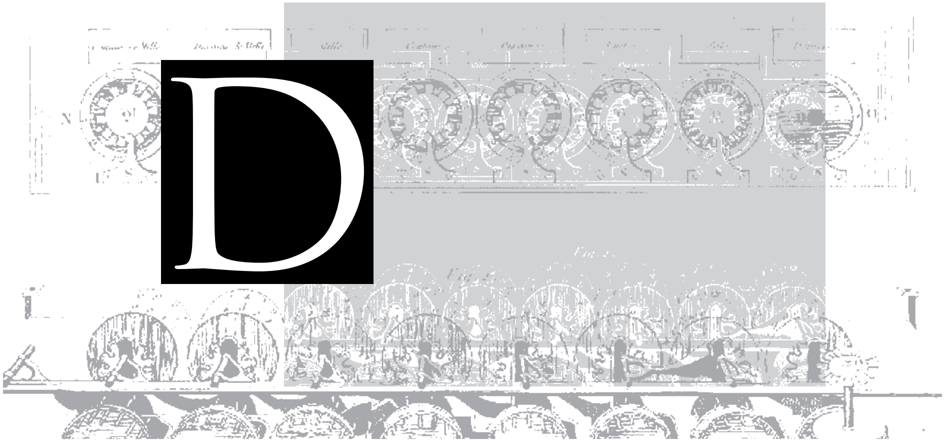

DETERMINABLES AND DETERMINATES

The terminology of "determinables and determinates" existed in scholastic philosophy, but the modern use of these terms originated with the Cambridge (U.K.) philosopher and logician W. E. Johnson, who revived the terminology in his *Logic* (1921). Johnson said, "I propose to call such terms as colour and shape *determinables* in relation to such terms as red or circular which will be called *determinates.*" Some other determinables are size, weight, age, number, and texture. The terminology has since passed into philosophical currency and is now used to mark both the relation between determinate and determinable qualities and the relation between the corresponding words.

The chief features of this relation that Johnson and his successors have found interesting are:

(1) It is logically distinct from the relation of genus to species. The denotation of a species term is marked off within the denotation of a genus term by the possession of properties known as differentia. The species is thus to be construed as formed by the conjunction of two logically independent terms, either of which can, depending on the purposes at hand, be construed as genus or differentia. For example, the species term *man* is defined as the conjunction of the terms *rational* and *animal.* However, the determinate term *red* is not definable by conjoining the determinable term *color* with any other independent term. To put this point another way: Whereas we can say, "All humans are animals which are rational," no analogous statement can be made beginning, "All red things are colored things which are." Any term that could fill the gap would have to be synonymous with *red.* Red things do not possess some trait other than their redness that, when conjoined with their coloredness, makes them by definition red. Both the genus-species relation and the determinable-determinate relation are relations of the less specific to the more specific; but in the former case the specification is provided by some property logically independent of the genus, whereas in the latter case the determinate cannot be specified by adding additional independent properties to the determinable.

This characteristic has been emphasized by Johnson, John Cook Wilson, A. Prior, and John R. Searle; and it is

this feature that chiefly justifies the introduction of this terminology as an addition to the traditional arsenal. Attempts have been made—by Searle, for example—to give a rigorous formal definition of the determinable relation utilizing this feature; but it is not clear to what extent they have succeeded.

(2) Determinates under the same determinable are incompatible. For example, the same object cannot be simultaneously red and green at the same point; and a man six feet tall cannot be simultaneously five feet tall. It might seem that counterexamples could be produced to this point since, for example, an object can be both red and scarlet, and red and scarlet are both determinates of color. However, such counterexamples are easily disposed of on the basis of the fact that scarlet is a shade of red, and hence red is a determinable of scarlet.

We must distinguish the relation in which red stands to scarlet from the relation in which color stands to either red or scarlet. Both are cases of the determinable relation, but they are significantly different. We may think of color terminology as providing us with a hierarchy of terms, many of which will stand in the determinable relation to each other as the specification of shades progresses from the less precise to the more precise. But at the top of the hierarchy stands the term *color,* which we may describe as an absolute determinable of all the other members of the hierarchy, including such lower-order determinables as "red" and their determinates, such as "scarlet."

Our original point can then be restated by saying that determinates under the same determinable are incompatible unless one of the determinates is a lower-order determinable of the other. In the literature of this subject, the counterexamples are usually avoided by saying that any two exact determinates—for example, exact shades of color—are incompatible. However, it is not clear what *exact* is supposed to mean in this context.

(3) Absolute determinables play a special role vis-à-vis their determinates. This role may be expressed by saying that, in general, for any determinate term neither that term nor its negation is predicable of an entity unless the corresponding absolute determinable term is true of that entity. For example, both the sentence "The number seventeen is red" and the sentence "The number seventeen is not red" sound linguistically odd because numbers are not the sort of entities that can be colored. Lacking the appropriate absolute determinable, neither a determinate term nor its negation is true of the entity in question.

To have a convenient formulation of this point, we may say that the predication of any determinate term or its negation of an object *presupposes* that the corresponding absolute determinable term is true of that object. We define presupposition as follows: A term *A* presupposes a term *B* if and only if it is a necessary condition of *A*'s being either true or false of an object *x* that *B* is true of *x.* Thus, in short and in general, determinates presuppose their absolute determinables. No doubt certain qualifications would have to be made to account for the operation of this principle in a natural language. For example, perhaps what is presupposed by *red* is more accurately expressed by *colorable* rather than by *colored.*

Aside from the intrinsic interest of these distinctions, they have proved useful in other areas of philosophy. John Locke's very puzzling discussion of primary and secondary qualities can be illuminated by pointing out that he fails to make sufficient use of the distinction between determinable and determinate qualities. When, for example, he says the primary qualities of a material body are inseparable from it in whatever state it may be, he clearly does not mean that a body must have this or that determinate shape or size as opposed to some other shape or size, but rather that it must have the absolute determinables of the primary qualities: It is a necessary condition of something's being a material object that it have some shape or other, some size or other, and so on.

Again, it is useful to point out that absolute determinables are closely related to categories. The notion of a category (or at least one philosophically important notion of a category) is the notion of a class of objects of which a given term can be significantly predicated. Thus, for example, correlative with the notion of *red* is the notion of things that can significantly be called red; these things are the members of the category associated with *red.* But a necessary condition of something's being a member of the class of things that can be significantly called "red" is that the absolute determinable of *red* must be true of that thing since, as we saw above, determinates presuppose their absolute determinables. Because a category (of the sort we are considering) is always a category relative to a certain term, and because a determinate term presupposes its absolute determinable, the absolute determinables provide a set of necessary conditions for category membership relative to the determinate terms.

Where the absolute determinable provides not only a necessary but also a sufficient condition of predictability of the determinate term, the absolute determinable will

simply denote the members of the category associated with the determinate term. Thus, assuming *colored* (or *colorable*) is the only presupposed term of *red,* the category associated with *red,* and with any other determinate of *color,* is only the class of objects that are (or could be) colored.

See also Categories; Locke, John; Negation; Primary and Secondary Qualities; Prior, Arthur Norman; Properties; Searle, John.

Bibliography

Johnson, W. E. *Logic,* 3 vols. Cambridge, U.K.: Cambridge: University Press, 1921–1924.

Körner, S., and J. R. Searle. "Determinables and the Notion of Resemblance: Symposium." *PAS,* Supp. (1959): 125–158.

Prior, A. "Determinables, Determinates, and Determinants." *Mind* 58 (1949): 1–20, 178–194.

John R. Searle (1967)

DETERMINABLES AND DETERMINATES [ADDENDUM]

The relation between determinates and determinables has certain interesting formal and modal features. It is controversial whether these features are to be explained in terms of something more basic or whether they are primitive.

Formally speaking, the determinate-determinable relation is transitive, asymmetric, and irreflexive. Because *scarlet* is a determinate of *red* and *red* is a determinate of *color*, *scarlet* is a determinate of *color*. Because *scarlet* is a determinate of *red*, *red* is not a determinate of *scarlet*. And nothing is a determinate of itself.

Modally speaking, three features are worthy of note. First, if anything has some property, *p*, then it also has every property, *q*, of which *p* is a determinate. Thus, of necessity, scarlet things are red and colored. Second, the relation guarantees the exclusion of codeterminates. Nothing can have two determinates under a single determinable (provided the determinates are not themselves determination related). Thus, nothing can be both scarlet and crimson, because both are codeterminates of color. Third, and more controversially, any object with a determinable property must have a property that is a determinate of that property. Furthermore, there must be an exactly determinate property under every determinable.

It may be that the modal and formal structure of the determinate-determinable relation is brute, but two theories suggest otherwise. According to David M. Armstrong (1997), codeterminates under a single determinable are partially identical. Having five grams of mass just consists in having one gram of mass five times over. So, the exclusion relation is neatly explained by appeal to familiar facts about identity. Nothing can be five grams of mass and one gram of mass for the same reason that no room can have exactly one lectern and exactly five lecterns. However, the notion of partial identity for properties, as opposed to individuals, remains unclear.

Sydney Shoemaker (1984, 1998) holds that properties are individuated by the causal powers they bestow on objects that instantiate them. This theory of properties provides a ready explanation of the nature of the determinate-determinable relation: The powers endowed by a determinable property are a proper subset of the powers endowed by a determinate of that property (2001). For example, *scarlet* bestows the power to trigger scarlet detectors as well as red and color detectors. Some of the modal and formal features of the relation are then explicable simply by appeal to set theory, with its transitive, asymmetric, and irreflexive relation of proper subsethood. For example, if anything is scarlet, then it is also red, because if anything has the set of causal powers endowed by *scarlet*, then it has every subset of the causal powers in that set, and one of those subsets corresponds to *red*. The exclusion of codeterminates requires another explanation, however, which appeals to the individuation of powers. If an object were both scarlet and crimson, it would have incompatible causal powers, that is, it would be disposed to act in contradictory ways in the identical circumstances.

This reduction of the determinate-determinable relation would be more satisfying were the causal theory of properties that underwrites it less controversial. Among the more surprising consequences of the theory is that the laws of nature are strictly metaphysically necessary. Moreover, the theory is not perfectly general, but applies only to certain properties. The causal relation itself, along with purely formal properties like self-identity, cannot be correlated with a unique set of powers, but such noncausal properties may nevertheless stand in determinate-determinable relations.

One other characteristic is worthy of note: Determinates and determinables do not compete for causal efficacy. If a scarlet patch sets off a red detector, it is appropriate both to say the detector was triggered by the red and that the detector was triggered by the scarlet. The

overdetermination here is harmless, which raises the possibility that the relation may be appropriated by nonreductive physicalists seeking a way to preserve the causal efficacy of the mental in a physical world; perhaps physical properties are determinates of mental determinables (Yablo 1992).

The fit is not quite right, however. To repeat the point made earlier, the determinate-determinable relation is not the genus-species relation, nor is it merely one of greater and lesser generality. A perfectly determinate shape may be realized in different materials, but the conjunctions of that shape with different types of material do not form further determinates of shape. Likewise, mental properties may still admit multiple physical realizations even if they are perfect determinates of thought (Funkhouser).

See also Armstrong, David M.; Properties; Set Theory; Shoemaker, Sydney.

Bibliography

Armstrong, David M. *A World of States of Affairs*. New York: Cambridge University Press, 1997.

Funkhouser, Eric. "The Determinate-Determinable Relation." Forthcoming, *Nous*.

Shoemaker, Sydney. "Causal and Metaphysical Necessity." *Pacific Philosophical Quarterly* 79 (1998): 59–77.

Shoemaker, Sydney. "Causality and Properties." In *Identity, Cause, and Mind: Philosophical Essays*. New York: Cambridge University Press, 1984.

Shoemaker, Sydney. "Realization and Mental Causation." In *Physicalism and Its Discontents*, edited by Carl Gillett and Barry Loewer. New York: Cambridge University Press, 2001.

Yablo, Stephen. "Mental Causation." *Philosophical Review* 101 (1992): 245–280.

Troy Cross (2005)

DETERMINISM, A HISTORICAL SURVEY

Determinism is the general philosophical thesis that states that for everything that ever happens there are conditions such that, given them, nothing else could happen. The several versions of this thesis rest upon various alleged connections and interdependencies of things and events, asserting that these hold without exception.

There have been many versions of deterministic theories in the history of philosophy, springing from diverse motives and considerations, some of which overlap considerably. We shall consider these in the order in which they have been historically significant, together with certain alternative theories that philosophers have proposed. There are five theories of determinism to be considered, which can for convenience be called ethical determinism, logical determinism, theological determinism, physical determinism, and psychological determinism.

ETHICAL DETERMINISM

ADVOCATES. It seemed to Socrates that every man always chooses what seems to him best, that no man can set as the object of his choice something that seems evil or bad to him. Plato had much the same view, arguing that no man who knows what is good can possibly choose anything else. They drew the obvious corollary that wrongdoing or the pursuit of evil must always be either involuntary or the result of ignorance.

A thirsty man, for example, might choose to drink from a certain cup in ignorance of the fact that it contains poison, or, knowing its contents, he might be forced to drink from it. But he could not, knowing that it contained poison and that this would bring upon him a great evil, voluntarily drink from it. Socrates and Plato thought that similar reasoning applies to any choice whatsoever. Hence, the Socratic doctrine that virtue is knowledge and vice ignorance. If one knows the good, he automatically seeks it; if one seeks something else, it can only be because he is pursuing an apparent, but specious, good—in other words, because he is ignorant of what is in fact good. An obvious corollary to this, and one that was drawn by Plato, is that the best commonwealth would be one governed by philosophers—that is, by men who know the good and can intellectually distinguish it from its counterfeits.

It is evident that in this ethical intellectualism, which is so central to Platonism, there is a theory of determinism. Men's voluntary actions are invariably determined by an apparent good; hence, all their actions are determined by this, if by nothing else. Philosophers who have been convinced by this teaching have nevertheless without exception insisted that it enhances rather than debases man's freedom. Freedom, they have maintained, is precisely the determination of the will by what is good. To have one's will or choice determined by what is bad is to be enslaved; to have it determined by something less than the highest good is, to that extent, to be less than perfectly free. Thus, Plato described the wicked tyrant, who pursues what is evil because he is ignorant of the true good, as enslaved and an object of pity.

René Descartes believed that no man who knew his true "end" or highest good could reject it in favor of something less and maintained that man's freedom consisted precisely in knowing that good and being thereby determined to seek it. St. Thomas Aquinas spoke similarly, with qualifications, concerning man's knowledge of his true "end" or highest good. Gottfried Wilhelm Leibniz similarly took for granted the fact that God could not possibly be guided by anything except the true good, which he must surely know, and that in creating a world, for example, he therefore could not create any but the best possible world. Still, Leibniz maintained, this is no derogation of God's freedom; on the contrary, it is the most perfect freedom to have one's will thus determined.

OPPONENTS. Aristotle rejected this theory of ethical determinism, mostly because it conflicts with what he took to be the evident fact of incontinence. It seemed clear to him that sometimes a man's desires or appetites are in conflict with his reason, precisely in the sense that he desires something bad even while knowing that it is bad, which is the very essence of incontinence. John Locke took the same position. A drunkard, Locke pointed out, well knows that his use of spirits is bad for him, but the mere knowledge of this cannot be depended upon to extinguish his desire for them.

Most contemporary thinkers incline to the same view. The moral and intellectual determination of men's choices and the consequent impossibility of genuine incontinence are no longer considered a plausible view by very many. Nevertheless, it is not easy to see just what is wrong with it. Surely, men do prefer the better to the worse in some sense—not what is absolutely better, perhaps, but what at least seems better; otherwise, why would any man choose it? It is the very nature of things bad to be shunned, and that is precisely why they are called bad.

Perhaps the real issue here is the more general opposition between rationalism and voluntarism. If one assumes the primacy of man's reason and supposes his will, or what the Greeks called his appetite, to be naturally subordinate to it, then the Socratic thesis of the determination of the will by the reason is difficult to refute. If, on the other hand, one presupposes the primacy of man's will or appetite and assumes the intellect to be at least sometimes subordinate to the will, then there is no difficulty in accounting for incontinence. Furthermore, there have been many philosophers—for example, Benedict (Baruch) de Spinoza, Thomas Hobbes, and William James—who have insisted that all it means to describe something as good is that it is the object of one's will—that is, of his desire or interest. If this is so, then the Socratic thesis becomes utterly trivial. It amounts to saying nothing more than that the object of a man's will is always an apparent good—that is, something that is the object of his will. This is certainly true but not significant.

LOGICAL DETERMINISM

Very early in the development of Western philosophy it occurred to certain thinkers that logic alone suggests that men's wills are fettered, that nothing is really in their power to alter. This thesis was developed by Diodorus Cronus and others of his school, whom Aristotle sometimes referred to as "the Megarians," and more importantly by the highly influential school of the Stoics. Such views were associated by the ancients with the idea of fate, an idea that has, however, the same implications as certain forms of determinism with respect to human freedom. Thus, if no man's destiny is in any degree up to him, if everything that he ever does is something he could never have avoided, then in the clearest sense it is idle to speak of his having a free will. The Stoics thought that the most elementary consideration of logic shows this to be true.

The consideration in question is simply the supposition that every statement whatsoever is true or, if not true, false. This ultimately came to be expressed in the dictum *tertium non datur,* meaning that no third truth-value, besides true and false, can be assigned to any statement. If this is so, then it must hold for statements about the future as well as any others, for statements about individual men's future actions and even for statements or propositions that are never asserted. It must also, of course, apply to statements believed by the gods. The last idea eventually became very important when the belief in an omniscient and infallible god became theological dogma.

What apparently led certain ancients, such as Chrysippus, Posidonius, and the Stoics generally to take the idea of logical determinism seriously was a consideration of signs, omens, and portents, which were then widely believed in. If there are signs from which it can be discovered what is going to happen, especially what a certain man is going to do at a certain time, and if, moreover, such signs are vouchsafed to men by gods, then it seems that such predictions must unavoidably, in the fullness of time, be fulfilled. Any such prediction that was not fulfilled could not have been true when made, contradicting the supposition that it was true. If such a prediction must be fulfilled, then it seems to follow that it is not within

anyone's power to confute it. The extension of this thought to all actions of all men leads quite naturally to the view that no man's actions are ever free or that nothing any man ever does was ever avoidable, it having always been true that he was going to do whatever he eventually did.

ARISTOTLE'S OPINION. A penetrating discussion of this problem is contained in some much disputed passages of Aristotle's *De interpretatione.* Aristotle there considers the question whether every true proposition, asserting that a certain event has occurred at a certain time, was true before the event in question took place and whether every false proposition, asserting that a certain event has occurred at a certain time, was false before the event failed to take place at that time.

Suppose, for example, a naval battle took place yesterday. This would seem to entail that it was already true, prior to yesterday, that it was going to occur. If anyone had said a thousand years earlier that such a battle was going to occur that day, then it would seem that his prediction was true, and if anyone had denied it a thousand years earlier, then the events of that day would have shown him to have been wrong. Aristotle, however, seemed reluctant to make this seemingly obvious inference. He suggested that it is inconsistent with the fact that men sometimes deliberate about whether to make certain things happen and with the belief all men have that it is sometimes up to them whether the events about which they deliberate will occur. If it is true a thousand years before a naval battle occurs that it is going to occur on a certain day, then whether or not anyone actually makes the prediction, it is difficult to see how, when that day arrives, it can still be up to the naval commander whether the battle will occur or what point there could be in anyone's deliberating about whether to precipitate it. The same difficulty arises if one supposes it to have been false a thousand years earlier that a naval battle would later occur. Aristotle therefore seems to suggest that some propositions—namely, those which assert or deny the future occurrence of certain deliberate actions of men or of events which are dependent upon these—are sometimes neither true nor false until the actions have either occurred or failed to occur.

SUBSEQUENT CONTROVERSY. This whole question was highly vexing to the early thinkers who followed Aristotle. It was even more troublesome to the Scholastics, many of whom felt bound to affirm the freedom of the human will but also bound to affirm that God knows from the beginning of time everything that will ever happen in his creation. Most of the Stoics, whose philosophy was highly fatalistic anyway, embraced the view of logical determinism or fatalism, while many of the Epicureans, who from moral considerations had always set themselves against any theories of fatalism, sometimes defended the view that statements about the future need not be either true or false and hence could not be known in advance even by the gods.

Diodorus Cronus was perhaps the most polemical of the early advocates of logical determinism. His fundamental principle, which is obviously a very strong one, was that it always follows from the fact that something *has* happened that it was *going* to happen and, hence, that it was true that it was going to happen before it did happen. Applying this seemingly incontestable dictum, Diodorus concluded that nothing is ever possible except what actually happens, from which it follows that it is never within any man's power to do anything except what he actually does.

Among the problems to which this conclusion gave rise was one called "the idle argument," which states that there is never any point in any man's ever taking any precautions or making any preparations. If, for example, a man is ill, then it follows from Diodorus's principle that he is either going to recover or he is not going to recover. If he is going to recover, then he will recover whether or not he summons a physician; similarly, if he is going to perish, then he will perish whether or not he summons a physician. Hence, there is no point in his summoning a physician in either case because the outcome is already inevitable. The philosopher Chrysippus sought to resolve this evident absurdity by inventing the notion of "condestinate" facts, facts whose truths are dependent upon one another. Thus, it may be true that a man is going to recover from his illness and also true that he is going to recover only if he summons a physician, from which one cannot conclude that he will recover whether or not he summons a physician. The two facts are, in this case, "condestinate."

CONTEMPORARY ANALYTICAL DISTINCTIONS. Contemporary philosophers have for the most part tried to resolve the problems of logical determinism by distinguishing between modal concepts, such as necessary, impossible, and so on, and the nonmodal concepts of true and false and by refusing to make certain inferences from one kind of concept to the other. Thus, from the fact that something happens of necessity, it follows that it happens, and from the fact that it is impossible for something to happen, it follows that it does not happen. The

reverse of these inferences cannot be made, however; something might happen without being necessary, and something might fail to happen without being impossible. This permits one to say without contradiction that it is true, without being necessary, or false, without being impossible, that a certain man is going to perform a certain action.

The difficulty that some writers have found in this seemingly obvious solution is that "necessary" and "impossible," as applied to human actions, do not mean logically necessary and impossible. (As Gilbert Ryle and others have noted, the only things that can be logically necessary or impossible are propositions, not events or actions.) When the ancients described an event or action as necessary, they simply meant that it was unavoidable, and when they described it as impossible, they meant that it was not within the power of an agent to bring it about. This is still what men mean by such locutions. It is surely not obvious how an action can be avoidable on the supposition that it has been true from the beginning of the world that it would be performed by a certain man at a certain time and place, and it is not obvious how it can be within the power of an agent to perform a given action on the supposition that it is eternally false that he will. Still, as critics of this line of thought have forever pointed out, we must take for granted that men are often able to do many things which they never do and to forgo many things which they do all the time. It is perhaps just this that has always been at issue.

Following the suggestions of Aristotle, some contemporary philosophers, such as Charles Hartshorne, have maintained that predictions concerning a man's future voluntary actions are always false, the truth being expressed only by a statement to the effect that he might and might not perform them. Others have argued that such predictions are neither true nor false when made, though they eventually become either true or false. In this connection Ryle has suggested that "correct" and "incorrect," as applied to predictions of this sort, are more like verdicts than descriptions and thus convey more the idea of "fulfilled" and "unfulfilled" than of "true" and "false." It would be always wrong to call a prediction fulfilled as long as it is a prediction, and similarly, Ryle suggests, it is misleading to speak of predictions as having been true. Ryle and others have also noted the error of thinking of predictions as the causes of the events they predict, though essentially the same error was pointed out by St. Augustine and many of the Scholastics, who noted that God's prescience is never by itself the cause of anything.

Perhaps the most significant upshot of this whole problem, however, has been the considerable contemporary philosophical discussion concerning the status of future things, particularly future contingent or undetermined things. Do they exist "in the future," awaiting only the lapse of time in order to become present, or do they have the more nebulous status sometimes referred to as possible existence? Ryle has suggested that predictive statements are not true or false in the same way that statements about past things are, precisely because the things to which they ostensibly refer do not have the same determinate existence, and that some descriptive statements therefore cannot make sense until the things ostensibly described really do exist. He thus compares certain predictive statements, such as the statement that a given man is going to cough at a certain future time, with statements about "past" things which might have been but never were—for example, certain automobile accidents that were prevented. All these suggestions have raised some of the most vexatious questions in contemporary metaphysics, and they are very far from being resolved.

THEOLOGICAL DETERMINISM

With the development of Christian theology there arose the concept of a God who is, among other things, perfectly good, omniscient, and omnipotent and upon whom, moreover, the entire world and everything in it, down to the minutest detail, are absolutely dependent for existence and character. This idea is obviously loaded with possibilities for deterministic theories, and there have been many philosophers and theologians who have developed them into extensive systems, some of which have formed the basis for theological doctrines having an extremely wide and abiding influence.

MORAL DETERMINATION OF GOD'S WILL. If, for example, we consider first the absolute goodness of God, it seems incongruous not only to think of him as choosing or by his action inflicting evil, but equally of his being able to choose, inflict, or even permit evil. Since, moreover, the world is the result of his act of creation, it seems to follow that it is the only world that was ever possible, being of necessity the best that was possible. Many of the Stoics affirmed this conception, identifying the world or "nature" with God or Zeus and also with fate. The world, they thought, is the only possible world, and nothing in it could be different from what it is. It is nevertheless good, and so the aim of a wise man should simply be to find and accept his place in it. Spinoza's philosophy contains essentially the same idea. In the first book of his *Ethics* he affirms that nothing in nature is contingent, that there is

no free will in God, and, hence, that things could not have been produced by God in any other manner, though Spinoza was led to these conclusions by considerations other than the mere goodness of God.

Perhaps it was Leibniz who tried hardest to reconcile the moral determinism implied by God's absolute goodness with the existence of alternative possibilities. Leibniz distinguished two senses of necessity, which he called absolute and hypothetical. Given the absolute goodness of God, he said, then the world that exists must be the only possible world, because it is of necessity the best possible one. But this is only on the hypothesis that God is good; hence, the exclusive necessity of this world is only hypothetical. In the absolute sense, not taking into account God's goodness, this world is only one of many possible worlds, contrary to what Spinoza maintained. Something is necessary in an absolute sense only if its negation involves a contradiction, and in this sense neither God's acts nor men's are necessary. The actions of men are necessary only in the sense that there is a sufficient reason for them, as for everything else. This is consistent with their being free, considered in themselves, Leibniz thought, since in no absolute sense are they necessary.

It is doubtful, however, whether Leibniz's distinctions supply more than a verbal solution to the problem of theological determinism. One can grant that this must be the only possible world given the hypothesis that it is the creation of an absolutely good creator and thus agree that apart from that hypothesis it is not the only possible world. But as soon as one affirms God's goodness, which traditional theology considered beyond doubt, then it is difficult to see in what sense alternative worlds are still "possible." Leibniz's concept of hypothetical necessity has nevertheless had the most far-reaching significance in the subsequent development of the ideas of determinism and free will, for it became a cornerstone for generations of later philosophers, like David Hume, in their attempted reconciliations of physical and psychological determinism with free will.

DIVINE OMNISCIENCE AND DETERMINISM. The omniscience of God has likewise seemed to many thinkers to imply the inevitability of everything that happens. The philosophical arguments involved in this kind of determinism, resting on the idea that all truths are eternal, are essentially the same as those which led Diodorus and others to assert fatalism, but the addition of the premise that there is a being who knows all truths from the beginning of time gives these arguments an especially powerful appeal to the imagination.

St. Augustine. An omniscient being knows everything. St. Augustine and virtually every other theologian who contributed greatly to the development of Christian thought assumed without question that God, as thus conceived, must know in advance every action that every man is ever going to perform, including, of course, every sin he will ever commit. If this is so, then the question arises of how men can behave otherwise than God knows they will—how, for example, a man can forgo those sins that God, when he created the man, knew he would commit. The strongest concise way of expressing this point is to say that (1) if God knows that I shall perform a certain act at a certain time and (2) if I am nevertheless able to forgo that act when the time for performing it arrives, then (3) it follows that I am at least able to confute an item of divine knowledge, whether or not I actually do so. That conclusion, of course, is absurd. The second premise, accordingly, must be false if the first is true.

Carneades, a pre-Christian defender of human self-determination and freedom, maintained that even Apollo could not know in advance what men were going to do. Such a view, however, seemed so inconsistent with the notion of omniscience that hardly any Christian thinker entertained it. St. Augustine, in considering this question independently of the idea of God's power, maintained that God's foreknowledge constitutes no threat whatsoever to man's free will. God, according to St. Augustine, foresees all events because they are going to occur; they do not occur just because he has foreseen them. Thus, he compared God's prescience to a man's memory. The fact that someone remembers an event does not render that event necessary or involuntary, and the same is true with respect to God's foreknowing an event. Again, St. Augustine pointed out, there is no difficulty in the notion of God's foreknowing that someone will be happy, from which one can hardly conclude that such a man must therefore be happy against his will. And whether or not we do anything else voluntarily, it can hardly be denied that we *will* things voluntarily, and this constitutes no reason why God should not know what we are going to will. Many of the other events God foreknows are things that, as God knows, depend upon our wills for their happening, from which it follows that they are both foreknown and willed—that is, voluntary. Most of the apparent difficulties in reconciling divine prescience with human freedom seemed to St. Augustine to evaporate in any case as soon as one comprehends the nature of God's eternity. The distinctions of "before" and "after," which

are essential to the formulation of this kind of theological determinism, have no application to God, according to St. Augustine. His eternity is not an everlastingness but, rather, an existence that is altogether independent of time. God therefore sees the whole of history in a manner similar to that in which we view the present, and from this point of view one is not easily tempted to suppose that God's knowledge imposes any determination on things to come.

SUBSEQUENT VIEWS. St. Augustine's reflections on this problem have for the most part been followed by subsequent thinkers. St. Thomas Aquinas, for example, similarly emphasized the eternity of God's vision and argued that God's knowledge is not by itself the cause of anything. Boethius, in *The Consolation of Philosophy,* defended the same view, adding numerous analogies to increase the plausibility of his arguments. Thus, he noted, a sign shows that to which it points without thereby producing it. In the same way God knows what will come to pass, but his knowing does not cause anything to happen. Again, a man might at one and the same time see another man walking and the sun rising; yet the man's walking can be voluntary, whereas the sun's rising is not. This, Boethius maintained, is the manner in which God views all things from the perspective of eternity. Boethius was thus led to his famous definition of eternity as "the simultaneous and complete possession of infinite life." In such a conception there is no suggestion of succession in time, and God must thus see all things in a manner similar to that in which we view things spread out in a given moment.

This Augustinian solution to the problem, echoed so often in the subsequent history of thought, has not been without dissenters, however. In the fourteenth century Peter Aureol reaffirmed what he took to be the arguments of Aristotle, maintaining that propositions concerning particular future contingent events, such as men's acts of free will, cannot be either true or false. This would seem to imply, of course, that God cannot foreknow them, but Peter Aureol seemed reluctant to draw that heterodox conclusion. He observed that God's foreknowledge does not make anything true or false and is to that extent consistent with the lack of either truth or falsity in some such propositions. He apparently did not observe that in order to be known by God, a proposition must nevertheless be true when foreknown, since God obviously cannot know something to be true that is in fact neither true nor false. William of Ockham expressed similar doubts but, unlike Peter Aureol, was unwilling to reject either the law of excluded middle or the doctrine of divine omniscience. God, according to William of Ockham, is omniscient and hence knows all future contingent events. In the case of any disjunction to the effect that a given contingent event either is going to occur at a given time or is not going to occur at that time, God knows which of the mutually inconsistent propositions is true since he is omniscient. It follows that one of them is true and the other one false. But, according to this thinker, no one knows how this is possible, and no philosophical arguments, such as St. Augustine's, can render it really intelligible. Ockham's position thus consisted essentially of simply affirming what he thought was required by both logic and faith and refusing to render either intelligible in terms of the other.

The attempts of St. Augustine and many others to reconcile God's omniscience with the indetermination of men's actions were entirely rejected by the eighteenth-century American theologian Jonathan Edwards, who maintained that divine prescience imposes the same necessity upon things as does predestination, a doctrine that had been taught by St. Augustine. Foreknowledge, Edwards agreed, does not cause those things that are foreknown, but it nonetheless renders them certain and therefore inevitable. Indeed, such foreknowledge could not exist if determinism were not true, for there can be no certainty with respect to any contingent things. To say that things are foreknown with certainty by God and are nevertheless contingent and thus uncertain struck Edwards as an evident absurdity.

Similar doubts are expressed, among contemporary philosophers, by Charles Hartshorne. Hartshorne has defended indeterminism and free will, and defending also the belief in God, he has proposed an exceedingly interesting revision of the idea of omniscience. An omniscient being, according to him, is one who knows everything that it is possible to know. There can, however, be no antecedent truth with respect to particular future free actions of men other than that they might and might not occur. God, accordingly, cannot know whether they will be performed until the time for the performance arrives. He is nevertheless omniscient, since only those things that are inherently unknowable are unknown to him. It is significant and rarely noted that this is precisely the position taken by St. Thomas Aquinas with respect to God's omnipotence. God, according to St. Thomas, is omnipotent not in the sense that he can do anything whatsoever but, rather, that he can do anything that it is possible to do.

DIVINE POWER AND PREDESTINATION. It was earlier noted that the three chief sources of theological determinism are God's presumably unlimited goodness, knowledge, and power. It is undoubtedly the third of

these alleged attributes that has been the richest source of such theories. Even St. Augustine, although he defended human freedom on other grounds, felt obliged to relinquish it in the light of his conception of God's power. Thus arose the doctrine of predestination and all the baneful consequences it has wrought in the history of Christendom.

A man's power, St. Augustine thought, is nothing in comparison to that of his maker. Indeed, a man is helpless to do anything except sin unless he is assisted by the power and grace of God—"God worketh in us both to will and to do." Adam, our first ancestor, was, to be sure, free and, hence, free not to sin, but he sinned anyway and thereby cast the entire race of men into a morass of sin from which it is unable to lift itself by its own power. God as well as the blessed are unable to sin, but men are unable to avoid it. Accordingly, no man can be saved by the exercise of his own will, which can lead him only to damnation. He can be saved only by being chosen by God.

The same opinions were promulgated by Martin Luther and John Calvin, particularly in Luther's dispute with Desiderius Erasmus and Calvin's dispute with the Arminians on the issue of man's free will; they formed a considerable part of the theological basis of the Protestant Reformation. Both Luther and Calvin stressed the power, sovereignty, and righteousness of God, subordinating to these the belief in his love and mercy. God, according to Luther, does not merely foreknow what will happen. He foreknows, purposes, and does everything according to his eternal, changeless, and infallible will. To affirm any power or freedom on man's part, particularly any freedom to perform meritorious actions, seemed to both Luther and Calvin to compromise the power of God and even to set men in competition with him. Without God's grace everything we do is evil and therefore determined. It is not within any man's power to do any good thing. Even actions which would otherwise be right and proper, such as acts of charity, are, according to Luther, without merit if not accompanied by faith and prompted by grace. Luther thus compared the human will with the will of a beast of burden, which is ridden by either God or Satan. If ridden by God, it goes where God wills, and if by Satan, where Satan wills; in neither case, however, does it choose the rider. The riders, God and Satan, vie over who shall control it. Such views as these were once, of course, the source of persecutions and upheavals, but they are rarely enunciated with seriousness now, even by theologians, for the idea of divine power no longer has the reality in men's minds that it once had.

PHYSICAL DETERMINISM

Modern theories of determinism were inspired mainly by the development of physical science, particularly in the seventeenth and eighteenth centuries. Scientists then discovered that the motions of the heavenly bodies were not only regular but also "obeyed" certain laws that could be expressed with mathematical exactness. Gradually, the whole approach to the study of nature, which had been philosophical, speculative, and heavily influenced by Aristotle, gave way to observation, experiment, and the search for laws. The idea slowly took hold that all things in nature, men included, behave according to inviolable and unchanging laws of nature. In the philosophical tradition there was a great deal that made this idea plausible, reasonable, and almost inevitable. Theories of determinism were about as old as philosophy. The rise of physical science only prompted philosophers to revise somewhat the content of deterministic theories to which they were already thoroughly accustomed. They more or less ceased thinking of human actions and other events as determined by moral considerations or by an eternal and immutable God and began thinking of them as determined by eternal and immutable laws of nature.

THE EPICUREANS. Of course, this idea was by no means new. The view that everything is composed of matter or, more precisely, of minute and impenetrable atoms or invisible material particles had been elaborated by Leucippus and Democritus before the Christian era and had been perpetuated in the teachings of the Epicureans for centuries. Such a conception of nature gave rise to the idea that if everything that happens is resolvable into the motions and combinations of atoms, then men's behavior, too, must be reducible to and understandable in terms of the motions of atoms. The early atomists assumed that this must be true even of men's thoughts and desires, since, according to them, even the "soul" is composed of atoms. The behavior of atoms, in turn, was thought to be a function of their speed, direction of motion, and sometimes their shapes. Atoms changed the direction of their motion simply by being struck by other atoms. Material bodies arose from the combination of atoms into groups or clusters and perished as a result of their dispersion. The atoms themselves, however, were individually indestructible and indivisible.

The Epicureans who took over this theory of nature were not long in discovering its implications with respect to human freedom. These philosophers were concerned mostly with discovering the means to the attainment of the highest good for man, which they took to be happi-

ness and freedom from pain. It would be idle, however, to work out the means for the attainment of this if men had no freedom to choose those means. If the theory of atomism were true, then it would seem that what became of a man and whether he attained a good life were simply matters of how physical bodies and, ultimately, the atoms of which all bodies are composed behaved, and no man would have any hand in what became of him. The Epicureans accordingly modified the theory by claiming the atoms to have the power of occasional spontaneous motion, which they referred to as the capacity to swerve. Ordinarily, an atom would change its direction only by being driven from its path by impact with another atom, but occasionally, they maintained, an atom alters its path spontaneously, without any cause for this change at all. This enabled the Epicureans to maintain that there is an element of contingency and uncertainty in nature, that not everything is determined by physical laws, and that men can therefore intelligibly be thought of as free to some extent or, in modern terms, as having free will. The Epicureans' opponents never tired of waxing merry with the doctrine of the swerve, however. Indeed, that doctrine did enable the Epicureans to avoid determinism, but there appeared to be nothing else in its favor, and it seemed, moreover, to be plainly irrational.

HOBBES'S MATERIALISM. Perhaps the best example of physical determinism in modern philosophy is the system of Thomas Hobbes. His philosophy represents a thoroughgoing attempt to interpret human nature according to the basic presuppositions of the science of bodies—that is, physics—and although it is no longer novel, it is probably fair to say that the generations of thinkers since Hobbes who have shared his aim and purpose have not significantly modified or improved upon his fundamental ideas. Modern materialistic theories differ from Hobbes's basic system only in details and mode of expression and share equally with it such purely philosophical merits and defects as it may possess.

Hobbes denied the existence of any immaterial soul or spirit in men, maintaining, as do some contemporary materialists, such as J. J. C. Smart, that ideas, sensations, and all psychological processes are motions or modifications of matter in the brain. From this it at once follows that human behavior is the behavior of matter and is to be understood according to the same general principles that we apply to matter. The idea that men might be the original sources of their own voluntary motions or that acts of will might arise without causes was rejected as unintelligible; nothing, Hobbes said, "taketh a beginning from itself." Whatever happens, whether in the realm of human behavior, human thought, or elsewhere is caused and hence causally determined by changes of material particles. Voluntary actions are therefore no less necessitated than anything else.

Hobbes nevertheless insisted that such complete physical determinism is consistent with human liberty, for he defined liberty as simply the absence of external restraint or impediment and, hence, as something that even inanimate things can possess. He said that, properly understood, liberty is simply the "absence of all the impediments to action that are not contained in the nature and intrinsical quality of the agent." Hobbes concluded that any unobstructed moving body can be considered free. The unobstructed water of a flowing stream, for example, descends freely, though it is not at liberty to ascend or to flow across the riverbed. It is part of the "nature and intrinsical quality" of water to flow downward, and it flows freely.

Hobbes interpreted human nature according to such analogies. All voluntary human action, he thought, is caused by the alternate operation of the general motives of desire and aversion, which he took to be similar to, and, indeed, varieties of, physical forces. The proximate or immediate cause of a voluntary motion is an act of the will, but an act of the will is never free in the sense of being uncaused. It is caused by some kind of desire or aversion. Deliberation was described by Hobbes as an alternate succession of contrary appetites, a kind of vacillation between competing impulses, in which the appetites are of such approximately equal force that neither immediately overcomes the other. Deliberation ceases when one of them comes to outweigh and thus to prevail over the other. An "act of will," accordingly, is simply the "last appetite"—that is, the desire or aversion upon which one finally acts. To speak of an agent's act of will as "free" would be equivalent to saying that the agent is able to perform it if he wills to perform it, and this Hobbes dismissed as an "absurd speech." To say a man is free to do a given action means only that he can do it if he wills—that is, that his will or "last appetite" is sufficient to produce that action—but it is obviously nonsense to speak of an act of will itself being free in any such sense. Any other sense of freedom, however, seemed to Hobbes inherently incoherent. It is, for example, a fairly common conception of liberty among the advocates of free will that a free agent is one who, when all things necessary to produce a given action are present, can nevertheless refrain from that action. This, according to Hobbes, is equivalent to saying that conditions might be sufficient to

produce a given effect without that effect's occurring, which is a contradiction.

It is noteworthy that Hobbes, though he claimed all human behavior to be physically determined and necessitated, did not conclude that men are not responsible for their actions. In this his theory represents an important departure from some of his predecessors. The Epicureans took for granted that behavior that is physically determined is unfree, and they therefore denied, in the face of their own presuppositions, that all human behavior is physically determined. But Hobbes maintained that a voluntary act is simply one that is caused by an act of will. It is rendered no less voluntary by the fact that acts of will are caused. Generations of philosophers, while for the most part rejecting Hobbes's materialism, have nevertheless followed him in this and in his conception of liberty. Arthur Schopenhauer, for example, declared it nonsense to ask whether acts of will are free, giving the same reason that Hobbes had given; defined freedom as the absence of impediments and constraints; and, like Hobbes, found no incongruity in speaking of inanimate bodies, such as a flowing stream, as acting freely. In the twentieth century Moritz Schlick, A. J. Ayer, and many others made the point that freedom is not opposed to causation but to constraint. The significance of these ideas is enormous, for they appear to offer the means of once and for all reconciling the apparent opposition between determinism and freedom, thus dissolving the whole problem of free will. Many philosophers are still convinced that this insight is entirely correct and that there really is therefore no problem of free will.

PSYCHOLOGICAL DETERMINISM

Most philosophers since Socrates, and even those before him, have, unlike Hobbes, distinguished between men's minds and their bodies, taking for granted that men are not just collections of material particles. Descartes distinguished minds and bodies as two entirely distinct substances whose essential properties are utterly different. Most philosophers since have rejected much of Descartes's philosophy but have nevertheless preserved the distinction between minds and bodies. In contemporary philosophy minds and bodies are not often described as distinct substances, but an absolute distinction is nevertheless often drawn between "psychological" predicates and verbs, on the one hand, and "physical" ones, on the other, and this amounts to much the same thing. Because of this, most modern theories of determinism, as applied to human behavior, can suitably be called theories of psychological determinism. Most of these theories are in complete agreement with Hobbes's concept of free and voluntary behavior as the unconstrained and unimpeded behavior that is caused by an act of will, a motive, or some other inner event. The only significant difference is that acts of will and other inner causes are conceived of as psychological or mental events within the mind of the agent rather than as modifications of matter in his brain.

CARTESIAN INDETERMINISM. Descartes stands out in modern philosophy as a defender of free will, which is conceived of as indeterminism with respect to the voluntary operations of the mind. In his *Meditations* he described such freedom as infinite, meaning that no limitation whatsoever is put upon the mind's power of choice. His theory was essentially that willing consists of assenting or dissenting to some conceived object of choice or to some proposition. By the understanding one is enabled to entertain certain propositions, but understanding by itself neither affirms nor denies, neither chooses nor rejects. This role is reserved for the will. Accordingly, human understanding can be of limited scope, as it is, without in any way limiting the freedom of the will. The understanding sometimes represents things in an obscure and confused manner, sometimes even falsely, as in the case of various illusions and deceptions, but it sometimes represents them clearly and distinctly. Intellectual error results from the precipitous use of the will—that is, from assenting to things that are not clearly and distinctly perceived by the understanding. Moral error results from a similar unrestrained use of free will—that is, from men's assenting to or choosing objects that are only speciously good, without a clear and distinct apprehension of their true worth. Thus, error is always avoidable. To know what is true, attain genuine knowledge, and choose rightly, one needs only to confine the assent of the will to what is clearly and distinctly perceived by the understanding as true or good. God cannot therefore be blamed for men's errors. He endowed men with understanding adequate for the perception of truth and with a will that is absolutely unlimited in its freedom to accept what is true and reject what is doubtful or false.

This way of conceiving of the human will has provided what is virtually a standard solution to the problem of moral evil—that is, to the problem of reconciling the occasional turpitude of men with the presumed goodness of their creator—but beyond that hardly any philosophers have agreed with it. Probably no other indeterminist, for example, has described the freedom of the human will as unlimited. The theory was also quickly subjected to criticism on epistemological grounds. With great perception Spinoza, for example, challenged the basic dis-

tinction between the understanding and the will. It is quite impossible, Spinoza said, to have a clear and distinct understanding of some truth without at the same time assenting to it. The perception of truth is one and the same thing with the knowledge of it, and one cannot therefore have a true idea without at the same time knowing that he has a true idea.

Much more important, however, were the implications of Descartes's idea of a "free" will, conceived of as a will that is not determined by anything else. It appeared to imply that men's choices are completely random and capricious, utterly mysterious and inexplicable. In fact, this has always been the overwhelming stumbling block for all theories of indeterminism, whether in the Epicurean notion of spontaneous swerves of atoms or Descartes's notion of uncaused assents, dissents, and choices. If such things are really free in the sense of being causally undetermined and if human behavior is to be explained in terms of such things, then human behavior itself would have to be random, capricious, and utterly inexplicable. Since, however, human behavior does not appear to be exactly what these theories suggest, there has always been a powerful incentive to reject indeterminism in favor of some conception of determinism that does not do violence to men's conceptions of liberty.

Innumerable philosophers have thought that this is accomplished in the manner suggested by Hobbes—that is, by conceiving of a voluntary action as one that is caused by such an inner event as volition, motive, desire, choice, or the like; conceiving of an involuntary action as one that is caused by some state or event external to the agent; and then defining a free action not as a causally undetermined one but as one that is not involuntary or constrained. This kind of determinism has been advocated by so many philosophers, including many contemporary writers, that it would be tedious to list them. The basic idea was suggested by Aristotle, although Aristotle did not discuss the problem of free will as such. It was lengthily defended by John Locke, who was, however, aware of some of the difficulties in it, which he never entirely resolved except by enormous equivocations. Probably the most famous classical defense of it was presented by David Hume, who is still thought by many to have solved the problem of free will.

LOCKE'S THEORY OF LIBERTY. Locke, like Descartes, distinguished between a man's mind and his body and described both as substances. Changes in a man's body, including voluntary motions, are, he thought, all caused, but the causes are within the mind in the case of voluntary motions. Unlike Descartes, however, Locke did not suppose that anything within the mind is causally undetermined, nor did he think it necessary to suppose this in order to preserve the belief in human freedom, which he thought misleading to label "freedom of the will."

Locke defined liberty or freedom as "a power in any agent to do or forbear any particular action, according to this determination or that of the mind, whereby either of them is preferred to the other." One acts freely, then, provided he is acting according to the preference of his own mind, and this is perfectly consistent with his action's being causally determined. It might, for instance, be determined by that very preference. Locke also defined freedom as "being able to act or not to act, according as we shall choose or will," and this again, far from implying that free actions are uncaused, implies that they are caused by the agent's choice or will. In the light of this, Locke, like Hobbes, dismissed the question whether men's wills are free as "improper" or meaningless, like asking whether a man's sleep is swift or whether virtue is square. Liberty, he said, is something that can be possessed only by agents, not by their wills.

That an action can be perfectly voluntary and nevertheless unavoidable was, Locke thought, borne out by clear examples. Suppose, for instance, that a man went to a certain room because there was someone he had a strong desire to see and suppose that while he was there conversing with him, someone secretly bolted the door behind him so that he could not leave. Now, Locke pointed out, his action of remaining in the room, entirely in accordance with his own preference and desire, would not cease to be voluntary just because he could not, unbeknown to him, leave if he wanted to.

One acts voluntarily and freely, then, in doing what one wills, prefers, or chooses. Locke distinguished, however, between desires or preferences and volitions, noting that men can prefer certain things they can by no means will. Thus, a man might prefer to fly than to walk, but he cannot will it. Locke defined a volition as "an act of the mind knowingly exerting that dominion it takes itself to have over any part of the man, by employing it in, or withholding it from, any particular action." Elsewhere he defined a volition as "an act of the mind directing its thought to the production of any action, and thereby exerting its power to produce it." A volition, then, is a psychological act that sometimes figures causally in the production of voluntary motion. It is itself causally determined by the mind, and the mind, in the determination of its volitions, is, Locke thought, causally deter-

mined by the satisfaction of doing or continuing a given action or by feeling uneasy in doing or continuing it.

There is, then, throughout Locke's involved, tortuous, and sometimes equivocating discussion the general presupposition that determinism is true and that indeterminism is irrational and unintelligible. The philosophical problem, as he understood it, is simply that of showing that determinism is compatible with what all men believe concerning human liberty. He seemed to believe that once certain crucial concepts, such as "voluntary," "free," and the like, are rightly defined and understood, the problem of free will would evaporate.

HUME ON FREEDOM AND NECESSITY. The defining of the concepts was, in any case, precisely what David Hume set out to do in his celebrated discussion of liberty. According to Hume, all men have always been of the same opinion on this subject, believing both that men are free and that all their actions are causally determined. There is therefore no philosophical problem of free will, and the whole dispute, he thought, has heretofore been purely verbal in character, involving only confusions in the meanings of words.

It was a fundamental point of Hume's philosophy that causation is essentially constant succession, that there is no necessary connection between causes and their effects. Causes, therefore, do not compel the occurrence of their effects; they only precede them. The question of whether human actions are caused, then, is simply the question of whether there is anything with which they are constantly joined. Hume claimed that no one has ever been in any doubt about this. Throughout history certain actions have always been associated with certain motives with the same constancy and regularity that one finds between any causes and their effects. Human actions are caused, then, in the same way that everything else is caused.

Far from concluding from this, however, that no human actions are free, Hume concluded the opposite, for he considered it the very nature of a free action that it springs from the motive of the agent. He therefore defined freedom as being able to act according to the determinations of one's own will—that is, of one's motives—a definition that presupposes that one's free actions are caused. One's actions are not unfree if they are caused but if they are caused by something other than the determinations of one's own will.

Nor does this conception of liberty, according to Hume, vitiate a man's responsibility for what he does. On the contrary, responsibility depends upon the causation of actions by motives. All laws are based on rewards and punishments and thus rest on the assumption that men's motives can be relied upon to have a regular influence on their behavior. There would be no point in appealing to such motives as fear and hope if nothing could be predicted from their operation. Justice, moreover, requires such an operation of motives, for no man can be a fit object of punishment if his actions are in no way traceable to his motives. Indeed, if one could not rely upon the constant and predictable operation of motives, all intercourse with one's fellows would be hazardous or impossible. One could not even invite a guest to his table with any confidence of not being robbed by him, for the knowledge of his honesty and friendliness would in that case provide no assurance. Sometimes, to be sure, men are robbed or murdered when they had every reason to expect otherwise; however, men are also sometimes destroyed by earthquakes and the like when they had no reason to expect it. No one concludes from this that earthquakes are without any causes. Determinism, then, does not imply that all human behavior is predictable in the most straightforward sense of the term, for many unpredictable things are nevertheless causally determined. A man might not know why his watch has stopped and might not have been able to predict that it was going to stop, but this is only because the cause is hidden from him. He does not suppose that there was no cause at all. Similarly, a normally genial man might on occasion be peevish, but this is only due to some cause—some intestinal disorder, for instance—that is hidden from others and perhaps even from himself.

The important question for Hume, then, was not whether all human actions are causally determined, since all men have always been convinced that they are, or whether any human actions are free, since all men have always been of the same opinion on this, too. It is simply the question of how these two beliefs, so universally shared, can both be true, and Hume found the answer to this in analyzing what is meant by saying that one's action may be caused and also free.

DETERMINISM AND RESPONSIBILITY. What is essentially Hume's argument has been repeated by other philosophers and is still vigorously pressed by many of them. There have nevertheless always been doubters who have contended that this is a superficial conception of liberty, that the actions of a causally determined agent can be "free" only in a technical sense that does not at all correspond with the notion of freedom that men in fact have and that moral responsibility requires. A genuinely free action, according to this point of view, is not merely one

that is in keeping with one's preferences, desires, and volitions, but one that is avoidable or, in C. D. Broad's terminology, "substitutable." To say that a given action was free means at least, according to these writers, that the agent could have done otherwise given the very conditions that obtained, not just that he could have done otherwise if something within him had been different. This thought was expressed by Immanuel Kant, who rendered it in the formula "ought implies can." What Kant had in mind was that whenever one rightly judges that a given agent is morally obligated to perform a certain action, he must logically presuppose that the agent can perform it—not just that he can if he wants, prefers, or wills to, but that he can in some absolute sense. This kind of freedom has been aptly called "categorical," as opposed to the "hypothetical" freedom defended by Hume and others, for it is a freedom both to do and to forbear doing a certain action under the same set of conditions.

The difficulty in deterministic theories that all these critics have felt can perhaps be illustrated with an example. Suppose that a given man is often motivated to steal and that in accordance with determinism he always does steal when, prompted by that motive, his efforts to do so meet with no impediment. According to the determinist theory, these actions are then free and voluntary, and he is responsible for them. Suppose further, however, still in keeping with determinism, that he has no control over the occurrence of this motive, that it arises, let us suppose, as a result of an abominable background and deprivation in his youth, that, in short, he is the product of precisely those influences that nourish and perpetuate that motivation. One's inclination may be to say that even given such a background, he did not have to become a thief, but that would not be in keeping with the thesis of determinism. According to that thesis, it was causally determined and, hence, inevitable and unavoidable that he should become whatever he is. It follows from these suppositions, then, that he cannot help being whatever he is and performing just the actions he does perform. We can indeed still say that if he were not the kind of man he is or if he were motivated otherwise than he is or if something had been different, he could then act otherwise than he does; however, any point to ascribing this merely hypothetical kind of freedom to him seems to vanish when we add, as the determinist must, that nothing could have been different, that he could not have been any other kind of man, that he could not have been motivated differently, and that, hence, he could not have acted otherwise than he did.

It was with this sort of thing in mind that Kant, contrary to what he acknowledged to be the requirements of reason, postulated what he called a "causality of freedom" and insisted that the theory of determinism cannot be applied to men. Their freedom, he thought, must be categorical or such that their actions are not entirely determined by factors over which they have no control. The same point was pressed by G. W. Fichte, Thomas Reid, Samuel Clarke, and William James, and among contemporary writers it has been eloquently urged by C. A. Campbell and many others. It was essentially the point that was skillfully made by Henry Mansel in his criticisms of J. S. Mill's determinist theories. Mill defended a theory that was in all basic respects identical with Hume's—that causation is constant conjunction; that men, when acting voluntarily, always act in accordance with their strongest desires or aversions; that justice, morality, and the administration of laws all require such causal determination of behavior, and so on. Mansel argued that when pressed to its ultimate conclusions, this theory did not differ in its consequences from what he called "Asiatic fatalism," or the view that all men are helpless to do anything except what they actually do. Mill denied this by arguing that although one's actions are determined by his will, his will by his desires, his desires by his motives, and his motives by his character, his character is itself amenable to his will. Mill did not, however, succeed in explaining how, according to his theory of determinism, a man's character, which he evidently thought of as the ultimate determinant of his conduct, could be "amenable" to or within the control of his "will," which is merely the expression of his character.

"HARD" AND "SOFT" DETERMINISM. William James is among the relatively few philosophers who, impressed by the kind of argument Mansel directed against determinism, have defended a theory of outright indeterminism or chance. He was, like the Epicureans, led to do so by what he thought were the requirements of morals. Determinism, he said, implies that the world we have is the only possible world and that nothing could have been other than it was; he declared this to be incompatible with the reasonableness of regret and other basic moral sentiments. In the course of his argument he drew a very useful distinction between what he called "hard" and "soft" determinism. By soft determinism he meant all those theories, like those of Hobbes, Hume, and Mill, which affirm that determinism is true and then, by means of what he considered sophistical and contorted definitions, somehow manage to preserve a semblance of certain moral notions like liberty, responsibility, and so on that, accord-

ing to James, are plainly obliterated by any theory of determinism. Hard determinists, on the other hand, are those who affirm what their theory entails—namely, that no man can help being what he is and doing what he does and that moral distinctions are therefore irrational and ought never to be applied to men or anything else.

There have been relatively few defenders of hard determinism, most philosophers preferring instead to try reconciling determinism with morals. Certain materialist philosophers of the French Enlightenment, such as Baron d'Holbach, are exceptions, for they did maintain that men are only helpless products of an impersonal nature who govern neither themselves nor anything else but are simply carried along to whatever destinies the circumstances of their lives inflict upon them. Arthur Schopenhauer sometimes defended the same thought, emphasizing the irrational forces that govern human behavior. The American lawyer Clarence Darrow applied this hard determinism in courts of law with the most devastating effect, saving many men from the gallows not by pretending they were legally innocent but by the simple and eloquent plea that they could not help being what they were and doing what they had done. Among contemporary philosophers the claim that men are not morally responsible, as an implication of determinism, has been vigorously defended by John Hospers, and many others have pointed out the dubious character of soft determinism. The standard "solution" to the problem of free will, embodied in the writings of Hume, Mill, and many others, is as a result no longer considered to be as obvious as it once was, and a decreasing number of philosophers are now willing to speak blithely of free and voluntary behavior's being caused by motives, desires, volitions, and the like.

DETERMINISM AND MODERN PSYCHIATRY. Contemporary psychiatrists are for the most part highly impatient with theories of human freedom, particularly the theories with which philosophers are familiar. Whether all or most human behavior is causally determined is, after all, an empirical question of fact, and psychiatrists profess to know with considerable assurance not only that it is but to some extent what the causal factors are, particularly in cases of deviant behavior. Philosophers have largely been content to speak in general terms of motives, volitions, desires, and the like as the springs of action, but psychiatrists speak of specific unconscious fears, defenses, and hostilities. One finds in their writings, in fact, an extensive and elaborate terminology for the identification and description of hitherto undreamed of forces that are supposed to be the real determinants of behavior, including certain typical human behavior that both the learned and unlearned have long been accustomed to thinking of as rational, deliberate, and free. Philosophical speculations on the problem of free will have, as a result, come to appear rather superficial to many of those who are familiar with psychiatry.

Hospers's opinion. Perhaps no contemporary philosopher has done more toward viewing these problems in the light of modern psychiatry than John Hospers. One can, according to this writer, agree with the philosophers who maintain that freedom is opposed not to causality but to restraint and compulsion and also think of human behavior as being typically caused by human desires and even volitions. He nevertheless advances impressive empirical evidence, drawn from typical cases of the kind long familiar to psychiatry, to show that our very desires, volitions, and even deliberations are the product of unconscious forces, compromises, and defenses that are not only not within our control but whose very existence is usually unsuspected by those—all of us—who are their victims; that they were for the most part implanted in us in our earliest years, to which our memory does not even extend; and that our after-the-fact explanations or reasons for our behavior are mostly illusions and wishful thinking. "It is not," Hospers claims, "as if man's will were standing high and serene above the flux of events that have moulded him; it is itself caught up in this flux, itself carried along on the current." Spinoza compared a man with a conscious stone which thinks it moves freely through the air only because it does not know the cause of its motion, and Baron d'Holbach compared him with a fly riding on a heavy wagon and applauding itself as the driver. Hospers similarly says that a man is "like the hands on the clock, thinking they move freely over the face of the clock," a comparison that is particularly apt in the light of the psychiatrists' claim that the forces that move us lie within us and are normally deeply hidden.

Philosophers almost entirely agree that if a man's behavior is the effect of a neurosis or inner compulsion over which he has no control and of which he usually has no knowledge, then in a significant sense he is not morally responsible, and in any case he certainly is not free. The most common illustration of this is kleptomania. What is philosophically significant about kleptomania is that its victim does act according to his own volition and desire but that the volition and desire are themselves the product of a neurosis. The profound significance of Hospers's view lies in his claim, which with considerable justification he believes is empirically sup-

ported by psychiatry, that virtually all significant behavior is of the same order as kleptomania and other familiar compulsions, having its sources in the unconscious. The issue is accordingly not a philosophical one but an empirical one. It is simply whether, in fact, as Hospers graphically expresses it, "the unconscious is the master of every fate and the captain of every soul." His defense of this claim is an array of fairly typical cases that are quite well understood by psychiatrists—the compulsive gambler who always plays until he loses, the man who inwardly loves filth and so washes his hands constantly, the mother who lets her child perish of illness on the train because she "must get to her destination," and so on. In case histories like these, Hospers believes, we can, if we are honest and sophisticated, see our own lives and conduct partially mirrored and perhaps begin to have some inkling of the unconscious, deeply hidden but powerful forces that almost entirely determine what we are and what we do. If Hospers is right and if psychiatrists do actually know what they confidently claim to know—and it would be very rash to suggest that they really do not—then the problem of determinism versus free will is not, as Hume thought, resolved in a way that accommodates both views. It is, rather, solved, and it is solved on the side of hard determinism with all the enormous and, to some minds, shocking implications that theory has for morals and law.

THE THEORY OF SELF-DETERMINATION. The great difficulty of indeterminism, as previously noted, is that it seems to imply that a "free" or causally undetermined action is capricious or random. If one's action is strictly uncaused, then it is difficult to see in what sense it can be within the control of an agent or in any way ascribable to him. The difficulty with determinism, on the other hand, is that it seems to render every action ultimately unavoidable. The implications of determinism do not therefore significantly differ from those of pure fatalism.

It is partly in order to meet both of these difficulties that some philosophers have defended a theory of self-determination or agency. The essential elements of all such theories are that men are the sources or causes of their own actions; that their being the source or cause distinguishes those bodily motions that are actions from those that are not, the latter being caused by something other than themselves; and that free actions are those that an agent performs or produces but that he is not caused by anything else to perform or produce. This theory thus distinguishes "action," or "agency," as a basic philosophical category, treating actions as different in kind from other events and as not in any way describable in terms of the latter.

The theory of self-determination is most fully and clearly set forth by Thomas Reid in his *Essays on the Active Powers of Man,* though he does not call his theory by that name. The basic idea, however, was, according to Cicero's essay *On Fate,* advocated by Carneades. It has also been defended by G. W. Fichte and Samuel Clarke. Aristotle seems to have had some such conception in mind when he spoke of men and other animals as self-moved, and Kant also seemed to when he ascribed to men a special causality of freedom and distinguished this sharply from ordinary causality. Perhaps its best-known advocate among contemporary philosophers is C. A. Campbell, who ascribes a "creative activity" to "selves"—that is, to minds or persons—and argues that men are capable of originating their own actions in opposition to the inclinations of their characters.

Carneades on causality and freedom. Carneades, in trying to resolve the problems begotten by the Epicurean theory of uncaused swerves of atoms, on the one hand, and the fatalism of their opponents, on the other, suggested that the idea of being uncaused is ambiguous, like the idea of something's being empty. When one describes a vessel as empty, one does not ordinarily mean that it is absolutely empty—that it does not contain even air, for example. One means only that it does not contain oil or wine or whatever one might expect. Similarly, when one says that a man's action was uncaused, one does not mean that it was without any cause at all but only that it had no antecedent cause. This is compatible with its having been caused by the agent himself. Carneades noted, moreover, that the Epicureans themselves ascribe the power of motion to atoms, giving no account or cause of why they should be in motion other than that it is their nature to move. Why, then, may not men be thought of as having a similar original power of motion without supposing that some antecedent force must set them going? When men act freely, he thought, they are simply the sources of their own behavior, which is therefore caused, though not caused by anything external to themselves. One acts unfreely when one is caused to act as one does by some antecedent and external force. This way of viewing the matter, Carneades suggested, does not imply any fatalism, nor does it imply that a man's actions are random, like the swerves of the atoms. To say that a man is the cause of his own action does not imply that he was unable to cause any other action, nor does it imply that his action was uncaused.

Reid's theory. Reid developed many arguments against determinism, which he sarcastically called "the glorious system of necessity," but his own positive theory is remarkably similar to that of Carneades. Reid argued that determinism is inconsistent with a whole range of beliefs that are shared by all mankind and maintained that we have far more reason for adhering to these than for affirming any philosophical theory with which they are inconsistent. In particular, he maintained that determinism is incompatible with deliberation, with morality, and with the pursuit of ends. When, for example, a man deliberates about some possible course of action, he assumes that the proposed end, as well as the means to its attainment, is within his power to accept or reject—that is, that it is up to him whether the end shall be sought and if so, how. Without this belief he could not deliberate. The belief itself, however, is incompatible with determinism, for determinism entails that no act that is performed was avoidable and that in this sense it is never up to any man what he does. Again, all men believe that a basic distinction can be made between acts that are blameworthy, praiseworthy, and neither. Determinism, however, implies that every act that is performed is ultimately unavoidable and, hence, that no such basic distinction can be made. Finally, all men believe they can pursue, sometimes over a long period of time, certain ends that they have previously conceived. This implies, however, that their actions in pursuit of such ends are within their own power and control, which is inconsistent with determinism.

Reid therefore defined the liberty or freedom of a moral agent as "a power over the determinations of his own will," a definition that contrasts interestingly with Hume's definition of freedom as "a power of acting or not acting according to the determinations of the will." In rejecting determinism, Reid did not, however, affirm that human actions are uncaused. On the contrary, he maintained that nothing happens without a cause, that everything that changes is changed either by some other thing or by itself. Not all causes, then, are antecedent and external causes. Some things, such as men, are sometimes the causes of their own behavior. Indeed, Reid took this to be the very reason for calling a man an agent—namely, that he is a being who acts, not merely one that is acted upon. To speak of an agent being caused to act by something other than himself was for Reid a contradiction, so that acting and acting freely amount to the same thing, whereas the idea of a necessary agent amounts to a contradiction.

It is evident that Reid employed the concept of causation differently from Hume. A cause, he said, is not merely some change that always accompanies another. It is always something that has the power to produce a change, whether in itself or in something else, and no man can define it beyond this. In fact, he maintained that no man would even understand any philosophical definition of a cause if he did not first have the idea of causation from the awareness of himself as an agent. There is, then, no reason why men may not be the original causes of their own voluntary actions, which is precisely what all men believe themselves to be. This way of viewing the matter permits us to say that determinism, defined as the thesis that everything that happens is the result of some antecedent cause or causes, is false and, further, that nothing occurs without any cause whatsoever. Reid's philosophy thus overcomes the chief difficulties of both determinism and simple indeterminism. It accomplishes this, however, only by introducing what many philosophers have thought to be an enormous difficulty of its own—namely, understanding how anything can be the cause of its own changes. One is reminded of Hobbes's dictum, "Nothing taketh a beginning from itself." Alexander Bain pressed this difficulty in both Reid's and Samuel Clarke's philosophies, maintaining that it rendered their claims quite unintelligible, and Patrick Nowell-Smith has made the same point against C. A. Campbell's similar views. The idea of something's being self-moved in the sense understood by Carneades, Reid, Clarke, and Campbell is obviously entirely unlike any concept of physics. Accordingly, Nowell-Smith has suggested that it should be understood in the way such physical concepts as self-regulating, self-propelled, self-starting, and the like are understood, thus rendering it less esoteric. It was Reid's view, however, that this seeming difficulty is only a fact, that all men really do consider themselves to be the causes of their own voluntary actions in a sense in which no inanimate things are ever causes, and that we should be guided in our opinions not by what this or that system of philosophy requires but by what the common sense of mankind universally affirms.

THE "STRONGEST MOTIVE." It is fairly common to suppose that a man invariably acts—in fact, must act—in response to his "strongest motive" and that voluntary behavior is therefore always causally determined by such motives. Philosophical determinists frequently fall into this line of thought, sometimes substituting "strongest desire" for "strongest motive," though it is now less common than it once was. It is well illustrated in one of Alexander Bain's discussions of the free will controversy,

in which he writes that "in the absence of prohibition, [an agent's] decision follows the strongest motive; being in fact the only test of strength, of motive on the whole." Again, Bain notes that "any supposition of our acting without adequate motive leads at once to a self-contradiction; for we always judge of strength of motive by the action that prevails" and, further, that the action that follows upon deliberation "testifies which motive has in the end proved the strongest."

It is to the credit of Thomas Reid, with whose writings Bain was familiar, that he exhibited both the source of the considerable persuasiveness of such reflections as these and at the same time their fallaciousness. The reason this kind of claim has seemed so compelling to so many philosophers is that it has functioned as an analytic statement or one that is rendered true by definition of the concept of a "strongest motive." As such, it sheds no light whatsoever on any fact of human nature and leaves entirely unanswered the question of whether voluntary actions are really caused.

What, Reid asked, is the test of whether the motive that is strongest is the one acted upon? It is simply the motive that prevails. The claim that a man acts upon his strongest motive therefore means, Reid noted, only that he acts upon that motive upon which he acts, which is hardly a significant philosophical claim. If, however, we apply any other criterion for distinguishing which motive is strongest, then there is nothing at all to suggest that we always act on our strongest motives. On the contrary, it is a fairly common experience to feel strongly motivated to do something from which we nevertheless refrain from purely rational considerations, for example, or perhaps from moral ones. The temptation here, of course, is to say that the fact that one refrains from a given action only shows that some contrary motive is "stronger," but this indicates that we are again using as our concept of the strongest motive the motive that prevails and saying nothing more than that a man acts upon the motive upon which he acts.

Reid, however, went further than this by denying that motives can be likened to forces and that varying "strengths" can be ascribed to them in the first place. A motive, he said, is not a cause but a rational consideration of a reason. As such, it is something purely abstract, which has "strength" or "weakness" only in the sense of expressing wisdom, prudence, or the opposites. A "conflict of motives" is nothing at all like the conflict of opposing forces, one of which overcomes the other by superior force. It is more to be likened to the conflicting pleas of contending attorneys. One of these can be "stronger" or have more "force" or "weight" than the other only in the sense that it is more reasonable and persuasive. When, accordingly, we speak of rational or intelligible considerations as having "force," "weight," or "strength," we are not using these notions in the sense they have for physics but as metaphors borrowed from physical nature. It is, Reid thought, largely from mixing these literal and metaphorical meanings that some persons are led into theories of determinism and into supposing that human nature bears a greater resemblance to inanimate bodies than it actually does.

CONTEMPORARY PROBLEMS

The problems of determinism are still very lively in philosophy and have recently gained powerful momentum from detailed philosophical analyses of peripheral questions. Most current philosophical discussion bearing on the problem of free will is not aimed directly at whether men have free will, but at a whole host of questions that have been begotten by this long controversy. Ludwig Wittgenstein's reflections have made it evident, for example, that philosophers do not even know what it means to call something an action in the first place or just how some of men's bodily motions qualify as actions while others do not. It is an elementary distinction that is constantly made by common sense, but philosophers have thus far been unable to analyze it. Obviously, as long as this ignorance prevails, there is little point in discussing whether men's actions are ever free. Certain recent writers, such as Arthur Danto, have suggested that the concept of an action is basic and unanalyzable and that it corresponds to nothing that is found in physical science. Previous generations of philosophers often took for granted that an action is a bodily motion caused by some such inner episode as a volition, motive, desire, or choice, but these terms are now used with much greater care.

Gilbert Ryle, in his *The Concept of Mind,* declared volitions to be a fabrication of philosophy, corresponding to nothing that has ever existed, and since his devastating critique of this whole notion there has been great reluctance among scholars even to employ the word. The concepts of desire, motive, choice, and kindred notions have been similarly subjected to criticism, so that fewer philosophers are still willing to speak blithely of them as causes. A. I. Melden, for example, maintained that no particular motive can be described at all independently of the action of which it is allegedly the cause and that its connection with an action is therefore a logical one, not, as Hume and so many others supposed, a causal one. Moreover, Melden pointed out that if an action is con-

ceived of as a bodily motion together with its motive in order to distinguish actions from bodily motions that are not actions, then it is plainly impossible to explain any action in terms of its motive, as philosophers were once so ready to do.

The interpretation of statements expressive of human ability as either disguised or incomplete conditional statements has likewise been considerably unsettled by the precise and detailed analyses of J. L. Austin. In his celebrated essay "Ifs and Cans" this writer maintained that statements involving the locution "I can" cannot possibly require, for their complete sense, the addition of some such hypothetical as "if I choose" but are, instead, to be understood in some absolute sense. Accordingly, they do not, as so many philosophers since Hume have supposed, express the idea of a causal condition at all. "I could have if I had chosen," is similarly claimed by Austin to express a past indicative rather than a conditional despite its grammatical form, for it normally expresses the idea of having had an opportunity or ability rather than the idea of a causal connection between one's choice and one's action. In statements involving the locution "I shall if I choose," the word *shall,* according to Austin, is normally expressive of an intention rather than a simple future tense and thus also differs essentially from other conditionals in the future tense. Such painstaking analyses as Austin's, although not pursued with the explicit aim of supporting or disconfirming any theories of determinism or free will, have nevertheless considerably weakened some of the strongest defenses of determinism since so many of them have more or less presupposed that statements expressive of human ability, which are so central to any discussion of free will, are simply disguised statements of causal conditions and thus are not only consistent with, but actually imply, a theory of determinism for the very understanding of them.

The highly refined and critical inquiries of contemporary philosophy have brought into further question the whole concept of the will. Is willing to do something an act, for instance, or not? If it is, then how does it shed any light on the concept of acting? If it is not, then how does an action differ from any other bodily change having an inner psychological cause? Clearly, no difference is marked merely by applying different names to such things. Furthermore, if there are such things as acts of will, do they or do they not require antecedent causes? If not, then why should any action require an antecedent cause? If so, then how are deliberate or willed actions to be distinguished from simple compulsions?

Closely associated with the notion of the will is that of intending. Doing something intentionally is now seldom thought of as merely undergoing some change as the result of an inner intention, intentions currently being thought of more in the manner in which Reid described motives—namely, as reasons and purposes having a rational content. Again, it is fairly common practice among contemporary philosophers to distinguish sharply, as Reid did, between the causes of an action and the reasons for it. If this is a real distinction, then it follows that whether some human acts are reasonable and intelligible is quite independent of whether they are caused, and there is no absurdity in describing an action as both free, in the sense of being avoidable and not the effect of antecedent conditions, and rational. This line of thought has raised anew the whole problem of understanding purposeful behavior. Men often do certain things in order to achieve certain results, and this appears to distinguish human behavior from the behavior of inanimate things in a fundamental way. When philosophers were more eager than they are now to interpret human behavior within the framework of determinism, many of them assumed that purposeful behavior was simply behavior that is caused by purposes, desires, or intentions, but this conception harbors the same difficulties as the volitional conception of action that Ryle, Melden, and others have so severely criticized. If one is acting in acting purposefully and if action can be distinguished from such other bodily behavior as digestion, perspiration, and the like only in terms of concepts like purpose, desire, or intention, then one can hardly explain purposeful activity as action that is caused by one's purpose, desire, or intention. The connection is conceptual rather than causal. Desires, purposes, and intentions are, moreover, desires for this or that, purposes or intentions to do this or that, and their objects or aims may never be realized. Thus, they are what we sometimes call "intentional" concepts, and there seems to be nothing that completely corresponds to them in the realm of physical science. No inanimate thing, for example, can without metaphor be spoken of as behaving as it does in response to its desire for something which perhaps never has and never will exist, and no engineer who spoke in that manner of even the most sophisticated machine would ever suppose that he had thus given a causal explanation of anything.

More and more philosophers are inviting attention to certain fundamental differences between the way men view the past and the future. The future, some have wanted to suggest, is a realm of possibilities in a sense in which the past is not. This idea is at least as old as Aristo-

tle's philosophy, but the renewed interest in whether men's actions might be free in some sense not countenanced by determinism has quickened interest in it. It is, for example, sometimes contended that there is a fundamental difference between finding that something is true and making something become true, a contention that renders the concept of action more fundamental than it was once supposed to be and raises anew the question of what is meant by acting freely.

The question, then, of whether determinism is true or of whether men have free will is no longer regarded as a simple or even a philosophically sophisticated question by many writers. Concealed in it is a vast array of more fundamental questions, the answers to which are largely unknown.

See also Aristotle; Arminius and Arminianism; Augustine, St.; Austin, John Langshaw; Ayer, Alfred Jules; Bain, Alexander; Boethius, Anicius Manlius Severinus; Broad, Charlie Dunbar; Calvin, John; Carneades; Causation; Chance; Chrysippus; Clarke, Samuel; Descartes, René; Determinism and Freedom; Determinism and Indeterminism; Determinism in History; Determinism, Theological; Diodorus Cronus; Edwards, Jonathan; Epicureanism and the Epicurean School; Erasmus, Desiderius; Fichte, Johann Gottlieb; Hobbes, Thomas; Holbach, Paul-Henri Thiry, Baron d'; Hume, David; James, William; Kant, Immanuel; Laws of Thought; Leibniz, Gottfried Wilhelm; Locke, John; Luther, Martin; Mansel, Henry Longueville; Mill, John Stuart; Peter Aureol; Plato; Platonism and the Platonic Tradition; Posidonius; Rationalism; Reid, Thomas; Responsibility, Moral and Legal; Ryle, Gilbert; Schlick, Moritz; Schopenhauer, Arthur; Smart, John Jamieson Carswell; Socrates; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Time; Voluntarism; William of Ockham.

Bibliography

The literature on determinism and free will is so vast that only a sampling can be given here.

A good though not recent critical history of the controversy is outlined in Alexander Bain's *Mental and Moral Science* (London, 1872), Book IV, Ch. 11. More recent general studies include Sidney Hook, ed., *Determinism and Freedom in the Age of Modern Science* (New York: New York University Press, 1958), which is a collection of papers by contemporary philosophers, and Sidney Morgenbesser and James Walsh, eds., *Free Will* (Englewood Cliffs, NJ: Prentice-Hall, 1962), which brings together carefully selected discussions from classical and modern writers and is intended mainly for students. A widely read but superficial discussion of the problem is contained in D. F. Pears, ed., *Freedom and the Will* (New York: St. Martin's Press, 1963), which is in part the transcription of a series of discussions by contemporary philosophers most of whom are connected with Oxford University.

ETHICAL DETERMINISM

The ethical determinism associated with Plato and Socrates is a theme of Plato's *Protagoras* and *Gorgias,* and certain elements of this theory are treated rather unsatisfactorily in his *Hippias Minor.* Aristotle discusses the theory and related problems in the *Nichomachean Ethics,* Book VII, Ch. 2.

LOGICAL DETERMINISM

The most frequently cited reference in discussions of logical determinism is the ninth chapter of Aristotle's *De Interpretatione.* Among the many more recent discussions of the problems arising from those passages are A. N. Prior's "Three-Valued Logic and Future Contingents," in *Philosophical Quarterly* 3 (1953): 317–326; R. J. Butler's "Aristotle's Sea Fight and Three-Valued Logic," in *Philosophical Review* 64 (1955): 264–274; G. E. M. Anscombe's "Aristotle and the Sea Battle," in *Mind* 65 (1956): 1–15; Richard Taylor's "The Problem of Future Contingencies," in *Philosophical Review* 66 (1957): 1–28; R. Albritton's "Present Truth and Future Contingency," ibid.: 29–46; and C. Strang's "Aristotle and the Sea Battle," in *Mind* 69 (1960): 447–465.

One of the best sources for the ancients' views on both determinism and fatalism and the only source for some of them is Cicero's *De Fato,* translated by H. Rackham for the Loeb Classical Library (London, 1942). The problem of fatalism, conceived of essentially as it was by ancient philosophers, has been extensively discussed in recent literature—for example, in Gilbert Ryle's provocative essay "It Was to Be," which is Ch. 2 of his *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954), and by A. J. Ayer, "Fatalism," the concluding chapter of his *The Concept of a Person* (New York: St. Martin's, 1963). Richard Taylor's "Fatalism," in *Philosophical Review* 71 (1962): 56–66, was followed by many critical discussions by various British and American authors in subsequent issues of the same journal and in *Analysis* 23 (1962) and 24 (1963), and in the *Journal of Philosophy* 61 (1964) and 62 (1965).

THEOLOGICAL DETERMINISM

Leibniz's claim that God could create no world except the best one possible and the implications he drew from this are found in his *Discourse on Metaphysics* and his *Theodicy.* St. Thomas Aquinas's opinions on the moral determination of God's will are set forth in the *Summa Theologiae,* Part I, Q. 19, especially Articles 9 and 10.

The question whether determinism and fatalism follow from the conception of God as an omniscient being has been discussed by countless authors. St. Augustine's views, for example, are reproduced in a selection titled "On Free Will," in Morgenbesser and Walsh, op. cit., and also in *The City of God,* Book XI, Ch. 21. Boethius's famous treatment of the problem is given in *The Consolation of Philosophy,* Book V. St. Thomas Aquinas discusses it in the *Summa Theologiae,* Part I, Q. 14, Article 13. His views and the views of various other Scholastics are given in Frederick Copleston's excellent *History of Philosophy,* Vols. II–III (London, 1950–1953). An extensive defense of theological determinism and

predestination on various grounds is given by Jonathan Edwards in his famous *Freedom of the Will,* edited by P. Ramsey (New Haven, CT: Yale University Press, 1957). Charles Hartshorne's rather novel and perceptive reconciliation of free will with certain theological presuppositions is found in Ch. 3 of his *Man's Vision of God* (Chicago: Willett Clark, 1941). Although some of the foregoing sources raise the question of predestination, this doctrine, developed specifically as an implication of God's power, is more fully developed in St. Augustine's *Treatise on the Predestination of the Saints,* in the Nicene and Post-Nicene Fathers, first series, Vol. V, edited by Philip Schaff (New York, 1902); see also Augustine's *Enchiridion on Faith, Hope and Love,* edited by Henry Paolucci (South Bend, IN: Regnery/Gateway, 1961). Martin Luther's uncompromising denial of human free will is set forth in his polemic with Erasmus, under the title *Discourse on Free Will,* translated by Ernst F. Winter (New York: Ungar, 1961). John Calvin's defense of the same doctrine can be found at the close of the third book of his *Institutes of the Christian Religion.*

PHYSICAL DETERMINISM

The materialism of the Epicureans and the manner in which they tried to reconcile this with free will are beautifully exhibited in Lucretius's *On the Nature of Things*; an excellent source for earlier Epicurean arguments is Cicero's *De Fato.* Thomas Hobbes's materialism and arguments in favor of determinism are most fully expressed in *On Human Nature.* A more readily available source of Hobbes's important writings on this question is a paperback book of selections edited by Richard S. Peters, *Body, Man and Citizen* (New York: Collier, 1962). Arthur Schopenhauer, though he was not a materialist, defended a theory very similar to that of Hobbes in his *Essay on the Freedom of the Will,* translated by K. Kolenda (New York: Liberal Arts Press, 1960).

PSYCHOLOGICAL DETERMINISM

Most discussions of determinism and free will in modern philosophy have been within the framework of psychological determinism, which assumes that human behavior has its origins in psychological causes of various kinds. Descartes's defense of free will within this context is expressed in the fourth of his *Meditations* and also in *The Principles of Philosophy,* Part I, Sections 32–39. John Locke's extremely vacillating but influential discussion is found in *Essay concerning Human Understanding,* Book II, Ch. 21, where he discussed at length the idea of power. The classic attempt to reconcile determinism and liberty was achieved by David Hume in Section 8 of his *Enquiry concerning Human Understanding.* A defense along similar lines has been given, among numberless others, by C. J. Ducasse, in Ch. 11 of *Nature, Mind and Death* (La Salle, IL: Open Court, 1951). A now famous essay expressing essentially the same view was written by Dickinson Miller under the name R. E. Hobart and titled "Free Will as Involving Determinism and Inconceivable without It," in *Mind* 43 (1934): 1–27. J. S. Mill defended Hume's theory in his *Examination of Sir William Hamilton's Philosophy,* the relevant excerpts from which are reprinted in Morgenbesser and Walsh, op. cit.

Problems of moral responsibility are involved in almost every discussion of determinism and are central to most of them. Immanuel Kant's treatment of the problem and his defense of the idea of a causality of freedom are given in his *Critique of Pure Reason,* under the section "Transcendental Dialectic," particularly in his discussion of the third "antinomy," and, more fully, in his *Critique of Practical Reason.* C. D. Broad's influential and highly elaborate analysis, "Determinism, Indeterminism and Libertarianism," appears in his *Ethics and the History of Philosophy* (London: Routledge, 1952) and has been reprinted in Morgenbesser and Walsh, op. cit. Problems of determinism and responsibility are discussed by several authors in Hook, op. cit., particularly in the essays by Paul Edwards, "Hard and Soft Determinism," and John Hospers, "What Means This Freedom?" Both authors vigorously defend determinism and the claim that determinism and moral responsibility cannot be reconciled with each other.

William James's essay "The Dilemma of Determinism," in which the distinction between hard and soft determinism was first made, is included in almost all of the many collections of his popular essays. Most modern and contemporary writers who have defended deterministic theories have also defended some version of soft determinism, though they have seldom used the term itself. Examples, in addition to most of those already mentioned, are Patrick Nowell-Smith, in the last two chapters of his *Ethics* (Baltimore: Penguin, 1954), and A. J. Ayer, in Ch. 12 of his *Philosophical Essays* (London: Macmillan, 1954).

The most thoroughgoing defense of the theory of self-determinism was given by Thomas Reid, in his *Essays on the Active Powers of Man,* of which there have been many editions. A contemporary defense of what is essentially the same theory is given by C. A. Campbell, in Ch. 9 of *Selfhood and Godhood* (London, 1957). The same book contains an appendix in which the opinions of Patrick Nowell-Smith are subjected to a most thoroughgoing criticism. A similar concept is defended by Richard Taylor in "Determinism and the Theory of Agency," in Hook, op. cit. The same theory underlies Taylor's "I Can," in *Philosophical Review* 69 (1960): 78–89, reprinted in Morgenbesser and Walsh, op. cit. Another article that indirectly suggests such a view is Arthur Danto's "What We Can Do," in *Journal of Philosophy* 60 (1963): 435–445. Determinism is also attacked at great length in Konstantin Gutberlet, *Die Willensfreiheit und ihre Gegner* (Fulda, Germany, 1893), and in Ch. 9 of M. Maher, *Psychology* (London, 1940). These two works are written from a Catholic point of view.

A. I. Melden's *Free Action* (London, 1961) offers fairly elaborate and penetrating analyses of a wide range of concepts that have always been central to the free will controversy, such as those of wants, motives, actions, and so on; although the author does not try to prove directly that men have free will, he attacks the bases of certain widely held determinist theories. Gilbert Ryle's *The Concept of Mind* (London: Hutchinson, 1949) contains a chapter, "The Will," which amounts to a devastating critique of the idea that voluntary actions are caused by volitions. J. L. Austin's "Ifs and Cans," which is included among his *Philosophical Papers,* edited by J. O. Urmson and G. J. Warnock (London: Oxford University Press, 1962), is a painstaking inquiry into what is meant by saying of an agent that he could have done otherwise; although it is directed at claims made specifically by G. E. Moore and Patrick Nowell-Smith, it actually attacks the foundations of theories that have been widely held for over a century.

A detailed and annotated bibliography of works on determinism and free will can be found in Paul Edwards and Arthur Pap, eds., *A Modern Introduction to Philosophy,* 2nd ed. (New York: Free Press, 1965).

Richard Taylor (1967)

DETERMINISM, THEOLOGICAL

Theological determinism or predestination is the belief that events are determined or necessitated by God. One form of the traditional belief insists that owing to his omnipotence, God controls the occurrence of things. Another form asserts that his omniscience, making possible his foreknowledge of future events, affects the occurrence of such events. There are also nontraditional forms. Throughout the history of Islamic and Jewish philosophy, the debate over predestination was central.

When Islamic philosophy emerged in Baghdad in the ninth century CE, the religious and intellectual circles in the city had been witnessing a heated debate over the issue of predestination (*al-qadar*). There were three main Islamic views at the time: events in the universe, including human actions, are not predestined (Muʿtazila); all such events are predestined (Jabriyya); some aspects of such events are predestined, whereas others are humanly "acquired" (Ashʿariyya). In treating this issue, Muslim philosophers tried to reconcile Greek rationalism with Islam.

Abū Yūsuf al-Kindī (c. 801–873) and Abū'l-Walīd Ibn Rushd (Averroes, 1126–1198) denied predestination. They interpreted the Islamic revelations to assert that God does not, for example, control human actions. They both believed that at the moment God desires or wills something to happen, it happens. However, neither God's power nor his knowledge necessitates that he desire or will everything that happens to happen. If one reads Ibn Rushd carefully, though, one discovers that for him, God determines all events, because his omnipotence means that he fulfills all possibilities. Such fulfillment includes that of the natures of things and the laws that govern them. The conduct of any being is consequent upon its nature and its laws. In some of his writings, Ibn Rushd also stresses that God's knowledge of things is the cause of those things.

Abū al-Naṣr al-Fārābī (870–950) and Abūʿ Alī al-Husayn ibn Sīnā (Avicenna, 980–1037) adhered to neoplatonic tendencies, according to which everything necessarily follows from God's nature. Even God's nature itself is necessitated to act in certain ways. There is no room for God's will or choice, let alone the will or choice of any other being. This is despite the fact that al-Fārābī and Ibn Sīnā speak of God's omnipotence and omniscience, and even of human free will. However, they do not use these terms in the traditional sense. "Omnipotence," for example, is the ability to fulfill all possibilities, and omniscience is knowledge of universals.

Abū Ḥāmid al-Ghazālī (1058–1111) attacked such philosophical views in his famous work *The Incoherence of Philosophers* (1184). He considered such ideas non-Islamic and classified some of them, for example, God's inability to know particular events, as heretical. In the absence of such knowledge, reward and punishment, which are essential to Islam, become meaningless, especially in light of the Islamic concept of God's absolute justice.

Reward and punishment did not pose a problem for al-Kindī, because he believed that human beings have free will and that God knows particular events. Therefore, reward and punishment are not in conflict with his justice. The other three philosophers mentioned were not concerned about the issue either. For them, God does not reward and punish people. According to al-Fārābī and Ibn Sīnā, following death, bodies eventually disintegrate and souls become close to or distant from God, based on their degree of knowledge. Their closeness is their reward; their distance is their punishment. Reward and punishment are necessary consequences of the souls' conduct in life. To Ibn Rushd, there is no reward and punishment after death. The bodies disintegrate and the individual souls merge with the universal soul.

Moses Maimonides (1135–1204) asserts the Judaic belief that the human soul is intrinsically free, and agrees with the Greek and Muslim philosophers that matter is the source of natural evil. Thus, he absolves God from moral and natural evil, and justifies reward and punishment for the former, because God does not predetermine human action. However, God can intervene under certain circumstances. Maimonides was criticized by many Jewish thinkers for his rational approach to Judaism, which they feared denies some of its basic ideas, for example, that God wills whatever happens according to his knowledge of the natures of things.

See also al-Fārābī; al-Ghazālī, Muhammad; al-Kindī Abū-Yūsuf Yaʾqūb ibn Isḥāq; Averroes; Avicenna; Determinism, A Historical Survey; Islamic Philosophy; Jewish Philosophy; Maimonides; Universals, A Historical Survey.

Bibliography

Fakhry, Majid. *A History of Islamic Philosophy*. 2nd ed. New York: Columbia University Press, 1983.

Hourani, George F. "Averroes on Good and Evil." In his *Reason and Tradition in Islamic Ethics*, 249–269. Cambridge, U.K.: Cambridge University Press, 1985

Inati, Shams. *The Problem of Evil: Ibn Sina's Theodicy*. Binghamton, NY: Global, 2000.

Shams Inati (2005)

DETERMINISM AND FREEDOM

Determinism is the family of theories that takes some class of events to be effects of certain causal sequences or chains, more particularly certain sequences of causal circumstances or causally sufficient conditions. One of these theories, universal determinism, associated with much science and philosophy, concerns the class of all events without exception. Another theory concerns physical events. Determinism in a third and important sense is human determinism. It is the theory that our choices and the many other antecedents of our actions, and the actions themselves, are effects of certain causal sequences. Lesser theories, usually associated with Freud and given no philosophical attention to speak of, concern themselves with particular sorts of conscious or otherwise mental causes of choices and actions, notably early sexual desires.

There are various relations between these four determinisms, depending on how they are additionally characterized. The most important relation, perhaps, is that universal determinism entails human determinism. That is not to say, however, that human determinism cannot be asserted, supported, or proved independently of universal determinism.

It is explicit or implicit in any of the above theories that the events in question are effects as more or less standardly conceived. An effect is an event such that an identical event follows every counterpart of the causal circumstance in question, or an event such that because the circumstance occurred, the event was in a stronger sense necessitated or had to happen (Sosa and Tooley 1993). A theory of our choices and actions, in contrast, that has to do with effects so-called—say, for example, effects conceived as events preceded by merely necessary conditions, or events merely made probable by antecedents—would not ordinarily be taken as a determinism. Indeed, weaker ideas of effects have often enough been introduced by philosophers precisely in order to avoid something else explicit or implicit in determinisms—that they may be inconsistent with or pose a challenge to beliefs in human freedom.

HUMAN DETERMINISM

This entry's concern will be with human determinism. It involves three large problems or enterprises.

The first is the formulation of a conceptually adequate theory. Human determinism has traditionally been thought about without reference to the philosophy of mind. Still, an adequate treatment of it must rest on a theory of the mind that is conceptually adequate: clear, consistent, and something like complete. Also, it must surely be that the theory of the mind, perhaps in what it rejects, say a puzzling power of originating choices, should be consonant with the philosophy of mind generally (Priest 1991, Heil 1998, Lowe 2000, Crane 2001).

The second problem with human determinism is its truth, whether or not this is considered in relation to universal determinism. The third problem is what can be called the human consequences for our existence of a human determinism. Is there in fact the consequence that we are not free? The philosophy of determinism and freedom, except in the philosophy of science and philosophical ruminations by scientists, has mainly concerned itself with this problem of consequences.

If these three problems are not the only ones that have been raised about determinism and freedom (Adler 1958), they have become the main ones (Kane 2002; Campbell, O'Rourke, and Shier 2004; Clarke 1995).

The formulation of a conceptually adequate theory is simple in terms of a truly physicalist or materialist philosophy of mind—one that takes conscious or mental events to have only or nothing but physical properties, however additionally conceived. In this case, human determinism becomes part of physical determinism. However, relatively few philosophies of mind are truly physicalist. Anomalous Monism, to mention one, is fairly typical in denying "nothing-but materialism" (Davidson 1980).

All other determinist theories face considerable problems of formulation. They encounter the problem of actually characterizing their primary subject matter—conscious or mental events. There is also the problem of the psychoneural relation, traditionally called the mind-body problem. If mental events are taken not to be in space, how can they be lawlike correlates or effects or causes? Further difficulties include the avoidance of epiphenomenalism, the nineteenth-century doctrine that actually makes conscious antecedents no part of the causation or explanation of our actions.

It is my view, seemingly now shared with most philosophers of determinism and freedom in the early twenty-first century, that despite these difficulties a conceptually adequate theory of human determinism can be formulated. This used to be doubted (Austin 1961, P. F. Strawson 1968).

Is any theory of human determinism true? A conceptually adequate theory has the support of much ordinary rationality, philosophy, and much science. It is notable that the ordinary philosophy of mind has no indeterminism in it. This most flourishing part of philosophy, much of it concerned with exactly the explanation of behavior, contains nothing at all of origination, an uncaused or uncausing initiation of choices and actions. Contemporary neuroscience, as distinct from philosophizing by retired neuroscientists and the like, plainly proceeds on the assumption of a human determinism. A reading of any of the main textbooks of neuroscience confirms this (Kandel et al. 1991) It is worth remarking, about what was called ordinary rationality, that in the end, which may be a long way down the line, it sits in judgment on science itself. That is to say, first of all, that inconsistency is not an option.

DENIALS OF HUMAN DETERMINISM

Despite these considerations, many or most of us do not take human determinism to be true. We deny or more likely doubt it. There may be an explanation of this, as distinct from a ground or justification, in our culture, at any rate European and North American culture.

One familiar ground used for this denial or doubt has been interpretations of quantum theory—applications to the world of the formalism or mathematics in which this part of physics can be said actually to consist. According to these interpretations, there are things at a microlevel of reality that are not effects. These things, well below the level of neural events in the brain, the events of ordinary neuroscience, are taken as made probable by antecedents but not necessitated by them. They are not chance events in the sense of being events of which it is true in advance that they are as likely not to occur as to occur. However, each one is certainly a chance event in that its actual occurrence or existence, no matter the antecedent probability, is such that there exists no causal explanation to be found for it. This is a matter of what is in the world, not our capabilities of knowing it.

Perhaps there is no strong consensus within science as to the truth of such indeterminist interpretations of quantum theory, despite an inclination in that direction. Something of the same sort may be true within physics itself. It is notable that outstanding treatments of the question in the philosophy of science may be agnostic (Earman 1986, 2004).

Opposition to indeterminism, some of it by philosophers, is strengthened by the fact, too often glossed over, that no satisfactory interpretation of quantum theory's application to reality has ever been achieved, although the theory is now getting on for a century old. It is possible to try to explain an ascendancy of an indeterminist understanding of quantum theory, say among other philosophers who would not tolerate contradiction, obscurity, and mystery elsewhere, by the fact of a cultural and institutional ascendancy of science in general and physics in particular. It is unclear to me why indeterminist interpretations have persisted within physics in the absence of any direct and univocal experimental evidence (Bohm and Hiley 1993, van Frassen 1991, Bub 1997).

One opposition to the idea that indeterminist interpretations of quantum theory prove or indicate the falsehood of determinism has to do with the supposedly undetermined things. Are they in fact events, which is to say things that happen; perhaps understood as ordinary things having properties at or for a time (Kim 1973)? Determinism has no concern with anything other than events. Numbers or propositions or other abstract objects, for example, are not part of its subject matter of effects. It does not say five is an effect. A reading of accounts of quantum theory quickly establishes that it is not clear that the things denied to be effects, about which there is real and wide disagreement, are indeed things asserted to be effects by a determinism. Some of these have been probabilities, features of a calculation, and waves in abstract mathematical space.

There is another uncertainty about any undetermined microevents, assuming such real events to exist. What is their relation to macroevents, and in particular to the neural events ordinarily taken to be in some intimate connection with such conscious or mental events as choices? Does the microdeterminism issue in macrodeterminism? Does it "translate up"? Or does the microdeterminism, instead, "cancel out" (Weatherford 1982)?

It is difficult indeed to resist the proposition that there is no indication at all of macroindeterminism in the physical world. Taken together with the previous uncertainty about amplification, this appears to issue in a kind of dilemma. Either microindeterminism if it exists does not translate up, in which case it does not matter to the problem with which we are concerned—or, because it would translate up if it existed, and there is no macrodeterminism, it follows that microindeterminism does not exist.

Answers or attitudes with respect to the question of the truth of a determinism do indeed affect responses to the third problem, that of the consequences of human determinism. Someone inclined to the truth of determinism may then be inclined, partly as a result of the further inclination that we have some freedom or others, to the response that we must have a freedom that goes with determinism. Still, the problem of the human consequences of determinism can be considered on its own, as usually it has been by philosophers.

Traditionally those consequences have been taken as having to do with freedom or free will, moral responsibility, and the justification of punishment. The central question is whether determinism is compatible or consistent with free choices and actions, with holding people responsible for and crediting them with responsibility for actions, and with imposing justified punishments on people and rewarding them. Compatibilists, who can be traced back at least to the seventeenth century (Hobbes 1839), answer yes. Incompatibilists, with Hobbes's great adversary in their history, answer no (Bramhall 1844).

The stock in trade of compatibilists has been the conception of freedom as voluntariness. That, in a rudimentary account, is the conception of a free and responsible action as in accordance with the desire of the person in question rather than against his or her desire. It is the conception, they say, that issues in the seemingly indubitable judgment that a man chained to the wall is not free, and that a woman whose life is under real and immediate threat by someone with a gun is not free.

The stock in trade of incompatibilists has been the idea of freedom as origination. This, in a rudimentary account, is the conception of a free action as one that the person was not caused to perform, but which was up to the person or in his or her control. This is the conception, incompatibilists say, that is familiar to all of us in that most common thing in our lives: holding people responsible for things. We hold people responsible only, as we say, when they are not literally caused to do what they do, but have a choice. We take a man to have been free exactly when he could have done otherwise than he did.

DEALING WITH OBJECTIONS TO HUMAN DETERMINISM

The rudimentary conception of freedom as voluntariness, as well expressed as the absence of ordinary constraint or compulsion, has been enriched in order to deal with objections. One objection was that people in the grip of an addiction are not acting against their own desire for heroin, but nonetheless are not free. A response in defense of compatibilism has been that voluntariness consists in someone's acting according to a desire that they desire to have. There is the possibility, indeed, of thinking of a hierarchy of desires (Frankfurt 1971).

Other objections, or perhaps the reaction that both the rudimentary and the amended ideas of voluntariness do not do justice to the fullness of our reactions to people in their actions, may call up other developments. A free choice or action, it may be said, is not only in accordance with the desired desire of the agent rather than against it, but grows out of the personality, character, history, and indeed the very being of the person. Who can object, compatibilists ask, to the idea that such a choice or action, so autonomous, is what we take to be a free and responsible one?

The conception of freedom as origination has also been given much attention, again in response to objections, usually about obscurity. It has long been insisted that an originated decision, although not a standard effect, is not merely that. It is not merely a chance or random event. Hobbes's adversary Bramhall in the seventeenth century explained originated choices and actions as owed to the elective power of the rational will. It has become common to try to explain such choices by assigning them to what is called agent causation as against standard causation (Chisholm 1976, O'Connor 1995). Agent causation, whatever else is said of it, does not give rise to effects that had to happen or were necessitated. Other attempts to further clarify origination are in terms of teleology, in particular that the occurrence of choices and actions are somehow explained by their goals (O'Connor 1995), and in terms of a mixture of determined and undetermined events (Kane 1985, 2002), and in terms of reasons rather than causes (Ginet 1990).

It is clear that a determinism can be true and there can still be voluntary choices and actions. There is full compatibility. There is nothing in a theory of determinism that rules out choices and actions being according to someone's desire. Determinism is evidently never the theory that all choices and actions are against the wills of the agents. Compatibilism, indeed, is best seen as based on the proposition that free choices and actions have certain causes, causes somehow internal to rather than external and somehow opposed to the agent.

It is equally clear that if a decent theory of determinism is true, there can be no originated choices and actions. There is clear incompatibility. An originated choice or action, by rudimentary definition, is an event that is in a standard sense uncaused. The question of whether determinism is compatible with freedom has

been the question of whether our freedom consists in voluntariness or origination, not the question of whether determinism is compatible with origination.

HUME, KANT, AND COMPATIBILISM

To come to the principal arguments of the two traditions of philosophers, Hume was typical of compatibilists in maintaining that anyone who actually thinks of what he or she means in speaking of a free and responsible action will immediately see that it is an unconstrained or uncoerced one—a voluntary one. What is needed is no more than some self-reflection, unconfused by religion or the like (Hume 1955).

Kant, although in fact not an incompatibilist, certainly not an ordinary incompatibilist, was as positive in declaring that to think of one's idea of a free and responsible action is not to think merely of one that was necessitated in a certain way. To go along with Hume and suppose otherwise, he said, is to engage in no more than a little quibbling with words (Kant 1949). With these philosophers, there was already a kind of stalemate about determinism and freedom.

Near the beginning of the twentieth century, it was taken as established, by some, that compatibilism was proved by a simple consideration. If a person acted freely on some occasion, it was true that the person could have acted otherwise. But, it was said, the latter means that the person would have acted differently if he or she had chosen differently, which is consistent with determinism (Moore 1912). By the mid-twentieth century, however, it became clear to some that "could have acted otherwise" is inconsistent with determinism (Austin 1961).

Subsequent twentieth- and indeed twenty-first-century compatibilists, undaunted by the failure of their predecessors to prove it, have somehow stuck to the conviction that our common idea of freedom, our common idea of what is necessary for moral responsibility and right punishment, is voluntariness (Ayer 1973, Magill 1997). One further contention is that the idea of origination, despite the seemingly clear rudimentary description of it, is actually incoherent, and so the field is left to the tolerably clear ideal of voluntariness (G. Strawson 1986).

Another compatibilist argument, widely discussed, begins from a thought experiment about moral responsibility (Frankfurt 1969). What it amounts to is the idea of a person subject to the control of a neuroscientist with some apparatus who will secure that the person will act in a certain way if it happens that the person is not on the way to doing so. Those are the causal facts. Suppose, however, that the person actually is on the way to and absolutely committed to doing A—wants it, wants to want it, and so on. It remains true, given the neuroscientist in the background, that he cannot do anything else. But it is clear, surely, that he is morally responsible for A. It follows, we are told, that freedom does not require being able to do otherwise than we do in a strong sense—it does not require origination and is not itself origination. Other recent compatibilist argumentation has been the elaboration of the idea of voluntariness by seeing its growth and extent in terms of evolution (Dennett 2003). Our human freedom is favorably contrasted with the lesser freedom of other animals.

Twentieth-century incompatibilists gave much attention to an argument well-developed from its beginning in Kant's philosophy (van Inwagen 1986). Here we have it that a free action is one that is up to us. Suppose now that an action is subject to determinism—the effect of a causal sequence, a series of lawlike connections leading back to some causal circumstance prior to the birth of the agent. Can such an action be up to us? The answer given is that it can only be up to us if the lawlike connections and the first causal circumstance are within our control—which definitely they are not. Hence free actions cannot be effects of certain causal sequences but must be originated.

Given the unbroken history of the philosophical debate on determinism and freedom until recently, must there be a presumption that either compatibilism or incompatibilism is true? Can that respectful attitude survive certain troublesome questions and alternatives?

If you reflect on the compatibilist case of the desiring and committed agent but with the neuroscientist around the corner, or indeed on any of many cases, say the simple one of the man chained to the wall, one thing you must be persuaded of is that there certainly is *an* idea of freedom—voluntariness. Quite as clearly, if you reflect on the incompatibilist case of the agent about whom it is supposed that a causal circumstance before his birth was not up to him, one thing you must allow is that there is *an* idea of freedom such that he does not have it—origination.

Does it follow from either speculation, however, that each of us has *only* the idea of freedom in question? That we all have and use only that single settled idea? That is exactly what is intended by each speculation, exactly what it is supposed to prove.

To ask the question, perhaps, is to become at least worried. Recall the first agent doing what he wants and

responsible although in the toils of the neuroscientist. Is it just the philosophers who can readily think that there still *is* a sense in which he is not free—he cannot do otherwise in a sense of the words inconsistent with determinism? And is it just the philosophers who can readily think of the second agent, who indeed does not have a causal circumstance in the distant pas in his control, that there still *is* a clear sense in which his action may indeed be in his control? It may be wholly in accord with his desires and character and his whole existence, not pushed on him by anyone else or anything else or any conflict within him. Do we not have and use both conceptions?

What may lead someone to assent to one of the two speculations, and to either compatibilism or incompatibilism, is of course the proposition that freedom either is or is not compatible with determinism. That is a logical or necessary truth, is it not? Well, it is a truth only on a certain ordinary assumption or presupposition. The presupposition of course is that freedom is one thing, that we in general have only one idea of freedom. Evidently this presupposition needs thinking about, and it has been thought about in additional ways.

DEFENSES OF COMPATIBILISM

An original defense of compatibilism prepared the way by making more explicit the fact that determinism is not best seen as raising a question of consistency or inconsistency, but rather as affecting attitudes directed at certain facts or propositions having to do with moral responsibility—and also such personal and nonmoral attitudes as gratitude and resentment (P. F. Strawson 1968). Subsequently it was proposed that determinism affects more attitudes than these, including the important attitude to the future that is hope and the important attitude to inquiry and conclusions that is confidence.

It was argued that it is plain that we are all subject to two kinds of hope, one for an open future where all has not been fixed by the past, one for a future in which we get what we want, maybe a whole kind of life. To this attitudinal argument, a behavioral one was subsequently added. What we secure by enacting and benefiting from bills of rights and political liberty is evidently an absence of compulsion. What we punish for in part is an action of which we take it that it could have been otherwise despite the past, and we have the same thought in various personal relations (Honderich 1988, 1993).

Such considerations also bear nearly as sharply on weaker positions to which compatibilists and incompatibilists may be retreating. These positions are that voluntariness is our more important conception of freedom (Dennett 1984, 2003), the freedom more worth having, or that origination has these recommendations (Kane 1985, 2002).

THE WIDER DEBATE

The ensuing wider debate—wider than compatibilism and incompatibilism—has included the idea that our being free requires origination but our being responsible requires only voluntariness (Fischer 1994). A different inquiry into what is called autonomy also accepts that we do not have to choose between compatibilism and incompatibilism (Mele 1995). It has been argued, against compatibilism's way of saving our responsibility from determinism, that we must give up our real idea of responsibility (Pereboom 2001). There has been the more radical contention that ascribing freedom and responsibility to people is a matter of attitudes that do not depend on objective facts or propositions at all (Double 1991, 1996).

Against another thought, that of giving up the set of attitudes inconsistent with determinism and taking satisfaction in the set of consistent ones, it has been argued that despite the truth of determinism we must maintain the illusion that we have the power of origination (Smilansky 2000). The thought of giving up the inconsistent attitudes and being satisfied by the others has also been followed by another radical idea. It is that roughly our attitudes to ourselves previously associated with origination can survive acceptance of determinism, and so must be owed to something else entirely different. This could be the nature of our consciousness, or the explanatory nature of certain causal lines of events within sequences of causal circumstances (Honderich 2002).

It is too early to say, but it may be that a consensus is emerging that determinism and freedom can no longer be the protracted and tired battle between compatibilism and incompatibilism. It is not possible to conjecture about the outcome of an alternative discussion.

See also Action; Causation: Metaphysical Issues; Determinism, A Historical Survey; Freud, Sigmund; Hobbes, Thomas; Hume, David; Kant, Immanuel; Philosophy of Mind; Quantum Mechanics; Responsibility, Moral and Legal; Strawson, Peter Frederick.

Bibliography

Adler, M. J. *The Idea of Freedom*. New York, 1958; Westport, CT: Greenwood, 1973.

Albert, D. *Quantum Mechanics and Experience*. Cambridge, MA: Harvard University Press, 1992.

Austin, J. L. "Ifs and Cans." In his *Philosophical Papers*, edited by J. O. Urmson and G. Warnock. Oxford: Clarendon, 1961.
Ayer, A. J. *The Central Questions of Philosophy*. London, 1954; New York: Morrow, 1975.
Bohm, D., and B. Hiley. *The Undivided Universe*. London: Routledge, 1993.
Bramhall, John. *The Works of John Bramhall*. Oxford: John Henry Parker, 1844. Originally published in Dublin, 1675.
Bub, J. *Interpreting the Quantum World*. Cambridge, U.K., and New York: Cambridge University Press, 1997.
Campbell, Joseph Keim, Michael O'Rourke, and David Shier. *Freedom and Determinism*. Cambridge, MA: MIT Press, 2004.
Chisholm, R. M. "The Agent as Cause." In *Action Theory*, edited by Myles Brand and D. Walton. Dordrecht, Netherlands: Reidel, 1976.
Clarke, R. "Freedom and Determinism." In *Philosophical Books*. Oxford: Basil Blackwell, 1995.
Crane, Tim. *Elements of Mind: An Introduction to the Philosophy of Mind*. Oxford: Oxford University Press, 2001.
Davidson, Donald. "Mental Events." In his *Essays on Actions and Events*. Oxford: Clarendon, 1980.
Dennett, Daniel. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press, 1984.
Dennett, Daniel. *Freedom Evolves*. New York: Viking, 2003.
Double, R. *The Non-Reality of Free Will*. Oxford: Oxford University Press, 1991.
Double, R. *Metaphilosophy and Free Will*. Oxford: Oxford University Press, 1996.
Earman, John. *A Primer on Determinism*. Dordrecht, Netherlands: D. Reidel, 1986.
Earman, John. "Determinism: What We Have Learned and What We Still Don't Know." In *Freedom and Determinism*, edited by J. Campbell, M. O'Rourke, and D. Shier. Cambridge, MA: MIT Press, 2004.
Fischer, J. M. *The Metaphysics of Free Will*. Cambridge, MA: Blackwell, 1994.
Frankfurt, H. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66 (1969): 829–839.
Frankfurt, H. "Freedom of the Will and the Concept of a Person." In *Free Will*, edited by G. Watson, 81–95. Oxford: Oxford University Press, 1982.
Ginet, C. *On Action*. Cambridge, U.K.: Cambridge University Press, 1990.
Heil, John. *Philosophy of Mind: A Contemporary Introduction*. Oxford and New York: Routledge, 1998.
Hobbes, Thomas. *The English Works of Thomas Hobbes*. Vol. 5, edited by W. Molesworth. London: London, J. Bohn, 1839. Originally published in 1650.
Honderich, T. *A Theory of Determinism: The Mind, Neuroscience, and Life-hopes*. Oxford: Clarendon, 1988.
Honderich, T. *How Free Are You?* Oxford: Oxford University Press, 1993; 2nd ed., 2002.
Hume, David. *An Enquiry concerning Human Understanding*. Edited by L. A. Selby-Bigge. Oxford, 1955.
Kandel, E. R. R., J. H. Schwartz, and T. J. Jessell. *Principles of Neural Science*. New York: Elsevier, 1991.
Kane, R. *Free Will and Values*. Albany: SUNY Press, 1985.
Kane, R. *The Oxford Handbook of Free Will*. Oxford and New York: Oxford University Press, 2002.
Kant, Immanuel. *Critique of Practical Reason*. Translated by L. W. Beck. Chicago: University of Chicago Press, 1949. Originally published in 1788.
Kim, J. "Causation, Nomic Subsumption, and the Concept of Event." *Journal of Philosophy* 70 (1973): 217–236.
Lowe, E. J. *An Introduction to the Philosophy of Mind*. Cambridge, U.K.: Cambridge University Press, 2000.
Magill, Kevin *Freedom and Experience*. London and New York: St. Martin's Press, 1997.
Mele, Alfred. *Autonomous Agents*. Oxford: Oxford University Press, 1995.
Moore, G. E. *Ethics*. Oxford: H. Holt, 1912.
O'Connor, T. *Agents, Causes and Events*. Oxford: Oxford University Press, 1995.
Pereboom, Derk. *Living without Free Will*. Cambridge, U.K.: Cambridge University Press, 2001.
Priest, Stephen. *Theories of the Mind*. London: Penguin, 1991.
Smilansky, Saul. *Free Will and Illusion*. Oxford: Clarendon, 2000.
Sosa, E., and M. Tooley, eds. *Causation*. Oxford: Oxford University Press, 1993.
Strawson, G. *Freedom and Belief*. Oxford: Clarendon, 1986.
Strawson, P. F. "Freedom and Resentment." In his *Studies in the Philosophy of Thought and Action*. Oxford: Oxford University Press, 1968.
van Frassen, Bas. *Quantum Mechanics: An Empiricist View*. Oxford: Clarendon, 1991.
van Inwagen, P. *An Essay on Free Will*. Oxford: Clarendon, 1986.
Weatherford, Roy. *The Implications of Determinism*. London: Routledge, 1982.

Ted Honderich (1996, 2005)

DETERMINISM AND INDETERMINISM

Determinism is a rich and varied concept. At an abstract level of analysis, Jordan Howard Sobel (1998) identifies at least ninety varieties of what determinism could be like. When it comes to thinking about what deterministic laws and theories in physical sciences might be like, the situation is much clearer. There is a criterion by which to judge whether a law—expressed as some form of equation—is deterministic. A theory would then be deterministic just in case all its laws taken as a whole were deterministic. In contrast, if a law fails this criterion, then it is indeterministic and any theory whose laws taken as a whole fail this criterion must also be indeterministic. Although it is widely believed that classical physics is deterministic and quantum mechanics is indeterministic, application of this criterion yields some surprises for these standard judgments.

FRAMEWORK FOR PHYSICAL THEORIES

Laws and theories in physics are formulated in terms of dynamical or evolution equations. These equations are taken to describe the change in time of the relevant variables characterizing the system in question. Additionally, a complete specification of the initial state referred to as the initial conditions for the system and/or a characterization of the boundaries for the system known as the boundary conditions must also be given. A state is taken to be a description of the values of the variables characterizing the system at some time *t*. As a simple example of a classical model, consider a cannon firing a ball. The initial conditions would be the initial position and velocity of the ball as it left the mouth of the cannon. The evolution equation plus these initial conditions would then describe the path of the ball.

Much of the analysis of physical systems takes place in what is called state space, an abstract mathematical space composed of the variables required to fully specify the state of a system. Each point in this space then represents a possible state of the system at a particular time *t* through the values these variables take on at *t*. For example, in many typical dynamical models—constructed to satisfy the laws of a given theory—the position and momentum serve as the coordinates, so the model can be studied in state space by following its trajectory from the initial state (q_o, p_o) to some final state (q_f, p_f). The evolution equations govern the path—the history of state transitions—of the system in state space.

However, note that there are important assumptions being made here. Namely, that a state of a system is characterized by the values of the crucial variables and that a physical state corresponds to a point in state space through these values. This cluster of assumptions can be called the faithful model assumption. This assumption allows one to develop mathematical models for the evolution of these points in state space and such models are taken to represent (perhaps through a complicated relation) the physical systems of interest. In other words, one assumes that one's mathematical models are faithful representations of physical systems and that the state space is a faithful representation of the space of physically genuine possibilities for the system in question. Hence, one has the connection between physical systems and their laws and models, provided the latter are faithful. It then remains to determine whether these laws and models are deterministic or not.

LAPLACEAN DETERMINISM

Clocks, cannon balls fired from cannons, and the solar system are taken to be paradigm examples of deterministic systems in classical physics. In the practice of physics one is able to give a general and precise description of deterministic systems. For definiteness the focus here is on classical particle mechanics, the inspiration for Pierre Simon Laplace's famous description:

> We ought to regard the present state of the universe as the effect of its antecedent state and as the cause of the state that is to follow. An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in one single formula the motions of the largest bodies as well as the lightest atoms in the world … to it nothing would be uncertain, the future as well as the past would be present to its eyes. (Translation from Nagel 1961, pp. 281–282)

Given all the forces acting on the particles composing the universe along with their exact positions and momenta, then the future behavior of these particles is, in principle, completely determined.

Two historical remarks are in order here. First, Laplace's primary aim in this famous passage was to contrast the concepts of probability and certainty. Second, Gottfried Wilhelm Leibniz (1924, p. 129) articulated this same notion of inevitability in terms of particle dynamics long before Laplace. Nevertheless, it was the vision that Laplace articulated that has become a paradigm example for determinism in physical theories.

This vision may be articulated in the modern framework as follows. Suppose that the physical state of a system is characterized by the values of the positions and momenta of all the particles composing the system at some time *t*. Furthermore, suppose that a physical state corresponds to a point in state space (invoking the faithful model assumption). One can then develop deterministic mathematical models for the evolution of these points in state space. Some have thought that the key feature characterizing this determinism was that given a specification of the initial state of a system and the evolution equations governing its states, in principle it should be possible to predict the behavior of the system for any time (recall Laplace's contrast between certainty and probability). Although prima facie plausible, such a condition is neither necessary nor sufficient for a determinis-

tic law because the relationship of predictability to determinism is far too weak and subtle.

Rather, the core feature of determinism is the following condition: "Unique evolution: A given state is always followed (and preceded) by the same history of state transitions." This condition expresses the Laplacean belief that systems described by classical particle mechanics will repeat their behaviors exactly if the same initial and boundary conditions are specified. For example, the equations of motion for a frictionless pendulum will produce the same solution for the motion as long as the same initial velocity and initial position are chosen. Roughly speaking, the idea is that every time one returns the mathematical model to the same initial state (or any state in the history of state transitions), it will undergo the same history of transitions from state to state and likewise for the target system. In other words, the evolution will be unique given a specification of initial and boundary conditions. Note that as formulated, unique evolution expresses state transitions in both directions (future and past). It can easily be recast to allow for unidirectional state transitions (future only or past only) if desired.

UNIQUE EVOLUTION

Unique evolution is the core of the Laplacean vision for determinism (it lies at the core of Leibniz's statement as well). Although a strong requirement, it is important if determinism is to be meaningfully applied to laws and theories. Imagine a typical physical system *s* as a film. Satisfying unique evolution means that if the film is started over and over at the same frame (returning the system to the same initial state), then *s* will repeat every detail of its total history over and over again and identical copies of the film would produce the same sequence of pictures. So if one always starts *Jurassic Park* at the beginning frame, it plays the same. The tyrannosaurus as antihero always saves the day. No new frames are added to the movie. Furthermore, if one were to start with a different frame, say a frame at the middle of the movie, there is still a unique sequence of frames.

By way of contrast, suppose that returning *s* to the same initial state produced a different sequence of state transitions on some of the runs. Consider a system *s* to be like a device that spontaneously generates a different sequence of pictures on some occasions when starting from the same initial picture. Imagine further that such a system has the property that simply by choosing to start with any picture normally appearing in the sequence, sometimes the chosen picture is not followed by the usual sequence of pictures. Or imagine that some pictures often do not appear in the sequence, or that new ones are added from time to time. Such a system would fail to satisfy unique evolution and would not qualify as deterministic.

More formally, one can define unique evolution in the following way. Let S stand for the collection of all systems sharing the same set L of physical laws and suppose that P is the set of relevant physical properties for specifying the time evolution of a system described by L: A system $s \in S$ exhibits unique evolution if and only if every system $s' \in S$ isomorphic to s with respect to P undergoes the same evolution as s.

TWO CONSTRUALS OF UNIQUE EVOLUTION

Abstracting from the context of physical theories for the moment, unique evolution can be given two construals. The first construal is as a statement of causal determinism, that every event is causally determined by an event taking place at some antecedent time or times. This reading of unique evolution fits nicely with how a number of philosophers conceive of metaphysical, physical, and psychological determinism as theses about the determination of events in causal chains, where there is a flow from cause to effect that may be continuous or have gaps. The second construal of unique evolution is as a statement of difference determinism characterized by William James as "[t]he whole is in each and every part, and welds it with the rest into an absolute unity, an iron block, in which there can be no equivocation or shadow of turning" (1956, p. 150). This reading of unique evolution maintains that a difference at any time requires a difference at every time.

These two construals of unique evolution are different. For example, consider a fast-starting series of causally linked states (Sobel 1998) where every state in the series has an earlier determining cause, but the series itself has no antecedent deterministic cause (its beginning—the first state—is undetermined by prior events or may have a probabilistic cause) and no state in the series occurs before a specified time. The principle that every event has an earlier cause would fail for a fast-starting series as a whole though it would hold for the events within such a series. This would be an example where causal determinism failed, but where difference determinism would still hold.

However, the causal construal of unique evolution is unsatisfactory. Concepts like *event* or *causation* are vague and controversial. One might suggest explicating causal determinism in terms of the laws L and properties P, but concepts like *event* and *cause* are not used in most physi-

cal theories (at least not univocally). In contrast, unique evolution fits the idea of difference determinism: any difference between s and s' is reflected by different histories of state transitions. This latter construal of unique evolution only requires the normal machinery of the theoretical framework sketched earlier to cash out these differences and so avoids controversies associated with causal determinism.

DETERMINISM IN CLASSICAL MECHANICS

Most philosophers take classical mechanics to be the archetype of a deterministic theory. Prima facie Newton's laws satisfy unique evolution. After all, these are ordinary differential equations and one has uniqueness and existence proofs for them. Furthermore, there is at least some empirical evidence that macroscopic objects behave approximately as these laws describe. Still, there are some surprises and controversy regarding the judgment that classical mechanics is a deterministic theory.

For example, as Keith Hutchinson (1993) notes, if the force function varies as the square root of the velocity, then a specification of the initial position and velocity of a particle does not fix a unique evolution of the particle in state space (indeed, the particle can sit stationary for an arbitrary length of time and then spontaneously begin to move). Hence, such a force law is not deterministic. There are a number of such force functions consistent with Newton's laws, but that fail to satisfy unique evolution. Therefore, the judgment that classical mechanics is a deterministic theory is false.

NEWTONIAN GRAVITY. One might think that the set of force functions leading to violations of unique evolution represents an unrealistic set so that all force laws of classical mechanics really are deterministic. However, worries for determinism await one even in the case of point-particles interacting under Isaac Newton's force of gravity, the paradigm case of determinism that Laplace had in mind.

In 1897 the French mathematician Paul Painlevé conjectured that a system of point-particles interacting only under Newton's force of gravity could all accelerate to spatial infinity within a finite time interval. (The source of the energy needed for this acceleration is the infinite potential well associated with the inverse-square law of gravitation.) If particles could disappear to "spatial infinity," then unique evolution would break down because solutions to the equations of motion no longer would be guaranteed to exist. Painlevé's conjecture was proven by Zhihong Xia (1992) for a system of five point-masses.

Though provocative, these results are not without controversy. For example, there are two interesting possibilities for interpreting the status of these particles that have flown off to spatial infinity. On the one hand, one could say the particles have left the universe and now have some indefinite properties. On the other hand, one could say that the particles no longer exist. Newton's mechanics is silent on this interpretive question. Furthermore, are events such as leaving the universe to be taken as predictions of Newton's gravitational theory of point-particles, or as indications that the theory is breaking down because particle position becomes undefined? Perhaps such behavior is an artifact of a spatially infinite universe. If the universe is finite, particle positions are always bounded and such violations of unique evolution are not possible.

DIAGNOSIS. Other failures of unique evolution in classical mechanics can be found in John Earman's (1986) survey. What is one to say, then, about the uniqueness and existence theorems for the equations of motion, the theorems that appear so suggestive of unique evolution? The root problem of these failures to satisfy unique evolution can be traced back to the fact that one's mathematical theorems only guarantee existence and uniqueness locally in time. This means that the equations of motion only have unique solutions for some interval of time. This interval might be short and, as time goes on, the interval of time for which such solutions exist might get shorter or even shrink to zero in such a way that after some period solutions cease to exist. So determinism might hold locally, but this does not guarantee determinism must hold globally.

DETERMINISM IN SPECIAL AND GENERAL RELATIVITY

Special relativity provides a much more hospitable environment for determinism. This is primarily due to two features of the theory: (1) no process or signal can travel faster than the speed of light, and (2) the space-time structure is static. The first feature rules out unbounded-velocity systems, while the second guarantees there are no singularities in space-time. Given these two features, global existence and uniqueness theorems can be proven for cases like source-free electromagnetic fields so that unique evolution is not violated when appropriate initial data are specified on a space-like hypersurface. Unfortunately, when electromagnetic sources or gravitationally

interacting particles are added to the picture, the status of unique evolution becomes much less clear.

In contrast, general relativity presents problems for guaranteeing unique evolution. For example, there are space-times for which there are no appropriate specifications of initial data on space-like hypersurfaces yielding global existence and uniqueness theorems. In such space-times, unique evolution is easily violated. Furthermore, problems for unique evolution arise from the possibility of naked singularities (singularities not hidden behind an event horizon). One way a singularity might form is from gravitational collapse. The usual model for such a process involves the formation of an event horizon (i.e., a black hole). Although a black hole has a singularity inside the event horizon, outside the horizon at least determinism is okay, provided the space-time supports appropriate specifications of initial data compatible with unique evolution. In contrast, a naked singularity has no event horizon. The problem here is that anything at all could pop out of a naked singularity, violating unique evolution. To date, no general, convincing forms of hypotheses ruling out such singularities have been proven (so-called cosmic censorship hypotheses).

DETERMINISM IN QUANTUM MECHANICS

In contrast to classical mechanics philosophers often take quantum mechanics to be an indeterministic theory. Nevertheless, so-called pilot-wave theories pioneered by Louis de Broglie and David Bohm are explicitly deterministic while still agreeing with experiments. Roughly speaking, this family of theories treats a quantum system as consisting of both a wave and a particle. The wave evolves deterministically over time according to the Schrödinger equation and determines the motion of the particle. Hence, the particle's motion satisfies unique evolution. This is a perfectly coherent view of quantum mechanics and contrasts strongly with the more orthodox interpretation. The latter takes the wave to evolve deterministically according to Schrödinger's equation and treats particle-like phenomena indeterministically in a measurement process (such processes typically violate unique evolution because the particle system can be in the same state before measurement, but still yield many different outcomes after measurement). Pilot-wave theories show that quantum mechanics need not be indeterministic.

DETERMINISTIC CHAOS

Some philosophers have thought that the phenomenon of deterministic chaos—the extreme sensitivity of a variety of classical mechanics systems such that roughly even the smallest change in initial conditions can lead to vastly different evolutions in state space—might actually show that classical mechanics is not deterministic. However, there is no real challenge to unique evolution here as each history of state transitions in state space is still unique to each slightly different initial condition.

Of course, classical chaotic systems are typically considered as if there is no such thing as quantum mechanics. But suppose one considers a combined system such that quantum mechanics is the source of the small changes in initial conditions for one's classical chaotic system? Would such a system fail to satisfy unique evolution? The worry here is that, since there is no known lower limit to the sensitivity of classical chaotic systems, nothing can prevent the possibility of such systems amplifying a slight change in initial conditions due to a quantum event so that the evolution of the classical chaotic system is dramatically different than if the quantum event had not taken place. Indeed, some philosophers argue that unique evolution must fail in such circumstances.

However, such sensitivity arguments depend crucially on how quantum mechanics itself and measurements are interpreted as well as on where the cut is made distinguishing between what is observed and what is doing the observing (e.g., is the classical chaotic system serving as the measuring device for the quantum change in initial conditions?). Although considered abstractly, sensitivity arguments do correctly lead to the conclusion that quantum effects can be amplified by classical chaotic systems; they do not automatically render one's classical plus quantum system indeterministic. Furthermore, applying such arguments to concrete physical systems shows that the amplification process may be severely constrained. For example, investigating the role of quantum effects in the process of chaos in the friction of sliding surfaces indicates that quantum effects might be amplified by chaos to produce a difference in macroscopic behavior only if the fluctuations are large enough to break molecular bonds and are amplified quickly enough.

BROADER IMPLICATIONS

Finally, what of broader implications of determinism and indeterminism in physical theories? Debates about free will and determinism are one place where the considerations in this entry might be relevant. One of the most discussed topics in this regard is the consequence argument, which may be put informally as follows: If determinism is true, then our acts are consequences of laws and events in

the remote past. But what went on before we were born is not up to us and neither are the laws up to us. Therefore, the consequences of these laws and events—including our present acts—are not up to us. Whether or not the relevant laws satisfy unique evolution is one factor in the evaluation of this argument.

What of broader philosophical thinking about psychological determinism or the thesis that the universe is deterministic? For the former, it looks difficult to make any connection at all. One simply does not have any theories in the behavioral sciences that are amenable to analysis under the criterion of unique evolution. Indeed, attempts to apply the criterion in psychology do not lead to clarification of the crucial issues (Bishop 2002).

With regards to the universe, it has been common practice since the seventeenth century for philosophers to look to their best scientific theories as guides to the truth of determinism. As one has seen, the current best theories in physics are remarkably unclear about the truth of determinism in the physical sciences, so the current guides do not appear to be so helpful. Even if the best theories were clear on the matter of determinism in their province, there is a further problem awaiting their application to metaphysical questions about the universe as a whole. Recall the crucial faithful model assumption. In many contexts this assumption is fairly unproblematic. However, if the system in question is nonlinear—that is to say, has the property that a small change in the state or conditions of the system is not guaranteed to result in a small change in the system's behavior—this assumption faces serious difficulties (indeed, a strongly idealized version of the assumption, the perfect model scenario, is needed but also runs into difficulties regarding drawing conclusions about the systems one is modeling). Since the universe is populated with such systems—indeed, it is likely to be nonlinear itself—one's purchase on applying the best laws and theories to such systems or the universe as a whole to answer the large metaphysical question about determinism is problematic.

See also Determinism, A Historical Survey; Determinism in History; Philosophy of Physics; Quantum Mechanics.

Bibliography

Relevant Historical Material on Determinism:

James, William. "The Dilemma of Determinism." In *The Will to Believe and Other Essays in Popular Philosophy, and Human Immortality*. New York: Dover Publications, 1956.

Laplace, Pierre Simon de. *A Philosophical Essay on Probabilities*. Translated by Frederick Wilson Truscott and Frederick Lincoln Emory. New York: Dover Publications, 1951.

Leibniz, Gottfried Wilhelm. "Von dem Verhängnisse." In *Hauptschriften zur Grundlegung der Philosophie*. Vol. 2, edited by Ernst Cassirer and Artur Buchenau. Leipzig, Germany: Meiner, 1924.

Nagel, Ernst. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace, and World, 1961.

Sobel, Jordan Howard. *Puzzles for the Will: Fatalism, Newcomb and Samarra, Determinism and Omniscience*. Toronto: University of Toronto Press, 1998.

Laplace's vision expressed in the modern framework of physical theories, as well as discussions of chaos, prediction, and determinism, may be found in:

Bishop, Robert C. "On Separating Predictability and Determinism." *Erkenntnis* 58 (2) (2003): 169–188.

Bishop, Robert C., and Frederick M. Kronz. "Is Chaos Indeterministic?" In *Language, Quantum, Music: Selected Contributed Papers of the Tenth International Congress of Logic, Methodology, and Philosophy of Science, Florence, August 1995*, edited by Maria Luisa Dalla Chiara, Roberto Giuntini, and Federico Laudisa. Boston: Kluwer Academic, 1999.

Hobbs, Jesse. "Chaos and Indeterminism." *Canadian Journal of Philosophy* 21 (1991): 141–164.

Stone, M. A. "Chaos, Prediction, and Laplacean Determinism." *American Philosophical Quarterly* 26 (1989): 123–131.

There are a number of able discussions of problems for determinism in physical theories. The following all discuss classical physics; see Earman (1986, 2004) for discussions of determinism in relativistic physics:

Earman, John. "Determinism: What We Have Learned and What We Still Don't Know." In *Freedom and Determinism*, edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier, 21–46. Cambridge, MA: MIT Press, 2004.

Earman, John. *A Primer on Determinism*. Dordrecht, Netherlands: D. Reidel, 1986.

Hutchinson, Keith. "Is Classical Mechanics Really Time-Reversible and Deterministic?" *British Journal for the Philosophy of Science* 44 (1993): 307–323.

Xia, Zhihong "The Existence of Noncollision Singularities in Newtonian Systems." *Annals of Mathematics* 135 (3) (1992): 411–468.

Uniqueness and existence proofs for differential equations are discussed by:

Arnold, V. I. *Geometrical Methods in the Theory of Ordinary Differential Equations*. 2nd ed. Translated by Joseph Szücs; edited by Mark Levi. New York: Springer-Verlag, 1988.

For a discussion of deterministic versions of quantum mechanics, see:

Bohm, David. *Causality and Chance in Modern Physics*. London: Routledge and Paul, 1957.

Cushing, James T. *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*. Chicago: University of Chicago Press, 1994.

Possible consequences of determinism for free will in terms of the consequence argument may be found in:

Kane, Robert, ed. *The Oxford Handbook of Free Will*. New York: Oxford University Press, 2001.

Van Inwagen, Peter. *An Essay on Free Will*. Oxford, U.K.: Clarendon Press, 1983.

For a discussion of difficulties in applying determinism as unique evolution to psychology, see:

Bishop, Robert C. "Deterministic and Indeterministic Descriptions." In *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*, edited by Harald Atmanspacher and Robert C. Bishop. Thorverton, U.K.: Imprint Academic, 2002.

Elements of the faithful model assumption have received some scrutiny in recent physics literature. In particular, there is evidence that perfect models are not guaranteed to describe system behavior in nonlinear contexts:

Judd, Kevin, and Leonard A. Smith. "Indistinguishable States I: Perfect Model Scenario." *Physica D* 151 (2001): 125–141.

Judd, Kevin, and Leonard A. Smith. "Indistinguishable States II: Imperfect Model Scenarios." *Physica D* 196 (2004): 224–242.

Smith, Leonard A. "Disentangling Uncertainty and Error: On the Predictability of Nonlinear Systems." In *Nonlinear Dynamics and Statistics*, edited by Alistair I. Mees. Boston: Birkäuser, 2001.

Robert C. Bishop (2005)

DETERMINISM IN HISTORY

Philosophical reflection upon history has always been impressed by the limited extent to which individuals and groups seem to be able to mold events to their purposes. In the case of some events at least, there seems to be an inexorable necessity—an inevitability or unavoidability—about what happens. The "necessity" of historical events, however, has been asserted by historians and philosophers of history in at least three fundamentally different senses.

SENSES OF DETERMINISM

FATE AND PROVIDENCE. The first sense is the notion that events are "fated" to occur, a notion familiar to Greek as well as Oriental thought. The central concept is of an agency external to the historical process itself, sometimes, but not always, personified, determining events somewhat in the way a human agent may be said to determine, through his will, what happens in a process he monitors and manipulates. It is generally assumed, however, that the means by which fated events are brought about lie outside the mechanism of ordinary causal connection: they are "transcendent." This clears the way for a characteristic expression of fatalism—the assertion that what is fated will occur no matter what we do to try to prevent it. To many critics, such a claim has appeared unintelligible. For historical events are surely, in some sense at least, constituted by what we do. A revolution, for example, could hardly occur if nobody revolted. The fatalist claim thus looks self-contradictory. What fatalism really denies, however, is the preventive efficacy of anyone's actions prior to the fated event, a refinement that leaves the claim coherent, if unbelievable. Nor is the doctrine necessarily involved in the incoherence of representing prior actions as both within our power to have performed otherwise and, at the same time, fated in their turn. For fatalism, unlike some other forms of historical determinism, has generally been asserted selectively. It is the doctrine that certain things will necessarily come to pass, not that everything happens necessarily.

Many theological philosophies of history are fatalistic in the indicated sense because of the role they assign to the will of God in their accounts. Unlike most of their pagan predecessors, however, these accounts generally make some attempt to rationalize and even to moralize interventions hitherto conceived as arbitrary, and usually also as menacing. In this way a fatalistic conception of history becomes "providential." Theological interpretations, of course, leave little for philosophers to argue about; for the workings of Divine Providence can be discerned only through some extrarational insight or source of revelation. And as G. W. F. Hegel complained about providential theories generally, the overarching purpose or plan is usually conceded, even by those who claim insight into it, to be partly "concealed from our view." Some theological interpretations have tried to meet this sort of objection by identifying the workings of providence, tentatively at least, with certain standing conditions and even with historical laws. A comparison between Reinhold Niebuhr's twentieth-century *Faith and History,* with its confidence in the "providential structure of existence," and Bishop Jacques Bénigne Bossuet's seventeenth-century *Discourse on Universal History,* which still envisages God ruling the course of empire by "decree," is instructive in this connection. Yet even Niebuhr confessed in the end that, to a finite human mind, both the plan and mode of operation of God in history remain mysterious.

HISTORICAL INEVITABILITY. Any attempt to make fate or providence immanent in the ordinary processes of history is a move toward a second major conception of the necessity of historical events, one often referred to in contemporary discussion as the doctrine of "historical inevitability." In this conception, the course of history has a necessary overall direction, whether it be attributed to an active but impersonal "force," a nisus toward some

ultimate goal, or a "dynamic" law of development. The necessary direction of history has been variously conceived by various philosophers. Thus the Greeks tended to envisage it as cyclical and repetitive, while most philosophers of the Enlightenment found an equally simple but linear pattern of inevitable progress. According to Giambattista Vico, history traces a spiral path as civilization after civilization, each in its own unique way, follows the curve from heroic age to neobarbarism. According to Hegel, the spiral proceeds dialectically toward the actualization of a potential human freedom, each regress contributing to an ultimate spiritual synthesis. Just how deterministic such interpretations of history's direction were actually intended to be is, in fact, a disputable matter. Almost none assert that every historical event happens necessarily; the claim is usually limited to the main trend or the more significant events. And many speculative theorists do not seem to claim even that much. Oswald Spengler, for example, in his *Decline of the West* left the origin, by contrast with the development, of historical cultures unaccounted for; Hegel's lectures on the philosophy of history can be interpreted as having held that the stages of freedom succeed each other only with "rational," and not with "natural" necessity; and Arnold Toynbee's *Study of History* discovered historical "laws" so accommodating that they appear to be compatible with an almost indefinite number of exceptions.

Yet the discovery of inevitability is generally taken to be a major goal of speculative theories of history. And historians themselves often refer to "underlying tides and currents" (A. L. Rowse) or "great social forces" (E. P. Cheyney) in a way which seems to call for a more literal interpretation than the references they also occasionally let slip to the "fate" or "destiny" of historical individuals. Recent polemical works like K. R. Popper's *The Poverty of Historicism* (Boston: Beacon, 1957) and Isaiah Berlin's *Historical Inevitability* (London: Oxford Univ. Press, 1955) certainly assume that the doctrine of inevitability is still a live option for many people. Like fatalism, it is regarded by its critics as morally and politically dangerous. But it has also been subjected to a logical and conceptual critique, the major complaint of which is that insofar as historical inevitability is asserted on empirical grounds, the notion of "necessity" is employed in a way that is scientifically indefensible. According to Popper, inevitability theories confuse genuine laws, which assert conditional and hypothetical necessities, with statements of historical *trends,* which are not necessities, but facts. Laws license prediction whenever the conditions specified in their antecedent clauses are satisfied. The lack of corresponding empirical justification for the social "prophecies" obtained by merely extrapolating trends is often obscured by the "force" metaphors characteristically used in describing them.

A speculative theorist who wished to claim metaphysical rather than scientific status for his conclusions might perhaps remain unmoved by such considerations. Yet almost all inevitability theorists at some point cite empirical evidence; and in the nineteenth century particularly, such theories were often thought to provide models for social science itself. The belief that the extrapolation of trends is a scientifically respectable procedure, Popper observed, may well be traceable to the fascination that untypical sciences like astronomy have had for philosophers of history. The temptation is to say that if eclipses can be predicted by projecting the observed behavior of the solar system, then revolutions and the like ought similarly to be predictable by projecting the tendencies of the social system. Such reasoning ignores the fact that the cyclical "direction" of the solar system is not just observed; it is explained. And the explanation is in terms of initial conditions obtaining, together with laws of motion that are conditional and hypothetical. The same could be said of the so-called directional law of evolution in biology, which is sometimes cited as a paradigm for linear theories of historical inevitability. No corresponding attempt is usually made to derive the alleged necessity of observed historical trends from more fundamental considerations. For to represent the large-scale pattern as "resultant" in such a way, especially if the relevant initial conditions included individual human actions, might undermine the thesis of unavoidability.

SCIENTIFIC DETERMINISM. The notion of explaining historical trends in terms of the operation of scientific laws brings us to a third generic conception of necessity in history, the "scientific" sense. To put it most simply, an event might be said to be determined in this sense if there is some other event or condition or group of them, sometimes called its cause, that is a sufficient condition for its occurrence, the sufficiency residing in the effect's following the cause in accordance with one or more laws of nature. The general assertion of historical determinism then becomes the assertion that for every historical event there is such a sufficient condition. Whether, in consequence, history manifests a unitary pattern or direction is a further and separate question.

Race and climate. Many historical determinists who would claim to be "scientific" in the above sense have gone a step further. Like the inevitability theorists, they have sought a simple clue to the historical process, in this

case in causal factors of a limited range. Typical of such single-factor theories are those that fasten on certain biological or psychological conditions, such as the alleged racial characteristics of certain groups, or on features of the physical environment, such as topography, climate, soil, or natural resources. The writings of Joseph Arthur de Gobineau and of Houston Stewart Chamberlain, with their concept of Aryan superiority, are notorious examples of the first of these, although few serious attempts have been made to write detailed and scholarly histories (rather than propaganda) on their principles. The search for geographical determinants, on the other hand, has a reputable record going back at least to Baron de Montesquieu and Jean Bodin, and it received classic expression in the work of Henry Thomas Buckle in the nineteenth century and of Ellsworth Huntington in the twentieth. Both types of theory, however, oversimplify the diversity of history. It is one thing to point out that civilizations originated in river valleys or that the decline of Rome was accompanied by race-mixing. It is quite another—even if some features of events can properly be ascribed to such factors—to say that all significant historical change is determined by geographical or biological causes.

Social causes. Racial and environmental interpretations locate the explanatory factors outside the course of historical events themselves. Social interpretations offer single-factor accounts that seek causes in one kind of historical condition by contrast with others. According to Karl Marx, for example, the explanation of political, religious, legal, and other "ideological" features of a society is to be found in that society's mode of economic life and in the relations of production that its human elements consequently take up toward each other. In extreme forms of the theory at least, a one-way causal relation is asserted to hold at any time between economic and noneconomic factors, as well as between economic conditions at different times. Such an economic interpretation of history, with its more variable explanatory factor, has a far richer potential than racial or environmental ones for explaining the details of historical change. As with all single-factor theories, however, any attempt to defend its monistic causal claims generally either fails to carry conviction or runs afoul of a basic distinction between sufficient (determining) and merely necessary (conditioning) conditions. Thus, in a crude but revealing lapse, often cited, Friedrich Engels argued that because a man cannot engage in politics, science, religion, and art if he lacks the basic material conditions of life, the latter *determine* the former.

Multiple-factor theories. More considered statements of single-factor theories try to provide for a degree of interaction between the chosen factor and others. This leaves the difficult problem of explaining the sense, if any, in which the special factor is the fundamental one. It also leaves the problem—which bedeviled inevitability theories as well—of the relation between large-scale social causes and effects and the actions of participating individuals. "Great man" theories like Thomas Carlyle's are rightly out of fashion, but it is difficult to deny the historical importance of a Vladimir Lenin or a Napoleon Bonaparte. Georgii Valentinovich Plekhanov's classical Marxist discussion of this problem, in *The Role of the Individual in History,* adopts the uneasy compromise that individual causes can make a difference to a historical outcome, but only to its less significant features or to its timing. Such legislation as to the "spheres of influence" of various sorts of conditions, all conceded to be necessary, often seems highly arbitrary; and under pressure, single-factor theories tend to develop into "interpretations" only in the sense of directing attention to one factor in historical change that is deemed especially noteworthy, often for pragmatic reasons. The claim that historical events are determined then ceases to have any special connection with the claims made for the chosen factor. It reverts simply to the assertion that for every event there is a sufficient condition, no matter how disparate the causal elements that may sometimes be required to constitute it.

In the broad sense thus indicated, the contention that historical events are all determined may seem quite unproblematic. And when one considers the thoroughly causal language of historical accounts, the contention may seem also to be in accordance with historical practice. It is true that what historians actually call a cause is seldom itself a sufficient condition. But it is generally assumed by determinists that its claim to be a cause depends upon its completing a sufficient set of such conditions, some of which may not have been overtly specified. Yet the assumption of scientific determinism in history has been disputed on a number of grounds, the three set forth below being among the most frequently cited. These arguments have a common feature: all claim that this assumption contradicts others that the historian normally and properly makes. In consequence, the notion is represented as importing an incoherence into historical thinking as a whole.

OBJECTIONS TO DETERMINISM

CHANCE. It has been objected, first, that history is a realm in which events sometimes occur "by chance"—it

being assumed that what happens by chance cannot happen of necessity. Certainly, historians often report what happened in such terms. And chance has been regarded by some of them almost as a principle of historical interpretation. Thus J. B. Bury, in his *Later Roman Empire,* represented the success of the barbarians in penetrating the Roman Empire as due to a succession of coincidences—the "historical surprise" of the onslaught of the Asiatic Huns, which drove the Goths west and south; the lucky blow that killed a Roman emperor when the Goths engaged a Roman army that just happened to be in their way; the untimely death of that emperor's talented successor before he had arranged for the assimilation of those tribesmen who had settled within the imperial border; the unhappy fact that the two sons who subsequently divided the empire were both incompetent, and so on. Bury's example does at least afford a strong argument against the notion that history is a *self*-determining system—one of the assumptions of the doctrine of historical inevitability. It illustrates the intrusion of nonhistorical factors into the historical process—an untimely death, for example—Bury's awareness of which led him to object to any search for what he called "general" causes. Bury's example makes clearer, too, the inappropriateness of a science like astronomy as a model for social and historical explanation. For the solar system, unlike human society, is virtually isolated from such external influences. This makes it possible for us to make astronomical predictions without taking into account anything but the description of the state of the system itself at any time and to predict accurately for long periods ahead. In history the situation is very different. The sufficient conditions of historical events are seldom to be found in other historical events.

But does the admission of chance, as Bury described it, count against the whole doctrine of historical determinism in the scientific sense? In support of their claim that it must, historical indeterminists sometimes cite parallels in physical inquiry. Modern subatomic physics, for example, whether correctly or not, has often been said to be indeterministic precisely because it regards certain aspects of the behavior of single electrons as matters of chance. Yet it may be questioned whether any of the contingencies, accidents, or unlucky "breaks" mentioned by Bury were matters of chance in the physicist's sense. For there is no reason to think of any of them as uncaused. What is peculiar about them is that they occur (to use a common phrase) at the intersection of two or more relatively independent causal chains. But there is nothing in such coincidences, determinists will maintain, that enables us to say that what occurs at the "intersections" could not be deduced from prior statements of conditions and appropriate laws, provided we took all the relevant conditions into account.

In practice, of course, a historian may not be in a position to explain why a given coincidence occurred; at least one relevant chain—the biological one leading to the emperor's death, for example—may be beyond the scope of his kind of inquiry. What happened may consequently be represented by him as something unforeseen—perhaps even as the intrusion of the "irrational" into the course of events. Here the notion of chance is extended from the paradigm case where an event is said to have no cause at all to one where the cause is simply unknown because nonhistorical.

The notion is commonly extended further (as Bury's example illustrates) to events whose causes, although not beyond the range of historical inquiry, are beyond the immediate range of the historian's interests—the appearance of the Huns, for example. This makes it misleading to define "chance event" in history, as some have done, as an event that has historical effects but lacks historical causes. The causes of the invasion of the Huns simply lie outside the story the historian is telling. The judgment that a historical event happened by chance is thus a function of what the historian (and his readers) are concerned about. (This also covers the case where "by chance" seems chiefly to mean "unplanned.") It follows that, from one standpoint, an event may properly be judged to be a chance occurrence, while from another it clearly could not be: the activities of the Huns, for example, were scarcely a matter of chance from their own standpoint. Speculative philosophers of history, if they aim to take the additional standpoints of God or "History" into account, will obviously have further problems when deciding whether something was a chance occurrence. The issues thus raised are doubtless of considerable interest for a general account of the logic of historical narration. It is difficult to see, however, that they have any important bearing on the acceptability of historical determinism.

NOVELTY. A second consideration often advanced against the determinist assumption is that history is a realm of novelty and that its course must therefore remain not only unforeseen but unforeseeable, even if we take into account the broadest possible range of antecedent conditions. The fact that what the historian discovers is often surprising is thus held to have an objective basis in human creativity, from which periodically there emerge events and conditions with radically novel characteristics. Such "emergence," it is often claimed,

rules out the possibility of scientific prediction before the event because prediction is necessarily based on laws and theories that relate types of characteristics already known. In this connection it is interesting to note a "proof" offered by Popper that some historical events at least are unpredictable in principle. If we accept the common assumption that some historical events are dependent in part on the growth of human knowledge, Popper pointed out, then it is logically impossible that we should be able to predict them before they occur. For ex hypothesi, one of their conditions must remain unknown to us.

Confronted by such an argument, determinists would want to make clear that, as they conceive it, determinism does not entail predictability, even though it has, unfortunately, sometimes been defined in terms of predictability. An event can be determined even though it is not known to be so. Popper himself did not regard the argument cited above as counting against historical determinism; indeed, his own statement of it strongly suggested that the unpredictability of the events in question actually follows from their being determined in a certain way, that is, by a set of conditions that are less than sufficient in the absence of as yet unattained human knowledge. All that is required by the doctrine of determinism, however, is that events *have* sufficient conditions, whether or not they can be known before the fact. It would thus be better, perhaps, to define the notion in terms of explicability rather than predictability. Determinists often point out that the emergent characteristics of natural things can be explained in the scientific sense, although they could not have been predicted before they first emerged. In his "Determinism in History," Ernest Nagel cited the emergence of the qualities of water out of a combination of hydrogen and oxygen. These are emergent and novel in the sense of not being possessed by the original elements and not being deducible from information about the behavior of these elements in isolation. Yet we have been able to frame laws governing the emergence of these originally novel attributes under specifiable conditions that allow us to deduce and now even to predict the attributes.

A likely reply is that whereas the emergence of the characteristics of water is a recurring, experimentally testable phenomenon, the emergence of novelty in the course of history is not. At least some historical events and conditions, it may be said, are unique and hence not subject to scientific explanation even after the fact. In considering this rejoinder, however, it is important not to misunderstand the claims of scientific determinism. For these do not include the deducibility in principle of the occurrence of historical events "in all their concrete actuality." Only events as historians represent them in their narratives are said to be so deducible. And their descriptions of events, it will be argued, are necessarily phrased in terms that apply, although not necessarily in the same combinations, to events at other times and places.

It may of course be doubted that we shall ever actually discover the determining conditions of such historical novelties as Alexander's use of the phalanx, Caesar Augustus's imperial policy, or the organization of the medieval church, under descriptions as highly detailed as historians customarily apply to them—a problem scarcely touched by the consideration, advanced by Nagel, that social science has sought, with some measure of success, to discover the conditions under which men act creatively. Yet determinists will regard these as merely "practical" difficulties, not bearing on the basic issue. That issue, they will maintain, is whether the novelties that can be recognized by historical inquiry are such as to rule out their subsumability under laws "in principle." Unless historians' knowledge can be said to go beyond any description of such novelties in terms of a unique conjunction of recurring characteristics, the argument from historical novelty will be deemed to have missed its mark.

In fact, this further, and highly debatable claim is one that some historical theorists would be quite prepared to make. They would point out, for example, that we can *listen* to Wolfgang Amadeus Mozart's music and *read* Isaac Newton's scientific writings—two examples of creativity cited by Nagel—and, by thus enjoying direct acquaintance with radical historical novelty, discover more than could be conveyed by any description in terms of recurring characteristics. Ordinary historical knowledge of novel military tactics, imperial policies, or institutional organizations, they would maintain, would similarly go beyond what could be expressed without reference, either explicitly or implicitly, to named individuals, groups, or periods. They would consequently represent historical narrative as employing concrete universals—like "Renaissance" or "Gothic"—as well as abstract ones. And since scientific laws can be framed only in terms of abstract universals, they would claim that warranted assertions of novelty expressed in terms of concrete universals do undermine the assumption of determinism.

FREEDOM. A third and even more common argument against accepting a determinist view of historical events turns on the claim that history is a realm not only of chance and novelty but of human freedom. The subject

matter of history, it is sometimes said, is not mere "events" but human "actions," in a distinctive sense quite familiar to plain men who deliberate and decide what to do. If the historian is not to misrepresent such a subject matter, the argument goes, then he must take seriously the notion of choosing between alternatives. As Johan Huizinga expressed it, in his "Idea of History" (in *The Varieties of History,* edited by Fritz Stern), "the historian must put himself at a point in the past at which the known factors still seem to permit different outcomes. If he speaks of Salamis, then it must be as if the Persians might still win." In *Historical Inevitability,* Isaiah Berlin gave a further and even more familiar reason for adopting the standpoint of "agency." "If determinism were true, …" he wrote, "the notion of human responsibility, as ordinarily understood, would no longer apply." For an ascription of responsibility requires the assumption that the agent was "in control," that he could have acted otherwise than he did. Historical accounts, in other words, like the moralistic ones plain men ordinarily give of their own and others' actions, presuppose "freedom of the will." And this is held to be incompatible with the assumption of determinism.

Few philosophical problems have been discussed as exhaustively (or as inconclusively) as the problem of freedom of the will, and it is quite impossible in this context to do justice to the subtleties involved. There are, however, two chief ways of handling the present objection. Historical determinists can try to explain away the problem of freedom by arguing that, although moralistic accounts properly regard historical agents as free, the sense in which they must do so is quite compatible with the deterministic assumption. Libertarians, correspondingly, can try to give an account of historic causation that does not rule out an action's being both caused and undetermined. For historians, either of these ways out of the difficulty would presumably be more acceptable than the outright denial of the legitimacy of either moral appraisal or causal explanation in historical accounts. For, with no obvious sign of strain, historians generally offer both.

The determinist case often turns on the contention that the sense of freedom involved in attributing responsibility to a moral agent is not the "could have done otherwise" of absolute indeterminism; that sense implies only that the agent would have done otherwise if certain antecedents—his circumstances or his character, for example—had been a little different. Indeed, it is often argued that the test of whether the agent is really "in control," and hence responsible, is whether he acts differently on another occasion when the conditions have been changed—say, by his having been praised or blamed, rewarded or punished. It is therefore not the agent's freedom in the sense of his action's being uncaused that is at stake. The determinist, in arguing this way, conceives himself, furthermore, as accepting, not rejecting, the notion that the moral categories the historian uses are those of the plain man. What is denied is that the "ordinary" sense of "free" is the unconditional "freedom of the will" of the metaphysicians. As for Huizinga's claim that the historian must think of the agent's problem as if there were real possibilities open to him, this would be regarded as a purely methodological point. What is brought out thereby is the applicability to actions of a concept of understanding that requires us, quite properly, to view them in relation to what the agents thought about their situations, including any illusions they may have had about them.

Many libertarians might accept the latter contention. But most would surely repudiate the claim that responsibility requires freedom only in a sense compatible with determinism. To ascribe responsibility to a person whose actions necessarily follow from antecedent events, Berlin declared, is "stupid and cruel," and he meant rationally incoherent, not just foolish. In a sense alleged to be central to our notion of responsibility, such a person could *not* have done otherwise. Must a libertarian who takes such a stand, then, abandon the possibility of explaining actions causally? Some, at least, would say, No, provided we recognize that the term *cause,* when applied to human actions, bears a special sense. Thus, according to R. G. Collingwood, the causes (in a distinctively historical sense) of "the free and deliberate act of a conscious and responsible agent" are to be sought in the agent's "thought" about his situation, his reasons for deciding to act (*Essay on Metaphysics*). What a libertarian will deny is that any combination of such "rational" causes that excludes the agent's decision to act—since the latter falls into the historian's explanandum, not his explanans—is a sufficient condition of his action. Such causes become "effective," it might be said, only through an agent's deciding to act upon them. Yet when he does so, reference to them as his "reasons" will explain what he did in the sense of making it understandable. What such reference will not and need not do is explain his action in the sense of showing its performance to be deducible from sufficient antecedent conditions.

It is generally agreed that the conflict between historical determinists and indeterminists cannot be resolved by the offering of proofs or disproofs. Modern scientific determinists, in any case, seldom state their position dog-

matically. According to Nagel, for example, all that can be claimed is that the principle of determinism has "regulative" status as a presupposition of the possibility of scientific inquiry—a principle that must therefore govern the scientific study of history as well. What is particularly interesting about theories of rational causation is the conceptual foundation they offer for denying that the principle of determinism is a necessary presupposition even of seeking explanations when the subject matter is human action: they show at least the conceivability of explanatory inquiry on libertarian principles. It must be conceded, however, that few contemporary philosophers regard indeterminism as an acceptable assumption to carry into historical or social investigation.

See also Berlin, Isaiah; Bodin, Jean; Bossuet, Jacques Bénigne; Buckle, Henry Thomas; Carlyle, Thomas; Chamberlain, Houston Stewart; Chance; Collingwood, Robin George; Determinism, A Historical Survey; Determinism, Theological; Determinism and Freedom; Determinism and Indeterminism; Engels, Friedrich; Gobineau, Comte Joseph Arthur de; Hegel, Georg Wilhelm Friedrich; Lenin, Vladimir Il'ich; Marx, Karl; Montesquieu, Baron de; Nagel, Ernest; Newton, Isaac; Niebuhr, Reinhold; Paradigm-Case Argument; Philosophy of History; Plekhanov, Georgii Valentinovich; Popper, Karl Raimund; Providence; Spengler, Oswald.

Bibliography

For examples of determinist or near-determinist views of history, see H. T. Buckle, *A History of Civilization in England* (London, 1899) or E. Huntington, *Mainsprings of Civilization* (New York, 1945). The works of various speculative and single-factor theorists mentioned above may also be consulted: Patrick Gardiner's *Theories of History* (Glencoe, IL, 1959) contains relevant extracts from the works of Vico, Hegel, Marx, Plekhanov, Buckle, Tolstoy, Spengler, Toynbee, Croce, and Collingwood. For a contemporary attack on deterministic views, of both the scientific and metaphysical kinds, see Isaiah Berlin, *Historical Inevitability* (London: Oxford University Press, 1954) and the reply offered by E. H. Carr in *What Is History?* (London, 1961). For a moderate defense of the deterministic assumption against such attacks, see Ernest Nagel, "Determinism in History," *Philosophy and Phenomenological Research* 20 (1960): 291–317. The viability of indeterministic historical and social scientific inquiry is argued for in Alan Donagan, "Social Science and Historical Antinomianism," *Revue Internationale de Philosophie* 11 (1957): 433–449. The role of the individual in history is discussed in Sidney Hook, *The Hero in History* (New York: John Day, 1943). Johan Huizinga's "Idea of History" is included in English translation in *The Varieties of History,* edited by Fritz Stern, pp. 290–303 (New York: Meridian, 1956). The claim that historians use "cause" in a special sense is developed by R. G. Collingwood in *An Essay on Metaphysics* (Oxford: Clarendon Press, 1940), which should be read in conjunction with his *The Idea of History* (Oxford: Clarendon Press, 1946).

W. H. Dray (1967)

DEUSSEN, PAUL

(1845–1919)

Paul Deussen, the German philologist and philosopher, was the son of a Protestant clergyman in the village of Oberdreis in the Westerwald. He received a thorough classical training in the old secondary school of Pforta, where he developed a close friendship with Friedrich Nietzsche. Both Deussen and Nietzsche enrolled in the theological faculty at the University of Bonn, but Nietzsche soon shifted to classical philology and followed his teacher Ritschl to Leipzig. Deussen remained in Bonn for four semesters, then also shifted to classical philology and earned his doctorate at Berlin in 1869 with a dissertation on Plato's *Sophist.* After a brief period of teaching in secondary schools, he became the tutor for a Russian family in Geneva in 1872. There he intensified his study of Sanskrit, began a study of the Indian philosophical classics, and became an enthusiastic follower and interpreter of Arthur Schopenhauer (after having long resisted Nietzsche's enthusiastic endorsements). In 1881 he qualified to lecture in Berlin under Eduard Zeller on the basis of his work *The System of the Vedanta,* and became an extraordinary professor in 1887. Appointed full professor in Kiel in 1889, he retained this post until his retirement.

Deussen's major work, on which he labored for more than twenty years, was the *Universal History of Philosophy,* consisting of two large volumes in six parts. The first volume was devoted to Indian thought and the second to the thought of the West from the Greeks to Schopenhauer, with a section on the philosophy of the Bible.

For Deussen the history of philosophy was a discipline indispensable not only for the understanding of life but for its religious interpretation as well. Its task was to strip off the "mythical vestments" or "hulls" of the various philosophical and religious systems in order to discover the single unified truth that all share.

This unified, permanent truth was made clear in the philosophy of Immanuel Kant as completed by Schopenhauer, but it also embraced insights from the Vedanta, Plato's doctrine of Ideas, and Christian theology. Schopenhauer, Deussen said, had "freed the essentials of

Kant from the weight of traditional misunderstanding" and offered "the completion of a unified doctrine which is grounded in experience, internally coherent in its metaphysics, and which appears, in its practical part, as a Christianity renewed throughout its whole depth on scientific foundations, and which will become, and for the predictable future remain, the foundation of all human scientific and religious thought" (*Geschichte der Philosophie,* Vol. 1, Part 1, p. 22). Rightly understood, Schopenhauer was the *philosophus Christianissimus* (the most Christian philosopher). The affirmation of the will to live is the egoism of our natural existence; its denial is "disinterested righteousness, the love of man, and the willingness to sacrifice for great causes—all great, heroic, overindividual striving and creating" (*Erinnerungen an Friedrich Nietzsche,* p. 105). But the divine, in this synthetic conception, cannot be understood theistically. The highest Being is beyond all personality, and all will eventually confess, "I believe in one living, but not one personal God."

Deussen was one of the early interpreters of Jakob Boehme (1897). He edited a critical edition of Schopenhauer in fourteen volumes (Munich, 1911), and he founded the Schopenhauer Society and edited its yearbook from 1912 until his death.

See also Boehme, Jakob; Continental Philosophy; History and Historiography of Philosophy; Indian Philosophy; Kant, Immanuel; Nietzsche, Friedrich; Plato; Schopenhauer, Arthur.

Bibliography

Deussen's chief work was *Allgemeine Geschichte der Philosophie, mit besonderer Berücksichtigung der Religionen,* 2 vols. (Leipzig: F. A. Brockhaus, 1894–1917); Vol. I, Part 2 was translated by A. S. Geden as *The Philosophy of the Upanishads* (Edinburgh: T. and T. Clark, 1908). *Die Elemente der Metaphysik* (Aachen: J. A. Mayer, 1877), was translated by C. M. Duff as *The Elements of Metaphysics* (London, 1894).

Deussen was the first Western philosopher to include Eastern thought in a general history of philosophy in any scientific way. Among his publications in this field are *Das System des Vedanta* (Leipzig: F. A. Brockhaus, 1883), translated by Charles Johnston as *The System of the Vedanta* (Chicago: Open Court, 1912); *Die Sūtra des Vedānta,* translated from the Sanskrit (Leipzig, 1887), translated by H. Woods and C. B. Rumble as *The Sutras of the Vedanta with the Commentary of Cankara* (New York, 1906); *Sechzig Upanishads des Veda,* which he translated from the Sanskrit (Leipzig: F. A. Brockhaus, 1897); *Vier philosophische Texte des Mahâbhâratam* (Leipzig: F. A. Brockhaus, 1906); *Bhagavadgītā. Der Gesang des Heiligen* (Leipzig, 1911); and *Die Geheimenlehre des Veda* (Leipzig: F. A. Brockhaus, 1907–1909.

Three volumes of an autobiographical nature are *Mein Leben* (Leipzig, 1927); *Erinnerungen an Friedrich Nietzsche* (Leipzig: F. A. Brockhaus, 1901); and *Erinnerungen an Indien* (Leipzig: Lipsius and Tischer, 1904). Bound together with the *Erinnerungen an Indien* is a lecture, "On the Philosophy of the Vedanta in Its Relations to Occidental Metaphysics," delivered and first published in Bombay in 1893.

On Deussen, see "Erinnerungen an Paul Deussen," which is Vol. 20 of *Jahrbuch der Schopenhauergesellschaft* (1920).

L. E. Loemker (1967)

DEUSTUA, ALEJANDRO O.
(1849–1945)

Alejandro O. Deustua, the Peruvian educator, aesthetician, and philosopher, was born in Huancayo. He was a professor at the University of San Marcos, rector of the University, and director of the National Library in Lima. Deustua contributed greatly to the development of Peruvian education at all levels. His philosophical writing was done at an advanced age. It reflected the influence of K. C. F. Krause and Henri Bergson.

Running through the thought of Deustua are the polar ideas of liberty and order. Their interplay extends to a philosophy of civilization, but it is most clear in his major interest, aesthetics. It may be introduced through his definitions of beauty and art. Beauty is "a conciliation of liberty and nature, through the mediation of an ideal order created by the imagination." Since an internal image is not sufficient, external forms are created by art, which is the "graceful expression of the conciliation between nature and liberty, a conciliation imagined by the artist and translated by means of adequate or expressive forms."

The element of nature is furnished by human sensibility, including sensation and emotion. Liberty is found in absence of resistance, which in turn allows development from within to take place. It belongs to spirit and is paramount in that function of spirit called imagination, which is defined not as imaginal but as creative. Liberty is manifest only in an order, and it is fully realized only in an order entirely of its own making, an artistic order or harmony. This order is created by the imagination, using sensuous elements and acting in close relation with emotion. Harmony is a unity in variety: aesthetic pleasure is opposed to monotony and to excessive complexity. Types of harmony are symmetry and rhythm. Related to these are an outward order of parts and whole in space, charac-

teristic of classical art, and an inward order of causes or purposes in time, characteristic of romantic art. When liberty is realized in order, the result is grace.

In addition to beauty there are several other types of value, to all of which imagination can contribute in one degree or another. These values may in turn contribute to the aesthetic experience, but they fall below beauty in freedom. Logical truth is characterized by demonstrative necessity. Economic value is subject to the imperative of desire, in contrast to the disinterestedness of aesthetic experience. Although moral value presupposes a free agent, it requires that the will submit to duty and law. Religious revelation and myth are aesthetic in nature; but they demand submission to the divine will. Only in the aesthetic sphere is liberty sovereign, unbound by orders or norms external to it. For this reason, aesthetic value is "the value of values."

See also Aesthetics, History of; Beauty; Bergson, Henri; Imagination; Krause, Karl Christian Friedrich; Latin American Philosophy; Liberty.

Bibliography

WORKS BY DEUSTUA

"Las ideas de orden y libertad en la historia del pensamiento humano" (The ideas of order and liberty in the history of human thought). *Revista universitaria* (Lima), 1917–1922.

Estética general (General aesthetics). Lima: E. Ravago, 1923.

Estética aplicada. Lo bello en el arte: escultura, pintura, música (Applied aesthetics: the beautiful in art: sculpture, painting, music). Lima: Americana, 1935.

WORKS ON DEUSTUA

Salazar Bondy, Augusto. *La filosofía en el Perú* and *Philosophy in Peru.* Washington, DC, 1954. This is a single book, in both Spanish and English, published by the Pan American Union. The Spanish text is on pp. 35–40 and the English on pp. 77–82.

Arthur Berndtson (1967)

DEWEY, JOHN
(1859–1952)

The American philosopher, educator, and social critic John Dewey was born in Burlington, Vermont. A shy youth, he enjoyed reading books and was a good but not a brilliant student. He entered the University of Vermont in 1875, and although his interest in philosophy and social thought was awakened during his last two years there, he was uncertain about his future career. He taught classics, science, and algebra at a high school in Oil City, Pennsylvania, from 1879 to 1881 and then returned to Burlington, where he continued to teach. He also arranged for private tutorials in philosophy with his former teacher, H. A. P. Torrey. Encouraged by Torrey and W. T. Harris, the editor of the *Journal of Speculative Philosophy* who accepted Dewey's first two philosophical articles, Dewey applied for the graduate program at the newly organized Johns Hopkins University. He was twice refused fellowship aid, but he borrowed $500 from an aunt to begin his professional philosophical career.

The external events of Dewey's Vermont years were relatively unexciting, and there is very little to indicate that he would become America's most influential philosopher and educator as well as one of the most outspoken champions of social reform. Yet the New England way of life left a deep imprint on the man and his thought. His modesty, forthrightness, doggedness, deep faith in the workings of the democratic process, and respect for his fellow man are evidenced in almost everything that he did and wrote.

Under the imaginative guidance of Daniel Gilman, the first president of Johns Hopkins, the university had become one of the most exciting centers for intellectual and scholarly activity. Dewey studied with C. S. Peirce, who taught logic, and with G. S. Hall, one of the first experimental psychologists in America. The greatest initial influence on Dewey, however, was G. S. Morris, whose philosophical outlook had been shaped by G. W. F. Hegel and the idealism so much in vogue on the Continent and in England.

Dewey was an eager participant in the controversies stirred up by Hegelianism. He dated his earliest interest in philosophy to a course in physiology that he took during his junior year at the University of Vermont, where he read T. H. Huxley's text on physiology. Dewey discovered the concept of the organic and developed a sense of the interdependence and interrelated unity of all things. He tells us that subconsciously he desired a world and a life that would have the same properties as had the human organism that Huxley described. In Hegel and the idealists, Dewey discovered the most profound philosophical expression of this emotional and intellectual craving. From this organic perspective, which emphasized process and change, all distinctions are functional and relative to a developing unified whole. The organic perspective could be used to oppose the static and the fixed and to break down the hard and fast dichotomies and dualisms that had plagued philosophy.

Dewey's writings during his Hegelian period are infused with an evangelical spirit and are as enthusiastic

as they are vague. Whatever issue Dewey considered, he was convinced that once viewed from the perspective of the organic, old problems would dissolve and new insights would emerge. Long after Dewey had drifted away from his early Hegelianism, his outlook was shaped by his intellectual bias for a philosophy based on change, process, and dynamic, organic interaction.

After completing his doctoral studies at Johns Hopkins with a dissertation on the psychology of Immanuel Kant, Dewey joined Morris at the University of Michigan in 1884. He remained there for the next ten years, with the exception of one year (1888) when he was a visiting professor at the University of Minnesota. At Michigan, Dewey worked with G. H. Mead, who later joined Dewey at Chicago. During his years at Michigan, Dewey became dissatisfied with pure speculation and sought ways to make philosophy directly relevant to the practical affairs of men. His political, economic, and social views became increasingly radical. He agreed to edit a new weekly with a socialist orientation, to be called *Thought News,* but it never reached publication. Dewey also became directly involved with public education in Michigan. His scientific interests, especially in the field of psychology, gradually overshadowed his interest in pure speculation. He published several books on theoretical and applied psychology, including *Psychology* (New York, 1887; 3rd rev. ed., 1891), *Applied Psychology* (Boston, 1889), and *The Psychology of Number and Its Applications to Methods of Teaching Arithmetic* (New York, 1895). The latter two books were written with J. A. McLellan.

Dewey's appointment in 1894 as chairman of the department of philosophy, psychology, and education at the University of Chicago provided an ideal opportunity for consolidating his diverse interests. In addition to his academic responsibilities, Dewey actively participated in the life of Hull House, founded by Jane Addams, where he had an opportunity to become directly acquainted with the social and economic problems brought about by urbanization, rapid technological advance, and the influx of immigrant populations. Dewey mixed with workers, union organizers, and political radicals of all sorts. At the university, Dewey assembled a group of sympathetic colleagues who worked closely together. Collectively they published the results of their research in a volume of the Decennial Publications of the University of Chicago titled *Studies in Logical Theory* (Chicago, 1903). William James, to whom the book was dedicated, rightly predicted that the ideas developed in the *Studies* would dominate the American philosophical scene for the next twenty-five years.

Shortly after Dewey arrived in Chicago, he helped found the famous laboratory school, commonly known as the Dewey School, which served as a laboratory for testing and developing his psychological and pedagogic hypotheses. Some of Dewey's earliest and most important books on education were based on lectures delivered at the school: *The School and Society* (Chicago, 1900) and *The Child and the Curriculum* (Chicago, 1902). When Dewey left Chicago for Columbia in 1904 because of increasing friction with the university administration concerning the laboratory school, he had already acquired a national reputation for his philosophical ideas and educational theories. The move to Columbia, where he remained until his retirement in 1930, provided a further opportunity for development, and Dewey soon gained international prominence. Through the Columbia Teachers College, which was a training center for teachers from many countries, Dewey's educational philosophy spread throughout the world.

At the time that Dewey joined the Columbia faculty, the *Journal of Philosophy* was founded by F. J. E. Woodbridge, and it became a forum for the discussion and defense of Dewey's ideas. There is scarcely a volume from the time of its founding until Dewey's death that does not contain an article either by Dewey or about his philosophy. As the journalistic center of the country, New York also provided Dewey with an opportunity to express himself on pressing political and social issues. He became a regular contributor to the *New Republic.* A selection of Dewey's popular essays is collected in *Characters and Events,* 2 vols. (New York, 1929).

Wherever Dewey lectured he had an enormous influence. From 1919 to 1921, he lectured at Tokyo, Beijing, and Nanjing, and his most popular book, *Reconstruction in Philosophy* (New York, 1920), is based on his lectures at the Imperial University of Japan. He also conducted educational surveys of Turkey, Mexico, and Russia. Although he retired from Columbia in 1930, he remained active and wrote prolifically until his death. In 1937, when Dewey was seventy-eight, he traveled to Mexico to head the commission investigating the charges made against Leon Trotsky, during the Moscow trials. After a careful investigation, the commission published its report, *Not Guilty* (New York, 1937). In 1941 Dewey championed the cause of academic freedom when Bertrand Russell—his arch philosophical adversary—had been denied permission to teach at the City College of New York, Dewey collaborated in editing a book of essays protesting the decision.

Although constantly concerned with social and political issues, Dewey continued to work on his more technical philosophical studies. M. H. Thomas's bibliography of his writings comprises more than 150 pages. Dewey's influence extended not only to his colleagues but to leaders in almost every field. The wide effects of his teaching did not depend upon the superficial aspects of its presentation, for Dewey was not a brilliant lecturer or essayist, although he could be extremely eloquent. His writings are frequently turgid, obscure, and lacking in stylistic brilliance. But more than any other American of his time, Dewey expressed the deepest hopes and aspirations of his fellow man. Whether dealing with a technical philosophical issue or with some concrete injustice, he displayed a rare combination of acuteness, good sense, imagination, and wit.

EXPERIENCE AND NATURE

The key concept in Dewey's philosophy is experience. Although there is a development from an idealistic to a naturalistic analysis of experience and different emphases in his many discussions of the concept, a nevertheless coherent view of experience does emerge. In his early philosophy Dewey was sympathetic to the theory of experience developed by the Hegelians and the nineteenth-century idealists. He thought of experience as a single, dynamic, unified whole in which everything is ultimately interrelated. There are no rigid dichotomies or breaks in experience and nature. All distinctions are functional and play a role in a complex organic system. Dewey also shared the idealists' antipathy to the atomist and subjectivist tendencies in the concept of experience elaborated by the British empiricists. But as Dewey drifted away from his early Hegelian orientation he indicated three major respects in which he rejected the idealistic concept of experience.

First, he charged that the idealists, in their preoccupation with knowledge and knowing, distorted the character of experience. Idealists, Dewey claimed, neglected the noncognitive and nonreflective experiences of doing, suffering, and enjoying that set the context for all knowing and inquiry. Philosophy, especially modern philosophy, had been so concerned with epistemological issues that it mistook all experience as a form of knowing. Such bias inevitably distorts the character of both man's experience and his knowing. Man is primarily a being who acts, suffers, and enjoys. Most of his life consists of experiences that are not primarily reflective. If we are to understand the nature of thought, reflection, inquiry, and their role in human life, we must appreciate their emergence from, and conditioning by, the context of nonreflective experience. There is more to experience, Dewey believed, than is to be found in the writings of the idealists and, indeed, in the writings of most epistemologists.

The second major departure from his early idealism is to be found in Dewey's rejection of the idea of a single unified whole in which everything is ultimately interrelated. In this respect, he displayed an increasing sympathy with the pluralism of the British empiricists. He insisted that life consists of a series of overlapping and interpenetrating experiences, situations, or contexts, each of which has its internal qualitative integrity. The individual experience is the primary unit of life.

The third shift is reflected in Dewey's increasingly naturalistic bias. The Hegelians and the nineteenth-century idealists did have important insights into the organic nature of experience, but they had overgeneralized them into a false cosmic projection. Dewey discovered in the new developing human sciences, especially in what he called the anthropological-biological orientation, a more careful, detailed, scientific articulation of the organic character of experience.

Dewey thought of himself as part of a general movement that was developing a new empiricism based on a new concept of experience, one that combined the strong naturalistic bias of the Greek philosophers with a sensitive appreciation for experimental method as practiced by the sciences. He was sympathetic with what he took to be the Greek view of experience, which considers it as consisting of a fund of social knowledge and skills and as being the means by which man comes into direct contact with a qualitatively rich and variegated nature. But Dewey was just as forceful in pointing out that this view of experience had to be reconstructed in light of the experimental method of the sciences. One of his earliest and clearest discussions of the nature of experience as an organic coordination is to be found in "The Reflex Arc Concept in Psychology" (*Psychological Review,* Vol. 3, 1896).

Dewey's interest in developing a new theory of experience led many critics to question the exact status of experience within nature, and some objectors charged him with excessive anthropomorphism. Sensitive to this type of criticism, Dewey, particularly in *Experience and Nature* (Chicago, 1925; 2nd ed., New York, 1929), attempted to deal with this criticism and to sketch a metaphysics, "the descriptive study of the generic traits of existence."

Nature, according to Dewey, consists of a variety of transactions that can be grouped into three evolutionary plateaus, or levels. Transaction is the technical term that Dewey used to designate the type of action in which the components and elements involved in the action both condition and are conditioned by the entire coordination. The elements of a transaction play a functional role in the developing coordination. The three plateaus of natural transactions are the physicochemical, the psychophysical, and the level of human experience. There are no sharp breaks or discontinuities within nature. But there are distinctive characteristics of the different levels of natural transactions that are reflected in their patterns of behavior and in their consequences. From this perspective, human experience consists of one type of natural transaction, a type that has been the latest to evolve. The distinguishing characteristics of this level of natural transaction are to be located in the type of language, communication, and social living that humans have developed. Experience is all-inclusive in the sense that man is involved in continuous transactions with the whole of nature, and through systematic inquiry he can come to understand the essential characteristics of nature. Some of the more specific areas of Dewey's philosophy can be investigated against this panoramic view of experience and nature.

ART AND EXPERIENCE

The ideas contained in Dewey's *Art as Experience* (New York, 1934) provided a surprise for many readers. Popular versions of his philosophy had so exaggerated the role of the practical and the instrumental that art and aesthetic experience seemed to have no place in his philosophical outlook. More perceptive commentators realized that Dewey was making explicit a dimension of his view of experience that had always been implicit and essential to an understanding of his philosophy. The meaning and role of art and aesthetic quality are crucial for understanding Dewey's views on logic, education, democracy, ethics, social philosophy, and even technology.

Dewey had persistently claimed that knowing, or more specifically, inquiry, is an art requiring active experimental manipulation and testing. Knowing does not consist of the contemplation of eternal forms, essences, or universals. Dewey argued that the "spectator theory of knowledge," which had plagued philosophy from its beginnings, is mistaken. He also objected to the sharp division between the theoretical sciences and the practical arts that had its explicit source in Aristotle and had influenced so much later philosophy. Dewey maintained that Aristotle's analysis of the practical disciplines is more fruitful for developing an adequate theory of inquiry than is his description of the theoretical sciences of knowing. Not only is inquiry an art, but all life is, or can be, artistic. The so-called fine arts differ in degree, not in kind, from the rest of life.

Dewey also gave a prominent place to what he called immediacy, pervasive quality, or aesthetic quality. This immediacy is not restricted to a special type of experience but is a distinctive feature of anything that is properly called "*an* experience." The primary unit of life, we have mentioned, is *an* experience, a natural transaction of acting, suffering, enjoying, knowing. It has both temporal development and spatial dimension and can undergo internal change and reconstruction.

But what is it that enables us to speak of an individual experience? Or, by virtue of what does an experience, situation, or context have a unity that enables us to distinguish it from other experiences? Dewey's answer is that everything that is an experience has immediacy or pervasive quality that binds together the complex constituents of the experience. This immediacy or pervasive quality can be directly felt or had. But this qualitative dimension of experience is not to be confused with a subjective feeling that is somehow locked up in the mind of the experiencer. Nor is it to be thought of as something that exists independently of any experiencer. These qualities that pervade natural transactions are properly predicated of the experience or situation as a whole. Within an experiential transaction we can institute distinctions between what is subjective and what is objective. But such distinctions are relative to, and dependent on, the context in which they are made. An experience or a situation is a whole in virtue of its immediate pervasive qualities, and each occurrence of these qualities is unique. As examples of such pervasive qualities, Dewey mentions the qualities of distress or cheer that mark existent situations, qualities that are unique in their occurrence and inexpressible in words but capable of being directly experienced. Thus, when one directly experiences a frightening situation, it is the situation that is frightening and not merely the experience.

These pervasive, or "tertiary," qualities are what Dewey calls aesthetic qualities. Aesthetic quality is thus an essential characteristic of all experiences. Within an experience, the pervasive quality can guide the development of the experience, and it can also be transformed and enriched as the experience is reconstructed. Aesthetic quality can be funded with new meaning, ideas, and emotions. A situation that is originally indeterminate, slack,

or inchoate can be transformed into one that is determinate, harmonious, and funded with meaning; this type of reconstructed experience Dewey called a consummation. Such experiences are reconstructed by the use of intelligence. For example, when one is confronted with a specific problematic situation that demands resolution, one can reconstruct the situation by locating its problematic features and initiating a course of action that will resolve the situation. Consummations are characteristic of the most mundane practical tasks as well as the most speculative inquiries. The enemies of the aesthetic, Dewey claimed, are not the practical or the intellectual but the diffuse and slack at one extreme and the excessively rigid and fixed at the other. The type of experience that philosophers normally single out as aesthetic is a heightened consummation in which aesthetic qualities dominate.

Dewey viewed human life as a rhythmic movement from experiences qualified by conflict, doubt, and indeterminateness toward experiences qualified by their integrity, harmony, and funded aesthetic quality. We are constantly confronted with problematic and indeterminate situations, and insofar as we use our intelligence to reconstruct these situations successfully we achieve consummations. He was concerned both with delineating the methods by which we could most intelligently resolve the conflicting situations in which we inevitably find ourselves and with advocating the social reforms required so that life for all men would become funded with enriched meaning and increased aesthetic quality.

LOGIC AND INQUIRY

Early in his career, Dewey started developing a new theory of inquiry, which he called instrumental or experimental logic. Dewey claimed that philosophers had lost touch with the actual methods of inquiry practiced by the experimental sciences. The function of instrumental logic is to study the methods by which we most successfully gain and warrant our knowledge. On the basis of this investigation, instrumental logic could specify regulative principles for the conduct of further inquiry.

The central themes of Dewey's conception of logic were outlined in *Studies in Logical Theory* (Chicago, 1903), applied to education in *How We Think* (Boston, 1910), and further refined in *Essays in Experimental Logic* (Chicago, 1916). Dewey also wrote numerous articles on various aspects of logic, but his most systematic and detailed presentation is in *Logic: The Theory of Inquiry* (New York, 1938), in which he defines inquiry as "*the controlled or directed transformation of an indeterminate situation into one that is so determinate in its constituent distinctions and relations as to convert the elements of the original situation into a unified whole*" (p. 104). By itself, this definition is not sufficient to grasp what Dewey intends. But his meaning can be understood when the definition is interpreted against the background of what we have said about the individual experience or situation and the way in which it is pervaded by a unifying quality.

We find ourselves in situations that are qualified by their indeterminateness or internal conflict. From the perspective of the experiencer or inquirer, we can say that he experiences a "felt difficulty." This is the antecedent condition of inquiry. Insofar as the situation demands some resolution, we must attempt to articulate the problem or problems that are to be solved. Formulating the problems may be a process of successive refinement in the course of the inquiry. The next logical stage is that of suggestion or hypothesis, in which we imaginatively formulate various relevant hypotheses for solving the problem. In some complex inquiries we may have to engage in hypothetico-deductive reasoning in order to refine our hypotheses and to ascertain the logical consequences of the hypothesis or set of hypotheses. Finally, there is the stage of experimental testing in which we seek to confirm or disconfirm the suggested hypotheses. If our inquiry is successful, the original indeterminate situation is transformed into a unified whole. Knowledge may be defined as the objective of inquiry. Knowledge is that which is warranted by the careful use of the norms and methods of inquiry. When "knowledge" is taken as an abstract term related to inquiry in the abstract, it means warranted assertibility. Furthermore, the knowledge gained in a specific inquiry is funded in our experience and serves as the background for further inquiry. By reflecting on this general pattern of inquiry, which can be exhibited in commonsense inquiry as well as the most advanced scientific inquiry, we can bring into focus the distinctive features of Dewey's logic.

First, this pattern of inquiry is intended to be a general schema for all inquiry. But the specific procedures, testing methods, type of evidence, and so on, will vary with different types of inquiry and different kinds of subject matter. Second, a specific inquiry cannot be completely isolated from the context of other inquiries. The rules, procedures, and evidence required for the conduct of any inquiry are derived from other successful inquiries. By studying the types of inquiry that have been most successful in achieving warranted conclusions, we can abstract norms, rules, and procedures for directing further inquiry. These norms may themselves be modified in

the course of further inquiry. Third, all inquiry presupposes a social or public context that is the medium for funding the warranted conclusions and norms for further inquiry. In this respect, Dewey agrees with Peirce's emphasis on the community of inquirers. Inquiry both requires such a community and helps to further the development of this community. Dewey attempted to relate this idea of a community of inquirers to his view of democracy. The essential principle of democracy is that of community; an effective democracy requires the existence of a community of free, courageous, and open-minded inquirers. Fourth, inquiry is essentially a self-corrective process. To conduct a specific inquiry, some knowledge claims, norms, and rules must be taken as fixed, but no knowledge claim, norm, or rule is absolutely fixed; it may be criticized, revised, or abandoned in light of subsequent inquiry and experience.

Dewey's theory of inquiry as an ongoing self-corrective process and his view of knowledge as that which is warranted through inquiry both differ radically from many traditional theories of inquiry and knowledge. Dewey thought of this theory as an alternative to the views of those philosophers who have claimed that there is an epistemological given that is indubitable and known with certainty. According to this epistemological model, some truths are considered to be absolutely certain, indubitable, or incorrigible. They may be considered self-evident, known by rational insight, or directly grasped by the senses. On the basis of this foundation, we then construct the rest of our knowledge. From Dewey's perspective, this general model that has informed many classical theories of knowledge is confused and mistaken. There are no absolute first truths that are given or known with certainty. Furthermore, knowledge neither has nor requires such a foundation in order to be rational. Inquiry and its objective, knowledge, are rational because inquiry is a self-corrective process by which we gradually become clearer about the epistemological status of both our starting points and conclusions. We must continually submit our knowledge claims to the public test of a community of inquirers in order to clarify, refine, and justify them.

DEMOCRACY AND EDUCATION

Dewey is probably best known for his philosophy of education. This is not a special branch of his philosophy, however, for he claimed that all philosophy can be conceived of as the philosophy of education. And it is certainly true that all the concepts we have discussed inform his thinking about education. He returned again and again to the subject of education, but the essential elements of his position can be found in *My Pedagogic Creed* (New York, 1897), *The School and Society* (Chicago, 1900), *The Child and the Curriculum* (Chicago, 1902), and especially in his comprehensive statement in *Democracy and Education* (New York, 1916).

It is essential to appreciate the dialectical context in which Dewey developed his educational ideas. He was critical of the excessively rigid and formal approach to education that dominated the practice of most American schools in the latter part of the nineteenth century. He argued that such an approach was based upon a faulty psychology in which the child was thought of as a passive creature upon whom information and knowledge had to be imposed. But Dewey was equally critical of the "new education," which was based on a sentimental idealization of the child. This child-oriented approach advocated that the child himself should pick and choose what he wanted to study. It also was based on a mistaken psychology, which neglected the immaturity of the child's experience. Education is, or ought to be, a continuous reconstruction of experience in which there is a development of immature experience toward experience funded with the skills and habits of intelligence. The slogan "Learn by Doing" was not intended as a credo for anti-intellectualism but, on the contrary, was meant to call attention to the fact that the child is naturally an active, curious, and exploring creature. A properly designed education must be sensitive to this active dimension of life and must guide the child, so that through his participation in different types of experience his creativity and autonomy will be cultivated rather than stifled.

The child is not completely malleable, nor is his natural endowment completely fixed and determinate. Like Aristotle, Dewey believed that the function of education is to encourage those habits and dispositions that constitute intelligence. Dewey placed great stress on creating the proper type of environmental conditions for eliciting and nurturing these habits. His conception of the educational process is therefore closely tied to the prominent role that he assigned to habit in human life. (For a detailed statement of the nature and function of habit, see *Human Nature and Conduct,* New York, 1922.) Education as the continuous reconstruction and growth of experience also develops the moral character of the child. Virtue is taught not by imposing values upon the child but by cultivating fair-mindedness, objectivity, imagination, openness to new experiences, and the courage to change one's mind in the light of further experience.

Dewey also thought of the school as a miniature society; it should not simply mirror the larger society but

should be representative of the essential institutions of this society. The school as an ideal society is the chief means for social reform. In the controlled social environment of the school it is possible to encourage the development of creative individuals who will be able to work effectively to eliminate existing evils and institute reasonable goods. The school, therefore, is the medium for developing the set of habits required for systematic and open inquiry and for reconstructing experience that is funded with greater harmony and aesthetic quality.

Dewey perceived acutely the threat posed by unplanned technological, economic, and political development to the future of democracy. The natural direction of these forces is to increase human alienation and to undermine the shared experience that is so vital for the democratic community. For this reason, Dewey placed so much importance on the function of the school in the democratic community. The school is the most important medium for strengthening and developing a genuine democratic community, and the task of democracy is forever the creation of a freer and more humane experience in which all share and participate.

ETHICS AND SOCIAL PHILOSOPHY

In order to understand Dewey's moral philosophy, we must again focus on his concept of the situation. Man is a creature who by nature has values. There are things, states of affairs, and activities that he directly enjoys, prizes, or values. Moral choices and decisions arise only in those situations in which there are competing desires or a conflict of values. The problem that a man then confronts is to decide what he really wants and what course of action he ought to pursue. He cannot appeal to his immediate values to resolve the situation; he must evaluate or appraise the situation and the different courses of action open to him. This process of deliberation that culminates in a decision to act is what Dewey calls "valuation." But how do we engage in this process of valuation? We must analyze the situation as carefully as we can, imaginatively project possible courses of action, and scrutinize the consequences of these actions. Those ends or goods that we choose relative to a concrete situation after careful deliberation are reasonable or desirable goods. Our choices are reasonable to the extent that they reflect our developed habits of intelligence. Choices will be perverse or irrational if they are made on the basis of prejudice and ignorance. Dewey is fully aware that there are always practical limitations to our deliberations, but a person trained to deliberate intelligently will be prepared to act intelligently even in those situations that do not permit extended deliberation. When we confront new situations we must imagine and strive for new goals. As long as there is human life, there will always be situations in which there are internal conflicts that demand judgment, decision, and action. In this sense, the moral life of man is never completed, and the ends achieved become the means for attaining further ends. But lest we think that man is always striving for something that is to be achieved in the remote future, or never, Dewey emphasized that there are consummations—experiences in which the ends that we strive for are concretely realized.

It should be clear that such a view of man's moral life places a great deal of emphasis on intelligence. Dewey readily admitted his "faith in the power of intelligence to imagine a future which is a projection of the desirable in the present, and to invent the instrumentalities of its realization." It should also be clear that ethics conceived of in this manner blends into social philosophy. Valuation, like all inquiry, presupposes a community of shared experience in which there are common norms and procedures, and intelligent valuation is also a means for making such a community a concrete reality. Here, too, ends and norms are clarified, tested, and modified in light of the cumulative experience of the community. Furthermore, it is the objective of social philosophy to point the way to the development of those conditions that will foster the effective exercise of practical intelligence. The spirit that pervades Dewey's entire philosophy and finds its perfect expression in his social philosophy is that of the reformer or reconstructor, not the revolutionary. Dewey was always skeptical of panaceas and grand solutions for eliminating existing evils and injustices. But he firmly believed that with a realistic scientific knowledge of existing conditions and with a cultivated imagination, men could ameliorate the human condition. To allow ourselves to drift in the course of events or to fail to assume our responsibility for continuous reconstruction of experience inevitably leads to the dehumanization of man.

PHILOSOPHY AND CIVILIZATION

Dewey presented a comprehensive and synoptic image of man and the universe. The entire universe consists of a multifarious variety of natural transactions. Man is at once continuous with the rest of nature and exhibits distinctive patterns of behavior that distinguish him from the rest of nature. His experience is also pervaded with qualities that are not reducible to less complex natural transactions. Thus, Dewey attempted to place man within the context of the whole of nature. In addition, Dewey was sensitive to the varieties of human experience. He

sought to delineate the distinctive features of different aspects of experience, ranging from mundane practical experience to the religious dimension of experience. Within the tradition of philosophy Dewey may be characterized as a robust naturalist or a humanistic naturalist. His philosophy is both realistic and optimistic. There will always be conflicts, problems, and competing values within our experience, but with the continuous development of "creative intelligence" men can strive for and realize new ends and goals.

This synoptic view of man and the universe is closely related to Dewey's conception of the role of philosophy in civilization. Philosophy is dependent on, but should attempt to transcend, the specific culture from which it emerges. The function of philosophy is to effect a junction of the new and the old, to articulate the basic principles and values of a culture, and to reconstruct these into a more coherent and imaginative vision. Philosophy is therefore essentially critical and, as such, will always have work to do. For as the complex of traditions, values, accomplishments, and aspirations that constitute a culture changes, so must philosophy change. Indeed, in pointing the way to new ideals and in showing how these may be effectively realized, philosophy is one of the means for changing a culture. Philosophy is continually faced with the challenge of understanding the meaning of evolving cultures and civilizations and of articulating new projected ideals. The motif of reconstruction that runs throughout Dewey's investigations dominates his conception of the role of philosophy in civilization. He epitomized the spirit of his entire philosophical endeavor in his "plea for casting off of that intellectual timidity which hampers the wings of imagination, a plea for speculative audacity, for more faith in ideas, sloughing off a cowardly reliance upon those partial ideas to which we are wont to give the name facts." He fully realized that he was giving philosophy a more modest function than had been given by those who claimed that philosophy reveals an eternal reality. But such modesty is not incompatible with boldness in the maintenance of this function. As Dewey declared, "a combination of such modesty and courage affords the only way I know of in which the philosopher can look his fellow man in the face with frankness and humanity" (*Philosophy and Civilization,* p. 12).

See also Aesthetic Experience; Aesthetic Qualities; Aristotle; Experience; Harris, William Torrey; Hegel, Georg Wilhelm Friedrich; Hegelianism; Huxley, Thomas Henry; Idealism; James, William; Kant, Immanuel; Mead, George Herbert; Naturalism; Peirce, Charles Sanders; Philosophy of Education, History of; Pragmatism; Value and Valuation; Woodbridge, Frederick James Eugene.

Bibliography

The most exhaustive bibliography of John Dewey's writings is M. H. Thomas's *John Dewey: A Centennial Bibliography* (Chicago: University of Chicago Press, 1962). This excellent guide includes a comprehensive listing of Dewey's writings, translations and reviews of his works, and a bibliography of books, articles, and dissertations about Dewey.

A less comprehensive bibliography of Dewey's writings can also be found in *The Philosophy of John Dewey,* edited by P. A. Schilpp (Evanston, IL: Northwestern University Press, 1939).

The following secondary sources are helpful as general introductions to Dewey's life and philosophy. Richard J. Bernstein, *John Dewey* (New York: Washington Square Press, 1966) focuses on the concept of experience and nature. George R. Geiger, *John Dewey in Perspective* (New York: Oxford University Press, 1958) stresses the role of aesthetic experience as the key to Dewey's philosophy. Sidney Hook, *John Dewey: An Intellectual Portrait* (New York: Day, 1939) captures both the spirit and letter of Dewey as a man, social reformer, and philosopher. Robert J. Roth, S.J., *John Dewey and Self-Realization* (Englewood Cliffs, NJ: Prentice-Hall, 1962) shows the importance and meaning of religious experience for Dewey.

Richard J. Bernstein (1967)

DEWEY, JOHN [ADDENDUM]

John Dewey has undergone an extraordinary renaissance of scholarly and public concern with his thought. Dewey (1859–1952) was encyclopedic in both his interests and achievements. The full and startling range of his written reflections is now apparent with the completed publication of his *Works* in a critical edition of thirty-seven volumes. Commentaries and critical interpretations have followed apace.

In the mediated public mind, prior discussion of Dewey's thought for the most part was devoted to his work on education, both in theory and practice. Unfortunately, these discussions of Dewey's approach to pedagogy and to schooling as an institution in a democratic society were often disconnected from his metaphysics, aesthetics, and social and political philosophy. This interpretive mishap is now being rectified with the appearance of many perceptive studies of Dewey's thought, including his previously neglected thoughts on religion and logic.

Fundamentally, John Dewey is an unregenerate philosophical naturalist, one for whom the human jour-

ney is constitutive of its own meaning and is not to be rescued by any transcendent explanations, principles of accountability, or posthumous salvation. Obviously, this position is both liberating and baleful, in that it throws us back on our own human resources, for better and for worse. In effect, we are responsible for our actions, for the course of human history, and we are called upon to navigate between the shoals of supine obeisance and arrogant usurpation. In *A Common Faith* (1934), Dewey warns of the danger to human solidarity when we do not accept this responsibility. "Weak natures take to reverie as a refuge as strong ones do to fanaticism. Those who dissent are mourned over by the first class and converted through the use of force by the second."

Leaving no philosophical stone unturned, Dewey addresses the pitfalls and possibilities of the human condition from a wide array of vantage points. His central text is *Experience and Nature,* in which he probes the transactions of the human organism with the affairs of nature. These transactions are to be understood and diagnosed as experiential oscillations between the "precarious" and the "stable." The settings for this trenchant discussion include communication, mind, art, and value. In retrospect, Dewey offered that he should have titled this work *Culture and Nature,* an appropriate reconsideration, for it is helpful to read Dewey as a philosopher of culture, with an eye toward his grasp of human institutions, social, political, and educational.

Since the 1980s the focus of commentaries on the work of Dewey has been directed to his social and political philosophy, particularly his writings between 1927 and 1935, namely, *The Public and Its Problems, Individualism Old and New,* and *Liberalism and Social Action.* Although Dewey's thought was indigenous to American culture, it is nonetheless remarkable that themes found in Marxist and existentialist traditions are present in these writings, cast differently but equally telling. Of special note is the renewed admiration for Dewey's philosophy of community and his deep grasp of the complex relationships of individuals in communities. For Dewey the irreducible trait of human life is found in the activity of face-to-face communities. Their quality is the sign of how we are faring, humanly. At the end of *Human Nature and Conduct,* he writes a message for his time and for our time as well.

> Within the flickering inconsequential acts of separate selves dwells a sense of the whole which claims and dignifies them. In its presence we put off mortality and live in the universal. The life of the community in which we live and have our being is the fit symbol of this relationship. The acts in which we express our perception of the ties which bind us to others are its only rites and ceremonies.

See also Existentialism; Feminism and Pragmatism; Marxist Philosophy; Social and Political Philosophy.

Bibliography

WORKS BY DEWEY

Works of John Dewey. Edited by J. A. Boydston. Carbondale: Southern Illinois University Press, 1969–1990. Divided into Early Works (5 vols.), Middle Works (15 vols.), and Later Works (17 vols.). This is a critical edition.

The Correspondence of John Dewey, 1871–1952. Edited by Larry A. Hickman. Charlottesville, VA: InteLex Corporation, 1999–2005. Vol. 1 (1871–1918), 1999. Rev. ed. 2001, 2005; Vol. 2 (1919–1939), 2001. Rev. ed. 2005; Vol. 3 (1940–1952), 2005.

The Essential Dewey. 2 vols. Edited by Larry A. Hickman and Thomas Alexander. Bloomington: Indiana University Press, 1998.

The Philosophy of John Dewey. Edited by J. J. McDermott. Chicago: University of Chicago Press, 1981. Complete selections from Dewey's major writings.

WORKS ON DEWEY

Campbell, J. *Understanding John Dewey.* Chicago: Open Court, 1995. The most intelligent and accurate interpretation of Dewey's thought overall.

Hickman, L. A. *John Dewey's Pragmatic Technology.* Bloomington: Indiana University Press, 1990.

Hickman, L. A., ed. *Reading Dewey: Interpretations for a Postmodern Generation.* Bloomington: Indiana University Press, 1998.

Martin, Jay. *The Education of John Dewey—A Biography.* New York: Columbia University Press, 2003.

Rockefeller, S. C. *John Dewey: Religious Faith and Democratic Humanism.* New York: Columbia University Press, 1991.

Schilpp, P. A., and L. E. Hahn, eds. *The Philosophy of John Dewey,* 3rd ed. La Salle, IL: Open Court, 1989. Contains a bibliography of Dewey's publications with entries and corrections until 1989.

Shook, John R. *Dewey's Empirical Theory of Knowledge and Reality.* Nashville, TN: Vanderbilt University Press, 2000.

Sleeper, R. S. *The Necessity of Pragmatism: John Dewey's Conception of Philosophy.* New Haven, CT: Yale University Press, 1986. An insightful presentation of the relationship between Dewey's thought and major currents in contemporary philosophy.

Welchman, J. *Dewey's Ethical Thought.* Ithaca, NY: Cornell University Press, 1995.

Westbrook, R. *John Dewey and American Democracy.* Ithaca, NY: Cornell University Press, 1991. A synoptic and especially perceptive book on Dewey's social thought.

John J. McDermott (1996)
Bibliography updated by John J. McDermott (2005)

DIALECTIC

The term *dialectic* originates in the Greek expression for the art of conversation (*διαλεκτικὴ τέχνη*). So far as its great variety of meanings have anything in common, it is perhaps that dialectic is a method of seeking and sometimes arriving at the truth by reasoning, but even this general description, which to fit the variety of cases is so vague as to be valueless, fails to do justice to the Hegelian and Marxist notion of dialectic as a historical process. However, among the more important meanings of the term have been (1) the method of refutation by examining logical consequences, (2) sophistical reasoning, (3) the method of division or repeated logical analysis of genera into species, (4) an investigation of the supremely general abstract notions by some process of reasoning leading up to them from particular cases or hypotheses, (5) logical reasoning or debate using premises that are merely probable or generally accepted, (6) formal logic, (7) the criticism of the logic of illusion, showing the contradictions into which reason falls in trying to go beyond experience to deal with transcendental objects, and (8) the logical development of thought or reality through thesis and antithesis to a synthesis of these opposites. Meaning (2) is notably still current, and the term is often used in a pejorative sense.

In the following discussion the different kinds of dialectic will be elucidated in their historical order.

SOCRATES AND HIS PREDECESSORS

Dialectic perhaps originated in the fifth century BCE, since Zeno of Elea, the author of the famous paradoxes, was recognized by Aristotle as its inventor (Diogenes Laërtius, *Lives* VIII, 57). Aristotle presumably had Zeno's paradoxes in mind, as they are outstanding examples of dialectic, in the sense of refutation of the hypotheses of opponents by drawing unacceptable consequences from those hypotheses. For example, it is unacceptable that Achilles never overtakes the tortoise; therefore, the hypothesis that leads to this conclusion must be rejected. Insofar as this method relies on the law of formal logic known as *modus tollens* (if p implies q, and q is false, then p is false), Zeno was a pioneer of logic, but there is no evidence that he could formulate the law itself; it was left to Aristotle later to state explicitly the principles that underlie this kind of dialectic, and thus to create the science of formal logic.

Dialectic as the use of such indirect logical arguments to defeat an opponent seems to have been used by Zeno for serious philosophical purposes, but it later became, in the hands of the Sophists, a mere instrument for winning a dispute. For example, the Sophist Protagoras claimed that he could "make the worse argument appear the better"; such an aim belongs rather to rhetoric than to logic or philosophy. This degenerate form of dialectic was named "eristic" by Plato (for example, in *Sophist* 231E) and others, from the word *ἔρις* (strife). Eristic came to make deliberate use of invalid argumentation and sophistical tricks, and these were ridiculed by Plato in his dialogue *Euthydemus,* which takes its name from an actual Sophist who appears in it as a user of eristic arguments. Aristotle, too, thought the Sophists worth answering in his book *De Sophisticis Elenchis* (Sophistical refutations), although he sharply distinguished eristic from dialectic, dialectic being for him a respectable activity.

If, however, the lost work of Protagoras did begin, as several subsequent writers attest, with the claim that on every subject two opposite statements (*λόγοι*) could be made, and if the book continued with a content of statement and counterstatement, then Protagoras deserves to be considered the ancestor of the medieval or of the Hegelian dialectic rather than the father of eristic.

Socrates stands in contrast to the Sophists. Unlike them, he professed to be seeking the truth. But he was not above winning the argument, and what is called the *elenchus* was a major element in dialectic as practiced by him, if we are to accept as accurate the presentation of him in Plato's earlier dialogues. The Socratic *elenchus* was perhaps a refined form of the Zenonian paradoxes, a prolonged cross-examination that refutes the opponent's original thesis by getting him to draw from it, by means of a series of questions and answers, a consequence that contradicts it. This is a logically valid procedure, for it corresponds to the logical law "if p implies not-p, then not-p is true (that is, p is false)." Dialectic seems to have been, for Socrates, literally the art of discussion, a search for truth by question and answer; but the definition of a concept is the sort of truth that was typically sought by him, and he supplemented his *elenchus* with another technique, later called epagoge (*ἐπαγωγή*) by Aristotle. This consisted in leading the opponent on to a generalization by getting him to accept the truth of a series of propositions about particular cases. It may now be seen why, in discussing dialectic, Aristotle says "there are two innovations that may justly be ascribed to Socrates: epagogic arguments and universal definition" (*Metaphysics* M 4, 1078b). For Aristotle had a different conception of dialectic, and since *elenchus* goes back to Zeno, the two features he mentions are the only contributions made by

Socrates to dialectic as Aristotle understood it. The Socratic irony, or pretense not to know anything and not to be conducting a refutation, was a personal feature of Socrates' dialectic and contributed nothing to later developments.

PLATO

In the middle dialogues of Plato there occurs a development of the notion of dialectic beyond what we take to be typical of the historical Socrates. Even though Socrates is the protagonist, the views he is portrayed as putting forward are presumably those of Plato. Dialectic is regarded there as the supreme philosophical method, indeed the highest of human arts: it is "the coping-stone, as it were, placed above the sciences" (*Republic* 534E). In the *Cratylus* Plato had described the dialectician as "the man who knows how to ask and answer questions" (390C), and this view of dialectic as question and answer is the Socratic element that forms the single thread running through his altering conceptions of the method. Furthermore, dialectic always had the same subject matter: it sought the unchanging essence of each thing. But the kind of reasoning that Plato regarded as involved in dialectic seems to change: In the middle dialogues it was some kind of operation on hypotheses, whereas in the later ones (for example, *Phaedrus* and *Sophist*) there is, instead, an emphasis on division (*διαίρεσις*) as a method. Division in effect consists of a repeated analysis of genera into species, of more general notions into less general ones, as a way of arriving at a definition when no further division is possible. This process is complemented by the opposite process of synthesis or collection (*συναγωγή*).

Although Plato always spoke of dialectic in an extremely favorable manner, his discussion of it in *Republic* VI–VII marks a high point, as it is there made to be the distinguishing feature in the education of the philosopher-kings and is to be concerned eventually with the supreme Form, that of the Good. It is to reach certainty and overcome the need for hypotheses (*Republic* 511B). But the elevation of the sentiments expressed is matched by suitable vagueness as to the exact process involved, and the interpretation of the few words that are at all precise has been greatly disputed.

It may seem that if dialectic is a process of discussion, then it cannot be of any use for private thought. For Plato, however, there was no difference between the two: "Thought and speech are the same thing, but the silently occurring internal dialogue of the soul with itself has been specially given the name of thought" (*Sophist* 263E; see also *Theaetetus* 189E). However, Plato's most important pupil, Aristotle, was already taking a different view of the nature of thought and hence assigning a merely secondary role to dialectic: "Deception occurs to a greater extent when we are investigating with others than by ourselves, for an investigation with someone else is carried on by means of words, but an investigation in one's own mind is carried on quite as much by means of the thing itself" (*De Sophisticis Elenchis* 169a37). Dialectic was no longer to be the method of science.

ARISTOTLE

The practice of dialectic was probably a major activity in Plato's Academy, to which Aristotle belonged from 367 BCE until Plato's death in 347. Aristotle's *Topics* was apparently intended as an aid to this dialectical debate. It is a handbook for finding arguments to establish or demolish given positions, or *theses,* such as "Every pleasure is good," and while the particular theses used as examples in the *Topics* are no doubt borrowed from the debates in the Academy, the methods provided for dealing with them are completely general, that is, applicable to any thesis of the same form. The *Topics* is therefore the first systematic account of dialectic, and Aristotle indeed boasted that prior to his own treatment of the subject "it did not exist at all" (*De Sophisticis Elenchis* 183b36), and criticized the Sophists for giving teaching that was unsystematic (*ἄτεχνος*). His own trend toward generality and system had the effect that in the *Topics* Aristotle discovered many basic principles of formal logic, including some in the propositional calculus and in the logic of relations, but he hardly reached an explicit formal statement of them. A large part, at least, of this work was written before his discovery of the (categorical) syllogism, a type of argument for which he developed, in his *Analytics,* an elaborate system—the earliest system of formal logic—that superseded dialectic as a theory of demonstration. But even if Aristotle's formal logic developed as an alternative to his dialectic, it may still have arisen out of dialectic in some sense, since it has been argued that he discovered the syllogism as a result of reflection on Plato's method of division.

The distinguishing feature of dialectic for Aristotle was not so much the type of reasoning as the epistemological status of the premises. Reasoning is dialectical if its premises are opinions that are generally accepted by everyone or by the majority or by philosophers; if the premises merely *seem* probable, or if the reasoning is incorrect, then it is "eristic." Aristotelian dialectic is thus quite respectable; it has even been called a "logic of probability," a name that could be misleading because dialec-

tic does not in fact involve inductive reasoning. However, dialectic is not good enough, Aristotle believed, to be a method of acquiring knowledge proper, or science. For that we require demonstration, which is valid reasoning that starts out from true and self-evident premises. The value of dialectic, according to Aristotle, is threefold: It is useful for intellectual training, for discussions with others based on their own premises, and for examining the unprovable first principles of the sciences. "Dialectic, being a process of criticism, contains the path to the principles of all inquiries" (*Topics* 101b3).

STOICS AND MEDIEVALS

Euclides of Megara (a contemporary of Plato) and his successors in that town were logicians of note, and the Megarian tradition in logic was continued by the Stoics. The Stoic logic was known as dialectic, perhaps because the initiators of their tradition had an interest in the Zenonian paradoxes and related reasoning. Under the headship of Chrysippus, who lived from 280 to 206 BCE, the Stoic school reached its zenith, and it was still going strong four centuries later. A saying is recorded from this period, that "if the gods had dialectic, it would be the dialectic of Chrysippus" (Diogenes Laërtius, *Lives* VII, 180). By "dialectic" the Stoics primarily meant formal logic, in which they particularly developed forms of inference belonging to what we now call the propositional calculus. But they applied the term *dialectic* widely: for them it also included the study of grammatical theory and the consideration of meaning-relations and truth. This widened scope, reflecting the special interests of the early Stoics, remained typical of the school; it was accepted by Cicero and perhaps overemphasized by Seneca, who wrote that dialectic "fell into two parts, meanings and words, that is, things said and expressions by which they are said"—*διαλεκτική in duas partes dividitur, in verba et significationes, id est in res quae dicuntur et vocabula quibus dicuntur* (*Epistulae Morales* 89, 17).

In the Middle Ages "dialectic" continued to be the ordinary name for logic: for example, the first medieval logical treatise was the *Dialectica* of Alcuin. But the word *logica* was also used; in fact, Abelard wrote a *Dialectica* and more than one *Logica.* As the works of Plato and Aristotle became known, the Scholastics took over various conceptions of dialectic, and the medieval disputation, by which university degree examinations were conducted, can be regarded as a remote descendant or revival of the debates in the Platonic Academy. The disputants maintained theses and antitheses, arguing mainly in syllogisms; the most significant difference from ancient practice was that the class of unacceptable consequences now included those propositions that were inconsistent with divine revelation.

KANT AND HIS SUCCESSORS

In his *Critique of Pure Reason* (A61, B85) Immanuel Kant asserted rather sweepingly that the actual employment of dialectic among the ancients was always as "the logic of illusion (*Logik des Scheins*)." He explained that he applied the term to logic as a critique of dialectical illusion. He titled the second division of his Transcendental Logic "Transcendental Dialectic." This new kind of dialectic was concerned with exposing the illusion of transcendental judgments, that is, judgments that profess to pass beyond the limits of experience; but the illusion can never, he thought, be dispelled entirely, as it is natural and inevitable.

Although Kant, in his Transcendental Dialectic, had set out the antinomies of pure reason as four sets of thesis and antithesis, he did not call his resolution of the antinomies a synthesis. It was his successor Johann Gottlieb Fichte who, in his *Grundlage der gesamten Wissenschaftslehre* (Jena and Leipzig, 1794), first introduced into German philosophy the famed triad of thesis, antithesis, and synthesis. In this he was followed by Friedrich Schelling, but not in fact by G. W. F. Hegel. Fichte did not believe that the antithesis could be deduced from the thesis; nor, on his view did the synthesis achieve anything more than uniting what both thesis and antithesis had established.

HEGEL AND HIS SUCCESSORS

Hegel is commonly supposed to have presented his doctrines in the form of the triad or three-step (*Dreischritt*) of thesis, antithesis, and synthesis. This view appears to be mistaken insofar as he did not actually use the terms; and even though he evinced a fondness for triads, neither his dialectic in general nor particular portions of his work can be reduced simply to a triadic pattern of thesis, antithesis, and synthesis. The legend of this triad in Hegel has been bolstered by some English translations that introduce the word *antithesis* where it is not required.

However, there is indeed a Hegelian dialectic, involving the passing over of thoughts or concepts into their opposites and the achievement of a higher unity. But if it is a process that arrives at a higher truth through contradictions, it does not constitute a new conception of dialectic. Hegel actually showed his awareness of the traditional notion by paying tribute to "Plato's *Parmenides,* probably the greatest masterpiece of ancient dialectic."

And even the doctrine that dialectic is a world process—not merely a process of thought but also found in history and in the universe as a whole—was not wholly new, but goes back to Heraclitus and the Neoplatonist Proclus. Here again Hegel, with his interest in the history of philosophy, was aware of his predecessors. What seems to be genuinely new in Hegel's view of dialectic is the conception of a necessary movement. Dialectic was said to be "the scientific application of the regularity found in the nature of thought." The "passing over into the opposite" was seen as a natural consequence of the limited or finite nature of a concept or thing. The contradictions in thought, nature, and society, even though they are not contradictions in formal logic but conceptual inadequacies, were regarded by Hegel as leading, by a kind of necessity, to a further phase of development.

Hegel has had an enormous influence not only on willing disciples but even on thinkers nominally in revolt against him, such as Søren Kierkegaard. One of the most important offshoots of the Hegelian dialectic was the Marxist dialectic, in which, of course, "matter" was substituted for Hegel's "spirit."

See also Abelard, Peter; Aristotle; Chrysippus; Cicero, Marcus Tullius; Dialectical Materialism; Diogenes Laertius; Fichte, Johann Gottlieb; Greek Academy; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heraclitus of Ephesus; Infinity in Mathematics and Logic; Kant, Immanuel; Kierkegaard, Søren Aabye; Marxist Philosophy; Medieval Philosophy; Neoplatonism; Plato; Proclus; Protagoras of Abdera; Schelling, Friedrich Wilhelm Joseph von; Socrates; Sophists; Stoicism; Zeno of Elea.

Bibliography

GENERAL DISCUSSIONS

For general discussions of dialectic, see Paul Foulquié, *La dialectique* (Paris, 1949); Eduard von Hartmann, *Über die dialektische Methode* (Berlin, 1868; 2nd ed. 1910); Karl Dürr, "Die Entwicklung der Dialektik von Plato bis Hegel." *Dialectica* 1 (1947); and Jonas Cohn, *Theorie der Dialektik* (Leipzig, 1923).

SOCRATES AND HIS PREDECESSORS

On Zeno, see H. D. P. Lee, *Zeno of Elea* (Cambridge, U.K., 1936); G. E. L. Owen, "Zeno and the Mathematicians," in *PAS* 58 (1957–1958): 199–222; and John Burnet, *Early Greek Philosophy* (London, 1892; 4th ed., 1930). For the period before Socrates as a whole, see Burnet's *From Thales to Plato* (London, 1914). For criticisms of the Sophists, see Plato, *Euthydemus,* and Aristotle, *De Sophisticis Elenchis,* translated by W. A. Pickard-Cambridge in *The Works of Aristotle,* Vol. 1 (Oxford, 1928). On this whole phase of dialectic, see the extremely reliable work of William and Martha Kneale, *The Development of Logic* (Oxford: Clarendon Press, 1962), Ch. 1.

PLATO

The most helpful book on Plato's use of dialectic is Richard Robinson, *Plato's Earlier Dialectic* (Ithaca, NY: Cornell University Press, 1941; 2nd ed., Oxford: Clarendon Press, 1953), but the word *earlier* occurs in its title because it contains no examination of the methods of division and synthesis that appear in some of Plato's later dialogues. For the original descriptions of the Platonic dialectic, see the Dialogues of Plato, *passim* (in any translation), but above all the *Republic,* Books VI–VII, especially 510–540 in the standard numbering. See also James Adam, *The Republic of Plato,* Vol. II (Cambridge, U.K., 1902; reissued, 1963), pp. 168–179; Richard Lewis Nettleship, *Lectures on the Republic of Plato* (London: Macmillan, 1901), pp. 277–289; Julius Stenzel, *Studien zur Entwicklung der Platonischen Dialektik von Sokrates zu Aristoteles* (Breslau, 1917; 2nd ed., 1931), translated by D. J. Allan as *Plato's Method of Dialectic* (Oxford: Clarendon Press, 1940; reissued, New York: Russell and Russell, 1964); Francis M. Cornford, "Mathematics and Dialectic in the *Republic* VI–VII," in *Mind* 41 (1932): 37–52, 173–190; R. S. Bluck, "*ὑποθέσεις* in the *Phaedo* and the Platonic Dialectic," in *Phronesis* 2 (1) (1957): 21–31; D. W. Hamlyn, "The Communion of Forms and the Development of Plato's Logic," in *Philosophical Quarterly* 5 (1955): 289–302; Georges Rodier, "Les mathématiques et la dialectique dans le système de Platon," in *Archiv für Geschichte der Philosophie* 15 (1902): 479–490, and his "L'évolution de la dialectique de Platon," in *Année philosophique* (1905): 49–73; D. S. Mackay, "The Problem of Individuality in Plato's Dialectic," in *University of California Publications in Philosophy* 20 (1937): 131–154; and Juan A. Nuño Montes, *La dialéctica platónica* (Caracas, 1962).

ARISTOTLE

For the original account, see Aristotle, *Topica,* translated by W. A. Pickard-Cambridge in *The Works of Aristotle,* Vol. I (Oxford, 1928). The most accessible scholarly discussion of this phase is Ernst Kapp, *Greek Foundations of Traditional Logic* (New York: Columbia Press, 1942). Important treatments of the controversial issues involved are Friedrich Solmsen, *Die Entwicklung der Aristotelischen Logik und Rhetorik* (Berlin, 1929), his "The Discovery of the Syllogism," in *Philosophical Review* 50 (1941): 410–421, and his "Aristotle's Syllogism and Its Platonic Background," in *Philosophical Review* 60 (1951): 563–571; and Paul Wilpert, "Aristoteles und die Dialektik," in *Kant-Studien* 68 (1956): 247–284. There is also Livio Sichirollo, *Giustificazioni della dialettica in Aristotele* (Urbino, 1963).

STOICS AND MEDIEVALS

For a discussion of Stoic use of dialectic, see Benson Mates, *Stoic Logic* (Berkeley: University of California Press, 1953). For a general account of the medieval phase and references to particular works, see William and Martha Kneale, op. cit., Ch. 4. For the postmedieval rejection of dialectic, see Duane H. Berquist, "Descartes and Dialectics," in *Laval théologique et philosophique* 20 (2) (1964): 176–204.

KANT AND HIS SUCCESSORS

Kant and his successors are discussed in Graham H. Bird, *Kant's Theory of Knowledge* (London: Routledge and Paul, 1962); John E. Llewelyn, "Dialectical and Analytical Opposites," in *Kant-Studien* 55 (1964): 171–174; and Richard Kroner, *Von Kant bis Hegel,* 2 vols. (Tübingen, 1921–1924).

HEGEL AND HIS SUCCESSORS

For general discussions of the Hegelian phase, see K. R. Popper, "What Is Dialectic?," in *Mind* 49 (1940): 403–426, reprinted in Popper's *Conjectures and Refutations* (London: Routledge, 1963); Sidney Hook, "What Is Dialectic?," in *Journal of Philosophy* 26 (1929): 85–99, 113–123, his *From Hegel to Marx* (London, 1936), and his "Dialectic in Social and Historical Inquiry," in *Journal of Philosophy* 36 (1939): 365–378; and Siegfried Marck, *Die Dialektik in der Philosophie der Gegenwart,* 2 vols. (Tübingen, 1929–1931). On dialectic in Hegel, see John M. E. McTaggart, *Studies in the Hegelian Dialectic* (Cambridge, U.K., 1896); G. R. G. Mure, *An Introduction to Hegel* (Oxford: Clarendon Press, 1940); and above all, John N. Findlay, "Some Merits of Hegelianism," in *PAS* 56 (1955–1956): 1–24, and his *Hegel: A Re-examination* (London: Allen and Unwin, 1958), and a valuable chapter by him on Hegel in *A Critical History of Western Philosophy,* edited by D. J. O'Connor (London, 1964). See also Gustav E. Mueller, "The Hegel Legend of Thesis-Antithesis-Synthesis," in *Journal of the History of Ideas* 19 (1958): 411–414; Carl J. Friedrich, "The Power of Negation: Hegel's Dialectic and Totalitarian Ideology," in *A Hegel Symposium,* edited by D. C. Travis (Austin: University of Texas, 1962), pp. 13–35; and Walter Kaufmann, *Hegel* (New York, 1965). On Marx, see Harold B. Acton, *The Illusion of the Epoch: Marxism–Leninism as a Philosophical Creed* (London: Cohen and West, 1955).

Roland Hall (1967)

DIALECTICAL MATERIALISM

Marxism-Leninism is the name given to the form of Marxist theory that was accepted and taught by the Russian and Chinese Communist parties and the Communist parties associated with them. Marxism-Leninism is both a view of the world as a whole and of human society and its development. The view of human society is called historical materialism, the name bestowed upon it by Friedrich Engels. The view of the world as a whole is called dialectical materialism, a title devised by G. V. Plekhanov, the Russian Marxist, and first used by him in an article published in 1891. Marxist-Leninists regard dialectical materialism as the basis of their philosophy and generally begin comprehensive expositions of that philosophy with an account of it. One might say that dialectical materialism constitutes the logic, ontology, and epistemology of Marxism-Leninism, and historical materialism its ethics, politics, and philosophy of history. Sometimes, however, the term *dialectical materialism* is used for the fundamentals of Marxism-Leninism as a whole. When dialectical materialism is thus conceived, the natural sciences are the working-out of dialectical materialism in the nonhuman sphere and historical materialism its working-out in the sphere of human society. But these slight differences do not affect the content of the theory.

MARX'S MATERIALISM

Approving references to materialism are prominent in Karl Marx's writings, especially in the early works. In *The Holy Family* (1845), for instance, he argued that one branch of eighteenth-century French materialism developed into natural science and the other branch into socialism and communism. Thus he regarded "the new materialism," as he called it, as a source of the social movement that he believed was destined to revolutionize human life. Materialism, as Marx understood it, was very closely connected with social criticism and social development. One aspect of materialism that Marx supported was its rejection of idealist attempts to undermine and belittle sense experience. He held that there is something dishonest and irresponsible in philosophies which deny that sense experience reveals the existence of an independent material world; hence his view of knowledge was realist, both on philosophical and moral grounds. In taking this view he was much influenced by Ludwig Feuerbach. Like Feuerbach, Marx rejected speculative philosophy, or metaphysics, as we should call it today, on the ground that the truth about the world and society can only be discovered by the use of empirical scientific methods. In a broad sense of the term, therefore, Marx was a positivist, in that he denied the possibility of any knowledge of the world that is not based on sense experience. Hence, Marx's view of the world was naturalistic and opposed to any form of religion or supernaturalism. Again under the influence of Feuerbach, Marx held that belief in God, in an afterlife, and in heaven and hell cannot be rationally justified, but may be explained (indeed, explained away) in terms of the unfulfilled needs and hopes of men whose lives are frustrated by an oppressive social order. Marx held, too, that men are not immaterial souls conjoined with material bodies. In his view, psychophysical dualism is a relic of supernaturalism and must be rejected with it. Marx did not systematically develop this view as part of a philosophical argument but took it as the basis of his view, expressed in *The Holy*

Family and in *The German Ideology* (1845–1846), that repression of the instincts and natural desires is bad. Marx, therefore, thought that thinking is inseparable from acting and that scientific advance and practical improvement are in principle bound up with one another. Marx's materialism, therefore, is very wide in scope, combining empiricism, realism, belief in the use of scientific methods pragmatically conceived, rejection of supernaturalism, and rejection of mind-body dualism. Animating these aspects of his view is the conviction that they support and justify the socialist diagnosis of social ills and the prediction that a communist form of society must come.

Marx was very much influenced by the philosophy of G. W. F. Hegel. For example, in *The Holy Family* he borrowed almost verbatim some arguments from Hegel's *Encyclopedia* against abstract and unrealistic thinking, and his earliest, unfinished sketch of his theory of man and society, the so-called *Economic and Philosophical Manuscripts* (1844), was both a critique of political economy and a critique of the philosophy of Hegel. Marx's interest in Hegel continued throughout his life. In a letter to Engels in 1858 Marx wrote that he had been looking at Hegel's *Logic* and would like, if he had time, to write a short work setting out what was wrong and what was valuable in Hegel's method. Later, in the Preface to the second edition of Volume I of *Capital,* Marx referred to "the rational kernel" of Hegel's dialectical method and said that in *Capital* he had "toyed with the use of Hegelian terminology when discussing the theory of value." This sentence does not indicate a very strong attachment to Hegel's dialectic, for "toyed with" (*kokettierte sogar hier und da*) is appropriate to a superficial liaison, and the word *terminology* (*Ausdrücksweise*) might be meant to contrast with the substance of what is being said. But although Marx was as much opposed to the speculative element in Hegelianism as any professed positivist could have been, he was deeply influenced by the Hegelian dialectical method. Jean Hippolyte has shown in his *Études sur Marx et Hegel* (Paris, 1955) how very closely the structure of *Capital* is linked with Marx's earlier, more consciously Hegelian writings, so that some of the Hegelian substance persists, although the Hegelian terminology is less apparent. One important Hegelian legacy is the view that social development takes place through struggle and opposition. Another is that the transition from one important form of society to another is by means of sudden leaps rather than by merely gradual stages. Thus Marx considered that different social laws applied at different historical epochs. Again, Marx shared Hegel's aversion to abstraction and his predilection for total views, but in this he was at one with Auguste Comte as well as with Hegel. These views of Marx's, however, related to the theory of human society. He showed little inclination to linger over questions of ontology. There is a reference in Volume I of *Capital* to the "law discovered by Hegel in his *Logic,* that at a certain point what have been purely quantitative changes become qualitative," and at this point Marx said that some chemical changes take place in accordance with this law. However, Marx left it to Engels to pursue the matter.

ENGELS AND DIALECTICAL MATERIALISM

Engels took up the law of quantity and quality in his *Herr Eugen Dühring's Revolution in Science* (1878), generally known as *Anti-Dühring,* which had appeared as a series of articles in the Leipzig *Vorwärts* in 1877. Engels's work was directed against Eugen Dühring, a well-known non-Marxist socialist and publicist, who had vigorously criticized some Hegelian features in Marxist writers as being speculative, metaphysical, and unscientific. Thus Engels, like Marx, felt called upon to defend the Hegelianism of his youth, although, again like Marx, he claimed to have purged it of its speculative and idealist elements. In the Preface to the second edition of *Anti-Dühring* (1885) Engels stated that he had read the whole of the manuscript to Marx before it was printed and that Chapter 10 of Part II (on economics and its history) had been written by Marx himself and abridged by Engels. This chapter has no direct relevance to dialectical materialism and thus has some significance as an indication of Marx's own interests.

PHILOSOPHY OF NATURE. Engels apologized in a general way in the preface to the second edition of *Anti-Dühring* for inadequacies in his knowledge of theoretical natural science, although he retracted nothing. He also spoke with approval of "the old philosophy of nature." By this he meant a philosophical examination of the phenomena of the natural world claiming to be more fundamental and general in scope than the particular researches of individual men of science. Such inquiries were more frequent at a time when the term *philosopher* was applied to philosophers and scientists alike and the role of the natural scientist was less definitely specified than it became in the nineteenth and twentieth centuries. Engels alluded to Hegel's contributions to the philosophy of nature in the second main triad of the *Encyclopedia* and called attention in particular to Section 270 in which Hegel criticized Isaac Newton's theory of forces. Hegel,

like Johann Wolfgang von Goethe and Friedrich Schelling (and William Blake), was highly critical of Newton's cosmological theories, and Engels believed that Hegel, at any rate, was being justified by subsequent researches. It should be noted, therefore, that Engels had no objection to the practice of philosophizing about the nature of the physical world but, on the contrary, was consciously reviving an older, and apparently abandoned, intellectual tradition. By doing this, he introduced into the Marxist theory of nature one of its most characteristic features: the claim that the specialized sciences of nature need to be supplemented by a unified philosophy of nature and that as they develop, the natural sciences are constantly verifying the views first propounded by Hegel in his *Logic* and in his *Encyclopedia.*

From 1873 onward Engels had been studying the natural sciences with a view to writing a comprehensive work on the dialectical characteristics of the material world. Part of what he did was incorporated into *Anti-Dühring,* but much of his more detailed work remained unpublished until 1925, when an edition of the surviving manuscripts was published by the Marx-Engels Institute in Moscow under the title *Dialectics of Nature.* This edition was found to be faulty in various ways, and corrected versions were subsequently published and translated. The work contains, inter alia, an essay on electricity (a subject much favored by Schelling and other romantics), in which Engels says that the basic thought of Hegel and Michael Faraday is the same; an attack on parapsychology as "the shallowest empiricism" and a proposal that it be rejected outright on general grounds of theory; notes on infinite series and infinite numbers, which he takes to prove that the world is both infinite and contradictory; and sketches for an attack on Ludwig Büchner and other nonsocialist, nondialectical materialists popular during the second half of the nineteenth century. Engels's criticism of Büchner is particularly interesting since, among a series of passages probably intended to document *Anti-Dühring,* there is a quotation from Büchner's *Kraft und Stoff* in which, while attacking supernaturalism and idealist philosophy, Büchner wrote: "It is needless to observe that our expositions have nothing in common with the conceptions of the old 'philosophy of nature.' The singular attempts to construe nature out of philosophy instead of from observation have failed, and brought the adherents of that school into such discredit that the name 'philosopher of nature' has become a bye-word and a nickname." Engels regarded this as an "attack on philosophy" and accused Büchner of "shallow materialist popularisation." Engels made his own attitude quite clear by appending passages from Hegel's "Philosophy of Nature."

ENGELS ON MARXIST MATERIALISM. After Marx's death in 1883 Engels was occupied in editing the unpublished parts of *Capital,* but in 1886, in some articles that appeared in the Social Democratic journal *Die Neue Zeit,* he turned his attention once more to fundamental philosophical issues. These articles were published in 1888 in book form under the title *Ludwig Feuerbach and the Outcome of Classical German Philosophy.* In this work Engels set out to explain what sort of materialism Marxist materialism is and to show how it is related to the Hegelian philosophy. Engels renewed his support for the dialectical structure of Hegel's philosophy, although, of course, he rejected its idealist aspects. There is an account of Engels's epistemology, in which a pragmatistic point of view is emphasized.

Mind and matter. According to the argument of *Ludwig Feuerbach* there are two and only two fundamental but opposing philosophical alternatives: idealism, according to which mind is primary in the universe and matter is created by, or dependent upon, mind; and materialism, according to which matter is the primary being and mind the subordinate and dependent feature of the world. It will be seen that in stating this view Engels extended the term *idealism* beyond its usual philosophical meaning to comprise not only such views as George Berkeley's immaterialism and Hegel's absolute idealism but also any form of theism. Thus, in Engels's classification, St. Thomas Aquinas and René Descartes would both be regarded as idealists because they both held that an immaterial deity created the material world. It should be noted that in this view mind is held to be secondary but not nonexistent. Engels took the widely held natural-scientific point of view that there was once a time when only matter existed and that mind evolved from it and must remain dependent upon it. He did not hold the theory of reductive materialism, according to which mind is just a form of matter.

Knowledge and perception. In *Ludwig Feuerbach* Engels also gave a brief account of knowledge and sense perception. He considered that in sense perception the material things in the neighborhood of the percipient's body are somehow "reflected" in his brain "as feelings, instinct, thoughts, volitions." Engels recognized that the theory that in perception the immediate object of awareness is a "reflection" could lead to agnosticism or idealism, for a skeptic could question whether we can ever know of the existence of material things at all if all that we directly apprehend are reflections of them. This, indeed, is a line of thought that Berkeley developed in criticizing John Locke's theory that it is ideas, not physical things, that are directly apprehended. Engels's answer was that

what must dispel any such doubts is "practice, viz. experiment and industry." His discussion is vague, but he appears to have thought that skeptical doubts about the existence of material things are rendered untenable by a consideration of what we do to and with things. A skeptic's or idealist's practice belies his theories. Furthermore, Engels held that the truth of scientific theories about the material world is established by the power they give men to manufacture new substances and things and to bring the forces of nature under human control. "If we are able to prove the correctness of our conception of a natural process by making it ourselves, bringing it into being out of its conditions and using it for our own purposes into the bargain, then there is an end of the Kantian incomprehensible 'thing-in-itself'" (*Ludwig Feuerbach,* pp. 32–33). Engels appears to have conflated the problem of our perception of the external world with the problem of how scientific laws are established, but it is clear that he believed that the notion of practice can help to solve them both. In the Preface to *Ludwig Feuerbach* Engels printed for the first time, under the title *Theses on Feuerbach,* some jottings made by Marx in 1845. The doctrine of the philosophical importance of practice is stated in these theses, particularly in the first, second, fifth, and eleventh. One of the things that Marx appears to have been asserting in them is that perception is a deed or activity of the perceiving corporeal man and not merely a passivity of an immaterial mind. In 1892, in the introduction to some chapters from *Anti-Dühring* published separately under the title *Socialism: Utopian and Scientific,* Engels developed this view, arguing that perception is a more or less successful action on the world.

Attack on "vulgar materialists." Another feature of Engels's materialism is its opposition to the theories of those whom he called in *Ludwig Feuerbach* "vulgarising pedlars," and who, in later Marxist philosophy, are called "vulgar materialists." These were a group of German writers and lecturers, of whom Büchner was one, who argued that materialism was the inevitable consequence of natural science in general and of physiology in particular. Engels objected that they wasted too much time arguing that God does not exist. He also objected that they identified thought with brain processes. Furthermore, they failed to recognize the social, indeed the socialist, implications of materialism. But primarily he objected that theirs was a mechanical materialism. A consideration of this objection brings us to a central feature of Engels's dialectical materialism.

By mechanical materialism Engels meant the type of materialism current in the eighteenth century, when the most highly developed natural science was mechanics. According to this view, all the most complex phenomena of nature, including life and mind, can be reduced to the arrangement and rearrangement of material particles. The most complex beings can be nothing but arrangements of the ultimate simple ones, so that chemical combination, life, mind, and thought are no more than increasingly elaborate applications of mechanical principles. According to Engels, in saying that everything is reducible to the interaction of forces, the vulgar materialists were anachronistically upholding this eighteenth-century view, whereas the natural sciences of the nineteenth century, in developing chemistry and biology, went beyond those of the eighteenth century. In merely mechanical mixtures the original components remain side by side with each other, but in chemical combinations new substances result from the joining of their ingredients. The theory of biological evolution showed that new forms of life have emerged from the simpler forms, not merely more complex ones.

Mechanical materialism itself is a form of what Engels, following Hegel, called the "metaphysical" attitude of thought. Engels's source in Hegel is the phrase "the former metaphysics," by which Hegel referred to the philosophical method used by Christian Wolff and others in the eighteenth century in trying to prove important truths about the world and the human soul by the use of definitions and axioms and allegedly strict deductions. Engels agreed with Hegel that this quasi-mathematical method was inappropriate in philosophy and added that it was inappropriate in science too. In *Anti-Dühring* Engels said that in the metaphysical mode of thinking, "things and their mental images, ideas" are regarded as isolated and fixed; things either exist or do not exist; and positive and negative exclude one another. But this, he held, is to overlook the changefulness and interconnections of things. Collecting distinct items of information and neglecting the aspect of process helped natural science to get started but was only a preliminary stage toward grasping the world in all its interconnections, processes, beginnings and endings, and contradictions. Mechanical materialism is a fruit of metaphysical thinking. Metaphysical thinking was, in the Hegelian philosophy, and then in the writings of Marx, superseded by dialectical thinking; and this was, in Engels's view, another way of saying that mechanical materialism must be superseded by dialectical materialism. Engels believed that nineteenth-century biology and chemistry had developed along lines that Hegel had foreseen and required. In particular, he referred to passages in Hegel's *Logic* and *Encyclopedia* according to which a fuller under-

standing is gained when the category of mechanism is left behind and replaced by the higher categories of life. In Hegel's "Philosophy of Nature," to which Engels's *Dialectics of Nature* so often refers, the mechanical forms are succeeded by physical ones that include "chemical process" and electrical phenomena, and these by "the organic." It is this sequence that provided the framework for Engels's philosophy of nature.

Engels on dialectics. Since dialectical thinking is, in Engels's view, opposed to metaphysical thinking, it is thinking that attempts to grasp things in their interrelationships and in the totality to which they belong, in the process of change, of being born and of dying, in their conflicts and contradictions. Furthermore, it is thinking that recognizes the emergence of novelty and that sees such emergences as sudden, even catastrophic. Dialectical thinking, he also held, was becoming more and more apparent as the natural sciences progressed. Scientific discoverers were dialecticians without knowing it.

CONTRADICTIONS IN NATURE. In *Anti-Dühring* Engels expounded his dialectical philosophy of nature in some detail. Dühring had criticized the Hegelian elements of Marx's thought. In particular he had argued that contradiction is a logical relationship and that it is absurd to suppose that it can be a relationship between things or events in the natural world. In Part I, Chapter 12 of *Anti-Dühring* Engels endeavored to defend the dialectical theory against this objection. First, he said that the view that there could be no contradictions in nature rests upon the assumption of "the former metaphysics" that things are "static and lifeless." Then he argued that when we consider things in movement and in their effects upon one another, the dialectical view has to be adopted. "Movement itself," he wrote, "is a contradiction: even simple mechanical change of place can only come about through a body at one and the same moment of time being both in one place and in another place, being in one and the same place and also not in it. And the continuous assertion and simultaneous solution of this contradiction is precisely what motion is." Engels also maintained that what is true of mechanical change of place is "even more true of the higher forms of motion of matter, and especially of organic life and its development." Engels had argued in Part I, Chapter 8 that in absorbing and excreting nutriment living matter at each moment is "itself and at the same time something else." Engels also held that there are real contradictions in "higher mathematics," where straight lines and curves may be identical. (He probably had in mind Section 119 of Hegel's *Encyclopedia.*) Similarly, Engels said that the square root of minus one is not only a contradiction but "a real absurdity."

Engels's claim that movement is in itself contradictory is based on a passage from Hegel's *Science of Logic* in which it is argued that it is not sufficient, if something is to move, for it to be *here-now* and then, after that, *there-then,* for this would merely be for it to be at rest first in the one place and then in the other. For it to move, Hegel concluded, a body must be "here and not here in the same now" and must "be and yet not be in the same here" (*Science of Logic,* Book II, Sec. 1, Ch. 2, C). Hegel was discussing Zeno, who had argued that since movement is contradictory, what is real cannot move. Hegel in this passage accepted Zeno's arguments that movement is contradictory, but unlike Zeno concluded that since there is movement, movement "is an existing contradiction." Hegel's views on contradiction are difficult to understand and have been interpreted in various ways. If intended to argue that contradictory propositions could both be true, that "both *p* and not-*p,*" then he was wrong and so was Engels in following him. For it can be proved that from any pair of contradictory propositions any conclusion we like can be deduced and hence that if contradictories are true, *anything* can be true. In this logical sense the term *contradiction* has its appropriate use in thought or discourse, as Dühring had argued. In saying that something both is and is not in the same place at the same time, that it is true both that it is in *P* at time *t* and that it is not in *P* at time *t,* the whole negating force of the word *not* is lost. Either, then, Hegel's philosophy has no value or he must have meant by "contradiction" something different from what formal logicians mean by it. It is likely enough that it is the second alternative that is correct. In attacking Dühring, Engels seems to have committed himself to the first alternative. He adopted a speculative, nonempirical thesis, for whereas movement is something that can be observed in natural things and events, contradiction is not observable in them. What Engels did in his argument about contradiction in the nature of things was to provide one of Zeno's paradoxes with a merely verbal, and indeed absurd, "solution."

It appears that Engels's doctrine on this matter is now being reinterpreted or abandoned. This process began with an article on Zeno's paradoxes by the famous Polish logician Casimir Ajdukiewicz. When this article appeared in Poland in 1948, dialectical materialists were forced to take account of his arguments. In order to do so they granted that "contradiction" does not mean "logical contradiction" when applied to what exists in nature. This view is adopted by the Russian authors of *The Fun-*

damentals of Marxism-Leninism: Manual (English translation, Moscow, no date, but later than 1960), who write: "Contradictions due to incorrect thinking should not be confused with objective contradictions existing in objective things. Although the word 'contradiction' is the same in both cases, it means different things" (pp. 99–100).

QUANTITY AND QUALITY. Another dialectical law of nature that Engels made much of in his *Anti-Dühring* is that according to which certain of the changes in nature take place suddenly and abruptly rather than by gradual accretion. The simplest instances of this sort of change are the changes of water into ice as its temperature is lowered to the freezing point and into steam as its temperature is raised to the boiling point. The ice and steam do not come into existence gradually and pari passu with the gradual lowering and raising of the temperature, but appear all at once as soon as the freezing or boiling point has been reached. Other examples of the principle were given by Engels: the sudden transformation of one chemical substance into another in the course of chemical combination; the melting points of metals; the transformation of mechanical motion into heat; the necessity for a sum of money to exceed a certain amount before it can become capital; the fact, reported by Napoleon, that whereas two Mamelukes were more than a match for three Frenchmen, a thousand Frenchmen were more than a match for fifteen hundred Mamelukes. One very general idea in all this is that gradual alterations in the quantity of something are not necessarily accompanied by a merely gradual alteration in its characteristics. Apart from this, Engels had in mind the evolutionary scheme of development from simpler forms of matter, like gases, to more distinctive and varied forms, like the many kinds of solids, plants, and animals. This development is not a mere rearrangement of otherwise unchanging particles or elements but is the emergence of new features out of the old, even though the later qualities could not have emerged unless the earlier and simpler ones had first existed. The emerged qualities, however, are not reducible to those from which they have emerged. The point at which changes in a single quality transform it into a new one Engels called a "nodal line." He also said that there is a "leap" from one quality to another.

Once again Engels was following Hegel very closely. The account in *Anti-Dühring* is based upon Sec. 108 of the *Encyclopedia* and Book I, Division 3, Chapter 2, *B* of the *Science of Logic,* where Hegel discussed the category of "measure." In these passages Hegel tried to show the part played by proportion in the constitution of things. He gave the examples of water turning, at critical points or nodal lines, into ice or steam, and of chemical combinations and constant proportions, which Engels and Marx repeated later. He also instanced birth and death, the acquisition of new properties by numbers as the series of natural numbers develops, and the acquisition of new features by the notes of a musical scale. He gave a moral example, based on Aristotle, of slight changes that turn virtues into vices, carelessness into crime, and so on. He even gave a political example, borrowed from Baron de Montesquieu, of the relation of a type of constitution to the population of a state. In the *Encyclopedia* Hegel also referred to the ancient Greek puzzles about the point at which a man becomes bald or at which a number of grains of wheat become a heap. Interesting as these examples are, they are extremely disparate. The grains of wheat example is partly a question of how many grains we shall *call* a heap, and this is to some extent a matter of decision. The concepts of a heap or of baldness are rather vague. The examples of a series of gradual physical changes succeeded by a total transformation of quality are clearly of interest to Engels because of the analogy to revolutionary social change by contrast with gradual alteration. Undoubtedly the social examples had impressed Hegel, who had called attention to the gradual steps that lead up to an explosive revolutionary break in the Preface to the *Phenomenology of Mind,* where he wrote: "This gradual disintegration which did not alter the general look and aspect of the whole is interrupted by the sunrise which, in a single flash, brings to view the form and structure of the new world."

In itself, whether there are or are not nodal lines and constant proportions in the physical world would seem to have no logical connection with the way in which the social order changes, unless, indeed, it is held that human society really is, or is reducible to, physical events—and this is in conflict with Engels's general rejection of reductive materialism. If, then, this law is not an expression of a view that is inconsistent with Engel's main view, it would seem to serve an almost animistic purpose. Sudden revolutionary change, he seems to be suggesting, is a fundamental character of the universe as a whole, so that when we urge revolution, we have the universe behind us. That the view at any rate serves this purpose may be seen from Joseph Stalin's subsequent impatience with it. When socialism is established, it is natural for the socialist leaders not to wish to think in terms of their own disappearance and of the emergence of still further social revolutions. Hence, Stalin, in his famous article on linguistics, wrote scornfully of "comrades who have an infatuation for explosions."

INTERPENETRATION OF OPPOSITES. In addition to the law of transformation of quantity into quality, Engels mentioned two other laws of dialectics, the law of the interpenetration of opposites and the law of the negation of the negation. The first of these laws was already touched upon in the exposition of the theory of contradictions in nature and of the deficiencies of the metaphysical point of view. Although Engels mentioned it in the *Dialectics of Nature,* he did not discuss it as such, and in *Anti-Dühring* his emphasis was on the other two laws, to each of which he devoted a chapter. The law of the interpenetration of opposites (which was later called the law of the unity and struggle of opposites) seems to have been intended to provide an explanation of why there is any change or development at all. An idea behind it is that in the absence of all tension everything would remain exactly as it is, since there would be nothing to provoke any change. Change takes place because the world does not consist of isolated, self-sufficient, independent particulars, but of opposing forces overcoming or being overcome. Contradiction, or opposition, is in this view the motive force both of natural and of human history.

NEGATION OF THE NEGATION. The law of the negation of the negation was more specifically emphasized by Engels. He was able to quote from a passage in Marx's *Capital* in which it is said that when, as a result of competition between capitalists, the few remaining giant capitalist enterprises find themselves confronted by a poverty-stricken proletariat, the latter will rise and expropriate the former, the expropriators will be expropriated. "Capitalist production," wrote Marx, "begets, with the inexorability of a Law of Nature, its own negation. It is the negation of the negation." According to Engels in *Anti-Dühring,* the law of the negation of the negation is "an extremely general—and for this reason extremely comprehensive and important—law of development of nature, history and thought, a law which … holds good in the animal and plant kingdoms, in geology, in mathematics, in history and in philosophy." The law is illustrated, according to Engels, by every case in which a plant has seeds that germinate and result in the growth of further plants. "But what is the normal life-process of this plant? It grows, flowers, is fertilised and finally once more produces grains of barley, and as soon as these have ripened the stalk dies, is in its turn negated. As a result of this negation of the negation we have once again the original grain of barley, but not as a single unit, but ten, twenty or thirty fold" (*Anti-Dühring,* p. 152). One idea in this very famous passage is that out of what looks like death and destruction there arises something better and more various. (Engels in fact wrote of "qualitatively better seeds which produce more beautiful flowers.")

In his early book *The Poverty of Philosophy* (1847) Marx had quoted the Latin phrase *mors immortalis,* that is, "deathless death," and Engels similarly regarded progress as taking place through continual destruction and amplified renewal. What holds for plants obviously holds for animals. Geology illustrates the law, too, for it describes "a series of negated negations, a series arising from the successive shattering of old and depositing of new rock formations." The same law appears in mathematics. A is negated by $-A$, and "if we negate that negation by multiplying $-A$ by $-A$ we get A^2, i.e., the original positive magnitude but at a higher degree, raised to its second power" (*Anti-Dühring,* p. 153). Engels even found the law operating in the history of philosophy. In early philosophy, he held, there is a simple, natural form of materialism according to which matter is the source of everything. This form of materialism was negated by idealism, which rightly showed that mind is not the same as matter, but wrongly held that matter is dependent upon mind. In its turn, idealism is negated by "modern materialism, the negation of the negation," which contains in itself two thousand years of philosophical development. Engels believed that in "modern materialism," that is, dialectical materialism, philosophy as previously understood is destroyed and yet preserved in the positive sciences.

This law, like the law of the transformation of quantity into quality, draws together some extremely disparate types of being. Is it likely, indeed, does it make sense to say, that the same principle is exemplified in a rule for operating on algebraical symbols and in the relationship of natural materialism, idealism, and dialectical materialism? One instance of the law that has given rise to much discussion is that of the grain of barley. What is it that negates what, and what is comprised in the negation of the negation? This problem was discussed by the Russian Marxist G. V. Plekhanov in his *The Development of the Monist View of History* (1895), in which he defended Engels's view against the criticisms of another Russian, N. K. Mikhailovskii, who had made fun of the idea that, as he put it, "oats grow according to Hegel." In his account of Engels's argument, Mikhailovskii took it that it is the stalk which negates the seed, and Plekhanov accused him of misquotation and asserted that it is the whole plant which does the negating. Plekhanov argued further that Engels's account of this botanical negation of the negation was supported by an authoritative textbook of botany, Philippe Van Tieghem's *Traité de botanique* (Paris,

1891), which had recently appeared. The whole discussion is entertaining but ludicrous. For the main difficulty about the law of the negation of the negation is that it can be made to fit almost anything by carefully choosing what are to count as the negating terms. The prime interest in the law is that it is intended to give support to the view that human progress is by means of destruction that leads to better things.

ENGELS'S PHILOSOPHICAL LEGACY. Engels was deeply interested in the advances of the sciences and believed that as a result of them nineteenth-century materialism had to be very different from earlier types of materialism. But Engels was drawn in two different directions. On the one hand, he sought to establish a naturalistic, scientific view of the world, and this led him in the same direction as the positivists. On the other hand, he was attracted by Hegel's dialectical method and by the romantic dream of a philosophy of nature, and this led him to regard the positivist outlook as thin and unadventurous. Like Marx, he deplored the conservative social tendencies of Auguste Comte and considered Hegel by far the better philosopher. Nevertheless, Engels did adopt one important positivist thesis, the thesis that knowledge of the world can be obtained only by the methods of the special sciences, so that all that can survive of philosophy is logic and the philosophy of the sciences. Thus, at the beginning of *Anti-Dühring* he wrote: "What still independently survives of all former philosophy is the science of thought and its laws—formal logic and dialectics. Everything else is merged in the positive science of nature and history." It should be noted that Engels here used the very adjective "positive" that had been formerly used by Comte de Saint-Simon and Comte. Although the positivists said nothing of "dialectics," Engels's point of approach from Hegelianism to positivism was his claim that the positive sciences make use of the dialectical method. But Engels, as we have seen, searched the sciences for examples of the dialectic and so applied his terms that he could not fail to find them there. This association of a positivist view of philosophy with what positivists would describe as a "metaphysical" view of the sciences was to remain a permanent feature of dialectical materialism.

Engels also bequeathed a problem about the nature of logic. Was formal logic disproved or rendered nugatory by the dialectical logic that was coming to fruition in the nineteenth century? In holding that there are existent contradictions Engels seemed willing to go against formal logic, but he also thought that formal logic would remain as a part of philosophy alongside dialectics. His position was complicated by the fact that in *Dialectics of Nature* he criticized formal logic as being "metaphysical" in the Hegelian sense already considered. As a result, controversy among exponents of dialectical materialism about the status of formal logic—by which they generally mean traditional Aristotelian logic—has been constantly renewed.

LENIN'S CONTRIBUTIONS

Lenin's great political achievements, as well as his deep philosophical interest, secured a respectful acceptance for his own philosophical views. And there is some appropriateness in the fact that Lenin's name, rather than Engels's, accompanies that of Marx in the name of the whole doctrine of Marxism-Leninism, since Lenin absorbed and reemphasized Engels's views before superseding him as a founding father.

Lenin's main contributions to dialectical materialism are the doctrine of *partiinost* ("party spirit" or "partisanship"), his elaborations of the Marxist theory of knowledge and of matter, and his renewed emphasis upon dialectics.

"PARTIINOST." Lenin briefly formulated the doctrine of *partiinost* as early as 1895, in the course of a controversy with the nonorthodox Marxist reformer Peter B. Struve, who had said that philosophical views were not a matter of controversy between parties but could be shared by members of opposing parties. Lenin wrote that *partiinost* is included in materialism and that no genuine adherent of materialism could remain uncommitted to the proletarian cause. In this particular context Lenin seems to have been thinking primarily of historical materialism; it is clear from his later writings, however, that he thought that the Marxist should never approach philosophical theories with detachment but should adopt or reject them in the light of their effects on the attainment of socialism. There are several points to be noted in Lenin's view. In the first place he held that dialectical materialism is not merely a theory but a form of action for the establishment of socialism. Thus, a dialectical materialist is necessarily a socialist, and his view of the world is inseparable from his efforts to promote the proletarian cause. In the second place, Lenin held that a socialist intellectual cannot be indifferent to philosophical matters. He is not a complete socialist unless he is a materialist, and a materialist of the right kind. Hence, the leaders of the socialist movement must always be on the alert to protect its doctrines against contamination by philosophical idealism. (This last is a doctrine that Stalin strictly enforced.) A fourth point on which Lenin laid great stress is that ide-

alism is fundamentally supernaturalistic, however tenuous the connection between certain forms of it and religion may appear to be on the surface. In attacking idealism, wherever and however it appears in the socialist literature, what is really being attacked is religion and the antisocialist class forces that uphold it.

The doctrine of *partiinost* derives from Marx's and Engels's theory of ideologies. Ideologies, in their view, are systems of ideas whose function is to defend and to justify the class interests of those who believe in them and teach them, and philosophical systems are ideologies in this sense. Bourgeois ideologies serve to promote bourgeois interests, and the way to criticize them is not primarily by intellectual refutation—this will have little or no effect as long as bourgeois class interests remain—but by unmasking the motives behind them. This view is supported by the Marxist doctrine of the unity of theory and practice. In writing a philosophical book a man is taking part in the social struggle, and in a society divided into classes he is of necessity promoting or endeavoring to promote some class attitude. Lenin considered that Marxists, who understand what is going on in the ideological sphere, should do deliberately and consciously what is so often done unknowingly. This attitude was powerfully expressed in his *Materialism and Empirio-Criticism* (1909). Lenin thought that certain members of the Russian Social Democratic party were spreading what were essentially idealist philosophical views, and he set out to put them right. These Marxists (false Marxists, as Lenin thought) were adopting, under the title of empiriocriticism, the phenomenalist theories of Ernst Mach and Richard Avenarius. In doing so, according to Lenin, they were adopting a cryptoidealist philosophy that could weaken the Marxist movement by dissipating its materialism. "Marx and Engels," wrote Lenin, "were partisans in philosophy from start to finish; they were able to detect the deviations from materialism and concessions to idealism and fideism in each and every 'new tendency'" (p. 352). Thus, *Materialism and Empirio-Criticism* was largely a diatribe intended to crush a view held to be dangerous to the party.

KNOWLEDGE AND MATTER. Lenin's *Materialism and Empirio-Criticism* is not only a partisan polemic but also the book in which Lenin expounded his views about knowledge and the nature of matter. It was pointed out above that some Russian social democrats had taken up ideas from the writings of Mach and Avenarius. Mach and Avenarius had tried to put forward as consistently empiricist a view as possible. Mach sought to eliminate from physics all notions that were not capable of direct or indirect verification in sense experience, and Avenarius sought for the terms in which the simplest and most economical explanations can be given. They both concluded that fundamentally the statements of science are statements of what people do experience or will experience and that scientific laws state how such experiences are correlated with one another. To the most elementary of these experiences Mach gave the name "sensations," and empiriocriticism amounted to phenomenalism, the view that material things are actual or possible sensations. Mach's theory of scientific knowledge is not unlike that of the idealist philosopher George Berkeley, who also sought to eliminate from the body of scientific knowledge any conceptions that could not be referred to sensations, or "ideas" (as he called sensations). Mach recognized the similarity between his view of science and that of Berkeley but pointed out that his view differed from Berkeley's in that he did not hold, as Berkeley did, that sensations were produced by God.

Lenin made the most of the fact that Mach's phenomenalist theory had affinities with that of Berkeley. Berkeley, Lenin said, was honest about his religious aims, whereas "in our time these very same thoughts on the 'economical' elimination of 'matter' from philosophy are enveloped in a much more artful form." Lenin objected that these phenomenalistic views run counter to our everyday practice, in which we come across material things and act upon them. We might call this the argument from common sense. He also objected that the theory that the material world is an orderly correlation of sensations is incompatible with the well-established scientific theory that there was once a time when matter existed but beings capable of having sensations did not. Berkeley, if he had known of this argument, could have countered it by saying that God could somehow have experienced the material world. If Mach had taken this course, Lenin claimed, he would have revealed his idealism.

Having rejected idealism and phenomenalism, Lenin had to give his own account of the material world and of our knowledge of it. He adopted Engels's theory that in perception material objects are "reflected" in the percipient and produce "copies" there. From this it would seem that the material world is much as we see and hear it to be, and Lenin seems to have emphasized this. Plekhanov, following Herrmann von Helmholtz, had argued that sensations are not exact copies of objects outside us but that they possess the same structure and might more accurately be termed "symbols" or "hieroglyphs." Lenin claimed, however, that Helmholtz's view undermines its

materialist basis, "for signs or symbols may quite possibly indicate imaginary objects, and everybody is familiar with the existence of such signs and symbols" (p. 239). Lenin did not see that a similar objection applies to "copies" or "reflections" as well, for unless we have independent knowledge of that from which the copy is made, we cannot know that it is a copy. Furthermore, Lenin held (*Materialism and Empirio-Criticism,* Ch. 5, Sec. 7) both that sensations copy what is in the physical world and that what is in the physical world is shown by science to be very different from what it appears to be. Thus, he wrote that sensations of red reflect "ether vibrations" of one frequency and sensations of blue, "ether vibrations" of another frequency, but he did not say how sensations can copy or be like the vibrations. Elsewhere he said that it is "beyond doubt that an image cannot wholly resemble the model" and went on to say that "the image inevitably and of necessity implies the objective reality of what it 'images'" (p. 240). By putting "images" in quotation marks, he seems to have been denying its literal force, and by saying that the images "cannot wholly resemble the model," he raised doubts about what it was he really meant to assert.

The basic thing that Lenin wanted to say about the nature of matter was that it exists objectively and independently; therefore, he actually defined matter as "that which, acting upon our sense-organs, produces sensations." This would apply to Berkeley's God as well as to material objects. Still, Lenin called this his "philosophical" account of matter, contrasting it with the "scientific" conception of matter, which changes as scientific knowledge advances. In Lenin's view, the philosophical conception of matter remains unaffected as the scientific view of it changes from atomist theories to theories of electromagnetism. In *Materialism and Empirio-Criticism* Lenin argued, probably correctly, that the electromagnetic theory of matter is no less materialistic than atomic theories. Indeed, he held that it is in closer accord with dialectical materialism. "Modern physics is in travail," he wrote, "it is giving birth to dialectical materialism" (pp. 323–324). Like Engels, he was attracted to theories of matter that "dissolve" the rigid substances and hard atoms of the older views. He believed that such theories were substituting dialectical concepts for metaphysical and mechanistic ones.

DIALECTICS. In 1894, in *What the "Friends of the People" Are*, Lenin quoted approvingly from Engels's *Anti-Dühring.* In *Materialism and Empirio-Criticism* he frequently referred to dialectics, without, however, making it the center of his discussion. But while he was in exile in Switzerland during World War I, he renewed his study of philosophy, particularly of its dialectical aspects. His *Philosophical Notebooks* (first published in 1933) show the wide extent of his reading during those years, particularly his detailed study of Hegel's *Science of Logic,* in which he noted some germs of historical materialism. Lenin's reading of this book led him to conclude that it was not so much opposed to materialist modes of thought as had previously been supposed. On the one hand, Lenin approved of the Marxist commonplace that Hegel's system is materialism turned upside down. On the other hand he wrote that in the final chapter of the *Science of Logic,* on the Absolute Idea, there is scarcely a mention of God and that "it contains almost nothing that is specifically *idealism,* but has for its main subject the dialectical method" (*Collected Works,* Moscow, 1961, Vol. 38, p. 234). It is apparent from Lenin's notes that his respect for the *Science of Logic* increased as he read it. Not only did he conclude that it transcended idealism but also that idealism itself has virtues. Two notes in particular may be referred to. Among his comments on Hegel's *Lectures on the History of Philosophy* he said: "Intelligent idealism is closer to intelligent materialism than stupid materialism" (p. 276). And at the end of a short paper titled "On the Question of Dialectics," written in 1915, he wrote that idealism "is a *sterile flower* undoubtedly, but a sterile flower that grows on the living tree of living, fertile, genuine, powerful, omnipotent, objective, absolute human knowledge" (*Philosophical Notebooks,* p. 363). Many of Lenin's jottings in his *Notebooks* are of this character, in marked contrast to the rancorous anti-idealism of *Materialism and Empirio-Criticism,* in which any approach toward idealism is regarded as treachery. Perhaps it is of significance that the one thesis common to Berkeley and Lenin is the thesis that nothing is substantial that is not active.

MAO ZEDONG (MAO TSE-TUNG)

Mao Zedong's writings on dialectical materialism are referred to here mainly because of the political eminence of their author. Apart from his poems, his writings are mostly on political subjects, and his chief excursions into philosophy are two short articles written in 1937, "On Practice" and "On Contradiction." It has been suggested that Mao has introduced an empiricist element into dialectical materialism, but this is not borne out by a study of these two writings. In the first, it is true, Mao stated that knowledge begins with sense perception in practical contexts, passes on to rational knowledge, which enables the world to be "molded" for human purposes,

and then leads to more rational knowledge at a higher level. It is not clear from the article whether the author was thinking of induction or of the testing of hypotheses or of both. But it is clear that, in Mao's view, in passing to this higher level "a leap" is made. In thus utilizing the law of the transformation of quantity into quality Mao was asserting that certain sorts of rational knowledge are different in kind from sense knowledge, and this can hardly be described as empiricism.

In "On Contradiction" Mao Zedong argued that in a contradiction each contradictory aspect "finds the presupposition of its existence in the other aspect and both aspects co-exist in one entity." As examples of this he mentioned life and death, above and below, misfortune and good fortune, landlords and tenant-peasants, bourgeoisie and proletariat, imperialists and colonies. He also argued that "each of the two contradictory aspects, according to given conditions tends to transform itself into the other," and as examples of this he cited the revolutionary proletariat becoming the rulers instead of the ruled, peace and war, landlords becoming landless tenants and landless tenants becoming smallholders.

It is easy to see the incongruities in both sets of examples. The opposition between life and death, for instance, is different from those between above and below and misfortune and good fortune, for there is nothing intermediate between life and death, whereas between above and below there is the relation of being at the same level and between good and bad fortune there is the condition of having neither the one nor the other. As to the second set of examples, the transformation of revolutionaries into rulers is not a logical transformation, but something that sometimes happens and sometimes does not. The example of peace and war is trivial. Mao wrote: "War and peace transform themselves into each other. War is transformed into peace; for example, the First World War was transformed into the postwar peace. … Why? Because in a class society such contradictory things as war and peace are characterised by identity under certain conditions." We know, of course, that wars end and that peace is often followed by war, but nothing is added to this by saying that a contradictory aspect transforms itself into its opposite, as if peace were one entity and war another. These writings of Mao Zedong's are, in fact, mainly concerned with immediate practical issues and contribute little to the philosophy from which they derive. It was in Soviet Russia that dialectical materialism was most fully elaborated after Lenin died.

See also Aristotle; Avenarius, Richard; Berkeley, George; Blake, William; Communism; Comte, Auguste; Descartes, René; Dühring, Eugen Karl; Engels, Friedrich; Faraday, Michael; Feuerbach, Ludwig Andreas; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Helmholtz, Hermann Ludwig von; Historical Materialism; Idealism; Infinity in Mathematics and Logic; Lenin, Vladimir Il'ich; Locke, John; Logical Paradoxes; Mach, Ernst; Marx, Karl; Marxist Philosophy; Materialism; Matter; Mikhailovskii, Nikolai Konstantinovich; Negation; Newton, Isaac; Phenomenalism; Plekhanov, Georgii Valentinovich; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Schelling, Friedrich Wilhelm Joseph von; Socialism; Thomas Aquinas, St.; Wolff, Christian; Zeno of Elea.

Bibliography

MARXIST WORKS

Marx

"Oekonomische–philosophische Manuskripte." In *Marx-Engels Gesamtausgabe,* edited by D. Riazanov and V. Adoratski. Berlin, 1927–1932. Division I, Vol. III. Translated by Martin Milligan as *Economic and Philosophic Manuscripts of 1844.* Moscow and London, 1959.

"Theses on Feuerbach," appended to *The German Ideology.* Edited by R. Pascal. London, 1938.

Misère de la philosophie. Brussels and Paris, 1847. Translated by H. Quelch as *The Poverty of Philosophy.* Chicago, 1910.

Marx and Engels

Die heilige Familie. Frankfurt, 1845. Translated as *The Holy Family.* Moscow and London, 1956.

Die deutsche Ideologie. Edited by V. Adoratski. Vienna, 1932. Translated as *The German Ideology,* edited by R. Pascal. London, 1964.

Engels

Herr Eugen Dührings Umwälzung der Wissenschaft. Leipzig, 1878. Translated by E. Burns as *Herr Eugen Dühring's Revolution in Science.* London, 1934.

Die Entwicklung des Sozialismus von der Utopie zur Wissenschaft. Zürich, 1883. Translated by E. Aveling as *Socialism: Utopian and Scientific.* London, 1892.

Ludwig Feuerbach und der Ausgang der klassischen deutschen Philosophie. Stuttgart, 1888. Translated as *Ludwig Feuerbach and the Outcome of Classical German Philosophy.* London, 1935.

Dialektika Prirody. Moscow, 1925. Translated as *Dialectics of Nature,* with an introduction by G. B. S. Haldane. London, 1940; new translation, London, 1954.

Plekhanov

Selected Philosophical Works. Moscow and London, 1961. Vol. I.

K Voprosu o Razvitii Monisticheskago Vzglyada na Istoriyu. St. Petersburg, 1895. Translated by Andrew Rothstein as *In Defense of Materialism: The Development of the Monist View of History.* London, 1947.

Lenin

Materializm i Empiriokrititsizm. Moscow, 1908. Translated by A. Fineberg as *Materialism and Empirio-Criticism.* Moscow and London, 1948.

"Karl Marx." In his *Collected Works.* Moscow and London, 1960–. Vol. 21, 1964.

Filosofskie Tetradi. Moscow, 1933. Translated as *Philosophical Notebooks,* in his *Collected Works.* Moscow and London, 1960–. Vol. 38, 1962.

Others

Cornforth, Maurice. *Dialectical Materialism,* 3 vols. London, 1954.

Kedrov, B. M. *Classification of Sciences. Book I: Engels, His Predecessors.* Moscow, 1961.

Stalin, Joseph. *Dialectical and Historical Materialism.* First published as Chapter 4 of Stalin's *History of the Communist Party of the USSR.* Moscow, 1939.

WORKS BY NON-MARXISTS

Acton, H. B. *The Illusion of the Epoch: Marxism-Leninism as a Philosophical Creed.* London, 1955.

Acton, H. B. "Karl Marx's Materialism." *Revue internationale de philosophie* nos. 45–46 (1958): 265–277.

Bochenski, J. M. *Der sowjetrussische dialektische Materialismus (Diamat).* Bern and Munich, 1950. Translated by Nicolas Sollohub as *Soviet Russian Dialectical Materialism.* Dordrecht, Netherlands, 1963.

Hook, Sidney. *Reason, Social Myths, and Democracy.* New York: John Day, 1950. Contains an excellent discussion of Engels's dialectical laws.

Joravsky, David. *Soviet Marxism and Natural Science.* London, 1961. Important on "mechanism" and the relation of dialectical materialism to positivism.

Jordan, Z. A. *Philosophy and Ideology. The Development of Philosophy and Marxism-Leninism in Poland since the Second World War.* Dordrecht, Netherlands, 1963.

Paul, G. A. "Lenin's Theory of Perception." *Analysis* 5 (1938): 65–73.

Wetter, Gustav. *Der dialektische Materialismus: seine Geschichte und seine Systeme in der Sowjetunion.* Vienna and Freiburg, 1953. Translated by Peter Heath from the fourth German edition as *Dialectical Materialism. A Historical and Systematic Survey of Philosophy in the Soviet Union.* London, 1958. Discusses Marx, Engels, Plekhanov, Lenin, and later Soviet writers.

H. B. Acton (1967)

DIALECTICAL MATERIALISM [ADDENDUM]

The term *dialectical materialism*, commonly used to describe the philosophy of Karl Marx, is suggested by certain statements of Marx, but was not a term that he himself used. In the afterward to the second German edition of *Capital*, Marx says, "My dialectic method is not only different from the Hegelian, but is its direct opposite" (1996a, p. 1:19). For Georg Wilhelm Friedrich Hegel, the Idea is an independent power, a "demiurge," for which the real world is an external, phenomenal form. For Marx, "the ideal is nothing else than the material world reflected by the human mind, and translated into forms of thought" (19).

SPECIES BEING

Marx does not here directly call his method "materialist," however. In his early *Economic and Philosophic Manuscripts of 1844* he rejects the antithesis of materialism and idealism as terms usually applied to separate individuals, for a social ontology of the human being. The human being is a *species being*—that is, a being that is not just a member of a species, as are individual animals, but one that takes the life of the species as an object of concern. The animal "is one with its life activity," whereas the human being "makes his life activity itself the object of this will and of his consciousness" (1996b, p. 3:276). (No doubt Marx, in his disparagement of individualism as animalistic, underestimated the extent to which such animals as gorillas, chimps, whales, and so on create real communities comparable to those of humans.) Rather than take human consciousness as a given, and then relate it in some way to the body, Marx explains human consciousness as the result of the way in which human individuals relate to their species. "[I]t is only because he is a species-being that he is a conscious being" (276). So he concludes in general terms that go beyond the opposition of materialism and idealism that "*just as* society itself produces *man as man*, so is society *produced* by him.… Thus *society* is the complete unity of man with nature—the true resurrection of nature—the accomplished naturalism of man and the accomplished humanism of nature" (278).

In his first thesis on (Ludwig Andreas) Feuerbach, Marx says that for "all previous materialism" things are regarded as objects of contemplation, and so the active side of human practice "was set forth abstractly by idealism" (1996d, p. 5:3). In the third thesis he rejects "the materialist doctrine that men are products of circumstances and upbringing." Instead, he asserts "the coincidence of the changing of circumstances and of human activity or self-change," which is "revolutionary practice" (4). The problem with Feuerbach's materialism is that he conceives of the individual in isolation, and of the species as only "an inner, mute, general character which unites the many individuals *in a natural way*," whereas "the essence of man is no abstraction inherent in each single

individual. In its reality it is the ensemble of the social relations" (4).

Language is the initial form of human species consciousness. Language is not primarily a mode of expressing the brain activity of the separate individual, but a vehicle of the social intercourse or species being that constitutes distinctively human consciousness. "[L]anguage is practical, real consciousness that exists for other men as well, and only therefore does it also exist for me" (1996c, p. 3:44). While consciousness presupposes the activity of the brain, it exists primarily outside of the individual's head in the linguistic interchange between human beings. Language-mediated consciousness is the direct presence of the species to the individual while at the same time the individual is always creatively reproducing the species in new forms. "Consciousness is, therefore, from the very beginning a social product, and remains so as long as men exist at all" (44). In distinguishing such socially mediated human consciousness from animal consciousness Marx supposes a nonreductivist, emergentist, or "dialectical" materialist conception of the relation of conscious activity to the brain. Brain activity is nature presupposed by human social activity but then dialectically transformed and uplifted or "resurrected" by it. This is Hegel's dialectical sublation (*aufhebung*).

LABOR AND THE CUNNING OF REASON

This relational conception of the human individual is given specific expression in the different historical forms of social existence and in terms of different levels of analysis within these social forms. At the most basic level, and within every social form of existence, socially related individuals transform the natural world and "humanize" it through "labor." In *Capital* Marx again compares human conscious activity with that of animals, "[W]hat distinguishes the worst architect from the best of bees is this, that the architect raises his structure in imagination before he erects it in reality" (1996a, p. 35:188). Thus, the idea, purpose, or goal has primacy in relation to the materials of labor. Marx approvingly cites Hegel's own idealist analysis of the labor process in his *Logic of Hegel*, where he writes, "Reason is just as cunning as she is powerful. Her cunning consists principally in her mediating activity, which, by causing objects to act and re-act on each other in accordance with their own nature, in this way, without any direct interference in the process, carries out reason's intentions" (1968, 350). Thus, Marx directly incorporates Hegel's demiurge, the Idea, into his analysis of human labor. There are certainly material things or complexes of things, tools, and materials involved in the labor process. But thanks to the "cunning of Reason," those ideal constructions of language that constitute the presence of the species to the individual, the human agent powerfully channels the forces of nature in ways that lead to the intended goal.

Human activity is mediated by historically evolved systems of tools, material and ideal, spiritual, or cultural. While consciousness is extended through language and other means of communication (such as books, newspapers, and the Internet), practical activity on the natural environment is not merely the activity of the physical body, but of the body extended by tools. Naturalism and reductive materialism abstracts the human individual from his or her intrinsic connection to humanly produced tools and reduces the individual to the naked body alone. The human being for Marx is not "the naked ape" (Morris 1967) but the ape who wears clothes and extends his or her natural existence by humanly produced organs of thought and action existing outside of the biological organism. These organs of thought and action constitute the presence to the individual of the being of the species and are in turn objects for the individual's creative and transformative thought and action.

ALIENATION OF SPECIES BEING

Whereas in other social systems the social connections tying individuals to one another are evident on the surface of social life—as in the direct communal relations in early societies, or as the personal relations of master and slave, lord and serf—in capitalism the social relations are hidden, while the seemingly separate individual comes to the forefront of empirical awareness. Here is the historical basis of those separate, abstract individuals of both idealist and materialist philosophies that Marx attributes to the alienation of human individuals from one another. In capitalism the social relations that essentially underlie the activity of individuals take the specific existential form of separate individuals exchanging their products and labor in the market. In appearance individuals confront other individuals existing outside of them, competing with them, and serving as means to the achievement of their goals, as they are to the others. In essence there is a holistic system of division of labor that makes possible the highly specific activities of each individual and requires their interdependence. It is this underlying social interdependence that constitutes the reality of communism that progressively emerges in specific ways within the womb of capitalism. Marx's employment of the dialectical categories of essence and existence, or reality

and appearance, reflects his adoption of Hegel's dialectical logic.

FETISHISM OF THE COMMODITY

The product of labor in this context has a dual nature. It is an individual object of some kind that can be described in its own terms: an automobile or a software program embodying the current state of technology and specific skills of the workers who produced it. At the same time it has an economic value that cannot be explained by its material qualities and that enables it to be equated somehow with a qualitatively different object:

> A commodity is therefore a mysterious thing, simply because in it the social character of men's labour appears to them as an objective character stamped upon the product of that labour; because the relation of the producers to the sum total of their own labour is presented to them as a social relation, existing not between themselves, but between the products of their labour. (Marx 1996a, pp. 35:82–83)

Materialists and idealists battle interminably over the explanation of this and other mysteries of philosophy because they preserve the standpoint of the independent, separate individual that gives rise to them. Behind the mystery of economic value is the social nature of human labor, the fact that each product embodies a certain proportion of the combined labor of society. Because the people whose interdependent labor is responsible for the product have organized themselves as separate, disconnected individuals, their underlying social connection takes the form of a mysterious, nonmaterial property of their products. In the value form of the commodity spirit and matter confront one another as irreducible opposites: for the "value-relation between the products of labour … [has] absolutely no connection with their physical properties and with the material relations arising therefrom" (Marx 1996a, p. 5:83). Consequently, "[t]here it is a definite social relation between men, that assumes, in their eyes, the fantastic form of a relation between things" (83).

This complex relationship produces the "fetishism" (Marx 1996a, p. 5:83) of the products of labor when they become commodities. The combined power of human beings appears before them as an external power ruling over them—the market and the quasi-omnipotent power of money. The mystery of the nonmaterial characteristics of the product can ultimately be explained in one of two ways: (1) As the expression of the social relations between the producers, seen in essentially cooperative activities that belie the capitalist form of private ownership. This is the kind of social-historical and dialectical "materialism" that Marx espouses. (2) Or it can be approached by reference to "the mist-enveloped regions of the religious world," in which "the productions of the human brain appear as independent beings endowed with life, and entering into relation both with one another and the human race" (83). Marx thought that Hegel's idealism, for all its advances over previous materialism, did not escape this religious, other-worldly, appearance of alienated human activity.

See also Communism; Marx, Karl.

Bibliography

Marx, Karl. "Capital, Vol. 1." In *Marx, Engels, Collected Works*. Vol. 35. New York: International, 1996a.

Marx, Karl. "Economic and Philosophic Manuscripts of 1844." In *Marx, Engels, Collected Works*. Vol. 3. New York: International, 1996b.

Marx, Karl. "The German Ideology." In *Marx, Engels, Collected Works*. Vol. 5. New York: International, 1996c.

Marx, Karl. *The Logic of Hegel*. New York: Oxford University Press, 1968.

Marx, Karl. "Theses on Feuerbach." In *Marx, Engels, Collected Works*. Vol. 5. New York: International, 1996d.

Morris, Desmond. *The Naked Ape: A Zoologist's Study of the Human Animal*. New York: McGraw-Hill, 1967.

James M. Lawler (2005)

DIALECTIC IN ISLAMIC AND JEWISH PHILOSOPHY

In these closely related traditions dialectic is primarily associated with the science of *kalām*, commonly translated as "theology," but literally meaning "word," "speech," or "discussion." *Kalām* began in the eighth century as an intellectual defense of Islam against external critics and quickly developed into an internal debate over doctrinal issues concerning the legitimacy of political authority, the necessary conditions of religious belief, predestination and free will, the ontological status of the Qur'ân, and the relation of God's attributes to His essential Unity. *Kalām* was subsequently appropriated by Arabic-speaking Jews living in the Islamic realm, who shared some of its concerns and employed its distinctive techniques and formulas in the defense and systematic explanation of their own faith.

Kalām in general is marked by its dual reliance on revelation and reason. The *kalām* theologians, or *mutakallimūn*, took scripture as their primary data but

employed rational argumentation to produce the most robust and coherent interpretations thereof. This distinguished them on the one hand from traditionalists and literalists who saw logical disputation and interpretation as leading to heresy, and on the other hand from the Greek-influenced Islamicate philosophers, or *falāsifa*, who were more fully committed to the demands of reason and thus wary of their theological brethren's residual dogmatism. *Kalām*'s method of reasoning and argumentation was dialectical in at least two respects. The first recalls the Aristotelian concept of dialectic, insofar as the *mutakallimūn* based their arguments on merely probable or generally accepted beliefs—specifically, the revealed truths of Islam or Judaism—rather than rationally self-evident first principles or premises that necessitated consent. The *falāsifa*, who appropriated Aristotle's hierarchical distinction between dialectic and demonstration, considered this approach insufficiently rigorous. While their own adoption of the demonstrative syllogism held out the prospect of certitude, they saw the *mutakallimūn* as hobbled by the questionable epistemic status of their faith-based premises. However, the *falāsifa* did not reject dialectic altogether. They generally recognized its value as a propaedeutic for honing intellectual skills, as well as a tool for communicating crucial truths to those unequipped for philosophical discourse. The *mutakallimūn*, for their part, remained dubious about the philosophers' claims to apodictic certainty.

The second sense in which *kalām* was dialectical recalls certain aspects of the Socratic method. First, it was dialogical: It typically took a question and answer form, in effect presupposing the existence of an intellectual adversary to drive the discourse forward. Its method was thus parasitic: The *mutakallimūn* tended to establish their own conclusions indirectly, by teasing out inconsistencies or internal contradictions in the opponent's position. This strategy often involved the use of dilemmas, where the adversary would find himself trapped between two unacceptable consequences that could be avoided only by adopting the questioner's position. The *mutakallimūn* commonly fashioned their arguments with an eye to the specific concerns, presuppositions, and methods of their opponents as well, advancing internal critiques of their adversaries to refute them on their own terms. Ironically, their assault on the *falāsifa* in the eleventh and twelfth centuries, which effectively brought an end to the classical period of Islamic philosophy, required the instrumental adoption of Aristotelian logic, specifically, the demonstrative syllogism.

Although the presence of dialectical methods within the Islamic and Jewish traditions is often attributed directly to Greek influences, a number of contemporary scholars and historical figures have made the case that versions of these argumentative strategies in fact predate exposure to Christian, Greek, or Syriac sources.

See also Aristotle; Dialectic; Islamic Philosophy; Jewish Philosophy.

Bibliography

Black, Deborah L. *Logic and Aristotle's Rhetoric and Poetics in Medieval Arabic Philosophy*. Leiden, Netherlands: E. J. Brill, 1990.

Haleem, Muhammad Abdel. "Early Kalam." In *History of Islamic Philosophy*, edited by Seyyed Hossein Nasr and Oliver Leaman. London: Routledge, 1996.

Ess, Josef van. "The Logical Structure of Islamic Theology." In *Logic in Classical Islamic Culture*, edited by G. E. von Grunebaum. Wiesbaden, Germany: O. Harrassowitz, 1970.

Ess, Josef van. *Theologie und Gesellschaft im 2. und 3. Jahrhundert Hidschra: Eine Geschichte des religiösen Denkens im frühen Islam*. 6 vols. Berlin: Walter de Gruyter, 1991–1996.

Gardet, Louis. "La dialectique en morphologie et logique arabes." In *L'ambivalence dans la culture arabe*, edited by Jean-Paul Charnay. Paris: Éditions Anthropos, 1967.

Gardet, Louis. "'Ilm al-kalam." In *Encyclopedia of Islam*, Vol. III, edited by H. A. R. Gibb et al. Leiden, Netherlands: E. J. Brill, 2003.

Leaman, Oliver. "Logic and Language in Islamic Philosophy." In *Companion Encyclopedia of Asian Philosophy*, edited by Brian Carr and Indira Mahalingam. London: Routledge, 1997.

Stroumsa, Sarah. "Saadya and Jewish *kalam*." In *The Cambridge Companion to Medieval Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman. New York: Cambridge University Press, 2003.

Watt, W. Montgomery. *Islamic Philosophy and Theology: An Extended Survey*. 2nd ed. Edinburgh, Scotland: Edinburgh University Press, 1985.

Wolfson, Harry Austryn. *The Philosophy of the Kalam*. Cambridge, MA: Harvard University Press, 1976.

Wolfson, Harry Austryn. *Repercussions of the Kalam in Jewish Philosophy*. Cambridge, MA: Harvard University Press, 1979.

Peter S. Groff (2005)

DICTIONARIES

See *"Philosophy Dictionaries and Encyclopedias" in Volume 10*

DIDEROT, DENIS

(1713–1784)

Denis Diderot, the French encyclopedist, philosopher, satirist, dramatist, novelist, and literary and art critic, was the most versatile thinker of his times and a key figure in the advancement of Enlightenment philosophy.

LIFE

Born in Langres, son of a master cutler, Diderot was a brilliant student in the local Jesuit schools. He was sent to college in Paris and received his master's degree at the age of nineteen. Afterward, he refused to adopt a regular profession and, when his allowance was cut off, lived for many years in poverty and obscurity. His great ambition was to acquire knowledge. In this he was eminently successful, for he emerged from this period of self-education with an excellent command of mathematics and considerable proficiency in the Greek, Italian, and English languages. He first came into public notice as a translator of English works—a history of Greece, the earl of Shaftesbury's *Inquiry concerning Virtue and Merit* (1745), and Robert James's *Medicinal Dictionary* (1746–1748). He was secretly married in 1743; and his wife bore him a number of children, all of whom died in childhood except a daughter, Angélique, who lived to perpetuate the memory of her distinguished father.

In 1746 he published his first original work, the bold and controversial *Pensées philosophiques.* In that year, too, he became associated with the *Encyclopédie,* the greatest publishing venture of the century, of which he soon became editor-in-chief, with the aid of Jean Le Rond d'Alembert for the mathematical parts. This enterprise was his chief occupation and source of income until 1772. The boldness of his thought, in spite of the dexterity with which he attempted to conceal it, met almost instant opposition, resulting in the seizure of manuscripts, censorship, and temporary suppression. Only a man of Diderot's indomitable courage and determination could have brought the project to a successful conclusion.

In 1749, while manuscripts for the *Encyclopédie* were being prepared for the printer, Diderot published his *Lettre sur les aveugles* (Letter on the blind), in which he questioned the existence of purpose or design in the universe. For this and other suspect works he was seized by the police and spent a few uncomfortable months in the prison of Vincennes. His reputation in his parish as a materialistic atheist was catching up with him. The subsequent *Lettre sur les sourds et muets* (Letter on the deaf and dumb; 1751), equally original, was mild enough to escape persecution. His *Pensées sur l'interprétation de la nature* (Thoughts on the interpretation of nature; 1754) was both a plea for strict adherence to the scientific method and an exposition of results of that method, including definite evidence in support of evolutionary transformism.

After the official suspension of the *Encyclopédie* in 1759, Diderot prudently withheld his most important philosophical works for the use of posterity. The *Rêve de d'Alembert* (D'Alembert's dream), written in 1769, and the *Réfutation de l'ouvrage d'Helvétius* (Refutation of Helvétius) first became public in the nineteenth century. *Le neveu de Rameau* (*Rameau's nephew*), a scathing satire of eighteenth-century society, and the novels *La religieuse* (The nun) and *Jacques le Fataliste* (*Jacques the Fatalist*), which saw the light of day only after the French Revolution, as well as various short stories and dialogues, were all of ethical import. Two bourgeois dramas, *Le fils naturel* (The natural son; 1754) and *Le père de famille* (The father of the family; 1758), accompanied by critical essays, could, however, be safely published, though the *Paradoxe sur le comédien* (The Paradox of the actor), important for its aesthetic insights, was withheld. Diderot's *Salons,* replete with brilliant criticism of art and literature, were also published posthumously, although in manuscript copy they formed an important part of Friedrich Grimm's *Correspondance littéraire,* written only for foreign consumption. Diderot knew that his ideas were too advanced for his own generation, but he maintained the conviction that he would some day be appreciated at his true value.

When, in 1772, his long labors on the *Encyclopédie* were ended, Diderot set off for St. Petersburg by way of Holland and spent some months in 1773 in intimate conversations with Catherine the Great. Persuaded of his merit through Grimm, she had not only paid in advance for his library (he desperately needed the money as a dowry for his daughter) but also gave him a salary as its custodian until his death. Baron d'Holbach's *System of Nature* (1770), frankly atheistic and materialistic, had sharply drawn the line between atheism and deism, and both Catherine and Frederick II took the side of the less revolutionary Voltaire. Since Diderot supported Holbach in this controversy, his political *Observations* on Catherine's plan to recodify Russian law were deemed too radical and suppressed by his royal patron.

Returning to France in 1774, Diderot spent the remaining years of his life in semiretirement, enjoying at least a semblance of domestic felicity. His letters to his mistress, Sophie Volland, form, next to Voltaire's, the

most interesting correspondence of the century. His final work, the *Essai sur les règnes de Claude et Néron* (Essay on the reigns of Claudius and Nero; 1778–1782), was a eulogy of Stoic virtue, as illustrated by Seneca, and also a reply to charges of treachery and immorality made against Diderot in the *Confessions* of Jean-Jacques Rousseau, his former friend and coworker.

Diderot died in Paris six years after Voltaire and Rousseau, with whose names his is inextricably linked as a leader of the French Enlightenment.

GENERAL PHILOSOPHICAL ATTITUDES

Diderot's philosophy was remarkably undogmatic. He advocated the open mind and believed that doubt was the beginning of wisdom and often its end; he continually questioned his own theories and conclusions, developed extreme theses, or paradoxes, in ethics and aesthetics, and decided that "our true opinions are those to which we return the most often." Nevertheless, after passing briefly through a period of deistic belief (a deist, he finally concluded, was a man who had not lived long enough—or wisely enough—to become an atheist), he became an unabashed and enthusiastic materialist and developed a theory of materialism much less vulnerable than that of his forebears. His main contribution was a philosophy of science that looked far into the future and upon which his aesthetic and ethical theories were firmly and inseparably founded.

SENSATIONALISM. Like Voltaire, Rousseau, and Étienne Bonnot de Condillac, Diderot was early preoccupied with the theory of sensationalism. At weekly dinners with the latter two, John Locke's psychology was thoroughly discussed. Between Diderot and Condillac influence was undoubtedly mutual. But Condillac, having taken holy orders and being therefore more circumspect, worked out a more systematic and more abstract philosophy and left it to Diderot to direct French sensationalism into definitely materialistic channels.

Diderot's philosophical thought was clarified by his constant distrust of abstractions. Abstractions, he declared in *Rêve de d'Alembert,* are linguistic signs, which are useful in speeding up discourse and upon which the abstract sciences are built; but as symbols emptied of their ideas, they are obstacles to clear thinking. Those who use abstractions must have constant recourse to examples, thus giving them perceptibility and physical reality. The mind is nothing but the brain functioning; the will is the latest impulse of desire and aversion. The naming of things is purely conventional.

Diderot's early philosophical publications were especially concerned with problems of communication. His empirical mind could not be satisfied with speculative studies, such as Condillac's theoretical experiment of endowing a statue with one sense at a time. He chose rather to study the actual cases of individuals deprived of the sense of sight or the sense of hearing. His *Lettre sur les aveugles* (1749) dealt first with case histories and the problems of "reading" through touch, illustrated by the methods of Nicholas Saunderson, the blind professor of mathematics at Oxford. This first truly scientific study of blindness led to Diderot's imprisonment. The passage that provoked the authorities was an imaginary deathbed conversation, in which the blind professor, unable to appreciate the alleged perfection of the order and beauties of nature, expressed his consequent doubts as to the existence of an intelligent God. The treatise on the deaf and dumb, two years later, was also based on scientific observation, but proceeded to discuss aesthetic theories, especially the importance of gesture to communication. In his later posthumous works, sensationalism played an important role in the development of his materialistic monism.

EMPIRICISM. As early as 1748, in the libertine novel *Les bijoux indiscrets* (The indiscreet toys), Diderot showed himself a pronounced empiricist, a firm believer in the efficacy of the scientific method. In an important chapter of that work, Experience (the word meant both observed fact and experiment) figures first as a growing child, who discovers with the aid of a pendulum the velocity of a falling body, calculates the weight of the atmosphere with a tube of mercury, and with prism in hand, decomposes light. The child visibly grows to colossal stature and, like a Samson, crumbles the pillars of the Portico of Hypotheses.

Diderot's *Pensées sur l'interprétation de la nature,* taking its title and inspiration from Francis Bacon, again extolled the experimental method above purely rationalistic theory. Following the work of Pierre-Louis Moreau de Maupertuis and Comte de Buffon—and especially in studying Louis Daubenton's anatomical comparison of the foot of the horse and the hand of man—Diderot arrived at principles of transformism and natural selection that were to influence greatly his mature philosophy. He surmised that "there had never been but one animal, prototype, through differentiation, of all other animals."

The dawning of the age of biological science, he believed, would usher in the great discoveries of the future.

IMAGINATION. Observation and the classification of natural phenomena was the first and essential step, but the great scientist must perceive relationships and form hypotheses, subject to experimental verification. Diderot closely associated the poetic imagination with the scientific, both in theory and practice. This theory is clearly expounded in the first of the three "conversations" of *Rêve de d'Alembert.* This section discusses the role of analogy, which is merely the working out of the rule of three by the feeling instrument that is man. To the genius, whether poet or scientist, will come the sudden perception of a new relationship, resulting in poetic metaphor or useful hypothesis.

STYLE. Diderot's own mind worked in sudden flashes of perception. His best philosophical works are random or loosely associated thoughts or observations—or dreams. His satirical narrative, *Rameau's Nephew,* and his novel, *Jacques the Fatalist,* are apparently loosely constructed, much given to dialogue, with digressions and intercalated stories after the manner of Laurence Sterne. They follow the pattern of general conversation, in which one idea gives birth to another, and so on, until the thread is difficult to retrace. The theory of associationism was firmly based, however, on his theories of sensationalism and memory (to be discussed below).

SCIENTIFIC BACKGROUND. Diderot's inquisitive and encyclopedic mind equipped him admirably to comprehend the great advances that the sciences were making in the middle of the century. From mathematics he turned to chemistry and for three years studied assiduously under Guillaume-Francois Rouelle, forerunner of Antoine Lavoisier. He was well acquainted with the work of the Dutch biologists Niklaas Hartsoeker and Bernard Nieuwyntit, who laid the foundations for the still unknown science of genetics. He was familiar with Abraham Trembley's experiments with the freshwater polyp, and with Joseph Needham's discovery of Infusoria, in apparent proof of the theory of spontaneous generation. These experiments influenced his development of the concepts of the sensitivity of matter and the essential identity of its organic and inorganic forms.

As translator of Robert James's *Medicinal Dictionary,* Diderot was well informed in the science of medicine. Characteristically, he sought (in vain), before writing his *Lettre sur les aveugles,* to be admitted to an operation for cataract, and he consorted with doctors, many of whom were contributors to the *Encyclopédie.* While in prison at Vincennes, the recently published first three volumes of Buffon's *Natural History* received his careful scrutiny, and from all possible sources he collected case histories of injuries to, and surgical operations on, the brain.

By 1769, when he composed *Rêve de d'Alembert,* Diderot was adequately prepared to develop an original philosophy of science, a monistic theory that has been described as naturalistic humanism and dynamic, or "energetistic," materialism, which far surpassed the mechanistic theories of his forebears, from Lucretius to Julien Offray de La Mettrie, and foreshadowed Charles Darwin. In this work, first published in 1830, Diderot showed himself at once a great and an imaginative philosopher and writer. In its pages, his mature philosophy, presented fantastically but seriously, was best illustrated.

MATERIALISM—MATTER IN MOTION

Diderot adopted the Heraclitean theory of flux. The universe, for him, was a single physical system, obeying the immutable laws that René Descartes assigned to matter in motion; it was dynamic or "becoming," rather than static or created. Unlike Descartes, however, Diderot followed John Toland in believing that motion was not added but was essential to matter. He gave the idealistic monad of Gottfried Wilhelm Leibniz a positive content. Diderot maintained that not only are bodies affected by external force but that the atom contains internal forces, a form of kinetic or potential energy. All things carry with them their opposites; being and not-being are part of every whole. "Living," he wrote, "I act and react as a mass; dead, I act and react in the form of molecules. Birth, life, decay, are merely changes of form." No knowledge was gained, no solution reached, in postulating a Creator or supernatural agency to account for material phenomena. All change, including the transformation of the universe from chaos to order, was to be explained by the interaction of the elementary material particles. What man perceives as order is simply his apprehension of the laws of motion as enacted by material bodies.

SENSITIVITY OF MATTER. An additional and very important hypothesis upon which Diderot's construction was built was the sensitivity of all matter, both inorganic and organic. By postulating both motion and sensitivity as inherent in matter, he felt that the entire range of natural phenomena (both physical and mental) and the full variety of experience could be adequately explained. All that nature contains is the product of matter in motion,

subject to the processes of fermentation produced by heat; through eons of time growth, increasing complexity, and specialization have occurred.

Diderot believed that there were no inexplicable gulfs between the various kingdoms. The known facts concerning the inorganic, the organic, plant, animal, and man, were like islands jutting out of a sea of ignorance. As the waters receded through scientific investigation, the missing links would be discovered. "How d'Alembert differs from a cow," he admitted, "I cannot quite understand. But some day science will explain." He nevertheless attempted to trace the development of his friend, from the earth mold to mathematician, from the unconscious through the subconscious to the conscious life.

BIOLOGY AND EVOLUTION

During Diderot's lifetime the biological sciences were in their infancy. The scope and profundity of his insights are therefore all the more amazing. When scientific facts failed him, he had recourse to hypotheses that he was convinced would some day be verified. It was in consideration of this conviction that he presented his mature philosophy as a dream, a dream that, with the passage of time, can truly be called prophetic.

The crucial problem that confronted Diderot was to account for the emergence and behavior of the living individual. The coordinated behavior and continuous identity that characterize the organism seemed to transcend any possible organization of discrete material particles. It was difficult to see how merely contiguous material parts could form an organic whole capable of a unified and purposeful response to its environment. Traditionally, the existence of unique species and individuals was explained by recourse to supernatural design and metaphysical essence.

Contemporary science offered Diderot a choice between preformation, a Lucretian theory accepted at times by La Mettrie, and epigenesis, which explained organic formation in terms of juxtaposition and contiguity. Diderot rejected preformation, and in support of epigenesis he developed the concept of molecular combinations endowed with specialized functions and organic unity. In *Rêve de d'Alembert,* Diderot employed the image of a swarm of bees in an attempt to bridge the gap between contiguity and continuity in the production of a whole that is qualitatively unique and different from the sum of its parts. He pointed out that although the swarm consists simply of numerous separate individuals in physical contact, it does, as a whole, possess the characteristic of purposeful, unified behavior that is associated with the individual organism. It is possible to mistake the swarm of thousands of bees for a single animal. The unity of the organism is derived from the life of the whole, and Diderot thus affirmed the continuity of the kingdoms and refuted the metaphysical principle of essences. A half century later the discovery of the organic cell and the principles of cell division confirmed his views.

Diderot found support for his theories in the embryological ideas that he had gathered from his reading, especially of Albrecht von Haller's *Elements of Physiology,* and from Dr. Bordeu, his friend and the protagonist in the conversations of *Rêve de d'Alembert.* In the conversation with d'Alembert, which gives rise to the dream, Diderot attempts briefly to trace d'Alembert from the parental "germs." He then describes how, under the influence of heat, the chicken develops within the egg. Excluding all animistic hypotheses, he declares that this development "overthrows all the schools of theology; … from inert matter, organized in a certain way and impregnated with other inert matter, and given heat and motion, there results the faculty of sensation, life, memory, consciousness, passion, and thought."

HEREDITY. Diderot's conviction of the importance of hereditary factors constitutes the main argument of his refutation of Claude-Adrien Helvétius's work *On the Mind,* in which education and law, purely environmental factors, were proposed exclusively as causes of the development of a moral society. Diderot agreed with Bordeu ("organs produce needs, and reciprocally, needs produce organs") on the Lamarckian principle of the inheritability of acquired characteristics. Moreover, he clearly stated his belief that the individual recapitulates the history of the race and that certain hereditary factors may crop up after many generations.

To explain how parental factors are inherited (cells and genes were as yet unknown), Diderot resorted to a hypothesis of organic development through a network or bundle of threads (or fibers or filaments), which strongly suggested the nervous system. Any interference with the fibers produced abnormalities, or "monsters." (He was one of the first to seek to understand the normal through the abnormal, both in embryology and psychology.) In his careful description, in *Rêve de d'Alembert,* of the embryological differentiation between the male and female sex organs, he was led to surmise that man is perhaps the "monster" of the woman, and vice versa. His theories clearly foreshadow not only the phenomena of recessive genes but also the fundamental role of chromo-

somes. One of his chief arguments against design in the universe was nature's prolific production of "monsters," most of which were too ill adapted to their environment to survive. Their elimination was the closest he came to the principle of natural selection.

MATTER AND THOUGHT

Diderot believed that once it is granted that sensitivity is a property of matter and that matter thereby develops increasing complexity and specialization, it then follows that thought can best be understood as a property of that highly complex and specified material organ, the brain. He accepted Bordeu's theory of the individual life of the various bodily organs. All were linked, however, through the nervous system to the central organ, which, depending upon circumstances and temperaments, exerted more or less control over them. Personal identity, the unified self, was thus assured by the nervous system, and the brain played the role of both organ and organist.

MEMORY. Self-awareness, however, depends entirely on the remembering function of the human brain. Quite characteristically, Diderot assigned a neural mechanism to Locke's theory of the association of ideas. In his investigations of the physical substrata of memory, he read all he could find on the anatomy of the brain and injuries to the brain and consulted doctors and specialists in brain surgery. A number of case histories were reported in *Rêve de d'Alembert.* In the preliminary conversation with d'Alembert, however, he used La Mettrie's metaphor of vibrating strings and harmonic intervals to explain the association of images and memory, the passage from sense perceptions to comparisons, reflection, judgment, and thought. Memory furnishes the continuity in time, the personal history that is fundamental to self-consciousness and personal identity. In Diderot's mind, memory was corporeal, and the self had only material reality. He thus attempted to give psychology a scientific, physiological basis, which was further developed in the nineteenth and twentieth centuries.

In the midst of notes taken mostly from his reading and published later as the *Éléments de physiologie,* Diderot included an eloquent passage in support of his theory: "I am inclined to believe that all we have seen, known, perceived, or heard—even the trees of a great forest … all concerts we have ever heard—exists within us and unknown to us." He could still see in his waking hours the forests of Westphalia, and could review them when dreaming—as brilliantly colored as if they were in a painting. Moreover, "the sound of a voice, the presence of an object … and behold, an object recalled—more than that, a whole stretch of my past—and I am plunged again into pleasure, regret, or affliction."

DREAMS AND GENIUS. The concept of the greater or lesser control exerted by the central organ over the other organs of the body was applied by Diderot not only to dreams but also to the phenomenon of genius. In sleep, control is relaxed and anarchy reigns. A random recall in the central organ may then be referred to the subordinate organ, or the procedure may be reversed, from organ to brain. In dreams, random combinations may be formed and dragons created. Only personal past experience is available, however, for such imaginings. The one impossible dream is that the dreamer is someone else.

Applied to genius, the explanation of which was of great concern to Diderot and an important aspect of *Rameau's Nephew,* the concept of central control ran into difficulties. In the early *Pensées philosophiques,* in opposition to Blaise Pascal, he championed the strong emotions as the chief source of the good, the true, and the beautiful. Later, his acquaintance with David Garrick led him to write a paradox on the acting profession, in which he claimed that the great actor, with complete command of his emotions, makes his audience laugh or weep by coolly calculated gesture and intonation; he must register the emotions, but not feel them at the same time. In *Rêve de d'Alembert* he explained that dominating control by the center produced wise and good men but that genius was the result of the strongest emotions under almost complete control, a theory that could be illustrated by the horseman, Hippolytus, in firm command of the most spirited horses that Greece produced. In Diderot's hands, genius was not a mere talent produced, as Helvétius had claimed, by education and chance, but a psychophysiological phenomenon, and in that respect akin, when central control is lost, to madness.

ETHICS

The fundamental principles of Diderot's ethics may be found most readily in *Rêve de d'Alembert.* Will and liberty (free will) he described as senseless terms, abstractions that obscured the facts. The will of the waking man is the same as that of the dreamer: "the latest impulse of desire and aversion, the last result of all that one has been from birth to the actual moment." "There is only one cause … and that is a physical cause." But Diderot clearly distinguished between fatalism and determinism. Man is not, like the lower animals, a prey to the bombardment of the senses. The self, the brain with its properties of memory

and imagination, intervenes between the external stimulus and the act.

Diderot was tempted, but refrained from writing a treatise on ethics. Many critics have attributed this failure to the moral dilemma posed by his determinist convictions. It is more probably that he felt his ideas were too advanced for the age and society in which he lived. Moral problems were foremost in his mind throughout his career. A letter of 1756 stated clearly his deterministic beliefs. Heredity played a dominant role, for some, happily, are endowed with moral or socially acceptable propensities, while others, unfortunately, are not. Moral monsters must be eliminated, but in general, man is modifiable. *Rameau's Nephew* is, among other things, the story of the dilemmas that confront moral man in an immoral society, in which honesty is not necessarily the best policy.

Diderot's imaginary *Supplément au voyage de Bougainville* (1796) describes and extols the primitive customs of Tahiti. Unlike Rousseau's, Diderot's "primitivism" was not a plea for a return to a less civilized society. Not nature or natural law, but the fundamental laws of nature, were uppermost in Diderot's mind. The conventions of modern society, it seemed to him, unnecessarily restricted the basic biological needs of man. Before Sigmund Freud, he sensed the dangers of sexual repression, a theme developed in the final section of *Rêve de d'Alembert* and fundamental to his novel *La religieuse.* Celibacy, in his view, led too often to mental or sexual aberration. He ended his Tahitian tale, however, with the admonition that, though we should try to change bad, or "unnatural," laws, we must obey the laws that our society has imposed.

Diderot frankly admitted his enjoyment of sensual pleasures—books, women, pictures, friends, and toasting his toes before a fire. But in the preface of *Le père de famille,* addressed to the princess of Nassau, he declared that "he who prefers a voluptuous sensation to the conscience of a good act is a vile man." He felt certain that through education and knowledge we could recognize what was good, and that virtue, or beneficence, was the one and only path to happiness. There are intimations in his works of a belief that the good and wise man, in a corrupt society, should at times rise above a bad law, a theme illustrated in his last play, *Est-il bon? Est-il méchant?*

Toward the end of his life, in his praise of Seneca, he extolled the Stoic concept of virtue as its own reward. He summed up his natural, humanistic ethics in a brief pronouncement: "There is only one virtue, justice; one duty, to be happy; one corollary, neither to overesteem life nor fear death."

AESTHETICS

In the theory and practice of the arts dependent on the imagination—literature, music, and the fine arts—Diderot also introduced innovations. His approach to the theory of Beauty was through the perception of relationships and the arts of communication. An unusual perception of relationships, through analogy and associative memory, was the mark of the genius, whether scientist or poet. The artist first experiences an emotional or aesthetic stimulus strong enough to fire his imagination. A second moment of enthusiasm, which comes from the ability to communicate his vision through his special technique, sounds, colors, lines, or words, is essential, however.

His *Encyclopédie* article "Beau" (1751) gave evidence of a thorough acquaintance with French and English aestheticians. That same year he launched out on his own in his *Lettre sur les sourds et muets.* Here he discussed the importance of gesture and expression in communication. The great actor is one who paints in gestures what he expresses in words, just as the great poet paints in sounds and rhythms what he means in words. Likewise, the beauty of a painting depends on its inner rhythm and structure. The sublime in painting and poetry is derived from the emotions imparted through the harmonies of sound and color, the wedding of sense and sound. Poetry, he declared, is therefore essentially untranslatable.

ART AND MORALITY. A strong moralistic tone pervaded Diderot's aesthetic theories and criticism. The painter must have morals as well as perspective. The bourgeois drama, a genre that he originated and illustrated, though not very successfully, should compete with the law in persuading us to love virtue and hate vice.

There was more than a touch of sentimentality in the art criticisms of the *Salons,* which he wrote biennially from 1759 to 1781. For a period, the bourgeois pathos of Jean-Baptiste Greuze held a strong appeal for him. A notable connoisseur of the arts, he was not, however, fooled. He recognized the masterly compositions of François Boucher, but condemned his allegorical subjects and depiction of the loves of the gods. Pierre Teilhard de Chardin's use of color, he knew, was far superior to that of Greuze, though his subject matter was too often "ignoble." Yet Teilhard de Chardin taught him that painting was not, as the classical theorists long held, the imitation of beautiful nature. He stood in awed amazement before

Teilhard de Chardin's painting of the skate and called it magic.

CRITICISM. Diderot created modern art criticism as a literary art. The *Salons,* especially of 1765 and 1767, still make fascinating reading and contain the best of his literary criticism. That he was himself a great writer is now at last being generally recognized. First and foremost, he was a master of dialogue; written for the ear, his dialogues are artistic transpositions of reality. His dislike of abstractions made him an early champion of realism. He never ceased to admire Molière and Jean Racine—and William Shakespeare—but believed that the theater was destined to follow new paths. His romantic spirit was revealed by his advocacy of strong emotions and his streak of sentimentality. He therefore foreshadowed the romantic-realistic revolt against classicism, delayed in France until the nineteenth century by the political revolution.

Diderot's trinity was truth, goodness, and beauty. In his aesthetic order, first place was given to that which was both useful and agreeable; second, to the merely useful; and third, to the purely agreeable. Since the essence of the arts was not subject matter, but the perception and communication of relationships, he felt it was advantageous to add a moral subject, the useful, to technical beauty.

SOCIETY AND POLITICS

Diderot made his *Encyclopédie* a major weapon for upsetting the social and political institutions of the Old Regime. In the first volume his article "Autorité politique" boldly proclaimed, before Rousseau's *Contrat social,* that sovereignty resided in the people, who alone should determine how and to whom it should be delegated. There, too, appeared the first discussion of the "general will." In an often vain effort to evade censorship, he chose out-of-the-way places, sometimes seemingly harmless definitions of terms, to point out the danger that lay before both the state and the church unless they were strictly separated.

In his *Observations* on the instructions of Catherine II to her deputies in the recodification of Russian law, he was even more forthright: "The only true sovereign is the nation," he wrote; "there can be no true legislator except the people." He also chided Catherine for submitting political institutions to religious sanction: "Religion is a support that in the end almost always ruins the edifice." He did not hesitate to call her a tyrant and refuted her arguments in favor of benevolent despotism. Her suppression of his manuscript was so thorough that parts of it were coming to be known only in the twentieth century.

Rameau's Nephew was a sweeping satire of French eighteenth-century society, especially of the often ignorant and very wealthy general tax collectors, who, with their hordes of parasites, were a menace to the development of the arts, as well as powerful enemies of the *Encyclopédie.* In a dialogue with Diderot, the parasitic nephew of the great Jean-Philippe Rameau defended his debasement and moral corruption, quite shocking to his moralistic interlocutor, as the only means of satisfying the pangs of hunger in a thoroughly corrupt society. Throughout Diderot's works—in his dramas, his short stories and novels, in his art and literary criticism, as well as in his social and political theories—his sympathies were with the Third Estate.

Because he was forced to withhold his best and most forthright works for publication by future generations, the growth of Diderot's fame has been a very slow process. Rousseau declared that it would take two centuries for the realization that he was the great genius of his century. His first enthusiasts were also men of genius, Johann Wolfgang von Goethe, Honoré de Balzac, Charles-Pierre Baudelaire, and Victor Hugo.

It can hardly be a cause for wonder that Diderot is receiving special attention in Marxist societies and that many excellent editions and translations have come from Marxist presses. Yet it was to the scientist and philosopher in Friedrich Engels, rather than the social economist, that Diderot's work most greatly appealed. His philosophical determinism was in no sense economic determinism; his sturdy bourgeois qualities give small comfort to Marxist sociology; and his views of the importance of hereditary traits are in sharp opposition to behavioristic theory. He would seem to qualify most readily as a naturalistic humanist.

See also Alembert, Jean Le Rond d'; Atheism; Buffon, Georges-Louis Leclerc, Comte de; Clandestine Philosophical Literature in France; Condillac, Étienne Bonnot de; Darwin, Charles Robert; Deism; Descartes, René; Doubt; Empiricism; Encyclopédie; Engels, Friedrich; Enlightenment; Ethics, History of; Freud, Sigmund; Goethe, Johann Wolfgang von; Helvétius, Claude-Adrien; Holbach, Paul-Henri Thiry, Baron d'; La Mettrie, Julien Offray de; Lavoisier, Antoine; Leibniz, Gottfried Wilhelm; Locke, John; Lucretius; Marxist Philosophy; Materialism; Maupertuis, Pierre-Louis Moreau de; Pascal, Blaise; Rousseau, Jean-Jacques; Scientific Method; Seneca, Lucius Annaeus; Sensationalism; Stoicism; Teilhard de Chardin, Pierre; Toland, John; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY DIDEROT

Oeuvres complètes. 20 vols, edited by J. Assézat and M. Tourneux. Paris: Garnier, 1875–1877.

Oeuvres, edited by A. Billy. Paris, 1935. Essential works in one volume.

Correspondance, edited by G. Roth. Paris, 1955–.

Le rêve de d'Alembert, edited by Jean Varloot. Paris, 1962.

English Translations

Diderot's Early Philosophical Works. Translated and edited by Margaret Jourdain. Chicago and London: Open Court, 1916.

Dialogues. Translated and edited by Francis Birrell. New York: Brentano, 1927.

Diderot, Interpreter of Nature; Selected Writings. Translated and edited by Jean Stewart and Jonathan Kemp. London: Lawrence and Wishart, 1937. Best translations of *Rêve de d'Alembert,* etc.

Rameau's Nephew and Other Works. Translated by Jacques Barzun and Ralph H. Bowen. Garden City, NY: Doubleday, 1956.

Jacques the Fatalist and His Master. Translated by J. Robert Loy. New York: New York University Press, 1959 and 1961.

Selected Writings. Translated and edited by Lester G. Crocker. New York: Macmillan, 1966.

Diderot: Political Writings, edited by John Hope Mason and Robert Wokler. New York: Cambridge University Press, 1992.

WORKS ON DIDEROT

Anderson, Wilda. *Diderot's Dream.* Baltimore: Johns Hopkins University Press, 1990.

Blum, Carol. *Diderot: The Virtue of a Philosopher.* New York: Viking Press, 1974.

Bonneville, Douglas A. *Diderot's Vie de Seneque: A Swan Song Revised.* Gainesville: University of Florida Press, 1966.

Bremner, Geoffrey. *Order and Chance: The Pattern of Diderot's Thought.* Cambridge, U.K.: Cambridge University Press, 1983.

Cabeen, D. C., ed. *A Critical Bibliography of French Literature.* Vol. IV. Syracuse, NY: Syracuse University Press, 1951. Especially valuable for researchers.

Clark-Evans, Christine. "Charles de Brosses and Diderot: Eighteenth-Century Arguments concerning Primitive Language, Particular Natural Languages and a National Language." *History of European Ideas* 16 (1–3) (1993): 183–188.

Creech, James. *Diderot: Thresholds of Representation.* Columbus: Ohio State University Press, 1986.

Crocker, Lester G. "Diderot as Political Philosopher." *Revue Internationale de Philosophie* 38 (1984): 120–139.

Crocker, Lester G. *Diderot's Chaotic Order: Approach to Synthesis.* Princeton, NJ: Princeton University Press, 1974.

Crocker, Lester G. *The Embattled Philosopher.* East Lansing: Michigan State College Press, 1954. Good general introduction to Diderot's life and works.

Davis, Colin. "Backward, Forward, Homeward: Encounters in Ithaca with Kant and Diderot." In *Proximity: Emmanuel Levinas and the Eighteenth Century,* edited by Melvyn New. Lubbock: Texas Tech University Press, 2001.

Dynnik, M. A. "On the Esthetics of Diderot." *Soviet Studies in Philosophy* 3 (1964–1965): 48–53.

Edmiston, William F. *Diderot and the Family: A Conflict of Nature and Law.* Saratoga, NY: Anma Libri, 1985.

Eisenberg, Jose. "The Theater and Political Theory in Rousseau and Diderot." *Kriterion* 41 (101) (2000): 86–108.

Fellows, Otis E. et al., eds. *Diderot Studies.* Geneva: Librairie Droz, 1949–. Critical essays and monographs in English and French by contemporary scholars. Five volumes had been published as of 1965.

France, Peter. *Rhetoric and Truth in France: Descartes to Diderot.* London: Clarendon Press, 1972.

Holt, David K. "Denis Diderot and the Aesthetic Point of View." *Journal of Aesthetic Education* 34 (1) (2000): 19–25.

Ibrahim, Annie. "The Life Principle and the Doctrine of Living Being in Diderot." *Graduate Faculty Philosophy Journal* 22 (1) (2000): 107–122.

Lough, John. *Essays on the "Encyclopedie" of Diderot and d'Alembert.* London: Oxford University Press, 1968.

Luxembourg, Lilo K. *Francis Bacon and Denis Diderot, Philosophers of Science.* New York: Humanities Press, 1967.

Niklaus, Robert. "Denis Diderot: Search for an Unattainable Absolute of Truth." *Ultimate Reality and Meaning* 3 (1980): 23–49.

Pucci, Suzanne L. *Diderot and a Poetics of Science.* New York: Peter Lang, 1986.

Rey, Roselyne. "Diderot and the Medicine of the Mind." *Graduate Faculty Philosophy Journal* 22 (1) (2000): 149–159.

Richards, Joyce A. *Diderot's Dilemma: His Evaluation Regarding the Possibility of Moral Freedom in a Deterministic Universe.* New York: Exposition Press, 1972.

Schmidt, James. "The Fool's Truth: Diderot, Goethe, and Hegel." *Journal of the History of Ideas* 57 (4) (1996): 625–644.

Simon, Julia. *Mass Enlightenment: Critical Studies in Rousseau and Diderot.* Albany: State University of New York Press, 1995.

Strugnell, Anthony. *Diderot's Politics: A Study of the Evolution of Diderot's Political Thought after the Encyclopedie.* The Hague: Nijhoff, 1973.

Vasco, Gerhard M. *Diderot and Goethe: A Study in Science and Humanism.* Geneva: Slatkine, 1978.

Waldauer, Joseph L. *Society and the Freedom of the Creative Man in Diderot's Thought.* Geneva: Droz, 1964.

Wilson, Arthur M. *Diderot: The Testing Years (1713–1759).* New York: Oxford University Press, 1957. Best biography to date and best critical studies of early works. The first of two volumes.

Norman L. Torrey (1967)

Bibliography updated by Tamra Frei (2005)

DIKĒ

Dikē is the old Greek word for "law, justice." By the fourth century BCE it was largely replaced by its cognate *dikaiosynē*, Plato's cardinal virtue, justice.

In early Greece (Homer, Hesiod), *dikē* ranges in meaning from a specific claim by one party to a dispute, to a judgment or settlement, or to the personified force or goddess Justice/Law. In Homer's *Iliad*, the trial scene on Achilles's shield (18.497–508) depicts the elders (as judges) in a competition to see who can propose the straightest *dikē* (the best judgment/settlement). In Hesiod's *Works and Days* animals eat one another, but Zeus gave humans *dikē*—law, judicial process—which is far better (276–280), and *Dikē* sits beside her father Zeus and punishes those who corrupt the judicial process with crooked *dikē* (256–262).

The sixth-century lawgiver Solon promotes *dikē*—law-abiding conduct—as part of a general program of *eunomia* (good order, law and order). He also speaks of his legislation as providing a straight *dikē* (judicial process) for every Athenian. For the fifth-century thinker Heraclitus, *dikē* becomes a cosmic force of order and balance. Heraclitus's *dikē* is not static, however, but—as in a lawsuit—a balance of opposing forces, so that, as he says paradoxically, *dikē* is *eris* (strife).

Fifth-century tragedians regularly see *dikē* as a cosmic force, justice, largely in the sense of punishment or retribution. All the characters in Aeschylus's *Oresteia* claim to seek *dikē*—justice—primarily in the sense of punishment or revenge for previous wrongs, though in some passages the chorus suggest a larger sense of justice as cosmic and social order. Plato's *Protagoras* pictures the sophist Protagoras telling a story in which the gods give *dikē*, law or justice, together with *aidōs* (respect) to all humans; he concludes from this that *dikē* is necessary for the survival of human society.

Through the fifth century, *dikē* in all its meanings—from judicial process to cosmic force—remains something external to human beings. Not until the fourth century does Plato make justice a personal virtue of individuals, and then it is no longer *dikē* but *dikaiosynē* (see especially *Republic*, Book IV).

See also Justice; Plato.

Bibliography

Gagarin, Michael. "*Dikē* in the *Works and Days*." *Classical Philology* 68 (1973): 81–94. Argues that in the beginning meanings of *dikē* are restricted to the realm of law and the legal process.

Havelock, Eric A. *The Greek Concept of Justice*. Cambridge, MA: Harvard University Press, 1978. Argues that in a preliterate, oral culture, justice is primarily procedure.

Lloyd-Jones, Hugh. *The Justice of Zeus.* 2nd ed. Berkeley: University of California Press, 1983. Broad discussion of divine justice—often seen as equivalent to *dikē*—in Greek literature.

The works of Hesiod, including his poem*Works and Days*, are available in the Penguin series and several other translations.

Michael Gagarin (2005)

DILTHEY, WILHELM

(1833–1911)

The German philosopher and historian Wilhelm Dilthey was born in Biebrich on the Rhine, the son of the preacher to the Duke of Nassau. He studied theology and philosophy in Heidelberg and Berlin and combined both of these interests in his early work on the ethical and hermeneutical writings of Friedrich Schleiermacher. Dilthey's first major publication, a volume on the life of Schleiermacher, appeared in 1870 while he was teaching in Kiel. In 1871, Dilthey received a professorship in Breslau (now Wrocklaw, Poland). It was around this time that he met Count Yorck of Wartenburg, and their friendship produced an intellectual correspondence about the nature of life and the meaning of history that has inspired thinkers such as Martin Heidegger and Hans-Georg Gadamer. In 1882, Dilthey was called back to Berlin to fill the chair that George Wilhelm Friedrich Hegel had once held. The University of Berlin and the Prussian Academy would be the locus of his world for almost thirty years, until his death in 1911. This is the period in which he published most of his writings about the human sciences (*Geisteswissenschaften*), a covering term for both the humanities and social sciences. These writings consider how the human sciences contribute to the understanding of life and history.

CRITIQUE OF HISTORICAL REASON

Dilthey saw his overall project as a Critique of Historical Reason examining the conditions that make possible the respective cognitive results of the natural and the human sciences. Although influenced by both Immanuel Kant and Hegel, he rejected the transcendental and formal limits of the former and the metaphysical absolutes of the latter. His task was to translate the insights of idealism into a more open empirical approach to what it means to experience reality.

Although the natural sciences are about nature and the human sciences about history, this does not justify hypostatizing history as a spiritual domain separate from nature. The spiritual life of human beings is conditioned—but not determined by—natural processes. Even

when human beings set themselves free purposes, the realization of these purposes requires that the laws of nature be obeyed. In Book 1 of his *Introduction to the Human Sciences* (1883), Dilthey grants the human sciences a relative cognitive independence from the natural sciences. Yet he assigns the human sciences a greater reflective scope in that they express more aspects of human experience. They not only ascertain what is—as do the natural sciences—but also make value judgments, establish goals, and prescribe rules.

For the human sciences, theory is always framed by practical considerations instigated by historical life. Therefore, philosophical reflection about their conditions of possibility makes it necessary to regress behind the logical and epistemological foundations of the natural sciences to establish the more encompassing life-nexus of all human experience. This reflective turn initiated in Book 2 of the *Introduction to the Human Sciences* and worked out in the posthumously published drafts for Book 4, shows the human sciences to have an important advantage over the natural sciences in that they preserve some of the intuitive access to the reality of experience as it is lived. The natural sciences merely construct a phenomenal or ideal world that abstracts from the overall nexus of life so that human beings stand as impartial intellectual observers of this abstractly represented nature.

By contrast, the world that is formed by the human sciences is the historical-social reality in which human beings participate. It is a fuller world that is accessible not merely as conceptually mediated cognition (*Erkenntnis*), but also as immediate knowledge (*Wissen*) found in lived experience. Conceptual cognition is representational and objectifying. Lived experience provides a prerepresentational self-presence that involves a direct knowing. Any state of consciousness is implicitly present to itself in what Dilthey calls "reflexive awareness" (*Innewerden*). This does not require an explicit consciousness of being conscious—such an act of self-consciousness would be more than reflexive, namely, reflective. At the basic level of reflexive awareness there is not yet a self as an object of reflection.

According to Dilthey, there is no self underlying consciousness. Instead, the self arises out of consciousness as the correlate of the world. Within the nexus of consciousness as a function of life, reflection can differentiate between facts of inner perception and facts of outer perception, thereby producing a distinction between self and world. This world is not a product of an inference, but is felt primarily through resistance to the practical impulses of the will. Rather than grounding the objectivity of the world on a transcendental "I think," Dilthey claims that its reality is given in the reflexive awareness of the relation between efficacy and resistance involved in willing. Through this expanded reflexive awareness, the life-nexus in which the self participates discloses things and other selves that can resist its will. These modes of reflexive awareness are as basic to Dilthey's theory of hermeneutical understanding (*Verstehen*) as the transcendental and empirical ego were to Kant's theory of intellectual understanding (*Verstand*). Whereas Kant sought an explanative mode of understanding for natural phenomena by deriving them from the most general laws of scientific cognition, Dilthey seeks to understand the meaning of things in terms of their own inherent context. Hermeneutical understanding provides a kind of situated understanding that receives its bearings from the reflexive awareness of lived or prescientific experience.

DESCRIPTION AND STRUCTURAL UNDERSTANDING

In 1894, Dilthey published another important work, the *Ideas for a Descriptive and Analytic Psychology* (Dilthey 1977). Here he works out the implications of his philosophical views about lived experience for psychology as a human science. Hitherto, psychology had been treated as a kind of natural science that synthetically constructs mental phenomena from atomistic elements such as sense-data by using hypothetical laws of association. This assumes that psychic life comes in discrete states that must be connected. Dilthey argues, however, that psychic life presents itself as a continuum in which states are already connected. It is the task of psychology to attempt to describe this general nexus of psychic life and to analyze specific states on its basis.

Dilthey's descriptive and analytic psychology has three main parts. The first delineates the general structural systems of consciousness that can be differentiated at the levels of cognition, feeling, and volition. The cognitive system relates the acts of perception, imagination, and memory on the basis of which we conceptually represent the world. The felt and instinctual aspects of consciousness can be related to form a distinct structural system whereby we coordinate the value of things. A volitional structural system functions to link and rank the purposes we set. A cross-sectional analysis of any lived experience will manifest aspects of each of these three functional structures. Indeed, the structural systems manifest a degree of interdependence belying the traditional hierarchical assumption that the cognitive level is fundamental and that feeling and willing merely respond

to what has been perceived. Thus we do not perceive impressions of sense unless there is a felt interest in them and the will is stirred enough to attend to them.

The second main part of psychology as a human science traces the development of psychic life. It examines how psychic structures are defined and articulated over time. Here Dilthey stresses the importance of treating each phase in the teleological development of a psychic life-course as having its own inherent worth. Every phase has its immanent purposiveness and is to be treated as a kind of epoch. Although an epochal phase may contribute to its successor, it should never be treated as a mere means. The values of childhood, for example, should never be sacrificed for the goals of adulthood.

The third, concluding part of Dilthey's descriptive and analytic psychology integrates these structural and developmental approaches by showing how an acquired psychic nexus is gradually produced and informs future experiences. The acquired psychic nexus becomes the individualized framework according to which each self tends to specify its own experiences. It provides a historicized apperceptive mass that influences what will be perceived. It is like an implicit worldview that can regulate further experiences and actions.

Dilthey initially formulated his conception of the acquired psychic nexus as part of an effort to understand artistic creativity. In his 1887 essay "The Imagination of the Poet: Elements for a Poetics" (Dilthey 1985), Dilthey argues that what distinguishes artists from other human beings is the capacity to articulate their acquired psychic nexus in typical ways. In ordinary life, our experience and behavior reflect contingent local conditions as well as our acquired psychic nexus. Playwrights and novelists can establish fictional contexts that limit the extent to which characters will be distracted by local contingencies. By more adequately reflecting the acquired psychic nexus of their creators, the actions of fictional characters can also address more general aspects of life. The literary imagination produces typical situations and characters that help focus the meaning of human existence. Individuals manifest creativity when the perspective that informs their acquired psychic nexus becomes more than regulative, but constitutively typical.

The self-givenness of reflexive awareness and the self-presence of lived experience provide an implicit kind of understanding of life that psychological description and literary expression can make explicit. The inherent connectedness of consciousness renders it unnecessary to introduce hypothetical explanative links into the foundation of psychology. On this basis, Dilthey claims that the natural sciences are mainly about causal explanation and the human sciences about description and structural understanding. But this contrast is not absolute. Sometimes natural sciences must be content with description and interpretation, and sometimes human sciences cannot rely on general descriptions to account for significant details and must appeal to hypotheses. The difference is that the natural sciences tend to begin with explanative hypotheses, whereas the human sciences may end up with explanative hypotheses.

HERMENEUTICS

Unlike the natural sciences, the human sciences do not abstract from ordinary life, but analyze it. Analysis is compatible with understanding because, unlike abstraction, it need not isolate things from their overall context. The hermeneutical task of analysis is to enable us to recognize the whole in its parts and the parts in the whole. There is always this circularity in coordinating parts and wholes when reading a text. Hermeneutics as a human science reflects on what it means to apply the art of exegesis from texts to the experience of life in general.

The essay "The Rise of Hermeneutics," published in 1900 (Dilthey 1996), represents an important phase in Dilthey's development. Here he begins to sketch out a position that would define his final work. While he does not abandon the project of describing and analyzing lived experience, he came to view description and analysis as limited in their ability to capture the full meaning of life. The inner connectedness of our own experiences may provide a kind of self-understoodness or self-evidentness (*Selbstverständlichkeit*), but we do not achieve real self-understanding (*Selbstverständnis*) until we have manifested ourselves objectively. To truly understand ourselves is to be able to see ourselves as others see us.

One of the most revealing ways in which we manifest ourselves is through linguistic expression and communication. But Dilthey defines hermeneutics as the theory of interpreting *all* human manifestations, including actions that are not intended to communicate. The range of objectifications needing interpretation is broad. It includes impersonal theoretical judgments, abstract mathematical formulas, concrete poetic expressions of lived experience, personal correspondence, journal entries, works of art, historical monuments and archives, and political deeds and their aftereffects. They are important because only that which is publicly accessible and has been objectified in a common medium can produce determinate meaning.

The work that best articulates this hermeneutical approach to the human sciences is *The Formation of the Historical World in the Human Sciences* (1910). This most mature formulation of Dilthey's Critique of Historical Reason revisits many of the themes of the *Introduction to the Human Sciences*. The human sciences form the historical world, not by producing it, but by giving it a multifaceted discursive shape. Determinate meaning will never be found by confronting the course of history monolithically. The human sciences can give a cognitive form to various strands of history that we knowingly participate in. They allow use to analyze the overall stream of history and direct it, as it were, into a variety of structural systems in which selected currents can be examined for specific interacting forces.

Some of these historical structures had already been identified in the *Introduction to the Human Sciences* as cultural and social organizational systems. Cultural systems were conceived as purposive systems that bring individuals together to achieve certain voluntary goals. These purposive systems are not limited to the goals of high culture—the sciences, the arts, and religion—for they also include economic and social cooperation. Dilthey distinguished these cultural systems from institutional structures which make up the external organization of society. Institutions such as families, tribes, and nation-states are also interactive, but not primarily voluntary. We do not choose our parental family but are born into it. One of the advances of *The Formation of the Historical World* is that all these historical structures are no longer subsumed under the concept of "purposive system." Dilthey introduces the covering term "productive system" (*Wirkungszusammenhang*) to capture the ways in which the forces of historical life can become structurally organized. The efficacy of history is to be understood in terms of productivity before any causal or teleological account is given. The carriers of history, whether they be individuals, cultures, institutions, or communities, can all be considered as productive systems capable of exerting influence, and in some cases, realizing purposes. Each productive system of history should be approached as being centered in itself.

Individuals too are productive systems when they appropriate new impressions into their acquired psychic nexus: They cognize the present on the basis of past evaluations and future goals. The productivity of the psychic nexus lies in the ways the cognitive, evaluative, and volitional aspects of experience interact. As productive systems, individuals are centered in themselves, but far from self-sufficient. They are also dependent on other more inclusive productive systems. In the *Introduction to the Human Sciences*, Dilthey was unwilling to conceive these larger systems as subjects or carriers of history. In *The Formation of the Historical World* he qualifies his opposition to transpersonal subjects by treating them as logical rather than real subjects—they are now considered co-carriers of history. Although individuals cooperate in terms of cultural systems and other encompassing productive systems, they never engage more than a part of themselves to any of such systems and therefore cannot be defined by them. Yet the engagement can become so intensive that an individual can put his or her stamp on its mode of productivity. As a consequence, more than the agreed-upon functions of a cultural system will be achieved. For instance, in relation to the classical conventions established by Joseph Haydn (1732–1809) and Wolfgang Amadeus Mozart (1756–1791), a composer such as Ludwig van Beethoven (1770–1827) charts a new course. As a consequence, more than the expected purposes of the system will be achieved. In addition to accommodating new purposes, productive systems provide a meaning framework for expressing a variety of human values.

Dilthey states that he is not offering a philosophy of history that would establish a final purpose of human history. This is because he does not find any justification for the belief that there is a law of overall historical development. Yet there is good reason to think that there can be lawlike development within specific productive systems. Dilthey's theory of history is meant to provide the critical tools to articulate history into the productive systems that can provide an orderly understanding of history. Today, Dilthey's approach would be considered a philosophy of history of the critical rather than of the more traditional speculative kind.

THE CATEGORIES OF THE HUMAN SCIENCES

Whereas Kant's *Critique of Pure Reason* defined the categories or fundamental concepts of the natural sciences, Dilthey set out to explicate the categories of the human sciences. He distinguishes between formal and real categories. Formal categories relate to all experience, whether it be prescientific or scientific. They arise from elementary operations of thought such as comparing, differentiating, and relating that bring out what is inherent in experience. The formal categories of unity and plurality, identity and difference are shared by the natural and human sciences.

Real categories organize the content of experience more concretely. The natural and human sciences both organize their subject matter in terms of formal part-whole relations and locate them in space and time. In temporal location we can see a transition from the formal to the real. For the natural sciences, time is an infinite form that unfolds uniformly. For the human sciences, time is a finite structure that projects the future based on what is remembered from the past. The time of the human sciences is a lived reality and can be articulated in ways that allow us to understand historical development and the productive force of cultural systems.

Causality is a real category of the natural sciences. While Dilthey does not rule out its applicability to the events that are recounted in human history, he makes it clear that for the understanding of history, the Aristotelian categories "of agency and suffering, of action and reaction" are more appropriate (Dilthey 2002, p. 219). They express how human beings experience the productive force of the historical world and allow them to conceive purposiveness as an agency that stems from within and causality as a force coming from without.

Among the real categories that are distinctive for the human sciences, the three most important are value, purpose, and meaning. From the perspective of value, life is judged as a multiplicity of prized moments that can be juxtaposed. From the perspective of purpose, everything in a life-course tends to be subordinated to some future moment. According to Dilthey, the category of meaning can overcome the juxtaposition and subordination of value and purpose. Meaning articulates the connectedness of life on the basis of the relation between past and present. It is the main category of historical thought and is assigned to memory.

We resort to memory when we orient our experience to the past. On the private level, Dilthey had articulated meaning in terms of the workings of the acquired psychic nexus. At the public level, Dilthey now explicates meaning in terms of Hegel's concept of "objective spirit." Objective spirit stands for what the spirit of the past has left behind in the present and has preserved in objective form. It is the most basic framework for orienting us to the past. Objective spirit is the tradition-based sphere of commonality in which we grow up. The language we inherit, the conventions adopted, and the customs learned are all aspects of objective spirit that shape our childhood experiences. "Everything in which spirit has objectified itself contains something that is common to the I and the Thou. Every square planted with trees, every room in which chairs are arranged, is understandable to us from childhood because human tendencies to set goals, produce order and define values in common have assigned [them] a place…" (Dilthey 2002, p. 229).

Objective spirit represents the initial framework of reference for elementary understanding, not unlike the way a dictionary serves as our first resource when a word in a sentence is not understood. Objective spirit is the common historical medium by which we orient elementary understanding. But when problems arise in understanding that a common reference cannot resolve, we must resort to what Dilthey calls "higher understanding." Higher understanding attempts to account for cases when the normal convergence between an expression and the meaning it expresses is lacking. Instead of merely appealing to objective spirit as the common background for locating meaning, higher understanding can consider more specialized contexts to determine meaning. Thus, if an unclear sentence is uttered by an economist we can consult professional handbooks. Similarly, social circumstances, industrial conditions, and market forces can be considered when some economic claim is not fully intelligible.

Although higher understanding often concentrates on more restricted productive systems as focal contexts, it will at the same time seek to extract more general results. The universality aimed at by higher understanding may be in the form of an inductive generalization or it may be that of a larger context. Thus the attempt to understand a line of poetry in relation to the poem as a whole is also an act of higher understanding. Here again the attempt is to move from common meaning to universal significance. The important breakthrough for Dilthey is that he no longer requires the understanding of human products to be related back to the psyches of their producers. Although the possibility of referring a work of art to its creator is not ruled out, it is far from being the primary source of its understanding. Indeed, a great work of art can take on a life of its own and can become itself a productive nexus generating an ever deeper meaning over time, as Gadamer has also argued.

Historical understanding, however, requires the move from universality back to individuality. It is appropriate for higher understanding to turn into what Dilthey calls a "re-experiencing," where individual contributions to the productivity of life do count. To re-experience meaning is not to reproduce the state of mind of an author, but to understand an author better than he understood himself. This is achieved by the contextualizing and structural explication of life-situations made possible by the human sciences.

REFLECTION ON LIFE

It is never enough to consider an individual life by itself. As Dilthey writes: "The limit of biography lies in the fact that general movements find their point of transition in individuals" (Dilthey 2002, p. 269). Drawing on his own struggles to complete a second volume of the life of Schleiermacher, Dilthey concludes that a biographer cannot fulfill his task without also having broached universal questions about life and history. Notwithstanding the problematic status of biography, Dilthey considers autobiography an especially instructive mode of history because here "the work of historical narrative is already half done by life itself" (Dilthey 2002, p. 222). The narrative produced is never a simple copy of an actual life-course, but a retrospective judgment that depends on the way an individual reflects on his or her life. Here history is not just a human science but has reflective philosophical import.

In the later writings Dilthey often speaks of anthropological reflection as crucial for obtaining a unity of perspective on life. The sciences are radically pluralistic and cannot provide a comprehensive outlook or worldview (*Weltanschauung*). A worldview is not merely a cognitive picture of the world. It goes deeper in expressing a specific stance (*Stellung*) toward concrete life-concerns (*Lebensbezüge*) as well as to life as a whole. An individual's stance toward life can develop into a reflective worldview on the basis of certain more general moods (*Stimmungen*). These moods are more than states of mind; they orient us to the world in ways that anticipate what Heidegger says about moods as modes of attunement in *Being and Time*.

Worldviews have been articulated in literary, religious, and philosophical works. Philosophers have conceptualized worldviews metaphysically. Dilthey analyzes three main types of such metaphysical formulations: naturalism, the idealism of freedom, and objective idealism. Naturalism as found in Democritus, Thomas Hobbes, and others reduces everything to what can be cognized and is pluralistic in structure; the idealism of freedom as found in Plato, Kant, and others insists on the irreducibility of the will and is dualistic; objective idealism as found in Heraclitus, Gottfried Wilhelm Leibniz, and Hegel affirms reality as the embodiment of a harmonious set of values and is monistic. The three types of metaphysical worldviews are incommensurable in that each is reductive in some way. No metaphysical formulation can have more than relative success. But this conclusion does not make Dilthey a relativist, for he rejects all metaphysics as speculative. Metaphysical systems attempt to arrive at universal determinations that transcend experience. All that is humanly possible is to probe reality on the basis of life-experience and to seek a more limited reflective universality.

The influence of Dilthey's thought and writings is manifold. Husserl considered Dilthey's *Ideas for a Descriptive and Analytic Psychology* (Dilthey 1977) a genial anticipation of his own phenomenological psychology and credits a meeting with Dilthey as leading to his interest in questions concerning understanding in the human sciences. Heidegger's lecture courses from 1919 through 1925 are filled with declarations of Dilthey's importance for understanding history and make extensive use of such Diltheyan terms as "life-nexus" and "life-concern." Max Weber applies Dilthey's distinction between explanation and understanding to sociology and extends Dilthey's reflections on typicality to his theory of ideal types. Herbert Marcuse's early work on Hegel is indebted to Dilthey's highly original approach to Hegel in his *Jugendgeschichte Hegels*. Georg Lukács's Marxist counterpart to this is *Der junge Hegel*.

Dilthey's work continues to play a significant role in the development of hermeneutics. While critical of the Schleiermacher-Dilthey tradition, Gadamer's hermeneutics represents an extension of Dilthey's effort to relate interpretation to the productivity and efficacy (*Wirkung*) of history. In France, the underlying influence of Dilthey's views on understanding and objective spirit can be seen in the writings of Raymond Aron, Jean-Paul Sartre, Lucien Goldmann, and Paul Ricoeur. In Spain, Ortega y Gasset had called Dilthey the most important philosopher of the second half of the nineteenth century, with the result that Dilthey was widely translated into Spanish before any other language. Now extensive translations into English, French, Italian, Chinese, Japanese, and Russian are also becoming available.

See also Gadamer, Hans-Georg; Hermeneutics; Philosophy of History.

Bibliography

WORKS BY DILTHEY

Briefwechsel zwischen Wilhelm Dilthey und dem Grafen Paul Yorck von Wartenburg, 1877–1897. Halle (Saale), Germany: M. Niemeyer, 1922. Famous correspondence about the nature of life and history.

Descriptive Psychology and Historical Understanding. Translated by Richard M. Zaner and Kenneth L. Heiges, with an introduction by Rudolf A. Makkreel. The Hague: Martinus Nijhof, 1977.

Dilthey's Philosophy of Existence. Translated by William Kluback and Martin Weinbaum. New York: Bookman Associates, 1957. Mainly about the theory of worldviews.

Das Erlebnis und die Dichtung: Lessing, Goethe, Novalis, Hölderlin. Leipzig, Germany: Teubner, 1922. Influential literary essays, two of which are translated in *Poetry and Experience*.

Essence of Philosophy. Translated by Steven A. Emery and William T. Emery. Chapel Hill: University of North Carolina Press, 1954.

Gesammelte Schriften, edited by Karlfried Gründer and Frithjof Rodi. 24 vols. Göttingen, Germany: Vandenhoeck & Ruprecht, 1914–2004.

Selected Works, edited by Rudolf A. Makkreel and Frithjof Rodi. Vol. 1, *Introduction to the Human Sciences* (1989). Vol. 3, *The Formation of the Historical World in the Human Sciences* (2002). Vol. 4, *Hermeneutics and the Study of History* (1996). Vol. 5, *Poetry and Experience* (1985). Princeton, NJ: Princeton University Press.

WORKS ON DILTHEY

de Mul, Jos. *The Tragedy of Finitude: Dilthey's Hermeneutics of Life*. Translated by Tony Burrett. New Haven, CT: Yale University Press, 2004. A reconstruction of Dilthey's ontology of life, highlighting the interpretive character of human existence, contingency, and narrativity.

Ermarth, Michael. *Wilhelm Dilthey: The Critique of Historical Reason*. Chicago: University of Chicago Press, 1978. A comprehensive account of Dilthey's thought with a good historical background.

Makkreel, Rudolf A. *Dilthey: Philosopher of the Human Studies*. Princeton, NJ: Princeton University Press, 1975, 1992. A developmental examination of Dilthey's philosophy that focuses on its relation to Kant's first and third Critiques and highlights the role of reflection and judgment in historical understanding.

Makkreel, Rudolf A., and John D. Scanlon, eds. *Dilthey and Phenomenology*. Washington, DC: Center for Advanced Research in Phenomenology and University Press of America, 1987. Exploration of Dilthey's relation to phenomenology by ten international scholars.

Owensby, Jacob. *Dilthey and the Narrative of History*. Ithaca, NY: Cornell University Press, 1994. A topical study focusing on Books 4–6 of the *Introduction to the Human Sciences*.

Revue Internationale de Philosophie 57 (4) (2003). Issue featuring a collection of essays edited by Rudolf Makkreel, including essays by Jean Grondin, Hans Ineichen, Matthias Jung, Makkreel, Sylvie Mesure, Jos de Mul, Tom Rockmore, and Frithjof Rodi.

Rodi, Frithjof, ed. *Dilthey-Jahrbuch für Philosophie und Geschichte der Geisteswissenschaften*. Vandenhoeck & Ruprecht, 1983–2000. Many volumes have special themes such as the relation between Dilthey and the early Heidegger.

Rodi, Frithjof, and Hans-Ulrich Lessing, eds. *Materialien zur Philosophie Wilhelm Diltheys*. Frankfurt am Main: Suhrkamp, 1984. A collection of classical essays on Dilthey by such thinkers as Scheler, Landgrebe, Bollnow, Plessner, Marcuse, Misch, Habermas, and Gadamer.

Rodi, Frithjof. *Das strukturierte Ganze: Studien zum Werk von Wilhelm Dilthey*. Göttingen, Germany: Hubert & Co., 2003. A series of essays stressing the structured nature of life and experience, and the importance of articulation and expression for Dilthey.

Rudolf A. Makkreel (2005)

DINGLER, HUGO

(1881–1954)

Hugo Dingler, the German philosopher of science, was the most important representative of Continental operationism, as distinguished from the operationalism of the American physicist P. W. Bridgman. Dingler was also a main contributor to *Grundlagenforschung* (research on the foundations of the exact sciences). After studying under such teachers as David Hilbert, Edmund Husserl, Felix Klein, Hermann Minkowski, Wilhelm Röntgen, and Woldemar Voigt at the universities of Erlangen, Munich, and Göttingen, Dingler received a Ph.D. in mathematics, physics, and astronomy in 1906 and became *Privatdozent* in 1912. He was appointed professor at the University of Munich in 1920 and at the Technische Hochschule in Darmstadt in 1932. In 1934 he was dismissed on charges of philosemitism. He later resumed teaching but soon rebelled again against the political situation, and eventually he was put under the continuous watch of a Gestapo agent "who unfortunately"—as Dingler told the present writer—"was not gifted for philosophy and did not profit from my compulsory daily lessons." Such difficulties in the German political situation during Dingler's life contributed to the lack of awareness of his work, despite his some twenty books and seventy essays in exceptionally clear German. Perhaps a more decisive factor was Dingler's independence of all the main schools and trends in contemporary philosophy of science—positivism and empiricism, Neo-Kantianism, phenomenology, intuitionism, and formalism.

From the juvenile *Grundlinien einer Kritik und exakten Theorie der Wissenschaften, insbesondere der mathematischen* (Essentials of a critique and rigorous theory of the sciences, especially of the mathematical ones; Munich, 1907) to the posthumous *Die Ergreifung des Wirklichen* (The grasping of reality; Munich, 1955), Dingler's main concern was to give a new answer to the Kantian question "How is exact science possible?" He regarded arithmetic, analysis, geometry, and mechanics as the exact sciences par excellence; he called them "mental" (*geistige*), meaning that they cannot be derived from experience and must be synthesized operationally from a few univocal ideas used as "building stones" (*Bausteine*). In this way

scientific inquiry was to be made continuous with everyday life and viewed in terms of practical activity. The operational reconstruction of the foundations of science was to abolish the field of foundations as an independent territory open to philosophical disagreement or mystification. Dingler came to consider the given itself, as expressed in protocol, or basic, sentences, as a highly complicated kind of result.

To prevent any residues of previous theories from entering into the operational reconstruction, we must start from a "zero situation" in which we suppose only that the world is "simply there" and that we can operate on it. This is a methodological principle, not a metaphysical denial of reality: it is a voluntary suspension of rational processes which can be brought about at any moment. After 1907, under Husserl's influence, Dingler labeled the zero situation "the standpoint of freedom from presuppositions." In 1942 he described it as *das Unberührte,* the intact or untouched—"that which has not yet been operated upon."

The first univocal step out of the zero situation consists in entertaining an idea in which the sheer relation of difference (with equality and similarity as its special cases) is present, and is applied (*anwendet*) only once, as in the idea "something distinct without further specification," that is, the idea of an entity as distinguished from all the rest, as standing out from a background. This idea is not the description of anything existing in the world but rather is the first requirement for any such description. All we can say about it is that it is present and limited; we can then specify it as constant or variable, and in either case we can also give special attention to its limits. In this way we reach a purely qualitative fourfold scheme which precedes the concepts of number, space, and time. To this scheme correspond four rules of operation, which afford the starting points of the exact sciences: (1) something distinct without further specification, and constant, for arithmetic; (2) the same, but variable, for analysis (more generally for the doctrine of time and variables); (3) the same, but constant, considered with respect to its limits, for geometry; and (4) the same, but variable, considered with respect to its limits, for kinematics and mechanics.

By means of complications of this basic scheme Dingler was able to operationally derive and prove the axioms of the exact sciences and to construct their whole fabric. This painstaking and original construction is to be found chiefly in *Philosophie der Logik und Arithmetik* (1931), *Die Grundlagen der Geometrie* (1933), *Die Methode der Physik* (1938), and *Lehrbuch der exakten Naturwissenschaften* (1944).

See also Bridgman, Percy William; Continental Philosophy; Hilbert, David; Husserl, Edmund; Operationalism; Philosophy of Science.

Bibliography

WORKS BY DINGLER

Works on "Grundlagenforschung"

Philosophie der Logik und Arithmetik. Munich, 1931.

Geschichte der Naturphilosophie. Berlin, 1932. A history of the development of the idea of *Grundlagenforschung* in experimental science.

Die Grundlagen der Geometrie. Stuttgart, 1933.

Die Methode der Physik. Munich, 1938.

Lehrbuch der exacten Naturwissenschaften. Berlin, 1944. Only thirty copies printed. Parts reprinted with Italian translation and a commentary by Enrico Albani in *Methodos* 7 (1955): 277–287, and 8 (1956): 29–30, 122–137, 191–199.

Other Works

Metaphysik als Wissenschaft und der Primat der Philosophie. Munich, 1926.

Das System. Munich, 1933.

Das Handeln im Sinne des höchsten Zieles. Munich, 1935.

Von der Tierseele zur Menschenseele. Leipzig, 1941.

Grundriss der methodischen Philosophie. Füssen, 1949. A crystal-clear summary of Dingler's main views, but of lower technical quality than his main treatises.

Essays

"Methodik statt Erkenntnistheorie und Wissenschaftslehre." *Kant-Studien* 41 (1936): 346–379.

"Über die letzte Wurzel der exakten Naturwissenschaften." *Zeitschrift für die gesamte Naturwissenschaft* 8 (1942): 49–70.

"Das Unberührte. Die Definition des unmittelbar Gegebenen." *Zeitschrift für die gesamte Naturwissenschaft* 8 (1942): 209–224.

"Die philosophische Begründung des Deszendenztheorie." In *Die Evolution der Organismen,* edited by Gerhard Hebener. Jena, Germany, 1943.

WORKS ON DINGLER

Benini, Giorgio. *I concetti fondamentali della filosofia metodica di Hugo Dingler.* Unpublished dissertation, Catholic University of the Sacred Heart, Milan, 1953.

Ceccato, Silvio. "Contra Dingler, pro Dingler." *Methodos* 4 (1952): 223–265, with English translation 266–290 and reply by Dingler, 291–296, translated into English 297–299.

Kramps, Wilhelm. *Die Philosophie Hugo Dinglers.* Munich, 1955. The main study.

Kramps, Wilhelm, ed. *Hugo Dingler Gedenkbuch zum 75. Geburtstag.* Munich: Eidos, 1956. Contains 14 essays by various authors and a bibliography.

Sandborn, Herbert. "Dingler's Methodical Philosophy." *Methodos* 4 (1952): 191–220.

Ferruccio Rossi-Landi (1967)

DIODORUS CRONUS

(b. 4th century BCE)

Diodorus Cronus was born in Iasus, a port town in Caria (a region in the southwestern part of Asia Minor). He inherited his nickname 'Cronus' (old fogy) from his teacher Apollonius. All else that is known about his life must be inferred from anecdotal evidence, connecting him to Athens, where Zeno of Citium studied dialectic with him (cf. Diogenes Laertius 7.25), and to Alexandria, where he is acquainted with the physician Herophilus (cf. Sextus Empiricus, *Pyrrh. Hypotyp.* 2. 245) and where Callimachus mentions him in one of his *Epigrams* (cf. Diogenes Laertius 2.111) suggesting that Diodorus was known in the town. He may have died in Alexandria, some time after 290 BCE.

Since our sources attribute no writings to him, he probably left nothing written. Yet the reports on him show that he was an extremely influential figure in the generation that saw the founding of the Hellenistic schools of philosophy. He belonged to a philosophical sect known as the Dialecticians; these Dialecticianswere a school distinct from the Megarians. The name Dialecticians was not, as assumed in the older literature, another name for the Megarians (Sedley 1977). In physics, Diodorus was an atomist; he is said to have called atoms "partless" (Sextus Empiricus, *Adv. Math.* 9.363). One consequence of his atomism is that there are, according to him, no objects that move, only objects that have moved (Sextus Empiricus, *Adv. Math.* 9.363).

Diodorus's greatest impact was in the field of logic where, together with his pupil Philo the Dialectician, he seems to have laid the foundations of propositional logic. With Philo, he engaged in a controversy about the truthcriteria for the conditional; Philo favored a truth-functional analysis of the conditional, claiming that the conditional is true if and only if it is not the case that its antecedent is true and its consequent false (cf. Sextus Empiricus, *Adv. Math.* 8.113–114), Diodorus gave a different account: According to him, a conditional is true if and only if it was not possible and is not possible that its antecedent is true and its consequent false (cf. Sextus Empiricus, *Adv. Math.* 8.115–117).

Diodorus's repute as a logician, even to the present day, derives from his Master Argument, mentioned by several authors but reported explicitly only in Epictetus (cf. Epictetus, *Diss.* 2.19.1–5). Diodorus claimed that the following three propositions are incompatible: (1) Every past truth is necessary, (2) nothing impossible follows from what is possible, and (3) there is something possible that neither is nor will be true. Diodorus used (1) and (2) to argue for the falsity of (3), hence for a notion of possibility that defines the possible as that which is or will be true. Here again we find him contradicted by Philo, who defines the possible as that which, by the intrinsic nature of the proposition, is receptive of truth (cf. Boethius, *De interpretatione* ii, 234,10–235, 9). The Master Argument became a bone of contention for Hellenistic logicians; it is still a matter of controversy how exactly Diodorus thought he could deduce the falsity of (3) from (1) and (2).

See also Atomism; Epictetus; Hellenistic Thought; Logic, History of; Megarians; Possibility; Zeno of Citium.

Bibliography

TESTIMONIA

The testimonia on Diodorus are now conveniently brought together in Volume 1 of G. Giannantoni's *Socratis et Socraticorum reliquiae* (Naples: Bibliopolis, 1990, pp. 414–435). Yet notice that in the account of the Master Argument in Epictetus, a line has been omitted from the text in Giannantoni's collection, so that only the first of the three propositions is quoted in full. The testimonia on Diodorus can also be found in K. Döring, *Die Megariker* (Amsterdam: Grüner, 1972).

WORKS ABOUT DIODORUS CRONUS

Gaskin, R. *The Sea Battle and the Master Argument*. Berlin/New York: de Gruyter, 1995. Contains an extensive bibliography.

Prior, A. N. "Diodoran Modalities." *Philosophical Quarterly* 5 (1955): 205–213.

Sedley, D. "Diodorus Cronus and Hellenistic Philosophy." *Proc. Cambridge Philological Society* N. S. 23 (1977): 74–120. A pioneering study.

Vuillemin, J. *Necessity or Contingency: The Master Argument*. Stanford: CSLI Publications, 1996. Originally published in French as *Nécessité ou contingence*, 1984.

Weidemann, H. "Diodors Meisterargument und der Aristotelische Möglichkeitsbegriff." *Archiv für Geschichte der Philosophie* 69 (1987): 18–53. Discusses extensively most of the literature up to the mid-1980s.

See also the four articles by H. Weidemann, R. Gaskin, M. J. White, and N. Denyer on the Master Argument in *Logical Analysis and History of Philosophy* 2 (1999) 189–252. The articles are as follows: H. Weidemann, "'Aus etwas Möglichem folgt nichts Unmögliches'. Zum Verständnis der zweiten Prämisse von Diodors Meisterargument," 189–202; R. Gaskin, "Tense Logic and the Master Argument," 203–224; M. J. White, "The Lessons of Prior's Master Argument," 225–238; and N. Denyer, "The Master Argument of Diodorus Cronus: A Near Miss," 239–252.

Theodor Ebert (2005)

DIOGENES LAERTIUS

(c. 200 CE)

Diogenes Laertius is the author of *Compendium of the Lives and Opinions of Philosophers*, the only general book on philosophers and their philosophy that has been transmitted from classical antiquity. Diogenes is known from this work only—nothing is known about his life—and his date can only be fixed by the dates of the latest personalities mentioned in his text (second century CE), and because he seems to have written prior to the rise of Neoplatonism (c. 250 CE). His work was dedicated to a woman interested in Platonism (bk. 3 § 47).

Diogenes's work belongs to a type of ancient literature (often called *Diadocha* or *Successions*) in which accounts of the lives of philosophers were arranged as series of biographies so that teacher and student followed one another within each major philosophical school.

Diogenes's text is divided into ten sections, or "books":

1: Introduction and various "wise men," including Thales.

(*2–7: The Ionian Tradition*)

2: The Ionian physicists, Socrates, and the minor Socratic schools down to the early third century BCE.

3: Plato.

4: The Academy down to Clitomachus (late second century BCE).

5: Aristotle and the Peripatetics down to Lyco (late third century BCE).

6: Antisthenes and the Cynics down to the end of the third century BCE.

7: Zeno and the Stoics down to at least Chrysippus (late third century BCE), and in the missing end of the book perhaps even down to the first century CE.

(*8–10: The Italic Tradition*)

8: Pythagoras and his early successors; Empedocles.

9: Heraclitus; the Eleatics; the Atomists, Protagoras and Diogenes of Apollonia; Pyrrho.

10: Epicurus.

Diogenes's book is basically a compilation of excerpts from numerous sources; in the biographical sections he often tells which sources he is using, whereas the philosophical sections contain few such references.

The book is also uneven. Some lives contain nothing but anecdotes and aphorisms, whereas others are mainly doxographical reports; some have long, detailed sections on philosophy, whereas others have short, superficial sections. Diogenes is unlikely to have read many philosophical works. However, in book 10 he has preserved four long, original writings by Epicurus, which constitute the most important evidence for Epicurus's philosophy from before the period of Cicero. However, his many references to his predecessors give an impression of the Hellenistic tradition of philosophical biography. Because Diogenes seems to have had a predilection for old documents, he has preserved the testaments of four peripatetics and a number of book catalogs.

Most of Diogenes's biographies include a number of items such as birth, parents, name, appearance, relationship to other philosophers, travels, lifestyle, and circumstances of death, yet they are presented in no particular order. The dominating element in the biographies is the use of anecdotes. In antiquity it was impossible to find documentary evidence concerning a deceased person, unless that person was a famous public figure or had left written works. Often literary works were exploited without regard to the fact that the content of a fictional work is unlikely to apply to the life of its author. Therefore, Diogenes's factual information must be viewed with some skepticism: Notice that most of his dates are taken from a Hellenistic poem.

Diogenes's biographies may have been written with less artistic skill than, for example, Plutarch's; however, they are not unlike other ancient accounts of the lives of philosophers.

Diogenes devotes much space to present the doctrines of the major philosophical schools: Book 3, § 48–109, is a general introduction to the study of the *Corpus Platonicum*; as an account of Plato's philosophy it may be inadequate, but it resembles other Platonic writings of the second century CE. The section on Aristotelian philosophy (bk. 5, § 28–34) is far less satisfying, but all three parts seem to go back to the Hellenistic period. Book 7, § 38–160, is the most comprehensive account of Stoic philosophy from antiquity, the section on logic is especially important. The survey of the Skeptic tropes (bk. 9, § 79–105) is shorter than in Sextus Empiricus but otherwise comparable. The three Epicurean letters and his forty "Principal Doctrines" in book 10 are crucial to what is known about Epicurus; when Diogenes places these aphorisms at the end of his book, he indicates that he considers them a culmination of philosophical wisdom.

For the pre-Socratic philosophers, Diogenes has used a "doxographical" source similar to other accounts in late antiquity; ultimately, it derives from Aristotle and Theophrastus. In the case of Pythagoras, Diogenes presents two excerpts from Aristotle and from Alexander Polyhistor (first century BCE), thus presenting a much earlier expression of Pythagoreanism than is found in other sources from late antiquity.

Diogenes was no philosopher, but he has preserved much of philosophical significance. He seems to have had no influence in antiquity, but since Walter Burley's *On the Life and Manners of the Philosophers* (early fourteenth century), the Latin translation by Ambrosius Traversarius (1432), and the *editio princeps* of the Greek text in 1533, Diogenes has been the most important single source for the lives and often for the doctrines of ancient philosophers. Until around 1800, Diogenes was the main model for historiography of philosophy.

See also History and Historiography of Philosophy.

Bibliography

TRANSLATIONS

Diogène Laërce Vies et doctrines des philosophes illustres, edited by Marie-Odile Goulet-Cazé. Paris: La Pochothèque, 1999. French translation.

Lives of Eminent Philosophers. Translated by R. D. Hicks. London: W. Heinemann; New York: Putnam, 1925. English translation.

Vitae Philosophorum. Vols. 1 and 2 edited by M. Marcovich; vol. 3, *Indices*, edited by H. Gärtner. Leipzig/Munich, Germany: Bibliotheca Teuneriana, 1999–2002. Greek translation. A new Budé edition by T. Dorandi, in preparation, will become the standard edition of the future.

STUDIES

Giannantoni, G., ed. *Diogene Laerzio. Storico del pensiero antico*. Naples: Bibliopolis, 1986.

Haase, W., and Temporini, H., eds. *Aufstieg und Niedergang der Römischen Welt*. Berlin and New York: W. de Gruyter, 1992. (Part) II volume 36 (parts) 5–6 contains various studies in several languages.

Mejer, Jørgen. *Die Ueberlieferung der Philosophie im Altertum*. Copenhagen: Det Kongelige Danske videnskabernes selskab, 2000. Contains an extensive bibliography.

Mejer, Jørgen. *Diogenes Laertius and His Hellenistic Background*. Wiesbaden, Germany: Steiner, 1978.

Jørgen Mejer (2005)

DIOGENES OF APOLLONIA
(5th century BCE)

Diogenes of Apollonia was a Greek philosopher belonging to the last generation of the pre-Socratics (fl. around 440–430 BCE.) His native town was either Apollonia on Crete or, more probably, Apollonia on the Pontus. Nothing is known for certain about his life. It has been debated whether he wrote only one book called, in English, *On Nature* or, as Simplicius reported (in *On Aristotle's "Physics"* 151, 20), four (*On Nature*, *Meteorology*, *On the Nature of Man*, *Against the Sophists*). All the existing fragments seem to come from *On Nature*. His work had an effect in Athenian intellectual life toward the end of the fifth century BCE, and his influence is detectable also in some treatises of the Hippocratic corpus and in the Stoic doctrine of *pneuma* (literally breath; in Stoic philosophy, the mixture of the two active elements, fire and air, and the sustaining cause of all bodies.)

His philosophy was termed "eclectic" already by Theophrastus, and most modern commentators agree with this assessment. Theophrastus listed Anaxagoras, Leucippus, and Anaximenes as the main influences on Diogenes, and to this list we should certainly add Heraclitus. Diogenes' philosophical doctrine has three prominent aspects: his monism, the teleological traits in his cosmology, and his theory of cognition. Most of the pre-Socratic philosophers working after Parmenides adopted a pluralist ontology. Diogenes, on the contrary, returned to the monism of his Ionian predecessors. He argued that if the proper nature of apparently different types of matter were not the same, then these different types of matter could not causally interact with one another, and we could not explain such phenomena as the nutrition and growth of living organisms, in which apparently different types of matter transform into each other. Therefore, the four elements and the other types of matter of our world must have differentiated from the same primordial stuff, must retain their underlying identity, and must ultimately return to what they differentiated from (Diels and Kranz [DK], B2). Apparent things exist for a limited time, whereas the basic stuff is "an eternal and deathless body" (DK, B7). Yet it is not a passive substrate, but is "strong" and determines how things are formed from it and return to it (DK, B8, B7). Because it is active and eternal, it can also be considered a god.

Diogenes continued by arguing that the basic stuff must be intelligent. He wrote, "For without intelligence it could not have been divided up in such a way as to hold the measures of all things, of winter and summer and night and day and rains and winds and nice weather, along with the rest, which, if one is willing to consider them intelligently, one will find disposed in the finest possible way" (DK, B3). Scholars have disagreed how thorough Diogenes' teleology, as expressed in this frag-

ment, is. According to Willy Theiler, Diogenes is a full-blown teleologist and the immediate source of the teleological views that Xenophon ascribed to Socrates in his *Memorabilia*. Others have doubted that Diogenes' conception is original and that it is genuinely teleological. Diogenes' argument certainly differs from later, explicitly teleological views in that it remains unclear whether the action of the intelligent principle is directed at some well-defined goal or goals. It also differs from classic statements of the argument from design, with which it has sometimes been associated, in that Diogenes did not argue for the existence of an intelligent causal principle, but sought to show that the ultimate causal principle, the existence of which he established on independent grounds, must also be intelligent.

Diogenes identified the bearer of intelligence with air. He argued that because humans and animals live by breathing, air must be what brings life and intelligence to them (DK, B4). If so, the air, which inheres in, and steers, all things, must be the intelligent causal principle at the cosmic level too. Moreover, the qualitative differences of air explain the differences between species and individuals (DK, B5).

Diogenes' most original contribution was a detailed description of the system of veins, which originate in the head and through which blood and air to all parts of the body. Sensation is produced when air from the outside acts on the air in the sense organs which then reaches the head through the veins. The quality of the air and the veins determine the sharpness of perception. Air mixed with blood produces thought, and we feel pleasure when the appropriate mixture of air and blood pervades the whole body.

See also Pneuma; Stoicism.

Bibliography

Barnes, Jonathan. *The Presocratic Philosophers*. 2nd ed. London: Routledge and Kegan Paul, 1982.

Diels, Hermann, and Walther Kranz. *Die Fragmente der Vorsokratiker*. 7th, rev. ed. Berlin: Weidmann, 1954.

Kirk, G. S., J. E. Raven, and Malcolm Schofield. *The Presocratic Philosophers*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983.

Laks, André. *Diogène d'Apollonie: La dernière cosmologie présocratique*. Lille, France: Presses universitaire de Lille, 1983.

Theiler, Willy. *Zur Geschichte der teleologischen Naturbetrachtung bis auf Aristoteles*. 2nd ed. Berlin: De Gruyter, 1965.

Gábor Betegh (2005)

DIOGENES OF SINOPE

(4th century BCE)

Diogenes of Sinope, who lived in the fourth century BCE, was the prototype of the Cynics, who probably were so called from Diogenes' Greek nickname, the Dog (*kuon*; adjective form, *kunikos*). Tradition held that on coming to Athens in exile, he was influenced by Antisthenes' teaching; Diogenes' ascetic distortion of Socratic temperance gives some point to Plato's supposed remark that he was a "Socrates gone mad."

It is not easy to recover the philosopher from, on the one hand, the lurid fog of anecdotal tradition that represents the stunts of an eccentric tramp at Athens and Corinth defacing conventional human standards—as he or his father, Hicesias, was supposed to have defaced in some way the currency of Sinope—or, on the other, the idealized legend that grew after his death. But doxographic traces (for example, Diogenes Laërtius, VI.70–73) and, indeed, the tradition as a whole presuppose a serious teacher who, in disillusioned protest against a corrupt society and hostile world, advocated happiness as self-realization and self-mastery in an inner spiritual freedom from all wants except the bare natural minimum; and who, in a bitter crusade against the corrupting influence of pleasure, desire, and luxury, extolled the drastic painful effort involved in the mental and physical training for the achievement of a natural and inviolable self-sufficiency.

The anecdotes illustrate Diogenes' philosophy in action. Since for Diogenes virtue was revealed in practice and not in theoretical analysis or argument, the stories of, for example, his embracing statues in winter and his peering with a lantern in daytime for a human being, the tales of his fearless biting repartee and criticism of notables such as Alexander, however embroidered or apocryphal, correctly reflect his pointed teaching methods, which encouraged the development of a new didactic form, the *chreia,* or moral epigram. Some exaggeration here is due to the "dog-cynic" shamelessness pedagogically employed to discount convention, and some is no doubt inherent in the uncompromising extremes of Diogenes' doctrines.

He is credited with tragedies illustrating the human predicament and with a *Republic,* which influenced Zeno the Stoic, that was notorious for its scandalous attack on convention. His famous remark that he was a citizen of the world is more probably antinational than international, for he was concerned with the individual rather than the community. Diogenes sought to make any man king, not of others, but of himself, through autonomy of

will, and his own life was his main philosophical demonstration to this end.

See also Antisthenes; Cynics; Diogenes Laertius; Hellenistic Thought; Plato; Zeno of Citium.

Bibliography

WORKS ON DIOGENES OF SINOPE

Crönert, W. *Kolotes und Menedemos.* Leipzig, 1906.

Diogenes Laërtius. *Lives,* VI. 20–81.

Dudley, D. R. *History of Cynicism.* London, 1937.

Höistad, R. *Cynic Hero and Cynic King.* Uppsala, Sweden, 1949.

Sayre, F. *The Greek Cynics.* Baltimore, 1948.

I. G. Kidd (1967)

DIOGENES THE CYNIC

See *Diogenes of Sinope*

DIONYSIUS THE PSEUDO-AREOPAGITE

See *Pseudo-Dionysius*

DIRECT REALISM

See *Realism*

DISCOURSE ETHICS

"Discourse ethics" refers to an approach to moral theory developed by Jürgen Habermas. It is a reconstruction of Immanuel Kant's idea of practical reason that turns on a reformulation of his categorical imperative: Rather than prescribing to others as valid norms that I can will to be universal laws, I must submit norms to others for purposes of discursively testing their putative universality. "Only those norms may claim to be valid that could meet with the approval of all those affected in their capacity as participants in practical discourse" (Habermas, 1990, p. 66). Normative validity, construed as rational acceptability, is thus tied to argumentation processes governed by a principle of universalization: "For a norm to be valid, the consequences and side effects of its general observance for the satisfaction of each person's particular interests must be acceptable to all" (p. 197). Furthermore, by requiring that perspective taking be general and reciprocal, discourse ethics builds a moment of empathy or "ideal role-taking" into the procedure of practical argumentation.

Like Kant, Habermas distinguishes the types of practical reasoning and the corresponding types of "ought" connected with questions concerning what is pragmatically expedient, ethically prudent, or morally right. Calculations of rational choice furnish recommendations relevant to the pursuit of contingent purposes in the light of given preference. When serious questions of value arise, deliberation on who one is and wants to be yields insight into the good life. If issues of justice are involved, fair and impartial consideration of conflicting interests is required to judge what is right or just. Again like Kant, Habermas regards questions of the last type, rather than specifically ethical questions, to be the proper domain of theory. (Thus, discourse ethics might properly be called discourse morality.) This is not to deny that ethical discourse is rational or that it exhibits general structures of its own; but the irreducible pluralism of modern life means that questions of self-understanding, self-realization, and the good life do not admit of universal answers. In Habermas's view, that does not preclude a general theory of a narrower sort, namely a theory of justice. Accordingly, the aim of his discourse ethics is solely to reconstruct the moral point of view from which questions of right can be fairly and impartially adjudicated.

By linking discourse ethics to the theory of communicative action, Habermas means to show that our basic moral intuitions are rooted in something deeper and more universal than particularities of our tradition, namely in the intuitive grasp of the normative presuppositions of social interaction possessed by competent social actors in any society. Members of our species become individuals in and through being socialized into networks of reciprocal social relations. The mutual vulnerability that this interdependence brings with it calls for guarantees of mutual consideration to preserve both the integrity of individual persons and the web of their interpersonal relations. In discourse ethics respect for the individual is built into the freedom of each participant in discourse to accept or reject the reasons offered as justifications for norms, and concern for the common good is built into the requirement that each participant take into account the needs, interests, and feelings of all others affected by the norm in question. Hence, the actual practice of moral discourse depends on forms of socialization and social reproduction that foster the requisite capacities and motivation.

See also Habermas, Jürgen; Justice; Kant, Immanuel; Practical Reason.

Bibliography

Habermas, J. *Justification and Application: Remarks on Discourse Ethics.* Translated by C. Cronin. Cambridge, MA: MIT Press, 1993.

Habermas, J. *Moral Consciousness and Communicative Action.* Translated by C. Lenhardt and S. Nicholsen. Cambridge, MA: MIT Press, 1990.

Rehg, W. *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas.* Berkeley: University of California Press, 1994.

Thomas McCarthy (1996)

DISPOSITIONAL THEORIES

See *Response-Dependence Theories*

DISTANT PEOPLES AND FUTURE GENERATIONS

Only recently have philosophers begun to discuss the question of whether we can meaningfully speak of distant peoples and future generations as having rights against us or of our having corresponding obligations to them. Answering this question with respect to distant peoples is much easier than answering it with respect to future generations. Few philosophers have thought that the mere fact that people are at a distance from us precludes our having any obligations to them or their having any rights against us. Some philosophers, however, have argued that our ignorance of the specific membership of the class of distant peoples does rule out these moral relationships. Yet this cannot be right, given that in other contexts we recognize obligations to indeterminate classes of people, such as a police officer's obligation to help people in distress or the obligation of food producers not to harm those who consume their products.

Of course, before distant peoples can be said to have rights against us, we must be capable of acting across the distance that separates us. Yet as long as this condition is met—as it typically is for people living in most technologically advanced societies—it would certainly seem possible for distant peoples to have rights against us and us corresponding obligations to them.

By contrast, answering the above question with respect to future generations raises more difficult issues. One concerns whether it is logically coherent to speak of future generations as having rights now. Of course, no one who finds talk about rights to be generally meaningful should question whether we can coherently claim that future generations *will* have rights at some point in the future (specifically, when they come into existence and are no longer future generations). But what is questioned, since it is of considerable practical significance, is whether we can coherently claim that future generations have rights now when they do not yet exist.

Let us suppose, for example, that we continue to use up Earth's resources at present or even greater rates, and as a result, it turns out that future generations will face widespread famine, depleted resources, insufficient new technology to handle the crisis, and a drastic decline in the quality of life for nearly everyone. If this were to happen, could persons living in the twenty-second century legitimately claim that we in the twenty-first century violated their rights by not restraining our consumption of the world's resources? Surely it would be odd to say that we violated their rights more than one hundred years before they existed. But what exactly is the oddness?

Is it that future generations generally have no way of claiming their rights against existing generations? While this does make the recognition and enforcement of rights much more difficult (future generations would need strong advocates in the existing generations), it does not make it impossible for such rights to exist. After all, the recognition and enforcement of the rights of distant peoples is also a difficult task, but obviously such rights can exist.

Perhaps what troubles us is that future generations do not exist when their rights are said to demand action. But how else could persons have a right to benefit from the effects our actions will have in the distant future if they did not exist just when those effects would be felt? Our contemporaries cannot legitimately make the same demand, for they will not be around to experience those effects. Only future generations could have a right that the effects our actions will have in the distant future contribute to their well-being. Nor need we assume that, for persons to have rights, they must exist when their rights demand action. Thus, to say that future generations have rights against existing generations, we can simply mean that there are enforceable requirements upon existing generations that would benefit future generations or prevent harm to them.

Most likely what really bothers us is that we cannot know for sure what effects our actions will have on future generations. For example, we may, at some cost to ourselves, conserve resources that will be valueless to future generations who may develop different technologies. Or, because we regard them as useless, we may destroy or deplete resources that future generations will find to be essential to their well-being. Nevertheless, we should not

allow such possibilities to blind us to the necessity of a social policy in this regard. After all, whatever we do will have its effect on future generations. The best approach, therefore, is to use the knowledge we have and assume that future generations will also require those basic resources we now find to be valuable. If it turns out that future generations require different resources to meet their basic needs, at least we will not be to blame for acting on the basis of the knowledge we have.

Assuming then that we can meaningfully speak of distant peoples and future generations as having rights against us and us corresponding obligations to them, the crucial question that remains is exactly what rights they have against us and what obligations we have to them. While the answer to this question obviously depends on a substantial social and political theory, the expectation is that the rights and obligations that morally bind us to distant peoples and future generations will be quite similar to those that morally bind us to near people and existing generations.

See also Bioethics; Environmental Aesthetics; Environmental Ethics; Genetics and Reproductive Technologies; Philosophy of Technology; Responsibility, Moral and Legal; Rights.

Bibliography

Elfstrom, G. *Ethics for a Shrinking World.* New York: St. Martin's Press, 1990.

Hardin, G. *Promethean Ethics.* Seattle: University of Washington Press, 1980.

Partridge, E. *Responsibilities to Future Generations.* Buffalo, NY: Prometheus Books, 1981.

James P. Sterba (1996)

DIVINE COMMAND THEORIES OF ETHICS

The general perspective on ethics known as theological voluntarism usually appears in philosophical discussions in the specific form of divine command theories. As its title suggests, theological voluntarism is the view that ethics depends, at least in part, on God's will. In divine command theories the dependency is spelled out in terms of commands by God that express the divine will. The Hebrew Bible portrays God as establishing norms for human conduct by giving commands. Though some of them pertain exclusively to the regulation of religious rituals, others such as the prohibitions of murder and theft clearly have ethical content. Since the Hebrew Bible counts as authoritative scripture for all three of the major monotheistic religions, divine command theories are a live option within Jewish, Christian, and Islamic traditions.

As the historical research of Janine M. Idziak (1979) shows, many Christian thinkers have exercised this option. St. Augustine, St. Bernard of Clairvaux, St. Thomas Aquinas, and St. Andrew of Neufchateau claimed that divine commands determine the ethical status of particular actions when they dealt with issues in biblical exegesis. John Duns Scotus and William of Ockham endorsed divine command theories. Both Martin Luther and John Calvin advocated an ethics of divine commands. John Locke and William Paley are among the modern philosophers who argued for divine command theories. Søren Kierkegaard's *Works of Love* (1847/1995) contains a divine command theory. In short, over a period of many centuries divine command ethics has attracted support from major figures in both Catholic and Protestant branches of Christianity.

A strong cumulative case for the importance of God's will in ethics can be constructed from within a Christian worldview. As Kierkegaard emphasized, a central element in such a case comes from the Christian New Testament. It is a striking feature of its distinctive ethics of love (*agape*) that love is commanded. In Matthew's Gospel the command is stated in response to a lawyer's query. Jesus says, "You shall love the Lord your God with your whole heart, with your whole soul, and with all your mind. This is the greatest and the first commandment. The second is like it: You shall love your neighbor as yourself" (Matthew 22:37–39). Similar commands are endorsed or stated by Jesus in the other three Gospels. If Jesus is God the son, as traditional Christians believe, such commands derive from and express the will of God. Thus, the ethics of agapeistic love advocated in the New Testament can plausibly be interpreted as having its source in a divine command.

During the final third of the twentieth century a revival of interest in divine command ethics took place among philosophers of religion. Most of the philosophers who wrote on the subject in this period understood divine command theories to be accounts of the realm of moral deontology. This domain of ethics studies topics related to duty; its main concepts are requirement (obligation), permission (rightness), and prohibition (wrongness). Edward R. Wierenga (1989) proposes a causal divine command theory according to which by commanding actions God brings it about that they are oblig-

atory and by forbidding actions God brings it about that they are wrong. Robert Merrihew Adams (1999) advocates a theory in which an action's being obligatory consists in its being commanded by God and an action's being wrong consists in its being contrary to a divine command. Stated in general terms, the principle of obligation of a divine command theory of the type favored by these philosophers asserts that actions are obligatory if and only if, and just because, they are commanded by God. And the principle of wrongness of such a theory claims that actions are wrong if and only if, and just because, they are prohibited by God.

Adams argues that divine commands do not account for ethical goodness and related axiological characteristics. In his theistic Platonism God plays the role of the Form of the Good; God is the paradigm or standard of goodness. Other things are good in virtue of bearing a relation of resemblance to God. For Adams (1999), ethical goodness thus depends on God, but not on God's will or commands.

Philosophers who contribute to the revival of divine command ethics devote a good deal of time and energy to defending divine command theories against criticism. Perhaps the most famous objection has roots that trace back to a question Socrates raises in the *Euthyphro*. Altering it a bit to allow for the difference between Greek polytheism and monotheism, one may imagine a Socratic gadfly asking: Does God command truth-telling because it is obligatory, or is truth-telling obligatory because God commands it? No matter which way questions of this sort are answered, a difficulty for divine command ethics emerges.

If one supposes that God commands truth-telling because it is obligatory, one contradicts the claim of divine command theorists that truth-telling is obligatory because it is commanded by God. In other words, this response forces one to conclude that the obligatoriness of truth-telling is independent of God's commands. But if one insists that truth-telling is obligatory because God commands it, which is what divine command theorists are committed to doing, then one must confront a difficulty that was eloquently formulated by Ralph Cudworth in *A Treatise concerning Eternal and Immutable Morality* (1731/1976). As he notes, divine command theorists are committed to the view that lying rather than truth-telling would be obligatory if it were commanded by God.

However, divine command theorists can accept Cudworth's (1731/1976) point with equanimity if they embed their divine command account of moral deontology in an axiological theory that, like the theistic Platonism espoused by Adams, makes ethical goodness independent of God's will and commands. Understood in this way, goodness is determined by God's immutable nature and character; it is a matter of who and what God is. God's essential nature, which is paradigmatic of goodness, will then constrain what God can command. Hence, it is open to divine command theorists to hold that it is impossible for God to command lying and so is impossible for lying to be obligatory. This view is consistent with granting that lying would be obligatory if, *per impossible*, God were to command it.

Certain forms of divine command ethics can be shown to stand up well under philosophical scrutiny. Divine command accounts of obligation and wrongness deserve to be regarded as respectable options in ethical theory if the larger theistic worldviews of which they are components are themselves philosophically defensible.

See also Moral Principles: Their Justification; Religion and Morality.

Bibliography

Adams, Robert Merrihew. *Finite and Infinite Goods: A Framework for Ethics*. New York: Oxford University Press, 1999. Argues for a divine command theory within an ethical framework of theistic Platonism.

Cudworth, Ralph. *A Treatise concerning Eternal and Immutable Morality* (1731). New York: Garland, 1976. Classic source of objections to divine command ethics.

Idziak, Janine M., ed. *Divine Command Morality: Historical and Contemporary Readings*. New York: Edwin Mellen Press, 1979. Collects informative selections from writings by both defenders and critics of divine command ethics.

Kierkegaard, Søren. *Works of Love* (1847). Translated and edited by Howard V. Hong and Edna H. Hong. Princeton, NJ: Princeton University Press, 1995. Contains a divine command account of Christian agapeistic ethics.

Mouw, Richard J. *The God Who Commands*. Notre Dame, IN: University of Notre Dame Press, 1990. Defends divine command ethics within a theological perspective.

Quinn, Philip L. *Divine Commands and Moral Requirements*. Oxford, U.K.: Clarendon Press, 1978. Defends divine command ethics within a philosophical perspective and sets forth rigorous formulations of several versions of divine command theory.

Quinn, Philip L. "Divine Command Theory." In *The Blackwell Guide to Ethical Theory*, edited by Hugh LaFollette. Oxford, U.K.: Blackwell, 2001. Outlines a cumulative case, most of which is internal to Christianity, that lends support to divine command ethics.

Wierenga, Edward R. *The Nature of God: An Inquiry into Divine Attributes*. Ithaca, NY: Cornell University Press, 1989. Presents and defends a causal divine command theory.

Philip L. Quinn (2005)

DŌGEN
(1200–1253)

MAJOR WORKS OF DŌGEN

Dōgen was the founder of Sōtō Zen Buddhism and helped introduce to medieval Japan many features of Chan Buddhist theory and practice that developed during the Song dynasty in China. His major works include the *Shōbōgenzō* (Treasury of the true dharma-eye), a collection of sermons composed in vernacular Japanese from 1231 until the end of his life; the *Shōbōgenzō Zuimonki* (Miscellaneous talks), another collection of vernacular sermons compiled from 1234 until 1238; the *Eihei Kōroku* (Recorded sayings at Eiheiji Temple), a collection of sermons in Chinese compiled from 1236 to 1252; the *Fukanzazengi* (Universal recommendation for Zazen practice), a concise summary of his views on meditation composed in 1233; and the *Eihei Shingi* (Monastic rules at Eiheiji Temple), a collection of six essays dealing with monastic rules and regulations composed from 1237 to 1249.

Dōgen is often referred to as the leading philosopher in Japanese history. His writings on many Buddhist topics reflect an approach to religious experience based on a more philosophically oriented level of analysis than is found in the writings of most thinkers in Zen, which is known as a "special transmission outside the scriptures, without reliance on words and letters." Dōgen has been a major influence on modern Japanese philosophy, especially representatives of the Kyoto School such as Nishida Kitarō, Nishitani Keiji, and Abe Masao, and has been compared with a wide range of classical and modern Western philosophers and religious thinkers ranging from Aristotle and St. Thomas Aquinas to Martin Heidegger, Friedrich Nietzsche, and Jacques Derrida.

DŌGEN'S LIFE AND TEACHINGS

Some of Dōgen's major philosophical ideas emphasize that philosophy of religion must reflect personal experience of transient existence based on an awareness that the ultimate reality of the universal Buddha-nature is not beyond but is conditioned by impermanence. Impermanent reality is characterized by a fundamental unity of being-time (*uji*) in that all beings occur as temporal manifestations and time is manifested through each aspect of existence. Dōgen maintains that religious practice, or training, and spiritual realization, or the attainment of enlightenment, occur simultaneously and are inseparable in the experience of liberation known as "the casting off of body-mind" (*shinjin datsuraku*) that is achieved through the methods of zazen meditation and kōan interpretation, which are equally conducive to realization. He also stresses that the naturalist dimension of being-time and impermanence-Buddha-nature is expressible through poetry and aesthetics, but reminds that karmic causality or moral conditioning and retribution are inherent to, rather than outside of, the attainment of enlightenment.

Much of Dōgen's emphasis on impermanence is based on his own experiences. According to the traditional accounts Dōgen was born into an aristocratic family at a time when Japan was beginning to be plagued by repeated civil warfare. He experienced profound sorrow at an early age as his father and mother died by the time he was seven. It is said that when Dōgen saw the smoke rising from incense and vanishing during his mother's funeral, he was deeply moved by an awareness of the inevitability of death and the pervasiveness of ephemeral reality.

The orphaned Dōgen decided to renounce secular life in pursuit of the Buddhist dharma. At first, he studied on Mount Hiei outside the capital city of Kyoto in the dominant Japanese Tendai church, in which the central doctrine was an affirmation of "original enlightenment" (*hongaku*) or the inherent potentiality of all beings to attain the primordial Buddha-nature. However, at the age of thirteen Dōgen had a fundamental "doubt" about the doctrine of original enlightenment: If everyone is already enlightened in that they possess the Buddha-nature, he thought, then what is the need for sustained meditative practice as required by the Buddha's teaching?

Unable to resolve this doubt in Japan, Dōgen traveled to China, where the contemplative path of Zen had become the dominant movement. At first, Dōgen was disappointed in the laxity of the Chinese Chan monks, who failed to inspire him. Then, on the verge of returning to Japan unfulfilled, he met the teacher Rujing, who insisted on an unrelenting approach to meditation. Under the guidance of his new mentor Dōgen attained an awakening experience of the casting off of body-mind, or a continuing process of liberation from all intellectual and volitional attachments, which signified the resolution of his doubt about the necessity of continuously renewed training.

Before his breakthrough experience Dōgen apparently presumed the conventional dichotomies between past, present, and future, now and then, life and death, impermanence and nirvana, time and eternity, and finitude and Buddha-nature. He thought that human beings

were bound to a realm of death and impermanence and that enlightenment was beyond this realm. However, in casting off body-mind he realized that a single moment encompasses the unity of practice and attainment, so that practice is not before—nor does it lead up to—enlightenment and enlightenment is not a teleological goal reached only at the end of practice. Rather, as Dōgen writes in the *Shōbōgenzō*, "[p]ractice and realization are identical. Because one's present practice is practice in realization, one's initial negotiation of the Way in itself is the whole of original realization.… As it is already realization in practice, realization is endless; as it is practice in realization, practice in beginningless" (Dōgen, *Dōgen Zenji Zenshū*, vol. 2, pp. 546–547).

On returning to Japan, in 1233 Dōgen established the Sōtō sect at Kōshōji temple in the Kyoto area, but because of sectarian disputes with Tendai and other Zen factions he eventually moved in 1243 to the remote, pristine mountains of Echizen (now Fukui) Province, where Eiheiji temple was constructed. According to Dōgen's writings of the late period, every action generates a retributive consequence, and only authentic repentance and acknowledgment of one's guilt can offset the effects of evil karma. Still, by emphasizing the moment-to-moment cause-and-effect process of karmic retribution—which is inseparable from nirvana as part of the Bodhisattva's commitment to compassion—Dōgen is consistent with his earlier philosophy of being-time.

A central feature of aesthetic realization is Dōgen's use of poetic language, especially elaborate metaphors and philosophical wordplay, to convey emotional fulfillment that enhances rather than opposes the enlightenment experience of detachment from worldly, materialistic concerns. One of Dōgen's most eloquent poems was written near the end of his life as he returned from Echizen to the capital city for medical care. Making the journey to see Kyoto for the first time in ten years, but for what would prove to be the last time, Dōgen wrote in the five-line, thirty-one-syllable *waka* form:

> Like a blade of grass,
> My frail body
> Treading the path to Kyoto,
> Seeming to wander
> Amid the cloudy mist on the mountain path.
>
> *(Heine 1989, p. 85)*

See also Buddhism—Schools: Chan and Zen.

Bibliography

WORKS BY DŌGEN

Dōgen. *Dōgen Zenji Zenshū*, Edited by Kōdō, Kawamura, et al., 7 vols. Tokyo: Shunjūsha, 1988–1993.

Dōgen. *Dōgen's Pure Standards for the Zen Community*. Translated by Taigen Dan Leighton and Shohaku Okumura. Albany: SUNY Press, 1996.

Dōgen. *Dōgen's Extensive Record: A Translation of the Eihei Kōroku*. Translated by Taigen Dan Leighton and Shohaku Okumura. Boston: Wisdom, 2004.

Dōgen. *Dōgen Zenji's Shōbōgenzō* (The Eye and Treasury of the True Law). Translated by Nishiyama Kōsen and John Stevens. 4 vols. Sendai, Japan: Daihokkaikakum, 1975.

WORKS ON DŌGEN

Abe, Masao. *A Study of Dōgen: His Philosophy and Religion*. Edited by Steven Heine. Albany: SUNY Press, 1992.

Bielefeldt, Carl. *Dōgen's Manuals of Zen Meditation*. Berkeley: University of California Press, 1988.

Faure, Bernard. "The Daruma-shū, Dōgen, and Sōtō Zen." *Monumenta Nipponica* 42 (1) (1987): 25–55.

Heine, Steven. *A Blade of Grass: Japanese Poetry and Aesthetics in Dōgen Zen*. New York: Peter Lang, 1989.

Heine, Steven. *Dōgen and the Kōan Tradition: A Tale of Two Shōbōgenzō Texts*. Albany: SUNY Press, 1994.

Kim, Hee-Jin. *Dōgen Kigen—Mystical Realist*. Tucson: University of Arizona Press, 1975.

LaFleur, William R., ed. *Dōgen Studies*. Honolulu: University of Hawaii Press, 1985.

Tanahashi, Kazuaki. *Moon in a Dewdrop: Writings of Zen Master Dōgen*. Translated by Robert Aitken, et al. San Francisco: North Point Press, 1985.

Steven Heine (2005)

DOGMA

The Greek word of which "dogma" is a transliteration means "that which seems good." It was applied by Greek authors to the decrees of public authorities and to the tenets of various philosophical schools. In English the word can be used for any fixed and firmly held belief on any subject, but it usually suggests that the belief is a condition, or at least a sign, of belonging to either a secular or (more frequently) a religious group. The word can also imply that the belief rests on a special—often divine—authority; that any member of the group who attenuates or changes the belief is thereby a "heretic"; and that heresy is a moral, and perhaps also legal, offense that merits the strongest condemnation (and perhaps also punishment).

The clearest example of religious dogma in ancient philosophy comes from Plato. In the *Republic* (376Eff.) he lays down two "ways in which God is to be spoken of" (*tupoi theologias*). The first is that God is good and the

cause of good alone; the second is that God is true and incapable of change. In the *Laws* (887E–888D) he actually uses "dogma" to mean a correct "belief" about the gods. Everyone must believe that the gods are concerned with human affairs and that they cannot be appeased by sacrifice. Those who reject these beliefs must be duly punished by the state.

The primary sense of "dogma" is the one it has acquired in Christianity. Other religions have their distinctive tenets, but Christianity alone deserves attention on three grounds. First, its dogmas are far more numerous and complex than those of other faiths: Judaism requires only the recitation of the *Shema*, and Islam requires only assent to the *Kalima*. (Both these short creeds affirm the unity of God.) Second, Christian dogma has had many important points of contact with Western secular philosophy. Third, Christian theologians have given to the word *dogma* itself a technical, precise significance. (There is nothing that can properly be called dogma in the religions of the East. The eightfold path of Buddhism is a nontheistic way of salvation, not a creed. In Hinduism there are many divergent views of God and the Absolute, but none of them is "orthodox.")

All the main Christian bodies are agreed that dogma is essentially the formulation of belief on the basis of the Scriptures. God revealed himself both in the events to which the Bible testifies and in the biblical interpretation of them. The role of dogma is to express the meaning of this revelation in conceptual terms.

All would also agree that dogma does not add to the revelation that was complete with the apostles. Dogma merely makes explicit what is implicit in apostolic teaching. Hence, St. Vincent of Lérins affirms that the development of dogma is an "advance" (*profectus*), not a change (*permutatio*). Although a dogma can always be restated in a form that is either more exact per se or more comprehensible to a particular audience, its substance is immutable.

This point is clearly made by Hans Küng in his important book on the second Vatican Council, *The Council and Reunion* (London, 1961). On the one hand, "dogmatic definitions express the truth with infallible accuracy and are in this sense unalterable (as against Modernism)" (p. 163). On the other hand, "one and the same truth of faith can always be expressed in a still more complete, more adequate, better formula" (p. 163).

All Christian bodies, finally, would agree that the *ultimate* object of assent is not any statement about God, but God himself. Furthermore, dogmas do not render God intelligible; they symbolize a mystery that surpasses understanding. Therefore, we cannot assent to them without the gift of faith.

However, Christians differ in their views on both the number of and the authority for dogmatic definitions. Roman Catholic theologians hold that the definitions given by twenty ecumenical councils of the church are inerrant. They further hold that the pope alone, when he speaks *ex cathedra*, is infallible in matters of faith and morals. Finally, they hold that a dogma (for example, the dogma of the Immaculate Conception) can be justified as a logical "development" even though it lacks any scriptural support.

Non-Roman Christians oppose these claims. The Orthodox church holds that only seven councils are ecumenical and inerrant. Both Martin Luther and the Anglican reformers said that all councils are capable of error. All Protestants and Anglicans agree in denying both the infallibility of the pope and the validity of dogmas that are not explicitly supported by the Bible.

From the beginning, dogma has been stated through the terms of secular philosophy. One need mention only the use made of "substance" and "relation" in the doctrine of the Trinity. Such philosophical expressions were required both to make the faith intelligible and to safeguard it against heresy. Even those Protestants who reject scholastic terminology are forced to substitute other concepts (for example, those of existentialism).

In the theology of Thomas Aquinas, and in conciliar definitions, philosophy is instrumental. The content and authority of dogma are derived wholly from revelation, although some theologians have attempted to place dogmas in the context of a speculative system that is alien to the basic principles of Christian theism. Inevitably, the dogmas then lose their original, distinctive, and (above all) supernatural significance. Thus G. W. F. Hegel and his disciples held that Christ merely exhibits in a supreme mode the natural coinherence of the finite and the infinite.

At the other extreme, some post-Kantian thinkers, while remaining in the church, have denied that dogmas state objective truths concerning God. But we are to act "as if" they were true, and in so acting we shall find that the moral life is given both a meaning and a power that it cannot otherwise possess. This reduction of dogmas to the status of pragmatic postulates is the twenty-sixth proposition condemned by the decree *Lamentabili* (1907).

Bibliography

Bettenson, Henry. *Documents of the Christian Church.* New York: Oxford University Press, 1947.

Brunner, Emil. *The Christian Doctrine of God.* Translated by Olive Wyon. London: Lutterworth, 1949. Chs. 1–11.

Journet, C. *What Is Dogma?* Translated by Mark Pontifex. London: Royal Institute of International Affairs, 1946.

Newman, J. H. *An Essay in Development.* Several editions since 1845.

H. P. Owen (1967)

DONG ZHONGSHU

(c. 179–c. 104 BCE)

Dong Zhongshu, probably the most influential Confucian scholar of the Han dynasty (206 BCE–220 CE), laid an institutional basis for the Confucian orthodoxy and for the recruitment of able scholars as government officials through the examination system. He was an expert in the Gongyang commentary of the Confucian classic *Spring and Autumn,* and he gave the classic a new interpretation that combines the ethical and political teachings of Confucius with the supernatural view of the metaphysicians.

After having received the degree of eruditus (*boshi*) in the Confucian classics, Dong Zhongshu became a public instructor during the reign (156–140 BCE) of Emperor Jing. It has been recorded that he lectured from behind a curtain, and although he had many students, few were admitted to his presence. He was also said to have been so engrossed in his scholarly pursuits that for three years he did not even once visit his garden. As a result of his responses to the written inquiries addressed to the scholars of the realm by Emperor Wu (reigned 140–87 BCE), Dong Zhongshu attracted imperial notice and was appointed minister successively to two royal princes. However, he was not successful in his political career and spent the remaining years of his life in teaching and writing. In addition to his several memorials to the throne, he is known for his work on the *Spring and Autumn,* titled *Chunqiu Fanlu* (Copious Dew in Spring and Autumn), a curious admixture of moral and metaphysical essays in seventeen chapters. He had numerous followers and his influence lasted well beyond his lifetime.

Dong Zhongshu's main contribution as a Confucian philosopher lies in his study of the *Spring and Autumn,* which, according to him, teaches "compliance with Heaven's will and imitation of the ancients." To do so is "for the people to follow the sovereign, and for the sovereign to follow Heaven." Thus, the basic principle in government is to subject the people to the sovereign's domination, and the sovereign to Heaven's will. In Dong's concept, Heaven (*Tian*) is not the all-mighty anthropomorphic god of the ancient Chinese but the physical universe itself. Somewhat akin to the Western concept of nature, it is nevertheless endowed with intellect and purpose. The ruler, as Heaven's representative on earth, should administer his kingdom in accordance with Heaven's will. As Heaven is inherently good and benevolent, so should the sovereign be. His virtuous rule will be marked by order and harmony in the universe. On the other hand, any evil act of his will cause catastrophes (such as floods and fires, earthquakes and mountain slides) and anomalies (such as comets, eclipses, and the growing of beards on women) sent by Heaven as a warning to men. "The origin of catastrophes and anomalies," he wrote in "Copious Dew," "is traceable to misrule in the state. First, Heaven sends catastrophes to admonish the people. When this goes unheeded and no changes are made, Heaven would then frighten the people with prodigies. If men are still unawed, ruin and destruction will finally befall the empire."

Although he was an avowed monarchist, Dong Zhongshu's strange science of the catastrophes and anomalies had the effect of curbing misgovernment on the part of the ruler. The idea has so embedded itself in the minds of the Chinese people that even in more enlightened and rational times, Confucian scholar-officials found Dong's concept useful as a means of remonstrance against the ruler's misuse of despotic power. But Dong Zhongshu is remembered today chiefly for his historical role in exalting Confucianism as China's official state doctrine, which was to mold the nation for more than two thousand years from the Han dynasty to the present age.

See also Chinese Philosophy; Confucius; Ethics and Morality.

Bibliography

Chan, Wing-tsit. *A Source Book in Chinese Philosophy.* Princeton, NJ: Princeton University Press, 1963.

Fung Yu-lan. *A History of Chinese Philosophy.* Translated by Derk Bodde, Vol. 2. Princeton, NJ: Princeton University Press, 1953.

Gassmann, Robert H. *Tung Chung-shu Ch'un-Ch'iu Fan Lu: Üppiger Tau des Frühling-und-Herbst-Klassikers.* Frankfurt: Verlag Peter Lang, 1988.

Tain, Tzey-yueh. "Tung Chung-shu's System of Thought: Its Sources and Its Influence on Han Scholars." Ph.D. diss., University of California, 1974.

Liu Wu-chi (1967)

Bibliography updated by Loy Huichieh (2005)

DOSTOEVSKY, FYODOR MIKHAILOVICH

(1821–1881)

Fyodor Mikhailovich Dostoevsky was a famed Russian writer whose works reflect an intense interest in philosophical questions about the human condition. With some justification, Dostoevsky's thought has been linked with existentialism—it is unsystematic and sometimes paradoxical, and his fiction in particular is marked by a concern with the irrational in human behavior and with the burdens and blessings of free choice. In the full sweep of his writings, however—which included essays, notebooks, diaries, and letters in addition to fiction—Dostoevsky gave expression to a comprehensive Christian philosophy that cannot be classed as either existentialism or irrationalism, despite his influence on thinkers of both of those schools—the European (Friedrich Nietzsche, Albert Camus), as well as the Russian (Nikolai Berdyaev, Lev Shestov).

METAPHYSICS AND EPISTEMOLOGY

Dostoevsky's conception of the human situation is rooted most fundamentally in a traditional Christian dualism: Reality is divided into material and spiritual realms, at the intersection of which stands humanity. Matter and spirit are binary opposites for Dostoevsky, mutually exclusive in essence and attributes. And yet humans partake of both—a situation that generates metaphysical and epistemological puzzles.

As physical inhabitants of the material world, human beings are perishable entities, subject to laws of causal determination of the kind discovered by natural scientists. But as spiritual persons they are eternal and not fully determinable by natural causes. Dostoevsky's sympathies lay on the spiritual side, and accordingly the major part of his philosophizing was devoted to defending such idealist theses as the immortality of the soul (which he considered the basic tenet of Christian belief) and the doctrine of free will (the philosophical thesis with which he is most closely identified). At least six separate arguments for life after death can be found in his writings, beginning in 1864 in a lengthy diary entry on the death of his first wife—a passage of utmost importance for his philosophical outlook (Scanlan 2002, pp. 19–37). The significance of free will as a defining trait of humanity is memorably portrayed in his most pointedly philosophical work—*Notes from Underground* (1864)—in which he attacks the determinism of Nikolai Chernyshevsky and other Russian materialists, contending that human choices are radically unpredictable because people are capable of deliberately falsifying any prediction made. As Gary Saul Morson (1998) points out, the notion of an indeterminate future is central to Dostoevsky's narrative style as well as to his philosophical outlook.

The epistemological puzzle created by humanity's hybrid nature is how a spiritual soul mired in a material world, dependent on a physical brain and sensory apparatus, can fully understand either realm. At times Dostoevsky despaired of the mind's ability to comprehend reality at all, but more typically he stressed the partiality and tentativeness of human knowledge and the inability of science to fathom the human essence. He regarded reason as a limited capacity, denying that it could present conclusive proofs of such beliefs as personal immortality and the existence of God; at the same time, he accepted reason as consistent with and providing some support for those beliefs, as his own discursive arguments for them attest. In the voice of Father Zosima in *The Brothers Karamazov* (1879–1880) he also accepted mystical experience as a limited source of knowledge of reality: "Much on earth is concealed from us, but in place of it we have been granted a secret, mysterious sense of our living bond with the other world" (p. 320). Even this mysterious sense, however, tells us nothing more than that there is a "full synthesis of all being," which in the 1864 diary entry he identified with God (Proffer, vol. 1, 1973, p. 40). He did not reject the theistic notion of God as a person who created and rules the world, but he based that notion not on reason or mystical experience but solely on faith grounded in love.

ETHICS

Dostoevsky's ethical thinking was dominated by his opposition to egoism and defense of altruism as expressed in Christ's commandment to "love thy neighbor as thyself." His first major attack on egoism came in *Notes from Underground*, in the form of a devastating critique of the ethical theory (a form of enlightened egoism) espoused by Chernyshevsky and his followers. In the diary entry on the death of his first wife, Dostoevsky formulated the opposition between the Christian law of love and the egoistic force in human nature that opposed it, which he dubbed the law of personality. The struggle between these two laws, both rooted in the complex material-spiritual nature of humanity, remained central to Dostoevsky's writings—fiction and nonfiction alike—throughout his career. Despite his emphasis on free choice he did not regard freedom as the highest human

value; freedom is limited morally by the Christian law of love.

As the philosophical foundation for the law of love, Dostoevsky long relied on the idea that an inborn human conscience tells people authoritatively whether an action is right or wrong. Shortly before his death, however, he reluctantly admitted that conscience does not always speak univocally and that it may itself be evil; he concluded that morality has as its ultimate ground the religious faith that accepts the law of love as Christ proclaimed and lived it. Dostoevsky interpreted the law deontologically, as commanding or prohibiting actions as good or bad in themselves regardless of their results, thus rejecting utilitarianism. He vigorously opposed the idea, powerfully dramatized in both *Crime and Punishment* (1866) and *Demons* (1871–72) that an action abhorrent in itself may be justified by supposed future good consequences.

Two other recurring ethical themes in Dostoevsky's novels, particularly *Crime and Punishment* and *The Brothers Karamazov*, are also directly related to his devotion to the Christian moral ideal. These are the notions of universal moral responsibility ("I am responsible not only for my actions but for those of everyone") and the moral value of suffering. If essentially the ethical ideal is to be Christlike, it means freely accepting responsibility for others and suffering for their good, as Christ in the atonement took upon himself the sins of all humanity.

AESTHETICS

Dostoevsky's philosophy of art was laid out most fully in a polemical essay entitled "Mr. —bov and the Question of Art" (1861), directed against the so-called civic school of Russian criticism then represented most prominently by Nikolai Dobrolyubov. Just as Dostoevsky rejected utilitarian ethics, he had no sympathy for the view that art should be judged on the basis of its usefulness in promoting the satisfaction of basic human needs, such as the needs for food, shelter, and clothing.

Dostoevsky's argument against these critics was twofold. First, they failed to understand that human beings have aesthetic as well as material needs—specifically, a need for beauty, defined broadly in classical terms as "harmony and tranquility" (Magarshack 1997, p. 125), and a need to engage in creative activity—a notion reminiscent of the play theory of art advanced by Konrad Lange and Karl Groos. Second, Dostoevsky contended that utilitarian reasoning is a poor tool for determining the value of art, regardless of what needs it serves, for such reasoning rests on predicting the future impact of a work—something people cannot do with any confidence.

Dostoevsky did not deny that aesthetic values may have social and moral significance; beauty is not a narrowly aesthetic category for him. In *The Idiot* (1868) he describes Prince Myshkin as insisting that "beauty will save the world," presumably having in mind Beauty as producing harmony and tranquility in society (p. 382). But he vigorously denied that artists have a duty to engage in useful activity. Art, he argued, should be judged on the basis of its artistry, not its moral or social impact, and he defended the right of the artist to free scope for creativity.

SOCIAL PHILOSOPHY

A critic of Russian serfdom, Dostoevsky was drawn to European Enlightenment thinking in his youth and became active in clandestine revolutionary circles; in 1849 he was arrested and sentenced to nine years of imprisonment and exile in Siberia. He was never opposed in principle to the Russian imperial system of government, however, and upon his return to European Russia and the subsequent emancipation of the serfs in 1861 be became a champion of Russian autocracy and a severe critic of violent revolution, which he attacked most powerfully in the novel *Demons*. Through many journalistic articles, especially a long series entitled *A Writer's Diary* (1873, 1876–1881), he was an influential commentator on political, economic, and other social issues, writing from a Slavophile, nationalist perspective.

Dostoevsky's defense of autocracy was based on his conviction that the citizens of Russia willingly accepted a patriarchal hierarchy of social strata based on inequalities in talents and abilities. Such inequalities are not evils in Russia, he argued, because they are mutually acknowledged in an atmosphere of respect dictated by the Christian law of love. European political institutions designed to limit authority, he contended, were outgrowths of the history of the European states, which had their origin in the conquest of one people by another (such as the Gauls by the Franks) and were still characterized by hostility between rulers and ruled, unlike the harmony between the Tsar-father and his children that always existed in Russia. In Dostoevsky's idealized conception, an autocracy can be the freest state in the world, for its rulers need not fear their subjects.

Dostoevsky's aversion to the Russian revolutionaries extended to their economic program—socialism—because he considered it one of the great European evils threatening Russia's unique civilization. He called it, par-

adoxically, the height of egoism, because its appeal was to personal greed and the advancement of one's own rights against those of others. Above all, he saw socialism as destructive of human freedom: The revolutionary socialist, Dostoevsky argued, seeks the compulsory union of humanity by forcing economic change in the supposed interest of all. *Notes from Underground*, *Demons*, and *The Brothers Karamazov* all offer vivid treatments of this theme; the tale of the Grand Inquisitor in the latter novel is universally acclaimed as one of the most brilliant dramatic embodiments of philosophical ideas in world literature. Dostoevsky's remarkably prescient anticipation, in these and other works, of the aims and even the tactics of the twentieth-century Russian Bolsheviks has contributed to his reputation as a prophet.

PHILOSOPHY OF HISTORY

Scattered throughout Dostoevsky's published and unpublished writings are fragments of a nationalistic theory of world history that, although generally consonant with his ethical and religious views, has provoked much controversy because of the messianic mission it ascribed to Russia (particularly in later writings such as *A Writer's Diary*) and its seeming inconsistency with his conception of the future as radically undetermined and hence unpredictable.

In an early (1864–1865) notebook, Dostoevsky sketched three stages in the evolution of human society: (1) Primitive patriarchalism, in which humans live in unreflective community, lacking a concept of self; (2) Civilization, in which personal consciousness and egoism arise; community disintegrates and previously accepted patriarchal laws are questioned. This is a diseased condition, for it undermines faith in God and destroys the spontaneity of life; and (3) Christianity, in which there is a return to God, community, and spontaneity but on a conscious level: individuals voluntarily give themselves to others by accepting the law of love.

Dostoevsky's many discussions of national differences among peoples drew on this conception of levels of evolutionary progress. He believed that the Western European peoples, and even more the Jewish people wherever they resided, represented the diseased condition of egoism characteristic of the second stage of history. Russians, by contrast, as true Christians, are altruistic; furthermore they possess a unique trait he calls *universal responsiveness*, by virtue of which they comprehend and sympathize with the problems of all peoples of the world. The Russians, then, are the only nation firmly situated in the third stage of history—Shatov in *Demons* calls them "the only 'god-bearing' nation" (p. 247)—and it is their mission to raise others to that level by uniting them in a single loving community. As early as 1856 Dostoevsky had coined the expression *the Russian idea* for his nation's special role in world history. More than a century later, following the collapse of the Soviet Union, the term gained new life as the rallying cry of Russian nationalists.

See also Berdyaev, Nikolai Aleksandrovich; Camus, Albert; Chernyshevskii, Nikolai Gavrilovich; Egoism and Altruism; Enlightenment; Existentialism; Materialism; Nietzsche, Friedrich; Russian Philosophy; Shestov, Lev Isaakovich.

Bibliography

PRIMARY WORKS

The Brothers Karamazov: A Novel in Four Parts with Epilogue. Translated by Richard Pevear and Larissa Volokhonsky. New York: Vintage Classics, 1991.

Complete Letters. 5 vols, edited and translated by David Lowe and Ronald Meyer. Ann Arbor, MI: Ardis, 1988–1991.

Crime and Punishment. Translated by Richard Pevear and Larissa Volokhonsky. New York: Vintage Classics, 1992.

Demons: A Novel in Three Parts. Translated by Richard Pevear and Larissa Volokhonsky. New York: Alfred A. Knopf, 1994.

The Idiot. Translated by Richard Pevear and Larissa Volokhonsky. New York: Vintage Classics, 2003.

Notes from Underground. Translated by Richard Pevear and Larissa Volokhonsky. New York: Vintage Classics, 1994.

Complete Collected Works in Thirty Volumes. Leningrad, Russia: Nauka, 1972–1990.

A Writer's Diary. 2 vols. Translated by Kenneth Lantz. Evanston, IL: Northwestern University Press, 1993–94.

SECONDARY WORKS

Bakhtin, Mikhail Mikhailovich. *Problems of Dostoevsky's Poetics,* edited and translated by Caryl Emerson. Minneapolis: University of Minnesota Press, 1984.

Frank, Joseph. *Dostoevsky*. 5 vols. Princeton, NJ: Princeton University Press, 1976–2002.

Knapp, Liza. *The Annihilation of Inertia: Dostoevsky and Metaphysics*. Evanston, IL: Northwestern University Press, 1996.

Kostalevsky, Marina. *Dostoevsky and Soloviev: The Art of Integral Vision*. New Haven, CT: Yale University Press, 1997.

Lauth, Reinhard. *Die Philosophie Dostojewskis in systematischer Darstellung ("Ich habe die Wahrheit Gesehen")*. Munich, Germany: R. Piper, 1950.

Magarshack, David, ed. and tr. *Dostoevsky's Occasional Writings.* Evanston, IL: Northwestern University Press, 1997.

Morson, Gary Saul. "Dostoevskii, Fëdor Mikhailovich (1821–1881)." In *Routledge Encyclopedia of Philosophy*, edited by Edward Craig, vol. 3, pp. 114–119. London: Routledge, 1998.

Proffer, Carl R., ed. *The Unpublished Dostoevsky: Diaries and Notebooks (1860–81) in Three Volumes*. Ann Arbor, MI: Ardis, 1973–1976.

Scanlan, James P. *Dostoevsky the Thinker*. Ithaca, NY: Cornell University Press, 2002.

James P. Scanlan (2005)

DOUBLE TRUTH, DOCTRINE OF

See *Averroism*

DOUBT

To be in doubt about a proposition is to withhold assent both from it and from its contradictory. Although people sometimes withhold assent with no reason for doing so and persist in this even after conceding that they have no reason, doubt is rational only when one has a reason for it and reasonable only when the reason is a good one. Doubt may be accompanied by various feelings, but it seems unlikely that there are specific feelings uniquely associated with it; in general, the feelings associated with doubt are anxiety or hesitation, which are identified as feelings of doubt when they arise in contexts involving questions of belief. In any case, philosophers are not ordinarily concerned with psychological characterizations of a doubter's state of mind. Their attention is primarily devoted to understanding the conditions under which doubt is reasonable and to defining the limits of reasonable doubt.

EVIDENCE AND REASONABLE DOUBT

Whether it is reasonable for a person to doubt a proposition cannot always be decided solely by considering the evidence that the person possesses relevant to the proposition or, in a situation in which there is purportedly noninferential knowledge, by considering his ground for assent. Doubts that are unreasonable or absurd in one situation may be quite reasonable in another, although the available evidence or ground is the same in both cases. For example, special caution is appropriate when the penalties for error are particularly great; hence, an ordinarily acceptable basis for assent may be inadequate if much depends upon avoiding error, although the gravity of the risk does not in itself constitute evidence. Moreover, a basis for assent that would be entirely compelling in normal circumstances may be insufficient if otherwise remote possibilities of error must be taken seriously because of threats posed by a resourceful deceiver.

From the fact that someone has no reason to doubt a given proposition, therefore, it does not follow that the evidence he possesses is sufficient to render unreasonable *all* doubts concerning the proposition. It would seem quite worthwhile to explore the ways in which the reasonableness of doubt is affected by considerations other than the available evidence or ground for assent. Philosophers, however, on the whole, are interested only in very general principles that are not affected by contingencies of any sort. For this reason, perhaps, philosophical studies of doubt have usually been concerned with limiting cases in which the reasonableness of doubt depends only on the available evidence or ground for assent. In other words, they have dealt mainly with what is *indubitable*—with what it is never reasonable to doubt regardless of contextual variables of the sort described above. Accordingly, a philosopher's designation of certain propositions as *dubitable* is not generally to be understood as a denial that there are circumstances in which doubting these propositions would be absurd. The designation means only that given the evidence or ground for the propositions, there are conceivable circumstances in which doubt would be reasonable.

CONDITIONS OF INDUBITABILITY

Toward the end of the First Meditation, René Descartes invokes the distinction between what is indubitable and what, in normal circumstances, is open to no reasonable doubt. In defense of his decision to regard as dubitable many propositions which, in practice, it is unreasonable to doubt, he declares, "I cannot at present yield too much to distrust, since it is not now a question of action but only of meditation and of knowledge." In their usual concerns, individuals are not often required to decide whether a proposition is indubitable, as distinct from deciding whether there is any reason to doubt it. Questions of indubitability are theoretical: they concern only the relation between a proposition and the evidence or ground for it, and take no account of the other concrete circumstances in which a proposition is evaluated.

LIMITS OF RELEVANT EVIDENCE. When is one entitled to regard a proposition as indubitable? It might be maintained that one is not entitled to do so as long as anything which can serve as evidence relevant to the proposition remains unexamined, on the ground that when this evidence comes to be examined, it may turn out to require an alteration of belief. But by virtue of the empirical and logical connections among facts, the truth-value of any proposition affects the truth-values of an unlimited number of others: Hence, the truth-values of an unlimited

number of propositions are relevant to that of any proposition and may serve as evidence concerning it. Since it is impossible to examine each of these other propositions, no proposition could ever be regarded as indubitable if it were first necessary to examine everything that may serve as evidence relevant to its truth-value. On the other hand, it seems that this impasse can be avoided only if it is possible to settle in advance the import of matters that have not been examined.

IMMEDIATE EXPERIENCE. That it is in fact possible to settle the import of matters that have not been examined may be brought out as follows. The impossibility of checking all the consequences of an empirical proposition is often cited to support the view that empirical propositions must always remain dubitable. Nonetheless, many philosophers who employ this argument concede the indubitability of so-called "basic propositions," or a person's current reports of the immediate contents of his consciousness (for example, pains, sense data, thoughts). But however fragmented and ephemeral immediate experiences may be, they are not without innumerable conditions and consequences. Like those of empirical propositions (statements of fact about the world outside immediate consciousness), the truth-values of basic propositions are connected with those of an unlimited number of other propositions which may be construed as evidence relevant to them. Hence, if a person's current reports of the immediate data of his own consciousness are indubitable, it is not because he has surveyed everything that may serve as evidence relevant to them: rather, it is because his ground for making the report is such that he cannot reasonably acknowledge that any evidence could supersede it. Indeed, it is reasonable for him to require that all evidence be interpreted so as to be consistent with his report.

INCORRIGIBILITY. When one proposition serves as evidence relevant to a second, it does so by virtue of certain other empirical or logical propositions (laws or rules) by which the two are connected. The connection may be broken or its nature altered, however, if the intermediary propositions upon which it depends are rejected or revised. Thus, the possibility of coming upon contrary evidence can be excluded by requiring that this alternative be adopted whenever necessary.

But under what conditions is it reasonable to make such a requirement of incorrigibility—to arrange that nothing count as evidence against a certain proposition? In some cases (for instance, when a mathematical proposition is supported by a well-understood proof, or when a basic proposition is grounded in immediate experience) it may seem fairly clear that the conditions are satisfied. However, philosophers have failed to provide a general account of these conditions; instead, they have usually limited themselves to identifying particular instances of their satisfaction. Some philosophers have claimed with considerable plausibility that certain elementary mathematical propositions (such as that $2 + 2 = 4$) may be regarded as indubitable without proof, but they have done little to explain systematically why this should be so. With regard to empirical propositions, neglect of the problem of clarifying the conditions in which they may be accepted as indubitable has resulted in part from widespread controversy over whether the problem properly arises at all. That there are no such conditions is frequently maintained by philosophers (for example, Bertrand Russell, A. J. Ayer, C. S. Peirce, C. I. Lewis) who subscribe to certain popular epistemological doctrines—in particular, the doctrines that every empirical proposition is to be construed on the model of a scientific hypothesis, or that it is to be interpreted phenomenalistically as equivalent to an unlimited number of predictions.

LOGICAL CONTINGENCY AND NECESSITY

A more general obstacle to a sound understanding of the basis of indubitability lies in a tendency to look for it in the wrong place. A proposition is indubitable when there could be no reason to doubt it, but this impossibility is not in general inherent in the logical character of the proposition itself. Indubitability is an epistemic property that depends on the relation between a proposition and the evidence or ground for assent with which it is considered. In particular, dubitability and indubitability must not be confused with logical contingency and logical necessity. The logical contingency of a proposition does not as such entail that no one has conclusive evidence or ground for it, and a logically necessary proposition may reasonably be doubted by someone who is not in a position to appreciate its necessity and who therefore must concede the possibility that further inquiry will uncover evidence against it.

Moreover, it is a mistake to suppose that evidence for a proposition is not conclusive unless its conjunction with the denial of the proposition is self-contradictory. To be sure, a proposition is indubitable if and only if no basis for assenting to its alternative is conceivable, but something may be inconceivable even though it contradicts neither itself nor what has already been established.

CONDITIONS OF RATIONAL INQUIRY

The claim that a basis for doubt is inconceivable is justified whenever a denial of the claim would violate the conditions or presuppositions of rational inquiry. Avoidance of self-contradiction is perhaps the most familiar of these conditions, but it is not the only one. For instance, since inquiry is fundamentally an attempt to discriminate between what is to be accepted and what is to be rejected, nothing can rationally be conceived which involves denying the necessity for making these discriminations or undermining the possibility of making them.

A systematic explanation of dubitability and indubitability awaits, therefore, a general theory of the nature of rationality which illuminates the presuppositions and conditions that rationality requires. Furthermore, it awaits an account, developed from this theory, of the particular conditions in which propositions of various sorts must be regarded as indubitable if the possibility of rationality is to be preserved. Even if this were done, however, a further problem would remain. While an adequate theory of rationality would give a clear account of the conditions in which a proposition may reasonably be regarded as indubitable, it cannot of course guarantee that these conditions are correctly identified in any given case. To support the claim that a certain proposition is indubitable, it is not sufficient to understand the conditions in which such claims are justified; it is also necessary to know that the conditions are fulfilled in the particular case in question.

THE INDUBITABILITY REGRESS

A disturbing pattern of argument seems to develop, however, in considering the proposition that a given proposition is indubitable. The proposition that the conditions for the indubitability of a certain proposition have been satisfied cannot itself be regarded as beyond doubt unless the conditions for *its* indubitability have been satisfied; but the satisfaction of *these* conditions is dubitable unless …, and so on.

But acknowledging this regress does not require one to concede that it is never reasonable to regard a proposition as indubitable. Rather, the view to which the regress leads appears to be that while there are occasions on which it is reasonable to regard a proposition as indubitable, it is never altogether indubitable just which occasions these are. There is an air of paradox here, perhaps, but there is no logical difficulty. The regress does not interfere with the possibility of there being satisfactory logical relations between indubitability claims and judgments establishing that these claims are reasonable. It only interferes with our confidence in ourselves, suggesting that there is always room for doubt as to whether we are being reasonable. Or, to put the matter a bit differently, the regress supports no more than the mordant comment that it is never reasonable to insist that the question of whether one is being reasonable is entirely closed.

See also Ayer, Alfred Jules; Descartes, René; Error; Experience; Knowledge and Belief; Lewis, Clarence Irving; Peirce, Charles Sanders; Propositions; Questions; Russell, Bertrand Arthur William; Skepticism, History of.

Bibliography

In addition to loci classici in Descartes, *Meditations on First Philosophy*; David Hume, *An Enquiry concerning Human Understanding*; and Immanuel Kant, *Critique of Pure Reason*, there are particularly interesting and relevant discussions in the following more recent works: A. J. Ayer, *The Problem of Knowledge* (London: Macmillan, 1956); C. I. Lewis, *An Analysis of Knowledge and Valuation* (La Salle, IL: Open Court, 1946); N. Malcolm, "Knowledge and Belief," in *Knowledge and Certainty* (Englewood Cliffs, NJ: Prentice-Hall, 1963); G. E. Moore, "Four Forms of Scepticism," in *Philosophical Papers* (London: Allen and Unwin, 1959); and C. S. Peirce, *Collected Papers* (Cambridge, MA, 1935).

OTHER RECOMMENDED TITLES

Buryeat, M. F., ed. *The Skeptical Tradition*. Berkeley: University of California Press, 1983.

DeRose, Keith, and Ted Warfield, eds. *Skepticism*. New York: Oxford University Press, 1999.

Fogelin, R. *Pyrrhonian Reflections on Knowledge and Justification*. Princeton, NJ: Princeton University Press, 1995.

Greco, John. *Putting Skeptics in their Place*. New York: Cambridge University Press, 2000.

Hankinson, R. J. *The Sceptics*. London: Routledge, 1995.

Klein, Peter. *Certainty: A Refutation of Scepticism*. Minneapolis: University of Minnesota Press, 1981.

Lehrer, Keith. "Why not Scepticism?" *Philosophical Forum* 2 (1971): 283–298.

Roth, Michael, and Glenn Ross. *Doubting: Contemporary Perspectives on Skepticism*. Boston: Kluwer, 1992.

Sosa, Ernest. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Sosa, Ernest, and Enrique Villanueva, eds. *Skepticism*. Cambridge, MA: Blackwell, 2000.

Stroud, Barry. *The Significance of Philosophical Skepticism*. Oxford: Oxford University Press, 1984.

Williams, Michael. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. Oxford: Blackwell, 1992; corrected edition, Princeton, NJ: Princeton University Press, 1996.

Harry G. Frankfurt (1967)
Bibliography updated by Benjamin Fiedor (2005)

DRAMA

See *Greek Drama; Tragedy*

DREAMS

Almost all of us have had dreams, yet few could say with confidence what they are, beyond agreeing that they occur during sleep and have some likeness to waking experience. Yet most people would in all probability accept the kind of definition given by philosophers, for example Plato's "visions within us, … which are remembered by us when we are awake and in the external world" (*Timaeus,* 46A) or Aristotle's "the dream is a kind of imagination, and, more particularly, one which occurs in sleep" (*De Somniis,* 462a). Indeed, such notions seem to be summarized in the *Oxford Dictionary*'s definition: "A train of thoughts, images, or fancies passing through the mind during sleep; a vision during sleep." Dreams are striking phenomena, and the more superstitious see in them signs and portents of what is to happen; even today divination by dreams has not lost its popularity. A more sophisticated way of looking at dreams is to regard them as revealing something about the sleeper, either about his physical condition or about his mental state. An example of the former can be seen in the diagnostic technique used in the temple of Aesculapius; patients seeking a cure had to sleep all night in the temple precincts and would experience a "vision" that would indicate the disease or its cure. Many writers had suggested that mental states were revealed by dreams, but there was little serious study of the idea until the work of Sigmund Freud and his followers. Freud's doctrine of the unconscious, and the way in which it is revealed in dreams and other less rational activities, is important for psychiatry; but he had little to say about the nature of dreams that is of interest to the philosopher, though the fact that they had been found worthy of study may have resulted in an increase in philosophic concern about the problems they raise.

While we are having them, dreams often appear to be as real as waking experience; children have to be told that the object of their terror "was only a dream," hence not part of the world. William James expressed this well in his *Principles of Psychology*: "The world of dreams is our real world whilst we are sleeping, because our attention then lapses from the sensible world. Conversely, when we wake the attention usually lapses from the dream-world and that becomes unreal." This similarity has led philosophers to pose the question, "How can you prove whether at this moment we are sleeping, and all our thoughts are a dream; or whether we are awake, and talking to one another in the waking state?" (Plato, *Theaetetus,* 158). In perhaps the most famous example of the difficulty of distinguishing dreams from reality, René Descartes introduced his method of universal doubt. He concluded, "I see so manifestly that there are no certain indications by which we may clearly distinguish wakefulness from sleep that I am lost in astonishment" (*First Meditation*). Descartes finally resolved his doubts in this respect by appealing to a criterion of consistency: "For at present I find a very notable difference between the two, inasmuch as our memory can never connect our dreams with one another, or with the whole course of our lives, as it unites events which happen to us while we are awake" (*Sixth Meditation*). Such a consistency criterion has been adopted by several more recent writers on the topic. Unfortunately, this will not do the task required, for consistency can only be used as a test of a particular experience by waiting to see what happens in the future. It would enable me to tell that I had been dreaming, not that I am now dreaming; for however confident I am of the reality of my surroundings, something may happen in the future that will reveal them to be part of a dream. Further, the problem remains whether any consistency discovered is a real or a dreamed one.

The failure of consistency to provide a test need not be worrying, for the times in which genuine doubt arises are normally those involving memory—I am not sure if this event actually happened or whether I dreamed it. In such a case I would normally try to remember some part of the event that would have left a mark in the physical world, and then see if there is such a trace of the event; if there is nothing, I conclude that I had dreamed the occurrence. In spite of Descartes's remark, it is rare that we are in doubt about whether we *are* dreaming. The expression "I must be dreaming" is normally used in circumstances when I am quite sure that I am not dreaming, to express surprise at some pleasant occurrence, for example the arrival of a friend whom I thought to be somewhere distant. There are times when we are aware that we are dreaming, though normally a dream presents itself as real and no questions about its genuineness arise. It seems that the conviction that one is dreaming does not come from a previous doubt within the dream about the status of the experience; it just occurs, though sometimes accompanied with a feeling of relief. But in most cases the dream convinces us that it is reality, in that no doubt or questioning arises during its course. The difference between dreams and hallucinations lies in the fact that there is nothing external to dreams with which they can be compared, no tests that can be applied. For if we did

apply a test in a dream, the result would be to confirm its reality. Philosophers have sought for some mark or test that would solve this problem, but there is none available. Any suggested sign of reality could be duplicated in the dream, and if all dreams bore marks of unreality, then there could not even be confusion over the remembering of them.

It has been generally agreed that dreams are due to the workings of the imagination no longer under the control of the intellect or the senses, as can be seen from the quotations at the beginning of this article; but it would seem that in such contexts the meaning of the word "imagination" had been left vague, serving rather as an indication of puzzlement than as a solution to a problem. Some recent work by physiologists has led to the suggestion (by W. Dement and N. Kleitman) that dreaming is correlated with rapid eye movements during sleep. Such a suggestion would seem to confirm Aristotle's remark that "dreaming is an activity of the sensitive faculty, but of it as being imaginative" (459a). The use of a physiological criterion for dreams has been challenged by Norman Malcolm in his book *Dreaming* (1959), which is clearly the most important contemporary discussion of the whole topic. In the course of it he challenges virtually all the assumptions made by previous philosophers. In criticism of the physiological work, he asserts that waking testimony is the sole criterion of dreaming (p. 81). The obvious difficulties that arise from the common belief that external stimuli can cause or influence the course of a dream, or that observers can sometimes tell from bodily movements that a sleeper is having a violent dream, he dismisses by means of a definition that dreams can take place only when the subject is sound asleep and that a person who is sleeping cannot respond to external stimuli (pp. 25–26). It might be thought that Malcolm was here doing the same thing for which he criticizes the physiologists, namely introducing a new concept of dreaming, for surely the ordinary unsophisticated notion includes the possibility of our recognizing that someone asleep is having a dream, in some cases at least, as well as the possibility of the dreamer being aware that he is dreaming. If both of these beliefs are ruled out by a philosophical argument, then it would appear that the concept of dreaming held by most people has been changed in important ways. Most of the points made in the earlier part of this article would be understood by those with an unsophisticated notion of dreaming.

Malcolm's arguments are, however, powerful and subtle, and his critics, of whom A. J. Ayer is perhaps the most eminent, have found it not at all easy to refute them. Malcolm bases his reasoning on Ludwig Wittgenstein's *Philosophical Investigations,* in particular on the dictum that "an 'inner process' stands in need of outward criteria" (I. § 580). Malcolm argues that we can come by the concept of dreaming only by learning it from descriptions of dreams, "from the familiar phenomenon that we call 'telling a dream'" (II, p. 55). To talk of "remembering a dream" is to use the word *remember* in a sense different from the normal, for there is no external criterion by which we can check our memory, as there is in the paradigm cases of remembering, that of remembering an event in the public world, which can be checked by ourselves and others. What is told sincerely on waking is the dream, because there is no other way of finding out what, if anything, occurred while the teller slept. (This can be compared with Freud's reliance on the narration of the dream, but this was essential for its use in diagnosis. Nevertheless, Freud was willing to evaluate critically the veracity of actual dream accounts on the basis of his theory or as a result of previous analysis of its dreamer. For most purposes, it made no difference whether the dream account or the dream itself was being considered; Freud's concern was with different problems.)

Yet Malcolm rejects Ayer's suggestion that this theory amounts to saying that "we do not dream, but only wake with delusive memories of experiences we have never had." Malcolm is clearly correct in stressing the importance of the report of a dream and its difference from reports of public events; what the dreamer says on waking is final. Though we must learn the use of the word *dream* in the way Malcolm indicates, this does not rule out the possibility of its use being extended by further experience, for instance, correlating dream reports with observations of the dreamer, as Dement and Kleitman have done. The trouble is Malcolm's use of the term *criterion,* which is never clearly explained, and which seems to lead him into a crude verificationism; he even talks of "the senselessness, in the sense of the impossibility of verification, of the notion of a dream as an occurrence" (p. 83). A further consequence of Malcolm's use of the dream report as a criterion for dreaming is that it becomes impossible to talk of children having dreams before they have learned to speak (p. 59). If, as Malcolm apparently wishes to maintain, words can be used only if their application can be strictly verified, then many ordinary uses will be cut out. That we now have a particular concept of some mental activity does not make it impossible that further experience will lead us to introduce a modification of it, in which case the way in which we first learned it may have no bearing on the criterion of its use. For example, many words used in the sciences are first

learned in an approximate way and their criteria of application refined in the course of education. Malcolm claims that his argument applies only to words that refer to "inner" processes. What he seems to do, however, is extend Wittgenstein's argument, valid in the area Wittgenstein intended it for, beyond its legitimate sphere. The primary use of the word *dreaming* depends upon the notion of telling a dream, but this does not prevent an extended use. Peter Geach remarks that Wittgenstein mentioned in a lecture Lytton Strachey's description of Queen Victoria's dying thoughts: "He expressly repudiated the view that such a description is meaningless because 'unverifiable'; it has meaning, he said, but only through its connexion with a wider, public, 'language-game' of describing people's thoughts" (*Mental Acts,* p. 3). In fact it is only because we know what it is to dream that we can understand the difficulties raised by talk of "verifying" reports of dreams.

Ayer also criticizes Malcolm's denial that one can make assertions while asleep, but in this case with less effect. It does seem clear that the words "I am asleep" cannot be used to make a genuine assertion, because such an utterance would contradict what was asserted, just as the only possible truthful reply to the question, "Are you asleep?" is "No." An absence of reply is what would lead the questioner to assert that the man was really asleep.

In spite of Malcolm's statement (p. 66) that there is no place for an implication or assumption that a man is aware of anything at all while asleep, many would claim, and understand others' claims, that they had become aware that they were dreaming. This also implies that they were aware that they were asleep. As part of a dream narrative, such awareness could be reported by the words, "I suddenly realized that it was all a dream." Clearly, such an assertion could not be taught by ostensive means. However, there seems no reason why, having learned how to use the ordinary concept of dreaming and expressions such as "I suddenly realized that," we should not combine the two into an assertion that would be commonly understood to apply to a possible experience. Malcolm's claim that a person must be partially awake to be aware that he is dreaming (pp. 38–44) seems, as suggested above, a redefinition of the term for which no adequate reason is advanced.

Malcolm wishes to say that the problem of what dreams are is a pseudo problem; he refuses to allow that they can be called experiences, illusions, workings of the imagination, or anything else they have been thought to be by previous philosophers. Ayer concludes his criticism of *Dreaming* by maintaining that dreams are experiences and mostly illusions, and "are found to be so by the same criteria that apply to illusions in general." This remark is difficult to understand; here Malcolm's stress on the report of the dream comes into its own; in recounting it I am not claiming that these things happened. Because while dreaming there is no possibility of making assertions about my experiences to other people, to describe dreams as illusions makes no sense. Malcolm has clearly made out his case in this respect. On the other hand, it seems difficult to deny that dreams are experiences, if only because the description is sufficiently vague to cover almost any "mental" phenomena. The same may be said of talking of dreams as being composed of images; here dreaming is being used as one of the examples of mental imagery, a vague concept. In spite of Malcolm's work, the problem of the nature of dreaming is still open for philosophic discussion, but any future examination of the problem will have to take his book fully into account. Many philosophers would still wish to assert that dreams occur, that they take place during sleep, while admitting that the meaning and justification of such claims is by no means clear.

See also Aristotle; Ayer, Alfred Jules; Descartes, René; Freud, Sigmund; Imagery, Mental; James, William; Malcolm, Norman; Plato; Psychoanalysis; Unconscious; Wittgenstein, Ludwig Josef Johann.

Bibliography

HISTORICAL WORKS

Aristotle. *De Divinatione per Somnum* (*On Prophesying by Dreams*). Translated by J. I. Beare, in *Works of Aristotle,* Vol. III. Oxford: Clarendon Press, 1931.

Aristotle. *De Somniis* (*On Dreams*). *Ibid.*

Bradley, F. H. "On My Real World." In *Essays on Truth and Reality.* Oxford: Clarendon Press, 1914.

Descartes, René. *Meditationes de Prima Philosophia.* In *The Philosophical Works of Descartes,* Vol. I, translated by E. S. Haldane and G. R. T. Ross. Cambridge, U.K., 1934.

Freud, Sigmund. *The Interpretation of Dreams.* Edited and translated by James Strachey. New York: Basic, 1955.

Freud, Sigmund. *Introductory Lectures on Psycho-analysis.* Translated by J. Riviere. London, 1949.

James, William. *The Principles of Psychology.* New York: Dover, 1950.

Malcolm, Norman. *Dreaming.* London: Routledge, 1959.

Plato. *Theaetetus* (158A). Translated by B. Jowett. Oxford, 1871.

Plato. *Timaeus* (46A). *Ibid.*

Russell, Bertrand. *Our Knowledge of the External World.* London, 1949.

Sartre, Jean-Paul. *L'imaginaire. Psychologie phénoménologique de l'imagination.* Paris: Gallimard, 1940. Translated by B. Frechtman as *The Psychology of the Imagination.* London, 1949.

Wittgenstein, L. *Philosophical Investigations.* Translated by E. Anscombe. Oxford: Blackwell, 1953. Especially pp. 184, 222–223. Also relevant is Malcolm's review of this work in *Philosophical Review* (October 1954): 530–559.

ARTICLES

Ayer, A. J. "Professor Malcolm on Dreams." *Journal of Philosophy* 57 (1960): 517–535. Malcolm's reply and Ayer's rejoinder in *ibid.* 58 (1961): 294–299.

Chappell, V. C. "The Concept of Dreaming." *Philosophical Quarterly* 13 (July 1963): 193–213.

Dement, W., and N. Kleitman. "The Relation of Eye Movements during Sleep to Dream Activity: An Objective Method for the Study of Dreaming." *Journal of Experimental Psychology* 53 (1957): 339–346.

MacDonald, M. "Sleeping and Waking." *Mind* 62 (April 1953): 202–215.

Manser, A. R., and L. E. Thomas. "Dreams." *Proceedings of the Aristotelian Society* Suppl. Vol. 30 (1956): 197–228.

Putnam, H. "Dreaming and 'Depth Grammar.'" In *Analytical Philosophy,* edited by R. J. Butler. Oxford, 1962.

A. R. Manser (1967)

DRETSKE, FRED

(1932–)

Born in 1932, Fred Dretske received his PhD from the University of Minnesota. He is emeritus professor of philosophy at Stanford University and professor of philosophy at Duke University. Since the early 1970s Dretske's work has been at the center of a number of key disputes in epistemology and the philosophies of perception, mind, and consciousness. Despite their range, two basic motivations unify Dretske's writings: the need to understand the mind in relation to its environment and a steadfast naturalistic outlook on the mind and its operations.

In *Seeing and Knowing* (1969), Dretske emphasized a form of perception that he labeled "nonepistemic seeing." This is an direct relation between perceiver and object not involving any particular conceptualization of the perceived object nor requiring any particular beliefs about it. Dretske argued that the concept of nonepistemic seeing is fundamental to understanding perception and the place of the mind within the world. Without it we have no way of understanding how we can all experience the same world despite having widely divergent concepts and beliefs. Via the notion of nonepistemic seeing, we can strip away our cognitive interpretive faculties and be left with the content of perception: the *objects* of the world we perceive.

Attention to nonepistemic seeing also undercuts the old idea that seeing involves "direct acquaintance" with some mysterious mental object, from whose incorrigibly known features we can only *infer* the existence of the external world.

We might naturally ask, What is the basic enabling feature of nonepistemic seeing? The answer is that there is an internal state of the perceiver that "carries the information" about the seen object. In *Knowledge and the Flow of Information* (1981), Dretske developed a sophisticated, elaborate, and technical account of information and its role in knowledge, thought, and perception. Building on his earlier epistemological work, Dretske analyzed knowledge in terms of *informationally caused* beliefs. To take one of Dretske's famous examples (from 1970), someone at a zoo *knows* that there is a zebra in front of him if that very information is causing his belief. Whether the appropriate information is available depends on the context of its occurrence, since information is a function of the *relevant* alternative messages that a signal could deliver. If there are lots of cleverly disguised mules about, his belief may not be caused by the information that there is a zebra in front of him (since the presence of that information may depend upon how much the perceiver knows about how zebras look), and thus he may not know that there is a zebra in front of him.

Dretske's account has an infamous consequence: the denial of inferential knowledge via known entailments. If our subject knows that these (the creatures in front of him) are zebras and that it follows from *x*'s being a zebra that *x* is *not* a disguised mule, then it would seem he could infer that these are not mules and hence *know* this. But how could he know this when he is utterly unable to distinguish a painted mule from a zebra? Dretske asserted that someone could know that something is a zebra without knowing that it is not a painted mule. While the mechanics of information allow this "paradox," the general issue remains highly contentious.

How can information or content play a causal role in the world? This is a key issue for Dretske's project of naturalizing the mind, or as Dretske puts it, baking "a mental cake with physical yeast and flour." Crudely put, the problem is that all behavior appears to have purely physical explanations that need appeal not to any information but only to local causes. We know how charge, momentum, and gravity cause events; informational causation seems to be something else altogether and quite mysterious.

In *Explaining Behavior* (1988), Dretske addressed this problem via a distinction between "triggering" and "structuring" causes. If *C* is an efficient or local cause of *M*, it is a triggering cause. The structuring cause of *M* is

"the processes which explain why *C* causes *M*" (p. 91). In particular, the structuring causes of behavior are the historical processes that institute the triggering causal links between information-carrying mental states and behavior. Two aspects of this sort of explanation must be distinguished. The first comprises the historical processes—evolution, learning, or design—by which some internal state comes to have an "indicator function." The second is the deployment of the indicator to modify behavior *because* of what is indicated. Dretske maintains that, while a great many states serve to carry information of one sort or another and while these states certainly do enter into causal relations, only *learning* can bring about systems in which the carried information causally explains why these states cause the behavior they do. Only in learning do "we see meanings … doing some real work in shaping behavior" ("Dretske's Replies," p. 201).

The emphasis on learning leads to obvious difficulties. It seems to imply that innate mental states cannot explain behavior (perhaps cannot even *cause* behavior and maybe cannot even exist). In *Naturalizing the Mind* (1995), Dretske, elaborating his view, allowed that evolutionary processes can produce representational mental states that do not depend on learning for their efficacy. There he distinguished *systemic* and *acquired* representational states. The former are the experiential qualities of experience. Their content is nonconceptual and fixed by biology. Systemic representation underpins nonepistemic perception. And it enables acquired representations, a form of which constitute beliefs and the other propositional attitudes. This distinction allows for a more nuanced theory of mind and forms the basis for an ambitious representational theory of consciousness. In *Naturalizing the Mind*, Dretske also develops an intriguing theory of introspection in which our self-knowledge involves a special application of mentalistic concepts to our own experience.

Dretske continues to claim that representation is essentially linked to the external environment. In his theory of consciousness, the experiential nature of mental states depends on their representational properties (and all conscious states, including such "pure" sensations as pain or tickles, are conceived of as representational). While promising a complete naturalization of the most troublesome feature of the mind, representational properties have a downside. Since representational properties are determined and constituted by relations with the environment, Dretske's views have the consequence that a newly created *duplicate* of a person would utterly lack consciousness. Many find this less than plausible.

Be that as it may, Dretske presents an elegantly unified and comprehensive theory of mind that makes our mental lives fully causal in an entirely naturalistic way. His views, in their clarity, argumentative care, and intellectual honesty, exemplify the best features of modern analytic philosophy.

See also Consciousness; Content, Mental; Introspection; Mental Causation; Perception, Contemporary Views; Philosophy of Mind; Relevant Alternatives.

Bibliography

WORKS BY DRETSKE

Seeing and Knowing. Chicago: University of Chicago Press, 1969.

"Epistemic Operators." *Journal of Philosophy* 67 (24) (1970): 1007–1023.

Knowledge and the Flow of Information. Cambridge, MA: MIT Press, 1981.

"Misrepresentation." In *Belief: Form, Content, and Function*, edited by Radu J. Bogdan. Oxford, U.K.: Oxford University Press, 1986.

Explaining Behavior. Cambridge, MA: MIT Press, 1988.

"Dretske's Replies." In *Dretske and His Critics*, edited by Brian P. McLaughlin. Oxford, U.K.: Blackwell's, 1991.

"Conscious Experience." *Mind* 102 (406) (1993): 263–283.

Naturalizing the Mind. Cambridge, MA: MIT Press, 1995.

Perception, Knowledge, and Belief. Cambridge, U.K.: Cambridge University Press, 2000.

WORKS ON DRETSKE

McLaughlin, Brian P., ed. *Dretske and His Critics*. Oxford: Blackwell's, 1991.

William Seager (2005)

DRIESCH, HANS ADOLF EDUARD

(1867–1941)

Hans Adolf Eduard Driesch, perhaps the outstanding representative of neovitalism, was born at Bad Kreuznach, Germany. His father, Paul Driesch, was a merchant in Hamburg. From 1877 Hans Driesch attended the Johanneum (a humanist gymnasium) in his native city, graduating with honors in 1886. He then studied zoology, first under A. Weismann at Freiburg, then at Munich, and finally under Ernst Haeckel at Jena, receiving his Ph.D. in 1889; his dissertation was titled "Tektonische Studien an Hydroidpolypen" (Tectonic studies of hydroid polyps).

DEVELOPMENT OF DRIESCH'S THOUGHT

Reacting to arguments advanced by G. Wolff, W. His, and A. Goette, Driesch early became skeptical of Haeckel's mechanistic interpretation of the organism. The work of Wilhelm Roux, in particular, induced him to explore the whole vitalism-mechanism issue. Driesch's first publication, *Die mathematisch-mechanische Behandlung morphologischer Probleme der Biologie* (Mathematico-mechanical treatment of morphological problems of biology; Jena, 1890), led to a break with Haeckel. Then, following Roux's example, Driesch put the embryogenetic theory of His and Weismann to an experimental test. His and Weismann had held that morphogenetic development of the living organism could be explained by assuming that a specifically organized yet invisible structure of great complexity is contained in the nucleus of the germ cell and that the gradual unfolding of this structure, through nuclear division, determines the course of every ontogeny. Roux's experiments, in 1888, had seemed to confirm this theory of "tectonic preformation." When he destroyed one of the blastomeres at the two-cell stage, the remaining one would develop into a half embryo—either the left half or the right half, depending on which blastomere had been destroyed. Driesch merely intended to provide further confirmation of these facts. But where Roux had experimented with the egg of a frog, Driesch used eggs of the sea urchin. Against all expectations he found that each blastomere of the two-cell stage of a sea urchin egg developed into a whole embryo half the normal size. This was the opposite of Roux's results and was irreconcilable with the His-Weismann theory.

While at the Marine Biological Station in Naples, from 1891 to 1900, Driesch continued his experimental investigations, confirming and reconfirming in startling ways his earlier findings, and began to formulate his own theory. Relevant to the development of his ideas was a study of Otto Liebmann's book *Analysis der Wirklichkeit* (Analysis of reality) and of the writings of Immanuel Kant, Arthur Schopenhauer, René Descartes, John Locke, and David Hume. Alois Riehl's *Kritizismus* (Criticism) provided the springboard for Driesch's own theoretical efforts. The first results were published in 1893 under the title *Die Biologie als selbständige Grundwissenschaft* (Biology as an independent basic science; Leipzig). This book was followed by *Analytische Theorie der organischen Entwicklung* (Analytic theory of organic development; Leipzig, 1894), which contains the first formulation of Driesch's own teleologically oriented embryological theory. But as yet this was a theory of "preformed teleology," not a vitalistic interpretation of embryological development. Only in 1895 did it dawn on Driesch that mechanistic principles could not account for his experimental findings.

Up to this time Driesch had accepted a "machine" theory of organismic development. Now he realized that such a theory would not do. In an essay titled "Die Maschinentheorie des Lebens" (The machine theory of life; in *Biologisches Zentralblatt* 16 [1896]: 353–368) he formulated as precisely as possible the view he had held so far, a view that he did not yet regard as vitalism. His first formulation of a dynamically teleological, and therefore genuinely vitalistic, theory was published under the title *Die Lokalisation morphogenetischer Vorgänge, ein Beweis vitalistischen Geschehens* (The localization of morphogenetic processes, a proof of vitalistic developments; Leipzig, 1899). In this book Driesch introduced the concept of the "harmonious equipotential system" and the proof that such a system cannot be accounted for in terms of mechanistic principles. The publication of 1899 thus marked the end of one period in Driesch's intellectual development and the beginning of another.

Gradually his interest in experimental work ceased. He now searched the literature in the field of physiology for possible proof that a "machine" theory could provide an adequate explanation of the phenomena of life. He found none, as his two books *Die organischen Regulationen* (Organic regulations; Leipzig, 1901) and *Die "Seele" als elementarer Naturfaktor* (The "soul" as elementary factor of nature; Leipzig, 1903) show. However, the conception of the "autonomy" of life had now to be justified within the broader framework of natural science. Driesch provided this justification in a book titled *Naturbegriffe und Naturурteile* (Concepts of nature and judgments of nature; Leipzig, 1904). In 1905 he published *Der Vitalismus als Geschichte und als Lehre* (*The History and Theory of Vitalism*), in which he summed up his position against a historical background. That same year he "resolved to become a philosopher." His Gifford Lectures at the University of Aberdeen in 1907–1908, published in 1908 as *The Science and Philosophy of the Organism,* provided a splendid opportunity to present his position in systematic form.

From 1908 on, Driesch was concerned exclusively with philosophical problems. In 1909 he became a *Privatdozent* at Heidelberg and in 1912 a member of the university's philosophical faculty. In 1912, also, he published his basic philosophical work, *Ordnungslehre* (Theory of order). This was followed by *Die Logik als Aufgabe* (Logic

as a task; Tübingen, 1913) and, in 1917, by *Wirklichkeitslehre* (Theory of reality). These three books together—ranging as they do over the fields of epistemology, logic, and metaphysics—embody the whole of Driesch's philosophical system, but they do not mark the end of his intellectual development. In *Leib und Seele* (Body and soul; 1916) Driesch set forth his definitive arguments against every "psycho-mechanical parallelism," and in *Wissen und Denken* (Knowing and thinking; Leipzig, 1919) he clarified and expanded his epistemological position.

In 1919 Driesch accepted a chair of systematic philosophy at the University of Cologne and in 1921 assumed a similar post at the University of Leipzig. During 1922–1923 he was a visiting professor in China. In 1926–1927 he lectured in the United States and in Buenos Aires. Being out of sympathy with the Nazi regime, ideologically and politically, he was retired in 1933. Adolf Hitler could not tolerate a thinker who fervently believed that nationalism was but "an obstacle to the realization of the *one* State of God." During the time of changing appointments, Driesch became interested more and more in problems of psychology and parapsychology. Books published in 1932 and 1938 reflect this development.

DRIESCH'S PHILOSOPHY

Although known primarily as one of the leading neovitalists, Driesch was also a critical realist and an "inductive" metaphysician. His system as a whole is developed most fully and most systematically in his *Ordnungslehre* and his *Wirklichkeitslehre.*

In his Gifford Lectures Driesch had evolved the argument that the phenomena of ontogenetic development, as revealed in his own experimental work, can be explained only when we assume the existence and the efficacy of some nonmechanistic and "whole-making" factor in nature, which Driesch called *entelechy.* This entelechy, "lacking all the characteristics of quantity," is not some special kind of energy, not a "constant" or a "force." It is not in space or in time but acts into space and into time. Entelechy, Driesch confessed, is "entelechy, an elementary factor sui generis" that "acts teleologically." But even Driesch could not blind himself to the fact that such a definition of his key concept is essentially meaningless because it is defined only negatively. He therefore tried, in his *Ordnungslehre,* to show that the conception of entelechy is logically legitimate after all.

Starting with the "irreducible and inexplicable primordial fact" that "knowing about my knowledge, I know something," Driesch found in his experience "primordial concepts of order the meaning of which I, as the experiencing subject, grasp only 'intuitively'" (*Bedeutungsschau*), and that the experience as a whole presses on toward our "seeing everything in order." The method through which this "order" is revealed is that of "positing" or "discriminating" "objects of experience." It is necessary, however, to distinguish between "positing" (*setzen*) and "implicitly positing" (*mitsetzen*). What is "posited" may, in turn, "implicitly posit" something else. The whole procedure implies that the "object" is always "my" object (since I "posit" it), not some "thing-in-itself." To postulate an "objectivity" as a reality independent of, and separated from, "my" experience would involve a fallacy. Still, we must somehow transcend this "methodological subjectivism" by attempting to obtain a complete view of the totality of experience, actual and possible. In constructing this "whole" we are to be guided by the principle of economy: Only necessary steps should be taken, for "order" is perfect only when it includes everything necessary but nothing more. Now, upon inspection, I find that the experience I have is such that I can always select some specific part of it and identify it as "this," or as A. But as soon as I have posited a "this," all the rest of my experience has become a "nonthis," and the basic principle of noncontradiction—"this is not nonthis"—emerges. Moreover, when I posit a "this" and define it as *A,* I have before me (1) the *concept A* and (2) the *judgment* "*A* is there" or "*A* exists" (at least as an object for me). But let us now assume that some particular object *A* has the discernible attributes *abcd,* whereas some other object *A*' has the attributes *acd.* The objects are clearly different, but *A* includes *A*', or "*A* implicitly posits *A*'." Thus, the posit "wolf" implicitly posits "beast of prey," and any existing wolf implicitly posits an existing beast of prey. By extension, we obtain "*A* posits *A*', and *A*' posits *a*; therefore, *A* posits *a.*" The principles of logic, thus, have their basis in our intuitive experience of order. The same is true, of course, of arithmetic and geometry. In fact, it is the aim of Driesch's general theory of order to disclose all the primordial elements of order first given in basic intuition.

Among "my" experiences there are some that I "have had before"; I "remember" them. This fact opens up an entirely new dimension of experience. But given this new dimension, I can now establish a remarkable order in my experience if I regard some of the objects of my immediate experience as an indication of the "being" or the "becoming" of an *X* that behaves as if it were independent of my experience of it; that is, it behaves as if it were a self-sufficient "realm of nature" in which the bipolar "cause-effect" relationship prevails. However, since, on

the one hand, the effect cannot be richer in content than is its cause but, on the other hand, the living individual is a "whole" that is more than the sum of its parts, a close scrutiny of experience led Driesch to distinguish between a "merely mechanical causality" (*Einzelheitskausalität*) and a "whole-making causality" (*Ganzheitskausalität*) that involves more than merely additive changes. In ontogenetic development, for example, a mere sum of "equipotentialities" is thus transformed into the "wholeness" of the mature organism. "Restitution" and "adaptations," experimentally demonstrable, are manifestations of this "whole-making" causality. The living organism itself, in its indisputable wholeness, is the most obvious result of *Ganzheitskausalität.* Thus, vitalism finds its justification within Driesch's epistemology.

At the psychological and cultural levels, "whole-making causality" predominates, and Driesch posited "my soul" as "the unconscious foundation" of my conscious experience. The "soul," therefore, is also "posited in the service of order." "My primordial knowing of the meaning of order and my primordial willing of order … indicate … a certain primordial state and dynamics of my soul." "The *working* of 'my soul' [which guides my 'actions'] and certain *states* [of my soul] are 'parallel' to 'my conscious havings.'" "This sounds very artificial," Driesch admitted, "but logic is a very artificial instrument." When Driesch took up this theme again, in his *Wirklichkeitslehre,* he argued that "metaphysically," "my soul and my entelechy are One in the sphere of the Absolute." And it is at the level of the Absolute only that we can speak of "psycho-physical interaction." But the Absolute, so understood, transcends all possibilities of our knowing, and it is "an error to take, as did G. W. F. Hegel, the sum of its traces for the Whole."

All considerations of normal mental life lead us only to the threshold of the unconscious; it is in dreamlike and certain abnormal cases of mental life that we encounter "the depths of our soul." And in parapsychological phenomena—especially in telepathy, mind reading, clairvoyance, telekinesis, and materialization (all of which Driesch accepted as proved facts)—we find traces of a supra-individual wholeness. More important, however, our sense of duty also points toward a supra-personal whole, which, in the course of history, is continuously evolving. "In my experience of duty I am participating in the supra-personal whole of which I am an empirical embodiment, and it is *as if* I had some knowledge about the final outcome of the development of that whole." That is to say, my sense of duty indicates the general direction of the supra-personal development. The ultimate goal, however, remains unknown. From this point of view, history took on its particular meaning for Driesch.

Throughout his work Driesch's orientation is intended to be essentially empirical. Any argument concerning the nature of the ultimately Real will therefore have to be hypothetical only. It starts with the affirmation of the "given" as consequent of a conjectural "ground." His guiding principle in the realm of metaphysics amounts to this: The Real that I posit must be so constituted that it implicitly posits all our experience. If we can conceive and posit such a Real, then all laws of nature, and all true principles and formulas of the sciences, will merge into it, and our experiences will all be "explained" by it. And since our experience is a mixture of wholeness (the organic and the mental realms) and nonwholeness (the material world), Reality itself must be such that I can posit a dualistic foundation of the totality of my experience. In fact, there is nothing—not even within the ultimately Real—to bridge the gap between wholeness and nonwholeness. And this means, for Driesch, that ultimately there is either God and "non-God," or a dualism within God himself. To put it differently, either the theism of the Judeo-Christian tradition or a pantheism of a God continually "making himself" and transcending his own earlier stages is ultimately reconcilable with the facts of experience. Driesch himself found it impossible to decide between these alternatives. He was sure, however, that a materialistic-mechanistic monism would not do.

See also Continental Philosophy; Critical Realism; Descartes, René; Haeckel, Ernst Heinrich; Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Locke, John; Riehl, Alois; Schopenhauer, Arthur; Vitalism.

Bibliography

ADDITIONAL WORKS BY DRIESCH

Der Vitalismus als Geschichte und als Lehre, Leipzig, 1905. Translated as *The History and Theory of Vitalism.* London, 1914. Rev. German ed., *Geschichte des Vitalismus.* Leipzig, 1922.

The Science and Philosophy of the Organism, 2 vols. London: A. and C. Black, 1908. Translated into German as *Philosophie des Organischen.* Rev. ed., Leipzig, 1921.

Zwei Vorträge zur Naturphilosophie. Leipzig, 1910.

Die Biologie als selbständige Grundwissenschaft und das System der Biologie. Leipzig, 1911.

Ordnungslehre, ein System des nichtmetaphysischen Teiles der Philosophie. Jena, 1912; rev. ed., 1923.

The Problem of Individuality. London: Macmillan, 1914.

Leib und Seele, eine Prüfung des psycho-physischen Grundproblems. Leipzig, 1916; rev. ed., 1920; 3rd ed., 1923. Translated as *Mind and Body.* New York: Dial Press, 1927.

Wirklichkeitslehre, ein metaphysischer Versuch. Leipzig, 1917; rev. ed., 1922.

Das Problem der Freiheit. Berlin, 1917; rev. ed., Darmstadt, 1920.

Das Ganze und die Summe. Leipzig, 1921. Inaugural address at the University of Leipzig.

" Mein System und sein Werdegang." In *Die Philosophie der Gegenwart in Selbstdarstellung,* Vol. I, edited by R. Schmidt. Leipzig, 1923. One of the more than 100 articles that Driesch published.

Metaphysik. Breslau, 1924.

The Possibility of Metaphysics. London, 1924.

Relativitätstheorie und Philosophie. Karlsruhe, 1924.

The Crisis in Psychology. Princeton, NJ: Princeton University Press, 1925.

Grundprobleme der Psychologie. Leipzig, 1926.

Metaphysik der Natur. Munich, 1926.

Die sittliche Tat. Leipzig, 1927.

Biologische Probleme höherer Ordnung. Leipzig, 1927; rev. ed., 1944.

Der Mensch und die Welt. Leipzig, 1928. Translated as *Man and the Universe.* London, 1929.

Ethical Principles in Theory and Practice. London, 1930.

Philosophische Forschungswege. Leipzig, 1930.

Parapsychologie. Leipzig, 1932; 2nd ed., 1943.

Philosophische Gegenwartsfragen. Leipzig, 1933.

Alltagsrätsel des Seelenlebens. Leipzig, 1938; 2nd ed., 1939.

Selbstbesinnung und Selbsterkenntnis. Leipzig, 1940.

Lebenserinnerungen; Augzeichnungen eines Forschers und Denkers in entscheidender Zeit. Edited by Ingeborg Tetaz-Driesch. Basel, 1951. Posthumous.

WORKS ON DRIESCH

Child, C. M. "Driesch's Harmonic Equipotential Systems in Form-regulations." *Biologisches Zentralblatt* 28 (1908).

Fischel, A. Review of Driesch's Gifford Lectures, *The Science and Philosophy of the Organism,* Vol. I. *Archiv für Entwicklungs-Mechanik* 26 (1908).

Griffith, O. W. Review of *The Problem of Individuality* and *The History and Theory of Vitalism. Hibbert Journal* 13.

Haake, W. "Die Formphilosophie von Hans Driesch und das Wesen des Organismus." *Biologisches Zentralblatt* 14 (1894).

Heinichen, O. *Driesch's Philosophie.* Leipzig, 1924.

Jenkinson, J. W. "Vitalism." *Hibbert Journal* (April 1911).

Jourdain, E. B. P. Review of *Ordnungslehre. Mind* 23 (1914).

Morgan, T. H. Review of *The Science and Philosophy of the Organism,* Vol. I. *Journal of Philosophy* 6 (1909).

Oakeley, H. D. "On Professor Driesch's Attempt to Combine a Philosophy of Life and a Philosophy of Knowledge." *PAS,* n.s., 21 (1920–1921).

Oakeley, H. D. Review of *Wirklichkeitslehre. Mind* 30 (1921).

Russell, L. J. Review of *Die Logik als Aufgabe. Mind* 23 (1914).

Schaxel, J. "Namen und Wesen des harmonisch-äquipotentiellen Systems." *Biologisches Zentralblatt* 36 (1916).

Schaxel, J. "Mechanismus, Vitalismus und kritische Biologie." *Biologisches Zentralblatt* 37 (1917).

Schneider, K. C. "Vitalismus." *Biologisches Zentralblatt* 25 (1905).

Secerov, Slavko. "Zur Kritik der Entelechielehre von H. Driesch." *Biologisches Zentralblatt* 31 (1911).

Spaulding, E. G. "Driesch's Theory of Vitalism." *Philosophical Review* 15 (1906).

Spaulding, E. G. Review of *The Science and Philosophy of the Organism,* Vols. I and II. *Philosophical Review* 18 (1909).

Vollenhoven, D. H. T. "Einiges über die Logik in dem Vitalismus von Driesch." *Biologisches Zentralblatt* 41 (1921).

Wagner, A. "Neo-Vitalismus," I, II. *Zeitschrift für Philosophie und philosophische Kritik,* Ergänzungsband, 136 (1909).

William H. Werkmeister (1967)

DUALISM IN THE PHILOSOPHY OF MIND

Mind-body dualism is the doctrine that human persons are not made out of ordinary matter, at least not entirely. Every person has—or, on many versions of the view, simply is identical to—a soul. A soul is said to have little in common with human bodies and other material objects but is in one way or another responsible for a person's mental life.

Mind-body dualism is sometimes called "substance dualism," to distinguish the view from "property dualism"—the thesis that mental properties (such as being in pain, thinking of Vienna) are in some way significantly different from or independent of physical properties (such as having neurons firing in one's brain in a certain pattern). Property dualism is meant to allow for what is often called "dual-aspect theory": persons are material objects with a nonphysical, mental "aspect" but no nonphysical *parts*—that is, no immaterial soul.

The entry begins with a brief discussion of property dualism, only to set it to one side in order to examine substance dualism in detail: its varieties, the traditional objections to the view, and the most popular arguments in its favor.

PROPERTY DUALISM

Before considering ways in which mental and physical properties might be distinct or independent, one needs to know what is meant by the terms *mental* and *physical.* (The expressions *property* and *state* shall be used interchangeably; being in pain is a mental property or mental state, weighing 150 pounds is a physical property or physical state. Many different things can be in pain or have the same weight; so properties and states are, in some sense, universals.)

Phenomenal states, such as experiencing a reddish afterimage or feeling a sharp pain, are surely mental states, as are "intentional attitudes" such as believing, doubting, loving, and hating. There may be puzzles about how to classify the unconscious desires and fears probed by psychoanalysts; but otherwise, the boundaries of the mental seem fairly clear. The range of things one might mean by *physical property* is, however, broader and more problematic. A narrow reading of *physical* might include only properties that come in for explicit mention in current fundamental physics—or in an imagined "final, true physics." A more generous approach would include any property expressible given just the resources of physics, mathematics, and logic. Sufficient generosity along these lines would allow for physical properties corresponding even to infinite disjunctions of arbitrarily chosen, maximally precise microphysical descriptions (that is, "consisting of such-and-such fundamental particles arranged in precisely this way, or that way, or …").

If property dualism were simply the thesis that mental properties are not identical to physical properties, narrowly construed, the doctrine would be of little interest. Synthesizing bile is a state of the liver; reaching gale force is a state of the winds in a hurricane; and neither "synthesizing bile" nor "reaching gale force" is a term likely to appear in any fundamental physics, contemporary or idealized. If "pain" fails to show up in physics for similar reasons, the mental state it names may be no less physical than the synthesis of bile or the force of a hurricane.

Given the more generous understanding of "physical," synthesizing bile or reaching gale force might well be identical to, or at least necessarily coextensive with, a physical property—a property equivalent to all the possible ways to synthesize bile or reach gale force, described in extreme microphysical detail. Imagine a god surveying all the possible worlds it could create, with their many varieties of particles and fields and laws. Such a being could disjoin all the microphysical descriptions of livers synthesizing bile or hurricanes achieving gale-force winds and thereby define physical properties necessarily coextensive with the target biological and meteorological properties. The existence of such definitions would show that the functioning of a liver or the strength of a hurricane could not possibly come apart from the behavior of the matter constituting the liver or the air and water through which the hurricane moves. If the god could do the same for mental states, that would show that they, too, are firmly grounded in microphysical facts.

To arrive at a truly interesting version of property dualism, one might suppose that even godlike powers to exhaustively describe every possible microphysical system would fail to produce a physical property necessarily coextensive with each mental property. Many who use the term follow David Chalmers (1996) in identifying it with the following sort of thesis: For at least some mental states, it is not possible to define, in terms of microphysical properties alone, a physical property common to all individuals in that mental state, and only to them—even given the resources of arbitrarily complex definitions and infinite disjunction, and even when restricting the search to a property that is merely coextensive in worlds with the same fundamental physical properties.

Property dualism, so understood, is equivalent to the failure of a variety of supervenience—a notion first used in philosophy of mind by Donald Davidson (1970) and brought into focus by Jaegwon Kim (1990). In the technical sense of *supervene* that is relevant here, the mental properties of a thing supervene upon its microphysical properties if and only if, among all the possible individuals in all the possible worlds, there is no pair with all the same microphysical properties but different mental properties. Kim showed that if supervenience held, one could define a physical property coextensive with any mental property simply by disjoining all the sufficiently precise microphysical descriptions of possible individuals having that property.

Defining property dualism as a failure of the mental to supervene upon the microphysical seems to presuppose that the fundamental properties of anything worthy of the name "physics" will not include mental states. But, as Robert Adams (1987) and Richard Swinburne (1997) point out, if mental states really are fundamental, one might expect that experiencing particular kinds of pains or smells will have to figure in some of the most basic laws. Still, so long as the nonmental physical properties of matter could be the same while the envisaged brutely mental ones could have been different (had there been different natural laws relating the two kinds of property), there would be a failure of supervenience: The mental properties would fail to supervene upon the purely physical properties.

Unlike substance dualism, property dualism remains a respectable position within philosophy of mind, defended by Chalmers (1996) and others. It seems easy to imagine physically indiscernible zombies (animate human bodies with no consciousness) or people whose spectrum of color experiences is the reverse of one's own. If genuinely possible, these scenarios show that the mental does not supervene upon the physical.

Substance dualism is also inconsistent with supervenience. If souls lack the properties mentioned in physics, they cannot very well differ physically; but, because different people are obviously thinking different things, the dualist's souls must differ mentally.

Until the latter half of the twentieth century, a dualism of mental and physical properties was largely taken for granted, even among philosophers who called themselves materialists. The term "dualism" almost always meant a dualism of distinct substances—a practice to be followed in the remainder of this entry.

PURE DUALISM AND COMPOSITE DUALISM

Many dualists, like Plato, teach that persons are entirely immaterial; they are identical with souls and are related to their physical bodies as pilot to ship. Others—perhaps René Descartes (1984), certainly St. Thomas Aquinas (cf. Stump 2003) and Richard Swinburne (1997)—identify a person with a composite of soul and body. Among composite dualists, further differences emerge: most composite dualists ascribe one's mental properties to the soul and one's physical properties to the body. On this version of composite dualism, a person is identical with a psychophysical whole that includes the thinking soul as a part. Eric Olson (2001) has drawn attention to some of the drawbacks of this view. It suggests that the soul is the real thinker, and that a person only has mental states by courtesy. But how could something—the soul—think and not be a person? How could it think for someone else? If the composite dualist insists that the person and the soul are both thinkers and that neither is the subject of mental states in a more fundamental way than the other, then each person includes two thinkers, neither of which can distinguish itself from the other.

St. Thomas Aquinas advocated a very different sort of composite dualism (for exposition, cf. Stump 2003, Leftow 2001). Within Aquinas's Aristotelian metaphysics, "accidental forms" explain a thing's accidental properties, and a "substantial form" explains its being, or essence. Following Aristotle, Aquinas calls the substantial forms of living things "souls"; the soul of a human being is responsible for its entire complex physical and mental nature. But it is not the soul that thinks or acts, it is the whole human being—a composite of matter and the soul or form that gives the matter its distinctively human structure. Aquinas departed from Aristotle in supposing that the human soul is a "subsistent form," something that continues to exist after death while not "informing" any matter. It even manages to think in that truncated state.

The Thomistic doctrine of the soul is a borderline case of mind-body dualism—although, with Eleonore Stump (2003) and Brian Leftow (2001), one may well regard its intermediate status as a promising sign. Although body and soul are united, says Aquinas, the soul has no mental properties; it is not itself a mind. Nor is it responsible for a person's mental powers alone; it includes the physical nature of a human being as well. For present purposes, dualism will be restricted to theories like Plato's pure dualism or Swinburne's composite dualism: theories positing souls with mental states of their own, in this life.

THE SPECTRUM OF DUALISMS

One point of agreement among dualists of all stripes is that there are a great many things in the world that lack mentality of any sort; and that, associated with each human person, there is a thinking thing, a soul, not composed of the same kinds of stuff as these nonmental things. The animist and spiritualist may think of the soul as extended or composite (ghostlike, perhaps composed of "ectoplasm"); but they deny, at any rate, that it is made of stuff that can be found in objects completely devoid of mentality. To be a substance dualist, then, one must at least accept a doctrine one might call *compositional dualism*: There exist things that can think alongside things that cannot think; and the thinking things either have no parts at all, or else parts of a special kind, unique to thinking things.

One could be a compositional dualist but still be a materialist. Roderick Chisholm (1978) took seriously the hypothesis that a person might be a tiny physical particle lodged somewhere in the brain. Suppose someone claimed, in a similar spirit, that the soul is a point-sized thinking substance that has the same mass as a proton and the same charge as an electron; and that every substance with a similar mass and charge is capable of thought. This rather bizarre theory qualifies as compositional dualism—yet it seems also to be a kind of materialism. Since dualism has always been thought of as an alternative to materialism, there must be more to it than compositional dualism. The missing component is clear: The thinking thing cannot simply be a special kind of physical object, such as a new species of fundamental particle; but what is it to be "nonphysical"?

Daniel Dennett sees a fundamental incoherence in the very idea of a nonphysical soul: "A ghost in the machine is of no help in our theories unless it is a ghost that can move things around … but anything that can move a physical thing is itself a physical thing (although

perhaps a strange and heretofore unstudied kind of physical thing)" (Dennett 1991, p. 35). If one were to define *physical* as "able to produce effects in space," then of course a nonphysical soul could not interact with a body. When dualists have denied that the soul is physical, they have meant many things—but none has been so foolish as to mean that.

Every plausible version of compositional dualism implies that substances capable of thought (and their parts, if any) have some important properties in common with substances utterly incapable of thought. To call a thinking thing "nonphysical" is not to say it has absolutely nothing in common with the matter of nonsentient things; it is rather to deny that they have as much in common as one might have thought. But dualists disagree about which attributes of ordinary matter are not found in thinking substances—that is, they mean different things by "nonphysical." The result is a spectrum of dualisms.

The maximal difference a dualist might posit between soul and body would be to identify souls with necessarily existing abstract objects, outside of space and time, like numbers or Plato's Forms. Some have said that persons are to their bodies as programs are to the computers that run the programs. And, if programs are understood in a way that makes them quite independent of the particular computers running them, they become abstract objects, mathematical entities. But it is hard to take this analogy very seriously. Almost all dualists will agree that souls have this much in common with ordinary material things: They are concrete entities, existing in time, and capable of change.

René Descartes allowed at least that much similarity between souls and ordinary matter, but little more. Cartesian souls are not dependent upon the behavior of matter for their continued existence or ability to think. They have no position in space. Descartes also claimed that souls are "simple," or without parts. Since he believed that everything in space was infinitely divisible, this was another way in which souls were unlike anything made of ordinary matter (Descartes, 1984).

Few dualists are so far out along the spectrum of dualisms as Descartes, however. It has become harder to deny that the ability to think depends upon a properly functioning brain. William Hasker (1999), Charles Taliaferro (1994), and other contemporary dualists go further, denying the existential independence of souls: When an organism has a sufficiently complex nervous system, it then automatically also generates a nonphysical substance to be the subject of that consciousness—an "emergent substance" that remains radically but not completely dependent upon the brain for most of its operations and even for its continued existence. Hasker, W. D. Hart (1988), and—long before them—Samuel Clarke (1738) and Hermann Lotze (1885) have insisted that souls are located in space. Hart argues that mind-body interaction could even involve the transfer of a conserved quantity between soul and body. The "psychic energy" he describes makes souls even more like paradigmatic physical things. Still, Hart's souls lack charge, mass, spin, and all other interesting intrinsic properties characterizing physical particles. Furthermore, Hart defines measurable degrees of psychic energy in terms of the propensity to sustain beliefs, not in terms of physical effects; so even this quasi-physical quantity seems grounded in the mental nature of Hart's souls rather than in any features they share with ordinary matter.

Hart's view should surely qualify as a kind of dualism—his souls are immaterial enough—and the Chisholm-inspired particle materialism should not. If, as seems likely, there is no sharp line on the spectrum of compositional dualisms between the two, then the term "dualism" is vague. As with most vague yet useful terms, the region of indeterminacy is largely unoccupied.

The less extreme dualisms are of greater philosophical interest than Cartesianism. They make souls a part of the natural order, generated by any brain sufficiently complex to subserve conscious experience. One of the worst problems of interaction (the "pairing problem," discussed in the next section) is easily solved if souls are in space. Furthermore, few, if any, of the principal arguments for dualism (including the ones surveyed below) require Cartesian souls. Less radical dualisms are safer, positing no more differences between souls and material objects than are implied by the reasons for rejecting materialism.

PROBLEMS OF INTERACTION

Most objections to dualism fall under one of three heads: problems of interaction, epistemological worries, and application of Ockham's Razor. The most commonly cited "knockdown" objection to dualism is the impossibility of causal interaction between things as dissimilar as a physical body and an immaterial soul. The obvious rejoinder is that very dissimilar things do interact. For example, particles are certainly quite unlike the fields that push them around and that are, in turn, altered when particles are introduced into them. Attempts to make the objection more persuasive come in two versions.

The "pairing objection" begins with Ernest Sosa's observation: "What pairs physical objects as proper mates for causal interaction is in general their places in the all-encompassing spatial framework of physical reality" (1984, p. 275). Consider a series of duplicate guns, each of which hits a different target. Guns and targets are exactly alike; only differences in spatial relations explain why each gun hits a different target—the target at which it is aimed. Compare guns and targets to the bodies and souls of identical twins Joe and Moe. However similar they are, only Joe's body causes experiences in Joe's soul; only decisions taken by Joe lead directly to motions of Joe's body. According to the Cartesian, there can be no differences in the spatial relations between Joe's soul and the bodies of Joe and Moe; being outside of space, the soul cannot be closer to one body than to the other. But in what other respects could Joe's soul be "closer" to Joe's body than to Moe's body, and Moe's soul closer to Moe's body than to Joe's? Descartes's souls are all equally cut off from the physical world, so no answer comes readily to mind.

The pairing objection tacitly assumes that causal laws, and the dispositions and powers of objects described by such laws, are always general—an assumption some dualists reject. John Foster (1991) and Peter Unger (2006) think that souls and bodies could have not only dispositions to react to certain types of objects and situations but also dispositions to interact in special ways with particular individuals—individuals that need not differ in any qualitative or relational way.

Dualists like Clarke (1738), Lotze (1885), Hart (1988), and Hasker (1999) are in an even stronger position, since they assume that souls fall within the same spatial coordinate system as bodies. They make the natural assumption that, if souls are to be found in space at all, they must be located within the brains with which they interact. But one still wants to know exactly what sort of region a soul is supposed to occupy. Many dualists believe souls are simple, or partless. Must a simple thing occupy a geometrical point, on pain of being divisible into at least two parts, a left and right half? Some philosophers say no. Clarke (1738) and Lotze (1885) claim that the soul is spatially extended but simple. Lotze locates the soul within the brain wherever interaction takes place—which could be many different places at once, and different places at different times. Leibniz considers a mode of spatial occupancy the Scholastics called "definitive ubeity": there is a precise region in which the soul is located, but it is not true of any subregions that it is located precisely there (Leibniz, 1981, p. 221). Although these are difficult notions, they may represent ways (or perhaps two descriptions of the same way) for a soul to occupy more than a mere point while remaining a partless unity.

A second objection to interaction alleges that the mental states attributed to souls are of the wrong sort to enter into laws governing physical phenomena. If the "qualia" of phenomenal experiences (for example, the felt redishness of a red after-image, the sharp flavor of an acrid smell) could somehow be reduced to physical states of brains or analyzed in terms of functional roles that physical states could play, then they would pose little threat to a materialistic picture within which all causation is underwritten by laws of the sort one finds in physics. If they characterize the states of a nonphysical soul, however, they will have to be taken seriously as extra, fundamental features of the world, requiring causal explanation. Causation requires laws; but in order for the astonishing variety of phenomenal states, falling under several sense modalities, to enter into the kinds of laws familiar from the sciences, they must be susceptible of precise mathematical comparison. However, as Robert Adams points out, "[t]here is no plausible, non–adhoc way of associating phenomenal qualia in general … with a range of mathematical values.…" (Adams 1987, p. 256). Laws linking the phenomenal experiences of a soul to the physical states of a body are bound to be relatively unsystematic and staggeringly complex. Far better to suppose that phenomenal properties are merely complex physical states of the brain; and that, as such, they obey laws that can be derived from those of biology, chemistry, and, ultimately, fundamental physics.

This second interaction objection, however powerful it might be, applies not only to substance dualists but also to anyone who is a property dualist about phenomenal states. Many philosophers who are happy to suppose that persons are identical with physical objects (such as living, human bodies or brains) nevertheless heartily endorse property dualism with respect to the qualia of phenomenal states. Like substance dualists, these property dualists must admit that there are additional laws governing the production of phenomenal qualia—laws that are quite complicated and, to some extent, piecemeal. (David Chalmers, Gregg Rosenberg, and others have floated theories about the form such laws might take [Chalmers, 1996; Rosenberg, 2004.])

Property dualism remains a respectable position within contemporary philosophy of mind, with powerful arguments in its favor. In the circumstances, then, this second problem of interaction can hardly be the final nail in the coffin of substance dualism.

EPISTEMOLOGICAL WORRIES

After interaction objections, the most commonly voiced complaints about substance dualism are epistemological in flavor: Suppose persons are souls that merely happen to be associated with bodies. One cannot keep track of another's soul by keeping an eye on it, or holding it fast. How, then, does one know that souls are not constantly coming and going "behind the scenes"?

Immanuel Kant's analogy illustrates the problem: "An elastic ball which impinges upon another similar ball in a straight line communicates to the latter its whole motion, and therefore its whole state (that is, if we take account only of the positions in space)." A series of mental substances passing on "representations together with the consciousness of them" would end with one that is "conscious of all the states of the previously changed substance, as being its own states, because they would have been transferred to it together with the consciousness of them." But if we identify persons with individual mental substances, "it would not have been one and the same person in all these states" (1965 p. 342). Kant's scenario is often turned into an argument against dualism: If it were reasonable to suppose that each person is identical with a soul, then it would be reasonable to be skeptical about whether we are dealing with the same person from one minute to the next. Since this is not reasonable, neither is the supposition that a person is a soul.

The argument fails if one endorses John Locke's view (in the chapter "Of Identity and Diversity" in his *Enquiry* [1975]) that a person is not identical with a particular soul but is instead constituted by a soul, and possibly by different souls at different times. So long as the succession of souls pass on the right sorts of mental states (Locke emphasizes memories), the person survives, constituted by one soul and then another. To give this reply would require that one say, with Locke, that a person and the person's soul are distinct things, although the soul thinks whenever the person does. In that case, if a person always remains responsible for the things she has done, then one soul could justly be punished for the deeds of another soul. (Locke himself seems to have thought that, although such punishment would not be unjust, it would not be very nice, and so God can be counted on not to allow soul-switching.)

Locke's approach is surely not the only way to dispel the Kant-inspired epistemological worry. Another is simply: *to quo que*. If our knowledge of the persistence of physical objects—including human bodies—is just as vulnerable to similar skeptical doubts, then materialism has no advantage over dualism. But what sort of evidence supports the belief that a physical object observed at one time is the same as an object observed at another time—and not, say, an exact duplicate that has swapped places with the original due to random quantum-mechanical fluctuations or the whimsy of a powerful demon? Just as one can imagine one soul being replaced by a near duplicate without anyone's being the wiser, so one can imagine a physical object being replaced by a near duplicate with no readily detectable evidence that a switch was made. Does the ability to imagine such things require that one produce nonquestion-begging arguments against them if one is ever to claim knowledge of identity over time? Surely not. Is there some special problem with souls? If so, it needs more spelling out than it usually receives.

OCKHAM'S RAZOR

Some of the most frequently voiced objections to dualism—the ones based on problems of interaction and epistemological worries—may become less impressive upon examination. At least one formidable objection remains, however: that there is simply no need to believe in souls in addition to bodies; so the soul falls victim to Ockham's razor, the injunction to postulate no more entities than necessary. One has the evidence of one's own senses for a world of physical bodies. But even if property dualists are right and some psychological phenomena cannot be reduced to or exhaustively explained in terms of properties similar to those now ascribed to physical bodies and their parts, nothing would be gained by supposing that these irreducible mental properties belong to some new entity. And adding the extra entities requires many further ad hoc epicycles that undermine any explanatory value their addition might have had. For instance, one must now explain why the exercise of the soul's mental powers depends so heavily upon a properly functioning brain. Perhaps hard evidence of spirit possession, reincarnation, veridical out-of-body experiences, and the like would change the situation. But, in its absence, respect for parsimony in theory construction provides a powerful reason to reject souls.

MODAL ARGUMENTS

The two most famous styles of argument for dualism may be found, unsurprisingly, in Descartes. One is a modal argument (that is, an argument built around what is possible or necessary) from the possibility of disembodiment to the conclusion that every person actually has, or is, a soul. The other is an argument from the "unity of consciousness" to the conclusion that the subject of consciousness is a partless (and so, by Descartes's lights,

nonphysical) substance. Each sort of argument has been subjected to withering criticism, however; and, despite repeated attempts to revive them, the prognosis is not good.

Some of a thing's properties appear clearly to be contingent, while others seem essential. It is possible to lose a contingent property, but not an essential one—it characterizes the thing necessarily. It is possible for me to survive the loss of my leg; so having two legs is one of my contingent properties. If it were possible for me to survive the destruction of my entire body, without acquiring new bodily parts, I would be contingently embodied. If it were not possible, then having a body would be part of my essence.

Descartes develops a modal argument in his sixth meditation: "[T]he fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct. … Thus, simply by knowing that I exist and seeing at the same time that absolutely nothing else belongs to my nature or essence except that I am a thinking thing, I can infer correctly that my essence consists solely in the fact that I am a thinking thing" (1984, p. 54).

Swinburne (1997) defends a roughly similar argument. He points out that it is easy to imagine scenarios in which one survives the utter destruction of all the material parts of one's body at once, or the swapping of one body for another. There is nothing straightforwardly inconsistent in such stories, and Swinburne takes this to be strong evidence that the stories represent genuine possibilities. He also assumes, not unreasonably, that no mere material object could survive such adventures. On these assumptions, one should reason as follows: "I could survive the destruction, all at once, of all the matter in my body; my body could not survive this; so I am not identical with my body."

In the absence of a reduction of possibility to logical consistency, it is unclear where evidence for possibility could come from if not from the seeming coherence of various imagined states of affairs. So it is not unreasonable to grant that, if one can conceive of being unextended or of surviving the destruction of one's body, then this fact provides at least prima facie evidence for the possibility of these things. But prima facie evidence may be undermined, and in the arguments of Descartes and Swinburne, it is counterbalanced by the conceivability of states of affairs that are inconsistent with the possibility of the separation of person and body. Many find that they are able to imagine themselves as having nothing but extended or material parts just as easily and clearly as they can imagine persisting without parts or without a body. One can conceive of oneself as a mere organism, a brain, or even a rock. But if such things cannot possibly be unextended, or continue to exist after annihilation of their physical parts—an assumption required by the modal arguments for dualism—then one has prima facie evidence for the possibility of being identical with a thing that could not possibly survive in an unextended or disembodied state. But if some envisaged situation is possibly not possible, then it is simply not possible. So it is simply not possible that I be unextended or disembodied.

The plausibility of this widely accepted principle of modal reasoning (that what is possibly not possible is not really possible at all) may be more apparent when stated in the jargon of "possible worlds": If there is a world that is possible from our perspective (that is, from the point of view of the actual world, this other world represents a way things could have been); and if, from the perspective of that other world, some imagined state of affairs or circumstance is not possible; then that imagined state of affairs is not possible from the point of view of the actual world either—that is, it is simply not possible. Applied to the case in hand, this modal principle becomes: If, according to some possible world, I do not exist without a body in any possible world, then this remains true in the actual world—I do not exist without a body in any possible world.

If I find it just as conceivable to suppose that I am entirely physical as to suppose that I become disembodied, then I have the same sort of evidence for the possibility of each supposition. But they cannot both be possible. So the evidence from conceivability cuts both ways and cancels itself out.

There is more to be said on behalf of modal arguments for dualism, of course. Perhaps the way in which one can conceive of one's disembodiment is qualitatively better—more luminous or complete—than the way in which one can conceive of one's being a mere brain or organism. And perhaps the higher quality of the act of conception brings with it an "epistemic boost" for the possibility of the scenario thus conceived. But making a case for such a difference would require wading far into the murky waters of modal epistemology.

ARGUMENTS FROM THE UNITY OF CONSCIOUSNESS

Many dualists (such as Joseph Butler [1736], Samuel Clarke [1738], Lotze [1894], and, Hasker [1999]) would agree with Descartes about the importance of what came to be called "the unity of consciousness": an argument

based on the unity of consciousness alone is "enough to show me that the mind is completely different from the body, even if I did not already know as much from other considerations" (1984, p. 59).

The unity of consciousness may be illustrated by a person who sees a book fall, hears the sound of its impact, and feels a pain in her right toe where it struck. She can immediately infer that there is something that sees the fall, hears the impact, and feels a pain. The facts of experience do not simply imply the occurrence of three events, a "seeing of a book's fall," a "hearing of an impact," and a "feeling of a pain." Events of these types could occur to three different thinking things, no one of which is able to compare the sound with the sight and the pain. What must be added to capture the additional information is that the three events all occur to one and the same individual.

Thus the unity of consciousness supports the view that whatever is the bearer of psychological properties must be a single substance capable of exemplifying a plurality of properties. Its unitary nature consists in the impossibility of its having a "division of psychological labor" among parts. If a single thinker can recognize the difference between sounds and colors, this thinker does not enjoy the ability to compare the two simply by having one part that does its seeing and another that does its hearing, even if these parts are tightly bound together. As Franz Brentano remarks, this "would be like saying that, of course, neither a blind man nor a deaf man could compare colors with sounds, but if one sees and the other hears, the two together can recognize the relationship" (1995 p. 159).

Many dualists have claimed that the unity of consciousness requires that whatever is conscious must be a unity having no parts at all. Although Brentano believed the soul to be simple, he did not think the simplicity of the soul follows immediately from the unity of consciousness alone, and he was surely right. As Brentano points out, what is not ruled out as a subject of consciousness is an extended substance that exemplifies all of its psychological properties as a whole (1987). To use Brentano's metaphor, the psychological properties could be "spread equally" over all of the parts of this extended thinking thing. None of the many arguments that have been given to rule out this possibility has met with widespread acceptance, even among dualists.

ARGUMENTS FROM THE VAGUENESS OF MATERIAL OBJECTS

Arguments for dualism often take the form of objections to any normal sort of materialism. A materialism that identified a person with a single cell or proton would be at least as incredible as dualism (absent some sort of revolution in neurophysiology). What materialists want is a view according to which a human person may be identified with a reasonably normal physical object, one that already has a place in our commonsense conception of the world—an object with natural boundaries, such as those of an organism, a brain, or perhaps even a single hemisphere of a brain. But animals and their organs belong on a spectrum that includes bushes, branches, clouds, mountains, rivers, tidal waves, and all manner of ill-behaved entities. Familiar material objects such as these exhibit vagueness or indeterminacy in their spatial and temporal boundaries. And the strategies typically implemented to resolve puzzles posed by vague objects do not seem so satisfactory when applied to oneself.

Human bodies and brains appear surprisingly like clouds upon close inspection—blurry around the edges. Many particles are in the process of being assimilated or cast off; they are neither clearly "in" nor clearly "out." The temporal boundaries of living things—their coming into existence and passing away—also display a disturbing fuzziness. No one doubts that meteorologists have considerable freedom in deciding where exactly to draw the line between a hurricane and a mere tropical storm. But organisms and brains are not unlike storms in this respect; pressure to find the first and final moments in the life of a human body or brain can only force a decision like the one made by the meteorologists.

Sharper lines will not be found by those who, with Locke, dismiss biological boundaries for persons in favor of psychological ones. Neo-Lockeans must admit that psychological continuity, like biological life, is a matter of more and less; that personalities emerge, and frequently deteriorate, only gradually.

The materialist must, therefore, allow that the spatial and temporal indeterminacies of large-scale material objects infect human persons; and that the standard strategies for coping with fuzzy objects apply to persons as well. But application of these strategies to oneself can produce a disturbing sense of vertigo. The feeling is especially intense in the temporal case.

One group of botanists could establish the convention that no acorn is an oak tree, and another that oak trees are grown-up acorns; one meteorological society

could lay it down that hurricanes only begin when a tropical storm attains wind-speeds exceeding 74 miles per hour, another could choose 73. Similarly, one linguistic community could insist that persons exist at conception (twinning, they might say, is the generation of two "new" persons and the end of the first); while another community might talk as though persons come into existence as soon as twinning is impossible or differentiation of organs begins or rudimentary psychological states are detectable or the first breath is taken. Similar ranges of options lie open at the other end of life. If human persons are as much like trees and hurricanes as human bodies appear to be, such differences in usage would affect the extension of "person" and, with it, the reference of "I" in the mouths of speakers from different communities. The physical facts leave room for more than one perfectly acceptable refinement of the concept "tree" or "hurricane"; if human persons are entirely physical, the same must be true of human person.

If these refinements in the extension of "person" are to be genuine possibilities, there must already exist different physical objects corresponding to the different decisions that could be made about origins and deaths; and each of these preexisting objects must have what it takes, intrinsically, to be a conscious person. Speaking and thinking differently cannot make new physical objects spring into existence, nor can it turn objects with no phenomenal states into objects with the rich phenomenology of a human person. But then there must already be quite a few humanlike creatures located wherever a human person is located, each exactly like a person in every intrinsic respect. Although some philosophers (notably, the friends of temporal parts) have learned to live with this result, it raises dizzying possibilities. If the extension of a term like *person* is determined by present and past usage and the rule for determining the referent of *I* is something like "it refers to the person speaking," then a shift from one of the acceptable refinements of "human person" to another could render a conscious, self-referring creature no longer able to think for itself. If, instead, *I* is not tied to the actual meaning of "person" but rather refers ambiguously to each of the humanlike creatures associated with a given person, then there are many thinkers with slightly different pasts and futures, and none can tell which one he or she is (a result emphasized in Olson 1997).

The possibility of fission and fusion is a further source of indeterminacy and conventionality in spatiotemporal boundaries, one that Chisholm (1976) and Swinburne (1997) have exploited in arguments for dualism. When half of a bush is destroyed, one is tempted to say it survives; when it is merely split in two, and the halves successfully transplanted, one is tempted to say one of two things: either that there are two new bushes or that the bush survives as a scattered object, part in one place, part in another. If persons are thought to be middle-sized material objects with biological or psychological persistence conditions, similar circumstances of fission and fusion are conceivable and perhaps even physically possible. (Because a great deal of basic psychological continuity is preserved through the loss of either hemisphere, fission is probably a physical possibility on neo-Lockean accounts of personal identity.) If one takes the first approach to bushes, regarding fission as the end of the original plant, one should say the same thing about a purely physical human being.

There has been little need for precision about the fate of a divided bush. But a community of language users that felt the need could surely introduce a term for things exactly like bushes while decreeing that no such thing can survive loss of half its mass at once; another community could choose 49 percent; but neither group need fear making a mistake. Comparable freedom with respect to persons would require one to say things like, "If my linguistic community were to change its mind, either this would alter my persistence conditions—a strange power to change the nature of a physical object by talking differently—or else it would shift the referent of *I* in my mouth, rendering me no longer able to refer to myself in the first person." Neither alternative is attractive. The analogue to treating the divided bush as a scattered object would be to say that a person could be in two places at once, undergoing radically different experiences, thinking incompatible thoughts, and so on.

TENDER-MINDEDNESS AND ONTIC IGNORANCE

It is hard to apply to oneself the same strategies one would unhesitatingly use to deal with indeterminacy in the identity conditions and borders of ordinary physical objects. Chisholm and Swinburne take this discomfort as evidence that human beings are not ordinary physical objects. Stipulations about whether a person survives a certain borderline adventure are bootless if the person is in fact an immaterial substance whose identity over time is an all-or-nothing affair.

Resisting materialism because it is hard to accept that human beings are as fuzzy and conventional as ordinary physical objects will no doubt strike many philosophers as mere tender-mindedness. After all, they will insist, it

should be possible for philosophy to reveal something new about persons; and surely it is more certain that human beings are material objects than that they have perfectly adequate self-conceptions. (Derek Parfit [1984] takes this approach, emphasizing the radical morals to be drawn from the vagueness of human persons.)

On the other hand, it would be high-handed to dismiss as tender-minded anyone who allows the argument from vagueness to count against materialism. If the consequences of supposing that persons are vague material objects seem incredible, this might quite properly increase the weight that can be given to other considerations in favor of dualism: arguments from theological premises, for example, or more esoteric philosophical arguments (such as those of Peter Unger, J. R. Smythies, or John Foster) that would carry greater conviction if materialism were not thought to be utterly obvious and unproblematic. All by themselves, however, the foregoing arguments from vagueness ought probably be taken to support nothing stronger than (what George Graham [1999] calls) "ontic ignorance": "I know not what manner of thing I am."

See also Mind-Body Problem; Physicalism.

Bibliography

Adams, Robert Merrihew. *The Virtue of Faith*. New York: Oxford University Press, 1987. See esp. the chapter "Flavors, Colors, and God."

Brentano, Franz. *On the Existence of God*. Dordrecht: Martinus Nijhoff, 1987 (first German edition, 1929).

Brentano, Franz. *Psychology from an Empirical Standpoint*. London and New York: Routledge, 1995. First German publication of this edition, 1924.

Butler, Joseph. *The Analogy of Religion*. London: James, John, and Paul Knapton, 1736.

Chalmers, David. *The Conscious Mind*. New York: Oxford University Press, 1996.

Chisholm, Roderick. "Is There a Mind-Body Problem?" *Philosophical Exchange* 2 (1978): 25–34.

Chisholm, Roderick. *Person and Object*. La Salle, IL: Open Court, 1976.

Clarke, Samuel. "Four Defences of a Letter to Mr. Dodwell." In *The Works of Samuel Clarke*. London: John and Paul Knapton, 1738.

Corcoran, Kevin, ed. *Soul, Body and Survival*. Ithaca, NY: Cornell University Press, 2001.

Davidson, Donald. "Mental Events." In *Experience and Theory*, edited by L. Foster and J. Swanson. Amherst: University of Massachusetts Press, 1970.

Dennett, Daniel. *Consciousness Explained*. Boston: Little, Brown, 1991.

Descartes, René. *The Philosophical Writings of Descartes*. Vol. 2. Cambridge, U.K.: Cambridge University Press, 1984.

Foster, John. *The Immaterial Self*. London and New York: Routledge, 1991.

Graham, George. "Self-Consciousness, Psychopathology, and Realism about the Self." *Anthropology and Philosophy* 3 (1999): 533–539.

Hart, W. D. *The Engines of the Soul*. Cambridge, U.K.: Cambridge University Press, 1988.

Hasker, William. *The Emergent Self*. Ithaca, NY: Cornell University Press, 1999.

Kant, Immanuel. *Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: St. Martin's Press, 1965.

Kim, Jaegwon. "Supervenience as a Philosophical Concept." *Metaphilosophy* 21 (1990): 1–27.

Leftow, Brian. "Souls Dipped in Dust." In *Soul, Body and Survival*, edited by Kevin Corcoran. Ithaca, NY: Cornell University Press, 2001.

Leibniz, G. W. *New Essays on Human Understanding*. Cambridge, U.K.: Cambridge University Press, 1981.

Locke, John. *An Essay concerning Human Understanding*. Oxford: At the Clarendon Press, 1975.

Lotze, Hermann. *Microcosmus: An Essay concerning Man and his Relation to the World*. Vol. 1. New York: Charles Scribner's Sons, 1894.

Lotze, Hermann. *Outlines of Psychology*. Minneapolis: S. M. Williams, 1885.

Olson, Eric. "A Compound of Two Substances." In *Soul, Body and Survival*, edited by Kevin Corcoran. Ithaca, NY: Cornell University Press, 2001.

Olson, Eric. *The Human Animal*. New York: Oxford University Press, 1997.

Parfit, Derek. *Reasons and Persons*. Oxford: At the Clarendon Press, 1984.

Popper, Karl, and John C. Eccles. *The Self and Its Brain*. Berlin: Springer-Verlag, 1977.

Robinson, Howard. "Dualism." Available online from the Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/.

Rosenberg, Gregg. *A Place for Consciousness*. New York: Oxford University Press, 2004.

Smythies, John. *The Walls of Plato's Cave*. Aldershot: Avebury Press, 1994.

Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9 (1984): 271–281.

Stump, Eleonore. *Aquinas*. London and New York: Routledge, 2003.

Swinburne, Richard. *The Evolution of the Soul*. Rev. ed. Oxford: Clarendon, 1997.

Taliaferro, Charles. *Consciousness and the Mind of God*. Cambridge, U.K.: Cambridge University Press, 1994.

Unger, Peter. *All the Power in the World*. New York: Oxford University Press, 2006.

Dean Zimmerman (2005)

DUBOS, ABBE JEAN BAPTISTE
(1670–1742)

With his *Réflexions critiques sur la poésie et sur la peinture*, Abbe Jean Baptiste DuBos—diplomat, man of letters, member of the *Académie française*—had an essential influence on the aesthetic thought of the Enlightenment.

Réflexions critiques, published for the first time in 1719 and re-edited several times, is one of the founding texts of the new "aesthetics" that came into its own in the eighteenth century. DuBos defends a sense-based theory of aesthetic feeling that is set in motion by poetry, painting, and music. In his *Réflexions*, DuBos's successors saw, on the one hand, an aesthetic that stressed the effects of artworks on spectators and that favored the highly emotional or moving dimension of the aesthetic response to art, and, on the other hand, an attempt to base aesthetic judgment on nonrational bases—what DuBos called the "sixth sense" feeling, or the "heart." His strictly emotionalist interpretation of the "paradox" of negative feelings—which has it that the more we are afflicted by the artistic representation the more pleasure we derive from it—captured the attention of all eighteenth-century theorists of tragic emotion.

The method of the *Réflexions* aims for the "experimental," that is, it is founded on the observation of psychological, social, environmental (*climatiques*), and historical causes. In this respect, his empiricism is tinged with eclecticism, while with regard to the analysis of the mind and emotions, he belongs more strictly to the philosophical vein stemming from John Locke. In view of its subject, this work can be interpreted in somewhat anachronistic terms as a *metacriticism*; as a philosophical endeavor aimed at revealing the general principles of literary and artistic criticism. A certain theoretical distance from DuBos's aesthetic thought (the coherence of which is not always obvious at first glance) helps us to see that it tries to articulate three specific issues: the analysis of the emotional response to an artwork, the theory of aesthetic judgment, and the causes of the historical variations of genius.

1) The only aim of poetry and painting is to please and to arouse feeling by the imitation of subjects that are themselves moving. Art fills a specific need: that of the mind to be kept occupied in order to avoid tedium. The "artificial" passions art stirs have thus the emotional power of ordinary passions, without having their grievous consequences. DuBos proposes that the pleasure we derive from passions (even negative ones) comes uniquely from the emotional energy and intensity inherent in them, not from the reflective consciousness that we ourselves are not at risk or from the mere enjoyment of artistic imitation. After analyzing the nature and causes of viewers' aesthetic pleasures, DuBos explores the various means of producing these pleasures by examining the powers of artistic imitation and by comparing the relative force of different artistic forms and, within these forms, the different artistic genres. Thus, tragedy is superior to comedy, for example, just as painting (which uses "natural signs") touches us more directly than poetry (which uses the "artificial signs" of language). However, at the end of the day, absolute aesthetic primacy goes to staged tragedy, which articulates a succession of "paintings" or scenes in time and takes gradual control of our emotions.

2) Aesthetic sentiment also possesses an *evaluative* dimension; it functions as a principle of *judgment* concerning artistic and literary works; DuBos demonstrates, against the pretensions of a normative and professional criticism, that only the *sentiments* of the *public*, which become more and more assured as time goes by, can reliably decide the real merit of artworks.

3) All this analysis of emotional and evaluative modalities of the artistic experience are part of what could be called a "scientific" criticism that aims to reflect on the diverse historical "causes" (both physical and moral) that explain the *variations* in the production and reception of artistic and literary works; DuBos develops a theory of genius, the manifestations of which are essentially submitted to so-called "climatic" (including physical and environmental) conditions, while simultaneously founding a vein of historical criticism supported by a cyclical conception of history. Together these three elements sketch out an aesthetic theory that is clearly antirationalist, for which neither individual aesthetic responses, nor the evaluation and acknowledgment of a work's merit, nor the mechanisms of artistic and literary production, are subjected to the constraints of rules and normative prescriptions. DuBos thus holds an original place in the *Querelle des Anciens et des Modernes*: he refutes the rationalist pretensions of the "*Modernes*" while shifting the debate to the analysis of the feelings.

See also Aesthetic Experience; Aesthetic Judgment; Locke, John.

Bibliography

WORKS BY ABBE JEAN BAPTISTE DUBOS

Réflexions critiques sur la poésie et sur la peinture (1719). Paris: École Nationale Supérieure des Beaux-Arts, 1993.

WORKS ABOUT ABBE JEAN BAPTISTE DUBOS

Becq, Annie. *Genèse de l'esthétique française: de la raison classique à l'imagination créatrice, 1680–1814*. 2nd ed. Paris: Albin Michel, 1994.

Dumouchel, Daniel. "Le problème de Dubos et l'affect compatissant: l'esthétique du 18e siècle à l'épreuve du paradoxe tragique." In *De la sympathie sous l'Ancien Régime: discours, savoirs, sociétés*, edited by T. Belleguic and E. Van den Schueren. Québec: Presses de l'Université Laval, 2005.

Lombard, Alfred. *L'Abbé Du Bos, un initiateur de la pensée moderne*. Paris, 1913.

Saisselin, Remy G. "Ut Pictura Poesis: Du Bos to Diderot." *Journal of Aesthetics and Art Criticism* (20) (2) 1960.

Daniel Dumouchel (2005)

DUCASSE, CURT JOHN

(1881–1969)

Curt John Ducasse, philosopher and educator, was born in Angoulême, France. After attending schools in France and England, he came to the United States in 1900. He received his B.A. and M.A. degrees from the University of Washington and, in 1912, his Ph.D. from Harvard University, where he had served as an assistant to Josiah Royce. He taught philosophy at the University of Washington from 1912 until 1926, at Brown University from 1926 until his retirement in 1958, and elsewhere as visiting professor. He served as president of the Association for Symbolic Logic (1936–1938), which he had helped to found, and of other learned societies. He published extensively in all fields of philosophy.

PHILOSOPHICAL METHOD

Ducasse's views on method are worked out in detail in *Philosophy as a Science: Its Matter and Method* (New York, 1941), in his Carus lectures, published as *Nature, Mind, and Death* (La Salle, IL, 1951), and elsewhere.

He held that philosophy is a science and that it differs from other sciences not in the generic features of its method but by virtue of its subject matter, which consists of "spontaneous particular appraisals" (1941) or "standard evaluative statements" (1951) made by some person or group. The primitive problems of philosophy are to define the value predicates "good," "valid," "real," and so on, and their opposites, as used by the person or persons whose standard evaluative statements are taken as data. In the definitions will appear such terms as *necessary, fact*, and *possibility*, which are also in need of analysis, giving rise to derivative problems. Both sorts of problems are essentially semantical. Ducasse is thus squarely in the analytical tradition. However, he argued more explicitly than other contemporary analysts that a proposed analysis of a term as used in paradigm statements has the status of a hypothesis, and that it can be confirmed or disconfirmed by observing whether it is substitutable for the analysans in the paradigm statements without altering any of their standard implications.

CAUSATION

Ducasse had adumbrated the above views and had applied his method to the concept of causality in *Causation and the Types of Necessity* (Seattle, 1924). Ducasse had always regarded causality as a "fundamental category," and in subsequent works he continued to refine his original analysis.

According to Ducasse, causality is a relation between events, is essentially triadic, and is correctly defined in terms of J. S. Mill's method of difference. That "method" is not in fact a method for discovering causal connections but a description of the causal relation itself. If, in a state of affairs *S*, only two changes occur, one the change *C* at time T_1 and the other the change *E* at time T_2, *C* is the cause of *E*. Ducasse asserted that despite David Hume's definition of causation as regularity of sequence, Hume actually thought of it in terms of the advent of a single difference in a given state of affairs, as is proved by the way he formulated his rules for ascertaining causal connections by a *single* experiment.

Given the above definition, the supposition that some events have no cause implies a contradiction. Hence, indeterminism, the view that some events are matters of objective change, is self-contradictory, although people are "free" in the sense that, and to the extent that, they can do what they will to do.

MIND AND NATURE

In *Nature, Mind, and Death*, Ducasse went on to assert that nature is the material world, comprising all the things, events, and relations which are publicly perceptible. The mental, which is directly observable only through introspection, is not part of nature. Substances are analyzed as systems of properties and their relations. A property is a causal capacity. Thus,

> to say of carborundum that it is *abrasive* means that, under certain conditions, friction of it against certain other solids causes them to wear away. … More generally, to say that a substance *S* has a property or capacity *P* means that *S* is such that, in circumstances of kind *K*, an event of kind *C*, occurring in *S* or about *S*, regularly

> causes an event of kind *E* to occur in or about *S*. (*Nature, Mind, and Death*, p. 165)

Since *C* and *E* may stand for either a physical or a mental event, there are four kinds of properties: physicophysical, if *C* and *E* are both physical events; physicopsychical, if *C* is physical and *E* psychical; psychophysical; and psychopsychical.

The relation of a mind, a mental substance, to "its body," a material substance, is that of causal interaction. This is an analytic truth, for by "its body" can only be meant "the body with which that mind directly interacts." Many of the usual objections to interactionism presuppose a mistaken conception of causality.

In the case of physicopsychical properties ("bitter," "blue") it is important to distinguish between the sense quality in terms of which the property is defined and the property itself. "Bitter," for example, is equivocal. As applied to quinine, it is a disposition term designating the capacity of quinine to cause a certain taste experience when one places it on one's tongue. As applied to the experience itself, it is the name of a quality. With respect to the properties of material things, Ducasse is a realist. Quinine is bitter and roses are red, in the dispositional sense, even if the properties are not being exercised. Of properties, it is false that *esse* is *percipi*. But in the case of sense qualities, it is true that *esse* is *percipi*.

Now G. E. Moore, in his "Refutation of Idealism," had argued that since we can distinguish the sensum blue that is the object of a sensation from the sensing itself, sensa might exist without consciousness of them, and they might therefore be nonmental. Against Moore, Ducasse argues in *Nature, Mind, and Death* that a sensum is not an "object" of sensation but the "content" of it. When one sees some lapis lazuli, the lapis lazuli is the object seen. But the relation of the lapis lazuli to the seeing of it when "I see some lapis lazuli" is true is not the same as the relation of blue to the seeing of it when "I see blue" is true. (Compare "I taste quinine" with "I taste bitter," or "I am jumping a ditch" with "I am jumping gracefully.") After a meticulous examination of various hypotheses on what the relation of sensa to sensing might be, Ducasse concludes that sensa are species of experience. "I sense blue" means "I sense bluely," or, alternatively, "I sense in the manner blue," just as "I am dancing a waltz" means "I am dancing waltzily (in the manner of dancing called 'dancing a waltz')." Just as a waltz could not conceivably exist apart from the dancing of it, a sensum could not exist apart from the sensing of it.

On the basis of this analysis, Ducasse submits that the basic criterion of the mental may be expressed by saying that "*if something being experienced is connate with the experiencing of it, then it is a mental primitive*."

AESTHETICS

In *The Philosophy of Art* (New York, 1929), *Art, the Critics, and You* (New York, 1944), and many articles, Ducasse formulates and defends an emotionalist theory of art and aesthetic experience. His principal contentions are that art in the broadest sense is skilled activity; that fine or aesthetic art consists in the skilled objectification of feeling; that the fine artist judges the adequacy of the work he creates not by the degree to which it approximates to beauty but by the faithfulness with which it reflects back to him the feeling to which he attempted to give objective expression; that in the aesthetic attitude one "throws oneself open" to the advent of feelings; and that judgments of aesthetic value are relative to the taste of the critic.

PHILOSOPHY OF RELIGION

In *A Philosophical Scrutiny of Religion* (New York, 1953), Ducasse defines religion as essentially any set of articles of faith, with the observances, feelings, and so on, tied thereto, that has the social function of motivating altruism in individuals and the personal function of giving the believer inner peace and assurance. According to this definition, belief in a God or gods is not essential to religion. Ducasse himself is not a theist. He holds that orthodox theism is contradicted by the existence of evil, and that polytheism is more plausible than monotheism conceived in the orthodox manner.

PARANORMAL PHENOMENA

Throughout his career, Ducasse was interested in and wrote about the "wild facts" of mental telepathy, clairvoyance, precognition, and so on. His interest in them was manifold. If paranormal phenomena do occur, received theories about the mental and the physical must be revised to account for them. It is a gratuitous assumption that any theory capable of taming the wild facts would have to postulate supernatural entities or "spooks." It could well be as scientific as are current theories about hypnotism, which have more or less tamed the wild facts of mesmerism. One of the troubles of psychical research is the lack of a fruitful theory.

If paranormal phenomena do occur, there would be important implications for philosophy. How would

philosophers have to conceive of time, causality, perception if there were such a thing as precognition?

It is a logical possibility that a mind survives the death of its body (or, to allow for reincarnation, bodies), even when due account has been taken of current science. But is there any evidence that it does? If there is, it is likely to be found by objective sifting of the reports concerning paranormal phenomena. In *A Critical Examination of the Belief in a Life after Death* (Springfield, IL, 1961), Ducasse states that the conclusion about survival seemingly warranted at present is that "the balance of the evidence so far obtained is on the side of the reality of survival," but that the evidence is not conclusive.

See also Aesthetic Experience; Art, Expression in; Causation: Metaphysical Issues; Hume, David; Logic, History of: Modern Logic; Mill, John Stuart; Moore, George Edward; Parapsychology; Reincarnation; Royce, Josiah; Sensa.

Bibliography

A complete bibliography of Ducasse's writings up to December 31, 1951, is available in *Philosophy and Phenomenological Research* 13 (1) (September 1952): 96–102. This issue also contains "Symposium in Honor of C. J. Ducasse" by seven philosophers, a biographical note, and a portrait.

For George Santayana's response to Ducasse's views on causation, "ontological liberalism," art, and properties, see *The Letters of George Santayana*, edited by Daniel Cory, 213–215, 234–235, and 287–288 (New York: Scribners, 1955). For a careful review of *Nature, Mind, and Death* by H. H. Price, see the *Journal of Parapsychology* 16 (2) (June 1952).

OTHER RECOMMENDED WORKS

Ducasse, Curt John. *Current Philosophical Issues: Essays in Honor of Curt John Ducasse*. Compiled and edited by Frederick C. Dommeyer. Springfield, IL: Thomas, 1966.

Ducasse, Curt John. *Truth, Knowledge and Causation*. London: Routledge & K. Paul; New York: Humanities P., 1969.

Vincent Tomas (1967)
Bibliography updated by Michael Farmer (2005)

DUHEM, PIERRE MAURICE MARIE

(1861–1916)

Pierre Maurice Marie Duhem was noted for his original work in theoretical physics, especially thermodynamics, and in the history and philosophy of science. He was born and studied in Paris, and at the age of twenty-five published an important book on thermodynamics. In 1887 he went to the faculty of sciences at Lille University, where he taught hydrodynamics, elasticity, and acoustics. He married but his wife soon died, leaving him with a daughter. In 1893 he moved to Rennes and in 1895 to a chair at Bordeaux University, which he held until his death. Throughout his life he was a Catholic and a conservative.

His approach to physics was systematic and mathematical, and his interest in axiomatic methods undoubtedly determined to some extent the nature of his philosophical account of scientific theories, contained mainly in his book *La théorie physique: son objet, sa structure* (*The Aim and Structure of Physical Theory*), first published in 1906. He wrote a great deal on the history of science, especially in the fields of mechanics, astronomy, and physics, largely because he believed that a knowledge of the history of a concept and of the problems it was designed to meet was essential for a proper understanding of that concept. For the scientist, the history of his subject should be not a mere hobby but an essential part of his scientific work. Duhem's most important works in this field are *Les origines de la statique,* published in 1905–1906, and *Le système du monde,* an account of various systems of astronomy, in eight volumes, published between 1913 and 1958.

SCIENCE AND METAPHYSICS

Duhem's account of physical theory is positivistic and pragmatic, having clear connections with those of Ernst Mach and Henri Poincaré. It begins with, and takes its character largely from, his views on explanation. Indeed, one might say that it begins with a dogmatic and unsupported presupposition about the nature of explanation. He says that to explain is "to strip reality of the appearances covering it like a veil, in order to see the bare reality itself."

But the sciences depend upon observation, and observation shows us no more than the appearances: it cannot penetrate to the reality beneath. This reality is the province of metaphysics; only metaphysics can explain. Science merely deals with the relations between, primarily, our sensations (or the appearance of the world to us) and, ultimately, our abstract ideas of these appearances. A physical theory is somehow an abstract representation of the relations between appearances and not a picture of the reality lurking behind them.

Thus, as far as science alone is concerned, Duhem is as antimetaphysical as Mach and more so than Heinrich Hertz. But, in general, he is not antimetaphysical at all. In a sense, metaphysics is the most important of all studies

because it penetrates to the reality of things and explains the appearances; but when we are doing science, we must never import into it metaphysical aims or ideas. Science and metaphysics are both highly respectable, but they are utterly distinct and must be kept so on pain of confusion.

We may, Duhem thinks, penetrate to reality, not by the methods of science, but by pure reason. He attaches great importance to the doctrine that man is free, a statement that cannot conflict with any of the conclusions of science. His metaphysical views, which he did not work out in detail, are Aristotelian; properly understood—that is, stripped of its outmoded science—the Aristotelian physics contains an accurate picture of the cosmological order, whose appearance to human beings is studied by the sciences.

Scientists, according to Duhem, have seldom made the distinction between science and metaphysics, with the result that many theories have been seen as attempted explanations and so have been garnished with strictly superfluous "pictorial" and explanatory elements. These theories can be divided into two parts, called by Duhem "representative" and "explanatory." What is valuable in such theories, and hence what survives and what may be common to apparently different theories, is the representative part.

THE USES OF THEORIES

This conception of the representative nature of theories is linked with the various ways in which theories are useful to us. First, they promote economy by connecting large numbers of experimental laws deductively under a few hypotheses or principles; we need remember only these principles instead of a large number of laws. Second, by classifying laws systematically they enable us to select the laws we need on a particular occasion for a particular purpose. Third, they enable us to predict, that is, to anticipate the results of experiments. These are functions that can be performed by the representative parts of theories, which merely link general statements derived from observation and experiment in a practically convenient way, rather than in a way that corresponds to the underlying reality of things.

THE CONSTRUCTION OF THEORIES

Duhem's account of the way in which theories are constructed exhibits his conception of the nature of physical theories. There are four fundamental operations in their construction.

(1) Among the observable, measurable properties that we wish our theory to represent, we look for a few that can be regarded as simple and as combining to form the rest. Because they are measurable, we can represent them by mathematical symbols. These symbols have no intrinsic connection with the properties they represent: they are conventional signs for these properties. For example, temperature measured in degrees centigrade is a conventional and quantitative representation of the felt warmth and cold of sense experience.

(2) We construct a small number of principles, or "hypotheses," which are propositions arbitrarily connecting our symbols in a manner controlled only by the requirements of convenience and logical consistency. We may give as an example the definition of "momentum" as the product of mass and velocity.

(3) We combine these hypotheses according to the rules of mathematical analysis; again there is no question of representing the real relations between properties, and convenience and consistency are still our guides.

(4) Certain of the consequences drawn out by our third operation are "translated" back into physical terms. That is, we arrive at new statements about the measurable properties of bodies, our methods of defining and measuring these properties serving as a kind of "dictionary" to assist us in the translation. These new statements can now be compared with the results of experiments; the theory is a good one if they fit, a bad one if they do not.

THE NATURE OF LAWS AND THEORIES

Thus, a physical theory, for Duhem, is always mathematical and is a conventional system of linkages between propositions "representing" general statements or laws arrived at by experiment or observation. It is a device for calculating, and nothing matters except that the results of the calculations square with our observations. We might illustrate this in the following way. There are various routes by plane from city *A* (the known laws) to city *B* (the new laws), and it does not matter which route we take as long as we arrive at *B*: We are flying blind; the plane has no windows, and we cannot see the landscape, the sun, or even the clouds during the journey; we must not suppose that the interior of the plane resembles *A* or *B* or the country in between.

The idea that physical characteristics are analyzable into basic elements that are simple and ultimate has figured largely in empiricist and positivist accounts of the sciences. This idea involves numerous difficulties, not the least among them being that of giving any precise meaning to simplicity. Duhem avoids some of the difficulties. Because physical theories do not explain, his simple elements need not be ultimate in nature; they need not be *incapable* of further analysis. They may merely be properties that we *take to be* fundamental and that we have not succeeded in analyzing.

Duhem distinguishes between "practical facts" and "theoretical facts." A description of a phenomenon in ordinary ("observational") language states a practical fact, and its translation into the symbols of the theory states a theoretical fact. But the theoretical fact, as should now be obvious, is a "fact" only in a very odd sense; it has some kind of formal correspondence with the practical fact, but it is always an approximation or an idealization and always has many alternatives.

There is a similar relation between empirical or "commonsense" laws and scientific laws. Scientific laws state the relations between symbols that derive their meanings from the theories of which they are a part. These laws are approximations and idealizations and do not state the relations between actual physical properties. As an example, Duhem cites Boyle's law. This states the relations, not between pressures that may be felt and volumes that may be seen, but between their ideal representatives in a complex theory of gases. The same word, *pressure,* may stand for different concepts in different theories, and in its commonsense, everyday use it stands for a concept or concepts different again from all these.

A commonsense law, such as "Paper is inflammable," is correctly said to be either true or false. No scientific law, however, can be said to be true or false because every accepted scientific law has equally acceptable alternatives. None of these alternatives is any more correct than any of the others. There are two points here. To call the law we actually accept "true" is to suggest that the acceptable alternatives are false, which is misleading. Moreover, all the possible alternatives are idealizations: there is nothing of which they can be said to be strictly true. The symbols used in scientific laws are always too simple to represent completely the phenomena and their connections; hence, the laws must always be provisional.

Duhem distinguishes between observation and interpretation in a way that would now be questioned by certain philosophers. An observer looking at a spot of light on a scale *may* be merely observing this spot, or he may be doing this *and* interpreting it as the final step in measuring the resistance of a coil. Here, observing needs only attentiveness and reliable eyesight, but interpreting requires a knowledge of electrical theory as well. A boy who knew nothing whatever about electrical theory could be given the task of recording the movements of the spot on the scale; a physicist who had not seen these movements but who knew the theory and was prepared to rely on the boy could interpret the records appropriately.

It follows from Duhem's account that scientific laws and theories are not arrived at by induction. No experiment in physics involves mere generalizing from observations because the description of the experiment and its result, in the appropriate terms, involves the use of our physical symbols and, therefore, an interpretation of the phenomena depending upon the acceptance of a particular theory.

Duhem has important things to say about the testing of scientific hypotheses and theories. An empirical generalization of the form "All *A*'s are *B*" can never be conclusively established, because we can never be sure that we have examined all the *A*'s, but it may be conclusively falsified by finding one *A* that is not *B.* Thus, if we take such a generalization to be the pattern of scientific hypotheses, we must say that these hypotheses are open to conclusive refutation. But this is too simple, for a scientific hypothesis can never be tested independently of other hypotheses. This is a point that probably has to be made for any adequate account of scientific theorizing, but it is clearly an essential part of Duhem's account. For him, a hypothesis is always part of a theory, and it is used to make predictions only along with other parts of the theory and perhaps other theories. The failure of a prediction, then, indicates some inadequacy in the hypothesis in question *or* in some other hypothesis of the theory *or* in another theory that has been assumed in making the prediction, but it does no more than this to locate the inadequacy. It shows conclusively that something is wrong, but it tells us neither where to look for that something nor what we must reject or modify.

Thus, there can be no crucial experiments in physics. The pattern of a crucial experiment is this: we have two conflicting hypotheses about a given phenomenon and we design an experiment that will give one specifiable result if one hypothesis is acceptable and the other not, and another specifiable result if the other is acceptable and the first not. But hypotheses are not, as this suggests, independent and isolable. In fact, we must always confront a whole theory, of which one hypothesis is a part, with another whole theory, of which the other hypothesis

is a part. It is much more difficult to devise an experiment to choose between theories, and even if we could, it might be that a theory that conflicts with the experiment could be squared with it by making minor modifications whereby it would become as acceptable as the other theory under test.

This view may be criticized on the grounds that it is logically possible to find a crucial experiment that would enable us to choose between two *theories.* Of course, a theory that conflicts with experimental results may be capable of modification so that it does not conflict, but if it then gives exactly the same deductions as its rival, it is doubtful that they can be regarded as different theories, in Duhem's view. On the other hand, if they give different deductions covering the same field, it remains logically possible to devise a conclusive experiment to choose between *these two theories.* Karl Popper objects to Duhem's view on the grounds that the only reason Duhem thought crucial experiments impossible was because he stressed verification rather than falsification. It is not clear that Popper's objection is valid, for Duhem seems to have noticed the obvious fact that the aim of a crucial experiment is to eliminate one of the theories.

Although there is much in common between Duhem's and Poincaré's accounts of scientific theories, Duhem uses this last point about theory modification in criticism of part of Poincaré's view. According to Poincaré and others, certain important hypotheses of physical theory cannot be refuted by experiment because they are *definitions.* For example, the statement that the acceleration of a freely falling body is constant really defines "freely falling"; if an experiment appears to conflict with this, the most we can say is that the body was not falling freely. Nothing we observe can compel us to reject the original statement because it is not an empirical statement. Duhem, in reply, gives a different reason why we sometimes treat scientific statements in this way. It is not that the hypotheses we treat in this way are definitions but that they cannot be tested in isolation; thus, we are usually free, in the face of an unfulfilled prediction, to keep any given hypothesis and reject some other. This does not mean that we shall never be forced to reject that given hypothesis in consequence of some other modification we make to the theory, but only that the odds are against this happening on any given occasion.

See also Continental Philosophy; Conventionalism; Explanation; Hertz, Heinrich Rudolf; Laws, Scientific; Mach, Ernst; Philosophy of Science, History of; Poincaré, Jules Henri; Scientific Method.

Bibliography

WORKS BY DUHEM

Le potentiel thermodynamique et ses applications à la mécanique chimique et à la théorie des phénomènes électriques. Paris, 1886. His first book.

Le mixte et la combination chimique. Essai sur l'évolution d'une idée. Paris, 1902.

Les théories électriques de J. Clerk Maxwell: Étude historique et critique. Paris, 1902.

L'évolution de la mécanique. Paris, 1903.

Les origines de la statique. Paris, 1905–1906.

La théorie physique, son objet et sa structure. Paris, 1906. Second edition with new appendix (1914) translated by P. P. Wiener as *The Aim and Structure of Physical Theory.* Princeton, NJ, 1954. This book contains Duhem's most important philosophical work.

Études sur Léonard de Vinci, ceux qu'il a lus et ceux qui l'ont lu. Paris, 1906–1913.

"Physics—History of." In *Catholic Encyclopedia,* Vol. 12, pp. 47–67. New York, 1911.

Le système du monde. Histoire des doctrines cosmologiques de Platon à Copernic, 8 vols. Paris, 1913–1958. An enormous historical work of considerable importance.

WORKS ON DUHEM

Agassi, J. "Duhem *versus* Galileo." *British Journal for the Philosophy of Science* 8 (1957): 237–248.

Duhem, H.-P. *Un savant français: P. Duhem.* Paris, 1936.

Frank, P. *Modern Science and Its Philosophy.* Cambridge, MA: Harvard University Press, 1949.

Ginzburg, B. "Duhem and Jordanus Nemorarius." *Isis* 25 (1936): 341–362.

Jammer, Max. *Concepts of Force.* Cambridge, MA: Harvard University Press, 1957.

Launay, L. de. "Pierre Duhem." *Revue des Deux Mondes* (1918): 363–396.

Lowinger, Armand. *The Methodology of Pierre Duhem.* New York: Columbia University Press, 1941. See also a review of this work by B. Ginzburg, *Isis* 34 (1942): 33–34.

Picard, E. *La vie et l'oeuvre de Pierre Duhem.* Paris, 1922.

Poincaré, H. "Sur la Valeur objective des théories physiques." *Revue de métaphysique et de morale* 10 (1902): 263–293.

Popper, K. R. *The Logic of Scientific Discovery.* London: Hutchinson, 1959.

Popper, K. R. *Conjectures and Refutations.* London: Routledge, 1963.

Rey, A. "La philosophie scientifique de M. Duhem." *Revue de métaphysique et de morale* 12 (1904): 699–744.

Roy, E. le. "Science et philosophie." *Revue de métaphysique et de morale* 7 (1899): 503.

Roy, E. le. "Un positivisme nouveau." *Revue de métaphysique et de morale* 9 (1901): 143–144.

Peter Alexander (1967)

DÜHRING, EUGEN KARL

(1833–1921)

Eugen Karl Dühring, the German philosopher and political economist, was born in Berlin and died in Nowawes, near Potsdam. Dühring practiced law in Berlin from 1856 to 1859, but an eye ailment, eventually leading to total blindness, forced him to abandon this career. In 1861 he took his doctorate in philosophy at the University of Berlin, with a dissertation titled *De Tempore, Spatio, Causalitate Atque de Analysis Infinitesimalis Logica.* He became university lecturer in 1863, but his feuding with colleagues and his attacks on the university led to his dismissal in 1877. From then until his death he lived the life of a private scholar. In his later years, Dühring's attacks on religion (*Asiatismus*), militarism, Marxism, the Bismarck state, the universities, and Judaism became more and more virulent. Nevertheless, he retained a small group of loyal followers who founded a journal primarily devoted to his essays, the *Personalist und Emanzipator* (1899). Three years after Dühring's death, E. Döll founded the Dühring-Bund.

Dühring's early views, expressed in his *Natürliche Dialektik,* were Kantian. Eventually, however, he came to reject Immanuel Kant's phenomena-noumena distinction, with its corollary that we do not apprehend reality as it is in itself. Dühring maintained that the mind does grasp reality directly, and that the laws of thought are in some sense also laws of being.

KNOWLEDGE AND REALITY

While denouncing metaphysics and every sort of supernaturalism, Dühring formulated a theory of reality that is no less metaphysical than that of the philosophers whom he attacked. Philosophy, according to Dühring, should aim at a comprehensive account of reality, an account that will be consonant with the natural sciences. A complete knowledge of reality is possible if we restrict ourselves to what is given, using the "rational imagination" that is the organ for philosophizing. (This constructive imagination is used also in mathematics, Dühring held.) The outcome of this activity, an activity of passion guided by the understanding, will be a coherent and comprehensive world picture. Dühring praised Arthur Schopenhauer, Ludwig Feuerbach, and Auguste Comte for their efforts in this direction.

The fundamental law that we are to use in apprehending reality is the Law of Determinate Number. This law provides an easy solution to the antinomies in which reason finds itself when seeking knowledge beyond the realm of possible experience. It states that all thinkable numbers are complete or determined, and that the notion of an infinite or undetermined number is therefore impossible. Dühring suggested that the conception of an infinity of events or of units is somehow logically contradictory, as if one were to speak of a countless number that had been counted. For the theory of reality, the consequences of Dühring's law are that the number of events in time that preceded the present moment must be finite, and so too must be the number of objects in space. The history of the universe must have had an absolute beginning, and every object that exists or has existed must be divisible into a finite number of parts. It is nevertheless possible, Dühring maintained, that time and space extend infinitely from here and now.

A "primordial being" lies beyond the first event in time, though this being can be defined only by negating the properties of objects and events in time. Still, we can say of it that it contains the "roots" of every event and object, though it does not consist of events and is not an object. History develops out of this primordial being by an evolutionary process, from the more homogeneous to the more diversified.

What is actual must be here and now. The past is no longer real. The primordial condition of being no longer exists, though its traces are still evident. The laws of the physical universe, the atoms that make up matter—these are the unchanging aspects of the world, the persistent traces of the primordial being.

CHANGE AND EVOLUTION

The evolution of the universe involves the coming into being of genuinely new forms, and there exists the possibility that further novelty will emerge with the passage of time. The coming into being of motion, and of living creatures and conscious agents, are examples of new phenomena in the transition from the original condition of the world to its present state. Productive, creative activity is an essential fact about the universe, yielding new existences, new phenomena. The laws that describe such changes are nevertheless constant. We do not clearly understand how such genuine novelty occurs, and we ought not to construct speculative hypotheses. An honest philosopher will simply confess his ignorance.

How the world may evolve in the future is also beyond our knowledge. Either natural processes will continue mechanically without ever coming to an end or, what is more probable, there will emerge something radically different. Dühring accepted the latter alternative for the reason that he believed differentiation is a basic law of

nature. However, since the number of possible changes is finite, there must be either an eternal recurrence of the world process, as Friedrich Nietzsche suggested, or an end.

MIND AND CONSCIOUSNESS

Dühring's philosophy of mind is at first glance dualistic. Conscious activity is totally different from inanimate processes. The former is, however, an outcome of the clash of mechanical processes or forces. The sensation of resistance is the most basic sort of consciousness, and it reveals very clearly that its origin is the antagonism of physical forces.

While Dühring's position is positivistic in its emphasis on the limitation of human knowledge to the world described by natural science, and in its rejection of any independent philosophical knowledge of reality, he differs from some nineteenth-century positivists, such as Ernst Mach, in rejecting phenomenalism as the only valid basis for knowledge. Dühring maintained that although no disembodied spirits or souls exist, the world that is given to consciousness is one that contains not only matter and physical forces but also life and activity. Furthermore, he did not repudiate the concepts of cause and force or approve of a reductionism that would restrict intelligible discourse to phenomena, a restriction that he called "a morbid and skeptical aberration."

RELIGION

In his passionate opposition to religion and to every form of mysticism, Dühring is reminiscent of Lucretius. Religion is "a cradle of delusions," he maintained, and it is only by becoming free from its superstitions that man can become truly noble. The idea of an "other world" is a stumbling block to the proper appreciation of the real world that we encounter directly. We must find our values in this world.

Dühring's teleological optimism led him to reject Charles Darwin's theory that a struggle for existence is necessitated by the insufficiency of means to satisfy natural needs. The conditions for happiness are not impossible, he said. Even pain exists as an enhancement of our appreciation of pleasure. Only manmade institutions stand in the way of human happiness; religion is one of these institutions. Science, as carried on in the nineteenth century, is equally pernicious, since it involves "a hodge podge of superstition, skepticism and apathy."

ETHICS AND ECONOMICS

Dühring held that the feeling of sympathy is the foundation of morality. In applying this theory to the field of economics, Dühring came to a conclusion that Friedrich Engels and other Marxists have found highly objectionable. The interests of capitalist and worker, Dühring maintained, are not really opposed. By means of free competition there could be an ultimate harmony and compatibility between the two classes. Dühring's economic doctrines also supported the idea of a "national" political economy. He advocated tariff protection of national industries as a means of promoting the culture and morality of all citizens in the state. This goal could be realized most effectively when the economy of a nation was self-sufficient.

NATIONALISM AND RACISM

Dühring was an ardent German patriot, and some of the enormous popularity that his writings enjoyed in the latter part of the nineteenth century can be traced to this. He worshiped Frederick the Great. Along with his nationalistic zeal, however, Dühring betrayed a generous amount of prejudice, denouncing Jews, Greeks, and even Johann Wolfgang von Goethe, who was too cosmopolitan for Dühring's taste. Some conjecture that Nietzsche was influenced by Dühring's *Wert des Lebens.* But the joyous affirmation of life that Dühring shared with Nietzsche stands in sharp contrast to the vicious, embittered tone of many of Dühring's writings, and Nietzsche's rejection of pessimism stands on quite other grounds than that of Dühring.

See also Comte, Auguste; Continental Philosophy; Darwin, Charles Robert; Darwinism; Engels, Friedrich; Eternal Return; Feuerbach, Ludwig Andreas; Goethe, Johann Wolfgang von; Kant, Immanuel; Lucretius; Nationalism; Nietzsche, Friedrich; Positivism; Racism; Schopenhauer, Arthur.

Bibliography

PRIMARY WORKS

Kapital und Arbeit. Berlin, 1865.

Der Wert des Lebens. Breslau, 1865.

Natürliche Dialektik. Berlin, 1865.

Kritische Geschichte der Philosophie. Berlin, 1869.

Kritische Geschichte der Nationalökonomie und des Sozialismus. Berlin: Grieben, 1871.

Kritische Geschichte der allgemeinen Prinzipien der Mechanik. Berlin, 1873.

Kursus der National- und Sozialökonomie. Berlin, 1873.

Kursus der Philosophie. Leipzig, 1875. Later editions are titled *Wirklichkeitsphilosophie.*
Logik und Wissenschaftstheorie. Leipzig: Fues, 1878.
Die Judenfrage. Karlsruhe and Leipzig, 1881.
Sache, Leben und Feinde. Karlsruhe, 1882. Autobiography.
Der Ersatz der Religion durch Vollkommeneres. Karlsruhe: Reuther, 1883.

SECONDARY WORKS

Adamiak, Richard. "Marx, Engels, and Dühring." *Journal of the History of Ideas* 35 (1) (1974): 98–112.
Albrecht, G. *Eugen Dührings Wertlehre.* Jena, 1914.
Döll, E. *Eugen Dühring.* Leipzig: Naumann, 1893.
Drechsler, Wolfgang. "Herrn Eugen Dühring's Remotion." *Trames* 3 (3) (1999): 99–130.
Druskowitz, H. *Eugen Dühring.* Heidelberg: Weiss, 1889.
Engels, F. *Herrn Eugen Dührings Umwälzung der Wissenschaft.* Leipzig: Genossenschaftsbuchdruckerei, 1878. Translated by E. Burns as *Herr Eugen Dühring's Revolution in Science.* London, 1935. Commonly known as *Anti-Dühring.* Attacks Dühring's philosophy, politics, and economics.
Reinhardt, H., ed. *Dühring and Nietzsche.* Leipzig, 1931.
Small, Robin. "Nietzsche, Dühring, and Time." *Journal of the History of Philosophy* 28 (2) (1990): 229–250.
Vaihinger, H. *Hartmann, Dühring und Lange.* Iserlohn: Baedecker, 1876.

Arnulf Zweig (1967)
Bibliography updated by Philip Reed (2005)

DUMMETT, MICHAEL ANTHONY EARDLEY

(1925–)

Michael Anthony Eardley Dummett is perhaps the most important philosopher of logic of the second half of the twentieth century. Born on June 27, 1925, in London, England, Dummett completed his formal education at Christ Church, Oxford, and served for many years on the faculty of that university. A fellow of All Soul's College from 1979 to 1992, Dummett was the Wykeham Professor of Logic. His influential work has made commonplace (though not uncontroversial) the claim that philosophical matters concerning logic and truth are central to metaphysics, understood in roughly the traditional sense. Dummett has profoundly and permanently shifted the ground of debates concerning metaphysical realism.

Much of Dummett's work has taken place in the context of his commentaries on Gottlob Frege, at whose hands, Dummett claims, epistemology was supplanted by the philosophy of language as the fundamental field of philosophical investigation. Frege's reorientation of philosophy, comparable to the Cartesian installation of epistemology as the foundation of philosophical thinking, finally directed philosophers' attention to the proper focus: the relation of language to reality. Dummett is thus a leading advocate of the "linguistic turn." He is heavily influenced by Ludwig Wittgenstein's later work and by intuitionism in the philosophy of mathematics.

Dummett claims to have articulated a common structure embodied in a number of disputes pitting realists against their opponents. For example, the medieval debate over universals consisted of realists, who argued for the existence of mind-independent, objective properties, against various denials of realism (conceptualism, nominalism). Realism's claim about material objects contrasts with varieties of idealism, all of which share the general view that material objects do not exist objectively and independently of the mind. Positions that are antagonistic toward the positing of an objective, mind-independent realm are antirealistic positions. Dummett holds that the proper way to approach the dispute is to investigate what logical principles that are valid on the realistic view must be abandoned by antirealism. In particular Dummett claims that the law of bivalence, according to which every meaningful statement is determinately either true or false, is the mark of realism.

According to Dummett, the route to antirealism must be a meaning-theoretical one and thus focus on the role of the notion of truth in explicating meaning. His position on the theory of meaning has been called verificationism but, more properly, should be called neoverificationism to distinguish him from logical positivism. Dummett argues that truth, if conceived realistically, cannot be the fundamental notion of a theory of meaning—that is, if truth is conceived as satisfying the principle of bivalence. He recommends abandoning this classical notion of truth. His positive proposal can be put either of two ways: he sometimes suggests that the classical notion of truth must be replaced by a different concept of truth, one that does not include the bivalence principle; at other times he suggests that truth be replaced by verification as the central meaning-theoretical notion.

The theory of meaning is concerned with the relationships of truth, meaning, and use. Holding to a sophisticated reading of the "meaning is use" idea, Dummett argues that a theory of meaning based on the classical notion of truth cannot successfully analyze the ability of speakers to use their language. That is, the meaning of a sentence cannot be identified with—or, more weakly, connected with sufficient intimacy to—the sentence's truth conditions if truth is conceived classically, because the resultant theory of meaning attributes to a speaker a grasp of meaning that cannot be explained in terms of

her possession of recognitional skills pertaining to truth, i.e., her possession of certain epistemic capacities.

Dummett's key arguments concerning this conclusion have been called the acquisition and manifestation arguments. Because some of the sentences of the language in question are undecidable (their truth or falsity cannot be recognized by means of "decision procedures"), it is inexplicable how a speaker is able to learn their truth-conditional meanings through training. Grasping the truth conditions of these sentences is beyond the ken of finite beings. Similarly, since a grasp of a sentence's meaning must be conclusively demonstrable in one's actions, it is inconceivable that a speaker could display competence in the language if this means demonstrating his or her grasp of a sentence's recognition-transcendent truth conditions. Because of this sensitivity of the theory of meaning to such epistemological concerns, Dummett concludes that the central explanatory notion of a theory of meaning cannot be epistemically unconstrained. Thus, a notion of truth requires sensitivity to the epistemic limitations of language users.

This requirement leads to an intuitionistic concept of truth, whereby bivalence fails and not all sentences can be said to possess a truth value despite being meaningful. Failure of bivalence may concern past-tense and future-tense sentences, attributions of dispositional properties to no-longer-existent objects that never displayed possession or lack of the dispositions in question, and, crucially, sentences involving unrestricted quantification over infinite domains. Further pursuit of this line leads Dummett into consideration and rejection of meaning-theoretical holism and to an emphasis on the role of logical inference in verification.

Dummett presents a compelling case for the interrelatedness of metaphysical questions and meaning-theoretical ones; in particular, he argues that notion of truth—and, concomitantly, the logic that correctly formalizes the corresponding notion of valid or truth-preserving inference—depends on a prior investigation in the theory of meaning.

Dummett's importance to philosophy lies in his demonstration of the ways in which metaphysics relates to the philosophy of logic and how those two fields in turn relate to the philosophy of language.

See also Frege, Gottlob; Idealism; Meaning; Phenomenalism; Philosophy of Language; Realism; Truth; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY DUMMETT

"What Is a Theory of Meaning?" In *Mind and Language*, edited by S. Guttenplan. Oxford: Oxford University Press, 1974.

"What Is a Theory of Meaning? (II)." In *Meaning and Truth*, edited by G. Evans and J. MacDowell. Oxford: Oxford University Press, 1974.

"The Justification of Deduction." In *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978.

"The Philosophical Basis of Intuitionistic Logic." In *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978.

"The Significance of Quine's Indeterminacy Thesis." In *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978.

Frege: Philosophy of Language. 2nd ed. Cambridge, MA: Harvard University Press, 1981.

The Interpretation of Frege's Philosophy. Cambridge, MA: Harvard University Press, 1981.

"Frege on the Third Realm." In *Frege and Other Philosophers*. Oxford: Oxford University Press, 1991.

The Logical Basis of Metaphysics. Cambridge, MA: Harvard University Press, 1991.

Frege: Philosophy of Mathematics. Cambridge, MA: Harvard University Press, 1993.

The Seas of Language. Oxford: Oxford University Press, 1993.

Michael Hand (1996, 2005)

DUNS SCOTUS, JOHN
(c. 1266–1308)

As with many of the medieval Schoolmen, little is known of the early life of John Duns, the Scot (or Scotus), a theologian and philosopher. From the record of his ordination to the priesthood by Bishop Oliver Sutton at Northampton on March 17, 1291, it is inferred that he was born early in 1266. Rival traditions, neither of which can be traced to medieval sources, link him with each of the two main branches of the Duns family in Scotland. According to one account, he was the son of Ninian Duns, a landowner who lived near Maxton in Roxburghshire, received his early schooling at Haddington, and in 1277 entered the Franciscan convent at Dumfries, where his uncle was guardian. Another popular tradition, however, states that his father was the younger son of the Duns of Grueldykes, whose estate was near the present village of Duns in Berwickshire. As a bachelor of theology, Scotus lectured on the *Sentences* of Peter Lombard at Cambridge (date unknown), at Oxford about 1300, and at Paris from 1302 to 1303, when he and others were banished for not taking the side of King Philip the Fair against Pope Boniface VIII in a quarrel over the taxation of church property for the wars with England. The exile

was short, however, for Scotus was back in Paris by 1304 and became regent master of theology in 1305. In 1307 he was transferred to the Franciscan study house at Cologne, where he died the following year.

WORKS

Scotus's early death interrupted the final editing of his most important work, the monumental commentary on the *Sentences* known as the *Ordinatio* (or in earlier editions as the *Commentaria Oxoniensia* or simply the *Opus Oxoniense*). An outgrowth of earlier lectures begun at Oxford and continued on the Continent, this final version was dictated to scribes, with instruction to implement it with materials from his Paris and Cambridge lectures. A modern critical edition of the *Ordinatio*, begun by the Typis Polyglottis Vaticanis (Vatican Press) in 1950, is still in progress. Though less extensive in scope, Scotus's *Quaestiones Quodlibetales* are almost as important; they express his most mature thinking as regent master at Paris. Also authentic are the *Quaestiones Subtilissimae in Metaphysicam* on Aristotle's *Metaphysics*; some forty-six shorter disputations held in Oxford and Paris and known as *Collationes*; and a series of logical writings in the form of questions on Porphyry's *Isagoge* and on Aristotle's *Categories, De Interpretatione* and *De Sophisticis Elenchis*. The *Tractatus de Primo Principio* is a short but important compendium of natural theology; drawing heavily upon the *Ordinatio*, it seems to be one of Scotus's latest works. Like the *Theoremata*, a work whose authenticity has been seriously questioned, the *Tractatus* was apparently dictated only in an incomplete form and left to some amanuensis to finish.

THEOLOGY AND PHILOSOPHY

Like the majority of the great thinkers of the late thirteenth and early fourteenth centuries, Scotus was a professional theologian rather than a philosopher. One of the privileges accorded mendicant friars like the Franciscans and Dominicans was that of beginning their studies for a mastership in theology without having first become a Master of Arts. The philosophical courses they took in preparation were pursued in study houses of their own order and were, as a rule, less extensive than those required of the candidate for an M.A. As a consequence of this educational program their commentaries on the philosophical works of Aristotle were usually written later than those on biblical works or on the *Sentences* of Peter Lombard; also, the most important features of their philosophy are frequently found in the context of a theological question. This does not mean that they confused theology with philosophy in principle, but only that in practice they used philosophy almost exclusively for systematic defense or explication of the data of revelation. But in so doing, these theologians assumed that philosophy as a work of reason unaided by faith played an autonomous role and had a competence of its own, limited though it might be where questions of man's nature and destiny were at issue.

This critical attitude concerning the respective spheres of philosophy and theology became more pronounced around the turn of the fourteenth century. Thus, we often find Scotus not only distinguishing in reply to a particular question the answers given by the theologians from those of the "philosophers" (Aristotle and his Arabic commentators) but also pointing out what the philosophers could have proved had they been better at their profession. On the other hand, the genuine interest in the logical structure of "science" (*episteme*), as Aristotle understood the term, led to an inevitable comparison of systematic theology with the requirements of a science such as Euclid's geometry.

Paradoxically, it is in the attempt of the Scholastics to show to what extent theology is or is not a science that we find the most important expressions of their ideas of a deductive system. This is particularly true of the lengthy discussions on the nature of theology in the prologue of Scotus's *Ordinatio*. Similarly, if we look for the origin of some important and influential philosophical concepts that lie at the heart of Galileo Galilei's mechanics, we find them in the medieval discussions of "the intension and remission of forms" (that is, how qualities like hot and white increase in intensity). It was in his analysis of how a man might grow in supernatural charity, for instance, that Scotus introduced his theory of how variations in the intensity of a quality might be treated quantitatively. This key notion, developed by the Merton Schoolmen and extended to the problem of motion, made possible Galileo's description of the free fall of bodies.

Scotus was most concerned with what philosophy has to say about God and the human spirit. Though his ethical views and philosophy of nature are not without interest, Scotus was primarily a metaphysician.

METAPHYSICS

Scotus was thoroughly familiar with the writings of Avicenna, whose concept of metaphysics Scotus brought to the service of theology. Avicenna agreed with Averroes that Aristotle's metaphysics was meant to be more than a collection of opinions (*doxa*) and had the character of a science (*episteme*) or body of demonstrated truths, where

"demonstration" is understood in the sense of the *Posterior Analytics*. They also agreed that this science was in great part concerned with God and the Intelligences responsible for the movement of the planetary spheres. But Averroes believed that the existence of God is proved by physics or natural science (by Aristotle's argument for a prime mover), whereas Avicenna developed a causal proof within the framework of metaphysics itself. Scotus argued that the Averroistic view subordinates Aristotle's "first philosophy" to physics when it should be autonomous. Moreover—and more important—one needs a metaphysician to prove that the "prime mover" is the First Being, and metaphysics provides more and better arguments for God's existence than this particular physical proof. Part of the difficulty with the physical proof stems from Aristotle's axiom that "whatever is moved is moved by another." Scotus did not regard this as intuitively evident or deducible from any other such principles. Furthermore, he saw numerous counterinstances in experience, such as man's free will or a body's continued motion after external force is removed.

THE TRANSCENDENTALS. Scotus saw metaphysics as an autonomous science concerned with the transcendentals, those realities or aspects of reality that transcend the physical. Its subject matter, as Avicenna rightly maintained, is being as being and its transcendental attributes. In contrast with St. Thomas Aquinas, who restricted transcendental to such notions as have the same extension as "being," Scotus treated any notion applicable to reality but not included in one of Aristotle's ten categories as transcendental. At least four classes of such can be enumerated. Being (*ens*) is the first of the transcendental notions. It is an irreducibly simple notion of widest extension that is used to designate any subject whose existence implies no contradiction. "Existence" refers to the real or extramental world. Next come the three attributes coextensive with being—"one," "true," and "good"—for to be capable of existing in the extramental world, the subject must have a certain unity and be capable of being known and being desired or willed. Third, there are an unlimited number of attributes such as "infinite-or-finite," "necessary-or-contingent," "cause-or-caused," and so on, that are coextensive with being only in disjunction. Finally, there are many other predicates whose formal notion or definition contains no hint of imperfection or limitation. These are known as pure or unqualified perfections. In addition to being (*ens*), its coextensive attributes, and the more perfect member of each disjunction, this class of transcendentals includes any attribute that can be ascribed to God, whether it pertain to him alone (such as omnipotence or omniscience) or whether it also is characteristic of certain creatures (such as wisdom, knowledge, free will).

Disjunctive attributes. Like Avicenna, Scotus regarded the disjunctive transcendentals as the most important for metaphysics, but being Christian, he conceived these supercategories of being somewhat differently. Avicenna held that creation proceeded from God by a necessary and inevitable process of emanation, whereas for Scotus creation was contingent and dependent on God's free election. Therefore, for Scotus the less perfect member of each disjunction represents only a possible type of real being, whereas for Avicenna these possible types must all eventually be actualized, and therefore the complete disjunction is a necessary consequence of "being." Scotus expressed this difference in what might be called his "law of disjunction":

> In the disjunctive attributes, while the entire disjunction cannot be demonstrated from "being," nevertheless as a universal rule by positing the less perfect extreme of some being, we can conclude that the more perfect extreme is realized in some other being. Thus it follows that if some being is finite, then some being is infinite, and if some being is contingent, then some being is necessary. For in such cases it is not possible for the more imperfect extreme of the disjunction to be extentially predicated of "being" particularly taken, unless the more perfect extreme be existentially verified of some other being upon which it depends. (*Ordinatio* I, 39)

The task of the metaphysician, then, is to work out the ways in which the various transcendental concepts entail one another. One of the more important conclusions that will emerge from such an analysis is that there is one, and only one, being in which all pure perfection coexists. Such an infinite being we call God.

PROOF FOR GOD'S EXISTENCE. Scotus suggested that the metaphysician might use any pair of disjunctives to prove God exists (and here he seems to be in the tradition of William of Auvergne and the "second way" of St. Bonaventure). However, the one metaphysical proof he chose to work out in any detail seems to be a synthesis of what he considered the best elements of all the proofs of his predecessors. Henry of Ghent, whose writings so often served as the springboard for Scotus's own discussion of any problem, had tried to bring some order into the many proofs advanced during the Middle Ages by grouping them under two general headings, the way of causality

and the way of eminence. The first drew its inspiration from Aristotelian principles, whereas the second was Augustinian in tone and stemmed from the School of St. Victor and the *Monologion* of St. Anselm. The way of causality was further divided by Henry accordingly as God is treated as the efficient, the final, or the exemplar cause of creatures.

Scotus simplified the causal approach by eliminating the exemplar cause as a distinct category. He treated it as merely a subdivision of efficiency and implied that the cause in question is intelligent and does not act by a blind impulse of nature. As for the way of eminence, it was treated not simply in terms of its Platonic or Augustinian origins but as having a foundation in Aristotelian principles as well. The proof was developed in two principal parts, one dealing with the relative attributes of the infinite being—efficiency, finality, and eminent perfection—and the second with the absolute property of his infinity. Given the infinity of God, Scotus essayed to show there can be but one such being. Each section is a concatenation of closely reasoned conclusions, some thirty-odd in all.

The argument was perhaps one of the most elaborate and detailed proofs for God's existence constructed during the Middle Ages, and apart from any intrinsic merit as a whole it is of considerable historical interest. From the time Scotus first formulated it, he subjected the proof to several revisions, mainly in the direction of greater conceptual economy and logical rigor. In what seems to be the final version (in the *Tractatus de Primo Principio*), the proof is prefaced by two chapters that represent an attempt to formalize what a Schoolman at the turn of the fourteenth century must have regarded as the basic axioms and theses of the science of metaphysics. Other interesting aspects of the argument appear in answer to possible objections to the proof. One anticipates Immanuel Kant's causal antinomy. Aristotle and his Arabic commentators maintained that the world with its cyclic growth and decay had no beginning. How, then, can one argue to the existence of an uncaused efficient cause? Scotus's solution reveals the influence of Avicenna. On the ground that whatever does not exist of itself has only the possibility of existence as something essential to itself, Avicenna argued that this holds not only of the moment a thing begins to be but of every subsequent moment as well. The true cause of any effect, then, must coexist with and conserve the effect and therefore must be distinguished from the ancillary chain of partial causes that succeed one another in time.

Scotus developed this distinction in terms of what he called an essential versus an accidental concatenation of causes. A series of generative causes such as grandparent, parent, and child, or any sequence of events such as those later analyzed by David Hume, would be causes only accidentally ordered to one another in the production of their final effect. Where an essential ordering or concatenation exists, all the causal factors must coexist both to produce and to conserve their effect. This is true whether they be of different types (such as material, formal, efficient, and final) or whether they be a chain of efficient or final causes, such as Avicenna postulated for the hierarchy of Intelligences between God and the material world. While infinite regress in accidentally ordered causes may be possible, Scotus said, the chain as a whole must be essentially ordered to some coexisting cause that guarantees the perpetuity of what is constant or cyclic about such repetitive productivity. But no philosopher postulates an infinite regress where the concatenation of causes is essential and all must coexist. One does not explain how any possible effect is actually conserved, for instance, by assuming an infinity of links upon which it depends.

Technical demonstration. How is any proof that begins with factual propositions demonstrative or scientific in Aristotle's sense of demonstrative? Are not all such premises contingent? With Avicenna obviously in mind, Scotus explained that pagan philosophers could admit that every factual proposition is necessarily true because of the deterministic chain of causes that links it to the first creative cause, God. According to pagan philosophers, this is true not only of eternal entities like primary matter or the inferior or secondary Intelligences but also of all temporal events brought about by the clockwork motions of the heavenly bodies that these Intelligences cause. Empirical explanations of temporal events are required only because the human mind is unable to trace all the intricate links of causal efficacy that make any given event a necessary and inevitable consequence of God's essential nature.

If such a theory were true, Scotus argued, it would eliminate all genuine contingency from the world and thus conflict with one of the most manifest truths of human experience, namely, that we are free to act otherwise than we do. Should one deny such an obvious fact, it is not argument he needs but punishment or perception. "If, as Avicenna says, those who deny a first principle should be beaten or exposed to fire until they concede that to burn or not to burn are not identical, so too ought those who deny that some being is contingent be exposed to torments until they concede that it is possible for them

not to be tormented" (*Ordinatio* I, 39). If true contingency exists, however, it can only be because the first cause does not create the world by any necessity of nature. But if the whole of creation depends upon God's free will, then every factual or existential statement about it will be radically contingent. How, then, can any proof from effect to cause satisfy Aristotle's demand that demonstration begin with necessary premises? One could argue legitimately, but not demonstratively, from such an obvious fact as contingency. Yet, Scotus maintained, it is possible to convert the argument into a technical demonstration by shifting to what is necessary and essential about any contingent fact, namely, its possibility. For while one cannot always infer actuality from possibility, the converse inference is universally valid. What is more, Scotus added, statements about such possibilities are necessary; hence, he preferred to construct the proof from efficiency in the mode of possibility thus: Something can be produced, therefore something can be productive; since an infinite regress or circularity in essentially concatenated causes is impossible, some uncaused agent must be possible and hence actual, since it cannot be both possible and incapable of being caused if it is not actually existing.

One can argue similarly of the possibility of a final cause or of a most perfect nature. (Scotus's argument in this connection bears a curious parallel to Ludwig Wittgenstein's about simple objects in the *Tractatus Logico-Philosophicus*.) Scotus saw God as the necessary or a priori condition required to make any contingent truth about the world possible; these possibilities must be a part of God's nature, "written into him from the beginning"; as source of all possibility, he himself cannot be "merely possible." It is in God's knowledge of, and power over, these limitless possibilities that we discover what is fixed, essential, and noncontingent about not only the actual world but about all possible worlds as well. Since God is the fixed locus in which all possibilities coexist, he must be infinite in knowledge, in power, and therefore in his essence or nature. Since contradictions arise if one assumes that more than one such infinite mind, power, or being exists, there can be but one God.

THEORY OF KNOWLEDGE

After establishing the existence of an infinite being to his own satisfaction, Scotus undertook an analysis of the concepts that enter into statements about God, and in so doing he threw considerable light upon his own theory of knowledge, particularly upon how he considered notions that transcend the level of sensible phenomena to be possible.

UNIVOCITY AND THE TRANSCENDENTALS. Some of the earlier Schoolmen like Alexander of Hales, St. Bonaventure, and Henry of Ghent fell back upon various theories of innatism or illuminationism (in which elements from St. Augustine and Avicenna were grafted upon the Aristotelian theory of knowledge) to account for such knowledge as seems to have no foundation in the data of the senses. These hybrid interpretations of Aristotle had this in common: His theory was used to explain only how general or universal concepts applicable to the visible world are abstracted from sense images. But where any notion applicable to God was involved, some illumination from a transcendent mind was thought to be required. Not only did this hold for notions obviously proper to God—such as "necessary being" and "omnipotence"—but also for such seemingly common transcendentals as "being," "true," and "good." Although the latter terms were predicated of creatures as well as of God, their meaning was not univocal. Associated with each term were two similar, and hence often indistinguishable, meanings, both simple and irreducible to any common denominator. One was believed to be proper to creatures and to be abstracted from sensible things by the aid of an agent intellect; the other was proper to God, and since it transcended in perfection anything to be found in creatures, it must be given from above. It was maintained that these innate ideas, impressed upon the soul at birth, lie dormant in the storehouse of the mind, to be recalled like forgotten memories when man encounters something analogous in sensible experience. The discovery of God in created things, then, was explained much like Plato's account of how man recalls the transcendent world of ideas.

As Aristotle's own writings became better known, however, the popularity of such theories diminished. More and more Scholastics followed Thomas Aquinas in rejecting any special illumination theory to explain man's knowledge of God, but like Thomas they failed to see that this required any modification of the traditional doctrine of the analogy of being and other transcendental terms. Scotus seems to have been the first to see the discrepancy between the two positions. He pointed out that if all of our general notions (including those of being and its transcendental attributes) are formed by reflecting upon sensible things, as Aristotle explained, then some notions such as being must be univocally predicable of God and creatures, or all knowledge of God becomes impossible. Arguing specifically against Henry of Ghent, who claimed

we have either a concept of being proper to God or one common to finite creatures, Scotus insisted on the need of a third or neutral notion of being as a common element in both the other concepts. This is evident, he said, because we can be certain that God is a being while remaining in doubt as to whether he is an infinite or a finite being. When we prove him to be infinite, this does not destroy but adds to our previous incomplete and imperfect notion of him. The same could be said of other transcendental notions, such as wisdom or goodness. Indeed, every irreducibly simple notion predicable of God must be univocally predicable of the finite and created thing from which it was abstracted. Any perfection of God is analogous to its created similitude, but we conceive such a perfection as something exclusive or proper to God through composite concepts constructed by affirming, denying, and interrelating conceptual elements that are simple and univocally predicable of creatures. For even though every such element is itself general, certain combinations thereof may serve to characterize one, and only one, thing. Although such concepts are proper to God, they retain their general character and do not express positively the unique individuality of the divine nature. Hence the need for proving that only one God exists.

Scotus also held that the transcendental notion of being (*ens*) is univocal to substance and accidents as well as to God and creatures. We have no more sensible experience of substance than we do of God; its very notion is a conceptual construct, and we would be unable to infer its existence if substance did not have something positive in common with our experiential data.

THE FORMAL DISTINCTION. The concept of the formal distinction, like univocity of being, is another characteristic metaphysical thesis connected with Scotus's theory of knowledge. Though usually associated with his name, the distinction did not originate with him. It represents a development of what is sometimes called the "virtual distinction" or "conceptual distinction with a foundation in the thing." The latter is an intermediate between the real distinction and that which is merely conceptual. The difference between the morning star and evening star, for example, is purely conceptual. Here one and the same thing, the planet Venus, is conceived and named in two different ways because of the different ways or contexts in which it appears to us. The real distinction, on the contrary, concerns two or more individual items, such as Plato and Socrates, body and soul, or substance and its accidents. Though two such things may coexist or even form a substantial unity or accidental aggregate, it is logically possible that one be separated from the other or even exist apart from the other. The Scholastics generally recognized the need of some intermediary distinction if the objectivity of our knowledge of things is to be safeguarded. How is it possible, they asked, to speak of a plurality of attributes or perfections in God when the divine nature is devoid of any real distinction? How is it possible for a creature to resemble God according to one such attribute and not another? Similarly, if the human soul is really simple, as many of the later Scholastics taught, how can it lack all objective distinction and still be like an angel by virtue of its rational powers and unlike the angel by reason of its sentient nature? All agree that it is possible for the human mind to conceive one of these intelligible aspects of a thing apart from another and that both concepts give a partial insight into what is objectively present to the thing known.

To put it another way, there is a certain isomorphism between concept and reality, in virtue of which concept may be said to be a likeness (*species*) or picture of reality. This "likeness" should not be construed in terms of the relatively simply way a snapshot depicts a scene, but perhaps something more akin to Wittgenstein's "logical picture," being based upon what shows itself in both the world of facts and our thoughts about the world. In virtue of this intelligibility of form, we can speak of *ratio* (the Latin equivalent of the Greek *logos* or the Avicennist *intentio*) either as in things or as in the mind. To the extent that this *ratio* or intelligible feature is a property or characteristic of a thing, we are justified in saying that the individual possessing it is a so-and-so. Though such *rationes* can be conceived one without the other because their definitions differ and what is implied by one is not necessarily implied by the other, nevertheless, as characteristic of a specific individual, they constitute one thing. They are not separable from that individual in the way the soul can be separated from the body, or a husband from his family. Not even the divine power can separate a soul from its powers or the common features of the individual from what is unique (his *haecceity*).

Thomas spoke of this nonidentity as conceptual, with the qualification that it does not arise merely in virtue of thinking mind but "by reason of a property of the thing itself." Henry of Ghent called it an intentional distinction, but he added that the distinction is only potential prior to our thinking about it. Scotus, however, argued that if something has the native ability to produce different concepts of itself in the mind, each concept reflecting a partial but incomplete insight into the thing's nature, then the distinction must be in some sense actual.

Put in another way, there must be several "formalities" in the thing (where "form" is understood as the objective basis for a concept and "little form" or formality as an intelligible aspect or feature of a thing that is less than the total intelligible content of a thing). Here again Scotus argued (on a line later followed by Wittgenstein) that a thing's possibilities, unlike their actualization, are not accidental but are essential to it and must have some actual basis. If a thing is virtually two things inasmuch as it is able to be grasped in two mutually exclusive ways, this nonidentity of intelligible content must be prior to our actually thinking about the thing, and to that extent it exists as a reality (*realitas*) or in other words, objectively. This nonidentity of realities, or formalities, is greatest in the case of the Trinity, where the peculiar properties of the three divine persons must be really identical with, but formally distinct from, the divine nature they have in common. This formal nonidentity holds also for the divine attributes, such as wisdom, knowledge, and love, which although really one are virtually many.

The formal distinction was also used by Scotus to explain the validity of our universal conceptions of individuals, a Scotistic thesis that influenced C. S. Peirce. Unlike the "nominalists," Scotus did not believe that the common features of things can be accounted for fully in terms of their being represented by a common term or class concept. Some objective basis for this inclusion is required, and this similarity or aspect in which one individual resembles another he called its common nature (*natura communis*). This common nature is indifferent to being either individualized (as it invariably is in the extramental world) or being recognized as a universal feature of several individuals (as it is when we relate the concept of this "nature," such as "man," to Peter or Paul). The common nature is individualized concretely by what Scotus called its thisness (haecceity), which is a formality other than the nature, a unique property that can characterize one, and only one, subject.

Scotus consequently rejected the Aristotelian-Thomistic thesis that the principle of individuation is identified somehow with matter by reason of matter's quantitative aspect. This thesis would seem to make individuality something extrinsic to the thing itself, or at least the effect of something really other than the thing itself, since matter or matter signed with quantity is really distinct from the form. The requirement of haecceity is a logical one, according to Scotus, for in practice we do not differentiate individual persons or objects because we know their respective haecceity (that is, their Petrinity, Paulinity, their "thisness," or "thatness"), but because of such accidental differences as being in different places at the same time, or having different colored hair or eyes. However, this individuating difference, he insisted, is known to God and can be known by man in a future life, where his intellect is not so dependent upon sense perception.

KNOWING AS AN ACTIVITY. Though Scotus rejected illumination in favor of what is basically an Aristotelian theory of knowledge, his teaching on the subject shows the influence of some other of Augustine's ideas, notably the active role of the intellect in cognition. Scotus's position is midway between the Aristotelian passivism (the "possible intellect" as a purely "passive potency" receives impressions from without) and Augustine's activism (the intellect as spiritual can act on matter, but matter cannot act upon the spirit or mind). Scotus believed that the so-called possible intellect actively cooperates in concept formation and other intellectual operations. This activity is something over and above that which is usually ascribed to the "agent intellect." Intellect and object (or something that is proxy for the object, such as the intelligible *species* where abstract knowledge is involved) interact as two mutually complementary principles (like man and woman in generation) to produce concepts. Since these concepts reflect only common or universal characteristics of individuals rather than what is uniquely singular about them, it cannot be the singular object itself that directly interacts with the mind, but an intelligible likeness (species) that carries information only about the "common nature" of the object and not its haecceity. The formation of such a likeness or species is the joint effect of the agent intellect and sense image working together as essentially ordered efficient causes. It is in this way that Scotus interpreted the Aristotelian distinction of agent and possible intellect.

INTUITIVE VERSUS ABSTRACTIVE COGNITION. Although the above description accounts for man's abstract intellectual knowledge, Scotus believed that the human mind is capable of intuitive knowledge as well. By this he understood a simple (nonjudgmental) awareness of an object as existing. Where abstract cognition leaves us unable to assert whether a thing exists or not, one can assert that it exists from intuitive cognition of anything. In such a case no intelligible species of the object need intervene, for the mind is in direct contact with the thing known. While most Scholastics limited intuitive knowledge to the sense level, Scotus argued that if the human intellect is capable only of abstract cognition—what can be abstracted from sense encounters in the way described

by Aristotle—then the face-to-face vision of God promised to us in the afterlife becomes impossible. Consequently, our ideas of the proper object of the human intellect must be expanded to account for this.

Scotus thought that rational considerations also require us to admit some degree of intuitive power in man even if the full ambit of this power cannot be established by a philosopher. There are many primary contingent propositions of which we are absolutely certain (such as "I doubt such and such" or "I am thinking of such and such"). Since this certitude cannot be accounted for by any amount of conceptual analysis of the propositional terms, we must admit some prior simple awareness of the existential situation that verifies the proposition. This cannot be mere sensory knowledge, since the existential judgment often involves conceptual or nonsensory meanings, as in the examples given above. It is not clear that Scotus wished to assert that in this life we have intuitive knowledge of anything more than our interior acts of mind, will, and so on. This would seem to limit intellectual intuition to reflective awareness and would be consistent with his statements that we have no direct or immediate knowledge of the haecceity of any extramental object. However, he believed that in the afterlife man by his native powers will be able to intuit any created thing, be it material or spiritual, and to that extent man's mind is not essentially inferior to that of the angel. On the other hand, it is not merely because of man's lapsed state that his mind is at present limited to knowing the intelligible features of sense data but also because of the natural harmony of body and mind that would obtain even in a purely natural state.

CERTITUDE. The human capacity for certitude was also discussed, with Henry of Ghent as the chief opponent. Henry, Scotus explained, appealed to illumination, not for the acquisition of our everyday concepts about the world, since these are obtained by abstraction, but for certitude of judgment. Although the "mechanics" of the process are not fully clear, two "mental images" or species are involved, one derived from creatures, the other imparted by divine illumination from above. Since both the human mind and the sensible object are subject to change, no species or likeness taken from the sensible object and impressed upon the mind will yield invariant truth. Something must needs be added from above. Scotus made short shrift of this theory. If the conclusion of a syllogism is no stronger than its weakest premises, neither does a blending of an immutable and a mutable species make for immutability. Furthermore, if the object is so radically mutable that nothing is invariant under change, then to know it as immutable is itself an error. By way of contrast, Scotus set out to show that certitude is possible without any special illumination. This is certainly the case with first principles and the conclusions necessarily entailed by them. Such necessary truths assert a connection or disconnection between concepts that is independent of the source of the concepts. It is not, for example, because we are actually in sense contact with a finite composite that we can assert that a "whole" of this kind is greater than a part thereof. Even if we erroneously perceive white as black and vice versa, a judgment like "white is not black" precludes any possible error because it depends only on a knowledge of the terms and not on how we arrived at that knowledge.

A second type of certitude concerns internal states of mind or actions. That we are feeling, willing, doubting are experiential facts that can be known with a degree of certitude equal to that of first principles or the conclusions they entail.

A third category concerns many propositions of natural science where a combination of experience and conceptual analysis gives us certitude. Reposing in our soul is the self-evident proposition: "Whatever occurs in a great many instances by a cause that is not free is the natural effect of that cause." Even if the terms are derived from erring senses, we know this to be true, for the very meaning of nature or natural cause is one that is neither free nor acts haphazardly. If experience reveals recurrent behavior patterns where no free intelligent agent is involved, then we are evidently dealing with a natural cause. If the same situation recurs, we can be certain at least of what *should* result therefrom. That the effect expected actually does occur depends upon two further conditions: that the natural course of events is not interrupted by some unforeseen causal factor and that God does not miraculously intervene. Even sensory perception can be analyzed critically to exclude any reasonable doubt. Conflicting sense reports produce such illusions as the stick immersed in water that feels straight yet appears to be bent. Yet there is always some self-evident principle possessed by our mind that enables us to decide which sense perceptual information is correct. Here it is the proposition "Any harder object is not broken by something soft that gives way before it." There are many areas of knowledge, then, where humans are perfectly well equipped to arrive at certitude without any special divine enlightenment.

THE DOMAIN OF CREATURES

EXEMPLAR IDEAS. Scholastics generally accepted Augustine's theory that before creatures are produced, they preexist in God's mind as archetypal ideas. Scotus differed from Bonaventure and Thomas, however, by denying that God knows creatures through such ideas. Every creature is limited and finite as to intelligible content. To make God's knowledge of a creature dependent upon this limited intelligibility of any given idea denigrates the perfection of his intellect; if there is any dependence of idea and intellect, it must be the other way round. Only the infinitely perfect essence can be regarded as logically, though not temporally, antecedent to God's knowledge of both himself and possible creatures. Since possible creatures are written into the divine nature itself, in knowing his nature God knows each possible creature, and in knowing the creature he gives it intelligibility and existence as an object of thought. Like the creative painter or sculptor who produces an idea of his masterpiece in his mind before embodying it in canvas or stone, God, if he is not to act blindly but intelligently, must have a guiding idea or "divine blueprint" of the creature that is logically prior to his decision to create it. Creatures, then, are dually dependent upon God; they depend upon his infinitely fertile knowledge for their conception as exemplar ideas, and they depend upon the divine election of his omnipotent will for their actual existence. This tendency to distinguish various "logical moments" in God, and in terms of their nonmutual entailment to set up some kind of order or "priority of nature" among them, is characteristic of much of Scotus's theological speculation and became a prime target for William of Ockham's subsequent criticism.

THEORY OF MATTER AND FORM. The hylomorphic interpretation formerly attributed to Scotus was based on the *De Rerum Principio*, now ascribed to Cardinal Vital du Four. Scotus, unlike most of his Franciscan predecessors, did not accept the view of Solomon ben Judah ibn Gabirol (Avicebron) that all creatures are composed of matter and form. He considered both angels and human souls as simple substances, devoid of any real parts, though they differ in the formal perfections they possess.

Since Scotus did not equate matter with potency (as did St. Bonaventure), nor did he consider it in any way a principle of individuation (as did St. Thomas), there was no reason to postulate it in spiritual creatures either to explain why they are not pure act like God or to account for the possibility of a plurality of individuals in the same species. Hence, against Thomas, Scotus argued that even though angels lack matter, more than one individual of the same species may exist. More important, Scotus, like John Peckham and Richard of Middleton before him, insisted that matter must be a positive entity. Peckham's view grew out of his Augustinian theory of matter as the seat of the "seminal reasons," but Scotus rejected this germinal interpretation of inchoate forms and argued that if matter is what Aristotle thought it to be, it must have some minimal entity or actuality apart from form. It is true that primary matter is said to be pure potency, but there are two types of such passive potency; one is called objective and refers to something that is simply nonexistent but that can be the object of some productive creation. Matter as the correlative of form, however, is a "subjective" potency or capacity; it is a neutral subject able to exist under different forms and hence is not really identical with any one of them. Absolutely speaking, God could give matter existence apart from all form, either accidental or substantial. In such a case, matter would exist much like a pure spirit or the human soul.

William of Ockham followed Scotus on this point, as well as in his view that the primary matter of the sun and planetary spheres is not any different from that found in terrestrial bodies, though the substantial form in question may be superior to that of terrestrial elements and compounds.

THE HUMAN SOUL AS FORM. From man's ability to think or reason, Scotus argued that the intellective soul is the substantial form that makes man precisely human. But to the extent that reason can prove the soul to be the form of the body, it becomes correspondingly more difficult to demonstrate that the soul will survive the death of the body. While the traditional arguments for immortality have probabilistic value, only faith can make one certain of this truth. On the other hand, if the soul must be a spiritual substance to account for the higher life of reason, at least one other perishable "form of corporeity" must be postulated to give primary matter the form of a human body. Though to this extent Scotus agreed with the pluriformists against St. Thomas, it is not so clear that he would postulate additional subsidiary forms. A virtual presence of the lower forms (elements and chemical compounds) in the form of corporeity would seem to suffice. The form of corporeity has dimensive quantity, that is, it is not the same in each and every part of the body, as is the human soul. The same may be said of the "souls" of plants and animals. Though the human soul has the formal perfections of both the vegetative and the animal souls, these components are not really distinct parts. A

formal distinction between the soul's faculties or powers suffices to account for this.

FREE WILL. Particularly in his conception of free will, Scotus departed in many respects from contemporary positions. The will is not simply an intellective appetite, a motor power or drive guided by intelligence rather than mere sense perception. Freedom of will, in other words, is not a simple logical consequence of intelligence but is unique among the agencies found in nature. All other active powers or potentialities (*potentiae activae*) are determined by their nature not only to act but to act in a specific way unless impeded by internal or external causes. But even when all the intrinsic or extrinsic conditions necessary for its operation are present, the free will need not act. Not only may it refrain from acting at all but it may act now one way, now another. The will has a twofold positive response toward a concrete thing or situation. It can love or seek what is good, or it can hate or shun what is evil. Moreover, it has an inborn inclination to do so. But unlike the sense appetites, the will need not follow its inclination. Scotus rejected Thomas's theory that man is free only if he sees some measure of imperfection or evil in a good object and that the will is necessitated by its end (the good as such), though it is free to choose between several means of attaining it.

But Scotus saw a still more basic freedom in the will, one that Aristotle and Plato failed to recognize. Their theory of human appetites and loves can be called physical in the original sense of that term. All striving, all activity stems from an imperfection in the agent, whose actions all tend to perfect or complete its nature. *Physis* or "nature" means literally what a thing is "born to be" or become. Since what perfects a thing is its good, and since striving for what is good is a form of love, we could say that all activity is sparked by love. The peculiarity of such "love," however, is that it can never be truly altruistic or even objective. It is radically self-centered in the sense that nature seeks primarily and above all else its own welfare. If at times we find what appears to be altruistic behavior, it is always a case where the "nature" or "species" is favored at the cost of the individual. But nature, either in its individual concretization or as a self-perpetuating species, must of necessity and in all that it does seek its own perfection. This is its supreme value, and the ultimate goal of its loves. Such a theory presents a dual difficulty for a Christian. How can one maintain that "God is love" (I John 4.16) and how can man love God above all things if self-perfection is his supreme value? Thomas tried to solve the problem within the general framework of the Aristotelian system by making God the perfection of man. In loving God as his supreme value, man is really loving himself. Love of friendship becomes possible to the extent that he loves another as an "other self." This solution had its drawbacks, for certain aspects of Christian mysticism must then be dealt with in a Procrustean way. It leaves unexplained certain facets of man's complex love life. Finally, the theory commits Thomas, as it did Aristotle, to maintain that the intellect, rather than free will, is the highest and most divine of man's powers—a view at odds with the whole Christian tradition and particularly with Augustine.

Scotus tried another tack, developing an idea suggested by St. Anselm of Canterbury. The will has a twofold inclination or attraction toward the good. One inclination is the affection for what is to our advantage (*affectio commodi*), which corresponds to the drive for the welfare of the self described above. It inclines man to seek his perfection and happiness in all that he does. If this tendency alone were operative, we would love God only because he is our greatest good, and man's perfected self (albeit perfected by union with God in knowledge and love) would be the supreme object of man's affection; it would be that which is loved for its own sake and for the sake of which all else is loved.

But there is a second and more noble tendency in the will, an inclination or affection for justice (*affectio justitiae*), so called because it inclines one to do justice to the objective goodness, the intrinsic value of a thing regardless of whether it happens to be a good for oneself or not. There are several distinguishing features of this "affection for justice." It inclines one to love a thing primarily for its own sake (its absolute worth) rather than for what it does or can do for one (its relative value). Hence, it leads one to love God in himself as the most perfect and adorable of objects, irrespective of the fact that he happens to love us in return or that such a love for God produces supreme delight or happiness in man as its concomitant effect. Third, it enables one to love his neighbor literally as himself (where each individual is of equal objective value). Finally, this love is not jealous of the beloved but seeks to make the beloved loved and appreciated by others. "Whoever loves perfectly, desires co-lovers for the beloved" (*Opus Oxoniense* III, 37). Recall the tendency to make others admire the beautiful or the sorrow felt when something perfectly lovely is unloved, desecrated, or destroyed. If the *affectio commodi* tends to utter selfishness as a limiting case, the first checkrein on its headlong self-seeking is the *affectio justitiae*. Scotus wrote:

> This affection for what is just is the first tempering influence on the affection for what is to our

> advantage. And inasmuch as our will need not actually seek that towards which the latter affection inclines us, nor need we seek it above all else, this affection for what is just, I say, is that liberty which is native or innate in the will, since it provides the first tempering influence on our affection for what is to our own advantage. (Ibid., II, 6, 2)

The will's basic liberty, in short, is that which frees it from the necessity of nature described by Aristotle, the need to seek its own perfection and fulfillment above all else. Here is the factor needed to account for the generous and genuinely altruistic features of human love inexplicable in terms of the physicalist theory.

Scotus therefore distinguished between the will with respect to its natural inclinations and the will as free. The former is the will considered as the seat of the affection for the advantageous. It views everything as something delightful, useful, or a good for oneself and leads to the love of desire (*velle concupiscentiae*). As free or rational (in accord with right reason), the will is the seat of the affection for justice that inclines us to love each thing "honestly" or as a *bonum honestum*, that is, for what it is in itself and hence for its own sake. Since only such love recognizes the supreme value and dignity of a person and finds its highest and most characteristic expression when directed toward another, it is usually called the love of friendship (*velle amicitiae*) or of wishing one well (*amor benevolentiae*).

ETHICAL AND POLITICAL PHILOSOPHY

Although not primarily an ethicist, Scotus did solve enough specific moral problems from the standpoint of his general system of ethics to make it clear that his ethical system falls well within the accepted code of Christian morality of the day. Yet it does have some distinctive features, most of them growing out of the theory of the will's native liberty. Without some such theory, Scotus did not believe a genuine ethics is possible. If man had only a "natural will" (a rational or intellectual appetite dominated by the inclination for self-fulfillment), he would be incapable of sin but subject to errors of judgment. On the other hand, if the will's freedom is taken to mean nothing more than simple liberation from this inclination of nature, its actions would become irrational and governed by chance or caprice. What is needed is some counterinclination that frees man from this need to follow his natural inclination yet is in accord with right reason. This is precisely the function of man's native freedom. Man's reason, when unimpeded by emotional considerations, is capable of arriving at a fairly objective estimate of the most important human actions in terms of the intrinsic worth of the goal attained, the effort expended, the consequences, and so on. By reason of its "affection for justice" the will is inclined to accept and to seek such intrinsic values, even when this runs counter to other natural inclinations of self-indulgence. But being free to disregard the inclination for self-indulgence and to follow the higher dictates of justice, man becomes responsible for the good or evil he foresees will result from either course of action. It is the exercise of this freedom that is a necessary, though not a sufficient, condition for any action to have a moral value.

The other requisite conditions become apparent if we consider the nature of moral goodness. An action may be called good on several counts. There is that transcendental goodness coextensive with being which means simply that, having some positive entity, a thing can be wanted or desired. But over and above this is that natural goodness which may or may not be present. Like bodily beauty, this accidental quality is a harmonious blend of all that becomes the thing in question. Actions also can have such a natural goodness. Walking, running, and the like may be done awkwardly or with a certain grace or beauty. More generally, an activity or operation of mind or will can be "in harmony with its efficient cause, its object, its purpose and its form and is naturally good when it has all that becomes it in this way" (*Opus Oxoniense* II, 40). But moral goodness goes beyond this natural goodness. "Even as beauty of body is an harmonious blend of all that becomes a body so far as size, color, figure and so on are concerned," Scotus wrote, "so the goodness of a moral act is a combination of all that is becoming to it according to right reason" (II, 40). One must consider not only the nature of the action itself but also all the circumstances, including the purpose of its performance. An otherwise naturally good action may be vitiated morally if circumstances forbid it or if it is done for an evil end.

Right reason tells us there is one action that can never be inordinate or unbecoming under any set of circumstances: the love of God for his own sake. "God is to be loved" is the first moral principle or ethical norm. This and its converse, "God must never be hated or dishonored," are two obligations from which God himself can never grant dispensation. He is the one absolute intrinsic value, which cannot be loved to excess; but "anything

other than God is good because God wills it and not vice versa" (III, 19).

Scotus argued here as in the case of the divine intellect. The intelligibility of a creature depends upon God's knowing it, and not the other way around. So too its actual value or goodness depends upon God's loving it with a creative love and not vice versa. This obviously applies to transcendental goodness, which is coextensive with a thing's being, but it also holds for natural and moral goodness as well. If the infinite perfection of God's will prevents it from being dependent or necessitated by any finite good, it also ensures that creation as a whole will be good. God is like a master craftsman. For all his artistic liberty, he cannot turn out a product that is badly done. Yet no particular creation is so perfect, beautiful, or good that God might not have produced another that is also good; neither must all evil or ugliness be absent, particularly where this stems from a creature's misuse of his freedom. Nevertheless, there are limits to which God's providence can allow evil to enter into the world picture. He may permit suffering and injustice so that humankind may learn the consequences of its misbehavior and through a collective sense of responsibility may right its social wrongs.

While certain actions may be naturally good or bad, they are not by that very fact invested with a moral value; they may still be morally indifferent even when all circumstances are taken into consideration. Only hatred and the "friendship-love" of God are invested with moral value of themselves, and as the motivation for otherwise naturally good or indifferent actions they may make the actions morally wrong or good. Otherwise, the action must be forbidden by God to be morally wrong or commanded by him to be morally good. To that extent, moral goodness too depends upon the will of God. However, it is important to know that some actions are good or bad only because God commands or forbids them, whereas he enjoins or prohibits other actions because they are naturally good or bad, that is, they are consonant or in conflict with man's nature in the sense that they tend to perfect it or do violence to it. Such are the precepts of the natural law embodied in the Decalogue and "written into man's heart."

But note that what makes obedience to this instinctual law of moral value is that it be recognized and intended as something willed by God; otherwise, good as it may be naturally, the action is morally indifferent. This too is a consequence of man's native liberty, which can be bound only by an absolute value or the will of its author. To the extent that the first two commandments are expressions of the first moral principle and its converse, God can never make their violation morally right or a matter of indifference; the same does not hold of the last seven, which regulate man's behavior to his fellow man. God granted genuine dispensations from natural law, permitting polygamy to the patriarchs so that the children of God might be multiplied when believers were few. This might be permitted again if plague or war so decimated the male population that race survival was threatened. In such a case, God would reveal this dispensation to man, probably through his church.

HUMAN SOCIETY. Although Scotus wrote little on the origin of civil power, his ideas of its origin resemble John Locke's. Society is naturally organized into families; but when they band into communities they find some higher authority necessary and agree to vest it in an individual or a group, and decide how it is to be perpetuated—for example, by election or hereditary succession. All political authority is derived from the consent of the governed, and no legislator may pass laws for private advantage or that conflict with the natural or divine positive law. Private property is a product of positive rather than natural law and may not be administered to the detriment of the common good. More striking, perhaps, than Scotus's social philosophy was his theological theory (which influenced Francisco Suárez and, more recently, Pierre Teilhard de Chardin) that the second person of the Trinity would have become incarnate even if man had not sinned. Intended as God's "firstborn of creatures," Christ represents the alpha and omega not only of human society but of all creation.

Known to posterity as the "subtle doctor," Scotus is admittedly a difficult thinker. Almost invariably his thought develops through an involved dialogue with unnamed contemporaries. Although this undoubtedly delighted his students and still interests the historian, it tries the patience of most readers. His style has neither the simplicity of St. Thomas's nor the beauty of Bonaventure's, yet as late as the seventeenth century he attracted more followers than they. Like students who unconsciously mimic the worst mannerisms of their mentor, many of Scotus's disciples seemed bent more on outdoing him in subtlety than in clarifying and developing his insights, so that for both the humanist and reformer "dunce" (a Duns-man) became a word of obloquy. Yet there have always been a hardy few who find the effort of exploring his mind rewarding. Even a poet like Gerard Manley Hopkins regarded his insights as unrivaled "be rival Italy or Greece," and the philosopher C. S. Peirce considered Scotus the greatest speculative mind of the

Middle Ages as well as one of the "profoundest metaphysicians that ever lived." Even existentialists, who deplore the efforts to cast his philosophy in Aristotle's mold of science, find his views on intuition, contingency, and freedom refreshing. Scotus's doctrine of haecceity, applied to the human person, invests each individual with a unique value as one wanted and loved by God, quite apart from any trait he shares with others or any contribution he might make to society.

Despite his genius for speculation, Scotus considered speculation merely a means to an end: "Thinking of God matters little, if he be not loved in contemplation." Against Aristotle, he appealed to "our philosopher, Paul," who recognized the supreme value of friendship and love, which, directed to God, make men truly wise.

See also Alexander of Hales; Anselm, St.; Aristotle; Augustine, St.; Augustinianism; Averroes; Averroism; Avicenna; Bonaventure, St.; Galileo Galilei; Henry of Ghent; Ibn Gabirol, Solomon ben Judah; Kant, Immanuel; Locke, John; Medieval Philosophy; Peckham, John; Peirce, Charles Sanders; Peter Lombard; Plato; Richard of Mediavilla; Saint Victor, School of; Scotism; Socrates; Suárez, Francisco; Teilhard de Chardin, Pierre; Thomas Aquinas, St.; Universals, A Historical Survey; William of Auvergne; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

EDITIONS AND TRANSLATIONS

Opera Omnia, edited by L. Wadding, 12 vols. (Lyons, 1639), reprinted with L. Vivès, ed., 26 vols. (Paris, 1891–1895), contains most authentic and some spurious works, with commentaries by seventeenth-century Scotists. The seven volumes of the critical Vatican edition, edited by C. Balić and others (Vatican City, 1950–), contain only the first book of the *Ordinatio* and seven distinctions of the Oxford lectures. The edition may run to thirty or forty volumes.

For *Tractatus de Primo Principio*, see M. Mueller's edition (Freiburg im Bresgau, 1941) and new editions with English translations by Evan Roche (St. Bonaventure, NY, 1949) and Allan Wolter (Chicago, 1965); the latter is titled *Duns Scotus: A Treatise on God as the First Principle*.

Wolter's book contains translations of two questions from the first Oxford lectures; his *Duns Scotus: Philosophical Writings* (Edinburgh and London, 1962) is in Latin and English, and the paperback reprint (Indianapolis, 1964) appears without Latin. The question translated in *Free Will*, edited by S. Morgenbesser and J. Walsh (Englewood Cliffs, NJ: Prentice-Hall, 1962), is Scotus's earlier view, which he modified slightly; cf. C. Balić, "Une Question inédite de J. D. Scot sur la volonté," in *Recherches de théologie ancienne et médiévale* 3 (1931): 191–208. A translation of a question on the need for theology appears in *Medieval Philosophy*, edited by Herman Shapiro (New York: Modern Library, 1964), and in Nathaniel Micklem, *Reason and Revelation* (Edinburgh, 1953); on Christ as alpha and omega of creation in C. Balić, *Theologiae Marianae Elementa* (Sibenik, Yugoslavia, 1933), a Latin edition, and Allan Wolter, "D. Scotus on the Predestination of Christ," in *The Cord* (St. Bonaventure, NY), 5 (December 1955): 366–372, an English translation.

STUDIES AND BIBLIOGRAPHIES

The best introduction to the vast literature of the nineteenth and twentieth centuries is O. Shäfer, *Bibliographia de Vita, Operibus et Doctrina I. D. Scoti Saecula XIX–XX* (Rome, 1955); also see A. B. Emden, *A Biographical Register of the University of Oxford to A.D. 1500*, Vol. I (Oxford, 1957), pp. 607–610, and the annual *Bibliographia Franciscana* (Rome), especially Vol. XI (1962) on.

Most general histories and studies as late as C. R. S. Harris, *Duns Scotus* (Oxford, 1927), use the inauthentic *De Rerum Principio* or other spurious works. Recommended are the following more recent histories of medieval philosophy: P. Böhner and Étienne Gilson, *Christliche Philosophie von ihren Anfängen bis Nikolaus von Cues*, 3rd ed. (Paderborn, 1954); Frederick Copleston, *A History of Philosophy*, Vol. II, Part II (Westminster, MD, 1950); Armand Maurer, *Medieval Philosophy* (New York: Random House, 1962); and Julius Weinberg, *A Short History of Medieval Philosophy* (Princeton: Princeton University Press, 1964).

See Franciscan Institute Publications, Philosophy Series (St. Bonaventure, NY): Allan Wolter, *Transcendentals and Their Function in the Metaphysics of Duns Scotus* (1946), for his metaphysics; P. Vier, *Evidence and Its Function according to Duns Scotus* (1947), and Sebastian Day, *Intuitive Cognition* (1947), for his theory of knowledge; and R. Effier, *J. D. Scotus and the Principle "Omne Quod Movetur ab Alio Movetur"* (1962), for his theory of motion and of the will.

W. Hoeres, *Der Wille als reine Vollkommenheit nach Duns Scotus* (Munich, 1962), on the will; J. F. Boler, *Charles Peirce and Scholastic Realism* (Seattle: University of Washington Press, 1963), on Peirce's relation to Scotus; and especially the volume of essays commemorating the seventh centenary of Scotus's birth, J. K. Ryan and B. Bonansea, eds., *Studies in Philosophy and the History of Philosophy* (Washington, DC: Catholic University of America Press, 1965), Vol. III, may also be consulted.

Allan B. Wolter, O.F.M. (1967)

DUNS SCOTUS, JOHN [ADDENDUM]

Perhaps the most important recent area of research in Scotus's philosophy has been in modal theory. There are two fundamental questions: To what extent does Scotus develop an understanding of modalities that is fundamentally *logical*, independent of states of affairs in the actual world? And to what extent are modal concepts dependent on divine causal activity? The two questions are distinct, in the sense that the first is about the *defini-*

tion of the various modal terms, whereas the second is about the explanation for the *fact* that there are modalities at all. If God were to cause modalities, then he would also cause the property of conceivability that is the mark of logical possibility. The first question, in particular, has important ramifications for Scotus's theory of the freedom of the will.

As always with Scotus, these questions do not admit of straightforward answers. On the first question, Scotus's modal thought has, as one modern commentator puts it, a "Janus-faced character." On the one hand, Scotus often formulates modal notions as though consistency (*repugnantia*, in Scotus's Latin) is the relevant root concept: A proposition is possible if it is consistent, impossible if it is inconsistent, contingent if its contradictory is consistent, necessary if its contradictory is inconsistent. On the other hand, however, Scotus frequently talks as though consistency requires the existence of real powers and capacities, such that, for example, "possibly *p*" is true only if there is an agent with the power to bring it about that *p*. The second of these accounts clearly owes a great deal to the Aristotelian modal notions of Scotus's predecessors, according to which, for example, something is possible if and only if it is at some time actual. Scotus's proof of God's existence exploits this second account. The real possibility of there being causes in the world (entailed by the fact that there are causes in the world) requires, given the impossibility of an infinite regress of causes, the real possibility of the existence of a first cause. But such a real possibility requires that the causal conditions for the existence of such a first cause be satisfied. Now, a first cause is, according to Scotus, one whose causal explanation is somehow internal to itself. So such a being must exist, or else the conditions for its possible existence are not satisfied.

The argument clearly reveals the problems associated with the second of Scotus's modal theories. Of far more interest in the history of philosophy is Scotus's development of the notion of what he labels "logical possibility," understood as pure consistency or conceivability. Thus, he sometimes defines *possible* as "that which does not include a contradiction," and in line with this defines *contingent* as follows: "I do not call contingent everything that is neither necessary nor everlasting, but that whose opposite could have happened when this did" (Scotus 1982, p. 85). So the contingent is that whose nonexistence does not entail a contradiction. The significance of the simultaneity claim is that contingency—and modality in general—is on this account to be thought of not temporally or diachronically but synchronically, in terms of conceivable states of affairs considered as alternatives at the same time.

Scotus uses this account to undergird his radically libertarian account of human freedom. The human will is such that it can, in exactly the same set of circumstances, determine itself to act or not to act, or to do *a* or to do not-*a*. But this account requires the notion of alternative possibilities at one and the same time. Given the other aspect of Scotus's "Janus-like character" on this issue, however, an acceptance of the synchronic notion of contingency also entails the notion of libertarian freedom. For *real* contingency—contingency in the real world, as Scotus believes to be observable—requires a real free power. Scotus uses this insight as part of his argument for the claim that the first cause (God) must be a free agent. This does not mean, of course, that every logical possibility has to correspond to some real power in the world. So the new modal theory could coherently be developed without any of the residual Aristotelian apparatus—something that occurs in the generation after Scotus, and then most notably in the work of Leibniz, on which Scotus and his followers were tremendously influential.

Modern discussions of the second modal question consciously or unconsciously reflect discussions among seventeenth-century followers of Scotus. On one rather Platonist reading of Scotus, modalities are wholly independent of God; they are preconditions that govern even divine thought and action. On another reading of Scotus, God alone determines the reality, though not the content, of modalities. If there were no God, then there would be no modalities, even though the content of the modalities is not something over which God has any control. A middle position holds that, according to Scotus, modalities cannot obtain in the absence of any other reality whatever, but that Scotus does not hold God to be necessarily the required cause of modalities. In the absence of God (a counterpossible claim canvassed by Scotus for the sake of argument), there would be modal facts if and only if there were some nonmodal facts to be the bearers of the modal facts.

Scotus is well aware of the objection that a counterpossible premise entails any conclusion. But he holds that there are, as it were, degrees of conceivability about counterpossibles. The nonexistence of God seems coherently conceivable—its self-contradictoriness is not immediately evident—and in this respect is unlike the concept of a married bachelor, or, in Scotus's essentialist example, an irrational man. Scotus holds that, on this basis, principled conclusions can be drawn from such "moderate" counterpossible premises.

THE COMMON NATURE AND UNIVERSALS

Scotus is one of the most important writers on the question of universals. D. M. Armstrong explicitly notes Scotus as taking a position on universals different from that of modern writers such as Armstrong himself. For in modern theories of universals, a universal is *numerically one* in all of its exemplifications. The ancient and medieval tradition, springing in various ways from Aristotle, Alexander of Aphrodisias, and the Neoplatonists, denies this claim about universals, and Scotus provides the fullest development and explication of the ancient tradition on this question.

The Islamic philosopher Avicenna provided the clearest distillation of the ancient tradition available to the medieval West. Avicenna, echoing a common earlier distinction, distinguished a kind-nature as such from the nature existing as a concept in the mind and the same nature existing in particulars. The kind-nature as such is the *content* of the concept. According to Avicenna, the notions of numerical unity and/or multiplicity cannot be built into the kind nature—it is neither one nor many — because the nature (humanity, for example) must be able to exist both in one thing (Socrates) and in many (all human beings). On this view, the kind-nature as such is nothing more than a theoretical construct unifying the concept, on the one hand, with the particulars on the other.

Scotus accepts the threefold understanding of nature but holds that the kind-nature as such must be more than a merely theoretical construct. The kind-nature, in the account of Avicenna, is supposed to be the *subject* of both individuality (as existent in particular substances) and of universality (as existent in the mind). Scotus reasons that something that is supposed to be the subject of individuality and universality must have some real being or entity of its own. And this means that the nature as such must have such entity. The nature as such also has a certain identity or unity—Scotus calls it a "less-than-numerical" unity, compatible with divisibility into different particulars. The nature is thus identical in all its instances, but in a nonnumerical way. And this marks the way in which the developed ancient and medieval accounts are distinct from the modern accounts of, say, Armstrong who, as he puts it, "cannot understand what this second, lesser, sort of identity is" (Armstrong 1978, p. 112), (Scotus rejects views such as Armstrong's because he does not see how a numerically singular item could be the subject of different and incompatible properties in different particulars.) In line with this argument, Scotus holds that individuation is fundamentally a matter of explaining indivisibility: An individual is not divisible in the way that a nature is; this explanation must in turn be something that is intrinsically indivisible—namely, a haecceity or thisness.

See also Alexander of Aphrodisias; Aristotle; Aristotelianism; Armstrong, David M.; Avicenna; Leibniz, Gottfried Wilhelm; Modal Logic; Neoplatonism; Platonism and the Platonic Tradition; Socrates; Universals, A Historical Survey.

Bibliography

WORKS BY SCOTUS

A Treatise on God as First Principle ("De Primo Principio"), edited by Allan B. Wolter. Chicago: Franciscan Herald Press, 1982.

Philosophical Writings: A Selection, edited and translated by Allan B. Wolter. Indianapolis, IN, and Cambridge: Hackett, 1987.

Opera Philosophica, edited by Girard. J. Etzkorn, et al. St. Bonaventure, NY: The Franciscan Institute, 1997.

Frank, William A., and Allan B. Wolter, eds. *Duns Scotus: Metaphysician*. West Lafayette, IN: Purdue University Press, 1995.

Spade, Paul Vincent, ed. *Five Texts on the Mediaeval Problem of Universals: Porphyry, Boethius, Abelard, Duns Scotus, Ockham*. Indianapolis, IN, and Cambridge: Hackett, 1994.

Wolter, Allan B., ed. *Duns Scotus on the Will and Morality*. Washington, DC: Catholic University of America Press, 1986.

WORKS ABOUT DUNS SCOTUS

Armstrong, D. M. *Nominalism and Realism*. Cambridge, U.K.: Cambridge University Press, 1978.

Bos, E. P., ed. *John Duns Scotus 1265/6–1308: Renewal of Philosophy*. Amsterdam and Atlanta: Rodopi, 1998.

Cross, Richard. *The Physics of Duns Scotus: The Scientific Context of a Theological Vision*. Oxford: Clarendon Press, 1998.

Hoffmann, Tobias. *Creatura Intellecta: Die Ideen und Possibilien bei Duns Scotus mit Ausblick auf Franz von Mayronis, Poncius und Matrius*. Münster: Aschendorff, 2002.

Honnefelder, Ludger, Rega Wood, and Mechthild Dreyer, eds. *John Duns Scotus: Metaphysics and Ethics*. Leiden, New York, and Cologne: E. J. Brill, 1996.

King, Peter. "Duns Scotus on the Reality of Self-Change." In *Self-Motion from Aristotle to Newton*, edited by Mary Louise Gill and James G. Lennox, 229–290. Princeton, NJ: Princeton University Press, 1994.

Knuuttila, Simo. *Modalities in Medieval Philosophy*. London and New York: Routledge, 1993.

Sileo, Leonardo, ed., *Methodologica ad mentem Ioannis Duns Scoti*. Rome: Antonianum, 1995.

Tweedale, Martin M. *Scotus vs. Ockham: A Medieval Dispute over Universals*. 2 vols. Lewiston, Queenstown, and Lampeter: The Edwin Mellen Press, 1999.

Williams, Thomas, ed. *The Cambridge Companion to Duns Scotus*. Cambridge, U.K.: Cambridge University Press, 2003.

Wolter, Allan B. *The Philosophical Theology of John Duns Scotus*. Edited by Marilyn McCord Adams. Ithaca, NY: Cornell University Press, 1990.

Richard Cross (2005)

DURANDUS OF SAINT-POURÇAIN

(c. 1275–1334)

Durandus of Saint-Pourçain, the scholastic philosopher and theologian, bishop, and author (*Doctor Modernus, Doctor Fundatus*), was born in Saint-Pourçain-sur-Sioule in Auvergne, France. He entered the Dominican order at Clermont at the age of eighteen, and his philosophical studies were probably completed in his own priory of Clermont. By 1303 he was assigned to St. Jacques, Paris, to study theology at the university. There, according to some historians, he was influenced by his confrere James of Metz. The first version of Durandus's commentary on the *Sentences* of Peter Lombard represents his lectures as bachelor (1307–1308). In these lectures he strongly opposed certain views of Thomas Aquinas, whom the Dominican order had in 1286 commanded its members to study, promote, and defend. At Paris the nominalistic views of Durandus were immediately attacked by Hervé Nédellec and Peter of La Palu. Consequently, between 1310 and 1313 Durandus prepared a revision of his commentary, in which he mitigated many of his previous statements and omitted the more offensive passages. However, this was neither satisfactory to the order nor in accord with his own convictions. Nevertheless, he was granted a license by the university to incept in theology, succeeding Yves of Caen. Before completing his first year as master (1312–1313), he was called to Avignon by Pope Clement V to lecture in the papal *Curia,* replacing Peter Godin. Toward the end of that year the master general of the Dominicans, Berengar of Landorra, appointed a commission of nine theologians, headed by Hervé Nédellec, to examine the writings of Durandus. The commission singled out ninety-three propositions that were contrary to Thomistic teaching. Between 1314 and 1317, Durandus was continuously attacked in Paris by Hervé Nédellec, Peter of La Palu, John of Naples, James of Lausanne, Guido Terreni, and Gerard of Bologna. He replied to these in his *Excusationes* and in his Advent disputations *de quolibet* at Avignon (1314–1316). In the first *Quodlibet* he inveighed against "certain idiots" who charged him with Pelagianism or semi-Pelagianism.

Consecrated bishop in 1317, Durandus prepared a third and final version of his commentary on the *Sentences,* now free from all control by his order. He expressed regret that the first version had been circulated outside the order against his wishes, "before it had been sufficiently corrected" by him, insisting that only this new version was to be recognized as definitive. However, while some views are closer to the "common teaching" of the schools, the final version contains much that was taken verbatim from the first draft and from the first Avignon *Quodlibet.* It is, perhaps, not surprising that the final version, completed in 1327, abounds in compromises and contradictions.

In the jurisdictional dispute between Pope John XXII and Philip VI of France, Durandus sided with the pope in the treatise *On the Source of Authority* (1328), a work that later was published by Peter Bertrandi as his own composition. However, Durandus's reply to the pope's theological opinion concerning the beatific vision (1333) was promptly submitted to a commission of theologians, who found eleven objectionable statements. The reply of "the blessed master Durandus" was later vindicated by Benedict XII. But Durandus did not live to see himself vindicated, for he died at Meaux in 1334.

In philosophical matters Durandus manifested an independence of spirit more influenced by Augustine and Bonaventure than by Aristotle and Thomas. He has often been called a precursor of William of Ockham, but the similarities are only incidental; and it is most unlikely that either philosopher influenced the other. Besides denying the Thomistic distinction between essence and existence in creatures (as did Hervé Nédellec), he rejected the reality of mental species and the distinction between agent and possible intellect. For him, only individuals exist, receiving their individuality not from matter but from their efficient cause. Thus, in the act of knowing, the possible intellect is sufficiently active of itself to grasp individual existents directly and to create universal concepts by eliminating individual differences from consideration. In theology he manifested certain nominalist and Pelagian tendencies typical of the *moderni* of his day, tendencies that were to assume a more radical form in the teaching of Ockham.

In the later Middle Ages the prestige of Durandus was considerable. In the sixteenth century his final *Commentary on the Sentences* enjoyed an extraordinarily high reputation, particularly after its first printing (Paris, 1508). At Salamanca it was one of the alternative texts in the faculty of theology, the others being the *Summa* of Thomas and the *Sentences* of Peter Lombard, and the

chair of Durandus rivaled those of Thomas and John Duns Scotus.

Later writers have sometimes confused this Durandus with William Durand, Durandus Petit, or Durandus Ferrandi.

See also Aristotle; Augustine, St.; Bonaventure, St.; Duns Scotus, John; Medieval Philosophy; Pelagius and Pelagianism; Peter Lombard; Thomas Aquinas, St.; Thomism; Universals, A Historical Survey; William of Ockham.

Bibliography

WORKS BY DURANDUS

De Jurisdictione Ecclesiastica et de Legibus. Paris, 1506. Published under the name of Peter Bertrandi.

Commentaria in Quatuor Libros Sententiarum. Paris, 1508, 1515, 1533, 1539, 1547, 1550; Lyons, 1533 and 1569; Antwerp, 1567; Venice, 1571 and 1586 (Venice, 1571, reprinted by The Gregg Press, Ridgewood, NJ, 1964).

Questiones de libero arbitrio. In Prospero T. Stella, "Le 'Quaestiones de libero arbitrio' di Durando da S. Porciano." *Salesianum* 24 (1962): 450–523.

Avignon disputations *de quolibet.* In *Quolibeta Avenionensia Tria,* edited by Prospero T. Stella. Zürich, 1965.

Libellus de visione Dei. In G. Cremascoli, "Il 'Libellus de visione Dei' di Durando di S. Porziano." *Studi Medievali* 25 (1984): 393–443.

Votum de paupertate Christi et apostolorum. In Jürgen Miethke, "Das *Votum de paupertate Christi et apostolorum* des Durandus von S. Porciano." In *Vera lex historiae. Studien zu mittelalterlichen Quellen. Festschrift für D. Kurze,* 149–196. Cologne, 1993.

Kaeppeli, Th., O.P., ed. *Scriptores Ordinis Praedicatorum Medii Aevi.* Vol. I. Rome, 1970, #927–960, pp. 341–350 (a repertorium of Durandus's works, with selected bibliography; see also the supplement in ibid., vol. IV, pp. 73–74. Rome, 1993).

WORKS ON DURANDUS

Fumagalli, M. T. B.-B. *Durando di S. Porziano. Elementi filosofici della terza redazione del 'Commento alle Sentenze.'* Florence, 1969.

Glorieux, P. *Répertoire des maîtres en theologie.* Paris, 1933. Vol. I, No. 70, pp. 214–220. The bibliography is relatively complete up to 1933.

Henninger, Mark G. "Durand of Saint Pourçain (B. CA. 1270; D. 1334)." In *Individuation in Scholasticism: The Later Middle Ages and the Counter-Reformation, 1150–1650,* edited by J. J. E. Gracia, 319–332. Albany, NY, 1994.

Koch, Joseph. *Durandus de S. Porciano.* Beiträge zur Geschichte der Philosophie des Mittelalters 26. Münster, 1927.

Quétif, J., and Échard, J. *Scriptores Ordinis Praedicatorum Recensiti*. Paris, 1719. Vol. I, pp. 586–587.

Schabel, Chris, Russell L. Friedman, and Irene Balcoyiannopoulou. "Peter of Palude and the Parisian Reaction to Durand of St. Pourçain on Future Contingents." *Archivum Fratrum Praedicatorum* 71 (2001): 183–300 (including an introduction to Durand, his works, his troubles with his order, and his place in contemporary debate on future contingents).

Stella, Prospero T. "Le 'Quaestiones de libero arbitrio' di Durando da S. Porciano." *Salesianum*, Vol. 24 (1962), 450–523. Additions and corrections to Koch's work.

James A. Weisheipl, O.P. (1967)
Bibliography updated by Russell Friedman (2005)

DURATION

See *Bergson, Henri; Time*

DURKHEIM, ÉMILE
(1858–1917)

The French sociologist and philosopher Émile Durkheim was born in Épinal (Vosges). At an early age Durkheim decided not to follow the rabbinical tradition of his family. On leaving the Collège d'Épinal Durkheim went to Paris, first to the Lycée Louis-le-Grand, and then, in 1879, to the École Normale Supérieure. He was dissatisfied with what he saw as a too literary, unscientific style of education, connected with a superficial dilettantism in contemporary philosophy. On graduating in 1882, he decided to devote his career to sociology with the aim of establishing an intellectually respectable, positive science of society to replace, or at least supplement, speculative philosophy and provide an intellectual foundation for the institutions of the Third Republic. At an early stage, then, Durkheim developed a preoccupation which was to dominate his whole intellectual life—to establish a genuine science of social life, which would include a science of ethics and thus provide a reliable guide to social policy.

INFLUENCES AND INTELLECTUAL DEVELOPMENT

From 1882 to 1887 he was professor of philosophy at *lycées* in Sens, Saint-Quentin, and Troyes, during which time various intellectual influences helped him to fill out his conception of a social science. His study of Herbert Spencer instilled in him a predilection for biological models, which was most pronounced in his early work. His reading of Alfred Espinas, and later personal contact with him, led him to his central conception of the "collective consciousness" of a society and the related conviction that the laws of social life are sui generis and not reducible, for instance, to laws of individual psychology.

In "Individual and Collective Representations" (1898) he argued that we should not attempt to infer social laws from biological laws, but that the findings of biology should be compared subsequently with independently established social laws on the assumption that "all organisms must have certain characteristics in common which are worth while studying." His conception of a positive science of ethics received a powerful new impetus from a visit to Wilhelm Wundt's psychophysical laboratory in Leipzig while on a leave of absence during the school term of 1885–1886. In 1887 he was appointed *chargé de cours* at the University of Bordeaux, becoming the first to teach social science at a French university; he also taught pedagogy and thus began to develop an enduring interest in the relevance of sociology to educational questions.

In 1896 Durkheim was promoted to professor of social science at Bordeaux. In 1898 he founded and became editor of *L'année sociologique*, a journal designed to unify the social sciences and encourage specific research projects. He moved to the University of Paris as *chargé de cours* in 1902, becoming professor of education in 1906 and professor of education and sociology in 1913. The outbreak of war in 1914 moved Durkheim to write a number of pamphlets with a strongly nationalistic tone, not always easy to reconcile with the views developed in his earlier, more scholarly works.

THE COLLECTIVE CONSCIENCE

Durkheim's determination to establish an autonomous, specialized science of sociology led him to investigate the possibility of viewing human societies as irreducible, sui generis, entities. From there he was led to the central conception in his work, that of "collective representations," whose system in a given society constitutes its "collective conscience." Collective representations have both an intellectual and an emotional aspect. As examples Durkheim offered a language, a currency, a set of professional practices, and the "material culture" of a society; but he also included the phenomenon of group emotions, such as may be generated, for example, at a lynching, and which cannot be accounted for as a mere summation of the individual emotions of the several participants.

Durkheim said that collective representations are "collective" rather than "universal"; they "exist outside the individual consciousness," on which they operate "coercively." It is possible to determine collective representations directly—not merely via the thoughts and emotions of individuals—by examining their permanent expressions in, for instance, systems of written law, works of art, and literature, and by working with statistical averages. Thus, in *Suicide* Durkheim said that the "social fact" was the statistical suicide rate, not the circumstances attending individual suicides. His treatment of the relations between collective and individual representations, however, was often obscure, and he would pass from statements about the social determinants of the suicide rate to statements like this: "Human deliberations … are often only purely formal, with no object but confirmation of a resolve previously formed for reasons unknown to consciousness." His important conception of social forces thus took on a questionable, metaphysical complexion.

NORMAL AND PATHOLOGICAL SOCIAL TYPES

The conception of "social solidarity" went with that of collective representations and provided Durkheim with a means of distinguishing social types. The simplest form of social group is the "horde," which exhibits a "mechanical" solidarity in which individuals are attached directly to the group by adherence to a common set of powerful collective sentiments. The "clan" is the horde considered as an element in a more extensive group, and the most primitive form of durable social group is the segmental society organized in clans. More complex societies exhibit "organic" solidarity with extensive division of labor: the collective conscience is weak and individuals are attached to functional groups, while the society's cohesion is to be seen in the complex interdependence of these groups.

The distinction between social types led to a conception of "normal" and "pathological" forms, which provided a basis for Durkheim's account of the practical, ethical relevance of sociology. The normal is so only relative to a given social type at a particular stage of development. It may thus be difficult to determine, particularly during transitional phases. But once we have determined it in a particular case, the normal will merge with the average, though the sociologist must also attempt to show how the normal condition of a species follows logically from its nature. Durkheim believed that we can thus distinguish between social "health" and "disease" by means of "an objective criterion, inherent in the facts themselves"; for, he argued, on Darwinian lines, the dissemination of a characteristic throughout a species would be inexplicable if we did not suppose it to be on the whole advantageous. The sociologist, like the physician, should try "to maintain the normal state."

Durkheim applied this precept in the practical conclusions he drew from his study of suicide. It is important to maintain collective sentiment against suicide, at least those types of suicide most characteristic of organic soli-

darity, since the general ideal of humanity is the sole remaining strong collective sentiment, and the practice of suicide offends this sentiment. He advocated making use of the special nature of societies with organic solidarity in order to counteract suicide, by strengthening occupational groups and allowing them to take a firmer grip on the lives of individuals.

Durkheim's most influential discussion of a pathological social situation concerned "anomie." Anomie is characteristic of advanced organic societies and comes about when diverse social functions are in too tenuous or too intermittent mutual contact. Anomic division of labor exhibits itself in commercial crises, conflicts between capital and labor, and the disintegration of intellectual work through specialization. In relation to individuals the result of anomie is that "society's influence is lacking in the basically individual passions, thus leaving them without a check-rein." Durkheim used this concept to explain such phenomena as the high correlation between suicide and widowhood and between the suicide rate and the divorce rate.

FUNCTION AND CAUSE

Closely connected with his position on suicide and collective sentiments is Durkheim's concept of "function" as a mode of sociological explanation. He defined "function" as a relation between a system of vital movements and a set of needs. The prime need of any social collectivity is solidarity among its members, and Durkheim's main attempts at functional explanation, as in his treatments of the social division of labor, punishment, and primitive religion, were designed to show how such institutions or practices contribute to the type of solidarity peculiar to the societies in which they occur. The function of a practice is not to be confused with any aims of its practitioners; this would be to confuse sociology with psychology. But neither did Durkheim identify the function of a practice with its cause. The function of a fact does not explain its origin or nature: that would imply an impossible anticipation of consequences. Explanations of origins require the concept of an "efficient cause," though the persistence of a practice may be explained by the fact that its function helps to maintain a preexisting cause.

The causes of social facts are always to be found in preceding social facts, in the "internal constitution of the social group," or "social milieu." This concept, Durkheim held, is what makes sociology possible, by facilitating the establishment of genuinely social causal relations. Without it there could be only historical explanation, showing how events were possible, but not how they were predetermined. The social milieu was defined in terms of the volume of the group, the degree of communication between its members, and their concentration. Durkheim used this last concept to explain the development of the division of labor. Greater density of population brings with it a sharpened struggle for existence between individuals and this, in turn, makes necessary a greater degree of specialization. The division of labor is thus a "mellowed dénouement" of the struggle for existence.

Durkheim regarded causation as a species of *logical* relation; it was J. S. Mill's failure to recognize this, Durkheim held, that led him to speak erroneously of a possible plurality of causes. The most important method of establishing causal relations in sociology is that of concomitant variations, which can establish a genuine "internal bond" between phenomena as opposed to a merely "external" relation.

PRIMITIVE RELIGION AND CATEGORIES OF THE INTELLECT

In his treatment of primitive religion Durkheim was more immediately interested in functional than in causal questions, though he did not distinguish these as carefully as in *The Division of Labor in Society*, using apparently interchangeable phrases like "respond to the same needs" and "depend on the same causes." He also seems to have confused questions about the function of religions with questions about their meaning and truth. All religions "hold to reality and express it"; all "are true in their own fashion; all answer, though in different ways, to the given conditions of human existence." Durkheim rejected both the animistic account of primitive religions offered by Spencer and E. B. Tylor and the naturalistic account originating with Max Müller; both went astray, he felt, in masking such religions vast systems of error. Durkheim saw totemism as the most fundamental feature of primitive religions; he tried to show that the totem symbolizes not merely the totemic principle (or "god"), but also the clan itself, and this is possible because "the god and society are only one." Religion is "primarily a system of ideas with which the individuals represent to themselves the society of which they are members, and the obscure but intimate relations which they have with it." He thus regarded the explicit content of religious ideas as relatively unimportant. The reality they express is a sociological one, concealed from the worshipers themselves.

Durkheim regarded religion as the mother of thought. The categories of the intellect, such as "class," "force," "space," and "time," originate with religion. Moreover, since the reality expressed by religion is a social one,

these categories themselves originally correspond to forms of social organization and activity. Because totemism involves the idea of forces permeating both the natural and the human realms, it solves the Kantian problem of how men can apply these categories to nature. The a priori necessity of these categories is a reflection of society's coercive insistence on the ritual performances in terms of which such concepts are originally used.

See also Mill, John Stuart; Philosophy of Social Sciences; Social and Political Philosophy; Society; Sociology of Knowledge.

Bibliography

WORKS BY DURKHEIM

De la division du travail social. Paris: Alcan, 1893. Translated by G. Simpson as *The Division of Labor in Society*. Glencoe, IL, 1952.

Les règles de la méthode sociologique. Paris: Alcan, 1895. Translated by S. A. Solovay and J. H. Mueller as *The Rules of Sociological Method*. Glencoe, IL, 1950.

Le suicide. Paris: Alcan, 1897. Translated by J. A. Spaulding and G. Simpson as *Suicide*. Glencoe, IL: Free Press, 1951.

Les formes élémentaires de la vie religieuse. Paris: Alcan, 1912. Translated by J. W. Swain as *The Elementary Forms of the Religious Life*. London: Allen and Unwin, 1915; Glencoe, IL: Free Press, 1954.

Education et sociologie. Paris: Alcan, 1922. Translated by Sherwood D. Fox as *Education and Sociology*. Glencoe, IL: Free Press, 1956.

Sociologie et Philosophie. Paris: Alcan, 1924. Translated by D. F. Pocock as *Sociology and Philosophy*. Glencoe, IL: Free Press, 1953. Includes "Individual and Collective Representations."

L'éducation morale. Paris: Alcan, 1925. Translated by Herman Schnurer as *Moral Education*, edited by Everett K. Wilson. New York: Free Press of Glencoe, 1961.

Leçons de sociologie: physique de moeurs et du droit. Paris: Presses Universitaires de France, 1950. Translated by C. Brookfield as *Professional Ethics and Civic Morals*. London: Routledge, 1957. The last three books, published posthumously, contain the ideas developed in Durkheim's university lectures.

On Morality and Society: Selected Writings. Chicago: University of Chicago Press, 1973.

Durkheim and the Law. Edited by Steven Lukes and Andrew T. Steven. New York: St. Martin's Press, 1983.

Durkheim's Philosophy Lectures: Notes from the Lycée de Sens Course, 1883–1884. Edited and translated by Neil Gross, Robert Alun Jones, and André Lalande. Cambridge, U.K.; New York: Cambridge University Press, 2004.

WORKS ON DURKHEIM

Alpert, Harry. *Émile Durkheim and His Sociology*. New York: Columbia University Press, 1939.

Cladis, Mark Sydney. *A Communitarian Defense of Liberalism: Émile Durkheim and Contemporary Social Theory*. Stanford, CA: Stanford University Press, 1992.

Hall, Robert T. *Émile Durkheim: Ethics and the Sociology of Morals*. New York: Greenwood Press, 1987.

Jones, Robert Alun. *The Development of Durkheim's Social Realism*. Cambridge, U.K.; New York: Cambridge University Press, 1999.

Jones, Robert Alun. *Émile Durkheim: An Introduction to Four Major Works*. Beverly Hills, CA: Sage Publications, 1986.

LaCapra, Dominick. *Émile Durkheim: Sociologist and Philosopher*. Ithaca, NY: Cornell University Press, 1972.

Lehmann, Jennifer M. *Durkheim and Women*. Lincoln: University of Nebraska Press, 1994.

Lukes, Steven. *Émile Durkheim: His Life and Work, a Historical and Critical Study*. New York: Harper & Row, 1972.

Mestrovic, Stjepan Gabriel. *The Coming Fin de Siècle: An Application of Durkheim's Sociology to Modernity and Postmodernism*. London; New York: Routledge, 1991.

Mestrovic, Stjepan Gabriel. *Durkheim and Postmodern Culture*. New York: A. de Gruyter, 1992.

Nandan, Yash. *The Durkheimian School: A Systematic and Comprehensive Bibliography*. Westport, CT: Greenwood Press, 1977.

Nielsen, Donald A. *Three Faces of God: Society, Religion, and the Categories of Totality in the Philosophy of Émile Durkheim*. Albany: State University of New York Press, 1999.

Nisbet, Robert A. *Émile Durkheim*. Englewood Cliffs, NJ: Prentice-Hall, 1965.

Parsons, Talcott. *The Structure of a Social Action*. New York: McGraw-Hill, 1937; Glencoe, IL: Free Press, 1949.

Pickering, W. S. F. *Durkheim: Essays on Morals and Education*. London; Boston: Routledge & Kegan Paul, 1979.

Pickering, W. S. F. *Durkheim's Sociology of Religion: Themes and Theories*. London; Boston: Routledge & Kegan Paul, 1984.

Poggi, Gianfranco. *Durkheim*. Oxford; New York: Oxford University Press, 2000.

Schmaus, Warren. *Durkheim's Philosophy of Science and the Sociology of Knowledge: Creating an Intellectual Niche*. Chicago: University of Chicago Press, 1994.

Wallwork, Ernest. *Durkheim: Morality and Milieu*. Cambridge, MA: Harvard University Press, 1972.

Wolff, Kurt H., ed. *Émile Durkheim, 1858–1917; a Collection of Essays, with Translations and a Bibliography*. Columbus, OH, 1962.

Peter Winch (1967)
Bibliography updated by Michael J. Farmer (2005)

DUTY

In practical reasoning of an informal sort, the concept of duty plays a limited, relatively unproblematic role. In thinking about what to do, reasonable people try to see their wants in relation to their interests and to the interests of others; they evaluate alternatives in the light of their previous commitments and bear in mind their obligations and responsibilities. Duty is one among other factors to be taken into account. The reason is obvious: A person's duties are the things he or she is expected to do

by virtue of having taken on a job or assumed some definite office. One could say (although it sounds somewhat redundant) that believing that one's duties entail doing something or other is a reason, though not a conclusive one, for doing that thing, and believing that a possible line of action would count as a neglect of duty is a reason against adopting that line of action. How much weight such considerations have depends on what duties are in question and on the agent's obligations as they affect the particular situation. Duties, then, are counted as one of the considerations that guide and constrain rational choice.

The concept of duty in theoretical ethics is quite a different matter. Some moral philosophers (F. H. Bradley would be one example, Cicero another) have concerned themselves with duties of the everyday sort, those that go with being a parent, voter, teacher, or whatever. But many philosophers use "duty" quite indiscriminately to refer to particular obligations, moral principles, or indeed to anything that is held to be a requirement of conscience. "Duty" is a technical term in ethics and the rules for its use vary from one writer to another. For the most part, these differences are of no theoretical interest, but there is one important exception, the doctrine of Immanuel Kant. His views, set forth in the *Critique of Practical Reason* and in the *Foundations of the Metaphysics of Morals,* mark a radical break with traditional ethics, and since what he takes to be the central concept of morals he calls "duty," it is worthwhile finding out what he means by it.

ORDINARY DUTIES

As noted above, ordinary duties are tasks or assignments for which a man becomes responsible as a result of holding a particular job or office. When the tasks are intricate and have to be done just right, for example, the duties of an airplane pilot, then they are spelled out in detail; thus also for tasks that are relatively simple but for which applicants are unlikely to be highly motivated or imaginative, for example, the duties of a night watchman. In contrast, the duties that go with being a parent or with the practice of a profession are not codified, and responsibility for deciding what should be done is assigned to the individual.

Someone who neglects his duties deserves blame. Censure, if reasonable, is graduated to accord with the degree of neglect and with the importance of the task. A host who fails in his duties to his guests is inconsiderate but does not deserve to be pilloried. Negligence on the part of a pharmacist or a bus driver is a more serious matter. A characteristic of duties, as distinct from other constraints on conduct, is that a man who is delinquent loses, at some point, his title to the office that his duties define. He is court-martialed, unfrocked, disbarred, or fired (compare the euphemism "relieved of his duties"). Ceremonial dismissals are appropriate, of course, only when the duties in question are, in a broad sense, institutional and have been formulated explicitly. Not all duties fit this pattern; a man may become unfit for an office without being declared to be so, without his dereliction being so much as noticed by anyone, including himself. Someone who fails in the duties of friendship is simply no longer a friend, no matter what he or anyone else may think.

Legal penalties attach to neglect of duties where such neglect is held to be seriously detrimental to human welfare. Where a verdict has to be reached, an offense must be clearly defined. Parents, physicians, and legislators are among those to whom the greatest measure of discretion is granted in discharging their duties. It is an odd consequence that in matters of the greatest human importance only gross and flagrant derelictions of duty are punishable by law. Of course there are extralegal sanctions, and the threat of contempt and blame, of ostracism from one's group, may be a strong incentive to duty. The penalties of social disapproval, however, are distributed in a capricious and often unreasonable way, and a man may neglect all sorts of duties and yet, given discretion and a certain amount of luck, escape criticism altogether. Appreciation of this fact is what leads those concerned with moral education to try to instill in their charges a sense of duty. The attempt succeeds to the extent that the subject becomes habitually conscientious and carries out his duties without thinking about whether he might neglect them with impunity. A more primitive stratagem is to introduce the fiction of an all-seeing Providence in the hope of making the subject believe that no lapses go unnoticed and that all who neglect their duties will, on some unspecified future date, be punished.

Since duties are required minimal performances, no special merit accrues to someone who does his duty. A hero, one who does something that is both worthwhile and hazardous, acts "beyond the call of duty." A modest hero disclaims credit by saying that he did no more than his duty required. A man may be praised for carrying out some particular duty under difficult conditions. Such praise is sometimes justified and sometimes not; the claims of any duty may on occasion be outweighed by the claims of obligation or moral principle.

Although being conscientious is a virtue, it is not the only one, and unless it is mediated by intelligence and moral sensitivity, it may do more harm than good. A man

must learn, for example, how to deal with conflicting duties. If he is a jobholder, a parent, and a citizen, then he holds three offices concurrently. Even if his life is well organized, situations are likely to arise in which he has to determine which of two duties takes precedence. Such questions have to be worked out in particular cases; there is no formula or principle of ranking that can be applied. Moreover, as noted earlier, questions about duties are not independent of broad moral issues: if, as seems likely, there are offices one ought not, as a matter of moral principle, to hold, then there are duties no one ought to perform, even when called upon to do so.

KANT'S DOCTRINE

The idea of taking duty (*die Pflicht*) as the central moral concept originates with Kant. There are earlier doctrines that appear, especially when paraphrased, to be analogous, but the similarities are inconsequential in contrast with the differences. Kant himself maintained that his basic thesis is neither original nor esoteric and that, on the contrary, it is self-evident to the plain man. Everyone, he held, recognizes the difference between doing something because one wants to do it and doing something because one feels that one is morally obligated to do it. Moreover, it is universally acknowledged that only what is done from a sense of moral obligation is meritorious. Kant's theory is an exposition of what he took to be the consequences of these premises. He did not claim that the *theory* is easy and familiar. (In fact, he is often obscure and difficult to follow.) He did claim that his theory is the one philosophers must eventually accept if they are consistent and if they take seriously the intimations of the plain man.

The views Kant ascribed to common sense appear to be correct: people do not deserve credit unless they act from reasons of conscience, and we do believe that such reasons are, somehow or other, distinctive. Kant used the word *duty* (and here he diverged, at least from ordinary English usage) to refer very generally to features he took to be distinctive of conscientious conduct. At times this practice leads to rhetorical vagueness, and "duty" becomes synonymous with "whatever ought to be done." However, he also gave it a more precise sense, one that appears in the set of interdependent definitions which, taken together, provide the framework of his theory. In brief, he held that the only unqualified good is the "good will" and that to have a good will is always to act from a sense of duty.

Duty involves recognition of and submission to the "moral law" that is the "supreme principle" of morality. Since what the moral law prescribes goes (more or less) against the grain, that is, runs counter to inclinations, the law is expressed as an imperative. The imperative is described as being "categorical" and "unconditioned," and Kant meant these modifiers to reinforce the distinction mentioned earlier: objects of desire are variable and evanescent, and thus strategies for achieving such objects are applicable under some conditions and not under others. The moral law, however, applies to everyone and is unrestricted with respect to times, places, and particular situations.

The "categorical imperative" is formulated in three ways that Kant seems to have regarded as equivalent. They are as follows: "So act that the maxim of your will could always hold at the same time as a principle establishing universal law"; "Act so as to treat humanity, whether in your own person or in that of another, always as an end and never as a means only"; "Act according to the maxims of a universally legislative member of a merely potential kingdom of ends." Apart from the question of how to collate these formulas, difficult problems of interpretation arise for each of them taken separately. Nonetheless, one can see in a general way what Kant had in mind: A man is dutiful to the extent that he is seriously concerned with being equitable and fair, with treating other people like human beings and not like machines, and with trying to govern his own behavior by standards that could be adopted by everyone.

Kant believed that the concepts of duty, the good will, and the moral law are all such as can be apprehended a priori. Part of what he meant (and what is certainly true) is that no conclusions about what ought to be done can be derived directly from compilations of facts about what people do or have done. Although Kant was much concerned with distinguishing actual laws that depend on external sanctions from the moral law that the individual imposes on himself, he characterized the moral life by means of a set of juristic metaphors. The righteous man, for example, is said to "accuse himself before the bar of his conscience." This device suggests that Kant believed the "kingdom of ends" invoked in the third version of the categorical imperative to be an ideal beyond the hope of achievement. Human inclinations are apt to be anarchic, and as duty is a kind of inner law, so conscience is prefigured as a stern magistrate.

PRE-KANTIAN DOCTRINES

It is customary to cite the Stoics as the earliest philosophers to elevate duty to the status of a first principle. As far as one can tell from their writings, however, which

tend to vagueness, and from sketchy accounts of what they were reputed to believe, their views were quite different from Kant's. In fact, their word *kathēkon,* usually translated as "duty," appears to mean "what it would be suitable or fitting to do." At any rate, the supreme duty is to live "in accord with nature," but it is not clear what that entails or how, if at all, one could avoid living in accord with nature. Particular maxims have to do with ways of avoiding anxiety and frustration, a goal that Kant would have regarded as morally unworthy. The one genuine point of contact, and also the most interesting contribution of Stoic thought, is the idea that morality transcends national boundaries and class distinctions. The cosmopolitanism of the Stoics marks an advance over the views of Plato and Aristotle, both of whom thought that the demands of morality can be satisfied without taking any account of the claims of barbarians, slaves, or foreigners. On the other hand, the Stoic one-world concept ought, perhaps, to be seen not so much as a moral ideal than as an implicit recognition of the changes brought about by the conquests of Alexander and, in later writings, as an aspect of the ideology of Roman imperialism.

Theological ethics attaches importance to the concept of duty, and, in this context, what is meant is, unlike Kantian or Stoic duty, something parallel to the ordinary notion. To be a believer or a member of a congregation is to hold a particular office, often one that is defined by clearly formulated rules of conduct and ritual observance. In some religions the faithful are told that they are in some sense children of God, and to the extent that this belief is taken seriously, a set of quasi-filial duties with respect to the deity will come to seem important. Kant, despite his Pietistic background, was clearly opposed to such a view. It is crucial to his doctrine that men should regard themselves and others as adults rather than as hapless children.

Anticipations of particular Kantian theses can be made out in a number of earlier writers: Richard Cumberland, Ralph Cudworth, Samuel Clarke, and Richard Price maintained (in opposition to Thomas Hobbes) that moral duty is based on self-evident axioms and that the requirements of duty are universally binding. Jean-Jacques Rousseau had much to say about conscience, which he regarded as a sort of inner voice—one that speaks with unique authority on questions of duty. David Hume explicitly remarked on the logical gap between the concept of what is done and the concept of what ought to be done. Nonetheless, it is not clear that anyone before Kant succeeded in holding in focus the idea of a morality that is not, in some indirect way, dependent on considerations of prudence.

In his paper "Does Moral Philosophy Rest on a Mistake?" (1912), H. A. Prichard argued that traditional ethics (for example, the doctrines of Plato, Aristotle, Hume, Jeremy Bentham, and J. S. Mill) goes astray in trying to work out some general answer to the question of why it is reasonable or worthwhile to do one's duty. Prichard's point is that the question itself is the result of a confusion. That something is a duty is (or may be) a sufficient reason for doing that thing, and *if* it is, then no further reason is called for. If Prichard's historical thesis is right, and it seems quite plausible, then there is a sense in which Kantian doctrine and common sense agree and are jointly opposed to traditional ethics. Ordinary duties are not hierarchically ordered under a supreme moral principle; nor do the claims of duty (individually or collectively) provide a unique determination of morally right action. Nonetheless, and despite their untidy array, ordinary duties are "unconditioned" in that they provide us with reasons for acting such that if the reasons are accepted, there is no need for, indeed no room for, further justification.

See also Aristotle; Bentham, Jeremy; Bradley, Francis Herbert; Cicero, Marcus Tullius; Clarke, Samuel; Cudworth, Ralph; Cumberland, Richard; Distant Peoples and Future Generations; Hobbes, Thomas; Hume, David; Kant, Immanuel; Kantian Ethics; Mill, John Stuart; Modal Logic; Plato; Price, Richard; Responsibility, Moral and Legal; Rights; Ross, William David; Rousseau, Jean-Jacques; Stoicism.

Bibliography

For the Stoic conception of duty, see Cicero's *De Oficiis,* which has been translated by H. M. Poteat as *On Duties* (Chicago, 1950), and W. J. Oates, ed., *The Stoic and Epicurean Philosophers* (New York, 1940).

See also Immanuel Kant, *Critique of Practical Reason and Other Writings in Moral Philosophy,* translated by L. W. Beck (Chicago, 1949), which contains *Foundations of the Metaphysics of Morals*; H. A. Prichard, "Does Moral Philosophy Rest on a Mistake?" in *Mind* 21 (1912): 121–152, also included in his *Moral Obligations* (Oxford, 1949), pp. 1–17; G. E. Moore, *Ethics* (London: Williams and Norgate, 1912), Ch. IV; G. E. Hughes, "Motive and Duty," in *Mind* 53 (1944): 314–331; W. D. Ross, *Foundations of Ethics* (Oxford, 1947), pp. 14–27; and W. K. Frankena, "Obligation and Ability," in *Philosophical Analysis,* edited by Max Black (Ithaca, NY: Cornell University Press, 1950).

Mary Mothersill (1967)

DWORKIN, RONALD

(1931–)

Ronald Dworkin, born in Worcester, Massachusetts, has been a leading participant in debates central to legal and political philosophy in the wake of the 1960s. After graduating from Harvard Law School and clerking for legendary federal judge Learned Hand, he held a number of distinguished faculty appointments in the United States and England, including Professor of Jurisprudence at the University of Oxford.

During the early portion of Dworkin's career, social movements such as those connected with civil rights, women's equality, the environment, and the Vietnam War, confronted philosophers with the task of reassessing liberalism. Influential radicals, including Herbert Marcuse, held liberalism responsible for the injustices of the era. However, other philosophers sought to reformulate and defend liberal ideas. John Rawls was the leading figure in the reformulation of liberalism, but next to Rawls, no thinker writing in English has played a larger role than Dworkin. His work is informed by the conviction that the moral task of citizens and public officials is not to jettison liberal democracy but to make their society a more faithful realization of liberal ideals.

Dworkin argues that legal reasoning has an ineliminable moral dimension and defends a form of liberalism that regards the right to equality as the sovereign political principle. His argument about legal reasoning rejects the positivist view that the existence of laws depends ultimately on social facts that can be ascertained without resort to moral judgments. It also opposes those natural law theories that hold the legal validity of a norm to depend on its consistency with substantive justice. Dworkin's defense of liberalism rejects the radical view that liberal principles are complicit in the perpetuation of oppression. It opposes as well the conservative view that liberal ideas have a corrupting influence on society. Writing as a public intellectual, Dworkin has contributed to controversies over civil disobedience, free speech, campaign financing, affirmative action, physician-assisted suicide, abortion, and civil liberties. He has also addressed debates over constitutional interpretation in the United States, rejecting theories resting on the framer's intent and advocating interpretations informed by moral principles that protect individual rights.

The most widely discussed thesis in jurisprudence for a decade was Dworkin's *rights thesis*, defended in *Taking Rights Seriously* (1977). The thesis holds that, in almost all legal cases, one side has the legal right to win. Dworkin criticizes H. L. A. Hart's positivist classic *The Concept of Law* (1961) for claiming that in hard cases, where legal rules do not determine which side should win, judges have discretion to render decisions as social utility dictates. Dworkin argues that Hart neglects the moral principles that underlie legal rules and constitute part of the law. Such principles help to determine the legal rights of persons whereas rights function as "trumps" that an individual holds against the government and its efforts to promote utility or some other societal good at the individual's expense. Dworkin imagines a superhuman judge "Hercules," who knows all the best moral principles underlying the settled law. Though more limited in their cognitive capacities, human judges should, and characteristically do, seek out those principles that bear on the cases they decide.

The most comprehensive statement of Dworkin's legal philosophy is in *Law's Empire* (1986). The work of judges is presented as continuous with that of legal philosophers. Both involve "constructive interpretation," a way of understanding an object in light of the best purpose it can be seen to serve. Adjudication gives a constructive interpretation of the laws within the court's jurisdiction, with the aim of deciding cases under the law. Legal philosophy gives a constructive interpretation of law more generally, with the aim of determining the strongest justification for the existence of law. Dworkin argues that the strongest justification is that law serves the ideal of integrity: treating citizens according to a single, coherent scheme of moral principles.

Notable critics of Dworkin's legal philosophy include Joseph Raz and Jules Coleman, who counter his criticisms of positivism and develop their own versions of the positivist view. Although Dworkin has proved unable to dislodge positivism from its dominant position, it is widely agreed that his work has advanced legal philosophy by forcing positivists and natural lawyers alike to refine and elaborate their views.

Dworkin's political philosophy forms an integrated whole with his legal thought. He argues that a political community cannot have legitimate authority over its members unless it treats each of them with equal concern. He elaborates by developing a theory of distributive justice in which citizens have a right to an equally valuable bundle of resources with which to pursue their own conception of the good. The choices individuals make in utilizing their resources affect the value of their holdings. Resulting economic inequalities are justifiable, as they derive from the person's own values and tastes. Dworkin argues that a suitably regulated market is indispensable

for justice because markets provide the only acceptable measure of the value of the resources a person holds, namely, the opportunity costs of denying those resources to others.

Dworkin contends that equality demands that individuals be respected in the exercise of their liberties, including liberties to obtain sexually explicit materials, engage in homosexual relations, and voice publicly fascist and racist attitudes. He rejects the view that equality and liberty stand in tension. Equality is the ground for civil and political liberties; it is not a competing value. Equal respect entails that government must remain substantially neutral on questions concerning what makes a good life, leaving it up to individuals to decide such matters for themselves.

Raz formulates a liberal alternative to Dworkin, arguing that government fosters freedom not by remaining neutral on questions of the good but by supporting a social environment in which a wide variety of models of a good life are visible. John Finnis and Robert George criticize Dworkin's view of equality and liberty by invoking an account of basic human goods that derives from the conservative tradition of natural law theory. Other important critics include Rae Langton and Catharine MacKinnon, who mount feminist criticisms of Dworkin's position on pornography. G.A. Cohen rejects his theory of equal resources, arguing that market outcomes are morally arbitrary. Most sweepingly, Roberto Unger criticizes Dworkin's philosophy for rationalizing the shortcomings of liberal democracy and glossing over the need for radical changes in existing forms of democracy and the market.

Dworkin has addressed many criticisms of his work, refining and revising his views in the process. His lasting contribution is to have developed a liberal account of law and politics that is original, nuanced, and systematic.

See also Philosophy of Law; Political Philosophy, History of; Rawls, John.

Bibliography

Hart, H. L. A. *The Concept of Law*. Oxford: Clarendon Press, 1961.

WORKS ABOUT DWORKIN

Burley, Justine, ed. *Dworkin and His Critics: With Replies by Dworkin*. Malden, MA: Blackwell, 2004.

Cohen, Marshall, ed. *Ronald Dworkin and Contemporary Jurisprudence*. Totowa, NJ: Rowman and Allenheld, 1983.

WORKS BY DWORKIN

Taking Rights Seriously. Cambridge, MA: Harvard University Press, 1977.

A Matter of Principle. Cambridge, MA: Harvard University Press, 1985.

Law's Empire. Cambridge, MA: Harvard University Press, 1986.

Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom. New York: Knopf, 1993.

Freedom's Law: The Moral Reading of the American Constitution. Cambridge, MA: Harvard University Press, 1996.

Sovereign Virtue: The Theory and Practice of Equality. Cambridge, MA: Harvard University Press, 2000.

Andrew Altman (2005)

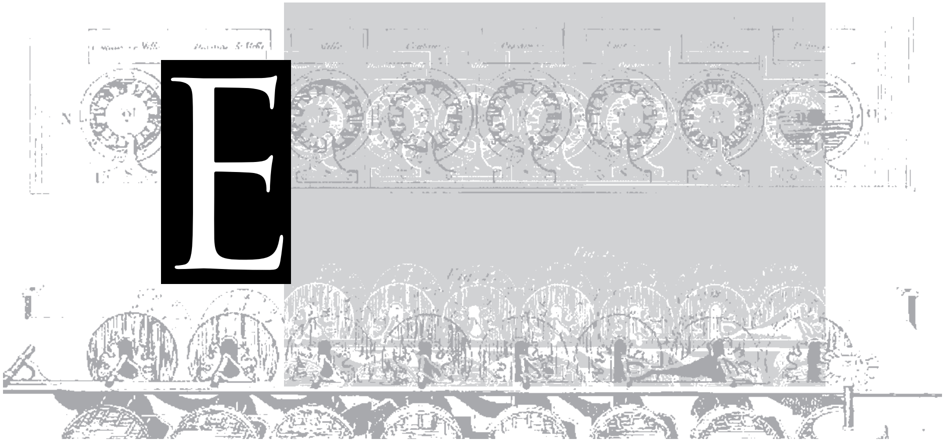

E

EARMAN, JOHN
(1942–)

John Earman is an American philosopher and professor of history and philosophy of science at the University of Pittsburgh. He is perhaps best known for contributions to the history and foundations of modern physics—especially space-time theories, and often with the question of determinism in view—and confirmation theory.

Earman completed his PhD at Princeton in 1968, under the direction of Carl G. Hempel. After brief appointments at University of California, Los Angeles, and the Rockefeller University, where he enjoyed tenure for a year before its philosophy department was disbanded in 1973, Earman spent twelve years at the University of Minnesota, where he was promoted to full professor in 1974. He moved to Pittsburgh in 1985.

SPACETIME AND DETERMINISM

A theme of Earman's earliest publications is that progress can be made on perennial philosophical problems by bringing modern physics and mathematics, thoroughly and properly understood, to bear on them. Through the late 1960s the reigning orthodoxy in the philosophy of space and time held the dispute between absolute and relational accounts to have been settled conclusively, and in favor of the relationalist, by the advent of relativity theory. Presenting Albert Einstein's theory in the language of differential geometry—the mode of presentation favored by mathematical physicists—Earman argues persuasively, in "Who's Afraid of Absolute Space" (1970), that traditional terms of debate are hopelessly ambiguous. The scientifically respectable disambiguations he devises enable him to turn orthodoxy on its head. Isaac Newton's arguments for absolute space succeed, Earman contends, and absolute kinematic quantities abound in relativistic space-times. Along with the contributions of Howard Stein, Michael Friedman, and Larry Sklar, this work helped drag the philosophy of space and time into its modern era.

As Earman aged, he aimed less to resolve perennial philosophical problems than to deploy them as a sort of dragnet in which to ensnare important issues in the foundations of physics. The philosophical problems typically emerge from this deployment considerably complicated. *A Primer on Determinism* (1986), which won the Lakatos Prize in 1989, recasts the question of whether the world is deterministic as a question about whether there are other physically possible worlds—that is, other worlds obeying the same natural laws as the actual world does—that

agree with the actual world at some times but not others. Subsequent chapters subject the doctrine of determinism to trial by a variety of prominent theories. Surprising verdicts are reached: Earman declares classical Newtonian mechanics, the physics that inspired Pierre Simon de Laplace's chilling statement of determinism, indeterministic. Admitting infinite signal velocities, classical physics admits as well possible worlds that agree up to a time *t*, but differ afterward due to the unheralded arrival at *t* in one world but not the other of "space invaders" that have traveled infinitely fast from spatial infinity. More often, the jury is hung and the fate of determinism is entangled with "sticky interpretations problems [that] resist narrowly scientific solutions" (p. 197). "We can't just read off the lessons for determinism from various branches of physics, for the implications we read will depend upon judgments about the adequacy of physical theories, and those judgments will depend in turn on our views about determinism" (p. 78).

In *World Enough and Space-Time: Absolute versus Relational Theories of Space and Time* (1989) determinism probes the doctrine of absolute space Earman so energetically rescued from ill repute in the 1970s. Space-time substantivalism—the thesis that spatiotemporal relations between bodies are "parasitic on relations among a substratum of … spacetime points that underlie events" (p. 12)—is a modernization of the doctrine with an impeccable pedigree: Newton himself, Earman argues, was a substantivalist. But Earman is not. He takes the lesson of Einstein's hole argument to be that anyone committed to substantivalism about general relativistic space-times is also committed to indeterminism (compare Earman and Norton 1987). On the principle that "if determinism fails, it should fail for a reason of physics" (Earman 1989, p. 181), Earman rejects substantivalism. He does not thereby embrace relationalism: "[M]y tentative conclusion is that a correct account of space and time may lie outside the ambit of the traditional absolute-relational controversy" (p. 4). The sample tertium quid he sketches—an interpretation mediated by Leibniz algebras—was later shown itself to imply indeterminism (Rynasiewicz 1992).

The hole argument turns on the fact that if one of any pair of space-times related by a *hole diffeomorphism*—roughly, a map between space-times that is the identity outside a region *h* (the hole) but is nontrivial inside that region—corresponds to a world possible according to general relativity, then so does the other. Supposing that substantivalists must take space-times related by a hole diffeomorphism to differ in properties assigned space-time points inside *h*, Earman and Norton (1987) conclude that substantivalists must take there to be worlds possible according to general relativity that agree at some times but not others. The hole argument launched a thousand responses. Many philosophers took exception to its accounts of reference to, or criteria for transworld identity of, space-time points, while some physicists credited the hole argument for raising interpretive questions pertinent to ongoing efforts to quantize gravity.

One way determinism might fail for a reason of (general relativistic) physics arises from space-time singularities. Space-time singularities are, roughly speaking, regions of space-time at which Einstein's equations become mathematically ill defined, so that imposing those equations is insufficient to prevent determinism-destroying emanations—Earman seems particularly worried about televisions playing Richard Nixon's "Checkers" speech—from those regions. *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes* (1995) discusses singularities and other eponymous acausalities. The book's topics—chronology horizons, inflationary cosmologies, and cosmic censorship—familiar to working physicists but less evidenced in philosophy journals, reflects a tendency, appearing in the mid-1980s and accelerating thereafter, for Earman to draw his problem agenda directly from contemporary physics.

BAYESIAN CONFIRMATION THEORY

The first half of Earman's *Bayes or Bust?: A Critical Examination of Bayesian Confirmation Theory* (1992) skillfully surveys the grounds supporting Bayesian confirmation theory: for example, the perspicuity of the analyses it offers of other accounts of confirmation, and its ability to provide some sort of solution to the Quine-Duhem problem and the new riddles of induction. The second half ruthlessly undermines those grounds, for example, it finds Bayesianism incapable of addressing the problem of old evidence or accommodating changes of belief in so-called scientific revolutions. Characteristically, Earman considers the point of the exercise not to reach a verdict on Bayesianism—in the introduction he admits to a diurnal oscillation between being an "imperialistic apostle" of Bayesianism and doubting its very viability—but to uncover worthwhile problems in the course of weighing the evidence.

These problems include historical ones—how to understand Thomas Bayes's essay in the context of eighteenth-century work on probability, for example. A

concern, and a knack, for matters historical informs much of Earman's work. *Hume's Abject Failure: The Argument against Miracles* (2000), his most recent book, situates David Hume's argument against miracles in a historical context. That Bayes and Hume were contemporaries licenses Earman to develop Bayesian analyses of Hume's central contentions and the notions (e.g., multiple witnessing) they involve. Although *Hume's Abject Failure* was not universally well received by Hume scholars or philosophers of religion, some of whom charged it with insensitivity to Hume's broader epistemology and with harboring too many equations, the work accomplishes its self-described aim: "not simply to bash Hume … but also to indicate how, given the proper tools, some advance can be made on these problems" (p. 4).

See also Bayes, Bayes's Theorem, Bayesian Approach to Philosophy of Science; Determinism, A Historical Survey; Space.

Bibliography

Earman, John, and J. Norton. "What Price Space-Time Substantivalism? The Hole Story." *British Journal for the Philosophy of Science* 38 (1987): 512–525

Rynasiewicz, Robert. "Rings, Holes, and Substantivalism: On the Program of Leibniz Algebras." *Philosophy of Science* 59 (1992): 572–589.

WORKS BY EARMAN

"Who's Afraid of Absolute Space?" *Australasian Journal of Philosophy* 48 (3) (1970): 287–319.

A Primer on Determinism. Dordrecht, Netherlands: D. Reidel, 1986.

World Enough and Space-Time: Absolute versus Relational Theories of Space and Time. Cambridge, MA: MIT Press, 1989.

Bayes or Bust?: A Critical Examination of Bayesian Confirmation Theory. Cambridge, MA: MIT Press, 1992.

Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes. New York: Oxford University Press, 1995.

Hume's Abject Failure: The Argument against Miracles. New York: Oxford University Press, 2000.

Laura Ruetsche (2005)

EBERHARD, JOHANN AUGUST
(1739–1809)

Johann August Eberhard, the German theologian and "popular philosopher," was born in Halberstadt. He studied theology at Halle, and became a preacher at Halberstadt in 1763 and at Charlottenburg in 1774. In 1778 Frederick II of Prussia appointed him professor of theology at Halle. Eberhard became a member of the Berlin Academy in 1786 and a privy councilor in 1805. He wrote on theology, epistemology, ethics, aesthetics, philology, and the history of philosophy.

Eberhard received a Wolffian education, but, under the influence of Moses Mendelssohn and Christian Friedrich Nicolai, he soon developed a personal point of view. As a popular philosopher, Eberhard was averse to abstract speculation and interested in natural theology, psychology, ethics, and aesthetics. He opposed enthusiasm, sentimentalism, and occultism, and favored the empirical approach.

In his *Neue Apologie des Socrates* (New Apology of Socrates; 2 vols., Berlin, 1772–1778) Eberhard denied that salvation depended on revelation, and asserted that there is no original sin and that a heathen could go to heaven. He rejected eternal punishment as a contradiction of its aim—the moral improvement of the sinner.

Eberhard's *Allgemeine Theorie des Denkens und Empfindens* (General theory of thinking and feeling; Berlin, 1776) was dominated by the thought of John Locke, and by Gottfried Wilhelm Leibniz's *Nouveaux Essais.* Like Immanuel Kant and Johann Nicolaus Tetens, Eberhard vindicated sensation against the earlier tendency to stress reason; and like Kant, Tetens, and Johann Heinrich Lambert, he developed a thoroughgoing phenomenalism. He held that sensation is passive and supported Locke's view that all ideas derive from sensation. He claimed that sensing is a transition from thinking to acting.

Eberhard held that Beauty is not an objective characteristic of things, but an adequacy of the object to the representative power of the subject (a view he called—as Kant did later—"subjective finalism"). Beauty excites this activity, and the aim of art is therefore the awakening of pleasurable passions (a doctrine rejected by Kant and later German aestheticians). The first appearance of aesthetic activity in man is represented, according to Eberhard, in children's play (a foreshadowing of Friedrich Schiller's aesthetics of play).

Eberhard, as editor of the *Philosophisches Magazin* from 1788 to 1791 and of the *Philosophische Archiv* from 1792 to 1795, published a large number of articles critical of Kant's *Kritik der reinen Vernunft,* most of them written by himself. He claimed that Kant's views were entirely derived from Leibniz, and that they were only a special kind of dogmatism. Kant answered Eberhard in his *Ueber*

eine Entdeckung, nach der alle neue Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll (Königsberg, 1790). It was one of the few times Kant deigned to answer unjustifiable criticism.

Bibliography

ADDITIONAL WORKS BY EBERHARD

Sittenlehre der Vernunft. Berlin, 1781.

Theorie der schonen Künste und Wissenschaften. Berlin, 1783.

Vermischte Schriften. 2 vols. Halle, 1784–1788.

Allgemeine Geschichte der Philosophie. Berlin, 1788.

Handbuch der Aesthetik. 4 vols. Halle: Hemmerde und Schwetschke, 1803–1805.

WORKS ON EBERHARD

Draeger, G. *J. A. Eberhards Psychologie und Aesthetik.* Halle, 1915.

Ferber, E. O. *Der philosophische Streit zwischen I. Kant und J. A. Eberhard.* Giessen, 1884.

Lungwitz, K. *Die Religionsphilosophie Eberhards.* Erlangen, 1911.

Nicolai, C. F. *Gedächtnisschrift auf J. A. Eberhard.* Berlin, 1810.

Giorgio Tonelli (1967)

ECKHART, MEISTER

(c. 1260–1327/1328)

Meister Eckhart, the German mystic, was born Johannes Eckhart at Hochheim in Thuringia. After entering the Dominican order at an early age, he pursued higher studies at Cologne and Paris. He became successively provincial prior of the Dominican order of Saxony, vicar-general of Bohemia, and superior-general for the whole of Germany (in 1312). During the last part of his life Eckhart became involved in charges of heresy. In 1329, twenty-eight of his propositions were condemned by Pope John XXII, eleven as rash and the remainder as heretical. Nevertheless, Eckhart was to have a lasting influence upon medieval mysticism.

Eckhart's account of God and the universe depended not only on theology and metaphysical speculation but also on his interpretation of mystical experience. Thus, he distinguished between *Deus* or God, as found in the three Persons of the Trinity, and *Deitas* or the Godhead, which is the Ground of God but is indescribable. The Godhead, through an eternal process, manifests itself as the Persons. In the same way, Eckhart distinguished between faculties of the soul, such as memory, and the *Grund* or "ground" of the soul (also called the *Fünklein, scintilla* or "spark"). By contemplation it is possible to attain to this *Grund,* leaving aside the discursive and imaginative activities that normally characterize conscious life. In doing this, one gains unity with the Godhead. Although Eckhart gave some sort of explanation for the ineffability of the Godhead (namely, that it is a pure unity and thus not describable), the main motive for his doctrine lay in a feature of mystical experience—that it involves a mental state not describable in terms of thoughts or images.

The need to give an account of contemplative knowledge led Eckhart to evolve a complex psychology. The soul operates at the lowest level, through the body; thus it has powers of digestion, assimilation, and sensation. At a higher level the soul functions through the powers of anger, desire, and the lower intellect (the *sensus communis* or "common sense," which combines what is given through the various senses in perception). At a third level the soul works through memory, will, and the higher intellect. At the fourth level it is possible in principle to know things in total abstraction, that is, as pure forms, which is therefore to know them as they preexist in God's intellect. Finally, the spark of the soul can possess a kind of knowledge in which God is known as he is.

In the development of these ideas, Eckhart certainly spoke in ways which might have offended his more orthodox contemporaries. The notion of the spark within the soul seemed to imply that the soul is uncreated. The notion of God's birth within the soul, through mystical experience, seemed to present the sacraments of the church as mere means of preparing for such experience, rather than as efficacious in themselves. Likewise, Eckhart's language of deification could easily have been construed to mean that the historical Christ has only an exemplary and symbolic value. Eckhart's teaching that God creates the world in the same "eternal now" in which the emanation of the divine Persons from the Godhead takes place could be understood as implying the eternity of the world—a doctrine that conflicts with the literal sense of biblical revelation. His statement that all creatures are a "mere nothing" could be held to imply a kind of monism. Recently, however, among Catholic historians of philosophy an attempt has been made to show that his theology is less unorthodox than the above doctrines might suggest, and as a Dominican, Eckhart certainly employed the language of Thomism.

This recent discussion serves to underline the degree to which Eckhart permitted changes and inconsistencies in the formulation of his ideas. Thus, at one time he held that the divine essence is *intelligere,* or understanding (a thesis original to Eckhart, and one which reinforced the doctrine of similarity of the soul to God), and only secondarily is God *esse,* or being. Later, however, he held, in

accordance with Thomist doctrine, that God's essence is *esse.* Various other fluidities and antinomies can be detected in Eckhart's thought; these were partly caused by the shifting way in which he used key terms. For example, he asserted that God is above being and yet also, that he is being. The first use of "being" could be taken to refer to finite existence; the second use could be taken in a Thomistic sense. At times he spoke of God as both Godhead and God, and at other times he spoke of God as distinguished from the Godhead.

Although on occasion Eckhart used the term *emanation* to describe the creation of the world, he in fact adhered to an orthodox account of creation out of nothing. But he stressed the continuous creativity of God, and in this and other respects he was influenced by Augustine. Even though his language about creation could be misinterpreted to imply the eternity of the cosmos, Eckhart was at pains to evolve a two-level theory of time. In a sense all events are simultaneous for God, who is timelessly eternal (so that to speak of a temporal gap between the procession of the Trinity and the creation of the world makes no sense). Temporal concepts, however, are properly applied within the created order, and therefore the creation can be dated retrospectively. Eckhart's two-level theory of time corresponded to his two-level theory of truth. The truths that we assert are limited and partial (or, as Eckhart asserted, there is untruth in them), but there is an absolute truth which can be realized existentially, namely, the pure being of the Godhead.

The general shape of Eckhart's beliefs, if we except his doctrines of the Godhead and of the soul, was fully in accord with contemporary belief (for example, in regard to angels and purgatory). What made his sermons and teachings popular was the way in which he reiterated the need to penetrate beneath the externals of religion, while his free use of homely, striking, and sometimes paradoxical examples and similes effectively conveyed his message.

There is a remarkable parallel between some of Eckhart's central ideas and the doctrines of the Indian theologian Śankara (d. c. 820)—a parallel first expounded by Rudolf Otto. In Śankara's system, too, there is a distinction between the Absolute and God conceived as personal and a similar claim that the divine can be found within the soul. The comparison may give a clue to the reason for the shape of Eckhart's teachings. It certainly suggests that there are experiential reasons for this kind of doctrine, even though they may be complicated reasons. They seem to be as follows. The experience of the introvertive mystic includes a state of consciousness in which there is both a sense of illumination and an absence of distinction between subject and object; that is, the contemplative is not having an experience like that of ordinary perception, where the thing perceived can be distinguished from the percipient. Consequently, if the mystic connects his experience with God (whom he believes in for independent reasons), he may be inclined to speak of merging with God. But since his experience is without differentiation and since the notion of God—and especially that of a Trinitarian God—includes the idea that he has attributes, it is not unnatural, although it appears unorthodox, to treat the entity experienced by the mystic as being "beyond" God conceived personally.

Indeed, Eckhart maintained that the true aristocrat (that is, the spark or ground of the soul) reaches beyond God, to the Godhead. It is likewise natural, in the Christian context in which Eckhart lived, to interpret this simple undifferentiated unity found in the Godhead as being the basis out of which the Persons of the Trinity proceed. In this way mystical experience, for Eckhart, was connected with the God of ordinary religion. Nevertheless, Eckhart endeavored to express himself in accordance with orthodox belief, despite the difficulties that he found in trying to do justice both to his experience and to the ordinary language of theism. Certainly, he did not seriously intend to deny orthodoxy.

Despite the papal condemnation of some of his propositions, Eckhart had a wide influence. Johannes Tauler, Heinrich Suso, Jan van Ruysbroeck, and the group known as the Friends of God were in different ways indebted to his teachings and example.

See also Augustine, St.; Mysticism, History of; Mysticism, Nature and Assessment of; Otto, Rudolf; Ruysbroeck, Jan van; Suso, Heinrich; Tauler, Johannes; Thomism.

Bibliography

WORKS BY ECKHART

Meister Eckhart. 4th ed, edited by F. Pfeiffer. Göttingen, 1924. Includes sermons, treatises, and fragments.

Meister Eckhart, a Modern Translation. Translated and edited by R. B. Blakney. New York, 1957. Includes bibliographical notes and the more important writings.

Selected Treatises and Sermons, edited by J. M. Clark and J. V. Skinner. London, 1958.

WORKS ON ECKHART

Clark, J. M. *The Great German Mystics.* Oxford: Blackwell, 1949. A good introduction.

Clark, J. M. *Meister Eckhart.* London: Faber and Faber, 1958. A biography.

Gilson, Étienne. *Christian Philosophy in the Middle Ages.* New York: Random House, 1955.
Otto, Rudolf. *Mysticism East and West.* London: Macmillan, 1932.
Wulf, Maurice de. *Histoire de la philosophic médiévale.* Vol. III. Louvain, 1947.

Ninian Smart (1967)

ECONOMICS, PHILOSOPHY OF

See *Philosophy of Economics*

ECONOMICS AND ETHICAL NEUTRALITY

See *Ethics and Economics*

EDDINGTON, ARTHUR STANLEY
(1882–1944)

Arthur Stanley Eddington was an English astronomer who was educated at Owens College, Manchester, and Trinity College, Cambridge, where he was Plumian professor of astronomy from 1913 to 1944. He never married, was socially rather diffident, and lived the quiet life of a Cambridge academic. He was elected a fellow of the Royal Society in 1914 and was knighted in 1930.

Eddington was one of the most brilliant theoreticians of his day, possessing an outstanding ability to survey complex and highly ramified subjects as wholes. His report to the Physical Society (1918) on the general theory of relativity, expanded into *The Mathematical Theory of Relativity* (London, 1923), contained important original contributions to the theory. Eddington's discovery of the mass-luminosity relation in stars and his explanation of white dwarf stars, which made possible the modern theory of stellar evolution, were published in *The Internal Constitution of the Stars* (London, 1926). These two books are considered to be his most substantial contributions to physics and astronomy. His interpretation of relativity theories led him to a belief in the profound importance of epistemology for physics. At first in semipopular books on modern physics—*Nature of the Physical World* (London, 1928) and *New Pathways in Science* (London, 1935) being the most important—Eddington argued for the view that physics could be almost entirely based upon investigations into the nature of sensation and measurement. A more elaborate and purely philosophical defense of his view was given in *The Philosophy of Physical Science* (London, 1939). Formal attempts actually to produce physics as derived in this way were presented in *Relativity Theory of Protons and Electrons* (London, 1936) and *Fundamental Theory* (London, 1946), published posthumously.

Eddington's real contributions to philosophy, if any, lie in his work on the epistemology of physics. However, he also defended idealism and mysticism, and he claimed that the indeterminacy of quantum physics solved the traditional philosophical problem of free will versus determinism in favor of free will. Particularly in his semipopular writings, Eddington was betrayed into philosophical excesses and, at times, gross confusion by a play of analogy and paradox, which, while part of his equipment as an immensely entertaining and brilliant writer, also served his love of mystery and obscurantism.

SELECTIVE SUBJECTIVISM

Eddington gave to his epistemological view the two names "selective subjectivism" and "structuralism." He accepted the causal theory of perception, and with this theory Eddington's own system stands or falls. From this theory it follows, first, that we know directly only the contents of our own consciousness (sense data) and, second, that these contents cannot be claimed to resemble elements of the objective world in any qualitative way. Our sensory apparatus selects from objective reality what we are able to observe and what is therefore the material for physical knowledge, just as, to use Eddington's own analogy, a net of a certain size mesh selects fish only of a size greater than the mesh. Just as we could generalize, prior to examining any catch of fish, about the size of fish the net would yield, so we can generalize in physics prior to the results of observation, merely by reflecting upon observational procedure, especially metrical procedures.

Despite distortions, mostly qualitative, in the picture that our senses thrust upon us, we may conclude that the picture has a structure in common with the unknowables that stimulate the senses. We notice patterns of recurrence in sensation, and it is the task of physics to elaborate the structure of these patterns. In particular, the structure of pointer-reading observations should be studied, since pointer readings—being merely observed coincidences—are minimally corrupted by the qualitative veils cast by our senses. However, Eddington denied the

pointer readings directly represent anything objectively real.

APRIORISM

Like Immanuel Kant, to whose system Eddington admitted that his own was distantly similar, he claimed that knowledge must conform to certain primitive rational patterns if it is to be intelligible. One of these forms of thought is that we believe in the existence of minds other than our own. The recognition of a common structure in the experience of many minds leads to a belief in an objective reality independent of these minds. There is no primitive belief in an objective reality. This route to the existence of an external world is an unobtrusive but significant part of Eddington's idealistic metaphysics.

Using the notion of structure as defined in the mathematical theory of groups, Eddington was able, out of highly generalized material from epistemology (for example, the claim that only relations between things are observable) and from the forms of thought, to build quite intricate group structures, for example, the structure found in Paul Dirac's mathematical specification of an elementary particle in an elementary state giving charge and spin. In addition to this a priori derivation of the formal structure of laws, Eddington also exploited the theory of groups in deducing a priori the basic natural constants, such as the gravitational constant and the fine structure constant, from various features of the group structure of the type of mathematics employed. In this, he compared himself with Archimedes, who deduced the nature of π from the axiom of Euclid, whereas previous determinations of its nature had relied upon merely empirical methods.

On this basis Eddington claimed that the mind fits nature into a pattern determined by the nature of the mind itself; that the discoveries made by the physicist are just what his sensory, intellectual, and metrical processes dictate that he shall find.

It is difficult not to share the general view that Eddington vastly overstated the extent to which convention enters into theory construction. Extensive criticism in this entry without more extensive elaboration of the complexities of his group structure derivations would be unjust. Some brief comments must suffice.

Eddington's view was that observation was required only for the purpose of identifying, on the one hand, the elements of the group constructed by pure mathematics with, on the other hand, the theoretical terms of, say, electromagnetism. It is far from clear where he thought the complete theoretical structure then stood from the point of view of its a priori status. If such "identification" demands that it be fully observed that the electromagnetic field is properly (that is, truly) described by Maxwell's equations, which have the group structure in question, then Eddington was requiring "observation" to add a very great deal more than he seems to have been prepared to admit.

Eddington fell into confusion that illustrates well his mistakes in general. This was his claim that the basis of the special theory of relativity may be deduced a priori because it depends on the fact that simultaneity of events at a distance from each other is not observable, that is, that it depends upon an epistemological fact. It is true that to decide a question about the simultaneity of spatially separated events, one must make assumptions as to the speed of the signals that inform one that the events have occurred. And it is also true that in the last resort these assumptions could be checked only if one could decide independently on the simultaneity of events spatially distant from each other. But this epistemological circularity is an insufficient basis for relativity theory. Moreover, further contingent facts, not deducible a priori (for example, the fact that in any inertial system light takes the same time round any closed paths of the same length, whatever their orientation) are required. Eddington claimed that the result of the Michelson-Morley experiment could have been foreseen on a purely epistemological basis. It seems quite clear that he was wrong.

IDEALISM

"To put the conclusion crudely—the stuff of the world is mind-stuff," Eddington wrote in *Nature of the Physical World.* The idealist conclusion was not integral to his epistemology but was based on two main arguments.

The first derives directly from current physical theory. Briefly, mechanical theories of the ether and of the behavior of fundamental particles have been discarded in both relativity and quantum physics. From this Eddington inferred that a materialistic metaphysics was outmoded and that, in consequence—the disjunction of materialism or idealism being assumed exhaustive—an idealistic metaphysics is required.

The second and more interesting argument was based on Eddington's epistemology and may be regarded as consisting of two parts. First, all we know of the objective world is its structure, and the structure of the objective world is precisely mirrored in our own consciousness. We therefore have no reason to doubt that the objective world, too, is "mind-stuff." Dualistic meta-

physics, then, cannot be evidentially supported. (The conclusion appears to be a valid deduction from its premises.)

But, second, not only can we not know that the objective world is nonmentalistic, we also cannot intelligibly suppose that it could be material. To conceive of a dualism entails attributing material properties to the objective world. However, this presupposes that we could observe that the objective world has material properties. But this is absurd, for whatever is observed must ultimately be the content of our own consciousness and, consequently, nonmaterial. This last argument confuses, among other things, the supposition that the objective world has certain properties with the supposition of our observing that it has them.

See also Determinism and Freedom; Epistemology, History of; Idealism; Mysticism, History of; Kant, Immanuel; Popular Arguments for the Existence of God; Stebbing, Lizzie Susan; Subjectivist Epistemology.

Bibliography

ADDITIONAL WORKS BY EDDINGTON

Space, Time, and Gravitation. Cambridge, U.K.: Cambridge University Press, 1921.

Science and the Unseen World. London: Allen and Unwin, 1929.

The Expanding Universe. Cambridge, U.K.: Cambridge University Press, 1933.

WORKS ON EDDINGTON

Dingle, Herbert. *The Sources of Eddington's Philosophy.* Cambridge, U.K.: Cambridge University Press, 1954.

Stebbing, S. *Philosophy and the Physicists.* London: Metheun, 1937.

Whittaker, E. T. *Eddington's Principle in the Philosophy of Science.* Cambridge, U.K.: Cambridge University Press, 1951.

G. C. Nerlich (1967)

EDUCATION, PHILOSOPHY OF

See *Philosophy of Education, History of*

EDWARDS, JONATHAN
(1703–1758)

Jonathan Edwards, the Puritan theologian and philosopher, was born in East Windsor, Connecticut. He was the only son of Timothy Edwards, the pastor of the Congregational Church at East Windsor; his mother was the daughter of Solomon Stoddard, pastor at Northampton, Massachusetts. About the age of twelve or thirteen he wrote several essays in natural science that reveal remarkable powers of observation and deduction. "Of Insects" describes the habits of spiders. Another essay, on the rainbow and colors, shows an acquaintance with Isaac Newton's *Opticks*. Around the same time Edwards wrote a short demonstration of the immateriality of the soul. These writings are the work of a precocious mind, deeply interested in nature and finding in it the marks of a provident God.

In 1716, Edwards entered Yale, where the world of philosophy opened up to him. For a short time his tutor was Samuel Johnson, who introduced him to the new philosophical ideas coming from England, especially those of John Locke. He read Locke's *Essay concerning Human Understanding*, from which, he claimed, he derived more enjoyment "than the most greedy miser finds, when gathering up handfuls of silver and gold, from some newly discovered treasure." His precocity in philosophy is proved by his notes "Of Being" and "The Mind," both probably written before his graduation in 1720.

There followed two years of graduate study in theology at Yale, in preparation for the ministry. During this period Edwards had a profound religious experience, which he described later, in his *Personal Narrative* (1739), as having given him a new awareness of the absolute sovereignty and omnipresence of God and of complete dependence on him. Edwards's religious philosophy grew out of this transforming experience.

In 1722 he became pastor of a Scotch Presbyterian congregation in New York, but the life of study and teaching attracted him, and two years later he was back at Yale as senior tutor. In 1727 he was ordained assistant minister to his grandfather Solomon Stoddard, and when Stoddard died, in 1729, Edwards took over the Northampton parish.

For almost twenty years Edwards preached and wrote in this parish. During that time he continued his boyhood custom of jotting down his reflections, which he called "Miscellanies" or "Miscellaneous Observations." They fill nine volumes and contain 1,360 entries. These journals, most of which are still unedited, were intended to be a first draft of a monumental book provisionally titled "A Rational Account of the Main Doctrines of the Christian Religion Attempted." This proposed summa of Calvinist theology was not completed.

Edwards's pervasive theme was the Calvinist doctrine of God's sovereignty and the complete helplessness of man to effect his own salvation by good works. In a famous sermon preached in Boston in 1731, titled "God Glorified in Man's Dependence," he opposed Arminianism—a doctrine derived from the Dutch theologian Jacobus Arminius (1560–1609) and then gaining ground in the colonies—which granted to men some part in their salvation through benevolence and good works. Edwards played a vigorous role in the revivalist movement known as the Great Awakening, which swept through New England in the 1740s, reaching hysterical peaks of religious enthusiasm. His own conception of religious experience is found in *A Treatise concerning Religious Affections* (1746).

Through sternness of doctrine and lack of prudence Edwards alienated his parishioners, and in 1748 he was dismissed from his parish. His next post was the missionary parish at Stockbridge, Massachusetts, where he preached to a small group of Indians and a few whites. He had plenty of leisure to write, and a major work, *Freedom of the Will*, defining and defending his Calvinist doctrine of human freedom, appeared in 1754. The sequel, *The Nature of True Virtue* (1765), places virtue in the emotions rather than in the intellect. His last completed work, "Concerning the End for Which God Created the World," is a speculative theological work on God's purpose in creation.

At Stockbridge, Edwards began a vast synthesis of theology called *The History of the Work of Redemption*, but this was interrupted by his election, in 1757, to the presidency of New Jersey College, now Princeton University. He died at Princeton the following year.

PHILOSOPHICAL ORIENTATION

In the language of the day, Edwards was a "philosophizing divine." His primary interests were religious, and his main writings were theological. Apart from his college notes he produced no purely philosophical works. However, his theological treatises abound in philosophical reflections, all of which were intended to clarify and defend his theological positions. For him the arts, sciences, and philosophy ideally had no status separate from theology; as they become more perfect, he said, they "issue in divinity, and coincide with it, and appear to be as parts of it."

Edwards's philosophical views reflect his college training in Puritan Platonism, itself an offshoot of Cambridge Platonism and the Platonism of Peter Ramus. He attempted to synthesize with this Christian Platonism elements from the English empiricists, especially Locke, Newton, and Francis Hutcheson (1694–1746), whose works were introduced into New England in the early 1700s. Puritan Platonism taught Edwards that the spiritual world alone is real, that the visible universe is but its shadow, created to lead the mind, under the divine illumination, to an awareness of the presence of God. Into this general idealistic philosophy he wove strands of doctrine from the empirically minded Locke and the scientist Newton, whose works were beginning to make a stir in the colonies. From Locke he took the notion that all our ideas originate in sensation; from Newton, the conception of space as the divine sensorium.

BEING

In his notes "Of Being," Edwards took up the Parmenidean thesis of the necessity of Being, arguing the impossibility of absolute nothingness on the ground that it is a contradictory and inconceivable notion. Since pure nothingness is an impossibility, he held, there never was a time when Being did not exist. In short, Being is eternal. He also established the omnipresence of Being, arguing that we cannot think of pure nothingness in one place any more than we can think of it in all places. Thus, Being possesses the divine attributes of necessity, eternity, omnipresence, and infinity. Consequently, Being is God himself.

Further attributes of Being deduced by Edwards are nonsolidity and space. Solidity, he argued, is resistance to other solids, and since there are no beings outside of Being, Being itself, or God, cannot be conceived as solid. That Being, or God, is identical with space Edwards proved by the impossibility of conceiving the nonexistence of space. We can suppress from thought everything in the universe but space itself. Hence, space is divine. Following the Cambridge Platonists and Newton, Edwards conceived of God's mind as the locus in which material things spatially exist.

NATURE OF MIND

Edwards's notes titled "The Mind" are heavily indebted to Locke. Like the English philosopher, he distinguished between two faculties of the mind, understanding and will. Understanding he defined as the faculty by which the soul perceives, speculates, and judges. Its first operation is sensation, for without the activity of the senses there can be no further mental operations. The mind needs the senses in order to form all its ideas. The objects of the senses are not real qualities of bodies but impressions and ideas given to us by God. Edwards agreed with Locke that secondary qualities, such as colors, sounds, smells, and

tastes, do not inhere in bodies but are mental impressions. Every intelligent philosopher, Edwards wrote, now grants that colors are not really in things any more than pain is in a needle.

IDEALISM. Edwards went beyond Locke in applying to primary qualities, such as solidity, extension, figure, and motion, the arguments against the reality of secondary qualities. All the primary qualities, he insisted, can be reduced to resistance. Solidity is simply resistance; figure is the termination of resistance; extension is an aspect of figure; motion is the communication of resistance from one place to another. Hence, a visible body is composed not of real qualities but of ideas, including color, resistance, and modes of resistance. Resistance itself is not material; it is "nothing else but the actual exertion of God's power." Consequently, the visible universe has only a mental existence. It exists primarily in God's mind, where it was designed by a free act of the divine will. It also exists in our minds, communicated to us by God in a series of united and regularly successive ideas.

Historians have debated whether Edwards owed his idealistic philosophy to George Berkeley or to his own precocious genius. At the time he formulated it, Berkeley's works were not yet available at Yale. Although it is possible that he heard reports of Berkeley's idealism, it is more likely that he arrived independently at his idealistic conclusions.

According to Edwards, minds alone are, properly speaking, beings or realities; bodies are only "shadows of being." Goodness and beauty belong to anything in proportion to its intensity of being. Hence, minds alone are really good and beautiful; the visible world has but a shadow of these perfections. Its value is to lead the mind to the enjoyment of spiritual and divine goodness and beauty.

CREATION. The created world depends entirely on God for its existence and preservation. He freely created it, and he constantly holds it in existence, as colors are continually renewed by the light of the sun falling on bodies. The universe constantly proceeds from God as light shines from the sun. Under the activity of God the universe is a revelation of the divine mind to created minds; it is a panorama of shadows and images exhibiting the divine mind and will. Edwards, in his notebook titled *Images or Shadows of Divine Things*, described nature as a symbol of God. God, he said, revealed himself in the Bible and also in the visible universe and the souls of men, which are made in the image of God. In order to interpret correctly the symbols of God in the created world, the mind has to be purified by a divine illumination. To Edwards there is no more sublime or delightful activity than to discover and to contemplate the traces of God in nature.

THE WILL. The second faculty of the mind described by Edwards is the will. The importance of the will lies in the fact that it is the seat of the passions or affections, the chief of which is love. According to Edwards, all the other passions originate in love and are for its sake. Love is the excellence and beauty of minds. In *A Treatise concerning Religious Affections* he argued that all human activities, especially those of religion, arise from affection. The affections, he said, are the "very life and soul of all true religion." The essence of religion lies in holy love, especially the love of God. Although Edwards's doctrine of religious experience, under the influence of pietism, gives ample scope to the emotions, and he appealed to them in his sermons, he generally maintained a Puritan sobriety of expression and avoided the sensationalism that marked the Great Awakening. He insisted that religion be centered in what he called the "gracious affections" that spring from the awareness of God and divine things.

RELIGION AND ETHICS

Religious experience is possible, according to Edwards, through a supernatural sense that the elect receive by divine grace. This new sense, which is different from the five bodily senses, gives humankind, reborn by grace, a new kind of sensation or perception by which he passively receives from God ideas and truths about divine things. By a kind of sense experience the elect enjoy an inward, sweet delight in God, which unites them to God more closely than all rational knowledge of him. The way to God is through the heart rather than through the head.

PROBLEM OF FREEDOM. Edwards regarded the will, like the intellect, as an essentially passive power, moved to action by external forces. As the intellect passively receives impressions and ideas from God, so the will is inclined to agreeable objects and repelled by disagreeable objects. The will is not a self-determining power; its actions are determined by causes. God alone is free in the sense that he can determine his own volitions. The principle of causality, according to which everything that happens has a cause, applies to the movements of the human will as it does to everything created. Of course, the will is moved not by physical causes but by motives or moral causes. These motives are presented to the will by the understanding, and the strongest of them determines the movement of the will.

Edwards opposed the Arminians of his day, who attributed to the human will an inner spontaneity and power of self-determination. In his view this kind of freedom is a divine prerogative; the human will does not have this kind of inner freedom. Its actions are determined not by being physically coerced but by being morally necessitated. A man cannot help willing as he does, given the motives presented to him. And since these motives are determined by God's providence, the movements of man's will are entirely within the divine power.

Although Edwards denied that the human will has freedom of self-determination, he granted that in a sense man is free. Like Thomas Hobbes and Locke, he defined human liberty as the ability to carry out what the will inclines man to do. Liberty is the absence of impediments to action. This denial of the essential freedom of the will harmonizes well with Edwards's Calvinist belief in the total depravity of man and in predestination.

VIRTUE. The third earl of Shaftesbury (1671–1713) and Hutcheson influenced Edwards's ethics. With them he denied that true virtue consists in the selfish pursuit of pleasure or in the utility of human actions. Rather, virtue is disinterested benevolence or affection; it is the intrinsic beauty of the dispositions of man's heart. An action is good not because it is advantageous to ourselves or to others but solely because it springs from a beautiful disposition of will. Virtue is a spiritual beauty or excellence that commends itself to us for its own sake. Any other motive for acting is based on self-love and consequently does not measure up to true virtue.

Edwards did not think that man has a natural impulse to such disinterested virtue. In his view man, owing to original sin, is totally depraved and given over to self-love. Only by the election of God and the gift of efficacious grace can man rise above his "dreadful condition" and perform truly virtuous actions. Without supernatural aid seemingly disinterested affections, such as the natural love of parents for their children, are accompanied by self-love and hence are not truly virtuous. At most they are secondary virtues or the shadows of true virtue.

Edwards was the most gifted and articulate theologian-philosopher in the New England colonies and perhaps in American history. He supported a losing cause in his defense of Puritanism, but for a while he gave it new life and spirit. The liberal theology that he combated all his life finally won the day; in the form of Unitarianism it dominated New England culture in the nineteenth century. But Edwards's powerful religious and philosophical stimulus remained. New England transcendentalists, such as Emerson, although rejecting all systematic theology and proclaiming the divinity of humankind, continued the Puritan's passionate search for the divine in the communion with nature.

See also Arminius and Arminianism; Being; Berkeley, George; Cambridge Platonists; Determinism and Freedom; Emerson, Ralph Waldo; Hutcheson, Francis; Idealism; Johnson, Samuel; Locke, John; New England Transcendentalism; Newton, Isaac; Platonism and the Platonic Tradition; Ramus, Peter; Shaftesbury, Third Earl of (Anthony Ashley Cooper).

Bibliography

WORKS BY EDWARDS

Works. 8 vols, edited by S. Austin. Worcester, MA: Isaiah Thomas, 1808–1809; reprinted in 4 vols., New York, 1844, 1847. An old but useful edition, found in some libraries.

Works. 10 vols, edited by S. E. Dwight. New York: Carvill, 1829–1830. This is the standard and best edition of Edwards's works, except for those newly edited. Vol. I contains a "Life of Edwards" by Dwight.

Representative Selections, edited by C. H. Faust and T. H. Johnson. New York: American, 1935; rev. ed. (paperback), with rev. and updated bibliography, 1962. Useful selections from Edwards's works and a good bibliography.

Images or Shadows of Divine Things, edited by P. Miller. New Haven, CT: Yale University Press, 1948.

Puritan Sage: Collected Writings of Jonathan Edwards, edited by V. Ferm. New York, 1953. Useful selection of Edwards's works.

Works, edited by P. Miller. New Haven, CT: Yale University Press, 1957–. Vol. I: *The Freedom of the Will*, edited by P. Ramsey. Vol. II: *Religious Affections*, edited by J. E. Smith. A new edition that will supersede Dwight's.

The Nature of True Virtue. Ann Arbor: University of Michigan Press, 1960. Foreword by W. Frankena. Best edition of this work.

Letters and Personal Writings. Edited by George S. Claghorn. New Haven, CT: Yale University Press, 1998.

Sermons and Discourses, 1730–1733. Edited by Mark R. Valeri. New Haven, CT: Yale University Press, 1999.

WORKS ON EDWARDS

Chai, Leon. *Jonathan Edwards and the Limits of Enlightenment Philosophy*. New York: Oxford University Press, 1998.

Daniel, Stephen H. *The Philosophy of Jonathan Edwards: A Study in Divine Semiotics*. Bloomington: Indiana University Press, 1994.

Davidson, Edward H. *Jonathan Edwards: The Narrative of a Puritan Mind*. Cambridge, MA: Harvard University Press, 1966, 1968.

Delattre, Roland André. *Beauty and Sensibility in the Thought of Jonathan Edwards: An Essay in Aesthetics and Theological Ethics*. New Haven, CT: Yale University Press, 1968.

Hatch, Nathan O., and Harry S. Stout. *Jonathan Edwards and the American Experience*. New York: Oxford University Press, 1988.

Holbrook, Clyde A. *The Ethics of Jonathan Edwards: Morality and Aesthetics*. Ann Arbor: University of Michigan Press, 1973.

Jenson, Robert W. *America's Theologian: A Recommendation of Jonathan Edwards*. New York: Oxford University Press, 1988.

Lesser, M. X. *Jonathan Edwards: A Reference Guide*. Boston, MA: G.K. Hall, 1981.

Marsden, George M. *Jonathan Edwards: A Life*. New Haven, CT: Yale University Press, 2003.

McClymond, Michael James. *Encounters with God: An Approach to the Theology of Jonathan Edwards*. New York: Oxford University Press, 1998.

McDermott, Gerald R. *Jonathan Edwards Confronts the Gods: Christian Theology, Enlightenment Religion, and Non-Christian Faiths*. New York: Oxford University Press, 2000.

McDermott, Gerald R. *One Holy and Happy Society: The Public Theology of Jonathan Edwards*. University Park: Pennsylvania State University Press, 1992.

Miller, P. *Errand into the Wilderness*. Cambridge, MA: Belknap Press of Harvard University Press, 1956. Both this and *Jonathan Edwards* (below) are first class.

Miller, P. *Jonathan Edwards*. New York: W. Sloane Associates, 1949; paperback ed., 1959.

Schneider, H. W. *A History of American Philosophy*, 11–31. New York: Columbia University Press, 1946; 2nd ed., 1963.

Schneider, H. W. *The Puritan Mind*. New York: Holt, 1930; paperback ed., Ann Arbor, MI, 1958.

Smith, John Edwin. *Jonathan Edwards: Puritan, Preacher, Philosopher*. Notre Dame: University of Notre Dame Press, 1992.

Winslow, O. E. *Jonathan Edwards, 1703–1758*. New York: Macmillan, 1940; paperback ed., 1962. Pulitzer Prize–winning biography—excellent and fascinating account of Edwards's life and times, with good bibliography.

Armand A. Maurer (1967)
Bibliography updated by Michael J. Farmer (2005)

EFFICIENCY

See *Philosophy of Technology*

EGALITARIANISM

See *Social and Political Philosophy*

EGOISM AND ALTRUISM

Why do we sometimes prefer to consult the interests of others rather than our own interests? What is the relationship between selfishness and benevolence? Is altruism merely a mask for self-interest? At first sight these may appear to be empirical, psychological questions, but it is obviously the case that even if they are construed as such, the answers will depend on the meaning assigned to such key expressions as "self-interest," "benevolence," "sympathy," and the like. It is in connection with elucidating the meaning of such expressions that philosophical problems arise—problems that are of particular interest because we cannot understand such expressions without committing ourselves, in some degree, to some particular conceptual schematism by means of which we can set out the empirical facts about human nature. That there are alternative and rival conceptual possibilities is a fact to which the history of philosophy testifies.

The problems with which we are concerned do not appear fully-fledged until the seventeenth and eighteenth centuries. That they do not is a consequence of the specific moral and psychological concepts of the Greek and of the medieval world. In neither Plato nor Aristotle does altruistic benevolence appear in the list of the virtues, and consequently the problem of how human nature, constituted as it is, can possibly exhibit this virtue cannot arise. In the *Republic* the question of the justification of justice is indeed raised in such a way as to show that if Thrasymachus's account of human nature were correct, men would find no point in limiting themselves to what justice prescribes, provided that they could be unjust successfully—and Thrasymachus's account of human nature is certainly egoistic. But Plato's rejoinder to Thrasymachus is a statement of a different view of human nature in which the pursuit of "good as such" and the pursuit of "my good" necessarily coincide.

In the medieval world the underlying assumption is that man's self-fulfillment is discovered in the love of God and of the rest of the divine creation. So although Thomas Aquinas envisages the first precept of the natural law as an injunction to self-preservation, his view of what the self is and of what preserving it consists in leads to no special problems about the relation between what I owe to myself and what I owe to others. It is only when Thomas Hobbes detaches the doctrines of natural law from their Aristotelian framework that the problem emerges in a sharp form.

INITIAL HOBBESIAN STATEMENT

Hobbes is the first major philosopher, apart from Niccolò Machiavelli, to present a completely individualist picture of human nature. There are at least three sources of Hobbes's individualism. First, there is his reading of political experience. His translation of Thucydides reveals his preoccupation with the topic of civil war, with the struggle of one private interest against another. Second, there is Hobbes's commitment to the Galilean resoluto-

compositive method of explanation: To explain is to resolve a complex whole into its individual parts and to show how the individual parts must be combined in order to reconstruct the whole. To explain the complex whole of social life is, therefore, to resolve it into its component parts, individual people, and to show how individuals must combine if social life is to be reconstructed. Since the individuals in terms of whose coming together social life is to be explained must be presocial individuals, they must lack those characteristics that belong to the compromises of social life and be governed only by their presocial drives. Third, there is the detail of the Hobbesian psychology, which insists that such drives must be competitive and aggressive because of the will to power over other men that ceaselessly and restlessly drives men forward.

Thus, from all three sources arises a picture of human nature as essentially individual, nonsocial, competitive, and aggressive. From this view it follows that the apparent altruism and benevolence of individuals in many situations need to be explained; the Hobbesian explanation is simply that what appears to be altruism is always in fact, in one way or another, disguised self-seeking. Undisguised, unmodified self-seeking leads to total social war. The fear of such war leads to the adoption of a regard for others from purely self-interested motives. John Aubrey in his sketch of Hobbes in *Brief Lives* tells of an exchange between Hobbes and a clergyman who had just seen Hobbes give alms to a beggar. The clergyman inquired whether Hobbes would have given alms if Jesus had not commanded it; Hobbes's reply was that by giving alms to the beggar, he not only relieved the man's distress but he also relieved his own distress at seeing the beggar's distress. This anecdote compresses the central problem into a single point: Given that human nature is competitive and self-seeking, why and how can altruism and benevolence be treated as virtues? One's immediate response to this brief and cryptic statement of the problem may well be to inquire why—if one does not share Hobbes's premises—one should take it as given that human nature is essentially self-seeking. To this one replies by posing another question: How can any actual or possible object or state of affairs provide me with a motive, appear to me as good or desirable, unless it appears to be what will satisfy some desire of mine? If the (necessary and sufficient) condition of an object's providing me with a motive is that it satisfy some desire of mine, then it will surely be the case that all my actions will have as their goal the satisfaction of my desires. And to seek only to satisfy my own desires is surely to have an entirely self-seeking nature.

EIGHTEENTH-CENTURY RESTATEMENTS

The root of the problem lies in the apparently egoistic implications of the psychological framework within which the questions of moral philosophy have been posed by a whole tradition of British thinkers from Hobbes on. Within this framework philosophers have oscillated between two positions: the Hobbesian doctrine of altruism as either a disguise or a substitute for self-seeking and the assertion of an original spring of altruistic benevolence as an ultimate and unexplained property in human nature.

On the one side we find, for example, the earl of Shaftesbury, who argues that men are so contrived that there is no conflict, but an identity, between what will satisfy self-interest and what will be for the good of others; the practice of benevolence is what satisfies man's natural bent. Bernard Mandeville, in *The Grumbling Hive, or Knaves Turn'd Honest* (later retitled *The Fable of the Bees: or, Private Vices, Public Benefits*), argues by contrast that the only spur to action is private, individual self-seeking and that it is for the public and general good that this is so. Francis Hutcheson, who treats benevolence as constituting the whole of virtue, provides no argument to back up his view, nor does he explain why we approve of benevolence rather than of self-interest.

BUTLER. Bishop Joseph Butler's position is at once more complex and more interesting than Hutcheson's or Mandeville's. Butler believes that we have a variety of separate and independent "appetites, passions and affections." Of these, self-love is only one, and it is not necessarily opposed to benevolence. We satisfy the desire for our own happiness in part, but only in part, by seeking the happiness of others. A man who inhibits those desires of his that find their satisfaction in achieving the happiness of others will not in fact make himself happy. By refusing to be benevolent, he damages his own self-interest and disobeys the call of self-love. Cool and reasonable self-love consists in guiding our actions by reference to a hierarchy of principles; supreme among these is moral reflection or conscience, by means of which human nature is defined and the good that will satisfy it discerned. Thus, self-love itself refers us to the arbitration of conscience, which in turn prescribes that extent and degree of benevolence that will satisfy the needs of self-love.

The chief objection to Butler is likely to arise from the apparently self-enclosed character of his account. In Butler's system the harmony between self-love and benevolence appears to reign by definition rather than in

fact, that is, in human nature itself. But this criticism misconstrues Butler's stand, although we can deduce from Butler's psychology empirical consequences of a testable kind that at first sight render it liable to refutation by the facts. For if Butler is correct, those who are benevolent to the required degree do not find their benevolence at odds with their self-interest. In this sense, at least, virtue and happiness may be required to coincide, and if they do not coincide, Butler's view of human nature is false. But Butler allows himself an escape clause. He concedes that in the world as we know it, the pursuit of self-interest and devotion to benevolence may not appear to coincide, but, he says, the divergence seems to exist only if we do not allow for divine providence, which ensures that the world to come will be such as to ensure that self-interest and altruistic benevolence required the same actions of us.

THEOLOGY AND THE LONG RUN

In contrast with Hobbes's view that altruistic behavior (or at least just behavior) is in our immediate interest as a means of preserving ourselves from the war of all against all and in contrast with Butler's view that benevolence and self-interest are two distinct springs of action that move us to the same actions, there is the view that benevolence is to our long-term, as opposed to our short-term, self-interest. Butler, as already noted, uses something like this view to supplement his basic position, but it is the stock in trade of a form of theological egoistic utilitarianism to be found in Abraham Tucker and William Paley.

In both writers the crucial psychological premise is that men are so constructed that they always pursue their own private and individual satisfaction. In both writers the fundamental moral rule is an injunction to universal benevolence, which is equated with the promotion of the greatest happiness of the greatest number. The problem is how, given the character of human nature, a motive can be found for obeying the fundamental moral rule. The solution is to say that God has so contrived the afterlife that only if we obey the fundamental moral rule will we in the long run, that is, in the eternal run, secure our own happiness. In Paley it is clear that we could find no good reason to be moral if God did not exist, but God's function in bridging the gulf between self-interest and morality is veiled in conventional theological terms. In Tucker's *The Light of Nature Pursued* the account of how God bridges the gulf is more explicit. God has arranged that all the happiness that men either have enjoyed or will enjoy is deposited in what he calls "the bank of the universe." By working to increase the happiness of others, I increase the amount of happiness so deposited. But by increasing the general stock of happiness, I also increase my own happiness, for God has arranged to divide this stock of happiness into equal shares, to be allotted one to a person, and so by increasing the size of the general stock, I also increase the size of my own share. I am, as it were, a shareholder in a cosmic bank of which God is at once the chairman and the managing director.

Tucker's absurdities, though unimportant in detail, do bring out how impossible is the task of reconciling an egoistic theory of human nature with a moral theory of benevolent utilitarianism. Of such impossibilities are absurdities born; to this the secular utilitarianism of David Hume, Jeremy Bentham, John Stuart Mill, and Henry Sidgwick is as much a witness as is the theological utilitarianism of Tucker and Paley.

HUME AND THE UTILITARIANS

Hume's initial approach to the problem is as flexible and undogmatic as that of any philosopher. In the *Treatise of Human Nature* Hume poses the question why we approve and obey rules that it is often in our interest to break. He makes no assumptions of the kind found in other eighteenth-century writers (men are entirely ruled by self-interest). He merely remarks, apparently on empirical grounds, that it is often the case that self-interest would, if it were followed, lead us to disregard the rules of justice. Nor does he invoke any compensating natural regard for the interests of others. We do have some regard for the interests of others, but it varies with the closeness of their ties to us, and we have by nature no regard for the public interest as such. "In general, it may be affirm'd that there is no such passion in human minds as the love of mankind, merely as such, independent of personal qualities, of services, or of relation to oneself" (*Treatise,* Bk. III, Part II, Sec. i).

If, then, self-interest would lead us to disobey the rules of justice and if we have no natural regard for the public interest, how do the rules come into existence, and what fosters our respect for them? The crucial fact is that did we not have respect for the rules of justice, there would be no stability of property. Indeed, the institution of property could not and would not exist. Now the existence of property and its stability is to all our interests, and we are always conscious of how much we are injured by others failing to observe the rules. So we have become conscious that although our immediate and short-term benefit rests in breaking the rules on a given occasion, our long-term benefit resides in insisting upon a universal observance of the rules.

By the time Hume came to write the *Enquiry concerning Human Understanding,* he had shifted his ground. He now sees self-interest and "a tendency to public good, and to the promoting of peace, harmony, and order in society" as two independent, coexistent springs of action; he sees the independent power of sympathy and of a sense of the public good, rather than a rational view of what is of long-term benefit to self-interest, as moving us to benevolence and altruism.

BENTHAM, GROTE, MILL, SIDGWICK. The utilitarians present the problem in terms differing somewhat from those of Hume because they were more rigidly committed to a psychology derived from David Hartley, according to which only pleasure and pain ever move us to action. In this psychology both "pleasure" and "pain" are the names of sensations. Clearly in this view the only pleasure whose prospect attracts me is *my* pleasure, and the only pain the prospect of which repels me is *my* pain. It seems to follow that all action is egoistically motivated, yet all four utilitarian writers make "the greatest happiness of the greatest number" either the only criterion of action or at least a central criterion. How can so egoistically motivated an agent as the utilitarians assume consult the general happiness? That he will have to learn to do so is what Bentham takes for granted in his legal and political writings. Bentham provides for inducements that will counteract the self-interest of legislators, for example. He affirms expressly that "the only interest which a man is at all times sure to find adequate motives for consulting is his own." But in the *Deontology* he seems by contrast to take it for granted that the pursuit of *my* pleasure and the pursuit of the greatest happiness of the greatest number will always as a matter of fact coincide.

This assumption of coincidence is abandoned by John Grote, who tries to minimize the difficulties by reducing our obligation to consult the general happiness to an injunction to consult the general happiness insofar as to do so will ensure our own happiness. Yet even Grote presupposes that, for the most part and generally, my happiness and that of the greatest number will not conflict.

Mill's arguments are of two kinds. He first argues that pleasure and the absence of pain are desired by all; here what is meant is clearly that each desires his own pleasure. The proof, and the only possible proof, that pleasure is desirable is that all people desire it, and since all people do desire it, it must be admitted to be desirable. Hence, everyone must acknowledge that it is desirable to produce as much pleasure as possible, and here what is clearly meant is that each ought to desire the pleasure of all. The fallacy in the transition from the premise that each desires his own pleasure to the conclusion that each ought to desire the pleasure of all is usually thought to reside in the transition from fact to value, but it lies, rather, in the transition from an assertion about the agent's own pleasure to conclusions about the general happiness.

However, elsewhere in *Utilitarianism* Mill faces the difficulties in such a transition explicitly. He reproduces familiar arguments in an interesting form. The feelings of sympathy that Hume stressed in the *Enquiry* reappear as a man's "feeling of unity with his fellow-creatures." A man who has this feeling has a "natural want" to live in harmony with others. It is often overshadowed by selfish emotions, but those who do possess it know that they would be worse off if they did not possess it. The reason for this conviction is that the best prospect of realizing such happiness as is attainable is a willingness to sacrifice the prospects of one's own present and immediate happiness to an ascetic devotion to altruism and benevolence. Sidgwick became conscious of the difficulties that Mill brushes aside in this account. In the *Methods of Ethics,* however, Sidgwick could find no way to make the transition from the desire for one's own pleasure to that for the general happiness, and these remain for him independent goals, as they had been for some eighteenth-century philosophers.

THE PROBLEM IN EMPIRICAL PSYCHOLOGY

The philosophers from Hobbes to Sidgwick who analyze the concepts of egoism, altruism, and sympathy often write as if they were empirical students of human nature, disputing the facts of human action and motivation. But it is more illuminating to read them as offering conceptual accounts of what it is to have a good reason for action and of what the limits upon the range of possible good reasons are. But so closely allied are conceptual and empirical issues at this point in the argument that it is not surprising to find that the would-be empirical accounts that psychologists claim to have derived from observation should sometimes turn out to be a rendering of conceptual schemes which have already been encountered in philosophy. So it is with Sigmund Freud, most strikingly in his earlier writings. The important place in Freudian theory held by the pleasure principle, the concepts of gratification and of libido, and the consequent view of socialization all lead to a theory in which the gratification of the self is primary and in which altruism and benevo-

lence are interpreted as secondary phenomena that acquire the regard that they do because they are originally associated with forms of self-gratification. Freud's genetic account differs in detail from that given by Mill, but the form of the account is the same. Nor is this accidental; the pre-Freudian psychologies of Hartley, who influenced Mill, and of Alexander Bain, Mill's contemporary offer associationist accounts in which the genetic order is the same as it is in Freud. There is, therefore, not only the task of clarifying the concepts involved in these accounts, but also the task of settling how far the issues raised are genuinely empirical and how far genuinely conceptual. The concepts in need of clarification are of five kinds: the nature of desire; self-interest; altruism and benevolence; motives, actions, and sympathies; and the genetic fallacy.

NATURE OF DESIRE. If I want something, it does not follow that I want it because it will give me pleasure to have it or because it is a means of getting something further which will give me pleasure. It is, of course, true that if I get what I want, I have thereby satisfied one of my wants. Having any of my wants unsatisfied is certainly less satisfactory than having them satisfied, but it is not necessarily painful or even unpleasant. So it is neither true that I necessarily desire pleasure nor true that in seeking to satisfy my desires, I necessarily seek pleasure or the avoidance of pain.

Moreover, if I do something, it does not follow that I do it because I want to, let alone that I do it because I shall get pleasure from it. It has sometimes been suggested that the performance of an action is itself an adequate criterion of the agent's wanting to do whatever it is, and those who hold this view interpret such an expression as "doing what one does not want to do" when it is applied in cases of action under duress as meaning that the agent would not want to perform that particular action normally but does want to do it on this occasion rather than endure the threatened consequences of not doing it. This contention is less than self-evident. Moreover, if there is a sense of "want" such that if I do something, it is thereby true that I want to do it, that sense is a weaker and a different one from that given when I explain what I do by citing as a, or the, reason that I want to do it. For it is precisely because we have independent criteria for asserting that the agent did or did not want to do what he did that the want can be cited as an explanation for the action.

Action, desire, and pleasure, then, do not stand in so close a conceptual relationship that we cannot ask as a matter of contingent fact on any given occasion whether a man acted to get pleasure or whether he did what he did because he wanted to or not. To understand this is a necessary preliminary to understanding the notion of self-interest.

SELF-INTEREST. What is to my interest depends upon who I am and what I want. This elementary but too often unnoticed truism underlies one of Socrates's implied answers to Thrasymachus in Plato's *Republic.* The question "Is justice more profitable than injustice?" will, as Plato makes clear, be answered differently depending upon whether it is answered by a just man or an unjust man. For what the just man wants is not what the unjust man wants. Thus, there is not a single spring of action or a single set of aims and goals titled "self-interest" that is the same in every man. "Self-interest" is not in fact the name of a motive at all. A man who acts from self-interest is a man who allows himself to act from certain motives in a given type of situation. The same action done from the same motive in another type of situation would not be correctly characterized as done from self-interest. So if I eat to sate my hunger or do my job well in order to succeed, I do not necessarily act from self-interest. It is only when I am in a situation where food is short or my rising in the world requires a disregard for the legitimate claims of others that to consult only my hunger or my ambition becomes to act from self-interest. The notion of self-interest therefore has application not to human behavior in general but to a certain type of human situation, namely, one in which behavior can be either competitive or noncompetitive. Equally, in this type of situation alone can the notions of benevolence and altruism have application. Therefore, it is to the elucidation of these that we must next turn.

ALTRUISM AND BENEVOLENCE. The question canvassed in the eighteenth century whether benevolence might not be the whole of virtue could have been raised only in an age in which the concept of virtue had been greatly narrowed or the concept of benevolence had been greatly widened or both. For in most of my dealings with others of a cooperative kind, questions of benevolence or altruism simply do not arise, any more than questions of self-interest do. In my social life I cannot but be involved in reciprocal relationships, in which it may certainly be conceded that the price I have to pay for self-seeking behavior is a loss of certain kinds of relationships. But if I want to lead a certain kind of life, with relationships of trust, friendship, and cooperation with others, then my wanting their good and my wanting my good are not two independent, discriminable desires. It is not even that I have two separate motives, self-interest and benevolence,

for doing the same action. I have one motive, a desire to live in a certain way, which cannot be characterized as a desire for my good rather than that of others. For the good that I recognize and pursue is not mine particularly, except in the sense that I recognize and pursue it.

We can now diagnose one major cause of confusion in the whole discussion. All too often from Hobbes on, a special type of human situation has been treated as a paradigm of the whole moral life—that is, a situation in which I and someone else have incompatible aims and my aims are connected only with my own well-being. Of course, such situations do arise, but the clash between self-interest and benevolence that characterizes them is only one case out of many in which incompatible aims have to be resolved.

MOTIVES, ACTIONS, AND SYMPATHY. We can now understand that at the root of the confusions lies a belief in the possibility of a purely a priori characterization of human motives. From Hobbes on there has been a tradition, shared by empiricists as well as by their critics, which seeks to discuss human motivation almost entirely in the light of general conceptual considerations about desires, the passions, and pleasure and pain. What evades this tradition is not only the variety of aims and motives that can inform action, a variety to be discovered only by empirical inspection, but also the specific and particular character of certain motives.

The difficulties in the notion of sympathy, for example, are such that one cannot inquire straightforwardly whether there is or is not a sympathy for humankind as such. To say that a man acted from sympathy is always to refer to a set of particular occasions when sympathy was aroused for particular people in some particular plight. How wide the range of a man's sympathies is, is an empirical fact, and there is no conceptual limit to the possibilities. But it is a conceptual point that just as a generalized ambition can be manifested only in particular aspirations, so a generalized sympathy can be manifested only in particular acts of charity and benevolence. Now, suppose a man to perform a charitable and benevolent action; we would be wrong to suppose that we can always answer the question whether he was sympathetic to them because they were his relations (or his countrymen or his next-door neighbors) or whether he would have been equally sympathetic if they had been strangers or foreigners. A man can act out of sympathy without the range of his sympathies being determinate. Thus, the eighteenth-century question whether there is, as such, a general benevolence toward humankind implanted in human breasts is misleading.

GENETIC FALLACY. The question of innate benevolence toward humankind is also misleading because the eighteenth-century view disregards both the variety and the variability of human nature. Philosophers discuss what passions men have and not what passions they might acquire. Learning is, at best, peripheral to their inquiry; insofar as it does enter, there is another fallacy in writers from Hobbes on—that of confusing the question of what motives there were originally (for Hobbes, in the state of nature; for Freud, in early childhood) with the question of what the fundamental character of motives is now, in adult life. Because the instinctual drives and desires of young children have to be socialized, it does not follow that adult attitudes and emotions are only masks for such drives and desires. This is not to say that they cannot be such masks, but if the notion is to have any content, whether they are must be an empirical question.

See also Altruism; Aristotle; Bain, Alexander; Bentham, Jeremy; Butler, Joseph; Ethical Egoism; Freud, Sigmund; Grote, John; Hartley, David; Hobbes, Thomas; Human Nature; Hume, David; Hutcheson, Francis; Machiavelli, Niccolò; Mandeville, Bernard; Mill, John Stuart; Paley, William; Plato; Self-Interest; Sidgwick, Henry; Thomas Aquinas, St.; Thucydides; Utilitarianism.

Bibliography

Broad, C. D. "Certain Features in Moore's Ethical Doctrines." In *The Philosophy of G. E. Moore,* edited by Paul A. Schilpp, 43–57. Evanston, IL: Northwestern University Press, 1942.

Broad, C. D. *Five Types of Ethical Theory,* 161–177. London: Kegan Paul, 1930.

Brunton, J. A. "Egoism and Morality." *Philosophical Quarterly* (1956): 289–303.

Ewing, A. C. *Ethics.* London: English Universities Press, 1953. Ch. 2.

Medlin, B. "Ultimate Principles and Ethical Egoism." *Australasian Journal of Philosophy* 35 (1957): 111–118.

Moore, G. E. *Principia Ethica.* Cambridge, U.K.: Cambridge University Press, 1903; paperback, 1959.

Rashdall, Hastings. *Theory of Good and Evil,* Vol. I, 44–63. London: Oxford University Press, H. Milford, 1924.

Sharp, F. C. *Ethics.* New York: Century, 1928. Chs. 22–23.

Alasdair MacIntyre (1967)

EHRENFELS, CHRISTIAN FREIHERR VON
(1859–1932)

Christian Freiherr von Ehrenfels, the Austrian psychologist and philosopher, was born in Rodaun near Vienna. He studied at the University of Vienna under Franz Brentano and Alexius Meinong, and took his doctorate at Graz in 1885. He taught at Vienna as a *Privatdozent* from 1888 to 1896, when he became extraordinary professor at the German University of Prague. He was a full professor at Prague from 1900 until 1929. Besides his professional work, Ehrenfels wrote two essays on Richard Wagner and several plays.

GESTALT PSYCHOLOGY

In psychology, Ehrenfels is best remembered for inaugurating gestalt psychology in his article "Über Gestaltqualitäten" (1890). Starting from Ernst Mach's thesis in his *Beiträge zur Analyse der Empfindungen* (Jena, 1886), that we can sense (*empfinden*) spatial and temporal forms ("wholes," *Gestalten*), Ehrenfels argued that sensing is limited to the present but that the apprehension of a complex datum requires recollection and so seems to lack the immediacy of sensing. This is particularly evident in the case of acoustic data, but it also holds for visual data perused successively. The immediate apprehension of a melody or a figure must therefore be otherwise accounted for than by sensing. Discussing acoustic complexes, Ehrenfels showed that what is in fact apprehended differs from the complex or sum of the component elements, since these vary while the gestalt remains unchanged. This is corroborated by the fact that acoustic forms (melodies) are more easily remembered than are tonal intervals or absolute pitch. Similarly, figures do not depend for their apprehension on absolute location. This implies that gestalt qualities are positive representational contents bound up with the occurrence in consciousness of complexes consisting of separable elements. In Meinong's language (adopted by Ehrenfels in a later paper), they are "founded contents" (*fundierte Inhalte*).

Ehrenfels's notion of gestalt was essentially developed from a differential analysis of data, complex, and unity, unity being regarded as a quality. The phenomenological account of a gestalt in terms of contrast, background, and poignancy—features essential to subsequent gestalt psychology—was secondary in Ehrenfels's analysis, although he did mention such features.

Ehrenfels extended the notion of gestalt to numbers and to the field of logic. He viewed the contradiction in such concepts as that of a round square as a temporal gestalt quality of the psychic process of attempting to form a representation of the concept, an attempt that proves unfeasible. Ehrenfels also used the notion of gestalt in cases, such as phenomena of style and behavior, in which an analysis into component elements is practically impossible. In general, a gestalt is a novel and creative feature with respect to its component elements (in contrast to David Hume, who admitted only the composition of impressions or ideas and imaginative interpolation within the continuum of sensory qualities).

VALUE THEORY

Ehrenfels made important contributions to value theory and ethics. His series of articles, "Werttheorie und Ethik," although inspired by Meinong's lectures, was published before Meinong's ethical works and possessed at least partial originality. Ehrenfels's subsequent *System der Werttheorie* (1897–1898) discussed points of difference with Meinong's first publications on value theory. Ehrenfels defined value as "the relation, erroneously objectified by language, of a thing to a desire directed towards it" ("Werttheorie und Ethik," in *Vierteljahrsschrift für wissenschaftliche Philosophie,* Vol. 17, p. 89) or to a disposition of desire or feeling (ibid., pp. 209–210). "The value of a thing is its desirability" (*System der Werttheorie,* Vol. I, p. 53). Ehrenfels took value not simply as instrumental to the promotion of one's happiness but insisted that instrumental value (*Wirkungswert*) is valuable only relative to intrinsic value (*Eigenwert*). We desire the existence or nonexistence of something, and do not necessarily strive for its possession as a means to our happiness. The valuable object is not bound up with utility (*Nutzen*) but possesses a more general fittingness (*Frommen*) for us. Ehrenfels adapted the economic theory of marginal utility to explain the strength of any desires possessing a fittingness for us (*Grenzfrommen*). He thus introduced a quantitative element of valuation: Values and valuation are conditioned by the prior existence of other value objects.

In view of their dependence on emotional dispositions, values have a certain relativity, but there exists wide agreement among human beings as to the value of pleasure and pain and of certain other psychic phenomena, both in ourselves and in others. We value those valuational dispositions of others that are directed toward objects valued by us. In fact, Ehrenfels restricted intrinsic values to psychic realities.

The relativity of values is also apparent in changes in valuation brought about by various causes. Ehrenfels also

distinguished trends of valuation, for which he offered a theoretical scheme. Means may turn into ends, as when the satisfaction of feelings of hunger replaces nourishment as the end of eating. By contrast, superior values may feature as ends, as when in the interest of nourishment we suppress our feelings of hunger in the presence of poisonous food. A third factor in trends of value is survival, which is best assured if the object serving it coincides with it. Ends transcending mere survival are exemplified in cultural progress, in which values become nonindividualistic. Superior nonindividualistic values are transmitted through example and suggestion, and cause further value promotion in a value milieu. Ehrenfels found reason to believe that with the increasing integration of human knowledge an upward trend toward superior values could be expected.

SOCIAL ETHICS

Ehrenfels's theory of value formed the basis for his ethics, which he subdivided into social and individual ethics. Social ethics is concerned with ethical valuation, that is, valuation of psychic (or supposedly psychic) objects that are causally related to certain actions. These objects are intrinsic values, and we demand that a plurality of individuals coincide in their valuation of them. The ultimate object of ethical valuation is not action, or its means or ends, but the desiderative and emotional disposition behind it. It is then called moral (or immoral) disposition, and its valuation moral (or immoral) valuation. (Accordingly, morality is distinguished from law and custom, which do not consider disposition.) Moral dispositions are the emotional dispositions of taking pleasure in others as intrinsic values, that is, as individuals themselves possessing a disposition toward actions serving intrinsic values, particularly the dispositions of love of one's neighbor, of humanity, of God. Such pleasure in others psychologically depends on an awareness of them in thought or in more or less vivid representation. There is a perspective of comparative closeness or distance in valuation. Among other moral dispositions are justice, constancy, and honesty, and their negative counterparts.

INDIVIDUAL ETHICS

Individual ethics is concerned with man's response, through "mystical" or "tragic elevation," to his fate as a finite body. The craving for such elevation is the source of the valuations (ethical sanction, conscience) of whatever goes to promote it. These private valuations do not strictly encompass the socioethical ones, but do as a matter of fact coincide with them. Ehrenfels's individual ethics thus was a separate strain centering on an aesthetic desire for psychic harmony. To reach such a state, belief in God or metaphysical convictions are helpful though not indispensable.

SEXUAL AND RACIAL VIEWS

Ehrenfels's tendency to emphasize biological factors led him in later writings ("Sexuales Ober- und Unterbewusstsein," 1903–1904; *Sexualethik,* 1907; "Sexualmoral der Zukunft, 1930; cf. the earlier statement in "Werttheorie und Ethik," *Vierteljahrsschrift für wissenschaftliche Philosophie,* Vol. 17, p. 354) to question moral restraint on sexuality and to advocate greater frankness, honesty, and delicacy in marital relations. He won Sigmund Freud's praise for his pioneering work in this field. His biological tendency also led him to recommend selective breeding practices for man (cf. "Die sexuale Reform," 1903–1904) and to embrace ideas bordering on race prejudice ("Leitziele zur Rassenbewertung," 1911).

METAPHYSICS

In his *Kosmogonie* (1916) Ehrenfels contributed to metaphysics a theory of the origin of the world. Rejecting a monism that admits only the cumulative effects of accidental events, he regarded the origin of the world as the result of the interaction of two principles, a principle of chaotic disorder and a principle of psychoid unity of gestalt that, with infinite improbability but with infinite time to allow for its incipience, has been solicited by the opposing principle. Once the principle of unity has been engaged, the resulting gestalt survives because it is infinitely improbable that chaos is capable of continuous destructive action of its own even in infinite time. The gestalt principle, in turn, is credited with creativity, making for further development. Ehrenfels's cosmogony can be taken as a speculative abstraction intended to put the theory of evolution on a new footing in that it tries to give a plausible account of emerging nonrandomness in the universe.

See also Brentano, Franz; Ethics, History of; Freud, Sigmund; Gestalt Theory; Mach, Ernst; Meinong, Alexius; Metaphysics, History of; Racism; Value and Valuation.

Bibliography

WORKS BY EHRENFELS

"Über Gestaltqualitäten." *Vierteljahrsschrift für wissenschaftliche Philosophie* 14 (1890): 249–292.

"Werttheorie und Ethik." *Vierteljahrsschrift für wissenschaftliche Philosophie* 17 (1893): 26–110, 200–266, 321–363, 413–425; 18 (1894): 22–97. Five consecutive articles.
System der Werttheorie. 2 vols. Leipzig: O.R. Weisland, 1897–1898.
"Sexuales Ober- und Unterbewusstsein." *Politisch-anthropologische Revue* 2 (1903–1904): 456–476.
"Die sexuale Reform." *Politisch-anthropologische Revue* 2 (1903–1904): 970–994.
Sexualethik. Wiesbaden, 1907.
"Leitziele zur Rassenbewertung." *Archiv für Rassen- und Gesellschaftsbiologie* 8 (1911): 59–71.
Kosmogonie. Jena, Germany, 1916.
"Sexualmoral der Zukunft." *Archiv für Rassen- und Gesellschaftsbiologie* 22 (1930): 292–304.
Philosophische Schriften. Munich: Philosophia Verlag, 1982.

WORKS ON EHRENFELS

Eaton, Howard O. *The Austrian Philosophy of Values.* Norman: University of Oklahoma Press, 1930.
Meister, Richard. "Ehrenfels." In *Neue deutsche Biographie,* Vol. IV, pp. 352–353. Berlin, 1959.
Orestano, Francesco. *I valori umani* (Vols. XII and XIII of his *Opere complete*), 2 vols. Vol. I, pp. 69–102, 123–126; Vol. II, pp. 46–101. Milan, 1942.
Smith, Barry, C. F. Ehrenfels, et al, eds. *Foundations of Gestalt Theory*. Munich: Philosophia Verlag, 1988.
Varet, Gilbert. *Manuel de bibliographie philosophique*. Vol. II, p. 877, note. Paris: Presses universitaires de France, 1956.

Klaus Hartmann (1967)

EINSTEIN, ALBERT

(1879–1955)

Albert Einstein was born in Ulm, in the south German kingdom of Württemberg on March 14, 1879. Following his graduation from the Federal Polytechnical Institute (*ETH*) in Zurich in 1900 he obtained a job as a patent examiner, ("technical expert, third class") in the Swiss patent office in Bern, starting in the summer of 1902. In January 1903 he married his first wife, Mileva Maric, a fellow student of physics at the *ETH* and, with Mileva's support, continued his investigations in physics, earning a PhD from the University of Zürich in 1905.

That was Einstein's "miracle year." In 1905 Einstein published the founding papers of the special theory of relativity, including a version of the famous $E = mc^2$. Also in 1905 he developed the light quantum hypothesis to treat the photo-electric effect, a work important in the subsequent development of the quantum theory and the official basis of his 1922 award of the Nobel Prize. There were also two papers on Brownian motion he produced that year, which helped demonstrate the reality of molecules.

Einstein left the patent office in 1909, moving to Berlin in 1914 to assume the directorship of the new Kaiser Wilhelm Institute for Physics. There his marriage quickly dissolved and his wife moved back to Zürich with their two sons, Hans Albert and Eduard. Einstein had been working on extensions of relativity since 1907 and in 1916 he published an account of what he called the "general theory" of relativity, which is essentially the modern theory of gravity. It predicted the bending of light rays around the sun. When that was confirmed during the solar eclipse of 1919, Einstein became a worldwide celebrity overnight, the first scientific superstar.

Einstein's celebrity status made him a target of growing German antiSemitism. His own interest was growing in Zionist and pacifist causes. Amidst this turmoil, in 1917 Einstein became ill and was cared for by his cousin Elsa Einstein Löwenthal, recently divorced and with two daughters, Ilse and Margot. Following his own divorce in 1919, he married Elsa, whose daughters also took the name Einstein.

In the period from 1914 to 1919 Einstein's scientific work continued to flourish. He began investigations into gravitational waves and cosmology, where he reluctantly introduced the cosmological constant, which he subsequently rejected, but which has come back to represent what now appears to be substantial density and pressure associated with empty space. Einstein also worked on statistical aspects of the quantum theory, developing the coefficients of spontaneous and induced emission and absorption that provided the theoretical opening for laser technology.

In the 1920s Einstein traveled extensively in aid of science and of Zionism. His scientific contributions slowed down in this period, although he made some preliminary attempts at tying together the electromagnetic and gravitational fields geometrically in a unified field theory. He made important contributions to the quantum theory of gases, developing Einstein-Bose statistics to treat radiation as a quantum gas of indistinguishable particles. This led to his discovery of the Bose-Einstein condensation, a low temperature phenomenon displaying quantum behavior at nearly macroscopic scale. In the 1927 Solvay conference Einstein began to "debate" with Niels Bohr over the foundations of the emerging quantum mechanics.

Einstein left Germany in 1932 for the Institute for Advanced Study in Princeton. He became a United States citizen in 1940. A year earlier he had signed a letter drafted by Leo Szilard advising President Franklin D. Roosevelt of the military potential of atomic energy. Later

he was an advocate for the control of atomic energy and for institutions supporting world peace. He was also a prominent critic of McCarthyism and a defender of civil liberties, as well as an outspoken opponent of racism and a defender of civil rights. Einstein died on April 18, 1955, of complications following a ruptured aortic aneurysm. His last scientific work was on the unfinished project for a unified field theory. His last phrase, written a few days before his death, was in a document with the pungent title, "Political Passions, Aroused Everywhere, Demand Their Victims."

PHILOSOPHY OF PHYSICS

Throughout his life Einstein read deeply in philosophy, where he was influenced both by David Hume and Immanuel Kant, as well as by Spinoza. His views in turn influenced the development of neo-positivism, whose more extreme doctrines he rejected in his criticism of the quantum theory. His philosophical reflections on the epistemology of science, as well as on metaphysical issues relating to space, time and causality, constitute an important chapter in twentieth century thought.

RELATIVITY. Einstein was a critic of the spatio-temporal framework of Newtonian physics, following a path marked out by Ernst Mach, who attacked Newton's introduction of "absolute" space and time not as unobservable (a fact advertised by Newton himself) but as unnecessary for doing physics. In Einstein's hands, however, Mach's critical method became a tool for positive theory construction.

Newton argued that acceleration must be absolute in order to explain inertial effects, such as the way water crawls up the sides of a rotating bucket. From absolute acceleration Newton moved (questionably, it turns out) to absolute space and time. Mach countered that inertial effects, like others, could be seen as purely relational. In particular, if one could rotate large enough masses and leave the bucket alone the same water-crawling effects would occur. This idea came to be known as Mach's Principle, which strongly influenced Einstein's development of the general theory of relativity. Sympathetic to Mach's relational conception of space and time, Einstein criticized an asymmetry built into the Newtonian framework. There space and time affect the behavior of bodies in so far as, in the absence of impressed forces, bodies move inertially, along spatially straight lines with temporally constant speeds. But there is no reciprocity. If space and time are absolute, bodies cannot affect spatio-temporal structure. Once space and time were merged into a unified spacetime, the field equations of general relativity allowed a reciprocal interplay between spacetime and matter.

The introduction of four dimensional spacetime, however, comes from the 1905 special theory of relativity. There Einstein dealt with an apparent conflict between the principle of relativity (any inertial frame is suitable for the representation of electrodynamic as well as mechanical phenomena) and the constancy of the speed of light for inertial observers. In the 1905 paper Einstein approaches this conflict by applying a technique of conceptual analysis that he learned from Mach (and from David Hume). He asks what is time and quickly shifts, epistemologically, to how one tells time in reading a clock. Telling time involves a spatially local judgment of simultaneity (where are the hands, when?); that is, it involves events in more or less the same place. What about events that happen very far apart?

The suggestion is that here one reaches the limit of applicability of the concept of simultaneity. In Mach's hands (or Hume's) one might stop here, with skepticism about the very meaningfulness of assertions of distant simultaneity. But, as Einstein commented later, although he respected Mach's hobbyhorse of seeking the limits of concepts he felt that it does not give rise to anything living. To employ conceptual analysis constructively one needs a theory. In the 1905 paper that theory is grounded on a quasi-operational definition using light signals to determine when distant events happen at the same time. Armed with that definition of simultaneity one can not only reconcile the principle of relativity with the constancy of the velocity of light, one can go on to develop a spacetime framework in which descriptions of events in different inertial frames are tied to one another by Lorentz transformations that leave the so-called spacetime "interval" invariant. Einstein had wanted to refer to this work as a theory of invariants. Ironically, Max Planck coined the term "relativity," and it stuck.

One of the conceptual innovations in special relativity is the variation of relativistic mass with velocity, which no longer appears to be a constant property of matter. This shift in the conception of *mass* prompted Thomas Kuhn and Paul Feyerabend to feature an "incommensurability" between Newtonian and relativistic physics. Einstein was unequivocally against the idea that the so-called "relativistic mass" is a proper notion at all. He rejects it as coordinate dependent and, hence, merely perspectival and thinks "the—unhappily—often mentioned concept of a mass which depends on speed is quite misleading." Instead, in keeping with his emphasis on invariance (as a

touchstone for scientific objectivity), Einstein says, "It is better to use the word mass exclusively for rest mass … which is always the same, independent of the speed …" (Earman and Fine 1977, p. 538). It is worth noting that the mass term m in $E = mc^2$ denotes precisely the rest mass.

Einstein was an early supporter of logical empiricism. He was also one of its icons, in part because his positivistic analysis in special relativity seemed evident also in the general theory. In his 1916 account, Einstein defends the relativity of all motion (not just inertial) by requiring that laws of nature be expressed by equations "valid for all coordinate systems." Called general covariance, this requirement, he says, "takes away from space and time the last remnant of physical objectivity." (Einstein 1987, Vol. 6 [1996], p. 287 and 291). In support he appears to offer a straightforwardly verificationist analysis. "All our space-time verifications invariably amount to a determination of space-time coincidences." (Einstein 1987, Vol. 6 [1996], p. 287 and 291). Thus it is only space-time coincidences ("to which all our physical experience can ultimately be reduced"), that is, the coordinate systems, that the laws of nature need respect.

Recent scholarship suggests that this positivist reading is mistaken (Einstein 1987, Vol. 6 (1996), p. 287 and 291). For in these passages Einstein is probably reacting to an earlier argument of his own (called the "hole argument") posing a conflict between general covariance and determinism (Einstein 1987, Vol. 6 (1996), p. 287 and 291). The key to unraveling that argument was his recognition that, by themselves, coordinates (the bare mathematical points) have no physical significance. Significance comes from the fields of the theory, as determined from given sources by the theory's field equations. That's what makes space-time coincidences observable. Einstein later held that, in general, scientific theory determines what one can observe. Thus the positivist reading has things exactly back to front. Whereas in special relativity Einstein follows a positivist line in grounding theoretical notions (simultaneity) in what is observable, here he entheorizes the observable and takes an anti-positivist line in grounding the observable in the theory itself.

QUANTUM THEORY. Einstein made fundamental contributions to the early understanding of quantum phenomena and his ideas, which emphasized the problem of wave-particle duality, influenced all subsequent developments. However Einstein became the foremost critic of the quantum mechanics that emerged from 1926 to 1930.

His dissatisfaction is often portrayed as a last ditch longing for determinism or causality ("God does not throw dice"), as against the essentially probabilistic character of quantum physics. To be sure, although Einstein was a master at statistical physics, he was certainly troubled by a science where probability occurs fundamentally. Nevertheless his problem with the quantum theory was not about determinism alone, nor even primarily about determinism at all. In a 1930s letter to his old friend and translator, Maurice Solovine, Einstein expresses his concerns this way. "I am working with my young people on an extremely interesting theory with which I hope to defeat modern proponents of probability-mysticism and their aversion to the notion of reality in the domain of physics" (Solovine 1987, p. 91). This is a typical linkage in Einstein's thought. In almost every context in which Einstein expresses reservations about quantum indeterminism he couples it with reservations about the irrealism of the theory; that is, giving up the ideal of treating individual events, or what he referred to as real states of affairs.

As usually understood, the quantum theory does not treat real states of affairs at all, not even probabilistically. It does not tell us whether an electron is likely (even) to be here or there, spinning up or down. Quantum theory only gives the probability for finding the electron here, or finding it spinning up, if one actually measures it for that particular property. This is the irrealism that Einstein found so disturbing. That there could be laws, even probabilistic laws, for finding things if one looks, but no laws of any sort for how things are independently of whether one looks, was mysticism, a "mindless" (1987, p. 119) form of empiricism.

Einstein responded with a program just as in the development of relativity. First he set out to establish the limitations of the concepts used in the quantum domain and then he explored the possibility of transcending those limitations with a positive theory. He began by challenging the uncertainty formulas. He accepted that they limit the simultaneous, precise measurement of conjugate quantities (like position and linear momentum) but he questioned the ontological reading in which they limit what is simultaneously real. He went on to examine the rationale offered, especially by Bohr, both for the statistical character of the quantum theory and for its irrealism.

Bohr postulated an uncontrollable interaction introduced in every act of measurement that, he argued, made a statistical treatment necessary and also prevented states of affairs being defined independently of the measurement. In a series of thought experiments Einstein devel-

oped the concept of indirect measurement as a challenge to Bohr's postulate. This culminated in a 1935 paper, co-authored with his research assistants Boris Podolsky and Nathan Rosen, and composed by Podolsky. Usually referred to as *EPR* this paper involved the idea that Schrödinger dubbed "entanglement" (*Verschränkung*).

Entanglement occurs when, after quantum systems interact, certain quantities become linked among the systems. In the *EPR* case, for a pair *A*, *B* of previously interacting particles—now far apart—both position and momentum are so linked that determining the position of one automatically determines the position of the other, and similarly for momentum. By directly measuring, say, the position of *A* one can determine *B*'s position and apparently without any "uncontrollable interaction" or disturbance of *B*, contra Bohr's postulate. Moreover by assuming a principle of local action according to which, provided the systems are sufficiently far apart, the "reality" at *B* is not affected by the measurement carried out at *A*, it follows that the position determined for *B* must have been *B*'s all along. Thus, contrary to Bohr, one can define a coordinate of position for *B* that is independent of measurement or observation there – a real state of affairs. Unfortunately, in *EPR* it is difficult to track these considerations clearly. It appears that Einstein never saw Podolsky's text before publication. When he did he expressed misgivings that it obscured his central concerns.

EPR has been seen as suggesting the possibility of a "hidden variables" account of quantum theory, an account that would introduce simultaneous values for both position and momentum, along with other quantities, and still, somehow, respect the uncertainty relations. But Einstein, who had toyed with and abandoned a hidden variables approach in 1927, was never again interested in such an account. In the context of *EPR*, he told Schrödinger explicitly that he "couldn't care less" about simultaneous values for position and momentum (Fine 1996, p. 38). In fact Einstein thought these point-particle concepts were not appropriate for the quantum domain. He hoped to introduce different concepts and explored how they would emerge from a unified field theory. Einstein pursued that quest unsuccessfully for many years. In the end he questioned whether even a field theory would do the job and speculated about the need for a purely algebraic kind of physics, one not based on a spatio-temporal continuum. He sometimes despaired, however, that this was like trying to breathe in empty space.

GENERAL PHILOSOPHY OF SCIENCE

Einstein's attempt to develop new concepts for the quantum domain accords with his anti-inductivist principle that ideas (or concepts) are free creations. By "free" he meant both that concepts are not innate and also that they are neither given in nor logically derived from experience. The only test for scientific concepts is whether they can be organized in a logically simple system that finds fruitful empirical applications. This highlights logical simplicity as a paramount factor in theory choice. It also represents a holistic attitude to theories, gleaned perhaps from Pierre Duhem. Holism is apparent in Einstein's acute analysis of the testability of geometry where, while rejecting Henri Poincaré's conventionalist defense of Euclidean physical geometry, he ultimately agrees with Poincaré that only the whole system of physics plus geometry is testable.

Einstein's work in relativity and his project for a unified treatment of gravity and quantum phenomena shows unification as central to his scientific outlook. His study of Baruch Spinoza, whom he read and re-read over the years may have influenced this attitude (or reinforced it). Certainly realism was another central feature. This is evident in his introduction of the light quantum and in his use of the kinetic-molecular picture in treating Brownian motion. It is evident as well in his worries over the instrumentalist understanding of the quantum theory. Nevertheless he ridiculed "assertions" of realism as meaningless, like chiming "cock-a-doodle-doo." For Einstein realism was not a doctrine but rather a motivational program. The program was to develop scientific theories that describe individual events themselves, without reference to conditions of observation. That is what he believed science had always done, and with great success. It was motivational because, at the personal level, he thought individuals would have no motivation to pursue science unless they felt that in doing so they were unlocking the secrets of nature. Clearly this program conflicts with the enormously successful but irrealist quantum theory, which is why Einstein struggled to make room for the possibility, at least, of a realist reinterpretation.

Determinism (or causality—he hardly draws a distinction) is another important item in Einstein's outlook. Here, again, he did not advocate a doctrine like, "The world is deterministic." Characteristically, he favored a program to entheorize determinism; that is, to build deterministic theories. His reaction to the dilemma between determinism and general covariance posed by the hole argument shows this concern, as does his sense that the probabilistic quantum theory involves a retreat

into statistics. Nevertheless, in reluctantly accepting that one might have to move to an algebraic physics, he did acknowledge that science might abandon the ideal of representing events in spacetime altogether, and hence move beyond causality (or determinism).

Einstein's views are sometimes described in terms of the philosophical "isms": holism, realism, determinism and so forth. While there can be some truth to these descriptions (provided one entheorizes them), he generally regarded philosophical positions pragmatically. He saw them as tools that may be useful at certain moments for building better scientific theories, judged by the criterion of empirical success. His sometimes strong statements for or against one of the "isms" are best be seen in the terms of the dialogism described by Mara Beller in her *Quantum Dialogue*, a dialectical view that highlights the creative role of scientific disagreement in shifting contexts. Einstein himself described it this way:

> I do not feel comfortable and at home in any of the "isms." It always seems to me as though such an ism were strong only so long as it nourishes itself on the weakness of its counter-ism; but if the latter is struck dead, and it is alone on an open field, then it also turns out to be unsteady on its feet. *So, away with the squabbling*.
>
> HOWARD 1993, P. 225

See also Quantum Logic and Probability; Relativity Theory; Space; Space in Physical Theories; Time.

Bibliography

WORKS BY ALBERT EINSTEIN

"Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" With Boris Podolsky and Nathan Rosen. *Physical Review* 47 (1935): 777–780.

The Evolution of Physics. With L. Infeld. New York: Simon & Schuster, 1938.

Relativity: The Special and General Theory. 15th ed. London: Methuen, 1938.

Ideas and Opinions. New York: Bonanza Books, 1954.

The Collected Papers of Albert Einstein. Princeton, NJ: Princeton University Press, 1987. Ongoing project that will include most of Einstein's writing and correspondence.

Letters to Solovine. Translated by Wade Baskin. New York: Philosophical Library, 1987.

Out of My Later Years. New York: Bonanza Books, 1990.

WORKS ABOUT ALBERT EINSTEIN

Beller, Mara. *Quantum Dialogue: The Making of a Revolution*. Chicago: University of Chicago Press, 1999.

Earman, John, and Arthur Fine. "Against Indeterminacy." *Journal of Philosophy* LXXIV (1977): 535–538.

Einstein Archives Online. Available from http://www.alberteinstein.info/.

Fine, Arthur. *The Shaky Game: Einstein, Realism and the Quantum Theory*. 2nd ed. Chicago: University of Chicago Press, 1996.

Howard, Don A. "Was Einstein Really a Realist?" *Perspectives on Science* 1 (1993): 204–251.

Norton, John D. "General Covariance and the Foundations of General Relativity: Eight Decades of Dispute." *Reports on Progress in Physics* 56 (1993): 791–858.

Pais, Abraham. *Subtle Is the Lord: The Science and the Life of Albert Einstein*. New York: Oxford University Press, 1982.

Ryckman, Thomas. *The Reign of Relativity: Philosophy in Physics 1915–1925*. Oxford, U.K.: Oxford University Press, 2005.

Schilpp, Paul Arthur, ed. *Albert Einstein: Philosopher-Scientist*. La Salle, IL: Open Court, 1949.

Arthur Fine (2005)

ELIMINATIVE MATERIALISM, ELIMINATIVISM

"Eliminative materialism" espouses the view that our commonsense way of understanding the mind is false, and that, as a result, beliefs, desires, consciousness, and other mental events used in explaining our everyday behavior do not exist. Hence, the language of our "folk" psychology should be expunged, or eliminated, from future scientific discourse.

Two routes have been taken to get to the eliminativist's position. The first and less popular stems from a linguistic analysis of mentalistic language. Paul Feyerabend argues that the commonsense terms for mental states tacitly assume some version of dualism. Insofar as materialism is true, these terms cannot refer to anything in the physical world. Thus they should not be used in discussing ourselves or our psychologies since we are purely physical beings.

The second and better-developed approach comes out of the philosophies of science developed by Feyerabend, David Lewis, Willard Van Orman Quine and Wilfrid Sellars. Two suppositions are important for eliminativism. (1) There is no fundamental distinction between observations (and our observation language) and theory (and our theoretical language), for previously adopted conceptual frameworks shape all observations and all expressions of those observations. All observations are "theory-laden." These include observations we make of ourselves; in particular, observations we make

about our internal states. There are no incorrigible phenomenological "givens." (2) The meaning of our theoretical terms (which includes our observational vocabulary) depends upon how the terms are embedded in the conceptual scheme. Meaning holism of this variety entails that if the theory in which the theoretical terms are embedded is false, then the entities that the theory posits do not exist. The terms would not refer.

Two more planks complete the eliminative argument. (3) Our way of describing ourselves in our everyday interactions comprises a rough and ready theory composed of the platitudes of our commonsense understanding. The terms used in this folk theory are defined by the platitudes. (4) Folk psychology is a radically false theory.

In support of this position. Patricia Churchland and Paul Churchland argue that belief-desire psychology wrongly assumes sentential processing; moreover, belief-desire psychology is stagnant, irreducible to neuroscience, and incomplete. Stephen Stich argues that our very notion of belief and, by implication, the other propositional attitudes is unsuitable for cognitive science. Patricia Churchland, Daniel Dennett, Georges Rey, Richard Rorty, and others argue that our notion of consciousness is confused. They all conclude, as do other eliminativists, that folk psychology should be replaced by something entirely different and more accurate, though views differ on what this replacement should be.

Attacks on eliminative materialism generally have come from four fronts, either on premise two, premise three, or premise four of the second approach, or on the eliminativist position itself, without regard to the arguments for it. Premise two asserts meaning holism and a particular theory of reference. If that theory were false, then the eliminativist's second argument would be undermined. There are alternative approaches to reference that do not assume holism; for example, causal-historical accounts do not. If meaning is not holistic, then even if folk psychology were incorrect, the terms used in that theory could still refer, and elimination of folk psychological terms would not be warranted.

Arguments that our folk psychology is not a true theory deny premise three. Here some detractors point out that even if a completed psychology did not rely on the propositional attitudes or consciousness, that fact would not entail that those sorts of mental states do not exist; instead, they just would not be referred to in scientific discourse. Nevertheless, they could still be used as they are now, in our everyday explanations of our behavior.

Others charge that premise four is false; folk psychology might be a rudimentary theory, but it is not radically false. While agreeing that belief-desire explanations or explanations involving conscious events might not be entirely empirically adequate or complete, champions of folk psychology argue that no other theory is either. In addition, our folk psychology has developed over time, is coherent, and its status with respect to neuroscience is immaterial. These arguments are generally coupled with the claim that no other alternative, either real or imagined, could fulfill the explanatory role that the propositional attitudes play in our understanding of ourselves. And until the eliminativist's promise of a better conceptual scheme is fulfilled folk psychology is here to stay. At least some properly revised version of folk psychology would remain.

Lastly, some supporters of folk psychology argue that any eliminativist program would be fatally flawed, regardless of whatever particular arguments are given, for the very statement of eliminative materialism itself is incoherent. In its simplest form, the argument runs as follows: Eliminative materialism claims that beliefs do not exist. Therefore, if eliminative materialism were true, we could not believe it. Therefore, no one can believe eliminative materialism on pain of inconsistency.

Replies to the four sorts of attacks are ubiquitous. However, answering the first three turns on (primarily empirical) issues yet to be settled. Which theory of reference is correct, whether folk psychology is actually a theory, and what revisions are required to make it adequate depend upon facts we do not yet know about ourselves or our linguistic practices.

The last point is more conceptual. In responding to it, eliminative materialists hold that something else will replace "belief," or some instances or aspects of "belief." Call this "schmelief." It is true that eliminative materialists cannot believe that eliminative materialism is true on pain of inconsistency. But, eliminativists maintain, they can "schmelieve it." Defenders of a revised folk psychology answer that, as used in this context, "schmelief" seems to be some other intentional operator or relation, a mere revision of belief. Without better exposition of what the replacement for folk psychology will be (and how it will be radically different), we simply cannot tell what the future holds for our commonsense theory of self: simple revision, peaceful coexistence, or outright replacement.

See also Consciousness; Dennett, Daniel C.; Folk Psychology; Lewis, David; Materialism; Philosophy of

Mind; Quine, Willard Van Orman; Reference; Rorty, Richard; Sellars, Wilfrid.

Bibliography

Baker, L. R. *Saving Belief: A Critique of Physicalism.* Princeton, NJ: Princeton University Press, 1987.

Boghossian, P. "The Status of Content." *Philosophical Review* 99 (1990): 157–184.

Churchland, P. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78 (1981): 67–90.

Feyerabend, P. "Mental Events and the Brain." *Journal of Philosophy* 60 (1963): 295–296.

Horgan, T., and J. Woodward. "Folk Psychology Is Here to Stay." *Philosophical Review* 94 (1985): 197–225.

Rey, G. "A Reason for Doubting the Existence of Consciousness." In *Consciousness and Self-Regulation,* edited by R. Davidson, S. Schwartz, and D. Shapiro, Vol. 3. New York: Plenum Press, 1982.

Rorty, R. "Mind-Body Identity, Privacy, and Categories." *Review of Metaphysics* 19 (1965): 24–54.

Stich, S. *From Folk Psychology to Cognitive Science: The Case against Belief.* Cambridge, MA: MIT Press, 1983.

Valerie Gray Hardcastle (1996)

ELIOT, GEORGE

(1819–1880)

Born Marian (or Mary Ann) Evans, George Eliot was the assumed name of the English novelist, poet, essayist, and translator. She was reared near Coventry and in her early years attended a school run by a fervent evangelical mistress. From this woman she acquired intense religious beliefs, but she gradually lost her faith. In 1842 she wrote that she thought Christian dogmas "dishonorable to God" and pernicious to human happiness. Within a few months, however, she had come to regard the dogmas in themselves as of little importance. "Speculative truth begins to appear but a shadow of individual minds, agreement between intellects seems unattainable, and we turn to the *truth of feeling* as the only universal bond of union," she wrote in a letter in October 1843; a belief in the importance of feeling remained central to her life and work.

In Coventry she had a group of friends with literary and philosophical interests, and under their influence she undertook, in 1844, a translation of D. F. Strauss's *Das Leben Jesu*; the translation was published in 1846. She went to London in 1851 to work for John Chapman as assistant editor of the *Westminster Review.* She published occasional essays and read much. Among her numerous friends in London were Herbert Spencer, to whom she was falsely rumored to be engaged, and George Henry Lewes, the philosopher and critic. Lewes was married but separated from his wife. In October 1854 Eliot and he decided to live together. They never married, but they lived a life of exemplary domesticity until Lewes's death, in 1878. On May 6, 1880, to everyone's surprise, she married John W. Cross, long a family friend. She died that same year, after a short illness.

In 1854 Eliot's translation of Ludwig Feuerbach's *Das Wesen des Christentums* was published. She also translated Benedict (Baruch) de Spinoza but did not publish the translation. Upon Lewes's urging, she tried writing fiction; her first story was published in *Blackwood's Magazine* in 1857. She was immediately successful as a writer of fiction. To her fiction—notably *Adam Bede* (1859), *The Mill on the Floss* (1860), *Silas Marner* (1861), *Middlemarch* (1871–1872), and *Daniel Deronda* (1876)—rather than to her poetry or her essays, she owed her fame and her considerable influence as a moral teacher.

Eliot's views on moral, religious, and metaphysical problems pervade and profoundly shape her writings, but they are never presented in abstract, systematic form. She had no faith in general moral principles: "to lace ourselves up in formulas," she wrote, is to repress the "promptings and inspirations that spring from growing insight and sympathy." Like Strauss, Feuerbach, and Auguste Comte, she thought of religious and metaphysical doctrines as projections and symbols of feelings, and as valuable only to the degree that the feelings they express and reinforce are valuable. Her "most rooted conviction," she told a friend in 1859, was that "the immediate object and the proper sphere of all our highest emotions are our struggling fellow-men in this earthly existence," and she declared that one of her main aims in her writing was to show that human fellowship does not depend on anything nonhuman. Christianity can foster many valuable emotions, she held, but the insistence of some Christians that all action must be for the glory of God stifles benevolence and love and directs feelings away from men. The idea of God has been beneficial only insofar as it has been "the ideal of a goodness entirely human."

Eliot thus belongs with those Victorian writers who tried, in different ways, to work out a humanistic morality capable of satisfying the deep human needs that they thought the older, religiously based morality could no longer satisfy. Her view is naturalistic and deterministic; men are seen as being as much under the dominion of the laws of nature as are other parts of the world, though the comparisons are usually with organic growth and decay rather than with purely mechanical processes. Hereditary

and social influences on character are heavily emphasized, as is the effect one's repeated actions or evasions will have on one's own character and hence on one's future actions.

The morality that springs from this view is primarily one of sympathy and compassion. The complexity and obscurity of motives and the mixture of good and evil in personality and in deed are constantly displayed in the novels. It is usually difficult, Eliot suggested, to know what one ought to do in particular cases; one must rely ultimately on one's deepest feelings when these are enlightened by sympathy and by knowledge of circumstances and consequences. Wrongdoing is usually traced to stupidity, callousness, or thoughtlessly excessive demands for personal satisfaction, rather than to deliberate malice or conscious selfishness. Vice and crime are shown as eventually bringing retribution, but the reward of virtue is at best the peace that comes with acceptance of one's lot. Eliot saw quiet renunciation and patient selflessness as the chief virtue. She frequently traced the career of an unusually sensitive and intelligent person who hopes to do great things for others but after painful defeats ends by settling into a life of unheroic and routine benevolence. She suggested that this is the only feasible way of achieving lasting good. In the thought that what we do will have some good effect on future generations and we shall be remembered by them with love, she held, there was a sufficient motive to virtue and a sufficient replacement of the belief in personal immortality and personal reward.

See also Comte, Auguste; Feuerbach, Ludwig; Religion and Morality; Spinoza, Benedict (Baruch) de; Strauss, David Friedrich.

Bibliography

Two essays reprinted in George Eliot's *Essays and Leaves from a Notebook*—"Evangelical Morality: Dr. Cummings" (1855) and "The Poet Young" (1857)—are especially relevant to her moral views.

The standard biography is John W. Cross, *The Life of George Eliot,* 3 vols. (London, 1885–1887); it is composed mainly of her letters, heavily censored. *Marian Evans and George Eliot,* by Lawrence and Elisabeth Hanson (London: Oxford University Press, 1952), is more accurate and contains a good bibliography. *The George Eliot Letters,* edited by Gordon Haight, 7 vols. (New Haven, CT: Yale University Press, 1954–1955), is a masterpiece of scholarship. Two essays by R. H. Hutton, reprinted in his *Modern Guides to English Thought* (London: Macmillan, 1887), give an assessment by a younger contemporary from an orthodox Christian standpoint.

There are numerous studies of Eliot's life, intellectual development, and writings. See especially Joan F. Bennett, *George Eliot: Her Mind and Her Art* (Cambridge, U.K.: Cambridge University Press, 1948); Gordon Haight, *George Eliot and John Chapman* (New Haven, CT, and London, 1940); Barbara Hardy, *The Novels of George Eliot* (London: University of London, Athlone Press, 1959); and Leslie Stephen, *George Eliot* (London, 1902).

J. B. Schneewind (1967)

ELIOT, THOMAS STEARNS

(1888–1964)

Thomas Stearns Eliot is best known as a poet and literary critic (he received the Nobel Prize for literature in 1948), but his work in social and cultural theory has also been widely influential. His principal works of this kind are *After Strange Gods* (London, 1934), *The Idea of a Christian Society* (London, 1939), and *Notes Towards the Definition of Culture* (London, 1949).

Eliot was born in St. Louis but lived in London from 1915 on and became a British subject in 1927. He was graduated from Harvard University in 1909 and engaged in advanced studies in philosophy there, at the Sorbonne, and at Oxford until 1915. In the year 1913/1914 he served as an assistant in philosophy at Harvard, studying methodology with Josiah Royce and logic with Bertrand Russell. Eliot and Russell, despite enormous differences in political, social, and religious outlooks, became close friends. Eliot's Harvard doctoral dissertation, completed at Oxford in 1915, was published as *Knowledge and Experience in the Philosophy of F. H. Bradley* (London and New York, 1964). Francis Herbert Bradley's idealism influenced Eliot's critical doctrines, and in 1926 Eliot published an essay on Bradley, reprinted in *Selected Essays* (London, 1951). In this essay he praised especially Bradley's critique of utilitarianism: "He replaced a philosophy which was crude and raw and provincial by one which was, in comparison, catholic, civilized, and universal." But even before completing his studies, Eliot had finished some of his finest early poems, and he never produced any technical philosophical studies aside from his thesis.

In his early poetry and criticism, Eliot was a considerable innovator, but it was a main goal of his experiments to try to recover the sense of a fruitful tradition. In particular, this meant rejecting the literary theory and practice of romanticism and finding earlier sources. In a famous comment in 1921, he argued that there had been, in the seventeenth century, a major change in the English

mind, which he called the "dissociation of sensibility"—the separation of feeling and thought. He came later to stress a loss of a sense of order, both internal and external, and to associate it with the decline of the Christian and classical cultural framework. To counteract this loss, the poet and critic must strive to recover a sense of the whole European tradition. At the end of this phase of his development Eliot described himself as a classicist, and he was to write henceforward as a declared and orthodox Christian.

After Strange Gods is the bridge from his mainly literary to his mainly social and cultural criticism. The book's subtitle is *A Primer of Modern Heresy.* Its argument is that modern writers, deprived of tradition, have constructed private or esoteric systems of belief, and, deprived of a common language and imagery, they have been forced to experiment. The struggle for common meanings, always difficult, is now even more difficult. This failure of communication is profoundly damaging to the whole society. The writer's task is to develop the full potential maturity of the language of his society. Paradoxically, therefore, the most creative work is that which begins from and is most aware of the full tradition and history of the language in which it is written. The loss of this tradition makes the modern writer's task overwhelmingly difficult.

In *The Idea of a Christian Society,* Eliot applied and extended this argument to social questions. He argued that the Western democracies, although nominally Christian, in fact live by quite other values. The idea of a Christian society is at best an understanding of the social ends that would deserve the name of Christian, but in the modern world there is an unusually wide gap between such ends and the main principles of social organization. Many of the driving forces of modern society—especially its false emphasis on profit, its substitution of exploitation of men and things for right use, and its general adoption of commerce as the central human concern—are in fact hostile to any Christian life in the world. It is therefore not surprising, Eliot claimed, that society is far from being Christian; what is surprising is that people retain as much Christianity as they do.

In *Notes towards the Definition of Culture,* Eliot's most substantial theoretical work, he distinguished three senses of "culture"—the culture of the individual, of the group, and of the whole society. He argued that it is false to set as the goal of the group what can be the aim of the individual alone, and to set as the goal of the whole society what can only be the aim of a group. This argument became Eliot's main theoretical justification for what is ordinarily called "minority culture," and for his critique of egalitarian doctrines in education: It is false to educate the whole society to perform the cultural tasks of a particular group. At the same time, culture in each sense is necessarily connected with culture in the other senses. The group depends on the whole way of life of the society, as social organization depends upon tradition. Likewise, the culture of the individual cannot be isolated from the culture of the group.

Eliot further emphasized the extent to which the culture of a whole society is a matter of custom and behavior and is often unconscious: It is all the characteristic interests and activities of a people, whether or not some of these are thought of as "culture" in the narrower sense. What is often called "culture"—religion, arts, laws, and intellectual activity—is the conscious expression of the total culture, the whole way of life.

It follows from this, Eliot argued, that the maintenance and extension of the conscious culture of a society cannot be delegated to an elite, a group of specialists selected by merit. However skilled an elite may be in the special activities themselves, its members will necessarily lack the continuity with the rest of the society that is ultimately necessary for the health of the conscious culture. An elite, newly selected in each generation, will inevitably lack a sense of tradition. Eliot therefore saw no alternative to the maintenance of classes in society, and in particular to the maintenance of a governing class with which the specialists will overlap and interact. The need for continuity in culture, and for a tradition as opposed to a group of specialists with unrelated skills, argues, finally, for a social conservatism that will keep a proper relationship between continuity and change. This last phase of Eliot's social thinking has been especially influential since World War II.

See also Belief; Bradley, Francis Herbert; Philosophy of Social Sciences; Social and Political Philosophy; Royce, Josiah; Russell, Bertrand Arthur William; Utilitarianism.

Bibliography

Brooker, Jewel Spears. *Mastery and Escape: T.S. Eliot and the Dialectic of Modernism*. Amherst: University of Massachusetts Press, 1994.

Childs, Donald J. *From Philosophy to Poetry: T.S. Eliot's Study of Knowledge and Experience*. New York: Palgrave, 2001.

Cooper, John Xiros. *T. S. Eliot and the Ideology of Four Quartets*. Cambridge, U.K.: Cambridge University Press, 1995.

Dale, Alzina Stone. *T.S. Eliot: The Philosopher Poet*. Wheaton: H. Shaw, 1988.

Freed, Lewis. *T. S. Eliot: Aesthetics and History*. La Salle: Open Court, 1962.
Freed, Lewis. *T. S. Eliot: The Critic as Philosopher*. West Lafayette: Purdue University Press, 1979.
Gardner, Helen. *The Art of T. S. Eliot.* London: Cresset Press, 1949.
Habib, Rafey. *The Early T. S. Eliot and Western Philosophy*. Cambridge, U.K., and New York: Cambridge University Press, 1999.
Jain, Manju. *T. S. Eliot and American Philosophy: The Harvard Years*. Cambridge, U.K.: Cambridge University Press, 1992.
Kearns, Cleo McNelly. *T. S. Eliot and Indic Traditions: A Study in Poetry and Belief*. Cambridge, U.K.: Cambridge University Press, 1987.
Leavis, F. R. *The Common Pursuit.* London: Chatto and Windus, 1952. See "Mr. Eliot, Mr. Wyndham Lewis and Lawrence" and "Approaches to T. S. Eliot."
Lockerd, Benjamin G. *Aethereal Rumours: T. S. Eliot's Physics and Poetics*. Lewisburg: Bucknell University Press; London; Cranbury, NJ: Associated University Presses, 1998.
Lucy, Seán. *T. S. Eliot and the Idea of Tradition.* London: Cohen and West, 1960.
Moody, David. A., ed. *The Cambridge Companion to T. S. Eliot*. Cambridge, U.K.: Cambridge University Press, 1994.
Rajan, B., ed. *T. S. Eliot: A Study of His Writings by Several Hands.* London; Dennis Dobson, 1947.
Roeffaers, Hugo. *Eliot's Early Criticism: Philosophical Explorations into the Sacred Wood*. Frankfurt am Main; New York: P. Lang, 1988.
Schuchard, Ronald. *Eliot's Dark Angel: Intersections of Life and Art*. New York: Oxford University Press, 1999.
Shusterman, Richard. *T. S. Eliot and the Philosophy of Criticism*. New York: Columbia University Press, 1988.
Skaff, William. *The Philosophy of T. S. Eliot: From Skepticism to a Surrealist Poetic, 1909–1927*. Philadelphia: University of Pennsylvania Press, 1986.
Spurr, David. *Conflicts in Consciousness: T. S. Eliot's Poetry and Criticism*. Urbana: University of Illinois Press, 1984.
Warner, Martin. *A Philosophical Study of T. S. Eliot's Four Quartets*. Lewiston: Edwin Mellen Press, 1999.
Williams, Raymond. *Culture and Society.* London: Chatto and Windus, 1958. Part 3, Ch. 3.

Raymond Williams (1967)
Bibliography updated by Desirae Matherly Martin (2005)

ELISABETH, PRINCESS OF BOHEMIA

(1618–1680)

Elisabeth Simmern van Pallandt was born in Heidelberg on December 26, 1618, the third child and eldest daughter of Frederick V of Bohemia and Elisabeth Stuart, daughter of James I of England. Her parents' marriage represented the rising political strength of Protestantism. In August of 1620, Elisabeth's father, Frederick, departed Heidelberg for Prague to assume the position of Emperor of the Holy Roman Empire. In November 1620, Frederick lost the battle of White Mountain and with it his empire; he was forced into exile. This event is usually taken as the onset of the Thirty Years War. In the late 1620s Elisabeth joined her parents in The Hague. There, she was tutored by the Dutch humanist Constantijn Huygens and the mathematician Johan Stampioen. She also interacted with Anna Maria van Schurman. She was accomplished in Greek, Latin, German, English and French. Throughout her life, she was involved in her family's political affairs. In 1660, Elisabeth entered the Lutheran convent at Herford in the Rhine Valley. She died on February 8, 1680, as abbess of the convent.

Several of her siblings were accomplished as well. Her older brother Charles Louis rehabilitated Heidelberg University after the Thirty Years War. Her brother Rupert was known for his chemical experiments, his soldiering, and his role in founding the Hudson's Bay Company. Her sister Louise Hollandine was an accomplished painter. Her youngest sister Sophie through her marriage became Electress of Hanover and corresponded with Leibniz and Diderot among others.

In 1643 Elisabeth began a correspondence with René Descartes that continued until Descartes's death in 1650. This exchange constitutes the whole of Elisabeth's extant philosophical work. However, record of Elisabeth's intellectual interests predates this correspondence. Edward Reynolds dedicates his *Treatise of the Passions and the Faculties of the Soule of Man* to Elisabeth, suggesting that she had seen and commented on a draft manuscript. In the 1660s the British mathematician John Pell contacted Elisabeth, through Theodore Haak, regarding her solution to Appolonius's Problem (that of finding a fourth circle whose circumference touches three given circles) undertaken in her correspondence with Descartes. In the 1670s, after Elisabeth had become abbess at Herford, she was contacted by English Quakers and corresponded with William Penn and Robert Barclay. She was also in contact with Nicholas Malebranche, Francis Mercury van Helmont, and G. W. F. Leibniz.

In the seven years of their correspondence, Elisabeth and Descartes address the full scope of philosophical inquiry. They discuss metaphysics, as well as topics in natural philosophy, including physics, geometry, and medicine. Equally, their exchange includes discussions of moral psychology, ethics, and political philosophy. Because all we have of Elisabeth's philosophical writings are her letters to Descartes, and those letters principally involve reactions to his work, it is hard to determine Elis-

abeth's own positions. Nonetheless, by considering the presuppositions of her questions and objections, it is possible to adduce her philosophical commitments.

Elisabeth, in her letter of May 6, 1643, begins the exchange by asking Descartes how the two really distinct substances of mind and body can causally interact with one another to effect voluntary action. That is, she poses the problem of mind-body interaction. Elisabeth's problem lies in understanding the nature of the causation at work between an immaterial substance (mind) and a material one (body). It is clear from her posing of the question, and her subsequent pressing of Descartes about his answers, that Elisabeth is willing to accept only efficient causal explanations of mind-body interaction. Insofar as she is skeptical that any such explanation can be offered of the interaction between an immaterial mind and body, she is inclined to think that the mind is material, but nonetheless has a capacity for thought.

Elisabeth's questions about mind-body interaction demonstrate her commitment to a mechanist account of the natural world and shows her to be well-versed in the varieties of mechanist accounts of causation available to adopt. This interest in natural philosophy is perhaps best reflected in Descartes's dedication to her of his *Principles of Philosophy*, the work in which he lays out his physics most clearly. It is also reflected in her remarks regarding human physiology and observed natural phenomena later in the correspondence.

In 1645, in part to help Elisabeth find some comfort from the effects of the English Civil War on her family, Descartes undertook to outline his views on moral psychology—the regulation of the passions—and the nature of the sovereign good. For him, the sovereign good consists simply in virtue, which Descartes takes to be simply a firm and constant will to do all that we judge to be the best. Once again, Elisabeth raises pointed objections. Here she is concerned with preserving the traditional tie between virtue and contentment. On Descartes's account, she charges, virtue would be insufficient for contentment. Given that our knowledge is incomplete, our best judgments would inevitably be wrong sometimes, and on those occasions we would regret our actions. Elisabeth takes this regret to be incompatible with virtue. Our incomplete knowledge also raises another problem for her, that of measuring the value of things. While Elisabeth admits the passions to be sources of value, she also recognizes that different individuals evaluate things differently. For her, the central problem of ethics is not achieving the sovereign good but rather reconciling competing evaluations of things. Her interest in the passions as sources of value leads her to request Descartes to enumerate and describe all the passions. In response, Descartes drafted his last work, *The Passions of the Soul*. Descartes sent this portion of the correspondence, including Elisabeth's letters, to Queen Christina of Sweden when she requested his views on the sovereign good.

Elisabeth and Descartes also address the problem of reconciling free will with determinism. Whereas Descartes asserts that human freedom is consistent with divine providence, though how it is so might escape us, for Elisabeth simply asserting that the two are consistent is insufficient. In addition, Elisabeth's request that Descartes lay out some maxims for civil life results in an extended discussion of Machiavelli's *The Prince* and the obligations of a good ruler to his subjects.

See also Descartes, René; Metaphysics.

Bibliography

Descartes, René. *Oeuvres*. Vols. 3–5, edited by Charles Adam and Paul Tannery. Paris: Vrin, 1996.

Elisabeth, Princess of Bohemia. *Elisabeth, Princess of Bohemia and René Descartes: Their Correspondence*. Translated by Lisa Shapiro. Chicago: University of Chicago Press, forthcoming (expected publication in 2006).

Penn, William. *An Account of W. Penn's Travails in Holland and Germany, Anno MDCLXXVII*. London: T. Sowle, 1695 and 1714.

Reynolds, Edward. *Treatise of the Passions and the Faculties of the Soule of Man*. London: Robert Bostock, 1640. Facsimile reproduction, edited by Margaret Lee Wiley, Gainseville, FL: Scholars' Facsimiles and Reprints, 1971.

Broad, Jacqueline. *Women Philosophers of the Seventeenth Century*. Cambridge, U.K.: Cambridge University Press, 2002.

Shapiro, Lisa. "Princess Elizabeth and Descartes: The Union of Soul and Body and the Practice of Philosophy." *British Journal for the History of Philosophy* 7 (3) (1999): 503–520.

Tollefson, Deborah. "Princess Elisabeth and the Problem of Mind-Body Interaction." *Hypatia* 14 (3) (1999): 59–77.

Lisa Shapiro (2005)

EMANATIONISM

Emanationism explains the origin and structure of reality by postulating a perfect and transcendent principle from which everything is derived through a process called emanation (Greek *aporroia, probolē, proodos*; Latin *emanatio*), which is comparable to an efflux or radiation. Emanation is timeless and thus can be called a process only figuratively. It leaves its source undiminished, so that

the source remains transcendent; but as the process continues, each of its products is less perfect.

In these three respects emanationism is opposed to evolutionism because evolution is a temporal process in which the principle itself is involved (immanent) and in which an increase in perfection is usually conceived. Emanationism is also opposed to creationism, according to which the principle creates the rest of reality (from which it differs absolutely), either out of nothing or by transforming a preexisting, chaotic matter into a cosmos. There is some affinity between emanationism and pantheism, except that the latter teaches the immanence of the principle in its product. Some philosophers characterize emanationism as panentheism.

Emanationism forms an important part of several philosophic and religious doctrines, though it is somewhat elusive in the latter.

PHILOSOPHIC EMANATIONISM

A theory of emanation can be found to a certain extent in the philosophy of Plato and the Old Academy as presented by Aristotle. Out of two highest principles (usually called the One and the Indefinite Dyad), ideas, in some way identified with or comprising mathematicals (numbers; geometrical entities, i.e., point, line, plane, solid) evolve; out of solids, the physical world evolves. But the nature of the process (for which Aristotle used the term *genesis*) remains unclear. The Stoa, Neo-Pythagoreanism, and Philo contributed some ideas to emanationism, but the philosophy first appears in full clarity in the system of Plotinus. His supreme principle, because it is transcendent, ineffable, and absolutely simple (One), must "overflow," just as what is mature must beget. The first product of this overflowing is intelligence (*nous*), which roughly corresponds to Plato's idea. From intelligence emanates psyche (corresponding to Plato's mathematicals) which becomes, by degrees, less and less perfect, more and more multiple. From the psyche emanates matter that, when "illuminated" by the psyche, becomes the physical world.

Often, although not always, Plotinus describes emanation as a necessary, involuntary, "natural," and therefore blameless process, somewhat like a point of absolutely intense light that emits a cone of light without any loss of its own substance. As the cone of light expands in volume, it grows dimmer, finally passing into complete darkness, on which the light produces images as on a screen. But just as the ontic status of darkness is ambiguous (Is it a minimum of light or its complete absence and therefore not its product?), so the status of matter in Plotinus is never quite clear.

The emanationism of Plotinus was taken over by all Neoplatonists, but among them, Proclus deserves particular mention. By subdividing Plotinus's emanative steps, Proclus made the process more continuous; and to the "vertical" emanation he added something like a "horizontal" one, fully articulating the realms of intelligence and psyche. From Neoplatonism, emanationism passed into the Christian, Muslim, and Jewish philosophies of the Middle Ages (Dionysius the Pseudo-Areopagite, John Scotus Erigena, Nicholas of Cusa, al-Farabi, Avicenna, Averroes, the book of Zohar), often with pantheistic or creationistic modifications. In modern times, evolutionism has obliterated the emanationist philosophy.

RELIGIOUS EMANATIONISM

In religion, emanationism appears in many Gnostic systems, most conspicuously in *Pistis Sophia* (Faith-wisdom) and in some writings of Valentinus. But in neither of these is it the exclusive principle explaining the origin of everything outside the highest principle. Furthermore, emanation appears in these writings as the result of some reflection and will. It produces, not abstract principles, as in Plotinus, but a host of mythological characters—the first products of emanation according to Valentinus are thirty Aeons—performing a cosmic drama. In addition, what remains entirely in the background in Plotinian theory becomes prominent in Gnosticism; namely, that some acts of the will, which produce emanations, are the result of error or shortcomings. The physical world is created by one of the products of emanation, the Demiurge (identified with the Mosaic creator, the Platonic divine craftsman). The Demiurge is evil himself, and his creation, the world, is an evil place in which man finds himself entrapped and from which gnosis shows the elect ones a way to salvation. Although soteriology plays some part in Plotinian theory, it does not occupy a central place in the system. According to Plotinus, the efflux is balanced by a reflux, which takes place *pari passu* with the efflux. For humankind, the enactment of this reflux remains the most important task; and every person is, by nature, capable of performing it. Gnostic emanationism is ultimately motivated by a feeling of complete hostility to and estrangement from the material world—a feeling that the emanationism of Plotinus, in spite of some ascetic and pessimistic strains, explicitly refuses to countenance.

See also al-Fārābī; Aristotle; Averroes; Avicenna; Erigena, John Scotus; Neoplatonism; Nicholas of Cusa; Pantheism; Plato; Plotinus; Proclus; Pseudo-Dionysius; Valentinus and Valentinianism.

Bibliography

Dodds, E. R. "Proclus." In *The Elements of Theology.* 2nd ed., 212–214, 230. Oxford: Clarendon Press, 1963.

Eisler, R. *Wörterbuch der philosophischen Begriffe.* 3 vols., 4th ed. 321–322. Berlin, 1927–1930.

Faggin, G. "Emanatismo." In *Enciclopedia filosofica,* 1861–1864. Venice and Rome, 1957.

Heinze, M. "Emanation." In *Schaff-Herzog Encyclopedia of Religious Knowledge,* 117ff. New York and London, 1909.

Mora, J. Ferrater. "Emanación." In *Diccionario de filosofia,* 4th ed., 400–401. Buenos Aires, 1958.

Ratzinger, J. "Emanation." In *Reallexikon für Antike und Christentum,* 1219–1228. Stuttgart, 1959.

Philip Merlan (1967)

EMANATIONISM [ADDENDUM]

The sort of philosophy that the Islamic world discovered when it came into contact with the main centers of civilization in the Middle East was neoplatonism. This represented a long tradition of philosophy that had as one of its main planks a theory of how the world is linked with its ultimate cause. Emanation is an important neoplatonic concept that provides an account of this relationship. The existence of this world and the other worlds that exist along with it are taken to flow from the ultimate cause, the One that is the cause of multiplicity, and a main difference from a normal causal relationship is that the cause and the effect are often taken to occur at the same time as each other. This is because emanation represents an eternal process, not only bringing other things into existence but also sustaining their continuing existence.

This concept was taken up with enthusiasm by most of the main Islamic philosophers, and subsequently by Jewish philosophers in the Islamic world. Emanation was often identified with grace, in that God's grace could be seen as eternally influencing lower forms of existence. The identification of the One from which everything else flows is not difficult to link with the God of Islam and Judaism, and the emphasis on the unity of the deity must have struck a chord with these two monotheistic religions. What is problematic from a theological perspective is that emanationism is different from the notion of ex nihilo creation that does seem to be mentioned in both Islamic and Jewish religious texts. On this notion of creation God decides at a particular time to create the world, so first of all there was nothing in existence except God, and subsequently the world came about through his fiat. Yet with emanation there never was nothing except God. God eternally thinks, and from that thought the worlds are produced, and the worlds always existed because God has always existed and thought. Moreover, the notion of emanation implies that God is not aware of what comes about as a result of his thought, because anything lower than the abstract level at which he thinks is beneath his dignity to contemplate.

The language of emanationism is useful for mystical thinkers in both Sufism and kabbalah. The notion of God being in constant contact with the world and everything in existence being connected to everything else provides theoretical underpinning for an immanent view of God's relationship to his creation. God is then in a radical sense *in* the world, and everything that exists is an aspect of him, even what looks insignificant. Mystics tended to argue that it is possible to come close to God by following the emanationist process back to where it starts, with God, but this is a tortuous route that only a few adepts are likely to entirely follow. For most philosophers, however, the route to perfection only goes as high as the active intellect, the most abstract form of thought of which we are capable. Once we go beyond this form of thinking we enter into realms of the emanationist cosmology that we cannot properly grasp except in general terms.

See also Islamic Philosophy; Jewish Philosophy; Kabbalah; Neoplatonism; Sufism.

Bibliography

Frank, Daniel, and Oliver Leaman, eds. *History of Jewish Philosophy*. London: Routledge, 1997.

Leaman, Oliver. *Brief Introduction to Islamic Philosophy*. Oxford: Polity, 1999.

Leaman, Oliver. *Introduction to Classical Islamic Philosophy*. Cambridge, U.K.: Cambridge University Press, 2003.

Nasr, Seyyed H., and Oliver Leaman, eds. *History of Islamic Philosophy*. London: Routledge, 1996.

Oliver Leaman (2005)

EMERGENCE

Emergence is, broadly speaking, the fact that there are features of the world—objects, properties, laws, perhaps other things—that are manifested as a result of the existence of other, usually more basic, entities but that cannot be completely reduced to those other entities. Theories of emergence tend to fall into two basic types: ontological emergence and epistemological emergence—with conceptual emergence serving as a subcategory of the latter.

Advocates of ontological emergence consider emergent phenomena to be objective features of the world, their emergent status being independent of human existence and knowledge; advocates of epistemological emergence consider emergent features to be a result of the limited abilities of people to predict, to calculate, to observe, and to explain; and advocates of conceptual emergence consider emergent features to be a product of theoretical and linguistic representations of the world.

Emergence has considerable philosophical importance because the existence of certain kinds of ontologically emergent entities would provide direct evidence against the universal applicability of the generative atomism that has dominated Anglo-American philosophy in the last century. By generative atomism is meant the view that there exist atomic entities, be they physical, linguistic, logical, or some other kind, and all else is composed of those atoms according to rules of combination and relations of determination. The failure of various reductionist programs, especially that of physicalism, would have significant impact on this program. In addition, the various accounts of epistemological emergence pose difficulties for the long established Cartesian requirement of completely transparent access to evidential relations.

Although there is no consensus upon what counts as an emergent entity, a cluster of features tends to recur in philosophical accounts of emergence. Emergent phenomena are irreducible, they are novel, they are usually unpredictable on the basis of theory, they are often unexplainable, they frequently involve global rather than merely local properties, and an emergent entity must emerge from something. This last feature separates emergent features from those entities whose existence does not depend upon anything else, such as the objects of fundamental physics or certain abstract entities. It also allows for two distinct kinds of emergence: static or synchronic emergence within which the emergent entities exist simultaneously with the entities from which they emerge; and dynamic or diachronic emergence, within which the emergent entities temporally develop from antecedent entities. Although it is rarely stated explicitly, dynamic emergence is generally held to result from more than causal processes alone.

At one time, life and chemical compounds were considered to be good candidates for emergent features, covered by what John Stuart Mill in book III of his *A System of Logic* (1843, ch. VI, pts. 1–2) termed heteropathic laws, but with advances in molecular biology and an understanding of the chemical bond that view fell into disfavor. Perhaps as a consequence, emergence came to be viewed with a certain degree of suspicion, apparently requiring a commitment to occult qualities that was at odds with the analytic methods of science and philosophy. It is thus ironic that emergence has reemerged as a vigorous and lively field of investigation, has shed much of its air of mystery, and plausible candidates for emergent phenomena have been discovered in fundamental areas of physics as well as in other areas of science such as complexity theory. As a result, it is important when considering emergence not to restrict one's range of examples to the widely discussed cases of mental properties.

This entry will emphasize contemporary positions on emergence, although occasional historical references will be made to illustrate conceptual continuities. For a history of the area, the reader is referred to Brian McLaughlin's 1992 survey article, "The Rise and Fall of British Emergentism."

ONTOLOGICAL APPROACHES TO EMERGENCE

One influential ontological approach to emergence uses supervenience relations to account for emergent features. An early version of this approach by James van Cleve (1990) asserted that a property P of a system is emergent if and only if P supervenes with nomological necessity but not with logical necessity on properties of parts of the system, and some of the supervenience principles linking the basal properties with P are fundamental laws. That is, once the features of the most fundamental level are fixed, so—via laws of nature—are the features of all higher levels. Advocates of supervenience approaches, especially the widely canvassed position known as Humean supervenience, generally hold that supervenience is all that is required to account for higher-level features of the world. David Lewis provided an influential statement of this position in the second volume of his *Philosophical Papers* (1986, pp. ix–xvi). Supervenience approaches usually contain the irreducibility and novelty aspects of emergence. Whether the global, unpredictability, and unexplainability features are present depends upon the type of supervenience involved.

A different ontological position, developed by Jaegwon Kim in his article "Making Sense of Emergence" (1999), begins with the idea that a higher level property P is reducible if: (a) P can be functionalized—that is, defined in terms of its causal role; (b) realizers of P can be found at a lower level; and (c) there is a lower level theory that explains how the realizers operate. In contrast, a property is emergent if it is neither a physical property nor reducible to physical properties in the sense just

described. Kim's position retains the irreducibility, novelty, theoretical unpredictability, and unexplainability features of emergent phenomena but apparently has the consequence that there is little scope for their existence, except perhaps in the case of qualia or consciousness.

The novelty of emergent features is usually captured in the idea that an emergent entity E must be qualitatively different from the entities from which it emerges. A popular version of this idea asserts that a property P is emergent if it has novel causal powers not possessed by entities at lower levels. The causation involved can be horizontal (to entities at the same level), upwards (to a higher level), or downwards (to a lower level). When downwards causation is involved, one of the most difficult problems facing advocates of supervenience emergence and many other ontological accounts of emergence occurs. This is the problem of causal exclusion, of explaining how emergent features can influence lower levels via downwards causation if one subscribes to the causal closure of lower levels as, for example, do most physicalists. For if the lower level is casually closed, any downwards influence is redundant, unless causal overdetermination is allowed. A clear statement of this argument can be found in Kim's 1992 article "'Downward Causation' in Emergentism and Nonreductive Physicalism."

A third ontological approach to emergence, found in Paul Humphreys' 1997 article "How Properties Emerge," addresses this problem. It has as its core idea the view that in certain cases of dynamic emergence the original elements or their properties fuse together in a way that the identities of those elements are lost in forming the new emergent entity. This feature allows emergent phenomena to avoid the causal exclusion argument because the lower level entities no longer exist and a fortiori cannot be causal competitors to the emergent entity. The position entails the irreducibility, novelty, and holistic features of emergent phenomena, but allows their predictability and explainability. Certain holistic quantum systems possessing states of joint systems but not states of individual components seem to be examples of fusion.

EPISTEMOLOGICAL APPROACHES TO EMERGENCE

Turning to epistemological accounts of emergence, one of the oldest approaches emphasizes the essential unpredictability of emergent phenomena. It is sometimes loosely and unhelpfully characterized in psychological terms by noting that emergent phenomena are surprising. A more precise version asserts that a property P belonging to domain E is emergent relative to a domain D, where E is at a higher level than D, if it is impossible to predict the occurrence of instances of P on the basis of any ideal theory about D. Early accounts of emergence based on unpredictability can be found in Stephen Pepper's article "Emergence" (1926) and C. D. Broad's book *The Mind and Its Place in Nature* (1925).

This unpredictability approach conforms to Ernest Nagel's well-known approach to the reduction of one theory to another in chapter eleven of his *The Structure of Science* (1961). Within Nagel's account, one theory is irreducible to another if the laws of a higher level theory cannot be deduced from those of a more fundamental theory by employing bridge laws connecting the two levels. Thus, in a somewhat crude manner the essential unpredictability approach to emergence captures the idea that if biology is Nagel-irreducible to physics then biological phenomena are emergent from physical phenomena. It satisfies the novelty, irreducibility and, trivially, unpredictability aspects of emergence and also accommodates nomological emergence, the view that entities of type B are emergent from entities of type A if and only if entities of type B have type A entities as constituents and there is at least one law that applies to type B entities that does not apply to type A entities. A statement of nomological emergence can be found in the physicist P. W. Anderson's much cited 1972 article "More Is Different."

A diachronic version of the unpredictability approach to emergence is widely used within the field of complexity theory and rests on the idea of stable patterns spontaneously emerging in a system. Although these patterns are, simply in virtue of being patterns, nonlocal, they are not the result of a central organizing principle but result from local, often nonlinear, interactions between members of a population. Examples of pattern emergence abound in what are commonly termed self-organized systems, one simple example of which is the formation and maintenance of bird flocks. The general area of agent-based or individual-based models, which include many examples of self-organizing systems, is of interest to philosophy because it combines a bottom-up commitment to individualism with the dynamic emergence of higher level structures possessing the features of novelty and holism. Such models can illuminate the traditional philosophical issue of methodological individualism, an issue that divides those who hold the view that there are sui generis facts in the social sciences from the individualists who deny this.

Because the dynamic emergence of the patterns can often be modeled only via computer simulations, an

important aspect of these systems is captured by Mark Bedau's concept of weak emergence (2003). A weakly emergent property P is one possessed by a structured system S, where P is incapable of being possessed by components of S, and S's possessing P is a fact that can be derived only by a step-by-step simulation of S. Despite its connection with prediction via computer simulations, weak emergence is ultimately a metaphysical rather than an epistemological account of emergence. The structure of the system places objective constraints on the possibilities of computation and complex physical and biological systems must step through their own development, thus making weak emergence a claim about the world itself.

A particularly interesting kind of weak emergence occurs when a pattern P exists independently of the nature of the specific components of the system exhibiting the pattern so that the structure is in that sense autonomous. There are connections here with the multiple realizability of higher level properties, a topic that has played an important role in arguments against reduction. One approach to emergence that explicitly considers multiple realizability is Robert Batterman's asymptotic emergence (2002). This sort of emergence involves a relation between two theories, one of which is a limiting case of the other and it is unusual in not relying on the part/whole relationship upon which most other theories of emergence are based.

CONCEPTUAL APPROACHES TO EMERGENCE

Running parallel with the issues of epistemological and ontological emergence is the phenomenon of conceptual emergence, based upon the idea that theories employed at different levels of the hierarchy employ different concepts and that these concepts require the introduction of distinctive, irreducible, predicates and relations. This approach is captured in Paul Teller's characterization: An emergent property of a whole is one that is not explicitly definable in terms of the nonrelational properties of the object's proper parts (1992). Because definability depends upon the linguistic resources available in a given language or theory, this criterion for emergence is relative to the theory or language employed and reflects a common feature of linguistic development. If psychological and sociological features, to take two examples, are ontologically emergent, one should expect the resources of explicit definability to fail and to force the invention of new vocabulary. It is not difficult to see how each of the approaches to emergence described above can necessitate this kind of linguistic innovation, and it calls into question various enterprises of linguistic reduction. Although it is not couched in terms of emergence, the influential arguments found in Jerry Fodor's 1974 article "Special Sciences" against reduction and in favor of the autonomy of the special sciences can be construed as reasons in support of conceptual emergence.

OTHER ISSUES

In contemporary philosophy, a commitment to emergent entities is generally held to violate physicalism, the position that the world's ontology contains nothing but the ontology provided by physics. What "nothing but" means differs from one version of physicalism to another, as does what is included within the scope of physics, but the core idea is that anything not required by fundamental physics is in principle redundant, even though one may employ a nonphysicalist vocabulary for practical reasons. Thus, mental entities such as beliefs are mere façons de parler on the reductionist view, and the social sciences have no genuine subject matter of their own. Strict versions of reductionism maintain similar views about biological and chemical entities.

All three approaches to emergence—ontological, epistemological, and conceptual—tend to appeal, implicitly or explicitly, to a layered view of the world that is divided into levels, with features at higher levels emerging from those at lower levels. This appeal to levels is usually grounded in the idea that larger entities such as molecules spatially include as parts smaller constituents such as atoms, this inclusion relation resulting in the familiar hierarchy of elementary particle physics, solid state physics, chemistry, biochemistry, biology, neurophysiology, and so on. Although this levels picture serves as a natural image within synchronic emergence, it can be a seriously misleading metaphor for diachronic emergence.

There is much casual talk in the literature on emergence about the difference between aggregate features and emergent features, the latter, in contrast to the former, being more than "mere sums" of the features of their components. It has turned out not to be informative to try to precisely capture what constitutes a "mere sum" but traditionally, holism—summed up in the slogan that the whole is greater than the sum of the parts—remains a core part of what is wanted from emergent phenomena. It is preferable to replace "greater than" by "different" and if this is done one has the suggestion that a property P is emergent only if it is a property of an entire system S that is composed of subsystems $S_1, \dots S_n$ but none of the S_i possess P. This feature is possessed by, at least, the fusion,

asymptotic, weak, and nomological approaches to emergence.

The principal aim of any philosophical account of emergence should be to make emergence intelligible and nontrivial. It is a separate matter, one with which science is properly concerned, whether the universe contains any examples of emergence. It is, nevertheless, a matter of considerable interest to philosophy whether examples of ontologically emergent phenomena exist because, if they do, our universe is more than an ontologically modest combinatorial device.

See also Chaos Theory; Physicalism; Reduction; Supervenience.

Bibliography

Anderson, P. W. "More is Different." *Science* 177 (1972): 393–396.

Batterman, Robert. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press, 2002.

Beckerman, Ansgar, Hans Flohr, and Jaegwon Kim, eds. *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*. Berlin: Walter de Gruyter, 1992.

Bedau, Mark. "Downward Causation and Autonomy in Weak Emergence." *Principia Revista Internacional de Epistemologica* 6 (2003): 5–50.

Broad, C. D. *The Mind and Its Place in Nature*. London: Routledge, 1925.

Fodor, Jerry. "Special Sciences, or The Disunity of Science as a Working Hypothesis." *Synthese* 28 (1974): 97–115.

Holland, John. *Emergence: From Chaos to Order*. Reading, MA: Addison-Wesley, 1998.

Humphreys, Paul. "How Properties Emerge." *Philosophy of Science* 64 (1997): 1–17.

Kim, Jaegwon. "'Downward Causation' in Emergentism and Nonreductive Physicalism." In *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, edited by Ansgar Beckerman et al, 119–138. Berlin: Walter de Gruyter, 1992.

Kim, Jaegwon. "Making Sense of Emergence." *Philosophical Studies* 95 (1999): 3–36.

Lewis, David. *Philosophical Papers.* Vol. II. Oxford: Oxford University Press, 1986.

McLaughlin, Brian. "The Rise and Fall of British Emergentism." In *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, edited by Ansgar Beckerman et al, 49–93. Berlin: Walter de Gruyter, 1992.

Mill, J. S. *A System of Logic: Ratiocinative and Inductive*. London: Longmans, Green and Company, 1843.

Nagel, Ernest. *The Structure of Science*. New York: Harcourt, 1961.

Pepper, Stephen. "Emergence." *Journal of Philosophy* 23 (1926): 241–245.

Teller, Paul. "A Contemporary Look at Emergence." In *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, edited by Ansgar Beckerman et al, 139–153. Berlin: Walter de Gruyter, 1992.

van Cleve, James. "Mind-Dust or Magic? Panpsychism versus Emergentism." *Philosophical Perspectives* 4 (1990): 215–226.

Paul Humphreys (2005)

EMERGENTISM

See *Emergence*

EMERSON, RALPH WALDO

(1803–1882)

Ralph Waldo Emerson, the American author and leader of New England transcendentalism, was born in Boston, Massachusetts. His father, a locally distinguished Unitarian clergyman, died in 1811 leaving Emerson and five other children in the care of a pious mother and a very learned aunt on the father's side. From 1813 to 1817 Emerson attended the Boston Latin School; then, after four undistinguished years at Harvard, he became a schoolmaster while he continued to study extramurally at Harvard Divinity School. "My reasoning faculty is proportionally weak," he confessed in his *Journal* in 1824, on deciding to become a minister, "nor can I ever hope to write a Butler's Analogy or an Essay of Hume. … [But] the preaching most in vogue at the present day depends chiefly on *imagination* [italics added] for its success, and asks those accomplishments which I believe are most within my grasp." Made just before he was twenty-one, this acute piece of self-analysis marks the stage in Emerson's life when he really began to understand himself and gain a genuine premonition of his future role as literary artist. For Emerson is, more than anything else, an imaginative writer. (Thus Friedrich Nietzsche, who was at an early stage influenced by Emerson—admiring his "manifoldness" and "cheerfulness"—recognized him as one of the nineteenth century's few great masters of prose.)

FORMATIVE EXPERIENCES

Unitarianism was at first the main formative influence on Emerson, but it was not the most far-reaching, and the sort of preaching he was eventually to excel in had little to do with any established church or, for that matter, with Christianity as such. A trip to Florida for health reasons, in the winter of 1826–1827, brought about a chance meeting with the aristocratic Achille Murat, whose "consistent Atheism" Emerson found combined, to his sur-

prise, with moral perspicuity. By the late 1820s the young theological student had already got through a prodigious regimen of philosophical and occult reading that included (as the most important authors for his maturer orientations) Zoroaster, Confucius, Muḥammad, the Neoplatonists, Jakob Boehme, Gottfried Wilhelm Leibniz, Baron de Montesquieu, Jean-Jacques Rousseau, Edmund Burke, the Scottish philosophers, Emanuel Swedenborg, Johann Gottfried Herder, and—above all—Madame de Staël (the *De l'Allemagne*). Emerson's attention was being irresistibly drawn to the new cultural movement in Germany. The disturbing advances in German biblical criticism were beginning to penetrate to him via his brother William's enthusiastic letters from Göttingen (William had also met and talked with Johann Wolfgang von Goethe). Soon Emerson was absorbed in Thomas Carlyle's pioneering essays on German literature, and in Samuel Taylor Coleridge's *Aids to Reflection* (1825)—in which Emerson discovered the pseudo-Kantian distinction between "Reason" and "Understanding."

In 1829 Emerson was appointed pastor of the Second Church of Boston; shortly afterward he married Ellen Louisa Tucker. Ellen's tragic death of tuberculosis early in 1831 had a deeply anguishing and yet strangely liberalizing effect upon Emerson. He questioned himself about immortality; preached sermons that expounded embryonic versions of his own later doctrines of "self-reverence" (or "self-reliance," as he sometimes called it), "compensation," and "correspondence"; found he was bored with weekday Bible classes; and eventually gave up his pastorate.

On January 2, 1833, he sailed for Europe. This first European tour (he made two more, one in 1847–1848 and one in 1872–1873) was crucial in helping him shape into something like a whole the new philosophical outlook he had been consciously groping toward since at least 1824 and to which he ultimately gave poetic expression in his major works. During a short stay in Britain he managed to get an interview with Coleridge at Highgate, met William Wordsworth, and spent twenty-four hours with the Carlyles at Craigenputtock. Carlyle immediately became a lifelong friend.

The conversations with Coleridge and Carlyle, the two men who were to the disenchanted young American living embodiments of all that was viable in contemporary European culture, had simply the effect of confirming Emerson's old belief: As a guide to solving the problem of life's meaning, there is "really nothing external, so I must spin my thread from my own bowels." He reasoned to himself that "the purpose of life seems to be to acquaint a man with himself" and "the highest revelation is that God is in every man." In his *Journal* entry for September 8, 1833, written while sailing back to America, Emerson included with the above affirmation of his maxim of "self-reverence" two other by then quite explicit convictions: (1) "There is a *correspondence* [italics added] between the human soul and everything that exists in the world," and (2) since "a man contains all that is needful to his government within himself," it must be that "nothing can be given to him or taken from him but always there is a *compensation* [italics added]." Here were brought together the key notions that Emerson was to elaborate for the rest of his life, first in his original transcendentalist manifesto, *Nature* (1836), and then in practically all the later works, including *Essays* (First Series, 1841; Second Series, 1844), *Representative Men* (1850), *English Traits* (1856), *Conduct of Life* (1860), *Society and Solitude* (1870), and *Letters and Social Aims* (1875).

In 1835 Emerson married Lydia Jackson, with whom, he soberly remarked to William, he had found a "quite unexpected community of sentiment and speculation." Soon he was settled in unusual domestic serenity with his wife and his mother in Concord, which remained his home for the rest of his life. Emerson's writings, his sage-like personality, and his roles as the leader of New England transcendentalism and the editor of the *Dial* gradually brought him an international reputation as perhaps America's leading man of letters.

MATURE WRITINGS

If propounded by a philosopher, Emerson's assertions concerning "correspondence" and "compensation" would demand further explication and defense. But to expect anything resembling epistemological lucidity, or even concern, in a writer like Emerson would be to approach him with misconceptions. Indeed, those who read him as one would a philosopher like Immanuel Kant, Friedrich von Schelling, G. W. F. Hegel, or even Coleridge (all of whom certainly had a great influence upon Emerson), largely miss the peculiar merits and significance of his works. For Emerson was neither a critical philosopher nor an idealist metaphysician, but an intuitive sage-poet: "In Emerson," wrote Nietzsche to Overbeck, "*we have lost a philosopher.*"

Like his artistic models Michel Eyquem de Montaigne, Blaise Pascal, and the Goethe of the *Maximen und Reflexionen,* Emerson was a virtuoso of the *pensée,* in which style and content, symbol and "meaning," are inseparably conjoined. His meditations are exploratory rather than defining or definitive, and the nonproposi-

tional, revelatory use of language with which Emerson alternately enraptures and ensnares his reader renders inappropriate the conventional task of giving a systematic conspectus of his leading ideas. The analysis to be applied to any work by Emerson is that of the literary critic rather than the philosopher. His method of exploration consists in the cumulative and often dialectical juxtaposition and attempted coalescence of *aperçus* relating to a single broad theme—"Nature," "Friendship," "Wealth," "Immortality"—usually in the form of an essay, lecture, or address. In fact, all Emerson's prose works are homiletic: They are secular sermons that differ from the sermons of his ancestors, the New England Puritan divines, largely by virtue of a greater breadth and subtlety of message and the intense personalism of their inner soliloquy.

Yet, despite the epistemological imprecision of his views, Emerson is philosophically interesting in at least two ways. First, because of the very full *Journal* he kept throughout his life, he affords an extremely well-documented record of a major writer who found it urgently necessary to struggle with philosophical ideas in order to achieve personal (and artistic) integration in an age "destitute of faith, but terrified at scepticism," as Carlyle characterized it. (The ideological perplexities of his age, moreover, lead directly to our own.) Emerson strove to discover for himself "an original relation to the universe": a kind of personal *Weltansicht* that would somehow keep vital his essentially religious sensibilities and give succor to his pressing emotional needs. Since Christianity could no longer do either of these things, he meditated upon his own experience in the light of those pieces of philosophy that seemed most accommodating. That Emerson found the Germanic philosophical tradition more to his liking than the Anglo-Saxon was the natural result of his individualism, his belief in the primacy of personality, and his closely related admiration for the hero, genius, or great man, in which he joined Johann Gottlieb Fichte, Carlyle, and Nietzsche (see especially *Representative Men*). He expressed these fundamentally anthropocentric and aristocratic orientations quite succinctly: "No object really interests us but man, and in man only his superiorities; and though we are aware of a perfect law in nature, it has fascination for us only through its relation to him, or as it is rooted in the mind."

Both Schelling and Hegel influenced Emerson in profound and clearly traceable ways—Schelling first, through Coleridge, and Hegel later, particularly through W. T. Harris and the St. Louis School of Hegelians, with whose *Journal of Speculative Philosophy* Emerson was closely associated in the late 1860s and early 1870s. The primacy of "personality," or "self-consciousness," as it was usually called, was already an established axiom with the Germans. And if the all-embracing dichotomy between mind and nature—with its innumerable manifestations in the troublemaking divisions of "reality and illusion," "religion and science," "moral law and physical law," "the eternal and the temporal," in effect, the division of "the transcendental ideal and the banal actual"—could be shown to be only an immature stage in the development of Absolute Spirit whose final blossoming would exhibit all as one: Then, indeed, there would be not only "a correspondence between the human soul and everything that exists in the world" but, even better, a coalescence.

Much in the manner of Hegel, Emerson came to see History, or God, or the Oversoul as a kind of primordial schizophrenic, originally split into mind and nature and now victoriously struggling to personal integration in and through the creative achievements of human culture. Metaphysically speaking, human culture is identical with mind's reintegration with nature. Indeed for Emerson science itself becomes the handmaiden of transcendentalism: Man's conquest of the material environment shows nature to be not alien but fully transparent to mind, and since whatever is intelligible must somehow be itself intelligence, mind and nature are in reality one. But in such a panspiritualistic universe every apparent evil can only be for the greater universal good; the "compensation" for evil lies in the ultimate self-harmony of mind. This is the tortuous metaphysical hallucination that forms the basis of Emerson's optimism. As far then as it can be discerned, his *philosophia prima* is that of the German idealists, and one sympathetic way of characterizing him would be to say that where Schelling and the rest made the fundamental mistake of attempting to give rational and systematic expression to the mythology of romanticism, Emerson put the whole thing into poetry—which was exactly where it belonged.

But Emerson's individualism had a further and more practical consequence. He could never reconcile himself to the values of a civilization that, as he saw it, was "essentially one of property, of fences, of exclusiveness"; and the incisive manner in which this dissatisfaction with the prevailing social reality found expression in his writings gives Emerson a special place in the great line of romantic critics of mass society from Rousseau to Karl Jaspers. Brilliantly critical of emergent American commercialism, which necessarily seemed to involve cultural superficiality, Emerson was particularly virulent against the species of democracy that in fact often demands only conformity to depersonalizing custom, and a consequent sacrifice of

individual autonomy, of "self-reliance." He did not limit his criticism to America; *English Traits* is still, among other things, a major indictment of European cant, Philistinism, and materialism by an American.

The second reason why Emerson is philosophically interesting is his influence on philosophers. Nietzsche has been mentioned; so also should be Henri Bergson. A number of Bergson's fundamental concepts often seem in part to be systematizations of Emerson's eclectic intuitions (compare, for example, the *élan vital* with Emerson's "vital force" in the essay "Experience"); perhaps the most noteworthy is the decided interest in Emerson shown by the pragmatists William James and John Dewey.

Emerson's most pervasive influence, however, was not so much on professional thinkers or writers, but on the public, through the great popular sale of his works. His highly personal yet persuasive and accessible form of romanticism insinuated itself into the general intellectual consciousness of America, and to a lesser extent into that of Europe. "His relation to us is … like that of the Roman Emperor Marcus Aurelius," said Matthew Arnold in *Discourses in America* (published in 1885, three years after Emerson's death); "he is the friend and aider of those who would live in the spirit."

See also New England Transcendentalism.

Bibliography

The Complete Works of Ralph Waldo Emerson in the twelve-volume Centenary Edition (Cambridge, MA: Riverside Press, 1903–1904) is the standard edition of Emerson's works. Emerson's *Journals* were originally edited by E. W. Emerson and W. E. Forbes (Boston: Houghton Mifflin, 1909–1914). A more recent version is *Journals and Miscellaneous Notebooks of Ralph Waldo Emerson,* edited by William H. Gilman and others, 3 vols. (Cambridge, MA: Belknap Press of Harvard University Press, 1960–1963). For more bibliographical details consult *Eight American Authors: A Review of Research and Criticism,* edited by Floyd Stovall (New York: Modern Language Association of America, 1956; reprinted with a bibliographical supplement extended to 1962, New York: Norton, 1963). An informed and brilliantly perceptive account of the role of German thought in Emerson's intellectual development is contained in H. A. Pochmann's *German Culture in America* (Madison: University of Wisconsin Press, 1961), pp. 153–207. Among recent studies of Emerson's mind and art, the most illuminating is Jonathan Bishop, *Emerson on the Soul* (Cambridge, MA: Harvard University Press, and London: Oxford University Press, 1965).

OTHER RECOMMENDED WORKS

Cameron, Kenneth Walter. *Emerson's Developing Philosophy: The Early Lectures (1836–1838)*. Hartford: Transcendental Books, 1996.

Cameron, Kenneth Walter. *Emerson's Philosophic Path to a Vocation*. Hartford: Transcendental Books, 1996.

Emerson, Ralph Waldo. *The Vision of Emerson*. Edited with an introduction by Richard Geldard. Rockport, MA: Element, 1995.

Goodman, Russell B. *American Philosophy and the Romantic Tradition*. New York: Cambridge University Press, 1990.

Gray, Henry David. Emerson: *A Statement of New England Transcendentalism as Expressed in the Philosophy of its Chief Exponent*. New York: Ungar, 1958.

Hopkins, Vivian Constance. *Spires of Form: A Study of Emerson's Aesthetic Theory*. New York: Russell & Russell, 1965.

Jacobson, David. *Emerson's Pragmatic Vision: The Dance of the Eye*. University Park: Pennsylvania State University Press, 1993.

Levin, Jonathan. *The Poetics of Transition: Emerson, Pragmatism, and American Literary Modernism*. Durham: Duke University Press, 1999.

Lopez, Michael. *Emerson and Power: Creative Antagonism in the Nineteenth Century*. DeKalb: Northern Illinois University Press, 1996.

Porter, David T. *Emerson and Literary Change*. Cambridge, MA: Harvard University Press, 1978.

Robinson, David. *Emerson and the Conduct of Life: Pragmatism and Ethical Purpose in the Later Work*. New York: Cambridge University Press, 1993.

Van Leer, David. *Emerson's Epistemology: The Argument of the Essays*. New York: Cambridge University Press, 1986.

Michael Moran (1967)
Bibliography updated by Desirae Matherly Martin (2005)

EMOTION

Over the centuries, the emotions have proven to be a notoriously recalcitrant philosophical subject, defying easy classification and stubbornly straddling accepted philosophical distinctions. With changing conceptions of the mind and its powers, categories such as emotion, desire, appetite, passion, feeling, and sentiment come and go. The general term *the emotions* is a relatively recent arrival to the English language, first gaining prominence in the nineteenth century, long after terms such as *fear*, *shame*, and *joy* were in common use. Its introduction was an attempt to clump together states that were supposedly marked by a degree of "emotion," a metaphorical extension of the original sense of the word, namely, agitated motion, or turbulence. Only the vagueness of the metaphor allows it to stretch far enough to cover typically quiescent "emotions" such as being pleased or sad about something.

Probably one influence on the extension of the term is the older category of "passions," in the sense of ways of being acted upon. In many languages nearly all emotion adjectives are derived from participles: for example, the English words amused, annoyed, ashamed, astonished, delighted, depressed, embarrassed, excited, frightened, horrified, irritated, pleased, terrified, surprised, upset, and worried—and even sad (from "sated"). When people are, for example, frightened, something acts on them, that is, frightens them: typically, something of which they are aware. However, even if the terms commonly used for the various emotions suggest that the notion of passivity is central to the ordinary concept of emotion, that notion seems irreparably vague, at best reflective of a prescientific picture of a person (or, for that matter, a physical object) as acting and acted on, as doing and "suffering." Indeed, it is not obvious that the states we call emotions have anything interesting or important in common that distinguishes them from all other mental states. Some philosophers and scientists have argued that what we call "the emotions" do not belong to a "natural kind" or class, and even that the concept of emotion should be banished entirely, at least from scientific discourse. These issues will be taken up in a later section.

THE PHILOSOPHICAL TRADITION

Since William Alston published his seminal article on "Emotion and Feeling" in the first edition of this Encyclopedia, philosophical scholarship in the area has undergone tremendous growth and variation. Among the major catalysts for change in philosophical thinking about the emotions have been new developments in psychology and neuroscience. However, the medium within which this ferment has largely been taking place is linguistic and conceptual analysis. Although analytic philosophers of emotion use relatively sophisticated logical and linguistic tools, their task has not been much different from that of the many classical philosophers who attempted definitions of various emotions: for example, Aristotle in the *Rhetoric*, Descartes in *The Passions of the Soul*, Hobbes in the *Leviathan*, Spinoza in his *Ethics*, and Hume in *A Treatise on Human Nature*. Moreover, the most important outcome of the analytic thrust was a view that had been at least implicit in traditional accounts, namely, cognitivism. Although there are several varieties of cognitivism, perhaps the most influential versions hold that the various emotions are distinguished in part by the types of situation that evoke them; or, more exactly, by the types of situation the awareness of which evokes them; more exactly still, by the content of the beliefs and other propositional attitudes that cause them. Note the importance of situational and cognitive features in Spinoza's definitions, for example:

> Fear: an inconstant pain arising from the idea of something past or future, whereof we to a certain extent doubt the issue.
>
> Regret: the desire or appetite to possess something, kept alive by the remembrance of the said thing, and at the same time constrained by the remembrance of other things that exclude the existence of it.

The classical philosophers contributed more than definitions, of course. For example, some declared certain emotions to be the primary or basic emotions. However, the philosophers remained armchair theorists, putting forward at best introspective or anecdotal data. The scientific advances of the nineteenth century, particularly in biology, made it possible to move beyond this.

BODILY RESPONSES AND FEELINGS: DARWIN AND JAMES

In *The Expression of Emotions in Man and Animals*, published in 1872, Charles Darwin investigated the various, mostly involuntary physiological changes, especially in the facial and skeletal muscles, which constitute the so-called "expressions" of emotions (1998 [1872]). Others broadened the investigation to include the internal visceral phenomena associated with various emotions. Still, these were thought to be investigations into mere manifestations or accompaniments of emotions. As John Dewey pointed out, "The very phrase 'expression of emotion,' … begs the question of the relation of emotion to organic peripheral action, in that it assumes the former as prior and the latter as secondary" (p. 553). It was left to the introspectionist psychologists, most notably Wilhelm Wundt and Edward Titchener, to offer a systematic account of what they regarded as "the emotions themselves," namely the subjective feeling qualities characteristic of the various emotions, an account that relied heavily on what subjects reported.

To William James, these descriptive studies of "what it is like" to feel the various emotions afforded no insight or understanding. Turning instead to the causes of these feelings, he argued, in his classic 1884 paper, "What is an Emotion?" that they were actually the felt awareness of precisely those physiological "manifestations" of emotion that Darwin and the biologists had been studying. Thus, according to James's theory (also known as the James-Lange theory), an emotion is the felt awareness of reverberations of the "bodily sounding-board," that is, of

bodily reactions to something perceived or thought: reactions such as trembling, quickening of pulse, crying, running, or striking someone. It is this perception of one's own bodily responses that endows each type of emotion, such as fear, anger, and joy, with its special feeling quality. From this premise James drew the radical conclusion that emotions or emotional states were effects rather than causes of these bodily reactions. Thus common sense has it backwards: The truth is that "we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful" (James 1884, p. 190). However, this conclusion drew on the further assumption, which James inherited from the introspectionists and from Darwin himself, that the various emotions or types of emotional state were nothing but particular feeling qualities. That is, *if* the emotions are just a subclass of feeling qualities, and *if* these feeling qualities are caused by the bodily reactions that commonsense regards as manifestations or expressions of emotion, *then* common sense has it backwards.

Whatever the merits of his arguments, the influence of James brought about a major shift in philosophical and scientific thinking about the emotions. Most important, the study of emotional feelings could no longer be regarded as a special introspective science, insulated from our general theory of human beings as biological organisms.

PHILOSOPHICAL COGNITIVISM

Opposition to the first of the Jamesian premises, which treats the various emotions as just so many feeling qualities, was a major impetus to the cognitive turn in the philosophy of emotion. However, it was above all the intentionality of emotions that put the cognitive revolution on its positive course. Starting with Anthony Kenny (1963), various authors endeavored to show that, unlike the brute physiological feeling states celebrated by James, emotions and their associated feelings had the characteristic of being *about* things and events. Thus, people sometimes are (and feel) scared of snakes or angry about the fact that their car was stolen. In this respect, emotions were thought to differ from mere bodily feelings, which, if they are about anything at all, are about bodily phenomena, rather than snakes or car thefts.

The intentionality of emotions also distinguishes them from moods, which are general response templates that are not about anything in particular, even though they may have been precipitated by the awareness of particular facts or events. In many languages the same term may be used for both a mood and an emotion: one may be sad (or: depressed, euphoric) about something, or simply in a sad (depressed, euphoric) mood or frame of mind; or both at once. How deep the distinction between emotion and mood goes is debatable, as many so-called emotions tend to spill over from one category to the other: Initially about their original precipitant ("He's angry about the theft of his car."), they develop into a general response template ("Don't go near him, he's in an angry mood!").

Among cognitive theorists, the notion that the content of emotions takes a propositional form assumed special importance. Suppose I believe that John stole my car. I may say that my anger is about the car, about the car theft, or about what John did. However, fully parsed, my anger attribution can be logically reformulated by the phrase: I am angry about the fact that "John stole my car." Because propositions are the primary vehicles of logic and cognition, the propositional nature of emotion and its intentional objects made it easy to think of emotions in both cognitive and logical terms. It was now possible to articulate and debate what were termed the logic and structure of the various emotions and even their inferential ties to one another.

An early and forthright propositional theory is due to Robert Solomon (1976), who, with a strong emphasis on phenomenology, revived the Stoic view that emotions were themselves simply normative judgments of an urgent kind. Ronald de Sousa (1987) argued that emotions are better assimilated to perceptions. Emotions of a given type, such as fright, represent what they are about as having the corresponding property—for example, as being frightening. They also impose "determinate patterns of salience" on our thought processes: guiding our attention, our lines of inquiry, and our inferential strategies. De Sousa's view in some ways anticipates Jesse Prinz's "embodied appraisal" theory, described in the section, "Somatic Wisdom" (2004). Robert Gordon (1987) argued that most emotions are propositional attitudes that are identified by their causal relations to other propositional attitudes, especially beliefs and wishes. Most emotions are "factive," that is, about a fact (or what the subject takes to be a fact) that frustrates or satisfies a wish; others, such as being afraid or hopeful, are uncertainty emotions.

Some critics argue that propositional accounts would exclude animals and infants lacking language. This criticism would seem committed to the controversial thesis that animals without language do not have any propositional attitudes, including desires or beliefs.

Nonetheless, it would be a mistake to try to force all emotions into a propositional framework. It is hard to think of a *that*-clause that describes what love or hate is about. In some cases, what is called the same emotion (or emotion type) has both propositional and nonpropositional forms. Although one may be startled (to discover) that something is the case, one may also be startled by a sound—with no associated proposition or cognition at all. Even in the case of an emotion about a fact, it is not obvious that its content is exhausted by its propositional content.

Finally, even where propositionality is not in dispute, one may not be able to explain it in terms of standard states such as beliefs and desires. For example, to sustain the claim that fearing (hoping) that p depends on being uncertain whether p, one needs to allow for compartmentalization, for example, to distinguish between emotional and intellectual certainty; otherwise one could not account for cases where the fear (hope) that p persists despite a belief that it is not at all possible (epistemically) that p. To make such a distinction with any clarity, however, may be beyond the competency of analytic philosophy. It may require reference to the underlying neural architecture. For example, Joseph LeDoux (1998) discovered that there are distinct pathways by which the amygdala may be activated, a cortical "high road" that is cognitive, and a "low road" that bypasses the cortex and is strictly perceptual. This hypothesis nicely complements the claim that some emotional states and processes might be *modular*, that is, "hardwired" in a manner that makes them impenetrable by changes in beliefs and desires. The "quick and dirty" low road often alerts us to emergencies that our cortex "knows" do not exist. These examples suggest an analogy with perceptual illusions, which a correct belief sometimes fails to dispel.

VALENCE

Emotions are often classified by their *valence*. Theorists and laypeople tend to readily agree that emotions, or most of them, are either positive or negative. The agreement evaporates, however, as soon as they are asked, "In what respect?" One point of disagreement concerns what is being evaluated: Is it what the emotion is characteristically about that makes it positive or negative (intentional valence), or is it one's having or experiencing the emotion (experiential valence)? What it is about may be good, or something the subject appraises favorably or would wish to be the case, as in pride, delight, and hope; or it may be bad, or something the subject appraises unfavorably or would wish not to be the case, as in shame, regret, and fear. Having or experiencing the emotion might also be judged positive or negative in any of several respects. It might be characteristically positive or negative in affect (i.e., pleasant or unpleasant), or even unconsciously aversive or attractive, and it might be beneficial or harmful, or morally good or bad. Because of such disagreements, some argue that the idea of emotional valence is of dubious value and should be abandoned.

However, it may be an important feature of emotions that they have multiple dimensions of valence. If an emotion's experiential valence is of the same sign (positive or negative) as its intentional valence—for example, an aversive emotion that is about something that is bad for you or goes against your wishes—then it is likely to promote rational decision-making and action. The actions people take to alleviate the unpleasantness or aversiveness of fear (a negative aspect of having or experiencing the emotion) tend to reduce the risk of bad things happening (a negative aspect of what the emotion is about): for example, fear of a flood leads the inhabitants to retreat to high ground, thereby averting disaster. (There are of course thrill-seekers for whom the very aversiveness of fear has a second-order attractiveness, and, within the safe confines of dramatic art, many people can enjoy the fear or "as-if" fear they empathetically experience.)

Likewise, the possible negative consequences of a decision tend to be amplified in our minds by our anticipation of regret and remorse: For example, if I buy this appealing but unreliable car, I may kick myself if anything goes wrong with it. These premonitory influences may on the whole guide us to useful behavior, in roughly the way that hunger, thirst, and sexual feelings lead us, wittingly or not, to behavior that is conducive to biological fitness. Add to this theme of doubly valenced emotions a revival of James's second premise, that emotional feelings are perceptions of bodily reactions, and we are led to the topic of the next section.

SOMATIC WISDOM

From Plato onward, European and North American philosophers have thought the regulation of emotion essential to a rational life, and a similar view was promoted even earlier in Buddhism and other Asian religions. The underlying supposition was that unregulated emotions are impediments to the rational life. However, this is compatible with the thesis held by a number of philosophers that emotions, or at least some of them, make a positive and possible indispensable contribution to rational decision-making. According to De Sousa, for example,

emotions are indispensable for guiding our attention, our lines of inquiry, and our inferential strategies.

It was suggested earlier in this article that if an emotion's experiential valence matches its intentional valence, then it is likely to promote rational decision-making. A similar view has received support from findings in neurology and neuroscience, most prominently by the cognitive neuroscientist Antonio Damasio (1994) and his coworkers. Damasio's somatic marker hypothesis holds that successful and unsuccessful decision outcomes produce differing bodily responses—for example, as measured by skin conductance—and the accumulation of such responses over time leads to anticipatory bodily responses that guide future decision-making. One need not be aware (phenomenally conscious) of these responses in order for them to influence decision-making. However, a part of one's frontal cortex (functionally, the somatosensory or body-sensing cortex) must keep track of them. Most of the supporting data have come from observations of decision-making deficits in people with prefrontal damage and comparison with normal subjects in experimental gambling tasks. Additional data suggest that the capacity to recognize and to name certain emotions in others on the basis of their facial expressions also depends on the capacity to monitor one's bodily responses when observing them. Damasio's theory goes far beyond this, and some of it is controversial; but this brief statement makes it clear why Damasio thinks the aversiveness or attractiveness of undergoing certain emotions can be a premonitory influence that sometimes "knows" better than pure reason does which decision paths are likely to lead to preferred outcomes.

Prinz is probably the first philosopher to build a general theory of emotion on a broad and richly detailed account of empirical research. Although sympathetic to the somatic theories of James and Damasio, he argues that our emotional "gut reactions," unlike pains, tickles, and feelings of fatigue, are representational states. Applying Dretske's thesis that a state may be representational in virtue of having an evolved function of carrying a certain class of information, he argues that these bodily changes constitute perceptual appraisals or evaluations of our relationship to the environment with respect to well-being. He calls his view a non-propositional appraisal theory, because he holds that emotions need not involve propositional attitudes such as belief, judgments, and desires (2004).

THE NATURE OF THINGS

The classical definitions of emotions were answers to questions of the traditional Socratic form: "What is regret?" "What is fear?" and so forth. The aim was not to capture the nuances of ordinary usage, but rather to be telling us something about ourselves: to explain, as Spinoza said, "not the meaning of words, but the nature of things" (1883 [1677], p. 178). However, if this is the ambition of the philosophy of emotion, then some philosophers would reply that it is up to science and not philosophy to tell us about the nature of things. In particular, we have to look beyond the terms of ordinary language and the concepts embedded in our everyday "folk" psychology, beyond even the best philosophical attempts to regiment these terms and concepts, if we are to discover what the emotions *really* are. This appears to be a special application of a more general view in philosophy of mind, that of eliminative materialism. However, whatever the merits of that general indictment of everyday psychology and any philosophy that attempts to build on it, there may be special reasons to be skeptical of traditional philosophical thinking about the emotions in particular.

Paul Griffiths maintains that we should use biological evolutionary principles of classification to determine what emotions *really* are. Following Paul Ekman (1992), a leading innovator on the role of facial expression in emotion, Griffiths posits surprise, anger, fear, disgust, sadness, and joy as the basic emotions. These adaptive responses are *evolutionary homologues*, discrete genetically ordained behavioral syndromes that are a legacy of our shared mammalian heritage. Appearing in all cultures, these adaptive responses are associated with the same facial expressions in each culture. The classification here is by descent and homology, rather than by resemblance and analogy, which is more typical of analytic approaches. The special evolutionary status of these basic emotions is reflected in Griffith's philosophical declaration that they are *natural kinds*. That is, they are *projectible* kinds: They share causal properties that are sufficiently well correlated to sustain generalizations from known to unknown cases. However, the term *emotions* does not designate a natural class of kind, for it would serve no scientific purpose to group them with the so-called higher cognitive emotions, such as envy, regret, and shame. Predictably, this thesis has sparked controversy, just as the general thesis of eliminative materialism did two decades earlier. Prinz counters that all emotions are valenced appraisals that exploit common aversive or appetitive mechanisms. Louis Charland (2002) suggests that there is a natural kind of organism that might be called an *emoter*, in virtue of having a brain that meets certain criteria of functional organization.

NORMATIVITY AND CULTURE

Evidently all cultures have implicit rules governing at least some emotions: not just whether and how they should be expressed, but also whether and under what conditions and in what degree one should have them. In European and North American cultures, at least, emotional responses are commonly measured by standards of rationality, appropriateness, and morality. A particular instance of an emotion may be thought irrational if it is based on an irrational belief or desire. However, it is common to think an instance of emotion may be irrational even if it is not based on an irrational belief or desire; typically, because it is not suited to what it is about. We also judge instances of emotions as too little or too much, for example, in the case of grief or remorse.

It was suggested earlier that the notion of passivity, of being acted on, may be an important feature of the ordinary concept of emotion. However, it is widely assumed that people have some control over how the environment acts on them and are to some degree responsible, not only for the expression of emotion, but for their having the emotion. Aside from regulating one's exposure to eliciting situations, it is supposed that one can in many cases alter the course of the emotion—for example, by intervening cognitively to reappraise the eliciting situation. Indeed, attending to one's emotional state and labeling it may alter the state. It is plausible that when we use emotion labels in giving expression to our emotion, as in, "I'm angry!" Or "I'm in love!" We are shaping as well as describing our emotional state. Emotion kinds would thus be what Ian Hacking (1995) calls "interactive kinds," like race, ethnicity, and gender: To classify one's own state as of a particular interactive kind, or to be so classified by others, tends to alter the state and to influence one's feelings and behavior accordingly.

Social constructionists would emphasize that we are shaping our emotions to fit it into an acceptable cultural mold. The psychologist James Averill argued that the various emotion concepts are merely cultural creations that shape our assessment of certain transitory syndromes. While pretending to be passively moved to behave in certain ways, people are actively adjusting their behavior to fit these cultural categories. Although this theory is a valuable counterbalance to the widely held assumption that our emotions simply "are what they are," it would be extreme to assert that our emotion categories simply create our emotions *ex nihilo* or to deny that the categories themselves are, perhaps in some societies more than in others, flexible and open to change (Reddy 2003).

Emotions seem a particularly nuanced category, varying in uncharted ways from instance to instance. They also vary in the course they follow from moment to moment and day to day. For reasons such as these, as Iris Murdoch (1970), Martha Nussbaum (2001), and Jon Elster (1999) have emphasized, often the best way of defining an emotion type is by reference to literary examples. Literary examples also make it clear that conceptions of emotion vary over time as well as from one present-day culture to another.

The issues addressed in this section may seem hopelessly tender-minded to philosophers who prefer to focus on biological mechanisms and natural kinds. In turn, philosophers drawn to the issues of this section may find the naturalistic focus excessively narrow. What is chiefly at issue is the proper equipment to bring to philosophical thinking about emotions. Should we allow ourselves to conceive human organisms as people and to employ the full panoply of concepts, learned or biologically preordained, that appear to be indispensable for everyday social perception and understanding? Or should we lay aside these concepts and steadfastly conceive human beings only as complex biological systems?

Retaining our everyday tools of social perception, we will find normative questions, matters of passivity and freedom, and the richness and perspectivity of narrative understanding coming to the fore. Laying these tools aside, we can focus on purely naturalistic explanations of emotional phenomena and the natural kinds that enter into these explanations. Partisans of the naturalistic approach may be tempted to assert that only by laying aside the accustomed tools can we discover what emotions "really" are. Partisans of the other approach might argue that to lay down these tools of social perception is precisely to forego understanding people. One important challenge task for the philosophy of emotion will be to determine whether and how to reconcile these two approaches.

See also Alston, William P.; James, William.

Bibliography

Ben-Ze'ev, Aaron. *The Subtlety of Emotions*. Cambridge, MA: MIT Press, 2000.

Charland, Louis C. "The Natural Kind Status of Emotion." *British Journal for the Philosophy of Science* 53 (2002): 511–537.

Damasio, Antonio. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam, 1994.

D'Arms, J., and D. Jacobson. "Expressivism, Morality, and the Emotions." *Ethics* 104 (1994): 739–763.

Darwin, Charles. *The Expression of the Emotions in Man and Animals*. London: John Murray, 1872.

Descartes, Rene. *The Passions of the Soul*. 1649. In *The Philosophical Writings of Descartes*. Vol. 1. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge, U.K.: Cambridge University Press, 1984.

de Sousa, Ronald. *The Rationality of Emotion*. Cambridge, MA: MIT Press, 1987.

Dewey, John. "The Theory of Emotion. (I) Emotional Attitudes." *Psychological Review* 1 (1894): 553–569.

Dixon, Thomas. *From Passions to Emotions: The Creation of a Secular Category*. Cambridge, U.K.: Cambridge University Press, 2003.

Ekman, Paul. "An Argument for Basic Emotions." *Cognition and Emotion* 6 (1992): 169–200.

Elster, Jon. *Alchemies of the Mind: Rationality and the Emotions*. Cambridge, U.K.: Cambridge University Press, 1999.

Goldie, Peter. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press, 2000.

Gordon, Robert. *The Structure of Emotions*. Cambridge, U.K.: Cambridge University Press, 1987.

Greenspan, Patricia. *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge, 1989.

Griffiths, Paul E. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press, 1997.

Hacking, Ian. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Cambridge, MA: Harvard University Press, 1995.

Jaggar, Alison. "Love and Knowledge: Emotion in Feminist Epistemology." *Inquiry* 32 (1989): 151–176.

James, William. "What is an Emotion?" *Mind* 9 (1884): 188–205.

Kenny, Anthony. *Action, Emotion and Will*. London: Routledge, 1963.

Lazarus, Richard. "The Cognition-Emotion Debate: A Bit of History." In *Handbook of Cognition and Emotion*, edited by Tim Dalgleish and Mick J. Power. Chichester, NY: John Wiley and Sons, 1999.

LeDoux, Joseph. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster, 1998.

Lyons, William. *Emotion*. Cambridge, U.K.: Cambridge University Press, 1980.

Murdoch, Iris. *The Sovereignty of Good*. London: Routledge, 1970.

Neu, Jerome. *A Tear is an Intellectual Thing: The Meanings of Emotion*. New York: Oxford University Press, 2000.

Nussbaum, Martha. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge, U.K.: Cambridge University Press, 2001.

Oatley, Keith. *Emotions: A Brief History*. Oxford: Blackwell, 2004.

Ortony, A., and T. J. Turner. "What's Basic About Basic Emotions?" *Psychological Review* 97 (3) (1990): 315–331.

Panksepp, Jaak. *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press, 1998.

Prinz, Jesse. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press, 2004.

Reddy, William. *The Navigation of Feeling: A Framework for the History of Emotions*. Cambridge, U.K.: Cambridge University Press, 2003.

Roberts, Robert C. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge, U.K.: Cambridge University Press, 2003.

Robinson, Jenefer. "Startle." *The Journal of Philosophy* 92 (1995): 53–74.

Rorty, Amelie. "Explaining Emotions." In *Explaining Emotions*, edited by Amélie Rorty, 103–126. Los Angeles: University of California Press, 1980.

Russell, James A. "Core Affect and the Psychological Construction of Emotion." *Psychological Review* 110 (1) (2003): 145–172.

Solomon, Robert. *The Passions*. New York: Doubleday, 1976.

Sorabji, Richard. *Emotion and Peace of Mind: From Stoic Agitation to Christian Temptation*. Oxford: Oxford University Press, 2000.

Spinoza, Benedict. *Ethics*. Translated by R.H.M. Elwes. London: George Bell & Sons, 1883.

Wollheim, Richard. *On the Emotions*. New Haven, CT: Yale University Press, 1999.

Zajonc, Robert, B. "The Interaction of Affect and Cognition." In *Approaches to Emotion*, edited by Klaus Scherer and Paul Ekman. Hillsdale, NJ: Erlbaum, 1984.

Louis C. Charland (2005)
Robert M. Gordon (2005)

EMOTIVE THEORY OF ETHICS

The term *emotivism* refers to a theory about moral judgments, sentences, words, and speech acts; it is sometimes also extended to cover aesthetic and other nonmoral forms of evaluation. Although sometimes used to refer to the entire genus, strictly speaking *emotivism* is the name of only the earliest version of ethical noncognitivism (also known as expressivism and nondescriptivism).

Classical noncognitivist theories maintain that moral judgments and speech acts function primarily to (a) express and (b) influence states of mind or attitudes rather than to describe, report, or represent facts, which they do only secondarily if at all. For example: To say "Stealing is wrong" is not primarily to report any facts about stealing but to express one's negative attitude toward it. Emotivists also deny, therefore, that there are any moral facts or that moral words like *good*, *bad*, *right*, and *wrong* predicate moral properties; they typically deny that moral claims are evaluable as true or false—at least in respect of their primary meaning. The attitudes expressed by moral judgments are held to be "conative"

(that is, they have a motivational element) and not "cognitive" (that is, they are not beliefs/do not have representational content). Species of noncognitivism are differentiated by the kinds of attitude they associate with moral thought and discourse: emotivism claims that moral thought and discourse express emotions (affective attitudes, sentiments, or feelings) or similar mental states, typically of approval and disapproval, and is therefore sometimes called the "boo-hurrah" theory of ethics.

To understand emotivism, it is important to contrast it with subjectivism, the view that moral judgments and utterances represent, report, or describe someone's attitudes (for example, that we can translate "Stealing is wrong" as "I disapprove of stealing"). Noncognitivist theories deny that moral expressions of attitude take the form of report or description: They are often vague about the expressive mechanism, but it is supposed to bear a family resemblance to that of ejaculations (for example, uttering "Ouch!" to express being in pain) and performatives (for example, saying "Thank you" to express gratitude). Saying "Stealing is wrong" is therefore like saying "Boo to stealing!"

The significance of this difference is apparent, to the advantage of noncognitivism, when one examines what the strategies have to say about moral disagreements. Subjectivists must accept—whereas noncognitivists deny—that moral claims are made true or false by facts about people's attitudes. If A asserts "Stealing is wrong," and B responds "Stealing is not wrong," it is possible, from a subjectivist view, for A and B to be expressing compatible judgments—if they are reporting the attitudes of different people—and therefore not actually to be disagreeing at all. Although noncognitivism does not portray A and B as disagreeing about any fact, it does claim a "disagreement in attitude": A opposes stealing, and B does not.

According to emotivists, we engage in moral discourse in order to influence the behavior and attitudes of others. They claim, therefore, that moral utterances have a psychological function of arousing emotions in others, based on a human susceptibility to emotional influence by exposure to the emotional expressions of others. Charles L. Stevenson even identifies a statement's emotive meaning with this causal tendency.

Almost all emotivist theories acknowledge that moral judgments possess some content that is descriptive and truth-apt. Consider first "thick" evaluative terms such as the names of virtues or vices (for example, *brave*) and pejoratives (for example, *geek*); here it is easy to distinguish a descriptive meaning and an emotive meaning. But most emotivists also ascribe descriptive content to "thin" evaluative terms like *good* and *right*. One common account of this content (Stevenson 1944, Edwards 1955, Hare 1952, Dreier 1990, Barker 2000, Gibbard 2003) is that the property predicated of an object T by *wrong*, for example, is the property for which the speaker disapproves of T. Suppose Elizabeth declares "Stealing is wrong" and disapproves of stealing because she believes it typically causes misfortune to its victims; then the descriptive meaning of her utterance is that stealing typically causes misfortune to its victims. However, this meaning is deemed secondary because (a) it depends upon the emotive meaning—the descriptive meaning of *wrong* will differ from context to context, speaker to speaker, and even occasion to occasion, according to what arouses speakers' emotions, and (b) it has little or no moral significance. A and B will argue over whether stealing is wrong if they differ in attitude toward stealing but not if they differ only with regard to which properties arouse their disapproval of stealing or over whether stealing has some particular property.

HISTORY AND DEVELOPMENT

Although suggestions of emotivism can be found throughout the history of philosophy (David Hume and other early modern sentimentalists have particularly close affinities), the emergence of the theory is usually attributed to a series of short suggestions by British philosophers in the 1920s and 1930s (Ogden and Richards 1923, Barnes 1933, A. S. Duncan Jones as reported in Broad 1933–1934, Ayer 1936); however, earlier formulations appear in German/Austrian value theory from the late nineteenth century (Lotze 1885, Windelband 1903, Marty 1908, and see Satris 1987 for this influence on Anglo-American emotivism). The British emotivists were reacting, in part, to the metaethical theory of nonnaturalism (or intuitionism) advocated by G. E. Moore, H. A. Pritchard, W. D. Ross, and others.

Moore had persuasively argued that moral words could not be defined except in terms of other moral words and inferred (invalidly, as was revealed by the discovery that nonsynonymous terms could be coreferential) that moral words could not refer to "natural" or empirical properties and that moral sentences could not describe natural or empirical facts. Any such attempted definition left out something essential. (This claim is closely related to the alleged is/ought distinction, or "fact-value gap"). Emotivists were convinced by these arguments, but some, influenced by logical positivism—the doctrine that only sentences which are empirically verifi-

able are meaningful—balked at the notion of "nonnatural," nonempirical moral properties and facts. In their diagnosis, the essential something that cannot be captured by any naturalistic analysis of moral language is the expression of speakers' emotions.

Emotivism found its greatest and most dedicated champion in the person of the American philosopher Charles L. Stevenson (1937, 1944) and enjoyed its heyday in the 1940s and 1950s (Nowell-Smith 1954, Edwards 1955) before being largely supplanted by forms of noncognitivism that were thought to be less vulnerable to objection (especially the prescriptivism of Hare 1952, 1963). To philosophers seeking to condemn the horrors of World War II in absolute terms, the claim that moral judgments merely express feelings appeared inadequate. Emotivism's legacy is a widespread recognition today of the significance of emotions for ethical thought, and the efforts of a number of contemporary philosophers since the 1980s—most notably Simon Blackburn (1993, 1998)—who continue to argue for its central tenets.

THE CASE FOR EMOTIVISM

The philosophical stature of emotivism has risen from a number of solidly argued foundations: the apparent failures of efforts to give naturalistic definitions of moral words or to identify natural properties as their referents, epistemological scruples about the existence of nonnatural properties, and the reliable link between moral judgment and emotion. Philosophers still vigorously disagree about whether or not it is possible to find objective referents for moral terms, however, and there are alternative explanations of the connection between moral judgment and emotion: perhaps moral words name properties that reliably arouse emotional responses in us, perhaps they name the dispositional properties of reliably arousing emotional responses, or perhaps their use conversationally communicates speakers' approval and disapproval without in any strict sense "meaning" it.

Further, many philosophers maintain that it is possible and not very unusual for people to make sincere moral judgments without feeling or expressing the relevant emotion (this discussion centers on a figure known as the "amoralist") and that emotive meaning is, therefore, not an essential element of moral judgment. Emotivists commonly respond with the claim that these are not genuine moral judgments but are made in "inverted commas"—i.e. that they merely mimic the practice of moral judgment. The case for emotivism is not bolstered by this claim, however, unless grounds can be found for accepting the "inverted commas" diagnosis that are independent of emotivist convictions themselves.

The emotivist explanation of moral language also provides simple answers to a number of puzzles in metaethics: First, it explains the fact that people are typically motivated to behave in accordance with their moral judgments. Cognitivists have some difficulty explaining this motivational connection because they identify moral judgments with beliefs. On an orthodox view, a belief is not enough to motivate action by itself; it needs to be combined with a desire or similar conative attitude. But, according to emotivism, moral judgments consist in favorable and unfavorable attitudes, and people are likely to perform the actions they feel favorably toward and likely to avoid actions toward which they feel unfavorably.

Second, emotivism explains the synthetic a priori character of moral judgment stressed by nonnaturalists: that is, that despite the fact that an empirical description of a state of affairs or action entails neither by logic nor by meaning the goodness or badness or rightness or wrongness of that state of affairs or action, its description alone nonetheless suffices for us to be confident in passing moral judgment on it. Although it may seem mysterious how anyone could know just from description of a state of affairs or action that it necessarily possesses some further, unspecified property, we have no such need for further information in order to respond emotionally.

Third, emotivism explains the supervenience of the moral on the empirical: why moral characteristics are such that if two states of affairs differ in any moral respect, they must also differ in some nonmoral or empirical respect. If a person is disposed to have a certain emotional response to some state of affairs, then he or she is disposed to have the same response to any qualitatively identical state of affairs. A person will be disposed to make the same moral judgment about two states of affairs, therefore, unless there is some difference between those states that arouses different emotions. While emotivism has an easier task offering solutions to these problems than most descriptivist theories, it must contend with noncognitivist rivals that offer similar explanatory resources.

PROBLEMS

Most of the objections to emotivism in particular are also objections to noncognitivism in general and focus on respects in which moral thought and discourse behave like ordinary, factual, truth-evaluable cognitive thought and discourse. These objections have been widely believed to refute noncognitivism of all varieties, and

accordingly the emphasis in recent noncognitivist writing is on the "quasi-realist" project (Blackburn 1993) of explaining how nondescriptive thought and discourse can mimic ordinary descriptive thought and discourse. The treatment here focuses on the significance of these objections for emotivist theories.

THE EMBEDDING (OR FREGE-GEACH) PROBLEM. Emotivism purports to tell us the meaning of moral sentences; however as P. T. Geach (1960, 1965) and John Searle (1962) have pointed out, it and other forms of noncognitivism appear to succeed at most at explaining one kind of use of simple moral sentences: their use in direct assertion (for example, saying "Stealing is wrong"). Emotivism claims the descriptive form of simple moral sentences is merely a disguise. However simple moral sentences are also given many other uses in which they also behave like descriptive sentences and for which emotivist explanations seem inappropriate or impossible. Consider embedding of simple moral sentences into complex sentences and indirect contexts: disjunctions ("Either stealing is wrong, or Robin Hood was a saint"), belief ascriptions ("Elizabeth believes that stealing is wrong"), conditionals ("If stealing is wrong, then Joe ought not take Mary's lunch"), predications of falsehood ("It is not true that stealing is wrong"), and interrogatives ("Is it true that stealing is wrong?). In each case, a speaker uses the simple moral sentence "Stealing is wrong" but does not express emotions or unfavorable attitudes towards stealing. The emotivist proposal therefore is not helpful in understanding the simple moral sentence in these uses, which is reason to doubt whether it has captured its meaning at all.

It is possible to extend the emotivist account by assigning meanings in each of these contexts, but doing so introduces a further difficulty. Consider a simple moral argument: *P1. If stealing is wrong, then Joe ought not take Mary's lunch; P2. Stealing is wrong; P3. Therefore, Joe ought not take Mary's lunch*. This looks like a standard instance of *modus ponens* and therefore a straightforwardly valid argument. But if we attribute different meanings to "stealing is wrong" as it occurs in each premise, then the argument equivocates, and the conclusion doesn't follow. (Indeed, if P2 is interpreted as a mere expression of emotion without truth value, nothing can logically follow from it). Emotivism therefore casts doubt on the possibility of drawing inferences to or from moral claims—something we do all the time.

Emotivists as early as Stevenson made use of minimalist theories of truth to argue as follows: to claim that *p is true* is simply to claim that *p*, so anyone who is disposed to claim "Stealing is wrong" is entitled to claim that "Stealing is wrong is true." But as the discovery of the embedding problem postdates emotivism's heyday, we do not find solutions to it from self-identified emotivists. Contemporary noncognitivists, however, devote much attention to the problem (especially Blackburn), and there are two broad strategies available: First, if some meaning can be found for the simple moral sentence that is common to these various embeddings and is compatible with emotivism, then arguably standard logic will allow moral inferences. There are two possibilities here. (a) Some seek to identify a noncognitive content that is common to all uses of moral sentences and that plausibly can be embedded in different sentential contexts. These efforts are characteristically found outside of the emotivist tradition (particularly in the work of Hare and Allan Gibbard), and the strategy does not seem so compatible with the emotivist doctrine that simple moral sentences express emotions; (b) Emotivists can turn to the supposed secondary descriptive content of moral claims to explain moral inferences. Because these descriptive contents have truth values, there is no difficulty in forming valid arguments with them. The success of any such explanation depends on the plausibility of the emotivist's claim to have identified the truth-conditional content of the premises and conclusions of moral arguments; it is also arguable that any success must come at the cost of abandoning genuine emotivism and noncognitivism.

Second, even if it is granted that there are no truth relations between the premises of moral arguments and between the contents of moral judgments, it is arguable that there are relations of coherence or consistency between the judgments or states of mind that express those contents. Blackburn accordingly proposes and develops a "logic of attitudes," a system of norms governing the consistency of combinations of attitudes. The conditional premise P1 above, on this view, expresses approval of disapproval of Joe's taking Mary's lunch in the circumstance that one disapproves of stealing. A's attitudes are then allegedly inconsistent if A holds both this second-order attitude and the attitude of disapproval towards stealing expressed by P2 but does not also disapprove of Joe's taking Mary's lunch, the attitude allegedly expressed by P3. Accused by a number of critics of conflating logical inconsistency with pragmatic incoherence (Hale 1986, Schueler 1988, Brighouse 1990, and Zangwill 1992), Blackburn suggests that we can expand the concept of consistency to encompass pragmatic and logical forms. Critics argue that this strategy is not successful: because there is no form of merely pragmatic incoher-

ence that exactly mimics logical inconsistency, Blackburn must claim that some apparently valid moral arguments are actually inconsistent (Hale 1993 and Van Roojen 1996), but noncognitivists have not been deterred.

REASONS AND JUSTIFICATION. Emotivism is charged with being unable to accommodate the important role of rational argument in moral discourse and dispute. Although it emphasizes moral discourse's function of influencing others' behavior, it is thought to characterize this efficacy wrongly, as similar in kind to that employed in manipulation, intimidation, and propaganda. According to emotivists, we engage in moral argumentation with the immediate aim of arousing emotions in others, and moral utterances accomplish this by direct psychological causation. Their opponents object that genuine moral discourse involves furnishing others with reasons, as rational agents, to recognize as correct and thereby accept one's moral views (Hare 1951 and Brandt 1959).

It is true that conscientious moral debaters offer factual considerations as evidence or justification for their positions, and emotivists do not deny it. According to Stevenson, moral argument can take both "rational" and "nonrational" (or "persuasive") forms. On Stevenson's view, by a "reason" for a moral judgment we mean any factual consideration that might influence someone's emotions in the direction of that judgment, and therefore "rational" means of moral argument consist in offering such considerations. Protagonists in a debate over the morality of legalized abortion, for example, might dispute the facts about its consequences. "Persuasive" argumentation, on the other hand, consists in the use of emotive language for its direct psychological effects.

One line of objection, spearheaded by Richard Brandt, observes that it is possible to be emotionally influenced by considerations that are morally irrelevant, and argues that emotivism cannot accommodate the distinction between what is morally relevant and morally irrelevant. Stevenson's reply exhibits a typical noncognitivist strategy: he insists that we can meaningfully distinguish between morally relevant and irrelevant influences on people's attitudes but that when we do so, we are making further moral (and hence emotive) judgments. To judge a consideration morally irrelevant is therefore to express disapproval of being emotionally influenced by it.

OTHER OBJECTIONS. Clearly not just any emotional response constitutes a moral judgment. Emotivists therefore distinguish moral judgments from other kinds of affective or conative reaction by appealing to a distinctive kind (or kinds) of moral emotion. Some critics object that moral approval and disapproval cannot be adequately differentiated from other kinds of affective and conative states without invoking the very moral concepts that emotivists seek to explain by them—and therefore that moral emotions are in fact cognitive attitudes. Moral approval, for example, can arguably only be adequately characterized as the attitude of judging something to be morally good. If this is correct, then emotivism puts the cart before the horse in attempting to explain moral judgments by appeal to emotional states. However, if moral attitudes are not cognitive and are simply affective or conative responses, then it is questionable whether they have the sort of first-person authority that moral judgments purport to possess. If Gary's judgment that homosexuality is morally wrong rests on nothing more than a disposition to have an unpleasant feeling when he contemplates homosexuality, then he may have as good or better reason to resist, suppress, or work to change his emotional sensibilities as he has to oppose homosexuality.

Another concern addresses whether emotivism has the resources to distinguish between accepting the negation of a moral claim and not accepting that moral claim. Believing that the next president of the United States will not be a woman is not the same mental state as not believing that the next president of the United States will be a woman; likewise it seems that accepting that abortion is not wrong is not the same mental state as not accepting that abortion is wrong. Critics charge, however, that emotivism has to explain both in terms of not feeling disapproval toward abortion.

See also Brandt, R. B.; Ethical Relativism; Ethical Subjectivism; Ethics, History of; Ethics, Problems of; Hare, Richard M.; Hume, David; Intuitionism and Intuitionistic Logic, Ethical; Logical Positivism; Moore, George Edward; Noncognitivism; Ross, William David; Searle, John; Stevenson, Charles L.; Value and Valuation.

Bibliography

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1936.

Barker, Stephen J. "Is Value Content a Component of Conventional Implicature?" *Analysis* 60 (2000): 268–279.

Barnes, W. H. F. "A Suggestion about Value." *Analysis*1 (1933): 45–46.

Blackburn, Simon. *Essays in Quasi-Realism*. New York: Oxford University Press, 1993.

Blackburn, Simon. *Ruling Passions*. New York: Oxford University Press, 1998.

Brandt, Richard. *Ethical Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1959.

Brighouse, M. H. "Blackburn's Projectivism—An Objection," *Philosophical Studies* 59 (1990): 225–233.

Broad, C. D. "Is 'Goodness' the Name of a Simple, Non-natural Quality?" *Proceedings of the Aristotelian Society* 34 (1933–1934): 249-268.

Dreier, Jamie. "Internalism and Speaker Relativism." *Ethics* 101 (1990): 6–26.

Edwards, Paul. *The Logic of Moral Discourse*. Glencoe. IL: Free Press, 1955.

Geach, P. T. "Ascriptivism." *Philosophical Review* 69 (1960): 221–225.

Geach, P. T. "Assertion." *Philosophical Review* 74 (1965): 449–465.

Gibbard, Allan. *Thinking How to Live*. Cambridge. MA: Harvard University Press, 2003.

Hale, Bob. "Can There Be a Logic of Attitudes?" In *Reality: Representation and Projection*, edited by J. Haldane and C. Wright. Oxford: Oxford University Press, 1993.

Hale, Bob. "The Compleat Projectivist." *Philosophical Quarterly* 36 (1986): 65–84.

Hare, R. M. "Freedom of the Will." *Proceedings of the Aristotelian Society* 25 (1951): 201–216.

Hare, R. M. *Freedom and Reason*. Oxford: Clarendon Press, 1963.

Hare, R. M. *The Language of Morals*. Oxford: Clarendon Press, 1952.

Lotze, Hermann. *Outlines of Logic and the Encyclopedia of Philosophy*, edited and translated by G. T. Ladd. Boston: Ginn, 1885.

Marty, Anton. (1908). *Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie*. Halle: Niemeyer.

Nowell-Smith, P. H. *Ethics*. Baltimore: Penguin Books, 1954.

Ogden, C. K., and I. A. Richards. *The Meaning of Meaning*. New York: Harcourt, 1923.

Satris, Stephen. *Ethical Emotivism*. Dordrecht: Martinus Nijhoff, 1987.

Schueler, G. F. "Modus Ponens and Moral Realism." *Ethics* 98 (1988): 492–500.

Searle, John. "Meaning and Speech Acts." *Philosophical Review* 71 (1962): 423–432.

Stevenson, Charles L. *Ethics and Language*. New Haven: Yale University Press, 1944.

Stevenson, Charles L. "The Emotive Meaning of Moral Terms." *Mind* 46 (1937): 14–31.

Urmson, J. O. *The Emotive Theory of Ethics*. London: Hutcheson, 1968.

Van Roojen, Mark. "Expressivism and Irrationality." *The Philosophical Review* 105 (1996): 311–335.

Windelband, Wilhelm. "Was ist Philosophie?" In *Präludien: aufsätze und reden zur philosophie und ihrer geschichte*. 2nd ed. Tübingen: J. C. B. Mohr, 1903

Nick Zangwill. "Moral Modus Ponens." *Ratio* 5 (1992): 177–193.

Stephen Finlay (2005)

EMPEDOCLES

(5th century BCE–after 444 BCE)

Empedocles, the Greek poet, prophet, and natural philosopher, was the originator of the doctrine of four elements that dominated Western cosmology and medical thought down to the Renaissance. Empedocles was born in Acragas (Agrigento), Sicily, in the early fifth century BCE and died sometime after 444 BCE. He played a political role in his native city, apparently as a democratic leader, was later exiled, and traveled through other Greek colonies in southern Italy. In one of his poems he describes himself as a "deathless god, no longer a mortal," surrounded wherever he goes by admiring crowds asking for advice, for prophecy, and for a "healing word" to cure them from disease (Fr. 112). A number of anecdotes illustrate his reputation for supernatural powers (including the raising of the dead), and the legend that he died by throwing himself into the crater of Etna gives us an idea of the charismatic impression he left behind in the popular imagination. Modern scholars have often found it difficult to reconcile the scientific and the religious sides of Empedocles' thought. He expounded his views in powerful hexameters, of which considerable fragments are preserved from two distinct poems, *On the Nature of Things* (*Peri Physeōs*) and *Purifications* (*Katharmoi*).

NATURAL PHILOSOPHY

Theophrastus said that Empedocles was much influenced by Parmenides and even more by the Pythagoreans. Pythagorean influence must be seen in his religious teaching and probably also in the role that he assigns to numerical proportion in the natural combination of the elements. From Parmenides he accepted the fundamental principle that nothing can arise out of nothing, nor can anything perish into nonentity. But whereas for Parmenides this meant that all motion and change must be illusory, Empedocles admits that there is real process in nature: "the mixture and separation of things mixed."

By accepting four distinct elements, or "roots of all things," in place of Parmenides' monolithic Being, Empedocles is able to explain natural change as a result of the combination, separation, and regrouping of indestructible entities. There remains, of course, something illusory about the kaleidoscopic appearance of change. Since there can be no generation or annihilation of anything real, Empedocles insists that to describe natural processes in terms of birth and becoming or death and destruction is to follow a linguistic usage which is systematically misleading (Frs. 8–12). In reality there is only the mixing, unmixing, and remixing of permanent entities.

One generation later a similar view of the discrepancy between the appearance of continual change and the reality of unchanging entities led Democritus to distinguish between primary (or true) and secondary (or conventional) sense qualities. However, there is no reason to believe that Empedocles envisaged any such distinction. He assigns the qualities of color, heat, and moisture to the elements themselves and describes the formation of compounds by analogy with the action of a painter mixing his colors. He seems not to have faced the difficult question posed by such analogies: In what does the indestructibility of the elements consist if their essential properties are those that are seen to change?

Nevertheless, the simplicity of this tetradic scheme and its direct application to the great cosmic masses of land, sea, atmosphere, and celestial fire (that is, sun, stars, and lightning) led Plato, Aristotle, and most of their successors to adopt the doctrine of four elements in variously modified forms. Empedocles himself developed the doctrine in a grandiose cosmology that can be reconstructed only in part. The four elements interact under the influence of two cosmic powers, Love (or Aphrodite), on the one hand, and Strife (or Quarrel), on the other. These powers function respectively as forces of attraction and repulsion, but they are also conceived of concretely as ingredients in the mixture. They operate as a kind of dynamic fluid, comparable in some respects to the concept of phlogiston in early modern science. The power of Love or attraction acts first by bringing like together with like—for instance, earth to earth, fire to fire—but it also assimilates the elements to one another, so that what were originally unlikes become like and are united in a new, homogeneous compound (Fr. 22). Love thus represents the power of organic unity and creative combination.

The process of world formation occurs in a cycle that may be said to begin with a totally homogeneous fusion of the elements in a primordial sphere under the exclusive influence of Love. The process of differentiation is set off when Strife makes its entry into the sphere, in accordance with some fixed periodic scheme. It would seem that the cosmic sphere is always saturated with one or the other of these powers or, more frequently, with both of them in a variable ratio; the quantity of Love present in the world varies inversely to that of Strife (Frs. 35 and 16). The life cycle of the universe thus oscillates between the poles of unity and diversity: "Now there grows to be one thing alone out of many; now again many things separate out of one; there is a double generation of mortal beings, a double disappearance" (Fr. 17). This has generally been taken to imply that the creation of things occurs twice, first in the passage from unity under Love to complete diversity under Strife and again in the reverse process from separation of all things to total fusion. (The standard interpretation has recently been challenged by Jean Bollack, who denies that Empedocles intended a double cosmogony. See bibliography.) The present phase of the world cycle is apparently regarded as one of the increasing prevalence of Strife.

Empedocles gave some account of the structure of the heavens and also of the phenomena of earth, sea, and atmosphere which the Greeks studied under the title of meteorology, but the remains of his physical poem show an equal or greater concern with zoology and botany. In the microcosm of plants and animals he discovered the same principles of elemental mixture, harmony, and separation at work. Following up an idea of Anaximander's, he imagined several phases in the emergence of living things from the earth (in combination with other elements), plants preceding animals, and he describes earlier, monstrous forms of animal life. As in Anaximander sexual reproduction appears only in the latest phase of the development. But the details of his doctrine are obscure, and it is difficult to say how far there is any significant anticipation of the theory of evolution.

PHYSIOLOGY AND PSYCHOLOGY. Empedocles shows a keen interest in embryology and physiology, explaining the structure of the eye by analogy with that of a lantern (Fr. 84) and comparing the process of respiration (including the movement of the blood) with the siphon effect of the clepsydra or water pipe, which retains or releases fluid by means of air pressure (Fr. 100). The notion of elemental combination is specified in numerical terms for certain living tissues. Bones are formed by earth, water, and fire in the ratio 2:2:4. The blend of the elements is most equal in flesh, especially in blood (Fr. 98).

Physiology passes over into psychology without a break. (It is clear that as a doctor Empedocles would have practiced psychosomatic medicine.) Blood is the primary seat of thought and perception (Fr. 105) precisely because it is here that the elements are most equably blended. Fundamental in Empedocles' psychology, as in his physics, is the principle of like to like. We see earth with the earth that is in us, water with water, love with love, strife with strife (Fr. 109). This and other passages in Empedocles suggest a one-to-one correspondence between the corporeal elements as such and our conscious experience of them. More precisely, his view seems to be that of a radical panpsychism in which, on the one

hand, all elemental bodies are endowed with thought and sensation (Frs. 102–103) and, on the other hand, knowledge itself is treated like a physical thing obeying the laws of combination, attraction, and repulsion. Thus, Empedocles announces that his own teachings, if carefully assimilated, will form part of the character and elemental composition of the student, whereas, if neglected, "they will leave you in the course of time, yearning to return to their own dear kind; for you must know that all things have intelligence and a share in thought" (Fr. 110). Hence, all our conscious thought and feeling has its direct counterpart in the elemental blend within us (Frs. 107–108), which is itself continually being altered by the stream of incoming and outgoing material (Frs. 89, 106).

RELIGIOUS TEACHING

The religious views stated in the *Purifications* are so strange and so dogmatically presented that some scholars—H. Diels and Ulrich von Wilamowitz, for instance—have supposed that this poem dates from a later, less scientific period in Empedocles' life, reflecting some religious conversion after the bitter experience of exile. Now, the *Purifications* may, in fact, have been composed later than the physical poem, but no biographical development can resolve the alleged contradiction between the scientist and the mystic in Empedocles, for the physical work also presupposes a religious point of view.

In particular, *On Nature* proclaims the immortality and preexistence of the soul (or life principle) as a special case of *ex nihilo nihil.* In Empedocles' view the Parmenidean law of conservation for all real entities guarantees the indestructibility of life in exactly the same way as it guarantees the imperishability of the elements. Hence, only fools can "imagine that men exist merely during what we call life, but that they are nothing at all before being composed or after they are dissolved" (Fr. 15; compare with Fr. 11). Since it is precisely the doctrine of immortality that is supposed to contradict the psychophysics of Empedocles, this contradiction, if it exists, must be located within the physical poem. Furthermore, the same poem implies a developed theology in the description of the primordial cosmic sphere as a "god" (*theos,* Fr. 31), in the reference to the four elements as immortal deities (*daimones,* Fr. 59; compare with Fr. 6), and in the apocalyptic pronouncement of the power of Love-Aphrodite (Fr. 17). Some readers might be inclined to discount such expressions as mere features of poetic style, but such a literary interpretation of theological language, which may be appropriate in the case of Lucretius, seems unconvincing for Empedocles, who appeals to principles of piety and purity throughout the poem (Frs. 3–5, 110, and so on).

The religious views thus alluded to in the physical poem receive emphatic statement in the *Purifications.* Here Empedocles proclaims his own divinity and traces his career as an immortal *daimōn,* banished from the company of the other gods for some prenatal crime; passing through a series of vegetable, animal, and human incarnations; at last attaining the purified life of "prophets, poets, doctors, and leaders"; and now ready to escape from human misery altogether and return once more to the blessed fellowship of the gods. Part of the process of purification consists in the ritual abstinence from meat and certain other foods, such as beans and laurel leaves. This joining of the belief in transmigration with the religious practice of vegetarianism is distinctly Pythagorean. If one adds Empedocles' notion that birth in human form means that the *daimōn* is clothed in an alien garment of flesh (Fr. 126) as a result of a lamentable fall from bliss (Fr. 118), one has a particularly striking example of that otherworldly tendency in Greek religion that is generally known as Orphic and that exercised such a profound influence on Plato as well as on the religious thought of late antiquity.

Remote as this view may seem from the biology and physics of the poem *On Nature,* Empedocles has taken care to preserve a sense of continuity between his religious teaching and his cosmology by a number of parallels, in particular by identifying the primeval sin of the *daimōn* (for which it is punished by incarnation in the cycle of rebirth) as "reliance on Strife." The fellowship of the purified spirits is conceived by contrast as a realm of Love and affection. Thus, the precosmic sphere of the physical poem is paralleled in the *Purifications* by an account of a bygone golden age in which war and bloodshed were unknown, affection prevailed between man and beast, and Aphrodite was queen (Frs. 128–129). Although both poems (which are addressed to different audiences) probably cannot be fitted together at every point, Empedocles clearly thought of the two as compatible, perhaps as complementary views of the world of nature (or physical transformation) and the world of spirit (or divine life). As a result of his panpsychism, Empedocles was able to conceive of nature and spirit as forming two aspects of a single whole rather than as constituting two entirely distinct realms. In any case the essential structure of both worlds is characterized by the same, almost Manichaean rivalry between the beneficent force of Love and the destructive power of Strife. If one sees Love in the physical poem as the cosmic counterpart

of the immortal *daimōn* and his extramundane homeland, Empedocles' whole cosmology will appear as a construction designed to find a place for the Pythagorean doctrine of the transmigrating soul within the shifting and unstable world of elemental strife that had been described by the Ionian natural philosophers.

This reconciliation of the two poems is possible only if one admits the identification of the transmigrating *daimōn* with the element of divine Love—that is, with the unifying principle of intelligent organization present within each one of us but also present throughout nature. This identification has been accepted by Francis Macdonald Cornford and by others, and there is much to be said for it. But it is only fair to add that the identification cannot be proved from the extant texts and that some responsible scholars have denied that there is any possibility of reconciling the doctrine of immortality with the physical psychology of *On Nature.*

One should note Empedocles' clear statement—the first by any Greek—of the notion of an invisible, incorporeal, nonanthropomorphic deity, characterized as a "holy mind [*phrēn*] alone, darting through the whole cosmos with rapid thoughts" (Frs. 133–134). Before Empedocles, Xenophanes had insisted that the "greatest god" must be nonanthropomorphic, but he did not specify its incorporeality. On the other hand, Anaxagoras' principle of mind is clearly noncorporeal, but it is not described as a deity. Empedocles seems to have worked the Anaxagorean principle into his own theology. The phrasing of his account of the spiritual deity recalls the verses concerning Aphrodite as well as the description of the divine sphere. All three principles—the sphere in which the elements are joined, the attractive force of Aphrodite, and the "holy mind" of the cosmos—must somehow have been related in Empedocles' theology, perhaps as three different expressions of the universal power of Love. If so, Empedocles' theology forms the direct continuation of his psychology, since (on the interpretation offered above) it is this same power of Love that figures in the human microcosm as the transmigrating *daimōn.*

See also Anaximander; Leucippus and Democritus; Parmenides of Elea; Psychology; Pythagoras and Pythagoreanism; Theophrastus.

Bibliography

Remains of the poems and other ancient evidence are in H. Diels and W. Kranz, *Die Fragmente der Vorsokratiker,* Vol. I, 6th ed. (Berlin: Weidmann, 1951), Ch. 31.

There are two major studies of Empedocles: Ettore Bignone, *Empedocle* (Turin, 1916), and Jean Bollack, *Empédocle,* 3 vols. (Paris: Les …ditions de Minuit, 1965).

See also Eduard Zeller, *Die Philosophie der Griechen,* edited by Wilhelm Nestle, Vol. I, 6th ed. (Leipzig, 1920), Part 2; John Burnet, *Early Greek Philosophy,* 4th ed. (London, 1930); F. M. Cornford, "Mystery Religions and Pre-Socratic Philosophy," in *Cambridge Ancient History,* Vol. IV (Cambridge, U.K., 1939), Ch. 15; W. K. C. Guthrie, *A History of Greek Philosophy,* Vol. II (Cambridge, U.K.: Cambridge University Press, 1962); Werner Jaeger, *Theology of the Early Greek Philosophers* (Oxford: Clarendon Press, 1947); and G. S. Kirk and J. E. Raven, *The Presocratic Philosophers* (Cambridge, U.K.: Cambridge University Press, 1957).

Special studies include J. Bidez, *La biographie d'Empédocle* (Gand, 1894); Ulrich von Wilamowitz-Moellendorff, "Die *Katharmoi* des Empedokles," in *Berlin Sitzungsberichte* (1929): 626–661; Friedrich Solmsen, "Tissues and the Soul," in *Philosophical Review* 59 (1950): 435–441; D. J. Furley, "Empedocles and the Clepsydra," in *Journal of Hellenic Studies* 77 (1957): 31–34; Charles H. Kahn, "Religion and Natural Philosophy in Empedocles' Doctrine of the Soul," in *Archiv für Geschicte der Philosophie* 42 (1960): 3–35; and E. L. Minar Jr., "Cosmic Periods in the Philosophy of Empedocles," in *Phronesis* 8 (1963): 127–145.

Charles H. Kahn (1967)

EMPEDOCLES [ADDENDUM]

The philosophy of Empedocles remains the subject of widely diverging interpretations. This is so despite the discovery of important new evidence, which, far from dousing old debates, has instead further inflamed them. The following account seeks to chart the impact of the new material on a number of these still-open questions, without, however, ignoring significant contributions made to scholarship before it. Because the assessment of this material is still in its early days, the debate on many points may yet shift in one or the other direction.

THE NEW EVIDENCE

Notwithstanding the addition of a few elements to the corpus since Diels's edition, the study of Empedocles truly entered a new era in 1999 with the publication, by Alain Martin and Oliver Primavesi, of the Strasburg papyrus of Empedocles. The papyrus, assembled from numerous smaller pieces, consists of four larger sections, called by the editors sections a, b, c and d, and a few leftover scraps. These comprise a total of seventy-four hexameter lines, some very partial, of which twenty overlap with lines already known, making the identification of the text certain. By a stroke of luck, the largest section, a,

continues the thirty-five-line Fr. 17, for another thirty-four lines, and thanks to a line-numbering mark in the margin of the papyrus, we can establish that the whole of Fr. 17 section a spanned lines 232 to 300 of its book. That book, as we know from Simplicius, the Aristotelian commentator, was book I of the work he calls the *Physics*, or *On Nature*. In these new lines, Empedocles moves from the broad cosmic cycle described in Fr. 17 to assert the capacity of the six principles to generate "all things," including men and women, trees, and "long-lived gods." The elements as cosmic bodies are described next, including a possible reference to a "we"—that is, humanity—within the churning world masses, but the reconstruction of these lines is controversial, the papyrus being poorly preserved. The section ends with a ten-line address to the disciple in which Empedocles promises to put before his eyes and ears "truthful proofs of my words" by showing Love and Strife at work in all manner of living creatures.

The second-longest new passage, however, section d, is the most significant for the interpretation of Empedocles. Its importance comes from the fact that, for the first time, we can see Empedocles moving from the theme of death and reincarnation, in lines 5–9, to cosmology and the origins of life, in lines 11–18, through a one-line transition at d 10. (Notably, d 5–6 overlap with the previously known Fr. 139, linked by its source to "purifications," and where Empedocles laments "shameful deeds" for the sake of food [that is, meat-eating]. The new text shows that these deeds were wrought "with claws," presumably in an earlier incarnation; the previous text had "for my lips," a much weaker reading.) Thus, unless we are willing to imagine that section d is from a different poem than the other papyrus sections, we are forced to admit that the poem that Simplicius called the *Physics* dealt with reincarnation as well.

THE NUMBER OF WORKS

Before the papyrus, the most important development in Empedoclean scholarship was the attempt to reject the older division of the fragments and assign them all to a single work, forcefully argued by Catherine Osborne in 1987 and Brad Inwood in 1992. But if section d now stands against any neat division of the fragments between religious and scientific content, it does not directly prove that there was only one work. Here one may speak, rather, of a shift in probabilities. In favor of the thesis that there was but a single work section d shows that *Purifications* material featured in the other supposed poem, raising the possibility that when Diogenes Laertius gives both titles (*Lives of the Philosophers* 8.77, the only source where they occur together), he might in fact be giving one long title, like Hesiod's *Works and Days*. In that respect, it is noteworthy that Diogenes gives a single line total for *both* supposed works. Further, the position of Fr. 17, well into Book I, combined with the testimony of Plutarch (who, at *De exilio* 607 c, says that Fr. 115, on the exile of the *daimōn*, was "proclaimed in the beginning of [Empedocles'] philosophy"), argues for a long opening section, or proem, on more traditional themes, as in Parmenides or Lucretius.

But against the single-work thesis there still stands the difficulty that Empedocles has two sets of addressees, the "Friends from Acragas" in Fr. 112, to whom he declares himself a god, and his single disciple Pausanias, to whom he imparts the *On Nature*. A critique of the single-work approach was made by Denis O'Brien (1969), before the discovery of the Strasburg papyrus, whereas the case for a single work has been renewed by Trépanier (2004) on the basis of the new evidence. But the debate on the number of works should not obscure the more fundamental contribution of section d to our understanding of Empedocles: the renewed emphasis it places upon the unity of his thought.

THE UNITY OF EMPEDOCLES' THOUGHT AND THE DAIMŌN

The unity of Empedocles' thought seems to be the one area of emerging consensus. This is in part the result of section d but was also a trend before the papyrus was discorvered. Yet if it now seems likelier than ever that Empedocles had only one philosophical system, this may also make the apparent also seem to drive the contradiction between the reincarnated *daimōn* —or more strongly, its immortality—and Empedocles' physics all the more potent. The contradiction seems to be the following: If the transmigrating *daimōn* is a compound of elements, then even if it could survive from one reincarnation to the next, it could not survive that phase of the cosmic cycle when Strife dominates and no stable compounds endure—at least on the traditional reading of the cycle. To deal with this, a number of alternatives have been put forth. One may downplay the apparent contradiction in various ways: (1) a developmental scheme, no longer favored, posits that Empedocles changed his mind—he wrote two poems, at different stages in his life; 2) less charitably, it has been suggested that, as a poet, Empedocles did not see a contradiction, or if he did, did not care; 3) more subtly, some propose that the *Purifications*constituted the exoteric, popular version of his philosophy, meant to thrill the crowd with promises of personal sal-

vation, whereas the esoteric *On Nature* reinterpreted that salvation as a more stringent and impersonal elemental immortality.

Alternatively, one can try to remove the contradiction by reconciling his conception of the *daimōn* with his physics. One version of this argument identifies the *daimōn* with the first principle Love, thereby denying that is a compound but granting it immortality. Or one may allow that the *daimōn* is a compound, admitting its reincarnation but denying its survival beyond one full cosmic cycle—that is, denying it full-blown immortality. To long-lived but not immortal being one can compare the "long-lived gods" of Fr. 17. The again, some interpretations of the cosmic cycle have held that Strife, while still powerful and active, will never again have complete sway. Thus, immortality might be possible for some compounds.

THE COSMIC CYCLE

That Empedocles held a doctrine of cyclical cosmic history has been generally accepted, but the actual form it took has been the subject of a surprisingly vast debate. The traditional account was the object of several challenges in the 1960s, but also of several powerful rehabilitations, of which O'Brien (1969) is the most comprehensive. Although the traditional version stresses the symmetry of the cycle and thus the equality of Love and Strife, in most alternative versions Strife is denied any creative and hence positive role in the world. Instead of the full pendulum swing found in the traditional version, and the dual creation and destruction it implies, Love would always be creative, Strife always destructive. This view was argued at greatest length by Jean Bollack (1969). To be sure, the constructive role Love plays in Empedocles' biology is far more prominent than that of Strife, whereas in his cosmology Strife is more conspicuous. But this imbalance may be no more than a difference of depiction in the original, for, as Daniel Graham (1988) has well shown, in the passages where the two powers are described together, waxing and waning over the whole macrocosm, they are systematically portrayed as equals.

See also Diogenes Laertius; Love; Lucretius; Parmenides of Elea; Plutarch of Chaeronea; Pre-Socratic Philosophy; Simplicius.

Bibliography

WORKS BY EMPEDOCLES (1962–2004)

Bollack, Jean. *Empédocle*. 3 vols. Paris: éditions de Minuit, 1969. Reprinted by Gallimard, 1992.

Inwood, Brad. *The Poem of Empedocles* (1992). 2nd ed. with new material. Toronto: University of Toronto Press, 2001.

Martin, Alain, and Oliver Primavesi. *L'Empédocle de Strasbourg*.Berlin/New York/Strasbourg: de Gruyter, 1999.

Wright, M. R. *Empedocles. The Extant Fragments*. New Haven, CT: Yale University Press, 1981.

Zuntz, G. *Persephone*.Oxford: Oxford University Press, 1971.

WORKS ABOUT EMPEDOCLES

Barnes, Jonathan. *The Presocratic Philosophers*. 2nd ed. London: Routledge, 1982.

Cantilena, Mario, ed. *Aevum Antiquum*1(2001): Forum -Sul Nuovo Empedocle. Contains seven articles on the Stasburg papyrus; for specialists.

Curd, Patricia. *The Legacy of Parmenides: Eleatic Monism and Later Presocratic Thought*. Princeton, NJ: Princeton University Press, 1988.

Furley, David J. *The Greek Cosmologists*. Cambridge, U.K.: Cambridge University Press, 1987

Graham, Daniel W. "Symmetry in the Empedoclean Cycle." *Classical Quarterly* 38 (1988): 297–312.

Kingsley, Peter. *Ancient Philosophy, Mystery and Magic: Empedocles and the Pythagorean Tradition*. Oxford: Oxford University Press, 1983.

Kirk, Gordon S., John Raven, and Malcolm Schofield. *The Presocratic Philosophers*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983.

Long, A. A. "Empedocles' Cosmic Cycle in the Sixties." In *The Presocratics: A Collection of Critical Essays*. 2nd ed., edited by A. P. D. Mourelatos. Princeton, NJ: Princeton University Press, 1993.

Mourelatos, Alexander P. D. "Quality, Structure and Emergence in Later Pre-Socratic Philosophy." In *Proceedings of the Boston Area Colloquium in Ancient Philosophy.* Vol. 2, edited by John J. Cleary, 127–194. New York: University Press of America, 1987.

O'Brien, Denis. *Empedocles' Cosmic Cycle*. Cambridge, U.K.: Cambridge University Press, 1969.

"Empedocles Revisited." *Ancient Philosophy* 15 (1995): 403–470.

Osborne, Catherine. "Empedocles Recycled." *Classical Quarterly* 37(1987): 24–50.

Sedley, David. *Lucretius and the Transformation of Greek Wisdom*. Cambridge, U.K.: Cambridge University Press, 1998, esp. ch. 1 "The Empedoclean Opening."

Stokes, Michael. *One and Many in Presocratic Philosophy*.Washington: Center for Hellenic Studies / Harvard University Press, 1971.

Trépanier, Simon. *Empedocles: An Interpretation*. London/New York: Routledge, 2004.

Simon Trépanier (2005)

EMPIRICISM

Empiricism is the theory that experience rather than reason is the source of knowledge, and in this sense it is opposed to rationalism. This general thesis, however, can

receive different emphases and refinements; hence, those philosophers who have been labeled empiricists are united only in their general tendency and may differ in various ways. The word *empiricism* is derived from the Greek *ἐμπειρία* (*empeiria*), the Latin translation of which is *experientia,* from which in turn we derive the word *experience.* Aristotle conceived of experience as the as yet unorganized product of sense perception and memory; this is a common philosophical conception of the notion. Memory is required so that what is perceived may be retained in the mind. To say that we have learned something from experience is to say that we have come to know of it by the use of our senses. We have experience when we are sufficiently aware of what we have discovered in this way. There is another, perhaps connected, sense of the term *experience* in which sensations, feelings, and so on, are experiences and in which to perceive something involves having sense experiences. These are experiences because awareness of them is something that happens to us. Indeed, the suggestion of passivity is common to uses of the word. To go into refinements here would not be relevant; one need only appreciate that the statement that experience is the source of knowledge means that knowledge depends ultimately on the use of the senses and on what is discovered through them. Sense experience may be necessary for the attainment of experience, but for present purposes that is unimportant.

The weakest form of empiricism is the doctrine that the senses do provide us with "knowledge" in some sense of the word. This could be denied only by one who had so elevated a conception of knowledge that the senses cannot attain to it. Plato, for example, held at one stage that because of the changeability of the world of sense, sense knowledge lacks the certainty and infallibility that true knowledge must possess. Hence, knowledge cannot be derived from the senses, but only from some other kind of awareness of what he called Forms. The most that sense perception could do would be to remind us of this genuine knowledge. This conception of knowledge demands an infallibility that sense perception cannot provide. Normally, we do not demand such high standards of knowledge, nor do we succumb to this kind of skepticism about sense perception. The commonsense view is that the senses do provide us with knowledge of some sort, and most people, when philosophizing, adopt this kind of empiricist view.

This weak form of empiricism can be generalized into the thesis that *all* knowledge comes from experience, The extreme form of this thesis would be the claim that no source other than experience provides knowledge at all. But this formulation is ambiguous, because there could be various reasons why all that we know might be dependent in some way upon experience. One reason might be that every proposition that we know is either a direct report on experience or a report whose truth is inferred from experience. A prima facie exception to such a thesis is provided by the propositions of mathematics; they have usually been thought to be a priori, not a posteriori—that is, we can know their truth independently of experience. There have, however, been philosophers who have denied the a priori nature of mathematical propositions. J. S. Mill, for example, maintained that the propositions of mathematics are merely very highly confirmed generalizations from experience and, consequently, all propositions are either reports on experience or generalizations from experience. This view has not been widely accepted.

A second reason for maintaining that all knowledge is dependent on experience would be that we can have no ideas or concepts that are not derived from experience, that is, that all concepts are a posteriori, whether or not the truths which can be asserted by means of these concepts are themselves a posteriori. It may be that we know some propositions without having to resort immediately to experience for their validation; for their truth may depend solely on the logical relations between the ideas involved. Yet these ideas may themselves be derived from experience. If all our ideas are so derived, then knowledge of any sort must be dependent on sense experience in some way. According to this thesis, not all knowledge is derived immediately from experience, but all knowledge is dependent on experience at least in the sense that all the materials for knowledge are ultimately derived from experience. St. Thomas Aquinas was an empiricist in this sense. He thought that all our concepts are derived from experience, in that there is "nothing in the intellect which was not previously in the senses" (a doctrine supposedly derived from Aristotle). He did not think, however, that all knowledge either consists of sense experience or is inferred inductively from experience. Similarly, John Locke held and tried to show that all our ideas are derived from experience, either directly or by way of reflection on ideas of sense. He did not hold, however, that all knowledge was sense knowledge.

It is possible to argue an even more complex thesis. It may be held that while there are ideas which are not derived from experience—a priori ideas—and while there are a priori truths which may or may not involve a priori ideas, such ideas and truths only have application on the precondition that there is experience. That is to say

that—for human beings at any rate—reason can function only by way of some kind of connection with experience; "pure" reason is impossible. This was, in effect, Immanuel Kant's position, and although he did not call himself an empiricist *simpliciter,* he was certainly opposed to what he called dogmatic rationalism. He held that there is no place for forms of knowledge of reality which are derived from pure reason alone.

It is possible, then, to maintain a general empiricist thesis that all knowledge is derived from experience on the grounds either that (1) all that we know is directly concerned with sense experience or derived from it by strictly experiential means, that is, learning, association, or inductive inference; or (2) all that we know is dependent on sense experience in that all the materials for knowledge are directly derived from sense experience; or (3) all that we know is dependent on sense perception in that even though we can know some things a priori, this is only in a relative sense, since the having of experience is a general precondition for being said to have such knowledge. None of these theses demand any more than the ordinary conception of knowledge. They do not demand that the knowledge in question should possess absolute infallibility so that the possibility of error is logically excluded. For none of the theses in question is essentially designed to be an answer to skepticism.

EMPIRICISM AND SKEPTICISM

Some forms of rationalism, for example, the Platonic theory already referred to, are meant to be answers to skepticism. They presuppose that an adequate reply to philosophical skepticism can be given only by showing that reason can provide forms of knowledge where error is logically excluded. The search for certainty, so intimately associated with seventeenth-century rationalism in general and René Descartes in particular, aimed at showing that knowledge is possible because there are some things about which we cannot be wrong. Empiricism can be a rival to rationalism, not just in the sense already noted—that it may reject the supposition that reason by itself, without reference to sense perception, can provide knowledge—but also in the sense that it proposes an alternate way of arriving at certainty. Empiricism, in this sense, is the thesis that the certainty required to answer the skeptic is to be found in the deliverances of the senses themselves and not in the deliverances of reason. Rationalism and empiricism, in this sense, are agreed that some such certainty must be found if skepticism is to be answered. They disagree about the sources of that certainty and about the method by which the rest of what we ordinarily call knowledge is to be derived from the primary certainties. Whereas rationalism seeks to derive knowledge in general from certain primary axioms (the truth of which is indubitable) by means of strictly deductive procedures, empiricism seeks to build up or construct knowledge from certain basic elements that are, again, indubitable. The clearest expression of this point of view is probably to be found in twentieth-century empiricism, especially that associated with the logical positivist movement. This point of view is also found in the British empiricists of the seventeenth and eighteenth centuries, Locke, George Berkeley, and David Hume, but in their case it is overladen with other elements and other forms of empiricism, some of which have already been noted. A short historical survey may serve to pinpoint the main issues.

EMPIRICISM IN GREEK AND MEDIEVAL PHILOSOPHY

It is often said that, in one sense, Aristotle was the founder of empiricism. Certainly Thomas Aquinas believed that he had Aristotle's authority for the view that there is nothing in the intellect which was not previously in the senses. It is not clear, however, that Aristotle ever raised this question. When he spoke of the relations between reason and the senses, he was concerned with issues in the philosophy of mind rather than with epistemology. Certainly Aristotle seems to have believed that knowledge is possible outside the immediate sphere of the senses and that reason can and does furnish us with necessary truths about the world. Aristotle's place in the development of empiricism, then, remains unclear.

Perhaps the first declared empiricist was Epicurus, who maintained that the senses are the only source of knowledge. Epicurus was an extreme atomist and held that sense perception comes about as a result of contact between the atoms of the soul and films of atoms issuing from the bodies around us. By this means *phantasiae* (appearances) are set up. These are all veridical. All sensations are true, and there is no standard other than sensation to which we may refer our judgments about the world. Sensations are set up in the soul by external stimuli, and for this reason Epicurus takes them to be "given." They constitute *phantasiae* when they occur in bulk. There is no further evidence that can be adduced in order that their veridicality may be assessed, either from other sensations or from reason. This is not to say that we cannot be in error concerning objects of perception; the films of atoms may become distorted in transit or the *phantasiae* caused by them may be fitted to the wrong *prolep-*

sis (conception). The last is a kind of abstract idea built up from successive sensations; the fitting of a *phantasia* to a *prolepsis* is what corresponds to judgment in Epicurus. It would appear that what Epicurus meant by his assertion that all sensations are true was that since they are caused in us, we can go no further in seeking information; they may not make us have true knowledge of objects, but in themselves they are incorrigible. Precisely how all knowledge was to be built up from these sensations is not clear, and it has often been remarked that the axioms on which Epicurus's metaphysical system rests are far from the data of sense and are often based on more or less a priori arguments. Nevertheless, Epicurus's ideal of knowledge is one which not only depends on experience for its materials but is based on basic truths of experience.

A theory of knowledge similar in many ways to that of Epicurus may be found in St. Thomas Aquinas, although the main sources of Thomas's philosophy are to be found in Aristotle. Thomas was not a complete empiricist, for he did not think that all knowledge was derived from truths of experience. Knowledge of God, for example, could be obtained in other ways, and his existence could be proved by logical argument. Yet Thomas did think that the materials for knowledge must be derived from sense experience, and he gave an account of the mechanism by which this comes about. Roughly, when the sense organs are stimulated, there also results a change in the soul, which is the form of the body; this is a phantasm, a kind of sensory image. In order for sense perception to occur, the universal character of the phantasm must be seen as such. For this purpose, Thomas resorted to Aristotle's distinction between an active and a passive reason. The active reason has to make possible the acquisition by the passive reason of the sensible form of the object of perception by a process which Thomas—probably adapting an analogy used by Aristotle—described as the illuminating of the phantasm. The active reason reveals the sensible form of the object by abstraction from the phantasm. This form is imposed upon the passive reason, which produces a *species expressa,* or verbal concept, which in turn is used in judgment. This process is called the *conversio ad phantasmata*; all concepts are arrived at in this way, by abstraction from phantasms. Hence, in applying them to entities that cannot be objects of perception, we must do so by means of analogies of various kinds with sensible objects. Thomas's empiricism is, therefore, limited to concepts, and it is only in this limited sense that he held "there is nothing in the intellect which was not previously in the senses."

THE BRITISH EMPIRICISTS

When thinking of empiricism, one tends to think, above all, of the British empiricists of the seventeenth and eighteenth centuries.

LOCKE. John Locke was an empiricist in roughly the same sense that Thomas was, and he set the tone for his successors. His "new way of ideas," as it was called, had as its purpose "to inquire into the original, certainty, and extent of human knowledge, together with the grounds and degrees of belief, opinion, and assent." The reference to certainty makes it appear that he was concerned with skepticism or with skeptical arguments similar to Descartes's method of doubt. Locke's solution to this problem, however, was by no means consistently empiricist. His main target for attack was the doctrine of innate ideas, the doctrine that there may be ideas with which we are born or, at any rate, which we do not have to derive from sense experience. The first book of his *Essay concerning Human Understanding* is devoted to a biting attack on this doctrine. In the rest of the book he sets out a positive account of the way in which ideas are built up, explaining that by "idea" he means that which the mind "is applied about whilst thinking." Ideas may be either of sensation or of reflection upon those of sensation; there is no other source. Ideas are also classified as simple or complex, the latter being built up out of the former. The mind has a certain freedom in this process, which may lead to error. (Locke later admitted ideas of relation and general ideas alongside the simple and complex.) The second book of the *Essay* is an exhaustive account of the way in which all objects of the mind are built up from ideas of sense. In this respect, then, Locke's philosophy may be considered an attempt to show in detail the truth of the kind of view which Thomas had embraced, without accepting the same view of the mechanism whereby ideas come into being.

But Locke wanted to assess the certainty of our knowledge as well as its extent. The mind's freedom in forming complex ideas is a source of error, but in the case of simple ideas the mind, to Locke, was like a great mirror, capable of reflecting only what is set before it. Nevertheless, he did not maintain that all our ideas reflect the exact properties of things nor that all knowledge is of this character. In the fourth book of the *Essay* he asserts that all knowledge consists of "the perception of the connection of and agreement, or disagreement and repugnancy, of any of our ideas," but he goes on to distinguish three degrees of knowledge—intuitive, demonstrative, and sensitive. We can have intuitive knowledge of our own exis-

tence, demonstrative knowledge of God's existence, and sensitive knowledge of the existence of particular finite things. Intuition and demonstration bring certainty with them; they provide in effect a priori knowledge. The question of how there can be a priori knowledge of the existence of anything and how this can be a matter of the agreement or disagreement between ideas presents many problems.

These problems become acute in connection with sensitive knowledge. Locke tried to argue at one point that knowledge of the existence of particular finite things is a matter of the perception of the agreement of our ideas with that of existence. This will not do; to know that something exists is not to know merely that the idea of it fits in with the idea of existence. Hence, Locke admitted that this knowledge has not the certainty of the other two, although he insisted that it goes beyond mere probability and is commonly thought of as knowledge. He also tried to argue for the claim that we do have knowledge of sensible things, maintaining that simple ideas are caused in us in such a way that the mind is passive in receiving them. Moreover, the senses may cohere in their reports. None of these considerations really show that we do have knowledge of sensible things, and Locke admitted that they did not amount to proof.

Locke did not claim that *all* our ideas correspond to the properties of things. He felt this claim was true in the case of the so-called primary qualities, for example, bulk, figure, and motion, qualities without which, he maintained, a thing could not exist. It was not true of secondary qualities—for example, color and taste. In this case, the properties of things cause us to have ideas that are not representative of those things; the term "secondary *quality*" is thus a misnomer. Locke's denial of the real existence of secondary qualities turns on his assimilation of our ideas of them to feelings like pain. (His acceptance of primary qualities was probably influenced by the success of physics in his time and its preoccupation with these properties of things.) As for things themselves, Locke maintained that we have little or no knowledge of their real essence, only of their nominal essence—their nature as determined by the way in which we classify them. This is due to the weakness of our senses. We cannot penetrate to the real essence of things, and our ideas of substances are mostly those of powers—the powers that things have to affect us and each other. It can be seen from all this that Locke was an empiricist in a very limited sense. In his view all the materials for knowledge are provided by sense perception, but the extent and certainty of sensible knowledge is limited, while on the other hand, there is nonempirical a priori knowledge of nonsensible things.

BERKELEY. One aim of Berkeley, the second of the British empiricists, was to rid Locke's philosophy of those elements that were inconsistent with empiricism, although Berkeley's main aim was to produce a metaphysical view which would show the glory of God. According to this view, there is nothing that our understanding cannot grasp, and our perceptions can be regarded as a kind of divine language by which God speaks to us; for God is the cause of our perceptions. The *esse* of sensible things is *percipi*—they consist in being perceived and they have no existence without the mind. There exist, therefore, only sensations or ideas and spirits that are their cause. God is the cause of our sensations, and we ourselves can be the cause of ideas of the imagination.

Berkeley argued against those elements of Locke's philosophy that presupposed a physical reality lying behind our ideas. He attacked Locke's conception of substance and the distinction between primary and secondary qualities, pointing out that there was no distinction to be made between them in respect of their dependence on mind. He also attacked the doctrine of abstract ideas which Locke had held, the doctrine that we have general ideas of things abstracted from the conditions of their particular existence—Locke's theory of universals. This Berkeley did because he believed that Locke's theory might provide a loophole for asserting the existence of an idea of substance. The outcome of this was Berkeley's claim that there are no restrictions on the extent of our knowledge. We have knowledge of the existence of God and ourselves to the extent that we have notions of these spirits. We have knowledge of everything else, since the existence of everything else is a matter of its being perceived. There is nothing further beyond our ken. Even subjects such as geometry, which might be supposed to involve knowledge of nonempirical matters, had to be limited in scope in order to rule out nonempirical objects of knowledge. Thus, Berkeley maintained that there is a least perceptible size; hence, there can be no ideas of infinitesimals or points.

In addition to claiming unrestricted scope for our knowledge, Berkeley asserted that knowledge is entirely dependent on sensations for all its materials other than the notions we have of God and ourselves. Berkeley claimed that this view "gives certainty to knowledge" and prevents skepticism. At the same time it defends common sense, he argued, because it does not involve the postulation of a reality behind ideas. His view gave certainty, he

held, because sensations are by definition free from error; for error can arise only from the wrong use of ideas in judgment. The certainty of our sensations is due to the fact that there can be no question whether they actually represent a reality behind them; and this is the basis of Berkeley's claim to deal with skepticism. In general, all knowledge apart from that of our own existence and of God must, for Berkeley, ultimately be derived from sense perception. With these exceptions, therefore, Berkeley was an empiricist not only in respect of the scope and materials of knowledge but also in respect of its foundations. All truths must be founded on the truths of sense experience. The relations between ideas, which Locke had found a source of knowledge, were, for Berkeley, the result of the mind's own acts.

The mind operates upon the ideas given to it, comparing or contrasting them; it does not merely record what is there. Formal disciplines like mathematics, which might be thought to turn on the relations between ideas, thus depend on the ways in which the mind arbitrarily puts ideas together. Hence, to put the matter in terms more familiar today, mathematics is as much a matter of invention as discovery.

HUME. In respect to relations between ideas Hume perhaps went back to Locke, but in other respects much of Hume's philosophy may be represented as an attempt to rid empiricism of the remaining excrescences of nonempiricist doctrine in Berkeley. As to the materials for knowledge, Hume tried to improve on his predecessors with attempts at greater precision. He distinguished first between impressions and ideas, the former being the contents of the mind in perception, the latter those in imagination, and so on. He further subdivided ideas into those of sense and those of reflection, and again, into those which are simple and those which are complex. Like Berkeley, he denied the existence of anything behind impressions, and a cardinal point of his empiricism, to which he returned again and again, was that every simple idea is a copy of a corresponding impression. The understanding is therefore limited to these mental contents. Hume's main method in philosophy was what he called the "experimental method," the reference in all philosophical problems to the discoveries of experience. In effect, the conclusions which he drew from this are the opposite of Berkeley's. They can produce only skepticism. No justification can be given for belief in the existence of the self and an external world, for example. Reason cannot justify such beliefs, for all that we are given is a bundle of impressions and ideas. Only a psychological explanation can be given to account for our having such beliefs. Hume gives such an explanation in terms of the constancy and coherence of our impressions and ideas, and the principles of the association of ideas.

Hume's theory of knowledge is based on a distinction between two kinds of relations of ideas. In the *Treatise of Human Nature* he makes the distinction between relations that depend completely on the related ideas and those that can be changed without changing the ideas. The former, in effect, constitute necessary connections, the latter factual ones. In the later *Enquiry Concerning Human Understanding* he short-circuited the discussion by distinguishing simply between relations of ideas and matters of fact. Mathematics depends entirely on relations of ideas and is thus concerned with necessary truths, the denial of which involves a contradiction. Matters of fact may rest simply on observation, but in the causal relation Hume finds the only case of a matter-of-fact relation that can take us from one idea to another. He shows that statements of causal connection cannot be logically necessary truths, in spite of the fact that we do attach some necessity to causal connections. After a long discussion he finds the explanation for this in the fact that causes precede their effects, are contiguous to them, and are such that there is a constant conjunction between them. As a result, the mind, through custom, tends to pass from one to the other. The feeling derived from this, which is an impression of reflection, constitutes the feeling of necessity that we find in the causal connection.

Hume denied any real connection between cause and effect but tried to explain why we think that there is such. His demonstration that the causal connection is a contingent one is of the utmost importance, but his conclusions about it are skeptical. He held that there can be no real or objective justification for inference from cause to effect. He did allow, it is true, that certain rules can be provided which, when followed, will give some kind of probability to those inductive inferences which we actually do make. The aim of these rules is to make custom reliable and to avoid superstition. Hume has really no right, according to his own principles, to allow so much, and in doing so, he deserts skepticism in favor of a reductionist positivism, which seeks only to deny any necessary connection among things, while retaining belief in inductive inference. The concept of causal connection is thus in effect reduced to that of constant association of events contiguous in space and closely related in time. This is a position incompatible with his general skepticism. Apart from this, Hume's philosophy is of a piece. In Hume, then, extreme empiricism led to skepticism. Apart from relations of ideas, he held, the only *knowledge* we can have is

of what we can directly observe, and any attempt to palliate this conclusion can produce only inconsistency.

In British empiricism, therefore, the gradual weeding out of anything inconsistent with empiricism, either in the form of the claim that the materials for knowledge must be derived from experience or in the form of the claim that knowledge cannot go beyond experience in its objects, resulted in skepticism about most of the things which we ordinarily claim to know. Kant proposed a reconciliation between this thesis and rationalism, maintaining that the rationalist claim of a priori knowledge about reality must be restricted to its application to experience. There is no room for a priori knowledge of anything that is not an object of experience. Pure reason can provide no real knowledge, despite the claims of rationalist metaphysicians. Such nonanalytic propositions as we do know a priori constitute principles that lay down the conditions to which experience must conform if it is to be objectively valid and not just a product of the imagination. A priori truths other than mere analytic truths have validity only in reference to experience; hence, while all knowledge is based on experience, it is not all derived from experience. This is scarcely empiricism in any recognized form, nor did Kant claim that it was; but it is a thesis that gives an important role to experience in knowledge.

One final point may be made about the British empiricists: They all employed a common method of trying to build up the body of knowledge from simple building blocks. The model for this method may have been the empirical science of the day. (Hume claimed to derive his experimental method from Isaac Newton.) The rationalists claimed more for reason and sought to reveal sources for knowledge and its materials other than experience; but they were also opposed to the empiricists in their choice of method, finding their inspiration in the method of axiomatic geometry.

JOHN STUART MILL. J. S. Mill, the main figure in nineteenth-century empiricism, followed directly in the tradition of Hume. Mill's account of our knowledge of the external world, for example, was in part phenomenalist in character; it maintained that things are merely permanent possibilities of sensation. But it was mainly an account of the way in which we come to believe in such a thing as an external world and thus followed Hume in its psychological character. In one respect, however, Mill was more radical than Hume. He was so impressed by the possibilities of the use of induction that he found inductive inference in places where we should not ordinarily expect to find it. In particular, he claimed that mathematical truths were merely very highly confirmed generalizations from experience; mathematical inference, generally conceived as deductive in nature, he set down as founded on induction. Thus, in Mill's philosophy there was no real place for knowledge based on relations of ideas. In his view logical and mathematical necessity is psychological; we are merely unable to conceive any other possibilities than those that logical and mathematical propositions assert. This is perhaps the most extreme version of empiricism known, but it has not found many defenders.

TWENTIETH-CENTURY EMPIRICISM

Empiricists in the twentieth century generally reverted to the radical distinction between necessary truths, as found in logic and mathematics, and empirical truths, as found elsewhere. Necessity is confined by them, however, to logic and mathematics, and all other truths are held to be merely contingent. Partly for this reason and partly because it has been held that the apparatus of modern logic may be relevant to philosophical problems, twentieth-century empiricists tended to call themselves "Logical Empiricists" (at least those who were connected in one way or another with logical positivism). Bertrand Russell, however, who derived something from the positivists, but who owes equally much to the British empiricists, always claimed that there are limits to empiricism, on the grounds that the principles of inductive inference cannot themselves be justified by reference to experience.

In general, twentieth-century empiricists were less interested in the question of the materials for knowledge than in that of the empirical basis for knowledge. Insofar as they considered the former question, the tendency has been, as in other matters, to eschew psychological considerations and to raise the problem in connection with meaning. All descriptive symbols, it is maintained, should be definable in terms of other symbols, except that ultimately one must come to expressions that are definable ostensively only. That is, there must ultimately be terms which can be cashed by direct reference to experience and to it alone; ostensive definition consists of giving the term together with some direct act of pointing, such that no other understanding of meaning is required. In regard to nondescriptive terms the situation is less clear, but the general tendency is to assume that the only possible source of ideas which might be called a priori is logic and mathematics. Following Russell, twentieth-century empiricists assumed that mathematical notions can be reduced to logical ones or can at least involve similar fea-

tures and that logical notions are concerned only with relations between symbols and can be defined accordingly. Russell, it is true, suggested that terms such as *or* might also be defined ostensively, for example, by reference to feelings of hesitation, but this suggestion has not been generally accepted.

If the views on the question of the materials for knowledge are not clear-cut, there has not been the same indefiniteness over the basis of knowledge. Although some positivists, the so-called physicalists, have maintained that the language of physics should be taken as providing the basic truths, most philosophers of positivist persuasion have gone to direct experience for the truths on which knowledge is taken to rest. These truths are to be found in sense-datum propositions—propositions that are a direct record of experience and which are for this reason incorrigible, consisting of ostensively definable terms, that is, names of sense data. It is not clear what would constitute an example of this. (Russell, for example, suggested "Red here now," where every expression is what he called a "logically proper name," such that its reference is guaranteed.) Nevertheless, it has been assumed that all propositions except logical ones must be reducible to these "basic propositions," which are about sense data.

However, propositions about physical objects are not incorrigible. Yet to suppose that such propositions deal with entities that lie behind the immediate data of the senses and that can only be inferred from those data would be to suppose that there is a gap between us and physical objects, the crossing of which is problematical. This would allow an opening for the skeptic. An alternative view is phenomenalism, the doctrine that the meaning of our statements about physical objects can be analyzed in terms of propositions about sense data. Physical objects are logical constructions out of sense data ("logical" because the issue concerns the correct logical analysis of propositions about physical objects and not the question of how, as a matter of psychological fact, we construct our ideas of physical objects). In general, according to positivists, all propositions other than those that are logically necessary must be verifiable by reduction, either directly or indirectly, to propositions about sense data. Anything which is not so reducible is nonsense. In epistemological terms, any contingent truth that we can be said to know must be founded on and reducible to propositions concerning sense experience. Necessary truths, it is generally held, are true by convention or in virtue of the meaning of the words involved. They tell us nothing about the world as such.

This program has run into difficulties of two main kinds. First, there have been difficulties in actually carrying out the analysis demanded. It would be almost universally agreed that propositions about physical objects cannot be analyzed in terms of propositions about actual and possible sense data, since the analysis would have to be infinitely long. This is an objection of principle. Second, the criterion of verifiability tends to exclude some kinds of propositions that we ordinarily think that we understand. There have been difficulties in this respect, for example, over propositions of natural law, as well as propositions of ethics, etc. There has been widespread dissatisfaction with attempts to justify empiricism of this sort.

It should now be possible to offer some assessment of empiricism. As an answer to skepticism it claims that the certainty and incorrigibility that knowledge demands can (apart from logical truths) be found only in immediate experience and that the rest of knowledge must be built upon this. In this sense, the theory is misguided as well as unsuccessful in carrying out its program. The lack of success can be seen in the fact that eighteenth-century empiricism led to skepticism, while the twentieth-century program of reduction was very widely admitted as a failure. The attempt was misguided in that knowledge does not require this kind of certainty and incorrigibility. Skepticism is not to be answered by providing absolutely certain truths, but by examining the grounds of skepticism itself. According to our ordinary conception of knowledge, what we claim to know must be true and based on the best of reasons. But by the best of reasons is not meant proof. Experience certainly provides justification for belief in, for example, physical objects, but if this belief is to amount to knowledge, it is not necessary that the justification should amount to proof. It is futile to argue whether experience or reason alone can provide proof of what we ordinarily claim to know. No one could have knowledge of the world unless he had experiences and could reason, but this does not mean that either experience or reason by themselves could provide the kind of absolute certainty which would constitute proof. Nor is it required that they should provide proof in order that knowledge may be possible.

What of the thesis that, whether or not experience can provide certainty, all knowledge is derived from experience? In Mill's sense, that all truths, of whatever kind, receive their validation from experience, the thesis is obviously false and need be considered no further. The thesis that all the materials for knowledge are derived from experience may seem more plausible. Yet, despite

the number of philosophers who have maintained this thesis, it is not altogether clear what it means. The version of the doctrine held by Locke and Thomas looks like a psychological account of the origin of our ideas; in logical dress it amounts to the view that all our concepts or all the words which we use are definable in terms of those which are ostensively definable. Whether or not there are any a priori notions outside logic and mathematics, it certainly seems implausible to say that logical and mathematical notions may ultimately be definable ostensively.

More important, the notion of ostensive definition is itself suspect. How could one understand what was going on when a noise was made, accompanied by a pointing to something, unless one knew the kind of thing which was being indicated and, more important perhaps, was aware that it was *language* that was being used? In other words, much has to be understood before this kind of definition can even begin. The notion that words can be cashed in terms of direct experience without further presuppositions is, thus, highly suspect. This is not to say that there are no distinctions to be made between different kinds of concepts or words, but merely that the distinctions in question cannot be made by means of any simple distinction between empiricism and rationalism.

There remains the Kantian point that the having of experience is a condition for any further knowledge. This would certainly be the case for creatures of our kind of sensibility, as Kant would put it. Yet the logical possibility of the possession of knowledge by nonsensitive creatures remains, whether or not any such creatures exist in fact.

See also A Priori and A Posteriori; Aristotle; Berkeley, George; Descartes, René; Epicurus; Hume, David; Kant, Immanuel; Locke, John; Logical Positivism; Logic, History of; Mill, John Stuart; Plato; Positivism; Pragmatism; Rationalism; Russell, Bertrand Arthur William; Sensationalism; Skepticism, History of; Thomas Aquinas, St.

Bibliography

EPICURUS

Bailey, Cyril. *Epicurus, the Extant Remains.* Oxford: Clarendon Press, 1926. In Greek, with English translation.

Bailey, Cyril. *The Greek Atomists and Epicurus.* Oxford: Clarendon Press, 1928.

Zeller, Eduard. *Stoics, Epicureans and Sceptics.* Translated by O. J. Reichel. London and New York: Longmans, Green, 1892.

THOMAS AQUINAS

Copleston, F. C. *Aquinas.* London: Penguin, 1955.

Thomas Aquinas, St. *Summa Theologica,* Ia, 78ff., in Vol. IV of the English translation by the Fathers of the English Dominican Province. London, 1922.

THE BRITISH EMPIRICISTS

Ayer, A. J., and Raymond Winch, eds. *British Empirical Philosophers.* London: Routledge, 1952. A collection of writings by the British empiricists.

Mill, J. S. *System of Logic.* 8th ed. London: Longmans, Green, Reader, and Dyer, 1872.

See also Immanuel Kant, *Critique of Pure Reason,* the translation by Norman Kemp Smith. London, 1953.

TWENTIETH-CENTURY EMPIRICISTS

Anderson, John, *Studies in Empirical Philosophy.* Sydney, Australia: Angus and Robertson, 1962.

Ayer, A. J. *Foundations of Empirical Knowledge.* London: Macmillan, 1940.

Ayer, A. J. *Language, Truth and Logic.* 2nd ed. London: Gollancz, 1946.

Ayer, A. J. *Philosophical Essays.* London: Macmillan, 1954.

Ayer, A. J. *Problem of Knowledge.* London: Macmillan, 1956.

Ayer, A. J., ed. *Logical Positivism.* Glencoe, IL: Free Press, 1959.

Lewis, C. I. *An Analysis of Knowledge and Valuation.* La Salle, IL: Open Court, 1946.

Price, H. H. *Thinking and Experience.* London and New York: Hutchison's University Library, 1953.

Russell, Bertrand. *Human Knowledge.* London: Allen and Unwin, 1948.

Russell, Bertrand. *Inquiry into Meaning and Truth.* London: Allen and Unwin, 1940.

For other writings dealing with Empiricism, see the bibliographies to the entries Logical Positivism; Positivism; Pragmatism; and Sensationalism.

D. W. Hamlyn (1967)

ENCYCLOPEDIAS

See *Encyclopédie; "Philosophy Dictionaries and Encyclopedias" in Volume 10*

ENCYCLOPÉDIE

Encyclopédie, or the French Encyclopedia, is a famous and controversial work of reference embodying much of what the French Enlightenment liked to call "philosophy."

PURPOSE, HISTORY, AND INFLUENCE

Begun simply as a commercial undertaking to translate and adapt Ephraim Chambers's *Cyclopaedia* (1728), the *Encyclopédie* was first entrusted to the Englishman John Mills and the German Godefroy Sellius, and then to the Abbé Gua de Malves of the French Academy of Sciences. Denis Diderot became chief editor in 1747 and, with Jean

Le Rond d'Alembert as his principal colleague, greatly expanded the scope of the enterprise. Diderot's prospectus (1750) promised, as a principal and novel feature, a description of the arts and especially the crafts in France, with numerous illustrative engravings, and was accompanied by an elaborate "Chart of the Branches of Human Knowledge," which Diderot referred to as "the Genealogical Tree of All the Arts and Sciences." This *Système figuré des connoissances humaines* was avowedly inspired by the work of Francis Bacon, whose empiricism greatly influenced the entire work. Assuming that all knowledge comes originally from sensations, the *Système figuré* subsumed all branches of learning under either memory, reason, or imagination, to which corresponded, respectively, history, philosophy, and poetry. The correlation of philosophy with reason, while history was associated merely with memory, was very characteristic of the Enlightenment.

The first volume of the *Encyclopédie,* which included d'Alembert's influential "Discours préliminaire," was published in 1751, and revealed at once that the work would be carried on in the spirit of John Locke's sensationalistic psychology and epistemology. Pierre Bayle, in addition to Francis Bacon and Locke, also served as a model and inspiration for the *Encyclopédie,* though its editors rarely found it expedient to admit the fact. The *Encyclopédie* was greatly influenced by Bayle's skepticism, while falling short of his thoroughgoing Pyrrhonism. The work went much beyond him, however, in its attention to natural science, to the nascent social sciences, to economic processes, and to social reform.

The first volume established the *Encyclopédie* at once as a work that was both controversial and indispensable. It was much more comprehensive than previous works of reference, and even included copious articles on grammar, synonyms, and gazetteer-like articles concerning countries and cities. It constantly attempted to explode vulgar errors (see the article "Agnus Scythicus"), to be as precise in definition as possible, to make exact technological explanation an accepted part of the language, to suggest social reforms (see the article "Accoucheuse") or greater civil liberties (see "Aius Locutius"), and to weaken dogmatisms. In biblical criticism (for example, see "Arche de Noé") or in articles touching upon political theory (for example, "Autorité politique") or materialism (for example, "*Âme*"), the *Encyclopédie* proved itself to be adventuresome and bold.

As a result, the *Encyclopédie* encountered much opposition and suspicion, especially from orthodox religious groups. In particular, the Jesuits, whom Diderot and d'Alembert suspected of wanting to take over the editing of the work for themselves, delighted in exposing plagiarisms in the *Encyclopédie* and in insinuating that it was subversive. In 1752, just after the publication of the second volume, the Royal Council of State prohibited further publication, although, a few months later, this decree was tacitly rescinded. Thereafter, the *Encyclopédie* was published at the rate of a volume per year until 1757, when it had reached through the letter *G.* By this time it was evident, as Diderot himself had stated in his remarkable article "Encyclopédie" in volume five, explaining the intentions and editorial policies of the work, that the object of the *Encyclopédie* was "to change the general way of thinking."

In 1757 there commenced a long and complicated crisis that resulted in d'Alembert's retiring from his part in the editing and finally in the suppression of the work by royal decree, on March 8, 1759.

Nevertheless, through the courage and tenacity of Diderot and the publishers, and as a result of the authorities studiously looking the other way, the work continued to be written, edited, and printed in secret, pending the time when it might once more be authorized. In 1765–1766, the rest of the alphabet (ten volumes of letterpress) was published. Meanwhile, the 11 volumes of plates were also being prepared and published under Diderot's supervision, the first appearing in 1762 and the last in 1772. About 4,225 sets of the original edition were sold, the price being 980 livres (326 for the 17 volumes of letterpress and 654 for the 11 volumes of plates). Inasmuch as the purchasing power of a livre was roughly equivalent to rather more than a dollar in current (1966) purchasing power, it is evident that this was a large commercial undertaking.

Each of the first seven volumes of the *Encyclopédie* had been subjected to previous censorship, but this was impossible with the last ten volumes, because they were edited secretly. There was, therefore, a considerable risk that the government might outlaw the whole edition if the articles were too forthright on theology and politics. In the end, there was little difficulty: By 1765–1766, when the final volumes were distributed, the order of the Jesuits had been suppressed and public opinion generally was moving irresistibly toward the point of view represented by the *philosophes.* But Andre-François Le Breton, the printer and chief publisher of the *Encyclopédie,* had meanwhile surreptitiously altered many of the most controversial articles after Diderot had edited them and read the proofs. Diderot discovered this treachery in 1764, too late to undo it. The subsequent discovery of a volume of

proof sheets permits a before-and-after comparison of some of the articles mutilated by Le Breton; a study of these shows that the changes were substantial. The exact number of Le Breton's alterations is not known even yet, though Diderot always remained convinced that the publisher's depredations had been extensive. In spite of the maiming of the text, however, the articles in the last ten volumes are rather more sharp and critical about religious, social, and political topics than the first seven volumes had dared to be.

One of the novel features of the *Encyclopédie* was that it identified many of its contributors, the most famous being Diderot, d'Alembert, Voltaire, Jean-Jacques Rousseau ("Économie politique" and articles on music), Baron de Montesquieu ("Goût"), François Quesnay ("Fermiers," "Grains"), Baron de L'Aulne Turgot ("Étymologie," "Existence"), Jean-François Marmontel, Baron d'Holbach, and Louis de Jaucourt. After the suppression of the work in 1759, many of the contributors (a total of 160 have been identified) discontinued their collaboration, thus greatly increasing the burden on Diderot. The *Encyclopédie* represented the greatest feat in the technology of printing and publishing up to that time. It was a symbol of the intellectual preeminence of France in the eighteenth century. But it was also the symbol of a new public philosophy; and its final publication, with editorial policies and practices consistent and unchanged, was a triumphant vindication of the energy and moral courage of Diderot and even, though to a lesser extent, of his publishers.

PHILOSOPHY IN THE ENCYCLOPÉDIE

The numerous and lengthy articles in the *Encyclopédie* concerning philosophers or schools of philosophy, from "Aristotélisme" to "Zend-Avesta," constituted in themselves a stage in the development of recording the history of philosophy. Most of these articles were written by Diderot himself. In the compilation of them, he avowedly relied upon works by Thomas Stanley and Boureau Deslandes and, very heavily, upon Johann Jacob Brucker's *Historia Critica Philosophiae* (Leipzig, 1742–1744). But Brucker's work, relaxed in style and blandly deistic, was changed by Diderot into a history of philosophy that was nervous and sometimes edgy in style and, in its implicit challenging of idealism and in its inclination toward materialism, very representative of the point of view of the Enlightenment in France. Some of the articles not written by Diderot are flabby or conformist in their thought (for example, "Aristotélisme," "Spinoza"), but Diderot's own most famous ones ("Chaldéens," "Cyniques," "Cyrénaique," "Éclectisme," "Éléatique," "Épicuréisme," "Hobbisme," "Leibnitzianism," "Platonisme," "Pyrrhonienne") substantiate the claim that through the *Encyclopédie* Diderot was one of the creators of the history of philosophy in France.

ONTOLOGY AND EPISTEMOLOGY. It was a favorite sport of the Encyclopedists to inveigh against "metaphysics." This criticism was primarily an expression of their dislike for the great rationalistic constructions of the seventeenth century, the systematic philosophy of René Descartes, Nicolas Malebranche, Benedict Spinoza, and Gottfried Wilhelm Leibniz. In reality, since the Encyclopedists—like the logical positivists of the twentieth century—had a theory of being and a theory of knowledge, they were more metaphysical than they acknowledged or perhaps realized. The Encyclopedists predicated a real world of brute fact, and steadfastly resisted the Berkeleian philosophy, although they were familiar with it (see d'Alembert's article, "Corps"). This real world was knowable, according to the Lockean system of epistemology, through the testimony of the senses and reflection thereon. Diderot stated, for example, in the article "Inné" that "there is nothing innate except the faculty of feeling and of thinking; all the rest is acquired." Such reference to external reality interpreted by reason, led to the great emphasis given by the Encyclopedists to *expérience,* which in the French of their day had the double meaning of experiential and experimental (see d'Alembert's article, "Expérimental").

With this empirical approach to the problems of reality and knowledge, the *Encyclopédie* contributed greatly to the strengthening of the rationale of scientific hypothesis and scientific method (see, for example, "Hypothèse"). In this respect, especially noteworthy in the articles written by d'Alembert (for example, "Cosmologie" and "Cartésianisme"), the *Encyclopédie* was a forerunner to the development of positivism. Nor were the Encyclopedists lamed by Humean skepticism. They knew David Hume personally and loved him and had read his books, but they simply overlooked the implications of Hume's philosophy in respect to their own ontology and epistemology. The sensationalistic psychology of the Encyclopedists, in combination with their view of the world, strengthened them in their faith in reason, by which it was deemed possible to know and evaluate objective reality, while making it unnecessary for them to have much faith in faith. The philosophy of the *Encyclopédie* was about as far from fideism as it is possible to be.

OPPOSITION TO RELIGIOUS DOGMATISM. The *Encyclopédie* was often accused by its enemies of favoring a philosophy of materialism. This it never did outright, yet many of its articles pointed that way, especially those that had to do with the mind-body problem (for example, "Spinosiste," "Âme"). Moreover, the Encyclopedists were constantly eager to undermine dogmatic and intolerant religious orthodoxy. This function they considered as one of their most "philosophical," and it is in this connection that they helped to establish a new historiography. The Encyclopedists often wrote as though they were historical pessimists and indeed distrustful of history: "One can scarcely read history without feeling horror for the human race," wrote Voltaire in "Idole, idolatrie." Nevertheless, in their desire to shake religious dogmatism, they used criteria of historical criticism, for example, in trying to establish the correct chronology of the Bible (see "Chronologic sacrée"), and explored the nature of historical evidence (for example, as to miracles) in a way that secularized and modernized historical techniques. (In this respect the articles "Bible," "Certitude," "Mages," "Syncrétistes," are of particular interest.) As for the philosophy of history, the Encyclopedists' convictions regarding the spread of enlightenment led to a faith in progress which became one of the conspicuous features of eighteenth-century thought.

ETHICS. The *Encyclopédie* was much concerned with ethics, especially because of its insistence, as expressed by Diderot in "Irréligieux," that "morality can exist without religion; and religion can coexist, and often does, with immorality." In ethical theory many of the articles still spoke in terms of *jus naturae,* and sometimes, as in "Irréligieux," identified this moral law as "the universal law that the finger of God has engraved upon the hearts of all." But this rather conventional ethics was constantly being blended with, or superseded by, utilitarianism. The articles in the *Encyclopédie* advanced a theory of ethics that was founded not so much in the will of God as in the nature of man. And inasmuch as man was conceived of as being by nature sociable, it logically followed that an ethic grounded in man's nature was also socially conscious and other-regardful. The *Encyclopédie* also endeavored to undermine notions of free will, teaching that man, precisely because he is modifiable and educable, is capable of virtue even in a deterministic universe (see "Liberté," "Modification," "Malfaisant").

SOCIAL AND POLITICAL THEORY. The social philosophy of the *Encyclopédie* was shaped in like manner by the conviction that man by his nature is sociable (see "Philosophe"). As a result, the *Encyclopédie* was much interested in theories of social origins, and devoted a good deal of attention to the ethnography of primitive peoples, using travel books as a principal source. The article on "Humaine espèce" is a remarkable exercise in physical anthropology; and articles such as "Laboureur," "Journalier," and "Peuple" are examples of a groping toward a recognizable sociology. Thus, the *Encyclopédie* figured importantly in the development of the social sciences, as well as in the dissemination of a utilitarian social philosophy. The *Encyclopédie* had a passion for improvement and constantly applied to institutions the criterion of social usefulness.

The *Encyclopédie* also possessed a quite clearly articulated political theory, even though it was difficult to discuss political philosophy critically in a country that was professedly an absolute monarchy and exercised censorship. This political philosophy was, as might be expected, greatly influenced by Locke. Articles such as "Droit naturel" and "Égalité naturelle" spoke of "inalienable rights" and continued, as Locke and Samuel von Pufendorf had done, to explore the implications to political philosophy of new and emerging insights into the nature of man. In articles such as "Autorité politique" and "Loi fondamentale," the *Encyclopédie* praised limited monarchy and suggested that proper government rests upon consent (see "Pouvoir"). In the article "Représentants" a theory of representative government was advanced, and numerous articles suggested the guarantee of civil liberties (for example, "Habeas corpus," "Aius locutius," "Libelle") or advocated reforms ("Impôt," "Vingtième," "Privilège"). An English writer, reviewing the *Encyclopédie* in 1768, remarked that "whoever takes the trouble of combining the several political articles, will find that they form a noble system of civil liberty."

LINGUISTIC THEORY. The *Encyclopédie* was much engrossed in theories regarding the origin of language, and devoted a great deal of space to articles on grammar and on synonyms. In part this was social philosophy, in the sense that it was hoped that such speculation would throw light upon social origins; even more, it was an early manifestation of scientific and philosophical interest in the nature of language. In articles such as "Étymologie," "Élémens de science," and "Encyclopédie," Turgot, d'Alembert, and Diderot, respectively, analyzed problems of definition, semantics, and nomenclature in the attempt to explore accurately the relationship between words, concepts, and things. The Encyclopedists were remarkable for realizing that knowledge itself depends upon the correct use of language.

AESTHETICS. Aesthetic theory was not systematically developed in the *Encyclopédie,* although there were numerous articles on belletristic subjects, especially those contributed by Jean-François Marmontel (see his article "Critique") and Voltaire. Special mention should be made of Diderot's articles "Beau" and "Beauté," which reviewed extensively the aesthetic theories current in the first half of the eighteenth century and argued that it is the perception of relationships that is the basis of the beautiful.

HUMANISM. The philosophy of the *Encyclopédie* was strongly humanistic in tone. Oriented toward science, and progressive (in the sense of believing in progress), the work was integrated by the particular philosophy of man that underlies the whole. It was a philosophy, Protagorean in savor, that made man the measure of all things. This point of view was summed up by Diderot in the article "Encyclopédie": "Man is the sole and only limit whence one must start and back to whom everything must return."

See also Aesthetics, History of; Alembert, Jean Le Rond d'; Bacon, Francis; Bayle, Pierre; Berkeley, George; Descartes, René; Diderot, Denis; Enlightenment; Epistemology; Ethics; Holbach, Paul-Henri Thiry, Baron d'; Humanism; Hume, David; Language; Leibniz, Gottfried Wilhelm; Locke, John; Logical Positivism; Malebranche, Nicolas; Metaphysics, History of; Montesquieu, Baron de; Ontology; Pufendorf, Samuel von; Rousseau, Jean-Jacques; Semantics; Spinoza, Benedict (Baruch) de; Turgot, Anne Robert Jacques, Baron de L'Aulne; Utilitarianism; Voltaire, François-Marie Arouet de.

Bibliography

EDITIONS

Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers, par une société de gens de lettres …, 35 vols. in folio. Paris, 1751–1780. Diderot was editor-in-chief for 17 volumes of letterpress (1751–1766) and 11 volumes of plates (1762–1772); these 28 volumes were reprinted in folio at Geneva (1772–1776). The remaining volumes consist of four volumes of *Supplément,* one volume of supplementary engravings, and two volumes of index. Other editions, all published in French, appeared at Lucca (28 vols. in folio, 1758–1771), Livorno (33 vols. in folio, 1770–1779), Yverdon (58 vols. in quarto, 1770–1780), Geneva (45 vols. in quarto, 1777–1781), and Lausanne and Berne (36 vols. in octavo, 1778–1781). Reproductions of 485 of the original engravings are available in the admirably edited and inexpensive *Diderot, Denis, Pictorial Encyclopedia of Trades and Industry,* edited by Charles C. Gillispie. 2 vols. New York: Dover, 1959.

HISTORY OF THE ENCYCLOPÉDIE

Gordon, Douglas H., and Norman L. Torrey. *The Censoring of Diderot's Encyclopédie and the Re-established Text.* New York: Columbia University Press, 1947.

Grosclaude, Pierre. *Un audacieux message. L'Encyclopédie.* Paris: Nouvelles Editions Latinos, 1951.

Kafker, Frank A. "A List of Contributors to Diderot's Encyclopedia." *French Historical Studies* 3 (1963–1964): 106–122.

Le Gras, Joseph. *Diderot et l'Encyclopédie.* Amiens: Malfère, 1928.

Lough, John. "Luneau de Boisjermain v. the Publishers of the *Encyclopédie.*" *Studies on Voltaire and the Eighteenth Century* 13 (1963): 115–177.

Venturi, Franco. *Le origini dell'Enciclopedia.* Florence: U Edizioni, 1946; 2nd ed., 1963.

Wilson, Arthur M. *Diderot: The Testing Years, 1713–1759.* New York: Oxford University Press, 1957.

INTELLECTUAL AND PHILOSOPHICAL ASPECTS

Barker, Joseph E. *Diderot's Treatment of the Christian Religion in the Encyclopédie.* New York: King's Crown Press, 1941.

Delorme, Suzanne, and René Taton, eds. *L'Encyclopédie et le progrès des sciences et des techniques.* Paris, 1952.

Havens, G. R., and D. F. Bond, eds. *A Critical Bibliography of French Literature.* Syracuse, NY: Syracuse University Press, 1951. Vol. IV, *The Eighteenth Century,* pp. 139–141.

Hubert, René. *Les sciences sociales dans l'Encyclopédie.* Lille: Au Siège d'Université, 1923.

"Numéro spécial à l'occasion du 2^2 centenair de l'*Encyclopédie.*" *Annales de l'Université de Paris* October 1952.

Proust, Jacques. *Diderot et l'Encyclopédie.* Paris: A. Colin, 1962. Especially valuable.

Schalk, Fritz. *Einleitung in die Encyclopädie der französischen Aufklärung.* Munich, 1936.

Schargo, Nelly. *History in the Encyclopédie.* New York: Columbia University Press, 1947.

Weis, Eberhard. *Geschichtsschreibung und Staatsauffassung in der französischen Enzyklopädie.* Wiesbaden: F. Steiner, 1956.

Arthur M. Wilson (1967)

ENERGETICISM, ENERGETISM

See *Ostwald, Wilhelm*

ENERGY

Energy, from the Greek *energeia* (*en,* in; *ergon,* work), originally a technical term in Aristotelian philosophy denoting "actuality" or "existence in actuality," means, in

general, activity or power of action. In the physical sciences it is defined as the capability to do work, as accumulated work or, in the words of Wilhelm Ostwald, as "that which is produced by work or which can be transformed into work." Energy is measured in terms of units of work, to overcome a resisting force of one dyne over a distance of one centimeter. (The joule = 10^7 erg = the watt-second; the kilogram-meter = 9.81×10^7 erg. In atomic physics the unit is the electron volt; $ev = 1.6 \times 10^{-12}$ erg.

In physics, energy is either kinetic or potential. A body of mass *m* moving with a velocity *v* possesses, owing to its motion, the kinetic energy $\frac{1}{2}mv^2$, which is the work necessary to overcome the inertial resistance in accelerating the body from rest to its final velocity and which is again transformed into work if the body is brought to rest. The energy that a system of bodies possesses by virtue of the relative geometrical position of its constituent parts, if subjected to gravitational, elastic, electrostatic, or other forces, is its potential energy. If, for example, a stone is raised from the surface of the earth, the potential energy of the system stone-and-earth is increased; if an elastic spring is expanded, its potential energy increases with increase of length. The attribute "potential" thus merely characterizes the latency of temporarily stored energy and does not call into question the reality of this kind of energy. With the recognition of the principle of the conservation of energy, it became apparent that the concept of energy applies to all branches of physics and to all physical sciences. Because of the at least partial convertibility of any energy into mechanical work, the aforementioned units of work also serve as measures of such energies as thermal, electric, magnetic, acoustic, and optical. For thermal energy (heat) it proved practical also to retain as a separate unit the caloric unit of heat, the calorie (equal to 4.18×10^7 erg).

HISTORY OF THE CONCEPT

In spite of its universality, the general notion of energy as a basic concept in science is a relatively recent result of a long and intricate conceptual process. From the scientific point of view this process may conveniently be divided into five consecutive stages: (1) early conceptions of energy as a source of force, (2) the rise of the concept of mechanical work, (3) the recognition of different forms of energy, of their interconvertibility, and of the conservation of their sum total, (4) the emancipation of energy as an autonomous existent, and (5) the mathematization of energy as an integral invariant. From the philosophical point of view—that is, with respect to the ontological and epistemological status of the concept of energy—one may speak of (1) accidental, (2) substantial, (3) relational, (4) causal, and (5) formal conceptions of energy.

ENERGY AND FORCE. Aristotle was the first to use *energeia* as a technical term in his conceptual scheme, where it often signified the progressive "actualization" of that which previously existed only in potentiality. He also seems to have formed, though in an implicit manner, the idea of energy in the sense of accumulated force or accumulation of force. Force, for him, was not only the cause of motion but also the factor determining the duration or extent of motion. In the *Physics* he formulated the fundamental law of his dynamics, which, in modern terminology, states that the velocity, *D/T* (distance divided by time), of a mobile is proportional to the ratio of the magnitude of the moving force, *A,* and the resistance, *B,* a relationship that he described by enumerating exhaustively all possibilities under which *AT/BD* remains constant (with the exception of doubling the distance, *D,* as well as the time, *T*). He argued that a given finite force cannot move a mobile over an infinite distance or for an infinite time. Aristotle thus associated with every force a capacitative limitation, or, in modern terms, an energy content.

The implications of this statement for cosmology—in particular, for the motion of the celestial spheres, which derive their eternal motion ultimately from the "first mover" in accordance with the axiom "all things that are in motion must be moved by something else"—called for further clarifications. Thus, for example, Averroes, in his "Commentary on the Physics," distinguished between the primary motive force, the *motor separatus,* and the secondary forces, the *motores coniuncti*; the latter, in direct contact with the spheres, corresponding to the medieval "intelligences," draw finite quotas of force from the inexhaustible supply of the former. By this process, according to which only finite amounts are subtracted from an infinite accumulation of force, Averroes thought he was able to explain both the eternity of celestial motion and the fact that this motion does not occur instantaneously (*in instanti*), as motion under the effect of an infinite cause should do.

Considerations of this kind, which engaged Aristotelian commentators until the times of Thomas Aquinas, show clearly that the notion of force signified not only the immediate cause of motion or acceleration but also its cumulative determination, or energy content. Thomas considered the possibility of a finite and yet invariable moving force, which, being immutable, acts

always in the same manner (*vis infatigabilis*), and thus he conceived of force as a moving agent independent of and separated from a constantly rejuvenating source, a notion essential for the future conception of the universe as a clockwork in action without the need of a constant supply of additional energy. Early in the fourteenth century the nominalist Peter Aureoli, in *Liber Sententiarum,* distinguished explicitly between two different aspects of force: its velocity-determining property and its capacity of consumption, or measure of exhaustibility. His differentiation can rightfully be regarded as the first ontological distinction between force and energy.

This, of course, does not imply that allusions to particular forms of energy are not found in early scientific writings. In fact, already in the *Mechanica,* commonly ascribed to Aristotle, the notion of kinetic energy is clearly referred to when it is asked:

> How is it that, if you place a heavy axe on a piece of wood and put a heavy weight on the top of it, it does not cleave the wood to any considerable extent, whereas, if you lift the axe and strike the wood with it, it does split it, although the axe when it strikes the blow has much less weight upon it than when it is placed on the wood and pressing on it? It is because the effect is produced entirely by movement, and that which is heavy gets more movement from its weight when it is in motion than when it is at rest.

MECHANICAL WORK. The modern concept of energy, as the definition shows, is a generalization of the notion of work in mechanics. The concept of work can be traced back to the principle of virtual displacements, or virtual velocities, which, in turn, has its ultimate origin in Aristotelian dynamics. Aristotle's conclusions (in *De Caelo*) concerning one single force (under whose action "the smaller, lighter body will be moved farther …; for as the greater body is to the less, so will be the speed of the lesser body to that of the greater") were soon generalized for the case of a force counteracting a load, as exemplified in simple machines such as the wheel and the axle. In particular, the study of the law of the lever, as mentioned in the *Mechanica,* in Archimedes' *On the Equilibrium of Planes,* in the writings of Hero of Alexandria, and in the *Liber Karastonis,* a Latin version of the ninth-century Arabic text by Thabit ibn Kurrah, contributed to the gradual establishment of the principle of virtual displacements for which finally, in the thirteenth century, Jordanus Nemorarius tried to give a theoretical proof. The Renaissance formulation of this law—namely, that the ratio between force and load is reciprocal to that of the spaces (distances) traversed within the same time—as pronounced by Guidobaldo del Monte (*Mechanica,* 1577), by Simon Stevin (*Hypomnemata Mathematica,* Leiden, 1608, Book 3), and by Galileo Galilei (*Opere* 2), formed the basis for the definition of work as force times distance traversed.

Pierre Varignon, in his *Nouvelle Mécanique ou statique* (Paris, 1725), reported a letter from Johann Bernoulli, dated January 26, 1717, in which the term *energy* appears in this connection, apparently for the first time in the modern period: "For all equilibrium of forces in whatever manner they are applied to each other, whether directly or indirectly, the sum of the positive energies will be equal to the sum of the negative energies taken positively." Although some historians, referring to this letter, have ascribed to Bernoulli the definition of energy as "force times distance," a critical study of the text shows undoubtedly that he still defined energy as "force times virtual velocity." In spite of the fact that this notion and its derivative, namely, the notion of work defined as "force times distance," played at least implicitly an important part in the establishment of classical mechanics—Joseph Louis Lagrange saw in the principle of virtual velocities the fundamental basis for his *Mécanique analytique* (1788), the highlight of classical mechanics—energy considerations were rarely found in theoretical or even practical mechanics prior to the middle of the nineteenth century. Before the development of the steam engine and the rise of thermodynamics, industry had little interest in energy calculations: Force, not its integrated form, counted in the use of simple machines. The primary object of theoretical mechanics, moreover, was still celestial dynamics, where, again, energetics was of little avail. This certainly is also one of the reasons why Isaac Newton's *Principia* contains practically no reference to the concept of energy or to any of its applications.

According to Ernst Mach, in *Die Mechanik in ihrer Entwicklung* (Leipzig, 1883; translated as *The Science of Mechanics,* La Salle, IL, 1942), the delay of the development of energetics as compared with that of general mechanics stemmed from what he called "trifling historical circumstances," namely, the fact that in Galileo's investigations of free fall, the relationship between velocity and time was established before the relationship between velocity and distance, so that, as multiplication with mass shows, the notions of quantity of motion or momentum and force gained priority and were regarded as more fundamental than the concept of energy, which thus appeared as a derived conception. Whatever the reason for energetics' lagging behind Newtonian mechanics,

it is an indisputable fact that the concept of energy became a subject of discussion among philosophers rather than among physicists or mechanicians.

THE MEASURE OF "FORCE." Foremost among the philosophical discussions was the controversy between the Cartesians and Gottfried Wilhelm Leibniz over whether the true measure of "force" (i.e., energy) is momentum (the product of mass and velocity) or *vis viva* (as defined by Leibniz, the product of mass and the square of velocity). René Descartes, having shown in his *Principles of Philosophy* that the (scalar) quantity of motion or momentum (the vectorial nature of this quantity was recognized only by Christian Huygens) is conserved, concluded that momentum is the measure of energy. Leibniz, in "A Short Demonstration of a Remarkable Error of Descartes" ("Brevis Demonstratio Erroris Memorabilis Cartesii," in *Acta Eruditorum,* 1689), opposed this view. Lifting a load of 1 pound, he claimed, to a height of 4 feet requires the same work as lifting 4 pounds to the height of 1 foot. Since, according to Galileo, the velocities (of free fall) are proportional to the square roots of the heights (of fall), the velocity of the first object is twice that of the second before reaching ground, or $v_1 = 2v_2$. Assuming that the "forces" (energies) are proportional to the masses (moles), Leibniz concluded that $m_1 \cdot f(v_1) = m_2 \cdot f(v_2)$, where $f(v)$ is an as yet unknown function of the velocity, v. Substituting $m_2 = 4m_1$ and $v_1 = 2v_2$ yields $f(2v_2) = 4 \cdot f(v_2)$, which shows that the unknown function is quadratic in its argument, v. What is conserved and hence is the measure of "force," Leibniz concluded, is mv^2.

This controversy between the Leibnizians, among them Johann Bernoulli, Willem Jakob Gravesande, Christian von Wolff, Georg Bilfinger, and Samuel König, and the Cartesians, among them Colin Maclaurin, James Stirling, and Samuel Clarke, was essentially only a battle of words, since the Leibnizians considered force acting on bodies traveling over equal distances and the Cartesians considered force acting on bodies during equal intervals of time, as Jean Le Rond d'Alembert in *Traité de dynamique* (1743) and Lagrange in *Mécanique analytique* (1788) made clear.

CONSERVATION OF "FORCE." The interesting aspect of the Leibnizian-Cartesian controversy is the fact that both sides argued on the basis of the conservation of their respective "measures": for the Cartesians it was the conservation of momentum, for the Leibnizians that of "living force" (kinetic energy). Both contentions, as we know today, were correct, since both measures are integrals of the equations of motion. One of the most ardent supporters of Leibniz was his disciple Christian von Wolff, who in the *Cosmologia Generalis* (1731) declared: "In all the universe the same quantity of living force is always conserved." Johann Bernoulli, in the essay "De Vera Notione Virium Vivarum" (in *Acta Eruditorum,* 1735), was probably the first to treat this statement of the *conservatio virium vivarum* as a fundamental principle in mechanics. The apparent loss of "living force" in inelastic collisions was usually explained away by the hypothesis that the invisible small parts of matter gain in *vis viva* just as much as the macroscopic bodies seem to lose, a view Leibniz had already expressed in *Essai de dynamique* and reaffirmed in a letter to Samuel Clarke (Fifth Letter, August 18, 1716), where he stated that "active forces are preserved in the world" and continued: "'Tis true, their wholes (unelastic colliding bodies) lose it with respect to their total motion; but their parts receive it, being shaken by the force of the concourse. And therefore that loss of force is only in appearance. … the case here is the same, as when men change great money into small." Johann Bernoulli, in contrast, explained this apparent loss as an absorption of force required for the compression of the colliding bodies.

TRANSFORMATION OF POTENTIAL ENERGY. What Bernoulli had in mind was obviously the so-called latent force, subsequently to be called potential energy, and his is the earliest description of transformation of kinetic energy into potential. The idea of such "latent force" was soon generalized to nonmechanical processes. Already in 1738 Daniel Bernoulli, in his *Hydrodynamica, sive de Viribus et Motibus Fluidorum Commentarii,* spoke of the "latent force" of combustible coal, which "if totally extracted from a cubic foot of coal and used for the motion of a machine, would be more efficient than the daily work of eight or ten men." But the measure of this "latent living force" was still mv^2.

Strictly speaking, the notion of potential—that is, a function whose space derivatives yield the force components and which therefore equals the potential energy for a unit of mass, charge, etc.—preceded the idea of potential energy. For in 1777, Lagrange, in "Recherches sur l'attraction des spheroides homogènes" (*Mémoire de l'Académie,* Paris), calculated the potential for an arbitrary discrete distribution of mass particles, and in 1782, Pierre Simon de Laplace calculated the potential for a continuous distribution. Potentials were still spoken of as "force functions"; the term *potential function* was introduced for the first time in 1828 by George Green in his *Essay of the Application of Mathematical Analysis* and later (1840), independently, by Karl Gauss.

When, in 1788, Lagrange derived the principle of the conservation of mechanical energy, or what subsequently was generally called the "theorem of the living force," as an integral of the equation of motion, he asked himself how many such integrals exist and under what conditions. The question, however, whether a similar principle exists also for nonmechanical processes did not occur to him.

The first clear and consistent terminology of energy conceptions, still in the domain of mechanical processes, was used by the Paris school of practical mathematicians and mechanicians, not by the purely analytical school headed by Lagrange and Laplace. It was Lazare Carnot who, in his *Essai sur les machines en général* (1783; republished in 1803 in a revised and enlarged edition under the title *Principes fondamentaux de l'équilibre et du mouvement*), declared that the "living force" can manifest itself either as mv^2 or as *Fd* (force times distance), the second being a measure of the "latent living force." Jean V. Poncelet, in *Mécanique industrielle* (1829), finally introduced for this quantity the term *mechanical work* and stated distinctly that it is the inertia of masses that serves for the accumulation of work and thus enables the transformation of work into "living force" and vice versa. Poncelet also measured this quantity by the kilogrammeter, a unit of energy universally adopted since then.

We thus see how at the beginning of the nineteenth century the notions of work and living force and their transformability became firmly established within the confines of mechanics proper. Even the *energy* was used in this connection. In *A Course of Lectures on Natural Philosophy* (London, 1807), Thomas Young, though an adherent of the Cartesian measure of force, admitted that "in almost all cases of the forces employed in practical mechanics, the labour expended in producing any motion, is proportional not to the momentum, but to the energy which is obtained." But it took another fifty years until the term *energy* in its present meaning acquired full citizenship within the vocabulary of the physical sciences. This was brought about from quite a different quarter. It derived from the study of those phenomena where heat and chemical change are the characteristic features.

CONVERSION PROCESSES. Although Francis Bacon, in his *Novum Organum,* had already stated that "the very essence of heat, or the substantial self of heat, is motion, and nothing else," and although similar statements had been made even before the seventeenth century, the late eighteenth century, in general, interpreted heat as a fluidum, in the spirit of the phlogiston theory. Still Jean B. J. Fourier, in his *Théorie analytique de la chaleur* (1822) declared: "Thermal processes are a special kind of phenomena which cannot be explained by the principle of motion and of equilibrium." Although Joseph Black's doctrine of latent heat accounted for the disappearance of heat on the basis of the fluidum theory, the appearance of heat, as Count Rumford's experiments, at Munich in 1796 and 1798, with the boring of cannon clearly showed, was incompatible with this theory. Having eliminated all sources from which the heat produced during the boring could have originated, Rumford concluded that "it appears to be extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner the heat was excited and communicated in those experiments, except it be motion."

At the same time (1799) Humphry Davy performed at the Royal Institution in London his famous experiment in which two pieces of ice were rubbed together by a clockwork mechanism in a vacuum, the whole apparatus being maintained at the freezing point of water. Davy concluded that heat was "a peculiar motion, probably a vibration of the corpuscles of bodies" (*Essay on Heat, Light, and the Combinations of Light,* London, 1799). Rumford's and Davy's experiments, though in their quantitative aspects not yet fully explored, suggested the interchangeability of heat and motion and thus led to the more general idea of an interconvertibility, or "correlation," of the forces of nature, previously regarded as disparate and incommensurable.

Approaching this problem from a chemical and biological point of view, Justus von Liebig, one of the earliest investigators of the economy of living organisms, advanced the theory that the mechanical energy of animals, as well as the heat of their bodies, originated from the chemical energy of their food. Such physiological experiments as those carried out in Liebig's laboratory made possible the study of conversion processes and together with increased concern with engines and natural philosophical considerations, seem to have been responsible for the independent discoveries, between 1837 and 1847, of the principle of energy conservation. In fact, Liebig's pupil Friedrich Mohr, adopting the mechanistic view that all forms of energy are manifestations of mechanical force, wrote as early as 1837: "Besides the known fifty-four chemical elements there exists in nature only one agent more, and this is called 'Kraft' ['force']; it can under suitable conditions appear as motion, cohesion, electricity, light, heat, and magnetism."

ENERGY CONSERVATION PRINCIPLE. Robert von Mayer, a physician from Heilbronn, Bavaria, who had served on a ship in the tropics, had noted that the venous blood of his patients there was redder than it had been in Europe. He explained this difference by an excess of oxygen due to a reduced combustion of the food that provided the heat of the body. He thus concluded that chemical energy, heat of the body, and muscular work are interconvertible, an idea that he pursued upon his return by a quantitative investigation of the mechanical equivalent of heat. The first enunciation of the energy conservation principle, combined with the determination of the mechanical equivalent of heat, is found in Mayer's article "Bemerkungen über die Kräfte der unbelebten Natur" (in Liebig, ed., *Annalen der Chemie und Pharmacie,* 1842, Vol. XLII, pp. 233–240). His calculations, as explained in greater detail in his *Die organische Bewegung* (1845) were based on the difference of the specific heats of air at constant volume and at constant pressure, as measured by F. Delaroche and others, yielding, in modern units, 3.65 joule per calorie; had Mayer employed Henri Regnault's more accurate results he would have arrived at 4.2 joule per calorie, the currently accepted value. The amount of heat liberated by the expenditure of mechanical or electrical work was systematically measured by James Prescott Joule, a Manchester brewer and amateur scientist. In heating liquids by the rotation of paddle wheels, forcing water through narrow tubes, or compressing masses of air, Joule demonstrated that the expenditure of the same amount of work, irrespective of the manner in which this work was done, resulted in the development of the same amount of heat. His measurements of such conversion processes gave a firm quantitative support for the conservation principle.

The discovery of the physical principle of the conservation of energy was soon found to be in full agreement with the principal tenets of the prevailing natural philosophy, the German *Naturphilosophie,* whose early proponent, Friedrich Wilhelm Joseph von Schelling, had declared in 1799, in *Einleitung zu dem Entwurf eines Systems der Naturphilosophie,* "that magnetic, electrical, chemical, and finally even organic phenomena would be interwoven into one great association … [which] extends over the whole of nature." Mayer supported his own conclusions by the metaphysical argumentation that forces are essentially causes and "causes equal effects"; since causes are indestructible and convertible into effects, forces must likewise be indestructible and interconvertible. Even the experimentalist Joule, in an article "On the Calorific Effects of Magneto-electricity, and on the Mechanical Value of Heat" (*Philosophical Magazine,* series 3, 23 [1843]: 442), declared: "I shall lose no time in repeating and extending these experiments, being satisfied that the grand agents of nature are by the Creator's fiat indestructible." In another paper (in *Philosophical Magazine,* series 3, 26 [1845]: 382) he stated: "Believing that the power to destroy belongs to the Creator alone, I entirely coincide with Roget and Faraday in the opinion, that any theory which, when carried out, demands the annihilation of force, is necessarily erroneous." The conduciveness of the philosophical climate toward the enunciation of the energy principle can most clearly be recognized from the arguments of A. Colding, who arrived at the principle independently of Mohr, Mayer, and Joule:

> The first idea I conceived on the relationship between the forces of nature was the following. As the forces of nature are something spiritual and immaterial, entities whereof we are cognizant only by their mastery over nature, these entities must of course be very superior to everything material in the world; and as it is obvious that it is through them only that the wisdom we perceive and admire in nature expresses itself, these powers must evidently be in relationship to the spiritual, immaterial, and intellectual power itself that guides nature in its progress; but if such is the case, it is consequently quite impossible to conceive of these forces as anything naturally mortal or perishable. Surely, therefore, the forces ought to be regarded as absolutely imperishable. ("Nogle Soetninger om Kraefterne," 1843, in *Philosophical Magazine,* series 4, 27 [1864]: 56–64).

Even the classic paper of Hermann von Helmholtz, the physiologist turned physicist, "On the Conservation of Force" (*Über die Erhaltung der Kraft,* Berlin, 1847), shows clearly the impact of contemporaneous philosophy, with its renunciation of Hegelianism and its reversion to an idealistic rationalism, when it declares:

> The final aim of the theoretic natural sciences is to discover the ultimate and unchangeable causes of natural phenomena. Whether all the processes of nature be actually referrible to such—whether changes occur which are not subject to the laws of necessary causation, but spring from spontaneity or freedom, this is not the place to decide; it is at all events clear that the science whose object it is to comprehend nature must proceed from the assumption that it is comprehensible.

The requirement of referring the phenomena of nature back to unchangeable final causes was interpreted by Helmholtz as reducing physical processes to motions of material particles possessing unchangeable moving forces that are dependent on conditions of space alone. Thus, Helmholtz, starting with the eighteenth-century dynamics of bodies acting under mutual attraction, generalized the Newtonian conception of motion to the case of a large number of bodies and showed that the sum of force and tension (what we now call kinetic and potential energies) remain constant during the process of motion. Applying conventional analytical mathematics, Helmholtz proved that the principle of the conservation of living force not only can be derived from Newtonian dynamics but may also serve as an equivalent point of departure for the deduction of theoretical mechanics.

This fundamental assumption may be formulated as the principle of the impossibility of a *perpetuum mobile.* When a system of particles acting under central forces passes from one configuration to another, the velocities acquired can be used to perform some work; in order to draw the same amount of work a second time from the system, one would have to restore its initial conditions by expending on it forces or energy from outside the system. The principle now requires that the amount of work gained by the transition from the first position to the second and the amount of work lost by the passage of the system from the second configuration to the first be equal, no matter in what way or at what velocity the change has been effected; otherwise a *perpetuum mobile* could be constructed on the basis of this cycle, contrary to the principle. So far Helmholtz's reasoning is but a paraphrase of the arguments used by Sadi Carnot and Benoît Clapeyron in their foundations of the thermodynamics of heat engines. By replacing the concept of work by that of "tensions" (*verbrauchte Spannkräfte*), which are equal but of opposite sign to the work performed, Helmholtz transformed the equation between living force (kinetic energy) and work into the statement that the sum of living force and tension is a constant, the tension being a function of the instantaneous state of the system. Although prima facie an insignificant change, this reformulation of the mechanical principle of the conservation of living force through the introduction of "tensions" opened up incalculable perspectives in that it could be applied to all branches of physics, not only to mechanics proper. Moreover, the new formulation was strikingly analogous to that of the principle of the conservation of matter, or mass, an accepted axiom in physical science since the times of Antoine Lavoisier. Exploiting the adaptability of the concept of "tension" to nonmechanical phenomena, Helmholtz not only reconciled the new doctrine of heat with the theory of mechanics, heat explicitly being treated as a form of energy, but also demonstrated the validity of the conservation principle for electrodynamics and other departments of physics. The recognition that mechanical work, heat, and electricity were only different forms of one and the same physical substratum—a result that can rightfully be considered the greatest physical discovery of the nineteenth century—found its analytical vindication in Helmholtz's paper.

At first, however, Helmholtz's memoir was hardly recognized, since its argumentation was based on mathematical reasoning, which at this time was accessible to but a small number of specialists. Another fundamental obstacle in the way of a just assessment of the new truth was the indiscriminate homonymous usage of the term *force* in both its Newtonian and its Leibnizian significations. Once the semantic difficulties had been removed, the principle of the conservation of energy found general acceptance and even popularity, owing to the writings of William Thomson (Lord Kelvin). In a discourse before the Royal Institution in 1856, Thomson distinguished carefully the significance of the Newtonian notion of force from what he called "energy." The term *energy*—apart from its early usage by Bernoulli and Young—had already been used three years earlier by William Rankine in his "On the General Law of the Transformation of Energy" (*Philosophical Magazine,* series 4, 5 [1853]: 106), but only Thomson's application led to its universal acceptance. "Any piece of matter or any group of bodies, however connected, which either is in motion, or can get into motion without external assistance, has what is called mechanical energy. The energy of motion may be called either 'dynamical energy' or 'actual energy.' The energy of a material system at rest in virtue of which it can get into motion, is called 'potential energy'" (*On the Origin and Transformation of Motive Power,* 1856). In 1893, in a footnote to a reprint of his 1856 lecture (in *Popular Lectures and Addresses,* London, 1894, Vol. II), Thomson wrote: "Shortly after the date of this lecture I gave the name 'kinetic energy' which is now in general use. It is substituted for 'actual' and for 'dynamical.'" Thus Helmholtz's "tension" was renamed "potential energy," and the sum total of kinetic and potential energies, the total energy of the system, was shown to be a constant that is characteristic of the system.

These innovations, however, had still to overcome some opposition. The Rankine-Thomson designation "potential energy" was rejected by John F. W. Herschel ("On the Origin of Force," in *Fortnightly Review and*

Familiar Lectures, 1857) as "unfortunate," being too common a name for such a "great truth." Even the term *conservation* of force or energy was subjected to severe criticisms, particularly by T. H. Huxley and by Herbert Spencer in his *First Principles* (1862), on the ground that "conservation" implies a conserver and an act of conserving and therefore the assumption that without such an act, force (energy) would disappear—an idea at variance with the conception to be conveyed. But in addition to the terminology, the conception itself, particularly that of potential energy, was still a matter of debate. An interesting testimony to these difficulties is Michael Faraday's paper "On the Conservation of Force" (*Philosophical Magazine,* series 4, 13 [1857]: 225–239), in which the following problem is raised: Is there creation or annihilation of force if the distance between two gravitating bodies is changed and the attractive force varies inversely with the square of the distance? "Gravitation," Faraday continued, "has not yet been connected by any degree of convertibility with the other forms of force… . That there should be a power of gravitation existing by itself having no relation to the other natural powers, and no respect to the law of the conservation of force, is as little likely as that there should be a principle of levity as well as of gravity." Rankine's answer to Faraday's objection (*Philosophical Magazine,* series 4, 17 [1859]: 250) seems to have had little effect, for as late as 1876, James Croll, in his paper "On the Transformation of Gravity" (*Philosophical Magazine,* series 5, 2 [1876]: 242–254), attempted to solve Faraday's query with the assumption that "a stone when in the act of falling [may] be acted upon by gravity with less force at any given moment than it would be were the stone at rest at that instant."

THE EMANCIPATION OF ENERGY. Although Croll's paper is full of misconceptions, which, interestingly, were clarified in an answer by the Viennese physiologist Ernst von Brücke, "On Gravitation and the Conservation of Force" (*Philosophical Magazine,* series 4, 15 [1858]: 81–90), it was of great importance for the subsequent development of the concept of energy. It connected the notion of energy for the first time with that of space. That space and change of position are necessary conditions for energy transformations Croll tried to demonstrate by the following consideration: four possibilities of energy transformations are conceivable—a change of potential energy into kinetic, of kinetic into potential, of kinetic into kinetic, and of potential into potential. Since, however, there "is evidently no such thing in nature, so far as is yet known, as one form of potential passing directly into another form" of potential energy and the existence of kinetic energy always implies change of position, the point is proved. Having thus associated energy with space, Croll went on to dissociate it from the material medium. "Our inability to conceive how force can exist without a material medium has its foundation in a metaphysical misconception," an idea he explained in greater detail in his book *Philosophy of Theism* (London, 1857). Croll's almost casual remarks, though scientifically rather objectionable and philosophically highly speculative, may be regarded as the earliest objection to the prevailing view, which still conceived of energy as an attribute, so to speak, of the dynamic system.

Meanwhile, James Clerk Maxwell's *Treatise on Electricity and Magnetism* (1873) appeared, opening the way for a field-theory treatment of electromagnetic phenomena. It showed, in particular, that the work necessary to build up an electromagnetic field can be regarded as equivalent to the energy produced in space with a certain density that depends on the squares of the magnitudes of the electric and magnetic fields. In the case of nonstatic fields these calculations lead to the conclusion, as was shown by J. H. Poynting in "On the Transfer of Energy in the Electromagnetic Field" (*Philosophical Transactions of the Royal Society* 175 [1885]: 343–361), that energy has to flow from one place in space to another in order to compensate for changes that occur in a particular region of space. A transfer of energy, it is true, had been associated with electricity before Poynting, but the energy flow was always considered as being confined to the conducting wires.

> But the existence of induced currents and of electromagnetic actions at a distance from a primary circuit from which they draw their energy, has led us, under the guidance of Faraday and Maxwell, to look upon the medium surrounding the conductor as playing a very important part in the development of the phenomena. If we believe in the continuity of the motion of energy, that is, if we believe that when it disappears at one point and reappears at another it must have passed through the intervening space, we are forced to conclude that the surrounding medium contains at least a part of the energy, and that it is capable of transferring it from point to point.

Thus the surrounding medium or empty space became the arena in which energy moves, and energy, disjoined from matter, was raised in its ontological status from a mere accident of a mechanical or physical system to the autonomous rank of independent existence: matter

ceased to be the indispensable vehicle for its transport. Mechanics, with its restricted conception of transfer of energy by matter, could proceed only as far as Gaspard de Coriolis's notion of "energy currents," described in his *Traité de la mécanique* (1844). The complete emancipation or reification of energy could be achieved only by a theory of action-at-a-distance, such as Maxwell's theory of electromagnetism. Here energy could be labeled and traced in its motion or change of form just as a piece of matter is ticketed so that it can be identified in other places under other conditions.

The recognition of the new ontological status of energy led to a result of great philosophical importance: It strengthened the position of those who opposed the prevailing kinetic-corpuscular theory of nature, according to which all processes are reduced to motions of particles and motion is the fundamental concept for physical explanation. Referring to the demonstrated equivalence of all forms of energy, the opponents claimed that kinetic energy is only one of the forms in which this quantity appears. In their view, energy was a much more general conception than motion, a conception that should not be narrowed down to mean only energy of attraction and repulsion of gravitational or electrostatic nature or energy of various forms of motion. One of the earliest exponents of this school of "energetics" was G. Helm, who, in a treatise, *Die Lehre von der Energie* (Leipzig, 1887), revived the term *energetics,* originally coined by Rankine, to characterize his position, according to which energy is the basic physical reality responsible for all natural phenomena. Helm referred to Gustav Zeuner, Ernst Mach, Josiah Gibbs, James Clerk Maxwell, A. J. von Oettingen, and Joseph Popper as advocating similar ideas. In particular, he claimed, energy can always be broken down into two factors, an intensity and an extensity factor, which characterize the quantity of energy as well as the direction in which changes of energy take place (the intensity factor always decreases).

In spite of further expositions, Helm's ideas did not attract much attention until Wilhelm Ostwald incorporated Helm's "factorization of energy" into the second edition of his treatise on physical chemistry, *Lehrbuch der allgemeinen Chemie* (1893), as the foundation of his theory of chemical affinity. In the period between the first and second editions of his treatise Ostwald embraced the new doctrine of energetics, and with his address in 1895 to the German Congress of Naturalists at Lübeck, "The Conquest of Scientific Materialism" (*Die Überwindung des wissenschaftlichen Materialismus*), he became the principal speaker of the new movement. In his view, not only was energy the universal currency of physics, but all phenomena of nature were merely manifestations of energy and of its manifold transformations.

In "Lectures on Natural Philosophy" (*Vorlesungen über Naturphilosophie,* Leipzig, 1901) Ostwald contended that since substance is by definition that which persists under transformations or changes, energy is substance. Methodological as well as epistemological considerations, he claimed, force us to see in energy the only substance—methodologically because the alternative view, scientific materialism, has failed to give an exhaustive explanation in even a single case of natural phenomena; epistemologically because "what we hear originates in work done on the ear drum and the middle ear by the vibrations of the air. What we see is only radiant energy which does chemical work on the retina that is perceived as light… . From this point of view the totality of nature appears as a series of spatially and temporally changing energies, of which we obtain knowledge in proportion as they impinge on the body, and especially upon the sense organs fashioned for the reception of the appropriate energies." Ostwald's conception of a physical object in terms of energy, of its volume in terms of compressibility, and of its shape in terms of elasticity is one of the final stages in a development that began with John Locke's sensationalistic conception and eventually put an end to the substantial conception of matter.

The "dissolution of matter" into energy was particularly welcomed by the adherents of the monistic school of thought in their search for a unified conception of the universe. Gustave Le Bon, for instance, in his *L'evolution de la matière* (Paris, 1905), spoke of the "dematerialization of matter into energy," a philosophical conclusion that in the same year found a far-reaching and profound scientific foundation. For in a paper titled "Does the Inertia of a Body Depend upon Its Energy Content?" ("Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?" in *Annalen der Physik* 18 [1905]: 639–641), Albert Einstein showed, on the basis of the Maxwell-Hertz equations of the electromagnetic field, that "if a body gives off the energy E in the form of radiation, its mass diminishes by E/c^2," where c denotes the velocity of light. Since then the mass-energy relation, $E = mc^2$, has been of fundamental importance, particularly in nuclear physics, where P. M. S. Blackett, G. P. S. Occhialini, O. Klemperer, and others showed that the total mass of a particle can be transformed into energy.

Whereas in classical mechanics differences of energy alone were of physical significance, so that energy could be determined only up to an additive constant, in mod-

ern physics energy lost this indeterminateness and became a physical quantity of absolute magnitude. Moreover, in the theory of relativity the principles of the conservation of energy, or mass, and momentum, the latter being the basis of the Cartesian measure of "force," revealed themselves only as different aspects of one and the same conservation law, the conservation of the momentum-energy four-vector. On the basis of the Einstein equation $E = mc^2$ the problem of the source of solar (or stellar) energy could be solved, the "packing effect" in nuclear physics could be explained, and the release of nuclear energy could be predicted. Energy was released mass, and mass was frozen energy, or as Bertrand Russell, in *Human Knowledge: Its Scopes and Limits* (New York, 1948), summarized the situation: "Mass is only a form of energy, and there is no reason why matter should not be dissolved into other forms of energy. It is energy, not matter, that is fundamental in physics."

CONSERVATION AND INVARIANCE. Although the theory of relativity threw new light on the conservational aspects of energy, or mass, the relationship between conservation and invariance found its final elucidation in Emmy Noether's article "Invariant Variational Problems" ("Invariante Variationsprobleme," in *Göttinger Nachricten* [1918], pp. 235–257), which demonstrates the conservation of certain quantities (for example, the canonical energy-momentum tensor) for dynamic systems that are invariant under continuous transformations of the coordinates or, more generally, of the field functions involved. Conservation thus appeared as a consequence of symmetry properties, a fact that was in part known already from the Hamiltonian formulation of classical mechanics. In particular, if homogeneity of space and time is assumed, that is, if it is postulated that the system is invariant under translational transformations of the origins of space-coordinates and time-coordinates, then the conservation of momenta and of energy is but a mathematical consequence. The principle of the conservation of energy of a given dynamic system is therefore ultimately a consequence of the invariance (or symmetry) of the system under changes in the zero-point of the time scale, that is, a consequence of the homogeneity of time.

See also Alembert, Jean Le Rond d'; Aristotle; Averroes; Bacon, Francis; Bilfinger, Georg Bernhard; Clarke, Samuel; Descartes, René; Einstein, Albert; Faraday, Michael; Force; Galileo Galilei; Gibbs, Josiah; Helmholtz, Hermann Ludwig von; Huxley, Thomas Henry; Lavoisier, Antoine; Leibniz, Gottfried Wilhelm; Mach, Ernst; Maxwell, James Clerk; Newton, Isaac; Ostwald, Wilhelm; Peter Aureol; Philosophy of Physics; Russell, Bertrand Arthur William; Schelling, Friedrich Wilhelm Joseph von; Thomas Aquinas, St.; Wolff, Christian.

Bibliography

Duhem, P. *L'évolution de la mécanique.* Paris, 1903.

Haas, A. E. "Die Begründung der Energetik durch Leibniz." *Annalen der Naturphilosophie* 7 (1908): 373–386.

Haas, A. E. *Die Entwicklungsgeschichte des Satzes von der Erhaltung der Kraft.* Vienna, 1909.

Helm, G. *Die Energetik nach ihrer geschichtlichen Entwicklung.* Leipzig, 1898.

Hiebert, Erwin N. *Historical Roots of the Principle of Conservation of Energy.* Madison: State Historical Society of Wisconsin for the Dept. of History, University of Wisconsin, 1962.

Jammer, M. "The Factorization of Energy." *British Journal for the Philosophy of Science* 14 (1963): 160–166.

Kuhn, T. S. "Energy Conservation as an Example of Simultaneous Discovery." In *Critical Problems in the History of Science,* edited by M. Clagett. Madison: University of Wisconsin Press, 1959.

Mach, E. *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit.* Prague, 1872.

Planck, M. *Das Prinzip der Erhaltung der Kraft.* Leipzig: Teubner, 1st ed., 1887; 2nd ed., 1908.

M. Jammer (1967)

ENERGY [ADDENDUM]

Force is among the most fundamental concepts in Newtonian physics. Energy became an important unifying concept in nineteenth-century physics. Energy and force take on somewhat different roles in relativity and quantum mechanics.

FORCE IN CLASSICAL PHYSICS

In classical physics, force is a vector quantity. Isaac Newton's second law of motion ($F = ma$) relates the net force (F) on a body to its mass (m) and acceleration (a) in an inertial reference frame. Newton's third law says that the force exerted by body A on body B is equal and opposite to the force that B exerts on A. To apply Newton's laws of motion in a non-inertial frame, correction factors with the dimensions of force ("pseudoforces") must be introduced, such as the Coriolis and centrifugal forces.

The constituents of a system of bodies (such as a macroscopic object) exert "internal forces" upon one another, whereas "external forces" are imposed on the system from without. By Newton's third law, the internal

forces cancel. Newton's second law then applies to the system as a whole: The net external force on the system equals the product of the system's total mass (the sum of its constituents' masses) with the acceleration of its center of mass. The system of bodies therefore itself constitutes a body in classical physics.

If the net force on a body is *defined* as (or is nothing over and above) the product of its acceleration and mass, then Newton's second "law" is true *trivially*. One way to avoid this result is to take "force" as defined partially by the various force laws (gravitational, electric, magnetic, etc.). Another way is to take forces as real entities existing alongside masses and accelerations and serving as simultaneous causes of accelerations. Philosophers who believe in the reality of forces have disagreed about whether component forces or only net forces are real. If component forces are real, then a zero net force may have real nonzero components. These components apparently cannot be understood as real parts of a nonexistent whole. Perhaps the components along arbitrary axes are unreal, whereas the components given by the various force laws are real. According to Nancy Cartwright, net forces are real but components are not. Hence, there is no component gravitational force between two bodies for the gravitational force law to relate to the bodies' masses and separation. The law must concern merely those (unreal) situations where a body feels only a single influence—from one other body's mass; the law then covers the net forces in those cases.

FORCE IN MODERN PHYSICS

Henri Poincare and Hans Reichenbach distinguished differential forces (such as electric and magnetic forces) from universal forces, which cannot be shielded against and would have the same effect on any body (whatever its charge, mass, chemical constitution …) in a given spatiotemporal location. A pseudoforce is a universal force, since it reflects the reference frame's acceleration, not the character of the body affected. Phenomena cannot reveal any universal forces acting on a given body, since measuring devices would be affected in the same way. Reichenbach argued that phenomena determine the geometry of spacetime only up to a conventional choice of universal forces.

According to Albert Einstein's general theory of relativity, gravity's "effects" on a body's trajectory result not from an external force, but from spacetime's geometry. To Einstein, this was suggested by the equality of inertial and gravitational masses, rendering a body's acceleration due to gravity independent of its own mass. That is, gravity functions as a universal force, and by adopting a non-Euclidean spacetime geometry, general relativity geometrises gravity away. Einstein's "principle of equivalence" says that in any sufficiently small spatiotemporal region, there is a reference frame in which no phenomena are attributed to gravity, and so the laws do not refer to gravity. (In the canonical example, the phenomena in an elevator falling freely in a gravitational field are explained, without appealing to gravity, in a reference frame falling with the elevator. In that frame, all phenomena are indistinguishable from those experienced in an inertial frame.) This principle thus treats gravity as a pseudoforce.

In quantum physics, a body's state can be affected even as the body passes exclusively through regions where it feels no force. This occurs in the Aharanov-Bohm effect, for example.

FIELDS OF FORCE

A force at some spatiotemporal location may be interpreted as having, apart from the affected body's charge, no causes nearer in space (time) than an appropriately charged body's being some distance away (sometime earlier). This would be (retarded) action at a distance. Although Newton regarded action at a distance as impossible, he failed to offer any local causal mechanism for gravity. Accordingly, gravity was later often interpreted as action at a distance.

Alternatively, a force may be interpreted as having an entirely *local* cause: the affected body's charge and the corresponding field at the body's spatiotemporal location. A field is a vector quantity equal at a given location to the force per unit charge that would be felt by a charged point body, were one present there. In the nineteenth century, Michael Faraday, James Clerk Maxwell, Oliver Heaviside, and Heinrich Hertz developed an electromagnetic field theory. Maxwell argued that the electric and magnetic fields are real, rather than mere mathematical devices (like the electric potential), on the grounds that fields possess energy (and momentum). According to Faraday, we can account for matter's impenetrability without positing that matter consists of hard "stuff"—namely, by positing that material objects are surrounded by short-range fields of repulsive force. Accordingly, we have no reason to believe in a solid body lying somewhere deep beneath its surrounding atmosphere of fields; Faraday speculated that matter is nothing but inertia-bearing centers of fields. Max Abraham, Wilhelm Wien, and Gustav Mie were among those who later tried to develop an "electromagnetic theory of matter," according to which

bodies are just local concentrations of the electromagnetic field.

ENERGY IN CLASSICAL PHYSICS

A system's energy is a scalar quantity reflecting the amount of "work" (force accumulated over distance) needed to assemble the system. Another way of putting the point is that a system's energy is its capacity to do work. The energy of a closed system is a conserved quantity.

Energy comes in two forms: kinetic and potential. A body moving with speed v has kinetic energy (equal to $\frac{1}{2} mv^2$). A system's potential energy reflects its configuration. For example, a pair of like electric charges (which mutually repel) has greater electric potential energy insofar as the two charges are nearer (and so required more work to be squeezed together to their current separation). A system's total energy is the arithmetic sum of its various kinetic and potential energies.

Energy conservation was the great unifying principle of nineteenth-century physics. All forms of energy—whether chemical (as stored in a battery), thermal, molecular (as in a chemical bond), elastic (as stored in a spring), kinetic, electric, magnetic, or gravitational—could be interconverted according to fixed rates of exchange, providing a common currency for nature's economy. A system's energy could be calculated—and some of the system's behavior thus predicted—even if a detailed mechanical model of the system was unavailable, either because the system was too complicated (e.g., a large collection of molecules) or because its presumptive inner workings were unknown (as in the case of the aether, the space-pervading medium purportedly responsible for long-distance electromagnetic interactions).

Potential energy is ascribed to a system *as a whole* (e.g., to the *pair* of charged bodies) but is assigned no definite distribution among the system's constituents. Prior to Maxwell's electromagnetic field theory, physical law also seemed silent on the *absolute* value of the system's energy; energy *differences* alone matter to energy's conservation. That is, the system's potential energy reflects the work needed to assemble it, but since we may take any configuration as constituting the initial "raw material" out of which the system is assembled, it is arbitrary which configuration is assigned zero potential energy. The arbitrariness of energy's spatial distribution and absolute value suggested that energy is not a real substance that is spread around space and "neither created nor destroyed." It suggests, rather, that energy is just an arithmetic combination of various physical quantities (e.g., charge, velocity) that is useful for predicting a system's behavior by virtue of its maintaining a constant value.

All this was greatly affected by Maxwell's electromagnetic field theory. The retarded character of electromagnetic action results in violations of Newton's third law in nonstatic cases and, therefore, in energy nonconservation—unless there are some additional energy terms beyond those used in calculating the system's total energy in a static case. Maxwell's theory supplied corrected terms for the electric and magnetic potential energies, restoring energy conservation. These corrected terms are most naturally interpreted as ascribing energy to the electric and magnetic fields—that is, to the apparently empty space surrounding charged bodies. The field energy density at a point is proportional to the square of the field's strength there. The field's zero level, as determined by the condition in which zero force per unit charge would be felt by a test body, designates a non-arbitrary condition of zero energy. Thus, by ascribing energy to the field, we find electromagnetic potential energy to have an absolute value and a determinate distribution in space. This result suggests that energy is a real substance and that the electric and magnetic fields, by virtue of containing energy, are real.

By assuming that parcels of energy have continuing identities as they move and, moreover, that energy obeys a "continuity equation" (in that parcels of energy must travel through space along continuous paths), John Henry Poynting found an expression for the energy flux density (the rate and direction of energy flow at each location in the electric and magnetic fields). This solution, the "Poynting vector," gives results that many (such as Hertz, James Jeans, and J. J. Thomson) found counterintuitive in certain cases (as when it entails that a tremendous flow of energy circulates around a stationary magnet near an unmoving charged body). Moreover, an individual parcel of energy cannot be marked in order to follow its trajectory. Furthermore, the Poynting vector is not the unique solution to the continuity equation for energy flow. All of the solutions have counterintuitive consequences like the above. They agree on the net energy flow across a closed surface. However, they disagree on the energy flow across an open surface (i.e., a surface that fails to completely enclose a volume). These results suggested to some (e.g., Heaviside and Hertz) that energy is not a substance. Perhaps it is a property (like velocity). In that case, its possession by fields might entail their reality. Alternatively, perhaps energy is merely a theoretical fiction (as Jeans believed). In that case, energy would lack

sufficient ontological status to underwrite the reality of fields. This issue is not resolved within classical physics.

ENERGY IN MODERN PHYSICS

In the special theory of relativity, energy is conserved but is not Lorentz invariant. Hence, a system's energy reflects not only its real state, but also the inertial reference frame in which we have chosen to describe it. Energy and momentum are the components of a 4-vector whose length is a Lorentz invariant quantity: the (rest) mass. That is, a system's mass appears in different inertial frames as different combinations of energy and momentum.

It is sometimes said that energy and mass are interconvertible according to Einstein's famous equation $E=mc^2$, as when a ball of gas is heated and the added thermal energy becomes additional mass. Such talk of energy's being "converted" into mass (or matter) is highly misleading, since energy and mass are radically different quantities: mass is real whereas energy is not. Mass (or matter) thus cannot be or be transformed into energy; bodies are not made of energy. Physical transformations of a closed system cannot result in the appearance or disappearance of some mass, since mass is a conserved quantity.

When we consider the ball of gas as a collection of bodies, we characterize the added heat as having boosted various molecules' kinetic energies, though not their masses. When we instead consider the gas as a single body, the kinetic energy contributed by the heat counts toward the gas's mass (which is not the sum of the masses of the gas molecules). This "conversion" of energy into mass is not a real physical transformation. It is an artifact of *our* shift from treating the gas as many bodies to treating it as a single body. As in classical physics, a system of bodies is itself a body.

In quantum mechanics, the value of a system's energy is more definite insofar as the moment at which it possesses that energy is less definite. This is a form of Werner Heisenberg's "uncertainty principle." Though the proper interpretation of this principle is controversial, it is generally held responsible for the brief departures from energy conservation required for the existence of virtual particles and states. A charged point particle's infinite self-energy also remains a source of controversy in quantum mechanics, though renormalization techniques allow it to be finessed in calculations.

See also Classical Mechanics, Philosophy of; Conservation Principle; Force.

Bibliography

Berkson, William. *Fields of Force*. New York: Wiley, 1974.
Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford: Clarendon, 1983.
Einstein, Albert, and Leopold Infeld. *The Evolution of Physics*. New York: Simon & Schuster, 1951.
Faraday, Michael. "A Speculation Touching Electric Conduction and the Nature of Matter." In *Experimental Researches in Electricity*, Vol. 2, 284–293. London: Bernard Quaritch, 1844.
Hesse, Mary. *Forces and Fields*. Totowa: Littlefield, 1965.
Imry, Yoseph, and Richard Webb. "Quantum Interference and the Aharonov-Bohm Effect." *Scientific American* 260 (4) (1989): 56–62.
Jammer, Max. *Concepts of Force*. Cambridge, MA: Harvard University Press, 1957.
Jeans, James. *The Mysterious Universe*. London: Macmillan, 1932.
Lange, Marc. *An Introduction to the Philosophy of Physics*. Malden: Blackwell, 2002.
Lodge, Oliver. "On the Identity of Energy." *Philosophical Magazine* 19 (series 5): (1885): 482–487.
Maxwell, James Clerk. "Action at a Distance." In *The Scientific Papers of James Clerk Maxwell*, Vol. 2., edited by W. D. Niven. 311–323. Cambridge, U.K.: Cambridge University Press, 1890.
Maxwell, James Clerk. *A Treatise on Electricity and Magnetism*. New York: Dover, 1954.
Nahin, Paul. *Oliver Heaviside*. Baltimore, MD: Johns Hopkins Press, 2002.
Poynting, John Henry. "On the Transfer of Energy in the Electromagnetic Field." In *Scientific Papers*, 175–193. Cambridge, MA: Cambridge University Press, 1920.
Reichenbach, Hans. *The Philosophy of Space and Time*. Edited and translated by Maria Reichenbach and John Freund. New York: Dover, 1958.

Marc Lange (2005)

ENGELS, FRIEDRICH
(1820–1895)

Friedrich Engels, the intellectual companion of Karl Marx, although generally considered inferior to his colleague as a thinker, contributed more than Marx to the development of the philosophical aspects of Marxism. Indeed he was the creator of orthodox Marxism as a system based on historical materialism and on dialectics. Engels was born in Barmen in the German Rhineland. His father was a textile manufacturer who had interests in England, and Engels went there to work in a cotton mill in Manchester, first as clerk, later as manager and part owner. Engels was a man of many talents, a scholar, linguist, pamphleteer, soldier, military commentator, and businessman. He was all those things with a thoroughness and distinction that would have brought him recognition

in his own right, but it was his intellectual partnership with a man of genius that brought him fame. Engels met Marx briefly in Cologne in 1842, became acquainted with him in Paris in 1844, and worked actively with him before and during the revolutionary ferment of 1848, when they wrote the *Communist Manifesto.* In 1850 Engels reluctantly returned to his business in Manchester, in part because he saw that Marx needed financial support in order to continue his researches. This help Engels gave unstintingly throughout Marx's life and for years after his death, to his surviving children. Outliving Marx by twelve years, Engels edited his friend's manuscripts, notably the two volumes of *Das Kapital* left unfinished by Marx. He also served as official interpreter of Marxist doctrine during the years when it was beginning to attain worldwide influence over workingmen's movements.

Beginning with works written during Marx's lifetime and with Marx's express approval—for example, *Anti-Dühring* (1878)—Engels emphasized the scientific, positivist component in their joint theories, which he compared with those of Charles Darwin. Engels believed that he and Marx had discovered a rigid system of historical laws that would lead with inexorable necessity to socialism. These laws, Engels held, were dialectic rather than mechanical in character. That is, instead of being like the laws previously discovered in natural science and extrapolated to social studies by men whom Engels called vulgar materialists, they were laws that took account of the contradictions in reality and of the fact that development occurred in revolutionary leaps to higher levels. Engels took from G. W. F. Hegel the doctrine, which he called the law of the interpenetration of opposites, that objective contradictions exist in reality. He enunciated the law of the transformation of quantity into quality, which asserts that change occurs abruptly, after a period of gradual progression. The last dialectical law, the negation of the negation, states that progress takes place by a series of detours, from position *A* to the opposite, position—*A,* and then back to the opposite of that position, which turns out to be position *A* "raised to a higher power." To give one of Marx's own examples, the industrial bourgeoisie generates its opposite, the miserable proletariat, which then negates bourgeois capital in a revolutionary leap to the higher stage of classless industrial society.

Engels adumbrated these theories in *Anti-Dühring* and stressed them in a special excerpt from that work, *Socialism: Utopian and Scientific* (1892), but the extent to which he carried them was not known until his *Dialectics of Nature* was published in 1925. In this work he extended materialist dialectics to the natural sciences, with results that are often held to be ludicrous, and implied that dialectics would supersede formal logic. The lengthy controversies that these questions provoked in Soviet philosophy arose, then, from the work of Engels rather than of Marx.

While it is certain that Engels stressed such questions more than Marx and that he lived on to formalize a Marxist tradition out of reverence for a friend who disliked just such formalism, one must be wary of attempts to set Engels, as a scientistic pedant, against Marx, as an existentialist or idealist. It is tempting for certain neo-Marxist philosophers, but in the end impossible, to purge Marxism of all its allegedly scientific content that has since been proven untrue and to lay all these errors at Engels's door, leaving only the "profound" (or ambiguous) speculations of the young Marx as true Marxism. For one thing, it was Engels who suggested those early speculations to Marx, in 1844. And decades later it was not Engels alone but the age and his own ambitions that led Marx to present his mature theory of history as a "scientific system" (decorated with some Hegelian flourishes). At all events, it was Marx's thought as understood by Engels that came to constitute Marxism and, in particular, Soviet dogma.

See also Communism; Dialectical Materialism; Hegel, Georg Wilhelm Friedrich; Marxist Philosophy; Marx, Karl.

Bibliography

PRIMARY WORKS

Die Entwicklung des Sozialismus von der Utopie zur Wissenschaft (Socialism: utopian and scientific). Translated by E. Aveling. New York: Scribner, 1892.

Der Ursprung der Familie, des Privateigentums und des Staats (The origin of the family, private property and the state). Translated by E. Untermann. Chicago: Kerr, 1902.

Die Briefe von Friedrich Engels und Eduard Bernstein. Berlin, 1925.

Gesamtausgabe. 12 vols. Berlin and Moscow, 1927–1935.

Herr Eugen Dührings Umwälzung der Wissenschaft (Herr Eugen Dühring's revolution in science). Translated by E. Burns. London: M. Lawrence, 1934.

Ludwig Feuerbach und der Ausgang der klassischen deutschen Philosophie (Ludwig Feuerbach and the outcome of classical German philosophy). New York: International, 1934.

Dialektik der Natur (Dialectics of nature). Translated by Clemens Dutt. New York: International, 1940.

Selected Works. 2 vols. London, 1942.

Grundsätze des Kommunismus (Principles of communism). Translated by Paul Sweezy. New York: Monthly Review, 1964.

Sochineniya, 32 vols. Moscow, 1955–.

Friedrich Engels–Paul et Laura Lafargue, Correspondance, 1868–95, 3 vols. Edited by E. Bottigelli. Paris: Editions Sociales, 1956–1959.

Selected Letters: The Personal Correspondence, 1844–1877. Edited by Fritz J. Raddatz. Translated by Ewald Osers. Boston: Little Brown, 1981.

SECONDARY WORKS

Arthur, Christopher J., ed. *Engels Today: A Centenary Appreciation.* Houndmills, U.K.: Macmillan, 1996.

Carver, Terrell. *Friedrich Engels: His Life and Thought.* New York: St. Martin's Press, 1990.

Dixon, Richard et al., trans. *Karl Marx, Frederick Engels: Collected Works.* 49 vols. New York: International, 1975–.

Hook, Sidney. *Reason, Social Myths, and Democracy.* New York: Humanities, 1950.

Lichtheim, George. *Marxism.* London: Routledge and Kegan Paul, 1964.

Meyer, Gustav. *Friedrich Engels.* 2 vols. The Hague, 1934.

Steger, Manfred B., and Terrell Carver, eds. *Engels after Marx.* University Park: Pennsylvania State University Press, 1999.

Neil McInnes (1967)
Bibliography updated by Philip Reed

ENGINEERING ETHICS

"[F]or all of its influence on our modern world, the engineering profession remains a mystery to many Americans." (ASEE *Action*). These words in President Bill Clinton's statement for Engineer's Week of 1999 capture the curious situation of engineering: its products shape the world, but engineers are virtually invisible.

The academic study of ethics and responsibility in engineering began in the United States in the mid-1970s at a time of social ferment and heightened public scrutiny of the professions. Scholars from philosophy and engineering, collaborating in workshops and conferences, teaching, and research, began to penetrate the mystery. They concentrated on engineering in the United States.

Engineering originated in France in the seventeenth century and led in France to the development of the first engineering curriculum during the eighteenth century. Subsequently, engineering took shape as an occupation elsewhere, notably in the United States, Britain, Germany, and Russia. The French curriculum, with its emphasis on mathematics, physics, and chemistry, became the model for the engineering curriculum in the United States and most other countries. Despite persisting differences among countries in the status of professions and of engineers, the academic study of engineering ethics spread to a number of other countries.

Engineering ethics critically examines the behavior of engineers and engineering institutions in light of the special standards of the profession and the common standards of morality. The discipline studies engineers' actions, practices, and workplace, focusing philosophical analysis on standards and concepts such as responsibility and loyalty, to help identify ethical problems and options for resolving them.

Cases or vignettes are essential starting points for teaching and research. For example, during an economic downturn, an engineer overseeing the testing of fuel pumps for a company receives instructions to curtail the testing process. The engineer's ethical concern is that he or she will not be able to ensure the life expectancy of the pumps relied on by the company's customers.

From the latter part of the nineteenth century, engineering in the United States organized as a profession, creating engineering professional societies and promulgating technical and ethical standards. The latter incorporate ordinary morality, for example, in requiring engineers to "issue public statements only in an objective and truthful manner." They include special standards, for instance, the canon requiring engineers to act "for each employer or client as faithful agents … and avoid conflicts of interest." (Accreditation Board for Engineering and Technology [ABET] 1977, p. 1).

In the ferment of the mid-1970s engineering societies revised their codes of ethics. Unsatisfied with a commitment to "due regard" or "proper regard" for the public, almost all the societies adopted as the first canon, "Engineers shall hold paramount the safety, health, and welfare of the public in the performance of their professional duties" (Florman 1986, p. 77–78).

A great majority of engineers are employees of large business organizations where they do not easily acquire authority or visibility. Still, through their professional societies, engineers profess a commitment to serve society and continue to promulgate technical and ethical standards supporting that commitment. Ethical standards articulate values underlying the technical standards, the core values—safety, reliability, and efficiency—that are also embedded in routines of engineering practice.

The engineering workplace features complexities and intricacies of large, generally hierarchical organizations. The role of engineers in business and in other organizations is elastic. They manage a range of responsibilities from narrowly technical to managerial and, while

often subordinate to managers, must cooperate with them in decision making. The major ethical challenge for engineers is to deal with these complexities (including cost constraints) as employees bound by the special ethical standards of a profession as well as by moral rules. For example, they must find ethically justified, practical options for coping with instructions to curtail testing or to drastically revise public statements.

Adding to the complexity of the engineering workplace is the legal environment, including contracts, the federal and state regulatory systems for health and safety, product liability litigation, and common-law adjudication involving expert witnesses. The legal framework both constrains engineers and generates questions about additional ethical responsibilities, for example, about the extent of their responsibilities to help formulate or implement government standards to control pollution.

Individual engineers' ethical obligations derive from requirements of morality, the obligation of everyone to exercise a reasonable standard of care, the special standards of the profession, and the duties they have as employees. All these ethical imperatives inform the exercise of practical judgment by engineers, the professionals who determine specifications for the design, development, testing, operation, maintenance, and disposal of technological products and systems.

Regarding concerns about safety, for instance, they have a duty to protect the public while avoiding injury to their employers. Engineers are thus subject to tension between the duty of loyalty to the employer (complicated by having to distinguish between interests of the company and what managers want) and the obligation to hold public safety paramount. An engineer's judgment that his or her company's environmentally damaging spill should be reported to the regulatory agency might encounter resistance challenging his or her loyalty. In handling the reporting obligation, the engineer must take due care to avoid injury to the company and to a manager perhaps more concerned with self-protection than other interests.

The moral status of loyalty and the idea of critical loyalty are central in research and teaching. Discussion focuses on a range of ways to express independent judgment, from disagreement and dissent to the extreme of whistle-blowing. Dissent, such as resisting assignment to a particular project out of safety concerns, may invoke the code of ethics as support. Disagreement and dissent require tact and sensitivity so as not to cause avoidable opposition or injury.

Whistle-blowing, that is, transmitting information outside normal channels, ruptures relationships and requires justification that trumps the harm it causes. Engineers blocked from obtaining images to assess the impact of foam debris on the space shuttle *Columbia* had justification for blowing the whistle. To help engineers perform responsibly without resorting to extreme measures, research and teaching focus on impediments to responsible conduct in organizations, for example, fear, deference to authority, and "group think."

The space shuttle *Challenger* disaster revealed another impediment: normalized deviance (Vaughan 1996). It is a form of complacency, the phenomenon of gradually accepting certain anomalous, originally unexpected occurrences that portend serious harm. As the occurrences continue without leading to actual serious harm, they come to be viewed as normal. Strategies to counter this relaxation of vigilance and other impediments to responsibility are current research subjects. This is preventive ethics, catching engineering ethics problems early before they ripen into disasters.

Canon one, the code provision that enjoins engineers to "hold paramount the safety, health and welfare of the public," (ABET 1977, p. 1) requires interpretation. Analysis begins with the question: Who is the public? Should the public include, for example, the crew on the *Columbia*, workers within the engineering workplace, or everyone who might be affected by an engineering product?

Michael Davis (1998) points out the need to determine a characteristic that identifies the relevant public, that is, the vulnerable parties who may be harmed by engineers' work. He suggests identifying members of the public by their ignorance and consequent helplessness in the face of dangers from engineers' work. On this interpretation, members of the *Columbia* crew, unaware of the extent of damage from the break off of insulation and therefore helpless to do anything about their perilous situation, were members of the public.

Analysis continues by asking: How can engineers translate the *paramountcy* provision into guidelines that are less vague? Kenneth Alpern (1983) draws attention to the importance of a standard or principle of due care that holds for everyone. Its corollary, a standard of care proportionate to the magnitude of harm and "the centrality of one's role" in producing the harm, further reduces the vagueness for engineers.

Mindful that this principle can demand moral heroism and that few people are capable of heroism, engineering ethics specialists focus on sources of support for

engineers and on constructing options for responsible problem solving within the capacities of most people. In constructing options for resolving ethics problems, engineers use methods resembling those for solving design problems (Whitbeck 1996).

Further analysis of the paramountcy provision addresses another problem: managers typically balance or trade off factors, such as cost, schedule, marketing, and safety. In their deliberations, managers include safety as a factor, but only as one factor that, like others, may have to be sacrificed. Because safety is a priority for engineers, they cannot treat safety in that way. Philosophers suggest interpreting canon one as requiring engineers to meet a threshold of safety before taking a balancing approach (Harris, Pritchard, and Rabins 2005). This interpretation can help engineers hold their ground with managers.

Employers' demands for secrecy and confidentiality give rise to a cluster of specific issues concerning disclosure and withholding of information and protection of intellectual property, including trade secrets and patents. Societal interests in open circulation of knowledge (and propagation of new technology) and engineers' interests in using their knowledge to advance their careers come into conflict with the interests of firms in protecting information perceived to be economically valuable.

Employment contracts generate ethical responsibilities for engineers and their employers and figure in the balancing necessary to reconcile these interests. These contracts commonly require engineers to keep information confidential even after moving to another company. Such contracts make employers responsible for clearly specifying information to be kept confidential over a reasonable period of time. Engineers become responsible for taking due care at a new job to protect specified information for an appropriately limited time. Drawing such lines between privately owned knowledge and public knowledge is an important practical issue for engineers as well as a subject for analysis in engineering ethics.

Among problems that readily arise for engineers is conflict of interest (COI). While specifying vendors, suppliers, contractors, materials, and components, engineers must be alert to affiliations, investments, and associations they have that can threaten the reliability of their engineering judgment. Philosophical investigation has explicated the concept of COI, the harm of COI, and appropriate responses for dealing with COI. Disclosure of the investment or affiliation that threatens reliable judgment is essential to avoid deceiving and betraying the party relying on professional judgment.

Because of the impact of engineers' work and the priority of safety, it is essential for engineers to acquire a sophisticated understanding of risk and approaches to dealing with risks to humans, other creatures, and the environment. One approach to fostering such understanding is to provide engineers an overview of important perspectives on risk and critical discussion of cost-benefit analysis.

Ethics specialists consider several perspectives alongside one another, those of risk experts (specialists in defining and assessing risks, usually relying on cost-benefit analysis), government regulators, and lay people. It is part of the engineering approach to provide knowledge about risks, for example, concentrations of pollutants in water. The engineering perspective also includes an understanding of cost-benefit analysis and its limitations, an orientation toward protecting the public (similar to, but not the same as that of the government regulator), and an appreciation of lay attitudes toward risks (e.g., those imposed as contrasted with those voluntarily assumed).

Accommodating lay attitudes introduces issues associated with informed consent, that is, explicit acceptance of risks by affected parties. Recognizing that many situations in engineering make it impractical to obtain voluntary informed consent, ethics analysis considers substitutes and compensatory policies. This and other engineering ethics topics encompass problems that arise for individual practitioners but point toward engineering responsibilities of the profession as a whole because they call on the collective capabilities of the profession.

Accordingly, engineers' responsibilities regarding the environment have begun to engage U.S. engineering societies as well as ethics specialists. Some societies have added provisions to their codes of ethics that provide a distinct place in decision making for attention to environmental implications.

Engineers work increasingly in an international environment. A decision-making situation may bring into play engineering standards and government regulatory standards of different countries. The tasks of finding common ground and making adjustments among differing standards consistent with morality and the paramountcy provision are appropriate responsibilities for the profession through its professional societies. For engineering ethics research, addressing international variations in standards is an important task. Advances in international law, which have been prompted by economic globalization, may encourage such research.

As radically innovative technologies have followed rapidly one after another, especially in the decades since World War II, issues associated with emerging technologies have come to the forefront. For individuals and the profession as whole, emerging technologies present issues not only regarding potential risks but also regarding the role of engineers (and the technologies they help create) in shaping the physical, social, and cultural world.

See also Duty; Ethics and Economics.

Bibliography

Accreditation Board for Engineering and Technology (ABET). *Code of Ethics of Engineers*. 1977. Available from http://ethics.iit.edu/codes/coe/abet-a.html.

Alpern, Kenneth. "Moral Responsibility for Engineers." *Business and Professional Ethics Journal* 2 (2) (1983): 39–48.

ASEE Action. "National Engineers Week" *American Society for Engineering Education Newsletter* (February 1999): 2.

Baron, Marcia. *The Moral Status of Loyalty*. Chicago: Center for the Study of Ethics in the Professions (IIT), 1984.

Davis, Michael. *Thinking Like an Engineer*. New York: Oxford University Press, 1998.

Florman, Samuel C. "Moral Blueprints." In *Engineering Professionalism and Ethics*, edited by James H. Schaub, Karl Pavlovic, and M. D. Morris, 76–81. Malabar, FL: Robert E Krieger, 1986. Originally published in the magazine *Harper's* (October 1978).

Harris, Charles E. "Explaining Disasters: The Case for Preventive Ethics." *IEEE Technology and Society Magazine* 14 (2) (Summer 1995): 22–27.

Harris, Charles E., Michael S. Pritchard, and Michael J. Rabins. *Engineering Ethics: Concepts and Cases*. 3rd ed. Belmont, CA: Thomson/Wadsworth, 2005.

Johnson, Deborah, ed. *Ethical Issues in Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

Ladd, John. "Collective and Individual Moral Responsibility in Engineering: Some Questions." *IEEE Technology and Society Magazine* 1 (2) (June 1982): 3–10.

Vaughan, Diane. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago: University of Chicago Press, 1996.

Weil, Vivian. "Engineering Ethics." In *Science and Technology Ethics*, edited by Raymond E. Spier, 59–88. London: Routledge, 2002.

Whitbeck, Caroline. "Ethics as Design: Doing Justice to Moral Problems." *Hastings Center Report* 26 (3) (May–June 1996): 9–16.

Vivian Weil (2005)

ENLIGHTENMENT

The term *enlightenment* is generally used to designate a period in European history stretching from the 1680s to the close of the eighteenth century, but this usage is not without ambiguities and controversy. During the eighteenth century the word *enlightenment* referred not to a period but to a process, a set of activities in which individuals engaged. These activities were viewed as involving the application of what was then termed *philosophy* to a range of concerns in what would subsequently be classified as the natural sciences, the humanities, and the social sciences. It was not until the nineteenth century that *the Enlightenment* came into general usage as a designation for the historical period defined by these various projects. Attempts to specify the character of the period have tended to spur reflection on the nature and scope of those projects and activities that are claimed to characterize the age. As a result, discussions of the Enlightenment typically slide into reflections on the nature and merits of the activity of enlightenment itself.

THE HISTORY OF THE CONCEPT

At the close of his 1784 essay in the *Berlinische Monatsschrift* in response to the question "What is enlightenment?" Immanuel Kant asked whether his might be characterized as an "enlightened age." He responded, "No, but it is an age of enlightenment" (p. 35). Kant's emphasis on enlightenment as an ongoing process, rather than as an achieved state, was typical of eighteenth-century usage, which favored such formulations as "century of philosophy" (Jean Le Rond d'Alembert), "age of critique" (Kant), or "age of reason" (Thomas Paine).

The question of what the process of enlightenment involved sparked an extended discussion in German journals during the 1780s, a discussion in which Kant's response would prove to be the most famous. The German *aufklären*—a word that had been used to designate a clearing of the weather and, metaphorically, a return to consciousness after a period of sleep—had been employed since the beginning of the eighteenth century as a translation for the French *eclairer* (an important term in the works of René Descartes and Gottfried Wilhelm Leibniz) and for the English *enlighten*. More generally, the use of light as a metaphor for knowledge had a long history in Western philosophy as well as a central place in religious discourse. Hence, the particular use to which these metaphors were put during the eighteenth century by those thinkers now associated with the Enlightenment had a polemical edge: True enlightenment, it was argued, resulted from the application of reason and philosophy, rather than appeals to revelation or to the mysteries of faith. Critics could, in turn, marshal the same metaphors and argue that what was proposed as enlightenment was instead a form of spiritual darkness.

The application of the term to a particular historical period was greatly influenced by Georg Wilhelm Friedrich Hegel's lectures on the history of philosophy and the philosophy of history from the 1820s, and his usage was widely imitated in German histories of philosophy and of literature. The French term for the period—*siècle des lumières*—suggested a more elastic understanding of the period: a century of "lights" rather than a single movement. English usage followed the German, but lagged behind it, with *the Enlightenment* replacing *the Illumination* as a label for the period only in the waning years of the nineteenth century. As late as 1910 the Princeton philosopher John Grier Hibben, in the first book in English to use the term consistently, treated the term as a neologism in need of explanation. Indeed, for much of the twentieth century *age of reason* remained a widely used alternative.

The seminal historical studies of the period date from the 1930s: Ernst Cassirer's *Die Philosophie der Aufklärung* (1932/1951), Paul Hazard's *La crise de la conscience europeén* (1935/1953); and Carl L. Becker's *The Heavenly City of the Eighteenth Century Philosophers* (1932), a work whose fame rests more on the novelty of its argument than on the quality of its scholarship. Peter Gay's *The Enlightenment: An Interpretation* (1966–1969) remains the most influential of the many subsequent studies. Some scholars criticize the tendency to exaggerate the unity of the Enlightenment and emphasize the diversity of enlightenments, sometimes distinguished by their "national context" (see Porter and Teich 1981), for example, the "Scottish Enlightenment," the "Berlin Enlightenment," and the "British Enlightenment." Still others (e.g., Israel 2001) maintain that a focus on national contexts ignores the cosmopolitan character of the Enlightenment, particularly in its more radical manifestations. Since the 1970s there has been a tendency for historical discussions of the Enlightenment to turn from the focus on prominent thinkers and their works that had been the defining feature of earlier studies in favor of approaches influenced by developments in social history and histories of publishing and reading. The work of the historian Robert Darnton (1995) has been particularly influential in this regard.

THE ROLE OF PHILOSOPHY IN THE ENLIGHTENMENT

In the earliest discussions the relationship between philosophy and the Enlightenment was pervasive and unproblematic: The Enlightenment was typically defined in terms of the philosophers who were said to have articulated its ideals. Some of the early controversial literature spurred by the French Revolution traced the origins of the Revolution to the writings of François-Marie Arouet de Voltaire, Denis Diderot, and the other *philosophes*, and terms such as *philosophism* and *Illumination* figured prominently in the writings of British opponents of the Revolution and in accounts (notably Augustin Barruel's [1743–1820] *Memoirs Illustrating the History of Jacobinism* [1798]) that traced the Revolution to a conspiracy of *philosophes* and Freemasons.

A more sober analysis could be found in Hegel's *Lectures on the History of Philosophy* from the 1820s, which tended to reserve the term *Aufklärung* for the German phase of the broader movement of modern philosophy that began with Descartes. In other lecture cycles Hegel extended the term to denote the modern attempt to deduce both the laws of nature and of morality from individual consciousness. In subsequent nineteenth-century German histories of philosophy the term (sometimes divided into French and German branches) was used to refer to both rationalist and empiricist tendencies in eighteenth-century philosophy, with Kant frequently portrayed as a thinker who managed to transcend the alleged limits of the movement and thus ushered in a new epoch. The early scholarship in English was heavily influenced by this tradition, with the work of Hibben (1910) representing one of the more sophisticated versions of this approach.

Cassirer offered an even more nuanced account, viewing the Enlightenment as the pivotal phase in the broader process through which "modern philosophic thought gained its characteristic self-confidence and self-consciousness" (1932/1951, p. vi). The book's opening chapter followed d'Alembert in distinguishing the *esprit de système* (the deductive system of seventeenth-century rationalism) from the *esprit systématique*, with its emphasis on induction and empirical analysis that marked the new era. In the discussions of approaches to nature, psychology, religion, history, politics, and aesthetics that followed, Cassirer (1932/1951) portrayed the Enlightenment as a European movement in which German thinkers such as Leibniz and Gotthold Ephraim Lessing stood on equal terms with their French counterparts. While Cassirer eschewed a historical account of various "individual doctrines" in favor of a study of "the form and manner of intellectual activity itself," Hazard (1935/1953) traced the history of responses to what he characterized as a "crisis of the classical mind" (i.e., seventeenth-century rationalism). If Hazard was less certain than Cassirer that this crisis had been resolved, his

account nevertheless saw the Enlightenment (though the term itself does not figure prominently in his work) as an attempt to respond to a philosophical problem: the problem of finding an alternative to religious belief as a foundation for normative judgments. In contrast, Becker (1932) held that far from providing an alternative to religious faith, the *philosophes* simply substituted one sort of faith for another, with a faith in the power of reason occupying the place previously occupied by religion.

However problematic as historical narratives, such studies capture one important feature of eighteenth-century discourse. In France figures such as Voltaire, Diderot, d'Alembert, Baron Paul-Henri Thiry d'Holbach, Claude-Adrien Helvétius, and others described what they were doing as philosophy and called themselves philosophers. Still, while accounts of "the philosophy of the Enlightenment" tend to emphasize the role of epistemological questions, the reach of the term *philosophy* during the Enlightenment was considerably more expansive. Isaac Newton published his laws of motion in a work that announced itself as a contribution to *natural philosophy* and the concerns of the American Philosophical Society, founded in Philadelphia by Benjamin Franklin in 1768, were closer to the modern natural sciences than to philosophy as it is now conceived. For much of the period, treatises on natural law provided thinkers with a context for exploring a wide range of issues in the areas of anthropology, the philosophy of language, political economy, and morality that were central concerns during the period.

The emergence of the salon and the coffeehouse spurred the growth of new forms of expression—for example, Diderot's remarkably open-ended dialogues and publications such the *Tattler* and the *Spectator*, journals edited by Joseph Addison (1672-1719) and Richard Steele (1672–1729) that aimed at improving the discourse and the mores of those who frequented coffeehouses. Many of the period's most influential works—for example, Pierre Bayle's *Historical and Critical Dictionary* (1697), Diderot's *Encyclopedia* (1751–1765), and Voltaire's *Philosophical Dictionary* (1764)—were lexicons, rather than philosophical treatises, while other important texts—including Jean-Jacques Rousseau's educational treatise *Emile* (1762) or Guillaume-Thomas-François de Raynal's influential *History of the Two Indies* (1770)—defy assimilation into familiar genres of philosophical writing.

The staggering variety of works labeled as philosophy is mirrored by the Enlightenment's conception of the vocation of the *philosophe*. The entry in Diderot's *Encyclopedia* (an abridgement of a text generally credited to the grammarian César Chesneau Dumarsais [1676–1756]) characterized the *philosophe* as an individual who is chiefly concerned with those "sociable qualities" that make individuals useful members of society, "For him, civil society is, as it were, a divinity on earth; he flatters it, he honors it by his probity, by an exact attention to his duties, and by a sincere desire not to be a useless or embarrassing member of it" (p. 510). Diderot's article on "Encyclopedia" stressed the differences between the "geniuses" of the seventeenth century, who engaged in solitary and unconstrained reflection on the nature of things, and the collaborative work of the *philosophes* of his own century, whose interest lay less in making new discoveries than in organizing and disseminating the knowledge that had already been attained by artisans and other useful members of society.

A similar view of the mission of the *philosophe* is found in the posthumously published work by the thinker who is often regarded as the last of the species: Marquis de Condorcet's *Sketch for a Historical Picture of the Progress of the Human Mind* (1793). He saw *philosophes* as "concerned less with the discovery or development of truth than with its propagation." Gathering under the banner of "reason, tolerance, humanity," they "made it their life-work to destroy popular errors rather than to drive back the frontiers of human knowledge—an indirect way of aiding its progress which was not less fraught with peril, nor less useful" (pp. 136-137).

Thus, while the Enlightenment saw the publication of works—for example, John Locke's *Essay concerning Human Understanding* (1689), David Hume's *A Treatise of Human Nature* (1739–1740), and Kant's *Critique of Pure Reason* (1781)—that are among the foundational texts of modern philosophy, the eighteenth-century *philosophe* engaged in activities that no longer occupy professional philosophers and a good many of the works that the eighteenth century classified as *philosophy*—for example, the political libels and philosophical pornography that were labeled philosophical books in the clandestine book trade—fall outside the discipline as it is now practiced. For this reason the Enlightenment invoked by philosophers and the Enlightenment studied by historians working in the area of eighteenth-century studies tend to diverge. For the former, the Enlightenment was a philosophical movement that emphasized the application of reason (defined for the most part in terms associated with modern science) to all aspects of life, a project that has been embraced by some (e.g., in Karl Raimund Popper's ideal of the "Open Society") and criticized by others (e.g., in Max Horkheimer and Theodor Adorno's

[1947/2002] account of the collapse of Enlightenment into totalitarianism). In contrast, scholars working in the area of eighteenth-century studies have tended to see the Enlightenment as a network of individuals and institutions, sometimes bound together by common interests or purposes, but in many cases diverging according to local contexts or their particular concerns and commitments.

ENLIGHTENMENT PROJECTS

THE PUBLIC USE OF REASON. As a general characterization of the movement's aims, there is much to recommend Kant's definition of enlightenment as "the freedom to make a *public use* of one's reason in all matters" (p. 36). Both the essay's demand that individuals make use of their own reason and its invocation of a cosmopolitan public sphere of readers and writers reiterated ideals that had accompanied the Enlightenment from the start. The requirement that the claims of religious, political, and other authorities be brought before what Kant called the "tribunal of reason" had, for example, been a point of honor for the deist John Toland, who opened his *Christianity Not Mysterious* (1696) by observing that he had "been very early accustom'd to Examination and Enquiry, and taught not to captivate my Understanding no more than my sense to any Man or Society" (p. 7). The idea that individuals might best carry out this project of thinking for themselves in the company of others had been central to Pierre Bayle's conception of a "republic of letters" consisting of readers and critics who were bound together, despite their separation in different countries, into a common endeavor.

The emergence of what the social theorist Jürgen Habermas termed the *public sphere*—the network of institutions, including coffeehouses, salons, Masonic lodges, and reading societies in which "private people come together as a public" (1989, p. 27)—is viewed by many historians as a defining feature of the period. Coffeehouses, particularly in England, provided a venue for the circulation and discussion of news, Parisian salons played an essential role in coordinating the activities of the *philosophes*, and the Masonic movement opened a space in which new forms of sociability, expressing the ideal of fraternal solidarity, were possible. No less significant was the emergence of an international book trade, with both legal and clandestine branches. Indeed, the most compelling evidence for the spread of enlightenment in eighteenth-century Europe may be the explosion of books and periodicals that made their way into new markets, the dramatic shift in the content of these books (with works on religious subjects eclipsed by a growing interest in science and literature by the end of the century), and the shift in reading practices from the repeated reading of a few texts (typically devotional in character) to the successive reading of a series of books, a practice that further increased the demand for new works.

TOLERATION AND RELIGIOUS HETERODOXY. Kant's suggestion that "religious matters" were central to the concerns of enlightenment and his insistence that restrictions on the public use of reason in this area were both "harmful" and "dishonorable" aptly summarized the views of those who saw themselves as engaged in efforts at enlightenment. The initial impetus behind the Enlightenment stemmed, in part, from Protestants' revulsion at Louis XIV's (1638–1715) campaign against the Huguenot minority (culminating in his Revocation of the Edict of Nantes in 1685) and reservations regarding the policies of the Catholic monarch James II (1633–1701) in England (culminating in his removal in the "Glorious Revolution" of 1688). Such concerns were particularly evident among the political and religious exiles from France and England who gathered in the Dutch Republic at the close of the seventeenth century, where they produced tracts on religious and political questions that ranged from such classic texts as Locke's *Letter concerning Toleration* (1689)—a work that had a pervasive influence throughout Europe and the New World on discussions of the proper roles of church and state—to the infamous *Treatise of the Three Imposters*, a clandestine manuscript that pieced together bits of Benedict (Baruch) de Spinoza, Thomas Hobbes, and various materialists to argue that Judaism, Christianity, and Islam owed their origins to the attempts of "imposters" (i.e., charlatans or magicians) to gain political power.

Toleration was the common cause of all those associated with the Enlightenment. In England Protestant dissenters such as Joseph Priestley and Richard Price drew on the arguments of Locke in their campaign against the limitations on political participation suffered by those who refused to swear conformity to central articles of the Anglican faith (e.g., the doctrine of the trinity). Similar arguments could be found, at the end of the century, in Moses Mendelssohn's *Jerusalem* (1783), a treatise on the relation between civil and ecclesiastical power. In France Voltaire—profoundly influenced by the diversity of religious practices he observed during his visit to England—waged a life-long campaign in support of toleration, culminating in an effort to clear the reputation of Jean Calas, a Huguenot executed under circumstances that, for Voltaire, epitomized the corruption of justice by religious fanaticism. By the end of the period the campaign for tol-

eration could claim such legislative achievements as Thomas Jefferson's *Virginia Statute of Religious Freedom* (1786) and Article X of the French *Declaration of the Rights of Man and Citizen* (1789).

The period was also marked by efforts at purifying Christian doctrine from what were seen as subsequent distortions and fabrications. Both Locke's *Reasonableness of Christianity* (1695) and Toland's *Christianity Not Mysterious* presented themselves as attempts to recover Christ's original teaching—which they argued contained nothing that contradicted what could be ascertained through "natural" reason—from what the more pugnacious Toland characterized as "the craft and ambition of Priests and Philosophers" (p. 96). More moderate versions of such arguments persisted to the end of the period in the so-called neologism embraced by Berlin clergy, whose sermons and writings denounced popular superstitions and religious enthusiasm as contrary to a conception of Christian doctrine and emphasized the importance of moral and civic responsibilities. Parallel efforts at reform could be found within the Ashkenazic Jewry in what came to be designated the *Haskalah* (the Hebrew term for *enlightenment*).

Projects of reform, however, easily crossed over into the advocacy of heterodoxy, with Socinian and pantheist doctrines having a broad appeal. For example, Toland's later writings, which hailed the Druids as practitioners of a "natural" religion, articulated positions that are difficult to reconcile with any established version of Christianity. The same is true for the work of Lessing, especially his *Education of the Human Race* (1777), a text that influenced Hegel's early writings. While explicit endorsements of atheism remained a minority position within the Enlightenment (Holbach's *System of Nature* [1770] was the famous notorious exception), Spinoza's writings held a particular interest for more radical free-thinkers, and various pantheist and materialist doctrines lent support to formulations in which references to the deity contributed rather little to the argument.

THE NEWTONIAN IDEAL AND THE RISE OF A SCIENTIFIC CULTURE

Though known chiefly by reputation or through popularizations, the work of Newton had a significant impact during the period. His influence was felt in England both in the increasing interest in experimental approaches to natural philosophy and in the popularity of his arguments among religious dissenters. On the Continent advocates of Newton's cosmology challenged Cartesian and Leibnizian approaches, with Newtonians eventually gaining the upper hand within the French Academy of Sciences and the Berlin Academy. Voltaire and Alembert were effective advocates of Newtonian positions before the broader reading public, as was Voltaire's mistress Émilie du Châtelet (1706–1749), a skilled translator of and commentator on Newton's work.

Attempts to extend Newton's approach to other areas were frequent, with the *Optiks* (1704) rather than the more daunting *Philosophiae Naturalis Principia Mathematica* (1687) serving as a paradigm. The study of electrical phenomena attracted a great deal of interest, with Franklin's contributions to the field enjoying a wide readership in Europe. There were also notable attempts to apply what were viewed as Newtonian approaches to moral philosophy. Hume subtitled his *Treatise of Human Nature* (1739) "an attempt to introduce the experimental method of reasoning into moral subjects," Adam Smith employed analogies to gravitational attraction in his *Theory of Moral Sentiments* (1759), and Condorcet attempted to bring mathematical approaches to bear on political decision making.

More generally, science and scientific reasoning came to enjoy an enhanced status among educated laypersons. The predictive success of Newton's laws in mapping the paths of celestial bodies—most notably Edmond Halley's (1656–1742) application of these laws to the path of the comet that now bears his name—played a role in this process, as did such practical innovations as Franklin's lightning rod. Scientific academies—both state sponsored and private—also had a significant impact in demonstrating the practical implications of scientific inquiry.

HUMAN NATURE AND CULTURAL DIVERSITY

The application of Newtonian approaches to the study of politics and society was but one example of a broader interest in the study of human nature. Accounts of the voyages of James Cook (1728–1779) and Louis-Antoine de Bougainville (1729–1811) brought reports of peoples whose social arrangements, moral practices, and views on religion differed radically from European norms and that posed significant challenges to assumptions regarding the uniformity of human nature. Theories that attempted to explain this diversity in terms of differences in modes of subsistence (hunting and gathering, pastoral, agricultural, and commercial) were particularly prominent among thinkers associated with the Scottish Enlightenment. There were also extended debates on the origin of different races (a term that had a much wider meaning

during this period than it would take on during the nineteenth century) between those who, like the French naturalist Comte de Georges-Louis Leclerc Buffon, maintained that all human beings descended from a common origin and that racial differences were the result of climate, and those who, like the Swede Carl von Linné (1707–1778), argued that the different races had descended from different ancestors.

Beyond these theoretical disputes, the literature on "savage peoples"—particularly accounts of the allegedly idyllic life of the natives of the newly discovered island of Tahiti—provided a means for criticizing European society. Rousseau's *Discourse on the Origins of Inequality* (1755) and Diderot's *Supplement to Bougainville's Voyage* (begun in 1772) can serve as examples of this mode of argument, which had an influential predecessor in Baron de Montesquieu's *Persian Letters* (1721).

EFFORTS AT "IMPROVEMENT"

The political views of those associated with the Enlightenment diverged widely. Some favored constitutional monarchies (with England representing one possible model), while others placed considerable hope in the efforts of reform-minded absolutists such as Frederick II of Prussia (1740–1786) and Joseph II of Austria (1741–1790). In the wake of the American Revolution republican ideas gained supporters in both England and France.

What was more pervasive than an allegiance to any particular political ideology was a concern with what was loosely characterized as "improvement." The interest of Scottish enlighteners in the promises of commercial development was reflected in Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations* (1776). In France Jacques Necker (1732–1804) and his protégé Condorcet wrestled with the worsening fiscal and political crises that plagued the French monarchy in its final decades. Throughout Europe various societies examined ways of improving agricultural production, fostering the growth of manufacturing, and increasing the circulation of commercial goods. For example, the Lunar Society of Birmingham—whose membership included the inventors James Watt (1736-1819) and Matthew Boulton (1728–1809), the manufacturer Josiah Wedgewood (1730–1795), and the polymaths Priestley and Erasmus Darwin—waged a wide-ranging campaign for political reform and commercial development.

Perhaps there is no more compelling testimony on the role of the Enlightenment in shaping the modern world than the emergence, since the 1940s, of critiques of the so-called Enlightenment Project that hold it responsible for the various alleged pathologies of modernity (for discussions, see Baker and Reill [2001] and Gordon [2001]). While this literature tends to be rather selective in its conception of what this alleged project involved, the diversity of charges that have been leveled against the Enlightenment speaks to the complexity of the movement and its perceived relevance for the present.

See also Addison, Joseph; Adorno, Theodor; Alembert, Jean Le Rond d'; Bayle, Pierre; Buffon, Georges-Louis Leclerc, Comte de; Cassirer, Ernst; Clandestine Philosophical Literature in France; Condorcet, Marquis de; Darwin, Erasmus; Descartes, René Diderot, Denis; Encyclopédie; Enlightenment, Islamic; Enlightenment, Jewish; Franklin, Benjamin; Habermas, Jürgen; Hegel, Georg Wilhelm Friedrich; Helvétius, Claude-Adrien; Hobbes, Thomas; Holbach, Paul-Henri Thiry, Baron d'; Horkheimer, Max; Human Nature; Hume, David; Jefferson, Thomas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lessing, Gotthold Ephraim; Locke, John; Mendelssohn, Moses; Montesquieu, Baron de; Newton, Isaac; Paine, Thomas; Popper, Karl Raimund; Priestley, Joseph; Rousseau, Jean-Jacques; Smith, Adam; Socinianism; Spinoza, Benedict (Baruch) de; Toland, John; Toleration; Voltaire, François-Marie Arouet de.

Bibliography

Baker, Keith Michael, and Peter Hanns Reill, eds. *What's Left of Enlightenment? A Postmodern Question*. Stanford, CA: Stanford University Press, 2001.

Becker, Carl L. *The Heavenly City of the Eighteenth Century Philosophers*. New Haven, CT: Yale University Press, 1932.

Cassirer, Ernst. *The Philosophy of the Enlightenment* (1932). Translated by Fritz C. A. Koelln and James P. Pettegrove. Princeton, NJ: Princeton University Press, 1951. Originally published under the title *Die Philosophie der Aufklärung*.

Condorcet, Marie Jean Antoine-Nicolas, Marquis de. *Sketch for the Historical Progress of the Human Mind* (1795). Translated by June Barraclough. London: Weidenfeld and Nicolson, 1955. Originally published under the title *Tableau Historique des Progrès de l'Esprit Humain*.

Darnton, Robert. *Forbidden Best-Sellers of Pre-Revolution France*. New York: Norton, 1995.

Dumarsais, César Chesneau. "Philosophe." In *Encyclopèdie, ou Dictionnaire Raisonné des Sciences, Arts et des Métiers*. Vol. 8, edited by Denis Diderot and Jean Le Ronde D'Alembert (Paris, 1757).

Gay, Peter. *The Enlightenment: An Interpretation*. 2 vols. New York: Knopf, 1966–1969.

Goodman, Deena. *The Republic of Letters: A Cultural History of the French Enlightenment*. Ithaca, NY: Cornell University Press, 1994.

Gordon, Daniel, ed. *Postmodernism and the Enlightenment: New Perspectives in Eighteenth-Century French Intellectual History*. New York: Routledge, 2001.

Habermas, Jürgen. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Translated by Thomas Burger and Frederick Lawrence. Cambridge, MA: MIT Press, 1989.

Hazard, Paul. *The European Mind: The Critical Years, 1680–1715* (1935). New Haven, CT: Yale University Press, 1953. Originally published under the title *La crise de la conscience europeén*.

Hibben, John Grier. *The Philosophy of the Enlightenment*. New York: Scribner's, 1910.

Horkheimer, Max, and Theodor Adorno. *Dialectic of Enlightenment: Philosophical Fragments* (1947). Translated by Edmund Jephcott. Stanford, CA: Stanford University Press, 2002. Originally published as *Dialektik der Aufklärung: Philosophische Fragmente*.

Israel, Jonathan I. *Radical Enlightenment: Philosophy and the Making of Modernity 1650–1750*. New York: Oxford University Press, 2001.

Jacob, Margaret C. *Living the Enlightenment: Freemasonry and Politics in Eighteenth-Century Europe*. New York: Oxford University Press, 1991.

Kant, Immanuel. "Beantwortung der Frage: Was ist Aufklärung?" In Kant, *Gesammelte Schriften*. Vol. 8, edited by Könglich Preussichen Akademie der Wissenschaften (Berlin: Walter de Gruyter, 1923).

Kors, Alan Charles, ed. *Encyclopedia of the Enlightenment*. 4 vols. New York: Oxford University Press, 2003.

Melton, James van Horn. *The Rise of the Public in Enlightenment Europe*. New York: Cambridge University Press, 2001.

Outram, Dorina. *The Enlightenment*. New York: Cambridge University Press, 1995.

Porter, Roy, and Mikulás Teich, eds. *The Enlightenment in National Context*. New York: Cambridge University Press, 1981.

Schmidt, James. "Inventing the Enlightenment: British Hegelians, Anti-Jacobins, and the *Oxford English Dictionary*." *Journal of the History of Ideas* 64 (3) (2003): 421–443.

Schmidt, James, ed. *What Is Enlightenment?: Eighteenth-Century Answers and Twentieth-Century Questions*. Berkeley, CA: University of California Press, 1996.

Toland, John. *Christianity Not Mysterious* (1696), edited by Philip McGuinness, Alan Harrison, and Richard Kearney. Dublin: Lilliput, 1997.

James Schmidt (2005)

ENLIGHTENMENT, ISLAMIC

BEGINNINGS

The Islamic *Nahḍah* (rebirth, renaissance) started in Syria and achieved its real momentum in Egypt in the nineteenth century, then as subsequently the intellectual engine room of Islamic intellectual life. The *Nahḍah* movement represented an attempt to do two things. One was to introduce some of the main achievements of European culture into the Islamic world. The other was to defend and protect the major positive features of Arab and Islamic culture and revive them despite the assaults of European imperialism. The important features of the movement were the attempt to combine these policies and the reaction to the apparent decadence of the Arab world not by rejecting Arab culture but by purifying it and introducing it to aspects of modernity from without that were seen as acceptable from an Islamic point of view.

DEVELOPMENT

The main *Nahḍah* thinkers were Sayyid Jamāl ad-Dīn al-Afghānī (1838–1897), and Muḥammad ʿAbduh (1849–1905), who in their different ways sought to confront modernity not by rejecting it nor by rejecting Islam, but by effecting some kind of synthesis. The Renaissance movement suggested that one could accept some European ideas and reject others, thus preserving tradition while adopting modernity at the same time. *Nahḍah* argued that Islam is itself a profoundly rational system of thought and has no problem in accepting science and technology. So there is no reason for Muslims to abandon their faith while at the same time accepting the benefits of European forms of modernity. Interestingly, the significance of reviving Islam or Arabism played a considerable part in the political rhetoric of the time.

The most important intellectual figure in this movement was undoubtedly al-Afghānī, who as his name suggested had close connections with Afghanistan, where part of his early education took place. He seems to have been deliberately unclear about his precise ethnic origins to prevent that from being a divisive factor in his attempts to address the whole Islamic community. At the age of around eighteen he moved to India, where he came across the modernist ideas of Sayyid Ahmad Khan (1817–1898), whom he later attacked in his *Refutation of the Materialists*. Ahmad Khan was intent on proving to the British rulers of India that Islam was a religion capable of accepting rationality, and it was this apologetical tone that al-Afghānī attacked. His arguments were not based on Islam alone; they also borrowed a great deal from what he regarded as science and philosophy. He argued that Islamic philosophy was perfectly compatible with modern science and technology and should encourage Muslims to acquire the necessary skills to resist the impact of

European imperialism. Part of the Islamic Renaissance ideology was that there should be a rebirth and rediscovery of the main intellectual and political achievements of the Islamic world during its high point.

Al-Afghānī's *Refutation of the Materialists* argues that the source of evil is materialism, the philosophical doctrine that argues that the world has developed out of a set of material preconditions. He also criticizes the theory of evolution, which he sees as denying God's role in designing the world. His critique also has a social aspect in that materialism is held to reject founding society on any common moral values, and in being critical of religion as such, and of Islam in particular. This sort of critique of what is seen as European culture has since the nineteenth century become common in the Islamic world.

The influence of his ideas was amplified by the efforts of Rashīd Riḍā (1865–1935), who founded in 1898 the journal *al-Manar* (The Lighthouse) in Cairo. The central theme of the journal was that there is no incompatibility between Islam on the one hand, and modernity, science, reason, and civilization on the other. Riḍā tended to emphasize religion and was a firm opponent of secularism, arguing that supporting modernity did not imply advocating secularism.

Muḥammad ʿAbduh used his position as head of al-Azhar, the leading theological university in the Sunni Islamic world, to propound the message of the *Nahḍah* that the Islamic world should accept modernity while at the same time not rejecting Islam. The period of stagnation that he identified with the tenth to the fifteenth centuries was a time when the early scientific and philosophical progress of the Islamic cultural world came to an end and the political and religious authorities had a mutual interest in maintaining control by restricting the intellectual curiosity of those over whom they ruled so effectively. What was now needed, he argued in the nineteenth century, was reform of all the institutions of the Islamic world, while preserving the timeless truths of Islam itself. He suggested that the connection between religion and modernity, in particular between Christianity and modernity, is entirely misplaced. Christianity itself advocates belief in the transience of everyday life, by contrast with the concern for possessions and comfort so characteristic of modern industrial societies. Still, Christianity found no inconsistency in combining religion with modernity, so this need not be a worry for Muslims either.

See also Evolutionary Theory; Islamic Philosophy; Materialism; Rationality; Renaissance.

Bibliography

Hourani, Albert. *Arabic Thought in the Liberal Age, 1798–1939*. New York: Cambridge University Press, 1983.

Keddie, Nikkie R. *An Islamic Response to Imperialism: Political and Religious Writings of Sayyid Jamāl al-Dīn "al-Afghānī"*. Berkeley: University of California Press, 1983.

Keddie, Nikkie, ed. *Scholars, Saints, and Sufis: Muslim Religious Institutions in the Middle East since 1500*. Berkeley: University of California Press, 1972.

Kedourie, Elie. *Afghani and ʿAbduh: An Essay on Religious Unbelief and Political Activism in Modern Islam*. New York: Humanities Press, 1966.

Kurzman, Charles. *Modernist Islam, 1840–1940: A Source Book*. New York: Oxford University Press, 2002.

Oliver Leaman (2005)

ENLIGHTENMENT, JEWISH

Growing emancipation of European Jews in the eighteenth century was matched by an intellectual movement that came to be called the Jewish Enlightenment or *Haskalah*. Jews started to enter the mainstream of European society, in particular in major German cities such as Berlin, and Jewish thinkers had to accomplish two tasks. They needed to show their Gentile peers that they were just as committed to rationality as anyone else, and they needed to persuade other Jews that they should establish links with the local non-Jewish cultures in which they lived.

The main embodiment of this movement was Moses Mendelssohn, who participated fully in German philosophy and culture, and lesser thinkers were Marcus Herz (1747–1803), Salomon Maimon (1753/4–1800), and Nachmun Krochmal (1785–1840). Mendelssohn first of all emphasized the importance of mastery of the local secular language, and of the contemporary culture. But this did not imply abandoning Judaism; he argued on the contrary that one could use modern ways of argumentation to explain and justify religious systems such as Judaism. He comes to argue that Judaism is a profoundly rational religion and so highly appropriate for those committed to reason. Mendelssohn was here reacting to the widespread view that Judaism was a rule-bound and legalistic system that only those stuck in a worn-out culture would persist in following. It came to be argued in German philosophy by Kant and Hegel that Judaism was a religion essentially superseded a long time ago in the past, fossilized and unsatisfactory, and Mendelssohn and other *maskilim* (Enlighteners) argued that these criticisms were misplaced.

The basis of *Haskalah* was respect for reason, as the word suggests (*sekhel* being *reason* in Hebrew) and this was to have longstanding effects on Jewish culture. It contributed to the start of Reform Judaism in Germany, its basis being a putative rational attitude to traditional Judaism. It also played a part in the secular nature of Zionism, the idea that the Jews, like other national groups, had a right to a homeland that was based on reason not tradition. After its growth and development in Germany, *Haskalah* moved east to affect the Jewish communities there, and produced a schism between the "modernizers" and those concerned to defend tradition. The *Haskalah* raised the issue of how far a religion upheld by a minority excluded from mainstream society could survive when that minority was allowed to join that society. If it could be argued that the traditional religion was as rational as anything else in society then the intellectual presuppositions of such a change might not threaten the survival of the religion. The *maskilim* were confident that both Judaism and secular European society would benefit from a more intimate relationship, because the basis of both is reason.

See also Maimon, Salomon; Mendelssohn, Moses.

Bibliography

Arkush, Alan. *Moses Mendelssohn and the Enlightenment*. Albany, NY: State University of New York Press, 1994.

Dubin, Lois. "The Social and Cultural Context: Eighteenth-Century Enlightenment." In *History of Jewish Philosophy*, edited by D. Frank and O. Leaman. London: Routledge, 1997.

Feiner, Shmuel. *The Jewish Enlightenment*. Translated by C. Naor. Philadelphia: University of Pennsylvania Press, 2004.

Harris, J. *Nachman Krochmal: Guiding the Perplexed of the Modern Age*. New York: New York University Press, 1991.

Rotenstreich, Nathan. *Jews and German Philosophy: The Polemics of Emancipation*. New York: Schocken, 1984.

Sorkin, David. *Moses Mendelssohn and the Religious Enlightenment*. Berkeley: University of California Press, 1996.

Oliver Leaman (2005)

ENTAILMENT, PRESUPPOSITION, AND IMPLICATURE

Entailment, as used by philosophers, is a term of art that, unlike *logical consequence*, lacks a precise definition that is consistently adhered to by those who employ it. Throughout much of the twentieth century, especially its early and middle years, many philosophers connected entailment with analyticity, requiring the material conditional ⌜A ⊃ B⌝ to be analytic when A entailed B. In later years, as conceptions of analyticity became less expansive, and philosophical uses of it more restricted, the presumption that entailment was to be understood in terms of analyticity waned. However, the relationship between entailment and necessity has remained robust. Standardly, when it is claimed that A entails B, B is taken to be a necessary consequence of A in the sense that it is impossible for A to be true without B's being true. Often, though not always, B is required to be apriori deducible from A, as well. The relata, A and B, are naturally thought of as propositions, or statements—in the sense of that which is stated by an assertive utterance of a sentence. However, sometimes theorists speak of sentences themselves as entailing other sentences. In such cases, it is natural to construe the relation holding between sentences as deriving from the primary entailment relation holding between the propositions they express.

A potentially more restrictive understanding of entailment requires that when A entails B, the falsity of A is a necessary consequence of the falsity of B. When entailment is understood in this way, it is sometimes contrasted with logical presupposition: A proposition A logically presupposes a proposition B if and only if the truth of B is a necessary condition for A's being either true or false. The most widely discussed (putative) examples of logical presuppositions are so-called existential presuppositions, corresponding to uses of singular terms. (These are also sometimes called *referential presuppositions*.) For example, according to a Fregean analysis of definite descriptions, the propositions expressed by (1a) and (1b) logically presuppose the proposition expressed by (1c).

1a. The person who proved Goldbach's conjecture is brilliant.

1b. The person who proved Goldbach's conjecture isn't brilliant.

1c. One and only one person proved Goldbach's conjecture.

For Frege, singular definite descriptions are complex singular terms, and the predicate *is brilliant* designates a function that assigns truth to some individuals and falsity to others. Because the sense of *the person who proved Goldbach's conjecture* fails to pick out any individual, the function designated by *is brilliant* has no argument to operate on, and (1a) is characterized as being neither true

nor false. The same is true of (1b), which is taken to be the negation of (1a). Because, for Frege, the negation function—which assigns truth to falsity, and falsity to truth—has no argument to operate on in this case, proposition (1B) is not assigned any truth value. On this analysis, the truth of the logical presupposition, (1c), is a necessary condition for (1a) and (1b) to be either true or false.

THE SEMANTICS OF FREGE AND RUSSELL

Although the compositional semantics of Frege (1891, 1892a, and 1892b) produce elegant results in cases such as this, they run into trouble when fully generalized. For Frege, *n*-place truth-functional operators designate *n*-place truth functions, and the truth value of a truth-functional compound is the value of the relevant truth function at the *n*-tuple of truth values of its sentential constituents. Hence, the argument used to show that the negation of a proposition is truth valueless if and only if the proposition negated is truth valueless can be generalized to yield the conclusion that a truth functional compound is truth valueless if and only if one of its constituents is. This result is demonstrably incorrect, as is shown by (2a) and (2b)—based on an example given by Bertrand Russell in "On Denoting" (1905). (Read *if, then*, in (2a) as material implication.)

2a. If one and only one person proved Goldbach's conjecture, then the person who proved Goldbach's conjecture is brilliant.

2b. Either it is not the case that one and only one person proved Goldbach's conjecture, or the person who proved Goldbach's conjecture is brilliant.

Far from being truth valueless, these examples are made true because no one has proved Goldbach's conjecture.

This was one of the considerations that led Russell to analyze the examples in (1) differently from Frege. On his analysis, the logical form of (1a) is (R1a), while (1b) is ambiguous between (R1bw), in which the description is said to have wide scope, and (R1bn), in which it takes narrow scope.

R1a. $\exists x\ [\forall\ y((y \text{ is a man } \&\ y \text{ proved Goldbach's conjecture}) \leftrightarrow x = y)\ \&\ x \text{ is brilliant}]$

R1bw. $\exists x\ [\forall\ y((y \text{ is a man } \&\ y \text{ proved Goldbach's conjecture}) \leftrightarrow x = y)\ \&\ \sim x \text{ is brilliant}]$

R1bn. $\sim \exists\ x\ [\forall\ y((y \text{ is a man } \&\ y \text{ proved Goldbach's conjecture}) \leftrightarrow x = y)\ \&\ x \text{ is brilliant}]$

When (1c) is false, (R1a) and (R1bw) are also false, but (R1bn) is true. On this analysis, (1c) is a necessary consequence of (1a), and of the reading of (1b) represented by (R1bw). However, on this reading, (1b) is not the (logical) negation of (1a). Hence, for Russell, these examples are not instances of logical presupposition.

STRAWSON'S THEORY OF PRESUPPOSITION

In "On Referring" (1950), Peter Strawson considered such cases, and presented his own analysis that included the following theses: (i) meaning is a property of expressions; referring, saying something, and being true or false are properties of uses of expressions in contexts; (ii) to give, or know, the meaning of a sentence S is to give, or know, a rule for determining the contexts in which S is used to say something true, and the contexts in which it is used to say something false; (iii) the primary referring use of a name, demonstrative pronoun, or singular definite description is one in which the term is used to refer to something that the rest of the sentence is used to say something about; the meaning of such a term, when used in this way, is a rule for determining its referents in different contexts; (iv) if a singular term b in a sentence ⌜Fb⌝ is used referringly in a context C, then ⌜Fb⌝ says something true (false) in C if and only if, in C, the referent of b has (does not have) the property that F is used to express; if b fails to refer to anything, then ⌜Fb⌝ fails to say anything true or false in C; (v) definition: S presupposes p relative to C if and only if the truth of p is a necessary condition for a use of S in C to say something true or false; and (vi) uses of ⌜The F is G⌝, ⌜All Fs are G⌝, ⌜Some F's are G⌝, ⌜No Fs are G⌝, and ⌜Some Fs are not G⌝ presuppose (in the sense of (v)) that expressed by ⌜There is at least one F⌝.

Thesis (ii) is problematic. As it stands, it does not rule out, and may even seem to suggest, that the meaning of a sentence is a function from contexts of utterance to truth values. According to a better picture, presented by David Kaplan in "Demonstratives" (1989), the meaning of a sentence is a function from contexts to propositions, where the latter determine functions from circumstances of evaluation to truth values. When this view is substituted for (ii), corresponding changes must be made in theses (iii) and (iv). Strawson's emphasis on referring as the semantic function of a singular term, plus his tendency to treat referring uses of demonstratives as prime examples of this function, suggest a reformulation of (iii)

and (iv) in which all referring uses of singular terms are, in Kaplan's words, *directly referential*. (iii_K) A referring use of a singular term b, as part of a sentence S, in a context C contributes the referent of b in C to the proposition expressed by S in C; the meaning of a singular term is a rule for determining the propositional constituents contributed by uses of b to the propositions expressed by sentences containing b in different contexts. (iv_K) If a referential use, in a context C, of a singular term b in a sentence ⌜Fb⌝ refers to o, and if F is used to express the property P, then ⌜Fb⌝ expresses a proposition in C that is true (false) in a possible circumstance w if and only if o has (doesn't have) P in w, if b fails to refer to anything in C, then there is no propositional constituent corresponding to b in C, and ⌜Fb⌝ fails to express a (complete) proposition in C.

The theory of presupposition that emerges from this reconstruction of Strawson's theses is a theory of what may be called *expressive presupposition*: A sentence S expressively presupposes a proposition p relative to a context C if and only if the truth of p is necessary for S to semantically express a (complete) proposition in C. This theory provides a plausible account of examples in which a pronoun, demonstrative, or demonstrative phrase is used referringly. However, the theory produces incorrect results when extended to the range of cases mentioned in thesis (vi). Such an extension also conflicts with Strawson's expressed intentions. In *Introduction to Logical Theory* (1952), he defines presupposition as follows: A statement (proposition) S presupposes a statement (proposition) S' if and only if the truth of S' is a necessary condition for S to be true or false. Because this is a definition of logical presupposition, Strawson's adoption of it belies any clear commitment to expressive presupposition, or any systematic analysis of the examples in (vi) along directly referential lines.

This suggests a second possible reconstruction of his position. On this interpretation, his theory of presupposition is substantially the same as Frege's, without the compositional semantics, but with the stipulation that statements involving certain generalized quantifiers bear existential presuppositions. This theory is broad in scope and has been historically influential. However, its leading ideas are not original with Strawson. As a historical point, it would be a mistake to attribute to him either an account of presupposition that is systematically Fregean (logical), or an account that is systematically expressive. His major discussions include elements of both, the conflict being masked by the flawed account of meaning given in thesis (ii), and the failure to articulate the more satisfactory picture later provided by Kaplan.

PRAGMATIC PRESUPPOSITION AND CONVERSATIONAL DYNAMICS

An important advance in the study of presupposition, signaled in Stalnaker (1973, 1974) and Lewis (1979), integrates presupposition into a broader theory of conversational dynamics. The crucial observation is that sentences are used in communication to contribute to an existing conversational record, which contains background assumptions shared by conversational participants. Because of this, it is natural for speakers to develop conventional means of indicating what assumptions they are making about the conversational record to which their utterances contribute. Pragmatic presuppositions may then be understood as requirements imposed on such records by utterances. Suppose, for example, that a use of S (e.g. "It wasn't Andrew who solved the problem") requires the set of background assumptions prior to the utterance to contain a specific proposition p (that someone has solved the problem). Imagine a conversation in which p is not already among the shared background assumptions prior to the utterance of S, but conversational participants are willing to accept p as uncontroversial. What response would be reasonable in such a case? The legalistic response would be to object to the speaker's remark on the grounds that p, which was required by the utterance of S, had not already been established. The speaker could then ask whether hearers were willing to accept p, and be told that they were. After adding p to the conversational record, the speaker could repeat the original remark, and continue.

But there would be no point to this. Because hearers are ready to accept p anyway, they may as well add it to the background, and let the speaker go on without objection. In short, the most efficient response is to accommodate the speaker by updating the conversational record so that it meets the requirements of the utterance. Knowing that hearers can work this out on their own, the speaker can exploit this strategy of accommodation by uttering sentences the presuppositional requirements of which are not already satisfied by the conversational record—provided the requirements are both recognizable and uncontroversial. One virtue of this pragmatic approach is its eclecticism regarding different factors that may give rise to presuppositional requirements. As indicated in Soames (1989), logical presupposition, expressive presupposition, conventional implicature, constraints on the interpretation of anaphora, and non-conventional pragmatic facts

have all been suggested as sources of pragmatic presuppositions. Further developments are found in Heim (1982, 1983) and Beaver (2001).

CONVERSATIONAL AND CONVENTIONAL IMPLICATURES

Closely related to pragmatic presuppositions, are conversational and conventional implicatures, introduced in Grice (1989) (originally delivered as the William James Lectures at Harvard in 1967). For conversational implicatures, the key insight is that the efficient and rational exchange of information by cooperative speakers is governed by maxims that include: (i) don't make your conversational contribution too weak (or too strong); (ii) don't say that which you believe to be false, or that for which you lack adequate evidence; (iii) be relevant; and (iv) avoid obscurity and ambiguity; be brief and be orderly.

Conversational implicatures are propositions that a speaker is committed to, above and beyond that which is said or asserted, by virtue of the presumption that the conversational maxims are being obeyed. According to Grice, a speaker s conversationally implicates q by saying p if and only if (a) s is presumed to be observing the conversational maxims, (b) the supposition that s believes q is required in order to make s's saying p consistent with that presumption, and (iii) s assumes that the hearers can recognize both this requirement and that s is assuming this. For example, if s assertively utters a disjunction ⌜A or B⌝, then standardly s conversationally implicates that there are non-truth-functional grounds for the assertion (because if s's grounds were truth-functional, and hence sufficient for asserting either disjunct alone, then s's utterance of the disjunction would be too weak, and hence violate maxim (i)). This shows that the simple truth-functional semantics for disjunction do not have to be complicated in order to explain why assertive utterances of disjunctive sentences standardly convey non-truth-functional information.

Another example of some philosophical significance involves the observation in Austin (1964) that it would be an abuse of language for a man who can see that there is a pig in front of him—without having to make any special investigation or to infer the presence of the pig from other propositions—to assert merely that it seems to him as if a pig is present, or that he has evidence of the presence of a pig. From this Austin concludes that such assertions would be false, and that, in a case such as the one imagined, a person can have empirical knowledge without having evidence for the proposition known. However, as pointed out in Ayer (1967), Austin's observation does not support his conclusion. Because the speaker in the imagined situation is in a position to make the stronger claim that a pig is present, the decision to make the weaker statement instead violates Grice's maxim (i). The abuse of language here is not that of stating a falsehood, but of conversationally implicating one.

Like conversational implicatures, Gricean conventional implicatures are propositions to which the speaker is committed, despite their not being parts of what is said by the speaker's utterance. The difference between the two is that the former arise from the conversational maxims, whereas the latter are due to aspects of meaning. For example, an utterance of "She is poor but honest" conventionally implicates—in virtue of the nonassertive meaning of "but"—that there is some contrast between poverty and honesty, an utterance of "He is an Englishman, and therefore, brave" conventionally implicates—in virtue of the nonassertive meaning of "therefore"—that being brave is an expected consequence of being an Englishman, an utterance of "Mary hasn't arrived yet" conventionally implicates—in virtue of the nonassertive meaning of "yet"—that Mary's arrival is expected, and an utterance of "It wasn't Andrew who solved the problem" conventionally implicates—in virtue of the nonassertive meaning attached to the construction "It was (wasn't) NP who VPed"—that someone solved the problem. (Contrast this with "Andrew didn't solve the problem.") A significant point, developed in Karttunen and Peters (1979), is that conventional implicatures such as these may plausibly be regarded as pragmatic presuppositions. This suggests that the nonassertive meaning that generates them may be presuppositional in nature.

See also Analyticity; Austin, John Langshaw; Ayer, Alfred Jules; Frege, Gottlob; Grice, Herbert Paul; Kaplan, David; Lewis, David; Presupposition; Propositions; Russell, Bertrand Arthur William; Semantics; Strawson, Peter Frederick.

Bibliography

Austin, John L. *Sense and Sensibilia*. New York: Oxford University Press, 1964.

Ayer, A. J. "Has Austin Refuted the Sense-Datum Theory?" *Synthese* 17 (1967): 117–140.

Beaver, David. *Presupposition and Assertion in Dynamic Semantics*. Stanford CA: CSLI, 2001.

Frege, Gottlob. "Function and Concept." An address given to the *Jenaisch Gesellschaft fur Medicin und Naturwissenschaft*, January 9, 1891. Translated by Max Black in *Translations from the Writings of Gottlob Frege*, edited by Peter Geach and Max Black. Oxford: Blackwell, 1970.

Frege, Gottlob. "On Concept and Object." *Vierteljahrsschrift fur wissenschaftliche Philosophie* 16 (1892a): 192–205. Translated by Max Black in *Translations from the Writings of Gottlob Frege*, edited by Peter Geach and Max Black. Oxford: Blackwell, 1970.

Frege, Gottlob. "On Sense and Reference." *Zeitschrift fur Philosophie und philosophische Kritik* 100 (1892b): 25–50. Translated by Max Black in *Philosophical Review* 57 (1948): 207–230.

Grice, Paul. "Logic and Conversation." In *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

Heim, Irene. "On the Projection Problem for Presuppositions." In *Proceedings of the West Coast Conference on Formal Linguistics: Volume 2*, edited by M. Barlow, D. P. Flickinger and M. T. Wescoat. Stanford, CA: Stanford Linguistics Association, 1983.

Heim, Irene. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. diss. University of Massachusetts, Amherst, 1982.

Kaplan, David. "Demonstratives." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press, 1989.

Karttunen, Lauri, and Stanley Peters. "Conventional Implicature." In *Syntax and Semantics 11: Presupposition*, edited by Choon-Kyu Oh and David Dinneen. New York: Academic Press, 1979.

Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.

Russell, Bertrand. "On Denoting." *Mind* 14 (1905): 479–493.

Soames, Scott. "Presupposition." In *Handbook of Philosophical Logic Volume IV*, edited by Dov Gabbay and Franz Guenthner. Dordrecht, Netherlands: Reidel, 1989.

Stalnaker, Robert. "Pragmatic Presuppositions." In *Semantics and Philosophy*, edited by Milton Munitz and Peter Unger. New York: New York University Press, 1974.

Stalnaker, Robert. "Presuppositions." *Journal of Philosophical Logic* 2 (1973): 447–457.

Strawson, Peter F. *Introduction to Logical Theory*. London: Methuen, 1952.

Strawson, Peter F. "On Referring." *Mind* 59 (1950): 320–344.

Scott Soames (2005)

ENTROPY

See *Philosophy of Statistical Mechanics*

ENVIRONMENTAL AESTHETICS

The term *environmental aesthetics* can apply to a variety of quite disparate sorts of cases—aesthetic appreciation of natural environments, of works of art situated in nature, of works of art—for example, landscape paintings—that are of or about nature, of works of art that take nature as their medium, and of gardens, a special category that seems to straddle the divide between culture and nature. In each case the philosophical challenge is the same: to determine the proper object and mode of appreciation. While these issues have not been definitively decided in the case of art appreciation, it remains helpful to use that example as a counterpoint against which an account of environmental appreciation can be constructed.

NATURE APPRECIATION

Nature scenes and natural items figure in our culture's most clichéd examples of aesthetic appreciation. Images of sunsets, rainbows, flowers, and baby animals are the stuff that enrich greeting card companies. But nature appreciation is also addressed by aestheticians and serious philosophers in the Western tradition. Immanuel Kant's examples of free beauty in *Critique of Judgment* (1790/1987) were natural items—flowers, birds, seashells. Beautiful items, Kant believed, provided a source of disinterested pleasure, their form alone triggering a pleasurable free play of imagination and understanding.

Nature appreciation is, of course, not confined to the beautiful. Kant's contemporary, Edmund Burke, indicates this even via the title of his 1757 work *A Philosophical Inquiry into the Origins of our Ideas of the Sublime and Beautiful* (1968). According to Burke, our attention is elicited not only by natural items that are small, lovely, and delicate, but also by those that are large, awesome, and terrifying. Surprisingly, such experiences are sought out. Kant concurs, offering the starry heavens, mountain peaks, and deep chasms as examples of the sublime. Certainly, nature is as much a repository of infinity and power as of delicacy and beauty.

Convinced that these two poles, the beautiful and the sublime, do not exhaust the grounds for aesthetic appreciation, eighteenth-century writers such as Sir Uvedale Price (1747–1829) and Richard Payne Knight (1750–1824) posited a third aesthetic category, the picturesque, situated midway between the beautiful and the sublime. Though the picturesque was initially defined as a species of beauty—that sort that would look pleasing in a picture—it soon came to be identified by an independently specified set of characteristics—roughness, sudden variation, and irregularity.

Additional factors of various sorts shape our nature preferences. Some are beliefs of which we are aware. Consider Thomas Burnet's (1635–1715) theory of the broken world. Promulgated in 1681 the theory impugned mountains as blemishes visited on the previously perfect

(smooth and spherical) earth in payment for humankind's Fall. In her classic study *Mountain Gloom and Mountain Glory* (1959/1997), Marjorie Hope Nicholson documented the changes that allowed Romantic poets to embrace mountain scenery. Less accessible instincts and emotions may also affect our attitudes toward nature. In the 1970s Jay Appleton formulated prospect-refuge theory according to which we all have a hard-wired preference for the savanna-type landscapes that afforded our long-ago ancestors crucially valuable opportunities to see yet not be seen. And in addition to such shared influences, we have each accumulated a vast store of personal experiences and associations that contribute to our landscape preferences.

CONTEMPORARY PHILOSOPHICAL DEBATES

Ronald Hepburn (1972–) is often credited with ushering in the current era of environmental aesthetics with his article "Contemporary Aesthetics and the Neglect of Natural Beauty" (1996/2004). Hepburn there pinpointed two crucial differences between the aesthetic appreciation of nature and the aesthetic appreciation of art: (1) The objects of nature appreciation are often unframed and unbounded, and (2) we are often immersed *in* those objects. Hepburn's rehabilitation of nature as an object of aesthetic appreciation has been welcome and effective. But it may be that in crafting his argument, he was focused on a particular subset of examples: macroscopic rather than microscopic objects of appreciation. We can savor entire panoramas: lofty mountain ranges, vast seascapes, all the sorts of scenes Donald Crawford (1938–) calls *postcardesque*, but we can also zoom in on tiny focussed delights: an alpine flower; a polished pebble; a single, wondrous insect. These are neither unbounded nor capable of immersing us. In addressing such objects we seem to adjust our focus at will; this may well counter the standard practice of the art world where conventional modes of appreciation are in place for each type of work.

Present-day philosophers have taken up Hepburn's challenge and examined the scope or proper objects of appreciation, its theory-ladenness, and the supporting roles of association, imagination, and emotion. Arnold Berleant's (1932–) 1991 theory of engagement proposes an approach to both nature and art in keeping with Hepburn's insights. Berleant emphasizes the participatory aspect of aesthetic experience, the reciprocity of perceiver and object in the aesthetic field. By contrast, Allen Carlson (2000, 2004) has built a distinctive theory of nature appreciation by rejecting at least part of the analogy between art appreciation and nature appreciation. Carlson argues that treating nature as a set of scenes or a collection of discrete but absorbing objects (e.g., the way we treat painting or sculpture) ignores just those hallmarks that were shown by Hepburn to set nature apart as unbounded and enveloping. Yet Carlson maintains that nature appreciation must be informed by some body of theory that plays the role that art theory and the history of art play in art appreciation. Carlson proposes that science fills this void in the case of nature appreciation. Thus geology, physics, astronomy, earth science, biology, and botany can all play a role in informing our appreciation of nature.

Just as it seemed that Hepburn's proposed hallmarks characterized some chunks of nature but not others (the macro rather than the micro), so, too, it seems that Carlson's theory works best for a certain subset of cases. Viewing the Rocky Mountains or the Grand Canyon, a Yellowstone geyser or a rampaging tornado, it seems that our appreciation can only be enhanced by knowing the forces that shape these expanses and events. Science here provides knowledge of origins. It is less clear that scientific knowledge is helpful in appreciating things that are small, or ordinary, or the sites of local, ongoing, yet invisible processes. Is my aesthetic appreciation of a forest path, of red maples in fall or of a spider's web glistening with dew enhanced by knowledge of the decomposition of leaf mold, of the loss of chlorophyll, or of the extrusion of spider silk?

Carlson's theory has generated a voluminous secondary literature. Among the challenges raised is the exact nature of the theories he urges appreciators to call upon—science only, or science mingled with common sense. Other critics challenge the exclusivity of Carlson's approach, suggesting that the appeal to scientific theory is one way to appreciate nature but that it can coexist with other ways. In this vein, Noel Carroll (1993) argues for the role of emotional responsiveness, insisting that it is often appropriate for people to be emotionally moved by natural scenes and events. Emily Brady (2003) argues for an expanded role for imaginative response and distinguishes four different kinds of imaginative activity—exploratory, projective, ampliative, revelatory—that are summoned up by nature. The first is a playful examination of form and its attendant associations, the second a deliberate exercise of seeing as, the third an inventive contextualizing that takes us beyond the perceptual image, and the fourth an arrival at aesthetic truth. These, too, seem compatible with the appeal to science—for exam-

ple, one might wander through a forest in spring, imagine it ablaze with fall reds and oranges, and acknowledge the mechanisms that would bring this about.

In her book *Aesthetics of the Natural Environment* (2003), Brady sorts various accounts of nature appreciation into cognitive and noncognitive camps. Since Brady basically elides the cognitive with the scientific, her taxonomy classes theories that appeal to association, imagination, emotion, or nonscientific information as noncognitivist. Thus she deems Hepburn, Berleant, and Carroll noncognitivists, along with Cheryl Foster (1998) who argues for an ineffable aspect of nature that she calls the ambient; Thomas Heyd (1956–), who champions the ascription of various narratives to natural goings on (the narrative is also how Foster labels the approach opposed to the ambient); and Yuriko Saito, who believes that nature appreciation should include a moral dimension—what she calls appreciating nature on its own terms.

The foregoing discussion has not touched on one profound, underlying problem, namely, the identification or definition of nature itself. There is good reason to think that there is no unsullied nature to be found on our planet. All nature has been intermixed with or affected by culture. Malcolm Budd (1941–) believes we are always able to abstract from such mixed cases and appreciate nature as nature even when, say, viewing an animal in a zoo (2002). The degree of mental/imaginative activity required here to arrive at an all-natural, intentional object of appreciation could be considerable. The water flowing from my kitchen faucet is natural only if I abstract away the changes wrung in the city treatment plant, or better yet, imaginatively travel back to the rainfall that was its source.

ART IN NATURE / ART FROM NATURE

This last topic of mixture lays the groundwork for considering cases where art and nature blend. The most innocuous in the continuum of such cases would be sculpture gardens and sculpture parks where works of art are simply arrayed in a natural setting. The effect would be very much that of an outdoor museum. Works of art in a sculpture garden can each be appreciated on their own. Additional insights arise from their juxtaposition.

While designers of a sculpture garden would of course take care to place each work in a setting conducive to its appreciation, there is no reason to think the arrangement could not be juggled just as curators can shuffle the order in which objects are arrayed in a museum. There is, however, one important feature that comes with outdoor exhibition. The works of art are viewed against an ever-changing background. Light and temperature are no longer controlled, as in a gallery, and the viewing experience is constantly affected by ongoing natural cycles: night and day, changing seasons, passing storms.

Additional complexities arise if the works of art exhibited outdoors connect with their setting in some way or other. One site might especially suit a given work on formal grounds. Alternatively, a work might comment on or interact with its setting. More explicit still are works that are *about* their settings. In such cases we enter the realm of site specificity, a trait made infamous by the controversy surrounding the removal of Richard Serra's (1930–) sculpture *Tilted Arc* from the site for which it was designed. Proper appreciation of site-specific works involves noting not only their formal, representational, and expressive properties but also their contextual properties. Thus the significance of Eero Saarinen's (1910–1961) *Gateway Arch* in St. Louis, Missouri, would be greatly compromised if it were not situated on the west bank of the Mississippi River marking the beginning of the western frontier brought about by the Louisiana Purchase. Nor would the work have the same significance if its legs were realigned to make the arch a portal for north–south rather than east–west travel. The environmental artist Robert Irwin (1928–) has codified four varieties of site specificity in his essay "Being and Circumstance." Irwin classifies works of art as site-dominant, site-adjusted, site-specific, and site-conditioned/determined. These four categories sort works whose meaning and purpose can be understood without reference to their site, works that make some concessions (such as placement and scale) to their site, works conceived with a specific site in mind, and finally, works that draw all their cues or reasons for being from their site.

A limiting case of the phenomenon of site specificity would be works of art that take (aspects of) their site as their medium. This would be true of some of the earthworks of the 1960s and 1970s. Michael Heizer's (1944–) *Double Negative*, Robert Smithson's (1938–1973) *Spiral Jetty*, and James Turrell's (1943–) *Roden Crater* are works that result from forceful gestures in the landscape; other environmental artists such as Andy Goldsworthy (1956–) and Michael Singer (1950–) make their art of more ephemeral stuff, taking walks and documenting them, making slight, nuanced adjustments to nature and then letting them dissipate. Both Crawford and Carlson have questioned whether the more bold types of environmental installations stand in an adversarial relation to nature as a result of creating aesthetic affronts.

GARDENS

When we turn to gardens, many of the topics already covered are still relevant. Gardens are in nature and their materials are often in large part natural. Japanese Zen gardens consisting of stones and raked sand are the most familiar counterexample to this expectation. And even more traditional gardens mix natural materials with a host of other components and features: paths, walls, benches, follies, fountains. Moreover, gardens bring to the forefront questions about *degrees* of naturalness. This has varied over garden history, with gardens that seemed utterly wild and untamed in one epoch coming later to be viewed as staid and artificial. Paradoxically, many gardens that are deemed natural in style achieve that effect through an intensive application of labor and care.

Unlike the sculpture parks and environmental works just discussed, the garden is a bona fide art form with its own history. Accordingly, Richard Wollheim's (1923–2003) notions of general and individual style take hold in gardening. Many garden styles are labeled in a way that includes a national designation—the French formal garden, the English landscape garden, the Italian villa garden, the Chinese scholar garden. This nomenclature flags the importance not only of cultural influences but also of topographical and climatological ones. (However, gardeners have always tried to trump nature with such aids as orangeries and greenhouses, trade in rare and exotic plants, breeding of entirely new species, and the bedding system—which allows several different gardens to succeed one another as the seasons unfold.) In addition to sustaining the notion of different general garden styles, the art of gardening also has practitioners whose individual style is recognizable. Thus we honor Andre Le Notre (1613–1700), Lancelot 'Capability' Brown (1716–1783), Gertrude Jekyll (1843–1932), Roberto Burle Marxe (1909–1994), and many, many more.

Garden appreciation must respond to this complexity. The sort of scientific knowledge that Carlson claims enhances our appreciation of natural scenes is also relevant to the garden—especially with regard to the plant species in place and the degree of skill or manipulation required to bring about various effects. Moreover, since all gardens are created rather than naturally occurring, their designers' intentions are always there to be retrieved. These intentions can range from trying to create a sensory delight to vastly ambitious promulgation of meanings. Not many gardens are what Mara Miller (1944–) calls grand gardens—that is, those that can claim to be great works of art. But exemplars have been produced in many different cultures. Gardens *can* convey complex meanings to those who view or walk through them. Through a judicious arrangement of plants, hardscape, topography, water features, statuary, buildings, inscriptions, and more, they can present disquisitions on matters of enduring interest and concern: politics, religion, love, war, the meaning of life, our place in the cosmos. Such gardens can sustain interpretive debates, with appreciators weighing in to defend alternative incompatible accounts of their meaning.

The sorts of garden content just discussed are pursued in much the same way that audiences track the meaning of works of art. Yet important aspects of nature appreciation also apply to gardens—especially the notions of unboundedness and surroundedness called to our attention by Hepburn. Though gardens are literally bounded, Miller (1993) has pointed out an important sense in which they cannot be controlled: since gardens are comprised of living things and are subject to natural forces, they are arenas of constant change. Plants grow; daily and seasonal cycles unfold; calamities occur. A garden designer's intentions are much less efficacious than those of other artists. The end result, as Miller puts it, is that gardens have no final form. There is no practical way to freeze a garden at a point in time and declare that to be the proper object of appreciation. In this regard, gardens truly do occupy a middle ground between nature and culture; wildness and art.

See also Burke, Edmund; Kant, Immanuel; Wollheim, Richard.

Bibliography

Appleton, Jay. *The Experience of Landscape*. London: John Wiley and Sons, 1975.

Berleant, Arnold. *Art and Engagement*. Philadelphia: Temple University Press, 1991.

Brady, Emily. *Aesthetics of the Natural Environment*. Tuscaloosa: University of Alabama Press, 2003.

Budd, Malcolm. *The Aesthetic Appreciation of Nature*. Oxford: Oxford University Press, 2002.

Burke, Edmund. *A Philosophical Inquiry into the Origins of Our Ideas of the Sublime and Beautiful*, edited by James T. Boulton. Notre Dame: University of Notre Dame Press, 1968.

Carlson, Allen. *Aesthetics and the Environment: The Appreciation of Nature, Art, and Architecture*. London: Routledge, 2000.

Carlson, Allen, and Arnold Berleant. *The Aesthetics of Natural Environments*. Ontario: Broadview Press, 2004.

Carroll, Noel. "On Being Moved by Nature: Between Religion and Natural History." In *Landscape, Natural Beauty, and the Arts*, edited by Salim Kemal and Ivan Gaskell. Cambridge, U.K.: Cambridge University Press, 1993.

Crawford, Donald. "Nature and Art: Some Dialectical Relationships." *Journal of Aesthetics and Art Criticism* 42 (1983): 49–58.

Foster, Cheryl. "The Narrative and the Ambient in Environmental Aesthetics." *Journal of Aesthetics and Art Criticism* 56 (1998): 127–137

Hepburn, Ronald. "Contemporary Aesthetics and the Neglect of Natural Beauty." In *British Analytical Philosophy*, edited by B. Williams and A. Montefiore. London: Routledge and Kegan Paul, 1996.

Kant, Immanuel. *Critique of Judgment*. Translated by Werner Pluhar. Indianapolis, IL: Hackett, 1987.

Miller, Mara. *The Garden as an Art*. Albany: State University of New York Press, 1993.

Nicholson, Margaret Hope. *Mountain Gloom and Mountain Glory: The Development of the Aesthetics of the Infinite*. Seattle: University of Washington Press, 1997.

Ross, Stephanie. *What Gardens Mean*. Chicago: University of Chicago Press, 1998.

Sonfist, Alan. *Art in the Land: A Critical Anthology of Environmental Art*. New York: E. P. Dutton, 1983.

Stephanie Ross (2005)

ENVIRONMENTAL ETHICS

Spurred by growing environmental concern in the 1960s, philosophers paid increasing attention to environmental ethics in the 1970s and 1980s. The field is dominated by dichotomies: anthropocentrism versus nonanthropocentrism, individualism versus holism, environmental ethics versus environmental philosophy, organic versus community metaphors, citizen versus consumer perspectives, scientific versus social scientific justifications, and tradeoffs versus synergism.

ANTHROPOCENTRIC ENVIRONMENTALISM

Traditional Western ethics is anthropocentric, as only human beings are considered of moral importance. Because people can help or harm one another indirectly through environmental impact, such as by generating pollution, destroying marshes, and depleting resources, environmental ethics can be pursued as a form of applied ethics in an anthropocentric framework.

Some anthropocentric issues concern the nature and relative importance of values. For example, does the beauty or inspirational quality of a canyon make the canyon intrinsically valuable? If so, is that value an objective feature of the canyon or an aspect of the evaluators' subjective experience or judgment? Finally, how important to people is such intrinsic, noneconomic value compared to economic considerations? Is the canyon's intrinsic value, assuming it has such value, sufficient to forgo its flooding to generate hydroelectricity that can power economic growth? Environmental ethics is a fertile testing ground for competing axiologies.

Environmental ethics also tests competing conceptions of the individual's relationship to society. A strictly economic approach views the individual as a consumer and directs government to regulate environmental matters to maximize the satisfaction of consumer demands. An alternative approach views the individual as a citizen concerned to promote individual excellence and to preserve and improve the community's best traditions and highest moral ideals. This dichotomy parallels that between liberalism and perfectionism in political philosophy. Just as many perfectionists would forgo the economic benefits of legal prostitution to protect the traditional family, many citizen-oriented environmentalists recommend preserving wilderness areas and species diversity to promote ideals of stewardship.

Environmental justice is primarily an anthropocentric ideal concerning the appropriate distribution of benefits and burdens among human beings affected by environmental decisions. Issues of resource depletion, nuclear waste, and population policy, for example, raise questions about intergenerational justice. Do future people have rights? Can a meaningful distinction be made between future people and possible people? Why should we care about future people if they can neither harm nor help us?

Issues of environmental justice arise when governments use cost-benefit analysis (CBA) to evaluate environmental policies. CBAs typically translate all values into monetary terms with the goal of identifying policies that maximize total social wealth. Exclusively monetary evaluations jeopardize future generations through the use of a discount rate that renders impacts 500 years from now insignificant. In addition, CBAs promote decisions that are unjust to poor people because the monetary value of items in a market economy, and therefore the total value of all such items, social wealth, depends on people's willingness to pay for things. Rich people can pay more than poor people, so the preferences of the rich are weighted more heavily in CBA than the preferences of the poor. Government policies guided by CBA therefore contravene the principle of justice that the interests of each individual person be considered equally.

Environmental racism concerns practices, in derogation of environmental justice, that subject people of non-European origin to disproportionate amounts of pollution and other negative side effects of economic

development. Within most industrial countries, such people are racial minorities. Internationally, such people reside in Third World countries. In both cases, many economists reject charges of environmental racism. They claim that people of color suffer disproportionate burdens not because of racist intent but because they are too poor to pay for better conditions. This is another area of tension between economic and noneconomic anthropocentric considerations.

MORAL EXTENSIONISM

Opposed to anthropocentrism are those who consider many nonhuman animals to be worthy of moral consideration in their own right. These views extend some traditional ethical theories, such as utilitarianism and neo-Kantianism, to include nonhuman individuals. Paul Taylor (2005) advocates further extension, according equal moral consideration to every living individual, amoeba included.

Many environmental philosophers consider moral extensionism too human-centered and individualistic. It is too human-centered because it justifies valuing nonhumans on the basis of similarities to human beings, such as sentience, consciousness, or merely life itself. Human traits remain the touchstone of all value. Moral extensionism is too individualistic for environmental ethics because some matters, such as species diversity, concern collectives, not individuals. From an individualist perspective, saving ten members of a common species is better than saving one member of an endangered species, other things being equal. Environmentalists concerned with maintaining species diversity reject individualism for this reason.

They reject individualism also as ecologically unrealistic. Ecology teaches that ecosystems depend on individuals eating and being eaten, killing and being killed. For example, predators must kill enough deer to avoid deer overpopulation, which would threaten flora on which deer feed. Reduced flora threatens soil stability and the land's ability to support life. So protecting individual deer from untimely death, which valuing deer as individuals may suggest, is environmentally harmful. Such harm threatens natural ecosystems, such as wilderness areas, that foster biological evolution, which is the focus of value for some nonanthropocentric environmentalists.

Tom Regan (2005) calls holistic views "environmental fascism." Sacrificing individuals for evolutionary advance or the collective good resembles Adolf Hitler's program, especially when human beings may be among those sacrificed. Human overpopulation threatens species diversity, ecosystemic complexity, and natural evolutionary processes, so consistent, nonanthropocentric environmental holism may be misanthropic.

Holists reply that human individuals, as well as environmental wholes and evolutionary processes, are intrinsically valuable, so individual humans should not be sacrificed to promote the corporate good. However, the casuistry of trade-offs among individuals and corporate entities of various species and kinds is not well developed by the holists. But Regan (2005), for his part, does not show how all individual nonhuman mammals, for example, can be accorded the equivalent of the human right to life without destroying wilderness areas and causing the extinction of carnivorous species.

Because they value not only nonhumans but holistic entities, many environmental philosophers believe their discipline calls for thorough review of the place of human beings in the cosmos. They reject the title "environmental ethics" in favor of "environmental philosophy" or "ecosophy" to emphasize that their views are not applications of traditional ethics to environmental problems but fundamental metaphysical orientations.

HOLISTIC ENVIRONMENTALISM

Holistic views tend to compare the environmental wholes they consider valuable in themselves with either communities or organisms. Aldo Leopold's "land ethic" (2005), for example, leans toward the community metaphor. Just as the benefits people derive from their human communities justify loyalty to the group, benefits derived from complex ecological interdependencies justify loyalty to ecosystemic wholes. J. Baird Callicott (2005) maintains that community loyalty is emotionally natural to humans, as our ancestors' survival during evolution depended on sentiments of solidarity. In this sense, ethics is based on Humean sentimentalism rather than on Kantian rationality or utilitarian calculation.

The Gaia hypothesis and deep ecology stress the similarity of holistic entities to individual organisms, thereby attempting to reconcile individualism with environmentalism. The Gaia hypothesis maintains that life on Earth operates as if it were a single organism reacting to altered conditions so as to preserve itself. This explains, for example, how Earth has maintained a relatively constant temperature over a 3-billion-year period while the energy emitted toward Earth from the Sun had increased 30 percent. Metaphorically, if Earth is alive, it is our mother because Earth's processes produce and sustain us. This metaphor justifies respect for Earth and Earth's processes analogous to respect for human mothers.

Deep ecology reflects the belief of the stoics and Benedict (Baruch) de Spinoza that reality is essentially one being. Accordingly, it questions the separateness of any individual from the environmental whole. Deep ecologists point out that the skin and other borders between individuals are really permeable membranes that connect as well as divide. Arne Naess (2005), deep ecology's founder, notes that human beings can possess the entire universe in their minds and suggests identifying one's real self with nature. Degrading nature is therefore unwise because it is a form of self-degradation. Deep ecology reconciles holistic environmental concern not only with individualism but also with individual self-concern. This metaphysical consideration is bolstered by observations about the lack of genuine fulfillment experienced by most people whose lives are dominated by consuming artifacts instead of appreciating nature.

ECOFEMINISM

Whereas the land ethic and Gaia hypothesis rely primarily on information drawn from science, other environmentalists stress social scientific information. Using the results of anthropological studies, especially of foraging (hunter-gatherer) societies, some environmentalists maintain that human life is better where people do not attempt to master nature in the human interest. Many indigenous societies practice an environmental ethic, similar to the land ethic, of reciprocal exchange with nonhuman environmental constituents such as water, sun, trees, and game animals. This enriches human life and preserves the environment.

Ecofeminists emphasize the relationship between mastering nature in the supposed human interest and the oppression of women, indigenous people, and other subordinated groups. Ecofeminists claim that much Western thinking is dominated by what they call "the master mentality," which is dualistic thinking that values one member of each dyad more than the other and relegates the inferior member to serve the superior. Such dyads include men versus women, heaven versus earth, mind versus body, reason versus emotion, culture versus nature, and progress versus stagnation. White Western men are associated with the superior member of each dyad: heaven, mind, reason, culture, and progress; whereas women, indigenous people, and other subordinated groups are associated with the inferior member: earth, body, emotion, nature, and stagnation. The master mentality thus justifies the continued subordination of women and non-Western, nonindustrialized men, because humanity in general flourishes, the masters claim, when the inferior serves the superior.

The master mentality's association of women with earth and emotion explains traditional exclusions of women from high religious offices and from professions emphasizing the use of abstract reason. The association of progress with economic growth explains Western insensitivity to the disruption of traditional patterns of life in many Third World countries. Traditional agriculture returns little money but produces a large variety of food for local consumption. It suffers in comparison with Western-inspired commercial agriculture that emphasizes remunerative monocultures when progress is associated with economic growth. The master mentality also generally undervalues work traditionally done by women because much of it is done free, whether it is childcare or tending a garden to feed the family.

Ecofeminists claim that the master mentality approach to the environment serves humanity badly. Worldwide, it marginalizes and impoverishes women and other subordinated people. Humanity would fare better if people's interactions with nature were guided more by thinking traditionally associated with women. Women tend to think more relationally, organically, and holistically than (Western) men, who favor individual rights, commercial success, and mechanistic processes. Whereas typical male patterns of thought and action precipitate ecocrises, typical female patterns ameliorate them. Empowering women can save ecosystems and species diversity.

Most anthropocentrists and nonanthropocentrists believe that in general a tension exists between protecting nature and serving humanity. Nonanthropocentric concern for nature as valuable in itself precludes actions that can make human life better, they think. This is the tradeoff perspective. The synergistic perspective, by contrast, is that valuing nature for itself most often precludes action that is mistakenly undertaken in pursuit of human welfare but that is actually counterproductive. For example, the Green Revolution attempted to improve human life by mastering nature but it harmed people more than it helped them because it disrupted traditional social systems and more holistically productive agriculture, argues Vandana Shiva (1991). Like the hedonic paradox, which claims that happiness is best achieved by not seeking it directly, synergists claim that human flourishing is best achieved by nonanthropocentrically valuing nature for itself rather than by trying anthropocentrically to maximize returns from nature for human beings. The land

ethic, deep ecology, and ecofeminism are compatible with environmental synergism.

Because environmental ethics/philosophy questions basic assumptions in economics, technology, metaphysics, ethical theory, moral epistemology, and gender relations, it approaches religion in its attention to the fundamental concerns of human existence.

See also Animal Rights and Welfare; Applied Ethics; Distant Peoples and Future Generations; Good, The; Intrinsic Value; Rights; Utilitarianism.

Bibliography

Callicott, J. Baird. "The Conceptual Foundations of the Land Ethic." In *Environmental Ethics: Readings in Theory and Application*. 4th ed, edited by Louis P. Pojman, 149–160. Belmont, CA: Thomson/Wadsworth, 2005.

Dobson, Andrew. *Fairness and Futurity: Essays on Environmental Sustainability and Social Justice*. New York: Oxford University Press, 1999.

Hargrove, Eugene C., ed. *The Animal Rights, Environmental Ethics Debate: The Environmental Perspective*. Albany: SUNY Press, 1992.

Leopold, Aldo. "Ecocentrism: The Land Ethic." In *Environmental Ethics: Readings in Theory and Application*. 4th ed, edited by Louis P. Pojman, 139–148. Belmont, CA: Thomson/Wadsworth, 2005.

Naess, Arne. "Ecosophy T: Deep versus Shallow Ecology." In *Environmental Ethics: Readings in Theory and Application*. 4th ed, edited by Louis P. Pojman, 192–200. Belmont, CA: Thomson/Wadsworth, 2005.

Partridge, Ernest, ed. *Responsibilities to Future Generations: Environmental Ethics*. Buffalo, NY: Prometheus Books, 1981.

Pojman, Louis P., ed. *Environmental Ethics: Readings in Theory and Application*. 4th ed. Belmont, CA: Thomson/Wadsworth, 2005. This is an excellent source of information on anthropocentric environmentalism, animal rights, biocentrism, the land ethic, deep ecology, social ecology, and ecofeminism.

Regan, Tom. "The Radical Egalitarian Case for Animal Rights." In *Environmental Ethics: Readings in Theory and Application*. 4th ed, edited by Louis P. Pojman, 65–72. Belmont, CA: Thomson/Wadsworth, 2005.

Sagoff, Mark. *The Economy of the Earth: Philosophy, Law, and the Environment*. New York: Cambridge University Press, 1988. Critiques economics and CBA.

Shiva, Vandana. *The Violence of the Green Revolution: Third World Agriculture, Ecology and Politics*. London: Zed Books, 1991.

Taylor, Paul. "Biocentric Egalitarianism." In *Environmental Ethics: Readings in Theory and Application*. 4th ed, edited by Louis P. Pojman, 117–131. Belmont, CA: Thomson/Wadsworth, 2005.

Warren, Karen J. *Ecofeminist Philosophy: A Western Perspective on What It Is and Why It Matters*. Lanham, MD: Rowman & Littlefield, 2000. A comprehensive introduction to ecofeminism.

Wenz, Peter S. *Environmental Ethics Today*. New York: Oxford University Press, 2001.

Peter S. Wenz (1996, 2005)

EPICTETUS

(55 CE–c. 135)

Epictetus became a slave of Epaphroditus, himself a freedman who was secretary to Nero. After being freed by his master, Epictetus studied with the Stoic Musonius Rufus, and he taught in Rome until Domitian banished the philosophers in 89 CE. He then established a school in Nicopolis in Epirus, a town in northwest Greece founded by Octavian to commemorate his victory at Actium. Epictetus was lame, perhaps because of his sufferings as a slave, but was a renowned teacher.

Like Socrates, Epictetus wrote nothing, but his pupil Arrian compiled a record of his oral teachings. Four books of these *Discourses* survive, together with a digest of central points known as the *Manual*. Although these works reveal that Epictetus taught his students through the careful study of Stoic doctrine (II 13.21, III 16.9–10), they also make it plain that the goal of philosophical learning is not to be an exegete of Chrysippus (I 4.6–9, I 17.13–18). In fact, the discourses themselves do not offer much exegesis, nor do they often develop heavily theorized explanations, careful distinctions, or involved arguments, as a usual philosophical treatise would. For years, scholars explained this by the hypothesis that Epictetus's teachings fit a particular genre of "diatribes"—the Greek title of the *Discourses* is *Diatribai*—but, because the evidence of such a genre is very thin, the hypothesis is no longer widely accepted.

It now seems, instead, that the discourses simply teach in the ways that Epictetus associates with his three philosophical heroes: They manifest the examining role of Socrates when Epictetus refutes unnamed interlocutors by question-and-answer; they display the rebuking and kingly role of Diogenes the Cynic when Epictetus hectors pupils and exemplifies haughty leadership; and they take on the teaching, doctrinal role of the founding Stoic, Zeno of Citium, when Epictetus offers terse, straightforward principles for the guidance of life (III 21.18–19). In all of these ways, the *Discourses* and *Manual* are concerned primarily with the concrete task of helping others live better lives.

So, when Epictetus outlines the "three topics" on which a person must train to become good, they are not abstruse philosophical matters or even the broad parts of

Stoicism, logic, ethics, and physics. Rather, he insists that one must study, first, desires and aversions; second, impulses, rejection, and in general, appropriate action; and third, infallibility in assent (III 2.1–2). He counsels that the first step is to extinguish desire—to rein in the passions and to free one to do what is appropriate (III 2.3–4, *Manual* 2). Epictetus here supposes the Stoic views that passions are defective judgments about what is good and bad for one, and that desire is an impulse for what is good. To eliminate desire, then, is to free oneself from making so many judgments of what is good and bad for one. This freedom from passionate attachments, in turn, frees one to consider coolly what is merely appropriate and act accordingly. Without passions and desires, one lives by weaker impulses, in terms of what is merely appropriate. The last topic is reserved for those who have already made substantial progress in taming their desire and managing their impulses (III 2.5).

This focused deployment of Stoic ideals without the full discussion of logic and physics recalls the Cynics, who traditionally offer the Stoic a "shortcut to virtue." Epictetus does in fact endorse a brand of Cynicism (III 22), and his Stoicism is much more austere than that of, say, Cicero's *On Duties*. Nevertheless, Epictetus is not hostile to all conventional roles and the activities appropriate to them, and he does not reject logic and physics so much as he keeps the focus away from them in the *Discourses*, to keep his pupils concentrated on bettering themselves.

Accordingly, the special features of Epictetus's Stoicism serve his practical aim of helping people, and most are probably due more to it than to any doctrinal disagreement with other Stoics. Among these features, perhaps the most prominent is the oft-repeated distinction between what is up to us and what is not. This distinction, which is highlighted in the first sentence of the *Manual*, tells one to care only for one's mind or soul. Often, Epictetus puts this by saying that our volition (*prohairesis*) is up to us (see, for example, I 1.23). Because the word *prohairesis* is common in Aristotle's ethics but not among early Stoics, who used it to pick out a limited sort of impulse, some scholars see Epictetus's concentrated concern for *prohairesis* as an especially innovative suggestion of a will, inspired by Aristotle and perhaps by debates about freedom and determinism. But this interpretation is hard to support, for the resources Epictetus uses to explain what he means by *prohairesis* do not much stretch the boundaries of earlier Stoicism, and the freedom that he connects with *prohairesis* is just the moral freedom familiar from earlier Stoicism.

Another special feature of Epictetus's Stoicism is its intensely personal theology. Stoics always locate divinity in the cosmos; they attribute the orderly workings of nature to the divine reason in which all humans have a share. Epictetus personalizes all of this. He considers the goal of living to be to follow the god or gods (I 12.5, I 30.4), and he considers himself a servant of god (IV 7.20). Moreover, he refuses the picture of servitude to a distant king: on Epictetus's view, Zeus has stationed a divinity within each of us, a "god within" (I 14.12–14, II 8.12–14).

According to this thought, which clearly reinforces the emphasis on what is up to us, we all have the resources we need to live well within us. Two additional ways in which Epictetus develops his appeals to our inner resources are among the most innovative features of his Stoicism. For one, he regularly insists that we have capacities of trustworthiness and self-respect that cannot be taken from us (see, for example, I 25.4). This appeal to personal integrity and to an ability to evaluate reflectively what is appropriate to oneself suggests the modern notion of conscience, and it clearly invokes a notion of self-respect (*aidôs*) that is distinct from what is attested for earlier Stoics. Epictetus makes another interesting departure from earlier Stoics when he insists that our notions of good and bad and the like are innate (II 11.1–8). Although earlier Stoics insist that human minds are blank slates at birth, Epictetus encourages us to take heart in our substantial inheritance from the gods.

Epictetus's Stoicism is fully realized for the purpose of encouraging others to progress as Stoics. His articulation of self-reliance has attracted many readers over the centuries, and his subtle moral psychology has deservedly found a wide audience, from the second-century emperor Marcus Aurelius to the sixth-century Neoplatonist Simplicius (who wrote a massive commentary on the *Manual*) and from the sixteenth-century neo-Stoic Justus Lipsius to the twentieth-century American prisoner of war James Stockdale.

See also Cynics; Diogenes of Sinope; Marcus Aurelius Antoninus; Simplicius; Socrates; Stoicism; Zeno of Citium.

Bibliography

WORKS BY EPICTETUS

Dissertationes ab Arriano Digestae, edited by H. Schenkl. 2nd ed. Stuttgart: Teubner, 1916.

WORKS BY EPICTETUS IN TRANSLATION

The Discourses, as Reported by Arrian, the Manual, and Fragments. 2 vols., edited and translated by W. A. Oldfather. Cambridge, MA: Harvard University Press, 1925–1928.

WORKS BY EPICTETUS IN PARTIAL TRANSLATION WITH COMMENTARY

Discourses: Book One. Translated, with an introduction and commentary, by R. F. Dobbin. Oxford: Clarendon, 1998.

WORKS ABOUT EPICTETUS

Bönhoffer, A. *Epictet und die Stoa*. Stuttgart: F. Enke, 1890.

The Ethics of the Stoic Epictetus. Translated by W. O. Stephens. New York: Peter Lang, 1996.

Long, A. A. *Epictetus: A Stoic and Socratic Guide to Life*. Oxford: Clarendon, 2002.

Eric Brown (2005)

EPICUREANISM AND THE EPICUREAN SCHOOL

The Epicureans perpetuated their founder's teaching with little change. Of Epicurus's immediate circle, the most distinguished was Metrodorus of Lampsacus (c. 330–277 BCE), who predeceased his master. Metrodorus was elevated by Epicurus to a position of eminence—he alone shared the appellation "wise" (*sophos*), and his works were regarded as authoritative statements of doctrine. He wrote on epistemology, ethics, religion, poetry, and rhetoric, and he composed polemics against Plato's *Gorgias* and *Euthyphro,* and against Democritus.

Colotes of Lampsacus, another member of the original circle, published a comprehensive refutation of other schools under the title "That the Doctrines of the Other Philosophers Actually Make Life Impossible." Our knowledge of it comes from Plutarch's *Reply to Colotes.* His other writings included attacks on Plato's *Lysis* and *Euthydemus* and on the myth of Er in the *Republic.*

Hermarchus of Mytilene (325–c. 250 BCE) was Epicurus's successor as head of the school. His chief work, in twenty-two books, was on Empedocles. He also wrote on the arts (including rhetoric), attacked Plato and Aristotle, and left a collection of letters.

Polystratus succeeded Hermarchus. Two of his works have been recovered in part from the library at Herculaneum; the better preserved is "On Unreasonable Contempt for Popular Opinion."

In the second and first centuries BCE the school continued to flourish. One member, Philonides, enjoyed the friendship of Antiochus Epiphanes (king of Syria, 175–164 BCE) and attained some standing as a mathematician. Later in the second century Zeno of Sidon lectured in Athens on logic, rhetoric, poetry, and mathematics; and he introduced into his ethical teaching many of the commonplaces of the popular moral essays developed by rival schools, including the use of moral examples drawn from literary sources. Zeno's older contemporary, Demetrius of Laconia, also composed popular moral essays and wrote on logic and poetics. These men's rivals were chiefly Stoics, and under the pressure of controversy they occasionally gave new formulations to Epicurean teaching. Whether they were concerned to any great extent with the atomic theory is uncertain; it appears that they did align themselves more closely than did their predecessors with the traditions of Greek *paideia,* perhaps for the added prestige it gave them as they spread their doctrine to the east and west. (Both Zeno and Demetrius counted Romans among their students.) Yet there were a few diehards; one of the rare schisms in the school developed over the question of whether rhetoric is an art. The use of literary embellishment as a means of persuasion was contrary to the principles of the strict Epicureans, who seem to have been influenced by Plato's *Gorgias.* Yet one group accepted epideictic oratory as a legitimate pursuit.

In the first half of the first century BCE Philodemus of Gadara (in Syria), who had attended the lectures of Zeno in Athens, founded at Naples an Epicurean group with liberal tendencies. The Epicurean library at Herculaneum has yielded extensive passages from his many writings, which included moral treatises, biographies of philosophers, a history of the philosophical schools, and such polemical works as "On the Gods" and *On Methods of Inference.* Among his followers were persons of considerable eminence, notably Piso Caesoninus, Roman consul in 58 BCE, who was his principal patron. To such as these, we may suppose, he addressed "On Wealth," "On the Management of Property," and "On the Good King in Homer." This last piece is remarkable, not so much for its concern with a political matter (Epicurus had written "On Kingship") as for its use of Homer as an authority. (The Epicureans' rejection of the traditional Greek education led them to minimize the importance of the Homeric poems and to challenge the wisdom of Homer.)

Philodemus's treatises "On Rhetoric," "On Poems," and "On Music" were orthodox to the extent of maintaining that these arts are not suitable media for philosophical teaching or moral training; yet Philodemus conceded to them a positive value as art forms. Indeed, he himself had literary pretensions and composed a number of short poems. As a philosopher he won a qualified

respect from Cicero. With Siro, his colleague, he attracted to the school a group of young Latin poets, among them Vergil; there is, however, no evidence to connect the school at Naples with the Roman Epicurean Lucretius.

Under the empire the Epicureans, true to their own precept, withdrew from public view. The last conspicuous member of the school was Diogenes of Oenoanda (a town in Lycia), who about 200 CE published the wisdom of Epicurus for his fellow townsmen by having a number of Epicurean writings inscribed on a wall at the entrance to the town. Most of the texts, apparently, he composed himself; two are on natural science, the remainder on ethics. He also included some of Epicurus's sayings and a letter from Epicurus to his mother.

See also Aristotle; Empedocles; Epicurus; Homer; Leucippus and Democritus; Lucretius; Philodemus; Plato; Stoicism.

Bibliography

Our knowledge of the Epicurean school is being improved by the constant publication and reediting of the Herculaneum papyri. Most of this material can be found, with good introductions and commentaries, in the series *La Scuola di Epicuro* (Bibliopolis, Naples) and also in the journal *Cronache Ercolanesi* (Naples, 1971–). For a history of the study of the papyri see Mario Capasso, *Manuale di Papirologia Ercolanese* (Galatina: Congedo, 1991). Diskin Clay, *Paradosis and Survival: Three Chapters in the History of Epicurean Philosophy* (Ann Arbor: University of Michigan Press, 1998), contains essays on a range of topics from different periods of Epicurean philosophy.

On Metrodorus see Alfred Koerte, "Metrodori Epicurei Fragmenta," in *Jahrbücher für classiche Philologie,* Supplementband 17 (1890): 531–570. (There is a 1987 facsimile reprint by Garland, New York, which also contains Vincenzo de Falco, *L'Epicureo Demetrio Lacone* [Naples, 1923].) For Hermarchus see: Francesca Longo Auricchio, *Ermarco: frammenti* (Naples, 1988).

For Polystratus see: Giovanni Indelli, *Polistrato: Sul disprezzo irrazionale delle opinioni populari* (Naples: Bibliopolis, 1971), and James Warren, *Epicurus and Democritean Ethics: An Archaeology of Ataraxia* (Cambridge, U.K., and New York: Cambridge University Press 2002), Chapter 5. On Demetrius of Laconia: de Falco (see above); Enzo Puglia, *Demetrio Lacone: Aporie testuale ed esegetiche in Epicuro* (Naples, 1988); Costantina Romeo, *Demetrio Lacone: La Poesia* (Naples: Bibliopolis, 1988). There is a commentary on Plutarch's reply to Colotes (*Adversus Colotem*) by Rolf Westman, *Plutarch gegen Kolotes* (Helsinki: Acta Philosophica Femica, 1955). See also Wilhelm Crönert, *Kolotes und Menedemos* (Leipzig, 1906).

The best introduction to Philodemus is: Marcello Gigante, *Philodemus in Italy* (Ann Arbor: University of Michigan Press, 1995). There are also useful discussions in recent editions of Philodemus's works: Dirk Obbink, *Philodemus: On Piety, Part I* (Oxford: Clarendon Press, 1996) and Richard Janko, *Philodemus: On Poems Book 1* (Oxford, 2000). There are useful articles in Dirk Obbink, ed. *Philodemus and Poetry* (Oxford: Clarendon Press, 1995) and Voula Tsouna, "Philodemus on the Therapy of Vice," *Oxford Studies in Ancient Philosophy* 21 (2001): 233–258, is a good introduction to his ethical approach.

For Lucretius: Diskin Clay, *Lucretius and Epicurus* (Ithaca, NY: Cornell University Press, 1983) and David Sedley, *Lucretius and the Transformation of Greek Wisdom* (Cambridge, U.K., and New York: Cambridge University Press, 1998) discuss the Latin poet's use of early Epicurean works. See also: Keimpe Algra et al., eds., *Lucretius and His Intellectual Background* (Amsterdam and New York: North-Holland, 1997).

For Diogenes of Oinoanda see the edition with introduction and commentary by Martin F. Smith, *Diogenes of Oinoanda: The Epicurean Inscription* (Naples, 1996; supplementary volume, 2003).

P. H. De Lacy (1967)
Bibliography updated by James Warren (2005)

EPICURUS

(341–270 BCE)

Epicurus was born on Samos to parents who were Athenian citizens. Evidence about his philosophical debts and development must be sifted from conflicting reports arising out of the agonistic context of ancient Greek philosophical rivalry and invective. While rivals charge him with merely plagiarizing his atomism from Democritus and hedonism from the Cyrenaics, his advocates praise his singular originality, probably encouraged in this by Epicurus himself. Like Parmenides, René Descartes, and other seminal figures in philosophy, Epicurus presented himself as a solitary herald of truth, creating his system *de novo* because of the inadequacy of his predecessors and teachers. Modern scholarship tends to split the difference, seeing a variety of possible influences—Democritean atomism, Cyrenaic hedonism, Aristotelian *eudaimonism*, skeptic impeturbablilty—while fully recognizing that however much Epicurus worked within an existing framework, he is responsible for a succession of remarkably influential innovations.

LIFE AND SOURCES

Epicurus spent most of his first thirty-five years in Asia Minor. There he began formulating his doctrines and collecting important adherents before embarking for Athens where he founded the Garden in 306 BCE. Alone among the major Athenian meeting places for philosophers (the Academy, Lyceum, and Stoa), it remained an authoritative center for the study and dissemination of its

founder's teachings late into the Roman period—largely, no doubt, because it alone continued to flourish as an institution with endowed property, stable traditions of teaching and doctrine, and generations of faithful advocates. It was distinguished as well by its admission of women and slaves. Epicurus's death, although physically painful, is portrayed in our sources as having been appropriately philosophical.

Diogenes Laertius (third century CE), our chief source for his writings (including his will), relates that Epicurus was the most prolific author of his time (some 300 papyrus rolls). Pitifully little survives. Diogenes himself preserves three short letters outlining Epicurus's physical theory, ethics, and explanations of celestial phenomena, though doubts exist that the last is from Epicurus's hand. *Kuriai Doxai*, a collection of excerpts quoted by Diogenes, and a parallel collection surviving in another manuscript, *Sententiae Vaticanae*, were apparently designed to remind adherents of Epicurus's key claims.

Although critical, the philosophical treatises of Cicero, written some two centuries after the time of Epicurus, offer our most articulate evidence for many Epicurean arguments. Other scattered citations are preserved, especially in Plutarch, Sextus Empiricus, Seneca, and the Aristotelian commentators, though it often is difficult to discern from them the original context and intent of his arguments. *De Rerum Natura*, by the Roman poet Lucretius (d. 50 BCE), renders into verse Epicurean atomic theory, epistemology, and social thought, relying especially on Epicurus's major treatise, *On Nature*. Badly damaged parts of *On Nature* and several works by Philodemus, an Epicurean roughly contemporary with Lucretius, were recovered in Herculaneum (1752) from the Casa dei Papiri, buried by the eruption of Vesuvius in 79 CE. New methods of reconstructing these texts are yielding important information about many facets of Epicureanism. Finally, in Oenoanda, in what is present-day southwestern Turkey, a large inscription erected by one Diogenes encapsulates several basic Epicurean doctrines.

PHILOSOPHICAL SYSTEM

Although not as insistent as the Stoics about the systematic coherence of all his philosophical doctrines, Epicurus believes that his arguments from the domains of physics, epistemology, psychology, and ethics are mutually supporting. So, for instance, a linchpin in his arguments against the fear of death is the claim that persons are material entities of the sort that no longer continue to exist upon death and are therefore no longer subject to harm. This being the case, we have no reason to fear a future state that can cause us no harm.

Such a view of persons similarly undergirds his theological claim that we have no reason to fear punishment from the gods in an afterlife. Since we do not survive our deaths, the gods can hardly mete out any post mortem punishment, even if they so wished. At the same time, the relation of his materialism to many of his other central doctrines is less immediately straightforward. Epicurus sometimes describes perceptual experiences, actions, and psychological properties in ways that, to many at least, do not look easily reducible to talk about atoms and the void, and one of the persistent problems in understanding his philosophical thought is gauging the extent to which he offers, or at least intends to offer, fully reductive material accounts of each of the primary domains of his philosophical system.

Gaps in our evidence, at least at present, preclude giving a precise accounting of his philosophical successes in coping with the demands of materialism within his general thinking. But one thing seems clear. By adopting a materialist physical theory and working through its implications, Epicurus formulates a series of questions about the material bases of perception and thought, the mechanisms of choice and avoidance, and the possibility of free agency in a world consisting of matter in motion that sets him on a path distinct from such predecessors as Plato and Aristotle. Moreover, he rejects their *polis* bound conceptions of ethics and politics and offers accounts of ethical motivation and political obligation strictly rooted in notions of individual agents and their mutual relations per se. In so doing, Epicurus and his followers develop both a professional philosophical vocabulary and ways of conceiving a broad range of philosophical issues that often appear distinctly modern. Indeed, it might be more historically precise to say that many modern ways of formulating arguments can strike one as being Epicurean—no doubt because a significant number of them in fact have their origins in ancient Epicureanism.

The philosophical challenges that Epicurus faced because of the convergence of his atomism, empiricism, hedonism, and politics of solitary individualism provided a basic conceptual framework for a whole range of thinkers who helped set the terms of modern philosophical debates. Michel de Montaigne, Thomas Hobbes, Pierre Gassendi, Robert Boyle, Jean-Jacques Rousseau, Baron d'Holbach, Jeremy Bentham, and John Stuart Mill—to name just a few who particularly felt his influence—all show basic debts across a wide variety of their

theoretical concerns—debts not only to more general Epicurean philosophical preoccupations, but also in many instances to Epicurus's particular methods of argument, his specific conclusions, and, above all, to the way that he originally devised the philosophical framework and individual terms in which they came to carry on their own debates.

Epicurus himself is frank about the role of materialism in his system. He asserts that if we were not made unhappy by the fear of death or suspicions that natural phenomena depend upon meddling gods, we would have no need to inquire into nature. He therefore rejects both Aristotle's defense of purely theoretical inquiry and the Socratic claim that ethical beliefs can be examined independently of a complete understanding of nature. But this does not commit him, of course, to a rigidly pragmatic conception of scientific truth. To the contrary, he insists that just as with respect to our health, what we want is not merely the semblance of health, in our scientific theories what we want is truth, not merely some beliefs that may appear true. Extra-scientific concerns may motivate our inquiries into nature, but that does not mean that our use of evidence or our procedures themselves need be compromised.

THE FEAR OF DEATH

Given, however, that Epicurus maintains that philosophical inquiry is driven by our need to understand the sources of unhappiness, it might be helpful to turn to his analysis of these, chief among them the fear of death. In contrast to most contemporary philosophers, Epicurus argues that we have no reason to fear death because it cannot harm us in the slightest. That said, however, he thinks that most people at some level do indeed fear death and that their mental lives, along with many of their actions, are affected in unhappy ways because of what turns out to be simply an irrational and readily eliminable misconception. So for instance, we find in Epicurean texts a version of the so-called symmetry argument.

Assume for the moment the Epicurean claim that death annihilates us. Before we were conceived, we also did not exist. Yet we typically are troubled only by our post mortem nonexistence. How are we to explain this asymmetry in our attitudes to two apparently similar states of our nonexistence? Epicurus argues that whenever we think about our future nonexistence, we find it difficult to view it as the total annihilation of our consciousness and to eliminate ourselves from our conception of it in the required way. This is because whenever I try to conceive of my death, I become a kind of conscious eyewitness to it in my imagination. I thus have the illusory experience of witnessing my ongoing annihilation, with death continually depriving me of things I value. When I look back at the time before I was conceived, however, I readily see it as the state of nothingness that it is. However explicable, holding asymmetrical attitudes to two equivalent states of our nonexistence, Epicureans argue, is irrational and we should come to view both death and the time before our conception with equal indifference.

This argument, no doubt, raises quandaries about our attitudes towards our past and future, questions about whether such general temporal attitudes apply to states that we do not consciously experience, and dilemmas about the contributions of past or future potential losses in fixing the identities of persons. It also is liable to backfire. We might instead begin viewing the time before our conception as another regrettable state of lost potentialities and thus merely duplicate our anxieties.

Epicurus's central argument against the fear of death, however, is an attempt to demonstrate that all such worries about states of our nonexistence are irrational. His claim is best illustrated by his deceptively simple observation, "When we are there, death is not, and when death is there, we are not" (*Epistula ad Menoeceum* 125). Anyone wishing to show that a state of nonexistence harms us, Epicurus insists, must show who is harmed, when the harm occurs, and how one is harmed. Although these questions make sense regarding harms to existent subjects, no meaningful conception of harm, he argues, can be applied to the nonexistent. Who is harmed by death? No one, since in death there is no longer any subject to be harmed. When are we harmed by death? At no time, since when we are alive, death is not there, and when we are dead, we are not there. How are we harmed by death? In no way, since harm, whether conceived of as deprivation or lost potential, can attach only to something that exists.

The Epicurean claim that any conception of the harm of death requires there to be an existent subject of that harm has challenged philosophers to explore and clarify the metaphysical status of the dead, the place of potential losses or deprivations in accounts of personal identity, the nature of counterfactual propositions about future and past persons, and the conception of time needed to justify the intuition that death harms us. However simple at first glance, the increasingly sophisticated literature generated in response to Epicurus's argument suggests that the verdict is hardly in.

HEDONISM

One might wonder, however, as did many of Epicurus's ancient critics, how his arguments about the harmlessness of death are compatible with another of his central claims, that pleasure is our final end. Shouldn't a hedonist fear the interruption and loss of pleasure that death threatens? To understand the particular nature of Epicurean hedonism it first must be placed in its ancient dialectical context. Epicurus argues that pleasure is our ultimate goal and the sole component of our *eudaimonia*, or happiness. He further supports his ethical hedonism with a version of psychological hedonism by appealing to a so-called cradle argument; that is, that the observation of infant behavior shows that we naturally seek pleasure and avoid pain. More surprisingly, however, Epicurus argues that the pleasures that comprise individual happiness can be specified objectively because they meet nonsubjective criteria and arise from pursuits limited by objective, natural constraints.

By way of contrast, for instance, take John Locke's account of the patient with sore eyes who prefers the pleasures of drink to those of sight and who remains the sole arbiter of his own pleasures, however self-destructive. (*Essay* 2.xxi.55). For the Epicurean, such subjectivity about pleasure and the good fails to meet the minimal demands placed on *eudaimonia* in ancient ethical arguments. So, for instance, desires are open to objective assessment, he claims, because they fall into three distinct classes. Some, for example desires for immortality or power, are incapable of being satisfied and depend on erroneous beliefs, many of which have been inculcated by society. They are therefore both unnatural and unnecessary. Moreover, since they have no natural limits and attempts to satisfy them inevitably lead to frustration, they are to be rejected as sources of unhappiness. Other desires, such as for sex or for particular types of pleasant food, are natural but unnecessary. They can be satisfied if the opportunity arises, but they are not necessary for happiness. Indeed, they can become sources of unhappy disturbance and pain if one becomes troubled by their unavailability or loss.

Finally, there are desires that are both natural and necessary. These have objective, natural limits and are easily satisfied. One needs only so much bread to satisfy one's natural and necessary desire for food, and that desire, unlike those for power or immortality, has a natural limit. By focusing on the satisfaction of natural and necessary desires and by adapting our desires to our circumstances, we can avoid the frustrations of pursuing pleasures that prove "empty." What we can hope to achieve instead is a natural state of satisfaction entirely free from both mental disturbance and bodily pain. Indeed, such a seemingly neutral condition, Epicurus insists, is the most pleasant state possible. Many have found this claim about pleasure to be paradoxical, arguing that such a state is merely intermediate between pleasure and pain. He denies, however, the existence of any intermediate states, maintaining that the lack of pain and disturbance is the highest pleasure possible.

It is easy to view this, perhaps, as a mere verbal ploy since it flies in the face of what most hedonists and nonhedonists alike have found salient about pleasure, its intensity and variability as a sensation. Epicurus, however, bolsters his argument by distinguishing *katastemmatic* from *kinetic* pleasures. Although its exact force is disputed, as is the extent to which it picks out two different types of atomic movement, this distinction seems to capture two readily identifiable, though perhaps not so readily separable, aspects of pleasure. Kinetic pleasures arise, he argues, during the process of satisfying a desire. Katastemmatic pleasures are the states of satisfaction that supervene when a desire has been satisfied. Most accounts of pleasure, Epicurus claims, make the mistake of focusing only on the pleasures of satisfying desires, when in fact the different ways one satisfies desires are simply interchangeable variants. What we value above all is the attendant state of satisfaction and the freedom from want or disturbance that it signals.

To the charge that he offers us the pleasures of a corpse, the Epicurean responds that the wise prefer no longer having an itch to the pleasures of scratching. More controversial still, Epicurus argues that the highest state of pleasure is itself complete and cannot be made more valuable by lasting longer. To his critics, this account of pleasure seems suspiciously tailor-made to his claim that death in no way harms the pleasant life. To his followers, it served as a powerful explanation of how pleasure, when properly understood, can meet the demands that our happiness be not only self-sufficient, but also complete and invulnerable to any harm at the hands of death.

THE VIRTUES AND FRIENDSHIP

An undeniable optimism about the power of reason pervades every facet of Epicurus's ethics. He thinks that reason can lead us to eliminate easily any fears based on mistaken beliefs and he lauds its sober power to focus each of our choices on our final good. Like Socrates, he denies that we can know the good and yet fail to pursue it either because of unmanageable desires or incorrigible weaknesses in our character. He does not go as far as the

Stoics in simply identifying irrational desires with mistaken beliefs, but he likewise eschews Aristotle's emphasis on the necessity of habituating desires and character in the ways of virtue. For Epicurus, the therapeutic benefit of rational argument transforms lives at any stage and in any condition. Unsurprisingly, his account of virtue is strongly cognitive. All of the virtues including justice, he insists, are a species of rational prudence, instrumentally useful in securing and maintaining a life of pleasure.

Yet while prudence, courage, and moderation arguably might be claimed as virtues useful to the hedonist, the other-regarding demands of justice seem more problematic. Surely, we might suppose, one might have prudential reasons for being unjust. Why restrict one's pleasures in the interests of others? Epicurean texts offer a panoply of arguments in defense of justice that probably work best if they are taken to have as their addressees Epicurean and non-Epicurean alike. For instance, some texts single out the fear of being caught as the key incentive in obeying laws. Such a motive for being just fits rather badly with the picture of the Epicurean who is supposed to eliminate all disturbing fears and live quietly, and who in any case, has little motive for transgressing laws given the limited range of his necessary desires. Other texts praise the psychologically calming benefits of justice—hardly a motive for the non-Epicurean searching for less tranquil experiences. The central and most lastingly influential component of Epicurus's theory of justice, however, is his account of the origins and nature of justice as a mutual contract among agents "neither to harm nor be harmed." Although not the first contractual theory in antiquity, it was the Epicurean account with its postulation of an original prepolitical state of nature that took center stage for Rousseau and his predecessors.

Interestingly, Epicurus's particular formulation of the contractual theory rejects the conventionalism of many later theorists, since he argues that contracts failing to reflect mutual usefulness no longer constrain. By placing his theory of justice in the larger context of his accounts of virtue and pleasure, he insists on the essential connection between agents' continuing interests and the contracts they have formed.

Although he denies that the virtues are valuable for their own sake, Epicurus insists that one can achieve happiness only by being virtuous. His ancient critics often doubted, however, whether purely instrumental motives were sufficient for maintaining one's commitments to virtue. A parallel problem arises in his account of friendship. Epicurus often speaks of friendship in the most extravagant terms and some later Epicureans proclaim that a hedonist can value his friends for their own sakes. But if one's motives in acquiring a friend are securely rooted in one's own pleasure, how can one maintain this seemingly altruistic commitment to friends?

Various later Epicureans struggled with the problem. Some conceded that if one properly focused on one's own pleasure, one could not treat a friend's pleasures as one's own, while others argued that friends might enter into mutually self-interested contracts to value each other's pleasures equally. A few went so far as to claim that one could come to value friendship in a way that went beyond motives of individual egoistic pleasure. All, perhaps, reflect a worry about what has come to be called the hedonist paradox. We can gain the pleasure we seek from friends only by maintaining the other-regarding values of friendship and by valuing friends' pleasures as much as our own. If we instead concentrate on our own pleasures, as hedonism seems to demand, we will undermine the very values that bring us pleasure. It is unlikely, however, that Epicureans would take this as evidence against their theory. Rather, for them, valuing friends' pleasures as much as their own represents an enlightened hedonist strategy that they need to fit, however awkwardly, within the confines of their overall theory.

INDETERMINISM, FREE WILL, AND THE MIND

Whether or not Epicurus, as many have claimed, was the first to formulate questions about determinism and free will in their modern form, it is clear that he attempts to find room in a mechanistic universe for our rational pursuit of pleasure and hence for our ability to rationally discriminate among alternative choices. Like Aristotle, he rejects logical determinism, as a threat to rational deliberation, and he denies that statements about future contingent events have truth value. Likewise, he contends that if the laws of atomic motion utterly determined everything in the universe, rational argument and choice could have no meaning and they would be robbed of their efficacy. What underwrites the manifest efficacy of reason and frees us from the bonds of both logical and mechanistic fatalism, Epicurus claims, is a slight indeterminacy or swerve in the movements of atoms at no fixed time or place. Such swerves break the bonds of necessity by interrupting the endless chains of causal interactions among atoms. Attempts at understanding the physics of these random swerves and their precise effects on human action abound. Some, for example, have seen the origins of libertarian defenses of free will in Epicurus's account,

with each random swerve of atoms underwriting a free and uncaused act of human volition.

Others, worried about the plausibility of such a strict correlation between micro and macroscopic events, have postulated more infrequent effects by swerves on human actions or character generally. Still others have argued that Epicurus's account is nonreductive or emergentist in a way that defuses the randomizing effects of atomic indeterminism at the macroscopic level. Swerves break causal bonds among atoms, but without generating randomness in emergent properties, thereby insuring the efficacy of rational deliberation and action. In the face of these disagreements, a few have concluded that Epicurus's main worries are innocent of such theoretical niceties and that ascribing to him either libertarian or emergentist views is merely anachronistic. Wherever the truth ultimately lies, it is certainly the case that from the early modern period onward, Epicurus was held to be the chief ancient proponent of libertarianism. And as is often the case in the history of philosophy, the subsequent reception of one's views can be far more influential than one's actual views.

Similar difficulties arise for Epicurus's philosophy of soul or mind. On the one hand, he regularly holds out the ambition of giving a strict identity theory of mind and explicates its materiality with an array of arguments showing that materialism alone explains the mutual causal interactions between mind and body. He claims, moreover, that the mind is composed of particular types of atoms whose specific properties are directly correlated to particular mental functions; for example, the smoothness of specific atoms accounts for the quickness of thought. It was the explanatory power of such instances of Epicurean reductionism that so influenced Julien de La Mettrie and d'Holbach and set the agenda for subsequent eliminativist theorists. Yet Epicurus also emphatically insists that mental properties are real properties and not mere epiphenomena. Some have taken his robust commitment to the reality of macroscopic properties as evidence for anti-reductionism or emergentism, while others have argued that his commitment to physicalist explanation not only entails reductionism, but that reductionism is entirely compatible with his endorsement of the reality of macroscopic entities. Perhaps if our own distinctions between explicative and nonreductive physicalist theories could be more confidently drawn, a choice among these options for Epicurus would be more easily forthcoming.

KNOWLEDGE AND ATOMS

However he intended to explain the relation between atoms and our world of perceptual experience, it is clear that Epicurus thinks we can attain certain knowledge of both. Epicureans share with Plato and Aristotle a conviction that skepticism is self-refuting both in theory and practice, but they disagree markedly about the criterion for knowledge immune to skeptical attack. Epicurus's epistemology, or "canonic," begins with the emphatic assertion that all sensations are true. What guarantee their truth are the mechanisms of their production. Films or images of inconceivable speed are continually streaming from bodies and striking our senses, which function simply as passive receptors that in no way distort the information they receive. Error can occur only if we extrapolate to the world from these sensations with mistaken assumptions or judgments. This we can avoid by applying to our sensations simple concepts (*prolepseis*) that, arising naturally from repeated sensations, are preserved in memory and embodied in the ordinary meanings of words.

Distrusting Aristotelian-style logics, Epicureans elaborate inferential methods for eliminating false judgments based on what they call confirmation or disconfirmation by the senses. For most subsequent empiricists, such an attempt to skirt formal logic became a dead end. However, Epicurus's arguments in defense of empiricism, especially his materialist accounts of the mechanics of perception and concept formation, enjoyed a long philosophical run, providing basic templates for a succession of empirical theories based on effluences, sense data, and other factors. Few, however, rival his ingenious attempts to work out the actual mechanics of perception in the face of what were to become standard objections to imagist theories of thought and perception.

Armed with empirical knowledge, the Epicurean believes, we can confirm the basic principles of atomism. For instance, from our perception of motion in the world, we can infer that atoms must move—an inference for which only the existence of void offers adequate explanation. Similarly, to explain the multiplicity of things we see in the word, we must postulate an indefinite number of atomic shapes. The existence of atoms themselves is confirmed by our observation that nothing comes from nothing and that division must stop somewhere for there to be something.

Interestingly, Epicurus allows multiple and competing theoretical explanations for some natural phenomena. But he indulges no such operationalist sensibilities toward the basic principles of his atomism. Collisions,

rebounds, and the compounding of atoms are the basic mechanisms of the material world. In an infinite, eternal world of atoms in continual motion, there is no room for teleology or interference from the gods, who in any case are indifferent to human concerns and, at least in some Epicurean sources, seem to have no stable constitution of their own but arise in relation to our mental conceptions of them. Atoms have the properties of shape, size, and weight. Explaining differences in atomic sizes and shapes presented Epicurus with difficult puzzles, however. Unlike later philosophers, such as Agostino Nifo, who constructed atoms from independently existing minima, Epicurus denies their physical divisibility. He postulates, instead, only conceptual divisibility, which required him, with rather mixed success, to theorize spatial and temporal quanta. His argument that atoms have weight, however, was fundamental. It changed the course of ancient atomism and gave the theory its essential shape down to the time of John Dalton.

See also Atomism; Cyrenaics; Epicureanism and the Epicurean School; Hedonism; Leucippus and Democritus; Lucretius; Philodemus.

Bibliography

WORKS BY EPICURUS

Epicurus: The Extant Remains, edited by Cyril Bailey. Oxford: Oxford University Press, 1926.

The Epicurus Reader: Selected Writings and Testimonia, edited by Brad Inwood and Lloyd P. Gerson. Indianapolis, IN: Hackett, 1994.

The Hellenistic Philosophers, edited by A. A. Long and David Sedley. Cambridge, U.K.: Cambridge University Press, 1987.

Epicurea, edited by Hermann Usener. Leipzig: Teubner, 1887.

WORKS ABOUT EPICURUS

Algra, Keimpe et al., eds. *The Cambridge History of Hellenistic Philosophy*. Cambridge, U.K.: Cambridge University Press, 1999.

Annas, Julia. *Hellenistic Philosophy of Mind*. Berkeley: University of California Press, 1994.

Annas, Julia. *The Morality of Happiness*. Oxford: Oxford University Press, 1993.

Asmis, Elizabeth. *Epicurus's Scientific Method*. Ithaca, NY: Cornell University Press, 1984.

Bailey, Cyril. *The Greek Atomistis and Epicurus*. Oxford: Oxford University Press, 1928.

Clay, Diskin. *Paradosis and Survival: Three Chapters in the History of Epicurean Philosophy*. Ann Arbor: University of Michigan Press, 1998.

Englert, Walter G. *Epicurus on the Swerve and Voluntary Action*. Atlanta: Scholars Press, 1987.

Festugiere, A. J. *Epicurus and His Gods*. Oxford: Oxford University Press, 1955.

Konstan, David. *Some Aspects of Epicurean Psychology*. Leiden: E. J. Brill, 1973.

Mitsis, Phillip. *Epicurus's Ethical Theory: The Pleasures of Invulnerability*. Ithaca, NY: Cornell University Press, 1988.

Nussbaum, Martha. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press, 1994.

O'Keefe, Timothy. *Epicurus on Freedom*. Cambridge, U.K.: Cambridge University Press, 2005.

Warren, James. *Facing Death: Epicurus and his Critics*. Oxford: Oxford University Press, 2004.

Phillip Mitsis (2005)

EPISTEMIC LOGIC

See *Modal Logic*

EPISTEMOLOGY

Epistemology attempts to explain the nature and scope of knowledge and rational belief. Its purview also includes formulating and assessing arguments for skeptical conclusions that we do not have knowledge of various kinds. In addition, epistemologists address topics that are closely related to these core concerns, including evaluations of thought processes and the relationship of science to philosophy. What follows is an overview of contemporary developments in epistemology.

THE ANALYSIS OF KNOWLEDGE

The traditional analysis of knowledge is that it is a combination of three conditions: truth, belief, and justification. The idea is that for someone to have factual knowledge, what is known has to be a fact and thus true; the person has to regard it as true, that is, believe it; and the person must have an adequate basis for believing it—that is, have sufficient justification for believing it. These conditions yield knowledge defined as a sufficiently justified true belief.

The publication by Edmund Gettier (1963) of one brief critical discussion of the traditional analysis brought about a flurry of activity in epistemology. Gettier refuted the traditional analysis by offering convincing counterexamples. He described examples in which someone forms a belief on the basis of strong justifying evidence, but the belief merely happens to be true as a result of a fortunate accident, independently of the evidence. Here is an example similar to Gettier's. Someone sees something that looks perfectly sheeplike in a nearby field.

On that basis the person justifiably believes that there is a sheep in the field. As it turns out, what the person sees is not a sheep. It is a highly realistic statue. However, the person's belief that there is a sheep in the field is true because of the fortunate coincidence that there is a real sheep hidden from view elsewhere in the field. Such a belief is clearly not a case of knowledge despite its being an instance of justified true belief. So justified true belief is not sufficient for knowledge.

Arguing that the person in the example does not have an adequate basis for believing that there is a sheep in the field seems to require taking the general position that few beliefs are justified. For if that person does not have an adequate basis and is not justified, then someone in a similar situation who actually does see a sheep would also be unjustified, given that her visual information would be no better. In almost all cases of actual knowledge of the world, there are possible, although unusual, cases in which one has the same belief on the basis of comparable reasons, yet that belief is only true in this accidental way. Therefore, responding to the Gettier cases by raising the standards for justification leads to the conclusion that we know very little.

Most epistemologists responded to Gettier's examples by seeking a fourth condition for knowledge in addition to justified true belief. Some proposed that to have knowledge, it is also required that the justification for one's belief be *undefeated*, meaning roughly that there is no truth that would undermine the justification for the belief (Klein 1976). Others have suggested that in cases of knowledge the justification does not involve a falsehood (Chisholm 1989). Still others have required that the reasons justifying a known belief be *conclusive*—roughly, reasons that would not exist unless the belief were true (Dretske 1971). Counterexamples refuted the original versions of these analyses, more complex analyses replaced the originals, and new counterexamples followed. (See Shope [1983] for a detailed summary of responses to Gettier's examples.)

Not all epistemologists accept the necessity of the three traditional conditions for knowledge. Some reject the justification condition. One proposed replacement requires a suitable causal connection between a known belief and the facts that make the belief true (Armstrong 1973, Goldman 1967). Another proposed replacement requires a known belief to vary counterfactually with the truth of that belief: if the belief were not true, it would not be believed by the same method, and if it were true, it would be believed by the same method (Nozick 1981). Others have taken the more drastic tack of denying that any set of necessary and sufficient conditions for knowledge can be given. An alternative explanation of knowledge is that it is the most inclusive *factive* mental state (Williamson 2000). A mental state is factive if the existence of the state guarantees its truth. Unlike the traditional analysis, this approach does not imply that the concept of knowledge can be decomposed into parts.

Although epistemologists have learned much about knowledge from this research, no consensus has emerged about the solution to the problem raised by examples like Gettier's.

JUSTIFICATION: FOUNDATIONALISM AND COHERENTISM

Justification itself has been investigated intensively in the wake of the Gettier problem. A central issue underlying views about justification is the infinite-regress problem. Typically, a belief is justified because it has support from other beliefs. For example, someone might be justified in believing that there are people in the next room by inference from the justified belief that Allen, Barbara, and Carol are in the next room. The supporting beliefs garner support from still other beliefs. The belief that Allen, Barbara, and Carol are in the next room might be justified by inference from the justified belief that they said they would enter the next room and then shouted that they had done so. However, given that our minds are finite, there cannot be an infinite regress of justifying beliefs. Therefore, either there are some beliefs—*basic* beliefs—that are justified without the support of other beliefs; or our beliefs form some sort of circle or web, with each justified/rational belief getting support from other beliefs within the system; or our beliefs are not justified at all. *Foundationalism* favors the first alternative, while *coherentism* favors the second. The third alternative, that no belief has any justification, seems indefensible.

The classic foundationalist view is that a belief is justified provided that it is a basic belief or rests upon a foundation of basic beliefs. Usually, the contents of basic beliefs are taken to be propositions about the mental states of the believer. For example, when someone observes an ordinary physical object in good viewing conditions, that person's visual system produces an experiential state. This is an internal mental state of the observer, knowable by introspection. Believing about oneself that one is in this experiential state is said to be a basic belief. Beliefs of this sort are supposed to provide a secure foundation for the rest of our justified beliefs. Classic foundationalists differ about the source of the security of basic beliefs. Candidate sources include the

alleged infallibility of our introspective capacities and the alleged immunity from doubt of some beliefs. According to classic foundationalism, we acquire whatever justified beliefs we can get about the external world by inference from our introspectively justified beliefs about our own states. Some foundationalists hold that only a deductive (logically necessary) connection can secure sufficient justification for knowledge, whereas others hold that inductive or explanatory relations also suffice. The question of what support is sufficiently strong for knowledge is central to the discussion of epistemological skepticism.

Some foundationalists have relaxed the requirements for basic beliefs (Chisholm 1989, Huemer 2001). The central foundationalist view is that each justified belief is basic or derives its justification from basic beliefs. This view does not require basic beliefs to be certain or infallible. A more modest level of independent support is enough to stop the regress of derived justification. Foundationalists can consistently hold that support from other beliefs gets the basic beliefs beyond this modest level. If the basic beliefs need not be maximally secure, then another departure from the classic view becomes attractive. Basic beliefs can include ordinary perceptual beliefs. For example, the belief that one sees a dog can be basic. It can gain some justification that is independent of other beliefs directly from an experience, which is visually just as though one is seeing a dog. Modest foundationalism is widely thought to be an improvement upon classical foundationalism.

Modest foundationalism has its share of critics, however. Its defenders have been challenged to explain how the basic beliefs can receive even modest support from experience (BonJour 1985). The main problem is that the best understood sort of epistemic support is the justification that is given by the premises of a strong argument for its conclusion, yet the experiences cited by modest foundationalists as providing foundational support do not seem to qualify as premises of arguments. This is because experiences are not statements, but the only kinds of things that can be premises are statements.

Coherentism is the chief rival to foundationalism (Lehrer 1974, BonJour 1985). Coherentists deny that there are any basic beliefs. The secure foundations that classic foundationalists have sought are, according to coherentists, impossible. They contend that all justified beliefs get their justification from a relation of coherence that holds among a body of beliefs. Coherentists have attempted to say what constitutes coherence, often appealing to explanatory relations among beliefs as the source of coherence. Some propose that justification arises from *reflective equilibrium*—a mutual adjustment of beliefs about particular cases and beliefs about general principles covering these cases that maximizes explanatory relationships among them (Goodman 1984).

Coherentists have been challenged to avoid the apparent implication of their theory that justified beliefs can have an implausible sort of detachment from sensory input. A body of beliefs can be internally coherent while the beliefs fail to take into account the person's experience, yet coherentism seems to imply that these cohering beliefs would be justified. Intuitively, however, such beliefs seem to be as unjustified as the beliefs formed by accepting as true everything in some well-crafted, elaborate, but fantastic story.

Not all philosophers agree that we must choose sides between foundationalism and coherentism. Several have argued that the central epistemological considerations on both sides can be reconciled (Alston 1989, Haack 1993, Sosa 1991).

JUSTIFICATION: OTHER ISSUES

In addition to formulating and assessing foundationalism, coherentism, and other theories of justification, epistemologists have addressed a variety of other questions about epistemic justification. Standard versions of foundationalism and coherentism share the presupposition that justification is determined by relations among the reflectively accessible contents of our minds—experiential states, beliefs, memories, inferences, and so on. Some philosophers, however, have opposed this *internalist* presupposition, engendering extensive discussion of the contrast between this view about justification and its *externalist* alternatives. (See Kornblith 2001 for essays on these issues.)

For internalists, justification is determined entirely by internal mental factors, whereas externalists assert that justification is at least partly determined by other things. Some internalists also require the believer to be aware of all justifying factors. A typical internalist theory is *evidentialism*, which holds that evidence held in mind determines the epistemic status of beliefs (Conee and Feldman 2004, Haack 1993). *Reliabilism* exemplifies the externalist viewpoint (Goldman 1979). Reliabilism maintains that a belief's justification is determined by a propensity to produce true beliefs of the process or mechanism leading to the belief. This reliability is not an internal factor because the truth of a belief is usually not an internal fact.

A good example to point out the difference between an internalist theory and reliabilism involves the victim of

a deceptive demon. The demon induces the victim to have the experiences like those a reasonable person might have through the perception of an ordinary environment. The demon's victim forms the same external world beliefs on the basis of these experiences. It is a further part of the example that the external world of the demon's victim is not at all an ordinary environment, and so her beliefs about her external world are not true. In such an example, the processes leading to the victim's external world beliefs seem to be unreliable because they produce her thoroughly false external world beliefs. So reliabilism seems to imply that such beliefs are not justified. The belief-forming processes of the counterpart person in a normal environment are presumed to be reliable, so that reliabilism implies that this person's beliefs are justified. In contrast, according to any internalist theory, the beliefs of both the normally situated person and the demon's victim are equally well justified if the individuals are in the same internal states.

Reliabilism has been a subject of intensive critical scrutiny since its introduction. Critics contend that reliabilists cannot plausibly specify the types of belief-forming processes or mechanisms on which the theory relies (Conee and Feldman 2004). For instance, the process of forming a typical visual belief can be classified as perception, visual perception, belief acquisition while relaxed, uninferred belief acquisition, and so on, indefinitely. The problem is to specify which of these process types has to be reliable in order for the resulting beliefs to be justified. Reliabilists must specify the relevant type for all of the processes that lead to justified beliefs. Critics have also charged that beliefs resulting from a reliable process can be unjustified when accompanied by a sufficient reason to think that the process is not reliable (BonJour 1985) and that beliefs resulting from an unreliable process can be justified when accompanied by reason to think that the process is reliable.

Some theories of justification require supplements to reliability. For instance, a proper functionalist theory holds that a belief is justified when the belief results from the operation of a generally reliable cognitive system that is functioning properly in an appropriate environment. One theistic variant of this view holds that the proper function of human cognitive systems is the result of the intentions of a creator (Plantinga 1993). In a nontheistic version, proper function is determined by natural selective forces. One prominent criticism of the proper functionalist approach is that it is possible for a cognitive mechanism to function improperly but felicitously. A perceptual mechanism might accidentally happen to work much better than it was designed to work. A resulting belief could be especially well justified by the acute perception.

Epistemologists also make comparisons between epistemic justification and ethical concepts such as obligation. Discussions of what a person is justified in believing easily slide into discussions of what the person *should* believe or is *entitled* to believe. Such talk is at least superficially similar to ethical evaluations of how a person should behave and what things the person is entitled to do. It can seem that the epistemic and ethical evaluations are fundamentally the same. However, there is some question about the applicability of ethical evaluations to beliefs. It is widely thought that what one morally should do is limited to those things that one can do. If something similar holds in epistemology, then what one should believe is limited to those things that one can believe. It apparently follows from this premise that beliefs must be under our voluntary control if we are to speak of our being justified in having them. Yet it seems that beliefs are not typically under voluntary control. Some philosophers respond by arguing that, contrary to appearances, we have sufficient control over our beliefs; some contend that it is acceptable to hold that we have justification for believing some propositions even though we are not able to control whether we believe them; and others conclude that few, if any, beliefs are justified since few, if any, are under our control. There is also concern about the connections between the epistemic justification of a belief and the moral or practical benefits of the belief. (Essays on this topic are collected in Steup 2001.)

Another widely discussed set of issues turns on a distinction between a priori justification and *a posteriori* justification. Justification of a belief is a priori when it does not derive from experience, and justification is a posteriori when it does. The leading candidates for a priori justification and knowledge are beliefs in basic truths of mathematics and logic. Other candidates include beliefs apparently made true entirely by conceptual relations, such as the belief that anything red is colored. These allegedly a priori justified propositions are, if true, necessarily true.

A priori justification seems mysterious to many philosophers, since it is difficult to understand what could justify beliefs independently of experience. A wide range of proposals has been made concerning how beliefs can have a priori justification. In the naturalistic approach a priori justification results from the operation of belief-forming processes that guarantee truth and justification (Kitcher 1980). The modal-reliability approach

holds that conceptual intuitions necessarily present us with mostly truths (Bealer 2000). And a resolutely traditional approach holds that humans have a capacity for rational insight that finds truth-making, necessary connections in some thoughts (BonJour 1998).

It appears that a belief could not be a priori justified or known unless its truth is somehow abstractly guaranteed. It also appears that if there is an abstract guarantee that a belief is true, then the truth of the belief must not be merely contingent. So a priori knowledge of contingent truths would be surprising. Yet some philosophers have argued that we can have such knowledge (Kripke 1980), advancing the following kind of argument: Suppose that there is a unique tallest spy; knowing nothing about this and reasoning entirely in our armchairs, we can stipulate that the name "Stretch" refers to whoever happens to be the tallest spy, if there is one. Having done this, it seems that we can logically infer from what we have done, and thereby know a priori, the following contingent truth: if there is a unique tallest spy, then Stretch is a spy. Perhaps this knowledge would not be strictly a priori, since we would be using the experience of our introduction of the name "Stretch." Nonetheless, it seems to be a way to know a contingent truth that is at least remarkably similar to a priori knowledge.

SKEPTICISM

Many traditional skeptical arguments appeal to the possibility of error. Skeptics often point out that it is possible for us to be wrong about even our most confident beliefs about the world external to our minds, perhaps because we are under the influence of a deceptive demon or some other source of deception. Skeptics typically make the further claim that this possibility implies that we lack knowledge of even the things about the world that we most confidently believe. (Many influential essays on skepticism may be found in DeRose 1999.)

Fallibilism is the heart of one influential response to skepticism (Chisholm 1989, Pryor 2000). Fallibilism is the view that knowledge is compatible with the possibility that the same belief on the same basis is false. For example, someone who has a clear view of a tree in the front yard and believes on a normal perceptual basis that there is a tree in the front yard is subject to some possibility of error. An experience that is visually just as though one is seeing a tree could have resulted from things like the efforts of a deceptive demon. However, a typical person who sees a tree has no reason at all to think that any such odd thing is actually occurring and every reason to think that there really is a tree present. Fallibilists hold that in such cases people often have sufficiently strong justification to know that there is a tree in the yard. According to fallibilists, a skeptical argument like the one about the possibility of error relies on setting the standard of justification for knowledge too high. We can have knowledge even though we cannot have the sort of absolute immunity from error that the skeptics wrongly associate with knowledge.

Fallibilism is not without problems. It is no easy task to explain what it is about our experiential evidence that makes it a good reason for thinking that we are in the presence of ordinary objects rather than the victims of some sort of deception. Some epistemologists contend that our justification for our external world beliefs depends upon an *inference to the best explanation* of our experiences (Vogel 1990), whereas others contend that there is something intrinsic to the character of experiences that makes them indicative of external world objects. Adequately spelling out just why our beliefs are even fallibly justified remains an unfulfilled task.

Some influential arguments for skepticism are updated versions of arguments based on possibilities of deception by dreams or demons. The newer arguments often appeal to the possibility of being a *brain in a vat*. The brain-in-a-vat arguments make use of the possibility that a fully functioning human brain, immersed in a life-sustaining vat of chemicals, receives computer-controlled neural stimulation that exactly matches the neural stimulation of an ordinary person in an ordinary environment. A premise of one brain-in-a-vat argument is that any of us might, for all we know, actually be such a brain in a vat. The argument also assumes that, since this possibility might be actual, we lack knowledge of the actual external world.

A much-discussed reply to such arguments employs a causal view of reference (Putnam). On one interpretation, the reply begins with the surprising contention that what a vat-entrapped brain would express by *I am a brain in a vat* would be a falsehood. A lifelong vat-entrapped brain would have learned the term *vat* from some computer-generated stimulations. The origin of the stimulation within the computer would have no causal connection to the brain's container of a sort that would be required for the brain's term *vat* to apply to the container. Hence, according to a causal view of reference, the brain's sentence *I am a brain in a vat* would not be true. Of course, what people in normal circumstances express by that same sentence is also false. Thus, the sentence *I am a brain in a vat* does not express a truth, whichever of these situations we are actually in. The antiskeptical reply

concludes that by this use of a causal view of reference, we can justify denying the brain-in-a-vat argument's premise that, for all we know, we might be brains in a vat.

The success of this sort of antiskeptical reply is in dispute. In any event, a notable limitation of the approach is that at best it refutes skeptical arguments that rely on only some brain-in-a-vat possibilities. For instance, one possibility that is unaffected by the reply is that we recently became brains in a vat, and our term *vat* refers to the vat containing us because proper causal connections were forged in our pre-vat situation.

Skeptical arguments frequently rely on an *epistemic-closure* principle that says that if a person knows one proposition and sees that another proposition follows immediately from it, then the person knows the latter proposition, too. If someone knows an ordinary fact such as that she is seeing a table, then the closure principle implies that she could know by deduction that she is not a mere brain in a vat. Since, according to some skeptics, she cannot know that she is no brain in a vat, the skeptics conclude that she does not know anything from which she could deduce this, such as that she is seeing a table. Some philosophers have denied the closure principle in an effort to argue against this case for skepticism about knowledge of ordinary facts. Most philosophers, however, contend that some version of the closure principle must be true and any mistakes in skeptics' arguments must lie elsewhere (Hawthorne 2004).

Another response to skepticism appeals to *epistemic contextualism* (Cohen 1999, Lewis 1996). Contextualists endeavor to account for the intuitive pull of the arguments for skepticism while allowing that many of our ordinary attributions of knowledge are correct. Their central thesis is about *truth conditions* for uses of sentences including the word *know* and kindred terms. A statement of the truth conditions for a particular use of a sentence specifies the conditions that have to be realized in order for that use of the sentence to state a truth. The main form of epistemic contextualism holds that the truth conditions of particular uses of any sentence including *know*, or cognate expressions, vary with the context in which the sentence is used.

Typically, the varying aspect of the truth conditions is said to be the strength of the epistemic position that is required of the subject of the sentence for a use of *know* to apply to the subject. Usually, contextualists assert that the required strength of epistemic position varies across a range that allows, at its low end, many true sentences that attribute "knowledge" to someone. Thus, what we say is often true when, in ordinary circumstances, we classify as "knowledge" beliefs that are based on perception, memory, testimony, and perhaps inductive generalization and inference to the best explanation. Contextualists typically also assert that some contexts, at the high end of the range of variation, are demanding enough to make true denials of "knowledge" of the external world. For instance, contextualists often claim that where issues concerning skepticism are salient, the standards for true attributions of "knowledge" are very high and that consequently, in those contexts, skeptical denials of "knowledge" are correct.

Some critics of contextualism deny that skepticism is true even when arguments for it are salient. Appealing to antiskeptical grounds such as the fallibilism discussed above, the critics contend that the arguments fail and that skepticism is wrong whether or not we are thinking about it (Conee and Feldman 2004). Other critics question the linguistic foundations of contextualism (Stanley 2004).

DEPARTURES FROM TRADITION

The philosophical study of knowledge, justification, and skepticism is the core of traditional epistemology. Some epistemologists have extended the discipline. One such extension involves connecting epistemology to scientific research about how people form beliefs and how they process information. Naturalism in epistemology is roughly the view that there is substantial overlap between epistemology and the sciences that study human cognition. Some philosophers endorse naturalism, whereas others find a reasonably clear distinction between the scientific/empirical questions about cognition and the conceptual questions at the heart of epistemology. A radically naturalistic epistemology advocates abandoning traditional epistemology and replacing it with the closest empirical discipline, cognitive psychology (Quine 1969). Few philosophers defend this extreme view. However, many urge close ties between epistemology and empirical studies of human cognition. For example, epistemologists who highlight the search for ways to improve our reasoning contend that the empirical study of how people actually reason is crucial for developing useful recommendations (Kornblith 1994). Philosophers who believe that the primary role of epistemology is to explain the concepts of knowledge, justification, and the like typically see less room for empirical input. Some advocate a less extreme form of naturalized epistemology that requires explaining central epistemic concepts in terms that they deem naturalistically legitimate.

Traditional epistemology has been largely individualistic in its emphasis on questions about knowledge and

justification as they apply to individuals. However, a social epistemology has arisen that raises questions about what it is for groups to have knowledge and how social factors influence the spread and development of knowledge (Schmitt 1994, Goldman 1999).

Another approach in epistemology highlights epistemic virtues (Sosa 1991). One version of virtue epistemology is a variant on the reliabilist view discussed earlier. This approach attempts to characterize knowledge or justification in terms of epistemic virtues that yield reliably true beliefs, such as open-mindedness and a willingness to consider new evidence. In a greater departure from traditional issues, other versions of virtue epistemology propose that epistemologists replace or supplement the traditional topics with that of virtuous epistemic conduct.

EPISTEMOLOGY AND RELATED DISCIPLINES

There has been extensive and significant epistemological work done in relation to issues in the philosophy of mind. Externalism in the philosophy of mind, usually called *content externalism*, is the widely held view that environmental factors can help to determine the identity of some mental states. One simple content-externalist claim is that the content of a person's thoughts formulated with natural kind terms, such as *elm* and *water*, depends on causal connections to the kind that was actually involved in the person's learning the term. If the connection had been to a different natural kind, then the person's thoughts formulated with the same term would have included a concept referring instead to the other kind. There need not be any distinguishing feature that displays to the person which kind the person's thoughts are about.

Seemingly, if this simple content externalist theory is true, then we can know it a priori. We can know that external causes help to determine some thought contents by just considering how the reference of our natural-kind terms intuitively varies in some causally different hypothetical situations. If this is correct, then the theory appears to be incompatible with the conjunction of two plausible epistemological doctrines. One of the doctrines is that we can know the contents of our own thoughts by just giving introspective attention to them. If so, then we could combine our a priori knowledge of the simple content externalist theory with our introspective knowledge of the content of one of our thoughts that is expressible using *water*. We could infer that water is causally connected to the thought and that water therefore exists. Yet, according to a second plausible epistemological doctrine, knowledge of our environment is not so easy. It requires empirical information. Thus, the simple content externalist theory seems to imply that either we cannot know the contents of our thoughts as easily as it otherwise seems we can, or that empirical knowledge of the existence of things in our environment is easier than it otherwise seems to be.

Critics of this line of reasoning have asked whether it can really be known, without empirical investigation, that content externalism applies to any of our concepts. The applicability of the version of content externalism described here to a concept is contingent on the existence of an appropriate causal connection between the concept and some natural kind. This dependence suggests that empirical information about the existence of a properly connected kind is needed to justify applying content externalism to our concepts. (For further discussion, see the essays in Nuccetelli 2003.)

Much that qualifies as epistemology has been done in other areas of philosophy. What follows is a brief inventory of some epistemic work in allied fields. One classic epistemological topic is the problem of induction. This is the problem of establishing whether or not people can use observation of some cases to draw justified conclusions about unobserved cases, and if this can be done, explaining when and why such inferences are reasonable. This problem has been pursued within the part of philosophy of science known as confirmation theory. Second, factual knowledge entails truth. Truth is a traditional topic in epistemology. Various theories of truth are also presented and discussed in metaphysics, the philosophy of language, and philosophical logic. Third, rational change of belief is closely related to the epistemological topic of justified belief. Rational belief change is a focus of probability theory, especially under the classification of Baysian epistemology. Fourth, epistemological issues are often important to issues of morality and religion. Epistemic concerns pertaining to morality, such as the question of how we can know what is morally right, are usually discussed in works that are primarily about moral philosophy. Similarly, epistemic issues pertaining to God are discussed primarily in works in the philosophy of religion. Finally, in the vicinity of the border between epistemology and cognitive science there has been considerable attention devoted to the nature of purported sources of knowledge and to the ways in which they do their epistemic work. Topics here include perception, memory, intuition, and testimony.

See also A Priori and A Posteriori; Basic Statements; Classical Foundationalism; Coherentism; Contextualism;

Doubt; Epistemology, History of; Evidentialism; Experience; Illusions; Inference to the Best Explanation; Internalism versus Externalism; Introspection; Intuition; Knowledge and Belief; Knowledge and Truth, The Value of; Knowledge, A Priori; Memory; Naturalized Epistemology; Perception; Propositional Knowledge, Definition of; Rationalism; Reason; Relevant Alternatives; Reliabilism; Self-Knowledge; Skepticism, History of; Social Epistemology; Solipsism; Subjectivist Epistemology; Testimony; Virtue Epistemology.

Bibliography

Alston, William. *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press, 1989.

Armstrong, David. *Belief, Truth, and Knowledge*. Cambridge, U.K.: Cambridge University Press, 1973.

Bealer, George. "A Theory of the A Priori." *Pacific Philosophical Quarterly* 81 (2000): 1–30.

Bennett, Jonathan. "Why Is Belief Involuntary?" *Analysis* 50 (1990): 87–107.

BonJour, Laurence. *In Defense of Pure Reason*. Cambridge, U.K.: Cambridge University Press, 1998.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Chisholm, Roderick. *Theory of Knowledge*. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1989.

Cohen, Stewart. "Contextualism, Skepticism and the Structure of Reasons." *Philosophical Perspectives* 13 (1999): 57–89.

Conee, Earl, and Richard Feldman. *Evidentialism: Essays in Epistemology*. Oxford, U.K.: Oxford University Press, 2004.

DeRose, Keith, and Ted A. Warfield, eds. *Skepticism: A Contemporary Reader*. Oxford, U.K.: Oxford University Press, 1999.

Dretske, Fred. "Conclusive Reasons." *Australasian Journal of Philosophy* 49 (1971): 1–22.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121–123.

Goldman, Alvin. "A Causal Theory of Knowing." *Journal of Philosophy* 64 (1967): 357–72.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Goldman, Alvin. *Knowledge in a Social World*. Oxford, U.K.: Oxford University Press, 1999.

Goldman, Alvin. "What Is Justified Belief?" In *Justification and Knowledge*, edited by George S. Pappas, 1–24. Dordrecht: D. Reidel, 1979.

Goodman, Nelson. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press, 1984.

Haack, Susan. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Cambridge, U.K.: Blackwell, 1993.

Hawthorne, John. *Knowledge and Lotteries*. Oxford, U.K.: Oxford University Press, 2004.

Huemer, Michael. *Skepticism and the Veil of Perception*. Lanham, MD: Rowman and Littlefield, 2001.

Kitcher, Philip. "A Priori Knowledge." *The Philosophical Review* 89 (1980): 3–23.

Kitcher, Philip. "The Naturalists Return." *The Philosophical Review* 101 (1992): 53–114.

Klein, Peter. "Knowledge, Causality and Defeasibility." *Journal of Philosophy* 73 (1976): 792–812.

Kornblith, Hilary, ed. *Epistemology: Internalism and Externalism*. Malden, MA: Blackwell, 2001.

Kornblith, Hilary, ed. *Naturalizing Epistemology*. 2nd ed. Cambridge, MA: MIT Press, 1994.

Kripke, Saul A. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Lehrer, Keith. *Knowledge*. Oxford, U.K.: Oxford University Press, 1974.

Lewis, David. "Elusive Knowledge." *Australasian Journal of Philosophy* 74 (1996): 549–67.

Nozick, Robert. *Philosophical Explanations*. Oxford, U.K.: Oxford University Press, 1981.

Nuccetelli, Susana, ed. *New Essays on Semantic Externalism and Self-Knowledge*. Cambridge, MA: MIT Press, 2003.

Plantinga, Alvin. *Warrant and Proper Function*. Oxford, U.K.: Oxford University Press, 1993.

Pryor, James. "The Skeptic and the Dogmatist." *Nous* 34 (2000): 517–49.

Putnam, Hilary. "Brains in a Vat." In *Reason, Truth, and History*. Cambridge, U.K.: Cambridge University Press, 1981.

Quine, W. V. O. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Schmitt, Frederick, ed. *Socializing Epistemology*. Lanham, MD: Rowman & Littlefield, 1994.

Shope, Robert. *The Analysis of Knowledge: A Decade of Research*. Princeton, NJ: Princeton University Press, 1983.

Sosa, Ernest. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Stanley, Jason. "On the Linguistic Basis for Contextualism." *Philosophical Studies* 119 (2004): 119–146.

Steup, Matthias, ed. *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford, U.K.: Oxford University Press, 2001.

Vogel, Jonathan. "Cartesian Skepticism and Inference to the Best Explanation." *Journal of Philosophy* 87 (1990): 658–666.

Williams, Bernard. "Deciding to Believe." In *Problems of the Self*. Cambridge, U.K.: Cambridge University Press, 1973.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford, U.K.: Oxford University Press, 2000.

Earl Conee (1996, 2005)
Richard Feldman (1996, 2005)

EPISTEMOLOGY, CIRCULARITY IN

Issues concerning circularity figure prominently in epistemology, finding a place in discussions of topics ranging from the Cartesian circle, to the problem of the criterion, to knowledge of the reliability of ways of forming belief.

DESCARTES AND CARTESIAN CIRCLES

Issues of circularity arise in the works of René Descartes. In his *Meditations*, in his search for items of certain knowledge (indubitable items given even the possibility of massive deception), Descartes finds that he is certain that he is a thinking being. But what makes this certain for him? The only explanation he finds is that he clearly and distinctly perceives this fact. Furthermore, he finds that clear and distinct perception could not be the source of such certainty if such perceptions could be false. So, he tentatively concludes, whatever is clearly and distinctly perceived is true. But does he really know this general principle? Could it not be that God has caused him to err even in what he clearly and distinctly perceives? Descartes then sets off to prove that a nondeceiving God necessarily exists.

In pondering the matter, Descartes seems to commit himself to the following two claims: (1) He can be certain that whatever he clearly and distinctly perceives is true only if he is first certain that God exists and is not a deceiver. (2) He can be certain that a nondeceiving God exists only if he is first certain that whatever he clearly and distinctly perceives is true. Accepting both these claims gives rise to the Cartesian circle. The problem is that if both (1) and (2) are true, then one cannot be certain of either the general principle or the view that a nondeceiving God exists. In general, if one must first know *A* to know *B* and one must first know *B* to know *A*, then it seems that one cannot know either *A* or *B*.

A related problem for Descartes concerns his use of his clear and distinct perceptions to support the general principle that whatever he clearly and distinctly perceives is true. To support this principle, he attempts to prove there is a nondeceiving God. Yet in his reasoning, he relies on premises that have no other support than his clear and distinct perceptions. Descartes thus relies on his particular clear and distinct perceptions to prove the general principle that whatever is clearly and distinctly perceived is true. To many critics, this is an epistemically unacceptable procedure. Seeing a similarity, the Scottish philosopher Thomas Reid objected that if a man's honesty were called into question, it would be ridiculous to trust his own testimony concerning his honesty. Epistemic circularity consists in using beliefs from source *A* to support the proposition that source *A* is reliable. Descartes's use of his clear and distinct perceptions to support the principle that whatever is clearly and distinctly perceived is true exhibits just such circularity. Whether such circularity is vicious is still debated. Below are some late-twentieth-century views on the issue.

CHISHOLM AND THE PROBLEM OF THE CRITERION

Another problem of circularity, one made prominent in contemporary epistemology by Roderick Chisholm (1973), is the ancient problem of the diallelus, or wheel. The problem is that to know which particular beliefs are instances of knowledge, one must know some criterion of knowledge. But among the many contenders, which criterion is the right one? To know that some proposed criterion is the right one, one must know that it picks out only instances of knowledge. Thus, to know which criterion is right, one must already know which beliefs are instances of knowledge. Chisholm formulated the problem in terms of a pair of questions: (A) What do we know? What is the extent of our knowledge? (B) How are we to decide whether we know? What are the criteria of knowledge? The problem of the criterion arises insofar as one must know the answer to B before one can answer A, and one must know the answer to A before one can answer B. As in the case of the Cartesian circle, if this is so, then one can answer neither A nor B.

Chisholm identified three responses to these questions: skepticism, methodism, and particularism. The skeptic says that to answer A one must first answer B and to answer B one must first answer A. Therefore, the skeptic concludes, one can answer neither question. One can neither pick out instances of knowledge nor identify a criterion for it. In contrast, the methodist holds that one can answer B first and then answer A. Unlike the skeptic, the methodist believes that he can know a criterion of knowledge. He holds that one must know a criterion of knowledge to know or pick out particular instances of it. Chisholm took the empiricism of Locke and Hume to be forms of methodism. One difficulty with this sort of methodism, according to Chisholm, is that it implies that we know nothing about the external world, other minds, the past, or the future. The third approach, the one favored by Chisholm, is particularism. Unlike the methodist and the skeptic, the particularist holds that one need not know a criterion of knowledge to pick out particular instances of it. The particularist holds that he can answer A and then work out an answer to B. Chisholm took Thomas Reid and G. E. Moore to be particularists. They held that we know pretty much what we ordinarily think we know, and if some philosophical criterion implies that we do not, then so much the worse for that criterion. Thus, if some criterion implies that I do not know that there are other thinking people or that I was alive yesterday, then that criterion must be mistaken.

RECENT DISCUSSIONS OF EPISTEMIC CIRCULARITY

The problem of epistemic circularity has received much attention in recent epistemology. As noted above, the problem of epistemic circularity arises for Descartes in his use of his clear and distinct perceptions to support the general principle that whatever is clear and distinctly perceived is true. Epistemic circularity also seems to be a feature of attempts to support the reliability of such doxastic sources as memory, sense perception, introspection, intuitive reason, and induction. To support the belief that one of these sources is reliable, one must appeal, it seems, to particular beliefs that issue from that source. For example, it seems that to support the belief that perception is reliable, one must appeal to beliefs produced by perception, and to support the reliability of memory, one must appeal to some memory beliefs. To many philosophers, epistemic circularity seems vicious and epistemically unacceptable. Other philosophers argue that it is not *necessarily* vicious or unacceptable.

To focus the discussion, consider the following track-record argument for the reliability of sense perception:

> At t_1, I formed the perceptual belief that *p*, and *p* is true.
>
> At t_2, I formed the perceptual belief that *q*, and *q* is true.
>
> At t_3, I formed the perceptual belief that *r*, and *r* is true.
>
> And so on.
>
> Therefore, perception is a reliable source of belief.

In this track-record argument for the reliability of perception, one reasons inductively from a wide sampling of perceptual beliefs, notes that the vast majority have been true, and concludes that perception is reliable. But how does one know that the second conjunct in each premise is true? How does one know, for example, that *p* is true? Let us suppose that it is known on the basis of perception. In this case, one is using perception to support the conclusion that perception is reliable. This makes the track-record argument epistemically circular. But is it therefore vicious or epistemically unacceptable?

A circular argument in which a premise is identical to the conclusion seems epistemically to carry no weight. Consider an argument of the form "*p*; *q*; *r*; therefore *p*." Arguments exhibiting this sort of circularity seem useless for conferring justification on the conclusion. If, on the one hand, one is not justified in believing the premise that *p*, then the reasoning does not justify the conclusion that *p*. If, on the other hand, one is justified in believing the premise that *p*, then it seems that the conclusion that *p* is already justified, and the reasoning or argument does not confer any justification on the conclusion that *p*. A defender of the track-record argument or epistemically circular arguments may concede that arguments of this form are epistemically without weight. He may still point out that the track-record argument above does not have this defect. The conclusion that perception is reliable is not identical with one of the premises, and so the argument is not for that reason unacceptable.

Another criticism of the track-record argument is based on a view about what is required for perceptual knowledge. A critic of the track-record argument might object that perceptual knowledge epistemically depends on one's knowing that perception is reliable. On this view, perceptual beliefs amount to knowledge only in virtue of one's knowing that perception is reliable. Knowledge of the premises of the track-record argument, says the critic, depends on knowledge of the conclusion, that perception is reliable. If this is so, then the premises of the track-record argument do not make the conclusion knowledge. Rather, they derive their positive epistemic status from one's knowing the conclusion. On this view of the nature of perceptual knowledge, a track-record argument would again seem unable to yield knowledge of the conclusion.

In response, one might argue that this objection rests on a mistaken view about the nature of perceptual knowledge. Perceptual knowledge, one might argue, requires that perception be reliable, but it does not depend on one's knowing that perception is reliable. In other words, *S*'s having perceptual knowledge that *p* requires that *S*'s perceptual belief that *p* be reliably formed, but it does not require that *S* know either that perception is reliable or that his perceptual belief that *p* is reliably formed. One might note that young children and brute animals can have much in the way of perceptual knowledge without knowing much about perception. They might even be unable to form the metabelief that perception is reliable. Indeed, one might maintain that perceptual beliefs are instances of immediate knowledge, and that they do not depend for their justification on any other belief.

ALSTON ON EPISTEMIC CIRCULARITY

William Alston, who has addressed the issue of epistemic circularity with both subtlety and care, finds that epistemic circularity does not always render an argument useless for justifying or establishing its conclusion. In "A 'Doxastic Practice' Approach to Epistemology," he writes, "Provided I can *be* justified in certain perceptual beliefs

without already being *justified* in supposing that perception is reliable, I can legitimately use perceptual beliefs in an argument for the reliability of sense perception" (1989, p. 3). Still, Alston himself worries that track-record arguments are not sufficiently discriminating. Part of the worry is that someone with clearly unreliable ways of forming beliefs could produce a track-record argument comparable to the simple track-record argument for perception given above. Imagine a crystal-ball gazer who reasons as follows:

> At t_1, I formed the belief that p on the basis of crystal-ball gazing, and p is true.
>
> At t_2, I formed the belief that q on the basis of crystal-ball gazing, and q is true.
>
> At t_3, I formed the belief that r on the basis of crystal-ball gazing, and r is true.
>
> And so on.
>
> Therefore, crystal-ball gazing is a reliable source of belief.

If the gazer is asked how he knows that p is true, he will reply that he knows it on the basis of gazing into his crystal ball. Alston worries that if we allow the use of epistemically circular arguments, then clearly unreliable sources of belief can be supported by such reasoning. In particular, it appears that the gazer's beliefs about the reliability of gazing would then be on a par with our beliefs about the reliability of perception. Alston believes that we need to try a different approach.

He argues that it is rational for us to form beliefs on the basis of certain sources such as perception and memory. The argument goes roughly as follows: (1) Many of our doxastic practices, our ways of forming beliefs, including perception and memory, are firmly established. (2) It does not seem to be in our power easily to avoid forming beliefs on the basis of these practices. (3) Moreover, even if there are alternative ways of forming beliefs, the very same problems of epistemic circularity that beset our attempts to support the reliability of our current practices would also confront these alternatives. (4) Therefore, it is rational for us to continue forming beliefs as we do, such as on the basis of perception and memory. But how does the fact that it is rational to continue to engage in these doxastic practices support the belief that they are reliable? Alston's view is that in taking it to be rational to form beliefs on the basis of our firmly established practices, I "commit" myself to judging that those ways of forming beliefs are reliable. I cannot reasonably judge that it is rational for me to form beliefs in those ways and deny that those ways of forming beliefs are reliable.

Alston's response to the problem of circularity is controversial. Some critics object that Alston's argument from firmly established doxastic practices is itself epistemically circular. Consider the claims that memory and introspection are firmly established practices and that one cannot easily avoid forming such beliefs. That certainly is true. But how does one know that? Clearly, one knows it on the basis of memory and introspection. Again, it does seem rational to form beliefs on the basis of reason. But in arguing that it is rational for one to form beliefs in that way, one must use reason itself. It does not seem, then, that Alston's strategy of appealing to our firmly established doxastic practices avoids epistemic circularity. If the track-record argument is unacceptable because it is epistemically circular, how would Alston's reasoning be any better? Furthermore, Alston worries that some clearly unreliable ways of forming beliefs can be supported if we allow epistemically circular arguments. How, critics object, does appeal to beliefs about firmly established practices help? Could not the gazer look into his crystal ball, form the belief that gazing is firmly established, and construct an argument analogous to Alston's? In short, if one problem with track-record and epistemically circular arguments is that clearly unreliable sources can be supported by them, the same seems true about arguments that appeal to beliefs about what is firmly established.

SOSA ON EPISTEMIC CIRCULARITY

Ernest Sosa (1994) holds that epistemic circularity is unavoidable if reflection on the reliability of our sources of belief is pushed far enough. In some cases, one might be able to support the belief that one source of beliefs is reliable by appealing to beliefs from another source. But we can always ask how we know that the second source is reliable. At some point, when reflection is pushed far enough, we cannot support the reliability of our sources except by appealing to beliefs that are the output of those sources. But if epistemic circularity is ultimately unavoidable, Sosa denies that it is necessarily vicious. Suppose, says Sosa, that *W* includes all our ways of forming beliefs, encompassing perception, memory, reasoning, etc. Suppose further (i) that *W* is reliable, indeed, that in our circumstances and with our nature, it is the most overall reliable way of forming beliefs we could have; (ii) that we are right in our description of W: it is exactly W that we use *W* in forming beliefs; and (iii) that we believe that *W* is reliable. Here our belief that *W* is reliable is formed by

means of *W* and is true. Sosa asks, how could we possibly improve our epistemic situation? Our belief that *W* is reliable is based on *W* itself, but Sosa does not see that there is anything epistemically vicious or unacceptable about our belief that *W* is reliable. Since our belief that *W* is reliable is formed on the basis of *W*, we have not avoided epistemic circularity, but in what way does our belief fall short epistemically? Recalling Descartes's initial, tentative reasoning concerning the truth of what he clearly and distinctly perceives, Sosa suggests that we might reason in a similar way for the reliability of other sources. Consider the following reasoning: (1) I know that there is a hand in front of me. (2) The best explanation for this knowledge is that I perceive the fact that there is a hand in front of me. (3) But perception could not be the source of such knowledge if it were generally unreliable. (4) Therefore, perception is not generally unreliable.

Sosa asks what is supposed to be so bad about epistemic circularity. Alston worries that if we allow epistemically circular arguments, then someone could give arguments in favor of their unreliable ways of forming beliefs that are analogous to those we might give in favor of perception and memory. Sosa grants that the crystal-ball gazer, for example, could construct arguments analogous to the track-record argument, and he notes that the gazer might also appeal to his crystal ball and "see" that gazing is a firmly established practice and thus is rationally engaged in. Sosa concedes that the gazer's belief in the reliability of his way of forming beliefs might cohere with his other beliefs and, more generally, that someone could have a coherent, yet false, view about the reliability of his sources. Yet Sosa denies that this fact puts the gazer's beliefs about the reliability of his doxastic practices on a par with our own. The fact that beliefs cohere with one another might provide some degree of epistemic justification, but for Sosa, it is not the only thing relevant to the epistemic status of belief. What makes our view about our doxastic practices epistemically superior is the fact that it was formed on the basis of reliable sources or intellectual virtues. Thus, our beliefs about the reliability of memory, introspection, perception, and reason are epistemically superior to the beliefs of the gazer about crystal-ball gazing in virtue of the fact that our beliefs are based on reliable sources or intellectual virtues, whereas the gazer's are not.

See also Alston, William P.; Belief; Cartesianism; Chisholm, Roderick; Criteriology; Descartes, René; Epistemology, History of; Hume, David; Locke, John; Moore, George Edward; Reid, Thomas; Sosa, Ernest.

Bibliography

Alston, William. "A 'Doxastic Practice' Approach to Epistemology." In *Knowledge and Skepticism*, edited by Marjorie Clay and Keith Lehrer. Boulder, CO: Westview Press, 1989.

Alston, William. "Epistemic Circularity." *Philosophy and Phenomenological Research* 47 (1) (1986): 1–30.

Alston, William. *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press, 1991.

Amico, Robert P. *The Problem of the Criterion*. Lanham, MD: Rowman and Littlefield, 1993.

Chisholm, Roderick. *The Problem of the Criterion*. Milwaukee, WI: Marquette University Press, 1973.

Cohen, Stewart. "Basic Knowledge and the Problem of Easy Knowledge." *Philosophy and Phenomenological Research* 65 (2) (2002): 309–329.

Descartes, René. *Meditations on First Philosophy*. 3rd ed. Translated by Donald Cress. Indianapolis, IN: Hackett, 1993.

Fumerton, Richard. *Metaepistemology and Skepticism*. Lanham, MD: Rowman and Littlefield, 1995.

Reid, Thomas. *Essays on the Intellectual Powers of Man*. Cambridge, MA: M.I.T. Press, 1969. Especially essay 6, chap. 5.

Sosa, Ernest. "Philosophical Scepticism and Epistemic Circularity." *Proceedings of the Aristotelian Society*, supp. vol. 68 (1994): 263–290.

Sosa, Ernest. "Reflective Knowledge in the Best Circles." *Journal of Philosophy* 94 (8) (1997): 410–430.

Van Cleve, James. "Foundationalism, Epistemic Principles, and the Cartesian Circle." *Philosophical Review* 88 (1979): 55–91.

Van Cleve, James. "Reliability, Justification, and the Problem of Induction." In *Causation and Causal Theories*. Midwest Studies in Philosophy, no. 9, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. Minneapolis: University of Minnesota Press, 1984.

Noah M. Lemos (2005)

EPISTEMOLOGY, HISTORY OF

Epistemology, or the theory of knowledge, is that branch of philosophy which is concerned with the nature and scope of knowledge, its presuppositions and basis, and the general reliability of claims to knowledge. The pre-Socratic philosophers, the first philosophers in the Western tradition, did not give any fundamental attention to this branch of philosophy, for they were primarily concerned with the nature and possibility of change. They took it for granted that knowledge of nature was possible, although some of them suggested that knowledge of the structure of reality could better come from some sources than from others. Thus, Heraclitus emphasized the use of the senses, and Parmenides in effect stressed the role of reason. But none of them doubted that knowledge of

reality was possible. It was not until the fifth century BCE that such doubts began to emerge, and the Sophists were chiefly responsible for them.

During the fifth century BCE human practices and institutions came under critical examination for the first time. Numerous things that had previously been thought to be part of nature were seen not to be. Thus, a general antithesis was drawn between nature and human convention or custom, and the question of where the line was to be drawn between them arose. The Sophists asked how much of what we think we know about nature is really an objective part of it and how much is contributed by the human mind. Indeed, do we have any knowledge of nature as it really is? Protagoras, for example, seems to have held, if Plato's report is to be believed, that everything is as it appears to a man, that appearances are the only reality. This is the meaning, or part of it, of his famous dictum "Man is the measure of all things." Gorgias was, if anything, more radical, claiming that there was no such thing as reality, that if there were, we could not know of it, and that even if we could know of it, we could not communicate our knowledge of it.

This general skepticism led to the beginning of epistemology as it has been traditionally known—the attempt to justify the claim that knowledge is possible and to assess the part played by the senses and reason in the acquisition of knowledge. Before Plato, Democritus, the Greek atomist, had already drawn a distinction between those properties ordinarily attributed to things which, in his view, really belong to them—for example, size and shape—and those which, as he put it, are a matter of convention (*nomos*), a function of the mind—for example, color. It was Plato, however, who can be said to be the real originator of epistemology, for he attempted to deal with the basic questions: What is knowledge? Where is knowledge generally found, and how much of what we ordinarily think we know is really knowledge? Do the senses provide knowledge? Can reason supply knowledge? What is the relation between knowledge and true belief?

THE NATURE OF EPISTEMOLOGY

Epistemology differs from psychology in that it is not concerned with why people hold the beliefs that they do or with the ways in which they come to hold them. Psychologists can, in principle, give explanations of why people hold the beliefs they do, but they are not necessarily competent, nor is it their province, to say whether the beliefs are based on good grounds or whether they are sound. The answer to these questions must be sought from those who are experts within the branches of knowledge from which the beliefs are drawn. The mathematician can give the grounds for believing in the validity of Pythagoras's theorem, the physicist can give the grounds for believing in, say, the indeterminacy principle, and an ordinary but reliable witness can provide the grounds for believing in the occurrence of an accident. Normally, when the beliefs are true and the grounds sufficient, it is permissible to claim knowledge, and whether a particular truth can be said to be known may be determined by reference to the grounds that are appropriate to the field from which the truth is drawn. The epistemologist, however, is concerned not with whether or how we can be said to know some particular truth but with whether we are justified in claiming knowledge of some whole class of truths, or, indeed, whether knowledge is possible at all. The questions that he asks are therefore general in a way that questions asked within some one branch of knowledge are not.

ROLE OF SKEPTICISM. To characterize the questions asked by the epistemologist as extremely general is not, however, sufficient. Interest in very general questions of this sort and in the nature of general concepts is typical of philosophy as a whole. What distinguishes epistemology other than the fact that its interests center on the concept of knowledge? When a philosopher asks whether something is possible, the question must be set against the consideration that this thing may not be possible. It must be set against a general skepticism concerning the matter in question. To be called upon to justify the possibility of knowledge or of certain kinds of knowledge makes sense only on the supposition that it or they may not be possible. It is no coincidence that epistemology began in the context of a form of the Sophists' general skepticism about knowledge, for until such doubts had been raised, the possibility of knowledge was bound to be taken for granted. Once the doubts had been raised, they had to be answered. How they were to be answered depended on the nature of the doubts and on the degree to which any particular philosopher was susceptible to them.

Views on the nature of knowledge. Perhaps the most characteristic form of skepticism about knowledge has been based upon the premise that we ought not to claim knowledge about anything unless we are absolutely sure about it, unless there is no possibility of our being wrong. Once given this, it is possible to point out that it is at least logically possible to be wrong about most, if not all, the things that we ordinarily claim to know. Philosophers who have been impressed by this argument have generally tried to show that there are at least some things that we

can claim to know, about which we cannot be wrong. Even so, most of the things that we normally think we know cannot, on this view, be said to be known at all. This consequence can be mitigated, although not removed, if it can be shown that the things accepted as known in the strict sense give reasons for believing the things that we normally take ourselves to know. Philosophers who have taken this course have differed both on what this "certain knowledge" is and on how it is connected with what we ordinarily suppose ourselves to know. Rationalists have generally attempted to show that the primary truths that constitute this certain knowledge are related to other truths somewhat as the axioms of a formal or geometrical system are related to the theorems.

Empiricists, on the other hand, have taken the view that the truths which constitute ordinary knowledge can be constructed out of the primary truths, as a building is built up from its foundations. They have differed again on the nature of the primary truths themselves. Rationalists have looked for them among the deliverances of reason, whereas empiricists have claimed that sense experience alone can provide such truths. Other philosophers have accepted part of the skeptical argument to the extent of denying the status of knowledge to some class of truths, reserving that status for some privileged class. Plato is a case in point in that for at least part of his life he maintained that sense experience never provides knowledge at all, this being reserved for a kind of awareness of or acquaintance with a world of quite distinct entities called Forms. In respect to the world of sense experience we have only opinion or supposition. This view of sense experience has not been uncommon among philosophers.

The concept of knowledge. A quite different way of dealing with the skeptical argument would be to question the initial premise that knowledge requires absolute certainty. One would not normally claim knowledge about something unless one were sure about it, but that is very different from asserting that a man could not be said to know something unless what he claimed to know was absolutely certain. Knowledge does not actually require this; it requires only that there be the best of grounds for what is claimed. To say this is to say something about what knowledge is, about the concept of knowledge itself. Hence, the skeptical arguments and the answers to them are not entirely independent of the conceptual question about the nature of knowledge. An understanding of the concept of knowledge is a prerequisite of embarking upon any attempt to answer other epistemological questions. Most philosophers have had something to say about the nature of knowledge, although many have taken its nature for granted. From this have stemmed a number of traditional epistemological difficulties.

GREEK PHILOSOPHY

PLATO. Plato (c. 428–347 BCE) was influenced by several views—the moral teaching and philosophical practices of his master Socrates, the views of the Sophists already mentioned, and such pre-Socratic views about the nature of reality as the Heraclitean view that the sensible world is in a state of flux and the Eleatic view that reality is one and unchanging. He came to hold that reality cannot be changing or imperfect and that it must therefore consist of a world of Forms or Ideas independent of the sensible world. The exact reasons why Plato postulated a world of Forms are not altogether clear. But probably, as Aristotle says, he was influenced by Socrates' search for the essences of, for example, moral virtues. But because justice, for instance, is never found in this world in a proper and perfect form, he postulated its separate, ideal existence in order to provide the standard by which sensible instances of justice might be judged. The Forms might be known by reason, not by the senses. Whether there was a Form for every sensible particular is arguable, with respect to Plato's earlier philosophy. However, it is clear at any rate that by the time he came to write the *Timaeus* he believed this to be the case. Thus, in the first place the Forms were probably standards or exemplars of which sensible things were imperfect copies. At the same time, however, they functioned as universals, entities meant to explain how it is possible to think generally about things of one kind and how it becomes possible to attach a meaning to common names. The fact that the Forms had to fill both these roles gave rise to certain logical difficulties which Plato himself pointed out in the dialogue *Parmenides* and which he tried to deal with in the later dialogues. The Forms were always objects of knowledge and, in his earlier thought, the *only* objects of knowledge. Sensible things were, in his view, objects of opinion only.

Knowledge and true belief. The distinction between knowledge and true belief is first made by Plato in the *Meno* in the context of another Platonic epistemological doctrine—the theory of recollection (anamnesis). In this dialogue Socrates claims to elicit from a boy without instruction the answer to a geometrical problem. Since the problem is a geometrical one, it is one that cannot be answered by an appeal to the senses. Socrates therefore claims that he is enabling the boy to recollect something that he had known in a previous existence. He maintains that it is a doctrine well known to priests and poets that

the soul has long ago experienced all things in its various existences. Hence, in a sense the soul knows all things, but because it has forgotten them, it has to be reminded of them. The example suggests that Plato may intend the doctrine, at least in part, as an explanation of our knowledge of a priori truths. Indeed, in the *Phaedo* he uses the doctrine to explain how we see things as instances of the Forms: Sensible things remind the soul of what it already knows and what it cannot know from sense experience—the perfect Forms. The *Meno* does not claim so much. Indeed, Plato goes on to suggest that merely arriving at the right answer to a problem may not constitute knowledge but only true belief. Knowledge requires an ability to give the grounds (logoi) on which the answer rests. Nevertheless, Plato says, true belief may sometimes be, in its practical effects, as good as knowledge.

This distinction between knowledge and true belief is retained by Plato in later dialogues, although he is not always so charitable to belief as such. At the end of Book 5 of the *Republic* Plato begins a long argument, involving the famous similes of the sun, line, and cave to show how the soul may be drawn up by education to a true knowledge of the Forms, the final stage in the process that Plato calls dialectic. At the outset Plato makes a threefold distinction between knowledge, ignorance, and an intermediate state that he calls belief. Each of these states of mind has, he says, an object. The object of knowledge is what exists; the object of ignorance is, paradoxically, what does not exist; and the object of belief is that which is between existence and nonexistence. The last seems to be identified with the sensible world. Plato appears to think of these states of mind as forms of acquaintance with some kind of object, although the allocation of the objects in question is puzzling on any account. He rejects the identification of knowledge and belief on the grounds that belief is liable to error, whereas knowledge can never be. His conception of knowledge is thus a strict one.

Knowledge and sense perception. Plato's reasons for maintaining that we cannot have knowledge of the sensible world are that we should be in error if we attributed properties to sensible things absolutely. A thing is beautiful, heavy, or good only in relation to other things; hence, Plato concludes, nothing is really beautiful, heavy, or good except the standards of Beauty, heaviness, and Goodness themselves, and they cannot be sensible things. When we judge that sensible things are beautiful, heavy, or good, we are in error and cannot therefore be said to have knowledge.

There are two objections to this argument. First, if we realize that terms like "beautiful" are relative terms, we shall not necessarily be in error in saying that sensible things are beautiful; error will arise only if we are tempted to think that they are beautiful absolutely. Second, not all terms are relative in this way; "red," for example, is not. Perhaps Plato eventually took account of both of these points, but it seems clear that he tried to deal with the second by reference to the Heraclitean doctrine that the sensible world is in a state of flux. If this doctrine is true, it is impossible to attribute any property to sensible things unequivocally. This position is put forward in the *Cratylus* and most fully expressed in the *Timaeus*. It depends, of course, on whether the Heraclitean doctrine is true and free from unpalatable consequences. It is examined in Plato's most extensive consideration of knowledge, the *Theaetetus*, a dialogue probably written at about the same time as the *Parmenides* with its criticism of the traditional theory of Forms.

In the *Theaetetus* Socrates engages in a discussion with a young mathematician, Theaetetus, concerning the nature of knowledge. Theaetetus first answers the question "What is knowledge?" in a manner typical of the dialogues, by giving examples of knowledge, but is then prevailed upon to give the answer that knowledge is esthesis (perception or sensation; the Greek word from which it comes is ambiguous). This view is identified with that of Protagoras, the Sophist, to the effect that everything is what it seems to a man and that esthesis is of what is and must be infallible—that is, what seems to a man is so and cannot be wrong. If knowledge is esthesis, it is thus an incorrigible awareness of something purely relative to the perceiver. Socrates enlarges on this view, indicating the extent to which our judgments about empirical things are relative in this way. Finally, the point of view is made absolutely general by the introduction of the Heraclitean doctrine of the flux. The joint effect of the doctrines is that all judgments about empirical things are relative—the classic Platonic point of view. This conclusion is reinforced by reference to various versions of what has become known as the argument from illusion—an argument stating that we cannot be said to perceive the real properties of things because of the possibility of illusion or because of the causal aspects of perception itself.

Having expounded the view that all empirical judgments are relative, Socrates sets out to refute it. He refutes the pure Protagorean view that appearances are the only reality by pointing out that there are acknowledged experts in different fields of knowledge concerning objective phenomena; moreover, the Protagorean view is in essence self-refuting because no absolute truth can be claimed for it on its own terms. He refutes the Heraclitean

doctrine of flux by claiming that if it were true, it would be impossible to say anything about anything, a consideration which he treats as a *reductio ad absurdum*. This section of the dialogue ends with the consideration that certain properties of things—the existence of things, their identity with themselves, and their difference from other things—are ascribable to them only as a result of a judgment by the mind. Knowledge of the sensible world cannot therefore be a simple matter of having sense impressions, of esthesis in the sense specified; it must also involve judgment by the mind.

Possibility of false beliefs. In the next section Theaetetus suggests that knowledge consists of true judgment or belief. This suggestion is eventually refuted by the consideration that it is possible to believe something truly when one's grounds are insufficient. Most of the section, however, is given up to a discussion of false belief or judgment, for Socrates wonders whether this is possible. False belief cannot be a belief in what is not, for, as Parmenides showed, there is no such thing as what is not. On the other hand, if false belief consists in erroneously taking one thing for another, it is difficult to see how it is possible, for the believer must know one or both of the things in question if he is to be in the position of taking one for the other. However, if he knows at least one of the things, how can he mistakenly take it for the other? Socrates considers various possibilities, but the only cases in which he will allow the possibility of error are those in which a man fails to recognize something correctly because he has fitted the wrong sense impression to the wrong memory impression. (It is in the context of this discussion that Socrates introduces the analogy of the wax and the seal to illustrate the nature of the mind; the mind literally receives impressions from things outside it.) Since this account will not cope with judgments like 7 + 5 = 11, where there is no question of erroneous recognition of a sensible thing, Socrates introduces another analogy, likening the mind to an aviary, with pieces of knowledge represented by the birds. A man may know something in the sense that the idea of it is in his mind as a bird is in the aviary, but he may not have it at hand. That is, he may know it implicitly but not explicitly. Even here the original difficulties recur, however. How can he mistake an explicit piece of knowledge for something else?

The difficulties in this section of the dialogue depend upon construing errors of judgment as mistakes of identity and equating knowledge with direct awareness. The mistakes allowed by Plato are not strictly mistakes of identity but mistakes in matching one thing with another, the sense impression with the memory impression. Plato returns to the matter in the *Sophist*, where he tries to provide a new logical analysis of the nature of judgment. He distinguishes between judgments of identity and existential judgments and probably between both of these and subject-predicate judgments (as they were later called). Judgments generally assert that one thing participates in another (at least the latter being a Form), but judgments of identity and existence assert the participation of something in certain especially important Forms—those of sameness and existence. These Forms are two of a list of five to which Plato gave special attention, the others being difference, rest, and motion. To say that something is identical with another thing is to say that it participates in the Form of sameness in relation to that other thing. Mistakes of judgment can arise over whether something really participates in another thing, and to that extent the difficulties of the *Theaetetus* are resolved, although it may be questioned whether Plato has given a really adequate account of judgment.

True beliefs and logoi. In the last section of the *Theaetetus*, it is suggested that knowledge may consist of true belief together with the giving of a logos. It will be remembered that knowledge was associated with the ability to give a logos in the *Meno*. Here Socrates recounts a "dream" according to which the elements of reality are perceptible but unknowable and without a logos, whereas the compounds which are formed from them are knowable and have a logos. The notion of a logos is a vague and possibly ambiguous one, but its connection with knowledge seems to imply at least that knowledge must be expressible in propositions. Socrates rejects the "dream view" (which may possibly be attributable to Antisthenes) on the ground that knowledge of compounds would not be possible unless there was already knowledge of the elements. Propositional knowledge must depend upon a knowledge by acquaintance of something, in Plato's view presumably a knowledge of the Forms. But what is the logos which, when added to belief, may turn it into knowledge? Three suggestions are considered. First, that a logos is simply the manifestation of the judgment in speech (a possible meaning of the word) is clearly insufficient. Second, that it consists of the recounting of the elements of the thing known is insufficient in that this may not actually amount to knowledge of how the elements are put together. Third, that it consists in the identification of the thing in question is rejected on the ground of circularity, for if being said to know something requires that one know the distinguishing mark of the thing in question, the account is manifestly circular. But nothing less than this is sufficient. The dialogue therefore ends inconclusively.

It is clear that in this dialogue Plato was working toward a view of knowledge which is not too far from the ordinary one. His thought, however, was never entirely free from ambivalence in this respect, and perhaps he never entirely lost his distrust of the senses. The ideal of knowledge as a kind of apprehension of a system of Forms remained. It was the task of the philosopher to investigate this system by means of dialectic, the techniques of logical division and classification. For Plato knowledge was always a state of mind and had to be accounted for accordingly. This presupposition lies behind the inconclusiveness of the *Theaetetus*. Yet most of the traditional epistemological problems arise in the course of Plato's argument, and it is worth attention for this reason alone.

ARISTOTLE. Aristotle (384–322 BCE) was not so affected by skeptical arguments as Plato was. He does, it is true, try to answer Protagoras (*Metaphysics* Γ 5ff.), and he does so in a way very much like that of Plato in the *Theaetetus*, by pointing to the standard case in each class of judgment. Even in his early (and now fragmentary) work the *Protrepticus*, he had emphasized the necessity of an appeal to the expert in deciding issues in any particular art or science. This remained his approach throughout his life. Aristotle's preoccupations with epistemology appear in two provinces in particular—in his theory of science and in his theory of the mind and its faculties. But his approach to epistemology was not so much the attempt to justify the claim that knowledge exists as the description of what knowledge and its presuppositions are.

Universals. Like Plato, Aristotle held that knowledge is always of the universal. Insofar as we can be said to know particular things, we know them as instances of a universal; we know the universal in the particular. But it must be emphasized that for Aristotle universals are inherent in particulars; he vehemently rejects the Platonic notion of a world of separate universals or Forms. (The only exceptions to the inherence of forms in matter and God and the most divine part of us, reason in the highest sense.) Knowledge therefore depends ultimately on the soul's or mind's reception of the forms of things. The soul itself, as he made clear in *De Anima*, is not a distinct, spiritual entity but the set of faculties possessed by the body insofar as it has organs to manifest them.

Means of knowledge. Sense perception is the receiving by the sense organ, the faculty of which is the respective sense, of the sensible form of a thing, as he puts it, without its matter. He also describes sense perception as an actualizing of the potentiality that the sense organ possesses as its faculty. It is not easy to see how this account of the matter can be worked out in detail. Granted that the hand becomes hot when it touches a hot object, what happens to the eye when we perceive color? Aristotle thinks that each sense is affected in a way peculiar to it, so that each sense has its own special sense object. The eye has color, the ear, sound; and so on. In addition to the special senses, there is a common sense, which has no sense organ peculiar to it. It is a faculty of all the sense organs or at least of those which are capable of perceiving the same qualities of objects; for example, size and shape are perceptible by both sight and touch. Aristotle speaks of both the special sensibles, such as color, and the common sensibles, such as shape, as essential to the respective senses. Apart from these there are the incidental sensibles, which are the things that possess these properties. Aristotle speaks of them as incidental sensibles because if we use our eyes, we are bound to see color; it is not essential that we see a particular object—to use Aristotle's example, the son of Diares. At any given time he may be the object of our vision, but he does not have to be. Some interpreters have spoken of these incidental sensibles as perceived indirectly, but there is no warrant for this interpretation in *De Anima*. As Aristotle points out, it is possible to see indirectly that sugar is sweet if we know that what we see directly is sweet. But this is quite different from perception of the incidental sensibles.

With this rather passive account of perception Aristotle gives a more active account when he stresses the role judgment plays in perception. Indeed, in the *Posterior Analytics* he speaks of the senses themselves as discriminative capacities. It is through such judgment that perceptual errors, such as mistaking the identity of a thing, can arise. Aristotle tends to say that when a sense is concerned simply with its own special object—for instance, sight with something white—there is the least chance of error. On occasion, however, he seems to imply that here error is impossible because the reception of the form in this case is something purely passive, so that there is no question of judgment. His position is not altogether clear, and there may be some confusion in his mind on this matter.

The same combination of passive impression and active judgment can be found in Aristotle's account of such faculties as imagination and memory, and to some extent there is a parallel account of the intellect itself. The persistence of the exercise of the sense faculty after its actualization by a sense object leads to the setting up of images. But imagination cannot exist in the full sense

without the exercise of some form of judgment. Likewise, memory must depend not only on having images but also on a referring to the past.

It is in the account of the intellect that the parallel with sense perception comes out most clearly. There is, first, a reception of form, in this case not sensible form but intelligible form. This corresponds to having concepts. Second, there is the combination of concepts in judgment, and it is only here that the possibility of error arises. Because the higher faculties depend for their existence on the lower, the exercise of the intellect, which is in itself nothing but a mere faculty, depends on the prior exercise of sense perception. Hence, Aristotle says, the soul never thinks without an image. Owing to the influence of St. Thomas Aquinas, this has often been interpreted as the basis of empiricism. If the issue had been raised by Aristotle, the outcome might have been this doctrine, but it is not clear that he did raise it. To say that the activities of the intellect are always dependent on the workings of the lower faculties is not in itself to say that the only ideas we can have are provided by sense experience. Finally, Aristotle distinguishes between an active and a passive intellect. The intellect thus far described is the passive intellect; the active intellect, something purely actual and without potentiality, is necessary in order to make possible the actualization of the faculties of the soul. In Aristotle's thought it is given no other function.

What knowledge is. Since knowledge is concerned with the universal—with form—any knowledge which is expressible in judgment must consist of an apprehension of an essential connection between forms. To know something about a thing is to be able to subsume it under species and genus and thus to know what is essential to it. It is matter which is responsible for what is accidental, and matter is in the last analysis—as prime matter—unknowable. To a large extent Aristotle's conception of knowledge in the full sense—that is, scientific knowledge—is coincident with Plato's. For Aristotle knowledge implies order; sense experience without this is something less than knowledge. This notion of order or organization is akin to Plato's notion of logos. Similarly, the idea that knowledge consists of classification in terms of genus and species and thus of a charting of the essential connections between forms is akin to Plato's conception of dialectic as concerned with the structure of the world of Forms.

Knowledge and definition. The connection in Aristotle's thought between knowledge and classification in terms of genus and species also entails a connection between knowledge and definition, for definition itself is in terms of genus and species. Aristotle distinguishes between nominal and real definition, the first being designed to give knowledge of terms only, the second giving knowledge of the essence of the thing itself. The difference turns largely on the fact that giving the essence of the thing involves the explanation of its cause. Thus, Aristotle says that we think that we have knowledge in the primary sense when we can give the cause of the thing. To give the cause of a thing involves the demonstration of its essence from first principles, and this is the function of science. The first principles themselves can be known only by a form of intuition; one sees their truth in their instances. It is possible to explain the principles of one science in terms of another science, but this process must come to a stop somewhere. It is the mark of a foolish man, Aristotle says, to think that everything can be proved. Principles such as the law of contradiction, which are implied in all demonstration, can be proved only dialectically. A dialectical proof is one that starts not from necessary first principles but from what is commonly accepted. In this case the proof consists in getting the man who denies the law of contradiction to say something and then to show him that what he says implies the law; he is thus convicted by his own testimony.

Aristotle thus presents us with a concept of an ideal of scientific knowledge and gives some account of what is presupposed by it. The difference between knowledge and true belief is, in his account, presumably dependent on whether what is claimed is essentially and necessarily true, a part of science as he sees it. But Aristotle gives little in the way of a justification of the claim that knowledge is possible at all, for he clearly feels no need to do so. To that extent he is out of the main stream of epistemological thinkers.

HELLENISTIC PHILOSOPHERS. If anyone in the ancient world was an empiricist, it was Epicurus, the leading Greek atomist.

Epicurus. In the view of Epicurus (341–270 BCE) all knowledge resulted from contact with the atoms of which the soul is composed by atoms from outside. It is true that he said atoms could sometimes stimulate the soul directly without affecting the senses, providing humans with visions of the gods, but in general the senses had to be involved. Atoms affecting the sense organs produced sensations; mass stimulation of the sense organs resulted in a presentation or appearance (*phantasia*) to the soul. Sense experience in the more general sense occurs when an incoming presentation is fitted to a general conception or abstract idea, which itself results from repetition of sensations. This is the nearest thing to judgment in Epicu-

rus's system, and this is the most usual source of error. Epicurus insists that all sensations are true and that they are the ultimate standard to which we must refer all our judgments; they admit of no other check. Since they are the ultimate standard of judgment, there is no other source for a metaphysical theory about the world. This, it has often been pointed out, does not fit in very well with the claims that Epicurus makes in order to give an atomistic picture of the world.

It is not altogether clear what Epicurus meant in saying that all sensations are true, especially since he also maintained that *phantasmata*, dreams or the delusions of the mad, are true. In the context of the atomist conception of the physical basis of perception, however, the view seems to imply a doctrine, common in the history of the subject, that anything in the mind which is caused cannot be liable to error. In reality questions of truth and falsity do not arise in such circumstances, a consideration implicit in Epicurus's statement that sensations are without a logos (not the sort of thing to involve judgment), but the conclusion that error is impossible has frequently been drawn.

Stoics. The rival Stoic school was founded by Zeno (fl. c. 300 BCE), but the main figure was, perhaps, Chrysippus (c. 280–c. 204 BCE). The Stoics were also empiricists to a large extent, although there is doubt whether at least some of the school did not admit innate ideas. The central notion of Stoic epistemology was intuition or apprehension (*katalepsis*). This, as is put by the Skeptic critic Sextus Empiricus (c. 200 CE), was their standard of truth. Like the atomists the Stoics thought that things make impressions on the soul, although they differed from the atomists over which physical processes were involved. They made no suggestion, however, that these impressions were necessarily veridical. This was true only of those impressions that were clear and distinct (*enarges*). Whenever an impression is received in the soul, the soul has to register it by a process known as assent, but there cannot be said to be knowledge until there is apprehension, until the soul is gripped by the impression (*katalepsis* literally means "gripping"). When this apprehension can properly be said to exist is clearly open to question, and this was pointed out by the Skeptics. Hence, later Stoics were forced to say that apprehensive impressions were a guarantee of truth as long as there was no objection.

Skeptics. Meanwhile, the Skeptics were making attacks upon the dogmatic schools, as they called them. The general tendency of this school was to accept the doctrine of impressions and *phantasiae*, but to maintain that there was no ground for going beyond them. Thus, it was necessary to be content with appearances and not to seek for the hidden truth about reality. The arguments against dogmatism were probably unsystematized initially, but they were gradually put into order. Probably under Aenesidemus (first century BCE) a list of ten (or eight) arguments (tropes) was drawn up. Some of these were forms of the argument from illusion, stressing the possibility of illusion and error in order to suggest that there was no reason to think that we ever gain knowledge of the real truth about things. Perceptions, they said, are always relative to the circumstances, the perceiver, and so on. Hence, no *phantasia* is a criterion of truth, and nothing else can be. A later Skeptic, perhaps Agrippa (first century CE), systematized the arguments even further by constructing general forms of skeptical argument. The final form stated that because of the earlier arguments nothing could be known in itself but only in relation to other things; however, something could be known relatively to other things only if these other things could be known absolutely. Because this had already been shown to be impossible, nothing could be known.

Some Skeptics came to see that this conclusion, put so baldly, was too dogmatic. When Arcesilaus of the Academy (the New Academy of the third century BCE) was sufficiently influenced by Skepticism to claim that knowledge was impossible, perhaps claiming Socratic practice as a precedent, the Skeptics still thought that this view was a species of dogmatism. Carneades (214–129 BCE), a later academic who tried to meet the arguments of Chrysippus of the Stoa, not only maintained that absolute knowledge was impossible but also tried to substitute a theory of probability for it. He distinguished three grades of probability in respect to perceptions: (1) the simply probable, (2) the probable and confirmed by its consistency with its concomitants, and (3) the probable, confirmed and tested for inconsistency with the system to which it belongs. The last grade is science as we ordinarily know it. But even this would have been too much for the pure Skeptic. Skepticism as a system received its fullest expression in the works of Sextus Empiricus in the second century CE.

Neoplatonists. In the third century CE Platonism was revived by Plotinus, the founder of the Neoplatonic school. This was Platonism in its more mystical aspects, although Plotinus often uses Aristotelian notions, sometimes with a Platonic twist. The soul, as opposed to the body, is given preeminence, so that perception and knowledge are made a function of the soul. The soul has its own activities, which are manifested in perception and

memory; the body and its impressions are merely instruments for the soul to use. The main function of the soul qua intellect is to contemplate the Forms, above which is the supreme principle or entity, the One. Unity with the One is the soul's goal.

MEDIEVAL THOUGHT

It was Neoplatonism which, according to St. Augustine (354–430), brought about his salvation from Manichaeism in theology and skepticism in philosophy. Neoplatonism offered a supposedly positive doctrine in both metaphysics and epistemology and one which St. Augustine could largely accept, thus ignoring the other, heterodox views. St. Augustine's thought is therefore Neoplatonic in its essentials. As a result he took it for granted that knowledge—and, most importantly, knowledge of God—was possible, and he felt no further need to question this assumption. The same is true of most other medieval thinkers. Since philosophy was closely linked with theology, it was axiomatic that knowledge of God was possible in some sense, and skepticism about knowledge in general was rejected by an appeal to whatever philosophical system was thought best able to explain that knowledge. Insofar as there was argument, it was about the presuppositions and sources of knowledge, not about whether it existed.

Knowledge of a thing involves, it is commonly thought, knowledge of its general characteristics and, therefore, its subsumption under a universal. Medieval thinkers differed according to their philosophical tradition, according to whether they were Platonists or, after the Aristotelian revival in the thirteenth century, Aristotelians. But the main dispute was over theories of universals. Since the dispute had theological implication, it was heated. The argument had its source in certain questions put by Porphyry, a disciple of Plotinus, about the exact status of species and universals. These questions, the answers to which Porphyry thought were obscure, were discussed by Boethius in a commentary. The main schools of thought on the subject were the realists, conceptualists, and nominalists. Realists thought that universals had an objective existence, although their view of this existence depended on whether they were Platonists or Aristotelians. Conceptualists held that universals existed only as concepts in the mind; nominalists held that the only universal things were words. Such theories, however, were rarely found in their pure form.

REALIST THEORIES OF UNIVERSALS. The division among realists is best seen in the differences between St. Augustine and St. Thomas Aquinas (1225–1274).

Augustine. St. Augustine gave preeminence to the soul, in Neoplatonist fashion. In his view the soul has its own functions and is not directly influenced by the body. Perception is based on the impressions produced by the soul when the body is stimulated. Experience, however, involves inference, as the soul subsumes its impressions under concepts. To have such concepts is, for St. Augustine, to be aware of Forms in the Platonic sense, the one difference being that in his view the Forms are thoughts in the mind of God. Thus, universals have a real existence in the mind of God, and all knowledge, even sense knowledge, involves some awareness of God. There is an ascent from lower forms of knowledge, like perception, to higher forms, with knowledge of God at the peak.

Thomas Aquinas. The Aristotelian revival in the thirteenth century led to St. Thomas Aquinas's acceptance of a more Aristotelian point of view than Augustine's. Like Aristotle, Thomas rejected self-subsistent universals, maintaining that universality is a function of the mind. Nevertheless, there are real similarities between things because of their common form. Hence, species have more than a mere mental existence.

The Thomist theory of knowledge consists largely of an explanation of how knowledge of Forms is possible. When the senses are stimulated, the soul's potentiality is actualized; a sensory image, or *phantasma*, is set up, corresponding to the object of perception. But since the universal aspects of such objects can be apprehended only by the intellect, they must be transferred from the *phantasma* to the intellect. Indeed, *phantasmata* as such are not objects of awareness on our part. The mind is aware only of the universal aspects of things, not their particularity, which is available only to the senses. Something has to illuminate the *phantasmata* in order to make clear their universal characteristics. Thomas employs Aristotle's distinction between the active and passive intellects here. The active intellect abstracts the universal or species from the *phantasma*, and this is imposed upon the passive intellect as a concept, which is then verbalized. Concepts thus exist only as the result of an abstraction of the universal aspects of things, and the essence of Thomas's empiricism is that all knowledge depends on sense experience in this way. Even knowledge of self-evident truths, which Thomas admits, as well as knowledge of the essential nature of things, is in the last resort dependent on sense experience, and all our thoughts must be based on experience. Thomas can be looked on as the founder of

empiricism in the sense that he held that all the materials for knowledge come ultimately from experience and from nowhere else. Unlike the later philosopher John Locke, however, he did not set out to justify the doctrine in detail.

CONCEPTUALIST THEORIES OF UNIVERSALS. Although Thomas may be classed as a realist in his theory of universals because he maintained that there are objective similarities between things by virtue of their common form. He could not do so without the notion of concepts in the mind.

Abelard. Peter Abelard (1079–1142) had previously held a conceptualist theory of universals, emphasizing the extent to which universality is a function of the mind. He held that universals are really concepts (*sermones*) involved in judgments that particular things have something in common. They are arrived at by abstraction from particular things, by attending to features of things considered in themselves. Abelard even used the notion of a generic image—that is, an image not of any particular thing but, supposedly, of what is common to a whole class of things—in order to account for our ability to think of things generally. He was anxious to reject both realism and the contemporary nominalism held by Roscelin of Compiègne (d. c. 1125), who maintained that universals were just words or even names. He did not, however, deny that concepts had a basis in things. Hence, in a sense Thomist realism and Abelardian conceptualism are very much two sides of the same coin.

NOMINALIST THEORY OF UNIVERSALS. Similar considerations apply to the great nominalist thinker of the fourteenth century, William of Ockham.

Ockham. Even Ockham (c. 1285–1349) did not quite maintain that the only universal things are words, for he held that words are conventional signs corresponding to concepts that are natural signs of things. Universality lies in the sign-significate relation, in the fact that both words and concepts can be signs of a class of things. To the question "What are universals?" Ockham initially replied that they had only a logical existence; they were meanings, the contents of the mind when thinking generally. For Ockham the term *universal* was what he called a term of second intention. A first intention is the state of mind involved in the apprehension of particular things; a second intention is that involved in the apprehension of first intentions. The term *universal* thus picks out the content of our apprehension of our first-order apprehensions of a class of things. "Redness" is not the name of an entity but of the content of the relation which exists between the sign or concept "red" and particular red things. Ockham later took another step toward conceptualism by holding that universals had a mental, not just a logical, existence. He then held that universals are the concepts that the mind has and which are the natural substitutes for things themselves. Ockham probably took this step for reasons of economy, for in the earlier account universals were a sort of intermediary between the mind and particular things. The place given to generality remained the same, however. There was nothing general in the world; generality depended on the relationship between the mind and particular things.

The real novelty of Ockham's approach lay in his holding the view, new to medieval thought, that the mind itself could have apprehension of particular things. Thomas, for example, had denied this, holding that the mind could be concerned only with universal characteristics abstracted from *phantasmata*. The consequent gap between the mind and the senses inevitably involved a representative theory of perception; the mind was confronted only with the representatives of things. Ockham denied all this. He held that the mind could be concerned directly with the particular by means of intuitions. Intuitive knowledge is a direct knowledge of a thing or its existence. The senses provide an intuition of a thing's existence, and the intellect provides an intuition of its nature. John Duns Scotus (c. 1266–1308) had held that intuition of a particular was a necessary condition of the abstraction of the universal from it, but he had also held that this intuition must, in this life at any rate, be confused. Ockham denied this. In his view intuitions may be perfect or imperfect respectively, according to whether they are dependent merely on present experience or whether they also involve past experience. The possibility of imperfect intuitions, however, depends on the possibility of direct, perfect intuitions.

Although Ockham thought that this kind of direct knowledge does exist, he did not claim that all intuition was equally clear; clarity, moreover, was not always enough to guarantee truth. In the first respect, he claimed, as St. Augustine had done earlier, that we have clearer knowledge of our own mind than of other things. In the second he maintained that God can give us an intuition of something that does not in fact exist (a consideration which looked forward to René Descartes's suggestion that God might be a deceiver). This is not the natural course of things, however.

Much of Ockham's thought is in the Stoic tradition, and to some extent this can be said about Descartes, the

first of the rationalist thinkers of the seventeenth century. By this time, however, the questioning of the accepted points of view of the Middle Ages had led to increased skepticism. Descartes's theory of knowledge is therefore in the fullest sense the beginning of that "search for certainty" whose elements were found in Plato but had not been prominent after him.

SEVENTEENTH-CENTURY RATIONALISM

DESCARTES AND CARTESIANISM. The emergence of science during the Renaissance and the disputes that it produced led to a certain skepticism about claims to knowledge and to the search for a method, like that of science, which would determine the truth once and for all. Descartes (1596–1650) was the pioneer in this new tradition, and although his roots were in the Middle Ages, he was to a large extent an innovator. Being a mathematician of distinction, he saw the solution to problems of epistemology in the systematization of knowledge in geometrical form, although he did not carry out the full program himself. This involved starting from axioms whose truth was clear and distinct. He describes the ideal method in the second chapter of the *Discourse on Method* as (1) not to accept as true anything of which we have not a clear and distinct idea, (2) to analyze the problem, (3) to start from simple and certain thoughts and proceed from them to the more complex, and (4) to review the field so thoroughly that no considerations are omitted. Of what do we have clear and distinct ideas? To deal with this problem, Descartes employs the method of doubt—a form of skepticism. This method involves setting on one side anything that can be supposed false until one arrives at something that cannot be supposed false.

That there is a goal to this skepticism is, it might be objected, prejudged, for Descartes points to the fact that he has often been deceived to suggest that he may always be deceived. This conclusion is not, however, admissible, since to establish his premises, he must at least know that he has sometimes been deceived. The truth is that Descartes has a definite conception of what knowledge must be, and most of what we ordinarily call knowledge does not fit that conception. His approach is therefore not strictly that of the general skeptic. It consists in setting on one side anything that does not possess the mark of genuine knowledge, this mark being that we should have clear and distinct ideas of the thing in question. We have a clear idea of a thing when it is open to the mind, when we are clearly aware of it; an idea is also distinct when we have a full knowledge of the nature of its object and of the means whereby that object can be distinguished from other things. Many philosophers have believed that we do not have certain knowledge of empirical truths but that we do have it of mathematical truths. Descartes agrees to the extent that he maintains that we can have clear and distinct ideas of the objects of mathematics, but he also maintains that if God were a deceiver, he might have caused us to have false beliefs even here. Hence, it is at least a possible hypothesis that there is an archdeceiver who brings it about that I am mistaken in all my beliefs. Is there anything which is free from this contingency?

"Cogito ergo sum." The result of Descartes's inquiry into this matter is that there is at least one proposition which is indubitable in the sense that I cannot be wrong in maintaining it. This is the proposition "I think, therefore I exist" (*Cogito ergo sum*). Descartes is definite that this is not to be treated as an argument despite its form; it is an indubitable proposition. (In a sense Descartes had been anticipated in this by St. Augustine's "If I am mistaken, I exist" [*Si fallor sum*], but St. Augustine had not used the proposition for exactly the same purposes.) It is reasonably clear that I cannot deny either "I think" or "I exist" without absurdity, although this is not enough to show that the *cogito* is in any way a logical truth. Yet for Descartes it must have the kind of necessity that is generally attributed to logical truths; it must be logically impossible for the proposition to be false. Moreover, it must have content such that its truth entails the existence of something with a specific nature—namely, a spiritual or thinking substance which has certain ideas, particularly those of God and material objects. Only then can Descartes go on to justify belief in such objects.

In effect, therefore, Descartes says that I can doubt everything except that I doubt. Since doubting is a form of thinking, I cannot doubt that I think, and since thought requires a thinker, I cannot doubt the existence of myself as the thinker. It might be objected that even if there is no reason to reject this position, it has not been shown to be necessary. If I cannot doubt that I doubt, this may be a contingent matter, not a logical necessity. Once given the *cogito*, however, Descartes can go on to use it as the premise of an argument whose outcome will be the justification of our belief in a world independent of ourselves.

The status of perception. Granted that we have ideas of a world of material things, what prevents those ideas being mere figments of the imagination? Ideas in themselves are purely mental entities (although Descartes is never very clear about their exact nature); they may or may not represent the things they purport to represent.

Ideas can be innate, adventitious, or fictitious. To say that they are adventitious is to say that they come from things outside us; to say that they are fictitious is to say that they are produced by the mind itself. Innate ideas are a priori, inborn. Which ideas these are, if any, may be disputed. But, at all events, our ideas of material things are clearly not innate. Why, however, are they not merely fictitious?

To say that an idea is fictitious is not to say that it is impossible for it to be an idea of something objective. To some of our ideas perhaps nothing could possibly correspond; these would be logical impossibilities and would have no objective validity in Cartesian terms. For an idea to have objective validity, the reality in it must be caused by something that has the same reality, either formally or eminently, in itself. A machine, to use Descartes's example, may be formally the cause of a man's idea of it; his idea may be a copy of the machine, the two having the same form. But if the man conceives of the machine himself, then he or his mind is eminently the cause of his idea; the idea is not a mere copy of its cause, for the source of the idea is something higher. If, then, our ideas of material things are to be objectively valid or have objective reality, they must either be copies of actual material things or be produced by something higher. If they are produced by something higher, they were produced either by our minds or by God. To show that our ideas of material things do correspond to those material things, Descartes has to show that they are not produced in this way either by our own minds alone or by God.

Now, ideas in themselves, Descartes maintains, cannot be strictly true or false; it is the use we make of them that can be true or false. Hence, truth and error are functions of judgment. Nevertheless, we have a natural disposition to believe that our ideas are veridical. In Meditation III, Descartes points to this natural disposition and to the fact that our perceptions do not depend on the will as reason for believing in the veridicality of our perceptions, although he rejects these considerations as insufficient. In Meditation VI, however, he has recourse to the same considerations, although they must now be viewed in the context of the view that perception is a faculty of the mind plus body and does not express the essence of the mind alone, which is concerned solely with thinking. The passivity of our perceptual ideas thus seems to be invoked in order to reject the notion that our ideas of material objects could be a product of our own minds. Ideas of material objects, however, might still be caused by God. Yet if this were true and if nothing answered to those ideas, God would be a deceiver, for he would be giving us a natural disposition to believe in the existence of things which do not exist in fact. God, Descartes maintains, is not a deceiver. This point is taken as axiomatic and provides the ultimate guarantee of our belief that we do perceive an objective world.

Existence of God. The existence of God is therefore a cardinal point in the chain of argument. Descartes produces two sets of arguments designed to demonstrate his existence as a necessary truth. The first argument, in Meditation III, is based on the same considerations about the causes of our ideas as those adduced in connection with ideas of material things. The idea of God, which Descartes takes to be objectively valid, could be produced only by something having the same reality formally or eminently in it. We, being inferior creatures, could not produce it, and it could not come from any other source except God himself. Hence, there must be a God. This argument is a version of the so-called Cosmological Argument.

The other argument, to be found in Meditation V, is a version of the Ontological Argument first invoked by St. Anselm. God must through his essence possess all positive attributes in perfection. Existence is a positive attribute; hence, God must exist. It is now generally recognized that neither argument is sufficient to demonstrate the necessary existence of God. However, the necessary existence of God must be demonstrated if the argument concerning the existence of material things is to have any cogency.

If Descartes's argument for the existence of God had been sound, he would have shown that all the reality in our ideas must be in their causes. More is required if it is to be shown that our ideas are, at least in some cases, copies of their causes. This is a problem for any representative or causal theory of perception, any theory that holds that our ideas and perceptions are mental entities which are, at best, only representatives of things outside them. From Descartes's point of view there is the general consideration that God is not a deceiver; any errors or illusions to which we are subject are the results of judgments we make because of the ideas we receive. This makes it incumbent on us, if we are to avoid error, to take due account of the clarity and distinctness of our ideas. We are right in judging that our ideas correspond to their causes only if those ideas are clear and distinct.

Primary and secondary qualities. In Descartes's opinion there is a big difference between primary qualities, such as figure, magnitude, and motion, and secondary qualities, such as color. Primary qualities can be known by an intuition of the intellect (*inspectio mentis*). Our ideas of them are clear and distinct because of the

part they play in mathematics, and in mathematics, therefore, the intellect can be regarded as having a proper knowledge of reality. This is not to say that we cannot make mistakes concerning the primary qualities of objects; judgment can be as liable to error here as elsewhere. However, since the ideas of them are clear and distinct, we have the assurance that in general objects do have such qualities. We have no such assurance in the case of secondary qualities. The intellect is not involved here, but since the senses were, Descartes maintains, provided only for the conservation of life, it does not matter very much whether our ideas of secondary qualities correspond to the actual qualities of objects.

It would not be generally admitted today that mathematics does provide the kind of knowledge of reality that Descartes requires. The question of the exact connection between mathematics and the world is a complicated one. Granted, however, that mathematical ideas have a precision not possessed by other ideas, it does not follow that we have a precise knowledge of any qualities of physical objects. For it remains an open question to what extent such ideas are applicable to physical objects. There is a genuine distinction between primary and secondary qualities in that the first are measurable in a way that the second are not. This, however, is not sufficient to justify the claim that knowledge of primary qualities is knowledge of reality in a way that knowledge of secondary qualities is not. In some places—for instance, *Principles*, Part IV, Section XI, and *Dioptric*, Section VI—Descartes tries to reinforce this view by arguing that since the effects in the brain caused by the stimulation of our senses possess only the properties of motion, figure, and extension, there is no means whereby we could come to apprehend any other properties of objects. There is a circularity in this argument, since its premise is that we do know of the primary qualities possessed by brain processes, whereas all ideas, being the effects of physical processes, should be in the same position.

In sum, Descartes's theory of knowledge is essentially one of a representative kind. It is based on the idea that the mind or soul, being something very different and distinct from the body, can have as its contents only ideas, which are, at best, representatives of physical things. The mind has its own activities, and its nature is to be active. Through these activities it can come to have knowledge of abstract mathematical truths. But all sense knowledge, as opposed to intellectual knowledge, can come only through the medium of ideas, and that these ideas correspond in any way to the physical objects presented to the senses is inevitably open to question. The justification of our belief that they do depends, in the long run, on the affirmation that there is a God and that he is not a deceiver. Descartes thinks that he can demonstrate that there is a God, taking as true by definition that he is no deceiver and that our natural disposition to trust our senses is therefore justified.

Occasionalism. Since Descartes conceived of the soul and the body as distinct substances with distinct essences—that of the soul being thought and that of the body extension—he was inevitably faced with the problem of how one could act on the other. He was never very clear on this point, although he came to insist that there must be some quasi-substantial union between the two. In some places—for example, the *Dioptric*—he speaks of brain processes giving "occasion" to the soul to have sensations or ideas. Some of his followers, who thereby became known as occasionalists, took up this notion and tried to explain the apparent link between soul and body by saying that God puts ideas into the soul on the occasion of the bodily processes. Arnold Geulincx (1624–1669) said that God puts the ideas there by means of the bodily processes; Nicolas Malebranche (1638–1715) said that God acts directly on the mind on the occasion of the bodily processes. Since Malebranche had leanings toward the views of St. Augustine, he interpreted this occasionalism in terms of the Augustinian doctrine that we know all things in God.

Malebranche. In other respects Malebranche tends to follow Descartes, although often with greater emphasis. He, too, insists that we have clear and distinct ideas of figure, extension, and movement, since these qualities, being conceivable in mathematical terms, are open to the intellect. He also lays great emphasis on our liability to error in anything connected with the senses, especially if we think that the senses provide us with knowledge of things as they actually are. In one respect, however, he adds a certain sophistication to the Cartesian distinction between the ideas or sensations that arise in the mind without any intervention on our part and the judgments that we make and which depend on our will. Sometimes what we see differs from what would be expected on the basis of the sensations resulting from the stimulation of our senses. We may see a thing in its right size, for example, although the actual pattern of stimulation received in the eyes is different; alternatively, we can be subject to illusion when the conditions of stimulation are not abnormal. In such circumstances we are not generally aware of making any judgment in order to correct the sensations received. Malebranche therefore distinguishes between natural judgments, or judgments of sense, and

free judgments. Free judgments depend on our will, but natural ones do not; they are, he says, a kind of complex sensation in that they do not depend on us. They are, he explains, made by God in us, in consequence of the laws of the union between soul and body. As judgments they can be true or false, but as sensations they may occur against our will and are certainly not due to our will. Malebranche is in an ambivalent position here, but his difficulties show a certain honesty.

SPINOZA. It has often been remarked that what makes the thought of Benedict (Baruch) de Spinoza (1632–1677) especially interesting is that it combines the quite different and, as generally conceived, quite disparate traditions of nominalism and extreme rationalism. In his nominalism he belongs to the tradition of William of Ockham and, more particularly, of Thomas Hobbes (1588–1679). Hobbes maintained that only names were universal. Although names were signs of images of things that constitute our conception of them, there was nothing universal in the conceptions themselves. Only the use to which names are put was universal, for they are taken as signs of many things. Hobbes used this to mount an aggressive attack on the paraphernalia of Scholasticism—essences, substance, and the like. He was a tough-minded mechanist who thought that reality consisted solely of corporeal bodies in motion.

Although he did not have the same motive, Spinoza was similarly opposed to the apparatus of Scholasticism. Indeed, he may have been influenced by Hobbes. In his *Ethics*, Part II, Spinoza held that as a result of the stimulation of our bodily senses by many things, confused, composite images arise, and it is these images that general words represent. There is no place for universals existing in things. Since these images may be set up differently in different men, "universal notions," as Spinoza calls them, may differ from man to man. The knowledge that is expressed by their means can only be confused. Spinoza is not content to leave knowledge there, however; he has the conception of knowledge of a much higher kind, and his working out of this conception is in effect the systematizing of Cartesianism. To make Cartesianism consistent, however, he had to change much in it.

Monism. Although according to Descartes's view the clarity and distinctness of an idea was a necessary condition of its truth, it was not a sufficient condition of its truth. It was always possible to raise the question of whether any particular idea did correspond to reality and, in particular, to its cause. This was a consequence of the dualism between the mind and its ideas and the physical world, a dualism inherent in Cartesianism. Spinoza replaced this dualism by a monism according to which the mental and the physical were two aspects of one thing—ultimately, God or Nature. In adopting this view, he was again carrying out the implications of Cartesianism, for Descartes had asserted that in the proper sense the concept of substance belongs only to God, for only God is self-subsistent, or *causa sui*. Hence, in Spinoza's view everything is a modification or mode of the one true substance, depending on God for its existence. The Cartesian distinction between mental and material substance becomes a distinction between the two infinite attributes of God in Spinoza's theory. Bodies are modifications of God qua extended, and minds are modifications of God qua thinking. They are not distinct things; they are merely parallel aspects of the one true substance. The order of ideas in the mind is the same as the order of things. Hence, the objects, or *ideata*, of ideas, insofar as these have reference to God, are always truly represented by them. No ideas can be absolutely false, and insofar as they refer to God, they must be true. Ideas can be considered false only from the point of view of what we ordinarily call a single human mind, not from the point of view of God.

Truth and falsity. Because everything can be deduced from the essence of God, everything is subject to necessity, and this applies to ideas as much as to everything else. There is no room in Spinoza's theory for the Cartesian distinction between ideas and the will; for him the will and the intellect are the same. Falsity cannot therefore arise from the exercise of our will in the employment of ideas. For Spinoza the exercise of the will in judgment is not something additional to having ideas, and he emphasizes that it is wrong to think of ideas as simple pictures that may or may not correspond to their objects. To have an idea of something and to make a judgment about it are one and the same, and it is by virtue of this composite conception that ideas, not only judgments, can be true or false. Considered as part of God's thinking, ideas cannot, of course, be false, for in that case they must always correspond, qua modifications of God as thinking, to the modifications of God as extended. But considered as ideas in a single human mind, they can represent their objects confusedly or inadequately. In having a confused idea of an object, we fail to see it as following necessarily from the nature of things. Such ideas, Spinoza says, are like consequences without premises.

In the *Ethics*, Part II, Definition 4, Spinoza defines an adequate idea as "an idea which, insofar as it is considered in itself without relation to an object, has all the proper-

ties or intrinsic marks of a true idea." If an idea is confused, it cannot be adequate. Therefore, truth must have an intrinsic criterion, not just the extrinsic criterion of the correspondence of an idea to its object. In other words, the clarity and distinctness that Descartes had looked to for the foundations of knowledge is the mark of every true idea. Truth and adequacy thus merge. For full truth an idea must be seen to be true, and this is possible only insofar as it is seen to follow from the nature of things. Truth, Spinoza says, is its own criterion. This is connected with another of Spinoza's epistemological views—that knowledge must ultimately be reflexive. Anyone who really knows something must know that he knows. If a man knows that something is necessarily so, he must know that he knows this, since the truth of what he knows must be manifest.

The doctrine of truth that Spinoza presents is commonly referred to as the coherence theory of truth, and it is normally associated with the doctrine of degrees of truth, knowledge, and reality. The distinctions between grades of reality that exist between the one true substance and its various modifications is paralleled by distinctions between kinds of knowledge. True knowledge, the having of adequate ideas, entails seeing things as following from the essence of God. Knowledge can be more or less inadequate or confused to the extent that a thing is not seen as following necessarily from that essence. Absolute truth consists in having adequate ideas, although every idea has some degree of truth since it must have a counterpart in the order of things. In other words, an idea, although necessarily true in some respect, has greater truth to the degree that it is adequate and to the degree that its object is seen as fitting in with the order of things. Like most versions of the coherence theory of truth, this is really not a theory of what truth is or what is meant in calling an idea or judgment true, but, rather, a theory of when or under what conditions an idea or judgment can be seen to be true. A judgment can be seen to be true if it coheres with the system of judgments that characterize reality. But coherence theorists tend to say that any judgment which can be seen to cohere with other judgments in this way is thereby "more true" than those which do not cohere. In effect, they tend to use the word *true* so that it is more or less equivalent to *verified*. A judgment that has a higher degree of verification by virtue of its coherence with other judgments is said to be ipso facto more true. The coherence theory of truth is thus a genuine epistemological theory, a theory of the conditions under which we can be said to know a proposition as true.

Kinds of knowledge. Spinoza distinguishes three grades of knowledge. Full knowledge Spinoza refers to as the third kind of knowledge and characterizes it as intuition. This kind of knowledge, he says in the *Ethics*, Part II, "proceeds from an adequate idea of the formal essence of certain attributes of God to an adequate knowledge of the essence of things." To have this knowledge is the goal of philosophy—to see things *sub specie aeternitatis*, as conforming to a kind of necessity. The right method in philosophy as set out in the *Treatise on the Correction of the Understanding* is to rid the mind of confused and inadequate ideas and to lead it to ideas which are adequate. It is significant that Spinoza calls this kind of knowledge intuition, because in its essence it consists of seeing the world as a coherent whole bound by necessary connections. Most rationalists have ended up with some such conception, and for them reason is inevitably second best. So it is for Spinoza.

What Spinoza calls reason is the second kind of knowledge, below intuition. This is best described by distinguishing it from the first kind of knowledge, which is knowledge derived from vague experience and is also called opinion or imagination. This corresponds roughly to sense experience. (Knowledge from hearsay, the fourth kind of knowledge, added in the *Treatise* to the bottom of the list, is of little importance for present purposes.) From sense experience we gain confused ideas of things without respect to their place in the general order of things. We may obtain knowledge of a similar status from signs—that is, from reading or hearing words which allow us to form ideas similar to those of the imagination. Both of these sources may lead to the setting up of the universal notions referred to earlier, notions that vary from person to person and cannot provide genuine knowledge. In the course of our experience, however, we may light upon notions that are common to all people, such notions as those of extension and other general attributes of reality. These notions Spinoza calls common notions, to be sharply distinguished from the universal notions already discussed. These common notions correspond to the ideas of primary qualities that Descartes had allowed to be clear and distinct because they were the objects of intellectual intuitions. For Spinoza, too, they provide the starting point of the sciences, and as such our ideas of them are adequate. They can be seen to be true of reality inasmuch as they reflect all-pervasive features of reality. It is for this reason that they are common to all humans; they are not subjective like the universal notions.

Reason or science thus consists in elaborating the essential features of the attributes of which we have common notions. Like Descartes, Spinoza conceived of science as based on the model of mathematics in general and geometry in particular. His conception of the right method in philosophy itself is modeled on the geometrical method. Indeed, Spinoza had tried to set out the Cartesian philosophy in a geometrical fashion according to the rules that Descartes had preached but had not practiced as well. Thus, although science, the systematization of knowledge, is ultimately derived from sense experience, it reflects the actual order of things more truly than experience does. Nevertheless, the goal of all knowledge is not just this systematization but the seeing of things as a whole, *sub specie aeternitatis*. For this reason intuition stands above reason.

Because the second and third kinds of knowledge involve adequate ideas, they cannot give rise to falsity. Sense experience alone can be the source of falsity. Through sense experience we can have only confused ideas, since ideas reflect particular modifications of reality in some finite respect, not in relation to the infinite attributes of God. Sense experience is ordinarily thought of as a passive form of knowledge, as opposed to forms of knowledge that demand the use of reason. This passivity, Spinoza thinks, is only a sign of the inadequacy of our ideas in this case. Activity on the part of the mind is, conversely, adequacy in its ideas. Spinoza points out that the human mind is part of the infinite intellect of God. Hence, "when we say that the human mind perceives this or that, we say nothing else than that God, not in so far as he is infinite, but in so far as he is explained through the nature of the human mind, or in so far as he constitutes the essence of the human mind, has this or that idea" (*Ethics*, Part II). Just as the ideas of sense experience are confused and inadequate only in relation to our mind, although when considered as God's ideas they are true, so the ideas of sense experience are passive in relation to our mind but are nevertheless part of God's active thoughts.

In sum, for Spinoza the goal of all knowledge is seeing the world as a single whole. On the way to this lies reason or science, which attempts to reveal things as subject to necessity by means of self-evident, necessary truths about things. All else, although not absolutely false, is the source of illusion. But, as in everything else in Spinoza, an adequate understanding of his theory of knowledge also involves an adequate understanding of his complete metaphysics or theory of reality. This is true of philosophers in general but never more so than in Spinoza's case.

LEIBNIZ. In many ways Gottfried Wilhelm Leibniz (1646–1716) is Spinoza with a strong dash of common sense. Spinoza's monism, especially its assertion of the necessity of things and the apparent consequence that free will is impossible, was anathema to Leibniz. In these respects Leibniz revolted against Spinoza, but in other respects he was very much like Spinoza. He, too, drew no distinction between the will and the intellect and made activity and passivity in the mind a function of the clarity and distinctness or otherwise of our ideas. Furthermore, although common sense told him that there must be a plurality of things, not just one, he had to conceive of each ultimate thing as a simple substance possessing all the properties of Spinoza's one substance.

Leibniz simple substances had to have a unity in plurality in that, although simple, each one had to be capable of reflecting the whole universe from its point of view. Since Leibniz took as axiomatic that in every true proposition the predicate is contained in the subject, everything that can be said about a substance is so because of the nature of that substance, and all the relations which it has to other things must arise from the nature of that substance and be internal to it. It is for this reason that every true substance must reflect the universe from its point of view and in this way be a microcosm of the macrocosm. The only thing, Leibniz thought, which could be both simple and capable of reflecting the universe in this way was something like the soul. In consequence, he postulated the existence of a plurality of simple substances, spiritual in nature, which he called monads. But for the monads, he said, Spinoza would be right.

Necessary and contingent truths. Since the properties of each monad were internal to it, it might be thought that Leibniz, like Spinoza, would have maintained that everything was a matter of necessity. However, although Leibniz maintained that all the properties of a substance are internal to it and thus follow from the nature of the substance, for other reasons he maintained a clear distinction between truths that are necessary in the logical sense and truths that are dependent on the facts. He was thus perhaps the first to draw a clear distinction between necessary or logical and contingent or factual truths. The first he called truths of reason, the second truths of fact. These truths had different principles as their basis. Truths of reason were dependent on the principle of contradiction, since their necessity turned on the fact that their denial would result in a contradiction. Leibniz thought that such truths, when their terms are defined, could be reduced to identical propositions of the form "*A* is *A*." The reduction to such identical propositions would

therefore proceed by means of chains of definitions. Mathematical truths are of this kind, and Leibniz was one of the first to seek a basis for mathematics in logic.

Truths of fact, on the other hand, could not be justified by reduction to identical propositions; their basis had to be found in a separate principle, the principle of sufficient reason. This principle received different formulations at different stages of Leibniz's thought. Insofar as it was meant to allow for the contingency of matters of fact while providing a rationale for them, Leibniz tended to formulate the principle by reference to the choice of God. In creating this universe, God could choose from a number of possible worlds each having a different order or structure. Since, as Leibniz thought, for various different reasons, the number of monads is infinite, the number of such possible worlds is also infinite. Any contingent truth about *this* world has for its justification the fact that in choosing this world, God chose it as the best of all possible worlds. The truth remains contingent because it is dependent on God's choice, but a sufficient reason for its truth is that God chose it as part of the best of all possible worlds. At other times, however, Leibniz's appeal to God's choice has fewer theological implications. For example, in his correspondence with Samuel Clarke, he tries to refute the idea of an infinite absolute space by saying that in such a space God would have no sufficient reason for putting the universe here rather than there. This means that there would be no way of telling where the universe was and that it would, in consequence, make no sense to speak of it as being in one place rather than another. This use of the principle of sufficient reason amounts to something like the use of the verification principle by logical positivists—the meaning of a proposition lies in the method of verification.

When Leibniz maintains, however, that in every true proposition the predicate is included in the subject, he seems to undermine the distinction between truths of reason and truths of fact. For this doctrine would make all propositions into what Immanuel Kant was later to call analytic propositions or judgments, propositions that are logically necessary. In consequence, it has been maintained (for example, by Louis Couturat) that in some of Leibniz's writings the principle of sufficient reason merely states that all true propositions are analytic, whereas the principle of contradiction states that all analytic propositions are true. In fact, this is probably a consequence of Leibniz's main views rather than a statement of it. If every proposition about a substance attributes to it a property that is part of its essence, then all such propositions must be analytic even if they are to be characterized as truths of fact on other grounds.

Leibniz accepted this conclusion but tried to evade the contradiction implicit in characterizing a truth of fact as analytic by explaining that the number of properties possessed by any substance must be infinitely great, as the points of view from which a thing can be regarded are endless. We, being finite creatures, cannot complete the analysis of any given substance. Hence, we cannot know for certain whether any given attribute really does belong to it; we cannot, without completing the analysis, even know whether it is possible for this substance to possess the property; it may be a contradiction to suppose it. God, being infinite, can complete the analysis, and so for him all propositions about things are analytic or logically necessary. We, on the other hand, can know only that if a proposition is true, it is necessarily true, but we cannot know for certain whether any given proposition is true. In our judgments about the truth of propositions, we have to depend on probabilities—that is, we have to estimate what reasons are sufficient for our conclusions. Thus, for us any judgment of fact is contingent. For God contingency enters only in that he has chosen what substances there should be, which of all possible worlds should exist. For him everything thereafter is necessary. Hence, the principle of sufficient reason comes into consideration in two related ways—first, in that it guides, without determining, God's choice of a world and, second, in that it guides our decision concerning which world God has chosen.

Perception and appetition of monads. So much for the logical consequences of Leibniz's point of view. Given the metaphysical system according to which there exist, an infinite number of spiritual entities or monads, other consequences also follow. According to Leibniz, every monad has perception and appetition—the apparent passive and active features of mental life. Since everything about a monad is internal to it, these features can indeed be only apparent. Appetition is that aspect of a monad responsible for internal change and development. No monad can affect or be affected by any other monad. A perception is any property of a monad that results from its development but that may reflect changes in other monads. (This use of the term *perception* is, although influenced by Leibniz's metaphysics, clearly very general, but its use was very general throughout the seventeenth century.) What may seem to be activity on the part of one monad with respect to another is really having distinct perceptions, whereas passivity is having confused perceptions. Here Leibniz sides with Spinoza.

Self-consciousness of monads. Leibniz's criterion of a distinct idea of a thing is the ability to list the characteristics which distinguish the thing from other things. This clearly involves a degree of self-consciousness, and this is possessed only by the monads constituting the human soul. Although all monads have perceptions in that other things are represented in them, not all have apperception. To have apperception, the monad must be conscious of what is involved in its perceptions, and those perceptions must therefore be distinct. The distinction between perception and apperception means that perceptions can sometimes be unconscious. Leibniz brings forward a number of arguments in support of this view, ranging from the argument that reflection upon perceptions must come to a stop with perceptions that are not self-conscious to the argument that there must be what he calls *petites perceptions*. When we hear the roar of the sea, he argues, we are not aware of hearing each little ripple despite the fact that the waves are made up of ripples. Since the overall perception must correspond in complexity to its object, he concludes that there must be perceptions corresponding to the ripples, and these little perceptions are therefore unconscious. This is not a psychological discovery but a philosophical analysis the premises of which are open to dispute. Like Descartes, Leibniz accepts the representative theory of perception in thinking that perception consists in having ideas which are, or may be, representative of objects. If this theory is rejected, Leibniz's argument about *petites perceptions* loses much of its force.

Error in perception. Just as Leibniz sides with Spinoza in maintaining that activity and passivity are to be explained in terms of the distinctness of our ideas, so he agrees with him, against Descartes, over the explanation of error. There is no room for the individual will in Leibniz's system. Appetition is only the impulse that provides the development of the monad's perceptions; it in no way corresponds to the will. Error is merely a matter of having confused ideas, and since these are correlated with passivity, the passive aspects of mental life—sense perception and the like—are the source of error.

Innate ideas. Yet although Leibniz can distinguish between ideas of perception and ideas of reason, or the understanding, it remains true that according to his view in a sense all ideas are innate. None is literally produced by things affecting our sense organs. Yet the distinction between ideas in terms of their clarity and distinctness does mean that it is possible to say that some ideas are what Kant called a priori—in no way derived from the senses. These are ideas such as those of mathematics, and Leibniz criticized his empiricist adversary Locke for failing to take sufficient account of these ideas. Indeed, to the empiricist principle that there is nothing in the intellect that was not first in the senses Leibniz replied, "Except the intellect itself."

Rationalism generally tends to emphasize the part played by the intellect in contradistinction to that played by the senses. It holds that real knowledge is that provided by the intellect, for only there is the certainty which knowledge requires. Moreover, it is by means of the disciplines that are peculiarly the province of the intellect that knowledge is to be obtained and preserved.

BRITISH EMPIRICISM

In general, empiricism stands in opposition to rationalism both in its views about the main source of our ideas and in its views concerning the source of true knowledge. Thus, it is often, historically speaking, a reaction against rationalism. The so-called British empiricists of the seventeenth and eighteenth centuries, however, were empiricists only in tendency. The first, Locke, was a complete empiricist concerning the source of our ideas, but he was often a rationalist in allowing other than empirical knowledge. Locke's new way of ideas, as it was called, was an attempt to show that all the materials for knowledge are derived from sense experience. Locke did not claim, however, that all knowledge was founded on experience in any other sense. George Berkeley, who carried on Locke's new way of ideas and even sharpened some of Locke's claims, especially on the subject of abstract ideas, was fundamentally a metaphysician with a special way of looking at the world. David Hume, the last of the trio, claimed to introduce the experimental method into philosophy, following in the steps of Newton, and of the three he had by far the best right to be counted an empiricist. Indeed, his empiricism led him extremely close to skepticism concerning a number of claims to knowledge; such skepticism, he believed, could be avoided only by "inattention" to philosophical issues. But all three of these philosophers were united in their opposition to any doctrine of innate ideas.

LOCKE. Book I of Locke's *Essay concerning Human Understanding* is devoted to an attack on the doctrine of innate ideas, and the positive doctrine begins only in Book II. At the outset Locke (1632–1704) had claimed that he was following the "historical plain method," the object of which was to "set down any measures of the certainty of our knowledge." This historical plain method consists in classifying our different ideas and plotting

their source as a prelude to an assessment of claims to different kinds of knowledge. Despite appearances this is not a psychological method; it is a method of philosophical analysis designed to discover the logical character of different ideas. In this way Locke distinguishes between ideas of sense and ideas of reflection. Ideas of reflection result from the operation of the mind itself upon ideas of the sense. There is no other source of ideas.

Locke also distinguishes between simple and complex ideas of both kinds, complex ideas being formed by the mind in compounding simple ideas. He seems to think that what it is to have a simple idea of sense is fairly obvious; it is to be aware of a particular quality of an object mediated by a single sense. The criterion of simplicity was, however, a problem for all British empiricists. In having simple ideas the mind is passive, but some activity is allowable in the forming of complex ideas.

Primary and secondary qualities. Among simple ideas of sense Locke makes an important distinction—already implicit in Descartes and others—between ideas of primary and ideas of secondary qualities. Primary qualities, such as bulk, number, figure, and motion, are, Locke thinks, inseparable from the bodies in which they are found. Bodies could not exist without them. Secondary qualities, such as color, sound, and taste, are, on the other hand, "nothing in the objects themselves but powers to produce the various sensations in us by their primary qualities." In other words, the primary qualities of things produce sensations in our minds that are ideas of secondary qualities, but "secondary quality" is a misnomer to the extent that there is really no such quality in things. Thus, our ideas of primary qualities actually correspond to the things that produce them, whereas our ideas of secondary qualities, although produced by things, resemble nothing in those things, being purely subjective. Locke brings forward a number of arguments for this conclusion, arguments based mainly on the assimilation of our perception of secondary qualities to sensations of pain. That is, he takes the perception of, for example, warmth or color to be the same kind of thing as feeling pain.

Account of perception. He thinks of perception in general as identical with merely having sensations, and he thus fully embraces the causal theory of perception according to which perceiving is having sensations caused by things. He goes further than this in respect to primary qualities, for here he also accepts the representative theory of perception according to which our ideas represent the things that cause them. This theory, as we have seen, was the stock in trade of seventeenth-century philosophy. Like most theorists of this pattern, Locke can give no good reason for the view that any of our ideas resemble their causes, and he cannot take the rationalist course of appealing to an intellectual intuition of some properties of things. It is a fair guess that he, like Descartes and others, was influenced by the success of physical science in maintaining that physical properties like extension—properties which are measurable—are *the* properties of things. There is also the connected fact that these properties are perceptible by more than one sense, as Aristotle had noted in his theory of common sensibles.

Modes, substances, and relations. Complex ideas may be exhaustively subdivided into ideas of modes, substances, and relations. Ideas of modes are ideas of things that depend on substances for their existence—for example, the idea of a triangle or a murder. Ideas of substances are ideas of particular things taken as existing by themselves—the complex idea of substance, he goes on to say, consists mostly of powers. Ideas of relation, finally, result from a comparison of one idea with another. Locke came to have some dissatisfaction with this classification of complex ideas, and in the fourth edition of the *Essay* he introduces a fourfold classification of ideas—simple ideas, complex ideas, ideas of relation, and general ideas. However this may be—and there is room for dissatisfaction with Locke's second classification, too—all ideas other than simple ones are in some way formed by the mind out of simple ideas.

Locke classified ideas of space, time, and number as ideas of modes. That is to say that they are ideas of entities which depend for their existence on particular things. We build up our ideas of these entities out of our ideas of particular things when seen in the appropriate relations. Kant later showed that such a view of the source of our ideas of space and time was untenable; Leibniz commented on the fact that Locke failed to take account of the a priori nature of the ideas of space and time and attributed the failure to Locke's inexperience with mathematics.

Locke maintained that our ideas of physical substances are mostly ideas of powers and that the idea of power is an idea of another mode. Since what we know of physical substances is largely due to their effects on us or on other substances, ideas of physical substances all mainly ideas of power. The effects, Locke thinks, are due to the motions of the invisible parts of things, but owing to the weakness of our senses, we are unable to perceive the nature of these causes. We have little or no knowledge of the "real essences" of things. What we do know of things is of their "nominal essences"—their nature

merely in respect of the classifications into which we fit them. Such classifications may not correspond to the real nature of things. Here Locke clearly shows how much he was influenced by the physical sciences. He thought that classifications show the way to the nature of things, but that owing to the weakness of our senses, we are unable to do more than gain a general impression of the nature of those physical processes. Therefore, we have to be content with an ordering of things according to their effects rather than as they are themselves.

Theory of general words. Locke adds to the account of ideas a discussion of language and of the words corresponding to the various ideas. It is in the context of this discussion that he puts forward his theory of the meaning of general words, a theory that was to come under attack from Berkeley. This theory—that the meaning of general words is given by the general or abstract ideas to which they correspond—is in effect Locke's theory of universals. He expresses the problem by asking, "Since all things that exist are only particulars, how come we by general terms; or where find we those general natures they are supposed to stand for?" His answer is that words are general by being signs of general ideas, and we form general ideas by abstraction, "separating from them the circumstances of time and place, and any other ideas that may determine them to this or that particular existence." Thus, words become capable of representing a number of individuals by standing for an abstract idea. Locke's view is therefore a form of conceptualism in that the universal or general element lies in our ideas or concepts, not in anything nonmental. Given a liberal enough interpretation of the word *idea*, there is perhaps no great difficulty in understanding what Locke is getting at, although the implication that the meaning of words must always consist in their standing for something (in this case an abstract idea) is certainly wrong. The idea terminology is a vague one, common though it was in the seventeenth and eighteenth centuries, but most of those who employed it would have denied that ideas were simple images of things. Moreover, such an interpretation is far from consistent with much that Locke says about ideas. Nevertheless, the use of the term "abstract idea" is not without its difficulties, especially since Locke says that such ideas must represent things.

Kinds of knowledge. In his account of ideas and their classification Locke is the strict empiricist, maintaining that all ideas must be ultimately derived from simple ideas of sense, either directly or as a result of the operations of the mind upon these. His account of knowledge, given in the last book of the *Essay*, is less empiricist in character; indeed, it owes an obvious debt to Cartesianism. He begins with the claim that knowledge is nothing but "the perception of the connexion of and agreement or disagreement and repugnancy of any of our ideas." This agreement or disagreement can be classified into four kinds: (1) identity or diversity, (2) relation, (3) coexistence or necessary connection, and (4) real existence. It is the fourth kind which presents the difficulties. How can the knowledge of the existence of a thing be a matter of the perception of the agreement or disagreement between our ideas alone? This could be so only if our knowledge of the existence of things could be a priori. Locke thinks that some knowledge of this sort can be shown to be a priori, but it is knowledge of the existence of sensible things that presents the greatest difficulties.

Locke distinguishes between three degrees of knowledge in a manner which is reminiscent of Spinoza's distinction between the three kinds of knowledge. There is, first, intuitive knowledge; second, demonstrative knowledge; and, third, "sensitive" knowledge of particular finite existences. The last Locke adds almost as an afterthought on the ground that it has by no means the certainty that belongs to the first two, although it goes beyond mere probability and is commonly thought of as knowledge. (Locke's conception of the standard to which knowledge must attain is noteworthy here.) Apart from the different degrees of certainty that are to be attached to these kinds of knowledge, they also differ in that intuitive and demonstrative knowledge can be concerned with relations between ideas (we can see that white is not black, and we can reason from one idea to another) but sensitive knowledge is concerned only with the existence of the objects of ideas. This is not to say that there cannot be intuitive and demonstrative knowledge of existence, too. Locke thinks that we have intuitive knowledge of our own existence (compare Descartes's *cogito*) and demonstrative knowledge of God's existence (by means of a version of the Cosmological Argument; Locke distrusts the Ontological Argument). But how can the existence of anything be known from ideas alone? Locke sometimes appears to say that such knowledge consists in the perception of the agreement of certain of our ideas with the idea of existence, but in general he acknowledges that more is required than this—real existence and not merely conceived existence. The difficulties here are especially evident in connection with sensible knowledge.

Existence of external world. In Book IV of the *Essay* Locke tries to justify the claim that we have knowledge of the existence of particular sensible things by showing that our ideas do correspond to the things that cause them.

Whereas complex ideas may not always correspond to things because of the part played by the mind in forming them, simple ideas receive no contribution from the mind; they are entirely passive. Unfortunately, it does not follow from this that they are necessarily veridical. He adduces further considerations, stressing the passivity of the mind in receiving ideas and the way in which the senses may cohere in their reports. None of these considerations is really sufficient, and Locke admits that they do not amount to proof. In fact, by simultaneously embracing a general empiricist approach and a representative theory of perception, Locke cannot provide a guarantee of, or even any general argument for, the veridicality of the senses. He cannot provide any independent access to the objects of perception other than that provided by the senses themselves. Like most empiricists, Locke accepts the correspondence theory of truth in that the truth of a proposition consists in its correspondence to the facts. (Truth, he says, signifies "the joining or separating of signs, as the things signified by them do agree or disagree one with another.") But he has no general warrant for the belief in the correspondence of ideas to things.

BERKELEY. The main aim of Bishop Berkeley (1685–1753), as he conceived it, was to defend common sense and religion against skepticism and atheism. But both his metaphysics and his theory of knowledge, connected as they are, can be partially regarded as attempts to rid Locke's theory of impurities. Locke's view of the world involved, besides minds and their contents (ideas), material substances, for the most part unknowable. Berkeley wished to get rid of material substances precisely because they were unknowable. In his view the existence of material objects consists only in their being perceived; their *esse* is *percipi* (their existence is to be perceived). In consequence, they must be regarded as complexes of ideas whose cause cannot be any substance underlying them but must be a spirit, the only active thing. Some of our ideas may be caused by ourselves qua spirits, but insofar as ideas have what we normally think of as an objective order, they must be caused by the supreme spirit, God. The laws of nature according to which ideas are ordered are guaranteed by God, and our ordinary way of looking at and talking about the world obscures this fact. Berkeley therefore thought that it was necessary to rid Locke's views of those elements which prevented this insight. The main issues concerned the notion of substance as the cause of our ideas, with the connected doctrine of the distinction between primary and secondary qualities. Berkeley also found fault with Locke's doctrine of abstract ideas, partly for its own sake but also because he thought it one of the main supports for the doctrine of substance. We might, that is, have an idea of substance by abstraction; it had to be shown that this was impossible.

Knowledge of the external world. At first, in the *New Theory of Vision*, a work on both optics and philosophy, Berkeley maintained that physical objects are primarily objects of touch. Vision, he asserted, could provide us with no direct perception of the distance of things, for the retina of the eye is only a two-dimensional surface. Our sensations of sight (and, like others of his time, Berkeley thought that perception fundamentally consists in having sensations) can only be of expanses of color. When we perceive things as at a distance from us, what really happens is that the visual sensations which we have suggest to us certain ideas derived originally from touch and connected with the visual sensations by experience. The *New Theory of Vision* consists largely in the working out of this theory in detail. Berkeley came to see, however, that there was no reason for making this distinction between sight and touch. All senses should be alike in these respects. Insofar as we have knowledge of what we take to be physical things, it is because we have, as the result of experience, so connected ideas that having certain sensations or ideas suggests other sensations or ideas. These ideas make up a collection that we identify as an object. Thus, for Berkeley objects are, in some sense, identical with collections of ideas.

In the beginning of the main part of his *Principles of Human Knowledge*, Berkeley flirts with the view, now known as phenomenalism, that all we mean when we say, for example, that there is a table in our study when we are not there is that if we were there, we should perceive it. But he adds as an alternative the suggestion that what we mean is that "some other spirit actually does perceive it," and this is his main view. What we ordinarily take to be things are really bundles of ideas in some spirit's mind; they have a certain stability even when we are not perceiving them because God still is. Indeed, it is God's having ideas according to a certain order that guarantees the order of our ideas.

Sensations and ideas generally (for although sensations are, strictly speaking, one species of idea, Berkeley often uses the terms interchangeably) are entirely passive. Their *esse* is *percipi*. On the other hand, spirits—God or ourselves—can be active. We cannot have ideas of spirits, although we have a notion of them, since we can understand the word *spirit* and at least know that we are the source of some of our ideas. There is no room in this for any substance's underlying our ideas, since we have no idea of such a thing. Furthermore, the special place that

Locke had given to primary qualities—that of being the properties essential to material substance—is untenable, and if what Locke said about secondary qualities is right, there are no grounds for making any distinction between them and primary qualities. They are equally dependent on the mind, so that if secondary qualities are subjective (and Berkeley accepts and adds to Locke's arguments for this conclusion), they must all be. All qualities are ideas in the mind.

Theory of universals. Berkeley's fiercest attack upon Locke was directed against his doctrine of abstract ideas. Berkeley interpreted Locke as asserting the existence of ideas or images that possess contradictory properties. The abstract idea of a triangle must be simultaneously scalene, isosceles, and equilateral. This is clearly impossible. It is doubtful whether Berkeley is right in this interpretation, but he clearly thought that if such ideas were admitted, there could be little objection to the admission of the idea of substance, too—that is, the idea of a physical but in principle imperceptible object.

In the place of abstract ideas, Berkeley introduced a theory of universals that was nominalist in character. Universals are merely particular ideas that are representative of other ideas in the same class in the way in which a particular man may be representative of other men; hence, their universality lies only in their power of representation. There is no need to assume the existence of general ideas since general words need not correspond to general ideas in order to have meaning. In other words, Berkeley challenged the theory of meaning that asserts that all words are names and refer to something—*unum nomen, unum nominatum* (one name, one thing named), as the Scholastics put it. In his view general words stand for a number of particular ideas belonging to the same class. General words are different from names in that general words represent a number of things indifferently. It must be confessed, however, that it is difficult to see clearly what, according to Berkeley's view, is involved in understanding a general word. Certainly, it involves having an idea which indifferently represents a whole class of things, but what is it to see that it does so?

Refutation of skepticism. Berkeley's general view has certain consequences. It means, for example, that important sections of mathematics have to be abandoned. There must, Berkeley believed, be a least perceptible size. Since all our ideas are ultimately derived from sensations, there can be no ideas of infinitesimals or points. For the most part, however, Berkeley considers himself to be defending common sense against the attacks of the metaphysicians. The vulgar, he maintains, believe that "those things they immediately perceive are the real things" (that is, not imperceptible substance), but philosophers believe that "the things immediately perceived are ideas which exist only in the mind." Berkeley characterizes his own view as the joining of these two notions in that he equated real things and ideas. He thinks that given his view that ideas, which are the objects of immediate perception, are the real things, there is no room for doubt concerning the real nature of things—a doubt which Locke had expressed. Moreover, since what is immediately perceived is by definition free from error, only the wrong use of ideas in judgment can give rise to error. Error is thus a product of the imagination. Insofar as we rely upon our sense perceptions as directly given, we must be free from error. Thus, Berkeley claims, his view prevents skepticism and "gives certainty to knowledge."

Concept of knowledge. Apart from the reference to God and spirits, Berkeley is a strict empiricist not only in the sense that he believes that all the materials for knowledge are derived from sense perception (as Locke, too, believed) but also in the sense that knowledge is itself founded on sense perception. Locke was not such a complete empiricist; he thought that knowledge in the strict sense is founded on intuition and demonstration, and he made skepticism possible to a certain extent over sense perception because he thought that its veridicality could not be completely shown. According to Berkeley, knowledge derived from reasoning must ultimately be founded on knowledge based on sense perception. Sense perception, in turn, is no longer conceived of as having ideas that are produced by objects and may not always represent their causes.

Berkeley has given up the representative theory of perception with its assumption that something so underlies our ideas that they may be representative of it. His rejection of the representative theory of perception is the basis of his claim to combat skepticism. Yet, as Hume asserted, it has often seemed a claim that fails to produce conviction, because the claim that what is directly perceived is free from error is true by definition. The question of how we know when we have direct perception still remains, however. Not all ideas are objects of immediate or direct perception; some are ideas of the imagination. According to Berkeley, these are less regular, vivid, and constant than ideas of perception, and they are more "dependent on the spirit"; they can be distinguished from ideas of sense by these criteria. But are all ideas of perceptible things ideas of things immediately perceived, and if not, how do we tell which are?

In the first of the three *Dialogues between Hylas and Philonous*, Berkeley argues that by sight we immediately perceive only light, colors, and figures; by hearing, only sounds; by taste, only tastes; by smell, only odors; and by touch, only tangible qualities. Here he appears to be arguing from the premise that these things are the special or proper objects of the senses. Although it is difficult to know what, if anything, is special to sight and touch, it is easy to see what is meant in the case of the other senses. Even if we grant that we hear only sounds, taste only tastes, and smell only odors, it does not follow, however, that we cannot be mistaken about the characteristics of these objects in particular instances. Are we necessarily free from error in hearing when we hear that a sound is loud or soft?

Nor is our attribution of colors necessarily free from inference as it should be if the perception of color is immediate. What, then, really counts as an object of immediate perception? In answering this question, Berkeley is subject to the same difficulties that have beset more modern philosophers when they have sought to base the philosophy of perception on the notion of sense data. If the foundations of knowledge are found in the deliverances of the senses, there must be certain perceptions that are incorrigible in the sense that they cannot logically be subject to doubt. But what counts as incorrigible perception? Berkeley tries to answer this question by assimilating perception to having bare sensations. Sensations, however, are not the sort of thing that can be right or wrong. The mere passivity of sensation, as opposed to the will, does not show that error arises from the will. If this criticism is valid, Berkeley's theory does not satisfactorily prevent skepticism in the way that he supposes.

HUME. Locke thought that his inquiry had revealed the limitations of the understanding by showing that there are parts of nature that our senses cannot discern. Berkeley, on the other hand, thought that there was nothing which our understanding could not grasp. Sense perception gives us complete knowledge of reality, and we have in addition notions of spirits, including God. Indeed, we could regard our ideas as a sort of divine language by means of which God speaks to us, so that our senses, if viewed correctly, continually reveal the glories of God. Hume (1711–1776) agreed with Berkeley in thinking that there is nothing in nature that lies beyond the reach of our senses, but, contrary to Berkeley, Hume reached the conclusion that our understanding is very limited and that skepticism is the only reasonable attitude toward knowledge. That Hume was intentionally a skeptic has been disputed, but there is no doubt that this is the logical outcome of his views. He thought that whatever Berkeley's claims, his arguments were in fact skeptical: "They admit of no answer and produce no conviction." In effect, therefore, Hume's position is that of following the principles of empiricism to their conclusion without any ancillary claim to knowledge of the inner workings of nature or of God. His conclusions are also something of a *reductio ad absurdum* of empiricism.

Nature of ideas. Hume begins by drawing a sharp distinction between impressions and ideas, impressions being the perceptions of sense and ideas the perceptions of the imagination or memory. In this he claims to be restoring the term *idea* to its original use. Every simple idea must have a corresponding impression—the idea of red, for example, resembling the impression of red—and complex ideas may be formed out of simple ideas. As with Locke, both impressions and ideas may be divided into those of sense and those of reflection, impressions and ideas of reflection being impressions and ideas of the mind's reflection on impressions or ideas of sense.

The criteria of the simplicity of an impression or idea are as much a problem here as with Locke. To have a simple impression is to have an elementary perception that cannot be further broken down into other perceptions, and this will function as a building block out of which the rest of knowledge may be constructed. Hume takes very strictly the principle that to every simple idea or perception of imagination or memory there must correspond an impression, or perception of sense, although at the very outset he admits a possible exception in the idea of a color in a series. We may have the idea of such a color from the principle of a series without ever having seen it. This possible exception, however, Hume refuses to take as important.

The principle that every simple idea must correspond to an impression is vital for a delimitation of the understanding and as a weapon against rationalism. Impressions and ideas can, however, be distinguished only by the superior force and vivacity of impressions; they cannot be distinguished in terms of their relations to physical objects or minds, for our knowledge of physical objects is derived solely from impressions, if at all. Likewise, among ideas, ideas of memory have a superior liveliness to ideas of the imagination. It is extremely doubtful whether this is always true, and this, in turn, casts doubt on any attempt to characterize remembering and imagining on empiricist lines by reference to the contents of the mind alone.

Theory of universals. Hume follows Berkeley in his theory of abstract ideas or universals. In his view there are

no abstract ideas, strictly speaking; however, ideas can be particular in their nature and general in their representation. Hume's only addition to Berkeley's account is his attempt to indicate how this can happen through the association of ideas. The occurrence of one idea may dispose the mind to call up all other ideas associated with it. Hence, the understanding of a general word lies in the disposition of the mind to have the ideas of those things to which it may be applied. This is not a very plausible account in itself since the notion of understanding cannot be analyzed in terms of habits or dispositions of minds, but it is at least an attempt to tackle the problem. The solution is in accord with Hume's general approach; his account of belief is similar.

Space and time are difficult notions for an empiricist to deal with, for, as Kant pointed out, particular phenomena seem to presuppose space and time rather than vice versa. Hence, it is difficult to see how our ideas of space and time can be derived from our ideas of particular phenomena. Locke had nevertheless classified our ideas of space and time as ideas of modes. Hume attempts to deal with our perception of spatial extension and temporal duration in terms of the order in which impressions or ideas appear. But in consequence he has to admit that ultimately the impressions that are ordered in this way cannot themselves be extended or of extended objects, nor can they take time. In general, Hume's treatment of space and time is one of the more puzzling parts of his work.

Causality. Hume's greatest reputation derives from his treatment of causality, although his approach to this subject is similar to his approach to the problem of our knowledge of the external world or of ourselves. His approach is founded upon a distinction between different kinds of relation. There are "relations of ideas," which depend completely on the ideas related, and factual relations, which can be changed without changing the ideas. This is a distinction between logical and matter-of-fact relations, and it leads to a distinction between logical truths and factual truths that parallels Leibniz's. Hume is interested in the causal relation because he believes that it is the only matter-of-fact relation that can lead us from one idea to another. Causality is not a logical or a priori connection, but it is a connection. This assertion is of the utmost importance. Why, however, do we think that there is some necessity in causal connections? It cannot be a logical necessity; also, it cannot be derived from a more general necessity such as might be provided by a principle of universal causality, for Hume believes that such a principle must be contingent and that the evidence for it must be derived from our knowledge of particular causal connections.

He therefore proposes to "beat about the neighbouring fields." He notes that we generally take a cause to be antecedent to its effect and contiguous to it in space. More important, in experience there is a constant conjunction between cause and effect. In a sense these factors provide the basis for our belief in the necessity of the causal connection. Hume takes belief to be a lively idea associated with a present impression, and here the principles of the association of ideas again play a part. What makes an idea a belief is the feeling of being determined by habit or custom to pass from the impression to the associated idea. This feeling is an impression of reflection. It is in such an impression that Hume finds the source of our idea of necessary connection between cause and effect, for the "experimental method"—the resort to experience—should show us that there is no impression of power as such. It is due to habit or custom that we pass from cause to effect, and our belief in the necessity of doing so arises from the impression of a reflection of being determined to do it.

It is important to note just what Hume has achieved here. He has not in any way justified our belief in the necessity of the causal connection; he has merely attempted to explain the origin of the belief by giving a psychological explanation, not a philosophical justification, of the belief. But he has rejected all theories of occult powers in things, so that in one sense he may be considered to have said that what we mean by calling one thing the cause of another is that it is a uniform and contiguous antecedent of another event. To this extent his account is a reductive analysis; he analyzes our notion of cause by reducing it to notions that we understand. Yet Hume can find no justification for inferring the occurrence of one event from that of another; certainly, one event does not logically imply the occurrence of the other, but what other justification is there? Hume's conception of justification fails just because he recognizes no other kind of justification except that one thing logically implies another. Although Hume is commonly said to have raised the problem of induction, he made no real attempt to solve it himself, nor could he within his framework.

Knowledge of the external world. In Hume's account of knowledge of the external world the skepticism already implicit in his account of causality comes to the fore. Like Berkeley, Hume distinguishes between the beliefs of the "vulgar" and of the "philosophical system." The vulgar believe that we are aware of perceptions only, but they

also believe that some of them—our perceptions of primary and secondary qualities—have permanent existence. The philosophical system holds that there is a distinction between objects and perceptions and that only objects are permanent. Hume claims to side with the vulgar, but he sees no reason to distinguish any perception from any other. The mind is like a theater in which scenes come and go. Yet he does admit that it is natural to believe in a world of permanent objects. Reason can provide no justification of this belief, but we can give a psychological explanation of it like the account of our belief in the necessity of causality.

Our impressions have a certain coherence and constancy—that is, they fit together and recur in the same order after intervals. As a result, the imagination tends to carry on by custom or habit, and it attributes more regularity to objects of perception than they actually possess. Thus, we come to believe in a world of permanent objects, and we tend to reconcile what reason tells us of the interrupted nature of perceptions and what our imagination suggests about their regularity by a "philosophical" (as opposed to a commonsense) belief in a world of permanently existing objects. Nevertheless, a "direct and total opposition betwixt our reason and our senses" remains. Hume often speaks as if objects were just bundles of perceptions, but he has to deal with the belief that they are more than this. For such a belief he can give no justification, although he offers an explanation of its origin. In the last resort he can only recommend inattention to both our senses and our understanding. This is nothing if not skepticism.

Personal identity. Very much the same account is given of our knowledge of ourselves, a fact which may seem even more paradoxical. Once again, Hume uses the appeal to experience to indicate that we have no impression of the self. He rejects once and for all Berkeley's suggestion that we have a notion of the self. Belief in the self must therefore be parallel to belief in an external world, and Hume proceeds similarly. Belief in our identity through time must result once again from the coherence and constancy that exists between our impressions and ideas, as a result of which the imagination takes them to be impressions and ideas of a single self. Once again, however, no reason can be given for this belief, a fact that worried Hume more than his other tendencies toward skepticism. He returned to the topic in an appendix to the *Treatise of Human Nature*, but in the end he could find no way of ridding himself of his worry except a game of backgammon and a good dinner.

Account of perception. Hume rejected Lockean substance, with the result that perceptions—impressions and ideas—become the substantial entities in his ontology (as he in effect admits in *Treatise*, Book I, Part 4, Ch. 5). In retaining the terminology of perceptions, especially that of impressions, Hume clung to the skeleton of the causal or representative theory of perception. But the skeleton no longer had flesh, despite the suggestiveness of the terminology. Thus, Hume is forced to take his starting point from perceptions that are logically independent of any owner and any object. In one place he says that there is nothing objectionable in the idea of an unperceived perception—a very odd notion. From this he has to build a world that fits the common supposition that there are physical objects and persons. The premises from which Hume derives his position are unacceptable, but given them, he can provide no reason whatsoever for belief in such a world and has to say that the belief is just a product of the imagination. This is skepticism with a vengeance, but it is the logical outcome of his approach.

REID'S CRITICISM. Hume's contemporary Thomas Reid (1710–1796) thought, rightly enough, that Hume's conclusions were manifestly absurd. Finding nothing wrong with the arguments presented by Hume, he concluded that the fault must lie in the premises and proceeded to attack the whole "way of ideas" which was the source of these premises. Reid maintained that it was necessary to make a strict distinction between sensation and perception, a distinction that the doctrine of impressions and ideas blurred. Reid was quite right about this, and his account of the nature of sensation and perception is interesting for its own sake. Sensation, he said, is an act of the mind that has no object distinct from that act, its prototype being pain. Perception is a much more complicated affair, involving a conception of an object and a belief in its existence. In many cases we fail to note the distinction because most of our perceptions are accompanied by sensations. Sensations provide in themselves no basis for inference about the nature of perceived qualities of things, although things cause the sensations. Reid expresses the relation between perception and sensations by saying that the sensations suggest the perceptions; sensations are natural signs of perceived qualities. This suggestion and the sign relationship involved are not a matter of experience, for we do not experience sensations in the same way that we perceive things (to have a sensation is not to perceive anything in itself). The relationship is a natural one like, he says, that between expressions of emotions and the emotions themselves.

Whatever the worth of this account, it is certainly a completely different analysis of perception from that of other philosophers of the period. In claiming to defend our ordinary beliefs by this account against the skepticism introduced by Hume, Reid set himself up as a philosopher of common sense, a position adopted by G. E. Moore in the twentieth century. Reid did retain some of the features of British empiricism, however, especially in a distinction between original and derived perceptions, a distinction that is in some respects very similar to that between simple and complex ideas or impressions. Moore similarly employed much of the apparatus of sense data invoked by modern empiricists.

KANT

Kant (1724–1804) represents the juncture of seventeenth-century rationalism and British empiricism. Brought up in the tradition of post-Leibnizian rationalism, he was, as he put it, awakened from his dogmatic slumber as a result of a reading of Hume. His critical philosophy, as expressed in the *Critique of Pure Reason*, can be characterized as an attempt to draw the boundaries between the proper use of the understanding and the improper use of reason in making assertions of speculative metaphysics and as an attempt to show how the understanding can provide objectively valid knowledge of those things which Hume left to the imagination.

CLASSIFICATION OF JUDGMENTS. Kant bases his approach upon a twofold distinction between types of judgment. There is, first, a distinction between a priori and a posteriori judgments, the first being judgments whose truth can be known independently of experience, the second being judgments whose truth can be known only through experience. A priori judgments are pure when they involve only concepts that are themselves independent of experience. On the other hand, it is not a necessary condition of a judgment's being a priori that it should involve only such concepts. The concepts that are involved in a posteriori judgments, however, must be derived from experience, must be empirical.

The second distinction, that between analytic and synthetic judgments, is different. Analytic judgments are judgments about a thing that give no information about the thing, although they may serve to analyze or explain the concepts involved. This is because the concept of the predicate is contained, albeit covertly or obscurely, in the concept of the subject. The denial of these judgments involves a contradiction; hence, they correspond to what Leibniz called truths of reason. Synthetic judgments, on the other hand, do give information about a thing; in them the concept of the predicate is not included in that of the subject, and their denial does not involve a contradiction.

Kant now combines the two distinctions. Analytic a posteriori judgments are clearly impossible, but there is no difficulty, Kant thinks, about analytic a priori judgments or about synthetic a posteriori judgments. There remains the class of synthetic a priori judgments. It is on these, Kant thinks, that the claims for metaphysics rest, and it is these in particular that empiricists refuse to admit. Kant's program is to show whether and to what extent such judgments exist. The outcome of the program is that although metaphysics in the traditional sense is impossible, synthetic a priori judgments are admissible—first, in mathematics and, second, in the form of the presuppositions of objective experience or science. This program, which constitutes the critical philosophy, is Kant's substitute for traditional metaphysics.

SYNTHETIC A PRIORI KNOWLEDGE. The possibility of synthetic a priori knowledge means that not all knowledge about things can be derived from experience. Nevertheless, Kant thinks that all such knowledge is based on experience. It starts from what he calls intuitions, but since knowledge involves the possibility of making judgments about things, it cannot consist of intuitions alone; it must also involve concepts. To have a sensible intuition is to have a simple awareness of something by means of the senses. This awareness Kant analyzes in a way that derives much from the British empiricists. A sensible intuition consists, first, of a sensation as the content of the intuition. Its form consists of spatial or temporal extension. Hume had to admit that impressions have an order, but he drew the consequence that the impressions themselves were unextended and nontemporal when at their simplest. Kant generally argues that sensations have only intensive qualities, qualities that can vary only in degree. Nevertheless, since the intuition consists of the sensation plus the form—that is, its relations to other sensations—spatiotemporal form is something which is a necessary part of our experience. One cannot, as Locke seemed to suppose, build up ideas of extension from first impressions. Spatiotemporal form is a necessary, a priori characteristic of experience.

Since Kant, however, has assumed a theory of perception similar to the representative theory, this a priori spatiotemporal form applies only to things as they appear to us—to phenomena. It does not apply to whatever may be thought to lie behind our experiences (things-in-

themselves). This fact Kant expresses by saying that spatial and temporal characteristics (and primary qualities in general) are empirically real but transcendentally ideal. The characteristics in question are not merely subjective; they are objective—valid for all men—but only in relation to phenomena, not to things-in-themselves. (Throughout, Kant's criterion of objectivity is the criterion of intersubjectivity—validity for all men; he is pointing out that from the point of view of the critical philosophy something may be objective in this sense without being a feature of something independent of the mind.)

Pure a priori intuitions. Kant goes on to argue that we have not only a sensible intuition involving a priori spatiotemporal features but also a pure a priori intuition of space and time themselves. It is by virtue of this that the science of mathematics is possible. In order to do geometry, for example, it must be possible to make constructions in space, an idea that presupposes that we have an intuition of space. (Kant insists that this is an intuition, not a concept, but his reasons for this are complex and difficult to understand.) Arithmetic similarly presupposes an intuition of time. It is for this reason that mathematical judgments are both synthetic a priori and possible. Kant took it for granted that Euclidean geometry was *the* geometry of space, so that his account provided the justification of that geometry. It has often been suggested that the discovery of other geometries has undermined his case, and to some extent it has. But Kant would still have insisted that some intuition of space is a necessary condition of the possibility of any geometry, and there is something to be said for this position. Our concepts of space and time are not concepts that can be simply abstracted from experience.

CATEGORIES OF UNDERSTANDING. Space and time, then, provide the form of all experience, just as sensation provides the content. What is given in this way must be subsumed under concepts in judgment if knowledge is to result. "Thoughts without content," Kant says, "are empty, intuitions without concepts are blind." But in itself the formation of judgments is not enough for knowledge. The judgments that we make might be just the work of the imagination, as Hume in effect supposed in considering our knowledge of the external world. What criteria, then, have to be observed in the case of objectively valid judgments? Kant's answer is that such judgments have to conform to certain principles of the understanding and that these principles are derived from the pure or formal concepts, which Kant calls categories, of the understanding. Only insofar as our judgments conform to these principles can the judgments that we make about appearances be intersubjective, true for all men. Objectivity can be a question of this intersubjectivity alone because no valid judgments can be made about things-in-themselves. What, then, are these principles, and what are the categories?

In the section of the *Critique* known as the "Transcendental Analytic" Kant puts forward two arguments for categories. The first, the "metaphysical deduction," tries to argue for the existence of the categories directly, by finding the key to the list of categories in the traditional table of judgments provided by formal logic. To each of the traditional headings under which judgments can be classified logically, there corresponds, Kant believes, a concept that provides the principle of construction of an objectively valid judgment. The second argument the "transcendental deduction," attempts to show that the existence of categories of the understanding is a necessary condition of possible experience. The two arguments are complementary in that the transcendental deduction depends upon the metaphysical deduction for the actual list of categories, while the metaphysical deduction does not really show that categories are necessary to objective experience.

It is true that later, in discussing the principles derived from the categories, Kant brings forward specific arguments in each case, so that it might be said that the case for accepting each of these principles does not depend entirely on an acceptance of the metaphysical deduction. Yet, it is the metaphysical deduction alone that provides the guide to which categories and which principles we should seek. Today, the metaphysical deduction is almost universally rejected. There can be no validity in the attempt to derive a table of categories for objectively valid judgments from a table of judgments classified according to purely logical principles. It remains, therefore, that if the transcendental deduction is valid, it is possible to accept some categories as necessary, but apart from arguments for specific cases one cannot determine which categories are necessary by reference to any general rule.

Transcendental deduction. The argument of the transcendental deduction is very complex, and only the most general outline will be given here. First, the senses provide us with a manifold of sensations set out in space and time. Second, in order that they may form a unity, the understanding, with the aid of the imagination, has to synthesize them. The imagination helps us to see the manifold as a manifold in space, and in the form of memory it ensures that we also see it as a unity over a tempo-

ral period. Kant calls these two forms of synthesis the synthesis of apprehension and the synthesis of reproduction. Third, in the synthesis of recognition the manifold has to be given a principle of unity by subsumption under a concept, so that we see the manifold as a such-and-such. The results of this, however, could still be only subjective. In order to attain objective validity, the understanding must enable us to conceive of the manifold as united in an object. What Kant calls the transcendental unity of apperception is the awareness of experiences as part of one consciousness and as having an object, although neither the owner nor the object of those experiences can be found in the experiences as such. Objective experience presupposes these features; otherwise, the situation would be, as Hume in effect supposed, a mass of experiences whose connection with a person or objective world is merely contingent.

Fourth, the judgments we make about the manifold of experience thus unified must themselves conform to certain principles of unity. It must be possible, for example, to see certain connections within experience as that of ground to consequent, and our judgments must presuppose such connections in the things joined in them. Thus, we arrive at the idea that if objective experience is to be possible, it must conform to such categories as ground and consequent. The categories are concepts of the principles of connection of things in judgment, if that connection is to be more than a mere subjective one. They are categories because they are applicable to anything.

Transcendental judgment. The categories derived from the logical table of judgments according to the metaphysical deduction are purely formal. For example, Kant believes that the category of ground and consequent is derivable from the logical notion of a hypothetical judgment. This purely formal category of ground and consequent can be given content only by being applied to phenomena in such a way that it emerges in more material form in terms of the particular relation of ground and consequent that is applicable in the case of phenomena—that is, in terms of the relation of cause to effect. Kant formulated a doctrine of schematism to explain how we can apply the purely formal categories to experience. A schema is a kind of principle for the construction in the imagination of anything that falls under a given concept. It is that which enables us to identify a given object as an instance conforming to the concept. Thus, the schema for each of the categories can be thought of as the principle for the application of the pure category to phenomena in time. The notion of ground and consequent applied to phenomena in time emerges, as we have seen, as the notion of cause and effect (an essentially temporal notion). It is only from these schematized categories that it is possible to derive the principles of the understanding according to which all objectively valid judgments must be viewed. These principles Kant discusses and argues for separately.

In all this Kant believes he has explained how judgments about mere phenomena can be objectively valid although they are confined to these phenomena. The judgments in question are by no means applicable to things-in-themselves, to whatever lies behind phenomena. Such notions as those of a world lying behind phenomena and of a real self that is aware of them are noumena, and as such they must be thought of as limiting concepts. To treat such concepts as if they were concepts of an ordinary kind and to use them in a systematization of knowledge, as is done in speculative metaphysics, is wrong and is liable to produce fallacies. A noumenon is merely a "nonphenomenon," and the concept of a noumenon is essentially a negative one, but reason tends to treat such concepts as positive, and from this the illusions of speculative metaphysics stem.

CRITIQUE OF METAPHYSICS. The judgments that the understanding allows us to make are conditional in that they are relative only to possible experience. Pure reason tends to assume an absolute, something unconditional, which provides the basis of the unity of all judgments of the understanding. It thus provides us with ideas whose proper use is only regulative, in that they are ideas of the goals or limits toward which the understanding may strive without being able to apply them directly to experience. To use these ideas, as speculative metaphysics does, as ideas under which we can directly subsume reality—such ideas as those of the absolute unity of the thinking subject, the absolute unity of the world in space and time, and the absolute unity of the conditions under which anything can be thought at all, the entity of entities or God—is the source of antinomies and other forms of contradiction or fallacy. Kant sets out to expound these contradictions and their resolutions in detail, but it is impossible to follow the argument here. The section of the *Critique* known as the "Transcendental Dialectic" is a critique of rational psychology, speculative cosmology, and metaphysical theology with its attempts at a demonstration of the existence of God. Although there is much other material in the *Critique*, this section essentially completes the critical philosophy in its attempt to show that the understanding can provide objectively valid

knowledge of phenomena and to reveal the limitations of the proper use of the ideas of reason.

An assessment of Kant's work in the theory of knowledge is difficult to give. It contains many extraordinary insights, although their detailed development often leaves much to be desired. Above all, perhaps, it takes as its starting point the analysis of experience provided by the British empiricists, and this undoubtedly limits it.

POST-KANTIAN IDEALIST PHILOSOPHY

FICHTE. German philosophy after Kant is in many ways a commentary upon Kant's philosophy, either as further development or opposition. The idealism which was so characteristic of nineteenth-century philosophy was begun when Johann Gottlieb Fichte (1762–1814) found fault with the Kantian view of things-in-themselves that are beyond the reach of knowledge and proceeded to reject the notion on grounds similar to those which are commonly used against any causal or representative theory of perception—there can be no good reason for believing in such things. With the rejection of things-in-themselves, even as a limiting concept, we are left merely with experiences or phenomena, and it is of these that, in the idealist view, reality must consist. The general problem of idealism that Fichte thus introduced was how it was possible to distinguish among experiences those which are purely subjective and those which are really objective. The problem is how we can distinguish between what is contributed by the mind and what is not, between the self and not-self, as Fichte put it. In Kant's view objectivity was equivalent to validity for all people, but that it was at all possible to distinguish between what was due to the mind and what was not seemed guaranteed only by the existence of things-in-themselves. With the rejection of the latter, experiences and experiencer became only two sides of the same coin. For this reason the general trend of idealism was toward the coherence theory of truth—the view that experiences and judgments are true to the extent that they cohere with one another, forming a coherent system. This view was naturally associated with the doctrine of degrees of truth—that judgments have varying degrees of truth to the extent that they cohere with each other. This more or less intelligible view was, however, complicated by being involved with the view that judgments about the empirical world have a very low degree of truth because they bring with them paradox and contradiction. The sensible world is therefore only appearance, and reality must be something else.

HEGEL. The belief that the sensible world is only appearance is perhaps less marked in G. W. F. Hegel (1770–1831) than in some of his idealist successors—for example, F. H. Bradley. Hegel was influenced not only by Kant but also by Greek thought, especially by Platonist and Neoplatonist conceptions of an intelligible world of Forms with a structure of its own. Nevertheless, Hegel's relation to Kant may be roughly characterized by saying that he attempted to restore the functions of reason that Kant had forbidden. Whereas Kant had tried to justify the processes of the understanding while underlining the contradictions involved in an improper use of reason, Hegel tries to show that the understanding involves its own paradoxes, which can be resolved only by the use of reason; this, in Hegel's view, is by no means improper.

Dialectical method. Contradictions arise during the application of philosophical categories like those of the One and the many, so that the philosopher finds himself asserting both a thesis and its antithesis, in a manner similar to that expounded in Kant's antinomies. There is, Hegel thinks, a method which reason can pursue in order to resolve any such contradiction. Reason has to find a synthesis, some category that will reconcile those which produce the apparent contradiction. But the resolution may, in turn, find itself opposed to a further antithesis which demands another synthesis and so on. This method Hegel calls dialectic. According to him, it provides the key to understanding how the ideas of reason may be charted. In the end they will be seen to be dependent on the ultimate, absolute idea that provides the ground for everything else. Thus, the idea of something absolute and unconditional which Kant had rejected is restored.

It must be confessed, however, that as a method in the strict sense, Hegelian dialectic is sadly defective in that there appear to be no rules for its use. Hegel presents a series of insights, sometimes real, sometimes imaginary, into the relationships between very general and abstract philosophical ideas, like those of being and essence or consciousness, self-consciousness, and reason. Dialectic provides the architectonic according to which these relationships may be charted, and Hegel is excessively thoroughgoing in its use. The result—the Hegelian system—is a complete map of all forms of knowledge and of all philosophical ideas, constructed on a single plan. The attempt is ambitious; the ground for its validity, slender in the extreme. It would be foolish, however, to deny the incidental insights.

Theory of knowledge. Hegel's theory of knowledge may be found partly in his *Science of Logic* and partly in

his *Phenomenology of Mind*. In the *Science of Logic* he explains his view of Kant, criticizing Kant's trust of the understanding and Kant's, to Hegel's own mind, undue restriction of the functions of reason. Then, through the dialectic he charts the notions most central to reason, beginning with the opposition between the categories of Being and Nothing, the synthesis of which he finds in Becoming. These are notions which reason finds indispensable for any account of the world and upon which logic must depend.

In the *Phenomenology* Hegel sets forth his view of perception most clearly. There Hegel begins by pointing out that consciousness appears to be an apprehension of what is immediate, of what *is*, which is, it appears, a confrontation of the ego with something else (as Fichte also supposed). But sense knowledge proper must involve a subsumption of this immediate consciousness under universals or concepts, and, moreover, there is no way of grasping the particular that is thus subsumed under concepts except by reference to other concepts. Proper names and even words like "this" are, in Hegel's view, general words, since they apply to a multitude of different things (Hegel here ignores or fails to appreciate the way in which they so apply); hence, they furnish us with no means of identifying a particular independently of universals. Sense knowledge thus turns out to be a mediated knowledge, a knowledge which is possible only through the medium of universals and which is not a direct knowledge of reality.

There is, however, Hegel argues, a contradiction between the fact that we take ourselves to perceive things which are unitary entities and the fact that our knowledge of them can exist only through a plurality of universals which are themselves unconnected. This contradiction is resolved only because the intellect provides us with a higher universal that constitutes the basis or condition for applying the lower-order universals in sense perception. This higher universal is force, the idea of "lawlikeness." The unity of the objects of perception is due to the lawlike connections that exist between the universals under which they are subsumed. This is something that can be discerned only by the intellect, which thus produces the synthesis of the contradictions apparent within consciousness. This, of course, does not end the matter for Hegel, as the phenomena of consciousness are equally phenomena of self-consciousness. The opposition between consciousness and self-consciousness requires a synthesis by reason.

BRADLEY. The kind of general argument that Hegel used can also be found in the English idealist Bradley (1846–1924), although he was far less attached to Hegel's method, referring to the dialectic as "the bloodless ballet of the categories." (In spite of being chronologically out of order, emphasis is placed on Bradley here because Bradley, although often difficult to understand, is generally easier than Hegel for English-speaking readers. The slight differences between Hegel and Bradley matter little in light of their essential similarity of purpose.) In his *Principles of Logic* Bradley argues that all judgments are only conditionally true in that the identification of the portion of reality which is their subject involves subsumption under universals. The judgment as a whole therefore says that some universal can be ascribed to reality only on condition that some other universal or universals may be ascribed to reality. All judgments, although categorical in that they are concerned with reality (reality being the subject of all judgments), are also hypothetical in this sense. Hypotheticals must likewise rest upon a ground; their truth is dependent upon connections within reality (Hegel's force).

In Bradley, however, there is greater emphasis on the idealism. For Bradley universals correspond to ideas, so that in judgment we are attaching an idea to reality. However, since what is known can be connections between ideas only, reality as we know it is a system of ideas joined by what he calls internal relations. A judgment is true to the extent that the ideas which it ascribes to reality cohere with the whole system of ideas. By an internal relation, as opposed to an external relation, Bradley means a relation that is more than a contingent one. The relations that form the system that constitutes what we know of reality are more than mere contingent relations, although they are not so close as to be logical entailments. All that we can know of reality apart from the ideas under which we subsume it is that it is experience; it is the bare fact of consciousness from which Hegel starts. Bradley is indeed Hegel made more palatable for English tastes.

It is only fair to add that in his most explicitly metaphysical work, *Appearance and Reality*, Bradley finds paradoxes in the notion of relations in general. The idea that two things may stand in a relation gives rise, he claims, to an infinite regress. What is relational must be set down as appearance only. It follows that judgment can never amount to absolute truth, for all judgment involves the setting of ideas in relation; all judgment, in other words, involves inference. No judgment can therefore furnish more than a limited degree of truth or be about

something which has more than a limited degree of reality.

Judgments about the Absolute, the sum total of reality, may have a certain intellectual incorrigibility, since they will in effect ascribe to the absolute reality what is merely part of itself, but they cannot add up to truth itself. This is inevitable, since Bradley takes all judgment to be asserting the identity of subject and predicate, of reality and idea, while maintaining an unbridgeable gap between them—between the "that" and the "what," as he puts it. The notion of judgment therefore involves a contradiction in itself, and for this reason the understanding is condemned. Intuition or immediate awareness gives us the bare fact of experience as constituting reality. What it is unconditionally and absolutely only reason can tell us. Whereas Kant had maintained that reason can tell us nothing of what is absolute and unconditional and that only paradox can result from the attempt to make it do so, the Hegelian doctrine espoused by Bradley is that the limitations of the understanding can be seen only by going beyond its limits to what is not finite and not conditioned but absolute. The claims for reason had never been pushed so far before, nor have they been since.

SCHOPENHAUER. With Arthur Schopenhauer (1788–1860) there was a partial return to Kant. Schopenhauer thought that Hegel's dialectical method was barren because it ignored Kant's insights into the nature of reason and the understanding. Yet he retained Hegel's idealist approach. Kant's phenomena became presentations or ideas in a sense similar to Berkeley's; that is, they became subjective experiences. Schopenhauer thought that the world consists of ideas or presentations and that necessary connections judged to exist between them are, as Kant thought, merely conditional; however, Schopenhauer does not accept the paraphernalia of Kant's categories for justifying such judgments. In his view all justifications for claiming objective experience rest on the principle of sufficient reason, which takes various forms according to the form of knowledge involved. It acts as a logical ground (the *ratio cognoscendi*) in connection with logical truths, as a ground connected with the features of space and time (the *ratio essendi*) in connection with mathematical truths, and as causality (the *ratio fiendi*) in connection with ordinary empirical phenomena. Thus, the notion of causality is made to play the role of all Kant's categories in relation to empirical phenomena; causality is their only ground for necessity in this sphere. Schopenhauer finally looks to the will as the only ground for action, for moral necessity. As one class of phenomenon, action finds its explanation only in the will.

Will as the thing-in-itself. The world as idea or presentation is only one half of Schopenhauer's philosophy (his main work is titled *The World as Will and Idea*). Although he accepted the idealist framework, he thought that the demand for a thing-in-itself as the basis of all our ideas was inescapable. He finds the nature of the thing-in-itself in the will. Reality consists of the manifestations of one force, the will, which uses consciousness as an instrument for its own self-promotion. Only in art is there anything like freedom from it, for only there does the mind achieve a state akin to the contemplation of Platonic Ideas, a sort of permanency which is foreign to the general manifestations of the will.

Although the last part of Schopenhauer's thought had some influence upon the romantic movement in nineteenth-century German thought, it has not received much welcome from the mainstream of philosophers. However, Schopenhauer's theory of knowledge, contained in the part of his philosophy devoted to the world as idea, contains much of interest along lines which are, in origin, Kantian.

LATE NINETEENTH-CENTURY PHILOSOPHY

Philosophical thought in Germany during the nineteenth century tended to be either romantic or neo-Kantian. Neo-Kantian philosophy came under empiricist influences from Britain, and at the end of the century under Franz Brentano and Alexius Meinong this finally led to a return to realism, a movement that not only produced phenomenology (perhaps the dominant philosophy in Europe today) but also to some extent influenced Bertrand Russell and other realist philosophers in the English-speaking world.

BRENTANO. Brentano (1838–1917) held that the objects of psychology were mental acts. Each mental act had an immanent object—what Brentano called an intentional object—thus reviving scholastic terminology. These objects were a kind of internal accusative to the relevant act, as a judgment is to the act of judgment. They provide the content of the act. But what is the status of these objects? Clearly, the question is especially pertinent because it is possible to think of or make judgments about things that do not exist. How can a real act have an unreal object?

MEINONG. This was the problem that Meinong (1853–1920) took up. He postulated nonexistent objects to explain the possibility of our thinking, for example, of

things that do not or cannot exist. Similarly, false judgments were said to correspond to what he called objectives—nonexistent states of affairs which would be facts if only the corresponding judgments were true. Objectives could not be said to exist, for they were not things, but they might subsist. From a linguistic point of view, this doctrine implied a realist theory of meaning, according to which the meaning of any expression was given by a corresponding entity. The fact that these entities were not themselves mental entities (although they gave content to what is mental) implied a return to realism in a more general sense. Objects could be real, according to Meinong, without being actual.

HUSSERL. Edmund Husserl (1859–1938), another disciple of Brentano, started from very much the same point of view as Meinong, maintaining that the proper philosophical task was to investigate the essence of mental acts and their objects. Philosophy consisted, in his view, in an inquiry into the essence of different manifestations of consciousness and the essences with which they are concerned. To study this, it was necessary to strip off all presuppositions, metaphysical or otherwise. Husserl later emphasized this aspect increasingly. He adopted the method of *epoch*—the bracketing of presuppositions—in a manner akin, as he pointed out, to the Cartesian method of doubt. This would lead to pure consciousness as the one absolute, the one firm thing, and from this the philosopher may turn back to investigate the essence of different phenomena as they appear to consciousness. Thus, in effect the initial realist point of view led back to one that was more like idealism. But this belongs, properly speaking, to the twentieth century.

J. S. MILL. Meanwhile, in Britain the predominant philosophy at the beginning of the nineteenth century was sensationalism with its attendant associationism. James Mill (1773–1836) took a radically empiricist point of view, trying to reduce perception to merely having sensations and other mental phenomena to sensations plus the ideas associated with them. His son J. S. Mill (1806–1873) brought greater sophistication to this point of view and, in so doing, led to its downfall. Like his father, J. S. Mill wished to reduce all knowledge to experience, to the association of certain ideas with basic sensations. He expressed a great admiration for Berkeley's *New Theory of Vision* and its explanation of how we come to see things as at a distance. He thought it possible to explain in a similar way how we come to think of ourselves as perceiving a permanent world of things. We have expectations that take us beyond the immediate sensations because of the associations built up in experience between our immediate sensations and ideas of "permanent possibilities of sensation." Our ideas of material things are simply ideas of these permanent possibilities of sensation. Like Hume, Mill approaches this problem psychologically; he seeks to explain why we believe in an external world. To the extent, however, that he is inclined to add that things are simply these permanent possibilities of sensation, his view is the extreme empiricist doctrine of phenomenalism, the doctrine that all we mean by "material object" is something about our experiences.

Mill's main contributions to philosophy perhaps lie in logic, ethics, and politics. His general approach, however, is psychological, based on a conception of experience as atomistic sensations which could be linked to derivative ideas by the processes of association. Mill's general point of view is perhaps nowhere more obvious than in his account of such necessary propositions as those of mathematics. These are, in his view, simply very highly confirmed generalizations. The only necessity is psychological necessity.

Mill's view of knowledge came under attack in the latter half of the nineteenth century from many sources in Britain and elsewhere. In Britain perhaps the main attack came from the returning idealism, particularly as represented by F. H. Bradley. His main line of thought, as already discussed, was Hegelian with less emphasis on the dialectic and greater emphasis on the idealist point of view, according to which reality consists of experience organized in thought by attaching ideas to it in judgment. In *Appearance and Reality* Bradley sought to show that all features of the empirical world are only appearance and that reality must consist of a form of experience which is absolute and unitary and transcends all the contradictions of appearance. His criticism of Mill is that the pure sensory given is a myth; all the content of our knowledge must come by way of ideas—that is, through thought—and association "marries only universals." Experience in itself is nothing.

BERGSON. Criticisms of a different kind came from France and America. In France there were few philosophical developments of interest during the nineteenth century until Henri Bergson (1859–1941). Bergson was an anti-intellectualist who emphasized life against thought. Much of his approach was therefore biological. The space and time of which we are conscious, Bergson thought, are continuous; the division of it into things and processes is due to the intellect, which carries out the division according to our biological needs. Perception involves an aware-

ness of the possible moves that our body can make in relation to an object; in contrast, sensation corresponds to a simple response to a stimulus. Like Bradley, Bergson thought that the atomistic sensations of the sensationalist were a product of intellectual analysis. There are actually no basic experiences of this sort. We perceive things as our biological needs cause us to do so. Similarly with memory; our body acts like a sieve. Without the body our mind might remember everything, and this would be both useless and even disastrous biologically. Our body saves us from this, causing us to select only that which is biologically useful.

Because of this emphasis on biological utility, there is a relativism inherent in Bergson's point of view, and it has much in common with American pragmatism as instituted by William James. Bergson went further than James in his emphasis on life, however. His starting point was a thesis about time that is really outside the scope of this article. Roughly, his view was that the time of consciousness (*la durée*) is continuous; the ideas which thought presents to us are or seem discontinuous. The continuity of consciousness must be due to an interpenetration of those ideas, and as a result, they form a developing series in which, given that each member developed from what has gone before, each member must be unique. The development itself is due to a vital spirit, and the same is true of the universe at large.

JAMES. Bergson's emphasis upon the continuity of consciousness has its counterpart in James's thesis of the stream of thought. In his *Principles of Psychology* William James (1842–1910) insisted, in opposition to the sensationalists, that there were no atomic sensations or ideas; distinct ideas are selected phases of one stream of consciousness, and these phases make up a continuity because each idea has a fringe which overlaps that of its neighbors. Thus far, however, James was content to argue that our ideas are determined by what things there are. We must, he maintained, distinguish between knowledge by acquaintance and knowledge about something (not to be confused with a somewhat similar but really different distinction made by Russell). The baby is acquainted with the universe but he has not yet selected anything from the mass of sensation with which he is confronted. Thus, he knows nothing *about* anything. James was later to go further in Bergson's direction. In *Essays in Radical Empiricism* he rejected the distinction between thought and things, embracing a thesis that is known as neutral monism, the thesis that reality consists of one stuff (in this case experience) out of which both the mental and the physical are to be constructed according to the principles which govern each. From this continuous experience a plurality of thoughts and objects can be developed, a plurality of things related by concatenation only, as experience tells us they are. In this James set himself in opposition to his other bête noire, idealism, with its emphasis on internal relations and its denial of plurality.

PEIRCE AND DEWEY. James's special claim to fame (although some would say notoriety) is perhaps his status as the founder of pragmatism, although this, too, has Bergsonian affinities. The original source for the pragmatist point of view was C. S. Peirce (1839–1914), a rather isolated figure. A man with a great wealth of ideas, Peirce, as a nonprofessional philosopher, was to some extent outside the main stream of philosophical thought. He came to philosophy as a mathematician and scientist. He was opposed to all intuitions of the Cartesian type, largely because of his belief in the power of hypothesis and his disbelief in ultimate inexplicables. Perhaps his greatest contribution, however, lies in his theory of signs and meaning. In this connection Peirce maintained that our conception of anything is determined by our conception of the practical bearings of that thing. In sum, meaningfulness is a question of practical utility, and the meaning of any given concept or expression is given by its precise utility.

James turned this theory of meaning into a theory of truth, much to Peirce's disgust. He maintained that the test of the truth of any belief was its fruitfulness. To say that a belief is true is to say that it is good in this sense. Such is the pragmatic theory of truth.

For John Dewey (1859–1952) it is knowledge that is successful practice; propositions are merely instruments which may take us to the goal toward which experimental inquiry is directed. There is no final truth; instead, there is "warranted assertability" when the judgments which we make lead us to the abstract goal of science in accordance with scientific method.

TWENTIETH-CENTURY REALISM

The course of twentieth-century philosophy was not smooth, and it is therefore not easy to chart. No more than a sketch will be attempted here. Undoubtedly, however, the main philosophical event at the start of the century was a swing from idealism to realism. In America there were the neorealists, such as E. B. Holt, W. B. Montague, and R. B. Perry, influenced by James and his theory of neutral monism; in England there was Samuel Alexander, and in Germany there was Meinong. The most important figures in this revolution were G. E. Moore and

Bertrand Russell at Cambridge, and of these Moore was the originator in this respect.

MOORE. G. E. Moore (1873–1958) began with a criticism of Bradley. Moore thought that Bradley had not taken far enough his rejection of the view that we abstract our ideas from experience. Insisting that concepts or ideas should be regarded as the objects, the meanings, of our thoughts, Moore went on to argue that there must be propositions as the objects of beliefs. Things are merely collections of concepts and as such enter into propositions as their constituents. Propositions are what we believe when we hold any belief, true or false. This amounts to an insistence upon a distinction between any mental activity or form of awareness and its object. Moore's article "Refutation of Idealism" in his *Philosophical Studies* is founded on just this point, for he finds the refutation of the doctrine that *esse est percipi* in a distinction, somewhat after the manner of Brentano, between the act of awareness in perception and the object of awareness, between consciousness and its object. Idealists, he maintained, failed to notice the distinction.

Moore was later to give up his doctrine of propositions, for in considering the problem of false belief, he said that "there do not seem to be propositions at all, in the sense in which the theory demands them." If there were, there would have to exist something corresponding to false beliefs, and the fact of its existence would make the beliefs true, not false. Belief cannot therefore consist in a relation between ourselves and an object. The rejection of propositions in this objectively existing sense may look like the abandonment of the very foundations of his realism, and so in a sense it is. But Moore did not give up the view that we do know of a reality independent of our minds, and when he abandoned the account of false belief which implied the existence of objectively existing propositions, he nevertheless maintained categorically that we must not give up the view that truth somehow consists in correspondence with reality.

Existence of the external world. Moore sometimes said that he never doubted that we do know things about reality that we ordinarily think we know. Therefore, he was not influenced by the usual skeptical arguments against this position. In his view the real philosophical problem was to analyze what we mean when we say that we have this knowledge. Moore has generally been one who raises the difficulties about this problem rather than one who gives the answers. He has definitely maintained, however, that we do have knowledge of many different kinds of thing, and in his notorious "Proof of an External World" (*Proceedings of the British Academy*, 1939) he gave as a good argument for the existence of such a world the fact that we can point to objects in it. Thus, he held up his hands, saying, "Here is one hand, and here is another," to prove the existence of an external world. Moore's thought moved toward the view that what requires defending is common sense, ordinary beliefs such as the belief that there are objective things, like his hand, in the world. When metaphysicians say such things as "Time is unreal," this is an affront to common sense and demands explanation. Ludwig Wittgenstein later said that the view defended by Moore was not strictly common sense, since it was a philosophical point of view. This seems correct; what Moore meant by common sense was a general realist point of view.

Account of perception. In his analyses of what we mean when we claim knowledge, however, Moore's discussion follows the lines of those which have been more influenced by skepticism. Thus, in his account of perception he brings in sense data as what we actually see or directly apprehend when we look at something. Characteristically, he distinguishes between the sense datum (the object that we actually see) and the sensation which we might be said to have of it. But in using the words "direct apprehension" and "actually seeing," he suggests that he wants that indubitability which other philosophers have sought as an answer to the skeptic. Direct apprehension cannot be of the physical object, for, he argues, when an envelope is held up, it cannot be this that we actually see, since some people may fail to identify it as such. There is room for error about the identity of the object but not, perhaps, about its color or even shape as seen, so it must be these that we actually see. Moore's main worry is about the exact relation that exists between sense data and physical objects, both of which, he thinks, certainly exist. The answer that he would like to give is that sense data are parts of the surfaces of physical objects, but the fact that different people may have conflicting views of these objects prevents him from giving this answer. The right answer and, for the same reason, the precise nature of sense data were always a puzzle to him.

It may be wondered why Moore felt it necessary to bring in sense data at all. The answer must be that for Moore there had to be things which we just know, and although we may know of the existence of physical objects, we may not be so sure of their exact qualities. Hence, the possibility of error plays a part in Moore's thought just as it has done in that of others, even if the influence of skepticism is not so explicit. Direct apprehension fills the same role in his philosophy of perception

as intuition in his ethics. The notion has its parallel in the conception of knowledge employed by his Oxford contemporary John Cook Wilson and his disciples, such as H. A. Prichard. They thought the trouble with idealism lay in its failure to see that there was a distinct state of mind, knowledge, in which there was no possibility of error.

RUSSELL. Bertrand Russell (1872–1970) was first an idealist but was converted to realism by Moore. From his early study of Leibniz, Russell took the view that philosophy consists in the analysis of propositions, and his interest in logic also brought him to a concern with language. During the early 1900s he became interested in Meinong, whose realism seemed to confirm Moore in Russell's line of thought. However, he came to think that Meinong's supposition that there had to exist objects to explain our ability to think of things which do not exist in fact, such as round squares, showed, as he put it, an insufficiently robust sense of reality. Partly in response to this and partly in response to a more complex theory of meaning put forward by the mathematical logician Gottlob Frege, Russell set forth the theory of descriptions.

According to the theory of descriptions, phrases of the form "the so-and-so" are incomplete in meaning. They have no meaning (in the form of an object of reference) in themselves; to give their meaning, it is necessary to analyze the meaning of the whole sentence in which they occur. Sentences of the form "The so-and-so is *F*" are really tantamount to composite sentences including as a part the sentence asserting that something exists corresponding to the description "the so-and-so." Where nothing of the kind exists, the whole proposition is false ("proposition" here being equivalent to "statement," not the objectively existing object of beliefs). The theory of descriptions became a tool of analysis, and Russell used it in many connections.

Theory of knowledge. The main importance of the theory in epistemology is connected with Russell's distinction between knowledge by acquaintance and knowledge by description. Knowledge by acquaintance is Moore's direct apprehension, but Russell has always been more concerned with the justification of claims to knowledge than was Moore. For Russell it was important that all knowledge be founded on knowledge by acquaintance, if it was to be possible at all, for only in knowledge by acquaintance is error absolutely impossible. In *Problems of Philosophy* Russell gave a list of entities of which we have knowledge by acquaintance—sense data, memory data, the self, and universals. Of physical objects we have only knowledge by description because here error is possible.

Russell also declares that "every proposition which we can understand must be composed wholly of constituents with which we are acquainted." (By "proposition" he meant the objectively existing entity in the early Moorean sense.) This is possible only if anything of which we have knowledge by description is reducible to things of which we have knowledge by acquaintance. Physical objects, for example, must be reduced to sense data or, since they are not always being perceived, to a combination of sense data and sensibilia—actual and possible sense data. They must, as Russell puts it, be considered as logical constructions from sense data; they are simply bundles of sense data and sensibilia.

From the logical point of view, names of physical objects are disguised descriptions, and the theory of descriptions shows that it is not necessary to suppose the existence of a special class of entities called physical objects in order to give propositions about them a meaning. What we are acquainted with when we perceive a physical object is a number of sense data; the physical object we know only by description, and any statement expressing a fact about it is a statement about a description. This statement is analyzable, so that it contains an existential proposition about something answering to that description, according to the theory of descriptions; it is, that is, about something falling under a set of universals, and these are objects of acquaintance as much as sense data are. The notion of an object of acquaintance is closely connected with that of a logically proper name, an expression that cannot fail in its reference. Descriptions can, of course, fail in that there may be nothing answering to them. With the reduction of physical objects to sense data, knowledge of physics is preserved, and this has always been a cardinal point in Russell's program. On the actual status of sense data, Russell's opinion varied, but when most impressed by physics, he made them, paradoxically enough, entities in the brain.

Logical atomism. At the beginning of World War I, when Russell had come under the influence of Wittgenstein, he held that the ultimate constituents of the universe are atomic particulars, which are terms of relations in atomic facts. All other facts were to be built up from atomic facts by the processes of logic. The atomic particulars are sense data, and the relations supply the universal element in the fact.

The theory of logical atomism is a metaphysical theory rather than an epistemological one, and this is even more obviously true of Wittgenstein's version in his *Trac-*

tatus Logico-philosophicus. Russell differs from Wittgenstein in his more explicit interest in the nature of the constituents of atomic facts, and it is clear that his choice of these constituents was determined by his desire to found all knowledge on knowledge by acquaintance. The one exception to the program that Russell found at this stage was that of mental states like belief, for he could not see how statements about beliefs could be reduced to more elementary statements. In order to deal with the problem, he has flirted with behaviorism, since if belief is analyzable only in terms of behavior, this, in turn, may conceivably be explained in terms of sense data.

Russell changed his mind over the details of his logical atomism, especially when he was influenced by the logical positivists, but the framework remained the same. His account of memory is similar to that of perception; it is founded on memory experiences only contingently related to the past, although they may have a feeling of familiarity. The memory data are objects of acquaintance, but the past itself is not. In one general respect, however, Russell acknowledges that empiricism fails—in our knowledge of the postulates on which, in his opinion, inductive inference and, therefore, science rest. Induction, he thinks, is founded on habit, and the principles implicit in such habits cannot themselves be derived from experience. Despite his belief in the limitations of empiricism, Russell never wavered in his defense of realism. He always embraced and keenly defended the correspondence theory of truth.

Nature of mathematics. Nothing has been said thus far of Russell's work on the foundations of mathematics, especially his great contribution, written with A. N. Whitehead, in *Principia Mathematica*, although this is in a sense part of his epistemology. In *Principles of Mathematics* Russell held that mathematical propositions were synthetic, but when, influenced by Frege, he embarked on the attempt to reduce mathematics to logic by deriving it from a small number of axioms containing only logical notions, he held mathematical propositions to be analytic. This is too complicated a matter for discussion here. The exact nature of mathematical propositions is still a matter of dispute, but the attempt to reduce them to logic may now be seen to be a great and splendid failure.

WHITEHEAD. It may be noted briefly that Russell's partner in *Principia Mathematica*, A. N. Whitehead (1861–1947), took a very different road in epistemology. He tried to explain the properties of things in terms of their relations to one another. In perception the mind tries to grasp—or in Whitehead's term to "prehend"—a part of the system of nature around it; it is reacting to the environment in biological fashion. There are no atomic sense data; to suppose that there are is to be liable to the "fallacy of misplaced concreteness"—the view that because science has a concept (such as that of an instant), there must be entities of this sort in experience. To some extent Whitehead's thought is in the idealist tradition, but it also contains a certain Platonism. The particulars of which nature is composed, he thought, are events; the permanent characteristics that we recognize in them are objects. The objects are, as he puts it, ingredient "into events," not into just one event but through an indefinite neighborhood of events. We do not, that is, see single things with isolated characteristics; we view them as part of a system. Whitehead's thought is, however, difficult, and little can be done to make it intelligible in a short space.

LOGICAL POSITIVISM

Wittgenstein's *Tractatus* influenced a group of philosophers in Vienna who were mainly interested in the philosophy of science after the empiricist fashion of Ernst Mach. Wittgenstein had said that to understand a proposition is to understand what it would be like for it to be true. The Vienna circle, as this group became known, wrongly interpreted this as a general criterion of significance, and so the verification theory of meaning was born. According to this theory, meaningful propositions must be either analytic or empirically verifiable. The propositions of mathematics and logic were thought to belong to the first class, and the propositions of science to the second. Metaphysical propositions, belonging to neither group, were declared meaningless.

The members of the group differed over the details of this scheme, and a progressive relaxation in its rigor gradually took place. One of the biggest problems was the status of the verification principle itself, for on the face of it it is neither analytic nor empirically verifiable. The eventual outcome for some members of the group was to view it as a recommendation only. There were also other problems. The initial aim of the movement was, above all, to lay the foundations of science. Scientific propositions had to be preserved and metaphysics excluded. It became apparent that it was difficult, if not impossible, to provide a formulation of the verification principle which fulfilled both goals.

Moritz Schlick (1882–1936), the original leader of the group, felt compelled to interpret scientific laws as rules rather than statements. Another problem lay in the meaning of the phrase "empirically verifiable." Schlick held that ultimately there had to be a direct confrontation

with experience. Other members of the movement—for example, A. J. Ayer (1910–1989)—held that there had to be basic propositions which were directly and strictly verifiable (and thus absolutely incorrigible), although others could be indirectly verified by reference to these. This led to a distinction between strong and weak verification; propositions about physical objects, for example, might be only weakly verifiable, since an indefinite number of propositions about immediate experience would have to be invoked in order to verify them conclusively. In this, positivism was associated with the thesis of phenomenalism that statements about physical objects are analyzable into a collection of statements about sense experiences. The fact that such an analysis must be indefinitely long has resulted in a progressive modification of the thesis on the part of its main proponents—for instance, Ayer.

Truth. Schlick's view brought with it the correspondence theory of truth. Otto Neurath (1882–1945), on the other hand, held that this involved an attempt to go outside the web of language and ran the risk of a lapse into metaphysics. Neurath maintained that there had to be "protocol propositions"—observation reports on which science might be founded—but that these need not be reports of immediate experience. The truth of a protocol proposition was determined by its coherence with other propositions making up the language of science. Thus, Neurath embraced a coherence theory of truth. This general line was developed by Rudolf Carnap (1891–1970) into a form of conventionalism. He put forward a "principle of tolerance," maintaining that logic had no morals; verificationism thus became a proposal for the best way of developing the language of science. Whereas Schlick and the earlier Carnap had maintained that basic propositions must be about immediate experience, Neurath had maintained the thesis of physicalism that protocol propositions must be in the language of physics. In line with his relaxation of the criteria, Carnap came to accept what he called the "thing-language"—the language of the commonsense world—as the basis for scientific language. Today, positivism in its strict form is more or less a thing of the past.

Nature of science. It is noteworthy that Karl Popper (1902–1994), who was not a member of the movement but who was influenced by it and influenced it, held that the key to an understanding of science lay not in verifiability but falsifiability. He put this forward, however, not as a theory of meaning but as a criterion for the demarcation of science from metaphysics. The generalizations of science are, because of their very form, unverifiable, but they *are* falsifiable, whereas the propositions of metaphysics are not. Popper developed these views into a thesis about science as based on the hypothetico-deductive method. The aim of science is to put forward bold hypotheses, the deductive consequences of which must be subject to rigorous testing and criticism. This view is associated with a form of skepticism, for Popper sometimes maintains that we can never *know* the truth. The best that we can do is to put forward hypotheses and subject them to rigorous tests, for this is the way in which science progresses. Truth itself is just an illusion.

CONTEMPORARY MOVEMENTS

Contemporary philosophy is in an untidy state of nonuniformity. In Europe perhaps the most prevalent philosophy is a phenomenology deriving from Husserl. This movement is also associated with existentialism, originally a reaction against the superrationalism of Hegel and, therefore, to some extent a form of irrationalism. Existentialists have added little to epistemology; they tend to take for granted the existence of an objective world, aiming only to present a picture of it and of man's place in it. Those existentialists who derive something from Husserl—for example, Jean-Paul Sartre and Maurice Merleau-Ponty—are concerned mainly with descriptions of forms of consciousness, with phenomenology as descriptive psychology.

ORDINARY-LANGUAGE PHILOSOPHY. Perhaps the most significant movement, apart from latter-day positivism and pragmatism, is the so-called ordinary-language philosophy, today most closely associated with Oxford. The leading spirit of the movement is, however, the Cambridge philosopher Ludwig Wittgenstein (1889–1951), who has had an immense influence. His work is not easy to summarize; it is in part a series of comments upon his earlier logical atomist views and the theory of meaning that it espoused. Only a small part of his work can be mentioned here. He has criticized the attempts implicit in much sense-datum philosophy to construct a private language by arguing that the results of such attempts would lack the essential conditions of a language. There would be no way of distinguishing between the occasions on which one was following a rule in applying an expression and those on which one was making a new decision so to apply it. He has also stressed the importance of bringing back terms to the language game (as he calls possible languages) that is their original home—ordinary language. This, he maintains, is perfectly in order as it is; the important thing is to examine the uses to which expressions are put, with the recognition that language is a form of life and must be treated

accordingly. Among other things this has led to the recognition of truths which are necessary but not analytic, truths which he calls "grammatical." These are truths which express nonanalytic connections between concepts. The emphasis upon such truths and the arguments which lead to them on the part of followers of Wittgenstein is in a sense a partial return to Kant. (The distinction between analytic and synthetic truths in general has in any case come under fire from several quarters, especially from the American logician W. V. Quine [1908–2000]. But his emphasis on the necessity of assessing the status of propositions only within a system is something more of a move toward the idealist point of view.)

The appeal to ordinary language has been used for many purposes. Gilbert Ryle (1900–1976), for example, has used it in order to plot, as he puts it, the logical geography of mental concepts such as mind, belief, or will. He, too, has attacked the notion of sense data, and he has made and emphasized an important distinction between knowing how and knowing that. For present purposes the main importance of the appeal to ordinary language lies in its confrontation of the skeptic. The most stringent appeals to usage (rather than merely to the functions of language, as in Wittgenstein's case) have been made by J. L. Austin (1911–1960), who has emphasized the extent to which many philosophers including skeptics, have departed from our ordinary use of words. It is clear, however, as Austin in effect admitted, that an appeal to what we ordinarily say cannot settle these issues, however much it may be a good first move. Arguments are first required in favor of our ordinary way of speaking.

It has been argued, for example, that anyone who says that we never know anything but only believe or suppose it robs the concept of belief of an essential contrast with knowledge, without which it would be meaningless. This argument—the so-called argument from polar concepts—is invalid, because a philosopher can use the concept of belief as long as he has the concept of knowledge, as long as he knows what it would be like to know; he does not have to admit that anyone knows anything as a matter of fact.

Another argument is that we can never deny that we have knowledge altogether, because this would be denying the existence of the paradigm case by reference to which we have learned the meaning of the word *knowledge*. This argument—the so-called paradigm-case argument—fails, in the opinion of the present writer, because it assumes that meaning is given by the applications of a term. Whereas it might be difficult to see how we could have come to attach meaning to a term unless we had learned some of its applications, it is not logically impossible that we should have done so. Hence, more complicated arguments are required.

This is the situation. Most philosophers would agree that if we are to be said to know a proposition *p*, we must believe *p*, *p* must be true, and we must have good reasons for believing in *p*. There is perhaps little argument over the first two conditions, although there might be some hesitation over the details. The problem is what counts as good reasons for believing in *p*. In the ordinary way we recognize different reasons according to the nature of the proposition involved. The skeptic denies that any of these are sufficient, and it is impossible to produce any knockdown argument which will dispose of the skeptic's claim. Each application of this claim must be assessed on its own merits, and the answer to the skeptic must therefore be a dialectical one in the Socratic sense. But the very existence of recognized forms of knowledge presupposes that there must be such knowledge. This is, however, only a presumption, not a proof.

Traditionally, and for good reason, skepticism has had biggest sway in connection with claims to knowledge of objects of perception, knowledge derived from memory, knowledge of other minds, and inductive knowledge. In each of these cases the skeptic may present too high a standard of knowledge, which cannot, in the ordinary way, be attained. But the temptation to accept such a standard may be increased by adopting certain views about the nature of—for example, perception or memory. If perception is thought of as merely having sensations or memory as merely having images or ideas in the mind, there is necessarily a gap to be crossed from our own minds if there is to be objective knowledge in these fields. Hence, the talk of knowledge of an external world, external to our own minds. With a different conception of memory or perception, in which an essential connection is recognized between memory and the past or perception and an objective world, this gap is removed. This is not to get rid of the skeptical problem at one fell swoop, since it remains a question whether and under what conditions these concepts of perception and memory can be applied. This, however, is merely the general problem of how knowledge is possible in different fields; it can be dealt with dialectically. In other words, an incorrect conceptual analysis can worsen the skeptical problem, but a correct one cannot solve it. But the presumption remains that objectivity is possible in these fields, even if it is incapable of proof. Each doubt can be alleviated only by argument; there is no overall answer. Finally, the supposition

that an answer can be provided by showing that there are forms of knowledge in which error is logically excluded and which are therefore absolutely indubitable is an illusion. First, there is no indubitable knowledge; second, it is not necessary for the general possibility of knowledge that there should be. Much of the history of epistemology has depended on this illusion.

See also Abelard, Peter; Alexander, Samuel; Antisthenes; A Priori and A Posteriori; Arcesilaus; Aristotelianism; Aristotle; Augustine, St.; Austin, John Langshaw; Ayer, Alfred Jules; Bergson, Henri; Berkeley, George; Boethius, Anicius Manlius Severinus; Bradley, Francis Herbert; Brentano, Franz; Carnap, Rudolf; Carneades; Cartesianism; Chrysippus; Cosmological Argument for the Existence of God; Couturat, Louis; Descartes, René; Dewey, John; Dialectic; Duns Scotus, John; Empiricism; Epicurus; Fichte, Johann Gottlieb; Frege, Gottlob; Geulincx, Arnold; Gorgias of Leontini; Hegel, Georg Wilhelm Friedrich; Hellenistic Thought; Heraclitus of Ephesus; Hobbes, Thomas; Holt, Edwin Bissell; Hume, David; Husserl, Edmund; Innate Ideas; Intuition; James, William; Kant, Immanuel; Knowledge, A Priori; Leibniz, Gottfried Wilhelm; Leucippus and Democritus; Locke, John; Logical Positivism; Mach, Ernst; Malebranche, Nicolas; Meinong, Alexius; Merleau-Ponty, Maurice; Mill, James; Mill, John Stuart; Montague, William Pepperell; Moore, George Edward; Neo-Kantianism; Neoplatonism; Neurath, Otto; New Realism; Ontological Argument for the Existence of God; Paradigm-Case Argument; Parmenides of Elea; Peirce, Charles Sanders; Perry, Ralph Barton; Plato; Platonism and the Platonic Tradition; Plotinus; Popper, Karl Raimund; Porphyry; Pre-Socratic Philosophy; Protagoras of Abdera; Quine, Willard Van Orman; Rationalism; Realism; Reid, Thomas; Roscelin; Ryle, Gilbert; Sartre, Jean-Paul; Schlick, Moritz; Schopenhauer, Arthur; Sextus Empiricus; Skepticism, History of; Socrates; Spinoza, Benedict (Baruch) de; Stoicism; Thomas Aquinas, St.; Universals, A Historical Survey; Whitehead, Alfred North; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

For the texts of individual philosophers see the separate articles devoted to them.

INTRODUCTORY BOOKS

Ayer, A. J. *The Problem of Knowledge*. London: Macmillan, 1956.

Russell, Bertrand. *The Problems of Philosophy*. London: Williams and Norgate, 1912.

Woozley, A. D. *Theory of Knowledge*. London, 1949.

GENERAL HISTORIES

Copleston, Frederick. *A History of Philosophy.* 7 vols. London: Burns, Oates and Washbourne, 1947–1975.

Hamlyn, D. W. *Sensation and Perception*. London: Routledge and K. Paul, 1961. A history of the philosophy of perception; contains bibliography.

O'Connor, D. J. *A Critical History of Western Philosophy*. New York: Free Press of Glencoe, 1964.

OTHER RECOMMENDED TITLES

Ayers, Michael. *Locke: Epistemology and Ontology*. London: Routledge, 1991.

Buryeat, M. F., ed. *The Skeptical Tradition*. Berkeley: University of California Press, 1983.

Fine, G. "Knowledge and Belief in Republic V." *Archiv fur Geschichte der Philosophie* 60 (1978): 121–139.

Fogelin, Robert. *Hume's Scepticism in the Treatise of Human Nature*. London: Routledge and Kegan Paul, 1985.

Hankinson, R. J. *The Sceptics*. London: Routledge, 1995.

Lehrer, Keith. *Thomas Reid*. London: Routledge, 1989.

Lovejoy, Arthur O. *The Revolt against Dualism*. La Salle, IL: Open Court, 1930.

Moser, Paul K. "Epistemology (1900–Present)." In *Routledge History of Philosophy.* Vol. 10: *Philosophy of the English Speaking World in the 20th Century*, edited by John Canfield. London: Routledge, 1996.

Moser, Paul K., and Arnold VanderNat, eds. *Human Knowledge: Classical and Contemporary Approaches*. Oxford: Oxford University Press, 1987; 2nd ed., 1995; 3rd ed., 2002.

Pears, David. *Hume's System*. Oxford: Oxford University Press, 1990.

Popkin, R. *The History of Scepticism from Erasmus to Spinoza*. Berkeley: University of California Press, 1979.

Shope, R. *The Analysis of Knowing*. Princeton, NJ: Princeton University Press, 1983.

Sosa, Ernest. "How to Resolve the Pyrrhonian Problematic: A Lesson from Descartes." *Philosophical Studies* 85 (1997): 229–249.

Stroud, Barry. *Hume*. London: Routledge and Kegan Paul, 1977.

Stroud, Barry. *The Significance of Philosophical Scepticism*. Oxford: Oxford University Press, 1984.

White, Nicholas. *Plato on Knowledge and Reality*. Indianapolis: Hackett, 1976.

Williams, Bernard. *Descartes: The Project of Pure Enquiry*. Atlantic Highlands, NJ: Humanities Press, 1978.

D. W. Hamlyn (1967)
Bibliography updated by Benjamin Fiedor (2005)

EPISTEMOLOGY, HISTORY OF [ADDENDUM]

Knowledge (*ʿilm*) has occupied a central place in the Islamic intellectual tradition. The religious incentive for this stems from the fact that the Islamic belief is

grounded in a knowledge claim about God's existence and His revelation. The theologians (*mutakallimin*) consider knowledge as a prerequisite for religious belief (*iman*). A related question is God's knowledge of things—how God as the knower (*al-ʿalim*) knows particulars, which are subject to change, without change in His essence. To address this issue, Ibn Sinā had absolved God of the necessity of knowing every particular thing and event because this might introduce change in his unchanging essence. Ghazālī, in turn, accuses Ibn Sinā of denying God the ability to know particulars. The general consensus on God's knowledge of things, however, is that His knowing is a generative act in that He knows things by creating them. In this sense, God's knowledge of things does not follow their existence, which would be to attribute ignorance to God, but precedes them.

In philosophy, four major theories of knowledge have developed. The first is the concept of knowledge as abstraction (*tajarrud*). Following Aristotle, the Muslim Peripatetics define knowledge as the abstraction of the intelligible forms of things from their material properties. We know things through their intelligible forms—that is, only as universals. When the mind encounters a particular object, it abstracts its form and turns it into a conception in the mind. This, however, raises the question of whether what we know is a universal or the things themselves.

Knowledge as abstraction leads to what we might call the representational theory of knowledge, according to which knowledge is an imprint or picture (*rasm*) of actually existing things in the mind. When there is a perfect correspondence between the thing and its representation in the mind, we arrive at true knowledge.

The second theory is based on knowledge as a relation (*idafah*) between the thing known and the knower. The knower intends things in the extramental world, and this intending creates a relation between a person and his or her object of knowledge. Defended by some theologians, knowledge as relation fails to account for self-knowledge where the knower and the known are one and the same thing. When I say, for instance, that I feel pain, object and subject are the same. Otherwise, we would have to admit a relation between myself and myself.

The third theory defines knowledge as a property of the knower. Knowledge belongs to the knower as a state of the mind (*hala nafsaniyyah*). Combining elements from the above theories, this view reduces knowledge to concepts in the mind. It also seems to suggest that what the mind knows is its internal procedures rather than the things in the external world. The goal of knowledge, however, is to know things as they are, not simply as they appear to us.

The fourth theory of knowledge proposed by Suhrawardi and developed by Mullā Ṣadrā argues for what is called knowledge-by-presence (*al-ʿilm al-huduri*). Suhrawardi defines knowledge as presence rather as absence or negation, as the Peripatetic term "abstraction" implies. Ṣadrā takes this point further and equates knowledge with existence (*wujud*). According to him, knowledge is a mode of existence and knowing is a cognitive act of unveiling an aspect of the all-inclusive reality of existence. This leads Ṣadrā to his celebrated defense of the unification of the intellect and the intelligible.

Bibliography

Aminrazavi, Mahdi. *Suhrawardi and the School of Illumination*. Surrey, U.K.: Curzon, 1997.

Chittick, William. *The Heart of Islamic Philosophy: The Quest for Self-Knowledge in the Teachings of Afdal al-Din Kashani*. Oxford and New York: Oxford University Press, 2001.

Rosenthal, Franz. *Knowledge Triumphant: The Concept of Knowledge in Medieval Islam*. Leiden: E. J. Brill, 1970.

Mehdi Ha'iri Yazdi. *The Principles of Epistemology in Islamic Philosophy: Knowledge by Presence*. Albany, NY: State University of New York Press, 1992.

Hossein Ziai. *Knowledge and Illumination: A Study of Suhrawardi's Hikmat al-ishraq*. Atlanta: Scholars Press, 1990.

Ibrahim Kalin (2005)

EPISTEMOLOGY, RELIGIOUS

The epistemology of religion, as practiced by philosophers, is seldom concerned with the sorts of epistemological questions that emerge on a practical level in ordinary religious life, such as how to determine the correct interpretation of a scriptural text or how to know whether someone's claim to special divine guidance is to be credited. Rather, it tends to focus on the epistemic evaluation of the most basic tenets of the religious worldview in question—the existence of God, the creation of the world and God's relation to it, and the possibility of recognizing divine action in the world and divine revelation. From the 1960s on, religious epistemology has been characterized by a marked decline of fideism, with a renewal of interest in evidentialism and an even more pronounced upsurge of what may be termed experientialism.

Fideism is best characterized as the view that one's basic religious beliefs are not subject to independent rational evaluation. It is defended by urging that religious

convictions are the most basic part of a believer's worldview and thus more fundamental than anything else that might be used to evaluate them. It is also said that to evaluate religious beliefs by standards other than the internal standards of the religious belief system itself is in effect to subject God to judgment and is thus a form of idolatry. In the mid-twentieth century fideism took two main forms, existentialism and Wittgensteinian fideism. In the succeeding decades philosophical existentialism has suffered a massive decline, as has its theological counterpart, neoorthodoxy. Wittgensteinian fideism, on the other hand, arose largely in response to the positivist contention that God-talk is cognitively meaningless; with the defeat of positivism it has lost much of its relevance. Many religious thinkers, freed from the need to defend religion's cognitive meaningfulness, have felt a renewed impulse to contend for the truth of their faith. And on the other hand, critics of religion have moved readily from the contention that belief in God is meaningless to the logically incompatible assertion that it is false and/or lacking in evidential support.

Evidentialism is the view that religious beliefs, in order to be rationally held, must be supported by other things one knows or reasonably believes to be true. Evidentialist defenses of religion typically rely heavily on theistic arguments, and all of the classical arguments saw renewed interest in the late twentieth century. Versions of the ontological argument propounded by Charles Hartshorne, Norman Malcolm, and Alvin Plantinga are clearly valid, though their premises remain controversial. William Rowe's work has directed renewed attention to the Clarke-Leibniz version of the cosmological argument, and new versions of the design argument teleological argument for the existence of God, focusing on God as the source of the basic laws of nature, have been developed by Richard Swinburne and others. Even the moral argument (Robert Adams) and the argument from religious experience (Gary Gutting) have come in for renewed attention. Two new arguments, or versions of arguments, are keyed to developments in cosmology. The "kalam cosmological argument" (William Craig) uses big-bang cosmology to argue that the physical universe as a whole has a temporal beginning and thus is in need of an external cause. And the anthropic cosmological principle is used by John Leslie, among others, to support a new version of the design argument: The apparent fact that the basic laws and initial conditions of the universe are "fine tuned" for life, with no apparent scientific explanation for this fact, is taken as evidence of intelligent design. Both of these arguments benefit from their association with cutting-edge science but also in consequence become vulnerable to future changes in scientific thinking on cosmology.

Evidentialist arguments against religion take a variety of forms. Most basically, evidentialists argue that the theistic arguments are unsuccessful and that theism fails for lack of evidential support. There are various challenges to the coherence and logical possibility of the traditional divine attributes. In most cases, however, these arguments, if successful, lead to a reformulation of the attributes in question rather than to the defeat of theism as such. But by far the most active area of consideration for antireligious evidentialism has been the problem of evil; the volume of writing on its various forms, by both critics and believers, has probably exceeded that on all of the theistic arguments taken together.

Along with the renewed consideration of the various arguments there have been reflections on the requirements for a successful argument. Traditional natural theology claimed to proceed from premises known or knowable to any reasonable person (e.g., "Some things are in motion") by means of arguments any reasonable person could see to be valid. By these standards it is not difficult to show that all of the arguments fail. But the standard is clearly too high; it is difficult to find significant arguments in any area of philosophy that meet it. No doubt a good argument should not be circular or question begging, and its premises must enjoy some kind of support that makes them at least plausible. But what seems plausible, or even evidently true, to one person may not seem so to another, equally rational, person; thus, the recognition emerges that arguments and proofs may be "person-relative" (Mavrodes 1970).

Furthermore, even a good argument is not necessarily decisive by itself, so it is necessary to consider the ways in which a number of arguments, none of them in itself conclusive, can lend their combined weight to establishing a conclusion. One model for this has been developed by Basil Mitchell, who compares arguments for religious beliefs to the kinds of cumulative-case arguments found in fields such as history and critical exegesis as well as in the choice between scientific paradigms. Richard Swinburne, by contrast builds a cumulative case for divine existence using the mathematical theory of probability. While it is not possible to assign precise numerical probabilities to the propositions involved in theistic argumentation, Bayes's theorem does provide insight into the way in which evidence contributes in a cumulative fashion to the support or defeat of a hypothesis such as theism Bayesianism. In addition, John L. Mackie and Michael

Martin have developed what are in effect cumulative-case arguments for atheism.

EXPERIENTIALISM

The most significant development in the epistemology of religion during the 1980s and early 1990s was the rise of a new type of theory distinct from both fideism and evidentialism. This theory, found in the writings of Richard Swinburne, Alvin Plantinga, and William Alston, lacks a generally recognized label (the term *Reformed epistemology* properly applies only to Plantinga's version) but may be termed experientialism in view of its emphasis on the grounding of religious belief in religious experience. Experientialism differs from fideism in that it does not seek to insulate religious belief from critical epistemic evaluation: rather, it affirms that religious experience can provide a sound epistemic basis for such beliefs. Experientialism is also importantly different from the evidentialist "argument from religious experience" in the following respect: The religious experience is not first described in ontologically neutral terms and then made the basis for an inference to the existence of the religious object. On the contrary, the religious belief is grounded directly in the religious experience, without mediation by inference, just as perceptual beliefs are grounded directly in perceptual experience.

This difference is important for a couple of reasons. For one thing it is more faithful to the phenomenology of both religious and perceptual belief: in typical cases neither form of belief involves such an inference. But more important, the direct grounding of belief in experience offers better prospects of a favorable epistemic status for the resulting beliefs than does the inferential approach. This is readily apparent in the case of perceptual experience: attempts at a "proof of the external world" have been notably unsuccessful, yet only those in the grip of philosophical theory doubt that we do in fact acquire a great deal of knowledge about the world through our perceptual experience. In the same way it is at least conceivable that believers acquire knowledge of God experientially even if no compelling inferential argument from religious experience is available.

Swinburne, Plantinga, and Alston share what may be termed a weak foundationalist approach to epistemology. That is to say, they accept the distinction between "basic" beliefs, which do not derive their rational acceptability from other beliefs, and "derived" beliefs, which gain their support from the basic beliefs. But they do not accept the traditional foundationalist restriction of basic beliefs to those that are nearly or entirely immune to doubt—beliefs that are self-evident, evident to the senses, or incorrigible. Each of them, furthermore, includes some religious beliefs in the category of basic beliefs. The epistemological task, then, is to show that this inclusion is epistemically proper—to show that such religious beliefs are among our "properly basic beliefs." (The terminology is Plantinga's, but the issue is the same for all three thinkers.) Each of them approaches this issue in a different way, though the approaches are ultimately compatible. Plantinga argues, following Roderick Chisholm, that the proper approach to the question of which beliefs are properly basic is inductive: one first conducts an inventory of the beliefs one takes oneself to hold rationally, then eliminates those that derive their epistemic support from other beliefs, and those that remain will be taken as properly basic. The typical Christian believer, Plantinga thinks, will find that she considers her belief in God to be rational but does not ground it inferentially on other beliefs she holds; thus, she will conclude that this is a properly basic belief. To be sure, atheists or believers in other religions will not concur in this, but Plantinga finds this to be unproblematic: "Followers of Bertrand Russell and Madalyn Murray O'Hair may disagree: but how is that relevant? Must my criteria, or those of the Christian community, conform to their examples? Surely not. The Christian community is responsible to *its* set of examples, not to theirs" (Plantinga and Wolterstorff 1983, p. 78).

In contrast with Plantinga's "internal" justification of the rationality of belief, both Swinburne and Alston attempt to show that religious experiences should have some epistemic weight, even for those who do not share the belief system the experiences ostensibly support. Swinburne appeals to the "principle of credulity," which states that "(in the absence of special considerations) if it seems (epistemically) to a subject that x is present, then probably x is present; what one seems to perceive is probably so" (1979, p. 254). He argues that a general denial of this principle lands us in a "sceptical bog" and that there is no justification for excluding religious experience from its scope.

Alston develops a "doxastic practice" approach to epistemology (indebted to both Thomas Reid and Ludwig Wittgenstein), which holds that all socially established doxastic practices are "innocent until proved guilty"; "they all deserve to be regarded as prima facie rationally engaged in … pending a consideration of possible reasons for disqualification" (1991, p. 153). Alston's delineation of the "Christian mystical practice" and his defense of its epistemic status constitute a systematic,

detailed, and highly sophisticated presentation of experientialism.

One major difficulty for experientialism is the existence of incompatible experientially grounded beliefs in different religions—in Alston's terms, the existence of a plurality of mutually incompatible mystical practices. Alston concludes that religious experience alone probably cannot resolve this ambiguity and that "the knowledgeable and reflective Christian should be concerned about the situation … [and] should do whatever seems feasible to search for common ground on which to adjudicate the crucial differences between the world religions, seeking a way to show in a non-circular way which of the contenders is correct. What success will attend these efforts I do not presume to predict. Perhaps it is only in God's good time that a more thorough insight into the truth behind these divergent perspectives will be revealed to us" (1991, p. 278).

Critics, however, have urged more far-reaching objections to the experientialist program. According to Richard Gale, the analogy between religious experience and sense perception is weak, with the dissimilarities far outweighing the similarities. He also argues that religious experience could not be cognitive—that is, could not provide independent grounds for belief in the existence of its object—and that religious objects such as God or the One are not possible objects of perceptual experience, even if they exist. Alston, on the other hand, has argued in detail that the phenomenological structure of religious experience is perceptual and that "mystical perception" constitutes a genuine species of perception along with sense perception.

See also Alston, William P.; Atheism; Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Chisholm, Roderick; Clarke, Samuel; Cosmological Argument for the Existence of God; Evil, The Problem of; Existentialism; Fideism; Leibniz, Gottfried Wilhelm; Mackie, John Leslie; Malcolm, Norman; Moral Arguments for the Existence of God; Ontological Argument for the Existence of God; Philosophy of Religion; Plantinga, Alvin; Positivism; Probability and Chance; Reid, Thomas; Religious Experience; Religious Pluralism; Teleological Argument for the Existence of God; Theism, Arguments For and Against; Wittgenstein, Ludwig Josef Johann.

Bibliography

Alston, W. P. *Perceiving God: The Epistemology of Religious Experience.* Ithaca, NY: Cornell University Press, 1991.

Gale, R. M. *On the Nature and Existence of God.* Cambridge, U.K.: Cambridge University Press, 1991.

Mackie, J. *The Miracle of Theism: Arguments for and against the Existence of God.* New York: Oxford University Press, 1982.

Martin, M. *Atheism: A Philosophical Justification.* Philadelphia: Temple University Press, 1990.

Mavrodes, G. *Belief in God: A Study in the Epistemology of Religion.* New York: Random House, 1970.

Mitchell, B. *The Justification of Religious Belief.* New York: Oxford University Press, 1981.

Plantinga, A., and N. Wolterstorff, eds. *Faith and Rationality: Reason and Belief in God.* Notre Dame, IN: University of Notre Dame Press, 1983.

Swinburne, R. *The Existence of God.* Oxford: Clarendon Press, 1979.

William Hasker (1996)

EPISTEMOLOGY, RELIGIOUS [ADDENDUM]

The most significant development in religious epistemology at the beginning of the new millennium was the completion of Alvin Plantinga's trilogy on warrant and religious knowledge. Plantinga's earlier work on Reformed epistemology focuses on the question of whether religious beliefs can be justified, in the sense that they can be accepted without violating any epistemic duties. His later work is concerned with *warrant*, defined as that which, added to true belief, enables such belief to qualify as knowledge. Plantinga argues convincingly that warrant is distinct from justification and also from rationality, in any of the several senses of the latter term. His own view is most akin to reliabilism, but he argues that standard versions of reliabilism face debilitating objections, and comes out instead for a definition of warrant in terms of the proper functioning of a person's epistemic faculties.

For these faculties to function properly, they must function as they were designed to function, and they must be functioning in an appropriate environment, the kind of environment for which they were designed. (The notion of *design* seems already to bring something like a theistic assumption into play. Plantinga, however, concedes provisionally that evolution may be thought of as "designing" people's epistemic equipment, though in the end he thinks this cannot be spelled out satisfactorily.) Furthermore, a person's faculties must be such that, in those circumstances, they function reliably, so as to produce as outputs a high proportion of true rather than false beliefs. Plantinga's formal definition of warrant is: "A belief has warrant for a person *S* only if that belief is pro-

duced in *S* by cognitive faculties functioning properly (subject to no dysfunction) in a cognitive environment that is appropriate for *S*'s kind of cognitive faculties, according to a design plan that is successfully aimed at truth" (Plantinga 2000, p. 56).

Given this epistemological framework, can belief in God be warranted when this belief is held in a basic way and not derived from other held beliefs? Plantinga's answer to this is yes. He holds that there is a component in the cognitive equipment of every person that is specifically designed to produce a belief in God, given certain "inputs" that are commonly available in our ordinary environment. Such inputs would include such situations as when people contemplate the majesty of the starry heavens, find God speaking to them in the Bible, or feel disgusted because of something they did wrong. This component in thehuman cognitive makeup Plantinga calls—following John Calvin—the *sensus divinitatis* (sense of divinity). When the *sensus* does its work, and produces in someone a belief in God, it is doing exactly what it is designed to do. Furthermore, the sensus is reliable because the belief that it regularly produces—namely, a belief that there is a God—is in fact true. (Malfunction, leading to a distorted conception of God, is of course possible.) Belief in God, produced in this way, satisfies all the conditions for being warranted, and when the belief is held with sufficient firmness the believer may be correctly said to know that there is a God.

Plantinga wishes to claim warrant also for the specific doctrinal beliefs of Christianity—for the "great things of the Gospel" (Plantinga 2000, p. 80). Considerably simplified, the model he presents is as follows: God, desiring to reveal himself, has become the principal author of the Scriptures of the Old and New Testaments, which contain his message to humankind. The individual who becomes aware of the teachings of the Gospel, through reading them or otherwise hearing of them, may come to have faith in these teachings; this faith is produced by the Holy Spirit, and has both a cognitive dimension (the teachings are "revealed to our minds") and a volitional/affective dimension (they are "sealed to our hearts"). One then comes to believe these teachings in a basic way; they are not inferred from anything else one believes, but as Jonathan Edwards said, the spiritually enlightened "believe the doctrines of God's word to be divine, because they see divinity in them" (Plantinga 2000, p. 259). A somewhat surprising (and perhaps implausible) consequence of this is that the fact that these doctrines are taught in Scripture does not constitute any part of the warrant they have for the typical believer. Plantinga argues, however, that beliefs so formed can be—and in typical cases are in fact—warranted when held in a basic way.

This model has been described on the assumption that the Christian faith is true, but that is an assumption Plantinga qua philosopher is not entitled to. Accordingly, his formal conclusion is not that belief in Christianity is warranted, but that, if Christianity is true, then belief in its truth is probably warranted. (If it is not true, then God has not endowed humans with the *sensus divinitatis*, nor is he the principal author of Scripture, nor is faith produced in believers by the Holy Spirit, as the model claims. In this case, belief in the truth of Christianity has other sorts of causes, and is probably *not* warranted.) Plantinga concludes that the de jure objection, which claims that Christian belief is unwarranted, cannot stand on its own without support from the de facto objection, that Christianity is in fact false.

As one may expect, all of Plantinga's principal claims have met with vigorous criticism. Internalist epistemologists such as Richard Fumerton regard his externalist proper functionalism as overly permissive in the beliefs it counts as warranted. Another criticism is directed at the somewhat negative and defensive character of his apologetic. Plantinga defends the propriety of holding certain kinds of basic beliefs, but this may be of little help to those (including many believers) who do not actually find themselves believing these things in a basic way. Richard Swinburne (2001), along with a number of others, has urged that Christian apologetics ought to go beyond this and present reasons that could be convincing to the inquiring nonbeliever. Plantinga, however, considers that determining the truth of Christian belief lies "beyond the competence of philosophy" (Plantinga 2000, p. 499).

See also Plantinga, Alvin.

Bibliography

Kvanvig, Jonathan, ed. *Warrant in Contemporary Epistemology; Essays in Honor of Plantinga's Theory of Knowledge*. Totowa, NJ: Rowman and Littlefield, 1996.

Plantinga, Alvin. *Warrant and Proper Function*. New York: Oxford University Press, 1993.

Plantinga, Alvin. *Warranted Christian Belief*. New York: Oxford University Press, 2000.

Plantinga, Alvin. *Warrant, the Current Debate*. New York: Oxford University Press, 1993.

Plantinga, Alvin, R. Douglas Geivett, Greg Jesson, Richard Fumerton, Keith E. Yandell, and Paul K. Moser. Symposium on *Warranted Christian Belief. Philosophia Christi* 3 (2) (2001): 327–400.

Swinburne, Richard. "Plantinga on Warrant." *Religious Studies* 37 (2) (2001): 203–214.

William Hasker (2005)

EPISTEMOLOGY AND ETHICS, PARALLEL BETWEEN

Usually, actions are taken and policies adopted to realize envisaged goals, and they are undertaken because of belief that they will probably realize the goals. Actions and policies may be criticized, then, on one of two grounds: that the goals are ill-chosen or that the belief that the actions or policies will probably achieve the goals is ill-founded. It is interesting, and perhaps indicative of the facts to be examined below, that many words of appraisal—such as "justified," "warranted," "reasonable," "right"—are used, although possibly in slightly different senses, to indicate the acceptability or unacceptability of both goals and beliefs. Moreover, such appraisals obviously have two features in common. First, the appraisal of a particular goal or belief can be made only in view of some general principle or standard; second, the standards and principles in question are not self-certifying, and their rational justification must be a serious question for a thoughtful person. The appraisal of actions and policies thus raises two questions about both goals and beliefs: What are the proper principles or standards to be used in appraisals, and what is the rational basis for regarding any principle or standard as proper?

Historically, ethics has been the philosophical discipline concerned with these two questions about goals. (However, "goals" must be taken broadly to include not only the question of ultimate values but also the question of moral obligations and moral rights.) And historically, epistemology has been the discipline concerned with the same questions about beliefs. (Again, "beliefs" must be taken broadly to go beyond mere predictions of consequences of the use of certain means to include theories, explanations, and systems of mathematical thought.)

In order to develop the parallel between ethics and epistemology, it is convenient to identify ethical and epistemological statements. Ethical statements are those that imply a statement that could be expressed by some English sentence containing essentially "is a good thing that," "is a better thing," or "is morally obligatory that" (on the assumption that "morally obligatory" does not introduce special concepts different from a phrase such as "it is morally wrong not to"). The class of statements thus specified will be identical with the class of statements that moral philosophers have traditionally been concerned with. Similarly, epistemological statements are those that imply a statement which could be expressed by some English sentence containing essentially "It is reasonable (or warranted) for a person *S* to place more confidence in *h* than in *i*," in which it is understood that for *h* and *i* can be substituted expressions of the form "its being true that. …" The class of statements thus specified is identical with the class of statements that epistemologists have been concerned with. It is useful to identify as ethical terms those phrases whose occurrence in a sentence distinguish it as an ethical statement and to identify as epistemic terms those phrases whose occurrence in a sentence characterize it as an epistemological statement.

PROBLEM OF ETHICS AND EPISTEMOLOGY

Moral philosophers have not, at least qua moral philosophers, been concerned with the acceptability of particular ethical statements such as "It would be a good thing if Jones learned to play the piano." Rather, moral philosophers have attempted to arrive at acceptable universal ethical statements which could serve as standards for the appraisal of particular situations. Thus, moral philosophers have defended or criticized such statements as "Enjoyment is always a good thing, abstracted from all consideration of consequences; nothing else is so" and "An action is morally right if and only if it will produce consequences as good as those of any other action the agent could have performed instead." The formulation of such generalizations, together with the proposal of reasons in support of them, is generally called normative ethics. Like moral philosophers, epistemologists are not concerned with the acceptability of particular epistemological statements such as "It is now highly probable that viruses are the cause of some forms of cancer." Instead, they have attempted to arrive at acceptable universal epistemological statements to be used as standards in appraising particular statements. Thus, they have examined the acceptability of such statements as "If at a time *t* person *S* seems to remember that he had the experience *E* at an earlier time, then he is initially warranted in believing that he did have the experience *E*" or "If at time *t* person *S* believes the statement *h* about his own experience at *t*, then it is reasonable for *S* to place at least as much confidence in *h* as in any other statement." The formulation of statements like these and the proposal of reasons in support of them have traditionally been the main

occupation of epistemologists. In order to distinguish this task from other concerns of epistemologists, we may call it epistemology proper.

SOME MINOR POINTS OF SIMILARITY. If it is morally wrong for a person to take action *A*, but he takes that action, then in the absence of any excuse we attribute to him a fault of character and say he is morally blameworthy. Similarly, if a person has good reason for believing a certain statement but he does not, then in the absence of any good excuse we attribute to him an intellectual fault and characterize him as intellectually open to criticism. (The excuse in either case might be much the same; for instance, a person might plead that he was very upset, not "himself.") It is sometimes said that the parallel extends further in another direction, that just as there are several senses of "morally obligated," so there are corresponding senses of "reasonable to believe." For instance, it is widely believed that "morally obligatory" is sometimes used to mean the act which a being omniscient about the facts of the case and about moral principles would be morally blameworthy for not doing if he were in the place of the agent; this is said to be a sense of "morally obligatory" different from that employed when a person with possibly faulty information about the facts and imperfect clarity about moral principles is said to be morally obligated to do something. Correspondingly, it is sometimes suggested that there is a sense of "reasonable to believe" in which we may say that it is reasonable for a person to believe that any statement is *true*; this sense is contrasted with the sense in which we say it is reasonable for a man to believe what is supported by the evidence he has. Whether there are such different senses in either case we must leave an open question here.

EPISTEMOLOGY REDUCIBLE TO ETHICS?

Some philosophers (R. Chisholm, for example) have thought that there is more than just a parallel between epistemology and ethics. They have thought that epistemic terms are properly defined by means of ethical terms. If this is correct, epistemological statements are complicated ethical statements, and, presumably, epistemology is a branch—doubtless a somewhat special one—of ethics. In accordance with this view, for example, the statement "It is warranted for *S* to place more confidence in *h* than *i*" might be taken to mean "For any good thing *G*, if *S* had to choose between risking it by a wager on the truth of *h* or risking it by a wager on the truth of *i*, he would be obligated to wager on *h*." If sound, this definition has the advantage of reducing the number of undefined terms in one's total system of concepts. The disadvantages of the definition are (1) that it is doubtful whether there is any useful sense of "obligated" in which the implied equivalence is true, (2) that the definition seems to be more obscure than the definiendum, and (3) that it does not seem that the meaning of "warranted belief" involves the notion of moral obligation but, conversely, that a person's being obligated to do something, in one ordinary sense, can be explained only by reference to the propositions it is reasonable for him to believe about his situation.

THEORIES OF MEANING AND VERIFICATION

If epistemology is not reducible to ethics, there is still a parallel between the higher-order questions and theories to which epistemology proper leads and those to which normative ethics leads. The discipline dealing with these epistemological questions and corresponding to the discipline of metaethics we may call metaepistemology. The central question of these disciplines, roughly, is how the statements of normative ethics and epistemology proper, respectively, can be supported, what is their "logic." The task of showing this is different, of course, from the task of producing a specific line of reasoning in support of specific ethical or epistemological principles. Can such principles be supported in the same way that propositions in the empirical and mathematical sciences can be? Whether they can be depends in part on what the meanings of the special epistemological or ethical terms are. Moral philosophers recognize three main views about the meaning of ethical terms and, correspondingly, about the ways in which ethical principles may be justified. Three very similar views have been given by epistemologists for the meaning and support of epistemological principles.

NONNATURALISM. The first view, which we may call nonnaturalism in epistemology and ethics alike, affirms two things. It affirms (1) that epistemic and ethical terms are meaningful and that epistemological and ethical statements are true or false and (2) that epistemic and ethical terms do not name observable qualities (such as color or shape) and that their meanings cannot be defined, even partially, by citing a relation between them and names of observable qualities. Epistemic and ethical terms can be explained only by way of other epistemic and ethical terms. Hence, neither epistemic nor ethical principles can be confirmed by observation in the way that principles in the empirical sciences can be. This

means that when we know ethical and epistemological statements that are not analytically true (as contrasted with ones like "A person ought to do his duty" or "One cannot know something unless it is reasonable for him to believe it"), our knowledge is synthetic a priori knowledge. A clear example of this view is the theory of probability held by J. M. Keynes, who thought that "probable" is an indefinable concept and that the axioms of probability theory are a priori synthetic knowledge.

NATURALISM. The second view can be called naturalism or "definism." It agrees with the first affirmation of nonnaturalism but denies the second. It holds that epistemic and ethical terms can be explained without the use of other epistemic or ethical terms, that they can be explained exclusively by use of empirical and logical concepts. As a result it holds that nonanalytic epistemological and ethical principles have the same logical status as the principles of the empirical sciences and can be appraised ultimately by reference to the data of observation or introspection.

For example, according to one such definition of an epistemic term, the statement "It is reasonable for *S* to believe *h*" means just "*S* believes *h*." A more plausible theory defines "know" as follows: "*S* knows that *h* at time *t*" means that *h* follows logically from the propositions *S* believes about his own experience at *t* (including what he seems to remember), plus the following (enumerated) principles of inductive logic and principles about the truth or probability of memory beliefs. Examples of parallel definitions in ethical naturalism are familiar. If the second definition given were accepted, it would be an analytic proposition that the principles of deductive and inductive logic are known by everyone, just as, given a utilitarian definition of "right," the principle of utilitarianism is analytic.

Parallel to the claim of the ethical relativist that conflicting basic ethical principles may be affirmed with equal warrant by different persons or social groups, is it possible that conflicting basic epistemological principles may also be affirmed with equal warrant by different persons, even if a naturalistic analysis of epistemological terms is adopted? For instance, just as the ethical relativist may affirm that different assessments of the ethical obligation of making a promise may be made with equal warrant by different persons, may someone claim with reason that different assessments, say, of the weight of an additional observation in support of some general law may be made with equal warrant by different persons? Such questions must be left unanswered here. It is obvious, of course, that given different evidence, it may be reasonable for a person *S* to believe some propositions which it would be unreasonable for person *S′* to believe.

NONCOGNITIVISM. The third view, which may be called noncognitivism, denies the thesis common to naturalism and nonnaturalism that epistemological and ethical statements are true or false but agrees with nonnaturalism that epistemic and ethical terms cannot be defined by means of empirical and logical concepts. Noncognitivism holds, however, that epistemic and ethical terms have a function and, in a sense, ideas in meaning. Ethical terms have been assigned various functions by different writers (functions like expressing the speaker's attitudes, changing the audience's attitudes, issuing prescriptions, declaring one's principles, giving advice, entreating, urging, exhorting, and so on).

Somewhat similar proposals have been made for epistemic terms. "It is probable that *h*," for example, is sometimes said to be a guarded way of affirming *h* or a cautious way of encouraging others to believe *h*. Again, "it is probable that *h*" is suggested not to assert that the speaker believes *h* but to express his belief in *h*. A more complex suggestion is to say that "it is reasonable to believe *h*" declares or expresses the speaker's own somewhat guarded inclination to believe and usually, at the same time, as a result of people's conditioning in the use of the language, strengthens the beliefs of others in *h*.

Further, parallel to ideas in C. L. Stevenson's ethical theory, one might say that epistemic terms also have some rather indefinite descriptive meaning, perhaps to the effect that acceptance of the proposition in question would conform to generally recognized standards—the standards, perhaps, of scientists. One could say, as P. H. Nowell-Smith does in his ethical theory, that epistemic terms have various functions in various contexts and that it is a mistake to look for some single function performed by them on every occasion. It could be added that the use of epistemic words (similar to the suggested parallel to Stevenson's theory) carries special contextual implications which distinguish them from nonepistemic terms. Since the noncognitivist does not think that epistemic and ethical statements are either true or false, his view does not contain any theory about how the truth of such statements may be established, although it often contains a descriptive account of various ways in which persons do or could try sensibly, in view of the kind of statement in question, to remove their disagreements. Defenders of the noncognitive view in epistemology include R. Chisholm, Stephen Toulmin, and J. N. Findlay. Although they have

been popular, noncognitive views, in epistemology and ethics alike, appear to face some difficulties. (For instance, in conditional clauses, such as "If it were known that …, then …" or "If it were a good thing to …, then …," the ethical and epistemological terms seem to be used in their normal sense, but obviously nothing is being urged or expressed. Again, whatever specific function one assigns to terms of either type, it seems possible to find affirmations which employ the terms and cannot plausibly be said to be performing that function.) Moreover, it is doubtful whether there are conclusive reasons for rejecting "definism" (naturalism) in all its possible forms.

SOME BROADER PERSPECTIVES

One feature of both ethical and epistemic terms is that even very educated people do not have any clear idea of their meaning or of how to support their applicability in, for example, the way they are able to support the assertion that someone is a bachelor. Nor is it clear how an appeal to whatever vague meaning there is could be used reasonably to require anyone, on pain of inconsistency, to accept a corresponding ethical or epistemological principle; the fact that a person rejected such a principle would always be good reason for saying that his use of the epistemic or ethical term does not correspond to the sense needed in order to require the principle. It would appear that one of the functions of moral philosophy and epistemology is, rather, to propose helpful and clarifying uses for these terms.

What would be a helpful and clarifying use for these terms must presumably be decided by a broad view of human nature and society, a view of action and the requirements of reliable prediction for the purpose of action, the need of informal social controls in a complex society, and the necessity for impartial and general rules for such control, given that human beings are for the most part intelligent, self-interested creatures. However this may be, it is likely that reflection on the functions of science (reasonable beliefs) and conscience in society may provide, in the case of science at least as well as in the case of ethics, a guide for the philosopher's reconstruction of the meaning of ethical and epistemic terms and a basis for the appraisal of epistemic and ethical principles.

It should be noted that metaethics and metaepistemology both have another part. That part is the explanation of various concepts necessary for the understanding of terms and statements of both epistemology and ethics and for an understanding of theories about how epistemological and ethical statements may be supported. Among the concepts that are epistemologically important are "meaning," "truth," "reference," "analytic," and "a priori." Some of these concepts are also important for ethics; other concepts important for ethics are "action," "choice," "free choice," "voluntary," "intention," "motive," and so on. Both metaethics and metaepistemology, then, have branches devoted to these auxiliary concepts.

If there is a parallel between ethics and epistemology, should we go on to say that there is a parallel between ethics and science since, after all, to assert any statement in science is at least to imply that there is evidence for what is stated, that it is generally known, and so on? Is not every scientific statement, at least covertly, an epistemological statement? It seems clarifying not to go this far. Epistemological statements are of the form "It is warranted for *S* to believe that …," and we should note that what comes after the "that" is the sort of statement that occurs in the textbooks of science. If the meaning of this statement also has to be construed as somehow epistemological, difficulties arise.

See also Analytic and Synthetic Statements; Chisholm, Roderick; Epistemology, History of; Ethical Relativism; Ethics, History of; Keynes, John Maynard; Metaethics; Naturalism; Noncognitivism; Stevenson, Charles L.

Bibliography

Adler, Jonathan E. *Belief's Own Ethics*. Cambridge, MA: Bradford Book MIT Press, 2002.

Alston, William. *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press, 1989.

Atkinson, R. F. "'Good' and 'Right' and 'Probable' in *Language, Truth and Logic*." *Mind* 64 (1955): 242–246.

Audi, Robert. *Belief, Justification, and Knowledge*. Belmont, CA: Wadsworth, 1988.

Austin, J. L. *Philosophical Papers*. Oxford: Clarendon Press, 1961. Pp. 67–71.

BonJour, Laurence, and Ernest Sosa. *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*. Malden, MA: Blackwell, 2003.

Chisholm, R. "Firth and the Ethics of Belief." *Philosophy and Phenomenological Research* 51 (1) (1991): 199–128.

Chisholm, R. *Perceiving*. Ithaca, NY: Cornell University Press, 1957.

Chisholm, R. "Evidence as Justification." *Journal of Philosophy* 68 (1961): 739–748.

DePaul, Michael. *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*. New York: Routledge, 1993.

DePaul, Michael, and Linda Zagzebski, eds. *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Clarendon Press, 2003.

Fairweather, Abrol, and Linda Zagzebski, eds. *Virtue Epistemology*. New York: Oxford University Press, 2001.

Feldman, Richard. "Epistemic Obligations." In *Philosophical Perspectives 2*, edited by J. E. Tomberlin. Atascadero, CA: Ridgeview Press, 1988.

Firth, R. "Chisholm and the Ethics of Belief." *Philosophical Review* 68 (4) (1959): 493–506.

Firth, R., and R. Chisholm. "Ultimate Evidence." *Journal of Philosophy* 53 (1956): 722–739. A symposium.

Greco, John. "Virtues and Vices in Virtue Epistemology." *Canadian Journal of Philosophy* 23 (1993): 413–432.

Lewis, C. I. *The Ground and Nature of the Right*. New York: Columbia University Press, 1955. Chs. 2–3.

Sinnot-Armstrong, Walter, and Mark Timmons, eds. *Moral Knowledge?* New York: Oxford University Press, 1996.

Sosa, Ernest. *Knowledge in Perspective*. Cambridge, U.K.: Cambridge University Press, 1991.

Steup, Matthias, ed. *Knowledge, Truth, and Duty*. New York: Oxford University Press, 2001.

Toulmin, Stephen. "Probability." In *Essays in Conceptual Analysis*, edited by A. G. N. Flew. London: Macmillan, 1956.

Urmson, J. O. "Some Questions concerning Validity." ibid.

Williams, Bernard. *Ethics and the Limits of Philosophy*. London: Fontana Press, 1985.

Zagzebski, Linda. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge, U.K.: Cambridge University Press, 1996.

Richard B. Brandt (1967)
Bibliography updated by Benjamin Fiedor (2005)

EQUALITY, MORAL AND SOCIAL

The proposition "*A* and *B* are equal" may be descriptive or normative, but in either case it is incomplete without a statement of the respects in which the objects or persons compared are deemed to be equal. In instances where this appears not to be so, either the context supplies the complement or the comparison is of pure quantities, as in pure mathematics. Two objects equal in weight, or height, or value may be unequal in other respects; apart from the abstractions of mathematics and logic, no two objects could ever be said to be equal in all respects, only in all relevant respects.

Correspondingly, to say that two candidates are equal in merit would usually mean that with respect to their performances in some understood competition or examination, they deserve to be treated alike; it does not rule out treating them differently if they are unequal in other respects. Aristotle's celebrated account of justice in Book III of the *Ethics* amounts to this: No distinction ought to be made between men who are equal *in all respects relevant to the kind of treatment in question*, even though in other (irrelevant) respects they may be unequal. On the other hand, in any matter in which they are in relevant respects unequal, they ought to be treated in proportion to their relevant inequalities.

These analytical distinctions are of considerable importance in dealing with equality as a moral and social ideal. Thomas Jefferson's claim that "all men are created equal" cannot be rebutted by pointing to the obvious fact that some are taller, stronger, or more clever than others. The claim is intelligible only as a prescription, as saying that there is some respect, at least, in which no difference ought to be made in the treatment or consideration given to all men, whatever differences there might be in their qualities and circumstances.

HISTORY OF EQUALITY AS AN IDEAL

Plato preached the political equality of the sexes, Aristotle that of all free citizens; nevertheless, both laid more stress on not treating unequals equally than on any general conception of equality. Aristotle believed that some men were slaves by nature, Plato that some souls were not merely capable of higher development than others but more valuable on that account. The political egalitarianism of Pericles' Athens, described by Thucydides, was concerned only with the equality of Athenian citizens and excluded slaves and foreigners. The first generalized egalitarianism was that of the Stoics, who stressed the natural equality of all men as rational beings with an equal capacity for virtue: "Virtue closes the door to no man; it is open to all . . . the freeborn, the freedman, the slave and the king . . . neither family nor fortune determines its choice—it is satisfied with the naked human being" (Seneca).

The New Testament doctrine of the equality of all souls in the sight of God (Galatians 3:26–29) is a religious expression of a similar principle: "Ye are all one in Christ Jesus." It was profoundly modified, however, by the Augustinian doctrine of election. Men were equal only in the sense that by sin all were totally, and therefore equally, unworthy; God in his mercy extended grace to some but not to others. Thomas Aquinas qualified the equality of all men in the sight of God by the doctrine that slavery is the consequence of sin. Though there are signs of a crude social egalitarianism in some of the protest movements of the later Middle Ages, such as the Lollards and the Hussites, medieval social theory was, on the whole, antiegalitarian, deeming hierarchy to be natural both to society and to the whole universal order.

Modern egalitarianism had its beginnings in the seventeenth century. It is related in part to Calvinist doctrine, which, although it admittedly drew a sharp distinction between the saved and the damned, insisted at the same time on the equality of the elect, whether clerical or lay. This view of equality came to be associated with

a theory of church government—and indirectly of secular government—that derived legitimate authority (i.e., the right of superiors to command inferiors) from the voluntary agreement of natural equals to submit to such of their number as they chose. These doctrines were given their first completely secular expression—associated with theories of natural right and social contract—by some of the Parliamentarians in the English civil war, particularly the Levelers. Colonel Rainborough's declaration in the General Council of the Army in 1647, that "the poorest he that is in England hath a life to live as the greatest he" and that no one can have a duty to obey a government that "he hath not had a voice to put himself under" is a classic expression of democratic political egalitarianism.

The idea of a natural equality of all men was a dominant theme from the seventeenth century on. Thomas Hobbes took it for granted that in the state of nature men are equal in right because roughly equal in strength and cunning. John Locke argued that by nature men are equally free, are subject only to natural law, and enjoy the same natural rights. This turns the problem of political authority and obligation into a search for reasons why free and equal men should accept the limitations of civil society. Political inequality, of ruler and ruled, must be justified as a conventional device for the better safeguarding of the rights and advantages that all men already possess but cannot securely enjoy, in a state of nature.

In eighteenth-century philosophy the idea of a natural equality of rights was reinforced by a theory of human nature, as put forth by Étienne Bonnot de Condillac and Claude-Adrien Helvétius, maintaining that all differences of character, talent, and intelligence are due to differences in environment and experience. By nature men are equal in the sense that at birth they have a limitless potentiality without natural characteristics to differentiate one from another. Consequently their diverse natures are, in fact, contingent; in principle all men are equally perfectible, given the appropriate social arrangements.

Jean-Jacques Rousseau explained social inequality by the pressures of a sophisticated way of life; in the state of nature men's needs are simple, none need rely on anyone but himself, so none can exploit another or make him subject. For Rousseau the key problem for social philosophy, to which the sovereignty of the general will could provide a solution, was to reconcile the natural equality and autonomy of men with the social condition and political authority. Without this reconciliation men cannot realize their potentiality as morally self-governing persons. Immanuel Kant offered a philosophically sophisticated version of a very similar moral position: All human beings must be treated as ends, not merely as means; all men are equally "legislating members of the kingdom of ends," because all are equally capable of realizing the good will, the only thing in the world good in itself.

These doctrines permeated the great revolutionary movements in America and Europe at the end of the century and were made explicit in their declarations of rights. In America the doctrine of equality was a denial that any authority imposed on unwilling subjects could be legitimate merely on grounds of law or prescription; in France *l'égalité* repudiated privileges of prestige and opportunity based solely on noble birth. But alongside these broad popular movements were others, such as François-Noël Babeuf's Conspiracy of the Equals, which challenged economic inequalities. Protests of this kind became increasingly important in the nineteenth century, in the evolution of socialist and communist thinking.

The target of modern egalitarianism, however, is by no means solely, or even primarily, economic inequality. Such inequality is objectionable to many socialists not so much as an unequal distribution of goods but as a source of unequal power, prestige, and regard. Other forms of differentiation have been as strongly attacked—in particular, differentiation by race and color and by sex. Again, egalitarians may make very general claims, such as that in the Universal Declaration of Human Rights (adopted by the United Nations General Assembly in 1948), that "all human beings are born free and equal in dignity and rights," or they may claim, more specifically, "equality before the law" or "equality of opportunity."

This history has two noteworthy features. First, there is a recurrent theme, the idea of a universal but imprecisely defined equality; behind all differences of talents, merits, and social advantages there is some characteristically human nature by virtue of which all men are equal. Second, the focus of egalitarianism has shifted continuously, now attacking the differential treatment of barbarian and Greek, now of freeman and slave, noble and commoner, black and white, rich and poor, male and female. Egalitarianism might be said not so much to assert equality as to deny the justice of some existing inequality of treatment based on some allegedly irrelevant differences of quality or circumstance.

UNIVERSAL EQUALITY AS AN IDEAL

The notion of universal equality as an ideal is difficult to pin down. Many egalitarians have tried to argue that despite the many points of inequality, all men are alike in possessing reason, or a soul, or some other essentially

human characteristic or nature, by virtue of which they stand equal. The difficulty, however, is to find an important characteristic that all men possess in *precisely the same degree*, so that whatever differences their other inequalities might justify, this fundamental equality would make them equal *qua* men. And even if one could identify such a characteristic, what would follow from it? If all men are alike in having souls, in what respect should they therefore be treated alike? After all, God is widely believed to punish wicked souls and to reward virtuous ones.

The ideal of universal equality can often be reduced to the principle that all men ought to be equally considered. This does not mean that there is any respect in which they are all alike and by virtue of which they should all be treated alike; it is rather a principle of procedure: that all men ought to be treated equally, despite all their differences, until a case has been made for saying that some particular difference between them is relevant to the matter at hand. The onus of proof rests on whoever wants to make distinctions. And up to a point this might be said to be implicit in the notion of rational decision, because it would be irrational, within a given class of cases, to treat some differently from others if no relevant grounds could be found for distinguishing between them.

Nevertheless, the principle of equal consideration does presuppose an initial commitment or decision, for it takes for granted whose interests are to count. No one claims equal consideration for all mammals—human beings count, mice do not, though it would not be easy to say why not. The Greeks made a similar distinction between themselves and barbarians, Aristotle between natural slaves and naturally free men, the slaves counting only as tools for the free men. It is not easy to see how anyone who seriously held that white men mattered but black did not could be reasoned out of this position, any more than one could argue for the equality of men and mice. Of course, many people who practice discrimination profess to believe that black men are in some way inferior to white, in intelligence, sensibility, responsibility, or some such quality, and on this account ought to be treated differently. But this is to admit the principle of equal consideration for all men, that all men count, and that an argument has to be made to justify discriminating against some among them. The man who denies that they count at all is not bound to show reasons, any more than we feel the need to show reasons for treating inanimate objects, plants, or primitive organisms, such as amoebas, according to our pleasure. Although we hesitate to inflict unnecessary pain on sentient creatures, such as mice or dogs, we are quite sure that we do not need to show good reasons for putting human interests before theirs. The boundaries of moral consideration are enlarged in practice by awakening sympathy and imagination; moral reasons presuppose an initial moral concern.

The principle of equal consideration may be more, therefore, than what is necessarily implied by the concept of rational action. The notion of acting with good reason does not in itself rule out any inequality of treatment, for it may always be possible to argue that there is some relevant difference between members of a given class. But the principle that all men should be treated as members of the class whose equality is procedurally presupposed is not necessarily implied by the notion of rational action.

However, to some philosophers universal equality has meant more than equal consideration for all men. John Plamenatz, for instance, has tied the notion closely to natural rights and has argued that there are some rights "so much more important than others that these others ought always, or nearly always, to be sacrificed to them, should the need arise" ("Equality of Opportunity," in Bryson et al., eds., *Aspects of Human Equality*, pp. 79–107). The purpose of this equality of rights is to ensure equality of freedom and opportunity: "the equal right of all men to live the kind of life that seems good to them … equality of opportunity to be oneself, to live as one pleases." This is attractive, but it hardly touches the problem of what is to be done when what pleases one man interferes or competes with what pleases another. Nor does it cope with the diversity of inclinations—can one be said to have, on a given income, an equal opportunity to become a collector of Picassos or of seashells? Or does equality of opportunity require differential provisions, so that the chance of fulfillment matches the aspiration? Does it envisage open competition or a handicap? Plamenatz has attached very great weight to the principle that every individual's view of where his own interests lie should be given equal consideration. He thus closely associates equality and freedom, denying both that one man's interest might legitimately be subordinated to another's and that anyone can be the proper judge of someone else's interest.

For some philosophers (D. D. Raphael and Gregory Vlastos, for example) the ideal of universal equality requires that the inequalities of nature be mitigated or rectified. By this view, precisely because men born with superior talents or social advantages can claim no merit on that account, it should be the aim of social policy to compensate for such advantages by differentiating between men to redress the balance. It is of course true

that modern welfare states commonly do provide special amenities, such as wheelchairs for the crippled or hearing aids for the deaf, to bring naturally handicapped people up to some minimum standard of well-being. But an account in terms of meeting needs or deficiencies is more accurate than one in terms of rectifying inequalities, for the policy is not so much to remedy a handicap that one man suffers in comparison with another (wheelchairs are not meant to enable handicapped persons to compete in races with runners) as to provide conditions necessary to his well-being, understood in the light of some presupposed standard of what a good life requires. This standard will no doubt be governed by the advantages commonly enjoyed by most people in the community, so that in an affluent society a person will be taken to have more needs than in an impoverished one; however, the claim will still be grounded on his own needs and interests, not on the greater advantages enjoyed by those more fortunate.

SPECIFIC EGALITARIAN IDEALS

The demand for equality is very often directed against some specific inequalities in social arrangements. It may take the form of a protest either against distinctions based on some specific ground (for example, racial equality, sexual equality) or against discriminations in a particular field (for example, equality before the law, economic equality). Each consideration necessarily involves the other; complaints of sexual inequality imply that sex is made a ground of distinction in some fields, unspecified but understood, where it is considered by the critic to be inappropriate (for example, salaries, jobs in the public service, voting rights). On the other hand, the claim to equality before the law implies that in legal relations or in relations between persons appearing before a court, some unspecified but understood difference (perhaps of sex, or of color, or of wealth) is made a ground of distinction and ought not to be.

The meaning of these ideals changes with their context. No one means by "equality before the law" that no distinctions should be legally recognized. A social system consists necessarily of different roles, such as father, son, tenant, landlord, and congressperson, each with its own appropriate qualifications and characteristic rights and duties, established and supported by law. A system is said to be unequal only if the differences in privileges are considered unjustifiable because they are irrelevantly grounded or because the qualifications for assuming a role are unduly restrictive (for instance, if a white skin is a necessary condition for voting rights). These ideals change their focus over time. "Equality before the law" in eighteenth-century France meant ending the disabilities of the members of the third estate as compared with the privileges of the nobles and clergy. Today it may mean abolishing racial disabilities, such as existed in South African law and with Jim Crow sanctions in the United States, or seeing that prejudice does not interfere with the administration of law. It may also mean eliminating the advantages of wealthy litigants over poorer ones, by public legal aid schemes, or making certain that no one is prevented by poverty from getting a fair trial (see Justice Hugo Black's opinion in *Griffin v. Illinois*, 351 U.S. 12, 1956).

Equality very rarely means treating everyone alike; usually it means getting rid of one system of distinctions and replacing it with another. Thus, equality of opportunity in education hardly ever means giving everyone exactly the same education; rather, it means eliminating some hitherto determining factor such as ability to pay school or university fees and substituting a test of proficiency. More ambitiously, it might aim at a system with various arrangements, each meant for an appropriate grade of intelligence or type of aptitude. Those who call this equality do so on the ground that the treatment accorded to each is equally appropriate to his needs. Thus, R. H. Tawney argued that "the more anxiously a society endeavours to secure equality of consideration for all its members, the greater will be the differentiation of treatment which, when once the common human needs have been met, it accords to the special needs of different groups and individuals among them" (*Equality*, p. 39). The greater the equality of consideration, the greater the differentiation in treatment. If the latter is not called "inequality in treatment" it is because the word *inequality* has acquired, in this sort of context, a pejorative force; "inequalities" have come to mean *indefensible* differences in treatment.

See also Impartiality; Justice.

Bibliography

HISTORY OF EGALITARIAN IDEALS

Abernethy, George L., ed. *The Idea of Equality*. Richmond, VA: John Knox Press, 1959. An anthology, from the Old Testament and Herodotus to the present. Contains bibliography.

Bouglé, Celestin. *Les idées égalitaires*. Paris: Alcan, 1899.

Lakoff, Sanford A. *Equality in Political Philosophy*. Cambridge, MA: Harvard University Press, 1964.

Locke, John. "Second Treatise on Civil Government" (1690). In *Two Treatises of Government*, edited by P. Laslett. Cambridge, U.K.: Cambridge University Press, 1960.

Rousseau, Jean-Jacques. *Discours sur l'origine et les fondements de l'inégalité parmi les hommes* (Discourse on the Origin of Inequality). Amsterdam, 1755. Various modern editions have been published.

Stephen, Sir James Fitzjames. *Liberty, Equality, Fraternity*. London: Smith Elder, 1873. A classic nineteenth-century criticism of radical and egalitarian ideals.

Thomson, David. *Equality*. Cambridge, U.K.: Cambridge University Press, 1949. Surveys the principal modern egalitarian ideals against a historical background—a useful, brief general introduction.

MORE-RECENT DISCUSSIONS

Benn, S. I., and R. S. Peters. *Social Principles and the Democratic State*. London: Allen and Unwin, 1959. Reissued as *Principles of Political Thought*. New York: Collier, 1964. See especially Chs. 5–7.

Brandt, Richard B., ed. *Social Justice*. Englewood Cliffs, NJ: Prentice-Hall, 1962. Essays by K. E. Boulding, Paul A. Freund, William K. Frankena, Alan Gewirth, Gregory Vlastos.

Bryson, L. et al., eds. *Aspects of Human Equality*. Fifteenth symposium of the Conference on Science, Philosophy, and Religion. New York: Harper, 1956. Nineteen papers on various aspects. See especially J. Plamenatz, "Equality of Opportunity."

Carritt, E. F. *Ethical and Political Thinking*. Oxford: Clarendon Press, 1947.

Hobhouse, L. T. *Elements of Social Justice*. London: Allen and Unwin, 1922.

Margolis, J. "'That All Men Are Created Equal.'" *Journal of Philosophy* 52 (13) (1955): 337–346. Argues a "fundamental and actual equality of all men."

Raphael, D. D. "Justice and Liberty." *PAS* 51 (1950–1951): 167–196.

Spiegelberg, Herbert. "A Defence of Human Equality." *Philosophical Review* 53 (2) (1944): 101–124. Interprets the concept in terms of compensation.

Williams, Bernard. "The Idea of Equality." In *Philosophy, Politics and Society*, edited by Peter Laslett and W. G. Runciman. 2nd series. Oxford: Blackwell, 1962.

Wollheim, R., and Isaiah Berlin. Symposium on "Equality." *PAS* 56 (1955–1956): 281. Reprinted in *Justice and Social Policy*, edited by F. A. Olafson. Englewood Cliffs, NJ: Prentice-Hall, 1961.

ECONOMIC AND SOCIAL EQUALITY

Dahrendorf, R. "On the Origin of Social Inequality." In *Philosophy, Politics and Society*, edited by Peter Laslett and W. G. Runciman. 2nd series. Oxford: Blackwell, 1962. A sociological approach.

Harris, R. J. *The Quest for Equality*. Baton Rouge: Louisiana State University Press 1960. "Equality" in American constitutional history and law.

Jouvenel, Bertrand de. *Ethics of Redistribution*. Cambridge, U.K.: Cambridge University Press, 1951. A critical analysis of egalitarianism.

Lampman, R. J. "Recent Thought on Egalitarianism." *Quarterly Journal of Economics* 71 (2) (1957): 234–266. Surveys economic doctrines.

Myers, H. A. *Are Men Equal?* Ithaca, NY: Great Seal, 1955. The American egalitarian tradition.

Tawney, R. H. *Equality*. Rev. ed. London: Allen and Unwin, 1952.

ADDITIONAL SOURCES

Anderson, Elizabeth. "What Is the Point of Equality?" *Ethics* 109 (2) (1999): 287–337.

Clayton, Matthew, and Andrew Williams, eds. *The Ideal of Equality*. New York: Palgrave Macmillan, 2000. Includes texts by the editors: John Rawls, T. M. Scanlon ("The Diversity of Objections to Inequality"), Tom Nagel, Derek Parfit ("Equality or Priority"), Larry Temkin, G. A. Cohen, Richard Arneson, and Ronald Dworkin. Essential contemporary texts.

Cohen, G. A. "On the Currency of Egalitarian Justice." *Ethics* 99 (4) (1989): 906–944.

Cohen, G. A. *If You're an Egalitarian, How Come You're So Rich?* Cambridge, MA: Harvard University Press, 2000.

Cohen, Joshua. "Democratic Equality." *Ethics* 99 (4) (1989): 727–751.

Cohen, Marshall, et al., eds. *Equality and Preferential Treatment: A Philosophy and Public Affairs Reader*. Princeton, NJ: Princeton University Press, 1977. Incluces essays by Tom Nagel, J. J. Thomson, Robert Simon, George Sher, Ronald Dworkin, Owen Fiss, and Alan Goldman.

Darwall, Steve, ed. *Equal Freedom*. Ann Arbor: University of Michigan Press, 1995. Includes essays by T. M. Scanlon, John Rawls ("The Basic Liberties and Their Priority"), Ronald Dworkin, Amartya Sen ("Equality of What?"), and G. A. Cohen.

Dworkin, Ronald. *Sovereign Virtue*. Cambridge, MA: Harvard University Press, 2000.

Dworkin, Ronald, et al. "Symposium on *Sovereign Virtue*." *Ethics* 113 (1) (2002): 5–143. Contains essays by Matthew Clayton, Andrew Williams, Michael Otsuka, Robert van der Veen, Marc Fleurbaey, and Ronald Dworkin.

Hurley, Susan. *Justice, Luck, and Knowledge*. Cambridge, MA: Harvard University Press, 2003.

Johnston, David, ed. *Equality*. Indianapolis: Hackett Publishing, 2000. With texts by Plato, Aristotle, the Levellers, Hobbes, Rousseau, Burke, de Tocqueville, Marx, Tawney, von Hayek, Rawls, Nozick, Sen ("Equality of What?"), Dworkin, Walzer ("Complex Equality"), Kymlicka, and Iris Young.

Kagan, Shelly. "Equality and Desert." In *What Do We Deserve?*, edited by Louis Pojman and Owen McLeod. Oxford: Oxford University Press, 1999.

Kekes, John. *The Illusions of Egalitarianism*. Cornell: Cornell University Press, 2003.

Koppelman, Andrew. *Antidiscrimination Law and Social Equality*. New Haven, CT: Yale University Press, 1996.

Nagel, Thomas. *Equality and Partiality*. Oxford: Oxford University Press, 1995.

Okin, Susan Moller. *Justice, Gender, and the Family*. New York: Basic Books, 1989.

Otsuka, Michael. *Libertarianism without Inequality*. Oxford: Oxford University Press, 2004.

Persson, Ingmar. "A Basis for Interspecies Equality." In *The Great Ape Project: Equality Beyond Humanity*, edited by Paola Cavalieri and Peter Singer. New York: St. Martin's Griffin, 1996.

Pojman, Louis P., and Robert Westmoreland, eds. *Equality*. Oxford: Oxford University Press, 1996. Includes texts by Aristotle, Hobbes, Rousseau, Hume, Babeuf and Marechal, Oppenheim, McKerlie, Temkin, Kant, Nozick, Lucas, Benn, Vlastos, Schaar, Fishkin, Westen, Galston, Rawls, Matson, Nielson, Hare, Arneson ("Equality and Equal Opportunity for Welfare"), Rakowski, Nagel, Frankfurt ("Equality as a Moral Ideal"), Pojman, Walzer ("Complex Equality"), and Vonnegut.

Rakowski, Eric. *Equal Justice*. Oxford: Oxford University Press, 1993.

Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge MA: Harvard University Press, 1999 [1971].

Raz, Joseph. "Equality." Chapter 9 in his *The Morality of Freedom*. Oxford: Oxford University Press, 1986.

Roemer, John. *Equality of Opportunity*. Cambridge, MA: Harvard University, 1998.

Scheffler, Samuel. "What Is Egalitarianism?" *Philosophy and Public Affairs* 31 (1) (2003): 5–41.

Sen, Amartya. *Inequality Reexamined*. Oxford: Oxford University Press, 1992.

Shiffrin, Seana Valentine. "Egalitarianism, Choice-Sensitivity, and Accommodation." In *Reasons and Values: Themes from the Work of Joseph Raz*, edited by Philip Pettit, Samuel Scheffler, Michael Smith, and R. Jay Wallace. Oxford: Oxford University Press, 2004.

Steiner, Hillel. "Capitalism, Justice, and Equal Starts." *Social Philosophy and Policy* 5 (1987): 49–71.

Tawney, R. H. *Equality*. Rev. ed. London: Rowman & Littlefield, 1952 [1931].

Temkin, Larry. *Inequality*. Oxford: Oxford University Press, 1993.

Vallentyne, Peter. "Equality, Efficiency and Priority for the Worse Off." *Economics and Philosophy* 16 (2000): 1–19.

Van Parijs, Philippe. *Real Freedom for All: What (If Anything) Can Justify Capitalism?* Oxford: Oxford University Press, 1995.

Waldron, Jeremy. *God, Locke, and Equality: Christian Foundations in Locke's Political Thought*. Cambridge, U.K.: Cambridge University Press, 2002.

Wilkinson, T. M. "Raz on Equality." *Imprints* 3 (2) (1998): 132–155.

Stanley I. Benn (1967)
Bibliography updated by Thomas Pogge (2005)

EQUALITY, MORAL AND SOCIAL [ADDENDUM]

Equality is a potent ideal that plays a major role in a wide-range of social, political, and moral debates. Unfortunately, equality defies easy characterization, and few ideals of such significance have been so poorly understood.

EQUALITY OF WHAT?

Much debate concerns what kind of equality is desirable: income, resources, primary goods, power, welfare, opportunity, needs satisfaction, capabilities, functionings, rights, or liberties. Should the chief concern be legal, social, or political equality? These are extremely important questions, as equality of one kind fosters inequality of another. Although many assume that we should only be concerned with one kind of equality, it is arguable that various kinds of equality matter, perhaps to various degrees in different contexts.

VARIOUS KINDS OF EGALITARIANISM

Philosophers have long distinguished between formal and substantive principles of equality. It is perhaps more useful to distinguish between equality as universality, as impartiality, or as comparability. A basic principle of rationality, equality as universality reflects the view that all reasons and principles must be universal in their application. Because it applies universally, even the view that all blondes should be rich and all brunettes paupers meets this egalitarian principle.

Equality as impartiality holds that all people must be treated with disinterested fairness. Of course, positions vary dramatically regarding what constitutes treating people impartially. For example, Kantians regard impartiality as treating people as ends in themselves and never merely as means, whereas for Utilitarians it requires neutrality concerning different people's interests when maximizing the good.

Although all plausible moral theories are committed to equality as universality and impartiality, equality as comparability reflects a deeper commitment to equality. Equality as comparability is concerned with how people fare relative to others. This is a distinctive substantive view that rivals nonegalitarian positions such as Utilitarianism and libertarianism.

Another important distinction is between instrumental egalitarianism, according to which equality is valuable only insofar as it promotes some other valuable ideal, and noninstrumental egalitarianism, which holds that equality is sometimes valuable in itself, beyond the extent to which it promotes other ideals. On noninstrumental egalitarianism, any complete account of the moral realm must allow for equality's value.

Many who favor significant redistribution from the wealthy to the poor are instrumental egalitarians; they favor such redistribution only as a way of reducing suffering, aiding the worst off, fostering solidarity, or

strengthening democratic institutions. Such reasons are morally significant and compatible with equality as universality and impartiality. But each is also compatible with the rejection of noninstrumental egalitarianism and equality as comparability.

Further distinctions: in person-affecting versions of egalitarianism, inequality only matters insofar as it adversely affects people; in impersonal versions, inequality can matter even when it does not adversely affect people. Similarly, deontic egalitarianism focuses on duties to address the legitimate complaints of victims of inequality by improving their situations, whereas telic egalitarianism focuses on removing objectionable inequalities as a means of improving the goodness of outcomes. Deontic egalitarianism focuses on assessing agents or actions in order to minimize the consequences of unavoidable inequalities for which no one was responsible, whereas telic-egalitarianism focuses on the goodness of outcomes in such a way that such inequalities may matter.

UNDERSTANDING EQUALITY

The notion of equality is widely assumed to be:

> holistic—concerned about (in)equality between groups or societies, blacks and whites, women and men, Ethiopians and Swiss, and so on. The aim is to address the factors accounting for objectionable inequalities between different groups or societies;
>
> simple—we all know what equality is, that where everybody has the same amount of x, for whatever x we are interested in; and
>
> essentially distributive—concerned with how certain acts or goods are distributed among various groups; *ceteris paribus,* an equal distribution is best.

The conventional assumptions are questionable. Arguably, the notion of equality is:

> individualistic—groups and societies are not the proper objects of moral concern, individuals in groups and societies are. For example, though on average whites may be richer than blacks, if inequality of wealth matters, then it matters between rich and poor blacks, as well as rich blacks and poor whites;
>
> complex—in judging outcomes regarding inequality, many considerations seem relevant, including how much deviation there is from a state of "pure" equality, how "gratuitous" the inequality seems, or the extent to which individuals have a "complaint" regarding equality. Moreover, the size of an individual's egalitarian complaint may depend on how he or she fares relative to the average person, the best off person, or all those better off than he or she; and, in addition, one might add individual complaints, focus on the worst-off's complaints, or add everyone's complaints, but give special weight to larger complaints. On reflection, there are many distinct "aspects" of equality that underlie and influence egalitarian judgments; and
>
> essentially comparative—equality is a relation that obtains between individuals, and the concern is for how individual's fare relative to each other.

LUCK EGALITARIANISM AND RESPONSIBILITY

Luck egalitarianism aims to rectify luck's influence in people's lives. Acknowledging the importance of autonomy and personal responsibility, luck egalitarianism holds that it is bad when one person is worse off than another through no fault or choice of his or her own. So luck egalitarians object when equally deserving people are unequally well off but not when one person is worse off than another due to his or her own responsible choices—perhaps to pursue a life of leisure or crime.

Some luck egalitarians distinguish between option luck—luck to which we responsibly open ourselves—and brute luck, luck that simply befalls us, unbidden. On this distinction, any option-luck inequalities such as those that result from people autonomously choosing to gamble or invest in the stock market are unobjectionable. By contrast, brute luck inequalities—such as those that result from some being born with less intelligence or to poorer parents or being struck down by lightning, cancer, or an accident—are objectionable.

Against luck egalitarianism, some claim that egalitarians should aid the worse off—for example lung-cancer victims or low-income earners—even if they were partly responsible for their predicament—say, because they smoked or dropped out of school. Against the option/brute luck distinction, some contend that drawing the line between them is difficult and that it is objectionable if Mary takes a prudent risk and John an imprudent one yet Mary fares much worse than John, because option luck frowns on Mary but smiles on John.

EQUALITY AND FAIRNESS

If I give one piece of candy to Andrea and two to Rebecca, Andrea might immediately protest, "Unfair!" This natural reaction suggests an intimate connection between equal-

ity and fairness. On one view, concern about equality is a matter of comparative fairness that focuses on how people fare relative to others. Specifically, concern about equality reflects the view that inequality is bad when and because it is unfair, where the unfairness consists in one person being worse off than another no more deserving.

The intimate connection between equality and fairness illuminates the relevance and limitations of luck egalitarianism's "through-no-fault-or-choice-of-their own" clause. On this view, among equally deserving people inequality is bad because it is unfair for some to be worse off than others through no fault or choice of their own. But among unequally deserving people inequality is not bad or unfair for someone less deserving to be worse off than someone more deserving even if the former is worse off through no fault or choice of his own. For example, egalitarians need not object if criminal John is worse off than law-abiding Mary, even if John craftily avoided capture and so is only worse off because, through no fault or choice of his own, a falling brick injured him.

Additionally, in some cases inequality is bad because unfair even though the worse off are responsible for their plight, as when people are worse off because they chose to do their duty or acted beyond the call of duty, in adverse circumstances not of their making. So, on reflection, luck itself is neither good nor bad from the egalitarian standpoint. Egalitarians object to luck that leaves equally deserving people unequally well off. But they can accept luck that makes equally deserving people equally well off or unequally deserving people unequally well off proportional to their deserts. Thus, luck will be approved or opposed only to the extent that it promotes or undermines comparative fairness.

ARGUMENTS AGAINST EQUALITY AND RESPONSES

Many arguments have been offered against equality. Some contend that a world in which everyone would be the same would be undesirable. On some accounts, equality conflicts with freedom because even if one had a perfectly equal outcome, one could only preserve such an outcome by preventing people from voluntarily engaging in beneficial exchanges. Some argue that equality requires that we level down the better off if we cannot benefit the worse off. For example, we might have to blind the sighted, handicap the athletically gifted or disfigure the beautiful even if no one benefited from such actions. Thus, many believe that we should accept prioritarianism, and give priority to people the worse off they are in absolute terms, rather than egalitarianism, which focuses on people's relative positions.

Some egalitarians soundly reject the radical egalitarian position that people should be equal in all respects. Only some inequalities are normatively significant, they argue, and they are compatible with vast inequalities in other respects. Regarding freedom, egalitarians may argue that genuine freedom involves the autonomous formulation and effective implementation of a meaningful life plan commensurate with one's capacities, a prospect that is incompatible with the levels of inequality prevalent throughout much of the world. Moreover, freedom is not all that matters; fairness does, too, so some tradeoffs may be necessary between freedom and equality when they conflict.

The leveling-down objection fails against person-affecting and deontic egalitarianism. Moreover, although it applies to telic versions of equality as comparability, it also applies to other impersonal moral principles to which many are committed, like proportional justice. Egalitarians can admit that worsening conditions for the better off might be bad but that this does not show that inequality doesn't matter, merely that it isn't all that matters. Equality is not the only ideal that would have terrible implications if exclusively pursued; the same is true of justice, utility, freedom, and perfection. The main lesson of the leveling-down objection is that we must be pluralists, a point readily granted by egalitarians.

REMAINING QUESTIONS

Many questions have not been addressed. For example, does inequality matter more at high or low levels? Is it affected by variation in population size? Does it matter between societies as much as within societies? Or between different species? Should egalitarians compare whole lives, simultaneous segments of lives (say, today's elderly with today's youth), or corresponding segments of lives (today's elderly with tomorrow's elderly)? Should egalitarians be neutral regarding space and time, or does inequality's significance depend on whether societies are connected? These questions are all important, and their answers may significantly bear on our understanding of morality and equality. It is no accident that appeals to equality are ubiquitous. Equality remains a powerful human ideal.

See also Libertarianism; Utilitarianism.

Bibliography

Anderson, Elizabeth. "What Is the Point of Equality?" *Ethics* 109 (1999): 287–337.

Arneson, Richard. "Equality and Equal Opportunity for Welfare." *Philosophical Studies* 56 (1989): 77–93.

Arneson, Richard. "Welfare Should Be the Currency of Justice." *Canadian Journal of Philosophy* (2000): 497–524.

Atkinson, A. B. *The Economics of Inequality*. Oxford: Clarendon Press, 1975.

Cohen, G. A. "On the Currency of Egalitarian Justice." *Ethics* 99 (1989): 906–944.

Crisp, Roger. "Equality, Priority, and Compassion." *Ethics* 113 (2003): 745–763.

Dworkin, Ronald. *Sovereign Virtue*. Cambridge, MA: Harvard University Press, 2000.

Dworkin, Ronald. "What is Equality? Part 1: Equality of Welfare," and "Part 2: Equality of Resources." *Philosophy and Public Affairs* 10 (1981): 185–246, 283–345.

Frankfurt, Harry. "Equality as a Moral Ideal." *Ethics* 98 (1987): 21–43.

Hurley, Susan. "Luck and Equality." *Supplement to the Proceedings of The Aristotelian Society* 75 (2001): 51–72.

Kolm, S. Ch. "The Optimal Production of Social Justice" In *Public Economics*, edited by J. Margolis and H Guitton, pp. 145–200. New York: Macmillan, 1969.

Kolm, S. Ch. "Unequal Inequalities, I and II." *Journal of Economic Theory* 12 and 13 (1976): 416–42, 82–111.

Lippert-Rasmussen, Kasper. "Egalitarianism, Option Luck, and Responsibility." *Ethics* 111 (2001): 548–579.

Lucas, J. R. "Against Equality." *Philosophy* 40 (1965): 296–307.

Lucas, J. R. "Against Equality Again." *Philosophy* 52 (1977): 255–280.

McKerlie, Dennis. "Equality and Time." *Ethics* 99 (1989): 475–91.

Nagel, Thomas. *Equality and Partiality*. New York: Oxford University Press, 1991.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Nussbaum, Martha, and A. K. Send, eds. *The Quality of Life*. Oxford: Clarendon Press. 1992.

Parfit, Derek. "Equality or Priority?" In *The Ideal of Equality*, edited by Matthew Clayton and Andrew Williams, pp. 81–126, New York: St. Martin's, 2000.

Rakowski, Eric. *Equal Justice*. Oxford: Oxford University Press, 1993.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Roemer, John. *Equality of Opportunity*. Cambridge, MA: Harvard University Press, 1998.

Schaar, John H. "Equality of Opportunity and Beyond." In *Nomos IX: Equality*, edited by R. Pennock and J. Chapman, pp. 228–249. New York: Atherton Press, 1967.

Scheffler, Samuel. "What Is Egalitarianism?" *Philosophy and Public Affairs* 31 (2003): 5–39.

Sen, A.K. *Inequality Reexamined*. Cambridge, MA: Harvard University Press, 1992.

Sen, A. K., with James Foster. *On Economic Inequality*. Oxford: Clarendon Press, 1997.

Temkin, Larry S. "Egalitarianism Defended." *Ethics* 113 (2003): 764–782.

Temkin, Larry S. *Inequality*. Oxford: Oxford University Press, 1993.

Temkin, Larry S. "Inequality: A Complex, Individualistic, and Comparative Notion." In *Philosophical Issues* 11, edited by Ernie Sosa and Enriquea Villanueva, 327–352. Blackwell, 2001.

Temkin, Larry S. "Justice and Equality: Some Questions about Scope." In *Social Philosophy and Policy*. Vol. 12, edited by Paul Miller, 72–104. New York: Cambridge University Press, 1995.

Tungodden, Bertil. The Value of Equality. *Economics and Philosophy* 19 (2003): 1–44.

Vallentyne, Peter. "Brute Luck, Option Luck, and Equality of Initial Opportunities." *Ethics* 112 (2002): 529–557.

Vonnegut, Kurt, Jr. "Harrison Bergeron." In *Welcome to the Monkey House*. New York: Dell, 1998.

Williams, Bernard. "The Idea of Equality." In *Philosophy, Politics, and Society*, Second Series, edited by Peter Laslett and W.G. Runciman, 110–131. Basil Blackwell, 1962.

Larry Temkin (2005)

ERASMUS, DESIDERIUS

(1466?–1536)

Desiderius Erasmus, the great Renaissance humanist and scholar, was born at either Rotterdam or Gouda in Holland, the illegitimate son of a priest. As a child he studied at Gouda, and from 1475 to 1483 he studied at Deventer with the Brethren of the Common Life, a pious, modernist-humanist order. Next, he studied at Hertogenbosch, became an Augustinian friar at St. Gregory's (near Gouda), and, in 1492, was ordained a priest. Disliking monastic life, in 1494 he became the Latin secretary to the bishop of Cambrai. The next year he went to the University of Paris to study theology, but he found both the life and the scholastic philosophy distasteful. In 1499 he went to England, where he became a close friend of the humanists John Colet and Thomas More and devoted himself to the study of the classics and sacred literature, desiring to combine the new humanistic spirit, based on the revival of interest in the classics, with Christian learning. In 1500 he returned to the Continent and devoted himself to the study of Greek. One of his first famous works was published in this period, the *Enchiridion Militis Christiani* (Handbook of a Christian soldier; 1501), an appeal for a return to the simple spirit of early Christianity.

In Belgium, in 1504, Erasmus came across a manuscript of Lorenzo Valla's *Annotationes* on the New Testament, in which Valla criticized the Vulgate (Latin) version of the Bible and set forth a critical method for arriving at a correct text of scripture. Erasmus was tremendously

impressed and published an edition of Valla's work in 1505, after which he returned to England and copied the Greek New Testament from the manuscripts available to him there. He then went to Italy as a tutor to the sons of Henry VIII's doctor and took his doctorate of divinity at Turin in 1506. He lived in various Italian cities for the next three years and began publishing the famous edition of his *Adagia*, a collection of 3,000 proverbs from classical writers, at Venice in 1508. As a result of this work, he was soon recognized as the foremost scholar of northern Europe. In 1509 he returned to England and stayed with Thomas More. There, he wrote the *Moriae Encomium* (In praise of folly), a witty satire on worldly learning and activities and a presentation of simple, pious, nontheological Christianity. While in England he lectured at Cambridge on Greek and on St. Jerome's [c. 347–419] epistles. In 1514 he went to Basel, Switzerland, to assist the publisher Johann Froben (c.1460–1527) in preparing an edition of his works. While there he received a summons to return to monastic life, which he resisted strongly, and finally Pope Leo X (1475–1521) granted him a dispensation allowing him to live in the world.

In 1516 he published one of his most influential works, the Greek edition of the New Testament. Comparing various manuscripts and citations from the church fathers, he presented a more accurate text than the Vulgate, along with his own elegant Latin version and many learned and critical notes. This edition became a model and inspiration for the new learning and for critical scholarship. Theologically, its omission of an interpolated passage in I John 5:7–8, stating the doctrine of the Trinity, greatly influenced liberal reformers like Michael Servetus, and its emphasis on St. Paul and the Greek fathers strongly affected those early reformers and those who antedated the Reformation who were anxious to turn from the opulence of the Church of Rome and from the intricacies of late Scholasticism to the spirit of primitive and early Christianity.

From 1517 to 1521 Erasmus stayed chiefly in Louvain, where he was involved with the new college for the study of the sacred languages: Greek, Hebrew, and Latin. He corresponded with humanistic scholars all over the world and became, perhaps, the leading figure of the northern Renaissance. His influence was great in all Europe, especially in southern France and Spain (where he was offered a chair at the new University of Alcalá). Liberal and reformist theologians and classical scholars looked to him for inspiration. In 1521 he went back to Basel, where, with Froben, he published a long series of works on the church fathers (editions on St. Jerome, St. Cyprian [third century], Pseudo-Arnobius [fifth century], St. Hilary [c. 315–c. 367], St. Irenaeus [c. 120 to 140–c. 200 to 203], Ambrose [339–397], St. Augustine, St. Chrysostom [c. 347–407], St. Basil [c. 329–379], and St. Origen [185?–?254]), all of which helped center attention on the theology of the early fathers rather than on that of the medieval Scholastics. His *Colloquies*, first published after 1518 and in many revised and expanded editions thereafter, is an excellent example of the revived and revitalized Latin style of the Renaissance; the several editions include biting and satirical attacks on various human institutions and beliefs, especially those connected with the church and with popular superstition.

Erasmus's merciless and witty critiques of church practices, monastic activities, Scholasticism, popular religion, and so on, as well as his scholarly efforts toward establishing the Greek text and the meaning of the New Testament and the doctrines of the early church fathers, had made him outstanding in the movement for church reform. As the reform movement became more revolutionary, however, Erasmus tried to stay aloof from the struggles. Both orthodox and reformist theologians pressed him to take a stand, while he sought means for mediation and reconciliation. When Martin Luther became more aggressive and violent in his words and actions, and when various early reformers criticized Erasmus for his refusal to join them, he, always hypersensitive to criticism, withdrew more and more.

Finally, in 1524 Erasmus spoke out against Luther in his work *De Libero Arbitrio* (On Free Will), in which he tried to show that Luther had dogmatically decided that man had no free will, even though (1) the issue was so complex that no human could really find a satisfactory solution to the problem and (2) the biblical texts were so obscure that no one could really tell what they asserted. Erasmus maintained that he preferred to recognize the inability of man to discover adequate answers to such theological problems and to rest content with the decisions of the church on such matters. Luther's furious reply, *De Servo Arbitrio* (The bondage of will; 1525), cried out against Erasmus's gentle humanistic skepticism and his willingness to accept church teachings uncritically. Christianity, Luther insisted, requires certainties, not opinions or probabilities. Salvation cannot be based on doubts. He concluded that if Erasmus wished to remain a skeptic, he should remember that *Spiritus sanctus non est scepticus* (the Holy Spirit is not a skeptic) and that judgment day is coming.

Erasmus wrote another answer, the *Hyperaspistes* (1526), and to a great extent broke with his former

reform-minded friends. When the reformers took control at Basel in 1529, Erasmus left for Freiburg im Breisgau, Germany, where he stayed almost until his death (which occurred in Basel seven years later as he was preparing to return to Holland). During his last years he continued to use his vast scholarship, his pen, and his influence to bring about religious and political peace. Attacked by both the radical reformers and the conservative churchmen, he tried to find a moderate solution before both sides became so rigid that a compromise maintaining the unity of Christendom was impossible. He advocated making sufficient internal reforms within the church to satisfy the less extreme reformers. Various popes and some Reformation leaders took him seriously (Pope Paul III [1468–1549] is supposed to have wanted to make him a cardinal), while the Sorbonne theologians condemned some of his works and views. The Spanish Inquisition stamped out the influence of his followers in Iberia, and the leading reformers attacked him both as a petty, self-serving person and as a heretical religious thinker.

THOUGHT

Erasmus's ambiguous position in the religious struggles was probably the result of his peculiar nondogmatic point of view and his cautious attitude toward developments in human affairs. He claimed to advocate the "philosophy of Christ," in contrast with the various kinds of Scholastic theories put forth by the Thomists, the Scotists, the Ockhamites, and others. Their technical discussions about the nature of baptism, grace, and the freedom of the will left him entirely unmoved. Rather than take their arguments and analyses seriously and present refutations, Erasmus attempted to undermine the whole Scholastic approach with the force of his ridicule.

In place of the philosophical and theological systems of the time Erasmus set forth his "philosophy of Christ," to be arrived at by pious study rather than by disputations. This "philosophy" was supposed to represent the simple and essential message of Christianity in its spirit rather than its letter; it was a message to be lived, not to be formulated in abstract systems. It was a nondoctrinal religion, a religion without a theology, that could be approached through the early church fathers and the morality of the New Testament but not through the morass of distinctions, terminology, and theory built up in the Middle Ages. This outlook had a great impact on the most liberal reformers and the nondoctrinal mystics.

Erasmus, who was so fully aware of the foibles of man, was also extremely cautious about the genuine possibilities for reform or constructive improvement in man and his institutions. This may account for his refusal to leave the Church of Rome (although he died without receiving the sacraments). Some have interpreted this refusal as due to personal fears, but it seems more probable that Erasmus remained within the church because he believed that it was better to preserve and improve what already existed than to risk the even greater abuses that might follow the destruction of the current order. Erasmus saw the Church of Rome as fossilized, in much the same manner that he portrayed the Jewish synagogue. On the contrary, he saw the reformers as revolutionists who, intentionally or not, were destroying the very fabric of human existence. He told Luther, "I always freely submit my judgment to the decisions of the Church whether I grasp or not the reasons which she prescribes."

He also declared, in the midst of the early Reformation struggles, "I will put up with this Church until I shall see a better." He apparently felt that, given the human condition, it was important to retain the (far from ideal) way that Christ's message had been institutionalized; at the same time he urged a revival of concern for the substance of this message and a revitalization of the church through the correction of as many abuses as possible and the encouragement of scholarly and moral efforts to recapture the original Christian spirit. Otherwise, he feared, the frail human world might be torn entirely asunder. But, for better of worse, the course of events carried the division of Christendom to a complete rupture; each side became more and more rigid and dogmatic rather than compromising on a vague or undefined Erasmian position.

INFLUENCE

Although Erasmus can hardly be classified as a professional philosopher, he influenced the course of philosophy in many ways. His humanistic scholarship greatly affected the European educational system and, both personally and through his many writings, Erasmus greatly encouraged the teaching and study of Greek, Latin, and Hebrew—the languages that were most important to intellectual achievement. The upheavals in the curricula that occurred in most of the major institutions of higher learning at that time were in large measure due to Erasmus's influence and spirit, and the study of the hitherto unknown or neglected classics of both the Greco-Roman and the Judeo-Christian worlds (many in the editions prepared by Erasmus himself) that resulted from this was the source of many new ideas and theories that became part of the intellectual revolutions of the Renaissance.

Erasmus's ridicule of Scholasticism, although hardly a philosophical refutation of either its methods or its doctrines, created the generally accepted view that the medieval approach to philosophical questions was trivial and useless. He made it difficult for many intellectuals to take seriously the views of St. Thomas Aquinas, John Duns Scotus, William of Ockham, and such later Scholastics as Francisco Suárez.

Besides teaching future generations to scoff at the achievements of the school philosophers, Erasmus also had a major role in creating the critical spirit that culminated in the Enlightenment. Through his satire, his critical scholarship, and his undogmatic spirit Erasmus popularized a critical and questioning attitude toward accepted mores, institutions, opinions, and texts that was to flourish in many forms in the next centuries, undermining confidence in almost every area of traditional achievement.

Thus, Erasmus, who was essentially conservative by nature and who shunned almost all theoretical or philosophical discussion, not even wishing systematically to oppose dogmatism with skepticism, as Michel Eyquem de Montaigne later did, was one of the most influential figures of the sixteenth century in changing the entire intellectual climate of opinion and in establishing the direction in which modern thought developed.

See also Augustine, St.; Colet, John; Duns Scotus, John; Luther, Martin; Medieval Philosophy; Montaigne, Michel Eyquem De; More, Thomas; Ockhamism; Origen; Patristic Philosophy; Reformation; Renaissance; Scotism; Servetus, Michael; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Valla, Lorenzo; William of Ockham.

Bibliography

WORKS BY ERASMUS

Omnia Opera. 9 vols., edited by Beatus Rhenanus. Basel, Switzerland: N.p., 1540–1541.

Desiderii Erasmi Roterodami Opera Omnia, Emendatiora et avctiora. 10 vols., edited by Jean Le Clerc. Leiden, Netherlands: Petri Vander, 1703–1704.

The Epistles of Erasmus. 3 vols., edited and translated by Francis Morgan Nichols. New York: Longmans, Greene, 1901–1918.

Opus Epistolarum. 12 vols., edited by P. S. Allen. Oxford, U.K.: Clarendon Press, 1906–1958.

The Praise of Folly, edited and translated by Leonard F. Dean. Chicago: Packard, 1946.

Inquisitio de Fide, edited and translated by Ernst F. Winter. New York: N.p., 1961.

The "Adages" of Erasmus: A Study with Translations, edited by Margaret Mann Philipps. Cambridge, U.K.: Cambridge University Press, 1964.

The Colloquies of Erasmus. Translated by Craig R. Thompson. Chicago: University of Chicago Press, 1965.

La correspondance d'Érasme: Édition intégrale traduite et annotée d'apres l'Opus epistolarum. 12 vols., edited by P. S. Allen, H. M. Allen, and H. W. Garrod. Brussels, Belgium: Presses Académiques Européennes, 1967–1984.

The Collected Works of Erasmus. 86 vols. Toronto, Canada: University of Toronto Press, 1974–2003.

WORKS ABOUT ERASMUS

Allen, Percy Stafford. *The Age of Erasmus*. Oxford, U.K.: Clarendon Press, 1914.

Allen, Percy Stafford. *Erasmus: Lectures and Wayfaring Sketches*. Oxford, U.K.: Clarendon Press, 1934.

Bataillon, M. *Érasme et l'Espagne*. Paris: E. Droz, 1937. A two-volume Spanish-language edition translated by Antonio Alatorre was published by Fondo de Cultura Económica (Mexico City, Mexico) in 1950.

Bouyer, Louis. *Autour d'Érasme: Études sur le christianisme des humanistes catholiques*. Paris: Éditions du Cerf, 1955.

Feugère, Gaston. *Érasme: Étude sur sa vie et ses ouvrages*. Paris: L. Hachette et cie, 1874.

Froude, James Anthony. *Life and Letters of Erasmus: Lectures Delivered at Oxford 1893–4*. London: Scribner, 1894.

Haeghen, Ferdinand van der. *Biblitheca Erasmiana: Bibliographie des oeuvres d'Érasme*. 12 vols. Ghent, Belgium: C. Vyt, 1897.

Huizinga, Johan. *Erasmus and the Age of Reformation*. Translated by F. Hopman and Barbara Flower. New York: Harper, 1957.

Kaiser, Walter J. *Praisers of Folly: Erasmus, Rabelais, Shakespeare*. Cambridge, MA: Harvard University Press, 1963.

Kunze, Johannes. *Erasmus und Luther: Der Einfluss des Erasmus auf die Kommentierung des Galaterbriefes und der Psalmen durch Luther, 1519–1521*. Hamburg, Germany: Lit, 2000.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Renaudet, Augustin. *Érasme: Sa pensée religieuse et son action d'après sa correspondance (1518–1521)*. Paris: F. Alcan, 1926.

Renaudet, Augustin. *Érasme et l'Italie*. Geneva: E. Droz, 1954.

Torzini, Roberto. *I labirinti del libero arbitrio: La discussione tra Erasmo e Lutero*. Florence, Italy: L.S. Olschki, 2000.

Richard Popkin (1967, 2005)

ERIGENA, JOHN SCOTUS

(c. 810–c. 877)

John the Scot (Irishman) or Erigena (of Irish birth) was active as a scholar in the court of Charles the Bald around 850 to 870. He intervened in the debate on predestination with a controversial treatise. At the request of the emperor, he made a Latin translation of the works of Dionysius the Pseudo-Areopagite (followed later by

translations of St. Maximus the Confessor [c. 580–662] and Gregory of Nyssa). The direct contact with the Greek theological tradition opened his mind for a more speculative neoplatonic interpretation of the Christian doctrine of creation than what he knew from his Latin authorities. Confronting both hermeneutic traditions with the requirements of "right reason," Erigena composed his own theophilosophical synthesis, *Periphyseon*.

Periphyseon is an attempt to understand the "division of Nature" and its "unification," thus offering a comprehensive interpretation of the Christian doctrine of creation, sin, and salvation as revealed in Genesis 1–3. Nature stands for the whole universe, encompassing both God and creation in all its divisions. It is the task of the philosopher to examine both the division of this Nature, that is, its articulation into a manifold of species from the most general to the most particular, and its unification from the utmost manifold to absolute simplicity. In the neoplatonic tradition *diairesis* (which divides a genus into specific forms) and *synopsis* (which brings a dispersed plurality under a single form) are not just two logical procedures of dialectic. They correspond to the movements of reality: the procession of multiplicity from the One and its return into the One; in Christian terms, creation and redemption.

At the start, Erigena introduces his famous fourfold division of nature, which will provide the main structure for the entire discussion. By applying the dialectical method of dividing a genus into species by differences, he presents a division that can be applied to the whole Universe, or Nature. The most fundamental difference is that between creating and being created. Applying four possible combinations of these differences one may discover the four fundamental species of Nature:

(1) That which creates and is not created

(2) That which creates and is created

(3) That which is created and does not create

(4) That which neither is created nor creates

The first species of nature is God, the uncaused cause of everything. The third species, which is diametrically opposite to the first, stands for the sensible world, comprehending the many species of animals and plants that come to be in times and places. The second species has attributes of both extremes: it is both created and creative. This is the level of the primordial ideas wherein God has from all eternity created all species (before they are manifested in time and place and individualized in matter). Finally, there is the fourth nature, which must be understood again as God. It is, however, God not as the creative cause from which all things proceed, but as the ultimate Good toward which of all things return.

In this division the divine nature is that which stands first and last. Still, God is not simply a species among many, because he "transcends everything that is or can be" and thus seems to fall outside all system. But one could as well say that God is the whole system in its unfolding and that all four divisions of nature are moments within the circular process whereby the divine nature proceeds from and returns to itself. In fact, the most fundamental distinction, that between creative nature and created nature, must itself be overcome. This is most true on the level of the primordial ideas, wherein the creator expresses in his Word his being as the being of the creatures: therefore, that nature is said to be both creative and created. As Erigena provokingly says, "God is the essence of all things" (*essentia omnium*). "It follows that we ought not to understand God and the creature as two things distinct from one another, but as one and the same" (Vol. III, 678C). In fact, this sensible world has no subsistence on its own but exists only through participation in the divine being and the primordial causes.

If the being of the creature is nothing but a participation in the being of its creator, one may also understand the creation of the world as God's creation of himself. "God is everything that truly exists because he himself makes all things and is made in all things." By creating the manifold species God reveals and makes himself known proceeding from his ineffable nature, where he is unknown even to himself. In this sense creation is revelation and the whole world must be understood as a theophany, that is, an "appearing of God." For everything that exists is nothing else but "the apparition of what is not apparent, … the comprehension of the incomprehensible, the materialization of the spiritual, the visibility of the invisible" (Vol. III, 633A). When it is said in the Christian creed that God creates "from nothing," it can only mean that God creates all things "out of the nothingness," which he is himself as transcending all beings. Only in his creatures "he begins to be."

In this cosmic process of emanation and return, human nature occupies a central place. Human nature, which comprehends body, vital powers, perception, imagination, reason, and intellect, is the "workshop of all things" (*officina omnium*), the intermediary connecting the whole universe, preventing its falling into separate sensible and intelligible realms. Apart from being created, human nature resembles the divine nature in all respects. Thus, as the divine mind, the human soul finds in itself

eternal *a priori* knowledge of all created things. In the divine wisdom, however, things exist as primordial causes or substantial forms, in human knowledge as the effects of those ideal forms.

Through the Fall, however, this connatural knowledge has been lost and the soul has fallen into ignorance of itself and of the content of its ideas. Human nature turned away from the creator, dishonoring its natural dignity and making itself similar to the beasts. This irrational nature does not belong to the image of God. In his original plan God had wanted to create humans similar to angels, not divided into male and female, without needing for their multiplication a sexuality similar to that of irrational beasts. But because God had foreseen from all eternity that humans would abuse their freedom and sin, from the first moment of their temporal existence, and thus fall from the status of equality with the angels to the level of the beasts, he introduced in the creation of the human being the consequences of sin before it occurred. Thus, the sexualized fleshy body (with all what it involves as pain, passion, sickness, and corruption) was created with the original rational nature, an addition required as a remedy and a penance for sin. It will be overcome when, at the resurrection, all shall rise in a perfect, sexless, spiritual body.

A philosopher must not only explain how creatures proceed from God but also how they return "by the same stages through which the division had previously ramified into multiplicity, until it arrives at that One which remains inseparably in itself and from which that division started" (Vol. II, 526A). Erigena makes a clear distinction between the general return to God, which is the common and natural destination of the whole creation (all corporeal things will return, that is, be resolved into their incorporeal causes), and the special return, which is only reserved to rational beings, the angels and the humans. At the end all human beings, blessed and damned alike, will return to the perfection of one and the same human nature. Still, they will be individually distinguished, not by differences in nature, body, or place, but by the different access each shall be granted to God's self-revelation. Those who led a righteous life will be beatified and allowed to see God in differing gradations of his theophanies. The damned, on the contrary, will be refused access to that vision and will be eternally tormented with the "vain dreams" of those things that incited their desires while still living.

Erigena stands apart from any of his contemporaries in his original speculations on creation and redemption, showing a great confidence in the harmony of reason and revelation. Still, he only exercised a limited direct influence in the Middle Ages, where he was mostly appreciated as a translator of Dionysius. *Periphyseon* was condemned as heretical in 1225 and copies of it were burned. From a philosophical point his greatest accomplishment is his understanding of creation as the self-creation of God. This doctrine attracted the admiration of idealist philosophers such as Friedrich Wilhelm Joseph von Schelling and Georg Wilhelm Friedrich Hegel, which led to his rediscovery in the twentieth century.

See also Neoplatonism; Pseudo-Dionysius.

Bibliography

WORKS BY ERIGENA

Periphyseon = The Division of Nature. Translated by I. P. Sheldon-Williams. Montreal: Bellarmin, 1987.

Periphyseon. 5 vols. Turnhout, Belgium: Brepols, 1996–2003.

Treatise on Divine Predestination. Translated by Mary Brennan. Notre Dame, IN: University of Notre Dame Press, 1998.

WORKS ABOUT ERIGENA

Brennan, Mary. *A Guide to Eriugenian Studies: A Survey of Publications 1930–1987*. Fribourg, Switzerland: Editions Universitaires, 1989.

Carabine, Deirdre. *John Scottus Eriugena*. New York: Oxford University Press, 2000.

McEvoy, James, and Michael Dunne, eds. *History and Eschatology in John Scottus Eriugena and His Time*. Louvain, Belgium: Louvain University Press, 2002.

Moran, Dermot. *The Philosophy of John Scottus Eriugena: A Study of Idealism in the Middle Ages*. New York: Cambridge University Press, 1989.

Van Riel, Gerd, Carlos Steel, and James McEvoy, eds. *Iohannes Scottus Eriugena: The Bible and Hermeneutics*. Louvain: Louvain University Press, 1996.

Carlos Steel (2005)

EROS

See *Love*

ERROR

When we engage in discursive thought and declarative speech, we may attain various forms of success: intelligibility, precision, correctness, and so on. These felicities are best explained by contrast with the corresponding mishaps that threaten our beliefs, assertions, and especially our claims to know something. A person's thinking may be inadequate because he is ignorant, and what he

says may be deficient because it is incoherent, rough, or, perhaps most important of all, downright false.

Many philosophers have been troubled in attempting to account for the occurrence of falsity in people's assertions and opinions, that is, in trying to understand how there could be such a thing as error at all. In examining these difficulties, we shall assume a man's statement is erroneous in case it is false and it reflects his belief. Thus, if a man lies, he may speak falsely, but not erroneously. We shall also assume that a person holds a false belief when he is inclined to express it in a statement that would be false. The statement would be erroneous; consequently we can say the belief it mirrors is erroneous as well.

Our inquiry will focus on a pair of famous knots:

(1) Do we believe anything when our belief is false? If a surgeon is convinced his patient will die, and he is correct, there is something he expects: the patient's demise. But if he is mistaken, there is no such event as the patient's death. Did the surgeon then expect nothing? Depicted thus, erroneous thinking seems impossible. Plato inherited this perplexity from Parmenides and finally resolved it.

(2) Granting that error can occur, is it ever voluntary? Clearly we are to blame for some of our mistakes, yet who knowingly and willingly goes in for false beliefs? This puzzle comes from René Descartes.

THE POSSIBILITY OF ERROR

Parmenides' maxim was that only being—what is—can exist. From this he argued that we cannot "know that which is-not … nor utter it; for the same thing can be thought as can be." In other words, "You will not find thought without what is, in relation to which it is uttered" (G. S. Kirk and J. E. Raven, *The Presocratic Philosophers*, Cambridge, U.K., 1957, pp. 269–278). According to Plato, the Sophists drew an incredible doctrine from these enigmas: They argued that error is asserting and believing "what is not," specifically, talking and "thinking contrary to the things that are" (*Sophist* 240C, D, F. M. Cornford translation). To state and think what is not is, however, to state and think nothing, which is not stating or thinking.

What could have led the Sophists to this paradoxical view? Naturally, it reinforced other logically quite independent doctrines of theirs regarding truth and falsity, for example, their claim that whatever seems true to any man is true for him and therefore *is* true (*Theaetetus* 161C–179C). The only direct support for their denial of error, however, derives from analogies. In *Euthydemus* (283E–288A) the stating of what is not is compared with impossibilities such as doing but doing nothing or gesturing although there is no gesture one performs. The parallel with sense perception in *Theaetetus* (188D–189B) is renowned. If you see, hear, or touch, then there is something you perceive. To see what is not, is not seeing.

PLATO'S THEORY OF ERROR. In reply, Plato suggests alternative models that will incline us to regard falsehoods as full-fledged (though incorrect) assertions. His most promising comparison is with spelling, where misspelling is the natural counterpart to stating falsely. This analogy is central to Plato's unsuccessful third definition of knowledge in *Theaetetus* (201D–208B), and it reappears in his successful treatment of falsity in *Sophist* (242D–253, 261D–264B). We should imagine a student who tries to spell syllables and words as his teacher says them or inscribes them. When he answers, aloud or in writing, with the right letters in the right order, he uses letters to represent a thing "that is," the teacher's utterance or inscription. According to conventions of spelling, the correct sequence of letters corresponds to what the teacher said or wrote. When the student misspells, he fails to represent the teacher's utterance or inscription, and nothing corresponds to the misspelling. Does the pupil therefore spell nothing? We describe his failure more accurately by saying: "When he misspells, what he spells is not; that is, not anything said or written by the instructor." How shall we describe cases where he gets some letters right and thus represents phonetic or graphological elements of what the instructor said or wrote? Clearly he does not *spell* those elements. For suppose he answers, "w-i-n" after the instructor says "wine." It makes no sense to claim: "He misspelled the word, but he correctly spelled some sounds the teacher made." Further, if the student happens to give the correct spelling of another word (as in this case), we must realize that the other word "is not," because the teacher did not say it. All this is secondary, however; the significant point is that when he misspells, the pupil is nevertheless engaged in the activity of spelling. This suggests that incorrect statements are statements after all.

Let us develop this parallel. If we restrict ourselves to the simplest paradigms—Plato used the true statement "Theaetetus is sitting" and the false statement "Theaetetus is flying"—we notice how words in declarative speech function like letters in spelling. Individual letters may represent sounds but do not spell them. Similarly, the name "Theaetetus" stands for a thing that is, but saying the name "Theaetetus" is not stating anything. Spelling words requires you to conjoin letters of different types, consonants and vowels, in appropriate patterns. Speech,

too, is fitting words together. "If you say 'lion stag horse' or any other names," Plato remarks, "such a string never makes up a statement"; but joining a noun with a verb "gets you somewhere" and of such a compound "we say it 'states,' not merely 'names' something" (*Sophist* 262B–E).

Now if you proclaim, "Theaetetus is flying," when he is not airborne, you refer by name to something that is, in the course of *stating* what is not. There is no aerial activity of Theaetetus to correspond to what you state. If you declare, "Theaetetus is sitting," and his posture is appropriate, then something corresponds to what you state. In Plato's words, the true statement "states about [Theaetetus] the things that are"; whereas "the false statement states about [him] things *different* from the things that are," and therefore states "things *that are not* as being" (*Sophist* 263B).

This correspondence theory of true and false statements may be extended to thought by our assumption that a man thinks falsely in case he is inclined to express his thought in a statement that is false. So formulated, the Platonic theory illustrates how error, even though it is believing what is not, hardly consists in believing nothing. Therefore, Sophists cannot maintain that thinking erroneously is not thinking.

Plato's account is, however, needlessly anchored to the type of counterexample he used against the Sophists. The falsehood "Theaetetus is flying" happens to be "about" an existing thing, but Plato makes this feature a prerequisite for every statement, true or false. He writes: "Whenever there is a statement, it must be about something" that exists (*Sophist* 262E; cf. 263C). Even if this ruling allows statements about things that do not exist now but did or will exist, it excludes too much. For instance, it would be impossible for me to state falsely, "There are flying saucers." By Plato's rule, saying this is not stating unless there are flying saucers for me to talk about. So if I *state* that there are flying saucers, I speak truly.

Here we should invoke Plato's orthographic model. The student does not cease to spell when upon occasion he misses every letter and thus fails to represent any sounds his teacher made. By analogy, why disqualify my utterance simply because I fail to refer to existing things? A correspondence theory still explains why it is false to state, "There are flying saucers": nothing corresponds to what is stated; that is, nothing corresponds to the existence of flying saucers, because none exist.

But the correspondence theory needs elaboration before it will transform Plato's view into a general account of correct and incorrect assertion. What "things that are" would I depict if I conceded "Flying saucers do *not* exist"? Does the nonexistence of flying saucers correspond to what I state? How can there be such a thing? Again, what "things that are" differentiate a true subjunctive conditional, for example, "If I had watered the lawn, it would not have died," from its false contrary, "Even if I had watered the lawn, it would have died"? Does the same withered grass make one statement true and the other false?

Because of these obscurities it appears that Plato has demonstrated how incorrect belief and assertion can occur, but he has not produced an exhaustive analysis of them. Plato's demonstration shows awareness of the distinctive features of assertion, sensitivity to the differences between referring and asserting, and perspicacity about the ontological status of what a person believes; indeed, treatments of error by such twentieth-century philosophers as G. E. Moore and Bertrand Russell are not more satisfactory in these respects.

MOORE AND RUSSELL. Moore's *Some Main Problems of Philosophy*, adapted from lectures he delivered in 1910–1911, appears to contain two incompatible theories: (1) The dyadic theory, according to which believing pairs believers and propositions. "Error," writes Moore, "always consists in believing some proposition which is false" (p. 66). (2) For complicated reasons Moore later contends that "there simply are no such things as propositions" for people to believe (p. 265), and he renounces his "attempt to analyse beliefs" (p. 266). Nevertheless, he characterizes truth and falsity of beliefs as follows: "To say that a belief is true is to say that the fact to which it refers is or has being; while to say that a belief is false is to say that the fact to which it refers is not" (p. 267). In technical terminology: "To say that [a] belief is … false is to say that there is *not* in the Universe any fact to which it corresponds" (p. 277).

In his 1953 preface, Moore comments that his two theories may after all be consistent; perhaps he used the term *proposition* in different senses when he first maintained and later denied their existence as targets for believing (p. xii). Apart from this problem, fundamental questions arise concerning Moore's treatment: Are there any facts for mistaken beliefs to be about? Are they nonexistent facts "in the Universe," or perhaps existent ones outside it? Besides, the very notions of "proposition" and "fact" are notoriously obscure.

Moore hints at a different analysis when he considers a person's belief that we are now listening to a brass band. What the person erroneously believes is a "combination

of us at this particular moment with the hearing of that particular kind of noise"—a combination which "simply has no being" (p. 255).

Russell's multiple relation theory, in his *Problems of Philosophy*, develops such an analysis. Concerning Othello's mistaken belief that Desdemona loves Cassio, Russell says: "The relation called 'believing' is knitting together into one complex whole the four terms Othello, Desdemona, loving and Cassio" (p. 126). This belief is mistaken because there does not exist another "complex unity, 'Desdemona's love for Cassio,' which is composed exclusively of the *objects* of the belief, … with the relation [loving] which was one of the objects occurring now as the cement that binds together the other objects of the belief" (p. 128).

The snags in this analysis are well known. How does loving cement things? In case Desdemona is indifferent to Cassio, how does believing sew them together with loving and Othello? Can believing stitch together any collection of objects? Russell noted the last two problems in his 1918 lectures, "The Philosophy of Logical Atomism" (in *Logic and Knowledge*). He recognizes that the structure of Othello's belief requires that "loves" should "occur as a verb"; but he is afraid that admitting the syntactical unity of "Desdemona loves Cassio" is tantamount to "assuming the existence of the non-existent … [namely,] a non-existent love between Desdemona and Cassio" (p. 225). But as Plato saw, there can be units of speech—statements—and thought without correspondents in reality.

ERROR AND VOLITION

Plato did not need to convince us that false belief is possible. But Descartes's thesis, that error is always voluntary, seems a contrived solution to an entirely gratuitous theological muddle. This appearance is deceptive.

Descartes begins his *Meditations* with a survey of dreaming, sensory illusion, and the errors they occasion; next he shows we can prove a few things for certain, including God's existence. Then he reasons: The deity "cannot have given me a faculty [of thinking] whose right employment could ever lead me astray"; however, "it seems to follow that therefore I can *never* go wrong" (*Meditation* IV, in *Descartes' Philosophical Writings*, translated by P. T. Geach and G. E. M. Anscombe, p. 93). Descartes's answer to this puzzle is that men have false beliefs, but through their own doing, not God's. Men are endowed by God with such power of will that they can assent to propositions they do not know to be true—that is, to "ideas" that are not "clear and distinct." Is God to blame for this disharmony between our limited capacity for knowledge and our unlimited power of assenting? No, "will is just a single thing; it is incompatible with its nature that anything should be subtracted from it" (*Meditation* IV, p. 99). Besides, although we are free to, we do not have to believe propositions for which we lack conclusive proof. In order to avoid "unsuspected error," we must restrain our desire for truth and withhold assent until we know for certain. (Descartes's *Principles of Philosophy* XXIX, XXXII, XXXV, XXXIX, XLII explain these points in detail.)

Now if we put aside Descartes's theological preoccupations, and his advice that we should only believe what is obvious, at least two genuine questions arise:

(1) Do we exercise any control over our convictions and opinions, as Descartes's concept of "assenting" requires? Clearly, people may decide to make statements. Some criminals voluntarily confess their misdeeds, and others are forced, against their will, to admit guilt. How about belief? Can we choose to reject a proposition that seems most likely, according to available evidence, and believe another that seems less plausible? Perhaps not. But we often make decisions as we form our opinions, as we collect or neglect data and seek or ignore expert testimony. Men who undergo brainwashing are deprived of this control over the formation of their beliefs. The same holds, incidentally, for knowledge. It is absurd to say the investigator decided *to know* but not absurd to say he resolved *to find out* for certain who robbed the grocer. Moreover, children are compelled *to learn* things. In acquiring knowledge and forming opinions, we pursue rather obvious goals: conclusive proof and correct information.

(2) Even so, is it intelligible to suppose that people act deliberately and knowingly when they settle upon *false* beliefs? One everyday case should dispel the appearance of contradiction: I study the racing form, mull over the evidence, and conclude that Wayfarer is bound to take the handicap. I willingly commit myself to this belief by wagering my paycheck. I realize, however, that even well-grounded expectations like mine can prove erroneous. Consequently, if my horse loses, it is true to say, "I formed my erroneous belief willingly, after deliberation, with the intention of predicting the handicap winner; furthermore, I was aware that I could be mistaken." It was not my goal to be wrong, but it was within the scope of my intention. Anyone who aims at truth is prepared for falsity, just as a marksman is prepared to miss the bull's-eye. Can we say I erred "knowingly"? A man who punches another is hardly ever certain that his victim will be injured, but

from a legal standpoint he knowingly inflicts harm if he has reason to think injury might result from his blow.

There remains another type of error, fortunately quite infrequent, where such awareness is impossible. This is the unusual situation where you are convinced you know something, banish doubt from your mind, and still turn out to be wrong. Perhaps you acted deliberately and followed your inclinations in pushing your investigation until you believed you could not be wrong. But with this degree of conviction, you cannot have the least awareness that you are mistaken. Your error, then, is not fully voluntary.

See also Correspondence Theory of Truth; Descartes, René; Moore, George Edward; Plato; Russell, Bertrand Arthur William; Sophists; Volition.

Bibliography

PLATO

Plato's *Theaetetus* and *Sophist*, along with excellent commentary, are translated by F. M. Cornford in *Plato's Theory of Knowledge* (London: K. Paul, Trench, Trubner, 1935). For further analysis of Plato's theory, see Gilbert Ryle, "Letters and Syllables in Plato," *Philosophical Review* 69 (4) (1960): 431–451; D. Gallop, "Plato and the Alphabet," in *Philosophical Review* 72 (3) (1963): 364–376, and Raphael Demos, "Plato's Philosophy of Language," with comments by J. L. Ackrill, in *Journal of Philosophy* 61 (20) (1964): 595–613.

MOORE AND RUSSELL

The views of Moore and Russell can be found in G. E. Moore, *Some Main Problems of Philosophy* (London: Allen and Unwin, 1953), especially Chs. 3 and 13–16, and Bertrand Russell, *Problems of Philosophy* (London: Williams and Norgate, 1912; reset 1946), Ch. 12, and "The Philosophy of Logical Atomism," in *Logic and Knowledge*, edited by R. C. Marsh (London: Allen and Unwin, 1956). For a view opposed to those of Moore and Russell, see Alexius Meinong, *Untersuchungen zur Gegenstandtheorie und Psychologie* (Leipzig, 1904), translated by Isaac Levi, D. B. Terrell, and Roderick Chisholm as "The Theory of Objects," in *Realism and the Background of Phenomenology*, edited by R. M. Chisholm (Glencoe, IL: Free Press, 1960).

There is a brilliant discussion of the connection between negative judgments and falsity in Gottlob Frege's essay "Negation," in *The Philosophical Writings of Gottlob Frege*, edited and translated by Max Black and P. T. Geach (Oxford: Blackwell, 1952).

DESCARTES

See *Descartes' Philosophical Writings*, translated by P. T. Geach and G. E. M. Anscombe (Edinburgh, 1954). Spinoza's objections to Descartes's theory that belief necessarily involves assent can be found in the *Ethics*, Part I, Props. 29, 32, and appendix, and all of Part II; see also the lucid analysis of Descartes and Spinoza by J. L. Evans, "Error and the Will," in *Philosophy* 38 (1963): 136–148.

OTHER RECOMMENDED TITLES

Almeder, Robert. "Recent Work on Error." *Philosophia* 27 (1999): 3–58.

Bennett, Jonathan. "Spinoza on Error." *Philosophical Papers* 15 (1986): 59–73.

Burgess, John. "Error Theories and Values." *Australasian Journal of Philosophy* 76 (1998): 534–552.

Devlin, John. "An Argument for an Error Theory of Truth." *Nous*, Supp. 17 (2003): 51–82.

Gomez-Torrente, Mario. "Vagueness and Margin for Error Principles." *Philosophy and Phenomenological Research* 64 (2002): 107–125.

Howson, Colin. "Error Probabilities in Error." *Philosophy of Science*, Supp. 64 (4): S185–S194.

Popkin, Richard. *Scepticism from Erasmus to Spinoza*. Berkeley: University of California Press, 1979.

Sievert, Donald. "Whose Fault Is It Anyway? Descartes's Doctrine That We, and Not God, Are Responsible for the Errors of the World." *Southwest Philosophy Review* 13 (1997): 33–41.

Stroud, Barry. "Explaining the Quest and Its Prospects: Reply to Boghossian and Byrne." *Philosophical Studies* 108 (2002): 239–247.

Wiland, Eric. "Psychologism, Practical Reason and the Possibility of Error." *Philosophical Quarterly* 53 (2003): 68–78.

Williams, Bernard. *Descartes: The Project of Pure Enquiry*. Atlantic Highlands, NJ: Humanities Press, 1978.

Irving Thalberg (1967)
Bibliography updated by Benjamin Fiedor (2005)

ERROR THEORY OF ETHICS

An "error theory of ethics" is the view that the ordinary user of moral language is typically making claims that involve a mistake. The concepts of ethics introduce a mistaken, erroneous, way of thinking of the world or of conducting practical reasoning. The theory was most influentially proposed by John L. Mackie in his book *Ethics: Inventing Right and Wrong* (1977). Mackie believed that ordinary moral claims presuppose that there are objective moral values, but there are no such things. Hence, the practice of morality is founded upon a metaphysical error.

Mackie's arguments against the existence of objective values are of two main kinds. One is the argument from relativity, which cites the familiar phenomenon of ethical disagreement. Another is the argument from "queerness." The moral values whose existence Mackie denies are presented as metaphysically strange facts. They are facts with

a peculiar necessity built into them: their essence is that they make demands or exist as laws that "must" be obeyed. In Kantian terms, the demands made by morality are thought of as categorical, "not contingent upon any desire or preference or policy or choice." The foundation of any such demands or laws in the natural world is entirely obscure. Hence, the right response of a naturalist is to deny that there can be such things. It should be noticed that this is not supposed to be an argument against any particular morality, for instance, one demanding honesty or fidelity, but against the entire scheme of thought of which particular ethical systems are examples.

Another influential theorist whose work bears some resemblance to Mackie's is Bernard Williams, whose *Ethics and the Limits of Philosophy* (1985) equally raises the doubt that ethics cannot possibly be what it purports to be, although Williams's own arguments are more specifically targeted on the morality of duty and obligation.

Responses to the error theory have taken several forms. Both the argument from relativity and that from queerness have been queried, the former on the grounds that, even if ethical opinions differ fundamentally, this does not prevent one from being right and the others wrong, and the latter mainly on the grounds that Mackie suffered from an oversimple, "scientistic" conception of the kind of thing a moral fact would have to be. Perhaps more fundamentally, it is not clear what clean, error-free practice the error theorist would wish to substitute for old, error-prone ethics. That is, assuming that people living together have a need for shared practical norms, then some way of expressing and discussing those norms seems to be needed, and this is all that ethics requires. Mackie himself saw that ethics was not a wholly illegitimate branch of thought, for he gave a broadly Humean picture of its function in human life. Even projectivists maintain that our need to express attitudes, coordinate policies, and censure transgressions is a sufficient justification for thinking in terms of ethical demands. Ethics does not invoke a strange world of metaphysically dubious facts but serves a natural human need.

See also Mackie, John Leslie; Metaethics; Williams, Bernard.

Bibliography

Blackburn, S. "Errors and the Phenomenology of Value." In *Essays in Quasi-Realism.* New York: Oxford University Press, 1993.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* Middlesex, U.K.: Penguin, 1977.

Williams, B. *Ethics and the Limits of Philosophy.* Cambridge, MA: Harvard University Press, 1985.

Simon Blackburn (1996)

ESCHATOLOGY

"Eschatology" is a doctrine or theory (*logos*) of the end (*eschaton*). "End" here can have two meanings. First, it can mean the end of each individual human life. Second, it can mean the end of the world—or, more narrowly, of the human race. In the first, the individualistic, sense eschatology is an account of the destiny that awaits each person after death. In the second, the cosmic or social, sense it is a description of a goal (*telos*) in which history will be fulfilled. This goal may be of either a this-worldly or an otherworldly kind.

The distinction between these two senses is important, for it is possible to have an eschatological doctrine in one sense without having any in the other. Plato held that the soul, being immortal, would face judgment after death, that it would receive rewards and punishments according to the goodness or badness of its earthly life, and that it would be given an opportunity to choose the condition of its next existence (*Republic* 608C to end). However, he did not believe that there was any purpose to history as a whole. Conversely, a Marxist believes in a purpose of history although he disbelieves in personal survival.

It is doubtful whether eschatology in the second sense is to be found anywhere outside Zoroastrianism and Judaism—together with the religious and philosophical systems that have drawn inspiration from them: Mithraism from the first, Christianity and Islam, and Western thought in general, from the second. According to Greek and Indian thinkers history moves in cycles. Just as the seasons recur within each solar year, so all events recur in a sequence of "Great Years." By contrast, the Persian Zend-Avesta and the Bible state that history is nonrepeatable and that it is destined for a divine fulfillment in which good will triumph over evil.

In the Bible the second sense predominates. The Old Testament contains only a few vague references to a personal afterlife. But it often refers to a future time when God will establish his everlasting reign of righteousness and peace (for example, Isaiah 11:1–9). The New Testament affirms that this divine end or goal has been reached by the exalted Christ, who defeated the powers of evil on

the cross (see, for example, Acts 2:14–36; Colossians 2:8–15; Ephesians 2:11–22; Hebrews 2:5–18). Those who believe in Christ have eternal life here and now (John 3:36; 5:24). While living in "this age," this spatiotemporal order that is still subject to sin and death, they have a foretaste of "the age to come," a renewed cosmos that will be wholly subject to the will of God.

This view of history stands in contrast, first, to the Greco-Roman theory of recurrent cycles—a theory condemned by Origen and Augustine—second, to the humanistic dogma of inevitable social progress, and, third to Marxism. Although the Marxist philosophy of history owes its form to G. W. F. Hegel's dialectic, its content has often been called a secularization of Christian eschatology. Materialistic determinism would be equivalent to a personal providence, the proletariat to the "chosen people," and the "classless society" to the kingdom of God.

During the early centuries of the church most theologians taught that there will be a universal resurrection of the dead for a final judgment at the end of history, when Christ will appear again "in glory." As a result of this judgment, it was also generally taught, some, the saved, will pass to paradise, where they will enjoy the beatific vision, but others, the damned, will be punished with everlasting torment. Four comments on this scheme are necessary:

(1) One must distinguish between belief in the immortality of the soul and belief in the resurrection of the body. The first belief is derived from Plato, who held that the soul will survive in an incorporeal state. The second belief is based on biblical revelation. Thomas Aquinas held both beliefs. He considered the immortality of the soul to be rationally demonstrable. He also thought that the dogma of a bodily resurrection could be rationally justified on the ground that since soul and body constitute (as Aristotle taught) a single substance, the soul requires the body for its self-expression and beatitude. (To account for the obvious fact that the flesh decays at death, Origen proposed the theory that although the resurrected body will have the same "form" as its earthly counterpart, it will have a different "matter.")

(2) Origen maintained that all spiritual creatures—angels, humans, and devils—will be saved in a final "restoration" (*apocatastasis*). But although his doctrine (known as Universalism), which was shared by Gregory of Nyssa, could claim biblical support, it was attacked by Augustine and formally condemned.

(3) Even orthodox Christian Fathers (such as Irenaeus), as well as Gnostics and Montanists, were millenarians. They believed that Christ would reign on Earth for a thousand years before the end of terrestrial history. But since the fifth century millenarianism has been almost wholly confined to minor sects.

(4) Although Clement of Alexandria and Origen spoke of a fire that would purge guilty souls, the full doctrine of purgatory (as a place of temporary punishment preparatory to the beatific vision) was not developed until the Middle Ages.

In the twentieth century there was a new attempt to understand the eschatological teaching of the New Testament (especially in the light of Albert Schweitzer's thesis that Jesus expected an imminent end of history and therefore intended his moral teaching solely for an interim). On the other hand, Rudolf Bultmann attempted to "demythologize" biblical eschatology, to restate it in existentialist terms that will make it intelligible to modern man. These instances indicate a twofold revival of interest in eschatology among professional theologians.

See also Augustine, St.; Bultmann, Rudolf; Clement of Alexandria; Death; Eternity; Gregory of Nyssa; Hegel, Georg Wilhelm Friedrich; Marxist Philosophy; Origen; Plato; Thomas Aquinas, St.; Zoroastrianism.

Bibliography

Bultmann, R. *History and Eschatology, the Presence of Eternity.* Edinburgh: Edinburgh University Press, 1957.

Charles, R. H. *Eschatology.* London, 1913. The standard work on Jewish and early Christian eschatology.

Corcoran, Kevin, ed. *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons.* Ithaca, NY: Cornell University Press, 2001.

Cullmann, O. *Christ and Time.* Translated by Floyd V. Filson. London, 1951. An excellent study on the New Testament.

Flew, Antony. *The Logic of Mortality.* Oxford: Blackwell, 1987.

Hick, John. *Death and Eternal Life.* New York: Harper and Row, 1976.

Penelhum, Terence. *Survival and Disembodied Existence.* New York: Humanities Press, 1970.

H. P. Owen (1967)
Bibliography updated by Christian B. Miller (2005)

ESP PHENOMENA, PHILOSOPHICAL IMPLICATION OF

See *Parapsychology*

ESSENCE

See *Definition; Essence and Existence; Universals, A Historical Survey*

ESSENCE AND EXISTENCE

It will avoid misunderstanding if the topic of essence and existence is expounded in an order other than chronological.

SEVENTEENTH-CENTURY VIEW

Thomas Hobbes and John Locke insisted that definitions are not of things but of names. In so doing, they conceived of themselves as breaking with Aristotelianism. Hobbes said that the essence of a thing is "that accident for which we give a certain name to a body, or the accident which denominates its subject"; he had earlier denied that "the definition is the essence of any thing"; and his example of an essence is that "extension is the essence of a body" (*De Corpore,* II, 8, 23).

Locke distinguished the real essence from the nominal essence; the nominal essence is the idea of the property or properties the possession of which justifies the application of a given name; the real essence is as it is understood by "those who look on all natural things to have a real but unknown constitution of their insensible parts, from which flow those sensible qualities which serve us to distinguish them one from another, according as we have occasion to rank them into sets under common denominators" (*Essay* III, 3, 17).

The mistake that Hobbes and Locke ascribed to Aristotelianism was that of confusing the meaning of an expression with the nature of the object which the expression characterizes. In the empiricist tradition this separation of questions of meaning from questions of characterization continued to be influential. One consequence is that the concept of the real essence is dropped altogether. Another is that philosophy itself becomes defined as the study of meaning, as a linguistic inquiry. But will the Lockean separation of the real and the nominal, from which so much of this derives, bear scrutiny? Did Aristotle commit the error ascribed to him? Is it an error?

ARISTOTLE AND THOMAS AQUINAS

For Aristotle, the essence of an object (*τὼ τίῆν εῖναι*) was what finds expression in the concept that the object embodies, the concept under which it must be identified if it is to be identified as what it is. The natural response of someone trained in the empiricist tradition is to question this concept of an object. In any particular case the question "What is *this?*" can have more than one correct answer—for instance, "a coat" or "a piece of cloth." To ask further what the essence of the thing indicated is, is to miss part of the Aristotelian point, which is best brought out by considering problems of identity. If I ask whether this is the same coat that you wore last year, I am not asking the same question that I would be asking if I inquired whether this is the same cloth made up into trousers that you used to wear in the form of a coat. "The same coat" and "the same cloth" pick out different identities. When I pick out "this" as an object, I can do so only by identifying it under some description, and the object does not have a nature apart from being identified under a description. For otherwise I could not identify what was to be characterized. In other words, we cannot identify an object solely by means of pronouns like "this" or by pointing.

It might be thought a fatal objection to this view that I can apparently identify an object without knowing what it is (a case which Aristotle allows for). Suppose I pick up something in your room and simply ask, "What is this?" The range of possible answers includes "a piece of stone," "a carving," "an image of a saint." My ignorance may extend as far down the range of specificity as you please; I must still be able to find some description to add to my use of "this" or to my pointing. For if I pick it up twice, I must be able to identify it as the same object; and it is a condition of my identifying it as an object at all that I should be able in principle to pick it up, or otherwise indicate it, more than once. Possible reidentification is a necessary condition of identification. But if this is so, then I must, in picking it up, be able to characterize it, even if only as "that small colorless lump" or some such description. There is a limit to vagueness, at which such purely formal concepts as "thing" and "object" lie.

Insofar, then, as Aristotle is concerned with the minimal conditions for identifying and characterizing objects, he is justified in a view which makes understanding what something is, inseparable from understanding

the meaning of the description which must be applied to it if it is to be identified as what it is. The nominal and the real cannot be entirely divorced. But Aristotle expresses all this in terms of the concepts of substance and of matter and form, and in so doing he appears to lay himself open to the Hobbes-Locke type of criticism. What Aristotle meant by *τί ἦν εἶναι* is the subject of disagreement among translators and commentators. Hugh Tredennick in translating *Metaphysics* 1031a15 ff.) uses "essence"; Joseph Owens invents an arbitrary phrase, "the What-Is-Being" of a thing, and explains it in terms of the being of a thing which is the being of its form. The form is the necessary and unchanging element in a thing, in contrast with the matter and the composite, which may change, and the generic characteristics, which may belong to some other species (*Doctrine of Being in the Aristotelian Metaphysics,* p. 94).

Aristotle thus made the concept, under which an object must fall if it is to be identified and characterized as what it is, express a timeless and necessary element in the nature of the object itself. And insofar as Hobbes, for example, wished to deny that this timeless and necessary element was what a definition could refer to, it would be difficult to disagree. But any further discussion of Aristotle could only proceed by analysis of the doctrine of matter and form.

What is clear is that Aristotle inherited from Plato the notion of a range of fixed and timeless Forms, natures or essences which are embodied in the changing physical world. Less pessimistic than Plato about the possibility of knowledge of the nature of particular material objects, he retained the view that what the intellect grasps is always a form which could have been embodied in other matter. The name given to the being that the intellect grasps is *οὐσία,* which W. D. Ross renders as *essence,* following Quintilian and Seneca, who translated it as *essentia. Essentia* comes to mean the nature of a thing, the answer to the question *quid sit.* Augustine used *substantia* and *essentia* without difference of meaning, and Boethius translated *οὐσία* as *substantia.* From then on the word *substantia* was used in this sense and *essentia* was reserved for a new context which was first found in explicit form in Giles of Rome. This contrast is that between essence and existence, which received its completest statement in the work of St. Thomas Aquinas.

THOMAS AQUINAS. A substance is composite; it is an essence upon which existence has been conferred. When existence is conferred on an essence, what was hitherto merely possible becomes actual. In the case of physical bodies, a form receives matter. Thus the concepts of essence and existence, potency and act, form and matter are mutually correlative. The notion of *esse* being conferred upon an *essentia* so that a substance is brought into being was foreign to Aristotle because the notion of creation was foreign to him. For Aristotle, analysis in terms of essence or substance was a way of approaching what already exists or is in the process of change. For Thomas, that anything at all exists must itself be explained. It is a purely contingent fact that any particular essence is an embodied existent. The only exception to this is God, in whom essence and existence are identical. But it does not follow that by grasping what God is, we can grasp that he is, as Anselm had supposed in his vision of the Ontological Argument. For we cannot grasp the divine essence.

MODERN VIEWS

The vocabulary of essence and existence was preserved after the seventeenth century both by late Scholasticism and by its intellectual first cousin, rationalist metaphysics. Christian von Wolff inherited, perhaps from Francisco Suárez, whose influence he acknowledged, a view of the universe as a system of essences on which God has chosen to confer existence. But his view of essence as what can be conceived as a clear and distinct idea points to the influence of René Descartes and in his version of the Ontological Argument we can see the confluence of John Duns Scotus and Gottfried Wilhelm Leibniz. Knowledge of essences is expressed in propositions which are necessary truths. But these necessary truths are truths about possibilities, and it is a contingent matter of God's will being what it is that these particular essences have been actualized.

A line of thought that is only superficially like that of rationalist metaphysics runs from Spinoza to G. W. F. Hegel. In Spinoza the essence that entails existence is that of the single substance. But this version of the Ontological Argument is only part of Spinoza's whole set of theorems determining the all-inclusive *Deus sive natura.* Hegel treated the transition from essence to existence as part of the logical play with concepts that is an essential preliminary to the world of becoming. Of course we cannot deny that being is; but that, for Hegel, is only because the assertion is so bare and empty. When we deal with the realm of becoming, we have the sharpest of contrasts between the *Was-sein* (essence) and the *Das-sein* (existence), as Friedrich von Schelling, the enemy of all clear distinctions, complained.

The notion of an essence as a fixed possibility whose character may be delimited apart from our acquaintance

with the existence which embodies it was inherited from Scholasticism by Franz Brentano and Edmund Husserl, whose phenomenology is concerned with essences insofar as it is a study of what is involved in *any* act of judgment, belief, feeling, or willing, independently of the context of particular acts. The use of "essence" and "existence" by Jean-Paul Sartre is partly derived from phenomenology and partly from Scholasticism. The latter influence is apparent in the way in which Sartre uses the formula that existence precedes essence in order to deny that men are created by God. Sartre identifies such a conception of creation with the notion of God creating beings with fixed, already determinate natures who would therefore be unfree. Nothing of this appears to be entailed by Thomas's use of "essence" and "existence," but Leibniz and Wolff could be more convincingly convicted on this charge. Sartre wishes to convey by his formula that men do not have determinate natures, fixed in advance of their choices. By this he means that Smith does not have an existence determined for him which if he did not live out he would not be Smith; so it is Leibniz or Wolff, and not Thomas, whose propositions he is in fact denying.

What, then, is Sartre asserting? The contention that existence precedes essence may be interpreted as entailing various consequences, not all of which were necessarily intended by Sartre. Sartre clearly does believe that his contention not only constitutes the denial of one species of determinism, as has already been noted, but also involves the invalidity of any version of the Ontological Argument, whether Anselmian, Cartesian, or Hegelian. That is, no essence is such that it is a necessary truth that there must exist some individual embodying that essence. But unfortunately the Sartrian contention is so loosely stated that he might also be taken to imply—what he certainly would not want to imply—that there are no essences, that is, no meanings, apart from existences. This is plainly false. Many meaningful expressions do not name or denote anything that exists, many descriptions do not characterize anything that exists, as the common examples of "unicorn" and "glass mountain" make clear. The chief difficulty with the Sartrian thesis, however, is not that it plainly entails absurd consequences. It is, rather, that the thesis is stated so generally, and is so inadequately developed, that it is not at all clear what does or does not follow from Sartre's contention.

See also Aristotelianism; Aristotle; Brentano, Franz; Descartes, René; Duns Scotus, John; Giles of Rome; Hobbes, Thomas; Husserl, Edmund; Leibniz, Gottfried Wilhelm; Locke, John; Ontological Argument for the Existence of God; Ross, William David; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von; Seneca, Lucius Annaeus; Spinoza, Benedict (Baruch) de; Suárez, Francisco; Thomas Aquinas, St.; Wolff, Christian.

Bibliography

PRIMARY SOURCES

Aristotle. *Metaphysics.* Translated by Hugh Tredennick. Cambridge, MA: Harvard University Press, 1944.

Aristotle. *Metaphysics.* Edited by W. D. Ross. Oxford, 1948.

Sartre, Jean-Paul. *L'être et le néant, essai d'ontologie phénoménologique.* Paris: Gallimard, 1943. Translated by H. E. Barnes as *Being and Nothingness, an Essay on Phenomenological Ontology.* New York: Philosophical Library, 1956.

Sartre, Jean-Paul. *L'existentialisme est un humanisme.* Paris: Nagel, 1946. Translated by P. Mairet as *Existentialism and Humanism.* London: Methuen, 1948.

Thomas Aquinas. *De Ente et Essentia.* Edited by M. D. Roland-Gosselin. Le Saulchoir, 1926. Translated by Armand Mauer as *Being and Essence.* Toronto, 1949.

Thomas Aquinas. *In Metaphysicorum Aristotelis Commentaria.* Edited by M. R. Cathala. Turin: Marietti, 1935.

Thomas Aquinas. *Summa Theologica.* Ottawa, 1941–1945. Translated by the Dominicans of the English Province, 3 vols. New York, 1947.

Wolff, Christian von. *Philosophia Prima Sive Ontologia.* Frankfurt and Leipzig, 1730.

SECONDARY SOURCES

Anscombe, G. E. M., and P. T. Geach. *Three Philosophers.* Ithaca, NY: Cornell University Press, 1961. Discussions of Aristotle, Thomas Aquinas, and Frege.

Ayer, A. J. "Some Aspects of Existentialism." *Rationalist Annual* (1948).

Bobik, Joseph. *Aquinas on Being and Essence, a Translation and Interpretation.* Notre Dame, IN: University of Notre Dame Press, 1965.

Owens, Joseph. *The Doctrine of Being in the Aristotelian Metaphysics.* Toronto: Pontifical Institute of Mediaeval Studies, 1951; 2nd ed., 1957.

Alasdair MacIntyre (1967)

ESSENCE AND EXISTENCE [ADDENDUM]

There was a lively and extended debate in Islamic philosophy over the relative status of essence and existence. Avicenna argued that existence is preceded by essence, in that everything that exists only comes into existence because it is brought into existence by something else, except for the ultimate existent, God, the necessary being. Many things might exist, they have essences, but unless something brings them to existence they will remain mere essences

without existence. So essence precedes existence. This view was challenged by Averroes, who argued that in an eternal universe anything that could exist would exist, and the existence of a thing is not just an attribute added onto it, but is an essential aspect of its meaning. In any case, if existence is an additional attribute, suggested Shihāb al-Dīn Yaḥyā al-Suhrawardī, then essence would have to exist before the attribute was applied for it to be an essence, and an infinite regress is started. He took this stance to show that essence precedes existence since the latter is only an idea with no reality attached to it, and it is essence that characterizes reality.

Despite Averroes's arguments, the principle that essence is the most basic concept in ontology was widely accepted in Islamic philosophy right up to the time of Mullā Ṣadrā. Mullā Ṣadrā entirely reversed al-Suhrawardī's thesis, arguing that existence is equivalent to reality. This is because existence is a necessary aspect of what it is for something to exist and so there is no regress in treating the concept as an attribute. Reality is existence, albeit manifested in a variety of different ways, and these different ways appear to one to be essences. What affects one are things that exist, and one forms ideas of essences after they impinge themselves on one, so there is no doubt that one sees here a theory in which existence precedes essence.

The significance of the debate lies in what it tells of the nature of philosophy. For Avicenna and al-Suhrawardī, philosophy is the study of the essences or ideas of things, and one then moves on to wonder whether and how far they exist. For Averroes and Mullā Ṣadrā, philosophy is a study of existing things, and as one knows more about them one knows more about their properties, but they can have no properties unless they first exist. Averroes criticized the doctrine of essentialism since it implies that something has to come from outside of something to bring it to existence, and this implies that the universe constantly requires an outside force to activate it. He saw Aristotle as arguing that the natural world consists of entities that have to have the properties they have, and that if they exist they have to exist since otherwise they would be different (i.e., nonexistent) things. Taking any other position makes the acme of Aristotelian science, the definition, vacuous, since it suggests that there are aspects of a thing (its existence) that might or might not be present, thus reducing the power of the definition.

The position that is taken on essence and existence also affects the way in which philosophy is done. An essentialist uses thought experiments in philosophy, since the imagination can rule on what notions are possible or otherwise. So Avicenna and his school accordingly used examples and potentialities to explore ideas and assess their possibility. If one's imagination cannot make sense of an idea, then that idea lacks possibility and so the state of affairs that it describes cannot exist. Those who are opposed to essentialism are critical of imagination in philosophy, since they argue that envisaging a possibility does not give one useful information about what is actually a possible existent.

See also Aristotle; Averroes; Avicenna; Islamic Philosophy; Mullā Ṣadrā; Suhrawardī, Shihāb al-Dīn Yaḥyā.

Bibliography

Leaman, Oliver. *Introduction to Classical Islamic Philosophy*. 2nd ed. New York: Cambridge University Press, 2003.

Nasr, Seyyed Hossein, and Oliver Leaman. *History of Islamic Philosophy*. 2 vols. London: Routledge, 1996.

Oliver Leaman (2005)

ESTHETICS

See *Aesthetics, History of; Aesthetics, Problems of*

ETERNAL RETURN

"Eternal return" is the doctrine that every event in the universe, in all its details and in its whole cosmic context, will recur an infinite number of times in exactly the same way that it has already occurred an infinite number of times in the past. This doctrine must be distinguished from the belief in the general periodicity of nature, according to which the main features—but not the specific details—of human and cosmic history recur.

THE PRE-SOCRATICS

The periodicity of various phenomena is a fact of daily experience; the alternation of day and night, of lunar phases and of the seasons of the year, and the rhythm of breathing and heartbeats were known to primitive people. Even the idea that cosmic history repeats itself in its general features appeared in various forms in mythological thought. Among the pre-Socratics the idea was held by Anaximander, Empedocles, and the atomists. The existing universe was regarded as a result of the differentiation of an original chaos—watery, fiery, or qualitatively undetermined—into which it would eventually return and from which a similar universe would emerge. This

idea of the periodicity of worlds soon became associated with the belief that not only the general features but also the most specific details would recur in the same order that they had occurred countless times in the past. According to Eudemus of Rhodes, this was the belief of the Pythagoreans: "Everything will eventually return in the self-same numerical order, and I shall converse with you staff in hand, and you will sit as you are sitting now, and so it will be in everything else, and it is reasonable to assume that time too will be the same" (H. Diels and W. Kranz, *Fragmente der Vorsokratiker,* 58B34).

The same idea of the cyclical nature of time was present, according to Pierre Duhem, in the thought of another Pythagorean, Archytas of Tarentum, who defined time as "the interval of the universe." The length of this cosmic cycle, called the Great Year or Perfect Number, was variously estimated by different thinkers who were influenced by Pythagoreanism. For Heraclitus it was equal to 10,800 years (according to another source 18,000 years). According to the testimony of the Stoics and of Simplicius (whose reliability on this point has been doubted by Friedrich Schleiermacher, Ferdinand Lassalle, John Burnet, and G. S. Kirk), it measured the period separating two successive conflagrations in which an old world perishes and a new one is reborn.

PLATO AND ARISTOTLE

Plato associated the period of the Great Year not with a periodically recurring cataclysm but with a return of all the celestial bodies to the same relative positions. Nor did Aristotle accept a universal conflagration, which was clearly incompatible with his idea of the incorruptible celestial realm. Nevertheless he did, if we accept his authorship of the *Problemata,* uphold eternal return in its most radical form: "Just as the course of the firmament and each of the stars is a circle, why should not also the coming into being and the decay of perishable things be of such a kind that the same things again come into being and decay?" (*The Works of Aristotle,* Vol. VII, p. 916a). Aristotle realized that the circularity of becoming would imply a relativization of succession: If the Trojan War will inevitably recur, in a sense we are living "prior" to it. The author of *Problemata,* however, was reluctant to accept the ultimate consequence of the idea of cyclical becoming: "To demand that those who are coming into being should always be numerically identical, is foolish" (ibid.).

THE STOICS

The problem of cyclical becoming was faced by the Stoics, who believed that at the end of each cosmic cycle a universal conflagration (*ἐκπύρωσέις*) that dissolves the universe into the original fire will occur. This will coincide with the beginning of another cycle; the events of the previous cycle will then be reconstituted in all their details and in the same order. But Stoics followed Aristotle by claiming that another Socrates who will marry another Xantippe and be accused by another Meletus will not be numerically identical with the previous Socrates, since numerical identity implies an uninterrupted existence. Some younger Stoics, in conceding small differences between successive Socrateses, gave up the circularity of becoming in all but name.

PLOTINUS

A curious argument for eternal return was given by Plotinus in the *Fifth Ennead* (Book VII, Chs. 1, 2). According to Plotinus, the intelligible world contains the ideal patterns not only of genera but also of all individuals, each of which successively finds its embodiment in the realm of change. But since the supply of these patterns is finite, a time will come when the same pattern—for example, of Socrates—will have to be incarnated again, and this will be possible only in the next identical cosmic cycle. Thus the successive cycles are identical, but there is no repetition within each cycle.

JEWISH AND EARLY CHRISTIAN THOUGHT

Both Judaism and Christianity, with their emphasis on the finiteness and irreversibility of cosmic history, were strongly opposed to the doctrine of eternal return. According to both the Jewish and the Christian view, the history of the world is bounded by two unique and unrepeatable events: its beginning (the Creation) and end (the Last Judgment). Every individual human life is similarly unique.

Origen, while accepting with the Neoplatonists the eternity of the world and even metempsychosis, rejected the identity of successive cosmic cycles because such a concept was incompatible with human freedom. Nemesius (*De Natura Hominis,* Ch. 38) and St. Augustine (*De Civitate Dei,* Book XII, Chs. 11, 13) rejected the doctrine, Nemesius on the ground that the Resurrection cannot take place periodically, Augustine because the incarnation of Christ occurred only once.

MEDIEVAL THOUGHT

A decree of 1277 threatening excommunication of those who accepted the Neoplatonic idea of a Great Year lasting

thirty-six thousand years demonstrates the survival of this belief into the Middle Ages. Although St. Thomas Aquinas rejected the cyclical view of time by claiming that the re-creation of numerically identical individuals would be contradictory, and as such was beyond even God's power, his view was not shared by John Duns Scotus and William of Ockham. Nicolas Bonet and François de la Marche explicitly insisted on God's power to restore any past motion, and therefore a corresponding past interval of time, since there was no difference between motion and time.

EARLY MODERN THOUGHT

Thus, at the threshold of the modern era two of the central ideas of the modern cyclical view of time were present—the reversibility of motion and the relational theory of time. The third essential ingredient of the cyclical theory—the finiteness of the material universe—was excluded by Giordano Bruno's vision of innumerable worlds and limitless space. This may explain the absence of the idea of eternal recurrence in Bruno's contemporaries despite their Neoplatonic leanings. For if the number of constituent parts of the universe is infinite, the number of possible combinations is also infinite, and recurrence of the same configuration is not inevitable.

In Isaac Newton and his successors there was an additional motive for not considering the cyclical view. They regarded time as absolute, as intrinsically irreversible, irrespective of its content. Even a complete restoration of the content of the past moment would not make this moment itself present.

René Descartes came very close to the cyclical view when he wrote that matter must successively pass through all its possible forms, but since matter to him was coextensive with infinite space, the number of its configurations was inexhaustible. Furthermore, the pagan and astrological associations of the ancient cyclical theory made it thoroughly suspect.

NINETEENTH-CENTURY VIEWS

Interest in eternal return was revived only with the development of modern cosmogony. The nebular hypothesis of Immanuel Kant (1755) and Pierre Simon de Laplace (1796) implicitly raised the question of the origin of any primordial nebula: Did it represent a truly initial stage preceded by an act of supernatural creation, or was it merely one of the countless stages in an unending cycle of successive worlds? The principle of the uniformity of nature in time, anticipated by Bruno's and Benedict (Baruch) de Spinoza's belief in the eternity of the universe, strongly favored the second answer. Although the law of entropy suggested the irreversibility of the whole cosmic process, because of its statistical character it did not exclude the general periodicity of nature. Various hypothetical mechanisms were invented to provide a "rewinding of the cosmic clock," at least on a local scale. The most popular one was that of cosmic clashes by which two stellar masses that had lost their heat could be transformed into another nebula, which would then develop into another world "ever the same in principle, but never the same in concrete results," as Herbert Spencer wrote in his *First Principles* (p. 550).

Such a new world could be the same even in concrete details only if the cosmic mass did not contain an infinite number of units. Eugen Dühring, in various writings (heavily annotated copies of which were found in Friedrich Nietzsche's library), rejected the concept of actual infinity as self-contradictory and inapplicable to the physical world.

In Nietzsche's thought the concept of a finite universe and of the discrete structure of matter implied a finite number of possible successive configurations, and therefore an inevitable recurrence of a configuration defining a state of the universe that had already occurred an infinite number of times in the past; and this recurring cosmic state must, according to the then accepted deterministic scheme, generate the series of the same events in the same order as in the previous cosmic cycles. This view, formulated by Nietzsche at the end of the fourth book of *Fröhliche Wissenschaft* (1881), became central to his philosophy. The intensely lyrical way in which this view was expressed in *Thus Spake Zarathustra* hid its intellectual origins, which are far more obvious in the posthumously published fragments of *The Will to Power* (see *The Complete Works of Friedrich Nietzsche,* Vol. XV, Ch. 2, esp. p. 430). Prior to Nietzsche only a few nineteenth-century thinkers held the same view: Louis A. Blanqui (*Éternité par les astres,* 1871), Gustave Le Bon (*L'homme et les sociétés,* 1881), Jean Marie Guyau (*Vers d'un Philosophie,* 1881). It was not held, however, by Dühring, who claimed that the continuity of space admitted an infinite number of configurations even if the number of atomic units was finite (*Cursus der Philosophie,* pp. 84–85). The same objection against eternal return was raised by Alois Riehl and Alfred Fouillée; against this view Franz Selety pointed out that concrete processes were discrete and not mathematically continuous, and therefore, he claimed, eternal recurrence was unavoidable; and G. N. Lewis claimed that the attempt at avoiding an exact recurrence by assuming a whole con-

tinuum of possible values is eliminated by the quantum theory.

Henri Poincaré, although he formulated the theorem of phases, according to which any mechanical system of a finite number of particles will in a sufficiently long time pass through a configuration infinitely close to the previous one, nevertheless dismissed the application of the theorem of phases to cosmogony in his *Leçons sur les hypothèses cosmogoniques* (p. 23) as "the dream of eternal return." C. S. Peirce (*Collected Papers,* Vol. I, pp. 498–500) held the cyclical view on the unusual ground that "since every portion of time is bounded by two instants, there must be a connection of time ringwise." Furthermore, this view was entirely incompatible with the rest of his philosophy. The arguments of Abel Rey in favor of the cyclical view were not essentially different from those of Nietzsche, since they were based on the classical corpuscular-kinetic scheme of nature.

CONTEMPORARY THOUGHT

The contemporary crisis of the classical scheme of nature makes the doctrine of eternal return extremely questionable. The doctrine was based on four fundamental assumptions: (*a*) that the universe is made up of distinct atomic units that persist through time without any intrinsic change, so that they may be identified in successive moments; (*b*) that the number of atomic units is finite; (*c*) that it is meaningful to speak of a definite "state of the universe" at each instant; (*d*) that one such particular state (embodied in a definite atomic configuration) causally determines all future states (Laplacean determinism).

Except for the thesis that the size of the universe is finite, which is favored by some cosmologists, none of these theses remains unchallenged by the recent developments in physics. The atoms of modern physics do not have the rigidity and permanence of classical atoms; and without permanent elements there can be no recurring configurations. The ontological status of "state of the universe at an instant" is challenged by the relativization of simultaneity, and the validity of rigorous determinism has been seriously questioned since the formulation of the indeterminacy principle in 1927.

Moreover, there are ambiguities and discrepancies within the theory of eternal return itself. The assumption of a completely identical repetition of cosmic situations makes the theory intrinsically unverifiable. Moreover, either the successive identical cycles are distinguished by their positions in time—which means that we surreptitiously introduce an irreversible time as their container—or we insist on the numerical identity of the cycles. But we then have only one cosmic cycle, and it clearly becomes meaningless to speak of a "succession of cycles" or of their "repetition." Although it is self-contradictory to speak of numerical identity of genuinely successive events, the two views have often been held jointly, as by the Scotists and Nietzsche.

The eternal return is rejected by all thinkers who insist on the irreversibility of becoming, genuine novelty, and the immortality of the past. Mircea Eliade regarded the theory as a manifestation of "ontology uncontaminated by time and becoming" (*The Myth of the Eternal Return,* p. 89); Émile Myerson saw in it an attempt to eliminate becoming (*L'identité et réalité,* Ch. 8). The emotional effect of the doctrine is equally ambiguous. Thus Nietzsche's mystical ecstasy over "the ring of eternity" was tinged by a note of anxiety and even despair. Gustave Le Bon compared the recurring cosmic cycles to the labors of Sisyphus, and Miguel de Unamuno, in *The Tragic Sense of Life,* regarded the doctrine as a poor substitute for personal immortality.

See also Anaximander; Aristotle; Atomism; Augustine, St.; Bruno, Giordano; Descartes, René; Dühring, Eugen Karl; Duns Scotus, John; Fouillée, Alfred; Heraclitus of Ephesus; Kant, Immanuel; Laplace, Pierre Simon de; Lassalle, Ferdinand; Nemesius of Emesa; Neoplatonism; Newton, Isaac; Nietzsche, Friedrich; Origen; Peirce, Charles Sanders; Plato; Plotinus; Poincaré, Jules Henri; Pre-Socratic Philosophy; Pythagoras and Pythagoreanism; Riehl, Alois; Schleiermacher, Friedrich Daniel Ernst; Simplicius; Spinoza, Benedict (Baruch) de; Stoicism; Unamuno y Jugo, Miguel de; William of Ockham.

Bibliography

CLASSICAL AND MEDIEVAL

Diels, H., and W. Kranz. *Fragmente der Vorsokratiker.* 6th ed. Berlin: Weidmann, 1951.

Duhem, Pierre. *Le système du monde.* Paris: A. Hermann, 1954. Vol. I, pp. 65–85, 164–169, 275–296; Vol. II, pp. 447–461; Vol. VII, pp. 441–461. On the idea of the Great Year in Greek and medieval thought.

Mugler, Charles. *Deux thèmes de la cosmologie grecque: Devenir cyclique et la pluralité des mondes.* Paris: Librarie C. Klincksieck, 1953.

Ross, W. D., ed. *The Works of Aristotle.* Oxford, 1923. Vol. VII.

Sambursky, Samuel. *The Physical World of the Greeks.* Translated by Merton Dagut. New York, 1956. Ch. 8. On Greek cosmogony.

NIETZSCHE ON ETERNAL RETURN

The Complete Works of Friedrich Nietzsche. Translated by A. M. Ludovici. Reissued, New York, 1964. Vol. XV, Ch. 2, esp. p. 430.

Andler, Charles. *Nietzsche, sa vie et sa pensée.* 5th ed. Paris: Gallimard, 1958. Vol. II, pp. 402–424; Vol. III, pp. 284–294. Vol. II discusses mythological, Greek, and modern influences on Nietzsche.

Batault, Georges. "L'hypothèse du retour éternel devant la science moderne." *Revue philosophique* 57 (1904): 158–167.

Bois, Henri. "Le 'retour eternal' de Nietzsche." *Année philosophique* 24 (1913): 145–184.

Danto, Arthur. *Nietzsche as Philosopher.* New York: Macmillan, 1965. Ch. 7.

Heidegger, Martin. *Nietzsche.* Pfullingen, Germany: Neske, 1961. Vol. I, Ch. 2; Vol. II, Ch. 3.

Horneffer, Ernst. *Nietzsches Lehre von der ewigen Wiederkunft und der bisherige Veröffentlichung.* Leipzig, 1900.

Löwith, Karl. *Nietzsches Philosophie der ewigen Wiederkunft des Gleichen.* Stuttgart, 1956.

Schlechta, Karl. *Nietzsches grosser Mittag.* Frankfurt, 1954.

Stambaugh, Joan. *Untersuchungen zum Problem der Zeit bei Nietzsche.* Haag: Nijhoff, 1959.

Steiner, Rudolf. *Friedrich Nietzsche. Ein Kämpfer gegen seine Zeit,* 3rd ed. Dornach, 1963. Pp. 187–197. On the relation of Nietzsche to Dühring.

GENERAL

Čapek, Milič. "The Theory of Eternal Recurrence in Modern Philosophy of Science." *Journal of Philosophy* 57 (1960): 289–296.

Couturat, Louis. "De l'évolutionnisme physique et du principe de la conservation de l'énergie." *Revue de métaphysique et de morale* 1 (1893): 564–572.

Denbigh, K. G. "Thermodynamics and the Subjective Sense of Time." *British Journal for Philosophy of Science* 4 (1953–1954): esp. p. 187.

Dühring, Eugen. *Cursus der Philosophie.* Leipzig: R. Reisland, 1876.

Eliade, Mircea. *The Myth of the Eternal Return.* New York: Pantheon, 1955.

Lewis, G. N. "Symmetry of Time in Physics." *Science* 31 (1930): 569–577.

Marbe, Karl. *Die Gleichförmigkeit der Welt.* Munich, 1916. Chs. 8–11.

Meyerson, Émile. *L'identité et réalité.* 5th ed. Paris, 1951.

Peirce, C. S. *Collected Papers.* Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks. 8 vols. Cambridge, MA: Harvard University Press, 1931–1958.

Poincaré, Henri. *Leçons sur les hypothèses cosmogoniques.* Paris: A. Hermann, 1911.

Rey, Abel. *Le retour éternel et la philosophie de la physique.* Paris: Flammarion, 1927.

Riehl, Alois. *Zur Einführung in die Philosophie der Gegenwart.* Leipzig: Teubner, 1903.

Selety, Franz. "Über Wiederholung des Gleichen im kosmischen Geschehen." *Zeitschrift für Philosophie* 5 (1914).

Spencer, Herbert. *First Principles.* 4th ed. New York: Appleton, 1896.

Weber, Louis. "L'évolutionnisme physique." *Revue de métaphysique et de morale* 1 (1893): 425–452.

Whitrow, G. J. *The Natural Philosophy of Time.* London: Nelson, 1961. Ch. titled "Cyclic Time," esp. pp. 40–41.

Zawirski, Zygmunt. *L'évolution de l'idée du temps.* Cracow, 1936. Pp. 336ff.

Milič Čapek (1967)

ETERNITY

The word *eternal* is derived from the Latin *aeternus,* a contraction of *aeviternus,* which, in turn, is derived from *aevum,* a word from the same root as the English words *ever* and *aye.* In Greek the corresponding adjectives are even more obviously connected with the notion of everlasting existence. This is the original sense of the word *eternal* and probably also the sense that is still the most common in ordinary language. But in certain philosophical contexts the notion of everlasting existence is expressed rather by "sempiternal," "eternal" being reserved for the sense of "timeless."

THE "TIMELESS PRESENT" IN SCIENCE

In English and other Indo-European languages there is a usage described by grammarians as the timeless present. When, for example, we say, "Seven is a prime number," we do not intend our use of the present tense to convey anything about the present as distinct from the past or the future. For this reason we find something very curious in the sentences "Seven was a prime number" and "Seven will be a prime number." Existential statements of a mathematical kind do not refer to the time of speaking. An assertion such as "There is a prime number between 5 and 10" could never be countered sensibly by the remark "You are out of date." For this reason the entities discussed in mathematics can properly be said to have a timeless existence. To say only that they have a sempiternal or omnitemporal existence (that is, an existence at all times) would be unsatisfactory, because this way of talking might suggest that it is conceivable they should at some time cease to exist, an absurdity we want to exclude.

Mathematics, however, is not the only study in which use of the timeless present is appropriate. The same idiom can be found in all studies that are concerned with necessary truths as distinct from matters of fact. It may occur, for example, even in empirical studies when the propositions we formulate involve the notion of necessary connection. Thus, we say "The hydrogen atom contains only one proton" because we do not wish to allow that hydrogen atoms may in the past have contained or may in the future come to contain more than one proton.

Here, however, our use of the timeless present is certainly not intended to suggest that hydrogen atoms exist out of time. What we wish to call timeless is simply the connection between being a hydrogen atom and containing a single proton. Sometimes such connections have been called eternal verities, most commonly when it has been thought they could be known a priori, as in mathematics.

GREEK THOUGHT

A different conception of timelessness appears in Parmenides' poem "The Way of Truth," where he says of the One, "It neither was at any time nor will be, since it is now all at once [*ὁμοῦ πᾶν*] a single whole." Since Parmenides and his pupil Zeno argued against the reality of change, we must suppose that this remark does not represent the One as existing merely for a moment but says rather that the One cannot be described in a language that employs tenses. The One exists all at once because it involves no temporal succession of earlier and later. But why should anyone talk in this way? Perhaps Parmenides accepted the religious teaching of Xenophanes, that the Whole is an everlasting god, and tried to defend it against Heraclitus's doctrine of universal flux by maintaining that the Whole is spherical in all respects—that is, temporally as well as spatially. For such a Whole could not itself be in time, and if we talk about it at all we must employ the timeless present. This is only a guess, but there is evidence to show that a conception of cyclical time order was current in the Pythagorean school with which Parmenides is said to have been connected in his youth.

From Parmenides the notion of a mode of existence that allows no distinction between past, present, and future passed to Plato, who applied it to his Forms, or Ideas. The most influential passage of his works dealing with this subject is in the *Timaeus* (37E6–38A6), where he contrasts the created world with the eternal living being, its timeless archetype.

The language of the passage is poetical, and it seems that we are not expected to take all the details seriously. In particular, Plato can scarcely have meant us to believe that time was an afterthought of the creator. Rather, we are to understand that time is to the perceptual world of becoming what eternity is to the intelligible world of being. For Plato said later (*Timaeus* 38B5), "Time was created with the heaven," and he seems to have held that it is identical with the movements of the heavenly bodies, which are commonly said to measure its passage. In many of his works Plato glorified the eternal and spoke of the temporal as something inferior, but he did not, like Parmenides and Zeno, deny the reality of time. The most he said in this regard is that temporal things never have being but are always in a state of becoming, as Heraclitus had argued. However, this seems to be no more than a recognition that we cannot talk of temporal things in the timeless present as we talk of Forms and mathematical objects.

The connection with necessity that Plato had claimed for timeless eternity Aristotle claimed for sempiternity. For having rejected Plato's doctrine of the creation of time (*Physics* 251b14), he did not wish to say that anything was wholly severed from it. Thus, in one place he explicitly associated the objects of mathematics with the universe, which he certainly did not regard as timeless (*Nicomachean Ethics* 1112a22). In his view the objects of mathematics are eternal (*ἄϊδια*) but only in the sense that they exist always—that is, are sempiternal. He held that among sempiternal things there is no difference between possibility and actuality and also that there is nothing merely accidental (*Physics* 203b30, 196b10). In one place he even said that sempiternal things, insofar as they are sempiternal, are not in time, because they are not bounded by time or subject to aging and the other conditions of time (*Physics* 221b30). Apparently he had in mind not only mathematical objects, such as numbers, but also God and the sun and stars and the whole heaven. For he said elsewhere that the heavenly bodies, unlike perishable things, are not wearied by their motion. The sun is active forever, and there is no danger that it will give out, as some philosophers feared (*Metaphysics* 1050b24). Benedict (Baruch) de Spinoza may have been influenced, even if only indirectly, by Aristotle's doctrine when he used the word *aeternitas* to signify both necessity and sempiternity.

CHRISTIAN THEOLOGY

The doctrine of Plato's *Timaeus* passed into Christian theology, with emphasis on the notion of timeless *life.* As early as the St. John Gospel (8:58) there is a curious passage in which Jesus is represented as saying, "Before Abraham was I am." But it is fairly clear that the author of this gospel knew something of Greek philosophy, possibly at second hand through the works of the Jewish theologian Philo of Alexandria, and also that his narrative is no mere historical record of the life of Jesus. By the end of the fifth century there was nothing at all strange in the use of Platonic thought for the exposition of Christianity, and St. Augustine, when commenting, in his *De Civitate Dei* (xi, 21), on the sentence in Genesis "God saw that it was good," referred to the passage of the *Timaeus* cited above. In his *Confessions* (xi, 13) he wrote also of God's "ever-

present eternity" and said that for God "all years stand at once" (*omnes simul stant*). A century later Boethius, in his *De Trinitate* (4), said that our "now," by running as it were (*quasi currens*), makes time and sempiternity, whereas the divine "now," by abiding, unmoved and immovable, makes eternity; in the final chapter of his *De Consolatione Philosophiae* he discussed this at greater length.

> Eternity is the complete possession of eternal life all at once—a notion that becomes clearer from comparison with things temporal. For whatever lives in time moves as something present from the past to the future, and there is nothing placed in time that can embrace the whole extent of its life at once. It does not yet grasp tomorrow, and it has already lost yesterday. And even in the life of today you do not live longer than in the transitory moment. That, then, which is subject to the condition of time, even if (as Aristotle thought of the world) it has no beginning or end and its life extends through endless time, is still not such as may be rightly judged eternal. For though its life be endless, it does not grasp and embrace the extent of it all at once [*totum simul*] but has some parts still to come. … And so, if, following Plato, we wish to give things their right names, let us say that God is eternal, but the world everlasting.

All these notions reappeared in the Middle Ages. St. Thomas Aquinas, for example, quoting Boethius as his authority, said in his *Summa Theologiae* (I, x, 1) that there are two marks of eternity, namely, that the eternal has neither beginning nor end and that eternity contains no succession, being all at once (*tota simul existens*). This last phrase, though Thomas could scarcely have known as much, is a rendering of words Parmenides had used over seventeen centuries earlier. Not content, however, with the old distinction of time and eternity, medieval theologians sometimes spoke of *aevum*—that is, everlastingness—as something intermediate that was appropriate to the heavens and to angels. This was conceived by some as having a beginning but no end and by others (probably influenced in part by Aristotle's account of God and the heavenly bodies) as possessing earlier and later, but without innovation and aging. Thomas regarded the latter view as self-contradictory, since, he held, there could be no succession without aging. *Aevum* does not necessarily include earlier and later, according to Thomas, though these can be joined with it, as is the case with angels, who have changeless being as well as the capacity of change according to choice (*Summa Theologiae* I, x, 5).

CRITICISM OF THE THEOLOGICAL USE

Anyone who, like Boethius, speaks of eternity as "the complete possession of eternal life all at once seems to be running together two incompatible notions, that of timelessness and that of life. For we can attach no meaning to the word *life* unless we are allowed to suppose that what has life acts. No doubt the word *acts* may be taken in a wide sense. Perhaps it is not essential that a living thing produce changes in the physical world. But life must at least involve some incidents in time, and if, like Boethius, we suppose the life in question to be intelligent, then it must also involve awareness of the passage of time. Purposeful action is action with thought of what will come about after its beginning. It is difficult to decide how much of this Plato was prepared to admit when he wrote the *Timaeus.* In his earlier works (for example, the *Meno* and the *Phaedo*) he tried to explain the possibility of a priori knowledge by a doctrine of reminiscence, which involves the hypothesis that before this life the human soul lived among timelessly existent Forms and contemplated them directly as in this life it sees things belonging to the realm of becoming. But he probably came to realize that there is something absurd in the suggestion that a soul can pass part of its time in a timeless realm and then at a certain date enter the temporal realm, for he appears to have dropped the doctrine of reminiscence in his later dialogues, where instead of glorifying the soul by treating it as something akin to the timeless Forms he praised it as the source of motion.

In the *Republic,* which belongs to the middle of his literary life, he spoke of God, who is presumably alive, as having created one and only one Form of each kind. But the wording of the passage (*Republic* 597C) seems to be obviously playful, and it is unlikely that Plato ever meant to suggest seriously that the Forms had been created. In the *Timaeus,* as we have seen, the Forms are said to be the timeless model used by the demiurge, or craftsman, who made the temporal world. Yet this same timeless model is said to be itself alive (*Timaeus* 37E6). Is this to be taken seriously? Unlike medieval theologians, for whom things predicated of the eternal were to be interpreted analogically, Plato maintained (*Timaeus* 29B) that discourse about the eternal is to be understood in the strict and primary sense of the words it employs.

How did the theologians come to commit themselves to talk about timeless life? The influence of Plato's style counts for a lot, but not for everything. One might say in the case of Thomas that the surprising thing is that he held to Plato's account of time and eternity though he

must have known it had been criticized by Aristotle. Probably the explanation is to be found in a peculiarity of Christian doctrine. Aristotle, though a theist of a sort, not only rejected the Platonic notion of the creation of time but also maintained the sempiternity of the heavens. To a theologian who had to produce a metaphysical scheme concordant with biblical revelation (which denied the sempiternity of the cosmos) this must have made Aristotle's criticism of Plato's doctrine of eternity seem unsatisfactory. But apart from that, Plato's doctrine had the positive merit of seeming to provide for the necessity of God's existence. If it is correct, once we have admitted that we understand the meaning of the word *God* and that it involves no inconsistency, we cannot sensibly deny that God exists. Another manifestation of this theological interest in the necessity of God's existence is Anselm's ontological argument. Admittedly, this was rejected by Thomas, but for epistemological reasons concerned with the limits of our capacities, not for the assumption it involves that divinity, by definition, entails existence. On the contrary, Thomas, following Boethius, said that God's essence and existence are one.

OTHER PHILOSOPHICAL USES

In later European thought Spinoza and various idealist philosophers used the word *eternity* to describe the existence of their God or Absolute. Spinoza, for example, said in his *Ethics* (I, Definition viii), "By *eternity* I mean existence itself in so far as it is conceived necessarily to follow solely from the definition of that which is eternal. Existence of this kind is conceived as an eternal truth, like the essence of a thing, and therefore cannot be explained by means of duration or time, though duration may be conceived without a beginning or an end." Here there is no longer any verbal connection of eternity with life, but there is still a wish to maintain that something concrete exists with the timeless necessity of which we speak in mathematics. Similar assertions have been made in Indian philosophy, which does not in any way derive from Parmenides or Plato, and we must therefore suppose that they correspond to a widespread demand of religious thought. In modern times even the Pythagorean notion of a cyclical time order has again been considered seriously, by Kurt Gödel.

See also Aristotle; Augustine, St.; Boethius, Anicius Manlius Severinus; Gödel, Kurt; Heraclitus of Ephesus; Indian Philosophy; Parmenides of Elea; Plato; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Time; Time, Consciousness of; Xenophanes of Colophon; Zeno of Elea.

Bibliography

Brabant, F. H. *Time and Eternity in Christian Thought.* New York: Longmans, Green, 1937.

Cushman, R. E. "Greek and Christian Views of Time." *Journal of Religion* (1953): 254–265.

Festugière, A.-J. "Le sens philosophique du mot *αἰών*." *La parola del passato.* Fasc. 11 (1949): 172–189.

Gödel, Kurt, "A Remark about the Relationship between Relativity Theory and Idealistic Philosophy." In *Albert Einstein: Philosopher-Scientist,* edited by P. A. Schilpp. New York: Tudor, 1951.

Guitton, J. *Le temps et l'éternité chez Plotin et chez Saint Augustin.* Paris: Boivin, 1933; 2nd ed., 1956.

Guthrie, W. K. C. "Time and the Unlimited." In *A History of Greek Philosophy.* Vol. I. Cambridge, U.K.: Cambridge University Press, 1962. Pp. 336–340.

Haldane, J. B. S. "Time and Eternity." *Rationalist Annual* (1946): 33–38.

Kneale, W. C. "Time and Eternity in Theology." *PAS* 61 (1960–1961): 87–108.

Mondolfo, Rodolfo. *L'infinito nel pensiero dei Greci.* Florence: F. Le Monnier, 1934. 2nd enl. ed. published as *L'infinito nel pensiero dell'antichità classica.* Florence, 1956.

Wolfson, H. A. "Duration, Time, and Eternity." In *The Philosophy of Spinoza,* Book I. Cambridge, MA: Harvard University Press, 1934. Ch. 10.

William C. Kneale (1967)

ETERNITY [ADDENDUM 1]

Since Kneale wrote his article, many writers have argued against divine timelessness by claiming that it is inconsistent with divine omniscience. If God knows everything, they reason, he knows what time it is now. But the token of "now" in (say) "it is noon now" refers to the time at which the speaker speaks. So if one knows that it is noon now, and one knows this only if one is able to assert it truly, one exists now (see Stump and Kretzmann 1981). A variant has it that if God is always omniscient, at noon he believes that it is noon and not 1 PM, and at 1 PM that it is 1 PM and not noon. Plausibly, what a person believes is an intrinsic fact about that person. But a timeless being cannot change intrinsically: What changes intrinsically has an intrinsic property at one time that it lacks at another. Some would reply here that one can only know what is true, and if God is timeless, for him it is not noon, or any other time: it is noon for God only if some part of his life which is located at noon is temporally present, and if God is timeless, no part of his life is located at noon or is temporally present.

This raises, of course, the question of how it can be noon for us but not God. And this question leads to an argument others (e.g. Craig 2002) press, that divine time-

lessness is inconsistent with our ordinary view of time. We ordinarily think that the present exists and the future does not. But a timeless being's life has no future part: lives with future parts are ipso facto lives in time. If your death is still to come for a timeless being, your death lies in a future part of its life. So for a timeless being, your death is not still to come. But for you, it is still to come. Events yet to come for us must not be yet to come for a timeless being: they must already be there, in some way. And so it seems to follow that the future as well as the present exists: If all of time later than noon exists for God, it exists, period, even for us. Either divine timelessness entails the reality of the future, counter to our usual way of thinking about time, or else if God is timeless, some parts of time are real relative to some persons but not others: For us, it is (say) noon, as all of time that is later than noon does not exist, but for God, all of time later than noon exists, and so it is not noon.

Stump and Kretzmann (1981) argued (contra Kneale) that talk of timeless action makes sense and that the move to make here has been known since Aquinas. God's acts consist of atemporal intendings plus effects of these that occur at particular times. God's contribution to the acts is not located in time, but nonetheless his life can "involve some incidents in time" (Kneale). Purposeful action involves "thought of what will come about after its beginning" (Kneale), but "after" can have a sense involving causal as well as temporal precedence and can also refer to temporal effects of an atemporal intention that occur after other such effects. Stump and Kretzmann also claim that an eternal being's life endures in its own way: "timeless duration" is no contradiction, and neither is "timelessly present." Further, they argue, events in an eternal life are in a sense simultaneous with events in time: "simultaneous" does not always mean "at the same time"; in the eternal-temporal case it has a complex sense involving the coexistence of eternal and temporal presents.

Some might see Stump and Kretzmann as working out a hybrid doctrine of divine eternity, one involving neither sheer timelessness nor ordinary temporality but presentness and duration without the full range of temporal features. Such "intermediate" theories have multiplied. Craig (2002) suggests that God is timeless "before" creating and became temporal by creating. It is hard not to think this view contradictory: being over seems a paradigmatic temporal property, one only temporal events have, yet according to Craig the timeless phase of God's life is over. Padgett (1992) argues that God is "relatively timeless"—that is, in some but not all ways in time. As he sees it, points in our time are points in the time God exists in, but length relations between these points do not exist for God. For us, noon is an hour before 1 p.m. but for God there is no definite length of time between the time we call noon and the time we call 1 p.m. Swinburne (1993) suggests that, whereas God exists in a time without definite length before he creates, once he creates, the lengths of time that exist for us exist for God.

See also Time, Consciousness of.

Bibliography

Craig, William. *God, Time and Eternity*. Dordrecht: D. Reidel, 2002.

Leftow, Brian. *Time and Eternity*. Ithaca, NY: Cornell University Press, 1991.

Padgett, Alan. *God, Eternity and the Nature of Time*. New York: Macmillan, 1992.

Stump, Eleonore, and Norman Kretzmann. "Eternity." *Journal of Philosophy* 79 (1981): 429–458.

Swinburne, Richard. "God and Time." In *Reasoned Faith*, edited by Eleonore Stump, 204–222. Ithaca, NY: Cornell University Press, 1993.

Brian Leftow (2005)

ETERNITY [ADDENDUM 2]

Islamic and Jewish philosophy emerged from a Hellenistic climate in which the universe was taken to be an everlasting emanation from a unitary source (Plotinus), so the debate that ensued among thinkers in these two traditions had to reconcile this philosophical conviction with the pronouncement of their respective revelational books: "In the beginning God created heaven and earth" (Genesis 1.1) and "God said 'be' and it is" (Qurʾan 2:117). An absolute beginning linked to an initial moment of time is conflated here with the freedom of the creator to create. Plotinus never denied emanation to be free, although that freedom appropriate to the One would be vastly different from creatures: not being faced with anything—including alternatives—freedom in the One (so far as humans can grasp it) would be more like pure assent.

Al-Fārābī (d. 950), the first of the Muslim philosophers to elaborate this subject, introduced *necessity* into the founding emanation by modeling it on logical deduction: everything that is derives from a single premise. Ibn Sīnā (979–1037), "Avicenna" to Europeans and North Americans, refined this scheme to align it with the "wandering" heavenly bodies (planets)—identified as succes-

sive spheres—to create a philosophical cosmology to articulate, after a fashion, the transition from one to many. On the Muslim side, al-Ghazālī (1058–1111) countered this necessary emanation (proposed as an adaptation of the Qur'an) with the charge of *unbelief* (*kufr*), whereas Moses Maimonides (d. 1204), the Jewish thinker imbued with Islamic philosophy, took these charges and elaborated them into a strict division of creator from creatures in order to safeguard the freedom and transcendence of the creator from creation.

Thus the crucial distinction between *everlasting* and *eternal* emerges: while what always was ("everlasting") might not have been, it is impossible for the One source of all that is not to be, so the One must be said to be "eternal." By connecting *eternal* with origins and with an adequate distinction of creator from creatures (which al-Farabi's logical model failed to do), these thinkers could affirm the creator God's eternity without further exploring the issue. Timelessness will be a feature of the eternal One because time itself is created; but eternity will comprise more than timelessness (which could also be said of mathematical entities) because its reality explains the existence of the universe. This discussion can be distinguished from the question of whether the origin of the universe requires an initial moment of time; the creator God would have to be eternal even if some creatures were everlasting. One can detect a Neoplatonic concern for the origin of all things in a unitary source, here adapted to a free creator whose transcendence can be removed from any hint of anthropomorphism by the assertion of *necessary existence* (Ibn Sīnā), and by distinguishing even everlasting things that are created from their uncreated *eternal* source.

Bibliography

Burrell, David. *Knowing the Unknowable God: Ibn-Sina, Maimonides, Aquinas*. Notre Dame, IN: University of Notre Dame Press, 1986.

Frank, Daniel H., and Oliver Leaman, eds. *History of Jewish Philosophy*. New York: Routledge, 1997.

Gerson, Lloyd. *God and Greek Philosophy : Studies in the Early History of Natural Theology*. New York: Routledge, 1990.

Nasr, Seyyed Hossein, and Oliver Leaman, eds. *History of Islamic Philosophy*. New York: Routledge, 1996.

David Burrell (2005)

ETHICAL EGOISM

Generally defined as the view that one ought to do whatever and only whatever is in one's own maximum interest, benefit, advantage, or good, "ethical egoism" contrasts with (1) psychological egoism, which says that people do in fact, perhaps necessarily, act in that way; and from (2) alternative ethical theories, which claim that we have other fundamental obligations such as to act for the sake of others, even at ultimate cost to ourselves, or in ways having no necessary relation to anyone's benefit.

Egoism strikes many as cutting through pretenses and getting down to fundamentals. This appearance soon dissipates when we make essential distinctions. Foremost is that due to the classic work of Bishop Joseph Butler (1692–1752). Is "self-interest" in that theory to be understood as one's interest in certain states unique to one's own self—as distinct from certain states of other people? Or is it merely interests of one's own self—the interests one happens to have, whatever they may be? Since action is necessarily motivated by interests of the agent motivated by them, the second interpretation is trivial: Whatever we do, we are somehow interested in doing it. But the first interpretation is implausible: People are notoriously capable of sacrificing themselves—for friends, loved ones, or causes.

Ethical egoism would also be vacuous if it said only that whatever we ought to do, we ought to do it only if we are motivated to do it. Only when self-interest is construed in the narrow sense, as describing certain of our interests—those focused specifically on oneself—but not others, does it make sense to say that we ought to act self-interestedly. Then the question "Why?" arises, for we have our choice.

This brings up the question of what is the ultimate good or interest of an agent. Alas, we must leave this important issue open in the present discussion. The next question, however, is crucial. What is meant by *ethical*? Here we must distinguish between a wide sense in which *ethical* means something like "rational" and a narrower sense in which specifically moral requirements are intended. I should choose Bordeaux 1989, but that isn't a moral matter; that I should refrain from cheating is.

If ethical egoism is understood in the wider sense, it is a theory about rational behavior; and construing self-interest in Butler's second way, egoism says that a rational agent acts so as to maximize the realization of whatever she or he is interested in attaining. This highly plausible idea is noncommittal about the content of our interests.

Now turn to the moral version. Moral rules call upon us all to do or refrain from certain things, whether we like it or not. Can there be a rational egoistic morality, then?

But the interests of different persons can conflict. This leads to a problem, which becomes clear when we distinguish two possible interpretations of moral egoism:

(1) "First-person" egoism appraises all actions of all persons on the basis of the interests of the propounder alone. What Jim Jones thinks, if he is this kind of an egoist, is not only that Jim Jones ought to do whatever, and only whatever, conduces to Jim Jones's best interests—but that everyone else should, too. This is consistent, to be sure, but from the point of view of anyone except Jim or his devotees, it is evidently irrational, if they too are self-interested.

(2) "General" egoism, on the other hand, says that each person ought to do whatever is in that person's interests. If Jim is an egoist of this type, he believes that Jim ought to do whatever is in Jim's interests, but Sheila ought to do whatever is in Sheila's interests, and so on.

Serious conceptual problems arise with general egoism. Suppose that Jim's interests conflict with Sheila's: Realizing his frustrates hers. Does Jim tell Sheila that it is Sheila's duty to do what is in Sheila's interests? Or what is in Jim's interests? Or both? Every answer is unacceptable! The first is unacceptable to Jim himself: How can he, as an exclusively self-interested person, support actions of Sheila's that are detrimental to himself? The second is unacceptable to Sheila: If she is exclusively self-interested, why would she take Jim's "advice"? And the third is flatly inconsistent: For their interests to "conflict" means that they cannot both do what is in their own best interest.

A standard reply is to hold that egoism tells each of the differing parties merely to try to do what is in their interests. But this is either just wrong or turns the theory into something else: "Here, all you ought to do is try to bring about your best interests—but it doesn't matter whether you succeed!" But self-interested agents are interested in results.

Or it might be held that the good life consists not in succeeding but in striving. This turns egoism into a game, and in conflict situations, a competitive game. And games are interesting, but also very special, requiring players to abide by certain game-defining rules. True chess-players do not cheat, even if they can—cheating is not really playing the game. They want opponents to do their best, even if they themselves lose. Of course, they prefer to win, but even if they do not, the game is worthwhile. This defense lacks generality. Ethical egoism is not about games, it is about life. Some people may make life into a game, but most people do not. They want results, not just effort; in conflicts, they are not about to cheer for the other side.

So egoism seems to be self-defeating. What to do? The answer requires, first, that we utilize the vital distinction between egoism as (1) a theory of rationality—of what is recommended by reason; and (2) as a theory of morality. The latter is interpersonal, and concerns rules for groups. Such rules require that people sometimes curtail their passions and conform to the rules.

If we view egoism as a theory of rationality, then whether agent *A* should aim only at bringing about certain states of *A* is an open question. But that *A* should aim at bringing about only those states of affairs that *A* values is not: We can act only on our own values—in acting, we make them our own.

But when we turn to the subject of formulating specifically moral principles, we must attend to the facts of social life. From the point of view of any rational individual, moralities are devices for securing desirable results not attainable without the cooperation of others. To do this, mutual restrictions must be accepted by all concerned. They will be accepted only if they conduce to the agent's interests. Therefore, moral principles, if rational, must be conducive to the interests of all, those to whom they are addressed as well as those of the propounder herself. Thus, egoism leads to contractarianism: Moral principles are those acceptable to each person, given that person's own interests, if all comply. Undoubtedly, some will not; but noncompliance, as Thomas Hobbes observed, leads to war, which is worse for all.

Rational egoism, then, leads to the abandonment of moral egoism. Sensible people will condemn egotism, and regard selfishness as a vice: We do better if we care about each other, engage in mutually beneficial activity, and thus refrain from one-sided activity that tramples upon others, such as killing, lying, cheating, stealing, or raping. The core of truth in egoism leads to a fairly familiar morality, whose principles must cash out in terms of the good of every agent participating in society. Narrowly egoistic moral principles cannot do this, and thus are the first to be rejected by rational egoists—another of those fascinating paradoxes of which philosophy is full.

See also Butler, Joseph; Egoism and Altruism.

Bibliography

Baier, K. *The Moral Point of View*. Ithaca, NY: Cornell University Press, 1958. Chapter 8, sections 1–4.

Butler, J. *Sermons*. London, 1726. Preface, I, and XI.

Fumerton, R. A. *Reason and Morality: A Defense of the Egocentric Perspective*. Ithaca, NY: Cornell University Press, 1990.

Hobbes, T. *Leviathan*. London, 1651. Chapters VI, XIII–XV.

Huemer, M. "Is Benevolent Egoism Incoherent?" *Journal of Ayn Rand Studies* 3 (2002): 259–288.

Hunt, L. "Flourishing Egoism." *Social Philosophy & Policy* 16 (1999): 72–95.

Kalin, J. "Two Kinds of Moral Reasoning: Ethical Egoism as a Moral Theory." *Canadian Journal of Philosophy* 5 (1975): 323–356.

Kelley, D. *Unrugged Individualism*. Poughkeepsie, NY: Institute for Objectivist Studies, 1996.

Long, R. *Reason and Value: Aristotle vs Rand*. Poughkeepsie, NY: The Objectivist Center. 2000.

Machan, T. *Generosity*. Washington, DC: Cato Institute, 1998.

Mack, E. "Egoism and Rights." *The Personalist* (Winter 1973): 5–33

Medlin, B. "Ultimate Principles and Ethical Egoism." *Australian Journal of Philosophy* 35 (1957): 111–118.

Plato. *Republic*. Books 1, 2.

Rand, A. *Virtue of Selfishness*. Toronto: Signet. 1961.

OTHER RECOMMENDED WORKS

On virtue-theoretic treatments of self-interest, see: Annas, J., *The Morality of Happiness* (New York: Oxford. 1995); Aristotle, *Nicomachean Ethics*; Badhwar, N., "Friends As Ends in Themselves," *Philosophy & Phenomenological Research* 48 (1987): 1–23; Cottingham, J., "Partiality and the Virtues," in *How Should One Live?* (New York: Oxford University Press, 1996); Den Uyl, D., *The Virtue of Prudence* (New York: Peter Lang, 1991); Foot, P., "Morality as a System of Hypothetical Imperatives," in *Virtues and Vices* (Berkeley: University of California Press, 1978); Hampton, J., "Selflessness and the Loss of Self," *Social Philosophy & Policy* 10 (1993): 135–165; von Humboldt, W., *The Limits of State Action* (Indianapolis: Liberty Fund, 1993); Rasmussen, D., "Human Flourishing and the Appeal to Human Nature," *Social Philosophy & Policy* 16 (1999): 1–43; Schmidtz, D., *Rational Choice and Moral Agency* (Princeton, NJ: Princeton University Press, 1995).

On contractarian treatments of self-interest: Gauthier, D., *Morals by Agreement* (New York: Oxford University Press, 1986); Narveson, J., *The Libertarian Idea* (Toronto: Broadview, 2001).

Other theories with egoistic components: Mack, E., "Moral Individualism, Agent-Relativity, and Deontic Restraints," *Social Philosophy & Policy* 7 (1989): 81–111; Nagel, T., *View from Nowhere* (New York: Oxford University Press, 1986); Scheffler, S., *The Rejection of Consequentialism* (New York: Oxford University Press, 1982); Wolf, S., "Moral Saints," *Journal of Philosophy* 79 (1982): 419–439.

On the connection between psychological and ethical egoism: Dawkins, R., *The Selfish Gene* (New York: Oxford University Press, 1976); Feinberg, J., "Psychological Egoism," *Reason and Responsibility* (Belmont: Wadsworth, 2005); Nagel, T., *The Possibility of Altruism* (Princeton, NJ: Princeton University Press, 1979); Ridley, M., *The Origins of Virtue* (London: Penguin, 1998).

Jan Narveson (1996)
Bibliography updated by David Schmidtz (2005)

ETHICAL NATURALISM

Philosophical naturalism, considered in general, is not a unified doctrine but a broad label applied both to methodological stances (e.g., "The methods of philosophy are continuous with those of empirical science") and to substantive positions (e.g., "For a belief to be epistemically warranted is for it to be the product of a certain kind of causal process"). The two are often combined, as when a naturalistic interpretation of a given domain of discourse is justified as "the best explanation" of associated practices. However, the two are in principle independent. In the moral case, for example, it has been argued that a projectivist or noncognitivist interpretation gives a better explanation of moral practice than any substantive naturalism (Blackburn 1984, Gibbard 1990).

But what makes a method or interpretation naturalistic? Attempts to give an explicit definition have largely been abandoned in favor of pointing. Roughly, naturalistic methods are those followed in actual scientific research (including—according to some but not all naturalists—mathematics and social sciences as well as natural sciences). And a naturalistic interpretation of a discourse is one based upon predicates or terms that play a role in the explanatory theories that research has generated.

This characterization of naturalism is informative but incomplete. There are vigorous debates within the philosophy of science over just what the methods, concepts, or posits of contemporary science are. Moreover, interpretation based upon naturalistic terms encompasses some quite different tasks. Some examples follow, but first we should ask, Why stay within naturalistic terms at all? Science is a theoretical, descriptive/explanatory enterprise while morality is held to be essentially practical and normative. One might think, no sooner did morality emerge from the shadow of religion than philosophers began trying to push it into the shadow of science. Is it never to be allowed to stand in its own right as a distinctive domain of inquiry?

An answer of sorts is possible. Morality by its nature cannot stand entirely on its own. Moral discourse is supervenient upon the nonmoral and, specifically, the natural—two actions or agents cannot differ in their moral qualities unless there is some underlying difference in their natural qualities. This and other truisms about morality, such as "Ought" implies "can," tie moral evaluation to the natural world in ways that no ethical theory can altogether ignore. Moreover, morality presents us with various epistemic and metaphysical puzzles. We

believe that we have come to possess at least some moral knowledge—but how? (See Harman 1977.) We treat moral statements as if they stated genuine propositions—but can this idea be sustained in light of the normative role of moral judgment? We freely make moral judgments, but do they have presuppositions or make claims that are incompatible with our understanding of the natural world?

Hard determinists, for example, have challenged intuitive attributions of moral responsibility by arguing that the notion of free agency they presuppose is incompatible with the world revealed by physics. And John L. Mackie is led to an "error theory" of morality by his diagnosis that moral evaluation attributes to states of the world an objective "to-be-pursuedness" that cannot be fit with any plausible empirical theory (Mackie 1977).

Immanuel Kant, for one, frankly accepted that he could see no way of reconciling the deliberative standpoint of morality with the causal perspective of science. Rational agents must, he held, postulate the compatibility of moral agency with the natural order, even though this remains inexplicable to them. But few philosophers have been willing to stop there. Empirical science affords the best-developed picture we have of ourselves and our world. Without the special authority of religion to back it up, morality inevitably becomes a focus of practical and theoretical concern.

Substantive moral naturalists in effect propose to overcome some of the mystery and potential conflict surrounding the relation of morality to our empirical self-understanding by showing just how much of morality might be found within the domain of the natural. This could be done by providing a naturalistic account of moral discourse that affords an analysis of moral terms (Lewis 1989), or permits a worthwhile revision of moral language that nonetheless can serve virtually all the same functions (Brandt 1979), or enables us to reduce moral properties to natural properties (Railton, 1993), or shows moral properties to be natural properties in their own right (e.g., thanks to their contribution to empirical explanation; see Boyd 1988, Miller 1985, Sturgeon 1985). Substantive ethical naturalism promises to explain such important features of moral discourse and practice as the applicability of notions of truth and falsity to moral claims, the supervenience of the moral upon the natural, the role of natural properties in justifying moral claims, and the possibility of semantic and epistemic access to moral notions through ordinary experience.

The first half of the twentieth century had not been kind to substantive ethical naturalism (for a brief history, see Darwall et al. 1992). Condemned by G. E. Moore (1903) for committing the "fallacy" of trying to close an "open question" by analytic means and rejected by nonfactualists (emotivists, prescriptivists, etc.) for failing to capture the special relation of moral evaluation to motivation and action, naturalism fell into disuse. But by mid-century naturalism had begun to win its way back. The initial steps were taken, independently, by Philippa Foot (1958–1959) and Geoffrey Warnock (1967), who argued that one could not be competent in moral discourse unless one possessed some substantive, contentful moral concepts. Moral evaluation is distinguished from aesthetic or prudential, for example, in part because it has a certain descriptive, arguably natural content—namely, a concern with the effects of our actions on the well-being of others. If we came upon a society in whose behavioral code the key notion was *guleb,* a term applied in the paradigm case to warriors who have killed an enemy bare-handed, we would certainly mislead if we translated *guleb* as "morally good" or "just" rather than "valiant" or "courageous."

Meanwhile, Peter Geach (1965) showed convincingly that existing nonfactualist views could not account for the full grammar of moral discourse, in particular, the logical behavior of unasserted moral claims in conditionals.

Foot (1972) took the next step as well, challenging the "internalist" conception of the relation of moral evaluation to motivation that served as the basis for nonfactualism. She argued that ordinary moral agents are able to see themselves as motivated by a rationally optional concern for others. Those who lack such a concern might lack moral character, but they do not make a linguistic mistake in using the moral vocabulary.

This sort of moral "externalism" offers an alternative explanation of why moral evaluation and motivation are so intimately related, at least in paradigm cases. Concern for others is a very basic part of normal human life. An Aristotelian would say that human nature itself is social; a Darwinian would emphasize the contribution of concern for others to inclusive fitness and to the possibility of benefiting from reciprocal altruism. Speculative biology apart, it is possible to see how social norms involving concern for others, keeping promises, and so forth might emerge and be sustained in virtue of their contribution to solving various serious coordination and collective-action problems. Such norms will function best only if well internalized by a major part of the population. It should therefore be unsurprising that moral judgment is usually accompanied by a positive attitude. Moreover, it

should not be forgotten that moral judgment is a species of assertion and that assertion itself involves, not only signaling a cognitive attitude of belief, but also various forms of active endorsement or encouragement, as well as associated claims of authority. Moral externalism, by drawing upon these ingredients (and others) for an alternative explanation of the evidence—such as it is—offered on behalf of internalism, has attracted a number of defenders (see, for example, Boyd 1988, Brink 1989, Railton 1986).

Another sort of naturalism, however, takes the opposite tack. It treats the purported relation to motivation as fundamental but interprets it in a subjectivist rather than nonfactualist manner. Subjectivist interpretations of moral discourse have historically faced difficulties in accommodating all the elements of an interconnected set of features of morality: the critical use of moral assessment, the nonrelativistic character of moral judgment and the possibility of genuine moral disagreement across social or cultural differences, the limits on empirical methods in resolving moral disputes, and the seemingly normative character of the relation between moral judgment and motivation. Can new forms of subjectivism succeed where others have failed?

Consider the simple subjectivist formula:

(1) Act *A* is morally good = *A* is such that one would approve of the performance of *A*.

Since approval is a positive attitude, (1) establishes a relationship with a source of motivational force. But is it the right relationship?

We do not typically regard our current tendencies to approve or disapprove as morally authoritative—they might, for example, be based upon hasty thinking or false beliefs. This has led naturalists to modify (1) to require that the approval be well informed and reflectively stable. (See, for example, Brandt 1979 and Firth 1952. For criticism, see Velleman 1988.)

Moreover, not all species of approval have a moral flavor. I can approve of an act because of its aesthetic or pious qualities, for example. Some naturalists therefore amend (1) to restrict the object of approval (e.g., to the set of rules one would—reflectively, informedly, etc.—approve for a society in which one is going to live [cf. Brandt 1979]). Others attempt to identify in naturalistic terms a specifically moral sort of attitude of approval or disapproval (e.g., an attitude of impartial praise or anger). Critics have argued that no noncircular characterization of this kind is possible (for a subjectivism without reductive ambitions, see Wiggins 1987).

Formulas like (1) also threaten to yield relativism. Since they introduce a necessary link to facts about motivation, moral attribution becomes tied to contingencies of individual psychology. That seems wrong, since moral evaluation purports to abstract from individual interest and motive and to prescribe universality. If one is not correspondingly motivated, that is a deficiency in oneself rather than an excusing condition or a limit on the reach of moral judgment. Each of us recognizes that he or she can in this sense be motivationally defective from a moral point of view. (But see Harman 1975 for a defense of a naturalistic moral relativism.) This has led naturalists to modify the formula away from the individualistic language of "one" or "I" and in the direction of a more inclusive "we" or "everyone" or even "normal humans" (see, respectively, Lewis 1989, Smith 1994, and Firth 1952). New problems arise. The notion of "normal human" threatens to introduce a term that itself requires naturalization—since we believe that statistically "normal" humans might be motivationally defective from a moral point of view—for example, in lacking sympathy with those from other groups. (Of course, one could at this point also embrace circularity.) If we insist that everyone approve, there is again a risk that contingencies of motivational idiosyncrasies will receive authority—this time, in preventing us from attributing moral value to states of affairs virtually all (but still not quite all) of us approve heartily on reflection. A less ambitious alternative is to replace "one" with "us" and seek moral consensus where we may. This would help explain the "outreach" function of moral discourse without altogether removing the account's relativism.

An alternative approach avoids relativism by "rigidifying" the subjectivist formula (cf. Wiggins 1987). One fixes the truth conditions of moral judgments by reference to the motivations *actually* prevalent in one's moral community (e.g., "*A* is such that we, with our actual motives and with full and informed reflection, would approve of it"). This secures the desirable result that changes in our motives will not in themselves change what is morally good. But it undermines some of the critical role of moral assessment in our own society (since, again, we can imagine that our actual motives are morally defective) and will have the result that those brought up in different social environments with different acquired motivations will lack a common subject matter even though they believe they are having a genuine moral disagreement (for discussion, see Johnston 1989).

No ethical naturalism has emerged that meets all the desiderata of an account of moral discourse and practice.

Nonnaturalists and nonfactualists attribute this to a mistaken starting point. But no alternative account has met all the desiderata, either. Moral naturalists have often been accused of "changing the subject"—shifting the locus of attention from the position of the agent involved in practical deliberation to that of the scientist engaged in theoretical description. But this criticism begs the question. Naturalists seek to explain, not ignore, moral experience; if they are right, the phenomena they study are the very stuff of moral thought and action.

See also Conditionals; Determinism and Freedom; Foot, Philippa; Kant, Immanuel; Mackie, John Leslie; Metaethics; Moore, George Edward; Naturalism; Philosophy of Science, Problems of; Projectivism; Supervenience.

Bibliography

Blackburn, S. *Spreading the Word.* Oxford: Clarendon Press, 1984.

Boyd, R. "How to Be a Moral Realist." In *Essays on Moral Realism,* edited by G. Sayre-McCord. Ithaca, NY: Cornell University Press, 1988.

Brandt, R. *A Theory of the Good and the Right.* New York: Oxford University Press, 1979.

Brink, D. O. *Moral Realism and the Foundations of Ethics.* Cambridge, U.K.: Cambridge University Press, 1989.

Darwell, S. et al. "Toward *Fin de siècle* Ethics." *Philosophical Review* 101 (1992): 317–345.

Firth, R. "Ethical Absolutism and the Ideal Observer." *Philosophy and Phenomenological Research* 12 (1952): 317–345.

Foot, P. "Moral Beliefs." *Proceedings of the Aristotelian Society* 59 (1958–1959): 83–104.

Foot, P. "Morality as a System of Hypothetical Imperatives." *Philosophical Review* 81 (1972): 305–316.

Geach, P. "Assertion." *Philosophical Review* 74 (1965): 445–465.

Gibbard, A. *Wise Choices, Apt Feelings.* Cambridge, MA: Harvard University Press, 1990.

Harman, G. "Moral Relativism Defended." *Philosophical Review* 84 (1975): 3–22.

Harman, G. *The Nature of Morality.* New York: Oxford University Press, 1977.

Johnston, M. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* 63 (1989): suppl., 139–174.

Lewis, D. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* 63 (1989): suppl. 113–137.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* New York: Penguin, 1977.

Miller, R. "Ways of Moral Learning." *Philosophical Review* 94 (1985): 507–556.

Moore, G. E. *Principia Ethica.* Cambridge, U.K.: Cambridge University Press, 1903.

Railton, P. "Moral Realism." *Philosophical Review* 95 (1986): 163–207.

Railton, P. "Reply to David Wiggins." In *Reality, Representation, and Projection,* edited by J. Haldane and C. Wright. New York: Oxford University Press, 1993.

Smith, M. *The Moral Problem.* Oxford: Blackwell, 1995.

Sturgeon, N. "Moral Explanations." In *Morality, Reason, and Truth,* edited by D. Copp and D. Zimmerman. Totowa, NJ: Rowman and Allanheld, 1985).

Velleman, D. "Brandt's Definition of 'Good.'" *Philosophical Review* 97 (1988): 353–371.

Warnock, G. *Contemporary Moral Philosophy.* London: Macmillan, 1967.

Wiggins, D. "A Sensible Subjectivism?" In *Needs, Values, Truth: Essays in the Philosophy of Value.* Oxford: Blackwell, 1987.

Peter Railton (1996)

ETHICAL NATURALISM [ADDENDUM]

Substantive ethical naturalists believe that the ethical is natural, that is, that ethical properties are natural properties. *Strong*, or *reductive*, ethical naturalists hold that there is an interesting further question of *which* natural properties the ethical ones are, just as there is an interesting question of which chemical property water is (Railton 1986, Jackson 1997). *Weak* ethical naturalists deny this; some hold that ethical properties, though natural, are irreducible (Boyd 1988), while others are in the business of *revising* moral language (Brandt 1979).

Moore's open-question argument was advanced against strong ethical naturalism, which he claimed committed the *naturalistic fallacy*. The basic idea of the argument is that we can test *cognitive significance* to test claims of property identity. So, for example, if *good* just is *pleasure*, then "Pleasure is good" should mean the same as "Good is good," and be equally informative to speakers who understand the two statements. There are different versions of the argument, according to which test of cognitive significance is used. It looks like a good argument against those who think that such property identities are *analytic*, or follows from the meanings of the words or the concepts that they express, but not against those who think that this identity thesis is *synthetic*, like the thesis that water is H_2O (Moore 1903, Brink 1989).

But synthetic naturalism may not have all of the same philosophical attractions as the analytic version. For example, naturalists have long held that naturalism can help explain how we can acquire moral knowledge. This seems easily true of analytic naturalism (Harman 1977), but synthetic naturalists have not yet explained how their

naturalism provides help in moral epistemology not available to nonnaturalists.

More striking is the fact that though it is widely known that synthetic naturalism escapes Moore's open-question argument, many still believe that the argument shows something against synthetic naturalism—a reaction that synthetic naturalists find puzzling. Nonnaturalists typically remain convinced, however, that Moore's argument illustrates how strong naturalism involves identifying ethical properties with something that they are not (Shafer-Landau 2003). And others maintain that naturalists do not really believe in ethical properties at all (Nagel 1985, McNaughton 1989).

It is not hard to see the force of these objections. If a friend tells you that he believes in God because God is love and he believes in love, you are bound to conclude that if he is really speaking literally, then he is an atheist. Whatever theists mean by saying that God is love, it is not simply that God is the relation that holds between two people when one loves the other. Naturalist theories about God seem bound to feel this way. They all seem to identify God with something else, and as a result they fail to be realist about God (Schroeder 2005).

By analogy, this can make ethical naturalism look hopeless. But that would be premature. Strong naturalism is clearly true about bachelors. Bachelors are just unmarried adult men. If your friend tells you he believes in bachelors because he believes in unmarried adult men, you will not conclude that he does not really believe in bachelors. Nonnaturalists think that ethical properties are more like God, but strong naturalists think that they are more like bachelorhood.

One salient difference between God and bachelors is that God is supposed to have features that love and other natural entities lack: omniscience and omnipotence, among others. If your friend does not believe that love is omniscient and omnipotent, then he does not believe in God. But in contrast, it is easy to see that unmarried adult men have the properties of bachelors (Schroeder 2005).

To find out whether ethical naturalism is more like naturalism about God or like naturalism about bachelors, then, we need to think about what features ethical properties have, and whether natural properties could have them. An old idea about the ethical is that ethical properties have to be related in some way to *motivation*. This is the thesis of *internalism*. Nonnaturalists have argued that natural properties could not motivate as internalism requires (Mackie 1977), but some naturalists have responded by explaining how they can (see especially Smith 1994).

More recently, nonnaturalists have insisted that internalism, even if true, is not what is special about ethical properties. Rather, they say, what is special about ethical properties is that they are *normative*. And if natural properties are not normative, it follows that ethical properties cannot be natural ones (Hampton 1998). But the evidence that natural properties cannot be normative is no better than the evidence that they cannot be ethical. If anything, it is worse. Nonnaturalists typically say that to be normative is to involve *reasons*, in some way. But if *reason* is a natural property, then natural properties could involve reasons, and hence be normative (Schroeder 2005).

A different worry about ethical naturalism is that if ethical properties are natural, then they must be highly disjunctive, and therefore uninteresting. This notion seems to force us to give up the idea that ethical discourse carves nature at its joints. This impression is reinforced by some naturalists who write at a very abstract level, such as Jackson (1997), but it is not sustained by looking in detail at most serious strong naturalist proposals. The problem is one of level of description.

A final serious objection leveled against ethical naturalism applies just as much to weak naturalism as to strong naturalism. It is that naturalism makes ethics out to be an a posteriori discipline, whereas surely ethical knowledge should be a priori (Shafer-Landau 2003). Since nonreductive naturalists hold that what makes ethical properties respectable is that they play a certain explanatory role, they may be particularly susceptible to this charge. But perhaps synthetic strong naturalism can escape it. The fact that empirical investigation was the right way to investigate what natural property samples of water have in common does not entail that empirical investigation is the right way to investigate what natural property right actions have in common. On the contrary, it seems that philosophers engage in the study of what right actions have in common all the time. Such study is ordinary normative ethical inquiry.

Bibliography

Boyd, Richard. "How to Be a Moral Realist." In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord. Ithaca, NY: Cornell University Press, 1988.

Brandt, Richard. *A Theory of the Good and the Right*. Oxford, U.K.: Oxford University Press, 1979.

Brink, David. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Hampton, Jean. *The Authority of Reason*. Cambridge, U.K.: Cambridge University Press, 1998.
Harman, Gilbert. *The Nature of Morality*. Oxford, U.K.: Oxford University Press, 1977.
Jackson, Frank. *From Metaphysics to Ethics*. Oxford, U.K.: Oxford University Press, 1997.
Mackie, J. L. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.
McNaughton, David. *Moral Vision*. Oxford, U.K.: Basil Blackwell, 1989.
Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.
Nagel, Thomas. *The View from Nowhere*. Oxford, U.K.: Oxford University Press, 1985.
Railton, Peter. "Moral Realism." *Philosophical Review* 95 (1986): 163–207.
Schroeder, Mark. "Realism and Reduction: The Quest for Robustness." *Philosophers' Imprint* 5 (1) (2005). Available from http://www.philosophersimprint.org/005001/.
Shafer-Landau, Russ. *Moral Realism*. Oxford, U.K.: Oxford University Press, 2003.
Smith, Michael. *The Moral Problem*. Oxford, U.K.: Basil Blackwell, 1994.

Mark Schroeder (2005)

ETHICAL OBJECTIVISM

See *Objectivity in Ethics*

ETHICAL PRINCIPLES

See *Moral Principles: Their Justification*

ETHICAL RELATIVISM

The term *ethical relativism* is always used to designate some ethical principle or some theory about ethical principles, but within this limitation different authors use it quite differently. Contemporary philosophers generally apply the term to some position they disagree with or consider absurd, seldom to their own views; social scientists, however, often classify themselves as relativists. Writers who call themselves relativists always accept the first and second and sometimes accept the third of the theses described as descriptive relativism, metaethical relativism, and normative relativism, respectively.

DESCRIPTIVE RELATIVISM

The first thesis, without which the others would lose interest, is that the values, or ethical principles, of individuals conflict in a fundamental way ("fundamental" is explained below). A special form of this thesis, called "cultural relativism," is that such ethical disagreements often follow cultural lines. The cultural relativist emphasizes the cultural tradition as a prime source of the individual's views and thinks that most disagreements in ethics among individuals stem from enculturation in different ethical traditions, although he need not deny that some ethical disagreements among individuals arise from differences of innate constitution or personal history between the individuals.

FUNDAMENTAL DISAGREEMENT. The most important and controversial part of the first thesis is the claim that diversities in values (and ethics) are fundamental. To say that a disagreement is "fundamental" means that it would not be removed even if there were perfect agreement about the properties of the thing being evaluated. (If disagreement is nonfundamental then it may be expected that all ethical diversity can be removed, in principle, by the advance of science, leading to agreement about the properties of things being appraised.) Thus it is not necessarily a case of fundamental disagreement in values if one group approves of children's executing their parents at a certain age or stage of feebleness whereas another group disapproves of this very strongly. It may be that in the first group the act is thought necessary for the welfare of the parent in the afterlife, whereas in the second group it is thought not to be. The disagreement might well be removed by agreement about the facts, and indeed both parties might subscribe, now, to the principle "It is right for a child to treat a parent in whatever way is required for the parent's long-run welfare." The disagreement might be simply about the implications of this common principle, in the light of differing conceptions of the facts. There is fundamental ethical disagreement only if ethical appraisals or valuations are incompatible, even when there is mutual agreement between the relevant parties concerning the nature of the act that is being appraised.

METAETHICAL RELATIVISM

A person might accept descriptive relativism but still suppose that there is always only one correct moral appraisal of a given issue. Such a position has been widely held by nonnaturalists and by some naturalists (see, for example, the interest theory of R. B. Perry). The metaethical relativist, however, rejects this thesis and denies that there is always one correct moral evaluation. The metaethical relativist thesis is tenable only if certain views about the meaning of ethical (value) statements are rejected. For instance, if "*A* is right" *means* "Doing *A* will contribute at

least as much to the happiness of sentient creatures as anything else one might do," it is obvious that one and only one of the two opinions "*A* is right" and "*A* is not right" is correct. Thus, the metaethical relativist is restricted to a certain range of theories about the meaning of ethical statements. He might, for instance, subscribe to some form of emotive theory, such as the view that ethical statements are not true or false at all but express the attitudes of the speaker. Or he might adopt the naturalist view that "is wrong" means "belongs to the class of actions toward which I tend to take up an impartial attitude of angry resentment" (held by the relativist E. A. Westermarck) or the view (suggested by the anthropologist Ruth Benedict) that the phrase "is morally good" means "is customary."

ETHICAL REASONING. At the present time metaethical relativists do not wish to rest their case solely on an appeal to what ethical statements mean; nor would their critics. The point of active debate is rather whether there is some method of ethical reasoning whose acceptance can be justified to thoughtful people with force comparable to the force with which acceptance of inductive logic can be justified. Is there any such method of ethical reasoning that can be expected in principle to show, when there is a conflict of values or ethical principles, that one and only one solution is correct in some important and relevant sense of "correct"? Metaethical relativists deny that there is any such method, and their denial may take either of two forms: They may deny that there is any method of ethical reasoning that can be justified with force comparable to that with which scientific method (inductive logic) can be justified. Or they may agree that there is such a method but say that its application is quite limited, and in particular that the fullest use of it could not show, in every case of a conflict of ethical convictions or of values, that one and only one position is correct in any important sense of "correct."

USE OF THE TERM *RELATIVISM.* Many writers, both in philosophy and in the social sciences, accept a combination of descriptive relativism and metaethical relativism. Philosophers who hold this view, however, seldom label themselves "relativists," apparently because they think the term confusing in this context. There is seldom objection to "cultural relativism" as a descriptive phrase, for it can be taken to mean that a person's values are "relative" to his culture in the sense of being a function of or causally dependent on it. But if "ethical relativism" is construed in a similar way, to mean that ethical *truth* is relative to, in the sense of being dependent on or a function of, something (for example, a person's cultural tradition), then this term is thought to be confusing since it is being used to name a theory that essentially denies that there is such a thing as ethical "truth."

One frequent confusion about what implies ethical relativism should be avoided. Suppose metaethical relativism is mistaken, and there is a single "correct" set of general ethical principles or value statements. It may still be true, and consistent with acceptance of this "correct" set of principles, that an act that is right in some circumstances will be wrong in other circumstances. Take, for instance, the possible "correct" principle "It is always right to do what will make all affected at least as happy as they could be made by any other possible action." It follows from this principle that in some situations it will be right to lie (for instance, to tell a man that he is not mortally ill when one knows he is, if he cannot bear the truth) and that in other situations it will be wrong to lie. Thus, even if metaethical relativism is false there is a sense in which the rightness of an act is relative to the circumstances or situation. The fact that the rightness of an act is relative to the circumstances in this way does not, of course, imply the truth of metaethical relativism.

NORMATIVE RELATIVISM

Neither descriptive nor metaethical relativism commits one logically to any ethical statement. The former is simply an assertion about the diversity of moral principles or values actually espoused by different persons; the latter is only a general statement about whether ethical principles are ever "correct." Nothing in particular about what ought to be, or about what someone ought to do, follows from them. Of particular interest is the fact that it does not follow that persons, depending on their cultural attachments, ought to do different things. In contrast, a person who holds to some form of what I shall call "normative relativism" asserts that something is wrong or blameworthy if some person or group—variously defined—thinks it is wrong or blameworthy. Anyone who espoused either of the following propositions would therefore be a normative relativist:

(*a*) "If someone thinks it is right (wrong) to do *A*, then it *is* right (wrong) for him to do *A*." This thesis has a rather wide popular acceptance but is considered absurd by philosophers if it is taken to assert that what someone thinks right really is right for him. It is held to be absurd because taken in this way, it implies that there is no point in debating with a person what is right for him to do unless he is in doubt himself; the thesis says that if he *believes* that *A* is right, then it *is* right, at least for him.

The thesis may be taken in another sense, however, with the result that it is no longer controversial, and no longer relativist. The thesis might mean: "If someone thinks it is right for him to do *A,* then he cannot properly be condemned for doing *A.*" This statement merely formulates the view, widespread in the Western world, that a person cannot be condemned morally for doing what he sincerely believes to be right. (In order to receive universal approval, some additions must doubtless be added to the thesis, such as that the person's thoughts about what is right must have been the product of a reasonable amount of careful reflection, not influenced by personal preferences, and so on.) The thesis is not relativist, since it is not asserted that any person's or group's belief that something is blameworthy is either a necessary or a sufficient condition of its being blameworthy.

(*b*) "If the moral principles recognized in the society of which *X* is a member imply that it is wrong to do *A* in certain circumstances *C,* then it *is* wrong for *X* to do *A* in *C.*" This principle says, in effect, that a person ought to act in conformity with the moral standards of his group. Like the preceding principle, this one has a good deal of popular acceptance, is espoused by some anthropologists, and has some plausibility; it will be discussed below.

DIFFICULTIES IN THE RELATIVIST POSITIONS

The following appear to be the most important questions about the various relativist theses: (*a*) Is descriptive relativism supported by the scientific evidence? There are methodological obstacles in the way of answering the question whether there is fundamental diversity of ethical views. Such diversity would be established by producing two individuals, or cultures, who attribute identical properties to an act but nevertheless appraise it differently. But it is not easy to be sure when one has produced two such individuals or cultures. First, it is difficult to demonstrate that an act is believed to have identical properties by individuals or groups who appraise it differently. Is theft the same thing in societies where conceptions or systems of property differ? Is incest the same thing in societies with different kinship terminologies, different ways of counting lineage, and different beliefs about the effects of incest? It is possible to question members of different cultural traditions in an abstract way so that such differences in conception are ruled out, but then it is likely that the informant will not grasp the question and that his answer will be unreliable. The second difficulty is that there is no simple test for showing that groups or individuals really conflict in their appraisals. We may know that we think it morally wrong to do so-and-so, but it is not clear how to determine whether a Navajo agrees with us. Perhaps we have first to determine whether the Navajo language contains an expression synonymous with "is morally wrong." Or can we show that a Navajo does not think it morally wrong to do *A* by the fact that he feels no guilt about doing *A?* Or perhaps a mere conflict of *preferences* is sufficient to establish a disagreement in personal values. These questions deserve more discussion among anthropologists than they have received.

The evidence for descriptive relativism consists mostly of reports from observers about what is praised, condemned, or prohibited in various societies, usually with only scanty information on the group's typical conception of what is praised or blamed. In some instances projective methods and dream analysis have been utilized, and discussions with informants have elicited fragments of the conceptual background behind the appraisals. On the basis of such material most social scientists believe there is some fundamental diversity of values and ethical principles. Several decades ago some investigators (among them W. G. Sumner and Ruth Benedict) supposed that the extent of diversity was practically unlimited, but by the 1960s it was believed that there is also considerable uniformity (for example, universal disapproval of homicide and cruelty). One reason for believing there is considerable uniformity is that it appears that it would be difficult for a social group even to survive without the presence of certain features in its value system. (A social system must provide methods for rearing and educating the young, for mating, for division of labor, for avoiding serious personal insecurity, and so on.) Uniformity of evaluation is the rule in areas that pertain to survival or to conditions for tolerable social relationships; in other areas there are apt to be fundamental differences. Psychology adds some support to this construction of the empirical data. It offers a theory of enculturation that explains how fundamentally different values can be learned, and it also suggests how some universal human goals can set limitations to diversity among value systems.

(*b*) Does descriptive relativism support metaethical relativism? It is evident that from descriptive relativism nothing follows about what ethical statements mean or could fruitfully be used to mean. It is also evident that nothing follows about whether there is some method of ethical reasoning that can correctly adjudicate between conflicting ethical commitments, at least in some cases. Descriptive relativism may very well have bearing, however, on whether a justifiable method of reasoning in

ethics (assuming there is one) could succeed in adjudicating between all clashes of ethical opinion. To be sure about this we would need to have an account of the reasonable method of ethical reflection before us. But let us take an example. Suppose we think that the only reasonable way to correct a person's ethical appraisal is to show him that it does not coincide with the appraisal he would make if he were vividly informed about all the relevant facts and were impartial in his judgment. Then suppose the descriptive relativist tells us that some people are simply left cold by the ideal of equality of welfare and that others view it as a basic human right, when both groups have exactly the same beliefs about what equality of welfare is and what its consequences would be. In this case we might be convinced that both parties were already impartial (if the views were not just typical of different social or economic classes) and that further information probably would not change their views. Doubtless much more analysis of the situation is necessary, but it is clear that given the described assumption about actual ideals and the assumption about the limitation of ethical argument, one might be led to a cautious acceptance of the view that not all ethical disputes can be resolved by this justifiable method of ethical reasoning.

(*c*) Are cultural and metaethical relativism necessarily committed to any form of normative relativism? Neither the cultural relativist nor the metaethical relativist is committed logically to any form of normative relativism. It is consistent to assert either of these positions and also to affirm any value judgment or ethical proposition one pleases. However, the second proposition cited under normative relativism (that a person ought to act in conformity with the moral standards of his group) at least presupposes the acceptance of cultural relativism. There would be no point in asserting this normative principle if cultural relativism were not accepted.

How strong are the arguments that can be advanced in favor of this form of normative relativism? Suppose that in *X*'s society it is a recognized moral obligation for a person to care for his father, but not for his father's siblings (at least to anything like the same degree), in illness or old age. Suppose also that in *Y*'s society it is recognized that one has no such obligation toward one's father or his siblings but does have it toward one's mother and her siblings. In such a situation it is hard to deny that *X* seems to have some obligation to care for his father and that *Y* seems not to have, at least to the same degree. (Some philosophers hold that there is no obligation on anyone, unless one's society recognizes such an obligation for the relevant situation.) So far, the principle seems intuitively acceptable. In general, however, it appears less defensible, for the fact that *X*'s society regards it as wrong to play tennis on Sunday, to marry one's deceased wife's sister, and to disbelieve in God does not, we should intuitively say, make it wrong for *X* to do these things. Thus, our principle seems valid for some types of cases but not for others. The solution of this paradox probably is that for those cases (like an obligation to one's father) where the principle seems acceptable, the reasons for which it seems acceptable are extremely complex and are not based simply on the fact that society has asserted that an obligation exists. When society recognizes a moral obligation, there are many repercussions that basically affect the types of responsibility the individual may have toward other members of his society. For instance, one result of a society's recognizing a son's moral obligation to care for his father is that no one else will take care of the father if the son does not. Another is that a kind of equitable insurance system is set up in which one pays premiums in the form of taking responsibility for one's own father and from which one gets protection in one's old age. So it is at least an open question whether it can seriously be claimed that the moral convictions of a society have, in themselves, any implication for what a member of the society is morally bound to do, in the way our suggested principle affirms that they do.

See also Emotive Theory of Ethics; Metaethics; Perry, Ralph Barton; Sumner, William Graham; Value and Valuation; Westermarck, Edward Alexander.

Bibliography

Aberle, D. F., et al. "The Functional Prerequisites of a Society." *Ethics* 60 (1950): 100–111.

Asch, S. E. *Social Psychology,* Chs. 12, 13. New York: Prentice-Hall, 1952. An excellent critique of the evidential basis for descriptive relativism.

Benedict, Ruth. "Anthropology and the Abnormal." *Journal of General Psychology* 10 (1934): 59–82. A graphic account of some ethnological data in support of a rather extreme type of relativism.

Brandt, R. B. *Ethical Theory,* Chs. 5, 6, 11. Englewood Cliffs, NJ: Prentice-Hall, 1959. A cautious review and assessment of the evidence.

Brandt, R. B. *Hopi Ethics.* Chicago: University of Chicago Press, 1954. An analysis of the values of the Hopi Indians with special reference to relativism. Methodologically oriented.

Firth, Raymond. *Elements of Social Organization,* Ch. 6. London: Watts, 1951. A learned survey of anthropological material by an anthropologist with leanings toward relativism.

Ginsberg, Morris. *Essays in Sociology and Social Philosophy.* Vol. I. New York, 1957. A learned survey of the evidence by a sociologist.

Herskovits, Melville. *Man and His Works,* Ch. 5. New York: Knopf, 1948. One of the most vigorous contemporary defenses of relativism by an anthropologist.

Kluckhohn, C. "Ethical Relativity." *Journal of Philosophy* 52 (1955): 663–677. A cautious assessment of the evidence by an anthropologist who stresses the uniformities of value systems.

Ladd, J. *Structure of a Moral Code.* Cambridge, MA: Harvard University Press, 1956. A fine study of the Navaho, emphasizing methodological problems.

Linton, R. "Universal Ethical Principles: An Anthropological View." In *Moral Principles of Action,* edited by R. N. Anshen. New York: Harper, 1952. An excellent statement, by an anthropologist of the middle-of-the-road position, emphasizing both uniformities and diversity.

Westermarck, Edvard A. *Ethical Relativity,* Ch. 5. New York: Harcourt Brace, 1932. A shorter defense of the relativist position, intended primarily for persons interested in philosophy.

Westermarck, Edvard A. *The Origin and Development of the Moral Ideas,* Vol. I, Chs. 1–5. London: Macmillan, 1906. One of the most influential statements of the relativist position, with encyclopedic documentation; however, it is methodologically somewhat unsophisticated.

Richard B. Brandt (1967)

ETHICAL RELATIVISM [ADDENDUM]

Accepting ethical relativism can make a big difference. Opponents often think that the doctrine arises from conceptual confusion and encourages indifference to moral enormities. Advocates think that the doctrine is an antidote to an oppressive moral imperialism that often rationalizes the more worldly forms of imperialism. Both sides can agree that ethical relativism has significant implications for how we should do normative ethics.

VULGAR AND NOT-SO-VULGAR RELATIVIST ARGUMENTS FOR TOLERANCE

A good example of debate over the normative implications of ethical relativism is Bernard Williams's (1972) criticism of "vulgar relativism," which he defined in three propositions: (1) "Right" means "right for a given society." (2) "Right for a given society" is to be understood in a functionalist sense (roughly, that conceptions of what is right are part of the social fabric). (3) It is wrong for people in one society to condemn or interfere with the values of another society. Williams pointed out that (3) employs a nonrelative sense of "right" and "wrong" excluded by (1). The vulgar relativist confusedly derives a universal moral principle of toleration from a sense of "right" that excludes any universal principle.

Relativists might deny that they are arguing for universal toleration. Rather, they may view tolerance as a value prominent within certain kinds of moralities—ones that emphasize respect for the autonomy of individuals and peoples and that require intervention in their affairs to be justifiable to them (at least when the intervention is rational and informed). To practice such restraint is to treat individuals as ends in themselves. Moreover, relativists may be addressing those who value autonomy to persuade them in particular that the existence of apparently irresolvable moral disagreements over basic moral values and norms is an occasion for not intervening (Wong 1984).

Indeed, one of the apparently irresolvable disagreements might concern the value of autonomy itself. Some moral traditions emphasize that individuals have morally legitimate interests that ought to be protected when they come into conflict with social interests. Other traditions emphasize that individuals flourish only in relationship to others. The former will provide a central place for the value of individual rights, such as rights to choose one's vocation and rights to privacy. The latter will not, or will have to provide a different basis for such rights, probably in terms of their value to desirable relationships. Reflective and knowledgeable members of both kinds of traditions might become aware of this value difference, and some of them might conclude that members of the other tradition are not necessarily mistaken. It may be, however, that only members of an autonomy-valuing tradition, such as the traditions dominant in the industrialized countries of the West, have a reason to refrain from intervening in other traditions on the grounds that such an intervention impedes members of the other traditions from acting on their reasonable values.

The no-so-vulgar relativist argument for tolerance, then, starts from the premise of metaethical relativism—the doctrine that there can be conflicting moral judgments about a particular case that are both fully correct (Harman 1978b). This metaethical premise coheres perfectly well with a normative premise that values autonomy, and their conjunction yields the conclusion, addressed to those who value autonomy, to refrain from intervening in the affairs of other societies.

NORMATIVE RELATIVISM

There is another way in which relativism can have implications for normative theory. Gilbert Harman (1978b) defined "normative moral relativism" as the position that

two people can be subject to conflicting moral demands and not be subject to some more basic moral demand that, in their situation, accounts for this. For example, some groups of people think that one ought not to cause animals needless suffering, and others do not. Normative relativism allows for the possibility that the former ought to refrain from causing animals needless suffering and that the latter are not subject to a similar constraint. Harman's path to normative relativism starts from two premises: Moral demands, when applicable, provide people with compelling reasons to act in the required way, and reasons to act are "internal." Reasons are internal if the relevant agents have available warranted practical reasoning leading to a decision to perform the relevant act, and if the practical reasoning is anchored in moral demands that the agents already accept or in other desires, intentions, and beliefs that they have. A controversial part of Harman's argument that reasons are internal is his rejection of the distinction between justifying moral reasons and motivating moral reasons. Justifying moral reasons, according to this distinction, warrant agents to do things without their necessarily being able to reason from existing desires, intentions, and beliefs to a decision to perform the actions.

IMPLICATIONS OF METAETHICAL RELATIVISM FOR DOING NORMATIVE ETHICS

Accepting metaethical relativism can shape one's conception of how to do normative ethics. Consider Harman's (1975) account of morality as constituted by the implicit agreements that structure relations among the parties. On this account, many moral disagreements must be understood as disagreements among members of a group bound together by an implicit agreement. They disagree in how to interpret the terms of their agreement or in how to solve or mitigate conflicts between different elements of their agreement. This conventionalist account of morality eliminates the possibility of resolving disagreements by appealing to ideal moral principles existing independently of any implicit agreement. Rather, people are left to work out their differences with each other in ways very much like the give-and-take of politics (Harman 1978a).

Harman applies this idea to explain why the duty not to harm others is commonly perceived to be stronger than the duty to help those in need. Everyone benefits equally from a stringent duty not to harm, but the poor and weak benefit much more from a duty of mutual aid. Since the poor and weak have much less bargaining power, the expected compromise is just what is commonly accepted: a much stronger duty not to harm. Harman's conventionalist conception of morality prompts an attitude toward moral phenomena that would be different under a robustly realistic conception of morality. Under a conventionalist conception, one comes to view the belief in a stronger duty not to harm not as based on a moral fact existing independently of social convention, but as resulting from the calculus of bargaining. If one desires to change the terms of agreement, one must engage in politics. To strengthen a duty to mutual aid, for example, one might look for points of leverage with which to strike a more advantageous bargain for the poor and weak, or one might appeal to the desires or interests of the rich and strong that go beyond narrow self-interest and are served by a stronger duty to mutual aid.

Metaethical relativism need not rest on such a purely conventionalist conception of morality as Harman's. Rather, one might hold that morality is socially invented under constraints arising from its function in human life of promoting social cooperation and from the features of human nature that make some ways of promoting social cooperation better than others (Wong 1996). The constraints might eliminate some moralities as unsuitable but nevertheless leave several acceptable moralities yielding conflicting, yet fully correct, moral judgments about particular cases. In such cases, when adherents of different and equally correct moral positions must deal with each other, a conception of morality as politics can again seem appropriate.

THE NEED FOR A MORAL VALUE OF ACCOMMODATION IN THE FACE OF SERIOUS DISAGREEMENT

The politics involved need not be unconstrained by moral values. One such value is accommodation, defined as a commitment to supporting noncoercive and constructive relations with others in spite of one's disagreements with them (Wong 1992). Such a value is likely to be present in a wide range of moralities because it is needed to manage the divisive effects of moral disagreement. Serious moral disagreement is common even in a group with relatively homogeneous moral beliefs. For example, the source of disagreement over abortion seems to be not so much a difference in ultimate moral principles held by the opposing sides, but partly a difference over the applicability of a commonly held principle requiring the protection of human life and partly a difference over the relative weight to be given to the widely held principle of protecting individual autonomy. Serious disagreement of this kind is

ubiquitous, and if it always threatened to become a source of schism, no society could survive for very long without brutal repression.

Even if accommodation is necessary for managing the divisiveness of disagreement, it does not have automatic priority over the moral values that are the source of serious disagreement. Abortion opponents may accept legal abortion in the early stages of fetal development to accommodate abortion-rights advocates, but be unwilling to accept compromise on abortion in the later stages of fetal development. And they may be unwilling even if they see their disagreement with those who advocate late-term abortion as irresolvable through the use of a common reason. Ultimately, there seems to be no useful general theory that specifies when to accommodate and when not to. It is a matter of judgment in the concrete situation, and also a matter of creatively devising courses of action that both honor one's own values and accommodate others who disagree. Some abortion-rights advocates and foes, for example, have joined common ground in efforts to prevent unwanted pregnancies.

ETHICAL RELATIVISM AND THE HIGH AMBITIONS OF MODERN MORAL THEORY

Ethical relativism has a deflating effect on modern normative theory, and especially on its highest ambitions. Consider Thomas Nagel's (1979) classification of five different sources of value: first, specific obligations to other people or institutions that depend on some special relation to the person or institution in question (e.g., motherhood); second, constraints on action deriving from general rights that everyone has, such as rights to liberty of certain kinds or freedom from assault or coercion; third, utility, the effects of what one does on everyone's welfare; fourth, perfectionist ends and values, such as the intrinsic value of scientific achievements and artistic creations; and fifth, commitment to one's own projects and undertakings. On Nagel's view, each of these sources is irreducible. The typical ambition of modern normative theory, however, has been to identify some of these sources as basic and the rest as derivative. Utilitarians, for instance, typically strive to reduce all moral considerations to utility, basing the assignment of individual rights and special obligations on utility. Absent reduction of moral considerations to one source, the next highest ambition is to specify general rankings of sources of value to settle cases of conflict between the several basic values. Deontologists, for example, might strive to specify how individual rights constrain the goal of maximizing utility.

The relativist views such attempts as preaching to the converted. There is no viable reduction of all moral value to one source or even a uniquely correct ordering of several irreducible values. The sources of value reflect the disorderly diversity of what human beings value and the various ways they have of structuring their relations with one another. Relativists hold that social convention plays an ineliminable role in selecting which values a group's morality emphasizes the most and in dealing with conflicts between important values. While relativists might disagree on the degree to which moralities are constrained by independent factors such as the functions of morality or the limits of human nature, they agree that independent constraining factors cannot validate the highest ambitions of modern moral theory. Moreover, even if we descend to particular cases of conflicts between values, the relativist will argue, contrary to Nagel, that there is no uniquely correct answer as to how to resolve such cases (or at least many of these cases) without collective and individual invention (see Walzer 1983, 1987; Wong 1996). If ethical relativism is true, much, if not most, of modern normative theory is wrong-headed, and this perhaps partly explains the intensely negative reaction that the doctrine elicits from many moral philosophers.

See also Metaethics; Noncognitivism; Objectivity in Ethics; Projectivism.

Bibliography

Harman, Gilbert. "Moral Relativism Defended." *Philosophical Review* 84 (1975): 3–22.

Harman, Gilbert. "Relativistic Ethics: Morality as Politics." *Midwest Studies in Philosophy* 3 (1978a): 109–121.

Harman, Gilbert. "What Is Moral Relativism?" In *Values and Morals*, edited by Alvin I. Goldman and Jaegwon Kim. Dordrecht, Netherlands: Reidel, 1978b.

Nagel, Thomas. "The Fragmentation of Value." In his *Mortal Questions*. Cambridge, U.K.: Cambridge University Press, 1979.

Walzer, Michael. *Interpretation and Social Criticism*. Cambridge, MA: Harvard University Press, 1987.

Walzer, Michael. *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books, 1983.

Williams, Bernard. *Morality: An Introduction to Ethics*. New York: Harper Torchbooks, 1972.

Williams, Bernard. "The Truth in Relativism." *Proceedings of the Aristotelian Society* 85 (1974–1975): 214–228.

Wong, David B. "Coping with Moral Conflict and Ambiguity." *Ethics* 102 (1992): 763–784.

Wong, David B. *Moral Relativity*. Berkeley: University of California Press, 1984.

Wong, David B. "Pluralistic Relativism." *Midwest Studies in Philosophy* 20 (1996): 378–400.

David B. Wong (2005)

ETHICAL SKEPTICISM

See *Emotive Theory of Ethics; Ethical Relativism*

ETHICAL SUBJECTIVISM

A subjectivist ethical theory is a theory according to which moral judgments about men or their actions are judgments about the way people react to these men and actions—that is, the way they think or feel about them. It follows that moral predicates are not possessed by actions or actors in the absence of people who pass judgments upon them or who respond to them with such feelings as admiration, love, approval, detestation, hate, or disapproval. It follows from this definition that nonpropositional or noncognitive ethical theories are not subjective, for according to them there are no moral propositions at all, and thus moral judgments cannot be propositions about people's feelings.

This definition is also intended to exclude views according to which moral judgments are judgments about how people behave; hence, views such as that "wrong" means "contrary to the accepted code of the society in which the action is performed" will not count as subjective, even if, as seems likely, statements about moral codes are statements about the prohibitory or permissory behavior of the communities possessing these codes. Although this distinction is not hard and fast (for statements about moral codes might often turn out to be statements about how the people possessing these codes think or feel), we shall maintain it in order not to trespass too far into the subject of ethical relativism, which is treated elsewhere. Elements of subjectivism can be found in so many ethical theories that it is almost impossible to give an account of them. The greatest, though not the most consistent, subjectivist was the eighteenth-century Scottish philosopher David Hume. The theory has also been popular among anthropologists, of whom Edward Westermarck was probably the most outstanding.

Subjectivist theories can provisionally be classified according to whether moral judgments are alleged to be about the speaker's thoughts or feelings, about the thoughts or feelings of some group of people, or about the thoughts or feelings of men as such.

MORAL JUDGMENTS STATE WHAT THE SPEAKER FEELS

The view that moral judgments are about the feelings of the person making the judgment—that what I mean when I say that an action is right or that a man is good is that the thought of that man or action evokes in me, personally, at this moment, a feeling of approval—has been subjected to an enormous number of objections. It has been argued that, so far from its being possible to identify moral judgments as those judgments that are about feelings of approval, it is in fact only possible to identify feelings of approval as those feelings that are evoked by the judgment that an action is right, and argued that if we feel approval of an action because we judge it to be right, our thinking that it is right cannot be identical with our approving of it. It has also been objected that if the theory is true, all I need do to settle any doubt I have concerning the rectitude of someone else's action is to introspect, and that it is very difficult, if not impossible, to be mistaken about one's feelings, although it is very easy to make a mistake about whether an action is right. More plausibly, it has been alleged that the theory implies that one can only criticize someone else's moral judgment on the ground that the other person is mistaken about how he himself feels.

Some of the worst difficulties for the theory arise from the fact that sentences offered as definitions of moral judgments contain such words as *I, now*, and *here*, whose reference depends upon who uses them, at what time, and in what place. From this it follows that one person's moral judgments can never be incompatible with any other person's moral judgments; the sentence "I do not feel disapproval of divorce," when used by one speaker, does not express a judgment incompatible with that expressed by the sentence "I do feel disapproval of divorce," uttered by a different speaker. It would appear, however, that when one person says, "Divorce is wrong," he does mean to say something incompatible with what someone else means when he says, "Divorce is not wrong." From the fact that moral judgments are alleged to have a covert reference to the feelings the speaker now has, it follows that if at one time he judges an action (say, Brutus's assassination of Caesar) to be right, his judgment is not incompatible with the judgment he may make at a later time when he judges Brutus's assassination of Caesar to be wrong. For according to this theory, what he meant on the first occasion was that he did not then feel disapproval of Brutus's assassination of Caesar, and what he meant on the second occasion was that he now does. Clearly there is no reason why both judgments should not be true.

G. E. Moore, in a famous argument, attempted to deduce that the subjectivist theory led to the paradoxical conclusion that the same action could be both right and

wrong, and that one and the same action could change from being right to being wrong. (It is important to remember that Moore thought classes of actions—for example, marrying one's deceased wife's sister—could change from being right to being wrong, that is, that an instance of a class of actions, performed at one time, might be right, while another instance of the same class of actions, performed at a later time, might be wrong.) First, Moore argued as follows. If Jones approves of Brutus's assassination of Caesar and says Brutus was right, it follows from the theory that Brutus was right. Similarly, if Smith disapproves of Brutus's assassination of Caesar and says Brutus was wrong, then Brutus was wrong. Hence Brutus was both right and wrong to assassinate Caesar. Second, to show that Brutus's assassination of Caesar can change from being right to being wrong, all Moore thought he need do was to point out that if Jones says (at a time when he approves of Brutus's action) that Brutus was right, then according to the theory, Brutus was right; if he later comes to disapprove of Brutus's action, then, if he says Brutus was wrong, according to the theory, Brutus was wrong. If Jones can truly judge at one time that Brutus was right and at a later time that Brutus was wrong, it must follow that Brutus's action has changed from being right to being wrong.

C. L. Stevenson has criticized Moore's argument in the following manner. Although Jones can truly say that Brutus was right to assassinate Caesar and Smith can truly say that Brutus's action was wrong, neither Jones nor Smith nor anyone else can say that this action is both right and wrong. For anyone to be able to say it is both right and wrong, someone would have both to approve of it and disapprove of it. Hence, although Jones, who approves of Brutus's action, can say it is right, and Smith, who disapproves of it, can say it is wrong, neither can say it is both right and wrong. Moore's mistake, perhaps, consisted in construing the theory we are considering as maintaining that "right" is a predicate like "disapproved of by *someone*" (from which it would follow that the same action can be both right and wrong), whereas "right" is alleged to be a predicate like "disapproved of by me, the speaker," from which it follows that the same action cannot be both right and wrong, since the speaker cannot both approve and disapprove, on the whole, of one and the same action.

A free exposition of Stevenson's criticism of Moore's argument (that if the view that moral judgments are statements about the speaker's feeling is true, one and the same action can change from being right to being wrong) might take the following form. Moore supposes that if ten years ago I could truly say Brutus was right to assassinate Caesar (because at that time the thought of this action did arouse approval in me) and now I can truly say that Brutus was wrong to do this (because at this time the thought of his action arouses disapproval in me), it follows that the action must have changed from being right to being wrong. Stevenson, however, points out that the statement "Brutus's action has changed from being right to being wrong" is equivalent to the conjunction of statements "Brutus's action was right a while ago" and "Brutus's action is now wrong." Although the truth of the second of these statements is entailed by the fact that I now feel disapproval of Brutus's assassination of Caesar, the first of them is not entailed by the fact that I earlier felt approval of Brutus's assassination of Caesar. Although Moore supposes "Brutus's assassination of Caesar *was* right" to mean "I *once* approved of Brutus's assassination of Caesar," what it actually means is "I *now* approve of Brutus's erstwhile assassination of Caesar," and, ex hypothesi, "I do not now approve of Brutus's action, I disapprove of it." Moore's mistake is to suppose that the word *was* in the sentence "Brutus *was* right to assassinate Caesar" shows that the sentence is about my past approval, whereas in fact its function is to show that it is the action I am disapproving of, not my disapproval, that is past.

Although Stevenson's detailed criticisms of Moore are valid, it is possible to restate Moore's arguments in a way that avoids them. To take the second of Moore's arguments first, it would plainly be absurd for me to say that Brutus was wrong to assassinate Caesar and at the same time to say that I was correct many years ago when I judged that Brutus was right to assassinate Caesar. If I now say he was wrong, I am bound to say that when, earlier, I said he was right, I was mistaken. In regard to the first argument, although it does not follow from this subjectivist theory that anyone can say that an action is both right and wrong, it does follow that if Jones says that Brutus was right to assassinate Caesar, he is not saying anything incompatible with what Smith is saying when he says that Brutus was wrong to assassinate Caesar.

Clearly, however, Jones and Smith think they are saying something incompatible, and it is unlikely that they have such a poor understanding of how to use their own language as to be mistaken on this point. In other words, according to the subjectivist theory, if Jones says Brutus was right to assassinate Caesar, Smith can say to Jones that Brutus was wrong and at the same time agree that Jones is making a true statement when he says that Brutus was right. This is absurd. Stevenson has tried to over-

come this particular difficulty by saying that although insofar as Jones and Smith are making assertions, their assertions are not incompatible with one another, they are doing something over and above asserting things—namely, Jones is trying to evoke in Smith an attitude of approval toward Brutus's action, and Smith is trying to evoke in Jones an attitude of disapproval toward Brutus's action. Hence, although their beliefs are compatible, their interests clash. Jones aims to achieve something that is incompatible with what Smith aims to achieve. A consideration of this view of Stevenson's, however, would take us away from subjectivism to a consideration of nonpropositional ethical theories.

The difficulties already mentioned may well be fatal to the type of subjectivist theory we are considering, at any rate as long as it is not bolstered by the nonpropositional theory. Moreover, there is a further difficulty that seems to settle the issue. Suppose Jones says that the death penalty for murder ought to be retained in Great Britain, and he says this because he wrongly supposes that abolishing the death penalty would lead to an increase in the number of murders. According to this kind of subjectivism, all Jones means when he says that the death penalty ought to be retained is that the thought of retaining it arouses in him feelings of approval, and since it does do this, his statement that it ought to be retained is true, however mistaken he may be about the facts of the situation. He is under no obligation to withdraw his statement, therefore, when he discovers his mistake. Again, this is absurd.

MORAL JUDGMENTS STATE THE SPEAKER'S THOUGHTS

So far we have considered the view that moral judgments are judgments to the effect that the action under judgment arouses certain feelings in the person making the judgment. It is possible, however, to think that what someone making a moral judgment is asserting is that the action or person being judged arouses in the person making the judgment certain thoughts or beliefs. The most natural view is that someone who asserts that an action is wrong is saying that the thought of the action arouses in him personally the belief that it is wrong, or in other words, that all we mean when we say that an action is wrong is that we personally think it is wrong. This view is circular because even though it offers a definition of "wrong" (that is, it maintains that "*X* is wrong" just means "I think *X* is wrong"), the word *wrong* still occurs in the definition (compare "'thoroughbred horse' means 'horse both of whose parents are thoroughbred horses'"). It is obviously impossible to get rid of the circularity by again substituting "thought wrong" for "wrong" in the definition, however many times we do it.

OBJECTIVE AND SUBJECTIVE SENSES OF "RIGHT"

It is quite commonly held that whenever one thinks one is acting rightly, one is acting rightly. Philosophers who hold this view, however, are not properly regarded as subjectivists. Clearly, if the word *right* is being used twice in the same sense when it is asserted that one is doing rightly if one does what one thinks is right, the view is contradictory. According to it, one would be acting rightly even if one did what one mistakenly thought was right: From the fact that one mistakenly thought it was right it follows that it is wrong, and from the fact that one is acting rightly if one does what one thinks is right, it follows that the act is right. Those philosophers who have held this view, however, have generally distinguished two senses of the word *right*, sometimes called an objective sense and a subjective sense, and have held that an action is right, in the subjective sense, if it is thought to be right in the objective sense. This removes both the contradiction and the suggestion that the property of being right depends on being thought to be right; the property of being subjectively right depends on the different property of being thought to be objectively right.

MORAL JUDGMENTS STATE WHAT A COMMUNITY FEELS

Next to be considered is the view that when individuals make moral judgments they are talking not about the way they themselves think or feel about the things they are judging, but of the way some group of people thinks and feels about these things. Presumably the group of people might be named by a proper name—for example, "Englishmen" or "Melanesians," or more plausibly (to avoid the difficulty that Englishmen cannot be supposed to be talking about the feelings of Melanesians and vice versa) by a descriptive phrase such as "my group" or "the community to which I belong." The theory that moral judgments are about the feelings of the speaker's own community is open to a large number of the objections to which the private reaction theory is open and a few more besides. Although two people, one of whom says a given action is right and the other of whom says that the same action is wrong, will be saying incompatible things if they belong to the same community, if they belong to different communities their statements will be perfectly compatible. Again, if a man says at one time that an action is right

and at a later time that the same action is wrong, there is no need for him to withdraw his first assertion, provided that the attitude of the community to which he belongs has changed during the interval. In that case there is no reason why both his first assertion and his second assertion should not be true. It follows, too, that however ignorant or mistaken a given community may be concerning the nature of the action being morally assessed, the statement by a member of that community that the action is right will be true as long as his community does approve of it. For example, if a community disapproves of giving eggs to pregnant women because it believes that this will cause them to give birth to chickens, the statement by a member of the community that it is morally wrong to give eggs to pregnant women will be a true one, because this community really does have the feelings it is alleged to have.

Over and above these quite fatal objections, the theory that moral judgments state how members of a community feel about the actions under judgment is exposed to two difficulties, to which the view that they state how the speaker feels is not exposed. Although one might just accept the conclusion, implied by the latter theory, that one discovers a given action is right by introspection (in Hume's language, by looking inside one's breast and discovering there a sentiment of approval or disapproval), it is quite impossible to accept the view that one discovers a given action is right simply by asking other members of one's community whether they approve of it. The theory also leads to the quite unacceptable consequence that anyone who believes, for example, that most of his community approves of retaining a law against homosexuality and at the same time believes that the law against homosexuality ought to be abolished believes two logically incompatible things; for, according to this theory, what he is believing when he believes that the law against homosexuality ought to be abolished is just that most of his community would feel approval of its abolition.

A variant of this last theory would be the view that what we mean when we say that an action is right is that it is approved of by the agent's community, not the speaker's. This is more plausible because it means that instead of condemning actions performed in distant places or times because they do not accord with the moral attitudes of our community, we may praise them because they do accord with the moral attitudes of the community to which the man who performed them belonged. Hence, it appeals to a tendency in some modern moral philosophers to be rather absurdly uncritical of the moral attitudes of communities other than their own. They may be less willing to accept this variant theory, however, when they realize that if it is true, they must be equally uncritical of the morals of their own community. It must be pointed out that according to this theory, if one says, "Introducing racial segregation into University *X* was wrong," one means that introducing segregation into the university was disapproved of by the community in which the action was performed. It is perfectly obvious, however, that one might perfectly well think that it was wrong and at the same time know that it was not disapproved of by the community in which it was performed. Again, this theory has the difficulty that even when a community disapproves of some practice only because the community is grossly mistaken concerning its nature (as in the case of the women and the eggs), we are still bound to say that the practice is wrong, providing the community does in fact disapprove of it.

MORAL JUDGMENTS STATE WHAT MOST PEOPLE FEEL

The objections that have already been raised against earlier types of subjectivism can be applied without much difficulty to the view that what we mean when we say an action is right is that most people approve of it. This theory does imply that any two individuals (whoever they are), one of whom condemns an action and the other of whom judges it to be right, really will be saying incompatible things, for it cannot be that most people approve of an action and at the same time disapprove of it. However, since people may change from approving of something to disapproving of it, the theory does entail that a man may judge an action to be right and later judge the same action to be wrong without having to retract his first judgment. The theory means that an action is wrong if most people feel disapproval of it, however ignorant or mistaken they may be about the nature of the action. It also means that it is impossible for a man to make up his mind concerning the rectitude of any action unless he has decided whether humankind in general would approve of it; it is obvious, however, that we can make up our minds on such questions without having the least idea what the attitude of most men would be. In view of what has already been said about the theory that what we mean when we say that an action is right is that the speaker thinks it is right, nothing need be said about the analogous views that an action is right if the speaker's community thinks that it is right, or that an action is right if most people think that it is right.

It is fairly obvious from what has been said that all subjectivist theories need to be amended, at least to

exclude the possibility that the attitude of the people we are alleged to be describing when we make moral judgments is not based on ignorance or mistake. Hence, a consideration of subjectivism may lead to the view that an action is right not if it is approved of by any actual person or group of people, but only if it would be approved of by a person of a very special kind—for instance, one who, at the very least, is never ignorant of or mistaken about any relevant matter of fact concerning the action toward which his attitude of approval is directed. Hence, a consideration of subjectivism must inevitably lead to a consideration of ideal observer theories; these, however, are best treated as a variety of ethical objectivism.

See also Emotive Theory of Ethics; Error Theory in Ethics; Ethical Naturalism; Ethical Relativism; Metaethics; Moral Skepticism; Noncognitivism.

Bibliography

Boyd, Richard. "How to Be a Moral Realist." In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord, Ithaca: Cornell University Press, 1988.

Brandt, Richard. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Brink, David. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Carritt, E. F. *Ethical and Political Thinking*. Oxford: Clarendon Press, 1947.

Carritt, E. F. *The Theory of Morals*. London: Oxford University Press, 1928.

Edwards, Paul. *The Logic of Moral Discourse*. Glencoe, IL: Free Press, 1955. An acute discussion of various kinds of subjectivism.

Ewing, A. C. *The Definition of Good*. New York: Macmillan, 1947. Ch. 1.

Firth, Roderick. "Ethical Absolutism and the Ideal Observer," *Philosophy and Phenomenological Research* 12 (1952): 317–345.

Foot, Philippa. "Does Moral Subjectivism Rest on a Mistake?" *Oxford Journal of Legal Studies* 15 (1995): 1–14.

Harman, Gilbert. "What Is Moral Relativism?" In *Values and Morals*, edited by Alvin Goldman and Jaegwon Kim. Dordrecht: Reidel, 1978.

Harman, Gilbert, and Judith Thomson. *Moral Relativism and Moral Objectivity*. Oxford: Blackwell, 1996.

Hume, David. *Inquiry concerning the Principles of Morals*, edited by C. W. Hendel. New York: Liberal Arts Press, 1957. Ch. 1, Appendix I. The classical criticism of objectivism.

Hume, David. *Treatise of Human Nature*, edited by L. A. Selby-Bigge. Oxford: Clarendon Press, 1888; reprinted 1955. Bk. III, Pt. I, Secs. I–II.

Johnston, Mark. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* Suppl. (1989): 139–174.

Kant, Immanuel. *The Moral Law, or Kant's Groundwork of the Metaphysics of Morals*. Translated by H. J. Paton. London: Hutchinson, 1948.

Lewis, David K. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* Suppl. (1989): 113–137.

McDowell, John. "Values and Secondary Qualities." In *Morality and Objectivity*, edited by Ted Honderich. London: Routledge & Kegan Paul, 1985.

Moore, G. E. *Ethics*. London: Williams and Norgate, 1912. Chs. 3–4. An important criticism of subjectivism.

Plato. *The Republic*. Translated, with an introduction, by A. D. Lindsay. New York, 1907; paperback ed., 1958. Bk. I.

Price, Richard. *A Review of the Principal Questions in Morals*, edited, with an introduction, by D. Daiches Raphael. Oxford: Clarendon Press, 1948. Excellent defense of objectivism.

Ross, W. D. *The Foundations of Ethics*. Oxford: Clarendon Press, 1939. This work and the following are two modern versions of Price's theory.

Ross, W. D. *The Right and the Good*. Oxford: Clarendon Press, 1930.

Shafer-Landau, Russ. *Moral Realism*. Oxford: Clarendon, 2003.

Smith, Michael. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* Suppl. (1989): 89–111.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.

Stevenson, C. L. "Moore's Arguments against Certain Forms of Ethical Naturalism." In *The Philosophy of G. E. Moore*, edited by P. A. Schilpp. Evanston, IL: Northwestern University Press, 1942. A reply to some of Moore's criticisms.

Westermarck, Edward. *Ethical Relativity*. New York: Harcourt Brace, 1932. A well-known version of subjectivism.

Wiggins, David. "A Sensible Subjectivism." In *Needs, Values, Truth*. Oxford: Basil Blackwell, 1987.

Jonathan Harrison (1967)
Bibliography updated by Mark van Roojen (2005)

ETHICAL THEORY

See *Metaethics*

ETHICS

Ethics is the branch of philosophy that tries to understand a familiar type of evaluation: the moral evaluation of people's character traits, their conduct, and their institutions. We speak of good and bad people, the morally right or wrong thing to do, just or unjust regimes or laws, how things ought and ought not to be, and how we should live. One part of the subject, metaethics, is concerned with what such judgments mean, what, if anything, they are about, whether they can be true or false, and if so what makes them true or false. The other part of the subject, normative ethics, is concerned with the content of those judgments: What features make an action right or wrong; what is a good life; and what are the characteristics of a just society? This entry will concentrate on

normative ethics, though some comments on metaethics will be unavoidable. And within normative ethics it will concentrate on general principles and foundations (what is usually called moral theory) rather than on applied ethics, the discussion of specific cases. Moral theory seeks a systematic understanding of the full range of moral convictions and disagreements and of their possible grounds.

WHAT IS MORALITY?

Morality of some kind seems to be a universal human phenomenon; it is a subpart of the broader domain of the normative, which seems also to be characteristically human. Normative questions and judgments are about what we ought to do, want, believe, or think (rather than just about what we actually do), and about the reasons for and against doing or believing one thing rather than another. Only rational beings, and probably only beings with language, are capable of normative thinking.

Many normative questions are not moral. If we ask whether we ought to believe on the basis of the available evidence that a painting is by Rembrandt, that is a normative question, but not a moral one. Moral questions are about what we ought to do and how we ought to live, not about what we ought to believe. In answering them we need to appeal to what are called practical reasons—reasons for doing or wanting something—rather than the purely evidential or theoretical reasons that determine the justification for factual or scientific belief.

But not all practical reasons and practical norms are moral, either. There are norms that tell you what you ought to do to keep your rose bushes healthy, the right way to make an omelet, or what to wear if you are going to be knighted by the queen, but these are not moral judgments. The moral is a subpart of the large normative domain of the practical.

Its further definition is the subject of controversy among different moral theories, but a rough approximation is this: Morality identifies certain norms that apply to everyone in a certain group and that should be recognized as valid for everyone by each member of the group although their separate individual aims and desires may differ and lead them into conflict with one another. In most, but not all modern moral theories, the group to which moral norms apply includes all mentally competent human beings. In such theories morality is conceived as consisting of universal norms.

Morality aims to provide us, in the practical domain, with a common point of view from which we can come to agreement about what all of us ought to do. This may be different from what each of us might want to do or want other people to do, if we looked at the question only from our own personal point of view. Morality tries to discover a more objective standpoint of evaluation than that of purely personal preference.

Much of the content of moral norms has to do with our relations to each other—how we treat each other in our individual conduct and how they are treated by collective institutions that we support. There are also moral norms and evaluations having to do with the way we conduct our own lives, norms that tell us how anyone may succeed or fail in living well. Virtues like prudence and self-control are examples of this universal but partly self-regarding aspect of morality. There are also important moral questions about our relations to the rest of the natural world, especially to other animals. But the bulk of the subject has to do with our lives as members of the human community.

Though there are ethical relativists who disagree, moral rules such as those that condemn murder, injury, lying, stealing, and betrayal and endorse kindness, honesty, and generosity are usually thought to apply universally. Whether this can be shown is one of the big questions of ethics, but such norms are not supposed to depend for their validity on the code of a particular society or group or the laws of a particular government. They are not like local codes of etiquette, taboos, or specific traffic or commercial regulations. Even the wrongness of a crime like murder does not depend on its being against the law. Rather, moral norms are supposed to be recognizable by a form of thought that is available in principle to any adult human being—even though some people may be better at it than others.

At the center of morality are standards that serve not only the interests of a particular individual who follows them but also the collective good of the community, by making it a safer place or otherwise promoting the general welfare of its members. But for those standards to do their work, most people have to adhere to them. Normative ethics tries both to identify such standards and to explain how individuals, even though their interests diverge in many respects, can be attached to universal norms that serve the common good.

OBJECTIVE REASONS OR SUBJECTIVE FEELINGS?

Most moral theories ascribe a kind of objectivity to moral judgments, because such judgments are supposed to issue from a point of view that different people can share and

that enables them to arrive at agreement about what should be done, what would be wrong, what would be fair, and so forth. Even when people disagree about such things, they usually share the belief that there is a right answer to the question, though they do not agree about what it is. They do not think of their moral disagreements simply as a divergence of personal preferences.

However, there is controversy about the exact nature of the convergence of judgments that we try to reach in moral thinking and about its source. This is one of the main questions of metaethics: Are moral judgments based on universally valid reasons, which would permit them to be objectively true or false, or do they merely express widespread subjective feelings that many people share? In the latter case, while we can come to agree in these judgments, they do not make claims that are either true or false; rather, they express certain attitudes or responses—though perhaps responses on which all human beings can converge.

One of the most important defenders of the subjective alternative is the eighteenth-century Scottish philosopher David Hume (1739, 1751). He argues for the claim that moral judgments express a special type of feeling, sentiment, or attitude on the ground that this is needed to explain how moral norms, like other practical norms, are capable of motivating people to act.

For example, if someone judges that it would be wrong to leave a campsite littered with garbage, this will probably move him or her to take the trouble to clean it up before he or she leaves. If we assume further that motivation must always start from some desire or feeling of the agent, it seems to follow that morality must get its motivational force from something of that kind—for example, from a sympathetic aversion to the unhappiness of others. For how could the recognition of any truth revealed by reasoning or thought alone, without the help of a desire, have the motivational consequences of a normative judgment?

Defenders of the objective alternative usually hold a different view of motivation. They are likely to maintain that while feelings and desires are often the source of motivation, there is also a form of practical reasoning that is capable of motivating rational persons on its own, through the recognition of existing reasons alone. If you decide, after considering the effect on others, that it would be wrong to leave the campsite a mess, you recognize that you have a reason to clean it up, and that will lead you, if you are a decent person, to do so. On this view the motive is produced by recognition of the norm, rather than the norm being the mere expression of a preexisting feeling or motive.

This opposition between the view that moral judgments express subjective feelings and the view that they express objective normative beliefs capable of being true or false has many different forms and subtle variations, but it is present everywhere in ethical theory, and in some cases it plays an important role in disputes over the normative content of morality, although it is primarily an issue of metaethics. It is also important in discussing the question whether moral standards have a universal basis, or whether they are really culturally relative. On the subjective or expressivist theory the relativist conclusion is not necessary, but it seems more of a possibility than on the theory of objective moral reasons.

MORALITY AND SELF-INTEREST

One of the great questions of ethics is whether, and if so to what extent, morality requires us to subordinate our individual self-interest to the general good. There is one view, whose most important representative is the seventeenth-century English philosopher Thomas Hobbes (1651), according to which morality does not conflict with self-interest because its requirements actually derive from self-interest in a subtle way.

Hobbes's argument is that certain rules of conduct are necessary for human beings to live at peace with one another and to enjoy the benefits of civilization, because if people do not abide by those rules they will fall into a miserable condition of insecurity and violence. They are the rules of morality, prohibiting murder, assault, theft, fraud, breach of contract, and so forth, and it is in every individual's self-interest to live in a society in which they are followed. General adherence to morality serves the collective self-interest of all the individual members of any community.

This alone is not enough, however, to show that private adherence to those rules is in the individual self-interest of each member of the community. That general adherence is in the collective interest of all does not imply that individual adherence is in the personal interest of each, because an individual cannot by his or her own conduct bring it about that others will act in the same way. What would serve the collective self-interest therefore does not necessarily coincide with what people would be led to do by their individual self-interest.

Hobbes thinks that reasoning allows us to see that collective self-interest would be served by general adherence to the rules of morality, but that individual self-

interest makes it irrational to follow those rules on our own, since that would simply permit others to take advantage of us. He concludes that to bring the rule of morality into effect, it is necessary to provide all individuals with incentives guaranteeing that individual and collective self-interest will coincide. This can be done only by a system of law, enforced by a sovereign with a monopoly of force over the members of the community. Only then will it be safe for each person to follow the rules, knowing others will also follow them because it is likewise in their personal interest to do so.

According to this theory self-interest does not motivate us to abide by the requirements of morality directly. If we could get away with it, self-interest would lead us to prefer that everyone else followed the rules, while we ourselves were exempt from them. But that alternative is not available, and it would be much worse for each of us if no one followed the rules. So the uniform solution that serves all of our interests best is that everyone follow the rules and that a system of incentives be set up to ensure that no individual can do better for himself or herself by breaking them.

There are also theories descended from that of Hobbes that preserve the connection between morality and self-interest but do not rely for stability only on external incentives produced by the enforcement mechanisms of a legal system. The Canadian American philosopher David Gauthier (1986) proposes that some of the work of bringing individual and collective self-interest to coincide can be done by internalizing the moral rules, so that individuals are inhibited against breaking them even apart from the threat of punishment. Feelings of guilt, for example, are a kind of emotional self-punishment that people who have internalized the rules inflict on themselves when they break them. It is in our collective interest for each member of the community to be subject to such feelings, because a community in which the moral norms have been internalized in this way provides its members with the benefits of mutual trust and peaceful cooperation.

In views of this kind there is already a departure from the reliance exclusively on self-interest to motivate moral conduct, even though morality is thought to serve the interest of each of its adherents. But many moral theories put a much greater distance than this between morality and self-interest. In different ways, most modern accounts of morality part company with Hobbes and base the appeal of moral norms on a concern for everyone, not just for oneself. This means that morality may sometimes require the individual to sacrifice his or her own interests for the good of others or to avoid transgressing the rights of others. That poses the question of the nature of the reasons or motives that can outweigh self-interest in these cases. If you can make a gain by harming someone else, why shouldn't you do it?

CONSEQUENTIALISM AND UTILITARIANISM

One important type of answer to this question is that everyone's life is just as important or valuable as everyone else's, and in particular your happiness is no more valuable than other people's happiness. Therefore, you have a reason to care impartially about what happens to everybody, and in your conduct should try to promote the general good and not just your own.

This depends on an important distinction between two ways in which things can be good or bad: They can be good or bad for someone, or they can be good or bad, period. If something is good for me, that obviously gives me a reason to want and promote it, but it does not obviously follow that anyone else has a reason to want and promote it, unless it is also good for him or her. But if something is simply good, period, then it is something anyone has a reason to want and promote.

Some things, like health and pleasure, are clearly good for the person who has them. And it is possible to hold the view that this is the only kind of value there is—value for someone. On this view there may be things, like the destruction of the ozone layer, that would be bad for everyone, but even that would not make it bad, period. Most ethical theories, however, hold that some of the things that are good or bad for individuals, like pleasure and pain, or happiness and unhappiness, are also good or bad without qualification—objectively good or bad, one might say. And each person has a common reason to promote what is good and to prevent what is bad in this way—not only what is good or bad for him- or herself. Some theories derive the content of morality entirely from such a conception of objective value—maintaining that morality emerges once we recognize the objective value and disvalue of the occurrence of all those things that are good or bad for individuals.

Utilitarianism is the most fully developed version of such a theory. A version of it is found in Hume; it was further developed by the English philosophers Jeremy Bentham (1798), John Stuart Mill (1863), and Henry Sidgwick (1907) and continues to be influential. Utilitarianism holds that morality is simply the specification of those forms of conduct that contribute most effectively to

the greatest overall happiness for all persons—or all sentient beings—impartially considered.

The basic value on which the whole theory depends is some measure of what is good in the lives or experiences of particular individuals—what is in itself desirable for them, and therefore also objectively valuable. This may be pleasure and the avoidance of pain, happiness and the avoidance of unhappiness, the satisfaction of their desires or preferences and the avoidance of their frustration, or perhaps some other measure, depending on the particular version of utilitarianism. Whatever the measure, it must be roughly quantifiable in a way that allows comparison between persons, and addition and subtraction of the amounts of the value to determine the total that is present in complex cases involving many people with different experiences.

This measure of value is called utility (a technical term—in this context the word does not mean "usefulness"). Utilitarianism is the theory that we should act, and organize our institutions, so as to promote the maximum total amount of utility, weighing the utility that arises in the lives of all persons impartially. This is sometimes crudely expressed by the formula "the greatest good for the greatest number."

What matters, according to this view, is not the quality of our actions themselves but the utility, as measured for example by general happiness, of the overall outcome of what we do, compared with the available alternatives. For this reason utilitarianism is an example of a consequentialist moral theory. It is results that matter, not the means by which we reach them. So an important feature of consequentialism is that it does not make a fundamental moral distinction between positive and negative responsibility for good or bad outcomes.

For example, one is positively responsible for someone else's suffering if one deliberately hurts that person, but negatively responsible if one fails to save him or her from being harmed. According to utilitarianism this alone makes no moral difference between the two cases: Morality does not require merely that you not harm people; it makes you equally responsible for the prevention of harm, from whatever cause. It even requires you to cause harm if that is the most effective way to bring about the greatest overall balance of good. If there is a moral difference between harming and failing to prevent harm, it must be due to some difference in the utility of the results.

Another significant feature of utilitarianism is that what matters in determining the rightness or wrongness of actions is the total utility that results, not how it is distributed among individuals. In calculating the total we add together or aggregate quantities of utility from different lives, and the sum of many small amounts of pleasure or pain from different people's experiences may add up to much more utility than the intense pleasure or pain of one individual.

Working out the details for principles of conduct and political, social, and economic institutions depends on estimates of the likely results of the various alternatives, and combining probabilities and utilities to arrive at a measure of what is called expected utility. This is often uncertain and difficult. But the ultimate moral foundation is simple: What matters is that people should have good experiences and avoid bad ones, and the higher the overall balance of good minus bad, the better. The aim of morality is to tell us how to maximize the amount of good in the world, where good is measured objectively and impartially, so that our own personal good is no more important a part of the total than anyone else's.

RIGHTS, OBLIGATIONS, EQUALITY, AND DESERT

Some familiar aspects of ordinary moral thought do not seem to conform to the utilitarian standard. One of the most controversial issues in moral theory is whether those aspects can be explained by utilitarianism, and if not, whether we should conclude that utilitarianism is false or that those aspects must be rejected.

Apparently, counterutilitarian moral norms are those that either require or permit a course of action or policy that fails to maximize utility. One type of example is found in the large and diverse category of individual rights, which include rights against certain kinds of interference or violation by others, and rights to do what one wishes so long as one does not violate the rights of others.

For example, it seems to be widely accepted that each individual has a right not to be killed, injured, enslaved, kidnapped, or imprisoned if he or she is not hurting anyone else—even when violating one of those rights would be useful as a means to producing a large net balance of benefits overall. Even if someone else's life could be saved by forcibly taking one of your kidneys and transplanting it to him or her, this would not be morally acceptable if you did not consent to it. For another example, each of us is generally thought to have a right to devote most of our resources and attention to our own life and the lives of our friends and family, even if we could do more good

overall by dedicating ourselves to the general welfare of everyone.

Other examples come from the sphere of special obligations, both those that arise from particular undertakings, like contracts and promises, and those that follow from standing conditions like citizenship and family membership. Conventional morality holds that one is obligated to keep a promise even if marginally more good than harm would be done by breaking it, that one should give special weight to the welfare of one's children, and so forth.

In ordinary moral reflection on social policy and public institutions, considerations of fairness seem to be sensitive not to the total aggregate welfare produced, but to the distribution of benefits and disadvantages among individuals. A distribution that produces greater total welfare at the cost of great inequality between rich and poor may be morally inferior to one with a lower total but less poverty and more equality of opportunity.

Finally, in thinking about the criminal law, the allocation of punishments seems to be justified not merely by what would produce the most utility, through deterrence and prevention, but also by the requirement that only those who are guilty be punished, and that the punishment deserved is proportional to the gravity of the offense. (The idea of moral desert brings up the large question of free will and moral responsibility. There are those who doubt that people can be responsible for their actions in a way that would mean they deserve punishment for wrongdoing, so that punishment can be justified only as a deterrent. But that issue is beyond the scope of this discussion.)

What these familiar moral ideas have in common is that they do not appear at first sight to interpret the right as what will maximize the overall good for individuals. They seem to rely instead on independent standards for what is right and wrong—standards that either permit or require certain types of actions even if we believe they will not produce the greatest impartial benefit. Such standards set certain moral limits on what we may do to other people and impose certain positive requirements as well, including moral requirements that must be met by the institutions of government. But they leave us morally free to lead our lives as we wish within those boundaries, without having to take the promotion of the general good into account in all our choices.

Standards of this kind (often referred to as deontological standards by contrast with the consequentialist variety) seem to require a foundation different from the impartial concern for the interests of all, which is the basis of utilitarianism. But before discussing what that foundation might be, it is necessary to consider the utilitarian response.

ACT-UTILITARIANISM AND RULE-UTILITARIANISM

Hume holds that all of morality can be accounted for by its tendency to promote utility, but that this works in two different ways. In some cases the relation of a morally good or bad act to utility is direct, as in the case of kindness or cruelty. These he describes as examples of the natural virtues and vices—types of conduct that increase or decrease utility act by act, through their direct causal effects.

But there is another set of moral requirements, which he calls the artificial virtues, where the effects on utility are not necessarily produced by each morally good act taken alone. Instead, the good effects are produced only by a general rule, convention, or practice, and it is one of the conditions of the utility of rules of this kind that they must be followed even in individual cases where the particular action they require is harmful to utility.

The utilitarian explanation of strict rights, obligations, and duties depends on this type of analysis, which is called rule-utilitarianism—by contrast with act-utilitarianism, which assesses the rightness of actions by their effects on utility taken one by one. For example, the institution of stable property rights, without which a functioning economy would be impossible, requires that property owned by one person should not be subject to appropriation by another person whenever the latter can get more utility from it than the former. A landlord has to be able to charge his or her tenants rent, even though they may need the money more than he or she does, or else no one would invest in rental property. The great utility of the general rules of property depends on their being consistently followed even in cases where violating them would advance utility, since that is the only way to ensure security and stability.

Likewise, the institution of promises has great utility because it makes it possible for people to rely on each other's future conduct and to create such reliance. But it can do so only because it is not permissible to break a promise whenever this would produce more utility than keeping it.

To some extent the utilitarian advantage of such rules can be obtained by embodying them in laws of property and contract that are enforced by the courts. But

the rules also seem to have moral weight apart from such enforcement: Violation of property rights and breach of promise seem wrong in themselves, and the rule-utilitarian explanation is that they are wrong because they violate valuable institutions or conventions.

Similar explanations can be offered of why it is morally permissible for individuals to live their lives without making every decision on the basis of how they can contribute the most to maximizing utility for humanity as a whole. The reason is that so much of human happiness depends on the pursuit of personal aims and fulfilling personal relationships, and a strict requirement that every act must strive to maximize general utility would make personal projects, friendships, and commitments impossible. In other words, a world governed by strict act-utilitarianism would be a world with much lower overall utility than a world in which not every action aimed to maximize impartial utility.

These are only some examples. The rule-utilitarian strategy can be applied to a wide variety of apparent exceptions to utilitarianism, including rights of bodily integrity, the requirement that punishment be deserved, the right to freedom of speech, the special obligation of parents toward their children, and the values of political, social, and economic equality.

Still, it is not clear just how much of the apparently counterutilitarian morality of rights, obligations, and permissions can be accommodated by rule-utilitarianism in this way. For example, the Australian philosopher Peter Singer (1972, 1979), a prominent utilitarian, argues that in the very unequal world in which we live, there is no justification for the moral latitude most well-off people in rich countries assume they have to favor themselves and their friends and families, when their resources could bring so much more benefit to the destitute in impoverished countries. In Singer's view utilitarianism should be seen as a radical position that cannot be used to underpin conventional morality, but requires that it be overturned.

There is also a theoretical problem about the relation between rule-utilitarianism and moral motivation. The problem is that, if a utilitarian is attached to property rights and the obligation of promises because of the contribution of those institutions to general utility, that does not explain what his or her reason is for abiding by the rule in an individual case that clearly does not serve utility. The utilitarian may say that there is a strong moral reason to want the institution of promises to exist. But if breaking a promise in a particular case will not cause the institution to disappear, or even weaken it noticeably, and if he or she can thereby produce more benefit than harm, why should the utilitarian's moral aim of maximizing utility not lead him or her to conclude that breaking the promise is the right thing to do in that case?

Some utilitarians are prepared to accept this conclusion. This is the act-utilitarian position, according to which laws, conventions, and practices may change the circumstances in ways that affect what acts will best promote utility, but can never make it right to do what one knows will not produce the most benefit.

Others believe that, since utility is best served if individuals have internalized a strict attachment to certain rules so that they are unwilling to break them even to promote utility, this creates an independent reason for adherence in such cases. In a sense, the utility of the rule provides a justification for the moral fiction that there is a reason to act contrary to utility in the particular case.

KANTIAN CONTRACTUALISM

The main rival to a consequentialist foundation for rights and special obligations is a theory that emphasizes the separate importance of each individual person instead of the value of maximizing the sum of benefits to the aggregate of all persons. According to this alternative the aim of morality is to find principles of conduct under which people are given equal consideration, not merely as elements in an aggregate, but as individuals. This would mean that the apparently counterutilitarian character of individual rights, for example, is real and cannot be explained away by rule-utilitarianism.

The most important representative of this type of theory is the eighteenth-century German philosopher Immanuel Kant (1785, 1788), who holds that moral principles can be identified directly by reference to a single standard, which he calls the categorical imperative. (He calls it *categorical* to indicate that its application to an individual is not conditional on what that person happens to want; the reasons provided by morality do not depend hypothetically on interests or desires, but apply categorically and unconditionally to all persons simply in virtue of their rationality.)

The categorical imperative says, roughly, that we should act only on principles that we would want everyone to act on. It is often referred to as the standard of universalizability, since it means that each of us should govern our conduct by principles that we would be willing to see followed universally. But if this test is to identify a single set of moral principles that apply to everyone, there must be a way to decide what principles we would want everyone to follow that will not give different

answers for different people depending on their interests and situation. That implies that in answering the question we must try to take into account the point of view of every person simultaneously, putting ourselves in the place of each of them, and rejecting those principles that could not be accepted by everyone.

The tradition deriving from the categorical imperative is sometimes called contractualism because it identifies moral principles through an imaginary agreement: They are the principles whose adoption by everyone would not be unacceptable to anyone. The results of such a test may be much less determinate than the utilitarian standard, but it does seem to imply some major differences from utilitarianism. First, the insistence on separate acceptability to each individual will rule out justifications that depend on aggregation of small benefits across many lives to outweigh a large cost to a single individual. Second, in deciding what principles are and are not universally acceptable, the determining factor will have to be some system of priorities among the things that matter in human life, and the effects that different principles would have on each person, as measured by these values.

MODERN CONTRACTUALISM

One result will be that in the application of moral standards to social policy, there will be a direct reason to concentrate on the relief of poverty and improvement in the condition of the worst off, not merely as a means of improving the total or average welfare, as in utilitarianism. This is a feature of the American philosopher John Rawls's (1971) theory of justice.

Another result is that the justification for individual rights will be different from that offered by rule-utilitarianism. The right not to be killed, injured, or deprived of liberty even if it would promote the general welfare will depend not on the overall balance of costs against benefits for all people affected by the existence or nonexistence of such a right, but on the importance for each separate individual of the security that such a right provides, by comparison with the advantages for each individual that its absence might make possible.

The right to pursue one's personal aims, interests, and attachments rather than the general welfare in most of what one does will depend not on the effect of such a right on the general welfare, but on the importance for each person of this kind of freedom by comparison with the value for each person of the possible benefits of its general restriction.

The emphasis is on providing certain protections and basic benefits to everyone equally rather than maximizing the overall sum of benefits. This is a fundamental difference in the approach to the foundation of morality, a difference in the way in which the interests of all persons are combined from a moral point of view.

Modern successors to Kant attempt to make the standard more precise in different ways. Rawls claims that what is wrong with utilitarianism is that it does not take seriously the distinction between persons. Writing not about morality in general but about social justice, he embodies the contractualist ideal in an imaginary choice called the Original Position, in which people are supposed to choose the principles of justice for their society without knowing who they are; this forces them to choose principles that would be acceptable whoever they turned out to be. Though influenced by Rawls, T. M. Scanlon (1998), another philosopher in the contractualist tradition, proposes a different test. He maintains that to identify standards of right and wrong we must search for principles that no one seeking to arrive at common standards of interpersonal justification could reasonably reject, knowing both his or her own situation and that of others.

Unlike a consequentialist theory, the contractualist method cannot proceed simply by calculating the total expected costs and benefits of different rules of conduct or forms of political and social organization. Rather, it must evaluate the priorities among different kinds of costs and benefits, for each individual, of living under alternative rules or systems. Which principles and practices are morally acceptable will depend on these priorities, applied equally to everyone.

For example, both utilitarianism and contractualism condemn slavery, but they do so for different reasons. Utilitarianism says slavery is wrong because the total misery of slaves vastly outweighs the total benefit to slave owners. Contractualism says slavery is wrong because any reasonable person thinking about his or her own or any other life must regard the avoidance of the possibility of being a slave as having strict priority over the possibility of enjoying the advantages of being a slave owner.

DEONTOLOGY: DOING AND ALLOWING

Not everyone who believes in rights and special obligations thinks they have to be justified by either contractualism or rule-utilitarianism. The general term for these apparently nonconsequentialist parts of morality is deontology, and there is an alternative ethical tradition, called

intuitionism (represented, for example, by the English philosopher W. David Ross [1930]), according to which the deontological elements of morality are fundamental. They do not derive from anything else, but they reveal themselves to reflection about what would and would not be the right thing to do in different cases.

On this view it is evident that we may not kill an innocent person to save five others (e.g., by harvesting the first person's organs for transplantation), and there is no more fundamental explanation of why we may not: It would be murder, that is all. The details of these moral requirements are sometimes complicated, but they can be discovered by exercising moral intuition in respect to real and imaginary cases that bring out the relevant distinctions.

One of the most important of these distinctions, mentioned earlier, contrasts the things we do to other people, for which we are positively responsible, and the things that happen to other people that we might have been able to prevent, for which we are only negatively responsible. If we kill an innocent person to transplant his or her organs to five others, for example, we would be positively responsible for the death of that person. But if we do not kill that person and the other five die of organ failure, we are not positively responsible for their deaths and have not violated their right to life. This means that the prohibition against murder must include some specification of the way in which one person's conduct has to be related to another person's death for it to count as wrong.

Different accounts have been offered of this relation. It might seem that what matters is whether your action causes the death or whether it is caused by something or someone else. But this turns out to be wrong in two ways. First, you may cause a death as an unavoidable side effect of something else you do, but if you were acting to save many more lives, you are not to blame. For instance, if you are the pilot of a plane that is about to crash, and you steer it from a densely populated area to a sparsely populated area, you are causally responsible for the deaths of a smaller number of people but you are not to blame, because it was a side effect of your aiming the plane away from the larger number.

Second, you may be to blame for a death that you didn't cause but could have prevented, if you deliberately failed to act to ensure that the death would occur. For example, if you let an otherwise healthy patient with asthma choke to death so that you can harvest his or her organs to save five others, you have intentionally allowed the patient's death—aimed at it even without causing it—in a way that makes the action wrong.

So the element of intention—intentionally causing or permitting someone's death either as an end in itself or as a means to something else—is an important part of wrongful killing. And rights in general have to be understood as rights against the intentional imposition of harms of various kinds.

In a morality of this kind, we are not generally responsible for preventing what is bad and promoting what is good. Morality is defined instead by a set of constraints against the intentional imposition of harm or violation of rights, plus some well-defined and limited positive obligations—like keeping our promises and taking care of our families.

Instead of deriving the content of morality from a point of view that tries to take everyone's interests into account—either a consequentialist or a contractarian point of view—intuitionism understands morality as setting a kind of boundary around each person, that protects us from intentional violation and interference by others. Positive obligations are also understood individually, as arising from the specific commitment undertaken by a promise or a contract, explicit or implied, with another person.

AGENT-NEUTRALITY AND AGENT-RELATIVITY

The difference between deontological and consequentialist moral theories can also be described in terms of a formal distinction between two kinds of principles or reasons: agent-relative and agent-neutral.

An agent-relative principle specifies what each individual should do in a way that involves an ineliminable reference to that agent himself or herself, or his or her situation—even when the principle is stated in its most general form. For example, the principles "Everyone may give priority to their own interests over those of a stranger," "Everyone should do what is best for their family," "Everyone should keep their promises," and "Everyone should refrain from killing innocent people" are all agent-relative.

The following principles, however, are all agent-neutral: "Everyone should promote the general welfare," "Everyone should promote the stability and devotion of families," "Everyone should try to minimize the breaking of promises," and "Everyone should aim to minimize the killing of innocent people." Agent-neutral principles depend on the objective value of certain kinds of hap-

penings or states of affairs, without regard to their relation to the agent. All that matters is whether the agent is in a position to affect the occurrence or nonoccurrence of the desirable or undesirable outcome. If the value attaches to a type of action, such as murder, an agent-neutral principle would not distinguish between a murder that the agent commits and one that someone else commits and that the agent could prevent. Accordingly, the principle that everyone should aim to minimize the occurrence of murders could authorize committing one murder to prevent several others.

For this reason, deontological principles naturally take an agent-relative form. They tell each individual what he or she may, must, and must not do, without giving all individuals a common outcome or state of affairs that they must try to promote. Deontological principles are universal, but the aims they assign to each individual always depend on his or her situation and are related to him or her. This logical feature unites the three aspects of deontology: deontological prohibitions—"Don't (you) commit murder"; deontological requirements—"Keep your promises"; and deontological permissions—"You can enjoy your life instead of devoting it to the service of humanity."

The exercise of moral intuition on different cases reveals a surprisingly detailed system of deontological principles on which many people can agree and that form a large part of conventional morality. But the view that there is no systematic foundation underlying these diverse principles, that their truth cannot be explained by something more basic, leaves many moral philosophers dissatisfied.

What they want is a general foundation for deontology to rival the clarity of consequentialism. Since it seems obvious that there is always a reason to prefer better results, deontologists need to explain in a clear fashion why morality often prohibits actions that would have the best overall results and permits other actions that would not have the best overall results. If promoting the best consequences is not, as utilitarianism maintains, the governing standard of morality, then it would be good to know what is.

VIRTUE

Contractualism is the most prominent foundational alternative to consequentialism, and it works by offering an alternative interpretation of what it is to treat all persons with impartial respect. But there is another way of criticizing consequentialism, and that is to attack its foundation directly, by denying the moral authority of the impartial point of view.

The criticism goes like this: Ethics is concerned with how people should live and what they should do, and the point of view from which we should seek an answer to that question is the point of view of the individual, not an impersonal point of view that takes into account all individuals at once. Even if this yields moral requirements on how one should treat other people, they must arise from considerations about how one has reason to live one's life and what kind of person one wants to be.

One version of this approach takes as basic the question: What is the difference between a good and a bad person? Once we know the difference between a virtuous and a vicious character, we can identify the morally right thing to do as what a virtuous person would do in the circumstances. This way of understanding the subject is found in Aristotle's *Nicomachean Ethics*.

The reason we all have to care about virtue, on this view, is not an impartial concern for others, but that being a good person is an aspect of being a good human specimen—analogous to physical health and being in good physical condition. To be virtuous is to function well with respect to feelings, desires, motives, and actions, including interactions with other people. Moral virtue, like good physical functioning, is part of the good for each individual, and it has as elements the distinct virtues such as courage, temperance, prudence, generosity, honesty, and justice. Each of these is a set of motivational dispositions and dispositions to choose that lead to virtuous conduct.

Some of the virtues, like courage and temperance, are good partly because they enable the individual to pursue his or her own aims effectively. But a virtue like justice is good for the individual because people are essentially social beings and must be able to live in harmony with others. This conception of ethics leaves the content of interpersonal morality rather vague. Instead of principles of conduct, it offers a rough indication of the types of motivational and behavioral dispositions, recognizable by example in the character of virtuous individuals, to which everyone has reason to aspire, simply in order to be a good person. But at least this account, even if it does not start from impartiality, offers a kind of harmony between the interests of the virtuous individual and the interests of the community to which he or she belongs.

RESISTANCE TO IMPERSONAL MORALITY

A more skeptical challenge to the impartial standpoint comes from the English philosopher Bernard Williams (1981, 1985). He argues that impersonal moral theories, whether consequentialist or contractualist, are incompatible with the integrity of an individual life, which is found in the unconditional commitment to particular projects and particular persons. Such commitments would be impossible if impersonal values were permitted to take precedence over them.

This is most forceful as a response to utilitarianism. Even if, from an impersonal standpoint, everyone's life is just as important as everyone else's, that is simply not true from your individual standpoint, and the impersonal standpoint has no authority on its own to overrule the standpoint of the individual. Ethics is supposed to govern individual conduct, so it must find its basis in the motivation of individuals. This may include some impartial values, but it also includes much else. For most people, life gets its substance and meaning from aims and attachments that are inseparable from the personal point of view. These cannot simply be abandoned when it turns out that there is something impersonally more valuable that one could do with one's life.

Utilitarians can reply in either of two ways. They may say that in rejecting the demands of impersonal value, Williams is simply rejecting morality, and that the whole point of morality is to replace the natural selfishness of individuals with an impartial perspective. Nobody said it would be easy. Alternatively, they may emphasize the ways in which utilitarianism takes into account the point of view of the individual, since it is the source of the happiness whose maximization over all persons utilitarianism takes as the aim of morality.

However, even after we take this second point into account, it is clear that utilitarianism, including rule-utilitarianism, will under some circumstances require the radical subordination of individual aims to the general welfare. There is an important difference of opinion here over what morality can reasonably demand of us.

The conflict between Williams's objection and contractualism is less stark, but here, too, he claims that it is incompatible with the nature of basic personal commitments to subordinate them to the test of what could be universalized, or what could be reasonably agreed to by everyone as a principle of conduct. Even to say, for example, that it is permissible to devote yourself to your children because you find it acceptable that everyone should favor their own children is inconsistent with the immediate and unconditional nature of your attachment to your own children. It is, in Williams's phrase, "one thought too many." Williams's resistance to the ultimate authority of the objective, impersonal standpoint is partly inspired by the more radical resistance to impartiality of Friedrich Nietzsche (1897), the great nineteenth-century German critic of Christianity and moral universalism. Contractualists like Scanlon reply that it does not denigrate the independent force of personal attachments and projects to require that they be embedded in a moral framework that sets limits to their pursuit, since the desire to live on mutually acceptable terms with others is such an important human value that it must be allowed to shape other, more personal values.

The question of the relative weight and interaction between personal motives and the claims of impartiality in determining the content of morality is a fundamental one, and it generates continuing controversy. Uncompromising utilitarians like Singer maintain that the commonsense morality that most people accept and that strictly limits their responsibility to sacrifice their own interests and aims for those in greater need is much too undemanding. If we really take seriously the undeniable fact that other people's suffering is just as bad as our own, we will have to change our lives.

In contrast, defenders of more conventional morality hold that while it is admirable to be self-sacrificing, it is also supererogatory—that is, it is morally praiseworthy but goes beyond what is morally required. They maintain that utilitarianism, by holding people morally accountable for anything that happens for which they are negatively responsible, leaves them with an unacceptably diminished control over their lives.

These disputes pose the question of whether or to what extent the content of morality should depend on a prior assessment of the human motives available to induce people to live in accordance with its requirements. There is a division of opinion between those who think morality has to rely on preexisting motives and those who think it can create new motives, by revealing specifically moral reasons we all have to act in certain ways.

Among the first group we find Hobbes, who derives morality from redirected self-interest, and Hume who derives it from a natural moral sentiment arising from sympathy for the happiness and unhappiness of others. Among the second group is Kant, who believes that the recognition that an intention cannot be universalized will itself motivate us to refrain from acting on it. He holds

that the recognition of moral principles, without an antecedent desire, is enough to create a motive.

Even if morality introduces new motives, it may be important to take into account human nature, including natural self-interest and natural personal attachments, in constructing a workable moral code. The question then becomes: How can humans who are not naturally impartial live together in a way that acknowledges that objectively none is of more intrinsic value than another?

RELATIVISM

So far this entry has discussed ethical theories that offer different accounts of morality conceived of as a single system—that is, as a set of general standards that will allow us to determine what is right or wrong for any person to do, in any society. That does not mean that the same specific actions will be morally required of everyone, since in different circumstances, different forms of conduct may satisfy or violate the same universal moral standards. For example, with overpopulation and environmental degradation, activities like deforestation and the unrestrained burning of fossil fuels, which on a small scale were once harmless, can become dangerous to future generations, and therefore wrong.

But is it also possible that the basic moral standards themselves should vary over time, or from culture to culture? The view that morality, even at the most fundamental level, is not universal but arises from local cultures or conventions that may vary is called ethical relativism. Relativism is not the view that there is a single overarching and universal moral principle, namely: "Follow the moral conventions of your culture"; nor is it the view that some other universal principle, such as utilitarianism, implies that it is always best to follow the conventions of the culture in which you find yourself. Relativism is the position that the true and ultimate source of moral standards is always a set of rules, practices, and attitudes shared by a historically situated community. While not everyone in the community will obey the rules, and some may reject them, there is often enough of a consensus about what the rules are to make it possible to identify them.

Morality, on this interpretation, is closer to etiquette or law than it is on the universalist interpretation. Naturally, there will be some overlap among moral systems—all of them can be expected to condemn murder and theft in some form, for example. But slavery, the subordination of women to their male relatives, polygamy, or homosexuality may be morally wrong in some societies and morally unquestionable in other societies, depending on the prevailing norms.

On this view it is probably a mistake to say that slavery in ancient Rome was wrong, that bullfighting in Spain is wrong, or that the subordination of women in Saudi Arabia is wrong. There is no universal, timeless standpoint from which to make these judgments. They would have to be defended from a standpoint internal to the cultures that they are about, and if that cannot be done, they should be abandoned.

That does not mean that it is impossible to criticize morally what a society does, for its conduct may sometimes violate its own moral principles. It may also sometimes be the case that there is no prevailing moral code in a particular culture, especially during periods of social transition or upheaval.

These qualifications mean that it will not always be easy for a defender of relativism to identify the standards that apply in a particular society. But relativism at least clearly rules out the attempt to appeal to universal standards. Even an internal moral critic of a society—someone who says, against the general consensus, that slavery is morally wrong—would be mistaken if he or she were making a universal claim. The critic has to be understood as trying to change the standards or as finding an inconsistency between one part of the prevailing standards and another part.

Relativism has the consequence that we must dismiss as confused certain judgments that we are strongly inclined to make, which implicitly or explicitly appeal to universally valid or objective standards in morality. They include judgments about societies other than our own, whose standards we think are mistaken, or judgments about our own society, whose present standards we think may be mistaken and may be rightly rejected by later generations.

Relativism cannot account for the apparent fact that when an individual rejects the moral standards that prevail in a culture, either from within or from outside, he or she may not be simply applying the standards of an alternative culture, but may be appealing to deeper moral reasons, such as unfairness to some members of the society or failure to give certain interests their true weight. Such arguments point to deeper and more general standards by which local conventions can be assessed.

This is connected with the question of the motivation for being moral. If morality is based on custom, the motivation for conforming to it is in a sense shallow. To be moral is to have internalized the patterns of conduct that prevail in one's surroundings. If, however, morality is not relative but universal, this means that the motives

that attach us to moral norms must be deeper, and theories of the foundations of morality must try to identify them.

MORALITY AND RELIGION

There is a way of defending the universality and objectivity of moral truth different from those that have been discussed so far. That is to claim that moral standards are laid down by divine command.

If this means that nothing would be right or wrong unless God declared it to be so—that "if God does not exist, everything is permitted"—then it is not a plausible view. Even if God does command that we not kill, lie, steal, and so on, it seems more plausible to hold that he forbids those things because they are wrong, not that they are wrong because he forbids them. (A polytheistic version of this point is made in Plato's dialogue *Euthyphro*.)

Though we can understand how divine command might establish specific requirements like dietary restrictions or forms of worship—where we are obliged to follow them simply out of obedience—the ordinary standards of morality seem different: They seem to depend on the intrinsic features and effects of certain kinds of conduct rather than on something external to them. We can understand what is wrong with murder without reference to God.

On the other hand, it may be possible to preserve a version of the divine command theory by referring to the characteristics of God, as all-knowing, all-good, and loving the world and his creatures. The rules that a divine being enjoins people to obey might in that case be said to be correct in virtue of the features of God's nature that lead him to choose those rules. But it also means that he chooses them for characteristics that themselves make them correct and that he could not have commanded a different morality.

Religion is sometimes thought to play another role, as the guarantor of an incentive to be moral through divine punishment and reward in an afterlife. The afterlife also serves a direct moral purpose in allowing us to hope that the world is not fundamentally unjust and that the virtuous will be rewarded, however much they may have suffered on this earth.

However, most modern moral philosophy has not depended on religion, but has tried to interpret ethics in secular terms. Those who believe that God commands our adherence to moral standards usually hold that we use our independent understanding of those standards in forming our idea of God's will. An exception is John Locke (1690), for whom the assumption that God gave the earth to human beings in common plays an important role in moral and political theory.

ETHICS, POLITICS, AND LAW

One of the main applications of moral theory is to evaluate political and social institutions—institutions like representative democracy or the market economy—as well as the more specific actions of government. But there are two different ways of thinking about politics from a moral point of view.

The first way is to start by identifying moral standards that apply to everything, and then to figure out what they imply for the special and complex case of political institutions and political life. This is the method favored both by utilitarianism and by the radical form of individual rights theory called libertarianism. Utilitarianism holds that the right way to evaluate anything, from an individual action to a form of government, is by reference to the value of its overall consequences for the total welfare of all persons, impartially considered. Libertarianism, on the other hand, determines the rightness or wrongness of individual actions and social institutions alike solely by reference to whether they violate or protect the natural rights of individuals not to be harmed, to exercise their freedom, and to acquire and transfer property. Because these theories hold that a single moral standard governs everything from individual conduct to the design of large social institutions, they are sometimes known as monist theories.

By contrast, other theories (sometimes called dualist—the terms are due to Liam Murphy [1999]) hold that ethics is more complicated than this and that different standards are appropriate for the regulation of different kinds of thing. According to this second approach it is a mistake to assume that the moral evaluation of institutions should derive from the same norms that govern individual conduct.

An important example of the dualist approach is Rawls's (1971) theory of justice. Rawls maintains that while private individuals should be free to pursue their own aims in life and favor their own and their families' economic interests, the basic institutions of a society must be much more impartial toward the interests of all its members. The social structure should be designed with the aim of providing equality of opportunity for all, and with the aim of reducing social and economic inequality by raising the condition of the worst-off class as much as possible. These strongly egalitarian values, according to Rawls, apply not to the personal interactions

of individuals, but to the design of the common institutions, imposed and sustained by state power, that provide the unchosen public framework for their private lives.

Whether one is a monist or a dualist, politics and law are important subjects for ethical theory. Politics poses in the starkest way questions about how to combine the conflicting interests of many different people who are affected by an institution, law, or policy. It poses questions about the relations between different values—the value of life, of liberty, of prosperity, and of freedom from coercion and violation of different kinds. It poses crucial questions about the possibility of outweighing harms to some by aggregate benefits to others.

The fundamental division between consequentialist and contractualist approaches shows up here. Consequentialists will not take the protection of individual rights as basic, but will regard it as an instrumental means for the promotion of the general welfare. Contractualists, by contrast, will find reasons to limit the power of the state over the individual in a separate and untradeable concern for each individual's autonomy and inviolability, regarded not merely as an element in the general welfare whose total is to be maximized.

Followers of Hobbes will hold that the only legitimate ground for state action is the provision of goods that are in the collective self-interest of all the citizenry, such as police protection, defense, economic stability, and public health. Utilitarians, on the contrary, will also favor policies that increase the total welfare, even if it means redistributing resources from the rich to the poor. Contractarians will give priority to the protection of individual rights, securing equal opportunity, and raising the social minimum. Libertarians will favor the minimum of government needed to keep the peace, protect individual rights, and secure private property. Therefore, many familiar political disagreements have a moral dimension and require that we ask how much and what kind of consideration we owe to the interests of our fellow citizens through our common institutions.

The distinction between monist and dualist theories comes up again when we ask whether the same principles that govern the moral acceptability of political institutions inside existing states should also be applied to our relations to people in other societies, indeed to the world as a whole. If our most fundamental moral duties to everyone are the same, then the division of the world into separate societies with special responsibility for their members is a historically understandable contingency, but it may or may not be morally acceptable. On the other hand, moral standards for the world as a whole may be different from those appropriate within a particular society. The question of the moral evaluation of the overall world order is a vital and wide open question.

BOUNDARIES OF THE MORAL COMMUNITY

This entry has been discussing moral standards as if they concerned our relations to other human beings, present and future. But there are other candidates for moral consideration: most notably, other sentient creatures who are not members of our species and human organisms not yet born—embryos and fetuses. We can leave aside the value of plants and other parts of nature because it seems separate from morality, just as aesthetic value seems to be in a different category. Morality is especially concerned with how we treat one another, but it probably goes beyond this to include our treatment of beings or creatures that are sufficiently like us in relevant respects.

The first question is whether sentience itself—the capacity to have conscious experience and to feel pleasure and pain—is sufficient to bring a creature under the protection of morality. On the utilitarian theory the answer is a clear yes. Pleasure is good and pain is bad, wherever they are found, so it is right to promote the first and avoid the second in all sentient creatures.

It may be difficult to compare the quantity, quality, and value of pleasures and pains across different species, so it is not always easy to calculate what actions or policies would maximize overall utility. There may also be, in some versions of utilitarianism, forms of human pleasure or happiness that are not available to other animals, and whose value counts heavily in calculating the total to be maximized. That is maintained by John Stuart Mill in the theory of higher and lower pleasures. But many utilitarians maintain that the widely prevalent treatment of animals in factory farming and slaughter for food, as well as in much scientific experimentation, is morally unacceptable.

It is doubtful that nonhuman creatures could be excluded from moral consideration entirely, except by an ethical theory based entirely on self-interest. Since other creatures do not threaten us and cannot enter into cooperative engagements with us, we have no reasons of self-interest to adopt ways of living at peace with them.

If we accept an other-regarding consequentialist basis for ethics, animals will certainly be included under its protection, but there may be different requirements on our treatment of animals from those on our treatment of people. Avoidance of suffering is likely to be the main

thing, and limits on killing or on infringement of liberty will probably depend on whether they lessen suffering.

It is less clear how contractarian theories can handle the moral status of creatures who cannot be imagined as participants even in a hypothetical agreement on standards of conduct. But perhaps this could be done through some system of imaginary representation of their interests (Scanlon [1998] discusses this issue).

The moral status of unborn humans is a different question. If we separate it from religious doctrine about when the soul enters the body, it becomes a question about whether the potential to develop into a fully conscious human being confers on an embryo or fetus some part of the moral protection due to such a being after it is born. Answers range from the position that the embryo has all the moral rights of an infant, specifically the right not to be killed, from the time of conception, when its genetic constitution is determined, to the position that it has no moral claims before the child is born alive and may therefore be disposed of at the discretion of the pregnant woman. In between are views that as the fetus develops toward viability it becomes gradually more and more difficult to justify interrupting the pregnancy deliberately—that is, the reasons for doing so have to be progressively stronger.

The difficulty of these boundary questions reveals uncertainty about the true foundations of ethics, but that is not surprising. Human morality is a constantly developing system of norms, and its philosophical investigation by ethical theory is an indispensable part of the process.

See also Applied Ethics; Consequentialism; Decision Theory; Deontological Ethics; Divine Command Theories of Ethics; Ethics, History of; Game Theory; Kantian Ethics; Metaethics; Teleological Ethics; Utilitarianism; Value and Valuation; Virtue Ethics.

Bibliography

Adams, Robert Merrihew. "A Modified Divine Command Theory of Ethical Wrongness." In *The Virtue of Faith and Other Essays in Philosophical Theology*, 97–122. New York: Oxford University Press, 1987.

Anscombe, G. E. M. *Collected Philosophical Papers*. Vol. 3, *Ethics, Religion, and Politics*. Minneapolis: University of Minnesota Press, 1981.

Aristotle. *Nicomachean Ethics*.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. London: Printed for T. Payne, 1789.

Brandt, Richard B. *A Theory of the Good and the Right*. New York: Oxford University Press, 1979.

Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1977.

Foot, Philippa. *Virtues and Vices, and Other Essays in Moral Philosophy*. Oxford, U.K.: Blackwell, 1978.

Fried, Charles. *Right and Wrong*. Cambridge, MA: Harvard University Press, 1978.

Gauthier, David. *Morals by Agreement*. New York: Oxford University Press, 1986.

Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press, 1990.

Hare, R. M. *Freedom and Reason*. Oxford, U.K.: Clarendon Press, 1963.

Hare, R. M. *Moral Thinking: Its Levels, Method, and Point*. New York: Oxford University Press, 1981.

Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press, 1977.

Hobbes, Thomas. *Leviathan, or, The Matter, Forme, and Power of a Common-Wealth Ecclesiasticall and Civill*. London: Printed for Andrew Ckooke, 1651.

Hume, David. *An Enquiry concerning the Principles of Morals*. London: Printed for A. Millar, 1751.

Hume, David. *A Treatise of Human Nature*. 1739.

Kagan, Shelly. *The Limits of Morality*. New York: Oxford University Press, 1989.

Kagan, Shelly. *Normative Ethics*. Boulder, CO: Westview Press, 1998.

Kamm, Frances M. *Morality, Mortality*. 2 vols. New York: Oxford University Press, 1993–1996.

Kant, Immanuel. *Critique of Practical Reason*. 1788.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. 1785.

Korsgaard, Christine M., with G. A. Cohen. *The Sources of Normativity*, edited by Onora O'Neill. Cambridge, U.K.: Cambridge University Press, 1996.

Locke, John. *Second Treatise of Government*. 1690.

Mackie, John L. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

Mill, John Stuart. *Utilitarianism*. London: Parker, Son, and Bourn, 1863.

Moore, George Edward. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Murphy, Liam. "Institutions and the Demands of Justice." *Philosophy and Public Affairs* 27 (4) (1999): 251–291.

Nagel, Thomas. *The Possibility of Altruism*. New York: Oxford University Press, 1970.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Nietzsche, Friedrich. *The Genealogy of Morals*. New York: Macmillan, 1897.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Clarendon Press, 1984.

Plato. *Euthyphro*.

Plato. *Gorgias*.

Quinn, Warren. *Morality and Action*. New York: Cambridge University Press, 1993.

Railton, Peter. *Facts, Values, and Norms: Essays toward a Morality of Consequence*. New York: Cambridge University Press, 2003.
Rawls, John. *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press, 1971.
Rawls, John. "Two Concepts of Rules." *Philosophical Review* 64 (1) (1955): 3–32.
Ross, W. David. *The Right and the Good*. Oxford, U.K.: Clarendon Press, 1930.
Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press, 1998.
Scheffler, Samuel. *Human Morality*. New York: Oxford University Press, 1992.
Scheffler, Samuel. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. New York: Oxford University Press, 1982.
Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1907.
Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1) (1972): 229–243.
Singer, Peter. *Practical Ethics*. New York: Cambridge University Press, 1979.
Smart, J. J. C., and Bernard Williams. *Utilitarianism: For and Against*. Cambridge, U.K.: Cambridge University Press, 1973.
Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge, MA: Harvard University Press, 1990.
Thomson, Judith Jarvis. *Rights, Restitution, and Risk: Essays, in Moral Theory*, edited by William Parent. Cambridge, MA: Harvard University Press, 1986.
Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.
Williams, Bernard. *Moral Luck: Philosophical Papers, 1973–1980*. New York: Cambridge University Press, 1981.
Wolf, Susan. "Moral Saints." *Journal of Philosophy* 79 (8) (1982): 419–439.

Thomas Nagel (2005)

ETHICS, DEONTOLOGICAL

See *Deontological Ethics*

ETHICS, HISTORY OF

The term *ethics* is used in three different but related ways, signifying (1) a general pattern or "way of life," (2) a set of rules of conduct or "moral code," and (3) inquiry *about* ways of life and rules of conduct. In the first sense we speak of Buddhist or Christian ethics; in the second, we speak of professional ethics and of unethical behavior. In the third sense, ethics is a branch of philosophy that is frequently given the special name of metaethics. The present discussion will be limited to the history of philosophical or "meta" ethics, for two reasons. First, because it is impossible to cover, with any degree of thoroughness, the history of ethics in either of the first two senses. Practices and the codification of practices are the threads out of which all of human culture is woven, so that the history of ethics in either of these senses would be far too vast a subject for a brief essay. Second, although ethical philosophy is often understood in a broad way as including all significant thought about human conduct, it can well be confined within manageable limits by separating purely philosophical thought from the practical advice, moral preaching, and social engineering that it illuminates and from which it receives sustenance. This distinction, while somewhat artificial, makes sense of the common opinion that philosophy in general, and ethical philosophy in particular, was invented by the Greeks.

The central questions of philosophical ethics are: What do we or should we mean by "good" and "bad"? what are the right standards for judging things to be good or bad? how do judgments of good and bad (value judgments) differ from and depend upon judgments of value-neutral fact? But when these questions are answered, it is important to find out the differences between specific types of value judgments that are characterized by such adjectives as *useful, right, moral,* and *just.* We may therefore divide our subject matter into the search for the meaning and standards of good in general, and of well-being, right conduct, moral character, and justice in particular. Needless to say, these are not watertight compartments. Many philosophers reject sharp distinctions between them. But provisional separation of these topics, subject to reunification in accordance with particular philosophical views, will prove helpful in disentangling the various issues on which philosophers have taken opposing stands, so that the history of ethics can be seen as irregular progress toward complete clarification of each type of ethical judgment.

GREEK ETHICS

Ethical philosophy began in the fifth century BCE, with the appearance of Socrates, a secular prophet whose self-appointed mission was to awaken his fellow men to the need for rational criticism of their beliefs and practices.

Greek society of the fifth century was in a state of rapid change from agrarian monarchy to commercial and industrial democracy. The religious and social traditions that had been handed down from one generation to the next through the natural processes of social imitation and household training were brought into question by the accession to power of a commercial class, whose mem-

bers were untrained in and scornful of the ancestral way of life. New rules of conduct were required by a market economy in which money counted more than noble birth and in which men had to be considered equals as buyers and sellers. Men who wished to be elected to public office, but had not been trained at home as rulers of serfs and household servants, needed a more explicit and general code of conduct than was embodied in the sense of honor and esprit de corps of the landed aristocracy. Occurring with the rapid political and social transformation of Greece, and interacting with it as both cause and effect, was the development of basic industrial arts and a scientific technology. These forces both expressed and intensified the developing interest in rational evaluation of beliefs. As Henry Sidgwick put it:

> This emergence of an art of conduct with professional teachers cannot thoroughly be understood, unless it is viewed as a crowning result of a general tendency at this stage of Greek civilization to substitute technical skill for traditional procedure…. If bodily vigour was no longer to be left to nature and spontaneous exercise, but was to be attained by the systematic observance of rules laid down by professional trainers, it was natural to think that the same might be the case with excellences of the soul. (*Outlines of the History of Ethics,* p. 21)

Early Greek thinkers drew frequent comparisons between medicine and ethics, describing ethics as the "art of living" and the "care of the soul." Socrates' motto, "A sound mind in a sound body," suggests the medical image of ethics as mental hygiene. Many thinkers took a special interest in medicine, and, recognizing the interdependence of mind and body, they practiced a rudimentary psychiatry. Alcmaeon of Croton, Empedocles, and Democritus were renowned for their psychotherapeutic skills. This biological conception of mind and soul led to a more critical and scientific approach to problems of ethical judgment. Philosophers began to search for reasons for established modes of conduct and, where no reasons were found, to suggest that action could be directed toward individual goals in defiance of tradition. The professional teachers known as Sophists, whose social role was to prepare the uncultivated nouveaux riches for positions of power in the rising democracies, employed the newfound weapon of logic with devastating effect against the code of honor of the declining aristocracy. Protagoras, Gorgias, and Thrasymachus taught methods of self-advancement and of attaining virtue. They stressed the difference between subjective values and objective facts, arguing that good and evil are matters of personal decision or social agreement (*nomos*) rather than facts of nature (*phusis*).

SOCRATES. Socrates stood midway between the unexamined, traditional values of the aristocracy and the skeptical practicality of the commercial class. Like the Sophists, he demanded reasons for rules of conduct, rejecting the self-justifying claim of tradition, and for this reason he was denounced as a Sophist by conservative writers like Aristophanes. But unlike the Sophists, he believed that by the use of reason man could arrive at a set of ethical principles that would reconcile self-interest with the common good and would apply to all men at all times.

The central questions of ethical philosophy were raised for the first time by Socrates and the Sophists, but only Socrates realized the difficulty, bordering on impossibility, of finding adequate answers. In this respect, Socrates may be regarded as the first philosopher, in the strictest sense of the term. While the Sophists, after exposing the impracticality of traditional rules of conduct, then offered glib formulas in their place—such as "Justice is the rule of the stronger" (Thrasymachus) and "Man is the measure of all things" (Protagoras)—Socrates applied the same logical criticism with equally devastating results to both aristocratic and marketplace morality. He did not find the universal and self-evident code he searched for, but it was his memorable achievement to have revealed to humankind that without such a code its actions will lack justification and that moral perfection is therefore an ideal to which we can only approximate. Perfect clarity about what constitutes moral perfection is no more of this world than is moral perfection itself.

Our knowledge of Socrates is primarily derived from the dialogues of Plato, so it is not possible to draw a sharp line between the ideas of the two men. But since Plato's early dialogues are considerably different in style and content from those that he wrote later in life, one may take the early as fairly representative of Socrates and the late as more expressive of Plato's own thought. The chief differences discernible are the following: The more Socratic dialogues are devoted to the criticism of conventional beliefs and to the demonstration of the need for further inquiry, while the later dialogues argue for positive conclusions; the early dialogues search for definitions of ethical concepts, while the later dialogues are concerned with justifying a contemplative way of life in which pleasures of the senses are spurned in favor of pleasures of the mind; finally, the Socratic style is conver-

sational and argumentative, while that of the later years is more didactic and abstract.

The Socrates of the early dialogues raises questions about the meaning of ethical terms, such as "What is justice?" (*Republic*), "What is piety?" (*Euthyphro*), "What is courage?" (*Laches, Charmides*), "What is virtue?" (*Protagoras*). The answers offered by others to these questions are then subjected to a relentless cross-examination (Socratic dialectic), exposing their vagueness and inconsistency.

Although Socrates did not separate judgments of value from judgments of fact, the negative results of his line of questioning suggest a distinction that was made explicit only in modern times by David Hume and G. E. Moore. In each of his discussions of ethical concepts such as courage or justice, Socrates refutes all efforts to define them in terms of ethically neutral facts. For example, when, in the *Protagoras, Laches,* and *Charmides,* courage is defined as resolute facing of danger, Socrates observes that a man who faces dangers that he would be wise to avoid is a fool rather than a hero. The generalization toward which Socrates points the way, although he does not arrive at it himself, is that ethical concepts can never be adequately defined in terms of observable facts alone. Many philosophers, beginning with the Sophists, have believed that this principle leads to ethical skepticism. Plato attempted to escape such skepticism by means of his theory of Forms, and the modern school of intuitionism proposes a similar way out. Indeed, all the ethical theories developed since Socrates may be considered as alternative explanations of the relation between facts and values, naturalistic theories stressing their interdependence and nonnaturalistic theories stressing their differences. Socrates, in demanding rational grounds for ethical judgments, brought attention to the problem of tracing the logical relationships between values and facts and thereby created ethical philosophy.

PLATO. Plato's thought may be regarded as an endeavor to answer the questions posed by Socrates. From the *Republic* on through the later dialogues and epistles, Plato constructed a systematic view of nature, God, and man from which he derived his ethical principles. The foundation of this metaphysical view was the theory of Forms, whose most succinct formulation may be found in the discussion of the Divided Line, toward the end of Book VI of the *Republic.* Plato divides the objects of knowledge into two main categories and each of these into two subcategories symbolized by unequal sections of the line. The main division is between the realm of changing, sensible objects and that of unchanging, abstract forms. Knowledge of sensible objects acquired by sense perception is inaccurate and uncertain, for the object of sense, like the river of Heraclitus, is in continual flux. In contrast, knowledge of timeless forms is precise and rigorously provable. The realm of sensible objects is subdivided into shadows and images, in the lower section, and natural objects in the upper section. The realm of forms is subdivided into mathematical forms and ethical forms. At the apex of this ascending line is the Form of the Good, in relation to which all other objects of knowledge must be defined if they are to be adequately understood. Thus, ethics is the highest and most rigorous kind of knowledge, surpassing even mathematics, but it is also the most difficult to attain. Mathematics leads us away from reliance on visual images and sense perception, and ethical philosophy demands an even greater effort of abstraction. The objects of ethical knowledge are even less visualizable than geometrical forms and numbers—they are concepts and principles ultimately unified under the all-encompassing concept of the Good.

Although Plato suggests in this and other passages that ethical truths can be rigorously deduced from self-evident axioms, and thus introduces the mathematical model of knowledge that has guided many philosophers ever since, he does not employ a deductive procedure in his discussions of specific ethical problems, perhaps because he did not feel that he had yet attained an adequate vision of the Good that would supply him with the proper axioms from which to deduce rules of conduct. His actual procedure follows what he calls an ascending dialectic, a process of generalization through the give and take of conversation and the consideration of typical cases, a process designed to culminate in an intellectual vision of the structure of reality, from which, by a "descending dialectic" or deduction from general principles, particular judgments of value can be deduced. Plato's main goal in his ethical philosophy is to lead the way toward a vision of the Good.

The Socratic-Platonic ethical theory identifies goodness with reality and reality with intelligible form and thus concludes that the search for value must lead away from sense perception and bodily pleasure. This suggests an ascetic and intellectualistic way of life that is spelled out in full detail in the *Republic,* in the description of the training of the guardians. Some difference in the degree of intensity of the preference for mind over body may perhaps be discerned in the increasing severity of tone from the early dialogues to the later. In the *Protagoras* and *Symposium,* Socrates argues for rational control over the

body for the sake of greater pleasure in the long run, but he does not oppose pleasure as such. In the *Symposium* the unity of body and mind is a luminous thread throughout the discussion. Love is regarded as a search for the pleasure that consists in possession of what is good, and it is shown to exist on many levels, the lowest being that of sexual desire and the highest that of aspiration toward a vision of eternity. While still under the influence of Socrates, Plato distinguishes noble pleasures from base pleasures, rather than condemning pleasure in itself. The image he draws of Socrates is of a man who eats and drinks heartily and enjoys himself on all levels of experience, but in rationally controlled proportions. Socrates enjoys the wine at the symposium as much as anyone else, but unlike the others he remains sober to the end. While the poet Agathon becomes drunk with his own rhetoric, Socrates employs richly sensual language and metaphor in a way sufficiently controlled to make a philosophical point and so remains master of his rhetoric as well as of his body.

In the extraordinarily beautiful dialogue *Phaedo,* which describes the day of Socrates' execution, the theme of superiority of soul to body is dealt with directly, as might be expected of a philosopher who is about to die. Here Socrates commits himself unequivocally to a rejection of the body and its pleasures, maintaining that a wise man looks forward to his own death, when the soul is freed from its corporeal prison. Whether this is an exact expression of Socrates' attitude toward life may, however, be doubted in view of other dialogues, such as the *Protagoras.* In any case, it is natural for a man confronting death to try to set the best possible light on it. But it was this more somber, otherworldly strain in Socrates that Plato in his later works elaborated into a mystical vision of a timeless higher world. Plato has Socrates say, in the *Philebus,* "no degree of pleasure, whether great or small, was thought to be necessary to him who chose the life of thought and wisdom" (translated by B. Jowett, New York, 1933, Para. 33).

In the *Timaeus,* where, significantly, the protagonist is no longer Socrates but the Pythagorean Timaeus, pleasure is described as "the greatest incitement to evil," and Timaeus places the "inferior soul" below the neck, separating it from the intellect. Plato's severe castigations of bodily pleasures, his sharp separation of soul from body and of the eternal from the temporal, and his mystical cosmology entail a more extreme asceticism than that preached or practiced by Socrates.

Plato's mistrust of bodily pleasure and perceptual judgment led him to take an unfavorable view of public opinion and, consequently, of democratic institutions. In the *Republic,* and still more emphatically in the *Laws,* he proposed that society be ruled by an intellectual elite who would be trained to govern in accordance with their vision of eternal forms. He proposed, in the *Laws,* a ruthless system of punishments and the propagation of ideologically useful myths that would preserve social harmony and class distinction. Yet despite his support of severe punishment for social transgressions, Plato followed Socrates in holding, in the *Protagoras, Timaeus,* and *Laws,* that evil is due only to ignorance or madness and that "no man is voluntarily bad," a paradox that Aristotle later tried valiantly to resolve.

ARISTOTLE. One might expect that Aristotle, who studied at Plato's Academy for many years, would take the same view of nature and human conduct as his mentor. But the differences between Plato and Aristotle are more fundamental than the resemblances. Although Aristotle naturally used a similar terminology and shared with Plato certain principles and attitudes expressive of the rationality of Hellenic culture, his method of inquiry and his conception of the role of ethical principles in human affairs were different enough from Plato's to establish a rival philosophical tradition. Plato was the fountainhead of religious and idealistic ethics, while Aristotle engendered the naturalistic tradition. Throughout the subsequent history of Western civilization, ethical views that looked to a supranatural source, such as God or pure reason, for standards of evaluation stemmed from the metaphysics of Plato, while naturalistic philosophers who found standards of value in the basic needs, tendencies, and capacities of man were guided by Aristotle.

Aristotle was born in Stagira, Macedonia, the son of Nicomachus, court physician to Amyntas II. He received early training in biology and physiology and in methods of careful observation and classification, a fact that may account for his later differences with Plato on the role of sense perception in the acquisition of knowledge. While Plato was guided by mathematics as a model of scientific knowledge, Aristotle modeled his system on biology, stressing the importance of observation of recurrent patterns in nature. Thus Plato's goal for philosophical ethics was to make human nature conform to an ideal blueprint, while Aristotle tailored his ethical principles to the demands of human nature.

Aristotle's ethical writings, consisting of the *Eudemian Ethics,* the *Nicomachean Ethics,* and the *Politics,* all edited by his disciples from his lecture notes, constitute the first systematic investigation of the foundations

of ethics. Since the *Eudemian Ethics* is superseded by the *Nicomachean Ethics* and the *Politics* is an extension of his ethical principles to social regulation, this discussion will be confined to the ideas contained in the *Nicomachean Ethics.*

In the latter work, Aristotle's main purpose was to define the subject matter and methodology of philosophical ethics. In doing so, he both drew upon and revised the beliefs and values of the Greek society of his time. Aristotle begins his study by searching for the common feature of all things said to be good and, in contrast with Plato, who held that there is a Form of Good in which all good things "participate," Aristotle concludes that there are many different senses of "good," each of which must be defined separately for the limited area in which it applies. Each such "good" is pursued by a specific practical art or science, such as economics, military strategy, medicine, or shipbuilding. But the ends of these particular disciplines can be arranged in order of importance, so that the supreme good can be identified with the goal of the most general practical science to which the others are subordinate. On an individual level, this all-inclusive science is ethics; on a social level, it is politics. The end of ethics is personal happiness and that of politics is the general welfare, and since the good of the whole ranks above that of the part, personal ethics is subordinate to politics. However, this principle does not entail, for Aristotle, that the individual must sacrifice his interests to those of the community, except under unusual conditions such as war, because he assumed that the needs of both normally coincide.

Aristotle identifies the supreme good with "happiness," which he defines as the exercise of natural human faculties in accordance with virtue. His next task is to define virtue as a skill appropriate to a specific faculty, and he distinguishes two classes of virtues—intellectual and moral. There are five intellectual faculties, from which arise art, science, intuition, reasoning, and practical wisdom. He offers a long list of moral virtues, defining each as the mean between the extremes of either emotion or tendencies to action. For instance, courage is the mean between the excess and the deficiency of the emotion of fear, temperance is the mean between the tendencies to eat and drink too much or too little, justice is the mean with respect to the distribution of goods or of punishments. The bulk of the *Nicomachean Ethics* contains detailed analyses of the criteria of specific moral virtues. The final result of Aristotle's investigations is the definition of happiness or the good life as activity in accordance with virtue, and thus as the harmonious fulfillment of man's natural tendencies.

SUMMARY: SOCRATES, PLATO, AND ARISTOTLE. Returning to the central problems of ethical theory, one may hazard an estimation of the contributions of Socrates, Plato, and Aristotle to their clarification. Socrates was the first to recognize the importance of analyzing the meaning of good, right, just, and virtuous, and of articulating the standards for ascribing these properties. Plato charted a spiritualistic direction for finding the answers in a realm of timeless ideals, while Aristotle located the answers in the scientific study of biology, psychology, and politics. Good, for Plato, means resemblance to the pure Form, or universal model of goodness, which serves as the standard for all value judgments. Actions are right, laws are just, and people are virtuous to the degree to which they conform to the ideal model. For Aristotle, good means the achievement of the goals at which human beings naturally aim, the balanced and rational satisfaction of desires to which he gives the name "happiness." Right action, just laws, and virtuous character are the means of achieving individual and social well-being. All three philosophers agree in identifying individual good with social good and in defining moral concepts such as justice and virtue in terms of the achievement of good.

Moral responsibility. The concept of moral responsibility that acquired crucial importance in later Christian thought was only obliquely considered by Plato and more fully, although inconclusively, dealt with by Aristotle. Plato, who identified virtue with philosophical understanding, concluded that "no one does evil voluntarily," so that wrong action is always due to intellectual error. Aristotle recognized that intellectual error must be distinguished from moral vice, since the former, unlike the latter, is involuntary. In order to distinguish punishable evil from innocent mistakes, he explained vice as due to wrong desire as well as poor judgment. The will, for Aristotle, is rationally guided desire, formed by moral education and training. But since even voluntary action is determined by natural tendencies and early training, Aristotle searched for an additional factor to account for the freedom of choice necessary for moral responsibility. He thought he found that factor in deliberation, the consideration of reasons for and against a course of action. The further question, as to whether, when an agent deliberates, he has any choice of and consequently any responsibility for the outcome of his deliberation, was not considered by Aristotle and remains an unsettled issue between determinists and libertarians. In general, the

concepts of free will and moral responsibility did not become matters of great concern until the rise of Christianity, when people became preoccupied with otherworldly rewards and punishments for moral conduct.

HELLENISTIC AND ROMAN ETHICS

During the two millennia from the death of Aristotle in the fourth century BCE to the rise of modern philosophy in the seventeenth century CE, the interests of ethical thinkers shifted from theoretical to practical ethics, so that little advance was made in the clarification of the meanings of ethical concepts, while, on the other hand, new conceptions of the goals of human life and new codes of conduct were fashioned. The philosophical schools of Skepticism, Stoicism, Epicureanism, and Neoplatonism that set the ethical tone of Hellenistic and Roman thought offered a type of intellectual guidance that was more like religious teaching than like scientific inquiry and paved the way for the conquests of Christianity. The popular conception of philosophy as an attitude of indifference to misfortune applies best to this period, in which philosophy and religion were nearly indistinguishable.

The subtlety of Socrates' thought is attested to by the variety of schools that developed out of his teaching. Plato and, through Plato, Aristotle probably represent the Socratic influence most completely. But the Stoics, Epicureans, and Skeptics also owed their guiding principles to Socrates. Aristippus of Cyrene, at first a disciple of Socrates, founded the school of Cyrenaicism, which followed the simple hedonistic principle that pleasure is the only good. Antisthenes, another Socratic disciple, founded the Cynic school on the apparently opposite principle that the good life is one of indifference to both pleasure and pain. The Cynics, of whom Diogenes was the most renowned, rejected the comforts of civilization and lived alone in the forests, like the dogs after whom they named themselves. Cyrenaicism developed into Epicureanism, and Cynicism into Stoicism. Soon after the death of Aristotle, Pyrrho of Elis initiated the philosophy of Skepticism, influenced by both the Sophist and the Socratic criticisms of conventional beliefs. According to Skepticism, no judgments, either of fact or of value, can be adequately proved, so that the proper philosophical attitude to take toward the actions of others is one of tolerant detachment, and toward one's own actions, extreme caution. In the second century BCE, the leaders of Plato's Academy, Arcesilaus and Carneades, adopted Skepticism, and Carneades developed a theory of probability that he applied to ethical judgments. During this period, the Peripatetic school at Aristotle's Lyceum continued the Aristotelian tradition until it merged finally with Stoicism.

EPICUREANISM. Epicurus (c. 341–270 BCE) founded one of the two dominant philosophical schools of the era between the death of Aristotle and the rise of Christianity. The other dominant school was, of course, Stoicism. These two traditions are often thought of as diametrical opposites, yet it may plausibly be argued that the differences between them were more verbal than substantial. Both views of life were fundamentally pessimistic, directed more toward escape from pain than toward the positive improvement of the human condition. Both encouraged individual withdrawal from the public arena of struggle for economic and political reform, in favor of personal self-mastery and independence of social conditions. The later Roman Stoics modified this extreme individualism and placed more stress on civic duties, but even they preached resignation to the imperfections of social organization rather than efforts at improvement.

Epicurus based his ethics on the atomistic materialism of Democritus, to which he added the important modification of indeterminism by postulating a tendency of the atoms that make up the human body—and particularly its "soul atoms"—to swerve unpredictably from their normal paths, resulting in unpredictable human actions. In this way, Epicurus thought he could account for freedom of the will. He assumed that freedom of choice of action is incompatible with the deterministic principle that all events are necessary results of antecedent causes. But this identification of freedom with pure chance seems to entail that a capricious person is more free than a rational and principled person, and such a conclusion would contradict Epicurus's own vision of moral life. For Epicurus's main difference with his Cyrenaic predecessors lay in his conviction that, by the use of reason, one could plan one's life and sacrifice momentary pleasures for long-run benefit. Like the Cyrenaics, Epicurus held that pleasure is the single standard of good. But he distinguished "natural pleasures," which are moderate and healthful, from "unnatural" satiation of greed and lust. His name for moderate and natural pleasure was *ataraxia,* gentle motions in the body that he regarded as the physiological explanation of pleasure. He proposed, as the ideal way of life, a relaxed, leisurely existence, consisting in moderate indulgence of the appetites, cultivation of the intellect, and conversation with friends, which is how Epicurus himself lived and taught in his famous garden. Two centuries later, Epicureanism was established in Rome by Lucretius (c. 99–55 BCE), whose influential

poem, *On the Nature of Things,* helped to spread Epicureanism among the Roman aristocracy.

STOICISM. Stoicism was by far the most impressive intellectual achievement of Hellenistic and Roman culture prior to Christianity, providing an ethical framework within which metaphysical speculation, natural science, psychology, and social thought could flourish to such a high degree that Stoicism has not unjustly been identified in the public mind with philosophy itself, that is, with the distinctively "philosophical" attitude toward life. Like every great tradition, Stoicism evolved through many stages and thus comprehends a great variety of specific beliefs. Historians generally distinguish three main stages of its development:

(1) The early Stoa—which derived its name from the portico, or porch, on which the early Stoics lectured—whose important figures were Zeno of Cyprus, Cleanthes, and Chrysippus. Chrysippus made the most substantial contributions to Stoic logic and theory of knowledge. The early Stoics remained close to Cynicism in recommending withdrawal from community life so as to render oneself independent of material comforts, social fashions, and the opinions of one's fellow men. Their ethical goal was the achievement of apathy, the state of indifference to pleasure and pain. They considered reason to be the distinctive nature of man and proposed that one should live "according to nature" and thus according to rational principles of conduct.

With the Stoics, the concept of duty acquired a central place in ethics, as conformity to moral rules that they identified with laws of human nature. The later Roman Stoics developed this doctrine into the theory of natural law on which Roman jurisprudence was largely based. Most of the Stoics were materialists, yet imbued with natural piety, and many identified God with the Logos of Heraclitus, as a universal "fire" or energy of nature embodied in its lawlike processes. Many were fatalists, maintaining that man can control his destiny only by resigning himself to it, a principle that contrasted vividly with their emphasis on rationality and self-control. They sought to reconcile this extreme determinism with freedom and moral responsibility by means of the Aristotelian distinction between external and internal causation, thus suggesting that the free man is one who, in understanding the necessity of what befalls him, accepts it and thus freely chooses it, a solution echoed in modern thought by G. W. F. Hegel's definition of freedom as the recognition of necessity.

(2) The middle Stoics, notably Panaetius and Posidonius, brought Stoicism to Rome, shaping the doctrine to the political-mindedness of the Romans by modifying its extreme individualism and stressing the importance of social duties.

(3) The late Stoics, Seneca, Epictetus, Marcus Aurelius, and, to some extent, Cicero—who accepted only certain parts of Stoic doctrine—developed the ideal of a "cosmopolis," or universal brotherhood of man, in which all men would be recognized as having equal rights and responsibilities, an ideal that Christianity absorbed into its conception of the "City of God" and which, in the modern age, Immanuel Kant made the cornerstone of his system of ethics.

NEOPLATONISM. Epicureanism offered a way of life that was open only to the leisure class. Stoicism appealed to highly reflective men of all classes, as evidenced by the fact that the two great figures of late Stoicism were the educated slave Epictetus and the emperor Marcus Aurelius. However, both philosophical views could interest only those of a sufficiently high level of education and thoughtful temperament to place intellectual values above all others. As the Roman Empire declined, and reason seemed powerless to solve the intense economic and social problems of the empire, an atmosphere of pessimism and disaffection with reason began to prevail, a situation that Gilbert Murray described as "a failure of nerve." Interest increased in finding supernatural routes to salvation of the kind offered by various religious cults, and even in the intellectual schools the study of logic and natural science declined in favor of a search for psychological means of escape from suffering. The philosophy of Neoplatonism fashioned by Plotinus (c. 204–270) offered an intellectual road to salvation, while early Christianity paved an emotional and ritualistic highway toward the same destination. Later, these two roads converged.

Plotinus lectured in Rome and, after his death, his notes were edited by his disciple Porphyry, forming the work titled *Enneads*—so called because of its division into chapters of nine sections each. Plotinus developed one strain of Plato's thought, the ascetic mysticism of the passages on the Form of the Good in the *Republic* and the *Symposium* and the pantheistic metaphysics of the *Timaeus.* According to Plotinus, the world is a series of emanations or overflowings of the One, the ineffable and ultimate reality of which every determinate thing is a part. The One is so transcendent as to be indescribable, "the One, transcending intellect, transcends knowing." But if the One cannot be described, it can at least be neg-

atively characterized in terms of what it is not, namely, that it is not limited by any finite properties. This negative characterization of the One was the source of Christian "negative theology," the description of God in terms of the denial of all modes of limitation.

The One emanates intelligible Forms or Platonic Ideas, out of which the World Soul produces individual souls that in turn emanate lower beings in a process that approaches, but does not quite reach, pure matter. Matter, as total formlessness, is so far from true being that it does not exist. Identifying evil with matter or formlessness, Plotinus concluded that evil does not exist in an absolute sense, but only as incompleteness or lack of good. This account of evil as having no positive existence was later adopted by Augustine and most subsequent theologians.

Since Plotinus, following Plato, equated goodness with reality and evil with unreality or distance from the One, it followed that virtue consists in purging the soul of reliance on sensual pleasures and imagery, so that it can ascend the ladder of being and return to its source in the One. The culmination of this process of purification through self-denial is the mystical experience of reunion with the One, which Plotinus describes—having experienced it himself at least four times—as "the flight of the Alone to the Alone." Thus virtue, for Plotinus as later for Augustine, is not its own reward but is a means to a metaphysical state of blessedness. In the words of the historian W. T. Jones, "Like other men of his time, Plotinus found this world a sea of troubles and a vale of tears; like them he sought to leave it; and like them he found perfect peace only in otherworldliness." How much of this view was absorbed into Christian, Islamic, and Judaic theology can hardly be overestimated, although the influence of Platonism on Judaism was mainly through Philo Judaeus (fl. 20 BCE–40 CE), an Alexandrian Jew and contemporary of Jesus, who combined elements of Stoicism with a Platonistic interpretation of Judaic theology and ethics.

(The above section on Hellenistic and Roman ethics was prepared in collaboration with Professor Richard O. Haynes of the University of Hawaii.)

MEDIEVAL ETHICS

The rise of Christian philosophy, out of a fusion of Greco-Roman thought with Judaism and elements of other Middle Eastern religions, produced a new era in the history of ethics, although one that was prepared for by Stoicism and Neoplatonism. The Stoic concern with justice and self-mastery, and the Neoplatonic search for reunion with the source of all being, were combined in early Christian philosophy with the Judaic belief in a personal God, whose commandments are the primal source of moral authority and whose favor is the ultimate goal of human life. Two sources of ethical standards, human reason and divine will, were juxtaposed in one system of ethics, and the tension between them was reflected in conflicting sectarian interpretations of theological principles.

From the second to the fourth century, Christianity spread through the Roman Empire, offering the poor and the oppressed a hope for otherworldly happiness in compensation for their earthly suffering, and thus a way of life with which the more pessimistic and intellectualist schools of philosophy could not compete. By the fourth century, Christianity dominated Western civilization and had absorbed the main ideas and values of the secular schools of thought, as well as rival religions such as Manichaeism, Mithraism, and Judaism. Having converted the masses, it was time to win over the intelligentsia, and doing this required the hammering out of an explicit and plausible system of metaphysical and ethical principles. This task was performed by the Church Fathers, Clement of Alexandria, Origen, Tertullian, Ambrose, and, most completely and authoritatively, by Augustine.

AUGUSTINE. St. Augustine (354–430), born near Carthage, the son of a pagan father and a Christian mother, was first a Manichaean and later became converted to Christianity. He rose in the church to become bishop of Hippo and helped to settle the doctrinal strife among the many Christian sects by constructing a system of theology, ethics, and theory of knowledge that soon became the authoritative framework of Christian thought, modified but not supplanted by subsequent church philosophers. Augustine's major works, *Confessions, The City of God, Enchiridion,* and *On Freedom of the Will,* wove together threads of Stoic ethics, Neoplatonic metaphysics, and the Judeo-Christian doctrine of revelation and redemption into a many-colored fabric of theology. With Augustine, theology became the bridge between philosophy and revealed religion, the one end anchored in reason and the other in faith, and ethics became a blend of the pursuit of earthly well-being with preparation of the soul for eternal salvation.

Like the Neoplatonists, Augustine rejected almost entirely the claims of bodily pleasures and community life, maintaining, as St. Paul had done, that happiness is impossible in this world, which serves only as a testing ground for reward and punishment in the afterlife. Augustine inherited the Neoplatonic conception of virtue

as the purgation of the soul of all dependence on material comforts in preparation for reunion with God. Against the Stoic and Aristotelian reliance on reason as the source of virtue, Augustine maintained that such apparently admirable traits as prudence, justice, wisdom, and fortitude—the four cardinal virtues identified by Plato and stressed by Stoics and Christians—are of no moral worth when not inspired by Christian faith. With the pessimistic view of life characteristic of an era of wars, political collapse, and economic decline—a view already apparent in the Stoic, Epicurean, and Neoplatonic modes of withdrawal from social responsibilities—intensified by his personal sense of guilt and worthlessness, Augustine saw life on Earth as a punishment for Adam's original sin. "For what flood of eloquence can suffice to detail the miseries of this life?" he laments in *The City of God.*

Nature. The tension between natural and supernatural values in Augustine's ethical thought shows itself most clearly in his ambivalent attitude toward nature. Nature, as God's creation, must be unqualifiedly good. Natural evils are only apparently evil, and in the long run they contribute to the fulfillment of divine purpose. Natural evil is simply imperfection that makes variety possible and thus, when viewed on a cosmic scale, does not exist at all. On the other hand, since man must be held morally responsible for his sins, human sin cannot be so easily explained away as incompleteness that promotes the cosmic good. Moreover, it is man's bodily desires that tempt him to sin. Without the aid of divine grace, the promptings of human nature, whether impulsive or rational, lead only to vice and damnation. Augustine resolves this paradoxical view of human nature by holding that man, unlike other natural species, was endowed by his Creator with free will and thus with the capacity to choose between good and evil. Through the original sin of Adam he has chosen evil, and it is for this reason, rather than because of any flaw in his original construction, that he is irresistibly inclined to further sin.

Free will and divine foreknowledge. If Augustine's dual conception of nature is explained by his concept of free will, the latter contains new difficulties. The problem of free will is critical in Christian ethics, which emphasizes responsibility and punishment. The Greek ideal of practical reason ensuring physical and mental well-being was supplanted by the ideal of purification of the soul through suffering, renunciation, and humble obedience to divine will.

Where the practice of virtue produces well-being as its natural consequence, as in the Greek view, virtue carries with it its own reward in accordance with the causal processes of nature, so that causal necessity and moral desert are not merely compatible; they normally coincide. But in the Christian view, causal necessity and moral responsibility seem incompatible, for the choice between good and evil is made by the soul, independently of natural processes, and its reward or punishment is independent of the natural effects of human actions. Man is punished or rewarded to the degree to which he voluntarily obeys or disobeys the commands of God. In the Greek view, man suffers from the natural consequences of his mistakes, but in the Christian view, no matter what the natural consequences of his actions, he is held to account for the state of his soul. It is his motives and not his actions that count in assessment of his moral responsibility, and the primary motive is his desire for, or his turning away from, God.

Responsibility is thus transferred from the consequences of a person's actions to the state of his soul. Yet if the soul is created by God, and not subject to its temporary owner's control, then in what sense can man be said to have freedom of choice between good and evil? Augustine describes the soul that chooses evil as "defective," but if so, is not the Creator of the defective soul responsible for its deficiency? In absolution of God, Augustine argues that a defect is not a positive entity, thus not a created thing and not attributable to a creator—a terminological escape that is vulnerable to the objection that, on such grounds, a man who stabs another produces in his victim a deficiency rather than a positive state and therefore is not responsible for his "nonexistent" product.

Augustine's concept of free will is further complicated by his support of the theological principle of divine omniscience, which entails foreknowledge by God of human decisions. The term *predestination,* used by later theologians and notably by the Protestant reformers, suggests a determinism that Augustine rejects in his criticism of fatalism. For Augustine, God knows what man will choose to do and makes it possible for man to act on his free choices but does not compel him to any course of action. To the obvious question of how God can know in advance what has not been destined or causally necessitated, Augustine replies by means of his subtle analysis of time. God has knowledge, not of what we are compelled to do but of what we freely choose to do, because his knowledge is not the kind of advance knowledge that is based on causal processes but is due to the fact that, in the mind of God, we have already made our decisions. All of past and future time is spread out in the specious present of the divine mind, so that what, from our limited stand-

point, would be prediction of the future is, for God, simply direct awareness of contemporaneous events.

Distinctions among ethical concepts. While Augustine's ethical writings are mainly concerned with the substantive problem of how to achieve redemption, rather than with the clarification of ethical concepts, much of his writing is philosophical in our strict sense, in that it suggests solutions to conceptual or metaethical problems of meaning and method. Augustine opposed the classical tendency to define the moral concepts of rightness and virtue in terms of individual and social well-being and interpreted moral right and virtue as obedience to divine authority. The concept of good is split into a moral and a practical sense. Good as fulfillment of natural tendencies is subordinated to eternal beatitude, the fulfillment of the aspirations of the virtuous soul. Freedom and responsibility are interpreted as internal states of the soul and as excluding, rather than (as for Aristotle) presupposing, causal necessity.

FOURTH TO THIRTEENTH CENTURIES. From Augustine in the fourth century to Peter Abelard (1079–1142) in the eleventh century, Christian, Islamic, and Judaic philosophy was dominated by Neoplatonic mysticism and preoccupied with faith and salvation. The outstanding figure of this period was John Scotus Erigena (c. 810–c. 877), whose conception of good was the Platonic one of approximation to timeless being and whose view of life as issuing from and returning to God bordered on heretical pantheism.

By the eleventh century, interest in rational philosophical speculation had revived, and even those Schoolmen like Bernard of Clairvaux (1090–1153), who continued to defend religious mysticism and denounced reliance upon reason as inimical to faith, nevertheless employed philosophical arguments to refute contrary opinions. Augustine had asserted that one must "believe in order to understand," and St. Anselm (1033–1109) took this to mean that faith is not incompatible with reason but, rather, prepares the soul for rational understanding. The main issues among philosophers of this time were the relation between faith and reason, and the nature of universals.

Abelard, however, an extraordinarily original and independent thinker whose vibrant personality reveals itself in his philosophical writings, rediscovered some of the unsolved problems of ethical philosophy. Abelard brought into clear view the distinctive features of Christian ethics implicit in Augustine's work, in particular, the split between moral and prudential concepts that sharply separates Christian ethics from Greek ethics. Abelard held that morality is an inner quality, a property of motive or intention rather than of the consequences of one's actions, a principle that was later stressed by the Reformation and attained its fullest expression in the ethical system of Kant. A somewhat heretical corollary follows from Abelard's principle, namely that, as Étienne Gilson put it, "Those who do not know the Gospel obviously commit no fault in not believing in Jesus Christ," and it seems clear from all this that Christian faith need not be the foundation for moral rules. Abelard concluded that one can attain to virtue through reason as well as through faith.

THOMAS AQUINAS. The towering figure of medieval philosophy is, of course, Thomas Aquinas (c. 1224–1274), whose philosophical aim was to reconcile Aristotelian science and philosophy with Augustinian theology. The way to this achievement had already been prepared by the revival in western Europe of interest in Aristotle, whose thought had been preserved and elaborated by Muslim and Jewish scholars such as Avicenna, Averroes, and Maimonides and had been brought to the attention of Christendom by the commentaries of Albert the Great. It remained for Aquinas to prove the compatibility of Aristotelian naturalism with Christian dogma and to construct a unified view of nature, man, and God. This he undertook with remarkable success in his *Summa Theologiae* and *Summa Contra Gentiles.*

To a large degree, Aquinas's union of Aristotelianism with Christianity consisted in arguing for the truth of both and in refuting arguments of his predecessors and contemporaries that purported to show their incompatibility. Aristotle's ethics was relativistic, rational, and prudential; Augustinian ethics was absolutist, grounded on faith, and independent of consequences. Now one of these views is totally misguided, or else there must be room for two different systems of ethical concepts and principles. Aquinas adopted the latter alternative and divided the meaning of ethical concepts into two domains, "natural" and "theological." Natural virtues, adequately accounted for by Aristotle, can be attained by proper training and the exercise of practical reason, while theological virtues—faith, hope, and love—require faith and divine grace. Similarly, he distinguished two highest goods, or paramount goals of life, worldly happiness and eternal beatitude (which has precedence); the former is achieved through natural virtue and the latter is achieved through the church and its sacraments. Aquinas thus expressed a considerably more optimistic attitude than did Augustine toward the possibility of improving man's

lot on earth through knowledge of nature and intelligent action. This helped to prepare the climate for the rebirth of natural science, whose first stirrings were felt in the thirteenth century.

Natural law. At the center of Thomistic ethics was the concept of natural law. The medieval doctrine of natural law, stemming from Aristotle's teleological conception of nature and from the Stoic identification of human reason with the Logos, was a fusion of naturalistic Greek ethics with monotheistic theology. On this view, the promptings of informed reason and moral conscience represent an inherent tendency in the nature of man, and conformity to this nature fulfills both the cosmic plan of the Creator and the direct commands of God revealed in the Scriptures. Natural law is the divine law as discovered by reason, and therefore the precepts of the church and the Bible, and scientific knowledge of the universal needs and tendencies of man, provide complementary rather than competing standards of ethical judgment. Where conflicts between science and religious authority arise, they must be due to inadequate understanding of science, since church authority and dogma are infallible.

The Thomistic unification of scientific and religious ethics in the doctrine of natural law—further elaborated in subtle detail by Francisco Suárez and other legalists—was an effective way of making room, within the religious enterprise of achieving salvation, for the practical business of everyday living in pursuit of personal and social well-being. The ideological supremacy of theology was maintained, but the doctrine of natural law purported to guarantee reliable knowledge of nature, psychology, and political economy. The weakness in this system was that it placed religious barriers in the way of scientific advance, tending to sanctify and render immune from revision whichever scientific principles seemed most congenial to theology, such as instinct theory in psychology, vitalistic biology, and geocentric astronomy.

Free will. Aquinas's account of freedom and moral responsibility was, in general form, similar to that of Augustine, maintaining the compatibility of free will with predestination or divine foreknowledge. Aquinas also maintained the compatibility of free will with causal determinism, thus dealing with the problem on the level of prudential ethics as well on as the theological level of grace and salvation. Aquinas's solution makes effective use of Aristotle's analysis of choice and voluntary action in terms of internal causality and deliberation, and it identifies free will with rational self-determination rather than with the absence of causal influences. On the other hand, Aquinas's concept of freedom is, as a result, more relativistic than Augustine's, and, while it explains the conditions under which an agent may be held responsible for his actions—namely, the conditions of desire, knowledge, and deliberation—it does not meet the further issue of whether these faculties that determine action are within the control of the agent, that is, whether a person can freely choose the habits and desires that determine his actions. Later writers, particularly Protestant theologians, tended to interpret Augustine as stressing predestination and Aquinas as stressing free will, but it may be argued to the contrary, that Augustine's conception of free will as an inexplicable and supernatural thrust of the soul allows the agent more independence of his formed character than does Aquinas's, but by that very token, Aquinas's account is more congenial to a scientific view of man.

Subsequent scholastic philosophy, from the fourteenth to the seventeenth centuries, added little to the clarification of metaethical problems, but it probed further into the relation between intellect and will as sources of human and divine action. John Duns Scotus (c. 1266–1308), William of Ockham (c. 1285–1349), and Nicolas of Autrecourt (c. 1300–after 1350) developed the voluntaristic doctrine that the will is free in a more absolute sense than that accounted for by Aquinas, in that it is independent both of external causality and of determination by the intellect—that is, by the agent's knowledge of what is right and good. Their view in one way strengthened the case for religious faith as against scientific reason, at least in matters of ethical judgment, but, in another way, it helped stimulate an attitude of individualism and independence of authority that prepared the ground for the secular and humanistic ethics of the modern age.

EARLY MODERN ETHICS

Philosophy seems to flourish best in periods of rapid social transformation, when the conceptual framework of a culture crumbles, requiring a reexamination of basic concepts, principles, and standards of value. The sixteenth and seventeenth centuries, which saw the demise of medieval feudalism and ushered in the modern age of industrial democracy, were, like the fifth and fourth centuries BCE, a period of intense philosophical ferment. In both cases, the preceding century witnessed the demolition of traditional beliefs, while the succeeding century was one of systematic reconstruction. The development of commerce and industry, the discovery of new regions of the world, the Reformation, the Copernican and Galilean revolutions in science, and the rise of strong sec-

ular governments demanded new principles of individual conduct and of social organization.

In the sixteenth century, Francis Bacon demolished the logic and methodology of medieval Scholasticism. Desiderius Erasmus, Martin Luther, and John Calvin, while attempting to strengthen the bond between religion and ethics, undermined the elaborate structure of canon law based on the moral authority of the medieval church, and Niccolò Machiavelli dynamited the bridge between religious ethics and political science. The task of reconstruction in philosophy was performed in the seventeenth century by René Descartes, Thomas Hobbes, Gottfried Wilhelm Leibniz, Benedict de Spinoza, and John Locke.

HOBBES. Modern ethical theory began with Thomas Hobbes (1588–1679). The advent of Galilean natural science had challenged the traditional notions, supported by authority, of purpose, plan, and value in the physical world; it cast into doubt the doctrine of natural law and nullified the anthropomorphic assumptions of theology. New standards of ethical judgment had to be found, not in the cosmic plan of nature or in scriptural revelations of the divine will but in man himself, either in his biological structure, or in his agreements with his fellow men, or in the social and political institutions that he creates. Thus were born, simultaneously and to the same parent, the ethical philosophies of naturalism, cultural relativism, and subjectivism, respectively.

Born in a time of international and domestic strife, Hobbes regarded the preservation of life as the paramount goal of human action and constructed his system of ethics and political science in his major work, *Leviathan,* with the principle of self-preservation as its cornerstone. His enthusiasm for Galileo Galilei's physics and his conviction that all fields of knowledge could be modeled on this universal science (following the method of Euclid's geometry) may have suggested to him that the drive to self-preservation is the biological analogue of the Galilean principle of inertia. Hobbes conceived of man as a complex system of particles in motion and attempted to deduce ethical laws from the principle of self-preservation. He offers, however, two formulations of this principle, the first of which is his foundation of ethics, while the second is, in effect, the repudiation of ethics.

The tendency to self-preservation, according to Hobbes, expresses itself in the quest for social harmony through peacekeeping institutions and practices or, alternatively, in the aggressive drive toward power over one's fellow men. Thus he formulates his "first and fundamental" principle in two parts, the "law of nature" to the effect that "Every man ought to endeavor to peace as far as he has hope of obtaining it," and the "right of nature," that "when he cannot obtain it, he may seek and use all the helps and advantages of war." Which of these two forms of the principle of self-preservation should be applied depends, for Hobbes, on whether the agent finds, himself in a well-organized society or in a "state of nature" in which he cannot expect cooperative behavior on the part of his fellow men. Thus, the concept of ethical law applies to social agreements and commitments, while that of rights applies to the exercise of natural powers. In the state of nature one has a right to do whatever one has the power to do.

From his fundamental law of nature, Hobbes derives a number of specific rules that prescribe the means of establishing and maintaining a peaceful society, the primary means being the willingness to make or, if already made, to maintain the social contract in which individual rights or powers are surrendered to a sovereign in return for the guarantee of personal security. The state is thus the artificial creation of reasonable men, a "Leviathan" that maintains peace by means of power relinquished to it by its citizens. Once such a commonwealth has been established by contract or conquest, other general rules of conduct follow in accordance with Hobbes's theory of psychology. To restrain the natural human tendencies to envy, mistrust, self-aggrandizement, and aggression, the virtues of accommodation, gratitude, clemency, obedience to authority, and respect for the equal rights of others are recommended by "laws of nature" as effective means of ensuring social harmony.

Reason and ethical laws. Hobbes's use of the term "laws of nature" in referring to ethical principles is to be distinguished sharply from the medieval concept of natural law that he rejected. There is, for Hobbes, no moral order in the cosmos, nor any natural prompting toward justice and sympathy for others in human nature. Man, like the rest of nature, is a system of particles perpetually moving and colliding in accordance with physical laws whereby direction and intensity of motion are determined solely by preponderance of force. Yet reason plays a role in human action that distinguishes man from the rest of the world machine. Ethical rules are "precepts, found out by reason, by which a man is forbidden to do that which is destructive of his life or taketh away the means of preserving the same."

In his mechanistic physiology, Hobbes explained reason as a mechanical process in the brain consisting in the combining and separating particles that serve as representations of objects and qualities; thus, cognitive

processes are a special type of physical process, governed by the same laws. But on this mechanistic view of man, it is difficult for Hobbes to account for the prescriptive character he attributes to ethical laws as distinguished from physical laws. Throughout his discussion, Hobbes vacillates between a conception of ethics as a branch of physical science that describes the behavior of human mechanisms and the quite different conception of ethics as rational advice on how to get along with one's fellow men by consciously restraining one's aggressive impulses. Both sides of the *nomos-phusis* controversy between the Sophists and Plato are represented in Hobbes's thought, and he cites both social authority and prudential reason as sources of ethical obligation. Moral virtue consists in conformity to custom and law, in opposition to the natural aggressiveness that equips a man for survival in the state of nature, yet the "precepts found out by reason" provide a natural basis for the establishment of customs and laws.

Desire and will. Hobbes's account of desire and will is designed to bridge the gap between rational directives and physical laws. He defines "good" as "any object of desire" and desire as the motion toward an object that results from physiological processes ("endeavors") within the body. To act rationally does not entail freedom to act contrary to one's physiological impulses, since rationality or deliberation is simply the mediating processes of the central nervous system. The will is not a supernatural power controlling desires but simply the last stage of deliberation that eventuates in overt action, and thus is itself a neurological process governed by laws of physics. Freedom of the will from causal influences is, for Hobbes, a senseless combination of concepts; freedom is the "absence of external impediments" to the will. It is the person who is free or unfree, and not his will, since his freedom consists in the determination of his overt actions by his will rather than by external forces. Yet this mechanistic account of the will seems in paradoxical contrast with his subjectivist account of civil law as deriving its obligatory force from the arbitrary will of the sovereign, an account that comes dangerously close to the Aristotelian and Augustinian notions of the will as a "first cause."

Naturalism and nonnaturalism. The importance of Hobbes to modern ethical theory is inestimable. In freeing ethics from bondage to revealed theology and its anthropomorphic view of nature, Hobbes brought philosophy back to the problems with which it had begun to wrestle in the time of Socrates and the Sophists, and of which it had lost sight for a millennium. At the same time, he raised the understanding of these problems to a higher level, profiting both from the Christian insight that moral principles have an obligatory force and from the refinements of scientific method introduced by Bacon, Galileo, and Descartes.

If ethics was to become a body of reliable knowledge, it must be grounded on objective laws of psychology and biology, rather than on tradition, sentiment, and church authority. On the other hand, if nature and its scientific description are ethically neutral, then ethics is to be contrasted with science and purged of references to nature, just as natural science must be purged of references to ethical values. In that case, ethical principles must be understood as subjective expressions of emotion and desire, and not as objectively verifiable laws. This dilemma has plagued philosophy ever since, and, if it was not resolved by Hobbes, at least his thought was not completely impaled on either horn but only a bit on both.

EARLY INTUITIONISTS. Reaction to Hobbes's attack on the objectivity of ethical judgment was immediate. The doctrine of natural law and its vision of nature as a moral system were defended in a new form by a group of scholars at Cambridge who became known as the Cambridge Platonists, principally Ralph Cudworth (1617–1688) and Henry More (1614–1687). They maintained that moral principles are self-evident truths, as certain and immutable as the laws of mathematics. Richard Cumberland (1631–1718) attempted to deduce all the principles of ethics from a single "Law of Nature" that later became the cornerstone of utilitarian ethics, namely, the law that all actions should promote the common good. Nicolas Malebranche (1638–1715) developed the Cartesian theory of ethics as a deductive system but gave it an Augustinian slant, attributing to God the sole power to translate knowledge of ethical truth into action. Malebranche realized that the analogy between ethics and mathematics fails to explain the connection between ethics and action, and so he made a virtue of this defect by means of his "Occasionalist" account of causality as divine intervention. Samuel Clarke (1675–1729) developed an intuitionist theory of "natural religion" similar to that of Cudworth and More, holding that the quality of right or "fitness" is an intrinsic property of actions that the mind can perceive as directly as it perceives geometrical relations.

SPINOZA. Born in the Netherlands of Jewish refugees from the Spanish Inquisition, Benedict de Spinoza (1632–1677) combined Descartes's faith in the capacity of reason to govern action with Hobbes's mechanistic the-

ory of psychology to express a scientific vision of nature as a unified system of laws. In his *Ethics Demonstrated in the Geometric Manner* Spinoza, like Hobbes but with more formal precision, derived the principles of physics, psychology, and ethics from metaphysical axioms.

The first principle of psychology for Spinoza, as for Hobbes, is the drive to self-preservation and self-aggrandizement, corresponding to the physical principle of inertia. But Spinoza's unique achievement was to derive, as the logical corollary of this egoistic psychology, a rational, humane, and cultivated way of life. A strict determinist in his metaphysics and a thorough naturalist in his ethics, Spinoza held that every event is deducible from antecedent causes and concluded that ethical right is identical with causal necessity. The rules of conduct are therefore laws of human nature, obeyed by all but obeyed blindly by the selfish person enslaved by his passions while understood and accepted by the free man who, in achieving a vision of the necessary order of all things, experiences the "intellectual love of God" that provides both happiness and moral virtue.

While Spinoza tried more consistently than Hobbes to reduce ethics to psychology and thus to make it a branch of natural science, it has often been contended that his program was self-defeating. For if men cannot help acting in accordance with their desires, and if nothing is objectively good or bad but only appears so to those who do not understand the necessity of all events, then what sense can there be to either prudential or moral rules of conduct? Having banished values from nature, Spinoza, like Hobbes, had to relocate them in human consciousness. But then consciousness must be either a supranatural force that interrupts the causal order of nature—as it was for Descartes—or a part of nature and thus ethically neutral, in which case ethics becomes senseless; or, finally, consciousness is an illusory reflection of physical processes in the body, in which case ethics, too, is illusory. Spinoza and Hobbes vacillated between the last two alternatives although, as we have seen, Hobbes's prescriptivist account of moral right as stemming from the will of an authority may be suspected of having slipped an element of supranatural agency back into the picture.

In their social and political theories, both Spinoza and Hobbes argued for the appraisal of institutions and policies in terms of the satisfaction of human needs rather than of conformity to religious tradition. But Hobbes's conception of force as the basis of law led him to support political authoritarianism, while Spinoza's identification of value and right with rational self-interest enabled him to argue, like Locke, for representative government and maximum civil liberty.

LOCKE. John Locke (1632–1704) is generally regarded as the founder of modern utilitarianism, although his applications of utilitarian ethics to social and political theory were more influential than his analysis of standards of individual conduct. He combined the mathematical model of ethical judgment suggested by Descartes and the Cambridge Platonists with a hedonistic theory of psychology according to which pleasure is the goal of all human action and consequently is the fundamental standard of evaluation. In his *Essay concerning Human Understanding,* Locke criticizes the doctrine of innate ideas of Descartes and Leibniz, in defense of the principle that all knowledge is founded on experience; he then, somewhat paradoxically, offers an account of ethics as a deductive science in which specific rules of conduct are derived "from self-evident propositions, by necessary consequences as incontestable as those in mathematics." The appearance of paradox dissolves, however, on noting that, for Locke, the formation of the ideas of goodness and justice is due to the sensations of pleasure and pain, and thus ethical concepts are derived from experience although their logical relations are then discoverable by reflective analysis.

Locke follows Hobbes in defining good as the object of desire, but then, assuming that the only property of things which provokes desire is their tendency to produce pleasure or reduce pain, he also defines good as "what has an aptness to produce pleasure in us." Again, like Hobbes, Locke defines moral virtue as conformity to custom and law, but he differs from Hobbes in maintaining that custom and law can in turn be evaluated by the more fundamental standards of utility and natural rights. It is in terms of these more basic standards that Locke justifies representative government and civil liberty.

Locke's main contribution to the clarification of the meaning of ethical concepts was in his distinction between "speculative" and "practical" principles. Speculative knowledge is independent of action, while practical principles (including ethical principles) can be said to be believed and known to be true only insofar as they are acted upon. This distinction accounts for the obligatory force of ethical principles and eliminates the need for a supernatural agency, "free will," to translate belief into action, although it makes it difficult to explain why, if practical principles are "self-evident propositions," we do not all behave in a morally impeccable way. Like Hobbes, Locke ridicules the notion of free will as a semantical

absurdity similar to the questions "whether sleep be swift or virtue square." Will is the power of the mind to decide on action, and freedom the power to carry out one's decisions, that is, to get what one wants.

MORAL-SENSE THEORIES. The seventeenth-century philosophers found the connection between self-interest and morality in the threat of punishment—divine, natural, or civil—that coerces the individual to be moral for the sake of self-interest. But it was soon noticed that this connection breaks down wherever the expected benefit to the individual of immoral conduct outweighs the likelihood of punishment and that, if morality is grounded in psychology, then human nature cannot be as aggressively self-centered as the apostles of self-preservation and pursuit of pleasure maintained.

The third earl of Shaftesbury (1671–1713) and Francis Hutcheson (1694–1746) proposed that moral obligation has its source in benevolent affections, such as love and pity, that are as natural and universal as the more aggressive tendencies ("self-affections"), such as envy, greed, and the impulse to self-preservation. Moreover, there is a "moral sense" in man that finds unique satisfaction in actions directed toward the common good. This moral sensibility turns us from the pursuit of pleasure toward the performance of duties toward others and explains our admiration of self-sacrifice independently of external reward or punishment.

Bernard Mandeville (c. 1670–1733), in *The Fable of the Bees,* defended egoistic psychology against this attack and ridiculed the concept of moral conscience as a hypocritical device for maintaining social privileges, a view later echoed by Baron d'Holbach, Karl Marx, and Friedrich Nietzsche. Bishop Joseph Butler (1692–1752), whose sermons in defense of Christian morality against the cynicism of Hobbes and Mandeville reveal extraordinary analytical power, argued that benevolence and conscience are as deeply rooted in human nature as is self-love. In adding conscience or intuition of duty to benevolence as the psychological source of moral obligation, Butler lessened the stress of earlier moral-sense theorists on emotion and gave more recognition to the role of rational judgment.

Moral-sense theory, refined further by David Hartley (1705–1757) and Adam Smith (1723–1790), who applied utilitarian ethics to economic theory, achieved its most persuasive formulation in the writings of David Hume.

HUME. David Hume (1711–1776), like Hartley and Smith, combined an emotional account of morality with a utilitarian theory of good. Hume's discussions of ethics in the third part of his *A Treatise of Human Nature* and, more fully, in his *An Enquiry concerning the Principles of Morals* are attempts to answer the metaethical questions of the meaning of good, right, justice, and virtue; by what standards they are attributed to persons and actions; how it is psychologically possible for men to admire and cultivate morality at the expense of self-interest; and by what rules ethical disputes can be decided in favor of one judgment against another. Despite the clarity and good sense that Hume brings to bear on these topics, his discussion shifts inadvertently from one type of question to another, particularly from questions of meaning to questions of motivation, a shift characteristic of moral-sense theories.

Hume begins his studies of ethical judgment with a search for the meanings of ethical terms. Finding no observable facts or logical relations that answer to our concepts of goodness, justice, and moral virtue, Hume concludes that the function of ethical terms is not to denote qualities or relations but to convey a "sentiment of approbation," so that their meaning is to be found in the feelings of the judge rather than in the object judged. We call things good for the same reason that we call them beautiful: because we find them agreeable. An object is good if it is immediately pleasant, or if it is a useful means for attaining something else that is pleasant. Virtues are qualities that render a person agreeable or useful to himself or to others, whether they are "natural virtues" such as talent, wit, and benevolence or "artificial virtues" like honesty and justice. While judgments as to what is useful in producing pleasure, insofar as they rest on knowledge of causal facts, are within the competence of reason, nevertheless they depend, for their distinctively ethical import, on feeling or taste, since rational knowledge alone is "not sufficient" to produce any moral blame or approbation. "Utility is only a tendency to a certain end; and were the end totally indifferent to us, we should feel the same indifference toward the means. It is requisite a certain *sentiment* should here display itself" (*Enquiry concerning the Principles of Morals,* Appendix I).

Thus, according to Hume, there are two possible grounds or standards of evaluation, utility and feeling, the one objective and subject to rational confirmation, the other subjective and personal. The objective standard, unfortunately, applies only to instrumental values and not to ultimate ends. However, the subjectivity of feelings is not cause for despair about achieving agreement on ethical judgments, since the sentiment that motivates them, the disinterested pleasure and approval that we feel in contemplating actions directed toward the welfare of

others, is, for Hume as for Butler, a universal tendency in human nature.

Moral reasons and psychological motives. In common with Hobbes and Locke, who justified moral conduct by the fear of punishment, and the earlier moral-sense theorists, who explained moral obligation in terms of the benevolent affections, Hume identifies the psychological motives that influence and often prejudice moral judgments with the logical grounds or reasons for moral judgments. From the premise that, were it not for our natural benevolence, we would not care enough about moral issues to make moral judgments, Hume draws the non sequitur that the only evidence which supports such judgments lies in the feeling of approval or disapproval that motivates them.

Hume tends to equate moral virtue with the artificial quality of justice, artificial because it is required only for the protection of property rights in a society in which goods are neither too scarce nor sufficiently abundant. The importance for social harmony of strict conformity to laws renders it dangerous and undesirable to make exceptions in the name of expediency. Consequently, the utility of strict justice outweighs the utility of any possible exceptions. But Hume realized that this rather abstract utilitarian consideration can hardly explain our sense of moral obligation and our admiration for those who demonstrate high moral character. He therefore supplements this account with the notion of "disinterested interest" that resembles the rational moral sense appealed to by Butler, Richard Price, and Thomas Reid (see below).

However, Hume is not positing any occult faculty, for he explains disinterested moral approbation as a combination of the natural quality of sympathy for others (pain at witnessing another's pain) and the habit of following rules. Since natural sympathy alone would lead us into injustices and considerations of utility alone would seem to justify exceptions to general rules, we come to agree on general principles of conduct and transfer to these principles the sentiment of approbation that we originally felt toward the happiness or release from pain usually produced by following such principles. Thus arises the sense of moral duty and the capacity for disinterested approval. Here again, Hume offers a psychological description of the motivating processes that cause us to approve of moral virtue as an answer to the question of what criteria we use to judge persons and actions to be worthy of moral approval. Once this identity of psychological motive and logical ground is presupposed, it becomes impossible to distinguish between correct and incorrect moral judgments. The question as to whether action that meets with general approbation actually merits such approbation cannot even be raised, since merit has already been identified with the mere fact of approbation.

Freedom. On the issue of free will and its relation to moral responsibility, Hume argued persuasively that responsibility presupposes the causal efficacy of threat of punishment. He developed further the arguments of Hobbes and Locke that freedom is not a quality of the will but a relation between desire, action, and environment, such that a man is free when his actions are caused by his own desires and unimpeded by external restraints, a view that William James later baptized "soft determinism."

COMMONSENSE INTUITIONISM. Hume's subjective account of moral judgment was countered by the commonsense intuitionism of Thomas Reid (1710–1796) and Richard Price (1723–1791), who explained the moral sense, or conscience, that enables man to distinguish right from wrong as a combination of benevolent emotion and rational intuition. Both argued, like Butler, that moral principles are not in need of utilitarian justification but are as natural to man as self-love and desire for pleasure. Reid argued that moral qualities are as directly perceived as physical properties are and thus exist in the object judged rather than in the feelings of the subject who judges. Ethics is as much a matter of objective fact as science is, except that its principles are self-evident and can be discovered by "common sense" alone, uncorrupted by bad philosophy. Reid also defended the belief in freedom of the will as the ground of moral responsibility, arguing that we are introspectively aware of our ability to choose between good and evil independently of our desires.

THE FRENCH ENLIGHTENMENT. Ethical thought in eighteenth-century France paralleled developments in Great Britain, although the French philosophers failed to establish as strong traditions as their British contemporaries. French thought subsequent to the eighteenth century added little to moral philosophy as compared with that of Germany and Great Britain. Due to their intense involvement in political issues, the French writers placed rhetorical effectiveness above clarity and consistency as a standard of philosophical value.

Voltaire (François-Marie Arouet, 1694–1778) and Jean-Jacques Rousseau (1712–1778) led the revolt against Cartesian rationalism as well as against political and religious superstition, so transforming philosophy into ideology that *idéologue* became a popular French synonym for *philosophe.* Voltaire employed acid satire in attacking religious and philosophical obscurantism in *Candide,*

Zadig, and his *Philosophical Dictionary,* while Rousseau inaugurated the romantic style of soul-stirring emotional intensity, in place of detached analysis and rigorous argument. Denis Diderot (1713–1784) raised philosophical writing to the highest level of literary grace and subtlety since Plato, criticizing conventional morality and religious beliefs in his remarkable essay-novels *Le neveu de Rameau, Jacques le fataliste,* and *Rêve de d'Alembert.* Yet while appreciating their extraordinary intellectual qualities and the permanence of their place in Western culture, it must be noted that they provided few new concepts and principles on which later ethical philosophers could build.

Rousseau. Rousseau's celebrated exaltation of untutored human nature in his two *Discourses* attributed genial and cooperative tendencies to man's innate disposition and aggressively self-serving tendencies to the harmful influence of civilization. This coincided with the British moral-sense theorists' attacks on Hobbesian egoism. However, unlike Hume (his friend and benefactor prior to their notorious public quarrel), Rousseau considered custom and law to be arbitrary restraints on natural impulses rather than rational methods of channeling self-interest toward the common good. Whatever justification can be given for control of the individual by social institutions lay, for Rousseau, in their claim to represent the "general will," that is, the desires of the majority, independently of whether what is so desired is good. While Rousseau argued forcefully, in *The Social Contract,* for popular sovereignty and the right of revolution, he justified the use by the state of extremely repressive measures, such as the death penalty for atheism. His rather mystical notion of the state as the embodiment of the general will helped to inspire the overthrow in France of absolute monarchy in favor of representative government, yet half a century later it was employed by Johann Gottlieb Fichte, and a century after that by V. I. Lenin, in the justification of authoritarianism.

Although Rousseau's religious mysticism and his preference for feeling over rational prudence were contrary to the general tone of the Enlightenment, his most lasting contribution to ethical philosophy was his insistence that good and evil tendencies are due to social causes, a principle that he shared with baron de Montesquieu, Voltaire, and the Encylopedists. The soundness of this principle is subject to question, but there can be no doubt that it served as a useful guide in the reform of social institutions.

Montesquieu. Charles Louis de Secondat, baron de la Brède et de Montesquieu (1689–1755), in *The Spirit of the Laws* founded the relativistic conception of moral and political principles as grounded in the traditions of particular societies. The "spirit of the laws" is the system of social practices in relation to which new laws are to be evaluated. Western European governments require a division of functions and compensating checks and balances to fulfill the partly republican, partly monarchical values of European society. In treating values as historical and sociological facts, rather than as divine principles or natural laws, Montesquieu developed further the scientific approach to ethics and politics begun by Machiavelli and Hobbes.

The Encyclopedists. Denis Diderot, Claude-Adrien Helvétius (1715–1771), and baron d'Holbach (1723–1789) derived, from a materialistic theory of nature, an ethical view based on the self-centered pursuit of pleasure as the sole rational motive for action. A well-ordered society, on their view, is one in which the pursuit of personal well-being is unhindered by social authority. Insofar as there are conflicts between morality and self-interest, these are due to defects of social organization and perverse education, rather than to the moral defects of individuals. These Encyclopedists, and kindred spirits in other countries, such as the Italian legal philosopher Cesare Bonesana Beccaria, employed utilitarian moral theory in political campaigns for representative government and humane laws and punishments.

KANT AND THE GERMAN ENLIGHTENMENT. The Enlightenment attack on tradition and authority in favor of individual reason took a nonutilitarian form in the philosophy of Immanuel Kant (1724–1804). The utilitarians identified reason with practical intelligence in the pursuit of happiness. Kant, however, inherited the Cartesian and Leibnizian conception of reason as the intellectual recognition of abstract truths. In fashioning an ethical theory that became the main rival of utilitarianism, Kant combined the Augustinian emphasis, revived by Butler, Price, and Reid, on the internal sense of moral obligation with the rationalistic ideal of knowledge as a deductive system. In his *Critique of Pure Reason,* he attempted to show that the laws of science are imposed by the mind on the objects of its perceptions and can thus be known with certainty through reflection on the a priori structure of knowledge. In his *Critique of Practical Reason* he applied the same analysis to ethics, founding morality on the a priori laws with which "practical reason" regulates action. While Kant defended religious faith against the utilitarian freethinkers, he shared their view that ethics is independent of theology, and he followed the deistic tradition of interpreting God as a scientific and

ethical ideal, rather than as a supernatural source of revelation and authority.

In his most influential work on ethics, *The Foundations of the Metaphysics of Morals,* Kant made the most thorough attempt by any philosopher to clarify and explain the difference between ethical principles and laws of nature. The difference lies both in our subjective sense of obligation to obey moral laws, as contrasted with laws of nature, toward which we feel no such obligation, and in the practical—that is, prescriptive—meaning of moral laws, in contrast with the "theoretical"—that is, descriptive—meaning of laws of nature. In virtue of this difference, moral rules are expressed in the imperative mood and laws of nature in the declarative mood. To account for this disparity, Kant distinguished two realms of knowledge dealing with two metaphysically distinct subject matters. Natural science, including scientific psychology, formulates laws of nature that the mind imposes on the objects of perception in accordance with the principle of causal determinism. Ethics articulates the "laws of freedom" that a rational being imposes on his own actions and expects other rational beings to recognize and obey. The justification for these rules lies in the logical fact that to be rational means to act in accordance with general rules and that moral rules are those which can be followed consistently by all rational beings. Thus, insofar as man is moral, he is rational and, in this sense, free; insofar as he is immoral, he is an irrational slave to his natural inclinations. The reward of virtue is not happiness but dignity and freedom.

Moral virtue: The supreme good. Kant's system of ethics is built on three pillars: the examination of the facts of moral experience, the analysis of the logic of ethical judgment, and the formulation of the metaphysical principles presupposed by ethical judgments, as distinct from scientific generalizations. In the first part of the *Foundations* Kant argues, like Reid, that commonsense reflection, uncorrupted by the dialectics of philosophers, informs us with unwavering certainty that duty is distinct from pleasure and utility, that moral virtue or "good will" is the supreme good to which all other values are subordinate, and that moral worth is not measured either by the consequences of a person's actions or by his natural benevolence but by the agent's intention to obey moral laws.

Categorical imperatives. In the second section of the *Foundations,* Kant attempts to explain the distinctive character of moral laws by clarifying the logical differences between three types of rules or imperatives: technical "rules of skill," prudential "counsels" as to how to achieve happiness, and moral duties. The first two, he argues, are "hypothetical imperatives" whose directives are contingent on the desires of the agent. Naturalistic ethics mistakes counsels of prudence for moral laws because the desire for happiness is so universal that directives toward this end have the superficial appearance of unconditional laws. But the generalization that all men seek happiness is a law of nature, not a rule commanding action, and the very possibility of a moral code entails that this psychological generalization is subject to exception. For moral duty requires that the agent sacrifice his personal happiness and even the welfare of his community rather than violate a "categorical imperative."

A moral or genuinely categorical imperative is a rule that commands a type of action independently of any desired end, including happiness. Kant accepts the utilitarian account of hypothetical imperatives but argues that the peculiar obligatoriness of moral principles can be explained only by their unrestricted universality and thus by their independence of any facts of human nature or circumstance. It is not in virtue of what satisfies human needs, but in virtue of the demand of reason that action be in accordance with universal law, that we feel obligated to obey moral principles.

Universalizability criterion. To the question of whether any rule of action can qualify as a moral principle, Kant's answer was in the negative. He maintained that there is one general or "fundamental" categorical imperative from which all specific moral duties can be derived: "Act only on that maxim which you can will to be a universal law." All maxims or specific rules of conduct can be judged morally right or wrong according to this general criterion. If universal obedience to a proposed rule would contradict the very purpose of the rule, as is the case for rules that under certain circumstances permit lying, stealing, or taking life (somewhat inconsistently, Kant approved of capital punishment), then the rule cannot be part of a true moral code. In contrast, a rule such as "Do not make false promises" can in principle be followed without exception and thus qualifies as a moral duty.

This criterion of universalizability, that is, the logical or psychological possibility of requiring universal obedience to a rule of action (logical for "strict" duties and psychological for "meritorious" duties), was undoubtedly Kant's most original and important contribution to ethical theory. It expresses more precisely and unambiguously the "golden rule" to be found in all the great religions, and it has been incorporated, in one form or another, in most modern systems of ethical theory. Countless writers since Kant have attempted to reformulate the criterion of universalizability in a way sufficiently

qualified to avoid reasonable objections, but without complete success.

The obvious objection to Kant's formulation is that no one would want any specific rule of action to be followed without exception. No one would want the truth to be told on occasions when unmitigated harm would result—for example, when a murderer demands to know where his intended victim is hiding. Kant's own reply to this objection is that, while one may not be psychologically inclined to tell the truth on such occasions, there is no logical contradiction in willing—that is, commanding—that it be told, come what may.

A second objection is that Kant assumes, for any rule of action, that either it or its negation must be a moral law, and yet there are few rules, if any, which we would care to have followed universally in either positive or negative form. Kant argues that, since it would be self-defeating to will that every person may make false promises when it suits his purposes, we ought to will that false promises never be made. Yet on the same reasoning one could justify all sorts of absurd laws, such as that everyone at all times wear heavy clothing, since we would not and could not will the universal prohibition of heavy clothing.

A third weakness of Kant's theory is that it provides no grounds for deciding what is right in a situation where apparent moral duties collide and one must be sacrificed in favor of another. With respect to this problem, utilitarianism seems clearly superior to Kantian ethics.

Autonomy of the will. The third part of Kant's ethical theory consists in the metaphysical account of the rational will as a source of action outside the sphere of causal determinism and thus not an object of scientific investigation. The autonomy of the will—that is, the capacity to obey laws of its own conception in defiance of natural causes—is, Kant argues, a necessary presupposition of any moral code. For if all actions were necessary effects of natural causes, then moral evaluation would be pointless. *Ought* implies *can,* that is, the obligation to do what is right entails the ability to do it and the ability not to do it. Since science rests on the regulative principle of universal determinism, there can be no scientific proof of freedom of the will. But this only shows the radical difference between science and ethics and the folly of attempting to derive ethics from psychology. Man as an object of scientific inquiry is an organic phenomenon obeying laws of biology and psychology. But man as an object of ethical evaluation is a noumenal being, free to obey or disobey the dictates of practical reason. From this dual conception of man as both inside and outside nature, Kant derives an ideal way of life impressive in its purity and its faith in human perfectibility. Man as a rational agent is a member of a "kingdom of ends" in which he is both subject and sovereign, legislating for himself and for others. The highest goal of human life is to realize this ideal "kingdom" in individual and social practice.

NINETEENTH-CENTURY ETHICS

Nineteenth-century ethical thought became a battleground for two rival traditions. Utilitarianism, stemming from Locke, Hume, and the French Encyclopedists, dominated British and French philosophy, while idealistic ethics was supreme in Germany and Italy. Both traditions took root in the United States, with idealism appealing to the religious vision of Ralph Waldo Emerson and Josiah Royce, while utilitarianism answered to the developing faith in technology that found philosophical expression toward the end of the century in the pragmatic ethics of James and John Dewey.

UTILITARIANISM. Christian ethics based on divine authority and natural law was given a utilitarian interpretation by William Paley (1743–1805) in his *Principles of Moral and Political Philosophy.* The source of moral obligation, he agreed with Hobbes, lies in the "violent motive resulting from the command of another," while the ground of goodness is pleasure or utility. But moral duty and self-interest coincide because God, as the paramount authority, commands us through the Scriptures and the promptings of conscience to seek the general good as well as our own happiness. Moral obligation is supported both by natural pleasure in the welfare of others and by the fear of divine punishment that provides the selfish but rational person with a good reason to sacrifice his pleasure for the common good. Paley's psychological account of morality, like that of earlier moral-sense theories, failed to explain why anyone who lacks natural benevolence ought to have it. His alternative justification of morality in terms of the fear of divine punishment equally fails to explain why such punishment would be just and why a nonbenevolent nonbeliever in Christian theology can nevertheless be expected to behave morally.

Bentham. The mainstream of utilitarian thought was anticlerical. Jeremy Bentham (1748–1832) and James Mill (1773–1836) formed a political movement that helped bring about legislative reforms by criticizing social institutions in terms of their utility in producing "the greatest happiness for the greatest number." In his influential *Introduction to the Principles of Morals and Legislation,*

Bentham formulated a theory of ethics and jurisprudence remarkable for its clarity and consistency. The great appeal of Bentham's theory lay in its apparent simplicity and ease of application, although these virtues may have been more apparent than real. Bentham attempted to make ethics and politics scientifically verifiable disciplines by formulating quantitative standards of evaluation. He began with the psychological generalization that all actions are motivated by the desire for pleasure and the fear of pain: "Nature hath placed mankind under the governance of two sovereign masters, *pain* and *pleasure.* It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne" (*Principles,* London, 1823, p. 1). From this equation between ethical obligation and psychological necessity, Bentham derived the general principle of utility that "approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question," happiness being understood as the predominance of pleasure over pain.

The most original but also the most dubious part of Bentham's theory is his "hedonic calculus" for measuring pleasures and pains, in computing the overall value of alternative policies. If such a procedure were feasible, ethical judgments would be as scientific as meteorological forecasts, even though both are subject to considerable error, due to the complexity of the factors involved. But Bentham's ideal of a science of ethics runs afoul of two internal difficulties, the resistance of pleasure to measurement and the impossibility of predicting the long-range consequences of actions. Aside from these internal defects, there remains the general objection that pleasure, unlike pain, is not a bodily sensation but a favorable response to an object grounded on the perception of value in the object, as Thomas Reid had argued. To conclude that an object is good from the fact that it pleases us involves the circular reasoning that it is good because it is judged to be good, a principle too vacuous to provide a guide to ethical judgment. If, on the other hand, pleasure is understood in a more narrow, technical sense as desirable bodily sensations, then Bentham's identification of happiness and welfare with pleasure is unacceptable because it reduces human experience to the level of animal existence. The plausibility of Bentham's theory may be due to the ease with which he shifts inadvertently from one of these senses of *pleasure* to the other.

Despite its theoretical defects, Benthamite utilitarianism, which was more socially oriented than that of Locke and Hume, had a salutary effect on social legislation. His analysis of pleasures into factors of intensity, duration, propinquity, certainty, fecundity, and "extent" (number of persons affected) offered reasonable criteria by which alternative social programs and laws can be evaluated and was a marked improvement over the sanctification of existing laws and customs by which Hobbes, Locke, and Hume had made the transition from self-interest to morality. But there is a missing link in Bentham's chain of reasoning that may not be reparable within the confines of his hedonistic psychology, namely, the link that should connect the desire for one's own pleasure with the willingness to consider "extent" or pleasure of others in deciding on a course of action. Is desire for the pleasure of others also a "sovereign master under which nature hath placed us?" If so, then desire for one's own pleasure cannot be sovereign as well. If not, then on what ground are we required to consider the factor of extent?

Mill. John Stuart Mill (1806–1873) recognized the defects in Bentham's formulation of utilitarianism, and in his essay "Utilitarianism" he offered a more sophisticated version that sought to incorporate the moral insights of rival ethical systems. Realizing that Bentham's emphasis on quantitative aspects of pleasure reduces pleasure to bodily sensation and tends to justify an uncultivated mode of life, Mill proposed a new factor by which pleasures could be compared, the factor of quality. Some pleasurable experiences, notably intellectual, aesthetic, and moral achievements, are qualitatively superior to the satisfaction of bodily needs: "Better to be Socrates unsatisfied than to be a fool satisfied." But like Epicurus's preference for "natural" over "unnatural" pleasures, Mill's criterion of quality introduces a standard of value other than pleasure, by which pleasure itself can be evaluated, and thus contradicts the principle of utility, that pleasure is the single standard of good.

Mill also tried to make room in utilitarian theory for the appreciation of the saintly virtues, renunciation and self-sacrifice, by arguing along Humean lines that such virtues are originally valued for their social utility but that we later become attached to them for their own sake, and that this psychological shift from appreciation of virtue as a social instrument to admiration of virtue for itself is a good tendency because it, too, is socially useful. For the appreciation of moral qualities independently of their immediate consequences ensures the social reliability of the agent and, in the long run, produces more good

than harm. This utilitarian defense of moral principles rested on an optimistic belief in the generally beneficial tendencies of man. In applying it to political theory, Mill argued for democratic institutions, minimum state interference in social life, and free economic competition. Assuming a general convergence of individual and social benefit, Mill, like Hume and Bentham, left unanswered the question why, in cases of conflict, one *ought* to place public over private interest and confined himself to explaining why we admire the person who does so. Yet if the social utility of moral self-sacrifice is the only rational ground for favorable judgment of it, then it would seem to follow that each of us has reason to approve of self-sacrifice in others but not in himself. If the step from individual happiness to the greatest good for the greatest number is justified only by the long-range coincidence of the two, then whenever we are assured that they will not coincide, we have no reason to prefer public welfare to our own other than the irrational habit of doing so, a habit that, in such case, it would be wise to break. In Kantian terms, utilitarianism, even in Mill's sophisticated version, fails to provide a logical bridge between inclination and obligation, between *is* and *ought.*

Later intuitionists, beginning with Henry Sidgwick, attempted to supply this bridge by combining the Kantian theory of rational duty with the utilitarian theory of value, maintaining that we are intuitively aware of the duty to obey moral principles at the expense of self-interest but that moral principles, in turn, are justified by their utility in promoting the common good.

IDEALIST ETHICS. Kant's distinction between man as noumenon, legislating and obeying "laws of freedom," and man as phenomenon, governed by laws of nature, was incorporated into new ethical systems by later German idealists, who assimilated the phenomenal side of the distinction to a part of the noumenal side, making natural science subordinate to ethics. Johann Gottlieb Fichte (1762–1814) extended the noumenal will into a universal force that creates the material world out of its own force and expresses itself partially in the free rational will of the individual conscience but more fully in social institutions and laws. The individual thus achieves self-realization in identifying himself with the universal will and voluntarily accepting his *Beruf* (vocation) as part of the social order.

Fichte. In his early work *Wissenschaftslehre* (*Theory of Science,* 1794) Fichte enlarged Kant's ethical concept of man into a metaphysical picture of the universe. Rejecting Kant's notion of things-in-themselves, Fichte reduced reality to the projections of an absolute mind, and he reduced mind itself to will. The criterion of reality became a practical one: That is real which it is right or good to believe and to act upon (the beginning of pragmatism). Fichte went even further than Kant in stressing moral duty as the goal of life. Kant had sharply separated duty from self-interest in criticizing positions of the kind later referred to as utilitarianism, but Fichte moved full circle by reidentifying moral duty with a higher form of self-interest, the self-realization of an absolute will of which each person is a temporary embodiment. The logical problem created by Fichte's voluntaristic idealism is caused by the fact that it begins with Kant's primacy of moral good over prudential good but concludes with a form of supernatural utilitarianism in which prudential good of a higher self reappears as the ground of morality.

Fichte explained the function of the state as the regulation of conflicts among individuals in protection of their natural rights, and on this basis he supported democratic government. But he advanced the view, later elaborated by Hegel, that governmental restraints on individual action are not limitations of personal freedom but expressions of the higher freedom of the absolute will.

In *The Vocation of Man* (1800) Fichte, who had been accused of atheism, developed a less rationalistic and more religious view of human life. He identified the absolute will with the personal God of Christianity and moral duty with the vocation imposed on man by God. In his later *Addresses to the German Nation* (1808) he applied his notion of divinely ordained vocation to the German nation, which he claimed was destined to raise civilization to a higher level. The evolution of Fichte's thought from austere moralism to religious mysticism and then to chauvinistic nationalism provides an instructive example of the lengths to which thought can go in denying the basic distinctions from which it begins, such as that between self-interest and moral duty or between individual rights and social restraints.

Hegel. G. W. F. Hegel (1770–1831) developed Fichte's social basis of ethics further and in more historical terms. For Hegel value, morality, and law are among the highest forms of self-realization of absolute spirit. The Enlightenment doctrine of abstract rights is only the first stage in the development of ethical consciousness. A higher stage is reached in the Kantian sense of moral duty, which recognizes the conflict between individual rights and social responsibilities, subordinating the former to the latter. But the highest stage of self-realization of "objective mind" involves the incorporation of rights and duties in

a rational system of social and political institutions which the individual citizen recognizes as the embodiment of the national will. The perfect freedom that consists in rational self-determination is achieved when individual conscience coincides with custom and law, so that will and reason, subjective motivation and objective necessity, become identical. But this is possible, according to Hegel, only in the modern age of the national state, Christian conscience, and constitutional law. In earlier stages of human history, whatever was necessary for historical progress was, for that age, necessary and therefore right, as, for example, the institution of slavery was necessary and right in ancient Greece. "World history," he declared, "is world justice."

POST-HEGELIAN THEORIES. The impact of Darwin's theory of natural evolution produced naturalistic echoes of Hegelian historical relativism in the utilitarian "survival of the fittest" doctrine of Herbert Spencer (1820–1903), the Marxist philosophy of class conflict, and the cultural elitism of Nietzsche.

Marx. Karl Marx (1818–1883) transformed Hegel's theory of the dialectical self-realization of mind into a doctrine of dialectical development of history through class conflict. In the Marxist theory, moral principles represent the sanctification of the interests of the ruling class at each stage in the development of progressively superior modes of economic organization. Marx criticized both utilitarian and Kantian ethics as variant expressions of bourgeois marketplace procedures. Subordinating rules of individual conduct to the historical imperatives of "revolutionary praxis," the *Communist Manifesto* of Marx and Friedrich Engels called for revolutionary action to achieve a classless society in which "the free development of each is the condition for the free development of all," a society that would require neither the internal repressions of conscience nor the external repressions of laws and punishments. Both morality and the state would "wither away."

Schopenhauer. Arthur Schopenhauer (1788–1860), like Fichte, located the source of both egoistic pursuit of pleasure and moral obligation in the universal will. The morality of equal rights for all represents a higher development of consciousness than that of self-interest, but a still higher stage is reached in the philosophical understanding that the will, in any form, produces illusion and suffering and that the extinction of desire is the only salvation. Schopenhauer gave the Stoic and Buddhist ethic of ascetic renunciation an idealistic metaphysical basis.

Kierkegaard. Søren Kierkegaard (1813–1855) rejected the rationalistic and socially oriented ethic of Hegel in favor of religious individualism. While, like Hegel, he regarded the conflict between self-interest (the "aesthetic attitude") and duty (the "ethical attitude") as reconciled and transcended in a higher stage of consciousness, he denied that this stage could be achieved by reason and described it as a "leap of faith" preceded by tragic anguish. As the contemporary existentialists who have rediscovered Kierkegaard have put it, "The world is absurd" because there are no objective grounds for human decisions. What is right, according to Kierkegaard, is what the individual asserts with the total commitment born of faith, but it is right only for him. Emotional authenticity rather than conformity to rules is the proper guide to action.

Nietzsche. Friedrich Nietzsche (1844–1900) proposed a less mystical but equally individualistic transcendence of moral codes. Like Hobbes and Mandeville, he regarded altruism as contrary to natural impulse and denounced moral restraint as a device created by religion to contravene the natural order of dominance of the strong over the weak. The true source of value lies in the creative self-assertion of the artist and the man of genius who produce new and positive forms of good, while moral prohibitions produce only resentment, envy, and dull conformity.

American developments. In the United States, the transcendentalists, led by Ralph Waldo Emerson (1803–1882) and the pragmatic idealist Josiah Royce (1855–1916), fashioned still other variations on the idealist theme of self-realization as the goal of human life. The transcendentalists identified the self with the creative force of nature, the "oversoul." Royce, following Hegel, defined the fully realized self as a unity of personal and community interests. All of these post-Hegelian philosophies rejected the Kantian morality of strict adherence to general rules of conduct and proposed ways of transcending the conflict between duty and self-interest through a higher mode of consciousness in which the conflict allegedly disappears.

Toward the end of the nineteenth century, William James (1842–1910) and John Dewey (1859–1952) developed the philosophy of pragmatism, in which all of human knowledge is regarded as essentially ethical. They rejected both the Kantian separation of ethics from natural science and the traditional conception of scientific knowledge as disinterested contemplation of value-neutral truths. The split between value and fact was bridged by reinterpreting both so that they became indistinguishable. James combined utilitarianism with a creative

individualism similar to that of Nietzsche and the prescriptivism of Hobbes, by identifying the source of value with the human act of making a claim, thus bestowing value on the object claimed. Ethical judgment is a rational process of determining by empirical investigation which policies are likely to satisfy the maximum number of such claims. James defended the indeterminist concept of free will, criticizing what he called the soft determinism of Hume and Mill as a purely verbal escape from the embarrassing consequences of scientific determinism.

BRITISH IDEALISM AND INTUITIONISM. In the last quarter of the nineteenth century the vitality of idealism began to attract even the sober British intellect, and the ethics of self-realization became a powerful rival to utilitarianism through the influence of Thomas Hill Green, Bernard Bosanquet, and F. H. Bradley.

Green. Thomas Hill Green (1836–1882) introduced Oxford students to the lofty vision of idealist metaphysics. In his *Prolegomena to Ethics* (published posthumously) Green derived liberal ethical and political principles from his conception of the individual self as part of a universal and divine self. He criticized both utilitarianism and moral-sense theories for downgrading the role of reason in moral judgment and for reducing human motives to natural causes. A motive, he argued, is a goal previsioned by a rational consciousness, not an event or process in the body. Value is therefore logically prior to desire rather than a product of desire. One can desire or find pleasure only in what one has judged to be good. The source of evil must therefore be found in defects of the understanding, in the failure of the human mind to realize its identity with the universal mind. The highest good is thus as much an object of self-interest as any other, but it is the kind of self-interest that also constitutes morality.

Green was active in social and political controversies, supporting the North in the American Civil War and supporting liberal legislation in England. Green rejected laissez-faire individualism, insisting on the more positive role of government in promoting social welfare.

Green's ethical theory was sharply criticized by Sidgwick in *The Ethics of Green, Spencer and Martineau* (1902). Sidgwick argued that Green's identification of morality with higher self-interest obliterates the all-important distinction between prudence and duty and thus fails to provide a basis for moral responsibility, a defect that, as we have seen, goes all the way back to Plato.

Bosanquet. Bernard Bosanquet (1848–1923), like Green, grounded ethics and politics on idealist metaphysics. Bosanquet stressed somewhat more than Green the uniqueness of individual values while at the same time taking a Hegelian view of the state as the embodiment of objective mind. Like Green, Bosanquet actively supported liberal political causes.

Bradley. Francis H. Bradley (1846–1924), generally considered the most distinguished ethical theorist among the British idealists, criticized both utilitarianism and Kantian formalism and favored a Hegelian conception of the community as an organic unity whose needs, expressed in social institutions, transcend those of individual citizens, a conception that he applied in the defense of conservative social policies. Bradley was probably more consistent than Green and Bosanquet. If law and custom are the expression of a higher self, then only internal inconsistencies can justify reforms, and individual rights are subordinate to group or national interests. In his *Ethical Studies* (1876) Bradley supported retributive punishment on the ground (which he held to be self-evident to common sense) that punishment is unfair unless it is deserved and that moral desert is independent of social utility. He attempted to reconcile freedom with causal determinism in the notion of an all-encompassing Reality that determines itself in accordance with rational laws. Recognizing that idealism faces the problem of accounting for evil and that its traditional solution—claiming that evil does not exist—is contrary to the judgment of common sense on which Bradley himself always relied, he employed a subtle distinction between existence and reality in holding that evil, though it exists, is unreal. From the standpoint of the totality of knowledge, evil may be seen to contribute to cosmic harmony. This "solution" was later castigated by Bertrand Russell as a morally untenable justification of evil.

Sidgwick. Henry Sidgwick (1838–1900) combined the social utilitarianism of Mill with the intuitionism of Butler and Kant. In *The Methods of Ethics* (1875), a work described by C. D. Broad as "the best treatise on Moral Philosophy that has ever been written," Sidgwick raised ethical analysis to a new level of precision and logical rigor. Setting aside practical moralizing as not the business of objective philosophical analysis, Sidgwick interpreted the task of moral philosophy to be the clarification of the logic of moral judgment, a conception of philosophy that was continued by the contemporary British school of linguistic analysis.

Sidgwick held that there are just three approaches to ethics worth philosophical consideration: egoistic hedo-

nism, utilitarianism, and intuitionism. He pointed out that neither the self-centered ethics of Hobbes and the French Encyclopedists nor the socially oriented ethics of Bentham and Mill can justify the step from psychology to ethics, that is, from the description of human motivation to judgments of moral obligation. Even those who declare that one ought to pursue one's own interests must justify their use of *ought,* and this cannot be done on the grounds of psychological facts alone. Sidgwick therefore insisted on distinguishing psychological hedonism from ethical hedonism and grounding the latter on intuition. His argument is reminiscent of Hume's claim that values cannot be deduced from facts, and it anticipates G. E. Moore's later analysis of the "naturalistic fallacy."

All three "methods of ethics" rest, according to Sidgwick, on principles held to be self-evident, and thus intuitionism is, to some extent, inescapable. The egoist must assume the self-evident rightness of pursuing one's own pleasure, and the social utilitarian must assume the rightness of maximizing the common good. Intuitionists differ from utilitarians and egoists only in holding many principles and duties to be self-evident as well, and thus they expose themselves to inevitable counterinstances. The more numerous and specific the rules claimed to be self-evident, the more subject to exception and vulnerable to disproof. Sidgwick concludes that social utilitarianism offers the correct standard of moral judgment but that this standard is in turn grounded on direct awareness of moral obligation. Thus at least one, and probably at most one, moral intuition is essential for moral judgment.

Sidgwick could not finally decide between the conflicting claims of self-interest and social utility. He leaned toward the latter as definitive of moral duty, but he recognized that one's self-interest rightly carries a special weight, other things being equal. Perhaps he would have been able to reconcile these two "intuitions" more easily had he considered utilitarianism in a somewhat weaker form, as the principle that one ought always to refrain from causing unnecessary suffering, rather than the stronger claim that one ought always to aim at maximizing happiness. For while one's own welfare seems naturally to outweigh that of others, it is very close to being self-evident to any morally sensitive person that he ought not to pursue his interests at the cost of substantial suffering to others.

It would appear from our brief glance over the history of ethics through the nineteenth century that philosophers failed to find any conclusive ethical truths and merely argued, more persuasively and with a more impressive display of learning than most, for whatever way of life and standards of conduct they happened to prefer. In some respects this impression would be justified, and it serves to remind us of the differences between scientific knowledge and ethical wisdom. The perennial character of the problems, the lack of general agreement on proposed solutions, and the return of later doctrines to principles advanced by earlier ones all contrast strikingly with the irreversible progress of scientific discovery. It has been suggested by some contemporary philosophers that the endless disputability of ethical issues is rooted in the very nature of ethical language, so that it is not a defect of philosophy to have failed to achieve general agreement on ethics. As W. B. Gallie put it (*Philosophy and the Historical Understanding,* New York, 1964), ethical concepts are "essentially contestable." It is essential to their meaning that they evoke continual disputes as to the correct standards for their application. But if we cannot find historical progress in the form of final settlement of issues, we can at least discern some degree of gradual, if irregular, advance toward greater clarity in the formulation of the issues.

On the central issue of the logical relation between facts and values, ethical theories have provided increasingly clear and sophisticated statements of two fundamental positions, naturalism and nonnaturalism (sometimes called teleology and deontology). Naturalistic theories relate values to facts by defining "good" and related concepts in terms of observable criteria, such as fulfillment of natural tendencies (Aristotle), satisfaction of desire (Hobbes and Spinoza), production of pleasure for the greatest number (utilitarianism), conduciveness to historical progress (Spencer and Marx), or efficiency of means to ends (Dewey). Nonnaturalistic theories stress the fact that the meaning of ethical terms goes beyond the observable facts on which ethical judgments are grounded, and they locate the additional component of meaning outside nature. Plato located it in a realm of abstract Forms, Christianity in the will of God, the intuitionists in the direct recognition of the quality of rightness, the moral-sense theorists in the feeling of approbation. Each of these accounts of value and moral right has revealed an additional dimension of the complex logic of ethical judgment. Naturalistic theories have brought to light various ways in which ethical judgment is grounded on the fulfillment of biological and social needs, while nonnaturalistic theories have revealed prescriptive aspects of moral concepts that are independent of prudential considerations. The main effort of twentieth-century ethical philosophy was to weave together in a consistent pattern all the threads, both nat-

uralistic and nonnaturalistic, that constitute our philosophical heritage.

CONTEMPORARY NONNATURALISM

In much of the English-speaking world G. E. Moore's *Principia Ethica* (Cambridge, U.K., 1903) is taken to be the starting point of contemporary ethical theory. But it is important to recognize that this primacy is to a considerable degree local and distinctive of the tradition of analytical ethics. On the Continent and in Latin America the work of Max Scheler and Franz Brentano has been a preeminent influence. For much of American thought until about the mid-twentieth century, the work of John Dewey or Ralph Barton Perry provided the starting point. But, for all that, it is reasonable to begin with G. E. Moore.

MOORE. It is the critical side of Moore's work in ethics that has had the most lasting effect. His delineation of the subject matter of ethics and his very careful effort to show that any form of ethical naturalism involves a fundamental conceptual mistake—the work of the first three chapters of *Principia Ethica*—has been the part of Moore's work that has deeply affected contemporary ethical thought. However, Moore's own positive nonnaturalistic cognitivism, with its reliance on nonnatural characteristics, has found few adherents. Most philosophers—C. L. Stevenson and R. M. Hare are typical—who have been convinced that in essence Moore's case against naturalism is sound have not followed Moore's lead but have adopted some form of noncognitivism.

It was Moore's belief that if moral philosophers simply interest themselves in good conduct, they are not really starting at the beginning, for we cannot know what good conduct is until we know what goodness is. Moore's concern was with a "general enquiry into what is good." Our first question must be "What is good and what is bad?" Such knowledge of good and evil, Moore claims, is the "goal of ethical investigation"; but, he stresses, "it cannot be safely attempted at the beginning of our studies, but only at the end." First we must consider how "good" is to be defined.

Moore clearly is not interested in giving a stipulative definition of "good," and from his disclaimers in *Principia Ethica* about being interested in a merely verbal point, it would seem that he is not interested in a lexical definition either. What he is after, in seeking a definition of "good," is just this: what property or set of properties is common to and distinctive of anything that could conceivably be properly called intrinsically good, for instance, "answering to interests." Moore thinks "good" stands for a property, and he seeks to determine what it is. Moore's answer, which he is aware will cause discontent, is that "good" is not definable. All we can finally say correctly is that good is good and not anything else. "Good," like "red," is, in the appropriate sense, indefinable. Good is a simple, unanalyzable, nonnatural characteristic. We are either directly aware of it or we are not, but there is no way of defining it or analyzing it so as to make it intelligible to someone who is not directly aware of it.

Such a radical claim on Moore's part would have little force if he could not thoroughly refute naturalistic and metaphysical theories that do purport to give the kind of characterization of intrinsic goodness that he takes to be impossible.

Moore's case against naturalism. Let us consider Moore's case against ethical naturalism. An ethical naturalist holds that moral judgments are true or false empirical statements ascribing an empirical property or set of properties to an action, object, or person. "Good" is defined in terms of this property or set of properties. But, Moore argues, we will not come to know what good is simply by "discovering what are those other properties belonging to all things which are good." Those who commit what Moore calls the naturalistic fallacy think that when they have "named those other properties they were actually defining good; that these properties, in fact, were simply not 'other,' but absolutely and entirely the same with goodness." But to identify good with any other property is to commit the naturalistic fallacy. The naturalists confuse the question of the meaning of the concept of good with the quite different question of what kinds of things are good.

In a famous argument, which has been dubbed the open-question argument, Moore points out that for whatever naturalistic value we substitute for the variable x in a proposed definition of "good," we can always significantly ask if it is good. If a man says "Happiness is good," or "Self-realization is good," or "The object of any interest is good," we can always significantly ask "Is happiness good?," "Is self-realization good?," "Is the object of any interest good?" Even though we agree, let us say, that happiness is good, it is an evident fact of language that these questions are not without significance. But they would be without significance if "good" did mean "happiness," or "self-realization," or "the object of any interest," just as it is pointless to ask if a father is a male parent or a puppy is a young dog. For whatever naturalistic definitions we offer—whatever naturalistic values replace the variable x—it always makes sense to ask if that thing

is good. Since this is so, these naturalistic definitions can be seen to be inadequate.

This can be seen in another way as well. If a statement like "The satisfaction of desire is good" were a definition of the sort Moore was searching for, it would be analytic and it would be self-contradictory to assert "This satisfies desire but it is not good." For whatever naturalistic definition one proposes, however, one can assert without self-contradiction "This is *x* but it is not good," but if *x* meant the same as "good" this would be impossible, for "*X* is good" would then be analytic. But since this is possible it is clear that the proposed statement is synthetic.

Moore's influence. The above arguments of Moore's, together with his famous argument in Chapter 3 of *Principia Ethica* against Mill's alleged naturalism, have provided the background for much of the controversy in contemporary ethical theory. While few have accepted all the details of Moore's case against ethical naturalism, it has been felt by many that Moore's essential case is well taken. R. M. Hare in his *The Language of Morals* (Oxford, 1952), P. H. Nowell-Smith in his *Ethics* (Harmondsworth, U.K., 1954), and A. C. Ewing in his *Second Thoughts in Moral Philosophy* (London, 1959) try to restate these Moorean insights in such a way as to present a decisive case against ethical naturalism.

It should be noted, however, that the reception of Moore's case against naturalism, even on the part of such eminent nonnaturalists as A. N. Prior and E. W. Hall, has not been that favorable. It is generally thought now that (1) the naturalistic fallacy is not, strictly speaking, a fallacy but is at best a mistake and (2) that it is not really distinctive of naturalism but should be called the definist fallacy, that is, the belief that moral terms are capable of definition in nonmoral terms.

Criticisms of Moore. It is easy to see that someone, though at a certain price, could be a consistent ethical naturalist and that Moore's naturalistic fallacy would not really point to anything necessarily fallacious in such a naturalist's reasoning. An ethical naturalist who is also a hedonist could argue: By "intrinsic good" I am just going to mean "pleasure." This is a stipulative definition on my part and I am making no claim that it squares with ordinary usage, but it will give a clear and consistent definition of "good" that fits well with my preanalytic insight that pleasure and pleasure alone is intrinsically good. It is indeed true that on my theory "Pleasure is good" is a tautology and "Is pleasure intrinsically good?" is a self-answering question. Still, there is a normatively vital question that I can and do ask with perfect conceptual propriety. The vital open question is this: Should an individual seek pleasure and only pleasure as the thing that, morally speaking, he ought always to do? If a man takes this position, Moore's arguments, given above, do not show anything fallacious in his thinking, that is, he has committed no formal or informal fallacy, though it can be shown by some additions to Moore's arguments that he has said something that is mistaken.

There is a further criticism of Moore that can be made with considerable plausibility. Though it is indeed true that *good* taken in isolation cannot be defined, the term *good* is in reality always used in specific contexts, with context-dependent meanings and with such riders as "good at" and "good for." But in such a context *good* can be defined. "A good car," "good teacher," "good at ballet," or even "good man" can be naturalistically defined, even though *good* sans phrase cannot. Finally, and perhaps most importantly, it has been pointed out that the open-question and noncontradiction arguments are not conclusive. At best they show why all the naturalistic definitions hitherto proposed do not work. They do not show that naturalistic definitions are impossible.

DEONTOLOGICAL NONNATURALISTS. There are other nonnaturalists who, while holding cognitive metaethical theories, reject Moore's ideal utilitarianism. Moore thought that Bentham and Mill were mistaken in trying to define *good* naturalistically, but that they were not mistaken in regarding good as the fundamental moral concept and were not mistaken in arguing that it is always our duty to seek to bring the greatest total good possible into being. H. A. Prichard, W. D. Ross, E. F. Carritt, and C. D. Broad all agree with Moore that intrinsic good is a unique, nonnatural quality that is indefinable and can only be known directly. But they reject Moore's claim that *right* means "productive of the greatest possible good." *Right,* they argue, is also sui generis; it is not reducible to *good* or to any teleological concept. To say "This is a right act" means, according to Ross, "This act ought to be done." Furthermore, even what makes an act right is not to be completely determined by teleological concepts. An act, even though it may be productive, everything considered, of the best consequences, may still not be the right thing to do. Even Broad, who makes the most concessions to the utilitarians of any of the deontologists (as they are called), argues that in determining what is suitable to the actual situation, we must consider both the total fittingness of the events that are relevant to the act in question and the utilities in question, and then without any precise measure of what is suitable to the situation, we must decide what we are to do. The utilitarians, including

Moore, the deontologists agree, oversimplify the situation here.

In 1909 H. A. Prichard, in his celebrated article "Does Moral Philosophy Rest on a Mistake?," set forth in perceptive but uncompromising form the deontological position. But it is W. D. Ross, taking Prichard's position as a starting point, who has been the most influential of these deontological nonnaturalists. Ross's *The Right and the Good* (Oxford, 1930) and his *Foundations of Ethics* (Oxford, 1939) present the classical statement of these views.

Prichard. In "Does Moral Philosophy Rest on a Mistake?" Prichard argued that it was an endemic mistake of moral philosophy to try to give reasons for our obligations. Moral obligation cannot be reduced to acts that ought to be done because by doing them, more good is likely to result than by doing any alternative act. We do not, Prichard contended, come to appreciate an obligation by argument, but in a particular situation we are either directly aware of what it is we ought to do or we are not. Moral philosophy cannot justify these obligations; it can only (1) help us to come to understand the nature of this immediate type of awareness and (2) help us to see through the confused attempts to exhibit the "truly rational foundations" of these obligations by showing how they are grounded in human interests.

Ross. Ross accepted the Prichardian belief that we have an intuitive insight into our obligations, but he went on from certain hints in Prichard to develop a concept of prima-facie duty. A prima-facie duty is a conditional duty of a very distinctive kind. What is meant by saying that it is "conditional" is that it is something that always would be an actual duty were it not for the fact that in certain circumstances there are more stringent moral considerations that outweigh it. But prima-facie duties are always actual duties unless such conditions obtain. Ross takes it as "self-evident that a promise, simply as such, is something that prima facie ought to be kept, and it does not, on reflection, seem self-evident that production of the maximum good is the only thing that makes an act obligatory." Like John Cook Wilson and Prichard before him, Ross takes as his data "the moral convictions of thoughtful and well-educated people." They serve as his point of departure and his check on all theorizing concerning morals.

Reasoning from this base, Ross can show that we do not always reason as utilitarian moralists would have us reason. We often have duties of special obligation that conflict with the utilitarian principle that we should always maximize good. If we carefully attend to the data of ethics—our actual moral experiences—we will note that we have prima-facie duties to fidelity, reparation, gratitude, justice, beneficence, nonmaleficence, and self-improvement. Some of these prima-facie duties are more binding than others. *Ceteris paribus,* the duty of nonmaleficence outweighs our obligation to keep a promise. But Ross stresses—as does Broad—that it is not always the case that we have a rule, a general principle, for deciding what to do when there is a conflict in prima-facie duties. Sometimes we simply have to appreciate or come to "see" what is suitable to the situation.

Criticisms of deontology. Many, though by no means all, philosophers would agree that the deontologists have shown that moral reasoning is not as simple as the classical utilitarians took it to be. But it has been thought by many that consequences play a far larger role in determining what makes an act right than the deontologists have been willing to admit. Their rather antiquated epistemology of intuitions, synthetic a priori judgments, and so forth, and their misleading use of mathematical analogies have stood in the way of an acceptance of deontology. It is, however, quite feasible to argue that such appeals are not essential to a deontological view.

It has also been repeatedly argued that a deontological position, with its list of prima-facie duties and its appeal to the convictions of the thoughtful and the well-educated, is thoroughly ethnocentric. To these objections it is reasonable to reply that most of Ross's prima-facie duties are very similar to the kind of generalities that the anthropologists Ralph Linton and Robert Redfield (among others) have claimed to be cross-culturally sanctioned "universal values." Moreover, the appeal to thoughtful and well-educated people surely need not and should not limit itself to people in one cultural circle.

Rather more important criticisms of deontology have been that it gives us no criteria for deciding what laws, practices, rules, or institutions are worthy of our acceptance. Here the kind of quasi-utilitarian reasoning concerning practices characteristic of the good-reasons approach seems to have decided advantage.

Ewing. It should be mentioned that A. C. Ewing in two closely reasoned books, *The Definition of Good* (New York, 1947) and *Second Thoughts in Moral Philosophy* (London, 1959), works out a theory that in many respects tries to find a middle ground between Moore and Ross. Ewing takes *ought* as his fundamental term, and in the second work he makes far more concessions to the naturalists and noncognitivists than in the first, without abandoning what he takes to be the core of his nonnaturalism.

PHENOMENOLOGICAL VIEWS. Moore, Ross, Broad, and Ewing are not the only nonnaturalists and intuitionists who have exerted a considerable influence on contemporary ethical thought. During a roughly comparable period, Franz Brentano, Nicolai Hartmann, and Max Scheler had a comparable influence on the Continent.

It is necessary to mention that in contemporary philosophical thought there is a fundamental cleavage that divides the English-speaking and Scandinavian countries, on the one hand, from the Continent, Latin America, and the Near East and Far East, on the other. In these latter areas of the world the influence, either direct or indirect, of the philosophers so far discussed has been slight, while the influence in intellectual circles of the philosophers to be discussed in this section and in the section on existentialism has been considerable. Even though Moore, Ross, and Ewing opposed empiricism, their techniques remained analytical, while the work of the philosophers about to be discussed is philosophy in the grand manner; that is, it is comparatively speculative and metaphysical.

Brentano. Franz Brentano's *The Origin of Our Knowledge of Right and Wrong* (Leipzig, 1889) and his later *Grundlegung und Aufbau der Ethik* (F. Mayer-Hillebrand, ed., Bern, 1952) mark the beginning of contemporary Continental ethical theory. In 1903 G. E. Moore remarked that Brentano's work more closely resembled his own than that of any writer with whom he was acquainted. Like Moore, Brentano rejected naturalistic definitions of ethical terms, regarded fundamental moral concepts as sui generis, and thought judgments of intrinsic value incapable of being proved.

To gain an adequate understanding of Brentano's ethical theory, it is essential to understand the rudiments of what he called descriptive psychology (the latter, in Edmund Husserl's hands, was to become phenomenology). Brentano classified mental phenomena into three fundamental classes: ideas and sensory presentations (images and the like), judgments, and emotions. That is to say, there are three fundamental ways in which one may be intentionally related to something. One may simply think of it, one may take an intellectual stance toward it by either accepting it or rejecting it, or one may take an emotional or attitudinal posture toward it. To do the last is a matter of loving or hating it. (Brentano, of course, uses *love* and *hate* here in a very stretched manner.) Brentano regarded emotions as intentional; he maintained that "certain feelings refer unmistakably to objects and language itself signifies this through expressions that make use of it." Moreover, emotions, like judgments but unlike ideas, can properly be called either correct or incorrect. In this way Brentano differed radically from the emotivists.

How do we decide whether a given emotion is correct or incorrect? Here Brentano, who like Ross was a careful student of Aristotle, was very Aristotelian. We can come to understand what a correct emotion or, for that matter, a correct judgment is only by contrasting actual cases of emotions and judgments taken to be correct by experienced and thoughtful people with cases that are not so regarded.

To say that something is good—where we are talking about "intrinsic good"—is to say that it is impossible to love it incorrectly. To say that something is intrinsically evil is to say that it is impossible correctly to love whatever is in question. "Good" and "evil" are what Brentano called synsemantic terms: They do not refer to concrete particular things, either physical or mental. But such ethical concepts were, on Brentano's view, objective because of the impossibility of loving correctly whatever is hated correctly and of hating correctly whatever is loved correctly. The truth of these fundamental moral judgments is directly evident to the mature moral agent. Any question about the empirical evidence for them is as impossible as it is unnecessary.

Scheler. Max Scheler attempted to apply Husserl's phenomenological method to moral concepts. His major works in ethics, *Formalism in Ethics and the Ethics of Intrinsic Value* (Halle, 1916) and *The Nature of Sympathy* (Bonn, 1923), are among his earlier writings (*The Nature of Sympathy* is simply a second and enlarged edition of the early *Zur Phänomenologie und Theorie der Sympathiegefühle,* Halle, 1913); but his later work in philosophical anthropology, *The Forms of Knowledge and Society* (Leipzig, 1926), also has important implications for his ethical theory.

Scheler's ethics is best understood by setting it in relation to that of Kant. Scheler accepted Kant's critique of naturalistic and utilitarian ethical theories. But while he took the categorical imperative as pointing to an essential feature of morality, he thought that such Kantian formalism was incomplete. Like Husserl, Scheler believed that Kant was mistaken in limiting the a priori to the purely formal. The phenomenological method shows that we have a *Wesensschau* (an intuition of essences) in virtue of which we know certain fundamental a priori but nevertheless nonformal moral truths, such as "Spiritual values have a higher place in the scale of values than vital values, and the Holy a higher place than the spiritual."

Given this very extended sense of "a priori," it is correct to say, according to Scheler, that there are objective nonformal moral judgments which are universal, necessary, and synthetic. These moral judgments are said to have an intrinsic content that is given in our intuition of essences.

Scheler argued that there is a hierarchy of objective values, all open to our intuitive inspection. There is, he would argue, nothing subjective about this ordering. In the hierarchy of values phenomenologically given to man, we have at the top religious values, then cultural values (aesthetic, speculative, scientific, and political), and finally, at the bottom, material values (useful things, things that satisfy needs, desires, etc.). All of these values are thought to have an ethical dimension. Questions concerning moral obligation arise when there is a conflict of values. Moral obligation is that which binds us, in such a situation, to take as the order of our incentives the values as they are ordered in the value hierarchy. Scheler was, however, sufficiently Kantian to believe that the ultimate ground of moral obligation lay not in the consequences of moral acts but in the intentions of moral agents. To someone who has studied Mill, Sidgwick, or Ross, this seems like a plain confusion between the moral "grades" we would give a person and an objective consideration of what acts are morally right.

There is another aspect of Scheler's moral theory that should be mentioned, namely, his claim that love and sympathy are the sole means by which we gain an intuitive insight into moral reality. Like Brentano, he thought that these feelings had intentional objects, and like Blaise Pascal, he thought that there was a "logic of the heart"—that through the feelings we gain a type of cognition into essential value structures that can be had in no other way.

Hartmann. Nicolai Hartmann's massive work *Ethics* was published in Berlin in 1926. It shows the influence of Scheler and Husserl and is without doubt the most extensive phenomenological discussion of value in the literature. Ethics, for Hartmann, is part of a general theory of value, though, as might be expected, ethical values are the highest values. "Value" for Hartmann, as for Scheler, is a general predicate, and under it there are more specific predicates for determinate values, for instance, "beauty" is to "value" as "red" is to "colored." Values are said to be essences, and we have a direct though emotionally tinged intuition of essences. Being essences, values, like numbers, are thought by Hartmann to have an ideal self-existence (*Ansichsein*). But unlike numbers, values have a "material essence."

Like Scheler, Hartmann believes that if we will but attend patiently to our feelings, we will be able to discern, though vaguely, some hierarchical ordering of those things that are valuable. Putting aside as far as possible our theoretical preconceptions concerning values, we should reflect carefully on our actual experience until we achieve a clear and evident insight into value phenomena. This, of course, is a desideratum that will never be completely achieved, for "morally no age entirely comprehends itself." The real ethical life is "a life deeper than consciousness." But there is a capacity on the part of the human animal to appreciate the valuable, and by ever more carefully attending to this, we can attain both a clearer view and a more purified form of the moral life.

Though values are material essences, they are not, as in Plato, identical with being. Hartmann, no more than Moore or Jean-Paul Sartre, will identify what is good or what has worth with what exists. That would destroy the autonomy of ethics and obscure the nature of value. But although values are independent of existence, they are related to existence by a "tendency to reality" that Hartmann calls the ideal Ought-to-Be. We have many different values, but it always remains the case that values ought to be. The criteria for what is good or for what is valuable vary from context to context, but the ought-to-be remains the same: "The ideal Ought-to-Be is the formal condition of value, the value is the material condition of the Ought-to-Be." In contrast with the ideal Ought-to-Be there is the more practical, more directly morally relevant "Ought-to-Do." Here "ought" implies "can," and here practical moral questions arise about making something the case that is not the case.

More recent developments in Germany. Finally, a brief note is in order about more recent developments in ethics among German philosophers. Martin Heidegger, whose influence is completely overshadowing in Germany, took a dim view not only of the relevance of logic to philosophy but also of philosophical ethics. This has impeded systematic work in ethics in Germany, but nonetheless it is going on. There has been a reaction against the work of Scheler and Hartmann. O. F. Bollnow has argued for a *Situationsethik* and Richard Schattländer has contended that the Scheler-Hartmann approach is too speculative and theoretical and does not adequately handle the moral agent's question: What ought I to do? But the Scheler-Hartmann school is hardly dead, for Hans Reiner, in his *Das Prinzip von Gut und Böse* (Freiburg, 1949), gives us a detailed and vigorous restatement of such a position. Against Heidegger, he defends the philosophical importance of a general theory of

value. But in an effort to blunt Heidegger's criticism that such investigations are morally and humanly irrelevant, Reiner concerns himself primarily with moral values. In his concern with moral value, he examines in some detail the problem of ethical relativism, and in this examination he stresses the importance of anthropological investigations to our understanding of morality.

NATURALISM IN AMERICA

While ethical naturalism seemed to have received its quietus in England from Moore and Ross and certainly could not be considered a major force on the Continent, in America in various forms it was, until shortly after World War II, the dominant form of ethical theory.

PERRY. R. B. Perry developed a general theory of value with specific applications to questions of normative ethics, law, politics, economics, and education in his *General Theory of Value* (Cambridge, MA, 1926) and *Realms of Value* (Cambridge, MA, 1954). "Value" is used by Perry in a very broad sense as a generic term to group together such terms as *desirable, good, worthwhile, right, beautiful, holy, obligatory,* and the like. Perry defines *value* as follows: "a thing—anything—has value, or is valuable, in the original and generic sense when it is the object of an interest—any interest." In an attempt to make his contention overtly verifiable, Perry in turn defined *interest* quasi behavioristically as "a train of events determined by expectation of its outcome." *Interest* for Perry was an umbrella term for such terms as *like, desire, preference,* and *need* and their opposites. For something to have positive value, it must be an object of a favorable interest; for something to have negative value, it must be an object of aversion, disapproval, or dislike: In short, it must be an object of negative interest.

It should be understood that this definition of *value* is not taken by Perry to be either a lexical or a purely stipulative definition. It is, rather, a reforming definition. That is to say, it is a deliberate proposal concerning the use of a term in the language, but the proposal is not simply a stipulation, for it has some antecedent basis in the usage in question. It is proposed that this use be adopted as the standard use in order to clear up what are taken to be confusions allegedly resulting from unclear and vacillating usage. By such maneuvers Perry hoped to escape from Moore's arguments concerning the naturalistic fallacy.

Such a theory, initially at least, is extremely attractive, for it holds out a promise for a genuine "normative science" and thus for some objective, if not absolute, knowledge of good and evil. It holds out the promise that we will eventually use the emerging sciences of man to gain some cross-cultural and interpersonally confirmed, and thus objective, knowledge of right and wrong.

The crucial problem for the naturalist is to show how all statements containing ethical terms can be translated into statements that do not contain such terms and are directly or indirectly confirmable or disconfirmable by empirical observation. What must be achieved to develop such a naturalism is to show the tenability of some set of naturalistic definitions of key moral terms.

Working from his initial definition of "value," Perry developed his system from the following definitions:

(1) "*X* has value" equals "*X* is the object of any interest."

(2) "*X* is bad" equals "*X* has negative value."

(3) "*X* is good" equals "*X* has positive value."

(4) "*X* is intrinsically good" equals "*X* is the object of a favorable interest for its own sake."

(5) "*X* is extrinsically good" equals "*X* is the object of a favorable interest because *X,* directly or indirectly, is the most efficient means to something which is intrinsically good."

(6) "*X* is morally good" equals "*X* is the object of interests harmoniously organized by reflective agreement."

(7) "*X* is the highest good" equals "*X* is the object of an all-inclusive and harmonious system of interests."

(8) "*X* is morally right" equals "*X* is conducive to the moral good."

(9) "*X* is morally obligatory" equals "*X* is a social demand that, of any alternative demand, is most clearly called for by the ideal of harmonious happiness."

A theory based on these definitions should, Perry would argue, provide us with a systematic account of our normative concepts and exhibit the rationale of our moral judgments. However, it would be queried by many, including many who are not intuitionists, just how it can be that all moral statements are really a subspecies of empirical statement and how they all could, even in principle, be empirically confirmed or disconfirmed. To take moral statements as empirical statements asserting that so-and-so is the case seems to miss their distinctive, dynamic, and guiding function in the stream of life.

DEWEY. For John Dewey, moral philosophy had a definite normative ethical function. Dewey wanted to criticize normative standards and hoped to indicate more reasonable moral goals. "Philosophy's central problem," he wrote, "is the relation that exists between the beliefs about the nature of things due to natural science and beliefs about values—using that word to designate whatever is taken to have rightful authority in the direction of conduct."

His basic proposal was that we should use what he called experimental intelligence in morals. This means that in moral inquiry we should use the same methodological principles we use in scientific inquiry. We should develop a scientific critique of our institutions and of the patterns of conduct designated "moral." In order to do this we must show the untenability of what Dewey took to be an unjustified but ancient philosophical preconception that injects a divorce or dichotomy between scientific knowledge, on the one hand, and moral, philosophical, or religious knowledge, on the other. There is but one kind of knowledge, with one reliable method of fixing belief, the experimental method, though this knowledge and method of fixing belief must be applied to different subject matters.

To most people, the use of the experimental method in ethics heralds a drop of any normative ethical standards. In trying to establish that this is a misconception, Dewey tried to establish a severe contextualism. A central mistake of traditional moral philosophies, both naturalist and nonnaturalist, was that of looking for one bedrock *summum bonum* or one ultimate moral criterion rather than realizing that there is an irreducible plurality of moral standards and that moral problems are fully intelligible and rationally resolvable only in a definite context. Moral standards are a part of a cultural context in which means and ends are qualitatively continuous and functionally interactive.

This reference to a continuum of means and ends leads to another main element in Dewey's moral philosophy. He argues against the specialist's conception of ethics. To hold this conception, which is traditional with philosophers as different as Plato and Russell, is to stress the distinction between intrinsic good and instrumental good and to contend that intrinsic good is the sole object of philosophical interest. This, according to Dewey, is a mistaken dichotomy rooted in the ancient Greek dichotomy between theory and practice. It is not only intellectually bankrupt but it can, Dewey argues, have vicious social consequences. It even makes for irrationalism in ethics, for given this conception, we are easily led to the assumption that while science can deal with mundane instrumental goods, the highest goods—the basic ends, namely, intrinsic goods—must be grasped by intuition, be vouchsafed by revelation, or be merely a matter of the whims of mortal will. Dewey argued that in concrete moral contexts, answers concerning means actually transform ends. In reasoning morally it is not a matter of discovering the most efficient means to attain a fixed end. If in considering the means it becomes apparent that our ends are utopian, we will, if we are behaving rationally, often give them up or modify them in view of this discovery. Here intelligence has a major role to play in morality. Ends cannot rationally be divorced from means. In fact, they are always functionally interactive. Furthermore, what is an end in one problematic situation is a means in another, and so on. There are never any actual normative goals or ends that are simply intrinsic goods. Ideals are always transformable in the light of what we discover about our world, and they are always imbedded in a network of other ideals.

Such considerations, it will surely be objected, hardly show that there are no intrinsic goods—but it could be contended that they effectively argue against Aristotelian final ends, or against the belief that in moral appraisal we can justifiably consider intrinsic goods independently of their consequences—and this, after all, is the major point Dewey wanted to establish.

Here we hardly have the metaethical concerns that are so distinctive of the work of Moore and Perry. But Dewey—though he did not call it that—also had a metaethical theory.

Dewey argued that moral judgments are judgments of practice. That is to say, they are made in problematic situations of choice in which a moral agent is trying to decide what to do. This gives them their distinctive normative or de jure force. But at the same time they remain de facto empirical statements. It is this puzzling amalgam that we must understand if we are to get clear what Dewey was claiming.

Dewey asserted that value judgments are not mere prizings and disprizings. They are predictions about the capacity or incapacity of actions, objects, or events to satisfy desires, needs, and interests. As such they are confirmable and disconfirmable. They predict that certain ends in view will satisfy certain vital impulses under certain conditions. Not everything that is desired is desirable, but those things which are desired "after examination of the relations upon which the object depends" are desirable. In short, to say of something that it is valuable, desirable, or good is to say that it is some-

thing which would be desired or approved after reflection upon its relevant causes and consequences.

Criticism of Dewey. Dewey's theory has been subject to some trenchant criticisms by Morton White and Charles Stevenson and has been staunchly defended by Sidney Hook, Gail Kennedy, and Gertrude Ezorsky. The basic considerations here are as follows: Even if *X* is desired after an examination of the causes and consequences of desiring *X,* it still does not follow that *X* is desirable or that *X* ought to be desired. However, to carry out Dewey's program of identifying moral statements as a subspecies of empirical statement, some such identity of meaning must be established.

But the admission that Dewey is wrong in claiming that moral statements are empirical statements or hypotheses is not destructive to his overall program about the place of reason in ethics. If we ask how we justify our ethical evaluations, it seems that much of Dewey's method of criticism, including much of his use of science, could still be reasonably instituted. Dewey's great failure in talking about morality was in not realizing how very different "values" and "facts" are; his great success was in seeing the extensive relevance of scientific knowledge and scientific method to the making of intelligent moral appraisals.

CONTEMPORARY NONCOGNITIVISM

Both naturalism and nonnaturalism are cognitive theories. That is to say, they regard moral utterances in the declarative form as statement-making utterances that assert the existence of certain moral facts and are thus either true or false. But first in Sweden, and later in England and America, a quite different kind of metaethical theory developed that has been called a noncognitive theory. According to this theory, moral statements do not assert moral facts; they are neither confirmable nor disconfirmable, and there is nothing to be known by "moral intuition." It is even characteristic of this view to argue that it is either mistaken or at least misleading to characterize moral utterances as true or false.

EMOTIVE THEORY. The noncognitive view, which has subsequently been called the emotive theory, received its first formulation in 1911, when the Swedish philosopher Axel Hägerström drew the outlines of such a theory in his inaugural lecture, "On the Truth of Moral Propositions." In 1917 Hägerström developed his ideas with particular attention to the concept of duty in his *Till Frågan om den Gällande Rättens Begrepp* (Uppsala, 1917). Similar statements of the emotive theory have been developed in Scandinavia by Ingmar Hedenius and Alf Ross. Independently of its Scandinavian formulation, the emotive theory was first stated in the English-speaking world by I. A. Richards and by Bertrand Russell, but it was developed in the Anglo-Saxon world by A. J. Ayer and by Charles Stevenson. There have also been interesting if somewhat atypical statements of it by Richard Robinson, Rudolf Carnap, and Hans Reichenbach.

The emotivists were convinced that moral statements are not a subspecies of factual statement, and they were further convinced that it was impossible to derive a moral statement from a set of purely factual statements. As Hägerström put it, "There is no common genus for the purely factual and the 'ought.' By using the predicate 'ought to happen' we refer an action to an altogether different category from the factual. That an action 'ought to be done' is regarded as something which holds true altogether without reference to whether it actually is done or not." The whole notion that there is a determinate character of an action that would make a moral statement true or false is, Hägerström argues, an illusion. There is nothing there for an "unmoved spectator of the actual" to observe that would either confirm or disconfirm his moral statements. Moral statements characteristically take a declarative form, but they actually function not to assert that so-and-so is true but to express an attitude toward an action or a state of affairs.

The emotive theory developed as a *via media* between intuitionism, on the one hand, and ethical naturalism, on the other. Both of these ethical theories displayed crucial difficulties. "Nonnatural qualities" and "nonnatural relations" were obscure, fantastic conceptions, to say the least, and the notion of intuition remained at best nonexplanatory. Furthermore, it was plain that moral judgments are closely linked to one's emotions, attitudes, and conations. But, as Moore in effect showed, neither "A cup of tea before bed is good" nor such general utterances as "Pleasure is good" and "Self-realization is good" are empirical or analytic.

The function of ethical statements. The emotivists maintained that while the grammatical function of a sentence like "A swim before bed is good" is indicative, its actual logical function is much closer to that of an optative or imperative utterance, such as "Would that we could go swimming before bed" or "Swim before bed." Because of this, emotivists have claimed that it is misleading to say that ethical sentences can be used to make statements: They do not function to assert facts.

Similarly, it is a mistake to treat all words as simply functioning to describe or designate some characteristic

or thing. Some words so function; but there are other words, like *nasty, saintly, graceful,* and *wise,* that function primarily or in part to express the attitudes of the utterer or to evoke reactions on the part of the hearer. The emotivists claim that *good, ought, right,* and the like are also emotive words. This gives them their normative function.

Ethical argument. Hägerström and Ayer contend that the fact that there are no moral facts carries with it the corollary that there can be no genuine moral knowledge. There are no moral facts to be learned; there is no moral information to be gained or forgotten. It makes clear sense to say "I used to know the difference between a pickerel and a pike, but by now I've forgotten it," but what is meant by "I used to know the difference between right and wrong, but by now I've forgotten it"? The word *forgotten* could hardly do its usual job here. The utterance is so deviant that without explanation and a very special context, we do not understand it. Considerations of this sort bring us to the realization that moral utterances are not used to state facts or assert truths; their essential role is a noncognitive one. They typically express emotions, attitudes, and conations and evoke actions, attitudes, and emotional reactions.

Because of this fact about the logical status of moral utterances, it always remains at least logically possible that two or more people might agree about all the relevant facts and disagree in attitude—that is, disagree about what was desirable or worth doing.

We do, however, as Ayer and Stevenson stress, give reasons for moral judgments. If I say "MacDonald did the right thing in killing Janet," it is perfectly in order to ask me to show why this is so. If I say "I don't have any reasons. There aren't any reasons, but all the same I just know that MacDonald did the right thing," I am abusing language. I am saying something unintelligible, for we cannot "just know" like that. The person who claims that an action is right must always be prepared to give reasons for his moral claim.

Ayer and Stevenson grant all that. This is indeed how we do proceed when we are being reasonable about a moral disagreement. But Ayer says: "the question is: in what way do these reasons support the moral judgments? They do not support them in a logical sense. Ethical argument is not formal demonstration. And they do not support them in a scientific sense either. If they did, the goodness or badness of the situation, the rightness or wrongness of the action, would have to be something apart from the situation, something independently verifiable, for which the facts adduced as the reasons for the moral judgment were the evidence." But this is just what we cannot do. There is no procedure for examining the value of the facts, as distinct from examining the facts themselves.

If we cannot demonstratively prove or inductively establish fundamental moral claims, then what can it mean to say that a factual statement *F* is a good reason for a moral judgment *E*? The emotivist's answer is very simple: If *F* causes the person(s) to whom *E* is addressed to adopt *E,* to share the attitude expressed by *E,* then *F* is a good reason for *E.* It is Ayer's and Stevenson's claim that whatever in fact determines our attitudes is ipso facto a good reason for a moral judgment.

Criticisms of emotive theory. It has been argued by many moral philosophers (W. D. Falk, Richard Brandt, Errol Bedford, Paul Edwards, and Kai Nielsen, among others) that so to characterize what is meant by "a good reason" in ethics is persuasively to redefine "a good reason" in ethics. As Bedford has well argued against the emotive theory, "we do use logical criteria in moral discussion, however inexplicit, unanalyzed, and relatively vague these criteria of relevance may be." Remarks like "It doesn't follow that you ought to" or "That's beside the point" are just as common and just as much to the point in moral argument as elsewhere. There is no reason to think that these remarks about relevance differ in any essential way from their use in nonevaluative contexts. We don't just seek agreement when there is a moral dispute, but we try to justify one claim over another and we rightly reject persuasion as irrelevant to this task of justification.

Stevenson has replied that to answer in this way is in effect to confuse normative ethical inquiries with metaethical ones. *Good* and *relevant* are normative terms and have their distinctive emotive force. To say that such and such are good reasons is to make a moral statement. Making such a statement involves leaving the normative ethical neutrality of metaethical inquiry. One answer to this is that to say what is meant by "good reasons" in ethics is to mention "good reasons" and not to use them.

EXISTENTIALISM. Noncognitivism is not limited to emotivism. The existentialists do not call themselves noncognitivists, nor do they write metaethical treatises. But reasonably definite metaethical assumptions are implicit in their writings. Their contention that "men create their values," their stress on decision, commitment, and the impossibility of achieving ethical knowledge, strongly suggests a noncognitivist metaethic. We shall limit the examination here to two major figures, Albert Camus and Jean-Paul Sartre.

Camus. Unlike Sartre, Albert Camus wrote no technical philosophy, but in his *Myth of Sisyphus* (Paris, 1942), *The Rebel* (Paris, 1951), and his plays and novels he did articulate an ethical view that has been called the ethics of the absurd. To read Camus is to be immediately thrown into normative ethics via what has been called philosophical anthropology. We are immediately confronted with a picture of man and man's lot. Man is divorced from the world yet is paradoxically thrust into it. The world as we find it—given our hopes, our expectations, our ideals—is intractable. It is incommensurate with our moral and intellectual demands. Life is fragmented. We seek to discover some rational unity amidst this diversity and chaos. We discover instead that we can only impose an arbitrary unity upon it. *L'homme absurde,* as distinct from *l'homme quotidien,* sees clearly the relativity and flux of human commitment and the ultimate purposelessness of life. Yet man has a blind but overpowering attachment to life as something more powerful than any of the world's ills or any human intellectualization. But the world is ultimately unintelligible and irrational, and man's lot in the world is absurd.

Given this situation, all moral commitments are arbitrary. There is no escaping this: Reason will only show us the arbitrariness of human valuations, and a Kierkegaardian leap of faith in the face of the absurd is evasive. It is evasive because it is to consent to absurdity rather than to face up to it, recognizing it for what it is. Man's dignity comes in his refusing to compromise. His very humanity is displayed in his holding on to his intelligence and in recognizing, contra Kierkegaard, that there is no God and, contra Karl Jaspers, that there is no metaphysical unity that can overcome the absurdity of human existence.

Yet paradoxically, and some would claim inconsistently, in his novel *The Plague* (Paris, 1947), and in his essays, collected and published in English under the title *Resistance, Rebellion and Death* (New York, 1961) Camus writes with passion and conviction in defense of human freedom and intelligence. Camus's rationale for this is that we become *engagé* because we see that life has no ultimate meaning and that, finally free from a search for cosmic significance, we can take the diverse experiences of life for what they are in all their richness and variety. Yet beyond that and perhaps because of that, Camus, as a humanist, is espousing the cause of man. By this is meant, as is very evident in *Resistance, Rebellion and Death,* that Camus repeatedly defends human freedom, equality, and the alleviation of human misery and deprivation. We must become involved, but in this involvement Camus urges a reliance on human intelligence in facing the problems of men.

What might be taken to be a conflict between the more theoretical side of Camus's thought and his more directly normative ethical side comes out in his fourth "Letter to a German Friend." Camus agrees with his "German friend" that the world has no ultimate meaning, but he does not and will not conclude from this, as his "German friend" did, "that everything was equivalent and that good and evil could be defined according to one's wishes." Camus then goes on to remark that he can find no valid argument to answer such a nihilism. His only "answer" is "a fierce love of justice, which after all, seemed to me as unreasonable as the most sudden passion." Camus felt he could only resolutely refuse to accept despair and "to fight against eternal injustice, create happiness in order to protest against the universe of unhappiness." Camus concludes with a cry of the heart that while "the world has no ultimate meaning … something in it has a meaning, namely man because he is the only creature to insist on having one."

Sartre. Jean-Paul Sartre's views on man's condition are in many important respects like those of Camus, but to a far greater degree than Camus, Sartre in *Being and Nothingness* (Paris, 1943) and *Critique de la raison dialectique* (Paris, 1960) sets his ethical theorizing in the murky atmosphere of metaphysics. The promised systematic work on ethics that was to follow *Being and Nothingness* has not been forthcoming, but in one way or another all of Sartre's works are concerned with ethics. It can be said that there are two Sartres, or at least that the Sartre of *Critique de la raison dialectique* has moved from his earlier existentialism over to a kind of Marxist materialism. Here we shall for the most part (except where specifically noted) be concerned with the earlier Sartre, whose philosophical endeavor centered on his massive *Being and Nothingness.*

Sartre, like Camus, finds man's lot in the world absurd. Since there is no God, life can have no ultimate meaning and there can be no objective knowledge of good and evil. We cannot "decide a priori," or find out by investigation, what we are to do. Man in his forlornness and freedom imposes values. The choices man makes, the projects he forms for himself, and the sum of his acts constitute his values. There is no good and evil to be intuited or in any way discovered by the human animal. Man in anguish creates his values by his deliberate choices, and, to add to his anguish, in making his choices "he involves all mankind." That is to say, Sartre stresses the Kantian claim that moral judgments, in order to be moral judg-

ments, must be universalizable, but, as Sartre adds in his lecture "Existentialism Is a Humanism" (1945), though their "form is universal … the content of ethics is variable" and there is no rational way of justifying the acceptance of moral principles with one content rather than another.

Sartre thinks this position is simply a matter of drawing out in a nonevasive manner the implications of a consistent atheism. Only if there were a God could values have an objective justification, but without God "everything is permissible" and "as a result man is forlorn, because neither within him nor without does he find anything to cling to." In this, Sartre is surely mistaken. It does not follow that if there is no God, nothing matters, or that everything is permissible. It is not a contradiction to assert, "Though there is no God, the torturing of children is still vile," and the nonexistence of God does not preclude the possibility of there being an objective standard on which to base such judgments.

Sartre asserts flatly, in good Moorean spirit, "Ontology itself cannot formulate ethical precepts. It is concerned solely with what is, and we cannot possibly derive imperatives from ontology's indicatives." (All the same, his account of morality in *Being and Nothingness* and his account of human action relevant to morality are immersed in "the language of being.") In fact, Sartre goes on to point out that ontology and what he calls existential psychoanalysis can in a given situation constitute "a moral description, for it presents to us the ethical meaning of various human projects." This method of description—though hardly the descriptions themselves—is very like the phenomenological method practiced by Scheler and Hartmann. Yet to proceed in this way hardly constitutes a violation of the is/ought distinction, since Sartre's descriptions of moral evaluations—descriptions of man's ethical life—need not themselves be evaluative, though given the language Sartre uses, they often are.

"Man," he tells us, "pursues being blindly by hiding from himself the free project which is this pursuit." Existential psychoanalysis can reveal to man the real goal of his pursuit. Horrified by the "death of God," man attempts in his anguish to be God. He flees from his freedom—he does not wish to be a creator of values—but in what Sartre ironically calls the spirit of seriousness, he seeks to deny human subjectivity and attributes to value some independent cosmic significance. To the extent that we are caught up in this spirit of seriousness, we will try to fuse "being-for-it-self" with the brute facticity of "being-in-itself." (The odd phrase "being-in-itself" is simply the label for the self-contained reality of a thing, while its mate, "being-for-itself," is the label for the realm of consciousness that perpetually strives to transcend itself.) But if we pursue this line, we still condemned to despair, for we "discover at the same time that all human activities are equivalent … and that all are on principle doomed to failure." Phenomenological analysis reveals to man that though he perpetually tries to become a thing, a brute existent, the fact that he has consciousness makes this impossible. Given this ability to think and to feel, man, whether he likes it or not, is slowly led to see that without God he can have no essential nature; that is, though he may form his own projects, there is and can be no purpose to life.

It should be noted that Sartre's view of man's lot is even grimmer than Camus's, for Sartre contends that even in community with others there is no surcease from suffering and alienation, for human relations are essentially relations of conflict and estrangement.

In *Critique de la raison dialectique,* Sartre tries to work out a new kind of Marxism and a new materialist conception of man. But he wishes to integrate his existentialist conceptions into a Marxist materialism in such a way that the latter can come to have a truly "human dimension." Marxism, he argues, must purge itself of its deterministic conceptions of man and acknowledge a rational conception of human freedom. Sartre, in a reversal from *Being and Nothingness,* now argues that there is nothing intrinsic in human nature that makes conflict, war, and a reign of terror inescapable, though, like a good Marxist, he does argue that conflict is a basic factor in human history. It is scarcity, scarcity of goods and materials, that triggers human conflict. Only under these conditions of scarcity is social conflict inescapable and a rational social order impossible. Men make their own history by the choices they make in the face of problems created by history. But man remains the rider, not the horse. Human choices—human projects—are still free choices for which men remain responsible.

RECENT VIEWS ON MORAL DISCOURSE

LINGUISTIC PHILOSOPHY. As has frequently been noted, there are at least superficial resemblances between the existentialists and the otherwise very different, self-consciously metaethical theories of such linguistic philosophers as R. M. Hare, P. H. Nowell-Smith, Bernard Mayo, Alan Montefiore, and John Hartland-Swann.

There is, indeed, this much similarity between these linguistic philosophers and the existentialists. All of the

former make the following contentions, all of which would be welcome to the latter:

(1) Moore was essentially right about the naturalistic fallacy. That is to say, moral statements cannot be deduced from any statement of fact, whether biological, historical, psychological, sociological, or religious.

(2) No moral choice or question of value can ever be guaranteed by logical rules.

(3) We are free, as far as language or logic is concerned, to apply evaluative or prescriptive terms to anything we wish to commend or condemn, criticize or approve, prescribe or forbid.

(4) Moral utterances are generalizable decisions, resolutions, or subscriptions.

Given that a man accepts certain moral principles, other moral principles can, together with certain factual statements, be derived from the above principles. But like Ayer and the existentialists, these linguistic philosophers hold that there must be some moral principles which are not derived from any other principles—moral or otherwise—and, being fundamental moral principles, they are not even verifiable in principle. They express moral commitments and can have no rational ground, for what is deemed worthy of acceptance ultimately depends on the very commitments (generalizable decisions, resolutions, or subscriptions) an agent is willing to make.

Many people have thought that such a view of morality is either directly or indirectly nihilistic—that both the linguistic philosophers and the existentialists espouse what is in effect an irrationalism that would undercut the very possibility of a rational normative ethic.

If we consider a reply linguistic philosophers typically make to such criticisms, we will become aware of a crucial dissimilarity between them and the existentialists and a fundamental defect in existentialist ethics.

Linguistic philosophers have frequently claimed that the existentialists have merely dramatized a logical point. That moral principles are expressions of commitment or choice, that man cannot simply discover what is good or evil or know a priori that a certain thing must be done but must "create his own values," is not a worrisome fact about the human predicament; it is a conceptual truth concerning the nature of moral discourse. It is not a fact of the human condition that man is born into a world alien and indifferent to human purposes. What is a fact is that the phrases "the universe has a purpose" and "value and being are one" are unintelligible phrases. To say "man creates his own values" is in reality only to say in a dramatic way that a judgment of value is an expression of choice. This statement, it is argued, is not an anguished cry of the human heart but is merely an expression of a linguistic convention.

To say "If *x* is a judgment of value, then *x* is an expression of choice" is not to say "Any choice at all is justified," "Anything is permissible," or "All human actions are of equal value." These latter statements are themselves value judgments and could not follow from the above-mentioned statement, for it is not itself a statement of value but a nonnormative metaethical statement about the meaning of evaluative expressions, and, as Sartre himself stresses, one cannot derive an "ought" from an "is." In general, Hare and Nowell-Smith, as well as Ayer and Stevenson, stress the normative neutrality of metaethical statements.

Hare. R. M. Hare in two very influential books, *The Language of Morals* (Oxford, 1952) and *Freedom and Reason* (Oxford, 1963), developed a very closely reasoned metaethical analysis of the type that has been discussed. In *The Language of Morals,* Hare views moral utterances as a species of prescriptive discourse, and he feels that we can most readily come to understand their actual role in the stream of life if we see how very much they are like another form of prescriptive discourse, namely, imperatives. Imperatives tell us to do something, not that something is the case. Moral utterances in their most paradigmatic employments also tell us to do something. Imperative and moral utterances do not, as the emotivists thought, have the logical function of trying to get you to do something. Rather, they tell you to do something. Furthermore, there are logical relations between prescriptive statements, just as there are logical relations between factual statements.

Moral judgments are viewed as a kind of prescriptive judgment but, unlike singular imperatives, moral judgments (as well as all value judgments) are universalizable. Hare means by this that such a judgment "logically commits the speaker to making a similar judgment about anything which is either exactly like the subject of the original judgment or like it in the relevant respects."

Hare stresses that while almost any word in certain contexts can function evaluatively, *good, right,* and *ought* almost always so function. The evaluative functions of these terms are distinct from their descriptive functions and are an essential part of their meaning. In fact, the distinctive function of all value words is that they in one way or another commend or condemn. But while *good* is a general word of commendation, the criteria for goodness

vary from context to context and are dependent on what it is that is said to be "good."

The meaning of *good* or any other value term is never tied to its criteria of application. There is nothing in the logic of our language to limit the content of a moral judgment. As far as logic is concerned, any universalizable prescription that expresses a deep concern or commitment is ipso facto a moral prescription, and we can decide without conceptual error to do anything that it is logically or physically possible to do. If we treat the resulting decision as a decision of principle, that is, a universalizable prescription, then it is a value judgment that is in good logical order. As Nowell-Smith has well put it in discussing Hare's theory, "Nothing that we discover about the nature of moral judgments entails that it is wrong to put all Jews in gas-chambers."

Criticism of Hare. Probably the most persistent dissatisfaction with Hare's theory has resulted from the belief that it makes moral reasoning appear to be more arbitrary than it actually is. To say "Nothing that we discover about the nature of moral judgments entails that it is wrong to put all Jews in gas-chambers" is, it will be argued, a reductio of such a position. Hare would reply that to argue in such a way is to fail to recognize that he is talking about entailment, and that he is simply making the point that from nonnormative statements one cannot deduce normative ones.

Hare argues that his thesis about the logical status of moral utterances does not commit him to the position that there can be no rational resolution of basic conflicts in moral principle. Returning, in *Freedom and Reason,* to a stress on decisions (though with a new attention to inclinations), Hare contends that to have a morality we must have freedom. Specifically, we must have a situation in which each man must solve his own moral problems. (This is not to moralize about what we should do but to state a logical condition for the very existence of moral claims.)

Philosophers who have criticized Hare, including someone as close to him as Nowell-Smith, have suggested that Hare still has a far too Protestant conception of moral discourse. He fails really to take to heart the Wittgensteinian claim that here, as elsewhere in human discourse, we must have public criteria for what could count as a logically proper moral claim. As F. E. Sparshott—whose book *An Enquiry into Goodness* (Chicago, 1958) deserves more attention than it has received—notes: Hare's individualism leads him to neglect the fact that a morality, any morality, will necessarily incorporate "those rules of conduct that seem necessary for communal living." It is not the case that just any universalizable set of prescriptions can constitute a morality or a set of moral judgments.

THE GOOD-REASONS APPROACH. The last metaethical theory we shall discuss has been dubbed the good-reasons approach. Stephen Toulmin, Kurt Baier, Henry Aiken, Marcus Singer, Kai Nielsen, A. I. Melden, A. E. Murphy, and John Rawls may be taken as representative figures of this point of view. It is an approach that obviously has been deeply affected by the philosophical method that we have come to associate with the work of the later Ludwig Wittgenstein. These philosophers have centered their attention on the logic of moral reasoning. Their central question has been "When is a reason a *good* reason for a moral judgment?" Accordingly, the crucial problems center on questions concerning the nature and limits of *justification* in ethics. These philosophers agree with the noncognitivists that moral sentences are used primarily as dynamic expressions to guide conduct and alter behavior. And they would also agree with ethical naturalists that moral utterances usually, at least, also make factual assertions. But they believe that the primary use of moral utterances is not theoretical or just emotive but practical. Hare and Nowell-Smith are right in stressing that they are designed to tell us what to do.

Yet while moral utterances typically tell us what to do, language with its complex and multifarious uses does not neatly divide into "the descriptive" and "the evaluative," "the constative" and "the performative," "the cognitive" and "the noncognitive." These are philosophers' specialized terms, and they do not help us to understand and clearly characterize moral discourse but actually distort our understanding of it. There can be no translation of moral terms into nonmoral terms, and the ancient problem of bridging "the is-ought gulf" is a muddle, for there is no clear distinction between such uses of language and no single function that makes a bit of discourse normative. Some moral utterances indeed bear interesting analogies to commands or resolutions, but they cannot be identified with them. It is a mistake to think ethical judgments are like scientific ones or like the judgments of any other branch of objective inquiry; yet cognitivist metaethicists were correct, not in pressing this analogy but in maintaining that there is a knowledge of good and evil and that some moral claims have a perfectly respectable objectivity. No matter how emotive or performative moral utterances may be, when we make a moral judgment, it must—logically must—satisfy certain requirements to count as a moral judgment. In making a moral judgment, we must be willing to universalize the

judgment in question, and it must be possible to give factual reasons in support of the moral claim.

The advocates of the good-reasons approach in the general tradition of the later Wittgenstein did not take it to be incumbent on the philosopher to translate moral utterances into some clearer idiom. They did not believe that there was some other favored discourse or form of life that moral discourse or morality should be modeled on. What was expected of the philosopher was that he should describe morality so as to perspicuously display the living discourse at work. In particular, philosophers should concern themselves with a conceptual cartography of the nature and limits of justification in ethics. Before we can reasonably claim that moral judgments are at bottom "all subjective" or that no moral claim can be "objectively justified," we must come to understand what can and what cannot count as a good reason in ethics and what the limits of moral reasoning are.

Toulmin. Two books, Stephen Toulmin's *An Examination of the Place of Reason in Ethics* (Cambridge, U.K., 1950) and Kurt Baier's *The Moral Point of View* (Ithaca, NY, 1958), have most single-mindedly attacked the problem of moral reasoning. They may be taken as paradigms of the good-reasons approach. Toulmin argues that moral rules and moral principles are to be justified by discovering which of these rules or principles, if consistently acted upon, will most likely lead to the least amount of avoidable suffering all around. Those social practices that probably will cause the least amount of suffering for humankind are the social practices that ought to be accepted. Classical utilitarians maintained that a moral rule is justified if it tends to produce greater happiness all around than any alternative rule, but Toulmin favors the negative formulation because (1) though it is very difficult to determine what will make people happy or what they want, it is less difficult to determine what causes suffering, and (2) it is less the function of morality to tell men what the good life is than to tell them what not to do so that their interests, including their differing conceptions of the good life, can be realized to the maximum extent. This theory about moral reasoning, while purporting to be metaethical, is very close to the normative ethical theory sometimes called rule utilitarianism.

Toulmin argues that if we examine closely the way moral reasoning is actually carried on, it will become evident that moral rules and practices are characteristically judged by roughly utilitarian standards, while many individual actions are judged by whether or not they are in accordance with an accepted moral rule or social practice. Utilitarians point out that it is of the greatest social utility that we characteristically judge moral acts in this seemingly nonutilitarian fashion. However, frequently a decision concerning how to act involves conflicting moral rules with no clear order of subordination, and in some situations there seems to be no moral rule—unless the principle of utility is taken as a moral rule—that is readily applicable. In such a situation, the thing to do is to act on a utilitarian basis when it is at all possible to make some reasonable judgment of the probable beneficial consequences to the people involved of doing one thing rather than another. If that is not possible in a given situation, then we should act as a reasonable man would act. (The concept of a reasonable man, we should not forget, is itself very much a moral concept.)

Criticism of Toulmin. There certainly are a host of objections that spring to mind concerning Toulmin's account. First, it will be said that this is normative ethics, not metaethics: It tells us what we should do, what a good reason is, and how we can justify basic moral rules. Moreover, why should we accept it? Once we see through its modish trappings, it will become apparent that it has all the difficulties attendant on classical utilitarianism.

It could be replied that though the speech is in the material mode and sounds like normative ethics, in reality it is a brief description of how moral reasoning is actually carried on. Even if this reply is accepted, there are difficulties here too, for viewed this way, Toulmin's account surely looks like an account of a basically sociological sort of how certain people in fact reason. That is to say, it appears to be an impressionistic bit of descriptive ethics and hardly a metaethical account of the logic of moral reasoning. It covertly and persuasively redefines as "moral" only a very limited pattern of reasoning—reasoning that expresses the historically and ethnographically limited views of a determinate group of people. The ethnocentric character of this linguistic analysis makes it implicitly, but surreptitiously, normative.

This contention will be rejected by many. It will be argued that moral reasoning, like any other mode of reasoning, is limited. To determine what the moral point of view is and what it is to reason morally, we need first to determine the function (purpose, overall rationale) of morality.

The function of morality, Toulmin tells us, is to adjudicate conflicting interests and to harmonize desires (that is, moderate our impulses and adjust our demands) so as to reconcile them with our fellows, in such a way that everyone can have as much as possible of whatever it is that, on reflection, he wants. Given this conception of the function of ethics, something like Toulmin's account of

moral reasoning is very plausible, but it has been objected that morality has no *one* such function. Many people have ideals of human excellence that have nothing to do with such a conception of the function of ethics: Many Jews and Christians, with their ideals of the love of God, do not conceive the function of moral living in this way, and the Buddhist community with its ideals of arhatship certainly would not accept, either in theory or in practice, such a conception of the function of ethics. Morality is a much more complicated and varied activity. There are diverse and often conflicting functions of morality. Any attempt to claim one function or rationale of morality as the function or the purpose of morality so circumscribes what can count as moral considerations that its effect is unwittingly to advocate one limited moral outlook as the moral point of view.

Finally, even if Toulmin could make out a case for claiming that the function of morality, or the primary function of morality, is such as he claims it to be, one could still ask, concerning this descriptive account of morality, "Why keep it as the sole or primary function of morality?" If altering the function of morality somewhat alters the meaning of "moral," then why should we be such linguistic conservatives? What is so sacred about that function of morality and its attendant conception of morality?

Toulmin could claim that now his critic has confused normative issues with metaethical ones. The issues here are complex and lead us into the heart of current discussion about the nature of moral reasoning. Yet a strong case can be made for the contention that there is more to be said for a general approach such as Toulmin's and Baier's than has commonly been thought.

It seems evident that much contemporary thinking about ethics, while devoted to Moore's exacting standards of making perfectly clear precisely what is being claimed, is concerned not with the very general question of the meaning of *good* or, for that matter, *right* or *ought* but with the rich texture of moral reasoning. This brings once more to the foreground the kind of detailed descriptions of the moral life distinctive of such phenomenologists as Scheler and Hartmann, but given the present care for actual conceptual distinctions, we may develop a kind of linguistic phenomenology that may be of major importance to an understanding of morality. Perhaps the most exciting endeavors from this point of view have been those of Rawls, Philippa Foot, and Georg von Wright. Rawls, in a series of distinguished essays, has shown the central role of considerations of justice in moral deliberation and the way such considerations modify utilitarian patterns of reasoning; Foot, also in a series of much-discussed essays, has shown the importance of a discussion of the virtues and the vices and has reinvigorated ethical naturalism. Wright's masterful discussion of the varieties of goodness in his *The Varieties of Goodness* (London, 1963) has contributed immensely to our understanding of morality.

See also Abelard, Peter; Albert the Great; Alcmaeon of Croton; Anselm, St.; Antisthenes; Aristippus of Cyrene; Aristotle; Augustine, St.; Averroes; Avicenna; Ayer, Alfred Jules; Bacon, Francis; Baier, Kurt; Beccaria, Cesare Bonesana; Bentham, Jeremy; Bernard of Clairvaux, St.; Bosanquet, Bernard; Bradley, Francis Herbert; Brandt, R. B.; Brentano, Franz; Broad, Charlie Dunbar; Butler, Joseph; Calvin, John; Cambridge Platonists; Camus, Albert; Carnap, Rudolf; Carneades; Chrysippus; Cicero, Marcus Tullius; Clarke, Samuel; Cleanthes; Clement of Alexandria; Cudworth, Ralph; Cumberland, Richard; Cynics; Cyrenaics; Descartes, René; Dewey, John; Diderot, Denis; Diogenes of Sinope; Duns Scotus, John; Emerson, Ralph Waldo; Empedocles; Engels, Friedrich; Epictetus; Epicurus; Epicureanism and the Epicurean School; Erasmus, Desiderius; Erigena, John Scotus; Ethics; Ethics and Morality; Fichte, Johann Gottlieb; Galileo Galilei; Gorgias of Leontini; Green, Thomas Hill; Hägerström, Axel; Hare, Richard M.; Hartley, David; Hartmann, Nicolai; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Hellenistic Thought; Helvétius, Claude-Adrien; Heraclitus of Ephesus; Hobbes, Thomas; Holbach, Paul-Henri Thiry, Baron d'; Hume, David; Husserl, Edmund; Hutcheson, Francis; James, William; Kant, Immanuel; Kierkegaard, Søren Aabye; Leibniz, Gottfried Wilhelm; Lenin, Vladimir Il'ich; Leucippus and Democritus; Locke, John; Luther, Martin; Machiavelli, Niccolò; Maimonides; Malebranche, Nicolas; Mandeville, Bernard; Mani and Manichaeism; Marcus Aurelius Antoninus; Marx, Karl; Mill, James; Mill, John Stuart; Montesquieu, Baron de; Moore, George Edward; More, Henry; Murphy, Arthur Edward; Neoplatonism; Nicolas of Autrecourt; Nietzsche, Friedrich; Origen; Paley, William; Panaetius of Rhodes; Pascal, Blaise; Perry, Ralph Barton; Philo Judaeus; Plato; Plotinus; Porphyry; Posidonius; Price, Richard; Prior, Arthur Norman; Protagoras of Abdera; Pyrrho; Rawls, John; Reid, Thomas; Religion and Morality; Ross, William David; Rousseau, Jean-Jacques; Royce, Josiah; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Scheler, Max; Schopenhauer, Arthur; Seneca, Lucius Annaeus; Shaftesbury, Third Earl of (Anthony Ashley

Cooper); Sidgwick, Henry; Skepticism, History of; Smith, Adam; Socrates; Sophists; Spinoza, Benedict (Baruch) de; Stevenson, Charles L.; Stoicism; Suárez, Francisco; Tertullian, Quintus Septimius Florens; Thomas Aquinas, St.; Voltaire, François-Marie Arouet de; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

GENERAL HISTORIES OF ETHICS

Historical and Social Context of Ethical Beliefs

Brinton, C. C. *A History of Western Morals.* New York: Harcourt Brace, 1959. A very readable history of ethical beliefs and practices.

Bruce, A. B. *The Moral Order of the World in Ancient and Modern Thought.* 4 vols. London: Hodder and Stoughton, 1899. A detailed study of the literary and religious background of ethical thought from ancient Greece to the Reformation.

Dittrich, Ottmar. *Geschichte der Ethik.* 4 vols. Leipzig: Meiner, 1926. A scholarly study of the literary and religious origins of Western ethical ideals.

Harkness, G. E. *The Sources of Western Morality.* New York: Scribners, 1954. An exploration of the ancient origins of modern ethics. Useful comparisons of primitive with Old Testament ethics and of Greek with Christian ethics.

Kropotkin, P. A. *Ethics, Origin and Development.* Translated by L. S. Friedland and J. R. Piroshnikoff. New York: Dial Press, 1924. A Marxist account of the development of ethics from primitive to modern Western society.

Lecky, W. E. H. *History of European Morals.* 2 vols. New York, 1919. A justly celebrated study of how moral values reflect, but fail to keep up with, actual practices in the Western world. Vol. I, Ch. 1, discusses modern British ethical philosophy.

Mencken, H. L. *Treatise on Right and Wrong.* New York: Knopf, 1934. Perceptive and often caustic comments on the evolution of moral ideas in philosophy and religion; a critique of authoritarianism.

Robertson, J. M. *A Short History of Morals.* London: Watts, 1920. A relativistic view of the history of ethics, stressing the sociological sources of morality.

Westermarck, Edward. *The Origin and Development of Moral Ideas.* 2 vols. London: Macmillan, 1908. A monumental work of historical and anthropological scholarship.

Histories of Ethical Philosophy

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. A critical study, by a distinguished intuitionist, of the ethical systems of Spinoza, Butler, Hume, Kant, and Sidgwick.

Kautsky, Karl. *Ethics and the Materialist Conception of History.* Translated by J. B. Agnew. Chicago: Kerr, 1907. A brief sketch of the evolution of ethical thought by a noted Marxist.

MacIntyre, Alasdair. *A Short History of Ethics.* New York: Macmillan, 1966. An important discussion by a contemporary analytic philosopher.

Mackinnon, D. M. *A Study in Ethical Theory.* London: A. and C. Black, 1957. A comparison of the theories of Kant, Butler, and the utilitarians as alternatives to religious ethics.

Maritain, Jacques. *Moral Philosophy.* Translated by M. Suther. New York: Scribners, 1964. A critical evaluation of secular ethical theories from the standpoint of Catholic theology.

Martineau, James. *Types of Ethical Theory.* 2 vols. Oxford: Clarendon Press, 1898. Vol. I evaluates the metaphysical and ethical views of Plato, Descartes, Malebranche, Spinoza, and Comte. Vol. II criticizes utilitarian and moral-sense theories.

Rogers, R. A. P. *A Short History of Ethics.* London: Macmillan, 1911. A brief survey of the history of ethical theory. Incomplete but clear and readable.

Sidgwick, Henry. *Outlines of the History of Ethics.* London, 1886. A brief but illuminating discussion of the most important schools of ethical thought from ancient Greece to nineteenth-century England.

Swabey, W. C. *Ethical Theory: From Hobbes to Kant.* New York: Philosophical Library, 1961. A useful summary of several philosophical systems.

PLATO

General Expositions

Dickinson, G. L. *Plato and His Dialogues.* New York: Norton, 1932.

Field, G. X. *The Philosophy of Plato.* London, 1951.

Friedländer, Paul. *Plato.* Translated by Hans Meyerhoff. New York, 1958–1964. A multivolume work of tremendous scholarship on the cultural background of Plato's writings.

Shorey, P. *What Plato Said.* Chicago: University of Chicago Press, 1933.

Taylor, A. E. *Plato, the Man and His Work.* London, 1908.

Specialized Studies

Lodge, R. C. *Plato's Theory of Ethics.* New York: Harcourt Brace, 1928. A thorough study limited to Plato's ethical ideas.

Nettleship, R. L. *Lectures on the Republic of Plato.* New York, 1962. An extraordinarily illuminating study of Plato's most important dialogue.

Relation of Plato to Other Greek Thinkers

Burnet, John. *Platonism.* Berkeley: University of California Press, 1928. Distinguishes sharply between Socratic and Platonic views.

Cornford, F. M. *Before and After Socrates.* Cambridge, U.K.: Cambridge University Press, 1960.

Demos, R. *The Philosophy of Plato.* New York: Scribners, 1939. Rejects any systematic interpretation of Plato.

Field, G. C. *Plato and His Contemporaries.* London: Methuen, 1948.

Gould, J. *The Development of Plato's Ethics.* Cambridge, U.K.: Cambridge University Press, 1955.

Hampden, R. D. *The Fathers of Greek Philosophy.* Edinburgh: A. and C. Black, 1862.

Merlan, Philip. *From Plato to Platonism.* The Hague, 1960.

Plato's Ethics and Politics

Crossman, R. H. *Plato Today.* New York: Oxford University Press, 1939. Critical.

Fite, W. *The Platonic Legend.* New York: Scribners, 1934. Critical.

Koyré, Alexandre. *Discovering Plato.* New York: Columbia University Press, 1945. Favorable.

Levinson, R. B. *In Defense of Plato.* Cambridge, MA: Harvard University Press, 1953.

Wild, J. D. *Plato's Modern Enemies.* Chicago: University of Chicago Press, 1953. Favorable.

ARISTOTLE

Aristotle's Ethics

Gauthier, R. A. *La morale d'Aristote.* Paris, 1958.

Grene, Marjorie. *A Portrait of Aristotle.* London: Faber and Faber, 1963.

Marshall, T. *Aristotle's Theory of Conduct.* London: Unwin, 1906.

Mure, G. R. G. *Aristotle.* London, 1932.

Oates, W. J. *Aristotle and the Problem of Value.* Princeton, NJ: Princeton University Press, 1963.

Randall, J. H. *Aristotle.* New York, 1960. Chs. 11–14.

Ross, W. D. *Aristotle.* London, 1930.

Veatch, H. B. *Rational Man.* Bloomington: Indiana University Press, 1962.

Walsh, J. J. *Aristotle's Conception of Moral Weakness.* New York: Columbia University Press, 1963.

Aristotle and Other Ethical Philosophers

Cornford, F. M. *Before and After Socrates.* Cambridge, U.K.: Cambridge University Press, 1960.

Hampden, R. D. *The Fathers of Greek Philosophy.* Edinburgh: A. and C. Black, 1862.

Merlan, Philip. *Studies in Epicurus and Aristotle.* Wiesbaden: Harrassowitz, 1960.

Taylor, A. E. *Aristotle.* London, 1943.

HELLENISTIC AND ROMAN ETHICS

Epicureanism

DeWitt, N. W. *Epicurus and His Philosophy.* Minneapolis: University of Minnesota Press, 1954.

Hadzsits, G. D. *Lucretius and His Influence.* New York: Longmans Green, 1935.

Merlan, Philip. *Studies in Epicurus and Aristotle.* Wiesbaden: Harrassowitz, 1960.

Stoicism

Arnold, E. V. *Roman Stoicism.* Cambridge, U.K.: Cambridge University Press, 1911. Middle and late Stoa.

Bréhier, Émile. *Chrysippe et l'ancien Stoicisme.* Paris: Presses Universitaires de France, 1951. Early Stoa.

Brussell, F. W. *Marcus Aurelius and the Later Stoics.* New York, 1910.

Murray, Gilbert. *Stoic, Christian and Humanist.* Boston: Beacon Press, 1950.

Wenley, R. M. *Stoicism and Its Influence.* Boston: Marshall Jones, 1924.

Neoplatonism

Elsee, C. *Neoplatonism in Relation to Christianity.* Cambridge, U.K.: Cambridge University Press, 1908.

Feibleman, J. *Religious Platonism.* London: Allen and Unwin, 1959.

Pistorius, P. V. *Plotinus and Neoplatonism.* Cambridge, U.K.: Bowes and Bowes, 1952.

Switalski, B. *Neoplatonism and the Ethics of St. Augustine.* New York: Polish Institute of Arts and Sciences in America, 1946.

Whittaker, T. *The Neoplatonists.* Cambridge, U.K., 1918.

Comparisons

Hicks, R. D. *Stoic and Epicurean.* New York: Scribners, 1910.

Zeller, Eduard. *Stoics, Epicureans and Sceptics.* Translated by O. J. Reichel. London: Longmans Green, 1880.

MEDIEVAL ETHICS

General Histories of Medieval Philosophy

Copleston, F. C. *A History of Philosophy,* Vols. II and III. Westminster, MD: Newman Press, 1950–1953.

Gilson, Étienne. *Reason and Revelation in the Middle Ages.* New York: Scribners, 1938.

Gilson, Étienne. *The Spirit of Medieval Philosophy.* Translated by A. Downes. New York: Scribners, 1936.

Hawkins, D. J. B. *A Sketch of Mediaeval Philosophy.* London: Sheed and Ward, 1946.

Husik, I. *A History of Mediaeval Jewish Philosophy.* New York: Macmillan, 1916.

Augustine

Gilson, Étienne. *Introduction a l'étude de saint Augustin.* Paris: Vrin, 1943.

Switalski, B. *Neoplatonism and the Ethics of St. Augustine.* New York: Polish Institute of Arts and Sciences in America, 1946.

Erigena and Abelard

Bett, H. *Johannes Scotus Eriugena.* Cambridge, U.K., 1925.

Sikes, J. G. *Peter Abailard.* Cambridge, U.K.: Cambridge University Press, 1932. Chs. 2 and 8.

Thomas Aquinas

Copleston, F. C. *Aquinas.* London, 1955. Ch. 5.

D'Arcy, M. C. *St. Thomas Aquinas.* Dublin, 1953. Ch. 9.

Gilson, Étienne. *The Philosophy of St. Thomas Aquinas.* Translated by E. Bullough. St. Louis, 1934.

Maritain, Jacques. *St. Thomas and the Problem of Evil.* Milwaukee, WI: Marquette University Press, 1942.

Mullane, D. T. *Aristotelianism in St. Thomas.* Washington, DC: Catholic University of America, 1929.

SEVENTEENTH-CENTURY ETHICS

Hobbes

Laird, John. *Hobbes.* London: Benn, 1934.

Peters, R. S. *Hobbes.* Harmondsworth, U.K.: Penguin, 1956.

Stephen, Leslie. *Hobbes.* Ann Arbor: University of Michigan Press, 1961.

Taylor, A. E. *Hobbes.* New York, 1908.

Spinoza

Bidney, D. *The Psychology and Ethics of Spinoza.* New Haven, CT: Yale University Press, 1940.

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. Ch. 2.

Hampshire, S. *Spinoza.* Harmondsworth, U.K.: Penguin, 1951. Chs. 4 and 5.

Joachim, H. J. *A Study of the Ethics of Spinoza.* London: Clarendon Press, 1901.

McKeon, R. *The Philosophy of Spinoza.* New York: Longmans Green, 1928.

Locke

Aaron, R. I. *John Locke.* Oxford: Clarendon Press, 1955.

Lamprecht, S. P. *The Moral and Political Philosophy of John Locke.* New York: Columbia University Press, 1918.

Cambridge Platonists

Cassirer, Ernst. *The Platonic Renaissance in England.* Translated by J. P. Pettegrove. Austin: University of Texas Press, 1953.

Powicke, F. J. *The Cambridge Platonists.* Cambridge, MA, 1926.

EIGHTEENTH-CENTURY ETHICS

Moral-Sense Theories

Bonar, J. *Moral Sense.* New York: Macmillan, 1930.

Raphael, D. D. *The Moral Sense.* London: Oxford University Press, 1947.

Clarke

Le Rossignol, J. E. *The Ethical Philosophy of Samuel Clarke.* Leipzig, 1892.

Butler

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. Ch. 3.

Duncan-Jones, A. E. *Butler's Moral Philosophy.* Harmondsworth, U.K.: Penguin, 1952.

Shaftesbury and Hutcheson

Fowler, Thomas. *Shaftesbury and Hutcheson.* London, 1882.

Hume

Basson, A. H. *David Hume.* Harmondsworth, U.K., 1958. Ch. 5.

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. Ch. 4.

Smith, Norman Kemp. *The Philosophy of David Hume.* London: Macmillan, 1941.

Stewart, J. B. *The Moral and Political Philosophy of David Hume.* New York: Columbia University Press, 1963.

Price and Reid

Raphael, D. D. *The Moral Sense.* London: Oxford University Press, 1947.

French Enlightenment

Becker, C. L. *The Heavenly City of the Eighteenth Century Philosophers.* New Haven, CT: Yale University Press, 1932.

Cassirer, Ernst. *The Philosophy of the Enlightenment.* Translated by F. Koelln and J. P. Pettegrove. Princeton, NJ: Princeton University Press, 1951.

Crocker, L. G. *Nature and Culture.* Baltimore: Johns Hopkins Press, 1963.

Halévy, Élie. *The Growth of Philosophic Radicalism.* Translated by M. Morris. New York: Macmillan, 1928.

Rouston, M. *The Pioneers of the French Revolution.* Translated by F. Whyte. Boston: Little Brown, 1926. A study of the social and political involvements of the Encyclopedists.

Kant

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. Ch. 5.

Jones, W. T. *Morality and Freedom in the Philosophy of Immanuel Kant.* Oxford: Oxford University Press, 1940.

Körner, S. *Kant.* Harmondsworth, U.K.: Penguin, 1955. Chs. 6 and 7.

Lindsay, A. D. *Kant.* London, 1934.

Paton, H. J. *The Categorical Imperative.* New York: Hutchinson, 1947.

Paton, H. J. *The Moral Law, or Kant's Groundwork of the Metaphysic of Morals.* New York: Hutchinson, 1948.

Teale, A. E. *Kantian Ethics.* London: Oxford University Press, 1951.

NINETEENTH-CENTURY ETHICS

Utilitarianism

Halévy, Élie. *The Growth of Philosophic Radicalism.* Translated by M. Morris. New York: Macmillan, 1928.

Stephen, Leslie. *The English Utilitarians,* 3 vols. London: Duckworth, 1900.

German Idealism

Dewey, John. *German Philosophy and Politics.* New York: Putnam, 1942. Highly critical.

Hook, Sidney. *From Hegel to Marx.* New York: Reynal and Hitchcock, 1936.

Löwith, Karl. *From Hegel to Nietzsche.* Translated by D. Green. New York: Holt, Rinehart and Winston, 1964.

Marcuse, Herbert. *Reason and Revolution.* London: Oxford University Press, 1941.

Royce, Josiah. *The Spirit of Modern Philosophy.* New York, 1896.

Fichte

Adamson, Robert. *Fichte.* Edinburgh, 1881.

Royce, Josiah. *The Spirit of Modern Philosophy.* New York, 1896. Part I, Ch. 5.

Hegel

Findlay, J. N. *Hegel.* London, 1958. Chs. 5 and 11.

Marcuse, Herbert. *Reason and Revolution.* London: Oxford University Press, 1941.

Royce, Josiah. *The Spirit of Modern Philosophy.* New York, 1896. Part I, Ch. 8.

Stace, W. T. *The Philosophy of Hegel.* New York: Dover, 1955.

Schopenhauer

Copleston, F. C. *Arthur Schopenhauer, Philosopher of Pessimism.* London, 1947.

Mann, Thomas. *Schopenhauer.* Stockholm, 1938.

McGill, V. J. *Schopenhauer, Pessimist and Pagan.* New York: Brentano, 1931.

Nietzsche

Brinton, C. C. *Nietzsche.* Cambridge, 1941.

Copleston, F. C. *Friedrich Nietzsche.* London, 1942.

Lefebvre, Henri. *Nietzsche.* Paris, 1939.

Marx

Berlin, Isaiah. *Karl Marx.* London: Butterworth, 1939.

Hook, Sidney. *From Hegel to Marx.* New York: Reynal and Hitchcock, 1936.

MacIntyre, A. C. *Marxism, an Interpretation.* London: SCM Press, 1953.

Marcuse, Herbert. *Reason and Revolution.* London: Oxford University Press, 1941.

Tucker, R. *Philosophy and Myth in Karl Marx.* Cambridge, U.K.: Cambridge University Press, 1961.

American Transcendentalism

Pochmann, H. A. *New England Transcendentalism and St. Louis Hegelianism.* Philadelphia: Carl Schurz Memorial Foundation, 1948.

Schneider, H. W. *A History of American Philosophy.* New York: Columbia University Press, 1946.

Royce and James

Santayana, George. *Character and Opinion in the United States.* New York: Scribners, 1920.

Smith, J. E. *Royce's Social Infinite.* New York: Liberal Arts Press, 1950.

British Idealism

Pfannenstill, Bertil. *Bernard Bosanquet's Philosophy of the State.* Translated by Bert Hood. Lund, Sweden: Hakan Ohlsson, 1936.

Sidgwick, Henry. *The Ethics of Green, Spencer and Martineau.* London, 1902.

Wollheim, Richard. *F. H. Bradley.* Harmondsworth, U.K., 1959. Ch. 6.

Sidgwick

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930. Ch. 6.

TWENTIETH-CENTURY ETHICS

Assessments of Moore

Adams, E. M. *Ethical Naturalism and the Modern World-View.* Chapel Hill: University of North Carolina Press, 1960. Careful discussion of naturalism-nonnaturalism debate.

Broad, C. D. "G. E. Moore's Latest Published Views on Ethics." *Mind* (1961).

Field, G. C. "The Place of Definition in Ethics." *PAS* (1932). Key discussion of difficulties in Moore's conception of definition.

Frankena, W. K. "The Naturalistic Fallacy." *Mind* (1939). Classical source for discussion of naturalistic fallacy.

Hall, E. W. *Categorial Analysis.* Edited by E. M. Adams. Chapel Hill: University of North Carolina Press, 1964. Important defense of Mill against Moore.

Prior, A. N. *Logic and the Basis of Ethics.* Oxford: Clarendon Press, 1949. Crucial for history of the naturalistic fallacy.

Taylor, Paul. *Normative Discourse.* Englewood Cliffs, NJ: Prentice-Hall, 1961. Is/ought carefully discussed and a general theory of value developed along linguistic lines.

Warnock, Mary. *Ethics since 1900.* Oxford: Oxford University Press, 1960.

Wellman, Carl. *The Language of Ethics.* Cambridge, MA: Harvard University Press, 1961. Clear statement of difficulties in ethical naturalism.

Deontological Nonnaturalism

Broad, C. D. "Critical Notice of H. A. Prichard, *Moral Obligation.*" *Mind* (1950).

Broad, C. D. *Ethics and the History of Philosophy.* London: Routledge, 1952.

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930.

Broad, C. D. "Imperatives, Categorical and Hypothetical." *Philosopher* (1950).

Broad, C. D. "Review of Julian S. Huxley's Evolutionary Ethics." *Mind* (1944).

Broad, C. D. "Some of the Main Problems of Ethics." *Philosophy* (1946).

Broad, C. D. "Some Reflections on Moral-Sense Theories in Ethics." *PAS* (1944–1945).

Carritt, E. F. "Moral Positivism and Moral Aestheticism." *Philosophy* (1938). Less important than Prichard or Ross but still influential.

Carritt, E. F. *The Theory of Morals.* London: Oxford University Press, H. Milford, 1928.

Ewing, A. C. "The Autonomy of Ethics." In *Prospect for Metaphysics,* edited by I. T. Ramsey. New York: Philosophical Library, 1961.

Ewing, A. C. *The Definition of Good.* New York: Macmillan, 1947.

Ewing, A. C. *Ethics.* London: English Universities Press, 1953.

Ewing, A. C. *Second Thoughts in Moral Philosophy.* London: Routledge and Paul, 1959.

Hall, E. W. *Categorial Analysis.* Edited by E. M. Adams. Chapel Hill: University of North Carolina Press, 1964.

Hall, E. W. *Our Knowledge of Fact and Value.* Chapel Hill: University of North Carolina Press, 1961.

Hall, E. W. *What Is Value? An Essay in Philosophical Analysis.* London: Routledge and Paul, 1952.

Prichard, H. A. *Moral Obligation.* Oxford: Clarendon Press, 1949. Posthumously collected essays. Primary source for deontology.

Assessments of Deontological Nonnaturalism

Edwards, Paul. *The Logic of Moral Discourse.* Glencoe, IL: Free Press, 1955.

McCloskey, H. J. "Ross and the Concept of a *Prima Facie* Duty." *Australasian Journal of Philosophy* (1963).

Monro, D. H. "Critical Notice of *Second Thoughts in Moral Philosophy.*" *Australasian Journal of Philosophy* (1960).

Schilpp, P. A., ed. *The Philosophy of C. D. Broad.* New York: Tudor, 1959. Note essays by Frankena, Hare, and Hedenius and Broad's reply. Broad's work, though largely critical of other philosophers, represents a key statement of deontology.

Strawson, P. F. "Ethical Intuitionism." *Philosophy* (1949).

Assessments of Phenomenological Nonnaturalism

Hook, Sidney. "A Critique of Ethical Realism." *International Journal of Ethics* (1929).

Jensen, O. C. "Nicolai Hartmann's Theory of Virtue." *Ethics* (1942).

Schlick, Moritz. "Is There a Factual *A Priori?,*" In *Readings in Philosophical Analysis,* edited by Herbert Feigl and Wilfrid Sellars. New York: Appleton-Century-Crofts, 1949.

Walker, Merle. "Perry and Hartmann, Antithetical or Complementary?" *International Journal of Ethics* (1939).

Ethical Naturalism

Dewey, John. *Essays in Experimental Logic.* Chicago: University of Chicago Press, 1916.

Dewey, John. "Ethical Subject Matter and Language." *Journal of Philosophy* (1945).

Dewey, John. *Human Nature and Conduct.* New York: Holt, 1922.

Dewey, John. *The Problems of Men.* New York: Philosophical Library, 1946.

Dewey, John. *The Quest for Certainty.* New York: Minton Balch, 1929.
Dewey, John. *Theory of Valuation.* Chicago: University of Chicago Press, 1939.
Findlay, J. N. *Language, Mind and Value.* London: Allen and Unwin, 1963.
Findlay, J. N. *Values and Intentions.* London: Allen and Unwin, 1961.
Foot, P. R. "Goodness and Choice." *PAS,* Supp. Vol. (1961).
Foot, P. R. "Moral Arguments." *Mind* (1958).
Foot, P. R. "Moral Beliefs." *PAS* (1958).
Pepper, Stephen. *A Digest of Purposive Values.* Berkeley: University of California Press, 1947.
Prall, D. W. *A Study in the Theory of Value.* Berkeley, CA, 1921.
Reid, J. R. "The Nature and Status of Values." In *Philosophy for the Future,* edited by R. W. Sellars, V. G. McGill, and Marvin Farber. New York: Macmillan, 1949.
Reid, J. R. *A Theory of Value.* New York: Scribners, 1938.
Rice, P. B. *On the Knowledge of Good and Evil.* New York: Random House, 1955.
Stace, W. T. *The Concept of Morals.* New York: Macmillan, 1937.
Westermarck, Edward. *Ethical Relativity.* London: Kegan Paul, 1932.

Assessments of Ethical Naturalism

Ezorsky, Gertrude. "Inquiry as Appraisal: The Singularity of John Dewey's Theory of Valuation." *Journal of Philosophy* (1958).
Frankena, William. "C. I. Lewis on the Ground and Nature of the Right." *Journal of Philosophy* (1964).
Frankena, William. "Ethical Naturalism Renovated." *Review of Metaphysics* (1957).
Frankena, William. "Lewis' Imperatives of Right." *Philosophical Studies* (1963).
Frankena, William. "Obligation and Motivation in Recent Moral Philosophy." In *Essays in Moral Philosophy,* edited by A. I. Melden. Seattle: University of Washington Press, 1958.
Frankena, William. "Three Comments on Lewis's Views on the Right and the Good." *Journal of Philosophy* (1964).
Hare, R. M. *The Language of Morals.* Oxford: Clarendon Press, 1952.
Kennedy, Gail. "Science and the Transformation of Common Sense: The Basic Problem of Dewey's Philosophy." *Journal of Philosophy* (1954).
Wellman, Carl. *The Language of Ethics.* Cambridge, MA: Harvard University Press, 1961.
White, Morton. *Social Thought in America.* New York: Viking Press, 1949.
White, Morton. "Value and Obligation in Dewey and Lewis." *Philosophical Review* (1949).
(See also the works by Moore, Ross, Ewing, and Hall previously cited.)

Contemporary Noncognitivism

Ayer, A. J. *Language, Truth and Logic.* London: Gollancz, 1935.
Ayer, A. J. "On the Analysis of Moral Judgments." In his *Philosophical Essays.* New York: St. Martin's Press, 1954.
Carnap, Rudolf. *Philosophy and Logical Syntax.* London: Kegan Paul, 1935.
Cassirer, Ernst. "Axel Hägerström: Eine Studie zur schwedischen Philosophie der Gegenwart." *Goteborgs Hogskolas Arsskrift* (1939).
Edwards, Paul. *The Logic of Moral Discourse.* Glencoe, IL: Free Press, 1955.
Hedenius, Ingemar. "Etikens Subjecktivitet." In *Tro och Moral.* Stockholm, 1955.
Hedenius, Ingemar. *Om Rätt och Moral* (On Law and Morals). Stockholm, 1941.
Hedenius, Ingemar. "Values and Duties." *Theoria* (1949).
Reichenbach, Hans. *The Rise of Scientific Philosophy.* Berkeley: University of California Press, 1951.
Richards, I. A. *The Meaning of Meaning.* New York, 1933.
Richards, I. A. *Speculative Instruments.* Chicago: University of Chicago Press, 1957.
Robinson, Richard. *An Atheist's Values.* Oxford: Clarendon Press, 1964.
Robinson, Richard. "The Emotive Theory of Ethics." *PAS,* Supp. Vol. (1948).
Ross, Alf. *Kritik der sogenannten praktischen Erkenntniss.* Copenhagen, 1933.
Ross, Alf. *On Law and Justice.* Berkeley: University of California Press, 1959.
Ross, Alf. "On Moral Reasoning." *Danish Yearbook of Philosophy* (1964).
Ross, Alf. "On the Logical Nature of Propositions of Value." *Theoria* (1945).
Russell, Bertrand. *Religion and Science.* London: Butterworth-Nelson, 1935.
Schilpp, P. A., ed. *The Philosophy of Rudolf Carnap.* La Salle, IL: Open Court, 1964.
Stevenson, Charles. *Ethics and Language.* New Haven, CT: Yale University Press, 1944.
Stevenson, Charles. *Facts and Values.* New Haven, CT: Yale University Press, 1963.
Stevenson, Charles. "The Scientist's Role and the Aims of Education." *Harvard Educational Review* (1954).

Assessments of the Emotive Theory

Bedford, E. "The Emotive Theory of Ethics." In *Proceedings of the Eleventh International Congress of Philosophy,* Vol. X. Amsterdam, 1953.
Brandt, Richard. "The Emotive Theory of Ethics." *Philosophical Review* (1950).
Falk, W. D. "Goading and Guiding." *Mind* (1954).
Foot, P. R. "The Philosopher's Defense of Morality." *Philosophy* (1952).
Nielsen, Kai. "Bertrand Russell's New Ethic." *Methodos* (1958).
Nielsen, Kai. "On Looking Back at the Emotive Theory." *Methodos* (1962).
Tegen, Einar. "The Basic Problem in the Theory of Value." *Theoria* (1944).

Assessments of Existentialism

Ayer, A. J. "Albert Camus, Novelist-Philosopher." *Horizon* (1946).
Ayer, A. J. "Novelist-Philosophers: Jean-Paul Sartre." *Horizon* (1946).
Ayer, A. J. "Philosophy at Absolute Zero." *Encounter* (1955).

Beauvoir, Simone de. *The Ethics of Ambiguity.* Translated by Bernard Frechtman. New York: Philosophical Library, 1948. Important in its own right and in relation to Sartre.

Cranston, Maurice. *Sartre.* Edinburgh and London, 1962. General discussion of the range of Sartre's work. Elementary but clear.

Cruickshank, John. *Camus and the Literature of Revolt.* New York, 1959. Contains a detailed and sympathetic account of Camus's thought.

Cruickshank, John, ed. *The Novelist as Philosopher.* London: Oxford University Press, 1962.

Hochberg, Herbert. "Albert Camus and the Ethic of Absurdity." *Ethics* (1965). Important analytical criticism.

Jeanson, Francis. *Le problème moral et la pensée de Sartre.* Paris, 1947. Sympathetic and informal discussion of Sartre's moral theory.

Murdoch, Iris. *Sartre: Romantic Rationalist.* New Haven, CT: Yale University Press, 1955. Relates Sartre to English thought.

Warnock, Mary. *Ethics since 1900.* Oxford: Oxford University Press, 1960.

Warnock, Mary. *The Philosophy of Sartre.* London: Hutchinson, 1966. Challenging but sympathetic assessment of Sartre.

Wollheim, Richard. "Modern Philosophy and Unreason." *Political Quarterly* (1955).

Wollheim, Richard. "The Political Philosophy of Existentialism." *Cambridge Journal* (1953).

Linguistic Noncognitivism

Hare, R. M. "Descriptivism." *Proceedings of the British Academy* (1963).

Hare, R. M. "Universalizability." *PAS* (1954).

Hartland-Swann, John. *An Analysis of Morals.* London: Allen and Unwin, 1960.

Mayo, Bernard. *Ethics and the Moral Life.* London, 1958.

Montefiore, Alan. "Goodness and Choice." *PAS,* Supp. Vol. (1961).

Montefiore, Alan. *A Modern Introduction to Moral Philosophy.* New York: Praeger, 1959.

Nowell-Smith, P. H. "Contextual Implications and Ethical Theory." *PAS,* Supp. Vol. (1962).

Nowell-Smith, P. H. *Ethics.* London: Penguin, 1954.

Nowell-Smith, P. H. "Morality: Religious and Secular." *Rationalist Annual* (1961).

Nowell-Smith, P. H. "Review of *Freedom and Reason.*" *Ratio* (1964).

Assessments of Linguistic Noncognitivism

Binkley, Luther. *Contemporary Ethical Theories.* New York: Philosophical Library, 1961.

Castañeda, Héctor-Neri. "Imperatives, Decisions and 'Oughts.' A Logico-Metaphysical Investigation." In *Morality and the Language of Conduct,* edited by Héctor-Neri Castañeda and George Nakhnikian. Detroit, MI: Wayne State University Press, 1963.

McCloskey, H. J. "Hare's Ethical Subjectivism." *Australasian Journal of Philosophy* (1959).

McCloskey, H. J. "Nowell-Smith's Ethics." *Australasian Journal of Philosophy* (1961).

Monro, H. D. "Are Moral Problems Genuine?" *Mind* (1956).

Monro, H. D. "Critical Notice of *Freedom and Reason.*" *Australasian Journal of Philosophy* (1964).

Monro, H. D. "Impartiality and Consistency." *Philosophy* (1961).

Sparshott, F. E. "Critical Study of *Freedom and Reason.*" *Philosophical Quarterly* (1963).

Good-Reasons Approach

Aiken, H. D. *Reason and Conduct.* New York: Knopf, 1962.

Baier, Kurt. "Decisions and Descriptions." *Mind* (1951).

Baier, Kurt. *The Meaning of Life.* Canberra, 1957.

Falk, W. D. "Action-Guiding Reasons." *Journal of Philosophy* (1963).

Falk, W. D. "Goading and Guiding." *Mind* (1953).

Falk, W. D. "Morality and Nature." *Australasian Journal of Philosophy* (1950).

Falk, W. D. "Morality, Self and Others." In *Morality and the Language of Conduct,* edited by Héctor-Neri Castañeda and George Nakhnikian. Detroit, MI: Wayne State University Press, 1963.

Falk, W. D. "Moral Perplexity." *Ethics* (1956).

Falk, W. D. "Morals without Faith." *Philosophy* (1944).

Falk, W. D. "Obligation and Rightness." *Philosophy* (1945).

Falk, W. D. "'Ought' and Motivation." *PAS* (1947–1948).

Gauthier, David. *Practical Reasoning.* Oxford: Clarendon Press, 1963.

Ladd, John. "The Issue of Relativism." *Monist* (1963).

Ladd, John. "Reason and Practice." In *The Return to Reason,* edited by John Wild. Chicago: Regnery, 1953.

Ladd, John. *The Structure of a Moral Code.* Cambridge, MA: Harvard University Press, 1957.

Melden, A. I. *Rights and Right Conduct.* Oxford, 1959.

Melden, A. I. "Two Comments on Utilitarianism." *Philosophical Review* (1951).

Murphy, A. E. "Blanshard on Good in General." *Philosophical Review* (1963).

Murphy, A. E. *The Theory of Practical Reason.* Edited by A. I. Melden. La Salle, IL: Open Court, 1965.

Murphy, A. E. *The Uses of Reason.* New York: Macmillan, 1943.

Murphy, A. E., William Hay, and Marcus Singer, eds. *Reason and the Common Good.* Englewood Cliffs, NJ: Prentice-Hall, 1963.

Nielsen, Kai. "Appraising Doing the Thing Done." *Journal of Philosophy* (1960).

Nielsen, Kai. "Can a Way of Life Be Justified?" *Indian Journal of Philosophy* (1960).

Nielsen, Kai. "Conventionalism in Morals and the Appeal to Human Nature." *Philosophy and Phenomenological Research* (1962).

Nielsen, Kai. "The Functions of Moral Discourse." *Philosophical Quarterly* (1957).

Nielsen, Kai. "The 'Good Reasons Approach' and 'Ontological Justifications of Morality.'" *Philosophical Quarterly* (1959).

Nielsen, Kai. "The Good Reasons Approach Revisited." *Archiv für Rechts- und Sozialphilosophie* (1965).

Nielsen, Kai. "Is 'Why Should I Be Moral?' an Absurdity?" *Australasian Journal of Philosophy* (1958).

Nielsen, Kai. "Justification and Moral Reasoning." *Methodos* (1957).

Nielsen, Kai. "Why Should I Be Moral?" *Methodos* (1963).

Rawls, John. "Constitutional Liberty and the Concept of Justice." In *Nomos*, Vol. VI: *Justice*, edited by Carl J. Friedrich and John W. Chapman. New York: Atherton, 1963.

Rawls, John. "Justice as Fairness." *Philosophical Review* (1958).

Rawls, John. "Legal Obligation and the Duty of Fair Play." In *Law and Philosophy*, edited by Sidney Hook. New York: New York University Press, 1964.

Rawls, John. "Outline of a Decision Procedure for Ethics." *Philosophical Review* (1951).

Rawls, John. "The Sense of Justice." *Philosophical Review* (1963).

Rawls, John. "Two Concepts of Rules." *Philosophical Review* (1955).

Singer, Marcus. *Generalization in Ethics.* New York: Knopf, 1961.

Singer, Marcus. "The Golden Rule." *Philosophy* (1963).

Singer, Marcus. "Negative and Positive Duties." *Philosophical Quarterly* (1965).

Taylor, Paul. "The Ethnocentric Fallacy." *Monist* (1963).

Taylor, Paul. "Four Types of Ethical Relativism." *Philosophical Review* (1954).

Taylor, Paul. *Normative Discourse.* Englewood Cliffs, NJ: Prentice-Hall, 1961.

Taylor, Paul. "On Justifying a Way of Life." *Indian Journal of Philosophy* (1961).

Toulmin, Stephen. *An Examination of the Place of Reason in Ethics.* Cambridge, U.K.: Cambridge University Press, 1950.

Toulmin, Stephen. "Is There a Fundamental Problem of Ethics?" *Australasian Journal of Philosophy* (1955).

Toulmin, Stephen. "Knowledge of Right and Wrong." *PAS* (1949–1950).

Toulmin, Stephen. "Principles of Morality." *Philosophy* (1956).

Wright, Georg von. *The Logic of Preference.* Edinburgh: Edinburgh University Press, 1963.

Wright, Georg von. *Norm and Action.* London: Routledge and K. Paul, 1963.

Wright, Georg von. "On Promises." *Theoria* (1962).

Wright, Georg von. "Practical Inference." *Philosophical Review* (1963).

Wright, Georg von. *The Varieties of Goodness.* London, 1963.

Raziel Abelson (1967)
(Ethics through the nineteenth century)
Kai Nielsen (1967)
(Twentieth-century ethics)

ETHICS, HISTORY OF: OTHER DEVELOPMENTS IN TWENTIETH-CENTURY ETHICS

Even setting aside the rich interplay between naturalism, nonnaturalism, and noncognitivism that is one of the hallmarks of twentieth-century moral philosophy, a very rich history remains, one that is impossible to even fully summarize here. Much of the story may be found elsewhere in the present volume in discussions devoted to particular moral theories and philosophers. The present entry examines just a few of the many themes that have occupied the attention of moral philosophers working within a diversity of traditions and that have thus exercised substantial influence on the shape of moral philosophy in the twentieth century.

MORAL PRINCIPLES

Whether the number of principles governing right conduct is one, or several, or even indefinitely many is a question that has animated the development of moral philosophy over the past century. Of course, this is by no means a new problem for moral philosophy, and the responses found to it in the twentieth century are themselves shaped by earlier debates between the nineteenth-century utilitarians and their intuitionist and idealist opponents.

Indeed, the *ideal utilitarianism* developed in George Edward Moore's *Principia Ethica* (1903) can usefully be seen as a possible rapprochement between utilitarianism and more pluralistic views. Moore maintained that there was but one ultimate principle of duty that one should act so as to promote as much good (intrinsic value) as possible, and in *Principia*, he maintained this principle to be analytic. It is puzzling that Moore did not take his own Open Question Argument to tell against this identity claim, but setting that aside, Moore appears to be an arch-monist about ultimate principles of right action.

If we take note of Moore's innovative and influential value theory, however, it becomes clear that Moore is in a position (assuming his view is otherwise sustainable) to accommodate many of the insights of pluralists. Two aspects of Moore's value theory are critical. First, Moore held that goodness (or intrinsic value) is not identical to any natural property. Consequently, the bearers of intrinsic value may form an ultimately heterogenous group, having nothing salient in common other than their goodness. Indeed, Moore is a pluralist about the bearers of intrinsic value. Second, Moore argued that the value of a whole need not be the same as the value of the sum of its parts; such wholes are *organic unities*.

This pluralism about the bearers of intrinsic value and the flexibility that the doctrine of organic unities affords when it comes to the value of a whole, yield a view that seems well poised to accommodate the concerns of pluralists about ultimate principles. For pluralists have long emphasized that there are seemingly many potential grounds of duty and have challenged utilitarians to show

that their view could leave in tact the seeming legitimacy of many moral rules that are not directly concerned with promoting utility. For the consequentialist who identifies goodness with a specific natural property and denies the doctrine of organic unities, these are very difficult challenges to meet. For one must then show that the apparently diverse grounds of duty really all involve the (naturalistically construed) property of goodness or else explain away the appearances. And one must show that apparently legitimate moral rules (e.g., rules governing the keeping of promises) really do serve to promote goodness. To a consequentialist of Moore's stripe, however, it is always open to identify further bearers of intrinsic value to accommodate our intuitions about the diverse grounds of duty and to appeal to the doctrine of organic unities to maintain the legitimacy of accepted moral rules even when these seem to lead to a universe with very disvaluable *parts*. Indeed, it may seem that provided a sufficiently flexible theory of value, the determined consequentialist will be able to say just about anything when it comes to duty.

Shortly, Moore's consequentialism was subjected to influential critique by William David Ross. Where earlier pluralists had identified many principles of (all things considered) duty, in *The Right and the Good* (1930), Ross sought principles of *prima facie duty*. For Ross, prima facie duties were acts of a type that tend to be our duty all things considered. Moreover, if someone had a prima facie duty to wash his neighbor's car (say because he had promised to do so), then this would be an all-things-considered duty if it did not pose any conflict with other prima facie duties. Ross argued forcefully that our apprehension of certain kinds of acts as prima facie duties does not depend upon our apprehension of them as being productive of more rather than less good.

In most, or perhaps all, cases, however, an agent will have conflicting prima facie duties. How is one to determine what duty requires, all things considered? To this, Ross answered that there are no principles (or at least no principles we have any prospect of identifying and using) that would determine the answer to this question, and that the best one can do is to exercise good judgment regarding which prima facie duty is, in the circumstances, most weighty. In taking this position, Ross appears to avoid an influential argument for monism about ultimate principles to which John Stuart Mill had given powerful voice in his *System of Logic* (1843/1874). Mill had argued that there could only be one possible ultimate moral principle because any set of several principles were liable to conflict about a given case, and there would need to be a higher principle to be an *umpire* between them. By construing principles as principles of prima facie duty and by denying that there are any principles determining final duty, Ross seems to sidestep Mill's argument.

At the same time, Ross cast serious doubt on whether the systematic advantages often credited to monism about principles (especially by the nineteenth-century hedonistic utilitarians) could really survive the death of naturalistic accounts of the good, a death that Moore's arguments were at the time widely held to have confirmed. Ross recognized that some may be dissatisfied that a system of prima facie duties leaves no clear method for determining final duty. Ross plausibly replied that his view was no worse off in this regard than was the *ideal utilitarianism* of Moore. While Ross's view provides no discernible method for determining which of several prima facie duties is most weighty, Moore's view provides no discernible method for determining which of the various goods (as well as combinations thereof) that we might produce through our action have the greatest intrinsic value. One consequence of this debate between Moore and Ross was that the debate between monists and pluralists was revealed to depend critically on views of value and moral conflict.

Despite Ross's influential case for pluralism, the two dominant normative theories of the twentieth century, utilitarianism and Kantian deontology, both hold that there is only one ultimate principle of duty though, of course, they disagree about what this principle is. A perennial challenge for such views is to explain and or to justify the seeming legitimacy of a diverse set of common moral rules. It is worth looking briefly, then, at some of the resources developed by principle monists to meet this challenge.

For the principle monist, it is critical to define some relationship between whatever principle is taken to be ultimate and more particular midlevel moral rules, such as rules against promise-breaking or against dishonesty. One possibility, discussed influentially by John Rawls in "Two Concepts of Rules" (1955) is that some rules might be constitutive of a practice while other rules serve to justify that rule-constituted practice. Thus, Rawls imagines that a utilitarian might use a consequentialist principle to justify a practice of punishment that is itself constituted by backward-looking retributive rules. For example, a practice of punishment might be constituted in part by a rule that only those who have committed a crime are to be punished (no matter how much good punishing an innocent might do) even while the practice is justified by the good it brings about. Despite what the example

of punishment might seem to suggest, this distinction does not entail that the only apt justifications of rule-constituted practices are consequentialist ones. Indeed, a Kantian could claim that the justification for a practice is that it best expresses respect for the dignity of each person even while that practice is constituted by rules that do not directly concern the dignity of persons. Rawls was quick to emphasize that his distinction was not a new one. Nevertheless, his clear and forceful discussion of it led moral philosophers to more steadfastly distinguish different levels of justification.

The latter half of the twentieth century saw the careful development of a variety of views about the relationship between ultimate standards and the more particular and diverse moral rules familiar from everyday life. Many of these developments were advanced by consequentialists. One influential view was put forward by Richard M. Hare in *Moral Thinking* (1981). According to Hare, there are two distinct levels of moral thinking. One, which Hare called *intuitive moral thinking*, involves the deployment of familiar and relatively simple moral principles in deciding how to act.

Hare claimed that intuitive moral thinking is characterized by a plurality of such principles and that such principles can conflict with one another by recommending different and incompatible courses of action. When this occurs, Hare argued, we can ascend to *critical moral thinking*. Doing so requires deploying a superior principle (for Hare, a version of the principle of utility). This principle is capable of both adjudicating the conflict between rival principles at the intuitive level and (in conjunction with facts about human psychology) justifying our everyday use of intuitive moral thinking. For example, instead of trying to determine which of two conflicting intuitive principles is more weighty (as Ross advocated), we can ask directly what course of action would best satisfy the utilitarian principle. Nevertheless, our everyday use of intuitive principles is justified because (for agents like us) directly applying the principle of utility to all of our decisions would be cumbersome, costly, and error prone.

Importantly, Hare's *two-level* view of the relationship between the principle of utility and more particular moral principles differs from classic versions of rule utilitarianism as well as from other forms of *indirect utilitarianism*. On the rule utilitarian view, the rightness of an act is defined in terms of its conformity to the best (i.e., best at promoting good consequences) rules whereas for Hare, it is possible both for an act to be in conformity with intuitive principles and yet wrong as well as for an act to be violative of intuitive rules and yet right. Rightness is determined by the principle governing critical moral thinking.

The resources of Hare's view can be deployed not only in considering familiar moral rules, but also in considering qualities of character. For just as the principle of utility might recommend the adoption of a range of more or less simple moral rules, so, too, it might recommend the cultivation of useful character traits. Again, though, one must be careful to distinguish what principle is actually the standard of rightness or duty. For some philosophers, such as Peter Railton, the act-consequentialist principle remains the ultimate moral standard even while it recommends that we develop the kinds of character traits that will sometimes lead us to act contrary to it. Others, however, suggest that the proper way for the utilitarian to evaluate acts is by reference to the motive that leads an agent to act with motives being evaluated by reference to their consequences.

Not surprisingly, the increased attention to indirect and two-level versions of principle monism also spawned more careful criticisms. In the case of indirect theories such as rule utilitarianism, one important worry—nicely and influentially discussed by David Lyons (1935–) in *Forms and Limits of Utilitarianism* (1965)—has been that such views inevitably *collapse* into their more direct progenitors. In the case of rule utilitarianism, for example, the collapse supposedly occurs because the really best rules would recommend the very same choices and actions as would a direct application of the principle of utility. In the case of two-level theories, an important set of worries has concerned the stability of the view. Some philosophers worry whether human beings really can smoothly ride the escalator between intuitive and critical thinking more or less as circumstances warrant.

In other cases the worry is more conceptual. For if (as Hare's view seems to require) it is sometimes advisable to set aside the ultimately correct standard, then it seems possible that there could be circumstances in which it is advisable to permanently set aside (or even to banish from thought entirely) the putatively correct standard. Whether a proper standard of conduct could be *self-effacing* (to borrow a term from Derek Parfit's influential discussion in *Reasons and Persons* [1984]), has been a matter of intense dispute, especially with regard to moral standards. Some, following philosophers such as Henry Sidgwick, welcome the possibility; others see in it a deep confusion. In some cases the thought has been that moral standards are essentially deliberative tools and so must have some place in the psychology of moral agents. In other cases the thought has been that moral standards

essentially play a role in interpersonal relations and so must remain *public*, where they can be appealed to, criticized, and defended.

Despite the development of new tools to defend principle monism, the last quarter of the twentieth century saw a notable (and ongoing) revival of pluralism, sometimes of a new and more radical variety. Some of this interest centered on the ways in which various values might conflict. Where Ross was cautiously skeptical of identifying principles that would systematically rank prima facie duties, some philosophers, such as Thomas Nagel (1970), began to argue that various values might be fundamentally incommensurable, perhaps because they arise or make sense only within certain standpoints or perspectives. Clearly incommensurable values would seriously complicate the prospects for reasoned and justified choice when such values conflict, and there is no general agreement on whether it would remain possible at all.

But the revival of interest in pluralism has also been inspired by careful reflection on the sheer variety of considerations that can acquire moral significance. Where Ross had countenanced a limited set of principles governing prima facie (but not final) duty, a new breed of pluralists (sometimes called particularists) seek to challenge whether there are any moral principles at all—or at least whether a proper understanding of morality must make reference to them. On such views our moral understanding is not best represented as the application of universal rules but, rather, should be seen as kind of direct appreciation of the moral relevance of the particular features before us. The advantages and liabilities of this position are still being explored and debated.

MORAL AND PERSONAL VALUE

According to a dominant view of moral value, what is of moral value should be an object of care and concern for any rational agent. This view goes back at least to Plato, and according to it, moral value is of universal appeal; one has reason to care about it no matter who one is. Though widely held this view raises a number of important questions, in part because there seem to be many values that do not necessarily have a claim on any rational person. We might call such values *personal values* to denote that whether (or perhaps how) one ought to care about them depends upon what sort of person one is. Such values are also commonly referred to as *nonmoral values*. Among the most commonly offered examples are the value we find in personal relationships, such as the value of one friend to another, the value of achieving a personal goal or project such as the goal of amassing the world's most comprehensive collection of glass paperweights, and the value of living up to a personal ideal such as being a good marine or a good writer.

It is now widely agreed that such personal values need not be self-interested in any intuitive sense since a person's projects and concerns might be directed outward to others and to the world. Intuitively, the person whose project is to save land for a bird sanctuary is working to benefit the birds or the environment and not, at least in the first instance, herself. It is often thought that personal values depend upon the particular relationship in which a person stands to the object of value (as the value a parent finds in their child may depend on it being *their* child) or upon the particular preferences or choices the person has made (amassing those paperweights may be of importance only because the agent has come, perhaps by choice, to care about doing so).

If we accept some distinction between moral and personal value, then a number of issues arise. Since distinguishing moral and personal value raises the possibility of their coming into conflict, one must wonder whether they really do so. One view, associated with Immanuel Kant, is that personal values cease to be of value if ever their pursuit runs afoul of moral value. Consistent with this, though, one might hold that moral values themselves must make room for the pursuit of nonmoral ends and that our understanding of what moral value requires of us should be shaped by our intuitions about the reasonable pursuit of other values. Then there is the possibility that there is a genuine conflict between moral and personal values, and this raises questions about what the appropriate response to such a conflict would be. Before turning to these issues in more depth, however, it is worth pausing to ask whether they might be sidestepped.

Given the immense energy devoted to understanding the relationship between moral and personal value, it is notable that the twentieth century opened with a classic rejection of the problem. Moore argued that the tendency to distinguish moral and personal value rested on confusion and that in fact there is no such thing as personal value. Though couched as an argument against egoism, Moore's reasoning, if sound, would undermine the possibility of something being valuable to me but not to others. For Moore held that any putative claim of the form *X* is good for *A* must be resolved into a claim of absolute value (either into the claim that *X* is absolutely good and so in having *X*, *A* would have something absolutely good or else into the claim that it is absolutely good that *A* have *X*) or else into a psychological claim, such as the claim

that *A* desires *X*, which Moore argues is not really a value claim at all. To be sure, Moore did not deny that many of the supposed examples of personal value are in fact valuable (if anything, he emphasized their value), but he argued that properly understood, such value must be *absolute* and so equally of value from any point of view. Few now accept Moore's argument or its conclusion. For a representative criticism see John Leslie Mackie's "Sidgwick's Pessimism" (1976); for a more recent defense of an argument that is similar in spirit to Moore's, see Brian Medlin's (1927–2004) "Ultimate Principles and Ethical Egoism" (1957).

Setting Moore's position aside, there remain a host of questions about the relationship between moral and personal value. During the heyday of Ross and his fellow intuitionists, these questions received comparatively less attention. But as Rossian pluralism receded from center stage and utilitarianism and Kantian ethics (and later virtue ethics) began to be perceived as the main alternatives, the relationship between moral and personal value became a central topic of debate for moral philosophers.

According to one influential critique initiated by Bernard Williams, some moral theories fail precisely because they do not give a proper place to personal values. In his "Critique of Utilitarianism" (1973), Williams charged that utilitarianism requires agents, in so far as they take up a moral point of view, to regard their own projects and values as of no greater importance than the values and projects of others. Indeed, the value of anyone achieving their personal aims depends not at all on whose aims they are but on what contribution they make to the overall amount of happiness in the world. However, having personal values in the first place seems to presuppose attaching some significant importance to them, and such values are part of what make us who we are.

The problem, Williams charged, is not that utilitarianism implies implausible moral verdicts but that it leaves us thinking about matters in the wrong way. And if utilitarianism asks us to treat our own values and projects as having no greater importance than anyone else's, then, as Williams put it, it amounts to an assault on our psychological health, on our *integrity*. Though officially directed at utilitarianism, Williams's argument was more broadly influential. It inspired further close attention to the character and variety of personal values and to the ways in which they are an essential element of any familiar picture of human life and agency. Some philosophers were also quick to attempt to extend Williams's point to moral theories other than utilitarianism and to take his critique to undermine any normative theory that embodied a very strict impartiality.

As moral philosophers gave greater attention to personal values in the 1970s and 1980s, there was at the same time a great increase in the moral discussion of concrete moral problems. None received greater attention than the grave problem of hunger and poverty. In light of these developments, a new interest in the *demandingness* of morality arose. It is important to distinguish the issue of demandingness from the question whether moral requirements represent categorical or hypothetical imperatives. While the latter is a matter of whether moral requirements represent actions as good (or bad) in virtue of the ends that may be brought about by acting on them, the former is a matter of the imposition that complying with morality makes upon the other interests and concerns of the agent. A moral requirement might be categorical (requiring an agent to perform certain actions *whether they serve certain purposes or not*) and yet still not be very demanding if complying with it would not frequently or profoundly interfere with an agent pursuing other aims. Conversely, an imperative might be hypothetical and still be very demanding. While the question of whether moral requirements represent categorical or hypothetical imperatives is a matter, in some sense, of the logic of moral requirements, the demandingness of moral requirements depends crucially upon both the content of moral requirements and also upon contingent facts about the aims of the agents to whom moral requirements apply, and also the state of the world.

Demandingness has come to be such an important issue, in part because, given the state of the world, it seems likely that virtually any plausible moral theory threatens to be extremely demanding. While this may be obvious in the case of utilitarianism, it may be no less true of any moral theory that recognizes a duty to aid others in need, as virtually any plausible view must do. As philosophers such as Peter Singer in his famous paper "Famine, Affluence, and Morality" (1972) have emphasized, most people in developed countries, even those of modest means, are in a position to take actions that would save the lives of many other people simply by sacrificing what in the light of comparison seem like trivial personal goods. And yet there is no obvious stopping point to this argument, and if there is not, then morality might require us to sacrifice nearly everything—at least until circumstances change and our aid is no longer needed (which is unlikely to say the least!).

Philosophers have explored various responses to the issue of demandingness. Three will be noted here. First,

for some, the problem is not with any account of morality that leaves morality demanding. Morality simply is demanding—or at least can be in a world such as ours—and if people are often unwilling to do what morality requires, this is because they do not care as much as they should about morality. As an alternative to such rigorism, many philosophers have argued that theories of morality that leave morality highly demanding are dubious for that reason. The task is then to find a plausible revision. In the case of our duties to the needy, some have suggested that those duties are less stringent when those in need are far away and unfamiliar. Others suggest that morality requires us only to do our fair share of the helping, even if this leaves many of those in desperate need without help at all. Still others argue that our duty to aid, especially when applied to those in faraway lands, is a collective not an individual one.

A third approach is to argue that morality may be very demanding but that morality should not be the sole or even dominating concern of a good person. On this view our ideal of a person is of one whose concern with morality is itself tempered by other (possibly conflicting) concerns. This last possibility shows further that the issue of demandingness must be distinguished from the question of whether moral demands are *overriding*. Just as a job might be very demanding even while only a fool would allow its demands always to override other concerns, so, too, some suspect that morality might be highly demanding but that one can also give it too high a place in one's life.

Though the view that moral value has a claim on all rational agents has been a dominant one, it has also been subjected to interesting and influential critique. In the first half of the twentieth century, much effort was made to understand what attitude is involved in the judgment that something is of moral value. Under the influence of Moore's Open Question Argument and along with the rise of the early noncognitivists, it was often thought that virtually anything could be coherently judged to have moral value. After all, if (as Moore held) moral value is not identical to any natural property, or (as the early noncognitivists maintained) to hold something to have moral value is just to take a special favorable attitude towards it, then there seems to be no conceptual bar to holding virtually anything to have moral value. In the latter half of the twentieth century, however, following the lead of philosophers such as Philippa Foot and Gertrude Elizabeth Margaret Anscombe, many moral theorists began to think that there was a necessary connection between judgments of moral value and particular human concerns. To claim, for example, that painting two of one's toenails green is morally good is, philosophers such as Foot argued, simply unintelligible (absent some special background story) since doing so has no discernible connection to human flourishing or well being.

Once philosophers began to look to the possibility that there might be a necessary connection between moral evaluation and specific human concerns and interests, they also began to look carefully at evaluative concepts that seem clearly to implicate such a connection, such as specific virtue concepts like courage and modesty. Where the first half of the twentieth century was dominated by discussion of ethical concepts such as *good* and *right* (sometimes called *thin* ethical concepts), the latter half of the century saw steadily increasing interest in *thick* ethical concepts, such as virtue concepts, that single out specific forms of action and practical orientation as worthy of esteem.

To think that moral evaluation must somehow be directed at specific human concerns and interests is clearly consistent with the possibility that the relevant concerns are (or at least can be) universally the concerns of any rational being. Thus, one might suppose that in seeing courage as a moral value one must suppose that it is worthy of the esteem of any rational agent. As the century progressed, however, philosophers such as Alasdair MacIntyre (1981) and Williams began to argue that the very way in which we make moral or ethical evaluations is shaped by particular institutions and practices—by what we might call culture. And since these differ from place to place and time to time, it is not clear that all rational agents must even share a mode of evaluation, much less that there is some thing (*moral value*) with which all rational beings should be concerned. Indeed, for such philosophers, the very idea that moral value is that value which has a claim on any rational being is itself a (dubious) product of Enlightenment culture, with Kant receiving a large share of blame. Though such philosophers are not skeptical about moral value, they are denying a particular way of demarcating moral value from supposedly other values. If such philosophers are correct, moral values would seem to have some of the hallmarks commonly associated with personal values: they depend upon the particular position and concerns of agents.

AGENCY

It is scarcely possible to imagine a moral theory that does not depend upon claims about the nature of human or rational agency, but the degree to which moral philosophers actively look to accounts of agency in developing

and defending normative theories fluctuates over time. Philosophers such as David Hume and Kant based their moral philosophy on sophisticated and original accounts of agency. By contrast, the beginning of the twentieth century saw a comparative neglect of agency. At least two factors may provide a partial explanation.

The first is that the early intuitionists, beginning with Moore, were highly interested in human agents as knowers of moral propositions or facts. If moral propositions were intuitive or self-evident, then the serious and difficult question became how human beings have such knowledge. On the surface this does not preclude a deep interest in agency (perhaps, in a Kantian vein, human beings know about the moral necessity of certain acts by knowing something about human agency). But the surrounding philosophical climate tended to eclipse this alternative. The nineteenth century saw the rapid and exciting development of psychology as an empirical science. Add to this the fact that Moore's Open Question Argument made it seem scandalous even to appear to derive moral conclusions from empirical observation. But if human agency is ultimately a matter of psychology and psychology is a matter to be settled by empirical methods, then the intuitive moral knowledge the intuitionists attributed to human beings could not be based upon knowledge of agency lest the dreaded naturalistic fallacy be committed. The second factor that may partially explain philosopher's relative neglect of agency is the fact that philosophers were busy looking elsewhere, especially to linguistic phenomena.

The latter part of the twentieth century, however, saw renewed and intense interest in the integration of normative theories with theories of agency. In part, this may be traced to the rise of Kantian ethics that treats moral principles as principles of rational willing. It was also no doubt influenced by the rise of *action theory* in the latter half of the twentieth century as well as by the renewed interest philosophers showed in problems of personal identity and by exciting work done on the issue of free will. All of these developments had the effect of drawing attention to human beings as actors with values, ideals, beliefs, emotions, evolving desires and interests, plans, habits, addictions, and more. This is quite different from looking to human beings as perceivers of value.

It is not possible to summarize all of the ways in which moral philosophy has been impacted by its renewed and increasing attention to agency. One important theme concerns the possible analogy between the interest agents are rational to take in their own future concerns and the interest agents are morally required to take in the concerns of others. We think of agents as unified across time. To reprise an example of Nagel's: If I expect that I will want to eat a persimmon next week, then I will be concerned to take the steps necessary to make this possible (even if I do not now care whether I eat a persimmon next week or care whether I come to want to eat a persimmon next week). Indeed, the ability to integrate one's desires and concerns this way is considered a hallmark of prudential rationality. In *The Possibility of Altruism* (1970), Nagel argued that, properly understood, vindicating the rationality of caring about the interests of your own future self also shows how one may vindicate the rationality of caring about the interests of others. Though he later abandoned this argument, Nagel's effort to integrate the justification of important moral norms with an account of an agent as a person persisting through time and potentially aware of other persons and their interests was of lasting importance. Philosophers such as Parfit, as well as Nagel himself, continued to develop these themes.

One such development was an increased attention to a distinction, first introduced by Nagel, between *agent relative* and *agent neutral reasons*. Though the proper way of drawing the distinction is a difficult technical matter, the intuitive heart of the distinction is between those reasons that are reasons for some agents but not for others and those reasons that are reasons no matter who you are. For example, I may have a reason to hold a birthday party for my child because it is my child. But many would doubt that you have a reason to hold a party for my child or even to help me hold one (though you might have a reason to hold a birthday party for *your* child). The reason depends on who I am. By contrast, if a stranger is about to step accidentally in front of an oncoming bus and I can pull him back, then I have a reason to do so. In this case, however, many people are inclined to agree that this is a reason anyone else has as well, or at least would have, provided only that they were in a position to do something about the matter.

The contrast between agent-relative and agent-neutral reasons is of importance in part because agent-relative reasons seem to be at play in the kinds of personal values discussed in the prior section. But many philosophers have also come to see them at work in important aspects of moral thinking, especially in the nature of *deontological* prohibitions on certain kinds of action. On one plausible interpretation, a person committed to a prohibition against, for example, killing the innocent, will care deeply about who does any killing. Though such a person may recognize any killing of the innocent to be

bad, it may matter greatly to them that they not be the one doing the killing. Because many philosophers are interested in defending the rationality of such moral norms, locating their place in an account of the kinds of reasons that agents may have has become a critical question.

THE RISE OF NATURALISM IN MORAL PHILOSOPHY

For much of the twentieth century, the empirical study of moral judgment, norms, and behavior was given little attention by moral philosophers. This state of affairs is undergoing a profound reversal. The present discussion will not attempt to summarize the state of developing research, but three types of inquiry are notable.

First, the rapid development of evolutionary biology spawned renewed interest in developing accounts of how moral norms might have arisen out of a process of evolution. Of course, the suspicion that moral norms do have some such history is an ancient one, but the techniques of modern evolutionary research, including the tools of game theory and the computer simulation, have made it possible to better develop and critically assess possible explanations. Of particular interest are norms regarding helping behavior (or altruism), cooperation, and fair dealing. There are often significant differences in the use of terminology between those conducting this empirical research and those engaged in more traditional moral philosophy. For example, in discussions of the evolution of moral norms, altruistic behavior is often any behavior that in fact benefits another whereas for many traditional moral philosophers, behavior is altruistic only if it is undertaken for the sake of benefiting another. So long as these terminological issues are treated with care, however, they do not seem to present any decisive obstacles.

Another area of rapidly developing research is the application of neuroscience and cognitive science to the topic of moral judgment. While moral philosophers have long deployed hypothetical cases as intuition pumps, the recent use of brain scanning techniques appears to reveal that moral thinking about different kinds of cases involves activity in distinct brain regions. Such results are of interest not least because many moral arguments depend upon claims that different cases are analogous and so merit comparable analysis. The science involved in these studies is both complex and rapidly evolving, and many philosophers and scientists expect increasingly fine-grained and thorough results.

A third area where empirical research has blossomed and impacted moral philosophy is in the area of social psychology. Moral theories of all stripes attribute to human beings beliefs in moral principles, or acceptance of moral norms, or possession of virtuous character traits. In each case the attribution typically brings along with it an expectation that an agent of whom it is true will be appropriately motivated to act accordingly. In this way, moral theories claim that agents have certain moral outlooks and that these outlooks impact the agents' behavior. In short, moral theories make claims about agency. Whether agents really are so motivated, however, appears to many to be an empirically testable hypothesis. Using the tools of social psychology, some scientists and philosophers have begun to emphasize the degree to which human beings of all stripes seem to be influenced by what would seem to be morally irrelevant factors.

For example, one might have thought that whether a person would help a stranger would depend largely on whether that person was a good person (perhaps because they have the virtue of kindness, or because they accept some moral principle that dictates helping others). Experimental studies, however, seem to reveal that whether people help each other is highly correlated with such factors as whether they are in a good mood, a factor that most would count as morally irrelevant and as precisely the kind of thing that sound moral commitment would get round. Thus, some have suggested that certain moral theories may rest upon dubious or false presuppositions about human agency.

In each of these cases, the precise relevance of empirical findings to more familiar questions of moral philosophy is disputed and uncertain. Few, if any, researchers believe that such empirical findings straightforwardly reveal which of our moral commitments are worthy of our endorsement and which are not. Brain scans do not tell us whether a moral intuition is to be trusted or not; evolutionary accounts of the development of norms do not tell us whether those norms are morally worthy or not. At least they do not do so in any simplistic way. Neither, however, do many philosophers assume that such empirical findings are ultimately irrelevant to the familiar normative questions of moral philosophy. For many moral philosophers, the question is not whether empirical science is relevant but how so. The absence of any consensus answer to this question may be due in part to the fact that the empirical sciences in question are not yet fully developed. Perhaps as we get a better scientific picture of the nature and history of moral norms and judgment, the relevance of this picture to normative questions will become clearer. Of note as well, though, is the fact that there is no agreed-upon epistemology for settling

normative questions, and unless there is, it seems unlikely that philosophers will be able to agree about the relevance of empirical studies. Indeed, one may suspect that questions of how empirical results are relevant to normative questions will itself become an important locus of dispute (if it is not so already) in assessing rival moral epistemologies. It is often remarked that Moore's writings set the stage for the development of twentieth-century moral philosophy, and in many ways, they did. But he surely would not have written this ending.

See also Anscombe, Gertrude Elizabeth Margaret; Consequentialism; Empiricism; Enlightenment; Epistemology; Foot, Philippa; Hare, Richard M.; Hedonism; Hume, David; Idealism; Intuition; Kant, Immanuel; MacIntyre, Alasdair; Mackie, John Leslie; Mill, John Stuart; Monism and Pluralism; Moore, George Edward; Nagel, Thomas; Naturalism; Parfit, Derek; Plato; Rawls, John; Ross, William David; Sidgwick, Henry; Singer, Peter; Utilitarianism; Virtue Ethics; Williams, Bernard.

Bibliography

Adams, Robert M. "Motive Utilitarianism." *Journal of Philosophy* 73(1976): 467–481.

Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33 (1958): 1–19.

Brandt, R. B. "Toward a Credible Form of Utilitarianism." In *Morality and the Language of Conduct*, edited by H.-N. Castañeda and G. Nakhnikian. Detroit, MI: Wayne State University Press, 1963.

Dancy, Jonathan. *Moral Reasons*. Oxford: Blackwell, 1993.

Hare, Richard M. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Clarendon Press, 1981.

Hooker, Brad. *Ideal Code, Real World: A Rule-consequentialist Theory of Morality*. Oxford: Oxford University Press, 2000.

Lyons, David. *Forms and Limits of Utilitarianism*. Oxford: Oxford University Press, 1965.

MacIntyre, Alasdair. *After Virtue: A Study in Moral Theory*. Notre Dame, IN: University of Notre Dame Press, 1981.

Mackie, J. L. "Sidgwick's Pessimism." *Philosophical Quarterly* 26 (1976): 317–327.

McDowell, John. "Virtue and Reason." *Monist* 62 (1979): 331–350.

Medlin, Brian. "Ultimate Principles and Ethical Egoism." *Australasian Journal of Philosophy*. 35 (1957): 111–118.

Mill, John Stuart. *System of Logic Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* (1843). 8th ed. New York: Harpers & Brothers, 1874.

Moore, G. E. *Principia Ethica*. Rev. ed., edited by Thomas Baldwin. Cambridge, U.K.: Cambridge University Press, 1903; 1993.

Nagel, Thomas. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press, 1970.

Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13 (1984): 174–131.

Rawls, John. "Two Concepts of Rules." *Philosophical Review* 64 (1) (1955): 3–32.

Ross, W. D. *The Right and the Good*. Oxford: Oxford University Press, 1930.

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1972): 229–243.

Williams, Bernard. "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and Bernard Williams. New York: Cambridge University Press, 1973.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Sean D. McKeever (2005)

ETHICS, TELEOLOGICAL

See *Teleological Ethics*

ETHICS AND ECONOMICS

Economics is linked to both ethics and the theory of rationality. Economics complements and intersects with moral philosophy in both the concepts it has constructed and in its treatment of normative problems.

RATIONALITY, UTILITY THEORY, AND WELFARE

At the foundation of economics lies a normative theory of individual rationality. The theory raises no questions about the rationality of ultimate ends and few questions about the rationality of beliefs. It maintains that an agent A chooses (acts) rationally if A's preferences are rational and A never prefers an available option to the option chosen. A's preferences are rational only if they are transitive and complete—that is, A can consistently rank all alternatives. If an agent's preferences are complete and transitive and satisfy some technical conditions, they can be represented by numbers. These numbers, which are arbitrary apart from their order, are "ordinal utilities." If an agent's preferences satisfy additional conditions concerning risky or uncertain alternatives, then they can be represented by a cardinal utility function. In contemporary economic theory, utility is merely an index locating alternatives in a preference ranking, not a substantive good.

Given economists' commitments to utility theory in explaining human choices, it is natural that they would look to levels of utility (preference satisfaction) as the

measure of welfare. It is, however, difficult to justify identifying welfare with the satisfaction of preferences, even as a simplification. Satisfying preferences that depend on false beliefs is often harmful, whereas satisfying preferences that do not concern oneself is typically irrelevant to one's welfare. Many philosophers endorse a more nuanced identification of well-being with the satisfaction of "informed" and self-interested preferences, and they can thus employ some of the framework of normative economics. Even with these qualifications, it is questionable whether taking well-being to be the satisfaction of preferences is suitable for assessing claims to scarce resources. Should one measure the well-being of those who have learned—possibly quite rationally—not to want what they have not gotten by the satisfaction of their preferences?

Some economists propose different conceptions of well-being. Of particular contemporary interest is Amartya Sen's capability approach. In Sen's view, a capability is the ability to achieve a certain sort of "functioning": literacy is a capability; reading is a functioning. People may value capabilities for their own sake as well as for the kinds of functioning they permit; someone who stays in his or her room may still be glad to know that the door is not locked.

One advantage of more objective approaches such as Sen's is that they avoid the problems of interpersonal comparison that derive from identifying well-being and preference satisfaction. Formerly, economists such as Pigou cited diminishing marginal utility of income to argue that a more equal distribution of income would increase total welfare. This argument compares the contributions income makes to the well-being of different people.

Once one takes seriously the preference-satisfaction view of well-being, these comparisons become problematic. Lionel Robbins argued that there is no objective way to compare the extent to which A's and B's preferences are satisfied, and most (though not all) economists maintain that economic evaluations must not rely on interpersonal-utility comparisons.

EFFICIENCY AND PARETO OPTIMALITY

Efficiency has a technical meaning within normative economics. Suppose that utility is the satisfaction of preferences. Consider some allocation of resources S—S is a "Pareto Improvement" over some other allocation R if and only if it increases the utility (preference satisfaction) of at least one person without decreasing anyone's utility. In other words, S is a Pareto improvement over R if and only if someone prefers to R, and nobody prefers R to S. S is "efficient" or "Pareto optimal" if no other allocation is a Pareto improvement over S. If S is Pareto optimal, then every alternative that satisfies someone's preferences better leads to someone else's preferences being less well satisfied. The Pareto concepts permit economists to rank some social states in terms of preference satisfaction without making interpersonal-utility comparisons.

If one is minimally benevolent and favors satisfying people's preferences, then, other things being equal, one will endorse Pareto efficiency. Moreover, it can be proved that competitive equilibria under certain idealized conditions (no externalities, no public goods, no informational limits, and so on) are Pareto efficient. Minimal benevolence then implies that competitive equilibria are (other things being equal) morally good economic states. A second theorem shows that an efficient economic outcome with any desired distribution of welfare can be attained by a competitive market, given the right initial distribution of endowments to agents. A preference-satisfaction view of well-being combined with minimal beneficence establishes the moral claims of efficiency.

Efficiency judgments capture only one moral dimension along which to assess economic policies, institutions, processes, and outcomes. Rather than pretending that efficiency judgments are conclusive or conceding that they reflect only one of a great many evaluative perspectives, economists generally regard economic evaluation as two-dimensional. In addition to questions of efficiency—with respect to which economists claim a special competence—there are also questions concerning distribution or equity, about which economists typically have little to say.

This view of economic evaluation is inadequate because the Pareto concepts have very limited applicability: Economic changes usually involve winners and losers. One way to extend the assessment of efficiency is via the notion of a potential Pareto improvement, where there are winners and losers in terms of preference satisfaction but the winners are able to compensate the losers. No compensation is actually required. Kaldor and Hicks thought that a potential Pareto improvement showed that the economic "pie" had grown larger, whereas questions about who wins and who loses concern equity not efficiency and should be left to the political process. This view is subject to technical difficulties, and the bottom line is that there is no way to judge changes that affect distributions while remaining neutral on distributive questions.

MORAL MATHEMATICS

The tools and theories of economists have contributed significantly to moral philosophy. Contemporary examples can be found in the literature on egalitarianism or on measures of freedom. The two branches of economics that have been most relevant to moral philosophy are social choice theory and game theory.

One can call any ranking of social states a "social-welfare function," and normative principles can be regarded as constraints on social-welfare functions. The Pareto principle, for example, picks out those social welfare functions that rank R over S if somebody prefers R to S and nobody prefers S to R. Economists have hoped to identify additional plausible normative principles relating individual and social welfare and from them to deduce some general method of evaluating outcomes, policies, and institutions. This framework is quite limited. Procedural matters such as fairness or due process apparently count only instrumentally. Furthermore, in investigating the implications of principles constraining acceptable relationships between individual and social values, Kenneth Arrow wound up establishing a striking impossibility theorem.

Social choice theory is the proof and interpretation of theorems concerning the aggregation of preferences, judgments, and interests. The relevance of social-choice theorems to morality depends on what is being aggregated and for what purpose. Does one seek principles for making social decisions or for carrying out social evaluations? Is one aggregating preferences or judgments? A good deal of social-choice theory, like John Harsanyi's derivation of utilitarianism, consists of formal arguments for moral conclusions, but the most important role of social choice theorems has been to reveal ambiguities and difficulties in apparently plausible moral principles. The interpretation of social-choice theorems is a subtle and complex task.

Game theory is concerned with strategic circumstances, where outcomes depend on the choices of several agents. The theory of games is particularly relevant to moral philosophy in its analysis of interaction problems of moral importance. Problems of social cooperation are often complicated, and it may be enlightening to think about recurring patterns. The most famous of these is the so-called "prisoner's dilemma," in which individuals who act in their self-interest reach a worse outcome for everyone than agents who do not. Prisoner's dilemmas vividly represent problems of social cooperation, free-riding, and public-goods provision.

Modeling interactions with game theory is a subtle task, because there are many different simple games and because simple models abstract from so much. Actual interactions, unlike prisoner's dilemmas, are rarely "one-shot" games, and game-theoretic analyses of repeated interactions are complicated and controversial. Even if game theory were in a more settled state, there would be grounds to hesitate before employing it to address ethical questions. One may have qualms about its focus on preference satisfaction and about whether the only perspective for individuals to adopt in social interactions is individual maximization. Nevertheless, some philosophers and economists such as Geoffrey Brennan, James Buchanan, David Gauthier, and Ken Binmore have used game theory to argue for views of justice.

CONCLUSION

Economists and moral philosophers share interests in rationality and in evaluating social institutions, processes, outcomes, and policies. Although the theoretical starting points of economists and moral philosophers differ, the two subjects have a great deal to offer each other. Moral philosophers who want their work to bear on social institutions, policies, and outcomes have much to learn from economists, who have studied the consequences of alternative policies and who have sought operational measures of theoretical concepts. Economists also offer moral philosophers formal and conceptual tools. At the same time, economists concerned with evaluating institutions and policies cannot avoid thinking about ethics.

See also Good, The; Utilitarianism; Value and Valuation.

Bibliography

Arrow, Kenneth. *Social Choice and Individual Values*. 2nd ed. New York: Wiley, 1963. This seminal contribution gave birth to social choice theory.

Binmore, Kenneth. *Game Theory and the Social Contract*. Cambridge, MA: MIT Press, 1998. A two-volume application of game theory to the theory of justice by a leading game theorist.

Brennan, Geoffrey, and James Buchanan. *The Reason of Rules: Constitutional Political Economy*. New York: Cambridge University Press, 1985. An application of game theoretic reasoning to problems of constitutional design.

Broome, John. *Weighing Goods*. Oxford: Basil Blackwell, 1991. A philosophical treatment of problems of rationality and well-being.

Dworkin, Ronald. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press, 2000. Essays on egalitarianism that reflect the influence of economics.

Elster, Jon. *Sour Grapes: Studies in the Subversion of Rationality*. New York: Cambridge University Press, 1983. A study of the notion of rationality and of failures of rationality.

Elster, Jon, and John Roemer, eds. *Interpersonal Comparisons of Well-Being*. Cambridge, U.K.: Cambridge University Press, 1991.

Gauthier, David. *Morals by Agreement*. Oxford: Oxford University Press, 1986. An ambitious attempt to derive a theory of justice from bargaining theory.

Hardin, Russell. *Morality Within the Limits of Reason*. Chicago: University of Chicago Press, 1988. An application of game theory and rational choice theory to the defense of utilitarianism.

Harsanyi, John. "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility." *Journal of Political Economy* 63 (1955): 309–321. A social-choice theoretical "proof" of utilitarianism.

Harsanyi, John. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge, U.K.: Cambridge University Press, 1977. A general development of rational-choice theory and game theory with a systematic discussion of interpersonal utility comparisons.

Hausman, Daniel, and Michael McPherson. *Economic Analysis and Moral Philosophy*. Cambridge, U.K.: Cambridge University Press, 1996. A comprehensive survey of literature on ethics and economics.

Hicks, John. "The Foundations of Welfare Economics." *Economic Journal* 49 (1939): 696–712. A classic statement of "new welfare economics" based on the compensation criterion.

Kaldor, Nicholas. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *Economic Journal* 49 (1939): 549–552. The first presentation and application of the notion of a potential Pareto improvement.

Kolm, Serge-Christophe. *Justice and Equity*. Translated by Harold See. Cambridge, MA: MIT Press, 2002. A pathbreaking book drawing on economic concepts in addressing philosophical questions.

Little, Ian. *A Critique of Welfare Economics*. 2nd ed. Oxford: Oxford University Press, 1957. A classic discussion of the difficulties of welfare economics without interpersonal comparisons.

McClennen, Edward. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge, U.K.: Cambridge University Press, 1990. A general development of the theory of rational choice including a critique of the standard theory and the development of an alternative.

Nussbaum, Martha. *Women and Human Development*. Cambridge, U.K.: Cambridge University Press, 2001. An alternative presentation of the capability approach.

Pigou, Arthur C. *The Economics of Welfare*. London: Macmillan, 1920. The classic statement of traditional utilitarian welfare economics.

Robbins, Lionel. *An Essay on the Nature and Significance of Economic Science*. 2nd ed. London: Macmillan, 1935. A major methodological overview with a classic argument against the possibility of interpersonal comparisons.

Samuelson, Paul A. "Evaluation of Real National Income." *Oxford Economic Papers* N.S. 2 (1950): 1–29. A demonstration of the impossibility of disentangling efficiency and distributional questions.

Scanlon, Thomas. "Preference and Urgency." *Journal of Philosophy* 72 (1975): 655–670. An argument that questions of justice cannot be decided only by considerations involving preference satisfaction.

Sen, Amartya. *Inequality Reexamined*. Cambridge, MA: Harvard University Press, 1992. An introduction to Sen's more recent work on inequality, capabilities, and welfare.

Sen, Amartya. *On Ethics and Economics*. Oxford: Blackwells, 1987. A concise but demanding introduction to the subject.

Sen, Amartya, and Bernard Williams, eds. *Utilitarianism and Beyond*. Cambridge, U.K.: Cambridge University Press, 1982. An influential anthology of essays on economics and ethics with particular reference to utilitarianism.

Taylor, Michael C. *The Possibility of Cooperation*. New York: Cambridge University Press, 1987. A technical and philosophical study of game-theoretical problems of repeated interactions.

Daniel M. Hausman (2005)

ETHICS AND EPISTEMOLOGY

See *Epistemology and Ethics, Parallel Between*

ETHICS AND MORALITY

Ethics signifies an aspect of human life that can also be called *morality*. There seem only minor differences in usage between the two terms. We speak more naturally of *professional ethics* than of *professional morals* to refer to virtues and codes of behavior of specific professions. This is not, however, because the word *moral* is restricted to human beings, or rational persons, as such, because we also speak naturally of *role morality*. A somewhat more substantial difference is that some forms of behavior especially related to sexuality, such as homosexuality, abortion, and premarital intercourse, are condemned (by some) as *immoral* where *unethical* would not be used. This usage may require a notion of a natural order (possibly of a religious character) that certain actions violate, even if they do not cause harm in some other way. At the same time, *immoral* and *unethical* are often used interchangeably, and both historically and contemporaneously, both can connote wrongness in actions and vice in character.

One influential attempt to get philosophical mileage out of the distinction between morality and ethics is Bernard Williams's *Ethics and the Limits of Philosophy* (1985). Williams proposed that ethics concerns how one should live (although excluding purely egoistic answers to

that question), and morality a systematic but narrower set of concerns that constitute one among many approaches to the ethical. Common usage does not support this linguistic suggestion, but the suggestion of broader and narrower ways of construing the subject with which moral philosophy deals has been influential in its own right.

What Williams refers to as *morality* is essentially Kant's view of it, although that view shares features with other philosophers, and contemporary Kantians have challenged aspects of Williams's reading of Kant. Prominent distinguishing features of the morality system, according to Williams, are the following:

(1) *Obligation is the fundamental moral notion*. However, considerations that render an action obligatory, such as its reducing the suffering of others, or involving defending an honorable person against attack, may also, in some circumstances, render an action good but not obligatory (sometimes called *supererogatory*). In *ethics*, by contrast, good actions can be those instantiating virtues, such as courage, justice, or compassion, without further assessment of the action as obligatory or supererogatory.

(2) *The source of moral demand lies within the agent's own autonomous self*. Yet most moral outlooks recognize some obligations and moral demands as arising, irreducibly, from outside ourselves, for example, from social or institutional roles we occupy, or relationships, such as familial ones, that are not simply voluntarily assumed.

(3) *Ethical assessment encompasses only that for which we are fully responsible*—that is, voluntary actions. Yet, Williams notes, we standardly treat as reflecting on an agent's ethicality responses (such as emotions and feelings) as well as actions, a dimension of the moral life especially emphasized by Aristotle and Iris Murdoch (1970). More generally, we see assessment of character, which is never entirely voluntary, as morally appropriate. Here Williams fails to note, and sometimes implies otherwise, that in the virtue tradition some degree of voluntariness is required for moral assessment. An emotional reaction or disposition over which (or over whose causal antecedents) the agent had absolutely no control whatsoever would not be a fit object of ethical assessment.

A comparable, but not equivalent, view of the difference in question here is that between universal moral requirements and the good life or personal flourishing. Like Williams's distinction between ethics or virtue and the morality system, this distinction is far from sharp.

See also Duty; Virtue Ethics; Williams, Bernard.

Bibliography

Murdoch, Iris. *The Sovereignty of Good*. London: Routledge, 1970.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Lawrence Blum (2005)

ETHICS AND RELIGION

See *Religion and Morality*

EUCKEN, RUDOLF CHRISTOPH

(1846–1926)

Rudolf Christoph Eucken, the German philosopher of life, was born in Aurich, East Friesland. He studied philology and philosophy at the University of Göttingen; after attaining his doctoral degree, he taught several years at Frankfurt Gymnasium. In 1871 he became professor of philosophy at the University of Basel, and in 1874 at Jena, where he remained until his death. In 1908 he received the Nobel Prize in literature.

Eucken was not a systematic philosopher. He began with life as man experiences it. Life inevitably tends to organize into "systems of life" that are organic or institutional. The function of philosophy is to make the meaning of each system explicit and, by explicating each, to raise the question, Which is to be preferred? But philosophy does not merely explicate; it also helps to transform existing life systems. Men assess these explications practically, in terms of their fruitfulness for life or for a particular life system. Each man chooses a life system, but he does not choose one simply for himself. Every act of such choosing inevitably involves other men. There is no escape for any man from this social involvement.

Life is a process, an evolution; it cannot be contained within the boundaries of any philosophy or life system. The strains and stresses created when life breaks its established boundaries raises the deep need for a new philosophy or new philosophies, and inevitably men develop them. Eucken believed that every significant new philosophy is more comprehensive and clearly defined than any past philosophy.

The elaboration of new philosophies comes only through action (i.e., activism), through man's relentless affirmation of life—an affirmation which recognizes both the good and evil inherent in life. No significant phi-

losophy is ever purely intellectualistic, for life is more than an idea or a theory. At its best, life is creative energy bursting into expression and molding past and present experience into a higher, more spiritual unity and order. For Eucken, life is neither noological nor psychological nor cosmological; its basis and meaning are to be found in man.

Life in man is self-conscious; as such, it goes beyond the subjective individual to bind together all conscious beings. Through this transcendence, it becomes the "independent spiritual life," or man reaching through action toward the absolute truth, beauty, and goodness. This "independent spiritual life" is attained only as personality is developed, but it is never a final achievement, since it is always a process that evolves as history. It is not rooted in the external world but in the soul, and it manifests itself more and more completely as the soul becomes independent of this world, self-willed yet subordinate to the ultimate trinity of truth, beauty, and goodness. These ultimates are not theoretical abstractions; they are concrete human experiences that push man beyond cosmic nature to something transcendentally spiritual.

Man has his beginning in nature, but through his soul evolves beyond it. His soul raises questions such as "Why?" and "Whence?" and opposes nature at all points. His soul seeks to become timeless and above nature, even as it feels helpless in the grasp of nature. In spite of this feeling of helplessness, it continues to seek freedom—a freedom realized through the creation of a consistent philosophy that makes possible man's physical and spiritual survival. For Eucken, thought is not something intrinsic to itself but a means, or organ, of life itself.

The need for a new philosophy, Eucken felt, arises from two social conditions—modern man's drive for a "broader, freer, cleaner life, a life of greater independence and spiritual spontaneity" and his drive for a "naturalistic culture … which limits all its activity to the world around us" (*Can We Still Be Christians?*, p. 51).

The first drive provides modern man with a basis for radically transforming classical Christianity. Man's new problems, created by science, transcend the theological and ritualistic solutions that Christianity offered for millenniums. The eternal contribution of Christianity is its religious affirmation of universal redemption. But redemption must be combined with new elements of faith (science as the true complement of religion; religious democracy, or the political equality of all religions before man; complete separation of church and state) if Christianity is to help give birth to the new spiritual philosophy needed by man.

Eucken was very critical of naturalism. A naturalistic culture imposes false limitations upon man's essential spirituality. The conception of a naturalistic culture is a result of the impact of science upon man's life—an impact that is essentially good, but dangerous if it leads to the restriction of man's potentialities to the realm of nature only. In two works, *Individual and Society* (1923) and *Socialism: An Analysis* (1921), Eucken clarified the grounds of his criticism of naturalism. The naturalistic approach opens the door to individual freedom, but it is unable to guide man in the proper use of his freedom, since it lacks an overarching conception of unity. It fails to understand the necessity of social cooperation and social cohesion. Intellectualistic idealism understands the necessity for cooperation, cohesion, and unity, but fails to understand the need for individual freedom. The only proper answer, Eucken believed, is spiritual autonomy. Autonomy gives primacy to the whole of which the individual is a part, but it never reduces him to a state of utter subordination to that whole. The individual realizes his own unique freedom through this whole.

In *Socialism: An Analysis* Eucken also offers six criticisms of socialism: It cannot give unity to the life process; it fails to understand man's need for an inner life; it makes the present the only significant moment in man's life and thus cuts him off from the past and future; by reducing men to mathematical equality, it fails to appreciate genuine cultural and spiritual differences among men; espousing no higher faith than naturalism, it reduces social life to a struggle of man against man; and by considering man in purely economic terms, it stunts and aborts his true nature.

Eucken illustrated the attainment of freedom in terms of science and the peaceful society. In science the primary objective is to give man control over nature, but this task can be accomplished only when scientists cooperate by working together. Science, in other words, is essentially social, but it accomplishes its task through the freedom to investigate that is given to scientists. The peaceful society, although not yet attained, plainly depends upon human cooperation, upon no man raising his hand against another.

However, spiritual autonomy is not possible in a naturalistic culture. It rests upon a faith that goes beyond naturalism—the spiritual belief that man can produce a better and a freer world for all of humanity. Such a belief cannot find support in external circumstances alone. It requires the presence in each man of an inner life, a life constantly struggling to attain the good.

See also Beauty; Determinism and Freedom; Naturalism; Truth.

Bibliography

WORKS BY EUCKEN

Geschichte und Kritik der Grundbegriffe den Gegenwart. Leipzig, 1878.

Die Einheit des Geisteslebens in Bewusstsein und Tat der Menschheit. Leipzig, 1888.

Der Sinn und Wert des Lebens. Leipzig, 1908. Translated by Lucy Judge Gibson and W. R. Boyce Gibson as *The Meaning and Value of Life.* London: A. and C. Black, 1909.

Grundlinien einer neuen Lebensanschauung. Leipzig, 1907. Translated by Alban G. Widgery as *Life's Basis and Life's Ideal.* London: A. and C. Black, 1911.

Können wir noch Christen sein? Leipzig, 1911. Translated by Lucy Judge Gibson as *Can We Still Be Christians?* New York: Macmillan, 1914.

Erkennen und Leben. Leipzig: Quelle and Meyer, 1912. Translated by W. Tudor Jones as *Knowledge and Life.* London: Williams and Norgate, 1913.

Sozialismus und seine Lebensgestaltung. Leipzig, 1920. Translated by Joseph McCabe as *Socialism: An Analysis.* London: T.F. Unwin, 1921.

Lebenserinnerungen. Leipzig, 1921. Translated by Joseph McCabe as *Rudolf Eucken: His Life, Work and Travels.* London: T. F. Unwin, 1921.

The Individual and Society. Translated by W. R. V. Brade. London, 1923.

WORKS ON EUCKEN

Gibson, W. R. B. *Rudolf Eucken's Philosophy of Life.* London: A. and C. Black, 1907.

Jones, W. T. *An Interpretation of Rudolf Eucken's Philosophy.* London: Williams and Norgate, 1912.

Rubin Gotesky (1967)

EUDAIMONIA

See Appendix, Vol. 10

EURASIANISM

(Classical) Eurasianism (Russian: *evraziistvo*) was an ideological-philosophical movement among Russian émigré intellectuals in the 1920s and 1930s. Its founders were the ethnologist and linguist Nikolai Sergeevich Trubetskoi, the geologist and economist P. N. Savitskii (1895–1968), the musicologist P. P. Suvchinskii (1892–1985), and the religious philosopher Georgii Vasil'evich Florovskii. Later in the 1920s major supporters and theoreticians joined the movement, including the historian G. V. Vernadskii (1886–1967) and the philosopher Lev Platonovich Karsavin.

Eurasianism was born as a reaction to the Russian revolution and the political situation in Russia after the crisis of World War I, revolution and civil war, and the Bolsheviks' rise to power. The Eurasians were not opposed to the revolution, as it put an end to the bankrupt European tsarist regime, but were against communism, which in their view was also a typical European product. As Russians were not Europeans according to the Eurasians, Russia needed its own ideology, which would do justice to its own particular historical and cultural development.

The Eurasians' anti-European attitude manifested itself from their first joint publication, *Ischod k vostoku. Predchuvstviia i sversheniia. Utverzhdenie Evraziitsev* (Exodus to the East: Forebodings and events: An affirmation of the Eurasians; 1921). It was, in a sense, a continuation of a book published by Trubetskoi in 1920, *Evropa i chelovechestvo* (Europe and mankind). In this book Trubetskoi warned against the imminent Europeanization of the world, which was the direct result of Romano-Germanic chauvinism. The colonialist countries of western Europe considered other cultures inferior to their own culture and tried to put their stamp everywhere in the world. Russia, according to Trubetskoi, had to stand up to this pernicious influence and follow its own path.

A central concept of Eurasianism was a geographical one. The Eurasianists did not divide the enormous surface of land of the Northern Hemisphere into two (Europe and Asia), but into three parts: (Western) Europe, including Poland, the Baltic states and the Balkans, Asia—the Far East, Southwest Asia, and Southeast Asia—and in between as a separate geographical world the relatively flat area extending from the Danube estuary to the Lena River basin. It is the area that was formerly controlled by the Mongol ruler Genghis Khan (c. 1162–1227) and that roughly corresponded with the territory of the Russian tsarist regime and with that of the Soviet Union. Contrary to Europe and Asia, Eurasia is not open to the sea; it has a continental climate with hot summers and cold winters.

Eurasia is inhabited by various peoples, Slavic as well as non-Slavic or, as Trubetskoi calls them, Turanic peoples, including the Finno-Ugrians, Samoyeds, Tartars, Turkmenians, and Kalmyks. For historical and geographical reasons all these Eurasian peoples form a unity, as do their languages. Culturally and spiritually—and in this respect the Eurasians differ from the nineteenth-century Slavophiles from whom they borrowed much of their ide-

ology—the Russians are closer to the peoples of Eurasia than to their Slavic brothers such as the Czechs and the Poles, who belong to an entirely different world: catholic Europe.

The geopolitical idea of Eurasia was for the first time realized by Genghis, who submitted the entire Eurasian territory for some 150 years to the mogul yoke. Generally, this period is considered as a time of stagnation and decline, but for the Eurasians it was the beginning of the Russian empire, triumphantly continued by the Muscovite state that, after having defeated the Mongols, was even better suited to establish the unity of Eurasia as it had a clear ideology: that of the Orthodox Church. The Eurasians considered Orthodoxy as one of the pillars of Eurasia's cultural identity. Different from the heretical rationalistic and individualistic Catholicism and Protestantism, it was based on brotherhood (*sobornost*) and therefore in an excellent position to unite all the Eurasian peoples, be they Slavic, Islamic, or heathen.

For the Eurasians the reign of Peter the Great (1672–1725), who wanted to bridge the gap between Russia and Europe and make his country one of the European powers, on a par with England and France, was disastrous for Russia's development. In particular, it led to the fragmentation of Russian society: on the one hand a Europeanized upper class, consisting of the tsar, his civil servants, and the intelligentsia; on the other hand the Russian people. Instead of stimulating cooperation between and assimilation of Russians and Turanians, the Russian monarchy aimed at enforced Russification of the Turanian peoples. Moreover, it considered the Orthodox Church primarily as a means of becoming a European power; it suppressed the independence of the church and submitted it to its own secular authority. Peter's policy was continued by the Romanovs who ruled after him. It resulted in the introduction into Russia of all kinds of notions that were strange to the Eurasianist mentality: imperialism, militarism, and capitalism by the state, and liberalism, parliamentarism, and socialism by other, progressive groups in Russian society. As the Soviet state was based on European ideas, it missed the opportunity to start a really new period, based on Russia's national character, history, and civilization. The only way to "save" Russia was to return to its roots, that meant to the spirit of its pre-Petrine past.

The Eurasians did not consider themselves utopians, but tried to start from historic and social facts. In the joint publication *Evraziistvo* (Eurasianism; 1926), which appeared when the movement was at its height, they developed the idea of a real people's party that would endorse their ideology and not that of the communists, but would leave intact many of the economic and political structures the Bolsheviks had introduced. In the 1930s the Eurasian movement gradually declined. The communist ideology proved to be much stronger than expected and in émigré circles many people were opposed to the growing contacts between members of the Eurasian movement and the Soviet regime.

Neo-eurasianism arose at the end of the twentieth century. It builds on the ideas of classical Eurasianism and the ethnogenesis theory of Lev Gumilev (1912–1992; the son of the famous poets Anna Akhmatova [1889–1966] and Nikolai Gumilev [1886–1921]). The movement holds particular geopolitical intentions and considers Eurasia as a separate civilization that under the leadership of the Russians as a young and fresh ethnos will break the dominance of the dying European ethnoses and in this way will change the political and cultural map of the world. Its main theoretician is the (nationalistic) philosopher Aleksandr Dugin (1962–), who is also the leader of the political party Partiia Evraziia (Eurasia Party).

See also Florovskii, Georgii Vasil'evich; Karsavin, Lev Platonovich; Trubetskoi, Nikolai Sergeevich.

Bibliography

Böss, Otto. *Die Lehre der Eurasier: Ein Beitrag zur russischen Ideengeschichte des 20. Jahrhunderts*. Wiesbaden, Germany: Otto Harassowitz, 1961.

Riasanovsky, Nicholas V. "The Emergence of Eurasianism." *California Slavic Studies* 4 (1967): 39–74.

Riasanovsky, Nicholas V. "Prince N. S. Trubetzkoy's 'Europe and Mankind.'" *Jahrbücher für Geschichte Osteuropas* 12 (1964): 207–220.

Shlapentokh, Dmitry V. "Eurasianism Past and Present." *Communist and Post-communist Studies* 30 (2) (1997): 129–151.

Trubetzkoy, Nikolai Sergeevich. *The Legacy of Genghis Khan and Other Essays on Russia's Identity*, edited by Anatoly Liberman. Ann Arbor: Michigan Slavic Publications, 1991.

Vinkovetsky, Ilya. "Classical Eurasianism and Its Legacy." *Canadian-American Slavic Studies* 34 (2) (2000): 125–139.

Vinkovetsky, Ilya, and Charles Schlacks, Jr., eds. *Exodus to the East: Forebodings and Events: An Affirmation of the Eurasians*. Marina del Rey, CA: Charles Schlacks Jr., 1996.

W. G. Weststeijn (2005)

EUSEBIUS

(265–339 or 340)

Eusebius, the church historian and Christian apologist, was bishop of Caesarea (Palestine) early in the fourth

century. He is best known for his enthusiastic support of the emperor Constantine and for his pioneering *Historia Ecclesiastica,* intended to show how the church expanded but always remained the same because of its leaders' fidelity to tradition. Though Eusebius was essentially a historian rather than a philosopher, he did produce one work of significance for the history of philosophy. This is his *Praeparatio Evangelica* (Preparation for the Gospel), probably written between 312 and 318. It consists of fifteen books, perhaps because Porphyry's treatise, *Kata Christianon* (Against the Christians), contained the same number. Eusebius claimed that his treatise went beyond earlier works of controversy or exegesis; the novelty seems to have lain in the method of quoting passages in which philosophers contradict one another, although he obviously found materials for this technique in some of his pagan sources.

The *Praeparatio* may be outlined thus: the earliest cosmogony (I. 7–8); the earliest theology (I. 9); Phoenician theology (I. 10); Egyptian theology (II. 1); Greek mythology (II. 2–8); Greek "physical" theology (III. 1–17); Greek oracles (IV–VI), leading on to the doctrines of Greek philosophers on fate, free will, and foreknowledge (VI); Hebrew doctrines (VII–IX); the chronological priority of Hebrew learning to Greek (X); the agreements and disagreements of Greek philosophy with the Hebrew oracles (XI–XIII); and the inconsistency of Greek philosophy, culminating with a transcription of part of the treatise "On the Doctrines of Philosophers" ascribed to Plutarch (XIV–XV).

The sources used by Eusebius reflect the predominantly Platonic character of the books assembled in church libraries, especially at Caesarea, by Origen and others. Eusebius made extensive use of Plato and Philo, but not of Aristotle. His other sources include the textbooks by Arius Didymus and Pseudo-Plutarch, as well as works by eclectic Platonists of the second Christian century (Atticus, Numenius, Plutarch, Severus) and a few of their contemporaries (the Peripatetic Aristocles, the Epicurean Diogenianus, the Cynic Oenomaus). From the third century he used the treatise "On Fate" by Alexander of Aphrodisias, the school text of Plotinus (earlier than that found in Porphyry's edition), several works by Porphyry ("On Abstinence," "Letter to Anebo," "Against the Christians," "Philological Lectures," "On the Philosophy to Be Derived from Oracles," "On the Soul against Boethus," "On Statues"), and a fragment by a Christian Neoplatonist named Amelius.

His basic viewpoint is that of a Christian ecclesiastic and a historian; he has considerable sympathy for his favorite philosophers (especially Plato), but he is not really at home with them. Indeed, in his later work, *Theophania* (On the theophany), written after 337, his attitude toward philosophy is markedly hostile.

In later times the *Praeparatio* was used as a mine of philosophical quotations by such Christian apologists as Theodoret and Cyril of Alexandria (Cyril often looked up Eusebius' sources and provided slightly different quotations). It would appear that this is its principal value.

See also Alexander of Aphrodisias; Apologists; Aristotle; Cosmology; Numenius of Apamea; Origen; Plato; Plotinus; Plutarch of Chaeronea; Porphyry; Precognition.

Bibliography

WORKS BY EUSEBIUS

The best Greek texts are in *Die griechischen christlichen Schriftsteller der ersten drei Jahrhunderte,* published by the Berlin Academy from 1902 on, although the first volume needs to be reedited (see F. Winkelmann, *Die Textbezeugung der Vita Constantini des Eusebius von Caesarea,* Berlin: Akademie_Verlag, 1962). Translations include H. J. Lawlor and J. Oulton, *Eusebius: The Ecclesiastical History and the Martyrs of Palestine,* 2 vols. (London, 1928), valuable commentary; W. J. Ferrar, *The Proof of the Gospel* (*Demonstratio Evangelica*), 2 vols. (London: Society for Promoting Christian Knowledge, 1920); and E. H. Gifford, *Eusebii Pamphili Evangelicae Praeparationis Libri XV,* 4 vols. in 5 (Oxford, 1903).

WORKS ON EUSEBIUS

Full bibliography in J. Quasten, *Patrology* (Utrecht: Spectrum, 1960), Vol. III, pp. 309–345 (on the *Praeparatio,* pp. 329–331). Later titles include D. S. Wallace-Hadrill, *Eusebius of Caesarea* (London: A.R. Mowbray, 1960) and J. Sirinelli, *Les vues historiques d'Eusèbe de Césarée* (Dakar, Senegal, 1961). On early use of Eusebius' works see Pierre Canivet, *Histoire d'une entreprise d'apologétique au Ve siècle* (Paris: Bloud & Gay, 1958); R. M. Grant, "Greek Literature in the Treatise *De Trinitate* and Cyril *Contra Julianum,*" *Journal of Theological Studies* 15 (1964): 265–279.

Robert M. Grant (1967)

EUTHANASIA

Euthanasia used to refer to an easy and gentle death, but it has come to refer to methods of inducing that kind of death, or more precisely, methods of bringing about death sooner and usually with less pain and suffering. Euthanasia used to be limited to patients in the terminal stage of an illness, but it is now thought to be appropriate in some cases of nonterminal patients, for example, those in a persistent vegetative state and those suffering from an

incurable and very painful chronic disease such as multiple sclerosis.

VOLUNTARY ACTIVE EUTHANASIA

Voluntary active euthanasia (VAE) is when a physician accedes to a rational *request* of an adequately informed, competent patient to be killed, for example, with a lethal intravenous injection of pentothal.

PHYSICIAN-ASSISTED SUICIDE

Physician-assisted suicide (PAS) is when a physician, at a rational *request* of an adequately informed, competent patient who plans to commit suicide, knowingly provides that patient with the medical means to commit suicide and the patient uses those means to commit suicide.

VOLUNTARY PASSIVE EUTHANASIA

Voluntary passive euthanasia (VPE) is when a physician abides by a valid rational *refusal* of treatment by an adequately informed, competent patient knowing that doing so will result in the patient dying, for example, complying with the refusal of a ventilator-dependent patient with motor neuron disease to receive further mechanical ventilatory support. Abiding by *patient refusal of hydration and nutrition* (PRHN) is another example of VPE, as is abiding by such refusals given in advance directives—either living wills or durable powers of attorney for health care—even though the patient is incompetent at the time the treatment is withheld or withdrawn.

Patients are *competent* to make a decision about their health care if they have the ability to make a rational decision. This requires both that they have no relevant mental disorder that prevents them from making a rational decision and that they have the capacity to understand and appreciate all the information necessary to make that decision. They are adequately informed when they have all the information necessary to make a rational decision. Patient competence, having adequate information, and no coercion by the health care team, are the elements of valid (informed) consent or refusal of treatment. If participation in research or donating an organ, rather than treatment, is involved, coercion by anyone invalidates consent.

Decisions by patients are *irrational* if they know the decisions will result in serious harm to them, for example, death, chronic pain, or significant disability, and they do not have adequate reasons for suffering that harm, for example, beliefs that some people, either themselves or others, will thereby avoid suffering an equal or greater harm by that decision. Only those decisions count as irrational that result in the person suffering significant harm and for which almost no one in the person's culture would rank the benefit gained or harm avoided as providing an adequate reason. Often, however, rational people rank harms in different ways; for example, it is rational to rank several months of suffering from a chronic or terminal disease as worse than death and it is also rational to rank death as worse.

INVOLUNTARY ACTIVE EUTHANASIA

Involuntary active euthanasia (IAE) is the killing of a patient who is suffering in order to relieve that suffering but without a request from the patient to be killed. This is most likely to occur with permanently incompetent patients who are unable to make such a request.

INVOLUNTARY PASSIVE EUTHANASIA

Involuntary passive euthanasia (IPE) is allowing a suffering patient to die by ceasing treatment, in order to relieve that suffering, when the patient has neither refused that treatment nor has an advance directive refusing that treatment. This is most likely to occur with permanently incompetent patients who are unable to refuse treatment and do not have an advance directive refusing that treatment. Ceasing treatment for permanently incompetent patients who do not have advance directives refusing that treatment but who have communicated to their families that they would not want to live in this kind of condition, is usually considered to be VPE rather then IPE.

As of 2005, VAE and IAE are illegal in every state in the United States, but PAS is legal in Oregon, and many people have begun to argue for its legalization in other states. IPE is also illegal, except when continuing treatment is considered futile, but almost all those who do bioethics hold either that the definition of futility be broadened so that all treatment of patients in a persistent vegetative state be classified as futile or that some other method be established that allows discontinuing all treatment, including hydration and nutrition, for those in a persistent vegetative state when there is no religious reason for the treatment to be continued.

The United States Supreme Court has explicitly distinguished between VPE and PAS, and by consequence, VAE, holding that only the former is based on a fundamental right to be left alone. The Court correctly regards terminal sedation, that is, being sedated to unconsciousness until one dies, as VPE for those patients who have refused hydration and nutrition and whose pain and suffering cannot be controlled in other ways. VPE is

approved by the American Medical Association and all other medical and legal organizations. Philosophers have attempted to provide an account of VPE that explains its almost universal acceptance. All of these attempts have identified VAE with killing and VPE with allowing to die. Two of the most common ways of distinguishing between VAE (killing) and VPE (allowing to die) are (1) acts versus omissions and (2) withholding versus withdrawing.

The philosophical distinction between acts and omissions seems a natural way to distinguish between killing and allowing to die. On this account, if a physician *does* something, for example, injects an overdose of morphine or turns off the respirator, that is an action and should count as VAE, should be considered killing, and should be prohibited. If the physician does nothing but, rather, simply fails to do something, for example, does not turn on the respirator or does not provide essential antibiotics, that is an omission and should count as VPE, should be considered allowing to die, and should be permitted. However, it seems pointless to distinguish between an authorized physician who turns a knob that stops the flow of life-sustaining antibiotics and one who omits filling the bag when it runs out of those antibiotics. Those who have used the distinction between acts and omissions to distinguish between VAE and VPE have usually concluded that the distinction has no moral significance.

The distinction between withholding and withdrawing treatment seems to have great appeal for some doctors as a way of distinguishing between killing and allowing to die. Some maintain that if patients validly refuse to start a life-saving treatment, doctors do not have a duty to force it on them and so are only allowing them to die. However, once treatment is started, if discontinuing it would lead to the patient's death, they have a duty to continue, and it is killing not to do so. Doctors are not required to force patients to go on the ventilator if they refuse, but once patients have accepted going on the ventilator, doctors have a duty to keep them on if taking them off would result in their death, even if they have had a change of mind.

As with the previous distinction between acts and omissions, there seems to be no morally significant difference between withholding and withdrawing treatment. Physicians do not have a duty to continue treatment if an adequately informed, competent patient rationally refuses to have it continued. Imagine two unconscious patients who are going to be put on a respirator; one becomes conscious *before* being put on and the other *after* being put on, but both are competent, adequately informed, and rationally refuse treatment. This accident of timing is morally irrelevant. Further, this way of distinguishing between active and passive euthanasia may create serious practical problems. Patients who had not been adequately evaluated (often at the scene of an accident) may be judged inappropriate for rescue efforts because the doctors believe that once the patient is on a ventilator they cannot legitimately withdraw it.

The inadequacy of these two attempts to distinguish between VAE and VPE has led many to doubt that there is a morally relevant distinction between them. However, closer attention to the way the distinction is actually made, both in law and medicine, shows that what was overlooked is the crucial role played by the patient. When a patient rationally and validly *refuses* what is offered, the physician is legally and morally required not to overrule that refusal. Abiding by a valid rational refusal, knowing that death will result, counts as VPE whether this involves (1) an act or an omission or is (2) withholding or withdrawing. That everyone acknowledges that a physician must abide by a valid refusal of treatment, whether this involves an action or is a case of withdrawing, explains why VPE is almost universally considered to be morally acceptable.

If a patient *requests* the physician to do something, however, the physician is not morally required to do it if in the physician's judgment it is inappropriate to do so. Physicians may accede to patient requests if they regard them as appropriate, but rarely are they required to do so. If a patient requests that a doctor do something illegal or that the doctor considers immoral, the doctor usually is not required to accede to that request. Killing patients at their rational request is illegal, and even if it were to be legalized, many physicians would still consider it to be immoral. Even granted that it is sometimes morally acceptable for physicians to kill patients at their request, it will never be legally or morally required for them to do so. This is sufficient to distinguish VAE from VPE, for it is legally and morally required for physicians to abide by the rational valid refusals of their patients.

Confusion sometimes arises because a patient's refusal is framed as a request. For example, a patient's *request* that no cardiopulmonary resuscitation (CPR) be attempted counts as a refusal of permission for CPR. Similarly, written advance directives requesting the cessation of other therapies or of hydration and nutrition, count as refusals. Any request for not starting or stopping a treatment is a refusal of treatment, and if the patient is competent and the request is rational, doctors are morally and legally required to abide by it.

Distinguishing between refusals and requests becomes more difficult when the life-prolonging treatment is provided by a device such as a pacemaker that has been implanted in the patient. Even though a pacemaker can be reprogrammed to cease functioning without any surgical procedure, some have considered the pacemaker to have become part of the person and so regard the request to turn it off as a genuine request and not as a refusal. However, the dominant view is that whether the artificial device that is keeping the person alive is inside or outside is not the important consideration. If the device can be turned off from the outside, then the patient's request can be counted as a refusal and should be honored. However, because it is not clear that the patient's request should be counted as a refusal, it is not clear whether the doctor is legally required to have the pacemaker turned off. On the other hand, if a surgical procedure is required, for example, to take out an implanted heart valve, no one would count the request to have it taken out as a refusal, and if any doctor agreed to such a request, that doctor would be regarded as having killed the patient.

Using valid refusal versus requests as the way of distinguishing VPE from VAE, while it explains the moral acceptability of VPE, does not make VAE morally unacceptable. Given present knowledge and technology, a physician can kill a patient absolutely painlessly within a matter of minutes. If there were no way for patients to shorten the time of their dying or for their pain to be controlled, VAE would seem to be clearly morally acceptable. However, PRHN, which, contrary to common belief, does not cause suffering, normally results in patients becoming unconscious within a week and dying within another week, and there is no limit on ways to control their other pain during that time. Because all proposals to legalize VAE or PAS involve at least a two-week waiting period, it seems pointless to argue for legalizing VAE or PAS, which are controversial, rather than providing education about PRHN, a form of VPE, which is already universally accepted. Failure to appreciate the available alternative of PRHN makes arguments such as those presented to the Supreme Court in The Philosopher's Brief in favor of declaring Washington and New York states prohibition of assisted suicide unconstitutional far less persuasive than they otherwise would be.

Abiding by the refusal of an advance directive of a competent patient, when that patient becomes incompetent, is also regarded as VPE. If competent patients explicitly state in advance directives that should they become permanently incompetent they want all life prolonging treatments to be discontinued, then the physician is morally required to abide by that refusal. However, some challenge this view, claiming that the views of the competent person who filled out the advance directive may not be the same as the views of the incompetent person to whom they are being applied. Some hold that advance directives need not be followed if the physician believes that the incompetent person would no longer choose to have life-prolonging treatment withdrawn. A public policy must be judged, however, in terms of the effects that this policy would have on everyone affected if all of them knew of the policy. Competent persons who fill out advance directives refusing life-prolonging treatment if they become permanently incompetent consider it distasteful and devoid of dignity to live as a permanently incompetent person, but after becoming permanently incompetent, the person, having no sense of dignity, does not view life with distaste.

If everyone knew that advance directives would not be honored in these cases, some permanently incompetent persons would live longer than they would if such advance directives were honored. This might be a positive result although it is not clear whether the incompetent person views it in that way. However, it is clear that another result of everyone knowing that their advance directives would not be honored would be anxiety, anger, and other unpleasant feelings by those competent persons who had made out such advance directives. This could result in an increase in deaths of such competent persons in order to avoid the unwanted prolongation of their lives as incompetent persons. That the public policy of honoring advance directives is likely to have better consequences than the public policy of not honoring them justifies the policy.

See also Applied Ethics; Bioethics; Medical Ethics.

Bibliography

Battin, Margaret P., Rosamond Rhodes, and Anita Silvers. *Physician Assisted Suicide: Expanding the Debate*. New York: Routledge, 1998.

Bernat, James L., Bernard Gert, and R. Peter Mogielnicki. "Patient Refusal of Hydration and Nutrition: An Alternative to Physician-Assisted Suicide or Voluntary Euthanasia." *Archives of Internal Medicine* 153 (1993): 2723–2728.

Brock, Dan W., *Life and Death* (especially pp 95–232). Cambridge, U.K.: Cambridge University Press, 1993.

Clouser, K. Danner, "Allowing or Causing: Another Look." *Annals of Internal Medicine* 87 (1977): 622–624.

Dresser, Rebecca S. and John A. Robertson. " Quality of Life and Non-treatment Decisions for Incompetent Patients." *Law, Medicine & Health Care* 17 (3) (1989): 234–244.

Dworkin, Ronald, Thomas Nagel, Robert Nozick, John Rawls, Thomas Scanlon, Thomas, and Judith Jarvis Thomson, as

amici curiae in Supreme Court cases: Washington et al. v. Glucksberg et al. and Vacco et al. v. Quill et al. (Known as *The Philosopher's Brief*), 1996.

Gert, Bernard, *Morality: Its Nature and Justification*. Rev. ed. New York: Oxford University Press, 2005.

Gert, Bernard, Charles M. Culver, and K. Danner Clouser. *Bioethics: A Systematic Approach*. Oxford University Press, 2006.

The Hastings Center. *Guidelines on the Termination of Life Sustaining Treatments and Care of the Dying*. Bloomington: Indiana University Press, 1987.

Lynn, Joanne, ed. *By No Extraordinary Means: The Choice to Forego Life-Sustaining Food and Water*. Bloomington: Indiana University Press, 1986.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. *Deciding to Forego Life-Sustaining Treatment*. Washington, DC: Government Printing Office, 1983.

Quill, Timothy E. *Death and Dignity*. New York: Norton, 1993.

Rachels, James. "Active and Passive Euthanasia." *New England Journal of Medicine* 292 (1975): 78–80.

Rachels, James. *The End of Life: Euthanasia and Morality*. New York: Oxford University Press, 1986.

Van er Maas, Paul J. et al. "Euthanasia and Other Medical Decisions concerning the End of Life." *Lancet* 338 (1991): 669–674.

Supreme Court Decisions: Washington et al. v. Glucksberg et al. and Vacco et al. v. Quill et al., 1997.

Wanzer, Sidney J. et al. "The Physician's Responsibility Toward Hopelessly Ill Patients: A Second Look." *New England Journal of Medicine* 320 (1989): 844–849.

Weir, Robert F. *Abating Treatment with Critically Ill Patients*. New York: Oxford University Press, 1989.

Bernard Gert (2005)

EVANS, GARETH

(1946–1980)

As an undergraduate from 1964 to 1967, Gareth Evans, a British philosopher of language and mind, studied for the PPE degree (philosophy, politics, and economics) at University College, Oxford, where his philosophy tutor was Peter Strawson. In 1968, less than a year after completing his degree, Evans was elected to a fellowship at University College. He took up the position in 1969, succeeding Strawson, who had become Waynflete Professor of Metaphysical Philosophy at Oxford. During the 1970s Evans and his University College colleague John McDowell played leading roles in developing a distinctive conception of truth-theoretic semantics, drawing on the work of Strawson, Michael Dummett, and especially Donald Davidson. Their coedited collection, *Truth and Meaning: Essays in Semantics*, appeared in 1976.

While philosophy of language enjoyed a central position in Oxford philosophy of that period, Evans did not share the view (regarded by Dummett as constitutive of analytic philosophy) that philosophy of language is foundational and so takes priority over philosophy of mind in the order of philosophical explanation. He attached particular importance to the mentalistic notion of understanding, and his work on the theory of reference was set within a theory of thought and especially thought about particular objects. Evans's published work ranged over philosophy of language, metaphysics, philosophy of mind, and philosophy of psychology. In 1979 he was elected to the Wilde Readership in Mental Philosophy at Oxford. He died in August 1980, at the age of thirty-four. His book *The Varieties of Reference* (1982), incomplete at the time of his death, was edited and brought to publication by McDowell. A collection of thirteen of his papers and two shorter notes appeared in 1985, and a further note was published in 2004.

NAMES AND REFERENCE

In his first published paper, "The Causal Theory of Names" (1973/1985) Evans contrasted two theories about the reference of names: the description theory and the causal theory. Evans agreed with Kripke (1972) in rejecting the description theory of reference, which he regarded as drawing support from a flawed account of what is involved in thought directed toward a particular object. In opposition to this description-theoretic account of object-directed thoughts, Evans maintained that a subject may think about a particular object in virtue of standing in a contextual relation to it and without being able to frame any description that the object uniquely satisfies. However, Evans did not accept the causal theory of reference suggested by Kripke's remarks. In the Kripkean picture, the reference of a name is established by an initial baptism and is then passed on from earlier to later users of the name. Evans challenged this picture by highlighting the fact that a name may change its reference over time and, more generally, he argued that a bare causal connection is not sufficient to underwrite reference. As against both the description theory and the Kripkean causal theory, Evans proposed that the bearer of a name is the object that is the dominant source of the body of information that speakers associate with the name.

Many of the themes of his early paper on names—including opposition to description-theoretic accounts of object-directed thoughts, rejection of causal theories as insufficiently demanding, and appeal to the notion of information—recur in *The Varieties of Reference*, set

against the historical background of Gottlob Frege and Bertrand Russell. Evans read Frege as committed to the principle that if a name has no reference, then a sentence containing the name has no truth-value and does not express a thought; a speaker using the sentence does not literally say anything. This "no reference, no thought" principle is in line with Frege's view that the semantic function of a name is to introduce an object, but it appears to rule out the possibility of names with sense but no reference, a possibility that Frege clearly allowed once his distinction between sense and reference was in place. Evans sought to reduce the tension that this reading finds in Frege's position by appealing to Frege's assimilation of the use of empty names to fictional uses of language, which express pretended senses or "mock thoughts."

Evans held that many singular terms—especially demonstratives such as "that ball" or "that vase"—conform to the "no reference, no thought" principle, and he called such expressions "Russellian" singular terms. He also held, following Russell, that definite descriptions, even though they appear superficially to occupy name positions, are not really referring expressions but rather quantifier expressions; "the *F*" semantically resembles "some *F*" and "every *F*." The contrast between a Russellian singular term, whose significance depends on its having a referent, and a definite description, whose significance can be grasped independently of whether it has a denotation, was fundamental for much of Evans's work on reference (Sainsbury 1985).

OBJECT-DIRECTED THOUGHT

Although *The Varieties of Reference* begins and ends with philosophy of language (returning to the topic of names and name-using practices in its final chapter), the central chapters address the issue of thoughts directed toward particular objects. According to the description theory of object-directed thoughts, thought about a perceived, remembered, heard-about, or recognized object, about an occupied place or about a present time, is a matter of the object, place, or time uniquely satisfying a descriptive condition that the thinker frames and deploys in thought. Alternative theories of *de re* (or object-directed) thought appeal to the causal relations implicated in perception, memory, and testimony and to contextual relations to places and times (Burge 1977). While Evans was opposed to the description theory, he was also concerned, here as in philosophy of language, that causal theories were liable to be insufficiently demanding. He was particularly critical of what he called "the photograph model of mental representation," according to which the causal ancestry of a mental state is sufficient to determine which object the state represents (as causal ancestry is sufficient to determine which object a photograph is an image of).

Evans's own theorizing about object-directed thoughts was guided by Russell's principle, which says that to think about a particular object, a thinker must know which object is at issue. Evans interpreted the principle as requiring discriminating knowledge, that is, the capacity to discriminate the object of thought from all other things, and this, at least initially, sounds so demanding as to make object-directed thought an extraordinary achievement. But Evans's examples of how to meet the principle make it seem more tractable: presently perceiving the object, being able to recognize it, knowing discriminating facts about it.

When a thinker meets the requirement of Russell's principle by having discriminating knowledge of a particular object, the thinker is said to have an adequate Idea of the object. In this technical use of the term, an Idea deployed in thought about an object is analogous to a concept deployed in thought about a property. Evans was particularly concerned with cases (centrally, cases of demonstrative identification) in which a thinker's Idea of an object depends on an information link between the thinker and the object, so that the Idea of the object, and thoughts in which the Idea is deployed, are information-invoking. The picture here is not that the information link contributes to the thought about the object, because the thinker frames a descriptive condition along the lines of "the object, whichever it is, that is the source of this information." It is the information link itself, and not a thought about the information link, that plays a role in making object-directed thought possible.

If, as a result of malfunction or hallucination, there is really no information link to an object, then the thinker has no adequate Idea of this information-invoking kind. A thinker who is unaware of the problem may essay a thought and yet fail to think about any particular object at all. Information-invoking thoughts (centrally, demonstrative thoughts) are object-dependent; where there is no object, there is no thought. Evans was especially interested in cases where understanding a singular term requires an information-invoking thought, and hence object-dependent thought, on the part of the hearer. For in such cases, it is possible to argue that the singular term is Russellian, that its significance depends on its having a referent.

DESCRIPTIVE NAMES

Despite the central role played by Russellian singular terms in *The Varieties of Reference*, Evans did not equate the categories of Russellian singular terms and referring expressions. In "Reference and Contingency" (1979/1985), he considered descriptive names (names whose referents are fixed by description). His example was the name "Julius," introduced with the stipulation "Let us use 'Julius' to refer to whoever invented the zip fastener [or zipper]." "Julius" behaves epistemically and modally like the definite description "the actual inventor of the zip." Evans offered "If anyone uniquely invented the zip, Julius invented the zip" as an example of a sentence whose truth can be known a priori, even though it is contingent. Evans argued that the thought expressed by the nonmodal sentence "Julius is *F*" is the same as the thought expressed by "The inventor of the zip is *F*"—a thought that can be grasped whether or not "Julius" refers to anyone. But he rejected the suggestion that descriptive names belong semantically with definite descriptions and maintained that, although the descriptive name "Julius" is not a Russellian singular term, it is still a referring expression. His argument for placing descriptive names in the category of referring expressions, alongside Russellian singular terms and separate from definite descriptions, involved two main points. First, as the introducing stipulation makes clear, the semantic contribution of "Julius" is stated using the relation of reference, no less than is the semantic contribution of a Russellian singular term ("John" refers to John). Second, even in a semantic theory for a modal language, the semantic contribution of a descriptive name, like that of a Russellian singular term, can be stated using a reference relation that is not relativized to possible worlds, but this is not generally so for definite descriptions.

INFORMATION AND NONCONCEPTUAL CONTENT

The notion of information, as Evans used it, is not the notion of what a subject believes. Indeed, Evans suggested that we should take the notion of being in an information state as a primitive notion, not to be explained in terms of belief, judgment, and reasons. Because perceptual-information states can be present in a creature that does not think or apply concepts, Evans maintained that the representational content of perceptual states is a kind of nonconceptual content. To be in states with such content (perhaps the nonconceptual content that a sound is coming from direction *d*) a creature does not need to apply, or even to possess, the concepts that we use to specify the content of the states (concepts such as those of sound and direction).

Evans held a distinctive view of the relationship between perceptual-information states and perceptual experiences according to which conscious perceptual experience requires that perceptual-information states should function as inputs to a system for thinking and reasoning. Thus, only a creature with concepts can enjoy perceptual experiences. Nevertheless, a perceiving, thinking, concept-applying creature need not possess all the concepts that would be required fully to specify the content of a perceptual experience and, in having the experience, need not employ even those concepts that are possessed. Evans allowed that the representational content of perceptual experience need not be conceptual content, and in subsequent work, the notion of nonconceptual content has played a major role in accounts of the representational content of perceptual experience (Crane 1992, Gunther 2003, Peacocke 2001).

FURTHER THEMES

Several of Evans's papers—beginning with "Identity and Predication" (1975/1985) and including "Semantic Structure and Logical Form" (1976/1985) and "Does Tense Logic Rest upon a Mistake?" (1985)—contributed to the foundations of semantics and particularly to constraints on semantic theories that show how the meanings of whole sentences depend on the meanings of their parts. In "Semantic Theory and Tacit Knowledge" (1981/1985), he connected the requirement that a semantic theory should reveal semantic structure in sentences with the idea that speakers of a language have tacit knowledge of such a theory. Evans developed a substantive account of tacit knowledge (see also Davies 1987, Peacocke 1989) and distinguished the nonconceptualized content of tacit-knowledge states from the conceptualized content of belief states.

Evans's account of the semantic properties of descriptive names, put forward in "Reference and Contingency" (1979/1985), led to developments in two-dimensional modal logic (Davies and Humberstone 1980; see also Evans 2004), and he made further use of the notion of a singular term with its reference fixed by description in seminal work on pronouns. In "Pronouns, Quantifiers, and Relative Clauses" (1977/1985) and in "Pronouns" (1980/1985), Evans developed an influential account of the semantic function of pronouns that depend for their interpretation on an earlier quantifier phrase yet without being interpretable as bound variables (Neale 1990; King 2005). Finally, "Things without the

Mind" (1980/1985) and "Molyneux's Question" (1985), along with the central chapters of *The Varieties of Reference*, have had a profound influence on subsequent work in philosophy of psychology, particularly work concerningthe perception and representation of space and, more generally, the conditions for an objective conception of a spatial world (Eilan, McCarthy, and Brewer 1993).

See also Davidson, Donald; Dummett, Michael Anthony Eardley; Frege, Gottlob; Kripke, Saul; McDowell, John; Philosophy of Mind; Proper Names and Descriptions; Reference; Russell, Bertrand Arthur William; Semantics; Strawson, Peter Frederick.

Bibliography

WORKS BY EVANS

With John McDowell, eds. *Truth and Meaning: Essays in Semantics*. Oxford, U.K.: Oxford University Press, 1976.

The Varieties of Reference. Oxford, U.K.: Oxford University Press, 1982.

Collected Papers. Oxford, U.K.: Oxford University Press, 1985.

"Comment on 'Two Notions of Necessity.' " *Philosophical Studies* 118 (2004): 11–16.

OTHER WORKS

Burge, Tyler. "Belief *de re*." *Journal of Philosophy* 74 (1977): 338–362.

Crane, Tim, ed. *The Contents of Experience: Essays on Perception*. Cambridge, U.K.: Cambridge University Press, 1992.

Davies, Martin. "Tacit Knowledge and Semantic Theory: Can a Five Per Cent Difference Matter?" *Mind* 96 (1987): 441–462.

Davies, Martin, and I. L. Humberstone. "Two Notions of Necessity." *Philosophical Studies* 38 (1980): 1–30.

Eilan, Naomi, Rosaleen McCarthy, and Bill Brewer, eds. *Spatial Representation: Problems in Philosophy and Psychology*. Oxford, U.K.: Blackwell Publishers, 1993. 2nd ed., Oxford, U.K.: Oxford University Press, 1999.

Gunther, York H., ed. *Essays on Nonconceptual Content*. Cambridge, MA: MIT Press, 2003.

King, Jeffrey C. "Anaphora." In *The Stanford Encyclopedia of Philosophy*, fall 2005 ed., edited by Edward N. Zalta. Available from http://plato.stanford.edu/entries/anaphora/.

Kripke, Saul. "Naming and Necessity." In *Semantics of Natural Languages*, edited by Donald Davidson and Gilbert Harman, 253–355. Dordrecht, Netherlands: D. Reidel, 1972.

Neale, Stephen. *Descriptions*. Cambridge, MA: MIT Press, 1990.

Peacocke, Christopher. "Does Perception Have a Nonconceptual Content?" *Journal of Philosophy* 98 (2001): 239–264.

Peacocke, Christopher. "When Is a Grammar Psychologically Real?" In *Reflections on Chomsky*, edited by Alexander George, 111–130. Oxford, U.K.: Basil Blackwell, 1989.

Sainsbury, R. M. Critical notice: *The Varieties of Reference*, by Gareth Evans. *Mind* 94 (1985): 120–142.

Martin Davies (2005)

EVENTS IN SEMANTIC THEORY

It is an ancient idea that many verbs are used to describe events—things that happen, in places and at times. Frank Plumpton Ramsey introduced an important twist, in the context of distinguishing events from facts. Suppose that Aggie hit Pat. Then on Ramsey's view, the fact reported with (1)

(1) Aggie hit Pat

is the general proposition that there was a hitting of Pat by Aggie. This existential generalization, unlike any event, has no specific spatiotemporal properties. But any event of Aggie hitting Pat verifies (1). So the action report seems to mean that an event of a certain sort occurred. Though at least initially, it is not clear how to square this with the compositionality of linguistic meaning. Let the invented monadic predicate Aghipatish_1 be satisfied by z if and only if (iff) z was a hitting of Pat by Aggie. Then plausibly, (1) is true iff $\exists z[\text{Aghipatish}_1(z)]$. But this biconditional reveals nothing about how the meaning of (1) is determined by the constituent words. And prima facie, the logical form of (1) is $\text{Hit}_2(\mathbf{a}, \mathbf{p})$; where **a** and **p** are names for Aggie and Pat, respectively, and Hit_2 is satisfied by an ordered pair $\langle x, y\rangle$ iff x hit y. Donald Davidson (1967, 1985) shows how to represent the meaning of (1) compositionally and "eventishly," and he offers an argument for doing so. Others develop his proposal and provide independent support for it.

ADVERBIAL MODIFICATION

Let Hit_3 be satisfied by an ordered triple $\langle x, y, z\rangle$ iff z was a hitting of y by x, so that $\forall x\forall y\{\text{Hit}_2(x, y) \leftrightarrow \exists z[\text{Hit}_3(x, y, z)]\}$. Then (1a),

(1a) $\exists z[\text{Hit}_3(\mathbf{a}, \mathbf{p}, z)]$

which is true iff $\exists z[\text{Aghipatish}_1(z)]$, has parts corresponding to the words in (1). If we represent the meaning of (1) with (1a), we can explain the apparent synonymy of (1) with (2):

(2) There was a hitting of Pat by Aggie

But one wants to see further evidence of the alleged covert variable.

An action report can be extended as in (3–8):

(3) Aggie hit Pat softly

(4) Aggie hit Pat with a red stick

(5) Aggie hit Pat in March

(6) Aggie hit Pat with a red stick in March

(7) Aggie hit Pat in March with a red stick

(8) Aggie hit Pat softly with a red stick in March

One might be inclined to represent the meaning of (3) with SoftlyHit$_2$(**a**, **p**); where the invented binary predicate is satisfied by ⟨x, y⟩ iff x hit y softly. But if an inference from (3) to (1) is of the form $\Phi_2(\alpha, \beta)$, so $\psi_2(\alpha, \beta)$, we need some other explanation for why the truth of (3) guarantees the truth of (1). One can stipulate that $\forall x \forall y$[SoftlyHit$_2$(x, y)→ Hit$_2$(x, y)]. However, one wants to know why the corresponding English sentences are related in this fashion. Furthermore, (3) seems to be synonymous with "There was a soft hitting of Pat by Aggie," which implies (2). Such implications are reminiscent of conjunction-reduction, as in inferences like the following: ∃x[Red(x) & Stick(x)], so ∃x[Stick(x)].

This invites Davidson's (1967) hypothesis that the logical form of (3) is (3a);

(3a) ∃z[Hit$_3$(Aggie, Pat, z) & Soft$_1$(z)]

where Soft$_1$(z) means that z was done softly. Furthermore, sentences like (1–8) exhibit a network of entailments. The truth of (4) or (5) also guarantees the truth of (1); (6) implies (7), which implies each of (4–6); and (8) implies each of (1–7). These facts, which illustrate that adverb-reduction is often a valid form of inference in natural language, go unexplained if we represent the meanings of (3–8) with binary predicates like *HitInMarchWithARedStick$_2$*. But we can explain the entailments by analyzing the adverbial modifiers as predicates conjoined with others, as in (7a);

(7a) ∃z[Hit$_3$(Aggie, Pat, z) & In$_2$(z, March) & With$_2$(z, a red stick)]

where In$_2$(z, March) and With$_2$(z, a red stick) mean that z occurred in March, and that z was done with a red stick. By analogy, "There was a red stick on the table touching the chalk" implies that there was a red stick on the table, a stick touching the chalk, a red stick, and so on.

There are also nonimplications to account for. Each of (3–5) could be true, even if (8) is false. Aggie may have hit Pat more than once, but never softly with a red stick in March. Let's suppose, though, that Aggie hit Pat exactly twice: once in March with a red stick, and once in April with a blue stick. Then (9–11) are true, like (4–6), but (12–13) are false.

(9) Aggie hit Pat with a blue stick

(10) Aggie hit Pat in April

(11) Aggie hit Pat with a blue stick in April

(12) Aggie hit Pat with a red stick in April

(13) Aggie hit Pat with a blue stick in March

The truth of (9) and (5) does not guarantee the truth of (13). Nor does the truth of (4) and (10) guarantee the truth of (12). This is what Davidson's (1967) account predicts.

If (4) and (10) are true, there were events z1 and z2, such that Hit$_3$(Aggie, Pat, z1) & With$_2$(z1, a red stick) & Hit$_3$(Aggie, Pat, z2) & In$_2$(z2, April). But it does not follow that there was an event z such that With$_2$(z, a red stick) & In$_2$(z, April). So (12) can be false, and likewise for (13). If we represent the meanings of (1–13) with predicates like *HitWithARedStickInApril$_2$*, we must add stipulations corresponding to the network of implications, and then explain why the English sentences exhibit these implications and not the others. (Note that appealing to times, instead of events, will not account for the facts. Aggie may have hit Pat simultaneously with a red stick softly and with a blue stick sharply.)

OTHER EVIDENCE

We can specify the meaning of (14) with (14a):

(14) Aggie fled after Pat fell

(14a) ∃z{Fled$_2$(**a**, z) & ∃w[(After$_2$(z, w) & Fell$_2$(**p**, **w**)]}

which is true iff an event of Aggie's fleeing occurred after an event Pat's falling. Words like *after* can thus be analyzed as devices for expressing relations between events. Perceptual reports are also relevant. In (15) as opposed to (16):

(15) Pat heard Aggie shout

(16) Pat heard that Aggie shouted

shout is untensed, and replacing *Aggie* with another name for the same individual is sure to preserve truth. Thus, one might render (15) as (15a),

(15a) ∃z∃w[Heard$_3$(**p**, w, z) & Shout$_2$(**a**, w)]

which is true iff there was a hearing by Pat of a shouting by Aggie. This lets us account for the ambiguity of (16), which can imply that the hearing took place in the hall or that the shout did:

(16) Pat heard Aggie shout in the hall

(16a) ∃z∃w[Heard$_3$(**p**, w, z) & Shout$_2$(**a**, w) & In$_2$(the hall, z)]

(16b) $\exists z \exists w[\text{Heard}_3(\mathbf{p}, w, z) \ \& \ \text{Shout}_2(\mathbf{a}, w) \ \& \ \text{In}_2(\text{the hall}, w)]$

Another kind of evidence concerns the intuitive unacceptability of certain adverbial modifications. While (18) sounds somehow wrong, (17) and (19) do not:

(17) Aggie ran for an hour

(18) Aggie ran in an hour

(19) Aggie ran to the store in an hour

If we represent the verb-phrase meanings in (17) and (19) with $RanForAnHour_2$ and $RanToTheStoreInAnHour_2$, not only do we fail to capture implications, we are left wondering why "ran in an hour" is a defective complex monadic predicate. By contrast, (17a) and (19a)

(17a) $\exists z[\text{Ran}_2(\mathbf{a}, z) \ \& \ \text{For}_2(z, \text{an hour})]$

(19a) $\exists z[\text{Ran}_2(\mathbf{a}, z) \ \& \ \text{To}_2(z, \text{the store}) \ \& \ \text{In}_2(z, \text{an hour})]$

suggest that "for an hour," unlike "in an hour," can be part of an event description that does not provide an independent way of saying when the events described are finished. This hypothesis is confirmed by examples like (20–21):

(20) Aggie painted the walls for/in an hour

(21) Aggie painted walls for/in an hour

While (20) is fine with either modifier, (21) is not. Furthermore, with "in an hour" (20) implies an event that ended when the walls in question were covered with paint. With "for an hour" neither sentence implies that Aggie finished painting any wall.

A striking generalization about action reports of the form "Subject Verb Object," with the verb in active voice, is that such reports invariably imply that the subject of the sentence was the actor. We can invent a predicate Tih_2 satisfied by $\langle x, y \rangle$ iff y hit x; and we can imagine a language in which a homophone of (1), with *Aggie* as the subject, means that $\text{Tih}_2(\mathbf{a}, \mathbf{p})$—or equivalently, $\text{Hit}_2(\mathbf{p}, \mathbf{a})$. However, there are no such expressions in natural human languages. And while the source of this fact is a matter of debate, a great deal of evidence suggests a constraint on how grammatical relations are related to thematic relations that hold between events and their participants. In which case, event variables (and thematic relations) are introduced somehow. However, there is more than one way to introduce them.

THEMATIC ELABORATION

Let Hitting_1 be satisfied by z iff z was an event of hitting, ignoring tense for simplicity. Let Agent_2 and Patient_2 signify thematic relations, without worrying here about how to get beyond intuitive specifications of these relations, so that $\forall x \forall y\{\exists z[\text{Hit}_3(x, y, z)]$ iff $\exists z[\text{Agent}_2(z, x) \ \& \ \text{Hitting}_1(z) \ \& \ \text{Patient}_2(z, y)]\}$.

This makes it easy to explain why it follows from (1) that Aggie did something, there was a hitting, and something happened to Pat. This view also preserves a sense in which the transitive verb *hit* is a binary predicate. For while the verb itself is associated with a monadic predicates of events, *hit* is also associated with two thematic relations. Correlatively, one can capture the distinction—independently motivated in many languages that mark nominative and accusative case—between intransitive verbs like *fled* that implicate action, and those that do not. Intuitively, events of falling (like deaths) are things that happen to individuals (not things done), even if such events are intended effects of actions. Besides, one can supplement Davidson's (1967) original proposal—as Davidson (1985) did—with hypotheses like the following:

$$\forall x\{\exists z[\text{Fled}_2(x, z)] \leftrightarrow \{\exists z[\text{Agent}_2(z, x)] \ \& \ \text{Fleeing}_1(z)]\};$$

$$\forall x\{\exists z[\text{Fell}_2(x, z)] \leftrightarrow \{\exists z[\text{Patient}_2(z, x)] \ \& \ \text{Falling}_1(z)]\}.$$

But there are at least two construals of such hypotheses.

One might view them as analyses of multiply unsaturated verb-meanings. From this perspective the verb *hit* is satisfied by ordered triples $\langle x, y, z \rangle$ such that z was a hitting whose Patient was y and whose Agent was x. In which case, given standard assumptions about semantic compositionality, *hit Pat* is satisfied by ordered pairs $\langle x, z \rangle$ such that z was a hitting whose Patient was Pat and whose Agent was x. Less standardly, one might say that *hit* is satisfied by events of hitting (period) and that *hit Pat* is satisfied by each event of hitting whose Patient is Pat. This is, in effect, to adopt the following hypothesis: combining *hit* with a direct object corresponds to predicate-conjunction, not predicate-saturation; and the thematic relation "being the Patient of" is expressed by a certain grammatical relation, between the verb and its object, not simply by the lexical meaning of *hit*.

Barry Schein (1993, 2002) and others argue that considerations involving plurality, along with the need for second-order quantification, favor the second perspective. On this kind of view, "Five boys ate two pizzas" has a (collective) reading according to which there some events of eating whose Agents were five boys, and whose Patients

were two pizzas; where this does not imply that any one event had all five boys and both pizzas as participants. The first view fits more naturally with the following formulation of the collective reading: There was an event whose (plural) Agent was a collection of five boys, and whose Patient was a collection of two pizzas. Still, however one thinks of thematic elaboration, it both extends the scope of event analyses and highlights difficulties.

As many authors have discussed, verbs like *boil* can appear in transitive and intransitive forms, with a characteristic entailment illustrated in (22):

(22) Aggie boiled the soup; so the soup boiled

Treating the two forms as independent predicates, $\text{Boiled}_2(x, y)$ and $\text{Boiled}_1(x)$, makes the implication mysterious. So one might analyze (22) as in (23) or (24):

(23) $\exists z\{\text{Agent}_2(z, \mathbf{a}) \ \&\ \exists w[C_2(z, w) \ \&\ \text{Boiling}_1(w) \ \&\ \text{Patient}_2(w, \mathbf{s})]\ \}$; so $\exists z[\text{Patient}_2(z, \mathbf{s}) \ \&\ \text{Boiling}_1(z)]$

(24) $\exists z\{\text{Agent}_2(z, \mathbf{a}) \ \&\ \exists w[M_2(z, w) \ \&\ \text{Boiling}_1(w)] \ \&\ \text{Patient}_2(z, \mathbf{s})\}$; so $\exists z[\text{Patient}_2(z, \mathbf{s}) \ \&\ \text{Boiling}_1(z)]$

Here, **s** stands for the soup, C_2 indicates a causal relation holding between an action and some of its effects, and M_2 indicates a merelogical relation holding between processes that start with actions and end with effects of those actions. Many linguists argue that some such analysis is required, especially given the constraints on how grammatical relations are mapped to thematic relations. But specifying the requisite causal/mereological relation has proven difficult. Moreover, (23) fails to represent Aggie and the soup as coparticipants in some event describable with an adverbial phrase; yet, if Aggie boiled the soup on Monday, both Aggie's action and the resultant boiling occurred on Monday. And given (24), a background premise is required to reveal the inference as valid: $\forall y \forall z \forall w[M_2(z, w) \ \&\ \text{Patient}_2(z, y) \rightarrow \text{Patient}_2(w, y)]$.

This raises hard questions about the individuation of actions and their relation to thematic relations. More generally, it is unclear what event variables range over, given the kinds of considerations that motivate such variables. Suppose that Aggie drank a pint of beer (and nothing else) in ten minutes. Then for those ten minutes Aggie drank beer. Let z1 be this event of Aggie drinking beer, and let z2 be the event of Aggie drinking the pint in question. Intuitively, z1 is z2; Aggie's beer drinking was none other than the drinking of that pint. In which case, z1 satisfies "in ten minutes" iff z2 does; and z2 does. However, if z1 satisfies "in ten minutes," why is "Aggie drank beer in ten minutes" anomalous? This kind of question arises often.

Consider two billiard balls, **b** and **c**, that came into contact exactly once. At that moment, **b** touched **c**, and **c** touched **b**. Perhaps *touched*, used in this way, does not mark its subject as an Agent, but letting Sub_2 and Ob_2 signify the relevant thematic relations, whatever they are: $\exists z[\text{Sub}_2(z, \mathbf{b}) \ \&\ \text{Touching}_1(z) \ \&\ \text{Ob}_2(z, \mathbf{c})]$, and $\exists z[\text{Sub}_2(\mathbf{c}, z) \ \&\ \text{Touching}_1(z) \ \&\ \text{Ob}_2(\mathbf{b}, z)]$. One might have expected the touching of **c** by **b** to be identical with the touching of **b** by **c**. But how can any one event of touching, z, be such that: $\text{Sub}_2(z, \mathbf{b}) \ \&\ \text{Ob}_2(z, \mathbf{c}) \ \&\ \text{Sub}_2(z, \mathbf{c}) \ \&\ \text{Ob}_2(z, \mathbf{b})$? Presumably, $\text{Sub}_2(z, \mathbf{b})$ implies that **b** is the unique individual that bears the relevant thematic relation to z—and likewise for $\text{Sub}_2(z, \mathbf{c})$—since Aggie touched/lifted Pat does not mean merely that there was a touching/lifting with Aggie as one of potentially many touchers/lifters, and Pat as one of potentially many things touched/lifted.

One can avoid the false implication that $\mathbf{b} = \mathbf{c}$ by denying that event variables in semantic theories range over language-independent occurrences. However, this has implications for the relations among meaning, truth, and ontology. Another option is to elaborate further, treating notation like $\text{Sub}_2(z, \mathbf{b})$ as shorthand for claims of the form: $\exists e[R_2(e, z) \ \&\ P_2(e, \mathbf{b})]$; where e ranges over language-independent spatiotemporal particulars individuated (at least roughly) in accordance with intuitions about events, R_2 signifies a relation that holds between such particulars and their grammatical presentations, and P_2 signifies a suitable participation relation. But if z ranges over things individuated partly in terms of the grammatical relations that verbs bear to their arguments, we still face questions about the individuation of events and their relation to any thematic relations appealed to in theories of meaning.

See also Event Theory; Semantics.

Bibliography

Davidson, Donald. "Adverbs of Action." In *Essays on Davidson: Actions and Events*, edited by Bruce Vermazen and Merrill B. Hintikka. Oxford, U.K.: Clarendon Press, 1985.

Davidson, Donald. "The Logical Form of Action Sentences." In *The Logic of Decision and Action*, edited by Nicholas Rescher. Pittsburgh: University of Pittsburgh Press, 1966.

Fodor, Jerry. "Three Reasons for Not Deriving 'Kill' from 'Cause to Die.'" *Linguistic Inquiry* 1 (1970): 429–438.

Higginbotham, James. "The Logical form of Perceptual Reports." *Journal of Philosophy* 80 (1983): 100–127.

Higginbotham, James, Fabio Pianesi, and Achille C. Varzi, eds. *Speaking of Events*. New York: Oxford University Press, 2000.

Landman, Fred. "Plurality." In *The Handbook of Contemporary Semantic Theory*, edited by Shalom. Lappin. Oxford, U.K.: Blackwell, 1996.
LePore, Ernest, and Brian P. McLaughlin, eds. *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford, U.K.: Blackwell, 1985.
Parsons, Terry. *Events in the Semantics of English*. Cambridge, MA: MIT Press, 1990.
Pietroski, Paul. *Events and Semantic Architecture*. New York: Oxford University Press, 2005.
Ramsey, Frank. "Facts and Propositions." *Aristotelian Society* (Supplement) 6 (1927): 153–170.
Schein, Barry. "Events and the Semantic Content of Thematic Relations." In *Logical Form and Language*, edited by Gerhard Preyer and Georg Peter. New York: Oxford University Press, 2002.
Schein, Barry. *Plurals*. Cambridge, MA: MIT Press, 1993.
Taylor, Barry. *Modes of Occurrence*. Oxford, U.K.: Blackwell, 1985.
Tenny, Carol. *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht, Netherlands: Kluwer Academic, 1994.
Thomson, Judith. *Acts and Other Events*. Ithaca, NY: Cornell University Press, 1977.
Vendler, Zeno. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press, 1967.

Paul M. Pietroski (2005)

EVENT THEORY

An event is anything that happens, an occurrence. The idea of an event began to take on a philosophical life of its own in the twentieth century, due to a reawakening of interest in the concept of change, to which the concept of an event seems inextricably tied, and to the growing use of the concept of an event in scientific and metascientific writing (see Broad 1933, McTaggart 1927, and Whitehead 1929). Interest in events has also been sparked by versions of the mind-body identity thesis formulated in terms of events and by the idea that a clearer picture of events will facilitate discussion of other philosophical issues.

Discussions of events have focused on whether there are events and, if so, what the nature of events is. Since whether there are events depends in part on what they would be like if there were any, the two issues have usually been treated together.

Some philosophers (e.g., J. J. Thomson) simply assume that there are events; others argue for that assumption. Donald Davidson has asserted that there are events (and actions) by arguing that, to explain the meanings of claims involving adverbial modifiers (e.g., "Jones killed Smith in the kitchen") and singular causal claims (e.g., "the short circuit caused the fire"), we should suppose that such claims implicitly quantify over, or posit, actions and events (e.g., killings, short circuits, and fires). Opponents of Davidson's analyses (e.g., Terence Horgan) have argued that alternative semantic theories, which do not posit events, are able to explain the semantic features of Davidson's target sentences.

While some singular terms purporting to refer to events are proper names (e.g., "World War I"), many are definite descriptions (e.g., "the killing of Caesar by Brutus"). The semantics of singular descriptions for events has been studied by Zeno Vendler and Jonathan Bennett. Of particular interest is the distinction between perfect nominals, such as "Quisling's betraying of Norway," which refer to events (or actions or states), and imperfect nominals, such as "Quisling's betraying Norway," which refer to factlike entities. Bennett has argued that much of what is wrong in Jaegwon Kim's theory of events can be traced to confusions involving these two sorts of nominals and to expressions (e.g., "the betrayal") that are ambiguous and can refer either to events or to facts.

Most philosophers take events to be abstract particulars: particulars in that they are nonrepeatable and spatially locatable, abstract in that more than one event can occur simultaneously in the same place. Some philosophers who think this way (e.g., Lawrence Brian Lombard) take events to be the changes that objects undergo when they alter. (Others, such as Bennett, have doubts about this; others, such as Kim and David Lewis, deny it outright.) Thus, the time at which an event occurs is the (shortest) time at which the subject of that event changes from the having of one to the having of another, contrary property. Since no object can have both a property and one of its contraries simultaneously, there can be no instantaneous events.

Events inherit their spatial locations from the spatial locations, if any, of the things in which those events are changes. Events do not get their spatial locations by occupying them; if they did, then distinct events, like distinct physical objects, could not occur in the same place simultaneously. But more than one event apparently can occur at the same time and place. However, some philosophers (e.g., W. V. O. Quine) hold that events are concrete and that events and physical objects do not belong to distinct metaphysical kinds.

Though it seems clear that some events are composed of others, it is not clear what the principles are that determine when events compose more complex events. Some views of events (perhaps A. N. Whitehead's) seem compatible with there being subjectless events, events that are not changes in anything whatsoever. However, subjectless events could not be changes, for it seems

absurd to suppose that there could be a change that was not a change in or of anything.

Theories about the nature of entities belonging to some metaphysically interesting kind must address the issue of what properties such entities essentially have. In the case of events, the issue is made pressing by the fact that certain theories concerning causation (e.g., Lewis's) require that judgments be made about whether certain events would occur under certain, counterfactual circumstances.

In the literature on events, attention has been given to four essentialist issues. The first is whether the causes (or effects) of events are essential to the events that have them; Peter van Inwagen has suggested that an event's causes are essential to the events that have them, while Lombard has argued that neither the causes nor the effects of events are essential to them. The second concerns the subjects of events; Bennett and Lewis suggest that the subjects of events are not essential, while Lombard and Kim argue that they are. The third is whether an event's time of occurrence is essential to it. Lombard has argued in favor of this essentialist claim, while Bennett and Lewis have argued against it. And the fourth is whether it is essential that each event be a change with respect to the properties to which it is in fact a change. Though the first three issues have received some attention, the fourth has attracted the most, due to the prominence given to debates between the defenders and opponents of Kim's and Davidson's views on the identity of events.

Theories about events typically contain, as a chief component, a "criterion of identity," a principle giving necessary and sufficient conditions for an event e and an event e' to be identical. Though there is no general agreement on this, such a principle is sought because, when it satisfies certain constraints, it becomes a vehicle for articulating a view about what it is to be an event and how events are related to objects belonging to other kinds.

Quine holds that events are the temporal parts of physical objects and thus that events and physical objects share the same condition of identity: sameness of spatiotemporal location. Kim's interest in events centers in part on the idea that they are the objects of empirical explanations. Since what is typically explained is an object's having a property at a certain time, Kim takes an event to be the exemplification of a property (or relation) by an object (or objects) at a time. This idea, combined with some others, led him to hold that an event e is the same as an event e' if and only if e and e' are the exemplifications of the same property by the same object(s) at the same time. Kim's view has been criticized, principally by Lombard and Bennett, on the grounds that what it says about events is more plausibly seen as truths about facts. Kim's view has also been criticized by those whose intuitions concerning the identity of events more closely match Davidson's.

Davidson once proposed that events, being essentially the links in causal chains, are identical just in case they have the same causes and effects. He has since abandoned this position in favor of Quine's.

Another view that places causation at the heart of the idea of an event is due to Lewis, who has tried to construct a theory in which events have just those features that would allow them to fit neatly into his counterfactual analysis of causation. In some respects, Lewis's view is like Myles Brand's in that both are moved in part by the idea that more than one event can occur simultaneously in the same place. Lewis takes an event to be a property-in-intension of a spatiotemporal region, so that events that in fact occur simultaneously in the same place but could have had different spatiotemporal locations are distinct.

Bennett thinks that the concept of an event is not precise enough to withstand much critical examination on its own and that events should be thought to be (only) whatever they need to be in order to make constructive use of them in the discussion of other philosophical issues. Like Lewis, Bennett takes an event to be a property; but, for Bennett, the property seems to be a property-in-extension and is a particular. That is, Bennett thinks that events are tropes.

Lombard's view is, like Kim's, a variation on a property exemplification account. Lombard's version is derived from the idea of events as the changes that objects undergo when they alter, and it takes events to be the exemplifyings of "dynamic" properties at intervals of time. Such alterations are the "movements" by objects from the having of one to the having of another property through densely populated quality spaces, where each quality space is a class of contrary properties, the mere having of any member of which by an object does not imply change.

See also Bennett, Jonathan; Davidson, Donald; Kim, Jaegwon; Lewis, David; Metaphysics; Mind-Body Problem; Quine, Willard Van Orman; Thomson, Judith Jarvis; Whitehead, Alfred North.

Bibliography

Bennett, J. *Events and Their Names.* Indianapolis: Hackett, 1988.

Brand, M. "Particulars, Events, and Actions." In *Action Theory,* edited by M. Brand and D. Walton. Dordrecht: Reidel, 1976.

Broad, C. D. *An Examination of McTaggart's Philosophy.* Cambridge, U.K.: Cambridge University Press, 1933.

Davidson, D. *Essays on Actions and Events.* New York: Oxford University Press, 1980.

Davidson, D. "Reply to Quine on Events." In *Actions and Events: Perspectives on the Philosophy of Donald Davidson,* edited by E. LePore and B. P. McLaughlin. Oxford: Blackwell, 1985.

Horgan, T. "The Case against Events." *Philosophical Review* 87 (1978): 28–47.

Kim, J. "Events and Their Descriptions: Some Considerations." In *Essays in Honor of Carl G. Hempel,* edited by N. Rescher et al. Dordrecht: Reidel, 1970.

Kim, J. *Supervenience and Mind,* Chaps. 1, 3. New York: Cambridge University Press, 1993.

Lewis, D. "Events." In *Philosophical Papers,* Vol. 2. New York: Oxford University Press, 1986.

Lombard, L. B. *Events: A Metaphysical Study.* London: Routledge and Kegan Paul, 1986.

McTaggart, J. M. E. *The Nature of Existence,* Vol. 2. Cambridge, U.K.: Cambridge University Press, 1927.

Quine, W. V. O. "Events and Reification." In *Actions and Events: Perspectives on the Philosophy of Donald Davidson,* edited by E. LePore and B. P. McLaughlin. Oxford: Blackwell, 1985.

Quine, W. V. O. "Things and Their Place in Theories." In *Theories and Things.* Cambridge, MA: Harvard University Press, 1981.

Thomson, J. J. *Acts and Other Events.* Ithaca, NY: Cornell University Press, 1977.

van Inwagen, P. "Ability and Responsibility." *Philosophical Review* 87 (1978): 201–224, esp. 207–209.

Vendler, Z. "Facts and Events." In *Linguistics in Philosophy.* Ithaca, NY: Cornell University Press, 1967.

Whitehead, A. N. *An Enquiry concerning the Principles of Natural Knowledge.* Cambridge, U.K.: Cambridge University Press, 1919.

Whitehead, A. N. *Process and Reality.* Cambridge, U.K.: Cambridge University Press, 1929.

Lawrence Brian Lombard (1996)

EVIDENTIALISM

"Evidentialism" is the view about epistemic justification that identifies the extent to which a person is justified in believing a proposition with the extent to which the evidence the person has supports the truth of the proposition. Other doxastic attitudes such as withholding judgment and denying are also justified by the character of the person's evidence.

A full-scale evidentialist theory would explain what constitutes evidence, what it means to have a certain body of evidence, and what it means for a body of evidence to support a proposition to any given extent. Ordinarily, people count as evidence external things such as fingerprints and bank records. However, according to evidentialists, our fundamental evidence is constituted by our perceptual experiences, our apparent memories, and other mental states. A full-scale theory requires an account of what we have as this ultimate sort of evidence: It is unclear, for example, whether someone's unactivated memories are part of the person's current evidence. The evidential support relation to which evidentialists appeal is not a familiar logical relation. Perceptual states can support beliefs about the external world, yet there is no familiar logical relation between those states and the beliefs they support. Furthermore, one's evidence on its own does not support its distant and unnoticed logical consequences. A complete evidentialist theory would clarify the justifying connection between a body of evidence and a proposition.

Leading skeptical controversies are usefully understood to concern what sort of evidence is required for knowledge. For example, if knowledge requires complete epistemic justification, and this requires having entailing evidence, then skeptics can cogently argue that we have no such evidence for any empirical proposition and that therefore we have no empirical knowledge. On the other hand, standard skeptical arguments fail if nonentailing evidence can completely justify belief. An evidentialist theory can resolve this dispute either way.

Diverse theories of justification can be understood as evidentialist views that differ on the nature of evidence, its possession, and how it supports belief. For instance, a typical coherentist theory in effect holds that a person has her beliefs as evidence and that support by evidence consists in coherence with it. A typical foundationalist theory in effect holds that justified beliefs must include some that are defended by a foundational sort of evidence—for example, by perceptual states—and that this evidence is had by the person by being consciously accessible.

Evidentialism entirely discounts factors that figure centrally in some theories of justified belief. These factors include the intellectual pedigree of the belief, the believer's capacity or intention to fulfill intellectual duties or to exemplify cognitive virtues, and the normal functioning of the operative belief-forming mechanism. Justifying evidence for a belief might happen to arise in an irresponsibly haphazard inquiry with no attempt to fulfill any epistemic duty, as a fluke result of some abnormal cognitive activity lacking in intellectual virtue. The evidentialist view is that regardless of all this, belief is justified because the evidence possessed supports the proposition.

See also Classical Foundationalism; Coherentism; Epistemology; Skepticism, Contemporary; Skepticism, History of.

Bibliography

Chisholm, R. "A Version of Foundationalism." In *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982.

Feldman, Richard, and Earl Conee. "Evidentialism." *Philosophical Studies* 48 (1985): 15–34.

Feldman, Richard, and Earl Conee. *Evidentialism*. Oxford: Oxford University Press, 2004.

Goldman, Alvin. *Epistemology and Cognition*, 87–93. Cambridge, MA: Harvard University Press, 1986.

Haack, Susan. *Evidence and Inquiry*. Oxford: Blackwell, 1993.

Moser, Paul. *Knowledge and Evidence*. Cambridge, U.K.: Cambridge University Press, 1989.

Plantinga, A. *Warrant and Proper Function*, 185–193. Oxford: Oxford University Press, 1993.

Earl Conee (1996)
Richard Feldman (1996)
Bibliography updated by Benjamin Fiedor (2005)

EVIL

For most twentieth-century philosophers, intent on dividing philosophy into discrete subdivisions, the problem of evil was a matter for philosophical theology, or—more rarely—for ethics and moral psychology. The theological question is as easy to state as it is hard to answer: How is a world full of evil and suffering compatible with the existence of an omnipotent, benevolent creator? The ethical question is altogether different: How can rational beings commit evil acts? The first question has preoccupied theists since *The Book of Job;* the difficulty of finding a satisfactory answer has served many as a reason for rejecting theism. The second question has been answered in some religious traditions by the appeal to original sin, but in recent years more it has often been viewed as outside the focus of traditional philosophy. The history of modern philosophy reveals that the problems are related, and part of a larger set of questions that precedes both: Can we make sense of the lives we are given? Does human reason have the ability to find or make the world intelligible? These are not questions that are driven primarily by theological or ethical concerns, but that drive those concerns, and arguably philosophy itself.

Aristotle claimed that philosophy begins in wonder; for Schopenhauer the main subject of this wonder is the world's evil and wickedness. Even if misery were visited only on the wicked or completely outweighed by goodness, alone might well question why it should exist at all. Idle curiosity alone might inspire such questioning about why things are as they are in general; but that questioning is likely to become urgent when things go wrong. If the principle of sufficient reason is the claim that nothing happens without a reason, then there are two choices: to seek an explanation for the evils in the world or to abandon the principle of sufficient reason itself.

Orthodox thinkers have often taken the latter route, maintaining that belief is not only a matter of faith but of absurd faith. Pierre Bayle (1647–1706), the French philosopher known as "the arsenal of the Enlightenment," took this view to its logical conclusion. He thought that Manicheism, the belief that the universe is controlled by equally powerful good and evil forces, is the scientific explanation that best conforms to the data. Insofar as faith prescribes monotheism, however, Manicheism is precluded—along with any attempt at scientific explanation altogether. After all, Bayle concludes, the new Cartesian philosophy teaches that properties like color are only secondary to mathematical properties, which we do not perceive but infer. With this great a gap between experience and scientific explanation, why take the latter seriously at all?

Bayle's "theory of the incomprehensibility of all things" was the target of Leibniz's *Theodicy* (published 1710). Concerned to reconcile faith and science by proving that both were based on the principle of sufficient reason, Leibniz argued that anguish over God's seeming tolerance for evil resulted from ignorance of His ways. Ptolemaic astronomy did not challenge the work of the divine creator but rather that of the early cosmologists. Similarly, Leibniz promises, later scientific discoveries will reveal our discontent with the world to be a function of our ignorance. God has reasons we do not yet understand, but if our knowledge were as infinite as the creator's, we too could recognize this world as the best possible one.

Voltaire's *Candide* ridiculed such arguments, and Hume's *Dialogues concerning Natural Religion* demolished them. The later Kant found Leibnizian attempts to verge on blasphemy and thought the appeal to God's unknown reasons to be a mockery of suffering that required "no refutation but the abomination of anyone with the least feeling for morality" (1968). What unites their rejection of Leibniz's theodicy is the rejection of a metaphysical tradition extending back to Plato. For this rationalist metaphysics, the appearances we see appear to reflect evil and corruption; the reality behind them is truer and better than what we experience. Against this

view, the more empirical outlook of Voltaire, Hume, and Kant can be seen as a moral imperative, for it implies keeping faith with the world's victims by acknowledging the reality of their suffering. But can that reality be acknowledged without entirely capitulating to it? Is it possible to maintain that evil is not essential to the world but rather an unhappy accident?

Jean-Jacques Rousseau (1712–1778) adopted just such an approach, and Kant found his solution so extraordinary that he called him "the Newton of the mind" (1942). Rousseau substituted the idea of history for the idea of original sin that had seeped into the most sober discussions of evil. Humankind was created morally neutral; a series of accidents based on minor instances of vanity and greed cascaded into the moral deterioration known as civilized society. At any particular point, we might have stopped the process, which has now gone so far that only the most radical measures hold promise of salvation. *The Discourse on Inequality* is Rousseau's diagnosis of evil, and his own replacement for the myth of the Fall; *Emile* is his recipe for a cure and the hope of salvation. The mixture of self-help manual and sacred text, science and literature is crucial: Rousseau argues that the problem of evil was so deep that it could only be approached on all fronts. Different forms of pedagogy, arts, political organization, religion, and metaphysics are all required to respond to it. Small wonder that Rousseau's plans for simultaneously reshaping individual human beings and their societies spurred the French Revolution.

The worry fueling debates about the difference between appearance and reality was not the fear that the world might be different than it seems, but rather the fear it might *not*. By acknowledging the reality of evil while maintaining that reality could be changed, Rousseau dislodged the problem of evil from the theological context in which it had been embedded.

That context is exemplified in Christian Wolff's work, which still divides evils into metaphysical, natural, and moral evils. The first evil was the imperfection in the substance(s) of which the world is made; the second was the suffering we experience through earthquakes, floods, plagues, and the like; the third was the cruelty and injustice we visit upon each other. After the mid-eighteenth century the two former evils were viewed as natural limits and natural catastrophes, devoid of significance. The only remaining evil is the moral evil committed by intentionally acting human beings. This absolves God of responsibility for evil while turning our attention to questions of ethics, psychology, history, education, and economics. With these issues in the forefront, nineteenth-century philosophy carried on the discussion of evil. While theistic discourse receded ever farther to the margins, modern thinkers remained preoccupied with the meaning of life and the intelligibility of a world full of evil. This was true not only for philosophers sometimes considered peripheral to the canon (for example, Rousseau, Voltaire, and Schopenhauer) but also for those central to it (for example, Spinoza, Leibniz, Hume, Kant, and Hegel). Nor it is a matter of national heritage: the sober Briton John Stuart Mill could write about the problem in terms almost as vehement as Nietzsche's.

The problem of evil was no longer central in twentieth century philosophy, but it persisted in different forms, retaining the bond between ethics and metaphysics. No one would take up Hegel's project of "theodicy, a justification of the ways of God (such as Leibniz attempted in his own metaphysical manner, but using categories which were as yet abstract and indeterminate)" (1975, p. 43). In the wake of Auschwitz, every form of theodicy was viewed with suspicion. But thanks to the work of two very different twentieth-century philosophers, the problem of evil remained a major concern of philosophy. Hannah Arendt's *Eichmann in Jerusalem: A Report on the Banality of Evil* was criticized for deflating the gravity of Eichmann's crimes by calling them banal. In fact she sought to justify a world in which criminals like Eichmann are possible, by showing they are not the result of deep or demonic impulses but of mindless and not entirely intentional behavior. With this work Arendt takes up a project going back to Rousseau: Explaining the existence of evil allows us to show that it does not belong to the essence of the world, and that it may be at least partially eradicable. In thereby reducing the role of intention in moral evil, Arendt challenged a conception of evil that had dominated modern thought.

The other twentieth-century philosopher in question, John Rawls, is known for his insistence that political philosophy is independent of metaphysics, but in later works and conversations he made clear that the problem of evil was a major concern behind his work. The author of the first major English book of substantive ethics since Mill, Rawls wrote in response to two metaphysical and moral problems that ground the problem of evil: the problem of contingency and the problem of reconciliation. In *Justice as Fairness* he invokes Hegel to stress political philosophy's role in providing reconciliation. Rawls's goal is to show that a realistic utopia, in which greatest evils are eliminated, is possible; without that hope "one might reasonably ask, with Kant, whether it is worthwhile

to live on earth" (1999, p. 128). Such passages encourage renewed attention to his main work, *A Theory of Justice*, the two principles of which show that human beings need not resign themselves to fate but can meet "the arbitrariness of fortune" with measures of their own (p. 102). If the problem of evil is evident in the work of two such different contemporary philosophers, it is likely to occupy major philosophers, whether theist or not, for the foreseeable future.

See also Arendt, Hannah; Aristotle; Bayle, Pierre; Cartesianism; Enlightenment; Evil, The Problem of; Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Mani and Manichaeism; Mill, John Stuart; Plato; Rawls, John; Rousseau, Jean-Jacques; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Virtue and Vice; Voltaire, François-Marie Arouet de; Wolff, Christian.

Bibliography

Arendt, Hannah. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Viking, 1963

Bayle, Pierre. *Historical and Critical Dictionary*. Translated by J. J. Knapton and P. Knapton, 1738. Reprint, New York: Garland, 1984.

Hegel, Georg Wilhelm Friedrich. *Introduction to the Lectures on the Philosophy of World History*. Translated by H. B. Nisbet. Cambridge, U.K.: Cambridge University Press, 1975.

Hume, David. *Dialogues concerning Natural Religion*, edited by J. C. A. Gaskin. Oxford: Oxford University Press, 1993.

Leibniz, Gottfried Wilhelm. *Theodicy*, edited by Austin Farrar. Translated by E. M. Huggard. La Salle: Open Court, 1985.

Kant, Immanuel. *Reflexionen*. *Kants Gesammelten Werken*, Band 20. Berlin: Akademie Ausgabe 1942.

Kant, Immanuel. "Über das Mißlingen aller philosophischen Versuche in der Theodizee." In *Werke*, Band 8, edited by Wilhelm Weischedel. Frankfurt/Main: Shirkamp, 1968.

Neiman, Susan. *Evil in Modern Thought: An Alternative History of Philosophy*. Princeton, NJ: Princeton University Press, 2002.

Rawls, John. *Justice as Fairness*. Cambridge, MA: Harvard University Press, 2000.

Rawls, John. *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Richardson, Henry, and Paul Weithman, eds. *The Philosophy of John Rawls*. New York: Garland, 1999.

Rousseau, Jean-Jacques. *Discourse concerning the Origins of Inequality*, edited and translated by Victor Gourevitch. New York: Harper and Row, 1986.

Rousseau, Jean-Jacques. *Emile*. Translated by Allan Bloom. New York: Basic Books, 1979.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne, Dover, 1966.

Voltaire, Francois-Marie Arouet. *Candide, Zadig, and Selected Stories*. Translated by Donald M. Frame. London: Penguin, 1981.

Susan Neiman (2005)

EVIL, THE PROBLEM OF

The problem of evil concerns the contradiction, or apparent contradiction, between the reality of evil on the one hand, and religious beliefs in the goodness and power of God or of the Ultimate on the other. In a very general classification, the religions of the world have offered three main types of solution: (1) There is the monism of the Vedanta teachings of Hinduism, according to which the phenomenal world, with all its evils, is *maya,* or illusion. A confused echo of this doctrine is heard in contemporary Western Christian Science, which affirms that "evil is but an illusion, and it has no real basis. Evil is a false belief" (Mary Baker Eddy, *Science and Health,* auth. ed., Boston, 1934, p. 480:23,24). Considered as a response to the problem of evil as stated above, this view is defective in that it redescribes the problem but does not attempt to solve it, for it leaves unexplained the evil of our suffering from the compulsive illusion of evil. (2) There is the dualism exemplified most dramatically in ancient Zoroastrianism, with its opposed good and evil deities, Ahura Mazdah and Angra Mainyu. A much less extreme dualism was propounded by Plato (*Timaeus* 30A and 48A) and is found in various forms in the finite deity doctrines of such modern Western philosophers as J. S. Mill (expounded in "Attributes," *Three Essays in Religion,* London, 1874) and Edgar Brightman (*A Philosophy of Religion,* New York, 1940, Chs. 8–10). (3) There is the distinctive combination of monism and dualism, or of an ethical dualism set within an ultimate metaphysical monism (in the form of monotheism) that has been developed within Christianity and that represents the main contribution of Western thought to the subject.

Since the terms of the problem of evil vary with the character of the religious beliefs which give rise to it, a separate study is required for each of the great religious systems. In the present article, however, the problem will be treated only in the context of the Christian tradition.

Christianity (like Judaism and Islam) is committed to a monotheistic doctrine of God as absolute in goodness and power and as the creator of the universe ex nihilo. The challenge of the fact of evil to this faith has accordingly been formulated as a dilemma: If God is all-powerful, he must be able to prevent evil. If he is all-good,

he must want to prevent evil. But evil exists. Therefore, God is either not all-powerful or not all-good. A theodicy (from *theos,* god, and *dikē,* justice) is accordingly an attempt to reconcile the unlimited goodness of an all-powerful God with the reality of evil.

The kinds of evil distinguished in the literature of theodicy are (1) the evil originated by human beings (and angels), that is, moral evil or sin; (2) the physical sensation of pain and the mental anguish of suffering, which may be caused either by sin or by (3) natural evil, that is, disease, tornado, earthquake, and so forth; and (4) the finitude, contingency, and hence imperfection of all created things which some have called metaphysical evil. The last two topics will be treated in the course of discussing the others in response to the questions: "Why has an infinitely powerful and good God permitted moral evil in his universe?" and "Why has an infinitely powerful and good God permitted pain and suffering in his universe?"

THE PROBLEM OF MORAL EVIL

THE TRADITIONAL AUGUSTINIAN THEODICY. The problem of evil was a lifelong preoccupation of Augustine (354–430), and the main lines of thought that he established have been followed by the majority of subsequent Christian thinkers. Before his conversion to Christianity, Augustine was attracted by Manichaeism, a powerful contemporary religious movement with Eastern and Gnostic roots, which affirmed an ultimate dualism of good and evil in the forms of light, or spirit, and darkness, or matter. In turning from this doctrine to Christianity, Augustine rejected a final dualism in favor of belief in a good God as the sole ultimate reality, and rejected the Manichaean disparagement of matter in favor of an acceptance of the material world as reflecting the goodness of its creator.

But if the sole ultimate power is unambiguously good, what is evil and whence does it come? In answer to this question, Augustine develops two interlocking lines of thought, presenting the *privative* and the *aesthetic* conceptions of evil.

Evil as privation. Augustine counters the Manichaean conception of evil as an independent reality and power coeternal with good by his analysis of evil (derived from Plotinus, *Enneads,* I, Eighth Tractate) as the privation, corruption, or perversion of something good. Evil, he taught, has no independent existence, but is always parasitic upon good, which alone has substantival being. "Nothing evil exists in itself, but only as an evil aspect of some actual entity" (*Enchiridion,* Ch. 4). Thus, everything that God has created is good, and the phenomenon of evil occurs only when beings which are intrinsically good (though mutable) become corrupted and spoiled.

Augustine expresses the same thought from another perspective when he equates being with goodness. God, as the highest, richest, and most intensely real being, is the supreme good, and everything that he has brought into existence is ipso facto good. For this reason the corruption that we call evil can never be complete; for if a thing becomes so vitiated in nature that it ceases to exist, the evil which is parasitic upon it must also cease to exist. Hence, there can be no wholly evil being.

How does this spoiling of God's initially good creation come about? Augustine's answer is that evil has entered into the universe through the culpable volitions of free creatures, angels and humans. Their sin consisted, not in choosing positive evil (for there is no positive evil to choose), but in turning away from the higher good, namely God, to a lower good. "For when the will abandons what is above itself, and turns to what is lower, it becomes evil—not because that is evil to which it turns, but because the turning itself is wicked" (*City of God,* XII, 6). Augustine holds that natural evils, such as disease, are divinely ordained consequences of the primeval fall of man, and thus traces all evils either directly or indirectly to a wicked misuse of creaturely freedom: "There are two kinds of evil, sin and the penalty for sin" (*Against Fortunatus,* 15).

When we ask what caused man to fall, Augustine's answer is his doctrine of deficient causation. There is no efficient, or positive, cause of evil willing. Rather, evil willing is itself a negation or deficiency, and to seek for its cause "is as if one sought to see darkness, or hear silence" (*City of God,* XII, 7). Perhaps the best way to interpret this obscure teaching is as an assertion of the inexplicability, in principle, of free volitions; for "what cause of willing can there be which is prior to willing?" (*Free Will,* III, xvii, 49). Augustine is saying that the origin of moral evil lies hidden within the mystery of human and angelic freedom. The freely acting will is an originating cause, and its operations are not explicable in terms of other prior causes.

Aesthetic conception of evil. The other main theme in Augustine is the aesthetic conception of evil, which is also derived from Plotinus (*Enneads,* III, 2, 17). According to this view, what appears to be evil, when seen in isolation or in a too limited context, is a necessary element in a universe that, viewed as a totality, is wholly good. From the viewpoint of God, who sees timelessly and as a whole the entire moving panorama of created history, the uni-

verse is good: "To thee there is no such thing as evil, and even in thy whole creation, taken as a whole, there is not" (*Confessions,* VII, 13).

The presupposition of this aesthetic view is the ancient conception, deriving from Plato (*Timaeus,* 41 B–C), which Arthur Lovejoy has called the principle of plenitude (*The Great Chain of Being,* Cambridge, MA, 1936). According to this principle, a universe in which all the varied potentialities of being are realized and which contains as many different kinds of entity as are possible (lower as well as higher, lesser as well as greater), is a better universe—one which more adequately expresses the infinite creativity of God—than would a universe which contains only the highest type of created beings. There is thus an immense hierarchy of forms of created existence, and each creature, in its own proper place in the scheme of nature, is good and glorifies its Maker. Those that are lower in the scale of being are not on that account evil; they are just different goods, contributing in their different ways to the perfection of the universe. Again, things that are transitory by nature, appearing and then perishing within the ever-changing pattern of nature's beauty, contribute, even by their death, to the perfection of the created order. As a very minor subtheme within this aesthetic conception, Augustine sometimes also uses the notion of evil as providing a contrast by which good shines the more brightly.

As an application of the principle of plenitude, Augustine holds that the universe must contain mutable and corruptible creatures, compounded of being and nonbeing. It is better that the universe should include free beings who may, and do, fall than that it should omit them. Thus, Augustine brings even moral evil within the scope of his aesthetic conception. In doing so, he employs the further principle (later invoked by Anselm, in his atonement theory) that as long as sin is exactly balanced by just punishment, it does not upset the moral harmony of the universe. "Since there is happiness for those who do not sin, the universe is perfect; and it is no less perfect because there is misery for sinners… . So, whatever a soul may choose, ever beautiful and well-ordered in all its parts is the universe whose Maker and Governor is God" (*Free Will,* III, 9, 26–27).

Influence of the Augustinian analyses. Both of these main Augustinian themes reappear in the thought of Thomas Aquinas in the thirteenth century (*Summa Theologiae,* I, 47–49).

Martin Luther and John Calvin, the Reformers of the sixteenth century, were not interested in developing a general theodicy, although they followed Augustine in his doctrine that all the evils of human life flow ultimately from the culpable fall of man.

Gottfried Wilhelm Leibniz, in his *Théodicée* (1710), employed the two Augustinian themes, the privative and aesthetic conceptions, in the course of his argument that this is the best of all possible worlds (or, more strictly, the best of all possible universes, for he uses "world" in its most comprehensive sense)—a notion pointedly satirized in Voltaire's *Candide* (1759). It is the best, not because it contains no evil, but because any other possible universe would contain more evil. The eternal possibilities of existence are individually present to the Divine Mind which, like an infallible calculating machine, surveys all possible combinations and selects the best, to which it then gives existence.

SUMMARY AND CRITICISM OF THE AUGUSTINIAN THEODICY. The traditional Augustinian theodicy in respect to moral evil asserts that God created man with no sin in him and set him in a world devoid of evil. But man willfully misused his God-given freedom and fell into sin. Some men will be redeemed by God's grace, and others will be condemned to eternal punishment. In all this, God's goodness and justice alike are manifested.

This traditional theodicy has been criticized for its accounts of the origin and of the final disposition of moral evil.

The origin of moral evil. It is urged that the notion of finitely perfect beings willfully falling into sin is self-contradictory and unintelligible (cf. Friedrich Schleiermacher, *Der christliche Glaube,* 2nd ed., Berlin, 1830–1831, par. 72). A truly perfect being, though free to sin, would in fact never do so. To attribute the origin of evil to the willful crime of a perfect being is thus to assert the sheer contradiction that evil has created itself ex nihilo.

There appears, further, to be a disharmony between this theodicy and Augustine's doctrine of predestination, which in effect sets the origin of moral evil within the purpose and responsibility of God. Augustine's doctrine (*City of God,* XI, 11 and 13, and XII, 9) refers to the fall of the angels. Calvin (*Institutes,* III, 23, 7 and 8,) has a parallel doctrine referring to the fall of man.

The assumption of the traditional theodicy that it is logically impossible for God to have created humans such that they would always freely make right moral choices has recently been attacked under the name of "the free will defense." Defining a free action as one that flows from the nature of the agent, without external compulsion, recent writers have claimed that, without logical contra-

diction, God might initially have given men a nature that would always freely express itself in right actions. (See J. L. Mackie, "God and Omnipotence," *Mind* [April 1955], and Antony Flew, "Divine Omnipotence and Human Freedom," *New Essays in Philosophical Theology,* 1955. For an important critical comment on these two articles, see Ninian Smart, "Omnipotence, Evil and Supermen," *Philosophy* [April 1961], and replies in the same journal by Flew [January 1962], and Mackie [April 1962].) Three of the questions involved in this debate are: (1) In denying that we do what we do because we are what we are, and that therefore we might have been made so that we would always freely act rightly, can we avoid equating free behavior with merely random behavior? (2) Is there any important difference between a good will that has been created ready-made as such, and one which has become steadfastly good as the outcome of a history of moral endeavor and struggle? (3) If God's primary purpose for humans is to evoke in them a free and uncompelled love and trust in relation to himself, would this purpose be frustrated by his creating people so that they cannot do other than make this response?

The final disposition of moral evil. The criticism of the eschatological aspect of the traditional Augustinian (and also, on this point, the Calvinist) theodicy has been expressed as a dilemma. If God desires to save all his human creatures, but is unable to do so, he is limited in power. If, on the other hand, he does not desire the salvation of all, but has created some for damnation, he is limited in goodness. In either case, the doctrine of eternal damnation stands as an obstacle in the way of Christian theodicy.

THE IRENAEAN TYPE OF THEODICY. Prior to Augustine and the development of his theology by the Latin fathers of the Christian church, a significantly different conception of the fall of man was prevalent among many of the Greek-speaking fathers, chief among them Irenaeus (c. 120–202). Whereas Augustine held that before his fall, Adam was in a state of original righteousness, and that his first sin was the inexplicable turning of a wholly good being toward evil, Irenaeus and others regarded the pre-Fall Adam as more like a child than a mature, responsible adult. According to this earlier conception, Adam stood at the beginning of a long process of development. He had been created as a personal being in the "image" of God, but had yet to be brought into the finite "likeness" of God. His fall is seen, not as disastrously transforming and totally ruining man's situation, but rather as delaying and complicating his advance from the "image" to the "likeness" of his Maker. Thus, man is viewed as neither having fallen from so great a height of original righteousness, nor to so profound a depth of total depravity, as in the Augustinian and Calvinist theologies; rather, he fell in the early stages of his spiritual development and now needs greater help than he otherwise would have required. (The contrast between the Latin and Greek doctrines of the Fall is most fully presented in N. P. Williams, *The Ideas of the Fall and of Original Sin.* London, 1927.)

In much of the British theology from the mid-nineteenth to the mid-twentieth century, which has been influenced by Friedrich Schleiermacher's discussion of evil, this earlier, less dramatic conception of the Fall has been carried further. The Fall is regarded as a virtually inevitable incident in man's development as a child of God. If man is to enter into a genuinely personal relationship with his Maker, he must first experience some degree of freedom and autonomy. For only a relatively independent being can enter into a relationship of love and trust with his Creator, and man's fall is seen as a fall into this independence. It is thus analogous to the phase of disobedience which signals a young child's assertion of his own individuality in relation to his parents.

This line of thought may be carried further on the basis of the awareness in much modern theology that the "Fall" does not refer to a historical or prehistorical, but rather a mythological event. That is to say, man has never actually existed in a state of pre-Fall perfection. The Fall story is an analysis of man's present condition of estrangement from God, but not an account of how he came to be in this condition. Using our knowledge of the early state of humankind, we may say that man, as he emerged from the lower forms of life, was endowed with only dim and rudimentary conceptions of his Maker. He existed at an epistemic distance from God, which allowed him to respond to modes of divine revelation that do not coerce the human mind but which preserve man's relative autonomy. Man's existence at this epistemic distance from God constitutes his fallen estate, and from this flows the moral and spiritual cleavage and estrangement which is traditionally called "original sin." In this type of theodicy, God bears the *ultimate* responsibility for (in other words, is the necessary and knowing cause of) man's existence as a fallen creature, although, on his own level, man remains individually responsible for his personal choices and actions. Further, though the significance of this cannot be pursued here, the God who has thus created man as imperfect but perfectible has also entered into human life, in Christ, to bring about man's redemption.

THE PROBLEM OF PAIN AND SUFFERING

HUMAN PAIN AND SUFFERING. Some instances of suffering—for example, those caused by war, injustice, and the many forms of "man's inhumanity to man"—are traceable to human wrongdoing, and thus fall within the problem of moral evil. But other sources of pain, such as disease, earthquake, flood, drought, and storm, are built into the structure of the world itself. Surely, it is urged, they make it incredible that the world should have been designed by a Creator who is both perfectly good and infinitely powerful. The theist's reply is that this reasoning presupposes that God's purpose in making the world must have been to produce a hedonic paradise for man to inhabit. (This is the assumption, for example, of David Hume in his *Dialogues concerning Natural Religion,* XI.) It is assumed by the critic of Christian theodicy that the Creator's problem was analogous to that of a human being who is making a cage for a pet animal. He will naturally make it as safe and agreeable as he can, and any remaining sources of danger or discomfort are evidences of either his want of care or want of means. But the Christian conception of the divine purpose in creation differs from the one which is presumed in such a criticism. According to the Augustinian and Calvinist theologies, nature was created free from defect, and its present perils and hardships are punishments which man has brought upon himself. According to the Irenaean type of Christian theodicy, the purpose of the world is to be a place of "soul-making," an environment in which the higher potentialities of human personality may develop. To this end, it is claimed, nature is an autonomous system operating by its own laws, which men must learn to obey.

If God had created a world in which natural law were continuously adjusted for the avoidance of all pain, the more heroic human virtues would never be evoked. Indeed, a great part of our present moral language would be meaningless. There would be no such thing as "doing harm," for no one would be able to suffer any kind of injury; there would be no such thing as "doing good," for there would be no needs, deficiencies or occasions for improvement; there would be no such thing as a crime or a benefaction, an act of generosity or of meanness, of kindness or unkindness; and there would be no situations to which such qualities as courage, fortitude, loyalty, honesty, and the caring and protective aspects of love would be appropriate responses. There would thus be no occasions for moral choice. Such a world would be ill-designed to evoke many of the human traits which we value most highly. Indeed, it would seem that the "rough edges" of the world—its challenges, dangers, tasks, difficulties, and possibilities of real failure and loss—constitute a necessary element in an environment which is to call forth humanity's finer qualities.

But might not men have been created by God already possessed of these virtues? This is one of the points of contemporary debate. On the one side it is argued that, in principle, there are no limitations to an omnipotent God's capacity to create beings endowed with specific personal characteristics. On the other side it is argued that a virtue which has been formed as a result of making real decisions in real situations of moral choice is of greater value than the analogous virtue created by divine fiat, and that it is reasonable to suppose that the Creator is not content to build into men the merely ready-made and unearned qualities.

However, the discernible connection between the more heroic human virtues and the kind of world in which we live remains a very general one. It does not by any means amount to a one-to-one correspondence between each item of evil in human experience and some moral gain accruing to those who undergo it. Further, it appears that evil has crushed the human spirit as often as it has developed it, and that men have collapsed before life's challenges and opportunities as often as they have risen triumphantly to meet them. Accordingly, this type of theodicy demands completion in an eschatology. It points toward the eternal happiness of human beings in society with one another and in communion with God, which is symbolized by the "Kingdom of Heaven"; and its fuller claim is that the final fulfillment of God's purpose for his creatures in his heavenly kingdom constitutes a good so great and enduring that it justifies all the pains which have been experienced in order to reach it. (At this point, again, theodicy excludes the notion of eternal torment, for such torment could never serve any good end beyond itself, and would thus constitute precisely the kind of unredeemed evil which would make a theodicy impossible.)

A fundamental objection that is raised against this appeal to eschatology is that there is a contradiction between justifying a first-order evil, such as danger, as being required for the second-order good of courage, and then justifying the process by which courage is produced out of evil by reference to a future heavenly state in which, presumably, there will no longer be any dangers, and hence no need to have developed the virtue of courage in the first place. More generally, if heaven is free from "rough edges," how will virtues, so dearly bought in this world, survive within it?

Possibly the difficulty might be met in terms of heavenly analogues of earthly virtues, created by the development of the latter but no longer requiring the situations which evoked them; or in terms of the transmutation of a particular virtue (courage, for example) into an aspect of faith in God. However, Christian theology has not developed any definitive answer to this question.

ANIMAL PAIN. Thus far this article has been concerned with evil as it directly affects humankind in the forms of sin, pain, and suffering. There is also, however, the baffling problem of animal pain beneath the human level. Throughout the animal kingdom, one species devours another, and painful accidents and lingering diseases disable and then kill. How is this spectacle of "nature, red in tooth and claw" to be reconciled with the religious belief in an omnipotent and perfect Creator?

Certain solutions of the problem have been proposed. It is claimed that the lower animals live wholly in the present moment and lack the high-level capacities of memory, anticipation, and conscience that give rise to the human experience of suffering as distinct from the experience of physical pain; that the pain mechanism is a necessary warning device in bodily organisms that move about within a material environment; and that an animal's life, even though violently terminated, is predominantly active and pleasurable.

Solutions of a more speculative nature have been sought in two main directions. From the viewpoint of the Augustinian type of theodicy, it has been suggested that the premundane fall of Satan has had cosmic consequences, perverting the entire evolutionary process to a savage struggle for existence (see for example, C. S. Lewis, *The Problem of Pain,* pp. 122–124). The criticisms that have been made of the Augustinian account of the origin of evil apply also to this extension of it.

From another point of view, which adopts a theme of Eastern thought, it has been suggested that there may be a continuous reincarnation of souls through the levels of animal existence up to self-consciousness in human life. Thus, the pain of the animals is not wasted, but is part of a long constructive process (see Nels Ferré, *Evil and the Christian Faith,* pp. 62–65). The aspect of this suggestion that is most readily open to criticism is its entirely speculative and unverifiable character.

See also Anselm, St.; Augustine, St.; Calvin, John; Evil; Indian Philosophy; Leibniz, Gottfried Wilhelm; Lovejoy, Arthur Oncken; Luther, Martin; Mani and Manichaeism; Mill, John Stuart; Monism and Pluralism; Moral Psychology; Moral Realism; Plato; Plotinus; Punishment; Thomas Aquinas, St.; Voltaire, François-Marie Arouet de; Zoroastrianism.

Bibliography

STATEMENTS AND HISTORIES

The problem of evil is forcibly stated in J. S. Mill, "Nature," *Three Essays on Religion* (London: Longmans, 1874) and David Hume, *Dialogues concerning Natural Religion,* edited by Norman Kemp Smith, 2nd ed., X and XI (London, 1947).

For historical treatments of the attempts to meet the problem, see Julius Müller, *Die Christliche Lehre von der Sünde,* 3rd ed., 2 vols. (Breslau: Josef Max, 1849), translated by William Urwick as *The Christian Doctrine of Sin* (Edinburgh: T. & T. Clark, 1868); F. Billicsich, *Das Problem des Übels in der Philosophic des Abendlandes,* 3 vols. (Vienna, 1952–1959); Charles Werner, *Le problème du mal dans la pensée humaine* (Paris, 1944); R. A. Tsanoff, *The Nature of Evil* (New York, 1931); A. G. Sertillanges, *Le problème du mal,* Vol. I (Paris, 1948); and N. P. Williams, *The Ideas of the Fall and of Original Sin* (London: Longmans, 1927).

PLOTINUS AND AUGUSTINE

Plotinus deals with the problem in the *Enneads,* First Ennead, Eighth Tractate, and Third Ennead, Second Tractate, which appear in *The Enneads,* translated from Greek by Stephen Mackenna (London, 1917–1921). See also B. A. G. Fuller, *The Problem of Evil in Plotinus* (Cambridge, U.K.: Cambridge University Press, 1912). Augustine's main discussions of evil are in *De Libero Arbitrio, De Vera Religione,* and *De Natura Boni,* among his anti-Manichaean works, translated by J. H. S. Burleigh in *Augustine: Earlier Writings* (London and Philadelphia, 1943); and in his *Confessions,* Bk. VII, Chs. 3–5 and 12–16, and *Enchiridion,* Chs. 3–5 in *Confessions* and *Enchiridion,* translated and edited by A. C. Autler (London and Philadelphia: Westminster, 1955); and *De Civitate Dei,* Bk. XI, Chs. 16–18 and Bk. XII, Chs. 1–9, translated and edited by Marcus Dods (Edinburgh, 1871). See also M. L. Burton, *The Problem of Evil: A criticism of the Augustinian point of view* (Chicago and London: Open Court, 1909); Régis Jolivet, *Le problème du mal d'après saint Augustin* (Paris: G. Beauchesne, 1936).

THOMAS AQUINAS

Thomas Aquinas deals with the problem of evil in his *Summa Theologiae,* First Part, Questions 47–49, and *De Malo* (second of the Quaestiones Disputatae); see A. C. Pegis, ed., *The Basic Writings of Saint Thomas Aquinas* (New York: Random House, 1945). See also Jacques Maritain, *St. Thomas and the Problem of Evil* (Milwaukee, WI: Marquette University Press, 1942).

LEIBNIZ

Leibniz's famous work is titled *Essais de théodicée sur la bonté de Dieu, la liberté de l'homme, et l'origine du mal* (Amsterdam, 1710), translated by E. M. Huggard as *Theodicy; Essays on the Goodness of God, the Freedom of Man, and the Origin of Evil* (London: Routledge & K. Paul, 1951). For a contemporary theological critique of Leibniz, see Karl Barth, *Die Kirchliche Dogmatik* (Munich: Kaiser, 1932), Vol. 3, Part 1, par. 42,3; translated into English as

Church Dogmatics, edited by T. F. Torrance and G. W. Bromiley (Edinburgh, 1959–1960).

HELLENISTIC THEODICY

For origins, see Irenaeus, *Adversus Haereses,* edited by W. W. Harvey, 2 vols. (Cambridge, U.K., 1857), Bk. IV, Chs. 37–38, also found in A. Roberts and J. Donaldson, eds., *The Writings of the Ante-Nicene Fathers* (Edinburgh, 1867; Grand Rapids, MI, 1885), Vol. I; and A. Harnack, ed., *Epideixis* (Leipzig, 1807), translated by J. P. Smith as *Proof of the Apostolic Preaching* (Westminster, MD: Newman, 1952).

JEWISH WRITINGS

Some major Jewish writings on the problem of evil are Maimonides, *Guide of the Perplexed,* translated by M. Friedländer, 3 vols. (London: Trubner, 1881–1885), Book III, and Martin Buber, *Bilder von Gut und Böse* (Cologne: J. Hegner, 1952), translated by Michael Bullock as *Images of Good and Evil* (London: Routledge & K. Paul, 1952).

NINETEENTH-CENTURY STUDIES

The Idealist treatment of the problem is found in Josiah Royce, *Studies of Good and Evil* (New York: D. Appleton, 1898). Arthur Schopenhauer's views are in his *Parerga und Paralipomena* (Berlin: A. W. Hahn, 1851), from which T. B. Saunders has assembled and translated passages under the title "On the Sufferings of the World," in *Studies in Pessimism* (London: Sonnenschein; New York: Macmillan, 1890).

TWENTIETH-CENTURY DISCUSSIONS

The following are some modern Augustinian discussions: A. G. Sertillanges, *Le problème du mal* (Paris: Aubier, 1951), Vol. 2; Paul Siwek, *The Philosophy of Evil* (New York: Ronald, 1951); François Petit, *Le problème du mal* (Paris, 1958), translated by C. Williams as *The Problem of Evil* (New York: Hawthorne, 1959); Charles Journet, *Le mal* (Paris, 1961), translated by Michael Barry as *The Meaning of Evil* (London: G. Chapman, 1963); Mark Pontifex, "The Question of Evil" in *Prospect for Metaphysics,* edited by Ian Ramsey (London and New York: Allen & Unwin, 1961); and among Protestant writers, C. S. Lewis, *The Problem of Pain* (London: Bles, 1940); Austin Farrer, *Love Almighty and Ills Unlimited* (New York: Doubleday, 1961).

The Irenaean type of theodicy finds classic expression in Friedrich Schleiermacher, *Der Christliche Glaube,* 1st ed. (Berlin: G. Reimer, 1821), Part I, 3, and Part II, translated and edited by H. R. Mackintosh and J. S. Stewart as *The Christian Faith* (Edinburgh: T. & T. Clark, 1928). More recent writers who tend in this direction include F. R. Tennant, *Philosophical Theology,* Vol. 2, Ch. VII (Cambridge, U.K.: Cambridge University Press, 1930); William Temple, *Nature, Man and God,* Ch. XIV (London: Macmillan, 1934); and Nels Ferré, *Evil and the Christian Faith* (New York: Harper & Row, 1947); J. S. Whale, *The Christian Answer to the Problem of Evil* (London, 1939). John Hick, *God and Evil* (London: Macmillan, 1966) traces the two types of theodicy, Augustinian and Irenaean, and presents an Irenaean theodicy for today.

Some philosophical discussions are: A. C. Ewing, *Fundamental Problems of Philosophy,* pp. 247–250 (London, 1951); J. E. MacTaggart, *Some Dogmas of Religion* (London: E. Arnold, 1906), Chs. 6 and 7; C. H. Whiteley, *An Introduction to Metaphysics* (London: Metheun, 1950), Ch. 11. In addition to the articles referred to in the body of this article, see John Wisdom, "God and Evil," in *Mind* 44 (173) (1935): 1–20, and H. D. Aiken "God and Evil," in *Ethics* 68, (2) (1958): 77–97, (1949), reprinted in Aiken's *Reason and Conduct* (New York: Knopf, 1962).

More purely theological discussions occur in P. T. Forsyth, *The Justification of God* (London: Duckworth, 1916) and Karl Barth, *Die Kirchliche Dogmatik* (Munich: Kaiser, 1932), Vol. 3, Pt. 3, par. 50, translated as *Church Dogmatics,* edited by T. F. Torrance and G. W. Bromiley (Edinburgh: T. & T. Clark, 1959–1960). A book discussing Karl Barth's views on the problem of evil is Kurt Lüthi, *Gott und das Böse* (Zürich: Zwingli-Verlag, 1961).

See also C. G. Jung, *Antwort auf Hiob* (Zürich: Rascher, 1952), translated by R. F. C. Hull as *Answer to Job* (London: Kegan Paul, 1954). On the problem of animal pain, see the relevant chapters in the works cited by Ferré, Farrer, and Lewis.

John Hick (1967)

EVIL, THE PROBLEM OF [ADDENDUM]

The problem of evil concerns whether the existence of an all-powerful, all-knowing, perfectly good creator is rendered unlikely (or less likely than it would otherwise be) given the horrendous evils that afflict the world. John Hick's soul-making theodicy is perhaps the best known of the attempts to provide a plausible account of the role that evils may play in the divine plan for human life. Two other important theodicies are due to Marilyn Adams and Richard Swinburne.

In "The Problem of Hell: A Problem of Evil for Christians" (1993), Marilyn Adams discusses the problem of evil from the perspective of Christian theism, acknowledging the distinctive values of Christian theism as well its dark side, "the postmortem evil of hell, in which the omnipotent creator turns effectively and finally against a creature's good" (p. 302). As a Christian philosopher, her own view is that God's goodness to the creatures he creates is such that he will provide to each person a life that is a great good to that person on the whole. Accordingly, she rejects the traditional doctrine of an eternal hell in favor of universal salvation. In developing her view Adams carefully discusses the alternative view that some creatures so misuse their free will that God has no choice but to condemn them to an eternal life in hell, a place of constant torment whose inhabitants would be better off had they not been born.

Adams's view is that a careful look at how some people exist in the world—such as kids brought up in crack

houses, or the abused—makes it simply unrealistic to suppose that each person freely chooses an evil life or a good life. She insists on seeing God as the loving, forgiving father, rather than as the vengeful lord bent on punishing those who disobey his rules. To the objection that withdrawing the threat of eternal punishment leads to moral and religious laxity, she replies that her pastoral experience as an Anglican priest suggests otherwise: "the disproportionate threat of hell produces despair that masquerades as skepticism, rebellion, and unbelief. If your father threatens to kill you if you disobey him, you may cower in terrorized submission, but you may also (reasonably) run away from home" (Adams 1999, p. 325). Because it is abundantly clear that the majority of humankind fail in this life to grow into true children of God, Adams must suppose that there are postmortem lives in which the slow progression in growth continues until all become true children of God. She also must suppose that undergoing suffering is somehow an important step to fully entering into a life with a God.

In "Some Major Strands of Theodicy" (1996) Swinburne cites certain good states of affairs—for example, enjoyment and pleasure owing to the satisfaction of desires—that God may bring about; he cogently argues that sometimes these good states of affairs cannot be brought about without certain evils occurring or its being in the power of some created beings to produce those evils. For example, Swinburne notes that compassion is a good state that requires the existence of the bad state of suffering. Moreover, the unique goodness of compassion may justify God's permission of some degree of suffering in the world. But may it reasonably be thought that the compassion of others for the victims of the Holocaust justifies a loving being's permission of that human tragedy?

Swinburne is aware of this common objection to his theodicy. His critics may agree with him that certain good states require the existence of bad states. They may also agree with him that one should not neglect the intrinsic value of being of help to those who suffer, as well as the intrinsic value of experiencing being helped and comforted. What his critics reject is the idea that these goods require God to permit the extraordinary amount of horrendous suffering that is known to exist in the world. Swinburne's response, "Yet it must be stressed that each evil or possible evil removed takes away one more actual good" (1996, p. 44), may strike many readers as doubtful. Surely, they may say, not every fawn's death by fire serves the good of teaching other deer to avoid fires. And would many great goods have been lost if only four million—rather than six million—perished in the Holocaust? The connections between evils and goods do not appear to be as fine-tuned as Swinburne takes them to be. In response, Swinburne suggests that the world his critics think God must bring about is a kind of "toy world" where every evil is clearly seen by everyone to directly result in some outweighing good.

See also Evil; Heaven and Hell, Doctrines of.

Bibliography

Adams, Marilyn. *Horrendous Evils and the Goodness of God*. Ithaca, NY: Cornell University Press, 1999.

Adams, Marilyn. "The Problem of Hell: A Problem of Evil for Christians." In *Reasoned Faith: Essays in Philosophical Theology in Honor of Norman Kretzmann*, edited by Eleonore Stump. Ithaca, NY: Cornell University Press, 1993.

Howard-Snyder, Daniel, ed. *The Evidential Problem of Evil*. Bloomington and Indianapolis: Indiana University Press, 1996.

Rowe, William ed. *God and the Problem of Evil*. Malden, MA: Blackwell, 2001.

Swinburne, Richard. *Providence and the Problem of Evil*. Oxford: Clarendon Press, 1998.

Swinburne, Richard. "Some Major Strands of Theodicy." In *The Evidential Argument from Evil*, edited by D. Howard Snyder. Bloomington and Indianapolis: Indiana University Press, 1996.

William L. Rowe (1996, 2005)

EVOLUTION

See *Darwinism*

EVOLUTIONARY ETHICS

Evolutionary theory came of age with the publication of Charles Darwin's *Origin of Species* (1859), in which he argued that all organisms, living and dead, including humans, are the end result of a long, slow, natural process of development from one or a few simple life forms. Believing this new world history to be the death knell of traditional ways of thinking, many were inspired to find evolutionary parallels in other fields, including ethics—in both evolution of appropriate guides for proper human conduct (substantive ethics) as well as the justificatory foundations for all such social behavior (metaethics).

At the substantive level the evolutionary ethicist's usual point of departure was Darwin's own suggested mechanism of change—the "natural selection" of the

"fittest" organisms in the struggle for existence—seeking to find an analogue in human conduct. Although this philosophy became known as Social Darwinism, its widespread popularity, especially in America, owed less to Darwin himself and more to the voluminous writings of his countryman Herbert Spencer, a notorious enthusiast for extreme libertarian *laissez-faire* social and economic policies.

In later writings Spencer tempered the harshness of his philosophy, seeing a definiterole for cooperation in society, and this ambiguity about his real position led to his followers making contradictory claims, all in the name of the same philosophy. At one end of the spectrum there were supporters like the sociologist J. B. Sumner, who saw a place only for the success of the successful, and at the other end were American Marxists who saw in biology, as interpreted by Spencer, the true rules of moral conduct. Softer and more subtle forms of Social Darwinism tried to combine social responsibility with enlightened capitalism.

In this century the debt to Spencer is ignored and unknown, and the term Social Darwinism, burdened by history, is avoided. Nevertheless, particularly among biologists and politicians, the tradition has continued of seeking rules of conduct in what are believed to be the sound principles of the evolutionary process. At the beginning of the century there was the exiled Russian anarchist, Prince Peter Kropotkin, who argued that all animals are subject to a cooperating tendency toward "mutual aid" and that this can and will function once we dismantle the apparatus of the modern state. Later, the English biologist Julian Huxley became the first director general of UNESCO and based his policies on a biologically oriented religion of humanity directed toward the survival of the human species. And today we have the Harvard entomologist and sociobiologist Edward O. Wilson, who urges the preservation of the rain forests lest humans, who live in symbiotic relation with the rest of nature, fade and die. It is less than obvious, from a historical or conceptual point of view, that some of the more racist ideologies of this century owe much to evolutionary biology. The Nazis, for instance, shrank from the implication that all humans have a common origin, ultimately simian (although they were happy with the idea that within the human species there were biological differences).

Evolutionary ethics has long fallen from favor in philosophical circles, chiefly because of its supposed metaethical inadequacy. In his *Principia Ethica* (1903), G. E. Moore penned the classic critique, complaining that systems like that of Spencer commit the "naturalistic fallacy," trying to define the nonnatural property of goodness in terms of natural properties, in Spencer's case the happiness supposedly produced by the evolutionary process. Psychologically, however, enthusiasts for evolutionary ethics find this critique most unconvincing. It is more effective to point to the earlier attack of Thomas Henry Huxley (Julian's grandfather), who argued that systems deriving morality from evolution invariably rely on the hidden—and dubious—premise that evolution is in some sense progressive and that value is thus increased as one goes up the scale. Recently, with the increased biological interest in the evolution of animal social behavior ("sociobiology"), there has been renewed interest by philosophers in the possibility of fruitful connections between biology and morality. In his influential *A Theory of Justice*, John Rawls suggested that social contract theorists might explore fruitfully the possibility that in real life morality is end result of the evolutionary process rather than the construct of a hypothesized group of rational beings. Rawls drew attention to the similarities between his own beliefs in "justice as fairness" and the results of such sociobiological mechanisms as "reciprocal altruism."

This position taken by Rawls and others is a naturalistic position on ethics. If the science fails, then so does the philosophy. Have we any reason to think that—even if we agree that a Rawlsian type of situation is that which could and would be maintained by selection—that this position would ever come into being? This position of Rawls is an option for intelligent agents rather than beings that are basically under the control of the genes and hence, in crucial respects, might not be planning at all for themselves. There are various ways in which one might start to approach the empirical questions. Much interest has been shown in our close relatives, the chimpanzees. Students of their behavior argue strongly that we do find actions strongly suggestive of cooperation that simulates the moral.

Another naturalistic approach focuses on game theory. Models drawn from game theory are now showing that some kind of justicelike reciprocation can evolve among humans, even when no prior planning is involved. To see this, let us introduce two important concepts. The first is the notion of a Nash equilibrium, which posits that if there are two players in a game who are fighting over a fixed sum, and if they together demand more than the sum, neither will get anything. Given that both players know what the other will do, what is the most rational move for this first player? Suppose, for instance, that there are 100 units to be divided and player 1 knows that player

2 will demand 70 units. Then the most rational demand for player 1 is 30 units. An equilibrium holds if the distribution is 30:70—player 1 cannot do better than this, and could do worse. The second notion is that of an Evolutionarily Stable Strategy, whereby no one mutantation or variation can gain predominate over or eliminate all others in the population. Selection for rarity will lead to such an equilibrium because if the variation gets more common, it will be under heavier selection pressure, and conversely.

Now fairness would seem to demand that the two players agree to divide 50:50, but why should this result evolve given that it could be rational to go 30:70, given the greediness (but not irrationality) of player 2? The philosopher Brian Skyrms has shown that in fact only a 50:50 distribution is an evolutionarily stable situation. His insight is that if anyone coming into a population asked for less than 50 units, where the inhabitants asked for 50 units, then the invaders would do less well. If they asked for more than 50 units in such a population, they would always get nothing. Conversely, if the inhabitants asked for less than 50 units, the invaders asking for 50 units would spread. In his conclusion, where everyone asked for less than 50 units, one would always get less than one might have had. But if one asks for more, then too often one will end up getting nothing at all. So a kind of justice as fairness result comes out of the evolutionary process.

Suppose we grant all of this. You may still complain, legitimately, that we do not have morality. We have beings behaving as if they were moral. Morality, however, involves a sense of moral obligation. At this point, obviously, the Darwinian ethicist supposes—that is, makes an empirical assumption—that this sense of obligation is something put in place by selection to make us work together, to make us altruists who respect fairness. Normally we are self-centered. That is the way that selection has made us. So we look to our own needs when it comes to food and sex and so forth. But we are social animals also, and there are advantages to being social. So we have this moral sentiment that makes us reach beyond ourselves. Morality in this sense is an adaptation, just like any other.

Work is now proceeding at an empirical level showing how moral sentiments emerge in games of strategy. But, at the general level, the most obvious empirical support for the suggestion that ethics (substantive ethics) is an adaptation is that it fits in with the general Darwinian picture. We do have biological inclinations to selfishness—we want food and mates for ourselves—and so, if cooperation is of value, we need adaptations to let us break through the selfishness. A moral sense is just what is needed. Substantive ethics is a kind of quick and dirty solution to the question of cooperation. It gets you to act quickly, even though (as with quick and dirty solutions) it might not always be the best answer.

Thinking of evolutionary ethics at the metaethical level also, we find that there has been renewed thought. Because the search for foundations seems so misguided – committing what Moore called the "naturalistic fallacy," could it not be that the evolutionist is directed toward some noncognitivist "ethical skepticism," where there simply are no foundations at all? This is the approach taken by Wilson collaborating with the philosopher Michael Ruse. Following up on the thinking of the late John L. Mackie, they suggest that ethics might be simply a collective illusion of our genes, put in place by natural selection to make humans into good cooperators. To this they add that the reason ethics works is that our biology makes us "objectify" our moral sentiments; thus, we are psychologically convinced that morality, despite its lack of real foundation, is more than mere subjective sentiment.

See also Altruism; Darwin, Charles Robert; Darwinism; Human Nature; Huxley, Thomas Henry; Kropotkin, Pëtr Alekseevich; Mackie, John Leslie; Metaethics; Moore, George Edward; Moral Motivation; Rawls, John; Self-Interest; Social Contract; Wilson, Edward O.

Bibliography

Huxley, T. H. *Evolution and Ethics* [1894]. Princeton, NJ, 1989.

Mackie, J. L. "The Law of the Jungle." *Philosophy* 53 (215) (1978), 553–573.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Richards, R. J. *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago, 1987.

Ruse, M. *Taking Darwin Seriously*. Oxford, 1986.

Ruse, M., and E. O. Wilson. "Moral Philosophy as Applied Science." *Philosophy*, Vol. 61 (1986), 173–192.

Skyrms, B. *Evolution of the Social Contract*. Cambridge, 1996.

Spencer, H. *The Principles of Ethics*. 2 vols. London, 1892.

Wilson, E. O. *On Human Nature*. Cambridge, MA, 1978.

Wright, R. *The Moral Animal*. New York, 1994.

Michael Ruse (1996, 2005)

EVOLUTIONARY PSYCHOLOGY

Human beings are evolved creatures. Our lineage stretches back through the first humans to have evolved roughly 150,000 years ago, through their hominid ances-

tors, all the way back to the common ancestor we share with all other forms of life on the planet. Many of our traits are the historical results of evolution. This holds as much for psychological traits such as the visual system, emotions, and some behavior-producing mechanisms as for physical traits such as the heart, eye, or hand.

In a broad sense, evolutionary psychology covers any inquiry that uses this fact about our biological heritage to illuminate our human psychology. Historically, Charles Darwin himself pursued this kind of inquiry, as did such disparate figures as Herbert Spencer, John Dewey, and Sigmund Freud. Contemporary scientific fields such as human ethology and evolutionary anthropology are also instances of evolutionary psychology in the broad sense.

More commonly, however, the term evolutionary psychology is used in a narrower sense to refer to a specific research program that deserves to be called a Kuhnian paradigm. This paradigm is most closely associated with the psychologist Leda Cosmides (1957–) and anthropologist John Tooby, who have been among its strongest and most vocal advocates; other prominent figures in this paradigm include David Buss (1953–), Martin Daly (1944–), Steven Pinker (1954–), Donald Symons (1942–), and Margo Wilson (1942–). The manifesto for this group is the 1992 volume *The Adapted Mind*. This specific paradigm has been the most controversial branch of the general evolutionary approach to psychology and is therefore the focus of this entry. Unless otherwise qualified, *evolutionary psychology* will here be used to refer to this specific paradigm. There are four distinctive theoretical commitments in this paradigm of evolutionary psychology.

COMPUTATIONALISM

In keeping with most cognitive science and much contemporary psychology, evolutionary psychology construes the mind as an information-processing machine, which can be described in cognitive and computational terms. What is important about the mind is not what it is made of but what it does, namely, take in information from the environment, operate on internal representations, and produce behavior. The physical properties of the brain, such as its size and the amount of energy it requires, may have played some role in our evolution; but at least as important in evolution is what the mind does, and this is to be characterized functionally.

ADAPTATIONISM

Organisms possess many traits that appear to have been designed to help them survive and reproduce—photosynthesis in plants, the vertebrate eye, and so on. Such traits increase the fitness of the organism, which essentially means they make it more likely for the organism to transmit its genes to future generations. These traits are adaptive.

Evolution by natural selection is the best explanation for the existence of complex and functionally integrated traits such as the eye. Natural selection works by preserving and modifying heritable mutations that increase their possessors' fitness. Suppose some organism is born with some novel and simple trait (due to random mutation) that gives it a slight fitness advantage over its conspecifics. The next generation will tend to have more such organisms, and so the new trait will spread throughout the population. The more common the trait becomes in the population, the more likely that some new, beneficial mutation will arise in organisms with that trait, in which case organisms with both mutations will become more frequent in the population, and so on. By accumulating many small, beneficial mutations, natural selection can build complex and well-designed traits. Traits that evolved because they increased their bearers' fitness are adaptations.

Two questions can be distinguished about any trait: first, whether it is an adaptation and, second, whether it is currently adaptive. The first is an historical question concerning the role of natural selection in the origin of the trait; the second concerns whether the trait at the present time *fits* the organism to its environment (strictly, whether the trait tends to increase the organism's genetic representation in later generations). Adaptations must have been adaptive when they evolved, but they need not be adaptive now. They may no longer fit the environment if it differs from the environment in which the trait evolved.

Evolutionary psychologists claim that the human mind contains many traits that are adaptations (but may no longer be adaptive in modern environments). The environment in which traits evolved is called the environment of evolutionary adaptedness (EEA). Note that *environment* is construed broadly in evolutionary theory, covering geographical, physical, biological, and social factors. In the case of human psychological evolution, the social environment must have been especially important. According to evolutionary psychologists, the human EEA was the Pleistocene era, which started about 1.8 million years ago and ended 10,000 years ago. They argue that there has not been enough time since then for selection to have produced any significant new adaptations. Adaptations take time to evolve, especially such complex adapta-

tions as psychological traits, and there have not been enough generations since the Pleistocene for new psychological adaptations to evolve.

Throughout the Pleistocene, human beings lived as hunter-gatherers in small-scale groups. Hence our adaptations are equipped to deal with this kind of environment but not necessarily modern environments, which are different in many salient ways. Our food preferences are a commonly cited psychological example. Humans enjoy and seek out foods high in sugar and fat. In the nutrient-poor environment of the Pleistocene, such preferences were adaptive since they helped our ancestors maximize their caloric intake. But they are no longer adaptive in modern environments in the developed world where such foods are all too readily available.

Since the mind/brain is an organ of tremendous complexity and sophistication, evolutionary psychologists argue that it must have evolved by natural selection. More than that, specific psychological mechanisms evolved to solve the suite of adaptive problems faced by our hunter-gatherer ancestors—problems of how to increase their genetic representation in future generations. This fact is crucial to understanding the mind, claim evolutionary psychologists, because it allows researchers to engage in reverse engineering. In an evolutionary functional analysis, evolutionary psychologists try to infer what adaptive problems our ancestors would have faced and what sorts of psychological mechanisms would be required to solve them on the basis of what is known about conditions in the human EEA. Through such an analysis they generate hypotheses about our psychological adaptations and then test for the presence of these adaptations in modern humans.

MODULARITY

The third main theoretical commitment of evolutionary psychology follows naturally from the previous one. The mind is not a single, monolithic adaptation, argue evolutionary psychologists. Rather, the mind is comprised of many functionally distinct units dedicated to solving specific adaptive problems faced in the EEA. These distinct psychological mechanisms are modules.

When Jerry Fodor first developed the notion of a psychological module in *The Modularity of Mind* (1983), he characterized them as sharing a cluster of nine distinctive features. Evolutionary psychologists have focused on only a subset of these. The modules they propose are supposed to operate fast and automatically (without conscious effort). They are more or less informationally encapsulated from other psychological mechanisms—they do not have full access to all the information stored elsewhere in the mind. Finally, they possess innate information about the adaptive problem they were designed to solve.

Fodor himself believed that modules would only be found at the functional periphery of the mind, handling input processes such as vision. More controversially, evolutionary psychologists claim not only that more central cognitive processes are modular but also that the mind is massively modular. Cosmides and Tooby (1992), for instance, claim that the mind must contain thousands of different modules, each of them dedicated to solving different adaptive problems (and subproblems) in the EEA.

Evolutionary psychologists have offered some general evolutionary arguments for why the mind should be largely comprised of modules rather than domain-general processes. First, the adaptive problems our ancestors faced were many and varied—foraging for food, selecting the best possible mate, avoiding incest with one's kin, and so on—and these require different sorts of solutions. A mind with domain-specific ways to solve these problems is faster, more efficient, and more reliable than a general-reasoning sort of mind. Therefore, modular minds would have been selected over general reasoners in our ancestral lineage, and our own evolved cognitive architecture should be massively modular.

The second argument for massive modularity is that only massively modular minds could have produced adaptive behavior. General reasoners could not have learned by themselves and in their own lifetimes the advantages of avoiding incest or helping kin, especially since what counts as error and success is not the same in all domains. Modular creatures with domain-specific knowledge of what to do and when to do it would have been fitter than general reasoners.

UNIVERSALITY

The last main theoretical commitment of evolutionary psychology is that psychological adaptations are part of our universal human nature, with two exceptions: where a person lacks the adaptation because of mutation, and some cases of sex differences (in particular, adaptations concerning sexual reproduction). Evolutionary psychologists believe this about adaptations in general: Any trait that increases its bearer's fitness will tend to spread to fixation through a population, given enough time. Since our psychological adaptations evolved during the Pleistocene, there was enough time for them to become fixed in the entire human species.

Evolutionary psychologists have several defenses against the obvious rebuttal that human psychological nature looks anything but universal. First, they tend to downplay the massive cultural differences that anthropologists in the early to mid-twentieth century claimed to have found. Second, and a less ad hoc defense, evolutionary psychologists claim only that our psychological adaptations are universal, not that all our psychological traits are universal. Given that they also view complexity as the mark of an adaptation, however, this concession does not really grant the possibility of variation in complex psychological traits.

Their third, most interesting defense is that, even if we grant significant diversity across and between cultures, this diversity may still be produced by a common underlying mechanism. Evolutionary psychologists are interested not in behavior but in the psychological adaptations that produce behavior. An adaptation exposed to one set of environmental cues might produce a different behavior if it were exposed to a different set of cues.

One way a universal mechanism can produce diversity is a common psychological mechanism responding differently to different environmental cues. The linguistic work of Noam Chomsky, himself not an evolutionary psychologist, provides a classic example of this. According to the Chomskian tradition, the world's various languages are all underwritten by a basic universal grammar. All normal humans possess a modular language-acquisition device that enables us to learn the language of our native environment during a certain critical period of development. Although two humans may speak two different languages, they acquired, comprehend, and speak their own language with the same mechanisms.

A second way to get diversity from a universal mechanism is where a common developmental program produces different psychological mechanisms in different environments. For instance, some mechanisms may only develop in the presence of certain environmental cues at certain stages of development. In environments where those cues are lacking, or where different cues are present, the mechanism will not develop. Evolutionary psychologists have proposed both types of explanation of how an underlying common human nature can produce behavioral diversity.

Strictly speaking, then, when evolutionary psychologists claim that our evolved psychology is universal, they mean this in a restricted sense. It is not behaviors, beliefs, or desires that are supposed to be universal but only our psychological adaptations. In some cases even the psychological mechanisms themselves are not universal but only the developmental programs that produce those mechanisms in the appropriate environments.

SPECIFIC ADAPTATIONS PROPOSED BY EVOLUTIONARY PSYCHOLOGISTS

Evolutionary psychologists have proposed too many psychological adaptations to list here, but two examples should suffice. Cosmides and Tooby (in "Cognitive Adaptations for Social Exchange") proposed a module dedicated to detecting cheaters in social exchanges. This module was postulated to explain a puzzling pattern of results on the Wason selection task—a psychological test. Humans tend to perform very badly on this task when it is framed as an abstract logical problem but perform much better when it is framed as a problem for detecting potential social violations. According to Cosmides and Tooby, we should predict that humans have a dedicated cheater-detection module because detecting cheats was a serious adaptive problem for our ancestors in the EEA, and this module is invoked by the second but not the first frame in the Wason task.

The second example of a proposed psychological adaptation is even better known and comes from Buss (particularly in *The Evolution of Desire*). According to Buss, different reproductive strategies would have been successful for men and women in the EEA, and so men and women should have evolved different mating preferences. Men who preferred to mate with younger, more fertile women would have been more reproductively successful than other men. Conversely, women who preferred to mate with high-status men would have been more reproductively successful than other women. In a massive cross-cultural survey, Buss claimed to have shown that these preferences exist to this day.

PROBLEMS WITH EVOLUTIONARY PSYCHOLOGY

Evolutionary psychology has been the subject of much critical scrutiny, from philosophy, psychology, and evolutionary biology. Each of its four main theoretical commitments is contentious, and the empirical case for many of its substantive claims has also been contested.

PROBLEMS WITH COMPUTATIONALISM. It is worth noting briefly that computationalism does have some critics in philosophy of mind. Such critics will thus be skeptical of evolutionary psychology since it assumes that the mind is computational in nature (at least, the parts of the mind of interest to evolutionary psychology). Critics of evolutionary psychology, however, have not tended to

challenge its computational assumptions since these are widely shared in cognitive science and contemporary philosophy of mind.

PROBLEMS WITH ADAPTATIONISM. Much more attention has been paid to the adaptationism of evolutionary psychology. Many biologists and philosophers of biology have looked upon adaptationist reasoning with suspicion since the biologists Stephen Jay Gould (1941–2002) and Richard Lewontin (1929–) published their famous critique *The Spandrels of San Marco and the Panglossian Paradigm*. Gould and Lewontin charged that adaptationist hypotheses about ancestral conditions are too speculative, often little more than *just-so* storytelling. Moreover, the dogmatic assumption that every trait must be an adaptation exaggerates the power of selection to overcome constraints imposed by development and population size. Finally, adaptationism neglects the other ways a trait might have evolved, in particular, that a trait might have evolved for one purpose and only later been co-opted for its current use.

For their part, adaptationists have denied the charge of dogmatism; their assumption that any particular trait is an adaptation is a heuristic one, which produces hypotheses about ancestral adaptive problems. These adaptationist hypotheses should be seen as forms of argument to the best explanation and, where possible, can and should be tested against the available empirical evidence.

The appropriateness of adaptationist reasoning is still a much debated question in evolutionary biology. Regardless of the answer to that question, however, the critique of Gould and Lewontin cannot be directly applied to evolutionary psychology. Evolutionary psychologists expressly admit that the original adaptive problem cannot be inferred from the present adaptiveness of a trait. They accept the standard distinction between the historical origin of a trait as an adaptation and its present status as adaptive or otherwise, and they believe that many adaptations are no longer adaptive.

Moreover, the reasoning in evolutionary functional analysis goes in the opposite direction to standard adaptationist reasoning. Adaptationism typically starts with an identifiable biological trait and works backward to hypotheses about the ancestral adaptive problems. By contrast, evolutionary functional analysis starts with hypotheses about the ancestral adaptive problems and predicts traits that should have evolved to solve them. If these traits can then be found in modern populations, the successful prediction corroborates the hypothesis about ancestral conditions, and we have also made some new discoveries about modern psychology.

This last point, however, highlights a legitimate theoretical concern about the adaptationism of evolutionary psychology. For evolutionary functional analysis to succeed, hypotheses about ancestral conditions must meet two conditions: First, they must be sufficiently likely to be true (or else there is no point in testing predictions derived from them), and secondly, they must be detailed enough to suggest testable predictions.

Evolutionary psychologists can draw on three main sources of evidence when developing hypotheses about ancestral conditions: direct prehistorical evidence of actual conditions in the Pleistocene, the conditions faced by still-extant groups of hunter-gatherers, and our close relatives among the nonhuman primates (primarily the chimpanzee). There is some reason to doubt that any of these sources can provide good enough evidence to meet the two conditions just mentioned, for the prehistorical record is sparse, nonhuman primates have undergone their own evolutionary trajectories since they diverged from our common ancestor, and the lifestyles of extant hunter-gatherer populations have probably changed significantly since the Pleistocene. It is also debatable whether humans in general have stopped accumulating adaptations since the end of the Pleistocene, as evolutionary psychologists claim.

If these concerns are well placed, then our knowledge of ancestral adaptive problems is at too coarse a grain to entail detailed predictions about psychological mechanisms. Granted, we can be sure of very general statements—for instance, that our ancestors faced the ancestral problem of securing a suitable mate—but their very generality robs them of predictive power. All sexually reproducing organisms face this problem, and the adaptations they evolve to solve it vary dramatically. Such coarse adaptive problems cannot provide any predictions about specific psychological solutions in human beings.

Of course evolutionary psychologists deny that the limits to our knowledge of ancestral conditions are so great; hence, a main point of contention is how much skepticism is warranted by these limits. Evolutionary psychologists think it is still possible to produce sufficiently detailed hypotheses even with such limited evidence; their critics claim otherwise.

PROBLEMS WITH MODULARITY. As with adaptationism, the concept of modularity has been the subject of general controversy, this time in psychology. Since Fodor's 1983 book, the notion of a module has been

highly influential in cognitive science and psychology. There is broad agreement on the existence of at least some modules, notably, modules for language and for visual processing. The disagreement is over the amount of modularity in the mind as a whole.

Fodor himself from the start denied that the mind could be modular, except at the periphery. He later expanded these arguments into an assault on massive modularity in *The Mind Doesn't Work That Way*. According to Fodor, a massively modular mind would not be able to entertain thoughts with contents that cross the domains of each module—it would be epistemically bounded. For instance, if the mind contained separate modules for thinking about numbers, physical objects, other minds, and so on, it could not entertain a thought about both numbers and objects. A fortiori, it could not integrate information about these various domains in reasoning.

The human mind, however, does not appear to be epistemically bounded in this way, at least for central reasoning processes. Our mind is flexible in the sorts of thoughts it can entertain; moreover, it can use information from different domains flexibly in abduction. Suppose we are trying to predict the outcome of an upcoming election; potentially, information about almost anything might be relevant—facts about geography; meteorology; economics; psychology. Human minds seem able to integrate relevant information from any domain of thought.

At most, however, Fodor's arguments show only that the mind cannot be completely modular. There is need for some central workspace where information from the various modules can be integrated. But this does not show that even central processes might not be substantially modularized. In particular, it fails to show that the mind could not contain modules dedicated to solving specific adaptive problems as well as nonmodular components downstream.

A more pressing criticism is offered by Richard Samuels (in "Evolutionary Psychology and the Massive Modularity Hypothesis") against the evolutionary psychologists' arguments for massive modularity. Samuels distinguishes between two types of module: computational and Chomskian. Computational modules are, so to speak, distinct computers with their own proprietary mental programs. Chomskian modules, by contrast, are mentally represented bodies of innate, domain-specific information that are supposed to underlie our cognitive abilities in various domains. Crucially, Chomskian modules are not computationally isolated but, rather, merely separate databases of information about the world. Various psychologists have posited the existence of such innate knowledge for domains such as intuitive physics, numbers, intuitive psychology, and universal grammar.

Samuels claims that the arguments from evolutionary psychology show the need for some domain-specific knowledge of the sort contained in Chomskian modules. Perhaps organisms do need substantial amounts of knowledge about the adaptive problems their ancestors faced in order to succeed at reproduction. This does not support the existence of the separate computational modules posited by evolutionary psychologists. All this domain-specific knowledge may be operated on by the same domain-general cognitive processes. It is one thing to argue that the mind must have a vast library of domain-specific information; it is another thing to show that it must also have a vast network of different computers dedicated to using that information.

PROBLEMS WITH UNIVERSALITY. Finally, there is room for debate about the evolutionary psychologists' argument that adaptations will generally be universal. There are known evolutionary mechanisms that can maintain alternative traits in a population, in particular, frequency-dependent selection. Frequency-dependent selection occurs when there is a set of alternative traits, no single one of which is the fittest overall. Rather, the fitness of any one of these traits depends on which traits are present in other organisms in the population and at what frequency. In some cases frequency-dependent selection can maintain polymorphism—that is, the presence of more than one alternative trait in a population—at a stable ratio in which each trait has equal fitness.

Evolutionary psychologists deny that such mechanisms would have produced true genetic polymorphism in humans. Rather than, say, two different sets of genes that produce two alternative traits, there would be a single set of genes that could itself produce the alternative traits (either randomly or in response to environmental cues, where these are available). Selection will favor this kind of adaptive plasticity over polymorphism.

Evolutionary psychologists have not provided a very good argument for this. They claim that sexual reproduction would disrupt complex adaptations unless both partners shared genes for all the adaptive traits in the population. The chief problem with this argument is that it is too strong, for it would disprove the possibility of complex genetic polymorphisms in any sexually reproducing species. Since there are several cases of genetic polymorphisms in different species, the argument cannot be sound.

THE EMPIRICAL CASE FOR EVOLUTIONARY PSYCHOLOGY

All these problems with the theoretical commitments of evolutionary psychology would mean little in the face of good empirical results. If evolutionary psychologists could point to universal psychological adaptations discovered by evolutionary functional analysis, the paradigm could be declared a success, regardless of any theoretical misgivings. Assessing the empirical case for particular evolutionary psychological claims is well beyond the scope of this entry; moreover, any such assessment would be out of date even before it went to print.

What can be said here is that the empirical case remains fiercely contested. Buss has put forth what evolutionary psychologists consider their *textbook cases* in his *Evolutionary Psychology*; David Buller (1959–) challenges the empirical case for three of these putative adaptations in *Adapting Minds*. The empirical terrain here is still up for grabs and will probably continue to be so for some time.

CONCLUSION

There are several reasons to be suspicious of the main theoretical commitments of the Cosmides and Tooby paradigm of evolutionary psychology. These reasons counsel caution about accepting uncritically the various empirical claims put forth by this paradigm. They do not prove that these claims are false, that they have not been adequately empirically supported, or that they will never be supported. To assess the claims of evolutionary psychology, our only recourse is to look to the data.

Finally, it must be stressed that the evolutionary psychology discussed here is only one paradigm within a broader field of inquiry that tries to integrate evolutionary and psychological research. Even if this specific paradigm is not entirely successful, this does not impugn the broader field itself. For human beings have evolved, and surely this fact should be relevant to psychology.

See also Chomsky, Noam; Computationalism; Darwin, Charles Robert; Darwinism; Dewey, John; Evolutionary Theory; Fodor, Jerry A.; Freud, Sigmund; Philosophy of Mind; Psychology.

Bibliography

Barkow, Jerome, Leda Cosmides, and John Tooby, eds. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press, 1992.

Buller, David J. *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature*. Cambridge, MA: MIT Press, 2005.

Buss, David. *Evolutionary Psychology: The New Science of the Mind*. 2nd ed. Boston: Allyn and Bacon, 2003.

Buss, David. *The Evolution of Desire: Strategies of Human Mating*. New York: Basic Books, 1994.

Cosmides, Leda, and John Tooby. "Cognitive Adaptations for Social Exchange." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, edited by Jerome Barkow, Leda Cosmides, and John Tooby. Oxford: Oxford University Press, 1992.

Fodor, Jerry. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press, 2000.

Fodor, Jerry. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

Gould, Stephen Jay, and Richard C. Lewontin. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London* Series B 205 (1979): 581–592.

Samuels, Richard. "Evolutionary Psychology and the Massive Modularity Hypothesis." *British Journal for the Philosophy of Science* 49 (1998): 575–602.

Kelby Mason (2005)

EVOLUTIONARY THEORY

While the fixity of species was the generally accepted view before Charles Darwin, he was not the first to propose that evolution, understood as the transformation of one species into another, occurred. The ancient Greek philosopher Anaximander maintained that people had evolved from fish, and the zoologist and botanist Jean-Baptiste Lamarck (1744–1829), as well as Darwin's grandfather, Erasmus Darwin (1731–1802), were also proponents of evolution.

Lamarck, for instance, argued, in his *Philosophie Zoologique* (1809), that life resulted from ongoing spontaneous generation and that each lineage, beginning with simple forms, was driven by an inner tendency to complexity and perfection. On his view, more complex creatures belonged to older lineages, with our own the oldest. Adaptation and diversity was explained by the inheritance of acquired characteristics. Different environments caused organisms to have different needs in response to which they would use or not use their various organs: Use would cause an organ to develop, enlarge, and strengthen, whereas disuse would cause it to shrink, deteriorate, and eventually disappear. Lamarck believed that these changes were inherited by offspring, who would in their turn continue to adapt to their environment, thus leading to transformation of the lineage. The term *Lamarckism* (or

Lamarckianism) is now used to refer to the idea that a trait that was not inherited, but was acquired within the life of an individual, could be inherited by that individual's descendants. For the most part, this idea has been discredited, but there are cases in which something that satisfies the description occurs.

DARWIN'S THEORY OF EVOLUTION

Darwin was not persuaded that evolution occurred by any of his evolutionist predecessors. The true history of his development of his ideas is controversial (Sloan 2005), but there were perhaps four main influences on him in this respect.

One was the *Principles of Geology* (1931), written by his mentor and friend, the geologist Charles Lyell (1797–1875), which Darwin read at the start of his famous five-year journey on the *Beagle* (1831–1836). Darwin was profoundly influenced by Lyell's methodological, as well as his factual claims. With respect to the former, Lyell was a uniformitarian. Broadly speaking, uniformitarianism is the view that the laws of nature have always been the same. For Lyell, this meant that geological features are to be explained by natural ("intermediate" not miraculous) processes that can still be observed to be in operation. Since he thought that these tended to bring about only slow and gradual change (e.g., a valley's formation from erosion), Lyell reasoned that the earth must be far older than the biblical 4,000 to 6,000 years. Although not a believer in evolution, Lyell also argued that investigation of the geological layers showed a continual introduction and extinction of species.

A second major influence on Darwin was his observation of the natural world, especially during his journey on the *Beagle*. His extensive collection of living and fossil animals, taken from many diverse parts of the world, and their analysis by experts in the relevant fields, convinced him (and through him, much of the scientific community) that, contra Lyell, neither fossil findings nor the present geographic distribution of species could be adequately explained other than by evolution. The task, as Darwin saw it, was to explain the evolution of species in a manner that was consistent with Lyell's uniformitarian principles.

At least by his own account (Darwin 1876/1958, p. 120) Darwin had help with this from a third major influence, *An Essay on the Principle of Population* (1798), written by the parson and social economist Thomas Robert Malthus (1766–1834). Malthus was no evolutionist; he believed that his understanding of population dynamics supported the view that populations could not change much. His concern was the possibility of social improvement, but his social theory was driven by an observation that applied to all species: Unchecked increases in population always outrun the means of subsistence. As Malthus says:

> Through the animal and vegetable kingdoms Nature has scattered the seeds of life abroad with the most profuse and liberal hand; but has been comparatively sparing in the room and the nourishment necessary to rear them. The germs of existence contained in this earth, if they could freely develop themselves, would fill millions of worlds in the course of a few thousand years.
>
> (I.I.5, 6TH EDITION)

Malthus's message for the poor was that if they were to reduce their struggle for existence they must reduce their fecundity. According to Darwin, this struggle for existence between members of the same species suggested to him a mechanism by which populations could evolve.

A fourth influence on Darwin that may have been important was his familiarity with the artificial selection of plants and animals for breeding. Such selection showed that differential reproduction could produce a change in the distribution of characteristics in a population. It was believed that this had never produced a new species, and others had used this fact to support the basic fixity of species. However, Darwin argued that if so much change could be produced in the short time since human cultivation began, vastly more change could be produced given vastly more time. Of course, artificial selection involved human intentions; differential reproduction was guided by our choices and design still had a designer. Darwin's task remained that of finding a mechanistic process that could achieve similar, only much more impressive, results.

It is impossible to do justice here to the argument that Darwin assembled in support of his theory, but the main outline of his theory is remarkably simple. It begins with the observation that the individuals of a species vary slightly one from another. Since there are many more offspring born or plants germinated than can possibly survive, there is a struggle for survival within each species. Some of this is direct competition (e.g., for food or mates), but some is indirect (e.g., some individuals are better able to withstand disease or drought). The individuals that have variants that give them an advantage in this struggle will tend to survive longer and leave more offspring. And since offspring tend to resemble their parents, this means that beneficial characteristics will tend to

be inherited more frequently than less beneficial characteristics. Over time, this causes a population to be better adapted to its environment and, especially if the environment changes, leads to a gradual change in the character of a population. The relevant periods of time are enormous ("we have almost unlimited time," "millions on millions of generations") so that, eventually, a species can be transformed to such an extent that it would be a new species.

According to Darwin, the main idea for his theory was formed in 1838 when he first read Malthus, but he did not publish his *Origin of Species* until 1859. Even then he was pushed to publish to avoid being preempted by the self-trained naturalist Alfred Russel Wallace, who in 1858 sent Darwin a letter that proposed a similar theory. Darwin's priority is well established, not only by the circle of scientists to whom Darwin had communicated his ideas but also in a summary of his theory sent in 1857 to the Harvard botanist Asa Gray. Darwin was the first to argue that natural selection was the principal cause of the diversity and adaptedness of organisms, and it was his extensive defense of the claim that evolution had occurred and could occur principally by means of natural selection that revolutionized biology.

Darwin's theory differs significantly from Lamarck's, but the views of the men were less distinct than those now attached to their names. Darwin did not believe in an inner tendency to complexity and perfection, and he argued that life had evolved just once or at most a few times. However, while he did not rate it as important as Lamarck did, he agreed that one mechanism of evolution was the inheritance of acquired characteristics.

Notice that Darwin's theory mentions only processes that can still be observed to be in operation. These processes, as he describes them, are also mechanistic: They do not involve a guiding intelligence. There is, as it is nowadays put, design without a designer.

Darwin's theory is also empirical and not, as sometimes alleged, tautological. The tautology problem was raised because the theory tells us that the fit will tend to leave more viable offspring than the unfit, although an individual's fitness is defined in terms of probable reproductive outcome. However, the theory is not tautological. Individuals within a species might not vary, they might not produce more offspring than can survive, and offspring might not tend to resemble their parents. Moreover, these facts might not lead to evolution, since the outcome also depends on any countervailing forces.

THE MODERN SYNTHESIS

In modern terms Darwin's main thesis was that when there is heritable variation in fitness within species, species tend to evolve. While Darwin appealed to natural processes that can still be observed to be in operation, he did not adequately explain all such processes: In particular, he did not adequately explain inheritance or the origin of new variation, both of which are crucial to his theory.

The mechanism of inheritance was a problem for Darwin. His (pangenesis) theory involved the idea, popular at the time, that the material responsible for inheritance was blended in offspring. If that were so, an advantageous new variant would be diluted—a popular metaphor here is that it is like a drop of white paint mixed in a can of red—with the result that its benefit, and selection for it, would probably be dramatically weakened. It was Darwin's concern over this that inclined him in his later years to give more credence to Lamarckian inheritance.

Unfortunately, Darwin never knew of the work of the Austrian monk and botanist Gregor Johann Mendel (1822–1884), which provided experimental support for a particulate theory of inheritance. According to Mendel the material responsible for inheritance consisted of discrete units (now known as genes) that could be passed unchanged from one generation to the next. Mendel's work was mostly ignored during his—and Darwin's—lifetime, and it was not until it was rediscovered in 1900 that this major difficulty with Darwin's theory was removed. The combination of Darwin's theory of evolution by means of natural selection, Mendelian genetics, and mathematical population genetics is often referred to as the modern synthesis. (Some major figures in the development of the modern synthesis were T. H. Morgan, Ronald Fisher, Theodosius Dobzhansky, Julian Huxley, and Ernst Mayr.)

Explaining the origin of variation was also important; without a new source of variation, a population cannot change much beyond a redistribution of already existing characteristics. Biologists now understand how, despite a high degree of fidelity, genes are sometimes altered. Biologists construe the word *gene* in different ways, but a common construal is that a gene is a functional segment of the DNA molecules that constitute chromosomes. Alterations to such genes can occur when there are errors in copying them or when there is a crossing-over of segments of genetic material between matching pairs of chromosomes.

Crucially, the origin of new variation is random, not in the sense that any is as likely to occur as any other, but because whether a given mutation occurs is insensitive to whether it would be adaptive if it occurred. (This leaves open the question of whether there might be selection for an increase in the rate of mutation under some circumstances.) In this sense, mutation is random but selection is not random. Whether there is selection for a characteristic is sensitive to whether or not that characteristic is adaptive. Thus, selection is thought to be mechanistic, but not random or merely a matter of chance.

Darwin's and Mendel's theories form the basis of modern evolutionary theory, but neither has survived without modification. Darwin's support of Lamarckian inheritance has already been mentioned and a number of Mendelian principles have also been revised. For example, Mendel proposed that the units of inheritance were independently sorted during the formation of gametes (sperm and eggs), but it is now known that adjacent genes on a chromosome tend to stay together when gametes are produced (this is known as gene linkage). Since the early twentieth century, however, biology has provided overwhelming confirmation of the dual ideas that evolution occurs by means of (although not exclusively) natural selection and that inheritance involves (although not exclusively) genes that are usually passed unchanged from one generation to the next.

Some developments sometimes touted as radical revisions are better seen as refinements: For example, the theory of punctuated equilibrium, which proposes that long periods of stases in a lineage are punctuated by periods of rapid change, is consistent with Darwin's thesis that evolution occurs primarily through the gradual accretion of small changes: the rapid change of punctuated equilibrium is only rapid relative to the periods of stases: no major saltations are proposed.

PHILOSOPHICAL ISSUES

No sharp line should be drawn between issues in theoretical biology and philosophy of biology. Some issues of interest to philosophers have already been touched on. The following is an outline of a few others of special interest to philosophers.

THE ADAPTATIONISM DEBATE

Biologists agree that natural selection is an important mechanism of evolutionary change, but there has been disagreement over how important it is. The biologists S.J. Gould and Richard Lewontin (1979) accuse some biologists of too readily assuming that every trait has an adaptational explanation (i.e., of assuming that each trait was selected because it was adaptive or contributed to fitness). Although the debate involves certain conceptual issues, and philosophers play a role in clarifying it (e.g., see Sober 1993, chapter 5), it is principally an empirical debate, though with widespread (including methodological) implications.

Evolution (at least genetic evolution) is now said to occur if there is a change in the proportional representation of genes or combinations of genes in a population, counting each individual's genetic makeup just once. Microevolution consists of such change within a species; macroevolution consists of such changes when they result in new species. Biologists agree that much genetic evolution is due to natural selection, but it can also be due to other causes. For example, mutation and migration can bring about a change in frequencies in a population. So can drift.

It is notoriously difficult to define the word *drift*, but the first thing to note is that both zygote (fertilized egg) formation and the selection operating on the resulting individuals are stochastic (probabilistic nor deterministic) processes, and it is this that makes room for drift. Just as a series of tosses of a fair coin can by chance deviate from a fifty-fifty ratio, genetic drift can occur either as a result of a chance disproportionate sampling of genes during fertilization, or as a result of a chance deviation from probable outcomes in survival and reproduction among the resulting individuals.

The potential for drift is increased when the population is small or the force of selection is weak. So it is, for instance, thought to have special importance in allopatric speciation, in which a small portion of a population becomes geographically isolated from the rest, and competition between almost equally or equally adaptive genes or nongene "junk" DNA (neutral selection). While drift is often spoken of as an alternative to selection, it is an aspect of its stochastic nature (Brandon 2005). Nonetheless, if a trait predominates due to drift alone, it is wrong to say that this was because there was selection of the trait, let alone selection for it.

Besides mutation, migration, and drift, there are other ways in which the evolution of a trait can require explanations other than or besides adaptive explanations. For example, even traits that were selected may not have been selected because they were adaptive. They might have been selected because of their special association with adaptive traits. Gene linkage is a way this can happen. Pleiotropy, in which a single gene has multiple phenotypic effects, is another. When a neutral or maladaptive

trait has been preserved or proliferated in a population because of its link to a beneficial trait, it is called a piggyback trait or free rider. There was selection of it, but not selection for it, and only in the latter case are traits considered adaptations (Sober 1984, pp. 97–102).

It is an issue to what extent natural selection has the power to produce ideally adaptive outcomes. How often, for instance, do gene-linked and pleiotropic traits get severed in the long run? To what extent is natural selection playing catch-up with an ever-changing environment? To what extent do developmental and phylogenetic constraints, or the necessity of climbing only local adaptive peaks, restrict its capacity to move around in design-space?

The answers to such questions have interesting methodological implications. Most obviously, if natural selection tends to produce ideally adaptive outcomes, it will be fruitful to try to understand evolutionary products as ideally adaptive solutions to problems posed by a selective regime. In contrast, to the extent that it does not, the fruitfulness of that strategy is more problematic, although the construction of what are known as optimality models could still be useful, for example, in determining to what extent natural selection was involved (Maynard-Smith 1978).

While important questions are engaged in the adaptationism debate, it has often been more rhetorical than substantial. So it is important to stress that behind the heat lays some basic agreement. Contenders agree that natural selection is not the only agent of evolutionary change but they also agree that it is the source of complex adaptive change. As Gould says, when trying to reverse the impression created by his rhetoric:

> May I state for the record that I (along with all other Darwinian pluralists) do not deny the existence and central importance of adaptation, or the production of adaptation by natural selection? Yes, eyes are for seeing and feet are for moving. And, yes, again, I know of no scientific mechanism other than natural selection with the proven power to build structures of such eminently workable design.
>
> (1997, p. 35)

THE SOCIOBIOLOGY DEBATE

The main reason the adaptationism debate has been so heated was its connection with attempts to explain human behavior and psychological characteristics by appeal to evolutionary history. A bitter debate over such attempts, one of the biggest scientific controversies of the twentieth century, began after the publication of *Sociobiology: The New Synthesis* (1975) by the Harvard entomologist Edward O. Wilson, and *The Selfish Gene* (1976/1989) by the English zoologist Richard Dawkins, which together marked the start of or brought into focus a new push by evolutionary theory into the domain of the social sciences.

Wilson's book discusses the social behavior of a wide range of species, beginning with ants and ending with humans. He suggests we should study ourselves as if we were anthropologists from Mars, bearing in mind evolutionary theory in doing so. He also offers bold and (as he acknowledges, speculative) adaptationist hypotheses regarding gender roles, the causes of war, religion, and such like. Dawkins explicitly distances himself from such claims, emphasizing (as Wilson also does to some extent) the significance of culture in our case. However, Dawkins does not refrain from colorful metaphors that undermine this distancing. In a famous passage, having talked about the origin of replicators in the primordial soup, he says:

> Four thousand million years on, what was to be the fate of the ancient replicators? … Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control. They are in you and in me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way, these replicators. Now they go by the name of genes, and we are their survival machines.
>
> (1976/1989, pp. 19–20)

The issues raised relate to what used to be known as the nature versus nurture debate. That debate, put crudely, concerned the extent to which our psychological propensities were due to nature (genes) or nurture (environment). So put, however, the debate is ill conceived, because every trait is necessarily the product of both genes and environment. A better way to understand the debate is that it concerns the extent to which differences among individuals are caused by differences in their genes or in their environments (or both). The suggestion was that we would find psychosocial differences among individuals in our species, as well as between our species and other species, that were due to differences in genes for which there had been selection.

The response was vitriolic. Even sympathizers were often concerned about political implications. Critics

blasted Wilson and Dankins—for being adaptationist, for proposing hypotheses that were neither tested nor testable, for being motivated by an ideological defense of the status quo, for being racist and misogynist, and for somehow being against free will and human dignity (e.g., see Rose, Lewontin, and Kamin 1984). They described the application of sociobiology to humans as "biological determinism" and proposed instead a position they called biological potentiality. The latter included the (patently true) claim that all our acts are within our biological potential and, something further, that there are no or virtually no significant task-specific psychological adaptations. On this view, evolution has endowed us with an impressive general-purpose intelligence and a capacity for culture and language, but it has done little else to shape our psychology. The latter is in a way the more absolute position: The claim that some social and psychological characteristics are (let alone may be) genetic adaptations is compatible with the claim that many are not. And those that are in part genetic adaptations might also be shaped by culture.

Evolutionary studies of human social and psychological characteristics now go under other names (e.g., evolutionary psychology). They remain controversial, but universities have increasingly devoted substantial resources to them. Today, the two sides have come together somewhat, with evolutionary theorists stressing the importance of culture, and with less emphasis on the other side on whether certain features are genetic adaptations as opposed to adaptations that can (whatever the basis of their heritability) usefully be understood by means of the concepts and methods developed in the context of evolutionary biology. Gene-culture coevolution has also become an important area of study. There is more discussion of how evolutionary studies should be conducted than whether they should be conducted (for more details, see Laland and Brown 2002). Nonetheless, the criticisms mentioned earlier are still repeated and are worth investigating.

The claim that some sociobiology is unduly adaptationist or inadequately tested is no doubt fair. However, there are poor practitioners in every field. Trivially, one should not too readily assume that a trait is a genetic adaptation, but one should not too readily assume that it is not either. Furthermore, the hypothesis, H1, that a given trait *t* is a genetic adaptation, competes with the hypothesis, H2, that *t* is not a genetic adaptation. So if one is not a scientific hypothesis because it cannot be tested, then the same must be true of the other.

Nor does it seem true that we cannot have evidence, one way or another, for such hypotheses. It is hard to assess claims regarding the evolutionary history of social behaviors and psychological characteristics, especially in the human case where ethical considerations constrain experiment more. However, relevant evidence can still be brought to bear. Consider the suggestion that male jealousy is an adaptation to the evolutionary problem posed by fertilization within the female ("Mama's baby, Papa's maybe"). To assess this claim, evolutionary psychologists appeal to analyses of fitness consequences, cross-cultural and cross-species comparisons, and relevant physiological findings (Barkow, Cosmides, and Tooby 1992). All such evidence can be put to poor use but it can also be put to good use. For example, cross-species physiological evidence relating to testes size and sperm competition suggests that human polyandry has been a significant factor. This evidence, while far from conclusive, helps (in combination with other evidence) to confirm rather than disconfirm the hypothesis, since it suggests that, had there been genes that predisposed human males to certain jealous behaviors, (*ceteris paribus*) there would have been significant selection pressure favoring those genes.

The criticism that sociobiologists are ideologically motivated attacks the scientists rather than the science. People try to fit new information to their preconceived ideas, and scientists are no exception. However, we can ask if, assuming we do not start from a racist or sexist perspective, an evolutionary study of racial or sexual differences will push us in that direction. Here it helps to distinguish between political consequences and logical implications. The former may be worrisome even if the latter are not, and this can muddy discussion. However, what is clear is that attempts to understand certain social and psychological characteristics as evolutionary adaptations need not be used to defend any racist or sexist status quo.

Suppose it were shown, for example, that women tend, on average, to give more priority to their children than to their careers than men do, and that this difference is in part due to a genetic adaptation. Or suppose it were shown that men tend, on average, to be more competitive and aggressive (even violent) in attempting to acquire power and status and that this difference is in part due to a genetic adaptation. It can be argued from this that men will continue to have more power in the public sphere. However, this is a prediction, not a justification, and it is based on an assumption of nonintervention. One could also argue on the basis of such claims that educators need to consider moderating such difference, that there ought

to be more work-based childcare, or that the human race would be better served if less aggressive women held more political power. It is not uncommon nowadays to see theorists from the left employ evolutionary theory to make their arguments (Singer 1999).

Finally, the issue of free will is a large one, but philosophers generally agree that it is a confusion to implicate it in this debate. It is a misunderstanding of the nature of the problem of free will to think that free will is enhanced by environmental as opposed to genetic causes of behavior. The problem of free will arises as soon as human choices are viewed in the context of their causes, whatever the nature of those causes. Nor is it right to see sociobiology or its descendants as committed to determinism. In general terms, determinism is the thesis that every event is causally necessitated by preceding conditions and the laws of nature. Neither sociobiology nor its descendants are committed to this or to variants that might plausibly be described as, more specifically, biological determinism. For instance, they are not committed to the view that if a person possesses a gene that was selected because it predisposes individuals to want multiple sexual partners then someone with that gene will be unable to resist the temptation to have multiple sexual partners. Desires need not be more irresistible for being genetic adaptations as opposed to cultural artifacts.

CONCLUDING REMARKS

A general defense of the study of social and psychological characteristics from an evolutionary perspective is not the same as a defense of particular claims about social or psychological adaptation. It is consistent with the claim that such a study will fail to establish that there are any significant social or psychological adaptations. However, many think that this research will provide (and has already provided) valuable insights, relevant to many areas in philosophy. For example, the study of the evolution of altruism and the evolution of emotions are of interest to ethicists, moral psychologists, decision theorists, philosophers of mind, and political philosophers.

Darwinian evolutionary theory has had a profound impact on our understanding of our species and on our worldview, even putting to one side its role in the social sciences. While it is remarkably well confirmed by innumerable findings, and now coheres with our understanding of genetics in innumerable detailed ways, it remains controversial in the public sphere for this reason. Philosophers have played an important role in this debate as well, particularly in discussions of the nature of scientific theories.

This entry has left many issues relating to evolutionary theory untouched. A great many other issues are important as well. For example, what is the role of teleology in Darwinian biology? How does a historical science, like the study of evolution, compare to the other natural sciences? What is the implication of Darwinian evolution for the idea that species are natural kinds? How should living things be classified? How best can the intertwined concepts of selection, fitness, and drift be understood? What is selected in selection? Can memes evolve by means of natural selection? Can cultures? A number of these issues are discussed elsewhere in these volumes.

See also Anaximander; Darwin, Charles Robert; Darwin, Erasmus; Darwinism; Determinism and Freedom; Evolutionary Ethics; Lamarck, Chevalier de; Malthus, Thomas Robert; Paley, William; Philosophy of Biology; Teleological Argument for the Existence of God; Wallace, Alfred Russel; Wilson, Edward O.

Bibliography

Barkow, Jerome H., Leda Cosmides, and John Tooby, eds. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press, 1992.

Brandon, Robert. "Natural Selection." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University, 2005.

Darwin, Charles. *The Autobiography of Charles Darwin* (1876), edited by Nora Barlow. New York: The Norton Library, 1958.

Darwin, Charles. *The Origin of Species by Means of Natural Selection*. London: John Murray, 1859.

Dawkins, Richard. *The Selfish Gene* (1976). New York: Oxford University Press, 1989.

Gould, Steven J. "The Darwinian Fundamentalist." *New York Review of Books*, 44 (10) (June 12, 1997): 34–37.

Gould, Steven J., and Richard Lewontin. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." In *Proceedings of the Royal Society of London* 205 (1979): 581–598.

Laland, Kevin L., and Gillian R. Brown. *Sense and Nonsense: Evolutionary Perspectives on Human Behavior*. New York: Oxford University Press, 2002.

Malthus, Thomas. *An Essay on the Principle of Populations: A View of Its Past and Present Effects on Human Happiness; with an Inquiry into Our Prospects Respecting the future Removal or Mitigation of the Evils Which It Occasions* (1798). 6th ed. London: John Murray, 1826.

Maynard-Smith, John. "Optimization Theory in Evolution." In *Annual Review of Ecology and Systematics* 9 (1978): 31–56.

Rose, Steven, Richard Lewontin, and L. J. Kamin. *Not in Our Genes: Biology, Ideology, and Human Nature*. New York: Pantheon, 1984.

Singer, Peter. *A Darwinian Left: Politics, Evolution, and Cooperation*. New Haven, CT: Yale University Press, 1999.

Sloan, Phillip. "Evolution." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University, 2005.

Sober, Elliott. *The Nature of Selection*. Cambridge, MA: MIT Press, 1984.

Sober, Elliott. *Philosophy of Biology*. San Francisco: Westview Press, 1993.

Wilson, Edward, O. *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard University Press, 1975.

Karen Neander (2005)

EXISTENCE

Philosophical discussion of the notion of existence, or being, has centered on two main problems that have not always been very clearly distinguished. First, there is the problem of what we are to say about the existence of fictitious objects, such as centaurs, dragons, and Pegasus; second, there is the problem of what we are to say about the existence of abstract objects, such as qualities, relations, and numbers. Both problems have tempted philosophers to say that there are inferior sorts of existence as well as the ordinary straightforward sort, and they therefore often suggest that we use the word *being* to cover both kinds but restrict "existence" to "being" of the common, nonfictitious, nonabstract sort. (Sometimes the term *reality* is proposed for "existence" or for "being.") The problems of fiction and abstraction are different, however, for there are both real and fictitious abstractions. For example, the integer between two and four is real, but the integer between two and three is fictitious. On the other hand, there are both concrete and abstract fictions; for example, the winged horse of Bellerophon and the integer between two and three. Accordingly, philosophers have often dealt with the two problems in quite different ways and perhaps ought to do so.

While these are the two main problems, there are others, for example, that of what we are to say of the being of objects which have not yet begun, or have now ceased, to exist. The history of this subject, moreover, has been tangled with theological issues, to which it will be necessary to refer at certain points. Most of what follows will concentrate on the question of fictitious existence, with some consideration of past and future existence and with a final section on the being sometimes ascribed to abstractions.

FICTIONS

ANCIENT AND MEDIEVAL. It is clear from Aristotle's earliest works that there was current among the Greeks a sophism to the effect that whatever is thought of must exist in order to be thought of. Countering this, Aristotle distinguished between "to be *A*" (for instance, to be thought of) and "to be" without qualification. He made the same remark about the "being *A*" of that which has "ceased to be"; for instance, from "Homer is a poet" it does not follow that "Homer is." Some such distinction seems necessary, since among the *A*'s that one can be is "dead" or "no longer existent." The Aristotelian view is not simple, however, for elsewhere he suggested that propositions equivalent to "Socrates is ill" and "Socrates is well" imply the plain "Socrates is"; "neither is true if Socrates does not exist at all."

The various facets of the Aristotelian position were reproduced by the Scholastics. They distinguished, for example, between *est secundi adiecti* ("is" added as a second element in a simple sentence, as in "Socrates is"), and *est tertii adiecti* ("is" added as a third element, as in "Socrates is ill"), a distinction made by Aristotle in substantially the same terms; they also had the rule that from "being *A*" we may infer plain "being." But the Scholastics questioned and qualified this rule with the other Aristotelian examples in mind; from *Chimaera est opinabilis,* "The chimera can be thought of," the plain *Chimaera est* does not follow, nor does *Caesar est* follow from *Caesar est mortuus.* Some predicates, they said, presuppose *esse simpliciter,* and some do not. (*Chimaera* itself, incidentally, they put in the first class, and were thereby led to say that even "The chimera is the chimera" is false—the chimera, since it "isn't," isn't anything, even the chimera.) In this connection they sometimes distinguished between "objective" and "formal" being; "objective" not having its modern sense but rather the opposite—for a thing to exist "objectively" it sufficed that it be an object of thought; "formal" being was, as it were, serious being.

Anselm's Ontological Argument for the existence of God hinged on this notion of "objective" being, which he called existence in the mind. The Psalmist's fool who says in his heart that God does not exist must in that very act be thinking of God, so that God exists in his mind. But by definition God is that than which nothing greater can be thought of (if this definition is objected to, we can simply say, "Never mind—that than which nothing greater can be thought of, whether you give this the name "God" or not, can be thought of, and therefore exists in the mind even of one who denies its existence anywhere else"). And to exist outside the mind as well as in it is a greater thing than to exist in the mind only; we can also think of a being that does this, so if it exists in the mind only, the *maximum cogitabile* is not the *maximum cogitabile* at all;

it is therefore self-contradictory to deny real existence to that than which nothing greater can be thought of. "Greater" in this argument means, in part, better, so it is no answer to say that of some things it might be better for them not to exist. If God is by definition everything that a being ought to be, then such a being ought to be, among other things, real; and to think of him as not being real is therefore to think of him as not being all that he ought, and thus is not to be really thinking of God. The thought of an unreal God is internally incoherent, so if God can be thought of at all (and the man who thinks he does not exist does think of him), he must, if we are to be consistent, be thought of as real. But it is clear that this type of argument, if valid, has other than theological applications; as Anselm's opponent Gaunilo pointed out, there would be a similar contradiction in ascribing merely mental existence to the greatest (and best) conceivable island.

The notion of "weak" forms of being entered into another medieval theory, that of the "ampliation," or widening, of the "supposition," or reference, of certain terms in certain contexts. If we say "Some men are running," it is understood that we mean some men now existing; but if we say "Some men were running" or "Some men will be running," it will be enough to verify the former if some formerly existing men were running, and the latter if some men who will then exist will be running; and if we say "Some men are thought of as running," it will do if some merely thought-of men (Sherlock Holmes and Dr. Watson, say) are thought of as running. The pool of objects on which we may draw to verify our statements may in some cases extend to objects whose being is comparatively shadowy yet is substantial enough for them to be genuine subjects of affirmative discourse.

It is noteworthy, however, that at least one medieval thinker, Thomas Aquinas, quite firmly refused to avail himself of the notion of substandard existence at a point where it might have been thought to be helpful: in dealing with objections to the doctrine of creation out of nothing. Creation out of nothing, Thomas insisted, is not creation out of a peculiarly tenuous material; to be created out of nothing is simply not to be created out of anything, and God himself is "created out of nothing" in this sense. But to be created (as God of course is not), yet not created out of anything, is to be given existence; and to what is existence given if there is literally nothing there to give it to? To give existence to nothing, surely, is just not to give existence to anything, and thus not to create at all. That there already are "creatables" whose capacity to exist in the full sense is made actual in their creation is explicitly denied; the only power involved is that of God to create, and creating is denied to be a genuine action on an object. All that Thomas could say positively is that the receiver of existence starts to be simultaneously with the giving of that existence—*Deus simul dans esse, producit id quod esse recipit: et sic non oportet quod agat ex aliquo praeexistenti* (*De Potentia Dei,* Q.2, A.1). Leibniz's later talk of God as conferring actual existence upon a selected few of an infinite number of *possibilia,* each with its own eternally complete individuality, seems much less perceptive. Creating out of *possibilia* is not creation out of nothing at all.

EIGHTEENTH AND NINETEENTH CENTURIES. The rejection of halfway points between existence and nonexistence that characterized Thomas's treatment of creation found wider applications in the works of the eighteenth-century philosopher Thomas Reid. In discussing "conception," Reid distinguished this "operation of the mind" from others by the fact that whereas "the powers of sensation, of perception, of memory, and of consciousness [introspection] are all employed solely about objects that do exist, or have existed … conception is often employed about objects that neither do, nor did, nor will exist" (*Essays on the Intellectual Power of Man,* Essay IV). It is important, he said, to "distinguish between that act or operation of the mind, which we call conceiving an object, and the object that we conceive." The former always exists; the latter need not. The notion that it must has led some philosophers to interpose between the act and the object an entity called an idea, which is the "immediate" object of conception and exists even when the "remote" object does not. According to Reid's view, there are no such entities, and "having an idea of" is just a circumlocution for "conceiving."

> The philosopher says, I cannot conceive a centaur without having an idea of it in my mind…. He surely does not mean that I cannot conceive it without conceiving it. This would make me no wiser. What then is this idea? Is it an animal, half horse and half man? No. Then I am certain it is not the thing I conceive…. This one object which I conceive is not the image of an animal—it is an animal. I know what it is to conceive an image of an animal, and what it is to conceive an animal; and I can distinguish the one of these from the other without any danger of mistake. (Ibid.)

Not only is there no reason to believe in the existence of "ideas" (apart from the sheer "prejudice" that all acts of

the mind must have existent objects); postulating them does not help, and only gives rise to an infinite regress.

> In every work of design, the work must be conceived before it is executed—that is, before it exists. If a model, consisting of ideas, must exist in the mind, as the object of this conception, that model is a work of design no less than the other, of which it is the model; and therefore, as a work of design, it must have been conceived before it existed. (Ibid.)

The point that the idea or copy theory just does not help seems unanswerable. But Reid's own position has serious difficulties. Every act of conception, he said, "must have an object; for he that conceives must conceive something"; thus, to conceive nothing—not to conceive anything—is not to conceive at all. But if what is conceived—for example, a centaur—does not exist in any sense, is this not simply conceiving without any second term? The object of conception in all cases is something, as Reid claimed; it would therefore seem that "somethings" ("beings," "objects," in a wide sense) are divisible into two sorts, ones with existence over and above their mere objecthood, and ones without. Reid said as much when he spoke of a man as acquiring sufficient judgment to "distinguish things that really exist from things that are only conceived." Reid was trying to slip this statement past as if it were not the theory of weak and strong modes of "being." But if it is not, what is it? He said that men six feet tall and men sixty feet tall are both "things," even if he said it under his breath. The theory of weak and strong being is not easily avoidable, and in the present century it was stated in these terms in Bertrand Russell's early work and in Reid's terms by Alexius Meinong.

There is a rather different strand in postmedieval philosophy, starting perhaps with Pierre Gassendi. Replying to the Cartesian version of Anselm's Ontological Argument, Gassendi said that existence is not a "perfection," not because it is sometimes an undesirable property but because it is not a property at all; it is, rather, a prerequisite for the possession of any properties. In itself this looks like merely accepting one side of Aristotle's dilemma and ignoring the other (must a thing exist in order to possess the property of being thought of?). Later writers, however, took the thought further. David Hume, in particular, held that existence is a nugatory notion. "When after the simple conception of any thing we would conceive it as existent, we in reality make no addition or alteration on our first idea" (*Treatise of Human Nature*, Book I, Part 3, Sec. 7 and note). The notion of an existent God, man, or centaur is simply the notion of a God, a man, or a centaur. The common view of judgment as "the separating or uniting of different ideas" is therefore mistaken; the judgment that God exists, for example, involves only one idea, that of God. The real difference is not that between a God and an existent God but that between conceiving God's existence (which is the same as conceiving God) and believing in his existence; this is a difference in our mode of thought, not in what is thought of. Immanuel Kant made the same point in his treatment of the Ontological Argument; existence, he said, is not a genuine predicate, and the conception of a hundred real dollars has no more "content" than that of a hundred merely possible dollars.

Hume's denial that the object of judgment is necessarily complex was taken further by Franz Brentano. Since the time of Aristotle, logicians had divided propositions or judgments into simple "existential" assertions and denials, of the forms "*X* is" and "*X* is not," and "predicative" ones, of the forms "*X* is *Y*" and "*X* is not *Y*," and had tended, when assimilating the two, to treat existential assertions as a special case of predicative ones—"*X* is" amounts to "*X* is existent" or "*X* is a being." Brentano reversed this, treating the difference between the two types as merely a difference in the complexity of what is asserted or denied to be—"*X* is *Y*" amounts to "*XY* is" and "*X* is not *Y*" to "*XY* is not" (for instance, "Horses don't fly" amounts to "Flying horses do not exist"). He avoided the infinite regress that would appear if "*XY* is" was further transformed into "Existent *XY* is," and "*XY* is not" into "Existent *XY* is not" by denying, with Hume, that existence is a feature of the object of thought. "*X* is" and "*XY* is" simply express the mind's acceptance of the concept *X* or *XY*, as opposed to the mere entertainment of it, and "*X* is not" and "*XY* is not" similarly express the mind's rejection of "*X*" and "*XY*."

An obvious objection to the theories of Hume, Kant, and Brentano is that they are concerned not with the notion of existence but with that of believing in something's existence, and the account given of this latter—that to believe that what is before the mind exists is simply to accept this object—is such that the notion of the object's existence apart from our belief disappears. It is like the account of "believing *X* to be good" that reduces it to "liking *X*"; the notion of *X*'s being good, outside of its being believed to be so, vanishes. However it may be with beliefs about goodness, beliefs about the existence of things are surely true or false; and they are true only if the things in question exist in fact. Brentano's answer was to define what really exists as the object of right affirmative judgment, but this seems to reverse the

matter—the judgment that *X* exists is a right one if and only if *X* does exist.

However, the treatment of predicative judgments as complex existential ones can easily be disentangled from Brentano's subjectivism; and it is also found in writers, notably John Venn, whose interests were less in metaphysico-psychological questions than in developing formal logic by algebraic means. If we are to do this, it is natural to write "No *X* is a *Y*" as "$XY = 0$" and to read this as "*XY*'s are nonexistent." Similarly "Some *X* is a *Y*" becomes "$XY \neq 0$"; "Some *X* is not a *Y*," "$X\overline{Y} \neq 0$" (writing "$\overline{Y}$" for what is not a *Y*; that is, "non-*Y*"); and "Every *X* is a $\overline{Y}$," "$X\overline{Y} = 0$," or "*X*'s that are not *Y* are nonexistent." This, as Brentano and John Venn observed, has the odd consequence of making both universal forms true of what does not exist, for if there are no *X*'s at all, there are none that are *Y* and none that are not *Y*, so that "No *X* is a *Y*" is true in the sense of "*XY*'s are nonexistent" and "Every *X* is a *Y*" is true in the sense of "*X*'s that are not *Y* are nonexistent." This is sometimes expressed as the dictum that the null (or empty) class is included in every class; the class of centaurs, for example, is included in that of tables because, since there are no centaurs at all, there are no centaurs that are not tables.

This leads to divergences from the scholastic rule that affirmative predications imply the existence of their subject while negative predications do not; and the late nineteenth-century discussions of this problem of the "existential import" of the Aristotelian predicative forms were very elegantly summed up by J. N. Keynes in his *Studies and Exercises in Formal Logic.* The scholastic rule, it should be noted, was subject to qualification in the light of the doctrine of "ampliation" already mentioned; "All centaurs are said to be musical," for example, was not falsified by there being no real centaurs because in this context the reference of "centaur" was extended to centaurs merely spoken of. A nineteenth-century counterpart of this doctrine was the theory of the "universe of discourse." "Existence," according to this theory, always means membership in some collection of objects taken as real for the purposes of the discussion, and this might be either the actual universe or, for instance, the world of Homeric mythology. Taken seriously, this would seem to be a variant of the theory of weak and strong modes of being.

Brentano's account of predicative judgments as complex existential ones is reversed in Alexander Bain's theory that all intelligible assertions of existence have complex subjects and therefore can be restated in the predicative form. For example, "when we say there *exists* a conspiracy for a particular purpose, we mean that at the present time a body of men have formed themselves into a society for a particular object" (*Logic,* Book I, Ch. 3, Sec. 23), and in general, "*XY*'s exist" = "Some *X*'s are *Y*'s." In the present century this view was enlarged upon by John Anderson. According to Anderson, statements are meaningful (and thus true or false) only if all terms occurring in them are real; that is, only if they have objects answering to them. This means, in view of the Bain-Anderson analysis of assertions of existence, that these terms must be implicitly complex and that the statements in which they occur must presuppose other statements in which this implicit complexity is made manifest; these in turn presuppose others, and so on without end. "All *X*'s are *Y*'s" (for instance, "All albinos are short-sighted") is meaningful only if there are *X*'s and *Y*'s (albinos and short-sighted individuals), and this condition is itself meaningful only if "*X*" and "*Y*" are implicitly complex—if an *X*, for example, is an *AB* (an albino is an animal with white hair and pinkish eyes), so that "*X*'s exist" means "Some *A*'s are *B*'s" ("Some animals with white hair have pinkish eyes"); and similarly for these statements in their turn.

The apparent predicate "exists" disappears from this system, without any subjectivism (the plain "*X*'s exist," or at all events "*XY*'s exist," has been given a meaning, and not just "*A* judges that *XY*'s exist"). But the system completely wrecks such simple rules of construction as the one that if "*X*" and "*Y*" can both figure as terms in a meaningful sentence, so can "*XY*" ("what is at once an *X* and a *Y*").

RUSSELL'S THEORY OF DESCRIPTIONS. The most extensive and fruitful discussions of existence in the present century have been those initiated by Russell, in works written after his abandonment of the theory of weak and strong modes of being. On the meaning of "all" and "some" Russell stands squarely in the tradition of Venn: he reads "Some *X*'s are *Y*'s" as "*XY*'s exist" and "All *X*'s are *Y*'s" as "*X*'s that are not *Y*'s do not exist"; but he also insists, in the tradition of Kant, that "exists" is not a genuine predicate. For Russell the fundamental form of prediction is the singular or "atomic" proposition, "$x\ \varphi$ s," where "*x*" is a proper name of an individual and "φ s" is a genuine predicate or verb. "Lions exist" and "Tame lions exist" (or "Some lions are tame") are respectively analyzed as "For some individual *x*, *x* is a lion" and "For some individual *x*, *x* is a lion and is tame"; the predicates are "is a lion" and "is a lion and is tame"—"exist" has disappeared into the prefix, or "quantifier," "For some individual *x* … ." To predicate "exists" of an individual directly

named is meaningless; but Russell makes a sharp distinction between genuine, or "logically proper," names and spurious ones. A genuine proper name (Russell usually takes the demonstrative "this" as an example) contributes nothing to the meaning of a sentence except the identification of some individual as its subject; in John Stuart Mill's terminology, it "denotes" but does not "connote." Anything it tells us about the individual, if it is being used as a genuine proper name, is no part of what is being said; and if there is no individual that it identifies, then nothing is being said at all. A singular, or "definite," description of the form "the so-and-so," or a grammatical proper name with the sense of such a description (such as "Scott" used to mean "the author of *Waverley*"), is quite a different matter. When analyzed, "the so-and-so," like "some so-and-so" and "every so-and-so," disappears into a complex of predicates and quantifying prefixes; for example, "The present king of France is bald" amounts to "For exactly one individual *x, x* is now king of France, and for any *x,* if *x* is now king of France, *x* is bald." "The so-and-so exists," unlike "This exists," has a perfectly clear meaning; it means "For exactly one individual *x, x* is a so-and-so."

This apparatus yields a neat solution to the Aristotelian problem of when "*X* is *Y*" entails "*X* is" and when it does not. When "*X*" is a genuine proper name, in Russell's view, the question does not arise ("*X* is" being simply meaningless). But when it is a definite description or has the sense of a definite description, and the predicate is complex, it is often possible to read "*X* is *Y*" in different ways; and whether it entails "*X* is" will depend on how it is read. For example, "The present king of France is not bald" may assert that someone is the present king of France but is not bald, which does entail that someone is the present king of France—that the present king of France exists; but if it means only that it is not the case that someone is the present king of France and is bald, then it does not entail either that there is or that there is not a unique present king of France. Similarly, "The present king of France is believed to be bald" may be read as asserting that someone is the present king of France and it is believed of this person that he is bald, and this does entail the existence of the present king of France; or it might merely mean that it is believed that someone is the present king of France and is bald, and this could be true even if there were in fact no present king of France. Where the whole means something of the form "Someone is the present king of France and ...." (the complexity, if any, coming after the "and"), the description "the present king of France" is said to have primary occurrence; where a whole sentence of this form is preceded by or embedded in some qualifying context, the occurrence of the description is said to be secondary. Only primary occurrences of descriptions entail that there really is something answering to them; that is, that the thing described exists. Russell, incidentally, agrees with Jean Buridan that "The chimera is the chimera" entails "The chimera exists" and thus is false, although where "*x*" represents a genuine proper name, "*x* is (identical with) *x*" is always true.

The Russellian apparatus removes the necessity for postulating weak and strong modes of being, at least where definite descriptions are involved. We do not need to suppose, when we imagine the Hydra, that there is a Hydra that we imagine, even in a weakened sense of "is," or that there is some real thing that we imagine to be the Hydra. Two prefixes, "I imagine that" and "For exactly one individual *x,*" are involved here; and which governs which greatly affects the interpretation. "I imagine that for exactly one individual *x, x* is a Hydra" makes one assertion, and "For exactly one individual *x,* I imagine that *x* is a Hydra" makes quite another; and neither assertion entails that for exactly one individual *x, x* is a Hydra. The second form, indeed, asserts a relation between me and a certain real individual (the one which I imagine to be the Hydra); but the first form does not assert any relation between me and any individual whatever, real or imaginary—it just asserts that I imagine that there is such a thing as the Hydra.

These forms are sufficient to describe what might be going on; there is no need whatever to suppose that if I imagine that there is a real individual of such and such a sort, then there really must be an "imaginary individual" of whom I imagine these things; that is, there is no need whatever to twist the first form around to "For exactly one imaginary individual *x,* I imagine that *x* is the Hydra," or to talk about "imaginary individuals" at all. There is, indeed, a merely imagined state of affairs involved, so that one might say that this solution merely removes the problem one step further. But the existence of states of affairs is part of a different problem, that of the existence of abstract rather than fictitious objects, and is, perhaps, soluble.

Russell's insistence that "exists" has no sense in which it can be a genuine predicate—that is, in which it can attach directly to a genuine proper name—seems a little arbitrary and not at all essential to his position as a whole. As G. E. Moore has pointed out, all that follows from this position is that a sentence like "This exists" is bound to be true; and one like "This does not exist" is bound to be false, if it says anything at all (for if "This exists" were

false, "this" would not be picking out any object, and nothing would be being said). Russell's own formal system does in fact contain predicates—for instance, "——is identical with itself"—which have just this property, and he uses such predicates to define the class of individuals, or "things," that is, the "universal" class. He himself, moreover, reads the proposition "For some *x, x* is in the universal class" (which follows from "For some *x, x* is identical with itself") as "At least one individual exists," and he considers it a defect of *Principia Mathematica* that this proposition is provable in it. Such a proposition is scarcely avoidable in a system which has symbols for logically proper names; and its derivation is not, as some might fear, a revival of the Ontological Argument. "It is necessary that there should be something" does not mean or entail "There is something that 'is' necessarily," any more than "There is bound to be someone who wins" entails "There is someone who is bound to win." Nor can the Ontological Argument arise in its original form even if existence is a predicate, for it is blocked by the illegitimacy of the passage from "I am imagining that God really is" to "There is an imaginary God (a being in the mind) of whom I imagine this."

Nevertheless, there is a genuine problem about contingent and necessary being in this context. One of Moore's arguments for the meaningfulness of "This exists" is that it must have some meaning because "This might not have existed" often does have meaning and in general is true. The suggestion is that "This might not have existed" is analyzable as "It might have been the case that (it is not the case that (this exists))," and this could have no meaning if its innermost component did not. The odd thing, however, is that when it is thus analyzed, the statement does not appear to be true. "It is not the case that this exists" is just "This does not exist," but under what circumstances might this have been true, if "this" is being used as a genuine proper name? Under none at all, it would seem, according to Moore's own argument above.

Some of Russell's followers—Ludwig Wittgenstein in the *Tractatus* and F. P. Ramsey—have argued that in fact the only possible states of affairs alternative to the actual one are ones in which the objects that in fact exist have properties and relations different from those which they in fact have. This is to make all individuals what the Scholastics thought God was, "necessary beings." However, the formulation "This need not have existed," read as "It is not the case that (it is necessary that (this exists))," does not seem open to these difficulties, since its verification does not require that there be states of affairs in which we could truly say of the object in question "This does not exist"; it is enough if there are states of affairs in which nothing can be said of the object in question at all (for then not all states of affairs will be ones in which we can say of it "This exists," that is, "This exists" will not be a necessary truth). This solution makes a distinction between "possibly not" and "not necessarily," which in most logical systems are equivalent; however, it is possible to develop systems of modal logic in which they are not equivalent.

Russell's distinctions also throw some light on Thomas's problems about creation out of nothing. Thus, P. T. Geach claims that "God has created a man out of something" amounts to "For some *x* (God has brought it about that (*x* is a man))"; but in the case of creation out of nothing we have only "God has brought it about that (for some *x* (*x* is a man))." In the first form, where "For some *x*" governs "God has brought it about that," what is said is that there is an already existing object on which God has so acted as to make it a man; in the second form, where "God has brought it about that" governs "for some *x*," it is not said that there is something that God makes into a man, but only that he brings to pass the state of affairs expressed by "Something is a man." But it is only the creation of something under a certain description ("a man") that could be construed in this latter way; the creation of individuals as such out of nothing (as opposed to their creation out of, say, "possible individuals") still seems impossible, since they have no identity until they are "there," and a divine "Let there be this man" would be senseless before that time.

ABSTRACTIONS

Turning from the supposed existence of fictitious objects to that of abstract ones, it will be useful to make some preliminary grammatical distinctions. Besides the proper names of individuals, most languages contain common nouns, adjectives, and verbs that apply to various individuals without actually naming any of them, and abstract nouns which seem to name the qualities, relations, states, and actions of individuals as if they were themselves individuals of a tenuous sort. We may say that a verb "φs" applies to an individual *x* if in fact *x* φs, that an adjective "*A*" applies to *x* if in fact *x* is *A*, and that a common noun "*C*" applies to *x* if in fact *x* is a *C*. Current logic tends to treat the verb form of a predicate as fundamental and adjectives and common nouns as always implicitly functioning as parts of verbal phrases like "is *A*" or "is a *C*." Where common nouns appear as grammatical subjects, as in "All men are liars," and adjectives as

directly qualifying nouns, as in "All white men are liars," a little twisting will usually set the word in the basic verb context ("Whatever is-a-man is-a-liar," "Whatever is-a-man and is-white is-a-liar"). As verbs do not even look like names of objects, there are no very serious problems.

Abstract nouns, however, do look like names of objects, and so do common nouns in certain rather special uses, as in "Man is the noblest of animals" and "Man is a species." The objects which they appear to name are sometimes called universals; here they shall be called abstract objects or abstractions. Numerals also appear to name abstract objects—numbers—and noun clauses like "that Caesar conquered Gaul" and equivalent phrases like "Caesar's conquest of Gaul" also seem to name abstract objects—states of affairs or (in one use of the term) "propositions." Other apparent names frequently used in recent times are phrases of the form "the class of so-and-so's."

Some of these linguistic suggestions were taken very seriously by Plato and others have been taken seriously by other writers, and it is now common to describe as Platonism the view that abstract objects "exist" in a perfectly literal sense as part of the "furniture of the universe" alongside tables and chairs. If the Platonist admits that their being has something peculiar about it, and perhaps calls it subsisting rather than existing, he will nevertheless say that there is a single being which both subsisting and existing objects possess. The opposed Aristotelian tradition is to say that there is no single being which is common to objects of all different categories (things, qualities, relations, etc.), and that they are not just different sorts of beings but rather that they "are" in different but related ("analogous") senses of "are." In the fundamental sense of "are," only substances or things really "are"; qualities "are" also, but only in a secondary sense—their reality consists simply in their qualifying real things; the reality of relations consists similarly in their relating real things, and so on. On the Aristotelian view, one might say there are weaker and stronger senses of "be" but no sense so weak as to cover, unambiguously, the subjects of all the others. The view that there is no necessity to attribute existence in any sense to anything but concrete individuals is generally called nominalism (after the medieval view that "universals" are no more than empty names).

Most abstract nouns have verbs, adjectives, or common nouns related to them in meaning, and the nominalist's problem could be described as that of paraphrasing statements ostensibly about abstract objects in such a way that the abstract nouns disappear into the corresponding verbs or allied forms. Often this can be done quite easily. "Caesar is a member of the class of men" and "Caesar has the attribute of manhood," or "Manhood characterizes Caesar," seem no more than pompous ways of saying that Caesar is human (or human and male). Similarly, "The class of men is included in the class of mortals" and "Manhood is always accompanied by mortality" seem to mean simply that whatever is human is mortal. But other cases are trickier—"Red is a color," for example. This means more than "Whatever is red is colored"; we may presume that all red things, and only red things, occupy a certain set of regions of space (or of space-time), so that it is true that whatever occupies these regions is colored (just as all red things are colored) but not that the occupation of those regions is a color—it is not a color but a location. "Being red is a way of being colored" is more like it, but "being red" and "way of being colored" are as nounlike in form (the one a name of one thing and the other a common noun applying to many) as "red" and "color" themselves. If we write "being red" as "redding" and "being colored" as "coloring," we can perhaps say that "_____ing is a way of _____ing" builds a sentence out of two verbs just as "Whatever _____s, _____s" does and that the suffix "-ing" does not form real names but is an inseparable part of complexes like " _____ing is a way of _____ing." But some might say that rather than be driven to ad hoc solutions such as this last, it is better to admit the existence of abstractions and be done with it.

In comparatively recent history both Platonism and nominalism have run into difficulties of a rather technical sort. Some of these, including special problems about the existence of classes and numbers, are discussed in the entry on Russell, Bertrand Arthur William.

See also Analysis, Philosophical; Anderson, John; Anselm, St.; Aristotle; Bain, Alexander; Brentano, Franz; Correspondence Theory of Truth; Essence and Existence; Existentialism; Gassendi, Pierre; Hume, David; Kant, Immanuel; Moore, George Edward; Ontological Argument for the Existence of God; Plato; Ramsey, Frank Plumpton; Reid, Thomas; Russell, Bertrand Arthur William; Thomas Aquinas, St.; Universals, A Historical Survey; Venn, John.

Bibliography

The key texts in Aristotle are *De Sophistici Elenchis* 167al ff.; *De Interpretatione* 19b18 ff. and 21a26 ff.; and *Categories* 13b16 ff. and 29 ff.

Anselm's Ontological Argument is stated in his *Proslogium* and *Monologium,* and Gaunilo's answer is in his *Pro Insipiente*;

Thomas's summary and answer are in his *Summa Theologiae,* 2; Descartes's variant is in his *Meditations*; Gassendi's criticism is in the Objections appended to Descartes's *Meditations*; and Kant's criticism is in his *Critique of Pure Reason,* Transcendental Dialectic, Book II, Ch. 3, Sec. 4. The key passages in Anselm, Descartes, and Kant are included in Charles Hartshorne and William L. Reese, *Philosophers Speak of God* (Chicago: University of Chicago Press, 1953).

For Thomas on creation out of nothing and problems arising out of the notion of "giving existence," see his *De Potentia Dei* Q. 2, A. 1, Objections 2, 7, 11, and 17 and answers. For comments on this, see Antonin Sertillanges, *L'idée de création et ses retentissements en philosophie* (Aubier, 1945), especially pp. 45–48; A. N. Prior, "Identifiable Individuals," in *Review of Metaphysics* 13 (1960): 684–696; and P. T. Geach, "Causality and Creation," in *Sophia* (1962): 1–8. Problems about the identifiability of as yet nonexistent individuals are also acutely discussed in Jonathan Edwards, *Freedom of the Will,* edited by Paul Ramsey (New Haven, CT: Yale University Press, 1957), Part IV, Sec. 8, and in G. E. Moore, *The Commonplace Books,* edited by Casimir Lewy (New York: Humanities Press, 1967), p. 329. Leibniz's solution in terms of preexistent *possibilia* is developed in his correspondence with Arnauld (edited by Geneviève Lewis, Paris, 1952), which arose out of his *Discourse on Metaphysics* (translated and edited by P. G. Lucas and L. Grint, 2nd ed., Manchester, U.K., 1963), and in the dialogue at the end of his *Theodicy* (translated by E. M. Huggard, London and New Haven, CT, 1952).

For Thomas Reid's views, see his *Essays on the Intellectual Powers of Man* (Edinburgh, 1785), Essay IV ("On Conception") and Essay VI ("On Judgment"). For Russell's early position, see *The Principles of Mathematics* (Cambridge, U.K.: Cambridge University Press, 1903), especially pp. 43 and 71. The best and most accessible treatment of Meinong's views is in J. N. Findlay, *Meinong's Theory of Objects* (Oxford, 1933; reprinted with additions, Oxford: Clarendon Press, 1963), Ch. 2.

For Hume, the key passage is in his *Treatise of Human Nature,* Book 1, Part 3, Sec. 7 and note; for Brentano, his *Psychologie vom empirischen Standpunkt* (Leipzig, 1874), Vol. I, Book II, Ch. 7, Sec. 15. J. N. Keynes's discussion of possible responses to the challenge of Brentano and Venn is in his *Studies and Exercises in Formal Logic* (London: Macmillan, 1884; 4th ed., London: Macmillan, 1906), Ch. 8. Bain's views are in his *Logic,* 2nd ed. (London, 1873), Book I, Ch. 3, Sec. 23, and are discussed in a footnote to J. S. Mill's *System of Logic,* 9th ed. (London, 1875), Book I, Ch. 5, Sec. 5.

Russell's later theories are most clearly developed in his 1918 lectures "The Philosophy of Logical Atomism" (included in *Logic and Knowledge,* London: Allen and Unwin, 1956), Lectures V and VI, and in Ch. 16 of his *Introduction to Mathematical Philosophy* (London: Allen and Unwin, 1919). The latter contains, on pp. 168–170, his splendid and rather Reid-like appeal for a "robust sense of reality," G. E. Moore's agreements and disagreements with Russell are in his Aristotelian Society paper "Is Existence a Predicate?," reproduced in his collected *Philosophical Papers* (London: Allen and Unwin, 1959). The same collection contains another relevant Aristotelian Society paper, "Imaginary Objects," which may be compared with pp. 243–245 of Moore's *Commonplace Books.* F. P. Ramsey's development of the consequence that Russell's theory allows for no contingency of existence is in *The Foundations of Mathematics* (London: K. Paul, Trench, Trubner, 1931), p. 285; see also the ingenious argument on p. 155.

On the "grammatical" solution to the problem of universals, see Russell's 1924 essay "Logical Atomism," also included in *Logic and Knowledge,* and Tadeusz Kotarbiński's article "The Fundamental Ideas of Pansomatism," translated by David Rynin, in *Mind* 64 (1955): 488–500.

Much current discussion takes its start from W. V. Quine's paper "On What There Is," included in his *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953).

A. N. Prior (1967)

EXISTENCE AND ESSENCE

See *Essence and Existence*

EXISTENTIALISM

Existentialism is not easily definable. Its protagonists have traced it back to Pascal, to St. Augustine, even to Socrates. It has been alleged in our time to be the doctrine of writers as various as Miguel de Unamuno and Norman Mailer. At first sight, characteristics of the doctrine are almost as various. That two writers both claim to be existentialists does not seem to entail their agreement on any one cardinal point. Consequently, to define existentialism by means of a set of philosophical formulas could be very misleading. Any formula sufficiently broad to embrace all the major existentialist tendencies would necessarily be so general and so vague as to be vacuous, for if we refer to a common emphasis upon, for example, the concreteness of individual human existence, we shall discover that in the case of different philosophers this emphasis is placed in contexts so dissimilar that it is put to quite different and incompatible uses. How then is existentialism to be defined?

EXISTENTIALIST THEMES

Existentialism may perhaps be considered most fruitfully as a historical movement in which connections of dependence and influence can be traced from one writer to another. Thus, even if two writers who are both rightly called existentialist differ enormously in doctrine, they can be placed in the same family tree. But this only throws the question of definition one stage back. How do we select our philosophical pedigrees? The answer must be in

terms of a number of recurrent themes that are in fact independent of one another but have, as a matter of philosophical history, been associated in a variety of patterns. The key themes are the individual and systems; intentionality; being and absurdity; the nature and significance of choice; the role of extreme experiences; and the nature of communication.

THE INDIVIDUAL AND SYSTEMS. Søren Kierkegaard chose for his own epitaph the words "that individual." The concept of the individual for Kierkegaard was contrasted both with the concept of philosophical system and with the concepts of the stereotype and the mass. Between these contrasts there is a connection. A philosophical system was for Kierkegaard an attempt to understand individual existence within a conceptual scheme of a kind that would exhibit a logically necessary connection between every individual part and the conceptual scheme of the whole universe. People in the mass, or those who live out a stereotyped role, are people who understand themselves in terms of some concept or concepts they happen to embody. In both cases the individual is secondary to the concept it embodies. In fact, however, what exists comes first; concepts are necessarily inadequate attempts to grasp individual existence, which always evades complete conceptualization. One of the difficulties in understanding what Kierkegaard and his later followers have meant by assertions of this kind is that none of their detailed arguments appear to entail their conclusion. Consider two of these arguments.

The first is a revival of Immanuel Kant's argument against the so-called Ontological Proof. Like Kant, Kierkegaard argued that existence is not a property and that no concept of a given object entails the existence of that object. Also, Kierkegaard anticipated some modern writers in arguing that action and choice can be understood only if viewed from the standpoint of the agent rather than from that of the spectator. What is puzzling, however, is that Kierkegaard assumed that the notion of philosophical system is inextricably bound up with the viewpoint of the spectator and the refusal to admit that existence is not a property. In consequence, he concluded that justice can be done to the nature of the individual only if philosophical system building is condemned. The explanation for this particular line of thinking is that Kierkegaard equated the construction of philosophical systems with Hegelianism, and he interpreted Hegelianism as a form of rationalist metaphysics. It is noteworthy that some kind of metaphysical rationalism is almost always the background for existentialism. In countries where empiricism has a long history existentialism does not seem to flourish, even in the form of a reaction to the prevailing moods of thought.

Thus, it is perhaps instructive to regard existentialists as disappointed rationalists. When they announce that reality cannot be comprehended within a conceptual system or, more particularly, that individual existence cannot be so comprehended, they identify the role of a conceptual system with the notion of an all-embracing set of necessary truths derived by deduction from some axiomatic starting point. It may seem, therefore, that existentialists are sometimes doing no more than reformulating the empiricist protest against rationalism (namely, that no matter of fact can be expressed as a necessary truth) in an unnecessary and misleadingly dramatic way. The drama, however, has at least one independent source.

The nineteenth century witnessed a series of very diverse protests against the notion that the universe is a total system, whether one presided over by a Creator God or a purely rational one developing in an evolutionary progress toward higher and higher goals. That the universe does not make sense, that there are no rational patterns discernible in it, is a theme central, for example, to Fëdor Dostoevsky's *Notes from the Underground* (1864). Dostoevsky is often cited as a forerunner of existentialism precisely because in his disillusionment with rationalist humanism he stressed the unpredictable character of the universe and because his individuals appear face to face with pure contingency. Any established connection between things may break down at any minute. Order is a deceptive mask that the universe, especially the social universe, wears. The individual thus confronts the universe with no rational scheme by means of which he can hope to master it. Reason will only lead him to formulate generalizations that will, if he relies upon them, let him down.

Existentialism sometimes gives expression to this kind of view of the limitations of reason. But it is not thereby necessarily committed to irrationalism. At least some existentialist philosophers have been prepared to argue the case for the limits of reason on rational grounds—indeed, on grounds that are partly Kantian. Moreover, when existentialist philosophers speak of the limits of reason they are usually careful to explain that they wish in no way to trespass upon the territory of the natural sciences or of mathematics. Karl Jaspers goes so far as to accept positivism as a valid account of the sciences, illegitimate only when it aspires to give an account of reasoning as such. Moreover, Jaspers would claim that the areas with which existentialism concerns itself are not

outside the competence of reason but only demand that reason be understood in new and less restrictive ways.

The claims, therefore, that the individual cannot be comprehended within a rational system and that the universe which the individual confronts is absurd turn out to have a less striking content than might at first sight have appeared. What has led to their exaggeration is perhaps in part an association with two other philosophical traditions, phenomenology and the kind of philosophy that treats ontology as a central philosophical discipline. Each of these provides existentialism with characteristic themes, which will be considered below.

INTENTIONALITY. With the exception of Kierkegaard, existentialist philosophers often make use of a conceptual scheme derived from the phenomenologists Franz Brentano and Edmund Husserl and, through them, from René Descartes. In attempting to answer such questions as What is belief?, What is an emotion?, and What is an act of will? phenomenologists wished to combat the associationist psychology that aspired to explain beliefs and emotions in purely naturalistic terms. In contrast, phenomenology emphasized that belief is always belief that … and anger is always anger about… . The object of belief or of emotion is not an object or a state of affairs in the external world. I may believe what is false or be angry about what did not in fact happen. So the object of belief or emotion is internal to the belief or emotion. It is, in the language of phenomenology, an intentional object.

Brentano concentrated on the isolated individual only in order to describe accurately the central features of believing, feeling, willing, and so on. Husserl treated the individual's consciousness of his own acts as having a primary role not unlike that which Descartes gave it. Among post-Husserl existentialists, notably Jean-Paul Sartre, the doctrine of intentionality is used to underline a fundamental difference between my knowledge of myself and my knowledge of others. Other people, so it is asserted, are viewed not as they are but as intentional objects of my perceptions, my beliefs, my emotions. But to myself I can never be such an object, nor am I in fact an object, and if they regard me as such their view of me is necessarily falsified. The obvious criticism of this is to say that the word *object* has been used as a pun. To say that my beliefs have intentional objects is to say neither that they are necessarily false nor that my beliefs about other people commit me to viewing them as things rather than people. But no existentialist writer is in fact making so simple a mistake. There is always some additional premise to the argument that provides a basis for the existentialist claim that to make others the object of my perceptions or beliefs is to view them as other than they are. In the writings of Sartre and Simone de Beauvoir, for example, specific theses about the character of love and hate play an important role.

What is clear, however, is that although the doctrine of intentionality need not be understood in an existentialist way, this doctrine does add a dimension to the existentialist concept of the individual. Only through the notion of intentionality could the themes in Kierkegaard (which were partly an inheritance from the individualism of Protestantism and partly a reaction against G. W. F. Hegel) have become in Martin Heidegger part of a theory of knowledge and of a metaphysics.

BEING AND ABSURDITY. Existentialists, believing as they do that reality always evades adequate conceptualization, are especially apt to treat "Being" as a name, the name, in fact, of the realm which we vainly aspire to comprehend. "What the philosophers say about Reality," wrote Kierkegaard, "is often as disappointing as a sign you see in a shop window which reads: Pressing Done Here. If you brought your clothes to be pressed, you would be fooled; for only the sign is for sale" (*Either/Or,* 1843).

In Kierkegaard we get little or no systematic treatment of this kind of theme. In some of his successors, however, we find a systematic ontology, which owes more to the influence of scholastic metaphysics and of rationalism than it does to Kierkegaard. Heidegger took up Gottfried Wilhelm Leibniz's question, Why are there the things that there are rather than nothing? For Leibniz this question could be answered only by producing the Cosmological Argument for the existence of God. For Heidegger the question itself is misleading, because the posing of it relies upon an inadequate analysis of the notions of being and of nothing. Heidegger treats "Being" and "Nothing" as if they were both names, sometimes the names of powers, sometimes the names of realms. It is not that he is entirely unaware of the logical difficulties encountered in so doing. But he treats such difficulties as evidence of the exceptionally elusive character of Being and Nothing rather than as a sign of his own mistakes. He also accepts the fact that scientific thought never uses such concepts or language, but this he treats as a testimony to the inadequacy of science as a method for understanding reality and to the need for poetry and philosophy. He distinguishes Being (*Sein*) from beings (*die Seiende*) and from modes of being. At times his writing is

reminiscent of scholastic ontology, but it is more often aphoristic and oracular.

In Sartre, too, there is an implicit relation to metaphysical rationalism of the kind mentioned above. The thesis that existence is absurd, which is especially important in French existentialism, turns out to be a denial of the principle of sufficient reason. There is no ultimate explanation of why things are as they are and not otherwise. What is curious here is that on the one hand the fact that this is so is seen as a flaw in the nature of things. It belongs to what Heidegger calls their "fallenness"; the experience of it arouses in us anxiety and perplexity. Yet on the other hand that it is so is the guarantee of human freedom. Both German and French existentialists distinguish sharply between the beings that exist for themselves (*pour-soi*), which have consciousness and freedom, and the beings that exist in themselves (*en-soi*), which are simply things. Now, for existentialism all the important possibilities of human life are bound up with the fact of human freedom, so that to lament the absurdity of existence is in a way odd. But what this lament does reflect is the ambiguous attitude of existentialists to human freedom.

FREEDOM AND CHOICE. If any single thesis could be said to constitute the doctrine of existentialism, it would be that the possibility of choice is the central fact of human nature. Even the thesis that existence precedes essence often means no more than that people do not have fixed natures that limit or determine their choices, but rather it is their choices that bring whatever nature they have into being. As existentialists develop this thesis, they are involved in at least three separate contentions.

The first is that choice is ubiquitous. All my actions imply choices. Even when I do not choose explicitly, as I may not do in the majority of cases, my action bears witness to an implicit choice. The second contention is that although in many of my actions my choices are governed by criteria, the criteria which I employ are themselves chosen, and there are no rational grounds for such choices. The third is that no causal explanation of my actions can be given.

The first thesis is given varying interpretations. For Kierkegaard a person's actions will always form part of a coherent way of life: the aesthetic, in which pleasure is pursued, or the ethical, in which principles are treated as binding, or the religious, in which God is obeyed. Between these one must choose, and it is in this sense that behind any action there lies a choice. For Sartre it sometimes appears as if each separate action expresses an individual choice. Even if I do not choose, I have chosen not to choose.

The second thesis is fundamental to existentialism. But it is plausible to hold that I am free to choose the criteria by which I discriminate true from false beliefs only if this contention is restricted to the field of morals and religions. Kierkegaard sometimes, although not always, allowed for this restriction.

The third thesis, which seems to be logically independent of the others, is often treated by existentialist writers as though it were entailed by the first two. This is less surprising when it is recognized that one of the impulses behind existentialism seems to be a dissatisfaction with the kind of nineteenth-century materialism which held that if human actions can be causally explained, then determinism is true in a sense that excludes the possibility of human agents' being responsible and free. However, instead of denying that causal explanation entails this kind of determinism, the existentialist takes the unnecessary step of denying the probability of causal explanations of human action.

ANXIETY, DREAD, AND DEATH. Kierkegaard argued that in certain psychologically defined moments truths about human nature are grasped. One such moment would be when we realize that we do not just fear specific objects but experience a generalized dread. Of what? Of nothing in particular. What is this nothing, this void we confront? Kierkegaard interpreted it in terms of original sin. Heidegger sees it as an ontological constituent of the universe. Sartre sees it as a confrontation with the fact of freedom, of our unmade future.

The variety of interpretations suggests that perhaps different experiences are being discussed or that the ratio of interpretation to experience may be too high. But stress on the extreme and the exceptional experience is common to all existentialism. Everyday experience, by contrast, is thought of as a conventionalized, predigested aid to complacency, conformity, and self-deception. Heidegger gives a very special place to the continuous awareness of one's own future death; Jaspers lays a more generalized stress on a range of situations in which the fragility of our existence is brought home to us.

THE FORM OF COMMUNICATION. Since the existentialist writer acknowledges the sovereignty of individual choice and the importance of the concrete situation, he cannot address himself to his audience in the manner of traditional philosophy, for ex hypothesi the reader has to make his own choices in the light of his own experiences.

Argument will be powerless unless the reader chooses to agree with the author's premises. As a matter of fact, existentialist writings do commonly argue with the reader. But Kierkegaard, for example, was usually careful to frame his arguments in a hypothetical way: "If you choose this starting point, then that logically follows … ." He was also in the habit of writing different works under different pseudonyms, so that what the reader was confronted with would be a continuing debate between rival standpoints rather than a single argued case.

Later existentialist writers have developed in two differing ways. All the major existentialist philosophers have written systematic treatises. But they have also made large contributions to imaginative literature, and the content of existentialist philosophy makes it clear that dramatic dialogue, whether in plays or in the novel, is probably a form of expression more consistent with the author's intentions than deductive argument would be.

Such, then, are the shared themes of existentialism. But at this point one ought also to stress, even if briefly, the large differences that are compatible with the thematic resemblances between individual authors.

EXISTENTIALIST AUTHORS

Since the major existentialist philosophers are all treated in separate articles, what is delineated here is their interconnections insofar as they influence one another and above all the way in which the same themes recur in quite different social and philosophical contexts.

Søren Aabye Kierkegaard elaborated all his fundamental doctrines in order to expound and to defend what he took to be true Christianity. The philosophers upon whom he drew were Hegel (though only to attack), Kant, Aristotle (purely as understood through the writings of Friedrich Trendelenburg), and the Platonic Socrates. In contrasting philosophy from Plato to Hegel with authentic Christianity, Kierkegaard emphasized the concepts of the individual, of choice, of dread, and of paradox. He thus originated all the fundamental themes of existentialism.

These themes have been put to a quite new use by Karl Jaspers, who is concerned with criticizing positivism rather than Hegelianism. He has undertaken this with a view to defending a generalized spirituality that Christianity shares with other religions, rather than to defending specifically Christian doctrines. Where Kierkegaard spoke of paradox, Jaspers speaks of contradictions, and in this he is influenced as much by Friedrich Nietzsche as by Kierkegaard.

Martin Heidegger, too, has felt the influence of Nietzsche. But St. Augustine and Husserl have also been important for his synthesis of existentialism and phenomenology. As a result of this synthesis Heidegger has outlined a systematic ontology which, as such, stands at the opposite pole to Kierkegaard's enterprise. Heidegger's world is one from which God is absent (in this, too, he contrasts with Kierkegaard), but he has denied that he is therefore an atheist. This has no doubt made it easier for theologians to utilize his writings but makes it all the more surprising that his key concepts should have been so easily integrated into yet another existentialist system, that of Jean-Paul Sartre.

In Sartre the concept of choice, which for Kierkegaard was a decision between fundamentally different ways of life, has become a ubiquitous presence behind every human action, and the being of people, which Heidegger has distinguished from the being of things in terms of the relationship of consciousness in its various modes to the world, is now defined essentially in terms of such choices.

Sartre brings together other threads from the earlier history of existentialism. He employs psychological analyses similar to Kierkegaard's analysis of dread but sets them out in terms borrowed from phenomenology. These analyses are carried through for their own sake in Sartre's philosophical writings but are put to work in his novels and plays. They are employed, too, in the novels of Simone de Beauvoir, whose moral and political writings also use the Sartrian concept of choice.

Of parallel psychological interest are the novels of Albert Camus, but the atheism that for Sartre is a consequence of his views of human nature and the world was basic in the thought of Camus. Human life is represented in the myth of Sisyphus, who was doomed eternally to roll up a hill a vast stone that would always fall back just as he was about to reach the top. The dignity of life derives from humankind's continual perseverance in projects for which the universe affords no foothold or encouragement.

Gabriel Marcel is linked to Sartre and Camus by his critique of their atheism. He is an existentialist in his stress on key experiences and on the impossibility of adequately conceptualizing the important features of human life. But the features upon which he lays stress are those of hope and relationship, and his philosophy derives from Josiah Royce's personal idealism and even from F. H. Bradley, rather than from any existentialist predecessors.

The range of views expressed by existentialist writers has made it all too easy for the most multifarious authors to claim the title and for the most widespread ancestry to be found for existentialism. Someone like Unamuno, whose book on the tragic sense of life belongs to the same climate as Kierkegaard and Dostoevsky, could scarcely for that reason be called an existentialist, but those influenced by him in Spain today might well make use of the term. Karl Heim, the German writer on the philosophy of physics, has defined existentialism so widely that almost everything not strictly in the area of science becomes the subject matter of existentialism. Such examples could be multiplied indefinitely. Therefore, it seems wise now to consider the diffused influence of existentialism in the fields of theology, politics, and psychoanalysis.

EXISTENTIALIST THEOLOGIANS

There is a variety of theological systems which in some way are in debt to existentialism. The multiplicity of conclusions which theological writers have drawn from existentialist premises is perhaps testimony both to the ambiguity of those premises and to an underlying failure to analyze adequately some of the basic concepts involved.

BARTH. The earliest theological developments are to be found in Kierkegaard's thought, not surprisingly, since he was a theologian in his own right. When Karl Barth repudiated the optimistic liberal theologies of pre-1914 Protestantism, he did so in a commentary on the Epistle to the Romans (*Der Römerbrief,* 1919), which draws quite as heavily on Kierkegaard and Dostoevsky as it does on St. Paul. From Kierkegaard, Barth took the view that God is totally other than man. Finite reason cannot hope to grasp or comprehend infinite deity. From both Dostoevsky and Kierkegaard, Barth inherited the thesis that nature and human life are enigmatic, that nothing in the world is reliable.

Barth used these doctrines in two ways. In one direction he repudiated all attempts to find a rational foundation for Christianity, whether in the rational theology of Roman Catholicism or in the philosophical idealism of nineteenth-century Protestantism. In another he used his arguments to revivify the orthodox Protestant theory of the Reformation period. It is worth noting that although Barth repudiates the possibility of any rational ground for revelation, he has, like Kierkegaard, used philosophical argument when it suited his purposes.

TILLICH. Paul Tillich, unlike Barth, used existentialist materials in constructing a system that has analogies with Heidegger's but, in contrast with Heidegger's, reaches theistic conclusions. As with Heidegger, the terms "Being" and "Not-being" or "Nothing" played a key role in his thought. God is Being-itself, but in Tillich's interpretation this characterization of God has a quite different sense from that which the same form of words would bear in medieval theology. For according to Tillich we discover "Being-itself" through self-affirmation; we discover that what we call "God" or "Being-itself" represents our ultimate concern with overcoming doubt and anxiety in the face of nothingness. The message of theology is that we can overcome the meaninglessness of contemporary existence by taking up certain types of attitudes to that meaninglessness. It is pertinent to ask whether Tillich was trying to provide Christian conclusions with a new set of existentialist premises from which they may validly be derived or was trying to provide those Christian conclusions with a new sense, which enabled him to repeat some of the traditional forms of language but gave them a quite unorthodox meaning. Support for the latter alternative can be derived from the fact that Tillich was quite content to admit that the God of traditional theism does not exist. What remains unclear is whether the word *God* is an appropriate name for the concept of Being-itself as it figures in Tillich's thought.

BULTMANN. Rudolf Bultmann, by contrast with Tillich, is avowedly concerned with reconstructing Christianity. Bultmann is a historical critic of the New Testament who believes that in the New Testament a genuinely existentialist message is distorted by being presented in terms of a prescientific cosmology. This cosmology, Gnostic in origin, is a myth from which the kernel of the gospel must be extracted. The Gnostic cosmology pictures a three-tiered universe with human life on the earth occupying a place midway between the divine realm above and the powers of darkness below. The message concealed is that men are poised between the possibility of an "authentic" (Heidegger's term) human existence, in which the individual faces up to the limits of human existence and especially his own death, and the possibility of inauthentic existence, in which the individual retreats from death and Angst and *Sorge* and so becomes their victim. The charge made against Bultmann by orthodox theologians is that he turns Jesus Christ into a mere precursor of Heidegger. Bultmann's reply is that his interpretation of the gospel is still distinctively Christian because of his insistence that the decision in which man chooses between authentic and inauthentic existence is one that the rational man

does not have the power to make for himself. But here either Bultmann is bringing in a supernaturalism that he otherwise disowns or he means simply that the choice of authentic existence is an action of which no account can be given in terms of the life of "rational man," of inauthentic existence. But to suppose that the traditional Christian assertion of the need for grace and the necessity of Christ's work is even a disguised version of the Heideggerian account of the choice of authenticity seems highly implausible.

Two of Bultmann's followers, Wilhelm Kamlah and S. N. Ogden, have argued that there is a deep inconsistency between Bultmann's Heideggerian themes and his Christian interpretations. Kamlah has argued that not only belief in the historical Jesus but also belief in a God who intervenes in history is inconsistent with Heidegger and draws atheistic conclusions. Ogden, who remains a Christian, believes that the role of the historical Jesus must be less important than either Bultmann or traditional orthodoxy suggests if justice is to be done to existentialism. It is notable that for all the writers of this kind, existentialism is above all else a characterization of the human condition as such, sharing much of the generality and the theoretical character of the Hegelian doctrines which Kierkegaard condemned.

Bultmann's references to God always appear to be external to his central concerns. When his critics ask him how he justifies belief in and speech about God, he tends to reply in traditional Christian terms that have little to do with existentialism. This perhaps provides some confirmation of the view that existentialism is in fact a theologically neutral doctrine. Its neutrality derives from its stress on ultimate commitment and the unjustifiable character of any particular commitment. If the only justification for *any* belief is, in the last analysis, that I have chosen to believe, then the same justification is equally available for all beliefs, whether theistic or atheistic. But insofar as existentialism is a doctrine about human nature, its themes are very close to those of traditional theology, and it is therefore not surprising, quite apart from any impulses originating from Kierkegaard's special concerns, that most existentialist philosophers have taken up well-defined positions in relation to theology.

An existentialist vocabulary is often used by theological writers who are not in any strong sense existentialists. So the Russian Orthodox thinker Nikolai Berdyaev and some Catholic theologians, in their discussions of anxiety, guilt, and man's relation to God, have used existentialist concepts. But these uses reflect the fashionable character of existentialism rather than any of its philosophical characteristics.

EXISTENTIALISM AND POLITICS

As in theology so also in politics existentialism appears to be compatible with almost every possible standpoint. Kierkegaard was a rigid conservative who viewed with approval the monarchical repression of the popular movements of 1848; Jaspers was a liberal; Heidegger was for a short time a Nazi; and Sartre was over a long period a Communist Party fellow traveler. However, at least three systematic political themes can be discerned in existentialism.

The first is a form of religious humanism designed to counteract what is believed to be an unsatisfactory value system at the basis of modern society. Both Jaspers and Marcel maintained that the growth in technology and bureaucracy was creating in Europe a cult of mediocrity, conformism, and loss of individuality, with the inner life of the individual sacrificed to external forms. Heidegger, too, saw the individual as threatened by impersonality. But although Jaspers and Marcel pleaded for a greater recognition of transcendent and religious values in general, neither had a specific program of social reform to offer.

Second, the existentialist stress on commitment and irrationality of choice has sometimes been used in support of irrationalistic extremism. The most notorious but not the only example is Heidegger's brief excursion into politics. Needless to say, advocates of Nazism tend to ignore the existentialist stress on the importance of the individual.

Commonly, existentialism may be associated with communism, and this is largely due to the influence of Sartre. However, Sartre has occupied more than one position. His prewar writings contain scarcely any reference to politics. During the war and immediately after, his political aims—those of a radical democrat—were expressed in terms that seem largely independent of his existentialism. At that time, in his analysis of political activity he found himself at odds with orthodox Marxism because Marxism offered causal explanations of behavior that Sartre wanted to explain in terms of choices and purposes. But in his later writings he has accepted a Marxist framework for both political theory and political practice and has presented existentialism as merely a corrective to a too rigid and too deterministic Marxism. Yet his account of political life is, in fact, still far more psychological than any a Marxist would give.

EXISTENTIALISM AND PSYCHOANALYTIC TOPICS

There are several points at which existentialism touches on psychiatric themes. Karl Jaspers originally practiced as a psychiatrist, and in *Allgemeine Psychiatrie* (1913) he criticized ordinary scientific psychology and the psychotherapy based upon it. He did so on the ground that what he regards as the positivistic approach of conventional psychotherapy is unnecessarily and misleadingly deterministic. It treats the actual outcome of the patient's life as the inevitable outcome. Jaspers concedes that scientific examination will not reveal the fact of human freedom of choice. The personality available for empirical scrutiny is simply what it is, but the assumption that there is nothing to personality but what empirical scrutiny will reveal is groundless and arbitrary. Behind the empirical self there is, in Jaspers's view, a true self of which we are made aware in what Jaspers calls "boundary-situations"—that is, in situations of an extreme kind where we confront despair, guilt, anxiety, and death. In these moments of awareness we realize our own responsibility for what we are, and the reality of freedom of choice is thrust upon us.

The name "existential psychiatry" has been taken, however, by another tradition of thought, which derives from Heidegger and whose most important exponent is Ludwig Binswanger. Binswanger, who calls his system of analysis *Daseinsanalyse,* criticizes two of Sigmund Freud's central concepts. Freud saw the neurotic symptom of the adult as caused by a past traumatic event, the memory of which has been repressed into the unconscious, from where it exerts its causal power upon present behavior. According to Binswanger, however, the neurotic symptom is to be explained not in terms of the content of the patient's unconscious but in terms of his mode of consciousness, and the key concept involved in the explanation is not that of causality but that of meaning. When an adult reacts to a situation neurotically it is because his consciousness confers upon that situation a meaning he does not recognize as deriving from the nature of his own consciousness. Certainly, past traumatic events are relevant. But they are relevant precisely because in them a like meaning was given to a like situation. Attention is thus focused upon the patient's whole mode of consciousness, the way in which he approaches, attends to, and comprehends the world. The explanation of behavior lies in the present, in the mode of consciousness, not in the past or in the unconscious.

Binswanger's understanding of the different possible modes of consciousness is derived directly from Heidegger. He speaks of "Being-in-the-world" and its modes and of the contribution of existentialist philosophy to psychiatry as consisting in the a priori analysis of all possible modes of "Being-in-the-world." He very largely discounts the biological determination of human behavior, although he allows it a minor role. But he tends to insist on interpreting behavior, even at the biological level, in terms of the meaning it has for the agent.

This emphasis is reiterated by Sartre, who uses the doctrine of intentionality to criticize all causal theories of emotion and behavior. Sartre attacks both the James-Lange theory of the emotions and the Freudian theory of the unconscious because he holds that they cannot allow for the intentional (in the Husserlian sense) and purposive aspects of emotion and behavior. It has already been suggested that it is unclear why Sartre believes that if emotions, for example, must be understood in terms of their intentional object and aim, they cannot also be explicable in causal terms. Like Binswanger, Sartre approves of much in Freudian technique, and in his writings on Charles-Pierre Baudelaire and Gustave Flaubert he has emphasized the formative experiences of early childhood. Perhaps his most extensive treatment of these themes is his book on Jean Genet (*Saint Genet: Comédien et martyre,* 1952).

Both Sartre's earlier and his later writings have been utilized by R. D. Laing in the study of schizophrenia (*The Divided Self*). Sartre's account of experiencing another person as a free agent for whom one exists only as an object and by whom one is reduced to the status of an object (in *L'être et le néant,* III, 1943) is used by Laing to throw light on case histories of the kind where the decisive actions of another person have resulted in a loss of identity on the part of the patient. Laing's work does, in fact, strongly suggest that Sartre has sometimes offered us not, as he purports to do, a description of what is basic to human consciousness as such but a description of certain abnormal types of consciousness, to which we are all sometimes prone but which become dominant in mental illness. However, Laing himself does not make this criticism of Sartre and has used in the study of normal family life some of the concepts that Sartre elaborates in the *Critique de la raison dialectique* (1960).

CRITICISM AND EXPLANATION

The suggestion that existentialism is a form of disappointed rationalism has already been made. It may now be extended to the charge that existentialism's dissatisfaction with the concepts of traditional rationalist metaphysics has been insufficiently radical. If the thesis that the universe is absurd is simply a denial that the universe

has a Leibnizian sufficient reason, then it relies as much as Leibniz did on the adequacy of the concept of a sufficient reason. When the existentialist could profitably have questioned the very terms with which the rationalist characterized the world, he has all too often simply taken over the rationalist scheme of concepts and denied what the rationalist affirmed. Moreover, he has mistaken his own denials for a positive characterization of the nature of things.

It has also been suggested that the existentialist often makes the same logical points against rationalism that the empiricist did but invests them with more drama. Perhaps the explanation of this is that the discovery that there are no sufficient reasons or ultimate justifications, of the kind offered by rationalist metaphysics and allied types of theology, is not private to existentialist philosophers. Questions of ultimate justification remain unimportant and unexamined by most people so long as social life is relatively stable and social conflict is not disruptive. When, however, the conventional supports of civilized life are withdrawn, as they have been too often in Europe since 1914, ordinary people are forced to ask questions about justification that normally do not arise for them. The loneliness and self-questioning of a Kierkegaard become far more common. Moreover, people find that their normal responses are put in question; deception and self-deception become pressing topics. What were publicly approved acts with established utilitarian justifications become signals into a darkness where there are no answering lights.

It is a commonplace that it is people living in loneliness and doubt who provide the characters for existentialist novels, but it is less remarked that the existentialist's conceptual psychology rests equally upon examples drawn from extreme situations. How, indeed, could it be otherwise for those who assert that it is only in extreme situations, in what Jaspers calls boundary situations, that authentic human nature is revealed? But existentialist writers remain open to the criticism that they treat the exceptional as the typical. Indeed, because the contrast between the exceptional and the typical has been obliterated, the force of the notion of the boundary situation tends to be lost.

When existentialists come to construct their own systems, the most obvious criticism they are subject to is that they are insensitive to the syntactic and semantic properties of the language they employ. So Kierkegaard spoke of a dread of nothing in particular as though this implied that such dread had an object whose name was "Nothing." So Heidegger hypostatizes Being and Nothing as substantial entities. So Jaspers discards the traditional framework for metaphysics but writes of "the transcendent" as though this were an expression whose meaning raised few difficulties. A. J. Ayer accused Sartre of a systematic misuse, in his ontology, of the verb "to be."

Ayer suggested that when a philosophical criticism of existentialism has been carried through it is not improper to ask for a sociological explanation of its use and vogue. He himself pointed to the fact that German existentialism followed on the defeat of 1918, whereas French existentialism is a sequel to 1940. But, in fact, Sartre took up all his main existentialist positions before 1939. And the purely philosophical ancestry of later existentialism must be allowed for.

This is not to say that we should look for an account of existentialism only in terms of philosophical antecedents. It would be more illuminating to see existentialism as the fusion of a certain kind of dramatization of social experience with the desire to resolve certain unsolved philosophical problems. The unsolved problems are those of traditional epistemology and metaphysics. In the period between Descartes and Francis Bacon, on the one hand, and Kant and Hegel, on the other, certain philosophical problems were posed but not solved. Within the framework of assumptions in which they were posed they could not, in fact, have been solved. Foremost among these assumptions is that the whole of knowledge has to be reconstructed out of the epistemological resources of the single, isolated knowing subject. Also, there is the search for first principles, based either upon an indubitable, because logically undeniable, proposition or upon an incorrigible set of reports of immediate experience. There is the treatment of the first principles as axioms and their employment as a basis for a deductive model within which all human knowledge is to be set forth. There is the invocation of God or Nature to bridge the gulfs too great for argument on its own.

Hegel abandoned all these assumptions, as Ludwig Wittgenstein did later on. But where Wittgenstein placed epistemological problems in the context of an understanding of language as a social phenomenon Hegel placed them, in the end, in the context of a metaphysical system. Those who rejected his system retreated to the epistemological assumptions of the earlier period, but with this difference: after David Hume and Kant they could no longer believe in guaranteed first principles. So Kierkegaard's choice between the ethical and the aesthetic reproduces the Kantian choice between duty and inclination but lacks its rational basis. More generally, Kierkegaard's individual resembles the Cartesian ego

without the *cogito.* Sartre inherited from phenomenology an explicit Cartesianism. In Sartre the individual as the knowing subject is the isolated Cartesian ego; the individual as a moral being is a Kantian man for whom rational first principles have been replaced by criterionless choices. Neither God nor Nature is at hand to render the universe rational and meaningful, and there is no background of socially established and recognized criteria in either knowledge or morals. The individual of existentialism is Descartes's true heir.

According to Marxist critics, especially Georg Lukács, the debased individualism of the existentialist is a symptom of the malaise of the bourgeois intellectual. Bourgeois man can no longer find his values incarnated in the social life that surrounds him; therefore, he makes a fetish of his own inner experience and tries by the fiat of his own choice to legitimate the values that in public life no longer appear to have validity. This theory of Lukács's has two central weaknesses: it appears to suggest a correlation between holding existentialist views in philosophy and having certain highly specific political and social attitudes, and it assimilates all existentialism to one rather restricted model. The suggested correlation is not warranted by the evidence, and the preceding discussion points to the dangers of assimilating different existentialisms too closely to one another.

A more relevant criticism might be phrased as follows. Certain philosophical attitudes are embedded in the matrix of existentialism; in general, existentialism embodies a distrust of metaphysical rationalism. Insofar as existentialist philosophers elaborate conceptual analyses in such fields as ethics and the philosophy of mind, their work can be understood and assessed in the same way as the work of analytical philosophers. Paradoxically, however, when they go beyond conceptual analysis it is usually not only to stress the inevitability of choice or the importance of dread but also to construct systems of the kind that existentialists originally protested against. The outcome of these systems on the whole lends further weight to the protests.

Finally, the doctrine of choice itself stands in need of closer scrutiny than existentialist philosophers have given it. This doctrine depends on the relationship between choice and criteria for judging between true and false and right and wrong. In existentialist writings this relationship remains, on the whole, unscrutinized.

See also Alienation; Augustine, St.; Ayer, Alfred Jules; Bacon, Francis; Barth, Karl; Beauvoir, Simone de; Being; Berdyaev, Nikolai Aleksandrovich; Binswanger, Ludwig; Bradley, Francis Herbert; Bultmann, Rudolf; Camus, Albert; Cartesianism; Cosmological Argument for the Existence of God; Cosmology; Death; Descartes, René; Dostoevsky, Fyodor Mikhailovich; Essence and Existence; Existential Psychoanalysis; Freud, Sigmund; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heidegger, Martin; Heim, Karl; Hume, David; Husserl, Edmund; Jaspers, Karl; Kant, Immanuel; Kierkegaard, Søren Aabye; Leibniz, Gottfried Wilhelm; Lukács, Georg; Marcel, Gabriel; Marxism; Nihilism; Nothing; Pascal, Blaise; Sartre, Jean-Paul; Socrates; Tillich, Paul; Unamuno y Jugo, Miguel de.

Bibliography

Works by individual existentialist philosophers are included only if they contain discussions of existentialism.

Ayer, A. J. "Some Aspects of Existentialism." *Rationalist Annual* (1948).

Barrett, William. *Irrational Man.* New York: Doubleday, 1958.

Binswanger, Ludwig. "Daseinsanalyse und Psychotherapie." *Acta Psychotherapeutica et Psychosomatica* 8 (4) (1960): 258. Translated as "Existential Analysis and Psychotherapy," in *Progress in Psychotherapy,* edited by Frieda Fromm-Reichmann. New York: Grune & Stratton, 1956.

Blackham, H. J. *Six Existentialist Thinkers.* London, 1951; New York: Macmillan, 1952.

Brock, W. *Introduction to Contemporary German Philosophy.* Cambridge, U.K.: Cambridge University Press, 1947.

Bultmann, Rudolf. *Glauben und Verstehen.* Vol II. Tübingen, 1952. Translated as *Essays: Philosophical and Theological.* London: SCM Press, 1955.

Collins, James. *The Existentialists: A Critical Study.* Chicago: H. Regnery, 1952.

Gilson, Étienne, ed. *Existentialisme chrétien.* Paris: Plon, 1947.

Grene, Marjorie. *Introduction to Existentialism.* Chicago: University of Chicago Press, 1959.

Grimsley, R. *Existentialist Thought.* Cardiff, 1960.

Heidegger, Martin. *Existence and Being.* Translated by D. Scott, R. Hull, and A. Crick. Chicago: Regnery, 1949. Translation of *Was ist Metaphysik?; Vom Wesen der Wahrheit; Höderlin und das Wesen der Dichtung*: and *Andenken an den Dichter: "Heimkunft—an die Verwandten."*

Kaufmann, Walter, ed. *Existentialism from Dostoevsky to Sartre.* New York: Meridian, 1956. Selections with introductions.

Kolnai, A. "Existence and Ethics." *PAS,* Supp. Vol. 27 (1963).

Laing, R. D. *The Divided Self.* London: Tavistock, 1960.

Lukács, Georg. *Existentialismus oder Marxismus?* Berlin, 1951.

MacIntyre, Alasdair. "Existentialism." In *A Critical History of Western Philosophy.* London and New York: Free Press, 1964.

Manser, A. "Existence and Ethics." *PAS,* Supp. Vol. 27 (1963).

Marcel, Gabriel. *The Philosophy of Existence.* Translated by Manya Harari. New York: Philosophical Library, 1949. Republished as *Philosophy of Existentialism.* New York, 1961.

Molina, Fernando. *Existentialism as Philosophy.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

Mounier, E. *Introduction aux existentialismes.* Paris, 1947.

Ogden, S. N. *Christ without Myth.* New York: Harper & Brothers, 1961.
Ruggiero, Guido de. *Existentialism.* London: Secker and Warbug, 1946.
Salvan, Jacques. *To Be and Not to Be.* Detroit, MI: Wayne State University Press, 1962.
Sartre, Jean-Paul. *L'existentialisme est un humanisme.* Paris: Nagel, 1946. Translated by P. Mairet as *Existentialism and Humanism.* London: Metheun, 1948.
Sartre, Jean-Paul. "Questions de méthode." In *Critique de la raison dialectique,* Vol. I. Paris: Gallimard, 1960. Translated by Hazel E. Barnes as *The Problem of a Method.* New York, 1963.
Sartre, J.-P., R. Garandy, J.Hyppolite, et al. *Marxisme et existentialisme.* Paris: Plon, 1962.
Tillich, Paul. *The Courage to Be.* London: Nisbet, 1952.
Weigert, E. "Existentialism and Its Relation to Psychotherapy." *Psychiatry* 12 (1949).

Alasdair MacIntyre (1967)

EXISTENTIALISM [ADDENDUM]

The development of "existentialism" in the last years of its leading French proponents, Jean-Paul Sartre and Maurice Merleau-Ponty, occurred in the areas of social philosophy and existential psychoanalysis in the case of Sartre and the philosophy of language and fundamental ontology for Merleau-Ponty. Partly in response to the latter's critiques, but chiefly as a result of his own political commitment, Sartre constructed a social ontology and a theory of history in his *Critique of Dialectical Reason.* Faithful to his existentialist emphasis on the primacy of the individual, but replacing his earlier philosophy of consciousness with one of praxis (roughly, purposive human activity in its historical and socioeconomic context), Sartre formulated a set of concepts, especially praxis, seriality, and the practico-inert, that respected the power of social forces to countermand, deviate, and reverse our undertakings without totally robbing the organic individual of existentialist freedom and responsibility. He allowed far greater play to the force of circumstance in assessing human action and underscored the determining power of family and early childhood experience in his massive existential biography of Gustave Flaubert, *The Family Idiot.* This last, combining the discourse of *Being and Nothingness* with that of the *Critique,* forms a kind of synthesis of Sartre's work.

At the time of his death in 1961 Merleau-Ponty was at work on a manuscript that has come to be known as *The Visible and the Invisible,* a work that some consider his version of Martin Heidegger's "What Is Metaphysics?" It reveals a growing interest in an ontology that avoids the pitfalls of "philosophies of consciousness," with their subject-object relation, which has defined and limited philosophy in the West for centuries. Inspired by the painter's articulation of the world and building on the concepts of chiasm and flesh, introduced in his earlier *The Phenomenology of Perception,* Merleau-Ponty was moving beyond the boundaries of phenomenology to elaborate an "indirect ontology" in which language questioning being questions itself.

See also Existential Psychoanalysis; Heidegger, Martin; Merleau-Ponty, Maurice; Ontology; Phenomenology; Philosophy of Language; Sartre, Jean-Paul.

Thomas R. Flynn (1996)

EXISTENTIAL PSYCHOANALYSIS

"Existential psychoanalysis" is a trend in psychology and psychiatry best understood as a reaction against the theoretical and philosophical presuppositions of the psychologies based on natural science in general and of Freudian psychology in particular. The phenomenology of Edmund Husserl and the existentialism of Martin Heidegger, Jean-Paul Sartre, and Martin Buber, rather than the mechanistic worldview of natural science, are seen by existential psychoanalysts as providing the proper philosophical and methodological route to a more complete understanding of man. In its original form, therefore, existentialist psychoanalysis was not a countermovement to Freudian psychoanalysis, unlike Jungian or Adlerian psychoanalysis, for example. Its criticism always focused on the philosophical theory of man implicit in Sigmund Freud's work, and it offered itself as a philosophical complement to Freud. The main burden of its criticism is that a full understanding of the patient's experience and world is impeded if the patient is approached on terms derived from the hypotheses of natural science rather than on his own specifically human terms.

The pioneer of existential psychoanalysis, Ludwig Binswanger, sought to describe the experiential world of his patients with the help of the conceptual scheme of Heidegger's ontology of man's being. However, his work contained few major differences from Freud in therapeutic technique. Indeed, another existential analyst, Medard Boss, has claimed that existential analysis "enables psychotherapists to understand the meaning of Freud's rec-

ommendations for psychoanalytic treatment better than does his own theory" (*Psychoanalysis and Daseinsanalysis*, p. 237). The implication is, of course, that a fuller understanding of the patient will result in more efficacious treatment, but that the methods of treatment will not differ fundamentally from Freud's. The result is to separate Freud's dealings with his patients from the mechanistic scientific constructs by which he sought to explain psychic functioning.

However, those persons in the United States who call themselves or who are called existential psychoanalysts are a very heterogeneous group. Perhaps the most significant change in the United States is that existential psychoanalysis is seen by many practitioners as a substitute for Freudian psychoanalysis. The phenomenological method of describing and hence understanding the patient's world is looked upon as itself a therapeutic measure, something far from the minds of the movement's originators. Certain notions of Sartre or Buber concerning the existential encounter between human beings are taken as replacements for the classical analyst-patient relationship. On this point, Binswanger had simply argued that the full human meaning of the doctor-patient relationship be realized in therapy; he did not seek to eliminate the classical relationship. The result of these changes has been that almost any therapy that departs from the general Freudian mold is called existentialist—particularly those which place emphasis on unusual therapeutic intervention and those which reject the scientific element in psychoanalysis, for good and bad reasons. It is therefore essential to understand the philosophical core of the original movement.

THE SUBJECT-OBJECT SPLIT

The shortcomings attributed to psychoanalysis emanate from the scientific tradition in which psychology has sought to place itself. Natural science since Galileo Galilei can be understood as a mode of approaching the world in which one aspect of the phenomenal world, the aspect of pure corporeality, is given the privileged position of basic substance, of primitive, irreducible fact. The notion of pure corporeality as the reality to which all phenomena are to be reduced is the concomitant of a dictate that the perceiver remove himself as much as possible from the world in the attempt to gain knowledge of that which is perceived. The roots of this dictate lie in the philosophy of René Descartes, which isolates the realm of consciousness from that of the body and the perceived world. Thus, the concept of pure corporeality is the product of a methodological dictate: Keep the self out of its world as it investigates its world.

This famous Cartesian sundering of the world into two isolated regions, *res cogitans* (the thinking substance, the world of consciousness, purpose, telos, will, quality) and *res extensa* (the world of pure extended matter, undifferentiated, quantitative), has been attacked by phenomenologists and existentialists as the most disastrous event in four centuries of Western thought. Nevertheless, this subject-object split had the immeasurable value of disciplining a new kind of human self-awareness, in the air since the Renaissance: man's awareness of his self-sufficiency and his urge to master nature, or the universe, which had revealed itself as a radical other. The split furthered man's alienation from his world, but at the same time it gave him, in the methods and the objective attitude of natural science, the means to bridge the separation in action if not in philosophical comprehension.

But psychology, the most recent child of this attitude, is in a strained position. On the one hand it seeks to be objective, to take its place as one of the natural sciences along with biology, physics, and chemistry. On the other hand it seeks to study that which science since Descartes and Galileo has demanded be ruled out of the field of investigation—the soul, psyche, consciousness. Psychology is thus faced with the apparently self-contradictory task of investigating consciousness as part of the realm of the *res extensa*, although the *res extensa* is that which exists independently of consciousness. To investigate consciousness scientifically, psychology must eliminate from consciousness its essential element. Freud's doctrine of the unconscious can be conceived of as an attempt to overcome this contradiction by viewing the essence of consciousness as that which lies in the realm of *res extensa*. His success is due to the fact that his scientific psychology, unlike others, does not reduce the experienced meanings in the field of consciousness to a level below the level of meaning. In Freudian psychology, meanings are reduced not to physiology or to objectively perceptible spatiotemporal processes, but to another kind of meaning, instinctual meaning.

All explanation involves the reduction of that which is explained to something taken as more basic. The ever-present danger is that that to which the phenomena are reduced may become so alien to the phenomena that there is no returning to them without circularly invoking previous knowledge. Freudian psychoanalysis seems to avoid this danger by focusing on a basic reality, instinctual meaning, which is not totally alien to the phenomena to be explained.

INTENTIONALITY

Existential philosophy denies the subject-object split that defines the whole attitude of natural science. The mind, consciousness, is not a strange and unprecedented thing whose workings are somewhat more puzzling than those of its neighboring objects, the things of this world. Nor is it, as Descartes held, a distant spectator, alien and sufficient unto itself, moving like a ghost on Earth. For existential philosophy, the problem of how the mind reaches over to the object is a pseudo problem that results from the gratuitous and erroneous presupposition that consciousness can be understood independently, apart from that which it intends or is conscious of. Mind, or consciousness, is to be defined as simply this intentionality, this reference-to. Consciousness is not viewed as something that intends an object; consciousness is the intention.

Existentialist-phenomenological psychology maintains, therefore, that the phenomena with which psychoanalysis is concerned are intentional acts, conscious phenomena, not the nonintentional, nonreferring phenomena of the world of objects. Whereas other psychologies, in emulating science, stripped consciousness of that very quality that constitutes its essence, namely intentionality, the minimum level of reduction in psychoanalysis is an intentional act—instinct, a psychic act that intends pleasure.

However, the implications of the definition of consciousness as intentionality militate against the psychoanalytic notion of the unconscious. If consciousness is always consciousness of something, then what appears to consciousness is all of consciousness; there cannot, by definition, be an intentional act below the level of consciousness. To speak of "unconscious" acts is therefore a contradiction if the intentionality of these acts is to be preserved. If intentionality is not to be preserved, there is no contradiction; but then, of course, psychoanalysis would slip below the acceptable level of reduction. Thus Freud, according to existential philosophy, did not avoid the contradiction of placing the *res cogitans* in the *res extensa*, however brilliantly and extensively he refined his definitions of instinctual meaning.

BEING-IN-THE-WORLD

The issue of intentionality springs directly from the work of Husserl, the founder of phenomenology. Husserl, as a pure phenomenologist, did not seek to discover anything about the natural world. He was concerned purely and simply with ascertaining the essence of phenomena as they appear to consciousness. The question whether these phenomena correspond to the natural and real world he left to the explanatory disciplines of philosophy and science. Relations to the real human self he left to psychology. His student Heidegger, however, was interested in that oldest of philosophical questions, the nature of Being. For Heidegger, the essential structure of human being turns out to be an extension of the concept of intentionality. Just as consciousness is defined as consciousness-of, human being is characterized by Heidegger as being-in-the-world. The hyphens are deliberate; they represent an effort to undercut the subject-object split. Just as consciousness is not a separate entity that subsequently relates to objects, so man is not a separate being who then encounters his world. Rather, he is essentially in-the-world, he is his disclosure of world.

One essential that differentiates this basically human mode of being from that of the objectively known world is, for Heidegger, the element of possibility. The essence of man is always his possibilities, which he "has" in a more inclusive sense than the way an object has properties. An account of the factual content of an object can never express the essence of man, because that essence has yet to be his being as his own. Human time and space differ from "objective" time and space in that they are essentially related to man's determination of himself and his world. The essence of man, for Heidegger, is his appropriation of his essence, his making it his own. Thus the categories that describe human being are not qualities, but matrices within which qualities are to be appropriated.

We have noted how psychoanalysis seeks to place the *res cogitans*, or intentionality, into the *res extensa*, or sphere of pure corporeality. Existentialists object to the way in which psychoanalytic theory attempts to give man an essence in the way that an object has an essence. For psychoanalysis instinct, or libido, constitutes a residue that is taken a priori as the irreducible limit of investigation. Man has instinct as an object has its essence. It is not to the credit of psychoanalysis that this instinct is an intentionality of a sort; the existentialists maintain that the whole notion of intentionality is perverted when one particular class of intentional acts is singled out a priori as the basis of all classes. The meaning of psychic acts, intentions, is to be arrived at on their own terms, phenomenologically; they are not to be given meaning, as are the objects of natural science. Since, for the existentialists, human existence precedes essence, the task of an existential psychoanalysis must be, in Sartre's words, to uncover in each individual "a veritable irreducible; that is, an irreducible which would not be presented as the postulate of

the psychologist and the result of his refusal or his incapacity to go further…. This demand … is based on the refusal to consider man as capable of being analyzed and reduced to original givens, to determined desires (or 'drives'), supported by the subject as properties [are] by an object" (*Being and Nothingness*, pp. 560–561). The task of existential psychoanalysis is to apprehend the essence of each individual's life and world. If existence precedes essence, the analyst must apprehend the matrix within which essence is yet to be determined in each individual. Sartre calls this matrix the original choice or original project; Binswanger calls it the transcendental category that is the individual's mode of being-in-the-world.

Thus the critique of Freudian theory offered by existential psychoanalysis differs from that of the various revisionists in that it questions the theoretical root of all major movements in contemporary psychology; the assumption that the study of man can be wholly a natural science, that the notion of *homo natura* (man as a creature of nature) most fully expresses the essence of human being. The practical implications for psychiatry involve, among other things, wresting the concepts of mental health and illness away from analogies with purely somatic medicine, and thereby redefining the overall goal of any psychotherapy.

See also Being; Binswanger, Ludwig; Buber, Martin; Cartesianism; Consciousness; Existentialism; Freud, Sigmund; Heidegger, Martin; Husserl, Edmund; Psychoanalysis; Psychoanalytic Theories, Logical Status of; Sartre, Jean-Paul.

Bibliography

SOURCES

Binswanger, Ludwig. *Ausgewahlte Vorträge und Aufsätze*. 2 vols. Bern: Francke, 1947–1955.

Binswanger, Ludwig. *Ausgewählte Werke in vier Bänden*. Heidelberg: R. Asanger Verlag, 2004.

Binswanger, Ludwig. *Drei Formen missglückten Daseins*. Tübingen: Niemeyer, 1956.

Binswanger, Ludwig. *Traum und Existenz* (with an introduction by M. Foucault). Bern: Gachnang & Springer Verlag, 1992.

Binswanger, Ludwig. *Grundformen und Erkenntnis menschlichen Daseins*. Zürich: Niehans, 1942; 2nd ed., 1953.

Binswanger, Ludwig. *Schizophrenie*. Pfullingen, 1957.

Boss, Medard. *The Analysis of Dreams*. Translated by A. J. Pomerans. New York: Philosophical Library, 1958.

Boss, Medard. *Existential Foundations of Medicine and Psychology*. New York and London: Jason Aronson, 1979.

Boss, Medard. *Grundriss der Medizin: Ansze zu einer phänomenologischen Physiologie, Psychologie, Pathologie, Therapie und zu einer daseinsgemässen Präventiv-Medizin in der modernen Industrie-Gesellschaft*. Bern: Benteli, 1971.

Boss, Medard. *Meaning and Content of Sexual Perversions*. Translated by L. L. Abell. New York: Grune and Stratton, 1959.

Boss, Medard. *Psychoanalyse und Daseinsanalytik*. Bern: Huber, 1957. Translated by Ludwig B. Lefebre as *Psychoanalysis and Daseinsanalysis*. New York: Basic, 1963.

Cohn, I. *Existential Thought and Therapeutic Practice*. London: Sage, 1997.

Cohn, I. *Heidegger and the Roots of Existential Theory*. London: Continuum, 2002.

Fromm, Erich. *The Art of Loving*. New York: Harper & Row, 1989.

Keen, E. *A Primer in Phenomenological Psychology*. New York: Holt,Reinhart & Winston, 1975.

May, Rollo. *The Courage to Create*. New York: Bantam, 1984.

May, Rollo. *The Discovery of Being*. New York: W. W. Norton, 1983.

May, Rollo, Ernest Angel, and Henri F. Ellenberger, eds. *Existence*. New York: Basic, 1958. Translations of important works by Binswanger, Erwin Straus, Minkowski, V. E. von Gebsattel, and Kuhn.

Merleau-Ponty, Maurice. *Phénomenologie de la perception*. Paris: Gallimard, 1945. Translated by Colin Smith as *Phenomenology of Perception*. London: Routledge, 1962.

Sartre, Jean-Paul. *Esquisse d'une théorie des émotions*. Paris: Hermann, 1939. Translated by B. Frechtman as *The Emotions*. New York: Philosophical Library, 1948.

Sartre, Jean-Paul. *L'être et le néant*. Paris: Gallimard, 1943. Translated by Hazel Barnes as *Being and Nothingness*. New York: Philosophical Library, 1956.

Sartre, Jean-Paul. *L'imaginaire: Psychologic phénomenologique de l'imagination*. Paris, 1940. Translated by B. Frechtman as *The Psychology of the Imagination*. New York, 1948.

Speigelberg, H. *Phenomenology in Psychology and Psychiatry*. Evanston: Northwestern University Press, 1992.

Straus, Erwin. *The Primary World*. Translated by Jacob Needleman. New York: Free Press of Glencoe, 1963.

Valle R. J., and M. King, eds. *Existential-phenomenological Alternatives for Psychology*. New York: Oxford University Press, 1978.

van Deurzen, Emmy. *Everyday Mysteries—Existential Dimensions of Psychotherapy*. London: Routledge 1997.

Yalom, I. *Existential Psychotherapy*. New York: Basic Books, 1980.

A helpful annotated bibliography is available at http://www.strath.ac.uk/counsunit/features-biblio.html.

CRITICAL STUDIES

Lyons, Joseph. *Psychological Reports*, Monograph Supplement 5. 1959. Extensive bibliography of works in English.

Needleman, Jacob. *Being-in-the-World*. New York: Basic, 1963. The first half is a comparative and interpretive study of Binswanger and Freud.

Sonnemann, Ulrich. *Existence and Therapy*. New York: Grune and Stratton, 1954. A searching presentation, in a difficult prose style, of the issues and their relevance to the several schools of psychology. Bibliography.

Van den Berg, J. H. *The Phenomenological Approach to Psychiatry*. Springfield, IL: Thomas, 1955. A simple, clear, and quite elementary introduction.

Jacob Needleman (1967)
Bibliography updated by Thomas Nenon (2005)

EXISTENTIAL PSYCHOANALYSIS [ADDENDUM]

Toward the conclusion of *Being and Nothingness* Jean-Paul Sartre proposes "existential psychoanalysis" as an alternative to the Freudian version that he had criticized. Respecting individual freedom and responsibility, its basic principle is that "man is a totality and not a collection" of drives and complexes. By a comparative hermeneutic of the complex, multilevel symbolic expressions of an agent's actions it uncovers bad faith and ferrets out that fundamental project that gives unity and direction to our lives. It thereby renders possible "conversion" to an authentic existence, in which one can resist the need to create a substantialized self. Rejecting the hypothesis of an unconscious, this method relies heavily on prereflective consciousness and the distinction between what we prereflectively comprehend and reflectively know. The analyst's empathetic understanding helps bring this comprehension to knowledge. Since all consciousness is practical, for Sartre, this transformation involves a reorientation of one's way of being-in-the-world. Unfortunately, he concedes, "this psychoanalysis has not yet found its Freud."

It has been observed that Sartrean analysis deals with a set of human needs that have nothing to do with Freudian drives, namely "relational needs for holding, mirroring, positive regard, and emotional responsiveness, and needs for the development of a coherent and flexible sense of 'self'" (Cannon 1991, p. 1). Sartre's later work underscores the enabling power of the third to mediate our identity and social efficacy. In social ensembles some third parties objectify and alienate, whereas others generate practical unity and effectiveness. Sartre argues that consciousness is intentional, not only in the traditional phenomenological sense that it constitutes objects, but in the existential sense that it sets goals and strategies to obtain them. Psychoanalytic interpretation unmasks these goals and strategies.

Sartre has been accused of writing as if the human were born adult. But, in trying to understand Flaubert, he claims that "everything took place in childhood" and devotes several volumes of *The Family Idiot* to chart the "spiral of personalization" by which the individual interiorizes and reexteriorizes the structural relations into which it is born (its protohistory), the experiences of infancy and early childhood (its prehistory), and resultant "constitution" of its character and dispositions. At each stage the individual is in process of "making something out of what he or she has been made into." Sartrean metatheory respects the objective possibilities of this individual's situated being as it establishes "the way in which the child lives his or her family relations inside a given society." The approach is familial.

Sartrean psychoanalysis assumed a social orientation in the 1950s and 1960s, as did his thought generally. It became a form of social critique, linked to the thesis that entire societies could suffer from "objective neurosis," making it extremely difficult for its members to live authentic lives. His psychobiographies were prolonged instances of existentialist psychoanalysis conjoined to a kind of historical materialism, which he called the "progressive-regressive method."

Although the concepts of bad faith and authenticity have entered popular discourse, Sartre's psychoanalytic metatheory has yet to be adopted by professionals the way Heideggerian categories have been employed by Ludwig Binswanger, Medard Boss, and others.

See also Bad Faith; Binswanger, Ludwig; Consciousness; Freedom; Freud, Sigmund; Historical Materialism; Psychoanalytic Theories, Logical Status of; Sartre, Jean-Paul; Unconscious.

Bibliography

Bugental, J. F. *The Search for Authenticity: An Existential-Analytic Approach to Psychotherapy.* New York: Irvington, 1981.

Cannon, B. *Sartre and Psychoanalysis: An Existentialist Challenge to Clinical Metatheory.* Lawrence: University Press of Kansas, 1991.

Frankl, V. E. *Psychotherapy and Existentialism: Selected Papers on Logotherapy.* New York, 1985.

Laing, R. D. *The Divided Self.* New York: Routledge, 1999.

Sartre, J.-P. *L'idiot de la famille.* 3 vols. Paris: Gallimard, 1971–1972. Translated by Carol Cosman as *The Family Idiot.* 5 vols. Chicago: University of Chicago Press, 1981–1993.

Van Kaam, A. *Existential Foundations of Psychology.* Garden City, NY: Image, 1969.

Thomas R. Flynn (1996)

EXPERIENCE

As the average person understands the term *experience*, it means no more than familiarity with some matter of practical concern, based on repeated past acquaintance or performance. The experienced doctor or soldier knows his trade, not by the book merely, but by long practice under a variety of circumstances. The older philosophical meaning of the word differs but little from this, denoting as it does the capacity to do something, learned in the habit of doing it and guided rather by rule-of-thumb precept than by theoretical understanding (cf. the well-known passage in Aristotle, *Posterior Analytics* II, 19). It is in this fashion—by retention of individual memories and their gradual hardening into principle—that the craftsman acquires his skill, the scientist his knowledge, and the practical man his wisdom. But (save in the last case, perhaps) it represents at best only a stage on the way to real understanding in terms of universals and is thus by most ancient writers despised as a makeshift and uncertain form of knowledge. The mere empiric cuts, and continues to cut, a poor figure even into the modern period, though by that time associated, rather (as in Francis Bacon), with the trial-and-error experimenter in alchemy or medicine who endeavors by persistence alone to filch Nature's secrets without first gaining insight into its laws. The preference for rational certainty over mere empirical generalization is in fact endemic among philosophers and can be seen not merely in avowed rationalists (such as René Descartes), but also, for example, in those whose uneasiness over the principle of induction has led them to seek ways of validating it as the major premise for a quasi-deductive treatment of science.

The experience from which the empiric draws his conjectures is, of course, the homely and substantial experience of a world of public objects, which forms for all sane and unreflective persons the basis of ordinary life. It has been regularly insisted, however, since the earliest times, that experience in this sense is nothing ultimate: the alleged paradoxes of motion and change and the more familiar facts of perceptual error and illusion are enough (it is thought) to show that it cannot be straightforwardly identified with the real. Hence, in addition to the rejection of habit-learning as a road to knowledge, there arises that further prejudice against the deliverances of the senses and in favor of necessary reasoning from first principles, of which the Parmenidean distinction between the "ways" of truth and opinion is an early and famous example.

THE "GIVEN"

The uncertainty of sense experience leads, by this route, to a further important conclusion. Since perceptual illusion and mistake seem essentially to be the fault of the observer, he must himself contribute something to his experience by way of inference, interpretation, or construction. Experience must, in part at least, be the work of the mind. For all that, the individual certainly does not create or invent his experience and in certain respects is powerless to alter it at will, it seems, therefore, equally undeniable that some part of it is simply "given" and is only thereafter subject to adulteration by its recipient. This given is generally referred to as the object of "bare" or "immediate" experience, in contrast to the more "solid" or developed experience of which it is held to be an essential ingredient. The legitimacy of the contrast is seldom, indeed, disputed, for though immediate experience has often been denounced as a myth, the usual motive for doing so has been to stigmatize it as a mere abstraction got by analysis and not something that could occur, psychologically, by itself. *All* experience, on this view, involves interpretation, and it is thus senseless to suppose any unvarnished, direct acquaintance with the given. But since it would be equally senseless to suppose an interpretation with nothing to interpret, it is commonly admitted that an "epistemic" given must nonetheless be present in experience, though impossible to view independently, since this would ipso facto be to construe it in some fashion under the auspices of thought.

For writers who accept either a psychological or an epistemic given, a number of problems arise. What does it consist of? What marks or features does it exhibit? How is it related to the everyday experience built upon it? And how, once the latter has been derived, is it possible to proceed from there to the realities that presumably underlie, occasion, and explain the whole? The last is essentially a metaphysical question, but the remainder (to which we here confine ourselves) are staple topics of epistemological dispute.

As to the content of immediate experience, there are characteristic differences of opinion. At one extreme lie the theories of direct realism, whose claim is that material objects are immediately given, so that no real difference arises between naked and clothed experience, sensation and perception, or for that matter appearance and reality; apart from perceptual error there is thus no "problem of knowledge" at all. At the opposite pole are the theories of William James and F. H. Bradley, for whom immediate experience presents only an undifferentiated mass of feeling or sensation in which nothing is discriminated or

related and in which even the contrast between subject and object has not begun to appear. Of this "blooming, buzzing confusion" (in James's well-known phrase), it is obvious that nothing can be said, even to distinguish its modes. Ex hypothesi, it is merely the residue left after elimination of all processes involving association, memory, judgment, thought, and language; it is free from error because it says nothing; but as such, however indispensable, it has little to contribute to knowledge.

Writers in the empiricist tradition, less anxious to make knowledge the sole work of the mind, have been correspondingly averse to such unstructured versions of the experiential given, though the argument from illusion has equally deterred the majority from claiming direct acquaintance with a world of things. For most of them the given includes at least simple sense-qualities of color, taste, sound, and so on, together with organic sensations, feelings, and images, it being generally assumed that these are presented individually and even as "atomic" cognitive units (impressions, ideas, sense data, etc.) to a consciousness distinct from themselves. Beyond this, however, there is little agreement. Some have asserted, while others have denied, that spatial, temporal, causal, or other relations are given in this fashion. Some have been prepared to admit associative or sign material as part of the given; others have not. Visual impressions, for example, have been held (most notably by George Berkeley) to be initially two-dimensional; but in other writers (for example, H. H. Price) the claim is that they are presented in, or as having, depth. Images, memory impressions, and feelings have all been ascribed to the given, but again it is disputed as to whether all images belong in this category or only those of simple qualities (as with impressions of sense). There is similar contention as to whether the "pastness" of the memory image is intrinsic to it or imputed on the strength of some other feature, and whether, in general, the "inner" and "outer" character of feelings and images, on the one hand, and sense impressions on the other, are marks of the data themselves, or a construction imposed upon them.

CRITERIA OF "GIVENNESS"

The foregoing differences and others like them are not, as they seem, due to want of regard for the facts, nor would closer attention suffice to dispel them. They arise from a failure in agreement as to the formal characteristics of "givenness" itself, and hence as to the criteria for its identification. What are these criteria? The commonest answers seem to be that the given is private; that it is adventitious (in Descartes's sense); that it is simple, as involving no element of thought or inference; and that it is incorrigible. Too often these tests are also assumed to coincide in yielding the same result or to be sufficient rather than necessary conditions for givenness. Many of the historic uncertainties surrounding the description of the given would appear to have arisen in this way.

Privacy, for example, is inconsistent with simplicity, inasmuch as every variety of thought and feeling, imagery, or sensory seeming is necessarily private, however obvious it may be that it belongs to a sophisticated rather than a primitive level of consciousness. Adventitiousness, as a criterion, would similarly include within the given all phenomena not under the subject's voluntary control, including the appearance and causal behavior of objects, but excluding some part of his thoughts and feelings—though how much it seems impossible to say. In both cases the given seems too generously defined to serve as a foundation for knowledge. If the given is limited to experience uncontaminated by "inference," the difficulty is to know what counts as such and hence where analysis is to stop. Even the lowliest amoeba can react to sensory cues and so "transcend the given," but who is to say if it thinks or not?

The psychological given, on this showing, may well be accounted a myth. The epistemological arguments, however, are harder to put aside. Their main support has been the belief that the data they point to represent the only foundations for knowledge that could be called certain or incorrigible. The judgments of perceptual experience, concerning the existence, properties, and relations of objects, are all (it is said) subject to error, and so open to correction. But reports of what seems merely, of the presence of sense qualities to consciousness, make no such claim about real existence and so run no risk of mistake. The "sense datum" theories popular from 1900 onward, like their eighteenth-century predecessors, attempted to erect on this basis a theory of knowledge more stable and concrete than that underwritten by the necessary truths of rationalism. Their vogue has declined, however, and that for two reasons. On the one hand it has been argued that no sense datum statement, however guarded, can fail (if it says anything at all) to make some conceptual commitment that might later call for retraction; and on the other, that a great many factual statements are quite as certain as those they are alleged to depend on. Given a sufficiently straightforward case of perceptual contact with an object, there is no ground for treating it as a judgment based on the "evidence" of sense data, since it is beyond the power of any future evidence to enforce its withdrawal. It is as certain as anything could

be, and nothing is gained, therefore, by an appeal to supposedly more primitive certainties to provide it with support.

ORDINARY EXPERIENCE

The above argument is not beyond question, but it illustrates the difficulties not merely of characterizing the sensory given but even of securing agreement as to its existence. It is more profitable, perhaps, to turn for a moment to those writers who, having accepted a given of some sort from the outset, have occupied themselves with the second stage of the problem, namely, the manner in which this given is elaborated into the fullness of ordinary experience. Here the issue lies chiefly between those who maintain (with John Locke, and still more, with Étienne Bonnot de Condillac) that the concepts employed in the construction of developed experience are themselves derived (by abstraction, association, composition, or induction) from immediate experience, and those who argue (as do all rationalists) that this elaboration depends on principles contributed by the mind a priori, and not first learned from experience itself. The rationalist does not thereby deny the fact of immediate experience any more than does the empiricist. His claim, rather, is that this experience does not make itself intelligible by any natural process and that it is only the logical activity of the mind that brings order and coherence into the result.

The most celebrated statement of this position is doubtless that of Immanuel Kant, for whom the "manifold" of sensory intuition, though spatiotemporally ordered insofar as it is presented at all, is unified into a world of empirical objects only insofar as it is brought under the a priori rules, or categories, of the understanding. Experience in the full sense is thus a synthesis, part given and part made, though in some of Kant's idealist successors the creative aspect is so far emphasized at the expense of the given as to tend toward that extreme of rationalism in which the world of experience is construed as an exclusively mental product, with no element of "brute fact" in it at all. For the modern "logical empiricist" the position is, in effect, reversed, his typical doctrine being (as already noticed) that the content of all empirical propositions can be reduced without remainder to "protocols" recording actual or possible fragments of immediate experience.

But the attempt to reconstitute ordinary experience out of a mixture of sense data and formal logic, though long and ably sustained, has met in the end with little more success than the search for certainty that led to the introduction of these data in the first place. Recent work on the subject has shown signs of impatience with this starting point and seeks to discredit it by attacking the whole distinction between sensation and perception—the two-level theory of experience—and the argument from illusion on which that distinction so largely depends. Whether this rejection of the traditional premises of the problem offers any hope of solving (or dissolving) it is a question that time alone can answer. If experience teaches anything, it is that success is unlikely; but then, even in philosophy (or so philosophy tells us), experience is not always or necessarily a reliable guide.

See also Aesthetic Experience; A Priori and A Posteriori; Aristotle; Bacon, Francis; Berkeley, George; Bradley, Francis Herbert; Consciousness; Descartes, René; Empiricism; Induction; James, William; Kant, Immanuel; Perception; Religious Experience; Sensa.

Bibliography

The literature on this subject is endless. The following is a list of representative works by Anglo-American authors, published in the twentieth and early twenty-first centuries.

Audi, Robert. "Experience and Inference in the Grounding of Theoretical and Practical Reasons: Replies to Professors Fumerton, Marras, and Sinnott-Armstrong." *Philosophy and Phenomenological Research* 67 (1) (2003): 202–221.

Austin, J. L. *Sense and Sensibilia*. Oxford: Oxford University Press, 1962.

Ayer, A. J. *Foundations of Empirical Knowledge*. London: Macmillan, 1940.

Ayer, A. J. *Problem of Knowledge*. London: Macmillan, 1956.

Bonevac, Daniel. "Sellars vs. the Given." *Philosophy and Phenomenological Research* 64 (2002): 1–30.

Brewer, Bill. *Perception and Reason*. Oxford: Clarendon Press, 2002.

Chisholm, Roderick. "The Problem of the Speckled Hen." *Mind* 51 (1942): 368–373.

Crane, T., ed. *The Contents of Experience*. Cambridge, U.K.: Cambridge University Press, 1992.

Hamlyn, D. W. *Sensation and Perception*. London: Routledge and K. Paul, 1961. A history of the problem.

Hirst, R. J. *The Problems of Perception*. London: Allen and Unwin, 1959.

Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.

Lewis, C. I. *Analysis of Knowledge and Valuation*. La Salle, IL: Open Court, 1946.

Lewis, C. I. "The Given Element in Empirical Knowledge." *Philosophical Review* 61 (1952): 168–175.

Lewis, C. I. *Mind and the World-Order*. New York: Scribners, 1929.

Moore, G. E. *Philosophical Studies*. London: Routledge, 1922.

Moore, G. E. *Some Main Problems of Philosophy*. London: Allen and Unwin, 1953.

Moser, Paul. "The Given in Epistemology." *Communication and Cognition* 18 (1985): 249–262.

Pasch, A. *Experience and the Analytic*. Chicago: University of Chicago Press, 1958.
Price, H. H. *Perception*. London: Methuen, 1932.
Price, H. H. *Thinking and Experience*. London: Hutchinson, 1953.
Russell, B. A. W. *Our Knowledge of the External World*. Chicago: Open Court, 1914.
Ryle, G. *Concept of Mind*. London: Hutchinson, 1949.
Sellars, Wilfrid. *The Metaphysics of Epistemology*. Atascadero, CA: Ridgeview Press, 1989.
Sosa, Ernest. "Mythology of the Given." *History of Philosophy Quarterly* 14 (1997): 275–286.
Stace, W. T. *Theory of Knowledge and Existence*. Oxford: Clarendon Press, 1932.
Walsh, W. H. *Reason and Experience*. Oxford: Clarendon Press, 1947.
Williams, Michael. "Are There Two Grades of Knowledge? Mythology of the Given: Sosa, Sellars, and the Task of Epistemology." *PAS*, Supp. 77 (2003): 91–112.
Yolton, J. W. *Thinking and Perceiving*. La Salle, IL: Open Court, 1962.

P. L. Heath (1967)
Bibliography updated by Benjamin Fiedor (2005)

EXPERIMENTATION AND INSTRUMENTATION

See Appendix, Vol. 10

EXPLANATION

The three cardinal aims of science are prediction, control, and explanation, but the greatest of these is explanation. It is also the most inscrutable: Prediction aims at truth, and control at happiness, and insofar as one has some independent grasp of these notions, one can evaluate science's strategies of prediction and control from the outside. Explanation, by contrast, aims at scientific understanding, a good intrinsic to science, and therefore something that it seems one can only look to science itself to explicate.

Philosophers have wondered whether science might be better off abandoning the pursuit of explanation. Pierre Duhem (1954), among others, argued that explanatory knowledge would have to be a kind of knowledge so exalted as to be forever beyond the reach of ordinary scientific inquiry: it would have to be knowledge of the essential natures of things, something that neo-Kantians, empiricists, and level-headed practitioners of science could all agree was neither possible nor perhaps even desirable.

Everything changed when Carl Gustav Hempel formulated his deductive-nomological account of explanation. In accordance with the previous observation, that one's only clue to the nature of explanatory knowledge is science's own explanatory practice, Hempel proposed simply to describe what kind of things scientists tendered when they claimed to have an explanation, without asking whether such things were capable of providing true understanding. Since Hempel, the philosophy of scientific explanation has proceeded in this humble vein, seeming more like a sociology of scientific practice than an inquiry into a set of transcendent norms. In keeping with its mission as a branch of philosophy, however, the study of explanation pursues a particular kind of sociological knowledge: It is concerned almost exclusively with the ideal at which scientists take themselves to be aiming, and barely at all with the steps and missteps taken on the way to realizing the ideal.

As Hempel saw it, scientific explanation was of a piece with prediction, requiring the same resources and giving a similar kind of satisfaction. No doubt this modest view of the explanatory enterprise played a part in making the study of explanation acceptable in the climate of postwar empiricism. The story of explanation in decades since Hempel's time, however, is an expansionist one. Over the years philosophers of explanation have gradually required more resources for, and made grander claims for the significance of, explanation's role in science. (For a comprehensive overview of the philosophy of explanation from 1948 to 1988, with a full bibliography, see Wesley C. Salmon [1990].)

THE DEDUCTIVE-NOMOLOGICAL ACCOUNT

Hempel's deductive-nomological (DN) account (Hempel and Oppenheim 1948) is intended to capture the form of any deterministic scientific explanation of an individual event, such as the expansion of a particular metal bar when heated, the extinction of the dinosaurs, or the outbreak of the U.S. Civil War.

According to Hempel such an explanation is always a deductive derivation of the occurrence of the event to be explained from a set of true propositions including at least one statement of a scientific law. (The event to be explained is called the explanandum; the set of explaining statements is sometimes called the explanans.) In other words, a deterministic event explanation is always a sound, law-involving, deductive argument with the conclusion that the explanandum event occurred.

Intuitively, the premises of a DN explanation spell out the relevant initial and background conditions, and the laws governing the behavior of the system in which the explanandum occurred. For example, Hempel cites the following argument as a typical DN explanation of the event of a thermometer's mercury expanding when placed in hot water:

The (cool) sample of mercury was placed in hot water, heating it,

Mercury expands when heated thus

The sample of mercury expanded

Because the law or laws that must be cited in a DN explanation typically cover the pattern of behavior of which the explanandum is an instance, the DN account is sometimes referred to as the covering law account of explanation.

One can see that the DN account is not intended to give the form of probabilistic event explanations; Hempel offers a separate account of probabilistic explanation, which will be discussed later on. The explanation of phenomena other than events is, by contrast, apparently amenable to the DN approach. Hempel suggests that a scientific law can be explained, for example, much like an event, by deducing it from premises including at least one other law. However, he finds himself unable to make good on this proposal, for reasons connected to the relevance problem discussed in the next section.

Many scientific explanations of events and other phenomena undoubtedly have the form proposed by the DN account: They are logical derivations from laws and other information. Hempel does not entirely satisfy himself, however, with answering questions of form. Taking one step beyond sociological humility, he advances a thesis as to why deductive, law-involving arguments should confer understanding, "[A DN explanation] shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred" (1965a, p. 337, emphasis in the original).

Scientific understanding, then, takes the form of retrospective expectation: One might say (loosely) that, whereas prediction is concerned with what one should expect in the future, explanation is concerned with what one should have expected in the past. Explanation is, then, put on a par with prediction and so made safe for empiricist philosophy of science. Hempel even goes so far as to say that the difference between explanation and prediction is merely pragmatic (Hempel and Oppenheim 1948) though the DN account does not in itself entail such a thesis.

OBJECTIONS TO THE DN ACCOUNT

Three kinds of objections to the DN account have been especially important for the subsequent development of the philosophy of explanation.

The first kind of objection, developed by Henry Kyburg, Salmon, and others, points to the DN theory's inability to account for judgments of explanatory relevance. The paradigm is the following argument, which satisfies all the DN account's criteria for a good explanation of the event of a particular teaspoon of salt's dissolving:

The teaspoon of salt was hexed (meaning that certain hand gestures were made over the salt),

The salt was placed in water,

All hexed salt dissolves when placed in water thus,

The salt dissolved.

The explanation appears to attribute the salt's dissolving in part to its being hexed, when in fact the hexing is irrelevant.

There are various responses to the counterexample that aim to preserve as much of the DN account as possible, for example, holding that the generalization about hexed salt is not a true law or imposing the requirement that a DN explanation use the most general law available.

Salmon's much less conservative reaction is to conclude that Hempel is wrong to think of explanation in terms of expectability, therefore of explanations as kinds of argument. The relation between the factors cited in an explanation and the explanandum itself, Salmon holds, is not epistemic, but ontic; it should be a physical relevance relation—a relation of statistical relevance, he first proposes (1970), or a relation of causal relevance, as he later comes to believe (1984). The faulty explanation of the salt's dissolving is to be discarded, argues Salmon, not because of some formal or logical defect, but because it cites an event, the hexing of the salt, that fails to bear the appropriate relevance relation to the explanandum.

Hempel himself declines (early in his career, at least) to give a DN account of the explanation of laws because of a related problem. Kepler's laws may be derived from a

single law that is simply the conjunction of Kepler's laws and Boyle's law. Such a derivation is clearly no explanation of Kepler's laws, writes Hempel, yet it satisfies the DN account's requirements: The premises are true and the argument is valid and law involving (Hempel and Oppenheim 1948).

The second important objection to the DN account is perhaps also the most famous. It shows, most philosophers would agree, that the DN account pays insufficient attention to the explanatory role of causal relations.

The height of a flagpole can be cited, along with the position of the sun and the law that light travels in straight lines, to explain the length of the flagpole's shadow. The DN account is well able to make sense of this explanation: It can be cast in the form of a sound, law-involving argument. However, now take this same argument and switch the premise stating the height of the flagpole with the premise stating the length of the shadow. One now has a sound, law-involving argument for the height of the flagpole that cites, among other things, the length of the shadow—thus, according to the DN account, one has an explanation of the height of the flagpole that cites, as an explainer, the length of the shadow. This consequence of the DN account—that the height of a flagpole can be explained by the length of its shadow—seems obviously wrong, and it is wrong, it seems, because a cause cannot be explained by its own effects.

A further famous example strongly suggests that effects can only be explained by their causes and the laws and background conditions in virtue of which they are causes. Suppose that the arrival of a certain kind of weather front is always followed by a storm and that a certain reading on a barometer is a sure sign that such a front has arrived. Then a barometer reading of this sort is always followed by a storm. The storm cannot be explained, however, by citing the barometer reading and that such readings are always followed by storms, though these two facts together satisfy the requirements of the DN account. A constant, robust correlation is not, it appears, enough for explanation. What is needed, as Salmon eventually concludes, is a causal relation.

At first Hempel resists the suggestion that facts about causation play any special role in explanation (e.g., see 1965a, §2.2). Over the years, however, due in part to the development of sophisticated empiricist accounts of causation, this has become a minority view.

The third class of objections to the DN account focuses on the account's requirements that every explanation cite a law and that (except in probabilistic explanation) the law or laws be strong enough to entail, given appropriate boundary conditions, the explanandum. One way to develop the objection is to point to everyday explanations that cite the cause of an event as its explanation, without mentioning any covering law, as when one cites a patch of ice on the road as the cause of a motorcycle accident.

More important for the study of explanation in science are varieties of explanation in which there is no prospect and no need for either the entailment or the probabilification of the explanandum. Perhaps the best example of all is Darwinian explanation, in which a trait *T* of some species is explained by pointing to the way in which *T* enhanced, directly or indirectly, the reproductive prospects of its possessor. Attempting to fit Darwinian explanation into the DN framework creates a host of problems, among which the most intractable is perhaps the following (Scriven 1959): For every trait that evolved because it benefited its possessors in some way, there are many other, equally valuable traits that did not evolve, perhaps because the right mutation did not occur, perhaps for more systematic reasons (e.g., the trait's evolution would have required a dramatic reconfiguration of the species' developmental pathways). To have a DN explanation of *T*, one would have to produce a deductive argument entailing that *T*, and none of the alternatives, evolved. In other words, one would have to be in a position to show that *T* had to evolve. Not only does this seem close to impossible but also it seems unnecessary for understanding the appearance of *T*. One can understand the course of evolution without retrospectively predicting its every twist and turn.

Hempel is aware of the problem with Darwinian explanation. His response is to argue that there is no such thing: Faced with a choice between the DN account and Darwinian explanation, one should opt for the former and consider Darwinian stories to be at best partial explanations of traits (Hempel 1965c). He advocates a similar deflationary treatment of functionalist explanation in sociology and of historical explanations that are not entailments.

THE INDUCTIVE-STATISTICAL ACCOUNT

Hempel's (1965a, §3) account of the probabilistic explanation of events, the inductive-statistical (IS) account, in many ways parallels the DN account of deterministic event explanation. Like a DN explanation, an IS explanation is a law-involving argument giving good reason to

expect that the explanandum event occurred. However, whereas a DN explanation is a deductive argument entailing the explanandum, an IS explanation is an inductive argument conferring high probability on the explanandum.

Hempel's example is the explanation of John Jones's swift recovery from a strep infection. The probability of a swift recovery without the administration of penicillin, Hempel supposes, is 0.1, while the probability with penicillin is 0.9. Citing Jones's infection, his treatment with penicillin, and the resulting high probability of recovery, then, confers a high probability on Jones's swift recovery; in the circumstances, one would expect Jones to recover swiftly. This inductive argument is sufficient, in Hempel's view, to explain the swift recovery.

Inductive soundness imposes one additional requirement with no parallel in deductive logic. Suppose one knows that Jones's strain of strep is resistant to penicillin. An inductive argument is said to be sound only if all relevant background knowledge is taken into account; consequently, an inductive argument for Jones's swift recovery must cite the infection's penicillin resistance. But once the new premise is added, the argument will no longer confer a high probability on its conclusion. This is what is wanted: There ought to be no inductive argument for swift recovery—one ought not to expect swift recovery—when the strep is known to be resistant.

Hempel imposes a similar requirement on IS explanations, which he calls the requirement of maximal specificity (for details, see Hempel 1965a, §3.4). In virtue of this requirement, it is not possible to explain Jones's swift recovery by citing treatment with penicillin when the infection is known to be penicillin resistant.

As with the DN account of explanation, a number of objections to the IS account have exerted a strong influence on the subsequent development of the philosophical study of explanation. Versions of both the relevance and the causal objections apply to the IS account as well as to the DN account. Two other important criticisms will be briefly described here.

The first is the complaint that it is too much to ask that explanations confer high probability on their explananda. In many ways, this is the analogue of the third objection to the DN account mentioned earlier; in the same paper that Michael Scriven (1959) expresses doubts about the existence of a DN treatment of Darwinian explanation, he describes the following example, best conceived of as an objection to the IS account. The probability that Jones contracts paresis, a form of tertiary syphilis that attacks the central nervous system, given that he has untreated secondary syphilis, is low. However, only syphilitics contract paresis. It seems reasonable to cite untreated syphilis, then, as explaining Jones's paresis, though the explanation confers only a low probability on the explanandum.

The proponent of the IS account is committed to rejecting such attempts at explanation, as Hempel does, arguing that in such cases one has only a partial explanation of why the patient contracted syphilis. This is perhaps one of the most convincing of Hempel's defenses, but the paresis example is nevertheless widely regarded as posing a serious problem for the expectability approach to explanation.

A second objection to the IS account focuses on the requirement of maximal specificity. The requirement insists that all relevant background knowledge must be included in a probabilistic event explanation, but it does not require that relevant but unknown information be taken into account. In particular, if Jones's infection is penicillin resistant, but this fact is not known to the explainer, then the IS account deems the explainer's appeal to the administration of penicillin as a perfectly good explanation of Jones's swift recovery.

As J. Alberto Coffa (1974) argues, this is surely not correct. If the infection is resistant to penicillin, then the administration of penicillin cannot explain the recovery, regardless of what the explainer does and does not know. The requirement of maximal specificity makes probabilistic explanation relative to the explainer's epistemic situation, then, in a way that it appears not to be. This objection hits right at the heart of the expectability conception of explanation, suggesting that explanation is not an epistemic matter in the least. A third objection that is applicable to many accounts of probabilistic explanation will be raised in the following discussion of the statistical relevance account.

THE STATISTICAL RELEVANCE ACCOUNT

In response to the DN account's relevance problem, Salmon suggests that the factors cited in an explanation must stand in a relation of statistical relevance (SR) to the explanandum. He does not intend this as a friendly amendment to Hempel's account, but as a radical reconceptualization of the nature of explanation: The function of an explanation, Salmon (1970) argues, is not to show that the explanandum was to be expected, but to describe factors—ideally, all the factors—statistically relevant to the occurrence of the explanandum.

From the beginning, statistical relevance is presented as an objective relation, that is, a relation holding independently of the explainer's background knowledge or other context. (Coffa's [1974] critique of the IS account, discussed earlier, discourages relativistic backsliding.) Salmon thus requires an account of probability that is both objective and broad enough to encompass any possible explanandum.

For breadth, he settles on frequentism, the view that the probability of an event type is equal to the frequency with which it occurs in a reference class of outcomes. For objectivity, he develops what he calls a homogeneity constraint on the reference classes that can be used as bases for explanatory probabilities. Such a constraint, he believes, is strong enough to determine a single, observer-independent probability distribution over any set of outcomes of interest. Salmon (1984) summarizes the theory of homogeneity; for further information, see the discussion of the reference class problem in the separate entry on probability and chance.

Statistical relevance is a comparative concept: To say that a factor *A* is statistically relevant to the occurrence of an event *E* is to say that the probability of *E* (or for the frequentist, of events of the same type as *E*) in the presence of *A* is greater than the probability of *E* in the absence of *A*. Thus, the determination of a relevance relation requires not only a reference class—a class of outcomes all of which occurred in the presence of *A*—but a contrast class, a class of outcomes all of which occurred in the absence of *A*. The contrast class is not normally homogeneous. Thus for Salmon, the contrast probability must be a weighted sum of different homogeneous probabilities, each corresponding to a different way that *A* might have been absent, and giving the probability of *E* when *A* is absent in that way.

Perhaps inevitably, if not inescapably, Salmon arrives at the view that a full SR explanation is a complete table of relevance, describing not only factors that are present and statistically relevant to the explanandum but also factors that are absent but would have been statistically relevant if they had been present. He further adds to the table all the alternatives to the explanandum *E* with respect to which there existed homogeneous probabilities, and a list of all the factors that would have been relevant to these alternatives, if they occurred. Consequently, the information proffered in an SR explanation of an event *E* not only explains the actual occurrence of *E* but would also explain any occurrence of an event of the same type, even if different relevant factors were present, as well as the occurrence of any alternative to *E*.

As something of a corollary to this view, Salmon holds that negatively relevant factors—factors that lower the probability of the explanandum—are as explanatory as positively relevant factors and that all factors should be mentioned regardless of their degree of relevance. Salmon's not discriminating among these factors is perhaps best understood as follows. Seeing that a factor is statistically relevant to the explanandum is an explanatory end in itself. That the factor makes a particular kind of change—positive or negative, large or small—to the total probability of the explanandum would be important only if appreciating the value of the total probability were also an explanatory end. However, it is not; knowing which relevance relations hold is all that matters.

Four objections to the SR account are considered here. First, for all Salmon's justifications, an SR account seems to contain too much information. To explain *E* when *A* was absent, why is it necessary to know that, had *A* been present, it would have been relevant? Why is it further necessary to know what would have been relevant to the occurrence of some alternative to a type *E* event that did not in fact occur? This information does not appear to be directly relevant to the explanatory task at hand, that of explaining *E* itself.

Second, the SR account seems vulnerable to the causal objection to the DN account; it seems to hold that *A* is explanatorily relevant to *E* whenever *A* is correlated with *E*, when in fact it is necessary that *A* be a cause of *E*. The barometer reading is statistically relevant to the storm in the example described earlier, but it does not thereby explain the storm.

Salmon is well aware of this problem and proposes that only certain kinds of statistical relevance relations are explanatory, namely, those that survive a screening off test. A factor *A* that is correlated with *E* is screened off from *E* by another factor *B* if, conditional on *B*, *A* makes no difference to the probability of *E* (just as, for example, conditional on the presence of the front, the barometer reading makes no difference to the probability of the storm), but conditional on *A*, *B* does make a difference to the probability of *E*. When there is some *B* that screens off *A* from *E*, Salmon says that *A* is not genuinely statistically relevant to *E*. And *A*'s relevance will indeed disappear in a relevance table that also cites *B*. Note that Salmon's treatment does not make an explicit appeal to causal facts. Whether all problems concerning the role of causation in explanation can be solved in this way is unclear.

A third objection dogs all the probabilistic accounts of explanation to be considered in this entry. Suppose that I strap a small but unreliable bomb to one wheel of

your car. The probability that the bomb detonates is 50 percent, in which case your tire goes flat. The trigger fails, but you drive over a nail and your tire does go flat. The bomb has increased the probability of the flat, but it plays no role in its explanation. (Does the presence of the nail screen off the presence of the bomb? No, if it is assumed that the nail's effect is, like that of the bomb, probabilistic.) Sometimes statistically relevant factors are explanatorily irrelevant. Finally, it is not easy to see how the SR account might be generalized to give an account of the explanation of phenomena other than events.

THE UNIFICATION ACCOUNT

Michael Friedman (1974) suggests that, while the logical empiricists' official account of explanation is the expectability account, they have an unofficial account, too, on which to explain a phenomenon is to see it as an instance of a broad pattern of similar phenomena. Hempel himself occasionally writes in this vein, "The understanding [explanation] conveys lies … in the insight that the explanandum fits into, or can be subsumed under, a system of uniformities represented by empirical laws or theoretical principles" (1965a, p. 488). Friedman formulates what he calls a unification account of explanation, a particularly global version of this conception of explanation as pattern subsumption, on which a phenomenon is explained by the system of subsuming laws that best unifies all the phenomena there are. Philip Kitcher (1981, 1989) amends and extends Friedman's account in various ways.

The unifying power of a theory is proportional, on both Friedman's and Kitcher's accounts, not only to the number of phenomena that can be subsumed under the theory but also to the simplicity of the theory. (Kitcher imposes some additional desiderata.) The theory that best unifies all the phenomena, then, might be said to yield the most for the fewest: The most derivable phenomena for the fewest number of basic principles. It is characteristic of the unificationist position to insist that only the most unifying theory has full explanatory power, but this view does not in itself preclude the possibility of partial explanation by more weakly unifying theories.

Why be a unificationist? Friedman suggests that the virtue of the most unifying theory is that it reduces to a minimum the number of fundamental incomprehensibilities, that is, unexplained explainers. Perhaps a more common justification for unificationism is that suggested by Hempel: To understand something is to fit it into a wider pattern. Add that the wider the pattern, the more powerful the explanation, and one is well on the way to unificationism.

Many of the virtues of the unification account stem from the great versatility of the pattern subsumption relation. A subsuming pattern need not be exceptionless, so not only probabilistic explanation but also other forms of nondeductive explanation fit the unification mold. Darwinian explanation, for example, can be seen as accounting for a trait by seeing it as part of a widespread pattern of adaptedness in the biological world—though Kitcher (1989, §5), for one, resists this view of evolutionary explanation, and indeed, argues that all explanations can be formulated as deductive arguments. More inclusively, Kitcher argues that unificationism supplies an effective account of mathematical, as well as scientific, explanation. For some further claimed advantages of the unification over the causal approach, see Kitcher (1989, §3).

Unificationism promises to give a powerful and subtle account of explanatory relevance. For example, an explanation of a teaspoon of salt's dissolving that cites the law "all hexed salt dissolves in water" is rejected as insufficiently unifying, because the law is both more complex and covers fewer phenomena than the law "all salt dissolves in water." More interesting, the unificationist can give an account of why many of the low-level details of the implementation of biological, psychological, economic, and social mechanisms seem to be irrelevant to understanding those mechanisms' behavior; the details, however, have yet to be worked out (Kitcher 1984).

Two important classes of objections stand in the way of the unification approach to explanation. First is the familiar question concerning the role of causation in explanation. Can the unification account explain why explanation so often, perhaps always, seems to follow the direction of causation? One might think not: The explanation of a flagpole's height in terms of the length of its shadow seems to cite just as unifying a pattern as the explanation of shadow length in terms of pole height—the same pattern, in fact.

Kitcher (1981) takes up the challenge, arguing that the unification account reproduces the asymmetries in explanation usually put down to something causal. On his view, a unifying pattern is an argument pattern. Since arguments have a direction, the pattern in which the pole height explains the shadow length is distinguished from the pattern in which the length explains the height. The unifying power of each must, therefore, be assessed separately. To solve the problem, the correct comparison is not between the unifying power of these two argument

patterns, but between the unifying power of the pattern that wrongly explains pole height in terms of shadow length and that of the pattern one usually cites to explain the height of a flagpole.

Kitcher calls this latter argument pattern an origin and development pattern and claims that it is instantiated by, and so subsumes, every account one gives of the properties of a thing that describes its origin and development, as when, for example, one tells the story of the construction and erection of the flagpole. The pattern is enormously general, then, and so easily wins the right to explain the height of the flagpole. Having argued, in effect, that unificationist explanation tends to proceed in the direction of causation, Kitcher then makes the dramatic claim that it is the order of explanation that determines the order of causation; one's causal beliefs depend on and reflect one's explanatory practice.

The second objection to explanatory unificationism is that it makes explanation an overly global matter. How one phenomenon is to be explained depends, according to the unificationist, on what best unifies all the other phenomena, therefore on what the other phenomena are. To many writers, it seems that finding an explanation does not require, even in principle, knowledge extending to all corners of the universe. A more moderate or local unificationism is possible, of course, but another natural place to look for locality is in the causal approach to explanation.

THE CAUSAL APPROACH

In 1965 Hempel could regard the idea that there is something causal to explanation over and above the exceptionless regularities cited by a DN explanation as lacking a "precise construal" (1965a, p. 353). Since that time philosophers have come to see claims about causal relations as having a rich empirical content that goes far beyond mere regularities and their instantiation (see Spirtes et al. [2000]), though the tradition began well before Hempel made his remark, with Reichenbach [1956]). Even metaphysical empiricists, then, can agree that there is a distinctive causal approach to explanation. Thanks to the development of sophisticated but wholly empiricist accounts of causation (again beginning with Reichenbach), they can go further and in good conscience endorse the causal approach.

Strong arguments suggest that the causal approach is correct. The first and most persuasive is the equation of causal and explanatory direction suggested by the flagpole/shadow and barometer/storm examples. The second is the observation that a requirement of causal relevance between explainers and the explained will deal with the problem of the hexed salt and similar cases. The third is that one can give a cause for a phenomenon without being able to predict it. In those counterexamples to the DN and IS accounts where grounds insufficient for prediction nevertheless seem to be sufficient for explanation—the explanation of paresis by syphilis and of a trait's evolution by its conferring a certain benefit—the force of the explanation might well be thought to lie in the aptness of the cited cause. The causal approach is now dominant in the philosophy of explanation.

The most important divide within the causal approach concerns the nature of the causal relation called on to do the explanatory work. Salmon (1984) invokes a notion of causation close to fundamental physics and declares the explanation of an event to consist of the sum total of causal influences on the explanandum in this fundamental level sense.

Such an account, however, appears to count far too many events as explanatorily relevant. As Salmon concedes, though a baseball causally influences the window that it shatters, and so rightly counts as a part of the explanation of the shattering, so do the shouts of the ball players, which cause the window to vibrate even as it is struck by the ball. The shouts, too, then, will be counted on Salmon's approach as a part of the explanation of the shattering. However, they are surely (except perhaps in some unusual cases) irrelevant.

A popular response to this worry begins with the observation that, while it is correct to say that the ball caused the window to shatter, it is not correct to say that the shouts caused the window to shatter. Such locutions suggest that there is another kind of causal relation, distinct from Salmon's fundamental physical relation, that holds between the ball and the shattering but not between the shouts and the shattering.

How can it be that Salmon's relation holds between the shouts and the shattering but the new causal relation does not? One response is that Salmon's relation is based on a faulty theory of causation, but this is not the answer normally given. Rather, the new causal relation is understood as relating events at all levels, whereas Salmon's causality relates events only at the lowest level.

The high-level event of the shattering is the event that would have occurred no matter what the physical details of the shattering, that is, no matter which shards of glass flew where. The low-level event is the event individuated by all the shattering's physical details; this event only occurred, then, because the window shattered in

exactly the way that it did. (Some writers call high-level events states of affairs or facts and hold that events proper are always low level.)

When one asks for an explanation of the shattering, one is normally asking for an explanation of the fact that the window shattered, not that it shattered in exactly the way it did. Thus, one asks for the causes of the high-level event, not the low-level event. Even though the low- and high-level events are coextensive in space and time, it seems that there are causes of the former that are not among the causes of the latter, namely, the events that determine, given that the shattering occurred, exactly how it occurred. These detail-determining events, because they are not causes of the explanandum, the shattering, do not explain it (for more on the potential for different causal relations between low- and high-level events, see Bennett 1988).

The idea, in short, is that there are many different levels of explananda, corresponding to different levels of eventhood, and different causal relations at all these different levels. Salmon's fundamental physical causation, then, is only one among many different levels of causation. Add this conception of causation as a multilevel relation to the causal approach to explanation, and one gets a theory on which the explainers of an event depend on the level of the event. (This level dependence of the explanation is also characteristic of the DN, IS, and SR accounts.)

The best-known multilevel theory of causation is the counterfactual account. If the shouting had not happened, the high-level shattering event would still have occurred, but because it would have happened in a different way, the low-level shattering event would not have occurred. Thus, the high-level shattering does not, whereas the low-level shattering does, counterfactually depend on the shouting. On a counterfactual approach to causation, this implies that the shouting is a cause of the low-level shattering but not the high-level shattering, and so, taking this multilevel relation as the explanatory causal relation, that the shouting does not explain the high-level shattering, even though—as its causation of the low-level shattering shows—it is connected causally to the shattering in Salmon's sense. For this approach to explanation, but based on a more sophisticated counterfactual account of causation, see David Lewis (1986); for a different though related multilevel approach, see James Woodward (2003).

An alternative to the multilevel approach is a two-factor approach to causal explanation, on which all explainers of an event must causally influence that event at the fundamental physical level, as prescribed by Salmon, but on which they must pass in addition a further test for explanatory relevance. Salmon (1997) himself suggests, late in his career, that the further test might be one of statistical relevance; only the causal influences that change the probability of an event explain the event. Michael Strevens (2004) suggests a different two-factor approach.

An advantage of the two-factor approach is the relatively modest demands it makes of the metaphysics of causation, transferring as it does much of the burden of determining explainers to the further test for relevance. What, then, to say about claims apparently stating the existence of high-level causal relations, such as "The ball's hitting the window, but not the players' shouting, caused the window to shatter"? Strevens (2004) suggests that locutions of this form are in fact causal-explanatory claims, asserting the explanatory relevance of certain causal influences (compare Kitcher's theory of causation mentioned earlier).

Despite the popularity of the causal approach, it is relatively undeveloped. For example, little has been written about the causal explanation of laws; it is usually said that they are explained by describing their underlying mechanisms, but not every law explicitly concerns causes and effects. Equally, not every event explanation appears to involve the delineation of causes. For examples of both kinds of worry, see Kitcher (1989, §3).

Work on the causal approach to probabilistic event explanation is more advanced. Two main currents can be distinguished in the literature. The first springs from the idea that probabilities themselves have the character of dispositions and are able to cause the events to which they are attached. The probability of one-half that a tossed coin lands heads, for example, is interpreted as a statistical disposition that causes the coin (in most cases) to land heads about one-half of the time (Fetzer 1981).

The second current flows from the idea that other events or states of affairs can cause events by making a difference to the probabilities of those events. This view is compatible with the dispositional view of probabilistic causality, but it is compatible also with its rejection. Paul Humphreys writes that "chance is literally nothing" (1989, p. 65) and so cannot cause anything itself, but that events nevertheless cause other events in an indeterministic world by making a difference to their probabilities. Because probability itself is impotent, Humphreys holds that the kind of difference a cause makes to the probability of its effect is irrelevant. It does not matter whether the change in probability is positive or negative, large or

small (compare with the SR account). Whatever the change, the factor responsible for the change is a cause and so ought to be cited in an explanation of the effect.

Peter Railton (1978) offers an account of probabilistic explanation that makes room for both conceptions of the relation between probability and causation. On what Railton calls his DNP account, an event is explained by deriving its exact probability from the appropriate initial conditions, background conditions, and laws. Formally, a DNP explanation resembles, as its name suggests, a DN explanation, except that it is the probability of the explanandum, not the explanandum itself, that is deduced. In contrast to Hempel's IS account of probabilistic event explanation, the DNP account does not require a high probability for the explanandum, and because it asks for an accurate derivation of the exact probability, it requires, like the SR account, that an explanation cite all factors probabilistically relevant to the explanandum, whether known or unknown, and (though Railton does not give a criterion for relevance) no irrelevant factors. Perhaps most important of all, the DNP account is, unlike Hempel's various accounts, open to a causal interpretation: The factors that make a difference to the probability, and even the probability itself, can be considered causes of the explanandum, and the explanation successful precisely because it specifies these causes.

An important lacuna in causal accounts of probabilistic explanation is a detailed treatment of probabilistic explanation in sciences such as statistical mechanics and evolutionary biology, where there is some possibility at least that the underlying processes producing the usual explananda are approximately deterministic. The consensus is to regard such explanations as not genuinely probabilistic; Railton (1981) suggests that they can be reinterpreted as reporting on the robustness of the underlying processes with respect to the event to be explained, that is, the processes' tendency to produce the same kind of outcome given a variety of initial and background conditions.

OTHER ISSUES

This entry will conclude with a brief sketch of some issues concerning scientific explanation not mentioned earlier. First is the question of pragmatics in explanation. Most writers hold that pragmatics affects the explanatory enterprise in only one, relatively minor, way: When an explanation is transmitted from one person to another, the act is subject to the usual pragmatics of communication. This position on pragmatics dovetails with the majority view that the explanatory facts are not essentially communicative; explanations exist independently of anyone's intention to explain anything to anyone else.

Both Bas C. van Fraassen (1980, chapter 5) and Peter Achinstein (1983) dissent from this majority, holding that there is no explanation without communication and finding in the pragmatics of communication an account of many facets of explanatory practice. However, this literature has yet to answer the question why science treats explanations as preexisting facts to be discovered, rather than as entities created in the act of communication.

Second, it is an open question whether there is a single standard for evaluating scientific explanations that has remained constant since the beginning of modern science, let alone for the entire history of human explanation. The accounts of explanation in this entry assume, of course, a positive answer, but most work on explanation lacks a substantial historical dimension.

A third issue is idealization in explanation: While almost every account of explanation surveyed earlier requires that explanations contain no false representations of reality, the practice of using idealized models in scientific explanation is widespread. These models deliberately misrepresent the nature of the systems they describe; the ideal gas model, for example, represents gas molecules as having zero volume, but despite this distortion of the facts, it is considered to explain certain behaviors of real gases. Some writers regard idealization as a temporary or practical measure, out of place in a perfected science. Strevens (2004) suggests that on both the unificationist and a certain causal approach to explanation idealizations can be seen as serving a genuine and enduring explanatory role.

See also Causation: Metaphysical Issues; Causation: Philosophical Problems in; Hempel, Carl Gustav; Laws, Scientific.

Bibliography

Achinstein, Peter. *The Nature of Explanation*. New York: Oxford University Press, 1983.

Bennett, Jonathan. *Events and Their Names*. Indianapolis, IN: Hackett, 1988.

Coffa, J. Alberto. "Hempel's Ambiguity." *Synthese* 28 (1974): 141–163.

Duhem, Pierre. *The Aim and Structure of Physical Theory*. Translated by Philip P. Wiener. Princeton, NJ: Princeton University Press, 1954.

Fetzer, James H. *Scientific Knowledge: Causation, Explanation, and Corroboration*. Dordrecht, Netherlands: D. Reidel, 1981.

Friedman, Michael. "Explanation and Scientific Understanding." *Journal of Philosophy* 71 (1974): 5–19.

Hempel, Carl G. "Aspects of Scientific Explanation." In his *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 331–496. New York: Free Press, 1965a.

Hempel, Carl G. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965b.

Hempel, Carl G. "The Logic of Functional Analysis." In his *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 297–330. New York: Free Press, 1965c. Revised version of a paper originally published in *Symposium on Sociological Theory*, edited by L. Gross (New York: Harper and Row, 1959).

Hempel, Carl G., and Paul Oppenheim. "Studies in the Logic of Explanation." *Philosophy of Science* 15 (1948): 135–175.

Humphreys, Paul. *The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences*. Princeton, NJ: Princeton University Press, 1989.

Kitcher, Philip. "Explanatory Unification." *Philosophy of Science* 48 (1981): 507–531.

Kitcher, Philip. "Explanatory Unification and the Causal Structure of the World." In *Minnesota Studies in the Philosophy of Science, vol. 13, Scientific Explanation*, edited by P. Kitcher and Wesley C. Salmon, 410–505. Minneapolis: University of Minnesota Press, 1989.

Kitcher, Philip. "1953 and All That: A Tale of Two Sciences." *Philosophical Review* 93 (1984): 335–373.

Kyburg, Henry Ely. "Comment." *Philosophy of Science* 32 (1965): 147–151.

Lewis, David. "Causal Explanation." In *Philosophical Papers*. Vol. 2, 214–240. New York: Oxford University Press, 1986.

Railton, Peter. "A Deductive-Nomological Model of Probabilistic Explanation." *Philosophy of Science* 45 (1978): 206–226.

Railton, Peter. "Probability, Explanation, and Information." *Synthese* 48 (1981): 233–256.

Reichenbach, Hans. *The Direction of Time*. Edited by Maria Reichenbach. Berkeley: University of California Press, 1956.

Salmon, Wesley C. "Causality and Explanation: A Reply to Two Critiques." *Philosophy of Science* 64 (1997): 461–477.

Salmon, Wesley C. *Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press, 1984.

Salmon, Wesley C. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press, 1990.

Salmon, Wesley C. "Statistical Explanation." In *Statistical Explanation and Statistical Relevance*, 29–87. Pittsburgh: University of Pittsburgh Press, 1970.

Scriven, Michael. "Explanation and Prediction in Evolutionary Theory." *Science* 30 (1959): 477–482.

Spirtes, Peter, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press, 2000.

Strevens, Michael. "The Causal and Unification Accounts of Explanation Unified—Causally." *Noûs* 38 (2004): 154–176.

Van Fraassen, Bas C. *The Scientific Image*. New York: Oxford University Press, 1980.

Woodward, James. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press, 2003.

Michael Strevens (2005)

EXTENSION

See *Matter; Primary and Secondary Qualities; Space*

EXTERNAL RELATIONS

See *Relations, Internal and External*

EXTRINSIC AND INTRINSIC PROPERTIES

An "intrinsic property" is one whose possession by an object at a time involves nothing other than the object (and its parts) at that time; an "extrinsic property" is one whose possession at a time involves something else. We might say, therefore, that the properties of being red and round are intrinsic to this ball, but the properties of being in Rhode Island, being less than five feet away from a tree, and having once been owned by my sister are extrinsic to it.

Peter Geach has made a corresponding distinction among changes. There is change whenever "F(x) at time t" is true and "F(x) at time t'" is false. Socrates will change when he puts on weight; he will also change when he comes to be shorter than Theaetetus merely in virtue of Theaetetus's growth. Changes of the second kind—intuitively less genuine—Geach calls "mere Cambridge changes," without proposing a rigorous criterion. We might define a mere Cambridge property as a property, change in an object's possession of which is a mere Cambridge change. Mere Cambridge properties are plausibly taken to be the same as extrinsic properties.

The matter is important, among other things, for the clear statement of a Humean view of the world. For a Humean there is in principle a description in intrinsic terms of the state of the world at any one time that is both complete and free of implications for the state of the world at any other time. "Solidity, extension, motion; these qualities are all complete in themselves, and never point out any other event which may result from them" (Hume, *Enquiry,* sec. 8, 1). It is not clear, however, that what Hume says can be true: The motion of an object is hardly free of implications about the state of the world at other times. (If an object at place p is said to be moving at time t, this is standardly in the sense that, at other times more or less near to t, the object is in other places more or less near to p.) We may have to decide between complete description and a purely intrinsic one.

Two extreme views are that all properties are really intrinsic and that all properties are really extrinsic. Gottfried Wilhelm Leibniz holds the first: "There are no purely extrinsic denominations." His insistence resulted in the drastic denial of the reality of relations and, most notably, of space and time; it has not been widely accepted. A moderate version of the opposite view, that all properties are really extrinsic, might be held by someone, like Karl Popper, who believes that physical properties are essentially dispositional. Both extremes, in different ways, represent a sense that the nature of one thing cannot be divorced from the nature of others. Confidence in a firm distinction between the intrinsic and the extrinsic, on the other hand, is more characteristic of an optimistic Humean.

It is not easy to give a precise characterization of intrinsic properties, and there may not even be a unique idea, so to speak, waiting to be characterized. We might try saying that extrinsic properties are relational properties and intrinsic properties nonrelational. But many intuitively intrinsic properties still in some way involve a relation—squareness involves a relation among the sides of an object. Can we say that intrinsic properties are those that do not involve a relation to anything that is not a part of the object? This is perhaps the clearest criterion, but it may still be incapable of capturing all our intuitions at once. The power to open locks of kind *k,* for example, apparently involves a relation to external things of a certain kind—which would seem to make it extrinsic. Yet it is a property that a key can have if it is, so to speak, alone in the world—which would seem to make it intrinsic.

It may be helpful to invoke a distinction between relational descriptions of a property and descriptions of a relational property. But that distinction is itself perplexing. Is "possessing what is actually Jane's favorite intrinsic property" a relational description of a first-order property or a description of second-level relational property?

Philosophers have argued in many cases that apparently intrinsic properties are in fact extrinsic. Terms such as *old, great,* and *imperfect,* John Locke says, "are not looked on to be either relative or so much as external denominations," but they conceal a tacit relation (*Essay*). More worrying are challenges even to the idea that primary qualities, like size and shape, are intrinsic. The size of the ball is, we may think, intrinsic to it. We can describe a scenario where everything else in the universe is twice its actual size while the ball remains the same. But can we properly distinguish this from a scenario where the rest of the world is the same but the ball is half its actual size? Some will argue that length is relational, and the two scenarios make a distinction without a difference: size, after all, is extrinsic. Others will argue instead that even if our descriptions of size are relative, for example, to standard measures, what is described is still an absolute and intrinsic property.

Are any or all of a person's mental properties intrinsic to her? The question is in part about the limitations of methodological solipsism. If Jane could not possess the property of thinking of Bertrand Russell if Russell did not exist, then that property must be extrinsic to her. Some will try to segment referential thought into an internal and an external component; but if that proposal fails, referential thought will typically be extrinsic to the thinker. (Another option is that the thinker, or her mind, extends more widely than her body—and actually includes Russell.) One might argue a similar point with respect to thought about properties as well as about individuals. (A brain that has never been out of a vat does not know what a meter is.) Maybe there are very few mental properties intrinsic to a person; or maybe we should think again about what the notion of the intrinsic is, and what exactly it is supposed to do for us.

See also Hume, David; Internalism versus Externalism; Leibniz, Gottfried Wilhelm; Locke, John; Metaphysics; Popper, Karl Raimund.

Bibliography

Geach, P. T. *God and the Soul.* London: Routledge and Kegan Paul, 1969.

Leibniz, G. W. "Primary Truths" and "Letters to Des Bosses." In *Philosophical Essays,* translated by R. Ariew and D. Garber. Indianapolis: Hackett, 1989. See esp. pp. 32, 203.

Lewis, D. K. "Extrinsic Properties." *Philosophical Studies* 44 (1983): 197–200.

Lewis, D. K. *On the Plurality of Worlds.* Oxford: Blackwell, 1986. Chaps. 1.5, 4.2.

Locke, J. *Essay concerning Human Understanding.* Bk. 2, Chaps. 25, 28.

Popper, K. *The Logic of Scientific Discovery,* 424–425. London: Hutchinson, 1959.

Justin Broackes (1996)

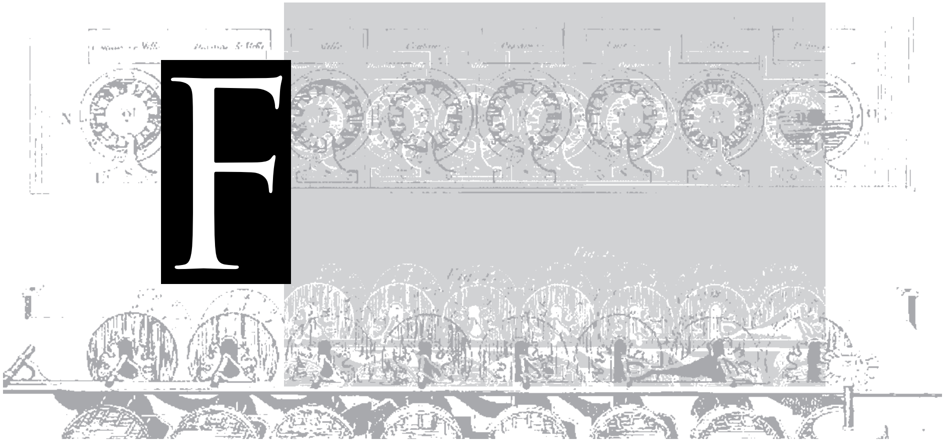

F

FACTS

See *Analysis, Philosophical; Correspondence Theory of Truth; Propositions*

FAITH

In discourse concerning religion, "faith" has two rather different meanings. As a trusting and confident attitude toward God, faith (*fiducia*) may be compared with trust in one's fellow human beings. As a cognitive act or state whereby men are said to know God or to have knowledge about him, faith (*fides*) may be compared with our perceptual awareness of our material environment or our knowledge of the existence of other persons. This article will deal with the notion of faith as putatively cognitive, as this has operated in Western religious thought.

FAITH IN CLASSIC CATHOLIC AND PROTESTANT THOUGHT

THOMAS AQUINAS. The key thinker for the discussion of faith in Roman Catholicism is Thomas Aquinas, who wrote on the nature of faith in his *Summa Theologiae.* Thomas's main points may be summarized as follows:

(1) Faith is belief in revealed truths. Ultimately the object of faith is God himself, who is not, however, known by the human mind in his divine simplicity but only discursively and by means of propositions. These revealed truths are authoritatively presented in the creeds. Thus, to have faith means to believe the articles of faith summarized in the credal affirmations of the church.

(2) In its degree of certainty, faith stands between knowledge (*scientia*) and opinion. It ranks below knowledge, for although the objective cause of faith—divine truth—is in itself more certain than the product of any human reasoning, yet faith's grasp of its object—since it lacks cogent demonstration—is less certain than rational knowledge. On the other hand, faith ranks above opinion, for while opinion is accompanied by doubt and by fear that the opposite opinion may be true, faith is firm and free from all such hesitations.

(3) The objects of faith on the one hand, and of sight and demonstration on the other, are different: "the object of knowledge [*scientia*] is something seen, whereas the object of faith [*fides*] is the unseen." There can thus be no faith concerning matters that are objects of rational knowledge, for knowledge excludes faith.

However, some truths may be objects of faith to one person and of knowledge to another. In particular, some of the preliminary articles of faith—such as the existence, unity, and incorporeality of God—are capable of being philosophically demonstrated and are revealed as objects of faith only for the sake of those many who are unable to follow the path of abstract reasoning. Those matters that are of faith absolutely are above reason—incapable of being arrived at by human reasoning, however expert.

Thomas's account of the relation between faith and reason is, accordingly, that they apprehend different sets of truths, the truths of faith being above reason. However, this statement must be qualified by adding that there is an area in which faith and reason overlap, since the basic theological propositions—those of natural theology—are held to be both demonstrable and revealed.

(4) Faith is "an act of the intellect assenting to divine truth at the command of the will moved by the grace of God." That is to say, whereas in knowledge the intellect is moved to assent by the object itself, known either directly or by demonstrative reasoning, in faith the intellect is moved to assent "through an act of choice, whereby it turns voluntarily to one side rather than to the other." Faith does not, however, represent an arbitrary or unmotivated decision. It is a response, under the influence of divine grace, to certain external evidences, particularly miracles. As such, it is sufficiently determined by the evidence to be rational and yet sufficiently undetermined and free to be meritorious.

MODERN CATHOLICISM. The doctrine of faith in modern Catholicism is essentially Thomist, although a fuller apologetic context is developed than was necessary in the medieval period. Faith is defined by the first Vatican Council (1870) as "a supernatural virtue, by which, guided and aided by divine grace, we hold as true what God has revealed, not because we have perceived its intrinsic truth by our reason but because of the authority of God who can neither deceive nor be deceived" (*Constitution on Faith,* Ch. 3). Such a definition provokes a query, for faith, characterized as belief in various truths on divine authority, presupposes a knowledge both that God exists and that he has revealed the propositions in question. How is this prior information gained? The question is answered by the doctrine of the *preambula fidei.* The preambles to faith consist in the acceptance of God's existence, established by philosophical proofs, and of the validity of the biblical revelation and the authority of the Catholic church as the divinely appointed guardian of revelation. These latter are authenticated by a variety of visible signs, such as miracles, fulfillments of prophecy, holy lives, and the growth and durability of the church. The believer's appreciation of the weight of this evidence is not an exercise of faith but of reason: "The use of reason precedes faith and must lead us to it" (Denzinger, *Enchiridion* No. 1626, cf. No. 1651). Thus, the whole structure of belief rests originally upon the historical evidences of miracles and other manifestations of divine activity that do not establish the articles of faith themselves, but rather the fact that the omniscient God has revealed these articles. Although it is denied by Catholic apologists, the comment of John Locke in his *Essay concerning Human Understanding* would still seem pertinent: "Though faith be founded on the testimony of God (who cannot lie) revealing any proposition to us, yet we cannot have an assurance of the truth of its being a divine revelation greater than our own [rationally acquired] knowledge; since the whole strength of the certainty depends upon our knowledge that God revealed it."

It should be noted that in some of the more recent Catholic discussions, such as that by Eugène Joly in the article "Faith" in the *Twentieth Century Encyclopedia of Catholicism* (Paris, 1956), there is a tendency to move beyond a narrowly propositional conception of faith and to be hospitable to the idea of an encounter with God mediated through man's religious experience.

PROTESTANTISM. For Martin Luther (1483–1546), the chief moving spirit of the Reformation, faith was not primarily belief in the church's dogmas but rather a wholehearted trust in the divine grace and love revealed in Jesus Christ. Thus, Luther considered faith as primarily *fiducia* rather than *fides.* Indirectly it included all the fundamental Christian beliefs, but Luther's main emphasis was upon faith as a total reliance upon God's omnipotent goodness. He was not concerned with the logically prior question of our knowledge that God exists. In this he was at one with the biblical writers, who were so vividly conscious of the reality and presence of God that their writings take his existence for granted. In the Bible, as in the thought of Luther, faith is not the belief that God exists, that he is three in one, and so on, but is an attitude of trust and self-commitment to him. In a distinction that Luther himself drew, it is not belief *that* but belief *in.*

John Calvin (1509–1564), the first and greatest systematizer of Reformed theology, gave greater prominence to the cognitive aspect of Christian faith, defining it in the *Institutes* as "a firm and certain knowledge of God's benevolence toward us, founded upon the truth of the freely given promise in Christ, both revealed to our minds

and sealed upon our hearts through the Holy Spirit." That to which faith responds is the Bible as the inspired Word of God: "there is a permanent relationship," Calvin says, "between faith and the Word." Thus, in Reformed theology acceptance of the authority of Scripture replaces the *preambula fidei* of Thomism.

The philosophical question raised by this conception of faith is similar to that raised by the Roman Catholic conception: what is the nature of our knowledge that the God whom we are invited to trust in fact exists, and that he has inspired the writings which he is alleged to have inspired?

Two subsequent Protestant contributions to some extent address themselves to this question. In the early nineteenth century Jakob Friedrich Fries, influenced by Friedrich Heinrich Jacobi in the previous century, described faith as *Ahnung* (or *Ahndung*), by which he meant an unconceptualized feeling, hunch, or presentiment as to the reality of the supernatural. Friedrich Schleiermacher also regarded faith as a kind of feeling (*Gefühl*), a sense of absolute dependence upon a higher reality. In a different vein altogether Søren Kierkegaard, the father of modern existentialism, emphasized the objective uncertainty of the religious realm, which can be entered only by a leap of faith. He stressed the tremendous risk involved, like being "out upon the deep, over seventy thousand fathoms of water."

MODERN THEORIES OF FAITH

The Thomist doctrine contains most of the elements that have, in varying proportions, characterized subsequent theories of faith. The Thomist analysis treats faith as (*a*) a form of propositional belief but as (*b*) belief that rests upon weaker evidence than scientific knowledge, and (*c*) regards it as requiring an act of will.

VOLUNTARIST THEORIES. Nearly all subsequent epistemological discussions of faith assume that it is a cognitive attitude directed toward religious propositions. Widespread in modern discussions is the rationalist definition of faith as (to quote a typical formulation) "very firm belief, either unsupported or insufficiently supported by evidence" (C. J. Ducasse, *A Philosophical Scrutiny of Religion,* New York, 1953, p. 74). Some such definition as this is used by a large number of religious philosophers as well as by many of those who reject religious belief.

How, from the believer's point of view, is the evidential gap supposed to be filled? Here the voluntarist theme, first stressed by Thomas, reappears.

Pascal. In the famous Wager passage in his *Pensées* (No. 233) Blaise Pascal (1623–1662) recommends a purely voluntarist route to religious belief, assuming that reason can find no grounds on which to determine whether there is a God. One must decide to believe or to disbelieve; and regarding the decision as a wager, it is prudent to decide to believe. One will then gain eternal life and felicity if God indeed exists and will lose nothing if he does not; whereas if one decides to disbelieve, one will gain nothing if he does not exist but will forfeit eternal life if he does.

William James. The idea briefly adumbrated by Pascal appears in an elaborated form in William James's well-known essay "The Will to Believe" (1895). He points out that there are cases in which we may come into contact with some aspect of reality only by acting, prior to any adequate evidence, as if it existed; in these instances our faith helps to bring its object into being. For example, in the realm of personal relationships faith in an individual's good will or honesty may on occasions elicit these qualities when otherwise they would have been wanting. Precursive faith of this kind is justified by its subsequent verification rather than by prior evidence.

James then proceeds to consider religious faith. Here we have what is for many people a living, momentous, and—James emphasizes—a forced option, for to refuse to say "Yes" to the claim of religion is in effect to say "No" to it. It is to miss the good that follows from believing the religious gospel, if it be true, as decisively as if one had positively rejected it. Therefore we have the right to choose for ourselves between the risk of falling into error by adopting a faith that may turn out to be false, and the risk of missing our highest good by failing to adopt a faith that may turn out to be true.

Furthermore, James adds, the Judeo-Christian religious hypothesis refers specifically to a personal God; and in our dealings with a cosmic Thou, as with our fellow humans, a venture of faith on our part may be necessary if we are to establish any positive relationship. To respond as a person to another person involves showing a certain trustfulness and willingness to "give the benefit of the doubt" and thereby anticipate proof and verification. It may be that God can or will disclose himself only to one who shows such an initial faith and is willing to venture in trust beyond what has been established by scientific proof or philosophical demonstration. In other words, it is possible that in order to gain the religious knowledge upon which our personal good depends, we must give rein to our "passional" desire to believe. Hence, James concludes, we cannot reasonably be required to adopt a

methodology that would prohibit us from finding this good: For "a rule of thinking which would absolutely prevent me from acknowledging certain kinds of truth if those kinds of truth were really there, would be an irrational rule."

James's argument has been criticized at a number of points, chief among them being the following:

(1) His basic assumption is that there are no grounds of either reason or evidence which might lead one to accept or reject the "religious hypothesis." There is nothing to make it significantly more probable either that there is or that there is not a God; and in such a situation, says James, we are entitled to follow our desires. But many, both theists and atheists, claim that there are substantial arguments or evidences for (or against) the existence of God, and that we ought to attend to these rather than to our personal predilections. Furthermore, whatever conclusion we arrive at should be held only with the degree of conviction that is warranted by the evidence.

(2) The "precursive faith" that helps to create that in which it believes, although a genuine phenomenon, is irrelevant to belief in the existence of God or in the reality of eternal values; for if these exist, they exist independently of man's belief or lack of belief in them. In the social situations James cites, our willingness to trust someone in advance of proof of his trustworthiness may help to make him trustworthy but does not bring him into existence, and faith in the existence of a divine creator of the universe cannot bring such a being into existence.

(3) James's argument ought not to be applied only to our current live options, since "live option" is a psychological category having no necessary relation to the truth or falsity of hypotheses. We ought to heed equally every momentous and forced option. However, we cannot act upon them all, since they demand incompatible responses. We shall act, then, only upon that which we should most *like* to be true. So stated, the "right to believe" argument stands revealed as an invitation to wishful thinking.

(4) From the side of religion, an unfavorable comparison is made between the kind of faith recommended by James and that already possessed by the religious believer. James presumes a complete absence of grounds for belief and, in this situation, he proposes a prudent gamble. However, the religious believer—as we meet him, for example, in the pages of the Bible—is convinced that he is aware of God acting toward him in and through the events of the world around him, so that at all times he is having to do with God and God with him. His concern is to draw others into this direct awareness of God, rather than to induce them to make James's gamble.

Tennant. F. R. Tennant (1866–1958) has provided the fullest recent voluntarist apologetic for theistic faith. Faith in general, according to Tennant, is the conative element in the acquisition of knowledge. In every advance from sense data to the perception of an ordered world or from the projection of a scientific hypothesis to its observational verification, as in every successful voyage of discovery or in the invention of some new kind of machinery, there must be not only an act of theorizing or of insight but also a sustained effort of will that carries the operation through to completion. In both of these respects religious cognition shares a common structure with knowledge in the sciences and in personal life. First, there is the creation of a hypothesis: Scientific hypotheses satisfy the inclination to explain the structure and order of the universe by quantitative laws, while theological and ethical hypotheses satisfy the inclination toward teleological explanation. Second, there is the volitional investment, the venture of faith, which may eventually be rewarded with a dividend of verified knowledge. The faith venture in secular contexts is continuous in kind with that of the religious prophets and apostles. Thus, faith is the indispensable volitional component within the process of acquiring knowledge, and it plays a basically similar role in both religion and nonreligious life.

However, the kinds of verification that are possible in science and religion are importantly different, although Tennant wavers between stressing their similarity and their dissimilarity. Scientific verification consists in observing that predictions deduced from a hypothesis are fulfilled in the experimenter's observations. Religious verification, on the other hand, consists in the valuable effects of faith in the life of the believer—in strengthening him as a moral agent and in his attainment of heroic life. Thus, while scientific verification leads to objective certainty, or at least to a high degree of objective probability, religious verification leads only to subjective certitude. "Nevertheless," Tennant adds, "verification such as religion claimed for its faith will satisfy most men."

It is noteworthy that the basic features of the classic Thomist analysis of faith reappear, although in a very different setting, in Tennant's theory: (1) Faith, as acceptance of the religious hypothesis, is propositional. (2) Faith is of the same cognitive order as scientific knowledge but is based upon a lower degree of evidence. (3) Faith is not concerned with the material world itself, which is an object of knowledge, but with its teleological

meaning. (4) Faith is distinguished by the conative element within it from ordinary belief and knowledge. Whereas the act of will can, in Thomism, appeal for rational justification to such external evidences as miracles and fulfilled prophecies, in Tennant's philosophy it appeals to a comprehensive teleological argument for the existence of God.

This propositional and voluntarist tradition, which has so largely dominated the scene since the time of Thomas, has been criticized on the following grounds: (*a*) Actual religious faith is not, from the believer's point of view, analogous to a scientific hypothesis but with a weaker verification. It is a direct awareness of God, with its own assurance that is not dependent upon philosophical argument. (*b*) As (putatively) a direct awareness of God, faith is not primarily a form of propositional belief; rather, it is a form of religious experience. Theological beliefs naturally grow out of it but are not themselves the primary objects or content of faith.

FAITH AND FREEDOM. A very important connection has long been recognized between faith and what may be called the cognitive freedom of the human mind in its relation to God. The first writer to note this connection was the second-century Christian writer Irenaeus, who said, "And not merely in works, but also in faith, has God preserved the will of man free and under his own control" (*Adversus Haereses,* IV, 37, 5). The theme is continued in Augustine and in Thomas's view that faith is a sufficiently free act to be meritorious. Pascal stated that God's self-revelation in the Incarnation took a deliberately veiled form, so that no one could be compelled to find God in Jesus Christ, and yet so that all who were willing to find God there might do so: "… willing to appear openly to those who seek him with all their hearts, and to be hidden from those who flee from him with all their hearts, he so regulates the knowledge of himself that he has given signs of himself, visible to those who seek him, and not to those who seek him not" (*Pensées,* No. 430). Søren Kierkegaard also spoke of the divine incognito in the Incarnation. The same theme is continued by the twentieth-century Protestant theologian Emil Brunner and by many other writers.

The basic thought behind this emphasis, at any rate in the modern writers, is that God, having created man as personal, always acts toward him in ways that respect and preserve man's freedom and responsibility. For this reason God does not reveal himself to man in his unveiled glory, for in a direct, unmediated awareness of infinite perfection man's frail moral autonomy would be destroyed. Therefore, the divine presence is always mediated through the events and circumstances of a world that God has created to be a relatively independent sphere of interaction with his human creatures. Man's personal autonomy is protected by the fact that he can become conscious of God's activity toward him only by an uncompelled response of faith. Thus, men are not only free to obey or disobey God; they also have the prior and more fundamental freedom to be conscious of God or to refrain from being conscious of him. The human mind displays a natural tendency to interpret its experience religiously, but this tendency acts only as an inclination that can be resisted and inhibited. Man is thus cognitively free in relation to God. Faith is the correlate of freedom and is related to cognition as free will is to conation.

FAITH AS INTERPRETATION. Closely related to this emphasis upon man's cognitive freedom is a contemporary theory that regards faith as the interpretative element in religious experience—that which constitutes it as *religious* experience in distinction from any nonreligious experiencing of the same field of data. Here "interpretation" does not mean intellectual interpretation or theory construction, but something more akin to the interpretative processes which take place in sense perception. From the point of view of epistemology, faith is thus analogous to the phenomenon of "seeing as," which was brought to the attention of philosophers by Ludwig Wittgenstein in his *Philosophical Investigations* (II, xi). We may look at a puzzle picture, seeing it now as a meaningless disarray of lines and now as the outline of, say, a human face. This is an instance of purely visual interpretation. But the concept of "seeing as" can be expanded into that of "experiencing as," referring to the way in which a situation apprehended through our sensory apparatus as a whole is experienced as having some particular kind of significance; that is, as rendering appropriate some particular dispositional response on our part. To cite religious examples, when the Old Testament prophets experienced the events of contemporary Israelite history as mediating the presence and activity of God and as speaking a divine imperative to them, they were undergoing a religious mode of "experiencing as." Again, the apostles whose witness constitutes the message of the New Testament saw, but were not compelled to see, Jesus as the Christ. Indeed, it is always true of the religious mode of "experiencing as" that the data in question are in themselves ambiguous and capable of being responded to either religiously or naturalistically. More strictly, the two types of interpretation are not alternatives on the same level but are different orders of significance found in the same field of data.

The religious significance of events includes and transcends their natural significance. Those events the prophets saw as acts of God can also be seen as having proximate natural or human causes; and the person of Christ, seen by Christian faith as divine, is depicted in the New Testament as being at the same time genuinely human. From a theological point of view, this systematic ambiguity, which is the precondition of faith, serves to protect man's freedom and autonomy as a finite personal being in relation to the infinite God.

See also Atheism; Augustine, St.; Bad Faith; Belief; Brunner, Emil; Calvin, John; Existentialism; Fries, Jakob Friedrich; Jacobi, Friedrich Heinrich; James, William; Kierkegaard, Søren Aabye; Luther, Martin; Miracles; Pascal, Blaise; Teleological Argument for the Existence of God; Tennant, Frederick Robert; Thomas Aquinas, St.; Traditionalism; Truth; Wittgenstein, Ludwig Josef Johann.

Bibliography

The article on *πίστις* by Rudolf Bultmann and A. Weiser in Vol. VI of Kittel's *Theologisches Wörterbuch zum Neuen Testament* (Stuttgart, 1959) treats authoritatively the various biblical concepts of faith: It was translated by Dorothea M. Barton as *Faith* (London, 1961). Historical treatments of the idea of faith occur in D. M. Baillie, *Faith in God* (Edinburgh, 1927) and W. R. Inge, *Faith* (London, 1909). The distinction between faith as trust and as cognition is developed in Martin Buber, *Zwei Glaubensweisen* (Zürich, 1950), translated by Norman P. Goldhawk as *Two Types of Faith* (London, 1951). The cognitive aspect of Christian faith is defined in John Calvin, *Institutio Christianae Religionis* (Basel, 1536; 5th ed., 1559), translated by F. L. Battles as *Institutes of the Christian Religion,* 2 vols. (London, 1861). Thomas's teaching on the nature of faith occurs in *Summa Theologica* (II–II, 1–7), translated by Anton C. Pegis in *The Basic Writings of Saint Thomas Aquinas,* Vol. II (New York, 1945). Thomas's teaching on the relation between faith and reason is contained in *Summa Contra Gentiles* (I, 3–8), which was translated by Anton C. Pegis as *On the Truth of the Catholic Faith* (Garden City, NY, 1955–1957).

Contemporary Roman Catholic discussions of faith include G. D. Smith, "Faith and Revealed Truth," in *The Teaching of the Catholic Church* (New York, 1956), Vol. I and Eugène Joly, *Qu'est-ce que croire?* (Paris, 1956); the latter was translated by Illtyd Trethowan as *What Is Faith?* (New York, 1956). Also see H. J. D. Denzinger, *Enchiridion Symbolorum* (Freiburg, 1952).

Protestant neoorthodox conceptions of faith are represented by Karl Barth, *Kirchliche Dogmatik* (Zürich, 1932), Vol. I, Part 1, translated by G. T. Thompson as *The Doctrine of the Word of God* (Edinburgh: Clark, 1936); H. F. Lovell Cocks, *By Faith Alone* (London, 1943); and F. Gogarten, *Die Wirklichkeit des Glaubens* (Stuttgart, 1957), translated by Carl Michalson and others as *The Reality of Faith* (Philadelphia, 1959).

The conception of faith as *Ahnung* occurs in F. H. Jacobi, *David Hume über den Glauben, oder Idealismus und Realismus* (1787), in *Werke* (Leipzig, 1815), Vol. II; and J. F. Fries, *Wissen, Glaube und Ahnung* (Jena, 1805), edited by Leonard Nelson (Göttingen, 1905). See also R. Otto, *Kantisch-Fries'sche Religionsphilosophie und ihre Anwendung auf die Theologie* (Tübingen: Mohr, 1909), translated by E. B. Dicker as *The Philosophy of Religion* (London: Williams and Norgate, 1931).

Christian existentialist views of faith occur in Søren Kierkegaard, especially in *Philosophical Fragments,* translated by David F. Swenson (Princeton, NJ, 1936) and *Concluding Unscientific Postscript,* also translated by David F. Swenson (Princeton, NJ, 1941); and in G. Ebeling, *Das Wesen des Christlichen Glaubens* (Tübingen, 1959), translated by Ronald Gregor Smith as *The Nature of Faith* (London, 1961).

The classic attempt to base faith on a moral foundation is in Kant's *Critique of Practical Reason.* W. R. Sorley, *Moral Values and the Idea of God* (Cambridge, U.K., 1918); D. M. Baillie, *Faith in God* (Edinburgh 1927); and J. Baillie, *The Interpretation of Religion* (Edinburgh, 1929) contain more recent endeavors to the same end.

Modern voluntarist theories of faith are found in William James, *The Will to Believe* (New York: Longman, 1897); James Ward, *Essays in Philosophy* (London, 1927); and F. R. Tennant, *Philosophical Theology,* Vol. I (Cambridge, U.K., 1928) and *The Nature of Belief* (London, 1943).

The view that faith operates not only in religion but also in many other spheres of life has an extensive literature, including Arthur Balfour, *The Foundations of Belief* (London, 1885); W. R. Inge, *Faith* (London: Duckworth, 1909); B. H. Streeter, ed., *Adventure: The Faith of Science and the Science of Faith* (London, 1927); Alan Richardson, *Christian Apologetics* (London, 1947); Raphael Demos in the symposium *Academic Freedom, Logic and Religion* (New York, 1953); and H. R. Niebuhr, *Radical Monotheism and Western Culture* (New York: Harper, 1960).

J. H. Newman's Illative Sense theory (*A Grammar of Assent,* London, 1870) is discussed in M. C. D'Arcy, *The Nature of Belief* (London, 1945). Paul Tillich's view of faith as "ultimate concern" occurs in *Dynamics of Faith* (New York: Harper, 1957).

The conception of faith as the interpretative element in religious experience is expounded in J. H. Hick, *Faith and Knowledge* (Ithaca, NY, 1957).

John Hick (1967)

FAITH [ADDENDUM]

This entry focuses on the various ways in which recent philosophers working within or in reaction to the scriptural traditions have construed the relationship between faith in God and the belief in God's existence. The entry will also touch on different views regarding the question of whether faith is, can be, or should be, *rational*. How-

ever, a full treatment of the issue of faith and reason is beyond the scope of this entry.

There are several different camps of views regarding the relationship of *faith in God* and *belief in God's existence*. One camp holds that faith is not belief, because it is, so to speak, higher than belief; faith is *knowledge* of God. Thus, Dewey Hoitenga (1991) holds that faith is a knowledge of God that comes through direct acquaintance. According to John Hick (1957), faith is the interpretive element within religious experience that results in an awareness or knowledge of God. On this view, a person who merely believes that God exists does not really have faith at all. And, on this view, faith is rational, in the sense that it is based on a religious experience of a certain sort.

A second camp follows Thomas Aquinas in treating faith in God as basically equivalent to belief in God's existence. In his book *What Is Faith?* (1992), Anthony Kenny recognizes that there are different senses of the term *faith*, but he suggests that in at least one of its senses, faith is equivalent to belief in God's existence. Hence, in discussing the issue of whether faith is rational, Kenny focuses primarily on the question of whether it is rational to believe in God's existence.

A third camp holds that faith in God has nothing to do with belief that God exists. Such philosophers sharply distinguish faith *in* God, from a belief *that* God exists. On this view, faith is a commitment of some sort, such as ultimate concern (Paul Tillich 1958) or hope in some divine end (Louis Pojman 1986). To a large degree, the motivation for such a conception is the claim that what is most crucial about genuine religious devotion is somehow fundamentally noncognitive, or perhaps, "nonpropositional," that is, having nothing to do with the proposition "God exists." On this view, faith is an affective or emotional matter, or a matter of the will; it is not a cognitive or intellectual affair. For Tillich, in some sense the question of the rationality of faith cannot properly arise, because faith is an ultimate commitment that cannot be adjudicated by anything prior or external to itself. And, for Pojman, it can be rational to have faith even if it is *not* rational to believe in God's existence.

In response to this camp, many writers have argued that belief or faith *in* something generally presupposes some kind of *belief that* some proposition is true. In order to commit oneself to God, or to trust God, or to have hope in some divine end, etc., one must have some kind of belief that God exists. Nicholas Wolterstorff (1983) points out that even if it is granted that what is most important about religious devotion is noncognitive, or even, nonpropositional, it does not follow that religious devotion is altogether noncognitive. Hence, the fourth camp, which holds that although faith is not identical with belief, faith involves or requires at least some kind of belief in God's existence. On this view, the issue of the rationality of faith cannot be completely divorced from the question of whether it is rational to believe in God's existence.

Philosophers within this camp differ on what precisely is involved in faith aside from belief in God's existence, and also on the nature and degree of that belief. Richard Swinburne holds that faith involves doing certain actions to achieve good purposes, whilst relying on the belief that God exists to provide what one wants or needs. For Swinburne, because the evidence that supports the belief in God's existence is strong but inconclusive, it is the role of faith to fill the gap between what the evidence warrants and what the religious believer believes. Paul Helm (2000) agrees that faith involves trust, but he also ascribes to a Lockean view that belief, as well as trust, should be proportionate to the evidence. Thus according to Helm, the more evidence one has that God exists, the greater degree of trust is appropriate.

Still others suggest that faith consists in a kind of practical assumption that God exists. Robert Audi (1991) proposes that to have faith is to assume that God exists as a practical rule for living. According to Audi, one can have this kind of faith so long as one does not actually believe that God does *not* exist. Joshua Golding (1990) proposes that faith be viewed as the pragmatic assumption that God exists, *for purpose of living a religious life*. For Golding, in order to have pragmatic faith, one must at least have the belief that there is at least some live probability that God exists. On either of these latter views, so long as God's existence has not been conclusively refuted or disproved, there might be pragmatic considerations that make it rational or justifiable to have faith. Such pragmatic considerations might include the potential benefits, spiritual, moral, or otherwise, that might accrue to the faithful, especially if it turns out that God does indeed exist.

OTHER VIEWS

Worthy of mention is another view that does not fit neatly into any of the above camps. William Lad Sessions (1994) claims that there is no single substantive concept of faith that applies univocally to everything reasonably labeled "faith" in all religious traditions (or even sometimes within certain traditions). Faith is an "analogical" concept whose various instances resemble one another in

various ways, without there being a single feature or set of features common to them all. According to Sessions, it is impossible to pin down what is "faith." However, this is almost a foregone conclusion if one takes under consideration the entire range of world religions. Sessions also concedes that adherents of different religions as well as adherents of the same religion will continue to argue over which sense of faith is the most true to a given tradition or set of texts, as well as over the question of whether faith is rational and if so in what sense. Thus despite Sessions's useful taxonomy, it appears that debates among the above camps are not likely to cease in the near future.

Finally, consider views of faith in Judaism and Islam. The Hebrew word for faith is *emunah*. Interpreting the notion of *emunah* (faith), some Jewish philosophers have emphasized intellectual assent to propositions, whereas other thinkers have emphasized volitional or affective commitment. Perhaps the most well-known discussion of this issue in the twentieth century is by Martin Buber. In his book *Two Types of Faith* (1951), Buber claimed that the Christian view of faith is "an acknowledgement of God's existence," and that the Jewish view is "trust in God." A superficial reading might lead one to think that, for Buber, the Jewish view of faith belongs in the third camp above, namely, that faith has nothing to do with a belief that God exists. However, such a reading is mistaken. While Buber claims that for Judaism the fiduciary aspect of faith is primary, he does not claim that beliefs are irrelevant to faith, or that somehow, one could have trust in God without believing that God exists. Hence, Buber more properly belongs in the fourth camp described above.

Another notable claim made by Buber is that, for Christianity, faith is primarily an *individual* matter; for Judaism, faith is primarily a *communal* matter. Buber argues that for Christianity, a relationship with God is initiated when and if a person decides to adopt Christian faith, that is, to acknowledge God. Of course, there can be a group of Christians who form a Christian community, but the primary act of faith is carried out by individuals. However, in the case of Judaism, God has already established a covenantal relationship with the people of Israel, in which he has made certain commitments to them as a people. Thus the primary act of faith, that is, trust in God, is the Jewish response to God's commitment to the Jewish people. Perhaps Buber's point is borne out by the fact that for centuries the primary liturgical expression of Christian faith has been in the singular, that is, "*credo in unum deo*," which translates as "I believe in one God." In the Jewish liturgy, the primary liturgical expression of faith is in the plural, "*Hear O Israel, the Lord our God, the Lord is One*."

Islam treats faith or *iman* as a religious duty or virtue. According to a well-known *hadith* (Islamic legend), Mohammed made a sharp distinction between *islam* (submission) and *iman* (faith). To accept *islam* is to testify that there is no God but Allah, and that Mohammed is his prophet; to pray, give charity, keep the fasts, and make the pilgrimage. To have *iman* is to believe in God, his angels, his messengers, and the last day of judgment. The Koran itself (49: 14) makes clear that one can be a *muslim* without having *iman*. Merely testifying to God's existence and Mohammed's prophecy makes one a *muslim*, but not one who has *iman*. The implication is that *iman* goes well beyond mere belief in God's existence, and that it involves an active trust in God. Islamic theologians have debated the issue of what kind of belief qualifies a person as a *muslim*. If one believes, but sins, or believes insincerely, does one truly belong in the Muslim community? Over the centuries, different schools of thought adopted different positions on these questions. However, it is difficult to find a discussion among recent Islamic writers directly on the question of what is the precise relationship between *iman* and belief in God's existence.

Bibliography

For a detailed bibliography on discussions of faith, consult William Lad Sessions, *The Concept of Faith*. See below.

Audi, Robert. "Faith, Belief, and Rationality." In *Philosophical Perspectives*. Vol. 5: *Philosophy of Religion*, edited by James E. Tomberlin. Atascerdo, CA: Ridgeview, 1991.

Buber, Martin. *Two Types of Faith*. New York: Macmillan, 1951.

Golding, Joshua L. "Toward a Pragmatic Conception of Religious Faith." Faith and Philosophy 7 (1990): 486–503.

Helm, Paul. *Faith with Reason*. Oxford: Clarendon Press, 2000.

Hick, John. *Faith and Knowledge: A Modern Introduction to the Problem of Religious Knowledge*. New York: Cornell University Press, 1957.

Hoitenga, Dewey. *Faith and Reason from Plato to Plantinga: An Introduction to Reformed Epistemology*. Albany: SUNY Press, 1991.

Kenny, Anthony. *What Is Faith?* Oxford: Oxford University Press, 1992.

Pojman, Louis. "Faith Without Belief." Faith and Philosophy 3 (1986): 157–176.

Sessions, William Lad. *The Concept of Faith: A Philosophical Investigation*. Ithaca, NY: Cornell University Press, 1994.

Swinburne, Richard. *Faith and Reason*. Oxford: Clarendon Press, 1981.

Tillich, Paul. *Dynamics of Faith*. New York: Harper and Row, 1958.

Waines, David. *An Introduction to Islam*. Cambridge, U.K.: Cambridge University Press, 1995. See esp. chap. 4.

Wolterstorff, Nicholas. "Introduction." In *Faith and Rationality: Reason and Belief in God*. Notre Dame: Notre Dame University Press, 1983.

Joshua L. Golding (2005)

FALLACIES

A fallacy, in the strict sense, is an invalid form of argument. Thus fallacy, or unsoundness in reasoning, is distinguished from simple falsity in that a single statement or belief may be false, but what is fallacious is the transition from a set of premises to a conclusion. However, this distinction is often slurred over; and we call other kinds of mistakes or confusion that are more or less closely related to faults in reasoning fallacies, in an extended sense. Indeed, we sometimes give the title of "fallacy" to what is little more than a particular type of false belief. At the same time, we usually count as fallacies only those invalid forms of argument, or related kinds of error, that are plausible and into which people frequently and easily fall. Fallacy is different from sophistry, which is the deliberate use of unsound reasoning or of related errors. A fallacy used with intent to deceive or to win an argument unfairly, to carry conviction without justification, or to defeat proper discussion becomes a sophistic device.

This article will survey and classify the main kinds of fallacies, explaining and illustrating many that have been traditionally recognized and named, and noting especially those that are of particular importance in philosophy; and it will touch on the conditions in which fallacies flourish and the means by which they may be avoided or detected.

In classifying fallacies, we shall take first fallacies in the strict sense, forms of argument in which the conclusion does not follow from the premise or premises. These are divided into formal fallacies, errors in the formal reasoning itself, and informal fallacies, in which the reasoner either argues invalidly without using any precise logical form or goes wrong in putting a thought or an ordinary language statement into logical form or in translating back from logical form into thought or ordinary language. (It is a consequence of this division that if anyone commits an informal fallacy, there would be a formal fallacy somewhere in the argument that would be obtained if his intended premises, conclusion, and intermediate steps were put correctly and consistently into some logical form; but it is useful to distinguish informal fallacies in order to indicate how the mistakes have occurred.)

Next we shall take fallacies in nondeductive reasoning and in observation. We cannot speak accurately of fallacies in this case, since we no longer have strictly valid arguments with which to contrast them; but in a looser way we can contrast good procedures and patterns of reasoning that confirm hypotheses with ones that fail to confirm or are likely to produce errors.

Third, we shall examine fallacies in discourse. Such faults as inconsistency, circularity, prejudice, irrelevance, and unfair interrogation—which include some of the best-known fallacies—are not mistakes in reasoning from premises, or evidence, to a conclusion but are to be condemned on some other ground. Philosophical fallacies do not constitute a special group apart from those already mentioned, but some of these have been singled out for special notice.

FALLACIES IN THE STRICT SENSE

FORMAL FALLACIES. Formal fallacies may be arranged by reference to the logical systems, or parts of a logical system, whose valid argument forms they mimic or distort.

Hypothetical and disjunctive reasoning. Hypothetical and disjunctive reasoning is systematized by the calculus of propositions. The *p*, *q*, and other terms in the forms given below stand for variables that range over complete statements or propositions, and the phrases "If … then" and "Either … or" stand either for the corresponding truth operators or for any operators that, with respect to the arguments into which they enter, obey substantially the same calculus. The following fallacies are common in reasoning of this kind.

> *Asserting the consequent*: If *p* then *q*, and *q*, therefore *p*.
>
> *Denying the antecedent*: If *p* then *q*, and not *p*, therefore not *q*.
>
> *Converting a conditional*: If *p* then *q*, therefore if *q* then *p*. For example, "If this equation holds, so does that one; therefore, if that equation holds, so does this one."
>
> *Negating antecedent and consequent*: If *p* then *q*, therefore if not *p* then not *q*. For example, "If the nations disarm, there will be peace; so if the nations do not disarm, there will not be peace."

These invalid forms of argument are plausible partly because they are distortions of valid forms. The first two are distortions of *modus ponens* (If *p* then *q*, and *p*, therefore *q*) and *modus tollens* (If *p* then *q*, and not *q*, therefore

not *p*). Similarly, the third and fourth both mimic the valid form transposition (If *p* then *q*, therefore if not *q* then not *p*). However, concrete arguments of these invalid forms may also be explained as informal fallacies due to ambiguity (discussed under "Ambiguous Words and Phrases" below). An expression that actually asserts only a proposition of the form "If *p* then *q*" may be wrongly taken as asserting "*q* if and only if *p*," and if each of the conditionals above were replaced by the corresponding biconditional, each fallacy would become a valid form of argument.

It is also easy to fall into these fallacies when one is working in a field in which corresponding statements of the forms "If *p* then *q*" and "If *q* then *p*" are frequently both true or both false. This is the case in certain areas of mathematics, and indeed this fact is used in the procedure for discovering proofs that is sometimes called geometrical analysis. We assume the truth of what we wish to prove, and work out its consequences; if among these we find something that is already known or that can be proved independently, we try to construct a proof by retracing the previous steps. We assume that *p*, we deduce in a series of steps that *q*, we prove independently that *q*, and hence, reversing the previous deduction, that *p*. However, this final proof will be valid only if each of the steps in the analysis is reversible. Geometrical analysis is a useful heuristic procedure because this is often the case, but this utility in many geometric arguments may tempt us to assume, wrongly, that such steps are always reversible and that wherever we have established "If *p* then *q*," we are entitled to infer, from this alone, "If *q* then *p*."

Another common fallacy is that of *asserting an alternative*: Either *p* or *q*, and *p*, therefore not *q*. This is a distortion of the disjunctive syllogism (Either *p* or *q*, and not *p*, therefore *q*). However, concrete examples may also be explained as due to the ambiguity of disjunctive expressions, for if "Either *p* or *q*" were replaced by the strong disjunction "Either *p* or *q* but not both," this would be a valid form of argument.

There are also fallacies that are distortions of De Morgan's rules. Thus, "Not both *p* and *q*" is equivalent to "Either not *p* or not *q*," but we may invalidly infer from it "Both not *p* and not *q*"; and from "Either not *p* or not *q*" we may invalidly infer "Not either *p* or *q*."

Use of arguments. If a conclusion follows validly from a premise or set of premises, we can use this fact correctly in either of two ways. Given that the premises are true, we can infer that the conclusion is true; or, given that the conclusion is false, we can infer that at least one of the premises is false. However, these correct inferences may be replaced by the following fallacious ones:

(1) The conclusion is true; therefore the premise is true (or therefore all the premises are true).

(2) The premise (or at least one of the premises) is false; therefore the conclusion is false.

The first of these can contribute to confusion between the confirmation of a hypothesis and a proof of it; for when a hypothesis is confirmed, a conclusion drawn from it as a premise is found to be true, and the fallacy would make us infer from this that the hypothesis is itself true.

(3) The conclusion is false; therefore all the premises are false.

We might take as a variant of the above inference a fallacy noted by Aristotle and inappropriately named *non causa pro causa.* In this variant, an assumption is rejected because an argument in which it is used as a premise leads validly to a false or self-contradictory conclusion. This unsatisfactory conclusion is not due to this assumption, however, and would have followed from the other premises used without this assumption. In practice, one may slip into such an improper *reductio ad absurdum* (or *ad falsum*) either through not noticing that other premises besides the assumption are used, or through too easily taking them to be correct.

There are also fallacious ways of using the fact that an argument is invalid, such as:

(4) The argument from this premise (or these premises) to that conclusion is invalid; therefore the conclusion is false.

Examples of the first and fourth fallacies in the use of argument can also be explained in another way. The correct inference in each case is that the conclusion is not supported by the proposed argument, and we may confuse "not supported" with "false." Indeed, where the conclusion is the subject of controversy and we have previously had both arguments tending to show that it is true and arguments tending to show that it is false, the demolition of a supporting argument will shift the balance between the opposing views and will leave our reasons for denying this conclusion relatively stronger than they were before.

Traditional logic. The *simple conversion of an A-proposition* (or universal affirmative) is a common fallacy having the form "all *P* are *Q*, therefore all *Q* are *P*." For example, having agreed that whatever is conceivable is logically possible, we are liable to infer from this that any-

thing that is logically possible is conceivable. An equivalent error is the negating of terms in an A-proposition, that is, arguing from "All *P* are *Q*" to "All not-*P* are not-*Q*": "Whatever is conceivable is logically possible; therefore, anything that is not conceivable is not logically possible."

A similar fallacy is the *conversion of an O-proposition*: "Some *P* are not *Q*, therefore some *Q* are not *P*." An example is "Some states with parliamentary government are not democratic; it follows that there are genuinely democratic states which lack parliamentary government."

We can give a complete list of the possible formal fallacies in the traditional syllogism and sorites because the following set of four rules (one of them in two parts) is such that every argument that has the form of a syllogism or a sorites is valid if it obeys all these rules and is invalid if it breaks any of them.

Rule I. Not more than one premise may be negative.

Rule II. If one premise is negative, the conclusion must be negative, and vice versa.

Rule III. Each middle term must be distributed at least once.

Rule IV. If a term is distributed in the conclusion, it must be distributed in the premise in which it occurs.

In interpreting these rules, we take the subjects of universal propositions and the predicates of negative propositions to be distributed, and the subjects of particular propositions and the predicates of affirmative propositions to be not distributed.

There are, then, the following possible formal fallacies:

Two negative premises.

Negative conclusion with no negative premise.

Negative premise with no negative conclusion.

Undistributed middle. A middle term is not distributed in either of the premises it is meant to connect.

Illicit major. The major term, the predicate of the conclusion, is distributed in the conclusion but not in its premise.

Illicit minor. The minor term, the subject of the conclusion, is distributed in the conclusion but not in its premise.

Fallacies of the last three kinds are the most common and important. The argument "All machines work in accordance with causal laws, and all human beings work in accordance with causal laws; therefore all human beings are machines" commits the fallacy of undistributed middle because the middle term, "things that work causally," is undistributed in each of the premises as the predicate of an affirmative proposition. This fallacy is more plausible if the reasoning is expressed hypothetically: "Machines are causally determined, so if human beings were causally determined, they would be mere machines."

The fallacy becomes yet more plausible if the argument is extended to form a sorites: "Machines are causally determined, and they are not morally responsible for what they do; therefore, if human beings were causally determined, they would be no more morally responsible than machines are." The syllogism "All matters of taste are subjective, and no moral judgments are matters of taste; therefore no moral judgments are subjective" contains the fallacy of illicit major, for the term *subjective* is distributed in the conclusion but not in its premise. The fallacy is not obvious here, and it is still less obvious in the sorites "Matters of taste are subjective, but we do not dispute about matters of taste; since we do dispute about moral judgments, they cannot be subjective." However, the fallacy may be easily seen in an argument of the same form on another subject, such as "All birds are egg-layers; no insects are birds; therefore no insects are egg-layers." Similarly, the argument "All Victorian Gothic buildings have nonfunctional features, and they are all ugly; therefore all buildings with nonfunctional features are ugly" is an example of the fallacy of illicit minor, for the term "buildings with nonfunctional features" is distributed in the conclusion but not in its premise.

There are fallacies that consist in the mishandling of complex (conjunctive and disjunctive) terms. These include distortions of the De Morgan rules for terms, corresponding to fallacies noted above. Thus, it is fallacious to argue from "No policy will both defend freedom and insure peace" (*PeDI,* which is equivalent by obversion to $Pa\overline{DI}$) to "Every policy both fails to defend freedom and fails to insure peace" ($Pa\overline{DI}$).

Two traditionally recognized fallacies are the *fallacy of the accident* and the *converse fallacy of the accident,* which is also called the fallacy *a dicto secundum quid ad dictum simpliciter.* The latter consists in going invalidly from a qualified statement to an unqualified one—for example, in arguing from "It is always wrong to take another person's property without his permission" to "It is always wrong to take another person's property." (It is similarly fallacious to go from a statement qualified in one way to a like statement qualified in another way; both

these errors, when they occur in moral reasoning, amount to neglect of the principle that circumstances alter cases.) Considered formally, the converse fallacy of the accident consists in invalidly dropping a conjoined term, in arguing from "All *PQ* are *R*" to "All *P* are *R*." It is always fallacious to drop a conjunct from a distributed term, and we might therefore extend the traditional name to cover all cases of this sort. But what, then, is the fallacy of which this is the converse? Adding a conjunct to a distributed term is generally valid, but it is always a fallacy to add a conjunct to an undistributed term—for example, to argue from "Some snakes are poisonous" to "Some snakes native to Madagascar are poisonous"—and we may give this fallacy the traditional name of the fallacy of the accident. However, supposed examples of this are often really examples of the converse fallacy.

Parallel with the fallacies of dropping a conjunct from a distributed term and adding a conjunct to an undistributed term are the fallacies of dropping a disjunct from an undistributed term (All *P* are *Q* or *R,* therefore all *P* are *Q*) and adding a disjunct to a distributed term (No *P* are *Q,* therefore no *P* are *Q* or *R*).

Relational arguments. We may recognize certain arguments involving relations as being valid on account of some formal feature of these relations, such as symmetry or transitivity. There will then be a kind of fallacy that consists in treating a certain relation as if it had some formal feature that it does not have. Thus, it is fallacious to argue "Even an experienced doctor may be unable to distinguish diphtheria at an early stage from tonsillitis, or tonsillitis from an ulcerated throat; even an experienced doctor, therefore, may be unable to distinguish diphtheria at an early stage from an ulcerated throat," because the relation "is indistinguishable from" is not transitive. This invalid argument is plausible because this nontransitive relation can be confused with the transitive one "is exactly like."

Multiple and nonextensional operators. In multiply quantified statements, the order of two successive universal quantifiers can be changed. Thus, "Every man is always selfish" (which we can symbolize as $\Pi m \Pi t Smt$—"For every man, for every time, that man is selfish at that time") is equivalent to "At every time all men are selfish" ($\Pi t \Pi m Smt$). Similarly, "Someone at some time is selfish" ($\Sigma m \Sigma t Smt$) is equivalent to "There is a time at which someone is selfish" ($\Sigma t \Sigma m Smt$). However, "Every man is sometimes selfish" ($\Pi m \Sigma t Smt$) is not equivalent to "Sometimes every man is selfish" ($\Sigma t \Pi m Smt$—"There is a time such that every man is selfish at that time"); the latter implies the former but not vice versa. It is, therefore, a fallacy to change the order of successive quantifiers from universal-particular to particular-universal. Aristotle would have been guilty of this fallacy if he had argued directly from "Every activity aims at some good" to "There is a good at which every activity aims."

There are similarly invalid ways of changing the order of successive operators one or both of which are not quantifiers. "It is certain that someone will win" (which may be symbolized as $V \Sigma x Wx$) does not imply "There is someone who is certain to win" ($\Sigma x V Wx$), although the invalid inference from the first to the second is facilitated by the fact that "Someone is sure to win" is ambiguous between the two. George Berkeley's central argument (in Section 23 of the *Principles of Human Knowledge* and in the first of the *Three Dialogues*) contains an example of this fallacy. He showed, correctly, that a statement which we can formulate as follows is necessarily false: "There is something which someone truly believes not to be thought of" ($\Sigma m \Sigma m B m N T x$). However, he thought he had demonstrated the necessary falsity of the different statement "Someone truly believes that there is something which is not thought of" ($\Sigma m B m \Sigma x N T x$). Berkeley argued invalidly from the denial of the former statement to the denial of the latter, and so to the conclusion that it is absurd to maintain that material objects exist unconceived.

We should recognize, then, a *fallacy of rearranging operators.* Indeed, we could bring under this heading many fallacious forms of argument. Thus, the fallacies due to distortions of De Morgan's rules noted above consist in reversing the order of negation and conjunction, or of disjunction and negation. The invalid argument from "You are not obliged to resign" to "You are obliged not to resign" reverses the order of the deontic operator and negation; the fallacious "logical" proof of determinism, "Necessarily either you will go or you will stay; so either you will go necessarily or you will stay necessarily," reverses the order of the modal operator and disjunction; and so on.

Some operators set up nonextensional contexts, contexts in which terms or propositions that are extensionally equivalent cannot be validly substituted for one another. Whereas "Mrs. Jones shot the man in her bedroom," together with the fact that the man in her bedroom was her husband, entails "Mrs. Jones shot her husband," "Mrs. Jones intentionally shot the man in her bedroom" does not similarly entail "Mrs. Jones intentionally shot her husband." And even if "*p*" is logically equivalent to "*q,*" "Smith believes that *p*" does not entail "Smith believes that *q.*" It is still a matter of dispute how such

contexts should be explained and classified, and what kinds of substitution are valid in each sort of context; however, we can recognize, as a further type of fallacy, *extensional substitution in nonextensional contexts.*

INFORMAL FALLACIES. Many informal fallacies are due to ambiguity or vagueness of expressions used to make statements. If the terms used are vague or ambiguous, the expressions in which they are used will be correspondingly vague or ambiguous. However, the whole expression may be vague or ambiguous even if the terms are not, principally because a sentence form may be indeterminate as to the logical form it represents. We may, therefore, distinguish fallacies that arise from the ambiguity or vagueness of expressions in representing logical form from those that arise from other sorts of ambiguity or vagueness. Ambiguity or vagueness is not in itself a fallacy, but it may lead to fallacy. For example, someone may move invalidly from one assertion to another, but not notice that he has made any move at all because he uses the same ambiguous expression for his premise and for his conclusion. Or he may use an ambiguous expression to assert a premise, and thus infer a conclusion that would follow from one possible sense of that expression but does not follow from the sense he intends to assert. Or, having validly inferred a certain conclusion, he may assert a different conclusion, using an expression ambiguous between the validly derived conclusion and the one asserted.

Indeterminacy of expressions. A sentence such as "Men are unwise" may be ambiguous between "All men are unwise" and "Some men are unwise." It suffers from suppressed quantification. Similarly, if someone says "Courage and wisdom go together" (or "always go together," or "are constantly conjoined"), is he saying that all the courageous are wise, that all the wise are courageous, or both of these? Some philosophical terminology is ambiguous in just this way. If we say that one thing is a criterion of another, do we mean that it is a necessary criterion, a sufficient one, or both? Such indeterminacy may facilitate an invalid move from one meaning to the other, and in actual cases we may be unable to decide whether an arguer has committed the formal fallacy of simply converting an A-proposition or the informal fallacy of going from one sense to the other of an ambiguous expression.

Conditional expressions are often similarly indeterminate. "You will succeed if you make an effort" may say what it would be literally taken as saying ($m \supset s$), but with a different emphasis or in a different context it may mean "You will succeed only if you make an effort" ($s \supset m$), or perhaps the conjunction of these two ($s \equiv m$). Disjunctive expressions, while they are commonly used to express a weak disjunction, can be ambiguous between weak and strong disjunction; but logicians have themselves often fallen into a fallacy in supposing that whenever two disjoined terms are mutually exclusive, either necessarily or as a matter of fact, the disjunction is itself a strong (exclusive) disjunction. The truth is that when the disjoined terms are known to exclude one another, it makes no practical difference whether the disjunction itself is weak or strong.

The name of the *fallacy of division* has been given, by some modern writers, to attempts to argue from the premise that something is true of some whole, or of some class considered collectively, to the conclusion that the same is true of the parts of that whole, or of the class considered distributively (that is, of each of its members); and the name of the *fallacy of composition* has been given to arguments in the reverse direction. Either of these fallacies may be covered by an ambiguity of the word *all* between its collective and its distributive sense. This ambiguity of *all* leads us to commit the fallacy of division when we argue, for example, from the fact that all the citizens are strong enough to resist a tyrant (meaning that the citizen body considered as a whole has sufficient strength to do this) to the conclusion that all the citizens are strong enough to resist a tyrant (meaning that every citizen, considered individually, has sufficient strength to do this). We are in this case arguing from the statement made by a sentence in which "all" is used collectively to the statement made by the same sentence when "all" is used distributively. We are committing the fallacy of composition when we argue from the premise that every man can decide how he will act to the conclusion that the human race can decide how it will act (for example, with regard to the rate of increase of population or the choice between war and peace). In this case we move from the distributive to the collective sense of "all" in "All men can decide on their actions." This, or a similar fallacy, is committed whenever we assume, without adequate reason, that we can speak about groups in the same ways in which we can speak about their members, that we can speak of a nation having a will or interests, or of a society having problems. Of course, it may be possible to do this; there may be predicates applicable (in the same sense) to a group and to its members, but this cannot be assumed without evidence. It may also be possible to introduce a different but useful sense in which a predicate normally

applied to individuals may be applied to a group; but if so, the new sense must be explained.

However, what Aristotle called the fallacies of division and composition are different from these. He was speaking about changing the ways in which words are combined; for example, from "John is able-to-write while he is not writing" to "John is able to write-while-he-is-not-writing." In all such cases there is an ambiguity that conceals a fallacy of rearranging operators (the former example may be symbolized as *ΣtKMWatNWat*—"At some time both it is possible that John is writing at that time and John is not writing at that time"—and the latter as *MΣtKWatNWat*—"It is possible that at some time both John is writing at that time and John is not writing at that time"). The ambiguity of "All the men pushed, but could not move the stone" is really of this sort; the first clause is symbolized in one sense as *ΣtΠmPmt*—"There is a time such that every man pushed at that time"—and in another sense as *ΠmΣtPmt*—"For every man, there is a time such that the man pushed at that time." There need not be any question of ascribing the activity of pushing to a totality of men. In either case there are only individual pushings; but the statement in one sense says that these were simultaneous and in the other sense it does not. This contrast might also be referred to as a distinction between collective and distributive senses. There are, therefore, at least two distinct pairs of fallacies that have been called fallacies of composition and division, but if we speak about collective and distributive senses we tend to run the two pairs together.

Ordinary language seems to lend itself to ambiguities about operator order. Does "You can fool all of the people some of the time" mean that there are times at which the whole populace can be deceived (*ΣtMΠmDmt*—using *M* for "It is possible that" and *Dmt* for "that man is deceived at that time")? Or that every person is occasionally foolable (*ΠmΣtMDmt*)? Does "You can fool some of the people all of the time" mean that some people are capable of being permanently deceived (*ΣmMΠtDmt*), or that at every time it is possible to fool some people (*ΠtΣmMDmt* or *ΠtMΣmDmt,* these two being perhaps equivalent)?

However, in all the cases considered here, and in some of those to be considered in the next subsection, it may be questioned whether we should say simply that the fallacy is due to ambiguity or vagueness. We may fail to distinguish two kinds (or forms) of situations because we use the same expression to describe them, but it could also be that we use the same expression because we commonly fail to distinguish the two things. Informal fallacies, as considered here, are due to confusion as much as to ambiguity. We can conveniently explain them in terms of the ambiguity of various expressions, but we should not assume that the linguistic fact of ambiguity (or vagueness) is the sole or the primary cause of these errors.

Ambiguous words and phrases. Ambiguity is extremely common, but it is likely to lead to fallacy only in cases in which the different meanings of a word or phrase are close enough to be confused. One fallacy that can then arise is that of the *ambiguous middle,* that is, an argument may appear to have the form of a syllogism, but the expression we take as standing for a middle term may have different meanings in the two premises. For example, an authority on theology is more likely than other people to be right about theology, and a learned divine is an authority on theology. Does it follow that a learned divine has a better than ordinary chance of being right about theology? Not if the phrase "an authority on theology" means in the second premise an authority on the body of theological assertions but in the first premise means an authority on that which theological assertions are about. In such cases there is really no term common to the two premises, and therefore there is no genuine syllogism. There are also similar fallacies in which an expression is used in different senses in a premise and in the conclusion. Ambiguity often gives rise to these fallacies when the meaning of a word is fixed by its context, and the two different contexts give the word two different meanings. All these are instances of equivocation.

Some words are systematically ambiguous in a troublesome way. An observation may be either what is observed or the observing of it; a perception may be either a perceiving or what is perceived. There are similar indeterminacies about "experience," "sensation," and "belief." Such ambiguities constantly create difficulties in epistemology, the philosophy of science, and philosophical psychology.

There are also forms of speech that tempt us to confuse what we can say about words with what we can say about the corresponding things. A cause necessarily produces an effect, but only in the sense that it would not be called a cause if it did not. Similarly, murder is necessarily wrong, but not in the sense that there is a necessary connection or a rationally discoverable link between the kind of act called murder and its being wrong.

Sometimes when words are not ordinarily ambiguous, we perversely make them so; for example, by giving a word, in addition to its ordinary meaning, another meaning that is borrowed from a cognate word or a sim-

ilarly formed word. If John Stuart Mill confused "is desirable" (meaning "ought to be desired") with "can be desired," deriving this second sense from the use of "is visible" to mean "can be seen" and "is audible" to mean "can be heard," he was making a mistake of this kind. Similar results are produced by an idiosyncratic use of language. It is hard to keep to a sense specially assigned to a word, and we are always liable to slip back into some more conventional use. When a psychologist has redefined "learning" in relation to some special procedure by which "learning" can be measured, he or his readers may think that what he then discovers is true also about learning in its ordinary, much broader sense.

Such unwarranted generalization, considered formally, exemplifies the fallacy *a dicto secundum quid*; in practice, however, it is aided by various ambiguities and confusions. Thus, the words *class* and *set* may be confined to finite collections or may embrace infinite ones as well. We are liable to argue from the fact that something holds for all finite classes or sets to the conclusion that it holds for all classes or sets, including infinite ones, partly because the words are ambiguous, partly because we fail to notice that the wider concept is a different one, and partly because we generalize from specimen cases and choose specimens that are more easily visualized but are not fully representative.

As we have noted, errors may arise not only from ambiguity as such but also from the confusing of things that, although similar or related, are different. A classic example of this, of great importance in philosophical discussion, is the confusing of separation with distinction. Thus the distinction between analytic and synthetic statements may be attacked, fallaciously, on the ground that actual statements are difficult to assign, without reservations, to one category or the other. Confusion here is due partly to failure to see what sorts of things are being distinguished—not verbal forms, not sentences, but ways of using sentences to make statements.

When this obscurity is removed, however, we may still have to defend the distinction against the critic who says that because of indeterminacies in the use of component words, every concrete use of a sentence in order to make a statement lies somewhere between being analytic and being synthetic. Even if this critic were right, this would in no way count against the distinction. Indeed, such a status makes it particularly important to draw the distinction, in order to expose the common fallacy of arguing from a statement in which words are so used as to make the statement analytically true to a synthetic statement made by the same words in a different sense (as might be done with the statement "A change in the moral code means social disintegration").

This confusion can also be used in the opposite way. It may be argued that because two things can be distinguished, they must be separate—for example, to argue that since we can distinguish a motive from a cause, things that have causes cannot have motives, or that a person's having a certain motive cannot be a cause of his action.

FALLACIES IN NONDEDUCTIVE REASONING AND IN OBSERVATION

Outside the sphere of deductive reasoning, we can speak of fallacies only in an extended sense. For example, we can contrast genuine confirmation of hypotheses with something that is mistaken for it, probable arguments that give some support to their conclusions with ones that do not, and, in general, techniques and procedures that tend to give correct results with ones that tend to produce error. However, it would be pointless and misleading to call a piece of inductive reasoning, say, fallacious, merely because its conclusion turned out to be false.

INDUCTION AND CONFIRMATION. We may note two fallacies *about* induction or confirmation: the mistaking of confirmation for proof, and the demanding of proof where no more than confirmation is possible. There are also fallacies *in* induction and confirmation. Where scientific or commonsense reasoning follows the lines of one of the eliminative methods of induction, failure to observe the requirements of that method will count as a fallacy. Thus, in reasoning along the lines of the method of agreement, it will be a fallacy to conclude that there is a causal relation between the phenomenon *P* and a certain feature *A,* merely because occurrences of *P* are repeatedly found to be accompanied or preceded by occurrences of *A,* without trying to discover other possibly relevant features common to these occurrences of *P* or, what amounts to the same thing, without trying to find occurrences of *P* that are as relevantly diverse as possible and then seeing whether *A* is present in them all. Thus, it is fallacious to conclude that William is allergic to strawberries from the evidence that his allergic symptoms have repeatedly appeared after he has eaten strawberries, if William has eaten strawberries only in one particular house, at a particular sort of gathering, and so on.

Similarly, in reasoning along the lines of the method of difference, it will be a fallacy to conclude that *A* is even an indispensable part of a sufficient condition for *P* from

a comparison of a case in which *P* and *A* are both present and a case where they are both absent, without checking that the two cases are otherwise relevantly alike, that no likely-to-be-relevant feature except *A* differentiates the case in which *P* occurs from the one in which it does not. In other words, it is fallacious to use a control case that differs from the experimental case in some unwanted respect. Thus, it is fallacious to infer that John's having recovered more rapidly than James is due to a drug that was given only to John, if John was also told that he was having a new treatment and the doctors and nurses all took special care of John because they were interested in the experiment. There can be correspondingly unsound experimental procedures, and corresponding errors in reasoning, in applications of the method of concomitant variation.

Post hoc ergo propter hoc is traditionally listed as a fallacy; but much respectable inductive reasoning would fall under this heading, and it is not to be condemned because it is not deductively valid. We argue, reasonably, that the one likely-to-be-relevant change causes the result that follows. We are, in effect, taking the "before" situation as the control case and the "after" situation as the experimental case. This is a fallacy only if we ignore other likely-to-be-relevant changes.

All such mistakes can be summed up as consisting in failures to test the hypothesis in question—that *A* is (in some sense) the cause of *P*—that is, in failure to look for what, if the hypothesis were false, would be most likely to reveal its falsity. If *A* is not the cause of *P*, we are most likely to reveal this by finding cases of *P* so diverse that *A* is not present in them all, or a control case so like the experimental case that *P* occurs in both, or occurs in neither, in spite of *A*'s being present in one and absent from the other.

Another inductive fallacy is to take a hypothesis as being confirmed by observations to which it is irrelevant, when without this hypothesis our other knowledge and beliefs would explain what is observed equally well. Further, since it is a basic principle of inductive reasoning that alternative hypotheses should be considered, and that to confirm one hypothesis we must eliminate its rivals or show them to be improbable, it is a fallacy to take a hypothesis as being confirmed by observations that are equally well confirmed by an intrinsically more probable alternative hypothesis—for example, to take the Michelson-Morley experiment as confirming the theory of relativity without eliminating the FitzGerald-Lorentz contraction and the emission hypothesis of the velocity of light.

We may add a fallacy of saving hypotheses. It is certainly a fault for a thinker to be so attached to a hypothesis that he notices only evidence that agrees with it and ignores or denies unfavorable evidence. Popular superstitions of all kinds are protected by this fallacy, but it is also common among scientists, historians, and philosophers. It may also be a mistake, when one finds evidence that is prima facie unfavorable, to introduce supplementary ad hoc hypotheses in order to protect the original one from falsification. Carried to an extreme, this procedure constitutes a linguistic change that makes the original hypothesis analytically true, and it can generate the fallacy described above of oscillating between an analytic and a synthetic use of the same expression. In less extreme cases, how can we systematically mark off this error from the respectable procedure of interpreting new observations in the light of an established theory? Perhaps in two ways: first, in the respectable procedure, we are working with a hypothesis that is already well confirmed, but it is a fallacy to "save" a hypothesis for which there is no strong independent support; and second, even if the original hypothesis was well confirmed, it may be appropriate to consider, after it has been "saved" by additional hypotheses (after the new observations have been interpreted in the light of the original hypothesis) or has been modified and qualified in various ways, whether some alternative hypothesis would account better for the whole body of evidence.

ANALOGY. All arguments from analogy are fallacious in the sense that they are not deductively valid. However, we often want further to distinguish weak analogies from strong ones and to suggest that a weak analogy is completely fallacious but that a strong analogy has at least some force. In an analogy we compare two things, *A* and *B*; we find some resemblances, say *X, Y, Z,* between them; and then we argue that since *A* has some further feature *P*, it is likely that *B* also has this feature. We are inclined to say that if the points of resemblance *X, Y, Z* are few or trivial, the analogy is weak or far-fetched, but that it is a strong analogy if there are many important points of resemblance. An alternative way of looking at the distinction is that to use this analogy is implicitly to frame and then use the hypothesis that all things that have the features *X, Y, Z* also have the feature *P*. The analogy will be weak if we already have evidence that falsifies this hypothesis or makes it implausible, but it will be strong if we have no such evidence and what we know about *A* somehow constitutes good inductive evidence for a connection between *X, Y, Z,* and *P*.

CLASSIFICATION. Faults in classification can in several ways give rise to fallacies in either the strict or the extended sense. If things are classified under headings where they simply do not belong, the classification implicitly asserts false propositions which may be used as premises in arguments that, even if formally valid, will therefore give no real support to their conclusions. If a classification is based on unimportant resemblances, this may give rise to weak analogies and to the framing of unlikely hypotheses, and inductive reasoning that uses such a classification—in the methods of agreement and difference, for example—will give an appearance of support to conclusions that are not really supported by the evidence as a whole. Again, if the division of a class into subclasses is not exhaustive, it may be wrongly taken to be so, and this will provide a false premise for a disjunctive argument. Thus, if we divide trees into conifers and deciduous trees, we may infer that since eucalypts are not conifers, they are deciduous. Similarly, a division that is not exclusive may be wrongly taken to be so; the same division of the class "trees" may lead us to infer that larches, being conifers, are not deciduous.

Two important fallacies concerned with classification arise from the attempt to draw sharp distinctions where the facts show a continuous (or near continuous) gradation. Is a man bald if he has one hair on his head? or two? or three? And so on. Just what degree of mental disorder is to count as insanity? One fallacy consists in assimilating every intermediate case to one or the other of the extremes and is exemplified in the black-and-white thinking that divides people into normal individuals and lunatics or states into peace-loving nations and warmongers. The contrary, and more subtle, fallacy consists in arguing that because there is no break in the gradations, there is no distinction even between the extremes—concluding, for example, that we are all insane—as if the problem about when a man is bald showed that there is no difference between a man with a completely smooth scalp and one with a full head of hair.

STATISTICS. We can deal here only with some elementary mistakes in statistical reasoning. One of these consists in paying attention to simple frequencies or proportions rather than to correlations. If a high proportion of atheists are honest, this in itself does not indicate any sort of causal connection between atheism and honesty; the first thing to discover is whether the proportion of honest people is higher among atheists than among nonatheists. Similarly, the frequency of persons who have both mathematical ability and artistic talent may be small in the population as a whole; but if only one person in ten has mathematical ability and only one in ten has artistic talent, then only one in a hundred would have both, even if there were no natural opposition between these gifts. Before we conclude that these abilities tend to occur separately, we must find whether artistic talent is more or less common among the mathematically able than among the rest of the population.

Another common statistical fallacy consists in directly inferring a causal connection from a positive correlation: given a positive correlation between cigarette smoking and lung cancer, it is a further question whether this is to be explained by a causal connection between them. An associated fallacy of confusion, which is becoming more common, is simply not to talk about causation but to use the word *correlation* as if it meant causal connection, for example, to infer predictions and practical recommendations directly from correlation statements. Another fallacy is the neglect of the requirements of significance. Essentially, this consists in taking as causally informative, or as representative of a similar correlation in a larger population, a correlation within a sample that could equally well be explained as a chance result. This is, therefore, an instance of the neglect of alternative hypotheses.

Even when there is good statistical evidence for a causal connection between two features *A* and *B,* it is a mistake to conclude immediately that one, say *A,* is the cause of the other without having considered and excluded the possibilities that *B* may tend to produce *A,* that *A* and *B* may be joint effects of some other cause, and that there may be causation in more than one direction. For example, a positive correlation between poverty and ill health might be due to the fact that poverty causes ill health, to the fact that ill health diminishes earning capacity and wastes resources, to the fact that stupidity, idleness, or drunkenness tends to produce both poverty and ill health, or to a combination of more than one of these causal tendencies.

PROBABILITY. Fallacies in reasoning about probability arise mainly from failure to attend to the fact that a probability is relative to certain evidence and changes as the evidence changes. The best-known is the gambler's fallacy. For example, since it is unlikely that a penny will fall heads up five times in a row, the gambler reasons, when it has fallen heads four times, that it is unlikely to fall heads at the next throw. But although the probability of five heads, relative to the knowledge that an unbiased penny is tossed in a random manner five times, is 1/32, the probability of this result, relative to the conjunction of this

knowledge with the knowledge that it has fallen heads on the first four throws, is 1/2.

OBSERVATION. It is questionable whether we should follow Mill and speak of fallacies of observation. Many of the items so described consist of errors in reasoning rather than in observation, and so fall under other headings. We may, however, note the following principles:

First, there are errors of nonobservation, which may be due to deficiency of one's senses or sense apparatus, to carelessness, or to the tendency to see only what we want to see. This may include the nonobservation that is one way of saving hypotheses.

Second, any of the above-mentioned causes may equally produce misobservation.

Third, it is impossible to separate, and difficult even to distinguish, observation from interpretation: we always have some conceptual framework, some expectations that determine how we shall observe what we observe. For example, we expect an object that looks like an adult human being to be between five and six feet tall, and we therefore tend to see any such object as being at a distance that would agree with this. The actual material to which our prior concepts are applied may not conform to them, however, and then we may make wrong judgments through using these concepts. Also, if we do not realize how observation and interpretation are mixed together, we may give the authority of an observed fact to a judgment that really rests on our preconceptions.

Fourth, our perceptual mechanisms automatically allow for factors that have been constant or to which it is inconvenient to attend, and errors arise when allowances are made for what is no longer there; for example, the illusion that the land is moving when we first go ashore after becoming used to the rolling of a ship.

Fifth, we may in perception confuse relations, say of comparison, with intrinsic qualities. This explains the illusions of contrast. For example, if after having had one's left hand in cold water and one's right hand in hot water, one puts both hands into lukewarm water, the lukewarm water feels hot to the left hand and cold to the right hand because it really is *hotter than* the left hand (or than what it has just been feeling) and *colder than* the right hand (or than what it has just been feeling).

Sixth, we may mislocate what we observe. In particular, we have a tendency to project and to treat as objective, as belonging to some external state of affairs, the feeling that the state of affairs arouses in us (the *pathetic fallacy*) or to mistake connections within our thoughts for connections between the corresponding objects. There is no room here for a full discussion of perceptual illusion and observational error, but it seems that many varieties of these can be explained by reference to one or more of these principles.

FALLACIES IN DISCOURSE

INCONSISTENCY. A position or a system of thought cannot be sound if it contains incompatible statements or beliefs, and it is one of the commonest objections to what an opponent says that he is trying to have it both ways. Inconsistency has many possible sources, but one that is of special importance in philosophy is the case in which a thinker, in order to solve one problem or deal with a particular difficulty, denies or qualifies a principle he has previously adopted, although in other contexts he adheres to the principle and uses it without qualification.

Inconsistency is a formal feature and can be formally checked, although it may also be concealed by the use of different expressions with a single meaning. It is not the same as invalidity, however; indeed, any argument with incompatible (or self-contradictory) premises will be formally valid. It is particularly important to detect inconsistencies in a set of premises, for an argument with inconsistent premises, even though valid, gives no support to its conclusion; and using one is not a satisfactory way of establishing anything or of convincing an audience.

On the other hand, it is a formal fallacy to suppose that because your opponent has tried to have it both ways, he cannot have it either way—that every part of an inconsistent position must be false.

PETITIO PRINCIPII. An argument that begs the question, that uses the conclusion as one of the premises, is always formally valid. A conclusion cannot fail to follow from a set of premises that includes it. This is also a fallacy only in the extended sense that such an argument gives no support to its conclusion. One kind of petitio principii consists in arguing in a circle, when one proposition is defended by reference to another, and the second is defended by reference to the first. For example, we may argue that a certain historian is trustworthy because he gives a balanced account of some episode, but also rely on that historian's account in order to decide what actually occurred in this episode, and hence to decide what would be a balanced account of it.

The larger and more complex a circle of argument is, the harder it is to detect the fallacy. One result of circularity is that the propositions that have been proved from

one another appear to have been conclusively established, although no empirical evidence has been given for either of them. This can create an illusion that there are synthetic propositions that have no need of empirical support. This may be combined with a fallacy of confusion, of failing to distinguish the coherence or consequential character of a system from its truth—a confusion that has developed into the coherence theory of truth and that is still encouraged by some eccentric uses of the word *true* or of such a phrase as "true within the system."

Circularity is common in moral reasoning, and here, too, it may make us think that moral conclusions can be rationally established without reliance on observations, intuitions, choices, or decisions. The exposure of such circularity compels us to give a more adequate account both of how moral judgments are to be supported and of how they are to be interpreted.

A PRIORI FALLACIES. Under the heading of a priori fallacies Mill listed a number of natural prejudices, including the popular superstition that words have a magical power and such philosophical dogmas as that what is true of our ideas of things must be true of the things themselves; that differences in nature must correspond to our received (linguistic) distinctions; that whatever is, is rationally explicable; that there is no action at a distance; that every phenomenon has a single cause; and that effects must resemble their causes. These are all errors, but we can go further and recognize a general a priorist fallacy, which consists in trying to base knowledge of fundamental synthetic truths on anything other than empirical evidence. These examples illustrate how once we start looking for a priori truths, we are led to try to distill them from language or from our ideas (giving each of these an authority to which it is not entitled), or to confuse continuity with intelligibility and necessity, or to dignify with the title of a priori truths what are no more than sweeping generalizations from the simplest and most familiar observations.

More generally still, we can recognize a fallacy of prejudice, which consists in believing without evidence, in adopting or adhering to views on any subject without any relevant reason. It is worth noting that adopting a method of argument (other than a deductively valid form) is tantamount to adopting an assertion. For example, regularly to judge the rightness of actions by their utility is tantamount to adopting the principle that whatever maximizes utility is right; and, again, regularly to argue that because a statement cannot be verified, it is meaningless is tantamount to adopting a verifiability theory of meaning. This is a particularly easy way of committing the fallacy of prejudice.

IGNORATIO ELENCHI. The fallacy of ignoratio elenchi consists in missing the point, in arguing for something other than what is to be proved. However, we can speak in this way only if the context somehow determines what is to be proved. In the first place, the context may be a discussion between *A* and *B,* and *B* will commit this fallacy if he claims to be replying to what *A* has said but fails to come to grips with *A*'s argument—for example, if he tries to disprove some proposition that *A* has not asserted either as a premise or as a conclusion. *B* is also guilty of this fallacy if he bluntly denies something that *A* has claimed to prove but does nothing to rebut *A*'s proof. Alternatively, it may be a thinker's general position or some long line of argument that makes it imperative for him to establish some point, and makes him guilty of irrelevance if he establishes something else instead.

There are a number of common and important types of irrelevance in discussion. If the question is whether a certain view is true or false, it is irrelevant to argue that adopting this view will be beneficial or pernicious. Thus, a body of religious doctrine may be irrelevantly defended on the ground that it makes people happier or better behaved. Similarly, the origin of a belief is in general irrelevant to the question of its truth; but if the fact that a belief is widely held has been used as evidence of its truth, then this reasoning may be relevantly rebutted by showing that the belief has come to be held for reasons or from causes that are independent of its truth. The truth of the belief and the account of its origin are in this case alternative explanatory hypotheses. That a view is held by certain people is also in general irrelevant to its truth, so that appeals to authority are usually examples of ignoratio elenchi. Cases in which the authority appealed to can be independently shown to be an authority in the sense of being likely to be well-informed about the point at issue are exceptions. Irrelevancy shades into prejudice; we may readily accept the doctrines of "our party" and reject those of "the enemy." In this, there may also be present a fallacy of confusion, in that we treat factual beliefs as if they were items of another category—principles to which we can adhere or subscribe, or which we can reject, by choice.

Another form of irrelevance is the tu quoque, or "two wrongs" technique. If some action or view of one's own is criticized, one may reply by attacking some action or view of one's critic that is equally hard to defend. The *argumentum ad hominem* is similar—we reject what

someone says on the irrelevant ground that he is in no position to say it. However, an *argumentum ad hominem* may quite properly point to an inconsistency, and may validly establish the limited conclusion that this man cannot consistently hold this view—a conclusion that may be of special interest in a moral discussion, where the problem may well be that of finding a policy that is both coherent and acceptable.

Related fallacies of irrelevance have been named *argumentum ad verecundiam* (appeal to authority or to feelings of reverence or respect), *argumentum ad personam* (appeal to personal interest), and *argumentum ad populum* (appeal to popular prejudice). Sometimes an *argumentum ad ignorantiam* or *ad auditores* is grouped with these, but these names seem to refer not to any specific fallacy but to the use of any unsound argument that is likely to deceive the actual audience.

FALLACIES OF INTERROGATION. There are two forms of the *fallacy of many questions.* In one, two or more questions are asked together, and a single answer is demanded to all of them. This is fallacious in that it unfairly prevents the person asked from giving different answers to the different questions. In the second form, the question asked has a presupposition that the answerer may wish to deny but which he would be accepting if he gave anything that would count as an answer. Thus, an answer of either "Yes" or "No" to the question "Have you left the party?" would be an admission of having been a party member, and any answer to the question "Why does such-and-such happen?" presupposes that such-and-such does happen. There is no fallacy, however, in merely asking a question that has a presupposition; the fallacy lies in demanding an answer in the narrow sense, in not permitting or in discouraging a reply that denies the presupposition. Again, it is an instance of the fallacy of prejudice to ask a question that has a presupposition without first investigating whether that presupposition is correct.

FALLACIES IN EXPLANATION AND DEFINITION. Just as a circular argument fails to give support, so a circular explanation fails to explain. There are concealed circularities of explanation; for example, some mental performance is explained by reference to a faculty, but further inquiry shows either that to say that this faculty exists is only to say that such performances occur or that, although more may be meant, there is, apart from such performances, no evidence for the existence of the faculty. Words like "tendency," "power," "disposition," and "capacity" lend themselves to circularities of this sort.

Similarly, a circular definition, in which the term to be defined recurs within the definiens, fails in its task. If it is intended as a stipulative definition, it fails to assign a meaning; and if it is intended as a reportive definition, it fails to inform anyone of the meaning with which the word is used.

Stipulative definitions can create ambiguity when we assign one meaning to a word but also retain another meaning. This amounts to an assertion that the two meanings go together, disguised as the innocent procedure of stipulation. Persuasive definition is an instance of this in which the retained meaning is an emotive one.

It is a fallacy, in the extended sense, to use words without meaning. But it is not a fallacy not to have defined one's terms, provided that they have a meaning that is known to the audience and is precise enough for the purpose in hand. On the contrary, since it is impossible to define all one's terms, it is a fallacy in discourse to demand that in all terms should be defined; a demand for definition can be a sophistic device for preventing the discussion of substantive issues.

PHILOSOPHICAL FALLACIES

THE NATURALISTIC FALLACY. What G. E. Moore called the *naturalistic fallacy* is the identifying of goodness with any natural characteristic, such as pleasantness or being the object of desire. If there is a distinct property, goodness, it will of course be an error to identify it with any other feature, even if the two are coextensive, and this would be an example of the refusal to distinguish what we cannot separate; however, it must first be shown that there is such a property as Moore's goodness. Alternatively, if it is a question of how the word *good* is commonly used, then it would be an error to say that it is used to convey some natural description. However, if the naturalist is not trying to report the ordinary use, but is saying that this ordinary use is somehow unsatisfactory (and also that there is no such property as the one of which Moore speaks) and is therefore proposing a different use, where is his mistake? It is true that if he redefines "good" as the name of some natural characteristic, but still also uses the word in its ordinary evaluative or prescriptive sense, he will be slipping into a fallacy of ambiguity; but a consistent ethical naturalist may be committing no fallacy at all.

ARGUING FROM "IS" TO "OUGHT." An error exposed by David Hume, but still frequently committed, is that of arguing from premises that contain only descriptive terms, and no copula except "is," to a conclusion that con-

tains an "ought." This is a fallacy in the strict sense; arguments of this sort cannot be valid, but they are often made plausible by the ambiguous use of such words as *reasonable, fitting, authority, desirable, beneficial, courageous, temperate, just, right,* and *good* itself, any one of which may be used first in a purely descriptive sense and then interpreted in a sense that is partly descriptive and partly prescriptive. A currently popular version of this fallacy combines it with one or more of the a priori fallacies. Since our concept of, say, courage or our ordinary use of the word *courage* combines a certain natural description with a certain prescription or evaluation, it is concluded that behavior that conforms to this natural description must be recommended or valued in just this way; that is, the move from *is* to *ought* is covered by an appeal to the supposed authority of our language or our ideas.

CONFUSING RELATIONS WITH THINGS OR QUALITIES. A group of philosophical errors that is less well known than the two just mentioned but at least as widespread and harmful consists in identifying a quality with a relation, in treating a relation as if it were an intrinsic quality of one of its terms, or in constructing fictitious entities out of relational situations. Presented linguistically, this means that a term is treated as standing both for a thing or a quality and for a relation, and this, like other ambiguities, can make synthetic connections appear necessary. Thus, an idea or a sense datum is supposed to be an object of which someone can be aware and to have this relation—someone's being aware of it—as part of its nature. This conflation generates a supposed matter of fact about which one can have infallible knowledge and thus gives rise to the pseudo problem of bridging the gap between this direct and infallible knowledge and ordinary fallible knowledge of objects that do not have being known built into their natures. Similarly, minds (or consciousness) have been treated as things that have as part of their nature the relation of being aware of something, and this generates difficulties in philosophical psychology. Also, errors that the naturalistic fallacy was meant to cover are better dealt with in this category. Goodness may be both treated as an intrinsic quality (natural or nonnatural) of, say, states of affairs and identified with or taken as logically including the relation of being pursued, aimed at, or recommended; indeed, it seems that it is just such a conflation of features that makes a quality nonnatural. Similarly, beauty may be both treated as an intrinsic quality and identified with or taken as logically including such relations as pleasing or being admired.

CATEGORY MISTAKES. Philosophers now carefully distinguish different uses of language, different "language games"; the contrasting error is to confuse different ways of using words, to treat a term that belongs to one category as if it belonged to another. However, the concept of a *use* of language is itself ambiguous. In distinguishing uses, we may be noting differences that lie within language, differences in the relations between words and things, differences in the things to which our expressions apply; and it will be a mistake to confuse one kind of distinction with another. There is also a tendency to think that, at least in philosophy, we cannot employ this distinction between words and things; this view is supported by a variant of Berkeley's fallacy: Since we cannot talk about something except by using words in relation to it, it is supposed that we cannot talk about things as they are, apart from relations to words.

AVOIDANCE AND DETECTION OF FALLACIES

Popular discussions of fallacies rightly lay great stress on the psychological or emotional aspect of fallacious arguments. Under the influence of violent passions, thinking becomes more purely associative and less consequential, and we are more than usually ready not only to employ arguments, however unsound, that appear to support whatever cause we espouse but also to extend our favor to anything linked, however loosely, with what we already like, respect, or admire, and to extend our hostility to anything linked with what we already dislike, despise, or fear. Ridicule can also be used to brush aside relevant considerations and to condemn a person or a view without a hearing. All sorts of attachments, passions, and emotional prejudices can foster fallacies, and one of the chief means for the avoidance or detection of fallacies is to consider a problem calmly.

Precise formal statement often helps in the detection not only of fallacies in the strict sense but also of inconsistency, circularity, and irrelevance. However, since it is too laborious a task to state all our reasonings formally, we can use this device only when we already have reason to suspect a fallacy. Also, in cases involving equivocation or a category mistake, there is a danger that inaccurate formulation will conceal the fallacy instead of exposing it.

As Richard Whately remarked, "a very *long* discussion is one of the most effective veils of Fallacy; … a Fallacy which when stated barely … would not deceive a child, may deceive half the world if *diluted* in a quarto volume" (*Elements of Logic,* p. 151). Consequently, an important weapon against fallacy is condensation,

extracting the substance of an argument from a mass of verbiage. But this device too has its dangers; it may produce oversimplification, that is, the fallacy *a dicto secundum quid,* of dropping relevant qualifications. When we suspect a fallacy, our aim must be to discover exactly what the argument is; and in general the way to do this is first to pick out its main outlines and then to take into account any relevant subtleties or qualifications.

See also Aristotle; Bentham, Jeremy; Berkeley, George; Conditionals; Definition; Hume, David; Induction; Logical Terms, Glossary of; Mill, John Stuart; Mill's Methods of Induction; Moore, George Edward; Probability; Truth and Falsity in Indian Philosophy; Whately, Richard.

Bibliography

The pioneer work on fallacies is the *De Sophisticis Elenchis* of Aristotle. Medieval and later logicians followed and expanded his account. Many textbooks on logic include a chapter on fallacies. Jeremy Bentham, throughout his writings, paid much attention to fallacious reasonings by which views that he opposed were supported, and he collected many of them in *The Book of Fallacies,* which is in Vol. II of his *Works,* edited by J. Bowring (Edinburgh, 1843). Richard Whately, in Ch. 3 of his *Elements of Logic* (London, 1826), gave a much improved classification and analysis of fallacies. John Stuart Mill devoted Book V of *A System of Logic* (London, 1843) to an account of fallacies, developing a new classification and concentrating on a priori fallacies (prejudices) and mistakes in observation and generalization.

Augustus De Morgan, in Ch. 13 of *Formal Logic* (London, 1847), rejected the attempt to list all possible ways of going wrong but gave a penetrating and well-illustrated analysis of many of the traditionally listed fallacies. Arthur Schopenhauer, in "The Art of Controversy," in *Essays from the Parerga and Paralipomena,* translated by T. B. Saunders (London, 1951), described stratagems that may be used in disputes, that is, both sophistic devices and ways of countering them. H. W. B. Joseph included in *An Introduction to Logic* (London, 1906) an appendix on fallacies based on the Aristotelian account.

M. R. Cohen and Ernest Nagel, in Ch. 19 of *An Introduction to Logic and Scientific Method* (New York, 1934), emphasized abuses of scientific method. R. H. Thouless in *Straight and Crooked Thinking* (London, 1930), Susan Stebbing in *Thinking to Some Purpose* (Harmondsworth, U.K., 1939), and W. W. Fearnside and W. B. Holther in *Fallacy—The Counterfeit of Argument* (Englewood Cliffs, NJ, 1959), gave lively and readable accounts, illustrated with many examples of popular errors and of sophistry in practice, concentrating on political and social debate and propaganda, and stressing the emotional basis of a great deal of fallacy.

J. L. Mackie (1967)

FARABI, AL

See *al-Fārābī*

FARADAY, MICHAEL

(1791–1867)

Michael Faraday, the British chemist and physicist, came from a poor family and had no formal schooling beyond the elementary level. While a bookbinder's apprentice, he became interested in chemistry and electricity. Faraday took notes on a series of lectures given by Sir Humphry Davy, the leading British chemist, presented them to Davy, and soon afterward, at the age of twenty-one, was appointed laboratory assistant to Davy at the Royal Institution (London). He became director of the laboratory in 1825 and Fullerian professor of chemistry at the institution in 1833. His early scientific work in chemistry included the discovery of several new compounds and the liquefaction of chlorine and other gases. In 1831, Faraday discovered electromagnetic induction, or the creation of electric currents in a conductor by changing currents or moving magnets in the vicinity; this phenomenon is the basis of the electrical generator. This was followed by a series of investigations demonstrating with greater certainty than had been previously achieved the identity of the nature of the electricity generated by friction, voltaic cells, electromagnetic induction, and other means. Extensive experiments in electrochemistry led Faraday to the enunciation of his laws of electrolytic decomposition in 1833. The source of the power of the voltaic pile, or battery, was the object of his subsequent research. He investigated the electrical properties of insulators, or dielectrics, in 1837. In 1845 he discovered that the plane of polarization of light was rotated on passing through a transparent diamagnetic substance in the direction of externally applied lines of magnetic force. At the same time he began his investigation of diamagnetism. In his last years he suffered from loss of memory, and he ceased his researches in 1855. He was a member of a small Christian sect, the Sandemanians, and was noted for his gentleness of character.

Faraday is generally regarded as one of the greatest of all experimental scientists. The truth of this, adequately attested to in the three-thousand-odd paragraphs of the *Experimental Researches in Electricity,* should not be allowed to obscure the fertility of his imagination and conceptualizing powers and the guiding role of theoretical principles in sustaining his persistent research. His most important contribution to physics is probably the

concept of lines of force, which was the beginning of the development of field theory. The accepted approach to electrodynamical phenomena at the time was to express the forces between charges mathematically as direct actions at a distance, an approach that was to prove unfruitful. Faraday was not trained in the mathematics necessary for this tradition. In order to represent the action of electromagnetic induction, he envisaged the space surrounding magnets to be filled with lines of magnetic force representing everywhere the direction of the force that would be experienced by a magnetic pole introduced from outside in the manner of the lines formed by iron filings sprinkled on a paper resting on a magnet. The lines of magnetic force have not only direction but also sense—that is, a north magnetic pole is pushed one way along them, and a south pole is pushed in the opposite sense; furthermore, their concentration near a given point represents the intensity of the magnetic force at that point. Each such line forms a closed loop, beginning or ending nowhere, but in the case of a magnet passing through its substance from one pole to the other. In these terms the law of electromagnetic induction may be expressed: "The quantity of electricity thrown into a current is directly as the amount of curves intersected" (*Experimental Researches in Electricity*). In James Clerk Maxwell's famous words:

> Faraday, in his mind's eye, saw lines of force traversing all space where the mathematicians saw centres of force attracting at a distance: Faraday saw a medium where they saw nothing but distance: Faraday sought the seat of the phenomena in real actions going on in the medium, they were satisfied that they had found it in a power of action at a distance impressed on the electric fluids. (Preface to the first edition of the *Treatise on Electricity and Magnetism*)

In most of Faraday's researches the concept of lines of force was used merely as a "representative aid" and was not meant to include "any idea of the nature of the physical cause of the phenomena." This cautiousness was a mark of Faraday's methods; in the choice of terminology to describe new phenomena, for example, he carefully attempted to avoid suggesting anything more than they warranted. However, at times Faraday allowed himself to speculate, and in 1852 he considered "the possible and probable *physical existence* of such lines" ("On the Physical Lines of Magnetic Force"). On the basis of arguments that can be characterized only as suggestive (such as that the magnetic lines are curved), he hypothesized that magnetic lines of force have physical existence and contrasted this with gravitation, where there was no evidence that the lines of force are anything but abstract and ideal. In a charming talk published in 1846 ("Thoughts on Ray-Vibrations") Faraday speculated that the atoms of matter might be simply point centers of force, as Roger Joseph Boscovich had suggested in the eighteenth century, or, in Faraday's terms, points from which lines of force spread into space. The extension of the atom may be identified with the extent of these lines, so that each atom would occupy all space and atoms would be mutually penetrable. Light might consist of vibrations in these lines, possibly obviating the need for an ethereal medium for its propagation; on the other hand, he suggested elsewhere that the lines might represent a condition of the ether, "for it is not at all unlikely that, if there be an ether, it should have other uses than simply the conveyance of radiations."

Faraday's geometric-intuitive representation was in particular rejected by the Continental electrodynamicists, and in 1846 Wilhelm Weber developed a theory of forces acting directly at a distance between charges that included the phenomena of electromagnetic induction. The validity of the lines of force concept was vindicated by the theoretical researches of William Thomson and particularly of Maxwell, who regarded his task to be putting Faraday's ideas into mathematical notation. It was with this motive that Maxwell developed his electromagnetic field theory, which, reinterpreted in quantum terms, remains the accepted theory of electromagnetic action and which was the prototype of all the field theories that dominate physics today. Maxwell's early representation of the lines of magnetic force as vortexes in the ether spinning about these lines as axes was in part suggested by Faraday's discovery of the magnetic rotation of the plane of polarization of light. The existence of this magnetic effect upon light had confirmed others in their speculations that light was some sort of propagated electromagnetic phenomenon, and the rotation of the plane of polarization suggested to Thomson and Maxwell that magnetism was in some way a rotatory effect, or, in contemporary terminology, the magnetic field is a pseudovector field.

Faraday was one among many who gave adumbrations of the generalized principle of conservation of energy, the clear expression of which is credited to Julius Mayer, James Joule, and Hermann von Helmholtz. His convictions regarding the interconvertibility of forces led him, from Hans Christian Ørsted's generation of magnetism by an electric current, to seek that generation of a current from magnetism that he found. In a lecture of 1834, Faraday spoke explicitly of this mutual convertibil-

ity, but he did not proceed further to specify how he conceived of the "forces" or "powers" that might be conserved or to discover quantitative relations. In connection with his investigations in 1840 of the source of the action of the voltaic pile he cited this principle against the contact theory, according to which the mere contact of two metals was the source of the current so that there would be a "creation of power" out of nothing, and in favor of the chemical theory, which found the source in the chemical actions occurring in the pile.

See also Boscovich, Roger Joseph; Chemistry, Philosophy of; Dynamism; Energy; Helmholtz, Hermann Ludwig von; Maxwell, James Clerk.

Bibliography

Faraday's work in electricity, including electrochemistry, is to be found in his *Experimental Researches in Electricity,* 3 vols. (London: R. and J. E. Taylor, 1839, 1844, 1855); his chemical work is collected in *Experimental Researches in Chemistry and Physics* (London: R. Taylor and W. Francis, 1859). Of particular interest are the speculative papers "Thoughts on Ray-Vibrations," in *Philosophical Magazine* 28 (1846): 345–350, reprinted in the *Experimental Researches in Electricity,* Vol. III, pp. 447–452, and "On the Physical Lines of Magnetic Force," in *Royal Institution Proceedings* June 11, 1852, reprinted in *Experimental Researches,* pp. 438–443. The standard biography is H. Bence-Jones, *The Life and Letters of Faraday,* 2 vols. (London: Longmans, Green, 1870). See also L. P. Williams, *Michael Faraday* (New York: Basic, 1965). A brief and lucid treatment of the development of his researches is given by John Tyndall in *Faraday as a Discoverer* (London: Longmans, Green, 1868).

Arthur E. Woodruff (1967)

FARIAS BRITO, RAIMUNDO DE

(1862–1917)

Raimundo de Farias Brito was the philosophic forerunner of Brazilian modernism. A profound sense of crisis underlies the work of Farias Brito. Individual existence is a precarious struggle against despair and death, and social order is threatened by moral disintegration and anarchy. Knowledge of man's role in his world is necessary to confront this crisis. Intellectual pursuit of the truth is the primary ethical obligation. Lacking certainty, man establishes a relative morality through conviction, involving both philosophy and religion. Metaphysics attempts to embrace the truth theoretically through formulation of worldviews; religion embraces the truth practically through acceptance and appropriation of a given worldview. This free acceptance of common convictions creates community, informing and giving thrust to the total culture of which it serves as a focus. The theoretical task is a permanent activity of the human spirit; the practical task, a permanent necessity. Convinced that modern philosophy was not adequate to its task, largely because of its predominantly skeptical mood grounded in phenomenalism, Farias Brito hoped to establish a new dogmatism capable of providing convictions that could give both courage for withstanding suffering and despair and the bases for reestablishing social order and direction. There was a transition in his thought from an early attempt to provide grounds for a naturalistic religion, inspired by German monism, to the articulation of his philosophy of spirit, influenced by French spiritualism. The naturalism is expressed in an incomplete series of volumes titled *A finalidade do mundo* (1895–1905). The new series, *Filosofia do espírito* (1912–1914), was initiated after Farias Brito had moved to Rio de Janeiro in 1909 to accept the chair in logic at Colégio Pedro II.

The spirit, "a live principle of action, capable of modifying … the order of nature; … of dominating itself; … of exercising dominion over things" is the "foundation of all reality and the basis of all experience." Psychological data are therefore indispensable to the metaphysician. Physiological psychology deals solely with the physical base of spirit; psychology proper ought to be concerned with subjective psychic phenomena. Its method is introspective, direct introspection supplemented by indirect introspection, a study of the manifestations of consciousness through which men achieve expression and communication. "Transcendent psychology" is the method employed for utilizing psychological data in metaphysics. From the felt fact of human existence, it is possible to rise to the level of transcendence, seeking knowledge concerning essence. The introspective operation of the individual consciousness reveals two facets of experience—consciousness itself and that which is presented to consciousness; both constitute existential reality. Through abstraction and analysis of each, pure consciousness is seen to have priority. Understanding man as essentially conscious spirit, the method of "transcendent psychology," leads to the postulation of divine spirit.

See also Introspection; Latin American Philosophy; Logic, History of; Naturalism.

Bibliography

Other works by Farias Brito are *Verdade como regra das acões* (Truth as the rule of action; Rio de Janeiro, 1905), *Base física do espírito* (Physical base of truth; Rio de Janeiro, 1912), and *Mundo interior* (Inner world; Rio de Janeiro, 1912).

For a study on Farias Brito, see Fred Gillette Sturm, "Farias Brito: Brasilian Philosopher of the Spirit," in *Revista interamericana de bibliografía* 13 (1963): 176–204.

Fred Gillette Sturm (1967)

FASCISM

"Fascism" was the ideology of the movement that, under the leadership of Benito Mussolini, seized power in Italy in 1922 and held power until the Allied invasion of Italy in World War II. Mussolini was a socialist until 1915, and fascism is a paradoxical but potent mixture of extreme socialist, or syndicalist, notions with a Hegelian or idealist theory of the state.

An attempt to provide fascism with a fully articulated theory was made by an Italian neo-Hegelian philosopher of some distinction, Giovanni Gentile, who was converted to fascism after Mussolini's coup. But as a former liberal and collaborator of Benedetto Croce, Gentile was opposed by the anti-intellectual wing of the Fascist Party, and his draft for a manifesto of fascist ideology was rewritten by Mussolini himself and published in 1932 in the *Enciclopedia italiana* as *La dottrina del fascismo.* However, no adequate conception of fascism could be derived from these theoretical sources alone; the actual behavior of the Italian fascists during their twenty years of power must also be taken into account.

The word *fascism* is often used, especially by left-wing writers, not only for the Italian doctrine but also for the similar, if more fanatic, national socialism of Adolf Hitler and for the altogether less coherent ideologies of Francisco Franco, Juan Perón, Ion Antonescu, and other such dictators. But however justifiable the wider and looser use of the word, the present article is confined to the system and ideology that called itself *Fascismo* and that flourished in Italy under Mussolini.

Gentile in his two books *Che cosa è il fascismo* (1925) and *Origini e dottrina del fascismo* (1929) stressed, as one might expect, the Hegelian elements in fascism. He argued that fascism was essentially idealistic and spiritual. Whereas liberalism, socialism, democracy, and the other progressive movements of the nineteenth century had asserted the rights of man, the selfish claims of the individual, fascism sought, instead, to uphold the moral integrity and higher collective purpose of the nation. And whereas liberalism saw the state simply as an institution created to protect men's rights, fascism looked on the state as an organic entity that embodied in itself all the noblest spiritual reality of the people as a whole. Fascism opposed the laissez-faire economics of capitalism and the bourgeois ethos that went with it. But fascism equally opposed socialism, which preached class war and trade unionism and thus served only to divide the nation. Fascism could tolerate no organized sectional groups that stood outside the state, for such groups pressed the supposed interests of some against the true interests of all. Hence, in place of trade unions, employers' federations, and similar organizations, fascism set up corporations that were designed to integrate the interests of particular trades, industries, professions, and the like into the wider harmony of the state.

Fascism, said Gentile, understood all the defects of bourgeois capitalism that had led to the rise of socialism, but fascism revolutionized society in such a way that the socialist critique was no longer relevant. For fascism replaced the old, competitive, hedonistic ethos of liberalism with an austere, stern, rigorous patriotic morality in which "the heroic values of service, sacrifice, indeed death itself were once more respected." Fascism did not deny liberty, but the liberty it upheld was not the right of each man to do what he pleased but "the liberty of a whole people freely accepting the rule of a state which they had interiorised, and made the guiding principle of all their conduct."

Fascism was proud of its comprehensive nature, of its totalitarian scope. For fascism, Gentile argued, was not just a method of government; it was a philosophy that permeated the whole will, thought, and feeling of the nation. "The authority of the state," Gentile wrote, "is absolute. It does not compromise, it does not bargain, it does not surrender any portion of its field to other moral or religious principles which may interfere with the individual conscience. But, on the other hand the state becomes a reality only in the consciousness of individuals." The state was "an idea made actual."

When Mussolini revised Gentile's draft for his *La dottrina del fascismo,* he retained most of the neo-Hegelian idealistic talk about the ideal nature of the state, but he had more to say about fascism's debts to the more extreme and fanatic elements of the nineteenth-century left wing. Mussolini named Georges Sorel, Charles Péguy, and Hubert Lagardelle as "sources of the river of Fascism." From these theorists, especially from Sorel, Mussolini derived the idea that "action is more important than thought"; by "action" he meant, as Sorel meant, vio-

lence. The extremists of the anarchist movement in the nineteenth century were obsessed by what they called *la propagande par le fait* (propaganda by deed); this "deed" tended to take the form of undiscriminating acts of revolutionary violence, such as throwing bombs into crowded cafés. The exhilaration of this policy soon blinded several of its champions to the end they were supposed to be pursuing—overthrowing the state—so that anarchism produced a movement of revolutionary disciplinarianism that Mussolini recognized as the source of his own inspiration.

Fascism was thus a movement that not only accepted, but also rejoiced in, violence. It had no patience with parliamentary or democratic methods of changing society. Indeed, Mussolini believed that the violent seizure of power, such as his own movement accomplished when it marched on Rome in 1922, was a necessary part of the moral rejuvenation of the nation; it was needed in order to create that "epic state of mind" (a phrase of Sorel's) that fascism prized so highly. Thus rejoicing in violence, fascism was, as Mussolini explained, hostile to all forms of pacifism, universalism, and disarmament. Fascism frankly acknowledged that "war alone keys up all human energies to their maximum tension, and sets a seal of nobility on those persons who have the courage to fight and die." The fascist state would have nothing to do with "universal embraces"; it "looked its neighbour proudly in the face, always armed, always vigilant, always ready to defend its integrity." Schemes such as that of the League of Nations were anathema to fascism.

With some reason Mussolini also claimed that fascism derived historically from the nationalistic movement of the nineteenth century. Nationalism, he insisted, owed nothing to the left. The German nation was not unified by liberals but by a man of iron, Otto von Bismarck. The nation of Italy, too, had been created by such men as Giuseppe Garibaldi, a man of revolutionary violence; the first great prophet of Italian unity was Niccolò Machiavelli, the archenemy of liberal, pacifist scruples. Mussolini had the highest regard for the author of *The Prince.* Machiavelli's desire to rekindle in modern Italy all the military virtues and military glory of ancient Rome was also Mussolini's ambition, but Mussolini's version of Machiavelli's dream was a much more vulgar one, and his achievements would have struck Machiavelli as tawdry, shabby, and corrupt.

Mussolini argued that it was the Italian state that had created the Italian nation. Indeed, it was the state, as the expression of a universal ethical will, which created the right to national existence and independence. Mussolini rejected the racism that was so central a feature of Nazi teaching in Germany. "The people," he wrote, "is not a race, but a people historically perpetuating itself; a multitude united by an idea." It must be recorded in favor of fascism that it never taught race hatred, and even when Mussolini entered the war on Hitler's side and introduced anti-Semitic legislation to please his ally, the Italian fascists were far from zealous in the enforcement of the laws against Jews.

Indeed, Mussolini's glorification of war and violence had never more than a limited success with the Italian people. Accustomed to rhetoric and appreciative of any kind of display, the Italians accepted the showier side of fascism more readily than the "austere, heroic way of life" that it demanded. Slow to conquer the backward Ethiopians in Mussolini's colonialist war against Abyssinia in 1935, the average Italian conscript soldier was even less eager to meet the Allied forces in World War II. Likewise, despite the cruelty of Mussolini's henchmen to his numerous political prisoners, there was never in Italy anything approaching the genocide that was faithfully enacted by Hitler's followers in Germany; even at its worst fascism never robbed the Italians of their humanity.

Mussolini earned a reputation, even among critical foreign observers, for the "efficiency" of his administration; he was popularly supposed abroad "to have made the Italian trains run on time." This achievement was largely mythical, for economic growth was minimal, but Mussolini was able, by forbidding strikes and subordinating industries to his state corporations, to prevent any of the more easily discernible manifestations of economic disorder. In any case his rule was never a mere personal dictatorship. He built up a powerful party with an elaborate hierarchy of command that served him much as the Soviet Communist Party served Joseph Stalin. Fascism was in a very real sense the dictatorship of a party, and the effectiveness of the party organization in a country by no means notable for good organization was one secret of fascism's twenty years of success.

See also Anarchism; Croce, Benedetto; Democracy; Gentile, Giovanni; Machiavelli, Niccolò; Marxist Philosophy; Nationalism; Political Philosophy, History of; Socialism; Sorel, Georges; Violence.

Bibliography

Ambrosini, G. *Il partito fascista e lo stato.* Rome: Instituto Nazionale Fascista di Cultura, 1934.

De Felice, Renzo. *Interpretations of Fascism.* Translated by Brenda Huff Everett. Cambridge, MA: Harvard University Press, 1977.
De Grand, Alexander J. *Fascist Italy and Nazi Germany: The "Fascist" Style of Rule.* London: Routledge, 1995.
Eatwell, Roger. *Fascism: A History.* New York: Allen Lane, 1996.
Finer, Herman. *Mussolini's Italy.* New York, 1935.
Gentile, Giovanni. *Genesi e struttura della società.* Florence: Sansoni, 1946.
Gentile, Giovanni. *Origins and Doctrine of Fascism: With Selections from Other Works.* Translated and edited by James Gregor. New Brunswick, NJ: Transaction, 2002.
Germino, Dante L. *The Italian Fascist Party in Power: A Study in Totalitarian Rule.* Minneapolis: University of Minnesota Press, 1959.
Gregor, James A. *The Ideology of Fascism: The Rationale of Totalitarianism.* New York: Free Press, 1969.
Griffin, Roger, ed. *Fascism.* Oxford: Oxford University Press, 1995.
Landini, Pietro. *La dottrina del fascismo.* Florence: Nuova Italia, 1936.
Larsen, Stein Ugelvik, Bernt Hagtvet, and Jan Petter Myklebust, eds. *Who Were the Fascists: Social Roots of European Fascism.* Bergen: Universitetsförlaget, 1980.
Lion, Aline. *The Pedigree of Fascism.* London: Sheed and Ward, 1927.
Mussolini, Benito. "*The Doctrine of Fascism.*" In *Social and Political Doctrines of Contemporary Europe,* edited by Michael Oakeshott, 2nd ed. New York: Macmillan, 1942.
Mussolini, Benito. *Scritti e discorsi,* 12 vols. Milan, 1934–1939.
Pitigliani, Fausto. *The Italian Corporative State.* London: King, 1933.
Rocco, Alfredo. *Political Doctrine of Fascism: Recent Legislation in Italy.* New York: Carnegie Endowment for International Peace, Division of Intercourse and Education, 1926.
Salvatorelli, Luigi. *Storia del fascismo.* Rome: Edizioni di Novissima, 1952.
Salvemini, Gaetano. *Under the Axe of Fascism.* New York: Viking Press, 1936.
Spencer, H. R. *The Government and Politics of Italy.* Yonkers, NY: World Book, 1932.

Maurice Cranston (1967)
Bibliography updated by Philip Reed (2005)

FATHERS OF THE CHURCH

See *Patristic Philosophy*

FECHNER, GUSTAV THEODOR

(1801–1887)

Gustav Theodor Fechner, the German philosopher, was the founder of psychophysics, and a pioneer in experimental psychology. He was born in Gross-Saerchen, Prussia, and studied medicine at the University of Leipzig, where he passed his examinations at the age of twenty-one. His interests, however, led him into physics, and by 1830 he had published more than forty papers in this field. He also wrote a number of poems and satirical works under the pseudonym of "Dr. Mises," which he also used for some of his later metaphysical speculations. A paper on the quantitative measurement of electrical currents (1831) led to his appointment as professor of physics at Leipzig. Fechner's incipient interest in psychology is shown in papers of 1838 and 1840 on the perception of complementary colors and on subjective afterimages. His experiments on afterimages, however, had tragic consequences. As a result of gazing at the sun he sustained an eye injury, and his subsequent blindness led to a serious emotional crisis. Fechner resigned his professorship in 1839 and virtually retired from the world.

A seemingly miraculous recovery, three years later, stimulated Fechner's interest in philosophy, particularly in regard to the question of the soul and the possibility of refuting materialistic metaphysics. In a work titled *Nanna oder das Seelenleben der Pflanzen* (Nanna, or the soul-life of plants; Leipzig, 1848) he defended the idea that even plants have a mental life. This book is indicative of the panpsychistic bent of Fechner's thought, which was the major cause of the direction taken by his further work.

PSYCHOPHYSICS

In 1848 Fechner returned to the University of Leipzig as professor of philosophy. His desire to substantiate empirically the metaphysical thesis that mind and matter are simply alternative ways of construing one and the same reality was the main motivation for his pioneering work in experimental psychology. His *Elemente der Psychophysik* (Leipzig, 1860) was intended to be an outline of an exact science of the functional relations between bodily and mental phenomena, with a view to showing that one and the same phenomenon could be characterized in two ways. Fechner divided his new science of psychophysics into two disciplines: inner psychophysics, which studies the relation between sensation and nerve excitation; and outer psychophysics, to which Fechner's own experimental work was devoted and which studies the relation between sensation and physical stimulus. Psychophysics became one of the dominant fields within experimental psychology.

Fechner's work on the relation between physical stimuli and sensations led to a mathematical formulation

that he called the law of intensity, which states that the intensity of a sensation increases as the logarithm of the stimulus, that is, by diminishing increments. When Fechner realized that his principle corresponded to the findings of E. H. Weber (1795–1878), he called it Weber's law, a name now reserved for the vaguer statement that a barely noticeable difference in stimulus has a constant ratio to the stimulus. Fechner's studies in psychophysics included a number of classical experiments on the perception of weight, visual brightness, and distance.

PANPSYCHISM

Fechner's psychological studies were meant to confirm his theory of panpsychism. He maintained that the whole universe is spiritual in character, the phenomenal world of physics being merely the external manifestation of this spiritual reality. That which *to itself* is psychical is *to others* physical. In his *Atomenlehre* he argued that physics requires us to regard atoms only as centers of force or energy, as Gottfried Wilhelm Leibniz had argued; it is not necessary to suppose them to be material or extended. These atoms are only the simplest elements in a spiritual hierarchy leading up to God. Each level of this hierarchy includes all those levels beneath it, so that God contains the totality of spirits. Consciousness is an essential feature of all that exists, but this assertion does not mean, as Leibniz had supposed, that every physical entity or phenomenon has its own soul. Only certain systems, namely, organic wholes, give evidence of possessing souls, and those bodies that do not are only the constituents of besouled bodies. The evidences of soul are the systematic coherence and conformity to law exhibited in the behavior of organic wholes. Fechner regarded Earth, "our mother," as such an organic besouled whole. The stars and the physical universe as a whole are also bodies of this kind. God is the soul of the universe; He is to the system of nature as that system is *to itself*.

To regard the whole material universe as inwardly alive and conscious is to take what Fechner called the "daylight view" (*Tagesansicht*). To regard it as inert matter, lacking in any teleological significance, is to take what he called the "night view" (*Nachtansicht*). Fechner ardently advocated the daylight view and hoped that it could be supported inductively by means of his psychophysical experiments. But he also argued for the daylight view on pragmatic grounds, offering the sort of arguments that William James later found highly congenial. Fechner urged that any hypothesis that cannot be positively proved but that does not contradict scientific findings be accepted if it makes us happy. The antimaterialistic daylight view is such a hypothesis. Fechner also defended his theory by means of analogical arguments. When certain qualities are found to be present in several types of objects, we are justified in assuming hypothetically that these objects share other, undetected qualities. Entities which exhibit the sort of order that our own bodies do may therefore be assumed to be alive and inwardly spiritual as we are.

IMMORTALITY

Fechner's argument for immortality is based on the observation that many individual experiences that are forgotten or unnoticed may later be recalled into consciousness. If the soul as a whole is treated on the analogy of its individual experiences, then, since these do not vanish utterly but often return in the form of memory, the soul itself may likewise continue to exist in God's memory. Mind and body are not parallel aspects of some third substance, as in Benedict (Baruch) de Spinoza; they are identical. The persistence of mind is therefore no more difficult to entertain than the persistence of the material universe itself, which is only the outward manifestation of an all-inclusive soul.

AESTHETICS

Between 1865 and 1876 Fechner turned his attention to aesthetics. He published a paper on the golden section, the supposedly ideal proportion, and several papers on the controversy over two Hans Holbein paintings of the Madonna. These two paintings, one in Dresden, the other in Darmstadt, were the subject of serious debate among art critics and aestheticians. Fechner hoped to settle the question of their relative excellence by means of a public preference poll when the paintings were exhibited together.

The desire to put aesthetics on an empirical, scientific footing and to bring philosophical speculation into some sort of accord with experimental science is shown further in Fechner's *Vorschule der Aesthetik* (Propaedentic to aesthetic; Leipzig, 1876), a work of considerable significance for the history of experimental aesthetics. In the preface to this work Fechner stated that previous aestheticians such as Friedrich Schelling and G. W. F. Hegel had theorized "downward" from universal principles to particulars. Fechner proposed to reverse this procedure, to build aesthetic theory "from below," on a foundation of empirical evidence. The word *beauty*, he maintained, denotes the approximate subject matter of aesthetics. It is a word applicable to everything that has the property of arousing pleasure directly and immediately. (Pleasure

aroused by thoughts of the consequences of an object is nonaesthetic.) Our experiences of aesthetic pleasure are simple, unanalyzable psychic atoms. The aim of an experimental aesthetics is to discover the objects that produce such atoms, that is, the causal laws connecting aesthetic experiences with the characteristics of outer objects.

EXPERIMENTAL METHODS. Fechner suggested three experimental methods for carrying out this program: the method of selection or choice, the method of production or construction, and the method of measuring common objects. The first of these methods is illustrated by Fechner's experiments with rectangles. Ten rectangles of varying dimensions but equal areas were spread at random on a table. The subject was asked to make a selection, ranking the rectangles in the order of his aesthetic pleasure and displeasure. A record was kept of his responses, with allowance being made for variation in hesitation of response. Fechner's results seemed to support the hypothesis that there exist certain ratios of length to width that possess specific aesthetic value. Most of the people tested tended to reject as unpleasant both the square or nearly square and the extremely elongated figures, with the largest number of favorable responses going to the rectangle whose proportions were 34:21. Fechner took this as empirical confirmation of the special aesthetic status of the golden section.

In the second of Fechner's methods, the subject was confronted, for example, with four vertical lines of various lengths and asked to place a dot over each line at the distance that seemed to him most aesthetically pleasing. The results were that the average distance was proportional to the length of the line. This experiment was referred to as the "inquiry into the letter 'i.'" Fechner's third experimental method involved measuring such objects as books, visiting cards, and so on, and here too he found the ratio of the golden section in a large percentage of cases.

LAWS OF PSYCHOLOGICAL AESTHETICS. A number of psychological laws formulated by Fechner are relevant to aesthetic experience. His principle of aesthetic threshold states that a stimulus must acquire a certain intensity before it can produce pleasure or pain. The effect will then increase gradually until it reaches a maximum point, whereupon it will decrease to the point of indifference. In the case of pleasure but not in that of pain, the effect may, after the maximum is reached, change to its opposite. Aesthetic reinforcement refers to the fact that several conditions of pleasure may, when combined, produce a total satisfaction greater than the sum of these conditions taken separately, for example, melody and harmony in music, meaning and rhythm in poetry. The principle of "uniform connection within the manifold" states that we prefer objects which are both unified and complex over objects which are homogeneous or excessively diverse. The principle of "absence of contradiction" claims that harmony and truth are aesthetically preferable to disagreement, contradiction, or error. Vagueness and ambiguity are aesthetically displeasing, as the principle of "clarity" announces. The recollection of an event portrayed in some aesthetic object may bring pleasure or displeasure, depending on whether the event reminds us of something pleasant or unpleasant: the principle of "aesthetic association." The principle of "minimum effort" states that pleasure is derived from the smallest possible expenditure of energy relative to a given end in view and not simply from the minimum expenditure of energy as such.

CONCLUSION

These "laws of the mind" illustrate the spirit of Fechner's philosophizing. He was one of the most versatile thinkers of the nineteenth century, laboring to reconcile an idealistic view of reality with the methodology of modern science and, in so doing, providing some of the groundwork for further developments in a number of areas of experimental psychology. His somewhat fantastic metaphysical speculations disclose a mind of poetic sensitivity, whose visions, however, he insisted on subjecting to scientific scrutiny.

See also Aesthetic Experience; Aesthetics, History of; German Philosophy; Hegel, Georg Wilhelm Friedrich; Immortality; James, William; Leibniz, Gottfried Wilhelm; Mind-Body Problem; Panpsychism; Psychology; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de.

Bibliography

ADDITIONAL WORKS BY FECHNER

Büchlein vom Leben nach dem Tod. Dresden, 1936. Translated by John Erskine as *Life after Death*. New York, 1943.

Zend-Avesta oder über die Dinge des Himmels und des Jenseits (Zend-Avesta, or concerning matters of heaven and the world to come). 3 vols. Leipzig, 1851.

Über die physikalische und philosophische Atomenlehre (On physical and philosophical atomic theory). Leipzig, 1850.

Über die Seelenfrage. Ein Gang durch die sichtbare Welt, um die Unsichtbare zu finden (On the question of the soul: a path through the visible world in order to find the invisible). Leipzig, 1861.

Die Tagesansicht gegenüber der Nachtansicht (The daylight view as opposed to the night view). Leipzig, 1879.

Die drei Motive und Grunde des Glaubens (The three motives and grounds of faith). Leipzig: Breitkopf and Härtel, 1863.

Elements of Psychophysics. Edited by Davis H. Howes and Edwin Garrigues. New York: Holt, Rinehart and Winston, 1966.

Zur experimentellen Aesthetik. Vorschule der Aesthetik. Hildesheim; New York: G. Olms, 1978, 1925.

Fix, Ulla, Irene Altmann, and Gustav Theodor Fechner. *Fechner und die Folgen ausserhalb der Naturwissenschaften: interdisziplinäres Kolloquium zum 200. Geburtstag gustav Theodor Fechners*. Tübingen: Niemeyer, 2003.

Tagebücher 1828 bis 1879. Leipzig: Verlag der Sächsischen Akademie der Wissenschaften zu Leipzig; Stuttgart: In Kommission bei F. Steiner, 2004.

Note: A bibliography of 175 of Fechner's writings can be found in *Elemente der Psychophysik*. 2nd ed. (Leipzig, 1889. Vol. 1).

WORKS ON FECHNER

In English

Boring, E. G. *History of Experimental Psychology*. New York: Century, 1929 and 1950. Ch. 14.

James, William. *A Pluralistic Universe*. New York: Longman, 1909. Lecture 4.

James, William. *Principles of Psychology*. New York: Holt, 1890. Vol. 1, pp. 553–549.

Külpe, O. *Die Philosophie der Gegenwart in Deutschland*. Leipzig, 1902. Translated by M. L. Patrick and G. T. W. Patrick as *The Philosophy of the Present in Germany*. London: G. Allen, 1913. Pp. 147–160.

Murphy, G. *Historical Introduction to Modern Psychology*. New York: Harcourt Brace, 1949. Pp. 82–94.

Perry, R. B. *Philosophy of the Recent Past*. New York: Scribners, 1926. Pp. 82–86.

In German

Hermann, J. *Gustav Theodor Fechner*. Munich, 1926.

Kuntze, J. E. *Gustav Theodor Fechner*. Leipzig, 1882.

Lasswitz, K. *Gustav Theodor Fechner*. Stuttgart: F. Frommann, 1896.

Meyer, F. A. E. *Philosophische Metaphysik und Christliche Glaube bei Gustav Theodor Fechner*. Goettingen, 1937.

Arnulf Zweig (1967)
Bibliography updated by Michael Farmer (2005)

FËDOROV, NIKOLAI FËDOROVICH

(1829–1903)

Nikolai Fëdorovich Fëdorov was a Russian religious philosopher. From 1854 to 1868 he taught history and geography at district schools in Russia. From 1869 to 1872 he worked at the Chertkovskaia Library in Moscow, and from 1874 to 1898 he worked at the libraries of the Moscow Public and Rumiantsev Museums. For a quarter of a century he defined the spiritual atmosphere of this latter library, infusing it, in the words of his contemporaries, with the traditions of the "philosophical school." Many talented men of Russian science and culture used to gather in the catalogue room of the library where Fëdorov served to converse with the "Moscow Socrates." In the 1880s and 1890s Fëdorov met with Vladimir Sergeevich Solov'ëv, who called Fëdorov's teaching "the first progress the human spirit has made on the way of Christ." In that time period Fëdorov also carried on his religio-philosophical dialogue and debate with Lev (Leo) Nikolaevich Tolstoy.

Starting in 1851 Fëdorov expounded his ideas first orally, and then, starting in the second half of the 1870s, in large works and articles. After Fëdorov's death, his disciples V. A. Kozhevnikov and N. P. Peterson prepared for publication a three-volume collection of the philosopher's works under the title *Filosoviia obshchego dela* (The Philosophy of the common cause; the first two volumes were published in 1906 and 1913, respectively; the third volume remained unpublished).

In the evolutionary process Fëdorov discerned a tendency to the birth of consciousness and reason, which, beginning with man, were called to become the instruments, no longer of an unconscious, but of a conscious and morally and spiritually oriented perfecting of the world. "In us, nature begins not only to be conscious of itself but also to control itself." Man is both the crown of evolution and its agent; the labor of the cosmicization of being lies on his shoulders. In opposition to the existing parasitical and exploitative relation of man to the natural environment, which is leading civilization to the brink of catastrophe ("A civilization that exploits but does not restore can have no other result than its own end"), Fëdorov advanced the idea of the *regulation of nature*, which unfolds in a series of tasks. This series comprises the prevention of natural disasters (earthquakes, floods, droughts, etc.), the regulation of climate, the control of cosmic processes, labor directed at the conquest of death, and—as the climax of this regulation of nature, the focus of all of its efforts—the return to a new transfigured life of all those who have departed into nonbeing, infinite creative work in a renewed Universe.

Fëdorov gave his teaching both a natural-scientific and a religious foundation. Basing his thought on the patristic tradition (St. Basil the Great, St. Gregory the Theologian, and Gregory of Nyssa), he developed an actively Christian anthropology: God, in creating man in His image and likeness, acts in the world first and foremost through man, and through him He will realize the

central ontological promises of the Christian faith, such as the raising of the dead, the transfiguration of their nature, and the entry into the immortal, creative eon of being, the Kingdom of Heaven. He propounded the idea of divine humanity, the collaboration of the divine and human energies in the work of salvation, and argued that the prophecies of the Revelation have only a conditional significance. Will the end of history be catastrophic, leading to the Last Judgment with the consequent division of humankind into a handful of the saved and a vast multitude of the eternally damned? Or will it be radiant, where all will be saved (the apocatastasis)? This depends on people themselves, on whether the world's movement will continue on its false, antidivine vector or whether it will redirect itself to the ways of God.

Fëdorov also gave the idea of the regulation of nature a religious interpretation. Based on the sense of the profound moral responsibility of man for the fate of the entire earth, of the entire cosmos, and of the entire creation, regulation represents the fulfillment of the biblical commandment that man be lord of the earth. "Restoration of the world to that splendid beauty of incorruption that it possessed before the Fall"—that is how the philosopher of the universal task defines God's assignment to the "sons of men."

A successful outcome of history, which becomes a "work of salvation," presupposes, according to Fëdorov, the necessity of a new fundamental choice that is associated with the imperative of the evolutionary ascent of humanity. He exposes the defects of a one-sided technological development that improves machines and mechanisms but that leaves man's nature untouched and vulnerable, entirely at the mercy of the vagaries of the external environment. As an alternative, he advances the idea of organic progress that is oriented toward the transformation of the physical substance of conscious beings. As a result of this transformation, man himself, without the aid of technology, will be able to fly, to see far and deep, to build his tissues from elementary materials of the environment like plants under the effect of sunlight (here, Fëdorov anticipates what V. I. Vernadskii would later call *autotrophism*), and to create necessary organs for himself or change his existing organs as a function of the medium of his habitation and action (the notion of "fullness of organs"). According to Fëdorov the body, the receptacle of the soul, must be made wholly subordinate to the consciousness; the body must be regulated and spiritualized. Spirit must achieve total power over matter, leading to a state where the forces of decay, corruption, and death are limited and finally expelled from being.

Fëdorov envisaged a radical change in philosophy. This change would consist in the rejection of abstract thought and passive contemplation, in a transition toward the definition of the values of the necessary order of things, toward the development of a plan for humanity's transformative activity. He proclaimed the inseparability of ontology and deontology ("truth is only the path to the good") and the necessity of a projective thought (the project connects the ideal and reality and seeks ways toward a practical realization of the supreme idea). He advanced the principle of the integrity and universality of knowledge ("all people must be knowers and all things must be an object of knowledge"), and he spoke of the transformation of gnoseology into *gnoseo-urgy*. He called his system *supramoralism*, establishing the foundations of a "mature," "filial" morality ("we are all brothers according to love for the fathers").

Here, he did not limit the laws of ethics to the sphere of human relationships, indicating the dependence of the moral principle in man and in society on the material and natural order of things. Unkindred and unbrotherly attitudes, he emphasized, are rooted in the depths of postlapsarian, mortal being, which is based on the law of the succession of generations, with mutual devouring, expulsion, and struggle. And therefore only one thing can guarantee the attainment of "universal kinship": the conquest of the forces of death in the external world (by means of natural-cosmic regulation) and in man himself (by means of psycho-physiological regulation). Convinced of the incompleteness of altruistic morality (where the self-sacrifice of some presupposes the eternal egotism of others), Fëdorov offers the formula, "[N]ot for oneself and not for others, but with all and for all." He resolved the antinomy of individualism and collectivism through the principle of sobornost (communalism or all-togetherness), affirming the latter as the foundation of the perfect social organization (society "according to the type of the Trinity").

Fëdorov also interpreted the meaning of culture in the light of the idea of immortality and the raising of the dead. He viewed culture as an attempt at an "imaginary raising from the dead," as an impulse to preserve the memory of that which had lived in the past. He put a high value on museums and libraries as centers of the universal human memory. He dreamed of a radical expansion of the activity of museums and libraries, of their transformation into centers of collection, investigation, education, and training, around which associations of scholars would be grouped, associations of "specialists in all domains of human knowledge." By becoming an instru-

ment of the universal task, the museum, according to Fëdorov, was to animate knowledge with a heartfelt feeling of kinship, with a spirit of love for fathers and ancestors, thus serving the restoration of the brotherly connection of people.

Fëdorov's philosophy is at the origin of the Russian religio-philosophical renaissance and helps to define the fundamental themes of the latter. His philosophy is the source of the actively evolutionary noospheric thought of the twentieth century (N. A. Umov, V. I. Vernadskii, and A. L. Chizhevskii). Various talented representatives of Russian literature were influenced, at different times and to different degrees, by *The Philosophy of the Common Task*: Fëdor Mikhailovich Dostoevsky and Tolstoy, Valerii Briusov and Vladimir Maiakovskii, Nikolai Kliuev and Velimir Khlebnikov, Mikhail Prishvin and Maksim Gorky, Andrei Platonov and Boris Pasternak. Fëdorov's theurgic aesthetics (the transition from an "art of imitations" to the creative work of life to the liturgical synthesis of the arts) exerted an influence on the philosophical-aesthetic quests at the end of the nineteenth century and the beginning of the twentieth century (Solov'ëv, Belyi, Viacheslav Ivanovich Ivanov, V. Chekrygin, P. Filonov, and others).

See also Aesthetics, History of; Consciousness; Darwinism; Dostoevsky, Fyodor Mikhailovich; Gregory of Nyssa; Patristic Philosophy; Reason; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich; Tolstoy, Lev Nikolaevich.

Bibliography

WORKS BY FËDOROV

Sochineniia (Works). Moscow, 1982.

Sobranie sochinenii (Collected works). Vols. 1–4. Moscow, 1995–1999.

What Was Man Created For? The Philosophy of the Common Task: Selected Works, edited by E. Kontaissoff and M. Minto. Lausanne, Switzerland: Henyglen/L'Age d'Homme, 1990.

WORKS ON FËDOROV

Kozhevnikov, V. A. *Nikolai Fedorovich Fedorov*. Part 1. Moscow, 1908.

Gorsky, A. K. *Nikolai Fedorovich Fedorov i sovremennost'* (Nikolai Fëdorovich Fëdorov and the present time). Issues 1–4. Harbin, 1928–1933.

Hagemeister, Michael. *Nikolaj Fedorov: Studien zu Leben, Werk, und Wirkung*. Munich: Sagner, 1989.

Koehler, L. *N. F. Fedorov: The Philosophy of Action*. Pittsburgh, PA, 1979.

Lukashevich, Stephen. *N. F. Fedorov (1828–1903): A Study in Russian Eupsychian and Utopian Thought*. Newark: University of Delaware Press, 1977.

Semenova, S. G. *Filosof budushchego veka Nikolai Fedorov* (Philosopher of the future age, Nikolai Fëdorov). Moscow, 2004.

Semenova, S. G. *Nikolai Fedorov: Tvorchestvo zhizni* (Nikolai Fëdorov: The creativity of life). Moscow, 1990.

Young, George M., Jr. *Nikolai F. Fedorov: An Introduction*. Belmont, MA: Nordland, 1979.

S. G. Semenova (2005)
Translated by Boris Jakim (2005)

FEELING

See *Emotion*

FEINBERG, JOEL

(1926–2004)

Joel Feinberg was a noted moral, social, political, and legal philosopher. He was born in Detroit, Michigan. After his military service in World War II, Feinberg earned bachelor's, master's, and doctoral degrees at the University of Michigan (Ann Arbor). His doctoral dissertation was titled "Naturalism and Liberalism in the Philosophy of Ralph Barton Perry" (1957).

It was not until 1960, when Feinberg was thirty-three years old, that he published his first philosophical essay. During the next four decades, while Feinberg taught at Brown, Princeton, UCLA, Rockefeller, and Arizona, his scholarly output was prodigious. Within a few years of his arrival at the University of Arizona, the philosophy department there attracted several other prominent philosophers and become one of the most highly regarded programs in the United States. Feinberg was honored by his philosophical peers in 1981 by being elected president of the Pacific Division of the American Philosophical Association. In 1988, he was one of the first individuals to be designated Regents Professor at the University of Arizona.

Liberalism was Feinberg's focus throughout his long and distinguished career. During the 1980s, he wrote his *magnum opus*, the four-volume, 1,397-page *Moral Limits of the Criminal Law*. Feinberg's aim in this work (which he called his "tetralogy") was "to make the best possible case for liberalism" with respect to the moral limits of the criminal law (*Harm to Others*, p. 15). He thought of himself as "vindicat[ing] the traditional liberalism derived from [John Stuart] Mill's *On Liberty* [1859]" (ibid.). Although Feinberg had no legal credentials (other than having been a Liberal Arts Fellow at Harvard Law School

in 1963–1964), he has already influenced American law. At least one state supreme court has cited him as a persuasive authority. (See *Armstrong v. Montana*, 296 Mont. 361, 989 P.2d 364 [1999] [holding that a Montana statute prohibiting physician assistants from performing abortions violated the privacy, equal-protection, and bill-of-attainder provisions of the Montana constitution].)

Feinberg begins his tetralogy with what he calls a presumption in favor of (individual) liberty. This presumption means that "[l]iberty should be the norm; coercion always needs some special justification" (*Harm to Others*, p. 9). He then sketches a number of "liberty-limiting principles," each of which states a reason—but not a necessary or a sufficient condition—for coercing individuals. The question he sets for himself is which of these principles, if any, are valid. Here, for example, is the harm principle:

> It is always a good reason in support of penal legislation that it would probably be effective in preventing (eliminating, reducing) harm to persons other than the actor (the one prohibited from acting) and there is probably no other means that is equally effective at no greater cost to other values. (*Harm to Others*, p. 26 [italics in original; footnote omitted])

Feinberg endorses two liberty-limiting principles: the harm principle and the offense principle. He rejects two others: legal paternalism and legal moralism. Volume one of his tetralogy, *Harm to Others*, elaborates and defends the harm principle. Volume two, *Offense to Others*, elaborates and defends the offense principle. Volume three, *Harm to Self*, elaborates and rejects legal paternalism. Volume four, *Harmless Wrongdoing*, elaborates and rejects legal moralism.

Legislators who are guided by Feinberg's liberalism, with its normative commitments to individual liberty and personal autonomy, would prohibit and punish only harmful or seriously offensive conduct (but not necessarily *all* of such conduct). An example of seriously offensive conduct would be a pornographic billboard that individuals cannot reasonably avoid. Feinbergian (ideal) legislators would not punish conduct solely on the ground that it is harmful to the actor. That is legal paternalism, which is an affront to personal autonomy. Nor would they punish conduct solely on the ground that it is immoral. That is legal moralism. It is important to understand that Feinberg's rejection of legal moralism does not rest on moral skepticism, nihilism, relativism, or subjectivism. One can be a moral objectivist—a believer in objective moral values—and still hold that it is improper for legislators to enforce a single "true" morality. Feinberg's aim is practical: to "guide the legislator by locating the moral constraints that limit his options" (*Harm to Others*, p. 4). It is "a quest not for useful policies but for valid principles" (*Harm to Others*, p. 4).

The four volumes together make a powerful case for "the liberal position" on the moral limits of the criminal law. Feinberg does not argue for liberalism directly by appealing to "moral primitives" or "self-evident truths" (*Harm to Others*, p. 17). Instead, he adopts the *argumentum ad hominem* technique. This type of argument consists in appealing to values, beliefs, and convictions his readers are presumed to have or to judgments they are presumed to make. Feinberg's objective is to persuade these readers that the liberal position on the moral limits of the criminal law systematizes their values, beliefs, convictions, and judgments better than any alternative. It is a search for coherence, not foundations. In effect, he is trying to show his readers that they are—already, unwittingly—liberals.

Among the areas in applied or practical ethics to which Feinberg made important contributions are abortion and animal rights. In his influential 1979 essay "Abortion," he sought to structure the debate over the morality and legality of abortion by (as he later put it) "locating crucial but implicit presuppositions, centrally affected interests, critical distinctions, and so on" (*Freedom and Fulfillment*, p. viii). In an essay published in 1971, four years before Peter Singer's celebrated *Animal Liberation* appeared, Feinberg argued that animals are "among the sorts of beings of whom rights can meaningfully be predicated and denied" (*Rights, Justice, and the Bounds of Liberty*, p. 166). Feinberg was not arguing that animals do in fact have rights. He was arguing that it is not incoherent—as many people had thought—to ascribe rights to them. This was an important step in what became a powerful case for including nonhuman animals in the moral community. By clarifying the concept of a right, Feinberg was able to show that certain denials of rights were ill-founded. To Feinberg, "[c]onceptual clarification is the most distinctively philosophical of enterprises" (*Harm to Others*, p. 17). Clear thought leads to or is an indispensable part of sound moral judgment.

Feinberg's work, taken as a whole, is best characterized as social philosophy—interpreted broadly to include moral, political, and legal philosophy. In addition to the moral limits of the criminal law, he was interested in and made original contributions to the understanding of responsibility, punishment, desert, mental illness, rights,

justice, liberty, civil disobedience, freedom of expression, paternalism, autonomy, and fulfillment. His textbook *Reason and Responsibility: Readings in Some Basic Problems of Philosophy*, which appeared in 1965 (the twelfth edition was published in 2005), is among the best-selling philosophy textbooks of all time. Feinberg proved that original, important philosophical work is compatible with textbook writing. He was ever the teacher. Late in life, he published a delightful little book entitled *Doing Philosophy: A Guide to the Writing of Philosophy Papers* (1997).

It is fitting that Feinberg wrote a book on writing, for his writing style is justly famous and much emulated. His writing is clear, simple, and penetrating—at times even beautiful—despite the complexity of the issues and concepts with which he grappled. Several generations of philosophers have admired and learned from Feinberg, both substantively and stylistically. Many of his students went on to prominent careers of their own, in law or philosophy or both. In 1994, one of his most accomplished students, Jules Coleman, and a former colleague, Allen Buchanan, published an aptly titled collection of critical essays devoted to Feinberg's work: *In Harm's Way: Essays in Honor of Joel Feinberg*.

Bibliography

WORKS BY FEINBERG

Doing and Deserving: Essays in the Theory of Responsibility. Princeton, NJ: Princeton University Press, 1970. This volume collects essays published between 1960 and 1969 (inclusive). The essays concern such concepts as act, cause, harm, punishment, desert, and blame. Feinberg thought of these essays as "straddling ethics, philosophy of mind, and philosophy of law."

Social Philosophy. Foundations of Philosophy Series, edited by Elizabeth Beardsley and Monroe Beardsley. Englewood Cliffs, NJ: Prentice-Hall, 1973. This highly regarded monograph concerns itself with "philosophical questions about social relations." Among the concepts investigated are freedom, coercion, legal rights, human rights, and social justice. This book is where Feinberg introduced the concept of a liberty-limiting principle that figured so prominently in his later work.

Rights, Justice, and the Bounds of Liberty: Essays in Social Philosophy. Princeton, NJ: Princeton University Press, 1980. This volume collects essays published between 1964 and 1978 (inclusive). The essays concern such concepts as liberty, harm, offense, legal paternalism, and rights. Feinberg thought of these essays as dealing with "hard cases for the application of the concept of a right."

The Moral Limits of the Criminal Law, Vol. 1: *Harm to Others*. New York: Oxford University Press, 1984. This is the first volume of Feinberg's "account of the moral constraints on legislative action." Feinberg discusses the concept of harm; its relation to such concepts as interests, wants, hurts, offenses, rights, and consent; hard cases for application of the concept of harm; and various problems involved in assessing, comparing, and imputing harms.

The Moral Limits of the Criminal Law, Vol. 2: *Offense to Others*. New York: Oxford University Press, 1985. Feinberg discusses the concept of offense (as a mental state distinct from harm) and some of its applications, including pornography, obscenity, and "dirty words."

The Moral Limits of the Criminal Law, Vol. 3: *Harm to Self*. New York: Oxford University Press, 1986. Feinberg discusses legal paternalism, personal autonomy, and the concept of voluntariness.

The Moral Limits of the Criminal Law, Vol. 4: *Harmless Wrongdoing*. New York: Oxford University Press, 1988. Feinberg discusses legal moralism: the view that "[i]t can be morally legitimate to prohibit conduct on the ground that it is inherently immoral, even though it causes neither harm nor offense to the actor or to others."

Freedom and Fulfillment: Philosophical Essays. Princeton, NJ: Princeton University Press, 1992. This volume collects essays published between 1975 and 1991 (inclusive). The essays concern such concepts as wrongful life, abortion, freedom of expression, bad samaritanism, moral rights, and absurd self-fulfillment. Despite the title, Feinberg thought of these essays as dealing with "problems about rights."

Doing Philosophy: A Guide to the Writing of Philosophy Papers. Belmont, CA: Wadsworth, 1997.

Problems at the Roots of Law: Essays in Legal and Political Theory. Oxford: Oxford University Press, 2003. This volume collects essays published between 1992 and 2003 (inclusive). The essays concern such concepts as natural law, moral rights, entrapment, criminal attempts, government subsidies for the arts, and evil. As the title implies, Feinberg thought of these essays as dealing with "basic questions" in the philosophy of law.

Works on Feinberg

Coleman, Jules L., and Allen Buchanan, eds. *In Harm's Way: Essays in Honor of Joel Feinberg*. Cambridge, U.K.: Cambridge University Press, 1994. This volume contains critical essays by Allen Buchanan, Shelly Kagan, Richard J. Arneson, David Lyons, David A. J. Richards, Thomas Morawetz, Jules L. Coleman, Jean Hampton, John Martin Fischer and Mark Ravizza, Jeffrie G. Murphy, Joan McGregor, Robert F. Schopp, Sanford H. Kadish, Holly M. Smith, and Hyman Gross.

Keith Burgess-Jackson (2005)

FEMINISM AND CONTINENTAL PHILOSOPHY

Continental philosophy has been a significant force in the development of contemporary feminist thought. Many feminists have turned to the work of continental philosophers because the topics explored by these philosophers are germane to the kinds of questions feminists pursue.

Since Hegel continental philosophy has been concerned with questions of ethics, metaphysics, consciousness, and experience. Continental philosophy has occupied a prominent position in contemporary feminist philosophy because it examines these issues so central to feminist concerns.

EXISTENTIALISM AND PHENOMENOLOGY

The publication of Simone de Beauvoir's *The Second Sex* in 1949 marks the beginning of the contemporary feminist movement. De Beauvoir's work is rooted in two prominent continental philosophies, existentialism and phenomenology. The theme of her book is summarized in her famous statement that one is not born a woman, one becomes one. This statement and the analyses ensuing from it reveal the influence of both existentialism and phenomenology at the very beginning of the contemporary feminist movement. Existentialists such as Jean-Paul Sartre emphasized the ontological complexity of our existence as consciousness in bodies. Existential philosophers explored the themes of freedom and oppression, objectification, and the social construction of consciousness. Feminists such as de Beauvoir adapted these themes to the analysis of the situation of women in society. Existential feminists have described female bodily experience as socially constructed. They have analyzed the structures of society that perpetuate patriarchy and the oppression of women.

The influence of phenomenological thought has also been decisive. Husserl's phenomenological philosophy was rooted in an examination of how phenomena appear to consciousness. The phenomenological approach of Husserl, Heidegger, and Merleau-Ponty grounded philosophy in lived experience. For feminists this approach has provided a means of challenging a conception of objectivity that many theorists believe grounds Western philosophical thought and that many feminist philosophers identify as masculinist. It has fostered the development of feminist theory that arises from the distinctive lived experiences of women. Feminist phenomenologists explore how living in a female body in modern society produces a consciousness unique to women. They emphasize human subjectivity and the role of language in creating social reality. Their goal is to develop a feminist consciousness of oppression (Bartky 1990). Exploring the boundaries of that consciousness is the hallmark of feminist phenomenology.

In contemporary feminist thought the approaches of phenomenology and existentialism have merged in feminist analyses of the body. Feminist philosophers such as Iris Marion Young (1990) examine the phenomenon of the female body in patriarchal society. Young explores aspects of women's lived experience—pregnancy, for example—that are unique to women. Her point is that women's bodily experience is different from that of men and that this difference effects women's consciousness under patriarchy. Young argues that existential phenomenology exhibits an adherence to the subject/object dualism. Young's goal is to replace this dualism with an understanding that erases the difference between the inner and the outer. She wants to develop a position that corrects this error without abandoning the advantages of existential phenomenology.

MARXISM

In the 1960s many feminists were attracted to Marxist philosophy as a vehicle for feminist theory. There were several reasons for the convergence of feminism and Marxism. First, Marxism was the oppositional philosophy of the time; to be opposed to the status quo in this time period almost necessarily entailed a Marxist stance. Second, Marxism, like feminism, was concerned with oppression. Although Marx was not himself concerned with the oppression of women, his theory of the oppression of the proletariat seemed to many feminists to have much to contribute to the attempt to develop a theory of the oppression of women.

The aspect of Marx's theory that became most influential in feminist thought was his theory of the standpoint. Marx argues that the standpoint of the proletariat in capitalism affords it a privileged understanding of its social structure; in his view the proletariat's position as the oppressed class allows it to see the true reality of capitalism. As a social determinist, Marx asserts that knowledge is governed by the subject's historical/material position. Yet he also claims that the knowledge produced by those in the oppressed class is the only true knowledge; the knowledge of other classes, in contrast, is "partial and perverse."

Feminists such as Nancy Hartsock (1983) and Dorothy Smith (1987) have used Marx's theory of the standpoint to analyze the position of women in society. They argue, first, that women, like the proletariat, are an oppressed class. Their thesis is that the bourgeoisie's oppression of the proletariat parallels men's oppression of women. Patriarchy, like captialism, is a system of oppression in which the dominant class, men, hold the oppressed class, women, in subjection. Second, feminist-standpoint theorists argue that the activity of women in

society—child-rearing, child-bearing, and housework—creates a particular reality for women. Like Marx, they argue that the social actor's activity creates her knowledge. Finally, they contend that the knowledge produced by the standpoint of women is truer than that produced by men. Following Marx, they argue that the knowledge of the oppressed class of women reveals the truth of patriarchy, whereas that of the ruling class of men is partial and perverse.

Feminist standpoint theory has been a major component of contemporary feminist thought. Hartsock's *Money, Sex, and Power* (1983) advanced the thesis that the distinctive activity of women in society provides them with a privileged access to reality. Her analysis of how the feminist standpoint is produced through the practices distinctive of women in society became the basis for extensive analyses of that standpoint. Dorothy Smith's analysis of the "lifeworld" of women extends the concept of the standpoint into an analysis of the everyday life of women. Combining standpoint theory with a phenomenological approach, Smith argues for an analysis of the everyday life of women as constitutive of their social reality.

But feminist standpoint theory has also raised questions for feminist thought. As feminists moved from a consideration of the difference between men and women to the differences among women, the concept of the feminist standpoint became problematic. Feminists questioned how one feminist standpoint could account for the variety of women's experiences. Feminists also began to question the epistemology of the standpoint. If, as Marx claims, all knowledge is perspectival, then how can one perspective be "truer" than another? Standpoint theorists have difficulty answering either of these questions.

POSTMODERNISM AND POSTSTRUCTURALISM

Since the 1990s one of the principal influences in feminist thought has come from the predominantly French philosophies of postmodernism and poststructuralism. Inspired by the work of Friedrich Nietzsche, postmodern and poststructuralist philosophers have questioned not just aspects of Western thought but its very foundation. Rejecting the Cartesian subject and the pursuit of universal knowledge, these thinkers have fundamentally altered the project of philosophy. Many feminists have been attracted to these theories because they provide a radically new way to understand the feminine and its place in Western philosophy. Postmodernism and poststructualism, by redefining truth as plural rather than universal, provide the possibility of overcoming the inferiority of women that has pervaded Western thought.

The widely acknowledged inspiration for postmodern thought is the work of Nietzsche. The object of Nietzsche's attack is the tradition of Western thought beginning with the Greeks. Two aspects of his thought have been particularly relevant to feminism. First, truth, for Nietzsche, is relational and perspectival. It is a "mobile army of metaphors" that is harnessed for use by those in power. Second, Nietzsche questions the centerpiece of modern Western philosophy, the subject. By undermining the subject/object dualism that provides the grounding for the subject, Nietzsche calls into question the autonomy of the subject and its place in the constitution of knowledge.

The radical quality of Nietzsche's thought has resonated with many feminists. For those feminist philosophers claiming that the "man of reason" informing Western thought has excluded women from the pursuit of truth, Nietzsche's approach provided a mean of further articulating this claim and of exploring an alternative.

Two theorists whose work is rooted in that of Nietzsche have played a significant role in contemporary feminist philosophy. The work of Michel Foucault, although controversial, has had a significant impact on contemporary feminism. Like Nietzsche, Foucault takes on the two pillars of Western thought: truth and the subject. For Foucault truth is constituted through discourses; it is specific to the discourse in which it operates. It follows that the universal truth of the Western tradition is a fiction created, itself, by a particular discourse. For Foucault standards for what constitutes truth are not universal but, rather, internal to particular discourses. The most radical element of Foucault's thought, however, is his declaration of "the death of man." Foucault argues that the autonomous, constituting subject of modern philosophy (the Cartesian subject) is a creation of a particular discourse at a particular time and, most significantly, is now in eclipse. For Foucault discourses create specific kinds of subjects; there is no universal subject but only the subjects constituted by particular discourses.

Feminists have found Foucault's work extremely useful. His theory of the death of man has obvious feminist implications even if Foucault did not explore them. "Man"— the rational, autonomous, self-constituting subject—has been a problem for many feminists. Exposing this concept as the product of historically located discourses and thus vulnerable to change eliminates these problems. Feminists have also used Foucault's work to explicate how the subject "woman" is created by the dis-

courses of patriarchal society. In the highly influential *Gender Trouble* (1990), Judith Butler uses a postmodern approach to explicate how the identity "woman" is constituted. Butler argues that this identity is a fiction created by the actions of women who perform that identity. She advocates a feminist politics that eschews the identity "woman" and instead creates "gender trouble," the destabilization of the gender structures of society. Feminists have also used Foucault's thought to challenge the "truth" about woman, enshrined in Western philosophy and science. Using a Foucaultian approach, feminists have explicated how truths are established and sedimented into their discursive foundation.

The work of Jacques Derrida has also provided the basis for feminist philosophical investigations. Derrida's "deconstructive" approach, like that of Nietzsche and Foucault, constitutes a fundamental critique of Western rationalism. Derrida's strategy of deconstruction focuses on language and its construction of a monolithic reign of truth. Derrida attacks what he calls the "metaphysics of presence," the presuppositions informing the tradition of Western philosophy. His goal is to examine the elements of Western rationalism and expose them as an elaborate construction rather than as absolute truth. He does so by "deconstructing" its basic concepts—that is, examining the presuppositions that inform those concepts and the consequences that flow from them.

Feminist philosophers such as Luce Irigaray, Julia Kristeva, and Helene Cixous have employed a Derridean perspective to deconstruct the dualisms that found Western philosophy. Questioning the masculine definitions of rationality and truth on which Western thought is grounded, these feminist philosophers have argued for a distinctively feminine way of writing as a counterweight to the norms of male-dominated discourse. If, as Derrida claims, we are constituted by language, then we need another language to resist this constitution. The goal of these philosophers is to redefine "woman" and the feminine in ways that are not structured by Western dualisms.

Postmodern and poststructuralist philosophy have provided a rich addition to feminist philosophy. They have allowed feminists to examine the relationship between language and the status of women in radically different ways. But postmodern feminism has also been strongly criticized within the feminist community. Its critics have argued that postmodernism, by rejecting absolute truth, is a form of relativism, even nihilism. Without some conception of truth, these critics claim, feminists cannot proclaim the truth of the oppression of women. They further argue that postmodernism deprives feminism of a political stance, a necessary component of feminism. The defenders of postmodern feminism counter that their outlook does not preclude politics but, rather, offers a different understanding of the political. They point to the revolutionary force implicit in Derrida's deconstruction and the "local" rather than universal resistance advocated by Foucault. But the controversies over postmodernism and feminism show no signs of abating.

CRITICAL THEORY AND HERMENEUTICS

Although they do not represent as pervasive an influence as postmodernism, both critical theory and hermeneutics have also found a following among feminist philosophers. The work of Jurgen Habermas has influenced the writings of both Nancy Fraser and Seyla Benhabib. These theorists find Habermas's philosophy attractive because, although it is critical of Enlightenment rationalism, it nevertheless provides a normative basis for an alternative conception. Partly inspired by Marxism, Habermas's approach entails both a critique of social norms and an alternative vision of a society without oppression. Focusing on the communicative basis of society, Habermas envisions a polity characterized by undistorted communication. Feminists who embrace this view argue that it provides an appropriate basis for feminist politics.

The hermeneutics of Hans-Georg Gadamer has also attracted feminist attention. Although Gadamer is usually viewed as a conservative, some feminists drawn on his writings. Like Habermas, Gadamer attacks the Enlightenment conception of a single path to truth, arguing that there are many paths other than that of reason and logic. Gadamer also challenges the hegemony of the autonomous, rational subject, emphasizing instead the way in which languages create the "horizon of meaning" in which we live. For Gadamer, "horizons" are perspectives in which we are all located, positions from which we understand the world. Like other approaches rooted in language, Gadamer's approach has allowed feminists to analyze the linguistic constitution of social reality, and, in particular, the historical context that informs that reality. Linda Alcoff (1996) and Lorraine Code, for example, argue that feminists can employ Gadamer's approach to articulate an understanding of knowledge that is engaged, situated, and feminist.

See also Beauvoir, Simone de; Cixous, Helene; Code, Lorraine; Continental Philosophy; Critical Theory; Derrida, Jacques; Enlightenment; Feminist Epistemology;

Feminist Metaphysics; Feminist Philosophy; Foucault, Michel; Gadamer, Hans-Georg; Habermas, Jürgen; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Hermeneutics; Husserl, Edmund; Irigaray, Luce; Kristeva, Julia; Marx, Karl; Marxist Philosophy; Merleau-Ponty, Maurice; Nietzsche, Friedrich; Postmodernism; Rationalism; Sartre, Jean-Paul; Structuralism and Poststructuralism.

Bibliography

Alcoff, Linda. *Real Knowing*. Ithaca: Cornell University Press, 1996.

Bartky, Sandra. *Femininity and Domination: Studies in the Phenomenology of Oppression*. New York: Routledge, 1990.

Butler, Judith. *Gender Trouble*. New York: Routledge, 1990.

Cahill, Ann, and Jennifer Hansen. *Continental Feminism Reader*. Lanham, MD: Rowman and Littlefield, 2003.

De Beauvoir, Simone. *The Second Sex*. Hammondsworth: Penguin, 1972.

Hartsock, Nancy. *Money, Sex, and Power*. New York: Longman, 1983.

Hekman, Susan. *Gender and Knowledge: Elements of a Postmodern Feminism*. Cambridge: Polity, 1990.

Nicholson, Linda, ed. *Feminism/Postmodernism*. New York: Routledge, 1990.

Smith, Dorothy. *The Everyday World as Problematic: A Feminist Sociology*. Boston: Northeastern University Press, 1987.

Young, Iris Marion. *Throwing Like a Girl and Other Essays in Feminist Philosophy and Social Theory*. Bloomington: Indiana University Press, 1990.

Susan Hekman (2005)

FEMINISM AND PRAGMATISM

Pragmatist feminists hold some or all the following conceptual commitments, which are rooted in classical pragmatism:

(1) A rejection of foundationalist and essentialist notions of reality and truth, in favor of an understanding of reality as the result of mutually constitutive transactions between agents and their environments and of truth as good knowing that it enables an inquiry to grow

(2) A recognition that chance and uncertainty are parts of one's world, not (necessarily) signs of one's incomplete understanding of that world

(3) A rejection of sharp dichotomies separating theory from practice, self from world, mind from body, fact from value, and reason from emotion

(4) A view of inquiry as experiential and experimental: Inquiry springs from experience, and its findings must have the capacity to improve on experience, for the individual or for society

(5) Respect for the philosophical value of ordinary, everyday experience—including experiences that characterize women's lives

(6) Cognizance that the community of inquirers plays a central role in inquiry and a commitment to improving the goodness of inquiry by actively increasing the perspectives represented in the community

(7) An understanding of ends, aims, and values as experimental: subject to revision in light of new experiences

(8) A recognition that democracy provides a model of intellectual and moral growth for society and the individual possessing the greatest capacity to promote justice

Feminist strands of pragmatism stand in the somewhat unusual position of having been part of their parent tradition virtually since that tradition emerged in the late nineteenth century. However, only in the 1980s did an explicitly, self-consciously pragmatist feminist philosophical movement emerge.

EARLY FIGURES

The pragmatist movement counted women and feminists among its members from the early days; many were associated with the classical pragmatist John Dewey as his colleagues and as his students. These theorists worked almost exclusively at the margins of academic philosophy, as educators and school administrators, policy makers, and social activists. While their outsider status was not always chosen, their decisions to work in the community as teachers, policy makers, and community workers nevertheless embody a pragmatist commitment to creating philosophy that works to ameliorate the problems of everyday life—not simply the problems of philosophy.

Among women who contributed to the emergence of pragmatist thought in explicitly feminist ways, perhaps none exerted greater influence on subsequent pragmatist feminism than Jane Addams (1860–1935), a social activist, theorist, and founder of the Hull House settlement. Her choice to theorize with, rather than about, the people of the neighborhoods surrounding Hull House embodied a pragmatist understanding that inquiry transforms both inquirer and inquired; she and the other resi-

dents of Hull House produced both theory and public policy that began in, and returned to, the problems of the people of their community. Addams's feminism was rooted in her understanding that women are, by enculturation if not by nature, different from men, and that such differences constitute actual assets—in city government, for instance. There, women's experiences as homemakers and mothers directly prepare them for the associated tasks of running a city. Addams's long association with Dewey significantly shaped the intellectual development of both theorists, particularly in the areas of democracy and education.

Charlotte Perkins Gilman (1860–1935), a contemporary of Dewey and an acquaintance of Addams, shared with them a debt to the evolutionary theory of Charles Darwin. All three understood evolutionary theory to assert that humans have both a capacity and an obligation to improve the conditions of their world through reflection on concrete experience; they also understood that value concepts like improvement, progress, and the good themselves evolve as a result of reflection and action; they are not fixed and timeless. Gilman's *Women and Economics* (1998) offers an evolutionary account of human social development that argues for the necessity (indeed, inevitability) of women's evolution as workers and public figures; only as women so develop will humans realize their social and intellectual potential. Gilman argued for transformations of domestic life to enable women to take their place in the world of work: public kitchens and day care centers, for instance. She saw these proposals echoed in organizations and programs for working women developed at Hull House.

Several early women pragmatists worked as educators. Ella Flagg Young (1845–1918), Elsie Ripley Clapp (1879–1965), and Lucy Sprague Mitchell (1878–1967) studied with Dewey—in some cases, when they were already mature thinkers who exerted an influence on him. All three worked actively to develop pragmatist models of education that emphasized experiential, student-oriented, community-based learning: Young, as general supervisor of the Laboratory School at the University of Chicago; Mitchell, as a researcher and founder of the Bank Street School in New York; and Clapp, as the head of a community school system in Arthurdale, West Virginia, the first New Deal community in the United States.

CONTEMPORARY THEORISTS

Contemporary pragmatist feminists, like feminists working in other traditions, have undertaken two separate but related projects: reclaiming forgotten or neglected work of early women/feminist pragmatists and advancing pragmatist thought by developing new, explicitly feminist versions of it. In the first category, *Pragmatism and Feminism* (1996), by Charlene Haddock Seigfried, presents a systematic exposition of the contributions of early women pragmatists, documenting the lines of influence running among Addams, Clapp, Mitchell, Young, and Dewey. As Seigfried points out, such recovery work transforms both the history of pragmatist philosophy (restoring important voices that were lost) and its conceptual frameworks (engendering a reconceptualization of pragmatist positions that incorporates feminist contributions) (p. 6). Illustrative of this transformation is the work of Marilyn Fischer and Judy D. Whipps (2003), who elucidate the importance of Addams's work for the pragmatist tradition in their edited collection of her writings on peace.

Theorists working on the second task—developing feminist versions of pragmatist thought—draw on the (implicitly and explicitly) feminist and antiracist thought of several earlier pragmatist thinkers, including Addams, Dewey, W. E. B. Du Bois (1868–1963), and Alain Locke (1886–1954). Seigfried's *Pragmatism and Feminism* also marks the most significant early contribution to this project; it lays out a broad, flexible research agenda in epistemology, ethics, and sociopolitical philosophy to be undertaken by pragmatists and feminists using "a pragmatic hermaneutics of cooperation" and aimed at "changing the theoretical analyses and concrete practices of both" (1996, p. 4).

Much pragmatist feminist development has been in the area of feminist epistemology. Theorists here ground their work in the pragmatist emphasis on the primacy of experience and the experiential nature of knowing. Inquiry begins in the problems of ordinary life and possesses a melioristic function; this naturalistic epistemology is grounded in pragmatist thinkers such as Dewey and should not be conflated with Willard Van Orman Quine's naturalized epistemology.

Pragmatist feminist theorists also emphasize the pragmatist commitment to undermine or dissolve traditional dualisms between self and world, mind and body, and theory and practice. Shannon Sullivan challenges the self-world dichotomy to develop a Deweyan feminist understanding of humans as "transactional," where transaction is understood as "an active and dynamic relationship between things such that those things are co-constitutive of each other" (2001, p. 12). This gives rise to a conception of truth as "transactional flourishing":

truth and objectivity are conceived not in terms of transparency to reality, but as characteristics of transactions that enable both humans and their environments to flourish.

Pragmatist feminists have developed conceptions of reason, rationality, and objectivity that recognize the inherently collective, relational nature of these concepts—and that thus acknowledge their ethical, social, and political dimensions along with the epistemological. Lisa Heldke (1990) conceives a "coresponsible model" of objectivity grounded in responsibility to the inquiry community; on this model, inquiry becomes more objective as it acknowledges, fulfills, and expands responsibility to an increasingly pluralistic community. Reflecting the pragmatist commitment to problems of ordinary life, she develops the model through an analysis of food making, conceived as a "thoughtful practice"—a categorization that eschews the traditional division drawn between theoretical and practical activity.

Another significant body of work has developed in social and political philosophy. Theorists here utilize pragmatist understandings that social and moral ends are themselves subject to revision in light of new experience and that intelligent inquiry has melioristic potential and the pragmatist commitment to democracy, understood as a way of living emphasizing collective experimentation to transform current social realities. In *The Task of Utopia* (2001) Erin McKenna develops a pragmatist feminist concept of utopia, which understands it not as a fixed state, but as a characteristic of a (democratic) community's collective inquiry and education process. Such a utopia is necessarily open-ended, its aims always in principle subject to revision.

Pragmatist feminists deepen classical pragmatist notions of community, which emphasize the importance of pluralism for democracy and inquiry; and of personhood, which reject liberal notions of the individual in favor of a relational, transactional model. Feminists show why the perspectives of marginalized persons must be explicitly sought, if people's democratic communities are to continue to grow, promote justice, and create more reliable understandings of social reality. Whipps (2004) draws from Addams a form of communitarianism that rejects the radical individualism characteristic of its contemporary forms and recognizes the (messy, multiplicitous) ways selves are constituted through the interactions of daily life in the diverse community. And in *Deep Democracy* (1999) Judith Green creates a model of democratic practice as experimental. Her "radical critical pragmatism … engage[s] with liberalism, communitarianism, postmodernism, critical theory, feminism, and cultural pluralism" (p. x), not simply to identify the weaknesses of these other traditions, but also to draw on these expanded resources to address concrete problems of democracy, most notably racial, economic, and sexual injustice.

See also Feminist Epistemology; Feminist Philosophy; Feminist Philosophy of Science.

Bibliography

Addams, Jane. *Democracy and Social Ethics*. Introduction by Charlene Haddock Seigfried. Urbana: University of Illinois Press, 2002.

Addams, Jane. *Jane Addams's Writings on Peace*. 4 vols., edited by Marilyn Fischer and Judy D. Whipps. Bristol, U.K.: Thoemmes Press, 2003.

Gilman, Charlotte Perkins. *Women and Economics: A Study of the Economic Relation between Men and Women as a Factor in Social Evolution*. Minneola, NY: Dover, 1998.

Green, Judith. *Deep Democracy: Community, Diversity, and Transformation*. Lanham, MA: Rowman and Littlefield, 1999.

Heldke, Lisa. "Foodmaking as a Thoughtful Practice" and "Recipes for Theory Making." In *Cooking, Eating, Thinking: Transformative Philosophies of Food*, edited by Deane Curtin and Lisa Heldke. Bloomington: Indiana University Press, 1990.

McKenna, Erin. *The Task of Utopia: A Pragmatist and Feminist Perspective*. Lanham, MA: Rowman and Littlefield, 2001.

Rorty, Richard. "Feminism and Pragmatism." *Michigan Quarterly Review* 30 (1991): 231–258.

Seigfried, Charlene Haddock, ed. *Feminist Interpretations of John Dewey*. University Park: Pennsylvania State University Press, 2002.

Seigfried, Charlene Haddock. *Pragmatism and Feminism: Reweaving the Social Fabric*. Chicago: University of Chicago Press, 1996.

Seigfried, Charlene Haddock, ed. "Special Issue on Feminism and Pragmatism." *Hypatia* 8 (2) (1993).

Sullivan, Shannon. *Living across and through Skins: Transactional Bodies, Pragmatism, and Feminism*. Bloomington: Indiana University Press, 2001.

Whipps, Judy D. "Jane Addams's Social Thought as a Model for a Pragmatist-Feminist Communitarianism." *Hypatia* 19 (3) (2004): 118–133.

Whipps, Judy D. "Pragmatist Feminism." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University, 2004.

Lisa Heldke (2005)

FEMINISM AND THE HISTORY OF PHILOSOPHY

The beginning of the twenty-first century was witness to an emergent transformation of the history of philosophy. While still the subject of intense debate within philosophy, the dominance of the image of the history of philosophy as a succession of "master thinkers" whose texts provide the historical background to contemporary philosophical debates has begun to wane. As philosophers come to embrace the historiography of philosophy and accept that attention to the past is not a simple process of reading past masters, methodological issues have become central to the history of philosophy and questions are being raised concerning the canonization of both theorists and texts, the conceptual role of history in philosophy, the accessibility of the past, and the role of interpretation.

Feminist history of philosophy has played a significant role in this transformation. From its outset, feminist historians of philosophy have raised issues of canon formation and have developed new and productive reading strategies in their efforts to attend both to women and to the role of the feminine in the history of philosophy. These efforts to understand the apparent absence or denigration of women and of the feminine have led to interpretive strategies that have value beyond feminist concerns and have contributed to the transformation of contemporary history of philosophy.

Feminist attention to gender in the history of philosophy has led to the recovery of lost or silenced women philosophers, as well as having called into question models of philosophy and philosophical concepts emerging from a privileging of the masculine. As feminists came to understand the extent to which privileged concepts such as reason and justice revolve around the denigration of so-called "feminine" traits, they began not only to question the division between reason, emotion, and imagination in the history of philosophy, but also to search for and develop interpretive strategies that would not perpetuate such divisions.

ATTENTION TO WOMEN

Feminist attention to women in the history of philosophy has raised issues concerning canon formation. Until the mid- to late twentieth century, much of contemporary history of philosophy proceeded along a model of "master thinkers" in which only the truly great minds of philosophy are considered worthy of attention. Admittedly there has been significant debate within the various traditions of philosophy as well as between different historical periods concerning which philosophers are indeed worthy. In addition, even when there is general agreement about the canonization of such philosophers as René Descartes, David Hume, Immanuel Kant, and Plato, there remains significant contestation concerning which aspects of their corpus are most central, with Descartes's *Meditations*, for example, receiving far more attention than his *Passions of the Soul* in twentieth-century analytic history of philosophy.

As feminist philosophers of history contest the "great man" model of history, they have begun to demonstrate the importance of a richer approach to the history of philosophy. The recovery of women philosophers like Elisabeth of the Palatine, Jane Addams, Mary Astell, Sor Juana Inés de la Cruz, Jacqueline Pascal, Anna Maria van Schurman, and Mary Wollstonecraft has begun to transform modern prejudices about the history of philosophy. Since there were hundreds of women who contributed to philosophy, their absence from contemporary histories brings to the foreground the complex values that inform the narratives of philosophy and determine which questions and styles count as philosophical and whose voices are sufficiently influential to be chronicled. Feminist historians of philosophy have demonstrated, for example, how the nineteenth century move to excise from the canon work judged to be motivated by religious faith resulted in numerous philosophical schools and philosophical styles, and with them the work of many women, being excluded from the domain of philosophy. Feminists have also pointed out that if we limit our definition of philosophy to that work done only in the academy and the seminary, then we will exclude those locations, such as the convent and the salon, where women are most likely to be found in certain historical periods.

These investigations of the roles of women in philosophy have led to an enriched appreciation of the workings of the canon. For instance, feminist attention to the philosophy of Princess Elisabeth and the impact of her philosophical influence on Descartes has led to a renewed appreciation not only of *Passions of the Soul*, but of Elisabeth's philosophy in its own right and of her influence on Descartes's philosophy. Such feminist work details Elisabeth's efforts to develop a unique philosophical position that does not divorce reason from the body, but defends a rich interaction between the body and the mind without reducing one to the other or denying Descartes's intuition that thought is not determined by extension. Thanks to such work feminist historians of philosophy have been able to uncover lines of influence between Elisabeth's

thought and Descartes's *Passions*, arguing for a subtle yet important shift in his ideas concerning the role of embodiment upon the mind resulting from their correspondence. In this way, recovery of the work of women philosophers and the feminist desire to undo the denigration of faculties and traits (such as the body) that have been associated with the feminine go hand-in-hand with a rereading of the canon.

Feminist attention to women has also included a chronicling of philosophers' perceptions of woman. Through this lens feminists have uncovered a systematic perception of woman and the feminine as inferior and man and the masculine as the true form. This has led philosophers of sexual difference such as Luce Irigaray to argue that woman has been defined not in terms of true difference, but in terms of lack according to an A (male) / -A (female) logic, a logic well illustrated by Hegel's claim that women while educable, are not capable of activities like science or philosophy that demand a universal faculty. In such a schema, woman and the feminine receive no positive definition, no true difference, but are merely an inferior inversion of the masculine. These investigations have led to the contention that the very concepts of philosophy—reason, justice, virtue—have themselves been inscribed by this conception of man and thereby by the masculine as the true form.

PHILOSOPHICAL IMAGINARY

Feminist attention to gender thus presents as an issue central to philosophical investigation the question of whether the central categories of philosophy are formed through an exclusion or denigration of the feminine. Genevieve Lloyd's early study of the "maleness" of reason demonstrated that conceptions of rationality have privileged traits historically associated with masculinity and required control or transcendence of those traits historically associated with the feminine such as the body, the emotions, and the passions. Michèle Le Dœuff has referred to the often unacknowledged linkage of concepts, images, and metaphors in philosophical texts as the philosophical imaginary. She argues that this imaginary often inscribes values historically associated with masculinity onto dominant philosophical conceptions of reason and argues that this is not an instance of an individual philosopher's sexism that can be ignored or excised for it is at the core of the values from which the category emerges.

This scholarship has led to various efforts to identify and refigure the role of "the feminine" in the texts of canonized philosophers and to examine the specifically feminine sites of philosophy. These reading strategies are diverse. Some, like Annette Baier's work on Hume or Barbara Herman's analysis of Kant, return to the canonical texts to tease out new or overlooked resources for revaluing the role of embodiment, imagination, and the affective life. Others turn to the work of "recovered" women philosophers to trace alternatives to dominant models of philosophy. Catherine Villanueva Gardner, for example, argues that a complex notion of sensibility and a rhetorical style that exemplifies sensibility can be found in the work of women philosophers such as Wollstonecraft, Catharine Macaulay, Christine De Pisan, George Eliot, and Mechthild of Magdeburg that provide a rich conception of the role the passions play in moral philosophy. Another reading strategy is to provide correctives to histories of philosophy that have ignored topics like the emotions or the imagination as does Susan James (1997) in her account of the passions in seventeenth-century philosophy. Yet another style of feminist reading can be found in the work of Luce Irigaray who focuses on the moments of instability in philosophical texts caused by the contradictory effort to achieve universality through a denial of sexual difference. It is her goal to open the historical texts of philosophy to contemporary feminist concerns not simply to confront what has been repressed, but to rethink it.

Feminist attention to the philosophical imaginary and the lessons learned from the canonization of particular philosophical styles, has led to sensitivity to the rhetorical dimensions of philosophical writings, as well as to an appreciation of their affective dimensions. But such attention to style also means a rich situating of the history of philosophy and a realization that the writings of the past are not transparent. The meanings and affective resonance of philosophical texts are neither in the control of the author nor the contemporary interpreter of the text, but involve a complex interplay between the author's cultural context and the concerns of the contemporary reader. In this way, mainstream efforts to excise the figural in order to uncover the literal truth of canonical texts give way in feminist rereadings to an appreciation of the role of imagination in philosophy and better understanding of how reason, imagination, and emotion are interwoven in the practice of philosophy. This attention to rhetoric and affect is another dimension of the feminist rejection of conceptions of reason divorced from the "feminine."

In such attention to neglected aspects of historical texts, feminists are motivated by our own feminist wonder at the relation between reason and emotion in the

play of the canon and a feminist inspired desire to find a place in-between mind and body. In this sense, our desires are enacted in our reading strategies.

See also Astell, Mary; Baier, Annette; Descartes, René Eliot, George; Elisabeth, Princess of Bohemia; Feminist Philosophy; History and Historiography of Philosophy; Hume, David; Irigaray, Luce; Kant, Immanuel; Plato; Wollstonecraft, Mary; Women in the History of Philosophy.

Bibliography

Atherton, Margaret, ed. *Women Philosophers of the Early Modern Period*. Indianapolis, IN: Hackett, 1994.

Baier, Annette. *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, MA: Harvard University Press, 1991.

Bordo, Susan. *The Flight to Objectivity: Essays on Cartesianism and Culture*. Albany: State University of New York Press, 1987.

Conley, John J. *The Suspicion of Virtue: Women Philosophers in Neoclassical France*. Ithaca, NY: Cornell University Press, 2002.

Deutscher, Penelope. *Yielding Gender: Feminism, Deconstructionism, and the History of Philosophy*. London and New York: Routledge, 1997.

Dykeman, Therese Boos, ed. *The Neglected Canon: Nine Women Philosophers First to the Twentieth Century*. Boston and London: Kluwer Academic Publishers, 1999.

Gardner, Catherine Villanueva. *Rediscovering Women Philosophers: Philosophical Genre and the Boundaries of Philosophy*. Boulder: Westview Press, 2000.

Herman, Barbara. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Irigaray, Luce. *An Ethics of Sexual Difference*. Translated by Carolyn Burke and Gillian C. Gill. Ithaca, NY: Cornell University Press, 1993.

Irigaray, Luce. *Speculum of the Other Woman*. Translated by Gillian C. Gill. Ithaca, NY: Cornell University Press, 1985.

James, Susan. *Passions and Action: The Emotions in Seventeenth-Century Philosophy*. Oxford: Clarendon Press, 1997.

Kofman, Sarah. *Socrates: Fictions of a Philosopher*. Translated by Catherine Porter. Ithaca, NY: Cornell University Press, 1998.

Le Dœuff, Michèle. *The Philosophical Imaginary*. Translated by C. Gordon. Stanford, CA: Stanford University Press, 1989.

Lloyd, Genevieve. *Feminism and the History of Philosophy*. New York: Oxford University Press, 2002.

Lloyd, Genevieve. *The Man of Reason: "Male" and "Female" in Western Philosophy*. Minneapolis: University of Minnesota Press, 1984.

Nye, Andrea. *The Princess and the Philosopher: Letters of Elisabeth of Palatine to René Descartes*. Lanham, MD: Rowman and Littlefield, 1999.

Oliver, Kelly. *Womanizing Nietzsche: Philosophy's Relation to "The Feminine."* New York: Routledge, 1995.

O'Neill, Eileen. *Women Philosophers of the Seventeenth and Eighteenth Centuries: A Collection of Primary Sources*. Oxford: Oxford University Press, 1998.

Schott, Robin. *Cognition and Eros: A Critique of the Kantian Paradigm*. Boston: Beacon Press, 1988.

Tougas, Cecile T., and Sara Ebenrick, eds. *Presenting Women Philosophers*. Philadelphia: Temple University Press, 2000.

Tuana, Nancy, series ed. *Re-Reading the Canon*. 26 vols. University Park: Pennsylvania State University Press, 1994–2005.

Tuana, Nancy. *Woman and the History of Philosophy*. New York: Paragon Press, 1992.

Waithe, Mary Ellen, ed. *A History of Women Philosophers*. Vols. 1–4. Boston and London: Kluwer Academic Publishers, 1987–1995.

Nancy Tuana (2005)

FEMINIST AESTHETICS AND CRITICISM

As artists in the late 1960s and early 1970s started to produce explicitly feminist works, critics and historians of the various arts began to examine a previously unnoticed gender bias in the Western artistic tradition. Feminists discern this bias on two levels.

First, feminist critics charge that canonical artworks represent women and men in markedly different ways, a difference evident in the organization and scenarios of the works themselves. Whereas men are typically portrayed as strong, active, heroic, and playing important historical roles, women are nearly always shown as weak, inert, and vulnerable; in domestic or nurturing roles; identified with nature; and as sexually available for men's needs. This is perhaps most evident in the visual arts where representations of passive, anonymous, and vulnerable female nudes dominate many historical periods. Drawing on semiotics, psychoanalysis, and Marxist theory, feminists sought to expose and analyze manifestations of gender bias in structural features of traditionally admired artworks. One of the most influential concepts developed in this early period of criticism is the notion of "the male gaze" (Mulvey 1975). Although it is sometimes mistaken for an empirical description of individuals' actual viewing practices, "the male gaze" in fact refers to the viewpoint that many pictures adopt toward women, portraying women as passive objects of sexual desire.

Second, feminists argue that fully addressing gender inequality in the arts also requires questioning the canon; that is, those works traditionally deemed artistically excellent that form the core of a given discipline. Feminists are skeptical of the canon for two reasons. First,

although women make up roughly half of the population, they are almost entirely absent from the pantheon of great artists. Second, the kinds of artifacts traditionally produced by women—for example, quilts, pottery, needlework, and weaving—have not been taken seriously as art but rather have been relegated to the diminished categories of "decorative arts" or "crafts." The coincidence of pervasive gender inequality in the world with the exclusion of women's artifacts from the canon suggests that the canon might be shaped by more than purely aesthetic concerns. But what exactly is the relationship between unequal social relations and women's lack of representation in the canon? What explains the paucity of great women artists and the underestimation of artifacts customarily produced by women?

Some feminists, most notably Linda Nochlin (1971), argue that social, economic, and institutional barriers have prevented women from making art. For instance, in much of Europe in the nineteenth century women were not allowed to attend life-drawing classes and so lacked the training and practice necessary to adequately represent the human form. Although such obstacles and lack of opportunity surely contributed to the canon's one-sided configuration, this explanation has difficulty accounting for two facts: First, despite these adverse conditions, some women have been making oil paintings, sculptures, and the like for centuries, yet none number among the canon of great artists, and second, women still encounter discrimination in the contemporary art world (Guerilla Girls 1998). The historical explanation also has trouble accounting for the exclusion of *kinds* of artifacts conventionally produced by women.

Such questions prompt a need to examine traditional understandings of art. Might the prevailing standards of artistic excellence be tainted by biases that help explain why women and the artifacts they customarily produce have been excluded from the ranks of artistic greatness? At this point feminist philosophers and theoreticians enter the conversation to scrutinize the philosophical canon itself and analyze established theories of art, artistic talent, and aesthetic experience and value.

In their critical examination of the Western philosophical tradition feminists uncover and analyze previously unnoticed gender biases in theories of art from Plato onward. Some contend, for instance, that central aesthetic concepts such as "genius" and "masterpiece" have been traditionally gendered male (Battersby 1989). Others argue that influential theories of aesthetic perception implicitly take men's experience as their model by favoring sight and hearing, which customarily play a prominent role in men's lives, and by underestimating the aesthetic importance of those senses integral to the social roles assigned to women, namely touch, smell, and taste (Korsmeyer 2004). Finally, many feminist philosophers are critical of a cluster of theories and concepts that assume or attempt to justify the autonomy of art and of aesthetic appreciation and evaluation (for an overview, see Devereaux 1998). For example, some maintain that the common insistence on art's segregation from practical concerns results in the art-craft distinction and hence in the systematic depreciation of the sorts of artifacts customarily produced by women. Others make the case that the related doctrine of aesthetic formalism, which restricts artistic value to a work's formal features, departs in practice from purely formal concerns by reflecting masculine preferences for particular themes (such as the female nude). In these ways feminists argue that the presumed disinterestedness and universality of aesthetic judgment in theories following Immanuel Kant mask standards of evaluation that are partial to men's experience, preferences, and sensibilities.

Once the sources of this undervaluation of women's artistic efforts have been uncovered and analyzed, feminists then aim to delineate the positive means to overcome it. Besides providing women with opportunities in the art world, the prevailing conceptions of art and standards of artistic excellence must be revised. On this point most agree, yet several different solutions can be distinguished.

PERSPECTIVISM

One approach calls for the outright abandonment of the problematic concepts, methods, and categorizations of traditional aesthetics. Artistic autonomy, aesthetic formalism, the art-craft distinction, presumptions of a disinterested aesthetic attitude, and concepts of talent or genius are all to be rejected in favor of a perspectivism that embraces a pluralistic conception of art and artistic value (Hein and Lauter 1993). This approach eschews all pretension to universal standards of aesthetic excellence, leaving no standpoint from which to adjudicate between differing understandings of art and aesthetic experience. In practice this has led some art historians and critics to reject the notion of artistic canons altogether and to replace talk of art with that of visual or material culture (Pollock 1999).

One concern is that this perspectival approach risks rendering any notion of artistic value meaningless, a result that is particularly unwelcome given feminists' efforts to demonstrate the artistic merit of women's arti-

factual efforts. Another worry is that one ought not mistake the discriminatory and faulty use of concepts such as genius or of methods like formalism for inherent features of these concepts, methods, or standards themselves. It does not follow from the fact that the so-called universal voice of aesthetic judgment has surreptitiously been biased toward masculine concerns that the very ideal of universality in aesthetic judgment is inherently gender biased. Indeed, that traditional theories of art have been criticized for their *bias* is evidence of feminism's reliance on the notion of impartial standards of artistic excellence.

REVISIONISM

Some feminists warn against the assumption that all of aesthetic theory has been tainted by gender bias (Felski 1998) and point to developments in philosophical aesthetics, such as the critique of disinterestedness, that are continuous with feminism's aims (Silvers 1998). Others show how at least aspects of certain ideals such as artistic autonomy are actually useful for feminism (Devereaux 1998). These developments suggest that feminism might be compatible with traditional theories of art and aesthetic experience, provided that these theories are purged of their masculine biases. This could motivate revaluation of those canonical works that cater to male-defined assumptions about women, on the one hand, and would allow these theories and their central concepts to be adapted to the kinds of objects customarily produced by women, on the other hand. In practical terms this approach would mean integrating women's artistic efforts into the canon, a process that some historians and critics have already begun (Guerilla Girls 1998).

DIFFERENCE AESTHETICS

Still, some insist, incorporating women into the canon misses what is distinctive about their art. Likewise, they contend, traditional aesthetic theories cannot be adequately modified to capture the uniqueness of women's experience, preferences, values, sensibilities, and modes of expression. Instead, a variety of alternative aesthetic concepts and theories of art indigenous to women is proposed (Battersby 1989, Frueh 1998, Robinson 2001, Barwell 1993, Donovan 1993, French 1993, Lorraine 1993). Some French feminists like Irigaray and Kristeva, for instance, argue that women imagine, express themselves, and experience art somatically or experimentally, and that these distinctive methods require standards, concepts, and definitions of art that differ radically from the traditional ones (See Korsmeyer, 2004, Chapter 6, for an overview). In practical terms, this could lead to the formation of separate women's canons in each of the arts.

Critics charge that this approach rests on false essentialist assumptions about woman's nature and overlooks important differences between women such as ethnicity, race, class, sexual orientation, ability, and age, to name only a few (Felski 1998). Some also worry that separate principles and criteria of artistic excellence and aesthetic experience risk leaving the canon with its biases in tact while ghettoizing women's art (Nochlin 1971, Pollock 1999).

The debate about how to deal with gender bias in artworks, canon formation, and traditional theories of art is lively and ongoing. Many of the disputes rest on the question of how, if at all, gender matters to the production, appreciation, and evaluation of art. Besides these unresolved questions, all approaches face new challenges such as the insistence that one cannot divorce feminist struggles from those of other disenfranchised groups. For these reasons, feminist aesthetics does not involve a particular stance or methodological commitment but, rather, unites a variety of approaches toward the common goal of ending women's subordination in the arts and discourses about the arts.

See also Aesthetics, History of; Feminism and Continental Philosophy; Feminist Philosophy.

Bibliography

Barwell, Ismay. "Feminine Perspectives and Narrative Points of View." In *Aesthetics in Feminist Perspective*, edited by Hilde Hein and Carolyn Korsmeyer, 93–104. Bloomington: Indiana University Press, 1993.

Battersby, Christine. *Gender and Genius: Towards a Feminist Aesthetics*. Bloomington: Indiana University Press, 1989.

Brand, Peg Zeglin, and Mary Devereaux. "Women, Art, and Aesthetics." *Hypatia: A Journal of Feminist Philosophy* (Special Issue) 18 (4) (2003).

Brand, Peg Zeglin, and Carolyn Korsmeyer, eds. *Feminism and Tradition in Aesthetics*. University Park: Pennsylvania State University Press, 1995.

Devereaux, Mary. "Autonomy and Its Feminist Critics." In *Encyclopedia of Aesthetics*. Vol. 1, edited by Michael Kelly, 178–182. New York: Oxford University Press, 1998.

Donovan, Josephine. "Everyday Use and Moments of Being: Toward a Nondominative Aesthetic." In *Aesthetics in Feminist Perspective*, edited by Hilde Hein and Carolyn Korsmeyer, 53–67. Bloomington: Indiana University Press, 1993.

Felski, Rita. "Feminism: Critique of Feminist Aesthetics." In *Encyclopedia of Aesthetics*, edited by Michael Kelly, 170–172. New York: Oxford University Press, 1998.

French, Marilyn. "Is There a Feminist Aesthetic?" In *Aesthetics in Feminist Perspective*, edited by Hilde Hein and Carolyn

Korsmeyer, 68–76. Bloomington: Indiana University Press, 1993.

Frueh, Joanna. "Towards a Feminist Theory of Art Criticism." In *Feminist Art Criticism: An Anthology*, edited by Arlene Raven, Cassandra L. Langer, and Joanna Frueh, 153–165. Ann Arbor, MI: UMI Research Press, 1998.

Guerilla Girls. *The Guerilla Girls' Bedside Companion to the History of Western Art*. New York: Penguin Books, 1998.

Hein, Hilde, and Carolyn Korsmeyer, eds. *Aesthetics in Feminist Perspective*. Bloomington: Indiana University Press, 1993.

Hein, Hilde, and Estella Lauter. "Re-enfranchising Art: Feminist Interventions in the Theory of Art." In *Aesthetics in Feminist Perspective*, edited by Hilde Hein and Carolyn Korsmeyer, 21–34. Bloomington: Indiana University Press, 1993.

Korsmeyer, Carolyn. *Gender and Aesthetics: An Introduction*. New York: Routledge, 2004.

Lorraine, Renée. "A Gynecentric Aesthetic." In *Aesthetics in Feminist Perspective*, edited by Hilde Hein and Carolyn Korsmeyer, 35–52. Bloomington: Indiana University Press, 1993.

Mulvey, Laura. "Visual Pleasure and Narrative Cinema." *Screen* 16 (3) (1975): 6–18.

Nochlin, Linda. "Why Have There Been No Great Women Artists?" *ARTnews* 69 (9) (1971): 22–39.

Parker, Rozsika, and Griselda Pollock. *Old Mistresses: Women, Art, and Ideology*. New York: Pantheon Press, 1981.

Pollock, Griselda. *Differencing the Canon: Feminist Desire and the Writing of Art's Histories*. London: Routledge, 1999.

Robinson, Hilary, ed. *Feminism—Art—Theory: An Anthology, 1968–2000*. Malden, MA: Blackwell, 2001.

Silvers, Anita. "Feminism: An Overview." In *Encyclopedia of Aesthetics*. Vol. 2, edited by Michael Kelly, 161–167. New York: Oxford University Press, 1998.

A. W. Eaton (2005)

FEMINIST EPISTEMOLOGY

Feminist epistemology emerges from reflection on feminist inquiry. Core themes in feminist epistemology can be understood by considering a prima facie tension between two distinct strands of feminist research, one critical and one constructive. The critical strand aims to expose male bias in research while the positive strand aims to construct theories that are avowedly feminist and that bring women's experiences and interests to the center of inquiry. Most disciplines have come under critical scrutiny for male bias. Forms of bias identified include:

(1) Marginalizing women or women's interests. For example, economic theory is charged with making women's economic contributions invisible, political theory with overlooking power relations in the family, and evolutionary theory and anthropology with privileging male activities.

(2) Producing theories that naturalize and thus reinforce oppressive gender relations. Primatology and sociobiology are among the disciplines that have been charged with such bias.

(3) Embedding gendered metaphors that bias theory selection.

(4) Presupposing cognitive styles that arise from male psychosocial development. This charge is laid against philosophy, scientific method, and theories of moral development.

A puzzle immediately arises, however: If such research is bad because biased, then how can the constructive strand of feminist research escape a similar charge of bias and hence of epistemic fault? The puzzle deepens still further: Epistemic norms, including norms of objectivity, have themselves been charged with male bias. A charge of bias seems, however, to require a commitment to the value of objectivity. This puzzle is called "the bias paradox" (Antony 1993, pp. 114–115) and provides the context in which core themes in feminist epistemology can be understood. These are: the ideological role of epistemic norms; the importance of situated knowledge; the role of values in inquiry; and the nature of objectivity.

THE IDEOLOGICAL ROLE OF EPISTEMIC NORMS

Feminists have charged epistemic norms with being male biased. MacKinnon's analysis of the stance of objectivity as involving two components—distance and aperspectivity—is representative: "To perceive reality accurately, one must be distant from what one is looking at and view it from no place and at no time in particular, hence from all places and times at once" (MacKinnon 1989, p. 97).

To the extent that a putative knowledge claim can be shown to be the product of the inquirer's social situation, that claim is undercut as knowledge: "If social knowledge can be interpreted in terms of the social determinates of the knower, it is caused. Therefore its truth value, in this definition of the test for truth, is undercut. If it has a time or place—or gender—it becomes doubtful because situated" (MacKinnon 1989, p. 98).

Aperspectivity is alleged to be a "strategy of male hegemony" (MacKinnon 1982, p. 57) that maintains gender relations in three ways: by being implicated in the objectification of women, by masking malebias in research, and by deauthorizing women as knowers. According to MacKinnon, aperspectivity lends support to the (false) belief that women are by nature fitted to the

position of eroticized subordination prescribed by current gender relations. Men project onto women the qualities (e.g., docility and submissiveness) that they desire women have. When such projection is accompanied by the social power to make women behave as desired and to silence contesting conceptions of social reality, women come to have the properties men ascribe to them. The stance of objectivity allows men to assume that the regularities they observe are objective and to overlook the exercise of power that produced them. In this way, aperspectivity masks the fundamentally prescriptive nature of gender norms and thus lends stability to the oppressive relations constitutive of gender.

Aperspectivity also enables mainstream research to evade critical scrutiny. Even though, given the theory-dependence of method, all research requires presuppositions, mainstream theoretical presuppositions will typically not need to be articulated and defended. Since the beliefs that feminists contest are relatively entrenched, it will tend to be feminists and not mainstream researchers who are called on to defend the presuppositions of their research. Thus, credibility is differentially apportioned between feminist and mainstream views on gender.

In addition, norms that disparage knowledge claims that can be explained as the result of the inquirer's social location are incompatible with feminist method, including the method of consciousness raising. While the formats of consciousness-raising groups—a grass-roots phenomenon chiefly of the 1960s and 1970s—differed, they focused on recounting women's day-to-day experiences, especially of intimate relationships, and on their emotional responses to those experiences. In women's often-inchoate responses to their day-to-day experiences were found the resources with which to understand women's social position. Given that this method starts out from a detailed examination of women's lived experience an experience both available because of and constitutive of women's gender subordination, it finds social location to be an epistemic asset rather than a liability. Different epistemic frameworks offer different accounts of when and how social location is an epistemic asset. This is the subject of the next section.

SITUATED KNOWLEDGE

Feminist standpoint theory begins from the Marxist assumption that material life shapes consciousness, and it draws an analogy between the position of the proletariat under capitalism and women under patriarchy. Just as the proletariat has a privileged standpoint from which to understand the nature of capitalist social relations, there is an epistemically privileged *standpoint* from which to understand the nature of patriarchal social relations. The basis of this standpoint lies in the sexual division of labor. Key features of women's relation to material life that Hartsock argues provide the grounds for the feminist standpoint are women's domestic labor and their role in childbearing and rearing; the experience of female embodiment, including pregnancy and lactation; and the relational self-conception that object relations theorists argue is the result of girls being raised by mothers with whom they can share gender identification. The standpoint is identified as feminist rather than as women's standpoint to signal that the understanding it embodies must be struggled for and does not arise simply in virtue of occupying a subordinated social position.

Patricia Collins defends a black feminist standpoint, which she argues generates its own epistemology that emphasizes experience over book learning, dialogue in assessing knowledge claims, and relations of care and personal accountability. She finds the grounds for a black feminist standpoint in black women's experience of multiple oppressions.

Standpoint theorists reject any conception of objectivity that disparages beliefs that are to be explained by the social location of the believer as merely caused and hence as not truth tracking. They thus resolve the bias paradox by claiming that feminist perspectives provide insight into social relations that are obfuscated by dominant nonfeminist perspectives.

Standpoint theory is charged with valorizing oppression, with being unable to explain which standpoints have epistemic privilege without circularity, and with presupposing an overly simple and exclusionary conception of gender. Feminist postmodern charges standpoint theory with essentialism; that is with making false and exclusionary generalizations about women and their experiences. Feminist postmodernism challenges the stability of the category of woman: One is never simply a woman, but always a woman of some particular race, ethnicity, class, sexuality or historical and national location. Gender is constructed differently at each of these intersectional nodes of identity: One cannot extract from these complex and shifting social categories the single variable *gender*. Destabilizing the category *woman* undercuts the possibility of *a* feminist standpoint; moreover, given there is no in principle limit to the fragmentation of social categories, positing a black feminist standpoint likewise risks making false and exclusionary generalizations about black women.

The epistemic resources and liabilities of social location and other aspects of situated epistemic agency (embodiment, specifically human cognitive architecture, and so on) can be recognized without embracing the notion of epistemic privilege characteristic of standpoint theory in its initial formulations. Sandra Harding argues for multiple standpoints and views, each as a source of questions rather than a source of privileged answers. Lorraine Code calls for an epistemology that takes subjectivity into account; that is, for an epistemology that studies the psychology, interests, and social–cultural locations of inquirers. Likewise, feminists influenced by naturalized epistemology call for the empirical study of those features of our situated epistemic agency that enable one to truth-track and those that prevent one from doing so. Naturalized epistemologists view epistemology as the empirical study of knowers; thus, instead of offering a priori defenses of epistemic norms, they defend a posteriori norms of inquiry designed to help human agents—that is, finite embodied, social, agents—reliably to truth track. These tailored epistemic norms might be different for dominants and for subordinates. Whether and when such norms must recommend insulating political values from inquiry is a question to be settled empirically.

VALUES AND INQUIRY

Values and interests are recognized as influencing the choice of research questions, as contributing to the ways knowledge is applied, and as constraining research methods, especially those used in research involving human subjects. There is, however, widespread skepticism about according values and interests any role in *justification*. Inquiry aims at the truth and, the skeptic presses, nonepistemic considerations can only distract from this truth-seeking goal. Permitting moral and political values to influence theory choice leads to wishful thinking and totalitarian constraint on free inquiry.

Feminist epistemologists respond that it is a mistake to see epistemic and nonepistemic values as in competition so that inquiry must be governed *either* by epistemic values *or* by nonepistemic values. Given the underdetermination of theory by evidence, so that a body of evidence counts in support of a theory only given background assumptions, and given the pragmatics of inquiry, which aims not just for truth but for *significant* truth (where significance is a function of the interests motivating the research question), inquiry will be porous to nonepistemic values. These can enter into choice of background assumptions, of explanatory concepts, and of methodological frameworks. What matters is whether the values and interests that enter contribute to the goal of discovering significant truths and whether they are themselves defensible.

Because of their commitment to transforming gender relations, feminists are alert to background assumptions about gender that shape inference from a body of data and that shape choice of explanatory categories (e.g., the use of *dominance* to name a unified trait in primate research). This awareness has provided the platform from which to mount successful critiques of sociobiology, among other disciplines. Helen Longino argues for framework assumptions, including preference for models that allow for ontological heterogeneity and for complex multifaceted interaction over linear relations because only such models can allow one to represent complex human potentialities. This is no defense of wishful thinking: The claim being made is not that humanist political commitments determine which of two equally empirically supported theories to accept but, rather, that these commitments enjoin one to have models that enable such potentialities to be represented *if* they exist.

OBJECTIVITY REVISITED

Even though it is generally accepted that the concept of objectivity has functioned ideologically to deauthorize women as knowers, feminist epistemologists are unwilling to abandon the notion. Some argue that the conception of objectivity found in mainstream epistemology must be radically overhauled and others that mainstream epistemology has the resources to develop a conception of objectivity that is fully compatible with feminist epistemological projects both critical and constructive. A number of alternative feminist accounts of objectivity have been developed in the literature.

Naturalized epistemology rejects any conception of objectivity as requiring presupposition or bias-free inquiry. Given the theory-dependence of method, the success of inquiry depends on presuppositions. Thus, not only is the injunction to eliminate bias impossible to meet, inquiry without presuppositions would get nowhere. Inquiry based on presuppositions can yet be objective: One needs to distinguish the good biases from the bad: Good biases enable one to truth-track; bad biases prevent one from doing so. Presupposition-rich methods can yield knowledge just in case the presuppositions are approximately true.

Working within standpoint theory, Harding (1993) defends “strong objectivity” based on the notion of reflexivity: Subjects of knowledge must themselves become objects of inquiry. Their interests and social posi-

tions must be acknowledged and the presuppositions that flow from them investigated. Communities of inquiry must be made democratic for *epistemic* as well as political reasons. Drawing on postmodernist perspectives, Sandra Haraway reaches similar conclusions claiming that "feminist objectivity means quite simply "situated knowledges" (Haraway 1991, p. 188). Only situated knowers who acknowledge the partiality of their perspectives and their responsibility in adopting them can be held accountable for their knowledge claims. To achieve objectivity, Haraway advocates combining these partial located perspectives though power-sensitive conversation and through a politics of solidarity.

Longino argues that objectivity is not a property of individuals and their methods of inquiry but, rather, of communities and their structure. A community of inquiry is objective just in case it facilitates transformative criticism. In order to do this, the community must be democratically structured: It must have publicly recognized forums for critique and change in response to that critique; it must have publicly recognized standards for evaluating theories and standards that respect both cognitive and social values; and it must be characterized by equality of intellectual authority. Longino's account is procedural: Communities structured in the right way generate knowledge. She claims that this enables her to avoid begging the question about which standpoints are privileged and to avoid the naturalized epistemologist's assumption that some knowledge claims can be taken for granted. It is controversial, however, whether an account of equality of intellectual authority can, without presupposing the truth of at least some contested claims, simultaneously rule out those holding "irrelevant positions"—Longino cites New Age "crystallology" and creationism (1993, p. 118)—recognize the legitimate authority of expertise and not exclude those whose expertise has been denied for economic and political reasons.

All four accounts of objectivity recognize the importance of social relations and institutions in the production of knowledge; thus, feminist epistemology makes an important contribution to social epistemology—that family of theories that investigates epistemic dependencies and the role of social factors in knowledge and justification—by drawing critical attention to the political dimensions of the social.

See also Feminism and Continental Philosophy; Feminist Metaphysics; Feminist Philosophy; Feminist Philosophy of Science.

Bibliography

Alcoff, Linda, and Elizabeth Potter, eds. *Feminist Epistemologies*. New York: Routledge, 1993.

Anderson, Elizabeth. "Knowledge, Human Interests, and Objectivity in Feminist Epistemology." *Philosophical Topics* 23 (1995): 27–58.

Antony, Louise. "Quine as Feminist: The Radical Import of Naturalized Epistemology." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by L. Antony and C. Witt. Boulder, CO: Westview Press, (1993): 185–225.

Antony, Louise. "Sisters, Please, I'd Rather Do It Myself: A Defense of Individualism in Feminist Epistemology." *Philosophical Topics* 23 (2) (1995): 59–94.

Bar On, Bat-Ami. "Marginality and Epistemic Privilege." In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter. New York: Routledge, (1993): 83–100.

Bleier, Ruth. *Feminist Approaches to Science*. New York: Pergamon Press, 1986.

Bordo, Susan. *The Flight to Objectivity*. Albany: State University of New York Press, 1987.

Boyd, Richard. "On the Current Status of the Issue of Scientific Realism." *Erkenntnis* 19 (1983): 45–90.

Butler, Judith. *Gender Trouble*. New York: Routledge, 1990.

Chodorow, Nancy. *The Reproduction of Mothering*. Berkeley: California University Press, 1977.

Code, Lorraine. *Rhetorical Spaces: Essays on Gendered Locations*. New York: Routledge, 1995.

Code, Lorraine. *What Can She Know?* Ithaca, NY: Cornell University Press, 1991.

Collins, Patricia Hill. *Black Feminist Thought*. Boston: Unwin, 1990.

Flax, Jane. "Political Philosophy and the Patriarchal Unconscious: A Psychoanalytic Perspective on Epistemology and Metaphysics." In *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and the Philosophy of Science*, edited by Sandra Harding and Merrill B. Hintikka. Dordrecht, Holland: Reidel, (1983): 245–282.

Gilligan, Carol. *In A Different Voice*. Cambridge MA: Harvard University Press, 1982.

Goldman, Alvin. *Knowledge in a Social World*. Oxford: Oxford University Press, 1999.

Haack, Susan. "Knowledge and Propaganda: Reflections of an Old Feminist." *Partisan Review* 60 (1994): 556–564.

Haraway, Donna. *Primate Visions*. New York: Routledge, 1989.

Harding, Sandra. "Rethinking Standpoint Epistemology: What is Strong Objectivity"? In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter. New York: Routledge, (1993): 49-82.

Harding, Sandra. *The Science Question in Feminism?* Ithaca: Cornell University Press, 1986.

Harding, Sandra. *Whose Science? Whose Knowledge?* Ithaca NY: Cornell University Press, 1991.

Hartsock, Nancy. "The Feminist Standpoint: A Specifically Feminist Historical Materialism." In *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and the Philosophy of Science*, edited by Sanadra Harding and Merrill B. Hintikka. Dordrecht, Holland: Reidgel, (1983): 283–310.

Hrdy, Sarah. *The Woman that Never Evolved*. Cambridge, MA: Harvard University Press, 1981.

Haslanger, Sally. "On Being Objective and Being Objectified." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by L. Anthony and C. Witt. Boulder, CO: Westview Press, 1993.

Keller, Evelyn Fox, and Helen E. Longino, eds. *Feminism and Science*. Oxford: Oxford University Press, 1996.

Keller, Evelyn Fox. *Reflections on Gender and Science*. New Haven: Yale University Press, 1985.

Lloyd, Genevieve. *The Man of Reason*. 2nd ed. Minneapolis: University of Minnesota Press, 1993.

Longino, Helen. *Science as Social Knowledge*. Princeton, NJ: Princeton University Press, 1990.

Longino, Helen. "Subjects, Power, and Knowledge: Description and Prescription in Feminist Philosophy of Science." In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter. New York: Routledge, (1993) 101–120.

Lugones, Maria, and Elizabeth Spelman. "Have We Got a Theory for You! Feminist Theory, Cultural Imperialism, and the Demand for 'The Woman's Voice.'" *Women's Studies International Forum* 6 (1983): 573–581.

MacKinnon, Catherine. "Feminism, Marxism, Method, and the State: An Agenda for Theory." *Signs* 7 (3) (1982): 515–544.

MacKinnon, Catherine. *Towards a Feminist Theory of the State*. Cambridge MA: Harvard University Press, 1989.

Martin, Emily. "The Egg and the Sperm: How Science has Constructed a Romance Based on Stereotypical Male–Female Roles." In *Feminism and Science*, edited by Evelyn Fox Keller and Helen E. Longino. Oxford: Oxford University Press, (1996): 103–117.

Nelson, Lynn Hankinson. *Who Knows*. Philadelphia: Temple University Press, 1990.

Okin, Susan. *Justice, Gender, and the Family*. New York: Basic Books, 1989.

Spelman, Elizabeth. *Inessential Woman: Problems of Exclusion in Feminist Thought*. Boston: Beacon Press, 1988.

Waring, Marilyn. *If Women Counted*. San Francisco: Harper Collins, 1990.

Karen Jones (2005)

FEMINIST ETHICS

Feminist ethics is a diverse and growing body of philosophical work, initially based in the recognition that most canonical accounts of morality neglected, distorted, and/or trivialized women's moral perspectives while either ignoring or defending unjust power imbalances between women and men. Feminist ethicists have largely agreed that women's invisibility in canonical ethical theories—even leaving aside the overtly misogynist statements that also litter the tradition—is not only morally objectionable in and of itself, but also profoundly distorts many of the arguments and conclusions therein. Perhaps the most nearly unanimous claim of feminist ethicists has been that what passes for a human ideal in much of mainstream philosophical ethics is in fact a male or masculine ideal—and that such bias leads us into error not simply about women, but about morality itself.

In general, feminist ethicists suspect that, in ethical theory as in other disciplines of thought and research, what has been portrayed as the human experience is in fact (at least in significant part) the distillation of a very specific experience—namely, that of highly privileged white men who relied on the exploited labor of others (typically men and women of lower economic classes and/or of despised ethnicities, as well as women of their own class and ethnicity) to enable them to pursue higher inquiry. These relationships of unjust privilege and group-based oppression, although they need not characterize human experience, in fact have done so throughout the period of time (including the present) during which Western moral philosophers have developed and refined their theories. These oppressive conditions shape people's moral beliefs, values, priorities, and characters at deep levels.

The task of feminist ethicists is to try to correct for existing biases in moral theory while also developing new theories, concepts, and strategies that will forge a path away from oppression and toward more just and humane social relationships. Bringing a feminist perspective to moral philosophy has included critiquing and reinterpreting both canonical male authors (such as Immanuel Kant, Plato, Friedrich Nietzsche, Aristotle, and David Hume) as well as reclaiming underappreciated female and/or feminist foremothers (including Simone Weil, Iris Murdoch, and Simone de Beauvoir). An early emphasis on criticizing sexist biases in traditional moral theories has given way to the formulation of new theories which, though their degree of engagement and continuity with canonical theories varies widely, all share an understanding of both gender oppression and women's perspectives as fundamental to human experience.

For feminist ethicists, where one stands in a social world pervasively structured by oppression always matters in understanding and evaluating one's moral beliefs and responsibilities. Such analysis is rendered more complex by the fact that gender is only one of many bases for oppression and privilege. Many feminist ethicists (again, like feminists in other disciplines) have devoted significant attention to the intersections among different forms of oppression, including but not limited to oppression on the basis of race, of economic class, of age, of physical and mental ability, and of sexual orientation. A central question for feminist ethicists is *how* one's positions within

these and other oppressive systems—especially the kinds and degrees of power, authority, privilege, and entitlement that these positions afford one in various particular contexts—shape both one's moral character and one's moral responsibilities. This focus on power relationships and on their effects on moral life means that the boundary between feminist ethics and feminist social and political philosophy is often a fluid one.

CARE, RELATIONSHIP, AND WOMEN'S LABOR

One vital step toward remedying any masculinist bias in moral theory is to investigate and understand women's points of view. The research of educational psychologist Carol Gilligan (1982) was an important early inspiration for feminist ethicists' efforts to take seriously and learn from women's moral perception and reasoning. Based on her research interviewing males and females about moral dilemmas (both real and imagined), Gilligan argued that there are two distinct moral perspectives (or "voices") loosely associated with men and women respectively.

The justice perspective begins with a conception of persons as separate individuals who need moral rules to govern their interactions with each other, and in particular to safeguard a realm of autonomy within which each individual may act and make decisions without undue interference from others. Moral decision making is most centrally a matter of impartially adjudicating conflicts between individual rights and interests, and of seeing to it that one's actions conform to certain universal rules of conduct. According to Gilligan, the justice perspective is more prominent in the moral voices of males than in those of females.

The care perspective, in contrast, begins with a conception of persons as embedded in social relationships in which they bear different and sometimes conflicting responsibilities to one another. Here, the priority is on creating and preserving connections and on avoiding and ending suffering. One's primary responsibility is to respond to the needs of individuals located in concrete, particular situations, often by strengthening the relationships that support those individuals. Gilligan found that the care perspective is expressed most prominently and most frequently by women and girls, and urged that theorists pay due attention and respect to this perspective, rather than seeing it as an inferior and immature form of moral reasoning.

Since the 1990s, an early tendency to identify feminist ethics with care ethics has receded as feminist ethics itself grows more diverse and wide-ranging. Nonetheless, some of the themes that Gilligan highlighted continue to occupy a central place in the thinking of many feminist ethicists. One such theme is what is sometimes called a relational conception of the person. Annette Baier (1985) usefully captures this concept by describing persons as essentially "second persons"; that is, beings whose subjectivities are formed and maintained in and through connections with others. Feminist ethicists typically focus on persons as participants in relationships both public and intimate, as inhabitants and co-constructors of social roles and identities. Many have sought to reconceive and expand vital moral concepts such as autonomy, rights, respect, responsibility, and equality in ways that centrally incorporate such a relational understanding of persons.

When theorizing begins with a vision of persons as inextricably located in and shaped by relationships, the fact that many of those relationships are oppressive ones naturally comes to play an important role in the theorizing. Feminist ethicists have emphasized not only how people ought ideally to behave, but also the personal, social, and political conditions that would enable people to develop their characters and behave responsibly—and in particular, to how relations of oppression can cripple and distort the moral capacities of persons (both those who suffer from oppression and those who benefit from it). Identifying and possibly repairing the moral damage of oppression has been an important theme in feminist ethics; in such work, a key challenge is always to distinguish the important (and often neglected) values and insights of oppressed people from the moral damage of oppression itself.

The centrality of relationship, the importance of valuing women's perspectives, and the question of oppression's moral damage all converge in feminist ethicists' discussions of the labor that has most centrally characterized women's experience over the centuries. This might be called the work of relationship itself—of caring and nurturance, of tending to others' intimate emotional and physical needs (including for love, food, cleanliness, clothing, and the like) both inside the home as wives and mothers and outside of the home in professions such as nursing and teaching. Thus, in feminist ethics, due respect for the role of emotion in moral reasoning has been supplemented by attention to emotional labor: its importance to human well-being, its invisibility in some received ethical theories, and its disproportionate and often exploitative allocation to women (Bartky 1990, Calhoun 1992).

Sara Ruddick's influential *Maternal Thinking* (1989) attempted to reclaim the work of mothers as involving

particular forms of moral reasoning that are vital not only to the work of raising children, but to efforts to create and sustain a just and livable world. Virginia Held argued in *Feminist Morality* (1993) that the relationship between mother (or "mothering person") and child—rather than contractual relations or market transactions—should be considered the central or paradigmatic human experience and the basis for a feminist account of morality. While other feminists have been more wary of taking mothering either as paradigmatic of women's experience or as a model for morality itself, most feminist ethicists grant that having primary responsibility for the intimate care and nurturing of children seems likely to shape women's moral perspectives in deep and pervasive ways that are worthy of philosophical attention.

Peta Bowden (1997) argues against attempts to formulate universal principles to govern caring. Instead, care must be understood and elaborated through detailed attention to examples; she discusses motherhood, nursing, friendship, and citizenship as substantively different caring practices. In contrast to the canon's highly idealized emphasis on relations among persons considered as equal in freedom and power, another area of feminist analysis in care-based ethics is the dependencies that accompany certain stages and conditions of life, including childhood, illness, old age, and various physical and mental disabilities. Feminist discussions of such dependencies (such as that of Kittay 1999) focus attention on the ineluctable facts of human vulnerability and interdependence, as well as on inequalities both between caregivers (or "dependency workers") and those for whom they care, and between caregivers and non-caregivers in various communities.

Feminist ethicists have also drawn on women's experiences challenging, or at least moving outside of, traditional feminine roles as nurturers of children and men. Important forms of ethical insight and practice emerge from alternative or resistant female lives, particularly from the bonding of women with each other in friendship (Friedman 1993) and/or love (Card 1995, Calhoun 2002) and from feminist networks and communities. Work in this vein tends to ask what values, virtues, and capacities are necessary for women to maintain their own well-being under patriarchy as well as to challenge and resist oppressive structures. While Marilyn Frye (1983, 1992) would likely resist a characterization of her work as part of ethics, her work on vital concepts such as arrogance, loving perception, whiteness and racism, oppression, humanism, and lesbianism has been enormously influential for many who are working to articulate resistant feminist moral values and practices.

ISSUES, CONCEPTS, AND METHODOLOGIES

Feminist ethicists have extensively discussed concrete normative issues that are clearly gender-related: abortion, rape and sexual consent, sexual harassment, marriage, pornography and hate speech, prostitution, surrogate or contracted motherhood, reproductive technologies, homophobia and heterosexism, domestic labor and intrafamilial justice, and welfare policy, to name only a few. These discussions have often focused not only on whether or not the practice in question is morally legitimate but also—for instance, in the case of rape and other forms of misogynist violence—on exposing its role in maintaining women's political subordination and in forming women's and men's moral subjectivities. They have also brought a feminist perspective to bear on other concepts and attitudes that are less obviously gender-related, but for which an understanding of gender and power is illuminating. These include gratitude (Card 1996), shame (Bartky 1990), trust (Baier 1994), paternalism (Sherwin 1992), self-respect (Dillon 1997), guilt (Bartky 2002), and evil (Card 2002).

The feminist ethics lexicon also includes novel concepts developed specifically as part of the project of analyzing and finding ways to move beyond oppression and privilege—for example, María Lugones's (1987) concept of "world-traveling," which she recommends to feminists and others who seek to replace arrogance with love, identification, and loyalty in their relations to women who occupy different social "worlds." Finally, feminist ethicists have developed ambitious new conceptions of morality's nature, purposes, and sources of authority, such as Margaret Urban Walker's (1998) "expressive-collaborative" model of morality (as distinct from the "theoretical-juridical" model that she thinks more typical of mainstream moral theory).

Whatever the specific topic at hand, certain methodological approaches and themes cut across much of what goes under the rubric of feminist ethics. Feminist ethics is typically characterized by a resistance to excessive idealizing in moral theory, especially to idealizing that obscures the pervasive relationships of dependence and of unequal freedom and power that moral life calls upon us to navigate responsibly. As Claudia Card puts it, feminist ethics generally errs on the side of "peeling back rather than donning veils of ignorance" (1991, p. 25). Relatedly, many (though certainly not all) feminist ethicists are wary of

attempts to formulate universal and highly articulable rules or principles in ethical theory, tending instead to draw more limited conclusions based on detailed analyses of particular socially located experiences. Particularly since the 1990s, feminist ethics has developed a fairly consistent focus on the *practices* of morality, on how moral concepts are actually used and deployed in various contexts: what we do with rights, how we take and assign responsibility, for what and to whom we hold ourselves and others accountable.

Not surprisingly, then, many feminist ethicists emphasize the necessity of ongoing real (rather than hypothetical or idealized) conversation and dialogue as important to revealing, justifying, and/or challenging people's moral practices and agreements. What matters is not only what is said, but who is thought to be entitled to say it: As Margaret Urban Walker puts it, "Feminist ethics pursues questions about *authority, credibility*, and *representation* in moral life and in the practice of moral theorizing itself" (1998, p. 54).

Some longstanding themes in feminist ethics continue to be refined and taken in new directions. Some feminist ethicists, like Joan Tronto (1993), have continued to develop and refine a care-based approach. In *Moral Boundaries*, Tronto urges that we renegotiate the boundary between morality and politics and endorse care not as a form of "women's morality," but rather as a political virtue that can aid in redistributing power and transforming the public sphere. Several themes—a relational conception of persons, the need to repair oppression's moral damage and to articulate practical modes of resistance—combine in Hilde Lindemann Nelson's (2001) discussion of identities as narratively constructed. Nelson argues for the importance of oppressed people developing "counterstories" that can resist and ultimately replace the damaging and undermining stories told about them by dominant groups. Such "narrative repair" is especially vital, in Nelson's view, because who one takes oneself to be, and who others take one to be, affects how freely one can act. Perhaps reflecting the maturation of the field itself, as well as its longstanding focus on persons as embodied beings proceeding through a life cycle, some feminist ethicists (Walker 1999, Bartky 2002) have turned their attention to aging—particularly to the strengths, natural and humanly arranged vulnerabilities, and specific forms of inequality that confront elderly women.

Finally, a global focus in feminist ethics, already well underway in the work of such feminists as Uma Narayan (1997) and Martha Nussbaum (2000), also finds expression in Alison Jaggar's (1998) attempt to enlarge the possibilities for egalitarian and inclusive global feminist dialogue. In discussing the challenges facing feminists who would respectfully communicate and cooperate with each other across vast global divisions of power, resources, and accorded authority, Jaggar exemplifies and develops several ongoing themes in feminist ethics. Among these are a suspicion of idealization (in Jaggar's case, of "romanticizing discursive utopias"), a corrective emphasis on actual dialogue and on questions of authority and silencing therein, and a relentless attention to the effects of power dynamics on women variously located in multiple matrices of domination.

See also Applied Ethics; Aristotle; Baier, Annette; Beauvoir, Simone de; Ethics; Feminist Legal Theory; Feminist Philosophy; Feminist Social and Political Philosophy; Hume, David; Kant, Immanuel; Murdoch, Iris; Nietzsche, Friedrich; Nussbaum, Martha; Plato; Weil, Simone.

Bibliography

Baier, Annette. *Moral Prejudices*. Cambridge, MA: Harvard University Press, 1994.

Baier, Annette. *Postures of the Mind*. Minneapolis: University of Minnesota Press, 1985.

Bartky, Sandra. *Femininity and Domination: Studies in the Phenomenology of Oppression*. New York: Routledge, 1990.

Bartky, Sandra. *Sympathy and Solidarity and Other Essays*. Lanham, MD: Rowman and Littlefield, 2002.

Bowden, Peta. *Caring: Gender-Sensitive Ethics*. New York: Routledge, 1997.

Calhoun, Cheshire. "Emotional Work." In *Explorations in Feminist Ethics: Theory and Practice*, edited by Eve Browning Cole and Susan Coultrap-McQuin. Bloomington: Indiana University Press, 1992.

Calhoun, Cheshire. *Feminism, the Family, and the Politics of the Closet: Lesbian and Gay Displacement*. Oxford: Oxford University Press, 2002.

Card, Claudia. *The Atrocity Paradigm: A Theory of Evil*. Oxford: Oxford University Press, 2002.

Card, Claudia, ed. *Feminist Ethics*. Lawrence: University of Kansas Press, 1991.

Card, Claudia. *Lesbian Choices*. New York: Columbia University Press, 1995.

DesAutels, Peggy, and Joanne Waugh, eds. *Feminists Doing Ethics*. Lanham, MD: Rowman and Littlefield, 2001.

Dillon, Robin. "Self-Respect: Moral, Emotional, Political." *Ethics* 107 (1997): 226–249.

Friedman, Marilyn. *What are Friends For? Feminist Perspectives on Personal Relationships and Moral Theory*. Ithaca, NY: Cornell University Press, 1993.

Frye, Marilyn. *The Politics of Reality: Essays in Feminist Theory*. Freedom, CA: Crossing Press, 1983.

Frye, Marilyn. *Willful Virgin: Essays in Feminism*. Freedom, CA: Crossing Press, 1992.

Gilligan, Carol. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press, 1982.
Held, Virginia. *Feminist Morality: Transforming Culture, Society, and Politics*. Chicago: University of Chicago Press, 1993.
Held, Virginia, ed. *Justice and Care: Essential Readings in Feminist Ethics*. Boulder, CO: Westview, 1995.
Jaggar, Alison. "Globalizing Feminist Ethics." *Hypatia* 13 (2) (1998): 7–31.
Kittay, Eva Feder. *Love's Labor: Essays on Women, Equality, and Dependency*. New York: Routledge, 1999.
Lugones, Maria. "Playfulness, 'World'-Travelling, and Loving Perception." *Hypatia* 2 (1987): 3–19.
Mackenzie, Catriona, and Natalie Stoljar, eds. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. New York: Oxford University Press, 2000.
May, Larry. *Masculinity and Morality*. Ithaca, NY: Cornell University Press, 1998.
Narayan, Uma. *Dislocating Cultures: Identities, Traditions, and Third World Feminism*. New York: Routledge, 1997.
Nelson, Hilde Lindemann. *Damaged Identities, Narrative Repair*. Ithaca, NY: Cornell University Press, 2001.
Nussbaum, Martha. *Women and Human Development: The Capabilities Approach*. Cambridge, U.K.: Cambridge University Press, 2000.
Ruddick, Sara. *Maternal Thinking: Toward a Politics of Peace*. Boston: Beacon Press, 1989.
Sherwin, Susan. *No Longer Patient: Feminist Ethics and Health Care*. Philadelphia: Temple University Press, 1992.
Tronto, Joan. *Moral Boundaries: A Political Argument for an Ethic of Care*. New York: Routledge, 1993.
Walker, Margaret Urban. *Moral Understandings: A Feminist Study in Ethics*. New York: Routledge, 1998.
Walker, Margaret Urban. *Mother Time: Women, Aging, and Ethics*. Lanham, MD: Rowman and Littlefield, 1999.

Rebecca Whisnant (2005)

FEMINIST LEGAL THEORY

Feminist legal theory is the study of the philosophical foundations of law and justice; informed by women's experiences, its goal is to transform the legal system and the understanding of it to improve the quality of jurisprudence and women's lives. Feminists working in law share the convictions that the historical and continuing exclusions of women from the law's protective domain have injured women and that the exclusion of women from the study of law has limited both the understanding of law and it ethical compass. Feminists have accordingly sought to transform the rules and principles governing particular areas of law—torts, criminal law, constitutional law—so as to make them more responsive to women's needs and more reflective of women's perspectives. Feminist legal theorists examine the consequences—both for women and for jurisprudence—of the exclusion of women's input into our shared understanding of the law's philosophical foundations. Toward that end feminists have examined competing philosophical understandings of the nature of law, have attempted to show how they fail to reflect women's perspectives, and have attempted in each case to reinvigorate them by centralizing rather than marginalizing women's experiences.

Some feminist legal theorists—sometimes called liberal feminist scholars—argue that women's lives will be most improved by simply extending to women what are widely regarded as two of the central promises of law in a liberal regime: first, the promise of "formal equality," the idea that the state's legal institutions will "treat like causes alike"; and second, the promise to each individual of a wide a sphere of individual autonomy. Women, liberal feminists argue, are "like men" in all the ways that should matter to the state and accordingly should be treated, wherever possible, in precisely the same way as men by the law. Women and men are the same in their abilities: Women, like men, can engage in the professions and trades, wage war, fairly serve on juries, administer estates, and vote responsibly, and the law must accordingly not discriminate on the basis of a false claim of difference and must also forbid discrimination against women in the private sector on the basis of such false claims (Williams 1984).

Similarly, women and men are the same in their needs: Women, like men, need protection against violence, meaningful work and civic participation, and, most important, the freedom to develop their individual life plans. The law should therefore extend to women the same protection against private violence and the same sphere of autonomy it extends to men (McClain 1992). By pursuing the logic of these applications of fundamental liberal principles to the law's treatment of women, liberal feminist legal theorists have contributed to widespread changes in the relations of women, men, and the state, ranging from the institution of bans on private and state discrimination on the basis of gender to the expansion of women's reproductive freedom and choices so as to maximize their social and political autonomy.

As critics of liberal feminism have pointed out, however, women are not "like men" in all ways, and as a consequence a rigid application of liberal premises to the sometimes distinctive situation of women will often backfire. Where women are unlike men, the blanket insistence on equal treatment will sometimes impoverish actual women, albeit toward the admirable end of a gender-blind utopian society (Becker 1987). Equal distri-

bution of property at the time of divorce, for example, will impoverish the majority of divorcing women who have less earning potential than their husbands. The equal refusal of an employer to grant maternity or parental leave upon the birth of a child will disproportionately hurt female workers, who, because of their greater biological role in the process of reproduction, will need more time out of the workplace than will men if they are to enjoy the same rights as men to be both workers and parents (Littleton 1987). The refusal of the state to extend the protection of social security to career homemakers treats women and men similarly but disproportionately harms women because women are disproportionately represented in the ranks of unpaid domestic labor.

At the professional level, tenure policies and partnership tracks, equally applied, hurt women more than men, because of the differing reproductive cycles of the two sexes. To take an extreme and only partly hypothetical example, a state that failed altogether to criminalize rape would on one level treat men and women similarly and thereby abide by the liberal mandate of equal treatment, but women would obviously be disproportionately harmed by such a regime. In all of these cases, the even-handed application of legal rules harms women because of the very real differences in women and men's economic, political, and social conditions.

Partly in response to the perceived theoretical and practical inadequacies of liberal feminist legal theory and partly as a response to work in other fields on the differences between men and women's psychological lives, a number of feminists in legal studies, sometimes called difference or cultural feminists, have sought to place at the center of inquiry not the many ways in which women and men are the same or similar but, rather the ways in which women and men are different. This focus on difference has in turn led to three promising areas of inquiry. First, difference feminists in legal studies have put forward a modified or quasi-liberal theory of equality sometimes called an acceptance theory (Littleton 1987). According to this view, the state's moral (and constitutional) obligation to treat citizens equally entails the state's obligation not only to provide equal treatment of the sexes wherever the sexes are similarly situated but also to provide different treatment wherever necessary to ensure an equal acceptance of differences, so that those differences, whatever their origin, do not cause women harm. Because women (but not men) get pregnant, bear children, and lactate, for example, the law must fashion rules of employment and civic engagement that will facilitate the acceptance of those differences in the public and economic spheres, whether or not that in turn requires different or similar treatment of the sexes in various legal regimes. Since women engage in more unpaid domestic labor, the liberal mandate of equality demands that family law, divorce law, and social security law should develop in ways that will render that difference harmless.

Other difference feminists have put forward a related critique of liberalism itself, sometimes called the "dependency critique" (Kittay 1999, Fineman 1995). The conception of human nature on which liberal norms of justice and equality (and the vast bodies of law they imply) typically rests is that citizens of a liberal polity should be treated as fundamentally independent and autonomous. But this conception of our nature is transparently and badly flawed: All human beings are dependent upon caregivers for their very lives for a good part of their early childhood and continue to require care throughout adolescence so as to become the autonomous citizens, independent entrepreneurs, moral agents, and free individuals so valued by various strands of liberalism and so vigorously protected by our fundamental, constitutional law. Further, all of us require care when we are elderly, likewise undercutting the dominant understanding of the independent individual at the heart of liberal theory. Almost all women and many men spend a very high percentage of their adult lives providing this care, in private and for no compensation when done within the family, or for very low wages when done through labor markets.

The disproportionately greater amount of caregiving labor done by women throughout history tends to be invisible within a liberalism that steadfastly insists on individual autonomy; hence, legal regimes that depend upon or aspire to those liberal values are often irrelevant or harmful to women and to the children and elders that depend upon them. The result in practical terms is often the impoverishment of women and dependents; the jurisprudential and philosophical result is a set of moral ideals for law and legal justice that badly undercuts the aspirations and needs of much of the world's populations (West 1996). A liberalism enriched with a feminist regard for the centrality of caregiving labor, for the moral and ethical perspectives such labor both demands and partly produces, and a fuller understanding of the dependencies and interdependencies of our social and biological lives would enhance women's well-being and the strength of both legal and political philosophy (McClain 1992).

Difference feminists have tried to explicate the distinctive harms women suffer that have little or no corre-

late in men's lives, on the assumption that by virtue of their difference, among other things, the harms that women suffer often go unnoticed as well as unaddressed (West 1996). Women suffer from sexual assault, sexual harassment, and sexual violence in greater numbers and in different ways than men do. Women suffer unwanted and nonconsensual pregnancies; men do not. Whatever the reason, women world-over are more engaged in childraising, and consequently are more harmed than men by the loss of children in custody disputes and are more vulnerable than men to the threat of such loss, which significantly weakens their economic bargaining position both in the family and at the point of divorce. If women are to enjoy legal protection against these and other gender-specific harms, the laws governing the social interactions that occasion these harms must be responsive to the existence and the different nature of the harms that women differentially and distinctively experience.

Radical feminist legal theory, sometimes called dominance feminism, is also an attempt to fashion a feminist theory of law that avoids the pitfalls of liberal feminist legal theory, but it does so in a different way. The central question for feminists working in law, according to radical feminist theorists, is not whether women and men are fundamentally alike or different but how the state might foster the greater empowerment of women. Women are unlike men in one significant respect: women as a group lack power (MacKinnon 1989). Liberal feminists are wrong to downplay or disregard that difference, and difference feminists are wrong to focus on any other differences. A focus on the differential treatment of women by the state, whether with the liberal feminists' aim of eradicating those differences, expanding upon them, as difference feminists wish, will be at best distracting. Disempowerment, not discrimination, and not difference, is the source of the problem, and patriarchy, not law, is the source of women's disempowerment. Law reflects patriarchal influences, but patriarchy also exists independent of law. Consequently, law can be and should be employed to end it.

Loosely reflecting the logic of critical legal scholars' Gramscian analysis of the relation of law and market capitalism, radical feminists have sought to highlight the nonlegal ways in which patriarchal power is created and reinforced in culture and then legitimated by legal rules and institutions. Women are disempowered, for example, by the violence done them through rape, sexual harassment, and street hassling as well as other forms of sexual assault. That disempowerment is then underscored through the distorting messages and the attacks on women's self-esteem occasioned by pornography, the culture of romance, and other societal influences, all of which aim to render that disempowerment in some sense voluntary and all of which render problematic the liberal feminist insistence on expanding individual autonomy as a means for improving women's well-being. Absent feminist intervention, the law's role in this process of disempowerment and cooptation is largely to legitimate those harms: The constitutional doctrine of privacy, laws governing and only partially regulating rape and domestic violence, and the constitutional protection accorded to even extremely damaging assaultive speech all trivialize or render invisible the harms women sustain and reinforce the tendencies that cause them. Law does not itself cause these harms, but it contributes to a culture that tolerates them.

There is, however, nothing necessary about the handmaidenlike role of law in sustaining patriarchy; it only reflects current distributions of sexual and gendered power. Arguably, all of these forms of patriarchal power, and certainly those employing violence, can and should be prohibited by law. The law legitimates a good bit of the disempowerment occasioned by rape by underregulating it, but that can be changed: Rape laws can be expanded, and enforcement of those laws strengthened; to do both would go a long way toward undermining patriarchy. The goal of radical feminist theory is to employ the law in precisely this utterly conventional way toward the unconventional goal of first prohibiting and then eradicating the violence that sustains a patriarchal cultural regime.

Finally, a number of feminists engaged in legal theory have sought to appropriate the tools of postmodern analysis to free liberal, difference, and radical feminist legal theory from the presumed dangers of their essentialist premises. Two distinct projects have emerged from this effort, one critical and one reconstructive. First, postmodernists have joined with African-American, lesbian, and other arguably marginalized feminist legal scholars in an attempt to criticize the consciously or unconsciously racist or heterosexist assumptions in feminist legal theory, thus laying the groundwork for the emergence of a feminist jurisprudence strengthened by its recognition of women's racial, sexual, ethnic, and cultural differences (Harris1990). Critical-race feminist legal theorists have contributed the most to this project. Theoretical and empirical scholarship has accordingly shown the ways in which, for example, feminist writing on rape and rape law has failed to attend to the experiences of African American women, whose understanding of rape is informed by a history of the use of rape law as an instru-

ment of terror by the white state and by rape as an instrument of terror by men. These critics also point to the ways in which feminist writing on difference, care, and caregiving has failed to attend to the extent to which African American women have provided such care to whites for no or little pay. Likewise, critical-race theorists and writers in the civil rights traditions fail to attend to the different experiences of women and men in communities of color: for example, the communal censoring of African American women who try to theorize or even describe experiences of domestic violence or sexual violence in communities of color. Feminist race scholars writing in law have urged the adoption of the perspective of persons at the "intersection" of various "axes of subordination" to best understand the ways in which these modes of social interaction injure those most vulnerable to multiple forms of marginalization (Crenshaw 1991).

Second, postmodernist feminists have joined with cultural critics and "queer theorists" from other disciplines in an attempt to highlight the ways in which perceived differences between the genders and between sexual orientations are themselves socially constructed rather than biologically mandated (J. Williams 1989, Halley 2002). The aim has been partly to free feminism from false and essentialist stories or metanarratives of women's disempowerment and partly to redirect feminist legal reforms. Postmodern feminists, for example, have been attempting to redirect the law of sexual harassment, largely a product of radical feminism theorizing, away from its current focus on sexuality and toward a more pluralistic understanding of the various harms, whether sexualized or not, that women and men suffer in the workplace (Shultz 1998). This is in part in response to the postmodern complaint that radical feminism and hence sexual harassment law have wrongly relied on a grand metanarrative of women's sexual disempowerment by men and in part a response to a concern that sexual-harassment law may encourage or rest on homophobic responses to what might be harmless socio-sexual gestures in workplaces (Halley 2002).

Both projects—the enrichment of traditional feminist theory with the perspectives of African American and other ethnic minority women, and the challenge to the narratives of female sexual disempowerment at the heart of sexual harassment law and radical feminism—both resonate with long-standing feminist (as well as postmodernist) goals: the first in its insistence on respecting and honoring the voices of outsiders, including those women who find themselves "outside" mainstream feminist discourse, and the second in its insistence on locating within culture, rather than nature, the causes of women's oppression and the key to ending it.

See also Feminist Ethics; Feminist Philosophy; Feminist Social and Political Philosophy; Gramsci, Antonio; Justice; Philosophy of Law.

Bibliography

Becker, Mary. "Prince Charming: Abstract Equality." *Supreme Court Review* 5 (1987): 201–247.

Crenshaw, Kimberlé Williams. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color." *Stanford Law Review* 43 (1991): 1241–1299.

Fineman, Martha Albertson. *The Neutered Mother, The Sexual Family and Other Twentieth Century Tragedies*. NY: Routledge, 1995.

Fineman, Martha Albertson. *The Autonomy Myth: A Theory of Dependency*. NY: New Press: 2004.

Halley, Janet. "Sexuality Harassment." In *Left Legalism/Left Critique*, edited by Wendy Brown and Janet Halley. Durham: Duke University Press: 2002.

Harris, Angela. "Race and Essentialism in Feminist Legal Theory." *Stanford Law Review* 42 (1990): 581–616.

Kittay, Eva. *Love's Labor: Essays on Women, Equality, and Dependency*. NY: Routledge, 1998.

Littleton, Christine. "Reconstructing Sexual Equality." *California Law Review* (1987): 1279–1337.

MacKinnon, Catherine. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press, 1989.

McClain, Linda. "Atomistic Man Revisited: Liberalism, Connection, and Feminist Jurisprudence." *Southern California Law Review* (1992): 1171–1264.

McClain, Linda. "Care as a Public Value: Linking Responsibility, Resources, and Republicanism." *Chicago-Kent Law Review* (2001): 1673–1731.

Shultz, Vicki. "Reconceptualizing Sexual Harassment." *Yale Law Journal* 107 (1998): 1683–1805.

West, Robin. *Caring for Justice* NY: NYU Press, 1996.

Williams, Joan. "Deconstructing Gender." *Michigan Law Review* 1989: 797–845.

Williams, W. "Notes From A First Generation." *University of Chicago Legal Forum* 1989: 99–113.

Robin L. West (2005)

FEMINIST METAPHYSICS

Metaphysics seems to be one of the least relevant, most foreign, and inhospitable disciplines of philosophy in relation to feminist projects and concerns. Traditional metaphysicians have tried to answer questions about the basic structure of reality, about what kinds of beings exist, about the nature of time and causation, and they have probed difficulties like free will and determinism, the

nature of universals and particulars and the like. None of these issues seems directly pertinent to feminism, and their abstract formulation and universalist perspective strike some feminists as deeply suspect. Nonetheless feminist metaphysics has emerged as a distinct and lively field in feminist theory. And there are important connections between feminist work on certain metaphysical issues and mainstream metaphysics.

Feminist metaphysics revolves around three core issues: essentialism and anti-essentialism about sex/gender, theories of the self or the subject, and realism versus social constructionism (a version of the realism/anti-realism controversy in mainstream philosophy). Each of these issues is central to Simone de Beauvoir's *The Second Sex*, which is rightly seen as the primary intellectual source for twentieth-century developments in continental and analytic feminist theory. Beauvoir oriented her pioneering work toward ontology and essentialism by defining woman as the Other (in relation to man). At the same time she sketched out the first detailed and comprehensive social constructionist account of gender. And she was centrally concerned to retrieve the possibility of subjectivity, agency, and transcendence for women.

Beauvoir's legacy has been developed in two major directions. In broad strokes continental feminist theory is anti-essentialist about sex and gender, and skeptical about the unity and coherence of the self or subject. Continental feminist theory tends to derive both anti-essentialism and anti-realism about sex/gender from a social constructionist view of sex/gender. In contrast analytic feminist theory tends to distinguish among these positions, holding, for example, that socially constructed categories and entities are real, and perhaps even constituted by essential properties. Analytic feminist theory is more hospitable to essentialism about sex and gender, and open to the possibility of non-androcentric theories of the self. These are generalizations, however, as we can see by considering the fact that Luce Irigaray, a pioneer of continental feminist theory, has developed an essentialist theory of sexual difference.

Feminist preoccupation with the questions of essentialism and the nature of the self rather than other metaphysical topics is neither coincidental nor arbitrary. Both of these issues are directly relevant to feminist politics because of their implications regarding the possibility of individual agency and effective shared activity toward political change. For example, Naomi Zack (1997, 2005) points out the consequences of anti-essentialism about gender for collective agency on behalf of women. If women share nothing in common as a group, then on what basis can they forge a group identity, and on what basis can they find common goals? Other feminists, like Diana Meyers (1997, 2002), are troubled by the claim that there is no self or subject because of the implications of that position for the possibility of individual resistance to patriarchal norms, and for collective political agency. Similarly the feminist debate concerning social constructionist and realism/anti-realism is intended to reveal the arbitrariness and contingency of oppressive social and political structures in order to allow for the possibility of political change and an end to oppression. This entry explores the development of feminist metaphysical thinking about sex/gender essentialism, the self or the subject, and social constructionism and realism.

ESSENTIALISM AND ANTI-ESSENTIALISM IN FEMINIST THEORY

The feminist discussion of essentialism usually begins with a distinction between sex differences, which are the biological markers that distinguish females from males, and gender differences, which are the cultural or psychological features that distinguish women from men. Some feminists question the distinction between biology (or nature) and culture underlying the sex/gender distinction. They argue that there is cultural intervention in the production of two sexes from a more complex biological reality. In making this argument, they reject an essentialist account of sex because there are no biological features that demarcate human beings into just two kinds that correspond to female and male. See the discussion in Anne Fausto-Sterling (2000), and the essentialist account of sex differences by Linda Alcoff (2005). In addition, Sally Haslanger (2000) has argued that a major project of feminist metaphysics is the unmasking of putatively natural categories or properties as social.

A similar argument is made against gender essentialism; namely, that there are no biological, psychological, or cultural properties that are common to all women and not shared by any men. Let's call this the commonality problem. Moreover, women of color (and others) have pointed out that the psychological and cultural properties that some feminists propose as essential to all women in fact exclude many women. Let's call this the exclusion problem. Elizabeth Spelman (1998) and bell hooks (1981) made important contributions in articulating these problems. The doctrine of intersectionality was developed by Kimberle Crenshaw (1991) to respond to both the commonality problem and the exclusion problem. Intersectionality is the idea that feminists need to attend to the multiplicity of identities that can and do

characterize individuals (race, class, and sexual orientation) in order to avoid the problems of exclusion and commonality. However, the concept of intersectionality is problematic to the extent that it fractures the unity of women, and leads to skepticism concerning whether any useful program for political change can reflect the interests of a heterogeneous collection of individuals.

Feminists have responded to anti-essentialist arguments by developing approaches to essentialism that respect the problems of commonality and exclusion without fracturing the unity of women. There are two basic approaches. A materialist approach to gender essentialism, developed in different ways by Haslanger and Monique Wittig (1997), among others, begins with the body, and the way that bodies are hierarchically ordered in and by patriarchal (and racist, ageist) societies. Gender is a material, embodied state and bodies are classified by societies into hierarchical relations. Being gendered is a relational property because gender categorization is dependent upon how bodies are perceived by others rather than upon the possession of any intrinsic biological or psychological property. Being gendered is also a political property in the sense that it carries with it a position in a hierarchical social structure. The materialist approach meshes with the intersectionality perspective because it allows that bodies can be classified in multiple ways according to overlapping social hierarchies; for example, racialized bodies that are men occupy a different social niche from racialized bodies that are women. Able-bodied women occupy a different position from disabled women and so on. On this approach the identities of being a woman and being a man necessarily have positions in a hierarchical grid of social power relations; if patriarchy did not exist then neither would women and men.

Alternatively Natalie Stoljar (1995) makes the case for understanding woman as a cluster concept rather than an Aristotelian universal. In a related development Naomi Zack (1997) argues that that being a woman is a relational, disjunctive property shared by all women. Like the materialist approach these accounts emphasize the features common to all women, but select features that are sensitive to the problem of exclusion. Unlike the materialist approach to gender essentialism these views do not make oppression intrinsic to being a woman. They also do not provide a conceptual grid for other identities like race, or sexual orientation as the materialist approach does. Tracing the similarities and dissimilarities between gender and other social categories and identities, like race and sexual orientation, is a major theme in feminist writing on essentialism and anti-essentialism. Although the question of gender essentialism remains contested within feminist theory, dogmatic anti-essentialism is no longer a criterion for adequate feminist theorizing.

Finally, some philosophers frame the discussion of gender essentialism in terms different from those we have been considering. Rather than try to determine whether or not there are any properties common to all women, we might wonder whether or not being gendered is essential to the identity of individual women and men. Essentialism in this sense is not about kind membership but rather concerns the issue of whether or not any of an individual's properties constitutes her as the individual she is, and if so, whether or not being a woman is one of an individual's constitutive properties. In different ways, Anthony Appiah (1990) and Charlotte Witt (1992, 1995) explore essentialist theories of gender by focusing on the relationship between an individual's identity and his or her gender rather than the question of what all women or all men have in common.

As mentioned in the introductory text of this entry, one reason for the persistence of the issue of essentialism in feminist theory is the political requirement that women be identifiable as a group with common interests, and who suffer shared injustices. Group identity is politically necessary; mere strategic essentialism does not seem to be sufficient as a basis for political change. For similar reasons the issue of the subject or the self is central to feminist metaphysical thinking. Despite the shortcomings of traditional accounts of subjectivity, it is hard to conceptualize a politically adequate view of agency without some account of the subject who acts.

FEMINIST ACCOUNTS OF SUBJECTIVITY

Traditionally to be an agent one must be a self or a subject, and not a thing or an object that is acted upon. But feminists have catalogued serious deficiencies with the way in which traditional philosophers have described the self. These deficiencies include the tendency to identify the self or subject with reason in contrast with the emotions and the body; the tendency to associate agency with autonomous individuals rather than connected, relational selves, and the characterization of the subject as unified and coherent. The last criticism is the most radical as it rejects the very notion of a consistent self or subject rather than pointing out deficiencies with traditional characterizations of unified subjects.

The rejection of the unified and coherent subject or self is related to one strand of anti-essentialist argument

as we see in Judith Butler's work. Not only does Butler (1990) reject all forms of sex/gender essentialism, but also she does so as part of a rejection of the metaphysics of substance. And the rejection of the metaphysics of substance, the denial that individual, persisting beings exist entails the rejection of subjects in so far as they are characterized as unified individuals that persist through time. Some feminists find the dissolution of stable subjectivities liberating because of the possibilities for innovation, creativity, and performance that this view endorses. Other feminists find the rejections of stable subjects inadequate to the requirements of political resistance and change. Recall that the possibility of agency is based upon the existence of subjects who are agents. Agents can resist patriarchal norms, and can band together to effect political change.

However even those feminists who accept the importance of unified and coherent subjects criticize traditional notions of the self. For example, Susan Babbitt (1996) is critical of the philosophical tradition that centers subjectivity on reason, and defines reason as exclusive of emotions, imagination, perception, and other faculties associated with the body. And theories of the subject, which are mentalistic, also have come under feminist scrutiny. In response feminists like Moira Gatens (1996) have worked to define a bodily notion of subjectivity, which is more adequate to feminist understanding of the importance of embodiment in explaining human agency. Feminists have also developed a relational theory of the self, which interprets agents as constituted by their relations to others, and as embedded in concrete historical and cultural horizons.

The idea of the subject as relationally constituted and historically embedded is more adequate to feminist projects than the traditional idea of subjectivity. However, it is also problematic in relation to the idea of autonomy, which is an important constituent of many theories of moral and political agency. Moral and political subjects or agents act autonomously in some sense of the term. There appears to be tension between the requirement of autonomy on the one hand, and the feminist notion of a relational and embedded subject. If subjects are formed in and by particular cultures, and if their being is determined by their relations to other subjects and also in and by their relations to cultural and historical institutions, then in what sense do they choose and act autonomously?

Feminists like Diana Meyers (1997, 2002) have worked to specify criteria for a notion of autonomy that both recognizes the concrete causal formations of subjectivity, and carves out a reasonable zone for autonomous decision making. In this way, feminists have absorbed the lessons of contingency from social constructionism without giving up the important ethical and political norm of autonomy. Other feminists like Marilyn Frye (1983, 1989, 1996, 2000, 2005) question whether the ideas of individual choice, individual autonomy and individual selves are the central notions that feminists need to understand the structures of patriarchy. They argue against the focus on individual subjectivity and choice not because there are no individual subjects but because focus on the individual subject and her choices obscures the horizon of oppression against which and within which choice operates. It reflects a political commitment to individualism, which does not provide an adequate framework for feminist politics.

GENDER, SOCIAL CONSTRUCTIONISM, AND REALISM

Most feminists reject a biological, deterministic conception of gender. Instead they see gender as constituted and defined by social norms, practices and institutions. Since social norms, practices and institutions vary in different cultures, and also differ in the same culture at different historical periods, it seems to follow that gender is indeterminate and variable rather than fixed and stable. As we have seen some feminists think that the social construction of gender, in itself, rules out the possibility of gender essentialism because of the variety of cultural norms and their fluctuations through history. We have seen that not all feminists agree with that position. A related issue concerns the reality of gender, which can be understood as a local dispute within the realism/anti-realism debate in the philosophy of science.

Some feminists, influenced by postmodernism and continental philosophy, hold that gender is not a real and determinate category, but a designation whose meaning is indeterminate and unstable. Both Butler (1990) and Drucilla Cornell (1993) have developed views along these lines. An antirealist view of gender has the positive attribute of allowing for immediate liberation for both individuals and groups through novel and creative performances of gender. If you think that gender is performed, enacted, created through behavior in unstable patterns and novel directions, then there is no difficulty in rejecting oppressive structures and stereotypes. Even those who choose to enact conventionally appropriate gender roles can miss the mark and fail to do so exactly. One tension in this position concerns the appropriate understanding of the subject, the agent who enacts liberatory behavior, since

anti-realists about gender tend to also reject the notion of the unified, coherent subject.

Other feminists accept the social constructionist thesis about gender, but do not conclude that gender categories are unreal, unstable or indeterminate in meaning. The division between natural entities and artificial or social entities (however we might wish to draw this distinction, and indeed even if we reject it) does not require us to place only natural entities on the side of reality. On the contrary, socially constructed identities like gender and race are fully determinate and very real in their effects on individuals and communities. One tension in this position concerns the autonomy of individuals who are the product of very real social norms and institutions. If we are constructed causally as women and men, then how can we act autonomously to resist patriarchal norms? One response to this issue is to distinguish between the social construction claim interpreted as making a causal claim (which raises the specter of determinism) and the social construction claim interpreted as a view about the social constitution of gender norms (which does not have any implications for determinism). Gender norms are socially constituted through cultural practices and social institutions, but it is up to the individual to accept or to resist them.

Feminist metaphysics is a robust field within feminist philosophy that also contributes in important ways to recent work in feminist social and political theory. Feminist metaphysics also contributes to mainstream metaphysical thought especially in the topics of subjectivity, autonomy and agency; and social ontology, social constructionism and essentialism.

See also Beauvoir, Simone de; Feminism and Continental Philosophy; Feminist Epistemology; Feminist Philosophy; Feminist Philosophy of Science; Irigaray, Luce; Metaphysics; Postmodernism; Social Constructionism.

Bibliography

Alcoff, Linda. "The Metaphysics of Gender and Sexual Difference." In *Feminist Interventions in Ethics and Politics*, edited by Barbara S. Andrew, Jean Keller, and Lisa H. Schwartzman. Lanham, MD: Rowman and Littlefield, 2005.

Appiah, Anthony K. "'But Would That Still Be Me?': Notes On Gender, 'Race,' Ethnicity, As Sources of 'Identity.'" *Journal of Philosophy* 87 (10) (1990).

Babbitt, Susan E. *Impossible Dreams: Rationality, Integrity, and Moral Imagination*. Boulder, CO: Westview Press, 1996.

Beauvoir, Simone de. *The Second Sex*, edited and translated by H. M. Parshly. New York: Alfred A. Knopf, 1993.

Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.

Cornell, Drucilla. *Transformations*. New York: Routledge, 1993.

Crenshaw, Kimberle. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women." *Stanford Law Review* 43 (6) (1991).

Dupre, John. "Sex, Gender, and Essence." In *Midwest Studies in Philosophy*, Vol. 11, edited by Peter A. French, Theordore E. Uehling, Jr., and Howard K. Wettstein, Minneapolis: University of Minnesota Press, 1986.

Fausto-Sterling, Anne. *Sexing the Body: Gender Politics and the Construction of Sexuality*. New York: Basic Books, 2000.

Frye, Marilyn. "Categories and Dichotomies." In *Encyclopedia of Feminist Theories*, edited by Loraine Code. New York: Routledge, 2000.

Frye, Marilyn. "Categories in Distress." In *Feminist Interventions in Ethics and Politics*, edited by Barbara Andrew, Jean Keller, and Lisa Schwartzman. Lanham, MD: Rowman and Littlefield, 2005.

Frye, Marilyn. "Essentialism/Ethnocentrism: The Failure of the Ontological Cure." In *Is Academic Feminism Dead? Theory in Practice*, edited by the Center for Advanced Feminist Studies at the University of Minnesota. New York: New York University Press, 2000.

Frye, Marilyn. "The Necessity of Differences: Constructing a Positive Category of Women." *SIGNS: Journal of Women in Culture and Society* 21 (3) (1996).

Frye, Marilyn. *The Politics of Reality: Essays in Feminist Theory*. Trumansburg, NY: Crossing Press, 1983.

Fuss, Diana. *Essentially Speaking: Feminism, Nature, and Difference*. New York: Routledge, 1989.

Garry, Ann, and Marilyn Pearsall. *Women, Knowledge and Reality*. 2nd ed. New York: Routledge, 1996.

Gatens, Moira. *Imaginary Bodies: Ethics, Power, and Corporeality*. New York: Routledge, 1996.

Haslanger, Sally. "Feminism and Metaphysics: Negotiating the Natural." In *The Cambridge Companion to Feminism in Philosophy*, edited by Miranda Fricker and Jennifer Hornsby. Cambridge, U.K.: Cambridge University Press, 2000.

Haslanger, Sally. "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Nous* 34 (1) (2000).

Haslanger, Sally. "Ontology and Social Construction." *Philosophical Topics: Feminist Perspectives on Language, Knowledge and Reality* 23 (2) (1995).

hooks, bell. *Ain't I a Women? Black Women and Feminism*. Boston: South End, 1981.

Meyers, Diana Tietjens, ed. *Feminists Rethink the Self*. Boulder, CO: Westview Press, 1997.

Meyers, Diana Tietjens. *Gender in the Mirror: Cultural Imagery and Women's Agency*. New York: Oxford University Press, 2002.

Spelman, Elizabeth V. *Inessential Woman: Problems of Exclusion in Feminist Thought*. Boston: Beacon Press, 1988.

Stoljar Natalie. "Essence, Identity and the Concept of Woman." *Philosophical Topics: Feminist Perspectives on Language, Knowledge and Reality* 23 (2) (1995).

Witt, Charlotte. "Anti-Essentialism in Feminist Theory." *Philosophical Topics: Feminist Perspectives on Language, Knowledge and Reality* 23 (2) (1995).

Witt, Charlotte. "Feminist Metaphysics." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by

Louise Antony and Charlotte Witt. Boulder, CO: Westview Press, 1992.

Wittig, Monique. "One is Not Born a Woman." In *The Second Wave: A Reader in Feminist Theory*, edited by Linda Nicholson. New York: Routledge, 1997.

Zack, Naomi. *Inclusive Feminism: A Third Wave Theory of Women's Commonality*. Lanham, MD: Rowman and Littlefield, 2005.

Zack, Naomi, ed. *Race/Sex: Their Sameness, Difference, and Interplay*. New York: Routledge, 1997.

Charlotte Witt (2005)

FEMINIST PHILOSOPHY

The rubric "feminist philosophy" applies to work in many philosophical subareas, often spanning several disciplines. The work is united by its authors' commitment to feminism in some form and by their belief that an engagement between feminism and philosophy will have both theoretical and practical benefits for everyone.

Some work in feminist philosophy focuses on philosophical issues that have arisen in the course of feminist political activism. Not surprisingly, much of this work is in political philosophy or ethics. Some work in feminist political philosophy consists of the articulation and defense of feminist theory, whereas other work examines the relationships between feminist political theory and other more general political theories, like liberalism and socialism. Much early work in feminist ethics dealt with issues in practical ethics that were of particular concern to women, such as abortion and affirmative action.

Throughout the 1980s and 1990s, however, feminist philosophers increasingly drew from other areas of philosophy to gain clarity about basic concepts in feminist theory and abstract foundational issues. Feminist work in metaphysics, for example, takes up such issues as the ontological status of categories like "gender" and "race," the basis of personal and cultural identity, the nature of truth, and the nature of freedom and autonomy. Feminist work in epistemology has been concerned, inter alia, with the relationship between practical and theoretical knowledge, the nature of intuition, the role of trust and other emotions in the achievement of knowledge, the social construction of expertise, and the nature of objectivity. Feminist philosophers of science ask such questions as why science has so often been enlisted on the side of sexism, and why so few women enter scientific fields, even today. Other burgeoning fields of feminist philosophy are feminist legal theory and feminist aesthetics.

Feminist philosophers have also been interested in understanding the ramifications of the historical exclusion of women from the discipline of philosophy. This exclusion has several forms: First, women have had very little opportunity, until very recently, to engage in systematic philosophical study; second, women and gender relations have received very little philosophical attention from the male authors who dominate the philosophical canon; and third, when women are discussed in the canonical literature, they are almost with exception represented as intellectually and morally inferior to men. Feminist philosophers have been concerned to document, analyze, and explain these various exclusions. Some feminist philosophers, including, prominently, many feminist philosophers of science, have concluded the methods and central concepts of traditional Western philosophy have been corrupted by an "androcentric" bias—a pervasive presumption that distinctively male characteristics and experiences provide appropriate normative standards for the whole human race. Other feminist philosophers argue the problem is a matter of grossly false assertions about women that can be excised without affecting traditional methods or concepts. Feminist historians of philosophy have also been engaged in the "uncovery" of female philosophers not properly recognized either in their own times or in the present.

As a result of these sorts of investigations, many feminist philosophers have concluded that there is a need for distinctively feminist methodologies and have been engaged, along with feminist theorists in other disciplines, in developing such methodologies. Typically, these methodologies focus on ways of knowing that have been denigrated or excluded by mainstream philosophy and thus emphasize the cognitive value of the emotions, of practical experience, and of social interaction.

Feminist philosophers come from a wide variety of intellectual backgrounds and invoke a variety of figures and texts. While feminist philosophers do not all agree about how deeply sexist the field is, they do agree that there is much in the institutional culture of academic philosophy that is inimical to women. Feminist philosophers work for reforms individually and collectively through informal professional networks and through such organizations as the American Philosophical Association's Committee on the Status of Women and the Society for Women in Philosophy.

See also Analytical Feminism; Feminism and Continental Philosophy; Feminism and Pragmatism; Feminism and the History of Philosophy; Feminist Aesthetics and Criticism; Feminist Epistemology; Feminist Ethics;

Feminist Legal Theory; Feminist Metaphysics; Feminist Philosophy of Science; Feminist Social and Political Philosophy.

Bibliography

Cudd, Ann, and Robin Andreasen, eds. *Feminist Theory: A Philosophical Anthology*. Oxford: Blackwell, 2005.

Fricker, Miranda, and Jennifer Hornsby, eds. *The Cambridge Companion to Feminism in Philosophy*. Cambridge, U.K.: Cambridge University Press, 2000.

Jaggar, Alison M., and Iris Marion Young, eds. *A Companion to Feminist Philosophy*. Oxford: Blackwell, 2000.

Kourany, Janet A., ed. *Philosophy in a Feminist Voice*. Princeton, NJ: Princeton University Press, 1998.

Kourany, Janet A., James Sterba, and Rosemarie Tong, eds. *Feminist Philosophies*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

Louise M. Antony (2005)

FEMINIST PHILOSOPHY OF SCIENCE

Feminist philosophy of science arises at the intersection of feminist interests in science and philosophical studies of science. Feminists have taken an active interest in the sciences both as a key resource in understanding and contesting sexist institutions and systems of belief, and as an important locus of gender inequality and source of legitimation for this inequality. Feminist practitioners in many sciences, especially in the life and social sciences, typically engage two lines of critique: They document inequalities in the training, representation, and recognition of women in the sciences, and they identify myriad ways in which, far from eliminating the contextual biases of a pervasively sexist society, standard scientific methodologies frequently reproduce them in the content of even the most credible and well-established scientific theories.

The work of feminist philosophers of science is continuous with these critiques. Some feminist philosophers contribute to the analysis of androcentrism in the content and practice of particular sciences, in some cases linking these to inequities in the role played by women in science. The form these analyses take necessarily varies with the type of science in question. Critiques of disciplines concerned with an overtly gendered subject matter—the social and behavioral sciences and some branches of the life sciences—draw attention to ways in which unexamined, often stereotypic, assumptions about gender roles, relations, and identities delimit the subject of inquiry, define categories of analysis and description, shape assessments of plausibility that define the range of hypotheses to be taken seriously (e.g., in comparative evaluation), and inform judgments about the bearing of evidence on these hypotheses. Women may be simply left out of account; behaviors, patterns of practice or development, and values and roles associated with men may be treated as normative for the population as a whole; where women diverge from male-defined norms they may be treated as deviant, immature, or anomalous; gender differences may be assumed irrelevant or, alternatively, taken as a given, a parameter for analysis rather than a variable; and the description and analysis of gendered subjects may be structured by conceptual categories that embody highly specific (enthnocentric) assumptions about the form that gender roles, identities, institutions, and values may take. In short, critiques in these domains call attention to ways in which the social and behavioral sciences (including ethology) are pervasively androcentric in content (see Bleier 1986, Haraway 1989, contributions to Harding and Hintikka 1983, Longino and Doell 1983, Tuana 1989, Wylie et al. 1990).

When the subject domain of a science is not overtly gendered, as in the case of most natural and life sciences, it may be projectively gendered, as when gendered categories are used to describe natural phenomena or when scientific categories have (gendered) social meanings (Potter 1988). And even when the subject is not characterized in gendered terms, feminist critics find that the enterprise and practice of science may be conceptualized in gendered terms, metaphorically characterized as the domain of men or as exemplifying masculine qualities of intellect and disposition (see Keller 1985). Whether or not these metaphors directly shape the content of science or, indeed, accurately characterize the practice of a majority of scientists, they do articulate and reinforce a conception of scientific inquiry that aligns it with attributes that are valorized as masculine (see Martin 1988).

The philosophical significance of these discipline-specific critiques lies in the questions they raise for our understanding of science, specifically, its objectivity, the role of values and interests in science, the status of scientific evidence and of extant methodologies for developing and evaluating scientific theory. If androcentrism is pervasive in much that is accepted as 'good," even exemplary, science—if it is by no means limited to examples of manifestly "bad" science (from Harding 1986)—then feminist critiques of science challenge us to rethink the relationship between what Longino has described as "contextual" and "constitutive" values (these correspond roughly to standard distinctions between cognitive or epistemic con-

siderations "internal" to science and the noncognitive, sociopolitical factors that many believe are properly "external" to science).

In taking up these questions the interests of feminist philosophers of science intersect with themes central to postpositivist philosophy of science. Feminist critiques of specific sciences illustrate, and draw attention to the implications of, central antifoundationalist claims about the complexity and contingency of scientific practice. If scientific theories are routinely (indeed, perhaps, necessarily) underdetermined by all available evidence, and if hypotheses are never evaluated independently of one another and the evidence supporting (or refuting) them is always itself richly interpreted (the theses of holism and the theory-ladenness of evidence), then it seem unavoidable that nonevidential values and interests, features of the "external" context of science, must play a role not only in the formulation but also in the evaluation of hypotheses. The contribution of discipline-specific feminist critiques is the insight that these contextual factors may include gendered interest, values, and social structures.

Although feminist philosophers are sometimes charged with advocating an untenable, "cynical," and self-defeating relativism (Haack 1993) because of their insistence that social factors such as gender shape the practice and results of science, in fact neither feminist critics within the sciences nor feminist philosophers of science show much sympathy for extreme forms of social constructivism or contextualism on which epistemic considerations are reduced to social, political factors. Harding's (1986) discussions of a "postmodern" epistemic stance and some of Haraway's (1989) reflections on hybrid constructions of nature may be seen to move in this direction. But Harding was explicitly "ambivalent" about postmodern options at the time she proposed them and has since elaborated a thesis of "strong objectivity" according to which an understanding of the standpoint (the social location, interests, values) of epistemic agents serves as a resource in producing and evaluating "less partial and less distorted" knowledge claims (1991). Haraway has likewise elaborated the concept of "situated knowledges" with the aim of capturing the sense in which it is reasonable to require "a no-nonsense commitment to faithful accounts of a 'real' world" while yet acknowledging the radical historical and social contingency of all knowledge production (1991).

In a similar vein, while Keller reaffirms the value of psychodynamic analyses of the masculine orientation of science (e.g., as elaborated in Keller 1985), she distances herself from strong sociological theses and argues the need for feminist analyses of science that attend to "logical and empirical constraints" and account for the "technological prowess" that makes scientific claims so compelling for scientists and for the world at large (1992, p. 3). The central preoccupation of feminist philosophers of science who elaborate a positive account of scientific inquiry is to understand the ways in which the (gendered) standpoint of epistemic agents and epistemic communities shapes inquiry while yet making sense of constraints imposed by constitutive values such as the standard requirements of epistemic adequacy, reliability, internal coherence, and consistency.

A number of positions have been explored in this connection. Feminist standpoint theory is one such approach. Harding's (1991) formulation draws on the earlier proposals of feminists, such as Hartsock (1983), who are influenced both by Marxist-derived epistemologies and by psychoanalytic theory, and on the work of black and minority feminist theorists who draw attention to the insights afforded by subdominant status (Collins 1991, Narayan 1988). The central thesis of standpoint theory, as developed by feminist theorists, is that the empirical evidence to which epistemic agents have access, their powers of discernment and breadth of understanding, may be both enhanced and limited by their social location and associated experience, values, and interests. For example, those who must understand a dominant world of privilege from which they are excluded as well as the subdominant world(s) of which they are members may well be better situated to understand both worlds, in empirical detail and with critical precision, than those who are beneficiaries of systemic privilege. The epistemic partiality and authority of knowledge claims, and therefore the effective assessment of their epistemic adequacy, is thus contingent on understanding the conditions under which they are produced and authorized, the standpoint of epistemic agents and communities.

A number of feminist philosophers of science have argued that the social dimensions of scientific practice (including but not limited to its gendered dimensions) can be understood in terms compatible with a modified empiricism. Longino's (1990) carefully worked distinction between contextual and constitutive values provides a framework for identifying the various points at which epistemic considerations leave room for the play of social factors, institutional context, political commitment, and personal interests in the formulation of descriptive categories, the interpretation of data as evidence, and the evaluation of hypotheses against evidence. At the same time she accords constitutive (epistemic) values a central

role, arguing that standards of rational acceptability can be identified that are independent of individual interests and that the social nature of science (e.g., institutional structures that encourage rigorous critical scrutiny of knowledge claims) serves as much to protect scientific knowledge from idiosyncratic bias as to render it vulnerable to such bias.

In a similar vein Nelson (1990) argues that an empiricist theory, which grounds knowledge in evidence and construes evidence in experiential terms, is compatible with a feminist reconceptualization of the agents of inquiry as communities, not abstract individuals, which are historically situated and of socially specified form. Sophisticated feminist empiricisms offer an account of epistemic virtues that transcend standpoint-specific interests—the virtues of empirical adequacy, reliability, scope of applicability, and explanatory power, which different standpoints help or inhibit us from realizing—without invoking an untenable (asocial) foundationalism.

Despite significant philosophical differences between proponents of these positions, feminist philosophers of science share an ambition to develop an account of science that resolves (or circumvents) the polarized debate between objectivists and rationalists on one hand and constructivists and relativists on the other. This is conceived both as a contribution to postpositivist philosophy of science, in which the terms of debate are most clearly articulated, and to feminist theory, where questions about the proper grounds for evaluating knowledge claims are a matter of immediate practical concern.

See also Feminist Epistemology; Feminist Metaphysics; Feminist Philosophy; Feminist Philosophy of Science: Contemporary Perspectives; Philosophy of Science; Sexism.

Bibliography

Bleier, Ruth. *Feminist Approaches to Science*. New York: Pergamon Press, 1986.

Collins, P. H. "Learning from the Outsider Within." In *Beyond Methodology: Feminist Scholarship as Lived Research*, edited by Mary M. Fonow and J.A. Cook. Bloomington: Indiana University Press, 1991.

Haack, S. "Knowledge and Propaganda: Reflections of an Old Feminist." *Partisan Review* 60 (1993): 556–564.

Haraway, Donna Jean. *Primate Visions: Gender, Race and Nature in the World of Modern Science*. New York: Routledge, 1989.

Haraway, Donna Jean. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." In *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge, 1991.

Harding, Sandra G. *The Science Question in Feminism*. Ithaca, NY: Cornell University Press, 1986.

Harding, Sandra G. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca, NY: Cornell University Press, 1991.

Harding, Sandra G., and Merrill B. Hintikka, eds. *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*. Boston: D. Reidel, 1983. 2nd ed., Boston: Kluwer, 2003.

Hartsock, N. "The Feminist Standpoint." In *Discovering Reality*, edited by Sandra G. Harding and Merrill B. Hintikka. Boston: D. Reidel, 1983.

Keller, Evelyn Fox. *Reflections on Gender and Science*. New Haven, CT: Yale University Press, 1985.

Keller, Evelyn Fox. *Secrets of Life, Secrets of Death: Essays on Language, Gender, and Science*. New York: Routledge, 1992.

Longino, Helen E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Longino, Helen E., and R. Doell. "Body, Bias, and Behavior: A Comparative Analysis of Reasoning in Two Areas of Biological Science." *Signs* 9 (1983): 206–227.

Martin, J. "Science in a Different Style." *American Philosophical Quarterly* 25 (1988): 129–140.

Narayan, U. "Working Together across Difference." *Hypatia* 32 (1988): 31–48.

Nelson, Lynn Hankinson. *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia: Temple University Press, 1990.

Potter, E. "Modeling the Gender Politics in Science." *Hypatia* 3 (1988): 19–35.

Tuana, Nancy, ed. *Feminism and Science*. Bloomington: Indiana University Press, 1989. 2nd ed., with Evelyn Fox Keller, New York: Oxford Unviersity Press, 1996.

Wylie, A., K. Okruhlik, L. Thielen-Wilson, and S. Morton, "Philosophical Feminism: A Bibliographic Guide to Critiques of Science." *Resources for Feminist Research* 19 (1990): 2–36.

Alison Wylie (1996)

FEMINIST PHILOSOPHY OF SCIENCE: CONTEMPORARY PERSPECTIVES

Feminists are a very diverse lot, but one thing they all share is a commitment to gender equality and a determination to bring it about. Feminist philosophers of science, along with feminist historians and sociologists of science and feminist scientists themselves, have focused especially on science, investigating both the ways science has helped to perpetuate gender inequality (their critical investigations) and the ways science can now help to eliminate it (their constructive investigations).

CRITICAL INVESTIGATIONS

Feminists' critical investigations have dealt with fields as diverse as primatology and molecular biology, economics and medical research, and their claims have been jarring. For example, feminists have documented a history of misogyny in both psychology and biology. In psychology, a dominant theme has been the inferiority—the intellectual, social, sexual, and even moral inferiority—of women to men (Marecek 1995, Wilkinson 1997). In biology, a host of research projects have aimed to "explain" the origins and manifestations of these presumed inferiorities in terms of what is largely unchangeable: genes, brain structure, and hormonal structure (Schiebinger 1989, Fausto-Sterling 1992, 2000). Feminists have argued that other sciences, as well, have supported this view of women's inferiority: for example, the historical sciences (such as archaeology), with their modes of representation of the past, modes of representation marked by heroic exploits and spectacular accomplishments of men counterpoised with lackluster doings or outright invisibility of women (Conkey and Williams 1991). And they have argued that still other scientific fields have perpetuated or added to the problems of inequality women confront, but in different ways than by documenting women's inferiority. Neglecting women's needs and priorities in the employment and household sectors in economic model-building, they have claimed, has had dire effects on public policy relating to women (Waring 1992, Ferber and Nelson 1993, Nelson 1996). And neglecting women in both basic and clinical research until well into the 1990s, they have added, has had dire effects on women's health care (Rosser 1994, Weisman and Cassard 1994, Schiebinger 1999). Other scientific fields that have figured prominently in feminists' critical investigations are anthropology, sociology, and political science, and even—with regard to their past and sometimes even present exclusionary practices—the physical sciences and mathematics (Kramarae and Spender 1992, Stanton and Stewart 1995, Schiebinger 1999, Kourany 2002).

CONSTRUCTIVE INVESTIGATIONS

Feminists' constructive investigations have been the site of considerable controversy, far more so than their critical investigations. It is agreed all around that science will aid the cause of equality for women if science works to replace prevailing ignorance and prejudice and misinformation about women with more adequate perspectives. But just how is this to be done?

THE METHODOLOGICAL APPROACH. Many feminist scientists have pointed out that a great deal of sexist science is, by the lights of traditional scientific methodology, simply *bad* science. Thus, they have taken to task mainstream authors of androcentric and sexist scientific work for failing to abide by accepted standards of concept formation, experimental design, interpretation of data, and the like (Bleier 1984, Hubbard 1990, Fausto-Sterling 1992). If such standards were rigorously followed, they have suggested, the problem of sexism and androcentrism in science would be, at the very least, much reduced. Feminist health researchers, for example, have pointed out that until the 1990s diseases such as heart disease that affect both sexes were defined as "male diseases," studied primarily in white, middle-aged, middle-class males, and clinically handled accordingly. As a result, heart disease in women (who, as it turns out, differ from men in symptoms, patterns of disease development, and reactions to treatment) was often not detected and not properly managed when it was detected. Such problems could be—and ultimately were—handled simply by following accepted methodological procedures such as designing clinical studies with groups of subjects that were more nearly representative of the patient population at large (see, for example, Rosser 1994 and the special report on "Women's Health Research" in *Science* 1995).

Other feminist scientists have explored ways of reforming traditional scientific methodology. Margrit Eichler (1988 and 1980), for example, has developed batteries of detailed sex- and gender-related guidelines concerning such aspects of research as concept formation, research design and instrumentation, and data interpretation to help scientists screen sexism and androcentrism out of their research, and the Biology and Gender Study Group (1988) conceptualizes such procedures as a new kind of experimental control to deal with gender bias. Feminist scientists have also explored new pedagogies to reform scientists themselves: for example, cooperative rather than competitive pedagogical methods and ones that take full note of the experiences of women and the contributions of women scientists (Rosser 1995).

All of these suggestions can be rationalized by appeal to the ideal of value-free science. According to this ideal, scientific investigations must be kept strictly free of ethical or political commitments. Since sexism and androcentrism embody social values, they simply do not belong in science. Indeed, they bias science and thereby jeopardize science as an impartial resource in the struggle for social justice. On this view of science, the *only* legitimate strategy for eliminating sexist and androcentric bias is to

press for stricter adherence to the canons of scientific inquiry on the part of individual scientists. This view, that good method will yield science undistorted by sexism or androcentrism, can be called the methodological approach.

THE SOCIAL APPROACH. Few feminist philosophers of science accept the individualistic and formalistic conception of science implicit in the methodological approach. Some of them, along with some feminist scientists, have opted instead for a social approach unallied with the ideal of value-free science. They argue that no scientific method, however rigorous and however rigorously applied, can be guaranteed to screen out the various values and interests that scientists from their different social locations (race, gender, class, and so on) bring to their research. Scientists' values and interests can and do determine which questions they investigate and which they ignore, can and do motivate the background assumptions they accept and those they reject, can and do influence the observational or experimental data they select to study and the way they interpret those data, and so on. As a result, changes must be sought in the communities that generate our scientific knowledge if the knowledge generated is to aid the cause of equality for women. After all, scientific communities have historically been dominated by men—men who have been raised within sexist and androcentric societies and trained within sexist and androcentric scientific traditions; men who, moreover, profit from this sexism and androcentrism. Small wonder, then, that sexist and androcentric values have shaped the scientific knowledge generated by these communities.

But if changes should be made in the communities that generate our scientific knowledge, exactly what should be the nature of these changes? Here advocates of the social approach differ. Feminist-standpoint theorists argue that women—who also have been raised within sexist and androcentric societies and trained within sexist and androcentric scientific traditions—are still in a better position than their male counterparts to uncover and critique sexist and androcentric scientific perspectives and replace them with more adequate perspectives (are still in a better position, for example, to uncover and critique sexist assumptions about the sexual division of labor in prehistory made by male archaeologists and replace them with questions and hypothetical answers suggestive of new lines of research). "They have less to lose by distancing themselves from the social order; thus, the perspective from their lives can more easily generate fresh and critical analyses" (Harding 1991, p. 126, and cf. 1986). The upshot is that women's perspectives should not only be welcomed into scientific communities, but they should also be privileged over men's perspectives, at least in gender-relevant areas of research, if the knowledge generated by those communities is to be an adequate basis for social justice.

Feminist empiricists such as Helen Longino and her followers, on the other hand, argue that standpoint theorists fail to take note of the diversity of perspectives of both women and men. There are women, for example, who have participated in research that is damaging to women, and there are men who have done just the opposite (see, for example, the diversity of perspectives in the special report on "Women's Health Research" in *Science* 1995). As a matter of empirical fact there simply is no one standpoint shared by all and only women, and hence, no "women's standpoint" especially conducive to uncovering and correcting prevailing ignorance and prejudice and misinformation about women. If science is to provide us with more adequate views about women, Longino urges, scientific communities must finally be made into inclusive places where women and feminist perspectives are given an equal though not a privileged hearing. More specifically, scientific communities will have to have public venues for criticism, publicly recognized standards by reference to which criticism can be made, "uptake" of such criticism (that is, the criticism will have to be taken seriously and responded to), and "tempered equality" of intellectual authority among all parties to the debate, among whom "all relevant perspectives are represented" (Longino 2002, pp. 128–135; and cf. 1990). Only if scientific communities are organized in these ways, says Longino, will the necessary "transformative criticism" of our current views of women be possible. But Longino gives us no reason to believe—and certainly no empirical evidence to suggest—that organizing scientific communities in these ways *will* issue in that transformative criticism, that is, will dispel the ignorance and prejudice and misinformation about women of which we are now possessed.

THE POLITICAL APPROACH. This motivates yet another approach different from both the methodological and social approaches—what might be called the political approach. Like the methodological approach, the political approach recognizes that sexism and androcentrism must be rooted out of science if science is to replace prevailing ignorance and prejudice and misinformation about women with more adequate perspectives, but unlike the methodological approach, the political approach also recognizes that rooting sexism and androcentrism out of science is tantamount to implanting egal-

itarian social values into science. Again, like the social approach, the political approach recognizes that social values inevitably enter into science, but unlike the social approach, the political approach recognizes that we as a society have a definite say—through funding priorities and restrictions, for example—as to what these social values will be. Indeed, given that science is both a profound shaper of society and a profound beneficiary of society, these social values should be chosen so as to meet the needs of society, including the justice-related needs of society.

Under the political approach, in short, our scientific views (and hence, ultimately, our generally accepted knowledge) of women would no longer be plagued by sexism or androcentrism simply because those would be the morally justified political conditions under which scientific research would be conducted (Kourany 2003, Anderson 1995, 2004). But would this political structure for science jeopardize science's objectivity? That is to say, would it render science's resultant "knowledge of women" not genuine knowledge at all?

THE NATURALIST APPROACH. Feminist naturalists provide a possible answer. A naturalist approach to the philosophy of science rejects a priori prescriptions about the conduct of inquiry or the composition of scientific communities. This approach advocates instead a close look at successful scientific practice in order to identify those of its features that contribute to and explain its success (Antony 1993, 1995). Such observation shows, feminist naturalists point out, that egalitarian social values need not compromise the objectivity of science any more than do other features of scientific communities such as competitiveness, deference to authority, or the desire for credit for one's accomplishments.

Indeed, such observation shows, feminist naturalists point out, that egalitarian social values can be *aids* in the acquisition of objective knowledge: when these values are allowed to influence science (for example, by motivating particular lines of research or the maintenance of particular social structures), that science can actually be more developed and more empirically adequate than before (Wylie and Nelson 1998, Campbell 2001). And when we reflect on the effects of feminism in science during the last three decades—the wide-ranging critiques of traditional science in such fields as psychology, sociology, economics, political science, archaeology, anthropology, biology, and medical research, and the new research directions and results forged in the wake of these critiques—when we reflect on the effects of feminism in science during the last three decades, the claims of these feminist naturalists seem especially convincing. Egalitarian social values in these cases have seemed to yield better rather than worse science, more objective rather than contaminated science (Schiebinger 1999; Creager, Lunbeck, and Schiebinger 2001).

Feminist naturalism, however, faces at least one large problem, one that stems from a problem for naturalized epistemology in general: It threatens to eliminate the normative in favor of the descriptive, and in doing so, eliminate the grounds for normative critique. It is impossible, after all, to say a priori which values will be aids and which will be hindrances to the acquisition of objective knowledge. Racism and sexism and egalitarian social values, all are possible aids or hindrances to the acquisition of objective knowledge, and all must be empirically tested to see which they are. There is at least the suggestion, therefore, that any of them will do if only they can prove their mettle in scientific research. So if, for example, a close comparative study of German science before, during, and after the Third Reich discloses that Nazi social values produced the best scientific results, the most abundant and most empirically successful science, then Nazi social values would be "good" values and should therefore be welcomed into science. Or if such a study discloses that Nazi social values produced a science just as good as the others, but no better, then it should be a matter of complete indifference whether Nazi social values or the other sciences' values should find their way into science. In short, feminist naturalists do not tell us what considerations, other than empirical adequacy, ought to govern our choice of social values. Some feminist naturalists emphasize that social values are empirically tested by an interrelated *system* of facts and values (Nelson 1990, 1993; Anderson 1995, 2004; Campbell 1998), but it is unclear whether this move is sufficient to address the general problem.

THE CONTRIBUTION TO PHILOSOPHY OF SCIENCE

Feminists have pursued still other approaches in their constructive investigations of science, but what do they, or the critical investigations that preceded them, have finally to do with *philosophy of science*?

Nearly a half-century ago, Thomas Kuhn, Paul Feyerabend, Stephen Toulmin, Norwood Russell Hanson, and others issued a challenge to philosophers of science to make their field more relevant to actual science. That challenge, over time, has elicited a number of useful responses: first, efforts to "historicize" philosophy of sci-

ence, to make philosophy of science relevant to the actual development of science, both past and present; and second, efforts to "socialize" philosophy of science, to make philosophy of science relevant not only to science's conceptual products but also to the actual knowledge-productive social practices that have led to those products. But very few efforts have thus far been made to "societize" philosophy of science, to make philosophy of science relevant to the ways in which science interacts with the wider society in which it occurs, the ways in which science both shapes and is shaped by that society. The unit of analysis for philosophy of science has tended to remain (an historical, social) science-in-a-vacuum. Feminist philosophers of science, in collaboration with feminist historians and sociologists of science and feminist scientists themselves, provide philosophers of science with a start to a societized philosophy of science.

First, feminists *have* situated science within its wider social context when philosophizing about science. Indeed, feminists have been especially concerned with the social consequences of science—in particular, the ways science has all too frequently perpetuated and added to the problems of inequality women confront. This concern with science's social consequences has led feminists to scrutinize those features of science that help to shape its social consequences—not only the research strategies of scientists but also their social location and training, the social as well as epistemic values that inform their practice, and the funding priorities that direct their research. What's more, in all this feminists have been motivated, not only by the need and desire for understanding, but also by the need and desire for social change, and they have explored social/political/epistemic initiatives intended to bring about that change—new funding priorities for science, for example, or new kinds of recruitment or training programs, or new social or epistemic values.

In short, feminist philosophers of science, in collaboration with feminist historians and sociologists of science and feminist scientists themselves, have been pursuing a comprehensive analysis of science-in-society and a comprehensive plan of action to bring about needed change in both science and society. This is the first way in which feminists have given us a start to a societized philosophy of science—by giving us a ready-made example of such philosophizing.

In addition, the ready-made example of societized philosophy of science that feminists have given us can be generalized—this is the second way in which feminists have given us a start to a societized philosophy of science. Indeed, science has all too frequently perpetuated and added to other kinds of inequality besides gender inequality—inequality relating to race and sexual orientation and physical ability and disability, for example. And science has all too frequently perpetuated and added to other kinds of social problems besides those relating to inequality—problems relating to the environment, for example, and problems relating to the inability to achieve peaceful coexistence among nations. What's more, with different kinds of funding priorities, or different kinds of recruitment or training programs, or different kinds of social or epistemic values, or the like, science can not only cease to put obstacles in the way of solutions to these problems, but more effectively help to bring those solutions about. So there is much descriptive and normative philosophical work to be done on many fronts, philosophical work that can, at least in part, be modeled on the work already done by feminists.

Finally, the work done by feminists provides not only a generalizable example of societized philosophy of science, but it provides important additional resourcesas well—insights concerning the relations between epistemic values and social values and the place of social values in science, for example, insights concerning what makes for scientific objectivity and what threatens it, insights concerning the ultimate goals of science and the methods that are appropriate to them, and the like. This is the third way in which feminists have given us a start to a societized philosophy of science.

See also Feminist Epistemology; Feminist Metaphysics; Feminist Philosophy; Feminist Philosophy of Science; Philosophy of Science; Sexism.

Bibliography

Anderson, Elizabeth. *Hypatia* 19 (2004): 1–24.

Anderson, Elizabeth. "Knowledge, Human Interests, and Objectivity in Feminist Epistemology." *Philosophical Topics* 23 (1995): 27–58.

Antony, Louise. "Quine as Feminist: The Radical Import of Naturalized Epistemology." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise Antony and Charlotte Witt, 110–153. Boulder, CO: Westview, 1993.

Antony, Louise. "Sisters, Please, I'd Rather Do It Myself: A Defense of Individualism in Feminist Epistemology." *Philosophical Topics* 23 (1995): 59–94.

Biology and Gender Study Group. "The Importance of Feminist Critique for Contemporary Cell Biology." *Hypatia* 3 (1988): 61–76.

Bleier, Ruth. *Sex and Gender*. New York: Pergamon Press, 1984.

Campbell, Richmond. "The Bias Paradox in Feminist Epistemology." In *Engendering Rationalities*, edited by Nancy

Tuana and Sandra Morgen. Albany: State University of New York Press, 2001.

Campbell, Richmond. *Illusions of Paradox: A Feminist Epistemology Naturalized*. Maryland and Oxford: Rowman and Littlefield, 1998.

Conkey, Margaret W., and Sarah H. Williams. "Original Narratives: The Political Economy of Gender in Archaeology." In *Gender at the Crossroads of Knowledge: Feminist Anthropology in the Postmodern Era*, edited by Micaela di Leonardo, 102–139. Berkeley and Los Angeles: University of California Press, 1991.

Creager, Angela N., Elizabeth Lunbeck, and Londa Schiebinger, eds. *Feminism in Twentieth-Century Science, Technology, and Medicine*. Chicago: University of Chicago Press, 2001.

Eichler, Margrit. *The Double Standard: A Feminist Critique of Feminist Social Science*. New York: St. Martin's Press, 1980.

Eichler, Margrit. *Nonsexist Research Methods: A Practical Guide*. Boston: Allen & Unwin, 1988.

Fausto-Sterling, Anne. *Myths of Gender*. 2nd ed. New York: Basic Books, 1992.

Fausto-Sterling, Anne. *Sexing the Body: Gender Politics and the Construction of Sexuality*. New York: Basic Books, 2000.

Ferber, Marianne A., and Julie A. Nelson, eds. *Beyond Economic Man: Feminist Theory and Economics*. Chicago: University of Chicago Press, 1993.

Harding, Sandra. *The Science Question in Feminism*. Ithaca, NY: Cornell University Press, 1986.

Harding, Sandra. *Whose Science? Whose Knowledge?* Ithaca, NY: Cornell University Press, 1991.

Hubbard, Ruth. *The Politics of Women's Biology*. New Brunswick, NJ: Rutgers University Press, 1990.

Kourany, Janet A., ed. *The Gender of Science*. Upper Saddle River, NJ: Prentice Hall, 2002.

Kourany, Janet A. "A Philosophy of Science for the Twenty-First Century" and "Reply to Giere." *Philosophy of Science* 70 (2003): 1–14, 22–26.

Kramarae, Cheris, and Dale Spender, eds. *The Knowledge Explosion*. New York and London: Teachers College Press, 1992.

Longino, Helen. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press, 2002.

Longino, Helen. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Marecek, Jeanne. "Psychology and Feminism: Can This Relationship Be Saved?" In *Feminisms in the Academy*, edited by Domna C. Stanton and Abigail J. Stewart, 101–132. Ann Arbor: University of Michigan Press, 1995.

Nelson, Julie A. *Feminism, Objectivity and Economics*. London and New York: Routledge, 1996.

Nelson, Lynn Hankinson. "Epistemological Communities." In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter. New York: Routledge, 1993.

Nelson, Lynn Hankinson. *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia: Temple University Press, 1990.

Rosser, Sue, ed. *Women's Health—Missing from U.S. Medicine*. Bloomington and Indianapolis: Indiana University Press, 1994.

Rosser, Sue, ed. *Teaching the Majority: Breaking the Gender Barrier in Science, Mathematics, and Engineering*. New York: Teachers College, Columbia University, 1995.

Schiebinger, Londa. *Has Feminism Changed Science?* Cambridge, MA: Harvard University Press, 1999.

Schiebinger, Londa. *The Mind Has No Sex?* Cambridge, MA.: Harvard University Press, 1989.

Stanton, Domna C., and Abigail J. Stewart, eds. *Feminisms in the Academy*. Ann Arbor: University of Michigan Press, 1995.

Waring, Marilyn J. "Economics." In *The Knowledge Explosion*, edited by Cheris Kramarae and Dale Spender, 303–309. New York and London: Teachers College Press, 1992.

Weisman, Carol S., and Sandra D. Cassard. "Health Consequences of Exclusion or Underrepresentation of Women in Clinical Studies (I)." In *Women and Health Research*. Vol. 2, edited by Anna C. Mastroianni, Ruth Faden, and Daniel Federman, 35–40. Washington, DC: National Academy Press, 1994.

Wilkinson, Sue. "Still Seeking Transformation: Feminist Challenges to Psychology." In *Knowing Feminisms: On Academic Borders, Territories and Tribes*, edited by Liz Stanley, 97–108. London: Sage Publications, 1997.

"Women's Health Research." *Science* 269 (1995): 765–801.

Wylie, Alison, and Lynn Hankinson Nelson. "Coming to Terms with the Value(s) of Science: Insights from Feminist Science Scholarship." Paper delivered at the Workshop on Science and Values, Center for Philosophy of Science, University of Pittsburgh, 1998.

Janet A. Kourany (2005)

FEMINIST SOCIAL AND POLITICAL PHILOSOPHY

Within the enormously varied and fluid field of feminist social/political philosophy and political theory, several foci can be identified: analyses of women's oppression; explorations of differences among women and their implications for feminism; critiques of political philosophers and retrieval of little-known women philosophers; reanalyses of central concepts in political philosophy; analyses and recommendations on practical political issues.

A COMMON THEORETICAL BASIS FOR FEMINISM?

Whether there is anything that cuts across these different areas of work and the varieties of theoretical perspectives is not entirely clear. If feminism is to have a common basis it would seem necessary to say that whatever the disagreements, all agree that women are oppressed, or at least subordinated to men, and that to eliminate this requires not only legal changes of a kind that have mostly been achieved in the developed world, but more pro-

found changes in society and consciousness. However, even these modest generalizations are suspect to postmodernist feminists who eschew talk of "women" because the term conceals so many differences among women, and who are skeptical of claims to truth and objectivity.

Indeed, the question of differences—both between women and men, and among women—has been a consuming issue throughout the history of feminism. In first-wave feminism, whether women and men had distinct natures (beyond the biological) was the dominant theoretical question, with early feminists like Mary Wollstonecraft and Harriet Taylor basing their call for women's rights on the claim that women had the same capacity for reason as men. Utopian socialists and Marxists agreed, deepening the critique of naturalistic justifications of the hierarchy between women and men, with a call to end class inequality as well. By the time of second-wave feminism, most educated people agreed that whatever differences exist between women and men were largely social in origin and certainly not sufficient to explain women's subordination. Even among feminists, however, this view was not universal and for a period this disagreement consumed considerable attention. Nevertheless, the question of "differences" that dominated second-wave feminism and beyond was differences among women and how they affected the feminist project.

The issue was not discussed directly in those terms at first. Feminists assumed that women could all be said to be treated unfairly, or to be oppressed, the particular word chosen reflecting different political theoretical perspectives, but most shared an optimistic assumption of commonality expressed in slogans like Sisterhood Is Powerful. The issue of differences among women emerged implicitly, however, in debates regarding how to understand women's subordination. The standard labels for the competing political and philosophical perspectives on the roots of oppression and how to end it, best explicated by Alison Jaggar, are liberal feminism, Marxist feminism, radical feminism, and socialist feminism. There has also been much rich discussion of how to conceive oppression that is independent of these labels, by Iris Young, for example.

CRITIQUES AND REVISIONS OF LIBERALISM

Liberal feminism is liberal theory as adapted by criticisms that women had been left out. By and large, liberal feminists in the United States and Western Europe accept the terms of the dominant political discourse such as methodological individualism, the centrality of the values of individual freedom and choice, the focus on legal and political change, such as securing the legal right to abortion and the passage of an Equal Rights Amendment, and a faith in education to eradicate prejudice. They believe that the dominant political and economic system, that is, capitalism, is compatible with equal opportunities for women, but that many existing social arrangements need to be changed. In particular, they argue that it is unjust that the care of children should be exclusively women's responsibility and they call for arrangements to make possible sharing of childcare, like part-time work. Liberal feminists accept sexual freedom as a matter of individual right, but it is not central to their concerns, nor are differences among women along the lines of race/ethnicity, class, or sexuality.

The extension of the concept of justice from the public sphere to the family, traditionally understood as private, is one of the most distinctive features of feminist thought. While feminists differ on the importance of the notion of privacy, they point out that what counts as the private depends on the public, that is, legislation, and question many aspects of this fundamental distinction. Independence is another central concept and value that feminists question, pointing out that humans are all interdependent and that some people's independence is actually dependent on the invisible or undervalued labor of others, usually women. Feminists have also reconceived the concepts of autonomy and obligation in more relational terms, have debated the adequacy of rights talk for feminist concerns, have proposed that rights be extended to groups, and have explored positive and negative dimensions of power. The social contract tradition within liberalism, particularly Thomas Hobbes, has been radically reconceived by Carole Pateman as in actuality a sexual contract.

Most of these criticisms of liberal political philosophy are still within liberal feminism, but a broader sense of liberalism that encompasses social-welfare liberalism. Since it is these latter types of liberalism that have been influential in East and Central Europe, along with strains of liberalism that recognize collective goods like the family and the nation, many of the Western feminist critiques of liberalism do not apply there. As Nanette Funk shows, in those contexts feminists have needed to insist on individual rights versus the common good and neutral universalistic rights versus gendered and nationalistic conceptions. Western European and American feminists have also disagreed on these issues because they have disagreed about the source and centrality to political theory of differences between women and men.

In the 1980s and 1990s, an approach known as "difference feminism" was very influential, according to which universalistic gender neutral ideals and policies did not do justice to women's specific roles and capacities. Some went so far as to hold that these differences between women and men were biologically based, but most accepted a psychoanalytic approach to understanding the origins of psychological sex differences rooted in the fact that women are the primary caretakers of children; they paid little attention to class, race/ethnic and historic variations among women and men. Given male/female differences, whatever the source, they held that citizenship should be reconceived and accommodations for women should be made in law and public policy, such as pregnancy and maternity leave. Other feminists favored gender neutral policies such as disability and parenting leave.

Feminist philosophers have brought to light little known women philosophers such as Christine Di Pisan who had the idea of the body politic before Hobbes, and have examined classic and contemporary political philosophers with feminist eyes. Their purpose is not simply to expose sexist assumptions but to explore how central these are to the theory. Sexism is seen as ineliminable from the political theories of Aristotle, Jean-Jacques Rousseau, and Georg Hegel, for example. John Locke is credited by some with opening the door to feminism, but others, such as Lorene Clark, argue that Locke's theory is fundamentally inconsistent; while political obligation is said to rest on free, equal, and rational individuals consenting to a limited government, Locke's theory requires that women be subordinate in the family and society in order to guarantee his other aim, the preservation of private property. Hence, Locke's theory cannot be rewritten in universal terms. Not all feminist critiques of political philosophers have been so devastating. For example, according to Susan Moller Okin, though Rawls assumed the traditional sexual division of labor in his theory of justice, and did not extend the sphere of justice into the family, his theory does not depend on this sexist limitation and would be stronger without it.

MARXIST, RADICAL AND SOCIALIST FEMINIST PERSPECTIVES

Many feminists, particularly outside the United States, have found Marxism a useful tool for understanding women's oppression. Although focused on economic exploitation, Marxism does not deny other forms of oppression, like sexism or racism, or reduce them to the economic, (except for the crudest of "Marxists"), but it gives them less explanatory primacy. According to Marxism each exploitative mode of production, such as feudalism or capitalism, is distinctive in its mode of exploitation and each gives rise to certain distinctive forms of government, religion, culture, and family. Thus relations between women and men will vary in different modes of production. While women's lot in life is better in capitalism than in feudalism or slave societies, Marxist feminists generally maintain that sexism has certain benefits for capitalism, such as allowing socially necessary caring labor to be unpaid, and hiding the (un)(der)employment of women. They have debated the relations between sexism and capitalism, such as whether housework is exploited in a Marxist sense, whether women can be said to constitute a class and how domination and alienation at work contribute to the hierarchical construction of gender. For a sample, see the debate between Wally Secombe and Margaret Coulson et al. in the *New Left Review* (1975). Some feminist uses of Marxism involve quite significant revisions of Marxism, and in Europe some call this radical feminism.

Implicit in a Marxist approach is that women share certain common interests, but that women of different economic classes also have fundamentally different interests; and, moreover, that these are likely to be most important to them. For example, all women need the legal right to birth control and abortion, but poor women need public funding to exercise this right. The greatest problem facing women around the world is extreme poverty, according to the World Health Organization, but women capitalists profit directly from this poverty, while many others benefit from poor women's cheap labor. The political and strategic implications are that all women should unite to secure their common cross-class interests, but that working class women need to work with working class men to secure their specific economic interests, and that ultimately the elimination of women's oppression requires the end of capitalism.

Radical feminism, the youngest and most fluid of feminist theoretical perspectives, was developed by feminists who saw liberalism's goal of equality for women as not nearly radical enough and Marxism's focus on the economic as blind to the specific oppression of women by men of all classes. The very notion of politics, they held, must be radically reconceived. "The personal is political," they proclaimed. Some radical feminists like Catherine MacKinnon attempted to develop a theory in which sex replaced class as the primary category with which to understand history and current societies, seeing most societies as profoundly misogynist. Whether they share this overarching theory or not, radical feminists see dif-

ferences among women such as race/ethnicity, class or nationality as less important than what unites them — oppression by men, particularly sexual violence, focusing attention on the outrageous prevalence of rape and its use as a weapon of war, on trafficking and sexual slavery, and on pornography, Andrea Dworkin's work being the most notable on the latter. Most radical feminists are deeply skeptical about the pleasures of sexual liberation for women, focusing instead on the dangers and coercion of heterosexual sex in a male dominated universe, although some sexual liberationists might also fall within the radical feminist camp. Many have connected violence against women to violence against other species and nature, universally associated with women, and some have evolved into "difference feminists," echoing those first-wave feminists who argued for women's suffrage on the grounds that women were more moral than men.

"Socialist," as distinct from Marxist, feminism is best understood as a synthesis of Marxism and radical feminism. Maintaining that women's oppression in capitalist society is a function of both the economic system, capitalism, and the sex/gender system, which they called patriarchy, socialist feminists like Heidi Hartman refused to give primacy to one over the other. Many saw as sexist the Marxist emphasis on wage labor rather than on all kinds of labor, especially women's unpaid caring labor, and on the relations of production, rather than on what they called the "relations of reproduction" (sexuality and parenting). To correct this deficiency Ann Ferguson proposed a concept of "sex-affective production." While its synthesis is attractive, the theory gives rise to questions as to whether the oppression of women requires a "system" (patriarchy) to explain it, and if so, why doesn't racism or heterosexism, require a system to explain them, and what exactly a "system" means anyway. Some socialist feminists tried to accommodate racism by adding a race/ethnicity system, but questions remain regarding the meaning of "system," how the systems are related, and how the theory differs from simple pluralism.

In the twenty-first century, as Nancy Holmstrom explains, "socialist feminism" is often used more broadly to refer to any theory that tries to integrate class and sex, as well as other aspects of identity such as race/ethnicity in a coherent way, however exactly they are related. On this broad definition, it would encompass perspectives that either go by other names such as materialist feminism, womanism and black feminism or that have no theoretical labels of any kind. Which term a feminist uses to describe herself indicates where she wishes to position herself within certain debates or else signals certain commitments, but is not necessarily a "grand theory" in competition with liberal, Marxist/socialist or radical feminism. Although "materialist feminism" was introduced by Christine Delphy and Colette Guillaumin as a competing grand theory, and the label has recently been used by feminists wishing to engage with postmodernism, it fits within this broad definition. "Womanist" was introduced initially by some women of color who felt that "feminism" is too one-dimensional and who wished to indicate solidarity with men of color as well as women. "Black feminist," particularly as developed by Patricia Hill Collins, is a position whose insights stem from the particular experiences of African-American women.

RETREAT FROM GRAND THEORY

Most feminists in the early twenty-first century, especially in the United States, eschew the word socialist, both because of negative associations and because of an anti-theoretical mood brought on by postmodern criticisms of "totalizing narratives." Instead of one overarching feminist theory, feminists prefer to rest on the concept of intersectionality, to use Kimberle Crenshaw's useful descriptive term. Racism, sexism, classism, and heterosexism are seen as overlapping forms of oppression, similar in some ways, different in others, none of which is more important politically or theoretically than the others. But if being a woman cannot be separated from being a particular kind of woman, black or white or gay or working class, then this seems to imply that there can be no theory of women's oppression as such. And this suggests there is no basis for feminism, a theory and political movement for all women, but rather only for particular kinds of women, for example for black women. Moreover, the same logic can be carried further. Black women are also of a particular nationality, class, sexual orientation, (dis)ability. Thus, this seems to lead to a dead-end theoretically.

A hopeful assumption widespread among twenty-first-century feminists is that while commonalities cannot be assumed, they can be found, unity can be forged, despite the differences among women, but only with strong political commitment and efforts to seek commonalities. It entails accepting, negotiating and transcending differences and first of all, it means really listening. Implicit in this approach is the assumption that the various kinds of differences—"identities"—are on a par: race/ethnic, class, sexual orientation, (dis)ability. To give any order of importance is mistaken and oppressive.

This sounds promising for feminist political philosophy in that it could provide a common basis for femi-

nism, without denying differences. However, whether this approach to overcoming or at least bridging differences is applicable to all the different kinds of difference depends on whether or not they are inherently antagonistic. A plausible example is sexuality. Despite what social conservatives say, heterosexuality is not threatened by homosexuality. Neither the existence of heterosexuals nor their happiness is compromised by acceptance of different kinds of sexual and emotional desire. On the other hand, class differences are more problematic. Imagine a conversation between two women, a sweatshop owner and her employee. However much they talk and negotiate and understand each other's position, how is the difference between them to be overcome? Since classes are socially constituted by their antagonistic relationship of interest and power, those relations between members of different classes will persist.

Other feminist philosophers have been more involved with ethical theory, particularly care ethics, than with wholesale analyses of oppression, assuming that sufficient commonalities exist among women to justify their analyses and policy recommendations. Nel Nodding's approach starts with a characterization of the best of familial relations and then applies the lessons learned there to broad social policies regarding welfare, education, and criminal justice. Especially since the September 11, 2001, terrorist attacks, issues of war, peace, and terrorism have received a lot of attention, but Sara Ruddick connected mothering to peace politics early on. Many feminist philosophers have turned their attention in recent years to global gender issues, and have debated whether human rights, capabilities, or a care ethics is the most fruitful approach. Postcolonial feminists like Chandra Mohanty pay particular attention to the ways in which colonialism and imperialism work together with patriarchal structures and ideology to subordinate women. Within global feminism, differences among women are again a problematic issue, as the controversy around Okin's critique of multiculturalism attests.

See also Aristotle; Feminism and Pragmatism; Feminist Ethics; Feminist Metaphysics; Feminist Philosophy; Ferguson, Ann; Hegel, Georg Wilhelm Friedrich; Heterosexism; Hobbes, Thomas; Locke, John; Marxist Philosophy; Racism; Rawls, John; Rousseau, Jean-Jacques; Social and Political Philosophy; Wollstonecraft, Mary.

Bibliography

Butler, Judith, and Joan Scott, eds. *Feminists Theorize the Political*. New York: Routledge, 1992.

Clark, Lorenne M. G., and Lynda Lange. *The Sexism of Social and Political Philosophy*. Toronto: University of Toronto Press, 1979.

Collins, Patricia Hill. *Black Feminist Thought*. 2nd ed. New York: Routledge, 2000.

Coulson, Margaret, Branka Magas, and Hilary Wainwraight. "The Housewife and Her Labour Under Capitalism." *New Left Review* 89 (1975).

Crenshaw, Kimberle. *Demarginalizing the Intersectionality of Race and Sex: A Black Feminist Critique of Anti-Discrimination Doctrine, Feminist Theory and Antiracist Politics*. Chicago: University of Chicago Legal Forum, 1989.

Daly, Mary. *Gyn/Ecology: A Metaethics of Radical Feminism*. Boston: Beacon, 1978.

Davis, Angela. *Women, Race and Class*. New York: Random House, 1983.

Delphy, Christine, and Diane Leone. *Close to Home: A Materialist Analysis of Women's Oppression*. New York: Harper-Collins, 1984.

Dworkin, Andrea. *Pornography: Men Possessing Women*. New York: Perigee, 1981.

Funk, Nanette. "Feminist Critiques of Liberalism, Can They Travel East." *Signs* 29 (3) (2004).

Ferguson, Ann. *Sexual Democracy: Women, Oppression, and Revolution*. Boulder, CO: Westview Press, 1991.

Guillaumin, Colette. *Racism, Sexism, Power and Ideology*. New York: Routledge, 1995.

Hartmann, Heidi. "The Unhappy Marriage of Marxism and Feminism: Toward a More Progressive Union." In *Women and Revolution*, edited by Lydia Sargent. Boston: South End Press, 1981.

Held, Virginia. *Rights and Goods: Justifying Social Action*. Chicago: University of Chicago Press, 1984.

Holmstrom, Nancy. *The Socialist Feminist Project: A Reader in Theory and Politics*. New York: Monthly Review Press, 2002.

Jaggar, Alison. *Feminist Politics and Human Nature*. Totowa, NJ: Rowman and Allanheld, 1983.

MacKinnon, Catherine A. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press, 1989.

Mohanty, Chandra Talpade. *Feminism without Borders: Decolonizing Theory, Practicing Solidarity*. Durham, NC: Duke University Press, 2003.

Noddings, Nel. *Starting at Home*. Berkeley: University of California Press, 2002.

Nussbaum, Martha, and Jonathan Glover, eds. *Women, Culture and Development: A Study of Human Capabilities*. Oxford: Oxford University Press, 1995.

Okin, Susan Moller,*Is Multiculturalism Bad for Women?* Princeton, NJ: Princeton University Press, 1999.

Okin, Susan Moller. *Justice, Gender and the Family*. New York: Basic Books, 1989.

Pateman, Carole. *The Sexual Contract*. Stanford, CA: Stanford University Press, 1997.

Peters, Julie, and Andrea Wolper. *Women's Rights, Human Rights: International Feminist Perspectives*. New York: Routledge, 1995.

Ruddick, Sara. *Maternal Thinking: Toward a Politics of Peace*. Boston: Beacon Press, 1989.

Sargent, Lydia, ed. *Women and Revolution*. Boston: South End Press, 1981.

Secombe, Wally. "The Housewife and Her Labour Under Capitalism." *New Left Review* 83 (1973).

Shanley, Mary Lyndon, and Carole Pateman. *Feminist Interpretations and Political Theory*. University Park, PA: Penn State University Press, 1991.

Shanley, Mary Lyndon, and Uma Narayan. *Reconstructing Political Theory: Feminist Perspectives*. University Park, PA: Penn State University Press, 1997.

Waters, Kristin. *Women and Men Political Theorists: Enlightened Conversations*. Oxford: Blackwell Press, 2000.

Young, Iris Marion. *Justice and the Politics of Difference*. Princeton, NJ: Princeton University Press, 1990.

Nancy Holmstrom (2005)

FÉNELON, FRANÇOIS DE SALIGNAC DE LA MOTHE

(1651–1715)

François de Salignac de la Mothe Fénelon, the French bishop and author, was born in Périgord of an ancient noble but impoverished family. He received his education in Cahors and then in Paris, where he entered the seminary of Saint-Sulpice and was ordained priest about 1675. First in Paris and then in Saintonge he was made responsible for securing the conversion of Protestants, and in this, especially after the revocation of the Edict of Nantes (1685), he had to offset the effects of brutal military repression. He was certainly firm and successful, but opinions vary on how gentle he was. By 1689 he enjoyed the favor of Bishop Jacques Bénigne Bossuet and Mme. de Maintenon and had been appointed tutor to Louis XIV's grandson, the duc de Bourgogne.

Fénelon's association with Mme. Guyon, the exponent of quietism, dramatically changed his career. In 1694, mainly on Bossuet's initiative, she was censured by an official inquiry and temporarily put under his supervision at Meaux. Both Fénelon and Mme. de Maintenon were implicated with Mme. Guyon in a devotional group, but when Bossuet consecrated Fénelon archbishop of Cambrai in 1695 it seemed that he had averted potential scandal by using promotion as a pretext for removal. Fénelon, however, had become personally committed to mysticism and the doctrine of pure love (the disinterested love of God, divorced from any act of will, or even concern for one's salvation). Learning that Bossuet planned a crushing (and unfair) attack on Mme. Guyon and, through her, on all mysticism, Fénelon tried to forestall him by publishing a reasoned defense of mystical spirituality, *Les maximes des saints* (1679). Bossuet then embarked on a campaign of slander, falsification, and corruption, which resulted in Fénelon's banishment from the court (1697) and his condemnation by the pope (1699). Fénelon, who had always been fragile in health, remained in exile at Cambrai, conscientiously ruling his war-ravaged diocese, earning a reputation for sanctity, and pursuing a relentless, and ultimately successful, struggle against Jansenism in high places.

Though he owed much of his early success to Bossuet, whom he had at first admired, Fénelon was by temperament so different that a subsequent breach was inevitable. In his attitude to the theater Fénelon had a breadth and humanity of outlook that led him to praise Jean Racine and even Molière, who had been mercilessly attacked by Bossuet (*Lettre à l'Académie,* 1714). Fénelon had been deeply influenced by Greek culture, and much of his thinking bore the mark of Plato. He combined sensitivity and idealism with a strong vein of practicality, but he echoed neither the authoritarianism nor the moral grimness of Bossuet.

In philosophy Fénelon was enthusiastic rather than original. In 1687 he undertook for Bossuet a *Réfutation du système de la nature et de la grâce* against Nicolas Malebranche, but he soon espoused a form of Cartesianism—best represented in his *Traité de l'existence de Dieu* (1712 and 1718)—that came very close to Malebranche's position. Fénelon also wrote *Lettres sur divers sujets de métaphysique et de religion* (1718).

His early *Traité de l'education des filles* (1687) is humane and sensible, arguing that to neglect the education of one half of the human race can only have adverse effects on the other. Basing his system firmly on Christian teaching, he emphasized the need for a moral education deriving from love of virtue, rather than from fear of punishment. In addition to general literacy and elocution, Fénelon advocated the teaching of such practical matters as sufficient knowledge of law to enable women to protect their much-abused interests.

Fénelon's principle of developing rather than repressing character appears in *Télémaque,* written for his pupil about 1694 and semiofficially condemned on publication (1699). The transparent veil of Homeric legend does nothing to conceal the author's detestation of royal absolutism in its contemporary manifestations. Wars of aggression fought in the name of national prestige, territorial aggrandizement and extravagant luxury at court are condemned, not only for the misery they cause for impoverished subjects, but also as evils in themselves. For Fénelon a good king is one whose people enjoy prosperity based on industry and commerce and who accepts the duty of ensuring not only their material but also, through

his example, their moral welfare. Fénelon's fundamental political axiom was that kings and their policies are subject to and judged by the moral law, as embodied in Christian teaching, and that the true interests of a state can never conflict with this law. Similar views occur in the *Dialogues des morts.* Had it not been for the premature death of the duc de Bourgogne (1712), Fénelon's teaching, so contrary to Louis XIV's practice, might well have become official policy.

See also Bossuet, Jacques Bénigne.

Bibliography

WORKS BY FÉNELON

Oeuvres complètes. 10 vols, edited by J. Gosselin. Paris, 1848–1852.

Numerous editions of separate works.

WORKS ON FÉNELON

Adam, A. *Histoire de la littérature française au XVII^{e} siècle.* Vol. V, Ch. 5. Paris, 1956. Excellent chapter on Fénelon.

Bremond, H. *Apologie pour Fénelon.* Paris, 1910.

Carcassonne, E. *Fénelon, l'homme et l'oeuvre.* Paris: Boivin, 1946.

Goré, Jeanne-L. *L'itinéraire de Fénelon.* Paris, 1957.

Hazard, P. *La crise de la consciense européenne.* 2nd ed. Paris, 1961. Translated by J. L. May as *The European Mind.* London: Hollis and Carter, 1953; Harmondsworth, U.K. (paperback), 1964.

Kearns, Edward J. *Ideas in Seventeenth Century France.* New York: St. Martin's Press, 1979.

A. J. Krailsheimer (1967)
Bibliography updated by Tamra Frei (2005)

FERGUSON, ADAM
(1723–1816)

Born in Logierait, Scotland, to a parish minister, Adam Ferguson was educated first at the local parish school, then at grammar school in Perth, then at St. Andrews (MA 1742), and finally studied divinity at the University of Edinburgh (1743–1745). In Edinburgh he befriended many leading figures in moderate circles, including fellow divinity students Alexander Carlyle (1722–1805), William Robertson (1721–1793), and Hugh Blair (1718–1800) and older members of the Select Society including his close friend, David Hume. In 1745 he cut his studies short, was ordained, and became deputy chaplain (eventually chaplain) preaching in Gaelic to the Highland Black Watch Regiment. He returned to secular nonmilitary life in 1754 and became a mainstay of the Edinburgh intelligentsia, succeeding Hume as the librarian of the Faculty of Advocates (1758–1759), then (also with Hume's assistance) became professor of natural philosophy at the University of Edinburgh (1759–1764) and finally professor of pneumatics and moral philosophy (1764–1785).

Ferguson's international reputation was secured with the publication of his masterpiece, *An Essay on the History of Civil Society* in 1767. The *Essay* was quickly followed by the *Institutes of Moral Philosophy* (1769), a popular textbook used in moral philosophy curricula in America, Germany, and Russia. Now famous, Ferguson traveled extensively and engaged vigorously with the philosophical and political issues of his day, particularly the American Revolution, which he criticized in its revolutionary practice in a pamphlet against Richard Price (*Observations on the Nature of Civil Liberty* [1776]) and the settlement of which he sought as secretary to the Carlisle Commission (1778). Ferguson continued his publishing successes with the philosophical history *History of the Progress and Termination of the Roman Republic* (1783) and later, after his retirement from Edinburgh, the *Principles of Moral and Political Science* (1792). His intellectual engagements hardly dampened until his death, and in addition to his books he published a significant number of pamphlets.

His contemporaries were impressed by his intelligence and his distinctive temperament. Carlyle described Ferguson as having "a dignified reserve" in conversation filled with "dark allusions," and as jealous yet with a "boundless sense of humor" in private company. A nineteenth-century biographer nicknamed Ferguson the "the Scottish Cato" due to these qualities of character appropriate to the Scots advocate of republican Stoical virtue.

Like many of his contemporaries, Ferguson brought a wide range of scientific, anthropological, and historical resources to bear on moral and politics in a characteristically Scottish fusion of mid- and late Stoicism, natural law theory, history, natural science, and the natural sciences of man (including pneumatics or the physical history of mind). His *Essay on Civil Society* was built on a stadial theory that divided human societies according to their means of subsistence, social organization, and equality (among other variables). At the same time, Ferguson stressed that although morals should be fully informed by natural science and social history, it had a special provenance: what one ought to do in regard to good and evil and virtue and vice.

So far, nothing in Ferguson's theory was unique and he drew on many of his Scottish contemporaries—

notably Hume, Adam Smith, and Thomas Reid—for his arguments. What was distinctive was how Ferguson used this framework to think about the relation between morals and politics. For Ferguson virtue was thoroughly intertwined with political *virtú* in the tradition of Niccolò Machiavelli and Baron de Montesquieu. Francis Hutcheson had stressed the civic and social character of morality, but Ferguson drew on Montesquieu's arguments in *Spirit of the Laws* (1748), that laws and social institutions create a virtuous citizenry, and on his definition of political liberty as virtuous action in and through good laws, to interweave civic morality with the new sciences of man. For Ferguson, like Montesquieu, the growth of virtue was neither isomorphic with material progress nor necessarily antithetical to it: Virtue can be found in different times and places. But unlike Montesquieu and like Smith, John Millar, and numerous other Scots, he always assumed in the background a theory of historical stages, not as linear progress but as a means to analyze nations and peoples both synchronically and diachronically, and as a species of conjectural history to be used as an analytic framework for comparing progress, wealth, equality, virtue, and other variables. On the one hand, the optimal setting for virtue and equality was a small, republican meritocracy of social and political equals actively contributing to the common good. On the other hand, Ferguson also stressed that ancient, simple military societies tended to be impoverished, violent, and "rude," lacking many of the sociable virtues admired in a commercial society. The problem was, then, given the different forces that can affect a nation morally—its size, its prosperity, its historical stage, and its laws—how to maximize virtue and minimize vice?

Ferguson's diagnosed this problem as endemic to his contemporaries thinking about morals and politics. Hume (and later Smith and Millar) argued that commerce was a fundamental civilizing force and gave rise to a liberal progressive society superior to societies that preceded it. Still, Hume recognized the virtue of small, egalitarian societies. Ferguson thought that Hume and Smith confused material prosperity with wealth and this showed in their moral recommendations. Obviously, material prosperity was desirable, and once attained it was difficult to forego, but prosperous nations are often corrupt and there was no guarantee from the progress of history that they would not become luxurious and despotic. The focus should be on a broader conception of wealth that included moral and political virtue.

So what sorts of laws and civic institutions prevent moral corruption and reinforce virtue in large, wealthy societies? Ferguson focused throughout his career in his books and pamphlets on the importance of citizens' militias, that is, defense by ordinary citizenry as opposed to professional soldiers. His service in the Black Watch during the Jacobite uprising of 1745–1746 made him aware firsthand of the difficulties a standing army in a commercial society had in quelling rude but fierce Highland militias. Most of the Edinburgh intelligentsia—including Smith and Hume—supported a Scottish militia. Ferguson thought that the issue was philosophically pivotal and that Smith's lukewarm support for the militia was a symptom of the conflict in his theory between virtue and wealth. He believed that militias are paradigmatic egalitarian, socially activist institutions. Any soldier can rise in a militia through merit, and military and social virtue are rewarded and reinforced in local organizations where citizenry know one another, rely on one another, and are responsible for their actions. Complex, prosperous societies need such invigorating, egalitarian social institutions to be wealthy in a broader sense, to avoid moral corruption, and so to be vigilant against tyranny. They also are a bulwark against the deadening effect of the division of labor, which is driven forward by commerce but not morality. Active social institutions allow the moral vigor of rude society, above all the early Roman republic, to be infused in commercial societies when people cannot, or even do not want to, return to a prior state.

Ferguson's works were particularly popular in Italy, France, and Germany and influenced, among others, Gottfried Lessing, Christian Garve, and Friedrich Schiller. He also influenced Karl Marx in particular (with his criticisms of progressivism and the division of labor) and modern sociology in general, above all through the proliferation of the idea of civil society.

See also Garve, Christian; Hume, David; Hutcheson, Francis; Lessing, Gotthold Ephraim; Machiavelli, Niccolò; Marx, Karl; Montesquieu, Baron de; Natural Law; Price, Richard; Reid, Thomas; Schiller, Friedrich; Smith, Adam; Stoicism.

Bibliography

Oz-Salzberger, Fania. *Translating the Enlightenment: Scottish Civic Discourse in Eighteenth-Century Germany*. New York: Oxford University Press, 1995.

WORKS BY ADAM FERGUSON

The Correspondence of Adam Ferguson, edited by Vincenzo Merolle. Brookfield, VT: William Pickering, 1995a.

An Essay on the History of Civil Society (1767), edited by Fania Oz-Salzberger. New York: Cambridge University Press, 1995b).

Collection of Essays, edited by Yasua Amoh. Kyōto, Japan: Rinsen Books, 1996.

WORKS ABOUT ADAM FERGUSON

Burton, John Hill, ed. *The Autobiography of Dr. Alexander Carlyle of Inveresk: 1722–1805*. London: T. N. Foulis, 1910.

Kettler, David. *The Social and Political Thought of Adam Ferguson*. Columbus: Ohio State University Press, 1965.

Oz-Salzberger, Fania. "Adam Ferguson." In *Dictionary of National Biography*. New York: Oxford University Press, 2004.

Sher, Richard B. "Adam Ferguson, Adam Smith, and the Problem of National Defense." *Journal of Modern History* 61 (1989): 240–268.

Aaron Garrett (2005)

FERGUSON, ANN

(1938–)

Ann Ferguson, a socialist-feminist philosopher (PhD, Brown University, 1965; BA, Swarthmore College, 1959) teaches philosophy and women's studies at the University of Massachusetts at Amherst. Her political support for a democratic socialism grew out of sustained involvement with the civil-rights movement, the anti–Vietnam War movement, the new left, and the women's liberation movement in the United States.

Ferguson is best known for her critique of male dominance and her formulation of the concept of sex/affective production (1989). She contends that Marxist accounts of class oppression and radical feminist accounts of heterosexist exploitation do not properly account for (a) the social energies involved in parenting, sexuality, and affective bonding and (b) the unequal, exploitative production and exchange of services between men and women in a patriarchal society (1991). Critiquing Sigmund Freud, Ferguson claims that affective bonding and sexual desires aim primarily not at biological reproduction but rather at connecting with other humans, queer or straight.

In early work, Ferguson highlighted women's potential as a revolutionary class. In *Sexual Democracy* (1991), she developed a materialist-feminist multisystems theory of oppression: that race, class, and gender function as dominant, semi-independent categories, and thus that the ideal of sisterhood is obstructed by race, caste, class, and sexual identities. Her advocacy of "gynandry" (1991), a play on "androgyny" (1977), not only critiques the ideology of the theory that gender roles naturally complement each other, but also calls for revaluing feminine strengths and for building a society free of patriarchal oppression. In her vision, the feminine is not a fixed gender trait. In her important aspect theory of the self, Ferguson noted that it is misguided to speak of one essential core self; it is more helpful to note that "one's sense of self and … values" are context-dependent (1991, p. 105).

Expanding on her aspect theory of the self, Ferguson (1996) proposed building bridge identities as a strategy to counter positive- and essentialist-identity politics. Bridge identities "attempt to refuse the fixed identities given us by gender, race, class, and sexual differences" (1998a, p. 207) and reconstitute identities politically (1998b). For instance, when a feminist researcher from the global North wishes to network with people in the global South who are relatively disadvantaged, by self-questioning she can put her privileged position in check even to the point of destabilizing her identity. But by building a bridge identity, she can begin to recognize participants as *subjects* of resistance rather than as *objects* of knowledge (1998b). Ferguson (1998a) argues for a transitional feminist morality in which prostitution is defined as a morally risky practice, rather than, as most feminists define it, as a morally forbidden practice. In formulating a viable feminist ethico-politics, she affirms the political stance of subjects of resistance: sex workers who demand unionization and decriminalization. With a bridge-identity politics that refuses fixed identities of race, gender, etc., a feminist coalition could consistently support sex workers' rights locally and oppose trafficking in women internationally.

Ferguson exudes a passion for feminist coalitional and solidarity work with people who face marginalization due to capitalist, racist, or patriarchal forces. Her work is informed par excellence by the rich dialectical interplay of theory and practice.

See also Feminist Philosophy; Feminist Social and Political Philosophy; Marxist Philosophy; Social and Political Philosophy.

Bibliography

WORKS BY FERGUSON

"Androgyny as an Ideal for Human Development." In *Feminism and Philosophy*, edited by Mary Vetterling-Braggin, Frederick Elliston, and Jane English. Totowa, NJ: Rowman and Allanheld, 1977.

Blood at the Root: Motherhood, Sexuality, and Male Dominance. London: Pandora, 1989.

Sexual Democracy: Women, Oppression, and Revolution. Boulder, CO: Westview Press, 1991.

"Bridge Identity Politics: A Feminist Integrative Ethics of International Development." *Organization* 3(4) (1996): 571–587.

"Prostitution as a Morally Risky Practice." In *Daring to Be Good: Feminist Essays in Ethico-politics*, edited by Bat-Ami Bar On and Ann Ferguson. New York: Routledge, 1998a.

"Resisting the Veil of Privilege: Building Bridge Identities as an Ethico-politics of Global Feminism." *Hypatia* 13 (3) (1998b): 95–113.

Mechthild Nagel (2005)

FERRARA, FRANCIS SYLVESTER OF

See *Sylvester of Ferrara, Francis*

FERRI, LUIGI

(1826–1895)

Luigi Ferri, the Italian epistemologist and historian of philosophy, was born in Bologna. He studied at Paris and was professor of the history of philosophy at Florence and at Rome. A self-styled disciple of Terenzio Mamiani, Ferri contributed to Mamiani's journal, *La filosofia delle scuole italiana,* and continued editing the journal, under the title *Rivista italiana di filosofia,* from the death of Mamiani in 1885 until his own death in Rome in 1895.

Ferri's philosophizing moved within the framework of Italian ontologism, which saw in man the capacity for a direct and "intuitive" relationship with the Absolute (Being or God), but his interest focused principally on the psychological conditions in which this relationship takes shape. His investigations, therefore, had as their object man's interior experience, the "inner (or intimate) sense" of which Maine de Biran spoke. To the latter Ferri owed his basic inspirations. Reproving associationist psychology for reducing the spirit, or self, to an associative mechanism that takes no account of the activity of consciousness, Ferri tried to bring to light the function of this activity. He saw this activity as a kind of force or energy that "by making itself its own object, determines its modes according to rules proper to itself, proposes goals, directs and oversees its own work, and frees itself finally from the influence of sensation and emotive impressions so as to find truth with the intellect and to reproduce in itself, with ideas and the evidence of experience, the world of phenomena."

Ferri used the term *dynamism* to refer to the conception that the substance of both the physical and the spiritual worlds is energy and that in both of these worlds energy is regulated by the same laws of conservation. Thus there is a "permanence in the quantity, quality, and relationships of the spiritual world" just as there is a permanence in the amount of matter and energy. Ferri also held that only the energy regulating the spiritual world is known or immediately given to man in the act of consciousness; the actions of energy operating in the external world are known to man only indirectly, that is, by the effects they have upon this act through sense perception. The unity of the universal energy is, however, the sole theme of metaphysics.

See also Absolute, The; Energy; History and Historiography of Philosophy; Italian Philosophy; Maine de Biran; Ontology.

Bibliography

WORKS BY FERRI

Essai sur l'histoire de la philosophie en Italie au dix-neuvième siècle, 2 vols. Paris, 1869. Traces the development of Italian philosophy from the sensationalism of the eighteenth century to the ontologism or idealism of the nineteenth, represented by Rosmini, Gioberti, and Mamiani.

La psicologia di Pietro Pomponazzi. Rome, 1877. Contains a previously unpublished commentary by Pomponazzi on Aristotle's *De Anima.*

La psychologie de l'association depuis Hobbes jusqu'à nos jours, histoire et critique. Paris, 1883.

In the *Atti dell'Academia dei Lincei,* of which Ferri was a member: *Analisi del concetto di sostanza e sue relazione con i concetti di essenza, di causa, e di forza; contributo al dinamismo filosofico* (1885); *Il fenomeno sensibile e la percezione esteriore ossia i fondamenti del realismo* (1886); *Dell'idea del vero e sue relazione con l'idea dell'essere* (1887); *Dell'idea dell'essere* (1888).

WORKS ON FERRI

Barzellotti, G. "Luigi Ferri." *Nuova Antologia* (1895).

Gentile, G. *Le origini della filosofia contemporanea in Italia,* Vol. 1, pp. 215–233. Messina, 1917.

Nicola Abbagnano (1967)
Translated by Nino Langiulli

FERRIER, JAMES FREDERICK

(1808–1864)

James Frederick Ferrier, the Scottish metaphysician, was born in Edinburgh into a wealthy family of lawyers. After studying at the universities of Edinburgh and Oxford, he spent some months in Germany. He settled in Edinburgh

in 1832 as an advocate, becoming active in the intellectual circle of Sir William Hamilton, which included Thomas De Quincey and "Christopher North" of *Blackwood's Magazine.* Under this stimulus Ferrier contributed to *Blackwood's* between 1838 and 1843 the eleven long articles that fill most of the second volume of his *Lectures and Remains* (2 vols., Edinburgh and London, 1866). In 1845 he was appointed professor of moral philosophy and political economy at the University of St. Andrews. Ferrier issued a drastically revised version of his philosophy in the *Institutes of Metaphysic* (Edinburgh and London, 1854; 2nd ed., 1856). The *Institutes* was to some extent affected by Ferrier's commitments in the political and ecclesiastical struggles that then divided Scotland. This social influence is still more marked in the pamphlet defending his position, *Scottish Philosophy, the Old and the New* (Edinburgh, 1856). Meanwhile, Ferrier elaborated, until incapacitated in 1861, on an impressive series of lectures on Greek philosophy, posthumously published as Volume I of *Lectures and Remains.*

The first seven *Blackwood's* articles constitute a unitary work on the philosophy of consciousness. Its starting point is a critique of Thomas Brown's doctrine that it is wrong to regard states of mind, such as emotions, as objects of consciousness. Brown argued that to speak of being conscious of feeling angry is the same thing as to speak of feeling angry. Ferrier pointed out that there is a marked difference between speaking of someone as boiling with rage and speaking of him as being conscious of the boiling rage within him. In the latter case, instead of looking outward at the injustice and brooding on the affront, he looks inward at the consequent irritation in his heart and ceases to brood.

Thus far Ferrier was merely making an intelligent use of the doctrine of the inverse variation of feeling and knowledge proposed by his friend Sir William Hamilton. But as Hamilton noted with approval, Ferrier then went beyond the customary limits of British philosophy by asking what is involved in the shift from unself-conscious anger to self-conscious anger. This self-knowledge does not arise straightforwardly out of ordinary experience. The use of the first personal pronoun, which is the mark of self-knowledge in the proper sense, is something that cannot be learned from the experience of other people and their talk in the same imitative way as the use of a word like *table* can. The indubitability of self-knowledge arises just because it is not based on observation in the same way that our knowledge of mountains is. Therefore, Ferrier concluded, there is something anomalous about the foundations of self-knowledge. What is it?

In his four *Blackwood's* articles on the subject of sense perception, contributed between 1841 and 1843, Ferrier gave his problem a definite form by limiting it. To gain light on the nature of self-knowledge he looked into the foundations of the ordinary distinction between act of sense and object of sense. Ferrier's discussion is brilliantly original. The key to the difficulty is that as long as we view each sense field in isolation, no proper distinction can be drawn between the act and the object of sense. Within the visual field alone vision does not stand out as empirically separable from the colors seen; within the tactual field the effort of feeling presents itself as indistinguishable from the solids felt. But when the sense fields are viewed in correlation with one another, seeing separates itself from the colors seen as being connected with something tangible but not visible: the eye. Similarly, feeling distinguishes itself from solidity by being vested in an organ of touch revealed by vision rather than by touch. Ferrier thus argued that the key to self-experience is the peculiar experience of appropriating one's own body in the sense of correlating one's own sense organs. This is reminiscent of Maurice Merleau-Ponty and Jean-Paul Sartre. Like them, Ferrier developed the theme of human freedom, first by reference to the contrast between reflective experience and prereflective experience, then by reference to the contrast between the experience of one's own body and the experience of foreign bodies.

Ferrier was stimulated by Friedrich Schelling and G. W. F. Hegel, but there is a distinctive originality to his position in his attempt to give life and definiteness to their ideas by viewing them in terms of the problems of philosophy posed by Hamilton and Thomas Brown. As De Quincey said, Ferrier's philosophy is "German philosophy refracted through a Scottish medium."

Ferrier's highly original early efforts have been overshadowed for posterity by the respectable academic contributions of his later life. In his *Institutes of Metaphysic* he moved from a "phenomenological" standpoint, inherited from Thomas Reid by way of Hamilton and Victor Cousin, to a narrowly a priori point of view which, distinguishing sharply between necessary and contingent truth, would restrict philosophy to necessary truth. As a result, the *Institutes of Metaphysic* omits the analysis of self-knowledge and the experience of one's own body that distinguishes the *Blackwood's* articles, confining itself to well-worn doctrines that can be expounded in an a priori way, such as the Cartesian *cogito* and a verifiability principle not unlike that of modern positivism. But Ferrier's later work should not be underestimated. It contains remarkably illuminating discussions of the relations of

universals and particulars (rather like that in Henry Mansel), which is carried further in the *Lectures on Greek Philosophy.* In this work there is also an extremely impressive analysis of the experience of change and movement that in one way anticipates Henri Bergson and in another way looks back to Hegel.

Ferrier's later work was very influential in the late nineteenth century in the English-speaking world and to some extent in France. In particular, the *Institutes of Metaphysic* provided Shadworth Hodgson with his starting point and most of his leading ideas. Ferrier's early work, unfortunately, escaped notice in the nineteenth century, but a reevaluation of it has begun.

See also Bergson, Henri; British Philosophy; Brown, Thomas; Cousin, Victor; Hamilton, William; Hegel, Georg Wilhelm Friedrich; Hodgson, Shadworth Holloway; Hume, David; Mansel, Henry Longueville; Merleau-Ponty, Maurice; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von; Verifiability Principle.

Bibliography

WORKS BY FERRIER

Ferrier's works are collected in the three-volume *Philosophical Works* (Edinburgh: Blackwood, 1875).

WORKS ON FERRIER

Arthur Thomson's article "The Philosophy of J. F. Ferrier," in *Philosophy* 39 (1964): 46–62, reevaluates Ferrier's early work. There is a short biography by E. S. Haldane, *James Frederick Ferrier* (Edinburgh, 1894). See also *The Democratic Intellect,* by G. E. Davie (Edinburgh: Edinburgh University Press, 1961).

George E. Davie (1967)

FEUERBACH, LUDWIG ANDREAS
(1804–1872)

Ludwig Andreas Feuerbach, the German philosopher, theologian, and moralist, was born in Landshut, Bavaria. He studied theology at Heidelberg and Berlin and then, in 1825, under the influence of G. W. F. Hegel, transferred to the faculty of philosophy. He received his doctorate in 1828 at Erlangen, where he remained to teach as docent until 1832. In 1830 he published anonymously at Nuremberg a work—*Gedanken über Tod und Unsterblichkeit*—that created a minor scandal by interpreting Christianity as an egoistic and inhumane religion. When his authorship of this book became known, he was dismissed from the faculty. In 1836 he retired to Bruckberg, where he lived on a modest pension from the Bavarian government, income from his writings, and revenue provided by his wife's interest in a pottery factory.

Between 1836 and 1843 he collaborated with Arnold Ruge on Ruge's *Hallische Jahrbücher für deutsche Wissenschaft und Kunst*, in which many of Feuerbach's most important early writings on religion and philosophy first appeared. He broke with Ruge when the latter began collaboration with Karl Marx on the *Deutsch-Französische Jahrbücher*, although he contributed to the one issue of that journal. He reappeared briefly in academic life in 1848–1849, lecturing to audiences of intellectuals and workers at Heidelberg at the request of students, for whom he had become a symbol of liberal thought.

With the failure of the Frankfurt Assembly and the defeat of liberalism in Germany, Feuerbach retired once more to Bruckberg, where he devoted himself to the study of the natural sciences, the composition of a monumental *Theogonie* (Leipzig, 1857), and a voluminous correspondence with friends and admirers all over Europe. In 1860 his wife's pottery factory failed, and Feuerbach removed his family to Nuremberg, where he was forced to live off the generosity of his friends. In 1867 he suffered the first of a number of strokes that finally killed him.

WORKS

Feuerbach's most important works—"Zur Kritik der Hegelschen Philosophie" (in the *Hallische Jahrbücher*, 1839), *Das Wesen des Christentums* (Leipzig, 1841; translated by M. Evans [George Eliot], London, 1854), *Grundsätze der Philosophie der Zukunft* (Zürich and Winterthur, 1843), and *Das Wesen der Religion* (Leipzig, 1846)—were produced in his early years. They were meant to expose the contradictions in Hegelian philosophy, to establish the "illusionistic" character of all religious belief, and to plead for a "new philosophy," based on anthropology and physiology, that would provide the foundation of a naturalistic-humanistic ethic. His criticism of Hegelianism served as the point of departure for the so-called left Hegelians, of whom Marx and Friedrich Engels were the most important representatives.

CRITICISM OF HEGELIANISM. Feuerbach's critique of Hegelianism proceeded not from sympathy for "obtuse materialism," under which term he grouped Newtonian science, empiricism, and positivism alike, but rather from his discovery of contradictions in Hegel's own system. The resolution of these contradictions would, he

believed, allow the establishment of a "new philosophy," which, while remaining thoroughly materialistic, would accommodate those insights into the operations of human consciousness that constituted Hegelianism's definitive contribution to human self-knowledge.

Feuerbach viewed Hegelianism as the culmination of modern rationalism, and he believed that "the secret of Hegel," as of all rationalism, lay in an essentially religious spirit concealed beneath an apparent denial of all transcendence. This hidden religious element accounted for the degradation of the material world, of man, and of the senses that was characteristic of Hegel's metaphysics, ethics, and epistemology, respectively. In Hegel's thought, however, the means were provided for finally transcending all of the religious residues in modern philosophy. For Hegel's attempt to sustain simultaneously the primacy of intellect and the necessity of reason's realizing itself in matter results in the negation of the Hegelian system itself in the interest of a materialistic metaphysics, a humanistic ethics, and a sensible (*sinnliche*) epistemology, the bases of the "philosophy of the future."

DEVELOPMENT OF MODERN PHILOSOPHY. Feuerbach believed that modern philosophy had followed a pattern of development set by theology. The attempt of theology to establish the relationship between the sensible attributes of God and the extrasensible sphere in which he exists necessarily led to pantheism, which makes matter an attribute of God or defines God (as did Benedict de Spinoza) as "extended essence" and thus ends by deifying matter itself. In fact, pantheism is "theological atheism," the discovery by theology that matter is the sole reality, and hence it foreshadows the ultimate self-dissolution of religion.

Empiricism had already discovered that matter was the sole reality, but only in a practical, not in a theoretical, sense, for in making "mere" matter the sole reality it was unable to deal with the data of human consciousness. Rationalism, however, of which idealism was the necessary outcome, underwent a secularized development from theism as a divinization of spirit to pantheism as the self-dissolution of spirit. Idealism was nothing but an attempt to salvage God by vesting full epistemic authority in consciousness, intellect, or reason at the expense of the senses. Yet because it was overtly secularist, rationalism had to account for the world discovered by the senses. It could do this only by affirming, as Immanuel Kant did, an absolute hiatus between the world of intellect, to which it ascribed all truth, and the world of sense, to which it granted reality. Hegel tried to close this gap between truth and reality, but he could do so only by extending the Cartesian divinization of Reason to the world as a whole. The result was a transition from Kantian "rational theism, theism rationalized" to Hegelian "pantheistic idealism."

REASON IN HEGEL. In affirming the rationality of the real and the reality of the rational, Hegel, according to Feuerbach, elevated reason to the status of "absolute essence." Then, to account for the existence of the spatiotemporal world, he had to hold simultaneously that matter is the negation of thought and that thought can only "realize itself" by becoming matter. To Feuerbach this showed that on Hegel's own terms "thought presupposes, without being aware of it, that truth is reality, sensibility independent of thought." On the one hand Hegel viewed sensibility as "an attribute of the idea," whereas on the other he maintained that it is "an attribute without which thought has no truth"; that is, he had to hold that it is "at one and the same time central and marginal, essence and accident."

According to Feuerbach, idealism knew implicitly that "truth, reality, and sensibility are identical," but it suppressed this truth in order to subordinate the sensible world to an absolute being endowed with the attributes of the human ego, that is, with consciousness and reason. This led idealism to assert that the thinking of the absolute being is real, whereas that of the finite sensible being, man, is not. According to Hegel, human reason is nothing but the self-revelation of the absolute being to itself. Thus, Feuerbach exclaimed, Hegel "alienates and expropriates from man his typical essence and activity!"

PRIMACY OF HUMAN CONSCIOUSNESS. Feuerbach's own "new philosophy" began with the axiom "Only a sensible being is a real, true being," standing the Hegelian position on its feet so that its truth could be seen aright. "The true relation of thought to being is only this," he wrote in the *Vorläufige Thesen*: "being is the subject, thought the predicate. Thought is a product of being, not being of thought. … The essence of being as being is the essence of nature." The consciousness deified by Hegel, like the reason deified by René Descartes and Kant and the Matter deified by Spinoza, "is our ego, our intellect, our essence: and this God is no God in itself, but only the appearance of ourselves to ourselves." Hence, the lasting contribution of idealism to philosophy is its analysis, under the aspect of an examination of the absolute being, of the operations of human consciousness, the reality of which is denied by simple empiricism. Hegelianism, like all metaphysics, is nothing but "esoteric psychology."

MATERIALISM AND IDEALISM. Unlike conventional materialism the new philosophy granted ontological and epistemological status to consciousness and intellect, and unlike idealism it accorded reality to matter. But it deified neither matter nor consciousness. For according to Feuerbach, it is wrong to say, with the materialist, that "man is distinguished from the brute *only* by consciousness"; in fact, "in a being which awakes to consciousness, there takes place a qualitative change, a differentiation of the entire nature." Yet this "qualitative change" in no way justifies the idealist contention that man is consciousness alone, "for as man belongs to the essence of Nature,—in opposition to common materialism; so Nature belongs to the essence of man,—in opposition to subjective idealism."

MAN. Every attempt to specify the essence of man by deriving his material from his spiritual nature, or vice versa, is therefore mistaken, in Feuerbach's view. The task of philosophy is to encounter man in his situation, as that part of nature endowed with consciousness which seeks to realize its own peculiar essence through specific kinds of relationships with the rest of nature and with other members of its species. Feuerbach's philosophy assumed only that "I am a real, sensible essence: the body is constituted of my essence; indeed the body in its totality is my ego, my existence itself." It recognized that man's essence reveals itself quintessentially in the impulse toward union with other men: "The essence of man is contained only in community, in the unity of man with man—a unity which however is founded only on the reality of the differences between I and thou." To comprehend human action and thought one must take account of man's capacity to transcend the limited responses of the lower animals to their environment.

Philosophy, properly studied, then, is "the complete, coherent, and absolute resolution of theology into anthropology.… "It takes man as the culmination of the natural process and defines him as "universal essence" and then concentrates on the study of the totality of his responses to the rest of the world. Among these responses will be found the passions, especially the emotion of love, the impulse toward "union" with the "other" that is peculiar to man. The capacity to create communities of shared emotive contents is the secret of man and therefore the secret of all thought and action; for what men are really seeking in every imagined absolute is nothing but the "unity of I and thou."

RELIGION. All of this is assumed in Feuerbach's studies of religion and lies at the base of his "unmasking" of Christian beliefs in *Das Wesen des Christentums*, his most celebrated work.

Feuerbach regarded religion as one of the forms of human thought and action by which man raised himself above the animal. Beginning with the assumption of D. F. Strauss that religion, myth, ritual, and dogma tell us more about the inner lives of individual people than about their presumed object of worship, Feuerbach tried to determine the purely human significance of all mythological thought. He professed to be a uniformitarian in religious matters—that is, he denied that past religious experiences differ from those that can be observed in the present—thus anticipating the approach to religious experience of both William James and Sigmund Freud. Like them, he claimed to be rigidly empirical in method. "I found my ideas on materials which can be appropriated through the senses," he wrote in the 1843 preface to *Das Wesen des Christentums*; "I do not generate the object from the thought, but the thought from the object; and I hold that alone to be a proper object which has an existence beyond one's brain.… I am nothing but a natural philosopher in the domain of the mind."

His study led him to conclude that religion is a form of the projective spirit in man, the means by which man "projects his being into objectivity, and then again makes himself an object to this projected image of himself thus converted into a subject; he thinks of himself not as an object to himself but as an object of an object, of another being than himself." Thus, religion is "the dream of the human mind"; properly understood, it is a dream of human, not divine, development: "it is and can be nothing else than the consciousness which man has of his own—not finite and limited—but infinite nature." Man, then, unlike the animal, is self-transcending, and religion is one of man's means of objectifying his own essence in ideal terms, of spinning out visions of what he might be. For example, the Christian idea of the Incarnation is nothing but a reflection of the dream of man to become God and the realization that this can be achieved only through a transcendent love of one's fellow man.

Religious feelings thus depend on an alienation of man from himself. Religion generates belief in an objective "other" in which all of man's best qualities are vested, his worst qualities being designated as the true human essence. Philosophy must therefore "destroy an illusion" that deprives man of the power of a free life as well as a genuine sense of truth and virtue, "for even love, in itself the deepest, truest emotion, becomes by means of religiousness merely ostensible, illusory, since religious love gives itself to man only for God's sake, so that it is given

only in appearance to man, but in reality to God." In short, for Feuerbach religion is the uncontrolled and unconscious exercise of a human faculty that with the aid of the sciences of anthropology, physiology, and psychology can be controlled, raised to consciousness, and turned toward the attainment of genuine health, well-being, and community here on earth. For "the consciousness of God is nothing but the consciousness of the species."

INFLUENCE

Feuerbach was little concerned with political polemics, for which Marx and Engels vehemently criticized him, but his work served as an inspiration for those who were trying to work out a realistic program of reform in Germany during the middle decades of the century. Many of his dicta became dogmata for the radical movement, as for example the 1850 statement: "The doctrine of foods is of great ethical and political significance. Food becomes blood, blood becomes heart and brain, thoughts and mind-stuff. Human fare is the foundation of human culture and thought. Would you improve a nation? Give it, instead of declamations against sin, better food. Man is what he eats" (quoted in Höffding, *History of Modern Philosophy*, London, 1900, Vol. II, p. 281). But his main concern remained the mystery of the transformation of "human fare" into human thought. This mystery was the basis of his naturalistic humanism, which Marx and Engels regarded as merely a vestige of the old idealism. According to Marx's "Theses on Feuerbach," Feuerbach resolved "the essence of religion into the essence of *man*," and Marx protested that "the essence of man is no abstraction inherent in each separate individual. In its reality it is the ensemble of social relations." The judgment was basically correct. Feuerbach, though he resolved Hegelianism into psychology, made of consciousness itself a mystery, if not a miracle.

By 1850 Feuerbach's star had already set. The future of materialism in Germany lay with mechanists such as Ludwig Büchner on the one hand and with Marx on the other. Engels was right in saying, "With one blow, [Feuerbach] pulverized the contradiction [of idealism] and without circumlocutions … placed materialism on the throne again." But he was also right in noting that Feuerbach "stopped halfway; the lower half of him was materialist, the upper half idealist." Feuerbach's "destruction" of Hegelianism was less important than the way he carried it out, since this destruction was the sport of almost every significant thinker in the Germany of his day. But because he generated materialism out of Hegel himself, Feuerbach provided the means by which German thought could become "scientific" while still indulging its overriding interest in historical processes. Thus, his work inspired both Marx and Engels, but it also laid the foundation for that phenomenological anthropology that has made him a source of information and insights for such modern philosophers as Martin Heidegger, Jean-Paul Sartre, and Karl Barth.

See also Alienation; Barth, Karl; Büchner, Ludwig; Empiricism; Engels, Friedrich; Freud, Sigmund; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heidegger, Martin; Idealism; James, William; Kant, Immanuel; Marx, Karl; Materialism; Pantheism; Philosophical Anthropology; Rationalism; Sartre, Jean-Paul; Spinoza, Benedict (Baruch) de; Strauss, David Friedrich.

Bibliography

WORKS BY FEUERBACH

Ludwig Feuerbach Sämtliche Werke. 10 vols., edited by Wilhelm Bolin and Friedrich Jodl. Stuttgart: Frommann, 1903–1910. New and augmented ed., 12 vols. Stuttgart, 1959–1960. The complete works of Feuerbach, accompanied by a full *apparatus criticus* and biographical and bibliographical information on writings both by and about him.

The Essence of Christianity, translated by M. Evans. New ed. New York: Harper, 1957. Contains a chapter on Feuerbach from Barth's *Theologie und die Kirche*.

WORKS ON FEUERBACH

Barth, K. *Die protestantische Theologie im 19. Jahrhundert*. Zürich, 1952. Translated by Brian Cozens as *Protestant Thought from Rousseau to Ritschl*. New York: Harper, 1959.

Barth, K. *Die Theologie und die Kirche*. Vol. II, pp. 212–239. Zürich, 1928. Translated by L. P. Smith as *Theology and Church*. London: SCM Press, 1962.

Chamberlain, W. B. *Heaven Wasn't His Destination: The Philosophy of Ludwig Feuerbach*. London: Allen and Unwin, 1941.

Cornu, A. *Moses Hess et la gauche hégélienne*. Paris, 1934.

Engels, F. *Ludwig Feuerbach und der Ausgang der klassischen deutschen Philosophie*. Stuttgart, 1888. Translated as *Ludwig Feuerbach and the Outcome of Classical German Philosophy*, edited by C. P. Dutt. New York: International, 1934.

Grégoire, F. *Aux Sources de la pensée de Marx*. Paris, 1947.

Grün, K. *Ludwig Feuerbach in seinem Briefwechsel und Nachlasse*. 2 vols. Leipzig, 1874.

Hook, S. *From Hegel to Marx*. New York: Reynal and Hitchcock, 1936.

Jodl, F. *Ludwig Feuerbach*. Stuttgart: Frommans, 1904.

Lévy, A. *La philosophie de Feuerbach*. Paris, 1904.

Löwith, K. *Von Hegel zu Nietzsche*. Stuttgart, 1941.

Marcuse, H. *Reason and Revolution*. London: Oxford University Press, 1941.

Nüdling, Gregor. "Die Auflosung des Gott-Menschenverhältnis bei Ludwig Feuerbach." In *Der Mensch vor Gott*. Düsseldorf, 1948.

Nüdling, Gregor. *Ludwig Feuerbachs Religionsphilosophie*. Paderborn, Germany, 1936.

Rawidowicz, S. *Ludwig Feuerbachs Philosophie*. Berlin: Reuther and Reichard, 1931.

Tucker, R. *Philosophy and Myth in Karl Marx*. Cambridge, U.K.: Cambridge University Press, 1961.

OTHER RECOMMENDED WORKS

Amengual, Gabriel. *Crítica de la religión y antropología en Ludwig Feuerbach: La reducción antropológica de la teología como paso del idealismo al materialismo*. Barcelona: Laia, 1980.

Barata-Moura, José, and Viriato Susamenho Marques. *Pensar Feuerbach: Colóquio comemorativo dos 150 anos da publicação de "A essência do Cristianismo (1841–1991)*. Lisboa: Edições Colibri, 1993.

Braun, Hans-Jürg. *Ludwig Feuerbach und die Philosophie der Zukunft: Internationale Arbeitsgemeinschaft am ZIF der Universität Bielefeld 1989*. Berlin: Akademie-Verlag, 1990.

Cabada Castro, Manuel. *El humanismo premarxista de Ludwig Feuerbach*. Madrid: La Editorial Católica, 1975.

Feuerbach, Ludwig. *Anthropologischer Materialismus. Ausgewählte Schriften.* Edited by Alfred Schmidt. Wien: Europa Verlag, 1967.

Feuerbach, Ludwig. *The Essence of Faith According to Luther*. New York: Harper & Row, 1967.

Feuerbach, Ludwig. *The Fiery Brook; Selected Writings of Ludwig Feuerbach.* Garden City, NY: Anchor Books, 1972.

Feuerbach, Ludwig. *Geschichte der neueren Philosophie: Von Bacon von Verulam bis Benedikt Spinoza*. Frankfurt am Main: Röderberg-Verlag, 1976.

Feuerbach, Ludwig. *Grundsätze der Philosophie der Zukunft*. Frankfurt am Main: Vittorio Klostermann, 1983.

Feuerbach, Ludwig. *Kleine Schriften*. Frankfurt am Main: Suhrkamp, 1966.

Feuerbach, Ludwig. *Lectures on the Essence of Religion.* New York: Harper & Row, 1967.

Feuerbach, Ludwig. *Philosophische Kritiken und Grundsätze.* Leipsig: Reclam, 1969.

Feuerbach, Ludwig. *Principles of the Philosophy of the Future*. Indianapolis, IN: Hackett, 1986,.

Feuerbach, Ludwig. *Principles of the Philosophy of the Future.* Translated, with an introduction, by Manfred H. Vogel. Indianapolis, IN: Bobbs-Merrill, 1966.

Feuerbach, Ludwig. *Schriften aus dem Nachlass.* Darmstadt: Wissenschaftliche Buchgesellschaft, 1974-1976.

Feuerbach, Ludwig. *Thoughts on Death and Immortality: From the Papers of a Thinker, along with an Appendix of Theological-satirical Epigrams*. Berkeley: University of California Press, 1980.

Feuerbach, Ludwig. *Werke* Frankfurt am Main: Suhrkamp, 1975-.

Harvey, Van Austin. *Feuerbach and the Interpretation of Religion*. Cambridge U.K.; New York: Cambridge University Press, 1995.

Kamenka, Eugene. *The Philosophy of Ludwig Feuerbach.* New York: Praeger, 1970.

Ludwig Feuerbach e la natura non umana: ricostruzione genetica dell'Essenza della religione con pubblicazione degli inediti. Firenze: La Nuova Italia, 1986.

Solidarität oder Egoismus: Studien zu einer Ethik bei und nach Ludwig Feuerbach: Sowie kritisch revidierte Edition "Zur Moralphilosophie" (1868) besorgt von W. Schuffenhauer. Braun, Hans-Jürg, 1927; Schuffenhauer, Werner; Feuerbach, Ludwig; Zur Moralphilosophie. Berlin: Akademie Verlag, 1994.

Wartofsky, Marx W. *Feuerbach*. Cambridge, U.K.; New York: Cambridge University Press, 1977.

Hayden V. White (1967)
Bibliography updated by Michael J. Farmer (2005)

FICHTE, JOHANN GOTTLIEB
(1762–1814)

Johann Gottlieb Fichte was a German philosopher. The most original and most influential thinker among the immediate successors of Immanuel Kant, Fichte was the first exponent of German idealism. He set the agenda for the philosophical work of the generation of Friedrich Wilhelm Joseph Schelling and Georg Wilhelm Friedrich Hegel and exerted tremendous influence on German cultural life in the final decade of the eighteenth century and the first decade of the nineteenth century. Fichte undertook pioneering philosophical work on a number of topics, including the primacy of the practical over the theoretical, the nature and development of self-consciousness, the status and function of one's own body, the original discovery of the other person, the integration of freedom and nature, and the separation of law and morality.

LIFE

Fichte was born on May 19, 1762, in the village of Rammenau in Saxony (in today's eastern Germany). Through the support of local benefactors, he received an education that would have been beyond the means of his family, who were ribbon weavers. He attended the Princely Latin School at Porta (Schulpforta) (1774–1780), studied theology and law at the universities of Jena, Wittenberg, and Leipzig under difficult financial circumstances and without taking a degree (1780–1784), and served as a private tutor in Leipzig, Eastern Prussia, and Zurich (1785–1793).

In 1790, upon studying Kant's *Critique of Pure Reason* (1781) and *Critique of Practical Reason* (1788), he became an enthusiastic adherent and supporter of Kant's critical philosophy. When Fichte's first publication, *Attempt at a Critique of All Revelation* (1792), appeared, in part, anonymously, it was widely assumed to be a work

by Kant, whose public clarification of the authorship launched Fichte's meteoric philosophical career. He was offered a professorship in philosophy at the University of Jena, where he began teaching in the summer semester of 1794. Fichte's widely attended lecture courses and the publications based on them turned German academic philosophy for a brief period into a world-historical movement on a par with the French Revolution and literary Romanticism.

In 1799 Fichte lost his professorship in Jena over charges of atheism, based on his published view that God was nothing but the moral order of the world. He spent most of the remaining years of his life in Berlin where he initially supported himself by giving private and public lecture courses and later received a professorship at the newly founded university (1810–1814), at which he also served as Dean (1810) and Rector (1811–1812). Between 1804 and 1808 Fichte gave several popular lecture series in Berlin, that were also published, in which he presented a scathing diagnosis of the cultural and moral ails of his time along with a fervent call for spiritual and political renewal. The most famous of these works, the *Addresses to the German Nation* (1807–1808, published in 1808), arose as an act of public resistance against the Napoleonic occupation of Prussia, Fichte's adopted homeland. The work's call for autochthonous culture and politics was repeatedly instrumentalized in the nineteenth and twentieth centuries for nationalist and socialist thought and politics. Fichte died on January 29, 1814, from hospital fever, which he had contracted from his wife of twenty years, who had been working as a nurse during the uprising against Napoleon.

"THE FIRST SYSTEM OF FREEDOM"

From his chance rise out of poverty and obscurity and his vehement early support of the French Revolution, which brought him a reputation as a Jacobin, through his daring breakaway from academic and religious traditions, to his eloquent agitation for liberation from Napoleonic rule, Fichte struggled all his life for freedom from tutelage of all kinds and for radical self-determination. The theoretical counterpart to this unrelenting project of self-liberation is what Fichte himself termed *the first system of freedom*—a comprehensive account of natural and cultural reality in which the concept of freedom serves to ground and integrate the key aspects of human existence (cognition and volition) as well as their corresponding worlds (the sensible or the natural and the supersensible or the spiritual). Unlike Kant, who had correlated and connected nature and freedom as different but complementary domains, each with its own principles, Fichte subordinates all of nature to freedom, turning the material world into nothing but the arena for the exercise of free self-determination under self-given laws of acting. With nature relegated to a merely instrumental status, the conditions and principles of social and cultural life receive primary consideration. Fichte's systematic treatment of law, morality, religion, history, and politics as the main spheres for the actualization of freedom is grounded in a detailed account of the deep structure of the human subject.

Throughout, Fichte follows Kant's transcendental or Copernican turn. But he deepens as well as widens his predecessor's dual focus on the conditions of the possibility of experience and the conditions of the possibility of morality into a highly integrated inquiry into the structural requirements of consciousness of all kinds and of all kinds of objects. In order to stress both the rigorous scientific character of his investigations and their merely preparatory status for everyone's own practice of freedom, Fichte abandons the traditional designation, *philosophy* or *love of wisdom*, replacing it with the coinage *Wissenschaftslehre*, or *Science of Knowledge*. The term is not a reference to epistemology in the modern sense but to the protoscience that is to achieve a metaknowledge of the conditions of the possibility of all object-knowledge and that then refers everyone to their own experience for the contingent content of such formally functioning consciousness. In a wider sense all parts of Fichte's projected and partially executed philosophical system are termed *Science of Knowledge*. But Fichte preferentially employs the term for his various presentations of the *first philosophy*, which contains only the basic principles of all knowledge and its objects.

Insisting on the freedom of genuine philosophical thought from any fixed letter and on the need for direct, oral philosophical communication, Fichte worked out some fifteen different presentations of his core philosophy over a period of twenty years, of which he himself published only the first one. As a result of this unique practice of continued production but discontinued publication of his main philosophical work, the full extent and content of Fichte's thinking after 1800 remained, for the most part, unknown to his contemporaries and was recognized and became influential only with the partial publication of Fichte's literary remains in the nineteenth century and their integral edition by the Bavarian Academy of Sciences over almost half a century starting in the early 1960s.

THE "I" AS THE PRINCIPLE OF PHILOSOPHY

In the early presentations of the Science of Knowledge, dating from 1794 through 1799, Fichte terms the unitary unconditional ground of theoretical and practical knowledge and of its object domains, *the I*. The nominalized first person pronoun serves to designate the principle for the derivation (*deduction*) of the basic features of the subject and its world or worlds. Fichte's strategy is to elucidate the necessary conditions under which the subject is able to achieve consciousness of itself, or self-consciousness. Among those conditions are the application of a set of categorial concepts (such as cause and effect) that assure the law-governed structure of the objects in space and time and the individuation of the subject as an intelligent being among other such beings. In particular, Fichte aims to show that the subject's practical relation to the world by way of volition and action is a necessary condition, even for its theoretical relation to the world through thinking and knowing. Fichte's defense of the systematic primacy of practice over theory is counterbalanced by the recognition that all practice in turn stands in need of some guidance through the cognition of the ends to be pursued.

In the original presentation of the Science of Knowledge from 1794–1795 (*Science of Knowledge with the First and Second Introductions*), the basic distinctions between the subject and the object and between the theoretical and the practical are generated by means of a transcendental dialectic involving the progressive but never completely achieved elimination of the contradictions to be found among the three chief capacities of the I as absolute I, theoretical I, and practical I.

As absolute I, the I is the unconditional ground of everything in the I and for the I, including everything that is not I (Not-I). Fichte employs the term *positing* for the generic, preconscious, and spontaneous activity of the I in bringing about the most basic structure of the subject as well as the object. He distinguishes the threefold absolute activity of the I's positing itself, positing its other (the Not-I), and positing the mutual determination of I and Not-I. As theoretical I, or as subject of cognition, the I posits itself as determined through the Not-I. The subject thereby conceives of itself as bound by the properties of the object to be cognized. The contradiction between the active nature of the absolutely positing I and the passive nature of the I of theoretical cognition is resolved through the I's third capacity as practical I, which consists in the I's striving to completely determine the Not-I and to have all determination of the I be the I's self-determination. To be sure, for Fichte, the striving of the practical I toward the status of the absolute I—to determine everything and to be determined only by itself—is an infinite process in which the absolute I serves as an unobtainable ideal (*idea*).

In Fichte's reconstruction of the principal constitutive features of consciousness, the key factors of Kant's transcendental philosophy (apperception, space, time, categories, imagination, ideas of reason) are gathered into a *history of consciousness* that stretches from minimal self-awareness in undifferentiated feeling through the workings of the imagination in theoretical understanding to the practical self-consciousness of striving reason. Fichte's completion of Kant's transcendental idealism does away with the existence of unknowable things in themselves and provides a maximally internalist account of the determination and self-determination of the I. The only remaining externalist concession is the appeal to the I's inexplicable experience of being held in *check* by what is subsequently objectified, according to the I's own laws, as a world of things seemingly existing independently of the theoretical I.

When his initial transcendental account of the I was widely mistaken for referring to an individual person rather than to the set of structural requirements for personhood, Fichte provided important methodological clarifications and doctrinal expansions in his *New Presentation of the Science of Knowledge* (*Foundations of Transcendental Philosophy [Wissenschaftslehre] nova methodo*; 1796–1799). In particular, he stressed the difference between the transcendental, supraindividual I of the Science of Knowledge and the empirical, individual I of ordinary cognition and life; he argued for the reconstructive, experimental, and hence artificial nature of the transcendental account of the I; and he maintained that the ultimate evidence for the transcendental–idealist reduction of everything to the I's clandestine absolute activity was the fundamental, extraphilosophical belief that absolute freedom from all foreign reality and complete self-determination were the essence and end of human existence.

Among the doctrinal additions of Fichte's alternative presentation of the Science of Knowledge are the systematically prominent position of the will and the foundational role accorded to interpersonal relations (intersubjectivity) in the constitution of the subject and its relation to the world. Fichte's transcendental philosophy of the I now presents itself as a theory of the principal forms and conditions of practical activity (willing and doing), into which the main features of cognitive activity

and the world of objects to be cognized are integrated. More specifically, Fichte argues that the mutual requirement of willing and knowing threatens to involve the I's theoretical–practical double nature (*duplicity*) in a vicious circle: Willing an end requires prior cognition of the object to be willed while knowing an object requires a prior engagement of the will in the course of which objects first come into view. Fichte resolves the circle by postulating a nonempirical, prepersonal, and hence predeliberative willing that comes with its own knowledge of what do—a type of willing modeled on Kant's notion of pure practical reason in which knowing the morally good and willing it are supposed to coincide. This move transposes the I from its embeddedness in the natural world into the moral realm of the pure will and entails its individuation among a community of finite rational agents.

The grounding of the I's theoretical as well as practical activities in original, self-determined volition points to the strictly moral core of human subjectivity in Fichte. What lends reality and objectivity to the I's pervasive activity of positing and determining is not some external physical or metaphysical entity but the I's own unconditional laws for the exercise of its spontaneity and freedom. In Fichte's ethical idealism the physical world has reality as the sphere for the exercise of our moral obligations.

In his most popular and accessible work, *The Vocation of Man* (1800), Fichte summarizes his philosophy of freedom in a dramatic portrayal of the course of human insight: from initial doubt about how to reconcile the competing claims of freedom and determination in human affairs through the intermediary stage of (merely theoretical) knowledge, for which everything and everyone is but a product of the I, to the concluding stage of practical knowledge and the faith associated with it, which reconciles freedom and determination by reconceiving the latter as moral self-determination.

THE I AND THE ABSOLUTE

Fichte's subsequent popular lectures and publications in the philosophy of history, culture, and religion (1804–1808) continued to stress the practical and specifically the moral dimension of human existence. In his continuing work on the Science of Knowledge (1801–1814) Fichte explored in ever-new attempts and with changing terminological and conceptual means the possibilities as well as the limitations of human knowledge and human freedom. In critical distance to the contemporary turn toward an affirmative philosophy of the absolute in philosophers such as Friedrich Heinrich Jacobi, Schelling, and Hegel, Fichte stressed the epistemological strictures of any ascent from the transcendental to the metaphysical. While de-emphasizing the self-sufficiency of the I and abandoning much of his earlier terminology of the I, he nevertheless insisted on the close linkage—and the ultimate identity—of the absolute and the absolute I and on the I's function as the basic mode (*I form*) of theoretical and practical subjectivity.

For the later Fichte, the absolute is not some higher being apart from our self-determined existence as knowers and doers but that which sustains and animates our theoretical and practical activities as the unfathomable ground of their dynamics and laws. In order to avoid any objectivist misunderstanding of the subject's origination in the absolute, Fichte replaces the latter's appellation as *being* with that of *life*, understood as sheer activity, without a distinct bearer and a resultant product. For the later Fichte, human existence—more specifically, its normative accomplishment of knowledge of what there is and ought to be—is the one and only manifestation (*appearance* or *image*) of the absolute while everything else has being only secondarily, as possible object of cognition and volition. Moreover, the authentic manifestation of the absolute is the absolute's self-manifestation as such. The ultimate knowledge to be achieved is the philosophical or metaknowledge that knowledge is but the appearance of the absolute and that the absolute appears only as knowledge.

For Fichte this ultimate insight, which completes the Science of Knowledge, involves at once the self-limitation of knowledge over and against the absolute, of which knowledge is but an image, and the self-affirmation of knowledge as being the absolute itself in the latter's external mode (*existence*). Accordingly, the insight achieved by the Science of Knowledge is not some abstract, rare cognition but results from the lived identification of the subject with its absolute ground and results in a manner of thinking and acting animated by the inner presence of the absolute. Moreover, on Fichte's account, the thinking and acting in light of the absolute does not occur automatically but depends on the subject's free decision and sustained effort to radical reflection and its decision and effort to engage in conduct corresponding to the insight achieved. Thus, the speculative efforts of the Science of Knowledge aim beyond science and knowledge at practical wisdom and at the moral activity resulting from it—an ultimate confirmation of the intellectual and moral freedom of the subject.

Despite some appearances to the contrary, which are due to occasional metaphysically charged terminology

(*God*, *being*), the late presentations of the Science of Knowledge, when considered in their entirety, show Fichte arguing for the essentially practical nature of the absolute as absolute will and as the animating principle of the moral order. Thus, the later Fichte exhibits a striking continuity with the ethical orientation of his earlier speculative philosophy and, beyond that, with the moral agenda of Kant's critical philosophy.

PHILOSOPHY OF LAW AND ETHICS

Given its unique combination of systematic rigor, argumentative concentration, and freely varied presentation, Fichte's foundational work on the Science of Knowledge initially met with incomprehension; soon became marginalized by the work of his followers, Schelling and Hegel; and even in the early twenty-first century, in the context of detailed historical scholarship and extensive textual analysis, defies summary assessment and doctrinal reconstruction. By contrast, Fichte's work on the applied part of the Science of Knowledge, which consists of the philosophy of law and ethics, has always been more widely appreciated and quite influential.

Fichte's *Foundations of Natural Law* (1796–1797) integrates the theory of right and political authority into a systematic account of the I's individuation and socialization. Fichte argues that a subject can only possess self-consciousness if a number of conditions are met that take the form of the subject's implicit self-ascription (*positing*) of increasingly specific nonrelational and relational properties. To begin with, the subject has to ascribe to itself the faculty of free efficacy along with a sphere of objects, the world of sense, in which the efficacy can be exercised by bringing about change in the objects. Moreover, the subject's practical activity in the world of sense requires its self-ascription of a material object (*body*), by means of which it can act upon the material word.

In a crucial and highly original next step, Fichte argues that a further requirement for the subject's self-conscious, practical activity in the empirical world is its initiation into the rational standards of knowing and doing, which in turn leads to the presupposition of another, already fully functioning, subject and specifically to the latter's influencing the first subject to discover and engage its potential for theoretical–practical rationality. Moreover, the required influence has to be such that the constitutive freedom of the subject to be influenced is not infringed upon but rather called upon and made to emerge. The required influence is not physical but a *determination to self-determination* or the encouraging appeal (*summons* or *solicitation*) to act freely and rationally.

Fichte terms the soliciting subject's attitude of acknowledgment and respect toward the solicited subject's full human potential an act of *recognition* and moves on to inquire into the necessary condition for the possibility of continued mutual recognition between individual subjects. This condition is the *relationship of law*, in which each subject freely limits the exercise of its free efficacy in the world of sense through the concept of the possible freedom of the other individual subject—under the condition that the latter does the same.

Unlike Kant, Fichte does not subordinate the sphere of law under that of morals but defends a strict separation of law and morality. To be sure, for Fichte, the concept of law—the concept of the mutual recognition of free agency—represents a necessary condition of self-consciousness. But becoming part of a political state and following its laws is not an unconditional command of practical reason, as in the "Metaphysical First Principles of the Doctrine of Right" in Kant's *Metaphysics of Morals* (Immanuel Kant, *Practical Philosophy*. Translated and edited by Mary J. Gregor. General introduction by Allen Wood. Cambridge, U.K.: Cambridge University Press, 1996: 455–456.), which was published only after Fichte's work (1797). In Fichte, the validity of the law and that of its subsequent specifications as state law, family law, and cosmopolitan law is contingent upon the agreed-upon and continued practice of recognitional conduct on the part of all individual subjects involved. Accordingly, Fichte's account of the powers of the state is designed to assure the continued observance of mutual recognitional conduct.

With the philosophy of law and its postulation of the transcendental conditions of sociality relegated to an extension of theoretical philosophy, practical philosophy in Fichte completely coincides with ethics or the doctrine of our unconditional moral duties as opposed to our contingent legal obligations. Moreover, Fichte's ethics, published as *The System of Ethics* in 1798, differs widely in scope and structure from the "Metaphysical First Principles of the Doctrine of Virtue" of Kant's *Metaphysics of Morals* published in the preceding year. While Kant had focused on the systematic presentation of particular duties and had limited more general considerations to comparatively brief introductory sections, Fichte provides a detailed derivation of the principle of morality along with the basic conditions of its application. The treatment of ethics in the narrow sense, or the presentation of particular duties, is limited to the work's brief concluding section.

Fichte's chief ambition in practical philosophy is to overcome what he perceives to be the emptiness and formalism inherent in a Kantian ethics focusing on the moral criterion (*categorical imperative*) of the possible universality of subjective principles of action (*maxims*). By contrast, Fichte integrates the formation and execution of moral willing into the overall structure of practical subjectivity. The factual starting point of Fichte's real or material ethics is the subject's original self-experience as willing or as engaged in conceptually mediated efficacy in the world of sense. Its normative end point is the absolute freedom of the subject or radical self-determination. Under conditions of human finitude, this goal can only be approximated. Morality provides the direction and the motivation of the finite subject toward its infinite destination.

In his effort to lend content and specificity to moral obligation, Fichte positions the free will of the practical subject under the influence of a unitary but twofold *drive*: the *pure drive* that represents the claims of pure practical reason to radical self-determination and the *natural drive* that represents the demands of our nonrational nature. Fichte considers the *moral drive* to be a mixed drive in which the natural drive provides the content and the pure drive contributes the impetus for acting or the motivation. Fichte further argues for a pre-established harmony of sorts between the natural drive and the pure drive such that in each and every situation there is one and only one action that is both proposed by the natural drive and sanctioned by the pure drive. According to Fichte, the specific duties are detected by a non-sensory feeling of immediate practical certainty (*conscience*). The principle of morality can therefore be put into the following formula, which is empty by itself and refers everyone to their own conscience for its completion: Do in each case your duty and do it for duty's sake.

PHILOSOPHY OF HISTORY, EDUCATION, AND RELIGION

Compared with the unprecedented systematic rigor and highly abstract reasoning pervading the presentations of the Science of Knowledge in its foundational as well as applied parts, Fichte's historically influential contributions to the philosophy of history, education, and religion are popular works conceived and executed with the explicit intent of exercising moral and political influence on listeners and readers—whose abilities, preconceptions, and contemporary experiences have therefore entered into the design of these works. Accordingly, Fichte's popular works call not only for philosophical analysis but also for historical knowledge and exegetical skill in assessing the complex relation between their claims and their contexts.

Fichte's philosophy of history, presented in the *Characteristics of the Present Age* (1804–1805, published in 1806) and supplemented by the *Addresses to the German Nation* (1807–1808), constructs the ideal course of history as a linear progress in the governance of humankind in five stages: from blind but clandestinely rational instinct through irrational authority to anarchical intellectual, moral and political freedom—the *present age*, according to Fichte—on to incipient, freely exercised rationality, and finally to the complete reign of rational freedom. The transition from the present age of *complete sinfulness* to genuine freedom and true enlightenment is to be brought about by education and specifically by educational reform at all levels—from instituting compulsory public primary schools to a structural and curricular reform of the university, of which Fichte was a major theoretician and practitioner. Fichte's high regard for public education is also reflected in the three lecture courses that he gave on the moral and political role of the public intellectual (*vocation of the scholar*) at the beginning, toward the middle, and toward the end of his academic career (1794, 1805, 1811).

Fichte's philosophy of religion, presented as *The Way Towards the Blessed Life* (1806), recasts the speculative core of the Science of Knowledge in the form of a popular ontology identifying life with love and bliss. Fichte distinguishes five world views, each correlated to a specific standpoint and associated with a specific affect: the standpoint of sensibility and its enjoyments: that of objective legality (merely formal morality) and its love of formal freedom; that of higher morality and the self-satisfaction it affords; that of religion and the blessed life it entails: and that of science (viz., the Science of Knowledge), which adds no viewpoint of its own but unites the preceding ones by lending them clarity and by transforming the mere faith in the absolute into envisioning it through the self-immersion of reflection into the absolute.

ASSESSMENT

The immediate, immense, but short-lived influence that Fichte had on the course of German culture and philosophy around 1800 is augmented by the long-term and more clandestine effects that his original thinking on the nature of subjectivity and the relation between theory and practice exercised on such diverse philosophers as Arthur Schopenhauer, Karl Marx, Martin Heidegger, and

Jürgen Habermas. With several of his later works only now available for the first time, Fichte is very much a philosopher still to be discovered. His early work on the system of freedom is a tour de force in radicalized Kantianism while his later work on the absolute and its appearance as knowledge and will is a serious competitor to Schelling's and Hegel's claims of having brought to completion German idealist philosophy.

See also Copernicus, Nicolas; Epistemology; Hegel, Georg Wilhelm Friedrich; Habermas, Jürgen; Heidegger, Martin; Jacobi, Friedrich Heinrich; Kant, Immanuel; Marx, Karl; Schelling, Friedrich Wilhelm Joseph; Schopenhauer, Arthur.

Bibliography

WORKS BY FICHTE: GERMAN EDITIONS

Johann Gottlieb Fichte's nachgelassene Schriften. 3 vols. Edited by Immanuel Hermann Fichte. Bonn: Adolph-Marcus 1834–1835. Reprinted as *Fichtes Werke*. 11 vols. (vols. 9–11). Berlin: de Gruyter, 1971.

Johann Gottlieb Fichte's sämmtliche Werke. 8 vols. Edited by Immanuel Hermann Fichte. Berlin: Veit, 1845–1846. Reprinted as *Fichtes Werke*, 11 vols. (vols. 1–8). Berlin: de Gruyter, 1971.

J. G. Fichte-Gesamtausgabe der Bayerischen Akademie der Wissenschaften. Edited by Reinhard Lauth and Hans Gliwitzky. Stuttgart-Bad Cannstatt: Frommann-Holzboog: 1962– (40 vols. planned; to be completed by 2010). Series I: *Published Works*; series II: *Unpublished Works*; series III: *Correspondence*; series IV: *Lecture Transcripts*.

Fichte im Gespräch: Berichte der Zeitgenossen. 6 vols. Edited by Erich Fuchs. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1978–1992.

Fichte in zeitgenössischen Rezensionen. 4 vols. Edited by Erich Fuchs, Wilhelm G. Jacobs, and Walter Schieche. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1995.

Ultima Inquirenda. J. G. Fichtes letzte Bearbeitungen der Wissenschaftslehre Ende 1813/Angang 1814. Edited by Reinhard Lauth. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 2001.

WORKS BY FICHTE: ENGLISH TRANSLATIONS

The Popular Works of Johann Gottlieb Fichte. 2 vols. Translated by William Smith. London: Trübner, 1889; 4th ed, introduction by Daniel Breazeale. Sterling, VA: Thoemmes, 1999. (Contains *The Vocation of the Scholar*, *The Nature of the Scholar*, *The Vocation of Man*, *The Characteristics of the Present Age*, *The Way towards the Blessed Life or the Doctrine of Religion*, and *Outlines of the Doctrine of Knowledge*.)

Addresses to the German Nation. Translated by R. F. Jones and G. H. Turnbull. Chicago: University of Chicago Press, 1922. Reprinted, edited by George Armstrong Kelly. New York: Harper, 1968.

"The Science of Knowledge in Its General Outline." Translated by Walter E. Wright. *Idealistic Studies* 6 (1976): 106–117.

Attempt at a Critique of All Revelation. Translated by Garrett Green. Cambridge, U.K.: Cambridge University Press, 1978.

Science of Knowledge with the First and Second Introductions. Translated by Peter Heath and John Lachs. Cambridge, U.K.: Cambridge University Press, 1982.

"Some Lectures Concerning the Scholar." Translated by Daniel Breazeale. In *Philosophy of German Idealism*. Edited by Ernst Behler, 4–38. New York. Continuum 1987.

The Vocation of Man. Translated by Peter Preuss. Indianapolis, IN: Hackett, 1987.

"A Crystal Clear Report to the General Public Concerning the Actual Essence of the Newest Philosophy. An Attempt to Force the Reader to Understand." Translated by John Botterman and William Rash. In *Philosophy of German Idealism*. Edited by Ernst Behler, 39–115. New York: Continuum, 1987.

Early Philosophical Writings. Edited and translated by Daniel Breazeale. Ithaca, NY: Cornell University Press, 1988.

Foundations of Transcendental Philosophy (Wissenschaftslehre) nova methodo 1796–1799. Edited and translated by Daniel Breazeale. Ithaca, NY: Cornell University Press, 1992.

"Reclamation of the Freedom of Thought from the Princes of Europe, Who Have Oppressed It Until Now." Translated by Thomas E. Wartenberg. In *What Is Enlightenment? Eighteenth-Century Answers and Twentieth-Century Questions*. Edited by James Schmidt, 225–232. Berkeley: University of California Press, 1996.

Introductions to the Wissenschaftslehre and Other Writings. Edited and translated by Daniel Breazeale. Indianapolis, IN: Hackett, 1994.

Foundations of Natural Right. Edited by Frederick Neuhouser and translated by Michael Baur. Cambridge, U.K.: Cambridge University Press, 2000.

"Annotated Translation J. G. Fichte's Review of L. Creuzer's *Skeptical Reflections on the Freedom of the Will* (1793)." Translated and edited by Daniel Breazeale. *Philosophical Forum* 32 (2001): 289–296.

"Annotated Translation of J. G. Fichte's Review of F. H. Gebhard's *On Ethical Goodness as Disinterested Benevolence* (1793)." Translated and edited by Daniel Breazeale. *Philosophical Forum* 32 (2001): 297–310.

"Annotated Translation of J. G. Fichte Review of I. Kant's *Perpetual Peace* (1796)." Translated and edited by Daniel Breazeale. *Philosophical Forum* 32 (2001): 311–321.

The Science of Knowing. J. G. Fichte's 1804 Lectures on the Wissenschaftslehre. Edited and translated by Walter E. Wright. Albany: State University of New York Press, 2005.

The System of Ethics. Edited and translated by Daniel Breazeale and Günter Zöller. Cambridge, U.K.: Cambridge University Press, 2005.

WORKS ON FICHTE

Breazeale, Daniel, and Tom Rockmore, eds. *Fichte. Historical Contexts/Contemporary Controversies*. Atlantic Heights, NJ: Humanities Press, 1994.

Breazeale, Daniel. "Check or Checkmate? On the Finitude of the Fichtean Self." In *The Modern Subject. Conceptions of the Self in Classical German Philosophy*. Edited by Karl Ameriks and Dieter Sturm, 87–114. Albany: State University of New York Press, 1995.

Breazeale, Daniel, and Tom Rockmore, eds. *New Essays on Fichte*. Atlantic Heights, NJ: Humanities Press, 1996.

Breazeale, Daniel, and Tom Rockmore, eds. *New Essays in Fichte's "Foundation of the Entire Doctrine of Scientific Knowledge."* New York: Humanity Books, 2001.

Breazeale, Daniel, and Tom Rockmore, eds. *New Essays on Fichte's Later Jena "Wissenschaftslehre."* Evanston, Ill.: Northwestern University Press, 2002.

Breazeale, Daniel, and Tom Rockmore, eds. *Rights, Bodies, and Recognition. New Essays on Fichte's Foundations of Natural Right*. Aldershot, U.K.: Ashgate, 2005.

Henrich, Dieter. "Fichte's Original Insight." *Contemporary German Philosophy* 1 (1982): 15–53.

Horstmann, Rolf-Peter. "The Early Philosophy of Fichte and Schelling." In *The Cambridge Companion to German Idealism*. Edited by Karl Ameriks, 117–140. Cambridge, U.K.: Cambridge University Press, 2000.

La Vopa, Anthony J. Fichte. *The Self and the Calling of Philosophy, 1762–1799*. Cambridge, U.K.: Cambridge University Press, 2001.

Martin, Wayne. *Idealism and Objectivity. Understanding Fichte's Jena Project*. Stanford, CT: Stanford University Press, 1997.

Neuhouser, Frederick. *Fichte's Theory of Subjectivity*. Cambridge, U.K.: Cambridge University Press, 1990.

Pippin, Robert. "Fichte's Contribution." *Philosophical Forum* 19 (1988): 74–96.

Williams, Robert. *Recognition. Fichte and Hegel on the Other*. Albany: State University of New York Press, 1992.

Wood, Allen W. "Fichte's Philosophical Revolution." *Philosophical Topics* 19 (1992): 1–28.

Zöller, Günter, ed. *The Cambridge Companion to Fichte*. Cambridge, U.K.: Cambridge University Press (in preparation).

Zöller, Günter. *Fichte's Transcendental Philosophy. The Original Duplicity of Intelligence and Will*. Cambridge, U.K.: Cambridge University Press, 1998.

Zöller, Günter. "German Realism: The Self-Limitation of Idealist Thinking in Fichte, Schelling and Schopenhauer." In *The Cambridge Companion to German Idealism*. Edited by Karl Ameriks, 200–218. Cambridge, U.K.: Cambridge University Press, 2000.

Zöller, Günter, "The Unpopularity of Transcendental Philosophy: Fichte's Controversy with Reinhold (1799–1801)." *PLI: The Warwick Journal of Philosophy* 10 (2000): 50–76.

Günter Zöller (2005)

FICINO, MARSILIO

(1433–1499)

Marsilio Ficino, the founder of the Florentine Academy, was born the eldest son of a physician in Figline, near Florence. He studied the humanities, philosophy, and medicine in Florence but apparently did not obtain an academic degree. About 1456 he began to study Greek. In 1462 he received from Cosimo de' Medici a house in Careggi, near Florence, and several Greek manuscripts; this is regarded as the date the Platonic Academy of Florence was founded. Having earlier taken minor orders, Ficino was ordained in 1473; he held several ecclesiastic benefices and became a canon of Florence Cathedral in 1487. After the expulsion of the Medicis from Florence in 1494, Ficino, who had been closely associated with several generations of the family, apparently retired to the country. He was honored after his death in a funeral oration delivered by a chancellor of the republic of Florence.

Ficino became interested in Platonist philosophy at an early age, presumably through studying Augustine. His earliest extant writings also show familiarity with Aristotle and his commentators and with Lucretius. Among Ficino's Latin translations from the Greek, the first that attained a wide circulation was his version (1463) of the works attributed to Hermes Trismegistus. Ficino's translation of Plato, the first complete rendering of all his dialogues in any Western language, was begun in 1463, probably completed in 1469, subsequently revised, and first printed in 1484. His influential commentary on Plato's *Symposium* was written in 1469; the other Platonic commentaries, some of them extensive, belong to different periods of Ficino's life. The translation of and commentary on Plotinus was begun in 1484 and printed in 1492. Translations of Porphyry, Iamblichus, Proclus, and other philosophers appeared in 1497. Ficino's chief philosophical work, *Theologica Platonica de Immortalitate Animarum* (Platonic theology—on the immortality of the souls) was written between 1469 and 1474 and was printed in 1482. Aside from this work and his commentaries, the most important source for Ficino's philosophy is his letters, which he began to collect around 1473 and finally published in 1495. Important also are his apologetic treatise *De Christiana Religione* (1474) and his work on medicine and astrology, *De Vita Libri Tres* (1489), which is often wrongly referred to as *De Vita Triplici.*

Ficino's work as a translator of and commentator on Plato and the Neoplatonists, and his avowed intention of reviving Platonism, led many older historians to treat his doctrine merely as a repetition of ancient Neoplatonism. More recently, however, a closer study of his known and unpublished works has shown that in restating the doctrines of Plato and his ancient followers, Ficino showed a good deal of originality. In addition, his writings show the influence of medieval and Byzantine Platonism, early Italian humanism, and also the tradition of scholastic Aristotelianism, which had a strong impact upon his terminology and method. He was familiar with Dante

Alighieri and other Italian poets and wrote or rewrote several of his own works in the Tuscan vernacular.

Ficino was the founder and for many years the head and guiding spirit of the Platonic Academy of Florence, which has remained famous as a symbol and institutional center of Renaissance Platonism. The academy was not a firmly established institution in the manner of later academies but a rather loosely organized spiritual community of friends. We hear of informal discussions between the older members of the circle and of philosophical banquets celebrated on Plato's birthday. There were recitals of edifying orations before a small audience, private readings of Plato and other texts given by one or a few younger disciples, and public lectures on Plato and Plotinus delivered in a church or auditorium. Distinguished visitors from other Italian cities and abroad called upon Ficino or participated in the meetings, and Ficino's correspondence served as a vehicle both for maintaining contact with the members of the academy and for arousing the interest of strangers in the academy's activities. The catalog of his pupils, which he gives in one of his letters, and the list of the persons with whom he was in correspondence, whom he mentions, or who owned the manuscripts and printed editions of his writings are ample evidence of the wide influence he exerted during his lifetime.

Ficino's writings present a highly complex system of ideas, embroidered with similes, allegories, and lengthy quotations from his favorite authors. We can mention but a few of his more important and influential doctrines.

HIERARCHY

In his description of the universe, Ficino took from Neoplatonic and medieval sources the conception of a great hierarchy in which each being occupies its place and has its degree of perfection, beginning with God at the top and descending through the orders of the angels and souls, the celestial and elementary spheres, the various species of animals, plants, and minerals, down to qualityless prime matter. In spite of Ficino's indebtedness to earlier schemes, it appears on closer examination that his hierarchy differs in significant details from those of his predecessors. It is arranged in a final scheme of five basic substances: God, the angelic mind, the rational soul, quality, and body. This scheme comes fairly close to that of Plotinus but differs from it in various ways. Above all, quality did not constitute a separate level of being for Plotinus, who instead assigned separate places to the sensitive and vegetative faculties of the soul. It can be shown that Ficino intentionally revised the Plotinian scheme, partly to make it more symmetrical and partly to assign the privileged place in its center to the human soul, thus giving a kind of metaphysical setting and sanction to the doctrine of the dignity of man, which he had inherited from his humanist predecessors. The soul is truly the mean of all things created by God, Ficino tells us. It is in the middle between higher and lower beings, sharing some of its attributes with the former and some with the latter.

Ficino was not satisfied with a static hierarchy in which each degree merely stands beside the others and in which the relationship of degrees consists only in a continuous gradation of attributes. He was also convinced that the universe must have a dynamic unity and that its various parts and degrees are held together by active forces and affinities. For this reason, he revived the Neoplatonic doctrine of the world soul and made astrology part of a natural system of mutual influences. Since thought for Ficino has an active influence upon its objects, since love is an active force that binds all things together (as in Plato's *Symposium*), and since the human soul extends its thought and love to all things, from the highest to the lowest, in Ficino's writing the soul becomes once more and in a new sense the center of the universe. The soul is the greatest of all miracles in nature because it combines all things, it is the center of all things, and possesses the forces of all things. Therefore the soul may rightly be called the center of nature, the middle term of all things, the bond and juncture of the universe.

CONTEMPLATION

Ficino's cosmology, which was very influential during the sixteenth century, offers some points of intrinsic interest; however, it constitutes only one side of his thought. The other and even more profound component is his analysis, based on direct inner experience, of the spiritual or contemplative life, and analysis that links him with some of the medieval mystics and, again, with Neoplatonism. In the face of ordinary daily experiences, the mind finds itself in a state of continuous unrest and dissatisfaction, but it is capable of turning away from the body and the external world and of concentrating upon its own inner substance. Thus purifying itself of things external, the soul enters the contemplative life and attains a higher knowledge, discovering the incorporeal world that is closed to it while it is engaged in ordinary experience and in the troubles of the external life. Ficino interpreted this contemplative life as a gradual ascent of the soul toward always higher degrees of truth and being, an ascent that finally culminates in the immediate knowledge and

vision of God. This knowledge of God represents the ultimate goal of human life and existence—in it alone the unrest of our mind is satisfied—and all other modes and degrees of human life and knowledge must be understood as more or less direct and conscious preparations for this end. In accordance with Plotinus, Ficino was convinced that this highest experience could be attained during the present life, at least by a few privileged persons and for a short while, although he never explicitly claimed to have attained this state himself.

In describing the various states and the ultimate goal of inner experience, Ficino used a twofold terminology, and in this he was influenced by St. Augustine and by the medieval philosophers. The ascent of the soul toward God is accomplished with the help of two wings, the intellect and the will; accordingly, the knowledge of God is accompanied and paralleled on each level by the love of God; and the ultimate vision, by an act of enjoyment. Ficino also considered the question of whether intellect and knowledge or will and love are more important in this process, and although he seemed to come to different conclusions in different parts of his writings, in general he leaned toward the superiority of will and love over intellect and knowledge. Yet the question was not so important for him as might be expected, since he regarded the knowledge of God and the love of God as merely two different aspects or interpretations of the same basic experience—namely, the contemplative ascent of the soul toward its ultimate goal.

This experience and the manner in which it is interpreted hold the key to both Ficino's metaphysics and his ethics. It is the inner ascent of contemplation, through which the reality of incorporeal things—of the ideas and of God himself—is discovered and verified. Since this inner ascent constitutes the basic task of human existence, Ficino was not interested in specific moral precepts or in casuistry, but only in the general identification of the human good and man's moral excellence with the inner life. His whole moral doctrine, as expressed in his letters, may be said to be a reduction of all specific moral rules to a praise of the contemplative life. He who has attained this life is exempt from the blows of fortune; and, animated by his inner certainty and insight, he will know and do the right thing under any given circumstance.

Intimately related to the doctrine of the contemplative life are two other theories of Ficino's, both of great historical importance: his theory of the immortality of the soul and his theory of Platonic love.

IMMORTALITY

Ficino's main work, *Theologia Platonica de Immortalitate Animarum,* consists for the most part of a series of arguments in support of the soul's immortality. It appears from a famous passage twice repeated in Ficino's writings that, in direct contrast with the teachings of the Aristotelian philosophers of his time, he considered this doctrine the central tenet of his Platonism. It is true that the immortality of the soul had been defended by Plato and Plotinus, by Augustine and many other Christian writers, and that Ficino borrowed many specific arguments from them. It may also be granted that Averroes's doctrine of the unity of the intellect in all people, which had been widely discussed and often accepted by Aristotelian philosophers from the thirteenth to the fifteenth century, made a defense of individual immortality imperative. In addition, the humanists had attached great importance to the individual human being, his experiences, and his opinions; and the belief in personal immortality was, as it were, a metaphysical counterpart of this individualism and an extension of it into another dimension.

Yet it seems evident that for Ficino the doctrine of immortality was a necessary complement and consequence of his interpretation of human existence and of the goal of human life. If it is our basic task to ascend, through a series of degrees, to the immediate vision and enjoyment of God, we must postulate that this ultimate goal will be attained, not merely by a few persons and for a short while but by a great number of human beings and forever. Otherwise, man's effort to attain this ultimate end would be in vain, and the very end for which he had been destined would remain without fulfillment. Thus, man would be unhappier than the animals, which do attain their natural ends, and this would be inconsistent with the dignity of the place man occupies in the universe. Moreover, if a natural end corresponding to a natural desire implanted in all men could not be attained, this would contradict the perfection of the order of nature and the wisdom of God, who created that order. In his "Platonic Theology," and in other parts of his writings, Ficino never tired of repeating these and similar arguments. It seems obvious that they reflect the real intent and motivation of his thought, for his whole interpretation of human life as a contemplative ascent toward God would lose its meaning unless this ascent were to find its permanent fulfillment in the eternal afterlife of the immortal soul. This alone would explain why the doctrine of immortality assumed such a central place for him. All other arguments are merely auxiliary to this central one.

Ficino's doctrine of immortality, and his arguments for it, made a profound impression on many thinkers of the sixteenth century; and it may very well be due to his indirect influence that the immortality of the soul was formally pronounced a dogma of the Catholic Church at the Lateran Council of 1512.

THEORY OF LOVE

Of equal historical importance, although different in character, is Ficino's doctrine of human love. In this doctrine, as in many of his others, Ficino combined elements from several different sources and traditions. He took over and reinterpreted Plato's theory of love as expressed in the *Symposium* and *Phaedrus,* and combined it with other ancient theories of friendship that were known to him primarily through Aristotle and Cicero; he also tried to identify it with the Christian love (*caritas*) praised by St. Paul. He even added some touches from the tradition of medieval courtly love as it was known to him through Guido Cavalcanti, Dante, and other early Tuscan poets. This doctrine of love, which exercised a tremendous influence during the sixteenth century, and for which Ficino himself coined the terms *Platonic love* and *Socratic love,* was first expressed by him in his commentary on Plato's *Symposium* and further developed in many of his letters and other writings. The term *Platonic love* means love as described by Plato, according to Ficino's interpretation; more frequently, he spoke of it as divine love. The basic point is that he regarded love for another human being as merely a preparation, more or less conscious, for the love of God, which constitutes the real goal and true content of human desire and which is turned toward persons and things by virtue of the reflected splendor of divine goodness and beauty that may appear in them. Ficino insisted that true love or friendship is always mutual. A genuine relationship between two people is a communion founded on what is essential in man, that is, it is based in each of them on his original love for God. There can never be only two friends; there must always be three—two human beings and one God. God alone is the indissoluble bond and perpetual guardian of any true friendship for a true lover loves the other person solely for the sake of God. True love and friendship between several persons is derived from the love of the individual for God; it is thus reduced to the basic phenomenon of the inner ascent, which constitutes the core of Ficino's philosophy.

It appears from Ficino's letters that he considered true friendship in this sense to be the bond that united the members of his academy with each other and with himself, their common master, and that he liked to think of the academy not merely as a school but as a community of friends. This conception of Platonic love was to exercise a strong influence on Italian and European literature throughout the sixteenth century. Many lyric poets spoke of their love in terms that reflected the influence of Ficino, as well as that of the old Tuscan poets and Petrarch; and there was a large body of treatises and lectures on love that derived much of their inspiration, directly or indirectly, from Ficino's commentary on the *Symposium.* In this literature the concept of Platonic love was separated from the philosophical context in which it had originated with Ficino, and so it became more and more diluted and trivial. For this reason, the notion of Platonic love has acquired a slightly ridiculous connotation for the modern reader. Yet we should try to recapture its original meaning, remembering that the true meaning of an idea is best understood in the context of the thought in which it originated and which, in a sense, made its formulation necessary. If we trace Platonic love back to its origin in Ficino—back to the context of an individual's love of God—it may still seem a strange and remote concept, but we shall at least understand that it had a serious content and that it was related to the central ideas of his philosophy.

A further aspect of Ficino's thought that requires mention is his conception of religion and of its relationship to philosophy. Ficino was a priest and a canon of Florence Cathedral; he had an adequate knowledge of Christian theology; and he even wrote an apologetic treatise on the Christian religion as well as several other theological works. There is not the slightest doubt that he intended to be orthodox, although some of his doctrines may seem to have dubious implications and although he was in danger of an ecclesiastical condemnation for the views on astrology and magic expressed in his work *De Vita* (1489). He insisted on his Christian faith and submitted to the judgment of the church. He was even willing to abandon the opinions of his favorite Platonist philosophers when they seemed to contradict Christian doctrine. Thus, we are not surprised to find that he regarded Christianity as the most perfect of all religions.

At the same time, he saw some merit in the variety of religions and insisted that any religion, however primitive, is related indirectly to the one true God. In his implicit tolerance toward other religions, Ficino came very close to a concept of natural religion, a position that made him a forerunner of Herbert of Cherbury, the deists, and other advocates of a universal religion. Divine worship, he said, is almost as natural for men as neighing

is for horses or barking is for dogs. A common religion of all nations, having one God for its object, is natural to the human species. This religion, which is again based on man's primary knowledge and love of God, is not shared by the animals but is peculiar to man, a part of his dignity and excellence and a compensation for the many defects and weaknesses of his nature.

As to the relation between religion and philosophy, Ficino was convinced that true religion (that is, Christianity) and true philosophy (that is, Platonism) are in basic harmony with each other; and he was inclined to treat them as sisters instead of trying to make one subservient to the other. He believed that it is the task of Platonic reason to confirm and support Christian faith and authority, and he even considered it his own mission, assigned to him by divine providence, to revive true philosophy for the benefit of true religion. He believed that those who will not be guided by faith alone can be guided toward truth only through reason and the most perfect philosophy.

In the light of this relationship, the continuity of the Platonic tradition assumed a new significance for Ficino. Since this tradition was thought by him to go back to Hermes and Zoroaster, whose apocryphal writings Ficino treated as venerable witnesses of early pagan theology and philosophy, he considered the tradition to be as old as the religious tradition of the Hebrews. Thus, the religious tradition of the Hebrews and Christians, and the philosophical tradition of the Hermetics and Platonists, seemed to run a parallel course in human history from the early beginnings through antiquity and the Middle Ages down to the modern period. It is in accordance with this view of Ficino's that Augustinus Steuchus, a Catholic theologian of the sixteenth century, wrote his *De Philosophia Perenni* (On the perennial philosophy; 1542).

INFLUENCE

Ficino's influence was considerable, both during his lifetime and for a long time afterward. As a metaphysician in the proper sense of the word, Ficino added an element to Florentine culture that had been largely absent from it before and left a new imprint on that culture that was to last for several generations. Among his associates and pupils we find Cristoforo Landino, author of the *Camaldulensian Disputations* and of an influential commentary on Dante's *Commedia,* and Lorenzo de' Medici, famous not only as a statesman but also as one of the best Italian poets of his century. Whereas Giovanni Pico della Mirandola developed an independent position, another pupil, Francesco da Diacceto, carried the Platonic tradition of Ficino into the first decades of the sixteenth century; and later in that century, Platonic philosophy was cultivated both at the new Florentine Academy of 1540 and at the University of Pisa. This Platonist climate of opinion in Florence and Pisa accounts for some of the opinions and preconceptions of Galileo Galilei. In the rest of Italy, poets and prose writers drew on Ficino's theory of love, and theologians and philosophers upon his doctrine of immortality as well as some of his other ideas. His influence appears in the works of such leading philosophers as Francesco Patrizi and Giordano Bruno: Even thinkers who opposed his views, such as Pietro Pomponazzi, were impressed with his learning and acumen.

During his lifetime, Ficino's influence was already growing, through his correspondence and through the circulation of his writings, in most European countries. His admirers included Johannes Reuchlin and John Colet, Gaguin and Jacques Lefèvre d'Étaples. During the sixteenth century his writings were reprinted, collected, read, and quoted all over Europe. His medical and astrological treatises were especially popular in Germany. In France, he was repeatedly quoted and plagiarized by Symphorien Champier, and admired in the circles of Queen Marguerite of Navarre and of the Pléïade. There, some of his writings and his Latin translations of Plato were translated into French. Elements of his Platonism appear in Carolus Bovillus and Postel, and not so much in Peter Ramus as in his mortal enemy Jacques Charpentier. Even in René Descartes there are strong elements of Platonism. Outside of France, Desiderius Erasmus, Thomas More, Sebastian Fox Morcillo, Paracelsus, Cornelius Agrippa, and finally Johannes Kepler exemplify the importance of Platonism in sixteenth-century thought, an importance that is closely linked with the writings, translations, and commentaries of Ficino.

In the seventeenth century, after Galileo and Descartes, the speculative cosmology of the Renaissance was no longer possible within the framework of a natural science based on experiments and mathematical formulas. The influence of Platonism persisted, however, in the metaphysics and epistemology of Benedict de Spinoza and Gottfried Wilhelm Leibniz, Nicolas Malebranche and George Berkeley; and it gained a new life in the school of Cambridge Platonists. And, since the authority of Plato himself remained a powerful force with many thinkers, we find even in Immanuel Kant and Johann Wolfgang von Goethe several theories associated with the name and prestige of Plato (and Plotinus) that actually belong to his Florentine translator and commentator. Samuel Taylor Coleridge wrote in his autobiography that as a youth he

read Plato and Plotinus, together with the commentaries and the *Theologia Platonica* of the illustrious Florentine. It was only in the nineteenth century that Ficino lost even this anonymous or pseudonymous influence, after a new school of philological and historical criticism had begun to make a rigorous distinction between the genuine thought of Plato and that of his successors and commentators in late antiquity and during the Renaissance. On the basis of this distinction, it has become possible again to understand Ficino's thought in its own right—to appreciate its indebtedness to sources other than Plato, its close connection with the thought and scholarship, art and literature of its time, and its own peculiar style and originality.

See also Florentine Academy.

Bibliography

Of Ficino's main works, only the commentary on Plato's *Symposium* has been reprinted in modern editions. See *Marsilio Ficino's Commentary on Plato's Symposium,* translated and edited by Sears R. Jayne, Vol. XIX, No. 1, in The University of Missouri Studies (Columbia, MO, 1944), and *Commentaire sur le Banquet de Platon,* edited by Raymond Marcel (Paris, 1956). For the others, the Basel editions of his collected works, *Opera Omnia,* 2 vols. (1561 and 1576), are still the only source; the second edition is slightly more complete and has been reprinted more recently (Turin, 1959). Also see *Supplementum Ficinianum,* edited by Paul Oskar Kristeller, 2 vols. (Florence, 1937), with lists of manuscripts and editions and a chronology of Ficino's works, and P. O. Kristeller, *Studies in Ficino's Thought and Letters* (Rome, 1956), with a number of papers discussing Ficino's life, sources, influence, and thought. Aside from the commentary on the *Symposium,* only the short *Five Questions concerning the Mind* (from Ficino's *Letters*) has been translated into English, by Josephine L. Burroughs, in *The Renaissance Philosophy of Man,* edited by Ernst Cassirer and others (Chicago: University of Chicago Press, 1948), pp. 193–212.

For Ficino's life and the history of the Platonic Academy, see Arnaldo Della Torre, *Storia dell'Accademia platonica di Firenze* (Florence, 1902), which is still indispensable, and Raymond Marcel, *Marsile Ficin* (Paris, 1958).

For discussions of Ficino's philosophy, see Giuseppe Saitta, *La filosofia di Marsilio Ficino* (Messina: Principato, 1923; 3rd ed., Bologna: Fiammenghi and Nanni, 1954); Sears R. Jayne, *John Colet and Marsilio Ficino* (London: Oxford University Press, 1963); Paul O. Kristeller, *II pensiero filosofico di Marsilio Ficino* (Florence: Sansoni, 1953), translated by Virginia Conant as *The Philosophy of Marsilio Ficino* (New York: Columbia University Press, 1943)—the Italian edition is superior because of its additional indexes and original text quotations.

Also see Michele Schiavone, *Problemi filosofici in Marsilio Ficino* (Milan, 1957), and Ernst Cassirer, *Individuum und Kosmos in der Philosophie der Renaissance* (Berlin and Leipzig: Teubner, 1927; reprinted Darmstadt, 1962), translated by Mario Domandi as *The Individual and the Cosmos in Renaissance Philosophy* (New York: Harper and Row, 1963).

Ficino's influence on the theory and iconography of Renaissance art has been discussed by Erwin Panofsky—*Idea* (Leipzig and Berlin: Teubner, 1924; 2nd ed., Berlin: Hessling, 1960) and *Studies in Iconology* (New York: Oxford University Press, 1939; reprinted, New York: Harper and Row, 1962)—and by E. H. Gombrich and others; see especially André Chastel, *Marsile Ficin et l'art* (Geneva: Droz, 1954). For a discussion of Ficino and magic, see D. P. Walker, *Spiritual and Demonic Magic from Ficino to Campanella* (London: Warburg Institute, University of London, 1958).

Literature on the influence of Ficino's theory of love includes A. M. J. Festugière, *La philosophie de l'amour de Marsile Ficin et son influence sur la littérature française du XVIe siècle* (Paris: Vrin, 1941); John C. Nelson, *Renaissance Theory of Love* (New York: Columbia University Press, 1958). See also Nesca A. Robb, *Neoplatonism of the Italian Renaissance* (London: Allen and Unwin, 1935), and Raymond Klibansky, *The Continuity of the Platonic Tradition during the Middle Ages* (London: Warburg Institute, 1939 and 1950).

OTHER RECOMMENDED TITLES

Allen, Michael J. B. "The Absent Angel in Ficino's Philosophy." *Journal of the History of Ideas* 36 (1975): 219–240.

Allen, Michael J. B. "The Ficinian Timaeus and Renaissance Science." In *Plato's Timaeus as Cultural Icon,* edited by Gretchen Reydams-Schils. Notre Dame, IN: University of Notre Dame Press, 2003.

Allen, Michael J. B. "Marsilio Ficino on Significatio." *Midwest Studies in Philosophy* 26 (2002): 30–43.

Allen, Michael J. B. *The Platonism of Marsilio Ficino: A Study of His Phaedrus Commentary, Its Sources and Genesis.* Berkeley: University of California Press, 1984.

Collins, Ardis B. "Love and Natural Desire in Ficino's Platonic Theology." *Journal of the History of Philosophy* 9 (1971): 435–442.

Collins, Ardis B. "The Secular Is Sacred: Platonism and Thomism in Marsilio Ficino's *Platonic Theology.*" The Hague: Nijhoff, 1974.

Devereux, James A. "The Object of Love in the Philosophy of Marsilio Ficino." *Journal of the History of Ideas* 30 (1969): 161–170.

Fogarty, Anne. "Mystical Modernism: Yeats, Ficino and Neoplatonic Philosophy." *Accademia* 3 (2001): 131–147.

Gandillac, Maurice de. "Neoplatonism and Christian Thought in the Fifteenth Century." In *Neoplatonism and Christian Thought,* edited by Dominic J. O'Meara. Norfolk, VA: International Society for Neoplatonic Studies, 1981.

Hendrix, John. *Platonic Architectonic: Platonic Philosophies and the Visual Arts.* New York: Peter Lang, 2004.

Jacob, Alexander. "Henry More's 'Psychodia Platonica' and Its Relationship to Marsilio Ficino's 'Theologia Platonica.'" *Journal of the History of Ideas* 46 (1985): 503–522.

Kuczyanska, Alicja. "The Third World of Marsilio Ficino." *Dialectics and Humanism* 15 (1988): 157–171.

Mojsisch, Burkhard. "The Epistemology of Humanism." *Bochumer Philosophisches Jahrbuch* 1 (1996): 127–152.

O'Rourke Boyle, Marjorie. "Pure of Heart: From Ancient Rites to Renaissance Plato." *Journal of the History of Ideas* 63 (1) (2002): 41–62.

Paul Oskar Kristeller (1967)
Bibliography updated by Tamra Frei (2005)

FICTIONALISM

A fictionalist is one who aims to secure the benefits of talking *as if* certain kinds of things exist—numbers, moral properties, possible worlds, composite objects, or whatever—while avoiding commitment to believing in their existence. This understanding of fictionalism is broad and ecumenical, and it should be noted that *fictionalism* is frequently used in the recent literature to refer to one or other of the more specific doctrines that this entry discusses.

1. FICTIONALISM AND FICTIONS

Consider paradigm cases of works of fiction such as J. R. R. Tolkein's *Lord of the Rings* or Charles Dickens's *Christmas Carol*. On some occasions, such works of fiction are taken as the object of philosophical enquiry and explanation. In that case, as one would expect, there are competing philosophical accounts of the nature of fiction—competing answers to such questions as the following: whether hobbits exist; whether it is true that Ebenezer Scrooge was visited by Marley; how, or why, people come to rejoice in Scrooge's redemption, or whether *Lord of the Rings* might have existed if Tolkein had not. On other occasions, however, works of fiction (in general) are invoked to explain (by analogy) various philosophically interesting discourses that are not obviously works of fiction or intended to be. In that case, certain answers to questions about the ontology and language of works of fiction are taken as given, and the workings of other discourses are accounted for by analogy with works of fiction so construed—hence the use of the term *fictionalism* to describe accounts of this sort. Fictionalists need not explicitly propose an analogy with works of fiction. The link between works of fiction and fictionalism is best forged as follows.

There are three natural and plausible theses about fictions that fictionalists typically wish to transfer to the discourses that are the targets of their explanations. Firstly, a thesis of vindication: Fictional discourses do not call for elimination or rejection, nor may they be simply ignored. We should neither discourage novelists from writing stories nor prohibit literary critics from discussing fictional characters: Fiction fulfills some function in our lives and calls for a philosophical account that acknowledges that function. Secondly, an ontological attitude: We should not accept the existence of characters and kinds that are paradigmatically fictional. For instance, we should not believe in the existence of hobbits or Ebenezer Scrooge. Thirdly, a semantic thesis: It is not the case that any sentence that appears to be about fictional entities both (a) entails the existence of a fictional entity and (b) is literally true. This thesis is particularly important in the case of those sentences of the fictional narrative that are paradigmatically correct or true according to the fiction—for example, "Some hobbits live underground" and "Scrooge is the employer of Bob Cratchit." These sentences appear to entail the existence of fictionalia such as hobbits and Scrooge, and they appear to be true, so that anyone who accepts both these appearances will seemingly be committed to believing in the existence of hobbits and of Scrooge.

Fictionalists turn these theses about fiction into claims about the discourse for which they are accounting. Characterizing fictionalism this way enables us to distinguish fictionalism about a discourse (*F*-talk, say) from the most eminent rival approaches to interpretation and ontology—namely, eliminativism, realism, and reductionism. Firstly, the fictionalist's thesis of vindication says that we are well motivated in persisting in our use of *F*-talk because it serves some characteristic function or purpose that cannot effectively or efficiently be replicated if we abstain from *F*-talk. In contrast, the *eliminativist* characteristically denies that *F*-talk is so vindicated and proposes to abstain from its use. Secondly, the fictionalist's ontological attitude is that we ought not to accept the existence of *F*s. Thus the fictionalist is no *realist*—assuming that realists about *F*-discourse must believe in *F*s. We should also expect the realist to reject the characteristic semantic thesis of fictionalism and so to insist that there are *F*-sentences that both (a) entail the existence of *F*s and (b) are literally true. Thirdly, the fictionalist does not accept the existence of *F*s and so, a fortiori, does not accept the conjunctive thesis that the *F*s exist and (for some *G*) are identical with the *G*s. This separates the fictionalist from ontological reductionists, who assert this conjunctive thesis.

2. FICTIONALIST STRATEGIES OF INTERPRETATION

As has been seen, the fictionalist wishes to exploit the thesis that no sentence of fictional discourse both (a) entails the existence of any fictional entity and (b) is literally

true. Different general strategies of fictionalist interpretation correspond to different ways of rejecting (a) or of rejecting (b). To spell out these strategies, let us focus on the kind of sentence that may look like a counterexample to the view of fiction on which the fictionalist is drawing—one that appears to entail the existence of a fictional entity:

(1) Some hobbits live underground.

The first kind of fictionalist strategy proceeds from denial of the entailment component (a) of the thesis. It claims that the sentences do not have the existential consequences that they appear to have. If this claim is secured, then in order to avoid unwanted existential commitment, the fictionalist need not deny that the sentences in question are true. This entry distinguishes three versions of this strategy, modeled on different accounts of the nature of fiction:

(A1) NONFACTUALISM (RELATED TERMS: NONCOGNITIVISM, EXPRESSIVISM, INSTRUMENTALISM). The sentences of the fictional narrative are not used with the kind of illocutionary force or intent that is required in order for them to state a propositional, genuinely truth-evaluable content. The proper use of the sentence is in a kind of illocutionary act that precludes assertion or presenting the content of a belief. Accordingly, sentences so used are not truth-apt and lack the kinds of content that can be, or properly entail, any existential proposition. In fictional narratives, perhaps the crucial illocutionary act is that of fictionalizing or pretending. The nonfactualist strategy of interpretation is familiar from the moral and aesthetic cases, where fact-stating is contrasted with evaluating or attitude-expressing, but is also applied in certain instrumentalist approaches to (portions of) mathematics and science. The locus classicus of ethical nonfactualism is the emotivism of A. J. Ayer (1936, ch. 6). For an instrumentalism about infinitary portions of mathematics see David Hilbert's article "On the Infinite" (1983). Simon Blackburn presents a list of nonfactualist (expressivist) suggestions about a further range of discourses as background to his own nonfactualism (quasi-realism) about morals and modals (1984, chs. 5–6).

(A2) NONEISM (RELATED TERM: MEINONGIANISM). Sentences such as (1) may be used assertorically, they have propositional content, but they do not entail the existence of hobbits because, generally, propositions of particular quantification (some *A*s are *B*s) do not entail the existence of those things over which they quantify. On this view, it is consistent to hold, in general, that there are some things that do not exist and, in particular, that among some of the things that do not exist are hobbits that live underground, or infinitely many prime numbers or worlds that have talking donkeys as parts. A comprehensive, noneist fictionalism would treat all apparently existential quantification in the true propositions of the discourse as particular quantification, and would treat all such particular quantification as quantification that is not existentially committing. The locus classicus of noneism is Richard Routley (1980; see also McGinn 2000, ch. 2).

(A3) PARAPHRASIS. Sentences of the fictional narrative are elliptical expressions of propositions that do not entail the existence of fictionalia; correspondingly, sentences of *F*-discourse do not entail the existence of *F*s. One prominent development of this thought has it that sentences such as (1) express propositions in which nonfactive operators take position of widest scope—operators of modality, conditionalization, or consequence operators invoking a story (or theory). For example:

(1*) According to the Tolkein stories, some hobbits live underground;

or,

(1**) The Tolkein stories entail that some hobbits live underground.

The modal fictionalist introduced in Gideon Rosen's article "Modal Fictionalism" (1990) claims that one can have all the benefits of talking about possible worlds without the ontological costs by interpreting apparently existential claims about a plurality of worlds as claims about what is the case according to the plurality of worlds hypothesis advanced by David Lewis (1986). (Armstrong 1989 puts forward a related view; for discussion, see Lycan 1993). In the philosophy of mathematics, Geoffrey Hellman's (1989) modal structuralism is a proposal to treat apparently existential claims about numbers as claims about what would follow from the hypothesis that certain structures are instantiated. The presentism of A. N. Prior (1957) incorporates a proposal to paraphrase away apparent reference to past and future times by translation into a medium of tensed operators.

The second kind of fictionalist strategy proceeds from refusal to accept the component thesis (b)—that the sentences in question are literally true. A choice then presents itself. One might proceed by multiplying kinds of truth and make out the case that there is a feature of nonliteral truth or relativized truth that can be attributed to the "correct" sentences of the discourse. Thus, it might be held that (1) is fictively true or (in a metalinguistic analogue of an earlier proposal) is true-in-the-Tolkein-

stories but not literally true or true simpliciter. Going the other way, one might stick with only one univocal notion of (literal) truth but refuse to ascribe (univocal) truth to (1) or any other correct existential sentence of the fiction. Here only the latter, better-charted path will be explored. If the fictionalist refrains from holding that any existential sentence is true, she may treat those sentences as expressing the existential propositions that they appear to express without having immediate cause to worry that unwanted existential commitment will ensue. Three versions of this strategy are as follows:

(B1) PRESUPPOSITIONISM. When sentences containing special fictional terms occur outside the scope of a story operator ("According to T …") they simply lack a truth-value. This view might be supported by the contention that in order for some such sentences to be true or false, the existential presuppositions invoked by the use of the narrative—presupposition of the existence of hobbits and other kinds of things—would have to be fulfilled. For such a view of sentences involving predications to empty definite descriptions see P. F. Strawson's essay "On Referring" (1971).

(B2) AGNOSTICISM. The problematic existential sentences of the discourse may properly be used to assert the existence of *F*s and do have a determinate truth-value. But people are not, and perhaps cannot be, in a position to judge what that truth-value is, and certainly not in a position to assert any such sentence. The most famous agnostic fictionalism is that of Bas C. Van Fraassen (1980) on the unobservables of microphysical theory. Rosen and Cian Dorr (2002) proposed agnosticism about the existence of the composite objects—entities that have other entities as parts—which people's ordinary talk is about.

(B3) ERROR THEORY. The problematic existential sentences are assertoric of the existence of *F*s and do have determinate truth-value, and one is justified in holding that these sentences are systematically false. J. L. Mackie (1977) famously reads Locke as an error-theorist about secondary properties and develops a parallel error theory of morals. Mackie interprets ordinary moral judgements as requiring the existence of "objective prescriptivity"—states of affairs that give reasons for action independently of the agent's motivations—and thus as systematically false. However, Mackie goes on to recommend that one persevere with moral discourse in order to secure the benefits of social cooperation. The moral fictionalism of Richard Joyce (2001) develops this position, suggesting that people should cease to believe moral claims while continuing to utter sentences such as "Kicking babies is wrong," provided they do not do so with assertoric force.

Error theories about mathematics abound. Hartry Field (1989) construes the existential sentences of mathematical theories as entailing the existence of abstract objects. On epistemological grounds one should hold these sentences false. But, Field maintains, because reference to mathematical entities can be removed from the best physical theories, one is entitled to continue using theories that contain mathematics and is motivated to do so because it offers significant shortcuts in inference and calculation. Other mathematical error theorists develop positions that are not committed to such dispensability of quantification over abstract entities. Joseph Melia (2000) claims that mathematical sentences can convey useful information about the concrete part of the world, even though they are often false through entailing the existence of abstract objects.

When the fictionalist proceeds along any of the (b)-route strategies, she will typically offer as an alternative to truth some other subsidiary norm that the "correct" sentences satisfy and in terms of which the success or characteristic function of the discourse is to be explained. What the norm is will differ from discourse to discourse. For example, Van Fraassen's refusal to hold true the existential sentences of microphysical theory, and the theories that entail them, is combined with the views that the good theories in which they feature are good because they are empirically adequate and that one can explain their success without ascribing truth to them. Similarly, Field holds that mathematical "goodness" is not truth but membership of a nominalistically conservative theory (compare Rosen and Dorr 2002 on the atomistic adequacy of ordinary composite object talk). However, the fictionalist who proceeds along the (b)-route in withholding the ascription of truth may not feel compelled to appeal to any norm other than truth. The alternative is to maintain that the discourse could still fulfill its function if people limited themselves to believing or holding true only its *nonexistential* claims.

It is not the case that all versions of fictionalism are presented explicitly as versions of these strategies. Often other explanatory resources are prominent (e.g., quasi-assertion, games of make-believe, and metaphor. See Yablo 2001, Walton 1990). However, it is suspected that in order to avoid relevant ontological commitment, fictionalists must eventually commit to one of the semantic strategies presented here.

3. ISSUES FOR THE FICTIONALIST STRATEGIES OF INTERPRETATION

A fictionalist strategy of interpretation will succeed only if:

(1) It avoids the ontological commitments that it is intended to avoid;

(2) it secures the benefits that motivate persistence with use of the discourse; and

(3) it makes intelligible whatever distinctive semantic devices that it invokes in orderto escape ontological commitment.

Nonfactualist fictionalism faces its primary challenges over points (2) and (3). In particular, if apparently existential sentences do not, at bottom, state propositions, then how are nonfactualists to square whatever linguistic properties they ascribe to those sentences with the role of the sentences as premises in inferences? This point bites hard when the target discourse is one that functions as an inferential medium as, notably, portions of mathematics and scientific theory do with respect to observation statements.

Noneist fictionalism faces the sharpest challenge of all under point (3), for it has been held that the pivotal claim of noneism—that there are things that do not exist—is unintelligible. Moreover, there is a particular methodological challenge to be faced by all fictionalists that the noneist in particular seems to invite. Suppose that a fictionalist succeeds in avoiding commitment to some problematic realm of objects. It would smack of absurdity if their strategy could be applied globally, to free everyone of all ontological commitments whatsoever—even in cases where those commitments were not undesirable. The fictionalist who holds that some but not all discourses can be treated fictionally needs a principled way of drawing the line.

Paraphrastic fictionalists who invoke nonfactive operators immediately face a dilemma over point (3). If the operators in question are taken as primitive, that is at least an ideological cost of the theory; but if the (modal, consequence, or conditional) operators in question are interpreted in standard ways, then they may generate commitment to entities such as models (sets) or possible worlds. But these semantically induced entities may either be, or share problematic features with, the entities that the fictionalist is trying to avoid—thus contravening (1). For example, a fictionalist about numbers who is suspicious of abstract objects had better not end up invoking other abstract objects in order to explain what "according to standard arithmetic" means. Along the first horn of the dilemma the question arises again, as it did for the nonfactualist, of how to account for the (perhaps crucial) inferential role of the sentences. Often it seems that one could do so if the existential sentences were interpreted as existence-entailing; the paraphrastic fictionalist has to demonstrate that her alternative construal can do the same work.

Another difficulty arises when the nonfactive operator chosen is of the sort "According to T," where T is a philosophical theory the fictionalist holds to be false. (Rosen-style modal fictionalism is an example of a fictionalist theory that invokes such an operator.) It seems reasonable to ask the fictionalist why any philosophical theory—especially one the fictionalist holds to be false—should play such a central role in her account of the discourse. In addition, it may be asked why the fictionalist chooses to use the particular theory that she does. What distinguishes that theory from other philosophical falsehoods?

The various route (b) strategies that withhold ascription of literal truth face their most immediate difficulty over (2). Can the characteristic function of the discourse really be secured if all of the intuitively correct sentences are held to have some feature that is weaker than truth or if only the nonexistential intuitively correct sentences are held to be true? A question that lies just beyond this one is whether proper strictures of charitable interpretation will permit an interpretation of the discourse that ascribes such powerful existential entailments and then sets the standard of truth so high (and so distinct from the operating standards of correctness) that the users of the discourse systematically fail to meet them.

4. HISTORICAL POSTSCRIPT

These contemporary fictionalist views have many historical antecedents. The phenomenalist strand of empiricist thought, as represented, for example, in John Stuart Mill's *An Examination of Sir William Hamilton's Philosophy* (1979 [1865]) is paraphrastically fictionalist when it suggests the translation of ordinary material object talk into counterfactual claims about sensations. The nonfactualist strand of empiricist thought, which is often discerned in Hume (1978), can be viewed as suggesting fictionalism about a wide range of cases from the external world to the self. More specifically the notion of paraphrase, and its function in avoiding ontological commitment, is prominent in the thought of Jeremy Bentham in the nineteenth century and in that of W. V. Quine in the twentieth century. Arguably, this style of fictionalism also surfaces in Bertrand Russell's doctrine that classes are "logical fic-

tions," eliminable through contextual definition, and in his famous general treatment of definite descriptions (1956). Finally, in the early twentieth century, Hans Vaihinger (1924) proposes that one should accept atomic theory, theology, and many other discourses, without believing them. Vaihinger presents his views as a reading of Kant, though recent commentators have stressed his affinities with pragmatist thought.

See also Agnosticism; Ayer, Alfred Jules; Eliminative Materialism, Eliminativism; Error Theory of Ethics; Field, Hartry; Hilbert, David; Hume, David; Kant, Immanuel; Lewis, David; Literature, Philosophy of; Locke, John; Mackie, John Leslie; Mill, John Stuart; Noncognitivism; Prior, Arthur Norman; Propositions; Quine, Willard Van Orman; Strawson, Peter Frederick; Vaihinger, Hans; Van Fraassen, Bas.

Bibliography

Armstrong, D. M. *A Combinatorial Theory of Possibility*. Cambridge, U.K.: Cambridge University Press, 1989.

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1936.

Blackburn, Simon. *Spreading The Word: Groundings in the Philosophy of Language*. Oxford: Clarendon Press, 1984.

Field, Hartry. *Realism, Mathematics, and Modality*. Oxford: Blackwell, 1989.

Hellman, Geoffrey. *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Oxford University Press, 1989.

Hilbert, David. "On the Infinite." In *Philosophy of Mathematics: Selected Readings*. 2nd ed., edited by Paul Benacerraf and Hilary Putnam. Cambridge, U.K.: Cambridge University Press, 1983.

Hume, David. *A Treatise of Human Nature*. 2nd ed., edited by L. A. Selby-Bigge. Oxford: Oxford University Press, 1978.

Joyce, Richard. *The Myth of Morality*. Cambridge, U.K.: Cambridge University Press, 2001.

Lewis, David. *On The Plurality of Worlds*. Oxford: Blackwell, 1986.

Lycan, William G. "Armstrong's New Combinatorialist Theory of Modality." In *Ontology, Causality, and Mind: Essays in Honour of D. M. Armstrong*, edited by John Bacon, Keith Campbell, and Lloyd Reinhardt. Cambridge, U.K.: Cambridge University Press, 1993.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth, U.K.: Penguin, 1977.

McGinn, Colin. *Logical Properties: Identity, Existence, Predication, Necessity, Truth*. Oxford: Clarendon Press, 2000.

Melia, Joseph. "Weaseling away the Indispensability Argument." *Mind* 109 (2000): 455–479.

Mill, John Stuart. *An Examination of Sir William Hamilton's Philosophy* (1865), edited by J. M. Robson. London: Routledge, 1979.

Ogden, C. K. *Bentham's Theory of Fictions*. London: Kegan Paul, Trench, Trubner, 1932.

Prior, A. N. *Time and Modality*. Oxford: Oxford University Press, 1957.

Quine, W. V. "Existence and Quantification." In his *Ontological Relativity, and Other Essays*. New York: Columbia University Press, 1969.

Rosen, Gideon. "Modal Fictionalism." *Mind* 99 (1990): 327–354.

Rosen, Gideon, and Cian Dorr. "Composition as a Fiction." In *The Blackwell Guide to Metaphysics*, edited by Richard M. Gale. Oxford: Blackwell, 2002.

Routley (Sylvan), Richard. *Exploring Meinong's Jungle and Beyond: An Investigation of Noneism and the Theory of Items*. Canberra, Australia: Research School of Social Sciences, Australian National University, 1980.

Russell, Bertrand. "The Philosophy of Logical Atomism." In *Logic and Knowledge*, edited by R. C. Marsh. London: Allen and Unwin, 1956.

Strawson, P. F. "On Referring." In his *Logico-Linguistic Papers*. London: Methuen, 1971.

Vaihinger, Hans. *The Philosophy of "As If": A System of the Theoretical, Practical and Religious Fictions of Mankind*. Translated by C. K. Ogden. London: Kegan Paul, Trench, Trubner, 1924.

van Fraassen, Bas C. *The Scientific Image*. Oxford: Clarendon Press, 1980.

Walton, Kendall. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press, 1990.

Yablo, Stephen. "Go Figure: A Path through Fictionalism." *Midwest Studies in Philosophy* 25 (2001): 72–102.

John Divers (2005)
David Liggins (2005)

FICTIONALISM, MATHEMATICAL

See *Nominalism, Modern*

FIDEISM

Fideism is the view that truth in religion is ultimately based on faith rather than on reasoning or evidence. This claim has been presented in many forms by theologians from St. Paul to contemporary neoorthodox, antirationalist writers, usually as a way of asserting that the fundamental tenets of religion cannot be established by proofs or by empirical evidence but must be accepted on faith. Some forms of fideism denigrate or deny the value of reason and science, and these amount to a kind of irrationalism, as indicated in David Hume's ironic statement at the end of his essay "Of Miracles":

> [The] *Christian Religion* not only was at first attended with miracles, but even to this day can-

> not be believed by any reasonable person without them. Mere reason is not sufficient to convince us of its veracity; and whoever is moved by *Faith* to assent to it, is conscious of a continued miracle in his own person, which subverts all the principles of his understanding, and gives him a determination to believe what is most contrary to custom and experience. (*Essay concerning Human Understanding,* edited by L. A. Selby-Bigge, Oxford, 1951, p. 131)

EXTREME FIDEISM

Extreme fideists such as J. G. Hamann and Søren Kierkegaard have praised Hume's formulation as a proper characterization of religious orthodoxy.

Starting with St. Paul's contention that the central doctrine of Christianity was nonsensical by Greek philosophical standards and with Tertullian's announcement *credo quia absurdum* (I believe that which is absurd), there have been theologians who have insisted that religious truths are contrary to those that might be supported or justified by reasonable evidence and that rational activities are not proper means to arrive at such truths. Some have insisted that there are suprarational or extrarational ways, such as mystical or revelatory experiences that provide the "knowledge" of fundamental truths. Such writers have tended to ignore rational arguments or standards, and often, as St. John of the Cross did, they have offered means by which one could train oneself to escape the confines of rationality in order to intensify religious experience and belief.

Others have tried to show the inability of reason to establish any fundamental or absolutely certain truth. Usually employing skeptical arguments, they have contended that ultimate principles are open to question or rational standards, and also that these standards can themselves be questioned. In view of this, they have contended, basic truths are to be accepted on faith. Michel Eyquem de Montaigne, Pierre Charron, and other so-called Christian skeptics set forth this form of fideism.

Others, such as Pierre Bayle, Kierkegaard, Félicité Robert de Lamennais, and Lev Isaakovich Shestov, went further and asserted that religious truths were of such a nature that they were contrary to the kinds of assertions that were probable, plausible, or even possible on rational or reasonable standards and that such truths could therefore be believed or accepted only on faith. Bayle insisted that religious tenets were not only above and beyond reason but also in opposition to it and that the strongest faith was that which denied the truths based on natural light and embraced those most incomprehensible to or contrary to reason. Kierkegaard first accepted the type of skepticism developed by Bayle and Hume about rational knowledge and then insisted that the fundamental tenet of Christianity, the Incarnation, is not only contrary to rational evidence but even a self-contradiction on rational standards: "No knowledge can have for its object the absurdity that the eternal is the historical" (*Philosophical Fragments, or A Fragment of Philosophy,* p. 50). Kierkegaard held Hamann's view that Hume had summed up the nature of religious belief—that it really is contrary to reason, custom, and experience. For Kierkegaard the very absurdity of the Christian claim makes it worthy of belief, and it is only by total commitment, or "the leap into faith," that it can be accepted. There can be no reasons for the leap, no justification for it. In the words of Bayle's opponent, Pierre Jurieu (also an irrationalist), all the believer can say is, "I believe it because I want to believe it."

In the twentieth century, among the fideists who advanced Kierkegaard's view, one of the most striking was the Russian Orthodox theologian Shestov, who insisted that the rejection of all rational standards is a part of true belief. In a commentary on Fëdor Dostoevsky he contended that the refusal to accept that $2 + 2 = 4$ and the ability to believe that $2 + 2 = 5$ are intimately connected with attaining religious truth.

MODERATE FIDEISM

In contrast to irrationalist or antirationalist fideism, a more moderate kind has developed, especially within the Christian Augustinian tradition. Rather than insisting that all ultimate certitude rests on faith in contrast to reason, this tradition has admitted that faith precedes reason in establishing certain fundamental truths but that reason and evidence can play some role both in the search for these truths and in the explanation and comprehension of them. The Augustinian slogan, *credo ut intelligam* (I believe in order to know), places the primary emphasis on faith. However, as Augustine's philosophical dialogues show, the recognition of the basic fideistic element may be (and perhaps must be) preceded by a rational search for truth. Once rational inquiry has revealed the need to accept some fundamental principles or beliefs on faith, then it may be possible to show that these commitments are reasonable, probable, or plausible. Purported proofs of the existence of God, metaphysical systems interpreting what is accepted on faith, and historical and psychological evidences about the nature of religion and its effects on believers can all be offered as rational explanations or even justifications of what has already been accepted on faith.

Blaise Pascal's presentation in the *Pensées* illustrates this form of fideism. He forcefully argued that the natural capacities of man are inadequate to lead him to any completely certain truth. A man can show that it is unreasonable or unwise to be an atheist but not that it is reasonable to be a theist. Once one has realized the human predicament—man's fundamental need for ultimate truths and his inability to find them—then one is ready to hear God and to accept his revealed Word on faith alone. Once one has faith, one can see the force of the apologetic and psychological evidence for the truth of the Christian religion. Such evidence might then constitute "good reasons" for believing what one has already accepted fideistically.

FIDEISM IN PHILOSOPHY

A nonreligious analogue of moderate fideism appears in various skeptical philosophical views, such as those of Hume, Bertrand Russell, and George Santayana. Hume's contention that it is belief that "peoples the world," and that everybody lives within his own private belief system, could be considered as a kind of fideism. The ultimate presuppositions by which we live cannot be justified by reason or evidence and are accepted not on religious faith but (to use Santayana's term) on "animal faith." Russell, in his *Human Knowledge,* insisted (on the basis of Hume's arguments) that the fundamental assumptions of science cannot be justified but must be accepted on faith. However, even if one has the mystical skeptical experience Santayana described, of seeing all in doubt, it is still rational investigation that led Hume and his successors to the recognition of the belief factor involved in rational activities. Having discovered this, Hume then showed that one can study the causes of beliefs and that beliefs can be explained even though they can never be justified; working from the basis of a set of "reasonable" beliefs, one can evaluate other beliefs in terms of psychological factors. The philosophical tradition emanating from Hume, then, can be considered as a kind of fideism, sharing some of the characteristics of the moderate fideism of the Augustinian tradition.

CONTEMPORARY DEVELOPMENTS

At present irrationalistic fideism, especially of the Kierkegaardian variety, is extremely popular, especially among Protestant theologians (partly in reaction to liberal, rationalistic theological tendencies of the nineteenth century). Many theologians have been concerned with man's apparent inability to find any ultimate answers through science, secular political movements, and so forth, and his need to base his ultimate commitment on faith alone. The existentialist stress on the fundamental absurdity of man's world is part of this movement. The official Catholic position from the time of the Council of Trent to the present remains opposed to the central fideistic thesis, that ultimate beliefs can be established not by reason or evidence but only by faith. However, in a world in which so many optimistic, "reasonable," scientifically supported views have been undermined by the cataclysmic events of the twentieth century, fideism may provide one of the main avenues to some kind of significant belief for the present age. William James's analysis, in his "Will to Believe," of the psychological need for commitment and belief despite the lack of evidence may represent much of the present mood. The religious fideists, however, find James's own faith too tepid, and they seem to be moving more and more to the irrationalist fideism of Kierkegaard, Dostoevsky, and Shestov.

See also Atheism; Augustine, St.; Bayle, Pierre; Charron, Pierre; Dostoevsky, Fyodor Mikhailovich; Faith; Hamann, Johann Georg; Hume, David; James, William; John of the Cross, St.; Kierkegaard, Søren Aabye; Lamennais, Hugues Félicité Robert de; Montaigne, Michel Eyquem de; Pascal, Blaise; Russell, Bertrand Arthur William; Santayana, George; Shestov, Lev Isaakovich; Skepticism; Tertullian, Quintus Septimius Florens.

Bibliography

Bartley, William W. *Retreat to Commitment.* New York: Knopf, 1962. A critique of some contemporary fideistic views.

Bayle, Pierre. *Historical and Critical Dictionary, Selections.* Translated by R. H. Popkin. Indianapolis: Bobbs-Merrill, 1965. "Pyrrho" and the "Third Clarification" deal especially with fideism.

Evans, C. Stephen. *Faith beyond Reason: A Kierkegaardian Account.* Grand Rapids, MI: Eerdmans, 1998.

Geivett, R. Douglas, and Brendan Sweetman, eds. *Contemporary Perspectives on Religious Epistemology.* Oxford: Oxford University Press, 1992.

Harent, S. "Foi." In *Dictionnaire de théologie catholique,* Vol. VI, cols. 55–514. Paris, 1920.

Herberg, Will. *Four Existentialist Theologians: A Reader from the Works of Jacques Maritain, Nicolas Berdyaev, Martin Buber and Paul Tillich.* Garden City, NY: Doubleday, 1958.

Herberg, Will. *Judaism and Modern Man.* New York: Farrar, Straus, 1951.

Kierkegaard, Søren. *A Kierkegaard Anthology.* Edited by Robert Bretall. New York: Modern Library, 1959.

Kierkegaard, Søren. *Philosophical Fragments, or A Fragment of Philosophy.* Translated by David F. Swenson. Princeton, NJ: Princeton University Press, 1936.

Penelhum, Terrence. *God and Skepticism: A Study in Skepticism and Fideism.* New York: Kluwer Academic Publishers, 1983.

Pojman, Louis P. *Religious Belief and the Will.* London: Routledge and Kegan Paul, 1986.

Popkin, Richard H. "Kierkegaard and Scepticism." *Algemeen Nederlands Tijdschrift Voor Wijsbegeerte en Psychologie* 51 (1959): 123–141.

Popkin, Richard H. "Theological and Religious Scepticism." *Christian Scholar* 39 (1956): 150–158.

Santayana, George. *Scepticism and Animal Faith.* New York: Dover, 1955.

Sauvage, G. M. "Fideism." In *Catholic Encyclopedia,* Vol. V, 68–69. New York, 1909.

Shestov, Lev I. *In Job's Balances.* Translated by Camilla Coventry and C. A. Macartney. London: Dent, 1932.

Shestov, Lev I. *Kierkegaard et la philosophie existentielle.* Translated by T. Rageot and Boris de Schloezer. Paris: Vrin, 1936.

Swinburne, Richard. *Faith and Reason.* Oxford: Oxford University Press, 1984.

Tillich, Paul. *Dynamics of Faith.* New York: Harper, 1957.

Richard H. Popkin (1967)
Bibliography updated by Christian B. Miller (2005)

FIELD, HARTRY

(1956–)

Hartry H. Field was born in Boston. He received his BA in Mathematics at the University of Wisconsin (1967) and his Ph.D. at Harvard (1972) working under Hilary Putnam and Richard Boyd. He has taught at Princeton, USC, CUNY Graduate Center, and NYU, where he is currently Silver Professor of Philosophy. Field is the recipient of, among other awards, a Guggenheim Foundation Fellowship (1979–1980) and the Lakatos prize (1986) for his book *Science without Numbers* (1980). He was elected in 2003 to the American Academy of Arts and Sciences.

Field has made significant contributions in a number of areas. He is best known for his work in philosophy of mathematics and on a variety of issues connected with realism and with the notion of truth. In philosophy of mathematics, Field has defended a version of fictionalism: a view according to which mathematics, which he takes at face value as asserting the existence of numbers, pure sets, and so on, is literally false and cannot be interpreted via a nonliteral reading in such a way that it works out true. Field sees the central argument in favor of realism about mathematics to be its indispensability for formulating and making use of scientific theories, and he proposes to answer this argument by giving an account of the use of mathematics in the sciences that does not require that the mathematics be true: If T is a nominalistic physical theory (roughly, one that makes no mention of mathematical entities), and M is a mathematical theory used to derive consequences from T (an example of such a theory might be a version of set theory that allows one to treat the objects of T as urelements and that allows the vocabulary of T to appear in the comprehension axioms) then M is said to be conservative over T if any such consequences, if entirely stated in the vocabulary of T, are already (semantic) consequences of T—that is, true in any model of T.

Field points out that people have always expected mathematics to be conservative over physical theories, and that in fact there is good reason to believe it is. The importance of this observation is the following. Suppose P is a physical theory that, like most such theories, is not nominalistic. It may be possible to find a nominalistic theory N, from which one can derive P via definitions and mathematics. It will then follow that P and mathematics are jointly conservative over N. This at least suggests that N captures all the physical content of P, and that the mathematics (together with P itself) is simply a convenient device for drawing out the consequences of N. Following (and significantly extending) techniques familiar to decision theorists and others under the title of "measurement theory," Field succeeded in constructing a natural nominalistic N for the case where P is a form of Newtonian gravitation theory.

Field's project of extending this result to all of physics has stimulated widespread interest in a number of issues. To name just one, Newtonian gravitation, and any theory remotely like it, requires an N that quantifies over sets of points, which may be identified with regions of space; the sense of consequence in which anything about N provable in P + mathematics is already a consequence of N is second-order consequence, thought of as the complete logic of the part-whole relation. This raises interesting questions, both about the extent to which first-order approximations to Field's result are available or convincing, and about whether one can speak about second-order consequence while continuing to be a fictionalist about mathematics. Indeed, the latter question arises for first-order consequence, despite that it is coextensive with a syntactic notion—because a fictionalist about mathematics ought to be a fictionalist about, for example, the claim that a given theory is syntactically consistent. Field has responded to this question with an interesting theory of (purely) logical necessity as a sui generis kind of necessity, one that is not explained in terms of models or possible worlds.

Field's earliest work on truth, the essay "Tarski's Theory of Truth" (1972), appeared at a time when Putnam and others were trying to argue for a form of scientific realism that stressed, as against, for example, Thomas Kuhn, the continuity of reference across changing scientific theories. Integral to this view was a conception of reference that

made it a nontrivial question how use of the word "water" brings it about that "water" refers to the particular chemical compound it does, and thereby a nontrivial question, what brings it about that "Water tastes good," as uttered by an American, is true (beyond the fact that it does taste good). This conception, which sometimes goes (as do many other views) under the name "correspondence theory" (of reference or of truth), contrasts with the "deflationist" idea according to which "'water' refers to water (in English)" is nothing more than a straightforward consequence of a natural definition of "refers in English." In this paper and later related essays, Field forcefully articulated what has turned out to be the most persuasive argument in favor of the need for a correspondence theory: namely, that human success in interacting with the world using language requires a systematic explanation of a kind a deflationist is unable to supply.

It turns out that deflationists have some at least initially plausible responses to this argument, and in fact Field has been increasingly sympathetic to deflationism. One topic he has addressed is what the theory of meaning looks like from a deflationist perspective, given that deflationism needs to sever the apparently intimate connections between meaning and reference. Another has been what a deflationist (or anyone else—but the problem is particularly pressing for deflationists) is to make of situations where it seems correct to say that "there is no fact of the matter"; these include not only areas where philosophers have traditionally debated about realism, but also in borderline cases involving vague expressions like "bald." Field has presented an appealing picture in which one both abandons excluded middle, and introduces a "determinately" operator into the language. The "determinately" operator is not given a semantics; it is rather understood both through its connections with degrees of belief, and through its relations to a natural non-truth-functional conditional. Field shows that such a language allows one consistently (despite the presence of the "determinately" operator) to introduce a truth predicate T such that the Tarski sentences (written using the new conditional) work out to be theorems; in fact "T(A)" is everywhere substitutable for "A."

See also Mathematics, Foundations of.

Bibliography

The fictionalist approach to mathematics was first set out in Field's *Science without Numbers: A Defence of Nominalism* (Princeton, NJ: Princeton University Press, 1980). Many of Field's papers can be found in the collections *Realism, Mathematics, and Modality* (Oxford: Blackwell, 1989; rev. ed. 1991) and *Truth and the Absence of Fact* (New York: Oxford University Press, 2001).

For the recent work on vagueness and the liar, see "A Revenge-Immune Solution to the Semantic Paradoxes," in *Journal of Philosophical Logic* 72 (2003): 139–177; and "No Fact of the Matter," in *Australian Journal of Philosophy* 81 (2003): 457–480. Some recommended papers not in the collections, and on topics not mentioned in this entry, are "Logic, Meaning and Conceptual Role," in *Journal of Philosophy* 74 (1977): 379–409; "A Note on Jeffrey Conditionalization," in *Philosophy of Science* 45 (1978): 361–367; "The A Prioricity of Logic" *Proceedings of the Aristotelian Society* 96 (1996): 359–379; and "Causation in a Physical World," in *The Oxford Handbook of Metaphysics*, edited by M. Loux and D. Zimmerman, 435–460 (Oxford: Oxford University Press, 2003; reprinted in *The Philosopher's Annual* 26 (2003), edited by P. Grim, K. Baynes, and G. Mar). See also "Tarski's Theory of Truth," *Journal of Philosophy* 69 (13) (July 1972): 347–375.

Stephen Leeds (2005)

FIELDS AND PARTICLES

Broadly speaking, a field is a collection of properties ascribed to regions of space (one might also speak of the region itself as being "the field"); if the properties are quantifiable then the field is a mathematical function of spatial coordinates, $\Phi(x,y,z)$. Examples include the temperature at each point of a room, the velocity at each point of a fluid, the gravitational potential, and the electromagnetic field. In contrast—and broadly speaking—particles are entities of which positions are ascribed (and which lack any relevant internal structure). While these will do as broad characterizations, they are inadequate in a number of ways.

CLASSICAL FIELDS

For instance, one could say that a field theory ascribes positions (and field strengths) to the parts of a field, as a particle theory treats particles. Worse, one can reformulate particle theories (e.g., Isaac Newton, 1642–1727, and Immanuel Kant, 1724–1804) as theories that ascribe mobile particle-sized regions of repulsion to space: as a field theory according to the intuitive distinction. Hence a useful formal characterization adopts the practice of physicists and takes the difference between field and particle theories to be that the former associates infinitely many "degrees of freedom" (kinematically independent variables—the values of Φ at each point) with finite regions of space, but the latter only finitely many (the positions and momenta of a finite number of particles in a finite region).

The problem with the broad and formal characterizations of the field is that they ignore historically important distinctions. For instance, Aristotle's (384–322 BCE) plenum (i.e., space full of matter, with no vacuum) ascribes different properties—gravity here, levity there—to regions of space, but one would like to distinguish the modern concept of a field from the ancient plenum. Newton's (1687) gravitational field ascribes to every point of space a quantitative disposition for bodies to move (absent other bodies, if a body were at a point a distance r from a body of mass M then it would have acceleration proportional to M/r^2), which distinguishes it from the ancient plenum. But understood literally, Newtonian gravity is a force that acts at a distance without mediation, hence Newton took it as a purely mathematical, "effective" description of some unknown underlying physics (which he sought in vain; in the early twenty-first century it is believed to be general relativity). Indeed, arguably the modern conception of the field is of something physical that mediates the long-range interactions between bodies. If so, Michael Faraday's (mid-nineteenth-century) arguments for the reality of the electromagnetic field are crucial. For instance, he distinguished physical from merely mathematical fields according to whether changes propagate at a finite speed or not (i.e., "through" the medium or not).

The atomists (especially Democritus, 460–370 BCE) rejected the plenum, arguing that the physical world could be understood in terms of atoms moving in the void. However, general rejection of the vacuum meant that this idea did not become the foundation of useful science until Descartes (1596–1650); and while he believed in the plenum, he envisioned it to be composed of particles of varying sizes. Although Descartes failed to derive quantitative consequences from atomism, his successors, up to the present, have found it one of the most fruitful ideas in physics.

QUANTUM FIELDS AND QUANTUM PARTICLES

In the twentieth century, quantum field theory (henceforth "QFT") was developed, and experimentally tested with unprecedented accuracy, particularly in particle accelerators. Classical fields can be decomposed into a sum of waves of different amplitudes (as a chord can be decomposed into different notes), which means, intuitively speaking, that quantum fields can be decomposed into a sum of waves with quantized (i.e, whole number) amplitudes. In quantum mechanics a wave(-function) represents a particle (its probability distribution in space), so there is a natural equivalence of a quantum field with a system of quantum particles, with the whole number amplitude of a wave in the decomposition representing the number of particles with that wavefunction. Thus because amplitudes become quantized, QFT is the best theory of both fields and subatomic particles.

However, the particle interpretation is only approximate: The field-particle distinction does not really dissolve in QFT. First, quantum mechanical superposition means that a quantum field may contain an indeterminate number of particles (e.g., two with some probability and three with another), which conflicts with the intuitive idea of a particle. Second, an accelerating observer will decompose a field into waves differently from a nonaccelerating observer; in particular, when the nonaccelerating observer says the field contains no particles, the accelerating observer will say that it does (these are known as "Rindler" particles). There is no contradiction, because if the accelerating observer captures a particle, he or she thereby changes the field to a state that all observers agree contains particles. All the same, the concept of a "particle" does not allow for the absence or presence of particles to be frame-dependent. Finally, there is a theorem that in relativistic QFT it is impossible to localize particles to any finite region; if so, they don't fit the intuitive idea of a particle at all.

According to formal definitions, QFT is a field, not particle theory, because it involves infinitely many degrees of freedom—a fact with profound consequences in quantum mechanics, which are obscured by the particle interpretation. Infinite degrees of freedom mean that there are many quantum versions of a field, some of which may not allow a particle decomposition at all (technically, there are unitarily inequivalent representations of the canonical commutation relations). One might think that observations of particles in the world show that the particle version is the correct one, but because of the Haag-Hall-Whiteman theorem there are reasons to think that realistic fields have no particle formulation (technically, there may be no Fock representation of an interacting field). If so, the appearance of particles is presumably explained by the correct version suitably approximating a system of particles. Specifically, there are field states arbitrarily close to states of particles infinitely far apart.

Quantum mechanics can also treat a system of particles, which is (modulo the previous discussion) a field for which the particle content is always definite. Beyond the fact that quantum particles are represented by wavefunc-

tions, there are important differences in the "identities" of classical and quantum particles that the following analogies illustrate. Classical particles are like badges with different pictures on them; the pictures make them distinguishable entities. Some quantum particles—bosons—are like money in the bank: Nothing distinguishes two of the dollars in an account from each other. Other particles—fermions—are like memberships in a particular club: Like money, one membership isn't any different from another; but unlike dollars, one can only have a single membership. (Technically, fermions satisfy the "exclusion principle": there can be at most one particle in any state.) To distinguish bosons and fermions, they are called "quanta"; however, these analogies fail to reveal that quantum mechanics allows other kinds of particles—"quarticles"—that differ from both quanta and classical particles.

See also Aristotle; Descartes, René; Faraday, Michael; Kant, Immanuel; Leucippus and Democritus; Newton, Isaac; Philosophy of Physics; Quantum Mechanics; Space.

Bibliography

Earman, John, and Doreen Fraser. "Haag's Theorem and Its Implications for the Foundations of Quantum Field Theory." Forthcoming.

Hesse, Mary. *Forces and Fields*. London: Greenwood, 1962.

Huggett, Nick. "Philosophical Foundations of Quantum Field Theory." *British Journal for the Philosophy of Science* 51 (2000): 617–638. Reprinted in *Philosophy of Science Today*, edited by Peter Clark and Katherine Hawley. Oxford: Oxford University Press, 2003.

Lange, Marc. *An Introduction to the Philosophy of Physics: Locality, Fields, Energy and Mass*. Oxford: Blackwell Publishers, 2002.

McMullin, Ernan. "The Origins of the Field Concept in Physics." *Physics in Perspective* 4 (2002): 13–39.

Pullman, Bernard. *The Atom in the History of Human Thought*. Translated by Axel R. Reisinger. Oxford: Oxford University Press, 2002.

Teller, Paul. *An Interpretive Introduction to Quantum Field Theory*. Princeton, NJ: Princeton University Press, 1995.

Nick Huggett (2005)

FILMER, ROBERT

(c. 1588–1653)

Robert Filmer, the English political writer and theorist of the divine right of kings, was an early expositor of the patriarchal account of the state and of society. He was a country gentleman of the county of Kent but also belonged to the intellectual society of London and had some connection with the Court. He was an associate of prominent lawyers and historians, such as John Selden and Sir Henry Spelman, of the orthodox clergy, and of the Jacobean poets and literati too, including George Herbert and possibly John Donne. His absolutist views on political matters may have been acquired while he was at Trinity College, Cambridge, or at the Inns of Court and were developed well before the outbreak of the Civil War between the king and Parliament in 1642. In this he resembles Thomas Hobbes, his contemporary, but Filmer wrote his works for circulation in manuscript among his London acquaintances and the manor houses of Kent rather than for publication in print. Although his family was engaged on the side of the king in the struggle with Parliament in the 1640s and although he himself suffered considerable losses, Sir Robert never actually fought with the royalist forces and even pleaded neutrality, which has since been looked upon as inconsistency in the conduct of an extreme defender of royalist claims. His neutrality did not prevent his being sent to prison for a time.

Filmer's importance in the history of thought rests almost entirely on the fact that John Locke's work on political theory, the famous *Two Treatises of Government,* was directed against him, though it was not published until 1689, nearly forty years after Filmer's death. It has only recently been shown how extensive was Locke's preoccupation with Filmer, in the second of his treatises as well as in the first. But the social theorists of the present day are also interested in Filmer's thinking as an expression of traditional patriarchal attitudes toward authority and social structure. The relationship between Locke and Filmer has become the classic example of a rationalist-critical political system (the Lockean) confronting an ideological-determinist outlook (the Filmer view).

It has not been possible, however, to see in Filmer simply a "codifier of unconscious prejudice," as he has been called. He was remarkably enlightened in some of his views, especially as to witchcraft, and wrote with surprising urbanity rather as a critical reviewer of the political works current in his time than as the solemn expositor of outraged orthodoxy. Those of his works he himself had printed, mainly reviews of Aristotle, John Milton, Hugo Grotius, and Hobbes, are brief and pointed, and it is, perhaps, significant that he refused to publish the only concerted exposition of his political theory, the famous *Patriarcha; or the Natural Power of Kings* (London, 1680), from which all the others derive. He may have thought his political theory too extreme in its earlier, positive form.

Patriarcha, which was composed for the gentry of Kent in the 1630s, asserts that every individual is absolutely bound to obey the political authority established in his country because that authority enjoys by divine decree the powers originally conferred on Adam at the creation over his wife, his children, and their descendants eternally. From this view of the Old Testament it follows that males are always superior to females, the elder to the younger, and that all humans are naturally—physiologically—related to each other. Society is a family, descended from one single male individual. All men are born, and always remain, unfree and unequal, and consent is irrelevant to political association. Political society is also universal, for there are no humans who are not descended from Adam. A prepolitical state of nature makes no sense at all, nor does any idea of a contract to replace such a condition by political society. Property as well as political power is distributed according to God's patriarchal decrees and belongs absolutely to the person who inherits it or to whom it has been given.

These social and political doctrines are original only in the sense that Filmer combined together many positions held by his predecessors, notably those of the French legal theorist Jean Bodin, those of the bishops of the Anglican Church, and especially those of its royal head, King James I. These views are acceptable only to a naively fundamentalist believer in the Christian scriptures, and Locke had no difficulty in demolishing all the "glib nonsense," as he called it, about the kingship of Adam and its descent to the Stuart kings, to the usurper Oliver Cromwell, to any man or group lucky enough to seize power. Nevertheless, there was rather more to Filmer's "rope of sand" than Locke wished to admit, and in Filmer's shrewd remarks about the historical absurdities of a state of nature and in his very acute analysis of majority rule he raised difficulties that Locke never satisfactorily overcame.

Filmer demanded to know how an assembly convened for the purpose of making a universal contract could ever proceed to a valid vote of everyone with the right to vote. There would be bound to be absentees, and when it came to original multitudes voting to set up a government, the rights of some individuals would inevitably be overrun. What about servants, women, children, and the sick? Locke blandly responded by dogmatically asserting that in "one Body Politick the Majority have the Right to act and conclude the rest" (*Second Treatise,* 95). Filmer's doctrine of property seems to have impelled Locke into the formulation of the labor theory of value, with all its enormous consequences in social thinking.

Filmer's doctrines by no means disappeared with the Glorious Revolution of 1688 and the victory of Lockean rationalism. His arguments were persuasively restated by Jonathan Boucher in championship of the Tories at the American Revolution, and again by George Fitzhugh in defense of the South in the 1850s. Filmer remains the most valuable literary source for traditional European preindustrial patriarchal political attitudes.

See also Aristotle; Bodin, Jean; Determinism, A Historical Survey; Grotius, Hugo; Hobbes, Thomas; Ideology; Locke, John; Milton, John; Political Philosophy, History of; Political Philosophy, Nature of; Social Contract.

Bibliography

PRIMARY WORK

Patriarcha and Other Writings. Edited by Johann P. Somerville. Cambridge, U.K.: Cambridge University Press, 1991.

SECONDARY WORKS

Daly, James. *Sir Robert Filmer and English Political Thought.* Toronto: University of Toronto Press, 1979.

Fitzhugh, George. *Cannibals All, or, Slaves without Masters.* Edited by C. Vann Woodward. Cambridge, MA: Belknap Press, 1960.

Laslett, Peter, ed. *Patriarcha and the Other Political Works of Sir Robert Filmer.* Oxford: Blackwell, 1949.

Schochet, Gordon J. *Patriarchalism in Political Thought.* New York: Basic, 1975.

Peter Laslett (1967)
Bibliography updated by Philip Reed (2005)

FINK, EUGEN
(1905–1975)

Eugen Fink was born and first educated in Konstanz, where his reading in philosophic classics (Giordano Bruno, Kant, and Nietzsche) began in the Gymnasium. He took up the formal study of philosophy at Freiburg in 1925 during a period of extraordinary richness: Edmund Husserl, in the chair previously held by Wilhelm Windelband and Heinrich Rickert, was at his peak in both philosophic labor and renown when he retired in 1928; he was succeeded by Martin Heidegger—Husserl's own choice—after several years at Marburg. Fink's dissertation under Husserl was completed in December 1929 with Heidegger as coevaluator (*Korreferent*), at a point in time when a

long-smoldering break between Husserl and Heidegger was fully manifest.

On Husserl's retirement he chose Fink as second research assistant, alongside Ludwig Landgrebe—who had been Husserl's assistant since 1923. As Landgrebe turned to his Habilitation, he relinquished his assistantship with Husserl (March 1930), and Fink, who was just then entering more closely into Husserl's regimen of work, became not only sole assistant with Husserl in his retirement, but indispensable. Their daily walks in the hillside park near Husserl's residence and the tasks Husserl had Fink take up in furthering and consolidating Husserl's manuscript studies made for a unique philosophic collaboration. Husserl drove himself to produce manuscript after manuscript in an effort to present new work to the public to demonstrate the breadth, solidity, and relevance of his phenomenology in the face of Heidegger's ascendancy, and Fink worked at projects of integration, critique, and recasting so as to bring Husserl's massive output to coherence and philosophic clarity. In particular, he was able—as Husserl was not—to come to terms in phenomenology with Heidegger's thinking, along with that of others such as Hegel and Nietzsche, who had not really figured into Husserl's consideration. At the same time, Fink worked on writings that would counter two misperceptions: that transcendental phenomenology was a brand of idealism not much different from neo-Kantianism, and that Husserl's logic-driven abstractness was incapable of dealing with the existential trenchancy of actual life in the world.

Fink's essay in *Kantstudien* (1933), "The Phenomenological Philosophy of Edmund Husserl and Contemporary Criticism" (Elveton 2003), attempted to counter the first misperception, and was widely influential upon the French grasp of Husserl's phenomenology, most notably in the work of Gaston Berger and Maurice Merleau-Ponty. Unfortunately National Socialism's taking power in January of 1933 cut short Fink's providing a similar defense against the second misperception, influentially expressed in Georg Misch's *Lebensphilosophie und Phänomenologie: eine Auseinandersetzung der Diltheyschen Richtung mit Heidegger und Husserl* (1931). At the same time, Fink's Habilitation project, "Sixth Cartesian Meditation: The Idea of a Transcendental Theory of Method," was prevented from being pursued in that it purveyed "Jewish" philosophy, namely Husserl's. Fink recounts that, as he was not of Jewish background, Nazi authorities tried to get him to abandon his work with Husserl. He would not, and as a result lost all prospects for an academic future in Germany. He remained with Husserl until Husserl's death in April 1938. During that time he managed to get but two articles published, the *Kantstudien* essay, appearing just as that journal was being "coordinated" to Nazi policy, and "What Does the Phenomenology of Edmund Husserl Want to Accomplish?" in *Die Tatwelt* (1934), a cultural review edited by one of the few resistance circles in Freiburg around the well-known political economist Walter Eucken. It was only in 1939, after Fink emigrated to Belgium subsequent to Husserl's death, that he was again able to publish his work; but that was not to last long.

Nevertheless, it was Fink's contribution to the ongoing final development of Husserl's phenomenological program that must be noted. The unpublished "Sixth Meditation" was read and reread by Husserl, bringing home to him the need to be theoretically self-critical about the status and character of transcendental assertions. More concretely, Fink's drafts of projects he was involved in with Husserl—paradigmatically exemplified in the two-volume German edition of the "Sixth Meditation" (Fink 1988)—showed how earlier formulations of transcendental phenomenology needed radical recasting in order for their philosophic sense to stand forth in coherence and relevance. Here one can see Fink's ability both to develop an integrative perspective on Husserl's work and to make the critical moves that would express the philosophic core of transcendental phenomenology, an ability for which Husserl valued Fink's "cophilosophizing" so highly.

Upon Husserl's death in 1938, the visit of the Belgian Franciscan, Herman Leo Van Breda, in search of materials for his dissertation, led to Van Breda's finding a way to move out of Germany all Husserl's manuscripts as well as his entire library. Van Breda also arranged for Malvine Husserl, now widowed, as well as Fink and Landgrebe, to emigrate to Louvain. This was accomplished by the spring of 1939, and the Husserl Archives were born; and, as it happened, Maurice Merleau-Ponty became the first visitor to consult Husserl's manuscript materials in its new home at the University of Louvain (April 1–6, 1939). Here Fink finally began university lecturing, and the work of transcribing and interpreting Husserl's shorthand manuscripts began anew, only to be ended in May 1940 with Germany's attack upon the Low Countries and the onset of a Europe-wide world war. By the end of that year, Fink and Landgrebe were back in Germany, excluded from university involvement.

After the war's end in 1946 Fink was given a beginner's position as lecturer at Freiburg University, and in 1948 finally took up there a newly established chair in *Philosophie und Erziehungswissenschaft*. Fink's work after the war was unlike that of others influenced by Husserl. Rather than explicate Husserl's writings, in the books of his own thinking he turned to developing the dimension of the phenomenological program that he found Husserl had left too implicit and unrealized, what he termed the "speculative" component of the program, the overarching philosophical sense of its findings (Fink 1957, 1958, 1960). He did, however, occasionally present essays on Husserl (Fink 1976) that were highly respected and accorded high authority—given his intimacy with Husserl's thinking; but these papers stood in some contrast to dominant interpretations other scholars made of phenomenology. Ultimately in his writings and lecturing Fink dedicated himself to ways of awakening listeners to philosophical questioning. Rather than establishing definitive theses, it was the ever-increasing radicality of realizing what lay in philosophical problems, what was at issue in them, that mattered most. He kept apart from the various postwar philosophical currents and avoided fostering a following of disciples. Heidegger was one who especially appreciated discussions with him, and his closest philosophical comrade was Jan Patoĉka, of unique renown and importance for his underground seminars in Prague under Communist rule.

See also Husserl, Edmund; Phenomenology.

Bibliography

WORKS BY FINK, INCLUDING THE ONLY TRANSLATIONS OF HIS WRITINGS INTO ENGLISH

VI. Cartesianische Meditation, Teil 1: Die Idee einer transzendentalen Methodenlehre, edited by Hans Ebeling, Jann Holl, and Guy van Kerckhoven. *Teil 2: Ergänzungsband*, edited by Guy van Kerckhoven, Husserliana Dokumenta II/1-2, Dordrecht: Kluwer, 1988.

Sixth Cartesian Meditation: The Idea of a Transcendental Theory of Method. Translated by Ronald Bruzina. Bloomington: Indiana University Press, 1995.

Nähe und Distanz: Phänomenolgische Vorträge und Aufsätze, edited by Franz-Anton Schwarz. Freiburg, Germany: Karl Alber, 1976.

"Operative Concepts in Husserl's Phenomenology." Translated by William McKenna. In *A priori and World: European Contributions to Husserlian Phenomenology*, edited by William McKenna, Robert M. Harlan, and Laurence E. Winters. The Hague: Martinus Nijhoff, 1981.

Sein, Wahrheit, Welt: Vor-Fragen zum Problem des Phänomen-Begriffs. The Hague: Martinus Nijhoff, 1958.

Spiel als Weltsymbol. Stuttgart: W. Kohlammer, 1960.

"The Phenomenological Philosophy of Edmund Husserl and Contemporary Criticism." In *The Phenomenology of Husserl, Selected Critical Readings*, edited by Roy Elveton. Chicago: Quadrangle Books, 2003.

"The Problem of the Phenomenology of Edmund Husserl." Translated by Robert M. Harlan. In *A Priori and World: European Contributions to Husserlian Phenomenology*, edited by William McKenna, Robert M. Harlan, and Laurence E. Winters. The Hague: Martinus Nijhoff, 1981.

"What Does the Phenomenology of Edmund Husserl Want to Accomplish?" Translated by Arthur Grugan. *Research in Phenomenology* 7 (1972): 5–27.

Zur Ontologischen Frühgeschichte von Raum-Zeit-Bewegung. The Hague: Martinus Nijhoff, 1957.

Ronald Bruzina (2005)

FIRST-ORDER LOGIC

First-order logic is a bag of tools for studying the validity of arguments. At base it consists of a family of mathematically defined languages called *first-order languages*. Because these languages are constructed to be "logically perfect" (in Gottlob Frege's phrase), we can guarantee from their grammatical form that certain arguments written in these languages are valid. Separately from this we can study how arguments in English or any other natural language can be translated into an appropriate first-order language. It was Gottfried Wilhelm Leibniz who in the 1680s first proposed to divide the study of arguments into a mathematical part and a translational part, though his notion of mathematical languages was barely adequate for the purpose. First-order languages first came to light in the work of Charles S. Peirce in the 1880s; his name for them was "first-intentional logic of relatives." It took some time to develop a satisfactory mathematical description of these languages. David Hilbert achieved this in his lectures at Göttingen in 1917–1922, which appeared in an edited form in his book *Grundzüge der Theoretischen Logik* with Wilhelm Ackermann. Many logicians reckon that the appearance of this book in 1928 marked the true birth of first-order logic.

LOGIC AND ARGUMENTS

For purposes of this article an argument consists of one or more sentences of English, then the word "Therefore," and then a sentence. The sentences before "Therefore" are called the *premises* of the argument and the sentence at the end is called its *conclusion*. We say that the argument is *valid* if the conclusion follows from the premises, and *invalid* otherwise. Logic is the study of valid arguments. Typical questions of logic are: Which arguments are valid

and which are invalid? How can we construct valid arguments?

We shall study these questions with the help of first-order languages. First-order languages differ from natural languages in several ways. One is that their vocabulary and grammar are precisely defined. A second and equally important difference is that they contain expressions that we can interpret in a range of possible ways, and the range is determined by the grammatical form of the expressions.

Take for example the first-order sentence

(1) $\forall x\ (P(x) \rightarrow Q(x))$.

The part in brackets,

(2) $(P(x) \rightarrow Q(x))$,

is read "If $P(x)$ then $Q(x)$" and for the moment we can read "$\forall x$" as "Whatever x may be." Here there is no object named "x." Rather, "x" is a variable ranging over a class of possible objects. We call this class the *domain of quantification*, or more briefly the *domain*. The domain is not supplied with the sentence itself; when somebody uses the sentence to make a statement, we have to be told what domain the user of the sentence intended.

We also have to be told how the expressions "$P(x)$" and "$Q(x)$" are interpreted. For the whole sentence (1) to make sense, each of these expressions must translate into a *predicate*, that is, an English sentence with the variable "x" standing where we could have put a name; for example

(3) x is a town in Italy.

(4) The number 5 + 3 is equal to x.

(5) The father of x is a pianist.

But there is a further requirement. It must be possible to paraphrase the predicate into the form

(6) x is a member of S

where S names a particular class of objects in the domain. For example the sentences (3)–(5) paraphrase as follows:

(3)' x is a member of the class of towns in Italy.

(4)' x is a member of the class of numbers that 5 + 3 is equal to.

(5)' x is a member of the class of individuals whose fathers are pianists.

To see why this restriction is needed, consider my father-in-law Marcus Ward, who was born on July 4th. He used to reason:

(7) Americans celebrate July 4th. July 4th is the birthday of Marcus Ward. Therefore: Americans celebrate the birthday of Marcus Ward.

This conclusion gave him constant pleasure. Unfortunately the argument is unsound. The predicate

(8) Americans celebrate x

allows two paraphrases into class form:

(9) x is in the class of events celebrated by Americans.

(10) x is in the class of days of the year celebrated by Americans.

On the first paraphrase, the second premise of the argument breaks down; the event that the Americans celebrate is not Marcus Ward's birthday. On the second paraphrase, the conclusion holds but not in the sense that pleased my father-in-law. Requiring a translation into classes is very effective for detecting ambiguities in arguments.

Returning to the sentence (1) we can interpret "$P(x)$" and "$Q(x)$" by saying what the classes in question are. So in an obvious notation, here is an interpretation of the first-order sentence above:

(11)

domain	the class of people now living
P	the class of pianists
Q	the class of musicians

On this interpretation the sentence (1) expresses that among all people now living, the pianists are musicians. Under this interpretation the sentence is true. Another possible interpretation is

(12)

domain	the class of people now living
P	the class of musicians
Q	the class of pianists

Under this second interpretation the sentence (1) is false; there are musicians who are not pianists. A crucial fact about first-order languages is that their sentences may change from true to false, or vice versa, when their interpretation is changed.

There are many first-order languages. Their common core consists of the following symbols, known as the *logical symbols*.

(13)

symbol	$\neg$	$\wedge$	$\vee$	$\rightarrow$
read as	'not'	'and'	'or'	'if…then'
alternative	~	&		$\supset$

(symbol)	$\leftrightarrow$	$\forall x$	$\exists x$	=
(read as)	'if and only if'	'for all x'	'there is x'	'equals'
(alternative)	$\equiv$	(x), $\wedge x$	(Ex), $\vee x$	

The third row of the table lists the commonest alternative notations, though in this entry we will use only the symbols on the top row.

Besides the logical symbols, each first-order language has its own collection of *nonlogical symbols*, sometimes known as *primitives*. These are symbols such as "*P*" and "*Q*" above. The nonlogical symbols need to be interpreted, and the language carries a rule specifying what kinds of interpretation are allowed. The set of all nonlogical symbols of a language, together with the information what kinds of interpretation are allowed, is called the *signature* of the language. An *interpretation* of the language consists of a domain and allowed interpretations of all the symbols in the signature of the language.

Below we shall see what kinds of nonlogical symbol a first-order language can have. But before we do that, we shall define a notion that brings us back to the difference between valid and invalid arguments.

Suppose L is a first-order language and $\phi_1, \ldots, \phi_n, \psi$ are sentences of L. Then the expression

(14) $$\phi_1, \ldots, \phi_n \vDash \psi$$

means that if I is any interpretation of L and all of $\phi_1, \ldots, \phi_n$ are true under I, then ψ is also true under I. We call the expression (14) a *semantic sequent*, or for short a *sequent*. We say that (14) is *valid* if it is true, and *invalid* otherwise. The sentences $\phi_1, \ldots, \phi_n$ are called the *premises* of the sequent, and ψ is called its *conclusion*.

An example of a valid sequent, using the kinds of symbol that we have already seen, is

(15) $$\forall x(P(x) \rightarrow Q(x)), \forall x(Q(x) \rightarrow R(x)) \vDash \forall x(P(x) \rightarrow R(x)).$$

In any interpretation I, the first sentence of (15) expresses that the class assigned to "*P*" is a subclass of the class assigned to "*Q*," and the second sentence expresses that the class assigned to "*Q*" is a subclass of the class assigned to "*R*." If these two sentences are true then it follows that the class assigned to "*P*" is a subclass of the class assigned to "*R*," and hence the conclusion of (15) is true under interpretation I. So (15) is valid.

Now suppose we have an argument written or spoken in English. Suppose also that we can find a suitable first-order language L, sentences $\phi_1, \ldots, \phi_n$ of L and an interpretation I of L, such that under the interpretation I the sentences $\phi_1, \ldots, \phi_n$ are translations of the premises of the English argument and ψ is a translation of its conclusion. Suppose finally that we have a proof that

(16) $$\phi_1, \ldots, \phi_n \vDash \psi.$$

is a valid sequent. Then via the translation from English to L our proof shows that if the premises of the argument are true, its conclusion must also be true. In short we have shown that the English argument is valid.

The following example appears as an exercise in Richard Whately's logic textbook of 1826:

(17) A negro is a man; therefore he who murders a negro murders a man.

This argument seems to defy the logical tools available in 1826; in fact some years later Augustus De Morgan challenged the logicians of his age to develop a logic that does recognize such arguments. There is no record of how Whately expected his students to solve (17), but at least for first-order logic it is straightforward. The main step is to introduce symbols whose interpretations are predicates with two variables:

(18)

domain	the class of living beings (for example)
$N(x)$	x is a negro
$M(x)$	x is a man
$R(x, y)$	x murders y

(We need the variables on the left side of (18) to distinguish between "x murders y" and "y murders x.") Again we insist that there is a translation into classes. For example we can paraphrase "x murders y" by

(19) The pair of objects (x,y) lies in the class of all ordered pairs of objects such that the first murders the second.

Classes of this kind, whose members are ordered pairs, are called *relations*. Now Whately's argument translates into first-order sentences as follows:

(20) $\forall x(N(x) \rightarrow M(x)) \vDash \forall x(\exists y(N(y) \wedge R(x,y)) \rightarrow \exists y(M(y) \wedge R(x,y)))$.

Here the premise has a form that we have already considered. We can read "$\exists y$" as "There is something, call it y, such that," and we can read " $\wedge$ " as "and." So the conclusion says: For every living being x, if there is a negro y such that x murders y, then there is a man y such that x murders y. Now one can show that under every interpretation if the premise of (20) is true then the conclusion is true. So Whately's argument is valid.

If we fail to find a translation of an English argument into a valid first-order sequent, this does not prove that the original argument was invalid. It could be that the argument is valid but a more powerful language than first-order is needed to show this. It could be that there is a suitable first-order sequent but we simply failed to find it.

PROPOSITIONAL LOGIC

Before we define the languages of first-order logic, we should examine a simpler logic called *propositional logic* or *sentential logic*. It uses the symbols $\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$ but not $\forall$ or $\exists$ or = .

In "classical" propositional logic we consider meaningful sentences that are either true or false. We say that their *truth value* is Truth (T for short) if they are true, and Falsehood (F for short) if they are false. There are also "many-valued" propositional logics that allow three or more truth values; we shall not consider these.

If ϕ is any sentence, we can form a new sentence $\neg\phi$ by writing the negation sign "$\neg$" immediately in front of ϕ. We stipulate that $\neg\phi$ is true if ϕ is false, and false if ϕ is true. For example

(21) $\neg$ Today is Tuesday.

expresses

(22) It is not true that today is Tuesday.

We read the symbol "$\neg$" as "not," and $\neg\phi$ is called the *negation* of ϕ. Likewise if ϕ and ψ are sentences, we can form a sentence $(\phi \wedge \psi)$, and we stipulate that $(\phi \wedge \psi)$ is true if and only if both ϕ and ψ are true. For example we can form

(23) (Today is Tuesday $\wedge$ The paper is not yet finished),

and this sentence expresses

(24) Today is Tuesday and the paper is not yet finished.

The sentence $(\phi \wedge \psi)$ is called the *conjunction* of ϕ and ψ, and the sentences ϕ and ψ are its *conjuncts*. We read the symbol " $\wedge$ " as 'and.'

The remaining logical symbols of propositional logic have similar explanations. They all form new sentences from old ones, and in each case we stipulate the truth value of the new sentence in terms of the truth values of the old ones. The following table records these stipulations:

(25)

	ϕ	ψ	$\neg\phi$	$(\phi\wedge\psi)$	$(\phi\vee\psi)$	$(\phi\rightarrow\psi)$	$(\phi\leftrightarrow\psi)$
(*i*)	T	T	F	T	T	T	T
(*ii*)	T	F		F	T	F	F
(*iii*)	F	T	T	F	T	T	F
(*iv*)	F	F		F	F	T	T

We read the table as follows. Suppose for example that ϕ is the sentence "Today is Tuesday" and ψ is the sentence "The paper is not yet finished." If ϕ is true and ψ is false, then we are in row (ii) of the table. In this row there is F below $(\phi \wedge \psi)$, and this tells us that the sentence

(26)(Today is Tuesday $\wedge$ The paper is not yet finished)

is false. Likewise for the other rows and formulas.

We call $(\phi \vee \psi)$ the *disjunction* of ϕ and ψ; the sentences ϕ and ψ are its *disjuncts*. The symbol '$\vee$' can be read as 'or.' But notice that $(\phi \vee \psi)$ is true when both ϕ and ψ are true; so in some circumstances a safer reading of $(\phi \vee \psi)$ is 'Either ϕ or ψ, or both.'

The symbol " $\leftrightarrow$ " can be read safely as "if and only if."

There remains the symbol " $\rightarrow$," sometimes known as *material implication*. The one case where $(\phi \rightarrow \psi)$ is false is where ϕ is true and ψ is false, and this suggests reading $(\phi \rightarrow \psi)$ as "If ϕ then ψ." In mathematical contexts this reading generally works well. But note that we also have T in the bottom two rows of the table below $(\phi \rightarrow \psi)$, so that this sentence counts as true whenever ϕ is false, whether or not there is any connection between ϕ and ψ. For example the following sentence is true on any day of the week:

(27) (Three plus three is two $\rightarrow$ Today is Tuesday)

Likewise $(\phi \rightarrow \psi)$ is true whenever ψ is true, and so the following sentence is also true on any day of the week:

(28) (Today is Tuesday → Three plus three is six)

These properties of " → " are sometimes referred to as the *paradoxes of material implication*—though really they are not so much paradoxes as puzzles about how to translate " → " into English.

The symbols "¬," " ∧ ," " ∨ ," " → ," and " ↔ " are known as the *propositional operators*. We can build up complex sentences by using the propositional operators any number of times. But first it is helpful to introduce so-called *propositional symbols*

(29) $p, q, r, p_0, p_1, p_2, \ldots.$

which can stand as abbreviations of any sentence. Thus we can form sentences

(30) $\neg p, ((p \rightarrow q) \wedge r), (p \vee \neg q), \neg\neg p_2$

and so on. The *sentences of propositional logic* are the propositional symbols and all the complex sentences that can be built up from them using the propositional operators. The table (25) tells us when each of these sentences is true, as soon as we know what truth values to assign to the propositional symbols in them. Take $((p \rightarrow q) \wedge r)$, for example. It uses three propositional symbols. Each of these three symbols could stand for a true sentence or a false one, and so there are eight possible cases that we can list as follows.

(31)

p q r
T T T
T T F
T F T
T F F
F T T
F T F
F F T
F F F

For each row we can evaluate the truth value of $((p \rightarrow q) \wedge r)$ by starting at the propositional symbols and working upwards to more complex sentences, reading values from the table (25). Thus:

(32)

p q r	((*p*	→	*q*	∧	*r*)
T T T	T	T	T	T	T
T T F	T	T	T	F	F
T F T	T	F	F	F	T
T F F	T	F	F	F	F
F T T	F	T	T	T	T
F T F	F	T	T	F	F
F F T	F	T	F	T	T
F F F	T	T	F	F	F
	(i)	(iv)	(ii)	(v)	(iii)

On the right the columns below the propositional symbols copy the truth values from the left side of the table. The column below a propositional operator gives the truth values of the sentence formed by introducing this operator; for example the table below " → " gives the truth values of $(p \rightarrow q)$. The numbers at the bottom of the table show a possible order for working out the columns. The final column calculated, number (v), gives the truth value of the whole sentence in each of the eight cases listed on the left. The table

(33)

p q r	((*p* → *q*) ∧ *r*)
T T T	T
T T F	F
T F T	F
T F F	F
F T T	T
F T F	F
F F T	T
F F F	F

is called the *truth table* of the sentence $((p \to q) \wedge r)$.

We say that a sentence of propositional logic is a *tautology* if its truth table has T in every row, and a *contradiction* if its truth table has F in every row.

Suppose ϕ is a tautology and suppose also that we replace each propositional symbol in ϕ by an English sentence (the same English sentence for each occurrence of the same propositional symbol), creating a sentence *S*. Then *S* must be true since the truth values of the English sentences will indicate a particular row of the truth table, and we know that the value in this row must be T since ϕ has T in every row. Generally *S* will be a mixture of English and propositional operators. But we can translate *S* into a sentence S' of English, for example translating the propositional operators as suggested above but with due caution. Since S' is a translation of *S*, it has the same truth value, and we saw that this value has to be Truth. In short S' will be a necessary truth in English.

Here are some tautologies.

(34) $((p \wedge q) \to p)$

$((p \wedge q) \to q)$

$(p \to (p \vee q))$

$(q \to (p \vee q))$

$(p \to (q \to p))$

$((p \to q) \to ((q \to r) \to (p \to r)))$.

$(((p \to q) \wedge p) \to q)$

$((p \leftrightarrow q) \to (p \to q))$

$(\neg\neg p \to p)$

$((p \wedge \neg p) \to q)$

$(((p \to q) \to p) \to p)$

A possible translation of the first of these tautologies into English is "If the light is broken and the switch is on, the light is broken." This sentence has to be true in any situation in which each of the sentences "The light is broken" and "The switch is on" has a truth value, regardless of what these truth values are.

Suppose ϕ, ψ and χ are sentences of propositional logic. As above, the expression

(35) $$\phi, \psi \vDash \chi$$

is called a (semantic) *sequent*. We say that it is *valid* if in every case where ϕ and ψ are true, χ is also true. It's easy to calculate from the truth tables of ϕ, ψ and χ whether or not the sequent (35) is valid. As with tautologies we can translate the sentences ϕ, ψ and χ simultaneously into sentences of English, by choosing sentences to replace the propositional symbols and then paraphrasing to remove the propositional operators. The result is an English argument, if we write "Therefore" in place of $\vDash$. Suppose we have a proof that the sequent (35) is valid (for example, by truth tables). Then this proof shows that if the premises of the English argument are true, its conclusion must be true too. In this way we justify the English argument.

For example we can justify the English argument

(36) If sending abusive e-mails is an offense, then Smith has just committed an offense. Sending abusive e-mails is an offense. Therefore: Smith has just committed an offence.

by proving that the following sequent is valid:

(37) $$(p \to q), p \vDash q$$

In fact it is valid, as truth tables quickly show. Since this sequent corresponds to indefinitely many other English arguments too, we should think of it as a rule of argument rather than an argument in itself. Logicians sometimes express this point by saying that in logic we study forms of argument rather than individual arguments.

TRANSLATING BETWEEN ENGLISH AND FIRST-ORDER LOGIC

Translations from first-order logic to English are generally straightforward; the problem is to make the English version digestible. But for assessing English arguments we need translation in the other direction, and this can be hazardous.

NOUN PHRASES. Proper names with singular meaning can go over into constants. For example the interpretation

(38)

a	the Pyrenees range
b	France
c	Spain
$R(x_1, x_2, x_3)$	x_1 is between x_2 and x_3

(with any suitable domain supplied) allows us to make the translation

(39) The Pyrenees lie between France and Spain.

$$R(a, b, c)$$

Complex singular noun phrases such as "my father" are in general more complicated to translate. In first-order logic we are allowed to use function symbols, as *F* in the interpretation and translation

(40)

a	Lloyd George
b	me
$F(x_1)$	the father of x_1
$R(x_1, x_2)$	x_1 knew x_2

(41) Lloyd George knew my father.

$$R(a, F(b))$$

But there is a catch. The requirement that first-order languages should be "logically perfect" implies that if "*a*" names any element of the domain of an interpretation, the expression *F*(*a*) should also name an element of the domain. So the domain to be supplied for (40) above must contain not only me but my father, my father's father, my father's father's father, and so on. Worse still, to adapt an example from Frege, if for any reason the domain contains the moon, it must also contain the father of the moon!

For such reasons, one hardly ever meets function symbols in first-order logic outside mathematical contexts. Even there caution is needed. For example in studying number fields one would like to have a "multiplicative inverse" function taking 2 to 1/2, 3 to 1/3, and so on; but 1/0 is undefined.

A common solution is to use, instead of a function symbol, a relation symbol with one more argument place:

(42)

a	Lloyd George
b	me
$P(x_1, x_2)$	x_1 the father of x_2
$R(x_1, x_2)$	x_1 knew x_2

(43) Lloyd George knew my father.

$$\exists x(P(x,b) \wedge R(a,x))$$

This raises a further problem: If the implication that I have exactly one father plays any role in an argument using this sentence, then our translation by (42) fails to convey this implication. Here we need to call in Bertrand Russell's analysis of definite descriptions.

According to Russell's analysis, a sentence of the form "The X is a Y" paraphrases as

(44) At least one thing is an X, at most one thing is an X, and everything that is an X is a Y.

We can translate this directly into first-order symbols, but the following paraphrase is neater:

(45) $\exists z(\forall x\ (x \text{ is an X} \leftrightarrow x = z) \wedge z \text{ is a Y})$.

A major problem with Russell's analysis is that it assumes we can choose the domain of the interpretation in such a way that it contains a unique *x*. But there are other requirements on the domain; all the quantifiers in the first-order sentence range over it. These requirements may clash. In everyday English we use the phrase "the X" in situations in which there is one "salient" X (see for example David Lewis), and to do justice to this in a first-order translation we need to make explicit what makes a certain individual "salient."

NOUN PHRASES THAT CONTAIN QUANTIFIER WORDS LIKE "EVERY." We can handle some cases by paraphrasing:

(46)

P	the class of prime numbers greater than two
Q	the class of odd numbers

(47) Every prime number greater than two is odd.

Every object, if it is a prime number greater than two, is an odd number.

$$\forall x(P(x) \rightarrow Q(x))$$

(48) Some prime numbers greater than two are odd.

$$\exists x(P(x) \wedge Q(x)).$$

If we wanted "some" to imply "more than one" in this example, we would need a longer paraphrase using " = ":

(49) $\exists x \exists y(P(x) \wedge Q(x) \wedge P(y) \wedge Q(y) \wedge x \neq y)$

("$x \neq y$" is a standard abbreviation for "$\neg(x = y)$.") Likewise we can express that there are at least three odd numbers:

(50) $\exists x_1 \exists x_2 \exists x_3 \, (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3 \wedge Q(x_1) \wedge Q(x_2) \wedge Q(x_3))$

Using " = " in analogous ways, first-order logic is equipped to express things like "There are exactly ten million Xs."

CONDITIONALS. The nub of the paradoxes of material implication may be that in real life as opposed to mathematics, sentences play many roles besides being true or false. Paul Grice argued that when we make the appropriate distinction between "what is said" and "what is implicated," there remains no difference in meaning between 'If ϕ then ψ' and $(\phi \rightarrow \psi)$. On the other side, Dorothy Edgington pointed out that

(51) $((p \rightarrow q) \vee (\neg p \rightarrow q))$

is a tautology, and drew the following consequence of reading " → " as "If … then":

(52) … if I reject the conditional "If the Conservatives lose, Thatcher [the leader of the Conservative party] will resign," I am committed to accepting "If the Conservatives win, Thatcher will resign"!

She found this consequence implausible.

ADVERBS. Consider the argument

(53) Nadia accidentally poisoned her father. Therefore: Nadia poisoned her father.

The obvious first-order translations of "Nadia poisoned her father" don't allow us to add further information like "accidentally" or "last Wednesday" or "with strychnine." Peirce in 1892 suggested a way around this problem, namely to talk explicitly about actions. Thus:

(54)

n	Nadia
f	Nadia's father
$A(x_1, x_2)$	x_1 is an action that was performed by x_2
$P(x_1, x_2)$	x_1 is an action of poisoning x_2
$U(x_1)$	x_1 is accidental

(55) $\exists x \, (A(x,n) \wedge P(x,f) \wedge U(x)) \vDash \exists x \, (A(x,n) \wedge P(x,f))$

Donald Davidson and others have independently revived Peirce's suggestion in connection with the semantics of natural languages.

Peirce comments that his translation consists in "catching one of the transient elements of thought upon the wing and converting it into one of the resting places of the mind." This is more than idle whimsy. Peirce's point is that in order to formalize arguments like (53), we sometimes need to introduce abstract objects—in his case, actions—into the domain.

A different kind of example to illustrate Peirce's point is the sentence

(56) For every drop of rain that falls a flower grows.

(George Boolos.) Taken literally, this statement implies that there at least as many flowers as raindrops. If we wanted to make this explicit in order to draw out consequences in an argument, we would need to incorporate some set-theoretic apparatus for talking about cardinalities.

Arguments about past, present and future are another example of the same general point. Since sentences of first-order logic lack tense, the normal way to handle such arguments in first-order logic is to add points of time (or sometimes intervals of time) to the domain. Then in general we need to add to the premises of an argument some basic facts about time, for example that the ordering of time into earlier and later is a linear ordering. (One can use the axioms for linear ordering, (76) below.)

NON-INDICATIVE SENTENCES. Sentences of first-order logic are all in the indicative. They are not designed for giving instructions or asking questions. One place where this matters is the formalization of mathematical reasoning. Mathematicians often use imperatives:

(57) "Draw a triangle ABC and consider the midpoint of the side AB."

"Assume there is a greatest prime."

"Let x be a number between 0 and 5."

First-order logic has no straightforward way of expressing these instructions. In 1926 Jan Łukasiewicz suggested we should regard such instructions as moves in a proof. For example the instruction "Assume ϕ" is an indication that we are going to prove the sequent (84) below by proving the sequent (83).

Most English sentences can be translated into first-order sentences in many different ways. The translation that we choose should be guided by the arguments that we are trying to formalize. Some philosophers have speculated that for each unambiguous English sentence *S*

there is a first-order translation ϕ that expresses the correct analysis of S into its most primitive components. If we knew these translations we could use them to formalize in first-order sentences any valid argument in English. A difficulty with this thesis is that (as we saw) analyses of English sentences for the purpose of justifying arguments may need to bring in whole ontologies of abstract objects: Sets, actions, points of time and space. Some philosophers would add possible worlds.

Most of the starred textbooks in the bibliography below give further advice about translating from English into first-order sentences.

FIRST-ORDER SYNTAX

The signature of a first-order language consists of symbols of four kinds, though not every first-order language has all four kinds. The kinds are as follows:

(i) *Propositional symbols* as in section "Propositional Logic."

(ii) *Relation symbols*, usually

$$P, Q, R, R_0, R_1, R_2, \ldots.$$

(iii) *Individual constant symbols*, or more briefly *constants*, usually

$$a, b, c, c_0, c_1, c_2, \ldots.$$

(iv) *Function symbols*, usually symbols such as

$$F, G, H, F_0, F_1, F_2, \ldots.$$

Each relation symbol and each function symbol has an *arity*, which is a positive integer. One normally requires that no symbol occurs in more than one of these kinds, and that no relation or function symbol occurs with more than one arity. If a function or relation symbol has arity n, we describe the symbol as *n-ary*. *Binary* means 2-ary.

A first-order language also has an infinite set of symbols called *variables*. Variables are usually chosen as lower case letters near the end of the alphabet:

(58) $$u, v, w, x, y, z, v_0, v_1, \ldots.$$

The variables are not in the signature.

Given any signature σ, we define a first-order language $L(\sigma)$ in terms of σ. We begin with the terms of $L(\sigma)$.

(a) Every constant of σ is a term of $L(\sigma)$.

(b) Every variable of L is a term of $L(\sigma)$.

(c) Suppose F is a function symbol of σ, n is the arity of F, and $t_1, \ldots, t_n$ are terms of $L(\sigma)$. Then the expression

$$F(t_1, \ldots, t_n)$$

is a term of $L(\sigma)$.

(d) $L(\sigma)$ has no terms except as given by (a)–(c).

This definition is an *inductive definition*. Clauses (a) and (b) together form its *base clause*; they say outright that certain expressions are terms. Clause (c) is the *inductive clause*; it says that if certain expressions are terms then certain other expressions are terms too. Clause (d) tells us that if t is a term of $L(\sigma)$ then t can be generated in a finite number of steps by using the base and inductive clauses.

A *metatheorem* of first-order logic is a theorem about first-order logic (as opposed to a theorem proved by means of first-order logic).

METATHEOREM 1 (UNIQUE PARSING OF TERMS). *If t is a term of L(σ) then exactly one of the following holds:*

(1) t is a constant of σ.

(2) t is a variable.

(3) t is $F(t_1, \ldots, t_n)$ *where F is a function symbol of σ, n is the arity of F and* $t_1, \ldots, t_n$ *are terms of L(σ).*

Moreover in case (3) if t is also $G(s_1, \ldots, s_m)$ *where G is a function symbol of σ and* $s_1, \ldots, s_m$ *are terms of L(σ), then F is G and n is m and* t_1 *is* s_1 *and … and* t_n *is* s_n.

See Stephen Kleene §17 for the proof. Thanks to unique parsing, we can distinguish three types of term. The first two types are the constants and the variables, and together they form the *atomic* terms of L. The third type of term consists of those of the form $F(t_1, \ldots, t_n)$; these are said to be *compound* terms. Broadly speaking the terms of $L(\sigma)$ correspond grammatically to the singular definite noun phrases of a natural language.

The unique parsing lemma is used to justify certain types of definition and proof by induction. For example we can define, for each term t, the set $V(t)$ of variables that occur in t, as follows:

(a) If t is a constant then $V(t)$ is $\emptyset$ (the empty set).

(b) If t is a variable then $V(t)$ is the set $\{t\}$.

(c) If t is $F(t_1, \ldots, t_n)$ then $V(t)$ is the union

$$V(t_1) \cup \ldots \cup V(t_n).$$

(This is the set of objects that are in at least one of $V(t_1), \ldots, V(t_n)$.) We say that a term is *closed* if it contains no variables.

Along similar lines we can define $t(s_1, \ldots, x_n/v_1, \ldots, v_n)$, which is the term got by taking the term t and putting the term s_1 in place of each occurrence of the variable v_1 in t, s_2 in place of each occurrence of v_2 and so on; the replacements should be made simultaneously. For example if t is $F(x,G(y))$ and s is $G(z)$, then $t(s/x)$ is $F(G(z),G(y))$ and $t(y,t/x,y)$ is $F(y,G(F(x,G(y))))$.

Having defined the terms of $L(\sigma)$, we define the formulas of $L(\sigma)$ as follows.

(a) Every propositional symbol of σ is a formula of $L(\sigma)$.

(b) If R is a relation symbol of σ, n is the arity of R and $t_1, \ldots, t_n$ are terms of $L(\sigma)$, then the expression

$$R(t_1, \ldots, t_n)$$

is a formula of $L(\sigma)$.

(c) If s and t are terms of $L(\sigma)$ then the expression

$$(s = t)$$

is a formula of $L(\sigma)$.

(d) If ϕ is a formula of $L(\sigma)$ then the expression

$$\neg\phi$$

is a formula of $L(\sigma)$.

(e) If ϕ and ψ are formulas of $L(\sigma)$ then the four expressions

$$(\phi \land \psi), (\phi \lor \psi), (\phi \to \psi), (\phi \leftrightarrow \psi)$$

are formulas of $L(\sigma)$.

(f) If ϕ is a formula of $L(\sigma)$ and v is a variable of $L(\sigma)$, then the two expressions

$$\forall v\phi, \exists v\phi$$

are formulas of $L(\sigma)$. ($\forall v$ is called a *universal quantifier* and $\exists v$ is an *existential quantifier*.)

(g) Nothing is a formula of $L(\sigma)$ except as given by clauses (a)–(f) above.

The obvious counterpart to Metatheorem 1 is true for formulas. It allows us to say that a formula is *atomic* if it comes from clauses (a)–(c) and *compound* if has one of the forms described in clauses (d)–(f). Also no expression of $L(\sigma)$ is both a term and a formula.

The next definition will get its full motivation when we come to discuss satisfaction of formulas. Roughly speaking, a variable x can serve to name an object, unless it appears in one of the contexts "For all objects x, … x …" and "There is an object x such that … x …." (Here we recall that in first-order logic, 'for all objects x' is written $\forall x$ and "there is an object x such that" is written $\exists x$.) An occurrence of a variable in one of these contexts is said to be *bound*; an occurrence that is not bound is *free*. We say a variable is *free in* ϕ if it has a free occurrence in ϕ. A definition by induction of the set $FV(\phi)$ of variables free in the formula ϕ runs as follows:

(a) If ϕ is atomic then $FV(\phi)$ is the set of all variables that occur in ϕ.

(b) $FV(\neg\phi)$ is $FV(\phi)$.

(c) $FV((\phi \land \psi))$, $FV((\phi \lor \psi))$, $FV((\phi \to \psi))$ and $FV((\phi \leftrightarrow \psi))$ are all equal to

$$FV(\phi) \cup FV(\psi).$$

(d) If ϕ is a formula and v is a variable, then $FV(\forall v\phi)$ and $FV(\exists v\, \phi)$ are both the set

$$FV(\phi) \setminus \{v\}$$

of all the variables that are in $FV(\phi)$ and are not v.

A formula ϕ is said to be a *sentence* if $FV(\phi)$ is empty, in other words, if no variable is free in ϕ. For example $\forall x\, (P(x) \to Q(x))$ is a sentence, but $(P(x) \to Q(x))$ is not a sentence since x has two free occurrences in it.

Unique parsing also allows us to define by induction the *complexity* of a formula ϕ, comp(ϕ), as follows:

(1) If ϕ is an atomic formula then $\text{comp}(\phi) = 0$.

(2) For every formula ϕ, $\text{comp}(\neg\phi) = \text{comp}(\phi) + 1$.

(3) If ϕ and ψ are formulas and n is the maximum of comp(ϕ) and comp(ψ), then $\text{comp}((\phi \land \psi))$, $\text{comp}((\phi \lor \psi))$, $\text{comp}((\phi \to \psi))$ and $\text{comp}((\phi \leftrightarrow \psi))$ are all equal to $n + 1$.

(4) If ϕ is a formula and v is a variable then $\text{comp}(\forall v\phi) = \text{comp}(\exists v\phi) = \text{comp}(\phi) + 1$.

There is a similar definition for the complexity of terms. The chief use of complexity is in *proofs by induction on complexity*, which run as follows. We want to show that all formulas of a first-order language have a certain property P. So we show first that all atomic formulas have the prop-

erty P, and then we show that for every positive integer n, if all formulas of complexity $<n$ have P then every formula of complexity n has P too.

One speaks of the *subformulas* of a formula ϕ in two senses. First a subformula of ϕ is a segment of ϕ that is a formula in its own right. Second a subformula of ϕ is a formula that occurs as a subformula of ϕ in the first sense. For example the formula $(P(x) \leftrightarrow P(x))$ has two subformulas of the form $P(x)$ in the first sense, but only one in the second sense. It is easy to give a formal definition of the set of subformulas of ϕ in the second sense, by induction on the complexity of ϕ. Metatheorem 12 below uses subformulas in the first sense.

There are several useful conventions about how to write first-order formulas. For example

(59) $$\phi \wedge \psi \wedge \chi$$

is strictly not a first-order formula, but we count it as an abbreviation of

(60) $$((\phi \wedge \psi) \wedge \chi).$$

In the same spirit the conjunction

(61) $$(\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_n)$$

is an abbreviation for

(62) $$(\ldots(\phi_1 \wedge \phi_2) \wedge \ldots \wedge \phi_n),$$

and the disjunction

(63) $$(\phi_1 \vee \phi_2 \vee \ldots \vee \phi_n)$$

is an abbreviation for

(64) $$(\ldots(\phi_1 \vee \phi_2) \vee \ldots \vee \phi_n).$$

We count " $\wedge$ " and " $\vee$ " as binding tighter than " $\rightarrow$ " or " $\leftrightarrow$," in the sense that

(65) $$(\phi \wedge \psi \leftrightarrow \chi), (\phi \rightarrow \psi \vee \chi)$$

are respectively abbreviations for

$$((\phi \wedge \psi) \leftrightarrow \chi), (\phi \rightarrow (\psi \vee \chi)).$$

Other useful abbreviations are

(66) $$(x \neq y) \text{ for } \neg(x = y),$$

$$\forall x_1 \ldots x_n\, \phi \text{ for } \forall x_1 \ldots \forall x_n\, \phi,$$

$$\exists x_1 \ldots x_n\, \phi \text{ for } \exists x_1 \ldots \exists x_n\, \phi.$$

Also we allow ourselves to leave out a pair of brackets when they stand at the opposite ends of a formula.

Another useful convention is based on mathematical notation for functions. If ϕ is a formula and all the variables free in ϕ are included in the list $v_1, \ldots, v_n$, we introduce ϕ as $\phi(v_1, \ldots, v_n)$. Then if $t_1, \ldots, t_n$ are terms, we write

(67) $$\phi(t_1, \ldots, t_n)$$

for $\phi(t_1, \ldots, t_n/v_1, \ldots, v_n)$, which is the result of putting t_i in place of each free occurrence of v_i in ϕ, simultaneously for all i from 1 to n. (We shall revise this definition later.)

INTERPRETATIONS OF FIRST-ORDER LANGUAGES

A first-order language is a language $L(\sigma)$ for some signature σ. The signature σ determines the language $L(\sigma)$, but equally if we know the formulas of $L(\sigma)$ we can recover σ. So if L is the language $L(\sigma)$, we could equally well say "formula of σ" or "formula of L." Likewise σ-structures, which we are about to define, can equally well be called L-structures.

We recall some set theory. If X is a set and n a positive integer, then an *n-tuple* from X is an ordered list $(a_1, \ldots, a_n)$ where $a_1, \ldots, a_n$ are members of X. We write X^n for the set of all n-tuples from X. An n-ary *relation* on X is a subset of X^n. An n-ary *function* on X is a function f: $X^n \rightarrow Y$, for some set Y, which assigns to each n-tuple $(a_1, \ldots, a_n)$ from X an element $f(a_1, \ldots, a_n)$ of Y.

Suppose σ is a signature. A σ-*structure* is a set-theoretic interpretation for the symbols in σ. More precisely a σ-structure A has the following ingredients:

(a) A set (usually required to be nonempty) which is the domain of a, in symbols dom (A).

(b) For each propositional symbol p of σ, a truth value (T or F) which we write as p_A.

(c) For each constant c of σ, an element c_A of dom (A).

(d) For each relation symbol R of σ, an n-ary relation R_A on dom $(A)^n$, where n is the arity of R.

(e) For each function symbol F of σ, an n-ary function F_A: dom$(A)^n \rightarrow$ dom (A), where n is the arity of F.

EXAMPLE 1 (ARITHMETIC AND ITS LANGUAGE). For talking about the natural numbers

(68) $$0, 1, 2, \ldots,$$

we use a first-order language called the *language of arithmetic*. Its signature $\sigma_{\mathbb{N}}$ consists of one constant symbol $\underline{0}$, two function symbols + and · of arity 2, a function symbol S of arity 1 and a relation symbol < of arity 2. The number structure $\mathbb{N}$ is the following $\sigma_{\mathbb{N}}$-structure. The domain is the set of natural numbers. The constant $\underline{0}$ stands for the number zero (i.e. , $\underline{0}_{\mathbb{N}} = \underline{0}$). The function symbols + and · stand for addition and multiplication of natural numbers, and S stands for the function "plus one." The binary relation symbol < stands for the relation "less than" (i.e. , $<_{\mathbb{N}}$ is the set of all ordered pairs of natural numbers (m, n) with $m < n$). Following normal mathematical usage we write $+ (x, y)$, $\cdot (x, y)$ and $< (x, y)$ as $x + y$, $x \cdot y$ and $x < y$ respectively.

The structure $\mathbb{N}$ interprets the closed terms of $L(\sigma_{\mathbb{N}})$ as names of numbers. For example the term $S(\underline{0})$ stands for the number 1, the term $S(S(\underline{0}))$ stands for the number 2, and so on; we write these terms as $\underline{0}$, $\underline{1}$, $\underline{2}$ and so on. Likewise the closed term $\underline{2} + \underline{3}$ names the number 5.

We can use our earlier explanations of the first-order symbols in order to read any sentence of $L(\sigma_{\mathbb{N}})$ as making a statement about $\mathbb{N}$. The following sentences are all true in $\mathbb{N}$:

(69) PA1. $\forall x\, (Sx \neq \underline{0})$.

PA2. $\forall x \forall y\, (S(x) = S(y) \rightarrow x = y)$.

PA3. $\forall x\, (x + \underline{0} = x)$.

PA4. $\forall x \forall y\, (x + S(y) = S(x + y))$.

PA5. $\forall x\, (x \cdot \underline{0} = \underline{0})$.

PA6. $\forall x \forall y\, (x \cdot S(y) = (x \cdot y) + x)$.

PA7. $\forall x\, \neg(x < \underline{0})$.

PA8. $\forall x \forall y\, (x < Sy \leftrightarrow x < y \vee x = y)$.

The following *induction axiom* is also true in $\mathbb{N}$, though it is not a first-order sentence because the variable 'X' ranges over sets rather than numbers.

(70) For every set X of numbers,
$((0 \in X) \wedge \forall x\, (x \in X \rightarrow S(x) \in X \rightarrow \forall x\, (x \in X))$.

Within $L(\sigma_{\mathbb{N}})$ the closest we can come to this axiom is to give a separate axiom for each set X defined by a first-order formula. Namely let $\phi(x, y_1, \ldots, y_n)$ be any formula of $L(\sigma_{\mathbb{N}})$. Then we write the sentence

(PA.9)

$$\forall y_1 \ldots \forall y_n\, (\phi(\underline{0}, y_1, \ldots, y_n) \wedge \forall x\, (\phi(x, y_1, \ldots, y_n) \rightarrow \phi(S(x), y_1, \ldots, y_n) \rightarrow \forall x\, \phi(x, y_1, \ldots, y_n)).$$

The sentences of the form PA9 constitute the *first-order induction schema for arithmetic*. The infinite set of sentences PA1–PA9 is called *first-order Peano arithmetic*, or PA for short.

The situation with $\mathbb{N}$ is typical. Given any signature σ, any σ-structure A and any sentence ϕ of $L(\sigma)$, we can read ϕ as making a statement about A. If this statement is true we say that A is a *model* of ϕ, and we express this fact by writing

(71) $A \vDash \phi$.

If ϕ is false in a we write $A \nvDash \phi$. It is an unfortunate fact of history that we use the symbol "$\vDash$" both in (71) (which is not a sequent) and in semantic sequents such as (16) above and (72) below. One can avoid confusion by noting that in (71) there is a structure to the left of "$\vDash$," whereas in sequents the space to the left of "$\vDash$" is empty or contains sentences.

THEORIES AND THEIR MODELS

Let L be a first-order language. A set of sentences of L is called a *theory*. An L-structure A is called a *model* of T if it is a model of every sentence in T. The semantic sequent

(72) $T \vDash \psi$

states that every model of T is a model of ψ; when this holds, we say the sequent is *valid* and that ψ is a *(logical) consequence* of T. When T is a finite set, say $\{\phi_1, \ldots, \phi_n\}$, we write the sequent (72) as

(73) $\phi_1, \ldots, \phi_n \vDash \psi$.

When T is empty, we also write the sequent as

(74) $\vDash \psi$.

(74) says that ψ is true in every L-structure; when it holds, we say that the sentence ψ is *valid*, and that it is a *theorem*. Finally the sequent

(75) $T \vDash$

expresses that T has no models.

We say that T is *consistent* if it has models, in other words if $T \nvDash$. We say that T is *complete* if for every sentence ϕ of L, at least one of ϕ and $\neg\phi$ is a consequence of

T. So T is consistent and complete if and only if for every sentence ϕ of L, exactly one of ϕ and $\neg\phi$ is a consequence of T.

Since a sentence of $L(\sigma)$ is also a sentence of $L(\tau)$ whenever τ includes σ, we should check that the validity of a sequent depends only on the sentences in it, and not on the signature—otherwise our notation for sequents would need to mention the signature. The following metatheorem assures this. (It requires that every structure has nonempty domain. In a structure with empty domain, $\exists x\ (x = x)$ is false; but adding a constant to the signature automatically excludes structures with empty domains and hence makes $\exists x\ (x = x)$ a theorem.)

METATHEOREM 2. *If* T *and* ψ *are a theory and a sentence of* L(σ)*, and* τ *is a signature extending* σ*, then (72) is valid for* σ*-structures if and only if it is valid for* τ*-structures.*

A theory T is said to be *deductively closed* if T contains all its consequences. If S is any theory then the set T of all consequences of S is deductively closed and contains S; we call T the *deductive closure* of S. When T is the deductive closure of S we say also that S is a *set of axioms* for T.

First-order theories commonly arise in one of two ways.

In the first way we have a first-order language L and an L-structure A, and we want to list the facts that we know about A. By the *complete theory* of A we mean the set T of all sentences of L that have A as a model. This set T is always deductively closed, consistent and complete. If we have a set S of sentences that are all true in A, then certainly S is consistent; if it is also complete, then S is a set of axioms for the complete theory of A. An ambition for logicians is to give sets of axioms for the complete theories of various mathematical structures. For many cases this is achieved. But in 1931 Kurt Gödel gave an indirect but astonishingly insightful proof that PA is not complete, so that it doesn't axiomatise the complete first-order theory of $\mathbb{N}$. (See entry on "Gödel's Theorem.")

The second common source of first-order theories is the definitions of classes of mathematical structures. For example a linearly ordered set is a structure in a signature containing a binary relation symbol <, which is a model of the three sentences:

(76)
$$\forall x \forall y \forall z\ (x < y \wedge y < z \rightarrow x < z)$$
$$\forall x \neg\ (x < x)$$
$$\forall x \forall y\ (x < y \vee y < x \vee x = y)$$

The structure $\mathbb{N}$ forms a linear ordering, since all of (76) is true in $\mathbb{N}$. This theory (76) is a direct translation into first-order notation of the usual informal definition of linear orderings.

FORMULAS AND SATISFACTION

The formula $x < \underline{3}$ is neither true nor false in $\mathbb{N}$, because the variable x lacks an interpretation. The interpretations of the symbols of $\sigma_{\mathbb{N}}$ are fixed in $\mathbb{N}$, but the interpretations of the variables are not. The same holds for any first-order language L, any L-structure A and any formula ϕ of L in which some variables are free.

By an *assignment in* the structure A we mean a function α whose domain is a set of variables and which assigns to each variable in its domain an element of A. If t is a term and α is an assignment in A whose domain includes all the variables in t, then A and α together tell us how to treat t as the name of an element of A, and we write this element $t_A[\alpha]$. For example if A is $\mathbb{N}$ and α is an assignment in $\mathbb{N}$ with $\alpha(x) = 4$, and t is the term $x + \underline{5}$, then $t_{\mathbb{N}}[\alpha]$ is 4 + 5, in other words 9.

When t is a closed term, $t_A[\alpha]$ is independent of α and we write it simply as t_A.

Similarly we can use assignments to interpret the free occurrences of variables in a formula. (The bound occurrences need no interpretation; they are part of the apparatus of quantification.) For example in $\mathbb{N}$, any assignment α with $\alpha(x) = 4$ interprets the formula $x < \underline{5}$ as making the statement that 4 is less than 5, which is true. We express the fact that this statement is true by writing

(77)
$$\mathbb{N} \vDash (x < \underline{5})[\alpha].$$

More generally if α is an assignment in A which makes assignments to all the variables free in ϕ, then

(78)
$$A \vDash \phi[\alpha]$$

states that ϕ, interpreted in A by means of α, is true. When (78) holds we say that α *satisfies* ϕ in A. We write

(79)
$$A \nvDash \phi[\alpha]$$

if α fails to satisfy ϕ in A.

There are two reasons for introducing the notion of "satisfying." The first is that it allows us to use formulas with free variables in order to express properties of individual elements or sequences of elements in a structure. For this application another notation is helpful. Suppose $\phi(x_1, \ldots, x_n)$ is a formula of L and α is an assignment that assigns elements to at least the variables $x_1, \ldots, x_n$. Write a_i for $\alpha(x_i)$. Then instead of "$A \vDash \phi[\alpha]$" we also write

(80) $$A \vDash \phi[a_1, \ldots, a_n].$$

We read this statement as "$a_1, \ldots, a_n$ satisfy ϕ in A."

The second reason for introducing satisfaction is that (as Alfred Tarski pointed out) it allows us to give a fully precise mathematical definition of the relation "$A \vDash \phi$," by first defining the relation "$A \vDash \phi[\alpha]$." The first step of the definition is to define $t_A[\alpha]$ by induction on the complexity of t; we omit details. This done, the definition of "$A \vDash \phi[\alpha]$" goes by induction on the complexity of ϕ. We give some typical cases and leave the remainder to the reader.

(a) For every propositional symbol p,

$$A \vDash p \text{ if and only if } p_A \text{ is T.}$$

(b) If R is a relation symbol of arity n and ϕ has the form $R(t_1, \ldots, t_n)$, then

$$A \vDash \phi[\alpha] \text{ if and only if } ((t_1)_A[\alpha], \ldots, (t_n)_A[\alpha]) \text{ is in } R_A.$$

(c) $A \vDash (\phi \wedge \psi)[\alpha]$ if and only if $A \vDash \phi[\alpha]$ and $A \vDash \psi[\alpha]$.

The clauses for quantifiers need some further notation. Suppose α is an assignment whose domain includes all the variables free in ϕ except perhaps v, and a is an element of the structure A. Then we write $\alpha(a/v)$ for the assignment β whose domain is the domain of α together with v, and such that for each variable x in the domain of β,

(81) $$\beta(x) = \begin{cases} \alpha(x) & \text{if } x \text{ is not } v, \\ a & \text{if } x \text{ is } v. \end{cases}$$

(d) $A \vDash \forall v\phi[\alpha]$ if and only if for every element a of dom(A), $A \vDash \phi[\alpha(a/v)]$.

(e) $A \vDash \exists v\phi[\alpha]$ if and only if there is an element a of dom(A) such that $A \vDash \phi[\alpha(a/v)]$.

Tarski's definition of "$A \vDash \phi[\alpha]$" goes by induction on the complexity of formulas, as above. But by standard set-theoretic methods we can convert it to an explicit set-theoretic definition and hence prove the following metatheorem:

METATHEOREM 3. *There is a formula θ of set theory such that*

$$\theta(\sigma, A, \phi, \alpha)$$

is true in the universe of sets if and only if σ is a signature, A *is a σ-structure, ϕ is a formula of $L(\sigma)$ and $A \vDash \phi[\alpha]$.*

We would like to know that $\phi(y/x)$ says the same thing about an object y as $\phi(x)$ says about an object x. More precisely, we would like to know the following:

METATHEOREM 4. *Suppose $\phi(x)$ is a formula of the first-order language L, $t(y)$ is a term of L, α is an assignment whose domain includes y, and β is an assignment with $\beta(x) = t_A[\alpha]$. Then*

$$A \vDash \phi(t/x)[\alpha] \text{ if and only if } A \vDash \phi[\beta].$$

Unfortunately this metatheorem is false unless we make certain adjustments. For example let $\phi(x)$ be the formula $\exists y\,(x \neq y)$ which says that there is something else besides x, and let t be the variable y. Then $\phi(t/x)$ is the everywhere false sentence $\exists y\,(y \neq y)$, not a formula saying that there is something else besides y. The quantifier $\exists y$ has captured the variable y when we substituted t for x in $\phi(t/x)$.

There is a remedy. Given any formula ϕ and any term t, we define $\phi(t/\!/x)$ as follows. For each variable v occurring in the term t, choose another variable v' that doesn't occur in either t or ϕ, taking different variables v' for different v. Form the formula ϕ' by replacing each bound occurrence of each variable v in ϕ by an occurrence of v'. Finally take $\phi(t/\!/x)$ to be $\phi'(t/x)$. (A more precise account would say how we choose the variables v' and would explain the relevance of the logical equivalence (90)(i) below.) Then $\phi(t/\!/x)$ is said to come from ϕ by substituting t for x "without clash of variables." For simplicity of notation we now redefine $\phi(t/x)$ to mean $\phi(t/\!/x)$, thus throwing away the ladder we climbed up. After this redefinition, Metatheorem 4 is true.

Some authors avoid this redefinition by forbidding the use of $\phi(t/x)$ when ϕ contains a quantifier that captures a variable in t.

SOME METATHEOREMS OF FIRST-ORDER LOGIC

The metatheorems in this section are mostly immediate from the definitions. We state them because they have useful applications.

METATHEOREM 5. *If ϕ is a first-order sentence then the sequent $\phi \vDash \phi$ is valid.*

METATHEOREM 6 (MONOTONICITY). *If ψ is a sentence and T a theory in a first-order language, and U is a subset of T such that the sequent $U \vDash \psi$ is valid, then the sequent $T \vDash \psi$ is valid.*

METATHEOREM 7 (CUT). *If T is a theory and ϕ, ψ are sentences, all in a first-order language, and the sequents*

(82) $$T \vDash \phi, T \cup \{\phi\} \vDash \psi$$

are both valid, then the sequent $T \vDash \psi$ is valid. (The sentence ϕ is "cut.")

There are a number of metatheorems expressing properties of particular logical operators. The three below are only a sample.

METATHEOREM 8. *Suppose T is a first-order theory and ϕ is a first-order sentence. Then the sequent $T \vDash \phi$ is valid if and only if the sequent $T \cup \{\neg\phi\} \vDash$ is valid.*

METATHEOREM 9. *Suppose T is a first-order theory and ϕ and ψ are first-order sentences. If*

(83) $$T \cup \{\phi\} \vDash \psi$$

is valid, then

(84) $$T \vDash (\phi \rightarrow \psi)$$

is valid. Also if

(85) $$T \vDash \phi, T \vDash (\phi \rightarrow \psi)$$

are both valid then

(86) $$T \vDash \psi$$

is valid.

The first half of Metatheorem 9 is sometimes called the Deduction Theorem. The second half is one form of a rule traditionally called Modus Ponens.

METATHEOREM 10. *Suppose T is a first-order theory, $\phi(x_1, \dots, x_n)$ is a first-order formula and $c_1, \dots, c_n$ are n distinct constants that occur nowhere in either T or ϕ. Then if either of the following sequents is valid, so is the other:*

$$T \vDash \phi(c_1, \dots, c_n).$$

$$T \vDash \forall x_1, \dots \forall x_n \phi.$$

Our remaining metatheorems describe important properties of first-order logic as a whole.

METATHEOREM 11. *Let L be a first-order language and $\phi(v_1, \dots, v_n)$ and $\psi(v_1, \dots, v_n)$ formulas of L. Then the following are equivalent:*

(a) For every σ-structure A and all elements $a_1, \dots, a_n$ of the domain of A,

$$A \vDash \phi[a_1, \dots, a_n] \text{ if and only if } A \vDash \psi[a_1, \dots, a_n].$$

(b) $\vDash \forall v_1 \dots \forall v_n (\phi(v_1, \dots, v_n) \leftrightarrow \psi(v_1, \dots, v_n))$.

When these conditions (a) and (b) hold, we say that ϕ and ψ are logically equivalent, *and we write this as $\phi \equiv \psi$.*

Logical equivalence is an equivalence relation on formulas. Here are some logically equivalent pairs:

(87)
(a) $(\phi \vee \psi) \quad \equiv \quad \neg(\neg\phi \wedge \neg\psi)$
(b) $\neg(\phi \wedge \psi) \quad \equiv \quad (\neg\phi \vee \neg\psi)$
(c) $\neg(\phi \vee \psi) \quad \equiv \quad (\neg\phi \wedge \neg\psi)$
(d) $(\phi \wedge (\psi \vee \chi)) \quad \equiv \quad ((\phi \wedge \psi) \vee (\phi \wedge \chi))$
(e) $(\phi \rightarrow \psi) \quad \equiv \quad (\neg\phi \vee \psi)$
(f) $(\phi \leftrightarrow \psi) \quad \equiv \quad ((\phi \wedge \psi) \vee (\neg\phi \wedge \neg\psi))$

Equivalences (a), (b) and (c) are examples of a group of equivalences that go by the name of *De Morgan's Laws*.

METATHEOREM 12. *Let L be a first-order language and suppose ϕ and ϕ' are logically equivalent formulas of L. Let ψ be a formula of L, and let ψ' come from ψ by replacing a subformula of ψ of the form ϕ by a copy of ϕ'. Then ψ' is logically equivalent to ψ.*

Metatheorem 12 together with equivalences (a), (e), and (f) tells us that, given any first-order formula ϕ, we can remove all occurrences of the symbol " $\leftrightarrow$ ", and then all occurrences of the symbols " $\vee$ " and " $\rightarrow$," and so find a formula logically equivalent to ϕ in which none of these symbols occurs. So there would be no loss of expressive power if we removed these symbols from the language. By a similar argument we could make do with " $\vee$ " and "$\neg$," discarding " $\wedge$," " $\rightarrow$ " and " $\leftrightarrow$."

Other choices of symbol are open to us. For example we can introduce the symbol "$\perp$" as an atomic formula; since it has no variables it is a sentence, and we stipulate that its truth value is *F* in all structures. This symbol "$\perp$" is logically equivalent to $\neg\forall x(x = x)$, or more generally to $\neg\phi$ where ϕ is any valid sentence. We pronounce "$\perp$" as

"absurdity"; some computer scientists read it as "bottom." Given the logical equivalences

(88)

$$\neg\phi \equiv (\phi \rightarrow \bot)$$
$$(\phi \wedge \psi) \equiv \neg(\phi \rightarrow \neg\psi)$$

we see that in the presence of "$\bot$" and "$\rightarrow$" we can drop "$\neg$," "$\wedge$," "$\vee$," "$\leftrightarrow$" from the language.

A formula with no quantifiers is said to be *quantifier-free*. By a *literal* we mean a formula that is either atomic or the negation of an atomic formula. By a *basic conjunction* we mean either a literal or a conjunction of literals; likewise a *basic disjunction* is a literal or a disjunction of literals. A quantifier-free formula is said to be in *disjunctive normal form* if it is a basic conjunction or a disjunction of basic conjunctions; it is in *conjunctive normal form* if it is a basic disjunction or a conjunction of basic disjunctions.

METATHEOREM 13. *Every quantifier-free formula $\phi(x_1, \ldots, x_n)$ is logically equivalent to a quantifier-free formula $\phi^{dnf}(x_1, \ldots, x_n)$ in disjunctive normal form, and to a quantifier-free formula $\phi^{cnf}(x_1, \ldots, x_n)$ in conjunctive normal form.*

We illustrate Metatheorem 13:

(89)

$$\neg(p \leftrightarrow q) \quad \equiv \text{(by (f))}$$
$$\neg((p \wedge q) \vee (\neg p \wedge \neg q)) \quad \equiv \text{(by (c))}$$
$$\neg(p \wedge q) \wedge \neg(\neg q \wedge \neg q) \quad \equiv \text{(by (b))}$$
$$(\neg p \vee \neg q) \wedge (p \vee q) \quad \equiv \text{(by (d))}$$
$$((\neg p \vee \neg q) \wedge p) \vee ((\neg p \vee \neg q) \wedge q) \quad \equiv \text{(similarly)}$$
$$(\neg p \wedge p) \vee (\neg q \wedge p) \vee (\neg p \wedge q) \vee (\neg q \wedge q) \equiv$$
$$(\neg q \wedge p) \vee (\neg p \wedge q)$$

Here follow some important logical equivalences involving quantifiers.

(90)

(g) $\forall x \phi \equiv \neg \exists x \neg\phi$

(h) $\exists x \phi \equiv \neg \forall x \neg\phi$

(i) $\forall x \phi \equiv \forall y \phi'$ if y doesn't occur in ϕ and ϕ' comes from ϕ by replacing x by y everywhere

(j) $(\forall x \phi \wedge \psi) \equiv \forall x(\phi \wedge \psi)$ if x is not free in ψ

(k) $(\forall x \phi \vee \psi) \equiv \forall x(\phi \vee \psi)$ if x is not free in ψ

A formula is said to be *prenex* if it is quantifier-free or consists of a string of quantifiers followed by a quantifier-free formula. The (possibly empty) string of quantifiers at the front of a prenex formula is called its *quantifier prefix.*

METATHEOREM 14 (PRENEX FORM THEOREM). *Let L be a first-order language and $\phi(x_1, \ldots, x_n)$ a formula of L. Then there is a prenex formula $\psi(x_1, \ldots, x_n)$ of L that is logically equivalent to ϕ.*

To prove the prenex form theorem, one establishes ways of moving a quantifier in a formula "outwards." Equivalences (j) and (k) above are typical examples, and there are corresponding equivalences with $\exists$. If the variable x does occur free in ψ, we first use equivalence (i) to change x to another variable not occurring in ψ.

CONSTRUCTION OF MODELS

One way to show that a theory *T* is consistent is to build a model of *T*. Depending on what *T* is, this can call for a good deal of ingenuity. A number of logicians have studied how, by analysing *T* itself, we can make the process more systematic. The approach described below follows suggestions of Jaakko Hintikka.

Let *L* be a first-order language and *T* a set of sentences of *L*. For simplicity we assume *L* doesn't have $\vee$, $\rightarrow$ or $\leftrightarrow$. We say that *T* is a *Hintikka set* if it has the following properties:

(H1) If $(\phi \wedge \psi)$ is in *T* then both ϕ and ψ are in *T*; if $\neg(\phi \wedge \psi)$ is in *T* then at least one of $\neg\phi$ and $\neg\psi$ is in *T*.

(H2) If $\neg\neg\phi$ is in *T* then ϕ is in *T*.

(H3) For every closed term t, $(t = t)$ is in *T*.

(H4) If $(s = t)$ and $\phi(s/x)$ are both in *T* then $\phi(t/x)$ is in *T*.

(H5) If $\exists x\, \phi(x)$ is in *T* then for some constant c, $\phi(c)$ is in *T*; if $\neg\exists x\, \phi(x)$ is in *T* then $\neg\phi(t)$ is in *T* for every closed term t.

(H6) If $\forall x\, \phi(x)$ is in *T* then $\phi(t)$ is in *T* for every closed term t; if $\neg\forall x\, \phi(x)$ is in *T* then for some constant c, $\neg\phi(c)$ is in *T*.

(H7) If ϕ is an atomic sentence then ϕ and $\neg\phi$ are not both in *T*.

METATHEOREM 15. *If L is a first-order language, A is an L-structure and every element of A is named by a constant, then the set of all sentences of L that are true in A is a Hintikka set.*

METATHEOREM 16. *If the first-order language L has at least one constant and T is a Hintikka set in L then T has a model.*

We sketch the proof of Metatheorem 16. Let C be the set of all closed terms of L. Define a relation ~ on C by: $s \sim t$ if and only if $(s = t)$ is in T. Then we can show, using (H3) and (H4), that ~ is an equivalence relation on C. Write $t^{\sim}$ for the equivalence class of the closed term t, and $C^{\sim}$ for the set of equivalence classes $t^{\sim}$. Since L has at least one constant, $C^{\sim}$ is not empty. We shall build an L-structure A whose domain is $C^{\sim}$. For each constant c we take c_A to be $c^{\sim}$. If F is a function symbol of arity n and $t_1, \dots, t_n$ are closed terms, we define $F_A(t_1^{\sim}, \dots, t_n^{\sim})$ to be the equivalence class

(91) $$F(t_1, \dots, t_n)^{\sim}.$$

We can use (H4) to justify this definition. An argument by induction on complexity shows that for each closed term t of L, t_A is $t^{\sim}$. If R is a relation symbol of arity n and $t_1, \dots, t_n$ are closed terms, we define

(92) $(t_1^{\sim}, \dots, t_n^{\sim})$ is in R_A if and only if $R(t_1, \dots, t_n)$ is in T.

Again this definition is justified by an argument involving (H4). This completes the definition of the structure A.

Now we prove, by induction on the complexity of ψ, that for every sentence ψ of L, if ψ is in T then $A \vDash \psi$, and if $\neg\psi$ is in T then $A \vDash \neg\psi$. A typical clause is where ψ is $\exists x\, \phi(x)$. If ψ is in T then by (H5) there is a constant c such that $\phi(c)$ is in T. Since $\phi(c)$ has lower complexity than $\exists x\, \phi(x)$, the induction hypothesis shows that $A \vDash \phi(c)$. So $A \vDash \exists x\, \phi(x)$. On the other hand if $\neg\exists x\, \phi(x)$ is in T, then by (H5) again and induction hypothesis, $A \vDash \neg\phi(t)$ for every closed term t, so that $A \vDash \neg\phi[t_A]$. Since all elements of A are of the form t_A, this shows that $A \vDash \neg\exists x\, \phi(x)$. Thus A is a model of every sentence in T, proving the metatheorem.

As an example we shall show that the sequent

(93) $\forall x\, \neg(P(x) \wedge \neg Q(x)), \forall x\, \neg(Q(x) \wedge \neg R(x)) \vDash \forall x\, \neg(R(x) \wedge \neg P(x))$

is not valid. ((93) is the sequent (15) but with the conclusion reversed and " → " removed in favour of "¬" and " ∧ .") We begin by noting that by Metatheorem 8, the sequent is valid if and only if the theory consisting of the sentences

(94) $\forall x\, \neg(P(x) \wedge \neg Q(x)), \forall x\, \neg(Q(x) \wedge \neg R(x)), \neg\forall x\, \neg(R(x) \wedge \neg P(x))$

is inconsistent. So we can show the invalidity of (93) by constructing a model of these three sentences. We aim to build a Hintikka set that contains the sentences.

Property (H5) of Hintikka sets and the hypothesis of Metatheorem 16 alert us that we may need to call on constants. Maybe L has no constants; maybe it has constants, but they are all used in sentences of T that make their use for (H5) and (H6) impossible. So as a first step we allow ourselves to add new constant symbols to the language when needed. Metatheorem 2 tells us that this expansion of L makes no difference to the consistency of T. The added constants are called *witnesses*, since in (H5) the sentence $\phi(c)$ serves as a witness to the truth of $\exists x\phi$.

We begin by writing the sentences (94):

(95)
$$\forall x\, \neg(P(x) \wedge \neg Q(x))$$
$$\forall x\, \neg(Q(x) \wedge \neg R(x))$$
$$\neg\forall x\, \neg(R(x) \wedge \neg P(x))$$

At this point we apply the second part of (H6) to the third sentence. This requires us to introduce a witness, say *c*. A Hintikka set containing $\neg\forall x\, \neg(R(x) \wedge \neg P(x))$ needs to contain $\neg\neg(R(c) \wedge \neg P(c))$, so we add this to the diagram. We notice that by (H2) a Hintikka set containing this new sentence must also contain $(R(c) \wedge \neg P(c))$, so we add this too. Next (H1) tells us that a Hintikka set containing $(R(c) \wedge \neg P(c))$ must also contain $R(c)$ and $\neg P(c)$, so we add these below. Then by the first part of (H6) we also need to add $\neg(P(c) \wedge \neg Q(c))$ and $\neg(Q(c) \wedge \neg R(c))$, so we add them.

(96)
$$\forall x \neg(P(x) \wedge \neg Q(x))$$
$$\forall x \neg(Q(x) \wedge \neg R(x))$$
$$\neg\forall x \neg(R(x) \wedge \neg P(x))$$
$$|$$
$$\neg\neg(R(c) \wedge \neg P(c))$$
$$(R(c) \wedge \neg P(c))$$
$$|$$
$$R(c)$$
$$\neg P(c)$$
$$|$$
$$\neg(P(c) \wedge \neg Q(c))$$
$$\neg(Q(c) \wedge \neg R(c))$$

Here we meet a problem. The Hintikka set that we are constructing contains $\neg(P(c) \wedge \neg Q(c))$, so by the second

part of (H1) it must contain either $\neg P(c)$ or $\neg\neg Q(c)$. But we don't yet know which will work; so we try both possibilities. The diagram will branch. To the left we try to construct a Hintikka set containing $\neg P(c)$, and to the right, a Hintikka set containing $\neg\neg Q(c)$ (and hence also $Q(c)$ by (H2)). The same rule (H1) applies to the sentence $\neg(Q(c) \wedge \neg R(c))$, and tells us to add either $\neg Q(c)$ or $\neg\neg R(c)$. We must try each of these choices in each branch, so we have a double branching. At this point we notice that the third branch from the left contains both the atomic sentence $Q(c)$ and its negation $\neg Q(c)$; so by (H7) there is no hope of extending this branch to a Hintikka set, and we close it with a horizontal line across the bottom.

(97)

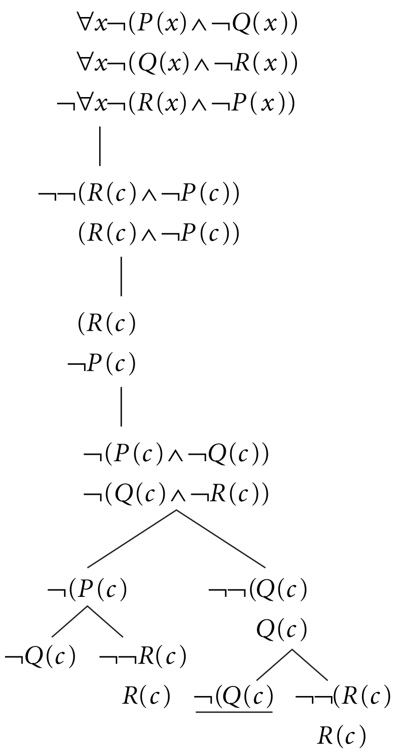

The first, second, and fourth branches are still open. Inspection shows that all that is needed to turn the contents of each of these branches into a Hintikka set is to add the equation $(c = c)$; so we do that. Now choose one of the open branches, say the first. Since its contents (reckoning from the top of the diagram) form a Hintikka set, it has a model, and this model is a model of the first three sentences in particular. So we have shown that the sequent (93) is not valid.

What would happen if we replaced the conclusion of (93) by $\forall x\, \neg(P(x) \wedge \neg R(x))$? The resulting sequent is a paraphrase of (15) from section 1, and we claimed that that sequent was valid. Here is the tree for the revised sequent:

(98)

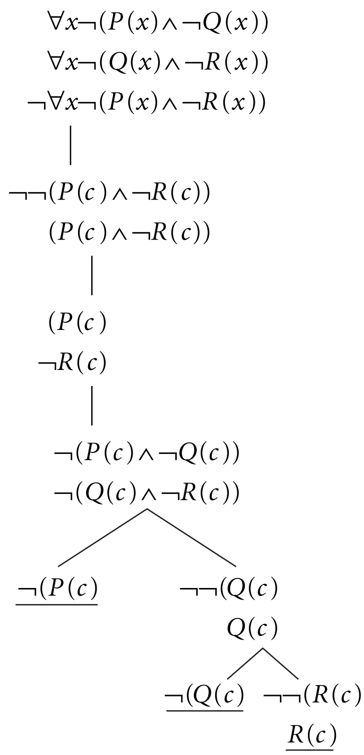

Here every branch contains an atomic sentence together with its negation, so they all close and we say that the diagram is *closed*. This closed diagram shows that there is no Hintikka set containing the top three sentences. Hence by Metatheorem 15 they have no model, and this shows that our revised sequent is valid, giving a formal justification of (15).

PROOF CALCULI

A proof calculus is a mathematical device for proving the validity of sequents. In any proof calculus a central notion is that of a *formal proof of a sequent*. For most proof calculi a formal proof is an array of symbols that can be written on a page. But some calculi are for use by computers and their formal proofs are not meant for visual inspection. On the other hand the proof calculus *Hyperproof* (Jon Barwise and John Etchemendy) allows formal proofs that consist of sequences of labelled pictures.

If a proof calculus $\mathscr{C}$ contains a formal proof of the sequent

(99) $$\phi_1, \ldots, \phi_n \vDash \psi,$$

we express this fact by writing

(100) $$\phi_1, \ldots, \phi_n \vdash_{\mathscr{C}} \psi.$$

The expression (100) is called a *syntactic sequent*. When we are discussing a particular proof calculus $\mathscr{C}$, we can drop the subscript from $\vdash_{\mathscr{C}}$ and write simply $\vdash$; the symbol "$\vdash$" is read as "turnstile."

In many proof calculi a formal proof of a sequent is a formalisation of an argument that one might use in order to persuade someone that the sequent is valid. For example the *natural deduction* calculus proposed by Gerhard Gentzen in 1934 is designed to make the same moves as are used in "natural" mathematical arguments. But in general a formal proof need not have any visible connection with arguments. The main requirements on a proof calculus $\mathscr{C}$ are as follows.

(a) Whenever a syntactic sequent (100) holds, the corresponding semantic sequent (99) is valid. A proof calculus satisfying this condition is said to be *sound*.

(b) Whenever a semantic sequent (99) is valid, the corresponding syntactic sequent (100) holds. A proof calculus satisfying this condition is said to be *complete*.

(c) A computer can identify those arrays of symbols that are formal proofs in $\mathscr{C}$, and for each formal proof and each finite semantic sequent, the computer can calculate whether or not the proof is a formal proof of the sequent.

Soundness says that $\mathscr{C}$ doesn't prove any sequent that it ought not to; completeness says that $\mathscr{C}$ does prove any sequent that it ought to.

All the proof calculi commonly taught to undergraduates are both sound and complete. There is one main exception: the resolution calculus is limited to proofs of sequents of the form

(101) $$\forall x_1 \ldots \forall x_m \, \phi \vDash$$

where ϕ is quantifier-free and in conjunctive normal form. Computer science students who study this calculus also learn how to reduce more general proof problems to this form. Also some Hilbert-style calculi are only able to prove sequents of the form $\vDash \psi$.

If a proof calculus $\mathscr{C}$ is sound and complete, then finite sequents with $\vDash$ are valid if and only if the corresponding sequents with $\vdash_{\mathscr{C}}$ are also valid. It follows that all the metatheorems of section 8 using $\vDash$ remain true when they are stated with $\vdash_{\mathscr{C}}$. But the versions with $\vdash_{\mathscr{C}}$ generally have direct proofs using syntactic properties of the proof calculus $\mathscr{C}$. Sometimes these direct proofs (particularly proofs of the Deduction Theorem) play a role in proving that $\mathscr{C}$ is complete. One deathtrap for unwary teachers is the Cut rule, Metatheorem 7. This is very easy to prove directly for some calculi, for example natural deduction. But for the truth tree calculus below, the truth of the Cut rule is a deep fact equivalent to Gentzen's cut elimination theorem; a syntactic proof of it is a major enterprise.

The entry "Proof Theory" contains much more information about proof calculi. For example it discusses how one can translate proofs in one proof calculus into proofs in another. If two proof calculi translate into each other in this way, then clearly soundness and completeness theorems for one calculus carry over to the other calculus too. This is probably the main reason why logicians often talk about "the completeness theorem" as if there was a single theorem for all proof calculi, when strictly each complete proof calculus has its own completeness theorem.

We introduced a kind of proof calculus when we discussed the construction of models. Suppose we have a tree diagram in the style of that section, showing that there is no Hintikka set containing the sentences

(102) $$\phi_1, \ldots, \phi_n, \neg\psi.$$

Then in view of Metatheorem 8 we count this diagram as a formal proof of the sequent

(103) $$\phi_1, \ldots, \phi_n \vDash \psi.$$

The proof calculus based on Hintikka sets in this way is called *truth trees*.

The truth tree calculus is sound by Metatheorem 15. In order to establish that the calculus is complete, we must show the following: If it is not possible to construct a closed truth tree in a finite number of steps starting from $\phi_1, \ldots, \phi_n, \neg\psi$, then it is possible to construct a truth tree starting with these sentences, in which one branch forms a Hintikka set (so that it has a model by Metatheorem 16). We have to bear in mind that a branch might go on forever. In fact if L has infinitely many closed terms, then (H3) implies that every Hintikka set in L must be infinite. In this case the conditions (H1)–(H7) impose infinitely many separate requirements on a Hintikka set, and we have to be sure that we construct our branches in such a way that each of these requirements will eventually be faced and met if possible. This can be arranged.

DECIDABILITY

When the signature σ is finite, we can assign natural number values to the symbols of $L(\sigma)$ and thereby

express each formula of $L(\sigma)$ as a finite sequence of natural numbers. (In fact this is possible, though less tidy, when σ is countably infinite.) Hence it makes sense to apply notions of computability theory to $L(\sigma)$. Thus we say that a set X of finite sequences of numbers is *computably enumerable* (abbreviated to *c.e.*) if a computer can be programmed to output all and only the sequences in X. Also we say that X is *computable* if a computer can be programmed to output Yes if a sequence in X is input, and No if a sequence not in X is input. These notions carry over immediately to theories in $L(\sigma)$. We say also that a theory T is *decidable* if the set of its logical consequences is computable. A procedure for computing whether any given sentence is a consequence of T is called a *decision procedure* for T.

Now suppose we have a proof calculus $\mathscr{C}$ for L that meets the conditions (a), (b) and (c) of section 10. Then $\mathscr{C}$ is sound and complete, so that we have the equivalence

(104) ϕ is a consequence of T if and only if:

There is a finite subset U of T and there is P such that

[P is a formal proof in $\mathscr{C}$ of the sequent $U \vDash \phi$].

By property (c) of $\mathscr{C}$, the relation in square brackets is computable. It follows that if T is computably enumerable, we can program a computer so that it (i) lists all possible finite subsets U of T, all sentences ϕ of $L(\sigma)$ and all formal proofs P in $\mathscr{C}$, (ii) tests, for each U, ϕ and P, whether or not P is a proof of $U \vDash \phi$, and (iii) outputs ϕ whenever the answer to (ii) is Yes. This computer will output all and only the logical consequences of T. We have shown:

METATHEOREM 17. *If T is a c.e. theory in a first-order language with finite signature, then the set of consequences of T is also c.e.*

An important corollary is:

METATHEOREM 18. *Suppose T is a complete, consistent, and c.e. theory in a first-order language with finite signature. Then T is decidable.*

To compute whether ϕ is a consequence of T, list all the consequences of T; eventually either ϕ or $\neg\phi$ will appear in the list. This is not a practical method. But when this abstract argument shows that T is decidable, one can usually find a much better decision procedure.

An important case is to determine whether a given sentence ϕ is a consequence of the empty theory, that is, whether ϕ is a theorem of first-order logic. By a result proved by Alonzo Church in 1936, the set of theorems of first-order logic (say, in a signature with at least one relation symbol of arity at least 2) is not computable. But if we restrict ϕ to come from some appropriate class of sentences, the picture changes. Suppose for example that ϕ is in propositional logic. Then we can construct a truth tree to determine whether $\neg\phi$ has a model, and after a finite number of steps the truth tree will have ground to a halt in the sense that there are no new sentences that we can add to it. If all its branches are closed, there is no model of $\neg\phi$; if at least one branch remains open, it gives us a Hintikka set and hence a model of $\neg\phi$. This provides a decision procedure for propositional sentences.

In fact the truth tree procedure allows us to test for theoremhood every first-order sentence of the form $Q_1Q_2\phi$ where Q_1 is a string of universal quantifiers, Q_2 is a string of existential quantifiers and ϕ is a quantifier-free formula with no function symbols.

A formula is said to be *universal* if it is quantifier-free or consists of a string of universal quantifiers followed by a quantifier-free formula. Thoralf Skolem proved that the problem of determining whether a given first-order sentence has a model can always be reduced to the problem of determining whether a certain universal first-order sentence has a model:

METATHEOREM 19. *Let L be a first-order language with finite signature σ. There is a computational procedure which, given any sentence φ, will find a universal first-order sentence* ϕ^{sk} *with the following properties:*

(a) ϕ^{sk} *is in a signature got by adding function symbols and constants to σ.*

(b) Every model of ϕ^{sk} *is a model of φ.*

(c) Every σ-structure that is a model of φ can be made into a model of ϕ^{sk} *by adding suitable interpretations for the new function symbols and constants.*

In particular φ has a model if and only if ϕ^{sk} *has a model.*

We can illustrate Skolem's idea. Let ϕ be the sentence

(105) $\forall x\, \exists y \forall z \exists w\, R(x,y,z,w)$.

Then for ϕ^{sk} we can take the sentence

(106) $\forall x\, \forall z\, R(x,F(x),z,G(x,z))$.

The function symbols F and G, and the functions they represent in a model of ϕ^{sk}, are called *Skolem functions*.

Thus if we want to determine whether ϕ is a first-order theorem, one way to proceed is to find $(\neg\phi)^{sk}$ and determine whether it has a model or not. There is no

guarantee that this approach will settle the question; the importance of Skolem's theorem is mainly theoretical.

See also Church, Alonzo; Conditionals; Davidson, Donald; De Morgan, Augustus; Frege, Gottlob; Hilbert, David; Hintikka, Jaakko; Leibniz, Gottfried Wilhelm; Lewis, David; Logic, History of; Logic, Non-Classical; Łukasiewicz, Jan; Mathematics, Foundations of; Peirce, Charles Sanders; Proof Theory; Russell, Bertrand Arthur William; Second-Order Logic; Tarski, Alfred; Whately, Richard.

Bibliography

In the citations below, modern introductory texts not requiring mathematical sophistication are marked with an *.

*Barwise, Jon, and John Etchemendy. *Hyperproof*. Stanford, CA: CSLI, 1994.

*Bergmann, Merrie, James Moor, and Jack Nelson. *The Logic Book*. New York: McGraw-Hill, 1997.

Boolos, George. "For Every A There is a B." *Linguistic Inquiry* 12 (1981): 465–466.

Church, Alonzo. *Introduction to Mathematical Logic, Volume 1*. Princeton, NJ: Princeton University Press, 1956.

Dalen, Dirk van. *Logic and Structure*. Berlin: Springer, 1980.

Davidson, Donald. "The Logical Form of Action Sentences." In *Essays on Actions and Events*, edited by Donald Davidson, 105–148. Oxford: Clarendon Press, 2001.

Edgington, Dorothy. "Do Conditionals Have Truth-Conditions?" *Critica* xviii, 52 (1986): 3–30. Reprinted in *Conditionals*, edited by Frank Jackson. Oxford: Oxford University Press, 1991.

*Forbes, Graeme. *Modern Logic: A Text in Elementary Symbolic Logic*. New York: Oxford University Press, 1994.

*Gabbay, Dov. *Elementary Logics: A Procedural Perspective*. Hemel Hempstead: Prentice Hall Europe, 1998.

Grice, Paul. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

Hilbert, David, and Wilhelm Ackermann. *Principles of Mathematical Logic*. Translated from the second edition of *Grundzüge der Theoretischen Logik*. New York: Chelsea, 1950.

Hintikka, Jaakko. "Form and Content in Quantification Theory." *Acta Philosophica Fennica* 8 (1955): 11–55.

Hodges, Wilfrid. "Elementary Predicate Logic." In *Handbook of Philosophical Logic*. 2nd ed., vol. 1, edited by Dov Gabbay and Franz Guenthner. Dordrecht: Kluwer, 2001.

*Hodges, Wilfrid. *Logic*. 2nd ed. London: Penguin, 2001.

*Jeffrey, Richard C. *Formal Logic: Its Scope and Limits*. New York: McGraw-Hill, 1994.

Kalish, Donald, and Richard Montague. *Logic, Techniques of Formal Reasoning*. New York: Harcourt Brace, 1964.

Kleene, Stephen Cole. *Introduction to Metamathematics*. Amsterdam, North-Holland, 1964.

Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.

*Mates, Benson. *Elementary Logic*. New York: Oxford University Press, 1965.

Mendelson, Elliott. *Introduction to Mathematical Logic*. 3rd ed. Belmont, CA: Wadsworth and Brooks/Cole, 1987.

Peirce, Charles S. "On the Algebra of Logic: A Contribution to the Philosophy of Notation." *American Journal of Mathematics* 7 (1885): 180–202.

Peirce, Charles S. "The Reader Is Introduced to Relatives." *The Open Court* 6 (1892): 3416–3418.

*Quine, Willard Van Orman. *Methods of Logic*. London: Routledge and Kegan Paul, 1952.

Shapiro, Stewart. "Classical Logic." *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu.

Shoenfield, Joseph R. *Mathematical Logic*. Reading, MA: Addison-Wesley, 1967. Republished Urbana, IL: Association for Symbolic Logic and A. K. Peters, 2001.

Smullyan, Raymond M. *First-Order Logic*. New York: Dover Publications, 1995.

*Suppes, Patrick. *Introduction to Logic*. Princeton, NJ: Van Nostrand, 1957.

Velleman, Daniel J. *How To Prove It*. Cambridge, U.K.: Cambridge University Press, 1994.

Whately, Richard. *Elements of Logic*. London: Mawman, 1826. Reprinted Bologna: Editrice CLUEB, 1987.

Wilfrid Hodges (2005)

FISCHER, KUNO
(1824–1907)

Kuno Fischer, the German philosopher and historian of philosophy, was born at Sandewalde in Silesia. He studied philology at Leipzig and theology and philosophy at Halle. In 1850 Fischer was appointed *Privatdozent* in philosophy at the University of Heidelberg, but his pantheistic views caused his dismissal three years later. In 1856 he qualified as *Privatdozent* at the University of Berlin, and in the same year he was invited to Jena as professor of philosophy. In 1872 he returned to Heidelberg, where he taught with great success until 1903.

Fischer's major work is his *Geschichte der neueren Philosophie* (1852–1877). This widely reprinted history of modern philosophy owed its success in large part to Fischer's splendid gift for exposition. Endowed with a remarkable capacity for sympathetic understanding, Fischer was able to reproduce the great philosophical systems in a literary form of exemplary brilliance and clarity, as well as to unravel their basic themes and subtlest ramifications and to illuminate and reconstruct them systematically. At the same time, he sought to place these systems in their larger cultural and historical context and thus to understand the historical development of philosophy as the progressive self-knowledge of the human mind.

Fischer was the author of the first large German monograph on Immanuel Kant, *Kants Leben und die*

Grundlagen seine Lehre (Mannheim, 1860), and it is from Fischer that Neo-Kantianism received its decisive impulse.

Apart from Kant, G. W. F. Hegel was the chief object of his interest. In its equating of logic and metaphysics, Fischer's *System der Logik und Metaphysik oder Wissenschaftslehre* (Stuttgart, 1852) exhibited the strong influence of Hegel. In this work an attempt was also made to bring Hegel's principle of dialectical development into harmony with modern evolutionism in the sense of a teleological idealism. Fischer held that the dialectical development ran from Being through Essence to purpose. The system of logical and, at the same time, metaphysical categories that he outlined culminated in the idea of finality, which guaranteed a purposeful development that goes beyond the merely given. In the second edition (Stuttgart, 1865), Fischer attempted to mediate between Kant and Hegel and to do justice not only to Hegelianism but also to Kantianism and empiricism.

Arthur Schopenhauer also influenced Fischer. In the study *Das Verhältnis zwischen Willen und Verstand im Menschen* (Heidelberg, 1896), Fischer distinguished between the will that is guided by knowledge and the unconscious volition that precedes all knowledge and conscious behavior. He also claimed that just as the essence of nature is "force," so the essence of man is "will" and the essence of the body is the manifestation of volition.

Fischer was also an extremely productive literary aesthetician. His conception of art is to be found in his early publication *Diotima: Die Idee des Schönen* (Diotima: the idea of the beautiful; Pforzheim, 1849). In this work Fischer defined the aesthetic attitude as one of "playing," characterized by a concentration and uniting of all our faculties. He devoted later works to the origin and development of humor and to the classical poetry of William Shakespeare, Gotthold Ephraim Lessing, Johann Wolfgang von Goethe, and Friedrich Schiller.

See also Aesthetic Attitude; Empiricism; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hegelianism; History and Historiography of Philosophy; Humor; Kant, Immanuel; Lessing, Gotthold Ephraim; Neo-Kantianism; Schiller, Friedrich; Schopenhauer, Arthur.

Bibliography

ADDITIONAL WORKS BY FISCHER

Fischer's chief work is the *Geschichte der neueren Philosophie* (Stuttgart, Mannheim, and Heidelberg, 1852–1877). Originally published in six volumes (on René Descartes, Benedict de Spinoza, Gottfried Wilhelm Leibniz, Kant, Johann Gottlieb Fichte, and Friedrich Schelling), it was later extended to ten volumes, with a second volume on Kant, one on Schopenhauer (1893), and one on Hegel (1901); and it also included Fischer's early work on Francis Bacon and his school (Leipzig, 1856). Other writings include *Kleine Schriften,* 8 vols. (Heidelberg, 1888–1898), which contains *Über den Witz* (On humor), *Über die menschliche Freiheit* (On human freedom), *Das Verhältnis zwischen Willen und Verstand* (The relationship between will and understanding), and other essays. See also *Philosophische Schriften* (Philosophical writings), 6 vols. (Heidelberg, 1891–1892; 6th ed., 1908–1909).

English translations of Fischer's works include the following: J. Oxenford, *Francis Bacon* (London, 1857); J. P. Mahaffy, *A Commentary on Kant's "Critick of the Pure Reason"* (London and Dublin, 1866); F. Schmidt, *Benedict Spinoza* (Edinburgh, 1882); J. P. Gordy, *Descartes and His School* (New York, 1887); W. S. Hough, *A Critique of Kant* (London: S. Sonnenschein, 1888).

WORKS ON FISCHER

For literature on Fischer, see Hugo Falkenheim, *Kuno Fischer und die literarhistorische Methode* (Berlin, 1892); Wilhelm Windelband, *Kuno Fischer* (Heidelberg, 1907), a memorial address; and Ernst Hoffmann, *Kuno Fischer* (Heidelberg, 1924).

Franz Austeda (1967)

Translated by Albert E. Blumberg

FISHER, R. A.

(1890–1962)

Ronald Aylmer Fisher was a titan who bestrode two signature disciplines of twentieth-century science: population genetics (or the mathematical theory of evolution), of which he was a cofounder and principal architect, and mathematical statistics, in which he played a pivotal role. On the one hand, he led a revolution that replaced the Bayesian approach of inverse probability with one based solely on direct probabilities (i.e., probabilities of outcomes conditional on hypotheses). On the other hand, he unequivocally rejected the conception of statistics as decision making under uncertainty that his own work inspired. This rift in the new statistical orthodoxy has never healed. Thus, Fisher's conception of probability was at once frequentist and epistemic, his approach to statistics at once inferential and non-Bayesian, and the chief question his life's work poses is whether a consistent theory can be built along these lines.

After excelling in mathematics at the secondary level, Fisher won a scholarship to Cambridge University in 1909 and graduated in 1912 as a Wrangler (i.e., with hon-

ors) in the Mathematical Tripos, then spent another year at Cambridge studying statistical mechanics and quantum theory under the astronomer James Jeans. In the 1911 paper (unpublished at the time) "Mendelism and Biometry," he pointed the way to a synthesis of Mendelian genetics and Darwinian evolution.

Fisher received two important job offers in 1919: one as chief statistician under Karl Pearson at the Galton Laboratory of University College, London, and the other a temporary position at the Rothamsted (Agricultural) Experimental Station. Fisher was already on famously bad terms with Pearson, so he accepted the Rothamsted offer, leaving him free to develop his own non-Bayesian approach to statistics free of Pearson's supervision. Over the next fifteen years Fisher developed a world renowned Department of Statistics at Rothamsted that became a training ground for many statisticians who disseminated his new methods far and wide. Fisher's "golden age of invention" at Rothamsted ended in 1933 when Karl Pearson retired and his department was split into a Department of Statistics, with Egon S. Pearson (Karl's son) as head, and a Department of Eugenics, with Fisher as head. Relations between the two departments were never cordial. Further details may be found in the biography by Fisher's daughter, Joan Fisher Box (1978), who gives excellent sketches of his many and varied contributions as well as his side of the many protracted debates in which he engaged.

FISHER AND THE BAYESIANS

Although weaned on inverse probability at school (Fisher 1950, 27.248), Fisher came to regard Bayesian solutions as vitiated by the arbitrary and subjective character of prior distributions not squarely based on frequency data. Replying to criticism by Karl Pearson of a Bayesian solution he had proposed in his earliest published paper, he noted that the solution favored by Pearson "depends almost wholly upon the preconceived opinions of the computer and scarcely at all upon the actual data" (Fisher 1971–1974, 14.17). This led him to emphasize the need to "allow the data to speak for themselves," an injunction some of his followers carry to the extreme of deliberately ignoring, for example, all prior information bearing on the efficacy of a new medical treatment. To the Bayesians' palliative that whatever errors of estimation arise from use of an inappropriate prior will become negligible with accumulating data, he retorts that "it appears more natural to infer that it should be possible to draw valid conclusions from the data alone and without *a priori* assumptions." Then he adds, "we may question whether the whole difficulty has not arisen in an attempt to express in terms of the single concept of mathematical probability, a form of reasoning which requires for its exact statement different though equally well-defined concepts" (Fisher 1950, 24.287).

Of the alternative measures suitable for "supplying a natural order of preference" among competing estimates or hypotheses, Fisher recommended the likelihood function (LF) or the data distribution qua function of the unknown parameter(s) of one's model. Or, when the LF is undefined (i.e., when the probability of the observed outcome conditional on the alternative hypotheses cannot be computed from the model), significance tests are in order. Now the LF provides only relative probabilities and is nonadditive, but the logarithm of the LF is additive and this allows one to combine evidence from different (independent) sources. The value of the unknown parameter that maximizes the LF—the so-called maximum likelihood estimate (MLE)—when it exists and is unique, must then be the best supported value. Fisher's first task was to provide a rationale for this evidential use of the LF, which Pierre Simon de Laplace and Carl Gauss had drawn as a corollary of Bayesian conditioning, but that, from Fisher's perspective, "has no real connection with inverse probability" (*Statistical Methods for Research Workers* in Fisher 2003, p. 22).

THEORY OF ESTIMATION

The first thing that struck him is that, unlike the uniform prior Thomas Bayes and Laplace seemed to conger out of ignorance, the MLE is invariant. That is, if a problem is reparametrized as $\zeta = g(\theta)$, then the MLE of the new parameter, ζ, is $g(\hat{\theta})$, writing ($\hat{\theta}$) (throughout) for the MLE of θ (De Groot 1986, p. 348). At the same time, he noted, unbiased estimators—those whose mean is equal to θ—are noninvariant, an unbiased estimator of θ being a biased estimator of θ^2 or θ^{-1}. His requirement of invariance is, in reality, a requirement of consistency, namely, that one's estimates and inferences do not depend on which of several equivalent forms of a problem one adopts. This already brings Fisher closer to the position of his protagonist, Harold Jeffreys, or that of Jeffreys's worthy successor, Edwin T. Jaynes. Nor did it ever occur to Fisher, as it did to Jeffreys and Jaynes, to use an invariant prior to represent, not pure ignorance, but a state of knowledge that is unaltered by a specifiable group of transformations. Knowing, for example, no more than that θ is a scale parameter, a suitable prior—the Jeffreys prior—would be one invariant under changes of scale.

However, Fisher was not satisfied with this justification of MLEs, but insisted that "the reliance to be placed" on one "must depend on its frequency distribution." (Fisher 1950, 10.327) Thus, Gauss had shown that the arithmetic average (or sample mean) of a set of normally distributed errors of known variance, $\frac{\sigma^2}{n}$, is itself normally distributed about the population mean, μ, with variance μ^2/n. Since a normal distribution is determined by its location parameter, μ, which locates the bell-shaped density curve along the x-axis, and its scale parameter, σ^2, which measures the spread, the variance presents itself as the uniquely suitable measure of the concentration of any estimator whose distribution is normal or asymptotically normal about the estimated parameter. What Fisher claimed to show in his seminal 1922 paper, "On the Mathematical Foundations of Theoretical Statistics" (Fisher 1971–1974: paper 18; Fisher 1950, paper 10), is that MLEs are the most concentrated. He dubbed such estimators of (asymptotically) smallest variance efficient.

One source of tension in Fisher is that his use of likelihood implies the irrelevance of outcomes that might have been but were not observed, and, at various places, he explicitly endorses this implication (*Statistical Methods and Scientific Inference* in Fisher 2003, pp. 71, 91; hereafter *SMSI*). For if as he says "the whole of the information supplied by a sample … is comprised in the likelihood" (p. 73), the LF of the outcome actually observed, all other points of the sample space must be irrelevant. However, the sampling (or frequency) distribution of an estimator, T, depends on the whole sample space, and its use to compare estimators therefore violates this likelihood principle.

In the course of his investigation of the large-sample properties of MLEs, Fisher uncovered a class of statistics a knowledge of which renders all other statistics irrelevant for inferences about θ, and so he termed them *sufficient* for θ. In the classic 1922 paper, he showed that sufficient estimators are asymptotically efficient, thus linking a purely logico-informational requirement—that of utilizing all the information supplied by the data—with a performance characteristic—that of having maximal precision. In fact, he virtually equated the property of not wasting information with efficiency. Then he could describe the statistician's job succinctly in purely cognitive terms as that of effecting the maximum information-preserving reduction of the data (Fisher 1950, 26.366). Such a maximal reduction is called a minimal sufficient statistic and is mathematically a function of every other sufficient statistic. Philosophers will recognize sufficiency as a close relative of Rudolf Carnap's requirement of total evidence, and Fisher remarks that "our conclusions must be warranted by the whole of the data, since less than the whole may be to any degree misleading" (Fisher 1950, 26.54).

Fisher's claim that maximal likelihood estimation is "unequivocally superior" to all other methods (Fisher 1950, 24.287) would then be vindicated, at least for large samples, by showing that MLEs are sufficient (hence, asymptotically efficient). His proof of this in the 1922 paper was less than rigorous, as he candidly admitted (Fisher 1950, 10.323), and he offered improved versions in sequels to that paper. In the 1934 paper "On Two New Properties of Mathematical Likelihood," CMS paper #.24, he presented a new criterion of sufficiency, namely, that the LF factors as

(1) $$p(x|\theta) = g(T, \theta)h(x)$$

which allows one to recognize a sufficient statistic at sight. This was of great importance because the property of utilizing all the information in one's data can be applied to estimators based on small samples. And Fisher's experimental work in genetics and agronomy (at Rothamsted) had impressed on him the great practical importance of statistical methods applicable to small samples, and, hence, of exact tests or estimates based on exact, as opposed to approximate, sampling distributions. In this he was also strongly influenced by W. S. Gossett's 1908 discovery of the exact distribution of the statistic,

$$n^{\frac{1}{2}}(\bar{x} - \theta)/s$$

where

$$s^2 = (n-1)^1 \sum_{i=1}^{n} (x_i - \bar{x})^2$$

is the sample variance, which could then be used to test hypotheses about normal means using a small sample when the variance of the measurements is unknown. Thus, he came to view large sample theory, concerned with the never-never world of asymptotic behavior, as a mere preliminary to the study of small samples (*SMSI*, p. 163).

To facilitate the study of small samples, he introduced a quantitative measure of information. His leading idea was to measure the information an experiment with outcome variate X conveys about an unknown parameter θ by the precision (or inverse variance) of an MLE of θ. Earlier work of Karl Pearson and Francis Ysidro Edgeworth, the two leading figures of the British school of statisticians of the generation preceding Fisher's, had

linked the precision of an estimator to the second derivative of the logarithm of the LF, ln $p(x|\theta)$, where $x = (x_1, \ldots, x_n)$, which one denotes $L(x|\theta)$, or even by $L(\theta)$. For example, to find the MLE of a binomial parameter, p, noting that the LF and its logarithm have the same maxima, one solves the likelihood equation,

$$0 = L'(p) = \frac{x}{p} - \frac{n-x}{1-p},$$

the observed relative frequency of successes. Taking the second derivative, one finds:

$$L''(p) = -\left(\frac{x}{p^2} + \frac{n-x}{(1-p)^2}\right)$$

whereupon replacing x by its mean, np, reduces this to:

$$-E[L''(p)] = \frac{np}{p^2} + \frac{n(1-p)}{(1-p)^2} = \frac{n}{p(1-p)}$$

the variance of p. "This formula," he declares, "supplies the most direct way known to me of finding the probable error of statistics," adding (with critical reference to Pearson) that "the above proof [not shown here] applies only to statistics obtained by the method of maximum likelihood" (Fisher 1950, 10.329).

Now one might hope to show that the Fisher information, defined by

$$(2) \qquad I_n(\theta) = -E[L''(x|\theta)]$$

imposes an upper limit on the precision of any estimator of θ for any given sample size n. To make a long tangled story short, Edgeworth proved special cases of this using the Schwarz inequality and Fisher extended his results (see Hald 1998, pp. 703–707, 716–719, 724–726, 734), offering a proof (again less than rigorous) that $V(T) \geq 1/I_n(\theta)$. The first rigorous proofs came in the 1940s (Cramer 1946, p. 475; De Groot 1986, p. 425) and a general form of this so-called Cramer-Rao inequality reads:

$$(3) \qquad \text{var}(T) \geq m'(\theta)^2/I_n(\theta)$$

where $m(\theta) = E(T) = \int T(x)p(x|\theta)dx$. One's assumptions are that the density is defined for a nondegenerate interval that does not depend on θ and has (finite) moments up to second order. When $m(\theta) = \theta$, so that T is unbiased, (3) simplifies to $\text{var}(T) \geq 1/I_n(\theta)$, as anticipated by Edgeworth and Fisher. Estimators that achieve this minimum variance bound are called MVB estimators, and this condition effectively replaces asymptotic efficiency since it applies to samples of all sizes. Cramer then proved (1946, pp. 499ff) that if an efficient (or MVB) estimator T of θ exists, then the likelihood equation has a unique solution given by T, and that if a sufficient estimator of θ exists, any solution of the likelihood equation will be a function of that estimator. These results round out Fisher's small sample theory of estimation.

Fisher used his factorization criterion (1) for sufficient statistics to show that the distributions admitting a sufficient statistic are precisely those of the form:

$$(4) \qquad p(x|\theta) = F(x)G(\theta)\exp[u(x)v(\theta)]$$

provided that the range of X does not depend on θ, as it does for the uniform distribution on $[0, \theta]$ with θ unknown. Called the exponential class, (4) includes almost all the other distributions that figure prominently in applied probability and statistics, including the normal, Poisson, beta, gamma, and chi-squared distributions (and there is also a multiparameter form of (4)). Thus, the class (4) occupies a position of central importance, akin to that of the central limit theorem. Using a clever change of variable in the condition for equality in (3), Jaynes (2003, p. 519) shows that the exponential class is also the class of maxent distributions, those yielded by the principle of maximizing the (Shannon) entropy subject to one or more given mean value constraints. Thus, as Jaynes proclaims, "if we use the maximum entropy principle to assign sampling distributions, this automatically generates the distributions with the most desirable properties from the standpoint of … sampling theory (because the sampling variance of an estimator is then the minimum possible value)" (520). Once again, the fruits of Fisher's own investigations drew him closer to the objectivist Bayesian position that he so vigorously opposed. Indeed, the maximum entropy formalism can be used to generate either data distributions or prior distributions and is supported by the kinds of consistency properties Fisher also endorsed. Mathematics makes strange bedfellows!

Fisher information defined by (2), or, equivalently, by $I_n(\theta) = E[L'(x|\theta)^2] = \text{var}[L'(x|\theta)$, also plays a prominent role, as one would expect, in Fisher's theory of experimental design. Given multinomial data with category counts $a_1, \ldots, a_k$ and category probabilities $p_1(\theta), \ldots, p_k(\theta)$ that depend on a parameter θ, the Fisher information for a sample of one is:

$$(5) \qquad I(\theta) = \sum \frac{1}{p_i}\left(\frac{dp_i}{d\theta}\right)^2$$

Examples arise in genetics, especially linkage. For example, one may wish to compare the information about the linkage parameter θ (the recombination fraction) yielded by a double backcross, *AB/ab* × *ab/ab*, with that given by a single backcross, *Abab* × *Abab*. Under the former mating, the genotypes *AB/ab*, *Ab/ab*, *aB/ab*, *ab/ab* occur among the offspring with probabilities ½(1 – θ), ½(θ), ½(θ), and ½(1 – θ), and so

$$I(\theta) = \frac{2}{1-\theta}\left(-\frac{1}{2}\right)^2 + \frac{2}{\theta}\left(\frac{1}{2}\right)^2 + \frac{2}{\theta}\left(\frac{1}{2}\right)^2 + \frac{2}{1-\theta}\left(-\frac{1}{2}\right)^2$$

$$= \frac{1}{\theta(1-\theta)}$$

while for the single backcross one similarly finds $I(\theta) = 1/2\theta(1 - \theta)$, or half the information yielded by the double backcross. Further refinements arise when there is dominance in one or both factors (see Edwards 1992, pp. 148–149). For more examples, see chapter 11 of *The Design of Experiments* (in Fisher 2003; hereafter *DE*) and Kenneth Mather's *The Measurement of Linkage in Heredity* (1938).

SIGNIFICANCE TESTS

One comes, at last, to Fisher's second important measure for ordering hypotheses, namely, significance tests. The earliest significance tests were aimed at distinguishing a hypothesis of chance from one of cause or design (Hald 1998, §4.1). For example, is the perfect agreement of the wrong answers of two students on a multiple choice test due to collusion or a mere coincidence? In the usage of Laplace, one compares the probability of such agreement on the two hypotheses and when this probability is "incomparably greater" on the hypothesis of design, "we are led," he says, "to disbelieve" that of chance. Laplace readily extended this reasoning to the separation of "real" from "spurious" physical causes, as when he concluded that "the actual disposition of our planetary system," by which he meant that all six planets and their satellites move in the same direction as the earth and have inclinations to the ecliptic within a small neighborhood of zero, "would be infinitely small if it were due to chance" and so indicates a "regular cause" (§4.4). In the same vein, Gustav Kirkhoff concluded that the perfect coincidence of the sixty dark lines in the solar spectrum of iron with sixty bright lines of the spectrum obtained by heating iron filings in a Bunsen burner could not be due to chance but indicated the presence of iron in the sun.

In such cases, the probability of agreement on the hypothesis of design may be only qualitatively defined, but the logic is essentially that of a likelihood ratio test. Nor did Laplace speak in terms of rejecting the hypothesis of chance or prescribe a threshold of improbability beyond which belief gives way (or should give way) to disbelief. He took as his test criterion the tail area probability, that is, the probability of a deviation at least as large as that observed (Hald 1998, p. 25). Moreover, a low probability of observing so large a deviation by chance points to some alternative explanation that, however, need not be formulated beforehand. Rather, "by letting the remarkable feature [of the data] determine the statistic used in the test, we concentrate implicitly on an alternative hypothesis" (p. 67).

Fisher embraced most but not all these features. The *locus classicus* of his account is the famous treatment of the tea-tasting lady who claims to be able to tell whether milk or tea was added first to a mixture of the two (*DE*, chapter 2). Every serious student of inductive reasoning should read and reread this chapter with infinite care. Of great importance, too, is the fourth chapter of *SMSI*, "Some Misapprehensions about Tests of Significance."

To begin with, a significance test is, emphatically, not a decision rule (*DE*, §12.1; *SMSI*, §4.1], the differences between them being characterized as "many and wide" (*SMSI*, p. 80). Thus opens Fisher's trenchant critique of the Neyman-Pearson theory of testing. In choosing a test statistic, "the experimenter will rightly consider all points on which, in the light of current knowledge, the hypothesis may be imperfectly accurate, and will select tests … sensitive to these possible faults, rather than to others" (p. 50).

However, Fisher is clear that the hypothesis one chooses to test may be suggested by one's data (p. 82). Thus, in tossing a coin, the outcome may lead one to test the hypothesis that the coin is fair, that the trials are independent, or that the same coin was tossed each time. Each test will require a different reference set and a different measure of deviation from the null hypothesis. This point is further illustrated by examples from genetics, where departures from posited 9:3:3:1 Mendelian ratios for a hybrid cross may be due to linkage, partial dominance in one of the factors, linked lethals, or other causes. In such cases, the partitioning of the chi-squared statistic into orthogonal components allows one to pinpoint the source(s) of such a discrepancy (for illustrations of this method, see Mather 1938, chapter 4). This practice is markedly at odds with the Neyman-Pearsonite insistence on predesignating all the elements of a test. Fisher goes on

to draw three more such contrasts between significance testing and the acceptance sampling paradigm that informs the Neyman-Pearsonite theory.

First, in acceptance sampling, the population of lots from which one is sampling is well defined and one has a real sequence of repeated trials, "whereas the only populations that can be referred to in a test of significance have no objective reality, being exclusively the product of the statistician's imagination through the hypothesis which he has decided to test" (*SMSI*, p. 81). Thus, a test is possible where no repetition of one's experiment is contemplated. However, Fisher's hypothetically infinite populations lead a shadowy existence and, as Jaynes (2003) remarks, it is hard to see how such imaginings can confer greater objectivity on one's methods.

Second, decisions are final, and conclusions are provisional. And, third, "in the field of pure research, no assessment of the cost of wrong conclusions … can conceivably be more than a pretence, and in any case … would be inadmissible and irrelevant in judging the state of scientific evidence" (*DE*, pp. 25–26; also see *SMSI*, pp. 106–107). Still, Fisher could easily have admitted the relevance of cost functions to the planning of an experiment and still deny their relevance to the weighing of the evidence that results.

The main thrust of Fisher's critique of the Neyman-Pearsonite theory, however, was to deny that the significance level, which measures the strength of the evidence against the null hypothesis of no difference, can be identified with the frequency with which the null hypothesis is erroneously rejected—with the Neyman-Pearsonite's "type I error probability" (*SMSI*, pp. 93–96). Varying Fisher's more complicated example, J. G. Kalbfleisch and D. A. Sprott (1976, p. 262) consider the composite hypothesis H that at least one of m coins is fair ($m > 1$). Each coin is tossed ten times, and if each shows 0, 1, 9, or 10 heads (with at least one showing 1 or 9), one can quote an exact significance level of $22 \times 2^{-10} = 0.0215$ against the fairness of each coin, hence evidence no stronger than this against H. (Intuitively, the evidence that all the coins are biased can be no stronger than the evidence that any particular one of them is biased.) However, the frequency of rejecting H using this criterion, even when H is "truest" (i.e., when all the coins are fair) is only $.0215^m$, which, even for moderately large m, is much smaller than .0215. This leads Kalbfleisch and Sprott to conclude, with Fisher, that "the frequency with which a true hypothesis would be rejected by a test in repetitions of the experiment will not necessarily be indicative of the strength of the evidence against H" (p. 263). More generally, it may be nearly impossible to obtain strong evidence simultaneously ruling out all the simple constituents of a composite hypothesis (*SMSI*, p. 93), which prompts Fisher to conclude that "the infrequency with which, in particular circumstances, decisive evidence is obtained, should not be confused with the force, or cogency, of such evidence" (p. 96).

Fisher, like Laplace, refrains from imposing a universal critical level of significance and almost always reports exact significance levels or tail area probabilities, but, unlike Laplace, he does speak of rejecting hypotheses, even though in most instances this is just shorthand for "regard the data as discordant or inconsistent with the hypothesis." Nevertheless, this language invited confusion with the different decision theoretic approach of Jerzy Neyman and Egon Pearson, and, in fact, misled generations of textbook writers, who regularly graft the Neyman-Pearson account of testing onto Fisher's and paper over the many and wide differences between them.

Fisher's crucial departure from Laplace is to construe significance levels as evidence against the null hypothesis. Like Karl Popper, he steadfastly refuses to concede that evidence sufficient to reject the null hypothesis at a stringent level of significance is evidence for the alternative hypothesis of interest. However, his own practice belies his precept. In testing for genetic linkage, rejection of the hypothesis of independent assortment is routinely followed by estimation of the recombination fraction, that is, the degree of association. And in the example of the tea-tasting lady, his language is that the lady "makes good her claim" when she classifies all the cups presented to her correctly (*DE*, p. 14). The reason he gives for denying that an experiment can do more than disprove the null hypothesis (p. 16) is that the alternative hypothesis that the lady can discriminate "is ineligible as a null hypothesis to be tested by experiment, because it is inexact." That reason is rather question-begging. The real reason, one suspects, is that Fisher wanted to be able to disprove a null hypothesis without providing evidence for any alternative hypothesis. The possibility of such purely negative significance tests has been at the heart of the controversies that have swirled about this topic (see Royall 1997, chapter 3, especially §3.9).

For Laplace, as it was seen, significance tests are extensions of likelihood ratio tests to rather amorphous ill-defined alternatives. And for Fisher, too, they come into play when the LF is unavailable—a point that seems to have been lost on Neyman and Pearson, whose methodology assumes that outcome probabilities conditional on the alternative hypotheses can be computed

from the model. However, for Fisher, the logic of a test is a probabilistic form of *modus tollens*. A hypothesis is rejected when the outcome it entails does not occur; similarly, it is rejected at a stringent level of significance when an outcome it predicts with high probability does not occur. And this eliminativist logic applies whether or not alternative hypotheses have entered the arena.

Kalbfleisch and Sprott (1976) also strongly insist that the alternative to, say, a null hypothesis of homogeneity may be too amorphous to admit specification. Significance tests allow one to postpone the hard work of formulating such an alternative until a significance test has demonstrated the need for one. No doubt, there are strong arguments on both sides and the issue may be considered unresolved. An interesting case in point is provided by the maximum entropy method wherein the signs and magnitudes of the deviations from expected values indicate a new mean value constraint that then leads to a new maxent distribution. The presence of such an additional constraint is indicated when the entropy of the current maxent distribution lies sufficiently far below the maximum allowed by the current mean value constraints. Ultimately, however, one must agree with Gossett (see Royall 1997, p. 68) that one cannot securely reject a hypothesis or a model unless or until one has a better fitting one to put in its place (compare de Groot 1986, p. 523).

Critics of significance testing have also questioned the use of tail areas, which as Fisher admits, "is not very defensible save as an approximation" (*SMSI*, p. 71), for it appears to make the import of what was observed depend on possible outcomes that were not observed. Actually, in cases where the measure of deviation is a continuous variate, like Pearson's chi square or Gossett's $n^{\frac{1}{2}}(\bar{x} - \mu)/s$, the probability of a deviation exactly as large as that observed is nil and so one has no choice but to use a tail area. However, more to the point, tail areas give (approximately) the proportion of possible outcomes that agree with the hypothesis of cause, design, or efficacy as well as that observed, and this provides a sort of absolute standard of comparison, one that even allows one to compare the strength of the evidence in favor of hypotheses in disparate fields. In any case, the Laplacean logic of significance testing, which views such a test as an index of the evidence in favor of some hypothesis of design, averts a host of interpretive difficulties and fits well with a form of argument—the piling up of improbabilities—that occurs across a broad spectrum of the sciences.

CONCLUSION

No article of reasonable length could hope to touch on more than a fraction of Fisher's vast output and the many thorny issues raised therein. Nothing has been said here, for example, about Fisher's notorious third measure of uncertainty, namely, fiducial probabilities. A good place to start is with the example of Gossett's t-test (*SMSI*, pp. 84–86). Turn next to the critique of the fiducial argument by A. W. F. Edwards (1992, §10.5), and then to the excellent papers by Teddy Seidenfeld (1992) and Sandy L. Zabell (1992). Oscar Kempthorne somewhere remarked that it would require at least ten years of preliminary study before attempting a definitive account of Fisher's work in statistics alone, but the effort would be well repaid. The same may be said of his work in genetics and evolution.

One may view Fisher as a "foiled circuitous wanderer," for his heroic attempts to construct a comprehensive alternative to the Bayesian account of inductive reasoning drew him ever more firmly back into the Bayesian position he started from and then rejected. The question one must address, however, is not whether Fisher would ultimately have returned to the Bayesian fold had he lived, say, another decade, but whether the consistency requirements he endorsed force one "back to Bayes." As it has been seen, his position is close to the objectivist Bayesianism of Laplace, Jeffreys, and Jaynes at many points (see Zabell 1992, p. 381 and notes 42 and 56). At the same time, it has to be admitted that Fisher created almost single-handedly the conceptual framework and technical vocabulary all statisticians, whether Bayesian or non-Bayesian, utilize. For sheer fertility of invention, Fisher has few equals in the history of the mathematical sciences.

See also Information Theory; Statistics, Foundations of.

Bibliography

Cramer, Harald. *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press, 1946.

De Groot, Morris. *Probability and Statistics*. Reading, MA: Addison-Wesley, 1986.

Edwards, A. W. F. *Likelihood*. 2nd ed. Baltimore, MD: Johns Hopkins University Press, 1992.

Hald, Anders. *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley, 1998.

Jaynes, Edwin T. *Probability Theory: The Logic of Science*. Edited by G. Larry Bretthorst. New York: Cambridge University Press, 2003.

Kalbfleisch, J. G., and D. A. Sprott. "On Tests of Significance." In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Vol. 2, edited by W. L.

Harper and C. A. Hooker. Dordrecht, Netherlands: D. Reidel, 1976.

Mather, Kenneth. *The Measurement of Linkage in Heredity*. London: Metheun, 1938.

Royall, Richard. *Statistical Evidence: A Likelihood Paradigm*. London: Chapman-Hall, 1997.

WORKS BY FISHER

Contributions to Mathematical Statistics. New York: Wiley, 1950.

Collected Papers of R. A. Fisher, 5 vols, edited by J. H. Bennett. Adelaide, Australia: University of Adelaide, 1971–1974.

Statistical Inference and Analysis: Selected Correspondence of R. A. Fisher. Edited by J. H. Bennett. Oxford, U.K.: Clarendon Press, 1990.

Statistical Methods, Experimental Design, and Scientific Inference. New York: Oxford University Press, 2003. Reprints of the latest editions of Fisher's three books, *Statistical Methods for Research Workers*, *The Design of Experiments*, and *Statistical Methods and Scientific Inference*. All page references are to this edition.

WORKS ABOUT FISHER

Box, Joan Fisher. *R. A. Fisher: The Life of a Scientist*. New York: Wiley, 1978.

Seidenfeld, Teddy. *Philosophical Problems of Statistical Inference: Learning from R. A. Fisher*. Dordrecht, Netherlands: D. Reidel, 1979.

Seidenfeld, Teddy. "R. A. Fisher's Fiducial Argument." *Statistical Science* 7 (1992): 358–368.

Zabell, Sandy L. "R. A. Fisher and the Fiducial Argument." *Statistical Science* 7 (1992): 369–387.

Roger D. Rosenkrantz (2005)

FISKE, JOHN

(1842–1901)

John Fiske, the American philosopher and advocate of evolutionary theory, was born in Hartford, Connecticut, and baptized Edmund Fisk Green. He changed his name to John Fisk shortly after his mother remarried in 1855 (the *e* was added in 1860). He grew up in Middletown and attended the Congregational Church, but he became dissatisfied with orthodox Christianity and found himself drawn to the philosophical and theological implications of modern science. He early declared himself an "infidel," meaning by the word "non-Christian" rather than atheist. While he was a student at Harvard, he was punished by the college faculty for reading Auguste Comte in church.

Fiske's main philosophical work, *Outlines of Cosmic Philosophy,* developed from lectures given at Harvard in 1869 and 1871, and was completed in London during 1873 and 1874. In it he acknowledged himself a disciple and expositor of the philosophy of Herbert Spencer, the importance of which, he believed, would in time be seen to surpass that of Isaac Newton. This judgment did not appear extravagant to Fiske, since Spencer's law of evolution was "the first generalization concerning the concrete universe as a whole." According to Fiske's formulation of this law, "The integration of matter and concomitant dissipation of motion, which primarily constitutes Evolution, is attended by a continuous change from indefinite, incoherent homogeneity to definite, coherent heterogeneity of structure and function, through successive differentiations and integrations." He illustrated the law's operation at great length with examples drawn from organic processes, the nebular origin of the solar system, comparative philology, and the development of civilization.

Fiske maintained that at some time in the past, human evolution had reached a stage in which man's brain alone continued to evolve; ultimately, a level was achieved at which the individual's brain continued to develop after his birth. This process, which necessitated a period of prolonged infancy accompanied by the evolution of strong parental affection, provided the physical setting for the evolution of the resulting family into clans and society; for the origin of morality in the altruism demanded by family care; and for cultural progress, through the enhanced receptivity of yet developing minds. Prolonged infancy was the cornerstone of an evolutionary explanation of civilization; indeed, Fiske believed that this theory was his most important contribution to philosophy.

Fiske aimed to show the unity of all knowledge, the inevitability of progress, and the ultimate harmony of science and religion. He appealed to the law of evolution to accomplish the first two aims and to "Berkeleian idealism" to accomplish the third. All knowledge is "relative" in the sense that it consists only of classifying and discovering regularities among phenomena. What underlies and creates our experience or phenomena Fiske calls the "Unknowable," "Deity," and "Absolute Power." This "Deity" is the only proper concern of religion, while the regularities discoverable among the phenomena in which Deity manifests itself are the scientist's laws of nature. Thus Fiske's "cosmic theism" reconciles religion and science. Religious dogmas that intrude upon the scientist's world of phenomena are vestiges of primitive, anthropomorphic stages of religious development. Miracles must therefore be rejected, and the doctrine of special creation must give way to Charles Darwin's theory of natural selection. Pantheism, which according to Fiske identifies Deity with the phenomenal world, is rejected, since Fiske's Deity is an "unconditioned existence" which is

"something more than the universe." He rejects Comte's Religion of Humanity as a mere conceit. Materialism is rejected because it is at least conceivable that matter is reducible to mind or feeling, but inconceivable that feeling should evolve from matter; thus, the view that Deity is "Spirit" is plausible.

The major difference between Fiske and Spencer is Fiske's greater emphasis on the religious implications of evolutionary philosophy. Whereas Spencer was guarded, Fiske was unambiguous in calling what lay behind the phenomenal world "Spirit," and he took pains to prove that it was a plausible object of earnest religious contemplation. A further difference between the two thinkers is that Fiske, unlike Spencer, brought evolutionary philosophy to the defense of social conservatism, in the belief that inevitable progress obviated the need for radical social and religious change.

Fiske greatly enjoyed living in Cambridge, Massachusetts, and as late as 1878 he retained the hope of gaining a permanent position at Harvard in either the department of history or that of philosophy. He declined job offers from other universities, but at Harvard he could obtain only temporary positions as a lecturer and as assistant librarian of the college. Early in life Fiske sought to make a living from his writing. Later he was always short of money and tried to make ends meet by going on the lecture circuit; however, he achieved genuine popularity both as a writer and lecturer only in the last decade of his life.

Throughout his life Fiske retained an earnest religious attitude, which he expressed in his later popular lectures in terms were more and more conciliatory toward New England Protestantism. Thus "The Unseen World" (1876), the title essay of his first collection of essays, merely urged that science could not refute the immortality of the soul and that "a simple act of trust" in immortality was not unreasonable. In "The Destiny of Man" (1884), another title essay, Fiske said that the human soul was not merely the end product, but the goal of the great evolutionary process contrived by God. Finally, in *Through Nature to God* (1899), Fiske argued, in reply to T. H. Huxley's Romanes lecture "Evolution and Ethics" (1893), that nature is not morally indifferent but, on the contrary, that evolution "exists purely for the sake of moral ends." He also argued that science offered confirmation of the existence of God and of immortality.

After 1887 Fiske wrote nearly twenty volumes on American history. He was never an original philosopher, but through his clear writing and well-phrased public lectures he helped to advance American religious liberalism. He was a competent popularizer of Darwin's theory of evolution at a time when most religious writers were attacking evolution with frenzy.

See also Comte, Auguste; Cosmos; Darwin, Charles Robert; Darwinism; Evolutionary Theory; Huxley, Thomas Henry; Newton, Isaac.

Bibliography

Fiske's more philosophical works include *Outlines of Cosmic Philosophy; Based on the Doctrine of Evolution, with Criticisms on the Positive Philosophy,* 2 vols. (Boston: Houghton Mifflin, 1874); *The Unseen World, and Other Essays* (Boston: J. R. Osgood, 1876); *Darwinism, and Other Essays* (London and New York: Macmillan, 1879); *The Destiny of Man Viewed in the Light of His Origin* (Boston: Houghton Mifflin, 1884); *Excursions of an Evolutionist* (Boston: Houghton Mifflin, 1884); *The Idea of God as Affected by Modern Knowledge* (Boston: Houghton Mifflin, 1886); *A Century of Science, and Other Essays* (Boston: Houghton Mifflin, 1899); and *Through Nature to God* (Boston: Houghton Mifflin, 1899). Josiah Royce examines Fiske's philosophy in the introduction to the 1902 edition of *Cosmic Philosophy.* Milton Berman's *John Fiske: The Evolution of a Popularizer* (Cambridge, MA: Harvard University Press, 1961) includes an extensive bibliography.

Andrew Oldenquist (1967)

FITNESS

See *Evolutionary Theory; Philosophy of Biology*

FLORENSKII, PAVEL ALEKSANDROVICH

(1882–1937)

Pavel Aleksandrovich Florenskii was a Russian philosopher, theologian, art theoretician, and scientist-polymath. He was born January 22, 1882, of a Russian father and Armenian mother, and died on December 8, 1937. Florenskii lived in the Caucasus, mainly in Tiflis (Tbilissi), Georgia, until 1898, when he entered the Department of Physics and Mathematics of Moscow University. Endowed with many talents, he was invited to stay at the university after graduation for further studies in mathematics, but declined and instead entered Moscow Theological Academy in 1904. By then he was already known as an active member of the Russian Symbolist movement; he published poetry, essays, and philosophical articles, and he corresponded with Andrey Bely—a leading Symbolist poet and theorist—regarding basic

theoretical problems of Symbolism. In 1908 Florenskii graduated from the Academy and stayed there as a professor of philosophy. In 1911 he was ordained a priest. Between 1906 and 1914 he wrote his magnum opus, *The Pillar and Ground of the Truth*, that became one of the principal texts of Russian religious philosophy. It is also a very unusual text, a modernist masterpiece that is at once a theological treatise (bearing the subtitle *An Essay of Orthodox Theodicy*), an exposition of a new philosophical system, a cycle of lyrical letters to a friend (it is divided into "Letters," not chapters, with each letter accompanied by a period engraving with a motto), and an endless chain of digressive studies on all kinds of subjects. Essentially, *An Essay of Orthodox Theodicy* is an itinerary of a spiritual journey; and because the journey is undertaken by a philosopher, it includes the creation of a philosophy.

FLORENSKII'S EARLY PHILOSOPHY: SOPHIOLOGY

The philosophy that is expounded in the book is a system of metaphysics of All-Unity. In Russian philosophy this kind of metaphysics was introduced by Vladimir Solov'ev. Its basic concept represents a specific transrational principle of inner form that ensures a perfect unity of a manifold such that any part of the latter is identical to the whole. The principle is an ancient aporetic philosophical paradigm that originated in pre-Socratics, was later articulated in Neoplatonism, and then carefully elaborated by Nicolas of Cusa. It made its last appearance in Western metaphysics in Schelling's thought. Solov'ev's contribution to this philosophical tradition consisted in making the concept of All-Unity the guiding principle of a comprehensive philosophical system. The main new element introduced by Solov'ev was the linking of All-Unity with the biblical and Gnostic mythologem of Sophia, the Wisdom of God. As a result, Solov'ev's metaphysics of All-Unity was articulated as a metaphysics of Sophia (sophiology). Florenskii accepts the connection between Sophia and All-Unity but otherwise does not follow Solov'ev and hardly ever mentions him. *The Pillar* presents a new, different kind of sophiology.

One may single out two lines in the history of the mythological Sophia, both deriving from the Wisdom books of the Bible. One included Gnostic and later Western mystical doctrines, with Valentine, Heinrich Seuse, Jacob Boehme, and Emmanuel Swedenborg as the chief exponents; whereas the other, found in Eastern Christianity, manifested itself in cultic forms, such as consecration of churches to Sophia and icon-painting. Solov'ev drew upon the Western tradition, whereas Florenskii upon the Eastern one. Further elaboration, turning Wisdom of God into a metaphysical concept, is also independent of Solov'ev. The association of All-Unity with Sophia is based on the fact that, ontologically, they both are intermediate realities between God and the empirical world. Such reality was traditionally conceived as "the world in God" or, according to Greek patristic writings, the system of divine logoi (ideas, designs) of all created things.

Florenskii made this conception of Christian Platonism still more Christian by linking each human person to God's love of this person. This love further coincides with the divine logos or God's idea of this person, and represents a monad of Liebnitzian type; thus there is a noumenal love-idea-monad corresponding to each person and implementing his or her connection to God. Love also connects all these love-idea-monads with one another, and taken together, they form a loving and *eo ipso* living being. Sophia is this noumenal, meta-empirical, living and loving being. Analyzing love, Florenskii finds that it means a certain kind of identity of lovers, their "consubstantiality in God." Thus any two monads belonging to Sophia are consubstantial by virtue of their love, which implies that all parts of Sophia are identical both to one another and to the whole, while at the same time retaining their individual differences. This means in turn that Sophia is the perfect unity of a manifold—that is, All-Unity.

The concept of Sophia as a noumenal loving being and community of monads connected by love forms the basis of the Platonist ontology in Florenskii's early philosophy. Florenskii's epistemology is also Platonist at this stage. The key to the epistemology of *The Pillar* is given in the epigraph of the book, which is a quotation from St. Gregory of Nyssa: "Knowledge is achieved by love." As in Florenskii's ontology, the main principle here is love: Cognition is a kind of communion of the knower and the known, it implies building up their unity and consubstantiality, and this implies, in turn, that genuine cognition is achieved only in love and by love. Such treatment, integrating epistemology into ontology, is opposite to the mainstream of Western metaphysics and especially to Kantian and post-Kantian philosophy that insisted on the primacy of epistemology and subsumed ontology under it. Accentuating this opposition to the extreme, Florenskii depicts the entire history of European thought as a dramatic conflict between Platonism and Kantianism.

FLORENSKII'S LATE PHILOSOPHY: "CONCRETE METAPHYSICS"

The Pillar and Ground of the Truth made Florenskii famous, but it was a milestone rather than an exhaustive treatment of his ideas. In it he tried to follow strictly the Church doctrine and thus had to put aside many themes that were important for him, above all, his ideas of the symbol and its role. Florenskii considered symbol as a constitutive element and building block of reality, and invariably defined his outlook as symbolist. He began to develop this view already in his early texts and returned to it after the publication of *The Pillar*. He conceives and nearly completes a broad project of symbolistic philosophy called *concrete metaphysics*, a kind of all-embracing synthesis resulting in a detailed symbolist picture of Being and the Universe. Originally, symbol was conceived by Florenskii in the classical Platonic and Schellingian way, as an inseparable union of the phenomenal (sensuous) and noumenal (intelligible), joined in perfect mutual expression (the Schellingian *Einbildung*).

In Florenskii's concrete metaphysics the concept of symbol acquired new features—along with his entire worldview that should now be described as Christian Neoplatonism. Firstly, the structure of symbol became layered, as a set of concentric spheres, with the one in the center corresponding to the perfect union of the symbol's phenomenal and noumenal components, whereas the outer spheres represented increasingly imperfect expressions of the noumenon in the phenomenon. Secondly, the inner mechanism of symbol was now seen as a dynamic union of phenomenal and noumenal energies. Energy became the basic new element in Florenskii's late philosophy, which he treated in unwavering Neoplatonist terms.

Concrete Metaphysics was intended to provide a systematic description of reality as formed by symbols of various kinds. In many aspects it resembles Ernst Cassirer's contemporaneous theory of symbolic forms. The description had to consist of studies devoted to definite kinds of symbols. Its basic criteria for distributing symbols into types or classes are anthropological and correspond to human perceptive modalities; the main classes of symbols analyzed by Florenskii are visual (spatial) and acoustic (verbal). The study of visual symbols includes, first of all, a specific model of the Cosmos. In this model, the physical Universe is complemented by a noumenal yet equally spatial world; contacts and transitions between the two worlds include death as well as phenomena of religious and mystical experience, and the boundary between the two worlds is regulated by the cult. Another vast domain of visual symbols is provided by the plastic arts. Florenskii made detailed studies of this domain; he developed a theory of reverse perspective used in icons and then more general theory of space as it figures in works of art; from 1921 to 1924 he expounded these theories in lecture courses in Vkhutemas, one of the main centers of the Russian avant-garde art of 1920s. As for the studies of acoustic symbols, they include mainly Florenskii's philosophy of language, a specific feature of which is the idea of occult energies present in the word. Other parts of Concrete Metaphysics that merit mention are the outlines of the philosophy of technics, based on the idea of the projection of human organs.

Florenskii's late philosophy is presented in numerous works, nearly all of which were created in the decade from 1914 to 1924. Many of the studies planned were not completed. After 1917 the Theological Academy was closed, and Florenskii started working in applied physics and engineering. As he never relinquished his Christian faith, he was persecuted, being arrested in 1928 and again in 1933. After the second arrest he was sent first to the Far East and then, in 1934, to the concentration camp in the Solovetsky islands in the White Sea. In 1937, in the campaign of mass murders, he was shot. Most of his works written after 1917 remained unpublished until the 1980s and 1990s; when they gradually became known, it was discovered that they contain pioneering ideas in many fields—for example, in semiotics—and some bold previsions, such as the existence of the genetic code.

See also Metaphysics; Philosophy of Religion; Russian Philosophy; Sophia.

Bibliography

FLORENSKII'S WORKS

Beyond Vision: Essays on the Perception of Art. Translated by Wendy Salmond. London: Reaktion, 2002.

Detyam moim. Vospominanya proshlykh dnei. Genealogicheskie issledovaniya. Iz solovetskikh pisem. Zaveshchanie. Moscow: Moskovskii rabochii, 1992.

Iconostasis. Translated by Donald Sheehan and Olga Andreev. Crestwood, NY: St. Vladimir Seminary Press, 1966.

The Pillar and Ground of the Truth. Translated by Boris Jakim. Princeton, NJ: Princeton University Press, 1997.

Sobranie sochinenii. Filosofiya kul'ta (opyt pravoslavnoi antropoditsei). Moscow: Mysl', 2004.

Sobranie sochinenii. Stat'i i issledovaniya po istorii i filosofii iskusstva i arkheologii. Moscow: Mysl', 2000.

Sobranie sochinenii v chetyrekh tomakh, Tt.1–4. Moscow: Mysl', 1994–1998.

OTHER WORKS

Franz, Norbert, Michael Hagemeister, and Frank Haney, eds. *Pavel Florenskij—Tradition und Moderne: Beitraege zum Internationalen Symposium an der Universitaet Potsdam, 5. bis 9. April 2000*. New York: Peter Lang, 2001.

Hagemeister, Michael, and Nina Kauchtschischwili, eds. *P. A. Florenskii i kul'tura ego vremeni*. Atti del Convegno Internazionale, Università degli Studi di Bergamo, 10–14 gennaio 1988. Marburg, Germany: Blaue Hoerner Verlag, 1995.

Khoruzhii, Sergei S. *Mirosozertsanie Florenskogo*. Tomsk, Russia: Vodolei, 1999.

Sergey Horujy (2005)

FLORENTINE ACADEMY

"Florentine Academy," or Platonic Academy of Florence, is the name usually applied to the circle of philosophers and other scholars who gathered around Marsilio Ficino, under the auspices of the Medici, in Careggi, near Florence, between 1462 and 1494. These scholars were engaged in the study and discussion of the works of Plato and his followers and of Platonic philosophy. The name "academy" was adopted in memory of Plato. There is no direct link between this Platonic Academy and other academies active in Florence at a later date.

According to Ficino, the academy was founded by Cosimo de' Medici because his enthusiasm for Plato had been aroused by the lectures of Gemistus Pletho at the time of the Council of Florence (1438). In 1462 Cosimo placed a villa in Careggi at the disposal of Ficino, the promising young son of his physician, and lent him several Greek manuscripts of Plato and other ancient philosophers, assigning him the task of studying, translating, and interpreting these writings. This event may be considered the founding of the Florentine Academy. Unlike most later academies, Ficino's Platonic Academy had no formal organization, charter, or fixed membership, and its activities must be inferred from contemporary sources, mainly the letters and other works of Ficino and his associates.

The chief products of the academy are the numerous writings of Ficino and his associates. Whether the public lectures given by Ficino on Plato, Plotinus, and St. Paul were considered part of the work of the academy we do not know. Its activities probably included some regular readings of the Platonic texts and some lectures about them, and surely Ficino gave individual instruction to some of his pupils. On many occasions he addressed edifying sermonlike speeches to his gathered friends and pupils, and this fact, along with a few others, suggests a link between the academy and some of the lay religious associations of the same period. The most famous events of the academy are the banquets celebrated on Plato's reputed birthday, November 7, in 1468 and in 1473, and perhaps in other years. The banquet of 1468 provided the setting for Ficino's commentary on Plato's *Symposium.* The academy also held discussions on philosophical and other subjects on numerous occasions, and it was customary for learned or otherwise distinguished visitors to Florence to attend some of the meetings. The study in which Ficino talked to his pupils contained a painting that represented the globe, with the crying Heraclitus and the laughing Democritus on either side. The often repeated story that Ficino kept an ever-burning lamp before a bust of Plato must be rejected as a legend.

There is no philosophical doctrine common to the Florentine Academy distinct from that of Ficino, but the thought of all its members was influenced to a greater or lesser degree by his teachings. The circle included such philosophers as Giovanni Pico della Mirandola and Francesco Cattani da Diacceto, and such philosophically inclined scholars and poets as Cristoforo Landino, Lorenzo de' Medici, Angelo Poliziano, and Girolamo Benivieni, to mention only some of the better-known members whose writing showed the impact of the academy. The meetings of the academy became famous during its own time, and its intellectual influence, through its visitors and through Ficino's correspondence, spread to the rest of Italy and Europe. Thus, in spite of its informal and fluctuating character, the academy became, and has remained in history and tradition, the most tangible center of Renaissance Platonism.

There is a close link between Ficino's philosophical doctrine and the structure of the academy as he conceived it. Following the model of the ancient philosophical schools, Ficino considered the academy as a community of friends, and his philosophy included an elaborate theory of friendship that he identified with Platonic love. The members of the academy were, he felt, linked with each other and with their master through a "divine" friendship that was based on their common concern with the contemplative life and with the spiritual ascent toward the knowledge and enjoyment of God.

The goal of the academy was philosophical and, in a broader sense, spiritual and cultural rather than political. Although Ficino and the academy were closely identified with four generations of Medici rulers, it cannot be proved that he was their political tool or that his personal and scholarly attachments were limited to their partisans. Nonetheless, although Ficino lived until 1499 and

remained active as a scholar throughout his later years, we hear next to nothing of the activities of the academy after 1494, the year in which the Medici were expelled from Florence, Poliziano and Pico died, and Ficino's published correspondence stopped. There is no direct evidence that the meetings of the academy were discontinued, but it is easy to understand that the illness and old age of its leader, the death of some of its most prominent members, and the political and religious climate that prevailed in Florence after 1494 must have put an end to the academy or at least reduced its activities to a strictly private character.

See also Ficino, Marsilio.

Bibliography

Barzman, Karen-edis. *The Florentine Academy and the Early Modern State: The Discipline of Disegno.* New York: Cambridge University Press, 2000.

De Gaetano, Armand L. *Giambattista Gelli and the Florentine Academy: The Rebellion against Latin.* Florence: L.S. Olschki, 1976.

Della Torre, Arnaldo. *Storia dell'Accademia Platonica di Firenze.* Florence, 1902. This work is still indispensable.

Kristeller, Paul Oskar. "The Platonic Academy of Florence." *Renaissance News* 14 (1961): 147–159. Reprinted in Kristeller's *Renaissance Thought II,* 89–101. New York: Harper and Row, 1965.

Marcel, Raymond. *Marsile Ficin.* Paris, 1958.

Wallach, John R. "The Platonic Academy and Democracy." *Polis* 19 (1–2) (2002): 7–27.

Paul Oskar Kristeller (1967)
Bibliography updated by Tamra Frei (2005)

FLOROVSKII, GEORGII VASIL'EVICH
(1893–1979)

Georgii Vasil'evich Florovskii was a Russian clergyman, theologian, patrologist, and historian of culture. A descendant of several generations of Orthodox priests, Florovskii graduated from the Department of History and Philology of Novorossiyskii University in Odessa (1916), and taught in Odessa until January 1920, when he emigrated to Sofia, Bulgaria. There he became a member of a group of five Russian émigrés who, in 1920 and 1921, founded the so-called Eurasianism. The Eurasian doctrine was first presented in the collective work "Exodus to the East," for which Florovskii wrote three articles.

"Exodus to the East" presented what was essentially a cultural morphology of Oswald Spengler's type (i.e., the botanized view of history as a process of development and interaction of ethnic and cultural organisms) and a geopolitical theory stating that, because Europe had exhausted its spiritual energies, Russia should break with it and cultivate cultural and political ties with Asia. Eurasianists sharply criticized European civilization and argued that in all principal aspects Russia belongs neither to the European nor to the Asian world; rather, it occupies its own continent "Eurasia" and has its own type of culture that borrowed much from the Mongols in the thirteenth to the fifteenth centuries—when Russia was their vassal state. In politics, they propounded the principle of *ideocracy*, very close to that of the one-party rule in Bolshevik Russia.

During the next several years, Eurasianism became popular among young Russian émigrés, evolving from a cultural theory into a political movement that had a pro-Soviet orientation and engaged in secret activities. Florovskii made a significant contribution to the initial form of the Eurasianist doctrine that gravitated toward philosophy of history and philosophy of culture. However, as early as 1923 he began to object strongly to his fellow Eurasianists' growing tendency toward ethnic and geopolitical concerns, favoring instead the opposite orientation toward Orthodox Christianity. This line was rejected by other Eurasianist leaders and, as a result, Florovskii left the movement. In 1928 he published the article "The Eurasian Temptation," in which he presented a profound critical analysis of Eurasianism, and in later years he invariably minimized the scale of the Eurasian episode in his biography.

NEOPATRISTIC SYNTHESIS

From 1922 to 1926 Florovskii lived in Prague, where he received a doctorate in philosophy for his study on Alexander Ivanovich Herzen's philosophy of history and wrote a number of essays on Russian cultural history, including works on Fyodor Dostoyevsky, Fyodor Tytchev, and Mikhail Gershenzon. In September 1926 he moved to Paris to become Professor of Patristics at the St. Sergius of Radonezh Theological Institute that was founded in 1925 and developed quickly into the leading Orthodox theological school. Though a self-taught theologian, Florovskii found in patristics his true calling. He focused on the Greek Fathers of the Church and developed brilliant survey courses, marked by a pioneering presentation of the subject. The teachings of the Greek Fathers also became the cornerstone of the theories that he started to develop in philosophy of culture, theology, and ecclesiology.

At the core of these theories was the concept of *Neopatristic Synthesis*. According to its main thesis, a permanent creative renewal of ties with patristic thought is a necessary condition for Christian culture; in fact, such a renewal defines the mode of the latter's existence. This thesis has manifold implications. Firstly, it is a restatement of the basic tenet of the Orthodox doctrine that establishes a permanent and all-embracing normative role of the Tradition of Fathers. Secondly, it is a polemic reformulation of the claim that Greek philosophy—rather than theology—is an eternal source for all subsequent philosophical thought, so that keeping ties with it is necessary for philosophy of all ages. And thirdly, it is a viable premise for constructing a new philosophy or theology of culture of the archeological type—that is, one based on the permanent generating and productive role of a certain source. Florovskii's theory stated that Greek patristics made Hellenism not simply Christian or Christianized, but "ecclesianised" (*votserkovlenny*)—integrated into the life of the Church with all its mystical and sacramental dimensions. This new transfigured Hellenism replaces the old pagan one and should serve as the generating source of a new "ecclesianised," Christocentric culture. The principle of this culture is a sui generis creative traditionalism, which is open to all contemporary problems and tries to solve them, drawing upon Fathers not for ready answers, but for the mental, cultural, and spiritual attitudes required to meet the challenge.

Seeing in the concept of Neopatristic Synthesis an original and universal criterion, Florovskii applies it to many cultural phenomena. He elaborates a critique of German Idealism and of the mainstream European metaphysical tradition in general, charging them with gnosticism and a mere continuation of the primordial pagan, untransfigured Hellenism. This critique of European thought (likened by himself and others to the philosophy of Charles Renouvier) did not gain much popularity.

By contrast, Florovskii's discussion of the Russian intellectual tradition in the *Ways of Russian Theology* (1937)—which he described as an "attempt at an historical synthesis"—became a widely acknowledged masterpiece. The book is a systematic and enormously erudite exposition of Russia's cultural and spiritual evolution from the fifteenth century until the Bolshevik revolution; many of its sections are of independent value as brilliant critical essays. In his conceptual analyses Florovskii strictly applies the criterion of Neopatristic Synthesis, which renders most of his assessments mercilessly critical. In particular, Russian religious-philosophical renaissance and its main figures—such as Pavel Aleksandrovich Florensky and Nikolai Aleksandrovich Berdyaev—are severely reprimanded. Another such figure, Fr. Sergius Bulgakov, the dean of St. Sergius Institute and author of a controversial teaching about Sophia Divine Wisdom, became Florovskii's target during the so-called "Dispute over Sophia," a heated debate over Bulgakov's teaching in émigré theological circles in the mid-1930s.

THE ECUMENTICAL MOVEMENT

Another important dimension of Florovskii's work was his participation in the Ecumenical Movement. During his Paris period, he was ordained as a priest and took an active part in inter-religious contacts. Beginning with the Edinburgh Conference of 1937, he was a member of various ruling bodies of the Movement, playing a significant role in its formative period and recognized as a leading Orthodox voice in theological discussions. In connection with this activity, he produced a significant number of texts on the Church, its nature and tasks. Taken together, these texts form a self-consistent ecclesiology that eventually became widely known and influential. In September of 1948 Florovskii moved from Paris to the United States to take up the position of Professor of Dogmatic Theology and Patristics at St. Vladimir's Orthodox Theological Seminary in Crestwood, New York. The concluding American period of his biography is chiefly that of brilliant teaching career: at St Vladimir's (until 1955), Columbia University (1951–1956), Harvard University Divinity School (1956–1964), and Princeton University (1964–1972). During these years, his reputation as a theologian and church historian reached its peak. Although Florovskii was not a founder of a school, his numerous disciples include many prominent personalities—and not only from the Orthodox world. He can be considered as the most influential Orthodox theologian of the last decades of the twentieth century.

See also Eurasianism; Philosophy of Religion; Religion; Russian Philosophy.

Bibliography

Blane, Andrew, ed. *Georges Florovsky: Russian Intellectual and Orthodox Churchman*. Crestwood, NY: St. Vladimir's Seminary Press, 1993.

Florovsky, Georges. *Collected Works of Georges Florovsy*, Vols. 1–5. Belmont, MA: Nordland, 1972–1979.

Florovsky, Georges. *Collected Works. Vols. V–XIV*. Vaduz (Lichtenstein), Germany: Academic Books, 1987–1992.

Florovsky, Georges. *Ways of Russian Theology*, edited by Richard S. Haugh, translated by Robert L. Nichols. Belmont, MA: Nordland, 1979.

Horuzhy, Sergey S. "Neo-Patristic Synthesis and Russian Philosophy." *St. Vladimir's Theological Quarterly* 44 (3–4) (2000): 309–328.

Sergey Horujy (2005)

FLUDD, ROBERT
(1574–1637)

Robert Fludd, or Flud, also known as Robertus de Fluctibus, was an English physician, author, and occultist. The son of Sir Thomas Fludd, paymaster to Queen Elizabeth I's forces in France and the Low Countries, Fludd was born at Milgate House, Bearsted, Kent. At the age of seventeen he entered St. John's College, Oxford, then a center of high Anglicanism. After taking his M.A. degree in 1598, Fludd spent some years abroad, studying medicine. On returning to Oxford, he entered Christ Church. He took the degrees of MB and MD in 1605, but had considerable difficulty obtaining from the College of Physicians the right to practice medicine, which was not granted until 1606. It was alleged that he had spoken with contempt of Galen. Nevertheless, he was admitted as a fellow of the College of Physicians in 1609.

As a London doctor Fludd prospered; he was able to provide himself with an amanuensis, to whom he dictated his numerous treatises. His first book, *Apologia Compendiaria Fraternitatem de Rosea Cruce* (1616), was a defense of the ideas of the "Fraternity of the Rosy Cross." About the origins and character of the Rosicrucian Fraternity there is considerable dispute. Although allegedly introduced into Europe in the fifteenth century, Rosicrucian ideas, in fact, derive from two anonymously published tracts written by the Lutheran theologian Johann Valentin Andreä in the early seventeenth century. These, for motives that are somewhat obscure, purported to be of fifteenth-century origin. Whether, even in the seventeenth century, there actually was a Rosicrucian Society as described by Andreä remains a matter of dispute. But these tracts provided a common point of reference for like-minded occultists.

It is impossible to take Fludd seriously as a philosopher; however, he did give expression to a system of ideas that was very influential in the seventeenth century. This can most succinctly be described as an attempt to uphold allegorical interpretation of the Bible, and the established pseudosciences—astrology, chiromancy, alchemy, and sympathetic magic—against the scientific spirit.

Fludd attacked scientific inquiry mainly in its Greek form, as represented in Aristotle and Galen, but certainly with an eye on what was happening around him. His point of departure was St. Paul's attack upon philosophers who try to discover the truth by their own efforts rather than by the interpretation of what God has chosen to reveal. Fludd's criticism of science can be summed up in the familiar phrase: "What is true isn't new; what is new isn't true." He argued that so far as science has any truth in it, it teaches doctrines that careful interpretation will reveal in Genesis (Like Henry More, Fludd was greatly influenced by cabalistic writings). For the most part, however, the teachings of science have to be rejected. Fludd attacked Aristotle's meteorological writings, for example, because Aristotle gives a naturalistic account of lightning and thunder; whereas lightning, according to Fludd, "is a fire burning from the face and presence of Jehovah."

Yet, strangely enough, there is a good deal of contemporary science incorporated into Fludd's work. His contemporaries, he complained, demanded "ocular demonstrations" of divine truths and he used the thermometer—the invention of which is sometimes ascribed to him—and the lodestone for that purpose. Like many of his fellow occultists, Fludd had a passion for diagrams, and some of his optical diagrams remained in physics textbooks up to the twentieth century.

His general approach, however, is cosmogonical, in the manner of the mythmaker, rather than cosmological, in the manner of the scientist. His ideas are most fully presented in *Utriusque Cosmi, Maioris Scilicet et Minoris, Metaphysica, Physica atque Technica Historia* (An account, metaphysical, physical, and technical, of both worlds, greater and lesser), which was published as a series of volumes from 1617 to 1621, and was even then left unfinished. Fludd makes great play with the general concepts of light (heat) and darkness (cold)—hence his interest in optics; rarefaction and condensation—hence the thermometer; sympathy and antipathy—hence the lodestone. His theory can be described in this way: in the beginning God created a void by withdrawing into himself (contraction), and the void appeared as darkness because God is light. Expanding again as light into the void, God created all the substances of the world. Thus, the world we live in is ruled partly by light (God) and partly by darkness (the kingdom of the devil). Since everything is of the same nature—that is, a mixture of light and darkness—there are secret sympathies and secret antipathies everywhere, marked by signs that the adept can discover with God's help. The practice of medicine depends entirely on understanding these forces, as do the practices of chiromancy and astrology.

Fludd's works were published in Latin, and circulated on the Continent, where they attracted a considerable amount of attention. In 1623 Marin Mersenne attacked Fludd as an "evil magician"; and when Fludd replied, Pierre Gassendi, at Mersenne's request, criticized his occultism at length. Fludd also engaged in controversy with Johannes Kepler, who had criticized Fludd in the appendix to his *Harmonice Mundi* (1619).

See also Aristotle; Cosmology; Galen; Gassendi, Pierre; Kabbalah; Kepler, Johannes; Mersenne, Marin; More, Henry; Philosophy of Medicine.

Bibliography

Fludd's works were brought together as *Opera* (Gouda, 1638); *Philosophia Moysaica* appeared posthumously in an English version—Fludd's own—in London (1659); his short alchemical essay "Truth's Golden Harrow" was published in *Ambix,* Vol. III, Nos. 3 and 4 (April 1949), 91–150, with an introduction by C. H. Josten.

James Brown Craven, in *Doctor Robert Fludd, the English Rosicrucian* (Kirkwell, U.K., 1902) summarizes Fludd's works enthusiastically, but without much insight. Craven's book includes details of the Fludd–Mersenne–Gassendi controversy.

Other writings on Fludd include Serge Hutin, *Robert Fludd: Le Rosicrucien* (Paris, 1953) with a bibliography; Walter Pagel, "Religious Motives in the Medical Biology of the Seventeenth-Century," in *Bulletin of the Institute of the History of Medicine* 3 (4) (1935): 265–312; Robert Theodore Gunther, *Early Science in Oxford,* Vol. XI (Oxford, 1937); Sherwood Taylor, "The Origins of the Thermometer," in *Annals of Science* 5 (2) (December 1942): 129–156; Lynn Thorndike, *History of Magic and Experimental Science,* Vols. VII and VIII (New York: Macmillan, 1958). See also Denis Saurat, *Milton: Man and Thinker* (London: Dent, 1944), Ch. 3, pp. 248–267.

John Passmore (1967)

FODOR, JERRY A.

(1935–)

Jerry Fodor is the most significant philosopher of mind in the last fifty years. A student of Hilary Putnam, he joined him, Noam Chomsky, and others at MIT in the early 1960s and became the philosopher most responsible for the "cognitive revolution" that replaced the behaviorism that had dominated much of philosophy and psychology since the 1920s, replacing it with a computational approach derived from the work of Alan Turing. In this way he hoped to provide a basis for a naturalist and realist account of mental processes that rendered them amenable to scientific study. Indeed, he is one of the few philosophers who has combined philosophical and empirical psychological research, publishing work in both domains, and developing at least two theories that have become highly influential in each: a computational/representational theory of thought processes ("CRTT") and a "modularity" theory of perception and linguistic processing.

CRTT is an effort to salvage what Fodor (1975) regards as essential to the familiar "belief/desire," or "(propositional) attitude" psychology with which people routinely explain each other's behavior, as when one explains someone's crossing a road in terms of a desire to meet a colleague. As the name emphasizes, the theory has two parts. According to the computational part, each attitude involves a computationally specifiable relation to a syntactically specifiable representation in a "language of thought" entokened in the agent's brain. For example, *judgment* might be the output of perceptual and reasoning systems that serves as the input to decision making. For the "representational" part, Fodor (1998) argues at length against popular "prototype," "conceptual role" and "holistic" theories of content, and defends instead an "atomistic," "informational," "asymmetric dependency" theory according to which, (i) *ceteris paribus*, tokenings of symbols causally co-vary with phenomena that they thereby mean; and (ii) tokenings caused by phenomena they don't mean *depend* upon (i), but (i) doesn't depend upon them. For example, "horse" means *horse* if (i) it's a *ceteris paribus* law that "horse" tokens are caused by horses, and (ii) nonhorses (e.g., distant cows) causing "horse" tokens depends on horses doing so, but not *vice versa* (1991). Thus, Jones's judgment that horses fly might consist in a sentence, "Horses fly," playing the aforementioned judging role in her brain, where "Horses" and "flies" are each asymmetrically dependent upon the respective phenomena in the world. In this way Fodor hopes to defend intentional realism, in contrast to the widespread eliminativism about the mental, and mere instrumentalism about psychology, that renders psychological ascription a matter of mere "interpretation," such as one finds in the work of Willard Quine, Donald Davidson, Daniel Dennett, and Paul and Patricia Churchland. (Fodor's [1983, 1998] account of representation also leads him to claim that virtually all concepts expressed by single morphemes in natural language are innate, rejecting the empiricism also associated with these figures.)

CRTT is a species of functionalism, or the view, due originally to Putnam, that mental states are to be individuated by their causal relations, for example, to inputs, outputs, external phenomena, and each other, in ways

analogous to the individuation of a program in a computer (1968). Since different physical phenomena can realize these relations, functionalism naturally gives rise to cross-classificatory *layers* of explanation: one level of causal relations may be "multiply realized" by different mechanisms at lower levels (1968, 1975). Specifically, the *intentional level* of a cognitive psychology may be implemented at a lower level by various computational *syntactic* processes, which in turn may be implemented by different physical mechanisms—brains in the case of people, transistors in the case of machines. For this reason, intentional psychology enjoys a considerable "autonomy" from levels of explanation closer to the brain, for example, neurophysiology. However, although the laws and explanations at the intentional level are not *reducible* to laws and explanations at the lower levels, Fodor presumes that they "*supervene*" on them.

One of Fodor's (1986) main arguments for CRTT is that it promises to account for the sensitivity of human beings to indefinitely many properties that are not "transducible" by sense organs, in particular, to arbitrary nonphysical and/or nonlocal properties, such as being a morpheme or a noun phrase, a crumpled shirt, a grieving widow, or a collapsing star. These sensitivities are particularly impressive given that they seem to be (i) *productive* and (ii) *systematic* (1987): that is, people seem capable of discriminating stimuli of indefinite logical complexity, such as *being a crumpled shirt that was worn by the thief who stole the cat that chased the rat* …; and anyone capable of thinking one logical form is capable of thinking logical permutations of it: for example, one can think *John loves Mary* if and only if one can think *Mary loves John* (1987). Fodor (1968, 1987, 1988) argues that non-CRTT accounts, such as behaviorism, Gibsonianism, and purely connectionist accounts are either vacuous or empirically inadequate for this task. What one needs is a system that can exploit internal processes of logical combination, inference and hypothesis confirmation, which presuppose at the least the resources of a CRTT.

However, Fodor (1983) has also been a critic of the "New Look" theories of perception, such as one finds in the work of Jerome Bruner, Thomas Kuhn, and Nelson Goodman, which emphasize how people's background expectations color their perceptions. Against this view, Fodor calls attention to the fact that the very perceptual illusions that New Look theorists prize actually tell against their case: for many of these illusions do not disappear even when one knows better, suggesting, along lines developed by Zenon Pylyshyn, that visual perception occurs in a "cognitively impenetrable module," that is "informationally encapsulated" from the "central" system whereby we reason and fix our beliefs. Fodor argues for a similar view of linguistic perception.

By contrast, the central system is "Quinian" (i.e., computed over the totality of beliefs, as when people settle on a theory that is, for example, simplest and most conservative overall) and "isotropic" (every belief is potentially relevant to the confirmation of every other, as when radio waves confirm the age of the universe). This leads Fodor (1983, 1999) to somewhat pessimistic conclusions regarding the tractability of central reasoning to a Turing-style CRTT, which depends upon exploiting *local* syntactic features of representations. Although CRTT is necessary for an adequate theory of mind, it may not be sufficient.

See also Behaviorism; Chomsky, Noam; Computationalism; Content, Mental; Davidson, Donald; Dennett, Daniel C.; Functionalism; Goodman, Nelson; Kuhn, Thomas; Language of Thought; Mental Representation; Putnam, Hilary; Quine, Willard Van Orman; Turing, Alan.

Bibliography

WORKS BY FODOR

For a full bibliography of Fodor's work up until 1991, see B. Loewer and G. Rey, *Meaning in Mind: Fodor and his Critics* (Oxford: Blackwell, 1991), which also contains a long introduction to his work and critical essays by a number of prominent philosophers and cognitive scientists.

Psychological Explanation. New York: Random House, 1968.

The Language of Thought. New York: Thomas Y. Crowell Co., 1975.

"How Direct is Visual Perception? Some Reflections on Gibson's 'Ecological Approach.'" With Z. Pylyshyn. *Cognition* 9 (1981): 139–96.

Representations: Essays on the Foundations of Cognitive Science. Cambridge, MA: MIT Press, 1981.

The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, MA: MIT Press, 1983.

"Why Paramecia Don't Have Mental Representations." In *Midwest Studies in Philosophy*, vol. 10, edited by P. French, T Uehling, Jr., and H. Wettstein. Minneapolis: University of Minnesotta Press, 1986.

Psychosemantics. Cambridge, MA: MIT Press, 1987.

"Connectionism and Cognitive Architecture." With Z Pylyshyn. *Cognition* 28 (1–2) (1988): 3–71.

A Theory of Content and other Essays. Cambridge, MA: MIT Press, 1990,

Holism: A Shopper's Guide. With E. Lepore. Oxford: Blackwell, 1992.

Concepts: Where Cognitive Psychology Went Wrong, Oxford: Oxford University Press, 1998.

The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology. Cambridge, MA: MIT Press, 2000.

Georges Rey (2005)

FOLK PSYCHOLOGY

Among the more remarkable qualities of human beings is that they describe and explain their own minds and behavior. People are self-explainers and self-understanders. By and large, though not invariably, of course, people's efforts to understand themselves are couched in a familiar language: the language of belief, desire, intention, hope, and so forth—the language of intentional mental states. Perhaps just as remarkable is that people are "mindreaders" (Nichols and Stich 2003). In everyday commerce people attribute—sometimes unself-consciously, sometimes painfully and with great difficulty—to others intentional mental states.

Humans are social creatures—competitive and cooperative—and the practice of attributing intentional mental states is, by and large, the vehicle whereby they come to understand others and others come to understand them, and so, this practice is fundamental to efforts to navigate the social world. One is often able to anticipate or to predict what another will do via the command of what one takes the other person to believe and to desire. Whether the arena is chess or rock-paper-scissors, arms negotiations or freeway driving, the human capacity to characterize others in terms of such intentional mental states is often what determines whether plans succeed or fail. Not surprisingly, this scheme of intentional characterization is applied retrospectively in the explanation of the behavior of others and in the explanation and, often, the justification of one's own behavior to others. So, for example, why did Achilles, after earlier refusing to return to the battle, suddenly rejoin it? One may say that Achilles wants desperately to avenge the death of Patroklos, and he believes that killing Hektor, and so reentering the fray, is the best way to accomplish this aim. In less lofty instances, one can explain why the dog lover ran upon seeing Fido (the dog lover believed the dog was rabid), and why a dear friend refused to return repeated phone calls (the dear friend is angry and wants to stew a bit longer). This commonsense framework of mentalistic understanding, this scheme of intentional description, explanation, and prediction, among many other uses, has come to be termed *folk psychology*.

UNDERSTANDING FOLK PSYCHOLOGY

An intuitively compelling and seductive understanding of the nature of folk psychology might be seen to be offered by René Descartes. At the end of the *Second Meditation*, he famously writes, "I know plainly that I can achieve an easier and more evident perception of my own mind than of anything else" (1988, p. 86). In short, nothing is more easily known by or to the mind than itself. Just by looking within, and by, as we say, introspecting, I can know that I believe certain things to be the case, that I desire this or that to be so, and that I behaved as I did because I believed and desired as I did. This mentalistic characterization, made manifest to me in the first person, is then applied to the characterization of the minds and behavior of others. One upshot of such a view is that the description and understanding of the mental and intentional action are unlike the efforts to understand other natural phenomena.

This view of the human capacity to describe and to understand the mental has proved exceedingly unpopular with both philosophers and psychologists over the past six decades or so. Wilfrid Sellars (1956) pioneered an alternative account of the mentalistic talk of human beings. Sellars tells a story, a *myth*, he terms it, according to which at a time in human prehistory people understood their conspecifics in purely behavioral or observational terms and without appeal to the language of intentional mental states. Then, something of a savant, named "Jones," came to posit unobservable theoretical entities that served to explain the behavior of others. *Belief*, *desire*, and *intention*, are explicitly introduced as theoretical terms to explain why it is that, for example, Tom is behaving as he is. Finally, this theory came to be applied in the first person, to oneself. The point of the myth is not, of course, that it is historically accurate; rather, the point is that people's mentalistic talk needn't be viewed, as it was on the Cartesian model, as picking out entities or states to which people have some special, privileged, access.

This Sellarsian picture has produced a conception of folk psychology that has come to be extremely influential: the "theory-theory" account of folk psychology. In part the result of dominance of functionalist accounts of the mental, the "theory-theory" has it that folk psychology is an empirical theory of mind and behavior (Lewis 1972, Morton 1980, Churchland 1981). And, in this way, folk psychology is "protoscience" (Rudder Baker 1999). *Belief*, *desire*, and *intention* are theoretical terms whose meaning and reference are secured by their place in a vast network

of implicit folk psychological laws. One such law might be stated as follows: If S desires or intends that v, and believes that k is necessary for v, then, ceteris paribus, S tries to bring it about that k. Another such law might be: If S believes that p and believes that if p then q, then, ceteris paribus, S comes to believe that q. (The entry thus far has emphasized intentional mental states, but it should be noted that qualitative states—pains, itches, visual imaginings—fall within the purview of folk psychology.) The human capacity to engage in folk psychologizing points, then, to a constellation of psychological laws that relate behavior, internal states, and stimuli. Belief, desire, intention and the rest of the mentalist vocabulary are theoretical posits of a folk theory of mind.

In an important essay, "What Is Folk Psychology?" (1994) Stephen Stich and Ian Ravenscroft point out that there is a good deal of ambiguity in discussions of folk psychology. They characterize a sense of "folk psychology" according to which that term picks out a theory that is implicit in the everyday talk about the mental. This sense they term an "external account" of folk psychology, because such a view (largely the conception described above) carries with it no commitment to the claim that folk psychology is "an internally represented knowledge structure or body of information" (1994, p. 460). Folk psychology in this sense "ain't in the head" (p. 460), and so is not implicated in an informative account of just how it is that people in fact have the capacities to predict, to explain, or to describe the minds and behavior of their fellows. A second account of folk psychology Stich and Ravenscroft term *internal*. In this sense of "folk psychology," it is an internally represented theory that explains how it is that people predict, describe, and explain in the psychological realm. Stich and Ravenscroft go on to muster powerful arguments for the claim that this distinction has important implications for the eliminativist-vindicationist debate.

THE ELIMINATIVIST CHALLENGE

Historically, many of the chief philosophical issues surrounding folk psychology have been engaged in the effort to characterize the nature and status of folk-psychological explanation. For it is one immediate consequence of the theory-theory that folk psychology might be false in the way that any empirical theory might be false. Vindicationists argue that folk psychology is, in broad terms at least, a correct theory of mind and behavior. Eliminativists argue that folk psychology is plausibly a false theory. As a causal explanatory account of mind and behavior, folk psychology awaits replacement by some nonintentional robustly scientific account of behavior (Churchland 1981, Stich 1983). The theory of mind and behavior implicit in one's everyday talk is just false, the eliminativist alleges.

Thus the eliminativist-vindicationist debate hinges upon the anticipated relationship between folk psychology and scientific psychology/neuroscience. Because both of these aim to explain what is intuitively—though, controversially—the same class of explananda, if people are to regard folk psychology as, by and large, a correct account of human behavior, then they are presumably committed to thinking that the cognitive sciences will, in some way, serve to vindicate the ontology and explanations of folk psychology (Kim 1989).

A notable advocate of this brand of vindicationism, Jerry Fodor (1987), has argued that a scientific psychology will count as vindicating folk psychology just in case it postulates states that (1) are semantically evaluable; (2) have causal powers; and (3) are found to conform to the tacit laws of folk psychology. Each of these has given rise to eliminativist complaint.

Insofar as intentional content figures essentially in folk psychological explanation, it may seem a quick matter to demonstrate that such explanations are not respectable:

(1) The causes of behavior supervene upon the current, internal, physical states of the organism.

(2) Intentional mental content does not supervene upon such states.

(3) The science of psychology is concerned to discover the causes of behavior.

(4) Therefore a causal explanatory psychology will not trade in the intentional idiom.

If this argument were correct, folk psychological explanations would be deeply suspect, because appeal to such explanations would be irrelevant to the causal explanation of behavior. The argument is, however, suspect on many fronts. One might dispute the sense of "behavior" in (1) and with it the notion the respectable explanation must be "individualistic" (Burge 1986). In addition, one might grant that whereas truth-evaluable content is "wide," and so fails to supervene upon internal states of the subject, there is a kind of content, "narrow content," that respects individualist scruples.

Content-based objections such as those above focus upon the puzzling status of intentional properties in a physical universe; many theorists point to the allegedly irreducible nature of intentional mental content as a way

of undermining the integrity of folk psychology (Churchland 1986). Another family of eliminativist worries points to matters structural. It is, for example, claimed that if certain connectionist models of humans' cognitive architecture are correct, then there will literally be no states or events that play the causal role intentional mental states are understood to play in folk psychology. Folk psychology appears committed to the view that mental states are "functionally discrete" internal states with a certain causal profile (Ramsey, Stich, and Garon 1991). Yet, on connectionist models there are no such discrete internal states with the causal roles that belief, desire, and so on are presumed to play in folk psychology.

If these objections are given some taste of the eliminativist assault, they serve as well to highlight an assumption held by many vindicationists and eliminativists alike: folk psychology possesses, in Fodor's terminology, "causal depth" (1987, p. 6). It posits unobservable states and events in aid of the causal explanation of observed phenomena. The explanations of folk psychology are, then, structurally informative insofar as they aim to offer information about the structure of causal relations that hold between behavior, stimuli, and unobservable internal states. Only on such a supposition is it plausible to suggest that folk psychological states and events will go the way of caloric and phlogiston. And this is why many vindicationists hold that the survival of folk psychology demands that there be some scientific level of the description of human cognitive architecture that mirrors the folk psychological one.

Much hinges upon the resolution of this dispute. If the eliminativist is correct, there are no beliefs and desires, and so no intentional actions. It is, for example, just false that human beings often intend to do what they most desire. Nothing would appear to remain of people's conception of themselves as deliberators and actors. While this may strike one as incredible, the eliminativist will insist that this is but another case in which what the folk have taken to be patently obvious turns out to be radically false.

Even so it has been argued that, more than incredible, eliminativism is self-refuting or pragmatically incoherent (Rudder Baker 1987). The charge here is not that the eliminativist thesis is self-contradictory or internally inconsistent. Rather the claim is that there is no perspective from which the doctrine can be coherently put forth. For if eliminativism is true, there are no actions. Yet the eliminativist asserts the truth of eliminativism, and assertion is certainly an action. Moreover, the eliminativist asserts eliminativism because she takes it to be a correct or true thesis, one amply supported by available evidence. But what sense can be made of the notion of justification or even truth without the intentional framework of folk psychology? This argument is sometimes developed in concert with the suggestion that folk psychological principles are not contingent regularities but are, rather, normative principles that are true a priori.

FOLK PSYCHOLOGY STRIKES BACK

Whatever the merits of the foregoing lines of argument, the prima facie oddity that attaches to eliminativism suggests that whereas it is one thing to assert that intentional mental states will not figure in the ontology of some ideal cognitive science, it is another to assert that there are no intentional mental states. In hopes of saving the folk psychological phenomena, an alternative conception of the nature of folk psychology rejects the assumption that folk psychology does offer such informative causal explanations. Rather, folk psychological explanations are silent about the internal mechanisms and processes of cognition and behavior. Because its explanations are not informative in the ways that a cognitive science aims to be informative, folk psychological explanation is not in competition with a scientific psychology, and so folk psychology might be regarded as immunized against scientific advances.

In an extremely influential series of papers, Daniel Dennett (1987) advocates something like this view. According to him, folk psychological explanation and prediction proceeds via the assumption of rationality. When one predicts what an agent will do in various circumstances, the question asked amounts to: What is it rational for her to do, given that she believes and desires as she does? To be, in Dennett's terminology, an "intentional system"—to be such as to have beliefs truly attributable to one—is to be a system whose behavior is so predictable. Folk psychological description, then, does not aim at the description of internal processes and mechanisms. And, whereas an empirically informative cognitive science will reject the intentional idiom, folk psychological explanation is adequate in its own preserve. Even so, it is not easy to see how this brand of instrumentalism about the intentional makes folk psychological description anything more than a façon de parler.

Other philosophers who offer various versions of this approach emphasize that many of the folk explanations that people regard as true bear no easy relationship to science (Chastain 1988, Horgan and Woodward 1985, Horgan and Graham 1991). One may, for example, explain why Ajax slipped by the ramparts by pointing out

that the ground was slimy. In such a case, one is in command of a tacit law to the effect that slimy surfaces are apt to produce slippings. But sliminess and slipperiness are certainly not scientific kinds; it seems likely that no science will make appeal to such kinds. Still, it would be mad to insist that such explanations are false, and that the description of surfaces as slimy is no more than a colorful way of speaking.

Such explanations can survive most any developments in the sciences. People, moreover, are likely to regard the more informative scientific account of the phenomena as a way of spelling out and so vindicating the folk "slimy/slippery" account. With such folk explanations all that is demanded is that there be some more basic account of the properties/processes one characterizes in terms of "sliminess" and "slipperiness." The source of the robustness of such explanations is precisely their relative uninformativenss. Indeed, folk recognize that sliminess and slipperiness do not play any deep or informative role in the causal explanations in which they figure. Rather, their role would appear to be something like the following: There's *something* about the surface picked out as "slimy" that causes events picked out as "slippings." So, just by virtue of their offering scant information about the relevant causal processes, they are insulated from any serious threat of elimination posed by developments in the sciences.

It is, then, urged that we adopt a similar position as to the status of folk psychological explanation. Just as there are slimy things, there are beliefs and desires. And just as it is true that Ajax slipped because the ground was slimy, so it is true that Achilles behaved so because he believed and desired as he did. It should, nonetheless, be noted that this appealing conclusion has been secured at some considerable price: folk psychological explanations, though serviceable for everyday purposes, are about as superficial as causal explanations can be. It is not all apparent that, for example, people's conception of ourselves as reasoners and actors—a conception that appears to implicate certain views as to the nature of mental processes—can withstand so deflationary a reading. One might well conclude that this gives everything to the eliminativist but what she wants.

Finally, Michael Bishop and Stephen Stitch argue that both eliminativists and vindicationists, in developing their arguments, make use of favored theories of reference to establish the conclusion that the terms of folk psychology either do or do not refer, and from this they draw the further conclusion that beliefs do or do not exist. Bishop and Stitch point out that neither the eliminativist nor vindicationist bothers to defend the claim that his or her favored theory of reference is the correct account, one that would sanction a transition from a claim about reference to a claim about existence or nonexistence. The upshot of this argument is that neither eliminativists not vindicationists have a right to make claims about the existence or nonexistence of folk psychological states and entities on the basis of the considerations they adduce.

SIMULATION VERSUS THEORY

In response to the unpalatable alternatives described above (folk psychology is gravely at risk of elimination, or folk psychology is exceedingly unlikely to be eliminated by virtue of its uninformativeness) some have suggested that it is the theory-theory account of folk psychology itself that demands reevaluation. This reevaluation of the nature and status of folk psychology can assume a number of different forms. By far the most influential of these accounts is the simulation account of folk psychology. Jane Heal (1986), Robert Gordon (1986), and Alvin Goldman (1989) have resuscitated the view that people's folk psychological capacities are mediated by the simulation of others. In the effort to understand others, people make adjustments for their cognitive and affective constitutions and, then, using these as inputs, allow their own psychological mechanisms to run "offline." In prediction, simulated beliefs and desires are attributed to the psychological subject of interest.

Advocates of the account claim that simulation is a far simpler and more psychologically plausible account of folk psychologizing. In this way, simulation is, in the language of Sitch and Ravenscroft, a response to an internal theory-theory account. What explains, according to simulationists, the human capacity to describe, explain, and predict the mental states and behavior of others is not an internally represented theory, but rather just the capacity to engage in simulation. In this regard, it is important to note the much of the original impetus behind the development of a competing simulation account of folk psychology was to blunt the force of eliminativist argument. For if psychology is not a theory it cannot be a false theory. So, it seems that on a simulationist account the eliminativist worry cannot be raised. But, as Stich and Ravenscroft (1994) point out, even if human folk psychological capacities may not be subserved by an internally represented theory, it may nonetheless be that eliminativism threatens folk psychology on an external reading.

See also Simulation Theory.

Bibliography

Bishop, Michael, and Stephen Stich. "The Flight to Reference." *Philosophy of Science* 65 (1998): 33–49.

Burge, Tyler. "Individualism and Psychology." *The Philosophical Review* 45 (1986): 3–45.

Chastain, Charles. "Comments on Baker." In *Contents of Thought*, edited by R. Grimm and D. Merrill. Tucson, AZ: University of Arizona Press, 1988.

Churchland, Patricia S. *Neurophilosophy*. Cambridge, MA: MIT Press, 1986.

Churchland, Paul. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78 (1981): 67–90.

Dennett, Daniel. *The Intentional Stance*. Cambridge, MA: Bradford Books, 1987.

Descartes, René. *Descartes: Selected Philosophical Writings*, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge, U.K.: Cambridge University Press, 1988.

Fodor, Jerry. *Psychosemantics*. Cambridge, MA: Bradford/MIT Press, 1987.

Goldman, Alvin. "Interpretation Psychologized." *Mind and Language* 4 (1989): 161–185.

Gordon, Robert. "Folk Psychology as Simulation." *Mind and Language* 1 (1986): 158–171.

Heal Jane. "Replication and Functionalism." In *Language, Mind and Logic*, edited by J. Butterfield. Cambridge, U.K.: Cambridge University Press, 1986.

Horgan, Terrence, and George Graham. "In Defense of Southern Fundamentalism." *Philosophical Studies* 62 (1991): 107–134.

Horgan, Terrence, and James Woodward. "Folk Psychology is Here to Stay." *The Philosophical Review* 44 (1985): 197–226.

Kim, Jaegwon. "Mechanism, Purpose, and Explanatory Exclusion." In *Philosophical Perspectives*, edited by James Tomberlin. Atascadero, CA: Ridgeview Publishing Company, 1989.

Lewis, David. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (1972): 249–258.

Morton, Adam. *Frames of Mind*. Oxford: Oxford University Press, 1980.

Nichols, Shaun, and Stephen Stich. *Mindreading*. Oxford: Clarendon Press, 2003.

Ramsey, William, Stephen Stich, and William Garon. "Connectionism, Eliminativism, and the Future of Folk Psychology." In *The Future of Folk Psychology*, edited by John Greenwood. Cambridge, U.K.: Cambridge University Press, 1991.

Rudder Baker, Lynne. *Explaining Attitudes*. Cambridge, U.K.: Cambridge University Press, 1995.

Rudder Baker, Lynne. *Saving Belief*. Princeton, NJ: Princeton University Press, 1987.

Rudder Baker, Lynne. "What Is This Thing Called 'Folk-Psychology.'" *Philosophical Explorations* 2 (1999): 3–19.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949.

Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." In *Minnesota Studies in the Philosophy of Science*, edited by H. Feigl and M. Scriven. Minneapolis: University of Minnesota Press, 1956.

Stich, Stephen. *From Folk Psychology to Cognitive Science*. Cambridge, MA: Bradford Books/MIT Press: 1983.

Stich, Stephen, and Ian Ravenscroft. "What Is Folk Psychology?" *Cognition* 50 (1994): 447–468.

Dion Scott-Kakures (1996, 2005)

FONSECA, PETER

(1528–1599)

Peter Fonseca, the neo-Scholastic Aristotelian philosopher, was born at Proença-a-Nova, Portugal, and died at Lisbon. He entered the Society of Jesus at the age of twenty, completed philosophical and theological studies in that order, and spent most of his life as a professor of philosophy at Coimbra, where he was the leader of a group of scholars who produced a famous series of textbooks (*Cursus Conimbricensis*). Fonseca has been called the Aristotle of Portugal. His *Institutionum Dialecticarum* (Eight Books on Logic; Lisbon, 1564), was widely used as a textbook throughout Europe, and in 1625 it was in its thirty-fourth printing.

Basically an interpreter of the philosophy of Aristotle, Fonseca corrected the Aristotelian text then in use, using Greek manuscripts, and started the process of improving the Renaissance Latin versions. His logic is the traditional syllogistic which continued to be taught in Europe until J. S. Mill and the nineteenth-century mathematicians broadened the scope of the subject. As a student Fonseca had, of course, been taught a modified form of Thomism, but he showed a great deal of independence on specific questions. In theory of knowledge he maintained that a singular thing is directly known by the human intellect (contrary to Thomas Aquinas), and he seems to have felt (with the later Ockhamists) that the theory of intelligible species as intellectual determinants of the process of conceptualization is useless.

Fonseca placed great emphasis on the unity of the formal concept of being (influencing Francisco Suárez) and taught that this concept is univocal and not analogical in its reference to individual realities. However, he approximated the Thomistic real distinction of essence and existence by treating essence as an ultimate intrinsic mode of the nature of a thing and existence as a contingent addition to this nature. He is, then, partly responsible for the introduction of the terminology of modes into early modern metaphysics. Fonseca abandoned Thomism in denying that matter is pure potency and in rejecting quantified matter as the principle of individuation in bodies. He explained individuation as due to a positive

difference (*differentia*) added to the essence of a thing, a theory reminiscent of John Duns Scotus.

See also Aristotelianism; Aristotle; Being; Duns Scotus, John; Mill, John Stuart; Ockhamism; Scotism; Suárez, Francisco; Thomas Aquinas, St.; Thomism.

Bibliography

WORKS BY FONSECA

Commentariorum in Libros Metaphysicorum Tomi IV (Commentary on Aristotle's Metaphysics). Vols. I and II, Rome, 1577–1589; Vol. III, Evora, 1604; Vol. IV, Lyons, 1612.

Isagoge Philosophica (Introduction to philosophy). Lisbon, 1591.

WORKS ON FONSECA

Ashworth, Earline J. "Petrus Fonseca and Material Implication." *Notre Dame Journal of Formal Logic* 9 (1968): 227–228.

Ashworth, Earline J. "Petrus Fonseca on Objective Concepts and the Analogy of Being." In *Logic and the Workings of the Mind,* edited by Patricia A. Easton. Atascadero, CA: Ridgeview, 1997.

Coombs, Jeffrey. "The Ontological Source of Logical Possibility in Catholic Second Scholasticism." In *The Medieval Heritage in Early Modern Metaphysics and Modal Theory, 1400–1700,* edited by Russell L. Friedman. Dordrecht: Kluwer Academic, 2003.

Giacon, C. *La seconda scolastica.* Vol. II, 31–66. Milan: Fratelli Bocca, 1946.

Gomes, Joaquim F. "Pedro Da-Fonseca, Sixteenth Century Portuguese Philosopher." *International Philosophical Quarterly* 6 (1966): 632–644.

Nedelhofen, M. *Die Logik des Petrus Fonseca.* Bonn, 1916.

Vernon J. Bourke (1967)
Bibliography updated by Tamra Frei (2005)

FONTENELLE, BERNARD LE BOVIER DE
(1657–1757)

Bernard Le Bovier de Fontenelle, the French author, forerunner of the Enlightenment, was born in Rouen and died in Paris, having lived one month short of a century. Schooled by the Jesuits, he also studied law, but soon abandoned the career of advocate to follow in the literary footsteps of his uncles, Pierre and Thomas Corneille. Neither then nor later was he to distinguish himself as a poet or dramatist but, in 1683, with the appearance of the *Dialogues des morts* (*Dialogues of the Dead*), he achieved immediate success as a man of letters. The witty paradoxes and sparkling conversations in these imaginary dialogues of illustrious and notorious figures of the past confirmed the reputation of their twenty-six-year-old author as a seventeenth-century *belesprit*; more important, they revealed him as a singularly independent thinker, skeptical of traditional values and, as such, a potential enemy of seventeenth-century orthodoxy. Judging his literary fame firmly established, Fontenelle turned to the study of mathematics, physics, and astronomy and published *Entretiens sur la pluralité des mondes* (*Conversations on the Plurality of Worlds*, 1686), a brilliantly successful popularization of the Copernican system which, until that time, had achieved very limited acceptance.

The following year his *Histoire des Oracles*, ingeniously adapted from the ponderous Latin of A. van Dale, appeared anonymously. Ostensibly an exposure of imposture and charlatanism in religious practices of pagan antiquity, the work was soon recognized for what it really was: a bold attack on credulity and superstition in all ages. Equally daring was *De l'origine des fables* (The Origin of Fables), composed by Fontenelle before 1680, but fear of persecution invited prudence, and it was not published until 1724. One of the first modern studies in the field of comparative religion, it based early man's belief in the supernatural on his ignorance of natural phenomena. But it was obvious that the criticism was intended to apply equally as well to Christianity and other revealed religions.

The quarrel over the relative literary merits of the Ancients and Moderns had been raging for some years when, in 1688, Fontenelle entered the fray with his *Digression sur les Anciens et les Modernes*, His thesis was that since the question also included the problem of man's progress, the recent accumulation, organization, and dissemination of scientific knowledge proved the superiority of the Moderns. Because of his position in the dispute, entry into the French Academy was denied him on four occasions; and he was not elected a member until 1691.

In 1697 Fontenelle was elected to the Academy of Science, and two years later he became its secretary. His clarity and intelligence, the cool impartiality of his judgment, his wide range of scientific knowledge, and his gift for expression made Fontenelle ideally suited for the post, and he came to be considered as spokesman for his fellow academicians. He contributed a great deal to the widespread popularization of the scientific spirit at home and abroad with his remarkable series of *Éloges* for departed academy members, written over a period of forty years. These essays provided an impressive, constantly renewed picture of accomplishments in science on various fronts, written with the same lucidity and ease of expression that

marked all of Fontenelle's serious writing. They were admirably complemented by the *Histoire de l'Académie royale des sciences* that alone, with its masterful preface and original views, would have assured Fontenelle's reputation throughout eighteenth-century Europe as one of the great historians of the philosophy of science.

In the field of mathematics, Fontenelle was particularly interested in the differential calculus of Isaac Newton and Gottfried Wilhelm Leibniz and the analytical geometry of René Descartes. One of his own mathematical treatises is the *Préface des éléments de la géométrie de l'infini* (Elements of Infinitesimal Calculus; 1727). The last book he wrote was also scientific in nature. Titled *Théorie des tourbillons cartésiens* (The Theory of Cartesian Vortices; 1752), it showed him to be a disciple of Descartes in physics, if not in metaphysics.

Concerning Descartes, Fontenelle said that he should be held in esteem at all times but followed only now and then. Nevertheless, Fontenelle can be considered a Cartesian in two respects. First, his own skepticism was closely bound up with Descartes's principle of methodical doubt. Second, as a stout believer in the purely mechanical philosophy of nature, he found the Cartesian theory of vortices far closer to reality than Newton's laws of attraction, according to which it was necessary to hold that some invisible, seemingly supernatural force operated across vast stretches of space.

Among a number of audaciously conceived, anonymous works on religion and metaphysics ascribed to Fontenelle is the *Traité de la liberté*, which appeared in 1745 together with four other pamphlets under the title of *Nouvelles Libertés de penser*, The work, a few copies of which escaped police seizure, purports to reconcile divine foreknowledge with human free will, but, in fact, casts doubt on the existence of either.

Immediately following Fontenelle's death in January 1757, the general opinion of his accomplishments was summed up by Frédéric-Melchior Grimm: "The philosophic spirit, today so much in evidence, owes its beginnings to M. de Fontenelle" (*Correspondance littéraire*, February 1, 1757).

Although there were serious lapses in Fontenelle's knowledge and, hence, in his scientific judgment, his works nevertheless served as the single most important bond between the philosophico-scientific revolution in progress during his life and the *philosophe* movement just getting under way. He was one of the great forerunners of the French Enlightenment, and no small part of his success in this role lay in the fact that he exploited, as had never been done before, a technique for the popularization of science that was still to have its effects some two centuries later.

See also Cartesianism; Clandestine Philosophical Literature in France; Descartes, René; Enlightenment; Leibniz, Gottfried Wilhelm; Newton, Isaac; Philosophy of Science, History of.

Bibliography

PRINCIPAL WORKS BY FONTENELLE

Oeuvres. 3 vols, edited by G.-B. Depping. Paris, 1818.

Entretiens sur la pluralité des mondes; Digression sur les Anciens et les Modernes, edited by Robert Shackleton. Oxford: Clarendon Press, 1955.

WORKS ON FONTENELLE

Adkins, Gregory M. "When Ideas Matter: The Moral Philosophy of Fontenelle." *Journal of the History of Ideas* 61 (3) (2000): 433–452.

Carré, J.-R. *La philosophie de Fontenelle ou le sourire de la raison*, Paris: Alcan, 1932.

Cohen, Bernard I. "The Eighteenth-Century Origins of the Concept of Scientific Revolution." *Journal of the History of Ideas* 37 (1976): 257–288.

Cosentini, John W. *Fontenelle's Art of Dialogue*, New York: King's Crown Press, 1952.

Kearns, Edward J. *Ideas in Seventeenth Century France*, New York: St. Martin's Press, 1979.

Maigron, Louis. *Fontenelle, l'homme, l'oeuvre, l'influence*, Paris: Librairie Plon, Plon-Nourrit, 1906.

Marsak, Leonard M. "Bernard de Fontenelle, the Idea of Science in the French Enlightenment." *Transactions of the American Philosophical Society*, n.s., 49 (7) (1959): 64.

Riley, Patrick. "The General Will before Rousseau." *Political Theory* 6 (1978): 485–516.

Vartanian, Aram. *Science and Humanism in the French Enlightenment*, Charlottesville, VA: Rookwood Press, 1999.

White, Reginald J. *The Anti-Philosophers: A Study of the Philosophes in Eighteenth-Century France*, London: Macmillan, 1970.

Otis Fellows (1967)
Bibliography updated by Tamra Frei (2005)

FOOT, PHILIPPA
(1920–)

In the last half of the twentieth century, few philosophers figured as prominently and persistently in the central debates of English-speaking moral philosophy as Philippa Foot. Née Philippa Ruth Bosanquet, she was born in 1920 in Owston Ferry, Lincolnshire, in the United Kingdom. She studied for the PPE (philosophy, political science, and economics) at Somerville College, Oxford,

from 1939 to 1942. After receiving an MA in 1947, she became the Sommerville's first philosophy tutorial fellow in 1949 and vice principal in 1967. Moving to the United States, she held positions at Cornell University, the University of California at Berkeley, the Massachusetts Institute of Technology, Princeton University, New York University, and Stanford University. She settled at the University of California at Los Angeles in 1976 and became the first holder of the Gloria and Paul Griffin Chair in Philosophy in 1988, which she held until her retirement in 1991. A founder of Oxfam, she has been instrumental in bringing philosophy to bear on practical issues.

Although her work on such practical topics as abortion and euthanasia has been widely and justly influential, Foot's fundamental contributions are to the foundational questions of moral theory. Her publications in moral theory concentrate on three interlocking themes: the notion that virtue is central to morality, naturalism in ethics, and the place of practical reason in the moral life. These themes are pursued in a set of forcefully argued, original essays, most of which are collected in Foot 2002 and 2003. Foot's thoughts on these topics culminated in her book *Natural Goodness* (2001).

Although these themes are a constant preoccupation of her writings, Foot's positions evolve in significant and unexpected ways. This evolution can usefully be divided into an early, middle, and late period (for an excellent discussion of the first two periods, see Lawrence 1995). In several early papers (notably 1958–1959/2002 and 1961/2002), Foot set herself in opposition to a dominant trend in moral philosophy toward noncognitivism, as represented by the emotivism of Charles L. Stevenson (1947) and the prescriptivism of R. M. Hare (1952). According to these philosophers, evaluative language, and moral language in particular, has a distinctive function or meaning that sets it sharply apart from empirical or factual discourse. On this view, the primary function of a moral utterance is not to describe human actions and choices but rather to express the speaker's attitudes or stances (e.g., emotions or commitments) regarding them. Hence moral judgment is not objective, because it is not answerable to the nature or properties of its subject matter.

Foot strenuously opposed this trend, arguing that the concept of morality concerns what is necessary for human flourishing, and therefore that the truth of moral judgments is fixed by facts about the needs of human beings in relation to one another. This naturalism is intimately linked to Foot's view that "a sound moral philosophy should start from a theory of the virtues and vices" (2002, p. xi). The ultimate standard of choosing and acting well is the natural needs of human beings. And the virtues are those traits that enable us to do so.

This virtue-centered naturalism, which Foot has never abandoned, reaches back to the ethics of Aristotle (1998), and sparked a resurgence of interest in virtue ethics in the last two decades of the twentieth century. Yet her naturalism was in tension with two further views to which she was drawn. If possessing and acting on the virtues is necessary for human flourishing, she thought, then having and acting on the virtues benefits their possessor. But common experience shows that in the case of at least some virtues, notably justice, acting virtuously might not benefit the agent, for justice restricts us from advancing our interests in certain ways. So either justice is not a virtue, or virtues are not necessarily good for us. In that case (as Thrasymachus was made to argue in Plato's *Republic*), we cannot honestly recommend justice as a virtue, and we have to concede that not everyone has reason to act justly.

Foot's initial response (1958–1959/2002) was to take Thrasymachus's challenge seriously, arguing, in effect, that the potential costs of committing injustice, and of being the kind of person who would commit injustice, are too steep to be worth it, that being unjust does not pay. But this response, Foot came to think, rested on a mistaken assumption. Justice is indeed a virtue because of its essential role in human happiness, but the mistake is to think (as Foot had tended to do) that the only way that virtues can serve well-being is by advancing the interests of those who possess them. Justice is concerned with the *common* good. Human life goes badly when individuals are prepared to cheat, lie, and steal. In this way, a deep connection between virtue and human well-being is retained, but it does not follow that every individual who acts contrary to justice disadvantages himself. This recantation (2002, pp. xii–xiii) marks Foot's transition to her middle period.

This reply to Thrasymachus prompted Foot to reconsider an orthodoxy to which she had previously been inclined to subscribe: that "moral judgments give reason for acting to each and every man" (1958–1959/2002, n. 6). One has reason to do something, Foot had argued, only if doing so contributes to one's ends or good. Since acting as justice or loyalty or charity does not necessarily promote my interests or ends, I do not necessarily have a reason to act in these ways. Foot concludes that the allegiance to morality derives not from the authority of practical reason (as followers of

Immanuel Kant (1998) argue) but from contingent attachments and devotions, such as love of the common good and hatred of cruelty. In this sense, Foot argues in a famous essay (1972/2002), moral reasons are "hypothetical," not categorical.

Although this provocative thesis deeply shaped the ensuing philosophical literature on the connection between morality and practical reason, Foot eventually rejected it. This rejection signaled the third period of her work, in which she sets forth an entirely novel conception of practical reason. A vice like injustice is a kind of natural defect, she comes to argue, analogous to the defect in a lioness who neglects her cubs. What makes it a moral defect is that it concerns the will, in a broad sense: the ways in which the individual recognizes and responds to reasons. The virtues are a form of goodness in choosing, that is, in taking certain considerations as reasons for acting and desiring.

This way of linking the concept of the virtues to that of practical reason stands the traditional account on its head. Traditionally, it was supposed that we could develop a robust theory of practical reason independent of an account of virtue, and then we could see how morality measures up by that standard of rationality. This is an error, Foot argues in *Natural Goodness* (2001), for practical rationality is reasoning well in matters of action, and that cannot be specified without a general conception of what it is to function well as a human being. The theory of practical reason thus depends on a naturalistic understanding of virtue and vice.

Whether this understanding can be developed without relying on an unconvincing Aristotelian conception of human function is a disputed question. One major challenge is to spell out the sense in which goodness is natural. Foot recognizes that assertions about what is and is not rational cannot be settled by the methods of the natural sciences. (For reflections on this challenge, see Thompson 1995.) A related challenge is to understand the role that culture plays in morality. Culture is, of course, natural to human beings, but particular cultures obviously shape the content and understanding of morality by their members. It remains to be seen how these points can be accommodated within a contemporary Aristotelian theory.

See also Abortion; Aristotelianism; Aristotle; Ethical Naturalism; Ethics, History of; Euthanasia; Hare, Richard M.; Kant, Immanuel; Metaethics; Plato; Stevenson, Charles L.; Virtue Ethics.

Bibliography

WORKS BY FOOT

"Moral Beliefs" (1958–1959). In her *"Virtues and Vices" and Other Essays in Moral Philosophy*. Oxford, U.K.: Oxford University Press, 2002.

"Goodness and Choice" (1961). In her *"Virtues and Vices" and Other Essays in Moral Philosophy*. Oxford, U.K.: Oxford University Press, 2002.

"Morality as a System of Hypothetical Imperatives" (1972). In her *"Virtues and Vices" and Other Essays in Moral Philosophy*. Oxford, U.K.: Oxford University Press, 2002.

Natural Goodness. Oxford, U.K.: Clarendon Press, 2001.

"Virtues and Vices" and Other Essays in Moral Philosophy. Oxford, U.K.: Oxford University Press, 2002.

Moral Dilemmas and Other Topics in Moral Philosophy. Oxford: Oxford University Press, 2003.

OTHER WORKS

Aristotle. *Nicomachean Ethics*. Translated by David Ross; revised by J. L. Ackrill and J. O. Urmson. Oxford, U.K.: Oxford University Press, 1998.

Crisp, Roger, and Michael Slote, eds. *Virtue Ethics*. Oxford, U.K.: Oxford University Press, 1997.

Darwall, Stephen, ed. *Virtue Ethics*. Oxford, U.K.: Blackwell Publishing, 2003.

Darwall, Stephen, Allan Gibbard, and Peter Railton. *Moral Discourse and Practice: Some Philosophical Approaches*. Oxford, U.K.: Oxford University Press, 1997.

Hare, R. M. *The Language of Morals*. Oxford, U.K.: Oxford University Press, 1952.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford, U.K.: Oxford University Press, 1999.

Hursthouse, Rosalind, Galvin Lawrence, and Warren Quinn, eds. 1995. *Virtues and Reasons: Philippa Foot and Moral Theory*. Oxford, U.K.: Oxford University Press.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Translated by Allen W. Wood. New Haven, CT: Yale University Press, 2002.

Lawrence, Gavin. "The Rationality of Morality." In *Virtues and Reasons: Philippa Foot and Moral Theory*, edited by Rosalind Hursthouse, Galvin Lawrence, and Warren Quinn. Oxford, U.K.: Oxford University Press, 1995.

MacIntyre, Alasdair. *After Virtue*. South Bend, IN: University of Notre Dame Press, 1985.

McDowell, John. "Two Sorts of Naturalism." In *Virtues and Reasons: Philippa Foot and Moral Theory*, edited by Rosalind Hursthouse, Galvin Lawrence, and Warren Quinn. Oxford, U.K.: Oxford University Press, 1995.

Stevenson, Charles L. *Ethics and Language*. New Haven, CT: Yale University Press, 1947.

Thompson, Michael. "The Representation of Life." In *Virtues and Reasons: Philippa Foot and Moral Theory*, edited by Rosalind Hursthouse, Galvin Lawrence, and Warren Quinn. Oxford, U.K.: Oxford University Press, 1995.

Gary Watson (2005)

FORCE

In the most general sense, force denotes the faculty of action or the power to overcome a resistance. In the physical sciences it is that entity that changes, or tends to change, the state of rest or of motion of a body. Consequently, it may also be defined as the cause of motion, or more precisely—assuming the validity of the principle of inertia, according to which unaccelerated motion and rest are dynamically and causally equivalent and correspond merely to different choices of the reference systems—as the cause of acceleration.

The metric unit of force in science is the *dyne,* which is the force necessary in order to give a mass of one gram an acceleration (increase of velocity) of one centimeter per second in each second. The British unit of force is the *poundal,* which is the force necessary to give a mass of one pound an acceleration of one foot per second each second. The practical unit is the *gram force,* that is, the force Earth exerts on one gram of mass at sea level and 45° latitude; it equals 980.616 dynes. Another common unit is the *newton,* which is the force necessary in order to give a mass of one kilogram an acceleration of one meter per second each second, and is therefore equivalent to 10^5 dynes.

Apart from being used in a figurative sense, such as "force of habit," "police force," or "economic forces," the word *force,* especially in the natural philosophy of the eighteenth and nineteenth centuries and in the early writings on the principle of conservation of energy (R. Mayer, H. von Helmholtz) signified action and energy. This homonymic use caused considerable confusion at the time.

Originally taken as an analogy to human will power, muscular effort, and spiritual influence, the concept early became projected into inanimate objects and played an important role in ancient thaumaturgy, occultism, and medieval sorcery.

CONCEPT OF FORCE IN ANCIENT PHILOSOPHY

The early Greek hylozoism of the Milesian school (Thales, Anaximander, Anaximenes) conceived nature as a living, animated, and self-moving being, and consequently did not see a problem in the origin of motion. The concept of force gained prominence only with Heraclitus's doctrine of opposing tensions, according to which force is a primary constituent of physical reality and a regulative element in the universe. In Empedocles's philosophy of love (*philia*) and strife (*neikos*), forces, although still conceived in analogy to human affections, became efficient causes of change and motion. In spite of the fact that Plato's natural philosophy relegates the principle of motion ultimately to the existence of a world soul and corresponds in this respect to early hylozoism rather than to the dynamistic teachings of Empedocles and Anaxagoras, the term *dynamis,* signifying not only transitive activity but also passive susceptibility or receptibility, plays an important role in his doctrine. Although Aristotle, in his conception of nature as "physis," still recognized the Platonic notion of force as something inherent in matter, in *De Caelo* he also approached the formulation of a more mechanical conception of force as a physical emanation from one substance to another: through push and pull, bodies affect each other and generate motion in extraneous objects. This Aristotelian notion of emanating kinematic effects, although restricted to contiguous modes of action, is the first instance of the modern dynamical conception of force. In his *Physics* Aristotle subjected this cause of compulsory motion to a quantitative investigation: A force A that moves a mobile B through a distance D during the time T could move half the mobile ($\frac{1}{2}B$) through twice the distance ($2D$) during the same time (T), or could move half the mobile ($\frac{1}{2}B$) through the distance D during half the time ($\frac{1}{2}T$), and so forth. In modern terminology Aristotle's dynamical law of motion may be stated as follows: The velocity of a mobile is proportional to the ratio of the motive force and the resistance of the medium. Nowhere did Aristotle employ units in which these quantities were to be measured. Although it is fairly obvious that forces were practically measured in terms of weight (the early use of the balance is an evidence of this), Aristotle's conception of weight as a manifestation of natural motion and not as a cause of compulsory motion precluded, on theoretical grounds, the possibility of using the units of weight as units of force. Since, according to Aristotle, contiguity between the motor and the mobile was an indispensable prerequisite for the occurrence of dynamical action, force as an action at a distance had no place in his conceptual scheme. Hence, an explanation of planetary motion required the assumption of an external agent or astral intelligence as a "motor" attached to the star, unless the star was thought to be endowed with a life of its own.

With Posidonius's investigations at Gades of the connection between the tides and the movements of the sun and the moon and his doctrine of a universal tension, the concept of force was generalized as something able to pervade all space. Stoic philosophy thus abandoned the Aristotelian restriction of an immediate linkage between the mover and the moved, and conceived force as a

mutual correspondence of action between objects, even when the objects were separated in space. In fact, the Stoics were probably the first to formulate the idea of a field of forces and to regard the universe as a vast system ruled by the interaction of forces.

MEDIEVAL PHILOSOPHY

Arabian and Christian medieval philosophy, in general, adhered to the Aristotelian conception of force. The exceptions were mostly inspired by Neoplatonic ideas. Thus, in Abū-Yūsuf Yaʿqūb ibn Ishāq al-Kindī's treatise *On the Tides* (*Fi-l-madd wal-jazr*), his notion of force is wholly Aristotelian except that he holds that force can be propagated by means of optical rays, a theory conducive to astrological exploitation. Roger Bacon's conception of forces as "species"—isolated entities, detached from their subject and spreading through space in accordance with specific laws of propagation—showed similar features.

The Aristotelian law of motion, already criticized by John Philoponus in the sixth century CE and by Avempace in the twelfth century, was shown by Thomas Bradwardine in the middle of the fourteenth century to contradict experience in the case of equality between the motive force and the resistance, so that the ratio is one but the velocity zero. Bradwardine consequently modified the law, claiming that the velocity, in modern terms, depends on the logarithm of the ratio between motive force and resistance.

In the fourteenth century the Stoic conception of a field of forces was also revived, probably independently of the ancient school. In his *Quaestiones Super Libris Quattuor de Caelo et Mundo* (Questions on the four books of the heaven and the Earth) John Buridan postulates a celestial force that permeates all space and exerts its influence on physical bodies, in contrast to the Peripatetic dictum, *Causa agens est simul cum suo effectu proximo et immediato.* However, the revolutions of celestial bodies, according to Buridan, are not the result of a constant activity of special intelligences, but rather of an original rotational impetus communicated to these bodies by the Creator at the beginning of time.

KEPLER

A decisive stage in the development of the concept of force was reached in Johannes Kepler's search for a quantitative determination of dynamic activity. In his early writings, such as the *Mysterium Cosmographicum* (1596), Kepler still refers to force as a soul animating the celestial bodies. His correspondence, however, and particularly his letters addressed to David Fabricius, show clearly that his use of the term *anima* ("soul") in his writings was merely a metaphor to express the immateriality of the principle that governs the mutual movements of celestial bodies. In 1605 Kepler was already convinced that the force of attraction could be subjected to a mathematical formalism. In the third part of his *Astronomia Nova* (1609), Kepler discusses the causes of planetary motion and insists for the first time on a mathematical definition of force, even if it is not a push or pull. "For we see that these motions take place in space and time and this virtue emanates and diffuses through the space of the universe, which are all mathematical conceptions. From this it follows that this virtue is subject also to other mathematical necessities." Having discovered that the planets move in their orbits with velocities that vary with the distance from the sun, Kepler inquired into the physical cause of this mathematical relation and was thus led to assume the existence of a regulative force whose magnitude decreases with the distance. However, attraction was not yet seen as a radial force, but rather as a tangential drag, and Kepler, under the influence of William Gilbert's *De Magnete* (1600), suggested an analogy with magnetism. But in spite of this, Kepler's conception of a gravitational force of attraction is a typical example of the fact that the existence of forces is, and has to be, inferred from the phenomenological aspects of regularities in the variations of motion. It also exemplifies the fact that the postulation of forces as causes of motions and their kinematic variations is a methodological process that finds its philosophical justification in the reduction of numerous cases of functional dependence to one single agency. Kepler's procedure thus became the prototype for the introduction of forces in the various branches of physics: gravitational, elastic, electromagnetic, nuclear forces, and so forth.

NEWTON

Isaac Newton's conception of force can be traced to two originally disparate classes of mechanical or dynamical phenomena which, however, finally found their logical unification in his *Principia* (1687), through its very definitions of force and mass. Documentary evidence seems to show that his earliest conception of force originated from the study of impact phenomena. Thus Newton's "Waste Book 1664" (Ms. Add. 4004, Portsmouth Collection, University Library, Cambridge, U.K.) starts with a definition of the quantity of motion of a body as the product of its "quantity" (mass) and its velocity, and continues: "Hence it appeares how & why amongst bodys moved some require a more potent or efficacious cause others a lesse to hinder or helpe their velocity. And ye

power of this cause is usually called force. And as this cause useth or applyeth its power or force to hinder or change ye perseverance of bodys in theire state, it is said to Indeavour to change their perseverance." In another document (Ms. Add. 3965, Portsmouth Collection), force is implicitly defined by the statement: "The alteration of motion is ever proportional to ye force by wch it is altered." Considering the exact text of Newton's second law of motion in the *Principia,* "The change of motion is proportional to the motive force impressed," one is led to the conclusion that "force" in these statements denotes more or less what we mean today by "impulse" (which, in fact, is equal to the change of momentum). Newton's original conception of force was consequently that of a thrust, a kick, or a push, as exhibited in collision phenomena, which at that time were the subject of extensive studies by Galileo Galilei, Marcus Marci, John Wallis, and Christian Huygens. On the other hand, in his search for a derivation of the phenomenological aspects of planetary motions from the hypothesis of an inverse-square law, Newton needed the time rate of change of momentum as the primitive notion, and thus identified the change of momentum with its rate of change for astronomical applications. Later commentators, therefore, interpreted Newton as stating that force is measured by the product of mass and acceleration, a product that for constant mass equals the time rate of change of momentum. Thus, although not rigorously impeccable, Newton's definition of force led to a unified treatment of terrestrial and celestial mechanics, and the notion of force became a fundamental concept of physics. Whereas Newton's first law of motion or law of inertia, according to which every body, unaffected by a force, persists in a state of rest or of uniform motion, may be regarded as a qualitative definition of force (namely, as change of state of motion), the second law quantified the concept and provided a meaning for the notion of mass. The Newtonian characterization of force is completed with the third law, which states, in essence, that every force manifests itself invariably in a dual aspect: It has a mirror-image twin. For it claims that if *A* acts on *B,* then *B* acts on *A* with equal magnitude in the opposite direction; or in other words, to every action there is always opposed an equal reaction. Forces, consequently, arise only as the result of a combined interaction of at least two entities. In a universe composed of only one body, no forces are conceivable.

Having thus explored the quantitative aspects of force, and of gravitational force in particular, Newton does not specify the metaphysical nature of force; as far as physical science is concerned, force is an ultimately irreducible notion. Newton's contribution may thus be regarded as the culmination of a conceptual development in a search for a quantitative determination of an otherwise obscure and indiscernible, yet necessary, notion—a development whose philosophical necessity had already been stressed by Bacon, Thomas Hobbes, and even René Descartes.

The scientific legitimacy of a force such as gravitation, which could act at a distance without the intermediacy of an intervening medium, was early called into question. Newton himself, particularly in his *Opticks* (1704), referred to certain speculations, primarily to the notion of an ether, in order to reduce such actions at a distance to contiguous effects compatible with the corpuscular-kinetic theory prevalent at that and later times. Yet, in spite of early opposition (as voiced particularly by Gottfried Wilhelm Leibniz, who rejected action at a distance as a scholastic obscure quality), the notion of force as conceived by Newton became the basic concept of classical theoretical mechanics. Pierre de Laplace, in his *Mécanique céleste* (1799–1805), considered the reduction of all mechanical phenomena to forces acting at a distance as the ultimate objective of the physical sciences, and Joseph Louis Lagrange's *Mécanique analytique* (1788), the highlight of classical mechanics, was written in the same spirit. The mechanics of action at a distance gained further support through its successful applications by Laplace, Siméon Denis Poisson, and Wilhelm Weber in the classical theories of electricity and magnetism. Even capillary phenomena—contact phenomena par excellence—were treated by Laplace and Karl Gauss as subject to actions at a distance.

CRITICISM OF ACTION AT A DISTANCE

The great mathematical success of these theories of force as an action at a distance did not suppress doubts as to the philosophical legitimacy of such conceptions, and alternative mechanistic or kinetic-corpuscular theories, especially for gravitation, were proposed in great number. One of the earliest attempts in this direction, George Louis Lesage's theory of "ultramundane particles" (1747), was typical of similar hypotheses that gained great popularity in the nineteenth century. Particles were assumed to move in all directions through space and to be rebounded by macroscopic bodies; the resulting screening effects were supposed to produce the mutual "attractions" of "gravitating" bodies. The main criticism of the Newtonian conception of force from the philosophical point of view, however, was directed against the hypostatization of force as a metaphysical entity of an autonomous ontolog-

ical status. George Berkeley, in his *De Motu* (On motion; 1721) opposed this approach and viewed the notion of force as a convenient auxiliary fiction with which to work; for the notion had the same status in science as the concept of epicycle has in astronomy. Such terms as *force, gravity,* and *attraction,* he admitted, are convenient for purposes of reasoning or computation; for an understanding of the nature of motion itself, however, Berkeley regards them as wholly irrelevant. They should not lead us to the fallacy that they could throw any light on the real efficient causes of motion, for the only objective of physical science is the establishment of the regularities and uniformities of natural phenomena; to account for particular phenomena means "reducing them under, and shewing their conformity to, such general rules" (*Siris,* 1744). David Hume, Pierre de Maupertuis, and especially the early proponents of modern positivism (Gustav Kirchhoff, Heinrich Hertz, Ernst Mach) followed Berkeley in asserting that force is merely a construct in the conceptual scheme of physics and that it should not be confounded with metaphysical causality. Most radical in this respect was Mach's antimetaphysical attitude, in accordance with which he tried to divest mechanics of all conceptions of cause and force and to adopt a purely functional point of view. Following Kirchhoff's *Lectures on Mechanics* (*Vorlesungen über Meckanik,* 1874–1876), Mach, in his *Science of Mechanics* (*Die Mechanik in ihrer Entwicklung, historisch-kritisch dargestellt,* 1883), identified force with the product of mass and acceleration and thus reduced it to a purely mathematical expression relating certain measurements of space and time.

But even after this process of purification and divestment of all causal or teleological implications, the concept of force was not eliminated from the conceptual scheme of physics. Its methodological justification lies in the fact that it enables us to discuss the general laws of motion irrespective of the particular physical situation with which these motions are associated. In contemporary physics the concept plays somewhat the same role as does the middle term in the traditional syllogism; it is a methodological intermediate in terms of which we can study the kinematical behavior of a physical body independent of the particular configuration in which it is found.

PSYCHOLOGICAL ORIGINS OF THE CONCEPT OF FORCE

The advancement of the critical attitude toward the concept of force, initiated by Berkeley and Hume and culminating in the logical and metaphysical point of view held by Kirchhoff and Mach, brought in its wake a study of the psychological origin of the notion. The first to deal at length with this problem was Thomas Reid, Hume's immediate successor and founder of the Scottish school. He derived the concept of force from the consciousness we have of the operations of our own mind, and especially from the consciousness of our voluntary exertions in producing effects. Reid concluded that if we were not conscious of such exertions, we would not have formed any conception of force and consequently would not have projected this notion into nature and the changes in it that we observe. Immanuel Kant's younger contemporary, Maine de Biran, whose personalistic philosophy has many points in common with Reid's empirical intuitionalism, considered our own will as the source of the notion of force; in his view, the resistance to muscular effort felt in the case of voluntary activity makes us aware that certain actions are not involuntary acts, but the results of our ego as a source of force. From the twofold nature of the ego as an individual source of action and as inseparably united to a resisting organism, we acquire the universal and necessary notion of force. While the Berkeley-Hume criticism led almost to the exclusion of the concept of force from science and natural philosophy, at the same time it supplied to the more psychologically and physiologically oriented philosophy important arguments to oppose such elimination. For it was claimed that the concept of force stands in the same relation to the sensation of muscular effort as the concept of motion to visual perception, and science without the concept of motion is inconceivable. Moreover, if one kind of sensation is to be preferred to the others, it should certainly be muscular sensation, the nearest to the psychological experience of volition. Even William James, who, in "The Feeling of Effort," in *Collected Essays and Reviews* (1920), rejected the so-called feeling of innervation and opposed the view that the resistance to our muscular effort is the only sense that brings us into close contact with reality, contended that reality reveals itself in the form of a force like the force of effort we exert ourselves. The concept of force, according to James, thus remains "one of those universal ideas which belong of necessity to the intellectual furniture of every human mind."

See also al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq; Anaximander; Anaximenes; Aristotle; Bacon, Roger; Berkeley, George; Bradwardine, Thomas; Buridan, John; Descartes, René; Empedocles; Energy; Galileo Galilei; Heraclitus of Ephesus; Hertz, Heinrich Rudolf; Hobbes, Thomas; Hume, David; Ibn Bājja; James, William; Kant, Immanuel; Kepler, Johannes; Kirchhoff,

Gustav Robert; Laplace, Pierre Simon de; Laws, Scientific; Leibniz, Gottfried Wilhelm; Mach, Ernst; Maine de Biran; Mass; Matter; Medieval Philosophy; Neoplatonism; Newtonian Mechanics and Mechanical Explanation; Newton, Isaac; Panpsychism; Philoponus, John; Plato; Power; Reid, Thomas; Stoicism; Thales of Miletus.

Bibliography

Ellis, Brian D. "Newton's Concept of Motive Force." *Journal of the History of Ideas* 23 (1962): 273–278.

Hesse, Mary B. *Forces and Fields.* London: Nelson, 1961.

James, William. "The Feeling of Effort." In *Collected Essays and Reviews.* New York: Longman, 1920.

Jammer, Max. *Concepts of Force (A Study in the Foundations of Dynamics).* Cambridge, MA: Harvard University Press, 1957; reprinted in paperback, New York, 1962.

Mach, Ernst. *Die Mechanik in Ihrer Entwicklung, historisch-kritisch dargestellt.* Leipzig, 1883. Translated by T. J. McCormack as *Science of Mechanics.* La Salle, IL: Open Court, 1960. With introduction by Karl Menger.

Margenau, Henry. *The Nature of Physical Reality.* New York: McGraw-Hill, 1950. Ch. 12.

Röhr, Julius. *Der Okkulte Kraftbegriff im Altertum.* Leipzig, 1923.

Weyl, Hermann. *Philosophy of Mathematics and Natural Science.* Princeton, NJ: Princeton University Press, 1949. Pp. 148ff.

M. Jammer (1967)

FORCE [ADDENDUM]

Forces, understood as pushes or pulls that are exerted (in the first instance) by particulars and that cause motions, have received little philosophical attention in recent decades, reflecting both that forces no longer play a role in fundamental physical theory and that even where they do play a role (e.g., in Newtonian mechanics), it has seemed advisable (following Jammer, above) to give them a purely instrumentalist interpretation. What attention has been paid however indicates that various aspects of the notion of force (or notions; see below) deserve further philosophical consideration.

One such aspect concerns the ontological status of forces. Jammer's deflationary account of force as a mere "methodological intermediate," enabling the kinematical behavior of particulars to be studied independent of the details of specific configurations, but not to be taken with ontological seriousness, was motivated by traditional empiricist concerns with forces as purely theoretical entities (of the sort that exercised Berkeley); such concerns also figure in van Fraassen's instrumentalist agnosticism about forces. Another source of concern about forces lies in the redundancy argument (of which Mill was an early proponent), according to which forces are not needed to explain motions (the usual non-force causes and effects being sufficient unto the task) and hence should (by Ockham's razor) be eliminated.

There are however ways of resisting or responding to such concerns. Hesse rejects Jammer's instrumentalism as inappropriately eliminating "the metaphysical, a priori, intuitive and anthropomorphic elements" of the classical notion of force. More straightforwardly one can deny that forces are purely theoretical on grounds that these are experienced in the course of ordinary events (of, for instance, liftings, pushings), in which case instrumentalist concerns with force are misguided. And in response to the redundancy argument, Bigelow et al. note that the appropriate application of Ockham's razor involves a ceteris paribus clause: Other things being equal, forces should be eliminated. But, they argue, other things are not equal: In particular, physics without forces does not explanatorily unify phenomena (in particular, motions) as well as does physics with forces. Indeed, one might maintain that, even if other entities unify motions, so long as forces unify these in a distinctive fashion (as they appear to do) Ockham's razor can be resisted.

It remains the case that forces do not play the role in contemporary physics that they once were thought to do. Even within the domain of classical (slow-moving, non-quantum) entities, Newton's force-based formulation of mechanics has been superseded for most explanatory and practical purposes by energy-based (e.g., Lagrangian and Hamiltonian) formulations. And while forces and Newton's laws (the third law being understood as a statement of conservation of momentum) are recognizably present in the relativistic extension of Newtonian mechanics, quantum indeterminacy appears to prevent Newton's theory (which presupposes that bodies have a determinate position and momentum) from being extended to treat quantum phenomena. (This is so, assuming the incorrectness of Bohm's deterministic, force-based interpretation of quantum theory, developed in Bohm and Hiley 1993, on which indeterminacy is given an epistemological spin, as uncertainty.) By way of contrast the concepts and operative principles of energy-based theories (energy, Hamilton's principle of stationary action) straightforwardly extend to both quantum and relativistic contexts. Moreover, in the General Theory of Relativity (GTR), the concept of force disappears altogether:

geometry plus inertial motion, rather than forces, guide motions due to gravity.

Upon closer examination however the above considerations do not show that the concept of force is ontologically obsolete. Concerning the classical domain: Force-based and energy-based formulations of mechanics are not only compatible but are also interderivable (under assumptions generally in place); as Feynman notes, Newtonian and Lagrangian dynamics are "exactly equivalent." This equivalence reflects, among other things, the fact that both potential and kinetic energies are initially defined in terms of the work done by a force; more generally, it appears that force-based and energy-based mechanics are, from a theoretical point of view, mutually supporting, compatible perspectives on the same phenomena. (Such a take is reflected in an intuitive ontological conception of the relation between forces and potentials or potential energies, according to which the latter are dispositions of which forces are the manifestations.) Moreover the restricted application of Newtonian mechanics needn't imply that forces don't exist—at least supposing that the similarly restricted application of special sciences such as chemistry and biology doesn't impugn the existence of their subject matters.

The question remains whether the posit of force is compatible with more fundamental theories. As mentioned, quantum indeterminacy poses a barrier to taking forces, as traditionally conceived, to exist at the quantum level; but if forces are special science entities, this is no surprise (plants don't exist at the quantum level, either). Compatibility might rather be indicated by noting that the deep connection between forces and energies persists in quantum theory, albeit at an analogical level; as Jammer says, "No one has ever directly demonstrated the force of attraction between, say, a proton and an electron. And yet, in writing Schrödinger's equation for such a system, we use the term e^2/r [associated with inverse-square attraction] for the potential energy, carrying it over, so to say, from classical dynamics as a generalization ultimately based on the concept of force." More to the ontological point, one might take the fact that quantum interactions involve exchanges of momentum to suggest that forces are constituted by quantum particle exchanges.

A greater difficulty from the perspective of common applications of force-based mechanics is GTR's denial of gravitational forces. It appears that *if* GTR is the correct theory of gravity, then the posit of gravitational forces cannot be maintained. For GTR and Newtonian mechanics agree that inertial motion does not involve forces; hence there is no way of arguing that an object's inertial motion along a geodesic "constitutes" the occurrence of gravitational forces. It is presently unclear, however, whether GTR is the correct treatment of gravity. In response to well-known problems in incompatibility between GTR and quantum theories, various attempts are underway to quantize gravity, which if successful might allow for gravitational forces after all.

Philosophers who agree that forces exist may yet disagree over metaphysical details. It remains unclear for example whether forces are independent intermediaries between non-force causes and effects (as Bigelow et al. suggest), or are rather dependent aspects of the latter entities. What (considered) ontological category do forces fall under—are they properties, manifested dispositions, relations, causal relations, sui generis? Another question concerns the status of component vs. resultant forces. In cases in which phenomena involve more than one sort of force (e.g., both an electromagnetic and a gravitational force—supposing the latter exist), do the associated component forces (whose occurrence is expressed by Coulomb's law and Newton's law of gravitation, respectively) exist alongside the resultant force input into Newton's second law? Cartwright maintains that only the resultant force exists, while the component forces are mere mathematical fictions; Creary argues that the need to explain by composition of causes (here, forces) indicates that it is better to keep component and reject resultant forces.

Besides what might be called "Newtonian forces," a distinct but related scientific notion of force also deserves philosophical attention: that of a "fundamental force" or interaction. Paradigmatic fundamental forces/interactions (electromagnetic, gravitational, nuclear) come in many of the same varieties as paradigmatic Newtonian forces; and as already indicated, there are interesting open questions here concerning the relationship between (e.g., electromagnetic) Newtonian forces and the lower-level mechanisms operative in the field-theoretic treatments of the corresponding fundamental forces/interactions. Besides these general metaphysical concerns, fundamental forces/interactions may shed new light on old metaphysical debates. For example, an appeal to fundamental forces/interactions provides what is arguably the best way of formulating physicalism and emergentism as viably contrasting views: With this approach, physicalists maintain that all phenomena are grounded solely in fundamental physical forces/interactions, whereas emergentists maintain that, at certain complex levels of organization (notably, those involved in the having of mental states), a new fundamental force/interaction comes into play.

See also Bohm, David; Cartwright, Nancy; Maxwell, James Clerk; Newtonian Mechanics and Mechanical Explanation; Relativity Theory; Schrödinger, Erwin.

Bibliography

Berkeley, George. "De Motu." In *The Works of George Berkeley*, edited by A. A. Luce and T. E. Jessop. London: Nelson.

Bigelow, John, John Ellis, and Robert Pargetter. "Forces." *Philosophy of Science* (1988): 614–630.

Bohm, David, and B. J. Hiley. *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge, 1993.

Cartwright, Nancy. "Do the Laws of Physics State the Facts?" *Pacific Philosophical Quarterly* 61 (1980): 75–84.

Creary, Lewis. "Causal Explanation and the Reality of Natural Component Forces." *Pacific Philosophical Quarterly* 62 (1981): 148–157.

Feynman, Richard. *The Character of Physical Law*. Cambridge, MA: MIT Press, 1965.

Goldstein, Herbert, Charles Poole, and John Safko. *Classical Mechanics*. San Francisco, CA: Addison Wesley, 2002.

Hesse, Mary. "Review of Concepts of Force." *The British Journal for the Philosophy of Science* 10 (1960): 69–73.

Jammer, Max. *Concepts of Force*. Cambridge, MA: Harvard University Press, 1957.

Mill, John S. *A System of Logic*. London: Longmans, Green, Reader, and Dyer, 1843.

van Fraassen, Bas. *The Scientific Image*. Oxford: Oxford University Press, 1980.

Jessica Wilson (2005)

FOREKNOWLEDGE

See *Precognition*

FOREKNOWLEDGE AND FREEDOM, THEOLOGICAL PROBLEM OF

Divine foreknowledge, like the other classical theistic attributes, raises philosophical problems of at least three kinds. First, there are problems with understanding the attribute itself. How should it be construed (assuming that it is even coherent)? And how might God come by such knowledge? (Are future events all present in their causes? Does God arrive at foreknowledge by inference from "middle knowledge"? Does he *see* the future as through a "time telescope"? Or does he just know it?) Second, there are questions about how this attribute can be compatible with the other divine attributes. As the creator, sustainer, and providential overseer of the world, for example, God is supposed to be the supreme agent—but how can God approach the future as an active agent if his foreknowledge presents to him everything, including his own decisions and engagements with the world, as a fait accompli?

Finally, there are problems reconciling God's possession of this attribute with other things that appear undeniable. Of these, the most important is surely human freedom. If God knows before a person is even born exactly what that person will do throughout life, how could this person nevertheless retain the power to do otherwise, as free agency apparently requires? This is the classic foreknowledge problem; efforts to solve it are often what drive proposed solutions to the other two problems.

HISTORICAL BACKGROUND

In *De Interpretatione*, Aristotle worried that accepting the truth of future contingents would result in a necessitarianism incompatible with human freedom; for if it is true either that there will be a sea battle tomorrow or that there will not be a sea battle tomorrow, the admiral on whose decision this event depends either cannot issue the requisite order (if there will not be a sea battle) or cannot refrain from issuing the order (if there will be a sea battle). A similar worry was later elaborated into the influential "Master Argument" of Diodorus Cronus, discussed by the Stoics. Because this threat to human freedom rests solely on logical principles, like the Law of Excluded Middle, it is often called "logical fatalism" in contrast to the "theological fatalism" generated by divine foreknowledge.

The subtheistic nature of the ancient divinities and the pluralism of pagan theology made the problem of theological fatalism avoidable, but this was to change with the advent of Christianity. Augustine provides a classic early exposition of the problem in *On Free Choice of the Will* (III.3):

> How is it that these two propositions are not contradictory and inconsistent: (1) God has foreknowledge of everything in the future; and (2) We sin by the will, not by necessity? For … if God foreknows that someone is going to sin, then it is necessary that he sin. But if it is necessary, the will has no choice about whether to sin …. [So:] either we draw the heretical conclusion that God does not foreknow everything in the future; or … we must admit that sin happens by necessity and not by will.

Augustine went on to offer his own solution to this problem; his medieval successors added further solutions and

contributed enormously to the understanding of the problem, especially its modal character. Recent interest in the problem, sparked by a 1965 article by Nelson Pike, is probably as strong as it has been since the problem's heyday in the Middle Ages.

FORMULATING THE PROBLEM

As Augustine notes, the argument for theological fatalism is designed to show that a certain assumption about God is incompatible with a certain assumption about free will, so that one of them must be rejected.

THE GOD ASSUMPTION. The theological assumptions that play an actual role in the argument concern God's existence and cognitive excellence. It is assumed in the first place that God knows all truths, or

(i) God is omniscient.

Moreover, God believes *only* truths; indeed, he not only *does* not but *could* not believe any falsehoods. So

(ii) God is essentially inerrant, that is, infallible.

The final assumption about God is

(iii) God exists "from eternity."

The phrase *from eternity* is purposely ambiguous, straddling the view of God as an everlasting temporal being existing at all points in time (sempiternity) and the view of God as an atemporal being whose existence transcends time altogether (eternity proper). If (iii) is read, "There is no time such that the proposition *God exists*, if asserted at that time, would be false," then both views are accommodated. This allows for disambiguation, if necessary, to occur in the argument itself.

THE FREEDOM ASSUMPTION. The assumption with which the God Assumption is supposed to be incompatible is simply this:

Someone sometime does something freely.

Freely should be understood here in whatever sense is required for morally responsible agency, but otherwise pretheoretically—that way the theory of freedom under which it is allegedly incompatible with the God Assumption can emerge as a premise in the argument, and rejection of that premise can count as a solution to the problem.

THE ARGUMENT. Suppose someone X performs an action A at a time T3. Let T2 be a time prior to X's birth and T1 any time prior to T2. Then

(1) It is true at T1 that X will do A at T3.

The principle underwriting this claim, sometimes called the omnitemporality of truth, is that a statement true at any time is (suitably modified) true at every time. This does not imply, in the case of (1), that anyone can know at T1 what X will do at T3, let alone that there are conditions at T1 sufficient for X's future action; it only says that, since X's doing A at T3 is an assumption of the argument, it is, at T1, *true* that X will do A at T3.

According to clauses (i) and (iii) of the God Assumption, an omniscient God who exists "from eternity" must know at T1 whatever is true at T1. So

(2) God knows at T1 that X will do A at T3.

And if (2) is true, then so is

(3); God believes at T1 that X will do A at T3.

This follows from the standard analysis of knowledge, according to which knowledge entails belief.

Once God holds this belief, it becomes part of the fixed past that he held that belief. It is no longer possible for him not to have held this belief. This is an instance of the "necessity of the past," conveyed in such maxims as "what's done is done." This is not logical necessity, since there are logically possible worlds with a different past; but it is arguably stronger than natural or causal necessity. Aristotle notes that "this alone is lacking even to God, to make undone things that have once been done" (*Nicomachean Ethics* VI.2.1139b10–11), and Aquinas comments, "As such it is more impossible than the raising of the dead to life, which implies no contradiction, and is called impossible only according to natural power" (*Summa Theologiae* I.25.4). Because what is necessary when past might have been nonnecessary or accidental when future, it is often called accidental necessity. The next step in the argument can therefore be stated this way:

(4) It is accidentally necessary at T2 that God believed at T1 that X will do A at T3.

Since T1 is past relative to T2, (4) is true, given (3).

Though accidental necessity was introduced as a modality characteristic of the past, it is more general in scope. For a proposition *p* to be accidentally necessary at a time T is for *p* to be true no matter how the world continues after T. The past is then accidentally necessary by default; but the future can also qualify as accidentally necessary if entailed by accidentally necessary facts about the past. One such fact is the following:

(5) It is accidentally necessary at T2 that X will do A at T3.

This follows from (4) combined with clause (ii) of the God Assumption, according to which God's believing that X will do A at T3 entails that X will do A at T3.

Since T2 is a time prior to X's birth, X comes into existence with it already being the case that he must do A at T3. It is therefore too late for X to bring it about that he fails to do A at T3—that is,

(6) X cannot refrain from doing A at T3.

But if X cannot refrain from doing A at T3, then

(7) X does not do A at T3 *freely*.

This last inference is sanctioned by a "freedom version" of the so-called Principle of Alternate Possibilities, according to which a person is morally responsible for performing an action only if the person could have refrained from performing it. If a person is not morally responsible, owing to an inability to refrain, this person is not free in the sense required for moral responsibility. This is precisely the sense of "free" that is relevant to the Freedom Assumption.

The foregoing argument does not turn on any peculiar features of X, A, or T3; the same argument can be given for any agent, action, and time. So no one ever does anything freely. If the God Assumption is true, the Freedom Assumption is false.

SOME COMMENTS ON THE ARGUMENT

Before canvassing possible responses to this argument, some explanatory remarks are in order.

First, some versions of the argument bypass (4), inferring the necessity of X's future action from (3) and divine infallibility alone. Such versions might succeed if clause (ii) of the God Assumption could be parsed this way:

(iia) If God believes that p, then necessarily p.

Unfortunately, the correct analysis of divine infallibility is

(iib) Necessarily, if God believes that p, then p.

And all that follows from (iib), given simply that God believes that X will do A at T3, is that X will do A at T3 (and will do so in any world in which God holds this belief). For the action to be necessary, based on (iib), God's belief must be necessary. The illusion that (5) can be derived without reliance on (4) is produced by an equivocation between (iia) and (iib). Boethius, who called the necessity in (iia) "simple necessity" and the necessity in (iib) "conditional necessity," and Aquinas, who termed these "the necessity of the consequent" and "the necessity of the consequence" respectively, diagnosed the problem accurately and rightly insisted on the ineliminability of (4).

Second, step (4) does not rest on the simplistic principle that all true statements indexed to the past through tense or temporal references like "at T1" are accidentally necessary. This principle is in fact false. Confident of victory in tomorrow's election, the candidate proclaims, "My campaign for President began two years before its successful completion." Having just been fooled, I vow, "That was the last time I'm falling for that trick!" Suppose these declarations are in fact true. Though both assert something about the past, neither one is accidentally necessary, since either could (though ex hypothesi it won't) turn out false: The candidate might lose, I might get fooled again. Statements like these, which are not *genuinely* and *strictly* about the past, are called "soft facts" about the past as opposed to the "hard facts" to which the necessity of the past is applicable. What justifies (4), then, is that (3) looks like a *hard* fact about the past. (Certainly there is little question about the human analogue: If Joe believed yesterday that he will shave tomorrow, it is a hard fact, and therefore accidentally necessary, today that he held this belief yesterday.) This also explains the apparently trivial move from (2) to (3). To say of God's cognition in (2) that it constitutes knowledge is to say, in part, that it is true; but its truth depends on how things go at T3. So (2) is not strictly about the past; unlike (3), it is not a hard fact relative to T2.

Third, some critics point out that the future-truth argument for logical fatalism also begins with (1), but then moves directly to (4') It is accidentally necessary at T2 that it was true at T1 that X will do A at T3, and thence to (5). Their claim is that the argument for theological fatalism is just a needlessly complicated version of this argument, and is equally fallacious. The problem with this critique is that (1) is a paradigmatic soft fact relative to T2, undermining the inference to (4'), whereas routing the argument through the theological premises (2) and (3) allows (4) to follow from a prima facie hard fact about the past. This gives the argument for theological fatalism a clear logical advantage.

RESPONSES

If the argument succeeds, either the God Assumption or the Freedom Assumption must be rejected. Those who

deny the Freedom Assumption in response to the argument are "theological fatalists." There appear to be very few theological fatalists in this sense. Calvinists would qualify if anyone would. But most Calvinists are compatibilists and would therefore affirm the Freedom Assumption, while those who do reject the Freedom Assumption tend to do so on grounds other than the argument for theological fatalism.

Denying the God Assumption does not entail atheism unless the falsity of just one of the three clauses constituting the God Assumption is sufficient for there being no God. Some theists, indeed, deny that clause (i) is essential to theism when omniscience includes future contingents. If the argument succeeds, such truths are logically unknowable and should be excluded from divine omniscience, just as the logically impossible is excluded from divine omnipotence. "Open Theists" sometimes take this position, maintaining that God willingly limits his foreknowledge to make space for human freedom. There are, however, a number of reasons for thinking that the argument does not succeed.

THE ARISTOTELIAN SOLUTION. Step (1) has been rejected on the grounds that a statement about the future is not (yet) either true or false; it acquires a truth value only when what is now future becomes present. This seems to have been the position Aristotle adopted in response to the "future truth" argument for logical fatalism. It is also the favored position of Open Theists who prefer not to deny the God Assumption: If future contingents lack truth value, a deity who fails to foreknow them will not thereby lack anything necessary to omniscience. Critics, however, have pointed to serious logical costs associated with this move.

THE BOETHIAN SOLUTION. Step (2) follows from (1) only if God exists at T1. But if God does not exist in time, a view famously associated with Boethius, (2) is false; what is true instead is:

> (2*) God (timelessly) knows that X will do A at T3.

Two questions may be raised here. The first is whether this view of God is coherent: Though it is the classical view, it has come in for increasing criticism in recent years. The second question is whether a timeless deity might succumb to a modified version of the argument. It has been claimed, for example, that (2) can be replaced by:

> (2#) It is true at T1 that God (timelessly) knows that X does A at T3.

(3) and (4) can be similarly modified, and (5) will then follow as before. It has also been claimed that what is fixed in eternity may be no less accidentally necessary than what is fixed in the past, so that (2*) leads to:

> (4*) It is accidentally necessary at T2 that God (timelessly) believes that X will do A at T3.

and thence again to (5). But intuitions are a fragile guide here, and the viability of the Boethian solution remains open.

THE OCKHAMIST SOLUTION. The most popular solution in the contemporary debate is the denial of (4). A radical critique might challenge the very idea of accidental necessity as a modality characteristic of the past; but this extreme position runs counter to deep intuitions about the necessity of the past. The principal assault has come from those who accept the necessity of the past but argue, following William Ockham, that (3) is really a soft fact about the past.

In his treatise *Predestination, God's Foreknowledge and Future Contingents*, Ockham distinguishes hard and soft facts this way: "Some propositions are about the [past] as regards both their wording and their subject matter. … Other propositions are about the [past] as regards their wording only and are equivalently about the future, since their truth depends on the truth of propositions about the future." Ockham's modern followers have cited at least four grounds for placing (3) among the latter propositions.

First, God's belief that X will do A at T3 is counterfactually dependent on X's doing A at T3; if X were to do otherwise, God would have believed otherwise. Unfortunately, this counterfactual dependence can obtain even if X *cannot* do otherwise; hence it provides no reason to think that God can still believe otherwise.

Second, one might develop necessary and sufficient conditions for hard facthood and show that (3) does not qualify. If this is done in terms of an "entailment criterion," it appears that (3) is a soft fact after all, since it entails the future fact that X will do A at T3. But analyses of the hard/soft distinction, most employing entailment criteria, have grown mind-numbingly complex in response to counterexamples, and none has won consensus. This strategy has fallen into disfavor.

Third, one might approach the question from the side of the divine beliefs. How should divine cognition be construed so that (3) can be a soft fact relative to T2? Perhaps the "narrow content" of God's belief is a hard fact about the past, but its "wide content" is determined by the

way the future actually unfolds—that the belief counts as the belief *that p* might then be a soft fact about the past. Or perhaps God's beliefs about future contingents are dispositional rather than occurrent in nature, and this makes a difference to their status as hard or soft; then God might be (dispositionally) omniscient at the same time as future contingents remain contingent. Or perhaps, as William Alston (1986) argues, God does not even have beliefs—a position which Linda Zagzebski (1991) terms Thomistic Ockhamism. Even if coherent, such proposals appear to make God's foreknowledge unavailable to him for action-guidance.

Fourth, one might finesse the difficulties of the above approaches with a direct demonstration that (3) is a soft fact, as suggested by Alvin Plantinga (1986) and (in another form) Ted Warfield (1997). If God exists necessarily, then (3) is true in all and only the worlds in which (1) is true, making (3) logically equivalent to (1). Since (1) is a paradigmatic soft fact relative to T2, (3) must be a soft fact as well, and (4) no longer follows. Critics, however, have charged this argument with question-begging.

THE SCOTIST SOLUTION. The inference from (4) to (5) has this form:

(4) It is accidentally necessary that M

(iib) Necessarily, if M, then N

Therefore:

(5) It is accidentally necessary that N

This is a so-called *transfer principle*, since it transfers necessity from one proposition to another. Whether the inference is valid depends on the logic of accidental necessity. The parallel inference for logical necessity is certainly valid; if accidental-necessity-at-T can be modeled as truth in all of some subset of logically possible worlds—for example, the set of all worlds that share the same past up to T—then the above inference should be valid as well. Nevertheless, some types of necessity appear not to work like this, and similar transfer principles, like Peter Van Inwagen's (1983) "rule β," have been disputed.

THE EDWARDSIAN SOLUTION. A compatibilist about free will and causal determinism will not agree that (5) is a reason for endorsing (6). The case for (and against) compatibilism is too large a subject to be broached here and is best pursued in connection with the problem of freedom vs. determinism, where it has received its most sophisticated development. Among theists who come to compatibilism from theological rather than causal determinism, Jonathan Edwards is notable for rejecting step (6).

THE AUGUSTINIAN SOLUTION. Augustine seems to have argued that the agent might remain free even if divine foreknowledge closes all alternatives, so long as the agent's action is self-initiated and God's foreknowledge does not cause, compel, or otherwise explain the action. (How this fits with what Augustine says about divine grace, sovereignty, and predestination is another question.) The moral Augustine draws from foreknowledge cases is arguably the same moral that Harry Frankfurt (1969) draws from cases in which a mechanism eliminates an agent's alternatives without interfering with the agent's actual course of action; indeed, when divine foreknowledge is the mechanism, the result appears to be a perfect "Frankfurt-type counterexample" to the Principle of Alternate Possibilities, on which step (7) rests. If, however, only a predetermined future can be foreknown, even by God, then this solution fails.

THE NATURE OF THE PROBLEM

There are a number of philosophical problems in the neighborhood of this one that can be approached simply as thought experiments, without regard to whether the world is arranged as the problem presupposes. These include Newcomb's puzzle, the paradoxes of time travel and retrocausation, and perhaps even causal determinism itself. The problem of theological fatalism might be one of these; if it is, certain solutions become irrelevant. If someone reflecting on Zeno's paradoxes of motion is puzzled about how Achilles could fail to pass the tortoise, the puzzlement is not addressed by denying Achilles' existence or by reconceiving his attributes (for example, by making him a cripple or supposing he is in a coma). Likewise, someone reflecting on the argument for theological fatalism might be puzzled about how a paradigmatic candidate for free agency might be rendered unfree simply by adding infallible foreknowledge to the situation. Reconceiving God or denying God's existence outright simply removes God from complicity in this puzzle; it does not solve the puzzle. The purely theological solutions—the Boethian and the third and fourth of the Ockhamist responses—fail to address this deeper puzzle, assuming that it is genuine.

See also Alston, William; Aristotle; Augustine, St.; Boethius, Anicius Manlius Severinus; Diodorus Cronus; Duns Scotus, John; Edwards, Jonathan; Frankfurt, Harry; Freedom; Plantinga, Alvin; Precognition; Stoicism; Thomas Aquinas, St.; William of Ockham.

Bibliography

Adams, Marilyn McCord. "Is the Existence of God a 'Hard' Fact?" *Philosophical Review* 76 (1967): 492–503.

Alston, William P. "Does God Have Beliefs?" *Religious Studies* 22 (1986): 287–306.

Craig, William Lane. *The Only Wise God*. Grand Rapids, MI: Baker Book House, 1987.

Fischer, John Martin. "Freedom and Foreknowledge." *Philosophical Review* 92 (1983): 67–79.

Fischer, John Martin, ed. *God, Foreknowledge, and Freedom*. Introduction by John Martin Fischer. Stanford, CA: Stanford University Press, 1989.

Fischer, John Martin. "Scotism." *Mind* 94 (1985): 231–243.

Frankfurt, Harry. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 46 (1969): 828–839.

Freddoso, Alfred J. "Accidental Necessity and Logical Determinism." *Journal of Philosophy* 80 (1983): 257–278.

Haack, Susan. "On a Theological Argument for Fatalism." *Philosophical Quarterly* 24 (1974): 156–159.

Hasker, William. *God, Time, and Knowledge*. Ithaca, NY: Cornell University Press, 1989.

Hunt, David P. "The Compatibility of Divine Determinism and Human Freedom: A Modest Proposal." *Faith and Philosophy* 19 (2002): 485–502.

Hunt, David P. "Dispositional Omniscience." *Philosophical Studies* 80 (1995): 243–278.

Hunt, David P. "On Augustine's Way Out." *Faith and Philosophy* 16 (1999): 3–26.

Hunt, David P. "What *Is* the Problem of Theological Fatalism?" *International Philosophical Quarterly* 38 (1998): 17–30.

Kvanvig, Jonathan. *The Possibility of an All–Knowing God*. New York: St. Martin's Press, 1986.

Pike, Nelson. "Divine Omniscience and Voluntary Action." *Philosophical Review* 74 (1965): 27–46.

Plantinga, Alvin. "On Ockham's Way Out." *Faith and Philosophy* 3 (1986): 235–269.

Purtill, Richard L. "Fatalism and the Omnitemporality of Truth." *Faith and Philosophy* 5 (1988): 185–192.

Stump, Eleonore, and Norman Kretzmann. "Eternity." *Journal of Philosophy* 78 (1981): 429–458.

Van Inwagen, Peter. *An Essay on Free Will*. Oxford: Clarendon Press, 1983.

Warfield, Ted. "Divine Foreknowledge and Human Freedom Are Compatible." *Nous* 31 (1997): 80–86.

Widerker, David. "Troubles with Ockhamism." *Journal of Philosophy* 87 (1990): 462–480.

Widerker, David. "Two Forms of Fatalism." In *God, Foreknowledge, and Freedom*, edited by J. M. Fischer, 97–110. Stanford, CA: Stanford University Press, 1989.

Zagzebski, Linda. *The Dilemma of Freedom and Foreknowledge*. New York: Oxford University Press, 1991.

Zemach, Eddy, and David Widerker. "Facts, Freedom, and Foreknowledge." *Religious Studies* 23 (1988): 19–28.

David P. Hunt (2005)

FORGIVENESS

Through the mid-twentieth century, academic treatments of forgiveness were largely theologically based. The latter part of the century saw the start of a secular discussion of forgiveness within analytic philosophy. The topic provides rich ground for philosophical reflection.

Participants in the discussion often focus on three issues: what forgiveness is, how it is accomplished, and when it is justified. Regarding the first, many appropriate Bishop Butler's claim that forgiveness is the overcoming of resentment. It is widely thought to be accomplished through compassion, perhaps by an imaginative process. The question of justification raises interesting issues about whether forgiveness can be required or whether it is always supererogatory.

There is, however, a prior question of considerable philosophical interest: How is forgiveness, so understood, even possible? Most would agree that not just any elimination of resentment counts as forgiving. You could not forgive by simply taking a pill that rendered you incapable of resentment. Nor does simply forgetting count as forgiving. Forgiveness requires overcoming resentment in the right way. However, it is not merely hard to say what that way is; it is unclear whether there could be such a way.

To keep forgiveness distinct from other responses, such as excuse or contempt, the forgiver must not deny (a) the seriousness of the wrong, (b) the moral standing of the wrongdoer, or (c) his or her own moral standing. Overcoming resentment by denying either the seriousness of the wrong or one's own claim against being wronged is excusing. Overcoming resentment by denying the standing of the wrongdoer is showing contempt for the wrongdoer, excluding him or her from the class of persons whose actions matter. To forgive, one must affirm the seriousness of the wrong and the importance of both oneself and the wrongdoer. Forgiveness must be uncompromising. The difficulty is that the three claims that forgiveness must not deny seem sufficient to ground the resentment that forgiveness must overcome. How, then, is forgiveness possible?

If resentment were necessarily vengeful or malicious, one could overcome it without compromise by achieving compassion. But resentment—that anger over a wrong that is incompatible with forgiveness—is not necessarily vengeful or malicious. One can empathize with the plight of the wrongdoer, have no desire to see him or her harmed, and still resent the wrong. Thus, in contrast with

a widely held view, compassion will not secure forgiveness.

If the three most obvious ways to overcome resentment—to discount the wrong, the wrongdoer, or oneself—were the only ways to overcome it, then forgiveness would be impossible. In order to understand an overcoming of resentment as a case of forgiveness, it needs to be distinguished from compromise. Here, then, lies a task for philosophy: to provide an articulate account of the way in which the overcoming of resentment can count as forgiveness. With that task completed, discussion can turn to how forgiveness is accomplished and when it is justified.

See also Moral Sentiments.

Bibliography

Butler, Bishop Joseph. *Sermons*. Boston: Hilliard and Brown, 1827.

Grover, Trudy. *Forgiveness and Revenge*. London: Routledge, 2002.

Hieronymi, Pamela. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62 (2001): 529–555.

Murphy, Jeffrie G., and Jean Hampton. *Forgiveness and Mercy*. Cambridge Studies in Philosophy and Law. Cambridge, U.K.: Cambridge University Press, 1988.

Pamela Hieronymi (2005)

FORM

See *Medieval Philosophy; Plato; Universals, A Historical Survey*

FORM, AESTHETIC

See *Aesthetics, History of; Aesthetics, Problems of*

FORMALISM IN MATHEMATICS

See *Mathematics, Foundations of*

FOUCAULT, MICHEL

(1926–1984)

Michel Foucault, though trained in philosophy, never considered himself a professional philosopher. Still, his research into the historical formation of truth, power relations, and modes of recognition regarded as self-evident in various disciplines—most notoriously the figure of man—is an important contribution to philosophy and is itself strikingly original philosophical thought. Born in Poitiers, France, Foucault studied at the École Normale Supérieure under Maurice Merleau-Ponty, Jean Beaufret (1907–1982)—Martin Heidegger's major interpreter in France—and Louis Althusser (1918–1990). Foucault earned his License de philosophie in 1948 and Diplôme de psycho-pathologie in 1952. He taught in Sweden, Poland, and Germany before his appointment as the head of the philosophy department at the University of Clermont-Ferrand. After two years in Tangiers following the publication of *Les mots et les choses* (*The Order of Things*) in 1966, Foucault returned to France and the university at Vincennes, France, just after the anti-authoritarian protests of May 1968. Foucault was elected to the Collège de France in the fall of 1970. Though he grew more engaged in political struggles in the late 1960s and early 1970s, his resistance to humanism made him an uneasy participant in organized movements. Still, his activism and writing earned him attention in the United States, where he became a popular lecturer. Foucault contracted AIDS at the outset of the epidemic and died of complications from the disease in June 1984.

Foucault's work is often divided into three periods, the earliest marked by his archaeological approach, the middle by a genealogy of the modern subject and the relation between power and knowledge, and the late identified with his turn to ethics and the "care of the self." This chronology is controversial: though it orients much of the secondary literature about Foucault, its value lies in its convenience more than in its philosophical or conceptual importance. Taken together, Foucault's works pursue critical inquiry into formative, elementary dimensions of knowledge, autonomy, and experience and are an important contribution to a process of critical engagement with the emergence and limitations of dominant forms of power and knowledge. His goal was to analyze the conditions under which forms of self-relation are created or modified so far as these relations constitute possible knowledge of oneself when such knowledge is referred to something other than an essential identity. Through a historical or genealogical approach to these conditions, Foucault challenges the traditional philosophical model

of the subject as having a nature or essence associated with ahistorical capabilities.

ARCHAEOLOGY

Folie et déraison (*Madness and Civilization*; 1965) is the first of Foucault's archaeological works. At the time it was published, Foucault's thinking ranged from psychology and the human sciences (in relation to Ludwig Binswanger, Gaston Bachelard, and Georges Canguilhem [1904–1995]) to Friedrich Nietzsche and avant-garde literature. The book is therefore a powerful introduction to the challenge posed to traditional philosophical practice (and the dominance of phenomenology and existentialism in France) by the growing interest in structuralism, psychoanalysis, and postmodernism. Combining a materialist historical approach associated with the Annales group (Ernst Bloch, Henri Lefebvre [1901–1991], and Fernand Braudel [1902–1985]) and an ontology of the subject derived from his engagement with literature and his critical approach to psychoanalysis, *Madness and Civilization* established Foucault as an important philosopher and social critic in France.

THE ASYLUM AND THE CLINIC

Madness and Civilization traces the emergence of a form of reason in reason's encounter with indications of its limits in unreason (in the Renaissance) and later in madness (in the classical age—the mid-seventeenth century to the beginning of the nineteenth century). Reason encountered its limits in the course of a transformation—at once administrative, moral, and epistemological—in which the exclusion of madness at the margins of community gave way to its confinement in hospitals and then in asylums. This confinement produced new objects of study—excluded populations marked by an inability to work, moral weakness, and disorder—displayed and subjected to emerging forms of knowledge and techniques for the disciplining of disorder and the cure of insanity. On the basis of these practices scientific psychology established the limits of the "normal," themselves a product of a moral, medical, and juridical synthesis made possible by an ascendant administrative capacity to confine populations marked by unreason.

Madness and Civilization comprises an examination of the historical a priori conditions of the emergence of classical reason and an imaginative account of the formation of an experience of reason that defines not only the classical age (particularly René Descartes) but also contemporary thought. Its archaeological approach supposes that discursive formations—statements that delimit and condition what can sensibly be said of madness—are governed by rules that are not reducible to subjective intentions or consciousness and that also govern what can be said or known. *Madness and Civilization* can also be understood as a preface to an analysis of discursive practices that produce relations of knowledge and power. It thus introduces readers to themes that traverse Foucault's work: the exclusion of difference in institutional contexts, the formation of knowledge of subjects on the basis of that exclusion, the relationship between knowledge and power, and the possibility of achieving distance from one's judgments, commitments, and philosophical prejudices through critique. Thus, these works are critical in the Kantian sense as Foucault understood it: they allow one to examine and transform the conditions through which the subject becomes an object of possible knowledge.

Foucault pursues a similar archaeological project in *Naissance de la clinique* (*The Birth of the Clinic*; 1963), an account of the formation of a mode of perception that makes possible medical knowledge of the body. Foucault shows that modern knowledge of disease is dependent on changing structures of perception and language that are sustained by practices and powers that inhabit the space of the clinic. Where standard histories of medicine portray medical knowledge as derived from an unstructured gaze and converging on objectivity, Foucault shows that accepted medical practices have their origins in something other than necessities of medical reason (e.g., the practice of the "round") or inference and pure observation in the context of steadily improving methods. The philosophical importance of the book is its analysis of the merging of clinical language and ways of seeing—a contingent form of a gaze and its links to institutional powers that sustain it—with the language of rationality.

WORDS AND THINGS

The most significant work of Foucault's archaeological period is *Order of Things*, in which Foucault again unearths and articulates the historical conditions for the possibility of knowledge in the human sciences in a given period: those knowledges associated with labor, life, and language. At the same time, *Order of Things* is a genealogy in an important sense: it traces the emergence of Foucault's own commitments and of the privileges and imperatives that accompany his own discourse. Thus, some critics accuse Foucault of engaging in criticism that leaves him with no standpoint from which to judge structures of power and knowledge that are evidently in question in his work, undermining his own ground and

promulgating relativism. Foucault called this charge "intellectual blackmail."

Order of Things is a genealogy of the Same, of the rules and conditions that make possible the perception and knowledge of order. It proceeds by way of an account of two profound breaks in the coherence of knowledge about man and of the way those breaks affect modern knowledge and give it resources with which to freely think new possibilities. The first break occurred between the Renaissance and the classical epistemes. Foucault uses the word *episteme* to designate the regularities that account for the coherence of knowledge in a given period. The Renaissance episteme was coherent—one could speak truly about nature and link one's speech to the world—because of its dependence on resemblance and similitude for the organization of what counted as knowledge and true perception. But this understanding of the relationship between language and the world, between the signifier and the signified, is ultimately broken—similitude becomes deceptive. The subsequent Renaissance episteme is oriented around the primacy of representation: the capacity of language to mirror the world and to correspond to it in a truthful way by virtue of its capacity to organize the multiplicity of identities and differences in a table or grid, making possible a new recognition of sameness. This is the first of two breaks.

Foucault's primary concern, though, is to document the second break, the "profound upheaval" that led to the disintegration of representation at the end of the eighteenth century. This disintegration was prompted in various domains by a growing recognition of the limits of representation, particularly of its ability to account for the act of representing itself and to adequately represent the being who represents. As a result of this disintegration, knowledge in the human sciences becomes an "Analytic of Finitude." Man appears for the first time as both the object of knowledge and the one who knows, an "empirico-transcendental doublet" understood in terms of his labor that can be alienated, his organism that is part of an evolutionary history, or his speaking a language that is no longer controlled by a representing subject but that has its own historicity, rules, and organic structures, while being utterly internal. Knowledge of man as this doublet is thus dependent on being able to account for man's being in those places or regions in which man is absent. One of the consequences of this analysis is that the centrality of the figure man is itself subject to questioning and overcoming, which Foucault hoped his work would both reflect and generate. This project is to a large extent shared by Nietzsche, Heidegger, and Sigmund Freud.

L'archéologie du savoir (*The Archaeology of Knowledge*; 1969) attempts to give a systematic account of his methodological assumptions and procedures in his archaeological works, formulating the rules that operate within a discourse "at a superficial level" and that constitute a discourse's coherence as a "game of truth." Foucault's work after *Archaeology of Knowledge* is usually understood as genealogical in scope and approach.

GENEALOGY

The word *genealogy* is associated with Nietzsche and is understood as a patient tracing of the descent of authoritative discursive practices that structure the application of power to bodies and subjects (e.g., in the school, the hospital, and the prison). Foucault studies *dispositifs*, practices that exclude and construct forms of experience as abnormal in various ways (e.g., criminality, madness, and sexual deviancy) and that construct forms of subjectivity on the basis of knowledge of normalcy (e.g., the soldier, the student, the guard, or the attendant). He examines practices and texts that are no longer part of received knowledge but that nevertheless were important in the formation of a practice or the exclusion of a form of experience, where genealogy is an attempt to remember those lost experiences and complicated formations. The genealogy of various formations of subjectivation led Foucault to the identification and articulation of forms of power, most importantly the power of surveillance—a "microphysics of power"—in *Surveiller et punir* (*Discipline and Punish*; 1975).

Discipline and Punish concerns the emergence of the modern power to punish in the prison and of the way in which the prison, through observation, examination, and normalizing judgment, produces the conditions for the recognition of delinquency. Thus, it is a genealogy of the way in which power divides the "normal" from the "incarcerated" and of the formation of self-relation around the axes of normalcy, lawfulness, and the careful monitoring of one's own excesses. Modern power encourages one to correct one's own deviance. The notion of power at work in *Discipline and Punish* applies to the practices and techniques that operate inside and outside of the prison that discipline subjects who show signs of disorder (e.g., children, soldiers, students, crowds, criminals, and workers). Those techniques aim to produce a moral subject capable of self-discipline and of being aware of the virtues of obedience.

On this conception of power there are no agents in whom power is concentrated, but only techniques, regimens, regulations, and measures that divide the normal

or average from the pathological or criminal. This power is not in the service or control of a dominant interest, class, or group, but dispersed throughout the social body and concentrated in various institutions that are simultaneously carceral and clinical. This dispersion makes resistance to power difficult, but Foucault thought resistance was possible by intensifying one's recognition of the intolerability of specific forms of power by attention to voices or discourses that cannot be adequately heard from within dominant regimes. He conceived of his work as tools for use in the strategic interruption of dominant discourses and practices.

ETHICS

While working on his genealogies and occasional politically incendiary essays in the 1970s (including lecture courses on the contemporaneous emergence of psychiatry and racism in *Abnormal* [2003a] and on discourse of and as war in *Society Must Be Defended* [2003b]), Foucault assembled his three-part *Histoire de la sexualité* (*History of Sexuality*). *La volonté de savoir* (*An Introduction*; 1976), the first volume, was an analysis of the "repressive hypothesis," the idea that sexual expression went through a period of repression in the Victorian era and subsequently was liberated by an increasing awareness of the naturalness of sex. Foucault argues instead that sex was an important and much discussed issue for the Victorians and that discourses of sexuality and techniques of sexual control and expression are important avenues through which power operates on the body (by encouraging subjects to work on themselves) and are not reducible to a single repressive power. To examine what he called subjectivizing practices at work on the formation of sexuality, he constructs a genealogy of the experience of sexuality. On Foucault's terms sexuality is not a constant, natural feature of human beings, but takes historically singular forms, the emergence of which can be traced through a genealogical account.

L'usage des plaisirs (*The Use of Pleasure*; 1984) and *Le souici de soi* (*Care of the Self*; 1984), the second and third volumes of the *History of Sexuality*, respectively, were published eight years after the first volume and after considerable revision of his overall project. Foucault turns his attention from relatively recent formations of sexuality in the eighteenth and nineteenth centuries to the problem of desire and the desiring subject in ancient Greek and Hellenic thought, though always in relation to the present. He conducts a genealogy of the problematizations—the ways in which certain practices and forms of knowledge become a matter of concern—and practices surrounding the formation of the subjects who can recognize and understand themselves in terms of the techniques, ethical concerns, and political relations that form around men who desire. *Use of Pleasure* focuses on the ways in which pleasure was a matter of concern for the Greeks and how it played a crucial role in the command that one "know thyself." Foucault then traces a change from a focus on pleasure and its use to a focus on desire and how to protect oneself from its dangers before the emergence of the Christian problematic of pleasure, desire, and ethics. The third volume is a genealogy of the emergence of the modern subject in Hellenic and Roman practices of self-control and asceticism.

Foucault made important contributions to discreet areas of philosophical research, including feminist philosophy and gender theory, social, political and legal philosophy, the philosophy of science, aesthetics, theories of knowledge, and especially ethics, which is a constant concern throughout Foucault's works. While Foucault resisted moral theory and insisted on its danger, and while he resisted the articulation of a solid moral stance on which one could found commitment or advocacy, he nevertheless insisted on the ethical value of his genealogical work. Through the investigation of the conditions under which subjects are formed and modes of recognition are validated or legitimated, Foucault intensified awareness of the subjugating powers that invest the practices and discourses that structure one's understanding of oneself and others and turned that awareness back on itself to promote the exploration of new and singular modes of self-relation.

See also Archaeology; Binswanger, Ludwig; Bloch, Ernst; Descartes, René; Feminist Philosophy; Freud, Sigmund; Heidegger, Martin; Merleau-Ponty, Maurice; Nietzsche, Friedrich; Renaissance; Structuralism and Post-structuralism; Subject.

Bibliography

WORKS BY FOUCAULT

Mental Illness and Psychology. Translated by Alan Sheridan. New York: Harper and Row, 1976. Originally published under the title *Maladie mentale et psychologie* in 1962.

Birth of the Clinic. Translated by A. M. Sheridan Smith. New York: Pantheon, 1973. Originally published under the title *Naissance de la clinique* in 1963.

Madness and Civilization. Translated by Richard Howard. New York: Pantheon, 1965. Originally published under the title *Folie et déraison* in 1966.

The Order of Things: An Archaeology of the Human Sciences. New York: Pantheon, 1970. Originally published under the title *Les mots et les choses* in 1966.

The Archaeology of Knowledge. Translated by A. M. Sheridan Smith. New York: Pantheon, 1972. Originally published under the title *L'archéologie du savoir* in 1969.

Discipline and Punish: The Birth of the Prison. Translated by Alan Sheridan. New York: Pantheon, 1977. Originally published under the title *Surveiller et punir* in 1975.

The History of Sexuality. 3 vols. Translated by Robert Hurley. New York: Pantheon, 1978–1986. Originally published in three volumes under the title *Histoire de la sexualité* in 1976 and 1984.

The Essential Works of Foucault, 1954–1984. 3 vols. Translated by Robert Hurley. New York: New Press, 1997–2000.

Fearless Speech. Edited by Joseph Pearson. Los Angeles: Semiotext(e), 2001.

Abnormal: Lectures at the Collège de France, 1974–1975. Translated by Graham Burchell; edited by Valerio Marchetti and Antonella Salomoni. New York: Picador, 2003a.

Society Must be Defended : Lectures at the Collège de France, 1975–1976. Translated by David Macey; edited by Mauro Bertani and Alessandro Fontana. New York: Picador, 2003b.

The Hermeneutics of the Subject: Lectures at the Collège de France, 1981–1982. Translated by Graham Burchell; edited by Frédéric Gros. New York: Palgrave-Macmillan , 2004.

WORKS ABOUT FOUCAULT

Bernauer, James W. *Michel Foucault's Force of Flight: Toward an Ethics for Thought*. Atlantic Highlands, NJ: Humanities Press International, 1990.

Eribon, Didier. *Michel Foucault*. Translated by Betsy Wing. Cambridge, MA: Harvard University Press, 1991.

Davidson, Arnold I., ed. *Foucault and His Interlocutors*. Chicago: University of Chicago Press, 1997.

Deleuze, Gilles. *Foucault*. Minneapolis: University of Minnesota Press, 1988.

Dreyfus, Hubert L., and Paul Rabinow. *Michel Foucault: Beyond Structuralism and Hermeneutics*. Chicago: University of Chicago Press, 1982.

Foucault.Info (May 2005), http://www.foucault.info.

Foucault Resources (May 2005), http://www.qut.edu.au/edu/cpol/foucault/.

Gutting, Gary. *Michel Foucault's Archaeology of Scientific Reason*. New York: Cambridge University Press, 1989.

Han, Béatrice. *Foucault's Critical Project: Between the Transcendental and the Historical*. Translated by Edward Pile. Stanford, CA: Stanford University Press, 2002.

Macey, David. *The Lives of Michel Foucault: A Biography*. New York: Pantheon, 1993.

McWhorter, Ladelle. *Bodies and Pleasures: Foucault and the Politics of Sexual Normalization*. Bloomington: Indiana University Press, 1999.

Sawicki, Jana. *Disciplining Foucault: Feminism, Power, and the Body*. New York: Routledge, 1991.

Visker, Rudi. *Michel Foucault: Genealogy as Critique*. Translated by Chris Turner. New York: Verso, 1995.

Benjamin Pryor (2005)

FOUCHER, SIMON

(1644–1696)

Simon Foucher was one of the foremost critics of Cartesian philosophy. He was born in Dijon, France, where, after taking orders, he was made honorary canon of the Sainte Chapelle. He took a bachelor's degree in theology at the Sorbonne and spent his adult life as a chaplain in Paris, where he died. His first published work is a long didactic poem commemorating the death of Anne of Austria (1601–1666). In another long poem he defends the compatibility of Greek and Christian moral principles. In Paris he attended the lectures on Cartesian physics given by Jacques Rohault, which inspired him to make original experiments in the science of hygrometry (humidity of the atmosphere) on which he published two pioneering works in 1672 and 1686. He also produced three major dissertations concerning the value of Academic skepticism in the search for truth. He was the first to publish criticisms of both Nicolas Malebranche's occasionalism and Gottfried Wilhelm Leibniz's monadism, and it is for these critiques that he is best known.

ACADEMIC METHOD

Foucher considered himself to be the reviver of Academic philosophy, by which he means Socratic ignorance combined with the reasonable doubt of Philo of Larissa and Antiochus of Ascalon, who say that they know some things and are ignorant of others; he argues that this is the middle way between dogmatism and Pyrrhonism. The primary maxim of his Academic philosophy is to recognize only *vérité evident* as a rule of truth. The Academic laws are:

(1) To proceed only by demonstrations in philosophy

(2) To avoid unanswerable questions

(3) To admit when one does not know

(4) To distinguish what one knows from what one does not know

(5) Always to seek after knowledge

There are three important axioms:

(1) True knowledge cannot come from sense experience

(2) Opinion is not knowledge

(3)Words must presuppose concepts

Foucher argues that the goal of philosophy is to find a criterion of truth with which to avoid error in judgment. The criterion can be used to obtain knowledge of the essence of things and to put this knowledge into a necessary order. But no criterion of truth can be adequate for attaining the absolute certainty that René Descartes seeks. Truth is basically human and fallible.

Foucher builds no system of his own; his talents are primarily critical. His method is that of the traditional skeptic: he assumes the suppositions of the system under analysis and then reasons by *reductio ad absurdum* to contradictory conclusions. But unlike the Pyrrhonian skeptics who wish to confute all knowledge claims, Foucher's Academic skepticism is meant to advance probable science and knowledge.

CRITICISMS OF MALEBRANCHE AND CARTESIANISM

Foucher claims that Descartes, to his credit, takes his rules of method from the Academics but that it is a major mistake on Descartes's part to assert that clear and distinct ideas can be certain and that they represent things external to one. Foucher follows Aristotle in professing that he cannot understand how one can have knowledge of the external world if no such knowledge comes through the senses. He further insists that both Descartes's claim that the knowledge of the essence of matter is innate and that knowledge of the properties of extension comes only through the reason, and not the senses, are unintelligible. Beyond this, Foucher makes four basic criticisms of Cartesianism.

First, Foucher argues that if mind and matter differ in essence, this allows no possibility of essential likeness between the two substances, which is necessary for causal interaction. Therefore, Cartesian mind and matter cannot interact.

Second, interaction between mind and matter obviously takes place, yet this interaction cannot be accounted for by Cartesian principles. Consequently, the Cartesians cannot know the true essences of mind and matter. The principle that likeness is necessary between cause and effect is self-evident, Foucher says, so mind and matter cannot be essentially different.

The third criticism concerns the ontological similarity between sensations and conceptual ideas, both of which are said by the Cartesians to be modifications of the mind. Both also are caused by the interaction of the mind with material things. However, ideas are said to represent objects external to the mind, whereas sensations do not. Foucher argues that if ideas are mental modifications representative of the material things that cause them, then why cannot sensations, which are also modifications of the mind, represent the material things that cause them? Or conversely, if sensations cannot represent material things, then how can ideas do so? This objection of Foucher's seems to be based on the Cartesian dictum that the cause of an idea must have at least as much formal or eminent reality as the idea has objective reality. Foucher argues that this means that it is necessary for there to be a likeness between the formal or eminent reality of the material thing and the objective reality of its idea. Because of this likeness, the material thing can cause the idea to resemble it and, hence, to represent it. Since sensations are caused by the same objects that cause ideas, why would sensations not also be like their causes, and hence representative of them? In this criticism Foucher basically ignores the Cartesian implication that conceptual ideas represent through description, not through resemblance, as sensations are ordinarily thought to do (although not by Descartes).

Fourth, if mind and matter are substances that differ in essence, then there can be no similarity or resemblance between them or their respective modifications. And, Foucher claims, it is obvious that if there is no resemblance, there can be no representation. Unextended ideas cannot represent extended material things or material modifications because ideas are mental modifications that can in no way resemble material things or material modifications. Hence, Cartesian ontology precludes an intelligible epistemology.

Such Cartesians as Rohault, Pierre-Sylvain Régis, Robert Desgabets, Louis de La Forge (1632–1666), and Antoine Le Grand (1629–1699) deny that ontological likeness or resemblance is necessary for an idea to represent its object. Foucher persists in asking for an explication of this nonresembling representation that is as intelligible as the notion that representation depends on resemblance, but he receives little more in reply than that God assures it. Foucher is himself a faithful Christian, but he insists against the Cartesians and Malebranche that declarations of faith in God's power and wisdom cannot be used as principles in philosophy.

Foucher takes Malebranche, as well as Descartes, to be saying that both sensations and ideas are modifications of the soul, which is a substance differing in essence from body. Malebranche denies that his ideas are mental modifications, but holds rather that they are beings in the mind of God. Foucher argues that Malebranchian ideas external to the mind, even if they are in the mind of God,

would be as difficult to know as are material objects external to the mind. Despite Malebranche's derision and that Foucher never takes him to be anything but a Cartesian, Foucher's criticisms bear on a vital point in Malebranche's system as well as in the systems of nonoccasionalist Cartesians. The epistemological failure of Cartesianism stems from the inability of Cartesians to give an explication of how ideas represent material things that is compatible with their dualist ontology.

CORRESPONDENCE WITH LEIBNIZ

In a correspondence noteworthy for the clearness with which each philosopher states his views, Leibniz agrees with Foucher that Academic principles are useful and that once in a lifetime a philosopher should follow his suppositions to their foundations. But Foucher insists that philosophy is primarily the examination and establishment of first principles, whereas Leibniz contends that very few philosophers are needed for this task; the important work is to follow out consequences in the development of knowledge. Foucher agrees that mathematics and hypothetical systems based on propositions of identity allow the deduction of truths internal to coherent systems, but he is concerned with the correspondence relation of these conceptual systems to the external world. Before a deductive natural philosophy is possible, it must be determined that the physical world is truly represented by one's concepts, axioms, and systems.

Extracts from the correspondence appear in the *Journal des Sçavans* from 1692 to 1696. In these Leibniz first places his new system before the public and Foucher gives it its first published critique. Foucher sees Leibniz's new system as little more than preestablished Malebranchian occasionalism, and he asks why God should go to such trouble to make it appear that mind and body interact if they really do not. Leibniz objects to occasionalism on the grounds that God should not continually be involved in making adjustments; Foucher argues that preestablished harmony, with all adjustments made at once, is no better. He says that Leibniz, like Malebranche, retains matter that is useless in his system because everyone experiences the interaction between mind and body. The task is to explain how interaction does take place, not merely how it seems to take place and how one can talk as though it does. For this, a monistic ontology in which mind and matter are metaphysically similar is required.

Foucher thus approves of Leibniz's denial of the Cartesian contention that extension is the essence of matter and his development of a monism of monads. The closest Foucher himself comes to outlining a monistic ontology is his suggestion to Leibniz that he should develop his ontology of monads to this end. Leibniz does not do this.

Foucher is not assured that any first principles apply to the world, and he criticizes Leibniz for building a system on uncertain foundations. Foucher reiterates that Descartes's criterion of certainty, clarity, and distinctness is useless and that the infallible mark of truth has not yet been discovered.

Foucher is important in the history of modern philosophy as a skeptic who originated epistemological criticisms that are fatal to the Cartesian way of ideas. Foucher's arguments against the distinction between ideas and sensations—that both are modifications of mind—were utilized by Pierre Bayle (*Dictionnaire historique et critique*, 5th ed. 1740, "Pyrrhon," Remark B), George Berkeley (*A Treatise concerning the Principles of Human Knowledge* 1710, 8–15; *Three Dialogues Between Hylas and Philonous* 1710, I), and David Hume (*A Treatise of Human Nature* 1739, I, IV, iv) to destroy the distinction between the primary qualities of size, shape, and position that John Locke says actually modify material bodies and the secondary qualities of sensible visual imagery, touch, taste, sound, and smell that Locke says do not modify bodies but are merely caused by them. The argument is that all these qualities are equally sensible.

See also Antiochus of Ascalon; Aristotle; Bayle, Pierre; Berkeley, George; Cartesianism; Descartes, René; Desgabets, Robert; Hume, David; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Philo of Larissa; Pyrrhonian Problematic, The; Régis, Pierre-Sylvain; Rohault, Jacques.

Bibliography

WORKS BY FOUCHER

Critique la recherche de la vérité, 1675.

De la sagesse des anciens, 1682.

Dissertation sur la reserche de la vérité, 1687.

GENERAL

Armour, Leslie. "Simon Foucher, Knowledge, and Idealism: Philo of Larissa and the Enigmas of a French 'Skeptic.'" In *Cartesian Views: Papers Presented to Richard A. Watson*, edited by Thomas M. Lennon, 97–115. Leiden, Netherlands: Brill, 2003.

Lennon, Thomas M. "Foucher, Huet, and the Downfall of Cartesianism." In *Cartesian Views: Papers Presented to Richard A. Watson*, edited by Thomas M. Lennon, pp. 117–128. Leiden, Netherlands: Brill, 2003.

Maia Neto, Jose R. "Foucher's Academic Cartesianism." In *Cartesian Views: Papers Presented to Richard A. Watson*,

edited by Thomas M. Lennon, 71–95. Leiden, Netherlands: Brill, 2003.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Watson, Richard A. *The Downfall of Cartesianism, 1673–1712: A Study of Epistemological Issues in Late Seventeenth Century Cartesianism*. The Hague: Nijhoff, 1965.

Watson, Richard A. "Foucher's Mistake and Malebranche's Break: Ideas, Intelligible Extension, and the End of Ontology." In *Nicolas Malebranche: His Philosophical Critics and Successors*, edited by Stuart Brown, 22–34. Assen/Maastricht, Netherlands: Van Gorcum, 1991.

Watson, Richard A. Introduction to *Critique de la recherche de la vérité* by Simon Foucher. New York: Johnson Reprint Corporation, 1969.

Watson, Richard A. "Introduction and Translation of Simon Foucher's *Critique [of Nicolas Malebranche's] of the Search for Truth*." In *Malebranche's First and Last Critics: Simon Foucher and Dortous de Mairan*, edited by Richard A. Watson and Marjorie Grene, 1–57. Carbondale: Southern Illinois University Press, 1995.

Richard A. Watson (1967, 2005)

FOUILLÉE, ALFRED

(1838–1912)

Alfred Fouillée, the French philosopher and sociologist, was a prolific writer, especially on political, social, and historical subjects. He was a lecturer in lycées at Douai and Montpellier, at the University of Bordeaux, and finally, from 1872 to 1875, at the École Normale in Paris. When he had to retire because of ill health, he devoted his time to his writings. Through most of his varied output there ran a common thread. This was a concern to reconcile the values of traditionally metaphysical or spiritualistic philosophy—above all, liberty and free will—with the deterministic and antimetaphysical findings of contemporary work in the natural sciences: a concern, that is, to reconcile philosophical idealism with scientific naturalism. Fouillée, who was not closely identified with any formal school of thought, thus represented a further step in the direction indicated before him by some of the later disciples of the spiritualistic school of Victor Cousin, such as Paul Janet and Étienne Vacherot, who aimed at absorbing or coming to terms with, rather than combating, the rising power of natural science and scientific philosophy.

Fouillée's outstanding and most original contribution to this enterprise was the idea that thought could lead to action, which he embodied in his concept of *idée-force,* or "thought force." This concept contains in itself the essence of Fouillée's consciously eclectic, conciliatory method and aim, for it borrows the notion of "force" from contemporary physical science and applies it to mental states, to consciousness. Force, defined as a tendency to action, becomes a universal fact of consciousness; conversely, every idea is a force that has a potential for realizing itself in action. Thus ideas, whether or not they are themselves caused, are causes; and since ideas are mental phenomena, mind is an efficient cause of physical action. The *idées-forces* are intermediaries between the private existence of consciousness and the objective existence of things. They enabled Fouillée to preserve spiritual values within the conditions imposed by natural science by developing what has been called a "positive metaphysics," that is, a metaphysics within the limits of the physically conceivable. Thus he undertook to refute the central tenet of materialism that mind or consciousness is merely an epiphenomenon. Specifically taking up the crucial concept of liberty, Fouillée argued that the consciousness of liberty amounts to the existence of liberty, since it gives rise to ideas formulated in terms of freedom of choice and since these ideas can in fact exert an effect on the outside world.

Fouillée's system is based primarily on psychological analysis, resembling, again, the spiritualism of the school of Cousin. This orientation was indicated by Fouillée himself when in his last work he labeled his philosophy "voluntaristic idealism." The will is the most immediate reality of consciousness, although not sharply separated from the intelligence or reason; ideas in Fouillée seem scarcely distinguishable from intentions. Yet, since he was attempting a comprehensive philosophical synthesis, Fouillée also constructed ontological categories on his psychological foundations. Causation, for instance, was established as an objective reality because it is one of the conditions necessary for the exercise of will, for the efficacy of the *idées-forces.* In like manner he developed an ethics with a strong social orientation. Consciousness, he taught, is aware not only of its own existence but of the consciousness of others (in this connection he suggested the emendation of René Descartes's famous dictum to read *Cogito ergo sumus*). Altruism is a necessity, since isolation is impossible; moral choice is explained in terms of the attractive or repulsive power of *idées-forces* in the form of ideals; and ethical conduct is defined in terms of social beneficence.

It is doubtful whether a system like Fouillée's, developed from a defensive posture, could ever prove generally acceptable. The concept of *idées-forces* is suggestive and useful as a tool of psychological analysis, but dubious if

elevated to the status of ontological reality. It is ultimately a merely verbal concept or device, seeking to bridge the gap between internal or mental processes and physical actions by, as it were, inserting a hyphen between them. But it will not bear the weight it is meant to carry, and as a result, the system as a whole remains merely suspended between idealism and naturalism. Though he struck a responsive chord and was widely read in his day, Fouillée had, in the end, few if any important followers.

See also Cousin, Victor; Descartes, René; Determinism and Freedom; Force; Idealism; Liberty; Naturalism.

Bibliography

Among the most important of Fouillée's works are *La liberté et le déterminisme* (Paris: Ladrange, 1872); *La science sociale contemporaine* (Paris: Hachette, 1880); *L'évolutionnisme des idées-forces* (Paris: F. Alcan, 1890); *Critique des systèmes de morale contemporains* (Paris: G. Baillière, 1883); *L'avenir de la métaphysique* (Paris: F. Alcan, 1889); *Psychologie des idées-forces,* 2 vols. (Paris, 1893), perhaps the central work; *Le mouvement idéaliste et la réaction contre la science positive* (Paris: F. Alcan, 1896); *Les éléments sociologiques de la morale* (Paris, 1905); *Morale des idées-forces* (Paris, 1908); *La pensée et les nouvelles écoles anti-intellectualistes* (Paris, 1911); *Esquisse d'une interprétation du monde* (Paris, 1913).

For literature on Fouillée, consult Augustin Guyau, *La philosophie et la sociologie d'Alfred Fouillée* (Paris, 1913), a laudatory view; Harald Höffding, *Moderne Filosofer* (Aarhus, 1904), translated by Alfred C. Mason as *Modern Philosophers,* pp. 82–88 (London, 1915); D. Parodi, *La philosophie contemporaine en France,* pp. 40–48 (Paris, 1919); and Elisabeth Ganne de Beaucoudrey, *La psychologie et la métaphysique des idées-forces chez Alfred Fouillée* (Paris: J. Vrin, 1936), a massive and judicious work.

W. M. Simon (1967)

FOUNDATIONALISM

See *Classical Foundationalism*

FOURIER, FRANÇOIS MARIE CHARLES

(1772–1837)

François Marie Charles Fourier, the French social critic, utopian socialist, and eccentric, was born into a merchant family in Besançon. Except during the French Revolution, Charles Fourier led a quiet and isolated life as a minor business employee and bachelor in Paris, Lyons, Rouen, and elsewhere in France, with occasional trips abroad. Shortly after the turn of the nineteenth century, Fourier began to develop his doctrine, publishing his first major work, *Théorie des quatres mouvements et des destinées générales,* in 1808. He continued throughout his life to elaborate and propagate his views with a single-minded devotion, acquired some followers, and was able to dedicate his last years entirely to his self-appointed task.

After a superficial classical secondary education in a Jesuit school in Besançon, Fourier was entirely self-taught. His reading was confined largely to contemporary periodicals and often apparently to bits of articles or merely to headlines. His views reflect many ideas of the Enlightenment and of the early nineteenth century, with strong Rousseauistic and physiocratic strains.

Fourier believed that, because the world had been created by a benevolent deity and yet wallowed in misery, men had obviously failed to carry out the divine plan. The plan was discovered by Fourier, and it had to be translated into practice. Happiness would then replace misery, unity would replace division, Harmony would replace Civilization. The transformation would occur through the release of man's thirteen passions, instilled by God but repressed in Civilization: the five senses; the four "group," or social, passions of ambition, friendship, love, and family feeling; the three "series," or distributive, passions, that is, the "cabalist," or passion for intrigue, the *papillone* (butterfly), or passion for diversification, and the "composite," or passion for combining pleasures; and, finally, the passion for harmony, which synthesizes all the others. With the passions released, existence would become intense joy, and a lifetime would seem but a moment.

To accomplish the release of the passions, humanity would have to be organized into phalanxes of about eighteen hundred men, women, and children. In each phalanx different characters and inclinations would be scientifically combined in a complex and finely graded system of groups and series so that each person could give full expression, in his work and in his other activities, to all his passions, tastes, and capacities, and avoid everything that did not suit him. The economies accomplished by communal work and living and by finding the right place for every talent, and the enormous enthusiasms and energies mobilized by the new order, would make phalanxes extremely successful economically as well as in terms of human happiness. Indeed, a single trial phalanx would prove its absolute superiority within a few weeks or, at the most, months and, through imitation, abolish Civilization in a year or two. Moreover, the savages and

barbarians who had stubbornly resisted Civilization would eagerly join Harmony. The result would be one world of happy phalanxes, linked vaguely by a hierarchy of monarchs and more effectively by temporary industrial armies for special tasks and similar touring bands of poets, actors, and musicians. Fourier's life became a constant search for the means to establish a trial phalanx, and his political, social, and other preferences were all subordinated to this one great purpose. Fourier believed his main enemies to be the *philosophes* of all sorts, with their "400,000 false volumes."

Fourier's ideas for transforming society were linked to peculiar views on man's past and to strange cosmological beliefs and "analogical" methods (Fourier argued "by analogy" in dealing with all elements of the cosmos). Because the world was one, the coming of Harmony would lead to new, beneficial creations on earth and would result in the appearance of new satellites, in the regaining of health by our planet, and in more distant desirable cosmic repercussions. At the moment, however, earth remained deplorably behind other planets, and Fourier hoped that sufficiently powerful telescopes would enable men to observe the system of Harmony as practiced by the Solarians or the inhabitants of Jupiter. Fantastic details of many kinds abound in Fourier's writings, and the very form of the writings is frequently bizarre.

In addition to giving rise to Fourierist communal experiments and anticipating cooperatives, Fourier has exercised a broad general influence as social critic, early socialist, and man of many insights, especially psychological ones. Fourier's criticism, appreciated by Karl Marx and Friedrich Engels among others, is notable for its fundamental character, its incisiveness, its richness, and its lack of compromise or nuance. It ranges from magnificent denunciations of exploitation and sham in family, society, church, and state, through striking discussions of fraudulent business practices (in particular of fraud in commerce, Fourier's bête noire), and of the appalling conditions of the masses, to a listing of dozens of different kinds of cuckoldry. Fourier was a moralist and believed that Harmony would establish truth as well as happiness among men, for truth rather than deception and hypocrisy would then become the profitable and accepted way of life.

Fourier's socialism is sui generis; he would have retained some private property, and he regarded inequality and discord as necessary for the construction of graded series and groups and the exercise of all passions. He stressed gastrosophy (the science of cuisine), opera, and horticulture rather than large-scale agriculture or industry. Far from desiring to mold man to a social purpose, he essayed to create a society where every individual whim would be satisfied. But Fourier did define man in social terms (the natural unit for lions, he said, is the couple, and for man, a phalanx, for only in a phalanx could man truly be man); and he charted an extremely complicated and interdependent socialist society, in which men own property, work, and live in common, in their specially built phalansteries, one for each phalanx.

This vision, together with his criticism of the existing system and many of his specific doctrines, places Fourier as one of the most inspired preachers and prophets of modern socialism. Fourier's remarkable psychological insights, such as his championing of brief sessions and variety in work, his quickness to see oppression no matter how veiled, and his at times penetrating concern with different character formations and problems, link him, for instance, to modern pedagogy, the emancipation of women, and personnel management. Fourier can also be described as a brilliant exponent of the idea of alienation or as a premature theoretician of the affluent society. Especially notable are his emphasis on the repression of passions as the source of all evil, as well as the foundation of Civilization, and his vision of that insane world of repressed passions.

See also Enlightenment; Philosophy of Social Sciences; Social and Political Philosophy; Socialism; Society.

Bibliography

PRIMARY WORKS

Théorie des quatre mouvements et des destinées générales, 2 vols. Lyons, France, 1808.

Théorie de l'unité universelle. Paris, 1822.

Le nouveau monde industriel et societaire. Paris: Bossange, 1829.

La fausse industrie morcelée, répugnante, mensongère, et l'antidote: l'industrie naturelle, combinée, attrayante, véridique, donnant quadruple produit, 2 vols. Paris: Bossange, 1835–1836.

Oeuvres Complètes de Charles Fourier, 12 vols. Paris: Anthropos, 1966–1968.

Utopian Vision of Charles Fourier; Selected Texts on Work, Love, and Passionate Attraction. Translated and edited by Jonathan Beecher and Richard Bienvenu. Boston: Beacon, 1971.

The Theory of the Four Movements. Edited by Gareth Stedman Jones and Ian Patterson. Cambridge, U.K.: Cambridge University Press, 1996.

SECONDARY WORKS

Beecher, Jonathan. *Charles Fourier: The Visionary and His World.* Berkeley: University of California Press, 1986.

Bourgin, Hubert. *Fourier Contribution à l'étude du socialisme français.* Paris: Société Nouvelle de Librairie et d'Édition, 1905.
Manuel, Frank E. *The Prophets of Paris.* Cambridge, MA: Harvard University Press, 1962.
Poulat, Émil. *Les cahiers manuscrits de Fourier.* Paris: Entente Communautaire, 1957.
Riasanovsky, Nicholas V. *The Teaching of Charles Fourier.* Berkeley: University of California Press, 1969.
Zilberfarb, I. I. *Sotsialnaia Filosofiia Sharlia Fure i Ee Mesto v Istorii Sotsialisticheskoi Mysli Pervoi Poloviny XIX Veka.* Moscow, 1964.

Nicholas V. Riasanovsky (1967)
Bibliography updated by Philip Reed (2005)

FRAME PROBLEM

A conundrum known as the frame problem within artificial intelligence concerns the application of knowledge about the past to draw inferences about the future. It requires distinguishing those properties that change across time against a background of those properties that do not, which thus constitute a frame (Charniak and McDermott 1985). From the point of view of philosophy it appears to be a special case of the problem of induction, which requests justification for drawing inferences about the future based on knowledge of the past. David Hume, in particular, suggested that one's expectations about the future are no more than habits of the mind and doubted that knowledge relating the future to the past was possible.

Bertrand Russell, a twentieth-century student of Hume's eighteenth-century problem, observed that this problem cannot be resolved merely by stipulation or by postulating that the future will be like the past. That the future will be like the past in every respect may be significant but it is also false. That the future will be like the past in some respect may be true but it is also trivial. The problem is to discover those specific respects in which the future will be like the past that provide justification for inferences to some outcomes rather than others, under the same initial conditions. That in turn suggests that the frame problem, like the problem of induction, depends for its solution on a defensible theory of natural laws that supplies a basis for linking the future to the past.

BACKGROUND

The first mention of a problem by this name was by John McCarthy and Patrick J. Hayes, who advanced a solution—the situational calculus—that depends on making assumptions about "the complete state of the universe at an instant of time," where "the laws of motion determine, given a situation, all future situations" (1969, p. 477). The reference to time raises concerns with relativity but, more important, not every feature of the universe makes a difference to every other feature at a later time. If one draws a distinction between global and local situations, where *global* concerns the complete state of the universe at a time and *local* only specific parts thereof, then local situations may prove tractable even if global situations should prove to be intractable.

Other characterizations of the problem include keeping track of the consequences of an action, including changes that they entail for representations of the world (Hayes 1973), and as a process of updating databases in response to changes that occur in the world (Barr and Feigenbaum 1981). Some claim it is not the problem of justifying inferences but of finding appropriate ways to express them (Hayes 1991), while others discuss the importance of the problem in relation to robots (Dennett 1984). As Robert Hadley notices, researchers in artificial intelligence tend to adopt narrower definitions of the problem, while philosophers tend to take the frame problem "to include any problem whose solution is *presupposed* by a solution to the narrow problem" (1988, p. 34). Some authors characterize the problem as less about knowledge than about knowledge representation.

WORLDS OF ROBOTS

The connection between actions, representations, and the problem of change arises in part from the desire to provide artificial human beings (or robots) with the directional capabilities to navigate their way around the world. If those robots act on the basis of maps—where the term is being used as a generic characterization for internal representations—then it becomes important to distinguish between permanent and transient features of those maps, which makes database updating important. And because robots may bring about changes in their environment through interaction, it becomes important to revise those maps to reflect those changes, to maintain the current relevance of those internal representations, where these concerns converge.

The same connections, however, also obtain for human beings as other things that act on the basis of beliefs as their internal representations of the world. When those beliefs are sufficiently accurate and complete, actions taken based on them may be expected to be more appropriate and less likely to fail than would otherwise be the case. Insofar as the frame problem revolves around knowledge of when things are going to change and when

they are not going to change, it possesses general significance for natural humans and for artificial humans alike. Beliefs are true when they correspond to reality (as the way things are or as everything that is the case), and when they correspond to reality they provide an appropriate foundation for human action as well as for robotic behavior.

The suggestion has been made that the frame problem concerns common sense as a product of everyday experience in interacting with the world, based on the fact that often the course of events conforms to one's expectations (Hayes 1991, p. 72). The existence of habits of the mind, however, does not resolve the problem with respect to justifying those habits on the basis of experience in the past nor explain how one's beliefs about the future ought to be represented. Presumably, the problem of knowledge must be resolved to have knowledge to represent. The kind of knowledge that holds promise for solving these problems derives from studying those features of the world that remain constant across time as the objects of scientific inquiries rather than as the products of common sense. These properties are known as the laws of nature.

LAWS OF NATURE

Laws of nature, unlike laws of society, cannot be violated or changed and require no enforcement. They must be distinguished from what are called accidental generalizations, which may be true as correlations that describe the history of the world but which could be violated and changed. If every Ferrari during the world's history happened to be red, then the generalization "all Ferraris are red" would be true, but it would not be a law, since there are processes and procedures, such as repainting a Ferrari, that would render it false. For a generalization to be lawlike, its falsity must be logically possible but not physically possible, precisely because there are no processes or procedures that could separate an attribute from its reference property, even though the possession of that attribute is not true merely as a matter of definition.

There appear to be several species of natural laws, including simple laws of nomic form and causal laws of different kinds (Fetzer 1981, 1990). That gold is malleable and that matches are ignitable are examples of simple laws, provided that those attributes are permanent. The selling price of gold, by comparison, at $500 an ounce, for example, is a transient attribute. These laws characterize properties that are possessed at one and the same time and do not explicitly imply changes across time. If the property of being malleable is a permanent property of gold, however, then gold has the causal properties that define *malleability*, including assuming different shapes at subsequent times as an effect of different forces at prior times. Thus, simple laws entail causal counterparts.

CAUSAL KINDS

The conception advanced by McCarthy and Hayes (1969), according to which complete states of the universe determine subsequent complete states according to laws of motion, presumes that those laws are exclusively deterministic, where given a complete state of the universe S1 at time t1, one and only one complete successor state S2 is physically possible at t2. Gottfried Wilhelm Leibniz and Pierre Simon de Laplace advanced similar conceptions. However, if any of the parts of the world are governed by causal processes that are indeterministic (or probabilistic), more than one successor state, S2, S3, … , Sn may be physically possible at time t2. Simple examples may include flips of coins, tosses of dice, and draws of cards from decks, but that depends on the specificity of the conditions attending those events.

Draws of cards from decks, for example, are ontically deterministic in the sense that, given specific arrangements of the cards in the deck, one and only one specific card can be drawn. These draws are epistemically indeterministic in the sense that, as long as one adheres to the rules of the game, one does not know the specific arrangements and is consequently unable to predict the outcome. The situation is different with the laws of radioactive decay, however. For example, an atom of polonium-218 has a half-life of 3.05 minutes, which means that, during any specific 3.05 minute interval, it has a probability of decay of one-half. This implies that, for collections of polonium atoms, one can expect that, during any 3.05-minute interval, about half will decay without knowing which ones.

TYPES OF SYSTEMS

Atoms of polonium-218 are closed systems for which there are no other properties that make any difference to their probability of decay than the length of temporal interval. Neither the weather, the day of the week, the presence or absence of observers—none of these factors affect the strength of this probability. In the case of flips of coins, tosses of dice, and draws of cards from decks there are other properties, such as the precise angular momentum imparted to a coin when flipped, which make it predictable with greater and greater precision, where condition F is relevant to outcome O under conditions C when it makes a difference to the probability out-

come O, given C. Increasingly precise specifications of the relevant conditions that affect outcomes thus allow instances of epistemic indeterminism to be established as ontically deterministic.

The probabilities of outcome depend on and vary with the complete sets of factors that are present on any specific occasion. When coins are bent, dice are loaded, or decks are stacked, the probabilities of various outcomes are no longer what they would have been under normal conditions. It follows that the truth of a lawlike sentence depends on taking into account the presence or absence of every property whose presence or absence makes a difference to the outcome on any specific occasion, which has been called the requirement of maximal specificity (Fetzer 1981, 1990). Closed systems are systems that satisfy this requirement, which is why their behavior across time can be systematically anticipated on the basis of corresponding maximally specific causal laws.

PREDICTION

For closed systems, it is therefore possible to predict—either invariably or probabilistically—precisely how that system will behave over an interval of time t to t* (when those properties are instantiated at time t and the outcome occurs at t*), so long as the laws of systems of that kind are known. When either (1) the laws of systems of that kind are not known or (2) the description available for that system is not closed, however, then precisely how that system would behave over a corresponding interval of time t to t* cannot be predicted with—invariable or probabilistic—confidence, because essential information remains unknown. In those cases the frame problem cannot be solved; but, even given knowledge of those kinds, the representation problem remains.

Indeed, there are at least two dimensions to the problem, where the first concerns whether the system under consideration qualifies as an open or as a closed system in relation to the outcome of interest. In either case, one needs to have predicates in one's language to describe each of its relevant properties. The second concerns whether or not the system under consideration, even if it happens to be a closed system, requires a finite or an infinite set of predicates for its complete description. When the complete description of states of the universe requires infinitely many predicates, for example, because infinitely many properties need to be described relative to successive states of the universe, there are no solutions to frame problems for global situations. Those are restricted to closed systems appropriately describable by finite sets of predicates.

SEMANTIC ISSUES

McCarthy and Hayes (1969) consider hypothetical situations that concern what would happen if specific events were to occur (such as the situation that would arise if Mr. Smith sold his car to Mrs. Jones, who has offered $250 for it). These situations are properly represented by subjunctive conditionals (concerning what would be the case, if something were the case) and counterfactuals (as subjunctive conditionals with false antecedents). However, this implies that, even envisioned primarily as a problem of representation, the solution to the frame problem entails solving some of thorniest issues in philosophical logic concerning intensional conditionals and possible-world semantics. A plausible solution involves distinguishing ordinary-language subjunctives from scientific conditionals elaborated in recent research (Nute 1975; Fetzer and Nute 1979, 1980).

The semantics that appears most appropriate for scientific conditionals and lawlike sentences is a form of maximal-change semantics rather than one of the varieties of minimal-change semantics proposed by Robert Stalnaker (1968) and by David Lewis (1979). Thus, while their semantics depend on assuming that possible worlds that differ from the actual world are as similar to the actual world as they could be, apart from the specific features being varied, the semantics assumed here—for the sake of exploring representational aspects of the frame problem—permits possible worlds to differ from the actual world in all respects except those specified by their maximally specific reference-property descriptions and the permanent properties that attend them. Subjunctives are true provided that, in every world in which their antecedents are true, their consequents are also true or would be true with constant probabilities.

LOGICAL FORM

An intensional calculus for the representation of lawlike sentences and causal conditionals of deterministic and probabilistic strength affords a possible framework for resolving the problem of representation (Fetzer and Nute 1979, 1980). Suppose that matches of kinds defined by chemical composition M are such that, when they are dry D, struck in fashion S, in the presence of oxygen O, then they light L. That could justify the lawlike claim, for every match x of kind M that is D and O, S-ing x at t1 would invariably bring about its L-ing at t2. That maximally specific antecedent could equally well be represented by various alternative formulations that included the same complete sets of relevant conditions, since adding oxygen when the other properties were present, for example,

would bring about the outcome just as the striking of the match, when those other properties were present, would bring it about.

Employing the double arrow, ___ ⇒ … , as the subjunctive conditional sign and the causal double arrow, ___ n⇒ … , as the (probabilistic) causal conditional sign—where values of n range over u for deterministic cases and p (from zero to one) for probabilistic cases—then these lawlike relations could be formalized by means of a generalized conditional, (x)(t)[(Mxt & Dxt & Oxt) ⇒ (Sxt u⇒ Lxt*)], which would be read, "For all x and all t, if x were M and D and O at t, then S-ing x at t would bring about (invariably) L-ing x at t*" (where t* is a specific interval after t). An instance of this generalization for a specific object c at a specific time t1 would have the following logical form, (Mct1 & Dct1 & Oct1) ⇒ (Sct1 u⇒ Lct2), which would have logically equivalent variations, such as (Mct1 & Dct1 & Sct1) ⇒ (Oct1 u⇒ Lct2), and so on.

SCOREKEEPING

The conception of conversational scorekeeping was introduced by Lewis (1973) as a helpful technique for keeping track of assumptions that have been made within the context of an ordinary conversation. Donald Nute (1980), for example, discusses its application relative to conditionals that occur during ordinary language conversations. Suppose, for example, that, at one point in their conversation, Bill and Hillary agree that either she will run for the Senate (again) or she will run for president. If they later conclude that she is not going to run for the presidency, they are entitled to infer that she is going to run for the Senate, even if that conversation occurs weeks later, assuming the premise has not been withdrawn.

Analogously, for a computerized system with the capacity for the representation of conditionals, such as LISP or Prolog, for example, developing programs that reflect the laws of systems that interest project managers should be relatively straightforward. No matter when specific data enters the program and regardless of the specific order in which it is received, once the antecedent of the conditional has been satisfied, the program draws the inference with deterministic certainty or probabilistic confidence that an outcome of kind O either has occurred or may be expected to occur, given the temporal parameters that apply. The function **cond** in LISP, for example, appears to be appropriately adaptable for this purpose (Wilensky 1984, Fetzer 1991). Hayes (1991) raises the objection that **cond** supports inferences of the form *modus ponens* but not of the form *modus tollens*, but that is sufficient for deriving predictions.

IMPLEMENTATION

It appears to be the case that the frame problem can be solved, at least in principle, for closed systems involving only finite sets of relevant properties. Whether or not it can be solved in practice, of course, depends on the state of science and one's knowledge of systems and laws of the kind under consideration. The solution that has been presented here, of course, presupposes an account of the nature of laws of which Hume would not have approved. Hume adopted an epistemic principle that precluded inference to the existence of properties and relations, including lawful and causal connections, that are not directly accessible to experience. His narrow form of inductivism cannot justify inferences to the existence of laws by contrast with mere correlations. Fortunately, a more robust epistemology based on inference to the best explanation accommodates the discovery of natural laws, where hypotheses are empirically testable by means of severe attempts to refute them (Fetzer 1981, 1990).

In spite of his emphasis on the representational aspect of the frame problem, even Hayes (1991) acknowledges that a theory of causation is a necessary condition for its solution. During the course of their review of a recent collection of studies of the frame problem, Selmer Bringsjord and Chris Welty (1994) suggest that the frame problem presupposes a solution to the problem of induction, which agrees with the position presented here. Whether or not the frame problem can be solved depends on whether or not the problem of induction can be solved, which in turn depends on deep issues in ontology and epistemology. If the considerations outlined earlier are well founded, however, then the problem of induction and the frame problem are both capable of successful resolution, even including its representational dimensions.

See also Artificial Intelligence; Computationalism; Connectionism; Induction; Laws of Nature.

Bibliography

Barr, Avron, and Edward A. Feigenbaum. *The Handbook of Artificial Intelligence*. Vol. 1. Reading, MA: Addison-Wesley, 1981.

Bringsjord, Selmer, and Chris Welty. "Navigating through the Frame Problem." *AI Magazine* (Spring 1994): 69–72.

Charniak, Eugene, and Drew McDermott. *Introduction to Artificial Intelligence*. Reading, MA: Addison-Wesley, 1985.

Dennett, Daniel. "Cognitive Wheels: The Frame Problem in AI." In *Minds, Machines, and Evolution*, edited by

Christopher Hookaway, 129–151. New York: Cambridge University Press, 1984.
Fetzer, James H. *Artificial Intelligence: Its Scope and Limits*. Dordrecht, Netherlands: Kluwer Academic, 1990.
Fetzer, James H. "The Frame Problem: Artificial Intelligence Meets David Hume." In *Reasoning Agents in a Dynamic World: The Frame Problem*, edited by Kenneth M. Ford and Patrick J. Hayes, 55–69. Greenwich, CT: JAI Press, 1991.
Fetzer, James H. *Scientific Knowledge*. Dordrecht, Netherlands: D. Reidel, 1981.
Fetzer, James H., and Donald Nute. "A Probabilistic Causal Calculus: Conflicting Conceptions." *Synthese* 44 (1980): 241–246.
Fetzer, James H., and Donald Nute. "Syntax, Semantics, and Ontology: A Probabilistic Causal Calculus." *Synthese* 40 (1979): 453–495.
Hadley, Robert. Review of *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, edited by Z. Pylyshyn. *Canadian Philosophical Reviews* 8 (1988): 33–36.
Hayes, Patrick J. "Commentary on 'The Frame Problem': Artificial Intelligence Meets David Hume." In *Reasoning Agents in a Dynamic World: The Frame Problem*, edited by Kenneth M. Ford and Patrick J. Hayes, 71–76. Greenwich, CT: JAI Press, 1991.
Hayes, Patrick J. "The Frame Problem and Related Problems in Artificial Intelligence." In *Artificial and Human Thinking*, edited by Alick Elithorn and David Jones, 45–59. New York: Elsevier, 1973.
Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.
McCarthy, John, and Patrick J. Hayes. "Some Philosophical Problems from the Standpoint of Artificial Intelligence." In *Machine Intelligence 4*, edited by Bernard Meltzer and Donald Michie, 463–502. New York: American Elsevier, 1969.
Nute, Donald. "Conversational Scorekeeping and Conditionals." *Journal of Philosophical Logic* 9 (1980): 153–166.
Nute, Donald. "Counterfactuals and the Similarity of Worlds." *Journal of Philosophy* 72 (1975): 773–778.
Pylyshyn, Zenon W., ed. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex, 1987.
Stalnaker, Robert. "A Theory of Conditionals." In *American Philosophical Quarterly Supplementary Monograph No. 2*, 98–112. Oxford, U.K.: Basil Blackwell, 1968.
Wilensky, Robert. *LISPcraft*. New York: Norton, 1984.

James H. Fetzer (2005)

FRANCK, SEBASTIAN

(1499–1542)

Sebastian Franck, also known as Franck von Word, was an outstanding figure among the spiritualists of the Reformation. His basic spiritualist concept of the conflict in each human being and in the world between the Inner Word (Son of God; eternal, invisible Christ), which is ultimate reality, and the outer word (law, flesh, selfishness), which is only appearance, shadow or phantom, was developed in all his philosophical, theological, historical, and cosmographical works. Franck was born in Donauwörth, Germany, and died in Basel. After studying at the University of Ingolstadt, Franck entered the Dominican Bethlehem College in Heidelberg in 1518. As a priest he officiated in the diocese of Augsburg. He turned to the Lutheran faith about 1526 and became Lutheran pastor in Buchenbach near Ansbach and then in Gustenfelden near Nürnberg. Franck resigned his pastorate in 1528 or early 1529 to become an independent writer and lived in Nürnberg until 1529 or early 1530.

Nürnberg, a cultural center, offered ample literary resources and personal contacts, especially with Theophrastus Paracelsus and the many followers of Hans Denck. Among Denck's followers were Albrecht Dürer's famous pupils the brothers Hans Sebald and Barthel Beham, whose sister Ottilie became Franck's wife. When Franck left Nürnberg, three of his controversial books were already written. Two of them were free translations from Latin into German (with many of his own unorthodox ideas injected) of Andreas Althamer's *Diallage* (1528), a Lutheran attack against Anabaptism, and of an unknown author's *Chronica und beschreibung der Türkey mit yhrem begriff* (Nürnberg, 1530), in which his ideas on the invisible church were already outlined. The first book wholly his own, *Von dem grewlichen laster der trunckenheyt* (1528), is a notable contribution to the literature on alcoholism.

From Nürnberg he moved to Strassburg, where he had occasion to meet Johann Buenderlin, Caspar Schwenckfeld, Melchior Hofmann, Jacob Ziegler, Michael Servetus, Johann Campanus, and, again, Paracelsus. There he wrote a universal history, extending from the creation of the world to the reign of Emperor Charles V and of Pope Clement VII, titled *Chronica, Zeytbuch und geschycht bibel* (Strassburg, 1531), famous for its numerous penetrating spiritualistic comments on many ecclesiastical and secular personalities and events. Its chronicle of heretics included Erasmus of Rotterdam as a Roman heretic. Because of this and his adverse remarks about Charles V, Franck was arrested and banned from Strassburg in 1531.

After living in Esslingen, Franck settled in Ulm as a printer and wrote most of his books there. His spiritualistic interpretation of the Scriptures can be found in his *Paradoxa ducenta octoginta …* (Ulm, 1534), *Die guldin Arch …* (Augsburg, 1538), and *Das verbütschiert mit sieben Siegeln verschlossen Buch* (Basel, 1539). *Die vier*

Kronbüchlein (Ulm, 1534) contains Erasmus's *Das thèur und kunstlich Buochlein Morie Encomion, das ist, ein Lob der Thorhait* and Cornelius Agrippa's *Von der Heylosigkeit Eitelkeit und ungewissheit aller Menschlichen Kunst und Weissheit* (both of which he freely translated from the original Latin texts), *Vom Baum des Wissens Guts and Böss …*, in which he tries to prove that awareness of good and evil can impair one's goodness, and *Encomion, ein Lob des Thorechten Gottlichen Worts …*. His *Weltbuch, Spiegel un bildtniss des gantzen Erdbodens …* (Tübingen, 1534), a cosmography with one of the first German descriptions of America and with one chapter dealing with the different religious movements of his time, which initiated systematic comparison of religion on Reformation soil, became one of his most popular books. His *Germaniae Chronicon* (Augsburg, 1538) has been used as an important source for historical research. In his *Das Krieg büchlein des Friedens …* (1539) Franck tried to prove that war not only contradicts Christ's teaching but is also "a devilish, inhuman thing, an abhorrent plague … an open door for all vices and sins and destruction of land, soul, body and honor." Most of these works made Franck the defendant in a trial before the city council that was instigated by Martin Frecht, main preacher in the cathedral of Ulm, Philipp Melanchthon, Martin Butzer, and Landgrave Philip of Hesse. It resulted in his expulsion from Ulm.

Franck, his wife, and their six children went to Basel in July 1539. There, after the death of his wife, he married Barbara Beck of Strassburg. His famous collection, with his interpretation, of *Sprichwörter …* (Frankfurt, 1541) was partially republished by G. E. Lessing. The last years of his life were devoted mostly to his Latin paraphrase of the *Deutsche Theologie,* which was never published, and to several posthumously published tractates (*Van het Ryke Christi,* Gouda, 1611; *Een Stichtelijck Tractaet van de Werelt des Duyvels Rijck,* Gouda, 1618; and *Sanctorum Communio,* Gouda, 1618), all of which survive only in the Dutch translations. They prove that dualism of God and world fully dominated his thoughts before his death.

Franck's worldview is primarily panentheistic, with heterogeneous elements drawn from Lutheranism, medieval mysticism, Neoplatonism, Renaissance speculation, humanism, Anabaptism, and rationalism, with ample citation of the Church Fathers and non-Christian philosophers. This comprehensive syncretism makes Franck an almost unique figure in the Reformation era and therefore a major figure in the history of ideas. As a religious philosopher he will be remembered for his radical spiritualistic tendency to replace exterior authority with inner illumination by God's spirit. The deep spiritual meaning of the Bible (outer word)—which is allegorical, not historical but typological, full of contradictions and merely testimonial to the eternal truth—can be comprehended only by those who have already accepted the Inner Word: "Unless we listen to the word of God within ourselves, we can make nothing of Scripture … for everything can be decked and defended with texts" (*Das verbütschiert mit sieben Siegeln verschlossen Buch*). In the light of his spiritualism none of the churches and sects, with their outgrown external disciplines, dogmas, sacraments, ceremonies, and festivals can be the true church. The true church is his *ecclesia spiritualis,* where only inward enlightenment is sufficient; it is the universal invisible church of the spirit, to which even those non-Christians who without knowledge of the incarnate Word have accepted the Inner Word can belong: "I love any man whom I can help and I call him brother, whether he be Jew or Samaritan … I cannot belong to any separate sect" (ibid.).

As a historian Franck placed the Reformation in the stream of historical development and thus relativized it. He is credited with recognizing the historic force that externalizes the spiritual ("The world must have a papacy even if it has to steal it."). He also observed the typical recurrent rise and fall of kingdoms and peoples, and by recognizing this change of fortune as God's punishment for disobedience of his Inner Word, saw history as interaction between God and the world, as the struggle between the spirit and the forces which resist it.

As one of the most ardent advocates of religious liberty in the sixteenth century, Franck insisted on toleration not only among the individual members of the different churches and sects in Christendom but also toward Jews, Muslims, heathens, and even heretics, since all men, created by God, descended from Adam, and accessible to the Holy Spirit, are equal.

Martin Luther, Philipp Melanchthon, and Martin Butzer were especially aware of the danger of Franck's unorthodox thoughts to the new Protestant position. Luther called him "the devil's most cherished slanderous mouth." The convention of Protestant theologians at Schmalkalden in 1540 issued a resolution of condemnation of both Franck and Schwenckfeld, which the latter called a (Protestant) papal bull.

Franck's extraordinarily well written books had a great influence on German prose style. They were widely read in German, Dutch, Swiss, and even English editions until the end of the seventeenth century. There exist at least ten editions of his *Chronica* and as many of his

Sprichwörter. His *Weltbuch* went through at least six editions, as did his *Vier Kronbüchlein* or parts of it.

While Franck's specific traceable influence was restricted in Germany to Valentin Weigel and Gottfried Arnold, and in Basel to Sebastian Castellio, his spirit and ideas found ardent followers in Holland (Dirk Volkerts Coonhert, Menno Simons, David Joris, and the Franckists or Sebastianists). Although he had strong roots in the late Middle Ages, much of Franck's thought carried the seed of what was to become important in modern thinking. Wilhelm Dilthey rightly testifies that "the ideas of Franck flow toward modern times in a hundred streamlets."

See also Dilthey, Wilhelm; Emanationism; Humanism; Ideas; Lessing, Gotthold Ephraim; Luther, Martin; Medieval Philosophy; Melanchthon, Philipp; Neoplatonism; Paracelsus; Reformation; Servetus, Michael; Toleration.

Bibliography

ADDITIONAL WORKS BY FRANCK

Klagbrieff oder supplication der armen dürftigen in Englandt an den könig daselbst gestellt wider die reychen geystlichen Bettler. Strassburg, 1529.

Dass Gott das einig ain und höchstes Gut sey …. Ulm, 1534.

Sechshundertdreyzehn Gebot und Verbot der Juden. Ulm, 1537.

Des Grossen Nothelffers und Weltheiligen Sant Gelts oder S. Pfennings lobgesang …. Ulm, 1537.

Was gesagt sei: Der Glaub tuts alles: Und warumb im die Rechtfertigung alleyn werde zugeschriben …. Ulm, 1539.

Schrifftliche und gantz grundtliche ausslegung des LXIIII Psalm. …. Ulm, 1539.

Handbüchlein Siben Haubtpunken aus der Bibel gezogen und zusammengebracht, darin angezeigt ist leben und todt, Himmel und Hell … durch Sebastian Franck gemacht. …. Frankfurt, 1539.

"A Letter to John Campanus by Sebastian Franck, Strassburg, 1531." Translated by G. H. Williams, in *Spiritual and Anabaptist Writers.* Library of Christian Classics, Vol. XXV. Philadelphia: Westminster Press, 1957.

WORKS ON FRANCK

Bainton, R. H. "Concerning Heretics (Castellio)." In *Records of Civilization XXV,* 93–104. New York, 1935. Republished New York, 1966.

Barbers, Meinulf. *Toleranz bei Sebastian Franck.* Bonn: Röhrscheid, 1964.

Becker, Bruno. "Nicolai's Inlasshing over de Franckisten." *Nederlandsch Archief voor Kerkgeschiedenes,* n.s. 18 (1925): 286–296.

Dilthey, Wilhelm. "Auffassung und Analyse des Menschen im 15. und 16. Jahrhundert." In his *Gesammelte Schriften,* Vol. II. Stuttgart and Göttingen, 1957.

Gritsch, E. W. "The Authority of the Inner Word. A Special Study of the Major German Spiritual Reformers in the 16th Century." Dissertation. Yale University, 1959.

Hayden-Roy, Patrick. *The Inner Word and the Outer World: A Biography of Sebastian Franck.* New York: Peter Lang, 1994.

Hegler, Alfred. "Beitrage zur Geschichte der Mystik in der Reformationszeit," in *Archiv für Reformationsgeschichte,* Vol. I. Berlin, 1906.

Hegler, Alfred. *Geist und Schrift bei Sebastian Franck.* Freiburg im Breisgau, 1892. The most important work on Franck.

Hegler, Alfred. *Sebastian Francks lateinische Paraphrase der deutschen Theologia und sein holländisch erhaltenen Traktate.* Tübingen, 1901.

Jones, R. M. *Spiritual Reformers in the 16th and 17th Centuries,* 46–64. London: Macmillan, 1914. Republished, Boston: Beacon, 1959.

Kaczerowsky, Klaus. *Sebastian Franck: Bibliographie.* Wiesbaden, Germany: G. Pressler, 1976.

Koyré, Alexandre. *Mystiques, spirituels, alchimists du XVI^e siècle allemand,* 22–43. Paris, 1955.

Krahn, Cornelius, and N. van der Zijpp. "Sebastian Franck." In *The Mennonite Encyclopedia,* Vol. II, 363–367. Scottdale, PA, 1956.

Muller, Jan. *Sebastian Franck, 1499–1542.* Wiesbaden, Germany: Otto Harrassowitz, 1993.

Oncken, Hermann. "Sebastian Franck als Historiker." In *Historisch-Politische Aufsätze und Reden,* Vol. I, 273–319. Munich and Berlin: Oldenbourg, 1914.

Peuckert, W. E. *Sebastian Franck, ein deutscher Sucher.* Munich: Piper, 1943. The only recent interpretation of all aspects of Franck's thought.

Quast, Bruno. *Sebastian Francks 'Kriegbüchlin des Frides': Studien zum radikalreformatorischen Spiritualismus.* Tübingen, Germany: Francke, 1993.

Racber, K. *Studien zur Geschichtsbibel Sebastian Francks.* Basel: Helbing and Lichtenhahn, 1952.

Raumer, Kurt von. *Ewiger Friede,* 23–60, 249. Freiburg im Breisgau: Alber, 1953.

Rieber, Doris. "Sebastian Franck (1499–1542)." *Bibliothèque d'humanisme et Renaissance* 21 (1959): 190–204.

Schottenloher, Karl. *Bibliographie zur deutschen Geschichte im Zeitalter der Glaubensspaltung,* Vol. I, 263–266. Leipzig: Hiersemann, 1933. Lists all books on Franck to 1932.

Teufel, Eberhard. "Die 'Deutsche Theologie' und Sebastian Franck im Lichte der neueren Forschung." *Theologische Rundschau,* n.s. 11 (1939): 304–319, and 12 (1940): 99–129.

Teufel, Eberhard. *Landraümig; Sebastian Franck, ein Wanderer an Donau, Rhein und Neckar.* Neustadt an der Aich: Degener, 1954.

Teufel, Eberhard. "Luther und Luthertum im Urteile Sebastian Francks." In *Festgabe für D. Dr. Karl Müller zum 70. Geburtstag dargebracht von Fachgenossen und Freunden,* 132–144. Tübingen, 1922.

Troeltsch, Ernst. *The Social Teaching of the Christian Churches,* Vols. I and II. Republished, London: Allen and Unwin, 1956.

Weigelt, Horst. *Sebastian Franck und die lutherische Reformation.* Gütersloh, Germany: Gerd Mohn, 1972.

Williams, G. H. *The Radical Reformation.* Philadelphia: Westminster Press, 1962.

Jacques J. Whitfield (1967)
Bibliography updated by Christian B. Miller (2005)

FRANK, ERICH
(1883–1949)

Erich Frank studied philology and classics at the universities of Vienna, Freiburg, and Berlin. In 1907 he turned to philosophy, which he studied in Heidelberg under Heinrich Rickert and Wilhelm Windelband. His philosophical career in Germany was brief but distinguished. In 1923 he became professor at Heidelberg, and five years later he was appointed Martin Heidegger's successor in Marburg. Three years after his dismissal from Marburg in 1936, he came to Harvard on a research fellowship and made America his second home. Almost all of Frank's works reflect his double interest in philosophy and history and his efforts to combine historical knowledge and philosophical thought: *Plato und die sogenannten Pythagoreer* (Halle, 1923); *Wissen, Wollen, Glauben* (Knowledge, will, belief), a collection of English and German historical and speculative essays, edited with an appreciation by Ludwig Edelstein (Zürich, 1955), of which the title essay represents Frank's most original contribution to philosophy; *Philosophical Understanding and Religious Truth* (New York, 1945).

As a student, Frank felt dissatisfied with current attempts to model philosophy on science and to eliminate the traditional questions of metaphysics, ontology, and religion. Nor was he long satisfied with the post-Kantian idealism that was offered as an alternative and which for a time attracted him. When in 1914 he discovered Søren Kierkegaard, at that time almost unknown in philosophical circles, he thought he had found the beginning of a new and fruitful approach to the problem of the subject-object dialectic. He shared his discovery with Karl Jaspers, and five years later, with the publication of Frank's essay *Wissen, Wollen, Glauben* and Jaspers's *Psychologie der Weltanschauungen* the foundations of German existentialism were laid. The major theme of Frank's essay is that the unity of the subject in self-consciousness is achieved not in the act of knowing or in the act of willing, but only in the act of faith. This knowledge of the self is logically unprovable but is also incontrovertible. The act of faith is neither blind belief nor a "will to believe," but arises out of the immediate awareness both of oneself as free and of a transcendence of oneself. Faith is thus both the condition and the result of the subject's freedom, and all theoretical and practical activity has its source in this freedom. Frank believed he had found in the act of faith the unity of the subject that Immanuel Kant sought but could not find in the act of judging.

Later Frank came to question the subjective direction in which existentialist philosophy was developing. In his review of Jaspers's *Philosophie* (1933) he not only criticized what he called the "*atheistischer Nihilismus*" (atheistic nihilism) of Heidegger, but also pointed out the insufficiency of Jaspers's existential ontology (the *Chiffre*), which, he claimed, bears no analogical relation to Being. Existentialism, he argued, has not succeeded in combining existential concerns with metaphysical objectivity. The freedom of the subject is not threatened by his encounter with the objective. Indeed, that freedom which does not express, historically and analogically, a truth concerning objective Being, is an empty, irrational freedom.

In *Philosophical Understanding and Religious Truth,* which grew out of the Flexner Lectures he gave at Bryn Mawr in 1943, Frank considered the question of analogical terms through which alone, in his view, philosophy could adequately express the subjective, existential experience of objective reality. All philosophical truth, he argued, is analogical in that it recounts, in and for each historical period, the relation of man to Being. Philosophical analogy is possible only because there is an objective reality to which our thinking bears a relation. Just as in *Knowledge, Will, Belief* Frank argued that the freedom of the subject always presupposes a transcendence of it, so he now maintained that philosophical thought presupposes an object beyond itself that is its content and its substance. Thus, philosophy shares with religion the belief that there is an objective reality to be known; the task of philosophy is, in part, the rational elucidation of religious truths. However, philosophy must not take the place of the revealed mystery of religion. In every historical period philosophical truths have a different starting point and find a different expression, but their content—Being—is eternal. Philosophical analogy is possible only because there is Being; and Being becomes part of our thinking only in analogy. The purpose of philosophy is to present in rational terms the existential dialectic of the subjective and the objective, the temporal and the eternal.

See also Being; Existentialism; Faith; Heidegger, Martin; History and Historiography of Philosophy; Jaspers, Karl; Kierkegaard, Søren Aabye; Rickert, Heinrich; Windelband, Wilhelm.

Bibliography

In addition to the works cited above, the reader may wish to consult Frank's "Das Prinzip der dialektischen Synthesis und die Kantische Philosophie," in *Kant-Studien,* Ergänzungsheft

No. 21 (1911), and his "Mathematik und Musik und der griechische Geist," in *Logos* 9 (2) (1920): 222–259. See also his editions of Fichte's *Die Anweisung zum seligen Leben* (Jena, 1910) and of the so-called *Nachtwachen von Bonaventura* by Clemens Brentano (Heidelberg, 1912), and his literary and philological studies of Schelling and Brentano in *Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Philosophische–Historische Klasse,* 1 Abh. (1912); and *Germanisch–Romanische Monatsschrift* 4 (1912): 417–440.

Eva Gossman (1967)

FRANK, SEMËN LIUDVIGOVICH

(1877–1950)

Semën Liudvigovich Frank, the Russian philosopher and religious thinker, was trained in law at Moscow University (1894–1898) and in economies and philosophy at the universities of Berlin and Munich (1899–1902). As a student in Moscow he was a member of a Marxist group headed by P. B. Struve; his first published work was a critique of Karl Marx's theory of value (1900). Between 1902 and 1905 (during which years he moved back and forth between Moscow and Germany) he was a principal contributor to Struve's journal *Osvobozhdenie* (Liberation), published in Stuttgart.

Frank joined a number of other young ex-Marxist intellectuals—among them Struve, Nikolai Berdiaev, and Sergei Bulgakov—in publishing three important symposium volumes: *Problemy idealizma* (Problems of Idealism; Moscow, 1903); *Vekhi* (Signposts; Moscow, 1909); and *Iz glubiny* (De profundis; Moscow, 1918). This last work was printed but because of Soviet censorship was never released.

In 1906 Frank settled in St. Petersburg; in 1912 he joined the Russian Orthodox Church and began to teach philosophy at St. Petersburg University. In 1915 (at St. Petersburg) he published, and in 1916 defended, his master's thesis, *Predmet znaniia* (The object of knowledge); in 1917 he published his doctoral dissertation, *Dusha cheloveka: Opyt vvedeniia v filosofskuyu psikhologiyu* (Man's soul: An introductory essay in philosophical psychology; Moscow), but was unable to defend it because of political events. From 1917 to 1921 Frank was professor of philosophy and dean of the newly organized faculty of history and philosophy at Saratov University. In 1921 he was named professor of philosophy at Moscow University. He was among the group of non-Marxist intellectuals expelled from the Soviet Union in the summer of 1922. He settled in Berlin, where he gave university lectures (in German) on Russian literature and culture. In 1937, forced to leave Germany, he moved to France. In 1945 he moved to London, where he died.

From Vladimir Solov'ëv—and ultimately from Plotinus—Frank took his central doctrine of positive "total-unity" (*vseedinstvo*). His epistemological intuitivism was close to that of his older colleague Nikolai Losskii. His characteristic emphasis on the "metalogical unity" of the real, and its transcendence of the Aristotelian laws of thought, was drawn mainly from Nicholas of Cusa. Frank always identified himself as a Platonist.

Although Frank's thought exhibits many Hegelian strands, and although he regularly used terms like *moment* (*das Moment*) in G. W. F. Hegel's special sense (as "dialectical phase" or "component of a totality"), he employed one crucial pair of terms in a very un-Hegelian way. To the absoluteness of *real'nost'* (reality) he opposed the relativity of *deistvitel'nost'*, not "actuality" in the sense of Hegel's *Wirklichkeit* (the common meaning of *deistvitel'nost'* in Russian philosophy) but the merely empirical or factual. Frank distinguished between conceptualizable and objectifiable "factuality" and the nonconceptualizable, metalogical "dual-unity" (*dvuedinstvo*) of "reality." The real is fully related and concrete; the factual is isolated and abstract: "Being is a total-unity, in which everything particular exists and is conceivable only in its relation to something else" (*Nepostizhimoe* [The unknowable], p. 51). We apprehend reality as a "monodual" coincidence of opposites, as both "distinct from all particular determinate contents" of knowledge and as "containing and permeating" every such content (*Real'nost' i chelovek* [Reality and man], pp. 93–94). The real is both "transdefinite" and "transfinite"; and in both respects it eludes conceptualization. "Everything finite," Frank declared, "is given against a background of infinity. … The knowable world is surrounded on all sides by the dark abyss of the unfathomable" (*Nepostizhimoe*, pp. 29, 35).

Frank agreed with René Descartes that, although the term *finite* is prior and positive in meaning, and the term *infinite* is derived from it by negation, "it is precisely the infinite as the 'fullness of all' that is given as primary and positive, while the concept of finitude is formed by negation of that fullness" (*Real'nost' i chelovek*, p. 57).

Forms, or "ideal elements," are determinate aspects of factuality. The totality of such determinations is grounded in what Frank, following Solov'ëv, called the primordial unity or total-unity of the real. Although reality is unfathomable, it is not hidden; rather, it is "entirely

evident, being mysterious only in the sense that it is inexplicable, irreducible to anything else, and inaccessible to logically analytic thought. It is what Johann Wolfgang von Goethe called *ein offenes Geheimnis*" (*Real'nost' i chelovek*, p. 78). "Objective factuality" is something alienated and abstract, a "rationalized, i.e., logically crystallized, part of reality." Like a nut's shell, it forms a "hard and relatively distinct outer layer, produced by the inner saps and energies of a living organism" (pp. 106ff.).

Reality in its wholeness is graspable only in the integral intuition of "living knowledge" (*zhivoe znanie*), of which conceptual knowledge is only a derivative product or superstructure: "All particular knowledge is partial knowledge of a whole."

The "I" of the Cartesian *cogito* is a "reality in which subject and object coincide—a "self-revealing" and "self-transparent being-for-itself," accessible to "living knowledge." Sounding rather like Martin Heidegger, whose general position he repudiated, Frank wrote, "We are conscious of ourselves only as a self-revelation of [being] in us" (*Nepostizhimoe*, p. 93). He also offered a more emphatic version of Heidegger's doctrine of *Mitsein*: "No finished 'I' exists prior to the encounter with the 'thou.' … It is in this encounter … that the 'I' in a genuine sense first comes into being" (pp. 148, 154). Frank also suggested Heidegger's category of impersonal "itness" (*das Man*): "The 'we' appears in the form of an 'it' … which constitutes the basis and first source of objective being" (p. 177). Although there can be no "I" apart from its relation to a "thou," "every 'I' has a special root of its own, lying in secret depths inaccessible to others" and "the most essential part of me remains solitary and inexpressible." The more one is aware of oneself as a person, the more one withdraws into "metaphysical solitude," for "we are wholly open only to ourselves and to God" (*Real'nost' i chelovek*, pp. 127, 129).

In religious—and especially mystical—experience, "I encounter God as a 'thou' for me, only in … that ultimate and essentially solitary stratum of my 'I' in which I am … inaccessible to everyone except myself—and God (as Kierkegaard rightly insisted). I encounter God in the utter solitude in which I encounter death" (pp. 215f.).

Like Solov'ëv, Frank generalized the notion of "Godmanhood" (Solov'ëv's term was *Bogochelovechestvo*; Frank's somewhat more abstract term, *Bogochelovechnost'*) beyond its Christian context. Its primary reference is not to the Incarnation, but to the basic ontological category of "divine-human reality." In Frank's words, "The dual-unity of Godmanhood is logically prior to the conceptions of both God and man" (p. 249).

See also Berdyaev, Nikolai Aleksandrovich; Bulgakov, Sergei Nikolaevich; Descartes, René; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heidegger, Martin; Kierkegaard, Søren Aabye; Losskii, Nikolai Onufrievich; Marx, Karl; Nicholas of Cusa; Platonism and the Platonic Tradition; Plotinus; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY FRANK

Man's Soul: An Introductory Essay in Philosophical Psychology. Translated by Boris Jakim. Athens: Ohio University Press, 1993.

Reality and Man: An Essay in the Metaphysics of Human Nature. Translated by Natalie Duddington. London: Faber and Faber, 1965.

The Spiritual Foundations of Society: An Introduction to Social Philosophy. Translated by Boris Jakim. Athens, Ohio: Ohio University Press, 1987.

The Unknowable: An Ontological Introduction to the Philosophy of Religion. Translated by Boris Jakim. Athens, Ohio: Ohio University Press, 1983.

WORKS ON FRANK

Boll, M. M. *The Social and Political Philosophy of Semen L. Frank: A Study in Prerevolutionary Twentieth-Century Russian Liberalism*. Madison, WI: author, 1970.

Boobbyer, P. *S. L. Frank: The Life and Work of a Russian Philosopher, 1877–1950*. Athens, OH: Ohio University Press, 1995.

Lossky, N. O. *History of Russian Philosophy*. New York: International Universities Press, 1951. Pp. 266–292.

Nazarova, O. *Ontologicehskoe obosnovanie intuitivizma v filosofii S. L. Franka* (The ontological grounding of intuitivism in S. L. Frank's philosophy). Moscow: Ideia-Press, 2003.

Swoboda, P. *The Philosophical Thought of S. L. Frank*. Ph.D. Diss., University of Columbia, 1992.

Zenkovsky, V. V., *Istoriya Russkoi Filosofii*. 2 vols. Paris, 1948, 1950. Translated by G. L. Kline as *A History of Russian Philosophy*. 2 vols. London and New York: Columbia University Press, 1953. Pp. 852–872.

Zenkovsky, V. V., ed., *Sbornik Pamyati Semëna Lyudvigovicha Franka* (A collection of essays in memory of Simon Ludvigovich Frank). Munich, 1954. Reminiscences and critical essays by Zenkovsky and others. Contains a bibliography of Frank's writings, compiled by L. A. Zander, on pp. 177–192.

BIBLIOGRAPHY

Frank, V. *Bibliographie des oeuvres de Simon Frank*. Paris: Institut d'études slaves, 1980.

George L. Kline (1967)
Bibliography updated by Vladimir Marchenkov (2005)

FRANK, SIMON LYUDVIGOVICH

See *Frank, Semën Liudvigovich*

FRANKFURT, HARRY

(1929–)

Harry Frankfurt grew up in Brooklyn, New York, and Baltimore, Maryland. He received his PhD in philosophy from Johns Hopkins University in 1954, and he taught in the philosophy departments at Ohio State University; State University of New York, Binghamton; Rockefeller University; Yale University; and Princeton University.

Frankfurt has made original and important contributions to various fields in philosophy, including history of modern philosophy (primarily René Descartes), philosophical psychology, and moral philosophy. He has explored such issues as the relationship between moral responsibility and free will, the nature of the self, the role of necessitation or inevitability in both constraining and constituting persons, and central phenomena such as care, love, and truth. His work has exerted a significant influence on philosophers working in these areas, and some of his writings (especially on the role of love and truth in our lives) have been read by a wide audience. It is perhaps not surprising that Frankfurt's work has been appreciated beyond the walls of academia, as it is both penetrating and elegant.

In one of his most influential papers, "Alternate Possibilities and Moral Responsibility," Frankfurt argued that moral responsibility does not require the sort of free will that entails *alternative possibilities* or genuine freedom to do otherwise. He offered a template for a kind of example that calls into question the *Principle of Alternative Possibilities*, (PAP), according to which moral responsibility requires alternative possibilities. The *Frankfurt-Style Counterexamples* (to PAP) have a distinctive structure that involves *preemptive overdetermination*, that is, the existence of a fail-safe device that plays no role in the causal sequence that issues in the relevant behavior, but which renders that behavior inevitable.

The examples can be seen to be extensions of an example presented by John Locke. Locke discussed a man who is transported into a room while asleep. When he awakens, the man considers whether to leave the room, but stays for his own reasons. Unbeknownst to him, the door was locked and thus he could not have successfully left the room. According to Locke, the man stayed in the room voluntarily although he could not have left the room.

Now it might be pointed out that although the man in Locke's example lacked a certain alternative possibility (the power to leave the room), he nevertheless had various important options available, including choosing to leave, trying to leave, turning the knob, and so forth. Frankfurt's distinctive contribution is the addition of a component to this sort of example which, as it were, *brings the locked door into the agent's brain*. That is, Frankfurt asks us to imagine someone who can secretly monitor an agent, even his brain activities; as things happen, no intervention by this kind of shadowy *counterfactual intervener* occurs. But if the agent were about to choose to do otherwise, this would trigger some process by which the intervener—say, a *nefarious neurosurgeon*—could ensure that the agent choose and behave as he actually does. Thus, Frankfurt has provided a more sophisticated version of Locke's example, one in which it is at least plausible to suppose that the agent in question chooses and acts freely and could legitimately be held morally responsible even though the agent literally could not have chosen otherwise and could not have done otherwise.

Frankfurt thus denied PAP. One who agrees with Frankfurt can thus contend that one of the main objections to compatibilism about causal determinism and moral responsibility can be blocked. That is, it is traditionally supposed that causal determinism threatens moral responsibility because it rules out the sort of free will that involves alternative possibilities; but if this sort of free will is not required for moral responsibility, then at least this sort of objection to compatibilism is rendered irrelevant. Of course, there may be other reasons to reject compatibilism. Frankfurt himself is officially agnostic about compatibilism, saying that we cannot be confident that causal determinism is compatible with being *active*, and thus we cannot be confident in the truth of the compatibility of causal determinism and moral responsibility.

In another seminal paper, "Freedom of the Will and the Concept of a Person," Frankfurt suggested that the distinctive feature of persons is a certain characteristic structure in their motivational states. We share preferences, beliefs, and so forth with mere animals. But we are unique in that we can step back from our preferences and form *second-order* preferences—preferences about our first-order preferences. Some of these second-order preferences are what Frankfurt called *second-order volitions*—the preference that a certain first-order preference lead one to act. According to Frankfurt it is not crucial what the basis for the second-order reflection is; it need not be

moral deliberation, for example. On his view, persons are distinctive in that they have the capacity to form second-order volitions; thus, they are the sort of entities for which freedom of the will can be a problem.

For Frankfurt, it is important to distinguish such notions of freedom to choose otherwise and freedom to do otherwise, on the one hand, from notions such as choosing freely and acting freely, on the other. The former involve alternative possibilities whereas the latter do not. In "Freedom of the Will and the Concept of a Person," Frankfurt gives an account of acting freely in terms of the hierarchical account of the structure of human motivation. When one acts freely, one acts on the preference one really wants to have as one's *will* (roughly, the actually motivating preference). In Frankfurt's terminology when one acts freely, one *identifies* with one's will, that is, one identifies with the first-order desire that actually motivates one to act. In contrast, one does not act freely when one does not identify with one's will—one acts (say, smokes another cigarette or eats another piece of chocolate cake) despite identifying with other first-order desires). Frankfurt suggests, additionally, that identification consists in forming the relevant second-order volition; he suggests that one identifies with a first-order desire insofar as one forms a second-order preference to be motivated by that first-order desire. So, acting freely consists in a kind of *mesh* or harmony in the hierarchical structure of one's mental economy. Of course, the existence of this synchronization of levels is entirely compatible with the agent's lacking alternative possibilities.

In further work Frankfurt has refined the analysis of the crucial notion of identification in light of various problems. Additionally, whereas the early papers were primarily addressed to issues pertaining to freedom, determinism, and moral responsibility, the later papers exhibit an evolution toward questions about the *true self*. In "Identification and Wholeheartedness," Frankfurt concedes that mere formation of the relevant second-order volition is not sufficient for identification, and he provides a more refined analysis, including the important notion of *decisive commitment* or *decision*. In a later paper, "The Faintest Passion," Frankfurt adds the component of *satisfaction* to the analysis of identification. The notion of identification is important both to the account of acting freely and the true self, and it is interesting to ask whether the same notion can play the required roles in both accounts.

Not only is a certain sort of inevitability (lack of alternative possibilities) compatible with moral responsibility, Frankfurt contends that certain *volitional necessities*—things we simply cannot bring ourselves to will—help to constitute the boundaries of our true selves. In a series of papers Frankfurt explores the way in which our selves are formed through the process of caring, identification, and volitional constraints. In "The Importance of What We Care About," Frankfurt identifies *caring* as a distinctive kind of motivation importantly different from morality. He denies that all-things-considered rationality needs to coincide with the deliverances of morality. In later work Frankfurt has built on his work on caring to give a nuanced account of the nature of love. For Frankfurt, love is central to the foundations of morality as well as to the formation of our selves.

Central themes in Frankfurt's work are as follows: the compatibility of moral responsibility, caring, and love with certain sorts of necessity or inevitability and the contention that morality, normativity, or rationality should not be *built into* our analyses of human motivation at the very foundational level. For Frankfurt, caring and love are more central or, perhaps, more fundamental notions than rationality and morality.

See also Descartes, René; Determinism and Freedom; Ethics, History of; Locke, John; Love; Responsibility, Moral and Legal; Truth.

Bibliography

Buss, S., L. and Overton, eds. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, MA: MIT Press, 2002.

WORKS BY FRANKFURT

"Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66 (1969): 829–839.

Demons, Dreamers, and Madmen. Indianapolis, IN: Bobbs-Merrill, 1970.

"Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1971): 5–20.

"The Importance of What We Care About." *Synthese* 53 (1982): 257–272.

"Identification and Wholeheartedness." In *Responsibility, Character, and the Emotions*, edited by F. Schoeman, 27–45. New York: Cambridge University Press, 1987.

The Importance of What We Care About. New York: Cambridge University Press, 1988.

"The Faintest Passion." *Proceedings of the American Philosophical Association* 66 (1992): 5–16.

Necessity, Volition, and Love. New York: Cambridge University Press, 1999.

The Reasons of Love. Princeton, NJ: Princeton University Press, 2004.

On Bullshit. Princeton, NJ: Princeton University Press, 2005.

John Martin Fischer (2005)

FRANKLIN, BENJAMIN
(1706–1790)

Benjamin Franklin, the U.S. statesman, scientist, and author, was born in Boston, where he attended school for less than a year. He learned the printer's trade, and at seventeen he ran away to Philadelphia. After two years in England (1724–1726) he returned to Pennsylvania, where, prospering in his trade, he began publishing the *Pennsylvania Gazette* in 1729 and *Poor Richard's Almanack* in 1732. He had already formed a tradesman's self-improvement club, the Junto, and soon began civic and educational promotions, including the founding of the American Philosophical Society.

Franklin retired from business in 1748, turned to science, and in 1751 published *Experiments and Observations on Electricity*. The same year he entered the Pennsylvania Assembly, where he was a leader in opposing the influence of Proprietor Thomas Penn and in advocating colonial union. In 1757, as agent for the assembly, he went to England, where, except for eighteen months, he lived until 1775, enjoying English society and the friendship of David Hume, Henry Home (Lord Kames), Richard Price, and other British philosophers. At first he worked loyally for the expansion of the British Empire and sought to exchange proprietary for royal government in Pennsylvania, but after 1765 he became the leading colonial spokesman in resisting British measures in North America. Although he opposed every act of oppression, he sought until the very end to reconcile differences; but in 1775 he returned home, signed the Declaration of Independence, and worked for a united war effort. In 1776 he went to France, where he signed the French Alliance (1778), secured loans and supplies for the Revolutionary War, and helped negotiate the Treaty of Paris (1783).

He was lionized by Voltaire, Madame Helvétius, Marquis de Condorcet, La Rochefoucauld d'Enville, and other *philosophes*, and returned home in 1785. He served for three years as president of the Pennsylvania Executive Council, attended the Constitutional Convention of 1787, sought the abolition of slavery, and worked on his *Autobiography* in the five years preceding his death.

Franklin's greatest popular fame is as a moralist. The aphorisms of Poor Richard and the example of his *Autobiography* have served as a philosophy of life for millions. In these two works Franklin sought deliberately to set down the rules of conduct that would enable anyone, however humbly born, to prosper and live more meaningfully. The emphasis was unashamedly on the mundane virtues: thrift, hard work, diligence, prudence, moderation, honesty, and shrewdness. For this, Franklin has been denounced by D. H. Lawrence and others as a "snuff-colored man" who impoverished life by "fencing it in" with a stifling, despiritualizing morality. In fact, Franklin knew the precepts of Poor Richard were but a partial philosophy; in his own career and in his other writings he showed abundantly how full and imaginative human life can be.

Like many deists of his day, Franklin believed "in one God, Creator of the Universe, that he governs it by his Providence … [and] that the soul of Man is immortal" (letter to Ezra Stiles, March 9, 1790).

As a scientist, Franklin formulated important and influential laws concerning the nature of electricity. By proving that lightning is an electrical discharge, he placed electricity beside heat, light, and gravity as one of the primordial forces in the universe and hypothesized a new dimension or quality possessed in some measure by all matter. Characteristically, Franklin turned readily from electrical theory to a useful invention, the lightning rod. His scientific attitude is summarized in the statement "Let the experiment be made," and in the observation that electrical experiments would "help to make a vain man humble."

As a public philosopher, Franklin assumed that the traditional personal values have political relevance. He shared the Aristotelian belief that government exists for the sake of the good life and that its powers can be used to that end. A good citizen, guided by the virtues Franklin encouraged in *Poor Richard's Almanack* and in his *Autobiography*, would undertake civic improvement and participate disinterestedly in government. In an expanding country filled with opportunity, Franklin saw individual initiative as the essential engine of progress, but he did not hesitate to seek whatever seemed required for the public good through government. His confidence in the virtue of the citizens of the United States caused him to favor government by consent, but he was not a simple democrat who believed majority will should be omnipotent. He accepted democracy because he thought it would yield good government; if it did not, he readily rejected it.

Franklin thought freedom's dynamism would cause its spread around the world, and therefore that the United States, as a leading free nation, would be influential without being predatory. At the same time he understood the anarchic character of international relations and counseled the nation to maintain its strength, protect its national interest, and act to maintain a balance between France and Great Britain. His essential faith was that,

from tradesmen's juntos to the court of Versailles, good men working together could improve the condition of humankind.

See also Aristotelianism; Condorcet, Marquis de; Deism; Home, Henry; La Rochefoucauld, Duc François de; Price, Richard; Voltaire, François-Marie Arouet de.

Bibliography

For Franklin's writings, see L. W. Labaree and others, eds., *The Papers of Benjamin Franklin*, 37 vols. (New Haven, CT: Yale University Press, 1959–) and *The Autobiography of Benjamin Franklin* (New Haven, CT: Yale University Press, 1964); and A. H. Smyth, ed., *The Writings of Benjamin Franklin*, 10 vols. (New York: Macmillan, 1905–1907).

For his life and thought, see Carl Van Doren, *Benjamin Franklin* (New York: Viking Press, 1938); Carl Becker, *Benjamin Franklin* (Ithaca, NY: Cornell University Press, 1946); and R. L. Ketcham, *Benjamin Franklin* (New York: Washington Square Press, 1965).

On his scientific thought, see I. B. Cohen, *Franklin and Newton, an Inquiry into Speculative Newtonian Experimental Science and Franklin's Work in Electricity as an Example Thereof* (Philadelphia: American Philosophical Society, 1956). For his political thought, see R. L. Ketcham, ed., *The Political Thought of Benjamin Franklin* (Indianapolis: Bobbs-Merrill, 1965). For his place as an Enlightenment philosopher, see Frank L. Mott and Chester E. Jorgenson, eds., introduction in *Representative Selections* (New York: American, 1936).

C. L. Sanford, ed., *Benjamin Franklin and the American Character* (Boston: Heath, 1955), reprints a good collection of critical essays on Franklin.

OTHER RECOMMENDED WORKS

Aldridge, Alfred Owen. *Benjamin Franklin and Nature's God*. Durham, NC: Duke University Press, 1967.

Lemay, J. A. Leo. *The Oldest Revolutionary: Essays on Benjamin Franklin*. Philadelphia: University of Pennsylvania Press, 1976.

Wright, Esmond. *Franklin of Philadelphia*. Cambridge, MA: Belknap Press of Harvard University Press, 1986.

Ralph Ketcham (1967)
Bibliography updated by Michael J. Farmer (2005)

FREEDOM

In the history of philosophical and social thought "freedom" has a specific use as a moral and a social concept—to refer either to circumstances that arise in the relations of man to man or to specific conditions of social life. Even when so restricted, important differences of usage are possible, and most of the political or philosophical argument about the meaning or the nature of freedom is concerned with the legitimacy or convenience of particular applications of the term.

ABSENCE OF CONSTRAINT OR COERCION

It is best to start from a conception of freedom that has been central in the tradition of European individualism and liberalism. According to this conception, freedom refers primarily to a condition characterized by the absence of coercion or constraint imposed by another person; a man is said to be free to the extent that he can choose his own goals or course of conduct, can choose between alternatives available to him, and is not compelled to act as he would not himself choose to act, or prevented from acting as he would otherwise choose to act, by the will of another man, of the state, or of any other authority. Freedom in the sense of not being coerced or constrained by another is sometimes called negative freedom (or "freedom from"); it refers to an area of conduct within which each man chooses his own course and is protected from compulsion or restraint. J. S. Mill's essay *On Liberty* is perhaps the best-known expression in English of this individualistic and liberal conception of freedom.

Some writers take the view that the absence of coercion is the sufficient and necessary condition for defining freedom; so long as a man acts of his own volition and is not coerced in what he does, he is free. Other writers wish to widen the concept in one or both of two ways. They argue that natural conditions, and not only the will or the power of other men, impose obstructions and restraints on our capacity to choose between alternatives and that therefore the growth of knowledge or anything else that increases our capacity to employ natural conditions for the achievement of our purposes ipso facto enlarges our freedom. They also sometimes argue that whether or not it is the will of other men or natural obstacles that are considered as limiting or constraining our actions, we cannot truly be said to be free to choose some preferred alternative unless we have the means or the power to achieve it, and thus the absence of means or power to do *X* is equivalent to absence of freedom to do it. For those who take this view the necessary conditions for the existence of freedom would be (*a*) the absence of human coercion or restraint preventing one from choosing alternatives he would wish to choose; (*b*) the absence of natural conditions preventing one from achieving a chosen objective; (*c*) the possession of the means or the power to achieve the objective one chooses of one's own volition. Many of the assertions frequently made about liberty in

recent political thought assume that possession of the means or power to realize preferred objectives is part of what it means to be free. For example, the contention that men who suffer from poverty or have a low level of education cannot really be free, or that they cannot be as free as the well-to-do and the well educated, relies on the assumption that "to be free to do *X*" includes within its meaning "to be able," "to have the means," and "to have the power" to do *X*.

What are the objections to thus connecting "being free to" with "having the capacity or the power to"? It can be said that, at least in many cases, equating freedom with possession of power will involve a distortion of ordinary language. If I ask, "Am I free to walk into the Pentagon?" the question will be clearly understood; but if I ask, "Am I free to walk across the Atlantic Ocean?" the appropriate answer will be "You are free to, if you can." This suggests the main argument: The linking of "being free to" with "having the capacity or power" deprives the word *free* of its essential and unequivocal function, which is to refer to a situation or state of affairs in which a man's choice of how he acts is not deliberately forced or restrained by another man. As Bertrand de Jouvenel points out, if we say that to be free to achieve chosen ends requires the possession of the power and the social means necessary for their achievement, then the problem of freedom coincides with (or becomes confused with) the quite different problem of how satisfactions are to be maximized. It may be true to say that the poor man is as free to spend his holidays in Monte Carlo as the rich man is, and true also to say that he cannot afford to do so. These two statements, it is argued, refer to two distinct states of affairs, and nothing is gained by amalgamating them.

MEANING OF "COERCION"

Even if we confine ourselves to saying that a man is free insofar as his action is not coerced by another, it is evident that the concept of coercion itself requires some consideration. An important point may be made by examining Bertrand Russell's often-quoted sentence: "Freedom in general may be defined as the absence of obstacles to the realization of desires." This hardly goes far enough. Let us imagine an authoritarian society in which rulers have for years been so successful in controlling and manipulating what members of the community read and what views they encounter, and in which the educators have been able so subtly and skillfully to mold the minds and dispositions of the very young, that almost all citizens naturally desire what their rulers desire them to desire, without its ever occurring to them that there are alternatives to what they are accustomed to or that their freedom to choose has been in any way circumscribed. They are not conscious of any obstructions to the satisfaction of desire and, indeed, no obstructions may exist to the satisfaction of any desires they experience. This is a limiting case, but it points to conditions that exist more or less in all societies. We would scarcely concede that the members of such a society enjoyed any or much freedom. The society described may be one in which coercion in the usual sense does not occur and has in fact become unnecessary.

Two important points follow from this. First, if absence of coercion is a necessary condition of being free, coercion must be understood as including not only the direct forms—commands or prohibitions backed by sanctions or superior power—but also the many indirect forms—molding and manipulation or, more generally, forms of control that are indirect because they involve control by certain persons of the conditions that determine or affect the alternatives available to others. This is an important extension of the notion of coercion. Second, if liberty means the right of individual choice between alternatives, then this right in turn implies that the alternatives can be known by those who are to choose; that individuals have the opportunity to understand the character of available alternatives and can make a deliberate or informed choice. The freedom that members of a society enjoy will be connected, therefore, with the extent to which competing opinions, objectives, modes of behavior, ways of living, and so on are, so to speak, on display; on how freely they can be recommended, criticized and examined; and thus on the ease with which men can make a deliberate choice between them.

For this reason, since literacy or education enlarges the capacity or faculty of choice and decision, it is an important precondition of the existence of freedom: knowledge extends the capacity for acting freely. Similarly, not only suppression but also distortion and misrepresentation, any kind of dishonest propaganda that gains its effect from privileged control over sources of publicity, may restrict the freedom of others; insofar as it succeeds in concealing or misrepresenting the character of certain of the available alternatives, it will tend to restrict or manipulate the range of choice no less effectively than direct coercion or constraint may; and thus it will also tend to limit the exercise of freedom in a particular society. It is not sufficient to consider only the presence or absence of coercion in the more literal and direct sense. Freedom in its positive aspect is the activity or process of choosing for oneself and acting on one's own

initiative, and choice can be manipulated as readily as it can be coerced.

Does it follow from this that the extent of freedom is related to the number of available alternatives, in that the more alternatives there are for choice, the freer a man is? Clearly there can be no simple or direct relationship between the range of available alternatives and the extent of freedom. However numerous the alternatives between which a man may choose, he will not admit himself to be free if the one alternative that he would most prefer is the one that is excluded. In a society that forbids the preaching of Catholic doctrine and the practice of Catholic forms of worship, Catholics will not concede that they are free just because they are still free to be either Anglicans, Methodists, or Buddhists. In certain circumstances the extent of the range of available alternatives may be relevant to a judgment of the extent of freedom; but in general we can talk profitably about both the existence and the extent of freedom in a particular society only by taking into account the individual and social interests, the capacities, the modes of behavior, and the ways of living on behalf of which freedom is claimed.

KINDS OF FREEDOM

When men speak of their being free or claim freedom for themselves, they are referring not only to the absence of coercion and restraint imposed by others (freedom *from*) but also to that on behalf of which freedom is being claimed (what they are claiming freedom *for*). This is another sense in which we can speak about a positive aspect of freedom. In political and social discussion a claim to freedom is almost invariably (albeit usually implicitly) a claim to a particular liberty, a claim to freedom for or in the exercise of some particular interest or form of activity. Although Russell says that freedom is the absence of obstacles to the satisfaction of desire, probably no serious philosophical or social thinker has defended freedom in the sense of absence of obstacles to the satisfaction of *any* desire; what has been defended, and what freedom has been identified with, is the absence of obstacles to the exercise and satisfaction of specific interests and forms of activity that are accepted as possessing special moral and social significance.

Thus, freedom in the abstract is a class comprising many species—freedom of thought and speech, freedom of association, freedom of assembly, freedom of worship, freedom of movement, freedom in the use or disposal of one's property, freedom in the choice of one's employer or occupation, and so on. In every case there is, of course, a reference to the absence of coercion or interference and to an area within which one can choose or act on one's own initiative; not to an abstract or indeterminate possibility of choosing but instead to a specific sphere of individual or social activity within which the right to make one's own choices and decisions, to follow one's own course, is regarded as being of particular importance in the moral life of the individual. This seems to be one way in which positive notions of freedom (as contrasted with the more abstract idea of bare immunity from coercion or interference by others) have emerged, namely, in the attempt to identify (and thus to identify with freedom) those specific spheres of human activity within which what Mill calls individuality, the right and capacity for individual choice and initiative, really matter.

Some of the particular freedoms that have been much emphasized in recent times (freedom from want and freedom from fear are important examples) seem at first sight to refer neither to the absence of coercion nor to any specific interest or form of activity for which freedom is being claimed. It might appear that what *is* being claimed is, rather, the institution of political and economic arrangements by means of which men may be made immune from feelings and circumstances that they find to be evil. If this is all that is meant, then this is to employ freedom in a sense different from the one we have been discussing; this is shown by the fact that freedom from want and fear could conceivably be attained by the setting up of political and social arrangements under which the amplitude of choice within important spheres of activity would be drastically restricted and under which there might be a considerable measure of coercion and constraint; in other words, freedom from want and freedom from fear might well be compatible with a very authoritarian regime, just as in contemporary China freedom from flies is said to have been achieved by very authoritarian methods. Thus, if "freedom from want" and "freedom from fear" are taken simply in that way, the freedom involved is logically and socially distinct from that which has so far been taken as being central and fundamental in the tradition of liberal thinking. However, this may be to interpret these two freedoms superficially. For a more sympathetic interpretation we must return to what has been said about manipulation.

FREEDOM AND POWER

In modern societies manipulation in various forms is at least as important as the processes we normally identify as coercive. It is well known that, within a society, a group of men may enjoy such control over property or the means of production, or over an educational system or

the media of communication, that they are able to determine within a fairly narrow range the alternatives between which their fellow citizens can choose. It is not only true that less privileged men often lack the means or the power to attain their preferred alternative but also that others can exploit their lack of power in order to prevent them from attaining what they would wish to attain; sometimes the less powerful can even be prevented from knowing what alternatives there are and from knowing that some of them might be capable or worthy of being pursued. It is this argument that can justify notions like "freedom from want" or "freedom from economic insecurity" and that links them with what has been taken to be the central sense of freedom, the absence of constraint. Even though we refuse to conclude that the mere absence of the means or the power to attain a preferred alternative goal is equivalent to not being *free* to pursue it, it is a different situation when means and power are controlled and manipulated by others in order to secure compliance with their demands. Thus, if "want" and "insecurity" describe a condition in which there is unequal control over the means and conditions of choice and action, in consequence of which some men can manipulate the range of choice available to others, then freedom from want and insecurity belongs with freedom from coercion; in that case, freedom from want and insecurity is the condition of the ability to act on one's own initiative, which is the positive side of liberty.

There is, then, this connection between freedom and power: When there is conflict between individuals and groups for possession or control of scarce means and conditions of action, control over means is a condition of the availability of alternatives, and hence of choice and freedom. It follows, therefore, that when men have unequal power, this will often mean that they will also be unequal with respect to the freedom they enjoy—not merely in the sense that the man who is better off has the means to choose more widely and live more abundantly than his poorer brother (although this is also true) but in the more relevant sense that the more powerful man can restrict the range of choice and the freedom of the less powerful in order to satisfy his own interests more fully. Obviously this relation between inequality of power and inequality of freedom provides one of the connections that exist between liberty and democracy. If we define democracy as being a form of political organization in which all adult members of the community share in making decisions about the common arrangements of the society (including those decisions about the use and distribution of the resources that affect the choices of acting available to men), then the right to participate in the making of these decisions is a liberty that will affect (or at least may very substantially affect) the range and character of the alternatives that are available in very important areas of social and private life.

POLITICAL PARTICIPATION. Thus, we may say that political participation, or sharing in the process of government, will enter into the meaning of "liberty" in society in at least two different ways. First, political activity and participation in government is an interest and mode of activity to which many men attach great importance, and thus the existence of the right and opportunity to engage in this form of activity is one of the liberties that some men cherish highly. Second, it is in addition a liberty that forms part of a wider structure of liberties because the extent to which this liberty is accorded and exercised will usually also affect the extent to which liberty is available in other areas of social life. This is not to say, of course, that the more democratic a society is (the less men are restrained or restricted in their participation in the activity of government), the more freedom there will be in other areas of social life; it is possible for democracies to be exceptionally coercive, restrictive, or intolerant in certain areas of living and, apart from this, it is also true that expansion of particular liberties (or of liberty in particular areas) often entails the curtailment of others. The point is, rather, that political liberty in the sense specified forms part of a more complex system of liberties in any developed society; both logically and causally, political liberty is connected with the liberties that are established in other spheres of individual activity.

FREEDOM AND CHOICE

We have seen that liberty has its negative and its positive sides—"negative" referring to the absence of obstructions, interference, coercion, or indirect control; "positive," to the processes of choosing and acting on one's own initiative, and more concretely and less formally to the general types of human interests or forms of activity for the expression and exercise of which liberty is claimed. Some writers, concentrating particularly on the positive aspect, have been inclined to assert that a man is being free only when he is actually choosing, exercising initiative, and acting deliberately or responsibly. Mill, in what he says in *On Liberty* about "individuality," "individual spontaneity," the "despotism of custom," and related matters, comes very close to asserting this, although he never quite does so. The same kind of view is hinted at in Graham Wallas's "Freedom is the capacity for continuous initiative," but it would be difficult to accept

this as a general position. For the devotee of a religious faith, the religious freedom he claims and believes himself to enjoy may be no more than the freedom to practice unmolested a form of worship he has inherited and which he has never felt the faintest temptation to question; in such a case it is a fiction to speak of a process of choice. The same can be said of the man who is content to follow narrowly, uncritically, and unadventurously the established customs and conventions of his society. Even though there may be a sense in which we can intelligibly talk of such men as being slaves to customs, habits, or orthodoxies, it would still be straining the point to maintain that they are not free.

On the other hand, the man who has been so molded and manipulated that he always wants what his ruler or superior wants him to want is scarcely free. This case suggests that freedom will exist only where there exists the *possibility* of choice, and the possibility of choice in turn implies not only the absence of direct coercion and compulsion but also that the availability and the characteristics of alternatives must be capable of being known. Thus, whatever the situation of any particular individual may be, it is most likely that there will be a large measure of individual freedom within a society when there exists what Mill calls a variety of conditions—where a wide variety of beliefs are in fact expressed and where there is a considerable diversity of tastes and pursuits, customs and codes of conduct, ways and styles of living. And, because of the connection between inequality of power and inequality with respect to the enjoyment of freedom, a society in which power is widely distributed is also likely to be the one characterized by the existence of wide possibilities for choice and individual initiative.

See also Authority; Censorship; Democracy; Determinism and Freedom; Liberalism; Liberty; Mill, John Stuart; Power; Rights; Russell, Bertrand Arthur William.

Bibliography

Adler, M. J. *The Idea of Freedom,* 2 vols. New York, 1958–1961; Westport, CT: Greenwood, 1973.

Bay, Christian. *The Structure of Freedom.* Stanford, CA: Stanford University Press, 1958.

Berlin, Isaiah. *Two Concepts of Liberty.* Oxford: Clarendon Press, 1958.

Cranston, Maurice. *Freedom: A New Analysis.* London: Longmans, Green, 1953.

Friedrich, C. J. *Man and His Government.* New York: McGraw-Hill, 1963.

Fuller, Lon. "Freedom: A Suggested Analysis." *Harvard Law Review* 68 (1955): 1305–1325.

Hayek, F. A. *The Constitution of Liberty.* London, 1960.

Jouvenel, Bertrand de. *Sovereignty: An Inquiry into the Political Good.* Chicago: University of Chicago Press, 1957.

Knight, Frank. *Freedom and Reform.* New York: Harper, 1947.

Malinowski, Bronislaw. *Freedom and Civilisation.* London: Allen and Unwin, 1947; Westport, CT: Greenwood, 1976.

Mill, J. S. *On Liberty.* London: Parker, 1859.

Oppenheim, F. E. *Dimensions of Freedom.* New York: St. Martin's, 1961.

Russell, Bertrand. "Freedom and Government." In *Freedom: Its Meaning,* edited by Ruth N. Anshen. New York: Harcourt Brace, 1940.

P. H. Partridge (1967)

FREE WILL

See *Determinism and Freedom*

FREGE, GOTTLOB

(1848–1925)

LIFE

After studying mathematics, physics, chemistry, and philosophy at the universities of Jena and Göttingen, the German mathematician, logician, and philosopher Gottlob Frege obtained his mathematical doctorate in Göttingen (1873) and his mathematical *habilitation* in Jena (1874). From 1874 to 1879 he taught mathematics at the University of Jena as a lecturer; in 1879 he was promoted to adjunct professor, and in 1896 to associate professor. Frege never obtained a full professorship. He retired from teaching in 1917 because of illness, becoming emeritus in 1918.

While he received little professional recognition during his lifetime, Frege is widely regarded in the early twenty-first century as the greatest logician since Aristotle, one of the most profound philosophers of mathematics of all times, and a principal progenitor of analytic philosophy. His writing exhibits a level of rigor and precision that was not reached by other logicians until well after Frege's death.

MAIN WORKS

In the monograph *Begriffsschrift* (1879) Frege introduces his most powerful technical invention, nowadays known as predicate logic. In his second book, *Die Grundlagen der Arithmetik* (1884), he discusses the philosophical foundations of the notion of number and provides an informal

argument to the effect that arithmetic is a part of logic (a thesis later known under the epithet *logicism*). The pamphlet *Funktion und Begriff* (1891) is an elucidation of Frege's fundamental ontological distinction between functions (with concepts as a special case) and objects; certain difficulties with the views expressed therein are discussed in the essay "Über Begriff und Gegenstand" (1892). Frege's most celebrated achievement in the philosophy of language, the distinction between the sense and the reference of an expression, is expounded in his landmark essay "Über Sinn und Bedeutung" (1892). *Grundgesetze der Arithmetik* (volume 1, 1893; volume 2, 1903), his magnum opus, constitutes his abortive (because of Bertrand Arthur William Russell's antinomy) attempt at rigorously proving the logicist thesis. The essay "Der Gedanke: Eine logische Untersuchung" (1918) is a conceptual investigation of truth and that with respect to which the question of truth arises (called thoughts by Frege).

FREGE'S LOGIC

By replacing the traditional subject-predicate analysis of judgments with the function-argument paradigm of mathematics and inventing the powerful quantifier-variable mechanism, Frege was able to overcome the limitations of Aristotelian syllogistics and created the first system of (higher-order) predicate logic. He thereby devised a formal logical language adequate for the formalization of mathematical propositions, especially through the possibility of expressing multiply general statements such as "*for every* prime number, *there is* a greater one."

The first presentation of his *begriffsschrift* (concept script—Frege's logical formula language) is contained in the 1879 monograph by the same name. At this time, the linguistic and philosophical underpinnings of begriffsschrift, as well as the description of the language itself, are still somewhat imprecise. There are, for instance, no formation rules given for the formulas of the language; functions seem to be identified with functional expressions; the meanings of the propositional connectives are specified in terms of assertion and denial rather than truth and falsity; and although Frege officially countenances only one inference rule, namely, *modus ponens*, he tacitly uses an instantiation rule for the universal quantifier as well. The first volume of *Grundgesetze*, however, presents a mature and amazingly rigorous version of the system, taking into account the various insights Frege had developed since the publication of *Begriffsschrift*. Unless otherwise noted, the following discussion pertains to this later system; for the time being, one should ignore the course-of-values operator, which is discussed later on in connection with Russell's antinomy.

The primitive symbols of Frege's begriffsschrift are then those for equality, negation, the material conditional, and the first- and higher-order universal quantifiers. In addition, there are gothic letters serving as bound variables (of first and higher orders), as well as Latin letters, whose role one would today characterize as that of free variables (again, of various orders). Disjunction, conjunction, and the existential quantifier are neither primitive, nor are they introduced as abbreviations, as would be customary today; rather, Frege notes that they can be simulated by means of the existing primitives.

Frege carefully distinguishes between basic laws (axioms) on the one hand, and inference rules on the other hand. With respect to a specified set of basic laws and rules of inference, he comes close to a rigorous definition of derivations in the predicate calculus.

The logical connectives, as well as the quantifiers, are taken to be denoting expressions, having as references the requisite truth functions and higher-order functions, respectively. Equality undergoes a radical change in interpretation between the time of *Begriffsschrift* and that of *Grundgesetze*. In the earlier system, assuming that the expression A refers to the object a, and the expression B to object b, Frege construes identities of the form $A = B$ metalinguistically, taking them to mean that the expressions A and B are coreferential, rather than that a and b are the same object. In *Grundgesetze*, however, identity is conceived of as a binary relation between objects, much as is standard today (this change in interpretation is, incidentally, accompanied by a switch in notation from the triple bar $\equiv$ to the now customary double bar $=$). Arguably, there is an analogous shift in the understanding of the universal quantifier; the formulations in *Begriffsschrift* suggest that it is to be interpreted substitutionally, whereas it is fairly clear in *Grundgesetze* that an objectual interpretation is intended. But the issue is difficult to judge, not only because the language of the earlier work is rather imprecise but also because it is not clear whether Frege was aware of the significance of the distinction between objectual and substitutional quantification.

Frege's perhaps most impressive achievement in pure logic is his celebrated definition (with the proof of its adequacy) of the ancestral (or transitive closure) R^* of a binary relation R with the help of second-order quantification, already contained in *Begriffsschrift* and central to the logicist enterprise. Informally, an object a bears the ancestral R^* of a relation R to an object b if b can be

reached from a in a finite (nonzero) number of R-steps. That is, whenever there are objects $a_1, a_2, \ldots, a_n$ $(n > 1)$ such that $a_1Ra_2, a_2Ra_3, \ldots, a_{n-1}Ra_n$, then a_1 bears R^* to a_n. For example, if R is the parenting relation (so that xRy holds if and only if x is a parent of y), then R^* is the ancestor relation (i.e., xR^*y holds if and only if x is an ancestor of y), because x is an ancestor of y if y is a child of x, or a child of a child of x, or a child of a child of a child of x, and so on. Frege's idea is to define R^* from R as follows: a stands in the relation R^* to b if and only if b has every property F such that (1) all objects to which a bears R have F, and (2) F is hereditary with respect to the relation R (meaning that, whenever something x has the property F, and x bears R to some y, then y also has F). Note that this definition employs second-order quantification (over all R-hereditary properties F).

It is clear that, if b can be reached from a in a finite nonzero number of R-steps, then Frege's definition correctly implies that aR^*b, for if F is any property and b can be reached from a in one step, then by clause (1) of the definition b must have F, and if b can be reached from a by some number of R-steps greater than 1, one must have passed through an object to which a bears R, and which thus has F by clause (1), and every further object through which one has passed, including the last object b, must have F by clause (2). On the contrary, if b cannot be reached from a in a finite nonzero number of R-steps, then b lacks just that property of being reachable from a in a finite number of R-steps (a property that fulfills conditions [1] and [2]). In modern notation Frege's formal definition is as follows:

$$aR^*b :\leftrightarrow \forall F((\forall x(aRx \rightarrow Fx) \ \&\ \forall x \forall y\ (Fx \ \&\ xRy \rightarrow Fy)\ ) \rightarrow Fb).$$

It should be noted, finally, that Frege did not regard the sentences of his begriffsschrift as mere forms, open to arbitrary interpretation. Rather, he took them to express definite thoughts (i.e., propositions). This is manifest in the presence of a special symbol, the vertical judgment stroke, whose occurrence before a begriffsschrift formula indicates that the formula's content is actually asserted (and not talked about or simply entertained without judgment as to truth and falsity). While Frege did discuss the formal character of logic in terms of preservation of consequence on substituting nonlogical expressions for others (witness his correspondence with David Hilbert and the 1906 essay series "Über die Grundlagen der Geometrie"), he showed little inclination to pursue such investigations himself. Frege also has little to say about the characterization of propositions as logical truths; there is no indication that he had anything like Alfred Tarski's model-theoretic criterion in mind. He occasionally remarks that logical axioms are required to be "obvious," but generally takes it for granted that the specific basic laws he lays down are in fact logical truths.

FREGE'S ONTOLOGY AND PHILOSOPHY OF LANGUAGE

Frege's mature ontology is characterized by the fundamental dichotomy between saturated entities or objects (*Gegenstände*) on the one hand, and unsaturated entities or functions on the other hand. Functions are unsaturated or incomplete in the sense that they carry argument places that need to be filled; an object is anything that is not a function. Concepts are special functions, namely, functions whose values are always one of the two truth-values: the True and the False (which Frege takes to be objects, as will be explained). The realm of functions is stratified: Unary functions mapping objects to objects are first level, unary functions mapping first-level functions to objects are second level (an instance being the concept denoted by the first-level existential quantifier, which maps every first-level concept under which some object falls to the True, and all other first-level concepts to the False), and so on. The stratification becomes more complicated with functions of more than one argument, since there exist, for instance, functions of two arguments with one argument place for unary first-level functions and one argument place for objects (an instance being the application function, which maps a unary first-level function f and an object a to the result $f(a)$ of applying f to a), and so on.

The saturated-unsaturated dichotomy has, for Frege, a parallel in the linguistic realm. Singular terms, such as proper names and definite descriptions, are (linguistically) saturated (or complete) and refer to objects; predicate and functional expressions are incomplete and refer to functions. In determining the ontological status of certain entities Frege often proceeds by analyzing the expressions used to refer to them and takes the saturated or unsaturated nature of the expressions as a reliable guide to their ontological saturation status.

Now since the expression "the concept *horse*" is grammatically a singular term, Frege takes it to refer to an object, which commits him to the paradoxical claim that the concept *horse* is not a concept (compare to "Über Begriff und Gegenstand"). In an attempt to resolve this predicament Frege proposes that with every concept F is associated a certain (proxy) object that serves as the referent of "the concept F" (some commentators believe that Frege intended the extension of F to be this proxy object,

but the interpretive issue remains contentious). There remains a fundamental problem, however, for on the one hand, objects and concepts belong to distinct ontological categories, so that no predicate can be meaningfully applied to both a concept and an object; but on the other hand, Frege's explanation of this categorial distinction requires him to use the predicates "is an object" and "is a concept" in just this way—as contrasting (nonempty) predicates that can be applied to the same items. This creates some famous difficulties, some of which are discussed in the essay "Über Begriff und Gegenstand," because singular terms such as "the concept *horse*" cannot, according to Frege, refer to concepts, but refer to certain (proxy) objects instead.

Frege's most famous invention is perhaps his distinction between the sense (*Sinn*) and the reference (*Bedeutung*) of a linguistic expression, first introduced in his short 1891 booklet *Funktion und Begriff*, and expounded in detail in the 1892 essay "Über Sinn und Bedeutung." In the case of a singular term its reference is the object denoted by the term, whereas its sense is determined by the way that object is presented through the expression (its mode of presentation). Frege conceives of complete (declarative) sentences, perhaps infelicitously, as peculiar singular terms, so that their references, the special logical objects the True and the False, respectively, are objects. The thought expressed by a sentence is then defined by Frege to be the sentence's sense. The sense of a sentence is thus the mode of presentation of its truth-value; that is, on a natural reading, the sentence's truth-conditions. In the case of incomplete expressions, such as predicates and functional expressions, the references are of course the corresponding unsaturated concepts and functions.

While not explicitly discussed in "Über Sinn und Bedeutung," it becomes clear from the Frege-Husserl correspondence that Frege intended the notion of sense to apply to predicates as well. Scholarly discussion continues whether Frege considered the senses of unsaturated expressions to be functions, or whether he regarded all senses as objects (a stance suggested by the fact that every sense can be referred to by means of a singular nominal phrase of the form "the sense of the expression X"). In the essay "Der Gedanke" Frege expounds a Platonistic view of senses as inhabitants of a "third realm" of nonperceptible, objective entities, as opposed to the (perceptible) objects of the external world and the subjective contents (ideas) of humans' minds.

Frege was motivated to introduce the sense-reference distinction to solve certain puzzles, chief among them (1) the apparent impossibility of informative identity statements and (2) the apparent failure of substitutivity in contexts of propositional attitudes. As for (1), Frege argued that the statements "the morning start is the evening star" and "the morning star is the morning star" obviously differ in cognitive value (*Erkenntniswert*), which would be impossible if the object designated constituted the only meaning of a singular term. The sense-reference distinction allows one to attribute different cognitive values to these identity statements if the senses of the terms flanking the identity sign differ, while still allowing the objects denoted to be one and the same.

Regarding (2), Frege noticed that the sentences "John believes that the morning star is a body illuminated by the sun" and "John believes that the evening star is a body illuminated by the sun" may have different truth-values, although the one is obtained from the other by substitution of a coreferential term. He argued that, in contexts of propositional attitudes, expressions do not have their usual reference, but refer to their ordinary senses (which thus become their indirect references); then since "the morning star" and "the evening star" differ in ordinary sense, they are not, in the context at hand, coreferential, having distinct indirect references. Debate continues as to Frege's intentions concerning indirect senses of expressions, in particular whether iterated propositional attitude contexts give rise to an infinite hierarchy of indirect senses.

In the introduction to *Grundlagen* Frege enunciates "three fundamental principles" for his investigations. The first of these is an admonition to separate the logical from the psychological (a motif that runs through all of Frege's works); the third demands observance of the concept-object distinction. But it is the second of these principles that has drawn most attention and interpretation: "never to ask for the meaning of a word in isolation, but only in the context of a proposition." Other (not obviously equivalent) formulations of the principle occur in sections 60, 62, and 106 of *Grundlagen*; some authors take Frege to express a precursor of this principle in section 9 of *Begriffsschrift*, and some see an echo of it in *Grundgesetze*, volume 1, section 29.

The proper interpretation of the context principle continues to be contentious. While some philosophers regard it as being of the utmost importance to an understanding of Frege's philosophy, others view it as a rather ill-conceived and incoherent doctrine that he appears to have given up in later works. Those who take the context principle seriously mostly take it to claim some sort of epistemological priority of sentences (or perhaps the thoughts expressed by such) over subsentential linguistic

items (or perhaps their senses). It is easy to see why one might have misgivings about such an interpretation; after all, it at least appears to conflict with another Fregean principle, namely, that of compositionality (according to which the sense/reference of a compound expression is determined by the senses/references of its constituent expressions), which he held in high regard throughout his life.

FREGE'S PHILOSOPHY OF MATHEMATICS

Frege was, first and foremost, a philosopher of mathematics. While he followed Immanuel Kant in taking the truths of (Euclidean) geometry to be synthetic and knowable *a priori* (forcefully defending this view against Hilbert's axiomatic method in geometry), he vigorously argued, against Kant, for the logicist thesis, that is, the claim that the arithmetic truths, presumably including real and complex analysis, are analytic. In comparing Frege's views with Kant's it is however important to keep in mind that Frege was operating with his own technical definitions of analyticity and syntheticity, which are not obviously equivalent to Kant's: According to Frege (*Grundlagen* §3), a mathematical truth is analytic if it is derivable by means of logical inference rules from the general logical laws (and definitions) alone, whereas it is synthetic if it cannot be proved without recourse to truths belonging to a particular area of knowledge. Thus, analyticity and syntheticity are, for Frege, logico-epistemic notions, while Kant took them to be part semantic (analytic judgments are those whose predicate is contained in the subject, they are true by virtue of the meanings of their terms) and part epistemic (synthetic judgments extend one's knowledge, analytic ones do not).

In the preface to *Begriffsschrift* Frege makes it clear that it was the question of the epistemic status of arithmetic truths that prompted him to develop his new logic. At this time, Frege still avoids outright endorsement of the logicist thesis, stating only that he intends to investigate how far one may get in arithmetic with logical inferences alone. But there can be little doubt that he already envisages a definite path along which the ultimate proof of logicism is to proceed. Thus, he notes in part 3 of this work that mathematical induction rests on the *Begriffsschrift* theorem that, if an object x bears the transitive closure R^* of a binary relation R to an object y, and if x has a property F that is inherited along R, then y has F as well. It therefore seems clear that Frege already understood the possibility of logically proving the mathematical induction principle once the number 0 and the successor relation among natural numbers had been suitably defined, for the natural numbers could then be given as just those objects following 0 in the transitive closure of the successor relation.

By the time of *Grundlagen* the doctrine of logicism is firmly in place. Having vigorously criticized a selection of philosophical views about the notion of number (notably John Stuart Mill's empiricist and Kant's transcendentalist views), Frege, in the second part of that work, provides an informal, yet rigorous outline of how the reduction of arithmetic to logic may actually be carried out. He begins this endeavor by insisting that (1) ascriptions of number involve assertions about concepts and (2) the numbers themselves must be construed as objects. Frege argues for (1) by noting first that certain statements, like universal categoricals such as "all whales are mammals" and existential statements such as "there are books on the shelf," predicate something of concepts (rather than individuals). The first example statement is clearly not about any individual whale, but says of the concept *whale* that it is subsumed under the concept *mammal*; the second example predicates nonemptiness of the concept *book on the shelf*. The point is even clearer with respect to negated existential statements; "there are no Venus moons" is obviously not about any moon of Venus (if the statement is true, there are none), but denies that something falls under the concept *Venus moon*. Indeed, Frege notes, saying that there are no Venus moons amounts to the same thing as ascribing the number zero to the concept *Venus moon*. And just as in these examples, the numerical statement "there are four books on the shelf" clearly does not predicate anything of any particular book; instead, it, too, is a statement about the concept *book on the shelf*.

The thesis that ascriptions of number are best understood, in analogy with these examples, as assertions about concepts, is further bolstered by the observation that everyday numerical statements invariably involve common nouns or predicates, which, according to Frege, refer to concepts. Moreover, faced with the fact that one may with equal justice say "there is one deck of cards on the table," "there are fifty-two cards on the table," and "there are four suits of cards on the table," one is led to the recognition that there are different standards of unit involved in these assertions, and it seems perfectly natural to identify the respective concepts as these standards of unit. Thesis (2) is a consequence of Frege's view that the ontological category of an entity may be read off reliably from the linguistic category of expression that denotes the entity: According to Frege number terms typically appear as singular terms in natural languages, for

example, as "*the* number of cards on the table" or "*the* number four." Furthermore, pure arithmetic number terms typically flank the equality symbol, positions that, in Frege's view, are reserved for singular terms. Hence, Frege concludes, numbers must be objects.

Thus on the one hand, numbers, qua properties of concepts, would seem to be (higher-order) concepts; yet on the other hand, they must be construed as objects. Frege solves this apparent difficulty by suggesting that attributive uses of number words, as in "Jupiter has four moons," can always be paraphrased away, as in "the number of moons of Jupiter is four" (or, even more explicitly, "the number belonging to the concept *moon of Jupiter* is four"). In the latter statement, Frege claims, the *is* must denote identity and cannot function merely as a copula, since *four* is a singular term, and singular terms cannot follow the *is* of predication. The paradigmatic ascription of number then has the form "the number belonging to $F = x$," where F represents a predicate and x a singular term. Thus, the number term only forms part of the (higher-order) property ascribed to the concept, so that the objectual nature of number and the attributive character of ascriptions of number are compatible after all.

Frege next identifies a constraint that his reconstruction of arithmetic will have to abide by. Of fundamental importance for arithmetic are judgments of recognition, that is, identities, and so the definitions of the number-theoretic notions required for a proof of the logicist thesis must ensure that, in particular, identities of the form "the number belonging to F = the number belonging to G" receive the proper truth conditions. For this special type of identity statement, the truth conditions can readily be formulated in (dyadic second-order) logical terms, namely, the number belonging to F is the same as the number belonging to G if and only if there exists a binary relation R that correlates the objects that are F one-one and onto with the objects that are G. Since Frege quotes a somewhat obscure passage from David Hume at this point in *Grundlagen*, the principle has, perhaps infelicitously, come to be known as Hume's principle (HP).

Frege rejects HP as a definition of "the number belonging to F" on the grounds that it fails to specify truth conditions for contexts of the form "the number belonging to $F = x$," where x is a term that does not have the form "the number belonging to G," for example, when x is an individual variable. (This objection is now usually referred to as the Caesar problem—somewhat inaccurately, as Frege uses Julius Caesar as an example in arguing against a slightly different proposal for a definition). Some commentators maintain that Frege's only point in bringing up this objection is to show how HP is inadequate as a definition of *number* as described earlier. Other commentators see Frege as struggling here to arrive at adequacy conditions for the introduction of new sortal concepts into a language. On such a reading, however, it is difficult to see why Frege was not troubled by the obvious analogous problem arising for extensions of concepts in the *Grundgesetze*.

In any case Frege proposes an explicit definition of "the number belonging to F" that in effect amounts to taking this number to be the equivalence class of F under the equivalence relation of equinumerosity (which is explained in terms of the existence of a one-one and onto correlation): the number belonging to F, Frege stipulates, is the extension of the concept "concept equinumerous with F." Frege relies on a naive understanding of the notion of extension (later, in *Grundgesetze*, extensions themselves would be governed by an axiom that was to prove fatal for Frege's project). Frege then defines an object a to be a (cardinal) number if there exists a concept F such that a is the number belonging to F.

From the explicit definition of the number belonging to a concept, Frege proceeds to show that HP becomes derivable by means of pure logic and defines 0 as the number belonging to the concept "is an object not identical with itself" and 1 as the number belonging to the concept "is an object identical with 0." The successor relation among cardinal numbers is defined as follows: n succeeds m if n is the number belonging to some concept F under which some object a falls, and m is the number belonging to the concept "is an object falling under F, but not identical to a." Without proof Frege mentions the theorems that every number has at most one successor and one predecessor, and that every number except 0 succeeds some number. Making use of his definition of the ancestral (transitive closure) of a binary relation (as developed in *Begriffsschrift*), he defines the finite or natural numbers as those objects standing to 0 in the transitive reflexive closure of the successor relation, that is, informally, as those numbers than can be reached from 0 by taking successors finitely many times. Frege observes that this definition allows for a rather straightforward proof of the mathematical induction principle for natural numbers.

At this point, he has effectively recovered all the axioms of (second-order) Peano arithmetic from his definitions, except the one requiring every natural number to have a successor. Frege sketches a proof for this remaining axiom, which ultimately consists in showing by means of induction that, for any natural number n, the number belonging to the concept "object to which n bears

the transitive reflexive closure of the successor relation" (i.e., informally, "natural number being less than or equal to *n*") succeeds *n* (a fully detailed proof is carried out in *Grundgesetze*, although it is not entirely clear whether this is the same proof Frege intended in *Grundlagen*).

While the exposition of *Grundlagen* is entirely informal, *Grundgesetze*, which Frege hoped to be the final word on the logical nature of arithmetic, carries out the earlier sketch with full rigor, containing pages and pages of formal deductions in begriffsschrift notation. The crucial element added in *Grundgesetze* is the rigorous treatment of extensions of concepts (more precisely, of courses-of-values of functions, of which concept extensions are a special case). These are governed by Frege's basic law V, whose special case for concepts says that the extensions of concepts F and G coincide if and only if the same objects fall under F as fall under G. The use of extensions allows for the technique of type-lowering: First-level concepts can be simulated by their extensions, second-level concepts H can be simulated by the first-level concepts under which fall precisely the extensions of concepts falling under H, and so on. Frege makes extensive use of this technique; in particular, instead of defining the number belonging to F as the extension of the second-level concept "concept equinumerous with F," he is now able to take numbers to be extensions of first-level concepts. Otherwise, he follows the sketch of *Grundlagen* closely.

As Russell pointed out in a letter to Frege in 1902, the theory expounded in *Grundgesetze* is inconsistent, since it allows for the derivation of Russell's antinomy: Letting R be the first-level concept "*x* is the extension of some concept under which *x* does not fall," and *r* its extension, it follows easily from Frege's rules of inference, together with basic law V, that *r* both does and does not fall under R. Frege immediately realized that the antinomy threatened to undermine his life's work. While the second volume of *Grundgesetze* was in press, he hastily devised a quick fix that has come to be known as Frege's way out and added an appendix to the book, expressing both confidence that the revised system would prove capable of reconstructing arithmetic and worries about the philosophical underpinning of his revised basic law V. Frege's way out proved not to be a way out, since it was inconsistent with the existence of more than one object. The genesis of the antinomy in Frege's system is by now well understood; it arises through interplay of two principles that are individually consistent, namely, basic law V as mentioned earlier and *impredicative* second-order comprehension (roughly, statements to the effect that there exists a concept with a certain property, where that property is itself specified with the help of quantification over concepts); Frege's system with basic law V but only predicative instances of comprehension is now known to be consistent, but too weak to allow for a reconstruction of substantial mathematics.

Frege's work on the logical foundation of real analysis remained fragmentary; the second volume of *Grundgesetze* contains only preliminary definitions and theorems. Presumably he had planned a third volume, which, however, never appeared. Toward the end of his life, Frege seems to have abandoned logicism altogether, suggesting that arithmetic was instead based entirely on geometry, and hence synthetic, as Kant had held. His ideas on how such a claim might be proved were, however, never worked out.

NEO-FREGEANISM

Frege himself, and generations of philosophers and logicians after him, considered the mathematical content of *Grundlagen* and *Grundgesetze* largely obsolete because of the inconsistency of Frege's theory of extensions of concepts. In the 1980s, however, it began to be recognized that Frege had indeed hit on an exciting fact: If one takes the framework of Frege's theory to be essentially second-order predicate logic and adopts HP (with a primitive operator "the number belonging to," attaching to concept expressions) as an axiom, all of second-order Peano arithmetic becomes derivable, using the exact definitions and proofs employed by Frege (who used the explicit definition of "the number of F" only to prove HP from it, obtaining all further results directly from HP). This fact has become known as Frege's theorem. Importantly, it was soon observed that Frege arithmetic (i.e., full axiomatic second-order logic plus HP) is consistent, in contradistinction to the system of *Grundgesetze* (indeed, consistent relative to second-order Peano arithmetic).

It is still being debated whether, and to what extent, these discoveries have any bearing on the validity of the logicist thesis (restricted to arithmetic proper). While no one has seriously suggested that HP could be regarded as a principle of logic, some argue that it nevertheless enjoys some privileged epistemological status akin to analyticity, the principle being, in some sense, "analytic *of*" number. There are, however, serious difficulties in defending Frege arithmetic as being analytic. To start with, there is the familiar problem about the status of second-order logic itself, quite independently of HP. But even granting that second-order logic may count as logic in the requisite sense, further objections apply to HP. First, the principle

is not ontologically innocent, since it requires the first-order domain to be infinite, which is usually taken to be incompatible with analyticity. Second, any attempt to ground a privileged logical status of HP on its logical form (of an abstraction principle) runs afoul of the "bad company objection": There are abstraction principles of the same general logical form as HP that are inconsistent (such as Frege's basic law V). What is more, there are abstraction principles (like Boolos's parity principle) that hold only in finite domains, which makes them incompatible with HP, and hence it cannot be the logical form of an abstraction principle alone that could make HP analytic. Research on abstraction principles has increased significantly as a consequence of this discussion, as has work on the general logical and mathematical features of Frege's systems.

FREGE'S INFLUENCE

Through his publications, as well as through personal correspondence, Frege exerted a profound influence on Russell, who appears to have been the first major thinker to appreciate Frege's achievements in logic. Russell took over the logicist torch from Frege, and although Alfred North Whitehead and Russell's *Principia Mathematica* differs in many ways from Frege's work (it is much wider in scope, considerably less rigorous, and, in view of Russell's antinomy, takes a different approach to classes), it is clearly also heavily influenced by Frege (e.g., in imposing a structure of levels, or types, on the underlying ontology, and in the definition of number, nowadays often referred to as the Frege-Russell definition of cardinal number). It is known that Russell had read "Über Sinn und Bedeutung" and at least parts of *Grundgesetze* when he developed his celebrated theory of descriptions; and while there is no direct evidence for such a claim, it seems plausible to assume that Frege's discussion of definite descriptions in these works (especially the fully worked out formal theory of *Grundgesetze*) provided a helpful foil for Russell's own theory.

The degree to which Frege influenced Edmund Husserl is a more contentious matter. It is known that Husserl read all of Frege's major works and that the two corresponded extensively (except in the aftermath of Frege's rather hostile review [1894] of Husserl's *Philosophie der Arithmetik* [1891]). It seems fair to say that Frege (in particular, through the aforementioned review, as well as the preface to the first volume of *Grundgesetze*) is at least partly responsible for Husserl's antipsychologistic turn.

While Frege met neither Russell nor Husserl in person, he did have personal interactions with both Rudolf Carnap and Ludwig Josef Johann Wittgenstein. As a student, Carnap enrolled in various classes on begriffsschrift taught by Frege in Jena between 1910 and 1914; surely it was Frege who instilled in Carnap the idea that mathematics was reducible to logic, a view that was to become central to the Vienna Circle's philosophy. More generally, Frege shaped Carnap's whole attitude toward philosophy. After his immigration to the United States, Carnap, with Alonzo Church, was instrumental in keeping Fregean ideas in logic alive in the United States (where they came to flourish, for instance, in the work in semantics of David Kaplan and Richard Montague). Wittgenstein first visited Frege in Jena in 1911, and then at least two more times, in 1912 and 1913, while he was Russell's student in Cambridge. In addition, the two corresponded rather extensively from 1911 to 1920; it is clear from this correspondence that Frege and Wittgenstein thought highly of each other (the end of the correspondence is marked by an exchange of rather critical remarks by Frege on the *Tractatus* and by Wittgenstein on "Der Gedanke"). Fregean themes pervade the work of both the early and the late Wittgenstein, and it appears that Wittgenstein's intellectual respect for Frege never subsided.

In spite of this illustrious group of correspondents, Frege was for many years regarded as a somewhat obscure and ultimately failed predecessor of Russell's, possibly because few philosophers fully acknowledged Frege's influence on them (of course, the extent of this influence may not have been clear to them at the time). In the 1930s Heinrich Scholz and his school in Münster, Germany, rediscovered Frege and began work on an edition of his works, but that never materialized. The situation changed somewhat in the wake of John Langshaw Austin's English translation of the *Grundlagen*, which appeared in 1950; Frege was read, at that time, mainly as a philosopher of language, and as such influenced, among others, the British philosopher Peter Geach. The originality and independence of Frege's work (especially from Russell's), as well as his important role as a progenitor of analytic philosophy, was brought to prominence through the writings of Michael Dummett in the 1970s, who was himself heavily influenced by Frege's methodology and interests. In the United States, besides those mentioned earlier, Donald Davidson's work also revived discussion of Fregean themes. Crispin Wright's neologicism, especially as subsequently articulated and criticized by George S. Boolos and others, caused a veritable renaissance of interest in Frege's logical and mathematical work, beginning in the 1980s and continuing to this day.

See also Analytic and Synthetic Statements; Analyticity; Aristotle; Austin, John Langshaw; Carnap, Rudolf; Categories; Church, Alonzo; Davidson, Donald; Dummett, Michael Anthony Eardley; Geometry; Hilbert, David; Hume, David; Husserl, Edmund; Identity; Kant, Immanuel; Kaplan, David; Logic, History of; Logical Positivism; Mathematics, Foundations of; Mill, John Stuart; Montague, Richard; Peano, Giuseppe; Propositions; Russell, Bertrand Arthur William; Scholz, Heinrich; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

ABBREVIATIONS

CP

Collected Papers on Mathematics, Logic, and Philosophy. Translated by Max Black et. al; edited by Brian McGuinness. New York: Blackwell, 1984.

PW

Posthumous Writings. Translated by Peter Long and Roger White; edited by Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach. Chicago: University of Chicago Press, 1979.

TPW

Translations from the Philosophical Writings of Gottlob Frege. 3rd ed, edited by Peter Geach and Max Black. Totowa, NJ: Rowman & Littlefield, 1980.

FR

The Frege Reader, edited by Michael Beaney. Cambridge, MA: Blackwell, 1997.

KS

Kleine Schriften, edited by Ignacio Angelelli. Hildesheim, Germany: Georg Olms, 1967.

FBB

Funktion, Begriff, Bedeutung: Fünf logische Studien, edited by Günther Patzig. Göttingen, Germany: Vandenhoeck and Ruprecht, 1962.

LU

Logische Untersuchungen, edited by Günther Patzig. Göttingen, Germany: Vandenhoeck and Ruprecht, 1966.

WORKS BY FREGE

Ueber eine geometrische Darstellung der imaginären Gebilde in der Ebene. Inaugural-Dissertation der philosophischen Facultät zu Göttingen zur Erlangung der Doctorwürde vorgelegt von G. Frege aus Wismar. Jena: A. Neuenhahn, 1873. Reprinted in *KS*, tr. as "On a Geometrical Representation of Imaginary Forms in the Plane," in *CP*, pp. 1–55.

Rechnungsmethoden, die sich auf eine Erweiterung des Größenbegriffes gründen. Dissertation zur Erlangung der Venia Docendi bei der Philosophischen Facultät in Jena von Dr. Gottlob Frege (1874). Reprinted in *KS*, tr. as "Methods of Calculation based on an Extension of the Concept of Quantity [Magnitude]," in *CP*, pp. 56–92.

Review of H. Seeger, *Die Elemente der Arithmetik, für den Schulunterricht bearbeitet. Jenaer Literaturzeitung* 1 (46) (1874): 722. Reprinted in *KS*, tr. as "Review of H. Seeger, *Die Elemente der Arithmetik*," in *CP*, pp. 93–94.

Review of A. v. Gall and Ed. Winter, *die analytische Geometrie des Punktes und der Geraden und ihre Anwendung auf Aufgaben*. *Jenaer Literaturzeitung* 4 (9) (1877): 133–134. Reprinted in *KS*, tr. as "Review of A. von Gall and E. Winter, *Die analytische Geometrie des Punktes und der Geraden und ihre Anwendung auf Aufgaben*," in *CP*, pp. 95–97.

Review of J. Thomae, *Sammlung von Formeln welche bei Anwendung der elliptischen und Rosenhain'schen Funktionen gebraucht werden*. *Jenaer Literaturzeitung* 4 (30) (1877): 472. Reprinted in *KS*, tr. as "Review of J. Thomae, *Sammlung von Formeln, welche bei Anwendung der elliptischen und Rosenhainschen Funktionen gebraucht werden*," in *CP*, p. 98.

"Über eine Weise, die Gestalt eines Dreiecks als complexe Grösse aufzufassen," *Jenaische Zeitschrift für Naturwissenschaft* 12 (1878) Supplement, p. XVIII. Reprinted in *KS*, tr. as "Lecture on a Way of Conceiving the Shape of a Triangle as a Complex Quantity," in *CP*, pp. 99–100.

Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle: L. Nebert, 1879. Reprinted in *Conceptual Notation and Related Articles*, tr. and ed. by T. W. Bynum (OUP, 1972), pp. 101–203. Also in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931* ed. J. van Heijenoort, tr. S. Bauer-Mengelberg (Harvard University Press, 1967), pp. 5–82. §1–12 in *TPW*, pp. 1–20, and in *FR*, pp. 47–78.

"Anwendungen der Begriffsschrift," *Jenaische Zeitschrift für Naturwissenschaft* 13 (1879) Supplement II, pp. 29–33. Reproduced in *Begriffsschrift und andere Aufsätze*, ed. I. Angelelli (Hildesheim: Georg Olms, 1964), tr. as "Applications of the 'Conceptual Notation,'" (1879), in *Conceptual Notation and Related Articles*, tr. and ed. by T. W. Bynum (OUP, 1972), pp. 204–208.

Review of Hoppe, *Lehrbuch der analytischen Geometrie*, *Deutsche Literaturzeitung* 1 (6) (1880). Reprinted in *KS*, tr. as "Review of Hoppe, *Lehrbuch der analytischen Geometrie I*," in *CP*, pp. 101–102.

"Ueber die wissenschaftliche Berechtigung einer Begriffsschrift," *Zeitschrift für Philosophie und philosophische Kritik* 81 (1882), pp. 48–56. Reprinted in *KS*, tr. as "On the Scientific Justification of a Conceptual Notation," in *Conceptual Notation and Related Articles*, tr. and ed. by T. W. Bynum (OUP, 1972), pp. 83–89, and tr. by J. M. Bartlett as "On the Scientific Justification of a Conceptual Notation," *Mind* 73 (1964), pp. 155–160.

"Über den Zweck der Begriffsschrift," *Jenaische Zeitschrift für Naturwissenschaft* 16 (1883) Supplement, pp. 1–10. Reprinted in *Begriffsschrift und andere Aufsätze*, ed. I. Angelelli (Hildesheim: Georg Olms, 1964), tr. as "On the Aim of the 'Conceptual Notation,'" in *Conceptual Notation and Related Articles*, tr. and ed. by T. W. Bynum (OUP, 1972), pp. 90–100. Also tr. by V. Dudman as "On the Purpose of the Begriffsschrift," *The Australasian Journal of Philosophy* 46 (1968), pp. 89–97.

"Geometrie der Punktpaare in der Ebene," *Jenaische Zeitschrift für Naturwissenschaft* 17 (1884) Supplement, pp. 98–102.

Reprinted in *KS*, tr. as "Lecture on the Geometry of Pairs of Points in the Plane," in *CP*, pp. 103–107.

Die Grundlagen der Arithmetik, eine logisch-mathematische Untersuchung über den Begriff der Zahl. Breslau: W. Koebner, 1884; reprints Breslau: M. & H. Marcus, 1934, Hildesheim: G. Olms, 1961, and Darmstadt: Wissenschaftliche Buchgesellschaft, 1961. Tr. as *The Foundations of Arithmetic* by J. L. Austin, with German text, 2nd edition (Blackwell, 1953). §55–91, 106–109 also tr. M. S. Mahoney, in *Philosophy of Mathematics: Selected Readings*, 2nd edition, eds. P. Benacerraf and H. Putnam (CUP, 1983) pp. 130–159. Introduction, §1–4, 45–69, 87–91, 104–9 with summaries of remaining sections also in *FR*, pp. 84–129. Also reprinted as *Die Grundlagen der Arithmetik*, German centenary critical edition, ed. C. Thiel (Hamburg: Felix Meiner, 1986).

Review of H. Cohen: *Das Princip der Infinitesimal-Methode und seine Geschichte*, *Zeitschrift für Philosophie und philosophische Kritik* 87 (1885), pp. 324–329. Reprinted in *KS*, tr. as "Review of H. Cohen, *Das Prinzip der Infinitesimal-Methode und seine Geschichte*," in *CP*, pp. 108–111.

"Erwiderung," *Deutsche Literaturzeitung* 6 (28) (1885): 1030. Reprinted in *KS*, tr. as "Reply to Cantor's Review of *Grundlagen der Arithmetik*," in *CP*, p. 122.

"Über formale Theorien der Arithmetik," *Jenaische Zeitschrift für Naturwissenschaft* 19 (1886) Supplement, pp. 94–104. Reprinted in *KS*, tr. as "On Formal Theories of Arithmetic," in *CP*, pp. 112–121.

"Über das Trägheitsgesetz," *Zeitschrift für Philosophie und philosophische Kritik* 98 (1891). Reprinted in *KS*, tr. as "On the Law of Inertia," in *CP*, pp. 123–136, and tr. by R. Rand as "About the Law of Inertia" in *Synthese* 13 (1961), pp. 350–363.

Function und Begriff. Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medicin und Naturwissenschaft. Jena: H. Pohle, 1891. Reprinted in *FBB*. Tr. as "Function and Concept," in *TPW*, pp. 21–41, also in *CP*, pp. 137–156, and *FR*, pp. 151–171.

"Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik*, NF 100 (1892): 25–50. Reprinted in *FBB* and *KS*. Tr. by H. Feigl as "On Sense and Nominatum," in H. Feigl and W. Sellars, eds., *Readings in Philosophical Analysis* (New York: Appleton-Century-Croft 1949), as "On Sense and Reference," in *TPW*, pp. 56–78, also in *CP*, pp. 157–77, and *The Philosophical Review* 57 (1948), pp. 207–230, and as "On Sinn and Bedeutung" in *FR*, pp. 151–71.

"Ueber Begriff und Gegenstand," *Vierteljahrsschrift für wissenschaftliche Philosophie* 16 (1892): 192–205. Reprinted in *FBB* and *KS*. Tr. as "On concept and Object," in *TPW*, pp. 42–55, also in *PW*, pp. 87–117, in *CP*, pp. 182–94, *Mind* 60 (1951): 168–180 and in *FR*, pp. 181–193.

Review of Georg Cantor: *Zur Lehre vom Transfinitien. Gesammelte Abhandlungen aus der Zeitschrift für Philosophie und philosophische Kritik. Erste Abteilung. Zeitschrift für Philosophie und philosophische Kritik* 100 (1892): 269–272. Reprinted in *KS*, tr. as "Review of Cantor's *Zur Lehre vom Transfinitien*," in *CP*, pp. 178–81.

Grundgesetze der Arithmetik. Begriffsschriftlich abgeleitet. Jena: H. Pohle, Band I: 1893, Band II: 1903. Repr. together, Hildesheim: Georg Olms, 1962, 1998; and Darmstadt: Wissenschaftliche Buchgesellschaft, 1962. Preface, introduction and §1–52 of Vol. I tr. as *The Basic Laws of Arithmetic: Exposition of the System*, tr. and ed. by M. Furth (University of California Press, 1964). Selections from both vols. also tr. in *TPW*, and in *FR*, pp. 194–233, 258–289. Selections also tr. by J. Stachelroth and P. Jourdain as "A Formal System of Logic and Mathematics" in *Readings on Logic*, eds. I. Copi and J. Gould (New York: Macmillan, 1964).

Review of Dr. E. G. Husserl: *Philosophie der Arithmetik. Psychologische und logische Untersuchung. Zeitschrift für Philosophie und philosophische Kritik* 103 (1894): 313–332. Reprinted in *KS*, tr. as "Review of E. G. Husserl, *Philosophie der Arithmetik I*," in *CP*, pp. 195–209. Extracts in *TPW*, pp. 79–85, and *FR*, pp. 224–226.

"Kritische Beleuchtung einiger Punkte in E. Schöders Vorlesungen über die Algebra der Logik," *Archiv für systematische Philosophie* 1 (1895): 433–456. Reprinted in *LU*, also in *KS*, tr. as "A Critical Elucidation of Some Points in E. Schröder, *Vorlesungen über die Algebra der Logik*," in *CP*, pp. 210–228, also in *TPW*, pp. 86–106.

"Le nombre entier," *Revue de Métaphysique et de Morale* 3 (1895): 73–78. Reprinted in *KS*, tr. as "Whole Numbers," in *CP*, pp. 229–233.

"Lettera del sig. G. Frege all'Editore," *Revue de Mathématiques (Rivista di Matematica)* 6 (1896–1899): 53–59. Reprinted in Giuseppe Peano. *Opere scelte, II* (Rome: Cremonese, 1958), also in *KS*.

"Über die Begriffsschrift des Herrn Peano und meine eigene," *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Mathematisch-Physische Klasse* 48 (1897): 361–378. Reprinted in *FBB*, tr. as "On Mr. Peano's Conceptual Notation and My Own," in *CP*, pp. 234–248.

Über die Zahlen des Herrn H. Schubert. Jena: H. Pohle, 1899. Reprinted in *LU*, also in *KS*, tr. as "On Mr. H. Schubert's Numbers," in *CP*, pp. 249–72.

"Über die Grundlagen der Geometrie." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 12 (1903): 319–324. Reprinted in *KS*, tr. in "On the Foundations of Geometry: First Series," in *CP*, pp. 273–284, also in *On the Foundations of Geometry and Formal Theories of Arithmetic*, tr. by Eike-Henner W. Kluge (Yale University Press, 1971), pp. 22–37, and as "The Foundations of Geometry," *The Philosophical Review* 69 (1960), pp. 3–17.

"Über die Grundlagen der Geometrie II." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 12 (1903): 368–375. Reprinted in *KS*, tr. in "On the Foundations of Geometry: First Series," in *CP*, pp. 273–284, also in *On the Foundations of Geometry and Formal Theories of Arithmetic*, tr. by Eike-Henner W. Kluge (Yale University Press, 1971), pp. 22–37, and as "The Foundations of Geometry," *The Philosophical Review* 69 (1960), pp. 3–17.

"Was ist eine Funktion?" In: *Festschrift Ludwig Boltzmann gewidmet zum sechzigsten Geburstage, 20. Feb. 1904* (Leipzig: J. A. Barth, 1904), pp. 656–666. Reprinted in *FBB* and *KS*, tr. as "What is a Function?" in *TPW*, pp. 107–16, also in *CP*, pp. 285–92.

"Über die Grundlagen der Geometrie I." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15 (1906): 293–309. Reprinted in *KS*, tr. in "On the Foundations of Geometry: Second Series," in *CP*, pp. 293–340, also in *On the*

Foundations of Geometry and Formal Theories of Arithmetic, tr. by Eike-Henner W. Kluge (Yale University Press, 1971), pp. 49–112.

"Über die Grundlagen der Geometrie (Fortsetzung) II." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15 (1906): 377–403. Reprinted in *KS*, tr. in "On the Foundations of Geometry: Second Series," in *CP*, pp. 293–340, also in *On the Foundations of Geometry and Formal Theories of Arithmetic*, tr. by Eike-Henner W. Kluge (Yale University Press, 1971), pp. 49–112.

"Über die Grundlagen der Geometrie (Schluß) III," *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15 (1906): 423–430. Reprinted in *KS*, tr. in "On the Foundations of Geometry: Second Series," in *CP*, pp. 293–340, also in *On the Foundations of Geometry and Formal Theories of Arithmetic*, tr. by Eike-Henner W. Kluge (Yale University Press, 1971), pp. 49–112.

"Antwort auf die Ferienplauderei des Herrn Thomae." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 15 (1906): 586–590. Reprinted in *KS*, tr. as "Reply to Mr. Thomae's Holiday *Causerie*," in *CP*, pp. 341–345.

"Die Unmöglichkeit der Thomaeschen formalen Arithmetik aufs Neue nachgewiesen." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 17 (1908): 52–55. Reprinted in *KS*, tr. as "Renewed Proof of the Impossibility of Mr. Thomae's Formal Arithmetic," in *CP*, pp. 346–350.

Notes to Jourdain, Philip E. B., "The Development of the Theories of Mathematical Logic and the Principles of Mathematics." *Quarterly Journal of Pure and Applied Mathematics* XLIII (1912): 237–269.

"Der Gedanke: Eine logische Untersuchung." *Beiträge zur Philosophie des deutschen Idealismus* 1 (1918): 58–77. Reprinted in *LU* and *KS*, tr. as "Thought," Part I of *Logical Investigations*, ed. P. T. Geach, tr. P. T. Geach and R. Stoothoff (Blackwell, 1977), included in *CP*, pp. 351–372, and in *FR*, pp. 325–345. Also as "The Thought: A Logical Inquiry," in *Mind* 65 (1956): 289–311, tr. by A. M. and Marcelle Quinton.

"Die Verneinung. Eine logische Untersuchung." *Beiträge zur Philosophie des deutschen Idealismus* 1 (1918): 143–157. Reprinted in *LU* and *KS*, tr. as "Negation," Part II of *Logical Investigations*, ed. P. T. Geach, tr. P. T. Geach and R. Stoothoff (Blackwell, 1977), included in *CP*, pp. 373–389, and in *FR*, pp. 346–361.

"Logische Untersuchungen. Dritter Teil: Gedankengefüge." *Beiträge zur Philosophie des deutschen Idealismus* 3 (1923): 36–51. Reprinted in *LU*, *KS*, tr. as "Compound Thoughts," Part III of *Logical Investigations*, ed. P. T. Geach, tr. P. T. Geach and R. Stoothoff (Blackwell, 1977), included in *CP*, pp. 390–406, and *Mind* 72 (1963): 1–17.

UNPUBLISHED WORKS BY FREGE

Wissenschaftlicher Briefwechsel. Edited by Gottfried Gabriel et al. Hamburg, Germany: Felix Meiner, 1976, abridged for English edition by Brian McGuinness and translated by Hans Kaal as *Philosophical and Mathematical Correspondence*. Chicago: University Press of Chicago, 1980.

WORKS ABOUT FREGE

Antonelli, G. Aldo, and Robert May. "Frege's New Science." *Notre Dame Journal of Formal Logic* 41 (3) (2000): 242–270.

Beaney, Michael, and Erich H. Reck, eds. *Gottlob Frege: Critical Assessments of Leading Philosophers*. 4 vols. London: Routledge, 2005.

Blanchette, Patricia. "Frege and Hilbert on Consistency." *Journal of Philosophy* 93 (1996): 317–336.

Boolos, George S. *Logic, Logic, and Logic*, edited by Richard Jeffrey. Cambridge, MA: Harvard University Press, 1998.

Burge, Tyler. "Frege on Knowing the Foundations." *Mind* 107 (1998): 305–347.

Burgess, John P. *Fixing Frege*. Princeton, NJ: Princeton University Press, 2005.

Demopoulos, William, ed. *Frege's Philosophy of Mathematics*. Cambridge, MA: Harvard University Press, 1995.

Demopoulos, William. "The Philosophical Basis of Our Knowledge of Number." *Noûs* 32 (1998): 481–503.

Dummett, Michael. *Frege and Other Philosophers*. New York: Oxford University Press, 1991.

Dummett, Michael. *Frege: Philosophy of Language*. 2nd ed. Cambridge, MA: Harvard University Press, 1981.

Dummett, Michael. *Frege: Philosophy of Mathematics*. Cambridge, MA: Harvard University Press, 1991.

Dummett, Michael. *The Interpretation of Frege's Philosophy*. Cambridge, MA: Harvard University Press, 1981.

Ferreira, Fernando, and Kai F. Wehmeier. "On the Consistency of the (11 -CA Fragment of Frege's *Grundgesetze*." *Journal of Philosophical Logic* 31 (2002): 301–311.

Goldfarb, Warren. "Frege's Conception of Logic." In *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy*, edited by Juliet Floyd and Sanford Shieh. New York: Oxford University Press, 2001.

Hale, Bob, and Crispin Wright. *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. New York: Oxford University Press, 2001.

Heck, Richard G., Jr. "The Consistency of Predicative Fragments of Frege's Grundgesetze der Arithmetik." *History and Philosophy of Logic* 17 (1996): 209–220.

Heck, Richard G., Jr., ed. *Language, Thought, and Logic: Essays in Honour of Michael Dummett*. New York: Oxford University Press, 1997.

Heijenoort, Jean van. "Logic as Calculus and Logic as Language." *Synthese* 17 (1967): 324–330.

Hodes, Harold. "Logicism and the Ontological Commitments of Arithmetic." *Journal of Philosophy* 81 (3) (1984): 123–149.

McFarlane, John. "Frege, Kant, and the Logic in Logicism." *Philosophical Review* 111 (2002): 25–65.

Quine, Willard Van Orman. "On Frege's Way Out." *Mind* 64 (1955): 145–159.

Reck, Erich H., ed. *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy*. New York: Oxford University Press, 2002.

Resnik, Michael D. *Frege and the Philosophy of Mathematics*. Ithaca, NY: Cornell University Press, 1980.

Ricketts, Thomas. "Frege's 1906 Foray into Metalogic." *Philosophical Topics* 25 (2) (1997): 169–187.

Sluga, Hans D. *Gottlob Frege*. Boston: Routledge and Kegan Paul, 1980.

Stepanians, Markus S. *Frege und Husserl über Urteilen und Denken*. Paderborn, Germany: Schöningh, 1998.

Tait, William W., ed. *Early Analytic Philosophy—Frege, Russell, Wittgenstein: Essays in Honor of Leonard Linsky*. Chicago: Open Court, 1997.

Tappenden, Jamie. "Metatheory and Mathematical Practice in Frege." *Philosophical Topics* 25 (2) (1997): 213–264.

Thiel, Christian. *Sense and Reference in Frege's Logic*. Translated by T. J. Blakeley. Dordrecht, Netherlands: D. Reidel, 1968.

Weiner, Joan. *Frege in Perspective*. Ithaca, NY: Cornell University Press, 1990.

Wright, Crispin. *Frege's Conception of Numbers as Objects*. Aberdeen, Scotland: Aberdeen University Press, 1983.

Zalta, Edward N. "Frege's Logic, Theorem, and Foundations for Arithmetic." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2005.

Zalta, Edward N. "Gottlob Frege." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2005.

Kai F. Wehmeier (2005)

FRENCH ENCYCLOPEDIA, THE

See *Encyclopédie*

FRENCH PHILOSOPHICAL LITERATURE

See *Clandestine Philosophical Literature in France; Encyclopédie*

FREUD, SIGMUND

(1856–1939)

Sigmund Freud was the father of psychoanalysis, but—contrary to much apocryphal lore that dies hard—certainly not the originator of the hypothesis that unconscious ideation is essential to explain much of human overt behavior.

The generic doctrine of an unconscious domain of the mind has a venerable, long pre-Freudian history. Indeed, many of the most important doctrines commonly credited to Freud as his creations were tenets of his intellectual patrimony. Thus, as we recall from Plato's dialogue *The Meno*, Plato was concerned to understand how an ignorant slave boy could have arrived at geometric truths under mere questioning by an interlocutor with reference to a diagram. Plato argued that the slave boy had not acquired such geometric knowledge during his life. Instead, he explained, the boy was tapping prenatal but unconsciously stored knowledge, and restoring it to his conscious memory.

At the turn of the eighteenth century, Gottfried W. Leibniz gave psychological arguments for the occurrence of subthreshold sensory perceptions and for the existence of unconscious mental contents or motives that manifest themselves in our behavior (Ellenberger 1970). Moreover, in his *New Essays on Human Understanding* (1981), Leibniz pointed out that when the contents of some forgotten experiences subsequently emerge in our consciousness, we may misidentify them as new experiences, rather than recognize them as having been unconsciously stored in our memory.

Historically, it is more significant that Freud also had other precursors who anticipated some of his key ideas with impressive specificity. As he himself acknowledged ([the abbreviation "S.E." will be used to refer to the Standard Edition of Freud's complete psychological works in English] S.E., 1914, 14:15–16), Arthur Schopenhauer and Friedrich Nietzsche had speculatively propounded major psychoanalytic doctrines that he himself reportedly developed independently from his clinical observations only thereafter. Indeed, in a 1995 German book, *Die Flucht ins Vergessen: Die Anfänge der Psychoanalyse Freuds bei Schopenhaeur*, the Swiss psychologist Marcel Zentner traces the foundations of psychoanalysis to the philosophy of Schopenhauer.

But, as Freud then pointed out illuminatingly, it is one of the greatest threats to human self-esteem to face that "*the [human conscious] ego is not master in its own house*" (S.E., l917, 17:143; emphasis in original). On the other hand, it is evasive to dismiss substantive criticisms of Freudian theory as being due to fears induced by psychoanalytic accounts of presumed unconscious motivations. Such a dismissal does not address the merits of the strictures directed against psychoanalysis.

Freud was born in Freiberg, Moravia, then part of the Austro-Hungarian Empire, in 1856. But when he was three years old, his family moved to Vienna, where he entered the University of Vienna in 1873 to study medicine. He lived there until he was expelled by the Nazis, when he moved to London, where he died in l939.

It is important to distinguish between the validity of Freud's work qua psychoanalytic theoretician, and the merits of his earlier work. The zealous Freudian partisan

Mark Solms has edited and translated a presumably forthcoming four-volume series, *The Complete Neuroscientific Works of Sigmund Freud*. One focus of these writings is the neurological representation of mental functioning; another is Freud's supposed discovery of the essential morphological and physiological unity of the nerve cell and fiber.

They also contain contributions to the histology of the nerve cell, neuronal function, and neurophysiology. As a clinical neurologist, Freud wrote a monograph on aphasia (Solms and Saling, 1990). As Solms claims furthermore in his preview *An Introduction to the Neuro-Scientific Works of Sigmund Freud* (unpublished), Freud wrote major papers on cerebral palsy that earned him the status of a world authority. And he was a distinguished pediatric neurologist in the field of the movement disorders of childhood. Besides, Freud did scientific work on the properties of cocaine that benefited perhaps from his own use of that drug. Alas, that elating intake may well also account for some of the abandon featured by the more bizarre and grandiose of his psychoanalytic forays.

In 1880, he published a (free) translation of some of John Stuart Mill's philosophical writings. Yet, as Paul-Laurent Assoun notes in his 1995 *Freud, La Philosophie, et les Philosophes*, Freud was often disdainful of philosophy, despite clearly being indebted to the Viennese philosopher Franz Brentano, from whom he had taken several courses. The marks of Brentano's quondam representationalist and intentionalist account of the mental in the 1995 edition of his *Psychology from an Empirical Standpoint* are clearly discernible in Freud's conception of ideation. And the arguments for the existence of God championed by the quondam Roman Catholic priest Brentano further solidified the thoroughgoing atheism of Freud, who has been called a "*godless Jew*" (Gay, 1987, pp. 3–4; Grünbaum, 1993, ch. 7).

PSYCHOANALYSIS

The most basic ideas of psychoanalytic theory were initially enunciated in Josef Breuer and Sigmund Freud's *Preliminary Communication* of 1893, which introduced their *Studies on Hysteria*. But the first published use of the word *psychoanalysis* occurred in Freud's 1896 French paper on *Heredity and the Aetiology of the Neuroses* (S.E., 1896, 3:151). Therein Freud designated Breuer's method of clinical investigation as "a new method of psycho-analysis." Astonishingly, the coauthored 1893 prolegomenon, which lays bare the logical foundation of the cornerstone theory of repression, has been overlooked and untutoredly neglected in the literature, both psychoanalytic and philosophical. Breuer used hypnosis to revive and articulate a patient's unhappy memory of a supposedly repressed traumatic experience. The repression of that painful experience had occasioned the first appearance of a particular hysterical symptom, such as a phobic aversion to drinking water. Thus, Freud's mentor also induced the release of the suppressed emotional distress originally felt from the trauma. Thereby Breuer's method provided a catharsis for the patient.

The cathartic lifting of the repression yielded relief from the particular hysterical symptom. Breuer and Freud (1893) believed that they could therefore hypothesize that the repression, coupled with affective suppression, was the crucial cause for the development of the patient's psychoneurosis (S.E., 1893, 2:6–7; 3:29–31).

Having reasoned in this way, they concluded, in Freud's later words: "Thus one and the same procedure served simultaneously the purposes of [causally] investigating and of getting rid of the ailment; and this unusual conjunction was later retained in psycho-analysis" (S.E., 1924, 19:194).

In his 1924 historical retrospect, Freud acknowledged the pioneering role of Breuer's cathartic method: "The cathartic method was the immediate precursor of psychoanalysis; and, in spite of every extension of experience and of every modification of theory, is still contained within it as its nucleus" (S.E., 1924, 19:194).

Yet Freud was careful to highlight the contribution he made himself after the termination of his collaboration with Breuer. Referring to himself in the third person, he tells us: "Freud devoted himself to the further perfection of the instrument left over to him by his elder collaborator. The technical novelties which he introduced and the discoveries he made changed the cathartic method into psycho-analysis" (S.E. 1924, 19:195). Later on, Freud regarded repressed wishes rather than forgotten traumata as the principal pathogens of neuroses. These extensive elaborations have earned him the mantle of being the father of psychoanalysis.

It is important to recognize that there are major differences between the unconscious processes hypothesized by current cognitive psychology, on the one hand, and the unconscious contents of the mind claimed by psychoanalytic psychology, on the other (Eagle, 1987). These divergences are such that the existence of the cognitive unconscious clearly fails to support, if not impugns, the existence of Freud's "dynamic" unconscious.

His so-called dynamic unconscious is the supposed repository of repressed forbidden wishes of a sexual or aggressive nature, whose reentry or initial entry into consciousness is prevented by the defensive operations of the ego-agency of the mind. Though socially unacceptable, these instinctual desires are so imperious and peremptory that they recklessly seek immediate gratification, independently of the constraints of external reality.

But, in the cognitive unconscious, there is great rationality in the ubiquitous computational and associative problem-solving processes required by memory, perception, judgment, and attention. By contrast, as Freud emphasized, the wish-content of the dynamic unconscious makes it operate in a highly illogical way.

Having populated the dynamic unconscious with repressions, Freud reasoned that the use of his new technique of free association could lift these repressions of instinctual wishes, and could thereby bring the banished ideas back to consciousness unchanged. But in the case of the cognitive unconscious, we typically cannot bring to phenomenal consciousness the intellectual processes that are presumed to occur in it, although we can describe them theoretically. For example, even if his/her life depended on it, a student of czarist history simply could not bring into his/her phenomenal conscious experience the elaborate scanning or search process by which he/she rapidly comes up with the name of the Russian czarina's confidant G.Y. Rasputin, when asked for it. In sum, the presumed psychoanalytic unconscious as such cannot derive any credibility from the hypothesized cognitive unconscious.

PSYCHOANALYSIS AND WESTERN CULTURE

The poet W.H. Auden claimed that psychoanalysis is a whole climate of opinion. And indeed, it has been argued dubiously that the supposed pervasive influence of Freudian ideas in our culture vouches for the validity of the psychoanalytic enterprise. But even the premise that Freudian theory has become part of the intellectual ethos and folklore of Western culture cannot be taken at face value. As the distinguished Swiss scholar Henri Ellenberger stressed in his major historical work of 1970 *The Discovery of the Unconscious*, the prevalence of vulgarized pseudo-Freudian concepts makes it very difficult to determine reliably the extent to which genuine psychoanalytic hypotheses have actually become influential in our culture at large.

For example, any slip of the tongue or other bungled action (parapraxis) is typically yet incorrectly called a *Freudian slip*. But, as Freud himself pointed out, what is required for a slip or so-called parapraxis to qualify technically as Freudian is that it be motivationally opaque rather than transparent, precisely because its psychological motive is repressed (S.E., 1916–1917, 15: 41). Once it is clear what is meant by a bona fide Freudian slip, we need to ask whether there actually exist any such slips at all, that is, slips that appear to be psychologically unmotivated but are actually caused by repressed, unpleasant ideas. It is very important to appreciate how difficult it is to provide cogent evidence for such causation, as shown by strenuous attempts to furnish it experimentally.

Thus, as long as good empirical support for the Freudian scenario is unavailable, we actually do not know whether any bona fide Freudian slips exist at all. Just this lack of evidence serves to undermine the thesis that cultural influence is a criterion of validity. After all, if we have no cogent evidence for the existence of genuinely Freudian slips, then Freud's theory of bungled actions (parapraxes) might well be false. And if so, it would not contribute one iota to its validity even if our entire culture unanimously believed in it and made extensive explanatory use of it: When an ill-supported theory is used to provide explanations, they run the grave risk of being bogus, and its purported insights may well be pseudo-insights.

THE CORNERSTONE OF PSYCHOANALYSIS

In his 1914 *On the History of the Psychoanalytic Movement*, Freud wrote: "The theory of repression is the cornerstone on which the whole structure of psychoanalysis rests. It is the most essential part of it" (S.E., 1914, 14:16) The pillars of the avowed cornerstone of Freud's theoretical edifice comprise several major theses: (1) Distressing mental states induce the operation of a psychic mechanism of repression, which consists in the banishment from consciousness of unpleasurable psychic states (S.E., 1915, 14:147); (2) once repression is operative (more or less fully), it not only banishes such negatively charged ideas from consciousness, but plays a further crucial multiple causal role: It is causally necessary for the pathogenesis of neuroses, the production of our dreams, and the generation of our various sorts of slips (bungled actions); and (3) the method of free association can identify and lift (undo) the patient's repressions; by doing so, it can identify the pathogens of the neuroses, and the generators of our dreams, as well as the causes of our motivationally opaque slips; moreover, by lifting the pathogenic repres-

sions, free association also functions therapeutically, rather than only investigatively.

Freud provided two sorts of arguments for his cardinal etiologic doctrine that repressions are the pathogens of the neuroses: His earlier one, which goes back to his original collaboration with Josef Breuer, relies on purported therapeutic successes from lifting repressions; the later one, designed to show that the pathogenic repressions are sexual, is drawn from presumed reenactments (transferences) of infantile episodes in the adult patient's interactions with the analyst during psychoanalytic treatment. The process of repression, which consists in the banishment of ideas from consciousness or in denying them entry into it, is itself presumed to be unconscious (S.E., 1915, 14:147). In Freud's view, our neurotic symptoms, the manifest contents of our dreams, and the slips we commit are each constructed as "compromises between the demands of a repressed impulse and the resistances of a censoring force in the ego" (S.E., 1925, 20:45; and 1916–1917, 16:301).

By being only such compromises, rather than fulfillments of the instinctual impulses, these products of the unconscious afford only substitutive gratifications or outlets. For brevity, one can say, therefore, that Freud has offered a unifying compromise model of neuroses, dreams, and parapraxes. Since the repressed impulse made a compromise with the repressing ego, compromise-formations are products of *unsuccessful* repressions!

But what, in the first place, is the motive or cause that initiates and sustains the operation of the unconscious mechanism of repression before it produces its own later effects? Apparently, Freud assumes axiomatically that distressing mental states, such as forbidden wishes, traumata, disgust, anxiety, anger, shame, hate, guilt, and sadness—all of which are unpleasurable—almost always actuate, and then fuel, forgetting to the point of repression. Thus, repression regulates pleasure and so called "unpleasure" or displeasure by defending our consciousness against various sorts of negative affect. Indeed, Freud claimed perennially that repression is the paragon among our defense mechanisms. As he put it dogmatically: "The tendency to forget what is disagreeable seems to me to be a quite universal one" (S.E., 1901, 6:144), and "The recollection of distressing impressions and the occurrence of distressing thoughts are opposed by a resistance" (S.E., 1901, 6:146).

Freud tries to disarm an important objection to his thesis that "distressing memories succumb especially easily to motivated forgetting". He says:

> The assumption that a defensive trend of this kind exists cannot be objected to on the ground that one often enough finds it impossible, on the contrary, to get rid of distressing memories that pursue one, and to banish distressing affective impulses like remorse and the pangs of conscience. For we are not asserting that this defensive trend is able to put itself into effect *in every case*.
>
> (S.E., 1901, 6:147, ITALICS ADDED).

He acknowledges as "also a true fact" that "distressing things are particularly hard to forget" (S.E., 1916–1917, 15:76–77).

Indeed, Freud himself told us as an adult that he "can remember very clearly," from age seven or eight, how his father rebuked him for having relieved himself in the presence of his parents in their bedroom. In a frightful blow to the boy's ego, his father said: "The boy will come to nothing" (S.E., 1900, 4:216).

But Freud's attempt here to uphold his thesis of motivated forgetting is evasive and unavailing: Since some painful mental states are vividly remembered, while others are forgotten or even repressed, it appears that factors different from their painfulness determine whether they are remembered or forgotten. For example, personality dispositions or situational variables may in fact be causally relevant. To the great detriment of his theory, Freud never came to grips with the unfavorable bearing of this key fact about the mnemonic effects of painfulness on the tenability of the following pillar of his theory of repression: When painful or forbidden experiences are forgotten, the forgetting is tantamount to their repression due to their negative affect, and thereby produces neurotic symptoms or other compromise formations.

The numerous and familiar occurrences of vivid and even obsessive recall of negative experiences pose a fundamental statistical and explanatory challenge to Freud that neither he nor his followers have ever met. Astonishingly, Freud thinks he can parry this basic statistical and explanatory challenge by an evasive dictum, as follows: "Mental life is the arena and battle-ground for mutually opposing purposes [of forgetting and remembering] (S.E. 1916–1917, 15:76) …; there is room for both. It is only a question … of what effects are produced by the one and the other" (S.E., 1916–1917, 15: 77). Indeed, just that question cries out for an answer from Freud if he is to make his case. Instead, he cavalierly left it to dangle epistemologically in limbo.

Freud's argument here is an evasive attempt to neutralize the ubiquitous refuting instances undermining his aforecited claim (S.E., 1901, 6:144) that "The tendency to forget what is disagreeable seems to me to be a quite universal one." And he tries to do so by peremptorily inventing ad hoc an opposing tendency to remember negatively charged experiences. But since this gambit clearly fails, he has forfeited his basis for his pivotal etiologic scenario that forbidden or aversive states of mind are usually repressed and thereby cause compromise formations, such as neurotic symptoms.

UNSUCCESSFUL REPRESSIONS AS PATHOGENS OF THE PSYCHONEUROSES

Let us articulate and scrutinize Breuer and Freud's 1893 argument, in their foundational *Preliminary Communication*, for the pathogenicity of unsuccessful repressions. There they wrote:

> For we found, to our great surprise at first, that each individual hysterical symptom immediately and permanently disappeared when we had succeeded in bringing clearly to light the memory of the event by which it was provoked and in arousing its accompanying affect, and when the patient had described that event in the greatest possible detail and had put the affect into words. Recollection without affect almost invariably produces no result. The psychical process which originally took place must be repeated as vividly as possible; it must be brought back to its status nascendi and then given verbal utterance.
>
> (S.E., 1893, 2:6–7).

Breuer and Freud make an important comment on their construal of this therapeutic finding:

> It is plausible to suppose that it is a question here of unconscious suggestion: the patient expects to be relieved of his sufferings by this procedure, and it is this expectation, and not the verbal utterance, which is the operative factor. This, however, is not so. (S.E., 1893, 2:7)

And their avowed reason is that, in 1881, that is, in the "'pre-suggestion' era," the cathartic method was used to remove separately distinct symptoms, "which sprang from separate causes" such that any one symptom disappeared only after the cathartic (abreactive) lifting of a particular repression. But Breuer and Freud do not tell us why the likelihood of a placebo effect should be deemed to be lower when several symptoms are wiped out seriatim, than in the case of getting rid of only one symptom. Thus, as is pointed out in Grünbaum (1993), to discredit the hypothesis of placebo effect, it would have been essential to have comparisons with treatment outcome from a suitable control group whose repressions are not lifted. If that control group were to fare equally well, treatment gains from psychoanalysis would then be placebo effects after all.

In sum, Breuer and Freud inferred that the therapeutic removal of neurotic symptoms was produced by the cathartic lifting of the patient's previously ongoing repression of the pertinent traumatic memory, not by the therapist's suggestion or some other placebo factor (see Grünbaum 1993). This claim can be codified as follows:

> *T. Therapeutic Hypothesis*: Lifting repressions of traumatic memories cathartically is causally relevant to the disappearance of neuroses.

As we saw, Breuer and Freud (S.E., 1893, 2:6) reported the immediate and permanent disappearance of each hysterical symptom after they cathartically lifted the repression of the memory of the trauma that occasioned the given symptom. They adduce this "evidence" to draw an epoch-making inductive etiologic inference, which postulates "a causal relation between the determining [repression of the memory of the] psychical trauma and the hysterical phenomenon" (S.E., 1893, 2:6). Citing the old scholastic dictum *Cessante causa cessat effectus* (When the cause ceases, its effect ceases), they invoke its contrapositive (S.E., 1893, 2:7), which states that as long as the effect (symptom) persists, so does its cause (the repressed memory of the psychical trauma). And they declare just that to be the pattern of the pathogenic action of the repressed psychical trauma. This trauma, we learn, is not a mere precipitating cause. Such a mere "agent provocateur" just releases the symptom, "which thereafter leads an independent existence." Instead, "the [repressed] memory of the trauma … acts like a foreign body which long after its entry must continue to be regarded as an agent that is still at work" (S.E., 1893, 2:6).

The upshot of their account is that their observations of positive therapeutic outcome from the abreactive lifting of repressions, which they interpret in the sense of their therapeutic hypothesis, spelled a paramount etiologic moral as follows:

> *E. Etiologic Hypothesis*: An ongoing repression accompanied by affective suppression is causally necessary for the initial pathogenesis and persistence of a neurosis.

Clearly, this etiologic hypothesis *E* permits the valid deduction of the therapeutic finding reported by Breuer and Freud as codified in their therapeutic hypothesis *T*: The cathartic lifting of the repressions of traumatic memories of events that occasion symptoms engendered the disappearance of the symptoms. And, as they told us explicitly (S.E., 1893, 2:6), this therapeutic finding is their evidence for their cardinal etiologic hypothesis *E*.

But this inductive argument is vitiated by what might be called the fallacy of crude hypothetico-deductive (H-D) pseudo-confirmation. Thus, note that the remedial action of aspirin consumption for tension headaches does not lend H-D support to the outlandish etiologic hypothesis that a hematolytic aspirin deficiency is a causal sine qua non for having tension headaches, although such remedial action is validly deducible from that bizarre hypothesis.

Wesley Salmon called attention to the fallacy of inductive causal inference from mere valid H-D deducibility by giving an example in which a deductively valid pseudo-explanation of a man's avoiding pregnancy can readily give rise to an H-D pseudo-confirmation of the addle-brained attribution of his nonpregnancy to his consumption of birth control pills. Salmon, in his coauthored 1971 book *Statistical Explanation and Statistical Relevance*, states the fatuous pseudo-explanation:

> John Jones avoided becoming pregnant during the past year, for he had taken his wife's birth control pills regularly, and every man who regularly takes birth control pills avoids pregnancy. (p. 34)

Plainly, this deducibility of John Jones's recent failure to become pregnant from the stated premises does not lend any credence at all to the zany hypothesis that this absence of pregnancy is causally attributable to his consumption of birth control pills. Yet it is even true that any men who consume such pills in fact never do become pregnant. Patently, as Salmon notes, the fly in the ointment is that men just do not become pregnant, whether they take birth control pills or not.

His example shows that neither the empirical truth of the deductively inferred conclusion and of the pertinent initial condition concerning Jones nor the deductive validity of the inference can provide bona fide confirmation of the causal hypothesis that male consumption of birth control pills prevents male pregnancy: That hypothesis would first have to meet other epistemic requirements, which it manifestly cannot do.

Crude H-D confirmationism is a paradise of spurious causal inferences, as illustrated by Breuer and Freud's unsound etiologic inference. Thus, psychoanalytic narratives are replete with the belief that a hypothesized etiologic scenario embedded in a psychoanalytic narrative of an analysand's affliction is made credible merely because the postulated etiology then permits the logical deduction or probabilistic inference of the neurotic symptoms to be explained.

THE PSYCHOANALYTIC METHOD OF CLINICAL INVESTIGATION BY FREE ASSOCIATION: IS IT BOTH INVESTIGATIVE AND THERAPEUTIC?

This method, the so-called "Fundamental Rule" of clinical investigation in the setting of psychoanalytic treatment, is the supposed microscope, and even X-ray tomograph, as it were, of the human mind. Freud devised it, when he became dissatisfied with the use of hypnosis, which Breuer and he had employed theretofore as their probe.

The rule of free association directs the patient to tell the analyst without reservation whatever comes to mind. Thus, it serves as the fundamental method of clinical investigation. We are told that by using this technique to unlock the floodgates of the unconscious, Freud was able to show that neuroses, dreams, and slips are caused by repressed motives. Just as in Breuer's cathartic use of hypnosis, it is a cardinal thesis of Freud's entire psychoanalytic enterprise that his method of free association has a twofold major capability, which is both investigative and therapeutic: (1) It can identify the unconscious causes of human thoughts and behavior, both abnormal and normal; and (2) by overcoming resistances and lifting repressions, it can remove the unconscious pathogens of neuroses and thus provide therapy for an important class of mental disorders.

But on what grounds did Freud assert that free association has the stunning investigative capability to be causally probative for etiologic research in psychopathology? Is it not too good to be true that one can put a psychologically disturbed person on the couch and fathom the etiology of her or his affliction by free association? As compared to fathoming the causation of major somatic diseases, that seems almost miraculous, if true at all. Freud tells us very clearly (S.E., 1900, 5:528) that his argument for his investigative tribute to free association as a means of uncovering the causation of neuroses is, at bottom, a therapeutic one going back to the cathartic method of treating hysteria.

In a nutshell, his argument for claiming that free associations are causally probative for etiologic research in psychopathology, as well as vehicles of therapy, is as follows: (1) As he and Breuer had contended, unsuccessful repressions are the pathogens of the psychoneuroses; (2) The supposedly free associations departing from the patient's neurotic symptoms uncover the pertinent repressions; (3) Hence the method of free associations can identify the pathogenic repressions, and in so doing, it lifts them and thereby provides therapy for the neurosis and its symptoms. But it behooves us to expand this argument with a view to then seeing why it fails in several respects, no matter how revealing the associative contents may otherwise be in regard to the patient's psychological preoccupations and personality dispositions.

Drawing on his joint work with Breuer, Freud first inferred that the therapeutic disappearance of the neurotic symptoms is causally attributable to the cathartic lifting of repressions by means of the method of free association. Relying on this key therapeutic hypothesis, he then drew two further major theoretical inferences: (1) The seeming removal of the neurosis by means of cathartically lifting repressions is good inductive evidence for postulating that repressions accompanied by affective suppression are themselves causally necessary for the very existence of a neurosis (S.E., 1893, 2:6–7), and (2) granted that such repressions are thus the essential causes of neurosis, and that the method of free association is uniquely capable of uncovering these repressions, this method is uniquely competent to identify the causes or pathogens of the neuroses.

But the argument fails for the following several reasons. In the first place, the durable therapeutic success on which it was predicated did not materialize, as Freud was driven to admit both relatively early and very late in his career (S.E., 1925, 20:27; 1937, 23:216–253). And indeed, over a century later, three currently practicing English psychoanalysts, (Fonagy et al., 2005, p. 367) conceded ruefully: "Notwithstanding a history of over 100 years, psychoanalytically informed psychological therapies have a poor evidence base." But even insofar as Freud achieved transitory therapeutic gain, it will be recalled that he had failed to rule out a rival hypothesis which undermines his attribution of such gain to the lifting of repressions by free association: the ominous rival hypothesis of placebo effect, which asserts that treatment ingredients other than insight into the patient's repressions—such as the mobilization of the patient's hope by the therapist—are responsible for any resulting improvement. (For a detailed account of the placebo concept in both psychiatry and medicine, see Grünbaum, 1993, chap. 3). Nor have other analysts ruled out the placebo hypothesis during the past century.

Last, but not least, the repression etiology is inductively ill-founded, as will be recalled, and will now be seen further. It is unavailing to the purported etiologic probativeness of free associations that they may lift repressions, because Freud failed to show that the latter are pathogenic. In sum, Freud's argument has forfeited its premises.

Long after the *Preliminary Communication* of 1893, Freud (S.E., 1914, 14:12) offered an argument in his theory of "Transference" for the pathogenic role of repressions, hailing that argument as the most unshakable proof for his sexual etiology of the neuroses. It is a commonplace that many, if not all, adults carry over (transfer) to their adult interactions with other people attitudes and notions that they had acquired in (early) childhood. In this vein, Freud elaborates on this phenomenon in the context of the interpersonal transactions between the psychoanalyst and the patient. Thus, we learn, the patient transfers onto his or her psychoanalyst feelings and thoughts that originally pertained to important figures in his or her earlier life. In this important sense, the fantasies woven around the psychoanalyst by the analysand, and quite generally the latter's conduct toward his or her doctor, are hypothesized to be thematically recapitulatory of childhood episodes. And by thus being recapitulatory, the patient's behavior during treatment can be said to exhibit a thematic kinship to such very early episodes. Therefore, when the analyst interprets these supposed reenactments, the ensuing interpretations are called transference interpretations.

Freud and his followers have traditionally drawn the following highly questionable causal inference: Precisely in virtue of being thematically recapitulated in the patient-doctor interaction, the hypothesized earlier scenario in the patient's life can cogently be held to have originally been a pathogenic factor in the patient's affliction. For example, in his 1909 case history of the "Rat-Man," Freud infers that a certain emotional conflict had originally been the precipitating cause of the patient's inability to work, merely because this conflict had been thematically reenacted in a fantasy the "Rat-Man" had woven around Freud during treatment.

Thus, in the context of Freud's transference interpretations, the thematic reenactment is claimed to show that the early scenario had originally been pathogenic. According to this major etiologic conclusion, the patient's thematic reenactment in the treatment setting is also

asserted to be pathogenically recapitulatory by being pathogenic in the adult patient's here and now, rather than only thematically recapitulatory. Freud extols this dubious etiologic transference argument in his *On the History of the Psycho-Analytic Movement* (S.E., 1914, 14:12).

On the contrary, the patient's thematically recapitulatory behavior toward his or her doctor does not show that it is also pathogenically recapitulatory. The etiologic belief that it does so commits the "thematic affinity fallacy" (Grünbaum, 1993, p. 129; 2002, p. 134). How, for example, does the reenactment, during treatment, of a patient's early conflict show at all that the original conflict had been pathogenic in the first place? Indeed, it is epistemologically circular to infer the occurrence of infantile episodes from the adult patient's reports, and then to claim that these early episodes are thematically recapitulated in the adult analysand's conduct toward the analyst. Quite generally, how do transference phenomena focusing on the analyst show that a presumed current replica of a past event is pathogenic in the here and now?

Freud went on to build on the quicksand of his etiologic transference argument. It inspired two of his further fundamental tenets: first, the investigative thesis that the psychoanalytic dissection of the patient's behavior toward the analyst can reliably identify the original pathogens of his or her long-term neurosis; second, the cardinal therapeutic doctrine that the working through of the analysand's so-called transference neurosis is the key to overcoming his or her perennial problems.

THE PSYCHOANALYTIC THEORY OF DREAMING

As we learn from Freud's opening pages on his method of dream interpretation, he extrapolated the presumed causally probative role of free associations from being only a method of etiologic inquiry aimed at therapy, to serving likewise as an avenue for finding the purported unconscious causes of dreams (S.E., 1900, 4:100–101; 5:528). And in the same breath, he reports that when patients told him about their dreams while associating freely to their symptoms, he extrapolated his compromise model from neurotic symptoms to manifest dream contents. A year later, he carried out the same twofold extrapolation to include slips or bungled actions.

But what do free associations tell us about our dreams? Whatever the manifest content of dreams, they are purportedly wish-fulfilling in at least two logically distinct ways: For every dream *D*, there exists at least one normally unconscious infantile wish *W* such that: (1) *W* is the motivational cause of *D*; and (2) the manifest content of *D* graphically displays, more or less disguisedly, the state of affairs desired by *W*. As Freud opined: "When the latent dream-thoughts that are revealed by the analysis [via free association] of a dream are examined, one of them is found to stand out from among the rest … the isolated thought is found to be a wishful impulse" (S.E., 1925, 20:44). But as Clark Glymour (1983) has emphasized, Freud manipulated and doctored the free associations to yield a distinguished wish motive. Thus, Freud had declared with categorical universality (S.E. l900, 4:134) "there cannot be any dreams but wishful [i. e., wish-generated] dreams"

Quite independently of Freud's abortive therapeutic argument for the causal probativeness of free association, he offered his analysis of his 1895 *Specimen Irma Dream* as a nontherapeutic argument for the method of free association as a cogent means of identifying hypothesized hidden, forbidden wishes to be motives of our dreams. But, in a detailed critique of that unjustly celebrated Irma Dream, it has been shown that Freud's account there is, alas, no more than a piece of false advertising for the following reasons:

- It does not deliver at all the promised vindication of the probativeness of free association.
- It does nothing toward warranting his foolhardy dogma that all dreams are wish-fulfilling in his stated sense.
- It does not even pretend that his alleged "Specimen Dream" is evidence for his compromise model of manifest-dream content.
- The inveterate and continuing celebration of Freud's analysis of his "Irma Dream" in the psychoanalytic literature as the paragon of dream interpretation is completely unwarranted, because it is mere salesmanship (Grünbaum 1984, pp. 216–239).

Moreover, careful studies have shown that the so-called free associations are not free but are strongly influenced by the psychoanalyst's subtle promptings to the patient (Grünbaum 1984). And recent memory research has shown further how patients and others can be induced to generate pseudo-memories, which are false but deemed veridical by the patients themselves (Goleman 1994). As a corollary of the latter epistemological defects of the method of free association, it appears that such associations cannot reliably vouch for the contents of presumed past repressions that are lifted by them.

Once Freud had clearly chained himself gratuitously to the universal wish monopoly of dream generation, his interpretations of dreams were constrained to reconcile wish-contravening dreams with the decreed universality of wish fulfillment. Such reconciliation demanded imperiously that all other parts and details of his dream theory be obligingly tailored to the governing wish dogma so as to sustain it. Yet Freud artfully obscured this dynamic of theorizing, while begging the methodological question (S.E., 1900, 4:135). Wish-contravening dreams include anxiety dreams, nightmares, and the so-called counter-wish dreams (S.E., 1900, 4:157). As an example of the latter, Freud reports a trial attorney's dream that he had lost all of his court cases (S.E., 1900, 4:152).

His initial 1900 statement of his dual wish fulfillment in a dream had been: "Thus, its content was the fulfillment of a wish and its motive was a wish" (S.E., 1900, 4:119). But the sense in which dreams are wish fulfilling overall is purportedly threefold rather than only twofold: One supposed motivating cause is the universal preconscious wish to sleep, which allegedly provides a generic causal explanation of dreaming as such and, in turn, makes dreaming the guardian of sleep (S.E., 1900, 4:234; 5:680); another is the individualized repressed infantile wish, which is activated by the day's residue and explains the particular manifest content of a given dream; furthermore, as already noted, that manifest content of the dream graphically displays, more or less disguisedly, the state of affairs desired by the unconscious wish. The disguise is supposedly effected by the defensive operation of the dream distortion of the content of forbidden unconscious wishes.

But this theorized distortion of the hypothesized latent content must not be identified with the very familiar phenomenological bizarreness of the manifest dream content! By achieving a compromise with the repressed wishes, the postulated distortion makes "plausible that even dreams with a distressing content are to be construed as wish fulfillments" (S.E., 1900, 4:159). Accordingly, Freud concedes: "The fact that dreams really have a secret meaning which represents the fulfillment of a wish must be proved afresh in each particular case by analysis" (S.E., 1900, 4:146).

THE HERMENEUTIC RECONSTRUCTION OF PSYCHOANALYSIS

In concert with the so-called hermeneutic German philosophers Karl Jaspers and Jürgen Habermas, the French philosopher Paul Ricoeur believed that victory can be snatched from the jaws of the scientific failings of Freud's theory by abjuring his scientific aspirations as misguided. Claiming pejoratively that Freud himself had "scientistically" misunderstood his own theoretical achievement, some hermeneuts misconstrue it as a semantic accomplishment by trading on the multiply ambiguous word *meaning* (Grünbaum 1999, 2002).

In Freud's theory, an overt symptom manifests one or more underlying unconscious causes and gives evidence for its cause(s), so that the sense or meaning of the symptom is constituted by its latent motivational cause(s). But this notion of meaning is different from the one appropriate to the context of communication, in which linguistic symbols acquire semantic meaning by being used deliberately to designate their referents. Clearly, the relation of being a manifestation, which the symptom bears to its cause, differs from the semantic relation of designation, which a linguistic symbol bears to its object.

The hermeneutic reconstruction of psychoanalysis slides illicitly from one of two familiar senses of the term "meaning" encountered in ordinary discourse to another. When a pediatrician says that a child's spots on the skin mean measles, the meaning of the symptom is constituted by one of its causes, much as in the Freudian case. Yet, when speaking of Freud's making sense of a patient's symptoms, the analyst Anthony Storr (1986) conflates the fathoming of the etiologic sense or meaning of a symptom with the activity of making semantic sense of a text, preposterously transmogrifying Freud into a semanticist: "Freud was a man of genius whose expertise lay in semantics" (p. 260). And Ricoeur even wrongly credits Freud's theory of repression with having provided, *malgré lui*, a veritable semantics of desire.

Relatedly, John R. Searle has noted illuminatingly in his 1990 book *Intentionality* that, unlike many mental states, language is not intrinsically intentional in Brentano's directed sense; instead, the intentionality (aboutness) of language is extrinsically imposed on it by deliberately decreeing it to function referentially. Searle points out that the mental states of some animals and of pre-linguistic very young children do have intrinsic intentionality but no linguistic referentiality.

Thus, it is a fundamental hermeneuticist error to slide illicitly from the intrinsic, non-semantic intentionality of (many, but not all) mental states to the imposed, semantic sort possessed by language. Moreover, some of the neurotic symptoms of concern to psychoanalysts, such as diffuse depression and manic, undirected elation even lack Brentano intentionality.

Yet some version of a hermeneutic reconstruction of the psychoanalytic enterprise has been embraced with alacrity by a considerable number of analysts no less than by professors in humanities departments of universities. Its psychoanalytic adherents see it as buying absolution for their theory and therapy from the criteria of validation mandatory for causal hypotheses in the empirical sciences, although psychoanalysis is replete with just such hypotheses. This form of escape from accountability also augurs ill for the future of psychoanalysis, because the methods of the champions of the hermeneutic reconstruction of psychoanalysis have not spawned a single new important hypothesis. Instead, their reconstruction is a negativistic ideological battle cry whose disavowal of Freud's scientific aspirations presages the death of his legacy from sheer sterility, at least among those who demand the validation of theories by cogent evidence.

FREUD ON THEISTIC RELIGION

In his 1933 essay *The Question of a Weltanschauung*, Freud appraised theism under the label of *religion* and wrote:

> Religion is an attempt to master the sensory world in which we are situated by means of the wished world which we have developed within us as a result of biological and psychological necessities. But religion cannot achieve this. Its doctrines bear the imprint of the times in which they arose, the ignorant times of the childhood of humanity. Its consolations deserve no trust. Experience teaches us that the world is no nursery.
>
> (S.E., 1933, 22:168).

And in his 1927 critique of theism entitled *The Future of an Illusion*, he stresses the logical priority of his atheism vis-à-vis his psychology of theism:

> Nothing that I have said here against the truth-value of religions needed the support of psychoanalysis; it had been said by others long before analysis came into existence. If the application of the psycho-analytic method makes it possible to find a new argument against the truths of religion, tant pis [so much the worse] for religion; but defenders of religion will by the same right make use of psycho-analysis in order to give full value to the affective significance of religious doctrines.
>
> (S.E., 1927, 21:37).

This avowed entitlement of religious partisans is presumably an allusion to Freud's friend Oskar Pfister, a Lutheran clergyman and avid champion of the use of psychoanalysis in pastoral work. Relatedly, though, like Freud, also a committed atheist, Karl Marx had expressed sympathy for the quest for solace in the face of the trials and tribulations of life. Marx wrote:

> "Religion … is … the protest against real distress. Religion is the sigh of the oppressed creature, the heart of a heartless world, just as it is the spirit of an unspiritual situation. It is the opium of the people."
>
> (FEUER 1959, P. 523).

Marx's use of the term *opium* here to characterize the consoling function of religion is descriptive rather than pejorative: In his time, opium was a commonly used anodyne, available without prescription.

Freud maintained that religious beliefs are engendered by the synergism of three significantly different sorts of powerful, relentless wishes. And for each of this trio of wishes, he conjectures a distinct scenario that specifies their content and mode of operation.

As he points out, the first set of these psychogenetic assumptions features wish motives that are largely conscious or manifest, instead of being the repressed wishes postulated by psychoanalytic theory. Accordingly, this component of Freud's triadic psychology of religion does not rely on any of his technical psychoanalytic teachings. But what are the relevant archaic conscious wishes? He explains eloquently in his 1927 book *The Future of an Illusion*:

> … the terrifying impression of helplessness in childhood aroused the need for protection—for protection through love—which is provided by the father; and the recognition that this helplessness lasts throughout life made it necessary to cling to the existence of a father, but this time a more powerful one. Thus the benevolent rule of a divine Providence allays our fear of the dangers of life; the establishment of a moral world-order ensures the fulfillment of the demands of justice, which have so often remained unfulfilled in human civilization; and the prolongation of earthly existence in a future life provides the local and temporal framework in which these wish-fulfillments shall take place. Answers to the riddles that tempt the curiosity of man, such as how the universe began or what the relation is between body and mind, are developed in conformity with the underlying assumptions of this system
>
> (S.E., 1927, 21:30).

Understandably, therefore, the protector, creator, and lawgiver are all rolled into one. No wonder, says Freud

(S.E., 1933, 22:163–164), that in one and the same breath, Immanuel Kant coupled the starry heavens above, and the moral law within as both being awe-inspiring. After all, Freud asks rhetorically, "what have the heavenly bodies to do with the question of whether one human creature loves another or kills him?" And he answers: "The same father (or parental agency) which gave the child life and guarded him against its perils, taught him as well what he might do and what he must leave undone" (S.E., 1933, 22:164).

Insofar as Freud's psychogenetic portrayal of religion depicts it as the product of conscious wishes, his account draws, not only on Ludwig A. Feuerbach, but also on commonsense psychology. After all, at least prima facie, it is rather a commonplace that people seek to avoid anxiety, and that they therefore tend to welcome the replacement of threatening beliefs by reassuring ones. Hence, for brevity, this component of Freud's triadic psychology of religion can be designated as the "*commonsense hypothesis*," which is not to say, however, that it is obviously true. Each of the other two components of this trinity is a set of psychoanalytic claims, asserting the operation of repressed motives. And yet they differ from each other, because one of them relies on Freud's theory of the psychosexual development of the human individual, while the other consists of ethnopsychological and psychohistorical averrals pertaining to the evolution of our species as a whole. Accordingly, the former psychoanalytic assumptions can be dubbed *ontogenetic*, while the latter can be labeled *phylogenetic*.

The legitimacy of any psychogenetic portrait of religious creeds depends on the evidential merit of the explanatory psychological hypotheses adduced by it. Even the commonsense component of Freud's triad is subject to this caveat. Invoking the criticisms of his great predecessors, he took it for granted that there is no cogency in any of the arguments for the existence of God offered by believers. But he coupled this philosophical judgment with the daring motivational claim that the faithful who nonetheless adduce such proofs had not, in fact, themselves been decisively moved by them, when giving assent to theism. Instead, he maintained, psychologically this assent is emotional or affective in origin.

Thus he is telling us that motivationally, the dialectical excogitations offered as existence proofs are post hoc rationalizations in which an elaborate intellectual façade takes the place of the deep-seated wishes that actually persuaded the theologians. Speaking epigrammatically in another context, Freud quotes Shakespeare's Falstaff as saying that reasons are "as plenty as blackberries" (S.E., 1914, 14:24).

It would seem to be basically a matter of empirical psychological fact whether the commonsense constituent of Freud's psychogenetic portrait of religion is sound. Yet, it is not clear how to design a cogent test even of this hypothesis. For note that the required design needs to have two epistemic capabilities as follows: (1) It needs to yield evidence bearing on the validity of the functional explanation of religious belief as being anxiety-reducing; presumably this explanation postulates some kind of stabilizing psychic servomechanism that reacts homeostatically to psychological threat; and, furthermore, (2) the required test needs to be at least able to rank-order the intensity of the wish to escape from anxiety, as compared to the motivational persuasiveness of the theological existence proofs. Perhaps oscillating anxieties of believers who went through cycles of doubt and belief have already gone some way toward meeting the first condition by Mill's method of concomitant variations. In any case, it would seem that an explicitly fideist belief in the existence of God—which avowedly is not based on any arguments—calls for psychological explanation in terms of wish motives!

The second requirement, however, seems to be a tall order indeed, although it does not warrant putting a cap on the ingenuity of potential empirical investigators. It, too, must be met, because of Freud's bold claim that even the best of the arguments for the existence of God would not have convinced the great minds who advanced them, unless stronger tacit wishes had carried the day, or had prompted these intellects to prevaricate. But note that, so far, Freud's portrayal of the motives for religious belief has studiously refrained from claiming that this belief is false, although he does avow its falsity later, after arguing that it is delusional. Hence whatever the empirical difficulties of validating his psychogenetic portrait, they are hardly tantamount to his commission of the hackneyed genetic fallacy, a mode of inference that he had explicitly rejected by means of disclaimers and qualifications.

In accord with his diagnosis of religion as an unwholesome childish fixation, Freud did advocate—as an experiment worth making—that children be given an irreligious education. But he took pains to say at once: "Should the experiment prove unsatisfactory I am ready to give up the reform and to return to my earlier, purely descriptive judgment that man is a creature of weak intelligence who is ruled by his instinctual wishes" (S.E., 1927, 21:48–49).

The two psychoanalytic components of Freud's triadic psychology of theism—its ontogeny and phylogeny—even more than its pre-psychoanalytic commonsense constituent, exigently require evidence for the existence of the two different sorts of wishes postulated by them. Insofar as even the very existence of these hidden desires is questionable, one remains less than convinced, when told that they contributed significantly to the initial genesis and later persistence of religious creeds. It is a corollary of the evidential scrutiny of the pertinent hypotheses that the psychoanalytic ontogeny of theism still lacks cogent evidential warrant (Grunbaum 1984, 1993).

But Freud was not content to confine himself to explanatory reliance on the conscious quest for anxiety reduction, and on his ontogeny of theism. Rather, he went on to develop a psychoanalytic phylogeny of theism (S.E., 1913, 13:100). In his view, this historical ethnopsychology is a valid extension of psychoanalysis.

As he sees it, by combining ethnography with psychoanalysis, he has discerned a third set of strong wishes that unite synergistically with the other two classes of this triad, and make the psychogenesis of belief in God the Father the more imperative. Therefore he proclaimed: "We now observe that the store of religious ideas includes not only wish-fulfillments but important historical recollections. This concurrent influence of past and present must give religion a truly incomparable wealth of power" (S.E., 1927, 21:42).

Daring and ingenious though it is, Freud's psychoanalytic phylogeny of theism is dubious, if only because it assumes a Lamarckian inheritance of repressed racial memories. Furthermore, contrary to the uniform evolution of religions required by his account, more recent historical scholarship seems to call for developmental pluriformity, as pointed out by Hans Küng in his 1979 book *Freud and the Problem of God* (p. 67).

Professor Edward Erwin's essay *Psychoanalysis: Theory, Therapy, and Method of Inquiry Created by Sigmund Freud (1856–1939)* herein covers the post Freudians.

See also Atheism; Brentano, Franz; Common Consent Arguments for the Existence of God; Dreams; Egoism and Altruism; Existential Psychoanalysis; Feuerbach, Ludwig Andreas; Habermas, Jürgen; Hermeneutics; Intentionality; Jaspers, Karl; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Marx, Karl; Mill, John Stuart; Nietzsche, Friedrich; Philosophy of Religion; Plato; Popular Arguments for the Existence of God; Psychoanalysis; Psychoanalytic Theories, Logical Status of; Ricoeur, Paul; Salmon, Wesley; Schopenhauer, Arthur; Searle, John; Unconscious.

Bibliography

PRIMARY WORKS BY FREUD

Breuer, Josef, and Sigmund Freud. "On the Psychical Mechanism of Hysterical Phenomena: Preliminary Communication" (1893). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 2, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"On the Psychical Mechanism of Hysterical Phenomena: A Lecture" (1893). In J. Strachey, *The Standard Edition of the Works of Sigmund Freud*, Vol. 3, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"Heredity and the Aetiology of the Neuroses" (1896). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 3, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"The Interpretation of Dreams" (1900). In *The Standard Edition of the Works of Sigmund Freud*, Vols. 4 (Part I) and 5 (Part II), edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"The Psychopathology of Everyday Life" (1901). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 6, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"The Claims of Psycho-Analysis to Scientific Interest" (1913). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 13, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"The Return of Totemism in Childhood" (1913). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 13, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"On the History of the Psycho-Analytic Movement" (1914). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 14, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"Repression" (1915). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 14, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"Introductory Lectures on Psycho-Analysis" (1915–1917). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 15 (Parts I and II) and Vol. 16 (Part III), edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"A Difficulty in the Path of Psycho-Analysis" (1917). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 17, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"A Short Account of Psychoanalysis" (1924). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 19, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"An Autobiographical Study" (1925). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 20, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"The Future of an Illusion" (1927). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 21, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

"New Introductory Lectures on Psychoanalysis" (1933). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 22, edited and translated by J. Strachey. London: Hogarth Press, 1933.

"Analysis Terminable and Interminable" (1937). In *The Standard Edition of the Works of Sigmund Freud*, Vol. 23, edited and translated by J. Strachey. London: Hogarth Press, 1953–1974.

SECONDARY WORKS

Assoun, Paul-Laurent. *Freud, La Philosophie, et les Philosophes*. Paris: Presses Universitaires de France, 1995.

Brentano, Franz. *Psychology from an Empirical Standpoint*. New York: Routledge and Kegan Paul, 1995.

Eagle, Morris N. "The Psychoanalytic and the Cognitive Unconscious." In *Theories of the Unconscious and Theories of the Self*, edited by R. Stern. Hillsdale, NJ: Analytic Press, 1987.

Ellenberger, Henri. *The Discovery of the Unconscious*. New York: Basic Books, 1970.

Fonagy, Peter. "The Outcome of Psychodynamic Psychotherapy for Psychological Disorders." *Clinical Neuroscience Research* 40 (2005): 367.

Gay, Peter. *A Godless Jew: Freud, Atheism, and the Making of Psychoanalysis*. New Haven, CT.: Yale University Press, 1987.

Glymour, Clark. "The Theory of Your Dreams." In *Physics, Philosophy, and Psychoanalysis: Essays in Honor of Adolf Grünbaum*, edited by R.S. Cohen and L. Laudan. Dordrecht, Netherlands: Reidel, 1983.

Goleman, Daniel. "Miscoding Is Seen as the Root of False Memories." *New York Times* May 31 (1994): C1 and C8.

Grünbaum, Adolf. "Critique of Psychoanalysis." In *The Freud Encyclopedia, Theory, Therapy, and Culture*, edited by E. Erwin. New York; London: Routledge, 2002.

Grünbaum, Adolf. *The Foundations of Psychoanalysis: A Philosophical Critique*. Berkeley: University of California Press, 1984.

Grünbaum, Adolf. "The Hermeneutic Versus the Scientific Conception of Psychoanalysis: An Unsuccessful Effort to Chart A Via Media for the Human Sciences." In *Einstein Meets Magritte, An Interdisciplinary Reflection: The White Book of Einstein Meets Magritte*, edited by D. Aerts. Dordrecht, Netherlands: Kluwer Academic, 1999.

Grünbaum, Adolf. *Validation in the Clinical Theory of Psychoanalysis: A Study in the Philosophy of Psychoanalysis*. Madison, CT: International Universities Press, 1993.

Küng, Hans. *Freud and the Problem of God*. New Haven, CT: Yale University Press, 1979.

Leibniz, Gottfried W. *New Essays on Human Understanding* (1705). Translated by P. Remnant and J. Bennett. Cambridge, U.K.: Cambridge University Press, 1981.

Salmon, Wesley. *Statistical Explanation and Statistical Relevance*. Pittsburgh, PA: University of Pittsburgh Press, 1971.

Searle, John R. *Intentionality*. New York: Cambridge University Press, 1990.

Solms, Mark, and Michael Saling. *A Moment of Transition: Two Neuroscientific Articles by Sigmund Freud*. New York: Karnac Books, 1990.

Storr, Anthony. "Human Understanding and Scientific Validation." *Behavioral and Brain Sciences* 9 (1986): 259–260.

Zentner, Marcel. *Die Flucht ins Vergessen: Die Anfänge der Psychoanalyse Freuds bei Schopenhauer*. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 1995.

Adolf Grünbaum (2005)

FRIENDSHIP

FRIENDSHIP AND ITS PLACE IN THE MORAL DEBATE

Friendship is a central theme in ancient ethics, most notably in Aristotelian ethics, with two of the ten books of Aristotle's *Nicomachean Ethics* (Books VIII and IX) (1985) devoted to the subject. But modern moral philosophy (from the mid-eighteenth century to the later part of the twentieth century) largely overlooked the role of friendship in moral life, in part because of the dominance of the impartialist stance of utilitarian and Kantian moral theory. Those theories also influenced the study of Aristotelian ethics. In the late 1970s and early 1980s, this trend shifted, in part due to a confluence of causes—renewed interest in Aristotelian ethics for its own sake, the development of modern *virtue ethics*, and the rise of feminist ethical theory. A seminal article by John Cooper on Aristotelian friendship (1977) helped to make Aristotle's account accessible, and especially emphasized the role of friendship in a morally reflective life. Aristotle's account remains the *locus classicus* for understanding the nature of friendship and its place in the moral life; however, before turning to that account, some background is important for understanding its resuscitation in the contemporary moral debate.

THE NEGLECT OF FRIENDSHIP IN MODERN MORAL PHILOSOPHY

From a classical utilitarian view, in the broad tradition of Jeremy Bentham (1748–1832), an agent is obligated to do that which promotes maximally desired outcomes for the greatest number of people, irrespective of standing commitments to friends and family or other personal projects and pursuits. One is to view oneself as a causal lever, Bernard Williams (1963) charged, of optimal outcomes. Thus, if one can save one's spouse or the next inventor of a cure for AIDS, one may be obligated, on a strict utility theory, to save the latter over the former. Rule utilitarians try to counter the unwelcome result, arguing that a general *rule* or *practice* of taking care of kith and kin is an

overall best way of promoting general welfare. But a strict act utilitarian (that is, one committed to assessing the overall good consequences produced by discrete acts) cannot consistently make this response.

From a Kantian view, drawn primarily from Immanuel Kant's early work *The Groundwork from the Metaphysic of Morals* (1785), motives of friendship may be acted upon in morally permissible ways when properly constrained by the impartial point of view of the Categorical Imperative. But even then, such motives, like those of sympathy or other inclinations, lack intrinsic moral worth of their own. So, to adapt a well-known example from Michael Stocker (1976) on a Kantian view, one acts in a morally worthy way when one visits a hospitalized friend not out of friendship, but out of duty. In later writings, Kant seems to soften his view, arguing that acting from friendship may be an important way of realizing the more general, obligatory end of beneficence. Still, Kant is ever wary that intimacy can undermine mutual respect; thus, friendship, is a constant teeter-totter between getting close and keeping at bay: "For we can regard love as attraction and respect as repulsion, and if the principle of love commands friends to come together, the principle of respect requires them to keep each other at a proper distance" (1976, p. 470).

The difficulty of fitting friendship squarely into modern moral theory led many to return to Aristotle's account. This renewal of interest coincides with a feminist push to take seriously the role of interpersonal relationships and caring in a moral point of view. In particular, the influential work by psychologist Carol Gilligan (1982) galvanized philosophers of various stripes to begin to look at friendship and attachment relations as important arenas of moral agency and moral development. Thus, in a sense, the renewed interest in friendship brought with it a rediscovery of the kind of moral psychology that is an integral part of ancient ethics.

FRIENDSHIP IN ARISTOTELIAN ETHICS

The framing question of Aristotelian ethics, like that of most ancient ethics, is what constitutes flourishing or happiness (*eudaimonia*) for human beings? Aristotle's answer is that happiness is a composite of virtuous activity and external goods; chief among those external goods is the relational good of friendship, or *philia*. Humans are by nature "social creatures," Aristotle says, and self-sufficiency is always relational. Even if it turned out that the kind of virtuous or excellent activity most fitting for humans was contemplative and not civic or practical, people would still contemplate best in the company of others (*NE* 1177a33).

According to Aristotle's definition, *philia* is a mutually acknowledged reciprocation of affection and good will on the basis of some ongoing specific interest, such as pleasure, utility, or virtue. Chosen friendship grounded in virtue or good character is the paradigmatic and most stable form of friendship. It is a friendship dedicated to the whole person and committed to the joint project of good living. The best sort of friends "live together" and "spend their days together," not as cattle grazing the same pasture, but "by sharing in argument and thought" (*NE* 1170b11–12). Given the intensity of these ideal friendships, one can reasonably expect to cultivate only a few at a given time. There is much good sense in these views: People are attracted to others on the basis of common pursuits and affinities and show mutual practical concern and good will within the context of the friendship. Were the friendship to dissolve, so, too, would the degree and nature of practical concern for the other.

Aristotle has sometimes been criticized for viewing friendship as a kind of mutual admiration society, and this, in part, because of his remark that a friend is "another self" or a "second self" (*NE* 1170b7). But in the context of his larger discussion, his claim is that people can rely on the best sort of friends to critically see themselves. Friends, he insists, are essential for the process of self-knowledge and for sustaining activities with a kind of zest and zeal that would be hard to muster individually (*NE* 1170a4–6). The best kind of friendship, he insists, is a sphere for moral growth and learning throughout life. And it is so, he concedes, even if friendship, as a kind of external good, exposes the individual to risk of loss and vulnerability. Kant's later worry that intimacy might erode self-sufficiency or autonomy is not Aristotle's concern. People's lives would lack luster without friends and loved ones. One misunderstands the nature of human happiness if one arms against the losses that attachments bring.

EMPATHY AND FRIENDSHIP

However, there is an aspect of friendship that Aristotle never fully articulates, though it is central to a viable conception of friendship. And this is the notion of empathy, or better, mutual empathy. Part of the craving of friendship is to be in synchrony with another. People want their closest friends to track their hearts and minds. They want to know that another can feel their joy or anguish and share concerns and wishes in a way that is psychologically deeper than just formally sharing ends or activities. They

want to know that without too much struggle, a friend can be on the same page and convey that fact in a way that makes it clear that they are understood.

Empathy is an early-twentieth-century psychology word, a Greco translation (from *empatheia*) of the German *Einfülhlung*, to feel one's way into another. A century and a half earlier, the Scottish moral sentiment theorists David Hume and Adam Smith used the term sympathy to mean something similar. For Hume sympathy is a kind of vicarious arousal, a congruent feeling that allows access into others' minds. His model is mechanical: One is connected as if by a cord. A tug at one end causes a reverberation at the other. In this way, one "catches" another's feelings, as if by contagion. Adam Smithproposed a more cognitive account: Feeling another's pain or anguish through an act of imagination; to trade "places in fancy" (1968, p. 4). And this requires some analogical reasoning. As he puts it, one brings another's experiences "home" to oneself; bring the case back to one's own "breast" (1968, p. 4-5). More precisely, one conjures up in one's own mind, through associations and memories, what it would be like to stand in the other's shoes. The process, while cognitive, is not emotionally flat. One must feel something of what the other is feeling, "beat time" with the other, as Smith says (1968, p. 140,146,167). Moreover, to really understand the other's mind, it is not enough for imagination to transport *oneself* into *the other's* shoes. One may have to become *the other* in *the other's* shoes. As Smith puts it in one point, one has to "become in some measure same person with him" (1968, p. 4).

Whether one thinks of empathy as congruent feeling or imaginative transport, one expects close friendship to have some degree of attunement of this sort. The demand is not for a friend to be a mind reader of one's most concealed thoughts; that would be both psychologically implausible and, moreover, an invasion of privacy and autonomy. The point is that one wants some sense of being in sync, of being understood by another in a way that truly makes a life shared. Granted, this can become narcissistic—reminiscent of what an infant demands of a parent and what a parent offers an infant as part of the basic formation of the parent-child bond. Thus, shared eye gaze and reciprocal smiling are part of the early moments of learning mutuality. But a touch of this is what most people still wish for into their adult years. The craving seems a reasonable part of close friendship.

See also Aristotle; Bentham, Jeremy; Hume, David; Kant, Immanuel; Love; Loyalty; Smith, Adam; Virtue Ethics; Williams, Bernard.

Bibliography

Aristotle. *Nicomachean Ethics*, translated by W. D. Ross, rev. J. O. Urmson. In *The Complete Works of Aristotle: The Revised Oxford Translation*, Vol. 2, Bks. VIII, IX, edited by Jonathan Barnes. Princeton: Princeton University Press, 1985. From the same set by Oxford, also see Aristotle's *Eudemian Ethics*. Translated by J. Solomon. VII. 1–2; *Magna Moralia*. Translated by St. G. Stock, II.16; *Rhetoric*, translated by W. R. Roberts, II.4.

Badhwar, N. K., ed. *Friendship: A Philosophical Reader*. Ithaca, NY: Cornell University Press, 1993.

Blum, Lawrence. *Friendship, Altruism, and Morality*. Boston: Routledge & Kegan Paul, 1980.

Cicero, Marcus Tullius. *De amicitia*. Translated by W. A. Falconer. London: Heinemann, 1923.

Cooper, John M. "Aristotle on the Forms of Friendship." *Review of Metaphysics* 30 (1977): 618–648. A revised and condensed version of both original articles appears as "Aristotle on Friendship" in *Essays on Aristotle's Ethics*, edited by A. O. Rorty. Berkeley: University of California Press, 1980.

Cooper, John M. "Friendship and the Good in Aristotle." *Philosophical Review* 86 (1977): 290–315. Reprinted in *Aristotle's Ethics: Critical Essays*, edited by Nancy Sherman. Lanham, MD: Rowman and Littlefield, 1999.

Friedman, Marilyn. *What are Friends For? Feminist Perspectives on Personal Relationships and Moral Theory*. Ithaca: Cornell University Press: 1993.

Gilligan, Carol. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press, 1982.

Hume, David. *A Treatise of Human Nature*, edited by L.A. Selby-Bigge. Oxford: Oxford University Press, 1968 (1739).

Pakaluk, Michael. *Other Selves: Philosophers on Friendship*. Indianapolis, IN: Hackett, 1991.

Kant, Immanuel. *Doctrine of Virtue*. Part II of *The Metaphysic of Morals*. Translated by Mary J. Gregor. Philadelphia: University of Pennsylvania Press, 1964 (1797).

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by H.J. Paton. New York: Harper and Row, 1964 (1785).

Sherman, Nancy. "Empathy and Imagination." In *Philosophy of Emotions*. Edited by Peter French and Howard Wettstein. Midwest Studies in Philosophy, Vol. 22. Notre Dame, IN: University of Notre Dame Press, 1998, p. 82–199.

Sherman, Nancy. "The Shared Life." Chap. 4 in *Fabric of Character* Oxford: Oxford University Press, 1989, p. 118–157; *Philosophy and Phenomenological Research* 47 (1987): 589–613.

Sherman, Nancy. "The Shared Voyage." Chap. 5 in *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge, MA: Cambridge University Press, 1997, 187–238.

Smart, J. J. C., and Bernard Williams. *Utilitarianism, For and Against*. Cambridge, U.K.: Cambridge University Press, 1973.

Smith, Adam. *The Theory of Moral Sentiments*. New York: Prome Thesus Books, 2000 (1759).

Stocker, Michael. "The Schizophrenia of Modern Ethical Theories." *Journal of Philosophy* 73 (1976): 453–466.

Stocker, Michael. "Values and Purposes: The Limits of Teleology and the Ends of Friendship." *Journal of Philosophy*

78 (1981): 747–765. Reprinted in *Friendship: A Philosophical Reader*, edited by N. K. Badhwar. Ithaca, NY: Cornell University Press, 1993.

Telfer, E. "Friendship." *Proceedings of the Aristotelian Society* Supp. (1970–1971): 223–241.

Nancy Sherman (2005)

FRIES, JAKOB FRIEDRICH

(1773–1843)

Jakob Friedrich Fries, the German critical philosopher, was born in Barby, Saxony. An avowed follower and elaborator of Immanuel Kant's philosophy, Fries emphasized the analytical, descriptive, and methodological aspects of the critical philosophy as against the constructive and speculative idealism of such contemporaries as K. L. Reinhold, Johann Gottlieb Fichte, Friedrich Schelling, and G. W. F. Hegel. He received his secondary and college education at the Moravian Academy in Niesky. From his Pietistic Moravian background Fries preserved a conviction of the importance of "pure feeling" as a manifestation of "the infinite in the finite." At Niesky he was given a thorough grounding in mathematics and in the natural sciences. There he was also introduced to a version of Kant's philosophy based on Reinhold's, which he early sought to correct and supplement by secretly reading in Kant's own writings. In 1795 he went to Leipzig, where he studied under the philosopher-physician Ernst Plattner.

The influence of Plattner and of F. H. Jacobi accounts for Fries's emphasis on the concept of self-observation. From 1797 on, Fries continued his studies in mathematics and physics at Jena, where he also attended Fichte's lectures. "I listened to Fichte, took notes, then rushed home and wrote rebuttals," he later recalled. These critical notes were incorporated into his polemical writings. As early as 1798, in the article "Über das Verhältniss der empirischen Psychologie zur Metaphysik" (On the relation of empirical psychology to metaphysics; in *Erhard Schmids Psychologische Magazin*, Vol. 3), he argued that the task of philosophy is essentially descriptive rather than speculative.

Following his studies at Jena, Fries served as a private tutor in Switzerland, then returned to Jena as a docent in 1801, submitting a habilitation thesis on intellectual intuition. The polemical tract *Reinhold, Fichte und Schelling* (Leipzig, 1803) established his reputation as a critic of the romantic orthodoxy in German philosophy. From Jena he was called to a professorship in philosophy and mathematics at Heidelberg. That year he published *Wissen, Glaube und Ahndung* (Knowing, faith and presage; Jena, 1805), a popular exposition of his doctrine of a threefold approach to reality. This was followed, during the years 1806–1807, by his chief work, the three-volume *Neue Kritik der Vernunft* (New critique of reason; Heidelberg, 1807; 2nd ed., *Neue oder anthropologische Kritik der Vernunft*, Heidelberg, 1828–1831), in which he attempted to correct and restate the Kantian critique of speculative and practical reason as a program of psychological self-observation or "anthropology."

PROGRESSIVE POLITICAL VIEWS

A decisive shift in Fries's career occurred in 1816, when he returned to Jena to a professorship in theoretical philosophy. Under the tolerant and liberal regime of Duke Karl August, he published his *Ethik* (Heidelberg, 1818), a work in which he stressed the ideal of individual liberty and political equality as a consequence of the Kantian doctrine of the dignity proper to a human being. In pamphlets and lectures and at student gatherings during this period, Fries argued for constitutional and representative government, extolled the political wisdom of the "people," opposed the conservatism of student secret societies, and advocated German unification. This activity, climaxed by his participation in the Wartburg Festival of October 18, 1817—a demonstration by student liberals that included a ceremonial burning of "reactionary" books—inevitably incurred the wrath of the Austrian and Prussian governments. It also elicited scornful comments from the politically more orthodox Hegel, who in his *Philosophy of Right* (translated by J. M. Knox, Oxford, 1942) downgraded Fries as "a ringleader of those hosts of superficiality, of these self-styled 'philosophers,'" and attacked his Wartburg speech as "the quintessence of shallow thinking … a broth of 'heart, friendship, and inspiration.'" By 1819 the conservative opposition had prevailed upon Karl August, and Fries was suspended from his position at Jena. He had earlier lost hope that he would be offered the chair of philosophy at Berlin, which in 1818 went to Hegel.

Although Fries was eventually allowed to resume teaching at Jena (he taught science from 1824 and philosophy from 1825 on) and held this post until his death, the 1819 suspension was the final turn in his estrangement and isolation from the intellectual currents of the period. From then on, supported by a small following, he devoted his life to studies of mathematics, physics, and psychology, to systematization of his metaphysics and ethics, and to a rewriting of the history of philosophy on the theme of "progress in scientific development." To this period

belong *Die mathematische Naturphilosophie* (Heidelberg, 1822); *System der Metaphysik* (Heidelberg, 1824); *Handbuch der psychischen Anthropologie* (2 vols., Jena, 1820 and 1821); and *Die Geschichte der Philosophie* (2 vols., Halle, 1837–1840).

APPROACHES TO REALITY

Fries followed Immanuel Kant in the overall architectonic of his philosophy and in specific doctrines. Corresponding to Kant's three Critiques, he distinguished three approaches or attitudes toward reality—knowing, faith, and presage, or presentiment (*Ahndung*). We know things only as appearances to a peculiarly human sensibility and understanding. But we have faith in the reality of a world of real moral agents under eternal moral laws. Our understanding is aware of this world only negatively, as a limitation of the empirical world, through the Ideas of Reason. Finally, through presage or presentiment, a pure and disinterested feeling akin to the experience of the beautiful and the sublime, we are given the assurance that the world of appearances and the real world are not two worlds but one, and that the former is a manifestation of the latter—a finite projection of the infinite into the finite.

TYPES OF KNOWLEDGE

Within the sphere of knowing, Fries distinguished two levels: original or immediate knowledge, and reflective or mediate knowledge. The types of mediate knowledge are given in the Kantian forms of judgment: analytic, synthetic a posteriori, and synthetic a priori. We must also distinguish three types of immediate knowledge. An empirical intuition is a direct awareness of the sensory given; a pure intuition is a direct awareness of space and time as empty containers of sensible entities; and an immediate metaphysical cognition is the direct but nonintuitive awareness of principles involving the categories of the understanding (for example, the principle of causality or the principle of the permanence of substance). No attempt to reduce cognitions of the second and third types to cognitions of the first type can ever succeed. Space and time are the forms of our empirical intuitions; the categories are the forms of human understanding. Fries thus shared with Kant the critical solution of the problem of a priori knowledge. He also shared with Kant the rejection of both the empiricist and the intellectual intuitionist solutions of the problem.

METAPHYSICAL KNOWLEDGE

Fries departed from Kant, however, in his interpretation of the basis for the critical solution in the case of a priori metaphysical knowledge. Fries found inconsistency and circularity in Kant's attempt to validate categories and to "prove" the principles of the understanding by referring them to "the possibility of experience." If these are indeed principles, no proof could be required and none would be sufficient. Kant succumbed, in Fries's judgment, to the ancient rationalist prejudice that everything can be proved and that all truths can be reduced to a single principle—in Kant's case, the concept of possible experience. All that is possible, Fries objected, is to display the status of certain cognitions *as* a priori and necessary. "I do not prove," he explained, "that all substance is permanent; rather I point to the fact that the principle of the permanence of substance lies in every finite mind" (*Neue oder anthropologische Kritik der Vernunft*, 2nd ed., Vol. I). In Kant's language, only a "metaphysical deduction" (the answer to the question, "What is the case?") is possible.

DISCOVERY OF METAPHYSICAL PRINCIPLES

The regression to metaphysical principles is not an easy task, for unlike empirical and pure intuitions, which are clear and readily available to consciousness, metaphysical principles lie "concealed and obscure" in the depths of human reason. Fries described this regression as a process of self-observation or "psychic anthropology," and likened it to experimental physics insofar as the latter aims to discover the general law involved in specific physical phenomena. Kant, accordingly, misunderstood the function of critical philosophy and the status of the judgments that constitute it, for whereas the truths that critical philosophy aims to uncover are nonempirical and necessary, the critique itself is empirical and fallible. Fries admired the long "subjective" deductions of the categories in the first edition of Kant's *Critique of Pure Reason* but was skeptical of the short "objective" deductions of the second edition.

PROOF

A complete theory of proof must, therefore, distinguish three kinds: (1) demonstration, or the reduction of a "reflective" or "mediate" cognition to an intuition (pure or empirical); (2) proof, or the reduction of one mediate cognition to another; and (3) deduction, or a regressive analysis that traces a given cognition to its ground in immediate metaphysical knowledge. Just as in the case of demonstration, in which no question can arise concern-

ing the validity of the intuitions themselves, so too, in the case of deduction, no question can possibly arise as to the validity of our immediate metaphysical knowledge. A deduction is, of course, something fallible; closer scrutiny may later reveal a disparity between a given cognition and its supposed ground. But the same danger exists for demonstration—in this case minimally.

TRUTH

Truth is a matter of correspondence between thought and object, but the object is not something transcendent; it is simply an immediate cognition. Truth is a relation between two levels of cognition. With regard to immediate knowledge itself we must accept the principle of "Reason's self-reliance" (*Selbstvertrauen der Vernunft*), that is, that we possess such knowledge and that it is intrinsically valid.

FRIES'S "PSYCHOLOGISM"

Fries's restatement of the Kantian deduction has often been attacked as psychologism. If psychologism is understood as the attempt to find the validity of human beliefs in their psychological causes and in the laws of association, the charge is unfair. Fries was not a proponent of psychologism in that sense: For him, the validity of immediate knowledge lay in its logical character, universality, and necessity, not in its causal origins. Indeed, Fries wrote critically against such contemporary advocates of psychologism as Friedrich Eduard Beneke, with whom he was sometimes mistakenly compared. At the same time, he did seem to suggest that logical character can be gathered from mere psychological observation of our mental processes. And in this connection he has been justly criticized for confusing a mental act with its logical content. Certainly the process that Fries described as "anthropology" would be more accurately described today as "logical" or "phenomenological" analysis. Fries was perhaps misled by the analogy between a logical regress to presuppositions in philosophy and the heuristic regress (induction) to general hypotheses and theories in physics.

SCIENCE AND MATHEMATICS

In the fields of mathematics, physics, and psychology, Fries's thought was highly original and expertly worked out. He had a clear conception of a philosophy of mathematics and physics as an independent discipline, and anticipated the modern distinction between a theory and a metatheory. In his theory of nature he attacked Kant's concessions to teleology and argued for a thoroughgoing mechanism that would also encompass the biological sciences. His psychological investigations extended into the study of pathological phenomena. He took note of the distinction between inherited and acquired, as well as between continuous and periodic, mental disorders and argued for the physiological basis of mental illness—concepts that were by no means as current in Fries's time as they are in ours and that were unfortunately ignored by the psychiatric practitioners of his day.

FRIES'S INFLUENCE

Fries was succeeded at Jena by his pupil E. F. Apelt, who published a masterly textbook of Friesian metaphysics and in 1847 established the journal *Abhandlungen der Fries'schen Schule*, which served for two years as a forum for critical, scientifically oriented philosophy. There was a revival of interest in Fries and in his approach to Kant in the years preceding and immediately following World War I, centering about Leonard Nelson at Göttingen, who shared Fries's scientific outlook and reacted to the idealist Neo-Kantian orthodoxy in the German universities of his time much as Fries had reacted to Fichte, Schelling, and Hegel. The theologian Rudolf Otto, an early associate of Nelson, developed Fries's concept of "presage" in his influential book *Das Heilige* (*The Idea of the Holy*, Gotha, 1917). In 1904 Nelson established a new series of the *Abhandlungen der Fries'schen Schule* of which six volumes appeared before publication was discontinued in 1937. National Socialism proved itself as inimical to Nelson's school as Klemens von Metternich's political reaction had been to Fries's. In 1958 Julius Kraft, a student of Nelson's, founded the philosophical journal *Ratio* as a continuation of the *Abhandlungen*.

Although Fries's influence was and remains limited, part of the interest that his philosophy holds for the modern reader lies in its analogues with, and anticipations of, positions and problems that were central in twentieth-century thought, especially in England and the United States. There is, first, an obvious but quite unexplored analogy between Fries's psychological method and Edmund Husserl's phenomenology. Moreover, the view that metaphysical principles can only be exhibited as such but not proved has been variously defended by R. G. Collingwood and by representatives of the Oxford school of linguistic analysis. There are also apparent counterparts of Fries's "self-reliance of Reason" in G. E. Moore's appeal to common sense, in the positivists' appeal to a level of incorrigible knowledge, and in Ludwig Wittgenstein's famous dictum that "the propositions of our ordinary language are in perfect order." Indeed, the question

of the status of the propositions employed by the critical or analytical philosopher, which was first raised by Fries, has come under much discussion in recent years, under the heading "the problem of analysis."

See also Collingwood, Robin George; Faith; Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; Jacobi, Friedrich Heinrich; Kant, Immanuel; Knowledge, A Priori; Moore, George Edward; Nelson, Leonard; Neo-Kantianism; Otto, Rudolf; Psychologism; Reinhold, Karl Leonhard; Schelling, Friedrich Wilhelm Joseph von; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY FRIES

A list of all Fries's writings will be found in *Abhandlungen der Fries'schen Schule*, n.s., 6 (1937): 473–495.

With G. König and Lutz Geldsetzer. *Sämtliche Schriften*. Aalen: Scientia Verlag, 1967.

With D. Z. Phillips. *Dialogues on Morality and Religion*. Totowa, NJ: Barnes & Noble, 1982.

With Frederick Gregory. *Knowledge, Belief, and Aesthetic Sense*. Köln: Jürgen Dinter, Verlag für Philosophie, 1989.

On Fries's life and writings, the fundamental work remains E. L. T. Henke, *Jakob Friedrich Fries* (Leipzig, 1867). The last detailed study of Fries's theory of knowledge is in Alfred Kastil, *Fries' Lehre von der Unmittelbaren Erkenntnis*, a monograph that constitutes *Abhandlungen der Fries'schen Schule* 4 (1912). For a short critical assessment of the same subject, see Ernst Cassirer, *Das Erkenntnissproblem in der Philosophie und Wissenschaft der neueren Zeit* (Berlin, 1923), Vol. III, Ch. 7. A stimulating but partisan account of Fries's philosophy will be found in Leonard Nelson, *Fortschritte und Rückschritte der Philosophie* (Frankfurt, 1962). In connection with Fries's theory of law and religion, especially interesting are Julius Kraft, *Die Methode der Rechtstheorie in der Schule von Kant und Fries* (Berlin, 1924) and Rudolf Otto, *Kantisch-Fries'che Religionsphilosophie und ihre Anwendung auf die Theologie* (Tübingen: Mohr, 1909), translated by E. B. Dicker as *The Philosophy of Religion Based on Kant and Fries* (London: Williams and Norgate, 1931).

Alexander P. D. Mourelatos (1967)
Bibliography updated by Michael Farmer (2005)

FROEBEL, FRIEDRICH
(1782–1852)

Friedrich Froebel, the German philosopher of education, was born at Oberweissbach in Thuringia. He studied forestry and related fields at the University of Jena, came in contact with Johann Heinrich Pestalozzi in 1808, and participated as a volunteer in the war of liberation against Napoleon Bonaparte. In 1816 he established a school, which soon was moved from Griesheim to Keilhau, and in 1837 he founded his first kindergarten at Blankenburg in Thuringia, which became the model of many similar institutions. However, these institutions had to be closed in Prussia in 1851 because the government, as well as the clergy, suspected Froebel of liberal political and religious leanings. The prohibition lasted for ten years, but afterward the kindergarten movement spread rapidly throughout the European countries.

Froebel's whole educational theory and practice was determined by his conviction of the ultimate oneness of life, of nature and spirit. According to him it is the destiny of all things to unfold their divine essence and to reveal God in their transient being.

As Froebel's autobiography shows, he was, as a child, deeply troubled by the contrast between "spirit" and "the flesh" in the Christian supernaturalism and moralistic dualism of his father, a pastor, until he discovered the pervasive beauty of nature and the mystery of sex life in the whole creation. His conviction about the inner unity of the cosmos was confirmed by his scientific studies, his reading of the Zend-Avesta, and his acquaintance with Friedrich Schelling's philosophy of identity.

In conformity with his metaphysics, Froebel conceived of education as a continuation of the world's unceasing evolution on the level of consciousness, with the child's play being the first sign of life's urge toward purposeful activity. Thus, he wrote in *The Education of Man* (pp. 1ff.), "Education consists in leading man, as a thinking, intelligent being, growing into self-consciousness, to a pure, unsullied, conscious and free representation of the inner law of divine unity, and in teaching him means thereto." Out of respect for the "inner law of unity" or for life as an "unbroken whole in all its operations and phenomena" (*The Education of Man,* p. 238), the educator should organize the instructional process in such a way that the order of the subjects to be taught supports the learner's inner development, while the whole program of studies should help the student to realize the reflection of the unity of life in the unity of knowledge.

Froebel's educational principles may be summarized as follows:

(1) That the development of nature reveals itself in the development of the individual mind should be demonstrated in the teaching of science, the humanities, and religion.

(2) Education should be structured to harmonize with the natural inner development of the pupil.

(3) Education should unfold the whole man in each person. Religion should be taught in order to cultivate the emotions, nature should be studied because it is the self-revelation of God, and mathematics should be appreciated as the symbol of universal order. Language, too, connects man with the order and rhythm of things and should therefore take its part in education.

(4) The arts should be taught, for art is a general human talent and conducive to the harmonious unfolding of a person's inner life.

The central theme in Froebel's educational work is most evident in *The Education of Man,* which presents a unique attempt to provide an ontological explanation of the process of human learning.

Historically, Froebel must be understood as being in the tradition of John Comenius, Jean-Jacques Rousseau, and Pestalozzi. His ideas have been criticized for many and sometimes contradictory reasons: for their pantheistic naturalism, their romanticism, their individualism and neglect of discipline, their sentimentality and their one-sided emphasis on early childhood. But there can be no doubt that the work with which his name is mainly connected, the kindergarten, has been an inestimable blessing to humankind, and many of his psychological insights, like those of Pestalozzi, have been increasingly confirmed by modern psychology.

See also Art, Value in; Comenius, John Amos; German Philosophy; Pestalozzi, Johann Heinrich; Schelling, Friedrich Wilhelm Joseph von; Philosophy of Education, History of; Rousseau, Jean-Jacques.

Bibliography

Froebel's *Gesammelte pädagogische Schriften* was edited by Richard Lange (Berlin, 1862–1874). Of the many translations of this work, the following are most important: *Froebel's Chief Writings on Education,* translated by S. S. F. Fletcher and J. Walton (New York, 1912); *Education by Development,* translated by J. Jarvis (New York: Appleton, 1899); and *The Education of Man,* translated by W. N. Hailmann (New York: Appleton, 1887).

For literature on Froebel, see Robert Ulich, *History of Educational Thought* (New York: American Book, 1950), and his *Three Thousand Years of Educational Wisdom* (Cambridge, MA: Harvard University Press, 1954).

Robert Ulich (1967)

FRYE, MARILYN
(1941–)

Marilyn Frye, American feminist philosopher, was born in Tulsa, Oklahoma. She earned her bachelor's degree in philosophy from Stanford University in 1963, and her doctorate in philosophy at Cornell University in 1969, where she worked under the supervision of the analytic philosopher Max Black. She taught at the University of Pittsburgh, the University of Michigan, and the University of Washington before taking up a position at Michigan State University, where she was tenured in 1978, promoted to professor in 1983, and named University Distinguished Professor in 2003, the position she currently holds. Frye has held fellowships at the Center for the Study of Women in Society at the University of Oregon, the Center for Advanced Feminist Studies at the University of Minnesota, and the National Humanities Center in North Carolina. In 2001 she was awarded the Distinguished Woman Philosopher Award by the Society for Women in Philosophy.

Frye's writings reflect the analytic philosophical style of conceptual analysis and display clear, concise, jargon-free writing, though she applies this to subjects beyond the pale of the narrow world of analytic philosophy. Frye's dissertation, "Meaning and Illocutionary Force," and her first several articles in philosophy were on topics in philosophy of language. Subsequently she turned to topics in feminist philosophy, especially sexism, lesbianism, and racism, and it is in this field that she has made her most important contributions to philosophy. Frye expresses unusual commitment to bringing about social change through her writings. Moreover, she expresses herself with a pragmatic urgency frequently lacking in most professional philosophy, and she also makes exceptionally clear the time-bound and culture-bound nature of such change.

Frye's book *The Politics of Reality* (1983) begins with one of her most important and most often reprinted essays: "Oppression." In this essay she seeks to clarify the term "oppression" and how women can be said to be oppressed. Oppression, on her analysis, is a network of (often microscopic) forces that bind and confine certain social groups within a defined place so as to benefit a privileged social group. She analogizes oppression to a birdcage, which is macroscopic and visible, even though each of the wires of the cage is itself small and seemingly inconsequential in itself. Frye describes two characteristic features of women's oppression. First, women hold positions that simultaneously make them responsible yet

powerless to effect decisions to carry out their responsibilities successfully. Second, women internalize and self-police their limitations and restrictions. While men also face social restrictions (e.g., they cannot cry in front of other men), their restrictions are a part of a system that oppresses women and privileges men. In her essay "Sexism," Frye defines "sexism" as an institutional term characterizing social structures that "create and enforce the elaborate and rigid patterns of sex-marking and sex-announcing which divide the species, along the line of sex, into dominators and subordinates" (1983, p. 38). She uses the term "male-chauvinism" to describe the personal relations that men engage in as dominators with women as subordinates. Most of the essays of the book are devoted to illuminating the social and personal relations that serve to oppress women.

In her writings, Frye illuminates the oppression of sexual minorities by heterosexuals and the oppression of minority races, and she connects these to the project of feminism. In two essays in her first book and in the majority of the essays of her book *Willful Virgin: Essays in Feminism* (1992), Frye takes up the theme of heterosexism as manifested in feminism and society at large. She carefully describes and analyzes the myriad ways in which heterosexuality is taken to be normative. In her essay "Willful Virgin, or Do You Have to Be a Lesbian to Be a Feminist," Frye argues, "The central constitutive dynamic and key mechanism of the global phenomenon of male domination, oppression and exploitation of females is near-universal female heterosexuality" (1992, p. 129). By the term "female heterosexism" she refers not to a preference to engage in heterosexual sex, but rather to the worship of men and maleness that heterosexuality has traditionally required of women. That is, sexism exists because most women willingly tolerate being subordinate to and serving men. Furthermore, because women are subordinate to "their" men, they often comply with whatever other oppression their men perpetrate, such as racism, classism, and ethnic oppression. Thus, not participating in the patriarchal institution of female heterosexuality is an important kind of resistance to oppression generally.

Frye also devotes particular attention to the struggle against racism. She notes that acting White is a way of being privileged, yet for women, acting White consists largely of conformity to white men's expectations of chastity, obedience, and decorum, does not offer any solace to white women, and serves only to separate them from other women. Thus for Frye, Whiteness, heterosexuality, and sexism are bound together in ways that institute and enforce patriarchy.

See also Feminist Philosophy.

Bibliography

WORKS BY FRYE

"Inscriptions and Indirect Discourse." *Journal of Philosophy* 61 (24) (1964): 767–772.

"Force and Meaning." *Journal of Philosophy* 70 (1974): 281–294.

The Politics of Reality: Essays in Feminist Theory. Trumansburg, NY: Crossing Press, 1983.

Willful Virgin: Essays in Feminism, 1976–1992. Freedom, CA: Crossing Press, 1992.

"The Necessity of Differences: Constructing a Positive Category of Women." *Signs: Journal of Women in Culture and Society* 21 (3) (1996): 991–1010.

Ann E. Cudd (2005)

FUNCTIONALISM

"Functionalism" is one of the major proposals that have been offered as solutions to the mind-body problem. Solutions to the mind-body problem usually try to answer questions such as: What is the ultimate nature of the mental? At the most general level, what makes a mental state mental? Or more specifically, what do thoughts have in common in virtue of which they are thoughts? That is, what makes a thought a thought? What makes a pain a pain? Cartesian dualism said the ultimate nature of the mental was to be found in a special mental substance. Behaviorism identified mental states with behavioral dispositions; physicalism, in its most influential version, identifies mental states with brain states. Functionalism says that mental states are constituted by their causal relations to one another and to sensory inputs and behavioral outputs. Functionalism is one of the major theoretical developments of twentieth-century analytic philosophy, and provides the conceptual underpinnings of much work in cognitive science.

Functionalism has three distinct sources. First, Hilary Putnam and Jerry Fodor saw mental states in terms of an empirical computational theory of the mind. Second, John Jamieson Carswell Smart's "topic neutral" analyses led David M. Armstrong and David Lewis to a functionalist analysis of mental concepts. Third, Ludwig Wittgenstein's idea of meaning as use led to a version of functionalism as a theory of meaning, further developed by Wilfrid Sellars and later Gilbert Harman.

One motivation behind functionalism can be appreciated by attention to artifact concepts such as *carburetor*, and biological concepts such as *kidney*. What it is for something to be a carburetor is for it to mix fuel and air in an internal combustion engine—*carburetor* is a functional concept. In the case of the kidney, the *scientific* concept is functional—defined in terms of a role in filtering the blood and maintaining certain chemical balances.

The kind of function relevant to the mind can be introduced via the parity-detecting automaton illustrated in the following figure, which tells us whether it has seen an odd or even number of "1"s. This automaton has two states, S_1 and S_2; one input, "1" (though its input can be nothing) and two outputs, it utters either the word "Odd" or "Even." The table describes two functions, one from input and state to output, and another from input and state to next state. Each square encodes two conditionals specifying the output and next state given both the current state and input. The left box says that if the machine is in S_1 and sees a "1," it says "odd" (indicating that it has seen an odd number of "1"s) and goes to S_2. The right box says, similarly, that if the machine is in S_2 and sees a "1," it says "even" and goes back to S_1.

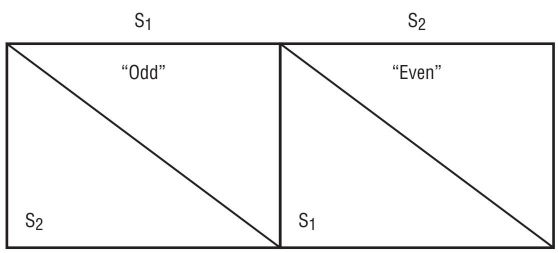

Now suppose we ask the question: "What is S_1?" The answer is that the nature of S_1 is entirely relational, and entirely captured by the table. We could give an explicit characterization of "S_1" as follows:

> Being in S_1 = being in the first of two states that are related to one another and to inputs and outputs as follows: Being in one of the states and getting a "1" input results in going into the second state and emitting "Odd"; and being in the second of the two states and getting a "1" input results in going into the first and emitting "Even."

Making the quantification over states more explicit:

> Being in S_1 = Being an x such that ∃P∃Q[If x is in P and gets a "1" input, then it goes into Q and emits "Odd"; if x is in Q and gets a "1" input it gets into P and emits "Even" & x is in P] (Note: Read "∃P" as "There is a property *P*.")

This illustration can be used to make a number of points. (1) According to functionalism, the nature of a mental state is just like the nature of an automaton state: constituted by its relations to other states and to inputs and outputs. All there is to S_1 is that being in it and getting a "1" input results in such and such, and so forth. According to functionalism, all there is to being in pain is that it disposes you to say "ouch," wonder whether you are ill, it distracts you, and so forth. (2) Because mental states are like automaton states in this regard, the illustrated method for defining automaton states is supposed to work for mental states as well. Mental states can be totally characterized in terms that involve only logicomathematical language and terms for input signals and behavioral outputs. Thus functionalism satisfies one of the desiderata of behaviorism, characterizing the mental in entirely nonmental language.

(3) S_1 is a second-order state in that it consists in having other properties, say mechanical or hydraulic or electronic properties, that have certain relations to one another. These other properties, the ones quantified over in the definitions just given, are said to be the realizations of the functional properties. So, although functionalism characterizes the mental in nonmental terms, it does so only by quantifying over realizations of mental states, which would not have delighted behaviorists. (4) One functional state can be realized in different ways. For example, an actual metal and plastic machine satisfying the machine table might be made of gears, wheels, pulleys and the like, in which case the realization of S_1 would be a mechanical state; or the realization of S_1 might be an electronic state, and so forth.

(5) Just as one functional state can be realized in different ways, one physical state can realize different functional states in different machines. This could happen, for example, if a single type of transistor were used to do different things in different machines. (6) Since S_1 can be realized in many ways, a claim that S_1 is a mechanical state would be false (at least arguably), as would a claim that S_1 is an electronic state. For this reason, there is a strong case that functionalism shows physicalism is false: If a creature without a brain can think, thinking cannot be a brain state. (But see the section on functionalism and physicalism below.)

The notion of a realization deserves further discussion. In the early days of functionalism, a first-order property was often said to realize a functional property in virtue of a 1-1 correspondence between the two realms of properties. But such a definition of realization produces far too many realizations. Suppose, for example, that at t_1

we shout "one" at a bucket of water, and then at t_2 we shout "one" again. We can regard the bucket as a parity-detecting automaton by pairing the physical configuration of the bucket at t_1 with S_1 and the heat emitted or absorbed by the bucket at t_1 with "odd"; by pairing the physical configuration of the bucket at t_2 with S_2 and the heat exchanged with the environment at t_2 with "even"; and so on. What is left out by the post hoc correlation way of thinking of realization is that a true realization must satisfy the *counterfactuals* implicit in the table. To be a realization of S_1, it is not enough to lead to a certain output and state given that the input is a "1"; it is also required that had the input been a "0," the S_1 realization would have led to the other output and state. Satisfaction of the relevant counterfactuals is built into the notion of realization mentioned in (3) above.

Suppose we have a theory of mental states that specifies all the causal relations among the states, sensory inputs, and behavioral outputs. Focusing on pain as a sample mental state, it might say, among other things, that sitting on a tack causes pain, and that pain causes anxiety and saying "ouch." Agreeing for the sake of the example to go along with this moronic theory, functionalism would then say that we could define "pain" as follows: being in pain = being in the first of two states, the first of which is caused by sitting on tacks, which in turn causes the other state and emitting "ouch." More symbolically

> Being in pain = Being an x such that ∃P∃Q[sitting on a tack causes P & P causes both Q and emitting "ouch" & x is in P]

More generally, if T is a psychological theory with n mental terms of which the 17th is "pain," we can define "pain" relative to T as follows (the "F_1" … "F_n" are variables that replace the n mental terms, and i_1, etc. And o_1, etc. indicates):

> Being in pain = Being an x such that $\exists F_1 \ldots \exists Fn$ $[T(F_1 \ldots F_n, i_1, \text{etc.}, o_1, \text{etc.})$ & x is in $F_{17}]$

In this way, functionalism characterizes the mental in nonmental terms, in terms that involve quantification over realizations of mental states but no explicit mention of them; thus functionalism characterizes the mental in terms of structures that are tacked down to reality only at the inputs and outputs.

The psychological theory T just mentioned can be either an empirical psychological theory or else a commonsense "folk" theory, and the resulting functionalisms are very different. In the latter case, conceptual functionalism, the functional definitions are aimed at capturing our ordinary mental concepts. In the former case, "psychofunctionalism," the functional definitions are not supposed to capture ordinary concepts but are only supposed to fix the extensions of mental terms. The idea of psychofunctionalism is that the scientific nature of the mental consists not in anything biological, but in something "organizational," analogous to computational structure. Conceptual functionalism, by contrast, can be thought of as a development of logical behaviorism. Logical behaviorists thought that pain was a disposition to pain behavior. But as Peter Geach and Roderick Chisholm pointed out, what counts as pain behavior depends on the agent's beliefs and desires. Conceptual functionalists avoid this problem by defining each mental state in terms of its contribution to dispositions to behave—and have other mental states.

FUNCTIONALISM AND PHYSICALISM

Theories of the mind prior to functionalism have been concerned both with (1) what there is, and (2) what gives each type of mental state its own identity, for example what pains have in common in virtue of which they are pains. Stretching these terms a bit, we might say that (1) is a matter of ontology and (2) of metaphysics. Here are the ontological claims: Dualism told us that there are both mental and physical substances, whereas behaviorism and physicalism are monistic, claiming that there are only physical substances. Here are the metaphysical claims: Behaviorism tells us that what pains (for example) have in common in virtue of which they are pains is something behavioral; dualism gave a nonphysical answer to this question, and physicalism gives a physical answer to this question.

Turning now to functionalism, it answers the metaphysical question without answering the ontological question. Functionalism tells us that what pains have in common—what makes them pains—is their function; but functionalism does not tell us whether the beings that have pains have any nonphysical parts. This point can be seen in terms of the automaton described above. In order to be an automaton of the type described, an actual concrete machine need only have states related to one another and to inputs and outputs in the way described. The machine description does not tell us how the machine works or what it is made of, and in particular it does not rule out a machine which is operated by an immaterial soul, so long as the soul is willing to operate in the deterministic manner specified in the table.

In thinking about the relation between functionalism and physicalism, it is useful to distinguish two cate-

gories of physicalist theses: One version of physicalism competes with functionalism, making a metaphysical claim about the physical nature of mental state properties or types (and is thus often called "type" physicalism). As mentioned above, on one point of view, functionalism shows that type of physicalism is false.

However, there are more modest physicalisms whose thrusts are ontological rather than metaphysical. Such physicalistic claims are not at all incompatible with functionalism. Consider, for example, a physicalism that says that every actual thing is made up entirely of particles of the sort that compose inorganic matter. In this sense of physicalism, most functionalists have been physicalists. Further, functionalism can be modified in a physicalistic direction, for example, by requiring that all properties quantified over in a functional definition by physical properties. Type physicalism is often contrasted with token physicalism. (The word *teeth* in this sentence has five letter tokens of three letter types.) Token physicalism says that each pain (for example) is a physical state, but token physicalism allows that there may be nothing physical that all pains share, nothing physical that makes a pain a pain.

It is a peculiarity of the literature on functionalism and physicalism that while some functionalists say functionalism shows physicalism is false, others say functionalism shows physicalism is true. In Lewis's case, the issue is partly terminological. Lewis is a conceptual functionalist about "having pain." Having pain on Lewis's regimentation could be said to be a rigid designator of a functional property. (A rigid designator names the same thing in each possible world. "The color of the sky" is nonrigid, since it names red in worlds in which the sky is red. "Blue" is rigid, since it names blue even in worlds in which the sky is red.) "Pain," by contrast, is a nonrigid designator conceptually equivalent to a definite description of the form "the state with such and such a causal role." The referent of this phrase in us, Lewis holds, is a certain brain state, though the referent of this phrase in a robot might be a circuit state, and the referent in an angel would be a nonphysical state. Similarly, "the winning number" picks out "17" in one lottery and "596" in another. So Lewis is a functionalist (indeed a conceptual functionalist) about having pain.

In terms of the metaphysical issue described above—what do pains have in common in virtue of which they are pains—Lewis is a functionalist, not a physicalist. What a person's pains and the robot's pains share is a causal role, not anything physical. Just as there is no numerical similarity between 17 and 596 relevant to their being winning numbers, there is no physical similarity between human and Martian pain that makes them pains. And there is no physical similarity of any kind between human pains and angel pains. However, on the issue of the scientific nature of pain, Lewis is a physicalist. What is in common to human and Martian pain in his view is something conceptual, not something scientific.

FUNCTIONALISM AND PROPOSITIONAL ATTITUDES

The discussion of functional characterization given above assumes a psychological theory with a finite number of mental state terms. In the case of monadic states like pain, the sensation of red, and so forth, it does seem a theoretical option simply to list the states and their relations to other states, inputs and outputs. But for a number of reasons, this is not a sensible theoretical option for belief-states, desire-states, and other propositional attitude states. For one thing, the list would be too long to be represented without combinatorial methods. Indeed, there is arguably no upper bound on the number of propositions, any one of which could in principle be an object of thought. For another thing, there are systematic relations among beliefs: for example, the belief that John loves Mary and the belief that Mary loves John. These belief states represent the same objects as related to each other in converse ways. But a theory of the nature of beliefs can hardly just leave out such an important feature of them. We cannot treat "believes-that-grass-is-green," "believes-that-grass-is-blue," and so forth, as unrelated primitive predicates. So we will need a more sophisticated theory, one that involves some sort of combinatorial apparatus.

The most promising candidates are those that treat belief as a relation. But a relation to what? There are two distinct issues here. One issue is how to state the functional theory in a detailed way. A second issue is what types of states could possibly realize the relational propositional attitude states. Hartry Field and Fodor argue that to explain the productivity of propositional attitude states, there is no alternative to postulating a language of thought, a system of syntactically structured objects in the brain that express the propositions in propositional attitudes. In later work, Fodor has stressed the systematicity of propositional attitudes mentioned above. Fodor points out that the beliefs whose contents are systematically related exhibit the following sort of empirical relation: If one is capable of believing that Mary loves John, one is also capable of believing that John loves Mary. Fodor argues that only a language of thought in the brain could explain this fact.

EXTERNALISM

The upshot of the famous "twin earth" arguments has been that meaning and content are in part in the world and in the language community. Functionalists have responded in a variety of ways. One reaction is to think of the inputs and outputs of a functional theory as long-arm as including the objects that one sees and manipulates. Another reaction is to stick with short-arm inputs and outputs that stop at the surfaces of the body, thinking of the intentional contents thereby characterized as narrow—supervening on the nonrelational physical properties of the body. There has been no widely recognized account of what narrow content is, nor is there any agreement as to whether there is any burden of proof on the advocates of narrow content to characterize it.

MEANING

Functionalism says that understanding the meaning of the word *momentum* is a functional state. On one version of the view, the functional state can be seen in terms of the role of the word *momentum* itself in thinking, problem solving, planning, and so forth. But if understanding the meaning of *momentum* is this word's having a certain function, then there is a very close relation between the meaning of a word and its function, and a natural proposal is to regard the close relation as simply identity, that is, the meaning of the word just *is* that function. Thus functionalism about content leads to functionalism about meaning, a theory that purports to tell us the metaphysical nature of meaning. This theory is popular in cognitive science, where in one version it is often known as procedural semantics, as well as in philosophy where it is often known as conceptual role semantics. The theory has been criticized (along with other versions of functionalism) by Putnam, Fodor, and E. LePore.

HOLISM

Ned Block and Fodor noted the "damn/darn" problem. Functional theories must make reference to any difference in stimuli or responses that can be mentally significant. The difference between saying "damn" and "darn" when you stub your toe can, in some circumstances, be mentally significant. So the different functionalized theories appropriate to the two responses will affect the individuation of every state connected to those utterances, and for the same reason, every state connected to those states, and so on. His pains lead to "darn," hers to "damn," so their pains are functionally different, and likewise their desires to avoid pain, their beliefs that interact with those desires, and so on. Plausible assumptions lead to the conclusion that two individuals who differ in this way share almost nothing in the way of mental states. The upshot is that the functionalist needs a way of individuating mental states that is less fine-grained than appeal to the whole theory, a molecularist characterization. Even if one is optimistic about solving this problem in the case of pain by finding something functional in common to all pains, one cannot assume that success will transfer to beliefs or meanings, for success in the case of meaning and belief may involve an analytic/synthetic distinction.

QUALIA

Recall the parity-detecting automaton described at the beginning of this entry. It could be instantiated by two people, each of whom is in charge of the function specified by a single box. Similarly, the much more complex functional organization of a human mind could "in principle" be instantiated by a vast army of people. We would have to think of the army as connected to a robot body, acting as the brain of that body, and the body would be like a person in its reactions to inputs. But would such an army really instantiate a mind? More pointedly, could such an army have pain, or the experience of red? If functionalism ascribes minds to things that do not have them, it is too liberal. W. G. Lycan suggests that we include much of human physiology in our theory to be functionalized to avoid liberalism; that is, the theory T in the definition described earlier would be a psychological theory plus a physiological theory. But that makes the opposite problem, chauvinism, worse. The resulting functional description will not apply to intelligent Martians whose physiologies are different from ours. Further, it seems easy to imagine a simple pain-feeling organism that shares little in the way of functional organization with us. The functionalized physiological theory of this organism will be hopelessly different from the corresponding theory of us. Indeed, even if one does not adopt Lycan's tactic, it is not clear how pain could be characterized functionally so as to be common to us and the simple organism.

Much of the force of the problems just mentioned derives from attention to phenomenal states like the look of red. Phenomenal properties would seem to be intrinsic to (nonrelational properties of) the states that have them, and thus phenomenal properties seem independent of the relations among states, inputs and outputs that define functional states. Consider, for example, the fact that lobotomy patients often say that they continue to have pains that feel the same as before, but that the pains do not bother them. If the concept of pain is a functional

concept, what these patients say is contradictory or incoherent—but it seems to many of us that it is intelligible.

The chauvinism/liberalism problem affects the characterization of inputs and outputs. If we characterize inputs and outputs in a way appropriate to our bodies, we chauvinistically exclude creatures whose interface with the world is very different from ours—for example, creatures whose limbs end in wheels or, turning to a bigger difference, gaseous creatures who can manipulate and sense gases but for whom all solids and liquids are alike. The obvious alternative of characterizing inputs and outputs themselves functionally would appear to yield an abstract structure that might be satisfied by, for example, the economy of Bolivia under manipulation by a wealthy eccentric, and would thus fall to the opposite problem of liberalism.

It is tempting to respond to the chauvinism problem by supposing that the same functional theory that applies to a person also applies to the creatures with wheels. If they thought they had feet, they would try to act like us, and if we thought we had wheels, we would try to act like them. But notice that the functional definitions have to have some specifications of output organs in them. To be neutral among all the types of bodies that sentient beings could have would just be to adopt the liberal alternative of specifying the inputs and outputs themselves functionally.

TELEOLOGY

Many philosophers propose that we avoid liberalism by characterizing functional roles teleologically. We exclude the armies and economies mentioned because their states are not for the right things. A major problem for this point of view is the lack of an acceptable teleological account. Accounts based on evolution smack up against the swamp-grandparents problem. Suppose you find out that your grandparents were formed from particles from the swamp that came together by chance. So, as it happens, you do not have any evolutionary history to speak of. If evolutionary accounts of the teleology underpinnings of content are right, your states do not have any content. A theory with such a consequence should be rejected.

CAUSATION

Functionalism dictates that mental properties are second-order properties, properties that consist in having other properties that have certain relations to one another. But there is at least a prima facie problem about how such second-order properties could be causal and explanatory in a way appropriate to the mental. Consider, for example, provocativeness, the second-order property that consists in having some first-order property (say redness) that causes bulls to be angry. The cape's redness provokes the bull, but does the cape's provocativeness provoke the bull? The cape's provocativeness might provoke an animal protection society, but is not the bull too stupid to be provoked by it?

Functionalism continues to be a lively and fluid point of view. Positive developments in recent years include enhanced prospects for conceptual functionalism and the articulation of the teleological point of view. Critical developments include problems with causality and holism, and continuing controversy over chauvinism and liberalism.

See also Armstrong, David M.; Behaviorism; Causation: Metaphysical Issues; Causation: Philosophy of Science; Chisholm, Roderick; Cognitive Science; Computationalism; Harman, Gilbert; Language of Thought; Lewis, David; Materialism; Meaning; Metaphysics; Mind-Body Problem; Ontology; Philosophy of Mind; Physicalism; Propositional Attitudes: Issues in Philosophy of Mind and Psychology; Propositional Attitudes: Issues in Semantics; Putnam, Hilary; Qualia; Sellars, Wilfrid; Smart, John Jamieson Carswell; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ben-Yami, H. "An Argument against Functionalism." *Australasian Journal of Philosophy* 77 (1999): 320–324.

Block, N. "The Mind as the Software of the Brain." In *An Invitation to Cognitive Science*, edited by D. Osherton, et al. Cambridge, MA: MIT Press, 1995.

Braddon-Mitchell, D., and F. Jackson. *Philosophy of Mind and Cognition*. Oxford: Blackwell, 1997.

Crane, T. *Mechanical Mind: A Philosophical Introduction to Minds, Machines, and Mental Representation*. London: Routledge, 2003.

David, M. "Kim's Functionalism." *Philosophical Perspectives* 11 (1997): 133–148.

Melnyk, A. *A Physicalist Manifesto*. Cambridge, U.K.: Cambridge University Press, 2003.

Polger, T. *Natural Minds*. Cambridge, MA: MIT Press, 2004.

Shoemaker, S. "Realization and Mental Causation." In *Physicalism and Its Discontents*, edited by C. Gillett and B. Loewer. Cambridge, U.K.: Cambridge University Press, 2001.

Ned Block (1996)

Bibliography updated by Alyssa Ney (2005)

FUNCTIONALISM IN SOCIOLOGY

In sociology and social anthropology the term "functional analysis" is used not only in the mathematical sense, where a function expresses a correspondence between two variables such that for every value of the one there are one or more determinate values of the other, and the second, or dependent, variable is, in a less technical use of the term, said to be a function of the first. Sociologists, of course, like all scientists, are interested in establishing such dependencies. The term "functional analysis" in their work also has a special connotation analogous to the use of the notion of "function" in describing biological systems or such artifacts as are self-organizing systems—for example, a heat engine with a thermostat. Such a system can be considered as a unitary whole; it is differentiated into elements, and the function of the elements can be said to be the part they play in maintaining the system in a persisting state or (in the case of artifacts) in maintaining the efficiency of the system for the purpose for which it has been set up. There are, however, differences between the use of the notion in sociology and the use as applied to biological and artificial systems, and these have become more apparent as sociologists have worked with and reflected on "functional methods." The differences hinge on the questions of whether a society should be taken to be a single integrated system or whether it may be so diversified that what is "functional" for one part may not be so for others, and whether the only "end" to which an element of a social system should be shown to contribute is the maintenance of the system as a whole in its environment.

FUNCTION AND CULTURAL FACTS

Functional notions were used by the pioneers of modern social anthropology and sociology, Émile Durkheim and W. Robertson Smith. The term *functionalism,* however, was first put forward as the name of a special method and approach by Bronislaw Malinowski in the article "Anthropology" in the *Encyclopaedia Britannica* (13th ed., supp. I). The article reads as something of a manifesto, in which functionalism is said to be "the right method" in social anthropology. Functional analysis is said to be "explanation of … facts … by the part they play within the integral system of culture, by the manner in which they are related to each other within the system, and by the manner in which this system is related to the physical surroundings. … The functional view … insists therefore upon the principle that in every type of civilisation, every custom, material object, idea and belief fulfils some vital function, has some task to accomplish, represents an indispensable part within a working whole" (ibid., pp. 132–133). Thus, the function of magic is said to consist in its being "a remedy for specific maladjustments and mental conflicts, which culture creates in allowing man to transcend his biological equipment" (ibid., p. 136), and myth is said to perform an "indispensable function" in strengthening the traditions on which a cultural life depends.

These claims for the "functional method" were both vague and grandiose. Later exponents and critics of a functional method in the social sciences have been concerned to state more precisely what it does and what it does not claim to assert. (See especially Merton, 1957, and Nagel, 1956.) Malinowski's account left the notion of the "needs" to which a function was said to be related insufficiently clear; his use of the word *indispensable* left it uncertain whether the "needs" themselves were indispensable to the society in question or whether the particular cultural item held to be the means of satisfying them was indispensable in the sense of not admitting of a substitute.

Malinowski's statement of the method was also far more than a recommendation to anthropologists to look for functions; it was a dogmatic assertion that "an object … appears as 'inessential,' 'arbitrary,' 'devoid of function' only as long as we do not understand the function of that detailed feature or object" (ibid., pp. 138–139). It also implied that every cultural item was necessary to the working of the social system as a whole. Of course, if the social system is defined as the total complex of all its cultural items, this becomes tautological. Malinowski avoided this by speaking of "vital needs" that the elements in the system are held to fulfill. But the notion of "needs," interpreted biologically and psychologically, is so extremely general that it is not shown why they can be fulfilled only by particular cultural arrangements.

FUNCTION AND SOCIAL STRUCTURE

The next leading anthropologist to use and also to write about "functional methods" was A. R. Radcliffe-Brown. (See especially his *Structure and Function in Primitive Society.*) Radcliffe-Brown worked with an even more "organic" notion of a society than did Malinowski, since the latter held that practices in a society should be seen as functional for the biological and psychological needs of its *members,* while Radcliffe-Brown was interested in seeing the function of a particular social usage as the contribution it makes to the *total social life,* which is unified as

a social system. Radcliffe-Brown regarded a social system as a set of interconnected features of social life, while he defined a social structure as "an arrangement of persons in institutionally controlled and defined relations." This definition appears insufficiently abstract: "Social structure" surely should not be used to refer to persons in relationships, but to the distinguishable pattern of recurrent sets of relationships described by social roles. But Radcliffe-Brown's account helped to link the notion of function with that of structure; that is, the uses studied were not those of separate cultural items, but of persistent forms of social relations, such as those shown in marriage arrangements.

The linking of function with structure helps to strengthen the biological, organic analogy behind this way of thinking. Thus, Radcliffe-Brown spoke of a social system as though it were a unitary whole in which every part is internally related to every other, and where it is possible to speak, by analogy with a biological organism, of the structure as serving a "total life." Following this analogy, the use so served is seen as the survival of the total society as an ongoing concern. This way of looking on a society was no doubt made more plausible by the fact that the societies so studied were small-scale primitive ones, where the society might seem to be a whole of integrated parts that, it was thought, could be exhaustively enumerated. It becomes much less plausible when applied to larger, more flexible societies comprising a number of subgroups that often are hostile to one another. This may also explain why it appeared that the "function" served was the survival of the society in its traditional given form and why, therefore, functional theory has been held to support a conservative ideology.

THE CONTEXT OF A FUNCTION

That functional theory need not be conservative was shown by Thomas Merton, who defined its central orientation as "the practice of interpreting data by establishing their consequences for larger structures in which they are implicated." Merton pointed out that although the notion of function is related to that of some end or need served, this end may not be the perpetuation of the existing social system. Subgroups may have radical interests served by certain social practices that would thus be functional within the context of those interests. Hence, it is not meaningful to speak of a cultural element or institutional practice simply as "functional." It must be shown to be functional in some specific context and in some specific respect; that is, it must have designated consequences for designated properties of designated units, but these units need not be "the society as a whole." However, the notion does presuppose some complex context in which it is possible to show how certain elements have certain consequences contributing to the complex being maintained in a certain state or to the furthering of some interest to which one "function" is related. That an element has such a function relative to such a context or interest can be stated as a matter of descriptive fact, and it need have no ethical implication to the effect that the interest itself (or the function) is thereby commended. Still less need it imply that persons or groups in a society are of consequence only because of such alleged functions.

EXPLANATIONS AND DESCRIPTIONS

How far can the direction of attention to consequences be a form not only of description but also of explanation? Malinowski spoke of such functional analyses as "explanations," though he also remarked that explanation, to the scientific thinker, is nothing but the most adequate description of a complex fact (*A Scientific Theory of Culture,* New York, 1944, p. 117). Whether or not this is a satisfactory view of scientific explanation, there remains the question of whether a functional analysis is the most adequate description of a complex fact *tout court,* or rather a description of the effects of certain elements in the complex on certain other elements; that is, a partial description from the point of view of a particular interest. Merton indeed used the word *interpretation,* which is presumably weaker than "explanation," and he spoke throughout his work of "functional analyses."

A functional analysis would be an explanation only if the answer to the question "What is the effect of *x* in context *a, b, c*?" could also be seen as an answer to the question "Why does *x* occur?" or "Why does *x* have the character it has?" It could be so put if the effect of *x* is the intended effect of an intentional action (the effect of my turning the key is to unlock the door, and the reason I turn the key is to unlock the door); that is, if the explanation is explicitly teleological, so that it is said that *x* occurs in order to produce the effect *y.*

The interest of sociologists is, however, largely directed to detecting the unintended and unanticipated consequences of actions (what Merton called their "latent" as distinct from their "manifest" functions). In such cases, can an effect *y* be cited as an explanation, or partial explanation, answering the question "Why does *x* occur?" Can functional statements in contexts where conscious purpose is presumably absent be looked on as explanations? Jonathan Cohen has defined a functional explanation as one in which the fact to be explained, for

example, the beating of the heart, is a necessary condition for that which is cited as explaining it, for example, "to circulate the blood" ("Teleological Explanation," *PAS* 51 [1950–1951]: 255–292; cf. D. M. Emmet, *Function, Purpose and Powers,* pp. 48ff.)

This definition describes the form of a functional statement, but does it show that it is a different kind of explanation from a causal one? Ernest Nagel claimed that the factual content of such functional statements can be exhaustively translated into causal terms (*Logic without Metaphysics,* pp. 250–251); for example, "the beating of the heart is a necessary condition for maintaining the circulation of the blood." Similarly Kingsley Davis, in "The Myth of Functional Analysis as a Special Method in Sociology and Anthropology," maintained that such statements simply assert that certain phenomena have certain consequences. To direct attention to consequences, especially unintended but interconnected consequences, is, he held, the distinctive approach of sociologists. "Functional" analysis is therefore not a special method in sociology, but just sociological method; and the name "functionalism," as supposedly that of a special movement or school, had better be dropped.

From this it would appear that functional statements can be explanations where they can be interpreted teleologically in terms of purpose; that is, where to say an element in a system has a function is to say that it is as it is because it has been so designed with reference to a purpose for which the system has been set up. Where this reference to purpose cannot be made, functional statements would be a form of causal statement in which the interest is directed not to the cause of a phenomenon itself, but to its effects considered as causes within a wider context. However, the reference to a wider context and the need for this to be a context within which some systematic interconnections can be shown distinguish such statements from those presenting unilinear sequences of cause and effect.

SELF-REGULATING SYSTEM

Functional statements are also particularly appropriate in those systematic contexts in which "return effects" on the cause itself can be shown; that is, where some of its consequences react back on it, so that the consequences can be invoked to explain, in part at any rate, why it is as it is. Thus, Nagel held that functional statements are most appropriately used in describing self-maintaining or self-regulating systems. His formulation for such systems can be briefly summarized as follows: Let *S* be a system and *E* its environment, and let *S* be functional, self-maintaining, or directively organized with respect to a trait (property, state, process), *G*. Let *S* undergo a series of alterations terminating in *G*. Let there then be some fairly extensive class of changes either in *E* or in certain parts of *S*. Then, unless *S* contains some mechanism that produces effects compensating for these changes, *S* will cease to exhibit *G* or the tendency to acquire *G*.

The system *S* must be specified to show how its parts are causally relevant to the state *G*, and if the "function" of a part in maintaining *G* against changes is to be cited as a cause of the state of *S*, the return effects of this part on other parts of *S* must be specified. The instantaneous values of the state coordinates must be independent at any given time, although the values of one set at one time will not be independent of those of another set at another given time (that is, the values in one set will change according to previous changes in another set). Nagel held that the relations between the elements in a functional system need to be thus precisely specified, and that very few "functional analyses" in sociology satisfy these requirements.

LATENT AND MANIFEST FUNCTIONS

Nevertheless, sociologists may be said to produce analyses in which they seek to approximate this model even if they do not entirely satisfy it. This is true particularly where the data studied are shown to have consequences in some larger context, and the consequences are return effects upon the data themselves, so that there is a mutual reinforcement. For instance, Malinowski claimed that the "function" of myths was to strengthen the traditions that help to maintain a social way of life. This may not have been the original reason for the creation of the myth (whatever this may have been, it was said by Malinowski to be sociologically unimportant). But it may be the case that the fact that the myth now performs this perhaps originally unintended function strengthens people's interest in the myth and its hold upon them, and so serves to perpetuate it. Perhaps in some cases what was a "latent" function of some activity, such as recounting a myth, can thus be made the "manifest" function, the explicit purpose of the activity, without disturbing the disposition of its practitioners to go on doing it. But in some cases this may not be true. When, for instance, Malinowski said that "the function of religion is to relieve anxiety," or others (such as Radcliffe-Brown) said that the function of religious ritual is to strengthen the will to maintain the common values on which the society depends, it is at least open to question whether the adherents of a religion would be able to go on practicing it if

they came to look on these functions as the "real reason" for doing so. Thus in some cases the change of a latent into a manifest function will be self-frustrating.

Certain conclusions can be drawn. First, it is misleading to speak of *the* function of a practice, belief, or institution *tout court.* It may have *a* function in relation to a certain interest in a certain context, and this itself may be a disfunction in relation to other interests. Thus, the fact that religion can sometimes relieve anxiety might be a disadvantage in contexts where interest lay in religion as a challenge to complacency.

Further, if such statements of "function" are to do more than merely describe consequences, it should be possible to show that the alleged function also reinforces the practice of the activity. But this must not be taken to imply that this is the sole or "real" reason for the practice. Thus, it may well be that, because of the complexity of human motivation, religious practices sometimes (not necessarily or always) relieve anxieties or promote loyalty to common values; if so, this can strengthen inducements to perform them.

The fact that activities performed with one interest in view can have unanticipated consequences in satisfying other interests can add to the survival value of these activities. Thus, Max Weber's well-known view that there was a nexus between the Calvinistic ethic and the pursuit of capitalist enterprise should not be taken to imply that "the function" of Calvinistic religion was to promote moneymaking (or vice versa), still less that the pursuit of the former was a hypocritical cloak for the latter. Rather, Weber's view implies that a particular kind of moral outlook, stressing diligence, thrift, and abstinence, was appropriate to the furtherance of capitalist enterprise, so that two independent and strong human interests, the religious and the economic, reinforced each other and thus helped to establish a way of life with considerable survival value. (It is worth noticing that this particular nexus could probably become established only under social and environmental conditions where there were opportunities for the entrepreneur who could save capital. But this is not to interpret these probably necessary conditions as sufficient conditions for explaining the Calvinistic way of life.)

A functional approach in sociology can therefore be taken not as the assumption that every cultural item has a function, but as a directive to watch for "functions," particularly in the unintended consequences of a form of social action, above all for those functions that react back on the form of social action itself, so as to produce a mutually reinforcing nexus. But the analogy with biological or with self-organizing systems must not be pressed too far, since behind forms of social activity are persons or groups capable of entertaining a variety of values and interests. Functional statements in sociology, even if they are not themselves teleological, carry an indirect teleological implication in that if something is said to have a function it has one in relation to some value, interest, or purpose held by some person or group within the society (though not necessarily by the sociologist himself, who may simply be reporting the fact that some form of activity promotes this value). Where no value is stated, the presumption tends to be that what is served is the preservation of the society as an ongoing concern. That it is desirable to preserve the society (though not necessarily just in its existing form) is taken for granted by almost everyone. Thus, when something is said to have a "function" in maintaining the society, although the point is not always recognized, one ingredient in the complex notion of function is a value judgment.

See also Durkheim, Émile; Functionalism; Nagel, Ernest; Sociology of Knowledge; Weber, Max.

Bibliography

There is an extensive literature discussing functional methods in sociology. Most of it is surveyed by Robert K. Merton, "Manifest and Latent Functions," in his *Social Theory and Social Structure* (Glencoe, IL: Free Press, 1949; revised and enlarged, 1957), pp. 19–84.

Discussions of the use of functionalism in anthropology up to 1954 are considered by Raymond Firth, "Function," in *Year Book of Anthropology* (Glencoe, IL, 1956).

See especially Bronislaw Malinowski, "Anthropology," in *Encyclopaedia Britannica,* 13th ed., supplement I (Chicago, 1926); A. R. Radcliffe-Brown, *Structure and Function in Primitive Society* (London: Cohen and West, 1952), which contains some papers and addresses written some 20 years before publication; and Talcott Parsons, *Essays in Sociological Theory, Pure and Applied* (Glencoe, IL: Free Press, 1957).

Other discussions include Ernest Nagel, "Formalization of Functionalism," in his *Logic without Metaphysics* (Glencoe, IL: Free Press, 1957), pp. 247–283, and *The Structure of Science* (New York: Harcourt Brace, 1961), pp. 520–535; D. M. Emmet, *Function, Purpose and Powers* (London, 1957); Kingsley Davis, "The Myth of Functional Analysis as a Special Method in Sociology and Anthropology," in *American Sociological Review* 24 (1959): 757ff.; C. G. Hempel, "The Logic of Functional Analysis," in *Symposium on Sociological Theory,* edited by Llewellyn Gross (Evanston, IL: Row, Peterson, 1959).

Dorothy M. Emmet (1967)

FUZZY LOGIC

"Fuzzy logics" are multivalued logics intended to model human reasoning with certain types of imprecision. The field of fuzzy logic originated with a 1965 paper by Lotfi Zadeh, a professor of engineering at the University of California, Berkeley. It is significant that the inventor of fuzzy logic was neither a philosopher nor a linguist. Since 1965 research in fuzzy logic has always had an engineering and mathematical bent, while the philosophical foundations of fuzzy logic have always been under attack.

Many different formal systems have been proposed under the general name of fuzzy logic, but there is wide acceptance that the fundamental principles of fuzzy logic are

(1) $t(A \wedge B) = \min\{t(A), t(B)\}$

(2) $t(A \vee B) = \max\{t(A), t(B)\}$

(3) $t(\neg A) = 1 - t(A)$.

In these axioms A and B represent arbitrary propositions. The truth value of A, a real number between 0 and 1, is denoted $t(A)$. The first axiom above says that the truth value of $A \wedge B$ is the lesser of the truth value of A and the truth value of B. The second and third axioms concerning disjunction and negation are to be understood similarly.

At the same time that Zadeh introduced fuzzy logic, he also introduced fuzzy set theory, a variant of naive set theory (i.e., everyday set theory as opposed to a foundational set theory such as the Zermelo-Fraenkel axioms) with the basic axioms

(1) $\mu(x \in P \cap Q) = \min\{\mu(x \in P), \mu(x \in Q)\}$

(2) $\mu(x \in P \cup Q) = \max\{\mu(x \in P), \mu(x \in Q)\}$

(3) $\mu(x \in P^c) = 1 - \mu(x \in P)$.

Here $\mu\ (x \in P)$ denotes the degree to which x is a member of the set P. Since 1965 many branches of mathematics have been generalized along fuzzy set theory lines.

There are two fundamental differences between fuzzy logics and conventional logics such as classical predicate calculus or modal logics. Although these differences are technical, they are of considerable philosophical significance. First, conventional logics (except intuitionistic logics) require for every proposition that either it or its negation be true, that is, that $t(A \vee \neg A) = 1$ in fuzzy logic notation. In fuzzy logics this "law of the excluded middle" does not hold. Second, there is no consensus about a semantics for fuzzy logic that is well-defined independently of its proof theory, that is, the inferential axioms given above. In contrast, conventional logics have well-accepted semantics, for example Tarskian model theory for predicate calculus, and Kripkean possible worlds semantics for modal logics.

Fuzzy logics are claimed to be capable of representing the meanings of intrinsically imprecise natural language sentences, such as "Many Texans are rich," for which the law of excluded middle fails. There is disagreement as to whether fuzzy methods successfully represent the complexities of concepts such as "many" and "rich." What is clear is that the rules of fuzzy logic cannot be used for reasoning about frequentist or subjective types of uncertainty, whose properties are captured by standard probability theory. The central issue here is that the probability of a compound proposition such as $A \wedge B$ is not a function just of the probabilities of the propositions A and B: The probability of $A \wedge B$ also depends on the relationship between the propositions A and B, in particular on their independence or correlation.

The tolerance for ambiguity found in fuzzy logic, and specifically the rejection of the law of the excluded middle, is a revolutionary idea in mathematical logic. Some advocates of fuzzy logic claim that tolerance for ambiguity is also revolutionary philosophically, since Western philosophy, from Plato through René Descartes, has supposedly been an intrinsically dualistic tradition. According to this argument, fuzzy logic has been better received in Japan and other Asian countries than in the West because of the holistic, subtle nature of the Eastern intellectual tradition. Apart from the dualistic oversimplification of the distinction between "Western" and "Eastern" thought, this claim also ignores the continuous holistic tradition in European philosophical thought, from Zeno through Blaise Pascal to Martin Heidegger and Ludwig Wittgenstein.

There has been much artificial intelligence research on using fuzzy logic for representing real-world knowledge, and there has been some recent convergence between this work and parallel work by a distinct research community on knowledge representation using classical logics, nonmonotonic logics, and probability theory. So far this research has remained almost exclusively theoretical. In contrast, engineering work on using fuzzy logic for controlling complex machines heuristically has been highly successful in practice.

A fuzzy controller is a device, usually implemented as software for an embedded microprocessor, that continually monitors readings from sensors, and makes decisions

about actuator settings. For example, a controller for the automatic transmission of a car monitors road speed, the position of the accelerator pedal, and other factors, and decides whether to shift gears down or up, or not to shift. The knowledge possessed by a fuzzy controller is typically represented as rules such as

$$\mu(\mathit{speed},\text{MODERATE}) \wedge \mu(\mathit{pedal},\text{FULL-DOWN}) \rightarrow \mu(\mathit{shift},\text{DOWN})$$

Here *speed* and *pedal* are sensory readings, *shift* is a possible actuator setting, and MODERATE, FULL-DOWN, and DOWN are fuzzy sets. Through inference rules for the fuzzy connectives $\wedge$ and $\rightarrow$, the degree of membership of *speed* in MODERATE and of *pedal* in FULL-DOWN determines the desired degree of membership of *shift* in DOWN. Given a set of rules, a fuzzy controller continually computes the degree to which the antecedents of each rule are satisfied, and selects a conclusion that is the weighted average of the conclusion of each rule, where rules are weighted using these degrees.

Fuzzy controllers are widely used for two basic reasons. First, since the action chosen at each instant is typically the result of interpolating several rules, their behavior is smooth. Second, fuzzy controller rule sets are easy for humans to read and understand intuitively, hence easy to construct by trial and error.

See also Artificial Intelligence; Descartes, René; Heidegger, Martin; Kripke, Saul; Logic, History of; Mathematics, Foundation of; Modal Logic; Model Theory; Pascal, Blaise; Plato; Probability and Chance; Proof Theory; Quantum Mechanics; Semantics; Set Theory; Tarski, Alfred; Wittgenstein, Ludwig Josef Johann; Zeno of Elea.

Bibliography

Elkan, C. "The Paradoxical Success of Fuzzy Logic." *IEEE Expert* 6 (August 1994): 3–8.

Mamdani, E. H. "Application of Fuzzy Algorithms for Control of Simple Dynamic Plant." *Proceedings of the Institution of Electrical Engineers* 121 (1974): 1585–1588.

Zadeh, L. A. "Fuzzy Sets." *Information and Control* 8 (1965): 338–353.

Zimmermann, H.-J. *Fuzzy Set Theory—And Its Applications.* Boston: Kluwer Academic, 1991.

Charles Elkan (1996)

2nd edition

ENCYCLOPEDIA OF

PHILOSOPHY

GADAMER – JUST WAR THEORY

4 *volume*

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

G

GADAMER, HANS-GEORG

(1900–2002)

Hans-Georg Gadamer, a Heidelberg philosopher and student of Martin Heidegger, is best known for his hermeneutic philosophy put forward in his *Wahrheit und Methode* (*Truth and Method,* 1960). Widely regarded as the most significant German philosopher after Heidegger, Gadamer wrote on Plato, Georg Wilhelm Friedrich Hegel, Heidegger, Aristotle's practical philosophy, reason in an age of science, aesthetics, poetics, Paul Celan, and other topics.

BIOGRAPHY

Gadamer was born in Marburg and grew up in Breslau. His mother died when he was four. His father was a well-known university research scientist in pharmacological chemistry. In 1919 Gadamer's father was called from the University of Breslau to a research chair at the University of Marburg. Gadamer entered Marburg as a second-year student with interests in literature, art history, and classical philology. But he was soon drawn to the great neo-Kantian philosopher and Platonist, Paul Natorp, under whom he completed his doctoral dissertation in 1922 on pleasure in the Platonic dialogues. In 1923 Gadamer journeyed for the summer semester to Freiburg to hear Heidegger, who was offering bold new interpretations of Aristotle and other philosophers. When Heidegger moved to Marburg in the fall of that year, Gadamer became his assistant and he remained so until 1928. During this time Gadamer also studied with Nicolai Hartmann, took seminars in classical philology under Paul Friedländer and others, and in 1927 was certified in classical philology. In 1928 he completed his habilitation under Heidegger on "Plato's dialectical ethics," based on the *Philebus.*

Gadamer remained another ten years in Marburg waiting for a call to a full-time teaching appointment. After 1933 his chances for a call were practically blotted out by his not being in good standing with the Nazis. But he remained active in the academic life at Marburg, which boasted some of Germany's leading intellectuals—Rudolf Bultmann in theology; Hartmann; Stefan George, the charismatic poet; Richard Hamann, the iconoclastic art historian; and finally, Friedländer and others, who represented the great philological tradition of Ulrich von Wilamowitz-Moellendorff.

In 1938 Gadamer was finally called to a chair in philosophy at Leipzig, where he was able to survive through the war years as a politically unthreatening classical

humanist. Because of his political integrity he was elected rector at Leipzig after the war. In 1947 he managed to escape the stultifying atmosphere of the new communist regime by being called to a position at Frankfurt University. He was at Frankfurt but two years when in 1949 he was called to fill Karl Jaspers's chair at the University of Heidelberg.

Gadamer remained in Heidelberg as chair in philosophy until his retirement in 1968. A gifted lecturer, he concentrated in the 1950s on topics that later became part of *Truth and Method.* At the same time, he worked to revive Hegel studies in Germany, and rebuilt a war-shattered department into one of the strongest in Germany. In 1952, along with Helmut Kuhn, he founded the *Philosophische Rundschau,* a journal dedicated to reviewing current books and discussing major issues in philosophy.

After 1968 Gadamer continued to lecture and offer seminars in Heidelberg as an honored emeritus professor, but now he allowed himself to accept invitations to speak in other countries and to serve as a guest professor at various universities, especially in the United States and Canada. This fed a growing interest in hermeneutics in the United States, an interest manifested in the number of dissertations and books being written on the subject. English translations of Gadamer's works began to appear: *Truth and Method* (1975), *Philosophical Hermeneutics* (1976), and *Hegel's Dialectic* (1976) being among the first.

WORKS

In *Truth and Method* Gadamer's concepts can be logically divided into those within *Truth and Method* and those in the shorter writings after it. The latter category includes further writings defending and defining hermeneutics, writings in modern and ancient philosophy, and in aesthetics and poetics.

In *Truth and Method* Gadamer articulated the most detailed and nuanced account of the "event of understanding" in the history of philosophy. He based much of his thinking on Heidegger, Hegel, and Plato. From Heidegger's *Origin of a Work of Art* he drew strength for a powerful reassertion of the "truth" of art, and from Heidegger's *Being and Time* and later writings he drew concepts that called into question the goal of objectivity in interpretation. From Hegel and Plato he drew emphases on tradition, history, and dialogue. From Wilhelm Dilthey and Heidegger he drew an emphasis on the horizonal character of consciousness and the operativeness of history in all understanding. Understanding, he argues, takes place in a consciousness in which history—that is, tradition—is always already at work, shaping, predisposing, predefining what the process of understanding involves. His term for this is *wirkungsgeschichtliches Bewußtsein,* "effective historical consciousness," and the encounter with the other, as person or as text, is a matter of *Horizontverschmelzung.*

In *Truth and Method* Gadamer shows the development after Kant of fateful conceptual turns in the course of eighteenth- and nineteenth-century philosophy, philology, and hermeneutics that have led to present presuppositions about understanding and the conditions for its possibility. He traces the dream of scientific objectivity in humanistic and social scientific knowledge in the nineteenth century, especially in Friedrich Daniel Ernst Schleiermacher and Dilthey, and the promising philosophical transformation of this "problematic" of understanding through Heidegger's phenomenological analysis of existential temporality and the historical situatedness of and the participation of history in understanding. He accepts Heidegger's description of the "forestructure" of understanding, adding to it his concept of an "anticipation of completeness" in all understanding. He argues that the process of understanding has the structure of a dialogue and can be likened to a game in that it follows rules and operates in a language that transcends it; thus, he emphasizes the "linguisticality" (*Sprachlichkeit*) of understanding and even ultimately its ontological character: "Being that can be understood is language," he asserted (*Truth and Method,* p. 432). Finally, one of the most distinctive and important of the contributions of *Truth and Method* is its insistence on a moment of "application" in all understanding.

The book's overarching goal, however, was to cause the artwork to be seen in a new way. While the title might lead one to expect it to be concerned with methods in the *Geisteswissenschaften,* Gadamer's professed aim is to defend the claim of artworks to be "true." In Gadamer's view the experience of encountering truth in great works of art demonstrates the limits of a science-oriented concept of understanding; the meaning and power of such artworks elude scientific modes of understanding. Gadamer wrote a good deal in explanation and defense of *Truth and Method.* These writings are now collected in volume two of his collected works.

Gadamer's writings on modern philosophy range through the Continental tradition since Kant and are influenced principally by Plato, who casts a shadow even over his modern writings; by Heidegger, about whom he wrote more than about any other modern philosopher; by Hegel, whose importance in modern philosophy

Gadamer repeatedly defended; and by Edmund Husserl whose phenomenology Gadamer used and treated as a major element in his thought. Most of his essays on ancient philosophy are directly or indirectly connected with Plato. From Plato he draws his model of dialogue, in which partners participate in quest of a truth that transcends the individual seeker. Gadamer's ethical thinking as well as his dialectical hermeneutics go back to Plato's "dialectical ethics" of respect for the other person, of openness, of seeking to strengthen the partner's case in order not merely to win a debate to one's own satisfaction but to move together toward truth, a result that benefits both sides and that both sides affirm.

Art and poetry were a major theme in Gadamer's writings throughout his career. In 1934 Gadamer wrote on "Plato and the Poets," and in the 1940s he was writing essays on Johann Christian Friedrich Hölderlin, Johann Wolfgang von Goethe, Karl Leberecht Immerman, and Rainer Maria Rilke. His articles after *Truth and Method* tend to select more sober and difficult poets such as Stefan George, Gottfried Benn, and Paul Celan. His essays on aesthetics and poetics continue to emphasize the truth of art, the need for dialogical openness, and the priority of the artwork's character of play. At the same time, another issue arises: What about basically nonrepresentational poetry? What about the "no longer beautiful" poetry of the modern (or postmodern) dark lyric? After a number of writings that struggle successfully with the dark lyric, such as *Wer bin ich und wer bist du?* (Who am I and who are you?; 1973), on poet Paul Celan, Gadamer poses the problem in somewhat different terms. For Gadamer it is a "task of philosophy" to develop a context within which one can still recognize and deal with—or "understand"—modern and postmodern art.

Gadamer's essay "The Relevance of the Beautiful" presents a twentieth-century defense of such art. In this essay experience becomes the reference point, even group experiences as one finds them in the historical record. Gadamer includes, not just experiences recorded in artworks or great poetry, which would create a circular argument, but in anthropological records of such things as (1) the role of play in human life, (2) the high experiences of festiveness in our own and other cultures, and (3) the power of participation in symbolic religious rites. In groping for an explanation of the power of art and a defense of its legitimacy, Gadamer offers an analysis of three categories—play, symbol, and festival.

In his essay "The Truth of the Artwork" (1960), Gadamer pointed to a threefold insufficiency of scientific thinking: (1) the insufficiency of scientific thinking, by itself and without recourse to standards outside itself, to grapple with ethical problems such as human rights, abortion, ecology, or planning the future; (2) its incapacity to account for the experience of beauty in art and poetry or to lay down principles for its creation; and (3) its insufficiency to meet, or even account for, the spiritual needs of human beings. All these suggest that a recourse to the absolute priority of scientific presuppositions cannot serve us well in dealing with the encounter with ethical problems, artworks, or the divine. Art, like ethics and the divine, seems to move beyond the competence of the categories of scientific thinking. And they can claim to be "true." This is a major theme both in *Truth and Method* and in later writings.

In "Wort und Bild" (Word and image; 1992) Gadamer takes the final step and attempts to articulate aesthetic categories that apply both to plastic/pictorial arts and arts of the word. Among the several concepts to which he turns are the Greek concept of the fine (*kalon*) and to our experience of the rightness and absoluteness of art.

See also Aristotle; Benn, Gottfried; Bultmann, Rudolf; Dilthey, Wilhelm; Goethe, Johann Wolfgang von; Hartmann, Nicolai; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Hermeneutics; Hölderlin, Johann Christian Friedrich; Husserl, Edmund; Jaspers, Karl; Kant, Immanuel; Natorp, Paul; Plato; Rilke, Rainer Maria (René); Schleiermacher, Friedrich Daniel Ernst.

Bibliography

MAJOR WORKS BY GADAMER

Platos dialektische Ethik: Phänomenologische Interpretationen zum "Philebos." 3rd ed. Leipzig: Meiner, 1931. Translated by R. M. Wallace as *Plato's Dialectical Ethics: Phenomenological Interpretations Relating to the "Philebus."* New Haven, CT: Yale University Press, 1991.

Wahrheit und Methode: Grundzüge einer philosophischen Hermeneutik. Tübingen: Mohr, 1960. Translated as *Truth and Method,* edited by G. Barden and J. Cumming. New York: Seabury Press, 1975; rev. ed., 1989.

Hegels Dialektik: Fünf hermeneutische Studien. Tübingen: Mohr, 1971. Translated by P. C. Smith as *Hegel's Dialectic: Five Hermeneutical Studies.* New Haven, CT: Yale University Press, 1976.

Wer bin ich und wer bist du?: Ein Kommentar zu Paul Celans Gedichtfolge "Atemkristall." Frankfurt: Suhrkamp, 1973.

Vernunft im Zeitalter der Wissenschaft: Aufsätze. Frankfurt: Suhrkamp, 1976. Translated as *Reason in the Age of Science.* Cambridge, MA: MIT Press, 1981.

Die Aktualität des Schönen: Kunst als Spiel, Symbol, und Fest. Stuttgart: Reclam, 1977. Translated by N. Walker as "The Relevance of the Beautiful." In *The Relevance of the Beautiful*

and Other Essays, edited by R. Bernasconi. Cambridge, U.K.: Cambridge University Press, 1986.

Die Idee des Guten zwischen Plato und Aristoteles. Heidelberg: Winter, 1978. Translated as *The Idea of the Good in Platonic-Aristotelian Philosophy.* New Haven, CT: Yale University Press, 1986.

Heideggers Wege. Tübingen: Mohr, 1983. Translated as *Heidegger's Ways.* Albany: State University of New York, 1994.

Gessamelte Werke. 10 vols. Tübingen: Mohr, 1985–. Contains all of the above works.

COLLECTIONS OF GADAMER'S WORKS IN ENGLISH

Philosophical Hermeneutics. Edited and translated by D. E. Linge. Berkeley: University of California Press, 1976.

Dialogue and Dialectic: Eight Hermeneutical Studies on Plato. Translated by P. C. Smith. New Haven, CT: Yale University Press, 1980.

The Relevance of the Beautiful and Other Essays. Edited by R. Bernasconi. Cambridge, U.K.: Cambridge University Press, 1986.

Dialogue and Deconstruction: The Gadamer-Derrida Encounter. Edited by D. Michelfelder and R. E. Palmer. Albany: State University of New York Press, 1989. Includes five essays by Gadamer.

Hans-Georg Gadamer on Education, Poetry, and History. Edited by Dieter Misgeld and G. Nicholson. Albany: State University of New York Press, 1992.

Literature and Philosophy in Dialogue: Essays in German Literary Theory. Translated by R. H. Paslick. Albany: State University of New York Press, 1992.

WORKS ON GADAMER

Hoy, D. C. *The Critical Circle: Literature, History, and Philosophical Hermeneutics.* Berkeley: University of California Press, 1978.

Schmidt, L. K. *The Epistemology of Hans-Georg Gadamer.* New York: P. Lang, 1985.

Warnke, G. *Gadamer: Hermeneutics, Tradition, and Reason.* Stanford, CA: Stanford University Press, 1987.

Weinsheimer, J. C. *Gadamer's Hermeneutics: A Reading of "Truth and Method."* New Haven, CT: Yale University Press, 1985.

Wright, K., ed. *Festivals of Interpretation: Essays on Hans-Georg Gadamer's Work.* Albany: State University of New York Press, 1990.

Richard E. Palmer (1996)

GALEN

(129—c. 216 CE)

Galen (Aelius or Iulius Galenus of Pergamum), a doctor and philosopher, was the son of a rich architect. Born in modern-day Bergama in western Turkey, he was introduced as a student to all the main philosophical theories of classical antiquity. On his own admission, this led him only into a confusion from which he was rescued by considering mathematics, which henceforth provided him with a paradigm for understanding truth and falsehood. From 145, following the appearance of the healing god Asclepius to his father in a dream, he turned to medicine. He sat at the feet of medical teachers in Pergamum, Smyrna, and Alexandria, as part of what is the longest recorded medical education from the ancient world. In 157 he returned to Pergamum as doctor to the gladiators of the high priest, but in 162 he traveled to Rome, the imperial capital. There he quickly established a reputation as a doctor, anatomist, and philosopher, not always to the delight of his many competitors. In 166 he left Rome hurriedly but was recalled in 168 by Emperor Marcus Aurelius to join him and his brother on campaign in northern Italy. After his return to Rome in 169, he seems to have spent the rest of his life in Italy as a physician to the emperor's household although he made at least one visit to Pergamum. The traditional date for his death, c. 200, is based on an early misunderstanding of a comment, preserved by Arabic authors, that divided his life into seventeen years as a student and seventy as a doctor. A date of death around 216 fits better with the internal evidence from his many treatises and would allow him to continue writing major treatises on medicine and pharmacology well into the first decade of the third century, or even later.

Galen was an enormously prolific author, credited with more than 350 treatises on subjects ranging from attic comedy to vivisection, and from logic to pharmacology. Roughly half of these survive today, in whole or in part, mainly in his native Greek but also in Arabic, Hebrew, Persian, and medieval Latin versions. These translations are of great importance, particularly when the originals have been lost, for they frequently deal with philosophical issues that seem to have held little interest for the Byzantine. New discoveries of previously unknown treatises can be expected as major libraries in the Muslim world are opened to scholars and more works of medieval Arabic and Jewish philosophy are published. The recent recovery of new fragments in Arabic of the lost *On Scientific Discovery* suggests that a complete copy of the work Galen thought his greatest contribution to philosophy may eventually be found.

Galen's interest in philosophy can be followed throughout his life, from his very early *On Medical Experience* to his last work, *On My Own Opinions*. He regarded philosophy as essential to the proper practice of medicine: The best doctor was also a philosopher, whether or not he realized it. Conversely, a knowledge of medicine was valuable for philosophers, a conjunction Galen

traced back to Plato whose notions of the body in the *Timaeus* Galen derived from a (unhistorical) friendship with Hippocrates. In turn, Galen visualized Hippocrates as a Platonic philosopher, a claim that contributed to the growing dominance of Hippocrates as the symbol of the medical profession.

Fundamental to medicine was logic, both for structuring accurate diagnosis and for distinguishing between various degrees of certainty. Some suppositions could be proved to be true, others shown to be false, others were merely plausible and could be adopted only provisionally. Still others, such as the nature of god or the eternity of world, were incapable of proof or refutation and were best left to idle sophists. But Galen often muddled these important distinctions, either by treating the merely plausible as if it were true or by choosing as *obvious* and *universally agreed* bases for discussion facts or ideas that themselves were disputed by some of his opponents.

Galen's formal logic, on which he wrote several books, is impressive in its rigor and clarity. The Arabs' attribution to him of the discovery of the fourth syllogism may be right, or it may simply reflect Galen's extension of earlier debates about argument. Throughout his writings he stressed the importance of accuracy and clarity of expression, to avoid confusion, and to allow discussion with those offering different points of view. Ambiguity, on which he wrote an extant tract, was harmful to medicine as well as philosophy, and a sound training in logic he considered necessary for everyone. His demands for a mathematical precision in debate are not, however, fully borne out by his own practice, and his overwhelming powers of rhetoric often obscure his unscrupulous representation of the illogicality of his opponents.

His philosophy and his medicine reinforced each other. Where empirical observation was not enough by itself, logic and understanding of the theories of other philosophers could bridge the gap. Conversely, the facts of medical life exemplified and justified the cosmological and psychological doctrines of philosophers. His discussions of the value of empiricism in relation to understanding the causes of disease, one of the goals of the true doctor, show an awareness of the epistemological difficulties involved and an understanding that an experienced practitioner might reach the correct conclusion without having to go through the necessary chain of causation. He might also be a swifter and safer option than a callow youth, no matter how brilliant the youth's reasoning abilities. Galen's entire approach was eclectic, rejecting the dogmatism of the philosophical schools of his own day in favor of the "twin legs" of reason and experience.

Galen's medicine was based on an Aristotelian physics combined with a Platonic psychology. His universe, made up of the four Aristotelian elements in various combinations or mixtures, had been overseen by a purposeful Creator, or Nature, and worked along the interconnected principles favored by the Aristotelians and Stoics. His explanations for the working of drugs, for instance, involved Aristotelian language and concepts. He was convinced that each part of the body had been designed teleologically, for a particular purpose, and any alteration or imbalance in its basic elements, qualities, or humors resulted in illness. Galen's defense of teleology, as evinced in the human hand and in the elephant's trunk, is arguably superior to that of Aristotle's, and his exposition of what he termed the "natural faculties" is far from the simplistic presentation familiar from later denunciations of Galenism.

Whereas he believed strongly in the existence of the soul, he refused to be drawn to any definitive statement about what the essence of the soul was. The Aristotelian and Stoic notions of an undivided controlling power within the body he vigorously rejected as being inconsistent with the facts of anatomy. His systematic dissections of a variety of animals convinced him that there were three almost independent systems within the body corresponding to the three parts of Plato's soul, as described in the *Republic* and *Timaeus*: the brain and nerves, concerned with thought and sensation; the heart and arteries, responsible for life and energy; and the veins and liver (a more precise rendering of Plato's *belly*), responsible for nutrition and growth. Galen never proclaimed a strict parallelism between the three systems, which was achieved only by later followers such as Avicenna and Averroes, and he devoted much more space to the first two than to the third. This lack of systematization was the result both of his enormous fecundity of ideas and his methods of composition, for most of his books were originally oral presentations, taken down by trained shorthand writers, and not carefully crafted treatises written at leisure. Not surprisingly, they are often repetitive and leave many knots untied.

The heart, for Galen, was the source of natural heat, and the place where a small amount of venous blood, mixed with air, was transformed into vigorous arterial blood. His repetition of the earlier experiments of Erasistratus (c. 304–250 BCE) proved convincingly that the arteries contained blood and not *pneuma* alone, as Erasistratus had argued. But his vitalist predilections con-

vinced him that the movement of arterial blood was not the result of any quasi-mechanical motion of the heart but brought about by the forcible contraction of the arteries controlled by natural powers within their thick coats. Just as most venous blood remained within the veins until it was absorbed as nutriment or excreted, so most arterial blood remained within the arteries. A tiny portion was transformed in the rete mirabile, a vascular plexus at the base of the skull (not found in humans but in some animals Galen dissected), to become psychic pneuma, which was refined still more in the networks of the brain to act as the means of transmission of sensation and the commands of the brain. Contrary to Aristotle and the Stoics, he could find no evidence for the heart as the seat of sensation and thought, especially since he could trace its nerves back to an origin in the innermost cavities of the brain. Galen's experimental dissections of the spinal cord in animals are among the most impressive ever performed, combining a precision of dissection with a careful planning and elucidation of what was to be achieved, and were not superseded until the mid-sixteenth century.

Galen's anatomical conclusions he believed far too important to be left entirely to doctors. In two of his longest treatises, *On the Usefulness of Parts* and *On the Opinions of Hippocrates and Plato*, he explained the consequences of his discoveries in Aristotelian and Platonic terms, respectively. Similarly, his comments on the *Timaeus* stressed the truth of many of Plato's observations and suggested that he must have gained his anatomical knowledge from the great Hippocrates himself. Plato was his favorite philosopher: Galen's writings are permeated with Platonic phraseology, and he wrote summaries of the *Timaeus* and other Platonic dialogues that are partially extant in Arabic.

A man of austere morality—Galen claimed to have read the *Golden Words* of Pythagoras nightly—he wrote extensively on ethics. He advocated a self-control brought about by an extensive philosophical training although he acknowledged that this might be doubly difficult for those who had been badly trained or whose psychic genetic makeup predisposed them to evil. He advocated a very strong interaction between body and soul, for just as overeating and drinking or pleasurable and painful sensations have an obvious effect on behavior, so, in turn, anger or grief can lead to physical illness and even death. Galen contrasts his own equanimity at the loss of most of his library in a fire with his mother's shocking irascibility and with the timorousness of a patient who worried himself to death after dreaming that he had replaced Atlas as the upholder of the world. Doctor and philosopher should cooperate in the search for health and wholeness.

Anatomy also helped to resolve some philosophical disputes about intentionality. Some actions, Galen showed, were under the direct control of the brain via the nerves and muscles; others were "natural," the result of our genetic makeup, and beyond rational control; others were more complex, such as speech, which required both the will and the modification of "natural" patterns of breathing. Others, such as winking and blinking, appeared to indicate the coexistence of voluntary and involuntary activity in the same organ whereas others, such as a penile erection, he explained by other notions such as sympathy. Only laughter defeated his attempts at explanation. Throughout, Galen sought to use his medical experience to illuminate contemporary philosophical debate, just as he used philosophical debates on physics or causation to explain his decisions as a practitioner.

Contemporary reactions to his philosophy were mixed. A sect of Christians attempted to recast their Christianity to take account of his logical objections to miracles, but others were less polite. Skeptics objected to his reliance on sensory data, and the Aristotelian Alexander of Aphrodisias thought him a great doctor but a poor philosopher, whose profession in *On My Own Opinions* of agnosticism about many philosophical questions was a confession of failure. But his views on creation and some of his Platonic commentaries were cited with respect in the fifth century, and the Christian philosopher Nemesius of Emesa built his Christian anthropology largely on Galen's discoveries. Although much of his philosophy had disappeared in Greek by 1000, the Arabs drew heavily on his work. New fragments of his ethics recovered from Spanish Jewish writers of the twelfth and thirteenth centuries show how much they valued his approach to morality. Renaissance biographers were equally impressed, some even viewing his life as exemplifying the cardinal virtues—a perspective hardly shared by modern scholars. Others hotly debated whether he had become a Christian at the end of his life or not. But after the sixteenth century, Galenic philosophy, like his medicine, was abandoned, not to be studied again in considerable detail until the 1970s.

See also Alexander of Aphrodisias; Aristotle; Ethics; Hippocrates and the Hippocratic Corpus; Logic, History of; Nemesius of Emesa; Philosophy of Medicine; Philosophy of Science, History of; Plato; Pythagoras and Pythagoreanism; Stoicism.

Bibliography

TEXTS

DeLacy, Phillip H., ed. *On the Opinions of Hippocrates and Plato/Galen*. Berlin, Akademie-Verlag. 3 vols. 1978–1984.

Kühn, Karl Gottlob. *Claudii Galeni Opera Omnia*, Leipzig: Knobloch. 20 vols. in 22. 1821–1832. The standard edition of the Greek Galen.

Kollesch, Jutta, and Diethard Nickel, "Bibliographia Galeniana: Die Beiträge des 20. Jahrhunderts zur Galenforschung." In *Aufstieg und Niedergang der römischen Welt, Teil II: Principat*. Band 37, 2, edited by Wolfgang Haase and Hildegard Temporini. Berlin: De Gruyter, 1994, p. 1351–1420, 2063–2070.

May, Margaret T., trans. *Galen on the Usefulness of Parts of the Body*. 2 vols. Ithaca, Cornell University Press, 1968.

Muhaqqiq, Mahdi, ed. *Kitab al-Shukuk 'ala Jalinus l-Muhammad ibn Zakariya al-Razi, Tehran, Mu'assasah*. 1993. New fragments of *On Scientific Discovery*.

Nutton, Vivian, ed. *Galen, On My Own Opinions*. Berlin: Akademie Verlag, 1999. A list of texts not included in this edition can be found in Vivan Nutton, ed., "The Unknown Galen," *Bulletin of the Institute of Classical Studies*, Supp. 77, 2002.

A good selection of philosophical oriented texts is available in English in Peter N. Singer, *Galen. Selected Works*, Oxford: Oxford University Press, 1997.

STUDIES

Barnes, Jonathan. "Galen on Logic and Therapy." In *Galen's Method of Healing: Proceedings of the 1982 Galen Symposium*, edited by Fridolf Kudlien and Richard J. Durling. Leiden, NY: Brill, 1991, p. 50–102.

Frede, Michael. *Essays in Ancient Philosophy*. Oxford: Clarendon Press, 1987.

Hankinson, R. James. "Galen and the Best of All Possible Worlds." *Classical Quarterly* 39 (1989): 206–227.

Hankinson, R. James. *Cause and Explanation in Ancient Greek Thought*. Oxford: Clarendon Press, 1998.

Hankinson, R. James, ed. *The Cambridge Companion to Galen*. Cambridge, U.K.: Cambridge University Press, forthcoming.

Lloyd, Geoffrey E. R. "Theories and Practices of Demonstration in Galen." In *Rationality in Greek Thought*, edited by Michael Frede and Gisela Striker. Oxford: Clarendon Press, 1996, p. 255–277.

Moraux, Paul. *Der Aristotelismus bei den Griechen*. Vol. 2. Berlin: De Gruyter, 1984.

Nutton, Vivian. *Ancient Medicine*. London: Routledge, 2004, p. 216–247.

Rescher, Nicholas. *Galen and the Syllogism: An Examination of the Thesis that Galen Originated the Fourth Figure of the Syllogism in the Light of New Data from Arabic Sources, Including an Arabic Text Edition and Annotated Translation of Ibn al-Sal ah's treatise on the Fourth Figure of the Categorical Syllogism*. Pittsburgh, PA: University of Pittsburgh Press.

Smith, Wesley D. *The Hippocratic Tradition*. Ithaca, NY: Cornell University Press, 1979.

Temkin, Owsei. *Galenism, Rise and Decline of a Medical Philosophy*. Ithaca, NY: Cornell University Press, 1973.

von Staden, Heinrich. "Body, Soul, and Nerves: Epicurus, Herophilus, Erasistratus, the Stoics, and Galen." In *Psyche and Soma: Physicians and Metaphysicians on the Mind-Body Problem from Antiquity to Enlightenment*, edited by John P. Wright and Paul Potter. Oxford: Clarendon Press, 2000.

Zonta, Mauro. *Un Interprete ebreo della Filosofia di Galeno: gli scritti filosofici di Galeno nell'opera di Shem Tob ibn Falaquera*. Turin: Zamorani, 1995.

Vivian Nutton (2005)

GALEN [ADDENDUM]

Galen's influence on medieval Islamic thought in general, and on Arabic philosophy in particular, can hardly be overestimated. Galen himself developed a system of philosophical and medical views that gained tremendous authority in Late Antiquity, notably in Alexandria and the Hellenized East. He emphasized the necessity that physicians be conversant in philosophy, and this idea was thoroughly enshrined, for instance, in the Late Antique medical curriculum in Alexandria. To give just one example, physiological and nosological processes were explained in terms of Aristotelian categories and the four causes. This medical tradition, aptly called "Galenism," shaped the Islamic notion of sciences and medicine to a large extent; it is therefore not surprising that many of the most famous Arabic philosophers such as al-Kindī, al-Rāzī (Rhasis), Ibn Sīnā (Avicenna) and Ibn Rushd (Averroes) were also prominent physicians.

Virtually all the works of Galen's medical and philosophical writings were translated into Arabic, and it is in this language that some of the most interesting philosophical works such as *On Medical Experience* survive. The idea of experience was hotly debated among medical authors in the medieval Islamic period, and treatises such as al-Rāzī's *Doubts concerning Galen* show that Arabic authors engaged critically with him. However, many of Galen's ideas, such as concepts about human physiology, which had already entered Late Antique Greek popular intellectual culture, became commonplace in the Islamic world.

See also al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotelianism; Averroes; Avicenna; Experience; Islamic Philosophy.

Bibliography

Galen. *Three Treatises on the Nature of Science*. Translated by Richard Walzer and Michael Frede. Indianapolis: Hackett, 1985.

Gutas, Dimitri. *Greek Thought, Arabic Culture: The Graeco-Arabic Translation Movement in Baghdad and Early ʿAbbāsid Society*. London: Routledge, 1998.

Ullmann, Manfred. *Islamic Medicine*. Edinburgh, U.K.: Edinburgh University Press, 1978.

Temkin, Oswei. *Galenism: Rise and Decline of a Medical Philosophy*. Ithaca, NY: Cornell University Press, 1973.

Peter E. Pormann (2005)

GALILEO GALILEI

(1564–1642)

Galileo Galilei, the Italian astronomer and physicist, was born at Pisa. Although he created no systematic philosophy, his influence on the trend of modern philosophical thought is very marked. To it may be traced the definitive separation of physical science from philosophy, the abandonment of authority as a criterion of scientific truth, the distinction between objective and subjective qualities in observable phenomena, and the introduction (or reintroduction) of empirical and skeptical elements into philosophical investigations. The seventeenth-century revival of atomism and the removal of occult qualities from the concept of causation owed much to Galileo. His writings marked the beginning of an antimetaphysical movement in philosophy, exemplified in later times by Positivism and operationalism, and they remained relatively free from such concealed ontological assumptions as are to be found in some ostensibly nonmetaphysical systems; for example, in philosophical empiricism, mechanism, and phenomenalism. The events of Galileo's personal life involved him in an active struggle for freedom of thought, and this in turn underlay those scientific and philosophical convictions for which he became a symbol to his contemporaries and followers.

After a meager, conventional preparatory education, Galileo was enrolled in the school of medicine at the University of Pisa in 1581. His interests turned to mathematics about 1583, and he left the university in 1585 without having received a degree. For a time he continued his studies alone, giving lessons at Siena and Florence, and in 1589 secured the chair of mathematics at Pisa. His early investigations in physics, particularly mechanics, set him in sharp opposition to the views prevailing among professors of philosophy, who, as followers of Aristotle, looked with disfavor upon the introduction of mathematics into physics. In 1591 Galileo left Pisa to become professor of mathematics at the University of Padua. Here he continued his mechanical researches, undertook the manufacture of mathematical instruments for sale, and composed several useful treatises that were circulated in manuscript among his pupils and friends.

In 1604 he disclosed, in a letter to Fra Paolo Sarpi, the mathematical law of freely falling bodies. He had made substantial progress toward a rational mechanics when, in 1609, his attention was diverted from those studies by the invention of the telescope in Holland. Galileo promptly duplicated and improved the device, producing an instrument suitable for astronomical observation, and early in 1610 he published *Sidereus Nuncius* (The starry messenger), in which he described the mountainous character of the moon, the existence of countless previously unobserved stars, and the discovery of four satellites of Jupiter. These discoveries provoked widespread controversy among philosophers and astronomers. Shortly after *Sidereus Nuncius* was published, Galileo resigned his post and returned to Florence as chief mathematician and philosopher to Cosimo II de' Medici, grand duke of Tuscany. Late in 1610, he detected the curious appearance of Saturn (although he could not account for it) and the phases of Venus, and shortly thereafter he noted the phenomenon of sunspots.

Strong philosophical opposition was aroused by his publication in 1612 of *Discourse on Bodies in Water,* in which he ridiculed the Aristotelian theory of the elements, advocated observation and experiment as the chief criteria of physical truth, and gave some support to the views of Democritus. Publication in 1613 of his *Istoria e dimostrazioni intorno alle macchie solari* (Letters on sunspots) aroused theological opposition by its open support of the Copernican theory, which appeared to contradict the Bible by asserting the motion of Earth and the stability of the sun. Toward the end of that year certain philosophers and priests, in the presence of members of the ruling family, attacked Galileo and Copernicanism on religious grounds. Galileo replied in a long letter addressed to his former pupil Benedetto Castelli, who had defended him in his absence. In this letter Galileo contended that biblical passages had no authority in scientific controversies, and that the language of the Bible should instead be interpreted in the light of man's knowledge of natural phenomena, gained through reason and observation.

Early in 1615, Galileo was violently attacked from the pulpit in Florence, and a copy of his earlier letter, together with a denunciation of the Galileists, was sent to the Inquisition at Rome. Rumors spread that the Catholic Church would soon officially condemn the Copernican theory and silence Galileo. He countered by greatly expanding the arguments of his previous letter to Castelli,

and widely circulated the new version in manuscript copies ostensibly addressed to Grand Duchess Christina, the mother of Cosimo de' Medici. At the end of the same year Galileo went personally to Rome in an attempt to stem the anti-Copernican tide. In this he was unsuccessful, for the church officially moved to condemn the views of the motion of Earth and stability of the sun as contrary to the Scriptures. Galileo was instructed not to hold or defend these views, but no action was taken against his person or his previously published books.

After a period of silence, Galileo entered into a polemic with a Jesuit professor at Rome over the nature of comets, and in 1623 he published *The Assayer* (*Il saggiatore*), ridiculing the Aristotelian philosophy and expounding his methodological ideas. This book was dedicated to the new pope, Urban VIII, who as cardinal had been very friendly toward Galileo. In 1624, Galileo visited Rome and obtained permission to write on the Copernican and Ptolemaic systems, provided that the treatment was impartial. The composition of his next book, *Dialogue concerning the Two Chief World Systems,* occupied Galileo intermittently for several years. After great difficulty in getting a license to publish it, he brought it out at Florence in 1632. Five months later the printer was ordered to issue no more copies, and Galileo was summoned to Rome to face the Inquisition. Despite remonstrances of the Tuscan ambassador and of Grand Duke Ferdinand II, he was eventually compelled to appear and stand trial. In June 1633 the book was condemned, and Galileo was sentenced to life imprisonment. He was, however, permitted to reside first at Siena with the archbishop, his friend and former pupil, and then under house arrest at his own villa near Florence. There he managed to compose and smuggle out his most mature work, a treatise on physics known as *Two New Sciences,* which was printed in Holland in 1638. Galileo died four years later.

PHILOSOPHICAL ROOTS

The conditions of Galileo's education and career led to his intimate familiarity with the works of Aristotle. There is little evidence, however, that he ever preferred to read the works of any particular philosopher. His personal library was scanty, and his correspondence is devoid of philosophical references or discussions. In polemic works he refers often to Aristotle, usually with disfavor. His occasional references to Democritus, Socrates, Plato, and Seneca are more favorable but superficial, and appear to express general anti-Aristotelianism rather than allegiance to any other philosopher. Near contemporaries such as Girolamo Cardano and Bernardino Telesio were dismissed by Galileo with the remark that he had read but little of their work. Although there is extant a Latin treatise in Galileo's handwriting that contains references to scores of ancient and medieval philosophers, this dates from his student days (1584) and is not demonstrably original or representative of his own views.

It is evident that during most of his life Galileo's favorite reading was literary rather than philosophical. He is said to have known Ludovico Ariosto by heart; he lectured on Dante Alighieri, annotated Torquato Tasso, delighted in the rustic dialect writings of Ruzzante (Angelo Beolco), and frequently cited Homer, Vergil, Pulci, and other poets in his works. Nowhere in his writings is there an overt expression of allegiance to any of his philosophical predecessors. However, the question of Galileo's true metaphysical position has been much debated in recent decades. His emphasis on the mathematical element in physics has induced many excellent scholars, led by Alexandre Koyré, to classify him unequivocally as a Platonist. On the other hand, Galileo's insistence on the power of observation to refute any reasoned conclusion has caused others, notably Ludovico Geymonat, to resist this conclusion and even to stress a strong Aristotelian element in Galileo's own work. In opposition to both these views, Edward Strong has questioned the propriety of reading any metaphysical position into Galileo's work and emphasizes his evident preoccupation with methodological considerations, to the exclusion of dogmatic philosophy. Finally, Alistair Crombie has aptly remarked that it is precisely the absence of systematic philosophy in Galileo that has made it possible for adherents of nearly every philosophical school to find some support for their views in his works.

Galileo's anti-Aristotelianism makes its first appearance in his early studies of motion. In order to defend a theory of motion (later abandoned) founded on Archimedean conceptions, he was obliged to demolish several prevailing Aristotelian assertions: that the speed of free fall is proportional to the weight of the falling body and inversely proportional to the density of the medium; that the motion of projectiles depends on some action of the medium; and that motion is impossible in the void. Galileo's attack on Aristotle widened with his adoption of the Copernican astronomy and his abandonment of the distinction between elemental and celestial matter, so essential to Aristotle's world view. In the end he questioned the reliability of Aristotelian logic and asserted that rigorous demonstration was to be found only in mathematics. Thus, it may be argued that Aristotle's physical errors led

Galileo to distrust logical deduction as a basis for physics, and his famous dictum that "the book of nature is written in mathematical characters, without a knowledge of which men cannot understand it" probably represents a methodological canon rather than a metaphysical position.

Galileo's opposition to Aristotle was also to some degree a literary pose rather than a true philosophical position. In later years he often declared himself a better Aristotelian than his contemporary Peripatetic adversaries, and in the opening sections of the *Dialogue* he made extensive use of Aristotelian arguments to secure assent to essential points in the Copernican theory. Perhaps the chief significance of Galileo's anti-Aristotelianism is its intimate relation to his consistent rejection of authority of any kind in matters of science. His unwillingness to accept any intermediary between himself and nature was the motivation of his bold warning to the church against the utilization of scriptural authority in scientific disputes. In short, Aristotle was not so much the philosophical opponent as the historical symbol in Galileo's unremitting battle against authority as a criterion of truth.

SEPARATION OF PHYSICS FROM PHILOSOPHY. Until Galileo's time, physical science (including theoretical astronomy) was regarded as a proper part of philosophy and was so taught in the universities. Aristotle's principles of motion supplied the axioms, and the science was purely deductive. Several of Galileo's predecessors had questioned those principles as being in apparent contradiction with experience; Galileo continued these attacks and undertook experimental investigations of the actual phenomena of motion. In this way he came upon some new results and sought principles from which both the old and new phenomena might be deductively established. If he was not entirely successful in this quest, that was not a matter of deep concern to him. René Descartes later criticized Galileo sharply for his investigation of physical effects without a prior knowledge of their causes, and Cartesian physics was made an integral part of Descartes's systematic philosophy. In the end, however, the example of Galileo, and not that of Descartes, was followed by scientists. Physics became first a distinct branch (the "natural philosophy" of Isaac Newton) and ultimately a separate discipline from philosophy. The philosophical effects of this separation have been enormous. The emphasis on physics that prevailed in philosophy at Galileo's time has vanished; in its place, the theory of knowledge has risen to preeminence in modern philosophy, where from time to time it has threatened to subordinate or even to expel metaphysics. It is very doubtful that this would have come about without the separation of physical researches from philosophical investigations, in which separation Galileo was the pioneer.

Nor was Galileo content merely to remove terrestrial physical phenomena from the realm of speculative philosophy. Telescopic observation suggested to him a direct analogy between terrestrial and celestial matter, a conception that was antagonistic to Aristotle's entire scheme. The mountainous character of the moon's surface refuted the axiom of perfect sphericity of celestial bodies; the appearance and disappearance of sunspots destroyed the axiom of celestial immutability and perfection. Galileo did not hesitate to attribute terrestrial qualities to all celestial bodies, thus laying the basis for physical astronomy, even though this had grave religious implications and challenged the traditional cosmological and cosmogonical assumptions of the Peripatetics, who dominated the philosophical thought of the time.

ABANDONMENT OF AUTHORITY

The age into which Galileo was born was one in which the power of authority was uppermost in every sphere of activity—political, religious, and philosophical. It was therefore virtually impossible to attack that power in one sphere without disturbing it in others. To Galileo it was clear that in matters of scientific investigation, authority as such could not be allowed any weight; observation, experiment, and reason alone could establish physical truth. Accordingly, he disputed the right of philosophers and theologians to exercise control over scientific investigations or even scientific theories. Confronted with almost overwhelming opposition in this dispute, he was compelled to adopt, if not to formulate systematically, an alternative criterion of truth that might have some hope of acceptance. In so doing he avoided the error of such philosophers as Bernardino Telesio and Francis Bacon, who in effect had called for reliance solely upon sensory evidence, for Galileo was well aware of the possibility of illusion or of misinterpretation. It appeared to Galileo that mathematics alone had the kind of certainty that could be completely trusted, and he took the position that only to the extent to which men can detect mathematical regularities in phenomena can they be certain they have reached the truth in physical matters: "Without mathematics, one wanders about in a dark labyrinth."

This dictum of Galileo's is often taken today as the expression of a fundamentally Platonic viewpoint, but there is no evidence that Galileo believed in a world of Platonic forms as distinguished from that of events. He appears to have been concerned with relationships rather

than essences, and it is in this sense that his mathematical conception of the world is to be taken. He expressly stated that the failure of a physicist to describe the real world was not the fault either of that world or of mathematics, but was merely a result of the limited competence of the physicist, analogous to the shortcomings of a merchant or an accountant who had failed to take into account the weights of the containers in computing the value of his merchandise. Galileo held that although we must be satisfied with limited objectives, we may achieve complete certainty with regard to them.

Galileo's battle for free inquiry, independent of the interference of authority, was, in his own time, doomed to defeat. Nevertheless, the practicability of his alternative criterion of truth gained him a substantial number of followers whose intelligent application of his suggestions was eventually beneficial not only to physical science, but to philosophy as well. The systematic search for solutions to physical problems within preestablished limits of inquiry, under the rule that the demands of sense data, reason, and mathematical interpretation must be simultaneously satisfied, led to the development of a body of dependable knowledge of the physical world that philosophers could no longer ignore in their speculations about underlying reality. At the same time, this complex criterion of physical truth gave rise to a more serious examination of epistemological problems that had been relatively neglected in previous ages of untrammeled speculative philosophy. "We must deal with the real world, and not one on paper," Galileo proclaimed in his *Dialogue.*

It should be noted that Galileo did not extend his demands beyond the domain of physical science. In order to reconcile his scientific position with his acceptance of religious authority, he distinguished sharply between the two uses of language, or even the two languages, of faith and science. This position was expounded at length in his *Letter to Christina* (1615, published 1636). Wherever natural phenomena are involved, the language of the Bible is to be interpreted by the findings of science, while the exposition of supernatural texts is to be left to theologians. This concept of duality of language was given an interesting extension in his *Two New Sciences*; the purely mathematical sections are written in Latin, while the physical and more general sections are in Italian. It is reasonable to assume that Galileo's attitude toward the difficult question of why mathematical relations prevail in physics, and indeed toward metaphysical questions in general, was similar to his expressed opinion with regard to supernatural assertions of any kind: that they were outside the domain of science.

OBJECTIVE AND SUBJECTIVE QUALITIES

It is well known that Galileo clearly set forth the distinction, later made into a fundamental principle by the philosophical empiricists, between primary and secondary qualities. In accordance with his conception that only mathematics afforded complete certainty, he believed that to the fundamental particles of matter one must attribute size, shape, number, and rate of motion; but that one is under no compulsion to invest them with color, sound, odor, and the like. This separation of subjective qualities from those capable of mathematical treatment was a decisive step in the removal of man from his traditionally central place in the entire scheme of things. It is also noteworthy that Galileo showed no interest whatever in pursuing an analysis of subjective (or secondary) qualities; he did not (as did Aristotle and Descartes) compose books on the mind, the spirit, or man in general. Thus, Galileo's treatment of this fundamental principle of empiricism is by no means indicative that he subscribed to an empiricist philosophy in the technical sense. His distinction of objective-subjective was simply an integral part of his separation of physics from philosophy; and if it had any metaphysical implications, they were lost to Galileo. The source of inspiration for this fundamental notion was certainly Greek atomism; but if any classifiable system of philosophy is to be found in Galileo's own writings, it is that of naive realism—a recurrent theme in physical science and the philosophy of science from Galileo's time through the nineteenth century.

Highly important philosophically in Galileo's physical investigations was his insistence on the doctrine of relativity of motion. On the one hand, this removed Earth from its privileged position in astronomy, and in this regard was little more than an extension of the Copernican revolution. On the other hand, it began a new revolutionary movement in which the investigation of natural laws as mathematical relationships was to replace traditional inquiries into the natures or essences of physical entities. Thus, the way was opened to a coherent science of dynamics, while the accepted world view based on the doctrine of "natural places" was rendered untenable.

SKEPTICISM AND THE SCIENTIFIC METHOD

It is worth noting briefly that Galileo introduced (or rather reintroduced) into Western philosophy certain traditions of skepticism which had lapsed during the reign of authority. He often said that it was easier for him to determine that something was false than to discover the

truth of any matter, and he openly contended that it was less shameful to be ignorant than to argue on the wrong side. He was accustomed to advise his pupils freely and candidly to confess "I do not know" rather than to offer merely verbal explanations. Thus, among his followers it became once more respectable for a philosopher to acknowledge ignorance. The effects of this were very noticeable in the activities of his disciples and in the Galilean school they created. Establishment of the first great experimental scientific society, the Academy of the Cimento, with its motto of "Test and test again," would scarcely have been possible before Galileo. It was his habit to stress the infinite amount that must remain unknown, no matter how deeply one might penetrate into any subject of inquiry.

METHODOLOGICAL VIEWS. Galileo's philosophical importance is nowhere clearer than in his contributions to the method of scientific investigation. His suggestions and his example in this field laid the basis for procedures in physical science that have enriched epistemology as much as science itself. Yet Galileo's conceptions of scientific method, like his other philosophic views, make their appearance in his books only implicitly and incidentally to other considerations. They are first apparent in his polemic work on floating bodies (1612), in which experiments are designed for and applied to the refutation of verbal explanations and arguments. This book marked a definite epoch in the philosophy of science, inasmuch as it is perhaps the first systematic exposition of physical experiments specifically designed to refute a philosophical position. In classical form, it presents the Peripatetic theory and Galileo's countertheory, with a set of experimental tests to show the falsity of the former and the truth of the latter. Although methodological considerations are not discussed in the abstract, the work is a model of the "experimental philosophy" carried on by the school of Galileo.

The work on sunspots published in the following year is also rich in methodological material; here Galileo destroyed the arguments of his anonymous Jesuit adversary by establishing the analogy of terrestrial phenomena to solar phenomena and by applying mathematical reasoning to the problem of the location of the spots. In this instance the use of experiment was precluded by inaccessibility of the phenomena, but observational data were correctly applied in its stead. It thus constituted a methodological continuation of the book on floating bodies, although the points at issue were in this case strictly scientific and not philosophical, at least in modern terms. Particularly noteworthy are certain semantic critiques directed against arguments based upon purely verbal deductions made by Galileo's adversary in this controversy.

Because of the prohibition against discussion of Copernicanism, ten years elapsed before the publication of Galileo's next acknowledged work, *The Assayer,* which differed markedly from his earlier works. Ostensibly a polemic over the nature of comets, it was in reality a detailed critique of the then prevailing treatment of astronomical phenomena. Instead of adopting a specific theory of comets, Galileo undertook to refute his opponent by showing that all his arguments depended upon assumptions that could not be demonstrated or upon confusions of a linguistic character. It is probable that he was motivated at least in part by a desire to place in the hands of his readers a method by which they might themselves arrive at conclusions that he had been forbidden to advocate. Among the principal themes of this work are the proper and improper use of observation and experiment, the distinction between primary and secondary qualities, the necessity for clarification of language in dealing with physical concepts, and the infinite scope of natural phenomena.

In his ill-fated *Dialogue* of 1632, Galileo developed the last-named theme at length. It was his view that physical truths are boundless in number but perfectly consistent; that human knowledge at any time can comprise but a finite part of this infinite whole; that sense experience, indispensable though it is to a knowledge of the world, can be deceptive or misleading in any given instance. Thus, the concept of physical science as essentially a process of successive approximations is already implicit in the teachings of Galileo. At each stage of inquiry, sense experience must be combined with reasoning and with mathematics to afford a sound basis of deduction. Galileo noted that the method used in proof is rarely the same as that used in making a discovery, and held that unless the proof is mathematical, it lacks absolute certainty.

The *Two New Sciences* of 1638, Galileo's chief contribution to physics, was of less direct philosophical importance. Its indirect importance lies in the fact that it definitively established physics as a distinct discipline on the basis of its own methods of investigation, methods that have persisted virtually unchanged. Of particular importance to later philosophical developments were the introduction in this work of the concept of one-to-one correspondence in the analysis of the arithmetical infinite and Galileo's suggestions relating to the roles of physical and mathematical indivisibles in the explanation of observable phenomena.

In conclusion, it should be observed that a profound difference existed between the methods by which Galileo and his contemporary, Johannes Kepler, applied mathematical reasoning to physical science, particularly to astronomy. Kepler's thought was pervaded by the conviction that numerical relationships determined the structure of the universe in the sense of Pythagorean mysticism. Accordingly he attempted, repeatedly, to deduce that structure from a priori numerical hypotheses. As a result, he was early led into fantastic speculations from which he was sometimes able to extricate himself only after years of labor. Galileo, on the other hand, regarded mathematics as an indispensable practical tool and as the definitive test in the quest for physical certainty; but he was not inclined to follow wherever mathematical deduction might lead. His errors, unlike Kepler's, are usually to be found in attempts to create mathematical proofs for physical laws of which he had previously made certain, in the desire to achieve the unique degree of certainty that he ascribed to mathematical demonstration.

See also Aristotle; Atomism; Copernicus, Nicolas; Dante Alighieri; Descartes, René; Empiricism; Epistemology; Kepler, Johannes; Leucippus and Democritus; Newton, Isaac; Operationalism; Peripatetics; Phenomenalism; Plato; Positivism; Scientific Method; Seneca, Lucius Annaeus; Skepticism, History of; Socrates; Telesio, Bernardino.

Bibliography

WORKS BY GALILEO

Edizione nazionale delle opere di Galileo Galilei. Edited by A. Favoro. Florence, 1890–1909; reprinted 1929–1939.

Discorso intorno alle cose che stanno in su l'acqua. Florence, 1612. Translated by T. Salusbury as *Discourse on Bodies in Water.* Urbana: University of Illinois Press, 1960.

Istoria e dimostrazioni intorno alle macchie solari (Letters on sunspots). Rome, 1613. Translated by S. Drake in *Discoveries and Opinions of Galileo.* Garden City: NY: Doubleday, 1957. Introduction and notes by S. Drake.

Il saggiatore. Rome, 1623. Translated by S. Drake in *Controversy on the Comets of 1618,* edited by C. D. O'Malley and S. Drake. Philadelphia: University of Pennsylvania Press, 1960.

Dialogo sopra i due massimi sistemi del mondo. Florence, 1632. Translated by S. Drake as *Dialogue concerning the Two Chief World Systems.* Berkeley: University of California Press, 1953.

Nov-antiqua … Doctrina … de Sacrae Scripturae Testimoniis. Strasburg, 1636. Translated by S. Drake as "Letter to the Grand Duchess Christina" in *Discoveries and Opinions of Galileo.* Garden City: NY: Doubleday, 1957.

Discorsi e dimostrazioni matematiche intorno a due nuove scienze. Leiden: Appresso gli Elsevirii, 1638. Translated by H. Crew and A. De Salvio as *Dialogues Concerning Two New Sciences.* New York: Macmillan, 1914.

WORKS ON GALILEO

Aliotta, A., and C. Carbonara. *Galilei.* Milan: Fratelli Bocca, 1949.

Banfi, A. *Galileo Galilei.* Milan: Casa Editrice Ambrosiana, 1949.

Burtt, E. A. *Metaphysical Foundations of Modern Physical Science.* London: K. Paul, Trench, Trubner, 1925.

Cassirer, Ernst. *Substance and Function.* Chicago: Open Court, 1923.

Crombie, Alistair. *Galilée devant les critiques de la posterité.* Paris: Librarie du Palais de la Découverte, 1956.

Fazio-Allmayer, V. *Galileo Galilei.* Milan, 1911.

Geymonat, Ludovico. *Galileo Galilei.* Milan, 1957. Translated, with notes, by S. Drake; foreword by G. de Santillana. New York: McGraw-Hill, 1965.

Giua, M. *Storia delle scienze ed epistemologia.* Turin, 1945.

Goldbeck, E. *Der Mensch und sein Weltbild.* Leipzig, 1925.

Grimaldi, V. *La mente di Galileo Galilei.* Naples, 1901.

Höffding, H. *History of Modern Philosophy.* London: Macmillan, 1900.

Koyré, Alexandre. *Études galiléennes.* Paris: Hermann, 1939.

Lasswitz, K. *Geschichte der Atomistik.* Hamburg: L. Voss, 1890. Vol. II.

Natorp, P. "Galilei als Philosoph." In *Philosophische Monatshefte.* Heidelberg, 1882. Vol. XVIII.

Olgiati, F., S. Rovighi, M. Campo, et al. *Nel terzo centenario della morte di Galileo.* Milan, 1942.

Olschki, L. *Galilei und seine Zeit.* Halle, 1927.

Romao, M. *Galileu e o método scientífico.* Lisbon, 1944. Vol. I.

Rossi, G. *Galileo Galilei ed il suo metodo.* Bologna, 1880.

Strong, Edward. *Procedures and Metaphysics.* Berkeley: University of California Press, 1936.

Wohlwill, E. *Galilei und sein Kampf für die Kopernische Lehre.* Hamburg: L. Voss, 1909. Vol. II.

S. Drake (1967)

GALLUPPI, PASQUALE

(1770–1846)

Pasquale Galluppi, the Italian epistemologist and moral philosopher, was born in Tropea, Calabria. He began the study of law in Naples but soon switched to theology and philosophy. At first Galluppi was strongly influenced by Christian Wolff. In 1800 he began to read Étienne Bonnot de Condillac and John Locke, and his first published work, *Sull'analisi e sulla sintesi* (On analysis and synthesis; Naples, 1807), was an attack on sensationalism. From 1807 until 1815 Galluppi studied Immanuel Kant. Although he was strongly attracted by Kantianism, he finally rejected it as "skepticism," and, through an exami-

nation of René Descartes and Locke, he arrived at a position strongly resembling that of the Scottish commonsense school as it had been interpreted by the French eclectics.

The publication in 1819 of the first volume of his *Saggio filosofico sulla critica della conoscenza* (Philosophical essay on the critique of knowledge; 6 vols., Naples, 1819–1823) brought Galluppi widespread recognition. Between 1820 and 1827 he published his best-known works: the *Elementi di filosofia* (4 vols., Messina, 1820–1827), in which he expounded his theories, and the *Lettere filosofiche sulle vicende della filosofia relativamente ai principî delle conoscenze umane da Cartesio sino a Kant inclusivamente* (Philosophical letters on the events in philosophy concerning the principles of human knowledge from Descartes to Kant inclusive; Naples, 1838), a remarkable history of human thought. In October 1831 Galluppi was named professor of philosophy at the University of Naples. He corresponded with Victor Cousin, whose *Fragments philosophiques* he translated into Italian (2 vols., Naples, 1831–1832), and in 1838 he was named foreign correspondent of the Academie des Sciences Morales et Politiques.

Galluppi held that the only method of philosophy is analysis, a regressive movement in which reflective thought goes back over its own development. The starting point is consciousness: The existence of the conscious ego is "an original experimental truth" and an immediate intuition. The conscious ego consists in the immediate apprehension which the existing ego has of itself. This apprehension simultaneously produces apprehension of the object (which is sensation) and apprehension of the subject that perceives the object (which is feeling). Galluppi expressed this originating act in the formula "I feel (*sento*) a me which senses (*sente*) something" outside of me. Consciousness, in other words, is the awareness that the ego has of itself and of a separate, independently existing reality. On the basis of this indisputable testimony of consciousness Galluppi proclaimed the reality of both the ego and things, in opposition to George Berkeley's idealism and David Hume's analyses.

Using the same procedure, and by means of the evidence provided by internal consciousness, Galluppi found in the ego the universal ideas which had been denied by the empiricists: these ideas are proved by inner experience, which affirms the existence of God and, by revealing that the conscious ego can only be the effect of a divine intelligent cause, invariable and absolute, also attests the validity of causal relations. True knowledge, knowledge that is adequate to reality, consists in rearranging, by a real synthesis, the objective unities of beings just as they are. The existence of God, proved in the same way that Descartes did, by means of consciousness ("I think, therefore I am; therefore God exists"), proves that the self-evident relationships are valid. (This last argument differs from the Cartesian principle of divine truth.)

Galluppi adhered to the same theory in moral philosophy. In moral philosophy also it is the testimony of consciousness that tells us we are free and that makes us feel the necessity of moral good and thus the presence of a natural moral law: Our duty is affirmed to us by our innermost sense.

See also Analysis, Philosophical; Condillac, Étienne Bonnot de; Consciousness; Cousin, Victor; Descartes, René; Ethics, History of; Hume, David; Kant, Immanuel; Locke, John; Skepticism, History of; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY GALLUPPI

Lezioni di logica e metafisica. 4 vols. Naples, 1832–1834.

Filosofia della volontà. 4 vols. Naples, 1832–1840.

Considerazioni filosofiche sull'idealismo trascendentale e sul razionalismo assoluto. Naples, 1841. A paper presented at the Institut de France in 1841.

Lettere filosofiche. Edited by Augusto Guzzo, 2nd ed. Florence, 1925. Contains an ample bibliography.

WORKS ON GALLUPPI

Di Napoli, Giovanni. *La filosofia di Pasquale Galluppi.* Padua: CEDAM, 1947. Contains a complete bibliography, including an exhaustive list of manuscripts and editions.

Gentile, Giovanni. *Dal Genovesi al Galluppi.* Naples, 1903.

Guzzo, Augusto. *Idealisti e empiricisti.* Florence, 1937.

Lastrucci, Vincenzo. *Pasquale Galluppi, studio critico.* Florence: Barbera, 1890.

Spaventa, Bertrando. *Prolusione e introduzione alle lezioni di filosofia nell'Università di Napoli.* Naples, 1862; 2nd ed. edited by Giovanni Gentile as *La filosofia italiana nelle sue relazioni con la filosofia europea.* Bari, 1908; 3rd ed. Bari, 1926. Pp. 133–145.

Eugenio Garin (1967)
Translated by Robert M. Connolly

GAME THEORY

GAMES OF COMPLETE INFORMATION

A game is an abstract, formal description of a strategic interaction. Any strategic interaction involves two or more decision makers (players), each with two or more

ways of acting (strategies), such that the outcome depends on the strategy choices of all the players. Each player has well-defined preferences among all the possible outcomes, enabling corresponding utilities (payoffs) to be assigned. A game makes explicit the rules governing players' interaction, the players' feasible strategies, and their preferences over outcomes. Game theory describes games by means of mathematical concepts (e.g., sets, functions, and relations).

NORMAL FORM. A possible representation of a game is in normal form. A normal form game is completely defined by three elements that constitute the structure of the game: a list of players $i = 1, \dots, n$; for each player i; a finite set of pure strategies S_i; and a payoff function u_i that gives player i's payoff $u_i(s)$ for each n-tuple of strategies $(s_1, \dots, s_n)$, where u_i: $\times_{j=1}^{n} S_j \rightarrow R$. A player may choose to play a pure strategy or instead to randomize over his or her pure strategies; a probability distribution over pure strategies is called a mixed strategy and is denoted by σ_i. Each player's randomization is assumed to be statistically independent of that of his or her opponents, and the payoffs to a mixed strategy are the expected values of the corresponding pure strategy payoffs. A different interpretation of mixed strategies, based on the idea that players do not always randomize over their feasible actions, is that the probability distribution σ_i represents other players' uncertainty about what player will do. A mixed strategy is thus thought of as other players' conjecture about a player's plans of action. The conjectures depend on the player's private information, which is left unspecified in the model. A problem with this interpretation is that if there are reasons behind the choices a player makes, they should be included in the model, since they are likely to be payoff relevant.

The two-by-two matrix in Figure 1 depicts the two-player normal form representation of the famous Prisoner's dilemma game, where *C* stands for *cooperate* and *D* for *defect*. The numbers in the cell of the matrix denote players' payoffs: the first number is the payoff for the row player, the second for the column player. Each player picks a strategy independently, and the outcome, represented in terms of players' payoffs, is the joint product of these two strategies. Notice that in the game of Figure 1, each player is better off defecting no matter what the other player does. For example, if the column player cooperates, the row player gets a payoff of 3 by defecting and a payoff of 2 by cooperating, while if the column player defects, the row player gains a payoff of 1 by defecting and of 0 by cooperating. When, regardless of what other players do, a strategy yields a player a (strictly) inferior payoff than some other strategy, it is called a *dominated strategy*. When a strategy yields the same payoff of another undominated strategy, but it has an inferior payoff against at least one opponent's strategy, it is called a *weakly dominated strategy*.

FIGURE 1

Prisoner's Dilemma	C	D
C	2,2	0,3
D	3,0	1,1

The game of Figure 1 is one of complete information, in that the players are assumed to know the rules of the game (which include players' strategies) and other players' payoffs. If players are allowed to enter into binding agreements before the game is played, one can say that the game is cooperative. Noncooperative games instead make no allowance for the existence of an enforcement mechanism that would make the terms of the agreement binding on the players. What strategies should rational players choose? What could be rightly called the central dogma of game theory states that rational players will always jointly maximize their expected utilities, or play a Nash equilibrium (compare Nash 1996). Informally, a Nash equilibrium specifies players' actions and beliefs such that (1) each player's action is optimal given his or her beliefs about other players' choices; (2) players' beliefs are correct. Thus, an outcome that is not a Nash equilibrium requires either that a player chooses a suboptimal strategy or that some players misperceive the situation.

More formally, a Nash equilibrium is a vector of strategies $(\sigma^*_1, \dots, \sigma^*_n)$, one for each of the n players in the game, such that each σ^*_i is optimal given (or is a best reply to) σ^*_{-i}. That is

$$u_i(\sigma^*_i, \sigma^*_{-i}) \geq u_i(\sigma_i, \sigma^*_{-i}) \text{ for all mixed strategies of player } i \; \sigma_i$$

Note that optimality is only conditional on a fixed? σ_{-i}, not on all possible σ_{-i}. A strategy that is a best reply to a

FIGURE 2

Matching Pennies	H	T
H	1,–1	–1,1
T	–1,1	1,–1

FIGURE 3

	c	d
a	2,4	0,1
b	0,1	0,1

given combination of the opponents' strategies may fare poorly vis-à-vis another strategy combination.

In a game like the one depicted in Figure 2 the row player gains a payoff of 1 if the toss of two coins results in two heads or two tails and loses 1 otherwise, and vice versa for the column player.

This game has no Nash equilibrium in pure strategies. Nash proved that—provided certain restrictions are imposed on strategy sets and payoff functions—every game has at least one equilibrium in mixed strategies. In a mixed strategy equilibrium, the equilibrium strategy of each player makes the other indifferent between the strategies on which he or she is randomizing. In particular, the game in Figure 2 has a unique Nash equilibrium in which both players randomize between their strategies with probability ½. Then, if the first player plays $\sigma_1 = (½ H, ½ T)$, his or her expected payoff is $½ \; 1 + ½ - 1 = 0$ regardless of the strategy of the second player.

The players (and the game theorist) can predict that a specific equilibrium will be played just in case they have enough information to infer players' choices. The standard assumptions in game theory are:

CK1. The structure of the game is common knowledge

CK2. The players are rational (i.e., they are expected utility maximizers) and this is common knowledge

The concept of common knowledge was introduced by David K. Lewis (1969) in his study on convention, which is arguably the first major philosophical work in which game theory plays a central role as a modeling tool. Simply put, the idea of common knowledge is that a certain proposition *p* is common knowledge among two players if both of them know *p*, both of them know that they know *p*, and so on ad infinitum. The previous assumptions may allow the players to predict an opponent's strategy. For example, in the prisoner's dilemma game of Figure 1 rational players would never choose the strictly dominated strategy *C*. CK1 and CK2, then, allow the players to predict that the opponent will play *D*. However (compare Bicchieri 1993), the previous CK assumptions do not always guarantee that a prediction of play can be made. For one, even if the game has a unique equilibrium, the set of strategies that, under the assumptions CK1 and CK2, players may choose need not contain the equilibrium strategies only. Moreover, predictability is hampered by another common problem encountered in game theory: multiple Nash equilibria.

Suppose two players have to divide $100 among them. They must restrict their proposals to integers, and each has to independently propose a way to split the sum. If the total proposed by both is equal or less than $100, each gets what he or she proposed, otherwise they get nothing. This game has 101 Nash equilibria. Is there a way to predict which one will be chosen? In real life, many people would go for the fifty-fifty split. It is simple and it seems equitable. In Thomas C. Schelling's (1960) words, it is a focal point. Unfortunately, mere salience is not enough to provide a player with a reason for choice. In this example, only if it is common knowledge that the fifty-fifty split is the salient outcome does it become rational to propose $50. Game theory, however, filters out any social or cultural information regarding strategies, leaving players with the task of coordinating their actions on the sole basis of common knowledge of rationality (and of the structure of the game).

A different approach to the problem of indeterminacy is to start by considering the set of Nash equilibria and ask whether some of them should be eliminated because they are in some sense unreasonable. This is the approach taken by the refinement program (Kohlberg 1990, van Damme 1987). Consider the game in Figure 3:

The game has two Nash equilibria in pure strategies: (*a*,*c*) and (*b*,*d*). The equilibrium (*a*,*c*) is *Pareto dominant*,

since it gives both players a higher payoff than any other equilibrium in the game. However, common knowledge of rationality and of the structure of the game does not force the column player to expect the row player to eliminate the weakly dominated strategy *b*, nor is the row player forced to conclude that the column player will discard *d*. Prudence, however, may suggest that one should never be too sure of the opponents' choices. Even if the players have agreed to play a given equilibrium, some uncertainty remains. If so, one should try to model this uncertainty in the game. R. Selten's (1965) insight was to treat perfect rationality as a limit case. His "trembling hand" metaphor presupposes that deciding and acting are two separate processes, in that even if one decides to take a particular action, one may end up doing something else by mistake. An equilibrium strategy should be optimal not only against the opponents' strategies but also against some small probability $\varepsilon > 0$ that the opponents make mistakes. Such an equilibrium is *trembling-hand perfect.*

Is the equilibrium (*b,d*) perfect? If so, *b* must be optimal against *c* being played with probability ε and *d* being played with probability $1 - \varepsilon$ for some small $\varepsilon > 0$. But in this case the expected payoff to *a* is 2ϵ whereas the payoff to *b* is?. Hence for all $\varepsilon > 0$, *a* is a better strategy choice. The equilibrium (*b,d*) is not perfect, but (*a,c*) is. Therefore, a prudent player would discard (*b,d*). In this simple game, checking perfection is easy, since only one mistake is possible. With many strategies, there usually are many more possible mistakes to take into account. Similarly, with many players one may need to worry about who is more likely to make a mistake.

EXTENSIVE FORM. A different representation of a game is the *extensive form*. It specifies the following information: a finite set of players $i = 1, \dots, n$; the order of moves; the players' choices at each move; and what each player knows when he or she has to choose. The order of play is represented by a game tree *T*, which is a finite set of partially ordered nodes $t \in T$ satisfying a precedence relation $<$. A *subgame* is a collection of branches of a game such that they start from the same node and the branches and the node together form a game tree by itself. A tree representation is sequential, because it shows the order in which actions are taken by the players. It is natural to think of sequential-move games as being ones in which players choose their strategies one after the other, and of simultaneous-move games as ones in which players choose their strategies at the same time. What is important, however, is not the temporal order of events per se, but whether players know about other players' actions when they have to choose their own. In the normal form representation, players' information about other players' choices is not represented. This is the reason a normal form game could represent any one of several extensive form games. When the order of play is irrelevant to a game's outcome, then restricting oneself to the normal form is justifiable. When the order of play is relevant, however, the extensive form must be specified.

In an extensive form game the information a player has when he or she is choosing an action is explicitly represented using information sets, which partition the nodes of the tree. If an information set contains more than one node, the player who has to make a choice at that information set will be uncertain as to which node he or she is at. Not knowing at which node one is means that the player does not know which action was chosen by the preceding player. If a game contains information sets that are not singletons, the game is one of *imperfect information*.

A strategy for player *i* is a complete plan of action that specifies an action at every node at which it is *i*'s turn to move. Note that a strategy specifies actions even at nodes that will never be reached if that strategy is played. Consider the game in Figure 4. It is a finite game of perfect information in which player 1 moves first. If he chooses *D* at his first node, the game ends and player 1 nets a payoff of 1, whereas player 2 gets 0. But choosing *D* at the first node is only part of a strategy for player 1. For example, it can be part of a strategy that recommends "play *D* at your first node, and *x* at your last node." Another strategy may instead recommend playing *D* at his first node, and *y* at his last decision node. Though it may seem surprising that a strategy specifies actions even at nodes that will not be reached if that strategy is played, one must remember that a strategy is a full contingent plan of action. For example, the strategy *Dx* recommends playing *D* at the first node, thus effectively ending the game. It is important, however, to be able to have a plan of action in case *D* is not played. Player 1 may, after all, make a mistake and, because of player 2's response, find himself called to play at his last node. In that case, having a plan helps. Note that a strategy cannot be changed during the course of the game. Though a player may conjecture about several scenarios of moves and countermoves before playing the game, at the end of deliberation a strategy must be chosen and followed throughout the game.

The game of Figure 4 has two Nash equilibria in pure strategies:(*Dx,d*) and (*Dy,d*). Is there a way to solve the indeterminacy?

Suppose player 1 were to reach his last node. Since he is by assumption rational, he will choose *x*, which guar-

FIGURE 4

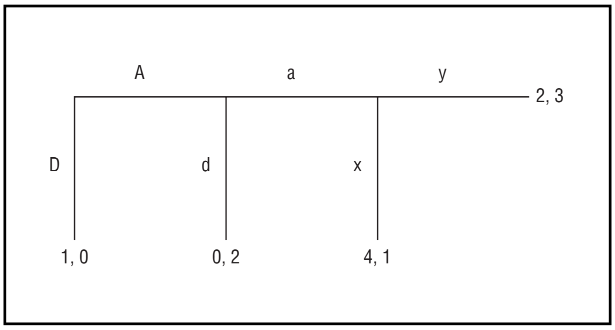

antees him a payoff of 4. Knowing (by assumption) that player 1 is rational, player 2—if she were to reach her decision node—would play *d*, since by playing *a* she would net a lower payoff. Finally, since (by assumption) player 1 knows that player 2 is rational and that she knows that player 1 is rational, he will choose *D* at his first decision node. The equilibrium (*Dy,d*) should therefore be ruled out, since it recommends an irrational move at the last node. In the normal form, both equilibria survive. The reason is simple: Nash equilibrium does not constrain behavior out of equilibrium. In this example, if player 1 plans to choose *D* and player 2 plans to choose *d*, it does not matter what player 1 would do at his last node, since that node will never be reached.

The sequential procedure one has used to conclude that only (*Dx,d*) is a reasonable solution is known as backward induction. In finite games of perfect information with no ties in payoffs, backward induction always identifies a unique equilibrium. The premise of the backward induction argument is that mutual rationality and the structure of the game are common knowledge among the players. It has been argued by Ken Binmore (1987), Cristina Bicchieri (1989, 1993), and Philip J. Reny (1992) that under certain conditions common knowledge of rationality leads to inconsistencies. For example, if player 2 were to reach her decision node, would she keep thinking that player 1 is rational? How would she explain player 1's move? If player 1's move is inconsistent with common knowledge of rationality, player 2 will be unable to predict future play; as a corollary, what constitutes an optimal choice at her node remains undefined. As a consequence of the previous criticisms, the usual premises of backward induction arguments have come to be questioned (compare Pettit and Sugden 1989, Basu 1990, Bonanno 1991). There are a number of further equilibrium refinements for games in extensive form. Their multiplicity makes it impossible to delve into details here. The interested reader can consult Bicchieri (1993, chapter 3).

GAMES OF INCOMPLETE INFORMATION

In games of incomplete information certain elements of the game are not common knowledge among the players. The knowledge and beliefs of the players have to be incorporated into the game-theoretic model, as one usually does in extensive form games, and an appropriate equilibrium concept has to be devised. The approach is based on the seminal work of John C. Harsanyi (1968). In the Bayesian approach adopted by Harsanyi, a player's uncertainty about variables that are relevant for his or her decision ought to be made explicit by means of probability distributions representing his or her beliefs. Moreover, second-order beliefs (beliefs about other players' beliefs) can be represented by further probability distributions, and third-order beliefs about second-order ones, and so on. The flexibility of Harsanyi's model allows one to incorporate all such infinite sequence of higher-order beliefs without an explicit representation of it.

The main idea is that the payoffs associated to each strategy profile depend on certain parameters $\theta_1, \dots, \theta_n$, one for each player 1, … , *n*. Each parameter is drawn from a set $\Theta_i = (a_i, b_i, \dots)$ associated with each player i. The composition of the sets Θ_i is known, yet the true value of the parameter θ_i is not (at least for one of the players). The parameter θ_i is called *i*'s type and the set Θ_i represents, intuitively, the other player's ignorance about *i*'s characteristics. A type amounts to a specification of certain variables: a player's strategy set, a player's preferences and payoff function, and so on, that make up the private information of a player. Although it is convenient to refer to "the type a_i of player *i*" as if it was a separate individual, one should keep in mind that types represent players' knowledge (and uncertainty about others) only. As mentioned earlier, in a Bayesian approach uncertainties are represented by probability distributions. Hence, each player has an initial probability distribution $\mu_i = (\mu_i(a_{-i}), \mu_i(b_{-i}), \dots)$ over the types of any other player other than i. Since in a Bayesian game the choices of a player depend on his or her type, the concept of Nash equilibrium has to be generalized accordingly.

Note that all that a player knows, except from the game itself (and the priors), is his own type, and the fact that the other players do not know his own type as well. As their best responses depend on the players' actual types, a player must see himself through his opponents' eyes and plan a best reply against the possible strategies of his opponents for each potential type of his own. Thus, a strategy in a Bayesian game of incomplete information must map each possible type of each player into a plan of

actions. Then, since the other players' types are unknown, each player forms a best reply against the expected strategy of each opponent, where he averages over the (well-specified) reactions of all possible types of an opponent, using his prior probability measure on the type space. Such a profile of type-dependent strategies which are unilaterally unimprovable in expectations over the competing types' strategies forms a Bayesian Nash equilibrium. In other words, a Bayesian Nash equilibrium is a Nash equilibrium "at the interim stage" where each player selects a best response against the average best responses of the competing players.

In the framework provided by Harsanyi (1968) it is possible to reduce a game of incomplete information to one of imperfect information. "Nature" is called to make the first move of the game, as if it was an actual player. Nature's random moves determine the type of each player, with a fixed probability that represents the prior probability attached to the events that player i is of type θ_i. Priors are assumed to be common knowledge, and players observe their own type only. Players then pick their strategies in this extended game, and it is possible to show that the equilibrium of such a game corresponds to the Bayesian Nash equilibrium of the game with incomplete information. In particular, the choice function s_i yields the action $\sigma_i(\theta_i)$ if and only if (iff) that is the action that player *i* chooses in the game with Nature when she observes her type θ_i.

EPISTEMIC FOUNDATIONS OF GAME THEORY

An important development of game theory is the so-called epistemic approach. In the epistemic approach to game theory strategic reasoning is analyzed on the basis of hypothesis about what players know about the game, about other players' knowledge, and about other players' rationality. Since Robert J. Aumann's (1976) formalization, the idea of common knowledge, and the analysis of what players choose depending on what their beliefs about each other are, began to play an increasingly important role in game theory. In particular, one can evaluate solution concepts by examining the epistemic assumptions and hypotheses from which they can be derived (compare Battigalli and Bonanno 1999). Such epistemic hypotheses are treated formally using the tools provided by interactive epistemology (compare Aumann 1999).

To formalize players' knowledge states, one considers a space set Ω whose elements are possible worlds. An event is then represented by a subset of Ω. For example, the proposition "it is sunny in Philadelphia" is represented by the set of all possible worlds in which it is sunny in Philadelphia. For each player, there exists an information function that partitions the space set. Intuitively, a player cannot distinguish among worlds belonging to the same cell of his or her information partition. Thus, in a possible world ω, player *i* knows an event *E* iff the set *E* (of possible worlds in which *E* obtains) includes the cell of his or her information partition containing ω. The intuition behind this is that if a player cannot distinguish among all the worlds in which *E* is true, then he or she knows that *E* is the case. It is possible to define a knowledge function K_i for each player *i* so that, when given *E* as an argument, it returns as a value the set of those worlds such that, for each one of them, the cell of *i*'s information partition that contains it is a subset of *E*. That is to say, K_iE is the event that *i* knows *E*.

By imposing certain conditions on the K_i's, one can force the epistemic functions to possess certain properties. For example, by requiring that K_iE be a subset of E, one requires that what players know is true, since in every possible world in which K_iE obtains, *E* obtains as well; similarly, by requiring that K_iK_iE be a subset of K_iE, one establishes that players know what they know, and by requiring that $Ki{-}KiE$ be a subset of $-K_iE$ that they know what they do not know (where – is the usual set-theoretical operation of complementation). The first condition is often referred to as the truth axiom, the second as the positive introspection axiom, and the third as the negative introspection axiom. Note that this setup has an equivalent formulation in terms of modal logics (compare Fagin et al. 1995, Meyer and van der Hoek 2004). To see the equivalence of the two approaches, consider that modal formulas express propositions whose semantic interpretation is given in terms of Kripke structures of possible worlds. It is then possible to establish a correspondence between formulas of the modal logic and events in the approach described earlier. In a Kripke model, then, an event corresponds to the set of those possible worlds that satisfy the formula expressing the proposition associated to that event.

Knowledge functions can be iterated, thus they can represent mutual and higher-order knowledge, and Aumann (1976) provides a mathematical definition of the idea of common knowledge in the setup sketched earlier. A proposition *p* is common knowledge between, say, two players *i* and *j* iff the set of worlds representing *p* includes the cells of *i*'s and *j*'s partitions meet that contain *p*, where the meet of two partitions is the finest common coarsening of them. An application of the definition is the

theorem proved in the same article, in which it is shown that if players have common priors, and their posteriors are common knowledge, then the posteriors are equal, even if the players derived them by conditioning on different information. Or, in other words, that one cannot "agree to disagree." As mentioned earlier, Aumann formalized Lewis's (1969) definition of common knowledge. However, it is currently debated whether Aumann's seminal definition is a faithful rendition of Lewis's informal characterization of common knowledge (compare Vanderschraaf 1998, Cubitt and Sugden 2003, Sillari 2005).

In such a framework it is possible to investigate which strategy profiles are compatible with certain epistemic assumptions about the players. For example, CK1 and CK2 imply that players would never choose strictly dominated strategies. The first contributions in this sense are David G. Pearce (1984) and B. Douglas Bernheim (1984), in which a procedure is devised to eliminate all the players' strategies that are not rationalizable, that is, not supported by internally consistent beliefs about other players' choices and beliefs. In general, it can be proved that certain epistemic conditions are only compatible with the strategy profiles yielded by a certain solution concept, hence providing an epistemic foundation for that solution concept. For example, Aumann and Adam Brandenburger (1995) proved that, for two-person games, mutual knowledge (i.e., first-order knowledge among all the players) of the structure of the game, of rationality, and of the players' chosen strategies implies.

CORRELATED EQUILIBRIUM. So far it has been assumed that players' strategies are independent, as though each player receives a private, independent signal and chooses a (mixed) strategy after having observed his or her own signal. However, signals need not be independent. For example, players can agree to play a certain strategy according to the outcome of some external jointly observed event, for example, a coin toss. If the agreement is self-fulfilling, in that players have no incentive to deviate from it, the resulting strategy profile is an equilibrium in correlated strategies or, in short, a correlated equilibrium (compare Aumann 1974, 1987). For any Nash equilibrium in mixed strategies, a correlation device can be set such that it generates a probability distribution over the possible outcomes of the game yielding such an equilibrium profile. Note, however, that the set of correlated equilibria of a game is much larger than the corresponding set of Nash equilibria. If the correlation signal is common knowledge among the players, one speaks of perfect correlation. However, players may correlate their strategies according to different signals (less than perfect correlation). The idea is that players have information partitions whose cells include more than one possible outcome, since they ignore which signals are received by other players. To represent the fact that players receive different signals (i.e., they ignore which strategies will be chosen by other players), it is required that in every cell of the information partition of player *i* his or her strategy does not change. It is then possible to calculate the expected payoff of playing the strategy indicated by the correlation device versus the expected payoff obtained by playing a different strategy. If the players have no incentive to deviate from the indicated strategy, the profile yielded by the correlation device is an equilibrium. Correlation by means of private signals may generate outcomes more efficient than those obtained by playing a Nash equilibrium. An important philosophical application of correlated equilibrium is found by Peter Vanderschraaf (1998, 2001), in which conventions as defined by Lewis (1969) are shown to be correlated equilibria of coordination games.

EVOLUTIONARY GAME THEORY

A Nash equilibrium need not be interpreted as a unique event. If one thinks of it as an observed regularity, one wants to know by what process such an equilibrium is reached and what accounts for its stability. When multiple equilibria are possible, one wants to know why players converged to one in particular and then stayed there. An alternative way of dealing with multiple equilibria is to suppose that the selection process is made by nature.

Evolutionary theories are inspired by population biology (e.g., see Maynard Smith 1982). These theories dispense with the notion of the decision maker, as well as with best responses/optimization, and use in their place a natural selection, a "survival of the fittest" process (with mutations) to model the frequencies with which various strategies are represented in the population over time. In a typical evolutionary model players are preprogrammed for certain strategies and are randomly matched with other players in pairwise repeated encounters. The relative frequency of a strategy in a population is simply the proportion of players in that population who adopt it. The theory focuses on how the strategy profiles of populations of such agents evolve over time, given that the outcomes of current games determine the frequency of different strategies in the future.

As an example, consider the game in Figure 5 and suppose that there are only two possible behavioral types: hawk and dove.

A hawk always fights and escalates contests until it wins or is badly hurt. A dove sticks to displays and retreats if the opponent escalates the conflict; if it fights with another dove, they will settle the contest after a long time. Payoffs are expected changes in fitness due to the outcome of the game. Fitness here means just reproductive success (e.g., the expected number of offspring per time unit).

Suppose injury has a payoff in terms of loss of fitness equal to *C*, and victory corresponds to a gain in fitness *B*. If hawk meets hawk, or dove meets dove, each has a 50 percent chance of victory. If a dove meets another dove, the winner gets *B* and the loser gets nothing, so the average increase in fitness for a dove meeting another dove is $B/2$. A dove meeting a hawk retreats, so his or her fitness is unchanged, whereas the hawk gets a gain in fitness *B*. If a hawk meets another hawk, they escalate until one wins. The winner has a fitness gain *B*, the loser a fitness loss *C*. So the average increase in fitness is $(B - C)/2$. The latter payoff is negative, since one assumes the cost of injury is greater than the gain in fitness obtained by winning the contest. One can also assume that players will be randomly paired in repeated encounters, and in each encounter they will play the stage game of Figure 5.

FIGURE 5

dove/hawk	D	H
D	B/2, B/2	0, B
H	B, 0	(B-C)/2, (B-C)/2

If the population were to consist predominantly of hawks, selection would favor the few doves, since hawks would meet mostly hawks and end up fighting with an average loss in fitness of $(B - C)/2$, and $0 > (B - C/2)$. In a population dominated by doves, hawks would spread, since every time they meet a dove (which would be most of the time) they would have a fitness gain of *B*, whereas doves on average would only get $B/2$. Evolutionary game theory wants to know how strategies do on average when games are played repeatedly between individuals who are randomly drawn from a large population. The average payoff to a strategy depends on the composition of the population, so a strategy may do well (in terms of fitness) in an environment and poorly in another. If the frequency of hawks in the population is *q* and that of doves correspondingly $(1 - q)$, the average increase in fitness for the hawks will be $q(B - C)/2 + (1 - q)B$, and $(1 - q)B/2$ for the doves. The average payoff of a strategy in a given environment determines its future frequency in the population. In this example, the average increase in fitness for the hawks will be equal to that for the doves when the frequency of hawks in the population is $q = B/C$. At that frequency, the proportion of hawks and doves is stable. If the frequency of hawks is less that B/C, then they do better than doves and will consequently spread; if their frequency is larger than B/C, they will do worse than doves and will shrink.

Note that if $C > B$ then $(B - C)/2 < 0$, so the game in Figure 5 has two pure-strategy Nash equilibria: (*H*, *D*) and (*D*, *H*). There is also a mixed strategy equilibrium in which hawk is played with probability $q = B/C$ and dove is played with probability $(1 - q) = C - B/C$. If the game of Figure 5 were played by rational agents who choose which behavior to display, one would be at a loss in predicting their choices. From common knowledge of rationality and of the structure of the game, the players cannot infer that a particular equilibrium will be played. In the hawk-dove example, however, players are not rational and do not choose their strategies. So if an equilibrium is attained it must be the outcome of some process very different from rational deliberation. The process at work is natural selection: High-performing strategies increase in frequency whereas low-performing strategies' frequency diminishes and eventually goes to zero.

One has seen that in a population composed mostly of doves, hawks will thrive, and the opposite would occur in a population composed mainly of hawks. So for example, if hawks dominate the population, a mutant displaying dove behavior can invade the population, since individuals bearing the dove trait will do better than hawks. The main solution concept used in evolutionary game theory is the evolutionarily stable strategy (ESS) introduced by John Maynard Smith and George R. Price (1973). A strategy or behavioral trait is evolutionarily stable if, once it dominates in the population, it does strictly better than any mutant strategy, and hence it cannot be invaded. In the hawk-dove game, neither of the two pure behavioral types is evolutionarily stable, since each can be invaded by the other. One knows, however, that a population in which there is a proportion $q = B/C$ of hawks and $(1 - q) = C - B/C$ of doves is stable. This means that the type of behavior that consists in escalating fights with

probability $q = B/C$ cannot be invaded by any other type, hence it is an ESS. An ESS is a strategy that, when it dominates the population, is a best reply against itself. Therefore, an evolutionarily stable strategy such as $(B/C, C - B/C)$ is a Nash equilibrium. Though every ESS is a Nash equilibrium, the reverse does not hold; in our stage game, there are three Nash equilibria, but only the mixed strategy equilibrium $(B/C, C - B/C)$ is an ESS.

Evolutionary games provide one with a way of explaining how agents that may or may not be rational and—if so—subject to severe information and calculation restrictions, achieve and sustain a Nash equilibrium. Philosophical implications and applications can be found in the works of Brian Skyrms (1990, 1996, 2004). When there exist evolutionarily stable strategies (or states), one knows which equilibrium will obtain, without the need to postulate refinements in the way players interpret off-equilibrium moves. Yet we need to know much more about processes of cultural transmission and to develop adequate ways to represent payoffs, so that the promise of evolutionary games is actually fulfilled.

See also Decision Theory; Philosophy of Biology; Philosophy of Economics.

Bibliography

Aumann, Robert J. "Agreeing to Disagree." *The Annals of Statistics* 4 (6) (1976): 1236–1239.

Aumann, Robert J. "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica* 55 (1987): 1–18.

Aumann, Robert J. "Interactive Epistemology I: Knowledge." *International Journal of Game Theory* 28 (1999): 263–300.

Aumann, Robert J. "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics* 1 (1974): 67–96.

Aumann, Robert J., and Adam Brandenburger. "Epistemic Conditions for Nash Equilibrium." *Econometrica* 63 (1995): 1161–1180.

Basu, Kaushik. "On the Non-existence of a Rationality Definition for Extensive Games." *International Journal of Game Theory* 19 (1990): 33–44.

Battigalli, Pierpaolo, and Giacomo Bonanno. "Recent Results on Belief, Knowledge, and the Epistemic Foundations of Game Theory." *Research in Economics* 503 (1999): 149–226.

Bernheim, B. Douglas. "Rationalizable Strategic Behavior." *Econometrica* 52 (4) (1984): 1007–1028.

Bicchieri, Cristina. *Rationality and Coordination*. New York: Cambridge University Press, 1993.

Bicchieri, Cristina. "Self-Refuting Theories of Strategic Interaction: A Paradox of Common Knowledge." *Erkenntnis* 20 (1989): 69–85.

Binmore, Ken. "Modeling Rational Players I." *Economics and Philosophy* 3 (1987): 179–214.

Bonanno, Giacomo. "The Logic of Rational Play in Games of Perfect Information." *Economics and Philosophy* 7 (1991): 37–61.

Cubitt, Robin, and Robert Sugden. "Common Knowledge, Salience, and Convention: A Reconstruction of David Lewis's Game Theory." *Economics and Philosophy* 19 (2003): 175–210.

Damme, Eric van. *Stability and Perfection of Nash Equilibria*. Berlin: Springer, 1987.

Fagin, Ronald, et al. *Reasoning about Knowledge*. Cambridge, MA: MIT Press, 1995.

Harsanyi, John C. "Games with Incomplete Information Played by 'Bayesian' Players, Parts I, II and III." *Management Science* 14 (1967–1968): 159–182, 320–334, 486–502.

Kohlberg, Elon. "Refinement of Nash Equilibrium: The Main Ideas." In *Game Theory and Applications*, edited by Tatsuro Ichiischi, Abraham Neyman, and Yair Tauman. San Diego, CA: Academic Press, 1990.

Lewis, David K. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press, 1969.

Maynard Smith, John. *Evolution and the Theory of Games*. New York: Cambridge University Press, 1982.

Maynard Smith, John, and George R. Price. "The Logic of Animal Conflict." *Nature* 246 (1973): 15–18.

Meyer, John-Jules, and Wiebe van der Hoek. *Epistemic Logic for AI and Computer Science*. New York: Cambridge University Press, 2004.

Nash, John F., Jr. *Essays on Game Theory*. Cheltenham, U.K.: E. Elgar, 1996.

Pearce, David G. "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica* 52 (4) (1984): 1029–1050.

Pettit, Philip, and Robert Sugden. "The Backward Induction Paradox." *Journal of Philosophy* 86 (1989): 169–182.

Reny, Philip J. "Rationality in Extensive Form Games." *Journal of Economic Perspectives* 6 (1992): 103–118.

Schelling, Thomas C. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press, 1960.

Selten, R. "Spieltheoretische Behandlung eines Oligopolmodels mit Nachfragetragheit." *Zeitschrift fur die gesamte Staatwissenschaft* 121 (1965): 301–324.

Sillari, Giacomo. "A Logical Framework for Convention." *Knowledge, Rationality, and Action*. 2005.

Skyrms, Brian. *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard University Press, 1990.

Skyrms, Brian. *Evolution of the Social Contract*. New York: Cambridge University Press, 1996.

Skyrms, Brian. *The Stag Hunt and the Evolution of Social Structure*. New York: Cambridge University Press, 2004.

Vanderschraaf, Peter. "Knowledge, Equilibrium, and Convention." *Erkenntnis* 49 (1998): 337–369.

Vanderschraaf, Peter. *Learning and Coordination: Inductive Deliberation, Equilibrium, and Convention*. New York: Routledge, 2001.

Cristina Bicchieri (2005)
Giacomo Sillari (2005)

GARRIGOU-LAGRANGE, RÉGINALD MARIE

(1877–1964)

Réginald Marie Garrigou-Lagrange notably influenced the revival of Thomism in some European and American philosophical circles. He was born Gontran-Marie Garrigou-Lagrange at Auch, France. His first university studies were in the faculty of medicine at the University of Bordeaux. After two years, however, he chose to embrace the priesthood and on May 20, 1900, made his profession as a Dominican, receiving the name Réginald Marie.

In addition to the regular course of philosophy as a Dominican, he pursued graduate studies at the Sorbonne, where he had the opportunity to attend the lectures of Henri Bergson. In 1909 Garrigou-Lagrange entered into what proved to be a long career as professor at the international university of philosophical and theological studies in Rome, now called the Universitas Studiorum Pontificia S. Thomae Aquinatis in Urbe. He remained in this position until 1959. Although his courses were primarily in the theological faculty, it is significant that throughout his teaching life he lectured each week on the metaphysics of Thomas Aquinas. Garrigou-Lagrange was also a founding member of the Academia Pontificia Academia Romanae S. Thomae Aquinatis.

An accurate view of the philosophical thought of Garrigou-Lagrange must take into account the fact that he was not simply a philosopher; his professional labors as well as his writings are preponderantly theological. However, because his concern was with the teachings of Thomas Aquinas, his work has a philosophical import on two counts. First of all, the Thomistic theological synthesis is characterized by its employment of the speculative resources of human intelligence. In this concentration on theology, then, Garrigou-Lagrange necessarily devoted himself to the exposition of the basic Thomistic philosophical positions. Second, from the beginning of his career Garrigou-Lagrange was faced with a challenge to the relevance and the validity of Thomism, or indeed of any metempirical assertions of the human mind. It is to this challenge that his purely philosophical labors and writings are principally addressed.

His first book, *Le sens commun, la philosophie de l'être et les formules dogmatiques,* is a rejoinder to the position taken by Édouard Le Roy in a series of articles (*Revue de métaphysique et morale,* 1899–1901). Le Roy alleged all expressions of truth by the human mind to be totally relative, mutable, and conditioned. Human thought is simply the expression of de facto acceptations, significant according to that natural and subjective orientation of the human mind which for Le Roy is the *sens commun.* Against this Bergsonian usage, Garrigou-Lagrange used the term *sens commun* to designate the commonly assumed character of the human mind, namely, its extramental orientation toward objectively existent and intelligible reality. He set himself the task of vindicating this realism, of defending the objective validity and transcendental range of human thought.

The basic themes of his position are readily discernible. The human intelligence has "being" as its connatural object. In its attainment of being the human mind surpasses sense knowledge, goes beyond mere phenomena. The first principles of human reason—identity, contradiction, causality, and finality—are not mere subjective thought patterns; they are grounded in being. The human evaluation of the data of experience in virtue of such principles, then, has an ontological validity; the human mind is capable of assertions concerning the real that are objectively true and absolute. Because in its attainment of being the mind goes beyond mere phenomena, the principles of philosophical inquiry have a transcendental validity. Man is able, consequently, to achieve true judgments, not only about the entitative structure of experienced reality, but also about the nonexperienced but necessarily affirmed primary cause of the beings of experience. The connaturally realistic orientation of human intelligence, therefore, provides the capacity for objectively valid metaphysical evaluations of reality and even for a true natural theology.

Garrigou-Lagrange maintained that Thomas presented a philosophy of being that was an effectively enunciated and developed expression of the natural metaphysical orientation of human intelligence in which the *sens commun* has its scientifically articulated realization. Garrigou-Lagrange's principal philosophical contribution, then, was a forceful and clear exposition of the basic Thomistic insights. In his writings there is a clear and honest confrontation of Thomistic realism with both nominalist empiricism and Kantian subjectivism.

An evaluation of the work of Garrigou-Lagrange must place it in relation to the so-called Neo-Scholastic movement. Since his career began well after the early attempts to reassert Thomism, his writings are free of the alien influences present in the work of the restoration's pioneers. His chief concern, the basic critical problem of the validity of human intelligence, is a central issue in all Neo-Scholastic philosophy. In the light of subsequent developments among Neo-Scholastic philosophers, and even among Thomists, concerning the critical problem,

the approach of Garrigou-Lagrange may be designated as somewhat simplified. He strove to set forth directly the positive statements of a philosophy of being against a philosophy of becoming, to manifest the human mind as a faculty of truth, not an amasser or coordinator of data. Later Thomists have sought by more reflective methods to show how being manifests itself in the very process of cognition as the evidential justification of human knowledge. Their efforts are a refinement of the task to which the efforts of Garrigou-Lagrange were directed. His work, then, was a necessary stage in a vital development. Because of his dedication to the thought of Thomas, he directed that development to a more fruitful use of Thomas's understanding of the problems of being and intelligence.

See also Bergson, Henri; Le Roy, Édouard; Thomas Aquinas, St.; Scotism; Thomism.

Bibliography

WORKS BY GARRIGOU-LAGRANGE

Books

Le sens commun, la philosophie de l'être et les formules dogmatiques. Paris, 1909; 4th ed., Paris, 1936.

Dieu, son existence et sa nature. Paris, 1915; 6th ed., Paris, 1950. Vols. I and II translated by Dom Bede Rose as *God, His Existence and His Nature, a Thomistic Solution of Certain Agnostic Antinomies.* St. Louis, 1934.

Le réalisme du principe de finalité. Paris: Desclée, de Brouwer, 1932.

La synthèse Thomiste. Paris: de Brouwer, 1946. Translated by Dom Patrick Cummins as *Reality, a Synthesis of Thomistic Thought.* St. Louis: Herdner, 1950.

Articles

Angelicum. 14 (1937): 5–37. A complete bibliography of Garrigou-Lagrange's numerous articles through 1937.

Bulletin Thomiste. For the discussion of the critical problem among Thomists during the height of Garrigou-Lagrange's career, see the following: 3 (1932): 451ff.; 4 (1935): 400ff.; 5 (1938): 365ff.; 6 (1940–1942): 236ff.

Thomas C. O'Brien, O.P. (1967)

GARVE, CHRISTIAN

(1742–1798)

Christian Garve, the German "popular philosopher," was born in Breslau. After studying at Frankfurt an der Oder, Halle, and Leipzig, he became extraordinary professor of philosophy at Leipzig in 1770, but in 1772 he resigned on account of ill health and moved to Breslau. In 1779 Frederick II called him to Charlottenburg, where he remained until his death.

Garve's interests were mainly in practical morality and empirical psychology. He sought useful knowledge and was averse to abstract speculation. He drew inspiration from Duc François de La Rochefoucauld and Claude-Adrien Helvétius, and especially from the British moralists. His translations of Adam Ferguson, Edmund Burke, Alexander Gerard, Adam Smith, and other British authors were important in popularizing British moral philosophy and aesthetics in Germany. He also translated and commented on the moral and political works of Aristotle and Cicero.

In his own writings Garve studied the individual characteristics and inclinations of different men, and their interrelation in society. He explained their differences by a difference in the degree of clarity and vividness of the ideas they possessed. Interest—the participation of an individual in the feelings, ideas, and actions of another—was a central notion in his psychology. It was derived from the "benevolence" and "sympathy" then current in British thought. In Garve's works psychology, sociology, and ethics were interwoven. His goal was that of a social psychologist, moralist, and educator. Immanuel Kant drew from Garve some elements of his moral and religious philosophy.

Bibliography

WORKS BY GARVE

Über die Verbindung der Moral mit der Politik. Breslau, 1788.

Über verschiedene Gegenstände aus der Moral, Literatur und dem gesellschaftlichen Leben, 5 vols. Breslau, 1792.

Vermischte Aufsätze. 2 vols. Breslau, 1796–1800.

Über Gesellschaft und Einsamkeit. 2 vols. Breslau, 1797–1800.

Übersicht der vornehmsten Principien der Sittenlehre. Breslau, 1798.

Einige Betrachtungen über die allgemeinen Grundsätze der Sittenlehre. Breslau, 1798.

Sammlung einiger Abhandlungen. 2 vols. Breslau, 1802.

WORKS ON GARVE

Manso, S. J. C. *Chr. Garve in seinem schriftstellerischen Charakter.* Breslau, 1799.

Müller, P. *Christian Garves Moralphilosophie und seine Stellungnahme zu Kants Ethik.* Erlangen, 1905.

Schelle, G. G. *Briefe über Garves Schriften und Philosophie.* Leipzig, 1800.

Stern, A. *Ueber die Beziehungen Christian Garves zu Kant.* Leipzig, 1884.

Van Der Zande, Johan. "The Microscope of Experience: Christian Garve's Translation of Cicero's *De Officiis* (1783)." *Journal of the History of Ideas* 59 (1) (1998): 75–94.

Giorgio Tonelli (1967)
Bibliography updated by Tamra Frei (2005)

GASSENDI, PIERRE
(1592–1655)

Pierre Gassendi, the leading French seventeenth-century skeptical and Epicurean philosopher and scientist, was born at Champtercier, a Provençal village in France. He studied at Digne and Aix-en-Provence and was appointed professor of rhetoric at Digne at the age of twenty-one. In 1614 he received his doctorate in theology at Avignon. He was ordained a priest in 1616 and was appointed professor of philosophy at Aix. From 1617 to 1623 he lectured on Aristotle's philosophy, developing a forceful critique of it. His first published work, *Exercitationes Paradoxica Adversus Aristoteleos* (1624), was intended to be followed by six more parts, of which only the second part, published posthumously, was written. It contains both an attack on Aristotle's thought and portions of Gassendi's mitigated skepticism.

After a year in Digne, during which he performed various ecclesiastical duties, Gassendi visited Paris for a brief period in 1625 and became friendly with such avant-garde thinkers as Francois de La Mothe le Vayer and Marin Mersenne. He continued the astronomical researches that he had begun in Provence, and, with the mathematician Claude Mydorge (1585–1647), observed a lunar eclipse. Gassendi's careful astronomical records from 1618 to 1655 were published after his death. He also engaged in many scientific studies with his patron, Nicolas-Claude Fabri de Pieresc (1580–1637). His discovery of the perihelion of Mercury was an important support for the Copernican theory at the time.

Gassendi returned to Paris in 1628, remained there until August 1629, and then spent nine months in Flanders and Holland, where he met many leading scientists and scholars. He spent the next years partly in Paris and partly in Provence, publishing scientific works on astronomy and physics and presenting skeptical attacks on Herbert of Cherbury and the Rosicrucian Robert Fludd. He undertook an intensive study of Epicurean atomism, a subject in which he had been interested for some time. The results of this study were to form a basic part of his later writings.

In 1634 Gassendi was elected provost of the Cathedral of Digne. In 1641 he was sent to the assembly of the French clergy in Paris and during this visit taught philosophy to the young Molière (1622–1673). Gassendi was appointed to the chair of mathematics at the Royal College (now the Collège de France) in 1645, but because of ill health he was away from his post from 1648 to 1653. He fell ill in 1654 and died the following year.

Except for his early attacks on Aristotelianism, Fludd's Rosicrucianism, and Herbert of Cherbury, Gassendi's philosophical works date from the 1640s onward. In 1641, at the request of Mersenne, Gassendi wrote his objections to René Descartes's *Méditations* ("Fifth Set of Objections"). Descartes's testy answer led Gassendi to expand his criticism into the bulky *Disquisitio Metaphysica*, finished in 1642 and published in Amsterdam in 1644. He published three works on Epicurus and his philosophy between 1647 and 1649: *De Vita et Moribus Epicuri* (1647), *Animadversiones in Decimum Libri Diogenis Laertii, qui est de Vita, Moribus Plascitisque Epicuri* (1649), and *Syntagma Philosophiae Epicuri, cum Refutationibus Dogmatum, Quae Contra Fidem Christianum ab eo Asserta Sunt* (1649). His most important philosophical writings appeared only posthumously, in the 1658 edition of his complete works. His overall treatment of philosophical problems appears in *Syntagma Philosophicum* (*Opera*, volumes 1 and 2) and in the second part of his first work, the *Exercitationes*, which presents his constructive or mitigated skepticism.

EXERCITATIONES

Gassendi's thought developed from a fairly thoroughgoing skepticism, strongly influenced by Sextus Empiricus, Michel Eyquem de Montaigne, Pierre Charron, and Francisco Sanches, to what he called a *via media* between skepticism and dogmatism. The *via media* involved both a fundamental epistemological skepticism and a hypothetical form of Epicurean atomism that was modified to eliminate those aspects of Epicurus's thought that conflicted with Christian doctrine. In the first part of the *Exercitationes* Gassendi, following in the tradition of Francesco Patrizi and Peter Ramus, tried to show all the erroneous or dubious aspects of Aristotelianism. The second part set forth an attack on all those who claimed to have discovered necessary and indubitable knowledge of the real nature of things. Our knowledge of the world, Gassendi insisted, comes only from sensory experience. We are unable to arrive at absolutely true first principles and real or essential definitions, since inductions from experience can never yield certain universal prepositions. No matter how much data are gathered, a negative instance may still turn up in the future. Even if we somehow managed to discover some genuine definitions and first principles, no further scientific knowledge about nature could be gained by employing syllogistic reasoning, since, as the Greek Pyrrhonists had shown, the truth of the premises of a syllogism depends on antecedent knowledge that the conclusion is true. Either the conclu-

sion is part of the evidence of the premises or the syllogism establishes nothing, since it is not known whether the premises are true.

In the concluding section Gassendi launched his strongest attack on the possibility of gaining necessary knowledge about the world. Using the arguments of the ancient skeptics, he tried to show that all that we can know is how things appear, not how they really are in themselves. We can know that honey seems sweet, but we cannot find out if it really is sweet. On the basis of appearances we cannot tell what the real nature of things must be that produces such effects on us. Sense experience varies too much to provide any means for determining what reality is like on the basis of what is perceived. We lack any means of reasoning from experience to what has caused it. We are not even able to establish any criterion of true knowledge. Hence, we can only conclude that nothing can be known about reality. However, in this early work Gassendi insisted that we can develop useful sciences about appearances. As long as we restrict our conclusions to the world of experience, we will neither come in conflict with divine truth, nor accept any dubious dogmatic theory about unperceived reality. Such theories, whether metaphysical or mathematical, are presumptuous conjectures that have no value whatsoever.

OBJECTIONS TO DESCARTES

In his middle period Gassendi challenged those who claimed to have discovered some means of knowing the real nature of things. He employed various skeptical arguments against Renaissance naturalists and against such leading "new philosophers" as Herbert of Cherbury and Descartes. Gassendi's two letters against Herbert's *De Veritate* (in which the latter claimed to have refuted skepticism) used arguments about diversity of experience and disagreements among individuals to counter Herbert's common notions and common consent theory. In the "Fifth Set of Objections" and the *Disquisitio Metaphysica*, Gassendi turned skeptical argumentation against Cartesianism. He tried to show that Descartes's method of doubt illustrated what the skeptics had claimed for centuries. Then Gassendi challenged Descartes's positive dogmatic conclusions. Gassendi contended that the vaunted Cartesian criterion of true knowledge (that of clarity and distinctness) was useless, since people often think that they clearly and distinctly perceive something and then turn out to be wrong. Hence, to employ this criterion, another criterion would be needed to determine when something really is clear and distinct and does not merely appear to be so. In addition, another criterion would be necessary to employ this second criterion, and so forth.

The culmination of Gassendi's attack, which Descartes called the objections of objections, was his posing the possibility that all knowledge, even if it were clear and distinct, might not be about anything outside of our minds. If this could be the case, then all the knowledge purported to be found by Descartes might be imaginary or fictitious. Descartes saw this suggestion as a fundamental challenge to his system and as a denial of the possibility of gaining knowledge about any reality other than our own thoughts. His reply consisted in refusing to take the objection seriously, since if one did, "it follows that there is nothing that we can in any way comprehend, conceive or imagine, that should be accepted as true, that is to say that we have to shut the door completely on reason, and be content to be monkeys, or parrots, and no longer men" (Descartes, *Oeuvres*, C. Adam and P. Tannery, eds. vol. IXA, 212).

MITIGATED SKEPTICISM

In his later writings Gassendi attempted to develop a mitigated skepticism that would show how we could possess worthwhile knowledge about the world of appearances and how a science of this world could be developed, using Epicurean atomism as a hypothetical model.

Gassendi's mature theory about our knowledge of the world appears in his *Syntagma Philosophicum*, published in 1658 after his death. The work is enormous, containing 1,600 folio pages, printed in double columns. It is divided into three general sections, the first dealing with logic and theory of knowledge, the second with the natural world, and the third with ethics. Because of his skepticism, Gassendi did not regard metaphysics as a serious subject and so he omitted it entirely from his book.

At the outset, Gassendi seeks to establish a way to knowledge that is between the doubts of the skeptics and the complete assurances of the dogmatists. Neither the view that we can know nothing nor the view that we can know everything is tenable. The skeptics admit that we can know how nature appears to us. But they deny that we can know more than this. On the contrary, the dogmatists claim that we can know the real nature of things, which are not apparent to us. This, Gassendi contends, is exaggerating the power of the human mind. However, between skepticism and dogmatism there is a third possibility, which has been called constructive or mitigated skepticism, an acceptance of the thesis that although in a fundamental sense we cannot gain certain knowledge of the nature of reality, we can nevertheless gain a type of

knowledge that we need have no reason to doubt and that will suffice to enable us to understand the world.

This limited knowledge is obtained first by accepting what is obvious to us, our sense experience, plus certain obvious conclusions from it, such as that things exist. Signs found in sense experience enable us to know about other matters not immediately obvious to the senses. The ancient Greek skeptics had admitted that, on the basis of the constant conjunctions found in experience, we could judge that certain things temporarily not apparent to us were the case, such as when we see smoke, we can judge that there is a fire. In addition, we are also able to judge, by means of our reasoning ability, that particular sense experiences indicate that the world has certain features, even though we are never able to perceive these features. Thus, we can judge from the appearance of sweat on the skin that it has pores.

Long before the invention of the telescope, Democritus was able to judge from the white color of the Milky Way that it is composed of an innumerable quantity of stars. This type of reasoning, which leads us to knowledge about the world, is based on a careful and cautious evaluation of our sense information by our reason, plus inferences, made from this information, based on careful reasoning and on certain general principles that we have learned from experience. The conclusions we reach in this way about the nature of the world are beyond doubt and are ultimately evaluated in terms of future information gained from experience (as in the case of the Milky Way) and from these conclusions in explaining the course of our experience. We do not discover the absolute truth in this way, but only a faint shadow of it. This faint shadow will turn out to be the most satisfactory scientific explanation that can be given of experience in terms of the hypothesis (confirmed by experience and reasoning) that the world is composed of atoms in motion.

In terms of this theory of knowledge Gassendi examines various logical systems, ancient and modern, to state the best method for attaining limited knowledge. Many of the classical devices, Gassendi finds, are practically useless. The philosophies of Francis Bacon and Descartes have serious defects, Gassendi claims. Our senses can err, and we cannot, no matter what we do, attain real knowledge of the inner nature of things. But a logical method that is based on sense information carefully analyzed on general, unquestionable principles gained from experience and careful reasoning, and constantly checked and verified, can serve as the instrument for attaining what truth is possible.

ATOMISM

According to Gassendi what we can know about the world consists of a modified form of the atomism of Epicurus, modified in terms of the science of the time and the religious principles Gassendi maintained that he accepted. (Whether Gassendi was a sincere Christian has been, and still is, debated among scholars.) After surveying and criticizing the views of various philosophers about the nature of the world, Gassendi offers as the most probable theory (but not as the necessarily true one) the view that the actual components of the universe are indivisible atoms, moving in empty space. The supposed atoms are assumed to have been eternally created by God, to have different shapes, and to be moving at different rates of speed. Gassendi did not want mathematically describable atoms, since he feared this would lead to some sort of mathematical metaphysics. His atoms had features like those of ordinary experience. The atoms collided and presumably the result of all the collisions is the world that we perceive. A mechanical model related to our experience can then allow us to find empirical laws, make predictions, and explain relationships between different kinds of phenomena. In so relating the phenomenal world to the atomic world, there is no longer any need for Aristotelian purposes.

Appealing to the recent findings of scientists such as Evangelista Torricelli (1608–1647) and Blaise Pascal, Gassendi insists that the essential feature of atoms is solidity. In addition, they have the properties of extension, figure, and weight. They are conceived of as having the kinds of configurations found in ordinary experience, like those of wagon wheels and houses, rather than mathematically describable sizes and shapes. Gassendi had a distrust of those who maintained that nature was to be described in mathematical terms, since he felt that they were probably advocating some type of Platonic metaphysical theory about the nature of reality.

God has created the atoms and given them an impulse to move downward. They move at different rates of speed, and for this reason they collide with one another. The collisions change the courses of the atoms, causing still further collisions, and so on. The various changes that take place in the world, both on the apparent and on the nonapparent level, can be accounted for by the movements of the atoms, their collisions, and their combinations. Thus, the real world is conceived as a mechanism made up of small moving parts, the atoms. The qualities and movements of the atoms suffice to account for changes in the real world and the way in which the world appears to us. The qualities that we per-

ceive, the colors, sounds, tastes, smells, and so on (the so-called secondary qualities), are not actually properties of real objects. Instead, they are the ways we perceive various atomic movements when they affect our sense organs.

Gassendi begins to modify his Epicurean theory when he discusses the cause of the movement of the atoms. He accepts the Scholastic thesis that the primary cause of motion is God. The evidence that God exists is the almost universal natural belief in a deity and the conclusion drawn from observing the order in the universe, namely, that there must be an orderer or designer of the world. That there are some atheists is dismissed by Gassendi as similar to the existence of blind people. That a few people lack the normal, natural human faculties and beliefs is no reason to doubt the reliability of the faculties and beliefs of the rest of humankind. Both the senses and our reasoning ability give us an adequate basis for accepting the view that there is a God.

One's conception of God is that of an omniscient and omnipotent being who is all-wise and all-good. He is the author and providential guide and cause of everything that exists and everything that happens in the world. Gassendi specifically rejects Epicurus's view that everything can be explained and accounted for solely in terms of the atoms and their motions. Where, he asks, do the atoms come from, and what makes them move? Furthermore, if the world were produced only by "the fortuitous concourse of atoms," why is it that the atoms never, by themselves, make a house, or a temple, or a book? Each of these seems to require a designer to organize the atoms in a specific way, and so does the universe in general.

Turning from physical events to mental ones, Gassendi attempts to give an atomic explanation of the nature of the soul. First, he exhibits his vast erudition by examining the opinions of many different ancient philosophers on the subject. Then he offers the theory that seemed most probable to him; namely, that the animal soul is a material object. Though we cannot see the soul, reason convinces us that it must exist. The various processes that occur in living beings, such as nutrition, sensation, and movement, could not take place were there not a soul. But what is the soul like? It is a tenuous material substance existing in the body. It is like a subtle fire, giving life to corporeal things somewhat as fire warms objects.

The human soul, however, is more complex than the animal soul, being composed of two parts. The first is the irrational soul, which is material and is like the soul of any other living thing. It accounts for the vegetative and sensitive processes that exist in man. This part of the human soul comes to us from our parents. Besides this, we possess another feature of our souls, the rational element that, Gassendi insists, contrary to Epicurus's view, is not corporeal and is not derived from other human beings, but only from God. The rational part of our souls, which is responsible for our higher intellectual activities, is also immortal. Epicurus had argued for the mortality of the soul, but Gassendi strongly insists that only the animal soul is mortal. As evidence for his belief in the immortality of the rational soul, Gassendi contends that the fact that it is immaterial suffices to show that it is immortal. Furthermore, the universal agreement of humankind on this point is offered as another proof, as well as the view that the divine and just government of the world would seem to require human immortality for a proper system of rewards and punishments to function.

Gassendi apparently believed that there was no conflict between his atomism and his views about man and God. Hence, Catholicism could be compatible with a strictly material account of the natural world. And although Gassendi was a heliocentrist he tried to present his astronomical views in such a way that they did not conflict with those of the church in its condemnation of Galileo Galilei.

Gassendi's atomism was as complete a scientific theory as any other offered in the first half of the seventeenth century. It rivaled Descartes's. However, as science developed later on, Gassendi's picture was replaced by that of Isaac Newton and others. No important discoveries are attributed to Gassendi's great scientific program.

HUMAN PSYCHOLOGY

In his discussion of human psychology Gassendi presents a theory to explain how the various mental processes take place. This section culminates with an examination of the sources of all of our knowledge, which, to some extent, anticipates the views that appear in John Locke's *An Essay concerning Human Understanding* (1690).

The faculties of sensation and imagination are common to humans and animals. Gassendi even asserts that sensation occurs to some extent in plants and minerals. Sensation occurs by means of a physical process involving material particles affecting a sense organ and causing a sensation, which is a physical event in the brain. The faculty of the imagination, which includes the memory as well, operates on traces or remains of the physical sense impressions. These traces are conceived of as waves in the brain that are actuated by other motions in the body and then cause further movements in the brain, giving rise to sensations or feelings similar to the original sensation

that caused the wave. Much of the account offered by Gassendi is close to that presented by his contemporary, the materialistic philosopher Thomas Hobbes.

The imagination has three functions: apprehension, judgment, and reasoning. We can apprehend, as a result of the wave motions, the exact experiences and sensations that have occurred. Because of movements inside and outside us, the various waves can be agitated at later times, so that we can now be aware of what we experienced yesterday. Also, different features of different experiences can be apprehended at the same time, giving rise to apprehensions of objects that have never, as such, been experienced. Thus, for example, our apprehension of a centaur results from our previous sense experience of a horse and a man, plus the simultaneous activation of part of the remaining wave that came from each of them. Judging and reasoning, which Gassendi insists takes place in both humans and beasts, involves comparing apprehensions and associating them together according to their relations in actual experience. The faculties of judgment and reasoning put various apprehensions into an ordered sequence based on the experienced sequences of sensations, plus the natural instinct that makes us expect certain consequences to follow from what we have experienced.

Up to this point the detailed psychological theory that Gassendi presents is much like that later developed by the British empiricists from Locke to John Stuart Mill. But Gassendi also insists that there is another mental faculty that exists in humans, but not in other animals, that of intelligence or understanding, which belongs to our rational souls. By means of intelligence we are able to know things that cannot be experienced in sensation, such as God, space, and time. By this faculty we are also able to know the abstract essences of things, which transcend the powers of the imagination. Thus, for example, the imagination can know what "man" is, in terms of the sensations received. But, the essence of man, what it is that makes him what he is, can be known only by the intelligence. Lastly, this highest mental faculty is capable of self-consciousness. It can reflect on its operations and those of the imagination and make us aware that we see, we think, and so on.

In terms of this theory of the nature of the soul Gassendi next offers his opinion about the origin of our ideas. He repudiates completely the theory of Descartes and of Herbert of Cherbury that we possess innate ideas. Instead, Gassendi insists on the principle accepted by Aristotle and Epicurus, that there is nothing in the understanding that was not first in the senses. At the outset, the mind is a tabula rasa, a blank tablet. All the particular ideas that the mind ever knows, such as that of the sun, either come directly from sense experience or result from combinations of elements furnished by the senses. General or abstract ideas are formed by the intelligence from the collection of sense materials. In this case the sense information is necessary, but not sufficient to account for general ideas, such as that of "man." The intelligence goes beyond the actual sense-data in forming a unique idea from all the particular sensations. With regard to ideas of incorporeal things, which cannot be known by the senses, sense experience and the imagination furnish the occasion for the understanding to gain this knowledge. Because of certain experiences the understanding thinks, reflects, abstracts, and arrives at ideas, such as that of God. The senses provide some of the basic materials for these ideas and provide the context in which the understanding reasons to reach a conception of an incorporeal being.

Thus, all ideas either come from the senses or result from intellectual activities that are either caused or occasioned by sense information. However, in the cases of abstract ideas and ideas of incorporeal things, the actual content does not derive from any particular sense experiences. General principles, such as "The whole is greater than the part," are formed by induction from various particular experiences. When all of our experiences exhibit the same characteristics, we reach a general conclusion, which then becomes the basis of all further reasoning.

ETHICS AND RELIGION

The last part of *Syntagma* deals with ethics. Gassendi's theory is only a slightly modified version of Epicurus's hedonism. Gassendi holds that every pleasure, considered in itself, is a good and that all things that are considered good have value only in terms of the pleasure they produce. A completely pleasurable life is one without pains and troubles. Ultimately, for Gassendi, such a life can be achieved only by God. We can mitigate the pains in our lives as much as possible and thus attain a relatively good life.

A major problem in interpreting Gassendi's contribution is that of assessing his intentions and actual beliefs. There has been great debate whether Gassendi was really a Christian. He has been seen both as the founder of modern materialism, a leading skeptic and libertine, and as a serious Christian trying to find a *via media* between his faith and the new science. There have been long debates, especially in French literature, about the so-called Le cas Gassendi. He was a close associate of

some of the leading French freethinkers and took part in retreats with them where they boasted of being able to speak freely on all subjects. Gassendi was also a close friend of some leading church figures such as Mersenne. Gassendi and Mersenne shared similar views about science and its foundations. They agreed that science could not refute skepticism and each offered a form of mitigated skepticism as a way of carrying on useful science without metaphysics. Nobody ever questioned Mersenne's religious sincerity, and he remained in closest touch with Gassendi. No charges were ever made at the time about any heretical opinion or activity on the part of Gassendi.

Arguments about how to evaluate Gassendi still go on. Researches into his few theological writings go one direction, his materialism points another way, and his associations with leading figures of the time, ranging from Hobbes to Pascal, allow for many interpretations. Gassendi was a priest all his life and he was friendly with the most orthodox and the most unorthodox figures of his time. His philosophical system represented a cautious and careful attempt to explain the world in keeping with both the results of the new science and the official views of the Catholic Church. He may have seen, as few others of his time did, the importance of the values of his religious tradition, of the classical heritage, and of the new science, and at the same time fully appreciated what the skeptics had shown about man's fallible nature. Unlike Montaigne, Charron, and La Mothe Le Vayer (all of whom he admired), he did not wish to destroy the fruits of human efforts along with man's presumptuous and dubious claims. More like Sanches, he wanted to find a constructive resolution to the skeptical crisis of the Renaissance, but not in the form of the new dogmatisms of Herbert of Cherbury or Descartes. Living within a major religious tradition, he tried to show that by discarding Aristotelianism and by accepting the wisdom of the skeptics along with certain elements of Epicureanism, faith and the new scientific discoveries could coexist.

Gassendi adapted various features of the philosophy of Epicurus to the state of knowledge of his day, and he modified certain portions of Epicurus's theory that were not in keeping with the Christian religion. The result was a semiskeptical, semiempirical theory that portrayed the world in terms of an atomic structure. Gassendi's philosophy remained important throughout the seventeenth century and was the chief modern alternative to Descartes's. It began to lose its appeal and importance after the development of Newton's scientific theories. Many of the basic elements of later English philosophy appear in Gassendi's views, and he probably had great influence on such thinkers as Hobbes and Locke.

INFLUENCE AND SIGNIFICANCE

Gassendi was one of the foremost philosophers and scientists of the early seventeenth century. He was the most important rival and critic of Descartes, and he had a crucial role in the revival of the ideas of the ancient Greek skeptics and atomists. Gassendi began his intellectual career as a skeptic; a staunch follower of Sextus Empiricus and Montaigne. Gradually, he mitigated his skepticism in the face of the scientific revolution of the time, in which he played a major role, and he adopted more and more of a materialistic explanation of the world based on the ancient theory of Epicurus. Though a prominent Catholic priest of his day, Gassendi developed one of the first completely mechanistic and materialistic theories of modern times.

Gassendi's ideas had much influence in the seventeenth century. Although he published his work in huge Latin tomes, a French abridgement was made in the latter part of the century and many portions of his work appeared in English. His ideas were being taught in Jesuit schools in France, English universities, and even newly founded institutions in North America. Because of the close similarity between Gassendi's skeptical empiricism and some of the major portions of Locke's *Essay concerning Human Understanding*, there has been a good deal of discussion about whether Locke was influenced by Gassendi or used some of his works. It has been discovered that one of Gassendi's main works on Epicurus appeared in English in 1659 in Thomas Stanley's *History of Philosophy*, a work that Locke knew. Locke had also met a few of Gassendi's disciples in France, so it is possible that some influence occurred.

Gassendi's atomism never connected itself with basic scientific findings, so that modern atomism had to start elsewhere. There has been more interest in Gassendi in recent years. Many conferences were held for his 400th birthday in 1992, with explorations of many aspects of his thought and activities, and his scientific researches are proving important in the history of botany, geology, and other fields. There is growing interest in his critique of Cartesian philosophy and he is now being seen as a genuinely original thinker of the first rank.

See also Aristotelianism; Aristotle; Atomism; Bacon, Francis; Charron, Pierre; Common Consent Arguments for the Existence of God; Descartes, René; Epicureanism and the Epicurean School; Epicurus; Fludd,

Robert; Galileo Galilei; Herbert of Cherbury; Hobbes, Thomas; La Mothe Le Vayer, François de; Locke, John; Materialism; Mersenne, Marin; Mill, John Stuart; Montaigne, Michel Eyquem de; Newton, Isaac; Pascal, Blaise; Patrizi, Francesco; Pyrrho; Ramus, Peter; Renaissance; Sanches, Francisco; Sextus Empiricus; Skepticism, History of.

Bibliography

WORKS BY GASSENDI

Exercitationes Paradoxica Adversus Aristoteleos (1624). Amsterdam: Apud Ludovicum Elzevirium, 1649.

Petri Gassendi Diniensis ecclesiae praepositi … Opera Omnia. 6 vols. Lyon, France: Sumptibus Laurentii Anisson, & Ioan. Bapt. Deuenet, 1658.

Lettres familières à François Luillier pendant l'hiver, 1632–1633, edited by Bernard Rochot. Paris: J. Vrin, 1944.

Recherches métaphysiques. Translated and edited by Bernard Rochot. Paris: J. Vrin, 1962. Originally published as *Disquisitio Metaphysica* in 1644.

WORKS ABOUT GASSENDI

Bernier, François. *Abrégé de la philosophie de Gassendi*. 7 vols. Paris: Fayard, 1992. French translation of some of Gassendi's major texts.

Bloch, Olivier R. *La philosophie de Gassendi: Nominalisme, matérialisme, et métaphysique*. The Hague: Martinus Nijhoff, 1971. Careful study of Gassendi's philosophical development.

Centre International de synthèse. *Pierre Gassendi, 1592–1655, sa vie et son oeuvre*. Paris: A. Michel, 1955. Lectures on Gassendi by Bernard Rochot, Alexandre Koyré, Georges Montgrédien, and Antoine Adam. Some excellent discussions of his philosophical and scientific views.

Gregory, Tullio. *Scetticismo ed empirismo: Studio su Gassendi*. Bari, Italy: Laterza, 1961. Important study emphasizing Gassendi's materialism and skepticism.

Hess, Gerhard. "Pierre Gassendi, Der Französische Späthumanismus und das Problem von Wissen und Glauben." *Berliner Beiträge zur Romanischen Philologie* 9 (1939): 1–99. Important study of Gassendi's views on faith and reason.

Joy, Lynn Sumida. *Gassendi the Atomist: Advocate of History in an Age of Science*. New York: Cambridge University Press, 1987. Study of Gassendi's humanistic development of his scientific theories.

Kroll, Richard W. F. *The Material Word: Literate Culture in the Restoration and Early Eighteenth Century*. Baltimore, MD: Johns Hopkins University Press, 1991. The impact of Gassendi's ideas in Restoration England.

Murr, Sylvia, ed. *Gassendi et l'Europe*. Paris: J. Vrin, 1997. Essays by many scholars on the 400th anniversary of Gassendi's birth.

Osler, Margaret. *Divine Will and the Mechanical Philosophy: Gassendi and Descartes on Contingency and Necessity in the Created World*. New York: Cambridge University Press, 1994. Contrasts Descartes's philosophical theology with that of Gassendi's.

Pintard, René. *Le libertinage érudite dans la première moitié du XVIIe siècle*. Geneva, Switzerland: Slatkine, 2000. New edition of Pintard's 1943 work. Gassendi is examined in relation to the *libertins*, and an irreligious interpretation of his views is offered.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003. This is a newly revised and enlarged edition of Popkin's 1979 work, *The History of Scepticism: From Erasmus to Spinoza*. Depicts Gassendi's place in the history of skepticism.

Rochot, Bernard. "Le cas Gassendi." *Revue d'histoire littéraire de la France* 47 (1947): 289–313. Defense of Gassendi's religious sincerity against the accusations of Pintard.

Rochot, Bernard. "Gassendi et le Syntagma Philosophicum." *Revue de synthèse* 47 (1950): 67–79. Study of Gassendi's philosophy of science and scientific method.

Rochot, Bernard. *Les travaux de Gassendi sur Epicure et sur l'atomisme, 1619–1658*. Paris: J. Vrin, 1944. Important study of Gassendi's development. Contains a critical bibliography.

Turner, Anthony, and Nadine Gomez. *Pierre Gassendi: Explorateur des sciences*. Catalogue from the exhibit on Gassendi at the Musée de Digne, Digne-les-Bains, France, May 19–October 18, 1992. Portrays the many scientific interests Gassendi had.

Richard Popkin (1967, 2005)

GAUGE THEORY

Gauge theory is concerned with the problem of comparing physical states at different space-time locations. To get a feel for the problem, it is best to begin with a simple example. Quantum chromodynamics is the theory of the force that binds quarks together. An initial presentation of the theory might begin by stating that the *color force* (a fanciful name having nothing to do with visual colors) comes in three *color charges*—red, blue, and green, and their *anticharges*—anti-red, anti-blue, and anti-green. Every quark has one of these charges, and a stable collection of quarks must have no net color. Thus, a stable three-quark object, such as a proton, can be formed from a red, a blue, and a green quark (red + green + blue = white, which is colorless), or a stable two-quark object, such as a pion, can be formed from a red quark and an anti-red quark. This would explain why quarks are never seen in isolation. Just as electric charge comes in two forms, positive and negative, color charges come in six species. The interaction between any pair of quarks will depend on their charges.

Note an immediate consequence of the little story just told. It suggests that given any two arbitrarily specified quarks, no matter where they happen to be, there is a fact about what color charge they have, and, a fortiori, a

fact about whether their color charges are the same or different. Most metaphysical accounts of properties have the same consequence. If there is a universal corresponding to the color charge red, for example, then there is a fundamental fact about any pair of quarks whether they both instantiate this universal or not. Or if one prefers a theory of tropes, then there is still a fundamental metaphysical fact about whether the trope that is part of one quark is qualitatively identical to the trope that is part of the other. These metaphysical facts would obtain no matter where the two quarks were located even if they were located in different space-times.

Gauge theory rejects this metaphysics. It is correct, according to gauge theory, that there is a variety of color charges possible for each quark and that the structure of these physical possibilities is that of a genus with six species (speaking roughly). But there is no natural identification between the particular color states available to one quark and the particular color states available to a distantly located quark. Here an analogy will help.

Consider the surface of a sphere. At any point on the sphere, there is a set of directions one can move in. And the generic structure of that set at any one point is exactly the same as the structure of the set at any other point: Wherever you are, you have a full 360 degrees of different directions available. But there is no fact about whether a particular direction at one point on the sphere is *the same* or *different* from a particular direction at another point on the sphere. One could, for example, lay down a circle with degree markings from zero to 360 at the North Pole and lay down an identical circle somewhere on the equator. But there is no further sensible constraint that the two circles be *oriented the same way* or that the zero-degree direction at the North Pole *point the same way* as the zero-degree direction at the equator. One cannot sensibly ask of two arrows, one at the North Pole and the other at the equator, whether they point the same way or not.

Having placed the circle of degrees at the North Pole in one orientation, one is still free to place the circle at the equator at any orientation one likes. Such an arbitrary choice of orientation for the degree numbers is called *picking a gauge*, that is, fixing on a convention for assigning numbers to different directions at different points. Once one has picked a gauge, one can talk about an arrow at the North Pole and one at the equator pointing the same way (e.g., both pointing at thirty-seven degrees), but since the gauge itself was an arbitrary choice, the *sameness* carries no ontological weight.

When considering distant points on the sphere, it is obvious that there is no sameness or difference of direction: The set of directions one can go in at one point are, as it were, specifically different from the set of directions one can go at another. And one might then be tempted to simply index any direction by the point it is attached to: There are a set of North Pole directions and a set of Eiffel Tower directions, and so on, with none of these being intrinsically comparable to any other. But the situation is not so simple. Suppose a person is standing at the North Pole holding a rod out in a certain direction and is told to walk forward *keeping the rod pointed the same way*, that is, the person is to walk forward *without letting the rod twist*. This is a sensible demand, and a physically meaningful one: If the rod is allowed to twist as the person walks, the force will be felt in the hands. But *twisting* is just *changing direction*. So if there is a fact about whether the rod is twisting, there must be a fact about whether it is changing direction even though at every moment the person is located at a different point on the surface of the earth.

There is a nice mathematical object that handles this situation. The set of directions one can go in at any given point of the sphere is called its *tangent space*. The tangent spaces are all generically identical (360 degrees around) but specifically different: Each tangent space is *glued to* a point on the surface of the sphere. The mathematical object that will now be introduced is called a *connection* on the tangent spaces, and what it allows, intuitively, is for one to make comparisons between directions in the tangent space at one point with directions in tangents spaces at points infinitesimally nearby. So, as one moves continuously from one point to another on the sphere, the connection will determine whether the direction of the rod is changing or not. There are no absolute comparisons of *distant* directions, but there are comparisons of *nearby* ones mediated by the connection. More precisely, the connection provides a notion of *parallel transport*, that is, of *carrying* a direction from one tangent space to another without twisting *along a specified path*. It does not underwrite any absolute comparison of directions in different tangent spaces.

This is the sort of structure used in the gauge theories of physics. There are various charge states available to a quark at one location, and a similarly structured set of charge states available to a quark at another location, but no absolute comparison between the two: There is no fact about whether the states of the quarks are the same or different. How, then, can there be any *forces* associated with the color charges? In the case of the electric force, it is critical to know whether two particles have the same or different charge: Like charges repel and unlike charges attract. How can one say, as was said above, that a stable

collection of quarks must have no net color if there is no fact about exactly which color charge each quark has?

It is not enough to say that in a proton or a pion the quarks are nearby so there is a way of comparing their charges: Nearby is evidently not a mathematically precise term. The story is rather this. In modern particle theory, every force is mediated by a set of particles. The electromagnetic force is mediated by the photon, and the color force is mediated by particles called *gluons*. Furthermore, unlike the case of the photon, which carries no electric charge, the gluons themselves carry the color charge. And the very same remarks about the impossibility of absolute judgments about *which* charge a quark carries can be made about which charge a gluon carries. For heuristic purposes, it helps to think of gluons as carrying two charges: a color and an anticolor.

Now, suppose there is a bound state of two quarks, as in a pion. The quarks are only bound to each other through the mediating effect of a gluon. Originally, it was said that the pion as a whole must have no color, so the quarks can be, for example, one red and the other anti-red. But the gauge freedom, the freedom to identify different states at different points as the red state or the blue state means that there must be an equally valid description according to which the one quark is blue and the other anti-red. This seems to violate the demand that the pion have no net color.

Here, the gluon comes to the rescue. When gauge is changed, it must be done in a locally smooth way, and this means that not only do the color charges ascribed to the quarks change, but the color charges ascribed to the mediating gluon will change, too. So, while in one choice of gauge the pion will be described as a red quark bound to an anti-red quark by (say) a mediating blue/anti-blue gluon, in another choice of gauge, the very same pion will be described as a blue quark bound to an anti-red quark by a mediating red/anti-blue gluon. In each case there is no net color charge even though the particular charges ascribed to the constituents change. Evidently, while there is a *gauge freedom* involved in ascribing charges to particles—free enough so that any particle can be ascribed any charge—there are global constraints on the choice of gauge, and changing the gauge for one particle will have to have consequences for the charges ascribed to others. Because the gauge can be changed in different ways at different points, this is called a *local gauge freedom*.

The key point is this: The gauge freedom is wide enough that there is no objective, gauge-independent fact even about whether two particles have either same or different color charge. So no metaphysics that tries to associate the color charges with universals or tropes in the usual way can succeed. This result is of particular significance for David Armstrong's project of justifying belief in universals by appeal to scientific accounts of the world. The fundamental structures employed by the best scientific theories simply do not correspond to the ontology of substance/universal that Armstrong proposes.

Gauge theories provide a novel approach to the fundamental ontological problem of sameness and difference. A metaphysics of universals or tropes entails that there be certain absolute facts about whether two individuals have similar or different qualities or properties, facts that obtain independently of where the individuals are located or even whether they are located in the same space-time at all. According to gauge theories, comparisons of *properties* are always mediated by a gauge connection. This means both that comparisons between individuals that inhabit disconnected space-time cannot be made at all and that even within a single space-time, ascription of charges to individuals always requires a somewhat arbitrary global choice of a gauge. The gauge connection itself is objective, but the particular charges assigned to individuals are not. This is an ontological structure that does not fit neatly in any traditional metaphysical category.

See also Metaphysics; Relativity Theory.

Bibliography

Armstrong, David. *Universals and Scientific Realism*. Cambridge, U.K.: Cambridge University Press, 1978.

Moriyasu, K. *An Elementary Primer for Gauge Theory*. Singapore: World Scientific, 1983.

Tim Maudlin (2005)

GAUNILO

(fl. 11th century)

Soon after St. Anselm circulated his *Proslogion*, it was the target of a vigorous rejoinder by an otherwise unknown Benedictine monk named Gaunilo. Although Guanilo's "Reply on Behalf of the Fool" raises a number of objections to the ontological argument, by far the best known is the Lost Island *reductio*, an argument intended to be exactly parallel to Anselm's that generates an obviously absurd conclusion. Gaunilo proposes that instead of "that than which nothing greater can be thought" we consider "that island than which no greater can be thought" (2001, p. 31). We understand what that expression means, so

(following Anselm's reasoning in the ontological argument) the greatest conceivable island exists in our understanding. But (again following Anselm's reasoning) that island must exist in reality as well; for if it did not, we could imagine a greater island—namely, one that existed in reality—and the greatest conceivable island would not be the greatest conceivable island after all. Surely, though, it is absurd to suppose that the greatest conceivable island actually exists in reality.

In order to defend himself against Gaunilo's criticism, Anselm would have to show why Gaunilo's argument about the island is not in fact analogous to his own argument about that than which nothing greater can be thought. Yet although his "reply to Guanilo" asserts more than once that the island example fails, he does not explain why it fails. The usual reply given on Anselm's behalf (and indeed often attributed to Anselm himself) is that the notion of a greatest conceivable island is incoherent; however great an island might be, one could always conceive of a greater. (For a reading of the argument that endorses a response of this sort, see Klima 2000.)

Gaunilo's reply does have its defenders, however, most notably Nicholas Wolterstorff, who argues that Anselm "realized the 'tellingness' of [Gaunilo's] points. … The sign of his realization, however, is not concession; Anselm does not concede. The sign is rather bluster" (Wolterstorff 1993, 87).

See also Anselm, St.; Ontological Argument for the Existence of God.

Bibliography

Gaunilo. "Reply on Behalf of the Fool." In *Proslogion, with the Replies of Gaunilo and Anselm*. Translated by Thomas Williams, 27–33. Indianapolis: Hackett, 2001.

Klima, Gyula. "Saint Anselm's Proof: A Problem of Reference, Intentional Identity and Mutual Understanding." In *Medieval Philosophy and Modern Times*, edited by G. Hintikka, 60–88. Dordrecht, Netherlands: Kluwer, 2000.

Wolterstorff, Nicholas. "In Defense of Gaunilo's Defense of the Fool." In *Christian Perspectives on Religious Knowledge*, edited by Stephen Evans and Merold Westphal, 87–111. Grand Rapids, MI: Eerdmans, 1993.

Thomas Williams (2005)

GAY, JOHN
(1699–1745)

John Gay, the English moral philosopher, was a fellow of Sidney Sussex College, Cambridge, and later vicar of Wilshampstead, Bedfordshire. His short "Dissertation concerning the Fundamental Principle of Virtue or Morality" was first published as a preface to Edmund Law's translation of William King's Latin *Essay on the Origin of Evil* (1731). (Law was bishop of Carlisle and King was archbishop of Dublin.) The "Dissertation" is one of the seminal works in the history of English utilitarianism. In the eighteenth century its influence may be found in the works of the theological utilitarians, Abraham Tucker (*The Light of Nature Pursued,* 7 vols., 1768–1778) and William Paley (*Principles of Moral and Political Philosophy,* 1785). David Hartley said that Gay's assertion of the importance of psychological association in human nature was the origin of his *Observations on Man* (1749).

Gay hoped to eradicate confusion in moral philosophy and to harmonize the competing theories about the criterion of virtue. In his survey of candidates for the criterion of virtue, Gay noticed acting agreeably to nature; acting agreeably to reason; conformity to the fitness of things; conformity with truth; promoting the common good; and conformity to the will of God. In opposition to the claim that a criterion of virtue can be stated, Gay noticed the protagonists of the moral sense who claim that our judgments of virtue and vice are but the instinctive determinations of a moral sense. Gay set himself the task of showing that all of the above-mentioned criteria of virtue are compatible and not inconsistent with our having a moral sense.

Gay insisted upon the difference between a definition and a criterion, claiming that one must know what a thing is before one can measure it. Therefore, he first defined virtue as conformity to a rule of life. He expanded on the concept of "rule of life" by saying that it is a rule directing the actions of all rational creatures with respect to each other's happiness and that the rule must be understood to be obligatory for everyone in all cases.

Gay next turned to the question, What is it that can oblige everyone in all cases to follow a rule of life? He argued that a full and complete obligation can only arise from the authority of God, because only God can in all cases make a man happy or miserable. Gay then said that the criterion of virtue is the will of God. But what rule of life does God will that we follow? Attending to God's nature, we find him supremely happy. From God's goodness we infer that he has designed men to be happy and that he has willed the means to human happiness. Therefore, a person should always behave so as to be a means to the happiness of humankind. Arguing from the will of

God, Gay thus arrived at a criterion of virtue once removed.

The above account covers what might be called the first part of Gay's system. In it he found the clues for harmonizing the several criteria of virtue he had collected from earlier writers. He found conformity to the will of God to be the fundamental criterion of virtue, but the other criteria are necessary to explicate this one. Thus the criterion of the will of God with respect to virtue is whatever promotes the happiness of humankind or the common good. Gay defined things that are fitting and agreeable to nature as those things or actions which may be used to bring about the happiness of humankind. He complained about earlier writers who left the phrases "fitness of things" or "agreeableness to nature" empty of meaning by not seeing that they must be used in relation to some end, namely, the happiness of humankind.

To account for agreeableness to reason as a criterion of virtue, Gay included under his notion of reason not only reason—that is, the foreseeing of the inconveniences of certain things and actions by contemplating their natures—but also experience, or the perceiving of these inconveniences when they happen. Reason in this extended sense is the criterion of the fitness and unfitness of things and actions, as they contribute to human happiness. Gay added that when reason conforms to things as they really are, we say that we have the "reason" of things, or the "truth" of things. Thus, he fit in conformity with truth as yet another criterion of virtue. But while he succeeded in fitting all these criteria into an account of virtue, he also warned that some are more remote criteria than others.

Gay brought the moral sense into his account of virtue by denying that it is innate, or that it operates instinctively. Men must acquire the moral sense, notably by learning to be pleased by those actions which promote human happiness and to be displeased by those which do the contrary. Gay allowed that once it is learned, the operation of the moral sense may be habitual. He also allowed that much of humankind may learn what virtue is by example and observation, without being able to reason out their judgments.

Gay also explained why a person may be virtuous. Curiously enough, he made little of man's obligation to obey the will of God. Rather he appealed to the universality of man's inclination to seek pleasure and to avoid pain; and he equated a person's happiness with his being pleased. There are two motives, then, for virtuous behavior. First, when I see that my own happiness depends on the happiness of others, I will seek to promote their happiness in the hope that they will in turn promote mine. Second, since esteem and merit are associated with virtue, I may behave virtuously in order to enjoy the pleasure of being esteemed. Similarly, I will esteem those who promote my happiness, in order to encourage them.

See also Ethics, History of; Hartley, David; Paley, William; Virtue and Vice.

Bibliography

The "Dissertation" is reprinted in *British Moralists,* Vol. 2, edited by L. A. Selby-Bigge (Oxford: Clarendon Press, 1897). Gay also wrote a preface to Edmund Law's *Enquiry into the Idea of Space, Time, Etc.* (London, 1734). For critical discussion, see Ernest Albee, *A History of English Utilitarianism* (New York: Macmillan, 1902).

Elmer Sprague (1967)

GEHLEN, ARNOLD

(1904–1976)

The German social psychologist Arnold Gehlen was born in Leipzig. In 1934 he succeeded his teacher Hans Driesch as professor of philosophy at the University of Leipzig. He went to Königsberg in 1938 and from 1940 to 1944 was at the University of Vienna. In 1948 he became professor of sociology and psychology at the Hochschule für Verwaltungswissenschaften at Speyer. After 1962 he was at the Technische Hochschule in Aachen. He died in Hamburg.

Gehlen, a leading representative of the movement known as philosophical anthropology, sought to reinterpret the concepts of mind and intelligence in biological and sociological terms. His eclectic thought has partial affinities with the pragmatism of G. H. Mead and F. C. S. Schiller, with the integrationalism of Rudolf von Ihering, Maurice Hauriou, and Carl Schmitt, and with the cultural criticism of Oswald Spengler, Hans Freyer, and Martin Heidegger. At the same time, he rejects ontology and metaphysics. He rejects the traditional dualisms of soul and body, mind and matter, theory and practice. He emphasizes the predominant role of collective, or institutional, values as against those of individuals. He discards rationalism and regards present-day civilization as one of late-period decline.

METHOD AND TASK OF PHILOSOPHY

Gehlen rejects the experimental methods of the natural sciences as leading to materialism and rejects the "understanding" approach of the advocates of the *Geisteswis-*

senschaften, because it employs contemporary intellectual standards in the analysis of heterogeneous situations. The method of philosophy, Gehlen claims, is the intuitive or phenomenological method that he himself uses to interpret the significance of sociocultural institutions. According to Gehlen the task of philosophy differs from that of science. Disregarding the factual inferences of the sciences as irrelevant, philosophers should "unravel" (*freilegen*) the realities that are their proper concern. These realities, or "categories," are the basic qualities of man and of institutions that remain intact after the fullest cultural, social, and historical analyses. Gehlen conceives of such a study of reality as empirical and thus envisages no complete system of categories.

MAN'S NATURE AND POWER

Gehlen defines man as an "acting, anticipatory, nondetermined, self-delimiting being—a product of culture." Like other philosophical anthropologists, Gehlen views man, compared with other animals, as a vulnerable, deficient being, lacking the powerful instincts and natural weapons of survival of other animals. Man's fabled power of thought is an artificial substitute for his weak instincts. He is reduced to dependence on technical means for his survival. For survival and to liberate himself from anxiety he has had to develop tools and techniques including language, myth, and magic, and has had to create a common, habitual, and stable cultural environment.

This cultural environment is perpetuated in institutions, the historically evolved realities of state, family, law, economy, and so forth. To be "legitimate" an institution need not be useful but must be derived from man's nature as expressed in the cultic, nonutilitarian experiences of ecstasy, trance, and asceticism. Institutions are comprehensive and abstract structures that, through their principle of order, impart autonomy to the individuals participating in the collective *entente secrète*. The utility of social and cultural institutions is a secondary by-product of their development. Gehlen contrasts unreflective, spontaneous, self-sacrificing action, which he describes as noble (*vornehm und edel*), with self-interested and utilitarian action (including its sublimated forms in art, philosophy, and literature), which he designates as base (*gemein*).

THEORY OF TRUTH

Like certain pragmatists, Gehlen stressed action as the determinant of valid thought. While defining truth in terms of inner coherence and correspondence with facts, Gehlen also distinguished another aspect of truth, which he calls "inner truth." "Essentially irrational, non-scientific and not directly controllable experience has *its* truth: that is certainty. And it has *its* form of acting: non-experimental action based on tradition, instinct, habit and conviction" (*Der Mensch*, p. 330). These illogical, ethical certainties are valid without rational or experimental justification—as a matter of mere "appositeness" or inner sanity. Rational knowledge (*Wissenschaft*) cannot take over the function of the *idées directrices* of society that are the product of *Urphantasie*, the divinity and energy of the animal component of man.

PESSIMISM

Gehlen's analysis of his age was unrelievedly somber. His times, according to Gehlen, were marked both by the dissolution of institutions and a shift in individual and social consciousness from irrational certainty to an anarchic intellectualization. This change took place against a historical background in which organic agrarian society was giving way to organized industrial society. The cultural rupture transforms social organisms into "colonies of parasites" riddled with subjectivism, mechanization, a turn toward abstract and mathematical methods in art and science (desensualization), and experimental thinking.

Rising living standards, far from representing progress, create new urges for limitless satisfactions. Such changes lead away from ethical obligation deriving from man's nature to goal-directed efficiency deriving from man's method. These changes entail making the spiritual sphere political and robbing the political sphere of its religious aura. Since science is esoteric, the mass of the people are condemned to be primitive. The eclipse of the nation-state and the trend toward supranational organization and peace will leave a legacy of unresolved conflicts that may lead to a complete loss of individual freedom. Only two very unlikely circumstances could reverse the trend: an unexpected return to legitimate, nonrational values that are not amenable to conscious volition or the rise of a creative personality to provide a new kind of institutional leadership.

According to Gehlen, the philosopher's task in such a world situation is to point to signs of decline and to emphasize the "legitimate" elements in national heritages as expressed in the institutions of state, church, and law. Although present-day society is increasingly alienated from these heritages, they alone represent society's legitimate "reality." Reality has therefore to be sought in the archaic forms of the past.

See also Action; Driesch, Hans Adolf Eduard; Experimentation and Instrumentation; Geisteswissenschaften;

Heidegger, Martin; Mead, George Herbert; Philosophical Anthropology; Schiller, Ferdinand Canning Scott; Spengler, Oswald.

Bibliography

WORKS BY GEHLEN

Theorie der Willensfreiheit. Berlin: Junker und Dünnhaupt, 1933.

Der Staat und die Philosophie. Leipzig: F. Meiner, 1935.

Der Mensch, seine Natur und Stellung in der Welt. Berlin: Junker und Dünnhaupt, 1940; 6th ed., Bonn, 1958.

Urmensch und Spätkultur. Bonn: Athenaüm, 1956.

Die Seele im technischen Zeitalter. Hamburg, 1957.

Anthropologische Forschung. Hamburg: Rowohlt, 1961. A collection of articles. Bibliography, pp. 144–145.

Zeitschrift für philosophische Forschung 6 (1951–1952): 589–593. A bibliography.

Zeit-Bilder; zur Soziologie und Ästhetik der modernen Malerei. Frankfurt am Main: Athenäum Verlag, 1965.

Theorie der Willensfreiheit und frühe philosophische Schriften. Neuwied, Luchterhand, 1966.

Der Mensch; seine Natur und seine Stellung in der Welt. Frankfurt am Main: Äthenäum Verlag, 1966 .

Moral und Hypermoral; eine pluralistische Ethik. Frankfurt am Main: Athenäum Verlag, 1969.

Studien zur Anthropologie und Soziologie. Neuwied, Luchterhand, 1971.

Soziales und instrumentales Handeln; Probleme der Technologie bei Arnold Gehlen und Jürgen Habermas [von] Wilhelm Glaser. Stuttgart: W. Kohlhammer 1972.

Standorte im Zeitstrom: Festschrift f. Arnold Gehlen z. 70. Geburtstag am 29. Jan. 1974. Edited by Ernst Forsthoff and Reinhard Hörstel. Frankfurt: Athenäum-Verlag, 1974.

Adornos Philosophie in Grundbegriffen: Auflösung einiger Deutungsprobleme. Frankfurt am Main: Suhrkamp, 1974.

Urmensch und Spätkultur: philosophische Ergebnisse und Aussagen.Frankfurt am Main: Athenaion, 1975.

Gesamtausgabe. Frankfurt am Main: Klostermann, 1978.

Die Frage nach der Technik bei Arnold Gehlen und Martin Heidegger. Aachen: Fotodruck J. Mainz, 1978.

Man in the Age of Technology. New York: Columbia University Press, 1980.

Zeit-Bilder: zur Soziologie und Ästhetik der modernen Malerei. 3rd ed. Frankfurt am Main: Klostermann, 1986.

Man, His Nature and Place in the World. New York: Columbia University Press, 1988.

WORKS ON GEHLEN

Ballauf, Theodor. Review of *Der Mensch. Zeitschrift für philosophische Forschung* 6 (1951–1952): 566–589.

Haering, Theodor. "Zu Gehlens Anthropologie." *Zeitschrift für philosophische Forschung* 6 (1951–1952): 593–598.

Mahn, Anneliese. "Über die philosophische Anthropologie von Arnold Gehlen." *Zeitschrift für philosophische Forschung* 6 (1951–1952): 71–93.

H. O. Pappé (1967)
Bibliography updated by Michael J. Farmer (2005)

GEISTESWISSENSCHAFTEN

Geisteswissenschaften, a term commonly used in German to denote disciplines referred to as "the humanities" in English, emerged in the course of a nineteenth-centry discussion about the proper designation for those disciplines whose topics and methodologies were different from those of the newly predominant natural sciences (*Naturwissenschaften*) such as physics, biology, and chemistry. A compound word, its second component—"the word *Wissenschaften*" or "sciences"—indicates that these disciplines are indeed legitimate sciences, but sciences of a different kind than the natural sciences. The assumption underlying the discussion out of which the term emerged is that there are valid scientific methods for studying topics such as literature, art, and history, but that the objects of these disciplines and their appropriate methods were significantly different from the objects and quantitative methods appropriate of modern natural science. Originally conceived as one side of a binary opposition between the realms of "nature" and those things that could not be subsumed under that heading, it uses the term "*Geist*" to provide a positive description of the general domain that is the proper field of study for those disciplines.

Many scholars have noted that the plural form of the term was used in 1849 in J. Schiel's German translation of John Stuart Mill's *Logic* as a translation for Mill's phrase "the moral sciences" and count that as the origin of the term *Geisteswissenschaften*. However, earlier uses of similar terms have been documented (see, e.g., Diemer) and the term *Geist* as a central term for historical and cultural manifestations of human mentality had become common in German romantic philosophy (Herder and Hamann), and German Idealism (especially Hegel) well before 1850.

The clearest formulation of the notion of *Geisteswissenschaften* as a group of disciplines united by a common method was presented by Wilhelm Dilthey in his *Einleitung in die Geisteswissenschaften* (1883). In this work, he identifies the common topic of these sciences as historical social reality that cannot be captured through the natural sciences. They find their ultimate basis in the structures of human experience, which is essentially historically and contextually situated. Hence the *Geisteswissenschaften* seek to do more than merely to explain, instead they seek to understand the expressions of human experience by situating them into broader personal, social, and historical contexts that provide insight into their "sense." For Dilthey, then, the fundamental disci-

plines for all of these others were anthropology and psychology, with psychology understood as a descriptive science aimed at understanding the structures of human experience.

During the late nineteenth and the first half of the twentieth centuries these fields included not only what would traditionally fall under the concept of the humanities in English, such as philosophy, history, philology, and the histories of art and music, but also traditional faculties such as law and theology whose methodologies were not consistent with those of the natural sciences. It also came to include areas that were just beginning to emerge as special disciplines such as political science and sociology, which at the time were paradigms of the *Geisteswissenschaften*. Hence in the second half of the nineteenth century and the beginning of the twentieth century, the term *Geisteswissenschaften* was in competition with the term "*Kulturwissenschaften*" or "cultural sciences" as another way of capturing the difference between all of the fields that were distinct from natural sciences, a term that was championed above all by members of the Southwest German school of Neokantianism such as Wilhelm Windelband (1915) and Heinrich Rickert (1986) for all of the areas Dilthey called *Geisteswissenschaften*. They stressed the unique and specific nature of the objects that these (idiographic) sciences seek to understand as opposed to the general laws that were the object of the (nomothetic) natural sciences.

During the first half of the twentieth century, Dilthey's preferred terminology predominated, but toward the end of the twentieth century, the social sciences have come to be generally grouped together under the heading of "*Sozialwissenschaften*" or "*Gesellschaftswissenschaften*," two different German words for "social sciences," as they have increasingly adopted the quantitative methodologies associated with the natural sciences. From the outset, questions about the status of psychology as a scientific discipline have played a pivotal role in the discussion of the nature and limits of the *Geisteswissenschaften*. At the end of the twentieth century, however, it too was increasingly grouped together with social sciences based on shared quantitative research methods; the notion of the *Geisteswissenschaften* became closer once again to what in English would be called the humanities. Law, economics, and the social sciences are becoming less commonly subsumed under the heading of the *Geisteswissenschaften* and an increasing number of departments concentrating on history, literature, art, and related fields often choose to refer to themselves as *Kulturwissenschaften* instead of as *Geisteswissenschaften* as they combine methodologies from both the humanities and the social sciences into their studies.

See also Dilthey, Wilhelm; Hamann, Johann Georg; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Historicism; Idealism; Mill, John Stuart; Neo-Kantianism; Rickert, Heinrich; Windelband, Wilhelm.

Bibliography

Diemer, Alfons. "Geisteswissenschaften." In *Historisches Wörterbuch der Philosophie*, edited by Joachim Ritter. Vol. 3, 211–215. Basel/Stuttgart, Germany: Schwabe Verlag, 1974.

Dilthey, Wilhelm. *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften*. Gesammelte Schriften 7. Stuttgart, Germany: Teubner, 1927.

Dilthey, Wilhelm. *Einleitung in die Geisteswissenschaften*. Gesammelte Schriften 1. Stuttgart, Germany: Teubner, 1922.

Hoffmeister, Johannes. "Geisteswissenschaften." In *Wörterbuch der philosophischen Begriffe*. Hamburg: Meiner, 1955.

Rickert, Heinrich. *The Limits of Concept Formation in Natural Science: A Logical Introduction to the Historical Sciences*. Edited and translated by Guy Oakes. Cambridge, U.K.: Cambridge University Press, 1986.

Rothacker, Erich. *Einleitung in die Geisteswissenschaften*. Tübingen, Germany: Mohr, 1920.

Rothacker, Erich. *Logik und Systematik der Geisteswissenschaften*. München/Berlin: Oldenbourg, 1927.

Von Bollnow, Otto Friedrich. *Studien zur Hermeneutik*. Bd. 1, *Zur Philosophie der Geisteswissenschaften*. Freiburg: Alber, 1982.

Windelband, Wilhelm. *Geschichte und Naturwissenschaft. Rede zum Antritt des Rektorats der Kaiser-Wilhelms-Universität Straßburg, gehalten am 1. Mai 1894*. In *Praeludien. Aufsätze und Reden zur Philosophie und ihrer Geschichte*. Bd. 2, 136–160. Tübingen, Germany: Mohr, 1915.

Thomas Nenon (2005)

GENERAL WILL, THE

The idea of the general will (*volonté générale*) forms the core of Jean-Jacques Rousseau's political philosophy. Others had introduced the term before him, and his use influenced many others, including Immanuel Kant and Georg Wilhelm Friedrich Hegel, but the general will is most closely associated with Rousseau's *Social Contract* (1762/1997). In that work, Rousseau argued that "the general will alone can direct the forces of the State according to … the common good" (II.1.1, p. 57) and that political rule is only legitimate when based on a social contract that establishes the general will as sovereign. This led Rousseau to hold that laws must be authorized by the people as a whole, since "only the general will obligates particulars, and there can never be any assurance

that a particular will conforms to the general will until it has been submitted to the free suffrage of the people" (II.7.7, p. 70). The general will, as Rousseau understood it, is impartial in that it "must issue from all in order to apply to all" (II.4.5, p. 62).

Prior to Rousseau, the term "general will" was introduced into seventeenth-century theological disputes by Antoine Arnauld and then discussed by Blaise Pascal and Nicolas Malebranche, among others. The issue was whether God has a *general will* to grant all people salvation, and if so, how it is possible and just for particular individuals to be condemned to hell. In the early eighteenth century, authors such as Pierre Bayle and the Baron de Montesquieu began to use the term in a secular context. In defending the separation of governmental powers, Montesquieu associated the legislative function with the general will and judicial power with a particular will. When Denis Diderot published an entry on "natural law" in his *Encyclopédie* in 1755, the general will held a central place. He wrote that only humanity, and not any individual, can "determine the nature of justice and injustice. … Private wills are suspect; they may be either good or bad. But the general will is always good" (1755/1992, pp. 19–20). He continued, "The general will is in each person a pure expression of the understanding, which in the silence of the passions calculates what every individual may demand from his fellow-man, and what his fellow-man has a right to demand of him" (pp. 20–21).

While clearly influenced by Diderot, Rousseau rejected his colleague's cosmopolitanism and focused instead on the general will of a society. Rousseau held that "each individual may, as a man, have a particular will contrary to or different from the general will he has as a Citizen" (I.7.7, p. 52). A person's private will directs him toward his own particular interests, while the general will aims at the common good of society. In addition, Rousseau introduced the crucial contrast between the general will and the will of all: "From the preceding it follows that the general will is always upright and always tends to the public utility: but it does not follow from it that the people's deliberations are always equally upright. … There is often a considerable difference between the will of all and the general will" (II.3.1–2, pp. 59–60). A simple aggregation of private wills may generate the will of all, but the general will requires a mutual adjustment of interests in light of what individuals can reasonably demand of one another.

There is no infallible procedure by which to determine the general will. Rousseau argued that the general will can only act when all the people are gathered together in the "people's assembly" to vote on whether a proposed law "does or does not conform to the general will, which is theirs" (IV.2.8, p. 124). However, when their private wills distort their assessment of the common good, individuals may be mistaken about the content of the general will. It is even possible for the majority to be mistaken, and Rousseau was especially concerned about two sources of corruption, not to the general will itself, but to a society's ability to identify it. The first was the existence of factions, which Rousseau believed would lead individuals to elevate their shared private interests above the general will. The second was large inequalities in wealth, which could allow the wealthy to replace the judgment of the poor with their own: "No citizen [should] be so very rich that he can buy another, and none so poor that he is compelled to sell himself" (II.11.2, p. 78).

Rousseau held that outside of society, individuals have "natural freedom," since they need not limit their ability to act on their private wills. However, because private wills may conflict, individuals may still be dependent on the private wills of others and therefore lack freedom. It is only when a society is guided by the general will that individuals are freed from their dependence on private wills and are able to achieve "civic freedom." Their natural freedom is then limited, since they may no longer act on their private wills when these conflict with the general will. However, since others are similarly constrained, no one is dependent on anyone's private will.

For example, it is only under the general will that mere possession is transformed into property, with the result that no one may take what is not theirs. Furthermore, in a passage that strongly prefigured the work of Kant, Rousseau wrote that being freed from the dictates of one's own private will also represents a kind of moral freedom, "which alone makes man truly the master of himself; for the impulsion of mere appetite is slavery, and obedience to the law one has prescribed to oneself is freedom" (I.8.3, p. 54). Understanding that freedom involves independence from arbitrary private wills and that such dependency can only be avoided by the general will helps to explain Rousseau's comment "Whoever refuses to obey the general will shall be constrained to do so by the entire body: which means nothing other than that he shall be forced to be free" (I.7.8, p. 53). For Rousseau, this merely meant that individuals should be constrained in their unconditional pursuit of self-interest by principles of justice, which make them independent of anyone's private will.

Beginning with Hegel, but especially in the twentieth century, many critics saw in Rousseau the origins of the

Reign of Terror of the French Revolution or an endorsement of unconstrained majority rule. For example, in 1945 Bertrand Russell wrote that Rousseau was "the inventor of the political philosophy of pseudo-democratic dictatorships" and that "Hitler is an outcome of Rousseau" (pp. 684, 685). Such interpretations, because they neglect the contrast between the general will and the will of all, typically reveal more about the ideological fears and commitments of the commentators than about Rousseau. In contrast, the final decades of the twentieth century brought a revitalization of liberal political philosophy, much of it under the influence of John Rawls, and with it came a renewed interest in the general will. Rawls's project can be understood as an attempt to reconcile the two elements that Rousseau identified as the central commitments of the general will: "If one inquires into precisely what the greatest good of all consists in, which ought to be the end of every system of legislation, one will find that it comes down to these two principal objects, *freedom* and *equality*" (II.11.1, p. 78).

See also Freedom; Justice; Liberty; Political Philosophy, History of; Rousseau, Jean-Jacques.

Bibliography

Bertram, Christopher. *Rousseau and "The Social Contract."* London: Routledge, 2004.

Cohen, Joshua. "Reflections on Rousseau: Autonomy and Democracy." *Philosophy and Public Affairs* 16 (3) (1986): 275–297.

Diderot, Denis. *Political Writings* (1755). Translated by John Hope Mason and Robert Wokler. New York: Cambridge University Press, 1992.

Mandle, Jon. "Rousseauian Constructivism." *Journal of the History of Philosophy* 35 (4) (1997): 545–562.

Neuhouser, Frederick. "Freedom, Dependence, and the General Will." *Philosophical Review* 102 (3) (1993): 363–395.

Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press, 1999.

Riley, Patrick. *The General Will before Rousseau: The Transformation of the Divine into the Civic*. Princeton: Princeton University Press, 1986.

Rousseau, Jean-Jacques. *The Social Contract and Other Later Political Writings* (1762). Translated by Victor Gourevitch. New York: Cambridge University Press, 1997.

Russell, Bertrand. *A History of Western Philosophy*. New York: Simon and Schuster, 1945.

Shklar, Judith N. *Men and Citizens: A Study of Rousseau's Social Theory*. New York: Cambridge University Press, 1969.

Sreenivasan, Gopal. "What Is the General Will?" *Philosophical Review* 109 (4) (2000): 545–581.

Jon Mandle (2005)

GENERICS

Generics are noun phrases (NPs) and sentences of certain types; the phenomena exhibited by these NPs and sentences are known as "genericity." Two rather different phenomena are embraced by this term, and they both are of interest to philosophers.

THE FIRST PHENOMENON: REFERENCE TO A GENUS (OR KIND)

An example of the first phenomenon, reference to a genus, is the sentence *The black-capped chickadee winters in central Alberta*, which refers to the genus, or kind, *The Black-Capped Chickadee*. The sentence may also do other things, such as make claims concerning individual black-capped chickadees and the things they do. But the way it accomplishes these other tasks is to employ its NP subject term to refer to the kind and then make a predication about this kind. There are various tests that one might employ to show that NPs like this really do refer to kinds. For instance, note that predicates like *is (not) extinct* are true only of kinds and not of individual instances of a kind. It makes no sense to assert that *Tweety is extinct* (as opposed to being dead). But it does make sense to say *The black-capped chickadee is not extinct*, thereby showing that in this sentence at least, *the black-capped chickadee* refers to a genus. Of course, not every occurrence of *the black-capped chickadee* refers to a kind. For example, it does not do so in *The black-capped chickadee in the far cage needs more seed*; here *the black-capped chickadee* refers to an individual instance.

A fundamental question concerning this type of genericity is the following: What types of expressions can refer to genera? As the previous example shows, definite NPs can do this in certain sentences. And since the sentence *Black-capped chickadees winter in central Alberta* has the same force as the previous example, most theorists take these "bare plural" NPs also to refer to kinds, at least in this sort of sentence. Another type of NP that is of the same nature contains mass terms ("bare singular" NPs they are sometimes called) such as *gold, furniture*, and *information*. The subjects of sentences like *Gold is a yellowish metal* refer to a kind. There is at least one case in English of a bare singular count NP that designates a kind: In *Man evolved from the great apes*, the NP *man* (as opposed to the common noun *man*, as it occurs in … *is a man*) refers to a kind. And there are some proper names of kinds also, such as Ursa arctos horribilis *is common in the mountains of Alberta*. On the other hand, indefinite NPs do not refer to kinds (with an exception to be men-

tioned just below): *A grizzly bear is common in the mountains of Alberta* seems nonsensical; the indefinite NP in *A black-capped chickadee winters in central Alberta* refers to some individual instance of the kind. The same is true for quantified NPs: The subject NPs in *Every/Most/Some/All/Each/Few black-capped chickadee(s) winter(s) in central Alberta* quantify over individual instances of the kind, and do not designate the kind. (Actually, there can be reference to kinds using indefinite NPs and quantified NPs, but then these NPs are given a *taxonomic* interpretation. We can say *All dinosaurs are extinct*, meaning thereby that every species of dinosaur is extinct; similarly, we can say *A whale has been labeled as endangered*, meaning thereby that a species of whale, perhaps the Blue Whale, has been labeled as endangered.)

Another fundamental question concerning this type of genericity is the following: What are the truth makers for such sentences? Some of these predications seem clearly to predicate a property *directly* of a kind—in cases such as *The dodo is extinct*. But consider the (true) generic sentence *Man landed on the moon in 1969*. In such cases the truth maker would seem to be the initial person who satisfies the predicate; then this property is attributed or projected to the kind. But of course not every property that is true of an individual person becomes true of mankind. For this type of *indirect reference to a kind*, it seems that the property in question must be "important" enough: For sentences like *Man pole-vaulted 6 meters in 1985* do not seem true, even though Sergey Bubka of Ukraine did so in Paris in 1985 (and he was the first person to have done so).

We can also sometimes use an individual exemplar of a kind as the truth-maker for things we (or other agents) do, as in *We photographed the grizzly in Alberta last summer*, when in fact it was only a few of the instances of *Ursa arctos horribilis* that were photographed. These and other types of indirect reference to kinds are discussed by Krifka et al. (1995). The fact that predicates that are primarily true of ordinary individuals are somehow projected to be true of kinds raises questions of both a logical nature (about the resulting "type mismatch") and a metaphysical nature (about the relation between kinds and their exemplars).

Yet a further fundamental question concerns what kinds there are. The examples thus far surveyed have been of "natural kinds," but clearly there are kinds of artifacts: *Schockley invented the transistor in 1957* employs *the transistor* as designating a kind. And *The Coke bottle has a narrow neck* employs *The Coke bottle* in this way also; yet *The green bottle has a narrow neck* seems not so much false as nonsensical, unless *the green bottle* is taken to designate a particular bottle. Intuitively, there just is no such kind as The Green Bottle. Of course, with sufficient background contextual buildup one can make *The green bottle* be, for instance, the salvation of all those stricken with some new disease. Considerations like these have suggested to some that the notion of kind that is relevant to genericity in this first sense is "conventional" or "social."

THE SECOND PHENOMENON: GENERIC CHARACTERIZATION

An example of the second phenomenon, generic characterization, is the sentence *Lions have manes*, which predicates the property of having a mane "generically" to lions. By this it is meant that it is *generally* true (plus some qualifications to be discussed below) that lions have manes. As we know, only male lions have manes, so this predication is not universally true of lions. This feature is usually described by saying that generic characterizations *allow for exceptions* while they nonetheless remain true. It is this feature of genericity that has aroused the interest of logically oriented philosophers of language; for, given this portrayal of generic characterization, some radically new logical techniques will be required in order to employ these sentences in arguments.

Note that this second notion of genericity is a feature of entire sentences, whereas the first notion was a feature of NPs. But to complicate matters, the two phenomena can occur together, as in *The rutabaga contains vitamins A and C*, where *The rutabaga* exemplifies genericity of the first sort and the sentence as a whole exemplifies genericity of the second sort. (There might be some rutabagas that are missing either vitamin A or C, yet the original sentence would be true.)

Generic characterizing sentences express regularities about specimens of a kind: some regularities concern properties that are exemplified by the typical member (such as in the rutabaga sentence), whereas others express regularities of action that an object engages in (such as in "habitual" sentences like *Mary plays tennis after lunch*, which again is true despite the existence of days where Mary must work after lunch). It is this ability to express regularities in the face of exceptions that explains why all languages allow the expression of generic characterization. People notice regularities in nature and form "folk laws" to codify these regularities and predict what the future might bring. Despite the existence of exceptions, they are intellectually satisfying and practically useful because the objects *typically* or *usually* or *normally* or

nomically perform those actions. And such regularities commonly have exceptions.

Most writers in the genericity literature have argued that it is wrong to view these characterizing generic sentences as "really false but acceptable despite the exceptions because they are close enough to being universally true." For, they claim, most of our knowledge of the world is encoded in these generic sentences, so this is not a useful attitude. And if it were correct, then we would expect that sentences with fewer exceptions are more acceptable. But this is not borne out by examples, as we will see shortly.

These writers also have tended to shun the view that generics are *neither* true nor false but are instead directions or rules. For, this would make most of our knowledge become neither true nor false but instead directions to guide our belief formation ability. Further, since generics would not have a truth value, they could not be embedded inside propositional attitudes or joined into longer generic statements. But *John knows that rutabagas contain vitamin A* and *It is common that countries that do not honor women's rights also do not honor general human rights* are in fact either true or false.

Consider this list of (true) characterizing generic sentences:

1. Snakes are reptiles.
2. Telephone books are thick books.
3. Guppies give live birth.
4. Italians are good skiers.
5. Crocodiles live to an old age.
6. Frenchmen eat horsemeat.
7. Unicorns have one horn.

Obviously these call for different proportions of the subject terms satisfying the predicate. In (1) it is all; in (2) most; in (3) some subset of the females; in (4) some small percentage, but a greater percentage than in other countries (or maybe the very best of the Italian skiers are often better than the very best from other countries); in (5) it is strikingly few, since of the hundreds born to one female at a time, most are eaten within a few weeks of birth; in (6) there need be only a very small percentage—somehow the culturally determined views of North America make it striking that it happens at all; and in (7) no unicorns have one horn. Such examples show that there is no univocal quantifier that will serve in all characterizing sentences.

Even attempts to employ vague, probabilistic quantifiers such as *most* or *generally* or *in a significant number of cases* are misguided. Consider such false characterizing sentences as the following:

8. Leukemia patients are children
9. Prime numbers are odd.

These false sentences would become true if prefixed by *In a significant number of cases*. Indeed, the actual number of cases has nothing to do with the truth of the characterizing generic sentence as opposed to what evidence we might have for the sentence's truth. We might use the preponderance of thick telephone books in the world as our evidence for the truth of (2); but we will be prepared to retract it when we discover other relevant background facts, as perhaps happens with (9). This is often put as "characterizing sentences are inherently intensional." A sentence like *Members of this club help one another in emergencies* can be true despite there never having been any emergency. What is required for its truth is intensional: the preparedness to act in certain ways in certain situations. This intensionality is often claimed to thwart any attempt to use an extensional quantifier in the analysis of characterizing generics.

TWO RELATED AREAS

The feature of allowing for exceptions while nonetheless remaining true raises interesting issues in logic. This general topic is called *nonmonotonic reasoning* in the artificial-intelligence literature. Although these researchers do not explicitly aim to provide a semantics for characterizing generic sentences, nonetheless results of their research might be pressed into service for this purpose (Pelletier and Asher, 1997).

The issue of how children can learn that some occurrences of NPs are universal, others existential, and still others generic has been investigated both from the point of view of English-speaking children learning language (Hollinger et al. 2002) and comparatively between English-speaking children and Mandarin-speaking children (Gelman and Tardif 1998). It seems that children learn the difference by the age of four and that there is a difference in the frequency of generic vs. nongeneric NPs and sentences encountered by English and Mandarin children.

PHILOSOPHICAL CONSIDERATIONS

Genericity has been seen by some as requiring an ontology of (abstract?) kinds, individuals, and (momentary) stages of individuals. Furthermore, it seems to some that the semantics of generics presumes that these kinds are

conventional in some way. And the issues surrounding natural laws that admit of exceptions has been seen by some as endorsing a kind of scientific antirealism. All of these considerations raise deep questions of the relation between semantic models of natural language and reality. Some have claimed that the relationship should be modest and that these ontological conclusions pertain only to the metaphysical presuppositions of natural language—a natural-language metaphysics, in the phrase of Bach (1986)—but not necessarily to reality. This is the deepest issue in philosophy of language.

See also Conventionalism; Non-Monotonic Logic; Plurals and Plurality; Propositional Attitudes: Issues in Semantics; Realism.

Bibliography

Bach, E. "Natural Language Metaphysics." In *Logic, Methodology, and Philosophy of Science, VII: Proceedings of the Seventh International Congress of Logic Methodology and Philosophy of Science, Salzburg, 1983*, edited by R. Marcus, G. Dorn, and P. Weingartner, 63–81. Amsterdam: North-Holland, 1986.

Carlson, G. *Reference to Kinds in English* New York: Garland Press, 1989.

Carlson, Gregory. "Generics and Atemporal 'When.'" *Linguistics and Philosophy* 3 (1979): 49–98.

Chierchia, Gennaro. "Reference to Kinds Across Languages." *Natural Language Semantics* 6 (1998): 339–405.

Cohen, Ariel. *Think Generic! The Meaning and Use of Generic Sentences*. Stanford: CSLI Press, 1999.

Gelman, Susan, and Twila Tardif. "A Cross-Linguistic Comparison of Generic Noun Phrases in English and Mandarin." *Cognition* 66 (1998): 215–248.

Heyer, Gerhard. "Semantics and Knowledge Representation in the Analysis of Generic Descriptions." *Journal of Semantics* 7 (1990): 93–110.

Hollinger, Michelle, Susan Gelman, and Jon Star. "Children's Interpretation of Generic Noun Phrases." *Developmental Psychology* 38 (2002): 883–894.

Koslicki, Kathrin. "Genericity and Logical Form." *Mind and Language* 14 (1999): 441–467.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia, Gennaro, and Godehard Link. "Genericity: An Introduction." In *The Generic Book*, edited by Gregory Carlson and Francis Jeffry Pelletier, 1–124. Chicago: University of Chicago Press, 1995.

Pelletier, Francis Jeffry, and Nicholas Asher. "Generics and Defaults." In *Handbook of Logic and Language*, edited by Johan van Benthem and Alice ter Meulen, 1,125–1,177. Amsterdam: Elsevier, 1997.

Schubert, Lenhart, and Francis Jeffry Pelletier. "Problems in the Representation of the Logical Form of Generics, Bare Plurals, and Mass Terms." In *New Directions in Semantics*, edited by Ernest Lepore, 385–451. New York: Academic Press, 1987.

Francis Jeffry Pelletier (2005)

GENETICS AND REPRODUCTIVE TECHNOLOGIES

Modern genetics and technological aids to human reproduction, like other advances in science and technology, have created ethical problems heretofore unencountered. Biomedical developments have also posed new conceptual, epistemological, and metaphysical problems. This entry addresses these philosophical concerns as well as the more widely discussed ethical implications of contemporary genetics and reproductive technologies. One conceptual and ethical link between these two fields is the prospect of "designing our descendants." This prospect has been viewed by some as a boon to humankind (Fletcher 1974) and by others as a fearsome possibility to be avoided at all costs (Ramsey 1970).

The Human Genome Initiative, a "big science" project launched by the U.S. government to map and sequence the entire human genome, has heightened concerns about the privacy and confidentiality of genetic information, the uses to which such information might be put, and the possibility of stigmatizing individuals or groups because of their genetic constitution. The knowledge the Human Genome Project can yield is massive in contrast to previous efforts to acquire information about human genetics.

The contemporary science of genetics provides, not only an understanding of heritable traits, but also the capability to diagnose the probability or certainty of transmitting to offspring genetic conditions such as sickle-cell anemia. Tay-Sachs disease, or cystic fibrosis. The ability to identify and locate specific genes that render a person likely to manifest heritable conditions, such as Huntington's disease and certain forms of cancer, raises profound questions about the wisdom and desirability of learning about future contingencies when no cure exists and preventive measures are of uncertain efficacy.

A conceptual question is prompted by the rapid advances in genetics: What constitutes genetic disease? The traditional concept of disease relies on the ability of medical scientists to identify deviations from the normal physiological functioning of an organism. Asymptomatic diseases, such as hypertension, can be detected by diagnostic instruments even though the individual feels no symptoms of illness. With the discovery of genes that render an individual with a family history highly likely to develop a particular disease later in life, how should the

individual who carries the gene be characterized? Does the person in whom the gene is found have a genetic disease or not? The individual has no symptoms and the disease may never express itself. Yet merely being susceptible opens the possibility of harm to the interests of such individuals, making them vulnerable to actions by others such as insurance companies who seek to deny insurance on grounds of a preexisting condition or employers who refuse to hire workers with a known propensity for illness.

Beyond the problems posed by diagnosis and prediction in genetics are those of intervention: Is gene therapy intrinsically different from traditional medical therapy? Even if gene therapy by means of manipulating somatic cells poses no special problem, what about altering germline cells, a procedure that would affect future generations? If genetic manipulation to correct defects is ethically permissible, what, if anything, would be wrong with alterations intended to provide genetic enhancement? Are efforts to improve human intelligence, appearance, or other attributes by genetic means essentially different from the traditional methods of education, physical or mental training, or behavior modification (President's Commission for the Study of Ethical Problems 1982)?

Attempts to improve the quality of the human gene pool, or "positive eugenics," have generally been viewed with disfavor, especially after the policies in Nazi Germany promoting racial hygiene (Proctor 1988). Yet eugenic practices remain at the level of individual choice. The recipients of donated sperm are typically given information about physical and other personal characteristics of donors, allowing them to choose sperm from a donor whose traits they hope to replicate in the child. The prospect of genetic enhancement using the techniques of recombinant DNA manipulation can allow for more precision and wider applications than older approaches such as selective sperm banking.

Knowledge that one carries a gene for a heritable disease can pose a profound dilemma for the individual. An early form of this dilemma arose when carrier screening was the only way to determine whether a couple would pass on a genetic disease to their offspring. A couple then had to decide whether to take the chance that a child would be born with the heritable condition. With the advent of various forms of prenatal diagnosis (amniocentesis, chorionic villus sampling, blood tests), the presence of some genetic diseases in a fetus can be detected. The ethical question in such cases is whether to abort an afflicted fetus. In the case of both carrier screening and prenatal diagnosis, trained genetics counselors have uniformly taken a nondirective approach. The norm in genetics counseling has generally been to provide unbiased information to enable individuals or couples to make an informed decision whether to initiate a pregnancy or to abort a fetus found to have a genetic disease (Lappe 1971, President's Commission for the Study of Ethical Problems 1983).

As the science of genetics yields an increasing amount of information, individuals are faced with making decisions about prophylactic medical interventions. For example, a woman who learns that she carries a gene for an inherited form of breast cancer may contemplate bilateral mastectomy before any clinical signs appear. The epistemological problem posed by such scenarios is a familiar philosophical one: decision making under risk and uncertainty. If the woman decides to undergo a major, disfiguring operation, she does so with the knowledge that she might escape the disease entirely. But if she forgoes the preventive step, she runs the risk of developing a dread disease that may be curable if detected early but that also has a high mortality rate.

The knowledge by individuals or couples that they are at risk for transmitting a genetic disease to offspring is one indication for embarking on the use of reproductive technologies. The couple may elect to use donated sperm or ova. A far more common indication for the use of reproductive technologies, however, is infertility or subfertility on the part of one or both members of a couple. Methods include in vitro fertilization (IVF)—fertilizing a human ovum outside the womb—the use of sperm or ova contributed by third parties or the womb of a woman not intended to be the rearing parent (surrogacy); cryopreservation (freezing) of fertilized ova, which are termed *preembryos*; and embryo splitting.

Frequently discussed ethical issues include concerns about destruction of the traditional family when third parties are used as gamete donors or surrogates (Macklin 1991); worries about the effect on children who learn that they were born as a result of these techniques; and the opposite worry about harmful effects of struggling to maintain family secrets. Prior to the first IVF birth in 1978, fears were expressed that IVF would produce a higher than normal incidence of birth defects, but scientific evidence gathered over the years has shown this concern to be unwarranted. The objection that being created with the aid of gametes from a third party can harm the interests of children is countered by the metaphysical observation that these are children who would never have existed but for the use of these techniques.

Different religions are opposed to the use of some or all of these reproductive technologies. The Roman Catholic Church has urged prohibition of virtually all forms of assisted reproduction (Congregation for the Doctrine of the Faith 1987). The church's opposition is based on the fact that these techniques separate the procreative and unitive functions of marriage. Some authorities in Orthodox Judaism allow insemination from non-Jewish sperm donors but prohibit donation from Jews, in order to prevent consanguinity; others oppose all third-party donations out of fear of consanguinity and also by analogy with adultery. Islamic law prohibits the use of sperm or eggs from anyone other than the married couple on grounds that the results are similar to adultery (Serour 1992). Since the identity of gamete donors is normally kept confidential, a secular concern is that a brother and sister may unwittingly mate or marry, unaware that they have a genetic parent in common.

Possibly the most intriguing philosophical issues posed by reproductive technologies are those that arise from the newfound ability to separate the genetic from the gestational procreative functions. IVF permits an ovum from one woman to be fertilized and the resulting embryo implanted in a different woman. This creates the entirely novel situation of two different "mothers": the genetic mother, who supplies the egg; and the gestational mother, who undergoes pregnancy and childbirth. Apart from the emotional or other psychological consequences that may result from such arrangements, the separation of the woman's procreative role into two distinct biological functions requires a conceptual decision of whether the individual who performs each function properly deserves the appellation "mother" (Macklin 1991).

A variation on this conceptual theme stems from research that demonstrates the capability of transplanting ovaries from an aborted fetus into an adult woman who lacks ovaries of her own. The woman into whom the ovaries are transplanted is a mother in the traditional sense of one who is pregnant and gives birth to the child. Is it appropriate to construe the aborted fetus as the "genetic mother"? The conceptual oddity of this construal suggests that "mother" is a concept laden with connotations that do not permit its expansion to include aborted fetuses. Although the aborted fetus is without question the source of the genetic material from which the new life was created, it is semantically odd to conclude that the aborted fetus is the genetic mother.

A persistent quandary relates to the status of extracorporeal embryos. The product of IVF is termed a *preembryo,* partly because of its early developmental stage but also because it is unimplanted. The ability to freeze embryos indefinitely and thaw them for use later poses both conceptual and ethical questions. When disputes arise concerning the ownership of embryos, should the embryos be construed as "people" or as "property" (Annas 1989, Robertson 1990)? Should anyone other than the couple who contributed the gametes have the authority to destroy frozen embryos? If it is permissible to destroy embryos that are not intended for implantation, is it permissible to do experiments on the embryos? Controversy exists over the splitting of embryos, a technique sometimes called cloning (Robertson 1994). One objection holds that such deliberate duplication destroys genetic individuality and thus devalues the uniqueness of each individual.

Genetics and reproductive technologies pose new philosophical questions about the scope and limits of such familiar concepts as disease, individuality, parent, mother, and the family. The importance accorded to human reproduction and lineage throughout history is a reminder that such questions are not merely abstract concerns of philosophers but deeply rooted in the lives of individuals and communities.

See also Abortion; Bioethics; Distant Peoples and Future Generations; Evolutionary Theory; Human Genome Project; Informed Consent.

Bibliography

Annas, G. J. "A French Homunculus in a Tennessee Court." *Hastings Center Report* 19 (1989): 20–22.

Congregation for the Doctrine of the Faith. *Instruction on Respect for Human Life in Its Origin and on the Dignity of Procreation.* Vatican City, 1987.

Fletcher, J. *The Ethics of Genetic Control: Ending Reproductive Roulette.* Garden City, NY: Anchor Press, 1974.

Lappe, M. "The Genetic Counselor: Responsible to Whom?" *Hastings Center Report* 2 (1971): 6–8.

Macklin, R. "Artificial Means of Reproduction and Our Understanding of the Family." *Hastings Center Report* 21 (1991): 5–11.

Milunsky, A., and G. J. Annas, eds. *Genetics and the Law III.* New York: Plenum Press, 1985.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. *Screening and Counseling for Genetic Conditions.* Washington, DC: President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1983.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. *Splicing Life.* Washington, DC: President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1982.

Proctor, R. N. *Racial Hygiene: Medicine under the Nazis.* Cambridge, MA: Harvard University Press, 1988.

Ramsey, P. *Fabricated Man: The Ethics of Genetic Control.* New Haven, CT: Yale University Press, 1970.

Robertson, J. A. "In the Beginning: The Legal Status of Early Embryos." *Virginia Law Review* 76 (1990): 437–517.

Robertson, J. A. "The Question of Human Cloning." *Hasting Center Report* 24 (1994): 6–14.

Serour, G. I. "Medically Assisted Conception: Islamic Views." In *Proceedings of the First International Conference on Bioethics in Human Reproduction Research in the Muslim World,* edited by G. I. Serour. Cairo, 1992.

Ruth Macklin (1996)

GENETICS AND REPRODUCTIVE TECHNOLOGIES [ADDENDUM]

Philosophical scholarship on genetics and reproductive technologies typically follows the development of these scientific fields closely, providing critical analysis of the major assumptions and implications of their emerging claims, often in advance of their realization. This addendum reviews three discussions that have become particularly prominent in the literature since Ruth Macklin's original entry appeared in 1996: debates over the design and interpretation of human genetic variation research, the prospect of human reproductive cloning, and the potential limits of human genetic modification.

HUMAN GENETIC VARIATION RESEARCH

The principal outcome of the Human Genome Project was a set of research tools for human genetic variation research. With improved genomic maps and DNA sequencing technologies, geneticists have been able to launch a new generation of projects comparing human genomes to better understand our similarities, differences, and patterns of relationship at the molecular level. These comparisons are critical to the development of successful medical applications of genomic research, as well as to the interests of anthropologists and paleontologists interested in the evolution, differentiation, and global migrations of our species (Risch et. al. 2002). On the other hand, these comparisons also raise two sets of important philosophical issues:

First, how should scientists define and identify the relevant comparison groups within our species? The initial attempt to use genomic tools in a large scale study of human variation, the so-called Human Genome Diversity Project, followed the accepted practice of physical anthropologists and epidemiologists of describing its target groups in ethnic, linguistic, and geographical terms, and was called to task by both biologists and social scientists for using socially constructed categories that would obfuscate rather than illuminate underlying patterns of gene flow within our species (Gannet 2001, Reardon 2005). Rather than reifying various human political histories by looking for "ethnic-affiliation markers" in human DNA, some suggested a random global sampling strategy blinded to social identifiers (National Research Council 1998). The U.S. National Institutes of Health followed this approach in developing a major genetic variation research resource—a databank of known single nucleotide variants in human DNA—and was in turn called to task by public health and pharmacogenomic researchers for omitting "phenotypic data" about the distribution of the DNA variants across different populations (Altshuler and Clark 2005).

As a result, the subsequent international effort to a variation-measuring "haplotype map" of the human genome intentionally collected samples from groups defined by their "continents of origin" (International Hapmap Consortium 2003). Critics charge that this strategy returns population genomics to a set of outmoded racial categories that human scientists of all stripes have repudiated as biomedically meaningless and socially pernicious (Duster 2005). Claims that, nevertheless, research framed in this way has identified patterns of genetic variation that cluster along racial lines, and that these variations may be the key to "population specific" public health interventions or even "race-based medicine," have only lent fuel to this conceptual debate.

The second issue follows from the first. Assuming that, for the foreseeable future, the definition of comparison groups in population genomics will be informed by socially constructed criteria at some level of resolution (either familial, tribal, ethnic, racial, or regional), how should the interests of group members be protected? Outside of groups with clear political sovereignty, like Native American nations, most targets of genetic analysis have ambiguous moral standing. Is it ethically important for scientists to attempt to discuss their plans with groups at the collective level before recruiting individual group members into genetic variation studies? Some argue strongly that a principle of "respect for community" needs to supplement our traditionally individualistic principles of research ethics in these contexts, if only

because individuals gain so much of their identity through their community memberships and their genetic lineages (Weijer 1999). Others argue that, at least for genetic studies, extensive efforts at community engagement are disingenuous and guaranteed to fail, given the mismatch between genetic populations and the politically defined communities available for consultation (Juengst 1998).

HUMAN REPRODUCTIVE CLONING

Philosophical questions about the nature of human identity have been raised from quite another angle, meanwhile, by the successful cloning of a sheep in 1997 (McGee 2000). In that case, the nucleus of a mammary gland cell from an adult ewe was transplanted into an enucleated egg, and, after 280 attempts, coaxed into developing into a genomic twin of the gland cell's donor. If the technique that produced Dolly, the ewe, could be used effectively with humans, it would open up the prospect of adding reproductive cloning to the repertoire of techniques available to those seeking to procreative assistance. This prospect has prompted questions about the personal identity and moral status of the resulting "delayed twin," but these seem easily answered by our experience with natural twins who also have identical genomes: Clearly, the cloned individual would be a distinct person with full moral standing, due full protection against exploitation or abuse by his or her progenitors.

However, anticipating parental expectations for cloned offspring raises more difficult questions. How similar might the cloned offspring be to the progenitor, and what impact should the foreknowledge (or assumption) of such similarities have on the rearing of the offspring? Should progenitors of clones assume special responsibilities to anticipate the health and behavioral challenges their delayed twins may face, or do they, instead, acquire unusual obligations to refrain as much as possible from prejudicing the life experience of their offspring? These questions are animating new work on the nature and ethics of parenting, procreative liberty, and the limits of genetic determinism. They have also given new energy to "natural law" arguments against reproductive technologies in general, by providing a case in which the slippery technological slope seems to lead us to a form of reproduction—asexual reproduction—that contradicts an element of human nature that has been fundamental to our species' identity to date (Lauritzen 2001).

Behind all of these concerns looms a follow-up question that links this discussion to back to the role of genetic lineage in human identity. Given the psychological complexities of cloning and their potential impact on the offspring, is the value of sheer genetic continuity important enough to ever warrant the inclusion of this option on the menus of fertility medicine? Some argue that prospective parents who feel the need for genetic connections with their offspring simply mistakenly essentialistic, and should be re-educated accordingly (Post 1997). Others, however, suggest that the interest in extending the limits of procreative liberty to defend technologies like cloning reflects something important about the role of lineage in human identity which philosophy has yet to fully unravel (Roberts 1995).

HUMAN GENETIC MODIFICATION

The philosophical status of intergenerational genetic connections also lies at the heart of a new generation of attempts to define the appropriate limits of human genetic modification. The provisional boundaries of such a practice were established in the 1980s to provide a window for human gene therapy research. On one axis, a line was drawn between using genetic interventions to treat disease, and using it to attempt to "enhance" human traits to achieve nonmedical goals. On the other genetic interventions that only affected somatic cells were distinguished from those that might lead to intergenerational transmission of modifications, through the "germ-line." Both distinctions have come under recent philosophical critique. The line between treatment and enhancement appears difficult to maintain as a conceptual matter and its moral implications diverge significantly between its personal, parental, professional and public policy applications (Parens 1998). Meanwhile, it begins to appear that the price to pay for successful somatic cell gene therapy will be the development of tools for safe and effective germ-line interventions, and in that light the arguments in favor of abandoning that boundary in the service of medicine are gaining strength (Chapman and Frankel 2003).

Animating these debates is another argument over the importance of our genetic inheritance. Those who strive to preserve the "common genetic heritage of humankind" and protect the rights of future generations to "inherit an untampered genome" argue that our genetic inheritance forms the limits of our "species integrity", the violation of which risks literally de-humanizing ourselves and our offspring (Annas, Andrews, Isasi 2002; Fukyama 2002). On this view, the contours of human nature, and thus the foundations of human moral status, are defined by the pool of genes—and their variants—that humans have collected over the course of our

evolution, and any manipulation of that legacy risks distorting or corrupting the human identity of our offspring.

This view is difficult to reconcile with what population genetics, developmental biology and gene transfer research itself teach us about the fluidity of the human genome and its relatively remote role in the human traits we prize most (Robert and Baylis 2003, DeGrazia 2005). Like the interest in race-based genomic medicine and concerns about the moral status of clones, it accepts a level of genetic essentialism that seems more indebted to Aristotle than to James Watson and Francis Crick. Nevertheless, all these views enjoy wide acceptance in both academic and public circles, which is intriguing. Perhaps this convergence signals the next challenge that genetics offers philosophy: to help clarify the role of genetic histories—the genes we inherit, the lineages we belong to, the peoples we spring from—in our identities as human beings and our experience of the human condition. If philosophical anthropology examines what it means to be a human individual, what genetics seems to call for now is a *philosophical genealogy*: the study of what it means to be a human descendant (Juengst 2004).

See also Aristotle; Distant Peoples and Future Generations; Human Genome Project; Medical Ethics; Natural Law.

Bibliography

Altshuler, David, and Angus Clark. "Harvesting Medical Information from the Human Family Tree." *Science* 307 (2005): 1,052–1,053.

Annas, George, Lori Andrews, and Rosario Isasi. "Protecting the Endangered Human: Toward an International Treaty Prohibiting Cloning and Inheritable Alterations." *American Journal of Law and Medicine* 28 (2002): 151–178.

Chapman, Audrey., and Mark Frankel, eds. *Designing Our Descendants: The Promises and Perils of Genetic Modifications*. Baltimore, MD: Johns Hopkins University Press, 2003.

DeGrazia, David. "Enhancement Technologies and Human Identity." *Journal of Medicine and Philosophy* 30 (2005): 261–285.

Duster, Troy. "Race and Reification in Science." *Science* 307 (2005): 1,050–1,051.

Fukuyama, Francis. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. New York: Farrar, Straus and Giroux, 2002.

Gannett, Lori. "Racism and Human Genome Diversity Research: The Ethical Limits of 'Population Thinking.'" *Philosophy of Science* 68 (2001): S479–S492.

International Hapmap Consortium. "The International Hapmap Project." *Nature* 426 (2003): 789–795.

Juengst, Eric. "FACE Facts: Why Human Genetics Will Always Provoke Bioethics." *Journal of Law, Medicine and Ethics* 32 (2004): 252-267.

Juengst, Eric. "Groups As Gatekeepers to Genomic Research: Conceptually Confusing, Morally Hazardous and Practically Useless." *Kennedy Institute of Ethics Journal* 8 (1998): 183–200.

Lauritzen, P. *Cloning and the Future of Human Embryo Research*. New York: Oxford University Press, 2001.

McGee, Glenn, and Arthur Caplan, eds. *The Human Cloning Debate*. Berkeley, CA: Berkeley Hills Books, 2000.

National Research Council. *Assessing Human Genetic Diversity*. Washington, DC: National Academy Press, 1998.

Parens, E., ed. *Enhancing Human Traits: Ethical and Social Implications*. Washington, DC: Georgetown University Press, 1998.

Post, S. "Adoption Theologically Considered." *Journal of Religious Ethics* 25 (1997): 149–167.

Reardon, J. *Race to the Finish: Identity and Governance in an Age of Genomics*. Princeton, NJ: Princeton University Press, 2005.

Risch, N., E. Burchard, E. Ziv, and H. Tang. "Categorization of Humans in Biomedical Research: Genes, Race, and Disease." *Genome Biology* 3 (7) (2002): 2007.1–2007.12.

Robert, J., and M. Baylis. "Crossing Species Boundaries." *American Journal of Bioethics* 3 (2003): 1–14.

Roberts, D. "The Genetic Tie." *University of Chicago Law Review* 62 (1995): 209–273.

Weijer, C. "Protecting Communities in Research: Philosophical and Pragmatic Challenges." *Cambridge Quarterly of Health Care Ethics* 8 (1999): 501–513.

Eric Juengst (2005)

GENOVESI, ANTONIO
(1713–1769)

The Italian philosopher and economist Antonio Genovesi (the name was originally Genovese), was born in Castiglione, Salerno. After studying literature and rhetoric and then philosophy, he attended the lectures of the aged Giambattista Vico. In 1741 he began to teach metaphysics at the University of Naples as extraordinary professor. In 1743 he published the first volume of his *Elementa Metaphysicae Mathematicum in Modum Adornata* (5 vols., Naples, 1743–1745), for which he was accused of rationalism and atheism. In 1745 he began to teach ethics. In that year he published his *Elementa Artis Logico-criticae* and an important historical introduction to the Neapolitan edition of Pieter van Musschenbroek's *Elementa Physicae*. In the same year his *Universae Christianae Theologiae Elementa* was accused of heterodoxy; it was not published until after his death (Venice, 1771). Discouraged, Genovesi turned to other, less philosophical studies. He was offered the new chair of civil economy

(economics), the first in Europe, by the University of Naples and began his lectures in 1754 (*Delle lezioni di commercio ossia di economia civile,* Naples, 1765–1767). The problems of practical philosophy which occupied his final years are discussed in *Diceosina o sia filosofia del giusto e dell'onesto* (2 vols., Naples, 1766–1777).

In Genovesi's judgment, modern philosophy began when Francis Bacon and Galileo Galilei freed Europe from abstract and sterile inquiry. "Dialectics and metaphysics," he proclaimed, "are the Don Quixote of the Republic of Letters." According to him, it is impossible to know true reality, substance, that which "underlies" the phenomena that we can observe. (He asked, "Who lifts the skirt of nature to see that which ὑπάρχει [underlies]?") Although his thought had some similarities to George Berkeley's idealism and Gottfried Wilhelm Leibniz's monadism, as time went on his interest turned from logic and metaphysics and was oriented toward the moral disciplines, particularly toward economics, which he considered as affecting "our present comfort and tranquility." He sought to determine in a rational system "the primary, simple, and universal laws" of economics. He arranged in a similar framework the *Discorso sopra il vero fine delle lettere e delle scienze* (Naples, 1753), in which he argued against all inquiries "that remain exclusively in the shadow of the school, and never transgress into the acquisition of something useful for mankind."

Bibliography

WORKS BY GENOVESI

Opere scelte. Milan, 1824.

Illuministi italiani. Vol. V: *Riformatori napoletani,* edited by Franco Venturi, 1–330. Naples and Milan, 1962. Contains an important introduction and bibliography on Genovesi's influence in Italy.

Autobiografia, lettere e altri scritti, edited by Gennaro Savarese. Milan: Feltrinelli, 1962.

WORKS ON GENOVESI

Bruni, Luigino, and Robert Sugden. "Moral Canals: Trust and Social Capital in the Work of Hume, Smith and Genovesi." *Economics and Philosophy* 16 (1) (2000): 21–45.

Garin, Eugenio. *La filosofia italiana,* Vol. II. 392–402. Milan, 1947.

Gentile, Giovanni. *Storia della filosofia italiana,* Vol. I: *Dal Genovesi al Galluppi,* 2nd ed., 1–23. Milan: Fratelli Treves, 1930.

Robertson, John. "Antonio Genovesi: The Neapolitan Enlightenment and Political Economy." *History of Political Thought* 8 (1987): 336–344.

Villari, Lucio. *Il pensiero economico di Antonio Genovesi.* Florence, 1959. Contains an excellent bibliography.

Zambelli, Paola. "Antonio Genovesi and Eighteenth-Century Empiricism in Italy." *Journal of the History of Philosophy* 16 (1978): 195–208.

Eugenio Garin (1967)

Bibliography updated by Tamra Frei (2005)

GENTILE, GIOVANNI

(1875–1944)

Giovanni Gentile was one of the major figures in the resurgence of Hegelian idealism in Italy at the beginning of the twentieth century. His "actual idealism," or "actualism," represents the subjective extreme of the idealist tradition in that the present activity of reflective awareness (*l'atto del pensiero, pensiero pensante*) is regarded as the absolute foundation on which all else depends. The act of thinking is the "pure act" that creates the world of human experience.

LIFE AND WORKS

Gentile was born on May 30, 1875, at Castelvetrano in Sicily. He began his university education as a student of Italian literature under Alessandro d'Ancona at Pisa in 1893, but was quickly drawn into the study of philosophy by Donato Jaja, a pupil of the Neapolitan Hegelian, Bertrando Spaventa. Of the two main threads that run through all of Gentile's work, one—his concern with the theory and practice of education—is rooted directly in his own temperament and his strongly felt vocation as a teacher; but the other—his almost chauvinistic interest in the Italian philosophical tradition and its relation to the general European tradition—reflects the lifelong influence of Spaventa on his mind. His degree thesis, *Rosmini e Gioberti* (Pisa, 1898), in which he emphasized points of contact and agreement between the native Catholic thinkers and the German Idealists, was meant to illustrate Spaventa's thesis regarding "the circulation of European philosophy."

His second book was a critical examination of Karl Marx (*La filosofia di Marx,* Pisa, 1899) from an orthodox Hegelian standpoint. While writing it, Gentile became acquainted with Benedetto Croce, who was similarly occupied at the time. Thus began a friendly alliance that lasted more than twenty years. Gentile was the younger by nine years, but it seems clear that in these early formative years it was he who influenced the development of Croce's philosophy rather than vice versa, as most of their contemporaries assumed. Gentile was always more of a Hegelian than Croce ever became, and was more exclusively interested in the traditional problems of philosophy.

In 1900 Gentile wrote his important essay "The Concept of Education" ("Il concetto scientifico della pedagogia") and began his long campaign for the reform of the Italian school system. He became *Privatdocent* at Naples in 1903 and professor of the history of philosophy at Palermo in 1906. But the "reform of the Hegelian dialectic" and the "method of immanence" that led to actual idealism (in a paper of 1912) were worked out amid controversies with Modernists and polemics for religious instruction in elementary schools; and Gentile's philosophy was first fully expounded in the two-volume work *Sommario di pedagogia come scienza filosofica* (Summary of Educational Theory; 2 vols., Bari, 1913–1914).

In 1914 Gentile succeeded to Jaja's chair at Pisa, where he wrote the one book through which he is internationally known, *Teoria generale della spirito come atto puro* (The General Theory of the Spirit as Pure Act; Pisa, 1916). In 1917 he moved to the University of Rome; and the first volume of his *Sistema di logica come teoria del conoscere* (System of Logic as Theory of Knowing; Pisa, 1917), the most systematic statement of his view, appeared. The second volume followed at Bari in 1923.

In 1922 Gentile became minister of education in Benito Mussolini's first Cabinet, and in this capacity he reformed and reorganized the whole Italian school system. After his resignation in 1924 he became the first president of the National Fascist Institute of Culture; he remained for the rest of his life the most prominent publicist of the regime and the self-styled "philosopher of fascism." Gentile continued until his death to lecture at Rome, but in the fascist period his only important philosophical work was the *Filosofia dell'arte* (Milan, 1931). He was directing editor of the *Enciclopedia italiana* from its inception in 1925 to its completion in 1937. After the fall of Mussolini in 1943, Gentile went into retirement and wrote a short but important book on the genesis and structure of society that was published only after his death (*Genesi e struttura della società,* Florence, 1946). Subsequently persuaded to return to public life as a supporter of the Fascist Social Republic set up by the Germans, Gentile was assassinated by Italian communist partisans at Florence on April 15, 1944.

CONCEPTION OF PHILOSOPHY

Gentile justifies his "theory of the spirit as pure act" in two ways. First, he strives to show that it is the logical outcome of the whole movement of Western philosophical thought since René Descartes; and, second, that the "method of pure immanence," when we arrive at it, provides an adequate and coherent way of explicating our actual experience. It is impossible to give more than the briefest indication of the line of his historical argument, although it bulks very large in most of his systematic works.

In any case, the significance of his theory emerges more clearly through an examination of his analysis of actual experience. The claim that actual idealism is the logical outcome of the main tradition of modern philosophy is interesting chiefly because it throws light on Gentile's conception of the essential problem of philosophy and the conditions for its solution. Philosophy for him, as for Johann Gottlieb Fichte, was *Wissenschaftslehre,* the science of knowledge, the science that, without presupposing anything itself, provides an a priori ground for the presuppositions actually made in other sciences. Descartes's method of universal doubt can quite naturally be viewed as the first approach to this problem, and George Berkeley's doctrine that *esse est percipi* is a vital step toward its solution. However, the genesis of actual idealism begins with Immanuel Kant; and although Gentile arrived at his view through the progressive elaboration of a "reform of the Hegelian dialectic" that had been initiated by Spaventa, he remains fundamentally a Kantian in his determination to confine philosophical speculation to the task of exhibiting the logical structure of actual experience. He is at one with Kant and Fichte in his resolute rejection of any "dogmatic metaphysics" that posits or presupposes a reality transcending actual consciousness.

THEORY OF SELF-CONSTITUTION

There is a temptation to say at once that it is a mistake to conceive of the task of philosophy in this way, as the exhibition of the logical structure of actual experience, and that the ideal of a "philosophy without presuppositions" is a chimera. The most primitive postulate of ordinary common sense is that a physical world exists prior to and independent of our consciousness of it. However, Gentile's theory is not meant to be taken as a denial of this assumption, but as a thesis about logical priority. The temporal preexistence of the object of awareness is itself something that we take ourselves to be aware of, and in this sense the commonsense assumption is a product of our attempt to organize our experience in thought. Actual idealism must properly be judged as a theory about this process of rational organization or "concrete logic."

The most primitive level of the process for which we have ordinary words is sensation. We normally distin-

guish the objective cause of a sensation from the subjective feeling (pleasant or unpleasant) that it arouses in us. According to Gentile, this is a mistake. The sensation as a whole is our act of self-awareness, and the pleasure or pain is an aspect of this whole, not a reaction of the self to an object. He agrees emphatically that there cannot be any actual consciousness without the distinction of the subject of the awareness from the object of which it is aware. But he holds that since what has to be understood is the integral unity of the self, it is a mistake to look for the cause of experience within the content of experience. At the ideal limit, pure sensation can be thought of as an encounter with something absolutely other than the self; but it can also be thought of as a spontaneous activity of self-affirmation. Gentile does in fact employ "sensation" in both ways. Spontaneous self-affirmation is in his view the ideal aim of the artist, and loss of self in the contemplation of an absolute object is the typical concern of religious experience. But actual experience is always a synthesis, so that pure art and pure religion are nowhere to be found; and the actual understanding of any type of artistic or religious experience will involve restoring the suppressed aspects of the synthesis, that is to say, discovering the philosophy behind it.

Actual sensation is a process of self-constitution (*autoctisi*) in which the subject preserves its own past and relates it to present sensation. Language is abstractly the instrument and concretely the form through which this is done. It is neither the clothing nor the vehicle, but the embodiment of our thought. But we are able to think of it abstractly, as an inheritance shared by all who are able to use it, because the thought embodied in it has universal import. Thus the self that comes to consciousness when we express our thoughts in language is a spiritual universe, a system of meanings in which all other thinking beings can share. This is the absolute subject of experience, the transcendental Ego whose being (like the God of Aristotle and St. Thomas Aquinas) is "pure act." The abstract form that Kant called the "transcendental unity of apperception" is given concrete existence, or brought to life, so to speak, in Gentile's conception of the "pure act" of "self-founding." My reflective awareness is on the one hand exactly what is essential to my existence as an independent personality; but on the other hand, so far as I achieve reflective awareness, I enter the world of thought in which nothing belongs, or can belong, to me personally. When I claim to think something, I must be able to communicate my thought; I must be able to show others the path by which I arrived at it so that, insofar as they can follow in my footsteps, they can share it. Reflective awareness is already communication, for my own thought is a dialogue within myself. The obvious fact that humans are social animals and that the peculiarly human institution of language is a collaborative production has its absolute or philosophical ground in the fact that the founding of the self is the founding of a transcendental society.

When we understand the fundamental concept of self-constitution in this way, Gentile's thesis about the unity of thought and action, which was the chief bone of contention between him and Croce, falls naturally into place and is easily understood. There appears to be a contrast between thought and action because in cognitive thinking we presuppose the reality we are concerned about, whereas our action is directed toward the creation of some object. However, all thinking and acting is in reality part of the same activity of self-conquest in which nothing is absolutely presupposed theoretically, and some things must be accepted (or presupposed) practically, if there is to be a line between the self and the not-self, the conquering subject and the nature or world that is to be conquered: "the spiritual act is never a self-creation that must be contemplated and watched over afterward; it is always simultaneously a self-creation that is self-awareness and vice versa" (*Opere* I, 84). The establishment of truth is the self-establishment of the transcendental Ego; and the establishment of the Ego is the establishment of an ideal community that Gentile, like G. W. F. Hegel, calls "the State."

The State is on the one side that complex of social institutions, cultural traditions, and ethical values that appears to the individual as the actual fabric of his own moral personality; on the other side it is all the ideals that have still to be striven for and achieved in the actual world in which he lives. Gentile often insists on this latter Mazzinian side of his doctrine, but in practice he tended to subordinate it to his conservative Hegelian faith in the rationality of the actual social structure. In his fascist apologias it often seems as if whatever is done in the name of the existing State must be patiently, even joyfully, accepted and endured as a condition for any further advance—an attitude that is more reminiscent of Thomas Hobbes than of Giuseppe Mazzini. There can be no question that this attitude is false to the spirit of his doctrine.

GENTILE'S LOGIC AND THE FORMS OF VALUE

The unity of theory and practice means that in Gentile's work "logic"—the concrete logic of the self-concept—becomes inseparable if not indistinguishable from ethics,

and philosophy itself is seen as the critical self-awareness of actual political life. His major theoretical problem was to show how the nonpolitical values of human experience could be integrated into his view. This problem came to Gentile in the form that Hegel gave it when he made art and religion the moments of the final triad of the Absolute Idea, subordinate only to philosophy itself. Gentile solved it by regarding art and religion as the moments of his own Absolute, the act of thought. Thus art and religion, instead of being ultimate, become primitive; they are the essential moments of all experience. They have their joint origin, as has been shown, in the opposite aspects of the sensation or "self-feeling" in which consciousness originates. As distinct modes of experience they are attempts to achieve the impossible by aesthetically recapturing or mystically losing oneself in that ideal point of origin.

Thus the seeming independence of aesthetic and religious values arises from the one-sided consciousness of the artist or worshiper. In reality the self-willed artist is dedicated to the production of an object of universal value and significance; art is not just the release of feeling, but the disciplined expression of it. And the proclaiming of the glory of God or the doing of his will is the work of a human voice or the task of a human hand. The "private" world of the artist and the "other" world of the believer get their meaning and fulfill their function in the actual society of the transcendental Ego. When we view the artist's work, we must strive to comprehend the ideal to which he has devoted his skill; and when we seek to interpret a religious doctrine, we must express its meaning for humanity and in terms of our own actual lives. It thus becomes the task of the critic to interpret the work of art or the religious doctrine philosophically. Gentile wrote a number of books and essays—mainly but not exclusively about literary artists—in which he endeavored to do just this; and on the religious side he maintained that his "humanistic conception of the world" was a philosophical expression of the Christian revelation.

In the concrete logic of the act of thought, the moment of spontaneous self-expression is prior to the consciousness of the object, which necessarily appears as a limit upon the self. Hence, in the progressive development of consciousness, which is the subject of Gentile's philosophy of education, an aesthetic phase of free, spontaneous play is succeeded by a religious phase that it is the main task of the elementary school to establish and govern. On this basis a properly philosophical conception of the world, a sense of the autonomous moral responsibility of the self-conscious citizen, should then be built up in secondary education.

ABSTRACT LOGIC. It should by now be clear that actual idealism can be interpreted as primarily a theory about the logical structure of our experience of values. But this theory does contain within it a theory about the ordinary logic of factual propositions. Formal logic, whether mathematical or conceptual, is the logic of presupposition, the logic of "nature," the abstract logic of any object that any actual concrete consciousness may assume as its content. Gentile thought of this *logo astratto* as being essentially static and unchanging. Benedict de Spinoza's system was for him the perfect philosophical expression and reductio ad absurdum of it; and his own conception of natural science was strictly Kantian. "Nature" was for him an a priori concept with a fixed logical structure, not an idea that evolves in the dialectic of actual research. But this is only a reflection of his own personal background and interests. The "idea of nature" has a history, and a full development of the theory of mind as pure act would seem to require that the history of science be incorporated as an essential aspect or complement of the history of theoretical philosophy.

Gentile's own use of the category of the *logo astratto* in the sphere of practical philosophy was quite fluid and dialectical. In ethics, for example, it appears as the nature that we must conquer and subdue, but it appears also as the abstract law to which we must submit. When we remember that the transcendental Ego itself, the *logo concreto,* is both the organic unity of all prior achievement and the ideal of a perfect harmony still to be achieved, this becomes quite easy to understand. The concrete self-concept has its abstract content under each aspect—there is sinful nature and there is the law in virtue of which we are aware of it as sinful; the act of self-constitution is the resolution of the conflict that produced the two opposed abstractions.

NEGATIVE VALUES. Finally, Gentile holds that error, pain, and sin are in some sense "unreal." This doctrine follows logically from the fact that they belong to the category of the *logo astratto.* They are things that we are conscious of, and they have already been overcome or surpassed in the very consciousness of them. It is quite easy to exhibit, as a mere matter of logic, how "truth" is the concrete category of which "error" is only the abstract content. For to be actually aware that some proposition is or may be an error is to hold that a proposition about that proposition is true—namely, the proposition that it is or may be an error.

In the case of sin, something more than a logical relationship of propositions is involved. If I say, "I am a sinner," I am setting myself up as a supposedly just judge of my own conduct; but I do not thereby cease to be a sinner. Rather, the question is posed of how a single self is to be constituted out of this divided consciousness. This is the key to the only defensible interpretation of Gentile's doctrine, which then asserts that when I truly say, "I am a sinner," I must be on the road to redemption and that the test of whether I do actually think I am a sinner is my consciousness of repentance.

Gentile's view that "pleasure" is the concrete and "pain" or "grief" the abstract category is more difficult to interpret. If I am conscious of being in pain, I have certainly "overcome" the pain; that is, isolated it and objectified it as a fact. But to argue that because of this it is not really I who am in pain is sheer sophistry. It is certainly true that the consciousness of pain is a complex activity (including, for example, the active seeking of a remedy or a distraction), while the pain in itself is an abstract element. But pain is "unreal" only in the sense in which art and religion are unreal. That is to say, there cannot be a pure pain-consciousness, for this is just the point at which consciousness disappears.

GENTILE'S INFLUENCE

Actual Idealism was the dominant philosophy in Italian state schools and universities throughout the last twenty years of Gentile's life. In this period his students, like Hegel's, split into two main parties. There was first the "right," led by Armando Carlini, who emphasized the Christian origin and Augustinian character of much of Gentile's thought, and identified the transcendental Ego with the God of Catholic theology. After Gentile's death this group joined with the religious existentialists to form the contemporary movement known as Christian Spiritualism.

On the other side, a group of younger disciples, led by Ugo Spirito, formed the Gentilian "left," which from the first devoted itself to social problems and provided much of the economic and political theory of the fascist corporate state. Since World War II this group has been aligned with the political left and has shown some affinities with orthodox Marxism. But in the current work of both groups it is the mystical spirit of Gentile's philosophy rather than the logical structure that has survived. Outside of Italy, Gentile's influence can be seen most notably in the work of R. G. Collingwood.

See also Absolute, The; Berkeley, George; Collingwood, Robin George; Croce, Benedetto; Descartes, René; Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Hegelianism; Hobbes, Thomas; Idealism; Kant, Immanuel; Marx, Karl; Marxist Philosophy; Philosophy of Education, Ethical and Political Issues in; Philosophy of Education, History of; Spaventa, Bertrando; Spinoza, Benedict (Baruch) de; Spirito, Ugo; Thinking; Thomas Aquinas, St.

Bibliography

The Gentile Foundation was established in Rome in 1947, with the *Giornale critico della filosofia italiana* (founded by Gentile in 1920) as its official organ. The foundation is issuing the definitive edition of Gentile's *Complete Works* in 55 volumes. Three of his books have been translated into English: *The Theory of Mind as Pure Act* (London: Macmillan, 1922); *The Reform of Education,* lectures to the schoolteachers of Trieste (New York: Harcourt Brace, 1922); and *Genesis and Structure of Society* (Urbana: University of Illinois Press, 1960). This last includes a complete bibliography and critical survey of everything by or about Gentile in English.

For critical studies of Gentile, see the following: H. S. Harris, *The Social Philosophy of Giovanni Gentile* (Urbana: University of Illinois Press, 1960) surveys the whole range of Gentile's practical philosophy, including his "philosophy of fascism," which has been largely ignored above. R. W. Holmes, *The Idealism of Giovanni Gentile* (New York: Macmillan, 1937) is a detailed and critical study of the *System of Logic* and a work of fundamental importance. P. Romanell, *The Philosophy of Giovanni Gentile* (New York: S.F. Vanni, 1938) provides a general survey.

H. S. Harris (1967)

GEOMETRY

Until 1800, mathematics was divided into two great branches: *geometry* and *arithmetic*. Both were commonly regarded as the more obviously secure repositories of human knowledge. At this stage, geometry could be suitably defined as "the science which investigates the properties and relations of magnitudes in space, as lines, surfaces, and solids" (*Oxford English Dictionary*). However, with the enormous enrichment of mathematics in the nineteenth century, the scope of geometry was greatly expanded and diversified, its content disrupted, and its epistemic standing called into question.

The word "geometry" comes from a Greek word that literally means measurement of the earth and was originally applied to the art of land surveying. But around 500 BCE or even earlier, the spatial properties and relations

that had been codified by land surveyors in Mesopotamia and Egypt became in Greece the starting-point of inquiries of a more abstract sort that soon took leave of their down-to-earth origins. In this guise, geometry appeared to Plato as a testimony of the other-worldly origin of the human soul (*Meno*) and was included by him as a compulsory item in the curriculum for would-be philosopher-kings (*Republic*, VII, 526c–528d). For more than twenty centuries, philosophers regarded the geometry created by Greek mathematicians from Eudoxus, through Euclid, to Archimedes and Apollonius as the standard of indubitable truth and cogent reasoning. As a result of later developments, geometry, with the rest of mathematics, came to be seen as a capital example of the loss of certainty that currently pervades most areas of civilized life (Kline 1980). In more than one sense, this enhances, rather than diminishes its philosophical significance.

This entry is divided into three sections. The first section touches on some philosophically noteworthy aspects of ancient geometry. The second section deals briefly with geometry and philosophy from 1600 to 1800. And the third section describes those episodes in the history of geometry since 1800 that had the greatest impact on twentieth-century philosophy.

TOPICS IN ANCIENT GEOMETRY

GEOMETRY IN THE MIDDLE EAST. According to Herodotus (2.109), the Greeks learned land surveying (*geōmētriē*) from the Egyptians, who used it to reassess taxes on properties partially washed away by the Nile. It appears that this art was first cultivated in the Middle East to cope with the consequences of floods in southern Mesopotamia. Archaeological evidence from both regions displays applications of the so-called theorem of Pythagoras, and a clay tablet now at Yale University (YBC 7289) gives the length of the diagonal of a unit square as 1.41421296, the same approximation to $\sqrt{2}$ that Ptolemy used some 2,000 years later. The Old Babylonian scribe who calculated it probably knew that he could improve on this figure, but it is highly unlikely that he suspected that no algorithm could ever yield a perfectly accurate one. No extant document from ancient Egypt or Mesopotamia contains the general statement of a geometric theorem or anything that even remotely resembles a geometric proof.

PYTHAGOREANS AND IRRATIONALS. Thales of Miletus, "the first to philosophize," supposedly was also the first to prove a geometric theorem (namely, that a triangle with two equal sides also has two equal angles). The earliest proofs probably consisted of diagrams that plainly displayed the relations they were meant to prove (see Plato, *Meno*, 80d–86c). But Greek geometers soon produced purely discursive proofs (like the one given later in this paragraph). The Pythagoreans, intellectually and politically active in southern Italy throughout the fifth century BCE, worked intensely on mathematical problems, as they thought that numbers (i.e., the positive integers) are the principles of everything. This suggestive belief was supported by their discovery that musical chords are associated with simple numerical proportions. It broke down, however, when a member of the school, possibly Hippasus of Metapontum, showed that there are geometric magnitudes of the same kind whose relative sizes cannot be conveyed by numbers. Presumably, this was first demonstrated for the diagonal and the side of the regular pentagon; but it is proved more easily for the diagonal and the side of a square by the following argument transmitted in an appendix to Euclid's *Elements*.

Take the side of the square as the unit of length. Then, by Pythagoras's theorem, the length of the diagonal equals $\sqrt{2}$. But there are no two integers a and b such that $(a/b)^2 = 2$. For suppose there are. Then, by simplification of the fraction a/b we should find two integers p and q, with no common divisor, such that $(p/q)^2 = 2$. Then $p^2 = 2q^2$ and p is an even number, equal to $2n$, say. (For the square of an odd number, say $2n + 1$, is always odd, that is, $4n^2 + 4n + 1$). But then $2q^2 = p^2 = 4n^2$, and $q^2 = 2n^2$, so that q is also even. But this is impossible, for we assumed that p and q do not have a common divisor. Therefore, one cannot find two integers a and b, no matter how large, such that the diagonal of a square exactly equals $a \times 1/b$ of its side. Awareness of the existence of incommensurable lengths cut short dreams of grasping nature through numbers and opened a chasm between arithmetic and geometry.

EUDOXUS'S THEORY OF PROPORTIONS. Eudoxus of Cnidos (c. 390–c. 337 BCE) invented a method for representing the visible motion of each planet in the sky (including Sun and Moon) as the resultant of the combined uniform rotations of several geocentric spheres. Eudoxus's planetary models are the earliest extant example of geometrical representation of natural processes for the sake of predicting their future evolution. Their moderate predictive success may have motivated Plato's change of mind from his early view that real planetary motions are essentially irregular and unpredictable (*Republic* VII, 529d7–530b4) to his later commendation of mathematical astronomy as an efficient servant of

theology (*Laws* VII, 822a4–c5; X, 897c4–9; XII, 966d6–967d2) and his endorsement of Eudoxus's program as the proper way of "saving the phenomena" of the sky (Simplicius, 7.492.30–35). Eudoxus also originated the method of exhaustion employed by Archimedes for calculating volumes enclosed by curved surfaces, which was the first step toward the creation of the integral calculus. But Eudoxus's chief contribution to geometry was his theory of proportions, preserved in book 5 of Euclid's *Elements*. With it, geometry recovered the computational powers it had lost when separated from arithmetic, and the road was opened for rigorously conceiving and handling physical quantities of all sorts.

Two magnitudes a and b are said to have a ratio $a{:}b$ to one another if there are integers m and n such that $m \times a > b$ and $n \times b > a$. (The assumption that any two lengths have a ratio to one another is known as the Archimedean postulate.) Eudoxus produced definitions by virtue of which ratios can be added and multiplied, yielding new ratios, and any two ratios $a{:}b$ and $c{:}d$ satisfy *trichotomy*, that is, either $a{:}b = c{:}d$, or $a{:}b > c{:}d$, or $c{:}d > a{:}b$. In this last case, there will always be an integer n such that $n(a{:}b) > c{:}d$. Thus, it is natural to regard all Eudoxean ratios as magnitudes that have ratios to one another. This paves the way for setting up equations that combine magnitudes of very different kinds, for example, masses, distances, and times, or volumes, temperatures, and pressures (as represented by their respective ratios with the appropriate units). However, it is not apparent that anyone saw this before the seventeenth century.

EUCLID'S *ELEMENTS*. In this, the most famous of mathematical textbooks, Euclid (c. 325–c. 265 BCE) organized the results and displayed the methods of fourth-century-BCE Greek geometry. It is usually taken for granted that the book is patterned after Aristotle's conception of a true science (*epistēmē*). This must consist of a collection of universal statements (theorems) obtained by deductive inference from self-evident premises (axioms) and definitions using a few self-explanatory terms (primitives). However, Euclid's book, though prima facie it may seem to prove every theorem from five postulates and a short list of so-called common notions, often resorts to unspoken assumptions. Moreover, Euclid's deductions do not all fit into the narrow frame of Aristotle's logic, and use forms of inference first codified by George Boole (1815–1864), Augustus De Morgan (1806–1871), and Charles Sanders Peirce (1839–1914). Also, his primary definitions (e.g., "A *straight* is a line which lies evenly with the points in itself") would have to be further supplemented by axioms to be of use in deductions. It seems more likely, therefore, that Aristotle based his idea of a true science on his own grasp of what contemporary geometricians were doing (textbooks similar to Euclid's had been around since Aristotle was a student in Plato's Academy) but did not set a paradigm that they or their successors actually followed.

EUCLID'S POSTULATES. The first three postulates are not statements, but requests to allow certain constructions. The third—"to describe a circle with any center and any radius"—would require an infinite drawing board, which is not self-evidently available. The fifth is a conditional existential statement: "If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, *if produced indefinitely*, intersect on that side on which are the angles less than the two right angles." Obviously, the condition here printed in italics can only be met on an endless plane. So in the finite world of Aristotelian and medieval cosmology, this postulate is vacuously true, and its existential consequent may be false (there may well not be any such intersection). Still, if Euclid's other postulates and the Archimedean postulate are true, denial of the consequent implies that a quadrangle with three right angles has an acute angle at the remaining corner, so there can be no rectangles. It also implies that polygons with the same shape also have the same size, in which case Aristotle's suggestion (*Physics* 207b29–34) that all geometrical theorems can be demonstrated in his bounded cosmos by suitably scaling down the diagrams employed would simply be wrong.

These seemingly counterintuitive implications kept geometers throughout the centuries trying to prove the fifth postulate from other principles until, shortly after 1820, Nikolay Lobachevsky and Janos Bolyai dared to deny it and independently published essentially the same system of non-Euclidean geometry. It might be a sign of Euclid's genius that he did not gloss over the fact that this assumption (without which the theorem of Pythagoras will not stand) is not self-evident.

GEOMETRY AND PHILOSOPHY AT THE ONSET OF MODERNITY

NATURE GEOMETRIZED. Aristotle taught that natural science, to adequately grasp its proper subject, must employ terms that connote the peculiar matter of each thing, for example, "snub," which only applies to fleshy noses, rather than "concave," which connotes merely a geometric shape (*Metaphysics* E, 1, 1025b30–1026a7; *Physics* II, 2, 194a2–27). Still, he agreed (*Metaphysics* Λ, 8)

with the purely geometric description of astronomical phenomena proposed by Eudoxus, presumably because he believed that ether, the stuff that the heavens are made of, can change only by rotation about the center of the Earth, and this is properly described in geometric terms. Anyway, Aristotle's strictures on science did not deter Archimedes (c. 287–212 BCE) from dealing mathematically with the equilibrium and the flotation of bodies. In the meantime, astronomers from Apollonius (third cent. BCE) and Hipparchus (second cent. BCE), through Ptolemy (second cent. CE), to Copernicus (1473–1543) developed ever more complex geometric models of planetary motion, involving diverse circular motions about different centers (none of which coincides with that of the Earth).

After Galileo Galilei's telescope showed that there are mountains on the Moon and fleeting spots on the Sun, the distinction between celestial and terrestrial physics became pointless, and each took cues from the other. Thus Johannes Kepler (1571–1630) sought to explain the motion of planets (including the Earth) by *forces* exerted on them from the Sun, while Galileo (1564–1642) proposed a chronogeometrical model of free fall on the surface of the Earth, which he conceived as uniformly accelerated rectilinear motion. Lasting success was finally achieved by Isaac Newton (1642–1727), by dint of his mathematical genius and his consummate command of geometry. In the course of these efforts, Kepler (1609) had the words "God is always doing geometry" printed on the front page of his masterpiece, and Galileo wrote that the book of nature "is written in mathematical language, and its characters are triangles, circles and other geometrical figures, without which we cannot understand a word of it" (1623, sec. 6). René Descartes's contention that extension is the one and only clearly and distinctly conceivable attribute of bodies surely called for a comprehensive and thoroughgoing geometrization of physics and might have led to it had geometry been ripe enough to deal with its strenuous demands.

DESCARTES'S REVOLUTION IN GEOMETRY. Except for his first law of motion (the principle of inertia) and his work on the refraction of light, Descartes's direct contribution to physics, subjected to unsparing criticism by Christian Huygens, Gottfried Wilhelm Leibniz, and Isaac Newton, failed to gain admission into the classical canon. But modern mathematical physics would not have been possible without Descartes's *indirect* contribution to it, through his two great inventions in geometry: *coordinates* (independently introduced also by Pierre Fermat) and the *algebra of lengths*.

Coordinates are quantitative labels employed for identifying points in space. By means of them the relations among the points can be quantitatively represented and investigated. Nowadays geometric coordinates are drawn from the field of real numbers $\mathbb{R}$, which we regard as a natural extension of $\mathbb{Q}$, the field of rationals, which, in turn, is constructed from the familiar integers. But this understanding of these matters was still far off in Descartes's time (although his geometric algebra was a decisive step toward it). To avoid anachronism, one must regard Descartes's original coordinates as oriented lengths or, more exactly, as Eudoxean ratios between such lengths and a conventionally chosen unit length. To assign so-called Cartesian coordinates to a point P in space, one takes the three distances x, y, and z from P to three mutually perpendicular planes (listed in a conventional order, the same for all points), and prefixes to each a plus sign or a minus sign, according to the side of the respective plane that faces P (again by convention). The Cartesian coordinates of P then form an ordered triple of oriented lengths, say $\langle +x, -y, -z \rangle$. (There are other ways of defining coordinates: Oblique coordinates depend on three planes not at right angles to each other. Polar coordinates label a point P by its absolute distance from a fixed point O, the angle made by OP with a fixed plane Γ through O, and the angle made by the perpendicular projection of OP on Γ with a fixed line through O on that same plane.)

In Euclid's *Elements* (1956), segments are added to segments in an obvious way to obtain new segments; multiplying a segment s by an integer n amounts to adding n copies of s end to end; a straight segment or a length is never multiplied by another one. Until not too long ago, it was usually understood that such multiplications do occur in Euclid's book, but then the product of such a multiplication had to be an *area*. Descartes followed Euclid on the addition of lengths and defined the multiplication of a length a by a length b so as to yield still another length ab. Here is how. Draw two straight lines from a point O. Mark points F and H on one line so that OF has unit length and OH has length a. Mark point G on the second line so that OG has length b. Draw the straight line segment FG. Let the parallel to FG through H cut the second line at K. Clearly, then, $OH/OF = OK/OG$. Therefore, $OG \times OH = OK \times OF$; in other words, OK has length ab (Figure 1).

By this procedure, entirely based on elementary geometrical knowledge available to Euclid, Descartes and his successors were able to represent all geometrical relations by equations or inequalities between given and unknown

FIGURE 1

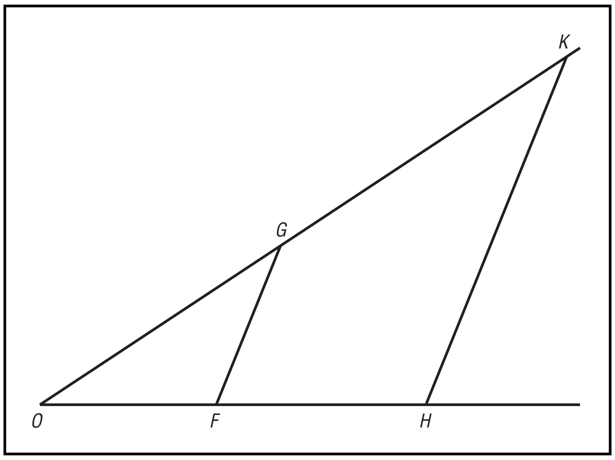

quantities, and to solve geometrical problems algebraically. In the algebra all such quantities were handled in the same way as the positive integers and they were therefore called *numbers*. Newton explains: "By a number we do not understand a multitude of units, but rather the abstract ratio of any quantity to another quantity of the same kind, which is taken as a unit. There are three varieties of number: integers, rationals and irrationals" (1707, p. 2). Eventually, they were called *real* numbers, to distinguish them from the imaginary ones, that is, the multiples of $\sqrt{-1}$, which also turned up as solutions of algebraic equations.

The method of coordinates soon suggested the idea of a space with n dimensions, whose points would be labeled by n quantities. In particular, if $\langle x, y \rangle$ denotes an arbitrary point on a plane, a straight line on that plane can now be defined as the set of points satisfying the linear equation $y = ax + b$, and a circle with radius r and center at $\langle 0, 0 \rangle$ as the set of points satisfying the quadratic equation $x^2 + y^2 = r^2$. These two equations take care of all points on the plane that can be constructed with a ruler and a compass, which were the only points contemplated by Euclid. But after Descartes, mathematicians felt free to consider any curve defined by an algebraic equation or indeed by a convergent series, such as $y = \sin x$, or $y = e^x$ (where e is the base of the natural logarithms). Even though Euclid never countenanced the plethora of points obtainable in such ways and it does not follow from his postulates, what we normally call "Euclidean space" comprises them all.

KANT'S PHILOSOPHY OF GEOMETRY. The overwhelming success of geometry in physics and astronomy induced some seventeenth-century philosophers to follow its example in ethics and metaphysics. The foremost instance of this is Benedict de Spinoza's *Ethica ordine geometrico demonstrata* (Ethics demonstrated in geometric order; 1677), but John Locke too believed that "if men would in the same method, and with the same indifferency, search after moral as they do mathematical truths," then "a great part of morality might be made out with that clearness, that could leave, to a considering man, no more reason to doubt, than he could have to doubt of the truth of propositions in mathematics, which have been demonstrated to him" (1690, IV.iii.20, xii.8).

Immanuel Kant, however, thought otherwise. Invidiously comparing geometry, as a science that "excels all others in certainty and distinctness," with metaphysics, which "has only just started out on the path to these goals" (1902–, 2: 168), he recommended, in 1763, that the latter stop imitating the former, in order to progress along that path. He soon went further. In his Latin dissertation of 1770, Kant taught that confusion and stagnation in metaphysics were due to the contamination of the human intellect with the sensuous notions of time and space. By thoroughly avoiding them, metaphysics will escape the temptations of materialism and determinism and become a secure science of God, freedom, and immortality. Yet in the *Critique of Pure Reason* (1781), Kant likened the purified understanding he advocated in 1770 to a bird that, tired by the resistance of the air, sets out to fly in a vacuum. In his mature view, the basic concepts of human thought—one and many, reality and negation, substance and cause—are not obtained from sensuous experience, but they can refer to objects only when applied to it, under the conditions of human sensibility, namely, space and time.

This decisive turnabout in the history of philosophy is closely related to Kant's reflections on geometry and its use in physics. In 1746 Kant spoke of a general or "supreme" geometry, adapted to a space with any number of dimensions. That the space we live in has only three dimensions is due to the empirical fact that all material particles are linked by forces governed by Newton's inverse-square law (1902–, 1: 34). But in 1768 he made a discovery that, he thought, put an end to all such explanations of space and spatial structure from the physical relations between bodies in space. No description of a shoe in terms of its different parts and the relations between them will allow us to tell a left shoe from the matching right shoe; the difference between the two shoes can be grasped only by considering their respective orientation in the space that embraces them. Kant understood this to imply that the bodily structure of bodies

depends on that of space as a whole, which therefore is presupposed by them, rather than being only an expression of their interactions.

Kant was then faced with the following dilemma: Either (a) space itself is a substance, of which bodies are modes (a position that, according to Kant, results in Spinoza's unchristian and immoral deification of space), or (b) space must be thought of as possessing a novel, hitherto unheard of manner of existence (which implies a corresponding adjustment of the ontological standing of bodies as such). About 1769 Kant lighted on alternative (b), which he described around 1791 as one of the two hinges on which metaphysics must turn (the other one being the reality of freedom). He claims that "*space is not something objective and real*, neither a substance, nor an accident, nor a relation; it is rather a subjective and ideal scheme, so to speak—which issues from the nature of the mind according to a stable law—for coordinating everything that is sensed externally" (1902–, 2: 403). Or, in mantra form, *space is one of the forms of human sensibility* (time is the other one). As a consequence of this, things are bodies only insofar as they are actual or potential objects of our sense perception, but not as they are in themselves. (Indeed, Kant figured out in the early 1770s that the standard assumption that things in themselves are spatial would make it impossible to solve the contradictions regarding the limit of the physical world and the divisibility of its content, which he later set forth in the first two items of the Antinomy of pure reason.)

Kant's conception of space is the key to his philosophy of geometry (and is in turn reinforced by it). The epistemological problem of geometry lies in explaining how it can furnish us with precise quantitative information about things we have never met in real life and which anyway we could not measure accurately, for example, the exact size of the angles of a trillion-sided regular polygon. Plato proposed that this knowledge is remembered from another life in which we had direct access to the intelligible "form" of things. The fact that geometry contains such knowledge nourished similar hopes for metaphysics and ethics, which, however, were crushed by Kant's approach. In his view, geometry rests on our natural awareness of the conditions under which alone the manifold appearances displayed through our external senses "can be ordered into certain relations" (1787, B 34) and thus shaped into corporeal phenomena. Such awareness is not intellectual but intuitive, as we may gather from the example of the pair of shoes, described above, and also from the fact that geometrical proofs proceed by the "construction of concepts." Kant explains this expression somewhat intriguingly as follows: "To *construct* a concept means to exhibit a priori the intuition that corresponds to it: the construction of a concept therefore requires a *non-empirical* intuition which … as intuition is a *particular* object, but nevertheless, as the construction of a concept (a general idea), must convey universal validity for all possible intuitions that belong under the same concept" (1781/1787, A 713/B 742).

Anyone not put off by these opaque notions could well regard them as a proper explanation of the amazing success of geometry in physics. For, if geometry spells out the ordering that is required for us to grasp external phenomena, then it is no wonder that all external phenomena comply in every detail with the teachings of geometry. Soon, however, innovations in geometry moved the ground from under Kant's position and made it untenable. Before we turn to them, it should be emphasized that among these innovations, the best known one—the derivation of a consistent geometry from the denial of Euclid's fifth postulate—does not challenge Kant's view but somehow corroborates it. For Kant, geometry provides information, conveyed by what he called *synthetic* propositions, and this implies that any of its unproven principles can be denied without self-contradiction.

FROM GAUSS TO HILBERT AND BEYOND

NON-EUCLIDEAN GEOMETRY. The fact that Euclid's fifth postulate is not self-evident prompted several mathematicians to try to prove it. John Wallis (1616–1703) succeeded in inferring it from the assumption that for any given figure there is another one, similar to it, of any arbitrary size. This assumption is neither necessarily true nor empirically obvious, but it does provide a perspicuous characterization of Euclidean space.

Girolamo Saccheri (1733[1986]) sought to prove the postulate indirectly. He devised a quadrilateral thus constructed on a plane: Draw straight lines m and n through points P and Q, making right angles with the segment PQ. Mark points A and B on m so that $AP = PB$. Mark points C and D on n so that CA and DB are both perpendicular to m (See Figure 2).

If one assumes the Archimedean postulate (which Saccheri tacitly does), the fifth postulate will hold if and only if $\angle ACQ$ and $\angle BDQ$ are right angles. Saccheri assumed that these angles are obtuse and easily proved that, if so, two points can be joined by more than one line, which he considered absurd. He then assumed that both

FIGURE 2

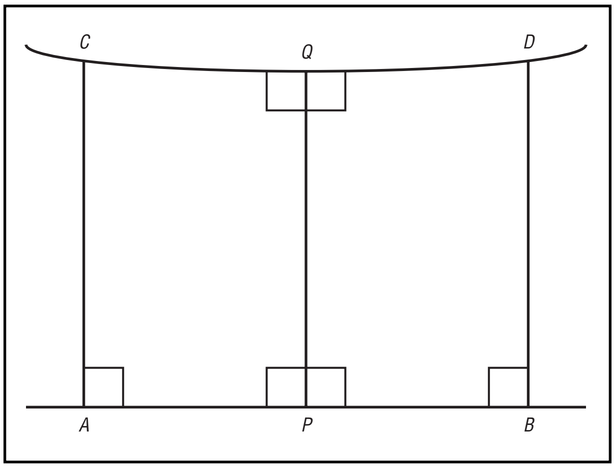

angles are acute and derived from this hypothesis many surprising propositions that did not appear to be contradictory, until at last he reached one he pronounced "repugnant to the nature of the straight line."

The consequences that Saccheri drew from the acute-angle hypothesis reappeared in the nineteenth century in the private papers of Carl F. Gauss (1777–1855) and in independent publications by Nikolay I. Lobachevsky (1793–1856) and Janos Bolyai (1802–1860). These authors treated these consequences as theorems of an alternative system of geometry, based on the straightforward denial of Euclid's fifth postulate (with the others retained). This system has received various names, but by priority of publication, it should be called *Lobachevskian geometry*. In this geometry, the three interior angles of a triangle add up to less than two right angles, the difference being proportional to the area of the triangle. Therefore, similar triangles are congruent. Consider again the segment *PQ* perpendicular to straight line *m* at *P*. By the denial of the fifth postulate, there is a set *S* of straight lines through *Q* that form an acute angle with *PQ* on one or the other side of it and yet do not meet *m* on that side (let alone on the side where they form an obtuse angle with *PQ*). Let α be the smallest of these angles. By symmetry, there are two lines in *S* that form angle α on either side of *PQ*. In Lobachevsky's terminology (independently adopted also by both Gauss and Bolyai), these two lines are called the *parallels* of *m* through *Q*, and α is the *angle of parallelism* for *PQ*. The size of α decreases as *PQ* grows. On any Lobachevskian plane, there is a unique length *h* such that the angle of parallelism for any segment of length *h* equals 45°. The length *h* provides an absolute standard of length for that plane. Kant's friend Johann Heinrich Lambert, who around 1766 worked on this subject along lines similar to Saccheri's, said there was "something alluring about this consequence which readily arouses the desire that the [acute angle] hypothesis be true!" (1786/1895, p. 162). Note that if, in the case discussed, $PQ = h$, the two parallels to *m* through *Q* are mutually perpendicular. Lambert thought this was an intolerable paradox.

The absence of contradiction in a long series of theorems inferred from the denial of the fifth postulate does not, of course, imply that Lobachevskian geometry is consistent. Lobachevsky proposed an argument for proving that his geometry is at least as tenable as Euclidean geometry. He showed that there is a logically formal correspondence between the equations of Lobachevskian trigonometry and the familiar equations of spherical trigonometry. By virtue of it, any contradiction derived from the former will be matched by one flowing from the latter. Such a contradiction would entail that the said standard trigonometric equations are false, and this in turn would entail the falsehood of the Euclidean principles from which these equations follow.

Lobachevsky also tried to ascertain whether his own geometry or Euclid's is true of physical space. He used astronomical data to calculate the sum of the internal angles of the triangle formed by three stars and concluded that the difference between the result obtained in a Lobachevskian space and the Euclidean value was well within the margin of observational error. Decades would pass before Hermann Lotze (1879, p. 774) pointed out that all such attempts are vain, for if astronomical measurements do not agree with Euclidean geometry, the disagreement can still be accounted for by a deviation of stellar light from its supposedly rectilinear trajectory.

GROUPS AND INVARIANTS. Shortly before Lobachevsky's earliest publication on his geometry, Jean-Victor Poncelet's *Treatise on the Projective Properties of Figures* (1822) started a way of doing geometry that seemed more intuitive and tame but which ultimately was much more radical and would have deeper consequences than the denial of Euclid's fifth postulate. It is based on adding to each straight line *m* a "point at infinity" that *m* shares with every straight line parallel to it and treating all such "points" as belonging to a single "plane." This assumption enormously simplified the statement and the proof of geometric theorems concerning relations of incidence, collinearity, and coplanarity among points, lines, and planes. Metric features like distance and metric relations like congruence were totally ignored. Natural (initially tacit) assumptions regarding the neighborhood relations

between the points at infinity and the standard points implied that ordinary space differed drastically from the new *projective space* in which it was now embedded. For example, a left shoe traveling indefinitely in a fixed direction would, after crossing the plane at infinity, return to its original location from the opposite side, in the guise of a right shoe. In this way, projective geometry disposed of Kant's claim about the irreducible difference between the two kinds of shoes, its ontological implications, and its intuitive roots.

Projective geometry grew in scope and sophistication at the hands of August Moebius (1790–1868), Julius Plücker (1801–1868), Karl Georg Christian von Staudt (1798–1867), Arthur Cayley (1821–1895), and others. Different sorts of numerical coordinates were introduced as sheer labeling devices, for in this metric-free context they plainly did not represent distances. The use of coordinates consisting of complex numbers made it possible to introduce more points, in addition to the familiar real points (those labeled by real numbers). These "complex points" are linked to real points and among themselves by relations of collinearity (if their coordinates satisfy the same linear equations) and vicinity (by dint of the neighborhood relations between real and nonreal numbers on the complex plane). The beautiful vistas opened by such developments inspired further flights of mathematical freedom leading to the creation of still other branches of geometry.

Moved by the confusing variety of geometrical methods and approaches, Felix Klein formulated his celebrated Erlangen Program, in which he seeks to unify all forms of geometry under a single overarching point of view. This is provided by the notion of *transformation group* and the related notion of *invariant*.

Let S be a set of points. A *transformation* (or *permutation*) T of S assigns to each point p of S one and only one point $T(p)$ of S, in such a way that every point of S equals $T(p)$ for some p. In other words, a transformation of S is a one-to-one mapping of S onto itself. We say that $T(p)$ is the *value* of T at p. T is said to *send* p to $T(p)$. If M is a subset of S, T is said to *send* M to the set $T(M) = \{T(p): p \in M\}$. For every transformation T, there is an *inverse transformation* T^{-1} that, for each p in S, sends $T(p)$ back to p. The identity transformation Id_S sends every p in S to itself. Given two transformations T_1 and T_2, their *product* T_2T_1 is the transformation that sends each p in S to the value of T_2 at $T_1(p)$. The product of transformations is clearly *associative*, that is, $(T_3T_2)T_1 = T_3(T_2T_1)$ for any three transformations T_1, T_2, and T_3. A set G is a *group of transformations* of S if every element of G is a transformation of S and G contains (1) the product of any two of its elements, (2) the inverse of every one of its elements, and (3) the identity transformation Id_S. Any subset of G that meets conditions (1) through (3) is said to be a *subgroup* of G.

Given a group G of transformations of a set S, let R be an n-adic predicate ($n \geq 1$) such that, for any points $p_1, \ldots, p_n$ in S and any transformation T in G, $R(p_1, \ldots, p_n)$ implies that $R(T(p_1), \ldots, T(p_n))$. We say then that R is an *invariant* of group G or that R is *G-invariant*. Likewise, a function f on S^n is said to be *G-invariant* if, for every n-tuple $\langle p_1, \ldots, p_n\rangle$ of elements of S, $f(p_1, \ldots, p_n) = f(T(p_1), \ldots, T(p_n))$. G is said to *preserve* its invariants.

Klein's Erlangen Program for systematically ordering geometries is based on the following simple idea: Each geometry is the study of the invariants of a group, and the relations of inclusion between groups and their subgroups determine a hierarchy of geometries. Starting from the group of all possible transformations of an arbitrary set, whose sole invariant is the cardinality of the set, one descends, through multiple branches, right down to the trivial group, which is a subgroup of every group and preserves every property and relation, for it only comprises the identity transformation. In particular, projective geometry studies the invariants of the group of collineations, that is, the set of transformations that send straight lines to straight lines. This is a subgroup of the group of continuous transformations, whose invariants are the *topological* properties of projective space. Drawing on work by Arthur Cayley (1859), Klein (1871, 1873) found a way of defining different real-valued functions on point pairs that behaved, on well-defined regions of projective space, precisely like the distance functions of, respectively, Lobachevskian geometry (which he called *hyperbolic*), Euclidean geometry (which he called *parabolic*), and a third geometry (which he called *elliptic*). Each of these functions was an invariant of a certain subgroup of the said group, comprising the collineations that map a specific quadric surface onto itself.

Klein's result led Russell (1897) to assert that the general "form of externality" is disclosed to us a priori in projective geometry, but its metric structure—which Russell wrongheadedly claimed can *only* be Lobachevskian, Euclidean, or elliptic—must be determined a posteriori by experiment. Poincaré took another view of this matter: If geometry is nothing but the study of a group, "one may say that the truth of the geometry of Euclid is not incompatible with the truth of the geometry of Lobachevsky, for the existence of a group is not incompatible with that of another group" (1887, p. 290). Euclidean geometry has

seemed preferable only because the rotations and translations of the Euclidean group reflect, to a comfortable approximation, the motions of ordinary hard bodies in our environment.

RIEMANNIAN MANIFOLDS. In his lecture "On the Hypotheses Which Lie at the Foundation of Geometry" (1867), Bernhard Riemann took an approach to geometry that did not fit into Klein's Erlangen Program. He noted that traditional geometry rested on assumptions summed up in Pythagoras's theorem (by which the distance between a point with Cartesian coordinates $\langle x, y, z\rangle$ and the origin $\langle 0, 0, 0\rangle$ equals the positive square root of $x^2 + y^2 + z^2$). These assumptions had been corroborated by using light rays to line up things and rigid bodies to measure the length of lines, and therefore were bound to break down on very small scales, where such physical objects are not available, and perhaps also on very large scales, where the errors of observation generated by using such instruments might become intolerable.

Riemann therefore proposed to proceed from more general assumptions toward a more flexible geometry that physicists could later resort to when they needed it. He took his cue from Carl Friedrich Gauss's work on the intrinsic geometry of surfaces (1828), which he extended to general spaces of n-dimensions (*n-manifolds*, for short). Though Riemann supposedly addressed his lecture to humanists (hence its meager use of mathematical symbolism), the meaning and reach of the lecture first became clear through its further elaboration by other outstanding mathematicians (e.g., Elwin Bruno Christoffel, Friedrich Schur, Wilhelm Killing, Élie Cartan, Hermann Weyl), and a whole century would pass before it was satisfactorily explained to undergraduate students (Spivak 1979). The following rough sketch of Riemann's breakthrough owes much to the light shed on it by such later developments.

An n-manifold M is furnished with coordinate systems or *charts*, by which different regions or *patches* of M are mapped continuously and one-to-one onto subsets of $\mathbb{R}^n$ (the set of all n-tuples of real numbers, endowed with the neighborhood relations it inherits from the real-number field $\mathbb{R}$). Two charts defined on overlapping patches are said to be *compatible* if the coordinate transformation between them is a smooth function from one open subset of $\mathbb{R}^n$ onto another (possibly the same) one. An *atlas* of M is a collection of compatible charts for M such that every point of M lies on the patch of at least one chart. Any atlas A of M determines a corresponding maximal atlas A_{max} comprising every conceivable chart of M compatible with those in A. The way the charts of A_{max} combine with each other in coordinate transformations reflects the overlapping and intertwining of the patches on which they are defined and thus specifies the global topology, the shape, of M. (In the realm of 2-manifolds, or surfaces, the atlas of a pretzel differs from that of a donut or a bun.) The lengths of curves drawn in M (which are best thought of as continuous mappings of an open interval of $\mathbb{R}$ into M) can then be defined in an endless variety of ways by rules that assign to each point p of M an appropriate function on the coordinate differentials at p, in a manner that varies smoothly from point to point.

Riemann was aware that this approach gave the mathematician enormous freedom, and he proposed restricting the admissible functions, for the time being, to quadratic functions on coordinate differentials, which, on a small neighborhood of each point of M, would yield a definition of length in optimal agreement with the n-dimensional version of the Pythagorean theorem. To conform to the standard concept of length, he also required his quadratic functions to be positive-definite, that is, to take their values only among the nonnegative real numbers. This requirement was subsequently relaxed in the spacetime geometries of the theory of relativity, which, for this reason, are often called *semi-Riemannian*.

From the standpoint of twentieth-century mathematics, the characterization of an n-manifold M by means of an atlas of the kind described makes it possible to assign an n-dimensional real vector space to each point p of M, the *tangent space* at p (thus called by analogy with the plane tangent to a smooth surface at any point of it). A Riemannian metric on M assigns to each p of M a tensor of rank 2, that is, a bilinear function on its tangent space, that varies smoothly from point to point. Riemann's quadratic functions on the coordinate differentials at each point can be naturally obtained as appropriate representations, relative to one or another locally defined coordinate system, of such coordinate independent objects.

A given Riemannian metric μ on M determines a smoothly varying assignment, to each p of M, of a tensor of rank 4, that is, a quadrilinear function on its tangent space. Such an assignment is in effect a *field of tensors* (one at each point), but, for short, this assignment is called the *Riemann tensor* of the Riemannian manifold $\langle M, \mu\rangle$. This is a natural generalization to n-manifolds of the analytically defined yet fairly intuitive concept of the Gaussian curvature of a surface (a real-valued function that is positive and constant on a sphere, variable but

always positive on an ovoid, alternatively positive and negative on the surface of a saddle, and equal to 0 on a plane or a cylinder), and is therefore also called the *curvature tensor*, a term that mathematicians wield nonchalantly but that among philosophers has been a source of endless worries.

On this same analogy, *n*-manifolds of constant 0 curvature are said to be *flat*. In particular, Euclidean space is a flat 3-manifold, while Lobachevskian space is a 3-manifold of constant negative curvature, and Klein's elliptic space is a 3-manifold of constant positive curvature. But a generic Riemannian manifold $\langle M, \mu \rangle$ has variable curvature, and therefore the only group of transformations of M that will preserve the metric μ is the trivial group consisting only of the identity transformation. Already for this reason, Riemannian geometry obviously cannot fall under the Erlangen Program. Another excluding reason is the fact that a geometric inquiry that considers general spaces of endlessly different shapes cannot be characterized by a group of transformations of one of these spaces onto itself. Still, there are Riemannian manifolds endowed with interesting symmetries, and group theory has been the tool of choice for studying them.

HILBERT'S *FOUNDATIONS*. Euclid's putative program for logically inferring the truths of geometry from a sufficient list of unproven premises was fondly imitated by scientists and philosophers in the seventeenth century, but it was first properly carried out by Moritz Pasch (1882). He gathered what he regarded as the empirical foundation of geometry into a few undefined concepts concerning the shape, size, and reciprocal position of bodies and a few axioms that linked these concepts among themselves and with other concepts defined in terms of them. Pasch's axioms "state what has been observed in certain very simple diagrams" (p. 43). All other geometric statements should be proved from the axioms by the strictest deductive methods.

Pasch dealt with projective geometry. The first rigorous axiomatization of Euclidean geometry was given in David Hilbert's *Foundations of Geometry* (1899), a book that had a major influence on twentieth-century mathematics and philosophy. Hilbert invited the reader to consider three arbitrary sets of objects, which he called points, straights, and planes; three undefined relations of *incidence* between a point and a straight line, between a straight line and a plane, and among three points; and two undefined relations of *congruence* between two pairs of points (*segments*,) and between two equivalence classes of point triples (*angles*). Hilbert linked these objects and relations through nineteen axioms, which—when supplemented with the "axiom of completeness" added in the second edition—are sufficient for characterizing the said objects and relations up to isomorphism. This means that if we have two threefold collections of points, straights, and planes having the prescribed relations of incidence and congruence in agreement with Hilbert's twenty axioms, there will always be a one-one mapping of the points, straights, and planes of one collection respectively onto the points, straights, and planes of the second that preserves all five sorts of relations. Such a structure-preserving mapping between structured sets is called an *isomorphism*. Evidently it can hold between two systems of intuitively very dissimilar objects.

Hilbert availed himself of this feature of axiomatic theories for studying the independence of some axioms from the rest. To prove such independence, he proposed actual instances (*models*) of the structure determined by all the axioms but one, plus the negation of the omitted one. Gottlob Frege complained that the geometric axioms retained in these exercises could be applied to Hilbert's far-fetched models only by tampering with the natural meaning of words. Hilbert replied, on December 29, 1899, "Every theory is only a scaffolding or schema of concepts together with their necessary mutual relations, and the basic elements can be conceived in any way you wish. If I take for my points any system of things, for example, the system *love*, *law*, *chimney-sweep*, … and I just assume all my axioms as relations between these things, my theorems, for example, the theorem of Pythagoras, also hold of these things.… This feature of theories can never be a shortcoming and is in any case inevitable" (Frege 1967, p. 412). Hilbert's declaration of independence from sense experience and ordinary usage concisely expresses the modern view of mathematics as a universal "science of patterns" (Resnik 1997), in which geometry is barely distinguishable from its other branches, except on historical grounds.

THE GEOMETRY OF THE UNIVERSE. This approach to pure geometry and mathematics gives physicists enormous freedom to choose the abstract structures they judge most suitable for representing (modeling) the phenomena under inquiry. Yet, as Albert Einstein (1921) pointed out, so long as physics remains unable to provide, from microphysical principles, an exact theoretical construction of the instruments it uses for measuring distances and times, it will continue to need a practical geometry, which must be suggested and corroborated by experience.

According to Einstein, the stability of sharp spectral lines justifies the postulate that any two ideal clocks, once running beside each other at the same rate, will always do so, no matter where and when they are brought together again for comparison. Under this postulate, and for spatiotemporal regions sufficiently small that gravity is practically homogeneous in them, experience has amply vindicated the validity of the flat semi-Riemannian geometry that Hermann Minkowski (1909) initially proposed for the whole world. For broader regions, Einstein's theory of general relativity assumes a semi-Riemannian spacetime geometry whose variable curvature reflects the variations in gravity.

By solving Einstein's equations of the gravitational field under cosmologically plausible special symmetry requirements, Alexander Friedmann (1922, 1924) produced big-bang models of the universe. These big-bang models were ready at hand to explain the systematic recession of the galaxies away from us when discovered by Vesto Slipher and Edwin Hubble before 1930 and the current low temperature of the background thermal radiation when discovered by Arno Penzias and Robert Wilson in 1964. It is worth emphasizing that the explanation of these phenomena from the Einstein field equations is purely geometrical—a consequence of the shape of the universe under the postulated symmetry requirements. Purely geometric arguments also support the proof by Stephen Hawking and Roger Penrose (1970) that, under physically very plausible assumptions, a generic relativistic space-time most likely contains black holes.

See also Aristotle; Black Holes; Boole, George; Copernicus, Nicolas; De Morgan, Augustus; Descartes, René; Einstein, Albert; Frege, Gottlob; Galileo Galilei; Hilbert, David; Kant, Immanuel; Kepler, Johannes; Leibniz, Gottfried Wilhelm; Locke, John; Lotze, Rudolf Hermann; Mathematics, Foundations of; Newton, Isaac; Peirce, Charles Sanders; Plato; Poincaré, Jules Henri; Pythagoras and Pythagoreanism; Russell, Bertrand Arthur William; Space; Spinoza, Benedict (Baruch) de; Thales of Miletus; Weyl, (Claus Hugo) Hermann.

Bibliography

REFERENCES

Aristotle. *The Complete Works*. Revised Oxford translation. 2 vols., edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.

Cayley, Arthur. "A Sixth Memoir upon Quantics." *Philosophical Papers of the Royal Society of London* 149 (1859): 61–90.

Einstein, Albert. *Geometrie und Erfahrung*. Berlin: Julius Springer. Reprinted in *The Collected Papers of Albert Einstein*. Vol. 7, 383–402. Princeton, NJ: Princeton University Press, 2002. Translated as "Geometry and Experience" in the matching supplementary volume.

Frege, Gottlob. *Kleine Schriften*, herausgegeben von I. Angelelli. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 1967.

Friedmann, Alexander. "Über die Krümmung des Raumes." *Zeitschrift für Physik* 10 (1922): 377–386.

Friedmann, Alexander. "Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes." *Zeitschrift für Physik* 21 (1924): 326–332.

Galileo Galilei. *Il saggiatore nel quale con bilancia esquisita e giusta si ponderano le cose contenute nella Libra astronomica e filosofica di Lotario Sarsi Sigensano*. Roma: Giacomo Mascardi, 1623. Partial English translation in Stillman Drake, *Discoveries and Opinions of Galileo*. New York: Doubleday, 1957 (*The Assayer*).

Gauss, Carl Friedrich. *Disquisitiones generales circa superficies curvas*. Göttingen, Germany: Dieterich, 1828. English translation by A. Hiltebeitel and J. Morehead: *General Investigations of Curved Surfaces*. Hewlett, NY: Raven Press, 1965.

Hawking, Stephen W., and Roger Penrose. "The Singularities of Gravitational Collapse and Cosmology." *Royal Society of London Proceedings* A 314 (1970): 529–548.

Herodotus, with an English translation by A. D. Godley. 4 vols. Cambridge, MA: Harvard University Press, 1946.

Hilbert, David. "Die Grundlagen der Geometrie." In *Festschrift zur Feier der Enthüllung des Gauss-Weber Denkmals*. Leipzig, Germany: B.G. Teubner, 1899, pp. 3–92.

Kant, Immanuel. (A), *Critik der reinen Vernunft*. Riga: J.F. Hartknoch, 1781. Repeatedly translated into English as *Critique of Pure Reason*.

Kant, Immanuel. (B), *Critik der reinen Vernunft*. Zweyte hin und wieder verbesserte Auflage. Riga: J.F. Hartknoch, 1787. Second edition of Kant (A); variants normally included in English translations.

Kant, Immanuel. *Kant's gesammelte Schriften*, herausgegeben von der Preussischen Akademie der Wissenschaften (Bde. 1–22), der Deutschen Akademie der Wissenschaften zu Berlin (Bd. 23), und der Akademie der Wissenschaften zu Göttingen (Bde. 24, 25, 27–29). Berlin, 1902–. Page numbers of this edition are printed on the margin of *The Cambridge Edition of Kant's Works in Translation*.

Kepler, Johannes. *Astronomia nova aitiologetos seu physica coelestis, tradita commentariis de motibus stellæ Martis*. Pragæ, 1609. English translation by William H. Donahue: *New Astronomy*. Cambridge, U.K.: Cambridge University Press, 1992.

Klein, Felix. "Über die sogenannte Nicht-Euklidische Geometrie." *Mathematische Annalen* 4 (1871): 573–625.

Klein, Felix. "Über die sogenannte Nicht-Euklidische Geometrie (Zweiter Aufsatz)." *Mathematische Annalen* 6 (1873): 112–145.

Klein, Felix. *Vergleichende Betrachtungen über neuere geometrische Forschungen*. Erlangen, Germany: A. Duchert, 1872.

Kline, Morris. *Mathematics: The Loss of Certainty*. New York: Oxford University Press, 1980.

Lambert, Johann Heinrich. "Theorie der Parallellinien." *Magazin für die reine und angewandte Mathematik* 2 (1786):

137–164; 3 (1786): 325–358. Reprinted in P. Stäckel and F. Engel, *Die Theorie der Parallellinien von Euklid bis Gauss*. Leipzig, Germany: B.G. Teubner, 1895, pp. 152–207; our quotation refers to this edition.

Locke, John. *An Essay concerning Humane Understanding*. London: Thomas Basset, 1690.

Lotze, Hermann. *Metaphysik: Drei Bücher der Ontologie, Kosmologie und Psychologie*. Leipzig, Germany: S. Hirzel, 1879.

Minkowski, Hermann. "Raum und Zeit." *Physikalische Zeitschrift* 10 (1909): 104–111.

Newton, Isaac. *Arithmetica universalis, sive De compositione et resolutione arithmetica liber*. Cambridge, England: Typis academicis, 1707.

Pasch, Moritz. *Vorlesungen über neueren Geometrie*. Leipzig, Germany: Teubner, 1882.

Plato. *Laches - Protagoras - Meno - Euthydemus*, with an English translation by W. R. M. Lamb. Cambridge MA: Harvard University Press, 1990.

Plato. *Laws*, with an English translation by R. G. Bury. 2 vols. Cambridge MA: Harvard University Press, 1984.

Plato. *Republic*, with an English translation by Paul Shorey. Cambridge MA: Harvard University Press, 1982.

Poincaré, Henri. "Sur les hipothèses fondamentales de la géométrie." *Bulletin de la Société mathématique de France* 15 (1887): 203–216.

Poncelet, Jean-Victor. *Traité des propriétés projectives des figures*. Paris: Bachelier, 1822.

Resnik, Michael D. *Mathematics as a Science of Patterns*. Oxford, U.K.: Clarendon Press, 1997.

Riemann, Bernhard. "Über die Hypothesen, welche der Geometrie zugrunde liegen." *Abhandlungen der Kgl. Gesellschaft der Wissenschaften zu Göttingen* 13 (1867): 133–152. For English translation, see under Spivak.

Russell, Bertrand. *An Essay on the Foundations of Geometry*. Cambridge, U.K.: Cambridge University Press, 1897. Unaltered reprint: New York, Dover, 1956.

Saccheri, Girolamo. *Euclides ab omni nævo vindicatus*. Mediolani: e typographia Pauli Antonii Montani, 1733. Reprinted as *Euclides vindicatus*, with facing English translation by G. B. Halsted, New York: Chelsea, 1986.

Simplicius. *In Aristotelis de Caelo Commentaria*, edidit I. L. Heiberg. Berlin: G. Reimer, 1894.

Spinoza, Benedictus. *Ethica ordine geometrico demonstrata*, in Opera posthuma. Amstelodami: J. Rieuwertsz, 1677.

Spivak, Michael. *A Comprehensive Introduction to Differential Geometry*. 2nd ed., Berkeley: Publish or Perish. 5 vols, 1979. Includes an excellent English translation, with mathematical commentary, of Riemann's lecture "On the hypotheses that lie at the foundation of geometry"; see Vol. 2, pp. 135ff.

SUGGESTIONS FOR FURTHER READING

Aaboe, Asger. *Episodes from the Early History of Mathematics*. New York: Random House, 1964.

Blumenthal, Leonard M. *A Modern View of Geometry*. San Francisco: W. H. Freeman, 1961.

Boi, Luciano. *Le problème mathématique de l'espace: Une quête de l'intelligible*. Berlin: Springer, 1995.

Bonola, Roberto. *Non-Euclidean Geometry: A Critical and Historical Study of Its Development*. English translation with additional appendices by H.S. Carslaw. New York: Dover, 1955. Includes English translations of works by Bolyai and Lobachevsky.

Buroker, Jill Vance. *Space and Incongruence: The Origin of Kant's Idealism*. Dordrecht, Netherlands: D. Reidel, 1981.

Coxeter, Harold S. M. *Introduction to Geometry*. 2nd ed. New York: John Wiley and Sons, 1969.

Descartes, René. *The Geometry*. Translated from the French and Latin by D. E. Smith and M. L. Latham, with a facsimile of the first edition of 1637. New York: Dover, 1956.

Ellis, George F. R., and Ruth M. Williams. *Flat and Curved Space-Times*. Oxford, U.K.: Clarendon Press, 1988.

Heath, T. L. *The Thirteen Books of Euclid's Elements*. Translated from the text of Heiberg with introduction and commentary. New York: Dover. 3 vols.

Heilbron, J. L. *Geometry Civilized: History, Culture, and Technique*. Oxford: Clarendon Press, 1998.

Hilbert, David and S. Cohn-Vossen. *Geometry and the Imagination*. Translated from the German book *Anschauliche Geometrie* by P. Nemenyi. New York: Chelsea, 1952.

Modenov, P. S. and Parkhomenko, A. S. *Geometric Transformations*. Translated and adapted from the first Russian edition by M. B. P. Slater. 2 vols. New York: Academic Press, 1965.

Nagel, Ernest. "The Formation of Modern Conceptions of Formal Logic in the Development of Geometry." *Osiris* 7 (1939): 142–224.

Neugebauer, Otto. *The Exact Sciences in Antiquity*. 2nd ed. New York: Dover, 1969.

Rosenfeld, B. A. *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*. Translated from the Russian by A. Shenitzer. New York: Springer, 1988.

Torretti, Roberto. *Philosophy of Geometry from Riemann to Poincaré*. Dordrecht, Netherlands: D. Reidel, 1978.

Roberto Torretti (2005)

GERARD, ALEXANDER

(1728–1795)

Alexander Gerard was professor of moral philosophy and divinity at the University of Aberdeen and a leading member of the Aberdeen Philosophical Club along with James Beattie and, most importantly, Thomas Reid. He is known primarily for his *Essay on Taste* (1759/1963), which was awarded a prize by the Edinburgh Society for the Encouragement of Arts, Sciences, Manufacture, and Agriculture. Gerard returned to the subject with *An Essay on Genius* (1774/1966). In addition to the primary influence of Reid, the work of David Hume is a principal influence, though, like Reid, Gerard disagrees fundamentally with what he takes to be Hume's skepticism.

Although Gerard writes in the tradition of eighteenth-century theories of taste, it is questionable whether he should be regarded as a taste theorist in a

strict sense. Gerard is responding to the theories of criticism of Francis Hutcheson and Hume who set the context of discussion in terms of taste and sentiment, but Gerard follows Reid in taking a more realist position regarding the qualities that produce a perception of beauty and relies more directly on rules and principles that are derived by induction. Thus, Gerard defends a position that is moving rapidly away from an essential dependence on taste.

Gerard depends on two fundamental principles. The first is a faculty of imagination. Imagination combines reflective ideas supplied by fancy. Gerard's faculty psychology posits internal senses that are "*reflexive*," that is, they refer to the workings of the mind rather than to external objects. However, whereas for Hutcheson an internal, reflexive sense is a direct intuition of beauty and virtue, Gerard, following Reid, treats internal senses as active principles of perception. Internal senses correspond to the qualities that they respond to. For example, there are senses of novelty, sublimity, beauty, imitation, harmony, ridicule, and virtue. The second fundamental principle, following Joseph Addison's *Spectator* essays (particularly no. 418), identifies the pleasures of the imagination as depending on mental activity. The faculty of imagination exercises the mind; and, when that exercise falls within a moderate range, it is experienced as pleasurable. If it is either too languid and easy or too excited and difficult, discomfort (or simply indifference) results. These two principles combine to explain judgments of taste.

The subordination of taste to imagination seems clear; for example in *Essay on Taste*, Gerard writes "Taste, therefore, though itself a species of sensation, is in respect of its principles, justly reduced to imagination" (1759/1963, p. 144); and later, "Taste, in most of its forms at least, [is] a derivative and secondary power. We can trace it up to simpler principles, by pointing out the mental process that produces it, or enumerating the qualities by the combination of which it is formed. These are found, on inquiry, to be no other than certain exertions of *imagination*" (1759/1963, p. 151). Gerard goes on to explain each of the aesthetic predicates in terms of the kind of pleasurable mental activity that they produce: "The sources of all the sentiments of taste ly [sic] in the mind. The qualities of objects affect, in a certain manner, some principles of human nature, which by their operation, either singly or several in conjunction, produce gratification or disgust. … Simplicity, for instance, occasions easiness of conception; novelty or variety, an effort to conceive; amplitude, an expansion of soul" (1759/1963, p. 260).

Gerard holds that sentiment can be judged false because the qualities of taste can be figured out empirically. If I perceive something as grand that lacks the necessary qualities of extensiveness and amplitude, I am mistaken in my sentiment just as I would be if I experienced motion in violation of its actual occurrence. Therefore, for Gerard, there can be only a limited appeal to sentiment: "the qualities of an object, which gratify us, are more fixed and definite than the sensation which they excite" (1759/1963, p. 288). Gerard is clearly committed to what he understands as a "scientific"—that is, Newtonian—model, but at bottom he is siding with Reid against Hume by holding that aesthetic properties must be really in the object and that principles of common sense are sufficient to provide standards of judgment when disagreement arises.

For both Reid and Gerard, active judgment is logically prior to sensory experience in the aesthetic process. The function of sensory experience is to supply the material; the aesthetic operation comes about only when the mind is actively engaged. For example, Gerard writes, "For all the objects that affect taste, and excite its sentiments, are certain forms or pictures made by fancy, certain parts or qualities of things which it combines into complex modes" (1759/1963, p. 157). "In order, therefore, to form an able critic, taste must be attended with a philosophical genius, which may subject these materials to a regular induction, reduce them into classes, and determine the general rules which govern them" (1759/1963, p. 171). That engagement is critical and judgmental. Gerard's theory points toward Archibald Alison's *Essay on the Nature and Principles of Taste* (1790) in that Alison, too, reduces taste to a form of mental activity. But when Gerard says that "The sources of all the sentiments of taste lie in the mind" (1759/1963, p. 290), his purpose is to deny Hume's division between external sense and passions and to side with Reid's dualism between mind and body. Gerard's theory of taste marks the beginning, therefore, of a break with the theories of taste that run from The Third Earl of Shaftesbury (Anthony Ashley Cooper) through Addison and Hutcheson to Hume.

See also Aesthetics, Problems of; Alison, Archibald; Beattie, James; Hume, David; Hutcheson, Francis; Newton, Isaac; Reid, Thomas; Shaftesbury, Third Earl of (Anthony Ashley Cooper).

Bibliography

WORKS

An Essay on Taste (1759). Facsimile of 3rd ed. (1780). Gainesville: Scholars' Facsimiles and Reprints, 1963.

An Essay on Genius (1774). Reprint edited by Bernhard Fabian, Munich: Wilhelm Fink Verlag, 1966.

ADDITIONAL SOURCES

Addison, Joseph. "The Pleasures of the Imagination." *The Spectator* 409 (1712): 411–421. In *Eighteenth-Century British Aesthetics*, edited by Dabney Townsend (New York: Baywood, 1999), 107–136.

Dickie, George. *The Century of Taste: The Philosophical Odyssey of Taste in the Eighteenth Century*. Oxford: Oxford University Press, 1996.

Gracyk, Theodore A. "Kant's Shifting Debt to British Aesthetics." *British Journal of Aesthetics* 26 (1986): 204–217.

Hipple, Walter J. *The Beautiful, the Sublime, and the Picturesque in Eighteenth-Century British Aesthetic Theory*. Carbondale: Southern Illinois University Press, 1957.

Kivy, Peter. *The Seventh Sense: Francis Hutcheson and Eighteenth-Century British Aesthetics*. New York: Oxford University Press, 2003.

McCosh, James. "Alexander Gerard." In *The Scottish Philosophy, Biographical, Expository, Critical, from Hutcheson to Hamilton*. New York: Robert Carter, 1875. Reprinted, Hildesheim: Georg Olms Verlagsbuchhandlung, 1966.

Stolnitz, Jerome. "'Beauty': Some Stages in the History of an Idea." *Journal of the History of Ideas* 22 (1961): 185–204.

Townsend, Dabney. "From Shaftesbury to Kant." *Journal of the History of Ideas* 48 (1987): 287–305.

Dabney Townsend (2005)

GERBERT OF AURILLAC

(c. 938–1003)

Gerbert of Aurillac, an educational reformer and pope (Silvester II) of the eleventh century, was born in Auvergne about 938, became a monk of St. Gérard d'Aurillac, and was educated there and in Catalonia. He later visited Rome, where Pope John XIII (965–972) introduced him to Emperor Otto I. Gerbert shortly left Otto's court to study at Rheims, where he later became master of the schools. His fame led Otto II to make him abbot of Bobbio about 980, but in 983 Gerbert returned to Rheims, where he engaged in political and antipapal controversies. In 991 he became archbishop, and after many vicissitudes he was transferred to Ravenna in 998. His old friend and patron Otto III secured his election as Pope Silvester II in 999; as pope he established the church in Hungary and strongly asserted papal claims. He died in 1003.

Despite the intrigues and restlessness of his later public life, Gerbert was—and was recognized as—the most learned, versatile, and influential master of his age. Rheims during his first stay (c. 966–980) became a principal center of the educational revival that was beginning to inspire the cathedral schools of France and that from them passed to the universities. Fulbert, founder of the school of Chartres, was Gerbert's pupil.

Gerbert's greatest achievement was to give new life to the skeleton of the ancient *trivium* and *quadrivium*. In rhetoric he restored the careful study of Terence and Vergil, the satirists Horace and Persius, Lucan, and the critics Seneca and Quintilian; in dialectic, which he reestablished as the goal of a literary education, he developed what was to become the classical syllabus of the *Isagoge* of Porphyry, the *Categories* and *De Interpretatione* of Aristotle, the *Topics* of Cicero, and the whole dialectical corpus of Boethius. He rescued the *quadrivium* from its bookish decadence and injected a real, practical orientation. In mathematics, his forte, Gerbert revived the ancient Greek tradition and replaced clumsy Roman numerals with the Indian numerals 1 through 9; he produced a simplified abacus, with instructions for its use; and he wrote at length on methods of multiplication and division. In astronomy he taught by means of a sphere showing the movements of the planets.

It is uncertain how much these innovations were the result of his early experiences in Spain and his contacts there with Arabic science and thought. Save for a short disputation on human reason, in which he showed an attraction toward the Platonic Ideas, he wrote no philosophical work. His only authentic scientific writings are mathematical. His letters, some of which contain discussions of mathematics, illustrate his political activity and the events of his age. In his later life he had little influence on the intellectual and spiritual life of his age. His earlier work as a teacher, however, marked an epoch.

See also Aristotle; Boethius, Anicius Manlius Severinus; Mathematics, Foundations of; Philosophy of Education, History of; Porphyry; Seneca, Lucius Annaeus.

Bibliography

WORKS BY GERBERT

"Opera." Edited by A. Duchesne. In *Patrologia Latina,* edited by J. P. Migne, Vol. 139, pp. 201–264. There is a more recent version edited by A. Olleris. Clermont-Ferrand, 1867.

Gerberti Opera Mathematica. Edited by N. Bubnov. Berlin, 1889.

Lettres de Gerbert. Edited by J. V. Havet. Paris: A. Picard, 1889. The best Latin text. Translated by H. P. Lattin as *The Letters of Gerbert with His Papal Privileges as Sylvester II.* Records of

Civilization. Sources and Studies, Vol. 60. New York: Columbia University Press, 1961.

WORKS ON GERBERT

Amann, E. "Silvestre II." In *Dictionnaire de théologie catholique.* Paris, 1939. Vol. XIV, cols. 2075–2083.

Leflon, J. *Gerbert: Humanisme et chrétienté au X^e siècle.* Paris: Editions de Fontenelle, 1946.

Picavet, F. *Gerbert, un pape philosophe.* Paris: E. Leroux, 1897.

Uhlirz, Mathilde. *Untersuchungen über Inhalt und Datierung der Briefe Gerberts von Aurillac.* Schriftenreihe der historischer Kommission bei dem bayerischen Akademie der Wissenschaft, Vol. II. Göttingen: Vandenhoeck & Ruprecht, 1957.

David Knowles (1967)

GERSHON, LEVI BEN

See *Gersonides*

GERSON, JEAN DE

(1363–1429)

Jean de Gerson was one of the most influential French intellectuals of the early fifteenth century. He studied under Pierre d'Ailly and received his doctorate in theology in 1392. He was elected the chancellor of the University of Paris in 1395. He used this key position for intense intellectual involvement in ecclesiastical politics. He was part of the University of Paris delegation to the Council of Constance and played an important role in the discussions there. After the council, he fell from political favor. Returning to France in 1419, he lived in Lyon for the rest of his life in relative obscurity. He was then engaged mainly in literary work, producing, for example, the well-known defense of Joan d'Arc.

During Gerson's lifetime, the emphasis in university work turned from research to teaching and social influence. The saying "everything necessary has already been written" was often used at the time, and accordingly university teaching was often directly based on canonical sources. Gerson was an active figure in developing the university away from "idle speculations" and toward applying learning for the larger world. His own philosophical work cannot be described as highly original. But he was very productive and very influential through his writings on popular topics.

In political philosophy, Gerson was close to his master Pierre d'Ailly. They worked in close cooperation on many issues. As *conciliarists*, they understood the church as a political society. Thus, they thought that a general council of the church would have the power to solve the papal schism, like in any political society the ruler may rightly be deposed if he fails to promote the welfare of the society. Gerson cannot be said to have promoted individual rights because he did not understand the welfare of society in terms of the welfare of the individual.

Gerson has been called both an opponent and a proponent of the nominalist movement of his time. In many contexts, he relied on nominalist positions. He was, however, an opponent of the idea that natural reason could solve metaphysical problems. Also, he acted with the Renaissance humanists against the increasing role of logic and natural reason in the theological faculties. This was a time in which the English tradition in nominalist logico-semantical work was gaining ground in continental universities, especially among the Scotists and the Ockhamists. Later on, achievements in this field were to prove crucial in the formation of what is today known as modern science. Gerson's opposition to this increasing emphasis on logico-semantical analysis in the theological faculties was not so much due to a disagreement about philosophical issues so much as a preference for what he saw as more applicable and experientially grounded knowledge.

Instead of speculative theology, Gerson encouraged mystical theology, and indeed many of his best known writings are from this field. His approach is that it is the duty of every person to acquire experiential knowledge of God. This did not mean a rejection of philosophical learning. Rather, Gerson sought for mutual support between devotion and learning. In his anthropological writings, he presents a threefold division both of cognitive potencies—simple understanding, reason, and sensitivity (*intelligentia simplex*, *ratio*, and *sensualitas*)—and of affective potencies—conscience, rational desire, sensitive desire (*synderesis*, *appetitus rationalis*, and *appetitus sensualis*). These divisions accord with neoplatonic models, but Gerson's special emphasis is upon the reciprocal relations between the affective and the cognitive powers. They must work together so that knowledge and love both contribute to the approach to God. In this way, the *unio mystica* can be achieved. Gerson says very little about that experience itself, claiming that it is known only through experience and cannot be described.

Gerson was a typical fifteenth-century Renaissance intellectual. He was deeply religious and committed his efforts to public affairs, concentrating on the papal schism and his duties at the University of Paris. Apart from the writings on mysticism, his philosophical views

are best understood in terms of the ecclesiastical situation and his position in the university politics of the time.

See also Ailly, Pierre d'; Bonaventure, St.; Luther, Martin; Platonism and the Platonic Tradition; Pseudo-Dionysius; Ruysbroeck, Jan van; Scotism; Thomism.

Bibliography

WORKS BY JEAN DE GERSON

Ouevres completes, edited by P. Glorieux. 10 vols. Paris: Desclée & Cie, 1960–1973.

Early Works. Translated by B. P. McGuire. Mahwah, NJ: Paulist Press, 1998.

The Consolation of Theology. Translated by Clyde L. Miller. New York: Abaris Books, 1998.

WORKS ABOUT JEAN DE GERSON

Brown, D. C. *Pastor and Laity in the Theology of Jean Gerson*. Cambridge, U.K.: Cambridge University Press, 1987.

Burger, C. *Aedificatio, Fructus, Utilitas: Johannes Gerson als Professor der Theologie und Kanzler der Universität Paris*. Tübingen: Mohr, 1986.

Ryan, John J. *The Apostolic Conciliarism of Jean Gerson*. Atlanta, GA: Scholars Press, 1998.

Mikko Yrjönsuuri (2005)

GERSONIDES

(1288–1344)

Gersonides, or Levi ben Gershon, also known by his acronym, "RaLBaG," was a French Jewish philosopher, biblical exegete, mathematician, and astronomer. He was born at Bagnols and died at Perpignan. He was the inventor of two astronomical instruments, the Jacob's staff ("baculus") and an improved camera obscura. Gersonides' literary contributions include biblical commentaries of a philosophical and moral tone, supercommentaries to Averroes's treatises on Aristotle, and his philosophical masterwork, *Milhamot Adonai* (Wars of the lord). Because of his knowledge of Averroes, Gersonides was exposed to a more authentic version of Aristotle than was available to his predecessors and was thus motivated to reexamine certain problems that he felt had previously been treated inadequately or incorrectly. These problems, corresponding to the six sections of the "Wars," are (1) the nature and immortality of the soul, (2) prophecy, (3) the nature of God's knowledge, (4) divine providence, (5) miracles and the structure of the universe, and (6) the creation of the world. Methodologically, he recognized the authority of the four roots of knowledge (as first formulated by Saadya Gaon), namely, reason, sensory perception, divine revelation, and rabbinic tradition, in that order of priority, although he seldom cited the last specifically.

The work begins with a detailed analysis of Aristotle's doctrine of the soul according to the interpretations of Alexander of Aphrodisias, Themistius, and Averroes. In agreement with Alexander, Gersonides maintained that the material or hylic intellect is a capacity inherent in the sensitive soul. Under the agency of the Active Intellect, the last of the separate intelligences, the material intellect is transformed, through the acquisition of ideas, into an actual or acquired intellect. Opposing the nominalism of Alexander and Maimonides, Gersonides maintained the reality of the ideational content of the acquired intellect. It is this acquired intellect that survives independently after the death of the individual.

Gersonides' account of the nature of God's knowledge is related to his theory of divine attributes. Maimonides' theory of homonymy, according to which attributes in general and the term *knowing* in particular refer to entirely different concepts when applied to God and man, allowed Maimonides to maintain both God's absolute omniscience and human free will. Rejecting this as absurd, Gersonides reaffirmed, in agreement with the Muslim philosophers Avicenna and Averroes, that attributes are to be treated as ambiguous terms, applied in a primary sense to God but in a derivative sense to man. Furthermore, the attributes imply no plurality in God's nature since they are subjects of discourse and not of essence, just as the terms *redness* and *color* imply no plurality in the concept "red."

Since God's knowledge is similar in nature to man's, he cannot know the contingent and consequently knows the particulars only insofar as they are ordered. This amounts to a virtual restriction of divine knowledge to the universals. Since men are endowed with free will, this restriction normally precludes special providence for individuals. However, some individuals enjoy special providence; this consists in a knowledge, received from the Active Intellect, of stellar configurations that determine events on earth. Modern scholarship has not generally noted that this explanation of special providence for the intellectual elite was foreshadowed in one of the two discussions of the problem in Maimonides' *Guide for the Perplexed* (III.51).

The communication of astrological information to the human intellect by the Active Intellect is known as prophecy. The prophet, to the extent of his ability, interprets the general information received, in the light of the particular circumstances with which he is concerned,

Gersonides' tendency to deny God's direct involvement in terrestrial affairs is further illustrated by his theory that the capacity for miracles was implanted in nature so that miracles do not represent any specific divine concern.

In his discussion of the origin of the world, Gersonides agreed with Maimonides that it was indeed created but, in opposition to him, maintained that *ex nihilo nihil fit.* Rather, he posited an absolutely formless matter (not eternal in time since time did not exist before the creation of the world) out of which the world was formed. Gersonides found this dualism useful in ascribing the origin of evil to matter.

See also Alexander of Aphrodisias; Aristotle; Averroes; Avicenna; Jewish Averroism; Jewish Philosophy; Maimonides; Saadya; Themistius.

Bibliography

Gersonides' *Milhamot Adonai* (Wars of the lord) was first printed at Riva di Trenta (1560) and reprinted at Leipzig (1866). The editions omit the rather technical astronomical treatise in the fifth book which is extant in manuscript and which was translated into Latin by order of Clement VI. The partial German translation by Benzion Kellerman, *Die Kämpfe Gottes von Lewi ben Gerson* (Berlin, 1914), should be read in conjunction with Isaac Husik's review, "Studies in Gersonides," in *Jewish Quarterly Review,* n.s., 8 (1917–1918): 113ff. and 231ff. For a complete bibliography, see *Encyclopedia Judaica* (Berlin, 1929), Vol. IV, p. 656; Vol. VII, pp. 328ff. and 338, to which may be added H. Wolfson, "Maimonides and Gersonides on Divine Attributes as Ambiguous Terms," in *Mordecai Kaplan Jubilee Volume,* edited by M. Davis, 515–530. (New York: Jewish Theological Society, 1953).

Frank Talmage (1967)

GERSONIDES [ADDENDUM]

The intensive and fruitful research conducted on Gersonides since the 1960s has increasingly led to a recognition of his exceptional stature as a philosopher-scientist. It is also appreciated now that the multifarious aspects of Gersonides' thought constitute a coherent unity that must be studied as such. Contrary to Maimonides, whom he venerated, Gersonides believed in the human power to discover God's blueprint for creation, namely through empirical study of nature and through the interpretation of God's revelation (the Scriptures). This belief triggered his various scientific activities, most of which have been the subject of recent research.

Following notably Charles Touati's French monograph on Gersonides' philosophy, many detailed studies have been devoted to specific issues. Several parts of the *Wars of the Lord* have been translated into English and French, followed by a full English translation by Seymour Feldman. The great significance and originality of the largely unpublished astronomical part of Gersonides' *Wars* (often referred to as his *Astronomy*) have been brought to light, notably by Bernard R. Goldstein. In addition, J. L. Mancha has shown that the Latin translation of the *Astronomy* had been realized with the collaboration of Gersonides himself at the papal court in Avignon, testifying to the importance ascribed to it there. (For these and all other bibliographical references see Kellner's "Bibliographia gersonideana.")

Scholars realize now that Gersonides' numerous super-commentaries on Averroes's commentaries on Aristotle's treatises are crucial for an adequate understanding of Gersonides' thought and its evolution. Recent scholarship (R. Glasner) established that Gersonides was the first to teach Averroes's commentaries—in their Hebrew versions—in the Jewish community of southern France. Recent editions of Gersonides' commentary of the Pentateuch unearthed new material, revealing that Gersonides was more interested in Jewish law than had previously been thought. More generally, the great importance of the biblical commentaries for the understanding of Gersonides' thought has become better appreciated.

Scholars—notably Ruth Glasner—point to the originality of Gersonides' thinking about natural philosophy and its central place in his work. The originality of Gersonides' work in logic has also been highlighted, notably by Charles Manekin. Other studies, especially by Sara Klein-Braslavy, emphasize the importance of Gersonides' scientific methodology. Gersonides depended essentially on literature available in Hebrew; the question why his work bears next to no trace of a familiarity with Latin philosophy—even though Gersonides is known to have had contacts with Christians—has been much, albeit inconclusively, discussed in recent research.

See also Aristotle; Averroes; Maimonides; Revelation; Scientific Method.

Bibliography

A comprehensive bibliography of works by and on Gersonides is given in Menachem M. Kellner's "Bibliographia gersonideana" published in two parts: (1) *Studies on Gersonides—A Fourteenth-Century Jewish Philosopher-Scientist*, edited by Gad Freudenthal, pp. 367–414 (Leiden,

Netherlands: Brill, 1992); and (2) *Aleph: Historical Studies in Science and Judaism* 3 (2003): 345–374.

ENGLISH TRANSLATIONS

ben Gershom, Levi (Gersonides). *The Wars of the Lord*. Translated with introduction and notes by Seymour Feldman. 3 vols. Philadelphia: The Jewish Publications Society, 1984, 1987, 1999. A full translation of the *Wars* (except the *Astronomy*).

Goldstein, Bernard R. *The Astronomy of Levi ben Gerson (1288–1344): A Critical Edition of Chapters 1–20 with Translation and Commentary*. Berlin: Springer, 1985. A partial translation of the *Astronomy*.

Manekin, Charles. *The Logic of Gersonides: A Translation of the Sefer ha-Heqqesh ha-Yashar*. Dordrecht, Holland: Kluwer, 1992. A full translation of the treatise on logic.

STUDIES

Freudenthal, Gad. "Levi ben Gershom (Gersonides), 1288–1344." In *The Routledge History of Islamic Philosophy*. Edited by S. H. Nasr and O. Leaman, pp. 739–754. London and New York: Routledge, 1996. Provides a concise overview.

MONOGRAPHS

Touati, Charles. *La pensée philosophique et théologique de Gersonide*. Paris: Éditions de Minuit, 1973.

COLLECTIONS OF STUDIES

Dahan, Gilbert, ed. *Gersonide en son temps. Science et philosophie médiévales*. Louvain-Paris: E. Peeters, 1991.

Freudenthal, Gad, ed. *Studies on Gersonides—A Fourteenth-Century Jewish Philosopher-Scientist*. Leiden, Netherlands: Brill, 1992.

Sirat, Colette, Sara Klein-Braslavy, and Olga Weijers, eds., *Les méthodes de travail de Gersonide et le maniement du savoir chez les scolastiques*. Paris: Vrin, 2003.

Gad Freudenthal (2005)

GESTALT THEORY

The Gestalt movement in psychology began early in the twentieth century; its founders were the German psychologists Max Wertheimer, Wolfgang Köhler, and Kurt Koffka. A Gestalt is essentially an organized whole whose parts belong together, as opposed to being simply juxtaposed or randomly distributed. As Wertheimer put it, "What happens to a part of the whole is determined by intrinsic laws inherent in this whole." The Gestalt theorists believed this principle to be of wide application and to be relevant to the psychology of perception in particular.

HISTORY

As early as 1890 Christian von Ehrenfels had pointed out that to appreciate a melody we need to be aware not of single tones in isolation but of a succession of tones which combine in a particular way. If notes of the same pitch as those of the original melody are presented in a different temporal order, there will be a completely different effect, whereas the same melody played in a different key is immediately recognizable, even though the notes are different in pitch from the original ones. The melody as a whole was said by von Ehrenfels to have a *Gestaltqualität* independent of the qualities of the separate notes. Wertheimer, Köhler, and Koffka were concerned to apply the concept of Gestalt over a wide area and thus give a new direction to psychological research.

FIGURE 1

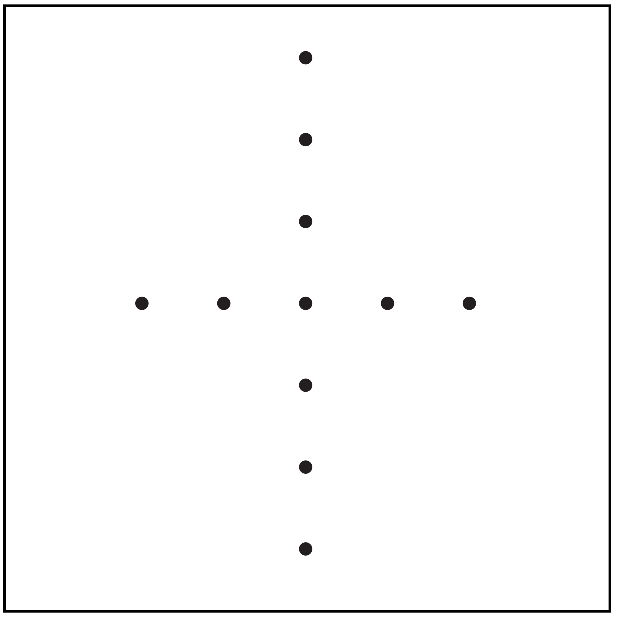

A central feature in their view was the doctrine of isomorphism, which asserts that our experiences have the same structure as the brain processes which underlie them. Thus, if the stimulus is a nearly complete circle which the subject sees as a complete circle the doctrine of isomorphism would assert that there must be some pattern in the brain that is isomorphic with the complete circle, as opposed to the incomplete one. The detailed neurological hypotheses which they put forward are of questionable value, but the general principle is still of interest, as is Köhler's demonstration that there are Gestalten in physical nature, for example, the soap bubble, whose spherical shape is the necessary result of the total forces in operation at any one time (see Koffka, *Principles of Gestalt Psychology,* p. 14).

The following are some typical examples of the Gestalt principle as applied to vision: Figure 1 appears as a cross; if, however, we consider the effect on the retina of

FIGURE 2A

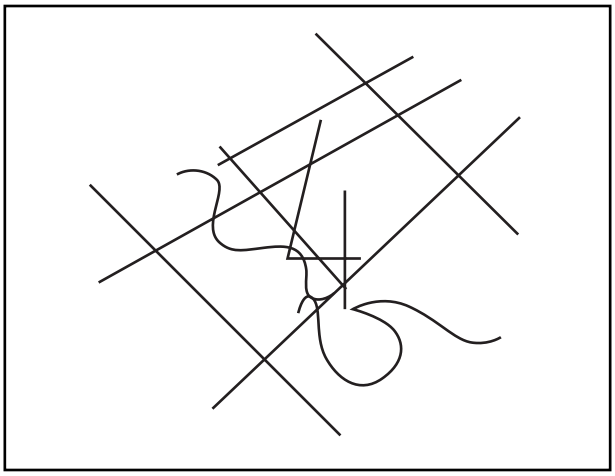

FIGURE 2B

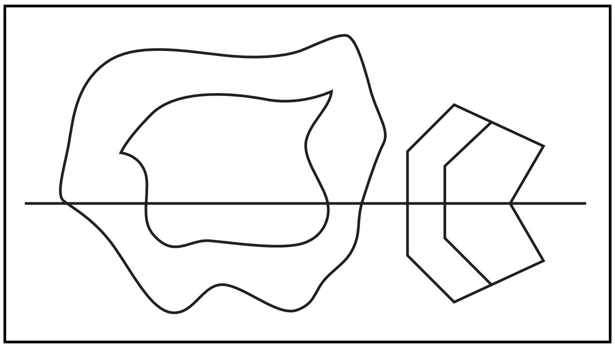

the eye of each of the dots in isolation, there is nothing to account for the way in which they are organized. Implicit in traditional thinking, according to Köhler and Koffka, is the so-called constancy hypothesis—the hypothesis that stimulation of a particular point on the retina has a constant effect regardless of the total pattern of stimulation. Yet if the constancy hypothesis were true, it would be hard to explain the obvious recognizability of the "4" in Figure 2a and its camouflage in Figure 2b, since the same retinal points are being stimulated in both cases. Similarly one cannot explain how a person who moves from twenty yards away to ten yards away continues to look approximately the same size, since the retinal stimulation must by the laws of optics be quite different. Indeed there are many characteristics of the perceived world (what Koffka terms the "behavioral," as opposed to the "geographical," environment) which do not bear a one-one relationship to anything in the pattern of stimulation. Thus, in Figure 3 we see the lines as four pairs, but the "togetherness" of each pair has no direct counterpart in the system of stimuli; and in Edgar Rubin's famous example (Figure 4), whether we see the white as "figure" and the black as "ground" or vice versa, there is no direct counterpart to the "thinglike" character of the figure and the absence of this character in the ground.

This thinglike character, the character of "productivity," which occurs in some causal transactions, the character of "being mine," which belongs to, say, my hand in contrast to an object on which my hand is resting, the character of anger present in someone's face—these are some of the many features which are present in the behavioral environment, even though they are necessarily absent from the world of physics because there is nothing in the stimulus situation directly corresponding to them.

On the Gestalt view what is all-important is the way in which the immediate, or "proximal," stimuli (for example, light waves or sound waves) combine in space and time; when these combinations are of a certain kind, certain perceptual organizations will arise (for example, two parts of a diagram will be seen as belonging together), and laws can be framed in terms of which such organizations can be predicted. Details of these laws have been set out by David Katz; they include the law of proximity, which states that, other things being equal, in a total stimulus situation those elements which are closest to each other tend to form groups, and the law of closed forms, which states that, other things being equal, lines which enclose a surface tend to be seen as a unit.

A law of a more general kind is that of *Prägnanz.* As formulated by Koffka (*Principles of Gestalt Psychology*, p. 110), this law states: "Psychological organisation will always be as 'good' as the prevailing conditions allow. In this definition the term 'good' is undefined. It embraces such properties as regularity and symmetry, simplicity and others." In other words, when the stimuli are of a certain kind, there are forces within the organism that operate in the direction of maximum simplicity; hence, we tend to see "good" figures—squares and circles, for example—rather than less regular ones. The word *Prägnanz* is of course ultimately connected with the Latin *impregnare.* The suggestion here, however, is not that of something being fertilized or made pregnant but rather of something being stamped or pressed into a particular shape (compare the word *prägen,* which is used primarily to refer to the minting of coins). Certain types of configurations, one might say, are particularly impressive; they carry a certain stamp or they strike us in particular ways.

Contrary to what has sometimes been said, the Gestalt psychologists did not dispute that past experience can influence perception; this is made plain by Katz in *Gestalt Psychology* (pp. 28–29). Their criticism was

FIGURE 3

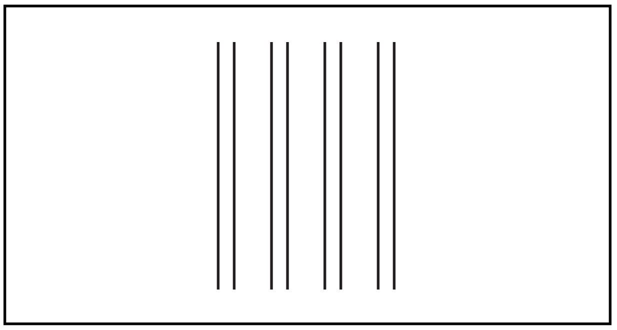

directed against the view of perception that invoked past experience as a deus ex machina when observed results did not fit the constancy hypothesis. Thus a penny, unless its flat surface is directly in front of us, might be expected from the laws of optics to look elliptical. Since it does not, one can preserve the view that the "basic datum" is an ellipse by postulating a rapid process of inference based on past experience. On this view we *infer* that the penny is round because of our alleged previous experience of round pennies. According to the Gestalt psychologists, however, not only is there nothing in introspection to suggest such an inference; they would also have questioned whether in fact it is particularly common in ordinary life for pennies to have their flat surfaces directly in front of the observer—a condition that their opponents' theory seems to require. In contrast, their view was that when the proximal stimuli combine in a certain way a particular perceptual organization is forced upon us; thus a circle is a "good" figure, and hence the "internal forces" will operate in the direction of a circle rather than an ellipse.

Interesting experimental studies include those of Wertheimer ("Experimentelle Studien über das Sehen von Bewegung," in *Zeitschrift für Psychologie*) on the perception of movement, those of Köhler (*The Mentality of Apes*) on problem solving in apes, those of Wertheimer (*Productive Thinking*) and K. Duncker ("On Problem Solving") on problem solving in humans, that of Katz (*The World of Colour*) on the perception of color, that of Rubin (*Synsoplevede Figurer*) on the figure-ground distinction, those of Kurt Lewin (*Principles of Topological Psychology*), who has attempted to apply Gestalt principles to the study of social situations, and those of Albert Michotte (*La perception de la causalité*)—although he was not a member of the original group—on the conditions in which we receive an impression of causality. A recent interesting development is the attempt to relate figural goodness to the amount of information (in the mathematical sense) needed to specify a particular pattern or figure (see especially Fred Attneave, *Applications of Information Theory to Psychology,* p. 82). The problem of perceptual Gestalten has arisen in an acute form in the programming of computers to carry out pattern recognition (see, for instance, Kirsch, "Computer Interpretation of English Text and Picture Patterns").

FIGURE 4

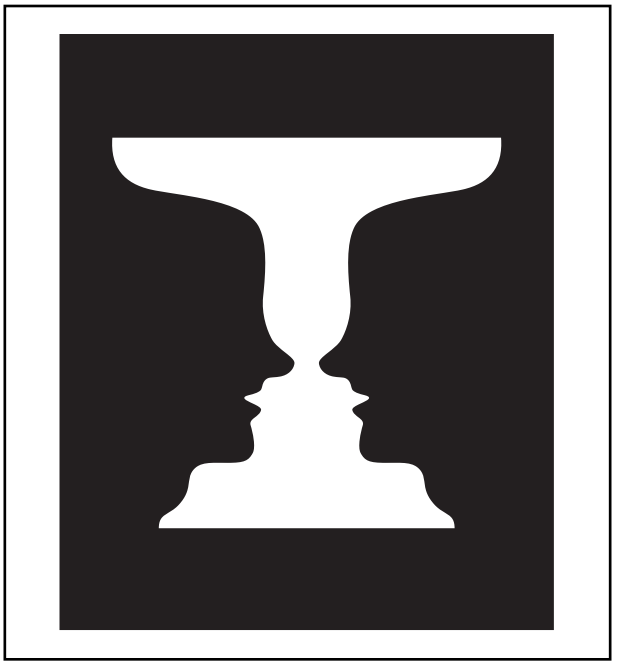

SCIENCE AND COMMON SENSE

The advance of science continually brings in its train a challenge to our commonsense beliefs about the world. At one time or another in the history of scientific thought it has been held, for example, that sense perception is unreliable, that the things around us are not really colored, that the floor on which we walk is not really solid, and that no two events are ever exactly simultaneous. In contrast with many other scientific systems, Gestalt theory involves the attempt to call us away from such paradoxes back to common sense; it invites us to consider the world as we in fact experience it, not as we might expect to experience it in the light of the latest scientific developments.

It does not, of course, follow, that philosophical paradoxes can be disposed of simply by pointing out that experience is in fact of such-and-such a kind. Thus, it is no argument against John Locke's account of substance or David Hume's account of causality to point out that

the behavioral environment is found by experience to consist of things in causal interaction, any more than it is an argument against the distinction between primary and secondary qualities to point out that we are aware of greenness as being in the grass and not in our heads. Similarly, a philosopher who wishes to defend the sense-datum terminology cannot be refuted simply by an appeal to Rubin's claim that what we perceive is organized into figure and ground and is not just "a mosaic of sense data." As a corrective against those who mistake the point of philosophical arguments, it may be helpful on occasions to make explicit exactly what we are aware of at the commonsense level, but this does not prove the philosophical arguments to be wrong.

Despite the emphasis on naive judgment, however, the Gestalt program does not involve an uncritical return to naive realism. Rather, its claim is that the gulf between what common sense tells us and what science tells us is not, after all, as great as might be supposed; the world of nature is *gestaltet* no less than are our experiences. Moreover, although in some of their discussions Köhler and Koffka speak in traditional terms about "the relation between mind and matter," their views do not fit easily into the traditional categories of interactionism, epiphenomenalism, and parallelism; indeed, like many modern philosophers, they are critical of a starting point which forces us to decide between theories couched in these terms.

POSITIVISM AND BEHAVIORISM

The prevailing scientific attitude of the time, which Koffka called "positivism," was mistaken, on the Gestalt view, because it allowed no place for the categories of meaning and value. The important fact for psychology is that the behavioral environment is organized—it is intelligible. Thus we are making *sense* out of a person's facial expression when we say that he is angry. Similarly, if a person listens to music, he is sometimes aware that a chord with the leading note (the seventh of the scale) at the top *requires* to be followed by the tonic chord of the original key; the cadence has its special meaning only if the second chord follows the first. This remains true even though such "requiredness," as Köhler terms it in *The Place of Value in a World of Facts,* can play no part in the world of physics. As physical science advances we are enabled to make continually more refined statements about the geographical environment; but in so doing, on the Gestalt view, we are in danger of losing sight of facts—those of the behavioral environment—which for the psychologist are of special interest. Koffka agreed that vitalism is "no solution but a mere re-naming of the problem"; but by taking seriously the concept of Gestalt one can, he held, be antimechanistic without being obscurantist.

There is also an attack on the allegedly "scientific" creed of behaviorism, whose development was almost contemporary with that of Gestalt theory. The term *behaviorism,* as Köhler understands it, implies a denial that there can be "a science of direct experience," either because there is no such thing as direct experience or because if it exists, it is not accessible to public scrutiny. In reply Köhler points out that no scientist can even begin to experiment unless he starts from his own experienced world. He also points out that one has as little or as much justification to be skeptical about the world of experience as one has to be skeptical about the world of physics; there is no good reason why the behaviorist should choose to ignore the world of experience while taking the world of physics on trust.

ISOMORPHISM

It is far from clear whether the doctrine of isomorphism constitutes a radically new discovery, as the Gestalt theorists supposed, or whether it is a somewhat high-sounding way of asserting the obvious. Most modern psychologists, if asked, would doubtless express the hope that complete explanations of perception and learning will eventually be found in terms of brain processes. If, therefore, the contribution of Gestalt theory is to be distinctive, clearly some more far-reaching claim must be involved.

Koffka expressly pointed out that Gestalt theory does not stand or fall with the correctness of a particular theory about perceived movement. According to Koffka its more general objective is to contribute to "the integration of value, life, and mind. … The Gestalt concept … cuts across the division of realms of existence, being applicable in each of them." That is, there are Gestalten in nature (for example, the soap bubble); there are Gestalten in the living brain; and there are Gestalten in our conscious experience. In traditional discussions about the relation between mind and body, according to Köhler, it was "tacitly assumed that only *microscopic* events in the cortex can be the correlates of mental life." In contrast, the doctrine of isomorphism invites us, in Koffka's words, to "think of the physiological processes not as molecular but as molar phenomena. … If they are molar, their molar properties will be the same as those of the conscious processes which they are supposed to underlie." As has already been pointed out, this is not just an answer within the context

of traditional mind-body dualism; it is an attempt to look at this whole family of problems afresh.

The distinction between the geographical and behavioral environments gives rise to difficulties of its own. One very reasonably asks, in the first place, what kind of duality is involved. Clearly, there are not two environments in the same sense that there are—or might be—two rooms in a country cottage. If a child specialist recommended a change of environment for a child, it would make no sense to reply "very well; we will change him from the geographical to the behavioral environment." One is reminded in this connection of Arthur Eddington's claim—which in fact involves the same kind of difficulty—that he is writing simultaneously at two tables. Second, if we take the idea of two environments at its face value, we are tempted to ascribe some kind of superior status or "reality" to one or the other. It is the geographical environment, according to Koffka, whose contents are "real": "The pen with which I am writing is a unit in my behavioural environment and so is the *real* pen in the geographical" (italics added). What is "real," however, does not apparently coincide with what is "given"; "every *datum* is a behavioural datum; physical reality is not a datum but a constructum." The suggestion that each of us somehow "constructs" a physical world out of his immediate experiences implies a phenomenalist view which in the last resort leads inevitably to solipsism. Clearly this was not Koffka's intention, but he gave no indication how such a conclusion can be avoided.

The important point, according to Köhler and Koffka, is that the concept of Gestalt cuts across these two different kinds of reality. In the words of Köhler, "Any actual consciousness is in every case not only blindly coupled to its corresponding psychophysical processes, but is akin to it in essential structural properties." The difficulty here is that *anything* can be regarded as "structurally akin" to anything else, provided enough rules are given. Many maps are structurally akin to landscapes in that they involve the same geometrical shapes; but if one is allowed sufficient rules for specifying what represents what, a map of England (as we now call it) could function without any misrepresentation as a map of France. What Köhler needs to argue for is some relatively clear-cut and uncomplicated structural relationship. Thus, it may well be that the shape of the areas stimulated in the cortex has something in common with the shape which we observe in an object, although it is hard to see how there can be any close parallel in the case of color, since, when *X* is looking at a green object, *Y* does not find anything green in *X*'s cortex. One must suppose that the use of the term *psychophysical* (instead of *physical*) to describe processes in the brain is intended to emphasize these relatively close structural similarities.

KNOWLEDGE OF OTHER MINDS

In reply to the charge of solipsism, Köhler and Koffka could point out that they both discuss the problem of knowledge of other minds at some length. The main feature of philosophical interest in these discussions is that structural similarities are pointed out between behavior which is noticed by others and so-called inner states, which are discriminated only by the person himself. A person's wincing may have precisely the same temporal properties as his twinges of pain, and the sound of his rising voice may have the same movement properties as his inner feelings of rising anger. Koffka pointed out that a character in a Mickey Mouse cartoon can quite well look exuberant or dejected; and if we can directly observe such exuberance or dejection in these cases, there seems to be no reason why we should not directly observe it in our friends. Similarly, Michotte has argued that visual experiences of live movement are structurally similar to kinesthetic and other experiences that we have when we make an effort. These considerations are not, of course, sufficient to remove all possible skeptical doubts; but they at least make clear the conditions in which we can justifiably say of a person that, for example, he is angry, dejected, or making an effort.

"SEEING" AND INFERENCE

Koffka suggested that the word *see* should be used in a "purely phenomenological" sense, that is, in such a way that the words which follow are a simple description of our experience. Thus, if we look at a table that is partly covered by a book, we are not aware of any gap or hole in the table in the area where the book is; and Koffka therefore wanted to say that what we *see* is a complete table. On his view it is necessary to describe our experience without being influenced by considerations of what we might expect to see on the basis of scientific knowledge—in this case, by our knowledge about the characteristics of the light waves striking the retina.

In ordinary usage, when we say that we or someone else saw something, this is normally not just the report of a visual experience; there is also an implicit claim to correctness. If the person is in fact deceived, then one is wrong to use the word *saw,* just as one would be wrong to say that a person has "proved" something if his argument contains a fallacy. One might therefore express Koffka's difficulty by saying that, as our knowledge of physics

increases, claims to *correctness* will force us further and further away from naive description. "Why," he asked rhetorically, "are we so hopelessly stupid as to call the colour of our table-cloth on the candle-lit dinner table white, when Helmholtz told us that it was yellow?" According to Koffka's proposed usage we actually see a white tablecloth, the evidence of the physicists notwithstanding.

This "purely phenomenological" sense of "see," however, is unnecessary. We already possess the expression "seeing-as" for situations where we do not wish to make claims to correctness. On this usage, we *see* the tablecloth *as* white even though we might have expected to see it as yellow. Similarly, a person who is exposed to an ambiguous diagram may *see* it *as* two-dimensional, even though on another occasion he sees it as three-dimensional. Unless one is an extreme skeptic and asserts that claims to correctness are never appropriate when visual perception is involved, it is surely useful to have a terminology which enables us to make such claims on some occasions and to withhold them on others.

Moreover, Koffka's proposed usage, if adopted consistently, carries the paradoxical consequence that we can never be sure or unsure about what we see, nor can we be right or wrong. If we look more carefully at something, or if others give us a verbal description or point to a contour line, we may, of course, see something new; but we were still neither right nor wrong about what we saw before.

There is the further paradoxical consequence that no two people can ever see the same thing. Köhler is apparently prepared to accept this, since he expressly tells us that no two scientific investigators ever see the same galvanometer. Moreover, if everything that we see (in Koffka's sense) is, by definition, part of the behavioral environment, this has the effect of turning the behavioral environment into a home for every erroneous perception which has been made. It is as bad as having to postulate false facts to ensure that false propositions refer to something. In general, the distinction between the geographical and behavioral environments involves many points of interest, but in philosophizing in this area, Koffka in particular was not successful in avoiding paradoxical consequences.

As far as psychology is concerned, the work of the Gestalt theorists has led to the discovery of a large number of new facts, particularly in the sphere of perception, and to a reinterpretation of facts which were already known. On the basis of the laws of structural organization, predictions can be made about what will be perceived when the proximal stimuli are of a particular kind, and these predictions normally work. The neurological explanations are inadequate by present standards, but even in this area the Gestalt theorists have at least called attention to problems which require to be solved, and, in particular, they have taken seriously the challenge presented by our ability to perceive spatial and temporal relationships.

On the broader theoretical issues, traditionally the province of philosophy, their main contribution has been to indicate the need for a change of emphasis. In the light of advancing scientific knowledge, it appeared to be the case that what we thought we were perceiving was not what we were really perceiving at all ("the grass is not really green," "the white tablecloth seen in candlelight is really yellow," and so on). For the Gestalt theorists, however, as for some modern philosophers, such claims were paradoxical and confused. Whatever physics tells us, the starting point, on the Gestalt theorists' view, must necessarily be the world as perceived by common sense; this is the world to which organisms respond, and it is therefore of special importance for the psychologist; moreover, if we ourselves did not perceive it in certain ways, physics could not even begin. This is to say, in effect, that language descriptive of ordinary experience can never be reduced to the language of physics—a thesis that has been held in many forms but which has seldom been defended in such a sustained and systematic way.

See also Behaviorism; Eddington, Arthur Stanley; Ehrenfels, Christian Freiherr Von; Hume, David; Koffka, Kurt; Köhler, Wolfgang; Locke, John; Perception; Phenomenological Psychology; Positivism; Psychology.

Bibliography

Attneave, Fred. *Applications of Information Theory to Psychology.* New York: Holt, 1959. For the relevance of information theory to Gestalt theory, see especially p. 82.

Duncker, K. "On Problem Solving." Translated by L. S. Lees. *Psychological Monographs* 58 (5) (1945). A short monograph with some interesting experiments on problem solving.

Ehrenfels, Christian von. "Über Gestaltqualitäten." *Vierteljarsschrift für wissenschaftliche Philosophie* 14 (1890): 249–292. Of historical interest.

Ellis, Willis Davis. *A Source Book of Gestalt Psychology.* London: K. Paul, Trench, Trubner, 1938. A helpful collection of the writings of Gestalt psychologists in one volume; recommended as a standard textbook.

Hamlyn, D. W. "Psychological Explanation and the Gestalt Hypothesis." *Mind* 60 (240) (1951): 506–520. A philosophical article in which, among other things, the concept of Gestalt is compared with that of substance.

Hamlyn, D. W. *The Psychology of Perception.* London: Routledge and Paul, 1957. Contains a critical examination of

some of the concepts of Gestalt theory in the light of recent British philosophy.

Hartmann, George W. *Gestalt Psychology; A Survey of Facts and Principles.* New York: Ronald Press, 1935. A readable account of the history of the movement and its discoveries up to 1935. Contains brief biographical details of Wertheimer, Köhler, and Koffka and a list of important dates. Recommended as a standard textbook.

Henle, Mary, ed. *Documents of Gestalt Psychology.* Berkeley: University of California Press, 1961. A collection of papers on Gestalt psychology by various authors, mostly dating from 1950 onward.

Katz, David. *Der Aufbau der Farbwelt.* Translated by R. B. Macleod and C. W. Fox as *The World of Colour.* London: K. Paul, Trench, Trubner, 1935. Interesting discussions of the perception of color.

Katz, David. *Gestalt Psychology; Its Nature and Significance.* Translated by Robert Tyson. New York: Ronald Press, 1950. An excellent introductory textbook.

Kirsch, R. A. "Computer Interpretation of English Text and Picture Patterns." *Institute of Electrical and Electronic Engineering Transactions on Electronic Computers, Electronic Devices* 13 (4) (1964). Illustrates some of the difficulties in programming computers to be aware of perceptual Gestalten.

Koffka, Kurt. *The Growth of the Mind.* Translated by R. M. Ogden. London: K. Paul, Trench, Trubner, 1924. An introduction to child psychology from the Gestalt viewpoint.

Koffka, Kurt. *Principles of Gestalt Psychology.* New York: Harcourt Brace, 1935. A comprehensive survey of all the relevant experimental work up to 1935, accompanied by theoretical discussion of wide scope. Not easy reading, but a standard text. See pp. 655–661 for discussion of knowledge of other minds.

Köhler, Wolfgang. *Dynamics in Psychology.* London, 1942. An amplified version of lectures given at the University of Virginia in 1958.

Köhler, Wolfgang. *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology.* New York: Liveright, 1947. An interesting although somewhat discursive discussion of general principles. See Ch. 17 for the problem of knowledge of other minds.

Köhler, Wolfgang. *The Mentality of Apes.* Translated from the 2nd rev. German ed. by Ella Winter. London: K. Paul, Trench, Trubner, 1925; paperback ed., Harmondsworth, U.K., 1957. A classic work on problem solving. This work and his *Gestalt Psychology* (see above) are recommended as standard texts.

Köhler, Wolfgang. *The Place of Value in a World of Facts.* New York: Liveright, 1938. William James lectures. A scholarly discussion of philosophical issues.

Lewin, Kurt. *Principles of Topological Psychology.* Translated by Fritz Heider and Grace M. Heider. New York and London: McGraw-Hill, 1936. An attempt to apply Gestalt principles to the study of social situations.

Madden, Edward H. *Philosophical Problems of Psychology.* New York: Odyssey Press, 1962. Contains a critical examination of some of the key concepts of Gestalt theory.

Michotte, Albert. *La perception de la causalité.* Louvain, Belgium: Institut Supérieur de Philosophie, 1946 and 1954. Translated by T. R. Miles and E. Miles as *The Perception of Causality.* London: Basic, 1963. A systematic study of moving Gestalten, with special reference to the conditions in which we receive an impression of causality. See Chs. 12–13 for the problem of knowledge of other minds.

Rubin, Edgar. *Synsoplevede Figurer; Studier i psykologisk Analyse.* Copenhagen and Oslo: Gyldendal, Nordisk, 1915. Translated by Peter Collett as *Visuell wahrgenommene Figuren.* Copenhagen, Berlin, and London, 1921. An abridged version, translated by Michael Wertheimer, appears in *Readings in Perception,* edited by D. C. Beardslee and Michael Wertheimer. Princeton, NJ: Van Nostrand, 1958. Pp. 194–203. Interesting pioneer work.

Wertheimer, Max. "Experimentelle Studien über das Sehen von Bewegung." *Zeitschrift für Psychologie,* Part I, 71 (1912): 161–265. Of historical interest; commonly regarded as marking the beginning of the Gestalt movement.

Wertheimer, Max. *Productive Thinking.* Edited by S. E. Asch et al.. New York and London: Harper, 1945. An important contribution to the study of problem solving.

T. R. Miles (1967)

GEULINCX, ARNOLD
(1624–1669)

Arnold (or Aernout) Geulincx, the Flemish metaphysician and moralist, was born in Antwerp. He studied philosophy and theology at Louvain and in 1646 was made professor of philosophy, a position he held for twelve years. Although information about his life at Louvain is limited and his important works date from a later period, it appears that as a student he was influenced by the Cartesian Guillaume Philippi, that in his teaching, as later, he attacked scholastic physics from a Cartesian point of view, and that he was also attracted by the doctrines of Cornelis Jansen.

In 1658, on charges that were not made public but that may have been prompted by his criticisms of scholasticism and accepted religious practices, he was deprived of his professorship and left Louvain for Leiden. At the same time, he renounced Roman Catholicism and became a Calvinist. Arriving in Leiden in distressed circumstances, he was assisted by the Cartesian Abraham van der Heyden (Heidanus) and set to work on a study of fevers, which he presented for the doctorate in medicine. Despite his precarious situation at first, Geulincx succeeded in publishing treatises on logic and method (*Logica Fundamentis Suis … Restituta,* Leiden, 1662, and *Methodus Inveniendi Argumenta,* Leiden, 1663) and the first part of his most accomplished work, the "Ethics" (*De Virtute et Primis Ejus Proprietatibus,* Leiden, 1665). He was appointed professor extraordinary of philosophy at

the university in 1665 and remained in Leiden until his untimely death, in 1669. Six years later the complete "Ethics" was published, under the title *Γνωθι σεαυτον, Sive … Ethica* (Leiden, 1675). His "Physics," taken from manuscripts used in his classes, appeared in 1688 (*Physica Vera,* Leiden); commentaries on René Descartes's *Principles of Philosophy* in 1690 and 1691 (*Annotata Praecurrentia, Annotata Majora,* Dordrecht); and the very important "Metaphysics," published apparently from a student's copy, in 1691 (*Metaphysica Vera et ad Mentem Peripateticam,* Amsterdam).

OCCASIONALISM

Geulincx is best known for his occasionalist theory of causation and his denial of the substantiality of particular created things. Following Descartes's order of procedure in his "Metaphysics," he considered at the outset the possibility and the limits of doubt and found that our first knowledge is of the self as a thinking thing. Consideration of the various states of the self or mind led him to formulate a principle, which he took to be self-evident though obscured by prejudices, that expresses a necessary condition implicit in our conception of an action: that something cannot be done unless there is knowledge of how it is done, or, as specifically related to activities of the self, that a person does not do what he does not know how to do (*impossibile est, ut is faciat, qui nescit quomodo fiat; quod nescis quomodo fiat, id non facis*).

The principle had far-reaching consequences in Geulincx's moral philosophy as well as in his metaphysics. Concerning the self, he contended that actions involving movements of the body cannot in truth be attributed to the self and that the mind or soul is not, as it is often supposed to be, the true cause of movements of the body. Not only are we unaware of changes in the brain, nerves, and muscles requisite for, say, moving the arm, but even if we know of these changes from a study of physiology, our knowledge is based on ex post facto observation of sequences of volitions and physiological happenings, not on awareness of a supposed mental activity producing these movements. Though we have, Geulincx maintained, immediate knowledge and understanding of internal actions—that is, of acts not involving bodily movements and consisting solely of changes in a state of mind—we are not in like manner cognizant of how movements are initiated in the body or how external actions come about. Accordingly the influence of the human mind is limited to its own states, and the mind is not the master of—that is, the true cause of movements in—the body.

The principle was also invoked against the assumption that bodies, or corporeal things, are capable of acting, either on minds or on other corporeal things. It is assumed, for instance, that a fire acts on a man's body and, affecting sense organs, nerves, and brain, produces sensations of light and heat in his mind. It is also assumed that in cases of impact one body striking another sets the second body in motion. But how, Geulincx asked, can a body produce these effects? To bring them about, according to his principle, it would have to know how. Yet admittedly a body is inanimate and, lacking consciousness, lacks the knowledge that on reflection we see is a necessary condition of acting. Bodies are *res brutae.* To suppose that they have the distinctively spiritual characteristic of acting is a signal instance of confusion involving the self-contradictory notion of corporeal action or causation. Arguing against the possibility of genuine corporeal causation as such, Geulincx, like the occasionalists Géraud de Cordemoy and Nicolas Malebranche, took it to be true a fortiori that bodies cannot act on minds. (There is no evidence that Geulincx was influenced by, or that he in turn influenced, the other occasionalists.)

Though the human mind does not act on the body and bodies do not act on the mind or on other bodies, changes obviously do take place, and in these changes we discern patterns or constant conjunctions of events. According to Geulincx, the agent responsible for these changes is God, and the patterns we observe are due to laws that God enacts and in accordance with which he operates. Explicating his theory of supernatural causation in the case of volitions and bodily movements, Geulincx iterated two analogies, the second of which was the subject of an important controversy among German historians in the nineteenth century. (1) When a child wants his cradle to move, it often happens that the cradle moves, not as a result of his willing it, but because the mother or nurse in attendance wills that it move. (2) Two clocks that are synchronized sound the hour in unison, not because one influences the other, but because they are fashioned in such a way that they keep the same time. The second illustration has been cited to show that Geulincx, like Gottfried Wilhelm Leibniz, conceived of a preestablished harmony between mind and body and that he was the unacknowledged source of Leibniz's famous analogy of concurrent clocks and, by implication, of Leibniz's view of the relation between mind and body.

Against this interpretation it can be argued convincingly (as was done by Eduard Zeller) that in Geulincx's view, God's actions, though in accordance with rules, are immediate or direct in the sense that there is nothing in

mind or body comparable to the internal natures which, according to Leibniz, account for their successive states and mediate the will of God and the course of events. It is not the case, however, that the actions of Geulincx's God are ad hoc or, as Leibniz accused the occasionalists, that Geulincx's God is a deus ex machina. The rules of his action are fixed, and he simply applies them, with no special volitions required, in particular circumstances.

SUBSTANCE

Geulincx's views about substance were roughly midway between Descartes's and Benedict (Baruch) de Spinoza's. In the Synopsis of the *Meditations,* Descartes, drawing a distinction between body taken generally (*in genere sumptum*) and the human body, suggested that the former, like a person's mind, is a substance or pure substance (*puram … substantiam*), whereas the latter, insofar as it is a particular body differing from other bodies, is not. Following Descartes's lead, Geulincx contrasted body in itself (*corpus ipsum, corpus simpliciter dictum*), which he identified with extension, and particular bodies, which he claimed are modes of body (*aliquid ipsius corporis simpliciter dicti, modi corporis*). Body in itself is simple, unique, individual, infinite, and indivisible. Particular bodies are limitations of, or abstractions from, body in itself. They are not, he explained, constituent parts, nor are they figments of the mind (*entia rationis*); rather, they are related to body in itself as the superficies, or surface, of a particular body is related to that particular body. In another analogy, as the country is not a collection of fields, orchards, and meadows but the land on which these divisions are imposed, so corporeal nature is not an aggregate of particular bodies but the matter or extension common to them all and specified in various ways. The analogy also explains Geulincx's conception of mind. Like Spinoza (though independently), he held that individual minds are themselves not substances but modes of mind (*modi mentis*) or of infinite thinking substance, which he identified with God. We are, he said, both from God and in God (*ex Deo et in Deo*). To the extent to which we can transcend the distorting forms of our limited understandings and see the eternal truths in ourselves as they are in the mind of God, we lose our status as limited beings and are one with God. Geulincx's reflections on problems about substance paralleled Spinoza's. However, he preserved the Cartesian distinction between thinking substance and extended substance, or matter.

ETHICS

In the letter prefaced to the first part of his "Ethics," Geulincx implied that his moral philosophy rounds out the system conceived by Descartes, who, though he proposed a provisional code of morality in the *Discourse,* did not bring this branch of the tree of knowledge to fruition. In Geulincx's view the subject matter of ethics is virtue, and virtue is located not in deeds but in a determination of the will—that is, in love of right reason or, since reason as prescriptive comprises laws imposed by God, in devotion to divine law. Though virtue is one and simple, there are four aspects, and these cardinal virtues are distinguished from and contrasted with the traditional cardinal virtues, which refer to actions or accomplishments, not to the locus of morality—namely, the condition of the will. (1) *Diligence* is attention to the voice of reason. Its issue is wisdom and prudence in conduct. (2) *Obedience* involves compliance with the dictates of reason. Though we are free to will in conformity to divine law or not, in the end we cannot but do what God wills. By obeying his prescriptions we attain freedom in the highest degree: We will what we can do and do not will what we cannot do, and our volitions are effective. (3) *Justice,* also, is a determination of the will: to will no more and no less than reason dictates. (4) *Humility* consists in knowledge, and denial, of self (*contemptio sui*) in the love of reason and of God. Contrasted with the virtuous man is the egoist, whose end in life is happiness. He is the slave of his passions and the creature of circumstance, whereas the virtuous man, not seeking happiness and resigned to what happens to him, is in a position to attain it.

See also Cartesianism; Cordemoy, Géraud de; Descartes, René; Ethics, History of; Leibniz, Gottfried Wilhelm; Malebranche, Nicolas; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY GEULINCX

Arnoldi Geulincx Antverpiensis Opera Philosophica. Edited by Jan Pieter Nicolaas Land, 3 vols. The Hague: Martinum Nijhoff, 1891–1893. The standard edition.

WORKS ON GEULINCX

Grimm, Eduard. *Arnold Geulinx' Erkenntnisstheorie und Occasionalismus.* Jena, Germany, 1875.

Haeghen, Victor vander. *Geulincx: Étude sur sa vie, sa philosophie et ses ouvrages.* Ghent: A. Hoste, 1886. A bibliography of Geulincx' works is on pp. 197–224.

Hausmann, Paul. *Das Freiheitsproblem in der Metaphysik und Ethik bei Arnold Geulincx.* Würzburg: R. Mayr, 1934.

Hoffmann, Georg. *Das Gottesproblem bei Geulincx und Malebranche.* Düren: Danielewski, 1931.

Land, Jan Pieter Nicolaas. *Arnold Geulincx und seine Philosophie.* The Hague: M. Hijhoff, 1895.

Nagel, Karl. *Das Substanzproblem bei Arnold Geulincx.* Cologne, 1930.

Terraillon, Eugène. *La morale de Geulincx.* Paris: Alcan, 1912.

Vleeschauwer, Hermann Jean de. "Three Centuries of Geulincx Research." *Mededelings van die Universiteit van Suid-Afrika* 1 (1957): 1–72. Survey of writings on Geulincx.

On Geulincx and Leibniz

Eucken, Rudolf. "Leibniz und Geulinx." *Philosophische Monatshefte* 19 (1883): 525–542.

Pfleiderer, Edmund. *Arnold Geulinx als Hauptvertreter der Okkasionalistischen Metaphysik und Ethik.* Tübingen, 1882.

Pfleiderer, Edmund. *Leibniz und Geulinx mit besonderer Beziehung auf ihr beiderseitiges Uhrengleichniss.* Tübingen, 1884. See also the review of this book by Rudolf Eucken in *Philosophische Monatshefte* 20 (1884): 417–421.

Pfleiderer, Edmund. "Noch einmal Leibniz und Geulinx." *Philosophische Monatshefte* 21 (1885): 20–39.

Zeller, Eduard. "Über die erste Ausgabe von Geulincx' Ethik und Leibniz' Verhältniss zu Geulincx' Occasionalismus." *Sitzungberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 29 (1884): 673–695.

On Early Editions of Geulincx's Works

Vleeschauwer, Hermann Jean de. "L'opera di Arnoldi Geulincx." In *Edizioni di filosofia.* Turin, 1958. Pp. 1–53.

Willis Doney (1967)

GEULINCX, ARNOLD [ADDENDUM]

LIFE AND WORKS

Arnold Geulincx was born in Antwerp, Belgium, in 1624. In 1641 he matriculated at University of Louvain, where he became a professor of philosophy in 1646. For reasons that never became clear but were probably of a religious nature (at Louvain there was much sensitivity over Jansenism), he was suspended from his duties and consequently dismissed in 1657/1658. He moved to Leiden, The Netherlands, and converted to Calvinism. After taking a degree in medicine he obtained permission to lecture on philosophy, but his position was regularized only in 1662, when he was appointed reader in logic. He became professor *extra ordinem* in 1665. In 1667 he died from the plague. Most of his works, dealing with logic, moral philosophy, physics, and metaphysics, were published posthumously.

LOGIC

Although the main merits of Geulincx's works on logic seem to be their elegance and precision, Karl Dürr (1939–1940, 1965) and Gabriël Nuchelmans (1983, 1984, 1986) show that he made some important steps toward a logic of propositions. According to him words like *est* and *non* are signs (*notæ*) by which to indicate the mental act performed with respect to a particular content. Every denial is the negation of an affirmative claim, and that means that an affirmation has been present to the mind (*affirmatio inclusa*). Accordingly, "Peter is not learned" must be interpreted as "It is not the case that Peter is learned" or also as "The sentence 'Peter is learned' is false."

Speaking of compound conditional sentences, Geulincx defines an antecedent as a statement that says that the whole of that which some other statement (the consequent) says to be the case, is indeed the case. Consequence is a form of containment (*continentia*): Between two statements A and B there obtains a relation of consequence if A says the dictum of B. Both the theory of containment and the theorem that every A implies the statement "A is true" are corollaries of Geulincx's idea that by making a statement one commits oneself to the truth of that statement and of everything entailed by that statement. For example, if one says "I am standing," this must be taken as an affirmation of whatever is entailed by that statement, such as, for example, "I am capable of standing." Accordingly, "I am standing" serves as the antecedent of any number of other statements to the truth of which one commits oneself.

METAPHYSICS

According to Geulincx metaphysics is first philosophy or first science. It deals with the human subject, body, and God, each of which is the basis of a separate science: *autologia*, *somatologia* and *theologia*. The autologia basically consists in an exploration of the Cartesian *cogito*, which, however, Geulincx does not see as the basic principle of his philosophy but rather as a way to gain access to the realm of necessary truths. In fact, the more fundamental principle is the axiom that one can truly be said to make or do something only if one knows how it is made or done (*quod nescis quomodo fiat, id non facis*). This axiom allows Geulincx to claim that one is a passive spectator of the world, one's only activities being to will and think, albeit in a purely immanent way. Indeed, the world cannot be the cause of one's seeing and perceiving, given the fact that, since it can neither think nor know anything, it cannot be active. The only true cause is God and the only truly causal relation is that between God and the world. In fact, all philosophy should start with the concept of God, and the only reason why one has to start with the cogito is that the Fall has obscured one's faculties.

The result of Geulincx's analysis is that God is Being *simpliciter* as well as Mind *simpliciter*. This implies not only that all reality is ultimately mental but also that whatever is neither God nor part of God is nothing but an appearance. In fact, there are only two things that really exist, namely, Mind, which is the creator, and Body, which is the created. One's mind is part of the Divine mind (*mentis quid*). One's body is part of a phenomenal world. Particular three-dimensional bodies can be understood as limitations of the archetypal Extension that was produced in the act of creation. However, that there is a world, extended in three dimensions, can be known through the sensations God causes one to have. Finally, since contingent facts cannot be accounted for by principles of metaphysics (which explain only what is necessary), physics makes use of hypotheses, which must consist of clear and distinct ideas that together with the principles of metaphysics must be sufficient to explain all phenomena.

MORAL PHILOSOPHY

Like his metaphysics, Geulincx's moral philosophy is based on a corollary of his fundamental axiom, namely, that where there is no possibility to act there can be no will either (*ubi nihil vales, ibi etiam nihil velis*). In whatever way one acts, it is God that makes one act in that particular way. Accordingly, virtue is not to act in a particular way but to internally yield to God's will. Morality lies in the intention, not in the act. As a result, the cardinal virtues are dispositions: diligence, obedience, justice, and above all humility. On the contrary, passions are like sense impressions. Although they belong to human nature, they are relevant only insofar as they prevent one from developing the right attitude toward God's will. The most dangerous passion in this respect is self-love. In any case, the reward of virtue is that, freed from self-love, one enjoys peace and tranquility in this life.

RELATIONS TO OTHER THINKERS

Although much in Geulincx's philosophy goes back to René Descartes, it would probably be wrong to call him a Cartesian. For not only are the various parts of his philosophy differently connected (his metaphysics is crowned by his moral philosophy, not by his physics, which is comparatively independent), his metaphysics is, as has been shown by Brian Cooney (1972, 1978), basically an attempt to provide a metaphysical account of Divine Creation.

Accordingly, his philosophy has more affinity with Nicolas Malebranche's, with whom Geulincx shares a basically occasionalist interpretation of causality. Also, there is some similarity (although no affinity at all) with Benedict (Baruch) de Spinoza's philosophy, except, of course, that Spinoza rejects the idea of creation. In fact, Geulincx's starting point is fundamentally different. Whereas Spinoza argues that on the basis of Cartesian metaphysics it is impossible to account for creation, Geulincx takes creation to be a fact and attempts to make sense of it in terms of Cartesian metaphysics. However, the only way to do this is, he believes, to assume that creation consists in producing a world of appearance.

Not only Geulincx's metaphysics but also his physics is different from Descartes's, not because it would involve different concepts but because the status of their concepts is interpreted in a different way: Geulincx's concepts are hypotheses that, even if they are clear and distinct, are not automatically and necessarily true. This doctrine, which involves an interplay of empirical and metaphysical principles, has often been associated with Immanuel Kant's theory of judgment (Cassirer 1971–1973), but this is a bit far fetched. The best characterization seems to be that he is a Christian philosopher trying to find his way in the world of post-Aristotelian philosophy and availing himself of the language and concepts of his contemporaries to provide an intelligible account of the mysteries of faith.

Bibliography

WORKS BY GEULINCX

Quæstiones quodlibeticæ. Antwerp, Belgium, 1653.

Logica suis fundamentis restituta. Leiden, Netherlands, 1662.

Methodus inveniendi argumenta. Leiden, Netherlands, 1663.

Van de Hooft-deuchden. Leiden, Netherlands, 1667.

"Gnoti seauton" sive Ethica integra, edited by "Philaretus" (Cornelis Bontekoe). Leiden, Netherlands, 1675.

"Physica vera." In *Metaphysica*, by Cornelis Bontekoe. Leiden, Netherlands, 1688.

Annotata præcurrentia ad R. Cartesii Principia. Dordrecht, Netherlands, 1690.

Annotata majora in Principia philosophiæ Renati Des Cartes. Dordrecht, Netherlands, 1691a.

Metaphysica vera et ad mentem peripateticam. Amsterdam, Netherlands, 1691b.

Opera philosophica. 3 vols., edited by J. P. N. Land. The Hague: Nijhoff, 1891–1893.

Sämtliche Schriften. 3 vols., edited by H. J. De Vleeschauwer. Stuttgart, Germany: Frommann, 1965–1968.

WORKS ABOUT GEULINCX

Cassirer, E. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit* (1911). 4 vols. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 1971–1973.

Cooney, Brian. "Arnold Geulincx: A Cartesian Idealist." *Journal of the History of Philosophy* 16 (3) (1978): 167–180.

Cooney, Brian. "The Development of Cartesian Metaphysics: Descartes, Malebranche, Geulincx." Ph.D. diss., McGill University, 1972.

De Vleeschauwer, H. J. *Three Centuries of Geulincx Research*. Pretoria, South Africa: Universiteit van Suid-Afrika, 1957.

Dürr, Karl "Arnold Geulincx und die klassische Logik des 17. Jahrhunderts." *Studium generale* 18 (1965): 520–541.

Dürr, K. "Die mathematische Logik des Arnold Geulincx." *Erkenntnis* 8 (1939–1940): 361–368.

Haeghen, Victor van der. *Geulincx: Étude sur sa vie, sa philosophie et ses ouvrages*. Ghent, Belgium: A. Hoste, 1886.

de Lattre, Alain *L'occasionalisme d'Arnold Geulincx*. Paris: Éditions de Minuit, 1966.

Nuchelmans, Gabriël. *Geulincx Containment Theory of Logic*. New York: North-Holland, 1986.

Nuchelmans, Gabriël. *Judgment and Proposition from Descartes to Kant*. New York: North-Holland, 1983.

Theo Verbeek (2005)

GEWIRTH, ALAN

(1912–2004)

Alan Gewirth was a twentieth-century moral philosopher best known for his attempt to complete the Kantian project and show that rationality requires morality. Gewirth took his BA at Columbia University in 1934, studying with John Herman Randall and Richard McKeon. After two years of graduate study at Columbia, he spend the academic year 1936–1937 on a Sage Fellowship at Cornell University and then followed McKeon to the University of Chicago as his research and teaching assistant. In June 1942 Gewirth was drafted into the army, and, without seeing combat, moved up the ranks from private to captain in four years. After World War II, he returned to Columbia and received his PhD in philosophy in 1948. From 1947 on, he was a regular member of the faculty at the University of Chicago, eventually becoming the Edward Carson Waller Distinguished Service Professor in the Philosophy Department in a career that lasted more than sixty years. He also taught as visiting professor at Harvard University, the University of Michigan, John Hopkins University, and the University of Santa Barbara.

Early in his career Gewirth did important work on Descartes's theory of knowledge; later he did notable scholarly studies of the medieval political philosopher, Marsilius of Padua, published as his *Marsilius of Padua and Medieval Political Philosophy* (1951) and a published a translation from the Latin of Marsilius's *Defensor Pacis* (1956) with a lengthy introduction. Gewirth is best known, however, for his attempt to develop a stringently rational foundation for morality in his *Reason and Morality* (1978). The central argument of this book begins with a claim that every rational agent must accept, which is that he or she prudentially ought to have freedom and well-being. Gewirth argues that when the logical implications of this claim are fully worked out, particularly when the claim is universalized, it follows that every rational agent must also accept the claim that all prospective, purposive agents *morally* ought to have freedom and well-being, although, of course, purposive agents may not act on this claim—that is, they may not act morally.

Most of the critical reaction to Gewirth's work has focused on this particular argument. Two book of critical responses, along with replies from Gewirth, have been published. key issue concerns whether the universalization of a rationally inescapable claim that "I prudentially ought to have freedom and well-being" leads to the claim that "we all morally ought to have freedom and well-being" or to the claim of universal ethical egoism that "we all prudentially ought to have (or pursue) freedom and well-being." Gewirth claims the former; many of his critics claim the latter. Yet even some of those who reject Gewirth's argument for morality, for example, Christine Korsgaard and myself have been inspired by him to develop somewhat different arguments that attempt to establish just the same conclusion that Gewirth wanted to establish—that morality is rationally required.

In *The Community of Rights* (1996), Gewirth hoped to add to enhance his defense of morality by establishing against libertarians that rights—especially the human rights that equally belong to all humans as such—are positive as well as negative, and that they therefore warrant serious and active governmental concern for protecting and promoting the freedom and well-being of all humans, especially those who are most deprived. To this end, Gewirth employs two independent arguments. The first appeals to a definition of freedom, but, unfortunately, not to a definition of freedom that libertarians are required to endorse. The second is dialectical, but this argument parallels and depends crucially on Gewirth's earlier argument for morality.

In his last completed book, *Self-Fulfillment* (1998), Gewirth develops an interesting notion of self-fulfillment that is either compatible with or required by his conception of morality. A new book manuscript, *Human Rights and Global Justice*, which focused on questions of international justice, remained unfinished at his death.

See also Descartes, René; Ethics and Morality; Justice; Marsilius of Padua; Metaethics; Rationalism in Ethics (Practical-Reason Approaches).

Bibliography

WORKS BY ALAN GEWIRTH

Marsilius of Padua and Medieval Political Philosophy. New York: Columbia University Press, 1951.

Reason and Morality. Chicago: Chicago University Press, 1978.

Human Rights: Essays on Justification and Applications. Chicago: Chicago University Press, 1983.

Community of Rights. University of Chicago Press, 1996.

Self-Fulfillment. Princeton: Princeton University Press, 1998.

WORKS ABOUT ALAN GEWIRTH

Beyleveld, Deryck. *The Dialectical Necessity of Morality: An Analysis and Defense of Alan Gewirth's Argument to the Principle of Generic Consistency*. Chicago: University of Chicago Press, 1984.

Boylan, Michael. *Basic Ethics*. Upper Saddle River, NJ: Prentice Hall, 1999.

Boylan, Michael, ed. *Gewirth*. Lanham, MD.: Rowman & Littlefield, 1999.

Korsgaard, Christine. *The Sources of Normativity*. New York: Cambridge University Press, 1996.

Regis, Edward, ed. *Gewirth's Ethical Rationalism*. Chicago: University of Chicago Press, 1984.

Sterba, James P. *The Triumph of Practice Over Theory in Ethics*. New York: Oxford University Press, 2005.

James Sterba (2005)

GEYSER, JOSEPH
(1869–1948)

Joseph Geyser, the German critical realist philosopher, was born in Erkelenz, in the Rhineland. He received a doctorate in philosophy from the University of Bonn in 1898. He became an extraordinary professor at the University of Münster in 1904 and a full professor there in 1911. In 1917 Geyser was called to Freiburg, and in 1924 he succeeded Clemens Baeumker, the distinguished historian of ancient and medieval philosophy, at the University of Munich.

From his youth, Geyser opposed what he regarded as two basic tendencies in recent philosophy, an intellectualism strongly tinged with historical relativism and an overly abstract, idealistic Kantianism. He devoted himself to recalling philosophy to the asking of questions that are largely independent of any temporary situation and to the answering of these questions in an objective, critically realist manner. This attitude, but not Geyser's attachment to Thomistic tradition, was shared by the realist Oswald Külpe and the philosopher of nature Erich Becher. Above all, Geyser was totally devoted to the *philosophia perennis* (Gottfried Wilhelm Leibniz), and through it he related himself to the older European intellectual traditions. Philosophy, he strongly believed, is not a constant new beginning. For Geyser, philosophy, in the words of the early Middle Ages, is like a dwarf perched on the shoulders of a great past in order to see farther. Classical philosophy, in the thought of Plato and Aristotle, was already approaching great truths, insights that have claims on the present as part of a constantly self-renewing stream of thought, a stream that has both enriched and been enriched by the Christian worldview.

Geyser held that answers to philosophical questions must be based on direct contact with a real actuality, understood in the Aristotelian sense, as an entity independent of consciousness, and not on the creative activity of an idealistic, theoretical thought. Only thus can we stand on firm ground. Philosophy is in this view a kind of middle position between the reality of experience and the ideality of a creative reconstruction of the forms of existence. Geyser sensitively expounded this basic attitude in his short but deeply probing book, *Eidologie oder Philosophie als Formerkenntnis* (Eidology, or philosophy as knowledge of form; Münster, 1921). In this programmatic work, whose basic ideas were to guide Geyser's thought from then on, philosophy is presented as a progressive penetration into the realm of possible essences of being insofar as they offer themselves to experience.

Geyser's inclusive *Lehrbuch der allgemeinen Psychologie* (Münster, 1908; 3rd ed., 2 vols., Münster, 1920) had already been written from this point of view, which also formed the basis of Geyser's acute critique of Neo-Kantianism and Edmund Husserl's phenomenology in *Grundlegung der Logik und Erkenntnistheorie* (Foundations of logic and epistemology; Münster, 1919) and *Auf dem Kampffeld der Logik* (On the battlefield of logic; Freiburg, 1926), as well as of his later exposition of ontology and metaphysics. In all of these works we see Geyser as a relentless logician who was honest and strict with himself and an epistemologist capable of critical observations.

We can now examine how Geyser deals with modern problems. Geyser's *Psychologie* combined philosophical and modern empirical psychology, for he found it impossible to separate philosophy from psychology without damage to both disciplines. This open attitude permitted him to develop a method for recording mental life in all its unconscious, organic, and even ontological aspects. Nevertheless, Geyser fought against and tried to refute psychologism, the claim that even logical phenomena depend on psychological structures of experience. He developed an unambiguous distinction between actual psychological events and the logical, objective analysis of

meaning (in the manner of Heinrich Rickert). With equal intellectual vigor he gave a firm basis to his logical objectivism bound to being and distinguished it from both epistemological idealism and the phenomenological theory of constitution.

Objectivism meant for Geyser that logical laws are not only inner relations of thought but that they also have a real ontological character and that they stand the test of analysis. According to Geyser, philosophy should therefore explain how man is capable of grasping the ideal logical order in reality itself. Here we come face to face with ultimate realities, which reveal themselves to the human intellect only after suitable deductive rational preparation. Nevertheless, Geyser attains a knowledge of essence akin to that of phenomenology. But one could say that his phenomenology is Aristotelian and realistic, and Geyser can be credited with showing a connection, through Bernhard Bolzano, between the thought of Husserl and that of Aristotle. Geyser's epistemology is logically rational and tied to reality, and it stresses discursive, genetic methods. Only as a last resort could Geyser justify to his intellectual conscience an encounter with pure immediate insights.

The same rational and empirically bound method is evident in Geyser's views on causation, which provoked a many-sided controversy. These views are of particular relevance to contemporary discussions of the bases of natural science. The principle of universal causation, further seen as the law of causation, has always been considered to be the solid foundation of any given truth that can be discovered by analysis. In this sense it was one of the most essential supports of one of the traditional proofs of the existence of God. It was assumed to be evident and analytically a priori provable that the contingent world must have a supercontingent cause in God. Geyser investigated the distinctions between the ground of knowledge, the principle of contradiction, and the principle of sufficient reason. In *Philosophia Perennis, Festgabe Joseph Geyser,* Kurt Huber gives an exact summary of Geyser's critical investigations. In the same book Aloys Wenzel points out that Geyser was the first philosopher to further develop Arthur Schopenhauer's investigation of the principle of sufficient reason. Geyser showed further, in *Das Prinzip vom zureichenden Grunde* (The principle of sufficient reason; Ratisbon, 1929) and *Das Gesetz der Ursache* (The law of cause; Freiburg, 1933), that only through experience can we discover the meaning of causality. Everything has its sufficient reason, including being: "Everything that comes into being does so through a cause." Such a notion is originally given to us in the mental experience of causation in willing. The notion contains a synthetic feature, but it nevertheless remains completely unconditional although it is not given to intuition as an analytic law of thought. The questions of whether the principle of causation is a priori, and of how it is related to matters of experience, is significant for any possible further epistemological and metaphysical construction that is in accord with experience. Controversy with Thomistic philosophers resulted from this statement.

Geyser's position was also clearly expressed in his metaphysics. He was committed to an inductive metaphysics, not to a purely speculative metaphysics derived from intellectual immediacy. He thus distinguished his own thought from metaphysics as practiced by such Neo-Thomists as Gallus M. Manser and Antonin-Dalmace Sertillanges.

Similarly, when Geyser, like Francisco Suárez, ascribed a concrete, individual spiritual essence to human existence, he did not do so primarily in the Thomistic sense of a universal spiritual essence which achieves individuality by being united with matter. In the foreground of Geyser's thought is the empirically unique real event. It is thus not surprising to learn that Geyser, although he recognized a rational metaphysical knowledge of God by analogy, critically denied any intuitive insight into God's existence. (Here he differentiated his thought from Max Scheler's philosophy of religion, which he criticized in *Max Schelers Phänomenologie der Religion,* Freiburg, 1929). It also shows why Geyser rejected any ontological proof of God, that is, any knowledge of God reached by even the concept of the most perfect being discovered by an a priori encounter with essence. Rather, he felt that the existence of God is to be discovered a posteriori by an interpretation of the "united facts of experience." Geyser's thought found its completion in a rationally founded metaphysical knowledge of God.

Geyser was one of the most inclusive systematic thinkers of modern times. Nicolai Hartmann, another great systematizer, once said that he had learned more from Geyser's criticisms of his ontology than from those of any other contemporary. Few recent philosophers can call such a consistently thought-out and complete worldview their own. Geyser's worldview was developed within and into an inner unity with his Christian conviction; this firmness of attitude toward the world was also expressed in his whole steady personality, which endured the unhappy experience of his homeland during the last years of his life under the perspective of hope.

See also Aristotle; Bolzano, Bernard; Critical Realism; Hartmann, Nicolai; Husserl, Edmund; Leibniz, Gottfried Wilhelm; Neo-Kantianism; Objectivity in Ethics; Plato; Rickert, Heinrich; Scheler, Max; Schopenhauer, Arthur; Suárez, Francisco.

Bibliography

ADDITIONAL WORKS BY GEYSER

Die Erkenntnistheorie des Aristoteles (1917). Aalen, Germany: Scientia, 1980.

Über Wahrheit und Evidenz. Münster, 1918.

Erkenntnistheorie. Münster: H. Schönigh, 1922.

Augustin und die phänomenologische Religionsphilosophie der Gegenwart. Münster: Verlagbuchhandlung, 1923.

Einige Hauptprobleme der Metaphysik. Freiburg: Herder, 1923.

"Selbstdarstellung." In *Deutsche systematische Philosophie.* Edited by Hermann Schwarz. Vol. II. Berlin: Junker, 1934.

WORKS ON GEYSER

Baur, Ludwig. "Joseph Geyser als Metaphysiker." In *Philosophia Perennis, Festgabe Joseph Geyser,* edited by F.-J. von Rintelen, Vol. II, 1173–1196. Regensburg: J. Habbel, 1930.

Ettlinger, Max. "Joseph Geyser als Psychologe." Ibid. Pp. 1131–1140.

Gabel, Herbert. *Theistische Metaphysik im Ausgleich von Idealismus und Realismus bei Joseph Geyser.* Freiburg, 1957.

Huber, Kurt. "Geysers Stellung in Logik und Erkenntnistheorie." In op. cit., edited by F.-J. von Rintelen, 1141–1172.

Rössli, J. *Das Prinzip der Ursache und des Grundes bei Joseph Geyser.* Fribourg, Switzerland: Paulusdruckerei, 1940.

Wenzel, Aloys. "Der Gestalt- und Ganzheitebegriff." In op. cit., edited by F.-J. von Rintelen, 657–684.

Fritz-Joachim von Rintelen (1967)
Translated by Tessa Byck

GHAZĀLĪ, ABU HAMID MUHAMMAD

See *al-Ghazālī, Muhammad*

GIBBON, EDWARD
(1737–1794)

Edward Gibbon, the English historian and man of letters, was born at Putney, Surrey, of a well-to-do family. Frail and constantly ill, the child owed the preservation of his life to an aunt, Miss Catherine Porten, who also acted as his teacher. After instruction by a series of tutors and much reading on his own, he entered Magdalen College, Oxford, at the age of fifteen, with, as he later confessed, "a stock of erudition which might have puzzled a doctor, and a degree of ignorance of which a schoolboy would have been ashamed." Fourteen months at college, "the most idle and unprofitable of my whole life," ended with self-conversion to Roman Catholicism. His irate father immediately packed him off to Lausanne, Switzerland, under the care of Daniel Pavillard, a Calvinist minister who soon led him back to Protestantism. Thereafter, he developed a decidedly skeptical bent. During his five years' stay in Switzerland, Gibbon learned French, Italian, and Greek, and read all the Latin classics. He also fell in love with Suzanne Curchod. When his father refused consent to their marriage, "I sighed as a lover, I obeyed as a son." Mlle. Curchod later married Jacques Necker, distinguished French financier and statesman, and became famous as a *salonnière.* Gibbon never married.

Gibbon's first publication was *Essai sur l'étude de la littérature* (1761). A later manuscript fragment of a "History of the Swiss Revolution," also in French, was shown to David Hume, who approved of the project but chided the author: "Why do you compose in French, and carry faggots into the wood?" Thereafter, Gibbon composed all his major works in English. For more than two years (1759–1762), Gibbon was a captain in the Hampshire militia, and a surprisingly good one. In 1763, with the end of the Seven Years' War, he returned to the Continent, visiting Paris, Lausanne, and finally Rome. He records that it was on October 15, 1764, while musing amid the ruins of the Capitol, that the idea of writing about the decline and fall of the city—later extended to the empire—first occurred to him.

Returning to England in 1765, he became a man of letters and man about town. In 1774 he was elected to Dr. Johnson's Literary Club, where he became the intimate friend of Adam Smith. In the same year he obtained a seat in parliament, where he earned the distinction of never making a speech. He was, however, hard at work on his great history, the first quarto volume of *The History of the Decline and Fall of the Roman Empire,* which appeared in 1776. A letter of congratulation from the dying Hume "overpaid the labour of ten years," but warned that a clamor would arise. It did, and Gibbon responded three years later with a *Vindication.* The sixth and last volume of the history was published in 1788. At least fifty British replies and refutations were published before Gibbon's death, and literally hundreds have been published in many languages since. At his death, Gibbon left behind six drafts of an autobiography, which were pieced together and published in his *Miscellaneous Works* in two volumes by Lord Sheffield (London, 1796).

THE HISTORY OF THE "DECLINE AND FALL OF THE ROMAN EMPIRE"

Received as a masterpiece on first publication, Gibbon's history is still regarded as such, and has never been superseded. Certain misinterpretations of facts, to be sure, have been detected and many additional facts have come to light; some prejudices have been revealed and some misjudgments have become apparent, "but in the main things he is still our master, above and beyond 'date'"—so acknowledged J. B. Bury in the modern standard edition of the work.

It is the famous fifteenth and sixteenth chapters of the first volume, however, that entitle Gibbon to an honored place in the history of philosophy. These are the two chapters that stirred up violent controversy in 1776, and they are still controversial. The problem that Gibbon set himself was to explain the progress of primitive Christianity and its influence upon the ultimate fall of the Roman Empire. Writing *en philosophe,* Gibbon comes to the conclusion that the fall of Rome represents "the triumph of barbarism and religion." He ironically dismisses the most commonly accepted causes of the triumph of Christianity, namely, the convincing historical evidence of the doctrine itself and the ruling providence of its great Author. He notes that through the course of time prejudice and passion have distorted and rendered ambiguous the meaning of the doctrine, while the providence of Deity remains inscrutable to man. The former cause, therefore, is unhistorical, while the latter is unphilosophical. Ruling out supernaturalism as a cause, Gibbon consequently confines himself in the fifteenth chapter to an analysis and discussion of the secondary causes of the rapid growth of the Christian church—causes that can be tested both by historical fact and by philosophical and psychological analysis.

With cool detachment of philosophical and historical inquiry, he examines the early history of the church in the same spirit that he would examine any period of secular history in which no assertions of supernaturalism had been made. He discusses five secondary causes of the rise of Christianity: (1) The inflexible zeal of the Christians was inherited in part from the Jews, who alone had broken the religious harmony of the ancient world, which was based upon mutual toleration of all creeds, and had insisted that theirs was the one and only true religion. The Christians turned this defensive zeal into both the proselytizing of all ranks of people and the persecution of all varieties of idolatry.

(2) Belief in immortality, uncertain and disputed among the ancient philosophers and not to be found in the law of Moses, gradually began to be accepted by the Jews after their servitude to Egyptians and Babylonians. Early Christians, contemptuous of their present existence and convinced of their immortality, believed in the near approach of the end of the world, which was to be preceded by the Second Coming of Christ. At this time, believers and unbelievers alike would receive judgment—the former, eternal bliss; the latter, eternal damnation. As for the tortures which awaited sinners and deluded philosophers, Gibbon finds it proper "to draw a veil over the rest of this infernal description."

(3) The early history of the church is replete with claims to miraculous powers and to divine inspiration. Such forms of superstition and "enthusiasm" made constant progress, until they became part of church tradition. But it remains the scholarly duty of the historian to examine such claims and to reject all pretensions to inspiration that are unacceptable in the light of reason. If the age of miracles once existed, all reasonable men, in contrast to the credulous and the fanatical, agree that at some time it either suddenly or gradually terminated.

(4) The pure and austere morals of the early Christians were enhanced by two laudable human motives: repentance for past sins and the desire for perfection. Converted sinners became saints, disdainfully rejecting the natural human propensities for pleasure and action in favor of the monkish virtues of humility, meekness, and patience. A state of celibacy came to be exalted as the nearest approach to divine perfection, and sensual pleasure was inexorably replaced by spiritual pride. Passive obedience to civil authority led to a refusal to partake in any form of civil administration or military defense of the empire, even when it became evident that such disregard of the public welfare guaranteed the triumph of barbarism. In sum, the morals, and the errors, of the primitive Christians were in reality the excess of their virtues.

(5) Though immune to both the business and the pleasure of the world, the primitive Christians took keen interest in the government of the church, an enthusiasm that gave rise to much religious contention. At first, the bishops were regarded as the equals of the people, but gradually took upon themselves arbitrary power, ultimately proclaiming themselves vice-regents of Christ. Thence arose the rigid distinction between clergy and laity. The early communion of goods among the Christians was soon relaxed, and the clergy adopted the tithe from the original Jewish code. Further clerical controls included excommunication, which involved not only spiritual but also temporal punishment. As to the actual numbers of Christians, nothing definite can be con-

cluded, the figures of the Fathers being at complete variance with those of the pagan historians, and neither providing accuracy. Seneca, the two Plinys, Tacitus, Plutarch, Galen, Epictetus, Marcus Antonius, great sages all, have little or nothing to say about the "perfection" of Christianity. Alleged miracles for the benefit of the church passed unnoticed.

In the sixteenth chapter, Gibbon examines the question of the persecutions of the primitive Christians by some of the Roman emperors. The blame, he indicates, rests chiefly upon the intolerant zeal of the Christians themselves, which drove the emperors reluctantly toward persecution. Even so, there were frequent peaceful intervals, and the detailed accounts of the sufferings of the "martyrs" were largely the inventions of later ecclesiastical writers. Gibbon estimates that no more than two thousand Christians were executed during the period of the most vigorous persecution, and suggests a comparison with the hundreds of thousands of Protestants executed during the relatively brief period of the Reformation, the latter figure far exceeding all martyrdoms over the course of many centuries of early Christian persecution.

GIBBON'S SOURCES

Among the many influences upon Gibbon's method and philosophy, the following should be mentioned: first, John Locke's commonsense approach to philosophy and religion; second, the rationalism of the deists; third, the philosophy of history presented in Baron de Montesquieu's treatise *Considérations sur les causes de la grandeur et de la décadence des Romains* (1734); fourth, the philosophical skepticism of Hume. From Hume he also learned the necessity of investigating the causes of historical events, and from Hume and Voltaire, the importance of cultural, social, and political history. The *Decline and Fall* has gone through multitudinous complete editions and condensations, both in English and in translation, and will continue to be read, not only as a great history, but also as a great piece of literature.

See also Epictetus; Galen; Hume, David; Johnson, Samuel; Locke, John; Montesquieu, Baron de; Philosophy of History; Plutarch of Chaeronea; Seneca, Lucius Annaeus; Smith, Adam; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY GIBBON

Gibbon's works include the *Autobiography,* edited by Dero S. Saunders (New York: Meridian, 1961), a more faithful version than Lord Sheffield's; *The Decline and Fall,* edited by J. B. Bury, 7 vols. (London, 1896–1900; 2nd ed., 1909–1913), the modern standard edition. See also *The Letters of Edward Gibbon,* edited by J. E. Norton, 3 vols. (London: Cassell, 1956).

WORKS ON GIBBON

For literature on Gibbon, see J. B. Black, *The Art of History: A Study of Four Great Historians of the Eighteenth Century* (London: Metheun, 1926); Harold L. Bond, *The Literary Art of Edward Gibbon* (Oxford: Clarendon Press, 1960); D. M. Low, *Edward Gibbon: 1737–1794* (London: Chatto and Windus, 1937); Shelby T. McCloy, *Gibbon's Antagonism to Christianity* (Chapel Hill: University of North Carolina Press, 1933); J. E. Norton, *A Bibliography of the Works of Edward Gibbon* (Oxford: Oxford University Press, 1940); George Sarton, "The Missing Factor in Gibbon's Concept of History," *Harvard Library Bulletin* XI (1957): 277–295.

Ernest Campbell Mossner (1967)

GIBBS, JOSIAH
(1839–1903)

Josiah Gibbs, a theoretical physicist, was born and died in New Haven, Connecticut, and, aside from a few years studying physics in Europe, spent his academic career at Yale. He is one of the few distinguished American theoretical physicists prior to the twentieth century. Gibbs made advances in vector analysis, and he made major contributions to thermodynamics including an insightful diagrammatic method, work on equilibrium and stability, the definition of free energy, and his famous phase rule regarding coexistent phases of a substance. In a vital contribution to thermodynamics Gibbs extended this theory to deal with the rules that describe how chemical interactions are to be integrated with the other thermodynamic processes. He is the inventor of the notion of chemical potential, the key concept of chemical thermodynamics.

For philosophers it is Gibbs's work in statistical mechanics that is of great interest. This work is contained in his elegant *Elementary Principles in Statistical Mechanics* (1902). James Clerk Maxwell and Ludwig Boltzmann had previously developed a method for calculating equilibrium values by taking them to be averages over functions of the microscopic phase of the system using a time invariant probability distribution over a constant energy subspace of the phase space of the system. This technique reappears in Gibbs in the form of his "microcanonical ensembles."

But Gibbs introduced other ensembles as well. Most important of these is the canonical ensemble that intro-

duces a time invariant probability distribution over the phase space that allows for different energies. For systems with a large number of degrees of freedom (a large number of molecules in a gas, for example), this probability distribution is highly concentrated about one specific energy. In these cases, averages calculated using the canonical distribution and those calculated using the microcanonical distribution will converge in the limit of an infinite number of degrees of freedom. Because calculations done using the canonical ensemble are much easier than those using the microcanonical, most practical statistical mechanics is done in the framework of Gibbs's canonical ensembles. Gibbs also developed the grand canonical ensemble whose use becomes necessary when chemical changes are part of the thermodynamic processes.

By showing how these ensembles and the features of them vary as constraints on the system are varied, Gibbs was able to show "analogies" between features of the ensembles and averages of features calculated with their probability distributions and standard thermodynamic quantities such as temperature and entropy. He is cautious in making any explicit "identification" of the latter with the former, possibly in part because of the known difficulties faced by standard statistical-kinetic reasoning at that time in correctly predicting such quantities as specific heats.

With the association of thermodynamic and statistical mechanical quantities, it is easy to understand the microcanonical ensemble as appropriate for a system energetically isolated from the rest of the world, and the canonical as appropriate for a system in perfect thermal contact with an infinite heat bath of constant temperature.

Gibbs's treatment of nonequilibrium is the source of one standard approach to that problem, but remains controversial to this day. Gibbs's ensembles can be thought of as a vast collection of systems identically prepared at the macroscopic level. Find the ensemble for such a collection of systems; now, change a constraint on the system (say by removing a partition in a box of gas): How will the ensemble, appropriate for equilibrium before the change of constraint, evolve? Will it evolve to the ensemble appropriate for equilibrium in the new constraint condition? This is what is most desirable because people want to show that, in some appropriate sense, the systems in the ensemble at a later time will be found, in general, to be ever closer to the equilibrium condition. But provably the Gibbs's ensemble cannot so evolve (Liouville's Theorem).

But, Gibbs argues, the ensemble may evolve in such a way as to approach the new equilibrium ensemble in a "coarse grained" sense. He uses the analogy of a glass mostly filled with water but partly filled with insoluble black ink. Stir the fluid. If one looks closely enough, the fluid always consists of pure water or pure ink, because the ink is insoluble. But looked at "coarsely," the fluid approaches a uniformly gray color. Gibbs was not able to show that such "mixing" would actually occur, but modern extensions of ergodic theory have been able to prove mixing theorems that hold under certain physical conditions. And idealized systems (such as molecules as "hard sphere in a box") have been shown to be mixing. It remains controversial, however, as to whether this model of an ensemble evolving in a coarse-grained sense is the appropriate one for characterizing the actual approach to equilibrium of nonequilibrium systems.

Gibbs is aware that "mixing" ought to be a time symmetric feature of his ensembles, given that it is driven by a time symmetric underlying dynamics of the molecules. But applying mixing in the past time direction would lead, incorrectly, to predict antithermodynamic behavior for systems. His solution is to argue that one ought to apply statistical inferences only into the unknown future, and that applying them to infer the a past that is already known is illegitimate. Paul and Tatiana Ehrenfest, in their important 1910 survey of statistical mechanics, called Gibbs's argument "incomprehensible." But it was later taken up and developed by Satosi Watanabe and Erwin Schrödinger. It also remains a subject of contemporary controversy in discussions of the relationship between the intuitive asymmetry of time and entropic features of the world.

See also Boltzmann, Ludwig; Philosophy of Statistical Mechanics.

Bibliography

Ehrenfest, Paul, and Tatiania Ehrenfest. *The Conceptual Foundations of the Statistical Approach in Mechanics*. Ithaca, NY: Cornell University Press, 1959.

Gibbs, Josiah Willard. *Elementary Principles in Statistical Mechanics*. New Haven, CT: Yale University Press, 1902.

Sklar, Lawrence. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge, U.K.: Cambridge University Press, 1993.

Wheeler, Lynde Phelps. *Josiah Willard Gibbs: The History of a Great Mind*. Woodbridge, CT: Ox Bow Press, 1998.

Lawrence Sklar (2005)

GILBERT OF POITIERS
(c. 1076–1154)

Gilbert of Poitiers (Gilbertus Porreta, Gilbert de la Porrée), the twelfth-century theologian and metaphysician, was born at Poitiers about 1076 and received his first schooling there. Next he went to study under Bernard of Chartres, and later (but before 1117) he devoted himself to theology under Anselm at Laon. He seems to have succeeded Bernard as chancellor at Chartres between 1126 and 1137 and, after a short period as a master in Paris, was elevated in 1142 to the bishopric of Poitiers. He died greatly esteemed in 1154, although in the 1140s he had been made to feel the hostility of other theologians, principally Bernard of Clairvaux, who brought him to trial to account for his opinions at Paris in 1147 and at Rheims in 1148.

Gilbert wrote much and acquired great fame for his scriptural and Boethian commentaries. The former were the fruit of his years at Laon and included major expositions of the Psalms and of the Epistles of St. Paul, as well as other biblical commentaries that have, with greater or lesser certainty, been ascribed to him. But the commentaries upon Boethius's four *opuscula sacra* (and especially that upon the *De Trinitate*) proved controversial. Although Gilbert was never officially condemned for theological error, after his trial in 1148 he appended a new preface to these commentaries professing his orthodoxy. In addition, the treatise *De Discretione Animae, Spiritus et Mentis* is now confidently ascribed to Gilbert. Highly uncertain, however, is Gilbert's authorship of the *Liber Sex Principiorum.* The six *principia* are the last six Aristotelian categories (place, time, situation, habit, action, and passion), which the writer of this treatise considered to be accessory forms (*formae assistentes*) or extrinsic circumstances of a substance. The first four categories, on the other hand, are either substance itself or necessarily inherent forms of a substance. This work enjoyed great authority in the Middle Ages as a completion of Aristotle's own *Categoriae.*

An understanding of Gilbert's authentic philosophical teaching must be based principally upon his Boethian commentaries and upon the literature inspired by his trial. Gilbert's doctrine of being and of the process of knowledge departs from a key distinction between substance and subsistence. A substance is an actually existing individual being that supports (*substat*) a number of accidents. Some beings, however—genera and species, for example—have no need of accidents and are more accurately described as subsistences than as substances. Forms or Ideas in themselves are subsistences and do not come into contact with matter. Only copies (*exempla*) descend into matter. The human mind arrives at the knowledge of the eternal Ideas by first "collecting" from concrete, individual things their substantial similarity, that is, their created or "native" forms (*formae nativae*), to which Gilbert attributed universality. By perceiving the similarity of forms within a group, the mind arrives at the concept of species and then, by the same process, it arrives at the concept of genus. Finally, transcending all created forms, it attains the primary forms, which are in God. Thus, Gilbert inquired why concrete forms agree with one another, and he focused his attention upon the *intellectus* of the universal which is abstracted from singulars. He based his theory of knowledge upon the Platonic doctrine of Ideas but also employed the Boethian-Aristotelian doctrine of abstraction.

The divine work of creation involved the production of forms, which are images of the divine Ideas, and the uniting of these forms to matter. Gilbert described the created being as a compound of the *id quod est* ("what it is") and the *id quo est* ("that by which it is"). Socrates is a man (*id quod est*), but he is what he is by virtue of his humanity and corporeity (*id quo est*). The origin of this distinction is the grammatical rule that, *in naturalibus,* every name signifies both a substance and a quality. But whereas all created being is compound, the divine being is absolutely simple. In God, essence (*id quod est*) and divinity (*id quo est*) coincide. Nonetheless, Gilbert applied the distinction to God, describing divinity as the form in God by which he is God. Gilbert's opponents, such as Bernard of Clairvaux, would not accept this separation of God and his divinity; they maintained that divinity is God, and not that by which he is. Gilbert's position was a difficult one to maintain, but he had no desire to compromise the divine simplicity or unity, and his writings support his claim that he had not established a real distinction between God and his divinity.

In a similar manner, Gilbert's application of logical and grammatical principles to the problem of the Incarnation of Christ aroused suspicions. Gilbert was reluctant to say that the divine nature became flesh, preferring to say that a person, Christ, took a human nature. Other logicians of the day were similarly concerned to test various traditional formulations of the divine Incarnation in the light of Boethian concepts. If Gilbert slipped in his analysis of the person and natures of Christ, he did not intend to deny Christ's divinity or his humanity.

Gilbert's school of disciples survived as a strong force in the twelfth century and included John of Salisbury,

Otto of Freising, Alan of Lille, Nicholas of Amiens, Radulphus Ardens, and John Beleth. It blended at times with the dialectical tradition stemming from Abelard, and, by its investigation of the character of essences, the school of Gilbert perhaps helped to prepare the way for the influx of Avicennian philosophy.

See also Abelard, Peter; Anselm, St.; Bernard of Chartres; Bernard of Clairveaux, St.; Boethius, Anicius Manlius Severinus; Ideas; John of Salisbury.

Bibliography

WORKS BY GILBERT

Liber de Sex Principiis Gilberto Porretano Ascriptus has appeared in two modern editions: that by A. Heysse (Münster-in-Westfalen: Aschendorf, 1929) and that by D. van den Eynde (Münster-in-Westfalen, 1953). Commentaries on Boethius may be found in *Patrologia Latina,* edited by J. P. Migne, Vol. 64 (Paris, 1847), cols. 1255–1412. N. M. Häring has edited new editions of the commentaries on the *De Trinitate* and on the *De Praedicatione,* which are in *Nine Mediaeval Thinkers,* edited by J. R. O'Donnell (Toronto, 1955); on the *De Hebdomadibus,* which is in *Traditio* 9 (1953): 177–211; and on the *Contra Eutychen et Nestorium,* which is in *Archives d'histoire doctrinale et littéraire du moyen âge* 21 (1954): 241–357. Häring has also edited the *De Discretione Animae, Mentis et Spiritus,* which is in *Mediaeval Studies* 22 (1960): 148–191, and "A Christmas Sermon by Gilbert of Poitiers," in *Mediaeval Studies* 23 (1961): 126–135. Gilbert's biblical commentaries are unprinted save for the prologue and commentary on Psalms 1 and 2, edited by M. Fontana, in *Logos* 13 (1930): 283–301.

See also *Commentarius Porretanus Anonymus in Primam Epistolam ad Corinthios,* edited by A. M. Landgraf, in Vol. 117 of the series *Studi e testi* (Vatican City: Biblioteca Apostolica Vaticana, 1945), and E. Rathbone, *Notae Super Iohannem Secundum Magistrum Gil*[*bertum*], in Vol. 18 of the series *Recherches de théologie ancienne et médiévale* (Louvain: Abbaye du Mont César, 1951), pp. 205–210.

WORKS ON GILBERT

The following works contain information on Gilbert: Étienne Gilson, *History of Christian Philosophy in the Middle Ages* (London: Sheed and Ward, 1955), pp. 140–144; A. M. Landgraf, "Untersuchungen zu den Eigenlehre Gilberts," in *Zeitschrift für katholische Theologie* 54 (1930): 180–213; A. Forest, "Le Réalisme de Gilbert de la Porrée dans le commentaire du *De Hebdomadibus,*" in *Revue néo-scholastique de philosophie* 36 (1934): 101–110; A. Hayen, "Le Concile de Reims et l'erreur théologique de Gilbert," in *Archives d'histoire doctrinale et littéraire du moyen âge* 10–11 (1935–1936): 29–102; M. H. Vicaire, "Les Porrétains et l'avicennisme avant 1215," in *Revue des sciences philosophiques et théologiques* 26 (1937): 449–482; M. E. Williams, *The Teaching of Gilbert Porreta on the Trinity,* Vol. 56 in the series *Analecta Gregoriana* (Rome, 1951); M. A. Schmidt, *Gottheit und Trinität nach dem Kommentar des Gilbert Porreta zu Boethius, De Trinitate* (Basel, 1956); S. Gammersbach, *Gilbert von Poitiers und seine Prozesse im Urteil der Zeitgenossen* (Cologne: Böhlau, 1959); and N. M. Häring, "The Case of Gilbert." in *Mediaeval Studies* 13 (1951): 1–40; "A Latin Dialogue on the Doctrine of Gilbert," in *Mediaeval Studies* 15 (1953): 243–289; and "Sprachlogische und philosophische Voraussetzungen zum Verstandnis der Christologie Gilberts," in *Scholastik* 32 (1957): 373–398.

David Luscombe (1967)

GILES OF ROME

(c. 1247–1316)

Giles of Rome, the scholastic philosopher whose real name was Aegidius Colonna Romanus, was born in Rome. Giles entered the Augustinian order of hermits in 1265 and subsequently studied at the University of Paris, where from 1268 to 1272 he was probably the pupil of Thomas Aquinas, who was then lecturing at the university as Dominican regent master. In 1277 the bishop of Paris made his far-reaching condemnation of 219 theses, mainly of Aristotelian origin but also including a number of Thomist propositions. Among these were Thomas's doctrine that each being contains only one substantial form, as opposed to the traditional Augustinian belief in a plurality of forms. Giles, a young scholar, joined in the ensuing controversy with the publication of a sharply worded defense of the Thomist view, the *Liber Contra Gradus et Pluralitatem Formarum.* He attacked the Augustinian doctrine as being contrary to both reason and faith. Upon his refusal of Bishop Tempier's demand for a retraction, Giles left Paris, perhaps for a cooling-off period, but returned in 1285 to take the first Augustinian chair in theology and to receive his license to teach. He remained a professor until 1292, when he was appointed prior general of his order. In 1295 Pope Boniface VIII appointed him archbishop of Bourges, in which office he remained until his death. In 1287 his teachings had become the official doctrine of the Augustinian order, although neither of the other great Augustinian thinkers of the fourteenth century, Thomas of Strasbourg and Gregory of Rimini (each a general of the order) followed his teachings.

METAPHYSICS

Giles's philosophical position still remains something of an enigma. The older view that he was strictly a disciple of Thomas has gradually been modified. While it is true that he reached substantially the same conclusions as Thomas on two of the burning issues of the day, the unity

of the substantial form and the distinction between essence and existence, neither of these seems to have been from Thomist premises and, in the case of the second issue, the conclusion did not even lead to the same doctrine. In particular, Giles seems to have been influenced to a far greater degree than Thomas by Neoplatonism, and especially Proclus, on whose *Liber de Causis* he commented in 1280. This affinity would explain his own treatment of the relation of essence to existence. Thomas had never made a real distinction between the two but had regarded them rather as a composition in which *esse* is the actuality of *essentia,* which is itself the source of a being's actuality; or put another way, a being is what it is in virtue of the actuality (*esse*) that derives from its form (*essentia*). For Giles, on the other hand, *esse* and *essentia* were distinct things (*res*) from the outset. He therefore treated as real what for Thomas were abstractions, an attitude confirmed in his Commentary on the *Liber de Causis,* where he thought in terms of a universe of intelligible beings. To attain intelligible knowledge, it suffices for the image of an object to act directly upon the possible intellect, which under the influence of the active intellect is able to conceive it as an intelligible species. This led Giles to the characteristically Platonic conclusion "that the same quiddity considered in things is particular, considered in the mind is universal."

SCIENTIFIC CONTRIBUTIONS

Although knowledge of Giles's scientific outlook is even less comprehensive than that of his philosophical system, his treatment, often Neoplatonic, of time, movement, gravity, quantity, the intensification and remission of forms, and matter is known to us. Giles made his most original and important contribution to later scholastic scientific discussion concerning the nature of quantity. He posited a twofold quantity (*duplex quantitas*) that corresponds to the modern distinction between mass and volume. On the one hand, a body contains a constant quantity of matter, which limits its possibilities of development; for instance, a barleycorn cannot become a mountain. On the other hand, the same quantity of matter can undergo various changes in dimension, and according to its volume it will be denser or rarer in structure—as with, say, water or air. Giles took this distinction to infer that mass and volume were thus two independent quantities.

Giles also distinguished sharply between form and matter in the structure of a material substance—the so-called problem of the *mixtum.* This raised the question of what happened to the forms of the four material elements—fire, earth, air, and water—which composed any material substance when they were combined with form of that substance, for example, wood. Did they continue to exist separately, or were they absorbed into the substantial form? This was one of the earliest scientific problems to exercise the Scholastics, and while Giles based himself upon what Thomas had already said, he also went further. He accepted Thomas's solution that the forms of the material elements, once included in a material substance, no longer remained formally and actually in being but, rather, virtually as part of the qualities of the substance. To this, however, he added the distinction between the material and formal qualities. The former (*ex parte materiae*) remained the same through all changes in the substance; the forms, on the other hand, could not remain numerically the same. Another aspect of Giles's *mixtum* theory was of a hierarchy (*ordo realis*) among substantial forms, in which each higher form virtually contained the lower forms, the higher form being able to do more perfectly whatever the lower form could do.

Giles was also the first among the high Scholastics to state explicitly the problem of the increased speed of a falling body, namely, that this was not caused by the approach of its destination but rather by the growing distance from its starting point. Again, concerning a falling body in a vacuum, a problem which was to exercise successive generations of fourteenth-century thinkers, Giles was the first to pose it directly, taking a standpoint different from that of Thomas. In his Commentary on Aristotle's *Physics,* Giles posed other original questions about movement: Whether the sole cause of why movement took place in time, and not instantaneously, was resistance to the mover from the medium in which it moved, and whether in a vacuum movement itself would be composed of a succession of instants which in themselves did not constitute time. To both Giles answered in the affirmative. Thus the difference between movement in a medium and movement in a vacuum was that in the first case it was successive as opposed to instantaneous, owing to the resistance encountered; in a vacuum, on the other hand, it was *motus discretus in tempore discreto.* There can be little doubt from what is already known of his scientific speculation that Giles was the forerunner of the scientific inquiry so characteristic of the fourteenth century.

POLITICAL PHILOSOPHY

Giles had been tutor to the future Philip IV (the Fair) of France, to whom he dedicated his *De Regimine Principium.* This work, stressing the Aristotelian view of a ruler, was based upon the *Ethics* and *Politics.* But with the out-

break of the struggle between Philip IV and Pope Boniface VIII in 1296, Giles took the pope's side. His *De Potestate Ecclesiastica* (probably written c. 1302) stated, in the most extreme form yet, the Augustinian view of society, in which the spiritual power is superior to the temporal and only the faithful can possess the just and righteous lordship derived from the universal lordship of the church. Lordship is a gift from God, and justice is submission to God through the church; hence, sin deprives the sinner of all right of lordship. Giles here sowed the seeds of the doctrine of dominion and grace that was to be developed by Richard FitzRalph and then turned against the church by John Wyclyf.

See also Aristotelianism; Aristotle; Augustinianism; Gregory of Rimini; Liber de Causis; Neoplatonism; Proclus; Thomas Aquinas, St.; Thomism; Wyclyf, John.

Bibliography

WORKS BY GILES

Commentaries on Aristotle

Physics. Padua, 1483; Venice, 1492.

Posterior Analytics. Venice, 1488.

De Anima. Venice, 1496.

Prior Analytics. Venice, 1499.

Expositio in Artem Veterem. Venice, 1507.

Metaphysics. Venice, 1550.

Other Works

Quodlibeta. Venice, 1496.

"De Intellectu Possibili" and "De Gradibus Formarum." In *De Anima.* Venice, 1550.

Super Authorem Libri de Causis Alfarabium. Venice, 1550.

Commentary on the Sentences. Córdoba, 1707.

De Ecclesiastica Potestate. Edited by R. Scholz. Weimar: Böhlaus Nachfolger, 1929.

Aegidii Romani Theoremata de Esse et Essentia, edited by Edgard Hocedez. Louvain, 1930.

WORKS ON GILES

Philosophy

Boffito, G. *Saggio di bibliografia egidiana.* Florence, 1911. A bibliography.

Egenter, R. *Die Erkenntnispsychologie des Aegidius Romanus.* Regensburg, 1926.

Hocedez, E. "Gilles de Rome et Henri de Gand sur la distinction réelle, 1276–87." *Gregorianum* 8 (1927): 358–384.

Hocedez, E. "Gilles de Rome et saint Thomas." In *Mélanges Mandonnet,* edited by P. Mandonnet. 2 vols. Paris: Vrin, 1930. Vol. I, pp. 385–409.

Makaay, S. *Der Traktat des Aegidius Romanus über die Einzigkeit der substantiellen Form.* Würzburg, 1924.

Mandonnet, P. "La carriere scolaire de Gilles de Rome (1276–1291)." *Revue des sciences philosophiques* 4 (1910): 480–499.

Nardi, B. "Egidio Romano e l'avverismo." *Rivista di storia della filosofia* 3 (1947): 197–220; 4 (1948): 1–22.

Paulus, J. "Les disputes d'Henri de Gand et Gilles de Rome sur la distinction de l'essence et de l'existence." *Archives d'histoire doctrinale et littéraire du moyen âge* 15 (1940–1942): 323–358.

Science

Maier, A. *An der Grenze zwischen Scholastik und Naturwissenschaft im 14 Jahrhundert.* Rome: Edizioni di Storia e Letturatura, 1951.

Maier, A. *Metaphysische Hintergründe der spatscholastischen Philosophic.* Rome: Edizioni di Storia e Letturatura, 1956.

Maier, A. *Die Vorläufer Galileis,* 2 vols. Rome, 1949.

Maier, A. *Zwei Grundprobleme der scholastische Naturphilosophie.* Rome: Edizioni di Storia e Letturatura, 1951.

Maier, A. *Zwischen Philosophic und Mechanik.* Rome: Edizioni di Storia e Letturatura, 1958.

Political Philosophy

Gwynn, Aubrey. *The English Austin Friars in the Time of Wyclif.* London: Oxford University Press, H. Milford, 1940.

Bibliography

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages,* 735–736. London: Sheed and Ward, 1955. Complete bibliography.

Gordon Leff (1967)

GILSON, ÉTIENNE HENRY

(1884–1978)

Étienne Henry Gilson, the French neo-Thomist philosopher, was born in Paris. His higher education was acquired at the University of Paris. In 1907 he received his agrégé and in 1913, after several years of teaching, his doctorate, publishing both his minor and major theses, *Index scolastico-cartésien* and *La liberté chez Descartes et la théologie.* The years 1914–1916 saw Gilson serving France as an officer on the battlefield. Captured at Verdun, he was a prisoner of war from 1916 to 1918. He spent two years as professor of philosophy at the University of Strasbourg and in 1921 became professor of the history of medieval philosophy at the Sorbonne, in which position he served until 1932, when he accepted the chair of the history of medieval philosophy at the Collège de France, where he taught until 1951. Gilson cooperated with members of the Congregation of Priests of St. Basil of Toronto, at their invitation, to found, in 1929, the Institute of Medieval Studies, in association with Saint Michael's College in the University of Toronto. He was a professor and director of studies at the institute from its foundation until 1956.

Numerous leading universities conferred honorary degrees on Gilson, and many invited him to deliver prominent lectureships, among them the Gifford Lectures at the University of Aberdeen (1930–1931), published as *The Spirit of Medieval Philosophy*; the William James Lectures at Harvard (1936–1937), published as *The Unity of Philosophical Experience*; the Richard Lectures at the University of Virginia (1937), published as *Reason and Revelation in the Middle Ages*; the Mahlon Powell Lectures at the University of Indiana (1940), published as *God and Philosophy*; and the A. W. Mellon Lectures in the fine arts (1955), published as *Painting and Reality.* Gilson founded and directed the famous *Études de philosophie médiévale* and the *Archives d'histoire doctrinale et littéraire du moyen âge* and was a director of *Medieval Studies,* the annual publication of the Pontifical Institute of Medieval Studies. Among the many academies and societies of which he was a member was the select French Academy, to which he was elected in 1947.

PHILOSOPHICAL POSITION

Gilson's main thoughts may best be appreciated in company with two parts of his own intellectual history. (1) The great Jewish scholar Lucien Lévy-Bruhl advised Gilson to study the relation between René Descartes and Scholasticism. From this research Gilson learned to read St. Thomas Aquinas and to recognize that the metaphysical conclusions of Descartes made sense only in the context of Thomas's metaphysics. (2) Further study of Thomas and other medieval thinkers from St. Augustine through William of Ockham proved for Gilson that there was no common philosophy employed within the theologies but, rather, there were different authentic philosophies.

To do the choosing, demonstrating, and judging that he considered one of the proper tasks of philosophy, Gilson gradually developed his personal philosophical position. The only philosopher, Gilson maintained, who made him clearly realize the full metaphysical implications of the major problems was Thomas, a fact that in no way lessened Gilson's intellectual freedom, for he always wanted to be free to agree with somebody when he thought that what was said was right. For him what characterized Thomism is the decision to locate the act of existence in the heart of the real as an act that can be grasped only by or in the essence whose act it is, as an act, therefore, that has primacy not over and above being but within being. Thus, Thomism as an authentic existentialism is opposed equally to the "Thomistic" essentialists, who deposit a dead essence in the mind as a quiddity without preserving its contact with the act of being, and to such existentialisms as those of Søren Kierkegaard, Martin Heidegger, Karl Jaspers, and Jean-Paul Sartre, which, although divergent from one another, commonly deal with existence only as an object of a possible phenomenology of human existence and are phenomenologies still in search of ontologies.

Gilson's personal commitment to the existentialism of Thomism was related to one of his most central philosophical doctrines—namely, the reality and philosophical validity of what he terms *Christian philosophy.* In *The Spirit of Medieval Philosophy* and in many other books and articles Gilson demonstrated that the Christian religion and its theologies have had the capacity to produce metaphysical conclusions and to transform philosophy itself. Several of its philosophical ideas the Greek philosophers never knew—for example, the existence of a unique God, the infinite, simple, supremely free Creator of the universe, as an all-powerful efficient cause, as well as the existence of man as a substantial composite of soul and body, free, made in God's image. Regarding the philosophical problem of how a speculation can be rational and philosophical if it is connected with religious beliefs, history as such is incompetent to answer, but philosophy provides the answer. History shows that the alliance of the two distinct orders of thought has produced positive philosophical results. Although Gilson recognizes, with Thomas and other medieval theologians, the distinction of philosophy and theology, he opposes their separation as practiced by Descartes and by numerous neo-Scholastics from the sixteenth century to the present day, for whom philosophy became no more than temporary and successively different alliances with any sort of currently fashionable philosophical position that could be reconciled with revelation.

As Gilson saw it, in the medieval theologians what could be philosophically demonstrated received in theological works the full benefit of rational demonstration. Such philosophical demonstrations were part of sacred doctrine and were also philosophy because they were reached by the human intellect through its own light. In the case of Thomas, who represents for Gilson the best in Christian philosophy, the philosophy is that of a theologian with the order of development required for theological ends; hence, one cannot release Thomistic philosophy from its theological moorings without running the risk of not knowing its origin and end, of altering its nature, and even of not grasping its meaning. Apart from the historical fact of the nonseparation of philosophy and theology, Gilson was convinced that the very nature of philosophy

does not demand that the philosophy of Thomas be extracted from the world of faith and the influence of revelation. Philosophy has been and can be authentically philosophy and Christian at one and the same time, for the orientation of Christian philosophy—to knowledge about God and man—entails no a priori exclusion of any area of philosophical research because nothing in the universe is irrelevant to knowledge about God and man.

This central theme in Gilson's philosophy—the nature and validity of Christian philosophy—has exasperated so-called Thomists seeking to develop a Thomism separate from theology; to rationalists it has seemed not to be philosophy at all. Gilson tirelessly represented the historical evidence and philosophical reasons to identify and justify Christian philosophy as the use that the Christian makes of philosophical reason when he associates religious faith and philosophical reflection. Rhetorically, Gilson asked why those who profess the Christian faith and its doctrines should see themselves excluded from philosophy simply because they prefer to philosophize about what they believe.

Other influential aspects of Gilson's philosophical doctrines concern education, social and political philosophy, the philosophy of art, and the history of modern and contemporary philosophy. In *Painting and Reality,* Gilson interprets the evolution of the art of painting, especially its most recent phases, in the light of his existential metaphysics. Because artistic beauty is made, not found, Gilson opposes mere imitation as artistic beauty; the function of any work of art *qua* art is solely to cause in us the contemplative pleasure of enjoying it. In a masterful defense Gilson analyzes the history of art from Leonardo da Vinci to the mid-twentieth century, demonstrates that representation is not of the essence of art, and argues for the legitimacy of abstraction and the necessity to sacrifice all elements of reality that do not contribute to the plastic structure of a work.

See also Art, Representation in; Augustine, St.; Descartes, René; Existentialism; Heidegger, Martin; History and Historiography of Philosophy; Jaspers, Karl; Kierkegaard, Søren Aabye; Leonardo da Vinci; Lévy-Bruhl, Lucien; Medieval Philosophy; Sartre, Jean-Paul; Thomas Aquinas, St.; Thomism; William of Ockham.

Bibliography

WORKS BY GILSON

The Christian Philosophy of St. Augustine. Translated by Lawrence Lynch. New York: Random House, 1960.

Elements of Christian Philosophy. Garden City, NY: Doubleday, 1960.

The Philosopher and Theology. New York: Random House, 1962.

Modern Philosophy, Descartes to Kant. New York: Random House, 1962. Written with Thomas Langan.

Introduction aux arts du beau. Paris: Vrin, 1963.

The Spirit of Thomism. New York: P.J. Kenedy, 1964.

Christian Philosophy: An Introduction. Toronto: Pontifical Institute, 1993.

For works published before 1960, see Callistus James Edie, ed., *Mélanges offerts à Étienne Gilson* (Paris: Vrin, 1959), a bibliography of all articles, monographs, books, and translations of books by Gilson.

WORKS ON GILSON

McGrath, Margaret. *Étienne Gilson: A Bibliography.* Toronto: Pontifical Institute, 1982.

Quinn, John. *The Thomism of Étienne Gilson.* Villanova, PA: Villanova University Press, 1971.

Redpath, Peter A., ed. *A Thomistic Tapestry: Essays in Memory of Étienne Gilson.* New York: Rodopi, 2003.

Shook, Laurence. *Étienne Gilson.* Toronto: Pontifical Institute, 1984.

Robert G. Miller, C.S.B. (1967)
Bibliography updated by Christian B. Miller (2005)

GIOBERTI, VINCENZO
(1801–1852)

Vincenzo Gioberti, the Italian philosopher, ecclesiastical polemicist, and statesman, was born in Turin. As a statesman he upheld federalism as the goal of the movement for Italian unity. Gioberti's *Del primato morale e civile degli Italiani* (Brussels, 1843) is one of the great documents of the Risorgimento. His most famous polemical work is *Il gesuita moderno* (5 vols., Lausanne, 1846–1847), attacking the allegedly reactionary influence of the Jesuits on church policy. Throughout his intensely active career, philosophy remained his dominant interest. A long political exile (1833–1845) provided the occasion for the composition of his most important philosophical works: *Teorica del sovrannaturale* (Brussels, 1838), *Introduzione allo studio della filosofia* (Brussels, 1840), and *Degli errori filosofici di Antonio Rosmini* (Brussels, 1841; 2nd enl. ed., 3 vols., 1843–1844).

"PROTOLOGIA"

In 1841 and 1842 Gioberti gave a course of lectures (published as *Cours de philosophie,* Milan, 1947). The second part of these lectures, "Protologie ou science première," was the first sketch of a subject of which many of Gioberti's works can be considered fragmentary studies.

The term *protologia* may derive from the title of a work by Ermenegildo Pini (1739–1825) that was published in 1803. Gioberti envisaged *protologia* as "the science of the creative act and of the ideal formula which expresses it completely." Its complement is *deuterologia,* the theory of the sciences constructed by reflection on the basis of being as it is intuited. *Protologia* has three divisions: theology, logic, and cosmology, which includes psychology. The division arises out of the three elements of the ideal formula, "Being creates the existent." *Protologia* escapes the subject-object dichotomy; it studies neither the subject nor the object but the intelligible principle that relates the two.

ONTOLOGISM

Because of his constant affirmation that being-in-itself is constitutively present to the human intellect, Gioberti's philosophical position is generally described as ontologism. Being is present to the intellect as thought, not as a sensible property of mind itself; that is, the being which is present to the mind is not merely the being of the mind but being itself. Gioberti rejected what he called sensism, by which he meant the view that the being present to the human mind is simply its own being apprehended by the senses. Gioberti asserted that the activity of the human mind "concreates" its object in conjunction with the presence of being. The being constitutively present to the human mind is not merely possible being but real being, and indeed the most real being. This being is indeterminate, not in the sense that it lacks all distinction but in the sense that all distinctions are so related and fused that the human mind does not immediately succeed in discerning them. For this reason the original intuition of being can make known the existent only in conjunction with sensible experience. Sensible experience makes the existent present, but the existent is known by virtue of being. The existent is not, however, a part, determination, or moment of being but a creation of being. Existents are present in being as elements of its creative possibility, not as its modes or qualities (on this point Gioberti thought that he was in disagreement with Benedict Spinoza). The act of thought renders the existent present, and since this act is being and is act only insofar as it is being, it concreates the existent. The act of judgment, which is the pure form of knowledge, has a particular form for thus establishing any particular existent. Its ideal form, or "ideal formula," informs every judgment independently of its particular concern. This ideal formula, "Being creates the existent," is the presence of the pure form of the judgment in its pure possibility.

LANGUAGE

Ontologism shows that thought is a creative act. The object of thought comes into being through the operation of thought in the word. In this view Gioberti was strongly influenced by Giambattista Vico, and like Vico he studied the problem of language as the general theory of the word. Language is the specific manner in which the concreative operation of thought and being is effected in the human matrix. The concreative act, to be actual and effective, must reflect the human condition—it must be psychologized, but not in the manner which Gioberti opposed as "sensism." Gioberti advocated a transcendent psychologism in which the transcendental operation is the transaction of a concrete existent subject. Language is transcendental in that it reflects the constitutive presence of being in accordance with the structure of the human mind while enabling the mind to transcend its own existential limits and achieve universal significance. In this process of transcendence the ideal formula is specified according to the form not only of the object but also of the subject. Language places knowledge beyond the subject-object split. It reduces the being mentioned in the ideal formula to the effective principle of concrete science.

The transcendental operation of language seems to be widely distributed throughout the numerical range of human subjectivity. Gioberti suggested that it is a transaction of particular groups and ultimately of the nation. This view was important in *Del primato morale e civile degli Italiani* and was given an aristocratic cast in the claim that within the nation the transcendental operation is the work only of an elite.

"PALINGENESIS"

Language constitutes the first movement of the life of the spirit, a movement implicit in the ideal formula, "Being creates existence." It is thus at the focus of the genesis of existence and the real as object. The pure form of human thought moves from the immediate givenness of the existent to its ideal ground. In experience we encounter the end product of the movement expressed in the ideal formula. Thought must return the existent thus encountered to its ground in being. Its movement is therefore a *palingenesis*—a return of existence to being.

Mimesis is the mode of existence of that which is encountered in experience. It is a state, but not a radical or irremediable one, of alienation. *Methexis* is the state of the thinking subject, the intelligence whose constitutive principle is the intelligible, that is being. *Methexis* is the

link that ends the alienation between being and existence in *palingenesis.*

Gioberti did not mean by *palingenesis* a dissolution of the distinctions of existence into the indeterminacy of ideal being. Rather, through *palingenesis* the being of the existent qua existent (that is, in its distinctness) is ideally grounded in being. Thus, being itself needs the movement of *mimesis* down into the world in order that it may come into its own actuality, or distinctness, through *methexis.* Hence, being is not absolutely transcendent. It reaches its actuality in the word and thus belongs inalienably to the region of culture and history. Gioberti's theory of language thus contains in germ a theory of culture and history. Culture and cultures are the historical forms of *palingenesis.*

THEOLOGY, POLITICS, AND ETHICS

The *palingenesis* of being is the central operation of the spirit and determines the actual form of the world. The fact that this one process can be studied from two points of view provided Gioberti with a basis for a correlative distinction and unity of theology and philosophy. Theology possesses a certain superiority deriving chiefly from its object, God. Supernatural theology does not, however, take possession of the internal, or constitutive, word of God; it must make use of analogies drawn from philosophy. Supernatural revelation makes use of the "natural revelation" of the word. Philosophy is therefore superior to theology in that it provides the interpretative categories of theology.

Palingenesis takes on deontological status as the supreme norm of action. In this aspect it is the axiological principle of both the moral and the political orders. In both morals and politics, in conscience and in law, the essential process is the return of existence to being. Similarly, Gioberti held that the church is the historical and institutional form of the *palingenesis* of being under the dispensation of revelation.

Gioberti's thought is still influential in two of the leading strands of contemporary Italian thought, Gentilean actual idealism and Christian spiritualism.

See also Being; Cosmology; Deontological Ethics; Language; Ontology; Psychologism; Vico, Giambattista.

Bibliography

WORKS BY GIOBERTI

Opere edite e inedite di Vincenzo Gioberti. Rome and Milan: F. Bocca, 1938–. Edited by the Istituto di Studi Filosofici. 8 vols. of works and 11 vols of correspondence to date.

Della protologia. Edited by G. Massari, 2 vols. Paris and Turin, 1857.

Nuova protologia. Edited by Giovanni Gentile. Bari: Laterza, 1912.

Scritti scelti di Vincenzo Gioberti. Edited by Augusto Guzzo. Turin, 1954. Contains an important introduction.

Saitta, Armando, Andrea Luigi Mazzini, and Vincenzo Gioberti. *Sinistra hegeliana e problema italiano negli scritti di A.L. Mazzini*. Rome: Istituto italiano per l'età moderna e contemporanea, 1968.

WORKS ON GIOBERTI

Bonafede, G. *Gioberti.* Palermo, 1942.

Bonafede, G. *Gioberti e la critica.* Palermo, 1950.

Bruers, A. *Gioberti.* Rome: Fondazione Leonardo per la Cultura Italiana, 1924. Contains an excellent bibliography.

Giusso, Lorenzo. *Gioberti.* Milan: Garzanti, 1948.

Ultima replica ai municipali e Preambolo dell'Ultima replica ai municipali. Edited by Claudio Vasale. Padova: CEDAM, 1976.

Introduzione allo studio della filosofia. Padova: CEDAM, 2001.

Stefanini, Luigi. *Gioberti.* Milan: Fratelli, 1947. The best modern treatment.

A. Robert Caponigri (1967)

GIVEN, THE

See *Basic Statements*

GLANVILL, JOSEPH

(1636–1680)

Joseph Glanvill was a skeptic, a prominent defender of the experimental research of the early Royal Society, a liberal rationalistic Anglican theologian and preacher, and a staunch and influential believer in witchcraft. He studied at Cambridge, where he came under the influence of Henry More. On first learning of René Descartes's work Glanvill became an advocate of Cartesianism but was quickly led to cast doubt on it as a metaphysical theory because of More's objections. He then treated Cartesianism as a working hypothesis and began analyzing how much certitude anyone could have about what is going on in the world. He came into contact with John Wilkins, the bishop of Chester, and began developing his case in terms of the categories employed by him.

Glanvill's first work, *The Vanity of Dogmatizing* (1661), was soon revised into the larger *Scepsis Scientifica* (1665), and began with a most laudatory "Address to the Royal Society," which led to Glanvill being elected as a fellow.

GLANVILL AND SKEPTICISM

Glanville saw the skeptical problem as one that could not be so easily set aside. He saw the reliability of one's faculties as central for avoiding any ultimate and overwhelming skepticism. But Glanvill saw that the kind of certainty one would need to be absolutely sure of one's faculties ("infallible certainty," in which one is assured, "'tis impossible things should be otherwise than we conceive them or affirm them") is unattainable—"for it may not be absolutely impossible, but that our Faculties may be so construed, as always to deceive us in the things we judg most certain and assured."

One may not be able to attain infallible certitude, but one can attain indubitable certitude that one's faculties are true. This is indubitable in two senses: first, that one finds that one has to believe them, and, second, that one has no reason or cause for doubting them. One has to believe one's faculties are reliable if one is to have any rational life at all, even though one has no evidence that one's faculties are, in fact, reliable.

Glanvill carried this on to base acceptance of historical data (and especially that of scripture) on the indubitable principle that "Mankind cannot be supposed to combine to deceive, in things wherein they can have no design or interest to do it." So, skepticism can be set aside in mathematics, science, history, and theology, because one has no actual reason to doubt the results in these areas. One has to believe various findings and act with confidence. But, having said this, Glanvill immediately made clear that he had not offered or provided any way of eliminating ultimate skepticism.

For Glanvill, reasons for doubting had to be reasonable. Descartes's reasons for doubting he dismissed as hyperbolic or metaphysical. No reasonable person would entertain them. On the contrary, there can be reasonable doubts about many things, but this does not prevent one from having a degree of certitude about other matters. Glanvill insisted that human beings are basically in a state of ignorance due to the original Fall. They cannot know the springs and principles by which the world is operating. They can only hypothesize about this and recognize that any hypothesis could be false. There is a reasonable basis for doubting in that one never has sufficient evidence or knowledge and one cannot be sure that things cannot be otherwise than one conceives them.

Glanvill introduced what was to be an important point in later scientific thought, namely, that one can never find necessary connections between events. Any causal hypothesis that one works out is always open to question and doubt, since one does not understand the inner workings of Nature. One can find concomitances of events (what David Hume later called constant conjunctions) but not necessary connections. Because of this analysis of one's causal reason, Glanvill has often been considered a precursor of Hume, although there is no evidence that Hume ever read any of his work.

REASON AND RELIGION

Glanvill's discussion of the relation of reason and religion is perhaps his most original contribution—that of offering a rational-skeptical fideism as a way of living with irremediable skepticism. Glanvill made the acceptance of the reliability of one's faculties a genuine act of faith. "The belief of our Reason is an Exercise of Faith, and Faith is an Act of Reason." He had preceded this by stating that "Reason is certain and infallible," which turns out to be based on one's knowledge "that first Principles are certain, and that our Senses do not deceive us, because God that bestowed them upon us, is True and Good."

Glanvill was not emulating Descartes in making true knowledge depend on the proof that God is not a deceiver. Rather, Glanvill was offering a kind of rational fideism. Faith, and faith alone, is the basis for one's belief in reason. One believes in reason because one believes in God's veracity. One does not try to prove that God is truthful; one believes this. Thus, faith in God gives one faith in reason, which in turn "justifies" one's belief that God is no deceiver.

Glanvill saw that the ultimate guarantee of one's certitude depends not on what one can prove, but on what one can believe. One can believe that God is truthful, and hence believe in the reliability of one's faculties. The first belief is reasonable, since one has no reason to doubt of it. This, then, enables one to avoid ultimate skepticism, by avoiding the fundamental skeptical problem of proving one's first principles.

Glanvill's rational fideism grows out of seeing the conditions requisite for certain and unquestionable reasoning (namely, that God is reliable), and is in sharp contrast to the irrational fideism being offered in the late seventeenth century by Pierre Bayle and Pierre Jurieu. Glanvill posed the possibility that rationality could be based on faith, and in terms of what human beings consider reasonable, accepting such faith is an exercise of reason. Using this rational fideism, Glanvill tried to show the reasonableness of religious belief and of Latitudinarian Christianity.

Glanvill provided an epistemology for a "mitigated" skepticism, which could delineate the kind of certitude that the new scientists could find. Instead of basing the "new science" on dogmatic metaphysical principles, he offered an undogmatic semiskepticism sufficient to encourage the nondogmatic inquiries of the scientists of the Royal Society, while opposing the dogmatism of Descartes, Thomas Hobbes, and Benedict (Baruch) de Spinoza.

Glanvill's belief in witches comes from his critique of the materialism of Hobbes and others. The question of whether evil spirits exist, Glanvill pointed out, is a factual question, not a metaphysical one and has to be answered by examining the empirical evidence. Glanvill compiled ample testimonials to convince any "reasonable" person that (1) it is possible that evil spirits or witches exist, (2) it is probable that they do, and (3) that the acknowledgment of their existence allows for the best explanation of various observed phenomena. Glanvill pointed out that various societies have laws against practicing witchcraft, so it seems likely that there is something of this sort that could be practiced. The possible existence of witches is also part of a larger and more significant question—that of the existence of spirits. If demonic or evil spirits cannot exist, then how can one be sure that good spirits—angels or God—can exist? To deny the possibility of the existence of witches is to deny the possibility of any sort of spiritual or divine world.

GLANVILL'S CARTESIANISM

Glanvill offered a nondogmatic, or deontologized, Cartesianism as the best scientific model of explanation for natural phenomena. In his continuation of Francis Bacon's *New Atlantis* Glanvill had his sage present Cartesianism as "the neatest Mechanical System of things that had appear'd in the world," though it was not certain or all encompassing. The sages could also accept the preexistence of the soul and the existence of spiritual agents, whose manner of operating may not be known or even knowable to one.

Glanvill was an eclectic philosopher, taking his views in part from More, Descartes, Bacon, Anne Conway, and the members of the Royal Society. Glanvill's world of natural science, spirits, and Christianity, based on the "plausible" testimony of historical documents, is one way these kinds of knowledge could be brought into harmony. Glanvill paid the price of having this all rest on a basically ineliminable skepticism. If one could find solace and comfort in a faith in a nondeceiving Deity, then a nice, harmonious world of science and religion could be accepted.

See also Cartesianism; Fideism; Skepticism, History of.

Bibliography

WORKS BY GLANVILL

The Vanity of Dogmatizing. London: Printed by E. C. for H. Eversden, 1661.

Scepsis Scientifica. London: Printed by E. Cotes for H. Eversden, 1665.

A Philosophical Endeavor in the Defense of the Being of Witches and Apparitions. London: Printed by E. Cotes for James Collins, 1668a.

Plus Ultra: or, the Progress and Advancement of Knowledge Since the Days of Aristotle. London: Printed for J. Collins, 1668b.

Essays on Several Important Subjects in Philosophy and Religion. London: Printed by J. D. for J. Baker, 1676.

WORKS ABOUT GLANVILL

Cope, Jackson I. *Joseph Glanvill, Anglican Apologist*. St. Louis, MO: Committee on Publications, Washington University, 1956.

Popkin, Richard H. "The Development of the Philosophical Reputation of Joseph Glanvill." *Journal of the History of Ideas* 15 (1954): 305–311.

Popkin, Richard H. "Joseph Glanvill: Precursor of Hume." *Journal of the History of Ideas* 14 (1953): 292–303.

Talmor, Sascha. *Glanvill: The Uses and Abuses of Scepticism*. Oxford, U.K.: Pergamon Press, 1981.

Van Leeuwen, Henry G. *The Problem of Certainty in English Thought, 1630–1690*. The Hague: Nijhoff, 1963.

Richard H. Popkin (1967, 2005)

GNOSTICISM

"Gnosticism" (from the Greek *gnosis*, "knowledge") designates a broad variety of religious teachings that were rife in the Hellenized Near East of the first centuries CE and purported to offer knowledge of the otherwise hidden truth of total reality as the indispensable key to man's salvation. Most of the schools or sects in question were ostensibly Christian by the time our earliest witnesses, the Church Fathers, were familiar with them, and in consequence the whole movement was long regarded as essentially an aberration from Christian doctrine. However, although Gnosticism provided the first chapter in the history of Christian heresies, the Christian veneer of the systems playing that role is often thin to the point of transparency; and clearly non-Christian writings have come to light that by all criteria of content must be classed as Gnostic as well. The details of the literary evidence point to highly syncretistic origins, in which Jew-

ish, Iranian, Babylonian, Egyptian, and other Oriental traditions were blended with one another and with Greek concepts in an extremely free manner. The results were as readily made to represent an alleged esoteric truth of the Christian message as to constitute a superior (Mani) or even hostile (Mandaeans) alternative to it.

This syncretism, pertaining mainly to the outer shell, does not preclude—in fact it tends to mask—a highly original inner unity of thought distinct from all the disparate historical elements employed in its representation. Massively mythological though this representation usually is, the substance thus expressed has philosophical significance as embodying a fundamental choice—the radical antithesis to the classical Greek choice—in the realms of universal theory and human practice at once. The powerful Gnostic impulse to elaborate its basic vision into grandly constructed, quasi-rational systems of thought where everything proceeds from an absolute beginning makes Gnosticism a landmark in the history of the speculative system as such; and it is the identity of that basic vision that defines what is Gnostic and alone justifies the classing of systems of such considerable diversity under one heading.

GNOSTIC TEACHERS AND SCHOOLS

A number of gnostic teachers and writers are known by name (mainly those listed as heresiarchs in the patristic refutations), but much of the surviving literature is anonymous or pseudepigraphic, in keeping with the revelatory style in which it is cast. Historical individuals whose thought is documented by either critical accounts or direct fragments of their works include the Samaritan Simon Magus and his spiritual descendants Menander, Saturninus, Cerinthus, and Cerdon (first and second centuries); the Alexandrians Carpocrates, Basilides and his son Isidore, and, foremost, Valentinus with his illustrious disciples Ptolemaeus, Heracleon, Theodotus, and Marcus (second century); the Pontian Marcion and the Syrian Bardesanes (second century); and the Persian-Babylonian Mani (third century). Major sects whose doctrines are well documented but not identified by individual authors or founders are, in the Christian camp, the Barbeliotes, Sethites, and Ophites (the last actually a cluster of sects); in the Hellenistic-pagan camp, the Hermetic religion (perhaps merely a literature and not an actual sect); in the Semitic East, the anti-Christian Mandaeans. Towering over the known thinkers are Valentinus, Marcion, and Mani; and Valentinianism and Manichaeism respectively represent the culminations of the two main alternative types of Gnostic speculation. The last two are here considered merely for their part in and exemplification of the wider context.

SOURCES

With the exception of that of the Mandaeans, Gnostic literature was denied direct tradition under the dominion of Christianity and Islam after the eclipse of the Gnostic communities themselves. Thus, until fairly recently, information was supplied almost solely by the abundant indirect sources. These were, in the main, the antiheretical works of the Church Fathers (Greek, Latin, and Syriac, from Irenaeus in the second century to Theodore bar Konai in the eighth century) with their diligent reports, summaries, and excerpts, and still later Islamic histories and compendia. However, for some time an impressive series of manuscript discoveries has been adding vastly to our store of original texts: Coptic-Gnostic papyrus codices from Egypt, belonging to the Christian branch of Gnosticism—the find in 1945 of a whole library at Nag Hammadi is revolutionizing the state of documentation in the area hitherto principally covered by the patristic testimony—Manichaean fragments in Persian, Turkish, and Chinese from Turfan in central Asia and in Coptic from Egypt, and the sacred writings of the Mandaeans of Iraq.

The Mandaeans are the one case of a Gnostic community surviving to the present with an unbroken written tradition of their voluminous Aramaic literature; it came to the attention of Western scholars in the nineteenth century, after it had escaped that of the Church Fathers in antiquity (probably because of the Fathers' predominantly Greek orientation). In all the other cases, the new original sources generally bear out, while greatly enriching, the testimony of the older indirect evidence. The following account, based on the entire, extremely varied material, is synoptic and selective, placing its emphases according to a conception of the whole as a system.

GNOSTIC DUALISM

A radically dualistic mood dominates the Gnostic attitude and unites its widely diversified expressions, whether doctrinal, poetical, or ethical. The dualism is between man and world, and between the world and God. In either case, it is a dualism of antithetical, not complementary, terms; and it is basically one: that of man and world mirrors on the plane of experience the primordial one of God and world and is, in Gnostic theory, deduced from it. The interpreter may hold conversely that the transcendent doctrine of a world-God opposition

sprang from the immanent experience of a disunion of man and world, that is, it reflects a human condition of alienation. In the three-term configuration, man and God belong in essence together against the world but are in fact separated by the world, which in the Gnostic view is the alienating, divisive agency.

The object of Gnostic speculation is to derive these basic polarities—the existing state of things—by way of genetic myths from the first things and through such genealogy to point the way to their eventual resolution. The myth, a conscious symbolical construction, is thus predictive by being genetic, eschatological by being explanatory. Accordingly, the typical Gnostic system starts with a doctrine of divine transcendence in its original purity; traces the genesis of the world from some primordial disruption of this blessed state, a loss of divine integrity that leads to the emergence of lower powers who become the makers and rulers of this world; then, as a crucial episode in the drama, it recounts the creation and early fate of man, in whom the further conflict becomes centered; the final theme—in fact, the implied theme throughout—is man's salvation, which is more than man's, since it involves the overcoming and eventual dissolving of the cosmic system and is thus the instrument of reintegration for the impaired godhead itself, the self-saving of God.

GOD AND THE DIVINE REALM. The transcendence of the supreme deity is stressed to the utmost degree in all Gnostic theology. Topologically, he is transmundane, dwelling in his own realm entirely outside the physical universe, at immeasurable distance from man's terrestrial abode; ontologically, he is acosmic, even anticosmic: To this world and whatever belongs to it he is the essentially "other" and "alien" (Marcion), the "alien Life" (Mandaeans), the "depth" or "abyss" (Valentinians), even "the not-being" (Basilides); epistemologically, because of the transcendence and otherness of his being, and because nature neither reveals nor even indicates him, he is naturally unknown, ineffable, defying predication, surpassing comprehension, and strictly unknowable. Some positive attributes and metaphors do apply to him: Light, Life, Spirit, Father, the Good—but not Creator, Ruler, Judge. Significantly, in some systems one of his secret names is Man. Mainly, the discourse about him must move in negations, and historically Gnosticism is one of the fountainheads of negative theology.

However, the Absolute is not alone but is surrounded by an aura of eternal, graded expressions of his infinitude, partial aspects of his perfection, hypostatized into quasi-personal beings (aeons) with highly abstract names (mostly of mental properties) and together forming the hierarchy of the divine realm, the pleroma (Plenitude). The emanation of this inner manifold from the primal ground, a kind of self-differentiation of the Absolute, is sometimes described in terms of subtle spiritual dialectics, more often in rather naturalistic (for instance, sexual), terms. Among the tenuously mythological entities that thus arise (such as Mind, Grace, Word, Knowledge, Life) are two more concrete ones with definite roles in the further evolution of the transcendental drama: Man as an eternal, divine, precosmic principle (sometimes even identified with the First Being himself) and Wisdom (Sophia), usually the last and youngest of the aeons. Extensive speculation about the diversity within the pleroma is the mark of advanced systems, but some degree of manifold on the upper reaches of being is requisite for all Gnostic metaphysics because it provides the condition for divine fallibility on which the movement into creation and alienation depends.

LOWER POWERS AND THE CREATOR. In the genuine Gnostic systems the downward movement starts from an internal crisis in the divine realm itself, whereas in those under Iranian influence it is occasioned by the action of dark forces from without, thus presupposing the very dualism that the typical speculation lets evolve from the one monistic root. We shall mainly follow this latter, more prevalent type, which is free from Iranian influence. Here, the protagonist of crisis and fall is most often the female aeon Sophia (or such equivalents as Thought and Conception) who, from some overstepping of bounds—assertion of self-will, creative presumption, even excessive desire to know the unknowable Father—is drawn into a history of passion and error that leads her outside the blessed pleroma. (In another family of systems, Primal Man assumes the role of the sinking part of divinity.) Although the upper powers immediately set about healing this breach in the divine order, the downward trend set in motion by the original lapse must take its course, and the counterplay of these two trends henceforth governs the process. There ensues, in a development too complex and too variously elaborated to recount here, a train of ever lower hypostases descended from the erring Sophia, episodically broken by certain archetypal salvations.

The Demiurge. Early in the descending series—and marked with all the deforming effects of the Fall whose fruit he is—appears the Demiurge, the monstrous and benighted archon (lord) of the nether powers. This widespread Gnostic figure, telling symbol of the Gnostic hos-

tility toward the world, is clearly a polemical caricature of the Old Testament God, and the identity is made explicit by frequent transference to him of well-known utterances and actions of God from the biblical text. Pride, ignorance, and malevolence of the Creator are recurring themes in Gnostic tales, as are his humbling and outwitting by the higher powers bent on thwarting his designs. However, over the whole range of Gnostic mythologizing the archon's image varies, and there are milder versions in which he is more misguided than evil and thus open to correction and remorse, even to final redemption. He is always a problematical and never a venerable figure.

Finding himself in the void or chaos outside the pleroma, possessed of the power inherited from his mother but ignorant of the divine worlds above him, he believes himself to be the only God and engages in creations chiefly designed to satisfy his ambition, vanity, and lust for dominion. Prominent among the host of lower powers that issue from him are six further archons whom he installs in six successive heavens; he occupies the seventh above them. Thus originate the cosmic order and its system of rule, the universe of Babylonian astrology with its seven planetary spheres and the almighty planetary deities. An eighth region beyond them (corresponding to the sphere of the fixed stars) is occupied by the mother Sophia, still exiled from the pleroma, who has no part in the creation and government of the world but intervenes in both for the purposes of salvation. The Valentinian version, the subtlest of all, depicts the Demiurge as trying vainly to imitate the perfect order of the aeons with his physical one, and their eternity with the counterfeit substitute of time—thus adding to the parody of the biblical Creator that of the Platonic Demiurge. However, the chief instance of illicit and bungling imitation is the creation of man.

The remaining part of creation is the joint work of the seven archons. Indeed, the early systems (such as that of Simon Magus) simply name the seven as the creators of the world; and the preeminence of one of them, growing into a kind of monotheism of cosmic (lower) divinity, seems to be characteristic of the mature stage of Gnostic speculation. There, an episode, told with almost identical words in the cosmogonies of many different schools, rings in the next act in the drama of creation: The First Archon (the Demiurge), exulting in his works with the Scriptural proclamation "I am God and there is none other than I," draws the retort from on high, "Thou art mistaken! Above thee is First Man."

CREATION OF EARTHLY MAN. Some such divulgence of superior godhead (here meant as no more than a humbling of the Creator's pride, elsewhere serving some other purpose in the divine strategy), and especially the appearance of a divine form with it, inspire the archons with the audacious plan to equal the upper perfection in a work of their own—to create terrestrial man—an effect not foreseen in the divine move. Letting them say on this occasion, "Come, let us make a man after the image we have seen," the Gnostics turned to account the puzzling plural of Genesis 1:26, and the resulting *imago Dei* character of created man, far from being a straight metaphysical honor, assumes an ambiguous, if not sinister, meaning. The motive for the archons' resolve is either simple envy and ambition, or the more calculating one of entrapping divine substance in their lower world by the lure of a seemingly congenial receptacle that will become its most secure bond. The imitation, presumptuous and blundering, is nevertheless effective. Although the mere creature of the archons—the body and a natural soul compounded from their several psychic powers—is not viable by itself, it becomes so through the injection of a spiritual element from beyond.

For this presence of transcendent spirit (pneuma) in psychophysical man—in itself a paradoxical, unnatural fact and the fulcrum of the whole soteriological drama—Gnostic speculation offers various explanations, their chief difference being whether the presence marks a success of the nether powers or a stratagem of the upper ones. In the first alternative, the causality operative on the divine side admits in turn the several variations of being a victim of violence (Mani), of deception, or of its own downward inclination (Poimandres). In the other alternative (the Valentinians), the divine seed is secretly deposited in the creature of the unknowing Demiurge in order to turn his work into an unintended vehicle of salvation. However, this variant is no more optimistic than the first, since the soteriological stratagem merely makes the best of a basic evil, of these divine portions' having become divorced from their source in the first place. In any case, the pneuma's innerworldly existence is a state of exile, the result of primeval divine tragedy; and its immersion in soul and body is the terminal form of that exile. For the archons, on the other hand, the incorporation of this transcendent element into their system is a condition of the system's existence, and its retention therefore becomes to them a matter of survival—their work's and their own. Hence, they must resist at all cost the spirit's extrication from the cosmic involvement, which the upper powers seek for the regaining of divine wholeness. The means of this extrication is knowledge.

HISTORY OF MAN. The process of conveying the saving knowledge to the world-imprisoned hostage of Light begins with Adam himself and runs through the history of humankind in a constant counterplay with the archontic powers. Human history is thus eschatological from the beginning. In the light of this scheme, the Scriptural account of early man, especially the Paradise story, is boldly recast, with all value signs reversed. The most significant of these reversals concerns the serpent, which, as the first bringer of knowledge in defiance of the Creator's mandate of ignorance, becomes the general symbol of the acosmic spiritual principle that works for the awakening of its captive kin in the world. The revelatory line thus started, and continued through the generations, ends in Christ (or may go beyond him to further revelations of the truth). Hence the cult of the serpent in a major group of Gnostic sects, the Ophites (from the Greek *ophis*, serpent). In the same spirit of reversal, Cain, Esau, and other rejected figures of the Old Testament became to certain sects (Cainites, Carpocratians, Perates) bearers of the pneumatic heritage, forming a secret lineage of gnosis and persecuted by the world god for this reason; their opposites, such as Abel and Jacob, his favorites, represent the unenlightened majority. Independently of the intention to scandalize that is evident here, the Gnostic scheme called for a prophetology in succession of the Adamitic revelation, for which Iranian tradition offered the idea of an eternal Messenger who moves through history in ever new incarnations. These messengers were variously identified with names from the religious past; in the final consolidation by Mani we find them reduced to four: Buddha, Zoroaster, Jesus, Mani. The significant omission of Moses from this list requires a comment on the anti-Judaism among the Gnostics.

The this-worldly spirit of the Hebrew religion combined with historical circumstance to make the Old Testament a prominent target of Gnostic dislike, to varying degrees. The extreme of hostility, even contempt, is found in Marcion, for whom this admittedly authentic revelation of the Creator and Lord of this world shares all the blemishes of its source: It is as opposed to the gospel of salvation as its divine author is to the God that saves and as this world, his work, is to the nonmundane realm beyond. Simon Magus and others are hardly less intransigent. A more qualified view is taken by the Valentinians: The law is at least partly prefigurative of the higher truth, and the prophets, although mainly inspired by the Demiurge, are sometimes (and unbeknown to him) used by his mother, Sophia, for her own messages, which thus are interspersed in the inferior bulk. There are other shades of opinion, but rejection of the whole body of Hebrew Scripture, joined with irreverent exegetical use, is by far the rule; and on this issue, and on the related one of the identity or nonidentity of the God of Moses with the Father of Jesus Christ, the main battle was fought between the church and the heretics.

COSMOS AND HUMAN NATURE. The material universe, the domain of the archons, is like a vast prison whose innermost dungeon is the earth, the scene of man's life. Around and above it, the cosmic spheres are ranged like concentric enclosing shells. Their number is usually seven, with a surrounding eighth that does not belong to the archontic realm proper but is intermediate between the cosmos and the upper world of the pleroma. There was, however, a tendency to multiply structures and to make the scheme more and more extensive: Basilides counted no fewer than 365 heavens. The religious significance of this cosmic architecture lies in the idea that everything that intervenes between here and the beyond serves to separate man from God, not merely by spatial distance but through active demonic force. Thus, the vastness and multiplicity of the cosmic system express the degree to which man is removed from God.

The spheres are the seats of the archons, whose ruling set of seven are the planetary gods of the Babylonian pantheon, now significantly renamed with synonyms for the Hebrew God—another sign of the latter's degradation. The archons collectively rule the world that they (or their overlord) made, and each individually in his sphere is a warder of the cosmic prison. Their tyrannical world rule, called Fate (*heimarmene*), is physically the law of nature, morally the law of justice, as exemplified in the Mosaic law, which issued from the Demiurge or the angels and, with its threat of retribution, aims at the enslavement of man as much as the first does with its force of necessity. As guardian of his sphere, each archon bars the passage to the souls that seek to ascend after death, in order to prevent their escape from the world and their return to God.

Man, the main object of these vast dispositions, is composed of flesh, soul, and spirit. Reduced to ultimate principles, his origin is twofold: mundane and extramundane. Both the body and the soul are products of the cosmic powers, who shaped the body in the image of the divine Primal Man and animated it with their own psychical forces: These are the appetites and passions of natural man, each stemming from and corresponding to one of the cosmic spheres, and all together making up the astral soul of man, his psyche. Through his body and his soul man is a part of the world and is subjected to

heimarmene. Enclosed in the soul is the spirit, or pneuma (also called the spark), a portion of the divine substance from beyond that has fallen into the world; the archons created man for the express purpose of keeping it captive here.

Thus, as in the macrocosm man is enclosed by the seven spheres, so in the human microcosm the pneuma is enclosed by the seven soul vestments originating from them. These psychical envelopments are considered impairments and fetters of the transmundane spirit, and its incarnation in the outer, material body merely completes the complex imprisonment. The resulting human constitution is, then, comparable to an onion with so many layers, on the model of the cosmos itself but with the order reversed; what is outermost and uppermost in the cosmos is innermost in man, and the innermost or nethermost stratum of the cosmic order, Earth, is the outer bodily garment of man. Only the innermost or pneumatic man is the true man, and he is not of this world, as his original in the total order, the deity, is external to the cosmos as a whole. In its unredeemed state the spirit, so far from its source and immersed in soul and flesh, is unconscious of itself, benumbed, asleep, or intoxicated by the poison of the world—in brief, it is ignorant. Its awakening and liberation are effected through knowledge.

ESCHATOLOGY: SALVATION THROUGH GNOSIS. The nature of Gnostic dualism determines the general concept of salvation, and the stratifications of cosmos and man condition its details. Its basic premise is that the transcendent God is as alien to this world as the pneumatic self is in the midst of it. The goal of Gnostic striving is the release of the inner man from the bonds of the world and his return to his native realm of light. The necessary condition for this is that he know about the transmundane God and about himself, that is, about his divine origin as well as his present situation, and hence, also about the nature of the world that determines his situation. Such knowledge is withheld from him by precisely the selfsame situation that requires it, for ignorance is the essence of mundane existence, just as it was the principle of the world's coming into being. In particular, the transcendent God is unknown in the world and cannot be discovered from it; therefore, revelation is needed. The necessity for revelation is inherent in the innercosmic condition; and its occurrence alters this condition in its decisive respect, that of ignorance.

Revelation, or the "call," is thus already a part of salvation. Its bringer is a messenger from the world of Light who penetrates the barriers of the spheres, outwits the archons, awakens the spirit from its earthly slumber, and imparts to it the saving knowledge from without. The mission of this transcendent savior begins even before the creation of the world, since the fall of the divine element preceded creation, and the archetypal redemption indeed takes place in the precosmic stage. It is the incompleteness of this initial restoration, whether of Sophia or of Mani's Primal Man, that leads to the genesis of the world and the protraction of the saving process throughout its history. The fact that in the discharge of his task the eternal messenger must himself assume the lot of incarnation and cosmic exile, and the further fact that, at least in the Iranian variety of the myth, he is in a sense identical with those he calls—the once lost parts of his divine self—give rise to the moving idea of the "saved savior" (*salvator salvandus*).

The knowledge revealed by the messengers, for short "knowledge of God," comprises the whole content of the Gnostic myth, with everything it has to teach about God, man, and world, including the history of the beginnings which alone offers the key to the secrets of existence; that is, the revelation contains the elements of a theoretical system. On the practical side, however, it is more particularly "knowledge of the way"—of the soul's way out of the world—comprising the sacramental and magical preparations for its future ascent and the secret names and formulas that force the passage through each sphere. Equipped with this gnosis, the soul after death travels upward, leaving behind at each sphere the psychical vestments contributed by that sphere; thus the spirit, stripped of all foreign accretions, reaches the God beyond the world and reunites with the divine substance. (The most circumstantial description of this ascent is found in the "Poimandres," the first treatise of the Hermetic corpus.) On the scale of the total divine drama, the individual ascent is part of the restoration of the deity's own wholeness, impaired by the events of the beginning. Only through the loss suffered then did the deity become involved in the destiny of the world, and only to retrieve his own does he intervene, through his envoys, in cosmic history. With the completion of this ingathering, the cosmos, deprived of its elements of light, will come to an end.

MORALITY. In this life the pneumatics, as the possessors of gnosis called themselves, are set apart from the mass of humankind. The immediate illumination that makes the individual sovereign in the sphere of knowledge (hence the great variety of Gnostic doctrines) also inspires superior rules of conduct. Generally, the pneumatic morality

is determined by hostility toward the world and contempt for all mundane ties. From this principle, however, two contrary conclusions could be drawn, and both found their extreme partisans: the ascetic and the libertine. The ascetic deduces from the possession of gnosis the obligation to avoid further contamination by the world and therefore to reduce the world's use to a minimum; the libertine derives from the same possession the privilege of unrestrained freedom. The libertine conclusion, more startling and more devious, is argued thus: The law, since it represents the will of the Demiurge and is one form of his tyranny, does not obligate the pneuma, which is "saved in its nature" and can be neither sullied by actions (which in themselves are morally neutral) nor frightened by the threat of archontic retribution which can affect only the body and the psyche).

Thus the pneumatic, since he is free from the power of fate, is also free from the yoke of the moral law, and all things are permitted to him. This freedom, however, is more than merely permissive; its practice is bidden by metaphysical interest. Through intentional violation of the demiurgical norm (for which the mythological vilification of the Demiurge prepares) the pneumatic thwarts the design of the archons and thus paradoxically contributes to the work of salvation. From the motive of defiance it is then only one step further to the teaching of the Cainites and Carpocratians that there is a positive duty to perform every kind of action, to leave no deed undone, no possibility of freedom unrealized, in order to render nature its due and exhaust its powers; only in this way can final release from the cycle of reincarnations be obtained. Gnostic libertinism thus spans the whole scale from mere negative license to positive Faustian obligation—at which point it loses again some of the contrast to its ascetic alternative.

The latter alternative, too, betrays the common root in Gnosticism from which both opposites spring. Although more obvious in the libertine choice, the element of defiance shows in the ascetic one as well; as much as it may serve purification or other perfectionist ends normally associated with asceticism, it often has the declared purpose of obstructing the cause of the Creator, even just to spite him, by refusing to use his works (a kind of metaphysical strike). This obstructive aspect is especially clear in the abstention from sexual intercourse and marriage when, as in Marcion and Mani, its purpose is not to help replenish the world of the Demiurge and further disperse in it the captive light—thereby prolonging its exile and making its ingathering more difficult. Indeed, according to Mani, the reproductive scheme was instituted by the archons with precisely this end in view. Asceticism is thus a matter less of ethics than of metaphysical alignment, and its common ground with libertinism is the determination not to play the Creator's game. The one repudiates allegiance to nature through abstention; the other, through excess. Both are lives outside the mundane norm. Freedom by use and freedom by nonuse are thus alternative expressions of the same acosmism.

ACOSMISM. Acosmism, the real basis of the Gnostic position, contains the seeds of nihilism; the very extremism of divine transcendence has nihilistic implications. As the totally other, alien, and unknown, the Gnostic God has more of the *nihil* than of the *ens* in his concept. For all purposes of man's relation to the reality that surrounds him, this hidden God is a negative term; no law emanates from him—none for nature, and thus none for human action as a part of the natural order. His only relation to the world is the negative one of saving from the world. Antinomianism follows naturally, even if not inevitably, from these premises.

TWO TYPES OF GNOSTIC DUALISM. This entry has kept mainly to the Syrian-Egyptian stream of Gnostic speculation, to which the majority of systems, especially the Christian ones, belong. There is, however, another, Iranian line of speculation that culminates in Mani.

Both types, being Gnostic, were evolved to explain the same facts of a dislocated metaphysical situation—both are dualistic concerning their common theme: the existing rift between God and world, world and man, spirit and flesh. The Iranian type, in a Gnostic adaptation of Zoroastrian doctrine, starts from a dualism of two opposed principles and then must explain how the original Darkness came to engulf elements of the Light—it describes the world drama as a war with changing fortunes; and the divine fate, of which man's fate is a part and the world an unwilled by-product, is explained in terms of mixing and unmixing, captivity and liberation. Here the knightly male figure of First Man, the warrior, assumes the role of the exposed and suffering part of divinity.

The Syrian speculation, with the female Sophia in that role, undertakes the more ambitious task of deriving dualism itself, and the ensuing predicament of the divine in the system of creation, from the one and undivided source of being. It does this by means of a genealogy of personified divine states evolving from one another that describes the progressive darkening of the original Light

in categories of guilt, error, and failure. This devolution within the divine being ends in the decadence of complete self-alienation that is this world. Both dramas start with a disturbance in the heights; in both, the existence of the world marks a discomfiture of the divine and a necessary, in itself undesirable, means of its eventual restitution; in both, the salvation of man is that of the deity itself. The difference lies in whether the tragedy is forced upon the deity from without by the trespass of an independent Darkness, which thus has the first initiative (the deity itself being in perfect tranquility), or is motivated from within itself, with Darkness and Matter the products of its passion, which they hypostatize in external terms. To divine defeat and sacrifice in the first case correspond divine guilt and error in the second; to compassion for the victimized Light, spiritual contempt for demiurgical blindness; to eventual divine liberation, reformation through enlightenment.

The Manichaean and Valentinian systems respectively exemplify the two types. The Iranian type, with its high-minded story of battle, defeat, and recovery, lends itself to more concrete and gripping dramatization. However, only the subtler Syrian type, by according metaphysical status to knowledge and ignorance as modes of the divine life and therefore as universal, cosmogonic categories, can do full justice to the redemptional claim made on behalf of knowledge in all Gnostic religion. Valentinian speculation inferred that the human individual event of pneumatic knowledge reverses the precosmic universal event of divine ignorance and is in its redeeming effect of the same ontological order. Thus the actualization of knowledge in the person is at the same time an act in the general ground of being.

See also Mani and Manichaeism; Valentinus and Valentinianism.

Bibliography

PRIMARY SOURCES

New Testament Apocrypha are translated in M. R. James, ed., *The Apocryphal New Testament* (Oxford: Clarendon Press, 1953), and in E. Hennecke and W. Schneemelcher, eds., *Neutestamentliche Apokryphen* (Tübingen: Mohr, 1959–1964). Editions of the original texts are listed there. The following Apocrypha are noteworthy for their Gnostic contents: the "Acts of Thomas," containing the Hymn of the Pearl; the pseudo-Clementine "Homilies and Recognitions"; and the "Odes of Solomon." See also the section on Coptic sources, below.

Patristic sources, in standard editions and translations, include Irenaeus of Lyon, *Against the Heresies*; Clement of Alexandria, *Excerpts from Theodotus* and *Stromateis*; Tertullian of Carthage, *Against Marcion, Against the Valentinians, On the Plea of the Heretics*, and *On the Soul*; Hippolytus of Rome, *Philosophoumena*; Origen, *Against Celsus* and *Commentary on the Gospel of John* (containing the fragments of Heracleon); and Epiphanius of Salamis, *Panarion*. Collections of patristic texts (Greek and Latin) may be found in A. Hilgenfeld, *Die Ketzergeschichte des Urchristentums* (Leipzig, 1884), and W. Völker, ed., *Quellen zur Geschichte der christlichen Gnosis* (Tübingen: Mohr, 1932).

Coptic sources include C. Schmidt, *Koptisch-gnostische Schriften* (Leipzig: Hinrichs, 1905) and *Pistis Sophia*, Coptica II (Leipzig, 1925); W. Till, ed., *Die gnostischen Schriften des koptischen Papyrus Berolinensis 8502* (Berlin, 1955). The newly discovered Nag Hammadi texts, so far as published, include the following: M. Malinine, H. C. Puech, and G. Quispel, eds., *Evangelium Veritatis* (Zürich, 1956 and 1961), translated in K. Grobel, *The Gospel of Truth* (New York and Nashville, TN, 1960); H. M. Schenke, "Das Wesen der Archonten," in *Theologische Literaturzeitung* 83 (1958): 661–670; A. Guillaumont, H. C. Puech, and others, eds., *The Gospel according to Thomas* (Leiden, 1959), translated in R. M. Grant and D. N. Freedman, *The Secret Sayings of Jesus* (London: Collins, 1960), and also translated by J. Leipoldt, J. Doresse, T. Säve-Söderbergh, H. Quecke, and R. McL. Wilson; M. Krause and P. Labib, *Die drei Versionen des Apokryphon des Johannes* (Wiesbaden, 1962), also translated by S. Giverson; A. Böhlig and P. Labib, eds., *Die koptisch-gnostische Schrift ohne Titel aus Codex II von Nag Hammadi* (Berlin: Akademie-Verlag, 1962), and *Koptisch-gnostische Apokalypsen aus Codex V von Nag Hammadi* (Halle: Martin-Luther-Universität, 1963); M. Malinine and others, eds., *De Resurrectione Epistula ad Reginum* (Zürich and Stuttgart, 1963); W. Till, *Das Evangelium nach Philippos*, Texte und Studien II (Berlin: De Gruyter, 1963), also translated by R. McL. Wilson. See also H. C. Puech, "Gnostische Evangelien und verwandte Dokumente," in *Neutestamentliche Apokryphen*, Vol. I, edited by E. Hennecke and W. Schneemelcher (Tübingen; Mohr, 1959), pp. 158–271.

Hermetica are collected in A. D. Nock, ed., *Hermès Trismégiste*, translated by A. J. Festugière, 4 vols. (Paris, 1945–1954), which contains text, translation, and notes and supersedes all earlier editions.

Writings of the Mandaeans are in the following editions and/or translations: M. Lidzbarski, *Das Johannesbuch der Mandäer* (Giessen, 1905 and 1915), *Mandäische Liturgien* (Berlin: Weidmannsche, 1920), and *Ginza. Der Schatz oder das grosse Buch der Mandäer* (Göttingen: Vandenhoeck and Ruprecht, 1925); and E. S. Drower, *The Book of the Zodiac* (London: Royal Asiatic Society, 1949), *Diwan Abatur* (Vatican City, 1950, Studi e testi 151), *The Haran Gawaita* (Vatican City, 1953, Studi e testi 176), *The Canonical Prayerbook of the Manaeans* (Leiden, 1959), *The Thousand and Twelve Questions* (Berlin: Akademie-Verlag, 1960), *The Coronation of the Great Šišlam* (Leiden: Brill, 1962), and *A Pair of Nasoraean Commentaries* (Leiden: Brill, 1963).

For writings of the Manichaeans, see the bibliography to the Mani and Manichaeism entry.

A selection of materials from the first four sections of this bibliography can be found in R. M. Grant, *Gnosticism: A Source Book of Heretical Writings* (New York: Harper, 1961).

SECONDARY SOURCES

General studies include F. C. Baur, *Die christliche Gnosis* (Tübingen, 1835); C. W. King, *The Gnostics and Their Remains*, 2nd ed. (London: Nutt, 1887); W. Bousset, *Hauptprobleme der Gnosis* (Göttingen, 1907); A. Harnack, *Lehrbuch der Dogmengeschichte*, 4th ed. (Tübingen, 1909), Vol. I; H. Leisegang, *Die Gnosis* (Leipzig, 1924; 4th ed., Stuttgart, 1955); E. de Faye, *Gnostiques et gnosticisme* (Paris, 1925); F. C. Burkitt, *Church and Gnosis* (Cambridge, U.K.: Cambridge University Press, 1932); H. Jonas, *Gnosis und spätantiker Geist*, 2 vols. (Göttingen, 1934 and 1954; 3rd ed., 1964), and *The Gnostic Religion* (Boston: Beacon Press, 1958; 2nd ed., 1963); J. Dupont, *Gnosis* (Louvain, 1949); G. Quispel, *Gnosis als Weltreligion* (Zürich, 1951); R. McL. Wilson, *The Gnostic Problem* (London: Mowbray, 1958); and R. M. Grant, *Gnosticism and Early Christianity* (New York: Columbia University Press, 1959).

Special studies include many important articles in periodicals, but for reasons of space, only books are listed here.

On the Valentinians, see the bibliography to the *Valentinus and Valentinianism* entry.

On Nag Hammadi, see F. L. Cross, ed., *The Jung Codex* (London: Mowbray, 1955); H. M. Schenke, *Die Herkunft des sogenannten Evangelium Veritatis* (Göttingen, 1959); J. Doresse, *The Secret Books of the Egyptian Gnostics* (New York: Viking Press, 1960); and W. C. van Unnik, ed., *Evangelien aus dem Nilsand* (Frankfurt, 1960).

On Hermetism, see R. Reitzenstein, *Poimandres* (Leipzig, 1904); J. Kroll, *Die Lehren des Hermes Trismegistos* (Münster, 1914); W. Scott and A. S. Ferguson, eds., *Hermetica*, 4 vols. (Oxford: Clarendon Press, 1924–1936); W. Gundel, *Neue astrologische Texte des Hermes Trismegistos* (Munich, 1936); A. J. Festugière, *La révélation d'Hermès Trismégiste*, 4 vols. (Paris: Lecoffre, 1944–1954); and G. van Moorsel, *The Mysteries of Hermes Trismegistos* (Utrecht, 1955).

On the Mandaeans, see W. Brandt, *Die Mandäer* (Amsterdam, 1915); R. Reitzenstein, *Das mandäische Buch des Herrn der Grösse* (Heidelberg, 1919); H. Odeberg, *Die Mandäische Religions-anschauung* (Uppsala, 1930); E. S. Drower, *The Mandaeans of Iraq and Iran* (Oxford: Clarendon Press, 1937), *Water into Wine* (London: John Murray, 1956), and *The Secret Adam* (Oxford: Clarendon Press, 1960); and K. Rudolph, *Die Mandäer*, 2 vols. (Göttingen, 1960–1961).

On other special aspects of Gnosticism, see R. Reitzenstein, *Das iranische Erlösungsmysterium* (Bonn, 1921), and *Die hellenistischen Mysterienreligionen*, 3rd ed. (Leipzig, 1927); R. Reitzenstein and H. Schaeder, *Studien zum antiken Synkretismus* (Leipzig, 1926); A. Harnack, *Marcion*, 2nd ed. (Leipzig, 1924); C. H. Kraeling, *Anthropos and Son of Man* (New York, 1927); W. Bauer, *Rechtgläubigkeit und Ketzerei im ältesten Christentum* (Tübingen, 1934); S. Pétrement, *Le dualisme chez Platon, les gnostiques et les manichéens* (Paris: Presses Universitaires de France, 1947); H. J. Schoeps, *Urgemeinde, Juden-christentum, Gnosis* (Tübingen, 1956); G. Scholem, *Jewish Gnosticism* (New York: Jewish Theological Seminary of America, 1960); A. Wlosok, *Laktanz und die philosophische Gnosis* (Heidelberg: Winter, 1960); J. Jervell, *Imago Dei* (Göttingen: Vandenhoeck and Ruprecht, 1961); and H. M. Schenke, *Der Gott "Mensch" in der Gnosis* (Göttingen: Vandenhoeck and Ruprecht, 1962).

OTHER RECOMMENDED WORKS

Sources

Grant, R. M. *Gnosticism: A Sourcebook of Heretical Writings from the Early Christian Period*. New York: 1961.

Horner, G. *Pistis Sophia*. London: Society for Promoting Christian Knowledge, 1924.

Layton, B. *The Gnostic Scriptures*. New York: Doubleday, 1987.

Schmidt, C., and W. Till. *Koptisch-gnostische Schriften* (Die griechischen christlichen Schriftsteller der ersten drei Jahrhunderte 45). 3rd ed. Berlin: Akademie-Verlag, 1959.

Till, W. *Die gnostischen Schriften des koptischen Papyrus Berolinensis 8502* (Texte und Untersuchungen zur Geschichte der altchristlichen Literatur 60). Berlin: Akademie-Verl,1955.

Völker, W. *Quellen zur Geschichte der christlichen Gnosis*. Tübingen: Mohr, 1932.

Studies

Baur, F. C. *Die christliche Gnosis oder die christliche Religionsphilosophie in ihrer geschichtlichen Entiwicklung*. Tübingen: 1835.

Colpe, et al. *Die Religion in Geschichte und Gegenwart*. 3rd ed. Tübingen: 1957–1963; 2.1648–1661.

Cornélis, H., and A. Leonard. *La gnose éternelle* (Je sais, je crois 146). Paris: 1959.

Doresse, J. *The Secret Books of the Egyption Gnostics*. New York: Viking Press, 1960.

Eltester, W., ed. *Christentum und Gnosis*. Beihefte zur *Zeitschrift für dir neutestamentliche Wissenschaft und die Kunde des Urchristentums*, 37. Giessen: 1969.

Grant, R. M. *Gnosticism and Early Christianity*. 2nd ed. New York: Columbia University Press, 1966.

Jonas, H. *The Gnostic Religion*. 2nd ed. Boston; 1963.

Logan, A. H. B. *Gnostic Truth and Christian Heresy: A Study in the History of Gnosticism*. Edinburgh; Peabody, MA: Hendrickson Publishers, 1996.

Peterson, E. *Frühkirche, Judentum und Gnosis*. Rome: 1959.

Prümm, K., et al. *Lexicon für Theologie und Kirche*. Freiburg: 1957–1966; 4.1021–1031.

Quasten, J. *Patrology*. Westminster: Newman Press, 1950; 1.254–277.

Quispel, G. *Gnosis als Weltreligion*. Zurich: Origo Verlag, 1951.

Sagnard, F. L. *La gnose valentinienne et le témoignage de saint Irénée*. Paris: J. Vrin, 1947.

Scholem, G. G. *Jewish Gnosticism, Merkabah Mysticism, and Talmudic Tradition*. New York: Jewish Theological Seminary of America, 1960.

Van Unnik, W. C. *Newly Discovered Gnostic Writings* (Studies in Biblical Theology 30). Naperville, IL: A. R. Allenson, 1960.

Wilson, R. McL. *Gnosis and the New Testament*. Oxford: Blackwell, 1968.

Wilson, R. McL. *The Gnostic Problem*. London: A. R. Mowbray, 1958.

Wolfson, H. A. *The Philosophy of the Church Fathers*. Cambridge, MA: Harvard University Press, 1956.

Yamauchi, E. M. *Pre-Christian Gnosticism: A Survey of the Proposed Evidence*. 2nd ed. Grand Rapids, MI: Baker Book House, 1983.

Hans Jonas (1967)
Bibliography updated by Scott Carson (2005)

GOBINEAU, COMTE JOSEPH ARTHUR DE

(1816–1882)

Comte Joseph Arthur de Gobineau was a French philosopher, historian, novelist, and diplomat. Gobineau's diplomatic duties during the Second Empire carried him to Switzerland, Persia, Greece, and Brazil, where he produced a number of historical and ethnographic works of considerable merit. He is best known for his *Essai sur l'inégalité des races humaines* (4 vols., Paris, 1853–1855; Vol. 1 translated into English by Adrian Collins as *The Inequality of Human Races,* London, 1915). This work is usually considered an important contribution to nineteenth-century racist thought; but Gobineau's racism was a by-product of his attempt to account for the decline of the European aristocracy in terms of the more general problem of the decline and fall of civilizations.

Gobineau presented his work as an essay in positivistic social theory; in the preface to the second edition (1884), he argued that Henry Thomas Buckle and Charles Darwin had merely proceeded along lines originally marked out by himself. Superficially, then, Gobineau's work resembled those positivistic theories of culture in which his century abounded. However, it differed from them in its categorical rejection of the doctrine of "progress." His work was profoundly pessimistic, and in the end Gobineau predicted the ineluctable decay, not only of Western civilization, but of the whole of humanity. Thus, Gobineau's racism differed from the later racism of the imperialist period. He was neither a nationalist nor a proponent of the idea of "the white man's burden." He was, rather, an apologist for a class that had come to feel that since it no longer had a genuine social role to fill, society itself was no longer possible.

Gobineau held that the human species was originally divided into three races as a result of environmental conditioning. The Negro race is dominated by "desire" and the need to gratify desire, and hence is the natural enemy of civilization. Driven by the need for sensual gratification alone, the Negro lacks both speculative and technical capability. The yellow race is the antitype of the Negroid, lacking in physical vigor but possessing a natural talent for technical accomplishment that allows it to create pseudo civilizations but prohibits it from developing any genuine science. The white race is superior to the other two because it combines energy and intelligence in just the right proportions. The white race has a genuine "love of life," but it is able to control and direct that love to culturally creative ends. The white man is a speculative thinker, which allows him to create both a science of nature and a science of politics. This makes of the white man the natural conqueror of the other two races. Whatever these other races have accomplished in the way of civilizational growth they owe to the superaddition of white blood, Gobineau held.

By the same token, however, racial intermixture results in the debilitation of the white race. Unlike the Darwinists, who saw survival itself as evidence of fitness, Gobineau held that in every racial mixture it is the weaker strain that predominates. In the long run, then, racial intermixture must result not so much in the elevation of inferior breeds as the mongrelization of the entire species. Thus envisaged, the white race is caught on the horns of a dilemma. Its inherent excellence drives it forth to world conquest, but that very conquest leads to its decline. Gobineau's theory of civilization, in short, was not so much an attempt to explain the facts of history as to justify his own overriding sense of *senescens saeculum,* a product of the breakdown of the social class to which he belonged.

Obviously, such a theory could not serve as a rationalization for imperialistic expansion, for if Gobineau were right, it would be better for the white race to cease expanding and seal itself off from contact with all other races. But Gobineau was valuable to the reactionary groups of his time even though he opposed imperialism, for he was an ardent enemy of liberalism. His theory explicitly designated the social egalitarianism of the radicals as an instrument of further mongrelization. The city of Paris, where the races mixed in perfect equality, proved his point, he wrote, because there "tradition is respected not at all."

As for nationalism, Gobineau regarded this phenomenon as another evidence of the breakdown of racial solidarity. He dreamed of an international aristocracy of blood to which the purest elements of all nations belong. His book *La renaissance* (Paris, 1877; translated into English by P. V. Cohen as *The Renaissance,* New York, 1913) was intended to demonstrate that as long as the white race had retained its internationalist sense of caste and eschewed expansion and intermixture, it had remained creative and productive. Neither libertarian nor expansionist, the aristocracy of the Renaissance, as represented by such figures as Cesare Borgia, Michelangelo, and Raphael, was able to produce masterpieces of art and politics. The problem of race did not intrude itself into Gobineau's handling of the Renaissance, because in *La renaissance* he was dealing with a preliberal, amoral, and creative example of the white race's power. But this book

does contain an implicit criticism of his own age, dominated, in his opinion, by weak-willed liberals and traditionless mongrels.

See also Buckle, Henry Thomas; Darwin, Charles Robert; Racism.

Bibliography

ADDITIONAL WORKS BY GOBINEAU

Trois ans en Asie. Paris: Hachette, 1859.

Les religions et les philosophies dans l'Asie centrale. Paris: Didier, 1865.

Histoire des Perses, 2 vols. Paris: H. Plon, 1869.

Correspondance d'Alexis de Tocqueville et d'Arthur de Gobineau. Paris, 1959. The correspondence is also available in an English translation by John Lukacs in Alexis de Tocqueville, *The European Revolution & Correspondence with Gobineau* (Garden City, NY: Doubleday, 1959), edited by Lukacs.

WORKS ON GOBINEAU

Combris, Andrée. *La philosophie des races de Gobineau.* Paris: Alcan, 1937.

Falk, Reinhold. *Die weltanschauliche Problematik bei Gobineau.* Berlin: Norm-druck, 1936.

Ferguson, W. J. *Renaissance in Historical Thought.* New York: Houghton Mifflin, 1948. On Gobineau as historian.

La nouvelle revue Française 42 (February 1934). Issue dedicated to Gobineau and his work.

Schemann, Ludwig. *Gobineau und die deutsche Kultur.* Leipzig, 1910; 7th ed. Leipzig: Teubner, 1934.

Sellière, Ernst. *Le Comte de Gobineau et l'aryanisme historique.* Paris: Plon-Nourrit, 1903.

Streidl, Rudolf. *Gobineau in der französichen Kritik.* Würzburg: R. Mayr, 1935.

Hayden V. White (1967)

GOD, ARGUMENTS FOR THE EXISTENCE OF

See *Common Consent Arguments for the Existence of God*; *Cosmological Argument for the Existence of God*; *Degrees of Perfection, Argument for the Existence of God*; *Moral Arguments for the Existence of God*; *Ontological Argument for the Existence of God*; *Religious Experience, Argument for the Existence of God*; and *Teleological Argument for the Existence of God*

GOD, CONCEPTS OF

It is very difficult—perhaps impossible—to give a definition of "God" that will cover all usages of the word and of equivalent words in other languages. Even to define God generally as "a superhuman or supernatural being that controls the world" is inadequate. "Superhuman" is contradicted by the worship of divinized Roman emperors, "supernatural" by Benedict Spinoza's equation of God with Nature, and "control" by the Epicurean denial that the gods influence the lives of men. Therefore, while the above definition satisfies a wide range of usages, it is not universally applicable.

This entry will deal with five problems: the transcendence and immanence of God, his relation to the world, his chief attributes, the extent to which he is "personal," and the ways by which he can be known. In discussing these problems it will be necessary to consult the data provided by both religion and philosophy. But purely religious data (in contrast with theological speculations based on them) will be mentioned only when they are relevant to philosophical understanding.

TRANSCENDENCE AND IMMANENCE

In Judaism and Christianity, God is unquestionably transcendent. He is "wholly other" than the world he made. In Judaism his transcendence was emphasized by, among other things, the prohibition of idols, the explicit teaching of Isaiah 40:12–26, the sacredness of the Tetragrammaton, and the speculations of Philo who, in a typical passage, speaks of God Platonically as "the pure and unsullied Mind of the universe, transcending virtue, transcending knowledge, transcending the good itself." The New Testament, in confirming the Old Testament, repudiates the claims of Hellenistic *gnosis* by affirming that "no one has ever seen God" (John 1:18) and that all our knowledge of him is like a confused reflection in a mirror (1 Cor. 13:12). Among later Christian thinkers this biblical attitude was reinforced partly by the influence of Neoplatonism and partly by the experience of the mystics (especially Dionysius the Pseudo-Areopagite). Hence, in the *Summa Contra Gentiles* (I, 14), Thomas Aquinas says that "the divine substance exceeds by its immensity every form which our intellect attains," so that while we can know *that* God is (*quod sit*) we cannot know his essence or *what* he is (*quid sit*). In recent times divine transcendence has been stressed by Søren Kierkegaard and Karl Barth, as opposed to Hegelian attempts to obtain a rational and synoptic understanding of ultimate reality. From a phenomenological point of view, Rudolf Otto, in his *Das Heilige* (Marburg, 1917), defined the object of worship as a *mysterium tremendum et fascinans* that is revealed to a suprarational faculty of the soul.

Christian theologians claim that this transcendent God can be spoken of either negatively by the *via negativa* or *via remotionis* (the apophatic way) or positively (by the cataphatic way). According to the negative way, we deny qualities to God by the use of such adjectives as "*in*corporeal" and "*un*created." Thus we come to know him by knowing what he is not. But we also speak positively of God (for example, by predicating goodness or wisdom of him). Thomas denied that positive predicates are definable in terms of negative ones. He also denied that they simply point to God as an indeterminate cause of finite properties. In his view, they refer to God in a positive manner through an "analogy of proportionality." Thus goodness exists in God in a "supereminent" form, proportionate to his infinite mode of being. Through this theory of analogical predication, Thomas hoped to steer a middle course between the anthropomorphism of univocal predication, on the one hand, and the agnosticism of equivocal predication on the other.

According to the main tradition of Christian thought, God is also immanent. Augustine held that the light of God's presence in the human mind enables it to recognize eternal truth. Thomas, while rejecting the Augustinian theory of illumination, affirmed God's omnipresence unambiguously. "God is in all things, not, indeed, as part of their essence, or as a quality, but in the manner that an efficient cause is present to that on which it acts. Hence God is in all things, and intimately" (*Summa Theologiae* Ia, 8, 1). Similarly, the mystics affirm that the transcendent God is present (even when unrecognized) at the "ground" or "apex" of the soul. But some philosophers have *identified* God's substance either partly or wholly with the world. The clearest exponent of this concept in Western thought is Spinoza, whose identification of God with Nature a paradigm of pantheism. Such later philosophers as Edward Caird and Sir Henry Jones, who equated the Christian God with the Hegelian Absolute, approximated pantheism in varying degrees. Many modern theologians, such as Barth and Rudolf Bultmann, who have followed Kierkegaard in reaffirming God's transcendence, have either denied or ignored his immanence. Paul Tillich is a notable exception. While he spoke of God "existentially" as the transcendent Object of our "ultimate concern," he also held that we could not know God without "participating" in him.

GOD AND THE WORLD

The degree to which God is transcendent or immanent depends on the view that is taken of his relation to the world. At least five views are possible.

GOD AS FINAL CAUSE. God can be viewed as a final, though not efficient, cause of the world. This view was held by Aristotle. According to him, God is the world's "prime mover." God "moves" the world in the sense that he educes form from its material structure by inspiring it, through a series of subordinate movers or "intelligences," to love him as its end or goal. Yet Aristotle expressly denied a creation of the world; he considered matter to be ungenerated and eternal.

WORLD AS EMANATION FROM GOD. The world may be regarded as in some way an emanation from, or self-expression of, God. This view has taken three main forms.

According to Plotinus, the One, or "first god," is beyond all thought and being. The One's simplicity would be violated if the world were a part of it. Its unchangeability would be violated if it were to create the world by an act of will. Therefore Plotinus propounded his theory of "emanation." Mind, Soul, and the material world flow from the One (as rays flow from the sun) without impairing its self-sufficiency.

According to Spinoza, the world *is* God (the only substance) under his attributes of thought and extension. Everything follows from his essence by a logical necessity. "Things could not have been produced by God in any other manner or order than that in which they were produced. All things must have followed of necessity from a given nature of God, and they were determined for existence or action in a certain way by the necessity of the divine nature" (*Ethics* I, prop. 33). Critics of Spinoza have continually pointed out that on these premises it is very hard to account for, first, the individuality which human persons *seem* to have; second, their apparent freedom, which Spinoza elsewhere attempts to analyze; and third, the fact of evil, especially in its moral forms.

The same type of relation between God and the world was posited by G. W. F. Hegel. Unlike Plotinus, he regarded God or the Absolute as in its essence a self-diversifying unity. Unlike Spinoza, he conceived of God's self-expression as a dynamic process that is discoverable in historical events. Hegel's thought is not free from ambiguity. He sometimes speaks of God as an independently existing entity. But his final and distinctive view is that the Absolute Spirit does not exist apart from the human spirits in which it is progressively evolved.

WORLD AS PREEXISTENT MATTER SET IN ORDER. The third way of relating God to the world was stated by Plato in his *Timaeus*. According to this dialogue (29E–30), God is bounded on the one hand by the world of Forms

and on the other by preexistent matter. His task is to impose the Forms on matter, and so construct a rationally ordered whole. Being wholly good, and therefore free from jealousy, he wished everything to be like himself. Since an intelligent being is superior to an unintelligent one, and since intelligence cannot be present in anything that is devoid of soul, "he put intelligence in soul, and soul in body, that he might be the creator of a work which was by nature best." (In the *Republic* 597, Plato implies that God creates the Forms, but this was not his usual view.)

CREATION EX NIHILO. In contrast with all the previous views, Christian theists since Augustine have held the doctrine of creation out of nothing. This phrase is meant to exclude both the idea that the world is a necessary emanation from God's nature and the idea that matter preexists his creative act. God brings the whole world into being by an undetermined choice. He does not need the world to complete his nature, for he is wholly self-sufficient. He is not confronted with an alien Necessity, for he is the efficient cause of all that is.

This conception of the relation between the Creator and the creature can be elucidated through the contrast between necessary and contingent being. God exists necessarily. In him essence and existence are identical. He is self-existent in a unique and incomprehensible way. Creatures, on the other hand, are contingent. Their essence, while preexisting ideally in the mind of God, would not have achieved independent being if he had not chosen to grant it by a free act of love. Therefore, while they participate in him both by nature and by grace, they never lose their created status. They can be deified (as the Greek fathers taught) within their finite limits, but they cannot become divine in the sense of sharing God's aseity.

The full Christian doctrine does not restrict God's creative act to an initial moment in the cosmic process. All things owe their being continuously to his power. He is a first cause in the order of existence, not of time, for he himself is supratemporal. Hence it is irrelevant to theology whether the world did or did not have a temporal beginning. Thomas held that while such a beginning was revealed through Scripture, it could not be rationally proved. All reason knows is that God is the eternal, ever-present, and creative source of anything that does (or can) exist. Creation and preservation are identical.

However, while no creature exists from itself (*a se*), every creature exists by itself (*per se*) or in itself (*in se*). Created substances have a relative independence, or derived autonomy. These paradoxical expressions are required in order to affirm the truth that while creatures owe their being to God as their first cause, they also act according to secondary causes that are appropriate to their natures. The distinction between these two types of cause is necessary for a true assessment of the relation between science and theology. Because finite things exist per se, their secondary causes are discoverable without the aid of faith. But the discovery of secondary causes does not, without a further, nonscientific, act of inference or intuition, either permit or prohibit belief in a first cause, God.

Yet God, as first cause, can suspend or transform secondary causes in order to perform his will. When he does so, his action is called a miracle. A miracle does not violate nature. It is a case of nature behaving in an abnormal way through a special act of the same creative power that is at work in the normal processes which can be subsumed under scientific laws. If the essence of finite being is to be dependent on God's will (and so to possess a *potentia obedientialis* in relation to it), miraculous acts are not less natural than nonmiraculous ones. But while the abnormal character of an event is empirically verifiable, its miraculous character as an act of God can be discerned by faith alone. (Many theologians readily admit that at least some of David Hume's skeptical objections have considerable prima-facie force.)

The relation between divine causality and the human will has been extensively discussed by theologians. The doctrine of predestination, in its rigid Augustinian form, would seem to be obviously incompatible with human freedom. Yet even those theologians who reject the doctrine are obliged to face the problem of the manner in which God acts on men both by nature (through his general providence) and by grace (through the supernatural gift of the Holy Spirit). While various attempts have been made to separate divine and human action so that, for example, the human will is left wholly autonomous in a strictly moral choice, many theologians (more recently, D. M. Baillie and A. M. Farrer) affirm, on grounds of Scripture and experience, that the divine and human wills act simultaneously throughout the Christian life, but that the manner of their interaction is a paradox, or mystery, that cannot be unraveled by the intellect.

GOD AS FINAL STAGE OF COSMIC PROCESS. Samuel Alexander held the eccentric view that God qua deity, so far from being the ground of the cosmic process, is (ideally) its final stage. The world evolves from space time through matter and life to mind. God exists wholly within the world, which is his "body," but he does not yet

exist as deity (that is, as an infinite, transcendent, Being). Moreover, he will never so exist. Deity, as a state of infinite perfection, is a goal to which the world (or God considered as the world) continually strives but which is unattainable.

Some philosophers have combined two or more of these views. Thus A. N. Whitehead, while rejecting the idea that God is the world's efficient cause, held, as did Aristotle, that he is a final cause who (like Plato's God) brings order into the world by ensuring the ingredience of eternal objects (which, however, do not exist independently) in the realm of temporal flux. But Whitehead also shows his affinity with Alexander by asserting that it is as true to say that the world creates God as it is to say that God creates the world.

THE DIVINE ATTRIBUTES

In most systems of religion and philosophy, God is endowed with characteristics that distinguish him from other forms of being.

INFINITY. The infinity of the Christian God was implied above in the accounts of his transcendence and creative power, and in most systems, God's infinity makes him free, in degree, if not in kind, from at least some human limitations. But he is not strictly infinite unless he is limitless throughout the whole range of his existence. He can be wholly limitless, however, only if he is self-existent and thereby self-sufficient. If (as Hegel thought) God needs the world as the sphere of his self-development, or if (as Plato thought) he copies an independent realm of Forms, he is *pro tanto* limited. He is strictly infinite only if his essence is identical with existence, as Thomas held when he said that the most appropriate name for God is the one disclosed to Moses according to the Vulgate text of Exodus—*Qui Est* ("He Who Is"). If God is thus infinite, he must possess all properties in a mode that is free from every limitation. He must be one, simple, incorporeal, immutable, impassible, eternal, good, omniscient, and omnipotent.

UNITY. The Greek philosophers were apt to speak interchangeably of "god" and "the gods" (as may be seen, for example, from Plato's *Laws* 900–905 and the *Discourses* of Epictetus 1,3,1). But in Judaism the belief that Yahweh is the only God became an unquestioned axiom that was inherited by Christians and defended by Thomas on the grounds that if there were two gods, one would possess what the other lacked, so that neither would be absolutely perfect (*Summa Theologiae* Ia, 11, 3). Similarly, Muslims hold as a primary article of faith that "there is no god but God." But Christians differ from Jews and Muslims in believing that the one God exists in a threefold form as Father, Son, and Holy Spirit. He is one substance (*substantia, ousia*) in three persons (*personae, hypostaseis*).

SIMPLICITY. According to Christianity and Neoplatonism, God is one also in the sense that he is absolutely simple; for the distinctions (such as those between essence and existence, substance and accidents) that make a finite being composite are inapplicable to him. Plotinus interpreted this simplicity as a bare, characterless, self-identity. But Thomas held both that God actually possesses the perfections we ascribe to him and that these coalesce in an unimaginable unity. Each of God's attributes is objectively distinct, but each expresses his whole being.

INCORPOREALITY. Those philosophers who regard the world as an aspect of God or an unfolding of his essence are obliged to think of him materially. Thus the Stoics identified him with nature's basic elements, air and fire. Similarly, Augustine learned from Manichaeism that God is a bright and very subtle substance. But the immateriality of God has constantly been taught by Platonists and Christians on the ground that matter, being a principle of limitation, is incompatible with his perfection.

IMMUTABILITY. That God's nature cannot change (for change implies imperfection) was affirmed by Plato and the Old Testament. It was reaffirmed by Christian theologians, especially Augustine.

IMPASSIBILITY. Impassibility is equivalent to immutability, if it means that God cannot suffer change from either an external or an internal cause. But it has also been taken to mean that God cannot experience pain. While there is an apparent contradiction between this last meaning and Biblical descriptions of God's love, it has been maintained by some theologians (but denied by others) that, although Christ experienced pain in his *human* nature, God cannot experience it in himself, for, being wholly perfect, he is pure Joy.

ETERNITY. In the Bible, God's eternity signifies an everlasting, endless time. In later Christian thought (through the influence of Platonism) it was understood as "timelessness." It is, in the famous definition of Boethius, *interminabilis vitae tota simul et perfecta possessio* ("eternal life possessed perfectly and simultaneously," *De Consolatione Philosophiae* V, vi). God, it is said, would not be perfect unless he possessed his whole being in a simultaneous act.

GOODNESS. The moral order has sometimes been interpreted nontheistically through such abstract ideas as *Rita* (in India), *Dao* (in China), and *Dike* (in Greece). The gods of Greco-Roman polytheism were notoriously immoral. But in Christian thought, Plato's affirmation that God is wholly good (*Republic* 379) was combined with the Hebraic vision of Yahweh's righteousness. Hence Thomas considered it to be axiomatic that "God is sheer goodness, whereas other things are credited with the sort of goodness appropriate to their natures" (*In Boethium de Hebdomadibus* 5).

OMNISCIENCE. Omniscience is entailed by infinity. But a special problem is created by the view that God now knows future *freely chosen* human acts. Those who hold this view urge, first, that since God is timeless it is, strictly speaking, incorrect to say that he "foreknows" events, and second, that even if we say this (speaking from our finite standpoint), we need not assume that a human act, because it is foreknown, is predetermined—by either God or any other factor outside the agent's will. To say that a human act can in principle be predicted is not to say that the agent has no control over it or is not really active and responsible for what he does; this, at any rate, is a view of human action widely held by philosophers at the present time. But other theists (notably James Ward and F. R. Tennant) consider it contradictory to say that a free choice can be known in any sense until it has been made. They affirm that God is ignorant of future human choices and that his ignorance is a "self-limitation" he deliberately incurred in granting man free will.

OMNIPOTENCE. Omnipotence too is entailed by infinity. It is important to note that in the Creeds, *Pantocrator* and *omnipotens* imply that God is ruler of all things, rather than that he can do anything. He cannot act against either reason or morality. But it is extremely difficult to explain the existence of evil in a world created by a God who is both infinitely powerful and infinitely good. Various explanations have been given. Thus, evil has been traced to the fall of a first man or World Soul. Again, it is said that God permits (even if he does not inflict) unmerited suffering as a means of purifying the soul for eternal life. But many theologians would endorse Friedrich Von Hügel's frank admission that no explanation is fully satisfying. It is therefore not surprising that some philosophers (notably J. S. Mill) have tried to relieve God of apparent responsibility for evil by supposing that he is finite both in knowledge and in power. (Christians believe that God displays his omnipotence by *overcoming* evil through the ministry of Christ; but an exposition of this belief would involve a study in the doctrines of Incarnation and Atonement.)

PERSONALITY

In the preceding sections it has been assumed that God is personal. The assumption is justified by the fact that, while in the primitive stages of religion he has often been conceived subpersonally, philosophers (in the West, at any rate) have nearly always described his nature to some extent by analogy with the human self. Thus, according to Plato, Aristotle, and Spinoza God has mental properties. But two conditions must be fulfilled if God is to be *fully* personal. First, it must be possible to speak of him as loving, or caring for, humankind. Second, it must be possible to speak of him truly through images drawn from human life. The Aristotelian and Spinozistic concepts of him fail to meet the first of these conditions. While Aristotle's First Mover contemplates himself, he does not have any knowledge of the world. Therefore, like Spinoza's God, he cannot return the love that he receives.

The second condition is not universally fulfilled either. Some thinkers have attempted to mediate between philosophy and religion by suggesting that concrete images of God are inadequate attempts to grasp a Reality that is suprapersonal. Thus Hegel held that Absolute Spirit can be adequately known only by the speculative intellect. Consequently, when he speaks of the Absolute as God he means by God (as Aristotle meant) self-thinking Thought. The personal God of theism is a prerational and imperfect representation (*Vorstellung*) of the Absolute. On the ascending scale of truth, religion occupies an intermediate place between art and philosophy.

This contrast between religion and philosophy becomes even more acute when the Absolute is equated with a suprarational Unity. Here there is a striking parallel between Indian monism and the thought of F. H. Bradley. Some Hindu scriptures (notably the *Bhagavad-Gita*) describe God as a personal being, the Lord of the universe, whose "grace" (*prasada*) requires the "loving devotion" (*bhakti*) of his worshipers. The *Gita* is especially significant. Through the theophany in the eleventh chapter, it declares that Krishna (the incarnate God, and friend of Arjuna) is "more to be prized even than Brahman." But Śankara, following the nondualistic strain in the Upanishads, held that the sole reality is the impersonal Absolute (Brahman) with which the soul is numerically identical. Personal concepts of the Absolute belong to the sphere of illusion (maya). They are forms under which the One appears to untutored minds. Likewise F. H. Bradley held that since Reality is nonrelational, a per-

sonal God is “but an aspect, and that must mean but an appearance, of the Absolute” (*Appearance and Reality,* Oxford, 1930, p. 397).

Christians, however, are obliged by revelation to identify the Absolute with a God who is fully personal, both in himself and in his dealings with humankind. Such primary images as Father, King, and Friend mediate a knowledge that cannot be surpassed by abstract speculation. During this century the personal nature of religious conviction has been stressed in varying terms by such writers as William Temple, John Oman, John Baillie, Karl Barth, Emil Brunner, Martin Buber, and the existentialists (especially Kierkegaard, Bultmann, and Gabriel Marcel). Buber’s distinction between an “I-Thou” and an “I-It” relationship and Kierkegaard’s contrast between subjectivity and objectivity have been widely used to express the difference between a personal and an impersonal attitude to God. At the same time, many theologians are aware that an unqualified application of personal categories to God results in anthropomorphism. Divine personality wholly transcends its finite counterpart. It is unique both because of the fact that essence and existence are identical in it and because of the mystery of its triune character.

THE KNOWLEDGE OF GOD. There are three main routes to God: reason, revelation, and religious experience.

Both Plato and Aristotle claimed that reason can obtain a certain knowledge of God’s existence and nature. This claim has been endorsed by many Christian theologians. Thus, St. Augustine, writing from within the Platonic tradition, affirmed that the human intellect by nature participates in eternal Truth. Furthermore, many theologians have held that God’s existence can be proved. These proofs may be divided between those which take the form of a priori reasoning from God’s essence and those which take the form of a posteriori reasoning from finite experience. The first type of proof is exemplified chiefly by the Ontological Argument, which was first formulated by St. Anselm and restated by René Descartes. In its Anselmic form it runs as follows: The idea of God is the idea of that than which nothing greater can be conceived; a being that exists is greater than a being that does not exist; therefore God exists. In view of the criticisms to which this proof has been subjected (especially by Thomas and Immanuel Kant), it is widely considered to be invalid by both theologians and philosophers today. The main a posteriori arguments received their classical formulation from Thomas.

He constructed five proofs based on the facts of motion, causality, contingency, relative perfection, and design. (The first, second, and third of these Five Ways are different forms of the Cosmological Argument—the argument that the world in all its aspects shows its dependence on self-existent Being.) Kant rejected *all* proofs based on the use of the “speculative reason.” But he maintained that the “practical reason” is obliged to postulate both God and immortality. Since World War I, natural theology has been vigorously attacked, on the one hand by Barth and, on the other, by those philosophers who deny the possibility of metaphysics. However, many twentieth-century philosophers (chiefly Roman Catholic Thomists—but also others, such as A. E. Taylor) held that the main a posteriori proofs can be presented cogently.

Thomas affirmed that in addition to a natural knowledge of God there is a supernatural knowledge revealed by Christ and received through faith. Thus, while reason can infer that God is the Creator, it cannot discover that he is Three-in-One. John Locke reproduced this distinction in his *Essay concerning Human Understanding* (Book 4, Ch. 18). But in his *Reasonableness of Christianity* he paved the way for the deists, who held that the Gospel merely “republishes” the basic truths of natural religion and morality. The supernatural character of revelation was also denied later by those Hegelians who regarded Christ as the highest instance of the Absolute’s universal presence in humanity.

Religious philosophers from Plato onward have claimed that it is possible to have a *direct* knowledge of divine reality. Among Christian thinkers, some hold that this knowledge is available (even if in a confused form) to everyone; others restrict it to the recipients of biblical revelation. Some regard it as the highest activity of ordinary mental powers; others assign it to a special faculty of the soul. Some describe it intellectually as an insight or intuition; others stress its volitional character by calling it a confrontation or encounter. Apart from these differences, it is necessary to distinguish between an experience that is mediated and one that is immediate. As many recent writers have stressed (notably, William Temple, John Oman, and H. D. Lewis), religious experience is normally mediated through secular experiences, including those which are formulated in the premises of the a posteriori proofs. Thus, we become aware of God as eternal through the contingency of finite things and as holy through the demands of the moral law. (Even the divinity of Christ is experienced, in the first place, through meditation on his human life and on the impact that it made on his disciples.) But there is also an immediate, purely spiritual

experience that is called "mystical." While Christian and non-Christian mystics often use the same terminology, the former (when they are orthodox) differ from many of the latter at two points. First, they affirm that God is transcendent as well as immanent. Second, and as a consequence, they claim, not an absorption into the Godhead, but a union of love and will in which the distinction between the Creator and the creature is permanently retained.

See also Absolute, The; Alexander, Samuel; Aristotle; Augustine, St.; Barth, Karl; Bradley, Francis Herbert; Brunner, Emil; Buber, Martin; Bultmann, Rudolf; Caird, Edward; Cosmological Argument for the Existence of God; Descartes, René; Emanationism; God/Isvara in Indian Philosophy; Hegel, Georg Wilhelm Friedrich; Hiddenness of God; Hügel, Baron Friedrich von; Infinity in Theology and Metaphysics; Kant, Immanuel; Kierkegaard, Søren; Locke, John; Mani and Manichaeism; Marcel, Gabriel; Mill, John Stuart; Neoplatonism; Oman, John Wood; Ontological Argument for the Existence of God; Otto, Rudolf; Plato; Platonism and the Platonic Tradition; Plotinus; Pseudo-Dionysius; Spinoza, Benedict (Baruch) de; Taylor, Alfred Edward; Tennant, Frederick Robert; Thomas Aquinas, St.; Tillich, Paul; Whitehead, Alfred North.

Bibliography

A. E. Taylor's article on theism in *Encyclopedia of Religion and Ethics,* edited by J. Hastings (New York, 1908–1926), is still authoritative. The best single book on Greek thought about God up to and including Plato remains James Adam's *The Religious Teachers of Greece* (Edinburgh: Clark, 1908). On the period from Augustine to Scotus, there are admirable accounts by F. Copleston in Vol. 2 of his *A History of Philosophy* (London: Burns, Oates & Washbourne, 1954). This book also contains an excellent bibliography. Most of the relevant passages in Thomas are assembled in the two anthologies, *Philosophical Texts* and *Theological Texts,* selected and translated by T. Gilby (Oxford: Oxford University Press, 1951 and 1955). On the modern period, there is much useful material in A. S. Pringle-Pattison's *The Idea of God* (Oxford: Oxford University Press, 1920) and in E. L. Mascall's *He Who Is* (London: Longmans, Green, 1943). The latter reexamines the essentials of theism, and contains a full bibliography. See also J. Collins's very comprehensive *God in Modern Philosophy* (Chicago: Regnery, 1959); Illtyd Trethowan's *The Basis of Belief* (London: Burns and Oates, 1960), a concise summary of recent work on natural theology; and F. Ferré's *Language, Logic and God* (London: Eyre and Spottiswoode, 1962), a full account of recent writing on religious subjects from an empiricist point of view. The classical analysis of religious experience remains Rudolf Otto, *Das Heilige* (Marburg, 1917), translated by J. W. Harvey as *The Idea of the Holy* (New York: Oxford University Press, 1925). The best modern defense of the miraculous element in Christianity is C. S. Lewis, *Miracles* (London: G. Bles, 1960). A full survey of recent writing on revelation is contained in John Baillie, *The Idea of Revelation in Recent Thought* (New York: Columbia University Press, 1956).

H. P. Owen (1967)

GOD, CONCEPTS OF [ADDENDUM]

Since H. P. Owen's entry there has been considerable work on Western theism's standard roster of divine attributes. One of the most-discussed, eternity, has its own entry. This entry notes developments on three others.

DIVINE FOREKNOWLEDGE: THE PROBLEM

Many biblical passages ascribe to God knowledge of what we will freely do in the future. But a now-standard argument (derived from Boethius) contends that no future creaturely action can be both foreknown and free. Suppose that for some act A, God believed yesterday that I do A tomorrow.

God is infallible. He cannot make a mistake. That is,

(1) Necessarily, for all P, if God believed yesterday that P, then P. So,

(2) Necessarily, if God believed yesterday that I do A tomorrow, I do A tomorrow.

(3) If God believed yesterday that I do A tomorrow and it is in my power not to do A tomorrow, then it is in my power to make it the case that yesterday God had a false belief, or it is in my power to make it false that God believed yesterday that I do A tomorrow.

(4) It is not in my power to make God have had a false belief (from 2).

(5) It is not in my power to alter the past. So,

(6) It is not in my power not to do A tomorrow. So,

(7) Tomorrow I do not do A freely.

Philosophers have revived approaches to this problem associated with the scholastic thinkers William of Ockham and Luis de Molina.

OCKHAMISM

Ockhamism rejects (4) in favor of

(4a) If it is in my power not to do A tomorrow, and God believed yesterday that I do A tomorrow, then: it is in my power to make it the case that yesterday God had a false belief; it is in my power to make it false that God believed yesterday that I do A tomorrow; or it is in my power to do something such that had I been going to do it, God would always not have believed that I would do A.

If it was the case yesterday that an infallible God had this belief, then I was going to do A. But it was also going to be the case that I have the power not to do A, even though I will not use this power. For Ockhamism, that I was going to do A determines what God believed, not vice versa: My future act constrains the past, rather than God's past belief constraining the future. Had I been going not to do A tomorrow, this would have determined what God believed, and so He would always have believed that I would not do A. Thus, since I have it in my power not to do A, it is in my power to do something such that had I been going to do it, God would always not have believed that I would do A.

But just how does what I do in the future determine what God believed? Ockhamists hold that what makes (1) true is not wholly in the past. Rather, as they see it, what makes a statement about God's past beliefs true is partly in the past and partly in the future. If (1) is true, for Ockhamism, what makes it true is partly God's mental state yesterday and partly my doing A tomorrow. One large question Ockhamism faces is how precisely to understand this. Perhaps God "sees" the future: Had I been not doing A in the future, that is what He would always have seen. Plausibly, when I see a tree, what makes this true is that I am seeing and that the tree is seen. So perhaps future events are part of what make it true that God sees the future.

But when I see a tree, this is because light reflected from the tree enters my eye—the tree sends a signal. So if God "sees" the future, future events send signals back to God in the present—there is backward causation. This is hard to defend.

MOLINISM

Molinism accepts (4a), not (4). For Molinism, God knows our future free actions by knowing "counterfactuals of freedom" (CFs), truths about what we would freely do in various circumstances, and knowing the circumstances we will be in. For Molinists, before all creation, it was true that

(8) were the snake and Eve in the garden, the snake eventually would freely tempt Eve, and

(9) were the snake to tempt Eve in the garden, she would freely fall.

God knew this. God placed Eve and the snake in the garden. As (8) was true, the snake freely tempted her. Since (9) was true, she freely fell. And so on. God decides who is created and what initial and later circumstances they face in light of His knowledge of all CFs. So God sets up the whole future, including the parts we would do freely. God knows the future by knowing the CFs and how He has set things up. But our freedom, say Molinists, is built into the CFs. Eve had it in her power not to fall. Had the CF

(10) were the snake to tempt Eve in the garden, she would not freely fall

been true instead of (9), God would have known that. And since Eve had it in her power not to fall, she had it in her power to do something such that had she been going to do it, God would always not have believed that she would fall.

One problem for Molinism concerns what makes CFs true. It cannot be God. If God determines both that (8) is true and that the snake and Eve are in the garden, God determines that the snake sins: Sin is God's fault. Furthermore, God simply determines all our actions, and so we are not free. Nor can it be our natures. Whatever our having our natures makes true is true necessarily. But CFs cannot be true necessarily. If we suppose that (8) is necessarily true, it is not so much as possible that the snake not tempt Eve, and so the snake does not do so freely—and so (8) turns out false. Nor can it be our actions. They come too late in the game. So there seems to be nothing at all in reality that can make CFs true.

Furthermore, arguably Molinism does not genuinely preserve our freedom. (8) is true from all eternity. It is also true from eternity that God has willed that Eve and the snake be in the garden. So from all eternity it is guaranteed that the snake tempts. It cannot be the case that God has willed that Eve and the snake be in the garden and that (8) is true and yet the snake does not tempt. Given conditions obtaining long before it existed, the snake cannot do otherwise: It is merely the case that it could have done otherwise (had these conditions not obtained). And there is a further worry: Even if God does not will that the snake tempt, does He not in fact initiate the snake's action, albeit indirectly? What initiates an action makes the first difference in the world that guarantees that (barring a miracle) the action is done. On the

Molinist scenario, we do not make this first difference at the moment of choice. God makes it from all eternity. If this is correct, then God removes our responsibility much as He would if He just directly caused our actions. Thus, it is not clear that Ockhamism or Molinism are successful approaches to the Boethian argument.

THE CONCEPT OF OMNIPOTENCE

The claim that God is omnipotent, that is, "all-powerful," concerns the range of God's power—how much He is able to do. At first glance, being all-powerful may seem to be having all powers. But this cannot be right. God has no body (leaving aside the Incarnation). So He cannot walk. He can take on a body. So He is able to walk, and He actually has a conditional power, the power to walk if He acquires a body. But having this conditional power does not entail being able to walk.

St. Thomas Aquinas gave a classic definition, that "God is omnipotent =df. God is able to bring about every absolutely possible state of affairs" (*Summa Theologiae* Ia 25, 3). The thought here is that God can make true any sentence stating something possible, but no sentence stating something impossible, for example, not "there is a square circle." In a sense this does not limit God's power. If a sentence describes something God can bring about, it describes something that can occur, so it states a possibility. So no sentence stating something impossible could describe something God can do. But what can seem to be a limit emerges when we add that some sentences really do state impossibilities, states of affairs God cannot bring about. One may wonder why there are any impossibilities if God is all-powerful. In response, some have wondered whether the sense that there being impossibilities limits God dissipates if we add that God's nature or activity accounts for these states of affairs being possible or impossible.

Discussions of omnipotence suggest that Aquinas's account is too sweeping—that there are some possible states of affairs God cannot bring about. These include

(a) Necessary states of affairs. Necessarily, 2 + 2 = 4. But this, some argue, has nothing to do with God. It is not something He or anyone else could bring about. Still, whatever is necessarily so is possibly so.

(b) Things God is too late to bring about, for example, that the Germans won World War II. It was possible that they do so, and it is now no contradiction to say "the Germans won the war." But it is now (some argue) impossible that anyone bring this about. Thus, some suggest that we relativize omnipotence to what it is possible to bring about at any given time.

(c) The free actions of creatures. It is possible that I finish writing this entry. But if God makes me finish, some say my doing so is not free. So while it is possible that God makes me finish, some say it is not possible that God makes me freely finish.

(d) States of affairs that would be evil to bring about. Many think that God cannot do evil. If this is so, then if it would be evil to kill you, God cannot do this.

(e) States of affairs that entail that no one brought them about. God can make an atom appear from nowhere. But He cannot make one so appear with absolutely no cause, because if He brings it about that the atom appears, its appearing has a cause.

(f) Molinist CFs.

It is hard to make definition building in all these exceptions seem smooth and natural. But two accounts are worth noting. One could say that God is omnipotent =df. God has the greatest range of power one individual can have. This lets one place outside God's power as many of (a) to (f) as one wants. Suppose that no one can bring about at t states of affairs earlier than t. Then this is not in the range of power of any individual. So if God cannot do this, God can be too late to bring some things about and yet have the greatest range of power one individual can have. Again, it is in my power to kill you when it is evil to do so, and perhaps it is not in God's. Even so, God can still have the greatest range of power one individual can have, because His range of power is overall larger than mine.

Another definition that also allows exceptions (a) to (f) begins from the thought that omnipotence cannot ever be powerful enough to do anything: God is omnipotent =df. there is nothing such that God cannot do because of a lack of power. Anyone other than God cannot do some things because of a lack of power. I cannot run a three-minute mile because my legs are not that powerful. A turtle cannot do mathematics because its mind is not that powerful. On this last definition, the "distinctive" of God's power is that nothing like this is true of Him. God cannot walk because of a lack of a body, not because of a lack of power. There is no power to walk-without-a-body to lack. God cannot sin because of His moral perfection, not because of a lack of any power needed, for example, to tell a lie. If He cannot make the Germans win World War II, this is because it is now too

late, not because of a lack of power. And so on. One may wonder, though, whether having a set of powers with so many limits, even if not because of a lack of power, amounts to being *all*-powerful.

DIVINE SIMPLICITY

Many medieval thinkers held that God is "simple," and this thesis returned to active discussion in the 1980s. The more complicated or complex something is, the more parts it has. The simpler something is, the fewer parts it has, and something is wholly simple if it has no parts at all. The doctrine of divine simplicity (DDS) asserts that God is wholly simple. Most theists think that God is not made of matter, spread out in space, or an event, with earlier and later parts. So they think that God has no material, spatial, or temporal parts. But Aquinas, for instance, also speaks of God as not "composed" of essence, accidents, and existence (*Summa Theologiae* Ia 3, 3–6). The theist mainstream is silent on these. And well it may be, for what Aquinas presupposes here is not at all a matter of common sense.

Parts compose. Wholes are composed. Wholes can consist completely of different sorts of parts at once—our bodies consist completely of both molecules and quarks. When Aquinas speaks of things other than God as "composed" of, for example, essences and accidents, he takes it that concrete things consist completely not only of concrete parts but also of abstract ones—essences, accidents, and so on This claim stretches the sense of "part." For the part-whole relation is transitive. If a bolt is part of a wheel and the wheel is part of a car, it follows that the bolt is part of the car. The subject-attribute relation (that between things and their accidents) is not transitive. Being an attribute is an attribute of my accidents but not an attribute of mine.

If DDS is true, then while what makes it true that Brownie is a donkey is that Brownie has an essential property, donkeyhood, what makes it true that God is divine is simply God. Again, while what makes it true that Brownie is brown is that Brownie has an accidental property, brownness, what makes it true that God is just is simply God. We could put this a bit crudely by saying that given DDS, all God's properties are identical with God.

This courts two obvious objections. One is that if all God's attributes are identical with God, then they are all identical: God has just one attribute. But that does not seem true. God is wise and omnipotent. These seem to be two attributes.

Still, this is a bit quick. *Wise* and *omnipotent* are two predicates. But it is not obvious that real attributes pair 1:1 with (almost all) predicates. And even if predicates usually apply in virtue of a thing's having distinct attributes, it would take some argument to show that they can never apply in virtue of the same thing. Perhaps what made it true at one time that someone was king of England was that he rightfully wore one crown, and what made it true at one time that someone was king of Scotland was that he rightfully wore another crown. When the monarchies united, both became true in virtue of rightfully wearing a single crown.

A second obvious objection is that if God is identical with His properties, then He is a property, yet surely He is not. This objection takes it that the result of identifying God with His properties is to eliminate God. But why think that? Perhaps the identification gets rid of the properties, leaving God to make it true that He is wise, good, and so on. If an identity-statement A = B is true, then where one could have thought there to be two items, A and B, there is only one. This one has all attributes A really has and all attributes B really has, but may have only some attributes A has been thought to have and some B has been thought to have. It may even have none of either. (Here's a partial analogy: Suppose a spy is also a bigamist. He is discovered, and his wives meet. Each wife may learn that the husband's real life story includes nothing he has told her about himself. In fact, by pooling clues that previously only one had possessed, the wives may eventually learn that he is not even human, but really a robot. So the wives may well wind up learning that almost nothing either had believed about the husband was true.) Given that A = B, which of the attributes we thought A and B had this one thing has remains to be determined. If God = justice, why think that God has only attributes we thought justice had? Some might reply that the property has to be there regardless—there has to be such a thing as justice, because others than God are just. But actually one can use God in place of the property in other metaphysical contexts, though this is a story too complex to tell here.

See also Foreknowledge and Freedom, Theological Problem of; Molina, Luis de; Ockhamism.

Bibliography

Fischer, John Martin, ed. *God, Foreknowledge, and Freedom*. Stanford, CA: Stanford University Press, 1998.

Flint, Thomas P. *Divine Providence: The Molinist Account*. Ithaca, NY: Cornell University Press, 1998.

Flint, Thomas, and Alfred J. Freddoso. "Maximal Power." In *The Existence and Nature of God*, edited by Alfred J. Freddoso. Notre Dame, IN: University of Notre Dame Press, 1983.

Hoffman, Joshua, and Gary Rosenkrantz. "Omnipotence Redux." *Philosophy and Phenomenological Research* 49 (1988): 283–301.

Hughes, Christopher. *On a Complex Theory of a Simple God: An Investigation in Aquinas' Philosophical Theology*. Ithaca, NY: Cornell University Press, 1989.

Morris, Thomas V. *Anselmian Explorations: Essays in Philosophical Theology*. Notre Dame, IN: University of Notre Dame Press, 1987.

Plantinga, Alvin. *Does God Have a Nature?* Milwaukee, WI: Marquette University Press, 1980.

Wielenberg, Eric. "Omnipotence Again." *Faith and Philosophy* 17 (1) (2000): 26–47.

Wierenga, Edward R. *The Nature of God: An Inquiry into Divine Attributes*. Ithaca, NY: Cornell University Press, 1989.

Zagzebski, Linda Trinkaus. *The Dilemma of Freedom and Foreknowledge*. New York: Oxford University Press, 1991.

Brian Leftow (2005)

GÖDEL, KURT

(1906–1978)

Kurt Gödel, a logician, was born in Brno, in what is now the Czech Republic, and educated at the University of Vienna, where he became privatdozent in 1933. In 1940 he joined the Institute for Advanced Study in Princeton, New Jersey, where he remained for the rest of his career. Following David Hilbert, Gödel was instrumental in establishing mathematical logic as a fundamental branch of mathematics, achieving results such as the incompleteness theorems that have had a profound impact on twentieth-century thought. In philosophy, by contrast, he represents the path not taken. Of his few writings in this area, including posthumous publications, those that focus on the more immediate ramifications of his own (and closely related) mathematical work have had the greatest impact.

GÖDEL'S INFLUENCE

A close student of the history of philosophy, Gödel follows Plato, Gottfried Wilhelm Leibniz, and Edmund Husserl as opposed to the more fashionable Aristotle, Immanual Kant, and Ludwig Wittgenstein. (On Kant, however, see Gödel 1946/9 and 1961.) Methodologically, two patterns in his thinking stand out. First, a tendency to move from the possible to the actual is reflected in his Leibnizian ontological argument for the existence of god (Gödel 1970). He relies here on the S5 modal principle, (possibly necessarily P ⊃ necessarily P). It can also, arguably, be discerned in his mathematical Platonism—because the distinction between the possible and the actual, relevant to material being, collapses in the formal realm of mathematics (see Yourgrau 1999). Finally, in relativistic cosmology (Gödel 1949, 1946/9) he concludes from the possible existence of rotating universes, where time is merely ideal, to its ideality in the actual world.

Second, he is preoccupied with probing mathematically the limits of formal methods in representing intuitive concepts. In his first incompleteness theorem, for example, by applying an ingenious arithmetization of metamathematics to a formal system of arithmetic, Gödel was able to construct a formula expressing its own unprovability, and thus to prove (as he made explicit later) the indefinability within the system of the intuitive concept of arithmetic truth (see Feferman 1984). Along the same lines one may view his results in cosmology as demonstrating the limits of the theory of relativistic space-time in representing the intuitive concept of time, although here, interestingly, his response was to abandon the intuitive concept (see Yourgrau 1999).

From a broader perspective Gödel isolates two basic philosophical worldviews: one with a "leftward" direction, toward skepticism, materialism, and positivism, the other inclined toward "the right," toward spiritualism, idealism, and theology (or metaphysics; Gödel 1961). He puts empiricism on the left and a priorism on the right and points out that although mathematics, qua a priori science, belongs "by its nature" on the right; it too has followed the spirit of the times in moving toward the left—as witnessed by the rise of Hilbert's formalism. With Gottlob Frege, Gödel resists this trend, pointing to his incompleteness theorems as evidence that "the Hilbertian combination of materialism and aspects of classical mathematics … proves to be impossible" (1961, p. 381).

FREGE AND GÖDEL

Frege's mathematical philosophy is held together by two strands that may appear to be in tension with one another: on one side his Platonism and conceptual realism, on the other his conception of arithmetic as analytic (that is, as resting on definitions and the laws of logic) and his "context principle" (which seems to put our sentences—hence language—at the center of his philosophy). This second aspect of Frege's thought, via Bertrand Russell and Wittgenstein, helped persuade the positivists of the Vienna Circle (whose meetings Gödel attended) that mathematics is without content, a mere matter of (more or less arbitrary) linguistic conventions concern-

ing the syntax of (formal) language. This conclusion was, however, rejected by both Frege and Gödel (1944, 1951, 1953–59), Frege hoping, contra Kant, "to put an end to the widespread contempt for analytic judgments and to the legend of the sterility of pure logic" (1884, p. 24; see also 1879, p. 55). Gödel, for his part, insists that "'analytic' does not mean 'true owing to our definitions,' but rather 'true owing to the nature of the concepts occurring therein'" (1951, p. 321). (See Parsons, 1994.)

Frege and Gödel are in further agreement against the spirit of the times, that the fundamental axioms of mathematics should be not simply mutually consistent but (nonhypothetically) true. They also reject Hilbert's conception of axiom systems as "implicit definitions," with Gödel insisting that a formal axiomatic system only partially characterizes the concepts expressed therein. Indeed, his Incompleteness Theorem makes the point dramatically: "Continued appeals to mathematical intuition are necessary … for the solution of the problems of finitary number theory.… This follows from the fact that for every axiomatic system there are infinitely many undecidable propositions of this type" (1947 [1964], p. 269). And it is in human ability—if indeed humans possess it—to intuit new axioms in an open-ended way that Gödel sees a possible argument to the effect that minds are not (Turing) machines (Gödel 1951; Wang 1996).

What kind of intuitions, however, are these? Gödel does, it is true, employ a Kantian term here, but he does not mean concrete immediate individual representations, and on just this point he faults Hilbert: "What Hilbert means by '*Anschauung*' is substantially Kant's space-time intuition.… Note that it is Hilbert's insistence on *concrete* knowledge that makes finitary mathematics so surprisingly weak and excludes many things that are just as incontrovertibly evident to everybody as finitary number theory" (1958 [1972], p. 272, n. b). (See also 1947 [1964], p. 258.) Note, further, that mathematical intuition, though a form of a priori knowledge, does not ensure absolute certainty, which Gödel rejects (Wang 1996); rather, as with its humbler cousin, sense perception, it too may attain various degrees of clarity and reliability (see Gödel 1951, his remarks on Husserl in 1961, and Parsons 1995, 1995a).

THE GÖDEL PHILOSOPHY

Frege and Hilbert, then, serve as useful coordinates in mapping Gödel's philosophy, in its tendency to "the right." What if one chooses Albert Einstein as a third coordinate? Note first that "idealistic" in the title of Gödel (1949) is not a gesture toward a subjective philosophy such as George Berkeley's. (In his final years, he became sympathetic with Husserl's later idealism, which does not exclude objectivism. See van Atten and Kennedy 2003.) Rather, Gödel is pointing to the classic Platonic distinction between appearance and reality. Though the world may appear (to the senses) as if temporal, this is in fact an illusion. Only reason—here, mathematical physics—can provide a more adequate cognition of reality (i.e., of Einstein-Minkowski space-time). Gödel makes a sharp distinction between intuitive time, which lapses, and the temporal component of space-time. By his lights, already in the special theory of relativity (STR) intuitive time has disappeared, because "the existence of an objective lapse of time means … that reality consists of an infinity of layers of 'now' which come into existence successively" (Gödel 1949, pp. 202–203), whereas the relativity of simultaneity in the STR implies that "each observer has his own set of 'nows,' and none of these various systems of layers can claim the prerogative of representing the objective lapse of time" (p. 203).

These observations, however, rely on the equivalence of all "observers" or reference frames in the STR, whereas in the general theory of relativity (GTR), of which the STR is an idealized special case, the presence of matter and the consequent curvature of space-time permit the introduction of privileged observers, in relation to which one can define a "world time" (which, one may say, objectively lapses). Gödel's discovery is that there exist models of the GTR—the rotating universes—where, provably, no such definition of a world time is possible. In particular, these worlds permit time travel, in the sense that, "for *every* possible definition of a world time one could travel into regions of the universe which are past according to that definition," and "this again shows that to assume an objective lapse of time would lose every justification in these worlds" (1949 p. 205). The idea here is clearly that if a time has "objectively lapsed," it no longer exists and so is not *there* to be revisited (in the future). Hence, by contraposition, if it can be revisited, it never did objectively lapse in the first place.

To describe the Gödel universe as static, however, as opposed to our own, would be misleading. The time traveler's rocket ship, for example, would move at a speed of at least $1/\sqrt{2}$ of the velocity of light! It would seem to observers, just as in this world, to be moving at great speed, and in general the denizens of Gödel's universe may well experience time much as we do in the actual world. Indeed, that is why Gödel moves from the mere possible existence of the Gödel universe to the ideality of time in the actual world, because "if the experience of the

lapse of time can exist without an objective lapse of time, no reason can be given why an objective lapse of time should be assumed at all" (p. 206; see Yourgrau 1999).

Here, then, is another example of the Janus-faced quality of Gödel's thinking, presaged already in his arithmetization of metamathematics—contributing mathematically to "the left" while at the same time, as he sees it, pointing to "the right."

See also Gödel's Incompleteness Theorems; Logic, History of; Mathematics, Foundations of.

Bibliography

PRIMARY SOURCES

"Russell's Mathematical Logic" [1944]. In *Collected Works*, vol. 2. New York: Oxford University Press, 1990.

"Some Observations about the Relationship between Theory of Relativity and Kantian Philosophy" [1946/9]. In *Collected Works*, vol. 3. New York: Oxford University Press, 1994.

"What Is Cantor's Continuum Problem?" [1947; 1964]. In *Collected Works*, vol. 2. New York: Oxford University Press, 1990.

"Some Basic Theorems on the Foundations of Mathematics and Their Implications" [1951]. In *Collected Works*, vol. 3. New York: Oxford University Press, 1994.

"The Modern Development of the Foundations of Mathematics in the Light of Philosophy" [1961], transl. E. Köhler and H. Wang. In *Collected Works*, vol. 3. New York: Oxford University Press, 1994.

"Ontological Proof" [1970]. In *Collected Works*, vol. 3. New York: Oxford University Press, 1994.

"On an Extension of Finitary Mathematics Which Has Not Yet Been Used" Translated by L. F. Boron and revised by K. Gödel. In *Collected Works*, vol. 2. New York: Oxford University Press, 1990.

"A Remark about the Relationship between Relativity Theory and Idealistic Philosophy" In *Collected Works*, vol. 2. New York: Oxford University Press, 1990.

"Is Mathematics Syntax of Language?" In *Collected Works*, vol. 3. New York: Oxford University Press, 1994.

SECONDARY SOURCES

Buldt, B., et al., eds. *Kurt Gödel: Wahrheit und Beweisbarkeit*. Vienna: Hölder-Pichler-Tempsky, 2002.

Dawson, J. *Logical Dilemmas: The Life and Work of Kurt Gödel*. Wellesley: A. K. Peters, 1995.

Feferman, S. "Kurt Gödel: Conviction and Caution." *Philosophia Naturalis* 21 (1984): 546–562.

Frege, G. *Begriffsschrift: A Formula Language, Modeled upon That of Arithmetic, for Pure Thought* [1879]. Translated by S. Bauer-Mengelberg. In *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by J. van Heijenoort. Cambridge, MA: Harvard University Press, 1967.

Frege, G. *The Foundations of Arithmetic* [1884]. Translated by J. L. Austin. Evanston, IL: Northwestern University Press, 1980.

Kreisel, G. "Kurt Gödel." *Biographical Memoirs of Fellows of the Royal Society* 26 (1980): 149–224.

Parsons, C. "Platonism and Mathematical Intuition in Kurt Gödel's Thought." *The Bulletin of Symbolic Logic* 1 (1) (1995a): 44–74.

Parsons, C. "Quine and Gödel on Analyticity." In *On Quine: New Essays*, edited by P. Leonardi., Cambridge, MA: Cambridge University Press, 1995.

van Atten, M., and J. Kennedy. "On the Philosophical Development of Kurt Gödel." *The Bulletin of Symbolic Logic* 9 (4) (December 2003): 425–476.

Wang, H. *From Mathematics to Philosophy*. New York: MIT Press, 1974.

Wang, H. *A Logical Journey: From Gödel to Philosophy*. Cambridge, MA: MIT Press, 1996.

Yourgrau, P. *Gödel Meets Einstein: Time Travel in the Gödel Universe*. Chicago: Open Court, 1999.

Yourgrau, P. "Review Essay: Hao Wang, *Reflections on Kurt Gödel*." *Philosophy and Phenomenological Research* 1 (1989): 391–408.

Palle Yourgrau (1996, 2005)

GÖDEL'S INCOMPLETENESS THEOREMS

The axiomatic method is at the heart of mathematics. The work of mathematicians is to derive the consequences of axioms. According to Euclid, axioms are evidently true, and deduction from them is a powerful method of learning new truths. The rise of non-Euclidean geometry disrupted the carefree connection between truth and proof and led many modern thinkers to adopt the formalistic attitude that the mathematician's sole endeavor is to work out the consequences of axioms, taking no professional interest in inquiring what, if anything, the axioms are true of.

In 1931 Kurt Gödel proved a deep theorem that showed that deduction from axioms cannot be all there is to mathematical understanding. Gödel showed that, for whatever system of truths of number theory we choose to regard as axiomatic, there will be statements of basic arithmetic that we can recognize as true even though they are not consequences of the axioms. That there are truths not derivable from our axioms is hardly surprising; nobody ever promised us omniscience. What is surprising is that there are arithmetical statements *we can recognize as true* even though they are not derivable, so that no system of axioms we can write down fully captures our arithmetical understanding. Moreover this situation holds not only for systems of axioms we are capable of producing today but also for whatever systems we may devise in the future.

Gödel's true, unprovable sentence is obtained by using strings of numbers to encode strings of symbols, thereby reducing statements about language to statements about numbers. Under such a coding Gödel's sentence says that the system of axioms is consistent. Of course if we accept the axioms, we regard the axioms as true, so we certainly regard them as consistent. But even though adopting the axioms means accepting their consistency, the statement that the axioms are consistent cannot be proved from the axioms. We could adopt the thesis that the axioms are consistent as a new axiom. This would give us a new, larger system of axioms that can prove the consistency of the old system but not the consistency of the new system. We can continue the process of adding consistency statements repeatedly, but however far we go we shall never catch up with Gödel. No consistent system that includes basic arithmetic can prove its own consistency.

Gödel's result has important corollaries, notably, Church's theorem (1936) that there is no algorithm for testing whether a sentence is logically valid and Tarski's theorem (1935) that the set of true sentences of a language cannot be defined within the language itself.

THE LANGUAGE OF ARITHMETIC

Gödel's results apply to the language of arithmetic, which is an artificial language for formalizing reasoning about the natural numbers, and to other languages into which the language of arithmetic can be translated. To state his results we need to specify the language exactly. As numerals, the language uses "0" and expressions obtained from "0" by repeatedly prefixing "*S*," which stands for the successor function. The numeral for 3 is "SSS0," which we abbreviate "$\underline{3}$." The language also contains function signs "+" "×" and "E," for addition, multiplication, and exponentiation, so that the *terms* of the language make up the smallest class that contains the numeral "0" and the variables v_0, v_1, v_2, v_3, …, and that contains $S\tau$, $(\tau+\rho)$, $(\tau\times\rho)$, and $(\tau E\rho)$ whenever it contains τ and ρ. In the exposition here we shall sometimes use other letters as variables in place of the official v_is, so as to reduce the proliferation of subscripts. Including "E" as a primitive operation is not strictly necessary, as we shall see below, but it enables us to get off to a fast start.

A term without variables is *closed*. Rules that we learned in elementary school enable us to calculate the numerical value of each closed term. A term with *n* variables represents an *n*-ary function, calculable by a grade-school algorithm.

The *atomic formulas* take the form $\tau = \rho$ or $\tau \le \rho$, where τ and ρ are terms, and the *formulas* constitute the smallest class containing the atomic formulas and containing $\sim \phi$, $(\phi \lor \psi)$, and $(\exists v_i)\phi$, whenever it contains ϕ and ψ. An occurrence within a formula of the variable v_i is *bound* if it occurs within some subformula that begins with $(\exists v_i)$, and it is *free* otherwise. A formula without free variables is a *sentence*; it is sentences that are either true or false. The symbols for conjunction ("$\land$"), the conditional ("$\rightarrow$"), the biconditional ("$\leftrightarrow$"), universal quantification ("$(\forall v_i)$"), and the less-than relation ("<") are treated as defined.

Where v_i does not occur within the term τ, we use $(\exists v_i{\le}\tau)\phi$ and $(\forall v_i{\le}\tau)\phi$ to abbreviate $(\exists v_i)(v_i{\le}\tau \land \phi)$ and $(\forall v_i)(v_i{\le}\tau \rightarrow \phi)$.These are *bounded quantifiers*, and a formula with no quantifiers that are not bounded is a *bounded formula*. For example 'v_0 is prime' is formalized by the bounded formula '$(SS0{\le}v_0 \land (\forall v_1{\le}v_0)(\forall v_2{\le}v_0)(v_0 = (v_1 \times v_2) \rightarrow (v_1 = S0 \lor v_2 = S0)))$'. A set or relation is said to be *bounded* if it is the extension of a bounded formula.

We can test whether an atomic sentence is true by grade-school algorithms; "true," that is, in the standard model consisting of the natural numbers 0,1,2,3, … Any bounded sentence is demonstrably equivalent to a truth-functional combination of atomic sentences, since bounded quantifiers can be cashed out as long but finite disjunctions and conjunctions. Thus we have an algorithm for determining the truth value of a bounded sentence. It follows that every bounded set or relation is decidable; that is, there is an algorithm for testing membership in the set or relation. If *S* is the extension of the bounded formula $\sigma(x_0)$, we can test whether $n \in S$ by asking whether $\sigma(\underline{n})$ is true.

The Σ *formulas* are obtained by prefixing a block of existential quantifiers to a bounded formula, and their extensions are *recursively enumerable* sets and relations. Any recursively enumerable set is the extension of a formula obtained by prefixing a single existential quantifier to a bounded formula, since $(\exists x_1)(\exists x_2)\ldots(\exists x_n)\phi$ is equivalent to $(\exists x_0)(\exists x_1{\le}x_0)(\exists x_1{\le}x_0)\ldots(\exists x_n{\le}x_0)\phi$. (The same goes for recursively enumerable relations; in the future we shall let this go without comment.) The union and intersection of recursively enumerable sets are recursively enumerable, since $((\exists y)\phi(x,y) \lor (\exists z)\psi(x,z))$ and $((\exists y)\phi(x,y) \land (\exists z)\psi(x,z))$ are, respectively, logically equivalent to $(\exists y)(\exists z)(\phi(x,y) \lor (\psi(x,z))$ and $(\exists y)(\exists z)(\phi(x,y) \land (\psi(x,z))$ (assuming bound variables have been chosen so as to avoid conflicts). If $\chi(x,y,z)$ is bounded and τ is a term, $\{x{:}(\exists y{\le}\tau)(\exists z)\chi(x,y,z)\}$ and $\{x{:}(\forall y{\le}\tau)(\exists z)\chi(x,y,z)\}$ are both recursively enumerable since they are the extensions of $(\exists z)(\exists y{\le}\tau)\chi(x,y,z)$ and $(\exists w)(\forall y{\le}\tau)(\exists z{\le}w)\chi(x,y,z)$, re-

spectively. If a Σ sentence is true, we can show it is true by providing an appropriate witness.

Σ FORMULAS AND DECIDABILITY

A set of numbers is *effectively enumerable* if there is a mechanical procedure for listing the set, so that every member of the set turns up on the list eventually and nothing appears on the list that is not in the set. Every recursively enumerable set is effectively enumerable. To see this, we introduce the pairing function. $Pair(x,y) = \frac{1}{2}(x^2 + 2xy + y^2 + 3x + y)$ is a one-one correspondence between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$ (where $\mathbb{N}$ is the set of natural numbers). Define the functions 1st and 2nd so that $Pair(1\text{st}(z), 2\text{nd}(z)) = z$. Given a recursively enumerable function $S = \{x_0 : (\exists x_1)\sigma(x_0, x_1)\}$, with σ bounded, we can list S by the following algorithm: At stage n, test whether the sentence $\sigma(\underline{1\text{st}(n)}, \underline{2\text{nd}(n)})$ is true; if it is, add $1\text{st}(n)$ to the list.

Every set that is known to be effectively enumerable is recursively enumerable. This striking fact, together with a large body of evidence obtained by examining idealized models of computation and examining structural properties of effectively enumerable and recursively enumerable sets, has led to the general acceptance of the *Church-Turing thesis*: A set of natural numbers is effectively enumerable if and only if it is recursively enumerable.

A set of natural numbers is *decidable* if and only if there is an algorithm for testing membership in the set. A set can be effectively enumerable without being decidable, since, if we have a procedure for listing an infinite set, there will be no stage at which, from the fact that a given number has not yet turned up on the list, we can conclude that the number will never appear on the list. On the other hand if a set and its complement are both effectively enumerable then the set is decidable, and conversely. Defining a set to be *recursive* if it and its complement are both recursively enumerable, the Church-Turing thesis tells us that a set is decidable if and only if it is recursive.

An *unary partial function* is a set of ordered pairs f with the property that, whenever $<i,j>$ and $<i,k>$ are both in f, we have $j = k$. If $<i,j> \in f$, for some j, we say that i is in the *domain* of f, and we write $f(i) = j$. (Partial functions of more than one variable are defined similarly.) f is said to be *calculable* if there is an algorithm that, for given input i, gives the output $f(i)$ if i is in the domain of f, and yields no output at all if i is outside the domain of f. A unary partial function is calculable if and only if, *qua* binary relation, it is effectively enumerable. It follows according to the Church-Turing thesis that f is calculable if and only if it is recursively enumerable. If so, f is said to be a *partial recursive function*. (The notation is confusing—a collection of ordered pairs can be a partial recursive function without being a recursive relation—but entrenched.) A *total recursive function*—a partial recursive function whose domain is all of $\mathbb{N}$—will be a recursive relation, since if f is $\{<i,j>: (\exists x)\theta(\underline{i},\underline{j},x)\}$, with θ bounded, the complement of f is $\{<i,j>:(\exists x)(\exists y)(\sim y = \underline{j} \wedge \theta(\underline{i},y,x))\}$.

ARITHMETIZATION OF METAMATHEMATICS

The set-theoretic paradoxes, particularly Russell's paradox, had on David Hilbert much the same effect that Zeno's paradoxes had on Aristotle. Both thinkers came to realize that the idea of the infinite held great intellectual peril with the risk of contradiction at every turn. Unlike Aristotle, however, Hilbert was unwilling to banish the actual infinite from mathematical reasoning. Instead he proposed to develop the theory of infinite sets in such a way that we could be assured that no contradiction would ensue, by treating mathematical proofs as the objects of mathematical study, in the same way that earlier mathematicians had treated curves, planes, and numbers as objects of mathematical study. A mathematical proof is, after all, a finite object, even if the sentences that appear in the proof talk about infinite objects, and Hilbert proposed that a new science of *metamathematics* could show by finite means that set theory was free of contradiction, by showing that there is no finite path that leads from the axioms to "~ 0=0."

The great breakthrough in metamathematics was Gödel's proof, which showed that it was not necessary to go outside set theory or even outside arithmetic to carry out metamathematical investigations. By assigning numerical codes to formulas and finite strings of formulas, and by reducing properties of proofs to properties of their code numbers, it was possible to develop proof theory as a branch of number theory. This technique led to a great flowering of metamathematics even though as we shall see, it derailed Hilbert's plan.

The arithmetization of metamathematics proceeded in two stages. In the first stage numerical codes are assigned to simple symbols more-or-less arbitrarily, so that a formula, which is a string of simple symbols, can be coded as a sequence of numbers. Second we devise a method for encoding a finite sequence of numbers as a single number. This enables us to encode a formula as a single number. In this way a proof, which is a sequence of formulas, is encoded as a sequence of numbers, which is, in turn, coded as a single number.

We attack the second stage first. We already know how to use the function *Pair* to code a pair of natural numbers by a single number. We can encode a finite set of natural numbers by a single number by setting the code number of the finite set F, $Code(F)$, equal to $\sum_{i \in F} (2Ei)$ *Code* provides a one-one correspondence between the set of finite sets of natural numbers and $\mathbb{N}$. The number n is the image under *Code* of the set of places in the binary decimal expansion of n in which "1"s appear. Finally, we encode the finite sequence $<k_0,k_1, \ldots ,k_m>$ as the number $Code(\{Pair(0,k_0), Pair(1,k_1), \ldots , Pair(m,k_m)\})$. Here we shall use an expression like "$<3,2,1>$" ambiguously to denote a sequence of length three and to denote the code number for that sequence, which is 448.

The relation that holds between k and n if k is an element of the set coded by n is defined by a bounded formula; abusing notation, we write "$\underline{k} \in \underline{n}$" to represent the statement that $(\exists i < (\underline{2}E\underline{k}))(\exists j < \underline{n})\underline{n} = (i + ((\underline{2}\ E\ \underline{k}) + (j \times (\underline{2}E(S\underline{k})))))$. The set of all code numbers of finite sequences is the extension of a bounded formula, as are the concatenation operation and the partial function that takes i and n to the ith member of the sequence coded by n (provided n codes a sequence of i or more elements). The simplicity of this technique for encoding a finite sequence of numbers by a single number is the motive for including exponentiation as a primitive operation.

The details of the assignment of numerical codes to terms and formulas are highly arbitrary. A motive for the particular choices here is to avoid fretting over parentheses. With each term τ, we associate a number $\ulcorner\tau\urcorner$, as follows: The numeral "0" is assigned $<0,0>$, and the variable x_i is assigned $<1,i>$. $\ulcorner S\tau\urcorner$ is $<2,\ulcorner\tau\urcorner>$, and $\ulcorner(\tau+\rho)\urcorner$, $\ulcorner(\tau\times\rho)\urcorner$, and $\ulcorner(\tau E\rho z)\urcorner$ are $<3,\ulcorner\tau\urcorner,\ulcorner\rho\urcorner>$, $<4,\ulcorner\tau\urcorner,\ulcorner\rho\urcorner>$, and $<5,\ulcorner\tau\urcorner,\ulcorner\rho\urcorner>$, respectively.

A number x is a the code of a term just in case it is an element of a finite set s with the following property: For any element y of s, either $y = <0,0>$; or $y = <1,i>$, for some $i \leq y$; or $y = <2,z>$, for some z in s; or y is equal to one of $<3,z,w>$, $<4,z,w>$, and $<5,z,w>$, for some z and w in s. s represents a finite tree, with each node labeled by the code of a term, so that when a node is labeled by a complex term, nodes beneath it are labeled by the term's constituents and so that each leaf of the tree is labeled either by the code of "0" or by the code for a variable. This characterization is naturally written out as a Σ formula, showing that the set of (code numbers of) terms is recursively enumerable.

The set of terms is, in fact, recursive. To see this, we note that, if x is not a term, then the attempt to construct a labeled tree with x at its trunk winds up with at least one branch that does not terminate in either $\ulcorner 0\urcorner$ or a variable. More precisely, x does not encode a term if and only if there is a sequence $<x_0,x_1,\ldots,x_n>$ of numbers $\leq x$ with the following properties:

$x_0 = x$.

If x_i has the form $<2,y>$, then $i < n$ and $x_{i+1} = y$.

If x_i has one of the forms $<3,y,z>$, $<4,y,z>$, or $<5,y,z>$, then $i < n$ and either $x_{i+1} = y$ or $x_{i+1} = z$.

If $i < n$, x_i has one of the forms $<2,y>$, $<3,y,z>$, $<4,y,z>$, or $<5,y,z>$.

x_n does not have either of the forms $<0,0>$ or $<1,k>$.

This can readily be written out as a Σ formula, showing that the complement of the set of terms is recursively enumerable.

The function Z that takes a number n to the code number for the numeral $\underline{n}$ can be described by a recursive definition:

$$Z(0) = <0,0> = 5.$$
$$Z(m+1) = <2,Z(m)> = 8 + (2E(Pair(1,Z(m)))).$$

We can convert this recursive definition into an explicit definition, using a quite general technique that Gödel obtained by refining an idea from Gottlob Frege's *Begriffschrift*. $Z(n) = k$ if and only if there is a sequence $<x_0,x_1,\ldots,x_n>$ with the following features:

$x_0 = <0,0>$.

For $m < n$, $x_{m+1} = <2,x_m>$.

$x_n = k$.

This characterization shows that Z is a total recursive function.

The function that associates a code $\ulcorner\phi\urcorner$ with each formula ϕ is again highly arbitrary. For τ and ρ terms, we let $<6,\ulcorner\tau\urcorner,\ulcorner\rho\urcorner>$ and $<7,\ulcorner\tau\urcorner,\ulcorner\rho\urcorner>$ be the codes of $\tau = \rho$ and $\tau \leq \rho$. For ϕ and ψ formulas, we let $<8,\ulcorner\phi\urcorner>$ be $\ulcorner\sim \phi\urcorner$, $<9,\ulcorner\phi\urcorner,\ulcorner\psi\urcorner>$ be $\ulcorner(\phi \vee \psi)\urcorner$, and $<10,i,\ulcorner\phi\urcorner>$ be $\ulcorner(\exists v_i)\psi\urcorner$. The proof that the set of codes of formulas is recursive is just like the corresponding argument for terms.

It is straightforward if somewhat laborious to verify, just by writing down an appropriate formula, that, for example, the arithmetical operations corresponding to forming the disjunction and the conjunction of two formulas, to prefixing a quantifier to a formula, and to substituting a given term for free occurrences of a variable in a formula are partial recursive functions. Also, for example, that the set of terms in which the variable v_{17} appears

and the set of formulas in which v_{123} appears free are recursive sets.

PROOFS AND COMPUTATIONS

Euclid's *Elements* deduces highly sophisticated geometric theorems as consequences of simple, intuitively obvious axioms. Aristotle, the father of logic, investigated the methods by which consequences are derived from axioms, identifying simple patterns of valid reasoning like the following so-called *syllogism*: "All men are animals. No stone is an animal. Therefore, no stone is a man." The methods of reasoning Euclid actually employed were far more sophisticated than the mere production of chains of syllogisms, however, and the ancients were generally content to take it as obvious that Euclid's deductions were legitimate, without demanding a detailed survey of deductive methods.

Meticulous nineteenth-century investigations revealed the surprising fact that, despite having been accepted by generations of scholars as the exemplar of deductive rigor, Euclid's proofs were often invalid. In proving a theorem he sometimes imported information from the accompanying diagram that was not justified by either the hypotheses of the theorem or the axioms. These investigations led to a search for fully precise methods of deduction in which one could have complete confidence. This search culminated in the widespread acceptance of a system of precise rules for the first-order predicate calculus—the logic governing the operators "$\vee$," "$\sim$," $(\exists v_i)$, and "$=$"—within which the deductions of classical mathematics can be formalized with scrupulous rigor.

With these rules in hand, we can capture the notion of logical consequence precisely, by pressing it from below and from above. It is clear that, if a sentence ϕ is a logical consequence of a theory (set of sentences) Γ, then it cannot be possible to choose a domain of discourse and semantic values for the nonlogical terms so as to make the members of Γ all true and ϕ false. Thus a necessary condition for a ϕ to be a logical consequence of Γ is that ϕ be true in every model of Γ. It is also clear, from examining the rules (for whichever of the standard textbook systems is convenient), that if ϕ derivable from Γ, ϕ is a logical consequence of Γ; this gives us a sufficient condition for logical consequence. Gödel's 1930 *Completeness Theorem* shows that these two conditions meet, so that if ϕ is true in every model of Γ then ϕ is derivable from Γ.

The Completeness Theorem applies equally well to any of many different logical calculi for first-order predicate logic. W.V. Quine developed a particularly convenient system with the following two properties: The (codes of the) axioms of logic form a recursive set; and each logical consequence of a theory Γ can be found at the end of a sequence of sentences, each member of which is either an axiom of logic, an element of Γ, or obtained from earlier members of the sequence by *modus ponens*, the rule that permits the deduction of ψ from ϕ and $(\phi \rightarrow \psi)$. (Such a sequence is a *proof* of the sentence from Γ.) Quine's axioms will not be written out here.

If Γ is recursive, the set of pairs $<s, \ulcorner\phi\urcorner>$ such that s is a proof of ϕ from Γ is a recursive relation, represented by a Σ formula we shall abbreviate "$\underline{s}\ B_\Gamma \underline{\ulcorner\phi\urcorner}$." (In terminology introduced below, "B_Γ" "binumerates" the relation.) We write "$Bew_\Gamma(\underline{\ulcorner\phi\urcorner})$" to abbreviate "$(\exists s)s\ B_\Gamma\underline{\ulcorner\phi\urcorner}$." Since "$Bew_\Gamma$," is Σ, the set of logical consequences of Γ is recursively enumerable.

William Craig noted a converse result: If the set of consequences of the theory Γ is recursively enumerable then Γ has the same consequences as some recursive set to axioms; Γ is, as they say, *recursively axiomatizable*. To see this, note that there is a bounded formula $\psi(x,y)$ such that the consequences of Γ constitute the set of sentences whose code numbers satisfy $(\exists y)\psi(x,y)$. Let Γ_{Craig} be the set of all sentences of the form $(\underline{m} = \underline{m} \wedge \theta)$, for which the pair $<\ulcorner\theta\urcorner,m>$ satisfies $\psi(x,y)$. Then Γ_{Craig} is recursive (bounded, in fact), and Γ_{Craig} and Γ are logically equivalent.

We would now like to see how any numerical computation by algorithm can be simulated by a logical deduction from basic arithmetical axioms. Q_E, a variant of *Robinson's arithmetic*, is the conjunction of the following nine statements:

$(\forall x)(x = 0 \leftrightarrow \sim (\exists y)x = Sy).$

$(\forall x)(\forall y)(Sx = Sy \rightarrow x = y)$

$(\forall x)(x + 0) = x.$

$(\forall x)(\forall y)(x + Sy) = S(x + y).$

$(\forall x)(x \times 0) = 0.$

$(\forall x)(\forall y)(x \times Sy) = ((x \times y) + x)$

$(\forall x)(x \text{ E } 0) = S0.$

$(\forall x)(\forall y)(x \text{ E } Sy) = ((x \text{ E } y) \times x)$

$(\forall x)(\forall y)(x \leq y \leftrightarrow (\exists z)(x + z) = y).$

Q, which we shall talk about later on, is obtained from Q_E by deleting the two clauses involving exponentiation.

A straightforward induction on the complexity of terms shows that, for every closed term τ, there is a number m such that the sentence $\tau = \underline{m}$ is a theorem of Q_E.

Another induction shows that every bounded sentence is decidable (either provable or refutable) in Q_E. Since every true bounded sentence is provable in Q_E, it follows that every true Σ sentence is provable in Q_E, since we can prove an existential sentence by providing a witness. If *S* is a recursively enumerable set, it is the extension of some Σ formula σ. Because every true Σ sentence is provable in Q_E and (because Q_E is true) no false Σ sentence is provable, we have (where "$\vdash$" is provability):

For any n, $n \in S$ if and only if $Q_E \vdash \sigma(\underline{n})$.

We shall say that σ *enumerates* *S* in Q_E. (The same observation holds for recursively enumerable relations.)

We shall say a formula ϕ *binumerates* a set *S* in Q_E if and only if, for each *n*, we have:

$n \in S$ if and only if $Q_E \vdash \phi(\underline{n})$.

$n \notin S$ if and only if $Q_E \vdash \sim \phi(\underline{n})$.

If *S* is recursive then there is a bounded formula $\chi(x,y)$ such that $(\exists y)\chi(x,y)$ enumerates *S* in Q_E, and there is a bounded formula $\theta(x,y)$ such that $(\exists y)\theta(x,y)$ enumerates the complement of *S* in Q_E. To show that *S* is binumerable in Q_E, we need to show that *S* is enumerable by a formula whose negation enumerates the complement of *S*. Developing an idea of J. Barclay Rosser, Tarski, Mostowski, and Robinson showed that the following Σ formula does the job:

$$(\exists y)(\chi(x,y) \wedge \sim (\exists z < y)\theta(x,z)).$$

Clearly if ϕ binumerates *S* in Q_E, it binumerates *S* in any consistent theory that entails Q_E.

A formula $\psi(x,y)$ *functionally represents* a total function *f* in a theory if and only if, for each *k*, the following sentence is a consequence of the theory:

$$(\forall y)(\psi(\underline{k},y) \leftrightarrow y = \underline{f(k)}).$$

If *f* is a total recursive function, we know that there is a formula $\phi(x,y)$ that binumerates *f* in Q_E. Tarski, Mostowski, and Robinson showed that the following formula functionally represents *f* in Q_E (and hence in any theory that entails Q_E):

$$(\phi(x,y) \wedge (\forall z < y) \sim \phi(x,z)).$$

THE FIRST INCOMPLETENESS THEOREM

We are now ready to see how to construct, for any recursively axiomatizable, true theory that includes Q_E, a true sentence that is not a consequence of the theory. The key to the construction is to see how to produce sentences that can talk about themselves so that we can construct a sentence that asserts its own unprovability. Such a sentence cannot be provable since if it were provable it would be a false consequence of the axioms. So the sentence must be true. To carry out this plan we use the following result, one of the masterpieces of modern mathematics:

GÖDEL'S SELF-REFERENCE LEMMA. For any formula $\psi(y)$, one can construct a sentence ϕ such that $Q_E \vdash (\phi \leftrightarrow \psi(\underline{\ulcorner\phi\urcorner}))$.

The hard part, the part that requires true genius, is to figure out what sentence to write down. The easy part is to verify that the sentence works. Here we shall only attempt the easy part.

Define a function *f* as follows: If *m* is the code of a formula $\chi(x,y)$ with only "*x*" and "*y*" free, let $f(m)$ be the code of the formula

$$(\exists x)(\exists y)((x = \underline{m} \wedge \chi(x,y)) \wedge \psi(y)).$$

Otherwise, $f(m) = 0$.

This definition can easily be written as a Σ formula, showing that *f* is a total recursive function. Consequently, there is a formula $\theta(x,y)$ that functionally represents *f* in Q_E. Let *m* be $\ulcorner(\theta(x,y)\urcorner$, and ϕ be the following sentence:

$$(\exists x)(\exists y)((x = \underline{m} \wedge \theta(x,y)) \wedge \psi(y)).$$

Then $\ulcorner\phi\urcorner = f(m)$, and so the following sentences are consequences of Q_E:

$(\forall y)(\theta(\underline{m},y) \leftrightarrow y = \underline{\ulcorner\phi\urcorner})$.

$((\exists x)(\exists y)((x = \underline{m} \wedge \theta(x,y)) \wedge \psi(y)) \leftrightarrow \psi(\underline{\ulcorner\phi\urcorner}))$.

$(\phi \leftrightarrow \psi(\underline{\ulcorner\phi\urcorner}))$.

Let Γ be a consistent, recursive set of sentences that entails Q_E. Using the Self-reference Lemma, we can find a sentence γ so that $(\gamma \leftrightarrow \sim Bew_\Gamma \underline{\ulcorner\gamma\urcorner})$ is a consequence of Q_E; γ is called the *Gödel sentence* for Γ. If γ were a consequence of Γ, $\sim Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})$ would be a consequence of Γ, and also $Bew_\Gamma(\ulcorner\gamma\urcorner)$ would be a true Σ sentence, hence a consequence of Q_E, hence a consequence of Γ. This contradicts the consistency of Γ. So γ is unprovable, so that $Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})$ is false, and γ is true. Thus γ is our example of a true, unprovable sentences.

If Γ is true then Γ does not prove $\sim\gamma$ because $\sim\gamma$ is false, so that γ is undecidable in Γ. Let us say that a theory Δ is *ω-inconsistent* if there is a formula $\chi(x)$ such that $(\exists x)\chi(x)$ is a consequence of Δ, and yet, for each n, $\sim \chi(\underline{n})$ is a consequence of Δ. Every ω-consistent theory is consistent, so if Δ is a recursive, ω-consistent theory that entails Q_E, the Gödel sentence γ for Δ is a true sentence not provable in Δ. Hence, for each *m*, the sentence

$\sim \underline{m}\ B_{\Delta} \ulcorner\underline{\gamma}\urcorner$ is true, hence provable in Q_E, hence provable in Δ. It follows by ω-consistency that $Bew_{\Delta} \ulcorner\underline{\gamma}\urcorner$ is not a consequence of Δ, and so ~γ is not a consequence of Δ. Thus the assumption of ω-consistency, rather than truth, is enough to ensure that γ is undecidable in Δ. Because γ is unprovable in Δ, $\Delta \cup \{\sim\gamma\}$ is consistent, although ω-inconsistent. So consistency does not imply ω-consistency.

Gödel used γ to show that every ω-consistent, recursively axiomatizable theory that entails Q_E is incomplete, that is, that there are sentences that the theory cannot decide; this is the *First Incompleteness Theorem*. Rosser went a step farther, showing that the assumption of ω-consistency can be weakened to consistency. Rather than examine Rosser's proof, we shall derive his conclusion from a stronger result, one due, in essentials, to Tarski, Mostowski, and Robinson:

RECURSIVE INSEPARABILITY THEOREM. There is no recursive set that includes the consequences of Q_E and excludes all the sentences refutable in Q_E.

Suppose *C* were such a recursive set, and take a formula μ(*x*) that binumerates *C* in Q_E. The Self-reference Lemma gives a sentence ν such that $(\nu \leftrightarrow \sim \mu(\ulcorner\underline{\nu}\urcorner))$ is a consequence of Q_E. We derive a contradiction by examining two cases:

Case 1. $\nu \in C$. Then $Q_E \vdash \mu(\ulcorner\underline{\nu}\urcorner)$, and so $Q_E \vdash \sim \nu$. Thus υ is a sentence refutable in Q_E, and so it is excluded from C. Contradiction.

Case 2. $\upsilon \notin C$. Then $Q_E \vdash \sim \mu(\ulcorner\underline{\nu}\urcorner)$, and so $Q_E \vdash \nu$. Thus υ is a consequence of Q_E, and so an element of *C*. Contradiction.

Corollary. No consistent theory that entails Q_E has a recursive set of consequences.

This follows from the fact that, if a consistent theory entails Q_E, it excludes the sentences refutable in Q_E.

Corollary **(*Rosser's Theorem*).** No consistent, recursively axiomatized theory that entails Q_E is complete.

If Γ is consistent, recursively axiomatized, and complete, then the complement of Γ is recursively enumerable, since it is the union of the set of non-sentences with the set of sentences whose negations are provable in Γ.

Corollary. No theory consistent with Q_E has a recursive set of consequences.

If Δ were such a theory then the set of sentences ψ such that $(Q_E \rightarrow \psi)$ is a consequence of Δ would be a consistent, recursive set of sentences, closed under consequence, that included Q_E.

Corollary **(*Church's Theorem*).** The set of logically valid sentences in not recursive.

The valid sentences are the consequences of the empty theory, which is consistent with Q_E.

MATHEMATICAL INDUCTION

Q_E is a weak axiom system. It cannot prove the associative law of addition or multiplication, nor can it prove the commutative law of addition or multiplication. The system is weak because it leaves out the essential feature of the natural number system, the principle of mathematical induction, according to which any collection of natural numbers that includes 0 and is closed under the successor operation has to include all the natural numbers. *Modulo* Q_E, the principle is equivalent to the thesis that the natural numbers are well-founded, that is, that any nonempty collection of natural numbers has a least element.

Richard Dedekind showed that the system one gets from Q_E by adding the principle of mathematical induction completely characterizes the natural numbers. The system is *categorical*, that is, there is an isomorphism—a one-one correspondence that preserves mathematical structure—between any two models of the system. Thus if $\mathfrak{A}$ and $\mathfrak{B}$ are models of Q_E plus the principle of induction, let *f* be the smallest class that includes the pair $<0^{\mathfrak{A}}, 0^{\mathfrak{B}}>$ and includes $<S^{\mathfrak{A}}(x), S^{\mathfrak{B}}(y)>$ whenever it contains $<x,y>$. It is easy to verify, using induction several times, that *f* is an isomorphism. It follows that the system is complete, since if it left ϕ undecided, it would have a model $\mathfrak{A}$ in which ϕ is true and a model $\mathfrak{B}$ in which ϕ is false; but then $\mathfrak{A}$ and $\mathfrak{B}$ could not be isomorphic.

Peano Arithmetic (PA), is the system used to formalize the principle of induction into a precise system of axioms. Its axioms are Q_E together with all instances of the *induction axiom schema*:

$$((R(0) \wedge (\forall x)(R(x) \rightarrow R(Sx))) \rightarrow (\forall x)R(x)).$$

An *induction axiom* is a sentence of the language of arithmetic obtained from the schema by substituting a formula of the language of arithmetic for "*R*," then prefixing universal quantifiers to bind all the variables other than "*x*" that appear free in the substituted formula.

In view of Dedekind's categoricity theorem, it is surprising to realize that PA is incomplete. But incomplete it must be, since it is a true, recursively axiomatized theory that entails Q_E. The explanation is that the induction axiom schema does not fully capture the principle of

mathematical induction. It tries to assure us that every nonempty collection has a least element, but only succeeds in telling us that every nonempty collection that is the extension of a predicate (with parameters) of the language of arithmetic has a least element.

Let γ be the Gödel sentence for PA. We know that γ isn't a consequence of PA, so that, by the Completeness Theorem, there is a model $\mathfrak{A}$ in which all the axioms of PA + $\sim\gamma$ are true. In $\mathfrak{A}$ there is an element g that satisfies "$x\mathrm{B}_{PA}$ $\underline{\ulcorner\gamma\urcorner}$." For each n, "$\sim\underline{n}\mathrm{B}_{PA}$ $\underline{\ulcorner\gamma\urcorner}$." is a theorem of PA, so g must be different from the referents of all the numerals $\underline{0}$, $\underline{1}$, $\underline{2}$, …. Instead, g is one of the *nonstandard numbers* that lie above all the standard numbers in the relation $\mathfrak{A}$ assigns to "$\leq$."

It is worth emphasizing because there has been some confusion on this score that the existence of nonstandard models of PA does not depend on the First Incompleteness Theorem. Their existence follows from the Compactness Theorem, according to which an infinite set of sentences has a model if every finite subset does, which Gödel derived from the Completeness Theorem. Let Γ be a consistent theory that entails Q_E. Add a new constant "c" to the language, and let Γ^c be the union of Γ with the set of sentences "$\sim c = \underline{n}$," for n natural number. Any finite subset of Γ^c has a model, obtained by taking a model of Γ and letting "c" denote a sufficiently large standard number. The Compactness Theorem gives us a model of Γ^c, which means we have a nonstandard model of Γ. This construction works even if we take Γ to be *true arithmetic*, the set of sentences true in the standard model, even though true arithmetic is complete. Because it is complete, the First Incompleteness Theorem tells us that true arithmetic is not recursively axiomatizable.

THE SECOND INCOMPLETENESS THEOREM

The proof of the First Incompleteness Theorem showed that, if Γ, a recursively axiomatized theory that entails Q_E, is consistent, then the Gödel sentence γ for Γ is unprovable in Γ. Using "$Con(\Gamma)$" as an abbreviation for

$$\text{“}\sim Bew_\Gamma(\underline{\ulcorner\sim \underline{0}=\underline{0}\urcorner}),\text{”}$$

we can formalize this result in a sentence of the language of arithmetic:

$$(Con(\Gamma) \rightarrow \sim Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})).$$

If we were able to prove this conditional in Γ, we could conclude that, if $Con(\Gamma)$ were provable in Γ, $\sim Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})$ would be provable in Γ. Since we already know that $\sim Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})$ is only provable in Γ if Γ is inconsistent, we could conclude that $Con(\Gamma)$ is only provable in Γ if Γ is inconsistent.

Can we prove the conditional in Γ? We certainly cannot do so if we take Γ to be Q_E, for we can scarcely prove any significant generalizations in Q_E. We can, however, prove the conditional if we take Γ to be PA. This is hardly surprising, since nearly all our reasoning about natural numbers can be formalized in PA. The details are, nonetheless, burdensome; so we only present a faint sketch here.

Let Γ be a recursively axiomatized theory that entails PA. M. H. Löb singled out the following three principles as central to Gödel's proof that, if Γ is consistent, it does not prove $Con(\Gamma)$:

(L1) If $\Gamma \vdash \phi$, then $\Gamma \vdash Bew_\Gamma(\underline{\ulcorner\phi\urcorner})$.

(L2) $\Gamma \vdash (Bew_\Gamma(\underline{\ulcorner\phi\urcorner}) \rightarrow Bew_\Gamma(\underline{\ulcorner Bew_\Gamma(\underline{\ulcorner\phi\urcorner})\urcorner}))$.

(L3) $\Gamma \vdash (Bew_\Gamma(\underline{\ulcorner(\phi \rightarrow \Psi)\urcorner}) \rightarrow (Bew_\Gamma(\underline{\ulcorner\phi\urcorner}) \rightarrow Bew_\Gamma(\underline{\ulcorner\Psi\urcorner}))$.

We have already seen why (L1) has to hold. If ϕ is a consequence of Γ, $Bew_\Gamma(\underline{\ulcorner\phi\urcorner})$ is a true Σ sentence, hence provable in Q_E, hence provable in Γ. (L2) is obtained, laboriously, by formalizing the proof of (L1). In fact, Γ proves $(\theta \rightarrow Bew_\Gamma(\underline{\ulcorner\theta\urcorner}))$, for each Σ sentence θ. (L3) is easy. If we have proofs of $(\phi \rightarrow \psi)$ and ϕ, we get a proof of ψ by concatenating the two proofs and tacking ψ on the end.

Given the Löb conditions, the proof of the *Second Incompleteness Theorem*, according to which, if Γ is a consistent, recursively axiomatized theory that entails PA, then Γ does not prove its own consistency, is straightforward. Let γ be the Gödel sentence for Γ. Because of the way γ was constructed, we have:

$$\Gamma \vdash (\gamma \rightarrow \sim Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})),$$

which is logically equivalent to:

$$\Gamma \vdash (\gamma \rightarrow (Bew_\Gamma(\underline{\ulcorner\gamma\urcorner}) \rightarrow \sim 0=0)).$$

One application of (L1) and two applications of (L3) give us this:

$$\Gamma \vdash (Bew_\Gamma(\underline{\ulcorner\gamma\urcorner}) \rightarrow (Bew_\Gamma(\underline{\ulcorner Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})\urcorner}) \rightarrow Bew_\Gamma(\underline{\ulcorner\sim \underline{0}=\underline{0}\urcorner}))).$$

(L2) gives us this:

$$\Gamma \vdash (Bew_\Gamma(\underline{\ulcorner\gamma\urcorner}) \rightarrow Bew_\Gamma(\underline{\ulcorner Bew_\Gamma(\underline{\ulcorner\gamma\urcorner})\urcorner})),$$

and these two results together give us:

$$\Gamma \vdash (Bew_\Gamma(\underline{\ulcorner\gamma\urcorner}) \rightarrow Bew_\Gamma(\underline{\ulcorner\sim \underline{0}=\underline{0}\urcorner})).$$

By contraposition,

$$\Gamma \vdash (\sim Bew_\Gamma \underline{\ulcorner\sim \underline{0}=\underline{0}\urcorner}) \rightarrow \sim Bew_\Gamma\ (\underline{\ulcorner\gamma\urcorner})),$$

that is,

$$\Gamma \vdash (Con(\Gamma) \rightarrow \sim Bew_\Gamma\ (\underline{\ulcorner\gamma\urcorner}))$$

Now assume

$$\Gamma \vdash Con(\Gamma).$$

Then

$$\Gamma \vdash \sim Bew_{\Gamma} (\ulcorner \underline{\gamma} \urcorner).$$

By the way γ was constructed,

$$\Gamma \vdash \gamma.$$

Hence, by (L1),

$$\Gamma \vdash Bew_{\Gamma} (\ulcorner \underline{\gamma} \urcorner),$$

and so Γ is inconsistent.

In accepting PA, we recognize that the axioms of PA are all true. If the axioms are all true then the theory is certainly consistent, and if the theory is consistent its Gödel sentence is true. So we have good reason to accept the Gödel sentence for PA, even though it is not a consequence of PA. If in this argument we replace PA with our total arithmetical theory—the (admittedly, vaguely defined) totality of arithmetical sentences we are willing to accept as true—we seem to get the curious result that, assuming that our total theory is recursively enumerable, we accept the Gödel sentence for our total theory even though it is not a consequence of the theory. But this contradicts the characterization of our total theory.

J. R. Lucas (1961) and Roger Penrose (1989) took this puzzling situation as reason to believe that the cognitive processes of the human mind cannot be simulated by any purely mechanical device, and that this conclusion undermines the prospects for a naturalistic conception of mind, according to which the human mind is a product of the orderly operation of the laws of nature, not in principle any more mysterious or less constrained by physical law than a player piano or a personal computer. Adherents to the computational theory of mind hold that the operations of the mind are usefully understood on the model of a sophisticated electronic computer, and even naturalists who are not advocates of the computational model will be inclined to say that the facts that the human body is produced by natural selection rather than conscious design and that its central processing unit is carbon-based rather than silicon-based will not affect its capabilities in any fundamental way, so that, according to a naturalistic conception, the cognitive activities of a human being can, in principle, be simulated by a purely mechanical device.

The connection between mechanism and recursive enumerability is given by a variant of the Church-Turing Thesis, supported by similar evidence, that declares that the set of numbers accepted by a mechanical input-output device is invariably recursively enumerable. This includes nondeterministic machines, whose operation is to some extent a matter of random chance, so that the set *S* is accepted by the machine just in case, for any *n*, *n* is in *S* if and only if there is some possible computation of the machine on input *n* that yields a positive outcome, as well as deterministic machines for which the course of a computation is uniquely determined by its input.

The argument that our total arithmetical theory is not recursively enumerable proceeds by *reductio ad absurdum*. If the theory were recursively enumerable, it would be recursively axiomatizable, so it would have a Gödel sentence. But we can see that the Gödel sentence is true, even though it is not part of the total theory.

The Lucas-Penrose argument is vulnerable to two criticisms. First, for naturalism to be correct, there has to exist a recursive axiomatization of our total theory. In order to construct the Gödel sentence, we have to be able to specify a recursive axiomatization by writing down a formula that binumerates it. However it is perfectly possible for a recursive axiomatization to exist without our being able to specify it.

Second, even if we were able to specify a recursive axiomatization, perhaps by analyzing a futuristic brain scan, it is hard to see how we could be justified in being completely confident that our total theory is consistent. If we decide to be strict about what arithmetical sentences we are willing to count as "accepted," so that we only regard a sentence as part of our total theory if we arrive at it by unimpeachably lucid reasoning, we shall increase our confidence that our total theory is consistent, but raising the bar this way will also heighten the hurdle that the Gödel sentence has to pass in order to count as "accepted." There are different standards we might use for when we are willing to count a sentence as proven, and each standard has a different Gödel sentence, but however high we set the standard the Gödel sentence corresponding to that standard cannot pass it, on pain of inconsistency.

THE LOGIC OF PROVABILTY

If we explicitly embrace a theory Γ, so that we are willing consciously to acknowledge that the axioms of Γ are all statements we regard as true then we surely ought to regard Γ as consistent. Yet (assuming that Γ implies PA and is recursively axiomatizable and consistent) the statement that Γ is consistent is not provable in Γ. Thus the arithmetical statements that we commit ourselves to in embracing Γ go beyond what Γ itself entails.

The disparity between what consciously accepting Γ commits us to and what Γ entails is even wider than the Second Incompleteness Theorem indicates. Accepting Γ means acknowledging that all the consequences of Γ are true. For a given sentence φ, we may not know whether φ is a consequence of Γ—there is after all no algorithm to tell us—but at least we accept that, if φ is a consequence of Γ, φ is true. Consciously accepting Γ commits us to the conditionals $Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner) \rightarrow \phi)$, but they are not in general consequences of Γ. In fact such a conditional is a consequence of Γ only if its consequent is a consequence of Γ.

> **Löb's Theorem.** Let Γ be a recursively axiomatized theory that entails PA. If $Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner) \rightarrow \phi)$ is a consequence of Γ, so is φ.

We can regard the Second Incompleteness Theorem as the special case of Löb's Theorem in which φ is taken to be the sentence "~ 0=0." Conversely we can derive Löb's Theorem from the Second Incompleteness Theorem. The argument, which is due to Saul Kripke, utilizes the observation that, for any ψ and θ, $\Gamma \vdash (\psi \rightarrow \theta)$ if and only if $\Gamma \cup \{\psi\} \vdash \theta$, and the fact that this observation is provable in PA.

Suppose that φ is not a consequence of Γ. Then $\Gamma \cup \{\sim \phi\}$ is consistent, which implies, by the Second Incompleteness Theorem, that $Con(\Gamma \cup \{\sim \phi\})$ is not a consequence of $\Gamma \cup \{\sim \phi\}$. Thus we have:

$$\Gamma \cup \{\sim \phi\} \vdash \sim Bew_{\Gamma \cup \{\sim\phi\}}(\ulcorner\underline{\sim 0=0}\urcorner)$$

$$\Gamma \cup \{\sim \phi\} \vdash \sim Bew_{\Gamma}(\ulcorner\underline{(\sim \phi \rightarrow \sim 0=0}\urcorner)$$

$$\Gamma \cup \{\sim \phi\} \vdash \sim Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner)$$

$$\Gamma \vdash (\sim \phi \rightarrow \sim Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner))$$

$$\Gamma \vdash Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner) \rightarrow \phi)$$

Conditionals of the form $Bew_{\Gamma}(\ulcorner\underline{\phi}\urcorner) \rightarrow \phi)$ are called *reflection principles*. We cannot obtain them by working within Γ. We get them from the outside by reflecting on the fact that Γ is a theory we accept.

We can describe the logic of provability precisely by utilizing the methods of modal logic. Modal sentential calculus has, in addition to formulas built up from atomic formulas by the familiar connectives "∨" and "~," a new connective "□." "□φ," usually read "It is necessary that φ," is here understood to mean, "It is provable in Γ that φ," where Γ is a consistent, recursively axiomatizable theory that implies PA. An *interpretation* of the modal sentential calculus is a function *i* that associates an arithmetical sentence with each modal formula, subject to the conditions that $i(\phi \vee \psi)$ be equal to $(i(\phi) \vee i(\psi))$, $i(\sim\phi)$ be equal to $\sim i(\phi)$, and $i(\Box\phi)$ be equal to $Bew_{\Gamma}\ulcorner\underline{i(\phi)}\urcorner$. A modal formula φ is *always provable* if, for each interpretation *i*, $i(\phi)$ is provable in Γ. φ is *always true* if, for each φ, $i(\phi)$ is true.

(L1) tells us, if $i(P)$ is provable, $i(\Box P)$ is provable, so that the set of always-provable formulas is closed under necessitation, the rule of modal logic that infers □θ from θ. (L2) tells us that $(\Box P \rightarrow \Box\Box P)$ is always true, and the formalization of (L2) tells us that it is always provable. (L3) tells us that $(\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q))$ is always true; it is easily seen to be always provable as well. Löb's Theorem tells us that whenever $i(\Box P \rightarrow P)$ is a theorem, $i(P)$ is a theorem. Formalizing his proof, we see that the formula $(\Box(\Box P \rightarrow P) \rightarrow \Box P)$ is always provable and always true.

Robert Solovay deployed an ingenious application of the Self-referential Lemma within the possible-world semantics for modal logic to show that, provided Γ does not prove any false Σ sentences, a formula is always provable if and only if it is derivable by *modus ponens* and necessitation from sentential-calculus tautologies (formulas that are assigned the value "true" by every function assigning truth-values to formulas that respects the meanings of "∨" and "~") and instances of the following schemata:

$$(\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi))$$

$$(\Box(\Box\phi \rightarrow \phi) \rightarrow \Box\phi)$$

Assuming Γ is true, a formula is always true if and only if it is derivable by *modus ponens* from always-provable formulas and instances of the reflection principle $(\Box\phi \rightarrow \phi)$.

BEYOND THE LANGUAGE OF ARITHMETIC

Gödel's results apply not only to the language of arithmetic but to any language into which the language of arithmetic can be translated. Thus any recursively axiomatized, consistent theory into which one can translate Q_E is incomplete. The appropriate notion of translation was made precise by Tarski, Mostowski, and Robinson. An *interpretation* (what they call a "relative interpretation") of an arithmetical theory Γ into a language $\mathfrak{L}$ is obtained by doing the following: First, having rewritten all the sentences in Γ so that the "+" sign only appears in the canonical form "$(v_i + v_j) = v_k$," pick a formula "$A(x,y,z)$" of $\mathfrak{L}$ and replace "$(v_i + v_j) = v_k$," by "$A(v_i,v_j,v_k)$," changing bound variables to avoid conflicts. Do the same thing for the other function signs and "0" and pick a formula $L(x,y)$ to replace "≤." Next pick a formula "$N(x)$" of $\mathfrak{L}$ to represent the members of the domain of $\mathfrak{L}$ that are to play the role of natural numbers, and restrict the quantifiers, writing "$(\exists v_i)(N(v_i \wedge \ldots$" in place of $(\exists v_i)$. Finally add an axiom ensuring that "$A(x,y,z)$" represents a function on

the set of things that satisfy "$N(x)$," writing " $(\forall x)(N(x) \rightarrow (\forall y)(N(y) \rightarrow (\exists z)(N(z) \wedge (\forall w)(N(w) \rightarrow (A(x,y,w) \leftrightarrow w = z))))$." Do the same thing for the other function signs and "0." If the theory thus obtained is a consequence of the theory Δ of $\mathfrak{L}$, Δ is said to *interpret* Γ.

We can translate the language of arithmetic into the language of set theory, identifying a number with the set of its predecessors, so that 0 corresponds to $\emptyset$, 1 corresponds to $\{\emptyset\}$, 2 corresponds to $\{\emptyset, \{\emptyset\}\}$, and so on, and defining set-theoretic analogues of "+," "×," "E," "*S*," and "≤" accordingly. The axioms of set theory, in any of its normal versions, interpret PA. We can arithmetize proofs in set theory just as we artimeticized proofs in PA, proving the Second Incompleteness Theorem for set theory. The axioms of set theory, if consistent, cannot prove their own consistency.

This result devastates the Hilbert program. Hilbert wanted to prove the consistency of set theory in a finitistic theory much weaker than set theory, and it turns out that proving the consistency of set theory requires a theory even stronger than set theory.

The standard way to prove that there is no algorithm for testing whether a given sentence is a consequence of a theory Γ—that is, for showing that Γ is *undecidable*—is to interpret an arithmetical theory strong enough to prove the First Incompleteness Theorem into Γ. As far as what we have looked at so far, we would need to take our arithmetical theory to be Q_E, but we can actually do much better. We can define exponentiation in terms of "0," "*S*," "+," and "×," and we can prove the First Incompleteness Theorem in the dialect of the language of arithmetic without "E," with Q in place of Q_E. In trying to prove undecidablity results, this improvement (which is due to Gödel) is an enormous practical advantage.

Let us define $\beta(u,v,w)$ to be the remainder obtained on dividing u by $(v \times w) + 1$. β can be defined by a bounded formula in the language of arithmetic. For $x > 0$, we have $(x\text{E}y) = z$ if and only if the following formula is satisfied:

$$(\exists u)(\exists v)((\beta(u,v,0) = 1 \wedge (\forall w < y)\beta(u,v,\text{S}w) = (\beta(u,v,w) \times x)) \wedge \beta(u,v,y) = z).$$

The right-to-left direction of this characterization is obvious. What is hard is to verify the left-to-right direction by finding an appropriate u and v. We make use of the Chinese Remainder Theorem, which says that, given $p_0, p_1, \dots, p_n$ relatively prime (that is, no two of the p_is have a common divisor other than 1), and given a sequence $a_0, a_1, \dots, a_n$, with each $a_i < p_i$, we can find a number b such that a_i is the remainder on dividing b by p_i, for each i. A proof of the theorem can be found in any number-theory textbook or in George Boolos's *The Logic of Provability* (1993).

Given x,y, and z with $(x\text{E}y) = z$, let $v = z!$, the product of the positive integers $\leq z$. If $s < t \leq z$, then $(s \times v) + 1$ and $(t \times v) + 1$ are relatively prime, since if p were a prime that divided both of them, p would divide $(t - s) \times v$, and so, since $(t - s)$ is one of the factors of v, p would divide v. But this enables us to conclude that the remainder on dividing $(t \times v) + 1$ by p is one, contrary to our assumption that p divides $(t \times v) + 1$. Use the Chinese Remainder Theorem to find u so that, for each $t \leq y$, $x\text{E}t$ is the remainder on dividing u by $(t \times v) + 1$.

Now that we have our Σ definition of exponentiation—Σ, that is, in the restricted language—we can apply our standard tricks for pulling quantifiers to the fronts of formulas to convert a Σ formula of the language with exponentiation to a Σ formula of the language without exponentiation. With this emendation, all the proofs go through.

The use of interpretations originates with Beltrami's proof of the consistency of non-Euclidean geometry. By interpreting non-Euclidean geometry (Euclid's axioms with the axiom of parallels replaced by its negation) into Euclidean geometry, Beltrami showed that if the latter is consistent then so is the the former. Beltrami's strategy was exploited by Alex Wilkie and Samuel Buss to obtain a dramatic strengthening of the Second Incompleteness Theorem, applying it to theories that merely contain Q rather than PA. The details are complicated, but the idea is to interpret into Q a theory that, while weaker than PA (the induction axiom schema being restricted), is just strong enough to provide the Löb conditions (L1)-(L3). The interpretation leaves the arithmetical symbols unchanged but restricts the domain of quantification to an initial segment, replacing "$(\exists x)$" by "$(\exists x)(J(x) \wedge \dots$," for artfully chosen "$J(x)$;" call the sentence thus obtained from ϕ "ϕ^J."

Where Γ is a recursively axiomatized theory that includes Q, let Γ^{-J} be the set of sentences ϕ for which Γ entails ϕ^J. Suppose that Γ entails Con(Γ). Con(Γ) entails Con$(\Gamma)^J$, so that Con(Γ) is in Γ^{-J}. The argument Beltrami used tells us that if Γ is consistent then Γ^{-J} is too. This proof can be formalized in Γ^{-J}, so that Γ^{-J} entails Con(Γ^{-J}). Because (L1)-(L3) yield the Second Incompleteness Theorem for Γ^{-J}, Γ^{-J} must be inconsistent. Consequently Γ is inconsistent.

TRUTH

There is a bounded formula of the language of arithmetic that defines the set of prime numbers, and there is a Σ for-

mula that defines the set of consequences of PA. Tarski proved that there is no formula of the language of arithmetic that defines the set of codes of true sentences. The difficult part of his argument was to say precisely what would be required for a formula to define truth; the easy part is to show that there is no such formula.

A proposed definition of truth is a formula of the form $(Tr(x) \leftrightarrow \tau(x))$, where $\tau(x)$ is a formula of the language of arithmetic. A proposed definition is *materially adequate*, Tarski tells us, if and only if it lets us derive all sentences of the form:

(T) $$Tr(\ulcorner\phi\urcorner) \leftrightarrow \phi.$$

To see that there is no materially adequate definition, apply the Self-reference Lemma to find a sentence λ so that $(\lambda \leftrightarrow \sim\tau(\ulcorner\underline{\lambda}\urcorner))$ is a consequence of Q. The argument here is a formalization of the paradox posed by Eubulides, who asked whether a man who says "I am lying" speaks truthfully.

We can define the set of true sentences of the language of arithmetic within, say, the language of set theory, but we cannot define it within the language of arithmetic. This negative result obtains for any language into which we can translate the language of arithmetic.

The question of what moral, if any, these formal results have for the notion of truth as applied to natural languages is deeply troubling. Tarski showed that there is no formula of the language of arithmetic that means (or even has the same extension as) "true sentence of the language of arithmetic." Manifestly there is a phrase of English that means "true sentence of English," and Tarski and Eubulides' reasoning would appear to apply to that phrase just as to the formal language. Is there in spite of this a coherent way to talk about the truth of an English sentence?

See also Analysis, Philosophical; Aristotle; Church, Alonzo; Computability Theory; Craig's Theorem; Geometry; Gödel, Kurt; Hilbert, David; Infinity in Mathematics and Logic; Kripke, Saul; Logic, History of: Modern Logic; Logical Paradoxes; Mathematics, Foundations of; Russell, Bertrand Arthur William; Tarski, Alfred; Turing, Alan M.; Wittgenstein, Ludwig Josef Johann; Zeno of Elea.

Bibliography

Aristotle. *Prior Analytics*. In *Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1995.

Beltrami, Eugenio. "Saggio di Interpretazione della Geometria Noneuclidea." *Giornale di Matematiche* 6 (1868): 284–312.

Boolos, George *S. The Logic of Provability*. Cambridge, U.K.: Cambridge University Press, 1993.

Buss, Samuel R. "First-Order Proof Theory of Arithmetic." In *Handbook of Proof Theory*, edited by Samuel R. Buss, 79–147. Amsterdam-Elsevier, 1998.

Church, Alonzo. "A Note on the Entscheidungsproblem." *Journal of Symbolic Logic* 1 (1936): 40–41, with a correction on pp. 101–102. Reprinted in *The Undecidable,* edited by Martin Davis, 110–115. Hewlett, NY: Raven Press, 1965.

Craig, William. "On Axiomatizability Within a System." *Journal of Symbolic Logic* 18 (1953): 30–32.

Davis, Martin. *The Undecidable*. Hewlett, NY: Raven Press, 1965.

Dedekind, Richard. *Was sind und was sollen die Zahlen?* Brunswick: Vieweg, 1888. Reprinted in Dedekind, *Essays on the Theory of Numbers*. New York: Dover, 1963, 29–115.

Euclid. *Elements*. 2 vols. Translated by Thomas L. Heath. 2nd ed. New York: Dover, 1956.

Frege, Gottlob. *Begriffschrift*. Halle: L. Nebert, 1879. Reprinted in *From Frege to Gödel*, edited by Jean van Heijenoort. 1–82. Cambridge, MA: Harvard University Press, 1967.

Gödel, Kurt. "Die Vollständigkeit der Axiome der logischen Funktionenkalküls." *Monatschefte für Mathematik un Physik* 37 (1930): 349–360. Reprinted in *From Frege to Gödel,* edited by Jean van Heijenoort (1967), 583–591. Also reprinted in Gödel, *Collected Works,* vol. I, 144–195. New York: Oxford University Press, 1986.

Gödel, Kurt. "Über formal unentscheidbare Sätze der *Principia mathematica und verwandter Systeme* I." *Monatshefte für Mathematik und Physik* 38 (1931): 173–198. Reprinted in *The Undecidable*, edited by Martin Davis, 4–38. Hewlett, NY: Raven Press, 1965. Also reprinted in *From Frege to Gödel,* edited by Jean van Heijenoort (1967), 596–628. Also reprinted in Gödel, *Collected Works,* vol. I, 102–123. New York: Oxford University Press, 1986.

Hájek, Petr, and Pavel Pudlák. *Metamathematics of First-Order Arithmetic*. New York: Springer, 1993.

Hilbert, David. "Über dan Undendliche." *Mathematische Annalen* 95 (1926): 161–190. Reprinted in *From Frege to Gödel*, edited by Jean van Heijenoort. 367–392. Cambridge, MA: Harvard University Press, 1967.

Löb, M. H. "Solution to a Problem of Leon Henkin." *Journal of Symbolic Logic* 20 (1955): 115–118.

Lucas, J. R. "Minds, Machines, and Gödel." *Philosophy* 36 (1961): 120–124.

Penrose, Roger. *The Emperor's New Mind*. New York: Oxford University Press, 1989.

Quine, Willard van Orman. *Mathematical Logic*. 2nd ed. Cambridge, MA: Harvard University Press, 1951.

Rosser, John Barclay. "Extensions of Some Theorems of Gödel and Church." *Journal of Symbolic Logic* 1 (1936): 87–91. Reprinted in *The Undecidable*, edited by Martin Davis, 231–235. Hewlett, NY: Raven Press, 1965.

Smorynski, Craig. "The Incompleteness Theorems." In *Handbook of Mathematical Logic*, edited by Jon Barwise, 821–865. New York: North-Holland, 1977.

Smullyan, Raymond M. *Gödel's Incompleteness Theorems*. New York: Oxford University Press, 1992.

Solovay, Robert M. "Provability Interpretations of Modal Logic." *Israel Journal of Mathematics* 25 (1976): 287–304.

Tarski, Alfred. "Der Wahrheitsbegriff in den formalisierten Sprachen." *Studia Philosophica* 1 (1935): 261–405. Reprinted

in Tarski, *Logic, Semantics, Metamathematics*, 2nd ed., 152–278. Indianapolis: Hackett, 1983.

Tarski, Alfred, Andrzej Mostowski, and Raphael M. Robinson. *Undecidable Theories*. Amsterdam: North-Holland, 1953.

Turing, Alan Mathison. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society*, 2nd series, 42 (1937): 230–265, with a correction at vol. 43 (1938): 544–546. Reprinted in *The Undecidable*, edited by Martin Davis, 115–154. Hewlett, NY: Raven Press, 1965.

Van Heijenoort, Jean. *From Frege to Gödel*. Cambridge, MA: Harvard University Press, 1967.

Wilkie, Alex J., and Jeff B. Paris. "On the Scheme of Induction for Bounded Arithmetic Formulas." *Annals of Pure and Applied Logic* 35 (1987): 261–302.

Vann McGee (2005)

GODFREY OF FONTAINES

Godfrey of Fontaines, the scholastic philosopher and theologian, was a native of Fontaines-les-Hozémont in the principality of Liège. He was born of a noble family about the middle of the thirteenth century, the exact date unknown. About 1270 he began studies at the University of Paris and became a *magister regens* in the faculty of theology there in 1285, having studied under Henry of Ghent and Gervais of Mt. St. Elias. His regency lasted until 1297, and during this period he produced fourteen of his *Quodlibets,* his most important works. There is evidence that he resumed teaching at Paris about 1303 or 1304, composing *Quodlibet* XV at this time. Canon of Liège, probably also of Paris, and provost of Cologne (1287–1298), Godfrey was chosen bishop of Tournai in 1300 but renounced his rights when the election was contested. He is cited among the senior members of the Sorbonne until 1306 and probably died about that time. The obituary at the Sorbonne dates his death October 29, but does not give the year.

Godfrey's doctrinal preferences generally favor the positions of St. Thomas Aquinas, but he manifests a marked independence of judgment on certain points and sometimes works out the logic of Thomas's principles to different conclusions. Some historians (M. De Wulf, E. Gilson) see Godfrey as an opponent of Thomas's distinction between essence and existence in finite being, and attribute Godfrey's stand to a hard-and-fast Aristotelianism that refused to admit an act of the form. Others see Godfrey as opposing the realism of Giles of Rome rather than Thomas. Godfrey held that in the divine mind there is no proper idea of individuals distinct from their species. On the hotly debated issue of the oneness or plurality of substantial forms in composite beings, Godfrey always remained hesitant. He would have favored the doctrine of the unicity of form were it not for the fact that it seemed to contradict theological truths.

Godfrey showed particular acumen in his treatment of psychological problems. Under the influence of Averroes, probably through Siger of Brabant, he espoused an Aristotelianism stricter than that of most of his contemporaries. Godfrey criticized and rejected the so-called Augustinian theory on the genesis of ideas, insisting on the close dependence of human concepts on sense experience. He insisted strongly on the passive nature of the human intellect—the abstractive function of the agent intellect does not consist in the production of any positive disposition in the sensible image upon which it works, but in disregarding in a merely negative way the concrete particularizations characteristic of the image. This outlook is intimately connected with an Avicennan realism of abstract essence, so that Godfrey held that the intellect does not produce intelligibility or universality either in things or in images, but that the agent intellect places the images under an illumination such that the quiddity or essence of the object can appear alone and act on the possible intellect and become known to us.

In his explanation of human free will Godfrey adhered closely to the Thomistic doctrine, but he insisted more than Thomas upon the freedom of the intellect as its foundation. Against the voluntarism of Henry of Ghent, Godfrey stressed the formal influence of the intellect upon the will to the point of making it an efficient cause, whereas Thomas, in different historical circumstances against the Averroists, minimized the formal influence of the object upon the will. In other respects Godfrey did not break cleanly with the Augustinian tradition. For example, he made an interesting equivalence of the active and passive intellects with Augustine's "memory," the passive intellect inasmuch as it conserves species and is a *habitus,* the active intellect inasmuch as it contributes to actual knowledge.

Godfrey was a lively controversialist, combating at length the opinions of his contemporaries, particularly Henry of Ghent, Giles of Rome, and James of Viterbo. Not only did he engage in an active dialogue with his contemporaries, but he also occupied himself with pressing problems—moral, legal, social, and political—arising from daily life. Among his admirers can be listed John the Wise, Peter of Auvergne, and Gerard of Bologna; among his critics, Bernard of Auvergne, Gonsalvus of Spain, and John Duns Scotus. His influence was widespread and lasted well into the fourteenth century but waned thereafter.

See also Agent Intellect; Aristotelianism; Augustine, St.; Augustinianism; Averroes; Duns Scotus, John; Essence and Existence; Giles of Rome; Henry of Ghent; Realism; Siger of Brabant; Thomas Aquinas, St.

Bibliography

WORKS BY GODFREY

Les philosophes Belges, Vols. 2–4 and 14: *Les quodlibets de Godefroid de Fontaines,* edited by M. De Wulf, A. Pelzer, J. Hoffmans, O. Lottin. Louvain: Institut Supérieur de Philosophie de l'Université, 1904–1937.

Wippel, J. F., ed. "Godfrey of Fontaines: *Disputed Questions 9, 10, 12.*" *Franciscan Studies* 33 (1973): 351–372. (See below: Wippel, 1982).

WORKS ON GODFREY

Arway, R. "A Half Century of Research on Godfrey of Fontaines." *New Scholasticism* 36 (2) (April 1962): 192–218.

Brown, S. F. "Godfrey of Fontaines and Henry of Ghent: Individuation and the Condemnation of 1277." In *Société et Église. Rencontres de philosophie médiévale,* Vol. 4, edited by S. Wlodek, 193–207. Turnhout, Belgium, 1995.

Duin, J. J. "La bibliothèque philosophique de Godefroid de Fontaines." *Estudios Lulianos* 3 (1959): 21–36, 137–160.

Hoffmans, J. "La table des divergences et innovations doctrinales de Godefroid de Fontaines." *Revue Néoscolastique* 36 (1934): 412–436.

Langevin, E. "Nécessité ou liberté chez Godefroid de Fontaines." *Sciences ecclésiastiques* 12 (1960): 175–203.

Tihon, P. *Foi et théologie selon Godefroid de Fontaines.* Paris and Brugge, 1966.

Wippel, J. F. "Godfrey of Fontaines and Henry of Ghent's Theory of Intentional Distinction between Essence and Existence." *Sapientiae procerum amore. Mélanges Médiévistes offerts à Dom Jean-Pierre Müller, O.S.B. Studia Anselmiana* 43 (1974): 289–321.

Wippel, J. F. "Godfrey of Fontaines and the Act-Potency Axiom." *Journal of the History of Philosophy* 11 (1973): 299–317.

Wippel, J. F. "Godfrey of Fontaines and the Real Distinction between Essence and Existence." *Traditio* 20 (1964): 385–410.

Wippel, J. F. "Godfrey of Fontaines at the University of Paris in the Last Quarter of the Thirteenth Century." *Nach der Verurteilung von 1277. Philosophie und Theologie an der Universität von Paris im letzten Viertel des 13. Jahrhunderts. Studien und Texte. Miscellanea Mediaevalia* 28 (2001): 81–101.

Wippel, J. F. "Godfrey of Fontaines' Disputed Questions 9 and 10 (Bruges 491): by Godfrey or by Giles of Rome?" *Franciscan Studies* 42 (1982): 216–247.

Wippel, J. F. "Godfrey of Fontaines: Divine Power and the Principle of Noncontradiction." *Les philosophies morales et politiques au Moyen Age. Actes du Ixe Congrès de Philosophie médiévale,* edited by B. C. Bazán et al. Ottawa, 1995, 1388–1398.

Wippel, J. F. "Godfrey of Fontaines on Intension and Remission of Accidental Forms." *Franciscan Studies* 39 (1979): 316–355.

Wippel, J. F. "Godfrey of Fontaines, Peter of Auvergne, John Baconthorpe, and the Principle of Individuation." *Essays Honoring Allan B. Wolter,* edited by William Frank and Girard Etzkorn, 309–349. St. Bonaventure, NY, 1985.

Wippel, J. F. *The Metaphysical Thought of Godfrey of Fontaines.* Washington, DC: Catholic University of America Press, 1981.

Wippel, J. F. "Possibile Sources for Godfrey of Fontaines' Views on the Act-Potency 'Composition' of Simple Creatures." *Mediaeval Studies* 46 (1984): 222–244.

Wippel, J. F. "The Relationship between Essence and Existence in Late-Thirteenth Century Thought: Giles of Rome, Henry of Ghent, Godfrey of Fontaines, and James of Viterbo." *Philosophies of Existence Ancient and Medieval,* edited by P. Morewedge, 131–164. New York, 1982.

Wippel, J. F. "The Role of the Phantasm in Godfrey of Fontaines' Theory of Intellection." *L'homme et son univers au moyen age. Actes du septième congrés international de philosophie médiévale,* Vol. 2, 573–582. Louvain-la-Neuve, 1986.

Wippel, J. F. "Some Issues concerning Divine Power and Created Natures according to Godfrey of Fontaines." *Diakonia. Studies in Honor of Robert T. Meyer,* edited by T. Halton and J. P. Williman, 158–181. Washington, DC, 1986.

Robert J. Arway, C.M. (1967)
Bibliography updated by Stephen F. Brown (2005)

GOD/ISVARA IN INDIAN PHILOSOPHY

This entry deals briefly with the Isvara concept in the six schools of philosophy in Hinduism, usually called the orthodox schools because they were thought to believe in the authority of the Vedas. The schools are Nyaya, Vaisesika, Samkhya, Yoga, Purvamimamsa, and Uttaramimamsa, also called Vedanta. This article is not a philosophical discussion of the nature of Isvara but is a description of how Isvara is viewed in these schools.

Each of the main schools has a foundational text called *sutras*. The word *sutra* means "a thread" and is usually a brief sentence of a few words that convey the basic philosophy of the respective school. As these *sutras* are difficult to follow without some explanation, commentaries called *bhasyas* emerged from erudite commentators, which in turn spawned commentaries on commentaries that went on for a long time up until the present day. One could generally assign the period between the second century to the fifith century CE as the point of origin for these schools of philosophy. The authors of the first *sutras* of Nyaya was Gautama, of Vaisesika Kanada, of Samkhya Kapila, of Yoga Patanjali, of Purvamimamsa Jaimini and of Vedanta Badarayana, also called Vyasa. These *sutras* are all written in the oldest lan-

guage of the world, Sanskrit, which was the language of religion and philosophy for most of Hinduism's history. Each of these schools has a unique approach to the understanding of Isvara.

It is useful to remember that in Hinduism, within which one has to view the six schools, Isvara cannot be equated with the concept of God as it is understood in the Abrahamic religions. Isvara does not have the role of creator because Isvara does not create the world and the selves from "nothing." The theory of karma and the cyclical evolution and dissolution of the universe in periodic cycles, in keeping with karma, does not allow Isvara the same role that is assigned to God in the Abrahamic religions. Hindu schools of philosophical thought have liberation or *moksa* as their highest value, and each school develops its ontology and epistemology in order to realize this eschatological value.

All the orthodox schools share in the belief that the self (also called variously as *atman, jiva, purusa*, and so on) is an eternal entity born into the world and associated with a body and other faculties in accordance with the karma that belongs to it from the past. The ego or sense of "I" that one normally associates with notions of one's identity is not the real self in these schools. The real self is the *atman*, which is the inner essence, and it is the quest for this inner self and its realization that constitutes ultimate freedom or *moksa*, which, in an extended sense, also means breaking the chain of subsequent births and deaths in the world. This quest for the true self is also situated within the inner efforts of individual earthly selves and so, on the surface at least, there is no role for Isvara in the way these *sutras* were initially formulated. As Isvara does not also have a role in the direct evolution of living beings in the world, the concept of Isvara is something that is sneaked into the *sutras* sometimes by later commentators in order to serve other needs. Since there is no uniformity of approach in the methodology followed for this purpose in the different schools, our task is to examine how this is done in the different schools and how the Isvara concept is made to fit into the general philosophy of the different systems.

NYAYA-VAISESIKA

Although *Nyayasutras* and *Vaisesikasutras* were composed by different authors, because of certain similarities in the way they viewed ontology and epistemology, they gradually came to be discussed jointly in all discussions of the philosophical schools. I shall also deal with them together in this entry.

Nyaya and Vaisesika are realistic schools and trace the origin of the real from basic atomistic principles. They have minor disagreements in the number of metaphysical categories and also in the emphasis that each brings into the discussion. Whereas Vaisesika concentrates on discussing in detail the metaphysical categories and the ultimate realistic principles, Nyaya is more concerned with developing the epistemology of gaining right knowledge of reality, which is to realize the true nature of *atman*.

The ultimate realities that explain the whole universe are the atoms of earth, water, fire, air, ether (*akasa*), space, time, mind, and self (*atman*). As can be seen from the above, the system is not purely materialistic. The self is considered to be eternal and many. Though omnipresent, it is confined to the body to which it is associated. The theory of causation is teleological; the karma called *adrsta* (unseen potential of past *dharma* and *adharma*) is sufficient to determine the coming together of the eternal atoms to form bodies for the selves to continue their cycle of lives until they attain liberation through a discrimination between the true nature of *atman* and the false identity it has with the body and other material substances. Thus the initial *sutras* of Gautama and Kanada did not really have a place for Isvara though Gautama refers to Isvara in a weak sense in one place (*Nyayasutras* IV 1.21)

Later commentators, however, found a place for Isvara in both Nyaya and Vaisesika by using various arguments. The eternal atoms and the eternal selves (*jivas*) are not created by Isvara. However, there was a need to bring together the *jivas* and their future embodied lives in consonance with past karma. Since karma itself was not a conscious category, there was room for the introduction of an intelligent, superconscious *atman* who could fulfill this task. Thus, some of the reasons for the existence of Isvara in Nyaya-Vaisesika are: (1) because the world as an effect needs an agent as an efficient cause equal to the task of coordinating the different phenomena of the world; (2) the atoms being basically inactive, Isvara enables them to combine in accordance with the past karma (*adrsta*) of *jivas*; (3) the manifestation and destruction of the world in cyclical rhythm is due to Isvara.

Liberation called *apavarga* in these two schools is, however, still an individual effort, and Isvara has no role to play in the achievement of the highest value of liberation (*moksa*) for the *atman*. It comes about by correct knowledge of things of which Isvara is also just one more thing.

SAMKHYA-YOGA

The *Samkhyakarika* of Isvarakrsna and the *Yogasutra* of Patanjali are used for this discussion. Like Nyaya and Vaisesika, Samkhya and Yoga also share some metaphysical ideas; they both believe in two ultimate realities—one called *prakrti*, the material reality and the building block of the world—and the other the spiritual reality called *purusa*, which is another word for *atman*. Even though there are many *purusas* in these two schools, they are not different from one another in essence. Whereas in Nyaya-Vaisesika *atman* only has knowledge as an adventitious property, in these schools it is also characterized as being pure consciousness.

The coming into being of the world and its properties in both the schools is from *prakrti* alone without the assistance of any outside agency. The proximity of *purusa* and *prakrti* is a sufficient condition for the evolution and involution of the world. *Prakrti* is viewed as constituted of the three *gunas* (characteristics) of *sattva, rajas*, and *tamas*. These *gunas* are not properties of *prakrti* but its very nature. *Prakrti* as constituted by the *gunas* is in constant motion. When the *gunas* are in equilibrium, there is no evolution of the world, and the world evolves when there is disequilibrium of the *gunas*. Thus evolution and involution is a teleological process governed by the past karma associated with *purusas*. The evolution is also explained as serving the twin purposes of *purusa*: experience in the world and gaining liberation or *kaivalya*.

Philosophically there are many difficulties, among them (1) the conception of many *purusas* who are all of the same nature of pure consciousness, (2) an insentient *prakrti* sufficient to explain the evolution of the world, (3) the problem of what initiates the disequilibrium in the first place, and so on. This article, however, confines itself to Isvara in the system. Thus, as seen above, it is clear that there is no role for Isvara in the *Samkhyakarika*. The final goal of liberation or *kaivalya* also comes through discrimination between the true nature of *purusa* and *prakrti* gained by correct knowledge. Thus Isvara does not figure either in the coming into being of the world or in the attainment of *kaivalya* for *purusa*.

Although Yoga shares with Samkhya the belief in the ultimate two realities of *prakrti* and *purusa*, there is a weak introduction of Isvara in the system, described as an excellent *purusa*. The excellent *purusa* (Isvara) is unaffected by karma in the past, present, and the future. By arguing from experience that there is a graded scale of knowledge, wisdom, power and so on, Patanjali describes Isvara as the one who represents the utmost excellence and who is also an aid to the practice of yoga by being an object of support (*alambana*) in meditation. But, at the same time, Isvara is only one among many supports in meditation. He is also called the first guru who teaches the Vedas to the sages. His symbol is Om, and he is one that brings the association and disassociation of *purusa* and *prakrti* to start the evolution and involution of the universe. Though philosophically these are weak arguments, Isvara has been accommodated in a backhanded manner into Yoga philosophy by Patanjali. However, because Isvara does not play any role in the manifestation of the world, in the evolution of individual *purusas*, or in the granting of liberation, one can conclude that Yoga also does not accommodate Isvara in the usual sense of the term.

PURVAMIMAMSA AND UTTARAMIMAMSA

The earlier sections (*purva*) of the Veda, the *mantra* and ritualistic or *brahmana* sections, deal with rituals and are therefore called the *karmakanda* (sections dealing with rituals) while the latter sections (*uttara*), the Upanisads, deal with knowledge of reality and so are called the *jnanakanda* (sections dealing with knowledge). Both the schools believe implicitly in their respective sections of the Veda. I shall first consider Purvamimamsa (PM) and then talk about Uttaramimamsa (UM) or Vedanta.

PM is a realistic school and considers the Vedas as an infallible authority. This discussion is based on Jaimini's PM *sutras* and on some later commentaries. PM primarily focuses on the right interpretation of Vedic statements and in the correct performance of rituals or karma. It differs from all other schools in not believing in the periodic evolution and involution of the world. According to PM, there never was a time when the world was different from what it is now. PM believes in the law of karma as an unseen power, in the individual selves that are ruled by the law of karma, and in *moksa* which, though initially the attainment of *svarga* (heaven), gradually changed to the attainment of the true nature of *atman* in later commentaries. It is attained by the exhaustion of *dharma* and *adharma* by the disinterested performance of one's own karma. The ultimate authority being the Veda, there was no need of an Isvara in the system. Karma classified in various ways was sufficient to explain the coming into being of the individual selves and their ultimate achievement of *moksa*.

Uttaramimamsa or Vedanta is based primarily on the Upanisads, which are the basis of Badarayana's Brahmasutras (BS), also called the Vedantasutras. Although there are many Vedanta schools based on differing interpreta-

tions of the BS and the Upanisads, this entry shall be dealing only with Advaita Vedanta and very briefly with Visistadvaita and Dvaita philosophies. Samkara (c. eighth century CE), Ramanuja (c. eleventh/twelfth century CE) and Madhva (c. thirteenth/fourteenth century CE) are the important commentators on the BS for Advaita (nondualism), Visistadvaita (qualified nondualism), and Dvaita (dualism), respectively.

Samkara declares the ultimate ontological reality as Brahman and identifies the individual self called *atman* with this Brahman. Because there is only one Absolute Reality, the so-called reality of the world and all other things is only an appearance, according to Advaita (nondual)Vedanta. Brahman is described as *nirguna* (without any qualities) and cannot be viewed in a personal way. However, because the world appears to be real, in order to reconcile this world-reality with the ultimate reality, Advaita views reality as a threefold entity that includes the illusory (*pratibhasika*, such as dreams), worldly experience (*vyavaharika*), and absolute reality (*paramarthikasatta*). Because Brahman is also without any properties, it cannot be an agent of manifestation. Therefore, the necessity of explaining the world forces Samkara to introduce *maya* (cosmic ignorance), which, when associated with Brahman, is called *saguna*-Brahman (Brahman with qualities) or Isvara, which is then considered to be both the efficient and material cause of the universe. There are many ways in which *maya* and its association with Brahman are explained in order to maintain the nondual nature of Advaita, but that need not concern us here. Because this Isvara is not free to manifest the world and the selves but is bound by the karma of the individual selves in the manifestation of the world, and because Isvara does not have a role to play in the attainment of *moksa* of the selves, it is only a device to explain the so-called reality of the world. Liberation is achieved when, through correct knowledge, the *atman* realizes its identity with Brahman.

By the times of Ramanuja and Madhva, a fundamental change has taken place in the religious sphere. Devotion (*bhakti*) has come to be valued as higher than knowledge in the attainment of *moksa*, and the highest entity Brahman is also now viewed in a personal manner. Brahman—variously called Narayana, Visnu, Gopala-Krsna, Vasudeva-Krsna, and so on—is capable of responding to the devotion of individual selves and even to mitigate the evil effects of karma, enabling the devotee to attain *moksa*. *Moksa* is also defined differently in a dualistic manner whereby the self retains its separation from the Supreme Brahman it worships. There are differences in the way in which the nature of the individual selves, the world, and the nature of Brahman are understood in Ramanuja's and Madhva's interpretation of the BS. But those are in the details. In both *bhakti* is a sufficient condition for *moksa*.

As long as the philosophical schools depended only on correct knowledge to attain the true nature of *atman*, there was no need for dependence on an outside agent called Isvara to enable *atman* to achieve its highest value. But when the religious atmosphere changed with the introduction of devotion as the paramount means for achieving liberation, it was possible for Isvara to play a role in many ways—like the reduction of karma, bestowing grace, and so on—for the *atman* to attain *moksa*.

See also Atomic Theory in Indian Philosophy; Brahman; Causation in Indian Philosophy; Indian Philosophy; Karma; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Meditation in Indian Philosophy; Self in Indian Philosophy.

Bibliography

Chakravarthi, Ram-Prasad. *Knowledge and Liberation in Classical Indian Thought*. New York: Palgrave, 2001.

Chandradhar, Sharma. *A Critical Survey of Indian Philosophy*. Delhi: Motilal Banarsidass Publishers, 1987.

Dadgupta, Surendranath. *A History of Indian Philosophy*, vols. 1–5. Cambridge: University Press, 1969.

Hiriyanna, M. *Outlines of Indian Philosophy*. London: Allen & Unwin, 1967.

Mohanty, J. N. *Classical Indian Philosophy*. Oxford: Rowman & Littlefield, 2000.

Radhakrishnan, S. *Indian Philosophy*. Vol. 1. London: Allen & Unwin, 1929. First published 1923.

T. S. Rukmani (2005)

GODWIN, WILLIAM

(1756–1836)

William Godwin, English political philosopher, novelist, and essayist, was born at Wisbech, in Cambridgeshire, where his father was a dissenting minister. He was educated at Hoxton, one of the dissenting colleges that had been founded because of the refusal of the established universities to admit nonconformists, and himself entered the ministry in 1778. By 1783, apparently as the result of reading Claude-Adrien Helvétius and Baron d'Holbach, he had lost his faith, and instead took to literature as a means of livelihood. Much that he wrote at this time was hackwork, including three novels, none of

which have survived. He did, however, gain some reputation as a political journalist, contributing regularly to such Whig publications as *The Political Herald* and *The New Annual Register.*

In 1791 Godwin managed to free himself from hackwork by persuading a publisher to subsidize him while he settled down to a serious treatise on political theory. The *Enquiry concerning Political Justice* (London, 1793) was the kind of book the intellectual radicals of the day had been waiting for, and Godwin soon became a celebrity. In this work, he set down, with passionate sincerity and a complete absence of compromise, the radical beliefs that were emerging from the French Revolution and the intellectual ferment that had preceded it. The book went through three editions (2nd ed., 1796; 3rd ed., 1798).

For a few years Godwin was happy and successful. His novel *Things as They Are; or, the Adventures of Caleb Williams* (London, 1794) was widely acclaimed as a masterpiece. In the same year, his *Cursory Strictures on the charge delivered by Lord Chief Justice Eyre to the Grand Jury* (London, 1794), which protested against the committal on charges of treason of twelve leading radicals, may have been partly responsible for the acquittal of three of the defendants and the dropping of the charges against the others. He also published a volume of essays, *The Enquirer* (London, 1797).

In 1797 Godwin married Mary Wollstonecraft, the author of *A Vindication of the Rights of Woman.* She died in the same year, a few days after the birth of their daughter Mary, who was to become the wife of Percy Bysshe Shelley. In 1798 Godwin wrote a memoir of his wife (*Memoirs of the Author of A Vindication of the Rights of Woman,* London, 1798), and in the following year another novel, *St. Leon* (London, 1799).

From then on his fortunes declined. Radicalism came into disfavor, and Godwin was fiercely attacked, sometimes by his former friends. One of these preached a sermon against him, to which he replied in *Thoughts Occasioned by Dr. Parr's Spital Sermon* (London, 1801).

In the same year, 1801, he married a widow, Mary Jane Clairmont. His second marriage was less happy than his first. Before long he was back at hackwork, and the last years of his life were spent in poverty.

Though he continued writing until his death, many of his later books were potboilers. The most important of these works (all of which, with one exception, were published in London) are his novels *Fleetwood* (1805), *Mandeville* (Edinburgh, 1817), *Cloudesley* (1830), and *Deloraine* (1833); his *Life of Chaucer* (1803); *An Essay on Sepulchres* (1809); *Of Population* (1820), a reply to Thomas Robert Malthus; and *Thoughts on Man* (1821). He had expressed his views on religion in a book that he called *The Genius of Christianity Unveiled,* but it was not published until long after his death, when it appeared under the title of *Essays Never before Published* (London, 1873).

ANARCHISM

Godwin's political theory is uncompromisingly anarchist. He was opposed to all kinds of coercion, including punishment, partly because of his determinism. "The assassin," he said in *Political Justice,* "can no more help the murder he commits than the knife in his hand." Such a view might be thought to lead to an authoritarianism based on the need to condition men rigidly so that their actions will not be antisocial. Godwin did indeed believe that it is society that molds men's characters and actions. He was one of the earliest proponents of what is now called cultural determinism, but he combined this view with a quite extreme liberalism and individualism.

TYPES OF SOCIETY. Before Godwin, Baron de Montesquieu (and indeed Plato) had already maintained that each type of government developed not only its own characteristic type of institution, but also its own characteristic attitudes and value judgments within the minds of its citizens. Montesquieu distinguished three main types of government, each with its own characteristic "spirit": despotism, whose spirit is fear; monarchy (the aristocratic semifeudal type of society still current in most of eighteenth-century Europe), whose spirit is honor; and the republic (which for Montesquieu suggested Sparta, oddly idealized in the eighteenth century, as much as the actual contemporary example), whose spirit is virtue, used not quite in its modern sense but rather to mean public-spiritedness.

Where Godwin differed from the modern anthropologist and, to some extent at least, from Montesquieu, was that he saw all three types of society as corrupting their citizens—even monarchy, with its ideal of honor, so much admired in Godwin's time. Honor demands that one shall do what is fitting to one's rank; this was thought a sufficient motive to keep the wheels of society turning and to ensure decent and at times even noble behavior. Falkland, in *Caleb Williams,* is a portrait of the Man of Honor. He is a thoroughly charming, accomplished, and benevolent man, but at the moment of crisis he is prepared to commit murder and further crime rather than see his good name disgraced. The moral is clear enough:

Honor is not enough to make men behave benevolently. Only benevolence will do that.

In some ways Godwin was more sympathetic to the republican ideal than to any other. Montesquieu's Republic is a state like Sparta, in which there is equality, frugality, and complete submission of the individual to the state. Godwin was strongly in favor of equality; he shared the republican objection to ostentation and luxury; and he agreed that it was the supreme duty of the individual to merge his own welfare with that of his neighbor. Nevertheless, Godwin rejected the republican ideal quite as decisively as the monarchic one—first, because it turns men's attention away from human beings toward a quite mythical entity called the State, and second, because it teaches men to merge their own judgment in that of the majority.

CORRUPTING INFLUENCE OF GOVERNMENTS. Behind these objections, there is a quite general criticism that would apply to any type of government. Godwin believed that social institutions corrupt because they create prejudice; they prevent men from seeing things as they are. Men in society see themselves and one another through a mist of preconceived ideas—as members of this or that social class, as fellow countrymen or foreigners, but never as the unique individuals they really are. To some extent this is inevitable, in society or out of it: All generalizations, according to Godwin, distort the particulars that are subsumed under them; and yet it is impossible to think at all without generalizing. Nevertheless, to see things as they are, though difficult, is not impossible. However, it is peculiarly difficult to see ourselves and our fellow men as they are unless we can get close enough to them to sympathize with them and to realize the true complexity of their motives.

Government perverts our judgment in three main ways. First, it creates artificial barriers between men, as the result of social inequality and of the insincerity that results from the perpetual effort to keep up with the Joneses.

Second, it encourages us to do the right things for the wrong reasons. Patriotism and social prestige are both wrong reasons for treating other men benevolently. Punishment, which leads men to keep the law from fear and not because they understand the reasons for keeping it, acts in the same way. The objection to doing the right thing for the wrong reason is that, since it results in the muddling of men's minds, they will become quite incapable of adapting their actions intelligently to changed circumstances; and consequently the things they do will not for long be the right ones.

Third, government encourages us to acquiesce in the opinions of others, whether of the majority or some minority of rulers. This means that we accept conclusions without really understanding the evidence upon which they are based. Consequently we are acting from prejudice, without any real understanding. Once again, this can only make us unfit to cope with a complex and changing world. "The history of mankind," Godwin said, "is little else than a record of crimes" (*Political Justice,* I, ii), crimes that are caused ultimately by man's inability to see things as they are and to think clearly about them.

THE IDEAL SOCIETY. For these reasons, Godwin rejected all three of Montesquieu's forms of society. Godwin did not, however, merely want to put a fourth type of government in place of the other three. He believed not merely that all existing governments have corrupted society, but also that government as such is necessarily corrupting.

In what kind of society, then, can one hope to escape prejudice? Obviously not in a large society, because every individual is unique, and one can avoid prejudice only by an intimate and sympathetic understanding of one's fellows. Indeed, it is rare enough to know even one person well enough not to misjudge him. In this connection, it is worth noticing that Godwin had an almost morbid obsession with friendship. In all his novels the central figure complains of being without a friend, and is cut off from the rest of his fellows—usually as the result of his own prejudices. This loneliness is, for Godwin, the central tragedy of the human situation.

The ideal community, then, must not be large and must not be highly organized. The citizen must never be a cog in a machine, unable to see the significance of his everyday activities. There must be no class distinctions that prevent us from seeing individuals as individuals. And there must be no formal rules and regulations, because these are rules of thumb that demand the acquiescence of the individual in propositions he does not really understand. For the same reason, there is to be no punishment. For Godwin, the ideal society is one in which individuals cooperate without any kind of compulsion because they like and understand each other and wish each other well.

THE NEED FOR GRADUAL CHANGE. Godwin's ideal society is usually criticized as absurdly unpractical and utopian. The truth is that he was not really a political

reformer in the ordinary sense. He was not very interested in blueprints for a "brave new world"; he did not believe in political organizations, and he had no program. He was primarily a moralist concerned with analyzing the causes of prejudice; once we understand these, according to him, the cure may very well be left to look after itself. We need to have some vague idea of the direction in which we wish to move, but we need have no more than that, because change can be brought about only very slowly and gradually. Godwin insisted again and again upon the folly of violent change. We can do nothing here and now but try to make a few small breaches in the wall of prejudice. If enough people can be brought to see what is wrong with society, society will right itself—but only by slow and gradual changes that will take generations. There is no question of a political program; political organizations are themselves a cause of prejudice. We are not even to point the way to the new society by setting an example of a better way of life. It is by reasoning and discussion that we must break down existing prejudice. The immediate task is to destroy the current ideals of honor and virtue. These ideologies have been created by existing institutions. They can be destroyed without destroying these institutions, however, because though prejudice is strong, it cannot entirely blind men to the facts. When supported by existing institutions, opinion can only be changed slowly—but it can be changed, and as it changes, the institutions will gradually be transformed.

MORAL THEORY

Society is to be transformed, then (even if only slowly and gradually), by means of a change in men's opinions—chief among them their opinions about what is desirable. But here a major difficulty presents itself. Even if men can be brought to see things as they are, how can their moral beliefs be changed thereby, since, as David Hume had pointed out, it is impossible to derive any conclusion about what ought to be the case from knowledge about what is the case? Godwin knew of Hume's views and agreed with them, at least in part. In a certain sense, Godwin believed that virtue is knowledge; but he also insisted quite emphatically that "moral reasoning is nothing but the awakening of certain feelings" (*Political Justice*). It is in order to reconcile these two positions that he introduced his concept of natural goodness.

GODWIN'S UTILITARIANISM. Godwin was a thoroughgoing utilitarian. For him, the right action is the one that makes for the greatest happiness of the greatest number. His utilitarianism, however, is unusual in two respects. First, it is not derived from egoism: The "greatest happiness" principle is ultimate, and cannot be derived from self-interest. Second, Godwin was quite prepared to push utilitarianism to its logical conclusion and openly embrace the consequences that many critics have regarded as fatal to it. He repudiated as immoral any obligations that cannot be derived from the general obligation to promote the general happiness—such obligations, for example, as promises and other contractual obligations, or the domestic obligation to prefer the happiness of one's friends and family to the greater happiness of others. He caused considerable scandal by the passage in *Political Justice* in which he said that one ought to save François Fénelon from a burning building rather than his chambermaid (supposing that the archbishop has more to contribute to the general happiness), even if the chambermaid is one's own mother.

KNOWLEDGE AND VIRTUE. For Godwin, then, our belief that *X* is desirable is true only when *X* is something that will make for the general happiness. Since "moral reasoning is nothing but the awakening of certain feelings," the virtuous man is the one who does desire whatever makes for the general happiness. How will seeing things as they are awaken this desire? The following may serve as an example: If someone says that it is a bad thing that millions of people in a distant part of the world are starving, I may very well agree, but the chances are that I won't do anything about it. But now suppose that one of them comes and starves on my doorstep. Almost certainly I shall be moved to feed him. When I see a man starving before my eyes, the proposition "starvation ought to be relieved" takes on a new meaning for me. I can now see in detail precisely how and why starvation is evil; I can see exactly how the generalization applies to the particular instance. It is not a question of perfect knowledge being reinforced by emotion, since my knowledge before was imperfect. When it becomes perfect, it necessarily brings the emotion with it. Thus, if I *know* in this sense that *X* is desirable, and if I not only accept this as a rule of thumb but also fully understand the evidence on which it is based, then I cannot but desire it.

THE CAUSES OF IMPERFECTION. Because he held that men are "perfectible" and "naturally good," Godwin has been accused of excessive optimism. He did not mean, however, that men are, or are likely to become, perfect. He was merely saying that imperfection has causes (usually social causes) that may be removed. To talk about original sin is to give up the search for the causes of sin. It is as if we were to say "disease is a natural phenomenon" and turn our backs on medicine. In general, no doubt,

wickedness, like disease, is always with us; but any particular piece of wickedness, like any particular disease, has specific causes, and it may be possible to remove them. That the causes are far-reaching and difficult to remove, Godwin did not deny. He was even prepared to grant that there may be "something in the nature of man incompatible with absolute perfection." Men can never fully understand the principle of universal benevolence, simply because they cannot hope to know all their fellows intimately. But this is an ideal toward which we can strive and which, even if it can never be reached, can always be brought a little closer.

RIGHTS. As a thoroughgoing utilitarian, Godwin, like Jeremy Bentham, denied that there are any natural rights. The only right, which is also a duty, is to do whatever makes for the general happiness. He would not concede the rights to life and liberty. Nevertheless, the individual has one right—the right of private judgment. The reason for this is simply that, in the final analysis, nothing will be gained if men do not understand the reasons for acting as we wish them to act. When men see things as they are, they will quite freely and without any kind of coercion do what makes for the general happiness. Any attempt to coerce them will hinder them from seeing things as they are and will therefore do more harm than good. In this way, Godwin was able to reconcile utilitarianism with the utmost insistence on individual freedom, and especially on freedom of thought and opinion.

THE BASIS OF GODWIN'S ETHICS. Godwin's basic beliefs may be summarized in three propositions—one about ethics, one about logic, and one about social psychology.

The nature of virtue. The ethical proposition is that to be virtuous is to feel the right emotions. The right emotions are those that men feel when they see all the facts clearly. When we analyze these emotions, we find that they are all consistent with the "greatest happiness" principle.

In the last analysis, statements about morals are expressions of feeling. But this does not mean that we cannot reason about morals. There are good and bad reasons for feeling frightened or angry. Fear is appropriate in a situation of danger. We do not doubt that if once we make a man see the full facts of the situation, the emotion of fear will come of its own accord. In the same way, if we want men to feel the appropriate emotions of benevolence, pity, affection, and so on, we can do it by making them realize the full facts about human beings. It is in this sense that men are naturally good.

Generalizations. The logical proposition is that true knowledge is of particulars and all generalizations are, if not false, at least seriously misleading. It is possible to know in a sense that a situation is dangerous without feeling the appropriate emotion. A man may ignore the danger out of bravado, but in that case we may say that he does not fully appreciate the danger. He knows, as a generalization, that the situation is dangerous, but he does not know the particulars that the generalization expresses—which is to say that he does not really know the generalization at all.

But we cannot, of course, do without generalizations. It is impossible to know every particular in all its particularity. Here, then, is an inescapable source of error. It is particularly likely to mislead us in our judgments of human beings and of human actions, for every human being is unique. Since we cannot know everyone intimately, we have to rely on generalizations, any one of which may be seriously misleading when applied to a given individual. Such generalizations form, as it were, a distorting glass through which we look at the world. And, since virtue depends on feeling the appropriate emotions toward other human beings, emotions that depend on a clear perception of all particulars, the logical proposition is an adequate explanation of human frailty. It also explains what Godwin meant by prejudice.

Political institutions influence beliefs. The proposition about social psychology is that the generalizations men believe depend on the political institutions under which they live. In practice, the particular distorting glasses we use are, so to speak, handed out to us by the governments under which we live. Our opinions about how human beings actually behave are influenced by concepts derived from legal institutions, like "thief" or "murderer," or concepts derived from social institutions, like "lord" or "pauper." These stereotypes come between us and the actual human beings around us. Our opinions about how human beings ought to behave are distorted by such concepts as "honor" and "virtue," which stem directly from political institutions, as Montesquieu had clearly demonstrated. This is what Godwin meant by the corrupting effect of government.

GODWIN'S SOLUTION. Two main conclusions follow from these three basic beliefs: First, that if we want to improve human beings, we must help them to see things, and particularly each other, as they are; and second, that

this can be done by simplifying society, by sweeping away social categories like rank and the legal categories that depend on punishment, and by encouraging individual judgment so that men will no longer trust to rules of thumb.

See also Anarchism; Bentham, Jeremy; Helvétius, Claude-Adrien; Holbach, Paul-Henri Thiry, Baron d'; Hume, David; Montesquieu, Baron de; Plato; Shelley, Percy Bysshe; Utilitarianism; Virtue and Vice; Wollstonecraft, Mary.

Bibliography

PRIMARY WORKS

Enquiry concerning Political Justice. Edited by K. Codell Carter. Oxford: Clarendon Press, 1971.

Collected Novels and Memoirs of William Godwin. 8 vols. Edited by Mark Philp, et al. London: Pickering, 1992.

Political and Philosophical Writings of William Godwin. 7 vols. Edited by Mark Philp, et al. London: Pickering, 1993.

SECONDARY WORKS

Bellamy, Richard. "Godwin and the Development of 'the New Man of Feeling.'" *History of Political Thought* 6 (1985): 411–432.

Brailsford, H. N. *Shelley, Godwin, and Their Circle.* London: Williams and Norgate, 1913.

Brown, Ford K. *Life of William Godwin.* London: Dent, 1926.

Clark, John P. *The Philosophical Anarchism of William Godwin.* Princeton, NJ: Princeton University Press, 1977.

Fleisher, David. *William Godwin.* London: Allen and Unwin, 1951.

Grylls, R. G. *William Godwin and His World.* London: Odhams, 1953.

Haakonssen, Knud, ed. *Enlightenment and Religion: Rational Dissent in Eighteenth-Century Britain.* Cambridge, U.K.: Cambridge University Press, 1996.

Kramnick, Isaac. "On Anarchism and the Real World: William Godwin and Radical England." *American Political Science Review* 66 (1972): 114–128.

Locke, Don. *A Fantasy of Reason: The Life and Thought of William Godwin.* London: Routledge, 1980.

Marshall, Peter H. *William Godwin.* New Haven, CT: Yale University Press, 1984.

Monro, D. H. *Godwin's Moral Philosophy.* Oxford: Oxford University Press, 1953.

Paul, C. Kegan. *William Godwin: His Friends and Contemporaries.* 2 vols. London: King, 1876.

Philp, Mark. *Godwin's Political Justice.* London: Duckworth, 1986.

Woodcock, George. *William Godwin.* London: Porcupine Press, 1946.

D. H. Monro (1967)
Bibliography updated by Philip Reed (2005)

GOETHE, JOHANN WOLFGANG VON
(1749–1832)

Johann Wolfgang von Goethe, the German poet, pantheist, novelist, and scientist, was born in Frankfurt am Main and died in Weimar. Goethe's literary genius disclosed itself early. He wrote numerous lyric poems, invariably inspired by love affairs, while still in his teens. University studies in Leipzig and Strasbourg were less important to his development than were his extracurricular interests: occult philosophy, astrology, and religious mysticism while in Leipzig; and his friendship with Herder at Strasbourg, a friendship that evoked Goethe's passion for William Shakespeare, nature, and German folk poetry. The historical drama *Götz von Berlichingen,* written while Goethe was a law student in Strasbourg, marks the start of his *Sturm und Drang* ("storm and stress") period. *Die Leiden des jungen Werthers* (The sorrows of young Werther, 1774), written to purge himself of the despair engendered by his love for Charlotte Buff, who married another man, marks the high point of this phase of Goethe's career. *Werther,* translated into numerous languages, made Goethe famous throughout Europe. Other works belonging to this period were the dramas *Stella, Egmont,* and the "Gretchen" episodes of *Faust.*

In 1775, at the invitation of Karl August, Duke of Saxe-Weimar, Goethe moved to the court at Weimar. Here, in addition to his work as chief of state and his continued literary activity, Goethe's interest in the sciences developed: His official duties involved such diverse matters as horticulture, mining, road inspection, and later the management of the state theater. In Weimar, Goethe's involvement with Frau Charlotte von Stein, an intellectual lady of refined tastes in the arts, lasted for twelve years. His writings during those years included some of his greatest lyrics. It is said that Stein exercised a humanizing, moral influence on Goethe.

Goethe's trip to Italy in 1786 was to his own mind the climax of his life. In his thinking about art and literature, the classical ideal of calm beauty replaced the representation of tempestuous emotion and rebelliousness characteristic of the *Sturm und Drang* movement. *Iphigenie auf Tauris* (1787), a verse reworking of an earlier play, and *Torquato Tasso* (1789) exemplify the new style.

Returning to Weimar, Goethe took a new mistress, Christiane Vulpius, who bore him a son in 1789 and whom he married in 1806. Many of Goethe's scientific studies were published in this period: *Versuch, die Metamorphose der Pflanzen zu erklären* (Essay on the meta-

morphosis of plants, 1790), *Beiträge zur Optik* (Contributions to optics; 1791 and 1792). Earlier he had published his discovery that a part of the human jawbone is analogous to the intermaxillary bone in apes (1784). Goethe returned to Italy in 1790 but did not find the excitement and inspiration of his earlier travels. In 1792 he accompanied Karl August in a battle against the French revolutionaries. In 1794 began Goethe's friendship—more literary and intellectual than personal—with Friedrich Schiller, which lasted until Schiller's death in 1805. Schiller was a sympathetic critic and he encouraged Goethe's work on *Faust.* It has been thought that Schiller's Kantian background stimulated Goethe's interest in Immanuel Kant, but Goethe was familiar with Kant's writings even before 1794.

While the political and social tumult of the Napoleonic era dominated the minds of his contemporaries, Goethe calmly concentrated his attention on optics and plant morphology. Perhaps as a result of Goethe's indifference to the popular causes of nationalism and democracy, his reputation declined somewhat, but the appearance of *Faust* (Part I) in 1808 and the psychological novel *Die Wahlverwandtschaften* (Elective affinities) in 1809 served to restore his stature. Some of Goethe's subsequent works were *Zur Farbenlehre* (Toward the theory of colors; 1810), which contains an extended attack on Isaac Newton's theory of light; *Dichtung und Wahrheit* (Poetry and truth, 1811, 1812, 1814, and, posthumously, 1833), a series of autobiographical essays; *Italienische Reise* (1816–1817), the record of his Italian travels; *Zur Morphologie* (1817–1824); and the second part of *Faust,* completed in 1831, just before his death. Goethe was buried in Weimar beside Schiller.

PHILOSOPHY

Although Goethe was not a systematic thinker and even asserted that philosophy only ruined him for poetry, he was aware of the philosophical and scientific tendencies and controversies of his time; and while he admitted his lack of a "proper organ for philosophy," he did not hesitate to express himself on numerous philosophical and scientific questions. In addition to specific essays and pronouncements, his poems and novels were often vehicles for expressing his intellectual convictions concerning God, man, and nature.

SPINOZA AND GOETHE. The influence of Benedict Spinoza on Goethe's overall *Weltanschauung* was considerable, although the importance of Kant, Gottfried Wilhelm Leibniz, and Friedrich Schelling is also evident. Goethe first became slightly acquainted with Spinoza's philosophy while in Strasbourg, but it was in 1774 that his acquaintance with Friedrich Jacobi (who regarded Spinoza's views as the only rational philosophy) drew his full attention to Spinoza. Goethe's commitment to pantheism is often cited to show his agreement with Spinoza. Yet when Goethe himself spoke of his relation to Spinoza, he emphasized the ethical as much as the metaphysical doctrines of Spinoza and Spinoza's "all-harmonizing peace," which contrasted with his own restlessness. Spinoza's rejection of final causes and his defense of determinism and of the view that praise, condemnation, and regret are attitudes reflecting an inadequate understanding of inexorable natural processes were accepted by Goethe and given expression in *Faust* (especially in the opening scene of Part II). Goethe said that Spinoza's mathematical method was the opposite of his own poetic way of feeling and expressing, and that Spinoza's orderly treatment of moral questions made Goethe his passionate disciple and convinced admirer. He defended Spinoza against the charge of atheism and claimed (without slavish regard for accuracy) that Spinoza was the most theistic and Christian of philosophers, since for him all existence is God and thus no proof of God's existence is needed.

The central thesis of Spinoza's system, Goethe thought, was that the universe contains and expresses a creative force which appears as a duality (*Zweiheit*) but is in fact a unity. God is not simply the cause but the indwelling spirit of the world, the all-embracing actuality. Goethe, however, questioned Spinoza's contention that reason can attain an adequate knowledge of God-nature. We cannot comprehend this infinite whole, and when we attempt to do so, even in a limited way, we must use imagination and intuition, not the method of mathematics.

LEIBNIZ AND GOETHE. While Goethe's view of nature was, like Spinoza's, deterministic and nonteleological, his mystical feeling for nature was more akin to Schelling, and he resembled Leibniz in maintaining that everything in nature is in some sense animate (*Beseelt*). The universe consists of an infinite number of unique beings—Leibnizian monads—each alive and harmonious with all others. The essence of these individuals is activity and creativity. Goethe's knowledge of Leibniz was probably derived from Goethe's friend Johann Kaspar Lavater, the Swiss theologian who linked his theory of phrenology with the theory of Leibniz's monadology, and from the Earl of Shaftesbury, whose presentation of Leibnizian optimism involved the belief, so congenial to Goethe, that nature's physical beauty expresses the divine harmony.

KANT AND GOETHE. Goethe was inclined to take from philosophers whatever elements or fragments fitted his intuitions and feelings. Thus, while he found Leibniz's confident optimism appealing, he also praised Kant for destroying the popular optimistic teleology of commonsense philosophers who with Philistine wisdom sought to demonstrate that everything in nature exists to satisfy some human purpose. Goethe's enthusiasm for Kant was mainly based on Kant's *Critique of Judgment.* He was pleased with Kant's claim that nature and art both resemble purposive agents but pursue no external goal. He maintained also that art mediates between nature and freedom, since it is produced by the artist in conformity with principles that operate in nature as well.

Goethe, like many of Kant's contemporaries (including Moses Mendelssohn and Johann Gottfried Herder), had little understanding of the *Critique of Pure Reason*; and while he praised Kant's ethics, he rejected most of Kant's central claims. In particular, he denied the opposition of duty and inclination, reason and sensuality, and regarded Kant's Calvinistic notion of a radical evil in human nature as a sad regression toward Christian orthodoxy. Goethe also took exception to Kant's view of knowledge. He insisted that imagination (*Phantasie*) was an avenue to knowledge distinct from and supplementary to Kant's faculties of reason, understanding, and sensibility. Furthermore, Goethe held, men are capable of intellectual intuitions, and with such nonsensuous insights they may hope to penetrate the heart of nature.

SCIENTIFIC THEORIES

Goethe thought his scientific theories were as important as his literary works. The concepts of primal phenomena (*Urphänomen*) and primal polarity (*Urpolarität*) were central to his conception of the world and were the foundations for both his scientific studies and his conception of man and existence.

PRIMAL PHENOMENA. Nature's secrets can only be understood by discovering, through intellectual intuitions, her ideal: ground phenomena. In optics the primal phenomenon is the opposition or antipathy of light and darkness. This *Urphänomen* (which in this instance is also an example of polarity) is the goal and limit of a scientific investigation of light. In mineralogy and geology the *Urphänomen* is granite, which Goethe believed to be the base of Earth's crust.

In the organic realm there are primal shapes and modes of development that nature repeatedly uses, like a theme and variations in music. The same organ is transformed manifoldly through metamorphosis. In plants the leaf is the organ that is varied to form all the parts of the plant. The study of the basic formations, morphology (*Gestalten*), would disclose the secret principles according to which nature operates. Seeking the primal image or idea by observing and comparing the metamorphosis of organisms, Goethe conjectured that a primal plant (*Urpflanze*), might be the basic model according to which all plants are patterned. This theory has sometimes been cited to show Goethe as a forerunner of the theory of evolution, but it is not at all clear that he believed in the historical evolution of species from a common ancestor. The doctrine of the *Urpflanze* is more Platonistic and, perhaps, mythical than Darwinian, Charles Darwin's reference to Goethe as a "path-maker" notwithstanding.

POLARITY. Goethe's distrust of mathematics and experimental instruments (such as prisms) was, unfortunately, great. He believed that numbers and equations only distort our vision of nature. Isaac Newton's physics was repellent to him; Newton's theory that white light contained the spectrum seemed to him absurd because light was an elemental entity, an inscrutable attribute of the world that could not be analyzed. Goethe attempted to explain the origin of color phenomena out of an original polar opposition of light and dark. If light and dark are mixed directly, the result is gray; but a "murky medium" (such as a prism, according to Goethe) produces a cooperation of the polar opposites, and this cooperation produces colors. The activity of the eye in color perception is explained by the rule that brightness is "demanded" when the eye encounters darkness. The perception of every color produces a "demand" for the complementary color.

Goethe used extensively the idea of polarity, of attraction and repulsion as basic cosmic forces. He explained the metamorphosis of plants in terms of the periodic alternation of contraction and expansion. Contraction (*systole*) produces specific differentiation; expansion (*diastole*) produces an "advance into the infinite." The importance of polarity is seen also in magnetism—another *Urphänomen*—and in the activity of the heart, the rhythm of life, and in man's moral activity where the good is brought into play by its contrary, evil. Schelling had said that there is no life without opposites, and Kant had claimed attraction and repulsion to be the only essential forces of matter. Goethe adopted these principles in both his science and his art. The whole of existence is "an eternal parting and uniting."

STRIVING. Polarity is one of the driving wheels of nature, and while Goethe's use of this idea may suggest a

cyclical view of life and history, his concept of gradation (*Steigerung*) is that of a constantly striving ascent. This upward striving Goethe believed to be a universal characteristic of nature. It discloses itself in the "higher intention" of every heavenly body and in the variations of similar organisms developing from a basic form. What Goethe meant by this is not very clear, but in *Faust* the idea is applied to man. Every man, said Goethe, innately feels an urge to strive upwards. This striving involves all his capacities, his creativity in every sort of action and experience. Faust's insatiable love of life and hunger for new experience are expressive of this natural longing.

RELIGION

Goethe early rejected positive religion. With Spinoza, he came to regard creeds and dogmas as irrelevant to the veneration of God-nature. Although Goethe was on friendly terms with many ardent Christians and although he even spoke at times of a providential God, he opposed the dogmatism of churches and theologians and regarded the idea of miracles as a "blasphemy against the great God and his revelation in nature." Since Goethe maintained that no set of concepts could be adequate to the unfathomable infinity of the divine, it is not surprising that his pronouncements concerning God are somewhat ambiguous and inconsistent. While the remark on miracles seems to imply a distinction between God and nature, this is, of course, not Goethe's usual position.

Goethe rejected asceticism and the tendency to devalue the physical in favor of a supernatural world. To Johann Kaspar Lavater he wrote that he could find a thousand pages of various books as lovely, useful, and indispensable to humankind as the Gospels. He claimed that he was unchristian rather than anti-Christian but declared the crucifix to be "the most repugnant thing under the sun." Although at one time he spoke of the Gospels as messages from God, he clearly did not intend this in the ordinary sense, since he held that God, being the inexorable order of nature, cannot have any personality or be in any sense outside the natural world. Thus God does not cause or control the world in the way that theists have believed. "What sort of God would it be, who only pushed from without?" (*Was wär' ein Gott, der nur von aussen stiesse?*, in *Weltanschauliche Gedichte,* 1815). The ambiguity (or richness) of Goethe's theology may be seen in what is perhaps his most famous remark on this topic: "We are pantheists when we study nature, polytheists when we poetize, monotheists in our morality" (*Wir sind naturforschend Pantheisten, dichtend Polytheisten, sittlich Monotheisten,* in *Maximen und Reflexionen,* No. 807).

Since every man is part of nature and, hence, of the divine, he shares the basic impulses of all natural things—specifically, as already noted, the urge to develop upward and outward, the striving for an ideal. Action and striving are not only means to some static goal but are also ends in themselves. Since there is no goal for man apart from his life, man struggles, like Faust, with the fear of life (*Lebensangst*) and is tempted by care (*Sorge*). Some have argued that Goethe saw man's fulfillment in activity itself, but perhaps it would be more accurate to say that there is no fulfillment—contentment means annihilation—so that man is destined to be dissatisfied, unfulfilled, no matter what he achieves.

See also Darwin, Charles Robert; Herder, Johann Gottfried; Jacobi, Friedrich Heinrich; Kant, Immanuel; Lavater, Johann Kaspar; Leibniz, Gottfried Wilhelm; Mendelssohn, Moses; Newton, Isaac; Pantheism; Pantheismusstreit; Schelling, Friedrich Wilhelm Joseph von; Schiller, Friedrich; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Spinoza, Benedict (Baruch) de; Spinozism.

Bibliography

Goethe's writings are available in many editions, one of the standard being K. Burdach and others, eds., *Jubiläums-Ausgabe,* 40 vols. (Stuttgart and Berlin, 1902–1907). Goethe's most important works, except *Zur Farbenlehre* and *Beiträge zur Optik,* have been translated into English. Goethe's *Botanical Writings* were translated by Bertha Mueller (Honolulu: University of Hawaii Press, 1952) and by Agnes Arber as *Goethe's Botany; the Metamorphosis of Plants* (Waltham, MA, 1946). Of the dozens of English translations of *Faust,* recent versions by Bertram Jessup (London, 1958) and by Walter Kaufmann (New York: Doubleday, 1961) are sensitive. Kaufmann's version includes sections of Part II and has the German text facing the English.

For an account of *Zur Farbenlehre* and *Beiträge zur Optik,* see Rudolf Magnus, *Goethe als Naturforscher* (Leipzig: Barth, 1906), translated by Heinz Norden as *Goethe as a Scientist* (New York: H. Schuman, 1949), a very sympathetic account. A more critical discussion of Goethe's contributions to science is found in C. S. Sherrington, *Goethe on Nature and on Science* (Cambridge, U.K.: Cambridge University Press, 1942).

Many philosophers have written on Goethe (Wilhelm Dilthey, Georg Simmel, Max Wundt, Heinrich Rickert, and Fritz-Joachim von Rintelen, to name only a few), and some may be read in English: Ernst Cassirer, "Goethe and the Kantian Philosophy," in his *Rousseau, Kant, Goethe,* translated by James Guttmann and others (Princeton, NJ: Princeton University Press, 1945); Benedetto Croce, *Goethe* (Bari:

Laterza, 1919), translated by Emily Anderson (London, 1923); and the beautiful and perceptive chapter on *Faust* in George Santayana, *Three Philosophical Poets* (Cambridge, MA: Harvard University Press, 1922). Thomas Mann's essays on Goethe, which include discussions of Goethe's political views, are in his *Freud, Goethe, Wagner,* translated by H. T. Lowe-Porter (New York: Knopf, 1937) and in his *Essays of Three Decades* (New York: Knopf, 1947). Karl Viëtor, *Goethe the Thinker,* translated by B. Q. Morgan (Cambridge, MA: Harvard University Press, 1950), discusses Goethe's science and philosophy comprehensively. See also Fritz-Joachim von Rintelen, *Der Rang des Geistes* (Tübingen: Niemeyer, 1955).

Arnulf Zweig (1967)

GOGARTEN, FRIEDRICH

(1887–1968)

Friedrich Gogarten, the German theologian, was born in 1887 at Dortmund. After serving as a pastor in Thuringia, in 1927 he became professor of systematic theology at Jena and in 1935 moved to the corresponding chair at Göttingen. He was early associated with the new dialectical theology and its revolt against liberalism and idealism. Within this movement he stands nearer to Rudolf Bultmann than to Karl Barth, but he worked out a distinctive position of his own. His thought shows the influence of existentialist philosophy, but he claimed that it also continues the insights of Martin Luther, on whom Gogarten was a recognized authority.

Gogarten believed that Luther delivered Christian theology from the hold of metaphysics. This achievement was obscured in the period of Protestant orthodoxy following the Reformation, but it is now time to revive his insights, which can be restated in terms of current existentialist philosophy. According to Gogarten, the major Christian doctrines were formulated under the domination of metaphysical categories, in an age when history was understood as a process that takes place within a stationary metaphysical framework and when the course of history was supposed to be determined by metaphysical factors. Deliverance from metaphysics makes it possible to take history with a new seriousness. Man is responsible for history and creates it by his decisions. So far are we from having an obligation to interpret history in the light of metaphysics that we must rather view metaphysical systems themselves as products of history. Christianity is not dependent on any metaphysical system but is rather the summons to a historical self-understanding, in which we accept responsibility for our own historical existence under the word of God, which addresses us in Christ.

These emphases, which Gogarten relates as much to the *sola fide* of Luther as to modern existentialism, are developed into a secular interpretation of the Christian gospel. The Christian faith brings man to maturity and strips the world of every mythical or numinous property. The world is deprived of its religious power and is handed over to man as the son who has come of age, the heir to whom God has entrusted the creation. These views are related by Gogarten especially to the teaching of St. Paul in Galatians 4:1–11.

See also Barth, Karl; Bultmann, Rudolf; Idealism; Liberalism; Luther, Martin; Metaphysics, History of.

Bibliography

WORKS BY GOGARTEN

Ich glaube an den dreieinigen Gott. Jena: E. Diederichs, 1926.

Illusionen, eine Auseinandersetzung mit dem Kulturidealismus. Jena: E. Diederichs, 1926.

Glaube und Wirklichkeit. Jena: E. Diederichs, 1928.

Politische Ethik. Jena: E. Diederichs, 1932.

Das Bekenntnis der Kirche. Jena: E. Diederichs, 1934.

Die Kirche in der Welt. Heidelberg: L. Schneider, 1948.

Die Verkündigung Jesus Christi. Heidelberg, 1948.

Verhängnis und Hoffnung der Neuzeit. Stuttgart, 1953.

Entmythologisierung und die Kirche. Stuttgart, 1953. Translated as *Demythologizing and History.* London: SCM Press, 1955.

Der Mensch zwischen Gott und Welt. Stuttgart, 1956.

Die Wirklichkeit des Glaubens. Stuttgart: F. Vorwerk, 1957. Translated by Carl Michalson and others as *The Reality of Faith.* Philadelphia: Westminster Press, 1959.

Gogarten also edited various works of Luther and contributed extensively to *Zwischen den Zeiten,* the journal of the dialectical theology group.

WORKS ON GOGARTEN

Siegfried, T. *Die Theologie der Existenz bei Friedrich Gogarten und Rudolf Bultmann.* Gotha, 1933. The most satisfactory critical exposition of Gogarten's thought.

John Macquarrie (1967)

GOLDEN RULE

One early use of the word *golden* in English is "most excellent, important, or precious." With reference to rules or precepts it was used to mean "of inestimable value," and the expression "the golden rule" was often specifically used with reference to the precept in Matthew: "Therefore all things whatsoever ye would that men should do to you, do ye even so to them: for this is the law and the prophets" (7:12). Thus, the principle that has come to be known as the golden rule has been so called presumably

because it has been regarded as being of inestimable value or importance. This regard was not derived solely from the fact that it was set forth in the sermon on the mount. The golden rule has been widely accepted, in word if not in deed, by vast numbers of greatly differing peoples; it is a basic device of moral education; and it can be found at the core of innumerable moral, religious, and social codes. So far as can be determined from available records, it was probably first formulated by Confucius some five hundred years before Christ—"What you do not like when done to yourself do not do to others"—and the multitude of different formulations testify to its widespread acceptance and influence.

There is probably no principle which has been so widely accepted and remained so controversial. Nonetheless, the golden rule has been the subject of comparatively little philosophical discussion. It is usually mentioned, when it is mentioned at all, only in passing, and it has generally received more attention in theological and inspirational literature. However, there are signs of increasing philosophical interest in it.

One of its commonest formulations today is "Do unto others as you would have them do unto you." It is commonly supposed that there are significant differences between this, the positive formulation, and the negative formulation, "Do not do unto others what you would not have them do unto you," and that the positive formulation "marks a distinct advance" since it "prescribes positive services rather than mere abstinences" and "sets forth an ideal which is higher and therefore more difficult to realize." It can be argued, however, that this is an error resulting from faulty analysis and perhaps also from theological bias. In connection with a specific action or object of desire, there is a considerable difference between a positive desire, a desire to do it or have it done to oneself, and a negative desire, a desire not to do it or not to have it done to oneself. But in the abstract, so the argument runs, there is only a difference in formulation, and a want, wish, or desire formulated in negative terms can always be reformulated in positive terms. For example, there is no difference between not wanting others to lie to oneself and wanting them not to lie to oneself, wanting them to tell one the truth and wanting them not to fail to tell one the truth. In general, "*A* wants *x* to happen" is equivalent to "*A* does not want *x* not to happen," and "*A* does not want *x* to happen" is equivalent to "*A* wants *x* not to happen." Thus, according to this line of argument, every desire formulated negatively, which would come within the scope of the negative golden rule, can be reformulated positively and will then come within the scope of the positive golden rule. It would follow, then, that there is no logical or moral difference between the negative and positive formulations, only a psychological or rhetorical one.

On either account the negative formulation of the golden rule is to be distinguished from the denial of the golden rule: "Do not do unto others as you would have them do unto you." Obviously, this is not a formulation of the golden rule at all but is, rather, its total rejection. The denial of the golden rule is usually supported by the claim that the golden rule presupposes a uniformity of human nature, in the sense of a uniformity of tastes, interests, needs, and desires, and the attendant claim that there is no such uniformity. One way of meeting this objection is to deny that the golden rule involves any such presupposition. It has been argued that it is necessary to make a distinction between the particular interpretation and the general interpretation of the golden rule. The particular interpretation implies that whatever in particular one would have others do to or for him, he should do to or for them. It is in the particular interpretation that, to take some of the standard objections, the golden rule "authorizes the quarrelsome person who loves to be provoked, to go about provoking others, and the person who hates friendliness and sympathy to be cold and unsympathetic in his dealings with others" (L. J. Russell). But these consequences, it has been claimed, do not follow from the general interpretation. On this interpretation what one has to consider is not what in particular one would have others do to or for oneself but, rather, the general ways in which one would have others act in their treatment of oneself. If one abstracts his general wishes from his particular desires, what one would have others do is to take account of his interests, needs, and desires, which may be quite different from theirs, and either satisfy them or not willfully frustrate them. What the golden rule requires a person to do, then, is to take account of the wishes of others and accord them the respect and consideration he would want them to accord to his. In other words, what the golden rule requires of each of us is that we should treat others in accordance with the same principles or standards that we would have others apply in their treatment of us. Thus, the golden rule, if this argument is sound, is compatible with differences in interests, needs, tastes, wishes, and desires and does not presuppose that human nature is uniform in the sense specified.

Another principle which should be distinguished from the golden rule is what might be called its inversion: "Do unto others as they would have you do unto them." The inversion of the golden rule has received some sup-

port, and it has even been urged that it replace the golden rule as a guide to conduct, mainly as a consequence of the same sort of objection as the distinction just outlined is intended to eliminate. It has been claimed that the inversion of the golden rule has "the merit of stressing the need for an understanding of other people as a basis of our behavior toward them" and does not presuppose any uniformity or identity of nature in the beings it is intended to govern.

One counterargument to this is that the implications of the inversion of the golden rule are more absurd than the alleged implications of the golden rule itself and that it is tantamount to a rule that would require everyone always to do whatever anyone else wants him to do, a rule it is impossible to follow in a world of conflicting interests. Once it is recognized, the argument runs, that the "uniformity of human nature," in the sense of an absolute identity of interests, needs, and desires, is not a presupposition of the golden rule, any temptation to substitute the inversion of the golden rule for the golden rule itself should disappear. For in its general interpretation the golden rule does require us to take account of and accord respect to the differing needs, interests, and desires of others, and it is just this that the inversion of the golden rule is intended to bring about. However, the question remains whether the inversion of the golden rule cannot be rescued from at least some of the more obvious objections to it by means of a distinction similar to that made between the particular and the general interpretation of the golden rule.

In the course of time a number of anomalous interpretations of the golden rule have found strong support. On the one hand, it has been said that the golden rule comprehends all the requirements of morality in a single formula; on the other, it has been said that the golden rule is only a guide, that it is far from complete, that it requires rules, a sense of justice, or even a whole system of morality for its proper interpretation and application. Again, the golden rule has been said to be not only consistent with but actually to comprehend all of utilitarianism; it has also been said to provide just that element, the requirement of justice or fairness, that is alleged to be most lacking in a utilitarian theory. On this interpretation the golden rule is regarded as being the basis of justice, sometimes also the basis or equivalent of Immanuel Kant's categorical imperative. Finally, it has been claimed that the golden rule is a perfect guide to conduct and that the only thing needed to make the world perfect is for everyone to follow it; at the same time it has been claimed that the golden rule leads to paradoxes and is misleading, false, or absurd.

Each of the points and issues mentioned here is discussed, more or less adequately, in one or more of the sources listed in the bibliography. But no one has yet dealt satisfactorily with the question of why this precept should have appeared in the codes and outlooks of so many diverse peoples and sages. The golden rule, in one version or another, has a prominent place in all the major religions and most minor ones; it has been enunciated by pagan philosophers both before and after Christ and by Sophists (Isocrates) and anti-Sophists (Aristotle). There are no detectable historical traces that could explain this, and the historical diffusion theory is worthless as an explanation here. The nearly universal acceptance of the golden rule and its promulgation by persons of considerable intelligence, though otherwise of divergent outlooks, would therefore seem to provide some evidence for the claim that it is a fundamental ethical truth.

See also Aristotle; Confucius; Ethics and Morality; Sophists.

Bibliography

Baier, Kurt. *The Moral Point of View.* Ithaca, NY: Cornell University Press, 1958. Pp. 202–211.

Browne, Lewis. *The World's Great Scriptures.* New York: Macmillan, 1946. P. xv.

Butler, Joseph. *Sermons.* Oxford, 1897. Sermon X, Sec. 18, pp. 154–155.

Cadoux, A. T. "The Implications of the Golden Rule." *International Journal of Ethics* 22 (1912): 272–287.

Dihle, Albrecht. *Die Goldene Regel, eine Einführung in die Geschichte der antiken und frühchristlichen Vulgärethik.* Göttingen, 1962. Vol. VII of *Studienhefte zur Altertumwissenschaft.*

Garnett, A. Campbell. *Ethics.* New York, 1960. Pp. 92, 126.

Goldstein, Morris. *Thus Religion Grows.* New York: Longmans, Green, 1936. Pp. 118, 127.

Gould, James A. "The Not-So-Golden Rule." *Southern Journal of Philosophy* 1 (1963): 10–14.

Hare, R. M. *Freedom and Reason.* Oxford: Clarendon Press, 1963. Passim.

Hertzler, J. O. "On Golden Rules." *International Journal of Ethics* 44 (1934): 418–436.

Hirsch, Emil R. "The Golden Rule." In *The Jewish Encyclopedia.* New York and London, 1904. Vol. VI, pp. 21–22.

Hume, Robert Ernest. *The World's Living Religions.* New York: Scribner's Sons, 1927. Pp. 265–267.

MacIver, Robert M. "The Deep Beauty of the Golden Rule." In *Moral Principles of Action,* edited by Ruth Nanda Anshen. New York: Harper, 1952.

Ramsey, Paul. *Basic Christian Ethics.* New York: Scribners, 1950. P. 61.

Russell, L. J. "Ideals and Practice." *Philosophy* 17 (1942): 109–110.

Sharp, Frank Chapman. *Ethics.* New York: Century, 1928. Pp. 140–141, 334–339.

Sidgwick, Henry. *The Methods of Ethics.* 7th ed. London: Macmillan, 1930. Pp. xvii, 379–380, 384–385.

Sidgwick, Henry. *Outlines of the History of Ethics.* 6th ed. London: Allen & Unwin, 1931. P. 167, note.

Singer, M. G. "The Golden Rule." *Philosophy* 38 (1963): 293–314.

Spooner, W. A. "Golden Rule." In *The Encyclopedia of Religion and Ethics,* edited by James Hastings. Vol. VI, 310–312. New York: Scribner's Sons, 1914.

Tasker, J. G. "Golden Rule." In *A Dictionary of Christ and the Gospels,* edited by James Hastings. Vol. I, 653–655. New York: Scribners, 1907.

Thomas, George F. *Christian Ethics and Moral Philosophy.* New York: Scribners, 1955. Pp. 48, 55, 429–430.

Tsanoff, Radoslav A. *Ethics.* 2nd ed. New York: Harper, 1955. Pp. 31–33.

Tsanoff, Radoslav A. *The Moral Ideals of Our Civilization.* New York: Dutton, 1942. Pp. 171, 234, 414, 421, 496, 512.

Votaw, C. W. "Sermon on the Mount." In *A Dictionary of the Bible,* edited by James Hastings. Vol. V, 41–42. New York: Scribners, 1904.

Weiss, Paul. "The Golden Rule." *Journal of Philosophy* 38 (1941): 421–430.

Weiss, Paul. *Man's Freedom.* New Haven, CT: Yale University Press, 1950. Pp. 138–154.

Westermarck, Edward. *The Origin and the Development of the Moral Ideas.* London: Macmillan, 1906. Vol. I, pp. 102–103.

Whately, Richard. *Lessons on Morals.* Cambridge, MA, 1857. Ch. 4.

Wheelwright, Philip. *A Critical Introduction to Ethics.* New York: Odyssey Press, 1949. Pp. 165, 207.

Marcus G. Singer (1967)

GOLDMAN, ALVIN

(1938–)

Alvin Goldman, an American philosopher, is best known for his contributions to epistemology, philosophy of mind, and related fields. His first paper, "A Causal Theory of Knowing," (1967, reprinted in *Liaisons* 1992), defends the view that an individual S knows a proposition p just in case p is causally related in the right sort of way to the individual's belief that p. Thus, for example, Sam knows that there is a cat on the mat because Sam is looking at the cat, and the fact that the cat is on the mat caused Sam to have that belief. This kind of account of knowledge breaks with the tradition that identifies knowledge with some sort of justified, true belief. While Goldman's account requires that a belief be true if it is to count as knowledge, the requirement of justification is replaced with a requirement that highlights the importance of the causal ancestry of the belief. Goldman further develops this view in "Discrimination and Perceptual Knowledge" (1976) and "What Is Justified Belief?" (1979) (both reprinted in *Liaisons*), coming to hold, in the latter paper, that knowledge does indeed require justification, where justification is to be identified with reliably produced belief rather than with any kind of ability to produce an argument, as traditional accounts require. This style of account has come to be known as "externalist" (because the factors in virtue of which a belief is justified may be external to the knower's mind), and is opposed to the more traditional "internalist" accounts on which the features in virtue of which a belief is justified are ones to which the knower inevitably has cognitive access. Goldman develops this view in tremendous detail in a series of papers, and ultimately in *Epistemology and Cognition* (1986).

Whereas Goldman's account of knowledge is offered as an analysis of the concept of knowledge, the substance of his account places a great deal of stress on the relevance of empirical work to epistemological issues. Thus, Goldman's approach prompts him to investigate the various psychological mechanisms by which belief is produced because it is upon the reliability of these mechanisms that people's status as knowers depends. This concern with the ways in which empirical work—and especially work in the cognitive sciences—may be brought to bear in advancing human understanding of traditional philosophical issues is characteristic of Goldman's work generally; his work in this area constitutes the most sustained development of naturalistic epistemology available.

Although his early work was concerned with the philosophy and psychology of individual cognition, Goldman has gone on to make seminal contributions to social epistemology. The mechanisms by which beliefs are produced and sustained include not only those inside the knower's head, but features of the social organization of the knower's epistemic community. In *Knowledge in a Social World* (1999), Goldman investigates the ways in which social structures may either contribute to, or interfere with, the discovery and dissemination of truths. This project includes work on the epistemology of testimony and argumentation, the social structure of scientific investigation, and the epistemology of education. Additionally, Goldman addresses questions about democracy, government regulation of speech, the role of truth in legal proceedings, and the economics of communication—all topics illuminated by his epistemological approach.

Goldman has also made important contributions to the philosophy of mind, especially in his elaboration and development of the "simulation" account of mental state attribution. A standard approach to mental state attribution, now known as the "theory-theory," sees human attribution of mental states to others as the product of theory construction. On this view, when one forms the belief that Jack will want the university president to resign, that person's belief about Jack's desire is derived from beliefs held about Jack's other mental states, together with theories the person holds about the laws governing interactions among mental states. On the simulation view, however, attribution of mental states does not derive from theory construction and need not involve any beliefs about psychological laws or regularities. Instead, the processes by which one's own mental states interact are brought to bear on the task of mental state attribution, being used to simulate the workings of the process by which the target mental state was produced. Goldman's *Simulating Minds: The Philosophy of Psychology, and Neuroscience of Mindreading* (forthcoming) develops this view in detail. He assembles evidence from psychology and especially neuroscience of low-level, automatic processes that mimic, mirror, or resonate with those of an observed other. Such processes play a crucial role in the facial mind-reading of emotions, for example. On the topic of the self-attribution of mental states, Goldman defends an introspectionist approach—in contrast with other simulationists. Elsewhere, he defends introspection as the basis for relying on subjects' verbal reports in the science of consciousness.

In his first book, *A Theory of Human Action* (1970), Goldman defended a fine-grained approach to the ontology of action and illuminated the relationship between determinism and fatalism. He has also explored the interface between metaphysics and cognitive science.

See also Epistemology, History of; Philosophy of Mind.

Bibliography

WORKS BY GOLDMAN

A Theory of Human Action. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Epistemology and Cognition. Cambridge, MA: Harvard University Press, 1986.

Liaisons: Philosophy Meets the Cognitive and Social Sciences. Cambridge, MA: MIT Press, 1992.

Philosophical Applications of Cognitive Science. Boulder, CO: Westview Press, 1993.

Knowledge in a Social World. Oxford: Oxford University Press, 1999.

Pathways to Knowledge. Oxford: Oxford University Press, 2002.

Simulating Minds: The Philosophy of Psychology, and Neuroscience of Mindreading. Oxford: Oxford University Press, forthcoming.

SECONDARY REFERENCES

"The Philosophy of Alvin Goldman." *Philosophical Topics* 29 (1, 2) (2001).

Hilary Kornblith (2005)

GONGSUN LONG

(320–250 BCE)

Gongsun Long was a logician in ancient China and a representative figure of the School of Names (*Ming-Jia*). What distinguishes Gongsun Long's work is his in-depth investigation into the relation between names and reality through conceptual analysis and rational argumentations. His thoughts are delivered in the *Gongsun Longzi*. Three brief essays in the text, "On the White Horse," "On Referring to Things," and "On Hardness and Whiteness," are considered most important in understanding his thoughts. The first one is considered the most philosophically interesting and influential in view of its substantial philosophical points, its articulate character of rational argumentation, and its sophistication.

Gongsun Long's well-known thesis "[the] white horse [is] not [the] horse" (*bai-ma-fei-ma*) is supported by several articulate arguments in the essay "On the White Horse." Modern scholars elaborate their substantial contents and philosophical significance through seemingly competing interpretations. Fung Yu-lan (1952–1953) renders Gongsun Long a Platonic realist; he considers that all of Gongsun Long's arguments are intended to argue that "white horse" and "horse" represent two distinct Platonic universals and thus the universal of white-horseness is not (identical to) the universal of horseness. One criticism is that Fung seems to impose his Platonic realist reading on the thoughts of a figure in the Chinese tradition whose general mentality and language characteristics have not tended to nourish a Platonic outlook of the universe.

Janusz Chmielewski (1962) takes a set-theoretic line: "white" and "horse" are used to denote distinct classes, and "white horse" denotes the intersection of the two classes, which is an empty class, instead of a subclass of either of the two classes. One major difficulty with Chmielewsky's interpretation is that it obviously deviates from the original text in which Gongsun Long clearly indicated that there are white horses.

Chad Hansen (1983) proposes a radical shift of interpretation based on mereology (part-whole logic) and his mass-noun hypothesis: The term "white horse" is a mass noun and refers to a mass *sum* whole of horse-stuff part and white-stuff part, distinguishing it from a mutually pervasive compound, like hard-white, that is a mass *product*; the whole of white part and horse part is not its horse part. Angus C. Graham (1990) endorses Hansen's mereological interpretation though without being committed to the mass-noun hypothesis. The Hansen-Graham radical mereological interpretation is to bypass class-member relation but resorts to whole-part division alone. Nevertheless, an interpretation that renders Chinese thinkers short of conceptual abstraction intrinsically involved in member-class relation is questionable.

The previous interpretations share one feature: Their interpretations of the semantic reference of those common nouns like "white horse" and "horse" seem to variously derivate from the semantic structure as embedded in actual language practice, in which Chinese common nouns are normally used to denote (a collection of) particular things (including particular properties) via their conceptual contents. A modest mereological interpretation with a collective-noun hypothesis might be reasonable for the sake of capturing the semantic structure. That is, (1) the denotational semantics and deep structure of Chinese common nouns are like those of collective nouns; their implicit ontology is a mereological one of collection-of-individuals with both part-whole structure and member-class structure. (2) The denotation of "white horse" is neither a Platonic universal nor a sum of horse stuff and white stuff nor an empty set, but a collection of white horses. (3) The collection of white horses is both a mereological whole and a class; the part-whole relation here is also the relation between subclass and class that accommodates conceptual abstraction and can be specified in terms of Fregean sense. From this point of view Gongsun Long argues for the thesis that what "white horse" denotes (the collection of white horses) is not identical to ("*fei*") or differs from ("*yi*") what "horse" denotes (the collection of horses) in view of their distinct conceptual contents, distinct extensions and ddistinct necessary-identity-contributors.

Although Gongsun Long emphasizes distinct aspects of things, he does not ignore common aspects and connections of things and thus explicitly indicates, "It is when what is pursued is their common aspect that the white horse [as a subclass] is [is included in the class of] the horse" The previously mentioned class-mereological nature of the denotation of collective nouns allows a flexible shift between the identity-relation and the class-inclusion relation between two collections, depending on whether the speaker's focus is on distinct aspects or on common aspects of things: The point of the referring subject's focus shift is related to one central point made in his essay "On Referring to Things": those relevant contributing elements involved in the referring subject's act of referring via a name (such as her purpose or focus) make their intrinsic contributions to the identities of things that are referred to.

In "On Hardness and Whiteness" Gongsun Long investigates the metaphysical status of properties themselves like hardness and whiteness, which are mutually pervasive in the hard white stone, by examining their separability. He thinks that the property itself, say, hardness, can stand alone in the form of *zi-cang* (being self-hidden) but also maintains that "there exists no such hardness in the world [except of its manifestations in particular things]" (the author's translation). The question is this: How could the property per se be self-hidden, in some realist way or in some conceptualist way or in a radical nominalist way? The text of that essay seems to be open to distinctive interpretations. This is a controversial issue that needs to be examined in the context of Gongsun Long's whole thought.

See also Logic, History: Chinese Logic; Mereology.

Bibliography

Chmielewski, Janusz. "Notes on Early Chinese Logic (I)." *Rocznik Orientalistyczny* 26 (1) (1962): 7–21.

Fung Yu-lan. *A History of Chinese Philosophy*. 2 vols. Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1952–1953.

Graham, Angus C. "Three Studies on Gongsun Lung." In *Studies in Chinese Philosophy and Philosophical Literature*. Albany: SUNY Press, 1990.

Hansen, Chad. *Language and Logic in Ancient China*. Ann Arbor: University of Michigan Press, 1983.

For further elaborations of the Platonic realist interpretation, see Chung-yin Cheng, "Kung-sun Lung: White Horse and Other Issues," *Philosophy East and West* 33 (1) (1983): 341–354; and Yiu-ming Fung, *On the Gongsun Longzi* [Gong-Sun-Long-Zi] (Taipei: Dongda, 2000) (in Chinese). Fung also gives a comprehensive discussion of the previously mentioned authors' various representative interpretations.

For a further elaboration of the collective-noun hypothesis for Chinese names to which the last interpretation resorts, see Bo Mou, "The Structure of the Chinese Language and Ontological Insights: A Collective-Noun Hypothesis," *Philosophy East and West* 49 (1) (1999): 45–62.

Bo Mou (2005)

GOOD, THE

For many evaluative terms it is useful to distinguish, following John Rawls, between concepts and conceptions of them; for instance, on the one hand there are questions about what it means to be good, on the other there are questions about which specific things actually are good. This distinction helps explain why many evaluative terms allow much more disagreement than other terms; two people can agree about what it means to be good (concept) and still disagree entirely about which concrete things are good (conceptions), whereas people could hardly agree on what it means to be, say, a knife and still widely disagree about which objects are knives.

Although it is not a distinction drawn by all moral philosophers, this split between concept and conception does set its mark on the literature in the sense that there are two main lines of inquiry about the good: inquiry into the meaning of good and inquiry into which things actually are good. Interest in the first question became prominent in the wake of the publication of G. E. Moore's *Principia Ethica* in 1903; with it moral philosophy took a turn toward conceptual analysis. When it comes to conceptions of the good there have been extensive discussions ever since antiquity and there are primarily two types of goodness that have been the focus of most of them, the prudentially good and the morally good (i.e., what constitutes a good life and what it takes to be a good person).

THE MEANING OF GOOD

Moore claimed that good is indefinable, that it is a simple, nonnatural property. He accused several of his predecessors of having committed what he called the naturalistic fallacy in trying to spell out what it meant to be good. His main argument was that of the open question: For any proposed analysis of good, it would seem that one can ask "But is x good?" and the openness of that question shows that the analysis has not succeeded. It is uncertain whether the philosophers discussed by Moore really had proposed analyses of good rather than simply presented conceptions of the good and it is difficult to find unambiguous examples of theories that are naturalistic in Moore's sense of the word. Nevertheless, Moore did set the agenda for the attempts at analyzing good that have since then followed.

Moore himself is not always clear about whether he discusses the concept of good or the property to which this concept refers. This can make an important difference. If one thinks that the reference is determined by the property or properties that a concept causally tracks, then they can see that, even for normative concepts, there is the possibility that they ultimately refer to natural properties even though they cannot a priori be analyzed in such terms. This kind of naturalism, so-called Cornell realism, was developed in the 1980s by David Brink and others and according to it there is nothing peculiar about the open question being open because the identity involved is synthetic, not analytic.

But there are also other alternatives. Around the same time that Moore developed his theory of the good, Franz Brentano (1969) developed an analysis of good that occupied a halfway point between Moore and naturalism. Moore contended that good was a simple, nonnatural quality; Brentano claimed that it was a complex, nonnatural property, that to be good was to be worthy of love. This kind of position was later elaborated in more detail by A. C. Ewing (1947), who claimed that to be good was to be a fitting object of a pro-attitude. He also thought that different types of goodness could be differentiated through kinds of pro-attitudes, so it was a theory that tried to capture an essential unity of good but also make sense of the variety of uses that good is put to. It is however not an analysis without problems; above all, the key notion of fittingness was not given a satisfactory elucidation by Ewing. Later philosophers, such as Thomas Scanlon (1998) with his buck-passing analysis of value, pursue a similar project, albeit framed in terms of reasons for pro-attitudes rather than the fittingness of them.

The largely Moorean approach of focusing on the meaning of particular evaluative notions was however out of fashion from the 1930s through to the 1960s. The dominant approach to ethical language then was instead non-descriptivist. This is an approach according to which the function of ethical language is not to describe a realm of values and norms, but rather to express sentiments or, as in the best worked-out version of this approach, namely that of Richard Hare (1952), to issue prescriptions. Thus, non-descriptivist analyses of evaluative language do not really provide analyses of the meaning of notions like good, rather what they provide is analyses of what it means to say or judge that something is good. And because ethical judgments appear to have a strong connection to action and motivation, one should perhaps doubt whether predications of good really work essentially like predications of round.

However, even if non-descriptivists eschewed talk about objective evaluative properties, it should be made clear that the position was not as such nihilistic. Rather, what these theories often did was to show that ethical

judgments could make sense even if there really were nothing in the world for them to be about. Hare especially argued at length that one could discuss moral matters rationally even though moral discourse was not descriptive. One problem, however, was that non-descriptivists tended to build their accounts on analyses of asserted evaluative judgments, but talk of good and similar notions does not just consist in that. Evaluative notions can also be embedded in nonevaluative claims, for instance, in conditionals such as "If Peter is good, then Mary is also good." When trying to explain how such sentences function even in very simple forms of deductive reasoning, non-descriptivists tend to be driven to suspiciously complex analyses. The popularity of non-descriptivism has waned considerably since its heyday, although new and sophisticated versions of the theory have been presented after that, most notably by Allan Gibbard.

THE UNITY OF GOOD?

When Brentano proposed his analysis of good it was with the intent of showing the essential unity of good. Many other philosophers have simply taken it for granted. However, even if there is one word that is used in a variety of contexts, that does not necessarily mean that a single concept is being dealt with. Some philosophers, such as Peter Geach (1956), have suggested that one needs to distinguish between a predicative and an attributive use of good. Take a sentence such as "X is an A B." If A is used predicatively, then this sentence can be split into "X is A" and "X is a B"; if this is not possible, then A is used attributively. The sentence "Jill is a dark-haired woman" can be split into "Jill is dark-haired" and "Jill is a woman," whereas "Jill is a tall woman" cannot be split that way. The difference is that there is no way of ascertaining whether Jill is tall without taking into account that she is a woman; it is qua woman that she is tall.

When it comes to goodness, judgments such as "that is a good knife" are clearly attributive, whereas judgments such as "that was a good event" seem predicative. Both uses of good have been prominent in the history of ethics, although some (like Aristotle) have leaned toward the attributive and others (like Moore) toward the predicative. Geach himself contended that, when it makes sense, good is always an attributive adjective; nothing is ever simply good. Relatively few philosophers have followed Geach in taking this stance although the distinction still highlights an important disunity of good. This can be seen if one compares good to valuable. The predicative use of good roughly corresponds to valuable, whereas attributive goodness has a much looser connection to value. For instance, if Jill shoots Jack in the head from a long distance, one might think that it was a good shot without finding the event valuable in any sense.

Although the distinction between the attributive and the predicative concerns good as a concept, it is also the case that when it comes to conceptions of the good, philosophers have tended to take different positions depending on how they have tended to use good. The advantage with attributive uses of good is that they hold the promise of naturalizing the good, thereby rendering its place in the world less mysterious, and this has led some, such as Philippa Foot (2001), to develop ethical theories in terms of the attributive good. The problem with such an approach is just that when it comes to the particular natural kind that is of most interest for ethical matters, namely that of human beings, it is very unclear whether there is any ideal way of leading a human life that one can simply distill from an understanding of what it is to be a human being similar to how one can know what is to be a good knife through an understanding of what is to be a knife.

CONCEPTIONS OF THE PRUDENTIALLY GOOD

The question of what is involved in leading a good life is one of the oldest in philosophy. And although there are many different theories about what is good for people, there are two main traditions of thought on the issue running through the history of ethics. The first is hedonism, the theory that what ultimately makes up a good life is pleasure or enjoyment. The second is perfectionism, that what ultimately is good for a person is to flourish as a human being. Of course, these two traditions do not exhaust the possibilities, but they stand apart in sharing an important strength, namely that they do not simply list a handful of things that are supposed to be good; they also provide an underlying idea about the point of it all. In the case of hedonism there is the appealing thought that if something is to be good for people, then it must feel good to them. In the case of perfectionism, there is another appealing thought, namely that as a human being, one has certain potentials and that it is a waste of one's life if one does not realize them. These two lines of thought do however pull in opposite directions, the first in a subjective one and the second in an objective.

Although there were philosophical schools during antiquity that tended toward hedonism (the Epicureans and, more clearly, the Cyrenaics), perfectionism dominated ethics during that time. As already noted, there was a tendency during antiquity to use good attributively: To

lead a good life is to excel at being a human being. Thus, to understand what the good life consists in, one must understand what lies at the heart of a distinctly human existence. The standard answer was that the exercise of reason was what lifted humans above the level of animals. For instance, Aristotle argued that rational thought was the function, or characteristic activity, of human beings and that, therefore, it must define wherein our good lies.

There are certainly problems with perfectionism having to do with its objectivist slant (what if one really does not want to achieve their potential but do other things instead, where does their good lie?), but the most significant worry is a metaphysical one. Does not perfectionism require a teleological conception of human nature and has not such a conception been put to rest by science? There are modern attempts to address this issue, one of the more interesting being Alasdair MacIntyre's (1984) attempt to understand humans as social beings and partly derive the contents of a good life from the social practices in which humans are embedded; but this remains a difficult issue for perfectionists.

Hedonism, by contrast, had its most prominent exponents in British moral philosophy of the 1700s and the 1800s. Its most emphatic proponent was Jeremy Bentham, although the version that has generated most discussion is that of John Stuart Mill (1998), who argued that there is a qualitative dimension to be considered when judging the value of pleasures: The pleasures of poetry are superior to the pleasures of pushpins. This can thus be seen as an attempt at a hybrid theory, introducing perfectionist elements into hedonism. Hedonism of all forms is however plagued by problems that are rooted in its subjectivist nature.

Say that two people lead lives containing equal amounts of pleasure or enjoyment, but in one of them the pleasures are all based in the subject's delusions; in that case it seems fair to say that the other life is at least somewhat better. So even if pleasure is a very important good, it can reasonably be doubted whether it really is, as hedonists would have it, the only one. Although there are later hedonists, a good example being Fred Feldman (2004), trying to address these problems, the second half of the twentieth century has seen those philosophers who are drawn to subjectivist conceptions of the good life largely abandoning hedonism for theories that emphasize the fulfillment of our preferences instead: One leads a good life when what one wants is realized. This has the advantage of involving the way things really are, but it is accomplished at the cost of moving away from the emphasis on how things feel that provided such an attractive rationale for subjectivist theories to begin with.

CONCEPTIONS OF THE MORALLY GOOD

In antiquity the standard conception of being a good person was to have the four cardinal virtues: wisdom, justice, courage, and temperance. Philosophers tended to conform roughly to this view. Aristotle was a notable exception and presented an extensive list of virtues. His theory is also original in that while many place virtue and vice as opposite poles on a moral spectrum, Aristotle conceived of virtue as a mean lying between two vices, one of deficiency and one of excess. It should, however, be pointed out that even if the ancients tended to list a number of traits as constituents of moral goodness, a common idea among them was that of the unity of the virtues, that one either has all the virtues or none. This is a controversial idea because many would say that it is quite possible, for instance, to be courageous without being just. Adherents of the unity thesis would respond that persons cannot really be deemed courageous if they do not have a fair appreciation of what is at stake—and for that they need to have all the virtues. Terrorists might be prepared to sacrifice their lives, but that alone does not make them courageous.

Although there was a renewed interest in the virtues, particularly in Aristotelian virtue theory, toward the end of the twentieth century, modern philosophers have focused more on the question of which actions are right than on what constitutes a good person. Kantianism is probably the modern moral theory with the best articulated vision of moral goodness. For Kant it is the will rather than character that is the potential bearer of moral goodness. And whereas for Aristotle one cannot be a good person unless acting well comes naturally to them—indeed, he even thinks that it is a mark of good persons that they take pleasure in acting well—for Kant virtue is essentially about self-control, about having moral determination.

This does not preclude that one takes pleasure in acting well, but the true test of virtue occurs when things do not come naturally: Does one then still put morality above their own inclinations? Kant also sharply distinguished between the moral and the prudential. This is very common among modern philosophers, whereas ancient philosophers often saw the virtues as constitutive parts of the good life. This tendency has led some philosophers, such as Bernard Williams (1985), to question whether the idea of a sui generis category of the

moral is not a modern artifice, thus partially echoing Friedrich Nietzsche's complaints about how the evolution of morality has involved a turn from a positive striving after excellence to a negative, prohibitive ethic of self-diminishment.

ETHICAL THEORIES AND THE GOOD

A common way of distinguishing ethical theories (i.e., theories of what one ought to do) with respect to the role played by the good in them is into teleological and deontological. This distinction was introduced by J. H. Muirhead (1932). In teleological theories actions are right because of the way that they contribute to the good, either, as is the case in utilitarianism, because it contributes to the common good or, as is the case in self-realization theories like those of ancient virtue theorists or Hegelians such as F. H. Bradley (1927), because it contributes, or is at least ultimately connected to, the agent's flourishing. Deontologists reject this direct link between the right and the good. In its simplest form, exemplified by W. D. Ross (1930), a deontological theory simply consists in a set of moral rules that are to be obeyed. Indeed, one main worry, voiced by Muirhead as well as others such as J. J. C. Smart (1973), about deontological theories is precisely that they inculcate almost a form of blind rule worship: One just obeys certain rules because they are the rules one should obey.

As a general charge against deontological theories, this is unfair. There can still be an underlying rationale in terms of the good; it is just that it does not take such a direct form. The most sophisticated form of deontology is probably Kantian ethics and while it is true that moral rules are given another kind of justification in Kant (in terms of what reason demands of one as an agent), he does still provide a picture of the place of morality in the life of a human agent, namely that it is a condition of the value of one's well-being: If one is not moral, one cannot reasonably view one's self as worthy of happiness. And in liberal rights theory it is a common idea that the rationale for the basic principles of right is provided at least in part by the fact that there is such widespread disagreement about the good: In the face of such disagreement certain principles of right make sense because they enable people to live together and pursue their own private conceptions of the good. It might even be seen as a weakness of teleological theories that they require that a conception of the good be spelled out in order to make it possible to tell right actions from wrong ones. A variety of teleologists, especially in the utilitarian tradition, have of course proposed a number of such conceptions, but none of them have won wide assent and there seems to be little reason to see such assent as forthcoming. In light of that, some might see deontology as a more viable approach.

See also Deontological Ethics; Ethical Naturalism; Intrinsic Value; Liberalism; Metaethics; Objectivity in Ethics; Teleological Ethics; Utilitarianism; Value and Valuation.

Bibliography

Aristotle. *Nicomachean Ethics*. Translated by T. H. Irwin. Indianapolis, IN: Hackett, 1999.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*, edited by J. H. Burns. Oxford: Oxford University Press, 1996.

Bradley, F. H. *Ethical Studies*. 2nd ed. Oxford: Oxford University Press, 1927.

Brentano, Franz. *The Origin of Our Knowledge of Right and Wrong*. Translated by Roderick Chisholm and Elizabeth Schneewind. London: Routledge, 1969.

Brink, David. *Moral Realism and The Foundation of Ethics*. Cambridge, MA: Cambridge University Press, 1989.

Ewing, A. C. *The Definition of Good*. New York: Macmillian, 1947.

Feldman, Fred. *Pleasure and The Good Life*. New York: Oxford University Press, 2004.

Foot, Philippa. *Natural Goodness*. Oxford: Oxford University Press, 2001.

Hare, R. M. *The Language of Morals*. Oxford: Oxford University Press, 1952.

Geach, P. T. "Good and Evil." *Analysis* 17 (1956): 33–42.

Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard University Press, 2003.

Kant, Immanuel. *The Cambridge Edition of The Works of Immanuel Kant: Practical Philosophy*. Translated by Marry J. Gregor. Cambridge, MA: Cambridge University Press, 1996.

MacIntyre, Alasdair. *After Virtue*. 2nd ed. Notre Dame, IN: University of Notre Dame Press, 1984.

Mill, John Stuart. *Utilitarianism*, edited by Roger Crisp. Oxford: Oxford University Press, 1998.

Moore, G. E. *Principia Ethica*. Cambridge, MA: Cambridge University Press, 1903.

Muirhead, J. H. *Rule and End in Morals*. Oxford: Oxford University Press, 1932.

Nietzsche, Friedrich. *On The Genealogy of Morality*. Translated by Carol Diethe. Cambridge, MA: Cambridge University Press, 1994.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Ross, W. D. *The Right and The Good*. Oxford: Oxford University Press, 1930.

Scanlon, Thomas. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Smart, J. J. C., and Bernard Williams. *Utilitarianism: For and Against*. Cambridge, MA: Cambridge University Press, 1973.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Johan Brännmark (2005)

GOODMAN, NELSON
(1906–1998)

Nelson Goodman, the distinguished American philosopher of science and language, was born in Massachusetts in 1906. He received a bachelor of science degree from Harvard in 1928 and took his Ph.D. in philosophy there in 1941. After an instructorship at Tufts College (1945–1946), he was appointed associate professor at the University of Pennsylvania (1946–1951) and then professor (1951–1964). From 1964 to 1967 Goodman was the Harry Austryn Wolfson professor of philosophy at Brandeis University. In 1967 he became a professor of philosophy at Harvard. He died in 1998.

Goodman's delineations of certain strategic problems in epistemology, philosophy of science, and constructional methods, as well as the results of his own inquiries, are fundamental in the areas in which he worked. Specifically, these include theories of inductive logic or confirmation, problems concerning the nature of causal or lawlike regularity, theories of the structural or logical simplicity of theories, and constructions of linguistic systems within which philosophical problems may be solved, as well as theories of the adequacy or accuracy of such systems. Because of his achievements any further significant contributions to these areas may be expected to rest, in some measure, upon his work.

In this brief compass no attempt will be made either to give a comprehensive account of Goodman's ramified views or to rehearse in full detail any one of his major achievements. Instead, we will give an account of a few aspects of his major contributions in just sufficient detail to make their general import intelligible and to show something of their interconnections.

Our order of presentation of topics is quite independent of their chronology in Goodman's philosophical development. We begin with those of his important studies with which there appears to be widest familiarity.

INDUCTIVE THEORY

One of Goodman's characterizations of the task of inductive theory is that it consists in "formulating rules that define the difference between valid and invalid inductive inferences." On this usage a set of rules for discriminating valid acceptances or nonacceptances of hypotheses from those which are invalid constitutes an inductive theory, or, alternatively, a theory of confirmation or a theory of projection.

Goodman's contribution to the provision of such inductive canons has been threefold. First, he provided an analysis of the character of philosophical problems about induction. Second, he furnished a critique of the problems still to be solved and of the versions of confirmation theory which have been at all fully elaborated (notably those of Rudolf Carnap and Carl Gustav Hempel; see *Fact, Fiction and Forecast,* especially pp. 24–34, 48–51, and 68–86, and also the published exchanges between Carnap and Goodman to which reference is made on p. 86). Third, he made advances, explicitly in the form of a discussion of a *theory of projection,* toward the solution of some of the problems thus delineated. Where *induction* is construed narrowly as inference about future cases on the basis of examined cases, *projection* is, by contrast, inference about any unexamined cases on the basis of examined ones. We will consider each of these three aspects of his contribution in turn.

THE "PROBLEM OF INDUCTION." Goodman argues that the so-called problem of induction, when it is construed as the problem of *justifying* induction, is one that may be "dissolved" as soon as we see what is at issue. Moreover, this "dissolution" highlights all the more clearly the bona fide problem that he calls the new riddle of induction. As he sees it the problem is not to justify induction but to be able to distinguish valid from invalid inductions. On Goodman's view the dissolution of the *old* problem of induction, that is, of the problem of justifying induction, is accomplished when we come to understand that a genetic or descriptive account of our inductive behavior, such as the one that David Hume *almost* brought off, furnishes the basis of such a justification. That this is a cogent view, he points out, can be seen when we raise the question of justifying *deduction.* How do we justify a deductive inference? By showing that it conforms to specific logical rules of deduction. By the same token, an inductive inference can be justified by showing that it conforms to a specific rule of induction.

One may immediately ask, however, what justification we have for adopting a set of rules of induction as valid. Of course, the same question might be asked concerning a set of deductive rules. The answer may be indicated by furnishing a parable.

Consider the situation of an imaginary philosopher to whom we may give the name "Aristotle." Aristotle has a keen interest in the area of deductive inference. In this area, he finds that although there is *already* an established practice among humans of making deductive inferences and although there is *already* a practice of discriminating,

among ostensible inferences of this type, those that are correct from those that are not, nevertheless no one has yet made explicit or systematically codified the implicit rules upon which such discriminations appear to be based.

Our imaginary philosopher decides to undertake this task and eventually comes forward with such a codification. Using his codification people are enabled to make explicit their reasons for discriminating valid from invalid deductions by referring to the explicit rules that Aristotle has placed conveniently at hand. Of course no one would have paid any attention at all to these rules if they did not, with fair accuracy, reflect established practice—this is indeed what constitutes their validity as a set of rules. In the course of many years, however, other philosophers come forward to point out anomalies in Aristotle's set of rules. They point out that in certain cases some of his rules yield unacceptable inferences, and these philosophers suggest amendments which will remove the anomalies. When the amendments are incorporated they, in turn, have the effect of modifying practice. As Goodman puts it:

> [Deductive] inferences are justified by their conformity to valid general rules, and … general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences." (*Fact, Fiction and Forecast,* p. 67)

If we return our attention to induction we see that an analogous situation obtains. Particular inductive inferences are justified by reference to rules of induction, and rules of induction are justified by reference to particular practices of inducing. Hume was on the right track in giving a descriptive account of inductive practice and in explicating rules of causal inference that he held to be in conformity with this practice. Those who have criticized him for this have been wrong.

We are thus quits with the old problem of induction, but the new, very formidable "riddle of induction" still remains. For although Hume was right in turning to description of actual practice, his description was insufficiently precise. He pointed out that observed regularities give rise to habits of expectation and that predictions based on such regularities are "normal or valid." But the defect in Hume's account, Goodman shows, lies in his failure to note "that some regularities do and some do not establish such habits; that predictions based on some regularities are valid while predictions based on other regularities are not. … To say that valid predictions are those based on past regularities is thus quite pointless" (ibid., pp. 81–82). Accordingly, the new riddle of induction consists in finding a set of rules of inductive logic that will do for us what Hume failed to do. Thus, the problem is not to justify induction but adequately to codify it. An adequate codification would presumably stand to inductive practice very much as the codification of deduction, accomplished by our mythical Aristotle, stood to deductive practice as described in our parable above. In particular, it would presumably consist of a set of rules the appeal to which would serve to validate specific acceptances or rejections of scientific hypotheses or theories.

CRITIQUE OF CONFIRMATION THEORY. In "The Problem of Counterfactual Conditionals" (reprinted without major change as Chapter 1 of *Fact, Fiction and Forecast*) Goodman was able to show that a solution to the problem of achieving an adequate interpretation of counterfactuals is intimately connected with many of the other crucial questions of the philosophy of science and that such a solution could be achieved only if various critical questions about the nature of scientific laws and of confirmation theory could be answered.

He shows, in particular, that the problem of furnishing adequate criteria for distinguishing true from false counterfactual conditionals has as a constituent the problem of adequately defining "scientific law," that this requires us to distinguish those hypotheses which are *confirmed* by their instances from those which are not, and that this, in turn, requires the fashioning of an adequate theory of confirmation. It is together the burden of the last part of "The Problem of Counterfactual Conditionals," of two brief articles on confirmation theory and of several passages in Chapter 3 of *Fact, Fiction and Forecast* that extant confirmation theories are defective, for they provide no means (except such as vitiate the theories through question-begging stipulations about what primitive predicates may be comprised in confirmable hypotheses) to distinguish the hypotheses to which such theories may be applied. Goodman, for example, points out that extant provisions of criteria for what constitutes a confirming instance in such defective theories either have the consequence that "any statement will confirm any statement" (ibid., p. 81) or make question-begging assumptions, mentioned before, about the recognizability

of "purely qualitative predicates" which are held to be the only permissible ones that may occur in (thus distinguishable) confirmable hypotheses. He shows, in short, that a desideratum of theories of confirmation is a definition of "confirmable hypothesis." In the final chapter of *Fact, Fiction and Forecast* he attempts to fill this need through advances on the problem of defining "projectible" as a predicate of hypotheses.

THEORY OF PROJECTION. In earlier discussions Goodman had shown that certain dispositional terms (other than *projectible*) may be adequately defined by *projecting* them over the extensions of (that is, by defining them in terms of) certain carefully specified nondispositional or *manifest* predicates. Such earlier successes provide important paradigms. If on their model the meaning of the term *projectible* can be clarified, it will be feasible to decide to which hypotheses the term applies, and a crucial desideratum of heretofore defective theories of confirmation will have been taken care of.

Inasmuch as the term *projectible* is itself a dispositional predicate we may expect that among the manifest predicates that will occur in any candidate definiens will be the *corresponding* manifest predicate: "projected." However, defining *projectible* in terms of "projected" offers some very special difficulties which do not arise in the case of many dispositional predicates. The predicate "projectible" is like "desirable." It is not the case that every hypothesis that has been actually projected ought to have been or ought to be projected. (A hypothesis is characterized as having been *actually projected* if "it is adopted after some of its instances have been examined and determined to be true, and before the rest have been examined"; ibid., p. 90.)

Goodman, perhaps unlike J. S. Mill in confronting "desirable" is explicitly aware of the trap, and although his task is thereby enormously complicated, he avoids falling into it. He proposes, eventually, an explication of "projectibility" that provides criteria for discriminating projectible hypotheses based on past projections and certain other characteristics of our actual linguistic habits. In particular, attention to actual projections of hypotheses enables Goodman to explicate a relevant sense of *projected predicate* (a predicate occurring in an actually projected hypothesis). This, in turn, leads to his explication of a concept that becomes pivotal to his theory of projection: the concept of "entrenchment"—more specifically, the concept "is a much better entrenched predicate than."

One predicate, *P*, is said by Goodman to be much better entrenched than another predicate, *Q*, if *P* and all predicates coextensive with it have actually been projected much more often than *Q* and all predicates coextensive with it. Thus, take the predicate "grue" (which applies to any blue thing not examined before some time, *t*, and also to any thing examined before time *t* and found to be green). This "highly artificial" predicate, occurring in the hypothesis "The next emerald to be examined (after time *t*) will be grue" allows that hypothesis to be *equally highly evidenced* with the more usual "The next emerald to be examined (after time *t*) will be green." But hypotheses employing "grue" (or any term applicable to exactly the things "grue" is applicable to) have, nevertheless, been much less frequently projected (for example, used in making predictions) than have hypotheses using "green" (or any term applicable to exactly the things "green" is applicable to). This is part of the basis upon which "green" is judged a much better-entrenched predicate than is "grue"; and Goodman's theory attempts to show how, although they are equally well evidenced, hypotheses containing much better-entrenched predicates are to be preferred to ones that contain much less well-entrenched predicates. Goodman points out that when we speak of the entrenchment of predicates we are really speaking of the entrenchment of habits of classification. This is to say that talk of the entrenchment of predicates is, in effect, talk of the entrenchment of their extensions. And, a little later on, still referring to his elucidation of entrenchment, he says:

> Like Hume, we are appealing here to past recurrences, but to recurrences in the explicit use of terms as well as to recurrent features of what is observed. Somewhat like Kant, we are saying that inductive validity depends not only upon what is presented but also upon how it is organized; but the organization we point to is effected by the use of language and is not attributed to anything inevitable or immutable in the nature of human cognition. To speak very loosely, I might say that in answer to the question what distinguishes those recurrent features of experience that underlie valid projections from those that do not, I am suggesting that the former are those features for which we have adopted predicates that we have habitually projected. (Ibid., pp. 96–97)

The import of these considerations is that what constitutes a valid projection, and consequently what comes to constitute a projectible hypothesis, is a result of how we have, as a matter of fact, come to classify.

If Goodman's attempt to define "projectible" is successful, we have at hand the means of solving the problem of distinguishing confirmable from nonconfirmable hypotheses and thereby of surmounting a major obstacle in the way of providing a logic of induction.

These results of Goodman's—both the critique of extant theories and the positive proposals put forward in 1955 (ibid.)—are clearly still being digested by people in the field, if one may judge by the discussions of them that (ten years later) appeared in print with increasing frequency.

THEORY OF STRUCTURAL SIMPLICITY

An early version of Goodman's calculus of simplicity (later extensively modified) occurs in Part I, "On the Theory of Systems," of his first book, *The Structure of Appearance.* There the calculus is exclusively connected with considerations somewhat more general than those involved in, for example, assessing the simplicity of scientific theories. In *The Structure of Appearance* interest in simplicity is interest in the simplicity of the primitive predicate basis of *any* constructional system; that is, any constructed linguistic system or axiomatic system which makes explicit what are the primitive (that is, the undefined) terms of the system. The main general problem that Goodman addresses is that of delineating criteria of adequacy for constructional systems generally, rather than for scientific theories in particular. For the constructor of such systems this problem is often posed—in part, at least—as the problem of choice among alternative primitive predicate bases. In choosing a primitive basis such considerations as antecedent clarity and "defining power" are obviously to be taken into account, but Goodman shows that the simplicity—the structural or logical simplicity—of such bases, is also an at least equally important consideration.

In his later writings on the subject (particularly in "The Test of Simplicity" and *Fact, Fiction and Forecast*) Goodman also made clear the relationship of measures of simplicity to the philosophy of science. He maintains that simplicity is a primary consideration guiding choices among scientific theories or systems of hypotheses. It is a mistake to believe that simplicity becomes a factor only after we have first sought a true system and then turn to matters of elegance. He maintains that, on the contrary, our concern with simplicity is an inevitable concomitant of our concern with system. For, he points out, we achieve systematization only to the extent that the basic vocabulary and principles we employ in dealing with some subject matter come to be simplified. The important thing to note is that "when simplicity of basis vanishes to zero—that is, when no term or principle is derived from any of the others—system also vanishes to zero. Systematization is the same thing as simplification of basis" ("The Test of Simplicity," p. 1064).

Goodman finds the key to the problem of measuring the structural simplicity of predicate bases in a "meagre and negative" but highly plausible principle: "If every basis like a given one can always be replaced by some basis like a second, then the first is not more complex than the second" (ibid., p. 1066). The relation "always replaceable by" between predicate bases holds in cases where the replacement is a matter of a purely routine procedure that can always be applied (presumably, for example, in case there is available some *decision procedure* for determining replaceability). Employing this key principle and some results in the theory of relations. Goodman provides a means of effecting the requisite measures. The *calculus of simplicity* is applicable only to theories that have been at least sufficiently formalized to enable discrimination of their primitive predicates. Its applicability (for example, as a factor in assessing the acceptability of some scientific theory) is thus severely limited for the present by the paucity of scientific theories that have reached this stage of formalization. On the other hand, this situation would be importantly alleviated if some means could be found either to bring more such theories to the requisite stage of formalization or to modify the calculus in such a way that useful applications of it may be made even to less fully formalized systems.

For the time being, applications of the simplicity measures may, however, be made to constructional systems devised for purposes of philosophical explication (for example, see Goodman's own system in *The Structure of Appearance*).

CONSTRUCTIONALISM

Whatever its importance for both philosophy of science and constructional methods, furnishing a way of measuring the simplicity of bases of any constructed systems by no means represents Goodman's only contribution to constructional methods. The first three chapters of *The Structure of Appearance* (for example) provide also a discussion of the problem of assessing the adequacy and the accuracy of definitional systems. Here an especially significant discussion (for example, in Ch. 1) provides both an illuminating critique of criteria which in the past have been adduced for assessing such systems and a newly developed criterion, *extensional isomorphism,* for assessing the *accuracy* of such systems. The development of this

criterion throws new light on the entire program of philosophical or logical analysis.

Although full elucidation of the criterion is beyond the scope of the present entry, some general inkling of its import may perhaps be conveyed by pointing out some of its differences from some of the criteria which have been previously offered for the adequacy of philosophical analyses. It has long been recognized that full synonymy of analysandum (the concept or term being subjected to philosophical analysis) and analysans (the concept or term constituting the product of the analysis) is too strong a requirement. Accordingly, weaker criteria (for example, *intensional identity* or *extensional identity* of analysandum and analysans) have been proposed. In *The Structure of Appearance* Goodman argues that even the weakest of these—extensional identity—is too strong a requirement to place on tasks of analysis, for none can totally fulfill such a condition. He proposes instead a criterion that does not "square" an analysandum with its analysans in any one-to-one fashion but rather tests the whole *system* of concepts to which the analysandum belongs against the whole newly constructed system to which the analysans belongs. The meeting of specified and relatively weak extensional correspondences between two such systems is sufficient—and indeed is the most that can cogently be required—to warrant the accuracy of the analysis.

The discussions of new constructional methods in the first chapter of *The Structure of Appearance* and the presentation of a version of the calculus of individuals which had been developed by H. S. Leonard and Goodman (in "The Calculus of Individuals and Its Uses") are well supplemented by the specific application of these and other devices to a detailed critique of Carnap's *Der logische Aufbau der Welt* (in Ch. 5). An important application is also provided by the construction (in Chs. 6–11) of his own systematic explication of phenomenal concepts or predicates.

PHENOMENALISM AND NOMINALISM

Goodman's actual work upon, and his defenses of work upon, phenomenalistic systems lead many observers to conclude that he subscribes to phenomenalism as a philosophical position. The fact is, however, that he wrote in full and explicit detail about the relative unimportance and the opacity of questions about the epistemological priority of the phenomenal (and "rival," for example, physicalistic) systems, and there seems to be no good reason to doubt the sincerity of his disavowals of that kind of philosophical commitment. (See *The Structure of Appearance,* Ch. 4 and passim, and "The Revision of Philosophy.") All of this is notwithstanding the fact that he made contributions to the solution of many very complex problems that are involved in the construction of a phenomenalistic system.

If phenomenalism represented, for him, no particular philosophical commitment, *nominalism,* on the other hand, surely did. His major writings on this topic (in his and W. V. Quine's "Steps Toward a Constructive Nominalism" and in his *The Structure of Appearance; Fact, Fiction and Forecast*; and "A World of Individuals") obviously constitute a fundamental philosophical conviction. Although Goodman's and Quine's nominalism coincide importantly (for example, in their mutual rejection of classes, see "Steps toward a Constructive Nominalism"—but note, however, that in later writings Quine appears no longer to embrace such views) it should nevertheless be observed that their nominalistic positions are quite disparate. Thus, Quine apparently rejects, so to speak, classes on account of their being abstract entities; whereas Goodman rejects, so to speak, classes not on account of their being abstract entities (his system in *Structure,* indeed, refers to abstract entities categorematically) but rather on account of their being *nonindividuals.* It is the notion of a *nonindividual* that Goodman finds unintelligible, and he is conscientious in avoiding any philosophical or logical method which presupposes or extorts the claim that there exist any nonindividuals. The consequent austerity in bases chosen and logical tools available to him have had, in fact, fruitful results in eliciting complex, ingenious, and far-reaching techniques or methods of constructional analysis.

We have indicated that there are differences between what might be called G-nominalism (Goodman's position)—the view, on the one hand, that there are no nonindividuals—and the position that might be called Q-nominalism—the view that there are no abstract entities, on the other. While it would, again, be beyond the scope of this entry to give a detailed account of G-nominalism, it may yet be illuminating to remind the reader that Goodman himself characterized his position as a sort of "super-extensionalism." The usual or classical extensionalist position prohibited some otherwise indiscriminate multiplication of entities by imposing a principle to the effect that two entities (say, two classes) that have, so to speak, the same proximate constituents are identical. G-nominalism goes further; it imposes the condition that any two things which have the same systematically *ultimate* constituents are identical.

Thus, consider the systematic atoms (things not having anything else in the system as possible constituents) *a, b, c,* and *d.* Suppose in a (classically) extensional system *A* we discriminated the classes of pairs {*a,c*} and {*b,d*}, and suppose in system *B* we discriminated the classes {*a,b*} and {*c,d*}. For classical extensionalism, systems *A* and *B* would *not* be identical; that is, the proximate constituents—the two classes of pairs—are different, and hence the world's population on this account is increased by two more classes. The G-nominalist, however, has a stronger condition for diversity. For him there are not, say, the eight different entities consisting of the four atoms and the four classes of pairs of them. Rather, there are only four entities—the ultimate atoms of the system themselves. The cogency of this view is argued with great vigor and clarity in "A World of Individuals."

WORK IN PROGRESS

Goodman's interest appears to be an analysis of representationalism in a very broad sense of this concept taken presystematically. Thus, the focus of his attention is not only upon representation as a phenomenon involving, for example, paintings in aesthetics but also upon the representational aspects or functions of maps, graphs, musical scores, and choreographic notations, and, in addition, theories and other descriptions. His deep and abiding interest in this topic is evidenced too as a recurrent thread in many of his works, from very early ones on. The articles in which this concern is most obviously expressed are "On Likeness of Meaning," "Sense and Certainty," "The Way the World Is," and "About." The concern is also dominantly present in his John Locke lectures (given at Oxford in 1962), published as *The Languages of Art.*

See also Aesthetics, History of; Aesthetics, Problems of; Aristotle; Carnap, Rudolf; Counterfactuals; Hempel, Carl Gustav; Hume, David; Induction; Kant, Immanuel; Mill, John Stuart; Nominalism, Modern; Phenomenalism; Philosophy of Science, Problems of; Projectivism; Quine, Willard Van Orman.

Bibliography

WORKS BY GOODMAN

"The Calculus of Individuals and Its Uses." *Journal of Symbolic Logic* 5 (1940): 45–55. In collaboration with H. S. Leonard.

"A Query on Confirmation." *Journal of Philosophy* 43 (1946): 383–385.

"On Infirmities of Confirmation-Theory." *Philosophy and Phenomenological Research* 8 (1947): 149–151.

"Steps toward a Constructive Nominalism." *Journal of Symbolic Logic* 12 (1947): 105–122. In collaboration with W. V. Quine.

"On Likeness of Meaning." *Analysis* 10 (1949): 1–7.

"An Improvement in the Theory of Simplicity." *Journal of Symbolic Logic* 14 (1950): 228–229.

The Structure of Appearance. Cambridge, MA: Harvard University Press, 1951.

"Sense and Certainty." *Philosophical Review* 61 (1952): 160–167.

Fact, Fiction and Forecast. London, 1954; Cambridge, MA: Harvard University Press, 1955.

"Axiomatic Measurement of Simplicity." *Journal of Philosophy* 52 (1955): 709–722.

"The Revision of Philosophy." In *American Philosophers at Work,* edited by Sidney Hook, 75–92. New York, 1956.

"A World of Individuals." In *The Problem of Universals,* 13–31. Notre Dame, IN, 1956.

"The Test of Simplicity." *Science* 128 (1958): 1064–1069.

"Recent Developments in the Theory of Simplicity." *Philosophy and Phenomenological Research* 19 (1959): 429–446.

"The Way the World Is." *Review of Metaphysics* 14 (1960): 160–167.

"Positionality and Pictures." *Philosophical Review* 69 (1960): 523–525.

"About." *Mind* 70 (1961): 1–24.

"Safety, Strength, Simplicity." *Philosophy of Science* 28 (1961): 150–151.

The Languages of Art. Indianapolis: Bobbs Merrill, 1968.

Richard S. Rudner (1967)

GOODMAN, NELSON [ADDENDUM]

Nelson Goodman (1906–1998) was one of the foremost philosophers of the twentieth century. His works reshaped epistemology, metaphysics, and the philosophy of art. *The Structure of Appearance* (1977), which grew out of his PhD dissertation, shows how to construct interpreted formal systems that solve or dissolve longstanding epistemological and metaphysical problems. *Fact, Fiction, and Forecast* (1983) poses the new riddle of induction and proposes a solution to it, arguing that to block the inference to "All emeralds are grue," one must consider the ways terms have been inductively used in the past. *Languages of Art* (1976) reconceives aesthetics, construing it as a branch of epistemology. It is tempting to say that Goodman worked in diverse branches of philosophy. A more accurate claim would be that he focused on issues that cut across philosophy, showing that the branches are not so diverse as they sometimes seem.

Goodman attended Harvard University both as an undergraduate and as a graduate student. During graduate school, he supported himself by running an art gallery in Boston. He spent most of his academic life as professor

of philosophy at the University of Pennsylvania. For the final decade of his teaching career he was professor of philosophy at Harvard, where he founded Project Zero, an ongoing research program in arts education, and the Harvard Summer Dance Program. He was an avid, eclectic collector of art.

MEANING

In the late 1940s, Goodman, W. V. Quine, and Morton White wrote a series of papers repudiating the analytic/synthetic distinction. Goodman's target is synonymy. He contends that synonymy requires that terms agree in primary extension and all parallel secondary extensions, where the secondary extension of a term is the extension of a compound containing that term. Although "unicorn" and "centaur" have the same (null) primary extension, because compounds such as "unicorn picture" and "centaur picture" differ in extension—because, that is, not all unicorn pictures are centaur pictures—"unicorn" and "centaur" differ in meaning. This fits human intuitions. But even seemingly synonymous terms differ in meaning according to Goodman's criterion. Although "spine" and "backbone" seem synonymous, one can readily contrive a spine description that is not a backbone description—for example, "spine that is not a backbone." In general, "*p* that is not a *q*" is a *p*-description but not a *q*-description. Such an all-purpose device for generating differences in meaning might seem illegitimate. Even if a person were to exclude its deliverances, the pictures and descriptions that belong to a secondary extension of one but not both of a pair of coextensive terms are easily found. The vast majority of apparently synonymous terms fail to satisfy Goodman's criterion. Although Goodman does not argue for his criterion, its justification is evident. Synonymous terms should be intersubstitutable in fiction and in statements of fact. Because nothing should count as a representation of the referent of the one that is not a representation of the referent of the other, divergence in the classifications of the descriptions or pictures marks a divergence in meaning.

Secondary extensions do more than discredit synonymy; they provide resources for recognizing degrees and kinds of likeness of meaning. To do this, one must limit focus. If, within a restricted range, all parallel compounds of a pair of coextensive terms are coextensive, the meanings of the coextensive terms agree within that range. The terms then may be alike enough in meaning to be intersubstitutable within that range, even if their meanings diverge elsewhere. If in medical discourse all and only instances of "spine representation" are instances of "backbone representation," then "spine" and "backbone" may be sufficiently similar in meaning to be intersubstitutable in purely medical contexts. If most parallel compounds are coextensive, or most important parallel compounds are coextensive, terms may be sufficiently similar in meaning to justify substituting one for the other. In place of a rigid, context-indifferent criterion of synonymy, Goodman provides a flexible, context-sensitive criterion of likeness of meaning (Goodman 1972, pp. 221–238).

The analytic/synthetic distinction is not unique. Other familiar dualisms—essence/accident, scheme/content, necessity/contingency, and the like—are vulnerable to similar objections. All must be rejected, Goodman, Quine, and White believe. Unlike Quine, Goodman devotes little subsequent effort to arguing against the dualisms. He simply jettisons them and does philosophy without them. He considers the demise of the dualisms not to deprive philosophy of resources, but to liberate it from unwarranted restrictions. Perhaps surprisingly, he finds that the rejection of the dualisms fosters progress in aesthetics.

ART

Goodman's trailblazing *Languages of Art* reorients aesthetics. Active engagement, rather than passive contemplation, marks the aesthetic attitude. Goodman believes that the arts function cognitively. He thus construes aesthetics as a branch of epistemology whose task is to explain how and what the arts contribute to cognition. The plausibility of such a position obviously depends both on the nature of art and the nature of cognition. If the cognitive function of art is simply to transfer information to passive receivers, resemblance might be the mechanism. Then works resemble their subjects and convey information about how their subjects appear.

There are, however, seemingly insuperable objections to such a position. At best it works only for representational art. Music, abstract painting, and architecture would not be accommodated under such an account. Nor can it feasibly accommodate anything except realistic works. Neither cartoons nor cubist portraits much resemble their subjects. Indeed, Goodman argues that the account does not even accommodate realistic works. The cognitive contribution of a painting could be to convey the way its subject looks only if there were such a thing to convey. Any thing looks many ways, and looks different ways to different people. A subject then has no canonical look that its painting, to be accurate, ought to convey. Moreover, some realistic pictures have fictive subjects—

unicorns, griffins, and the like. Clearly they do not resemble their subjects, because there are no unicorns or griffins to resemble. Nonetheless, one can readily interpret such pictures, recognize what they are of, and gain insights from them.

The problem, Goodman believes, is that the proposal rests on misunderstandings both of art and of cognition. Contrary to what classical empiricists thought, the mind is not passive in the reception of sensations. It actively searches, seeks, selects, and finds. Nor should works of art be construed as mere sensory surfaces. Rather, Goodman maintains, they are symbols with determinate syntactic and semantic properties. Art advances understanding and affords insight in much the way language does. Like linguistic symbols, the symbols that constitute a work of art require interpretation, and the symbol systems to which the works belong need to be learned. In *Languages of Art* Goodman develops a taxonomy of symbol systems used in the arts and elsewhere, detailing their powers and limitations.

Two modes of reference are basic. Denotation links names to bearers, predicates to instances, representations to the things they represent. "George Washington," "the first U.S. president," the figure on the U.S. dollar bill, and the Gilbert Stuart portrait all denote Washington. In exemplification, a symbol points up—hence refers to—properties it serves as a sample of. A fabric swatch exemplifies its pattern; a Mondrian painting, squareness; a blood test, the presence of antibodies. Ubiquitous in art, exemplification is also widespread in science, advertising … indeed anywhere people adduce samples and examples (Goodman 1976).

Reference need not be literal. Metaphorical reference, Goodman maintains, is real reference; metaphorical truth, real truth. "Bulldog" genuinely, albeit metaphorically, denotes Churchill. "Churchill is a bulldog" is genuinely, although not literally, true. Michelangelo's *Moses* genuinely, albeit metaphorically, exemplifies rage. Expression is metaphorical exemplification by a work of art functioning as such. *Moses* thus expresses the rage it metaphorically exemplifies (Goodman 1976, 1984).

Some reference is complex. In allusion, a referential chain composed of denotational and exemplificational links connects a symbol to its referent (Goodman 1984). Two chains figure in variation, one exemplifying features a variation shares with its theme, the other exemplifying features that contrast with the theme (Goodman and Elgin 1988).

Scientific symbols, Goodman maintains, are relatively attenuated, aesthetic symbols relatively replete. A scientific symbol is normally univocal, its full referential import readily apparent. An aesthetic symbol may bear multiple correct interpretations and symbolize along several dimensions simultaneously. Exactly what it symbolizes may never be settled. The same item may qualify as a symbol of either kind, depending on how it functions. So "When is art?" not "What is art?" is the crucial question. When, how, and to what effect does a symbol function aesthetically (Goodman 1978)?

Art advances understanding, not only because interpretation is a cognitive process. Encounters with art afford insights that extend beyond the aesthetic realm; discoveries made, orientations adopted, and patterns discerned in aesthetic contexts transfer and make sense of other aspects of experience; emotion transform from ends to means. The emotional reactions a work evokes are not ends in themselves but means of understanding the work and the light it sheds on other aspects of human experience.

WORLDMAKING

In *Ways of Worldmaking* (1978), Goodman returns to constructionalist themes first explored in *The Structure of Appearance*. Worlds, he contends, are made, not found. Because the elements of any group are alike in some respects and different in others, mere examination will not reveal whether two manifestations are of the same thing, or two things of the same kind. To settle such matters requires criteria of individuation and classification. Category schemes supply them. But category schemes are human constructs. In devising them, people demarcate the individuals and kinds that make up a world. Different demarcations yield divergent, but equally tenable world versions. One might characterize light as a stream of particles; another, as a sequence of waves. Each may be right relative to its own world-version, wrong relative to its rival's. Neither is right or wrong absolutely.

If overlapping world-versions all supervened on a single base, such differences would be ontologically innocuous. But world-versions do not supervene on a single basis. A physicalist version, for example, neither supervenes on nor underlies a phenomenalist version. Nor does any neutral version underlie them both. Because people can and do construct multiple, individually adequate but irreconcilable world-versions, there are, Goodman concludes many worlds, if any (Goodman 1978).

Worldmaking is not always deliberate. In *Ways of Worldmaking*, Goodman analyzes a series of psychological experiments and shows how, with only sparse cues, the visual system constructs the apparent motion it detects (Goodman 1978). Nor is worldmaking always discursive. Nonverbal schemes structure things in ways no description precisely captures. The arts as well as the sciences construct viable world-versions.

Despite Goodman's recognition of multiple ways of worldmaking and multiple worlds made, he does not contend that every version makes a world. Only right versions do. Rightness does not reduce to truth, for some truths are wrong, some falsehoods right, and some symbols right though neither true nor false. Rightness involves fitting and working—fitting with past cognitive practice and working to promote cognitive ends. Consistency, cogency, projectibility, and fairness of sample figure in the rightness of tenable world versions (Goodman 1978, Goodman and Elgin 1988).

Bibliography

WORKS BY NELSON GOODMAN

A Study of Qualities. New York: Garland, 1990. Goodman's PhD dissertation.

Esthetique et Connaissance: Pour changer de Sujet. With Catherine Z. Elgin. Combas: Editions de l'Eclat, 1990.

Fact, Fiction, and Forecast. 4th ed. Cambridge, MA: Harvard University Press, 1983.

Languages of Art. 2nd ed. Indianapolis: Hackett Publishing Co., 1976.

Of Mind and Other Matters. Cambridge, MA: Harvard University Press, 1984.

Problems and Projects. Indianapolis: Hackett Publishing Co., 1972.

Reconceptions in Philosophy and Other Arts and Sciences. With Catherine Z. Elgin. Indianapolis: Hackett Publishing Co., 1988.

The Structure of Appearance. 3rd ed. Dordrecht, Netherlands: Reidel, 1977.

Ways of Worldmaking. Indianapolis, IN: Hackett Publishing Co, 1978.

WORKS ON GOODMAN

Elgin, Catherine Z. ed. *The Philosophy of Nelson Goodman*. 4 vols. New York: Garland, 1997.

Elgin, Catherine Z. *With Reference to Reference*. Indianapolis, IN: Hackett Publishing Co., 1983.

McCormick, Peter J. ed. *Starmaking*. Cambridge, MA: MIT Press, 1996.

Scheffler, Israel. "The Wonderful Worlds of Goodman." In *Inquiries*. Indianapolis, IN: Hackett Pubishing Co., 1986.

Schwartz, Robert. "The Power of Pictures." *The Journal of Philosophy* 82 (1985): 711–720.

Stalker, Douglas, ed. *Grue: The New Riddle of Induction*. La Salle, IL: Open Court, 1994.

Catherine Z. Elgin (1996, 2005)

GORGIAS OF LEONTINI

(c. 485–c. 380 BCE)

Gorgias of Leontini (in Sicily) was a leading Greek rhetorician and Sophist of the fifth century BCE. He came to Athens on an diplomatic mission on behalf of Leontini in 427 B.C.E. and made an enormous personal success, delivering public orations as well as his official speech. He also toured the Greek cities as a celebrated teacher and public speaker, giving orations at the Olympic and Pythian games. Ancient sources associate him with the philosopher Empedocles (who may have been his teacher) and the rhetorician Isocrates (possibly his pupil). In addition to various sayings and fragments, three complete works by Gorgias have survived: *Encomium of Helen*, *Defense of Palamedes*, and *On Not Being or On Nature*. Gorgias is also depicted as a character in Plato's *Gorgias*, though how much evidence we can extract from this for his ideas or character is unclear. Whether Gorgias should be counted as a sophist is debatable. He was first and foremost a rhetorician, a teacher of public speaking, whereas the central subject of sophistic teaching was virtue. But the distinction was somewhat blurred, and Gorgias's ideas clearly belong to the sophistic movement, broadly construed.

Gorgias was not the first professional rhetorician; but his style was novel, and he was later seen as the real founder of the discipline. His language was notoriously elaborate, with a heavy use of antithesis and alliteration (see especially the *Helen* and the fragmentary *Funeral Oration*). At the same time, he specialized in improvisation: he would offer to answer any question posed by his audience or to speak extemporaneously on a suggested topic. He seems to have understood rhetoric as an all-powerful, value-neutral art (*technē*), consisting in a set of verbal techniques for the manipulation of an audience. As Plato reports it, he claimed that the art of persuasion was superior to all others because it enslaves all the rest—not by force, but with their own consent (*Philebus* 58a–b). Plato's *Gorgias* presents Gorgias as a genial, self-satisfied old gentleman, basically unreflective and blind to the morally problematic nature of such a craft. Plato's *Meno* also provides some intriguing scraps of information about Gorgias (who was Meno's teacher): he (a) disclaimed the teaching of virtue (95c), (b) held scientific

views, including a theory of how vision takes place (76c–d), and (c) held that the virtue of each kind of person (man, woman, child, slave, and so on) is different (71c–2b), suggesting that he may have advocated definition by an enumeration of species, in opposition to the Socratic search for a common denominator.

Of Gorgias's surviving works, the *Palamedes* is notable as an example of the rhetorical genre of *epideixis*: a set-piece speech presented as an advertisement, and perhaps used as a template for students to study. It argues on the basis of probability (*to eikos*), a characteristic rhetorical form of argument. Gorgias's other two surviving speeches might also have served as *epideixeis*: the sophists were traditionally described (for example, in Aristophanes' *Clouds*) as "making the weaker argument the stronger," and it is hard to think of a better way to display that skill than by proving that Helen of Troy was blameless (the *Helen*) or that nothing exists (*On Not Being*). But these texts are also more ambitious and philosophically interesting.

Gorgias's *Encomium of Helen* undertakes to free Helen of Troy from blame for having abandoned her husband for the Trojan prince Paris, triggering the Trojan War. His method is argument by the exhaustion of alternatives. Helen's action must have been caused either by fate, necessity and the will of the Gods, by force, by persuasion, or by erotic love. That an action caused by the force of another person cannot be blamed is hard to deny; Gorgias's strategy is to assimilate the other possible causes to cases of force. Divine forces are stronger than human will, so if Helen's action was caused by them she cannot be blamed. As for persuasion, Gorgias here launches into a hymn to the powers of speech (*logos*)—a passage that is outsized relative to the whole and may well give away the real purpose of the *Encomium*. *Logos*, he says, is a "mighty ruler," and though a small body it controls the actions of many larger ones. (Gorgias seems to assume a scientific account of speech as composed of tiny sound particles that physically enter the audience's body through sensory pores—as in Plato, *Meno* 76c–d.) Speech is to souls as drugs are to bodies, causing involuntary reactions: So persuasion is a kind of compulsion. What gives *logos* this power is, somehow, the reliance of the human mind on fallible opinion (*doxa*), which is necessitated by our limited access to the truth. Finally, *eros* is assimilated to involuntary perceptual reactions. We cannot help the way things appear to us: some sights terrify, others seduce, and actions driven by such reflexes are again compulsory.

The quality of argumentation in the *Helen* is inevitably uneven, but its ingenuity is remarkable. The upshot of the argument as a whole is much debated. The causes of action itemized by Gorgias, such as fate and the way things appear to us, are extremely general and able to cause a wide range of our actions. So the upshot seems to be that any one of our actions would appear as involuntary, if only its causal origins were known in full—a claim that still figures in arguments about determinism and free will.

But this claim is far from explicit in Gorgias's text. To complicate matters, Gorgias opens the speech by saying that the "adornment" (the virtue or best state) of a speech is truth; but he closes by describing the encomium as "a plaything for myself." This playful, self-subverting presentation leaves us to judge the arguments and their implications for ourselves. If anything about the *Helen* is unequivocally serious, it is the miniencomium to *logos*, with its conception of language as an instrument of manipulation, a conception the *Helen* itself aims to display.

The *On Not Being* has a complex structure, comprising three parts: Part I argues that nothing exists, Part II that if anything did exist we could not know it, and Part III that even if something existed and we could know it, we could not communicate it to one another. (In summarizing, I will freely combine points from the two somewhat garbled versions of the text that have come down to us: One is in the pseudo-Aristotelian *On Melissus, Xenophanes and Gorgias (MXG)*, the other in Sextus Empiricus, *Adversus Mathematicos* VII. The two differ substantially in places and neither can be exactly what Gorgias wrote.) Part I argues by the exhaustion of alternatives: for instance, Being (or "what is" or "the existent") must be eternal or generated or both, but each option leads to an impossibility; similarly, if Being exists, it must be either one or many, but each is argued to be impossible. Part II argues that things thought are not existent, and that, therefore, Being is not thought. Here Gorgias raises the perennial philosophical problem of reference to nonexistent objects: We can think of a man flying or chariots running over the sea, but it does not follow that any such things exist. Gorgias seems to infer, fallaciously, that existent things and objects of thought differ in the sense that nothing can be both. But perhaps his real point is just that thoughts and their objects are different in kind, and the connections between them are unreliable: contra both Parmenides and Protagoras, we can and do think what is not (Caston 2002). (The very obscure fragmentary saying

of Gorgias on being and seeming, DKB26, must be relevant here, but it is difficult to say what it adds.)

Part III of the *ONB* deals with language, and as with the *Helen*, we may here approach the real point of the exercise. Gorgias argues that just as sight and hearing have their own proper contents (colors, sounds), so speech is of words, which are different from sensory contents and from the things themselves. So how can speech make clear things different from itself? How can it reveal objects or sensations we are not already familiar with? And how can the same thought be shared by two different people?

The *ONB* has often been read as a parody of Parmenidean philosophy. There are clear echoes of certain arguments made by Parmenides and the other Eleatics, particularly in Part I. And the overall upshot of the *ONB* is, as Kerferd (1981) has noted, to sunder three things that Parmenides had argued must coincide: what is, what can be thought, and what can be spoken. The question, then, is whether the *ONB* is merely satirical, both satirical and serious (cf. DK82B12) but purely negative and critical, or intended as positive doctrine in its own right. As positive doctrine it seems to be self-refuting. Nevertheless, scholars have attempted to find interpretations of its conclusions which lend them some plausibility. Mourelatos (1987) has noted that Part III can be read as arguing for conclusions that complement those of the *Helen*. Language cannot communicate either the natures of things or the thoughts of the speaker; the remaining possibility is that it is to be understood not as a system of representations but simply as an instrument of behavioral manipulation. Alternatively, Parts II and III could perhaps be read as arguing only that mental and linguistic items are by nature distinct and different in kind from their referents (and from each other), and, therefore, are inherently fallible and defective in representing them.

Parts II and III are also often likened to Protagorean relativism as presented in Plato's *Theaetetus*, our other most important source for sophistic epistemology. However, the two positions are very different. There is nothing relativistic about Gorgias's conclusions; moreover Gorgias in effect denies the possibility of *true* opinion and speech, whereas for Protagoras their *falsity* is impossible. Nevertheless, there is a family resemblance insofar as both can be read as essentially critical positions. They repudiate the metaphysical ambitions of philosophers like Parmenides, denying the possibility of a knowledge distinct from opinion and a reality distinct from appearance.

See also Parmenides of Elea; Protagoras of Abdera; Sophists.

Bibliography

The standard edition of Gorgias's texts and fragments is Diels-Kranz (1952) (though see also MacDowell [1982] for the *Helen*). However, Diels-Kranz prints only the text of the Sextus Empiricus version of the *On Not Being*; Dillon and Gergel (2003) usefully give translations of both versions and a comprehensive assortment of other texts.

Caston, V. "Gorgias on Thought and its Objects." In *Presocratic Philosophy: Essays in Honour of Alexander Mourelatos*, edited by V. Caston and D. W. Graham. Aldershot: Ashgate, 2002.

Cole, T. *The Origins of Rhetoric in Ancient Greece*. Baltimore: Johns Hopkins Press, 1991.

Diels, H. *Die Fragmente der Vorsokratiker*. 3 vols. 6th ed. Revised by W. Kranz. Berlin: Weidmann, 1951–1952.

Dillon, J., and T. Gergel. *The Greek Sophists*. London; New York: Penguin, 2003.

Kerferd, G. B. "Gorgias on Nature or That Which Is Not." *Phronesis* 1 (1955): 3–25.

MacDowell, D. M. *Gorgias, Encomium of Helen*. Bristol: Bristol Classical Press, 1982.

Mourelatos, A. P. D. "Gorgias on the Function of Language." *Philosophical Topics* 15 (1987): 135–170.

Verdenius, W. J. "Gorgias' Doctrine of Deception. In *The Sophists and their Legacy*, edited by G. B. Kerferd. Wiesbaden: Steiner, 1981.

Wardy, R. *The Birth of Rhetoric: Gorgias, Plato, and Their Successors*. London: Routledge, 1996.

Woodruff, P. "Rhetoric and Relativism: Protagoras and Gorgias." In *The Cambridge Companion to Early Greek Philosophy*, edited by A. A. Long. Cambridge, U.K.: Cambridge University Press, 1999.

Rachel Barney (2005)

GOTTSCHED, JOHANN CHRISTOPH

(1700–1766)

Philosopher, literary critic, and theoretician, Johann Christoph Gottsched was Christian Wolff's disciple and one of the architects of the German Aufklärung. Particularly conscious of Germany's cultural shortcomings, compared to France and England, Gottsched worked vigorously to reform German theater and poetry. Taking the ancients (Aristotle, Horace) as models, but also the French "Grand Siècle" (Racine, Molière, Boileau) and some few national examples (such as Martin Opitz), he wrote his *Versuch einer Critischen Dichtkunst* (1729, but often reedited until 1751) as a normative poetic theory destined to help form the taste of German writers and public alike. Gottsched's project, however, did not reduce

itself to this pedagogical goal: His poetics was meant to ground the rules of poetic taste on systematic philosophical foundations inherited for the most part from Gottfried Leibniz and Christian von Wolff. He saw it as imperative for both the philosopher and the serious poet that they know not merely the rules inherited from antiquity and French classicism, but that they also understand the reason underlying these rules. For Gottsched, criticism was a philosophical task, a part of Wolffian rationalism. In this sense, the *Critische Dichtkunst* prefigured the new aesthetical science set forth by Baumgarten a few years later.

Gottsched's theoretical positions are utterly rationalist. In keeping with Wolff and Leibniz, he conceives beauty as the clear yet conceptually indistinct representation of a perfection in an object—whether this object is natural, technical, or the result of poetic imagination. Being the perception of a perfection, the apprehension of beauty is accompanied by pleasure. Gottsched therefore rejects the subjective account of beauty: Aesthetic pleasure reduces itself to the perception of a perfection, the components of which could be made explicit. In other words, this perception could lead to rational knowledge and could thus be reduced to knowable rules. Every category of beauty, and every type of poetic or artistic beauty, rests on specific rules (those of architecture, of music, of painting, of tragedy, of epic) that nonetheless share some common fundamentals, namely, the notions of order, proportion, correlation between the parts and the whole, and the appropriateness of the rules to the specific function of the object.

The rules of poetry and liberal arts are therefore neither subjective nor variable; they are brought out by the best specialists of each domain and confirmed by experience and reflection. In this context, aesthetic taste depends on understanding as it judges the sensation of a beautiful thing. Good taste (that is, correct taste) consists, according to Gottsched, of "judging adequately, from a simple sensation, of the beauty of a thing for which we lack clear and distinct knowledge." This knowledge is "indistinct" because the person for whom this thing is pleasing is incapable of explaining the causes of the pleasure. Here, Gottsched's rationalism almost forces him into a contradiction: If taste is an indistinct judgment, does its improvement—which is the avowed goal of Gottsched's normative poetics—lead to the development and enrichment of aesthetic sensibility or does it rather perfect judgment and, hence, dissolve taste into knowledge? Only with Baumgarten's *Aesthetica*, and its notion of sensible knowledge, will this problem, inherent to any aesthetic rationalism, find a credible answer.

In his analysis of the "poet's character," Gottsched applies the Wolffian theory of the mind's faculties to Boileau's classic conception of poetic production. For Gottsched, the "divine gift" traditionally attributed to the poet comes down to having a natural disposition for poetic imitation. Among the faculties the poet must have, wit (*ingenium, Witz*), or the capacity to easily perceive similarities between things, is the most important. But the mind must also be supported by a strong power of imagination, which Gottsched understands as the power to reproduce concepts we have already had on the occasion of present sensations and on the basis of the principle of resemblance, and perspicacity, which consists in perceiving nuances and differences within things.

Merely having these faculties, however, is insufficient: They must be the object of education. Moreover, imagination, perspicacity, and wit are not the poet's or the artist's only requisite talents; art (all the disciplines pertaining to the practice of a particular art), erudition (mythology, history, geography) and a profound knowledge of human psychology are also necessary to the artist's character. He must also develop his judgment (*Beurteilungskraft*), which serves reason as an instrument to control an overheated imagination; judgment keeps wit within the limits of verisimilitude and the natural. Finally, the poet's character rests on an honest and virtuous disposition of the mind that depicts morally wrong actions as ugly and revolting. On this issue, Gottsched's aesthetics concurs with one of the central tenets of the *Aufklärung*, which holds, drawing from a conception leading back to Horace, that poetry's mission is to please while providing moral instruction.

Even if imitation is the essence of poetry, the fable constitutes its "soul." There are three degrees of poetic imitation: the vivid portrayal of natural things, the imitation of characters, sentiments, and human passions, and the plot or "fable" (*Fabel*). Referring to Leibnizian metaphysics, Gottsched describes the fable as the tale of an event, rich in moral truth, that did not really happen but that could have taken place in some possible world. Poetic fiction is the "history of another world" that must nonetheless be submitted to the principle of verisimilitude, which Gottsched defines as concordance with the general order of nature. Ensues a tension between two principles, that of the fabulous (which satisfies the taste for novelty, strangeness, and remarkableness, but risks lapsing into the unconceivable and extravagant) and of

verisimilitude, on which rests poetry's credibility and its capacity to serve a morally edifying function.

Attempting to give more importance to the freedom of creative imagination, Swiss critics Johann Jakob Bodmer (*Critische Abhandlung von dem Wunderbaren in der Poesie*, 1740) and Johann Jakob Breitinger (*Critische Dichtkunst*, 1740) distanced themselves from Gottsched on this issue. Mobilizing Milton's *Paradise Lost* and Pseudo-Longinus' treatise on the sublime as guides for their reflection, they aimed to encourage the fabulous in poetry and to grant a certain autonomy to the "truth of imagination" vis-à-vis the "truth of understanding."

See also Aesthetics, History of; Aristotle; Baumgarten, Alexander Gottlieb; Boileau, Nicolas; Leibniz, Gottfried Wilhelm; Longinus (Pseudo); Milton, John; Wolff, Christian.

Bibliography

WORKS BY JOHANN CHRISTOPH GOTTSCHED

Versuch einer critischen Dichtkuns. Darmstadt: WBG, 1962.

"Versuch einer critischen Dichtkunst." In *Ausgewählte Werke*. Vol. 6. Edited by Von J. Birke and B. Birke, New York and Berlin: De Gruyter, 1973.

WORKS ABOUT JOHANN CHRISTOPH GOTTSCHED

Birke, Joachim. *Christian Wolffs Metaphysik und die Zeitgenössische Literatur- und Musiktheorie*. Berlin: Gottsched, Scheibe, Mizler., 1966.

Dahlstrom, Daniel O. "Die Aufklärung der Poesie: J. C. Gottsched, 1700–1766: Critische Dichtkunst, 1729." *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft* 31 (1) (1986): 139–168.

Freier, Hans, and Kritische Poetik. *Legitimation und Kritik der Poesie in Gottscheds Dichtkunst*. Stuttgart: Metzler, 1973.

Daniel Dumouchel (2005)

GOURNAY, MARIE LE JARS DE
(1565–1645)

Marie le Jars de Gournay was the editor of the first complete text of Michel Eyquem de Montaigne's *Essais*; author of feminist, moral, and religious tracts; and a literary writer and theorist. Born into an aristocratic family in Paris, she mastered Latin and translated Diogenes Laertius's *Life of Socrates* in her youth. At eighteen or nineteen, having read with enthusiasm Montaigne's *Essais,* books 1 and 2, she met with the author, which inspired her novel. Their friendship led to her becoming his "adopted daughter," which, in the sixteenth century, implied a literary partnership. Thus, in 1594, Montaigne's widow sent her the final manuscript of his *Essais,* which Gournay edited, later annotated, and published, together with a long "Préface," in 1595.

The "Préface" attempts to defend Montaigne against the main criticisms advanced by his contemporaries: (1) Against the charge that his Latinisms and neologisms did harm to the French language, Gournay stressed the importance of Montaigne's usages. Gournay would later make a name for herself as the protectoress of ancient French words and would defend the innovative, metaphorical use of language against Malherbe and other moderns. (2) In response to Dominique Baudius's and Étienne Pasquier's claim that Montaigne's frank discussion of love was indecent, a point Blaise Pascal would later take up, Gournay argued that the ancients rightly took such discussion as a prerequisite for the self-knowledge needed for virtue. (3) The charge of philosophical obscurity was countered with a skeptical attack against the critics' capacity for judgment: "The gift of judgment is the thing in the world that men possess in more varied proportion." (4) Gournay defended Montaigne's digressive style against the objection that it precluded treating a topic thoroughly and evidenced a lack of method. Since Gournay and Montaigne were steeped in skepticism, Gournay could hardly imagine Montaigne producing rigorous, linear proofs. (5) The accusation of heresy, leveled especially at the "Apologie de Raymond Sebond," was the criticism Gournay was most anxious to refute. Her defense of Montaigne's religious orthodoxy is of particular interest, since it rests on one of the clearer statements that we have of his fideism—a doctrine that she shared: "Who, likewise, could tolerate these new Titans of our century, these scalers of the heavens, who think that they will manage to know God by their own means?" "Judgment alone puts us in direct possession of God: which is to know nothing of Him and to worship Him on the basis of faith." (6) Montaigne's focus on the self and use of confessional autobiography had been attacked as vain and pointless. Gournay argued that Montaigne was instructing us in the Platonic art of self-examination; she was one of the first to see the epistemic and moral significance of the first-person philosophical voice, which would play such an important role in the works of René Descartes and Jean-Jacques Rousseau. (7) Beginning with the 1625 edition, Gournay countered the charge that Montaigne was ignorant of the sciences by providing a skeptical, humanist understanding of a "true science": That which aids us in conducting ourselves as "honnêtes hommes" and in leading a good life. The sub-

jects of which Montaigne might have been ignorant were "pure scholastic amusements."

After defending the Jesuits in a pamphlet, for which she was attacked in print, Gournay published a collection of classical translations, and a feminist tract, *Egalité des hommes et des femmes* (The Equality of Men and Women; 1622). *Egalité* is arguably the first modern philosophical response to the *querelle des femmes,* or "woman question." Gournay's innovative contribution was to combine (1) skeptical attacks, including the use of *reductio* arguments, against traditional views on the intellectual and moral inferiority of women with (2) evidence on behalf of the thesis of equality based on the authority of holy scripture, the early church fathers, and the ancient philosophers whom the church has recognized. As a Christian skeptic and fideist, Gournay saw (1) and (2) as consistent.

Gournay's moral essays reflect not only Pyrrhonism and fideism but the Christian stoicism that made up part of her *morale provisoire.* They appear in her collected works: *L'ombre de la Damoiselle de Gournay* (The Shadow of Mademoiselle de Gournay, 1626) and *Les advis ou Les presens de la Demoiselle de Gournay* (The Advice and Presents of Mademoiselle de Gournay, 1634; 1641).

She corresponded with Anna Maria van Schurman, Justus Lipsius, Saint Francis de Sales, La Mothe le Vayer. Abbé de Marolles, and Cardinal Richelieu. In her final years Gournay participated in the salons of the Duchesse de Longueville and the Comtesse de Soissons; her own salon was, arguably, the seed from which the French Academy grew.

See also Descartes, René; Diogenes Laertius; Feminism and the History of Philosophy; Fideism; La Mothe Le Vayer, François de; Lipsius, Justus; Montaigne, Michel Eyquem de; Pascal, Blaise; Rousseau, Jean-Jacques; Skepticism, History of; Women in the History of Philosophy.

Bibliography

WORKS BY GOURNAY

Le proumenoir de Monsieur de Montaigne… (Paris, 1594). Delmar, NY: Scholars' Facsimiles and Reprints, 1985.

"Préface." In M. Montaigne, *Les essais de Michel Seigneur de Montaigne* (Paris, 1595). Various editions of the preface appeared in subsequent editions of Montaigne's *Essais.*

Adieu de l'Ame du Roy de France… avec La Défence des Pères lésuites. Paris/Lyon, 1610.

Versions de quelques pièces de Virgile, Tacite et Saluste.… Paris, 1619.

Eschantillons de Virgile (n.p., n.d.).

Egalité des hommes et des femmes (n.p., 1622). First modern edition in Schiff (below). English translation by E. O'Neill in *Social and Political Philosophy in Perspective: Classical Western Texts in a Feminist and Multicultural Perspective,* edited by J. Sterba (Belmont, CA, 1994), and in Bijvoet (below).

Remerciment au Roy. Paris, 1624.

L'ombre de la Damoiselle de Gournay. Paris, 1626. Included *Proumenoir,* essays on education, morals, feminist issues, religion, poetry and literary and philological topics, translations from the *Aeneid* and the works of Tacitus, Salust, Ovid, and Cicero.

Les advis ou Les presens de la Demoiselle de Gournay. Paris, 1634. Included the material in *L'ombre,* an additional translation from the *Aeneid* and new moral essays; the 1641 edition included "La vie de la demoiselle de Gournay."

Correspondence with Schurman is found in Anna Maria van Schurman, *Opuscula* (Leiden, 1648).

Correspondence with Lipsius is in J.-F. Payen, "Recherches sur Montaigne: Correspondance relative à sa mort," *Bulletin du Bibliophile* (1862).

The correspondence with Lipsius, Dupuy, and Richelieu, along with autobiographical, feminist, moral and literary essays, appears in *Fragments D'un Discours Féminin,* edited by E. Dezon-Jones (n.p., 1988).

WORKS ON GOURNAY

Albistur, M., and D. Armogathe. *Histoire du féminisme français,* Vol. 1. Paris: Femmes, 1977.

Baillet, A. *Jugemens des Savans sur les principaux ouvrages des auteurs.* Paris, 1694.

Bayle, P. *Dictionnaire Historique et Critique.* Paris, 1697.

Bijvoet, M. "Editor of Montaigne: Marie de Gournay." In *Women Writers of the Seventeenth Century,* edited by K. Wilson and F. Warnke. Athens: University of Georgia Press, 1989.

Boase, A. *The Fortune of Montaigne: A History of the Essays in France.* London, 1935.

Bonnefon, P. *Montaigne et ses amis.* Paris, 1898.

Feugère, L. J. *Les femmes poètes du XVIe siècle.* Paris, 1860.

Ilsley, M. H. *A Daughter of the Renaissance: Marie le Jars de Gournay, Her Life, and Works.* The Hague, 1963.

La Forge, J. de. *Le cercle des femmes sçavantes.* Paris, 1663.

McDowell Richardson, L. *The Forerunners of Feminism in French Literature from Christine of Pisa to Marie de Gournay.* Baltimore: Johns Hopkins Press, 1929.

Menagiana. Paris, 1754.

Richards, S. A. "Feminist Writers of the Seventeenth Century." MA thesis, 1914.

Sainte-Beuve, C.-A. *Tableau historique et critique de la Poésie Française et du Théâtre Français au XVIe siécle.* Paris, 1828.

Schiff, M. *La fille d'Alliance de Montaigne, Marie de Gournay.* Paris, 1910.

Somaize, A. de. *Le grand dictionnaire des Précieuses.* Paris, 1660.

Zedler, B. "Marie le Jars de Gournay." In *A History of Women Philosophers,* edited by M. E. Waithe, Vol. 2. Dordrecht: Kluwer Academic, 1989.

Eileen O'Neill (1996)

GRACIÁN Y MORALES, BALTASAR
(1601–1658)

Baltasar Gracián y Morales was a Spanish Jesuit and author of several baroque, obscure, laboriously polished books in which he expounded and illustrated conceptism, or metaphysical wit. Conceptism (from *concepto,* thought) is the quest for fine, brilliant, subtle thoughts expressed in antitheses, ambiguities, new words, and elaborate conceits.

Gracián published only one book under his real name, *El Comulgatorio* (Sanctuary meditations for priests and frequent communicants; 1655). A book of devotion, it enjoyed great success in several languages until the nineteenth century but is little used today. All his other books were published under pseudonyms without the permission of his superiors, for which offense he was disciplined because their subjects were thought too worldly for a priest, especially at a time when the Society of Jesus was struggling against Jansenism. The first was *El héroe* (1637), a portrait of a Christian political superman, similar to scores of books printed in Europe in the sixteenth and seventeenth centuries in answer to Niccolò Machiavelli's *The Prince.* Governance, then in the hands of absolutist rulers, was thought to be an art that could be taught in analytical character studies of the "exceptional man." Unfortunately, Gracián's model was Philip IV of Spain. However, in *El político* (1640) he took Ferdinand the Catholic as a more plausible subject for another panegyric on the Great Man. Such works fell from favor as government came to be conceived more democratically, but romanticism revived the cult of the hero, and Gracián's books were annotated by Napoleon Bonaparte and admired by Friedrich Nietzsche.

In *El discreto* (1646) Gracián continued his portraits of perfect types, descending to the level of the man of the world to describe the perfect gentleman as seen by provincial Spanish society. This book is remembered for its formula for the ideal life: First converse with the dead, then with the living, finally with oneself. In other words, first book learning, then travel and worldly experience, and last, meditation and preparation for death. From these three books, and others like them that remained unpublished, Gracián extracted an anthology of 300 aphorisms, published as *El Oráculo manual* (1647), or *Art of Worldly Wisdom.* These wise sayings have enjoyed constant success. La Rochefoucauld echoed many of them in his *Maximes,* and Arthur Schopenhauer translated them into German.

La agudeza y arte de ingenio (1642–1648) is a treatise on rhetoric and aesthetics that codifies the taste of the baroque age with its thirst for conceits, subtlety, eloquence, and artifice. Composed in a tortured hermetic style—the style Gracián praised as literary perfection—the book has never been translated into English. Literary historians consider it the beginning of the decadence of Spanish literature.

Gracián then wrote a quite unexpected book, for which his uneventful, sheltered existence offers no explanation but on which his fame rests. After extolling heroism, kingliness, savoir-faire, and poetic beauty in the works so far mentioned, he composed *El criticón* (1651–1657), a bitterly critical satire of the very society he had been exploring so complacently. It is a long philosophical novel, painstakingly allegorical and overadorned to the point of obscurity, which has been compared with the paintings of Hieronymus Bosch. A noble savage, Adrenio, is brought to Europe and shown all the inventions and refinements of civilization. Critilo, a man of the world, directs an "experiment" that becomes an inventory of human knowledge at that date, reviewing most of the questions that then interested moralists, scholars, and statesmen. The idea, later used by Blaise Pascal and Voltaire, was not new; such didactic allegories were already known in Spain, perhaps because Indians really had been taken on "conducted tours" of civilization there. What was striking was the extent to which Gracián's characters came to pessimistic conclusions; their judgments on civilization were uniformly unfavorable.

Long before Jean-Jacques Rousseau, Gracián said that, although man was created pure in the state of nature and left God's hands perfect, civilization corrupted and debased him. Yet, he added, as man grows older, he can acquire wisdom to free himself from slavery to worldly illusions, to begin the hard apprenticeship of renunciation and preparation for death. Gracián's pessimism was redeemed by faith in salvation beyond life. The world is not wholly bad; it is a mixture of true and false values, of misleading images and authentic shadows of eternal life. This world is a profoundly ambiguous tragicomic farce, with a concealed sense that is to be sought in another world of eternal being. This combination of extreme pessimism and a confident religious faith introduced a curious ambivalence into Gracián's view of the world, notably of the things that he most admired: social success, worldly glory, and political power. Perhaps on only one subject was he utterly pessimistic—woman, whom he called "a Satanic creature, vile, inferior." Schopenhauer agreed with him here, but on other matters the German misrepre-

sented the Spanish Jesuit's pessimism by taking it out of its religious context. To be sure, some critics have argued that Gracián's piety was pretense, designed to get his work past the Inquisition's censorship. Voltaire knew *El criticón,* so resemblances to *Candide* might not be accidental; but Gracián's clearest philosophical influence was over Schopenhauer and Nietzsche.

Bibliography

WORKS BY GRACIÁN Y MORALES

The Heroe. Translated by J. Skeffington. London, 1652.

The Critick. Translated by Sir Paul Rycaut. London, 1681.

The Complett Gentleman. Translated by T. Saldkeld. London, 1726.

The Art of Worldly Wisdom. Translated by Joseph Jacobs. New York: Macmillan, 1943. Also translated by Otto Eisenschiml. New York: Essential, 1947.

Obras completas. Madrid, 1960.

The Oracle: A Manual of the Art of Discretion. Translated by L. B. Walton. London and New York, 1963.

WORKS ON GRACIÁN Y MORALES

Acker, Thomas S. *The Baroque Vortex: Velázquez, Calderón and Gracián under Philip IV.* New York: Peter Lang, 2000.

Bell, A. *Baltasar Gracián.* Oxford: Oxford University Press, H. Milford, 1921.

Correa-Calderon, E. *Baltasar Gracián.* Madrid: Editorial Gredos, 1961. Includes a bibliography of the extensive literature on Gracián.

Foster, Virginia R. *Baltasar Gracián.* Boston: Twayne, 1975.

Hafter, Monroe Z. *Gracián and Perfection: Spanish Moralists of the Seventeenth Century.* Cambridge, MA: Harvard University Press, 1966.

Romera-Navarro, M. *Estudios sobre Gracián.* Austin: University of Texas Press, 1950.

Sánchez, Francisco J. "An Early Bourgeois Literature in Golden Age Spain: Lazarillo de Tormes, Guzmán de Alfarache and Baltasar Gracián." PhD diss. University of North Carolina, Chapel Hill, 2003.

Serna, Emilio H. "The Philosophy of Ingenium: Concept and Ingenious Method in Baltasar Gracian." *Philosophy and Rhetoric* 13 (1980): 245–263.

Spadaccini, Nicholas, and Jenaro Talens, eds. *Rhetoric and Politics: Baltasar Gracián and the New World Order.* Minneapolis: University of Minnesota Press, 1997.

Neil McInnes (1967)
Bibliography updated by Tamra Frei (2005)

GRAMSCI, ANTONIO
(1891–1937)

Antonio Gramsci, the Italian socialist born in Sardinia, founded the Italian Communist Party in 1921. He turned from political action to philosophical study when the Fascists jailed him in 1926 in order, said the public prosecutor, "to stop that brain working for twenty years." That brain, one of the most gifted that communism has produced, could not be stopped even by the inhuman conditions of Benito Mussolini's prisons: Gramsci filled three thousand pages with writing on a vast range of philosophical and political subjects. His frail health undermined by ill treatment, he died in Rome a week after his commuted term ended.

From the thirty-two prison notebooks, devoted editors have extracted "books" by grouping fragments on connected topics. In addition, *L'ordine nuovo* (1954) is a collection of articles from a socialist newspaper of that name, *Letteratura e vita nazionale* (1950) contains book and drama reviews, and *Lettere dal carcere* (1947) are his letters from jail. The heroic conditions under which he worked, and his founding role in Italian communism, may be responsible for overestimation of Gramsci's contribution to philosophy, but there is no doubt of his erudition and critical powers.

His philosophical notes (they seldom attain essay length) go beyond defense of Marxist doctrine; they mean to be a refutation of the idealism of two eminent ex-Marxists, Benedetto Croce and Georges Sorel. His criticism of them is seldom hostile, and in fact implies a disparaging opinion of orthodox Marxist-Leninist thought. Gramsci's central thesis is that Karl Marx's "materialist" overturning of G. W. F. Hegel was not a once-for-all affair that left communism, in Friedrich Engels's phrase, the secure "heir of the classical German philosophy." It had to be a continuing effort, to be repeated by each generation. Better, it was a Giambattista Vico-style cycle in which the same work of philosophical synthesis recurred at ever-higher levels. Gramsci saw that official Marxist thought in his day was in danger of relapsing into that vulgar materialism from which Marx's Hegelian training had rescued socialism a century before. Thus, it needed a new blood transfusion from speculative philosophy, a synthesis with neo-Hegelian idealism, notably with Croce and Giovanni Gentile.

This diagnosis entailed departure from the standard Marxist view on how philosophy "culminated" in revolutionary action. The last of Marx's Theses on Ludwig Feuerbach—"the philosophers have only *interpreted* the world in various ways; the point however is to *change* it"—had been taken to mean that philosophy would be realized in, and replaced by, revolutionary action. The postrevolutionary world would have no room for mere speculation, and philosophy would become the ideological branch of the administration. Gramsci replied that

philosophy could not be realized in, and even less supplanted by, political action. If the proletariat were to be "the heir of the classical German philosophy" (and if it were not, the revolution would be a cultural failure), it would have to pursue some recognizably philosophical activity. Specifically, it would be bound to go on reckoning with speculative idealism, putting it back on its feet as Marx did with Hegel and as Gramsci hoped to do with Croce.

Any one philosophy or system could indeed "culminate," or be realized. In fact, if it were a significant cultural product and not reverie, it surely would be. Yet that realization, the passage from speculation to action, from theory to practice, was not the "end" of philosophy foreseen by many Marxists. Rather, it was the transposition of private thinking into historically effective mass beliefs and a new ethic. Thus, Gramsci's program was to synthesize V. I. Lenin and Croce, to produce a reinvigorated Marxist philosophy that could be translated into a mass faith "like the Protestant Reform or the French Enlightenment."

The victory of such a new ethic was the essence of revolution, which meant above all "a moral and intellectual reform" and the "creation of a new integral culture." In all this, Gramsci explicitly followed Sorel but, against him, he denied that a "revolution of ideas" could do without politics, as Lenin practices it. His defense of Leninism, and of political organization generally, was a genuine contribution to political theory. Incidentally, it involved a fresh assessment of Niccolò Machiavelli. Gramsci wanted to be "both Robespierre and Kant," and indeed he succeeded in combining—at least on paper and in jail—the tough-minded political practicality of communism with a liberal attachment to classical education and philosophical culture.

See also Communism; Croce, Benedetto; Engels, Friedrich; Feuerbach, Ludwig; Gentile, Giovanni; Hegel, Georg Wilhelm Friedrich; Lenin, Vladimir Il'ich; Machiavelli, Niccolò; Marx, Karl; Marxist Philosophy; Sorel, Georges; Vico, Giambattista.

Bibliography

PRIMARY WORKS

Opere, 12 vols. Turin: Einaudi, 1947–1972.
Gli intellettuali e l'organizzazione della cultura. 1949.
Il Risorgimento. 1949.
Note sul Machiavelli, sulla politica e sullo stato moderno. 1949.
Passato e presente. 1951.
Il materialismo storico e la filosofia di Benedetto Croce. 1948.
The Modern Prince and Other Writings. Translated by Louis Marks. London: Lawrence and Wishart, 1957.
The Open Marxism of Antonio Gramsci. Translated and edited by Carl Marzani. New York: Cameron Associates, 1957.
Prison Notebooks. Edited by Joseph A. Buttigieg. Translated by Joseph A. Buttigieg and Antonio Callari. New York: Columbia University Press, 1992–1996.

SECONDARY WORKS

Bellamy, Richard, and Darrow Schecter. *Gramsci and the Italian State.* Manchester, U.K.: Manchester University Press, 1993.
Cammett, J. McKay. *Antonio Gramsci and the Ordine Nuovo Movement.* Ann Arbor, MI, 1960.
Femia, Joseph V. *Gramsci's Political Thought: Hegemony, Consciousness, and the Revolutionary Process.* Oxford: Clarendon Press, 1981.
Lombardo-Radice, L., and G. Carbone. *Vita di Antonio Gramsci.* Rome: Edizioni di Cultura Sociale, 1951.
Matteuci, N. *Antonio Gramsci e la filosofia della prassi.* Milan: Giuffrè, 1951.
Ottino, C. *Concetti fondamentali nella teoria politica di Gramsci.* Milan: Feltrinelli, 1956.

Neil McInnes (1967)
Bibliography updated by Philip Reed (2005)

GRAY, ASA
(1810–1888)

Asa Gray was a leading American interpreter of Charles Darwin's theory of evolution. Born in Sanquoit, in central New York, he became deeply interested in botany as a youth. Although he received a medical degree from Fairfield Medical School in 1831, he decided to devote his life to botanical studies, in which field he soon gained an international reputation. Harvard University appointed him Fisher professor of natural history in 1841, a post he held for over forty years. His writings both popularized the subject of botany and advanced it scientifically. Through his correspondence with Charles Darwin in 1856 and 1857, Gray obtained a preview of the theory of evolution by natural selection. When the *Origin of Species* was published, Gray wrote one of the first reviews, in the *American Journal of Science and Arts* (March 1860). This review, with several other essays on evolution, was reprinted under the title *Darwiniana* (1876).

Gray's attitude to Darwin's theory was open-minded but cautious. He regarded it as a plausible scientific hypothesis, although far from conclusively proved. As an explanation of the diversification of species, it was markedly superior to the doctrine of special creation. However, it did not really explain the origin of species because it failed to give a satisfactory account of the cause of variations. Gray thought that Darwin was often rash in drawing conclusions that outran the evidence, as when he asserted that all species must have descended from "four

or five primordial forms" and when he contended that man's mental powers must have had an evolutionary origin. This last contention "accumulates improbabilities beyond belief."

Against those who said that the Darwinian theory implied atheism, Gray argued that "it is neither atheistical in statement nor in intent." The theory could be given a nontheistic interpretation, but it could equally be given a theistic one. A central question was the presence or absence of design in nature as a whole, and this question was one for the natural theologian or the philosopher, not for the biologist. Gray himself favored a theistic interpretation, since the idea of a Designer of the universe "is most natural to the mind." It was not even true to say that Darwin's theory was mechanistic. It assumed that adaptations produced by natural selection are useful to organisms, enabling them to achieve certain ends, and this assumption clearly reintroduced purpose or teleology into natural history. "If purpose in this sense does not itself imply design, it is certainly compatible with it and suggestive of it."

Gray's cool analysis of Darwinism coupled with his espousal of theism irked some of Darwin's militantly agnostic supporters, though not Darwin himself, who valued Gray as a friend and searching critic.

See also Darwin, Charles Robert; Darwinism.

Bibliography

WORKS BY GRAY

Darwiniana: Essays and Reviews Pertaining to Darwinism. Cambridge, MA, 1876. New edition with an introduction by A. H. Dupree, Cambridge, MA: Belknap Press of Harvard University Press, 1963.

Scientific Papers. Boston: Houghton Mifflin, 1889.

Natural Science and Religion. New York: Scribners, 1880.

Letters of Asa Gray. Edited by Jane Loring Gray, 2 vols. Boston and New York: Houghton, Mifflin, 1893.

WORKS ON GRAY

Dupree, A. H. *Asa Gray: 1810–1888.* Cambridge, MA: Belknap Press of Harvard University Press, 1959.

Robinson, B. L. "Asa Gray." *Science* 62 (1925): 45–46.

T. A. Goudge (1967)

GREEK ACADEMY

Academeca was the name of a public park, equipped with a gymnasium and lecture facilities, located about a mile northwest of the Dipylon Gate of ancient Athens. There, probably shortly after 387 BCE, Plato bought a house and estate and began to teach, so successfully that his school dominated the facilities of the area, was named after the park, and continued until Justinian's closure of the pagan schools of philosophy in 529.

CLASSIFICATION

The main philosophical contributions of the Academy had been made by the time of Antiochus's death (c. 68 BCE); the different phases in this period were classified into Old Academy and New Academy or, by some ancients, as Old (Plato and his immediate successors), Middle (marked by Arcesilaus in the middle of the third century BCE) and New (dominated by Carneades in the second century BCE). To this were sometimes added a Fourth Academy (Philo of Larissa, head 110/109–80 BCE) and a Fifth Academy (under his successor, Antiochus). Broadly speaking, the Old Academy was occupied with problems posed by Plato, Middle and New with aspects of skepticism, and the Fifth with the eclecticism introduced by Antiochus. The history of Middle Platonism and Neoplatonism is distinct from that of the Academy, which was, however, a notable Neoplatonic center under Proclus in the fifth century.

EVIDENCE

Plato is the only leading Academic whose published works have survived, and they are not primarily internal documents of the school; our knowledge of the doctrines of his successors and of the functioning of the Academy is tantalizingly limited to fragmentary references, opponents' criticisms, and later summaries.

ORGANIZATION

At first the organization may have been informal and fluid, with Plato's personality and interests forming the center of gravity. At some point this fused into a nonfee-paying (fees were instituted by Speusippus), nonresidential corporate society, possibly in the form of a religious guild (*thiasos*), joined in the common worship of the Muses and in pursuit of truth at their shrine of learning (*Mouseion*), which Plato built on his estate. There were regular dining and other formal ceremonies. Plato appointed his successor; thereafter, the members elected their head, who held the original estate for the society and who governed until his death.

TEACHING

Apparently the teaching varied to some extent among the junior and senior members of the society. Plato's own

writings indicated that he had a practical aim in training young men and inspiring them with his political ideals, but he also suggested that the process required much time and study; thus, we find that the most important members remained within the school, researching and teaching, for years or for a lifetime (for instance, Aristotle and Speusippus). Practical political contributions of the Academy therefore came from senior members who were advising rulers or drafting legal codes rather than from former students who had chosen political careers. See, for example, the details given by Plutarch in *Moralia* (1126 CD); however, the extent of Academic influence is in dispute (see E. Zeller, *Die Philosophie der Griechen* II, 420, n. 1). The *Republic* declares mathematics and dialectic to be the fundamental studies, thereby placing the Academy in opposition both to the literary education of the rival school of Isocrates and to the training in techniques characteristic of some professional sophists. These subjects were chosen not only because their object was real Being, as Plato thought, but also because their method forced students into a Socratic self-learning and willingness to follow a communal discussion wherever the rational argument led.

What is certain is that instruction did not consist in the propagation of orthodox doctrine. There is a hostility to the lecture system in Plato's dialogues, which, together with the notorious failure of his public lecture on the Good, suggests that the main activity of the Old Academy lay in discussions and seminars (of which there is also some contemporary evidence in a comic fragment of Epicrates). Aristotle's mention of unwritten tenets and ascription to Plato of certain doctrines which do not appear in the dialogues (such as the equation of Forms and Numbers, the principles of Unity and Indeterminate Duality, mathematical entities of intermediate status between Forms and phenomena) have led some scholars to posit an esoteric oral teaching of Plato's, and others to suggest that Plato did no oral teaching at all in the Academy.

Both extremes, however, run counter to evidence. Aristotle may be reflecting Plato's part in some exploratory debates with senior members, for while the young and inexperienced had to be nursed and stimulated along educational paths by Plato, it is clear that no consideration of orthodoxy froze research. For such stimulation Plato posed problems, such as accounting for the movements of the planets, which produced a variety of answers including, according to Simplicius, Eudoxus's famous hypothesis of concentric heavenly spheres. We have evidence of equally lively and uninhibited debate in the Old Academy on the theory of Forms, mathematical metaphysics, classification, soul, good, and pleasure. It was at this higher level that the Academy was most successful and influential; in the conflict of educational ideals for schooling the young, the literary education of Isocrates and the rhetorical schools completely defeated in the elementary and state schools the philosophical, scientifically based Academy. However, the latter throughout its history preserved the high ideals of a society dedicated to the disinterested and independent inquiry after knowledge, ideals that succeeding ages—from the Lyceum of Aristotle to the present day—have recognized as models and standards for their own institutes of advanced study.

OLD ACADEMY

The head of the society naturally influenced all members, so the history of the Old Academy is largely an account of Plato's pupils. The most brilliant was Aristotle, a member from his eighteenth year until Plato's death in 347 BCE; his philosophy stems from the Academy, where his earliest works were written. There is some evidence that he was still being considered for the headship at Speusippus's death in 339 BCE and that he finally broke away to found his own school four years later.

Mathematical research was particularly distinguished in the Old Academy. Theaetetus of Athens, tragically killed in battle in 369 BCE, succeeded in generalizing the theory of irrationals and in constructing and circumscribing the five regular solids, thereby laying the foundations of Euclid's solid geometry in Books 10 and 13. Still more important was Eudoxus of Cnidus, who with his pupils joined forces with Plato (c. 367 BCE) for a few years. Apart from the influential astronomical theory of concentric spheres, he is credited with a general theory of proportion and a method of exhaustion fundamental to Greek geometry. He criticized Plato's theory of Forms, arguing that Form, to be a cause, must be immanent in phenomena, as white is the cause of whiteness of that in which it is mixed. He apparently regarded Form as a kind of substance. This seems to have drawn reactions from Plato and Aristotle. Eudoxus's championship of hedonism likewise produced opposing arguments from Plato and Speusippus. Another astronomer, Philippus of Opus, is reported to have edited Plato's *Laws* and to have written its appendix, the *Epinomis*; whether written by him or by Plato, the latter is an important document of a stage in the Academy at which mathematical astronomy advanced from a propaedeutic science to the central science of Being and theology.

Heraclides of Pontus (c. 390–c. 310 BCE), temporary head of the Academy during Plato's third Sicilian journey, unsuccessful candidate in the election of 339 BCE, and later associated with the Lyceum, had a Pythagorean bias and was a prolific, learned, and elegant writer on a wide variety of subjects rather than an original thinker. He appears to have responded to the problem of the planets' motion, but it remains uncertain whether he partially anticipated Aristarchus's heliocentric system and posited the rotation of Earth on its axis, as was thought. He posited an atomic theory of irregular units of mass but attacked the mechanical atomic theory of his day, holding that matter was subject to a divine teleology.

Speusippus, who succeeded his uncle Plato in 347 BCE, may well have been the senior member or even a founder of the society. He and the third head, Xenocrates of Chalcedon (c. 396–314 BCE, head 339–314 BCE), another pupil of Plato's from his earliest years, were thought by Aristotle and others to be concerned with similar problems. Although neither was a great philosopher, a study of fragments of their writings reveals a development of trends apparent in Plato's later work, some positions between Plato and Aristotle, and some foreshadowing of later Platonism. Both, but especially Speusippus, were strongly influenced by the current Pythagorean fashions in the Academy. In the school debate over the division of Being into the three spheres of Forms, Mathematicals, and Physicals, Speusippus replaced Plato's Forms with mathematical entities and Xenocrates identified Forms and Mathematicals. Both were preoccupied with the derivation of a hierarchy of substances from mathematical first principles—Speusippus disjointedly, according to Aristotle, in that he abandoned the uniform interdependence of the whole universe on one set of first principles and assumed different principles for different kinds of substance in series. Xenocrates posited the Platonic Unity and Indeterminate Dyad (or continuum of opposites such as great and small). Good, which was derivative for Speusippus (coming at the end of becoming), was distinguished from both Unity and Being. Soul fell into the mathematical classification—as self-moving number for Xenocrates and seemingly as a form of extension for Speusippus, a theory of great importance for the later Stoic, Posidonius.

Another Pythagorean trait was the strong theological interpretation of their mathematical cosmology, also hierarchical in treatment, Xenocrates advocating the influential doctrine of *daimones,* animate beings between gods and men. In dealing with the Academic problem of real definition by *diaeresis* (division) and classification, Speusippus suggested in an important work, *Homoia* (Resemblances), that definition by division was impossible without knowing all existing things, the essential nature of any one concept being constituted solely in its relation of likeness and difference to every other concept. Xenocrates foreshadowed Aristotle by asserting the logical and ontological priority of species over genus. In epistemology both continued the trend of allowing more cognitive importance to perception; Xenocrates, who had a weakness for triadic systematization, worked out spheres of the universe corresponding to cognitive powers: the sphere within the heavens as perceptible, that outside the heavens as intelligible, the heavens themselves as a mixture of both and thus objects of opinion. Both men were prolific writers on practical ethics; they held that happiness can come into being from virtue alone but that virtue is not the only good. Speusippus campaigned against pleasure as being contrary to both pain and good.

The next head, Polemon of Athens (elected 314 BCE), concentrated on conduct, elaborating Xenocrates's conception of happiness as life "in accordance with nature," a phrase which, especially through the Stoa, became the center of ethical debate. Polemon had more personal influence than philosophical originality, but he is of some importance for the Academy of the first century BCE. His friend, Crantor of Soli in Cilicia (c. 335–c. 275 BCE), wrote a famous work, *Peri Penthous* (On grief), a prototype of an ethical genre later popular, directed against the extreme views of the Stoics on pain and the affections.

MIDDLE AND NEW ACADEMIES

Crantor's reaction to the Stoa heralded a major change in the school. Until Crates of Athens (elected head 270 BCE), the main topics of inquiry were Platonic questions and developments. The next head, Arcesilaus of Pitane (316–241 BCE), is reputed to have concentrated on an attack on the Stoic theory of knowledge. He was probably reacting not only to an ontology and epistemology inimical to the Platonic tradition but even more to the dogmatic character of Stoicism, which he countered by an exaggerated form of Socratic skepticism: Not only did he know, like Socrates, that he knew nothing, but he is also said to have doubted whether he could ever know that he knew the truth. He imitated Socrates in writing nothing and taught mainly in open debate, introducing to the school the system of arguing on both sides of a question. Nevertheless, he saw to it that Plato's works were studied; and in the controversy with the Stoa, the Academy maintained that by "suspension of judgment" (*epoche*) they

were not inhibited in their main philosophic task of a continual search for truth, however unattainable, or from moral action, as a guide for which Arcesilaus recommended "the reasonable." The dialectical influence of Arcesilaus set the Academy firmly in the main stream of Greek skepticism, but it was an age of formalization through controversy. Arcesilaus helped to produce Chrysippus, the great fortifier of Stoicism; Chrysippus was in turn the whetstone of the most brilliant figure of the second century BCE, Carneades of Cyrene (214–129 BCE), who systematized a comprehensive and devastating skeptical attack against the whole philosophy of the dogmatic schools. But while Carneades had penetrating observations on sense perception, probability, causation, fatalism, and anthropocentric theology, he seems, both in method and content, to have drifted some way from the original Socratic-Platonic tradition.

Reaction began with Philo of Larissa (160/159–80 BCE, head from 110/109), Cicero's teacher, who, while maintaining skepticism against Stoic epistemology, reclaimed his Platonic ancestry. It was completed with Philo's pupil and opponent, Antiochus of Ascalon (c. 130–c. 68 BCE, head from c. 79 BCE), who came to think that the representatives of the New Academy had perverted the teaching of the Old Academy by maintaining that no truth could be grasped, thereby destroying any criterion of true and false and, in Antiochus's opinion, inhibiting action through denial of certainty. Antiochus reinstated a dogmatism whereby a criterion of truth arising from but not contained in the senses was grasped by the mind as self-evident. His reinterpretation of Platonism was marked by an eclecticism in ethics that enabled him to propose that the Stoa and Peripatos had merely followed the lead of the Old Academy, differing from it in terms rather than in substance.

In fact, the theories of morals of the three schools were all naturalist, and Antiochus's distorted arguments were facilitated by certain shared areas of discussion, covered by Polemo and the early Peripatos and Stoa, based on "the things according to nature" (*ta kata physin*) to which human beings had a natural affinity (*oikeiosis*). Doubtless there had been interchange of ideas, which was fostered even by the New Academy's insistence on arguing on both sides of the question, but Antiochus blurred the quite different approaches of the schools to a common area of dispute.

His thesis may have been further encouraged by the early "Platonic" works of Aristotle which were then popular, and influenced by the Stoic Panaetius, who had admired Plato and who gave greater prominence in his teaching than had earlier Stoics to the "intermediate natural things" (health, wealth, etc.), which, although the material of ethics, were held by Stoics to be in themselves of only relative value and morally neutral compared with the absolute value of the rational operation of virtue. Antiochus, however, maintained that the end of action, the happy life (*beata vita*), although possible through virtue alone, was completed (*beatissima*) by bodily and external goods. Thus Antiochus shared a graded axiology with the Old Academy and Peripatos; for him the difference between virtue and other goods was one of degree, for the Stoa there was a difference of kind. In a manner similar to some Stoic arguments, he held that the chief good was based on natural instincts for self-preservation and self-development, so that from the germ of virtues in the impulses of childhood man gradually attained knowledge of his own nature; but for Antiochus the perfection of human nature involved all parts of it, not only the highest, and also man's relationship to others and to the community.

This attempted dogmatic synthesis of the three great schools was of minor philosophical interest in itself, but of major importance for subsequent Greek thought. Apart from professional Academics, Antiochus profoundly influenced the popular expositor Cicero, some Stoics with doxographic interests like Arius Didymus, who taught at the court of Augustus, and Middle Platonists such as Albinus of Smyrna, whose lectures Galen attended in 151/152. Albinus's markedly eclectic epitome of Platonism (the *Didaskalikos*) still survives. Indeed, without the growth of syncretism initiated by Antiochus, the fusion that created the final explosion of ancient thought in Neoplatonism would not have been possible.

Throughout the long history of the Academy, the founder's works were studied and his birthday revered with celebrations. It is a remarkable tribute to his personality, philosophy, and educational ideals that through the very different phases of the school all members considered themselves his true heirs, so that one man's ideas stimulated his pupils over nine hundred years without in any way rigidifying their thought.

See also Alcinous; Antiochus of Ascalon; Arcesilaus; Aristotle; Carneades; Good, The; Neoplatonism; Philo of Larissa; Plato; Platonism and the Platonic Tradition; Plutarch of Chaeronea; Proclus; Sextus Empiricus; Simplicius; Skepticism, History of; Stoicism.

Bibliography

Some continuous texts are Diogenes Laërtius, *Lives*, Books III and IV; Aristotle, *Metaphysics*, Books A, M, and N; Cicero, *Academica* and *De Finibus*, Books IV and V; Sextus Empiricus, *Outlines of Pyrrhonism* and *Against the Dogmatists*; Siegfried Mekler, *Academicorum Philosophorum Index Herculanensis* (Berlin: Weidmann, 1902).

Fragments include P. L. Lang, *De Speusippi Academici Scriptis* (Bonn, 1911); Richard Heinze, *Xenocrates* (Leizpig, 1892); Fritz Wehrli, *Herakleides Pontikos* (Basel: Schwabe, 1953); Georg Luck, *Der Akademiker Antiochus* (Bern: Haupt, 1953); and C. J. de Vogel, *Greek Philosophy*, Vol. II (Leiden, 1953; 2nd ed., 1960) and Vol. III (Leiden: Brill, 1959).

A standard work is Eduard Zeller, *Die Philosophie der Griechen in ihrer geschichtlichen Entwicklung*, 6th ed. (Hildesheim, 1963), II.1.2 and III.1.1 and 2.

Introductory works on the Old Academy include G. C. Field, *Plato and His Contemporaries* (London: Methuen, 1930; 2nd ed., 1948); Werner Jaeger, *Aristotle* (Berlin: Weidmann, 1923), translated by Richard Robinson (Oxford: Clarendon Press, 1934); A. E. Taylor, *Plato, the Man and His Work* (London: Methuen, 1926); Sir David Ross, the introduction to Aristotle's *Metaphysics* (Oxford, 1924; 2nd ed., 1953); Hans Herter, *Platons Akademie* (Bonn, 1952); and C. B. Armstrong, "Plato's Academy," in *Proceedings of the Leeds Philosophical Society* (1953): 89ff.

Important works of detail include Léon Robin, *La théorie platonicienne des idées et des nombres* (Paris: Alcan, 1908); Julius Stenzel, *Zahl und Gestalt bei Platon und Aristoteles* (Leipzig and Berlin: Teubner, 1933); H. F. Cherniss, *Aristotle's Criticism of Plato and the Academy* (Baltimore: Johns Hopkins Press, 1944) and *The Riddle of the Early Academy* (Berkeley: University of California Press, 1945); Sir David Ross, *Plato's Theory of Ideas* (Oxford: Clarendon, 1953); Philip Merlan, *From Platonism to Neoplatonism* (The Hague: M. Nijhoff, 1953) and "Zur Biographie des Speusippos," in *Philologus*, 103 (1959): 198ff.; Paul Wilpert, *Zwei Aristotelische Frühschriften über die Ideenlehre* (Regensburg: J. Habbel, 1949); and Shlomo Pines, *A New Fragment of Xenocrates and Its Implications* (Philadelphia: American Philosophical Society, 1961).

See also H. J. Krämer, *Arete bei Platon und Aristoteles* (Heidelberg: C. Winter, 1959), and Konrad Gaiser, *Platons ungeschriebene Lehre* (Stuttgart: Klett, 1968). The latter contains a very convenient collection of *Testimonia Platonica: Quellentexte zur Schule und mündlichen Lehre Platons*, on pp. 441–557.

For works on the Middle and New academies, see Victor Brochard, *Les sceptiques grecs* (Paris: Nationale, 1887); Albert Goedeckemeyer, *Die Geschichte des griechischen Skeptizismus* (Leipzig: Dietrich, 1905); M. M. Patrick, *The Greek Skeptics* (New York: Columbia University Press, 1929); Léon Robin, *Pyrrhon et le scepticisme grec* (Paris: Presses Universitaires de France, 1944); Olof Gigon, "Zur Geschichte der sogenannten Neuen Akademie," in *Museum Helveticum* 1 (1944): 47ff.; Max Pohlenz, *Die Stoa*, Vol. I (Göttingen: Vandenhoeck & Ruprecht, 1948) and Vol. II, 2nd ed. (Göttingen, 1955); R. E. Witt, *Albinus and the History of Middle Platonism* (Cambridge, U.K.: Cambridge University Press, 1937); and A. Weische, *Cicero und die Neue Akademie* (Münster: Aschendorff, 1961).

I. G. Kidd (1967)

GREEK DRAMA

When Plato spoke of an ancient quarrel between poetry and philosophy, and when Aristotle said that poetry is more serious and more philosophical than history, they were both thinking primarily of epic and dramatic poetry, especially of tragedy. The reason why the two great Greek philosophers paid so much attention to dramatic poetry is closely connected with the reasons why Greek tragedy continues to deserve the interest of philosophers today. An account of philosophical ideas in "Greek drama" can usefully begin with some consideration of ancient philosophical ideas about Greek drama.

PLATO AND ARISTOTLE

It is well known that Plato was hostile to what he regarded as the inflated pretensions of the poets as moral and religious teachers and to the arrogant claims made on their behalf by rhapsodes and expositors. In numerous dialogues (notably in *Apology, Ion,* and *Republic*) he reiterated the complaint that poets lack the knowledge that, he believed, can be achieved only by rigorous philosophizing and that is necessary for the understanding of the human situation and the ordering of human life. The poets pronounce on life's problems without being able to "give an account" (*λόγον διδόναι*) of themselves and their ideas. Plato might ironically allow that, like conventional statesmen, they have some divinely inspired glimpses of moral and political truth, but he insisted that they lack the true knowledge that is achieved by the philosopher after strenuous dialectical thought.

Tragedy is essentially a kind of rhetoric (*Gorgias* 502D), and Plato reviled it with all the passion that he displayed in his attacks on forensic and political oratory. All these are the arts of mere persuasion, and they are customarily used to persuade men of what is false. Plato explicitly held that most orators, politicians, and poets are dishonest or ignorant, or both, and even the most famous of them would be refused admission to the ideal republic.

One of Plato's most important grounds for despising literature was based on the theory of Forms. The poet deals in the concrete and particular; dialectic, like its mathematical archetype, is concerned with the abstract

and universal. It follows that even an honest poet must inevitably fail to achieve and convey knowledge and understanding, since he is operating at entirely the wrong level. He presents images of images; he and his audience are at two removes from the world of reality and truth (*Republic* 597E).

Aristotle, the philosopher of the concrete particular substance, with his keen interest in the actual particular specimen, was more sympathetic to poetry and literature. Poetry is philosophical because it portrays the nature of man in general by presenting particular individual men in such a way that each portrait throws light on other individuals, just as the biologist studies the genus or the species by attending in detail to actual particular specimens. A chronicle of "what Alcibiades did or suffered" tells us only about Alcibiades. Oedipus or Agamemnon is Man as well as *a* man. Aristotle regarded the poets as contributors to thought, knowledge, and understanding, not as mere entertainers.

The opposite views of Plato and Aristotle on the value of literature must not be allowed to hide the importance of a point on which they agreed. Plato's attack on poetry, like Aristotle's more sympathetic treatment, presupposes that there is an overlapping of function between philosophy and literature. Plato thought that the poets gave wrong answers to the questions and problems that he dealt with in his dialogues, but the very form of his attack implies his recognition that the poets are also concerned with those questions and problems. The old battle between philosophy and poetry could not take place at all unless the two parties shared at least a battleground.

This point is confirmed, and its importance is underlined, by further knowledge of the history of ancient literature and philosophy. It was only in the time of Plato and Aristotle themselves that there began to appear any very clear distinctions between philosophy, history, science, and imaginative literature. Homer had fulfilled all the functions that were later divided among historians, tragic and comic poets, philosophers, theologians, moralists, and scientists. Parmenides and Empedocles were poets as well as philosophers; they did not write both poetry and philosophy—their poetry was their philosophy. Heraclitus wrote in prose but in an oracular, literary manner. Hesiod is part of the history of philosophical and cosmological speculation as well as of the history of literature. Plato himself wrote philosophy that is also literature and, in spite of his own strictures, imaginative and dramatic literature.

MODERN VIEWS

Modern controversy about Greek tragedy has followed similar lines. Many scholars and critics have praised the Attic tragedians as religious and moral thinkers and prophets, thus accepting Plato's view of the nature of the aims and themes of the ancient plays while often strongly dissenting from his valuation of them. Others, by contrast, have denied that the tragedians either showed or meant to show any moral or religious depth or originality, and have presented them as "mere" poets and playwrights whose purpose was purely literary and dramatic and who used traditional mythological and religious material simply because it was traditional.

This dispute is misconceived and is based on a false dichotomy. It not only ignores the artificiality of any attempt to draw a sharp distinction between literature and philosophy in Greek times but also involves drawing a distinction between them that is too sharp to be faithful to the nature and function of literature and philosophy in any age, including our own. Both parties to the dispute share Plato's mistaken assumption that nothing can count as philosophical, religious, or moral thought unless it is explicitly and formally general and systematic. Aristotle's recognition that fifth-century tragedy illuminated morality and religion by a dramatic presentation of particular events and characters needs to be extended to literature in general. The themes of Greek tragedy are the themes of literature: Man, God, Nature, Chance, Freedom, Will, Fate, Necessity, and Good and Evil. Most, if not all, of these themes are also themes of great and permanent philosophical interest, and philosophers should not despise the contributions of dramatists, poets, and novelists to our understanding of them.

TRAGEDY

Aeschylus, Sophocles, and Euripides had much in common: they all drew their characters and plots from a common stock of religious mythology and historical legend inherited from Homer, and they all used their plays as means of presenting the relations between men and gods. The stories of the Trojan War, of Agamemnon and the House of Tantalus, of Oedipus and the House of Cadmus recur in the works of all three. (A "history play," such as *The Persians* of Aeschylus, is exceptional, although commentators have found historical and political allusions in many of the plays that are based on mythical themes.)

Aeschylus and Sophocles were relatively orthodox in their treatment of the traditional themes. The *Oresteia* of Aeschylus presents, through the story of the working out of a family curse, a study of the conflict between man's

efforts to choose and guide his own life and the almost irresistible weight of past events and external influences. Agamemnon "puts on the yoke of necessity" when he chose to sacrifice his daughter Iphigeneia rather than to risk the ruin of the Greek expedition against Troy. In that phrase and in that incident Aeschylus combines an awareness of the force of circumstance with a consciousness of the responsibility that a man bears for his own actions, however circumscribed they are by what lies outside his control. The yoke is a yoke of necessity; but it is Agamemnon who puts it on. In the same trilogy Aeschylus portrays the growth of revenge ("a kind of wild justice," as Bacon called it) into the cultivated plant of civil law. His *Prometheus Bound* is also concerned with conflict. The struggle between Prometheus and Zeus is also the struggle between man's aspiration after knowledge and power and the forces of nature and environment represented by the gods. Men pay in suffering for every step in understanding.

The *Oedipus Tyrannus* of Sophocles also shows the price that must be paid for knowledge and self-knowledge. The man who answered the riddle of the Sphinx finds, when he knows his own nature and his own circumstances, that not all knowledge is sweet, and blinds his eyes that have seen too much. And yet it was ignorance that led him to his tragic end. If he had known more and had known it sooner, he would have done better and fared better. In the *Antigone* we find the "conflict of right with right" that led to G. W. F. Hegel's definition of tragedy; the legitimate claims of Creon, the civil power, are set against Antigone's ardent loyalty to holy and unwritten laws.

Euripides used the same stock of mythical material but in a different spirit. He was a friend of Anaxagoras and a student of the sophists and orators, whose influence is seen in his set debates and rhetorical speeches. The sophistry of Hippolytus ("my tongue it was that swore; my heart is not forsworn") and the atheism of Bellerophon are only two examples of the "free thought" of some of his characters that shocked Aristophanes and other conventional Athenians. It was debated, and is still debated, whether Euripides was himself an atheist or a modernistic theologian. To the modern reader of the plays the question is of merely academic interest. In the *Hippolytus* and the *Bacchae* he vividly presents conflicts between Aphrodite and Artemis and between Dionysus and the forces of order and restraint. The impact of these conflicts on a modern reader is not much affected by questions about whether Euripides literally believed in the gods of the Greek pantheon or merely used them as personifications of forces in human nature that are as familiar to us as they were to his original audience.

In reading Greek tragedy, as in reading any work of imaginative literature, we must beware of attributing to the author the opinions and attitudes expressed by his characters. The best Greek tragedies are as dialectical as the works of Plato. The issues they deal with are too complex and subtle to allow a neat answer to be given to them in the speeches of any one character. The dramatist presents and portrays; he does not argue and declaim.

COMEDY

Attic comedy is of little more than historical interest from the philosophical point of view. The *Clouds* of Aristophanes pillories and parodies a "Socrates" who is made to represent all that was new and disturbing in contemporary Greek thought. Aristophanes shows here and in several other plays (especially in the *parabasis,* or address of the chorus to the audience on current topics) that he was a conservative who looked back to the golden days of Aeschylus and the other "men of Marathon." His satirical purpose could best be served by ignoring the great diversity in the movements of thought that he disliked: Pythagorean and Orphic mysticism, natural speculation, sophistic attacks on conventional morality and religion, and the revolutionary theology of Euripides.

Comedy, like tragedy, was religious in its origin, and Aristophanes, if read with caution, can contribute to our knowledge of the history of Greek religious thought. But no comic writer can be trusted very far as a source of information on philosophical or scientific thought, and in particular Aristophanes' account of Socrates needs more delicate handling than it has received from scholars preoccupied with the "Socratic question." What the *Clouds* does show is that philosophical speculation was of some interest to the general public in Athens, even if only as a butt of jokes and gibes. Socrates testifies in Plato's *Apology* that attacks by comic poets helped to foster prejudice against him.

The contest between Aeschylus and Euripides in the *Frogs* of Aristophanes is one of the earliest examples of literary criticism, and it preaches, as Plato did, that the poet's function is to edify and instruct his audience.

See also Anaxagoras of Clazomenae; Aristotle; Empedocles; Hegel, Georg Wilhelm Friedrich; Heraclitus of Ephesus; Homer; Humor; Literature, Philosophy of; Orphism; Parmenides of Elea; Plato; Pythagoras and Pythagoreanism; Socrates; Tragedy.

Bibliography

Adkins, Arthur W. H. *Merit and Responsibility: A Study in Greek Values.* Oxford: Clarendon Press, 1965.

Gill, Christopher. *Personality in Greek Epic, Tragedy, and Philosophy: The Self in Dialogue.* Oxford: Clarendon Press, 1996.

Halliwell, Stephen. *The Aesthetics of Mimesis: Ancient Texts and Modern Problems.* Princeton, NJ: Princeton University Press, 2002.

Havelock, E. A. *Preface to Plato.* Oxford: B. Blackwell, 1963.

Hornblower, Simon, and Antony Spawforth, eds. *The Oxford Classical Dictionary.* 3rd ed. Oxford: Oxford University Press, 2003. See the articles on Aeschylus, Sophocles, Euripides, tragedy, Aristophanes, comedy, and philosophers on poetry.

Kitto, H. D. F. *Greek Tragedy.* London: Methuen, 1939.

Lesky, Albin. *Greek Tragic Poetry.* Translated by Matthew Dillon. New Haven, CT: Yale University Press, 1983.

Lloyd-Jones, P. H. J. *The Greeks.* London: Watts, 1962.

Lucas, D. W. *The Greek Tragic Poets.* 2nd ed. London: Cohen and West, 1959.

Nussbaum, Martha C. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy.* Cambridge, U.K.: Cambridge University Press, 1986; 2nd ed., 2001.

Segal, Charles. *Interpreting Greek Tragedy: Myth, Poetry, Text.* Ithaca, NY: Cornell University Press, 1986.

Verrall, A. W. *Euripides the Rationalist.* Cambridge, U.K.: Cambridge University Press, 1895.

Williams, Bernard. *Shame and Necessity.* Berkeley: University of California Press, 1993.

Williams, Bernard. "The Women of Trachis: Fictions, Pessimism, Ethics." In *The Greeks and Us: Essays in Honor of Arthur W. H. Adkins,* edited by Robert B. Louden and Paul Schollmeier. Chicago: University of Chicago Press, 1996.

Renford Bambrough (1967)
Bibliography updated by Grace Ledbetter (2005)

GREEN, THOMAS HILL
(1836–1882)

Thomas Hill Green, the English idealist philosopher, was born the fourth son of a Church of England clergyman at Birkin in Yorkshire. His mother died when he was only a year old. Green received his early education from his father before going at the age of fourteen to Rugby School, which had been reorganized on distinctive lines by Thomas Arnold a few years earlier. The corporate side of life at Rugby had little appeal for Green, but his fellow scholars were already impressed by his seriousness and independence of mind. Academically, he was able but not outstanding. In 1855 he entered Balliol College, Oxford, where he was an undergraduate for the next four years. Green was only a moderate classical scholar, but he got first-class honors when he took the final examinations in *Literae Humaniores,* preparation for which gave him his first serious work in philosophy. He was elected a fellow of Balliol in November 1860 but did not get a regular teaching post there for several years. In 1863 he refused an offer of the editorship of the *Times of India,* then being started in Bombay; in 1864 he was an unsuccessful candidate for a philosophy chair at the University of St. Andrews. In 1865 and 1866 he served as assistant to a Royal Commission investigating school education in England and Wales, working mainly in and around Birmingham. From 1866 onward he was heavily engaged at Balliol, where he became the first nonclerical tutor; by 1870, the year in which Benjamin Jowett became master, much of the administration of the college had fallen on his shoulders. He continued to serve as a tutor until 1878, when he was elected Whyte's professor of moral philosophy at the university.

By this time Green had married Charlotte Symonds (1871) and had published his first major philosophical work, the long critique of empiricism which constitutes the introduction to the edition of David Hume's works, which he produced with T. H. Grose. He had also begun to take an active part in social work and in local politics. From 1872 on he was prominent in the temperance movement (one of his brothers was a hopeless drunkard), and in 1876 he became a member of the Oxford town council, being the first active teacher in the university to hold such an office. He also played a major part in a movement to found a new high school in Oxford. Unfortunately, however, his health deteriorated sharply during these years, and matters were not improved by the added lecturing duties of the professorship, which Green undertook with characteristic thoroughness, writing out his lectures in full. He had long planned a major work on moral philosophy, but his *Prolegomena to Ethics* was still incomplete when he died in 1882. It was published by A. C. Bradley the following year. Green's other philosophical and miscellaneous writings were collected in three volumes by R. L. Nettleship, who also wrote the long memoir printed in Volume III.

Green has been described as the first professional philosopher in the modern sense; he was certainly one of the first specialized teachers of the subject in Oxford. But he exercised influence in many spheres outside philosophy. His work as a Balliol tutor did much to produce a distinctive type of Oxford graduate, unknown in the mid-nineteenth century: hardworking, intensely serious, aware to a surprising degree of social problems and realities. In politics he was important not only because of what he did to ease relations between "town and gown" in

Oxford but also for his pronounced radicalism: He was a strong supporter of John Bright against Lord Palmerston and of the cause of the North in the American Civil War. His essay "Liberal Legislation and Freedom of Contract" (1881) is important for its criticisms of pure laissez-faire liberalism and can be seen as anticipating the doctrine of the welfare state. Theologically, Green was not strikingly original, but the low value he set on dogma and historical tradition was certainly not without its effect. By insisting on the independent authority of philosophy he may well have persuaded many intending ordinands to take up other careers. Although very different in his philosophical views and immediate disciples from his contemporary Henry Sidgwick, he made much the same contribution to the secularization of Oxford as Sidgwick made to the secularization of Cambridge.

CRITIQUE OF "POPULAR PHILOSOPHY"

A useful point of entry into Green's philosophical thought is to be found in his early essay "Popular Philosophy in its Relation to Life," originally published in 1868. The "popular philosophy" of the title was that professed by the advanced thinkers of the time, whom Green explicitly compared with the ancient Sophists. Like the Sophists they were superficially clear and rhetorically persuasive; again like them, they owed their apparent success to a refusal to examine their basic notions. Yet these notions, when applied in the concrete, turned out to be wholly inadequate; they could not successfully be brought to bear on life, as understood in art or religion or moral practice. In "Popular Philosophy" Green set himself to demonstrate this conclusion only in the case of ethics, reviewing for this purpose the doctrines of Joseph Butler, David Hume, and Jean-Jacques Rousseau, but it was obvious that he had wider considerations in mind. When he spoke of the need for an "adequate theory," of which the philosophy of G. W. F. Hegel might offer a foretaste, he was implying that the philosophers he was considering were wrong not only in their ethical distinctions but also in their whole method and metaphysical outlook. Following John Locke, they assumed that truth could be arrived at by simple introspection; they had no notion of the difference between image and concept and hence no tenable theory of thinking. The statement of their theories presupposed a continuing self-identical subject, but the theories themselves had no room for any such thing. Nor, in practice, had they anything like an adequate grasp of the workings of the human mind, which they looked on as an isolated automaton mechanically pursuing pleasure and seeking to avoid pain, instead of as an active agent whose interest and welfare were intimately bound up with those of others.

The corrective to popular philosophy, Green said at this stage, was to be found in "the deeper views of life which the contemplative poets originated" and in the notions of "evangelical religion," as well as in some of the better insights of Rousseau. It was not "'within his own breast'" that Wordsworth had looked to "read what he was," but to "the open scroll of the world, of the world, however, as written within and without by a self-conscious and self-determining spirit" (*Works,* Vol. III, p. 119). Similarly, the evangelical insistence on the sense of sin showed the superficiality of the moral philosophy of the Enlightenment, which could represent vice as an object of regret or distaste but never as an object of abomination. Much of Green's mature philosophy becomes intelligible if we bear these remarks in mind. He took neither poetry nor religion to be a substitute for philosophy, but nonetheless he felt deeply that both had important lessons to teach the philosopher.

CRITIQUES OF HUME AND NATURALISM

The more academic criticisms sketched in Green's early essay were elaborated in the introduction to his edition of Hume's works (1874). Green's view was that Hume was a major philosopher who had taken Locke's ill-thought-out assumptions to their logical conclusion and, in so doing, had revealed their absurdity. Hume's first principle was that nothing is real save feeling; Green attacked this view with the argument that to say that something is real is to relate it to other things and that relations are not given in feeling but are the work of the mind. Hume's attempt to ground "philosophical" in "natural" relations, that is, in what is given to sense, was a failure. So was his theory of the self as a succession of perceptions, for relating clearly demands an unchanging subject that relates. The argument of this passage was carried further in the first book of the *Prolegomena to Ethics,* where Green claimed that not only our consciousness of nature but also nature itself presuppose an "eternal," self-identical, and self-differentiating subject which is the source not only of the connections in thought but also of its material. A halfway position, such as Immanuel Kant had taken, was not intellectually defensible: The thing-in-itself and its empiricist counterpart, the sheerly given, remain unintelligible on this type of view.

Green's object in metaphysics was plainly to state an alternative to materialism, which struck him as both

insidiously attractive and intellectually preposterous. Much the same ambition informs his writings on ethics, in which field he again saw himself as developing an antinaturalist position. In his critique of moral sense theories, which forms part of his general criticism of Hume, he represented the supporters of those views as one and all hedonists, on the ground that they made the passions the only spring of action and thought of reason as practically inert. Hence, his own first aim in the *Prolegomena* was to establish that human actions spring from *motives* and to show that motives are not "natural phenomena." He defined a motive (Sec. 87) as "an idea of an end, which a self-conscious subject presents to itself, and which it strives and tends to realise." The vital point here was the connection between motives and a continuing subject consciously pursuing good; human action, for Green, was entirely different from animal behavior, for although much of it had animal impulses as its basis, these impulses were transformed in being brought into consciousness and thought of in relation to long-term aims. As for the good with which action is concerned, Green said (ibid., Sec. 92) that "anything conceived in such a way that the agent acts for the sake of it, must be conceived as *his own* good, though he may conceive it as his own good only on account of his interest in others, and in spite of any amount of suffering on his own part incidental in its attainment." But in practice he had little to say about the connection of good with the satisfaction of the agent: The moral ideal must be realized in persons, but one person's claims to moral self-expression were as good as those of another, and moral progress came about with the realization that more and more persons and types of person were entitled to have their claims considered. Green made much use of the phrase "the common good" in speaking of the ultimate aim of moral action, but his alternative description of the end as the attainment of "human perfection" is in some ways more appropriate, provided it is added that he wanted to see human perfection realized without distinction of persons.

POLITICAL PHILOSOPHY

In ethics Green had clearly learned a lot from Hegel, although his general outlook remained more Kantian than Hegelian—both in theory and in practice he must be counted as a liberal moralist. His political philosophy also is in the liberal tradition, despite its rejection of such elements of older liberal political theory as the doctrine of the contract. The state, according to Green, is the product of will, not of force, insofar as the system of rights and duties it operates rests on a moral as opposed to a merely natural basis. Green was as emphatic in his political as in his ethical theory that rights cannot be created out of nothing, in the way Thomas Hobbes and Benedict Spinoza supposed. But although he thus saw the state as, in a sense, a moral organism, Green had no inclination to endow it with positive moral authority. The state might sometimes have to inhibit the freedom of particular men to enable others to be free at all, but the end of political action could only be to put citizens in a position to lead the good life. The liberalism he favored was thus in the end a negative liberalism, concerned with creating the minimum conditions in which people could exercise moral choice and, for the rest, leaving matters to their voluntary efforts.

Compared with that of his younger contemporary F. H. Bradley, Green's literary style was flat and uninteresting. The moral earnestness that is apparent in so much of his writing also has had much to do with its neglect by more recent philosophers. But however earnest he was, he was at the least estimate an influential thinker; to describe him, as C. D. Broad did, as "thoroughly second-rate" is to forget the extent to which his articulation of problems is still accepted, for example in political philosophy. Nor are his solutions entirely without interest, if only we can divest them of the stiff Victorian garments in which he chose to clothe them.

See also Bradley, Francis Herbert; Broad, Charlie Dunbar; Butler, Joseph; Empiricism; Enlightenment; Ethics, History of; General Will, The; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Hume, David; Idealism; Kant, Immanuel; Locke, John; Metaphysics; Naturalism; Rousseau, Jean-Jacques; Self; Sidgwick, Henry; Sophists; Spinoza, Benedict (Baruch) de.

Bibliography

The Works of Thomas Hill Green, edited by R. L. Nettleship, 3 vols. (London: Longmans, Green, 1885–1888), contains all Green's surviving writings except his introductions to Hume (in Hume's *Works,* edited by T. H. Green and T. H. Grose, London, 1874) and the *Prolegomena to Ethics,* edited by A. C. Bradley (Oxford: Clarendon Press, 1883).

For literature on Green, see J. Pucelle, *La nature et l'esprit dans la philosophie de T. H. Green,* 2 vols. (Louvain: Nauwelaerts, 1961–1965), which is the fullest commentary; see also Pucelle's *L'idéalisme en Angleterre* (Neuchâtel: La Baconnière, 1955).

For Green's ethics, see W. D. Lamont, *Introduction to Green's Moral Philosophy* (London: Allen and Unwin, 1934). On his political views, compare A. J. M. Milne, *The Social Philosophy of English Idealism* (London: Allen and Unwin, 1962), and M. Richter, *The Politics of Conscience: T. H. Green and His Times* (London: Weidenfeld and Nicolson, 1964),

which also contains biographical details supplementing Nettleship's memoir in *Works,* Vol. III. Mrs. Humphry Ward's novel *Robert Elsmere* (1888) is dedicated to Green's memory, and one of its characters is said to portray him.

W. H. Walsh (1967)

GREGORY OF NAZIANZUS

(329/330–c. 390)

Gregory of Nazianzus, the poet, theologian, and bishop, was born at Arianzum in Cappadocia. Although his parents were Christians, he enjoyed an excellent classical education at Caesarea in Palestine and at the universities of Alexandria and Athens. He was from his earliest years a close friend of the other two Cappadocians, Gregory of Nyssa and Basil the Great. Baptized at about the age of thirty, he was ordained by his father, the local bishop—apparently against his will—and immediately fled. After his penitent return, Basil appointed him bishop of the isolated town of Sasima. However, Gregory refused to go and remained with his father at Nazianzus, staying on after his father's death in 374. After a period of monastic living he was approved as bishop of Constantinople under Emperor Theodosius, but distrust of his own administrative ability once again forced him to resign after a year. After a few years at Nazianzus, he finally retired to his estate at Arianzum and devoted his last years to writing; it was here, between 384 and 390, that he wrote his greatest poems.

Adequate study of Gregory is still hampered by the lack of a full critical edition of his works. The bulk is poetic (more than 16,000 lines). There are also 44 orations, including the important dogmatic ones (numbers 27–31, delivered in 380), and 244 authentic letters.

Gregory Nazianzen is the most literate, self-conscious stylist of the three Cappadocian Fathers, although perhaps not as profound as Gregory of Nyssa nor so immersed in ecclesiastical affairs as Basil the Great. Although he once compared philosophy to "the plagues of Egypt," his poetry shows the wide influences of all the Greek schools, and especially of the Stoic-Cynic. In morals Gregory reflects a sharply critical view of contemporary worldliness and sensuality; and his introspective poetry (especially the autobiographical "De Vita Sua") marks a new era in Christian self-awareness and is comparable to Augustine's *Confessions.* The bulk of his verse, however, is coldly classical and heavily didactic.

Gregory was fully aware (see *Oration* 20.17) of the role of speculation in theology. He contributed to Trinitarian theology by clearly defining the relations and properties of the three Persons of the Trinity. In Christology he insisted on the two distinct natures in Christ bound by a "union according to essence," copresent to each other by "circumincession"—a term later applied to the Persons of the Trinity. In developing traditional dogma, Gregory's discussion is sometimes sharper than either Basil's or Gregory of Nyssa's, although he is vague on the doctrines of hell and original sin.

See also Augustine, St.; Cynics; Gregory of Nyssa; Stoicism.

Bibliography

WORKS BY GREGORY

Works. Edited by C. Clemencet and A. B. Caillau. 2 vols. Paris, 1778 and 1840. Reprinted in *Patrologia Graeca,* edited by J. P. Migne. Paris, 1857–1866. Vols. 35–38.

Select Library of Nicene and Post-Nicene Christian Fathers. Series II. New York: Christian Literature, 1887–1892; Oxford, 1890–1900. Vol. VII, pp. 185–498. A translation of selected orations and letters by J. E. Swallow, with an introduction by C. G. Browne.

Five Theological Orations. Edited by A. J. Mason. Cambridge, U.K.: Cambridge University Press, 1899. Critical edition.

WORKS ON GREGORY

Campenhausen, Hans von. *The Fathers of the Greek Church.* Translated by S. Godman, 95–106. New York: Pantheon, 1959.

Pellegrino, M. *La poesia di S. Gregorio Nazianzeno.* Milan: Società editrice "Vita e pensiero," 1932.

Plagnieux, J. *S. Sainte Gregoire de Nazianze, Théologien.* Paris: Éditions franciscaines, 1951.

Quasten, Johannes. *Patrology,* Vol. III: *The Golden Age of Greek Patristic Literature,* 236–254. Westminster, MD: Newman Press, 1961.

Werhan, H. M., ed. *Lexikon für Theologie und Kirke.* 2nd ed., Vol. IV, 1210ff. Freiburg, 1960.

Herbert Musurillo, S.J. (1967)

GREGORY OF NYSSA

(c. 330–c. 394)

Gregory of Nyssa, the Christian theologian and Father of the Eastern church, was born in Cappadocia. Resisting the invitation of his brother, Basil the Great, to join his monastic community at Annesis, Gregory married and became a teacher of rhetoric. In 372 Basil, bishop of Caesarea, had Gregory appointed bishop of Nyssa; but Gregory was deposed in 374 by a local synod dominated by the Emperor Valens and the Arian party. Restored to his see in 377, Gregory began to grow closer to Basil's monas-

tic and theological ideals. After Basil's death in 379, Gregory engaged more and more in writing and in the vigorous administration of his diocese; he was an important figure at the councils convoked at Constantinople in 381, 383, and, just before his death, in 394. An ardent defender of the orthodox Trinitarian doctrine of Nicaea against the Arians and semi-Arians, he was also popular in court circles at Constantinople. Toward the end of his life, when his influence began to wane, he devoted himself to the deepening of the traditional Christian heritage of mystical theology; during this period, from about 390 until his death, he composed some of his most profound works, the *Commentary on the Song of Songs* and the *Life of Moses*, which represent the culmination of the process inaugurated in his earliest work, the *Treatise on Virginity* (c. 370).

Gregory's originality lay chiefly in the depth and mystical awareness he brought to the problem of human's knowledge of the Transcendent. Many of his works, such as the *Life of Moses*, can be understood on three levels: Moses represents the life of the true believer, the Christian philosopher, and the mystic attempting to find God in the universe. In his exposition of the Trinity and his discussion of God's nature, Gregory penetrated deeper than any other Eastern Father. The core of his theology is the historical perfection of humankind through the restoration of the divine image, regained by the Atonement and communicated through the church. In his doctrine of the *Apokatastasis*—the restoration of all people, even the damned, to the vision of God at the end of time—Gregory reveals his loyalty to Origen as well as his own attempt to create a harmonious structure of salvation history. Throughout his work we see the development of the doctrine of the spiritual or mystical senses (implying a direct intuition of God's presence) and an analysis of ecstasy that prepared the way for Dionysius the Areopagite, Maximus the Confessor, and later Byzantine mysticism.

In his epistemology Gregory is derivatively Neoplatonic, and his allegorical exegesis reflects the anthropology of Origen and Philo of Alexandria as well as the eclectic philosophy of Hellenistic Asia Minor. But Gregory never slavishly followed any master, and scholars like H. F. Cherniss go too far when they suggest that Gregory's theology was merely a question of giving Christian names to Plato's doctrines. Rather, the opposite view of Jean Daniélou and others seems closer to the truth: Gregory's theology represents a subtle transformation of Neoplatonism into authentic Christianity, whereby the intuitive vision and ethical achievement of the Christian mystic (in Daniélou's terminology, *epektasis*) was the culmination of the pagan philosopher's quest.

See also Mysticism, History of; Neoplatonism.

Bibliography

TEXTS AND TRANSLATIONS

Aubineau, M. *De Virginitate. Sources Chrétiennes* 119, 1966.

Forbes, G. H. *Apol. In Hexameron*. Burntisland, 1855–1861.

Heil, G. et al. *Sermones*. Leiden, 1967.

Jaeger, W. *Contra Eunomium*. Berlin, 1921; Leiden: E. J. Brill, 1960.

Jaeger, W., et al. *Opera Ascetica*. Leiden: E. J. Brill, 1952.

Krabinger, J. G. *De Anima et Resurrectione*. Leipzig, 1837.

Krabinger, J. G. *De Oratione Dominica*. Landshut, 1840.

Langbeck, H. *In Canticum Conticorum*. Leiden, 1960.

Lendle, O. *Encomium in Sanctum Stephanum Protomartyrem*. Leiden: E. J. Brill, 1968.

McDonough, J., and P. Alexander. *In Inscriptiones Psalmorum in Sextum Psalmum in Ecclesiasten Homiliae*. Leiden: E. J. Brill, 1962.

Migne, J. P. *Patrologia Graeca* 44–46. Paris, 1857–1866.

Mueller, F. *Opera Dogmatica Minora*. Leiden, 1958.

Mursillo, H. *De Vita Moysis*. Leiden, 1962; E. J. Brill, 1991.

Pasquali, G. *Epistulae*. Berlin, 1925.

van Heck, A. *De Pauperibus Amandis Orationes Duo*. Leiden: E. J. Brill, 1964.

English translation of selected texts in *Nicene and Post-Nicene Christian Fathers*. Series 2, vol. 5. Oxford, 1890.

STUDIES

Cherniss, Harold. *The Platonism of Gregory of Nyssa*. New York: B. Franklin; Berkeley: University of California Press, 1930.

Coakley, S., ed. *Rethinking Gregory of Nyssa*. Oxford; Malden, MA: Blackwell, 2003.

Coffigny, Daniel. *Grégoire de Nysse*. Paris, 1993.

Daniélou, Jean. *L'être et le temps chez Grégoire de Nysse*. Leiden: E. J. Brill, 1970.

Diekamp, F. *Die Gotteslehre des heiligen Gregor von Nyssa*. Münster, 1896.

Drobner, Hubertus, and Christoph Klock. *Studien zu Gregor von Nyssa und der christlichen Spätantike*. Leiden, 1990.

González, G. *La formula mia ousia treis hupostaseis en S. Gregorio de Nisa* (Analecta Gregoriana 21, 1939).

Jaeger, W. *Gregor von Nyssas Lehre vom heiligen Geist*. Leiden: E. J. Brill, 1966.

Laird, M. S. *Gregory of Nyssa and the Grasp of Faith: Union, Knowledge, and Divine Presence*. Oxford: Oxford University Press, 2004.

Meredith, Anthony. *Gregory of Nyssa*. London: Routledge, 1999.

Pellegrino, M. "Il platonismo di San Gregorio Nisseno nel dialogo 'Intorno all'Anima e alla Risurrezione." *Rivista di Filosofia Neo-Scolastica* 30 (1938) 437–474.

Turcescu, Lucien. *Gregory of Nyssa and the Concept of Divine Persons*. Oxford, 2005.

Vannier, Marie-Anne. *Les Cappadociens*. Montrouge, 1997.

Völker, W. *Gregor von Nyssa als Mystiker*. Wiesbaden: F. Steiner, 1955.

von Balthasar, H. U. *Présence et pensée: Essai sur la philosophie religieuse Grégoire de Nysse*. Paris: G. Beauchesne et ses fils, 1942.

Zachhuber, Johannes. *Human Nature in Gregory of Nyssa: Philosophical Background and Theological Significance*. Leiden: E. J. Brill, 2000.

Zemp, Paul. *Die Grundlagen heilsgeschichtlichen Denkens bei Gregor von Nyssa*. München: M. Hueber, 1970.

Herbert Musurillo, S.J. (1967)
Bibliography updated by Scott Carson (2005)

GREGORY OF RIMINI

(c. 1300–1358)

Gregory of Rimini, a member of the Augustinian friars and one of the foremost thinkers of the fourteenth century, was born in Italy and died in Vienna, where he spent the last eighteen months of his life as general of the Augustinian order. A large part of his active career was spent at Paris, where he studied from 1323 to 1329. After teaching in Italy, he returned to Paris in 1341 and remained there for ten years. During this second sojourn in Paris he wrote his main work, a *Commentary on the Sentences.* None of the other writings ascribed to him, ranging from biblical commentaries to a treatise on the remission and intensification of forms, has survived.

Gregory's system was a reassertion of St. Augustine's teachings in fourteenth-century terms. He shared the contemporary awareness of the radical contingency of the created order and the unbridgeable gulf between God and his creatures that it entailed. He thereby followed William of Ockham and his confreres in rigorously confining natural knowledge to what could be verified and in excluding theological truths and evidence for God's existence from ratiocination. On the one hand, God was sovereignly free and man had no means of knowing what He might do; on the other, the knowledge accessible to man dealt only with contingencies and was ever liable to be superseded if God so willed. In consequence, there was no guarantee that the world was not infinite or eternal or that there was only one world; and even if its finiteness was accepted, God could still transform it. Gregory parted from the Ockhamists, however, in his refusal to allow this distinction between natural experience and God's will to undermine the traditional certainties. Even if natural knowledge was confined to practical experience, there still remained an inner realm of knowledge that was the source of all necessary truths and nonsensory principles. Similarly, although God was unconstrained and his ways inscrutable, he still acted in accordance with his perfections.

Thus Gregory rebutted the Ockhamist assertions that God could cause a man to sin, or mislead him, or command a man to hate him: God's freedom could not violate his own nature as revealed in the Scriptures. Although Gregory subscribed to the current distinction between God's ordained power (*potentia ordinata*) and his absolute power (*potentia absoluta*), by which he could do anything without qualification, he never—unlike the Ockhamists—employed the latter to override dogma. Gregory accordingly adhered to the accepted dogmatic tenets whether they concerned God's foreknowledge, man's fallen state, or the theological virtues. Only in the case of the physical world did he acknowledge the possibility, both on epistemological and theological grounds, that the world could be other than it was: that it might be infinite or eternal. Gregory joined in the current reversion to the earlier view that theology was *sapientia* (wisdom) rather than *scientia* (scientific knowledge). It was distinguished by its inaccessibility to the nonbeliever, and faith, far from being communicable, was the barrier that divided the Christian from the infidel. In this, as in other ways, Gregory shared in the changed outlook of the time, while remaining true to the tradition of St. Augustine.

See also Augustine, St.; Augustinianism; Ockhamism; William of Ockham.

Bibliography

WORKS BY GREGORY

Gregorius Ariminensis. *Lectura super primum et secundum Sententiarum.* Vols. I–VI, edited by A. D. Trapp et al. Berlin and New York, 1979–1984.

WORKS ON GREGORY

Brown, S. "Walter Burleigh, Peter Aureoli, and Gregory of Rimini." In *Routledge History of Philosophy,* Vol. III: *Medieval Philosophy,* edited by J. Marenbon, 368–385. London and New York, 1998.

Courtenay, W. J. "John of Mirecourt and Gregory of Rimini on Whether God Can Undo the Past." *Recherches de Théologie Ancienne et Médiévale* 39 (1972): 224–256 and 40 (1973): 147–174.

Courtenay, W. J. *Schools and Scholars in Fourteenth-Century England.* Princeton, NJ, 1987.

Cross, R. "Infinity, Continuity, and Composition: The Contribution of Gregory of Rimini." *Medieval Philosophy and Theology*7 (1998): 89–110.

Smith, K. "Ockham's Influence on Gregory of Rimini's Natural Philosophy." *Dialexeis 1996–1997,* 107–144. Leukosia, Cyprus, 1999.

Zupko, J. "How It Played in the *Rue de Fouarre*: The Reception of Adam Wodeham's Theory of the '*compexe significabile*' in

the Arts Faculty at Paris in the Mid-fourteenth Century." *Franciscan Studies* 54 (1994): 211–225.

Zupko, J. "Nominalism Meets Indivisibilism." *Medieval Philosophy and Theology* 3 (1993): 158–185.

Gordon Leff (1967)
Bibliography updated by Stephen F. Brown (2005)

GRICE, HERBERT PAUL
(1913–1988)

Herbert Paul Grice was born and educated in England. He taught at St. John's College Oxford until 1968, when he moved to the University of California–Berkeley. He taught there until his death. He published little until near the end of his life, but had a great influence through students and the wide circulation of unpublished manuscripts. His earliest work dealt with perception, but he subsequently moved to problems in language, ethics, and metaphysics. A concern with reason and rationality is a subtle thread which unites these investigations. His historical idols were Aristotle and Kant.

One early topic was a defense of the causal theory of perception. This defense required separating the scientific or specialist's part of the task of analyzing perception from that of the philosopher. This distinction relies on an underlying notion of analysis closely related to the analytic–synthetic distinction for which Grice and Strawson provided a brief spirited defense. Three subsequent papers represent intricate attempts to define meaning using only common sense psychological concepts such as intention, belief, and desire. If this program is successful it would provide a more elaborate defense of the analytic synthetic distinction.

Grice's best known and most influential contribution is the concept of a conversational implicature. A conversational implicature of an assertion is something that is conveyed to a thoughtful listener by the mode of expression rather than by the meanings of the words. These arise from the fact that conversation is normally governed by principles including cooperation, truthfulness, and informativeness, and that both parties are aware of these. The two best known applications of this concept are to perception and logic. Grice was concerned to provide an account of sense data discourse in terms of how things seemed to the perceiver. A common objection to this is that it is odd to say in a normal case of the perception of a table that it *seems* to the subject that a table is present. Grice's concept of conversational implicature can be invoked to explain the oddity as a result of the fact that a stronger statement can be made, thus leaving room for the *seems* statement to be true.

The concept of conversational implication has been widely deployed in linguistics and artificial intelligence as well as in philosophy and is a continuing topic of research and debate. One major focus of discussion is the adequacy of the account when applied to quantitative statements, such as "John has two children." It is controversial whether this statement means that John has exactly two children, or whether it means that John has at least two children. In the latter case, interpreting an assertion of the statement as conveying that John has exactly two children is a matter of conversational implication.

Grice also scouts the possibility of defending the claim that the logician's material conditional is an adequate representation of the indicative conditional of English by explaining the apparent divergence as a matter of conversational implicatures. If one knows the truth values of P and Q then one can make a more informative statement than $P \supset Q$, so the only conversationally appropriate use of $P \supset Q$ is when the speaker does not know the truth of either component, but only that they are so connected that the truth of P guarantees the truth of Q. The appropriate conversational use of $P \supset Q$ requires a connection that is not part of the truth condition of the compound. The main objection to Grice's approach concedes that his account squares fairly well with the assertion of conditionals, but points out that it does nothing to ameliorate the implausible fact that on the material conditional account, to deny "if P then Q" implies both P and ~Q.

Part of the definition of a conversational implicature requires that the hearer should be able to reason out the intentions of the speaker and in conjunction with the conversational principles to discern the implicit message. This places an important role on reasoning, especially inasmuch as in typical cases the reasoning is not conscious in the hearer.

Grice devotes considerable energy to investigating rationality, reasoning, and reasons. Grice emphasizes that reasoning is typically directed to the goal of producing reasons relevant to some end in view. This intentional activity involves the ability to make reason-preserving transitions. Grice defines "reason preserving" analogous to the concept of "truth preserving" in deductive logic. A transition is reason-preserving just in case, for if one has reasons for the initial set of thoughts, beliefs, actions or intentions, then one does for the subsequent set as well.

Grice uses this general account of reasoning to investigate moral reasoning and moral reasons. He emphasizes the connections between reasons, actions, and freedom. Strong rational evaluation—which Grice sees as essential to freedom—involves the rational evaluation and selection of ends, including ultimate ends.

How do people choose ultimate ends? Grice answers that people should choose ends that have unrelativized value. Grice grants that the concept of unrelativized value requires defense. Typically, things have value only relative to ends and beneficiaries. A concern for the focus of relativization gives the value-concept a *bite* on a person; it ensures that the value-concept carries weight for that person. So how are people to understand unrelativized value?

Grice turns to final causation for a special kind of value. A tiger is a good tiger to the degree that it realizes the final end of tigers. Grice defines a good person as one who has, as part of their essential nature, an autonomous finality consisting in the exercise of rationality. Grice's philosophical psychology supports this conception of persons as end-setters. Freedom intimately involves the ability to adopt and eliminate ends. One does not (ideally) arbitrarily select and conform to ends; one does so for reasons. This makes being an end-setter an instance of unrelativized value; for to take a consideration as an ultimate justification of action is to see it as having value. Grice defines unrelativized value "in Aristotelian style [as] whatever would seem to possess such value in the eyes of a duly accredited judge; and a duly accredited judge might be identifiable as a good person operating in conditions of freedom." (*Aspects of Reason* 2001, p. 119) Of course, we are still talking about what is of value for and to persons. But the point was not to avoid this *relativization*; the point was to avoid relativization to this or that kind of person.

See also Aristotle; Conditionals; Kant, Immanuel; Perception; Rationality; Reason.

Bibliography

WORKS BY GRICE

The Conception of Value. New York: Oxford University Press, 1991.A posthumous publication of the John Locke Lectures, delivered 1979.

Studies in the Way of Words. Cambridge, MA: Harvard University Press, 1989. A collection including most of the important works published during his lifetime.

Aspects of Reason, edited by Richard Warner. Oxford: Oxford University Press, 2001. A posthumously published book exploring the nature of reasons and reasoning.

Grice, H. P., and P. F. Strawson. "In Defense of a Dogma." *Philosophical Review* Vol. 65 (1956) 141–58. A defense of the analytic-synthetic distinction, widely reprinted and discussed.

WORKS ABOUT GRICE

Avramides, Anita. *Meaning and Mind: An Examination of a Gricean Account of Language*. Cambridge, MA: MIT Press, 1989.

Cosenza, Giovanna, ed. *Paul Grice's Heritage*. Turnhout, Belgium: Brepols, 2001.

Davis, Wayne A. *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge, U.K.: Cambridge University Press, 1998.

Grandy, Richard E., and Richard Warner. *Philosophical Grounds of Rationality: Intentions, Categories, Ends*. Oxford: Oxford University Press, 1986. A festschrift celebrating Grice's work, with a lengthy editorial introduction and a response by Grice.

Richard E. Grandy (1996, 2005)
Richard Warner (1996, 2005)

GROSSETESTE, PSEUDO-

See *Pseudo-Grosseteste*

GROSSETESTE, ROBERT

(c. 1168–1253)

Robert Grosseteste was one of the most influential Englishmen of his day—initiator of the English scientific tradition, one of the first chancellors of Oxford University, a famous teacher and commentator on the newly discovered works of Aristotle, an important translator from the Greek, friend to the mendicant orders, first lecturer to the Oxford Franciscans, and zealous bishop of England's largest diocese. However, his life is imperfectly known and much of his work remains unpublished. He was born of humble parents in the county of Suffolk between 1168 and 1175 and by 1190 had become *magister in artibus* at either Oxford or Paris. Sometime between 1190 and 1198 he was a member of the household of William de Vere, bishop of Hereford, and may have taught in the Hereford schools. After the bishop's death in 1198, Grosseteste was a member of the arts faculty at Oxford or possibly at Cambridge. He probably studied theology at Paris during the *suspendium clericorum,* 1209–1214. At some time between 1214 and 1221 he became chancellor of Oxford University. In 1229 he became the first lecturer to the Oxford Franciscans, leaving this post only on his elevation to the see of Lincoln in 1235. He was bishop of Lincoln until his death eighteen years later.

Grosseteste lived at a crucial period in intellectual history: The scientific and philosophical writings of the Muslims were just becoming known in Latin Europe and the works of the Hellenistic writers and the recently rediscovered works of Aristotle were being translated, disseminated, and lectured upon. As teacher, commentator, and translator, he took an active part in this movement. Basically Augustinian in outlook and relying heavily on the standard authors, he was nevertheless deeply influenced by Muslim learning, especially Avicenna and the astronomers, by the Jew Solomon ben Judah ibn Gabirol (Avicebrón), and by the newly found Aristotelian works. He never wrote a comprehensive philosophical work or devised a system, but he developed many characteristic views that have had a profound influence on the later development of both philosophy and science. The most important of his many philosophical works are *De Luce* (Light), *De Motu Corporali et Luce* (Corporal Motion and Light), *Hexameron,* and commentaries on Aristotle's *Posterior Analytics* and *Physics.*

LIGHT METAPHYSICS

Basic to Grosseteste's view of the universe is his metaphysics of light. He held that in the beginning God created the first corporeal form (*lux*), which had the property of instantaneously multiplying itself infinitely in every direction, and simple matter, an unextended substance. The original point of light was joined to unextended matter (since matter and form never exist separately) and in its expansion drew matter out into spatial dimensions. The resulting universe was a sphere extremely rare at the periphery but dense and opaque near the center. It was finite because a simple substance multiplied an infinite number of times would result in a finite quantity, and the matter of the periphery (the firmament) was completely actualized and capable of no further change.

When this perfect first body, containing only first matter and first form, had been created, it diffused its reflected light (*lumen*) back to the center, where the *lumen* gathered together the mass existing below the first body, again rarefying the outermost parts and making the center more dense. The second sphere was thus formed, as were, by a similar process, all thirteen spheres, including the four elements. On the outside of our universe, matter is completely actualized and capable of no further change, while at the center the degree of actualization is less and matter remains susceptible of taking on a variety of forms. From first form (light) every subsequent form is generated, both substantial and accidental, and every privation derives from the privation of light.

Since all things have in common first form and first matter, they are, in a sense, one. But each thing includes a hierarchy of form superadded to the original form of corporeity, making it the individual thing it is. Most of Grosseteste's other views were either derived from or imply his light metaphysics. He considered light the cause of local motion, the means by which the soul operates on the body (he denied that the soul is the form or perfection of the body), and the principle of intelligibility in the created universe.

THEORIES OF KNOWLEDGE

Grosseteste had two distinct theories of knowledge. The first, in the Augustinian tradition and strongly influenced by Avicenna, held that men may acquire knowledge by virtue of the intellect alone, without recourse to sense. The second held that certain knowledge may also be gained through sense perception. Although sense turns toward matter and is therefore unstable, imperfect, and subject to imaginative embellishments, it also follows reason, even though confusedly, and does not obscure the species it provides. Reason, which understands the principles of nature in a single manner, either corrects or completes whatever was lacking in the senses.

Both these ways of knowing involve another of Grosseteste's key concepts, the purgation of the mind. It is not until the desires of one's mind (*affectus mentis*) are purged of error that the gaze of one's mind (*aspectus mentis*) can be raised to the eternal and true and can overcome the delusions caused by corporeal phantasms. "Many men," Grosseteste said, "can prove by sure reasons that the Intelligences exist and that God exists but they do not understand the Divine Essence or the non-corporeity of the Intelligences.… Aristotle and others, who firmly knew by discursive reasoning that eternity was simple but saw it under the phantasms of temporal extension, have affirmed many improper things such as the perpetuity of time and motion and consequently the eternity of the world." In this quotation from the *Commentary on the Physics,* we see Grosseteste at once as one of the foremost critics of the dangers latent in the works of Aristotle and yet also among the leaders in introducing Aristotle's natural philosophy into western Europe.

INFINITY

One of Grosseteste's most original and influential teachings concerns infinite aggregates. He believed that "one infinite number can be related in any proportion,

numeral or non-numeral, to another infinite number." To God, infinite numbers are finite, and he determines the primary cubit (and every other measure) by a certain infinite number and a half-cubit by another infinite number half (to him) that of the cubit, and so on. But such a manner of measuring is possible only to one to whom the infinite is finite. Being finite, we must necessarily adopt a different manner of measuring, that is, by commensurable magnitudes as accidents of matter.

SCIENTIFIC METHOD

It is as a scientist and innovator in scientific method, however, that Grosseteste attracted the interest of the twentieth century. In his most important scientific writings he progressively developed a characteristic method of investigating nature that employed analysis (*resolutio*) and synthesis (*compositio*) in physical inquiries, first breaking down a problem into its simplest parts, then framing a hypothesis that would show how these elements are to be combined in order to produce the phenomenon under investigation. He also held that an experimental universal of provisional truth might be obtained by observing that a given effect always results from a particular cause, if one controlled his observation by eliminating any other possible cause of the effect.

In addition to this framework, Grosseteste used experiments as an integral part of his investigation: as aids in accomplishing his analysis, as suggestions in framing his explanatory hypothesis, and most important, as tests of the truth or falsity of a hypothesis. He also employed mathematics in his researches, holding that since light is the cause of local motion and the means by which superior bodies act on inferior ones, and that since light behaves according to geometric rules, therefore all local motion can be described mathematically. He denied, however, that mathematical entities have any objective being and insisted that they are simply abstractions from physical bodies and exist only in the minds of mathematicians.

Another of his basic principles was that of the subordination of sciences. A superior science, he said, may provide the cause for which the inferior science provides the effect. In the study of heavenly bodies, for instance, the sciences of mathematics, astronomy, and physics are concerned. Mathematicians abstract magnitudes from motion and matter and demonstrate the accidents per se with respect to magnitudes. Physicists, on the other hand, demonstrate the figured magnitudes in the sense that they belong to physical bodies. Astronomers have much in common with physicists, but whereas they might both be studying the same body—for instance, the moon—the physicist demonstrates that the predicate belongs to the subject by nature, while the astronomer does not care whether it belongs to it by nature or not.

Grosseteste wrote four works on astronomy. His work *De Sphaera* (On the Sphere) is a theoretical treatise. The other three works are primarily concerned with reforming the Julian calendar, which was nearly four days in error at that time. Using the works of Ptolemy, al-Battani, and Ibn Thebit, he worked out a program for calendar reform that continued to find supporters until it was largely incorporated into the Gregorian reform of 1582.

Grosseteste was in many ways the hinge between the early and late Middle Ages. He had at his disposal the standard late Roman authors and the recently introduced Greek and Arabic sources. His powerful, resourceful, and disciplined mind assimilated and transformed this material. He left many loose ends and sometimes failed to think through his positions; and in his scientific works, despite his methodological triumphs, he was not a notable experimenter. Still, so powerful was his thought that he influenced an uninterrupted succession of philosophers and scientists throughout Europe for 300 years after his death.

See also Aristotle; Augustinianism; Avicenna; Ibn Gabirol, Solomon ben Judah; Scientific Method; Sensa.

Bibliography

BIBLIOGRAPHIES

Gieben, Servus. "Bibliographia universa Roberti Grosseteste ab anno 1473 ad annum 1969." *Collectanea Franciscana* 39 (1962): 362–418.

Gieben, Servus. "Robertus Grosseteste: Bibliographia 1970–1991." In *Robert Grosseteste: New Perspectives on His Thought and Scholarship,* edited by James McEvoy. Turnhout, Belgium: Brepols, 1995.

Thomson, S. Harrison. *The Writings of Robert Grosseteste, Bishop of Lincoln, 1235– 1253.* Cambridge, U.K.: Cambridge University Press, 1940.

EDITIONS

Baur, Ludwig, ed. *Die philosophischen Werke des Robert Grosseteste, Bischofs von Lincoln.* Münster: Aschendorff, 1912. The standard edition of Grosseteste's philosophical works.

Dales, Richard C., ed. *Roberti Grosseteste episcopi Lincolniensis commentarius in viii libros Physicorum Aristotelis.* Boulder, CO: University of Colorado Press, 1963.

Dales, Richard C., ed. "Robert Grosseteste's Treatise *De finitate motus et temporis.*" *Traditio* 19 (1963): 245–266.

Dales, Richard C., and Servus Gieben, eds. *Hexaëmeron.* Oxford: Oxford University Press, 1982.

Lewis, Neil, ed. "The First Recension of Robert Grosseteste's *De libero arbitrio.*" *Mediaeval Studies* 53 (1991): 1–88.

Panti, Cecilia, ed. *Moti, virtù e motori celesti nella cosmologia di Roberto Grossatesta.* Florence: Edizioni del Galluzzo, 2001. New editions and study of *De sphaera, De cometis, De motu supercaelestium.*

Rossi, Pietro, ed. *Commentarius in Posteriorum Analyticorum libros.* Florence: Leo S. Olschki, 1981.

TRANSLATIONS

Martin, C. F. J., tr. *On the Six Days of Creation: A Translation of the Hexaëmeron.* Oxford: Oxford University Press, 1996.

McKeon, Richard, tr. *Selections from Medieval Philosophers.* Vol. 1. New York: Charles Scribner's Sons, 1929. Includes translations of *De veritate, De veritate propositionis,* and *De scientia Dei.*

Riedl, Clare C., tr. *Robert Grosseteste on Light.* Milwaukee, WI: Marquette University Press, 1942.

Rossi, Pietro, tr. *Roberto Grossatesta: Metafisica della luce: Opuscoli filosofici e scientifici.* Milan: Rusconi, 1986.

STUDIES

Baur, Ludwig. *Die Philosophie des Robert Grosseteste.* Münster: Aschendorff, 1917.

Callus, Daniel A., ed. *Robert Grosseteste, Scholar and Bishop.* Oxford: Clarendon Press, 1955.

Crombie, A. C. *Robert Grosseteste and the Origins of Experimental Science, 1100— 1700.* Oxford: Clarendon Press, 1953.

Dales, Richard C. "Robert Grosseteste's Place in Medieval Discussions of the Eternity of the World." *Speculum* 61 (1986): 544–563.

Lewis, Neil. "Power and Contingency in Robert Grosseteste and Duns Scotus." In *John Duns Scotus: Metaphysics and Ethics,* edited by Ludger Honnefelder, Rega Wood, and Mechthild Dreyer. Leiden: Brill, 1996.

Lynch, Lawrence E. "The Doctrine of Divine Ideas and Illumination in Robert Grosseteste, Bishop of Lincoln." *Mediaeval Studies* 3 (1941): 161–173.

Marrone, Steven P. *William of Auvergne and Robert Grosseteste: New Ideas of Truth in the Early Thirteenth Century.* Princeton, NJ: Princeton University Press, 1983.

McEvoy, James. *The Philosophy of Robert Grosseteste.* Oxford: Clarendon Press, 1982; corrected reprint, 1986.

McEvoy, James. *Robert Grosseteste.* Oxford: Oxford University Press, 2000.

Rossi, Pietro. "Robert Grosseteste and the Object of Scientific Knowledge." In *Robert Grosseteste: New Perspectives on His Thought and Scholarship,* edited by James McEvoy. Turnhout, Belgium: Brepols, 1995.

Serene, Eileen F. "Robert Grosseteste on Induction and Demonstrative Science." *Synthèse* 40 (1979): 97–115.

Sharp, Dorothea E. *Franciscan Philosophy at Oxford in the Thirteenth Century.* London: Oxford University Press, 1930.

Southern, Richard W. *Robert Grosseteste: The Growth of an English Mind in Medieval Europe.* Oxford: Clarendon Press, 1986; 2nd ed., 1992.

Richard C. Dales (1967)
Bibliography updated by Neil T. Lewis (2005)

GROSSETESTE, ROBERT [ADDENDUM]

Since the mid-1970s, further research into Grosseteste's biography has produced more uncertainty than any new detail. His putative birth date and his humble beginnings are unquestioned, but the first twenty years of his life remain obscure. How he became learned in the liberal arts is open to speculation. Doubtless he was a *magister in artibus* by 1190, but there is no evidence that he studied (or taught) at either Oxford or Paris during this time. His association with Hereford may explain his lifelong interest in natural philosophy, although no evidence has emerged supporting a subsequent teaching career there, nor anywhere else, between 1198 and 1225. The scholarship is divided as to whether Grosseteste actually studied theology at Paris between 1209 and 1214 because of a compelling case made recently that he may have studied there between 1225 and 1229.

This latter dating would certainly explain how Grosseteste gained access to the commentaries of Averroes that were disseminated in Paris soon after 1225. It has been well-established that the 1220s was a highly productive period when he began to engage seriously the writings of Aristotle and his Arabic commentators. Some scholars have also called his Oxford chancellorship into question: The documentary evidence is open to interpretation and can be read as if he never were chancellor; or if he were, he was not the first. After 1229 or 1230, the first documented date of Grosseteste as a teacher, the biographical evidence is more abundant and the chronology from this point until his death is unchallenged.

See also Aristotle; Averroes.

Bibliography

Baur, L. *Die philosophischen Werke des Robert Grosseteste, Bischofs von Lincoln*. Münster: Aschendoff, 1912.

Callus, D.A.P., ed. *Robert Grosseteste, Scholar and Bishop*. Oxford: Clarendon Press, 1955.

Crombie, A.C. *Robert Grosseteste and The Origins of Experimental Science, 1100–1700*. Oxford: Clarendon Press, 1953.

Dales, R.C., ed. *Commentarius in VIII libros physicorum Aristotelis; e fontibus manu scriptis nunc primum in lucem*. Boulder, CO: University of Colorado Press, 1963.

Dales, R.C. and S. Gieben, ed. *Hexaëmeron*. London; New York: Published by the British Academy of Oxford University, 1982.

Ginther, James R. *Master of the Sacred Page: A Study of The Theology of Robert Grosseteste, ca. 1229/30–1235*. Aldershot, U.K.; Burlington, VT: Ashgate, 2004.

Lewis, Neil. "The First Recension of Robert Grosseteste's *De libero arbitrio*." *Mediaeval Studies* 55 (1991): 1–88.

Mackie, E. and J.W. Goering. *Editing Robert Gosseteste*. Toronto; Buffalo: University of Toronto Press, 2003.

Marone, Steven. *William of Avergne Robert Grosseteste: New Ideas of Truth in the Thirteenth Century*. Princeton, NJ: Princeton University Press, 1983.

McEvoy, J., trans. *Mystical Theology: The Glosses by Thomas Gallus and The Commentary of Robert Grosseteste on De mystica theologia*. Paris; Dudley, MA: Peters, 2003.

McEvoy, J. *The Philosophy of Robert Gosseteste*. Oxford: Clarendon Press; New York: Oxford University Press, 1982.

McEvoy, J. *Robert Grosseteste*. Oxford; New York: Oxford University Press, 2000.

Panti, Cecilia, ed. *Moti, virtú e motori celesti nella cosmologia di Roberto Grossatesta: studio ed edizonedei trattati de sphere, de cometis, de motu supercelestium*. Firenze, Italy: Sismel, 2001.

Rossi, Pietro, ed. *Commentarius in Posteriorum Analyticorum libros* Firenze: Olschki, 1981.

Southern, R.W. *Robert Grosseteste: The Growth of An English Mind in Medieval Europe*. Oxford U.K.: Clarendon Press; Oxford; New York: Oxford University Press, 1986.

Stevenson, F. *Robert Grosseteste, Bishop of Lincoln: A Contribution to The Religious, Political, and Intellectual History of the Thirteenth Century*. Dubuque, IA: Wm. C. Brown Reprint Library, 1962.

Thomson, S.H. *The Writings of Robert Grosseteste, Bishop of Lincoln, 1235–1253*. Cambridge, U.K.: Cambridge University Press, 1940.

James R. Ginther (2005)

GROTE, JOHN

(1813–1866)

John Grote, the English moral philosopher and epistemologist, was born at Beckenham in Kent. He was a younger brother of George Grote, the historian. Grote studied classics at Cambridge and became a fellow of Trinity College in 1837. He took orders in the Church of England and eventually obtained a church living at Trumpington, where he resided until his death. In 1855 he succeeded William Whewell as Knightbridge professor of moral philosophy at Cambridge. For a number of years an informal group, sometimes called the Grote Club, met regularly with him for philosophical discussion; Henry Sidgwick and John Venn were among its members.

Grote's writings were concerned primarily with ethics and theory of knowledge. He thought the former the more important study and intended the epistemological discussions in his *Exploratio Philosophica* to serve as prolegomena to his moral theory. Throughout his work he criticized the claim that only science or the "positive standpoint" could give us truth. Science treats perception simply as the action of one body on another, and it investigates the antecedents and concomitants of all thoughts and feelings indifferently. Hence it can give no adequate account of truth or falsity in thought. Philosophy, which is the study of thought and feelings as we are directly aware of them from within, can deal with truth and falsity, but it cannot give causal explanations. Hence the positive and the philosophical standpoints can lead us to truths that supplement each other. Grote argued with considerable acuteness that confusion of these standpoints was responsible for many of the difficulties of traditional theories of perception and knowledge, but he confessed himself unclear as to how they were related.

In ethics, Grote argued that utilitarianism overlooked the fact that man is essentially as active a creature as he is a sentient one. Concentrating only on human sentience, utilitarianism provides a theory of the good in feeling, but since it says little about right acting, it is unable to give an adequate account of the right distribution of good. The attempt to construct a positive science of morality is misguided and hopeless, since it omits the "ideal" element, or conception of what ought to be, which is central to morality. An ethical principle cannot be derived from facts alone, nor can it usefully be made true by definition; hence a basic intuition is required. There is, however, an important utilitarian element in morality, and that element provides a necessary check on possibly spurious intuitions. Grote suggested that the old conflict between utilitarianism and intuitionism should be seen as a conflict between partial views of a whole truth.

Grote held that the philosophical standpoint was more fundamental than the scientific. He gave a number of reasons for this. Underlying them is the view that the attempt to come to a rational understanding of the world implies the belief that the world is already rational, which implies in turn the belief that it is the creation of a mind. But mind, Grote held, can be understood as such only from the philosophical standpoint. The attempt to act morally in the world presupposes, similarly, a belief that the world is morally ordered, and this implies a belief in a moral governor. Grote interpreted these beliefs theistically. His development of them anticipated in many ways the absolute idealism of the generation after him. He argued that all truth is systematically interconnected; that truth is ultimately to be understood as coherence, rather than correspondence; and that the distinctions of perception and conception and of necessity and contingency are relative. In ethics he worked toward a view emphasizing self-development and man's duty in his station.

It has been said that Grote should be viewed as the first of the Cambridge analytic philosophers, and certainly his great respect for ordinary language and ordinary thought, his persistent attempts to find and remove logical confusions, his insistence on the importance of clarity, and his pursuit of it in detailed and painstaking criticism have obvious affinities with the work of that group. There is, however, little evidence to show that he had much direct influence on anyone, and his writing, which is difficult and prolix, has been very little studied despite its acuteness and considerable originality.

See also Epistemology; Ethics, History of; Intuitionism and Intuitionistic Logic; Sidgwick, Henry; Utilitarianism; Venn, John; Whewell, William.

Bibliography

With the exception of *Exploratio Philosophica,* Vol. 1 (Cambridge, U.K.: Deighton, Bell, 1865), Grote's main writings were edited from his papers and published posthumously by his literary executor, J. B. Mayor. They are: *Examination of the Utilitarian Philosophy* (Cambridge, U.K., 1870); *Sermons* (Cambridge, U.K., 1872); *Treatise on the Moral Ideals* (Cambridge, U.K.: Deighton, Bell, 1876); and *Exploratio Philosophica,* Vol. 2 (Cambridge, U.K., 1900). There are also some papers, of which the most interesting is a series called "Glossology" in the *Journal of Philology,* 1872 and 1874. Further details may be found in Mayor's prefaces to the volumes he edited. There are no detailed secondary studies, but G. Watts Cunningham, *The Idealistic Argument* (New York: Century, 1933) has a chapter on Grote.

J. B. Schneewind (1967)

GROTIUS, HUGO

(1583–1645)

The Dutch jurist and statesman Hugo Grotius, or Huig de Groot, was born at Delft, of a distinguished Calvinist family. He entered the University of Leiden when he was eleven, graduating with great distinction at the age of fourteen. At fifteen he served as a member of a Dutch mission to France and obtained the degree of doctor of law at the University of Orléans. In 1601 Grotius was asked for a legal opinion by the Dutch East India Company in an international case, which appears to have spurred the writing of his pamphlet in defense of freedom of the seas (*Mare Liberum,* 1609) and generally stimulated his enduring interest in international law. In 1607 Grotius was appointed advocate general of the fisc of the provinces of Holland, Zeeland, and Friesland. In 1613 he became the pensionary for Rotterdam and went to England as a member of a Dutch diplomatic mission. A bitter theological dispute, in which Grotius sided with the estates of Holland against orthodox Calvinism (supported by Prince Maurice of Holland) led to a special trial and the condemnation of Grotius to life imprisonment. During his imprisonment he wrote the famous pamphlet, *De Veritate Religionis Christianae* (Leiden, 1627). In 1621 Grotius escaped from prison and fled to France, where he wrote his great work *De Jure Belli ac Pacis* (1620–1625), dedicated to Louis XIII. Grotius later returned to Holland. From 1634 to 1644 he was Swedish ambassador to France. He was recalled to Sweden in 1644 and died in Rostock on his way back from Sweden to Holland shortly after.

Grotius's enduring influence upon legal science and, in particular, on the science of international law may be attributed to qualities somewhat comparable to those of John Locke. Both men formulated, articulated, and systematized, at a critical point in history, certain ideas and principles that suited the needs of a changing society.

But whereas Locke articulated the rights of the individual in a rapidly expanding, acquisitive, and increasingly antiabsolutist society, Grotius understood that the international society of burgeoning sovereign states had to find and abide by certain rules of conduct in war and peace, formalizing diplomatic relations and mutual respect for sovereignty. Since modern international society is still dominated by the legal and political supremacy of the national state, Grotius's classical treatise, *De Jure Belli ac Pacis,* is still an essential foundation for international law. The international order of the Middle Ages, based on the twin foundations of the ecclesiastical authority of the Church of Rome and the political authority of the emperors, had crumbled together with the social, economic, and spiritual conditions on which it was based. New kingdoms, dukedoms, principalities, and cities had emerged from the debris. Europe was torn by wars, big and small, motivated by religious, dynastic, political, and social conflicts. While Grotius wrote his principal work, the Thirty Years' War was raging in much of Europe, demonstrating the destructive effects of the lawlessness of a society that had not yet developed new rules of intercourse appropriate to the emerging society of sovereign nations. There was no prospect of reestablishing the international authority exercised by popes and emperors. There was no hope of abolishing or outlawing war. But there was an urgent need to establish a new code of behavior, and, more than that, a need to humanize the conduct of war even within modest limits. To lay the foundations for such a development was a gigantic task, a

task only for someone who could combine the qualities of philosopher, political scientist, jurist, humanist, and diplomat. That man was Hugo Grotius, a man of prodigious learning—theologian, philologist, historian, and poet, as well as jurist—who was also an active diplomat.

All his various interests are reflected in his great treatise, a rambling work ranging over many fields of human knowledge, studded with quotations and references to innumerable scholars and sources. *De Jure Belli ac Pacis* established a partially legalized system of international relations by blending certain general principles of political and moral philosophy with state practice. It is this combination that gives to Grotius's work the flexibility and durability that enabled subsequent generations to make use of it by emphasizing the one or the other aspect.

NATURAL LAW

Grotius was first a major exponent of the philosophy of natural law and of social contract. Second, he was an Aristotelian whose deepest and most abiding belief was in the power of reason and the rationality of man. Third, Grotius was a pragmatic diplomat who, through the observation and practice of diplomacy in a singularly disturbed and savage period, was fully aware of the practices of states in peace and war—and it was war that dominated both the life of the people and the thought of Grotius. But fourth, Grotius was a humanist in the spirit and tradition of his master, Erasmus of Rotterdam, a man who abhorred the brutality and lawlessness of war and violence, and whose principal purpose, therefore, was not only to civilize the conduct of war but also to place certain limitations upon its legality. He combined this last objective with his belief in reason and in natural order in the formulation of his famous theory of the *bellum iustum* (the just war).

As a natural law philosopher, Grotius was much closer to the Stoics than to the Scholastics. Like the former, he derived the postulates of natural law from principles of reason rather than of divine order. Such reason was founded in the human intellect. "Natural law is so immutable that it cannot be changed by God himself."

The natural law doctrine provided Grotius with the theoretical foundation for certain overriding principles of order in the relations between states. It also gave him faith in the rationality of man and in man's potentialities for developing a better society in accordance with the needs of social and international life. Grotius was, of course, well aware that there was in his time no law-giving authority superior to the will of the states. It was, therefore, necessary for him to find some principle that could bind the nations to a common standard of behavior. He found this principle in *pacta sunt servanda,* the respect for promises given and treaties signed. In the absence of an international sovereign authority, modern international lawyers, such as Dionisio Anzilotti and Hans Kelsen, have reaffirmed the same principle as the metalegal foundation of international law.

Grotius formulated a large number of other principles of natural law that inevitably share the weakness of all natural law teachings—that is, the sublimation of certain political postulates into immutable principles of order. Among Grotius's rules of natural law were respect for other people's property and the restitution of gain made from it, the reparation of any damage caused by a person's (or nation's) fault, as well as certain elementary principles of punishment. The political coloration of natural law is more evident in Grotius's postulate of the freedom of the seas. This postulate corresponded to the interests of the Netherlands as the world's leading maritime nation. It was opposed by the Englishman, John Selden (*Mare Clausum,* London, 1635), at a time when England was still struggling against stronger maritime nations.

SOCIAL CONTRACT

The other pillar of Grotius's legal philosophy was the theory of social contract, which also led him to emphasize the supremacy of the compact as the highest binding principle of law. Unlike later theorists of social contract, Grotius considered the contract as an actual fact of human history. In his view, the constitution of each state had been preceded by a social contract, by means of which each people had chosen the form of government they considered most suitable for themselves. While each people had the right to choose their own form of government, they forfeited the right to control or punish the ruler, however bad his government, once they had transferred their right of government to him. Generally, Grotius, like Thomas Hobbes, reflected not only the need of a disturbed society for strong governmental authority, but also the essentially absolutist and predemocratic character of government of that period. In his own official and diplomatic career Grotius represented autocratic governments.

INTERNATIONAL LAW

Aware of the insufficiency of natural law to supply more than certain general guiding principles, Grotius based the principal body of international law on *ius voluntarium* (the body of treaties and other engagements that form the

bulk of international state practice). Although a realist, Grotius was not a cynic. He believed not only in the essential rationality of man and peoples, but also in the necessity of progress from war to peace, from international anarchy to international order. His principal contribution in this respect was his theory of the *bellum iustum*. A major part of the second book of his treatise was devoted to the problems of the legality of war. For a war to be just, there must exist a legal cause for it. Essentially, there are only three types of just wars: Those that are conducted in defense against an actual or immediately threatening injury; those aimed at the recovery of what is legally due; and those inflicting punishment for a wrong done. Each of these categories allows for a great degree of latitude, especially in the absence of an impartial international judicial authority that can decide between conflicting claims.

Nevertheless, this emphasis on the need to justify war, and the limitation of its justification to causes that even today would be regarded as essentially defensive against wrongful injury, was a remarkable contribution to international order. It became obscured and forgotten during subsequent centuries of absolute national sovereignty, particularly during the nineteenth century when the aggressive national state celebrated its greatest triumphs, in practice as in theory. In modern time, the League of Nations Covenant and the United Nations Charter have attempted once again to distinguish between just and unjust wars. The future of humankind may well depend on the elaboration of an authoritative method of finding reliable and enforceable criteria for distinguishing between wars of aggression and wars of defense and on the establishment of an impartial forum to decide on claims for the reparation of wrongs alleged to be inflicted by one state on another.

International law in our own day is still essentially based on state practice as recorded in custom, treaties, and other international agreements; but these practices allow for the evolution of international law, not so much in the terminology of natural law as in the similarly conceived evocation of "general principles of law recognized by civilized nations." Some of these principles were applied, with dubious theoretical justification, in the Nuremberg and Tokyo trials of German and Japanese war criminals.

Grotius's doctrines were inevitably a mixture of reactionary and progressive principles. On the one hand he felt compelled to justify many barbarous practices of war, subsequently condemned in modern rules of warfare (yet surpassed in cruelty by modern war). Again, Grotius concurred with the great majority of legal and political philosophers in denying to the individual the right of resistance to an oppressive sovereign, although he affirmed the right of passive resistance of an individual against unjust wars. On the other hand, Grotius anticipated by centuries some of the principles of the Nuremberg Charter by regarding as justified a war waged to prevent the maltreatment by a state of its own subjects. And Grotius's concern with the individual stands in noble contrast to subsequent absolutist political theories.

In working and thinking within the limitations of his time, Grotius did not differ from any other philosopher, jurist, or political scientist. What is remarkable is that, in the midst of a war that threatened to undermine the whole fabric of European society, he developed principles and standards that can still serve as the basic themes for the struggle for international order in our time.

See also Aristotelianism; Hobbes, Thomas; Locke, John; Natural Law; Peace, War, and Philosophy; Philosophy of Law, History of; Political Philosophy, History of; Social Contract; Stoicism.

Bibliography

PRIMARY WORKS

De Jure Belli ac Pacis Libri Tres. Paris, 1625. Translated by F. W. Kelsey et al. Oxford, 1925.

SECONDARY WORKS

Balogh, Elemér. "The Traditional Element in Grotius' Conception of International Law." *New York University Law Quarterly Review* 7 (2) (1929): 261–292.

Basdevant, Jules. "Grotius." In *Les fondateurs du droit international,* edited by A. Pillet. Paris: V. Giard and E. Brière, 1904.

Bull, Hedley, Benedict Kingsbury, and Adam Roberts, eds. *Hugo Grotius and International Relations.* Oxford: Clarendon Press, 1990.

Edwards, Charles S. *Hugo Grotius: The Miracle of Holland.* Chicago: Nelson-Hall, 1981.

Gurvitch, George, "La philosophie du droit de Grotius et la théorie moderne du droit international." *Revue de métaphysique et de morale* 34 (2) (1927): 365–391.

Knight, W. S. M. *The Life and Works of Hugo Grotius.* London: Sweet and Maxwell, 1925.

Lauterpacht, Sir Hersch. "The Grotian Tradition in International Law." *British Year Book of International Law* 23 (1946): 51.

Miller, Jon, ed. "Stoics, Grotius, and Spinoza on Moral Deliberation." In *Hellenistic and Early Modern Philosophy.* Cambridge, U.K.: Cambridge University Press, 2003.

Nussbaum, Arthur. *A Concise History of the Law of Nations.* New York: Macmillan, 1954.

Onuma, Yasuaki, ed. *A Normative Approach to War.* Oxford: Clarendon Press, 1993.

Ter Meulen, Jacob, and P. J. J. Diermanse. *Bibliographie des écrits imprimes de Hugo Grotius.* The Hague: Nijhoff 1950.

Tuck, Richard. *Philosophy and Government 1572–1651.* Cambridge, U.K.: Cambridge University Press, 1993.

Tuck, Richard. *The Rights of War and Peace.* Oxford: Oxford University Press, 1999.

Van Eysinga, W. J. M. *Huigh de Groot, een Schets.* Haarlem: Tjeenk Willink, 1945.

Wolf, Eric. *Grotius, Pufendorf, Thomasius.* Tübingen, 1927.

Wolfgang Friedmann (1967)
Bibliography updated by Philip Reed (2005)

GUILT

There are two main forms of the idea of "guilt"—moral guilt and legal or quasi-legal guilt. Originally these were not sharply distinguished, but enlightened thought requires that they should be. In outward substance the two often coincide. In committing a crime one is usually morally at fault, but the degree of one's guilt is not likely to be the same in the two respects in such instances. We may in any case be morally guilty and legally innocent—and vice versa. Few who consult this book have committed a crime, but who is there who has never done anything for which he may be morally reproached? Some of the most vicious things men do are well within the law. Nor would it be wise to legislate against all forms of moral evil—much of that would defeat the purpose of morality. One may also break the law and incur no moral blame. This might be because of unavoidable ignorance (of the law or of some matter of fact), but we could be blameless even in committing a crime deliberately. That would come about if we broke the law on conscientious grounds. Some of the people we admire most (religious or political martyrs, for example) have put religious or moral scruples before the claims of the law. They were not in all cases outwardly justified. The outward justification of resistance to the law is greatest under oppressive government. The duty to conform is very great where there are constitutional means of seeking redress or reform. Upholding constitutional procedure is normally much more important than righting a particular wrong. Resistance (but not of course normal opposition) is very extreme medicine in a democracy, and it is not always justified under tyranny. Persuasion is the best means of reform. But whether outwardly justified or not, a person is free from moral blame (and perhaps worthy of much praise) if he breaks the law in obedience to his own conscience.

MORAL GUILT AND DETERMINISM

Legal responsibility means liability to punishment, and legal guilt thus means that one has merited some punishment. This may be understood in a retributive, reformative, or deterrent sense. On the latter view, the commonest today where strictly legal or social issues are concerned, absolute freedom of choice is not presupposed. It is of course pointless to seek by punishment to deter someone who in no sense controls what he does—or to make him an example for others. We restrain the insane or the delirious. We do not punish them, and it is absurd to punish people for what they do by accident. But punishment is not made pointless when we act in character and wittingly do certain things even though, being the persons we are, we could not help doing them. Punishment as deterrence is consistent with determinism, for our conduct on other occasions—and the conduct of other persons—could be affected by punishment or the threat of it. But in moral matters punishment is a secondary issue. To be morally guilty is to have incurred moral disvalue or to be morally blameworthy. This may call for punishment or some other outward censure, but that is a further question.

Moral guilt is a more basic notion than punishability. It is a unique moral concept not to be merged in associated social and legal notions. Moral guilt presupposes freedom of a more radical kind than legal guilt and is hard to reconcile with any form of determinism. The moral evil it involves must be distinguished sharply from nonmoral evils like sickness, pain, error, and stupidity. I am not to blame for being ill or for failing to win a race, compose a poem, or solve a mathematical problem. I am to blame for moral failure. I cannot help the former failures, provided I try, if it is my duty. But it is hard to see how there could be moral failure if there is any sense in which I could not help it.

THE ASSUMPTION OF ABSOLUTE FREEDOM. But how strictly are these last words to be understood in the case of moral guilt? Is absolute freedom presupposed? If it is, are we ever guilty in the strictly moral sense? Is there not some continuity of character and conduct? Plainly there is such a continuity, and advocates of absolute freedom of choice as a moral requirement have therefore argued that moral praise or blame only apply on the occasions where there is a conflict between "duty and interest"—that is, between what we most want to do and what we think we ought to do. During much of one's life there is no such conflict, and we can therefore anticipate one another's actions with much confidence: we know what to expect of

people we have come to know. But character and duty will sometimes draw apart. To that extent, it is maintained, nothing affects the outcome but the act of choice itself. If we fail to make the effort of will—an absolutely free one in this case—to overcome some weakness of character, and if we thus follow the line of least resistance rather than the call of duty, we incur moral guilt. The degree of the guilt depends not on the outward features of the situation and the magnitude of the ill we do, or at least not directly so, but on the effort of will that would have been required to do right. But it should be noted well that the more outwardly vicious an act may be, the less is the effort needed to resist a temptation to do it, for one can normally presuppose much natural resistance to the act in one's own character. The less the effort required, the more we are to blame for not making it; the greater the effort we do make, the greater our moral worth.

It follows from this view that while we may, for practical and kindred purposes, censure misdeeds in their outward form, we need to be very chary of passing strictly moral judgments on other persons. If we have reason to believe (as is often the case) that someone has acted contrary to his moral convictions, we can impute to him some measure of blame, but how much is much harder to assess than the outwardly objectionable features of a situation. It is also much harder to assess the positive moral worth of another person than to assess his moral guilt. For we know in the latter case that the effort required was not forthcoming; in the former case it is harder to know how much to ascribe to natural good qualities of character, a benevolent or naturally plucky nature, and so on, and how much to free effort. Only the agent himself and God can know the full inner story.

THE FEELING OF GUILT

We must, however, distinguish "guilt" in the strict moral meaning from the sense of guilt. The latter is the feeling that accompanies the consciousness of being guilty. It is appropriate that we should feel remorse for wrongdoing, the proper tone of the feeling being determined by its appropriateness to the situation of guilt. There are kindred feelings appropriate to the wrongdoing we encounter or suspect in others, feelings or attitudes of blame and indignation. The feelings we actually have are not always appropriate to the situation, and there may thus be a sense of guilt out of all proportion to the facts of the case. Some people seem even to enjoy the sense of guilt and to cultivate it. Psychologists have helped us a great deal to understand these deviations and that other curious aberration by which some people feel guilty for things they have not done at all.

PSYCHOLOGICAL REASONS. Some psychologists go further. They try to account for guilt entirely in terms of psychology. A common form of this attempt is that which ascribes guilt to an alleged "need for punishment." This need comes about through punishment or some other disapproval we suffer in infancy. Coming to expect punishment for certain acts, we feel distress when we wait for it without getting it over, and the strain and anxiety induced in this way is suppressed and operates subconsciously afterwards to produce the sense of guilt in mature experience. There is also the introjection into the "superego" of the relief experienced by those who punish us. These theories no doubt reflect states of mind which psychological investigation uncovers, and the layperson can appreciate much of them from common experience. But they seem nonetheless to be mainly concerned with aberrations and an unhealthy assumption of guilt, or perverse ways of dealing with it. The core of guilt is an ethical one, which psychology does not explain away.

COLLECTIVE VERSUS INDIVIDUAL GUILT

If guilt, in the proper sense, turns on deliberate wrongdoing, it seems that no one can be guilty for the act of another person—there can be no shared or collective or universal guilt. Guilt is incurred by the free choice of the individual. But many have questioned this. Among them are some sociologists who misrepresent in this way the dependence of the individual on society. But the main location of the idea of collective guilt is religion. Many forms of the doctrines of original sin and universal sin regard guilt as a pervasive state of humankind as a whole. It is the guilt of "man," not of this or that person as a whole. Others qualify this and speak of original sin which does not include original guilt. Others hold that while there can be no "great sin" and "little sin," there is inequality of guilt. But it is hard to reconcile the notion of universal sin or guilt, in any form, with elementary ethical convictions. Such notions can also do great harm, both by leading to victimization of the innocent—as in the treatment of Jews by the Nazis—and in undermining the sense of responsibility; for collective guilt is not the guilt of anyone in particular.

SOURCES OF DOCTRINES OF COLLECTIVE GUILT. Why then do such doctrines of collective guilt seem plausible? Mainly through religious confusions like the following. (1) The sense of religious unworthiness, the awe

felt in the presence of God, is mistaken for moral culpability. (2) Certain forms of religious experience are apt to be overwhelming, and the strain is eased at times by encapsulating the divine within the finite media or symbols by which it is known. This is the root of idolatry. The most grievous form of this is that by which the person himself becomes the idol—he aspires to make himself as God. But this distortion of religious experience tends to be conflated, in the heat of prophetic experience, with the expressly moral wickedness of putting one's own wishes before the proper claims of others. This encourages the notion of an unavoidable state of sin and guilt. (3) Guilt is what we seem most disposed to suppress, and at the unconscious level the confusions noted are apt to be intensified. (4) Religious doctrines have often been based on first-order religious utterances taken out of their full context and apart from the experience which prompted them. The figurative character of such utterances is also overlooked—for example, in interpretations of the metaphors of "bondage" or of "sin warring in all my members." (5) Wrongdoing has a cumulative influence that affects the state and the situation of persons irrespective of their own guilty actions; it thus tends to drive men in on themselves and hinder healthy relations with other persons—and with God. This also, or the misrepresentation of it, lies behind misleading doctrines of collective guilt. (6) The idea of universal guilt has often been made the pretext for evading the challenge of high ideals professed by religious people. This seems especially true of much Augustinian theology. (7) Religious confusions are deepened by confusions between the points of view of law—where the idea of corporated guilt has some place—and the point of view of morality.

Recent anthropology has thrown new light on the origin of the idea of guilt. There was at first little distinction between the points of view of law and of morality, both being merged in communal custom. Nor was heed adequately paid to whether the results of an act were those a person intended. The community was also more the bearer of guilt than the individual, and harsh judgments were thus passed on the innocent and bitter feuds perpetuated. But we should not allow this to determine for us how guilt must be understood in enlightened thought. Ethical notions are not jeopardized by having lowly and doubtful origins.

Religious thought today helps us to appreciate what is true and what is false in notions like collective guilt. But much recent sociology and some recent ethics go further and challenge the ultimacy of the ideas of guilt and responsibility. These are thought by some moralists and psychologists to be ideas we ought to have outgrown—"theological anachronisms." A sound ethical theory and better understanding of religion should correct these tendencies.

See also Determinism, A Historical Survey; Philosophy of Law, History of; Philosophy of Law, Problems of; Punishment.

Bibliography

For a general discussion of the subject, see the symposium on "Guilt" by H. D. Lewis, J. W. Harvey, and G. A. Paul, in *Proceedings of the Aristotelian Society,* Supplementary Volume XXI. See also the papers on Free Will, Responsibility, and Guilt by C. A. Campbell, P. H. Nowell-Smith, and Erich Fromm in *A Modern Introduction to Ethics,* edited by Milton K. Munitz (Glencoe, IL: Free Press, 1958), pp. 356–416. Sir Walter Moberly, *Responsibility* (Oxford: Oxford University Press, 1951) consists of three lectures dealing with the moral and religious aspects of guilt and sin. A classical discussion of the religious and biblical side of the subject is F. R. Tennant, *The Concept of Sin* (Cambridge, U.K.: Cambridge University Press, 1912). N. H. G. Robinson, *Concept of Sin* (New York, 1950) is a first-rate survey of the subject in the work of representative recent theologians like Karl Barth, Emil Brunner, and Reinhold Niebuhr. The liberal approach to this subject may be found in L. Harold DeWolf, *The Religious Revolt against Reason* (New York: Harper, 1949) and in H. D. Lewis, *Morals and the New Theology* (New York: Harper, 1948), Chs. 5–7.

The Freudian view is well presented in J. C. Flügel, *Man, Morals, and Society* (New York: Viking, 1961). This should be considered along with Sigmund Freud's *New Introductory Lectures on Psychoanalysis* (New York: Norton, 1933), pp. 78–106 and Susan Isaacs, *Social Development in Young Children* (London: Routledge, 1933). More recent psychoanalytic works are Melanie Klein and others, *Developments in Psychoanalysis* (New York, 1952); Melanie Klein, *Envy and Gratitude: A Study of Unconscious Sources* (New York: Basic, 1957); and E. Glover, *Technique of Psychoanalysis,* rev. ed. (New York, 1955). There are also helpful references to guilt in Walter J. H. Sprott, *Social Psychology* (London: Methuen, 1952), p. 168ff. The tradition of Ivan Petrovich Pavlov is continued in B. F. Skinner, *Science and Human Behavior* (New York: Macmillan, 1953), in which the conditioning effect of early training is stressed.

There is a short discussion of guilt in T. R. Miles, *Religion and the Scientific Outlook* (London: Allen & Unwin, 1959), Ch. 12.

The sociological approach is well exemplified in Barbara Wootton, *Social Science and Social Pathology* (London: Allen and Unwin, 1959). Criticism of this approach will be found in H. D. Lewis, *Freedom and History* (New York: Macmillan, 1962), and there is further discussion of the religious implications of the sociological approach in Fred Berthold, *The Fear of God* (New York: Harper, 1959).

H. D. Lewis (1967)

GUO XIANG
(c. 252–312 CE)

A champion of the Learning of the Mysterious (*Xuanxue*) or neo-Daoism that gained prominence in third-century China, Guo Xiang (c. 252–312 CE) is best known for his commentary to the *Zhuangzi*, which offers to reconcile orthodox teachings (*mingjiao*) with Daoist naturalness (*ziran*).

Like other neo-Daoist philosophers, notably Wang Bi (226–249), Guo recognizes the creative power of Dao; however, contrary to Wang, Guo rejects that "beings originate from nonbeing," which establishes Dao as the metaphysical ground of being (*Zhuangzi* commentary, chs. 2 and 23). The appeal to an anthropomorphic heaven or original substance as the source of creation should, according to Guo, be rejected, for it begs the question of the cause of its own being. Nonbeing, however, is not the answer, because nonbeing remains an abstraction and abstractions cannot bring about creation. Being and nonbeing are mutually exclusive, according to Guo, who writes "nonbeing cannot change into being" (*Zhuangzi* 22). Consequently, the only logical explanation of the origin of being is that "being spontaneously produces itself" (*Zhuangzi* 2).

This explanation introduces Guo's concept of "self-transformation," for which he is particularly famous. Whereas Wang Bi values nonbeing, Guo favors being. At the most basic ontological level, being is "so of itself" (*ziran*), and Guo believes that "we may know the causes of certain things and affairs near to us. But tracing their origin to the ultimate end, we find that without any cause, they of themselves come to be what they are. Being so of themselves, we can no longer question the reason of their being, but should accept them as they are" (*Zhuangzi* 14).

SELF-TRANSFORMATION AFFIRMS THE IMMANENCE OF DAO

Guo explains that the Dao pervades and informs nature as vital energy (*qi*) and that all beings are endowed with a "share" or "allotment" of the inexhaustible energy of Dao, and this defines their nature (*xing*) and capacity. Significantly, "benevolence and rightness" stem from nature (*Zhuangzi* 14); and, moreover, the state of *ziran* depicts an organized regime governed by principles and marked by interdependence and hierarchical order.

Given that individual *qi*-endowment varies, differences in capacity—for example, lifespan and intelligence—should be recognized. This is destiny (*ming*), in that "what one is born with is not something undue or vain" (*Zhuangzi* 5). This, then, begs the question: Is Guo—as many scholars hold—a fatalist?

Destiny dictates that one is born of sagely character or average capacity. Yet, Guo also attempts to distinguish *ming* as fact from value, and to affirm development in human flourishing. Fundamentally, differences in endowment do not constitute any basis for value judgment. Rather, as the *Zhuangzi* urges, what should be recognized is the "equality of things."

Unlike Wang Bi, who emphasizes the "one," Guo embraces the "many." Individuality and authenticity should be cherished (*Zhuangzi* 10). The Daoist goal can be defined as the realization of one's nature, and in particular the optimization of one's inborn capacity. As nature blossoms, "destiny" is fulfilled.

While this may not detract entirely from the charge of fatalism, Guo introduces a dynamic view of nature and destiny. The world of *ziran* is never static; it changes and renews itself constantly. Limits notwithstanding, one's potential should not be underestimated. The sage or person of Dao nourishes his nature and adapts to change, which brings out the meaning of nonaction (*wuwei*). Nonaction "does not mean folding one's arms and keeping quiet" (*Zhuangzi* 11). It is also not a technical skill; rather, nonaction stems from a discernment of *ziran*, which translates into a mode of being and a spirit of action, according to which one performs all functions.

Politically, nonaction means that the ruler enables the people to develop their nature and potential. Artificial restrictions and interference should be minimized, and because needs and circumstances change, sociopolitical practice should not be fossilized—timely adjustments ensure renewal and harmony in a dynamic realm. In this way Guo tries to reconcile the *mingjiao* (orthodox teachings) with *ziran*. Whereas the former refers to doctrines of propriety and government, the latter aspires towards transcendence and freedom from mundane concerns. Conflict arises, then, when orthodox teachings are seen to impinge on nature, or when transcendence is equated with renunciation. For Guo, however—because social and natural phenomena are governed by the same set of principles—*mingjiao* and *ziran* merge into one.

See also Chinese Philosophy: Daoism; Wang Bi.

Bibliography

Fung Yu-lan, tr. *Chuang Tzu: A New Selected Translation with an Exposition of the Philosophy of Kuo Hsiang*. Shanghai:

Commercial Press, 1933; reprinted, New York: Gordon Press, 1975.

Fung Yu-lan. *A History of Chinese Philosophy*. Trans. Derk Bodde. Vol. 2. Princeton, NJ: Princeton University Press, 1983.

Knaul, Livia. "Kuo Hsiang and the Chuang-tzu." *Journal of Chinese Philosophy* 12 (1985): 429–447.

Robinet, Isabelle. "Kouo Siang ou le monde comme absolu." *T'oung Pao* 69 (1983): 73–107.

Ziporyn, Brook. *The Penumbra Unbound: The Neo-Taoist Philosophy of Guo Xiang*. Albany: State University of New York Press, 2003.

Alan K. L. Chan (2005)

GURWITSCH, ARON

(1901–1973)

Aron Gurwitsch was one of the leading proponents of and contributors to phenomenology in the twentieth century. He was one of a small number of philosophers who brought phenomenology from Europe to the United States and led its growth into a significant presence there. Gurwitsch's main influence came through his expositions of Edmund Husserl's phenomenology and his original contributions that modified and supplemented Husserl's work.

Gurwitsch was born on January 17, in Vilnius, Lithuania (then a part of Russia), of parents who were descended from a long line of Jewish scholars. Following the pogroms of 1905 and 1906, the family moved in 1907 to Danzig where Gurwitsch received his early education. He began his university education at the University of Berlin in 1919, where he studied mathematics, physics, psychology, and philosophy; here he came under the guidance of the philosopher and psychologist Carl Stumpf. On Stumpf's suggestion, Gurwitsch went to the University of Freiburg in 1922 to attend some of Husserl's lectures. Gurwitsch was so influenced by Husserl's style of philosophizing that he decided to devote his life to the continuation and expansion of Husserl's phenomenology.

Gurwitsch left Berlin for the University of Frankfurt, where he studied with the psychiatrist Kurt Goldstein and the Gestalt psychologist Adhemar Gelb, whose studies of psychological pathologies stemming from brain injuries related to human capacity for abstraction, a topic in which Gurwitsch was interested. During this time Gurwitsch realized that Gestalt theory and phenomenology could benefit from one another. This led him to develop in his doctoral dissertation a field theory of sensory perception in which he rejected Husserl's concept that a nonworldly, transcendental ego was at the basis of the human ability to experience a world and developed a nonegological conception of consciousness that was like the one found later in Jean Paul Sartre's work. The dissertation, *Phenomenologie der Thematik und des reinen Ich*, was published in 1929. Husserl was impressed by this work and there were regular contacts between them until 1933.

In 1929 Gurwitsch returned to Berlin, with a Prussian habilitation grant, where he worked on and essentially completed *Die mitmenschlichen Begegungen in der Milieuwelt* (1976). This work concerned basic problems of social phenomenology and contained an original approach to social perception that combined phenomenological and gestaltist insights. Due to the rise of National Socialism that led him to flee Germany and go to France in 1933, Gurwitsch did not publish this book, and it was published posthumously in 1976.

In 1933 and 1934 Gurwitsch began lecturing on Gestalt psychology and phenomenology at L'Institut d'Histoire des Sciences at the Sorbonne in Paris. These lectures were attended by Maurice Merleau-Ponty, whose major work, *Phenomenology of Perception*, incorporates much that he acquired from Gurwitsch. Gurwitsch's posthumously published *Esquisse de la phénoménologie constitutive* (2002) is based on that latter parts of these lectures. In Paris, in 1937, Gurwitsch met the sociologist and philosopher Alfred Schutz. Their correspondence from 1939 to 1959 has been published as *Philosophers in Exile* (1989). In addition to discussing and exploring intellectual topics, the letters contain a fascinating look at the difficulties of the lives of emigré scholars at that time. With the help of Schutz, who preceded him in 1939, Gurwitsch emigrated to the United States in 1940, and took a position at Brandeis University in 1948, first in mathematics and then, in 1951, in philosophy.

In the United States Gurwitsch began work on his magnum opus, *The Field of Consciousness*, published first in French translation in 1957, and then in the original English in 1964. In this work, Gurwitsch related phenomenology to the thought of William James and others, offered a criticism of various dualistic theories of perception, and gave a masterful account of Gestalt theory. This is followed by what became the most influential part of the work, his account of perceptual consciousness wherein the field of what one is aware is articulated into theme, thematic field, and margin. Building on Husserl's work, but abandoning what he took to be a dualism in Husserl's theory where higher level cognitive functions worked on lower level sensations to produce the object as

experienced (what Husserl called the "noema"), Gurwitsch creatively employed Gestalt theoretical concepts to analyze the structure of the focally perceived object (theme) as well as its relationship to the wider experienced context (thematic field) and to other co-conscious items that are not relevant to the theme and thematic field (margin). His account of the object as experienced (noema) was an alternative to Husserl's theory and led to considerable discussion in the secondary literature. A part of the whole manuscript that Gurwitsch wrote that was not published in *The Field of Consciousness* was posthumously published as *Marginal Consciousness* (1985) and contains detailed analyses of human awareness of the margin.

When Alfred Schutz—who held a position at the Graduate Faculty of the New School for Social Research in New York City—died in 1959, Gurwitsch became his successor, joining the phenomenologist Dorian Cairns, and taught there until his retirement in 1972. This was a time when phenomenology attracted much attention in the United States and the New School was a major center for research and study. Gurwitsch, Cairns—and, earlier, Schutz—were major influences on a new generation of phenomenologists. Gurwitsch helped found the Society for Phenomenology and Existential Philosophy (SPEP) in 1962, and later the Husserl Circle, two major forums for the presentation of phenomenological research. During this time he republished eighteen essays in *Studies in Phenomenology and Psychology* (1966). Another group of previously published and unpublished essays dating from 1937 came out posthumously in *Phenomenology and the Theory of Science* (1974). Along with some of Gurwitsch's influential original and critical work these two volumes contain some of the authoritative accounts of Husserl's philosophy that made Gurwitsch such a leading exponent and interpreter of Husserl's philosophy.

Gurwitsch's interests went beyond Husserl and phenomenology. He also wrote and prepared for publication *Leibniz: Philosophie des Panlogismus*, published posthumously in 1974 and wrote what was posthumously published as *Kants Theorie des Verstehens* (1991). Gurwitsch died on June 6, 1973.

See also Gestalt Theory; Phenomenology.

Bibliography

WORKS BY GURWITSCH

"Phänomenologie der Thematik und des reinen Ich." In *Psychologische Forschung* 12 (1929): 279–381. Translated by Fred Kersten as *Phenomenology of Thematics and of the Pure Ego: Studies of the relation between Gestalt Theory and Phenomenology*, in *Studies in Phenomenology and Psychology*, 175–286.

Théorie du champ de la conscience. Paris: Desclée de Brouwer, 1957. Translated as *The Field of Consciousness.* Pittsburgh: Duquesne University Press, 1964.

Studies in Phenomenology and Psychology. Evanston, IL.: Northwestern University Press, 1966.

Leibniz: Philosophie des Panlogismus. New York: Walter de Gruyter, 1974.

Phenomenology and the Theory of Science, edited by Lester Embree. Evanston, IL: Northwestern University Press, 1974.

Die mitmenschlichen Begegungen in der Milieuwelt, edited by A. Métraux. Berlin: Walter de Gruyter, 1976. Translated by Fred Kersten as *Human Encounters in the Social World*, edited by Alexandre Métraux. Pittsburgh: Duquesne University Press, 1979.

Marginal Consciousness, edited by Lester Embree. Athens: Ohio University Press, 1985.

Schütz, Alfred, and Aron Gurwitsch. *Briefwechsel 1939–1959,* edited by von Richard Grathoff. München, Germany: Wilhelm Fink Verlag, 1985. Translated by J. Claude Evans as *Philosophers in Exile: The Correspondence of Alfred Schutz and Aron Gurwitsch, 1939–1959*, edited by Richard Grathoff. Bloomington: Indiana University Press, 1989.

Esquisse de la phénoménologie constitutive, edited by José Huertas-Jourda. Paris: Vrin, 2002.

SECONDARY WORKS

Embree, Lester, ed. *Life-World and Consciousness*. Evanston, IL: Northwestern University Press, 1972.

Embree, Lester, ed. *The Phenomenology of Gurwitsch*. Special Issue, *The Journal of the British Society for Phenomenology* 12 (May 1981). Contains secondary bibliography until time of publication.

Evans, J. Claude, and Robert S. Stufflebeam, eds. *To Work at the Foundations: Essays in Memory of Aron Gurwitsch*. Dordrecht, Netherlands: Kluwer, 1997. Contains further secondary bibliography.

Husserl Studies 19 (1) (2003). This issue is devoted to Gurwitsch's thought.

Other Work

Merleau-Ponty, Maurice. *Phenomenology of Perception*. London: Routledge & Kagan Paul, 1962.

William R. McKenna (2005)

HABERMAS, JÜRGEN

(1929–)

Jürgen Habermas, the German philosopher and leading representative of the Frankfurt school of critical theory, was born in Düsseldorf. After World War II he studied in Göttingen, Zürich, and Bonn, where he submitted a dissertation on Friedrich von Schelling in 1954. From 1955 to 1959 he was Theodor Adorno's assistant at the Institute for Social Research in Frankfurt. After habilitating at Marburg University in 1961, he taught philosophy and sociology at the universities of Heidelberg and Frankfurt before becoming codirector of the Max Planck Institute in Starnberg. In 1983 he returned to the University of Frankfurt, where he was professor of philosophy until his retirement in 1994.

Habermas's life and work have remained deeply influenced by the traumatic events of his youth under National Socialism. From the time of his involvement with the German student movement in the 1960s he has been one of Germany's most prominent public intellectuals, speaking out on a wide array of issues, from violations of civil liberties and the attempted "historicizing" of the Holocaust to immigration policy and the manner of German reunification.

Habermas's scholarly work, which aspires to a comprehensive critical theory of contemporary society, ranges across many of the humanities and social sciences. His early and influential *Strukturwandel der Öffentlichkeit* (1962) was a historical, sociological, and philosophical account of the emergence and transformation of the liberal public sphere as a forum for critical public discussion of matters of general concern. While the historical structures of that sphere reflected the particular constellations of interests that gave rise to it, the idea it claimed to embody, the idea of legitimating political authority through rational discussion and reasoned agreement, remains central to democratic theory. Habermas returned to these themes three decades later in *Faktizität und Geltung* (1992), where he applied the idea of justification by appeal to generally acceptable reasons to the deliberations of free and equal citizens in a constitutional democracy. The primary function of the system of basic rights, he argued, is to secure personal and political autonomy; and the key to the latter is the institutionalization of the public use of reason in the legal-political domain.

One might read Habermas's extensive writings in the intervening decades as a protracted examination of the cultural, psychological, and social preconditions of and

barriers to accomplishing this. The essays of the early 1960s, a number of which were collected in *Theorie und Praxis* (1963), introduced the idea of studying society as a historically developing whole for purposes of enlightening political consciousness and guiding political practice. The methodology and epistemology behind this approach were elaborated in the later 1960s in *Zur Logik der Sozialwissenschaften* (1967) and *Erkenntnis und Interesse* (1968). A principal target in both books was the neopositivist thesis of the unity of scientific method, particularly the claim that the logic of inquiry in the human sciences is basically the same as in the natural sciences. The former work started from an examination of the nature and role of *Verstehen* in social inquiry and argued that access to symbolically prestructured object domains calls for interpretive procedures designed to grasp the meanings on which social interactions turn. Intersubjective meanings constitutive of sociocultural lifeworlds can neither be wholly objectified, as positivism supposes, nor simply reappropriated, as hermeneutics proposes. Psychoanalysis suggests an alternative approach, in which explanatory and interpretive procedures are combined with a critique of ideology in a historically oriented theory with practical intent.

In *Erkenntnis und Interesse* Habermas undertook a historical and systematic study of "the prehistory of modern positivism" in an attempt to free the ideas of reason and rationality from what he regarded as a "scientistic misunderstanding." Tracing the development of the critique of knowledge from Immanuel Kant through German idealism to Karl Marx, and its transformation into the methodology of science in early positivism, he elaborated his own position in critical encounters with three classic but flawed attempts to overcome positivism from within methodology: Charles Sanders Peirce's reflections on natural science, Wilhelm Dilthey's on cultural inquiry, and Sigmund Freud's on self-reflection. In each case he examined the roots of cognition in life and argued for an internal connection of knowledge with "anthropologically deep-seated" human interests. A key feature of this "quasi-transcendental" theory of cognitive interests was the basic distinction between the interest in prediction and control of objectified processes and the interests in mutual understanding and distortion-free communication with speaking and acting subjects.

There followed a series of studies of basic structures of communication, organized as a three-tiered research program. The ground level consisted of a general theory of communication in natural languages, a "universal pragmatics," as Habermas called it. This served as the foundation for a general theory of socialization in the form of a developmental account of the acquisition of communicative competence. Building on both of these, Habermas sketched a theory of sociocultural evolution as the historical development of forms of communicative interaction and mutual understanding. These accounts of communication, socialization, and social evolution enabled him to anchor moral theory in the theory of social action. Arguing that our basic moral intuitions spring from something deeper and more universal than contingent features of particular traditions, his discourse ethics sought to reconstruct the intuitive grasp of the normative presuppositions of social interaction possessed by competent social actors generally.

The work of the 1960s and 1970s culminated in the monumental *Theorie des kommunikativen Handelns* (1981), in which Habermas developed a concept of communicative rationality freed from the subjectivistic and individualistic premises of modern social and political theory, together with a two-level concept of society that integrated the competing paradigms of "lifeworld" and "system." On this basis he then sketched a critical theory of modern society that focused on "the colonization of the lifeworld" by forces arising from the economy and the state: systemic mechanisms such as money and power drive processes of social integration and symbolic reproduction out of domains in which they cannot be replaced. The phenomena that Max Weber pointed to in his vision of an "iron cage" and that Marxists have dealt with in terms of "reification" arises from an ever-increasing "monetarization" and "bureaucratization" of lifeworld relations. This relentless attack on the communicative infrastructures of society can be contained, he argued, only by a countervailing expansion of the areas of life coordinated via communication, and in particular by the subordination of economic and administrative subsystems to decisions arrived at in open and critical public debate. Thus, the antidote to colonization is democratization, and the key to the latter is an effectively functioning cultural and political public sphere. What distinguishes this critique of modernity from the welter of counterenlightenment critiques during the last two centuries is Habermas's unflinching defense of enlightenment rationality—a defense, to be sure, that is itself informed by the critique of rationalism and that emphasizes the ongoing, unfinished character of the project of enlightenment.

See also Adorno, Theodor; Critical Theory; Democracy; Dilthey, Wilhelm; Discourse Ethics; Enlightenment; Freud, Sigmund; Hermeneutics; Holocaust; Kant,

Immanuel; Marx, Karl; Peirce, Charles Sanders; Positivism; Rationalism; Reason; Rights; Weber, Max.

Bibliography

WORKS BY HABERMAS

Knowledge and Human Interests. Translated by J. Shapiro. Boston: Beacon Press, 1971.

Theory and Practice. Translated by J. Viertel. Boston: Beacon Press, 1973.

Communication and the Evolution of Society. Translated by T. McCarthy. Boston: Beacon Press, 1979.

The Theory of Communicative Action. 2 vols. Translated by T. McCarthy. Boston: Beacon Press, 1984, 1987.

Autonomy and Solidarity: Interviews with Jürgen Habermas. Edited by Peter Dews. London, 1986.

The Structural Transformation of the Public Sphere. Translated by T. Burger and F. Lawrence. Cambridge, MA: MIT Press, 1989.

Moral Consciousness and Communicative Action. Translated by C. Lenhardt and S. Nicholsen. Cambridge, MA: MIT Press, 1990.

Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy. Translated by W. Rehg. Cambridge, MA: MIT Press, 1995.

WORKS ON HABERMAS

Bernstein, R., ed. *Habermas and Modernity.* Cambridge, MA: MIT Press, 1985.

McCarthy, T. *The Critical Theory of Jürgen Habermas.* Cambridge, MA: MIT Press, 1978.

Rehg, W. *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas.* Berkeley: University of California Press, 1994.

White, S., ed. *The Cambridge Companion to Habermas.* Cambridge, U.K.: Cambridge University Press, 1995.

Thomas McCarthy (1996)

HAECCEITISM

See *Modality, Philosophy and Metaphysics of*

HAECKEL, ERNST HEINRICH

(1834–1919)

Ernst Heinrich Haeckel, the German zoologist and monist philosopher, was born in Potsdam. He studied medicine and science at Würzburg, Berlin, and Vienna with such authorities as Johannes Müller, Rudolf Virchow, and R. A. Kölliker. After practicing medicine for a short time, he went to the University of Jena in 1862 to teach zoology.

Haeckel was the first noted German biologist to grant enthusiastic acceptance to organic evolution, and Charles Darwin gave him credit for propagating the theory of evolution in Germany. His views were the source of considerable controversy in biology, philosophy, and religion. He battled with his colleagues about their early hostility to Darwin's theory and their reluctance to include man and his consciousness in the evolutionary process. His dislike of the power of the church in social and political matters and his liberal opposition to Otto von Bismarck and other political figures resulted in many controversies; his rejection of free will, immortality, and the personality of God also antagonized many. Haeckel's achievements in zoology brought him academic offers from famous institutions, but he chose to remain at Jena, partly because of the academic freedom he found there.

His interests were broad; he published travel works and illustrated some of his own scientific essays. He founded the Monistic League to propagate his religious views. He had considerable popular success in science and was prominent in the movement to enlighten humankind about scientific developments.

SCIENTIFIC CONTRIBUTIONS

In biology Haeckel helped to publicize and promulgate what he called the "biogenetic law": "*Ontogenesis is a brief and rapid recapitulation of phylogenesis,* determined by the physiological functions of heredity (generation) and adaptation (maintenance)" (*The Riddle of the Universe,* New York, 1900, p. 81). He was a pioneer in drawing up genealogical schemata of the relationships between various orders of animals. Many of his major groupings are still accepted, although the finer divisions have undergone much revision. He was convinced of the essential unity of organic and inorganic nature, and argued that the simplest protoplasmic substances arose from inorganic carbonates through spontaneous generation. Individual primitive organisms, which Haeckel termed "monera," were differentiated out of these protoplasmic compounds. Haeckel believed that to reject this kind of spontaneous generation was tantamount to accepting a miraculous origin of life.

His theory of *gastraea* also received much attention. Haeckel argued that the entire animal world is made up of two groups: primitive unicellular animals, the protozoa, and multicellular animals with complex tissues, the metazoa. Haeckel believed that all the metazoa evolved, in accord with his biogenetic law, from one simple, long-extinct form, the *gastraea.* Although this theory of evolu-

tion from *gastraea* is no longer accepted, it influenced embryological research for nearly half a century.

DOCTRINE OF SUBSTANCE

Haeckel's conviction on the great importance of organic evolution led him into many other fields. His *Die Welträthsel* (Bonn, 1899; English translation by Joseph McCabe, *The Riddle of the Universe*) became a best-seller. The title derived from Emil Du Bois-Reymond's 1880 address to the Berlin Academy of Sciences on seven "world enigmas" (the nature of matter and force, the origin of motion, the origin of life, the order in nature, the origin of simple sensation and consciousness, rational thought and speech, and freedom of the will). Haeckel believed that his monistic outlook could resolve these problems, and others, leaving one "comprehensive riddle," the problem of substance. He insisted upon the essential unity of all substance, but also insisted that the "real character" of substance was as little understood as in the days of Anaximander and Empedocles. Indeed, it became "more mysterious and enigmatic" as more and more became known about its attributes and their evolutionary forms. Haeckel was especially opposed to theological dualism, but he also carefully distinguished his view from both materialistic and idealistic monisms.

Haeckel construed materialism as holding that atoms are "dead," and are moved only by external forces. He maintained instead that both matter and ether possess sensation and will in the lowest grade. They experience a dislike of strain, and struggle against it, and a liking of "condensation," for which they strive. Haeckel denied the existence of empty space and of action at a distance. Those parts of space not occupied with ponderable atoms are filled with ether; action is either the result of immediate contact or occurs through the mediation of ether.

On the other hand, Haeckel rejected any attempt to regard the world as immaterial or nonnatural. Infinitely extended matter and sensitive and thinking spirit, or energy, are two fundamental attributes of the all-embracing universal substance. Every living cell has psychic properties, and multicellular organisms have as their psychic functions the totality of the psychic properties of their parts. Although Haeckel insisted that his view of substance was Spinoza-like rather than materialistic, many of his specific views are similar to those of nineteenth-century materialism. His confidence that "consciousness, thought, and speculation" are "functions of the ganglionic cells of the cortex of the brain," his "hard" determinism, his mechanism, his complete rejection of the supernatural, and his enthusiasm for science all inclined his contemporaries to classify him as a materialist.

Haeckel, then, saw the world as an eternal evolution of substance, and man as part of that evolution. The "law of substance," a law of mechanical causality, established "the eternal persistence of matter and force, their unvarying constancy throughout the entire universe" (*The Riddle of the Universe,* p. 4). He regarded the laws of the conservation of energy and the conservation of matter as inseparable and as parts of his law of substance. Haeckel referred to "great eternal iron laws," and rejected all teleological views. The appearance of design in the world is a consequence of natural selection rather than of the action of a purposive agency.

Although Haeckel often emphasized the tentative nature of scientific conclusions and the necessity for the modification and improvement of hypotheses, on some issues he assumed the finality of certain scientific propositions, including many rejected today. He made this assumption most frequently in his polemics against the philosophic and religious views that he regarded as incompatible with science. Haeckel did not generally consider in detail the technical problems philosophers were debating, but tended rather to attack or to defend the conclusions of technical philosophers on the basis of the scientific results of his day and extrapolations from them.

THEORY OF KNOWLEDGE

Despite his insistence that much of philosophy was far too speculative and a priori, Haeckel held that both empiricism and rationalism are necessary to develop satisfactory knowledge. Although he was hostile to "pure metaphysics," he was also critical of those who advocated a "pure empiricism." The opposition between experimental science and philosophy must and can be overcome.

Haeckel held that the thing-in-itself lying behind knowable phenomena is unknown. He suggested that we need not trouble about this situation; we have no means for investigating the thing-in-itself, and are not even certain it exists. The only genuine knowledge is knowledge of nature, and it consists of "presentations" (combinations of sense impressions in the knowing subject) corresponding to external things. Comparative and critical observation tells us that normally the impressions received by the brain and sense organs from the outer world are the same for all rational people, and that normally the same presentations are formed. Those presentations are true that correspond to the knowable aspects of things, even though things-in-themselves" cannot be reached.

Haeckel's views on knowledge were closely connected to biological findings. He argued that the human sense activity, which forms the beginning of all knowledge, was slowly and gradually evolved from the other primates. The sense organs of all primates are structurally similar, and Haeckel insisted that these organs also function similarly, in a way describable by the same chemico-physical laws. The rod-shaped cells in the retina, the auditory cells in the ear, the olfactory cells in the nose, and the taste cells on the tongue were evolved from simple, undifferentiated cells of the skin. Invoking his biogenetic law, Haeckel concluded that man's higher sense organs were derived from the epidermis of lower animals. Our sense impressions are associated in the cortex of the brain so that isolated elements are united into integrated wholes. Haeckel called these integrated presentations "faith in the broad sense," because they go beyond our sense impressions. In this sense, science requires faith in the construction of both hypotheses and unifying theories. (In the main, Haeckel used "theory" to refer to hypotheses about a common cause for diverse phenomena.) However, he rejected religious faith, which he termed "faith in the narrower sense." He insisted that religious belief always means a belief in miracles and thus contradicts the "natural faith of reason." Even religious liberals, he contended, are forced into the acceptance of superstition, and their faith is no less irrational than the "crude spirit-faith of primitive fetichism."

PSYCHOLOGY

Haeckel attempted a scientific account of the soul. He regarded it as a natural phenomenon, so that psychology was a natural science, a part of physiology. Psychology was the "foundation and the postulate" of all the sciences, since knowledge of nature is "part of the life of the soul." The great difficulty in establishing a naturalistic psychology is that such a science presupposes a thorough knowledge of the human organism, especially the brain. Haeckel deplored the lack of biological training of the psychologists of his time. He insisted that psychic processes, like all others, are subject to the law of substance, and held that the prevalence of mind-body dualisms in psychology has led to a greater confusion of ideas there than in any other department of knowledge. Yet Haeckel did not insist on a nonintrospective psychology; he described the introspective method as "extremely valuable and indispensable." But it had to be supplemented by experimental methods.

Haeckel regarded consciousness as the "central mystery of psychology," and the citadel of all mystical and dualistic errors. He insisted that consciousness is a natural phenomenon, dependent upon a material substratum. He suggested that consciousness can perhaps best be conceived as "internal perception" and can be compared to the action of a mirror. The chief difficulty in the way of a scientific understanding of consciousness is that the subject and the object of knowledge are one and the same; our only source of knowledge of consciousness is consciousness itself. We can therefore only know the consciousness of others by comparing it with our own. This works rather well when the comparison is made between normal people, but the analogy may break down badly when a comparison is made between the normal and the abnormal, or between different evolutionary levels. However, the difference between the consciousness of humans and of other animals is a difference of degree only, not of kind. Haeckel thought it probable that consciousness arises with the centralization of the nervous system, and that the lower classes of animals lack that faculty. The province of unconscious psychic actions, reflex action, for example, is more extensive than that of conscious ones, but the two areas are closely connected.

The consciousness of man and of other mammals biologically close to man is changeable and is modified by both internal and external causes. Consciousness is dependent upon the normal development of certain organs and gradually develops in the child as those organs develop. Despite Haeckel's use of a "faculties" terminology, his views on psychology are often similar to those of recent functionalists.

ATTACK ON TRADITIONAL RELIGION

Haeckel's attack on supernatural religions had many facets. He unequivocally rejected revelation and theological faith. He was outspoken in combating the superstitions associated with the world's great religions. He scathingly attacked the influence of the church as an institution in politics and education. Indeed, he frequently coupled these problems, holding that the German government would not improve until it was free from church influence and its citizens received a better, more scientific, education. Haeckel even claimed that such questions as whether a monarchy is preferable to a republic and whether the constitution should be aristocratic or democratic are subordinate to the "supreme question": Shall the government be secular or dominated by the clergy? Haeckel was no respecter of religious heroes, prominent clerics, sacred myths, or widely held dogma. He tried to show that theological beliefs are

incompatible with scientific data, unreasonable, or merely dogmatic.

Haeckel's admiration for the views of Benedict Spinoza and Johann Wolfgang von Goethe and his belief that humankind's ethical aspirations needed some support led him to advocate a monistic religion. "The ethical craving of our emotion is satisfied by monism no less than the logical demand for causality on the part of reason." He had great respect for the ethical values of primitive Christianity, and felt that Christianity had been so influential in the social and political movements of civilized history that "we must appeal as much as possible to its existing institutions in the establishment of our monistic religion" (*The Riddle of the Universe,* p. 336). He maintained, therefore, that he sought a rational reformation, rather than a revolution, in religion. However, the extent of his criticisms of Christianity appear to be revolutionary.

Haeckel wanted to give rational support to the true, the good, and the beautiful, and he considered the relation of that trinity to prevailing Christian notions. Truth is to be found in the study of nature by means of critical observation and reflection, and hence revelation must be rejected. However, what "we call virtue, in our monistic religion coincides for the most part with the Christian idea of virtue," especially the Christianity of the first three centuries. Charity, toleration, compassion, and assistance are humanistic as well as Christian precepts, and are to be emphasized in the monistic religion.

On the other hand, Haeckel maintained that early Christianity preached the valuelessness of this-worldly things, because this life was merely a preparation for eternity. Hence the beautiful was of little consequence. Haeckel was especially interested in art forms in nature and believed that the microscope had newly aroused our aesthetic sense.

All forms of theism are to be opposed. A pantheism that identifies God and substance is necessarily "the world-system of the modern scientist." All scientists who think theism can be reconciled with science are, in Haeckel's view, either dishonest, or confused, or victims of sophistry. If atheism is construed as a denial of the existence of a personal and extramundane god or gods, then Haeckel agreed with Arthur Schopenhauer's remark that pantheism is only a polite form of atheism. In short, Haeckel's criticism of traditional religions was that their doctrines are often intellectually wrong; that they generate unrealistic hopes; and that the social, political, and educational consequences of supernaturalism are malignant. Haeckel's criticisms, especially of Roman Catholicism, are often strongly worded. Thus he wrote that the obligatory celibacy of the clergy, auricular confession, and the sale of indulgences were designed for the purposes of strengthening the rule of the church over the "credulous masses and making as much material profit as possible out of them."

ETHICS AND SOCIAL VIEWS

In ethics, Haeckel felt that traditional theories often either emphasize altruism too much (as in the case of many religious views) or emphasize egoism too much (hedonisms). He held that there should be an "equal emphasis" on self-love and love of one's neighbor. The "highest aim of all ethics" is to reestablish a "natural equality" of egoism and altruism. Along with this should go an emphasis on the body as well as the soul; an emphasis on fair treatment of animals as well as humans. Haeckel believed that a recognition of human evolution would incline us to be more sympathetic to animals, and that Christian attitudes easily lead to cruelty toward animals. Haeckel regarded the family as the foundation of society and as a necessity for humanity as well as for the higher social animals, whereas Christianity, he believed, tends to disparage the family as a this-worldly phenomenon. Haeckel also opposed the tendency that he found in Christianity to make woman subordinate to man and to regard sexual intercourse as "unclean." He was especially hostile to the hypocrisy he believed is often found in the church toward sex.

Haeckel was much interested in social reform, holding that progress is a law of nature. He compared the rapid progress made in the natural sciences with the lack of progress in government, the administration of justice, education, and social and moral organization. He gave special attention to justice. He believed that students of jurisprudence need much more education in science than they usually receive, and that their knowledge of human nature is sadly deficient. Politicians too make practical decisions of great import with no scientific grounding in the appropriate areas. He also decried the many impediments to free inquiry, whether they stem from political reaction or from theological superstition. He was highly optimistic about the consequences of an improved system of education.

Many of Haeckel's views that caused violent disputes in the past are today widely accepted by educated people. Much of the antagonism toward him was centered on his insistence that man is a part of nature and in the evolutionary stream. Although large portions of the scientific part of Haeckel's worldview have since been rejected, much is still regarded as sound. His views on religion

would still be challenged by many; some, of course, find them mild.

See also Altruism; Darwin, Charles Robert; Empiricism; Evolutionary Theory; Goethe, Johann Wolfgang von; Hedonism; Justice; Materialism; Progress, The Idea of; Psychology; Rationalism; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de.

Bibliography

ADDITIONAL WORKS BY HAECKEL

Generelle Morphologie der Organismen. 2 vols. Berlin: G. Reimer, 1866.

Natürliche Schöpfungsgeschichte. Berlin, 1868. Translated into English by E. R. Lankester as *The History of Creation.* London: H.S. King, 1876.

Anthropogenie. Leipzig, 1874. Translated into English by Joseph McCabe as *The Evolution of Man.* New York: Putnam's Sons, 1910.

Freie Wissenschaft und freie Lehre. Stuttgart: E. Schweizerbart, 1878. Translated into English as *Freedom in Science and Teaching.* New York: D. Appleton, 1879.

Der Monismus als Band zwischen Religion und Wissenschaft. Bonn: E. Strauss, 1893. Translated into English by J. Gilchrist as *Monism as Connecting Religion and Science: The Confession of Faith of a Man of Science.* London: A. and C. Black, 1895.

Die Lebenswunder. Stuttgart: A. Kröner, 1904. Translated into English by Joseph McCabe as *The Wonders of Life.* New York: Harper, 1905.

Der Kampf um den Entwickelungs-Gedanken. Berlin, 1905. Translated into English by Joseph McCabe as *Last Words on Evolution.* London: A. Owen, 1906.

WORKS ON HAECKEL

Bölsche, Wilhelm. *Ernst Haeckel: Ein Lebensbild.* Leipzig: H. Seeman, 1900. Translated into English by Joseph McCabe as *Haeckel: His Life and Work.* London: T.F. Unwin, 1906. Contains a useful bibliography to 1900 as well as detailed accounts of Haeckel's life and work. McCabe added a supplementary chapter, "The Crowning Years," which gives an account of the controversies in which Haeckel engaged after the publication of *The Riddle of the Universe.*

Merz, J. T. *A History of European Thought in the Nineteenth Century.* Edinburgh: Blackwood, 1903. Treats Haeckel and his contemporaries in some detail.

Perry, R. B. *Present Philosophical Tendencies.* New York: Longmans, Green, 1912. Gives a brief account of Haeckel and his contemporaries and also discusses the leading issues that engaged their interest.

Schmidt, Heinrich. *Ernst Haeckel, Leben und Werke.* Berlin, 1926.

Rollo Handy (1967)

HÄGERSTRÖM, AXEL
(1868–1939)

Axel Hägerström, the Swedish philosopher, was the son of an orthodox minister of the Swedish Lutheran church and grew up in an intensely religious atmosphere. With the intention of following his father's profession, in 1886 he began theological studies at Uppsala University, which was to remain his academic home throughout his life. His interests, however, were soon diverted from theology to philosophy. From 1893 to 1911 he was docent (roughly, assistant professor), and from 1911 to 1933 professor, of "practical philosophy" (philosophy of morals, law, and religion). During his student days the idealistic metaphysics of C. J. Boström was still influential in Uppsala, although this mode of thought was soon to be swept away by a kind of Neo-Kantianism. Hägerström's publications around the turn of the twentieth century mirror this situation. Under the influence of Immanuel Kant, he came to regard metaphysics as impossible and, going further than Kant, rejected the hypothesis of the *Ding an sich* (especially in *Kants Ethik,* Uppsala, 1902). Like Kant, he considered the pure Ego, the same in all individual minds, as somehow the principle of the reality given to us, as the source of the laws of logic, and also as the source of certain synthetic propositions a priori, such as the principle of causality. Gradually the role played by this pure Ego was taken over, in his thought, by "the concept of reality," which he treated in *Das Prinzip der Wissenschaft* (The principle of science; Uppsala, 1908) and *Botanisten och filosofen* (The botanist and the philosopher; Uppsala, 1910).

Speculation about the concept of reality was to remain a fundamental ingredient in his mature philosophy, but it gradually lost most of its original Kantian flavor. In 1909 Hägerström wrote his *Social teleologi i marxismen* (Social teleology in Marxism; Uppsala). Although this study is a sharp criticism of the Marxist philosophy of history, it seems evident that he was influenced by, or at least in strong sympathy with, certain other aspects of Marxism—its materialism and its views on the functions of ideologies. In his lectures Hägerström soon characterized his own outlook as "enlightened materialism." To give an adequate characterization of his philosophy in a few key words is difficult, for he himself never presented his views in a systematic fashion. His many philosophical writings are mostly devoted to rather special questions, and much of their space is taken by polemics against authors with whom he disagrees. The influences that molded his thought were diverse and seemingly somewhat incompatible. His final philosophi-

cal positions were, on the whole, as far to the left as possible of religion and of any philosophical system, such as that of Boström, that was akin to or gave support to religion. As his motto Hägerström once chose the Catonian paraphrase: "Besides, I think that metaphysics ought to be destroyed."

CRITIQUE OF "METAPHYSICS"

Like so many antimetaphysicians, Hägerström was prone to label any view opposed to his own as metaphysical. The word *metaphysics* as he used it remains somewhat vague as to connotation as well as to denotation. He held, however, that all metaphysical doctrines suffer from a common fundamental fault, that of (implicitly or explicitly) assuming that "reality itself is something real" (or "being is something that there is"). This assumption is as "absurd" as, for example, the assumption that triangularity is something triangular. Hägerström thought it possible to prove positively (1) that the spatiotemporal world of experience exists and (2) that nothing may exist outside this world. In his proof of (1) he made use of an "analysis of the concept of reality" and also of ideas reminiscent of René Descartes's *Cogito.* To deduce (2) from (1) he invoked the principle that two entities cannot exist "outside each other" except as parts of a spatiotemporal context. His materialistic conception of the world of experience does not exclude the existence of consciousness, but consciousness, in his opinion, is a quality of certain material bodies (the psychophysical organisms).

CRITIQUE OF "SUBJECTIVISM"

In an act of consciousness (awareness) we are always conscious of something. If *C* is a consciousness of *O,* then *C* and *O* are, according to Hägerström, always two distinct entities; and further, the fact that a consciousness of *O* exists does not imply that *O* is endowed with any special intrinsic quality (such as being "mental," being a "perception," or being an "idea"). To overlook this is, in his opinion, the fundamental "subjectivist" mistake, which he thinks he can trace in the majority of philosophical epistemologies. This mistake gives rise to a secondary "subjectivist" mistake, the assumption that our knowledge about our own acts of consciousness is the immediate knowledge from which our knowledge of the external world must be derived.

THEORY OF VALUE

Hägerström's first work in value theory was "Kritiska punkter i värdepsykologien" (Critical points in value psychology; in *Festskrift för E. O. Burman,* Uppsala, 1910), in which he raised objections to certain views of the Austrian school of value theory (Alexius Meinong, Christian von Ehrenfels, and others). He rejected especially their distinction between valuating emotive experiences and value judgments as theoretical judgments about the occurrence of such experiences. The value judgment, he claimed, is itself essentially emotive. By the time of his inaugural lecture, published as *Om Moraliska Föreställningars Sanning* (On the truth of moral ideas; Uppsala, 1911), Hägerström had arrived at the "value-nihilistic" doctrine that was to remain one of the most characteristic traits of his philosophic position. Statements of value, such as "To lie is bad," are neither true nor false: They lack truth value. Of the many arguments by which he tried to corroborate this view, the following is typical: A statement is true (or false) if, and only if, the judgment (as a mental phenomenon) expressed by the statement is true (or false); a statement of value, however, does not express any genuine judgment, but an "association" between an "idea" (for instance, the idea of lying) and an emotion. In his work *Till frågan om den objektiva rättens begrepp* (On the question of the notion of law; Uppsala, 1917), he elaborated this view also with respect to deontic statements. A statement such as "I ought not to lie," or "It is my duty not to lie," corresponds, not to a judgment with a truth value, but to an association between an "idea" and a "conative impulse." In this respect deontic statements are closely akin to imperatives. The persistent illusion that value statements and deontic statements have a truth value is caused by the relative stability of the underlying associations, which are built up and supported by the suggestive influence of a number of factors in the social system (such as early education by parents and teachers and the pressure of public opinion).

Hägerström nourished the hope that the spreading of his value-nihilism would contribute to the creation of a more tolerant, more humane, and less vindictive morality. Since he believed that the task of the moral philosopher is to analyze the mental phenomena expressed by, for example, statements of value, and since he took emotions to be an essential constituent in such phenomena, he became deeply interested in the nature of emotion. In order to substantiate his value-nihilism, he thought it important to demonstrate the subjective character of emotions; being "subjective," emotion can not be a source of knowledge—for instance, knowledge of values. Like all mental phenomena, an emotive experience is either an act of being conscious (aware) of something or some combination of such acts. What might commonly be called an emotive experience, such as enjoying the prospect of going to the cinema, is a combination of

intellectual and purely emotive ingredients. The purely emotive experiences, such as a mere feeling of pleasure, consist in being conscious of a certain emotive quality (here a pleasure quality). In his earlier publications, Hägerström seems inclined to regard emotion as "subjective" because emotive qualities are qualities of the Ego. Later he experimented with a variety of explanations. According to one, emotion is "subjective" because the emotive qualities are experienced without "localization." Here, Hägerström invokes his principle that localization in the spatiotemporal context is essential to reality and objectivity, but the form of his argument remains somewhat vague. According to another of his somewhat puzzling explanations, emotion is "subjective" because the emotive qualities inhere in the psychophysical organism that has the emotion, and not in "external" objects.

In some of his works he assumed a "projection of emotive qualities onto external objects." When I look at a painting that pleases me, in his opinion, I project the quality of pleasure experienced by myself onto the painting: I perceive the painting as pleasant, just as I perceive it as square or as dark. On this view, the epistemological distinction that he wished to maintain between emotive qualities and, say, colors, becomes rather problematical. Some Swedish critics of his value theory have taken this view as starting point for their criticism—Einar Tegen, "The Basic Problem in the Theory of Value," in *Theoria* 10 (1944): 28–52; and Søren Halldén, *Emotive Propositions* (Stockholm, 1952).

LEGAL PHILOSOPHY

Hägerström began his mature work in legal philosophy with a criticism of a doctrine, common in nineteenth-century "legal positivism" (*Rechtspositivismus*), according to which "positive law" (as opposed to "natural law," *Naturrecht*) is somehow the expression of a will actually existing in society. His essay "Är gällande rätt uttryck av vilja?" (Is positive law an expression of will?) in *Festkrift tillägnad Vitalis Norström* (Göteborg, 1916) and his previously mentioned book *Till frågan om den objektiva rättens begrepp* are largely devoted to a painstaking criticism of this doctrine in its many varieties. Hägerström devoted much energy to the attempt to clarify the nature of positive law and those factors in "the social machinery" that uphold the law. He maintained that our common view of legal phenomena is blurred by "magical ideas" that can be traced far back in history. In *Der römische Obligationsbegriff im Lichte der allgemeinen römischen Rechtsanschauung* (The Roman notion of obligation in the light of the general Roman view of law; Vol. I, Uppsala, 1927; Vol. II, Uppsala, 1941), he tried to demonstrate the magical element in ancient Roman law. He believed that such Roman concepts as *ius, dominium,* and *possessio* are magical ideas and that the old Roman legal acts, such as *mancipatio* and *stipulatio* are acts through which magical powers over things or persons are established.

CRITICAL HISTORY OF IDEAS

In his lectures (some of which have been posthumously published) Hägerström discussed, with a wealth of learning, the history of religious, philosophical, political, and legal ideas. The history of ideas appeared to him largely as the history of confusions and errors flowing from certain inborn mechanisms of the human mind. To explain them he used to point especially to certain thought processes that, in his opinion, almost inevitably take place when the emotions and the projection of emotive qualities interfere with intellectual operations.

INFLUENCE

In Sweden, and also in the neighboring Scandinavian countries, Hägerström has exercised great influence. With his pupil and colleague Adolf Phalén he became the founder of the so-called Uppsala school of philosophy, which flourished in the 1920s and 1930s and has had a lasting effect on the whole academic philosophical atmosphere in Sweden. Common to the members of this school—most of whom disagreed with much of Hägerström's own philosophy—were a distrust of metaphysical speculation and of epistemological subjectivism, a realistic (sometimes almost naively realistic) conception of the external world, an interest in the phenomenological analysis of mental acts and their contents, an emotive theory of value (of some kind or another), and an insistence on conceptual analysis as a central task of philosophy. Some of the original members of the Uppsala school became strongly influenced by the Cambridge school of analysis in England and by logical empiricism. Outside philosophy proper, Hägerström gave rise to a school of jurisprudence (Vilhelm Lundstedt, Karl Olivecrona, Alf Ross).

See also Appearance and Reality; Boström, Christopher Jacob; Ehrenfels, Christian Freiherr von; Kant, Immanuel; Marxism; Meinong, Alexius; Value and Valuation.

Bibliography

ADDITIONAL PRIMARY SOURCES

"Axel Hägerström." In *Die Philosophie der Gegenwart in Selbstdarstellungen.* Vol. 7. Leipzig: F. Meiner, 1929.

Socialfilosofiska uppsatser (Essays in social philosophy). Stockholm: Bonnier, 1939.

Religionsfilosofi (Philosophy of religion). Stockholm, 1946.

Moralpsykologi (Psychology of morals). Stockholm, 1952.

Inquiries into the Nature of Law and Morals. Translated by C. D. Broad. Uppsala: Almqvist and Wiksells, 1953. Translation of some of Hägerström's works.

Filosofi och vetenskap (Philosophy and science). Stockholm: Ehlins, 1957.

Rätten och staten (The law and the state). Stockholm: Almqvist and Wiksells, 1963.

Recht, Pflicht und bindende Kraft des Vertrages nach römischer und naturrechtlicher Anschauung. Stockholm: Almqvist & Wiksell, 1965.

Moralfilosofins grundläggning. Edited by Thomas Mautner. Stockholm: Almqvist & Wiksell, 1987.

Philosophy and Religion. London: Routledge, 2002.

SECONDARY SOURCES IN GERMAN AND ENGLISH

Broad, C. D. "Hägerström's Account of Sense of Duty and Certain Allied Experiences." *Philosophy* 26 (1951): 99–113.

Cassirer, Ernst. *Axel Hägerström.* Göteborg, Sweden: Göteborgs Högskolas Arsskrift XLV, 1939.

Ofstad, H. "Objectivity of Norms and Value-Judgments According to Recent Scandinavian Philosophy." *Philosophy and Phenomenological Research* 12 (1951): 42–68.

Passmore, J. "Hägerström's Philosophy of Law." *Philosophy* 36 (1961): 143–160.

Sandin, R. T. "The Founding of the Uppsala School." *Journal of the History of Ideas* 23 (1962).

SECONDARY SOURCES IN SWEDISH

Hedenius, I. *Om rätt och moral* (On law and morals). Stockholm, 1941.

Lang, Dieter, and Axel Hägerström. *Wertung und Erkenntnis: Untersuchungen zu Axel Hägerströms Moraltheorie*. Amsterdam: Rodopi, 1981.

Marc-Wogau, Konrad. "Axel Hägerström och kritiken av subjektivismen" (Axel Hägerström and the critique of subjectivism). In *Harald Nordensson 60 år.* Stockholm, 1946.

Marc-Wogau, Konrad. "Axel Hägerströms verklighetstheori" (Axel Hägerström's theory of reality). *Tiden* (1940): 286–299, 360–366.

Marc-Wogau, Konrad. "Känsla och värde enligt Axel Hägerström och Hans Larsson" (Emotion and value according to Axel Hägerström and Hans Larsson). *Ord och bild* 55 (1946): 245–260.

Marc-Wogau, Konrad. *Studier till Axel Hägerströms filosofi*. Stockholm, Prisma, 1968.

Marc-Wogau, Konrad. "Uppsalafilosofien och den logiska empirismen" (The Uppsala philosophy and logical empiricism). *Ord och bild* 53 (1944): 30–38.

Mautner, Thomas. *Vägledning till Hägerströmstudiet*. Stockholm: Almqvist & Wiksell, 1994.

Oxenstierna, G. *Vad är Uppsala-filosofien* (What is the Uppsala philosophy?). Stockholm, 1938.

Petersson, Bo. *Axel Hägerströms värdeteori*. Uppsala, Filosofiska f`reningen och Filosofiska institutionen vid Uppsala universitet, 1973.

Waller, Margit. *Axel Hägerström, Mannisker som fa Kände* (Axel Hägerström, the man whom few knew). Stockholm: Natur and Kultur, 1961.

A. Wedberg (1967)

HALEVI, YEHUDA

(c. 1075–1141)

Yehuda Halevi, or Judah ha-Levi, the philosopher of Judaism, was born in Toledo, Spain. In his youth he received an excellent grounding in biblical and rabbinic literature, as well as in the secular, particularly philosophic, disciplines. Halevi early displayed a marked poetic gift, which culminated in a body of Hebrew poetry noted for its adaptation of Arabic poetic forms to the Hebrew idiom and for its religious profundity. He practiced the profession of medicine for most of his life, residing in both Christian and Muslim Spain, a fact that may account for his excellent knowledge of Judaism's two descendant religions. His decision to leave for a perilous pilgrimage to Palestine was the result of his intense longing to see the Holy Land, a longing that is reflected in both his poetry and his philosophic work. Legend has it that he was killed in 1141 by an Arab horseman as he kissed the soil of the Temple Mount in Jerusalem, but there is no historic confirmation of this, since he cannot be traced beyond Egypt on his way to Palestine.

Yehuda Halevi's philosophic work *Kuzari: The Book of Proof and Argument in Defense of the Despised Faith*, written shortly before his departure for Palestine, uses as its framework the historically verified conversion to Judaism of the Khazar King Bulan II and most of his people about the year 740. This event had assumed almost legendary proportions by Yehuda Halevi's time, serving as a source of great religious satisfaction to the otherwise badly suppressed Jewish masses. In his work Yehuda Halevi reconstructed imaginatively the discussions that led to the king's conversion. At the beginning we are told that an angel has appeared to the king in his sleep and has informed him that the Creator was pleased with his intentions but not with his way of acting. In the hope of learning a better way of life the king calls in representatives of Aristotelian philosophy, Christianity, and Islam, but they all fail to satisfy him. The king did not originally plan to call on a representative of Judaism, judging this religion unworthy of serious consideration because of the misery of its adherents, but his dissatisfaction with the other presentations causes him to alter his decision and to

call for a rabbi. The discussion with the rabbi constitutes the rest of the volume.

The rabbi begins his presentation by asserting his belief in the God of Abraham, Isaac, and Jacob, who led the Israelites out of Egypt and whose intervention in the history of Israel has been continuous ever since. By beginning in this way, Yehuda Halevi broke sharply with the tradition of Aristotelian rationalism that characterized the bulk of medieval Jewish philosophy. He was very much aware of the profound abyss separating the God of the philosophers, who is self-contained, unmoved, and nonpersonal, from the personal and historic God of the Bible. For this reason he dispensed entirely with traditional proofs for the existence of God, the usual prolegomena of medieval Aristotelianism—whether Jewish, Christian, or Muslim—to the defense of the faith. For Yehuda Halevi it was history that was decisive. The God who reveals himself in the history of Israel could not have been reached by philosophical speculation but only by revelation. Similarly, Yehuda Halevi's interest in miracles reflected his view of history as the realm in which revelation takes place. The superiority of Judaism over its competitors follows, for Yehuda Halevi, from the public character of the Sinaitic revelation upon which Judaism is based. At Sinai, 600,000 men, women, and children were addressed by God, a mass revelation that no other religion can claim. This precludes the possibility of error or deception, a possibility that cannot be discounted in those instances where the revelation is restricted to one or to a few.

Yehuda Halevi's attitude toward the knotty problem of anthropomorphism also reflected his anti-Aristotelian orientation. Although he was not in sympathy with a literal interpretation of many of the terms applied to God by the biblical authors, realizing that this would lead to a humanization of God even to the extent of attributing corporeality to him, Yehuda Halevi was not willing to go the other extreme and strip God of all attributes, making it all but impossible to speak about him. There are events that can be experienced as proceeding from God directly. When biblical authors, such as the prophets, applied a term like *merciful* to God, they were referring to those actions of God that are experienced by man as merciful and as coming from God. Although the term *merciful* is therefore applicable more to the effects of God's actions than to his essence, to the extent that his actions are his discourse about God becomes possible.

The religious particularism that is fundamental to biblical religion was no source of embarrassment to Yehuda Halevi. The election of the people of Israel and of the land of Israel were fundamental concepts of his religious nationalism. This nationalism was based on the divine election of a people and a land for the proclamation to all humankind of those demands that God makes of all people, but for the special representation of which he has chosen one nation, whose suffering derives from its unfaithfulness to its mission. These themes permeate Yehuda Halevi's poetic works as much as they do his philosophical writings.

See also Aristotelianism; Nationalism; Rationalism.

Bibliography

WORKS BY YEHUDA HALEVI

The *Kuzari* or *Book of Refutation and Proof on Behalf of the Despised Religion* is Yehuda Halevi's magnum opus on religion, philosophy, and theology in general and Judaism in particular. The first critical edition of this work was based on Yehuda ibn Tibbon's medieval Hebrew translation and published together with a modern German translation and commentary by David Cassel, *Das Buch Kusari des Abū-l-Hasan Jehuda Hallewi*, Zweite verbesserte Auflage (Leipzig: Friedrich Voigt, 1869). A revision of this Hebrew translation with useful notes can be found in A. Zifroni, *Sefer ha-Kuzari* (Tel Aviv: Mahbarot le-Sifrut, 1948). The original Judeo-Arabic text was first edited by Hartwig Hirschfeld and published with Ibn Tibbon's translation as *Das Buch al-Chazari* (Leipzig: Otto Schulze, 1887). A new and much more comprehensive critical edition of the original Judeo-Arabic text has been published by David H. Baneth and Haggai Ben Shammai as *Kitāb al-Radd wa' l-Dalīl fī'l Dīn al-Dhalīl (al-Kitāb al-Khazarī)* (Jerusalem: Magnes Press, 1977); this edition is now the basis of all contemporary scientific translations. Hirschfeld eventually translated his Judeo-Arabic edition into English as the *Book of the Kuzari* (London: Routledge, 1905); this translation has been republished several times. Isaac Heinemann published a much abridged and revised English translation with notes, based on Hirschfeld's work, which originally appeared in *Three Jewish Philosophers*, edited by Hans Lewy, Alexander Altmann, and Isaac Heinemann (New York, 1945); this translation has likewise been republished several times. Modern Hebrew translations with notes include: Yehudah Even-Shmuel, *Sefer ha-Kozari le-R. Yehudah Halevi*, 2nd ed. (Tel Aviv: D'vir, 1972); and R. Joseph Qafih's combined edition and translation, *Sefer ha-Kuzari: Maqor Ve-Tirgum* (Qiryat Ono: Machon Mishnat Ha-Rambam, 1997), which is very close to the original Judeo-Arabic. Equally valuable is Charles Touati's French translation, *Le Kuzari: Apologie de la religion méprisée* (Louvain and Paris: Peeters, 1994). A new English translation of the *Kuzari* with notes, begun by the late Lawrence V. Berman, is being revised, continued, and prepared for publication by Barry S. Kogan under the auspices of the Yale Judaica Series of Yale University Press.

WORKS ON YEHUDA HALEVI

The single most valuable resource on Yehuda Halevi's life is Moshe Gil and Ezra Fleischer's *Yehudah Ha-Levi and His Circle* (in Hebrew; Jerusalem: World Union of Jewish Studies, 2001). A short biographical portrait, in English, reflecting recent research, is S. D. Goitein's "Judah Halevi: Poet Laureate, Religious Thinker, Communal Leader, Physician," in *A Mediterranean Society*, Vol. 5: *The Individual* (Berkeley: University of California Press, 1988). Book-length studies of Halevi's thought include Yochanan Silman, *Philosopher and Prophet: Judah Halevi, the Kuzari, and the Development of His Thought* (Albany: State University of New York Press, 1995); and Diana Lobel, *Between Mysticism and Philosophy: Sufi Language of Religious Experience in Judah Ha-Levi's Kuzari* (Albany: State University of New York Press, 2000). All chapter-length studies of Halevi in standard histories of Jewish philosophy remain useful: Isaac Husik, *A History of Medieval Jewish Philosophy* (New York: Macmillan, 1916); Julius Guttmann, *Philosophies of Judaism*, translated by D. W. Silverman (Philadelphia: Jewish Publication Society, 1964); Colette Sirat, *A History of Jewish Philosophy in the Middle Ages* (Cambridge, U.K.: Cambridge University Press, 1985); and L. E. Goodman, "Judah Halevi," in *History of Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman (London: Routledge, 1997). Articles and chapter-length studies on specific topics include: D. H. Baneth, "Judah Halevi and al-Ghazali," in *Studies in Jewish Thought: An Anthology of German Jewish Scholarship*, edited by A. Jospe (Detroit: Wayne State University Press, 1981; original Hebrew version with notes in *Knesset* 7 [1941–1942]: 311–329); Michael S. Berger, "Toward a New Understanding of Judah Halevi's *Kuzari*," in *Journal of Religion* 72 (1992): 210–228; Herbert A. Davidson, "Reverberations [of the Theories of Alfarabi and Avicenna] in Medieval Jewish Philosophy," in *Alfarabi, Avicenna, and Averroës on Intellect: Their Cosmologies, Theories of the Active Intellect, and Theories of Human Intellect* (Oxford: Oxford University Press, 1992); Kenneth H. Green, "Religion, Philosophy, and Morality: How Leo Strauss Read Judah Halevi's *Kuzari*," in *Journal of the American Academy of Religion* 61 (1993): 225–273; Barry S. Kogan, "Al-Ghazali and Halevi on Philosophy and the Philosophers," in *Medieval Philosophy and the Classical Tradition*, edited by John Inglis (Richmond, U.K.: Curzon, 2002); Barry S. Kogan, "Judah Halevi and His Use of Philosophy in the *Kuzari*," in *The Cambridge Companion to Medieval Jewish Philosophy*, edited by Daniel H. Frank and Oliver Leaman (Cambridge, U.K.: Cambridge University Press, 2003); Tsvi Langerman, "Science in the *Kuzari*," *Science in Context* 10 (1997): 495–522; Diana Lobel, "A Dwelling Place for the *Shekhinah*," *Jewish Quarterly Review* 90 (1999): 103–125; Shlomo Pines, "Shiʿite Terms and Conceptions in Judah Halevi's *Kuzari*," *Jerusalem Studies in Arabic and Islam* 2 (1980): 165–251; Eliezer Schweid, "The Artistry of the Dialogue within the *Kuzari* and Its Theoretical Meaning," in *Feeling and Speculation* (in Hebrew; Ramat Gan: Masada, 1970); Leo Strauss, "The Law of Reason in the *Kuzari*," *Persecution and the Art of Writing* (Glencoe, IL: Free Press, 1952); Harry A. Wolfson, "Maimonides and Hallevi: A Study in Typical Jewish Attitudes towards Greek Philosophy in the Middle Ages" and "Maimonides and Hallevi on Prophecy," in *Studies in the History of Philosophy and Religion*, Vol. 2, edited by Isadore Twersky and George H. Williams (Cambridge, MA: Harvard University Press, 1977.

Michael Wyschogrod (1967)

Bibliography updated by Barry Kogan (2005)

HAMANN, JOHANN GEORG

(1730–1788)

Johann Georg Hamann, the German Protestant thinker and critic of the Enlightenment, was born in Königsberg. In no sense a professional philosopher, and largely self-educated, he made his living as a secretary-translator and later as a government warehouse manager in Königsberg.

WRITINGS

Hamann's originality early caught the eye of such diverse figures as Johann Wolfgang von Goethe, G. W. F. Hegel, and Søren Kierkegaard, but his famous "darkness"—his opaque style—has daunted all but the most persistent investigators. Study of Hamann has long been dominated by Hegel's picture of him as an irrationalist and the paradigm of an individualist and also impeded by discouraging delays in the publication of complete editions of his works and letters. Following World War I, Hamann's influence on Kierkegaard began to be appreciated, but only more recently have scholars been able to expose enough of their subject so that the real dimensions of his thought could be guessed.

MEDITATIONS AND OCCASIONAL PIECES. Hamann's simplest writings were not intended for publication. These consist of his reflections following financial and spiritual crises he underwent on a business trip in 1758—*Biblische Betrachtungen* (Biblical meditations; 1758), *Gedanken über meinen Lebenslauf* (Thoughts on the course of my life; 1758–1759), and *Brocken* (Fragments; 1758). The *Sokratische Denkwürdigkeiten* (Socratic memorabilia; 1759) was his first public attack on the spirit of his century. A meditation on Socrates and his relation to Christ, it adumbrates the central concern of Hamann's intellectual career, the relation of philosophy to Christianity. Hamann saw himself as continuing the work of Martin Luther, under the different conditions of a later age. Whereas for Luther the problem had been the relation of faith to the "law," the established ecclesiastical and religious systems, the problem now concerned faith and philosophy.

Most of Hamann's writings were short occasional pieces. *Die Magi aus Morgenlande* (The Wise Men from the East; 1760), an essay on the symbolic meaning of eighteenth-century astronomical observations, earned him the sobriquet of the "Wise Man of the North." His reputation during his lifetime was largely based on such collections of essays as the *Kreuzzüge des Philologen* (Crusades of the philologian; 1762), which contains the influential "Aesthetica in Nuce" (Aesthetics in a nutshell), and on some political satires—*Lettre néologique et provinciale* (Neological and provincial letter; 1761), *Lettre perdue d'un sauvage du nord à un financier de Pe-kim* (Lost Letter of a savage of the north to a financier at Peking; 1773), and *Le kermes du nord* (The worm of the north; 1774). The sarcasm and irony characteristic of Hamann's style are readily apparent from some of his titles.

PHILOSOPHY AS CRITICISM

In what sense was Hamann a philosopher? Like Augustine, Anselm, Thomas Aquinas, Blaise Pascal, and Kierkegaard, Hamann is difficult to classify. His relation to philosophy was ambivalent and paradoxical. His thought moved between the twin figures of Socrates and the "Philologian." The figure of Socrates, the philosophical hero of the Enlightenment, Hamann adopted for his own purposes, to turn the symbol of the Enlightenment against itself and to call for a philosophical confession of ignorance in place of philosophical pretensions to knowledge. The term *Philologian* was selected for its ambiguity, in that it suggests both a "lover of the word" and a "lover of reason."

Like Socrates, Hamann considered man to be the crucial problem. (Hamann's famous simile compares self-knowledge to a "descent into hell," suggesting later explorations, by existentialism and depth psychology, into the anxieties and subconscious turmoil of the human psyche.) Like Socrates, he called for and practiced a critical and questioning philosophy; he especially appreciated the acid analyses of David Hume's reason. His answer to Immanuel Kant's criticism of metaphysics was a higher level of criticism, not "metaphysics" but "metacriticism." But as the Philologian, Hamann saw Socrates as a forerunner and prophet (although an unwitting one) of the Christ and philosophy as a discipline seen in its true light only in the context of Christianity.

CRITICISM OF THE ENLIGHTENMENT

Hamann's friends included many of the luminaries of the German Enlightenment, but personal relationships did not deter him from mounting the most severe criticism. (He believed friendship was like Mount Etna, "fire in the bowels" but "snow on the head.")

What were Hamann's objections to the philosophy of the Enlightenment? He viewed as "idealistic vanity" the attempts of leading Enlightenment thinkers to base philosophy upon undeniable rational truths (Moses Mendelssohn), to speak of "pure" reason (Kant), to discover a "natural religion" (the deists), to penetrate the mystery of man's constitution and isolate the origin of man's linguistic capacities (Johann Gottfried Herder and others), and to separate the knowledge of God from its provenance in historical revelation (Gotthold Ephraim Lessing). Hamann valued Hume's skepticism but insisted that it illustrated not the glory but the bankruptcy of reason. Skepticism is a paradigm of the ambiguity of all human powers—reason so conceived and directed is self-destructive. Hume, he felt, performed a service for philosophy in demonstrating what happens when reason is conceived as purely analytical, stripped of its functions of comprehension and intuition and removed from its orientation in religion and its foundations in historical experience.

Hamann's objections to Enlightenment philosophy can be illustrated by two sexual images he employed (such images were characteristic of his style). The rationalism of the age was trying to strip truth of her clothes, or, to change the figure, was trying to divorce what Nature had joined together, to attain reality by removing all excrescences, such as tradition, history, and experiential particulars. For Hamann truth appeared most authentically as "enfleshed" and therefore embodied in a unity of reason, faith, and sensual experience. He was skeptical of abstractions and saw language as the means by which the reason is confused as well as the means by which it expresses itself (the "seducer" as well as the "helpmeet" of man). He insisted on the wisdom and religious depth inherent in the naive vernacular, in imagery, and in myth. A "coincidence of opposites" was to be expected in the present world, even where the opposites seemed to be the most surprising and paradoxical—flesh and spirit, God and man, sensual language and transcendent conceptuality, history and reason. The most radical skepticism conceals a surreptitious credulity, and the most notorious agnosticism a covert religion.

QUARRELS WITH OTHER THINKERS

Hamann's belief that the Enlightenment was sacrificing the concreteness, historicity, and earthiness of reality to the paramount desire of rationalism to systematize initiated his quarrel with such thinkers as Herder,

Mendelssohn, Kant, and Lessing as well as obscure lesser lights. *Philologische Einfälle und Zweifel über eine akademische Preisschrift* (Philological ideas and doubts about a writing which received an academic prize; 1772) attacks Herder's prize-winning essay on the origin of language; *Golgatha und Scheblimini* (1784) is a criticism of Mendelssohn's theory of "natural religion" and its relation to church and state; *KONXOMPAX* (1779) contains a satire on Lessing's dichotomy between the knowledge of God and the knowledge of "accidents" of history.

CRITICISM OF KANT

Hamann's evaluations of Kant's *Critique of Pure Reason* are found in his letters and in several essays, one of which is called "Metacritik über den Purismum der Vernunft" (Metacritique of the purism of the reason; 1784). The issue between these thinkers is the extent to which reason can be "pure" (that is, devoid of experience). Hamann objected not only to Kant's overvaluation of formal knowledge but also to the belief that Kant demonstrated that in some respects reason can be separated from sense experience.

LANGUAGE

To Hamann, demonstration of a separation between reason and sense experience is impossible because it depends on language, or mental symbols, the "purity" of which is ambiguous. This ambiguity cannot be removed (and the old Platonic ideal of knowledge of forms refurbished) by the double maneuver of surrendering knowledge of reality itself and locating the forms of space and time in the knowing ego, known to be emptied of fallible sense experience. The mark of such an "emptiness" (absence of sense experience) would be a synthetic judgment a priori, and according to Kant this would be based on the pure forms of sensible intuition. However, to Hamann these forms of intuition, which he took to be types of *language*, cannot be demonstrated to be pure, because language contains the capacity to create what may be an illusion. "Not only does the entire capacity to think rest upon language … but language is also in the middle of the misunderstanding of reason with itself." The forms of intuition are not simply passive channels for the content of experience, but active forms of language (or of symbols) that have the power to deceive the mind and create the illusion that they are a priori and necessary.

In the new picture of being that Hamann offered as an alternative to his age, man is seen as a creature of flesh and blood ("the heart beats before the head thinks"), history as a living communication of the meaning of man's existence ("a continuing sign"), and the world as the "language" of God ("speech to the creature through the creation"). The metaphor of "language" points to the symbolic nature of the world, which is not to be exhausted in its material significance, deified in a pantheism, or transcended by reason in a Platonic dualism.

INFLUENCE

Scholarly interest in Hamann has focused upon his influence on the Sturm und Drang and Romantic movements and on such figures as F. H. Jacobi, Friedrich von Schelling, Hegel, and Friedrich Schleiermacher; upon his role as a forerunner of existentialism; upon his pioneering exploration of the nature of human sexuality (see his *Versuch einer Sibylle über die Ehe* [Essay of a sibyl on marriage; 1775] and *Schürze von Feigenblättern* [Skirts of fig leaves; 1777]); upon his influence on religious thought, such as the Movement of Awakening (*Erweckungsbewegung*, a revival of intellectual pietism) and neoorthodoxy; upon his contributions to a philosophy of language; and upon his reconception of reason as essentially historical.

See also Anselm, St.; Augustine, St.; Enlightenment; Existentialism; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hume, David; Kant, Immanuel; Kierkegaard, Søren Aabye; Lessing, Gotthold Ephraim; Luther, Martin; Mendelssohn, Moses; Pascal, Blaise; Philosophy of Language; Philosophy of Sex; Reason; Schelling, Friedrich Wilhelm Joseph von; Schleiermacher, Friedrich Daniel Ernst; Skepticism; Socrates; Thomas Aquinas, St.; Unconscious.

Bibliography

WORKS BY HAMANN

Werke, 6 vols. Edited by Josef Nadler. Vienna: Thomas-Morus-Presse im Verlag Herder, 1949–1957.

Briefwechsel, 8 vols. Edited by Walther Ziesemer and Arthur Henkel. Wiesbaden: Insel-Verlag, 1955–.

WORKS ON HAMANN

Alexander, W. M. "The Alpha and Omega of Hamann's Philosophy." *Ultimate Reality and Meaning* 4 (1981): 297–309.

Alexander, W. M. "Johann Georg Hamann: Metacritic of Kant." *Journal of the History of Ideas* 27 (1966): 137–144.

Alexander, W. M. *Johann Georg Hamann: Philosophy and Faith*. The Hague: Martinus Nijhoff, 1966.

Betz, John R. "Enlightenment Revisited: Hamann as the First and Best Critic of Kant's Philosophy." *Modern Theology* 20(2) (2004): 291–301.

Blanke, Fritz et al. *Hauptschriften erklärt*, 8 vols. Gütersloh, 1956–. Commentaries on Hamann's major works. Vol. I contains a comprehensive bibliography (pp. 145–176).

Butts, Robert E. "The Grammar of Reason: Hamann's Challenge to Kant." *Synthese* 75 (1988): 251–283.

Graubner, Hans. "Theological Empiricism: Aspects of Johann Georg Hamann's Reception of Hume." *Hume Studies* 15(2) (1989): 377–385.

Leibrecht, Walter. *Gott und Mensch bei Johann Georg Hamann*. Gütersloh, 1958.

Metzke, Erwin. *J. G. Hamanns Stellung in der Philosophie des 18. Jahrhunderts*. Halle, 1934.

Nadler, Josef. *Johann Georg Hamann, der Zeuge des Corpus Mysticum*. Salzburg: O. Müller, 1949. Biography.

O'Flaherty, James C. *Johann Georg Hamann*. Boston: Twayne Publishers, 1979.

O'Flaherty, James C. *The Quarrel of Reason with Itself: Essays on Hamann, Michaelis, Lessing, Nietzsche*. Columbia, SC: Camden House, 1988.

Redmond, Michael D. "The Hamann-Hume Connection." *Religious Studies* 23 (1987): 95–107.

Salmony, H. A. *Johann Georg Hamanns metakritische Philosophie*. Zollikon: Evangelischer, 1958.

Smith, Ronald Gregor. *J. G. Hamann, with Selections from His Writings*. London: Collins, 1960.

Swain, Charles W. "Hamann and the Philosophy of David Hume." *Journal of the History of Philosophy* 5 (1967): 343–351.

Wessel, Leonard P. "Hamann's Philosophy of Aesthetics: Its Meaning for the Storm and Stress Period." *Journal of Aesthetics and Art Criticism* 27(1969): 433–443.

W. M. Alexander (1967)
Bibliography updated by Tamra Frei (2005)

HAMELIN, OCTAVE
(1856–1907)

Ernst Heinrich Haeckel, the French idealist philosopher, carried further the neocriticism of Charles Renouvier and Jules Lachelier. Renouvier had criticized the categories of Immanuel Kant, maintaining that the category of relation really included all the others. He had also criticized the fact that Kant had not included personality in the categories, though he should have, since the will determines thought no less than does speculative reason. Thought, according to Renouvier, manifests relation in two ways—in the relations of its elements to each other, and in the relation of judgment to consciousness. The latter relation is always unique because consciousness displays its spontaneity in the synthesis of the objects which it posits.

Hamelin started from a similar position; in fact, he always declared himself a disciple of Renouvier. However, Renouvier had taken account of contingency and discontinuity, and for him the pursuit of truth involved a recurrent dilemma, in which a free choice or wager was presented to the seeker. Hamelin, by contrast, was much more intransigently rationalistic, and in this he was influenced by Lachelier. A priori thinking pervaded Hamelin's system, and he envisaged a reality made entirely transparent and intelligible by a process of rational deduction. "Knowledge will no longer be seen as the invasion of the subject by alien elements, but as a putting into action by the subject of his potentialities."

Such an idealistic system requires some principle that accounts for change yet allows change to remain compatible with the necessity of rational deduction. In Hamelin's system this principle is correlation, and following G. W. F. Hegel, Hamelin pictured a dialectical evolution of reality through the synthesis of complementary opposites. This movement is from abstract elements toward the constitution of concrete reality—toward the constitution, indeed, of conscious personality—and not, as in Hegel, toward the indefinite pursuit of an Absolute. Hamelin's philosophy was a highly ingenious, if perhaps unsuccessful attempt to do away with the dilemma that dogged nineteenth-century French philosophy, the dilemma of a necessary realm of thought and a contingent domain of occurrences. Hamelin brought contingency, freedom, and personal consciousness within his dialectical system, making them the necessary outcome of incomplete, abstract categories that invoke them. They make their appearance, moreover, as the product of the dialectical process and as the coming to fruition of a hitherto inchoate reality.

Hamelin deduced the categories, or elements of representation, according to this dialectical principle. Relation is the synthesis of being and nonbeing, as that which consists in interdependence. The antithesis of relation is what is essentially independent, number. Number and relation are synthesized in time. Space stands in antithesis to time since its parts, though separate, are also simultaneous and reversible. The space-time antithesis is transcended in motion, still a quantitative concept, which finds its opposite in what is unaffected by it, quality. Motion and quality are synthesized in modification (*altération*), which is the movement of quality. Modification is contrasted with a kind of resistance to change that tends to perpetuate the initial state, and this is specification, or the notion of class or species. Out of the interaction of modification and specification comes causality, or change brought about in beings through their sharing the world with other beings. Opposed to causality is a principle of persistence within the self, which is finality. The ultimate synthesis is in independence and self-sufficiency,

which is expressed as free becoming and is an active system, or conscious self.

The characteristic of Hamelin's philosophy is its strictly a priori derivation of the concrete and individual consciousness from general and abstract elements. There is a kind of dynamism of incomplete abstraction working toward its own fulfillment and specification as successive logical demands are met. The element of contingency, which is inescapable in reality, finds its way into the system in the freedom of the individual consciousness: "What provides the explanation of consciousness is the need to choose."

Hamelin proceeded as far as anyone could in the idealistic direction. His no longer fashionable metaphysical deduction of the world through a series of necessary relations has met with some criticism from more recent French philosophers, whose tendency has been rather to see meaningful experiences as involving a compromise between a world of brute facts, which is prereflective, and a mind that almost necessarily orders them, albeit according to its own requirements and in the light of its own tasks.

See also Appearance and Reality; A Priori and A Posteriori; Hegel, Georg Wilhelm Friedrich; Idealism; Kant, Immanuel; Lachelier, Jules; Renouvier, Charles Bernard.

Bibliography

WORKS BY HAMELIN

Essai sur les éléments principaux de la représentation. Paris, 1907.

Le système de Descartes. Paris: F. Alcan, 1911.

Le système d'Aristote. Paris: F. Alcan, 1920.

Le système de Renouvier. Paris: J. Vrin, 1927.

WORKS ON HAMELIN

Parodi, Dominique. *Du positivisme à l'idéalisme.* Paris: J. Vrin, 1930.

Parodi, Dominique. *La philosophie contemporaine en France.* 3rd ed. Paris, 1925.

Colin Smith (1967)

HAMILTON, WILLIAM
(1788–1856)

William Hamilton, the Scottish philosopher and logician, was born in Glasgow and was educated at Glasgow, at Edinburgh, and at Balliol College, Oxford, where he took his B.A. in 1811. After leaving Oxford he studied law and in 1813 was admitted to the Scottish bar. He was appointed professor of civil history at the University of Edinburgh in 1821 and was elected to the chair of logic and metaphysics in 1836. Hamilton, a man of stupendous erudition, was strongly influenced by Thomas Reid and Immanuel Kant.

PSYCHOLOGY

Hamilton divided "mental modifications, or phenomena" into three classes—the phenomena of knowledge or cognition; the phenomena of feeling, of pleasure and pain; and the phenomena of will or desire, the exertive or conative powers. Knowledge, feeling, and will or desire cannot exist independently of one another. Every state of mind is a combination of all three, although proportions may vary. We can conceive of a being who knows one thing or another but is totally devoid of feeling, desire, and volition; or of a being capable of knowledge and feeling only; but not of a being having the capacity for pleasure and pain and the capacity to will, but lacking the faculty of knowledge.

Mental phenomena are included under the phenomenon of consciousness. When one knows, he knows or is conscious that he knows; when one feels, he knows or is conscious that he feels; and when one desires, he knows or is conscious that he desires. Consciousness is not something in addition to knowledge, feeling, and desire, but the general condition of their existence. It is a relation between a knowing (or conscious, or intelligent) subject and an object of knowledge, in this case a modification of the mind. Although Hamilton sometimes denied the possibility of unconscious mental states, at other times he argued that "the mind exerts energies, and is the subject of modifications, of neither of which it is conscious" (*Lectures on Metaphysics,* Lecture 18).

PERCEPTION

In perception, according to Hamilton, we have an immediate or presentative rather than a mediate or representative knowledge of the object. In presentative cognition a thing is known in itself rather than via something other than itself. When I see a cat I come to know the animal in itself as contrasted with, for example, my knowledge of a past event, which is acquired through testimonials and other means distinct from the event thus cognized. I may have representative knowledge of the past, the future, and the merely possible, as in imagination. Immediate presentative knowledge is of that which exists here and now. Perception is the faculty presentative of the phenomena of the nonego (matter), and self-consciousness is the fac-

ulty presentative of the phenomena of the ego (mind). A thing is known in itself only if it be known as actually existing in its when (now) and its where (here). Perception has for its objects the primary qualities of bodies. Knowledge of secondary qualities is never immediate, in that all we can know of them is that some unknown external cause is responsible for the "present affections of the conscious subject." Thus Hamilton agreed with Thomas Reid that we have a "direct notion" of primary qualities but only relative notions of the secondary qualities of things.

In perception we are intuitively aware of the duality of ego and nonego. This is an immediate, primitive datum of consciousness, to whose existence the natural realist (or natural dualist—both terms were used by Hamilton to designate holders of views like his own) is implicitly committed. Perception is not inference. We do not first become aware of some mode of consciousness and then infer from this awareness the present existence of a physical object as cause of that modification. Nor are we aware of an inner representation or referent from which we conclude the existence of an object referred to or represented. Representative theories of perception presuppose what on their own terms could not be the case. In order to know that *A* refers to or represents or is a sign of *B,* it must be possible to gain knowledge of the existence and nature of *B* independently of our knowledge of the existence and nature of *A.*

RELATIVITY OF KNOWLEDGE

Granting that our senses inform us of the existence and the nature of physical objects, just what information do they provide us with? Hamilton held that our knowledge of mind and of matter is relative and conditioned and that "of existence absolutely and in itself, we know nothing" (*Lectures on Metaphysics,* Lecture 8). Our knowledge of the ego as well as of the nonego is purely phenomenal. The self is known to us only via the phenomena of the immediate introspective awareness of the flow of experience. In external perception we come to know about physical objects only as they appear to us through the senses. A physical object as known is that which appears to us as extended, solid, divisible, figured, colored, hot or cold. Thus, "matter" or "body" is a name for a certain set of appearances or phenomena, but these must be regarded as appearances *of* something. This something, however, is inconceivable apart from its phenomena, absolutely and in itself, out of relation to a knower. By virtue of a "law of thought" we are compelled to think of something absolute, unknown, and unknowable as the subject, substance, or substratum of the relative, the phenomenal, the known. The same reasoning applies to mind.

That a thing or a quality of a thing is known in itself does not mean that it is known in its "absolute existence" out of relation to the knower, for this is impossible. Hamilton meant, presumably, only that it is not known through a process of inference from signs or representations. All knowledge is relative in that in order to be known a thing must be related to the knower, the relation being precisely that of the knower to the known. But this is trivial. Hamilton pointed out that the way a thing appears to us in perception is relative in another sense—it is a function not only of the objective qualities of the thing, but of the medium and the sense organs as well. When I perceive a book, the phenomena or appearances of the external object are a resultant of the contributions of the book, of the intervening medium, and of the sense organs. Consequently my knowledge of the book is modified through certain intermediate agencies and must be relative. But, as J. S. Mill pointed out in *An Examination,* this conclusion does not follow; rather than entailing the relativity of all knowledge of physical objects, the considerations adduced show at most that that part of the knowledge that is not contributed by the book itself is relative.

PHILOSOPHY OF THE CONDITIONED

To think of a thing is necessarily to think of it as a thing of a certain sort, to classify it, to subsume it under a concept. Thought imposes conditions on its object. Therefore the conditioned is the only possible object of knowledge. The absolute, the nonrelative, the unconditioned is inconceivable; all we can know is *that* it is, not *what* it is. Although many things are inconceivable to us, nonetheless we know that some of them must be true. Hamilton claimed that, given the principles of contradiction and excluded middle, all actual thought lies between two extremes, each of which is inconceivable. The extremes represent that which is absolute or unconditioned. One of these absolutes we know must be true because they are mutually contradictory; but since both are inconceivable we cannot know which is true. "The Conditioned is the mean between two extremes—two unconditionates, exclusive of each other neither of which can be conceived as possible, but of which … one must be admitted as necessary" (*Discussions,* p. 22). The weakness of the human mind thus restricts its objects of positive thought to the mean. As illustration Hamilton argued that space must either be bounded or unbounded. One

alternative must be true, but both cannot be, even though we cannot positively conceive of either one. Similarly we cannot conceive of an absolute beginning of time or of an infinite regress. We cannot conceive of an absolute end of time or of an infinite prolongation, although one or the other must be admitted to be true.

LOGIC

Hamilton regarded his doctrine of the quantification of the predicate as his most important contribution to logic. The doctrine is based on the self-evident truth that we can operate rationally only with what we already understand. This in turn leads to the postulate that we ought to be able to state explicitly what is thought implicitly. When we make a judgment we always implicitly understand the quantity of the predicate as well as the quantity of the subject. Since the predicate is always quantified in thought, and since every quantity is either all or some or none, we always regard the predicate of a judgment as denoting all, some, or none of the objects in its extension. The proposition "All men are animals" must mean either that all men are all animals (all men and only men are animals), or that all men are some animals (all men, but not men only, are animals) "Some animals are carnivorous" becomes "Some animals are some carnivorous," which is to be understood as some and some only, that is, some animals are carnivorous and some are not. Among the advantages of this innovation, according to Hamilton, are the reduction of propositions to equations, the simplification of the doctrine of conversion, the abolition of the figured syllogism and the consequent manifestation of the absurdity of reducing syllogisms of other figures to the first.

Since in Hamilton's view logic is the science of the laws that of necessity govern all valid thought, criticisms of him drawn from psychological considerations are relevant. It comes as something of a surprise to the beginner in logic that conversion of the universal affirmative "All *S* is *P*" is only by limitation to "Some *P* is *S*." But in Hamilton's view this should be obvious to all and should not represent a new idea as in fact it does.

See also Kant, Immanuel; Logic, History of; Perception; Phenomenology; Psychology; Reid, Thomas.

Bibliography

WORKS BY HAMILTON

Original Philosophy

"On the Philosophy of the Unconditioned." *Edinburgh Review* 1 (1829): 194–221.

Discussions on Philosophy and Literature, Education and University Reform. London: Longman, Brown, Green and Longmans, 1852.

Lectures on Metaphysics and Logic. Edited by H. L. Mansel and John Veitch, 4 vols. Edinburgh and Boston: Blackwood, 1859–1860.

Editing and Commentary

The Works of Thomas Reid. 2 vols. Edinburgh, 1846. Hamilton edited this work and added important commentaries.

The Works of Dugald Stewart. 11 vols. Edinburgh, 1854–1856. Hamilton edited this work.

WORKS ON HAMILTON

Grave, S. A. *The Scottish Philosophy of Common Sense.* Oxford: Clarendon Press, 1960.

Mill, John Stuart. *An Examination of Sir William Hamilton's Philosophy.* 2 vols. London: Longman, Green, Longman, Roberts and Green, 1865.

Rasmussen, S. V. *The Philosophy of Sir William Hamilton.* Copenhagen, 1925.

Stirling, J. H. *Sir William Hamilton.* London: Longmans, Green, 1865.

Veitch, John. *Hamilton.* Edinburgh: Blackwood, 1882.

Wight, O. W. *Philosophy of Sir William Hamilton.* New York: Appleton, 1853.

Timothy J. Duggan (1967)

HAMPSHIRE, STUART NEWTON
(1914–2004)

Stuart Newton Hampshire, born in Lincolnshire, England, was a fellow of All Souls College and of New College, Oxford, and then Grote professor of mind and logic at University College, London (1960–1963); he also was professor of philosophy at Princeton University. From 1970 to 1984 he taught at Wadham College, Oxford; in 1984 he joined the faculty of Stanford University, becoming emeritus in 1990. Hampshire's contribution to philosophy, while clearly belonging to the main current of contemporary work in the English language, was highly individual. His work displays a broad and systematic outlook, concerned with bringing together views in the theory of knowledge, metaphysics, the philosophy of mind, ethics, and aesthetics. Among influences outside philosophy itself, it shows a particular awareness of psychoanalysis and of the history and criticism of literature and painting. His philosophical style is distinctive, a sensitive blend of the argumentative and the exploratory, which can be seen as the product of two contrasting influences: a sympathy with the outlook of Friedrich Waismann (himself influenced by Ludwig Wittgenstein) that there can be no proofs in philosophy, together with a respect

for the aim of J. L. Austin and other recent philosophers to reach definite results by definite methods.

Hampshire showed a constant interest in the connections between meaning and confirmation. To this extent, there are links between his concerns and those of logical positivism, and a relatively early paper, "Logical Form," shows a recognizably positivist spirit in explaining differences of form in terms of differences in methods of confirmation. However, Hampshire's views were never positivist. In particular, he was not so much concerned to assign a privileged possibility of certainty to some special class of statements but rather to explore the various certainty conditions of different classes of statement.

The connections between meaning and certainty conditions have been particularly explored with reference to psychological statements. Hampshire rejected the possibility of Cartesian statements of immediate experience, independent of any bodily conditions. He emphasized both the need of communication with other persons for self-knowledge ("The Analogy of Feeling") and the dependence of the subject's sense of his identity on his being a physical agent in a physical environment. This idea is treated in detail in *Thought and Action*, where some influence of Jean-Paul Sartre and of Maurice Merleau-Ponty can be seen. While stressing the connections between mental concepts and physical agency, Hampshire sought at the same time to give an intelligible place to introspection and to the possibility of incorrigible declarations by a speaker of his own mental states, particularly in the case of intentions: Besides *Thought and Action*, see "Self-Knowledge and the Will" and the important article "On Referring and Intending." In line with this is his rejection of any thoroughgoing behaviorist analyses of psychological concepts (review of Gilbert Ryle's *Concept of Mind*) and his interesting account of the notion of a *disposition* as applied to human character ("Dispositions"), an account later elaborated in psychoanalytical terms (in "Disposition and Memory"). Human dispositions must be distinguished logically from merely "dispositional properties," such as are possessed by material objects: Dispositional properties can exist without having been manifested, but ascription of human dispositions implies some actual manifestations of them in the life of the individual. Moreover, the understanding of a human disposition is of a different character, being basically historical or genetical. Psychoanalysis is taken to reveal a basic way of understanding the individual's disposition, as rooted in his early experiences and consisting in the generalization to a class of situations of primitive responses; the influence of the primitive situations is to be seen in terms of unconscious memory. These ideas provide a link between the concept of a disposition and those of rationality and freedom; control over one's dispositions may be increased by self-knowledge, the understanding of how they have come about.

The emphasis on the psychoanalytical type of account of dispositions—that is, a genetical account—is a particular application of the wider view that human activities must be understood historically. This view has had an important influence on Hampshire's outlook on ethics, which rests on two points—that any comprehensible system of ethics must be grounded in a view of human nature and that all views of human nature are historically conditioned and essentially revisable. However, the historical changes in views of human nature or "the powers of the mind" are comprehensible only against a background of something identified, under any view, as essential to human nature, and this Hampshire finds in the possibility of self-conscious intentional action. From this point of view, Hampshire seeks to illuminate two basic (and, he would hold, permanent) distinctions: that between art and other human activities and that between human actions and mere events. Art is connected with the absence of an intentional project (see "Logic and Appreciation"); the appreciation of art is a process of free exploration. The distinction between actions and mere events involves his theory of freedom, which turns on a basic distinction between decision and prediction, and on the claim that there is an ineliminable human power of "standing back" from any prediction of one's future actions, the situation thus being changed (see *Thought and Action* and "Decision, Intention and Certainty"). Whether this has the consequence that determinism is impossible is perhaps not entirely clear; it is notable that Hampshire treated these questions in an illuminating book on Benedict de Spinoza, and his sympathy for a Spinozist connection of freedom and knowledge, rather than a supposed freedom of the will, certainly continued (see "Spinoza and the Idea of Freedom").

See also Austin, John Langshaw; Freedom; Logical Positivism; Merleau-Ponty, Maurice; Perception, Contemporary Views; Philosophy of Mind; Propositions, Judgments, Sentences, and Statements; Ryle, Gilbert; Self-Knowledge; Spinoza, Benedict (Baruch) de; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY HAMPSHIRE

"Logical Form." *PAS* 48 (1947–1948): 37–58.

"Fallacies in Moral Philosophy." *Mind* 58 (1949): 466–482.
Critical notice of Ryle's *Concept of Mind*. *Mind* 59 (1950): 237–255.
"Freedom of the Will." *PAS*, Supp. 25 (1951): 161–215.
Spinoza. Harmondsworth, U.K.: Penguin, 1951; London: Faber and Faber, 1956.
"The Analogy of Feeling." *Mind* 61 (1952): 1–12.
"Logic and Appreciation." *World Review* (1952). Reprinted in *Aesthetics and Language*, edited by William Elton. Oxford: Blackwell, 1954.
"Self-Knowledge and the Will." *Revue internationale de philosophie* 7 (25) (1953): 230–245.
"Dispositions." *Analysis* 14 (1953): 5–11.
"On Referring and Intending." *Philosophical Review* 65 (1956): 1–13.
"Identification and Existence." In *Contemporary British Philosophy*, 3rd series, edited by H. D. Lewis, 189–208. London, 1956.
"Decision, Intention and Certainty." *Mind* 67 (1958): 1–12. Written with H. L. A. Hart.
Thought and Action. London: Chatto and Windus, 1959.
Feeling and Expression. London: H. K. Lewis, 1960. Inaugural lecture.
"Spinoza and the Idea of Freedom." *Proceedings of the British Academy* (1960): 195–215.
"J. L. Austin." *PAS*, Supp. 34 (1960): 1–14.
"Friedrich Waismann." *Proceedings of the British Academy* (1960).
"Disposition and Memory." *International Journal of Psychoanalysis* 43 (1962): 59–68.
Freedom of the Individual. London: Chatto and Windus, 1965.
For a discussion of Hampshire's views on freedom, see D. F. Pears, ed., *Freedom and the Will* (London: Macmillan, 1963), Chs. 6 and 7.
Modern Writers, and Other Essays. London: Chatto & Windus, 1969.
Freedom of Mind, and Other Essays. Princeton, NJ: Princeton University Press, 1971.
Morality and Pessimism. London: Cambridge University Press, 1972.
Kolakowski, Leszek, and Stuart Hampshire, eds. *The Socialist Idea: A Reappraisal*. New York: Basic Books, 1974.
Freedom of the Individual. 2nd ed. London, Chatto & Windus, 1975.
Knowledge and the Future. Southampton, U.K.: University of Southampton, 1976.
Two Theories of Morality. Oxford: Published for the British Academy by Oxford University Press, 1977.
Public and Private Morality. Cambridge, U.K.; New York: Cambridge University Press, 1979.
Thought and Action. Rev. ed., London: Chatto & Windus, 1982.
Morality and Conflict. Cambridge, MA: Harvard University Press, 1983.
Innocence and Experience. Cambridge, MA: Harvard University Press, 1989.
Justice Is Conflict. London: Duckworth, 1999.

Bernard Williams (1967)
Bibliography updated by Michael J. Farmer (2005)

HANAFI, HASSAN

(1935–)

Born in Cairo, Hanafi studied first in Egypt but obtained his PhD in Paris (1966), working with Robert Brunschvig on a thesis entitled *Les méthodes de l'exégèse: Essai sur les fondements de la compréhension, 'Ilm usūl al-fiqh*. Hanafi has been a professor of philosophy in Cairo University since 1967 and has written many substantial works, three of which are particularly significant. He produced a study in five volumes on political theory, *From Dogma to Revolution* (in Arabic, 1985), an eight volume inquiry (in Arabic) on the links between religion and revolution in Egypt that focuses also on contemporary Islamic trends (*Religion and Revolution in Egypt*, 1989). The perspective of the foundation of an "Islamic left" leading through religion to a reconstruction of politics is outlined in the book and represents the main political approach of the author. The collected papers of *Islam in the Modern World* (1996) (two volumes) cover a wide range of topics from philosophy and theology to sociology and politics.

Hanafi's philosophical method is grounded in phenomenology and hermeneutics and he is particularly effective in applying this method to Islam. Consciousness and history acquire great importance accordingly. Islam is not merely a religion, according to him, but above all an ideology that connects the temporal and the sacred. Thus the outward (social and practical) and the inner (related to conscience) dimensions of human reality are but two aspects of the same phenomenon. Theology must become anthropology in order to allow humanity to make faith the tool of transformation of economic and social relations. The translation of theology into anthropology needs firstly the Husserlian *epoché* on God's essence; and secondly a new orientation of the object of theology. The center of revelation as the science of God is no longer God but humanity. Revelation is the science of humanity because humans are its objects and interlocutors. In this transformation of theology into anthropology, God keeps his value as telos, the goal of human activity in front of which all are equal. God is not logos, but praxis; not an idea, but a form of practice. Consequently, in Hanafi's view, Islam is a religion of revolution and justice prompting everybody to refuse any subordination to oppressive power and to claim the liberation of the world and its people in the name of God.

Hanafi criticizes Orientalism as a science aimed at colonial submission. He believes that the Third World's peoples have to develop a science of Occidentalism in order to get a fresh cultural, political, and philosophical

stance. Then they will be able to join Europe and North America in its modernity and recover their role in universal history (these ideas are discussed in Hanafi's *Introduction to the Science of Occidentalism*, in Arabic, 1991). From the perspective of intellectual recovery, Hanafi believes a new interpretation of Islamic heritage (turāth) is vitally important, because the reconstruction of a historical consciousness—namely tradition—is the direct path to development. Among the applications of this new reading are an inquiry into the traditional science of *hadith* (the traditional sayings of the Prophet) which Hanafi acknowledges has a historical character, and into Qur'anic exegesis that is envisaged to require an interpretation linked not only to explanation, but also to understanding, and not only to knowledge, but also to awareness.

See also Hermeneutics; Islamic Philosophy; Justice; Revelation.

Bibliography

Akhavi, Shahrough. "The Dialectic in Contemporary Egyptian Social Thought: The Scripturalist and Modernist Discourses of Sayyid Qutb and Hasan Hanafi." *International Journal of Middle Eastern Studies* 19 (1997): 377–401.

Boullata, Issa J. *Trends and Issues in Contemporary Arabic Thought*. Albany: State University of New York Press, 1990.

Campanini, Massimo. "Hasan Hanafi e la fenomenologia: per una nuova politica dell'Islam." *Oriente Moderno* 4 (13) (1994): 103–120.

Esposito, John L., and John O. Voll. "Hasan Hanafi. The Classical Intellectual." In *The Makers of Contemporary Islam*. New York: Oxford University Press, 2001.

Hanafi, Hasan. *Islam in the Modern World*. Cairo: Anglo-Egyptian Bookshop, 1996.

Hanafi, Hasan "Théologie ou Anthropologie?" In *Renaissance du Monde Arabe*. Gembloux, Belgium: Duculot, 1973.

Salvatore, Armando. "The Rational Authentification of the Turath in Contemporary Arab Thought: al-Jabiri and Hasan Hanafi." *The Muslim World* 85 (1995): 191–215.

Massimo Campanini (2005)

HAN FEI

(c. 280?–233 BCE)

Han Fei, the outstanding authority of the Legalist school of Chinese philosophy during the Warring States Period (403–221 BCE), lived some two centuries or more later than Confucius, Laozi, and Mozi. His works consist of fifty-five treatises. He was an aristocrat of the state of Han, in the northwest, where a movement of political reform had arisen. In the rest of the Chinese kingdom, Confucianism, Mohism, and Daoism still maintained the theory of government by customary morality, religious sanctions, and personal virtuous example or even "inaction" (or "nonaction") by the ruler. Since the traditional feudal system had collapsed generally throughout China, new thinkers spoke of government by positive law, exaltation of the state at the expense of the individual, and possession of supreme power by the ruler. Out of these thoughts of his predecessors, Han Fei made a comprehensive synthesis and formed his unique doctrine of legalism. This doctrine was greatly admired by the ruler of the state of Qin, who by its adoption eventually became the first emperor of the unified Qin Empire. Han Fei had been invited to the court of Qin, but he was forced to commit suicide by Li Si, chief minister of Qin and former associate of Han Fei. Although they had studied together under Xunzi, the exponent of a reformed Confucianism, political jealousy overcame Li Si.

Han Fei accepted only one point of Xunzi's philosophy, that human nature is originally evil. He then insisted that all men act from selfish motives and that a system of rewards and punishments can provide an effective control for the government, for there is no one who does not fear punishment and welcome rewards. The ruler can then rule the state through a set of severe laws without his own interference. In economics, too, the government should adopt a laissez-faire policy, leaving people alone to carry on free competition; this will cause them to be more active and frugal, with the result that production will increase. Thus, Han Fei reinterpreted the inaction principle of Laozi in such a way that the sovereign, like the helmsman of a great ship, simply makes slight movements with the "two handles" of reward and punishment, and the whole state follows his desires and dictates.

While "statecraft" serves to keep the sovereign in power, laws are to be obeyed by the people. Han Fei said: "The intelligent ruler sees to it that his subjects do not allow their minds to wander beyond the scope of law; nor perform acts of favoritism within it" (*Works,* Treatise II). He also observed: "In the state of the intelligent ruler there is no literature of books and records, but the laws serve as teachings. There are no sayings of the early kings, but the officials act as teachers." Once laws have been established everyone should obey them; no heterodox doctrines of private individuals and traditional authorities should prevail. This led the first emperor to practice totalitarian regimentation, "burning the books and burying the literati" (*Works,* Treatise XIX).

It was Han Fei who supplied a rational explanation for revolutionary changes from the old order. He

asserted: "Affairs go according to their time and preparations are made in accordance with affairs.… The sage does not aim at practising antiquity and does not model himself upon what is considered to be permanently correct" (ibid.). Indeed, history does not repeat itself. Politics, therefore, must look always to the present and to changing circumstances rather than to any static idea or ideal. In a word, Han Fei can be regarded as a radical positivist or perhaps as the extreme realist of ancient China.

See also Chinese Philosophy; Confucius; Laozi; Mozi; Xunzi.

Bibliography

Landers, James R. "The Political Thought of Han Fei." Ph.D. diss. Indiana University, 1972.

Liao, Wen-kuei. *The Complete Works of Han Fei Tzu.* Vol 1: London: Arthur Probsthain, 1939. Vol 2: London: Arthur Probsthain, 1959.

Lundahl, Bertil. *Han Fei Zi: The Man and the Work.* Stockholm, Institute of Oriental Languages, Stockholm University, 1992.

Waley, Arthur. *Three Ways of Thought in Ancient China.* London: Allen and Unwin, 1939; Stanford, CA: Stanford University Press, 1982. Selected translations.

Wang, Hsiao-po, and Leo, S. Chang. *The Philosophical Foundations of Han Fei's Political Theory.* Honolulu: University of Hawai'i Press, 1986.

Watson, Burton. *Han Fei Tzu: Basic Writings.* New York: Columbia University Press, 1964. Selected translations.

Nai Z. Zia (1967)
Bibliography updated by Huichieh Loy (2005)

HAN FEI TZU

See *Han Fei*

HAN YU

(768–824)

Han Yu lived in a time when the Chinese Tang empire (618–907) was threatened by military separatism but enjoyed cultural creativity and economic expansion. He became a major writer in his youth and had a successful official career late in his life. He was an innovative poet and essayist and the chief champion of the *guwen* movement that paved the way for a fundamental change of prose style. More unusual for a writer, Han played a leading role in a crucial philosophical redirection.

As a thinker, Han's most important idea was that Confucianism is the sole legitimate teaching for human conduct, to the exclusion of Buddhism and Daoism. This was an extreme position in his own time, but it exerted profound influence throughout later Chinese history. Han presented this view most forcefully in his famous essay "Essentials of the Moral Way" (*Yuan Dao*). This essay asserts that the only Dao is the one based on everyday life, which is the Confucian Way discovered and developed by ancient sage-kings. What are the teachings of these sages? Han declares:

> To love universally, which is called humanity; to apply this in the proper manner, which is called rightness; to act according to these, which is called the Way; to [follow the Way and] become self-sufficient without seeking anything outside, which is called virtue. The *Book of Poetry*, the *Book of History* … are their writings; rites and music, punishments and government, their methods. Their people were the four classes of scholar-officials, farmers, artisans, and merchants; their relationships were those of sovereign and subject, father and son, teacher and friend, guest and host, elder and younger brother, and husband and wife. Their clothing was hemp and silk; their dwellings halls and houses; their food grain and rice, fruit and vegetables, fish and meat (de Bary, et al. 1960, pp. 378–379 with minor changes).

A key point here is the all-embracing and this-worldly nature of the Confucian Way. Han made his point by going so far as to include people's clothing and food as part of the Way.

In his treatise, Han not only rejects all teachings that attempt to find the meaning of life outside or beyond the social order prescribed by Confucian sages, but also asserts that Confucianism values spiritual life as well. However, the Confucian mode of self-cultivation is intrinsically linked to mundane life; spiritual purification should be a basis for bettering, not transcending, the world. In another essay, Han gives his picture of Confucian spirituality. It is essentially a reconfiguration of ideas prevalent mainly during the Han era (206 BCE–220 CE), and failed to win the approval of the later thinkers taking up his program of Confucian renewal.

The chief target of Han's intellectual campaign was Buddhism—the dominant religion of medieval China—which because of its foreign origins conflicted with Confucian values in many fronts. Yet the true significance of his thought shows more clearly in his criticism of Dao-

ism. Han categorically disagreed with the anticivilization attitudes in the *Laozi* and the *Zhuangzi*. He held that in constructing what is now called the Confucian order, ancient sage-kings saved humankind from a state of chaos and savagery; the Daoist calls for a return to the innocent primeval age, he believed, were absurd. Han also accused the religious Daoists, with their search for immortality and a secluded life, of deserting their inviolable duties as members of the human community. By taking an almost unique position against Buddhism, religious Daoism, and philosophical Daoism simultaneously, Han was assailing a view predominant in China since the early third century CE, that Daoism and Buddhism brought to light questions concerning the fundamental essence of the world and the spirituality of individuals whereas Confucianism had practical uses in building a proper sociopolitical order. Han put forward a new vision: that a proper human society can only be built upon Confucian principles in toto.

Han wrote only a few formal essays on philosophical issues, but often expressed his views in other genres and in highly literary manners. He was an original thinker who had effective weapons with which to send his message. Han challenged a fundamental intellectual premise of medieval China, and opened the way for the eventual formation of "the Learning of Principle" (*lixue*)—or neo-Confucianism—in the eleventh century. If the word "pioneer" means anything in historical account, Han was a most significant pioneer.

See also Confucius; Li Ao.

Bibliography

de Bary, William Theodore, et al., eds. *Sources of Chinese Tradition.* Vol 1, edited by Wing-tsit Chan and Burton Watson, 371–382. New York: Columbia University Press, 1960.

Hartman, Charles. *Han Yü and the T'ang Search for Unity*. Princeton, NJ: Princeton University Press, 1986.

McMullen, David. "Han Yü: An Alternative Picture." *Harvard Journal of Asiatic Studies* 49 (2) (December 1989): 603–657.

Jo-shui Chen (2005)

HAPPINESS

As an object of philosophical inquiry, the concept of happiness is as old as philosophy itself. It was central to the ethical thought of the Greeks, most famously Aristotle, and was restored to this position of prominence by the nineteenth-century utilitarians. Whether a principal theme or not, the pursuit of happiness plays some role in virtually all ethical traditions. Indeed, few would deny that happiness is one of the important goals in life, if not the only one. Through the twentieth century and into the twenty-first, both philosophers and psychologists have continued to ask questions about happiness. These questions fall into two broad categories: (1) The nature of happiness—in what does it consist? (2) The value of happiness—what is its role in a theory of ethics or of the good life?

WHAT HAPPINESS IS

The terms *happy* and *happiness* are used in many different ways, and it is important to identify the precise concept that is of philosophical interest. We often speak about being happy with or about something, where this means roughly regarding it favorably or having a positive attitude toward it. The object of such an attitude can, in principle, be anything: a state of oneself or a state of the world. We also speak about feeling happy, where this is an occurrent state of mind characteristically accompanied by energy, vitality, and buoyancy of spirit. Both of these notions need to be distinguished from that of being happy or having a happy life. When philosophers investigate the nature or value of happiness, it is this concept of a happy life that they have in mind.

Accounts of the nature of happiness can be partitioned initially along a subjective–objective dimension. An analysis is subjective if it makes a person's happiness depend, at least in part, on attitudes or feelings. Conversely, it is objective if happiness is taken to be entirely independent of these subjective states so that someone could be happy even if neither feeling happy nor having a positive attitude toward the conditions of life.

The best-known example of an objective conception of happiness is Aristotle's. According to Aristotle, eudaimonia consists in the excellent functioning of the soul, thus the exercise of virtue. While Aristotle allowed, as does common sense, that a person's happiness could also be affected by external goods and circumstances, the stronger view that virtue is not only necessary for happiness but also sufficient was affirmed by the Stoics. Both views, however, have seemed counterintuitive to many. If virtue is necessary for happiness, how is it possible (as it seems to be) for vicious people to be happy? And if virtue is sufficient, then would we not be compelled to call the virtuous happy even if they are consumed by torment and suffering? The seemingly paradoxical implications of linking virtue and happiness in either of these ways has led some commentators on Aristotle to the conclusion

that what he means by eudaimonia is not adequately captured by our concept of happiness, so that the Greek term he uses should really be translated in some other way, possibly as *well-being*.

However this might be, when the utilitarians revived interest in happiness as an object of ethical inquiry, they assumed a subjective analysis of its nature. For Jeremy Bentham, John Stuart Mill, and Henry Sidgwick, happiness consisted in pleasure and the absence of pain or a surplus of pleasure over pain. The hedonism they espoused represents one of the main options for a subjective analysis of happiness. On the simplest version of it, a person's hedonic state at a particular time is determined by the balance of agreeable and disagreeable feelings at that time. Happiness is then a matter of the longer-term tendencies of these hedonic balances: The greater and more enduring the balance of positive over negative states, the happier the person. In recent years this reduction of happiness to positive and negative affect has been one prominent theme in the emerging field of hedonic or positive psychology. Indeed, a hedonistic or affective state analysis of happiness is now most commonly defended by psychologists such as Daniel Kahneman, though it still has its philosophical advocates as well.

However, many philosophers attracted to the project of a subjective analysis of happiness have found the hedonistic account unsatisfactory. For one thing, the usual sources of pleasure seem too short-term and episodic to tell the whole story of whether a person is happy. The idea that happiness over a lifetime, or a considerable stretch of a life, can be computed simply by adding up episodes of positive and negative affect and finding the balance neglects the role of more global factors, such as the pursuit and achievement of long-term goals or projects. Indeed, there is good empirical evidence that even very intense pleasures on particular occasions add relatively little to a person's overall happiness compared with more stable and enduring sources of fulfillment or satisfaction. Furthermore, there seems to be a cognitive or judgmental aspect to happiness that is not captured by this exclusive focus on occurrent feelings. It seems plausible to think that how happy a person is must have something to do with how well that person thinks life is going, either as a whole or in important sectors (such as work, family, and health). Developing this line of thought has led philosophers and psychologists to develop a conception of happiness as life-satisfaction. This conception is still subjective since it takes someone's happiness to be a matter of how the person thinks life is faring given that person's interests and values. But it makes happiness more a matter of judgment than of feeling. Among psychologists Ed Diener has been the principal exponent of this life-satisfaction view.

So there are two principal subjective analyses of happiness: in terms of affective states or life-satisfaction. Each seems to capture a dimension of the phenomenon that eludes the other. On the one hand we would be reluctant to call anyone happy whose dominant state of mind tended toward the gloomy, dejected, or depressed. Here the affective state account seems to yield the intuitively right result. On the other hand we would be similarly reluctant if the subject were to report that in every important sector, life was failing to measure up to aspirations and expectations for it. This time the life-satisfaction account seems to be on the right track. Perhaps, then, the best theory of the nature of happiness will be a hybrid that takes both dimensions into account, looking for a preponderance of positive affect over time together with an endorsement of the conditions of one's life. Such a theory will not yield a determinate result if these dimensions can come apart, as seems both logically and psychologically possible. But in that case the right response might be to question whether our common notion of happiness is internally unified or whether it looks to both of these factors. If the latter is the case, then the hybrid theory might just be the best fit for it.

WHY HAPPINESS MATTERS: WELL-BEING

Whether construed objectively or subjectively, happiness has been thought to be normatively important either as a part of the good life or as an ethically valuable goal. When inquiring into the role of happiness in the good life, we must be mindful of the multiple ambiguities of this latter notion. A good life is a life high in some particular type of value, but there are many such types (such as aesthetic, perfectionist, and ethical). The dimension of value to which happiness seems most relevant is prudential: the value of a life for the person who is living it. But *prudential value* is a piece of technical philosophical terminology; its ordinary language equivalent is well-being.

Is happiness equivalent to well-being, thus to a good life, in this particular sense? There is some reason to think so. After all, we commonly wish people happiness at critical junctures in their lives, such as birthdays and anniversaries, seeming to imply that this is the best, most optimistic, hope we can have for them. Furthermore, there is much plausibility in the idea that any other of our life's conditions—health, income, job, family—is of little or no value if it does not make us happy. Perhaps then, as

Mill claimed, "happiness is desirable, and the only thing desirable, as an end; all other things being only desirable as a means to that end" (Mill 1969, p. 234).

But there is also reason to doubt the equation of happiness and well-being, especially if we assume a subjective conception of happiness. For one thing, since on either version of a subjective analysis happiness is a state of mind, it can in principle be synthesized in the manner made famous by Robert Nozick's science-fiction example of the experience machine. But we would be reluctant to think that the good life for a person could consist in a thorough-going illusion completely divorced from reality. In addition, people's subjective responses to their lives are notoriously subject to manipulation through such mechanisms as oppression and socialization. As a result, people's self-assessments of happiness will be dependent on their expectations for themselves, which may be artificially lowered through internalized conceptions of their social role or status. In order to correct for these distortions, some philosophers have suggested that self-assessed happiness is a reliable indicator of well-being only under conditions of information and autonomy. Others have taken the further step of suggesting that real or genuine happiness consists in the subject's endorsement of the right kinds of objects or states of affairs—those with independent value. Taking this route will lead to a hybrid subjective/objective theory, not of happiness (which is still interpreted subjectively), but of well-being.

WHY HAPPINESS MATTERS: ETHICS

As noted at the outset, happiness has been a central theme in ethical theories, both ancient and modern. The most natural route to treating happiness as an intrinsic ethical good is through its role in well-being. The argument would then be something like this: Well-being matters in its own right, happiness is at least an essential ingredient in well-being, therefore happiness matters in its own right. Some such argument, in one form or another, seems to have influenced both the Greeks and the utilitarians. However, it is also possible in principle (though perhaps less plausible) to hold that happiness is intrinsically valuable just in its own right, independently of its connection to well-being.

Where well-being is concerned, there is an easy answer to the question *whose* happiness matters: my happiness is central to my well-being, yours to yours, and so on. But the question takes on a more acute importance when we turn to ethics. For the Greeks, and for some contemporary versions of virtue ethics, the primary focus is on showing the agent how to live a good—that is to say, happy—life. The link to the happiness of others is then through the account of the virtues necessary for the agent's own happiness. This argumentative route, needless to say, is plausible only if we presuppose an objective conception of happiness.

For the utilitarians, by contrast, everyone's happiness is equally valuable: As Mill put it, citing Bentham, "everybody to count for one, nobody for more than one" (Mill 1969, p. 257). Another way of stating this contrast is that for the virtue theorists, happiness provides an agent-relative goal: Everyone has an ultimate reason to pursue their own happiness. For the utilitarians happiness provides an agent-neutral goal: Everyone has an ultimate reason to promote happiness, regardless of whose it is. In this respect Kant's deontological theory occupied the middle ground: We have no duty to pursue our own happiness (since Kant thought that we inevitably did that anyway), but we do have a duty to promote the happiness of others.

See also Aristotle; Bentham, Jeremy; Eudaimonia; Kant, Immanuel; Intrinsic Value; Mill, John Stuart; Nozick, Robert; Self-Interest; Sidgwick, Henry; Stoicism; Utilitarianism; Value and Valuation.

Bibliography

Annas, Julia. *The Morality of Happiness*. New York: Oxford University Press, 1993.

Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett, 1985.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Edited by J. H. Burns and H. L. A. Hart. London: Athlone, 1970.

Brandt, Richard B. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Frey, Bruno S., and Alois Stutzer. *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton: Princeton University Press, 2001.

Griffin, James. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press, 1986.

Hare, R. M. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Clarendon Press, 1981.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford: Oxford University Press, 1999.

Kahneman, Daniel, Ed Diener, Norbert Schwarz, eds. *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation, 1999.

Kant, Immanuel. *The Doctrine of Virtue*. Translated by Mary J. Gregor. Philadelphia: University of Pennsylvania Press, 1964.

Kraut, Richard. "Two Conceptions of Happiness." *Philosophical Review* 88 (2) (1979), 167–197.

McFall, Lynne. *Happiness*. New York: Peter Lang, 1989.

Mill, John Stuart. "Utilitarianism." In *Essays on Ethics, Religion and Society*. Edited by J. M. Robson. Toronto: University of Toronto Press, 1969.

Myers, David G. *The Pursuit of Happiness: Who is Happy—and Why*. New York: William Morrow, 1992.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Sidgwick, Henry. *The Methods of Ethics*. Chicago: University of Chicago Press, 1962.

Sumner, L. W. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press, 1996.

Telfer, Elizabeth. *Happiness*. London: Macmillan, 1980.

L. W. Sumner (2005)

HARDING, SANDRA

(1935–)

Sandra Harding is an American philosopher of science whose research interests include feminist and postcolonial theories, epistemology, and science studies. She received her PhD from New York University in 1973 and is a professor in the Graduate School of Education and Information Studies at UCLA. She is a former coeditor of *Signs: Journal of Women in Culture and Society* and former director of the UCLA Center for the Study of Women. Harding has authored four books and numerous articles and edited eight anthologies. She is best known for her work in developing feminist-standpoint theory. Initially focused on illuminating the gendered contexts of science, Harding has gone on to investigate other aspects of the social and cultural contexts of science, including its "racial" and colonialist contexts. Seeking to explore ways in which science can become a more significant force for human well-being, her work has analyzed various social and political contexts of science, including its implication in the exploitation of nature, non-Western cultures, and women.

Harding's work in the 1980s helped shape the landscape of developing feminist epistemology and feminist science. *Discovering Reality*, coedited with Merrill Hintikka (1983), and Harding's *The Science Question in Feminism* (1986) were groundbreaking efforts applying gender to epistemology and the philosophy of science. In *The Science Question*, Harding analyzes then-current feminist epistemologies and their ability to justify feminist science critiques. Although she urges ambivalence toward the frameworks, she suggests that feminist-standpoint theory is the most promising. Standpoint theory traces its roots to Hegel's argument, later developed in Marxist theory, that divisions in power yield corresponding divisions in worldviews: those in dominant positions have a distorted worldview that suggests their privilege is "natural," and those subordinated have the potential to achieve a less distorted view of the relevant social relations. Early feminist-standpoint theory proposed that men and women are, respectively, disadvantaged and potentially advantaged in this sense and stressed the role of the women's movement in helping women achieve a less distorted, feminist standpoint.

In *The Science Question* Harding identifies several problems in then-current versions of feminist-standpoint theory. One is that the theory assumes that there are experiences that are unique to women *qua* women, but Harding argues that is unlikely given differences in race, class, sexuality, and culture, among other factors. Another problem she notes is that there are as many standpoints as there are substantial divisions in power, unbridgeable chasms between the worldviews of those in dominant positions and those subordinated in a social hierarchy.

In *Whose Science? Whose Knowledge? Thinking from Women's Lives* (1991), Harding argues that those politically advantaged can and should come to understand the lives and perspectives of those not. This broadening of the notion of a standpoint has several consequences. It allows Harding to argue not that women as a group have an epistemic advantage over scientists but that, if scientists were to begin to research from the perspective of *women's* lives, new questions would emerge, along with data and theories that would prove more fruitful scientifically and socially. Harding also embraces the implication of multiple standpoints and contends that these are not unbridgeable. Each of us can work to "reinvent ourselves as 'others'" both to understand other standpoints and better understand the partiality and specificity of our own perspectives.

These several lines of argument come together in Harding's account of objectivity. In contrast to a traditional emphasis on a scientist or scholar's detachment from social contexts, Harding advocates what she calls "strong objectivity." To be objective in this sense requires a "robust reflexivity" that would oblige scientists and philosophers of science seek an understanding of the parochialism of the contexts within which their science and culture have coevolved. In *Is Science Multicultural? Postcolonialisms, Feminisms, and Epistemologies* (1998), Harding maintains that such reflexivity requires literacy in the sociology of science and social histories of science literatures, postcolonial and feminist science studies, and other critical science literatures. Only when scientists and science studies scholars achieve such reflexivity, Harding argues, will it be possible for the sciences to change in ways that will enable them to become an unproblematic and significant force for human well-being.

See also Feminism and the History of Philosophy; Feminist Epistemology; Feminist Philosophy; Feminist Philosophy of Science; Hegel, Georg Wilhelm Friedrich; Marxist Philosophy; Philosophy of Science, History of; Women in the History of Philosophy.

Bibliography

With Merrill Hintikka. *Discovering Reality* (1983). 2nd ed. Boston, MA: Kluwer Academic, 2003.

The Science Question in Feminism. Ithaca: Cornell University Press, 1986.

Whose Science? Whose Knowledge? Thinking from Women's Lives. Ithaca, NY: Cornell University Press, 1991.

Is Science Multicultural? Postcolonialisms, Feminisms, and Epistemologies. Bloomington: Indiana University Press, 1998.

ADDITIONAL EDITED COLLECTIONS BY HARDING

Can Theories Be Refuted? Essays on the Duhem-Quine Thesis. Boston: D. Reidel, 1976.

Feminism and Methodology. Bloomington: Indiana University Press, 1987.

With Jean O'Barr. *Sex and Scientific Inquiry*. Chicago: University of Chicago Press, 1987.

The "Racial" Economy of Science: Toward a Democratic Future. Bloomington: Indiana University Press, 1993.

With Uma Narayan. *Decentering the Center*. Bloomington: Indiana University Press, 2000.

The Feminist Standpoint Theory Reader: Intellectual and Political Controversies. New York: Routledge, 2003.

With Robert Figueroa. *Science and Other Cultures*. New York: Routledge, 2003.

Lynn Hankinson Nelson (2005)

HARE, RICHARD M.

(1919–2002)

Richard M. Hare, the White's Professor of Moral Philosophy at Oxford University from 1966 to 1983, is famous as the inventor of universal prescriptivism. This is a metaethical doctrine, a thesis about what moral words mean. But Hare uses his metaethic to generate an ethic. Anyone who employs the moral concepts consistently in full awareness of the facts must wind up a utilitarian. Hare claims that his utilitarianism is the product of conceptual analysis rather than of moral intuition. To rely on intuitions is a philosophical sin, since it leads to relativism (Hare 1981). His theory is developed in three books, *The Language of Morals* (1952), *Freedom and Reason* (1963), and *Moral Thinking* (1981).

Prescriptivism is a variant of noncognitivism. Moral judgments are action guiding, and the explanation of this is that they are prescriptive: They are not primarily designed to state facts but to prescribe actions. They are more akin to orders than statements or propositions. Nevertheless, moral judgments do have descriptive content, though this will depend upon the moral opinions of the speaker (Hare 1963). Thus, if Captain Bligh says that Burkitt is a scoundrel, we can assume he is disobedient. Indeed, even words such as ought have descriptive content, though this too will vary with the moral opinions of the speaker. Typically, the descriptive content of an ought judgment will consist in the factual considerations—the reasons—that can be advanced in its support. Thus, if Bligh asserts that Burkitt ought to be flogged, this will be because it would be an act of punishing disobedience. That the flogging would be such an act is the descriptive content of "Burkitt ought to be flogged." (Whence it follows that, if Burkitt has not been disobedient, the ought judgment will be factually false.) In Hare's view moral judgments are universalizable. Thus, if Bligh thinks that Burkitt ought to be flogged, he is committed to the view that anyone in relevantly similar circumstances—anyone who has been similarly disobedient to a king's officer—ought to be flogged likewise. He must assent to the imperative "Let me be flogged in the hypothetical case in which I am in Burkitt's position!"—which includes having committed Burkitt's heinous acts of disobedience (Hare 1963). Finally, moral judgments are overriding. They take precedence over any other imperatives the subject may accept. Thus, if Bligh thinks himself morally obliged to have Burkitt flogged, this takes precedence over his aesthetic obligation not to sully the pure air of the Pacific with Burkitt's distasteful groans. Sincere moral commitment entails action. Weakness of the will as traditionally conceived is not a genuine possibility. Thus, Hare reinstates the Socratic paradox that we cannot willingly do wrong (Hare 1952, 1963).

What about utilitarianism? Hare first points out that the metaethic generates a method for refuting moral "conjectures." Bligh considers the maxim "I ought to have Burkitt flogged." He universalizes this to derive the principle that anyone in relevantly similar circumstances ought to be flogged likewise. This in turn entails the imperative "Let me be flogged if I am in Burkitt's position!" But Bligh cannot assent to this unless he is a fanatic—someone who prefers flogging the disobedient to remaining unflogged himself. Thus, Bligh must rescind his original "ought" (Hare 1963). But this is only a method for vetoing moral maxims and a method, moreover, that leads to moral paralysis. As Hare himself points out, a guilty prisoner could challenge the judge to universalize the maxim that the accused ought to be put away and derive the imperative "Let me be imprisoned if I am

in the accused shoes!"—an imperative she could accept only if she had a fanatical preference for imprisoning the guilty rather than staying out of jail herself (Hare 1963). Nonfanatical judges would have to give up sentencing and justice would founder! But Hare offers a utilitarian solution. The correct course is to go the rounds of the affected parties and opt for the action that is subject to the weakest veto[es]. Thus, the judge must take into account the likely depredations of the prisoner and ask herself whether she can accept such imperatives as "Let me be robbed if the prisoner is released and allowed to carry on with his course of crime and I am one of his victims!" If not, and if the vetoes of the prisoner's potential victims outweigh his preference not to go to jail, then to jail he must go. The criminal-justice system can survive without fanaticism, and Hare's method becomes utilitarian. But does Hare derive utilitarianism from his conceptual analysis or assume utilitarianism to rescue that analysis from disaster (Roxbee Cox 1986)?

The fanatic remains a problem. She can consistently subscribe to a persecuting principle if she assents to the imperatives in which she is on the sharp end. In *Moral Thinking* Hare deprives her of this possibility. He claims it is a conceptual truth that if I fully represent to myself what an unpleasant experience is like for someone—an experience that they would prefer to stop—I now acquire an equally strong preference not to have that experience were I in their shoes. Hence, a fanatic who fully represents to herself the sufferings of her potential victims cannot assent to the imperative that she should suffer were she in their position. For she has a preference as strong as theirs that she should not. If, however, Hare's conceptual truth is neither conceptual nor a truth, then fanaticism remains an option (Seanor and Fotion 1991).

See also Metaethics; Noncognitivism; Utilitarianism.

Bibliography

WORKS BY HARE

The Language of Morals. Oxford: Clarendon Press, 1952. Hare sets forth his metaethic.

Freedom and Reason. Oxford: Clarendon Press, 1963. Hare develops his metaethic and devises an engine of moral argument to enforce utilitarian conclusions. But he admits the fanatic is immune.

Moral Thinking: Its Levels, Methods, and Point. Oxford: Clarendon Press, 1981. Hare draws a distinction between intuitive and critical thinking and invents a conceptual truth to dispose of the fanatic. Only amoralists can avoid utilitarianism.

Essays in Ethical Theory. Oxford: Clarendon Press, 1989a. Hare develops and defends his ideas and attacks his philosophical rivals. Contains a memorably savage critique of John Rawls for relying on intuition, a reply to J. L. Mackie, his chief Oxford opponent, and a bibliography of his extensive writings.

Essays on Political Morality. Oxford: Clarendon Press, 1989b. A collection on applied ethics.

WORKS ON HARE

Roxbee Cox, J. W. "From Universal Prescriptivism to Utilitarianism." *Philosophical Quarterly* 36 (1986): 1–15. Challenges Hare's derivation of utilitarianism.

Seanor, D., and N. Fotion, eds. *Hare and Critics: Essays on Moral Thinking with Comments by R. M. Hare.* Oxford: Clarendon Press, 1991. The best essays are by Brandt, Singer, Gibbard, and Vendler. Hudson's essay lists some of the important criticisms Hare has faced. Extensive bibliography.

Singer, P. *Practical Ethics.* 2nd ed. Cambridge, U.K.: Cambridge University Press, 1993. Hare's method is applied to practical affairs by his most gifted disciple.

Taylor, C. C. W. "Critical Notice of R. M. Hare's *Freedom and Reason.*" *Mind* 74 (1965): 280–290. Still perhaps the best critique of Hare.

Charles R. Pigden (1996)

HARMAN, GILBERT
(1938–)

Gilbert Harman was born in 1938, graduated from Swarthmore College in 1960, and received his PhD from Harvard in 1963, where W. V. Quine was his dissertation advisor. He is distinctive in being a leading contributor across a broad range of subdisciplines of philosophy: epistemology, ethics, philosophy of language, philosophy of mind, and metaphysics. This entry reviews only a few of his many important contributions.

Harman has been perhaps the most significant contemporary defender of moral relativism. According to Harman, moral right and wrong are akin to motion: They are *relative* to a framework and no framework is privileged. Harman appeals effectively to two sorts of consideration in developing his position. First, like J. L. Mackie, he is impressed by the degree of moral diversity across and even within populations. Second, complementarily, Harman defends moral naturalism—the view, roughly, that morality is fundamentally continuous with the natural sciences.

In rejecting moral nonnaturalism, Harman claims that the postulation of nonnatural moral properties is unjustified: They would be explanatorily impotent. Could not a sui generis fact of, say, torture's being bad explain, at least, our belief that torture is bad? Harman argues that there are other better ways of explaining such a belief, in terms of conventions and other social arrange-

ments—arrangements to which we are normally exposed as we develop.

Holding that moral disagreement is widespread, Harman accordingly argues—as a kind of inference to the best explanation—to the conclusion that there are no absolute moral facts, beyond the facts about what holds relative to one or another framework. Different people can, without ignorance (or related independently specifiable failings), find one or another ultimate moral demand inapplicable in their own case. But because a moral demand is said to apply only if the agent either accepts it or rejects it only out of ignorance, then ultimate moral demands may apply only selectively, to some agents and not to others.

A related aspect of Harman's moral relativism is thus his motivational internalism: If morality is understood as the product of a framework constituted by psychological states (and, in particular, by the agreement, plans, and conventions emerging therefrom), then it will be easier to understand how morality could have that motivational force.

It should be noted, too, that, although he distinguishes them, Harman embraces forms of each of *normative moral relativism*, *moral judgment relativism*, and *metaethical relativism*. Normative moral relativism holds, roughly, that people can be subject to different ultimate moral demands. Moral judgment relativism claims, in effect, that moral judgments implicitly refer to a person, group, or set of moral demands. And according to metaethical relativism, conflicting moral judgments about a particular case can in a way both be right.

In epistemology, Harman has long defended a view that has elements of foundationalism and elements of coherentism. Harman's is a kind of "foundations" theory in which everything a person accepts at a given time is foundational and needs no justification except when there are conflicts. Accordingly, knowledge is best understood "when skepticism is turned on its head": starting from what we know, we diagnose what goes wrong with arguments for radical skepticism. For Harman, a key insight is that knowledge is essentially inferential: Inference is a matter of increasing the coherence of one's overall state—reasoning consists in trying to obtain a reflective equilibrium (though he is concerned about possible instabilities in this process)—and coherence is partly a matter of explanation.

If inference to the best explanation is to have the central role in our cognition that it appears to have, it will have to be understood as a sort of explanatory inference. So Harman defends the adequacy of that sort of inference. In enumerative induction we generalize observed regularities; but according to Harman all such cases of induction are really cases of inference to the best explanation. So, at a minimum, inference to the best explanation is not in general any less legitimate than induction.

This emphasis on the role of inference in knowledge is related to Harman's focus on inference in its own right. He draws a sharp division between logic and inference. For Harman, there is, for example, no such thing as deductive inference and the search for an inductive logic is the product of confusion. Reasoning is change in view; logic is the theory of implication.

Consider *modus ponens*. This exceptionless rule of logic cannot serve as a principle about how to change one's view: sometimes, when one believes *P* and believes that *if P, then Q*, what one should do is to give up one's belief that *P*. Moreover, although according to the rules of logic inconsistent premises *imply* any proposition, it's not the case that inconsistent beliefs permit one to *infer* any proposition. The distinction between inference and logic coheres well with Harman's views about the fundamentally *explanatory* character of inference to the best explanation (see above).

In philosophy of mind, Harman was a seminal proponent of what has come to be known as intentionalism about experiential states. Harman holds that there is no phenomenal difference between such states without an intentional difference. They have accordingly come to be viewed as individuated by their representational or intentional character: what makes the state what it is is that it is about what it is about. That perceptual experience should not be understood as individuated by (what others would call) its *qualitative* character instead is seen as sustained by the *transparency* of experience. Introspection does not seem to reveal the nature of experiences themselves, only that of their external objects.

This brief review cannot do justice to the richness and range of Harman's work. We have not so much as touched on his important discussions of meaning and analyticity (and his brilliant exposition—really, development—of Quine's views on the subject), on his contributions to conceptual-role functionalism in philosophy of mind, or on his more recent argument that work in social psychology supports eliminativism about the central posits of any virtue theory. Still, some sense of the scope and significance of Harman's contributions should have emerged.

See also Epistemology; Ethics; Metaphysics; Philosophy of Language; Philosophy of Mind.

Bibliography

WORKS BY GILBERT

Change in View: Principles of Reasoning. Cambridge, MA: MIT Press, 1986.

"The Death of Meaning." In *Reasoning, Meaning, and Mind*. Oxford, U.K.: Clarendon Press, 1999.

Explaining Value and Other Essays in Moral Philosophy. Oxford, U.K.: Clarendon Press, 2000. Available from http://www.oxfordscholarship.com.

"The Intrinsic Quality of Experience." In *Reasoning, Meaning, and Mind*. Oxford, U.K.: Clarendon Press, 1999.

The Nature of Morality: An Introduction to Ethics. New York: Oxford University Press, 1977.

"Quine on Meaning and Existence, II: Existential Commitment." *Review of Metaphysics* 21 (1967): 343–367.

Reasoning, Meaning, and Mind. Oxford, U.K.: Clarendon Press, 1999. Available from http://www.oxfordscholarship.com.

Thought. Princeton, NJ: Princeton University Press, 1973.

Harman, Gilbert, and Judith J. Thomson. *Moral Relativism and Moral Objectivity*. Oxford, U.K.: Blackwell, 1996.

David Sosa (2005)

HARNACK, CARL GUSTAV ADOLF VON

(1851–1930)

Carl Gustav Adolf von Harnack, the German church historian and theologian, was born at Dorpat (now Tartu) in Estonia, where his father, Theodosius Harnack, was a professor of practical theology at the German-dominated university. Carl Harnack studied at Dorpat and then at Leipzig, becoming a *Privatdozent* there in 1876. He held chairs at Giessen from 1879 and Marburg from 1886 before going, in 1888, to Berlin, where he was a professor until his retirement, in 1924. He died at Heidelberg.

Harnack has come to be regarded as the typical representative of liberal theology. Following Albrecht Ritschl and the members of his school, Harnack stressed the ethical teaching of Christianity and avoided the more speculative flights of theology, but he went further than his predecessors in the direction of an undogmatic, practical statement of the Christian faith. Harnack's appointment to the chair at Berlin was opposed by conservative elements in the Lutheran Church, but by the time he retired he had trained a whole generation of students in the ways of liberal theology and in what he believed to be the unprejudiced pursuit of theological truth. His last years were spent in opposing the nascent "dialectical theology" of the school of Karl Barth, which he saw as threatening the scientific character of the discipline.

THE PROBLEM OF DOGMA

The vigor with which Harnack advocated the cause of liberal theology was matched by his vast erudition. Few Protestant scholars have equaled his knowledge of early Christian history and literature. One of Harnack's major works, the monumental *Lehrbuch der Dogmengeschichte* (3 vols., Freiburg, 1886–1889; many subsequent eds. published at Tübingen; 3rd ed. translated as *History of Dogma,* 7 vols., London, 1894–1899), not only gives a detailed account of the history of Christian dogma, especially in the early formative centuries, but also expounds a definite thesis concerning the nature and development of this dogma.

As Harnack understood it, religion is primarily a practical affair and aims at the right ordering of life. In Christianity, the power of achieving a well-ordered or blessed life had its origin in Jesus Christ and the revelation of God that he brought. But although religion has this practical character, it also implies certain beliefs concerning God, man, and the world; the religious man seeks to make his beliefs explicit and to formulate them in propositions. This happens especially when a religious community comes into being and subscription to the basic beliefs of the community is made the condition of membership—hence the rise of dogma in the early church.

However, Harnack regarded this development as a perversion of the original teaching of Jesus, obscuring its essentially practical character and destroying its spontaneity. On the whole, he saw the history of Christian thought as one of deterioration, a falling away from the original truth rather than an unfolding of it. The process began when the primitive preachers made Jesus himself, as the supernatural Christ, the center of their message, rather than simply repeating Jesus' teaching about the kingdom of God, which Harnack understood as an ethical ideal. The transformation of Christianity into dogma accelerated in the Hellenistic world; the extreme case can be seen in the Gnostic sects, where the supposedly original gospel of Jesus was altogether absorbed into Hellenistic philosophy. With the Reformation an attempt was made to emancipate Christianity from dogma, but it was only partially successful, and dogma persisted into Protestantism.

THE ESSENTIALS OF CHRISTIANITY

In a series of popular lectures, which attracted huge audiences at the university of Berlin in the winter of 1899/1900 and was subsequently published as *Das Wesen des Christentums* (Leipzig, 1900; translated as *What Is Christianity?*, London, 1901), Harnack expounded what he believed to be the core of the Christian religion, set free from the encrustations of dogma that had been laid down through the centuries. The core is to be reached by penetrating back to the teaching of Jesus himself, and Harnack represented this teaching as proclaiming the fatherhood of God, the infinite worth of the human soul, and the ethical ideal of the kingdom of God. The supposedly original gospel of Jesus is also claimed to be the only version of Christianity that can make sense for modern minds, since it is free from theological and metaphysical mystifications.

Harnack's views once commanded a wide following, but this, however, has declined sharply in more recent times, owing to the criticism of such scholars as Alfred Loisy, Albert Schweitzer, and Karl Barth.

See also Barth, Karl; Liberalism; Loisy, Alfred; Reformation; Ritschl, Albrecht Benjamin.

Bibliography

ADDITIONAL WORKS BY HARNACK

Harnack's literary output was immense, consisting not only of books and articles but also of critical editions of many early religious documents. In addition to the two books mentioned in this article, the following are important:

Das apostolische Glaubensbekenntnis. Berlin, 1892. Translated as *The Apostles' Creed.* London: A. and C. Black, 1901.

Die Geschichte der altchristlichen Literatur bis Eusebius. 2 vols. Leipzig, 1893–1904.

Beiträge zur Einleitung in das Neue Testament. 4 vols. Leipzig: J. C. Hinrichs, 1906–1911. Translations of parts of this work: *Luke the Physician.* London: Williams and Norgate, 1907. *The Sayings of Jesus.* London, 1908. *The Acts of the Apostles.* London, 1909.

Entstehung und Entwicklung der Kirchenverfassung und des Kirchenrechts in den zwei ersten Jahrhunderten. Leipzig: J. C. Hinrichs, 1910. Translated as *The Constitution and Law of the Church in the First Two Centuries.* London, 1910.

Marcion. Leipzig: J. C. Hinrichs, 1921.

Ausgewählte Reden und Aufsätze. Berlin, 1951.

Die Mission und Ausbreitung des Christentums in den ersten drei Jahrhunderten. 4th ed., Wiesbaden: VMA-Verlag, 1965.

The Expansion of Christianity in the First Three Centuries. Edited and translated by James Moffatt. Freeport, NY: Books for Libraries Press 1972.

History of Dogma. Gloucester, MA: P. Smith, 1976.

Das Wesen des Christentums. Gütersloh: Gütersloher Verlagshaus Mohn, 1977.

Adolf Harnacks Konflikt mit der Kirche: Weg-Stationen zum "Wesen des Christentums". Edited by Karl H. Neufeld. Innsbruck: Tyrolia, 1979.

Militia Christi: The Christian Religion and the Military in the First Three Centuries. Philadelphia: Fortress Press, 1981.

What Is Christianity? Philadelphia: Fortress Press, 1986.

Moderne Theologie: Der Briefwechsel Adolf von Harnack, Christoph Ernst Luthardt, 1878–1897. Neukirchen-Vluyn, Germany: Neukirchener Verlag, 1996.

Adolf von Harnack: Christentum, Wissenschaft und Gesellschaft: Wissenschaftliches Symposion aus Anlass des 150. Geburtstages. Edited by Kurt Nowak, et al. Göttingen: Vandenhoeck & Ruprecht, 2003.

Monasticism: Its Ideals and History; and, The Confessions of St. Augustine: Two Lectures. Eugene, OR: Wipf & Stock, 2004 .

WORKS ON HARNACK

Zahn-Harnack, Agnes von. *Adolf von Harnack.* Berlin: H. Bott, 1936. The standard biography of Harnack. See also her article on Harnack in *Die grossen Deutschen,* Vol. IV, 255–266. Berlin: Propyläen, 1957.

John Macquarrie (1967)

HARRINGTON, JAMES
(1611–1677)

The English political philosopher and publicist James Harrington was the eldest son of Sir Sapcote Harrington, of Rand, Lincolnshire. As such, he belonged to a junior branch of a family that had been prominent from the days of Richard I. An erudite man, Harrington must have acquired his great knowledge of languages, literature, and history largely independently, since he spent only two or three years at Oxford and in the Middle Temple and took no degree. During the 1630s he traveled extensively on the Continent and served in an English volunteer regiment in the forces of one of the palatine electors. From these experiences, and especially from a visit to Venice, he gathered much of the data that later formed the raw material for his political theory.

When civil war broke out in England, Harrington took a neutral position, despite his republican sympathies, because of his personal regard for the king, and at one point attempted the role of mediator between royal and parliamentary interests. But after Charles I's execution in 1649 he devoted himself to the construction of a republican political theory, which culminated in 1656 in the publication of his major work, *The Commonwealth of Oceana,* a blueprint for a perfect republic.

He was imprisoned by Charles II in 1661 on a false charge of treason. His mind became deranged while he was in prison, and he never fully recovered his faculties

after his release. He died at his Westminster home in 1677.

Although *Oceana* has the form of a utopia, Harrington stands squarely in the British empiricist tradition. Even Niccolò Machiavelli, whom he admired as "the only Polititian of later Ages," he criticized for violating the canons of empiricism by using such concepts as "virtue" and "corruption," which Harrington held to be meaningless as analytical tools.

Harrington's own concepts are sociological, rather than psychological or ethical. A stable governmental system always represents the dominant property-owning groups of a society. Where political and economic power are held by the same hands, and a single person controls three-fourths of all the property, the political system will be an absolute monarchy. If a few hold three-fourths of the property, it will be a mixed monarchy. If property is so dispersed that no monopoly vests in a single social interest, the system will be a republic, or "commonwealth."

Harrington made no moral ranking of the forms of government. Words such as *tyranny, oligarchy, anarchy* he used descriptively rather than evaluatively, to signify unstable governmental forms that do not match their foundations, those in which power is held incommensurately to the distribution of property. The theoretical question that preoccupied him was stability, and the chief cause of revolution and civil war he identified as an incongruence between social balance and form of government. Conflict he viewed as a mechanism for bringing the two into close proportion. He was not an economic determinist, however, for he thought it just as possible to "frame the foundation unto the Government" as the reverse.

Like Machiavelli, Harrington preferred the republican system to all others. He wrote of "that *Reason* which is the interest of mankind, or of the whole," as "a Law of Nature," and described "the publick interest of a commonwealth" as "nearest that of mankind." Since he espoused a radically hedonic view of human motivation, this must mean he preferred republics because in them the things men enjoy are more widely distributed than in systems with a narrower property base. Not absolute equality but a middle-class order is implied, however. For "leveling" impedes economic growth and the social accumulation of the riches humankind desire.

More sanguine than Machiavelli, Harrington thought it possible to create a perfectly stable and unchanging republic. It could be maintained by an "equal Agrarian" law, fixing forever a middle-class distribution of property, and by arranging a suitable balance of interests in the organization of the government through such devices as separation of powers, division of the legislature, and rotation in office.

See also Empiricism; Machiavelli, Niccolò; Social and Political Philosophy.

Bibliography

PRIMARY WORKS

James Harrington's Oceana. Edited by S. B. Liljegren. Heidelberg: Winter, 1924.

The Political Works of James Harrington. Edited by J. G. A. Pocock. Cambridge, U.K.: Cambridge University Press, 1977.

SECONDARY WORKS

Blitzer, Charles. *An Immortal Commonwealth: The Political Thought of James Harrington.* New Haven, CT: Yale University Press, 1960.

Downs, Michael. *James Harrington.* Boston: Twayne, 1977.

Fink, Zera S. *The Classical Republicans.* Evanston, IL: Northwestern University, 1945.

Fukuda, Arihiro. *Sovereignty and the Sword: Harrington, Hobbes, and Mixed Government in the English Civil Wars.* Oxford: Clarendon Press, 1997.

Russell-Smith, H. F. *Harrington and His Oceana.* Cambridge, U.K.: Cambridge University Press, 1914.

Tolan, John. *The Oceana of James Harrington and His Other Works.* 4th ed. London: 1771.

Wershofen, Christian. *James Harrington und sein Wunschbild vom germanischen Staate.* Bonn, 1935.

William T. Bluhm (1967)
Bibliography updated by Philip Reed (2005)

HARRIS, WILLIAM TORREY
(1835–1909)

The American philosopher and educator William Torrey Harris was born in North Killingly (now part of Putnam), Connecticut. He attended preparatory schools in his native state and entered Yale College. There he was led to philosophy by Bronson Alcott's "Conversations" on Platonism, which convinced him of "the ideality of the material world" through "insight and reliance on reason." He left Yale in his junior year, dissatisfied with the deficiency of modern science and literature in the curriculum, and went to St. Louis.

In St. Louis, where Harris taught school for eight years and was an administrator for fourteen, he met Henry C. Brokmeyer, a Prussian immigrant who had acquired an enthusiasm for G. W. F. Hegel from reading F.

H. Hedge's *Prose Writers of Germany* (1847) during some disputatious months at Brown University. In 1858, Harris, Brokmeyer, and a few friends began meeting informally as a Kant Club to find the root of Hegel's thought. Harris imported a copy of Hegel's larger *Logic* and encouraged Brokmeyer to undertake a translation, which was never satisfactorily finished but was circulated in manuscript. After the Civil War, adherents of the Kant Club joined the St. Louis Philosophical Society, organized in 1866 with Brokmeyer as president, Harris as secretary, and Denton Snider, G. H. Howison, A. E. Kroeger, and Thomas Davidson among the leading members.

When the editor of the *North American Review* rejected one of Harris's articles as "the mere dry husk of Hegelianism," Harris and the St. Louis Society founded the *Journal of Speculative Philosophy.* Edited by Harris from 1867 to 1893, the *Journal* published numerous translations of German philosophers, particularly Hegel, and original essays by Ralph Waldo Emerson, J. H. Stirling, James Ward, William James, John Dewey, and C. S. Peirce. In defending Hegel's views in America, Harris and Brokmeyer had been preceded by a group of Ohioans that included J. B. Stallo and August Willich, who became "auxiliaries" of the St. Louis Society, as did Emerson, Henry James Sr., Karl Rosenkranz, and Ludwig Feuerbach. But Harris was outstanding among American philosophers up to 1900 as an active public lecturer, a leader of the St. Louis movement and of the Concord School of Philosophy from 1879 to 1887, U.S. commissioner of education from 1889 to 1906, editor of America's first regular journal devoted to philosophy, and author of some five hundred articles and of a book on Hegel's *Phenomenology* and *Logic.*

Like Hegel, Harris saw philosophy as a science concerned with necessary factors in experience related systematically to a first principle. Reflection on sensible objects and their changes, he believed, immediately reveals two necessary factors with which philosophy is concerned, space and time. Both are "infinites" in that they are conditions of all experience. From a parallel analysis he concluded that there are three grades or stages of knowing. The first concentrates on the object and the surface of things as isolated and independent. The second sees how things exist only in relation to other things and thus concentrates on their dependence, on what they are not when taken by themselves as separate and isolated. The third "discovers the independence and self-relation underlying all dependence and relativity"; in discovering what is self-related it discovers "the infinite." These mutually related stages are to be found in every aspect of experience, and since there are no things-in-themselves behind experience, they characterize all aspects of our world. Harris thus attempted to put into plain English the main features of Hegel's dialectic. Through Brokmeyer, Harris came to believe that such dialectic illuminated the Civil War (legal right would be unified with moral right), American politics, and even problems of school administration—a use of philosophy that pleased the practical, institution-minded members of the St. Louis movement.

Proceeding dialectically from "seeming" to "truth," Harris analyzed causality and concluded that it incorporates space and time in a higher unity but also implies a "self-separation" of energy whereby a cause sends a stream of influence to other things. Without such self-separation a cause could not act upon something to bring about an effect. So conceived, causality must be grounded in "self-activity," which is necessarily self-related and thus independent, free, and creative. Ultimately, in Harris's view, the only authentic self-activity is God, conceived by Harris, following Aristotle and Hegel, as the unmoved motion and self-contained existence of Reason, which, as Reason, is also personal. Like Hegel, Harris believed that philosophy approaches Absolute Reason through conceptual analysis to first principles, whereas religion receives the Absolute "into the heart" through symbols.

As a corollary to the presupposition of relatedness in self-activity, Harris saw education as the self-development of the individual mediated through the salient traditions of civilization. With the self-development of the individual in view, he linked public schools with democracy, conceived of as self-government involving woman's suffrage and separation of religion from the state. With the traditions of civilization in mind, he criticized excessive vocationalism. Along similar lines, his social philosophy viewed civilized freedom as the will of the individual effectuated in such institutions as family, civil society, state, and the Invisible Church, the "absolute institution" uniting all people of all time. In spite of his stress on institutions, Harris apparently gave some kind of precedence to "self-activity" *simpliciter*; he admired the ruthless individualism of the "gilded age" and condemned socialism in all its aspects.

See also Absolute, The; Aristotle; Dewey, John; Emerson, Ralph Waldo; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Hegelianism; Howison, George Holmes; Idealism; James, Henry; James, William; Peirce, Charles Sanders; Rosenkranz, Johann Karl Friedrich; Ward, James.

Bibliography

WORKS BY HARRIS

Hegel's Logic, a Book on the Genesis of the Categories of the Mind. Chicago: S.C. Griggs, 1890. The preface outlines the genesis of the author's views in relation to Immanuel Kant and Brokmeyer.

Introduction to the Study of Philosophy. Edited by Marietta Kies. New York: Appleton, 1889. Topically arranged selections from Harris's numerous articles.

WORKS ON HARRIS

Leidecker, Kurt. *Yankee Teacher, the Life of William Torrey Harris.* New York: Philosophical Library, 1946. The most complete biography.

Pochmann, Henry A. *German Culture in America, Philosophical and Literary Influences,* 257–304. Madison: University of Wisconsin Press, 1957. Details on the St. Louis movement and the Concord School of Philosophy.

Schaub, Edward L., ed. *William Torrey Harris.* Chicago: Open Court, 1936. Essays by American philosophers on major aspects of Harris's life and thought.

Loyd D. Easton (1967)

HART, HERBERT LIONEL ADOLPHUS

(1907–1992)

Herbert Lionel Adolphus, professor of jurisprudence at Oxford University (1952–1968), was the most important and influential philosopher of law of the twentieth century. Bringing to bear the linguistic approach to philosophy championed by Wittgenstein and Hart's Oxford colleague, J. L. Austin, Hart transformed jurisprudence into the vibrant discipline it had been at the time of Jeremy Bentham and his student, John Austin. He revealed the law to be a fertile ground for addressing age-old philosophical questions on a wide range of topics, for example, the analysis of causation, human action and intention, responsibility and rights, and the very nature of morality. Ronald Dworkin, who succeeded Hart in the Oxford Chair, nicely expressed this feature of his predecessor's work in a speech delivered at Hart's memorial service: "Herbert showed how philosophy can be tutor to law, how lawyers' questions about punishment and cause and definition have philosophical dimensions that it is irresponsible to ignore. He also showed how law can be tutor to philosophy, how legal problems, discriminations and attitudes can help philosophers in formulating and attacking those same ancient philosophical puzzles." (Hart, Jenifer, p. 213). Although Hart's writings cover a wide range of topics, his most memorable and influential contributions were in four main areas: causation, the theory of punishment, the moral limits of the law, and the concept of law.

CAUSATION

Causation in the Law (1959), written with A. M. Honore, is an impressive and original analysis of causation as it figures in Anglo-American legal systems. The authors proceed from the premise that then-extant philosophical analyses of cause and effect were largely inadequate because they focused on causation in science, where the concern is to establish causal laws and generalizations. Within the domain of law (and analogously morals), however, causation is more particularistic in nature, concerned with whether, for example, the defendant caused the death of the victim. Assessing such claims requires appeal to an array of different principles and individuating features, and attention to central or paradigm cases in which causal responsibility is confidently assessed. These factors are all in various complex ways connected with ordinary understandings of causation as these are reflected in our linguistic practices. It is here, perhaps more than anywhere else, that Hart's debt to J. L. Austin's "ordinary language philosophy" is in evidence. This is a method in which, as Hart later said in *The Concept of Law*, "We are using a sharpened awareness of words to sharpen our perception of the phenomena" (Austin 1956–1957, p. 8).

THEORY OF PUNISHMENT

Punishment and Responsibility (1968) is a collection of essays in which Hart develops a distinctive theory of punishment. The account was motivated by two factors. First, there is the potential for abuse seemingly inherent in utilitarian theories of punishment. In the right circumstances, they appear to sanction excessive punishment, as well as punishment of the innocent. Second, there was Hart's utter rejection of Kantian retributivist theories that view punishment as warranted independently of any good that might be brought about through its exercise. In a characteristic effort to seek compromise, Hart sought the middle ground between the two theories. Retributivist principles should influence the distribution of punishment: Only the guilty should be punished and only to the degree that they deserve. But the utilitarian goal of general deterrence remains to justify the overall practice.

THE LAW'S MORAL LIMITS

Law, Liberty and Morality (1963) represents a brilliant statement of Hart's liberal views regarding the role of law in enforcing morality, views that echo those of his distin-

guished liberal predecessor, John Stuart Mill. The book summarizes and extends Hart's contribution to a famous debate with Lord Patrick Devlin who, in response to an official call for legalizing prostitution and homosexuality in England, argued that a society has a right to enforce its morals because a solid moral foundation is as essential to its survival as a firm political structure. In Devlin's view, society has as much a right to enforce its morals through legal means as it has the right to protect itself through laws against sedition and treason. Hart was thoroughly repelled by Devlin's legal moralism. There is little reason, he argued, to believe that failure to enforce widely shared moral beliefs inevitably leads to social disintegration. Devlin also fails to recognize the distinction between positive morality (the morality widely shared within a society) and critical morality (more enlightened, rational standards for assessing both human conduct and, crucially, a society's positive morality). Most importantly, Devlin fails to appreciate the important role of the latter in challenging positive morality and in keeping alive the animating spirit of morality—the belief that certain standards of behavior should be followed, not because they are widely accepted and enforced, but because adhering to them voluntarily is the right thing to do. In Hart's view, critical morality includes a mixture of principles and values, some of which have utilitarian roots, others of which are of a more deontological bent. These standards, he thought, reveal that the coercive hand of the law should be used only to prevent palpable harm to others, or for the sake of a limited set of paternalistic goals.

THE CONCEPT OF LAW

Hart's most memorable work was undoubtedly in the area of general legal theory, where "Positivism and the Separation of Law and Morals" (1957–1958) and *The Concept of Law* (1961) stand as monumental contributions to our understanding of law and legal systems. These works develop a modern version of the legal positivism espoused by Bentham and Austin. Much as those two theorists had done, Hart sought to explain law as it *is*, not as it *ought to be*. In his view, natural law theories confuse these two issues, thus leading not only to philosophical confusion, but to anarchism (this law is not as it ought to be; therefore it is not really law and I am free to disregard it) or reactionary thinking (this is the law; therefore it must be as it ought to be and I must obey). The latter mind-set was of particular concern to Hart because it led, he thought, to a dimming of the vitally important sense that, for all its aura of majesty and authority, law's demands must always remain open to moral critique and challenge.

Though he shared this overall positivistic approach to law and legal analysis with Bentham and Austin, Hart departed dramatically from them on a number of other fronts. Most importantly, he thoroughly rejected their command theory according to which law is comprised of the general commands of an habitually obeyed sovereign whose directives are backed up with threat of sanction. The command theory reduces law to "the gunman situation writ large" (1994, p. 7), that is, it views our situation under law as analogous to being obliged to surrender our money to a gun-toting robber. But we do not always view laws this way. Many of us take a more "internal point of view" (1994, p. 89) toward them. While we may feel obliged to hand over our money to the gunman, many—especially the officials of the legal system—view laws as imposing legitimate reasons for action. Furthermore, not all laws demand or prohibit conduct. Rather they facilitate our doing certain good things, such as entering into contracts and creating valid wills. As such, they are grossly mischaracterized if conceived as orders backed by threats.

CONCLUSION

So law is not the gunman situation writ large: It is the "union of primary and secondary rules" (1994, p. 99). Unlike the standards of morality and the social mores of etiquette and fashion, legal systems have a formal structure created by the interplay of primary rules (of duty and obligation) and certain fundamental secondary rules (rules about other rules). Every legal system contains a secondary rule of recognition that specifies criteria of validity—for example, parliamentary enactment or conformity with a Charter of Rights—which all other rules of the system must meet if they are to count among its binding laws. There will also be secondary rules of change, through which existing rules are altered or replaced, and secondary rules of adjudication, which regulate the enforcement and application of legal rules, most notably by judges. These fundamental secondary rules are social rules whose existence and content depend crucially on the behavior of the officials who use them in the everyday workings of the system.

Understood in this way, law can be seen to be a human, social creation. Its existence and content are matters of social fact, determined by what is, not by what ought to be. Despite his firm commitment to this positivistic view of law's content, Hart was prepared to concede that there is a kernel of truth in rival natural law

theories. Law is a social institution whose existence depends on its acceptance; but there would be no reason to accept law were it somehow devoid of "minimum forms of protection for persons, property and promises" (1994, p. 199). There is therefore a "natural necessity" (1994, p. 199) that legal systems contain this minimum content. Beyond this humble minimum, however, Hart was not prepared to venture, thus affirming his conviction that the law, by its very nature, may fail to meet our moral expectations of it.

See also Analytic Jurisprudence; Austin, John; Austin, John Langshaw; Bentham, Jeremy; Causation; Dworkin, Ronald; Historical School of Jurisprudence; Legal Positivism; Mill, John Stuart; Natural Law; Paradigm-Case Argument; Philosophy of Law, History of; Philosophy of Law, Problems of; Punishment; Rights; Wittgenstein, Ludwig Josef Johann.

Bibliography

Austin, J. L. "A Plea for Excuses." *Proceedings of the Aristotelean Society* 57 (1956–1957): 1–30.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*, edited by J. H. Burns and H. L. A. Hart. London: Athlone Press, 1970. 2nd ed., Oxford, U.K.: Clarendon Press, 1996.

Bentham, Jeremy. *Of Laws in General*, edited by H. L. A. Hart. London: Athlone Press, 1970.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

Dworkin, Ronald. *Taking Rights Seriously*. London: Duckworth, 1977.

Hart, Jenifer. *Ask Me No More*. London: Peter Halban Publishers Ltd., 1998.

Raz, Joseph. *The Authority of Law: Essays on Law and Morality*. Oxford, U.K.: Clarendon Press, 1979.

Raz, Joseph. *The Concept of a Legal System*. Oxford, U.K.: Clarendon Press, 1970.

WORKS BY HART

"Theory and Definition in Jurisprudence." *Proceedings of the Aristotelean Society* 29 (1955): 239–264.

"Are There Any Natural Rights?" *Philosophical Review* 64 (1955): 175–191.

"Positivism and the Separation of Law and Morals." *Harvard Law Review* 71 (1957–1958): 593–629.

"Legal and Moral Obligation." In *Essays in Moral Philosophy*, edited by A. I. Melden, 82–107. Seattle: University of Washington Press, 1958.

Causation in the Law. With A. M. Honore. Oxford, U.K.: Clarendon Press, 1959. 2nd ed., 1986.

The Concept of Law. Oxford, U.K.: Clarendon Press, 1961. 2nd ed., 1994. All references are to the second edition.

Law, Liberty and Morality. London: Oxford University Press, 1963.

Punishment and Responsibility. Oxford, U.K.: Clarendon Press, 1968.

Essays on Bentham: Studies in Jurisprudence and Political Theory. Oxford, U.K.: Clarendon Press, 1982.

Essays in Jurisprudence and Philosophy. Oxford, U.K.: Clarendon Press, 1983.

WORKS ABOUT HART

Fuller, Lon. "Positivism and Fidelity to Law—A Reply to Professor Hart." *Harvard Law Review* 71 (1958): 630–672.

Hacker, Peter, and Joseph Raz, eds. *Morality and Society: Essays in Honour of H. L. A. Hart*. Oxford, U.K.: Oxford University Press, 1977.

Lacey, Nicola. *A Life of H. L. A. Hart: The Nightmare and the Noble Dream*. Oxford, U.K.: Oxford University Press, 2004.

MacCormick, Neil. *H. L. A. Hart*. London: Edward Arnold, 1981.

Martin, Michael. *The Legal Philosophy of H. L. A. Hart: A Critical Appraisal*. Philadelphia: Temple University Press, 1987.

W. J. Waluchow (2005)

HARTLEY, DAVID
(1705–1757)

David Hartley, the association psychologist and moral philosopher, was born in Luddenden, Halifax, England, and was educated at the Bradford grammar school and Jesus College, Cambridge. He was elected a fellow of Jesus but lost his fellowship when he married. He did not take holy orders, probably because of doctrinal scruples. Although he never received a medical degree either, he became a physician and practiced medicine in, successively, Newark, Bury St. Edmunds, London, and Bath. He was a friend of bishops Butler, Law, and Warburton.

Hartley's contribution to philosophy is his treatise *Observations on Man, His Frame, His Duty and His Expectations* (London, 1749). The first part, called "Observations on the Frame of the Human Body and Mind," is Hartley's exposition of the doctrines of vibration, association, and the seven classes of intellectual pleasures and pains. The second part, called "Observations on the Duty and Expectations of Mankind," consists of arguments for the existence of God, a defense of the truth of Christianity, a set of rules of conduct, and an estimate of our legitimate expectations in this life and hereafter.

Hartley's merit lies not in innovation but in consolidation. Borrowing several doctrines from his predecessors, he offers a comprehensive account of human nature. He treats mind and body as parts of a coordinate system capable of influencing each other. Thus, his work is a mixture of speculative physiology and psychology. While his conclusions may be criticized for their lack of an experimental basis (although he appeals frequently to

experience), he nonetheless deserves credit for supporting the conceptual ideal of a unitary system of mind and body. Hartley's theory of knowledge is John Locke's, offered in a context of religious sentiment. Despite the role that association plays in Hartley's philosophy, there is no mention of David Hume in his pages. By freeing the doctrines of learning by experience and of psychological association from skeptical associations, Hartley gave them a respectability that assured them a general currency.

The three aspects of human nature that Hartley wished to explain are sensation, motion, and the generation of ideas. With regard to sensation, he wanted to account for the way in which impressions on the senses register perceptions in the mind. He postulated, first, that the "white medullary substance" of the brain, the spinal marrow, and nerves is the immediate instrument of sensation. He then claimed that when an external object is impressed on the senses, it occasions, first in the nerves linking the senses and the brain and then in the brain, vibrations of the infinitesimal medullary particles. These vibrations are the means of conveying the sensation to the brain.

From this account of sensation, Hartley moved on to his account of the origin of ideas. Sensations may remain in the mind for a short time after the sensible object has been removed from the vicinity of the senses. By being often repeated, sensations leave in the mind certain vestiges, types, or images of themselves. These images are the simple ideas of sensation, the materials from which complex ideas are made. Once the mind is supplied with simple ideas, the association of sensations and ideas may come into play. The first requirement is that we must have a given set of sensations "a sufficient number of times." These sensations then acquire such a power over their corresponding ideas that when any member of the set is impressed on the senses, it is able to excite in the mind the rest of the corresponding ideas that belong to the set. In this way simple ideas collect and become a complex idea.

In addition to arguing for an association of sensations and corresponding sets of ideas, Hartley also argued for a kind of association that depends on the sensory vibrations. Regularly occurring sensory vibrations leave behind in the nerves miniatures of themselves which he calls "vibratiuncles"; and even as a general sensation is able to call up a corresponding set of ideas in the mind, so a sensory vibration is able to call up a corresponding set of vibratiuncles in the nerves. Similarly, a complex idea may call into being the set of vibratiuncles appropriate to the complex of sensations with which the idea corresponds. Hartley claimed that some of the vibratiuncles attending upon complex ideas may be as vivid as any of the sensory vibrations excited by the direct action of objects.

In his account of human motion, Hartley again made use of the "white medullary substance" of the brain, spinal marrow, and nerves. He postulated this substance as the immediate instrument of motion. The motor nerves link brain and muscles; motion results as vibrations pass from the brain along the motor nerves and issue in muscular action. Briefly stated, then, Hartley's general theory of motion is that when objects are impressed in the senses, the vibrations excited in the sensory nerves spill over to the motor nerves by way of the brain and the higher ganglia; this process has a consequent effect on the muscles, and a motion results. According to Hartley, there are two sorts of motion, automatic and voluntary. They are distinguished by the fact that automatic motion depends on sensation, and voluntary motion depends on ideas.

Hartley makes "automatic motion" cover a varied class, which includes such motions as the heart's beating, crying, and voluntary actions that have become habitual through repetition. Heartbeats can be fitted into his theory only by the vaguest references to the spillover effect of the vibrations occasioned by sensation and by the additional suggestion that the circulation of the blood may also cause the heart to beat. The "motion" that best conforms to the theory is a fit of crying that results from a frightening experience or from the pain of being injured. In his theory of motion, Hartley did not intend "sensation" always to be the equivalent of "perception." For instance, the motion of breathing is excited in a newborn infant by cold air and the rough handling of the midwife, sensations which, for the infant, are not perceptions of anything.

In contrast to the automatic motions that depend on sensation, there are the voluntary motions which depend on ideas. A person's will consists of one of his ideas associated with sensory and motor vibratiuncles that are strong enough to excite the motor vibrations which, in turn, issue in muscular action. A voluntary action, for Hartley, is one that follows after an idea in the mind and not as a consequence of some outside force. He made it clear that "voluntary" must not mean "uncaused," and he argued that we have no "power of doing different things, the previous circumstances remaining the same." Indeed, he stigmatizes such an account of freedom as "philosophical freedom." He freely acknowledged that he was a mechanist, but he held that practical freedom does exist, in the sense that the causes of actions may sometimes

originate within a person. Nevertheless, he staunchly maintained that human action cannot be exempted from the reasonable and useful belief that everything has a cause. Indeed, he subscribed to this belief even though he knew that many of his readers would think him a greater friend to religion if he had not stated it explicitly.

Hartley distinguished seven different classes of pleasures and pains that may accompany our sensations (and consequently our ideas), and thus reinforce their affective power, namely the following: (1) those of Sensation, as they arise from impressions made on our external senses; (2) those of Imagination, as they arise from natural beauty and deformity; (3) those of Ambition, as they arise from the opinions of others concerning ourselves; (4) those of Self-Interest, as they arise from our possession (or want) of the means of happiness; (5) those of Sympathy, as they arise from the pleasures and pains of our fellow creatures; (6) those of Theopathy, as they arise from affections excited by our contemplation of Deity; and (7) those of the Moral Sense, as they arise from our awareness of moral beauty and deformity.

The classes of pleasures and pains are here arranged in an ascending order of value, from least to most valuable; and from this scale, Hartley derived the rule of life. The pleasures of sensation, imagination, ambition, and self-interest are not in themselves worthy of pursuit. But the pleasures of sympathy are worthy of pursuit in themselves and set a proper limit to our interest in the first four classes of pleasure. Moreover, the pleasures of sympathy are consistent with those of theopathy and the moral sense. Together, these last three classes of pleasure constitute, as a whole, the worthiest object of human pursuit that can be found.

See also Butler, Joseph; Locke, John; Pain; Pleasure; Priestley, Joseph; Psychology; Sensa.

Bibliography

The dissemination of Hartley's doctrines was aided by his enthusiastic admirer and younger contemporary, the scientist Joseph Priestley. Priestley published an abridged version of the *Observations* under the title *Hartley's Theory of the Human Mind on the Principle of Association of Ideas* (London: J. Johnson, 1775). While Priestley omitted the theory of vibrations, it is said that he was inclined to believe in it. Hartley is also the author of an *Enquiry into the Origin of the Human Appetites and Affections* (Lincoln, England, 1747). See also G. S. Brett, *A History of Psychology,* 3 vols. (London: G. Allen, 1912–1921), Vol. II, pp. 278–286, and Howard C. Warren, *A History of the Association Psychology* (New York: Scribners, 1921), pp. 50–64.

Elmer Sprague (1967)

HARTMANN, EDUARD VON
(1842–1906)

The German pessimistic philosopher Karl Robert Eduard von Hartmann was born in Berlin, the son of a Prussian artillery officer. Von Hartmann entered a school for artillery officers, but a knee injury in 1861 that aggravated older rheumatic ailments barred him from a military career and left him a lifelong semi-invalid. After two years devoted to musical composition and painting, he turned to an intensive study of philosophy. By 1867 von Hartmann had nearly finished his *Die Philosophie des Unbewussten* (Berlin, 1869; 9th ed. translated by W. C. Coupland as *The Philosophy of the Unconscious,* 3 vols., London, 1884). This work brought him prompt and widespread recognition, and the rest of his professional life was devoted to a long series of books that amplified and in some details modified its views, and applied them to various fields of philosophy and problems of contemporary culture. Before his death he had published the first volume of an eight-volume *System of Philosophy* (*System der Philosophie im Grundriss,* Bad Sachsa, 1907–1909). Unlike Arthur Schopenhauer's, von Hartmann's pessimism did not keep him from two happy and fruitful marriages.

INFLUENCES OF VON HARTMANN

Although he is generally regarded as a follower of Schopenhauer, von Hartmann found Schopenhauer's intense morbidity and his intuitive procedure temperamentally alien, and he corrected a basic incompatibility between Schopenhauer's Kantian phenomenalism and his Platonism by imposing upon the doctrine of the "will to live" a Hegelian but nondialectical doctrine of an intelligible categorial structure. Von Hartmann also acknowledged an indebtedness to the early Friedrich Schelling for his theory of the unconscious, to the later Schelling for the process in which nature and consciousness emerge from unconscious potencies, and to Gottfried Wilhelm Leibniz for the synthesis of individualism and monism. In *Die Philosophie des Unbewussten* von Hartmann built these influences into a system inductively grounded upon the data provided by the natural and historical sciences.

WILL AND IDEAS

Although Schopenhauer's concept of Will is needed to explain the dynamism of the world process, it cannot, according to von Hartmann, explain the world order. G. W. F. Hegel's dialectic is absurd, but his concept of the notion is required to explain world order, even though it

cannot account for the process by which, according to both Schopenhauer and Hegel, self-consciousness comes into being out of the unconscious. Ideas define the "what" of the world; Will determines its "that." The opposition of Will to the ordering of the Ideas brings about the emergence of consciousness and individuals.

Individuation results from the conflict of purposes into which the universal Will is driven through its resistance to its logical counterpart, the Ideas. Consciousness is required to emancipate the Ideas from bondage to the Will and its torments. Since space and time are the "sole principium individuationis known to us" (*Philosophy of the Unconscious,* Vol. II, p. 230), the result is a phenomenal but real evolutionary process of nature involving the greatest possible emergence of purposes. Consciousness, when it attains its maturity, will "suffice to hurl back the total actual volition into nothingness, by which the *process* and the *world ceases* … without any residuum whatever." (Von Hartmann suggested later, in Volume III, that the undifferentiated, substantial Will might continue to proliferate other orders of consciousness after this destruction.)

In this differentiation between the Unconscious Will and the Ideas, three orders of being must therefore be distinguished: (*a*) the metaphysical order of the unconscious; (*b*) the objective phenomenal-real order of nature; and (*c*) the subjective-ideal order of consciousness. The physiological unconscious in the second order ("the resting molecular predispositions of the central organs of the nervous system," *Die moderne Psychologie,* Bad Sachsa, 1901, p. 76), provides an unconscious ground for the total consciousness of an organism; conscious perception, in turn, is the bond by which knowledge of the transcendent but phenomenal-real order becomes possible.

CATEGORIES

In his *Kategorienlehre* (Berlin, 1896) von Hartmann distinguished between categorial concepts and the categorial functions of which they are the conscious representations. These unconscious rational functions assure that the concepts establish a relationship between phenomena and the thing-in-itself. There are innumerable categories, distinguishable as categories of sense and thought. There are two kinds of sense categories: sensations, which include quality, intensive quantity, extensive quantity (for example, temporality), and perceptions (that is, spatiality). The categories of thought include the primary category of relation, the categories of reflective thought (comparing, distinguishing, measuring, modality, and others), and the categories of speculative thought (causality, finality, substantiality).

TELEOLOGY

Von Hartmann rejected both the irrational intuitionism of Schopenhauer and the mechanistic and materialistic assumptions of much of the science of his day. His own view of nature was teleological, an interpretation he undertook to demonstrate mathematically by a calculation of probabilities (he estimated the probability of eyesight being produced by mere mechanical processes as less than $15/10^7$) and also by analogies with alleged facts of experience. Instincts are unconsciously purposive, for example, and unconscious ideation in the nerve endings must be assumed to explain the slightest voluntary bodily movements. It is noteworthy that von Hartmann was one of the first to criticize Darwinism, arguing that evolution requires a vitalism and a "heterogeneous generation" of new variations within the germ cells of existing forms of life.

PESSIMISM

Among the factors that account for the great popularity of *The Philosophy of the Unconscious,* the most important is von Hartmann's restatement and justification of philosophical pessimism in the third volume. He regarded Immanuel Kant, not Schopenhauer, as the father of his pessimism: This is not the worst of all possible worlds; indeed, the infinite purposiveness of the particulars in it makes it the best of all possible worlds. Nevertheless, it can be shown that it would be better if there were no world at all, and, paradoxically, the purposiveness of this world is moving to that end.

The argument for pessimism in *The Philosophy of the Unconscious* consists of a remarkable combination of neurological and psychological considerations with commonsense considerations about the misery in the world. Pessimism results from the successive dispelling of three stages of optimistic illusion. The first stage is that happiness has already actually been attained in the present stage of the world; the second is that happiness can be attained in a transcendent life after death; and the third is that happiness will be attained in a future state of this world.

In a later historical and critical essay on pessimism, *Zur Geschichte und Begrundung der Pessimismus* (Berlin, 1888), von Hartmann modified this sweeping argument for the misery of the world by setting up five criteria of value (*Wertmassstäbe*): pleasure, purposiveness, beauty, morality, and religiosity. His pessimistic theory of the

Weltlustbilanz, he now claimed, was based on only the first criterion, and he described his theory as a "eudaemonological pessimism" but a "teleological-evolutionary optimism" in the nonhedonistic fields of value. He still held that efforts to assess values always involve the subjective, hedonistic component, and therefore involve a balance of misery.

Von Hartmann was concerned with showing that, far from making ethics and religion impossible, pessimism is the only foundation for a tenable ethical system and that it provides as well the wider teleological perspective from which religion, including contemporary Christianity, can be evaluated. In his *Phänomenologie des sittlichen Bewusstseins* (Berlin, 1879) he tried to show that all previous efforts to provide a philosophical basis for ethics, whether hedonistic, or built upon formal principles (which, he held, inevitably collapse into an ethics of ends), or socially oriented like utilitarianism and social democracy, had failed because they are untrue of man and the universe. The proper goal, which unites all lesser ethical ends, can only be a cooperative participation in the cultural process contributing to the satisfaction of all particular wills and, therefore, contributing ipso facto to the termination of the universe.

This conclusion anticipates von Hartmann's religious thought. The ethics of pessimism becomes a cosmic drama of redemption. The goal of the absolute religion of the future must be to save God, as Will, from the agony involved in his own inevitable creativity. The essence of vital Christianity, according to von Hartmann, lies in its pessimism about the present world, and liberal Protestantism is the last dying phase of Christian ethics because, by adhering to a faith in social progress, it has lapsed into the first stage of optimism.

Although the unorthodox nature and clear forcefulness of von Hartmann's thought drew a popular following, much critical comment was directed at his paradoxical theory of the unconscious, his criticism of religion, and the incompatibility between his pessimism and his idealistic ethics and philosophy of religion.

Except for a brief attempt to revive interest in von Hartmann's work during the years after his death, it has been largely neglected. He has been hailed as the last of the great speculative idealists, as a philosopher of science who opposed the mechanistic materialism of his time and anticipated the vitalism of the twentieth century, and as a psychologist who introduced the unconscious as a decisive mental factor. His criticism of the human predicament, along with Schopenhauer's, prepared the way for more complete, intensified forms of pessimism and nihilism in the twentieth century.

See also Beauty; Categories; Darwinism; Determinism, A Historical Survey; Hegel, Georg Wilhelm Friedrich; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Nihilism; Pessimism and Optimism; Pleasure; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Teleology; Unconscious.

Bibliography

Additional works by von Hartmann include *Ausgewählte Werke,* 13 vols. (Leipzig, 1885–1901); *Das Unbewusste vom Standpunkte der Physiologie und Descendenztheorie* (Berlin, 1872); *Die Selbsterzetzung des Christentums und die Religion der Zukunft* (Berlin, 1874); *Wahrheit und Irrthum im Darwinismus: eine kritische Darstellung der organischen Entwickelungstheorie* (Berlin: Duncker, 1875), translated by H. J. Darcy as "The True and the False in Darwinism," in *Journal of Speculative Philosophy* 11 (1877): 244–249 and 392–399; *Kritische Grundlegung des transcendentalen Realismus* (Berlin, 1875); *Die Religion des Geistes* (Berlin, 1882), translated by Ernest Dare as *The Religion of the Spirit* (London, 1886), Part III separately translated by Thomas Hitchcock as *Religious Metaphysics* (New York, 1883); *Philosophic des Schönen* (Leipzig, 1882); *Das Grundproblem der Erkenntnistheorie* (Leipzig, 1889); *Geschichte der Metaphysik,* 2 vols. (Leipzig, 1889–1890); and *Das Problem des Lebens* (Bad Sachsa, 1906).

A chronological bibliography of von Hartmann's writings, prepared by Alma von Hartmann, appeared in *Kantstudien* 17 (1912): 501–520. See also Hans Stäglich, *Verzeichnis der Eduard von Hartmann-Literatur* (Leipzig, 1932).

For interpretation and criticism, see Arthur Drews, *Eduard von Hartmanns philosophisches System in Grundriss* (Heidelberg, 1902); Otto Braun, *Eduard von Hartmann* (Stuttgart: F. Frommanns, 1909); Leopold Ziegler, *Das Weltbild Hartmanns* (Leipzig, 1910); E. H. di Carlo, *La filosofia della storia de Eduard von Hartmann* (Palermo, 1906); R O. Petraschek, *Die Logik des Unbewussten,* 2 vols. (Munich, 1926); and W. Rauschenburger, *Eduard von Hartmann* (Heidelberg, 1942).

In English, J. W. Caldwell, "The Epistemology of Eduard von Hartmann," in *Mind,* n.s., 2 (1893): 185–207, and "Hartmann's Moral and Social Philosophy," in *Philosophical Review* 8 (1899): 465–483 and 589–603, are noteworthy.

Two popular but accurate accounts of von Hartmann's thought and influence are contained in G. Stanley Hall, *Founders of Modern Psychology* (New York: Appleton, 1912), pp. 181–246, and Radoslav Tsanoff, *The Nature of Evil* (New York: Macmillan, 1931), Ch. 12.

L. E. Loemker (1967)

HARTMANN, NICOLAI

(1882–1950)

Nicolai Hartmann, the German realist philosopher, was born in Riga, Latvia, and educated at St. Petersburg, Dorpat, and Marburg. He was a professor at Marburg from 1920 to 1925, at Cologne from 1925 to 1931, at Berlin from 1931 to 1945, and at Göttingen from 1945 until his death.

THE WORK AND THE MAN

The typical German philosopher since the mid-1850s gives generous assistance to anyone wishing to become acquainted with his main ideas. He will have published at least one work on a philosopher of the past, who, with the regrettable exception of a few Greeks, turns out to be either German himself or mediocre and, with no exception at all, proves to be someone who could have been the professor's disciple or apostate. By simply observing what the author lauds and damns, stresses and omits, one may gather in concentrated form the materials for a portrait, not of the sitter, to be sure, but of the artist himself. This is true of even so eminently fair a German philosopher as Nicolai Hartmann. In his essays on the history of philosophy, "Zur Methode der Philosophie-geschichte" (1910) and "Der philosophische Gedanke und seine Geschichte" (1936), Hartmann advocated an approach to the history of philosophy in line with that to the history of science (these essays, as well as all others referred to, are reprinted in *Kleinere Schriften*). The history of philosophy is to be presented not as the coming to be and passing away of personal systems but as the progressive accumulation of impersonal insights. Yet many of Hartmann's numerous studies in the history of philosophy show that what he valued as impersonal, objective clarifications and solutions of the past more often than not anticipated views of his own.

In writing his first historical work, *Platos Logik des Seins* (Plato's logic of being; Giessen, 1909), which was his earliest publication as well, Hartmann was so immersed in the neo-Kantianism of Hermann Cohen and Paul Natorp that he viewed Plato's ideas as absolute hypotheses in the neo-Kantian sense of foundational positions taken by thought in its work of constituting reality. His two-volume work on German idealism, on the other hand, particularly the volume on G. W. F. Hegel (1929), bears witness to his accomplished liberation from idealism and the emergence of his main anti–neo-Kantian positions. What he valued in Hegel's philosophy was not its systematic character but its aporetics; not its speculative idealist position but its being, as aporetics, prior to any position; not the Absolute and its self-realization—not even its dialectic, though Hartmann was fascinated and irritated by it, like a skilled craftsman in the presence of genius—but a sort of theory of emergence, describing and exploring, rather than constructing and deducing, basic strata and modes of Being and their interrelations; not the theological and teleological monism of the Spirit, with a capital *S*, but the discovery—or, rather, rediscovery, since Giambattista Vico had been forgotten—of the objective spirit, with a small *s*, that is, superindividual powers such as languages, moral customs, legal systems, into which individual consciousnesses are born and within which they carve their little niches.

Thus, in the end the Hartmannian Hegel is, in method, an aporetician who prefers careful analyses of problems to traditional solutions and, in subject matter, an ontologist engaged in describing a multitude of modes and strata of Being. In brief, what is alive in Hegel is Hartmann, for aporetics, particularly an aporetic epistemology "this side of realism and idealism," an ontological pluralism, and the categorial exploration of the real world, including the spirit as its highest stratum, describe all but one of the major commitments of Hartmann. The exception is his axiology, an exploration of the realm of values, a program that Max Scheler had designed in open battle against the Kantian formalism in ethics but whose execution had to wait for Hartmann's *Ethics*, the only major work of his translated into English.

In his *Grundzüge einer Metaphysik der Erkenntnis* (Outlines of a metaphysic of knowledge; Berlin, 1921), Hartmann presented in book form, for the first time, his aporetic and ontological epistemology. The book, published six years before Martin Heidegger's *Sein und Zeit*, caused quite a stir precisely because it heralded the Continental renaissance of ontology by asserting that epistemology is based on ontology and not the other way around. Almost all of Hartmann's subsequent books are in ontology, with his trilogy *Zur Grundlegung der Ontologie* (Foundations of ontology; Berlin, 1935), *Möglichkeit und Wirklichkeit* (Possibility and reality; Berlin, 1938), and *Der Aufbau der realen Welt* (The structure of the real world; Berlin, 1940) forming his ontological *opus maximum*. To these might be added his *Philosophie der Natur, Abriss der speziellen Kategorienlehre* (Philosophy of nature, outlines of the special doctrine of categories), which he began in 1927 but did not publish until 1950 (Berlin). In fact, his own philosophical work follows the plan he wished the history of philosophy to follow; there is much steady progress and expansion. The germs of his

central ideas—many of them images rather than concepts—can be seen even in his early neo-Kantian writings. For example, so typical an image as that of "the strata" and their hierarchy, basic to his later ontology, is already germinally active in his early *Zur Methode der Philosophiegeschichte* of 1909.

There was only one revolution in Hartmann's thinking. This was the revolution against the neo-Kantian idealism of his philosophical youth. It must have been a matter of profound travail for him, even though he was undoubtedly helped by certain select aspects of Edmund Husserl's phenomenology, in particular its Platonizing intuition of essences and its program for a merely descriptive reappropriation of experience. Whatever one may think of Hartmann's ontology philosophically, one cannot help being awed by his self-liberation from the grand German tradition of transcendental idealism, a liberation much more strenuous to a German than to Anglo-Saxons such as Bertrand Russell and George Edward Moore or Americans such as John Dewey, all of whom, to be sure, had somewhat similar conversions. These other conversions, however, were returns to the main current of their national philosophical traditions, whereas Hartmann's went counter to the main current of his. Indeed, even the severest critic of Hartmann's philosophy will respect the man himself. His philosophy reveals him to have been a careful, disciplined, honest, and sober conservative, kept by his common sense from philosophical extravaganza, but kept also from asking or appreciating radically revolutionary questions of either the existentialist or the new empiricist kind. Though as a person he was unmistakably German, his way of doing and writing philosophy was not at all typical of recent German philosophers. He cherished discussions and admitted to having learned from his students. He wrote not in the attitude of "the reader be damned" but with true courtesy toward his public, not to awe with profundity of learning but to guide with lucidity and thoroughness.

PHILOSOPHICAL POSITIONS

In presenting Hartmann's major positions it is advisable to begin with his conception of aporetics as philosophical method, not only because Hartmann himself put it at the beginning of his "Systematische Selbstdarstellung" (in *Deutsche systematische Philosophie nach ihren Gestaltern*, edited by Hermann Schwarz, Vol. I, Berlin, 1933; reprinted in *Kleinere Schriften*, Vol. I), but also because it is that part of Hartmann's philosophy that philosophers of the English tradition should find the most congenial.

APORETICS AND EPISTEMOLOGY. Aporetics is the unraveling of problems (aporia) into their strands; their presentation as clear-cut issues, preferably in the form of antinomies; and the weighing of the pros and cons of apparent solutions. There are some philosophical problems—the metaphysical problems—that will turn out to be in principle insoluble. Yet their unraveling is still useful, for as some part of the issues may turn out to be soluble, their discussion will contribute to the location and diagnosis of the unmanageable remainder. Aporetics is the central business of philosophy, all too often abandoned in favor of system building. Hartmann did not tire of pointing out that aporetics is what the Platonic dialogues and the best pages of Aristotle exemplify. However, this will hardly suffice in the age of science, when the nature of philosophy and philosophical problems is itself an aporia. Rather, one would wish to know what, if anything, distinguishes philosophical from logical or scientific problems. One would wish to know, besides, what it is that makes some philosophical issues insoluble and what the criteria are in terms of which some answers are solutions and some are not. Hartmann saw philosophical problems as arising from what he took to be the facts and from the contradictions they appear to harbor. His philosophical method, then, consisted really of two parts, a phenomenological presentation of the facts and an aporetic discussion of their implicit contradictions.

A typical example of Hartmann's descriptive-phenomenological and aporetic method may be found in his *Grundzüge einer Metaphysik der Erkenntnis*. No merely descriptive account of experience can plausibly deny that the objects known by a consciousness are experienced as existing independently of their being known. This fact, however, harbors in itself riddles in the form of flagrant contradictions: Consciousness, in knowing an object, transcends itself, yet anything known to consciousness is thereby a content of consciousness—that is, is immanent—and consciousness never transcends itself. The same riddle, but formulated from the side of the object, concerns the influence of the object on the subject. On the one hand, the object must break into a consciousness and produce an image of itself; on the other, the object must remain outside the subject, for it is, as object, something transcendent and indifferent to its being known by a subject.

Hartmann neither questioned the nature of the facts he supposed himself to be describing nor entertained any suspicion that the antinomies he found in those facts might be due to the sort of language he used in describing them. Instead, he proceeded from knowing to being.

The epistemic aporias are essentially ontic aporias, for both the object and the subject are beings (*Ansichseiendes*). The object is not exhausted in its being an object of a subject. Like a nocturnal thief caught in a sudden glare of light, it emerges out of an unredeemably transobjective and metarational background, a background that is in part beyond any human cognition, even beyond any possible sort of cognition. In knowing an object the subject knows "a thing that is," a being. In turn, the consciousness that knows the object is itself something not exhausted by its being a subject. It emerges out of a transsubjective and metarational medium; it is itself "a thing that is, a mode of Being." Hence, the epistemic relation between knower and known is really an ontic relation holding between one being and another, and the problems in epistemology are, or issue in, problems in ontology. As beings, both subject and object are ontologically homogeneous and are members of a context of Being (*Seinszusammenhang*). Within this context their relationship, so puzzling when taken in epistemological abstraction, becomes conceptually manageable, though an insoluble, and hence metaphysical, problem remains. This problem, however, concerns not the fact *that*, but rather how, subject and object stand in relation to each other.

In short, by seeing both subject and object as *Ansichseiendes*, Hartmann believed himself to have discovered that they are ultimately members of one matrix and context of Being. This is supposed to explain that they are related, though the how of their relation remains mysterious. Thus, the Hartmannian turn from epistemology to ontology looks suspiciously like a piece of verbal magic, as if a biologist, puzzled by the relation between males and females, proposed to solve the puzzle by calling both males and females "sexuals" and hence members of the sex context, thus "explaining" that they have sexual relations, though still wondering how they have them. Hartmann's reduction of epistemology to ontology is a piece of philosophical verbal magic if "subject" and "object" have empirical meaning, as "male" and "female" do. But if "subject" and "object" have no empirical meaning, what sort of meaning do they have? This basic question is unasked, and one cannot help wondering if the main use of the terms is not to engender the antinomies without which epistemologists would be out of work. In sum, Hartmann's phenomenological emphasis on descriptive facts seems to bring philosophical problems closer to empirical ones, whereas his aporetic emphasis on antinomies seems to bring them closer to logical ones. It is this basic ambivalence in his conception of philosophical method that cannot but be reflected in his conception of ontology.

ONTOLOGY. If there was anything twentieth- and twenty-first-century ontologists have had in common it is their unquestioning belief that the term *Being* is the name of something or other. What is debated is rather what *Being* is a name of: a quality or feature shared by all beings (and if so, whether this class, as *summum genus*, is distinguished from other classes merely by its higher degree of universality); some relation that any *x*, in order "to be a being," must have to be a subject, or man, or God; or an individual being who is the ground of all beings. Since they have not questioned that "Being" is a name, these ontologists, like their predecessors, have a problem concerning the unity of Being. Protons and principles, nations and numbers, salads and sentences are all said to have some sort of Being, and yet, because they *are* so differently, the ontologist is compelled to admit different kinds of Being. But this would make Being itself the genus of these kinds, just another class concept, albeit more abstract or universal, depriving ontology, in the process, of its metaphysical weight and attraction. In this predicament ontologists have chosen a linguistic escape. Instead of talking of kinds of Being, they prefer to talk of modes of Being. They thus believe themselves to be preserving the unity of Being in the variety of beings without prostituting Being to a mere class name—and a name, of course, it must be.

To a degree, Hartmann shared with the most outspoken ontologists of the mid-twentieth century, the existentialists, both the referential use of "Being" and the preservation of the unity of Being via modes. However, he was at once simpler and more confused than they. Heidegger, for example, made the most of the distinction between Being and beings, between *das Sein* and *das Seiende*, and, correspondingly, between ontological and ontic investigations. But Hartmann, at least in his pre-Heideggerian writings (such as *Grundzüge einer Metaphysik der Erkenntnis*), seems to have been rather uncertain about this difference and to have used the term *ontological* for any investigation concerned with beings. This makes the concept of ontology simpler, as it keeps the white whale of Being from perturbing the Ahab of beings, but it also makes the concept more confused, as one is now at a loss to distinguish between ontology and science, both of which have to do with beings. In his pre-Heideggerian *Grundzüge*, Hartmann was similarly apt to be very cavalier about the problem of the unity of Being, and he spoke of modes and strata of Being as if they were merely basic kinds of beings. Even though in his later works all this appears to have changed, presumably under the influence of Heidegger's *Sein und Zeit*, and Hartmann no longer slid terminologically from "beings" to "Being"

as he had in the *Grundzüge* (p. 182), conceptually the distinction between Being and beings and the problem of the unity of Being remain rather vague and were for him hardly the matter of primary philosophical concern they were for Heidegger.

Hartmann distinguished between two basic modes of Being (*Seinsweisen*), very much as the American new realists distinguished between existence and subsistence twenty years before him. One mode of Being consists of particulars, localizable in time and space, the other of universals—for example, essences, values, numbers. The former are real, the latter ideal; both are equally objective and independent of the subject. The ideals are logically prior to the reals, for a real is what it is only by virtue of an essence present in it (or valuable only by virtue of a value present in it). This apriority of ideal entities, however, does not exclude their being possible objects of experience, ideals being given in intuition just as reals are in perception. (Here "perception" and, it would seem, "intuition" must be used generously enough to include the emotional, for, following Max Scheler, Hartmann asserted valuables, if not values themselves, to be experienced emotionally rather than cognitionally.) Nor does the apriority of ideal entities exclude the possibility that the intuitional acts in which they are experienced are, *in ordo essendi*, grounded on the perceptual acts in which reals are experienced. As in Husserl, then, the a priori is not opposed to, but is rather part of, the empirical.

Within each of the two basic modes of Being, Hartmann distinguished between several strata of Being (*Seinsschichten*). The strata of reality correspond to the distinctions between inorganic nature, organic nature, consciousness, and superindividual culture (*Geist*)—all of them reals, but the last two also agents and carriers of ideals. Each stratum has basic, so-called categorial features, which it is the task of regional ontologies to lay bare. The strata form a hierarchy in which one stratum's dependence on the existence of another and partial freedom (autonomy) from the other's laws mark the higher from the lower. The working out of these regional ontologies through categorial analyses was one of Hartmann's central preoccupations, especially in *Der Aufbau der realen Welt*.

The distinctions between the two modes of Being and between the several strata within each mode were related by Hartmann to the traditional three modalities of possibility, reality, and necessity. Originally these are ontological modalities: It is beings that are possible, real, or necessary (*Seinkönnen, Sein, Seinmüssen*). Only derivatively are they distinctions concerning validity or certainty of knowledge. The many-dimensional relations of the ontological modalities to the modes and strata of Being, on the one hand, and to judgment and knowledge, on the other, are explored in *Möglichkeit und Wirklichkeit*, the most complex and difficult of Hartmann's works.

The revolution of the *Grundzüge einer Metaphysik der Erkenntnis*, making epistemology ontological, revenges itself upon Hartmann's ontology, making it epistemological. A being is primarily understood as that which is *an sich* (in itself), and this *an sich*, the traditional substance of ontology, is defined epistemologically as that which is indifferent to its being known by a subject. Moreover, this *an sich* is either that of reals or that of ideals. Reality and ideality, the two basic modes of Being, become two basic classes of beings. The genus common to both is the *an sich*. Thus, Hartmann appears to have slipped back into class concepts and some sort of taxonomy, half epistemological, half empirical. As with the modes, so with the strata: It is not at all clear what distinguishes strata of Being from kinds of beings. The several strata of reality seem to be related to reality as so many classes are to a genus. Finally, it is not at all clear what distinguishes the concept of a stratum from that of a mode. The two modes of Being, reality and ideality, are themselves related to each other in a multidimensional hierarchical order, very much like the strata, and it would therefore seem that the main difference between mode of Being and stratum of Being is, as with class concepts, the degree of abstractness or universality. In fact, in his *Ethics* Hartmann dealt with ideality as if it were just another stratum and not a mode. If the modes and strata of Being were just kinds of beings, it would follow that ontology is empirical and that the categories are concepts, like any other class concept but more universal. Here lies, as was hinted before, a crucial difficulty of Hartmann's ontology, a difficulty that is shared, mutatis mutandis, by Samuel Alexander's and Alfred North Whitehead's conceptions of categories and that reflects some really basic indecision on their respective conceptions of philosophy. Categories, as Hartmann conceived them, are descriptive of the behavior of different kinds of beings, yet are supposed to be different in kind, not just in the degree of universality, from both the class concepts of ordinary experience—such as "tree" or "rodent"—and the functional concepts of science. At the same time, categories are supposed to be related—and it seems, in some sense, necessarily related—to both. But neither difference nor relation is clearly worked out, and Hartmann's ontology, like Alexander's and Whitehead's cosmologies, continues to hover between the empirical and the a priori as well as between science and ordinary experience.

ETHICS. Indecision of this and other sorts haunts Hartmann's *Ethics*. Moral philosophy must not be casuistry; it must not try to teach what one ought to do in a particular situation. Rather, it should give the general criteria for a universal ethic. This sounds Kantian enough. But whereas Immanuel Kant used it as a steppingstone to philosophically central investigations concerning the logical nature and the transcendental foundation, if any, of these principles, Hartmann veered off in a very different direction. Somewhat like a course in art appreciation, ethics is supposed to make men sensitive to the wealth of values present in the world. This, however, makes Hartmann uncomfortable; it is not academically respectable. The task of moral philosophy is, rather, to present clearly, force into consciousness, and "establish" values, raising to the plane of science what was a mere affair of feeling. How are the two conceptions of moral philosophy—that of making explicit universal principles of what ought to be done and that of raising value feelings to the plane of science—to be united? What one ought to do can be gauged only if one has an insight into what is valuable in life.

In fact, however, this synthesis of the Kantian apriority of moral principles with the manifoldness of values, "which [Friedrich] Nietzsche had discerned only to let it melt away in historical relativism," is only a secondary aim of Hartmann's *Ethics*. Its central task is an analysis of the content of values, an elaborate axiology exploring the multitude of values and their relations to each other, to the ought, and to the real. This is the main body and the core of Hartmann's work in moral philosophy. It fills the second volume of the *Ethics*; the third is devoted to the problem of the freedom of the will. The first volume, besides developing his conception of moral philosophy, is a phenomenology of morality. Typical moral philosophies of the past are discussed with the aim of discovering in each of them a sound insight into some partial aspect of the moral phenomenon. Kant, for example, is said to have seen very clearly that ethical principles do not have the empirical sort of universality. They are a priori. Yet Kant's uncritical use of the Aristotelian form-matter dualism made him equate the a priori with the formal, and his epistemology made him equate the formal with the (transcendentally) subjective. Against this Kantian formalism and subjectivism Hartmann's axiology asserts an a priori of objective content—that is, of values as ideal entities that are intuitable.

It may, of course, be argued against this value objectivism that a judgment like "*x* is valuable" or "*x* is more valuable than *y*" will at some time and by some people be considered true and at some other time and by some other people false without there appearing to be any universal criteria of distinguishing, or any method of testing, the truth or falsity of these rival claims. But Hartmann answers this and other relativistic arguments by comparing the intuitional sense of values with a source of light. Light will penetrate darkness and illumine objects according to the strength of its source and will reveal the below and above of objects according to the position of its source, none of this preventing the objects and their spatial relations from being objective and knowable. This comparison assumes what was to be proved. It assumes that any particular sense of values is able to determine its own weaknesses with respect to all other possible senses of values, quite unlike a particular source of light, whose characteristics could not be objectively determined were it not for the sun and the light of day and the knowledge thus made possible. Hartmann's methodological reliance on an intuitional sense of values leaves his readers without a theoretical basis upon which to check and argue the truth or falsity of his observations in the realm of values. One cannot help being pleased with them as one might be with the descriptions of a foreign though somewhat familiar country, but since the recommended means of transportation is in no public domain, the country might as well be Cockaigne. There is no use arguing with the Baedeker of Cockaigne: Anyone may write his own. However, Hartmann's actual work in describing values is far superior to the account he gives of its method, and it is likely that this second volume of his *Ethics* will be found not only enjoyable but also useful.

Hartmann defined freedom of the will as independence from any determination that is constraint without this independence becoming indetermination. Indetermination is not only ontologically impossible; above all, it has once and for all been overcome in the Kantian conception of freedom as autonomy. Indeed, to Hartmann any determination that is not autonomy is constraint, and therefore autonomy fulfills the requirements of a definition of freedom of will. It is logically impossible for a will, insofar as it determines itself, to be either constrained or undetermined. If this is what freedom of will is, how can there be such a thing?

Hartmann's answer has two parts. The first is general, in terms of the hierarchy of ontological strata. Hartmann had asserted that the mark of a stratum's being higher than another is, in part, its autonomy—that is, the emergence of a new sort of determination or law. Autonomy is, then, a general ontological feature to be found, by definition, in any stratum but the lowest, and the auton-

omy called freedom of will is only a particular type of this general feature. Very much like theories of emergence, such as these of Alexander and Whitehead, which assume the miracle of emergence to be less miraculous if it is repeated, Hartmann's theory explains freedom of will as moral autonomy by making autonomy a universal ontological feature and then describing moral autonomy as a species of it.

The second part of his answer is concerned with this specific nature of moral autonomy. The will, in order to be free, must not be determined by the causal apparatus of nature, nor must it be determined finalistically by values and the corresponding oughts they confront us with. In either case the will would be determined by something outside it; that is, it would be constrained. However, if there were only these two kinds of determination, the causal determination of nature and the finalistic determination of values, then a will independent of nature would be one determined by the ought, and a will independent of the ought would be one determined by nature (indetermination being ontologically impossible); in either case the will would not be free. This is the most basic of Hartmann's aporias connected with the freedom of will. Hartmann proposed to solve it by positing a third kind of determination, whose nature he admitted to be completely inscrutable and metaphysical, a determination that belongs to the person itself, a self-determination through which an agent commits himself to the realization of value. Only such a third kind of determination, above both nature and value, explains the possibility of freedom of will.

Quite apart from the mistaken identification of determination and constraint, what seems particularly objectionable is Hartmann's suggestion of a kind of determination that is in principle inscrutable as a solution to the basic aporia of free will. Postulating an unknowable *x* as the solution to a problem is like shouting "victory" to undo defeat. In his escape from Kant's conception of autonomy as self-legislation of rational beings Hartmann fell under the spell of a supposed metaphysical ground in which a person and his decision making are taken to be rooted—a romanticism somewhat like Jean-Paul Sartre's "dreadful freedom," which on closer inspection turns out to be mere whim.

The absence of religious thought in Hartmann's philosophy is conspicuous. Value realism offers logical difficulties to theology, and it is, besides, more naturally connected with a life attitude whose religiosity—if this word can here be used—lies in value commitments and not in a personal relation to God. Thus, Hartmann's value realism, as well as his pro-scientific persuasions and his empirically colored ontology, makes his proximity to atheism quite understandable. Very much unlike his beloved German idealists, who expressed the main existential spring of their philosophical energy in the problem of the "relation of the infinite and the finite," Hartmann was energized by no such preoccupations. His were intellectual aporias, not existential quandaries.

Hartmann's influence on German philosophy, though for a while considerable, was unable to stem the tide of existentialism. With the mid-twentieth-century return of German philosophy to a more sober and rational style, a new esteem for Hartmann began to develop in Germany. In the English-speaking world in the same period his *Ethics* was greeted with respect and then allowed to disappear, leaving hardly a trace. His only notable influence in fields other than ethics seems to have been on W. M. Urban, but as Urban's books have been ignored, Hartmann's effect in English-speaking countries has been limited to such indirect sources as Mario Bunge's somewhat Hartmannian books. Perhaps the sobriety, carefulness, and common sense of his general philosophical style are too much in the English and American tradition to attract our interest, and his logical and analytic naïvetés too numerous to hold it. It will take some time before these naïvetés are overlooked for the sake of his insights in the realm of values.

See also Alexander, Samuel; A Priori and A Posteriori; Aristotle; Being; Cohen, Hermann; Dewey, John; Ethics, History of; Existentialism; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; History and Historiography of Philosophy; Husserl, Edmund; Idealism; Kant, Immanuel; Moore, George Edward; Natorp, Paul; New Realism; Nietzsche, Friedrich; Ontology, History of; Plato; Neo-Kantianism; Platonism and the Platonic Tradition; Realism; Russell, Bertrand Arthur William; Scheler, Max; Value and Valuation; Vico, Giambattista; Whitehead, Alfred North.

Bibliography

ADDITIONAL WORKS BY HARTMANN

Philosophische Grundfragen der Biologie. Göttingen: Vanderhoeck and Ruprecht, 1912.

Die Philosophie des deutschen Idealismus, 2 vols. Berlin, 1923–1929. Vol. I: *Fichte, Schelling und die Romantik*; Vol. II: *Hegel*. Related are "Aristoteles und Hegel," in *Beiträge zur Philosophie des deutschen Idealismus* 3 (1923), and "Hegel und das Problem der Realdialektik," in *Blatter fur deutsche*

Philosophie 9 (1935). Both are reprinted in *Kleinere Schriften* (see below).

Ethik. Berlin, 1926. Translated by Stanton Coit as *Ethics*, 3 vols. London, 1932.

Das Problem des geistigen Seins. Untersuchungen zur Grundlegung der Geschichtsphilosophie und Geschichtswissenschaften. Berlin, 1933.

"Neue Wege der Ontologie." In *Systematische Philosophie*. Stuttgart, 1943. Translated by R. C. Kuhn as *New Ways of Ontology*. Chicago, 1953. This summarizes *Der Aufbau der realen Welt* (Berlin, 1940).

Teleologisches Denken. Berlin, 1951.

Ästhetik. Berlin: De Gruyter, 1953.

Kleinere Schriften, 3 vols. Berlin De Gruyter, 1955–1958. Vol. I: *Abhandlungen zur systematischen Philosophie*; Vol. II: *Abhandlungen zur Philosophiegeschichte*; Vol. III: *Vom Neukantianismus zur Ontologie*. This collection reprints many articles originally published in journals.

WORKS ON HARTMANN

Baumgartner, H. M. *Die Unbedingtheit des Sittlichen; Eine Auseinandersetzung mit Nikolai Hartmann*. Berlin, 1962.

Cadwallader, E. *Searchlight on Values: Nicolai Hartmann's Twentieth-Century Value Platonism*. Washington, DC: University Press of America, 1984.

Feyerabend, P. K. "Professor Hartmann's Philosophy of Nature." *Ratio* 5 (1) (June 1963).

Werkmeister, W. *Nicolai Hartmann's New Ontology*. Tallahassee: The Florida State University Press, 1990.

Werkmeister, W. *Nicolai Hartmann's New Ontology*. Tallahassee: The Florida State University Press, 1990.

Heimsoeth, Heinz, and Robert Heiss, eds. *Nicolai Hartmann. Der Denker und sein Werk*. Göttingen, 1952. Fifteen essays, with bibliography.

Kanthack, Katharina. *Nikolai Hartmann und das Ende der Ontologie*. Berlin, 1962. A critique of Hartmann's ontology from the viewpoint of what the author believes to be the new, Heideggerian era of philosophy.

Tymieniecka, Anna-Teresa. *Essence et existence; Étude à propos de la philosophie de Roman Ingarden et Nicolai Hartmann*. Paris: Aubier, 1957.

Harich, W. *Nicolai Hartmann: Leben, Werk, Wirkung*. Würzburg: Königshausen & Neumann, 2000.

Wirth, Ingeborg. *Realismus und Apriorismus in Nicolai Hartmanns Erkenntnistheorie*. Berlin: De Gruyter, 1965.

Walter Cerf (1967)
Bibliography updated by Thomas Nenon (2005)

HARVEY, WILLIAM
(1578–1657)

William Harvey, the English doctor and anatomist, was the demonstrator of the principle of the circulation of the blood. He was born at Folkstone, Kent, and educated at King's School, Canterbury, and Gonville and Caius College, Cambridge. After taking his B.A. in 1597, he left Cambridge for Padua, where he worked with the anatomist Fabrizzi d'Acquapendente (often Latinized as Fabrizio of Aquapendente). Fabrizzi had observed the valves in the veins, although he had not understood their function; Harvey told Robert Boyle that he had developed his theory of the circulation of the blood by reflecting on the operation of these valves, perhaps while still at Padua. In 1602 Harvey graduated from Padua with a medical degree and incorporated as an M.D. of Cambridge. Taking up practice in London, he was married in 1604 to Elizabeth Browne, daughter of the physician to James I. He was elected fellow of the Royal College of Physicians in 1607, and two years later was appointed physician to St. Bartholomew's College, a position he held for thirty-four years. In 1616, when he began to lecture as Lumleian lecturer in surgery of the Royal College of Physicians (a post he assumed in 1615), he was already expounding his theory of the circulation of the blood, although he did not publish his *Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus* (An anatomical exercise concerning the motion of the heart and the blood) until 1628. Appointed Physician Extraordinary to James I in 1618, and in 1631 to Charles I, Harvey was identified with the royalist cause during the Civil War. In 1642 his London house was ransacked by Parliamentary troops who destroyed notes and specimens. Charles I appointed him warden of Merton College, Oxford, in 1645.

After the Civil War Harvey lived a secluded life, retiring from practice and devoting himself to embryological research. Pleading age, he declined the presidency of the Royal College of Physicians in 1654, and in 1656 resigned from his Lumleian lectureship. He died at Roehampton.

There has been considerable dispute as to whether the discovery of the circulation of the blood can properly be ascribed to Harvey. As early as 1543, Andreas Vesalius had expressed doubts about the traditional Galenic account, according to which blood was made in the liver, flowed through the veins, and was then excreted, except for a small part that passed through minute channels in the septum to the right ventricle and so into the arteries. Vesalius complained that he could not find the channels through the septum. Michael Servetus (1511–1553), Andrea Cesalpino (1519–1603), and especially Matteo Realdo Colombo (1516–1559) all gave a reasonably accurate picture of the flow of blood through the lungs—the so-called lesser circulation. But none of them recognized that the entire blood supply circulated through the body.

Harvey used the comparative method. Confused by the rapidity of movements within a living human body, he dissected such cold-blooded animals as toads and shrimps, in which movement is slower. Many of his con-

temporaries criticized him on the ground that what was true of the lower animals had no application to man. But for Harvey, as for Aristotle, man formed part of the animal kingdom.

Harvey's great importance lies in the fact that he used the concepts of mechanics in his analysis of physiological processes. He described the working of the heart in the language of pumps; he applied mathematical calculations to show that the body could not possibly manufacture the quantity of blood which, according to Galen's theory, would have to flow through it. The fact that blood still had a quasi-mystical significance made his matter-of-fact approach particularly significant.

Harvey's work greatly influenced many early modern philosophers, including René Descartes and Thomas Hobbes (otherwise so different), who both put Harvey on a level with Galileo Galilei. They saw that Harvey had broken down the barrier between the animal and the human body, and between the processes of the body and the processes of mechanics. Thus Harvey's discovery gave empirical support to their mechanistic hypotheses. Descartes objected, however, that Harvey had not shown from first principles that the blood must necessarily circulate; he had been content to say that the heart is in fact a pump and that the blood does in fact circulate. Harvey replied to Descartes in his letters to the French anatomist Jean Riolan, who had rejected Harvey's theory. These letters were included in *Exercitationes Duae Anatomicae de Circulatione Sanguinis, ad Johannem Riolanum Filium Parisiensem* (Two anatomical exercises concerning the circulation of the blood, addressed to Jean Riolan Jr., of Paris; 1649).

Scientific truth, Harvey argued, is to be discovered by direct observation. "No more certain demonstration or means of gaining faith can be adduced than examination by the senses, than ocular demonstration." In this respect, Harvey compares biology favorably with astronomy. The astronomer, he suggests, argues from appearances. He cannot see what happens in an eclipse; all he actually sees is one disc sliding across another, whereas the biologist can see the heart beating in a shrimp. Observation shows us that the blood circulates, and that is enough for the biologist. This is a classic statement of the attitude of the observational biologist, in opposition to the Cartesian mathematico-physical conception of science. The *Letters to Riolan* also contain Harvey's criticisms of the attempt to explain physiological functioning in terms of "spirits." "Persons of limited information when they are at a loss to assign a cause for anything, very commonly reply that it is done by the spirits."

Harvey's other major work is contained in his *Exercitationes Duae Anatomicae de Generatione Animalium* (Two anatomical exercises concerning the generation of animals; 1651). Although this work was important in developing the view that each living thing is produced from an egg, as opposed to the doctrine of spontaneous generation, it lacks the scientific assurance of the *De Motu Cordis.* The fact that Harvey had to rely upon the unaided eye very much limited his achievement in this area. The first work of any consequence carried out with a microscope was done in 1660 by Marcello Malpighi, who took Harvey's theory as his point of departure.

See also Aristotle; Biology; Boyle, Robert; Descartes, René; Galileo Galilei; Hobbes, Thomas; Philosophy of Biology.

Bibliography

WORKS BY HARVEY

The collected Latin works of William Harvey were published in London as *Opera Omnia* in 1766; the earliest English texts, dating from 1653, have been reprinted, edited by Geoffrey L. Keynes (London, 1928); a complete English translation was prepared by Robert Willis for the Sydenham Society, London, in 1847. The notes Harvey prepared in 1615 for his Lumleian Lectures were published in facsimile as *Praelectiones Anatomiae Universalis* (London, 1886) and translated by C. D. O'Malley, F. N. L Paynter, and K. F. Russell as *Lectures in the Whole of Anatomy* (Berkeley, CA, 1961). The manuscript *De Motu Locali Animalium* (1627) was first published in Cambridge, U.K., in 1959, edited by Gweneth Whitteridge. See also *De Motu Cordis,* edited and translated by Kenneth James Franklin as *Movement of the Heart and Blood in Animals* (Oxford: Blackwell, 1957); *The Circulation of the Blood,* containing Harvey's reply to Jean Riolan, edited and translated by Kenneth James Franklin (London: Dent, 1963).

WORKS ON HARVEY

For literature on Harvey, see Geoffrey L. Keynes, *A Bibliography of the Work of William Harvey* (London, 1928; 2nd ed., Cambridge, U.K.: Cambridge University Press, 1953), and *The Personality of William Harvey* (Cambridge, U.K.: Cambridge University Press, 1949); Henry P. Bayon, "William Harvey," in *Annals of Science* 3 (1938): 59–118, 435–456; 4 (1939): 65–106, 329–389; Étienne Gilson, *Études sur la rôle de la pensée médiévale dans la formation du système cartésien* (Paris, 1951), Ch. 2; A. Rupert Hall, *The Scientific Revolution, 1500–1800* (London: Longmans, Green, 1954); Louis Chauvois, *William Harvey* (New York, 1957); *Proceedings of the Harvey Tercentenary Congress,* edited by John McMichael (London, 1958); John Arthur Passmore, "William Harvey and the Philosophy of Science," *Australian Journal of Philosophy* 36 (2) (1958): 85–94.

John Passmore (1967)

HATANO SEIICHI

(1877–1950)

Hatano Seiichi, the Japanese historian of philosophy and philosopher of religion, was born in Nagano and died in Tokyo. He studied at Tokyo University, where his thinking was formed by Raphael von Koeber, a pupil of Eduard von Hartmann. He wrote his doctoral thesis, "A Study of Spinoza" (1904), in German. In 1901 he had published *Seiyo tetsugaku shiyo* (Outline of the history of Western philosophy; Tokyo), a book whose scholarship established his reputation. He went to Germany in 1904 and studied under Carl Gustav Adolf von Harnack and Otto Pfleiderer at Berlin for two years, then under Wilhelm Windelband at Heidelberg. He also developed his studies of Protestant theology (he had been baptized in 1902) under J. Weiss, Ernst Troeltsch, and A. Deissmann. Their lectures prepared him to be a temporary replacement for Anesaki Masaharu, Tokyo University's well-known historian of religion. From his lecture notes he published *Kirisutokyō no kigen* (The origin of Christianity; Tokyo, 1908). In a much later book, *Genshi kirisutokyō* (Primitive Christianity; Tokyo, 1950), he rose above the historico-textual criticism of his early days to present a more thorough study of the essence of Christianity. But at the beginning of the twentieth century this type of work was a novelty in Japan.

In 1917 he resigned from Waseda University, where he had taught for many years, and at the invitation of Nishida Kitarō, the leading philosopher of Japan, he joined the staff of Kyoto University; he taught there until he retired in 1947. Subsequently he became the president of Tamagawa University in Tokyo. At Kyoto he had the chair of science of religion, and from 1922 he held the chair of Christianity. Hatano developed his philosophy of religion in four books (all published in Japanese, in Tokyo): "The Essence and Fundamental Problem of Philosophy of Religion" (1920); "Philosophy of Religion" (1935); "Introduction to Philosophy of Religion" (1940); *Time and Eternity* (Japanese edition, 1943; English edition, 1963). His main idea is that the comparative study of religion presupposes a philosophy of religion because values are a necessary element of that science. As philosophy, religion must start from the reality of the religious experience of God; this experience is first an experience of the God of power, then a quest for the God of truth, and finally an experience of the God of love. Hatano distinguishes three kinds of time: the natural, encompassing the realm of nature; the cultural, characterized by eros; and the eternal, in which agape, or Christian love, triumphs. Original, too, are his observations on the "about-to-come" future (*shōrai*) and the distant future (*mirai*). The first is implicitly part of our present; it is the future we are making, the supplier of being. The second is time that will never be experienced by the subject. Clearly Hatano is much influenced by Christian ideas and takes almost nothing from the Oriental climate of thought; even so this type of philosophy of religion was and still is very influential in Japanese thought.

See also Harnack, Carl Gustav Adolf von; Hartmann, Eduard von; Japanese Philosophy; Nishida Kitarō; Philosophy of Religion; Troeltsch, Ernst; Windelband, Wilhelm.

Bibliography

Hatano's works have been published in Japanese as *Hatano Seiichi zenshū* (The complete works of Hatano Seiichi), 5 vols. (Tokyo, 1949). Vols. III–V include Hatano's works on the philosophy of religion. One of his works on that subject, *Toki to eien* (Tokyo, 1943), has been translated by Ichiro Suzuki as *Time and Eternity* (Tokyo: Japanese Government Bureau, 1963). For a discussion of Hatano's thought see G. K. Piovesana's *Recent Japanese Philosophical Thought, 1862–1962* (Tokyo: Enderle, 1963), pp. 123–131.

Gino K. Piovesana, S.J. (1967)

HAYASHI RAZAN

(1583–1657)

Hayashi Razan, the Japanese Confucianist, helped establish the Zhu Xi (Japanese: Shushi) school as the state doctrine of the Tokugawa government (1603–1867), which played an important role in shaping the national character. Hayashi, who was born in Kyoto, began studying Confucianism at the age of twenty-two, under Fujiwara Seika (1561–1619) and like his teacher abandoned Buddhism for the Neo-Confucianism of the twelfth-century Chinese philosopher Zhu Xi. Fujiwara recommended his talented pupil to Tokugawa Ieyasu as official adviser, a post Hayashi continued to fill under Ieyasu's successors. Through his son Gahō (1618–1680) and grandson Hōkō (1644–1732), both erudite Neo-Confucianists, Hayashi's influence spread. Gahō and Hōkō became hereditary heads of the Confucianist college (Shōheikō) of Edo (Tokyo), center of Japan's orthodox Zhu Xi-ism. Hayashi is credited with an important role in the various codes promulgated by the Tokugawa to reorganize the country under strict military rule. That he also determined educational policy is beyond dispute.

Hayashi, in contrast to his master Fujiwara, was very intolerant toward other doctrines—specifically, Wang Yangming Confucianism, Laozi, Buddhism, Christianity. Thus he is noted more for his negative polemics than for developing Zhu Xi's ideas, which he followed rather faithfully. An instance of deviation from Zhu Xi is Hayashi's almost monistic conception of *ri,* the principle, together with *ki,* the material-force. He came near to identifying these two basic concepts, thus approaching the rival school of Wang Yangming. Nevertheless, he sharply criticized Wang Yangming's "intuitive knowledge" and Kaibara Ekken's views. Hayashi disapproved of Laozi's emphasis on the "Nameless," or the Way understood as the indescribable Great One, intent as he was on stressing social relationships. For Buddhism's escape from society and neglect of loyalty and filial piety he had nothing but scorn, fighting until his death against influential Buddhist monks. His strictures on Christianity were many (Christians were then being persecuted and banished from Japan). He focused, however, on ethical questions and social differences neglected by Christianity. Only with Shintoism did he desire compromise and amalgamation.

See also Buddhism; Confucius; Japanese Philosophy; Kaibara Ekken; Laozi; Wang Yangming; Zhu Xi (Chu Hsi).

Bibliography

See W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 350–361, and Kyoto Shisekikai, ed., *Razan Sensei Shibunshū* (A collection of poems and essays by Master Razan), 4 vols. (Kyoto, 1918–1921). A biography in Japanese is Hori Isao, *Hayashi Razan* (Tokyo, 1964); this work has a good bibliography of Japanese references. For further bibliography see the Japanese Philosophy entry.

Gino K. Piovesana, S.J. (1967)

HAZLITT, WILLIAM
(1778–1830)

William Hazlitt, the English essayist, journalist, and critic, began his literary career as a "metaphysician," and the principles of his youthful philosophical writing survived to govern his thought during the years when a more brilliant prose style won him fame. Born at Maidstone, Kent, the son of a Dissenting minister, Hazlitt kept faith politically with his Unitarian heritage, but at an early age revolted against his father's rationalistic theology. After trying unsuccessfully to become a painter, he turned in his thirties to journalism and to popular lecturing, and until his death made his living in London as a writer for periodicals. Twice unhappily married, always the fierce defender of both the French Revolution and Napoleon Bonaparte, Hazlitt succeeded in alienating most of his friends and much of his public, although his critical influence on the literature of his time was perhaps second only to Samuel Taylor Coleridge's. Unlike Coleridge, his erstwhile friend and mentor, Hazlitt did not ground his thought in a version of the new Idealism; he stands alone in his age as a romantic thinker who developed a critique of empiricism that nonetheless supported the values and methods of the empiricist tradition.

Hazlitt continued the redefinition of the individual begun by William Godwin in *Political Justice* (1793). Four years before his first meeting with Coleridge in 1798, and while still a student at Hackney College in London, Hazlitt conceived his "metaphysical discovery"—a refutation of necessary egoism. Actually, his position had been anticipated by Joseph Butler and David Hume, but his arguments were original in his insistence on imagination—a power inseparable yet distinct from present sensation and past feeling—as the source of voluntary action, and even of self-consciousness.

His first book—*An Essay on the Principles of Human Action,* to which was added *Some Remarks on the Systems of Hartley and Helvétius* (1805)—argued that ideas of good determine conscious pleasure and self-interest, not the reverse, and that the same "reasoning imagination," which alone can unify sensations from moment to moment, is responsible for all the mind's "associations" except those arising from mere contiguity in experience. In his lectures at the Russell Institution in 1812 on the "Rise and Progress of Modern Philosophy," this line of thought inevitably led Hazlitt to challenge all epistemology, including George Berkeley's and Hume's, that relied on the Lockean premise of "simple" impressions in perception. To perceive the simplest object requires a "general idea," or some act of mind to "comprehend" objects in their sameness or wholeness before qualities can be differentiated. Failure to recognize an activity of mind inhering in sense perception itself had led, he believed, to the vain war between philosophies of "Necessity" and of "Liberty"—between a mechanism or materialism that reduced mind to sensation and an idealism like Immanuel Kant's that mistook man's formative consciousness for a power of will essentially free of sensory experience.

Hazlitt also came to oppose, then, the transcendentalism that Coleridge introduced from Germany. As is clear from a *Prospectus* (1809) for his projected history of English philosophy, Hazlitt saw himself as a loyal reformer of empiricism, although he admittedly left unresolved the central problem of the degree to which ideas are determined by the mind itself, on the one hand, and by "nature" on the other. In part, it was his belief that this dualism must remain intractable to reason which made him forsake formal analysis for the "familiar style" of his literary journalism. Averse to system and always more concerned with the cultural impact of ideas, he began, after 1812, to turn from an analysis of the formal problem to an exploration of the interaction of mind and world in experience as it is known by the self in life or realized by "genius" in the arts. Still affirming that "the mind is one," he made his theme the "everlasting contradiction" of man's nature—the "action and reaction" between the mind and the passional self as dialectical functions of the same unity of consciousness.

From his awareness of this conflict in consciousness Hazlitt forged no metaphysic of his own beyond a vague vitalist belief that "the spirit of life and motion" gave the mind a radical "sympathy" with the physical world. In religion he seems to have remained a modest agnostic, certain only that God is intellectually unknowable. Hazlitt thought that only in the aesthetic mode of imagination could the mind transcend experience, and even then it could attain to no intuition beyond "the soul of nature." The insistence that "passion" is the source both of man's freedom and of his bondage—a bondage to individual "character" that nonetheless implies the freedom of the self to sympathize with other selfhood—underlies Hazlitt's polemic on all fronts; it links his criticism of Thomas Robert Malthus and the utilitarians to his aesthetic theory that organic particularity is the basis of value in the arts.

In the England of 1830, when Hazlitt died impoverished in London, a humanism so darkly paradoxical found little favor; but his powers as a thinker have been increasingly recognized, and he appears today as the versatile Montaigne of his age, often prefiguring in his essays the dynamicist philosophies of Friedrich Nietzsche, Henri Bergson, William James, and Sigmund Freud.

See also Bergson, Henri; Berkeley, George; Butler, Joseph; Coleridge, Samuel Taylor; Empiricism; Freud, Sigmund; Godwin, William; Humanism; Hume, David; Idealism; James, William; Kant, Immanuel; Malthus, Thomas Robert; Montaigne, Michel Eyquem de; Nietzsche, Friedrich.

Bibliography

Hazlitt's philosophical writings are to be found mainly in *Complete Works,* edited by P. P. Howe, 21 vols. (London, 1930–1934), Vols. I, II (which includes the 1812 lectures), and XX. For studies on Hazlitt, see Elisabeth Schneider, *The Aesthetics of William Hazlitt* (Philadelphia: University of Pennsylvania Press, 1933); Herschel Baker, *William Hazlitt* (Cambridge, MA: Belknap Press of Harvard University Press, 1962); James Noxon, "Hazlitt as Moral Philosopher," in *Ethics* 72 (1963): 279–283. For further bibliography, see Elisabeth Schneider, "William Hazlitt," in *The English Romantic Poets and Essayists: A Review of Research and Criticism,* edited by C. W. and L. H. Houtchens (New York: Modern Language Association of America, 1957).

John Kinnaird (1967)

HEAT, SENSATIONS OF

See *Pain*

HEAVEN AND HELL, DOCTRINES OF

One of the most basic and existentially engaging of all questions has to do with the possibilities for human happiness and fulfillment on the one hand, and misery and loss on the other. Christian theology has returned the striking answer that the possibilities are truly extreme. According to Christian theology, the world is such that humans can experience perfect happiness, delight, and satisfaction, and do so forever. Indeed, that is just what human beings were created for: an eternal relationship with God that will fulfill humankind's best potentialities and aspirations. But the flip side of this is also possible—namely, that people may fail to achieve this relationship with God and thereby come to utter ruin and misery, a condition that is also believed to be eternal.

So understood, heaven and hell have provided an important moral source for European culture for the better part of two millennia. Not only have they served as moral sanctions that assure people that they are ultimately accountable for their actions, but heaven and hell also have been central to the majestic vision of life and its meaning that flows from belief in a God of perfect character and infinite power.

The fundamental logic of these beliefs is not unique to Christianity and European culture—it is common to Judaism and Islam as well. Of course the details differ in important respects, especially with respect to the crucial issue of the conditions for achieving heaven. In other

words, one's beliefs about the nature and conditions of salvation will be closely connected to one's conception of heaven and hell. But the point for emphasis is that belief in heaven and hell are not peripheral to theology, but are integral to traditional theistic faith, whether Jewish, Christian, or Muslim.

HISTORICAL DEVELOPMENTS OF THE DOCTRINES

Belief in an afterlife is either absent or ambiguous in early Jewish scripture. And even where that belief occurs, it is not always clear that there is a distinction between the fate of the righteous and that of the wicked. *Sheol*, the place of the dead, was conceived to be a place of shadowy existence without clear moral distinctions. A more developed view of the afterlife grew out the Jewish understanding of their covenant with God. While the possibility of punishment for disobedience to the covenant was always recognized, such punishment was understood as confined to this world. Increasing awareness of the injustices of this world led to calls for moral distinctions in the afterlife that would rectify the wrongs of this life, of which the book of Job is perhaps the most famous example. Belief in a double resurrection—of the wicked as well as the righteous, after which the wicked will be punished—emerged later in some Jewish scriptural texts. In extrabiblical literature, heaven and hell have been matters of considerable speculation among rabbis during both the time before the rise of Christianity and Islam, and after.

In the New Testament scriptures there is also significant diversity, and some texts appear to teach that the wicked will be annihilated whereas others appear to teach that all will eventually be reconciled to God. The view that came to predominate in Christian theology—also based on numerous New Testament texts—is that all will be resurrected, but that the wicked, perhaps constituting the majority of humanity, will be banished from the presence of God and forever lost in the misery of hell.

Many notable traditional theologians have conceived of hell as an eternal punishment that is justly imposed on sinners. In the Christian tradition, Augustine, Aquinas, Anselm, and Jonathan Edwards are among those who have formulated influential arguments in favor of this conception of hell. Anselm formulated his version of this argument in his famous account of the purpose of the atonement of Christ. People owe God total and perfect honor, Anselm argued, so any sin against him puts them in infinite debt to him that accordingly deserves infinite punishment. Edwards developed a similar argument by appealing to God's infinite nature. Because God is infinite in his loveliness, honor, and authority, Edward's believes that a person's obligation to love and honor God is likewise infinite. To fail in this obligation is to merit infinite consequences. Moreover, traditional theologians typically held that repentance after death is impossible, and consequently, no one may escape from hell.

In elaborating the punishment view of hell, traditional theologians often distinguished between the "pains of sense" and the "pain of loss." The former of these was typically understood to include literal fire of agonizing intensity, whereas the latter emphasized the unhappiness that naturally results from being separated from God, the true source of all joy and happiness. This picture of hell, along with its corresponding vision of heaven as a place of unbounded delight, has not only haunted the popular imagination but has also been a powerful source of inspiration for classic works of art, both visual and literary.

Despite its important role in both theology and the broader culture, belief in heaven and hell has been in decline in the European and North American world ever since the onset of modernity. The reasons for this decline are complex and are no doubt related to the more general defection from religious belief during this same period. The doctrine of hell, in particular, has lost credibility among believers as well as unbelievers largely because many see it as morally implausible.

CONTEMPORARY ACCOUNTS OF THE DOCTRINE

Both heaven and hell, however, have received renewed attention from contemporary philosophers as part of the revival of interest in philosophy of religion. A number of philosophers have moved beyond issues germane to generic theism to explore issues generated by distinctively Christian belief, including Trinity, incarnation, atonement, and the nature of salvation. Heaven and hell are closely connected to these beliefs, especially those pertaining to salvation. Heaven and hell have also played an important role in discussion of the perennial problem of evil and the project of theodicy. Whereas hell is typically seen as a particularly difficult aspect of the problem because it involves the prospect of eternal recalcitrant evil, heaven is often invoked as an essential component of a satisfactory theodicy. Only the hope of eternal life, it is argued, provides adequate grounds to believe the horrific tragedies of this life may be fully healed and redeemed.

Much of the contemporary discussion of hell has centered on the traditional arguments defending the claim that eternal torment is the just punishment for

human sin. Among those who have subjected these arguments to searching critical scrutiny are Marilyn Adams, Jonathan Kvanvig, and Charles Seymour. This critique begins by contesting the claim that human sin could ever be infinitely serious. Even the most notorious of sinners, such as Hitler, have done only finite evil and caused finite harm, however enormous it is. Next, it is contended that a just punishment should fit the crime. Thus, if God is perfectly just, he cannot punish human sin with infinite punishment. So eternal hell cannot be defended as a just punishment for sins committed in this life. There is a general consensus among contemporary philosophers that this critique is sound, so those who affirm the doctrine of eternal hell have turned to other arguments to make moral sense of it.

The most common strategy is to appeal to libertarian freedom to show how eternal hell can be compatible with God's perfect love and power. That is, it is contended that people have the freedom to reject God, even to the point of being forever separated from him. C. S. Lewis famously summed up the essence of this view in his remark that the doors of hell are locked from the inside. In the same vein, Richard Swinburne has defended the doctrine of hell on the grounds that people may, over time, form the sort of character that can no longer choose God and the good. Those who take this position thus typically affirm the pain of loss, but downplay or deny the pains of sense.

Kvanvig (1993) has defended a variation on this position that he calls the "issuant conception of hell." His position is so called because he believes the doctrine of hell should issue from the same character of God as the doctrine of heaven—namely, his love. It is a mistake, he thinks, to stress love only with reference to heaven, while emphasizing justice in connection with hell. The final choice everyone faces, according to Kvanvig, is either a relationship with God or annihilation, for to choose to live independently of God is in fact to choose annihilation, because living independently of God is actually impossible. Of course, God prefers that all persons accept his love, but he respects the freedom of those who reject a relationship with him.

However, not all who reject God choose annihilation in a clear and settled way. It is precisely because of his love that he allows them to remain in existence. Kvanvig's view is accordingly a "composite" view because it allows for both eternal separation as well as annihilation. God need not force people to choose either a relationship with him or extinction, so this allows the option of everlasting separation from him.

Seymour has focused on human choice in developing a defense of eternal hell that he calls "the freedom view." His fundamental definition of hell is that it is "an eternal existence, all of whose moments are on the whole bad" (Swinburne 1983). For this to be true of hell, he thinks it is not enough for hell to have the pain of loss—it must also include pains of sense. His appeal to freedom is crucial for he rejects the traditional arguments for the claim that sins committed in this life could be sufficiently serious to warrant eternal punishment. Rather, it is the continuing choice to sin that keeps sinners in the perpetual pains of hell.

Seymour believes that sinners can in principle repent and would be accepted by God if they did, so if they remain in hell it is due to their choice to persist in sin.

CONTEMPORARY CHALLENGES TO HEAVEN AND HELL

A growing number of Christian philosophers are challenging the doctrine of eternal hell in favor of a doctrine of universal salvation. Some Muslim thinkers have also advanced the speculation that all may be saved in the end. Not surprisingly, Christian philosophers who challenge eternal hell typically focus on libertarian freedom and the crucial role it plays in the contemporary defense of the doctrine.

Thomas Talbott (2003) has mounted a sustained attack on the doctrine of eternal hell, building his case on both biblical and philosophical grounds. In his biblical arguments, he has attempted to show that the New Testament is best interpreted as affirming that all will eventually be saved. He thereby aims to undermine one of the main pillars of the orthodox view of hell—namely, the contention that scripture requires Christians to believe it. Talbott's philosophical case against eternal hell largely focuses on his claim that the idea of choosing hell is finally incoherent.

His argument for this claim hinges on his account of what is involved in freely choosing an eternal destiny. In short, such a choice must be fully informed, and once the person making the choice gets what he or she wants, then it must be the case that the choice can never be regretted. This means that the person must be free from ignorance and illusion both in the initial choice as well as later. One must fully understand what has been chosen while freely persisting in that choice.

Given these conditions, Talbott thinks there is an obvious and important asymmetry between choosing fellowship with God as an eternal destiny, on the one hand,

and choosing hell as an eternal destiny on the other. Whereas the first of these obviously is possible, the latter is not. The reason for this is because there is no intelligible motive for choosing hell if one is free of ignorance and illusion. One may temporarily choose evil under the illusion that so choosing will make one happy. But God will eventually shatter this illusion by making one ever more miserable until the point is reached that one must repent and turn to God. Thus, Talbott affirms the view that universalism is necessarily true, in contrast to the more common claims that universalism is possibly true or probably true.

Marilyn Adams has also criticized the reliance on libertarian freedom in traditional theodicy, contending that its proponents exaggerate the dignity of human nature as something so sacrosanct that not even God may legitimately interfere with it. She sees this tendency particularly in the doctrine of hell, especially in the mild versions, which hold that hell is simply the natural consequence of freely choosing to reject God and the love he offers. Adams complains that advocates of mild hell tend to assume that God and human adults are moral peers in their insistence that they have the right to resist God and choose evil instead. As she sees it, this is not the appropriate sort of respect for God to pay to the likes of humans.

Indeed, the deeper difficulty here is that free will approaches underestimate what she calls the "size gap" between Divine and created persons. Whereas free will approaches picture the relationship between God and human persons with the analogy of parents and adolescent or adult children, Adams thinks it is better modeled by the relationship between a mother and an infant or a toddler. In the latter relationship, there is little if any sense that the child is free and responsible and that it would be wrong to interfere with his choices. This nicely serves Adams's view that God can save everyone in the end, and relieves her of the worry of how God may accomplish this without violating human freedom. If God needs to causally determine some things in order to prevent the everlasting ruin of some of his children, this should not be seen as an insult to our dignity.

The philosophical credibility of the doctrine of hell will largely depend on one's judgments about the nature and value of freedom as well as one's views of moral psychology. Those who disagree with Adams will argue that freedom is of sufficient value itself—or is the means to other goods of sufficient value—and that God will not override it to save us. In a similar vein, Talbott's critics, including the present writer, have argued that there are, contrary to his claims, intelligible motives for the choice of eternal damnation. Indeed, an essential component of freedom is people's ability to deceive themselves and turn away from the truth. If so, then God may not be able to shatter people's illusions without destroying their freedom.

Whereas the choice of heaven is easier to grasp from the standpoint of moral psychology because it is the choice of true happiness and fulfillment, some have argued that the notion of eternal joy is a dubious notion. Bernard Williams (1993) has made the case that the notion of eternal joy is incoherent because any life of endless duration would inevitably become boring, no matter how delightful the experiences it offered. Defenders of heaven have responded to this challenge in various ways, depending on how they conceive of the life everlasting. Two broadly different accounts of heaven have been prominent in the Christian tradition. On one end of the spectrum is the theocentric vision, which emphasizes the beatific vision as a timeless experience of contemplating the infinitely fascinating reality of God in all his aspects. On the other end of the spectrum is the anthropocentric view, which pictures heaven in terms familiar to this life, purged of course of the evil and suffering that currently mar human happiness.

However these debates continue and whatever resolutions may be achieved, it is apparent the renewed interest in heaven and hell brings into vivid focus some of the most profound issues that animate the philosophical enterprise. Not only the nature and ground of people's moral commitments, but their understanding of the meaning of their lives and their various configurations of joy and sorrow, hinge on what is believed about heaven and hell.

See also Immortality.

Bibliography

Adams, Marilyn McCord. "Hell and the God of Justice." *Religious Studies* 11 (1975): 433–437.

Adams, Marilyn McCord. *Horrendous Evils and the Goodness of God*. Ithaca, NY: Cornell University Press, 1999.

Bernstein, Alan E. *The Formation of Hell: Death and Retribution in the Ancient and Early Christian Worlds*. Ithaca, NY: Cornell University Press, 1993.

Kvanvig, Jonathan L. *The Problem of Hell*. New York: Oxford University Press, 1993.

Lewis, C. S. *The Problem of Pain*. New York: Macmillan, 1944.

Seymour, Charles. *A Theodicy of Hell*. Dordrecht, Holland: Kluwer Academic Press, 2000.

Smith, Jane Idleman, and Yvonne Yazbeck Haddad. *The Islamic Understanding of Death and Resurrection*. New York: Oxford University Press, 2002.

Swinburne, Richard. "A Theodicy of Heaven and Hell." In *The Existence and Nature of God*, edited by Alfred J. Freddoso. Notre Dame, IN: University of Notre Dame Press, 1983.

Talbott, Thomas. "A Case for Christian Universalism." In *Universal Salvation: The Current Debate*, edited by Robin Parry and Chris Partridge. Carlisle, U.K.: Paternoster, 2003.

Walls, Jerry L. *Heaven: The Logic of Eternal Joy*. New York: Oxford University Press, 2002.

Walls, Jerry L. *Hell: The Logic of Damnation*. Notre Dame: University of Notre Dame Press, 1992.

Williams, Bernard. "The Makropulos Case: Reflections on the Tedium of Immortality." In *The Metaphysics of Death*, edited by John Martin Fischer. Stanford, CA: Stanford University Press, 1993.

Jerry L. Walls (2005)

HEBBEL, CHRISTIAN FRIEDRICH

(1813–1863)

Christian Friedrich Hebbel, the German poet and playwright, was born in Wesselburen in the duchy of Holstein and died in Vienna. His father, an impoverished bricklayer who became destitute as a result of having guaranteed a loan that was defaulted, hated this son who showed no aptitude for earning a living. The boy's mother was more indulgent and protected him from the brutality of the father. It was thus possible for young Hebbel to keep alive his consuming passion for learning. At the age of fourteen he was employed as a clerk by a parish official named Mohr, who allowed him to use his library. Mohr treated Hebbel as a common servant, however, and for this Hebbel never forgave him.

Through the good offices of Amalie Schoppe, the editor of a popular magazine, Hebbel received enough money to go to Hamburg in order to try to complete his fragmentary education. There he met Elise Lensing, a seamstress ten years his senior who cherished an abiding love for him; over the years she gave him clothes, lodging, money, and two sons, both of whom died young. Hebbel, who was ridden by his demon to acquire learning and develop himself as a writer, refused to marry Elise. Instead, he went on to study at the universities of Heidelberg and Munich. In the late winter of 1839, he made the arduous trip from Munich back to Hamburg on foot. In the same year he completed his first play, *Judith,* which he cited as his chief accomplishment when he applied for a travel stipend to King Christian VIII of Denmark. The king granted the stipend, and Hebbel went to Paris and from there to Rome. Because his resources were dwindling, he struck out for Germany by way of Vienna, where he met the talented actress Christine Enghaus, to whom he became engaged after three months. Whether or not Hebbel was largely influenced in this decision by the prospect of financial and social security, the marriage was a happy one and enabled Hebbel to take a place of honor in artistic and intellectual circles. His early death must be attributed in large measure to the hardships he had endured in order to realize his genius.

With the desperate seriousness of the self-educated man, Hebbel dedicated himself to presenting in artistic form his solution, sometimes characterized as "pantragic," of what he considered the ultimate philosophical problem, the incomprehensible escape of the individual from the Absolute or Idea, man's freedom in relation to God.

In Hebbel's dualism individual forms exist only by virtue of having differentiated themselves from the Absolute. Their struggle to maintain themselves as separate entities is a rebellion, the primeval sin of individuation. The sinfulness of the individual consists merely in the fact that he exists, and it is in no way dependent upon the nature or direction of his individual will. For his sinfulness the individual must be punished; he will have to submerge his particular being in the undifferentiated whole. The more splendid, vigorous, and powerful he is, the greater is the threat he poses to the Absolute and the more tragic is the struggle, which can end in only one way. There is only one necessity—that the Absolute maintain itself. However, although the existence of individual forms threatens the Whole, it is precisely the process of individuation that gives life to this closed system. If it were not for the mysterious freedom of the individual forms, the Absolute would become rigid and lifeless. The total life process is dependent on the metabolic flow of individual forms, which may appear at one point; may be submerged forever; or may, whether they retain their identities or their elements enter into new combinations, reappear at another point only to lose individuality again in the never-ending compact flux of history, compact because nothing new enters the universe and nothing leaves it.

It is the common task of philosophy and art, particularly drama and more specifically Hebbel's drama, to describe and make understandable this supreme philosophical problem. Philosophy must fail in its part of the common task, to determine the original cause of individuation, because this ultimate cause is unfathomable. But

the drama is not concerned with this question. It accepts individuation as the prime condition of life and presents the tragic struggle of the All and the one in a way that makes it comprehensible to aesthetic intuition. In the drama the metaphysical breach is closed; the defeat of the tragic hero mirrors the cosmic process.

In order to achieve his aim, Hebbel sets the action of his plays at critical times in history, for at such times the relation of the individual to the Whole is most poignantly manifested. In some plays—for example, *Judith, Maria Magdalene, Herodes und Mariamne*—the prevailing form of the Idea is shown to be on the verge of breaking up. In others—*Genoveva, Agnes Bernauer, Gyges und sein Ring*—the prevailing form of the Idea, although threatened, affirms itself and persists. In both instances, whether the individual is opposed to the Idea or is an instrument of it, the end is tragic, and all individuals meet the same fate—they are crushed and absorbed by the Whole.

Hebbel always insisted that despite obvious parallels he had evolved his metaphysical truths independently of the romantic nature philosophy of German idealism. For a long time, in the absence of specific evidence to the contrary, many literary historians accepted this assertion. Recent research, however, has shown that Hebbel had early steeped himself in certain writings of Gotthilf Heinrich von Schubert, the natural scientist and philosopher, and Ludwig Feuerbach. The ideas he found there he experienced with such intensity that he incorporated them into his own psychic structure although his pride as a self-taught man did not allow him to acknowledge his debt.

See also Absolute, The; Feuerbach, Ludwig Andreas; Idealism.

Bibliography

WORKS BY HEBBEL

Sämtliche Werke, historisch-kritische Ausgabe. Edited by R. M. Werner. Berlin, 1904–1907. The major sources for Hebbel's theories are *Mein Wort über das Drama* and *Vorwort zu Maria Magdalene,* both in Part I of this edition. Equally important are various passages in the diaries (Part II) and letters (Part II).

WORKS ON HEBBEL

Campbell, T. M. *The Life and Works of Friedrich Hebbel.* Boston: R.G. Badger, 1919.

Flygt, Sten G. *Friedrich Hebbel's Conception of Movement in the Absolute and in History.* Chapel Hill: University of North Carolina, 1952.

Kreuzer, Helmut, ed. *Hebbel in neuer Sicht.* Stuttgart: W. Kohlhammer, 1963.

Kuh, Emil. *Biographie Friedrich Hebbels.* Vienna and Leipzig, 1907.

Liepe, Wolfgang. *Beiträge zur Literatur und Geistesgeschichte.* Neumünster, Germany: K. Wachholtz, 1963. Contains several essays illuminating the relationships between Hebbel's thought and the thought of Friedrich Wilhelm Joseph von Schelling, Ludwig Feuerbach, and Gotthilf Heinrich von Schubert.

Meetz, Anni. *Friedrich Hebbel.* Stuttgart: Metzler, 1962.

Purdie, Edna. *Friedrich Hebbel: A Study of His Life and Work.* Oxford: Oxford University Press, 1932.

Scheunert, Arno. *Der Pantragismus als System der Weltanschauung und Aesthetik Hebbels.* Leipzig, 1930.

Sten G. Flygt (1967)

HEDONISM

Hedonism (Greek, ἡδονή, "pleasure") is a term that refers to either of two distinct but related views, one a thesis in normative ethics, the other a generalization about human psychology.

ETHICAL HEDONISM

The first view, called "ethical hedonism," affirms that only pleasure is intrinsically desirable and that only displeasure (or pain) is intrinsically undesirable. More fully stated, it is the thesis that only pleasant states of mind are desirable in themselves; that only unpleasant states of mind are undesirable in themselves; and that one state of affairs is more desirable in itself than another state of affairs if and only if it contains more (in some sense) pleasant states of mind than the other (the quantity of value in a state of affairs being measured by the quantity of pleasure in it).

This thesis has been defended by a distinguished line of philosophers from the early Greeks to the present, including Aristippus, Epicurus, John Locke, Thomas Hobbes, David Hume, Jeremy Bentham, J. S. Mill, and Henry Sidgwick. Other philosophers have thought that happiness is the only thing that is intrinsically desirable; and if saying that a man is happy at a given time is the same as saying that he is experiencing pleasure at the time, then their names could be added to this roster. Many philosophers, however, have thought that happiness is different from pleasure, and there has been disagreement and confusion about what "happy" and "pleasant" mean.

The hedonist thesis was a part of traditional utilitarianism, as represented, for instance, by Bentham and Mill, with their "greatest happiness principle." These writers

combined the generic principle of utilitarianism—namely, that an act is morally right if performing it would produce, or could reasonably be expected by the agent to produce, at least as much intrinsic goodness in the world as any other act the agent could perform at the time—with the thesis of hedonism about what is intrinsically good. Traditional utilitarianism is thus a species of utilitarianism that is defined as asserting just the "generic thesis"; other kinds of utilitarianism (for example, that of G. E. Moore) reject hedonism. Unlike utilitarianism, ethical hedonism is not at all a proposal about which acts are morally right; it is only an affirmation about which states of affairs are intrinsically good or desirable.

What is meant by saying that a state of affairs is intrinsically desirable, as opposed to simply desirable, is that it is desirable, good, worthwhile, worthy of choice, when *taken by itself,* viewed abstractly, and in particular considered without reference to consequences. Many things (for example, a visit to the dentist) are worthwhile in view of their consequences, which nobody would say are intrinsically desirable. The hedonist does not deny that other things are desirable; he denies only that they are intrinsically worthwhile. He agrees that something can be desirable instrumentally—as a means to an end—even when it is not intrinsically desirable. (A thing can, of course, be both intrinsically and instrumentally desirable: pleasant experiences can be good in themselves and also instrumentally good, if, for example, they are relaxing and enable one to work better on the following day.) He does add, however, that something is instrumentally desirable only to the extent that it is a means to later pleasure, since a thing can be instrumentally desirable only if it is a means to attaining the intrinsically desirable.

When consequences are taken into account, the hedonist's view about what states of affairs are desirable is apt to differ very little from the view of the nonhedonist. In fact, if one reads various writers' accounts of the "good life," one finds that they are pretty much alike, whether the author professes to be a hedonist or a nonhedonist. Thus, Epicurus, for instance, advocated a simple life devoted to philosophical reflection, with a diet of bread, cheese, and milk, and with its tranquility unendangered by surging bodily passions. And J. S. Mill affirmed that having a good character is "part" of a person's happiness, so that according to him, character is intrinsically good after all by virtue of the fact that it is a part of happiness. Some hedonists, however, have advocated a more distinctive ideal for living: the Greek Aristippus thought that physical enjoyments are the richest source of pleasure and should be fully cultivated.

The meaning of the hedonist's thesis, of course, depends on what is meant by "pleasure." It is true that the associations of the word *pleasure* are such that if an English-speaking person says he favors a "life of pleasure," he is naturally taken to be advocating a life dedicated to the sensory enjoyments—wine, women, and song. Hedonists have not intended the term to carry this implication, however, and the strict meaning of the term does not. It is perfectly correct for a student to say, "I got a great deal of pleasure out of writing that paper." To say that an experience is pleasant (for example, "a pleasant evening"), is, in a strict sense, simply to say that one enjoyed it, or that one enjoyed himself during it. Thus, hedonism is done least injustice if it is taken as simply saying that an intrinsically desirable state of affairs is always a state of consciousness in which the person is enjoying himself in one way or another. Since reflection, reading, and creation are activities that people often enjoy, the hedonist means to include these activities, or states of mind, in the category of "pleasures," just as much as the so-called passive enjoyments, such as eating, drinking, and sex.

Hedonists have often disagreed about the proper analysis of "pleasure" or "enjoyment." Epicurus, for instance, said that pleasure is simply the absence of painful want or longing. Moreover, since the early 1900s, psychologists have also disagreed substantially on this point, some holding that pleasure is a special kind of sensation, others that it is a quality of certain kinds of feeling, and so forth. In recent years a considerable body of philosophical literature has accumulated on the subject of the analysis of "pleasant." While a generally accepted conclusion has not yet been reached, it is plausible to say that a person is enjoying himself (that is, his state of mind is pleasant) if and only if at the time he likes his experience or activity *for itself,* in the sense that, aside from moral considerations or considerations of consequences or of the possibility that something he likes even better could be substituted, he does not wish to change it and in fact would wish to avoid changing it if such a change impended. If this interpretation is accepted, the thesis of hedonism becomes the affirmation that a state of affairs is intrinsically desirable if and only if it is, or contains, an activity or experience which, at the time, the person likes for itself; and one state of affairs is more desirable intrinsically than another if it is, or contains, an experience or activity which, at the time, is liked better for itself. States of affairs which the hedonist thesis apparently rules as being only instrumentally rather than intrinsically desirable, from the point of view of a particular person, are such things as fame after his death and states of knowl-

edge and character (since the latter are not experiences or activities at all, but capacities or dispositions).

PSYCHOLOGICAL HEDONISM

Many (but not all) ethical hedonists have supported their ethical affirmation of hedonism by an appeal to a psychological doctrine known as "psychological hedonism." This theory historically has taken rather different forms; the significance of each for ethical hedonism must be assessed separately. The element common to them is the assertion that actions or desires are determined by pleasures or displeasures, whether prospective, actual, or past. The importance of the theory, however, transcends its relation to ethical hedonism: certain psychologists today are inclined to accept some form of it as a correct account of human motivation.

GOAL IS PLEASURE. The first and historically most important form of the theory of psychological hedonism may be called the "goal is pleasure" theory, according to which a person is motivated to produce one state of affairs in preference to another if and only if he thinks it will be more pleasant, or less unpleasant, for himself. This thesis, of course, is not intended to be a generalization about simple reflex or habitual behavior. The "belief" in question need not be explicit in the sense of having been verbally formulated before action; it may be an unformulated assumption. The theory is not simply about purposive action; it is also a theory about desire: a person is asserted to want one thing more than another if and only if he thinks its occurrence will be more pleasant for him.

The relation of this form of psychological hedonism to ethical hedonism may be explained by the following argument, often used by ethical hedonists. It is assumed as a major premise that something is intrinsically desirable if and only if it is something that people desire for itself. The minor premise is the "goal is pleasure" theory—namely, that people want only pleasure for itself. It is therefore concluded that pleasure is the only thing that is intrinsically desirable. The third-century writer Diogenes Laërtius said of Epicurus that "as proof that pleasure is the end he adduces the fact that living things, so soon as they are born, are well content with pleasure and are at enmity with pain, by the prompting of nature and apart from reason."

Contemporary ethical hedonists seldom appeal to the "goal is pleasure" theory to support their views, partly because the theory seems incompatible with obvious facts. For instance, political figures seem to take a strong interest in securing favorable notice in books on history that will appear after their death. This motivation obviously does not depend on the belief that the future event will be pleasant for them personally. Again, individuals often appear to risk personal loss for some moral principle or in order not to forsake a friend (this is illustrated by Dean Acheson's famous remark, "I will never turn my back on Alger Hiss"). Adherents of the "goal is pleasure" theory tend to explain such facts by saying that the individual would be unhappy in the future—and knows he would be—if he failed to live by his principle or forsook his friend; hence, the action is motivated by a calculation of personal pleasure after all. What the theory must hold, though, is that a belief to this effect, at least vaguely espoused by the agent, is a necessary condition of the motivation; and this seems implausible. Adherents of the theory may be confusing two things: the agent's belief that a certain future situation will be relatively more pleasant for him, and the agent's thought of that future situation being attractive or repugnant now. A person may say, "I *am unhappy with the idea* of dropping my friendship with *X*, in whose integrity I believe." This statement may be true and also an important clue to understanding his behavior. But this is very different from saying, "I'll continue my friendship with Mr. X *because I think I'll be less happy* if I don't," a kind of statement that would ordinarily be taken as proof that the person did not care about his friend. Adherents of the theory may always argue that the reasoning required by their theory takes place unconsciously, but the postulation of this is ad hoc, the only reason for it being that it saves the theory from conflict with observation.

MOTIVATION BY PLEASANT THOUGHTS. As suggested above, adherents of the "goal is pleasure" form of psychological hedonism sometimes confuse it with a different thesis which we may call the "motivation by pleasant thoughts" theory. This theory is the assertion that a person will choose to do *A* rather than *B* or will prefer *A* to *B* (whether an action or a situation), if and only if the thought of *A* (with its expected consequences) is more attractive, or less repugnant, than the thought of *B* (with its expected consequences). This theory is not obviously false: indeed, as a proposal about preference it could be an analytic proposition that sets forth one test we use to decide whether a person prefers one thing to another. As a proposal about action it is clearly a synthetic proposition. As such, it may not be able to explain the fact (if it is a fact) that sometimes the thought of doing *A* is not more attractive or less repugnant than the thought of doing *B*, but the agent simply decides to do *A* (perhaps he is required to make up his mind between the two).

Even if this form of psychological hedonism is true, however, it gives no support to ethical hedonism, since it sets no restrictions on the *kind* of goal which may be attractive or repulsive to a person. If support of ethical hedonism requires a demonstration that people desire only pleasure, then the present theory does not provide such support. For assuming that desiring a thing *means* finding the idea of it pleasant or attractive, it does not follow that only the idea of pleasure itself is attractive. Hence, it does not follow that only pleasure is desired, and it is therefore no part of the "motivation by pleasant thoughts" theory to assert that only pleasure is desired.

CONDITIONING BY PLEASANT EXPERIENCES. The third form of psychological hedonism, the "conditioning by pleasant experiences" theory, is a theory about the causal conditions of a person's wants or values. Roughly, it asserts that at least one's fundamental values can be correlated with past enjoyments or rewards, that these enjoyments are at least part of the causal explanation of the values, and that a person's values can be controlled by manipulating his enjoyments. If a person values ice cream, it is because in the past he has enjoyed ice cream (and not been made sick by it). The truth of this theory is hardly open to question insofar as it merely affirms that past enjoyments have some influence on likes and values; but its truth can be widely questioned if the theory is claimed to give a complete account of likes and values, which, according to experimental evidence, seem to be influenced by numerous factors. Acceptance of the theory, however, does not commit one to assert that persons desire only pleasure. The theory is consistent with saying that people want and value things such as posthumous fame or being a generous or courageous person. All the theory claims is that *whatever* values one has have been acquired because of past enjoyments or punishments of one sort or another—perhaps the enjoyment of parental praise or the punishment of parental reproaches.

FURTHER ARGUMENTS IN SUPPORT OF ETHICAL HEDONISM

Acceptable psychological theory, as we have seen, does not indicate that people desire only pleasure or things they think will be pleasant for them, or that people prefer *A* to *B* if and only if they think *A* will be more pleasant to them than *B.* Ethical hedonism, therefore, cannot appeal to psychological theory in support of its thesis.

Ethical hedonists sometimes rely on one or more of three other lines of argument in support of their view. The first line of reasoning is simply that ethical hedonism is an analytic truth that is true by definition. Locke, for instance, defined "good" as that "which is apt to cause or increase pleasure," and Benedict Spinoza defined it as "every kind of pleasure, and all that conduces thereto." The flaw in this contention, however, is that many people have at least thought either that some things other than pleasure are intrinsically good or that some kinds of pleasure are intrinsically bad. In the face of this, it is not easily claimed that "intrinsically good" simply *means* pleasant.

The second line of reasoning, which is more substantial, starts from the premise that it is usually agreed that at least some forms of pleasure are intrinsically good and proceeds by contesting the claim that anything else is intrinsically good. If the claims on behalf of other things are successfully refuted, it is concluded that ethical hedonism is left holding the field. The assessment of this line of reasoning is obviously a complex matter, since it presupposes conclusions about how to adjudicate ethical disputes. There is space here only to mention some examples frequently debated by hedonists and their opponents.

Critics of hedonism often urge that some kinds of pleasure are intrinsically bad—for example, malicious pleasure in the suffering of another person. And, they say, some unpleasant experiences are intrinsically good—for example, the punishment of one who has been cruel to another. Furthermore, it may be claimed that various things in addition to pleasure are intrinsically good: knowledge, certain traits of character, kindly or courageous deeds, life itself (at least the survival of mind with memory) even if it is not positively pleasant, being the object of respect or love on the part of other persons, being remembered after death, achievement, whether intellectual or aesthetic. Anyone who accepts any of these points cannot, strictly speaking, be an ethical hedonist.

A third, more practical, line of reasoning by hedonists has been the contention that their view makes possible scientific and objective evaluations of social planning which other views do not. For instance, if the question arises whether a certain tariff should be raised, the hedonist may say that his theory enables us (in principle at least) to decide objectively whether the tariff will do good, for we have only to decide whether a greater net sum of pleasures will be produced with or without the tariff.

This conception has come in for a great deal of criticism in recent decades, some of it unfair. One criticism, which appears repeatedly in the writings of economists, makes the point that we can know nothing about the

mental states of other persons, since there is no way of observing them directly; hence, the whole idea that theoretically an individual could determine the effects of a tariff on the happiness of anyone but himself is absurd. This criticism probably goes too far, but questions concerning other minds cannot be evaluated here. A more forcible objection is the following. If "is pleasant" is analyzed as meaning "is an experience liked at the time by the person, for itself," then presumably *A*'s experience can be said to be pleasanter than *B*'s, if *A* likes his experience more intensely. In theory, then, we might show that a tariff on bicycles would do more harm than good, if we could match every pleasant experience it would produce with an experience at least equally pleasant (one liked at least equally as intensely) and of at least equal duration, which the tariff prevents, *and* if in addition it costs us pleasures that are not matched with those it produces, or if some of the pleasures it costs us are more intense than the matching pleasures it produces, and the reverse is not the case. So far a decision could be reached, in principle at least. But it is possible that things might be too complex to permit such a simple matching. It might be that we would be forced to compare a more intense but brief pleasure with a less intense pleasure of greater duration; and it is not clear what would be meant by saying that one such experience "contains more pleasure" than the other. Thus, it is not clear that in principle the comparison could be made, except in special favorable situations. In this respect, however, the hedonist seems correct on one point: there is no other theory of the intrinsically desirable which makes such evaluations more scientific or more objective.

See also Aristippus of Cyrene; Bentham, Jeremy; Consequentialism; Diogenes Laertius; Epicurus; Happiness; Hobbes, Thomas; Hume, David; Locke, John; Mill, John Stuart; Moore, George Edward; Pleasure; Sidgwick, Henry; Utilitarianism; Value and Valuation.

Bibliography

PSYCHOLOGICAL HEDONISM

Duncker, K. "Pleasure, Emotion, and Striving." *Philosophy and Phenomenological Research* 1 (1940): 391–430.

Troland, L. T. *Fundamentals of Human Motivation.* New York: Van Nostrand, 1928. Ch. 16.

Young, P. T. "The Role of Hedonic Processes in the Organization of Behavior." *Psychological Review* 59 (1952): 249–262.

ETHICAL HEDONISM

Bentham, Jeremy. *Principles of Morals and Legislation.* Many editions.

Diogenes Laërtius. *Lives of Eminent Philosophers.* London: Heinemann, 1925. Bk. II (Aristippus) and X (Epicurus).

Ewing, A. C. *Ethics.* London: English Universities Press, 1953. Ch. 3.

Mill, J. S. *Utilitarianism.* Many editions.

Moore, G. E. *Ethics.* Oxford, 1949. Chs. 1–2.

Sharp, F. C. *Ethics.* New York: Century, 1928. Ch. 19.

Sidgwick, Henry. *Methods of Ethics.* London, 1922. Bk. III, Ch. 14.

MEASUREMENT OF PLEASURES

Alchian, A. A. "The Meaning of Utility Measurement." *American Economic Review* 43 (1953): 26–50.

Bentham, Jeremy. *Principles of Morals and Legislation.* Ch. 4. Many editions.

Cohen, Morris, and Ernest Nagel. *An Introduction to Logic and Scientific Method,* 293–298. New York, 1936.

McNaughton, R. "A Metrical Conception of Happiness." *Philosophy and Phenomenological Research* 14 (1954): 172–183.

Richard B. Brandt (1967)

HEGEL, GEORG WILHELM FRIEDRICH

(1770–1831)

Georg Wilhelm Friedrich Hegel, the German idealist philosopher, was born at Stuttgart and entered the theological seminary at the University of Tübingen in 1788. Among his fellow students were Friedrich von Schelling and the poet Friedrich Hölderlin. After graduating he became, in 1793, a resident tutor in the home of an aristocratic family at Bern, and in 1796 he took a similar post in Frankfurt. In 1800 he went to Jena, where Schelling had succeeded Johann Gottlieb Fichte as professor of philosophy and was developing an idealist philosophy of nature and metaphysics. Having been accepted as a teacher at Jena on the strength of his dissertation, *De Orbitis Planetarum* (1801), Hegel collaborated with Schelling in editing the philosophical journal *Kritisches Journal der Philosophie* and published his first book, *Differenz des Fichte'schen und Schelling'schen Systems der Philosophie* (1801). Notable articles by Hegel in the *Kritisches Journal* were "Glauben und Wissen" (1802) and "Über die wissenschaftlichen Behandlungsarten des Naturrechts" (1802–1803). At Jena, Hegel wrote his first major work, *Phänomenologie des Geistes* (*Phenomenology of Mind*, Würzburg and Bamberg, 1807). Completed about the time of Napoleon Bonaparte's victory over the Prussians at Jena in 1806, it was not published until 1807, after Hegel had left Jena to become editor of a daily paper at Bamberg in Bavaria.

In 1808, Hegel was appointed headmaster of a school in Nuremberg, a post he held until 1816. While at Nuremberg, Hegel published his *Wissenschaft der Logik* (*Science of Logic*)—Vol. I, *Die objective Logik* (2 vols., Nuremberg, 1812–1813, and Vol. II, *Die subjective Logik oder Lehre vom Begriff* (Nuremberg, 1816). From 1816 to 1818, Hegel was professor of philosophy at Heidelberg. There he published *Encyklopädie der philosophischen Wissenschaften im Grundrisse* (*Encyclopedia of the Philosophical Sciences in Outline*) in 1817. In 1818, Hegel was appointed professor at the University of Berlin, where he became famous and influential. *Naturrecht und Staatswissenschaft im Grundrisse* (*Philosophy of Right*) appeared there in 1821; a second edition, edited by E. Gans as *Grundlinien der Philosophie des Rechts*, was published in Berlin in 1833. In 1827 a second, much enlarged edition of the *Encyclopedia* appeared.

Hegel died during a cholera epidemic in 1831. After his death a group of his friends compiled an edition of his works in eighteen volumes (Berlin, 1832–1840). Several of Hegel's works were published for the first time in this edition: *Vorlesungen über die Aesthetik* (Lectures on aesthetics; translated as *The Philosophy of Fine Art*, edited by H. G. Hotho, 2 vols., 1835–1838); *Vorlesungen über die Philosophie der Geschichte* (*Lectures on the Philosophy of History*, edited by E. Gans, 1837); *Vorlesungen über die Philosophie der Religion* (*Lectures on the Philosophy of Religion*, edited by Philipp Marheineke, 2 vols., 1832); and *Vorlesungen über die Geschichte der Philosophie* (*Lectures on the History of Philosophy*, edited by K. L. Michelet, 2 vols., 1833–1836). This edition also contains notes taken by students of Hegel's comments on the *Encyclopedia* and on *Philosophy of Right*, which he was in the habit of using as textbooks.

In his biography, *Georg Wilhelm Friedrich Hegels Leben* (Berlin, 1844), Karl Rosenkranz referred to and quoted from the manuscripts of works written by Hegel prior to the publication of the *Phenomenology of Mind*. Not all the manuscripts known to Rosenkranz have survived, but toward the end of the nineteenth century Wilhelm Dilthey made a study of those that have and published an account and discussion of them in the *Proceedings of the Berlin Academy* in 1905. This has since received the title *Die Jugendgeschichte Hegels* and is reprinted in the fourth volume of Dilthey's collected works. Dilthey's pupil and editor, Herman Nohl, then published, under the title *Hegels theologische Jugendschriften*, the text of a great part of what Hegel had written while he was at Bern and Frankfurt. The chief of the writings unpublished during Hegel's lifetime are the essay "Das Leben Jesu" ("Life of Jesus," 1795), *Die Positivität der christlichen Religion* (*The Positivity of the Christian Religion*, 1796), and *Der Geist des Christentums und sein Schicksal* (*Spirit of Christianity and Its Destiny*, 1799). In 1915, Hans Ehrenberg and Herbert Link published, under the title *Hegels erstes System* (Heidelberg, 1915), an early version, written at Jena but never published by Hegel, of what later became the system sketched in the *Encyclopedia*. Since then Georg Lasson (*Hegels Jenenser Logik*, Leipzig, 1923) and Johannes Hoffmeister (*Hegels Jenenser Realphilosophie*, 2 vols., Leipzig, 1932) have published still other writings that Hegel had left unpublished. Thus, much more is now known about Hegel's writings and philosophical development than was generally known in the nineteenth century.

MAIN THEMES OF HEGEL'S PHILOSOPHY

MIND. In the preface to the *Phenomenology*, Hegel wrote that only mind (*Geist*) is real, and he constantly reiterated this view. (I have translated Hegel's *Geist* as "mind," in agreement with William Wallace's view that "to average English ears the word Spiritual would carry us over the medium line into the proper land of religiosity"—*Hegel's Philosophy of Mind*, Oxford, 1894, p. 1.) Thus, he must be regarded as a philosophical idealist. He wrote rather slightingly of George Berkeley, however, whose works he does not seem to have studied closely, and is sometimes described as an objective idealist in order to absolve him from suspicion of the subjective idealism that has often been attributed to Berkeley. Hegel's idealism presupposed the work of Immanuel Kant and was influenced by Fichte and Schelling, but his early unpublished writings show that he had preoccupations of his own, independent of his famous German predecessors.

When Hegel said that only mind is real, he did not mean that material things do not exist and that only minds do. Mind was not, in Hegel's view, a plurality of immaterial substances but a system of individuals actively developing their potentialities by embodying them in increasingly complex forms. A fundamental feature of mind, according to Hegel, is freedom, and nothing that is partial or finite can be wholly free. The mind that is the only reality is therefore infinite. Furthermore, no one is free unless he is conscious of what he is doing, and infinite mind is therefore self-conscious mind. Artists and statesmen, merchants and saints, all busy themselves with their more or less partial tasks without necessarily concerning themselves with what it is that they are doing. According to Hegel, it is the function of the philosopher

to make men conscious of what art and politics, commerce and religion, are, so that mind can exert itself to its utmost range and thus become absolute. Like Pythagoras, Plotinus, and Benedict de Spinoza, Hegel was a philosopher who held that philosophy is an activity that purifies and frees the mind.

DIALECTIC. Hegel is, of course, famous for his dialectical method, but it is enormously difficult to explain this in a brief compass. It should first be noted that Hegel set out his systematic writings in dialectical triads comprising a thesis, antithesis, and synthesis. Thus, he divided his *Encyclopedia*, in which he expounded his system as a whole, into three fundamental division sections—"Logic," "Philosophy of Nature," and "Philosophy of Mind." In the first he expounded the categories as developing forms of thought; in the second, he said "the Idea" is considered in its "otherness" (*Anderssein*) or externality; and in the third, mind is considered as existing "for itself," as conscious of itself and of the institutions it has given rise to. Within these main divisions there are further triadic subdivisions, although a very large number of subdivisions are not of this nature. It is therefore clear that Hegel himself regarded his whole work as a dialectical construction, with thought and nature as opposites united in mind and society, in the artistic and religious products of man, and, ultimately, in the activity of philosophical self-consciousness.

Hegel's system, then, has a dialectical structure, but what is his dialectical method? Hegel, like Spinoza, held that error resides in incompleteness and abstraction, but, unlike Spinoza, he held that the incompleteness and abstraction can be recognized by the contradictions they generate. It is the business of the philosopher, he held, to bring out the contradictions latent in partial or abstract views and to emphasize and elaborate them in such a way that less partial and less abstract views can be constructed that nevertheless retain in themselves what there was of truth in the original views. The same method is to be brought to bear on the less partial and less abstract views in their turn and to be pressed as thoroughly as it can be. This method of pressing and accentuating contradictions is not to be used merely to discard error but also to preserve truth. Because of the happy circumstance that in German *aufheben* means both "to cancel" and "to preserve"—its literal meaning is "to lift up"—Hegel was able to express this aspect of his view with brevity and acuity. The concept or view that is *aufgehoben* is transcended without being wholly discarded. Hegel's *Phenomenology of Mind* was an account of how various human attitudes—reliance on sense experience, the belief in substance, otherworldliness, strenuous moralism, and so on—all have some point and are yet contradictory, leading to the conclusion that "truth is a bacchanalian revel where not a member is sober," as Hegel put it in the Preface. His *Logic* gave an account of how the categories are related in this way. In his *Lectures on the History of Philosophy* he sought to show that the major philosophical outlooks from that of the Ionians on are, on the one hand, positive contributions that we could not do without and, on the other hand, contradictions that we have to overcome.

HISTORY. Another feature of Hegel's philosophy is its concern with history. Much as Hegel admired Plato's philosophy, he held that it was impossible to be a Platonist in the nineteenth century, when the philosophical context differed so greatly from that of Plato's day. In his *Lectures on the Philosophy of History*, Hegel argued that the history of man in the concrete was as much a progression as the history of his thought. This he deduced from the thesis that mind is of its very nature free. Thus, each historical epoch, according to Hegel, embodied some aspect of or stage in the development of man's free mind, and it would be absurd for an individual to go counter to his time except insofar as he was preparing the way for future epochs. Hegel borrowed this "progressivism," as it may be called, from the philosophers of the Enlightenment. It has greatly influenced Marxism.

CHRISTIANITY. Hegel thought his system provided a defense of Christianity, and both supporters and opponents of his system have taken this view of it. Those known as right Hegelians considered Hegel's apologetic successful, whereas the left Hegelians argued that his Christianity had been only superficial and his Christian terminology a disguise for something very different. In his system Hegel placed philosophy above religion in the dialectical scale, and this may give some support for the interpretations of the left. Yet there is ambiguity in Hegel's view on this, as on other important matters. On one hand, he held that only infinite mind is real; on the other hand, he held that infinite mind cannot be distinct from or beyond the finite and partial. He thought that these views were not incompatible, but it has been argued that the second is a denial of the first and, hence, a denial of any form of theism.

This entry will briefly describe Hegel's early works that were posthumously published in *Hegels theologische Jugendschriften*. It will continue with an account of the *Phenomenology of Mind*, Hegel's first important book,

and conclude with a brief discussion of the Hegelian system based chiefly on the *Encyclopedia*.

EARLY UNPUBLISHED WRITINGS

"LIFE OF JESUS." Even before he wrote his "Life of Jesus," Hegel had written some comments on Christianity in which he criticized it for its belief in the efficacy of prayer and had contrasted it, to its detriment, with the this-worldly, social religion of the Greeks. Jesus, he held, was obscurantist and narrow-minded in comparison with Socrates. In the "Life of Jesus" it almost seems as if Hegel had decided to rewrite the Gospels in the form of a Kantian manifesto. He began by claiming that God is pure reason. He described Jesus as the son of Joseph and Mary. The only miracles Hegel mentioned he interpreted naturalistically, bringing the work to an end with the death and burial of Jesus. The central theme is the conflict between the virtuous Jesus acting dutifully for the sake of the moral law and the Jewish priesthood calling for the meticulous observance of a set of irrational rules said to be commanded by God. Jesus is depicted as saying to the Pharisees, "When you regard your ecclesiastical statutes and positive commands as the supreme law given to mankind, you fail to understand the dignity of man and the power he has of creating out of himself the idea of the divinity and knowledge of his will." This improbable allocution is typical of the way in which this work denudes the Gospel narrative of what is individual and poetical.

***THE POSITIVITY OF THE CHRISTIAN RELIGION*.** The theme of *The Positivity of the Christian Religion*—the place in the Christian religion of the rational, on the one hand, and of the merely factual and historical, on the other—was already raised in the "Life of Jesus." Developing the implications of the then current distinctions between natural law and positive law and between natural religion and positive religion, Hegel argued that the positive element rested on authority and was not wholly based on the dignity of man. In Christianity, according to Hegel, the main positive element was provided by Judaism, a highly authoritarian religion. But Jesus himself brought elements of positivity into the rational morality that it was his prime aim to teach; he could not have obtained a hearing from the Jews of his day if he had not claimed God's authority for his teachings. "Jesus therefore demands attention for his teachings, not because they are adapted to the moral needs of our spirit, but because they are God's will" (*Early Theological Writings*, p. 76). In claiming to be the Messiah, Jesus was using the language his listeners would understand. His followers, from a natural interest in the details of his life, developed these positive elements into Christianity. They appealed to miracles as proofs of Jesus' divinity and virtue, and instead of revering him for his teaching about virtue, they revered his teaching about virtue because of the miracles he was supposed to have performed.

Hegel asked how it happened that the pagan religion of the Greeks and Romans was overcome by Christianity. His answer was that at the periods of their greatness the Greeks and Romans were free peoples each individual of which regarded his own good as inseparable from the good of his community. When they lost their freedom, they lost the motives that bound them to their fellows; government and authority were now imposed from without, weighing down upon isolated individuals who came to regard their lives as individual possessions to be preserved irrespective of the social whole that alone gave them meaning.

> Thus the despotism of the Roman emperors had chased the human spirit from the earth and spread a misery which compelled men to seek and expect happiness in heaven; robbed of freedom, their spirit, their eternal and absolute element, was forced to take flight to the deity. [The doctrine of] God's objectivity is a counterpart to the corruption and slavery of man. (ibid., pp. 162–163)

***THE SPIRIT OF CHRISTIANITY*.** In *The Spirit of Christianity* Hegel continued and sharpened his attack on Judaism, which he regarded as a religion of domination. He now criticized Kantian ethics as well, however, finding in it elements of the same positivity he had criticized in the Jewish religion and had seen as a contamination in the teachings of Jesus. Kant had contrasted his rational religion with the religion of the Siberian shamans on the ground that these primitive men, as well as some civilized prelates and puritans, irrationally worshiped alien forces that they regarded as exerting domination over men. But according to Hegel, the difference between the believers in these positive creeds and the follower of the religion approved by Kant is "not that the former make themselves slaves, while the latter is free, but that the former have their lord outside themselves, while the latter carries his lord in himself, yet at the same time is his own slave" (ibid., p. 211). Hegel here first used the word *morality* (*Moralität*) as a pejorative description of the Kantian morality, which he now considered to be a submission of man's inclinations, including his impulses and his feelings of love, to a universal reason held to be free from and above all passion. He held that virtue demands more than

this and that in the Sermon on the Mount Jesus made higher demands. "The Sermon does not teach reverence for the laws; on the contrary, it exhibits that which fulfils the law but annuls it as law and so is something higher than obedience to law and makes law superfluous" (ibid., p. 212). Thus, duty takes a lower place than love. "Jesus makes a general demand on his hearers to surrender their rights, to lift themselves above the whole sphere of justice or injustice by love, for in love there vanish not only rights but also the feeling of inequality and the hatred of enemies which this feeling's imperative demand for equality implies" (ibid., p. 218). Hegel here saw in the ethics of the Sermon on the Mount and in the conduct of Jesus something of the "beautiful soul" described by Johann Wolfgang von Goethe in *Wilhelm Meister*. Jesus retained his dignity by refusing to defend himself or to uphold his rights.

Hegel went on to discuss with subtlety the possible consequences for the individual and for other men of resistance to evil, on the one hand, and of withdrawal from conflict, on the other. In this part of the work the beginnings of dialectical method as it was used a few years later in *Phenomenology of Mind* may already be discerned.

PHENOMENOLOGY OF MIND

The *Phenomenology* is the most obscure and the most interesting of Hegel's works. On the title page it is described as a "System of Science, Part I. The Phenomenology of Mind," but this arrangement of Hegel's system was not continued in the *Encyclopedia*, where the section headed "Phenomenology of Mind" is contained in the third part and deals with only some of the topics of the original *Phenomenology*. Hegel put the *Phenomenology* together rather hastily and was uncertain what to call it. Different copies of the first edition have slightly differing titles, and what seems like a new title, "Science of the Experience of Consciousness," is placed after the preface and before the introduction. Insofar as there is a central theme, it consists of an account of the various stages of human consciousness from mere sense awareness to absolute knowledge, but there are many digressions into topics of current interest, such as Goethe's description of the "beautiful soul," the Reign of Terror, and F. J. Gall's phrenology. The difference between the dialectical progression of the *Phenomenology* and of the *Encyclopedia* was cited soon after Hegel's death as evidence of the inadequacy of the dialectical method (C. F. Bachmann, *Über Hegels System und die Nothwendigkeit einer nochmaligen Umgestaltung der Philosophie*, Leipzig, 1833). In the twentieth century Marxists preferred the *Phenomenology* to Hegel's other writings because Karl Marx himself admired it and because of its account of how man develops by transforming the natural world through his labor. Existentialists have preferred it to the later system because of its account of man as maker of himself; no doubt they are also impressed by Hegel's references to death and the fear of death.

The *Phenomenology* begins with a dialectical discussion of sense perception in which it is argued that knowledge of physical things presupposes the view that the physical world consists of forces interacting according to laws. Hegel maintained that knowledge of such a world is really a type of self-knowledge, since in penetrating to the forces behind phenomena we become aware of what we ourselves have devised and put there. "Behind the so-called curtain which is to hide the internal constitution of things, there is nothing to be seen unless *we ourselves* go behind." The physical world of scientific theory presupposes self-conscious beings. When he analyzed self-consciousness, Hegel argued that it presupposed a plurality of living and desiring beings each of whom seeks to subdue the world to his own wishes, to make it part of himself.

MASTER AND SLAVE. No individual will rest satisfied with a conquest that fails to secure the conscious acknowledgment of other men. Hence, there is a struggle for both power and recognition. In this struggle some will take greater risks than their competitors; those who risk the least will become the slaves or bondsmen of those who face death by risking their lives. In order to preserve his life, the slave submits to the master, who regards the slave as nothing but a means to his own designs. The slave is forced to work, whereas the master can enjoy leisure in the knowledge that the slave is reshaping the natural world to provide the products of his labor for the master to consume. Thus, the master's leisure protects him from experience of the negativity of nature, whereas the slave, in struggling with nature's recalcitrance, learns its secrets and puts mind into it. The master, in consuming, destroys; the slave, in working, creates. But the master's consumption depends upon the slave's work and is thus impermanent, whereas the slave's labor passes into things that have a permanent existence. Hegel argued, too, that the slave's work in transforming the natural world is a consequence of his fear of the master, who can kill him. Death is overcome by the works of civilization. The man who risks his life and becomes the first master breaks the bonds of nature and starts the process that will incorporate mind into it.

It is not surprising that this section in the *Phenomenology* has greatly interested Marxists. Both Georg Lukács, in *Der junge Hegel*, and Herbert Marcuse, in *Reason and Revolution* (2nd ed., London and New York, 1955), contrived to discuss it without mentioning Hegel's emphasis on the fear of death. In *Introduction à la lecture de Hegel*, Alexandre Kojève brought out the importance of the fear of death and showed, too, that Hegel was here concerned with the transition from nature to history, from mere life to thought, from animality to freedom.

THE UNHAPPY CONSCIOUSNESS. The next dialectical transition is from mind that is attempting to master nature to mind that seeks freedom and independence in itself, that says, "It is in thinking that I am free because I am not in another but remain completely with myself alone," an attitude exemplified in stoicism. But stoicism passes over into skepticism, for the stoic finds freedom in himself as a rational, thinking being, whereas the skeptic, pushing freedom still further, uses thought to dissipate its own categories. This, according to Hegel, was the state of mind that prevailed when the Roman Empire was dissolving. Christianity was an attempt on the part of men in intellectual despair to find stability in an eternal and infinite God.

Hegel called this frame of mind the unhappy consciousness. The individual is divided within himself, conscious of his own isolation, attributing all that is good to the activity of God. What Hegel said here was elaborated from a passage in *The Positivity of the Christian Religion* describing how the eternal and absolute in man had been "forced to take flight to the Deity." The unhappy consciousness was regarded by Hegel as a characteristic of both Judaism and Christianity and as the condition of all men at all times who believe in a transcendent God before whom they are as nothing. It is a stage on the way to higher forms of self-mastery.

It will be noted that in this part of the *Phenomenology* Hegel passed from epistemology through a sort of speculative sociology to an account of historical stages in human consciousness. According to Rosenkranz, Hegel, in his last years, used to refer to the *Phenomenology* as his philosophical "voyage of discovery," and it does seem that the course of the argument, although arresting, was not altogether foreseen. Josiah Royce was right when he said that in this book Hegel described "in serial order, some varieties of experience which … are at once characteristic of the general evolution of the higher intellectual life, and are examples of the transition from common sense naiveté to philosophical reflection and to the threshold of an idealistic system" (*Lectures on Modern Idealism*, edited by J. Loewenberg, New Haven, 1919, p. 139).

REASON AS "OBJECTIVITY." After discussing certain scientific theories of his time under the heading "Reason as Observer," Hegel went on to consider some of the ways in which reason becomes practical. He depicted the man who, like Faust, tries to make the passing moment stay. When this attempt fails, as inevitably it must, ideals are sought in a spirit of sentimental disillusionment, but such romantic crusades are never really serious. In reaction to this frivolity there develops a taste for the hard intellectual pursuits of disinterested scholarship, the concern for "objectivity," for facts, for "the thing itself." But these allegedly disinterested researchers actually go into a sort of intellectual jungle (*das geistige Thierreich*) where, deceiving one another and themselves, they tear one another to pieces in the service of truth. It soon emerges that it is not the facts that matter but a certain proprietorship that scholars working in their special fields claim over the facts.

THE DIALECTIC OF MORALITY. In the next part of the *Phenomenology*, titled "Mind," Hegel considered how the mind of man is embodied in his rules and institutions. This part constitutes both an account of the main types of moral attitude and a philosophy of history. These two lines of thought come together insofar as Hegel regarded the historical development from the Greek and Roman civilizations through early and medieval Christianity to Protestantism and the French Revolution as an unfolding of the main aspects and stages of freedom and, hence, as a dialectical actualization of what was merely latent and implicit in the morality of the ancient world. This unfolding is dialectical because it proceeds by oscillations and because it is made possible by conflict, in the ancient world by the conflict between the gods of the family and the laws of the city and in the modern world by the conflict between the claims of the individual and the demands of society.

In this part Hegel gave indications of the doctrine of alienation that attracted Marx in the 1840s. In building his civilization, man creates institutions and rules that are simultaneously his own products and alien constraints upon him. He may not even understand them, so that they appear strange to him. It was Hegel's view, of course, that without these institutions and rules and without the restrictions upon willfulness that they impose, mind could not reach its higher levels.

RELIGION AND ABSOLUTE KNOWLEDGE. In the last two parts of the *Phenomenology*, Hegel presented the dialectic of religion and the passage to absolute knowledge. In the earlier developments of mind the individual has to find his place in the natural world and in society, but in religion he gains consciousness of the Absolute Being. This is first approached in the primitive religions of nature, in which men worship trees, streams, or animals. Next come those forms of religion in which the Absolute Being is approached through such works of art as temples and statues. This type of religion reached a high level in ancient Greece, but when God was represented in human form, he came to be regarded as merely human and hence was lost sight of in the tragic heroes of Greek drama. As the religious element was discarded from tragedy, it gave way to comedy, in which the contingencies of human life were paraded and criticized, and God was completely ignored in favor of human self-knowledge. "The individual self is the negative force through and in which the gods … disappear."

This skeptical and sophisticated humanism is succeeded by Christianity, in which God is revealed to man in Christ. Here the human and the divine are no longer sundered, and God is seen to be present in the world. But it is easy to overemphasize the historical features of Christianity and, as Hegel put it, to neglect the spiritual revelation in the attempt to uncover the often commonplace ideas of the early Christians and to gain knowledge of the mere externality and particularity of Jesus. Thus, no religious experience, not even that of Christianity, can bring absolute knowledge. The historical element in Christianity, although necessary in order to avoid regarding the Absolute Being as apart from the world, is nevertheless inseparable from perception and imagination. The events of the Gospels are, so to speak, pictured or represented. Religion therefore leads on but is subordinate to the supreme form of knowledge, the philosophical, in which human history is "conceived history, the recollection of the Absolute Mind and its graveyard, the actuality, truth and certainty of its throne, without which it would be for ever alone and devoid of life." In these last words of the *Phenomenology*, Hegel made it clear that the course of history, philosophically conceived, was in his view the incarnation of the Absolute Mind. Apart from the history of man God would be alone and lifeless (*das leblose Einsame*). It would seem, indeed, that without the historical development of man and his freedom there would be no God.

THE HEGELIAN SYSTEM

It has already been mentioned that before writing the *Phenomenology*, Hegel had written but had left unpublished some attempts at a complete system of philosophy and that the *Phenomenology* was described on its title page as the first part of a system of science. It turned out that the *Science of Logic* (1812–1816) became the first part of Hegel's final system. A shortened and revised version of the *Science of Logic* appeared in 1817 as the first part of the *Encyclopedia*, a book intended for use at his lectures. A second, very much elaborated edition of the *Encyclopedia* appeared in 1827, and a third in 1830. This last edition was reprinted in the edition of Hegel's collected works published soon after his death, with inserted "additions" taken from the notebooks of students who had attended Hegel's lectures. These additions, which are most frequent in the first and second parts of the *Encyclopedia*, help greatly in the understanding of Hegel's argument but do not have quite the authority of the main text. Such additions are less frequent at the end, since the editors considered that the *Philosophy of Right*, first published in 1821, and some of the sets of lectures, provide commentary of this sort.

THE *ENCYCLOPEDIA*. The *Encyclopedia* starts with a discussion of "Logic"—a revision of *Science of Logic*—and proceeds to the sections "Philosophy of Nature" and "Philosophy of Mind." The transition from the "Logic" to the "Philosophy of Nature" is not easy to understand. There are statements that say that the idea decides to allow nature to go forth freely from itself (Sec. 244), that "Nature has come to pass as the Idea in the form of otherness" (Sec. 247), and that nature is "the unresolved contradiction" (Sec. 248). The last main heading in the "Philosophy of Nature" is "The Animal Organism." Toward the end of this section there is an account of the individual animal as having "an original sickness" and "an innate germ of death" (Sec. 375), which leads to the assertion that with the subjectivity of living organisms the "outside-itself-ness" (*Aussersichsein*) of nature is transcended by the "interiority" (*Insichsein*) of actuality (Sec. 376).

Hegel later claimed (Sec. 381) that mind presupposes nature but is "the truth [of nature] and its absolute ground [*deren absolut Erstes*]." He also stated that the essence of mind is freedom (Sec. 382). A fundamental comment on the dominating triadic division must be made before going further into the details of the system. The revised "Science of Logic" that appeared in the *Encyclopedia* was concerned with the categories of thought,

proceeding from the most inadequate and abstract to the most concrete and adequate, from being to the Absolute Idea. The inadequacies of the abstract categories show themselves through the contradictions they give rise to. Being is more abstract than becoming; becoming, more abstract than being-for-self; these early categories, more abstract than the latter categories of life, and so on.

But Hegel was always concerned with the categories of thought and their relations to one another. When he wrote that the idea decided to allow nature to go forth freely from itself, was he saying that thought is the Divine Being that created nature? The religious overtones that accompany Hegel's major transitions cannot be ignored, but those who wish to interpret him naturalistically—an interpretation his early writings and the *Phenomenology* may well justify—can take the view that the decision and the free going forth are meant to indicate that nature is not deducible from the categories of thought, that there is a contingency about it that no system of logic and no elaboration of concepts can eliminate. In *Subjekt-Objekt* (Berlin, 1951) Ernst Bloch suggested that the free decision of the Absolute Idea is reminiscent of the arbitrary act of an absolute monarch, and he quoted a passage from Schelling's *Philosophie und Religion* (Tübingen, 1804) which held that "the descent of finite things from the Absolute" is a "primal accident [*Urzufall*]."

In the third part of the *Encyclopedia*, Hegel described mind as it develops in the natural world, mind as it transforms the natural world in creating the works of civilization, and mind fully aware of itself in the complete self-consciousness of philosophical thought. The "Logic" culminates in the Absolute Idea, the most adequate category but still a category. In the "Philosophy of Nature," where there is no Absolute, the culminating point consists of mortal individuals belonging to persisting animal species. The "Philosophy of Mind" culminates in Absolute Mind, the consciousness man gains of himself through understanding his own history in a civilization that he has imposed upon the contingencies of nature.

"**LOGIC**." Like the Hegelian system as a whole, each of its three main sections—"Logic," "Philosophy of Nature," and "Philosophy of Mind"—is again divided into three. The "Logic" is divided into the "Doctrine of Being," the "Doctrine of Essence," and the "Doctrine of the Concept [*Begriff*]." The difficulties in presenting a comprehensible summary of Hegel's views are at their greatest in relation to the "Logic," and all that will be attempted is an indication of a few of Hegel's most characteristic views.

"Doctrine of Being." In the "Doctrine of Being" Hegel was concerned with the most abstract categories. Being itself, the most abstract of all, amounts to the same as nothing. Like Bertrand Russell in his theory of descriptions, Hegel held that nothing can be said to be unless some characteristic is attributed to it; hence, in Hegel's terminology being leads on to determinate being, which involves the notion of quality. On the ground that a quality is something distinct from other qualities, Hegel argued that quality implies the category of a unit (*das Eins*) and that this in turn leads on to quantity. This part of the "Logic" was completed by transitions to degree and measure.

Hegel's object in the "Doctrine of Being" was to show that these categories are not independent of one another but develop from one to the other in an ascending order of adequacy. We know more about something when we know the proportions of its parts than when we know only how many parts it has, that it is, or that it is something or other. An important element in this part of the "Logic" is Hegel's criticism of infinite numerical series as the false infinite and his contrast between the false and the true infinite, which is not an incompletable progression of similar items but a completed, complex whole of supplementary parts. The true infinite is not to be reached by attempting the impossible task of moving from one finite to the next but must comprise the finite.

"Doctrine of Essence." The "Doctrine of Essence" is concerned with such distinctions as that between a thing's nature and its appearances, forces and their manifestations, form and matter. Hegel exploited the difficulties ("contradictions") that arise when these oppositions are so accentuated that we are left with featureless essences, on the one hand, and unattached appearances, on the other. Typical of his treatment of these topics is his claim that "the explanation of an appearance in terms of a force is an empty tautology" (Sec. 136) and his assertion that as a man's outward actions are, so his inner aims and intentions must be (Sec. 140).

"Doctrine of the Concept." A prominent feature in the "Doctrine of the Concept" is Hegel's critical treatment and reorganization of the traditional formal logic. Thus, he classified judgments in terms of his own division of "Logic" into being, essence, and concept. The classification progresses from the mere factual attribution of a quality, through disjunctive and necessary judgments in which the predicate belongs essentially to the subject, to judgments of value that assert that a thing is good or bad just because it is that individual thing. Judgments gain in adequacy as they advance from mere factual attribution

to attribution for reasons contained in the subject. Hence, the more developed forms of judgment are indistinguishable from inferences. In his account of the syllogism Hegel placed inferences in which the terms are only contingently connected at the bottom of a scale leading up to the disjunctive syllogism, in which a genus is exhaustively specified.

Although Hegel retained the terms and distinctions of the traditional formal logic, the use he made of them was highly original. Instead of setting out the types of judgment and the figures and moods of the syllogism as equally valid forms, he regarded judgment as implicit inference and inference as ordered in a scale of ascending rationality. This conception of logic influenced such later writers as Christoff Sigwart and R. H. Lotze and was developed in both F. H. Bradley's *Principles of Logic* (London, 1883) and Bernard Bosanquet's *Logic: The Morphology of Knowledge* (2 vols., Oxford, 1888).

The argument of Hegel's "Logic" can be very briefly summarized. The least that can be said about anything is that it is. More is said about it when it is qualified, numbered, or measured; still more is said about it when it is explained in terms of essences, grounds, or causes. Most is said about it when it is placed in the context of life, purpose, will, and value.

"**PHILOSOPHY OF NATURE**." At the end of the eighteenth and the beginning of the nineteenth century there was a great deal of philosophizing about nature. Electricity was held to have cosmic significance, and Schelling made much of the opposition between positive and negative poles. Poets as dissimilar as William Blake and Goethe rejected what they regarded as the unduly quantitative physics of Isaac Newton. Spinoza was revived, and among German poets and philosophers much was said about the *έν και πᾶν*, the one and the all. It is not surprising, therefore, that Hegel's dissertation of 1801, *De Orbitis Planetarum*, was critical of Newton and sought to provide an a priori justification of Johannes Kepler's laws. At the end of the dissertation Hegel mentioned some numerological accounts of the distances and number of the planets and expressed the opinion that if Plato was right in the *Timaeus*, there could be no planet between Mars and Jupiter. Hegel did not then know that Ceres, an asteroid between these two planets, had been discovered at the beginning of the year. However, even after he had heard of this discovery and of the discovery of several other asteroids soon after, he continued to hope that philosophical reasons could be given for the positions of the heavenly bodies. In an addition to Section 270 of the *Encyclopedia*, Hegel tried to show that these asteroids filled a gap that would otherwise have been unreasonable. The addition ends with the words: "Specialists do not think about such matters. But a time will come when in this science there will be a demand for concepts of the Reason."

It should be mentioned here that Hegel accepted and developed Kant's distinction between the reason and the understanding. According to Hegel, the understanding, although a necessary stage of thought, is less philosophical than the reason. To think in terms of the understanding, as is done in mathematics, the natural sciences, and traditional metaphysics, is to think in terms of fixed and uncriticized categories, to think undialectically or in prephilosophical terms. The reason moves dialectically toward completeness in terms of fluid categories that constantly amend themselves. Thus, when Hegel wanted astronomers to pay attention to "concepts of Reason," he wanted astronomy to take its place within a system of philosophy. This place must be a subordinate one, for Hegel wrote in the Introduction to the "Philosophy of Nature" (Sec. 248): "Even if arbitrary will, the contingency of mind, leads on to wickedness, this is nevertheless something infinitely higher than the regular movements of the planets or than the innocence of the plants: for what goes wrong in that way is nevertheless mind." Here Hegel was emphasizing the gulf between mind and nature, even though he held that the understanding does not give a complete knowledge of nature.

Mechanics. The three main divisions of the "Philosophy of Nature" are concerned with mechanics, physics, and organic nature. The astronomical theories expounded in the first part have already been touched upon. This part also contains a brief discussion of space and time. Following Kant, Hegel regarded them both as "forms of sensibility," or, more strikingly, as "the non-sensible sensible." Although he regarded arithmetic and geometry as sciences of the understanding, he considered the possibility of a philosophical mathematics at the level of measure or proportion (mass).

Physics. The second part of the "Philosophy of Nature" moves through various triads from light, the elements, sound, heat, to electricity and chemical combination. Hegel commented upon the philosophical significance of each form of matter. The comment on heat is characteristic:

> Heat is the re-establishment of matter in its formlessness, its fluidity, the triumph of its abstract homogeneity over its specific determinations.… Formally, that is in relation to spatial

> determinations in general, heat therefore appears expansive, as cancelling the limitations which the specification of the indifferent occupation of space is. (Sec. 303)

That is, when heat spreads out from a heated thing, that thing is not confined to one place, as it would be if it were not heated. Or as Hegel put it in the next section, heat is the "real negation of what is specific and exclusive in body."

Organic nature. In the last main triad of the "Philosophy of Nature," Hegel passed from geological nature through vegetable nature to the animal organism. The most interesting part of this triad is the last, in which Hegel discussed animal species and their relationships. He seems to have thought that violent death is, in the animal world, "the natural fate of the individual" and that because of the contingency of nature animal life is "uncertain, anxious, and unhappy" (Sec. 369). But other members of the same species are not only hostile to the individual; they are also, like him, continuations of the species, and, hence, the individual feels a need to unite himself to the species (*Gattung*) and to continue it by copulation (*Begattung*)—the play on words is, of course, deliberate. Thus, Hegel seems to have held that animal sexual union is not merely a contingent affair. On the other hand, since the new individuals produced in this way only repeat the features of their parents and other ancestors, their constant reproduction is an instance of the false infinite, not of the true infinite in which completeness and perfection are achieved.

"**PHILOSOPHY OF MIND**." The major triad in the "Philosophy of Mind" consists of "Subjective Mind," "Objective Mind," and "Absolute Mind."

"Subjective Mind." Under the heading "Subjective Mind" and the subheading "Anthropology," Hegel dealt with the soul as a natural entity in the physical world; the soul as a sensitive, feeling being; and the soul as a being that can express itself and act upon the world through its body. The upright body, the hand "as the absolute tool" (Sec. 411), the mouth, and the power of weeping and laughing all enable man to express in nature—to externalize—his thoughts and feelings. Furthermore, the world has effects upon man's body that are internalized by him—Hegel here made a play on the word *Erinnerung*, which means "recollection" but, if taken in the literal sense of its German etymology, can be taken to mean "internalization." When the organism reacts to immediate stimuli in the light of its own experience, mind has evolved beyond the mere animal level and has reached the stage of consciousness.

Hegel discussed the next moment of subjective mind under the heading of the "Phenomenology of Mind," going through the main phases distinguished in the earlier chapters of his book with that title—namely, sense experience, perception, understanding, desire, the self-consciousness that recognizes others (containing the discussion of master and slave), reason.

The third triad of subjective mind, which is headed "Psychology," contains descriptions of such intellectual functions as intention, representation, recollection, imagination, memory, and thought and descriptions of the practical drives, impulses, and seekings after satisfaction.

This part ends with a brief section headed "Free Mind." Here it is asserted that the unity of theoretical and practical mind is free will. Hegel meant that human freedom is possible only on the dual basis of thought and impulse and consists of the rationalizing and systematizing of the impulses and passions. "This will to freedom," he said, "is no longer an impulse that demands satisfaction, but the character—the mind's consciousness grown into something non-impulsive" (Sec. 482).

"Objective Mind." At the very end of his discussion of subjective mind Hegel wrote that the freedom which is the culmination of subjective mind is only a concept, "a principle of mind and heart destined to develop into the objective phase, into legal, moral, religious and scientific actuality" (Sec. 482). The rest of the system is therefore concerned with the ways in which the human will, in which thought and impulse ("mind and heart") are combined in freedom, becomes effective (this is the idea behind the word *actuality*, which translates *Wirklichkeit*) in the public world, the world in which men act and in which their thoughts and deeds give rise to rules, institutions, and organizations. These rules, institutions, and organizations are independent of each man and thus may be regarded as kinds of objects, though not as physical objects. Men build up in the natural world a world other than the natural world by working on nature and transforming it and by creating systems of property, economic organizations, class differentiations, and the like. The triad that makes up objective mind comprises law (*Recht*), subjective morality (as Wallace translated *Moralität*), and social morality (as Wallace translated *Sittlichkeit*; T. M. Knox translated it as "ethical life"). The first part covers legal rights and duties as exemplified in property, contract, and punishment. The second is concerned largely with the morality of intention and conscience—the term *Moralität* was used by Hegel somewhat

pejoratively to mean a sort of ethics (of which Kant was, in his view, the chief exponent) in which the agent is unduly governed by the subjective and internal aspects of decision and action.

The third part is itself a triad. The first stage of social morality is the family, "the natural or immediate phase" of objective mind (*Philosophy of Right*, Sec. 152). When members of the family have matured, they detach themselves from it and enter the world of independent men who compete in an economic arena free from tribal allegiances. This phase of social life Hegel called "civil society." It is the world of intelligent, responsible individuals in their business relationships, free from irrational tribal loyalties, allowing their connections with one another to be formed by the coincidence of wants in a market of wide extent. Indeed, it is the aspect of human society that the classical economists, whom Hegel admired, had analyzed and justified. But civil society cannot exist as a mere market, for markets need to be policed, whereas trades and industries themselves find common concerns that unite the individuals in corporations of various kinds.

There is thus a double necessity for the state—as the upholder of fair dealing and as the ultimate curb on the selfishness of corporations within civil society. In the *Encyclopedia*, Hegel wrote of "the unification of the family principle with that of Civil Society" and described it as a unification of the love that is essential to the family with the conscious universality that is the mark of civil society (Sec. 535). In the *Philosophy of Right* (Sec. 257) the state was described as "the actuality [*Wirklichkeit*] of the ethical Idea"—that is, as its effective embodiment. In the same section of the *Philosophy of Right*, Hegel wrote that "the mind of a nation (Athene for instance) is the divine, knowing and willing itself," and in an addition to Section 258 is the famous phrase "The march of God in the world, that is what the State is." But this section has been misunderstood. In the sentence before that in which he had written that the state is divine, Hegel had said, with the family in mind, "The *Penates* are inward gods, gods of the underworld," so that it is not only to the state that he attributed divinity. Furthermore, in the same addition as that in which he claimed that the state is "the march of God in the world," he said that the state "stands on earth and so in the sphere of caprice, chance and error, and bad behaviour may disfigure it in many respects." Hegel's main concern was, as he stated, to analyze the state at its best. Although, like Aristotle, he regarded the state as the highest social achievement of man, he also held, again like Aristotle, that within the state there should be guarantees against arbitrariness and despotism. He did not take a favorable view of "popular suffrage" on the grounds that "in large states it leads inevitably to electoral indifference" and that "election falls into the power of a few, of a caucus" (*Philosophy of Right*, Sec. 311). He strongly believed that all important interests should be represented and thought that there should be a constitutional monarchy with considerable powers advised by an upper and a lower house.

This brings us to the most controversial part of Hegel's account of objective mind, his philosophy of history. Whatever else is involved in his view that the state is man's highest social achievement, it undoubtedly implies that there is no superior body or group by which its claims may be assessed. States are necessarily independent beings. Their relations are regulated to some degree by custom, and there is an international law that regulates dealings between subjects of different states and requires adherence to treaties, as if they were a sort of contract. When the vital interests of states clash, however, there is no alternative except war. War between states, Hegel had said in his "Die Verfassung Deutschlands" ("Constitution of Germany," 1802; first published in *Schriften zur Politik und Rechtsphilosophie*, edited by Georg Lasson, 2nd ed., Leipzig, 1923), does not decide which of the rights of the conflicting states is the true right—for both are—"but which right has to give way to the other." Hegel believed that war performs the function of keeping before the minds of men the realities of death and destruction. He held that states are individuals and that all individuals persist in their existence by ensuring that other individuals recognize them as they recognize the others. The very concept of a state therefore requires that there be a plurality of them, and this makes war a part of the system of states even though war is not their natural condition but an interruption of the normal state of peace. Hegel argued that since war is a relation between states and not a relation of individual men to one another, the rights and interests of noncombatants should be maintained to the utmost. For the same reason he was in favor of professional armies and against conscription or any form of levy en masse.

Each nation is limited by geographical and other accidental features and hence can build up only a particular culture and can have only a particular, not a universal, history. Thus, nations, when they reach the level of statehood, make their contribution to the whole in the part they play in world history (*Encyclopedia*, Sec. 548). World history is not wholly an affair of chance or contingency; as the work of mind it could not be. Therefore, the history of the world has a rational structure, and any his-

torical writing that ignored this "would be only an imbecile mental divagation, not as good as a fairy-tale" (Sec. 549). This rational structure, according to Hegel, is the development of freedom.

"Absolute Mind." The triad that completes the Hegelian system is composed of art, revealed religion, and philosophy. It will be remembered that at the end of the *Phenomenology* Hegel proceeded from the religion of nature to the religion of art and then to the philosophical knowledge of the history of the world. In the *Encyclopedia* art is given what seems to be a more independent status, but the details of the argument hardly bear out the general scheme, since the transitional sections describe a transition from objective mind to religion, as in the *Phenomenology*. Thus, in the concluding sections of the *Encyclopedia* art is regarded as an inadequate form of religion, religion as a more adequate form of art, philosophy as religion freed from picture thinking and wholly rationalized, and all three as manifestations of Absolute Mind. Art is the embodiment of Absolute Mind in material things fashioned by the artist, who, in a sense, is thus "the master of the God" (*Encyclopedia*, Sec. 560). In classical art the embodiment takes place without any antithesis between the embodiment and the mind that is embodied. In the art of the sublime, which preceded classical art, the Absolute Mind is regarded as something that defies embodiment and remains forever beyond and behind the sensible forms that succeed only in symbolizing it. The defect of artistic representation is that the sensible symbols may be taken to refer to another world beyond, which is as limited as this world is falsely taken to be. Thus, men worship idols or even bones, "which point to the unspiritual objectivity of that other world" (ibid., Sec. 562).

God is therefore not something grander and more powerful than the natural world yet fundamentally like it, nor is he something beyond the world that must remain forever inaccessible to man. God is manifested in the world, and this is the truth that revealed religion has expressed most adequately in the Christian doctrine of the Incarnation. Without this doctrine God would still be regarded as beyond the world and, thus, as incomplete and finite. Even with this doctrine he is conceived of through the medium of particular historical events that introduce an element of contingency and irrelevance into our conception of him. In philosophy the artist's external vision and the mystic's internal vision are united in a mode of thought in which there is no further conflict. The philosopher who achieves ultimate self-knowledge is freed from the conflicts that inevitably disturb the inferior levels of knowledge. By philosophizing to the end, he has made himself free (ibid., Sec. 576).

THE DIALECTICAL METHOD

CONTRADICTION. It is now necessary to give more detailed attention to Hegel's dialectical method. There are interpreters of Hegel who say that Hegel denied the principle of contradiction in that he held that contradictories can both exist and that contradictory propositions can therefore both be true. Others deny this interpretation, maintaining, instead that, according to Hegel, since contradiction is a mark of inadequacy and falsehood, contradictions are to be found in the lower categories but are absent from or resolved in the Absolute Idea. This view is summed up in Michael Oakeshott's reference to "the element of self-contradiction inherent in all abstraction" (*Experience and Its Modes*, Cambridge, 1933, p. 328). Those who take the first view can quote some convincing passages from Hegel's *Science of Logic*. For example, there he wrote that "all things are in themselves contradictory," that "movement is existing contradiction itself," and that "only insofar as something has contradiction in itself does it move, have impulse or activity."

If Hegel had rejected the principle of contradiction in the sense that that principle is understood by formal logicians, his case would indeed be serious, for it follows from the rejection of this principle that any proposition can be true and false and that there is thus no means of distinguishing truth from falsehood. It is important, therefore, to see whether Hegel did reject the principle of contradiction in this sense and whether its rejection is part of his dialectical method. That these questions are not easy to answer becomes apparent if we consult some of the commentators on the passages I have just quoted. J. M. E. McTaggart, in his *Commentary on Hegel's Logic*, was dissatisfied with the whole section and claimed that in it Hegel had allowed himself to be too much influenced by Schelling's view on polarity and opposition. "The whole point of the dialectic," McTaggart protested, "is that the perception of a contradiction is a reason for abandoning the category which we find contradictory." Indeed, he found this part of the *Logic* so unsatisfactory that he proposed to amend the sequence of categories by leaving out contradiction altogether.

McTaggart said nothing, however, about Hegel's statement that there are existing contradictions. G. R. G. Mure, in his *A Study of Hegel's Logic*, did not evade this difficulty. Examining Hegel's text more closely than McTaggart had done, he pointed out that on the ground that "the contradictory cannot be imagined or thought"

Hegel rejected the commonsense view that things cannot be self-contradictory but that thought can be. Mure called attention, too, to Hegel's statement that self-contradiction is not a mere disease of thought but something it must pass through on its way to truth. Furthermore, according to Hegel, it is finite things that are self-contradictory, and they are contradictory not in relation to one another but by virtue of their relation to what is infinite: Hegel "is not suggesting that Big Ben can now read both 9 p.m. and not 9 p.m." (p. 105). Although this is an improvement on McTaggart, it left out of account Hegel's statement that for something to move, it must be both here and not here at the same time. What Hegel said about movement is not altogether unlike Mure's example of Big Ben. So the difficulty remains.

In the "Logic" sections of the *Encyclopedia,* which was written later than the *Science of Logic,* contradiction is not a separate category at all. Perhaps the reason for this difference is that Hegel had second thoughts and gave up the idea of contradiction in the nature of things. But although contradiction is no longer a category in the *Encyclopedia*, Hegel still sometimes wrote as if there were contradictions in the nature of things. For example, he stated that although such concepts as "square circle," "many-sided circle," and "straight curve" are self-contradictory, geometers nevertheless regard circles as polygons composed of very short sides and "the center and circumference of a circle as opposite and contradictory to one another" (*Encyclopedia*, Sec. 119). Hegel also suggested that polarity in physics goes against the ordinary logic—but he used the word *opposition* (*Entgegensetzung*) rather than *contradiction* (*Widerspruch*).

In *Geschichte der neueren Philosophie* (Heidelberg, 1901, Vol. VIII, Part 2) Kuno Fischer tried to overcome the difficulty by distinguishing between two sorts of contradiction, "necessary contradiction" and "impossible contradiction." The example of a square circle illustrates the notion of an impossible contradiction, a *contradictio in adjecto*, for it is impossible for the same thing to be both circular and square. When a circle is regarded as a many-sided polygon, however, the contradiction is not *in adjecto* but in *subjecto*, for the circle is then being regarded as in the process of being formed or generated from these many sides. This, Fischer held, is the contradiction involved in all becoming (the first concrete category of the "Logic," the synthesis of being and nothing). Fischer's suggestion is therefore that there is not a vicious or stultifying contradiction involved in becoming or in movement, contradictory though they must in some sense be. But although this may be a correct exposition of Hegel's view, it is hardly a defense of it, since it merely repeats without explaining his claim that there are contradictions in the objective world.

By drawing this distinction, Fischer has nevertheless raised the question whether Hegel intended the word *contradiction* to be used in the way it is used in formal logic. The answer is clear enough. Hegel did not regard formal logic as a philosophical science, and he therefore rejected any idea that its categories should dominate philosophical thought. Thus, the fact that the word *contradiction* is used in a certain way by formal logicians was not for him a reason for confining himself to that meaning. When Hegel was advocating the dialectical method, he had in mind a method in which oppositions, conflicts, tensions, and refutations were courted rather than avoided or evaded. Hegel was a student of the classical, laissez-faire economists who held that wealth would be maximized by the free play of competition. In this view if traders and producers ceased to compete with one another, the whole level of economic life would be lowered. General prosperity could be reached only at the expense of labor and anxiety. So it is, Hegel believed, with the categories of our thought, the systems of philosophers, and the forms of life and society. There is no tranquillity to be had by withdrawal and isolation. Our categories compete with one another, and out of their competition emerges something better than either of them could have accomplished alone. But it is not possible for the superior category to go into retirement, for without the spur of competition it would fall into decay.

Furthermore, just as competition requires the competitors to continue in business—for if one destroys the others, there is monopoly and stagnation—so the competing categories cannot be swallowed up and lost in the Absolute Idea but must all play their part in maintaining its life and stability. There is nothing fanciful in this comparison. Indeed, it gains support from Hegel's "System der Sittlichkeit" of 1802 ("System of Morality"; in *Schriften zur Politik und Rechtsphilosophie*, edited by Georg Lasson, 2nd ed., Leipzig, 1923), in which it is quite clear that Hegel's systematic thinking was influenced by his understanding of economic theory. For example, in this essay he developed the triad need–labor–enjoyment and described labor as "the destruction of the object … but in such a way that another is put in its place." Here Hegel compared labor with knowledge and undoubtedly had in mind (in accord with his tendency to take German words in the sense of their roots) the element of negation (*nicht*) in the word for destruction (*Vernichtung*). The

destruction of the natural object is the creation of an artificial one.

NEGATION. Negation, indeed, is the vital notion in Hegel's account of the dialectic. In the Preface to the *Phenomenology* Hegel wrote, "The life of God and divine knowledge may, if we wish, be described as love disporting with itself; but this idea is degraded into mere edification and insipidity if it lacks the seriousness, the pain, the patience and the labour of the negative." "Seriousness," "pain," "patience," and "labor" would be strange words to use of the negative symbol of formal logic. Expressed in theological-economic terms Hegel's view is that God cannot be a mere consumer, for there is no consumption without labor, and labor has to face a recalcitrant nature that has to be understood and humored. Thus, there is no God apart from nature. In moral terms there is no good without evil, and in logical terms there is no truth without error. These, according to Hegel, are central truths of dialectics.

But surely, it will be said, this conflicts with such obvious facts as that there are some who consume without working, that in mathematics there are sequences of necessarily true propositions with no admixture of falsity, and that some things—for example, conscientious action—are good without qualification. As to the first point, Hegel argued in the *Phenomenology* that the master who consumes what his slave produces for him destroys what he consumes, whereas the slave shapes the external world in such a way that mind is embodied in it. Hence, the slave is on the road to freedom, whereas the master, who does not work, destroys without creating. As to mathematics, Hegel was inclined to hold it in contempt. There is no space here to consider the strange things he said about it, and it need only be remarked that he held that philosophical truth is utterly different from mathematical truth in that false philosophical views are taken up into true philosophy whereas false mathematics is not taken up into true mathematics. As to the alleged unmixed goodness of conscientious action (the Kantian "good will"), Hegel held that the morality of conscience contained in itself the seeds of willfulness and arbitrariness, for the most atrocious deeds can be defended on the ground that the man who committed them genuinely thought them right. Obedience to one's own conscience, Hegel thought, is an advance over obedience to the commands of an external lord but is nevertheless an unstable basis for morality.

Several ways in which the negative element is important in Hegel's method have been discussed. There is the conceptual competition without which thought must decay. Then, there is the polar character of certain fundamental notions that makes the one unthinkable without its opposite. At the prephilosophical level Hegel gave above and below, right and left, father and son, as examples. At the philosophical level his examples, were good and bad, master and slave, thought and nature. But not only do these opposites require each other; they also pass into each other. Good will can pass over into atrocity; philosophical truth is the result of errors that supplement each other; the master satisfies his desires but becomes dependent upon the labor of the slave in order to do so; and the slave, by work, controls his desires and develops a rational will. The life of thought in conceptual conflict, the mutual dependence of polar opposites, and the instability or oscillations of philosophical and moral attitudes are different sorts of dialectic that Hegel emphasized on different occasions. If they have anything in common, it is the activity of negation.

There are two other aspects of the dialectic to discuss, the role of reason and understanding and the role of skepticism.

REASON AND UNDERSTANDING. First, Hegel, following Kant, contrasted the reason, the source of dialectical thinking, with the understanding, the predialectical mode of thought. The understanding, as Hegel saw it, is the type of thinking that prevails in common sense, in the natural sciences, and in mathematics and those types of philosophy that are argued in quasi-scientific or quasi-mathematical ways. Fixed categories are uncritically adhered to, demonstrations are produced (only to be demolished), analyses are made, and distinctions are drawn. Analyzing and distinguishing are necessary foundations of philosophical activity but only to prepare the way for the more sinuous and subtle method of the dialectic. Once an analysis has been made, the elements of it are seen to conflict and collide as well as to cohere. First, the understanding isolates, then comes the Reason's negative moment of criticism or conflict, and after that its speculative moment of synthesis. It should be mentioned that distinctions somewhat similar to the distinction between the understanding and the reason had already been made by Plato when he distinguished between the highest knowledge and knowledge in the various sciences, by Spinoza in his second and third kinds of knowledge, by Blaise Pascal with his *esprit de géometrie* and *esprit de finesse*, by David Hume with his reason and imagination, and by Edmund Burke when he contrasted the abstract rationalism of the Enlightenment with the organic, evolutionary view of society that he preferred. These distinc-

tions are not all quite like that drawn by Hegel, but in his theory there is something corresponding to each of them.

SKEPTICISM. Second, Hegel thought that skepticism was an important forerunner and essential ingredient of the dialectical method. In a review of a book by G. E. Schulze that appeared in 1802, Hegel wrote appreciatively of the skepticism of Sextus Empiricus and of the skeptical features in the philosophy of Parmenides, of whom he wrote, "This skepticism, which in its pure *explicit* form comes forward in Parmenides, is to be found *implicit* in every genuinely philosophical system, for it is the free aspect [*die freie Seite*] of every philosophy" ("Verhältnis des Skeptizismus zur Philosophie," in *Kritisches Journal der Philosophie* 2 (1802): 1–74; quoted from *Sämtliche Werke*, edited by Georg Lasson, Vol. I, pp. 174–175). In the same essay Hegel wrote that when Spinoza held that God is the immanent but not the transcendent cause of the world, he was equating the cause with the effect, even though the very notion of an effect implies that it is distinct from the cause. Hegel agreed with Spinoza's equation but concluded that it shows that the reason can accept the principle of contradiction only as a formal principle. In "genuine" philosophy cause and effect are seen as both distinct and identical.

Hegel illustrated his comment that skepticism is "the free aspect" of philosophy in the following way. Dogmatists, he said, regard individual men as objects in the power of rules, laws, and customs. The more the dogmatists study man, the more they show him in subjection to these forces. When, however, the skeptics attack dogmatism, "they raise the freedom of Reason above this necessity of nature." An example of this is the way in which Europeans came to question their own concepts of law and morals when they were brought face to face with cultures very different from their own. When such skeptics as Montaigne mockingly insisted on these differences, men became more conscious of their own institutions and recognized the possibility of changing them. In theoretically breaking down men's traditional views and institutions, the skeptic frees men from the unconscious power of these views and institutions. Hegel repeated his general assessment of skepticism in the *Encyclopedia* (Sec. 81, addition 2) and in his *Lectures on the History of Philosophy*. In these lectures Hegel said that skepticism is "the demonstration that all that is determinate and finite is unstable." Hegel went on to say that "positive philosophy," by which he meant philosophy that is not content to remain in total skepticism, "has the negative to Scepticism in itself; thus it does not oppose, nor is it outside of it, for Scepticism is a moment in it" (Haldane and Simpson, 1955, Vol. II, p. 330).

FREEDOM

From what has just been said, it is clear that Hegel's account of dialectic and of reason is closely linked with his view of freedom. The exercise of thought in its most developed forms involves the negation of what had seemed firm and certain and the opening up of new possibilities. That mind is freedom applies both to the understanding and the reason, since both are spontaneous activities that interpret and arrange. But because the understanding is confined to a fixed system of categories, it is less free than the reason that criticizes, stretches, and transforms the categories of the understanding.

Freedom is, of course, logically connected with will, and according to Hegel, will is as essential to mind as intellect is. Reference has already been made to Section 482 of the *Encyclopedia*, in which Hegel asserted that the unity of theoretical and practical mind is free will. In the preceding sections he had argued that thought presupposes mind as practical, since classifying and explaining are activities through which the world is, so to speak, appropriated by the mind. In the sections of the *Encyclopedia* in which he expounded the categories of cognition and of will, Hegel endeavored to show that mere cognition is at a lower stage than will and that will is thus the actuality of what is only potential in knowledge. He also argued that the freedom and necessity that are opposed to each other are abstractions and that what is concrete must combine both. The very nature of necessity, he continued, presupposes a will on which it is a constraint.

At the logical-metaphysical level, therefore, Hegel held a view that implied that freedom is essential to mind, both the presupposition and outcome of intelligence, and in its concrete form inseparable from constraint and necessity. This view of the matter pervaded his account of freedom in the social and political sphere. Freedom is not something merely opposed to constraint; on the contrary, it presupposes and requires restraint. This is true of concrete freedom. However, abstract or negative freedom, when it is more than a moment in actual or positive freedom, is a purely destructive force. Hegel considered that this negative freedom played a large part in the French Revolution. The old corporations and institutions were destroyed in such a frenzy of annihilation that it took several years for new institutions to be created and recognized as authoritative. Furthermore, when the conflicting interests in society are overcome, individuals come to be treated as equal, undifferentiated, replaceable, and

expendable units. The events of the Reign of Terror thus led Hegel to hold that purely negative freedom was associated with force and death. The logical connections are not altogether clear, but it may well be that the links between egalitarianism, antinomianism, violence, and contempt for human life are not wholly accidental.

Freedom, according to Hegel, is something that has to be achieved, and it therefore would be impossible in the absence of opposition and negation. Hence, although negative freedom in its abstract form is a "fury of destruction," it is a necessary element in concrete freedom. Free will is not the liberty of indifference but the rational organization of the feelings and impulses.

Rationality is not a power that could reside in an isolated individual, however. To be rational, the individual must draw upon the resources of an organized and differentiated society and must be "formed" and educated to do this. His will is then in harmony with the ends of the various social groups by which he has been influenced and, in civilized societies, with the more complex ends of the state. In conforming to these pressures and in obeying the laws of the state, the individual is achieving his own rational ends and in so doing is free.

Hegel, like Jean-Jacques Rousseau, also held that an individual might be free even when he was being coerced, for although he might dislike the force applied against him, this dislike would be an expression of his particular whims, not of his rational insight, as can be seen when he approves of the imposition of a like force upon other people in like circumstances. Insofar as the criminal who is being punished would wish others to be punished who committed a like crime against him, he wills his own punishment.

FREEDOM IN HISTORY. Hegel considered that the history of the human race is a development from less to greater freedom and from less adequate forms of freedom to freedom in its perfection. Thus, his philosophy of history can be understood only in terms of his conception of freedom. In the Oriental world there was no freedom for the subjects and only an arbitrary, irrational freedom for the despot who ruled over them. In the classical world of Greece and Rome there was a more adequate conception of freedom, and more men achieved freedom than in the Oriental despotisms. In the Greek city-state the citizens often regarded themselves as finding their fulfillment in the achievements of their city, apart from which they conceived of no life for themselves. Indeed, they might accept personal defeat and misfortune and submit to what they called destiny and still regard themselves as free in so doing. Of course, there were slaves who had no part in this activity and had no freedom.

Christianity offered the prospect of freedom to all men, a freedom, furthermore, that transcended the given social order. In what Hegel called the Germanic world—that is, the Christian civilization that grew out of Protestantism—this latest form of freedom was being realized in the manifold institutions of Europe and America and in the states in which these institutions flourished and by which they were regulated and protected. In Christianity the individual is regarded as of infinite value, as a candidate for eternal salvation, and although the emphasis on subjective freedom can lead, as it did in the French Revolution, to contempt for social institutions, it comprises the form and aspect of freedom that gives its special quality to modern civilization, with its romantic art, romantic love, and support for the rights of conscience (*Philosophy of Right*, Sec. 124).

It is apparent from the foregoing that Hegel rejected the liberal view that man is free to the extent that he is guaranteed a sphere within which he can do what he wishes without interference from others who are guaranteed a like position. Such freedom he stigmatized as negative, abstract, or merely willful. Men enjoy concrete freedom when the various orders and groups of civilized life are maintained in and by the state. In this passage of the *Lectures on the Philosophy of History* (Hoffmeister, Vol. XVIIIA, p.111) Hegel also emphasized that in submitting their private wills to the laws of the state and to the rules of its subordinate but free institutions, men were submitting their passions to the control of reason. Thus, the argument comes full circle. The theoretical reason is inseparable from will and from freedom; necessity and negative freedom are only abstractions; in concrete freedom the negative, destructive element is held in check and rendered fruitful by being realized in institutions; the individual enjoys concrete freedom when he is educated to live in a civilized state and to be guided by the reason that permeates it.

There is no space here to criticize this view in any detail, for in a way it is a cross-section of the whole Hegelian metaphysic. It should be noted, however, that when a critic maintains that real freedom is what Hegel called negative or abstract freedom and when he goes on to maintain that "concrete freedom" is not freedom but indoctrinated submission, then he is criticizing Hegel's terminology rather than the substance of his view. To say that freedom consists of a willing acceptance of the tasks imposed by a civilized state is certainly to extend and perhaps to distort the ordinary senses of the term and to cap-

ture a word from the liberal vocabulary for use in a far from liberal scheme of concepts. It was Hegel's view, however, that the thoughts that the liberal phraseology expressed necessarily move in the directions he described and that societies themselves, the embodiments of men's thoughts and aims, move in these directions, too.

AESTHETIC THEORY

We have already seen that Hegel discussed the nature of art and of beauty toward the end of both the *Phenomenology* and the *Encyclopedia*. Art, according to Hegel, is one of the manifestations of Absolute Mind, of which religion and philosophy are the other two. Thus, although art presupposes the civilized life of the state, it also transcends it. In his lengthy *Vorlesungen über die Aesthetik* (Lectures on aesthetics) Hegel developed his theories of art and beauty in great detail. The lectures possess great power and attraction, and so much of their value resides in the details that a summary treatment is bound to be difficult.

THREE STYLES OF ART. Hegel's account of beauty is a modification of Friedrich Schiller's view, in his *Letters on the Aesthetic Education of Mankind* (1795), that beauty is the mediation between the sensible and the rational. According to Hegel, beauty is the rational rendered sensible, the sensible appearance being the form in which the rational content is made manifest. This sensible embodiment of the rational, he held, can take place in three principal ways: symbolic art, classical art, and romantic art.

Symbolic art. In the first and least adequate form, symbolic art, the sensible shape merely symbolizes the rational content without penetrating and transforming it. A lion may symbolize courage; a bird, the soul; or a temple, the presence of a god who nevertheless remains a mystery. Thus, in symbolic art the sensible object refers away from itself to a rationality that is enigmatically and mysteriously beyond it. In thus referring away from the sensible symbol to something vast and merely adumbrated, symbolic art sometimes achieves the sublime.

Classical art. In classical art, the second form of sensible embodiment, the sensible expression is adequate to the idea that it gives expression to and does not point vaguely beyond itself. This is typified in sculptures of the human body so formed that the divine ideal is realized in the stone, not merely hinted at. A temple makes us think of the god but is not the god. In a statue of Apollo the god is visible and tangible in the stone. Hegel pointed out that works of classical art have independence and completeness, so that when they have been created, it seems that there is nothing more left to do done. "Nothing more beautiful," he wrote, "can be or become."

Romantic art. Christianity, however, with its emphasis on the infinite value of the individual and upon subjective freedom, made classical art seem somewhat unsatisfactory. More is required than works of art in which reason, as Hegel put it, "stands in quiet and blessedness in bodily form." When the self and its inner life are regarded as of infinite value, the forms of art must move on from balance and harmony to the storm and turmoil of the subjective. According to Hegel, it is in romantic art that this progress to subjectivity and self-consciousness is achieved. Romantic art turns its back on the quiet and balanced beauty of the classical and "weaves the inner life of beauty into the contingency of the external form, and allows full scope to the emphatic features of the unbeautiful." In romantic art, as in symbolic art, there is much that is bizarre and even grotesque, but romantic art is on a higher level than symbolic art because the mind expressed in it is more complex and sophisticated. And in romantic art the mind has achieved a greater measure of freedom than in classical art because romantic art is less involved in and hampered by the sensible embodiment.

PRODUCTS OF ART. Hegel's view of the three main types of beauty is closely linked with his view of the main types of artistic product. Hegel divided the arts into architecture, sculpture, painting, music, and poetry. Works in any of these mediums may be produced in the symbolic, classical, or romantic styles, but, according to Hegel, architecture is particularly appropriate to symbolic art, sculpture to classical art, and painting, music, and poetry to romantic art.

Architecture. Architecture, Hegel held, is the basic art, the art that men first practice, for its material is mindless and its forms depend upon the weight and physical properties of this mindless medium. The architecture of early men, by bringing them together to worship the gods in temples, served to bring unity into their societies. Hegel imagined the men who built the first temples as they cleared the ground on which to build them, and he described this as "clearing the undergrowth of finitude."

Sculpture. In architecture a house is provided for the god, and the god is prepared for and expected. He is not, however, embodied or manifested in the stones of a mere building. In classical sculpture the god is embodied in the stone in such a way that all the parts of the statue combine in expressing and proclaiming him. Hence, it is not a mindless symbol of the mind beyond but a unified

expression of it. Hegel contrasted the stiff regularity of Egyptian sculpture with the harmonious independence of the Greek, the acme of classical art. In Christian sculpture this Greek ideal does not predominate, and even when, as with Michelangelo, it is fully understood and mastered, it is associated with "the kind of inspiration that is found in romantic art."

Painting, music, and poetry. The three romantic arts of painting, music, and poetry differ from the arts of sculpture and architecture, according to Hegel, by being more "ideal." One thing he seems to have meant was that the productions of these arts are not three-dimensional like the productions of architecture and sculpture. Painting, of course, is two-dimensional, and Hegel thought it is more ideal than sculpture because it is further removed from the solid substance of material things. He appears to have argued that the painter transforms to an extent that the sculptor has no need to do. In reducing the three dimensions to two, space is somehow rendered more "inward" and "subjective," and the first step has been taken on the road to poetry.

The next step toward subjectivity is taken by music, which abandons all the dimensions of space as well as the senses of sight and touch. Hearing, according to Hegel, is a "more subjective" sense than sight because it is less practical and more contemplative.

In poetry the sensible elements of music, the notes or tones, are replaced by words that stand for thoughts. "The art of poetry," Hegel wrote, "is the universal art of mind that has become free and is no longer dependent upon external sensible material for its realization." Within poetry as a whole he distinguished epic, lyric, and dramatic poetry. Hegel's account of dramatic poetry is particularly interesting. "In tragedy," he wrote, "individuals destroy themselves through the onesidedness of their upright will and character, or they are forced to resign themselves and identify themselves with a course of action to which they are fundamentally opposed." In comedy, on the other hand, there is no such reconciliation; the characters pursue courses of action that have only subjective significance. Indeed, in comedy, according to Hegel, the subjectivity characteristic of romantic art is taken to such an extreme that all unity is dissolved; with it goes beauty, too. In comedy there is merely a series of subjective interests playing against one another, as opposed to the aim of all art, which is the revelation of the eternal and divine in sensible form.

NATURAL BEAUTY. The discussion has thus far been confined to the beauty of works of art (*das Kunstschöne*). It is with this that by far the greater part of the "Lectures on Aesthetics" is concerned. In the second chapter, however, Hegel did say something about natural beauty (*das Naturschöne*). He discussed the notions of regularity, symmetry, harmony, and conformity to law and also the beauty claimed for plants, animals, and human beings. He concluded his discussion of the subject with some comments on how natural beauty falls short of artistic beauty. Plants and animals, he granted, are more beautiful than inanimate natural objects, but what we see of them is their outward coverings, not the soul that works within, for that is concealed by the visible feathers, hair, scales, fur, and the like that cover them. Hegel referred to natural beauty as the "prose of the world." Although Hegel did not altogether deny the beauty of nature, it is clear that he ranked it very low. Indeed, the structure of his system made this inevitable, for it is the self-conscious achievements of man that form its culmination.

It would seem that the triadic divisions of the "Lectures on Aesthetics" constrained and even corrupted Hegel's argument. An example of this occurs in his account of dramatic poetry, into which he introduced a species called "drama," the function of which was to add one species to tragedy and comedy and thus make three species of dramatic poetry.

Hegel also tended to confuse conceptual and historical relationships. For example, the distinction between symbolic, classical, and romantic art was intended to be made on conceptual grounds, but, on the other hand, Hegel had in mind historical progression. Here, as elsewhere, Hegel confused historical types, such as romanticism, with conceptual types, such as tragedy, which have no necessary temporal sequence. Perhaps the most interesting case of this is Hegel's suggestion that art comes to an end with the highest flights of romanticism. We have already seen that Hegel brought his account of dramatic poetry to an end with comedy, the most subjective of all art forms. At the very end of the "Lectures on Aesthetics" he said that "in this culmination comedy is leading straight to the dissolution of art in general." It is unlikely that Hegel believed that art was coming to an end, any more than he believed that with the Prussian state, history was coming to an end. Yet in each case he argued in such a way as to suggest that the culmination of a conceptual sequence must also be the conclusion of a historical progress. Insofar as he held that history was the movement of the Divine in the world, it was natural to make this identification, extravagant as it is. Bosanquet, who denied that Hegel believed that art was on the point of final dissolution, held that he did foresee that it was

about to suffer an eclipse in the new form of society. "But we must claim extraordinary insight for him, who, still under the spell of Schiller and Goethe, described the present exhaustion of the art-impulse and the conditions hostile to it in language approaching that of John Ruskin and William Morris" (*A History of Aesthetic*, 4th ed., London, 1917, p. 361).

PHILOSOPHY OF RELIGION

A few commentators have regarded Hegel's philosophy as atheistic, but most have considered it to be either theistic or pantheistic. Certainly religious expressions abound in his writings, even in the *Logic*. It has been shown how closely he associated art with religion and how he applied religious epithets to the state. It was also pointed out that the *Phenomenology* might with some justification be interpreted in atheistic terms. It would be obviously overstraining the evidence, however, to interpret Hegel's mature system in this way, for in the system religion is a form of Absolute Mind, along with art and philosophy, which is the supreme expression of the Absolute Mind. According to Hegel, religion represents or pictures the Absolute, whereas philosophy conceives or thinks it. The same truth, that is, expressed in quasi-imaginative form in one and in conceptual form in the other.

CHRISTIANITY. Since the concept is supreme and ultimate, philosophy surpasses religion to this extent, but in doing this, it finally and fully justifies Christianity, which is the absolute religion. The doctrine that elevates Christianity above all other religions is the doctrine of the Incarnation, which, according to Hegel, is the religious expression of the philosophical truth that the Infinite Being is not distinct from what is finite but is necessarily manifested in it. Hegel also interpreted the doctrine of the Trinity in philosophical terms. In the "Science of Logic" God is revealed as he is before the creation of the world; in the "Philosophy of Nature," in his material embodiment; and in the "Philosophy of Mind," as reconciling the finite and the Infinite. In this way the Father, the Son, and the Holy Spirit are explained in terms of the main themes of the Hegelian system. Again, Hegel interpreted the doctrine that God is love to mean that although the Infinite Being cannot exist without negation and opposition, the negation and opposition are finally reconciled. Finally, it should be mentioned that Hegel gave a series of lectures on the traditional proofs for the existence of God. He admitted the force of Kant's criticisms of these proofs but claimed to have reformulated the arguments so as to meet the criticisms. In particular, he held that the Ontological Argument, which Kant had regarded as vital but unsound, was valid when properly understood.

Undoubtedly, Hegel's later writings are much closer to orthodox Christianity than his earlier ones. The early "Life of Jesus" had nothing to say about the Resurrection, whereas in the *Lectures on the Philosophy of Religion* this doctrine was stated and defended. Hegel here wrote of "the death of death," of "the triumph over the negative," of mind as "the negative of this negative which thus contains the negative in itself," and of "the division of the divine idea and its reunion" that is "the whole of history." Although Hegel said that God appeared in the flesh at a particular time and in a particular individual, his account of the matter seems to be extremely general. In the Christian doctrine of the Incarnation, God became man in Jesus Christ at a particular time and place, whereas Hegel's God is incorporated in the finite world. It would seem that a highly specific historical view is replaced by a highly general metaphysical one. Hegel himself did not take this view of his own work, nor did a younger contemporary of his, Karl Friedrich Göschel, in his *Aphorismen über Nichtwissen und absolutes Wissen im Verhälnisse zur christlichen Glaubenserkenntnis* (Berlin, 1829). In the *Encyclopedia* (Sec. 564) Hegel recommended this book, which is generally regarded as giving a theistic account of the Absolute. Just before referring to Göschel's book, Hegel had written, "God is only God in so far as he knows himself; his self-knowledge is moreover his consciousness of himself in man and man's knowledge *of* God, a knowledge that extends itself into the self-knowledge of man *in* God."

What cannot be doubted is that Hegel's philosophy of religion contained elements that could easily be developed in ways that go counter to orthodox Christianity. Thus, when D. F. Strauss argued, in his *Life of Jesus* (1835), that the Gospel story was a set of myths, he was consciously working out what he thought was the consequence of Hegel's view that in religion the truth about God is understood in representative or pictorial terms. Again, Ludwig Feuerbach, in his *The Essence of Christianity* (1841), endeavored to interpret the Christian doctrines in human and psychological terms as the imaginary fulfillment of wishes that cannot be satisfied here on Earth. We have already referred to the passage in Hegel's *The Positivity of the Christian Religion*, in which he said that in the days of imperial Rome men who had been robbed of their freedom in this world sought for it in a heaven beyond. Feuerbach, who, of course, had not seen this work, could have read something similar in the *Phenomenology*. It is a very short step from Hegel's view

that the infinite is manifested in the finite to the view that it is a projection of it. Perhaps the truth of the matter is that the Christian religion, according to Hegel, is adequate in its own sphere and that the philosophy of religion is required to counteract false religious views and false views about religion but is not a substitute for it. This is the interpretation given by Lasson in the introduction to Hegel's philosophy of religion printed at the end of his edition of Hegel's *Lectures on the Philosophy of Religion*.

See also Dialectic; Philosophy of Mind.

Bibliography

WORKS BY HEGEL

The edition of Hegel's collected works, *Werke. Vollständige Ausgabe*, published soon after his death has already been mentioned. This was republished in 26 volumes under the editorship of Hermann Glockner (Stuttgart, 1927–1940) as the jubilee edition. Volumes 23–26 of this edition are an invaluable *Hegel-Lexikon* (1st ed., 1935; rev. ed., 2 vols., 1957). Mention must be made of Georg Lasson's edition, *Georg Wilhelm Friedrich Hegels Sämtliche Werke* (Leipzig, 1905–); this was to have been revised and reedited by Lasson and Johannes Hoffmeister, and some volumes were published (Hamburg, 1949–). As a result of the death of Lasson and of Hoffmeister, the plan was altered. The *Sämtliche Werke, neue kritische Ausgabe*, which Hoffmeister started to edit, will continue under the editorship of various scholars. This edition contains Hegel's letters in *Briefe von und an Hegel*, 4 vols. (1952–1960). The unpublished writings prior to 1800 will appear in this new critical edition under the editorship of Gisela Schüler. Another collection with still earlier documents is *Dokumente zu Hegels Entwicklung*, edited by Hoffmeister (Stuttgart: Fromanns, 1936).

Some of Hegel's early writings were translated from Herman Nohl, ed., *Hegels theologische Jugendschriften* (Tübingen, 1907) by T. M. Knox as *Early Theological Writings* (Chicago: University of Chicago Press, 1948). There is a translation of the *Phenomenology of Mind* by J. B. Baillie (London, 1910; 2nd ed., 1931). William Wallace translated the encyclopedia "Logic," with excellent prolegomena in *The Logic of Hegel* (Oxford: Clarendon Press, 1873); there are subsequent editions. He also translated the last part of the *Encyclopedia* (in *Hegel's Philosophy of Mind*, Oxford, 1894). See also *The Science of Logic*, translated by W. H. Johnston and L. G. Struthers, 2 vols. (London, 1929); *The Philosophy of Right*, translated and edited by T. M. Knox (Oxford, 1942), a very fine work of scholarship; *The Philosophy of Fine Art*, translated by F. P. B. Osmaston, 4 vols. (London, 1920); *Lectures on the Philosophy of Religion*, translated by E. B. Speirs and J. B. Sanderson, 3 vols. (London, 1895); and *Lectures on the History of Philosophy*, translated by E. S. Haldane and F. H. Simsen, 3 vols. (London 1892–1896; reprint, 1955). J. Sibree's translation of the *Lectures on the Philosophy of History* (London, 1857) has been reprinted several times. See also *Hegel's Political Writings*, translated by T. M. Knox with an introductory essay by Z. A. Pelczynski (Oxford, 1964). There is a translation of the *Encyclopedia* with notes by Gustav E. Mueller (New York, 1959). There are two useful books of selections, *Hegel: Selections*, edited by Jacob Loewenberg (London and New York, 1929), and *The Philosophy of Hegel*, edited by C. J. Friedrich (New York: Modern Library, 1953).

WORKS ON HEGEL

The System

Caird, Edward. *Hegel*. London and Edinburgh, 1883.

Copleston, Frederick. *A History of Philosophy*, Vol. VII: *From Fichte to Nietzsche*. London, 1963. Chs. 9–11.

Croce, Benedetto. *Ciò che è vivo e ciò che è morto della filosofia di Hegel*. Bari, Italy, 1907. Translated by Douglas Ainslie as *What Is Living and What Is Dead in the Philosophy of Hegel*. London: Macmillan, 1915.

Findlay, J. N. *Hegel: A Re-examination*. London: Allen and Unwin, 1958. Emphasizes the *Logic* and interprets Hegel in somewhat nonmetaphysical terms.

Glockner, Hermann. *Hegel*, 2 vols. Stuttgart, 1929. These are Vols. XXI–XXII of the jubilee edition of Hegel mentioned above.

Grégoire, Franz. *Études hégéliennes*. Louvain, Belgium, and Paris, 1958. Contains detailed and fruitful discussions of the central issues.

Haering, T. L. *Hegel: Sein Wollen und Sein Werk*, 2 vols. Leipzig, 1929. A leading authority.

Hartmann, Nicolai. *Die Philosophie des deutschen Idealismus*, Vol. II: *Hegel*. Berlin and Leipzig, 1929.

Kaufmann, Walter. *Hegel: Reinterpretation, Texts and Commentaries*. Garden City, NY: Doubleday, 1965.

Kline, George L. "Some Recent Re-interpretations of Hegel's Philosophy." *Monist* 48 (1964): 34–73.

McTaggart, J. M. E. *Studies in Hegelian Cosmology*. Cambridge, U.K.: Cambridge University Press, 1901. Stimulating and idiosyncratic.

Mure, G. R. G. *An Introduction to Hegel*. Oxford: Clarendon Press, 1940. Emphasizes Hegel's debt to Greek philosophy, Aristotle in particular.

Pringle-Pattison, A. S. *Hegelianism and Personality*, 2nd ed. London and Edinburgh: Blackwood, 1887. The first edition was published under the name Andrew Seth. See the article "The Absolute."

Seeberger, Wilhelm. *Hegel: Oder die Entwicklung des Geistes zur Freiheit*. Stuttgart, 1961. A new statement of the "orthodox" interpretation.

Serreau, René. *Hegel et l'hégélianisme*. Paris: Presses Universitaires de France, 1962. Brief but informative book with good introductory account of the various Hegelian schools.

Stace, W. T. *The Philosophy of Hegel*. London: Macmillan, 1924; New York: Dover, 1955. A detailed account based on the *Encyclopedia*.

Early Writings

Asveld, P. *La pensée religieuse du jeune Hegel*. Louvain, Belgium, 1953.

Dilthey, Wilhelm. *Die Jugendgeschichte Hegels*. Berlin, 1906. It is included as Vol. IV of Dilthey's *Gesammelte Schriften* (Berlin and Leipzig, 1921).

Hyppolite, Jean. *Introduction à la philosophie de l'histoire de Hegel*. Paris, 1948. A valuable account of Hegel's early views on history and politics.

Kaufmann, Walter. "The Young Hegel and Religion." In his *From Shakespeare to Existentialism*. Boston: Beacon Press, 1959. Ch. 8.

Lukács, Georg. *Der junge Hegel*. Berlin, 1948. An interesting Marxist interpretation.

Peperzak, A. T. P. *Le jeune Hegel et la vision morale du monde*. The Hague, 1960.

Walsh, W. H. "The Origins of Hegelianism." In his *Metaphysics*. London: Hutchinson, 1963. Ch. 9.

"Phenomenology of Mind"

Hyppolite, Jean. *Genèse et structure de la Phénomenologie de l'Esprit de Hegel*. Paris, 1946.

Kojève, Alexandre. *Introduction à la lecture de Hegel*. Paris, 1947. Gives the "atheistic" interpretation.

Loewenberg, Jacob. "The Comedy of Immediacy in Hegel's Phenomenology. II." *Mind*, n.s. 44 (1935): 21ff.

Loewenberg, Jacob. "The Exoteric Approach to Hegel's Phenomenology. I." *Mind*, n.s. 43 (1934): 424ff.

Wahl, Jean. *Le malheur de la conscience dans la philosophie de Hegel*. Paris, 1929; 2nd ed., 1951.

"Logic"

Hyppolite, Jean. *Logique et existence*. Paris, 1953.

McTaggart, J. M. E. *Commentary on Hegel's Logic*. Cambridge, U.K.: Cambridge University Press, 1910.

Mure, G. R. G. *A Study of Hegel's Logic*. Oxford: Clarendon Press, 1950.

Dialectic

Coreth, E. *Das dialektische Sein in Hegels Logik*. Vienna, 1952.

Grégoire, Franz. *Études hégéliennes*. Louvain, Belgium, and Paris, 1958. See Study II.

McTaggart, J. M. E. *Studies in the Hegelian Dialectic*. Cambridge, U.K.: Cambridge University Press, 1896; 2nd ed., 1922.

Mueller, Gustav E. "The Hegel Legend of Thesis, Antithesis and Synthesis." *Journal of the History of Ideas* 19 (1958): 411–414.

Popper, Karl. "What Is Dialectic?" *Mind*, n.s. 49 (1940): 403–426. Revised and reprinted in his *Conjectures and Refutations* (London: Routledge, 1963), Ch. 15.

Social and Political Philosophy

Foster, M. B. *The Political Philosophies of Plato and Hegel*. Oxford: Clarendon Press, 1935. An original and valuable discussion.

Grégoire, Franz. *Études hégéliennes*. Louvain, Belgium, and Paris, 1958. See Study IV.

Haym, Rudolf. *Hegel und seine Zeit*. Berlin, 1857; reprinted, Hildesheim, 1962.

Kaufmann, Walter. "The Hegel Myth and Its Methods." In his *From Shakespeare to Existentialism*. Boston: Beacon Press, 1959. Ch. 7 is a criticism of Ch. 12 of Popper's *The Open Society and Its Enemies*.

Knox, T. M. "Hegel and Prussianism." *Philosophy* 15 (1940): 51–63.

Plamenatz, John. *Man and Society*, 2 vols. New York: McGraw-Hill, 1963. Vol. II, Chs. 3–4, "The Social and Political Philosophy of Hegel."

Popper, Karl. *The Open Society and Its Enemies*, 2 vols. London: Routledge, 1945. See Vol. II, Ch. 12.

Rosenzweig, Franz. *Hegel und der Staat*, 2 vols. Oldenburg, Germany, 1920.

Weil, Eric. *Hegel et l'état*. Paris, 1950.

Aesthetics

Bosanquet, B. *History of Aesthetic*. London: Sonnenschein, 1892. Chs. 2, 12.

Teyssèdre, B. *L'esthétique de Hegel*. Paris, 1958.

Philosophy of Religion

Chapelle, A. *Hegel et la religion*, Vol. I: *La problématique*. Paris, 1964–. The first of four projected volumes on the subject.

Grégoire, Franz. *Études hégéliennes*. Louvain, Belgium, and Paris, 1958. Study III.

Iljin, I. *Die Philosophie Hegels als kontemplative Gotteslehre*. Bern, 1946.

McTaggart, J. M. E. *Studies in Hegelian Cosmology*. Cambridge, U.K.: Cambridge University Press, 1901. Especially Chs. 2, 3, 6 ("Sin"), and 8 ("Hegelianism and Christianity").

Mueller, Gustav E. "Hegel's Absolute and the Crisis of Christianity." In *A Hegel Symposium*, edited by D. C. Travis. Austin: University of Texas, 1962.

For further discussion see *Hegel-Studien*, Vol. 1 (Bonn, 1961), Vol. 2 (Bonn, 1963), and Supp. Vol. 1 (Bonn, 1964). This is edited by F. Nicolin and O. Pöggeler and contains articles, reviews, and bibliography. See also *Études hégéliennes* (Neuchâtel, 1955), containing articles by Alexandre Kojève, Georges Bataille, R. Queneau, Jean Wahl, and Eric Weil, and *A Hegel Symposium* (Austin: University of Texas, 1962), with articles by C. J. Friedrich, Sidney Hook, Helmut Motekat, Gustav E. Mueller, and Helmut Rehder.

Beiser, Frederick C., ed. *The Cambridge Companion to Hegel*. New York: Cambridge University Press, 1993.

Berthold-Bond, D. *Hegel's Grand Synthesis: A Study of Being, Thought and History*. Albany, NY: State University of New York Press, 1989.

Burbidge, J. W. *Historical Dictionary of Hegelian Philosophy*. Lanham, MD: Scarecrow Press, 2001.

Burns, Tony. *Natural Law and Political Ideology in the Philosophy of Hegel*. Brookfield, VT: Avebury, 1996.

Collins, Ardis B. *Hegel on the Modern World*. Albany, NY: State University of New York Press, 1995.

Cullen, Bernard, ed. *Hegel Today*. Brookfield, VT: Avebury, 1988.

Desmond, William, ed. *Hegel and His Critics: Philosophy in the Aftermath of Hegel*. Albany: State University of New York Press, 1989.

Fox, Michael Allen. *The Accessible Hegel*. Amherst, NY: Humanity Books, 2005.

Gascoigne, Robert. *Religion, Rationality, and Community: Sacred and Secular in the Thought of Hegel and his Critics*. Boston: M. Nijhoff, 1985.

Harris, Henry Silton. *Hegel's Ladder, Vol. 1 and 2*. Indianapolis: Hackett, 1997.

Henrich, Dieter, ed. *Kant oder Hegel?: Uber Formen der Begrundung in der Philosophie*. Stuttgarter Hegel-Kongress. Stuttgart: Klett-Cotta, 1983.

Hösle, Vittorio. *Hegel's System: Der Idealismus der Subjektivitat und das Problem der Intersubjektivitat Hegels*. Hamburg: F. Meiner, 1987.

Hospers, J. ed. “Hegel Today.” *Monist* (1991) 74: 3.

Hougate, Stephen, ed. *Hegel and the Philosophy of Nature*. Albany: State University of New York Press, 1998.

Houlgate, Stephen. *An Introduction to Hegel: Freedom, Truth, and History*. Malden, MA: Blackwell, 2005.

Hutchings, Kimberly. *Hegel and Feminist Philosophy*. Malden, MA: Blackwell, 2003.

Hyppolite, Jean. *Introduction a la philosophie de l'histoire de Hegel* [Introduction to Hegel's Philosophy of History]. Translated by Bond Harris and Jacqueline Bouchard Spurlock. Gainesville, FL: University Press of Florida, 1996.

Inwood, Michael. *Hegel*. Boston: Routledge & Kegan Paul, 1983.

Inwood, Michael, ed. *Hegel*. Oxford: Oxford University Press, 1985.

Inwood, Michael. *A Hegel Dictionary*. Cambridge, MA: Blackwell, 1992.

Kainz, Howard P. *An Introduction to Hegel: The Stages of Modern Philosophy*. Athens: Ohio University Press, 1996.

Kainz, Howard P. *G. W. F. Hegel: The Philosophical System*. New York: Twayne, 1996.

Kainz, Howard P. *Hegel's Phenomenology, Part I: Analysis and Commentary*. Tuscaloosa: University of Alabama Press, 1976.

Kainz, Howard P. *Hegel's Phenomenology, Part II: The Evolution of Ethical and Religious Consciousness to the Absolute Standpoint*. Athens: Ohio University Press, 1983.

Keenan, Denis King. *Hegel and Contemporary Continental Philosophy*. Albany: State University of New York Press, 2004.

Knowles, Dudley. *Routledge Philosophy Guidebook to Hegel and the Philosophy of Right*. New York: Routledge, 2002.

Lamb, David, ed. *Hegel and Modern Philosophy*. New York: Croom Helm, 1987.

Lamb, David. *Hegel—From Foundation to System*. The Hague: M. Nijhoff, 1980.

MacIntyre, Alasdair C. *Hegel: A Collection of Critical Essays*. Garden City, NY: Anchor Books, 1972.

Maker, William. *Hegel and Aesthetics*. Albany: State University of New York Press, 2000.

Maker, William. *Philosophy without Foundations: Rethinking Hegel*. Albany: State University of New York Press, 1994.

McCarney, Joe. *Routledge Philosophy Guidebook to Hegel on History*. New York: Routledge, 2000.

Moran, Philip. *Hegel and the Fundamental Problems of Philosophy*. Amsterdam: B.R. Gruner, 1988.

Mure, Geoffrey Reginald Gilchrist. *The Philosophy of Hegel*. Bristol, U.K.: Thoemmes Press, 1993.

O'Malley, Joseph J., K. W. Algozin, and Frederick G. Weiss, eds. *Hegel and the History of Philosophy: Proceedings of the 1972 Hegel Society of America Conference*. The Hague: Martinus Nijhoff, 1974.

O'Malley, Joseph J., et al, eds. *The Legacy of Hegel: Proceedings of the Marquette Hegel Symposium 1970*. The Hague: Martinus Nijhoff, 1973.

Ormiston, Alice. *Love and Politics: Re-interpreting Hegel*. Albany: State University of New York Press, 2004.

Petry, Michael John, ed. *Hegel and Newtonianism*. Boston: Kluver Academic Publishers, 1993.

Pinkard, Terry P. *Hegel's Dialectic: The Explanation of Possibility*. Philadelphia: Temple University Press, 1988.

Pippin, Robert B. *Hegel's Idealism: The Satisfactions of Self-Consciousness*. New York: Cambridge University Press, 1989.

Pippin, Robert B., and Otfried Hoffe, eds. *Hegel on Ethics and Politics*. Translated by Nicholas Walker. New York: Cambridge University Press, 2004.

Plant, Raymond. *Hegel*. New York: Routledge, 1999.

Plant, Raymond. *Hegel: On Religion and Philosophy*. London: Phoenix, 1997.

Reyburn, Hugh A. *The Ethical Theory of Hegel: A Study of the Philosophy of Right*. Oxford: Clarendon Press, 1967.

Rockmore, Tom. *Before and After Hegel: A Historical Introduction to Hegel's Thought*. Indianapolis: Hackett, 2003.

Rockmore, Tom. *Hegel, Idealism, and Analytic Philosophy*. New Haven: Yale University Press, 2005.

Rockmore, Tom. *Hegel's Circular Epistemology*. Bloomington: Indiana University Press, 1986.

Rosen, Stanley. *G. W. F. Hegel: An Introduction to the Science of Wisdom*. New Haven, CT: Yale University Press, 1974.

Sedgwick, Sally. *The Reception of Kant's Critical Philosophy: Fichte, Schelling, and Hegel*. New York: Cambridge University Press, 2000.

Singer, Peter. *Hegel*. New York: Oxford University Press, 1983.

Soll, Ivan. *An Introduction to Hegel's Metaphysics*. Chicago: University of Chicago Press, 1969.

Solomon, Robert C. *From Hegel to Existentialism*. New York: Oxford University Press, 1987.

Speight, Allen. *Hegel, Literature, and the Problem of Agency*. New York: Cambridge University Press, 2001.

Steinkraus, Warren E., and Kenneth I. Schmitz, eds. *Art and Logic in Hegel's Philosophy*. Atlantic Highlands, NJ: Humanities Press, 1980.

Stepelevich, Lawrence S., ed. *Selected Essays on G.W.F. Hegel*. Atlantic Highlands, NJ: Humanities Press, 1993.

Stern, Robert. “G. W. F. Hegel” In *An Introduction to Modern European Philosophy*. 2nd ed. Edited by Jenny Teichman and Graham White. New York: St. Martin's Press, 1995.

Stern, Robert, ed. *G. W. F. Hegel: Critical Assessments*. 4 vols. New York: Routledge, 1993.

Stern, Robert. *Hegel, Kant and the Structure of the Object*. London: Routledge, 1990.

Stern, Robert. *Routledge Philosophy Guidebook to Hegel and the Phenomenology of Spirit*. New York: Routledge, 2002.

Stewart, Jon, ed. *The Hegel Myths and Legends*. Evanston, IL: Northwestern University Press, 1996.

Stewart, Jon. ed. *The Phenomenology of Spirit Reader*. Albany: State University of New York Press, 1998.

Taylor, Charles. *Hegel*. 5th ed. New York: Cambridge University Press, 1987.

Taylor, Charles. *Hegel and Modern Society*. New York: Cambridge University Press, 1979.

Tristram, Engelhardt, Jr., and Terry Pinkard, eds. *Hegel Reconsidered: Beyond Metaphysics and the Authoritian State*. Boston: Kluwer Academic, 1994.

Tubbs, Nigel. *Contradiction of Enlightenment: Hegel and the Broken Middle*. Brookfield, VT: Ashgate, 1997.

Walker, John, ed. *Thought and Faith in the Philosophy of Hegel*. Boston: Kluwer Academic Publishers, 1991.

Weiss, Frederick Gustav, ed. *Beyond Epistemology: New Studies in the Philosophy of Hegel*. The Hague: Nijhoff, 1974.
Williams, Robert R. *Recognition: Fichte and Hegel on the Other*. Albany: State University of New York Press, 1992.
Wylleman, Luevan. *Hegel on the Ethical life, Religion, and Philosophy, 1793–1807*. Edited by A. Wylleman. Norwell: Kluwer Academic, 1989.

H. B. Action (1967)
Bibliography updated by Tamra Frei (2005)

HEGEL, GEORG WILHELM FRIEDRICH [ADDENDUM]

Georg Wilhelm Friedrich Hegel changed his major philosophical views very little from the publication of his first major work, the *Phenomenology of Spirit,* in 1807 until his death in 1831. This stability and continuity have not made it any easier for commentators to agree on what those views were. Disagreement about Hegel's basic position and its implications is still widespread, even more so after a great resurgence of Hegel studies after World War II.

In the Anglophone philosophical world, Hegel's position is still often summarized as an objective idealism, thanks largely to his influence on early twentieth-century British objective idealists such as Francis Herbert Bradley. He is said to have believed that only "mind" (the preferred translation of *Geist* until the A. V. Miller translation of *Phenomenology of Spirit* was first published in 1977) was "real"; or that no determinate individual object could be said to be real. Such an object was really a "moment" of the interrelated and temporally developing structure of the one true substance, the absolute, or absolute mind. Such a substance was said to develop over time; the nature of that development was a process of greater self-consciousness, and this development was reflected in, or the underlying basis of, the great social and political changes of world history, as well as intellectual changes in philosophy, art, and religion. Since Hegel appeared to have claimed a full and final "absolute knowledge," an "encyclopedic" account of such a structure, or the relation between "logic," philosophy of nature, and philosophy of spirit could be given. (A compelling demonstration that such an objective-idealist or "internal relations" view could not have been Hegel's position was published by the German Hegel scholar Rolf-Peter Horstmann in 1984, *Ontologie und Relationen.*)

Some aspects of such views of what Hegel really meant persist in many postwar interpretations but have not provoked much serious discussion or the interpretive variants that once characterized the work of John McTaggart Ellis McTaggart, G. R. G. Mure, Edward Caird, and Walter Terence Stace. Other interpretations and emphases have predominated. Many commentators have become interested in Hegel less as an object of purely historical research and more as a possible contributor to perennial and current philosophical controversies.

Charles Taylor's 1975 study, *Hegel,* while offering a comprehensive commentary on all aspects of Hegel's work, emphasized Hegel's insights into the emerging problems of the modern social and political world—problems such as social fragmentation, alienation, and the proper understanding of the modern goals of freedom and some sort of harmony with self. Taylor showed that many of Hegel's theoretical intentions could also best be understood against the backdrop of such concerns, and his approach became influential.

Hegel's understanding of the intellectual and social dimensions of modernization was also important in the work of many critical theory or Frankfurt school neo-Marxist philosophers (a group sometimes even designated as "neo-Hegelian" Marxists because of their attention to the social function of ideas and culture without a reliance on traditional Marxist versions of economic materialism). In the work of the most important "second generation" critical theorist, Jürgen Habermas, Hegel also plays a large role in what Habermas calls, in a book title, *The Philosophical Discourse of Modernity.* Hegel is called "the first philosopher who made modernity a problem" (p. 4)—this by raising many questions about the sufficiency of the modern notions of subjectivity and rationality.

In other developments, Klaus Hartmann in several influential articles proposed what he called a "nonmetaphysical" reading of Hegel, one that emphasized Hegel's category theory and the unusual "logic" of categorial relations, all as more or less autonomous philosophical problems, not necessarily wedded to any metaphysics of absolute mind. A group of German philosophers who came to be known as the Heidelberg school began to work in a more contemporary way on the single greatest problem that preoccupied the German idealists as a whole, and Hegel especially: the problem of self-consciousness, or "reflection," how the mind could be said to be both the subject of its own consciousness and object to itself at the same time. (The most important and influential work on this aspect of the idealists and Hegel in particular has been Dieter Henrich's.) Since, for many post-Kantian idealists, any possible cognitive or practical relation to the world was an active comporting of oneself

toward the world, or a "self-relation in relation to an other," the problem of self-relation was argued to be fundamental in any epistemology or account of human agency. These elements have also been emphasized by those who argue that Hegel should be read much more as a post-Kantian idealist, as much more decisively influenced by Immanuel Kant's founding arguments about the possibility of any self-conscious experience than by, say, Benedict de Spinoza or Friedrich von Schelling.

Hegel's contributions to all such problems—the nature and implications of modern social life, the possibility of self-consciousness and self-knowledge, the nature of the mind-world, and agency problems—reappear with great urgency in his ethical and social theory and in many interpretations. Debates about whether Hegel's 1821 *Elements of the Philosophy of Right* encouraged an accommodation of the conservative rulers of the Prussian state, or whether he was guilty of a kind of "organicist" anti-individualism, have been replaced by an emerging consensus that Hegel belongs within, if idiosyncratically, the modern liberal political tradition. This recognition has been somewhat complicated by "communitarian" writers and "traditionalist" writers suspicious of the modern reliance on claims of rationality as decisive in ethical life. Many such writers have occasionally enlisted arguments in a case against the classical liberal tradition. Hegel's position on the importance and "priority" of the ethical community in ethical life (*Sittlichkeit*) has sometimes been understood such that anyone who believes in the priority of prevolitional attachments or commitments in ethical deliberation (e.g., such attachments are necessary for deliberation to get started or have direction but cannot themselves be products of such deliberation) is labeled a neo-Hegelian. But Hegel believes that modern ethical life (the institutions and practices of modern social existence, the modern family, civil society, and the legal, constitutional state) are not just "ours" and "prior." He believes they are rational, raising the still much-debated question of how he distinguishes rational from nonrational ethical communities.

A great deal of scholarly work has been done in the postwar period on Hegel's texts, especially on the dating and organization of his Jena-period lecture materials. Karl-Heinz Ilting has compiled, edited, and published an extensive collection of Hegel's lecture notes on political philosophy, and a new critical edition of Hegel's works has begun to appear. New English translations of the *Phenomenology,* the *Logic,* the *Philosophy of Right,* the *Aesthetics* lectures, Hegel's letters, and many other works have also appeared.

See also Absolute, The; Alienation; Bradley, Francis Herbert; Caird, Edward; Communitarianism; Critical Theory; Freedom; German Philosophy; Habermas, Jürgen; Idealism; Kant, Immanuel; Liberalism; Marxist Philosophy; McTaggart, John McTaggart Ellis; Neo-Kantianism; Rationality; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de; Stace, Walter Terence; Subjectivity.

Bibliography

WORKS BY HEGEL

German

Gesammelte Werke. Edited by Rheinisch-Westfälischen Akademie der Wissenschaften. Hamburg: Meiner, 1968–. Of great value to serious students of Hegel in the original.

Philosophie des Rechts: Die Vorlesung von 1819/20. Edited by D. Henrich. Frankfurt: Suhrkamp, 1983.

Vorlesungen über Rechtsphilosophie 1818–31. Edited by K.-H. Ilting. Stuttgart: Frommann-Holzboog, 1974. Contains an invaluable collection of student notes and additions.

Werke in zwanzig Bänden. Frankfurt, 1971. Useful paperback version of the 1832–1845 Moldenhauer and Michel edition.

English

The Difference between Fichte's and Schelling's System of Philosophy. Translated by H. S. Harris and W. Cerf. Albany: State University of New York Press, 1977.

Elements of the Philosophy of Right. Edited by A. Wood; translated by H. B. Nisbet. Cambridge, U.K.: Cambridge University Press, 1991.

Faith and Knowledge. Translated by W. Cerf and H. S. Harris. Albany: State University of New York Press, 1977.

Hegel: The Letters. Translated by C. Butler and C. Seller. Bloomington: Indiana University Press, 1984.

Hegel's Aesthetics: Lectures on Fine Art. Translated by T. M. Knox. Oxford, 1965.

Lectures on the Philosophy of Religion: The Lectures of 1827. Edited by P. Hodgson; translated by R. F. Brown, P. C. Hodgson, J. M. Stewart, and H. S. Harris. Berkeley: University of California Press, 1984.

Phenomenology of Spirit. Translated by A. V. Miller. Oxford: Clarendon Press, 1977.

Science of Logic. Translated by A. V. Miller. London, 1969.

WORKS ON HEGEL

Hegel's Intellectual Development

Dickey, L. *Hegel: Religion, Economics, and the Politics of the Spirit.* Cambridge, U.K.: Cambridge University Press, 1977.

Görland, I. *Die Kantkritik des jungen Hegel.* Frankfurt: Klostermann, 1966.

Harris, H. S. *Hegel's Development.* 2 vols. Oxford: Clarendon Press, 1972–1983.

Henrich, D., and K. Düsing, eds. *Hegel in Jena.* Bonn: Bouvier, 1980.

Kimmerle, H. *Das Problem der Abgeschlossenheit des Denkens: Hegels System der Philosophie in den Jahren 1800–1804.* Bonn: Bouvier, 1970.

Peperzak, A. *La jeune Hegel et la vision morale du monde.* The Hague: Nijhoff, 1960.

Hegel's Basic Position and His System

Adorno, T. *Hegel: Three Studies.* Translated by S. W. Nicholsen. Cambridge, MA: MIT Press, 1993.

Beiser, F., ed. *The Cambridge Companion to Hegel.* Cambridge, U.K.: Cambridge University Press, 1993. Contains a comprehensive bibliography of recent work on Hegel.

Brockard, H. *Subjekt: Versuch zur Ontologie bei Hegel.* Munich: Pustet, 1970.

DeVries, W. *Hegel's Theory of Mental Activity.* Ithaca, NY: Cornell University Press, 1988.

Fackenheim, E. *The Religious Dimension in Hegel's Thought.* Chicago, 1967.

Harlander, K. *Absolute Subjektivität und kategoriale Anschauung.* Meisenheim: Hain, 1969.

Hartmann, K. *Die ontologische Option.* Berlin: de Gruyter, 1976.

Henrich, D. *Hegel im Kontext.* Frankfurt: Suhrkamp, 1971.

Hösle, V. *Hegels System.* 2 vols. Hamburg: Meiner, 1987. Includes a comprehensive bibliography of recent work on Hegel.

Houlgate, S. *Freedom, Truth, and History: An Introduction to Hegel's Philosophy.* London: Routledge, 1991.

Inwood, M. J. *Hegel.* London: Routledge and Kegan Paul, 1983.

Inwood, M. J., ed. *Hegel.* Oxford: Oxford University Press, 1985. Contains a comprehensive bibliography of recent work on Hegel.

MacIntrye, A., ed. *Hegel: A Collection of Critical Essays.* Notre Dame, IN, 1972.

McCumber, J. *The Company of Words: Hegel, Language, and Systematic Philosophy.* Evanston, IL: Northwestern University Press, 1993.

O'Brien, G. D. *Hegel on Reason and History.* Chicago: University of Chicago Press, 1975.

Peperzak, A. *Selbsterkenntnis des Absoluten: Grundlinien der Hegelsche Philosophic des Geistes.* Stuttgart: Frommann-Holzboog, 1987.

Pinkard, T. *Hegel's Dialectic: The Explanation of Possibility.* Philadelphia: Temple University Press, 1988.

Pippin, R. B. *Hegel's Idealism: The Satisfactions of Self-Consciousness.* Cambridge, U.K.: Cambridge University Press, 1989. Includes a comprehensive bibliography of more recent works on Hegel.

Plant, R. *Hegel.* Bloomington: Indiana University Press, 1973.

Rockmore, T. *Hegel's Circular Epistemology.* Bloomington: Indiana University Press, 1986.

Rose, G. *Hegel Contra Sociology.* Atlantic Highlands, NJ: Humanities Press, 1981.

Rosen, S. *G. W. F. Hegel: An Introduction to the Science of Wisdom.* New Haven, CT: Yale University Press, 1974.

Soll, I. *An Introduction to Hegel's Metaphysics.* Chicago: University of Chicago Press, 1969.

Taylor, C. *Hegel.* Cambridge, U.K.: Cambridge University Press, 1975.

White, A. *Absolute Knowledge: Hegel and the Problem of Metaphysics.* Athens: Ohio University Press, 1983.

Wohlfahrt, G. *Der spekulative Satz: Bemerkungen zum Begriff der Spekulation bei Hegel.* Berlin: de Gruyter, 1981.

Hegel's Phenomenology and Logic

Baum, M. *Die Entstehung der Hegelschen Dialektik.* Bonn: Bouvier, 1986.

Bubner, R. *Dialektik und Wissenschaft.* Frankfurt, 1962.

Düsing, K. *Das Problem der Subjektivität in Hegels Logik.* Bonn: Bouvier, 1976.

Flay, J. *Hegel's Quest for Certainty.* Albany: State University of New York Press, 1984. Includes a comprehensive bibliography of recent work on Hegel.

Forster, M. *Hegel and Skepticism.* Cambridge, MA: Harvard University Press, 1989.

Fulda, H. *Hegels Wissenschaft der Logik: Formation und Rekonstruktion.* Stuttgart: Klett-Cotta, 1986.

Fulda, H., ed. *Materialen zu Hegels Phänomenologie des Geistes.* Frankfurt: Suhrkamp, 1973.

Fulda, H. *Das Problem einer Einleitung in Hegels Wissenschaft der Logik.* Frankfurt: Klostermann, 1965.

Henrich, D. *Die Wissenschaft der Logik und die Logik der Reflexion.* Bonn: Bouvier, 1987.

Horstmann, R.-P. *Ontologie und Relationen: Hegel, Bradley, Russell, und die Kontroverse über interne und externe Relationen.* Königstein: Hain, 1984.

Labarriere, P. *La phénoménologie de l'esprit de Hegel.* Paris: Aubier-Montaigne, 1979.

Marx, W. *Hegel's Phenomenology of Spirit: Its Point and Purpose—A Commentary on the Preface and Introduction.* New York: Harper and Row, 1975.

Pinkard, T. *Hegel's Phenomenology: The Sociality of Reason.* Cambridge, U.K.: Cambridge University Press, 1994.

Pöggeler, O. *Hegels Idee einer Phänomenologie des Geistes.* Freiburg: Alber, 1973.

Puntel, B. *Darstellung, Methode, und Struktur: Untersuchungen zur Einheit der systematischen Philosophic G. W. F. Hegels.* Bonn: Bouvier-Verlag Grundmann, 1973.

Robinson, J. *Duty and Hypocrisy in Hegel's Phenomenology of Mind.* Toronto: University of Toronto Press, 1977.

Rohs, P. *Form und Grund: Interpretation eines Kapital der Hegelschen Wissenschaft der Logik.* Bonn: Bouvier Verlag Grundmann, 1969.

Rosen, M. *Hegel's Dialectic and Its Criticism.* Cambridge, U.K.: Cambridge University Press, 1982.

Shklar, J. *Freedom and Independence: A Study of the Political Ideas of Hegel's Phenomenology of Mind.* Cambridge, U.K.: Cambridge University Press, 1976.

Solomon, R. *In the Spirit of Hegel.* Oxford: Oxford University Press, 1983.

Theunissen, M. *Sein und Schein: Die kritische Funktion der Hegelschen Logik.* Frankfurt: Suhrkamp, 1978.

Westphal, M. *History and Truth in Hegel's Phenomenology.* Atlantic Highlands, NJ: Humanities Press, 1978.

Hegel's Practical Philosophy

Avineri, S. *Hegel's Theory of the Modern State.* Cambridge, U.K.: Cambridge University Press, 1972.

Brod, H. *Hegel's Philosophy of Politics: Idealism, Identity, and Modernity.* Boulder, CO: Westview Press, 1992.

Cullen, B. *Hegel's Social and Political Thought: An Introduction.* Dublin, 1979.

D'Hondt, J. *Hegel in His Time: Berlin 1818–1831.* Translated by J. Burbidge. Peterborough, ON: Broadview Press, 1988.

Fessard, G. *Hegel, le Christianisme, et l'histoire.* Paris: Presses Universitaires de France, 1990.

Fleischman, E. *La philosophie politique de Hegel.* Paris, 1964.

Hardimon, M. *Hegel's Social Philosophy: The Project of Reconciliation.* Cambridge, U.K.: Cambridge University Press, 1994.

Henrich, D., and R.-P. Horstmann. *Hegels Philosophie des Rechts: Die Theorie der Rechstformen und ihre Logik.* Stuttgart: Klett-Cotta, 1982.

Marcuse, H. *Reason and Revolution.* Boston: Beacon Press, 1960.

Pelczynski, Z. A., ed. *Hegel's Political Philosophy: Problems and Perspectives.* Cambridge, U.K.: Cambridge University Press, 1971.

Pelczynski, Z. A. *The State and Civil Society.* Cambridge, U.K.: Cambridge University Press, 1984.

Peperzak, A. *Hegel's praktische Philosophie: Ein Kommentar zur enzyklopädischen Dartstellung der menschlichen Freiheit und ihre objektive Verwirklichung.* Stuttgart: Frommann-Holzboog, 1991.

Peperzak, A. *Philosophy and Politics: A Commentary on the Preface to Hegel's Philosophy of Right.* The Hague, 1987.

Pinkard, T. *Democratic Liberalism and Social Union.* Philadelphia: Temple University Press, 1987.

Riedel, M. *Materialen zu Hegels Rechtsphilosophie.* 2 vols. Frankfurt: Suhrkamp, 1975.

Riedel, M. *Theorie und Praxis im Denken Hegels.* Frankfurt: Kohlhammer, 1965.

Ritter, J. *Hegel and the French Revolution.* Translated by R. D. Winfield. Cambridge, MA: MIT Press, 1982.

Siep, L. *Hegels Fichtekritik und die Wissenschaftslehre von 1804.* Freiburg: Alber, 1970.

Siep, L. *Praktische Philosophie im Deutschen Idealismus.* Frankfurt: Suhrkamp, 1992.

Smith, S. *Hegel's Critique of Liberalism.* Chicago: University of Chicago Press, 1989.

Stepelevich, L. S., and D. Lamb. *Hegel's Philosophy of Action.* Atlantic Highlands, NJ: Humanities Press, 1983.

Taylor, C. *Hegel and Modern Society.* Cambridge, U.K.: Cambridge University Press, 1979.

Theunissen, M. *Hegels Lehre vom absoluten Geist als theologisch-politischer Traktat.* Berlin: de Gruyter, 1970.

Tunick, M. *Hegel's Political Philosophy: Interpreting the Practice of Legal Punishment.* Princeton, NJ: Princeton University Press, 1992.

Walsh, W. H. *Hegelian Ethics.* London: Macmillan, 1969.

Wasczek, N. *The Scottish Enlightenment and Hegel's Account of Civil Society.* Dordrecht: Kluwer Academic, 1988.

Wildt, A. *Autonomie und Anerkennung: Hegels Moralitätskritik in Lichte seiner Fichte Rezeption.* Stuttgart: Klett-Cotta, 1982.

Wood, A. *Hegel's Ethical Thought.* Cambridge, U.K.: Cambridge University Press, 1990.

Aesthetics

Bungay, S. *Beauty and Truth: A Study of Hegel's Aesthetics.* Oxford: Oxford University Press, 1986.

Desmond, W. *Art and the Absolute: A Study of Hegel's Aesthetics.* Albany: State University of New York Press, 1986.

Fulda, H., ed. *Hegel und die Kritik der Urteilskraft.* Stuttgart: Klett-Cotta, 1990.

Robert B. Pippin (1996)

HEGELIANISM

The term *Hegelianism* is applied to a range of philosophical doctrines and traditions influenced by the nineteenth century German philosopher Georg Wilhelm Friedrich Hegel (1770–1831). Hegel's influence is as broad and diverse as his writings; moreover his legacy, like his philosophy, is characterized by tensions between dialectical opposites.

A central part of Hegel's philosophical reputation has always been in metaphysics, where he is seen as the leading proponent of absolute idealism: the thesis that reality as a whole—nature, humanity, history, and so on—is informed and shaped by (and indeed ultimately is a manifestation of) what Hegel famously called *Geist*: mind or spirit. For Hegel, Geist is both rational and rationally comprehensible, whether in logical structure, natural science, or historical progress. Hegel also held that Geist itself exhibits a distinctive self-consciousness or self-articulation, and that the manifestations of this self-consciousness can be found in psychology, history, religion, drama, art, and philosophy.

The earliest Hegelian movement comprised a core of adherents working to vindicate these claims in a diverse range of intellectual projects seeking to identify and exhibit the promised rational structure. Characteristic of the Hegelian position is the claim that rational structure is historical and dialectical: The rational structure of the real is not a static and self-consistent body of facts, but a dynamic process unfolding through the systematic resolution of dialectical contradictions. The earliest Hegelians sought to identify this structure in the areas of law, history, politics, and natural science. Among the first Hegelians were the members of the Society for Scientific Criticism, formed in July 1826 in Hegel's own home, and closely associated with the *Jarhbücher für wissenschaftliche Kritik*, a journal devoted to the dissemination and application of Hegelian ideas.

Prominent members of this old Hegelian movement included Leopold von Henning (*Prinzipien der Ethik in historischer Entwicklung*, 1824), who applied Hegelian ideas in ethics; and Eduard Gans (*Das Erbrecht in weltgeschichtlicher Entwicklung*, 1924–35), who was primarily concerned with issues in law and jurisprudence. The hallmark of this early Hegelianism was the emphasis placed

on historical approaches to traditional philosophical issues, together with the attempt to provide a critical justification of cultural configurations as the outcome of rational dialectical progressions. Hegelianism thus played an important role in the emergence of history as the central category and discipline of the developing human sciences of the nineteenth century.

Many of these early contributions were soon lost to memory as Hegel's influence waned in Germany after his death in 1831. In the case of theology and religious studies, however, self-professed Hegelians achieved a lasting prominence. Early Hegelian theologians also provided the first example of a pattern that was repeated later: the emergence of two traditions, each drawing explicitly on Hegelian teachings, yet developing those ideas in sharply opposed directions and soon coming into noisy confrontations, which themselves became points of reference in subsequent philosophical developments. Hegel's own theological position was that orthodox Judeo-Christian religious teachings were true yet ultimately inadequate articulations of ideas expressed more fully and exactly in the language of metaphysics. Hence in the Christian doctrines of the creation, divine incarnation, the Trinity, and human immortality, for instance, Hegel finds vivid thought-images (*Vorstellungen*) of rational concepts (*Begriffe*) fully graspable only in his idealistic metaphysics. On this model theism becomes a penultimate articulation of absolute idealism; the hope for immortality is satisfied in an individual's identification with a trans-individual cultural whole that survives the death of its constituents; and the doctrine of the Trinity is seen as a vivid representation of a metaphysical truth graspable only by a dialectical logic. The central Hegelian notion of *Aufhebung* (variously and inadequately translated as sublation, supercession, and so on) finds its exemplar in the Christian claim to complete the Hebraic law by negating it.

The tension in this position is manifest. The orthodox articles of faith are true and yet not true: true insofar as they articulate a thought that merits our assent, yet not true because cast in a language incapable of adequate articulation of the insights they express. The earliest conflicts in the Hegelian school emerged among theologians exploring the two opposed sides of this Hegelian contradiction. Philip Marheinecke (*Die Grundlehren der christlichen Dogmatik als Wissenshchaft*, 1827) and Karl Daub (*Die dogmatische Theologie jetziger Zeit*, 1833) defended and elaborated a Hegelian orthodoxy, and advanced the claim of Hegelianism to provide a middle way between an extra-rational fideism and the extreme atheism that had long been feared as the final outcome of Enlightenment rationalism. But their work was soon eclipsed by the more radical approach inaugurated by David Friedrich Strauss, whose influential and controversial *Leben Jesu, kritisch bearbeitet* (1835–36) proposed a radically revisionist and explicitly critical reading of religious texts (particularly the Gospels) and proved to be a watershed in the secularization of religious studies.

It was this split among the Hegelians that first gave rise to the language of left and right Hegelians—a description Strauss himself introduced in 1837 (*Streitschriften zur Vertidigung meiner Schrift*). The left Hegelian treatment of theological issues reached its culmination in the work of Ludwig Feuerbach. His *Thoughts on Death and Immortality* (1830) argued explicitly against the idea of personal immortality; his *Essence of Christianity* (1841) was directed forthrightly against religion, and inaugurated the influential nineteenth century movement that reinterpreted religious teachings in psychological and political terms. An 1853 English translation by George Eliot played a focal role in the nineteenth century reassessment of religion in British intellectual circles.

The terms left and right derived from the political rhetoric of the day (the division between the two sides of the French Chamber of Deputies) and it was quickly to be applied once again in the struggle over Hegel's legacy in political philosophy. This is an area where, once again, we find important tensions in Hegel's own position. Hegel's political theory was essentially an application of his social theory: He saw in human society an unfolding attempt to develop institutions that were rational and just, and capable of sustaining an unfettered critical self-examination. For Hegel, society provides not only the material fruits of social cooperation (self-defense and the power of cooperative labor), but also what Hegel called *Anerkennung* or acknowledgement: the mutual recognition by citizens as free and rational self-determining agents.

Hegel held that the basic demands of the just state were met adequately only in the modern era, in particular with the emergence of the modern sociopolitical institutions of the modern family, the free market, the republic, and private property, themselves resting on a guarantee of civil liberty. In political history Hegelians thus set out to trace the emergence of these modern rational institutions; in politics Hegelians defended a monarchist liberalism, and some (notably Karl Rosenkranz) even served in parliamentary assemblies. But the Hegelian framework was once again put to very different uses when it came to its application to concrete issues in politics and political economy.

Rosenkranz and others sought to use Hegelian ideas to justify and reform the major institutions of modern political life, particularly in mid-nineteenth century Prussia. In explicit rivalry to these old or right Hegelians there emerged what have come to be known as the young or left Hegelians, chief among them Feuerbach, Bruno Bauer, and the young Karl Marx, who sought to use Hegelian criticism as a device for advancing radical social change. The tension between these two movements brought out the opposed tendencies in Hegel's position, which on the one side was interpreted as a philosophy of reconciliation, and specifically as a justification of status quo Prussian institutions, whereas on the other inspiring what the influential Polish Hegelian, August von Cieszkowski, called a philosophy of action (*Prolegomena zur historiosophie*). The left Hegelians abandoned the claim that Hegelianism could vindicate orthodox religion and politics and instead turned to apply Hegelian social theory to provide orientation in the struggle for social change.

In the context of this dispute between left and right appropriations of Hegel, many features of Hegelian philosophy were appropriated and applied in abstraction from Hegel's own distinctive metaphysical commitments. The most dramatic example of this came with Marx, who sharply rejected Hegel's idealistic metaphysics and theory of sociohistorical development. Where Hegel placed primary emphasis and found the root of historical change in what he called simply the *concept* or the *notion* (*der Begriff*) that advanced toward self-completion in human history, Marx advanced a materialism that found the prime mover of history in the material conditions of human existence and the satisfaction of material economic needs.

But through this metaphysical reversal, important elements of the Hegelian position remained, notably Hegel's treatment of history as the overcoming of dialectical contradictions (tensions between opposed principles in historical configurations), in the ideal of a final resolution of these tensions and thus a certain kind of end of history, and in his appropriation and development of the Hegelian notion of alienation. In his early contributions to the *Deutsch-französischen Jahrbücher*, the young Marx argued that the social, economic, and political structures of human history have become alien powers that tyrannize human beings. Hence it was not enough to call, as Bauer and Feuerbach had, for the end of religion because religion is not the cause but the expression of self-alienated man. One must rather attack the real material conditions that create and sustain this condition.

The tension between left and right Hegelianism proved to be one of the most enduring aspects of Hegel's legacy, and it is a debate that has played out in a diverse range of historical, political, and theoretical contexts, of which the nineteenth century socialist movement is only the best known example. In Russia the appeal to Hegelian social theory figured prominently in the mid-nineteenth century debate over feudal institutions and Russian national identity, particularly in the works of Nikolai Stankevich, Vissarion Belinskii, and in the activism of Mikhail Bakunin, whose later fame as the leading figure in European anarchism followed an earlier period of intense engagement with Hegelian ideas. In more recent times Hegelian philosophy of history was invoked in the attempt to find new political orientation after the collapse of communism and the end of the cold war at the end of the twentieth century (Fukuyama 1992).

Hegel's influence was also strong in the emergence and development of the existentialist tradition. A number of the themes central to existentialism had been explored in Hegel's writings: the themes of death and immortality, alienation, nihilism, and so on. Already in the writings of S.A. Kierkegaard (*Concluding Unscientific Postscript to Philosophical Fragments*, for example), existentialist thinkers had defined themselves in relation to Hegel, albeit in that case in the form of an insistent negation. Kierkegaard was relentlessly critical of Hegelian rationalism, and ridiculed Hegel's claim to provide a rational reconstruction of religion that could eliminate the absurdity that Kierkegaard himself found both in religious consciousness and in the human condition. But later French existentialists, notably Alexandre Kojéve, Jean Hyppolite, and Jean-Paul Sartre, developed much more sympathetic appropriations, shifting the focus from Hegel's mature writings (*The Encyclopaedia of the Philosophical Sciences* and *The Philosophy of Right*) to the much earlier *The Phenomenology of Spirit*.

For Hegel, phenomenology was to be a study of structures of self-consciously lived experience as manifestations of Geist. Hegel's *Phenomenology* comprised a series of case studies that exercised great influence on the existentialists, most importantly in connection with the account of the confrontation with death in the dialectic of master and slave, which Kojéve (1934) in particular made central to his reading of Hegel. Kojéve followed Hegel in arguing that the encounter with an *other* was both an essential moment in the structure of autonomous self-consciousness and yet, at the same time, an essentially destabilizing confrontation with which human beings and human institutions must ultimately come to terms.

Thus while the existentialists sharply diverged from the orthodox Hegelian metaphysical position and sharply rejected Hegelian rationalism, their philosophical practice—focusing in particular on the unfolding, essentially narrative structures of self-conscious experience and its intrinsic tensions—followed no thinker more closely than Hegel. Sartre's *Being and Nothingness* (1943) provides the fullest development of this strand of existentialism, and in many passages closely imitates its Hegelian model.

In nineteenth century Britain and America, by contrast, it was metaphysics that was preeminent in Hegel's legacy. Hegel's influence in the English speaking world began with J.H. Stirling's influential study, *The Secret of Hegel* (1865). Stirling's work itself mainly took the form of a sympathetic but somewhat superficial synopsis of Hegelian texts and doctrines, but it nonetheless proved influential in forging a generation of British idealists. As the British idealist movement matured, it grew increasingly independent of its German models, and its leading figures—T.H. Green and F.H. Bradley—developed independent philosophical systems of considerable originality. Nonetheless, important traces of the Hegelian origins persist, particularly in Bradley's strategy of arguing for his monistic idealism by exhibiting the systematic contradictions hidden in the common sense assumptions of reality as plural (comprised of ontologically distinct individuals), empirically knowable and mind-independent.

Later figures in the British idealist movement included Bernard Bosanquet, who contributed greatly to the propagation and interpretation of Hegel's own philosophical writings, and John McTaggart, whose most influential legacy derived from his antirealism about time. When the tradition of logical analysis emerged in Britain at the beginning of the twentieth century, it began with a systematic critique of this dominant idealist orthodoxy, most notably in the early writings of G.E. Moore and Bertrand Russell (see Peter Hylton's *Russell, Idealism, and the Emergence of Analytic Philosophy* [1990]).

Hegelianism was also influential in the emerging philosophical traditions in America. Particularly in areas with strong German immigrant traditions, Hegelian schools thrived, notably in Cincinnati (John Bernard Stallo, August Willich, Moncure Conway), and in St Louis, where Henry Brokmeyer and William Harris formed *The Philosophical Society*, which was explicitly Hegelian in its orientation and sponsored the *Journal of Speculative Philosophy*, an influential journal on the late nineteenth century American philosophical scene. These early American Hegelians included not only academic philosophers but influential significant civic leaders who sought to apply Hegelian ideas in social and educational reforms, and appealed to Hegelian philosophy of history in coming to terms with the upheaval of the American Civil War. Indeed prior to the emergence of the pragmatist tradition, Hegelianism was arguably the most well-defined school of American philosophy.

Among the most influential American Hegelians was California native and Harvard philosopher, Josiah Royce, who had visited Germany as a student and returned to Harvard as a key conduit of Hegelian ideas. In *The Religious Aspect of Philosophy* (1885) Royce argued for a version of absolute idealism, proposing and defending a Hegelian theism, with the existence of God understood in Hegelian terms as a super-individual subject in which finite subjects figure as moments of an overarching organic totality. Religious and proto-existentialist themes dominated Royce's appropriation of Hegel, and his *Lectures on Modern Idealism* (1919) remained an influential introduction to idealistic philosophy well into the twentieth century. Although Royce's idealism was soon eclipsed in the American academy by the budding pragmatist movement, he set an important precedent for a number of later North American Hegelians (recent examples include Charles Taylor [1975] and Robert Pippin [1991]) who have looked to Hegelian philosophy to answer the charge that modern cultural forms lead inexorably toward a crisis of faith and nihilistic despair. In this respect these more recent North American Hegelians can be seen as developing a more secularized version of Royce's idealism.

Hegel's legacy in logic has been complex and somewhat diffuse. Hegel himself held logic to occupy a fundamental place in philosophical inquiry, providing not simply a theory or mechanism for inference but rather an articulation of the underlying rational structure of all reality. In this sense he can be understood as the leading advocate of a material (as opposed to a merely formal) construal of logic. His logic was also distinctive in its dialectical structure, taking the form of a series of dialectical transformations intended to unpack all the basic categories of the real from tensions inherent in the bare concept of being. But despite a few early adherents and defenders (Kuno Fischer being the most important), relatively few have followed Hegel's lead in these views, and the nineteenth century reform of logic grew much more out of Kantian themes and problems than from recognizably Hegelian doctrines.

Nonetheless, Hegel's influence in logic has been felt indirectly, particularly in connection with his account of

conceptual determinacy. On an orthodox empiricist construal, concepts receive their determinate content in virtue of a connection with some nonconceptual content of experience. But this empiricist doctrine has continually come under attack, both in the nineteenth century neo-Kantian movement associated with Hermann Cohen and Paul Natorp, and again more recently in the work of the seminal twentieth century American philosopher, Wilfrid Sellars. In looking for alternatives to the traditional empiricist doctrine, logicians and semantic theorists have repeatedly been drawn to Hegelian themes. Thus, for instance, Tyler Burge's "Individualism and the Mental" (1979) introduced his defense of social externalism with an invocation of Hegel, and recent semantic theorists such as John McDowell and Robert Brandom have explicitly turned to Hegelian themes for an alternative to the empiricist account of some foundational given content.

McDowell's *Mind and World* (1994) develops an essentially Hegelian thesis in arguing that the conceptual content of experience must reach all the way down to its most primitive content and indeed must ultimately be seen as reflective of conceptual structure inherent in the world itself. Brandom's Hegelianism in *Making it Explicit* (1994) is more nuanced and complex, but his inferentialist semantics adapts a recognizably Hegelian theme in arguing that conceptual determinacy must be traced to the inferential role played by concepts in an essentially social and pragmatic context of demanding and providing reasons. What is common to these approaches is the conviction that semantic or conceptual content is fixed holistically and (particularly in Brandom's account) in the context of unfolding social interactions. Although these accounts are quite distant from Hegel's ambitions for logic, they retain an essentially Hegelian logical moment in finding an essential appeal to a collective, diachronic social background in the fixing of even the most elementary concepts.

Hegel's influence on twentieth century continental European philosophy has been pervasive, and has taken many different forms. The work of Wilhelm Dilthey was seminal in this connection, in part because of his work in recovering Hegel's early theological writings and writing his biography, but mainly because Dilthey's own influential approach to the philosophy of the human sciences owed much to Hegel in arguing for the centrality of narrative, biography, and history in what Dilthey himself called the *Geisteswissenschaften* (literally, the sciences of Geist—the human sciences such as psychology, anthropology, jurisprudence, and so on).

In the twentieth century, the early works of Herbert Marcuse (1932 and 1960) sought to adapt Hegelian ideas to a new historical and cultural circumstance, combining a broadly Hegelian conception of history with elements of Martin Heidegger's existentialism. The Hegelian-Marxist conception of history as driven by dialectical tension found new voice in the writings of Max Horkheimer, Theodor Adorno, and other members of the Frankfurt School, albeit in this case without the Marxist and Hegelian optimism regarding the final resolution of such contradictions. And a broad array of thinkers followed a Hegelian lead in locating objectivity in configurations of intersubjective consensus. Rather than contrasting objective truth and subjective illusion, as had been common in the tradition stretching from Galileo to Immanuel Kant, these thinkers (including figures as diverse as Edmund Husserl and Jürgen Habermas) sought to reinterpret the notion of objective truth in terms of an ideal of a normative intersubjective consensus. Although the hints of this theory of objectivity can be traced back to Kant's aesthetics, it is perhaps the most pervasive legacy of Hegel's attempt to think through "the I that is we and the we that is I." Among contemporary European thinkers, Axel Honneth (1992) presents perhaps the clearest case of this Hegelian legacy in political theory.

Finally, Hegel's legacy can be found at work—diffuse but unmistakable—in the standing that the history of philosophy has acquired in the past two centuries. More than any other prominent philosopher since Aristotle, Hegel's philosophical practice was directly related to his appropriation of the history of his discipline. But where Aristotle's writings systematically surveyed the opinions of his predecessors, Hegel claimed to find in philosophy's history both a systematic order and the elements for his own philosophical synthesis. It is now a commonplace—albeit a commonplace that is sometimes challenged—to see the history of philosophy as directly relevant to philosophical inquiry generally. The emergence of this view, as of the conviction that philosophy is essentially unlike the natural sciences in this regard, can be traced to Hegel, who can without exaggeration be credited with inventing the very discipline of the history of philosophy. For Hegel, the history of philosophy is not merely the history of ideas; it is an attempt to reread and rethink the history of attempts to tackle philosophical questions. Its aim is ultimately not historical but philosophical: to uncover a rational order that will itself illuminate those questions themselves.

See also Hegel, Georg Wilhelm Friedrich.

Bibliography

Cieszkowski, August von. *Prolegomena zur historiosophie* (1838). Hamburg: Meiner, 1981.

Fukuyama, Francis. *The End of History and the Last Man*. New York: Free Press, 1992.

Green, T. H. *Prolegomena to Ethics* (1883). New York: Oxford University Press, 2003.

Honneth, Axel. *The Struggle for Recognition: The Moral Grammar of Social Conflicts* (1992). Translated by Joel Anderson. Cambridge, MA: MIT Press, 1996.

Hylton, Peter. *Russell, Idealism, and the Emergence of Analytic Philosophy*. New York: Oxford University Press, 1990.

Hyppolite, Jean. *The Genesis and Structure of Hegel's Phenomenology of Spirit* (1946). Translated by John Heckman. Evanston, IL: Northwestern University Press, 1974.

Kojeve, Alexandre. *Introduction to the Reading of Hegel* (1934). Translated by James Nichols. New York: Basic Books, 1969.

Marcuse, Herbert. *Hegel's Ontology and the Theory of Historicity* (1932). Translated by Selya Benhabib. Cambridge, MA: MIT Press, 1987.

Marcuse, Herbert. *Reason and Revolution: Hegel and the Rise of Social Theory*. Boston: Beacon Press, 1960.

Pippin, Robert. *Modernism as a Philosophical Problem: On the Dissatisfactions of European High Culture*. Oxford: Blackwell, 1991.

Royce, Josiah. *Lectures on Modern Idealism*. New Haven, CT: Yale University Press, 1919.

Stirling, James Hutchison. *The Secret of Hegel: Being the Hegelian System in Origin, Principle, Form, and Matter* (1865). Bristol: Thoemmes, 1990.

Strauss, David Friedrich. *The Life of Jesus, Critically Examined* (1835–36). 4th ed. Translated by George Eliot. Philadelphia: Fortress Press, 1972.

Taylor, Charles. *Hegel*. Cambridge, U.K.: Cambridge University Press, 1975.

Toews, John Edward. *Hegelianism: The Path toward Dialectical Humanism, 1805–1841*. Cambridge, U.K.: Cambridge University Press, 1980.

Wayne M. Martin (2005)

HEIDEGGER, MARTIN

(1889–1976)

Martin Heidegger (1889–1976) was born in Messkirch, a small town in the hills of southwestern Germany. The environment of his modest, middle class upbringing was that of a Catholic agrarian village where his father was the sexton of the local church. When Heidegger was fourteen he entered the Catholic seminary at Constance and began an education that appeared to be directed toward a vocation in the priesthood. He entered a novitiate with the Jesuits in 1909 but left that track after a short time and shifted out of clerical training altogether in 1911. He intensified his studies in philosophy, literature, and science, and for a time concentrated on mathematics. During this period (through 1915) he developed a conservative approach to neo-scholastic thought and published articles in conservative Catholic journals. He also read intensely the emerging phenomenological literature and neo-Kantian philosophy.

Heidegger's doctoral dissertation in 1913 had the title, "The Doctrine of Judgment in Psychologism: A Critical-Positive Contribution to Logic." He completed his habilitation dissertation in 1915, "Duns Scotus's Doctrine of Categories and Meaning."

With the emergence of a strong interest in historical development and in Edmund Husserl's thought, a significant counterforce to his Christian, transcendentally oriented convictions began to form. On the one hand he understood the basic structures of truth and meaning to have changeless validity. On the other, he saw that an act of mind requires time for syntheses and connections and that philosophical thought bears describable histories within it. Although academics identified him as a rising Catholic thinker, he was increasingly influenced by G. F. W. Hegel's historical, dialectical thought as well as by the "life-philosophy" of Friedrich Nietzsche, Wilhelm Dilthey, Henri Bergson, and Max Scheler. Søren Kierkegaard's and Martin Luther's writings also had a strong effect on his thought. The transcendental orientations of Thomistic, Husserlian, and neo-Kantian philosophy were increasingly challenged in Heidegger's thought as he devoured the art and philosophy—both religious and nonreligious—that influenced his moods and feelings as well as his thinking at least as powerfully as rational argument influenced them.

Heidegger's attention turned increasingly to issues of time, history, suffering, and unresolvable ambiguity. The regions of pure logic and transcendentally oriented morality and epistemology began to appear to him as desertlike and abstract. Metaphysical thought, if it is to count as important, must give clarity to and insight into lives and histories. Issues connecting phenomenology with time, history, and life formed a new horizon for the young philosopher. Whereas in his 1913 dissertation atemporal logic and its categories provided the way to understanding being, by 1915 the *question* of being, not being's static availability for conceptual grasp, began to take shape. Hegel and especially Husserl began to emerge as major transitional figures as Heidegger moved away from Thomism and neo-Kantian philosophy and toward a phenomenological approach that valorized description over speculation and practical life over categorical analysis.

When he married Elfride Petri in 1917 his departure from Catholicism, which became explicit in 1919, was well underway. Heidegger was in the process of a turn the momentum of which helped to define both his creativity and the movement of his thinking.

In 1918 he became, as a Privatdozent, an assistant to Husserl in Freiburg. The University of Marburg appointed him associate professor in 1923, and in 1928 he succeeded Husserl as Professor at Freiburg.

EARLY THOUGHT

Two of Heidegger's early insights are that thought takes place only within the particularities of cultural and communal lives and that particular lives are saturated with histories. This emphasis on temporal, historical particularity means that he began to place a primary importance on the situatedness of thought in the history of philosophy. This emphasis is particularly noteworthy because the historical emphasis added a dimension to phenomenological thought that was not clearly pronounced in Husserl's work and because it showed the particularity of Heidegger's own way: his early and deep engagement with ancient and medieval texts and the personal import of his traditional, historically oriented education. Even though he turned away from metaphysical theology, he did not turn away from the central importance for thought of the metaphysical tradition. It provided the site for philosophical transformation and departure.

Heidegger's recognition of the importance of temporal particularity for thought set in motion a conflict of values that would help to shape his thought for over a decade. He launched a task of learning to think with the traditions that formed his particularity in such a way that he could turn through their senses of timelessness by the temporal movement of his own thought: The temporal dimension of his thought began to define the meaning of claims to timelessness. To carry out such a project he needed to work through metaphysical thought, finding in it what overturns its predisposition toward unchanging truths, formal logic, and the priority of the knowing subject. "Temporal-historical occurrence" names the overturning element. "Phenomenology" names the approach by which Heidegger formulated the transforming power of time in European sensibility. Together, temporal-historical occurrence and phenomenology provided Heidegger with the elements that allowed him to reconsider the specificity and temporal palpability of life that he finds misconceived in his philosophical lineage and that takes the shape of the question of being.

This kind of turning also applies to his religious background. His movement from pretheology student to theology student, to religiousness without church or theology, to phenomenology, to a thinker of the question of being and of truth in the Greek sense of *aletheia*: this movement engaged a metamorphic turning through theology and religion. Within a few years of his appointment to Freiburg as a *Privatdozent* he would attempt to rethink such Christian words as "fall," "guilt," "word," "conversion," and "conscience," turning them out of contexts of faith and theological meaning to a contextual meaning without religious significance. Heidegger emphasized as his work progressed that such turning composes the way thought unfolds: A turning movement through and beyond a body of thoughts manifests the very life of thought. Without such turning thinking comes to its end. Further, a person cannot engage Heidegger's thought without encountering its metamorphic movement. In such movement claims about universal and timeless realities undergo for him a transformation specific to a particular engagement, and such engagement is defined by the singularity of its metamorphosis. Thought in its temporal-historical happening takes place as a living, particular, and self-transforming event. This orientation would mean that Heidegger was ill-disposed toward philosophical schools. When he was one of the leading philosophers in Europe, he did not encourage the formation of a Heideggerian school of thought. He believed that thinkers must find their own ways in their own settings and in their own life-worlds. Such thinking takes place in dialogue with the values, ideas, and beliefs that people find in other ways of thinking and living.

Heidegger experienced the beginning of transformative insights that distanced him from Thomistic and neo-Kantian thought in the late teens and early twenties, but he did not have an adequate way of bringing the insights to full thought. A "hermeneutics of facticity" names an early and landmark term for Heidegger. "Facticity" in this context refers to the irreducibility of things in their living events, and "hermeneutics" means interpretive explication. People recognize and interpret given things before they develop theoretical concepts about them. Interpretations arise as people live practically with things, encounter them in different contexts, use them, and feel their impacts. Prior to the distance invoked by theoretical reflection, people are enmeshed in their environments, and people's environments are filled with things that appear in their living, usually practical specificity. The ways in which things appear in those nontheoretical situations compose pretheoretical interpretations, and things in their appearing and facticity are nothing other than

their own events. How might philosophers think in and from a living situation that is filled with everyday and unconsidered interpretations and bring both those interpretations and themselves to conceptual elaborations that hold in mind the concreteness of things? What kind of language would be required? How would such thought transpire?

ARISTOTLE AND BEYOND

Heidegger came to the idea of a hermeneutics of facticity, that is, interpretations based on practical life, through his work on Aristotle and Husserl. His sympathy with Luther's attacks on scholasticism was consistent with his rejection of the scholastic interpretations of Aristotle that he learned as a student. Heidegger wanted to understand by intense reconsideration of many of his texts Aristotle's thought prior to Christian appropriations of it, the Aristotle whose concepts arose from his own Greek world. A radical departure from the Christianized Aristotle was required, Heidegger thought, in order to engage Aristotle's work in its vastly different manner of living when compared to that of later thinkers.

Heidegger's groundbreaking and influential interpretations of Aristotle provided him with a forceful return to European philosophy's Greek heritage. It also provided the occasion to rethink that heritage by means of an approach and vocabulary that he learned from Husserl. As he took away the Christian superstructure that encased the Aristotle he first studied, Heidegger reformulated Aristotle's thought with the eyes of a phenomenologist, eyes, he said, that he received from Husserl: he began to turn Aristotle's seemingly metaphysical thinking out of itself and into a way of thinking that moved decisively away from metaphysical formulation. In order to engage Aristotle well, Heidegger must preserve the vast difference between his own and Aristotle's spiritual environments. He would not overcome the difference; he would preserve it as he took his careful departure from Aristotle's way of thought by intense encounters with his texts.

Heidegger was arrested by, among other things, Aristotle's account of practical wisdom (*phronesis*). It described a kind of situational knowing that did not propose completion by reference to unchanging objects; it was intrinsically open to future development, and it functioned to open up future developments. Husserl's account of internal time consciousness also had, for Heidegger, the virtue of making impossible a *complete*, objective grasp of any thing. Heidegger found, however, in both Aristotle and Husserl an unquestioned prioritization of present time. This prioritization meant that neither saw clearly that futurity—coming to pass and yet to be—defined a nonobjectifiable dimension of presence or that presence is strangely modified—put in question—by its opening to futurity. This openness means that futurity defines presence, not as a categorical abstraction but as a constitutive indefiniteness and indeterminability in the lives of whatever happens. Time and its concept appeared to be the issues over which both Aristotle and Husserl stumbled—time and, for Husserl, the question of subjectivity.

THE PHENOMENOLOGICAL APPROACH

The sense of phenomenology, as Heidegger began to think of it during his years in Marburg, comes from the ancient Greek deponent, *phainesthai*, a middle voice form that means, "to show itself." A phenomenon is an event that shows itself. *Phainesthai* is formed from *phaino*, and that word means to bring something to light. The stem of the word is *pha*—*phos*, the light or shining whereby something is manifest. To give an account (*logos*) of phenomena meant for Heidegger to describe beings in their self-showing, to so speak and think that one is brought to things in their self-showing, and to give an account of the shining (the "light") that allows their manifestness. Self-showing composes the lives of individual beings.

Heidegger makes a sharp distinction between the specificity of a self-showing being and the enactment of that self-showing. Philosophers can give accounts of the ways beings show themselves, but they can also give accounts of the way the enactment of self-showing happens. This latter account addresses the being of beings and must not be confused with description of a highest being: being is not a being. This ontological difference between being (the occurrence of self-showing) and beings (a specific instance of self-showing) is basic in Heidegger's thought and persists in several forms throughout his career. It is a difference that characterizes the happening of phenomena: a phenomenon is a specific self-showing thing, and its being happens as the enactment of self-showing. Heidegger calls "ontic" the way a thing shows itself in its particularity. He calls "ontological" the happening of disclosiveness that is common for all phenomena.

The self-disclosive happening of phenomena, not a subjective state or action, thus becomes Heidegger's primary area for descriptive thought. Husserl too gave priority to the manifestness of phenomena and to ways things are manifest, and for him an intention may be

described as the direction of an appearing (manifest) event. "Direction" for Husserl suggested an unfulfilled, open-ended process of appearing that constitutes an event of transcendental subjectivity. It is located in transcendental, subjective acts of consciousness. Heidegger's way of engaging facticity, history, and time, however, turned him away from consciousness and toward the world. The unfulfilled directions of beings are not found in a proposed and lively structure of transcendental subjectivity but in the self-disclosive happening of beings-in-the-world. His thought turned through a characteristic modern priority given to subjective enactment and toward worldly structures that do not originate in human consciousness. He moved away from an epistemological orientation in his accounts of meaning, signification, and thought, and from a consequent emphasis on the subjects and objects of knowledge. He moved toward a way of thinking that is oriented by the disclosive, nonsubjective, and nonobjective enactments of things in the world. His aim was to show how those enactments do not begin with conceptual grasp or subjective appropriation, how the enactments are definitively historical and temporal, and how they happen self-disclosively in the world.

BEING AND TIME (1927)

Being and Time, one of the most influential books in the twentieth century, marks the culmination of Heidegger's years in Marburg. He had worked during this time especially on the history of the concept of time and brought together and honed ideas and preoccupations that began to form definitively as early as 1915. In lectures he elaborated his understanding of phenomenology and his departure from Husserl, as well as provided the conceptual scaffolding for *Being and Time* and many of its key terms. The availability in his Collected Works (*Gesamtausgabe*) of many of the lectures that he gave from 1923 through 1927 now allows scholars to follow the formations of *Being and Time*'s leading ideas and questions, a formation that this short discussion cannot pursue.

Being and Time appeared as a work in progress in the sense that its publication was hurried due to Heidegger's candidacy for Husserl's chair in Freiburg. It was projected as part of a much larger, multivolume series that Heidegger did not complete. The book made a huge cultural impact nonetheless, often due to interpretations of it that Heidegger found mistaken and at times offensive. Especially off the mark were those readings that turned *Being and Time* into a study in philosophical anthropology, existential humanism, or Husserlean phenomenology. The book's reception, in addition to Heidegger's own dissatisfactions with it, provided an occasion for Heidegger to see that he would have to turn through *Being and Time*'s concepts toward a different way of thinking if he were to carry out the book's mandate. He later understood *Being and Time* as an occasion in which he intensified a radical turn through his metaphysical inheritance toward a way of thinking that is based on that turn.

The book's mandate is found in reawakening the question of the meaning of being. Heidegger was persuaded that that question gave rise to European philosophical thought, although most traditions in European philosophy have obscured it. This question emerged for Heidegger when he was eighteen years old and read Franz Brentano's *On the Manifold Meaning of Being According to Aristotle*. Although Brentano's intentions were in part theological, Heidegger found through Brentano his entry into Husserl's *Logical Investigations*. This entry was in the context of the question of the meaning of being; and that question as well as a phenomenological way of thinking intensified for him and emerged together at the center of *Being and Time*.

Heidegger locates the question of being in the occurrence of *Dasein*. This word, *Dasein*, which has become a standard term in English among those who work within Heidegger's influence, names the located and disclosive occurrence of being in the world. It is not synonymous with "human being" but names the disclosive site of human lives. Dasein's way of occurring is the way things happen in their manifest availability for reference, recognition, and use. Dasein thus happens as the worldly region of disclosiveness. *Being and Time* provides a descriptive account of Dasein and shows that the being of worldly things is formed in their phenomenal quality, in their self-showing, not in any kind of creative or underlying substance. He further shows that the life, the being, of self-showing happens as temporal enactment and that its continuation is continuously in question: Dasein's being is able not to be. The question of the meaning of being thus arises in the prereflective occurrence of Dasein's mortality, not in a theoretical action by reflective subjects.

In his approach to this question Heidegger begins with what he calls the average, everyday understanding of being, that is, of the way beings happen in their practical lives. Usually we relate with things in terms of their usefulness and their standard identities in our environments. We have an operative, inchoate sense of what "to exist" means as we live with things. When we investigate something to know it better, we usually considers it as an object and work to make our statements and definitions appropriate to what we can find out about it. That means that

we usually do not question the meaning of nonobjective living occurrences and that we expect to discover something about the existence of things by treating them primarily as objects of use or knowledge. Their meaning and truth are found in our knowledge of them or in the appropriateness of the uses we find for them. A being is usually understood by reference to definitions of its objective presence; and if that presence is to be grounded in some way, philosophers usually look for a defining and continuously present reality that persists through the lives of changing and passing things. Such persistent and grounding presence might be found, for example, in such beings as God, Nature, Reason, or Transcendental Subjectivity.Transcendent beings such as these seem to provide a foundational meaning for finite things, and they embody the priority of presence for understanding the meaning of temporal passage: they are always present regardless of the changes and passage that beings undergo. The question of the meaning of being appears thus to be resolved by a presence that does not come to pass and that gives abiding meaning to passing beings.

Heidegger's account of Dasein, on the other hand, shows that temporality without a priority of presence defines the way beings are. In Part One of *Being and Time* he shows, first, that Dasein is intrinsically a caring occurrence. It is a way of being whose continuation is always in question, and consequently Dasein reverts to itself in the sense that it is always concerned with the preservation and continuation of beings and of itself. Being in the world is a passing occurrence, always situated in given histories and settings, always coping with uncertainties and transitions, always moving in the indetermination of the upcoming. The meaning of Dasein's being is care, Heidegger says—care, the inevitability of concern for whatever matters. Neither life nor world appears as guaranteed. Neither shows itself as supported by continuous presence. The disclosive happening of being in the world, in its *happening*, is always passing away. The meaning of care is thus found as the inevitability of losing presence, the inevitability of coming to pass, and the associated inevitability of taking care of whatever matters.

In the process of describing Dasein's temporality, Heidegger gives accounts in Part One of *Being and Time* of worldliness, relevance, spatiality, everyday superficiality, identity, worldly commonality, attunement, interpretation, and language. These accounts culminate with a section entitled, "Care as The Being of Dasein," and another, "Da-sein, Disclosedness, and Truth." In this part of the book he shows that our historical, situated, future-oriented being—our very life— is not at all like objective presence. Dasein happens as yet to be, as possibility to be. Individuals live in such possibility as in a "not yet" that is a dimension of any present moment. The completion that is sometimes attributed to definitive objects or identities is not a quality of living, worldly events. This constitutive, temporal incompletion describes at once Dasein's ontological disclosiveness and ontic worldly events in their specificity and concreteness.

Part Two of *Being and Time* intensifies the study of temporality around the axis of the question of the meaning of being. Whereas Part One began with accounts of the ways Dasein exists in an everyday way, Part Two shows that Dasein's existence is constituted fundamentally by a unifying structure of mortal temporality. The question of the meaning of being and of Dasein is founded in this structure. The guiding questions for this part address, on the one hand, the temporal, ontological unity of Dasein. On the other, they raise the possibility of living in fundamental and positive attunement with Dasein's ontological structure and of bringing together appropriately that structure and the specific way a person exists. He calls such living accord "authenticity." The possibility of authenticity is one of living in ways that affirm the unifying structure of mortal temporality. When such affirmation is achieved, people find a unity in their lives that is defined by finiteness, that is, by incompletion, indetermination, and being toward death.

In Part Two, Heidegger addresses such phenomena as the present occurrence of futurity, the draw of being for people and hindrances to alertness to that draw, ontological guilt, the ability for authenticity that is intrinsic to Dasein, and historicity. In that process he turns such words as "conscience," "call," and "guilt" out of their theological and religious heritage to an ontological and nontheological context. This part reconsiders the major phenomena addressed in Part One by what Heidegger calls a "primordial existential interpretation," that is, an interpretation that describes an ontological structure that is definitive for the occurrence of those phenomena. It develops the descriptive claim that temporality grounds care and is thus the meaning of care. The reader confronts again the thought that ontological grounding lacks substantial identity, presence, or necessity. The study ends with recognition of its own incompleteness.

THE ESSENCE OF TRUTH, TURNING OUT OF *BEING AND TIME*

The incompleteness of *Being and Time* was not due solely to the pressing circumstances under which it was submitted for publication. It was due also to Heidegger's con-

frontation with the inability of the book's language to say what needed to be said and with the limited range of his thought before the phenomena he addressed. Heidegger confronted the force of the metaphysical tradition in the way he used such words and phrases as "horizon," "structure," "the ontological condition for the possibility of something," "being," and even "Dasein." The book's manner of self-regulation and structure, its seemingly explanatory purpose, its conception of origin and history, and its inadequately conceived account of truth: these elements dissatisfied the author. He could see how the text could lead people to misunderstand his thought and its intentions. He also experienced the force of the movement of thought that had begun to uproot his metaphysical moorings. It was a force that he found turning him out of his own book toward a new beginning and in directions far more radical than he had foreseen.

The essay that most strikingly embodies the turning of his thought in the years shortly after the publication of *Being and Time* is *On the Essence of Truth,* which he wrote in 1933 and to which he returned over a period of nine years before publishing it. After *Being and Time* and prior to this essay he had lectured and written especially on Kant, Hegel, Schelling, and on basic concepts and problems in metaphysical thought. He turned down a professorship in Berlin, and enjoyed wide and growing recognition as a creative, leading philosopher.

In spite of his dissatisfactions with *Being and Time*, Heidegger had opened up the question of the meaning of being—or that question began to open up to him. It could be stated in several ways. Classically: Why are there beings instead of nothing at all? In terms of appearing: How is it that things appear, are present and manifest, and show themselves rather than not appearing at all? In terms of finitude: How does being happen in the passing presence of things? In general, the question of the meaning of being is at once a question of fundamental uncertainty in life, presence and passage, and of disclosure and closure in the occurrence of phenomena. Temporality is found in Dasein's having been now yet to be, a "structure" that seems to defy the meaning of "structure." At every turn as we consider this question we encounter the happening of manifest beings, and this—the happening of manifest beings—for Heidegger is essentially a question of truth. The question of being (of manifest happening or eventuation) is at once the question of truth. How is it that the temporality and disclosure of being inevitably raise the question of truth for him?

Long before Brentano pointed out the connection in Aristotle's thought of "true" and "being," and long before Aristotle himself, the word *aletheia*, usually translated as "truth," played a major role in ancient Greek civilization. The word, which Heidegger understands as combining the alpha privative with *lethe* (oblivion), names an occurrence when something is manifest, self-showing, and apparent. A being is exposed in its disclosiveness, is quite explicitly there where it happens. Its truth happens as its self-disclosure, as its own manifest presence. We have seen that for Heidegger the disclosiveness of something is not identical with what something shows itself to be, that there is a basic difference between *what* something is and disclosiveness as such. The disclosiveness (roughly, dis- = the alpha privative, and -closive = *lethe*) lets something be as it shows itself. Disclosiveness is, Heidegger says, an open region that, while apparent with manifest beings, is not limited to the specificity of what a being is, not limited to *a* being's time, place, and identity. "Truth" in this context means the free openness of disclosure, and a truth is found in the self-disclosure of a being. To know something in its truth is to engage it in a way appropriate to its "essence," to its own disclosive eventuation.

The factor of oblivion or complete lack of apparentness, however, elaborates what Heidegger considered in *Being and Time* as the mortality of phenomena: beings appear with nothing transcendent and specific to ground them or guarantee their lives; they are "grounded" in their disclosive eventuations. Their dis-closure carries oblivion with it as a strange and pervasive mortal factor, one that makes impossible a complete grasp of any phenomenon. It is as though oblivion protects a being from complete exposure, gives something other to its truth, removes it from availability. *Lethe* suggests concealment, withdrawal of being (i.e., of disclosiveness), and untruth (a complete absence of disclosure).

Heidegger's account of essence in *On the Essence of Truth* no longer struggles with what appear to be quasitranscendental structures of existence as he locates essence clearly in both the eventuation of disclosure and the history of the thought of truth. The untruth, the "non-essence," of concealment—the oblivion of being—suggests the inadequacy of "finitude" as a descriptive term for Dasein's temporality. It suggests a new departure in which existential uncertainty is compounded by impenetrable closure to the manifestness and "light," of life. This departure includes a strong sense of mystery, not the mystery of Pure Light or of a hidden fullness of being, but mystery in the sense of unsayable oblivion in the midst of disclosive openness. As his thought turns to them, both oblivion and truth appear to happen for Heidegger when he attempts to speak with alertness to them. Is it possible,

he wonders, that the turning he is undergoing describes a movement of disclosure and oblivion that is definitive as well as obscure in the history of western thought? Is his transforming movement toward early senses of *aletheia* following a path to the early beginning of metaphysical thought, one that makes apparent an oblivious departure of western thought from the questionableness and uncertainty of truth?

POLITICAL CATASTROPHE

At the time that this turning gave him new directions and possibilities for thinking, Heidegger became embroiled in a politics that belied the most promising of those possibilities. Ever a German nationalist, long persuaded of the unjust consequences of the Treaty of Verdun, convinced that Germany must resist communism at all costs, and disappointed in the inefficiency of democratic procedures, he embraced Hitler's National Socialism as Germany's best political hope. By 1933 Heidegger saw this party as a force toward revival of German culture and restoration of Germany's leadership in the transforming of European, materialistic civilization.

In April of that year the faculty at the University of Freiburg elected him rector, and in that role he supported Nazi ideology for German resurgence and helped to form university policy according to party interests. Nazi authorities, however, criticized him strongly for his failure to support anti-Semitic rhetoric and policy. Heidegger was not a gifted administrator. Sharp political and educational controversies intensified, and he resigned his post ten months after assuming office. His dream that Hitler, as a man of destiny, would transcend the foolish people around him and become a heroic, spiritual leader allowed Heidegger to support Hitler long after he became disillusioned with the National Socialist party. He began to undercut party interpretations of Nietzsche, and by the mid-1930s his classes were audited by suspicious party appointees. The party also restricted his freedom of movement and publications, and he was punished by means of hard physical labor when the authorities drafted him into the People's Militia. Heidegger never used his international stature as a base to criticize National Socialism, and although he privately admitted his errors after the war, he never publicly addressed German atrocities. After the war the French occupational authorities prevented Heidegger from teaching in the university until 1951.

There is considerable controversy around the question of whether Heidegger's philosophy led to his political hope, error, and naivete. Some people see a profound and causal linkage, while others see more distance and inconsistency between his thought and his politics of the 1930s. That decade, regardless of the way one assesses the controversy, constitutes a dreary segment in Heidegger's life. Responses to it have on occasion been ones of continuing outrage, whereas others find Heidegger's thought worthy of sustained and positive engagement. Perhaps the dangers of forgetting and those of a righteous condemnation should be foremost on our minds when we consider the importance of Heidegger's misjudgments.

SEARCHING FOR ANOTHER WAY TO THINK

During the 1930s Heidegger searched for language, conceptual movements, and rhythms of thought that could engage appropriately the disclosive happening of things. The systematic rationality of "onto-theology," that is, of traditional philosophy, seemed to constitute anxious attempts to overcome the questions of truth and of the meaning of being which gave European philosophy its inception. Approaches called materialistic, idealistic, empirical, and analytical seemed dedicated to forgetting those questions. Post-Cartesian thought gave forms of subjectivity and objectivity ontological priority, whether or not subjectivity was considered ahistorical or historical. In which writing and conceptuality might Heidegger find a degree of positive alertness to the questions that he found as the moving forces in European philosophy? This was a time of considerable isolation for Heidegger as he looked for alternative ways of thinking. He experienced disappointment in his own work, discouragement in its reception, political failure, and an uncertain future for himself and the kind of disciplined education that he thought necessary for the survival of western civilization. That would be an education in the classical origins of Europe and its many traditions, and it would be an education that recognized what he came to see as the devastation of instrumentally oriented culture and the desolation of contemporary spiritual accomplishment.

Heidegger had read widely in literature and especially in poetry since he was in his teens. The poetry of J. W. Goethe, Friedrich Hölderlin, Rainer Maria Rilke, Stephan Georg, and Georg Trakl, among others, helped to form his mind, and he turned to poetry, especially to Hölderlin's, with renewed intensity in the 1930s. He wanted to find ways to formulate and express what he sensed but could not say to his own satisfaction. He also read the ancient Greek tragedians with an emphasis on the two great questions that preoccupied him. He gave courses and lectures on Nietzsche's thought and found in

it a welcomed emphasis on the connections of art and thinking. In Nietzsche's thought he found as well a culminating and destructive fulfillment of that nihilism prepared by European metaphysics. It is a nihilism given partial expression by Hegel and carried out after Nietzsche by a technological society that is oriented around subjects and objects of use and knowledge.

His attention turned increasingly to thought and language as disclosive events. Truth, *aletheia*, names, as we have seen, the enactment of self-showing; the truth of thought and language is found as thought and language, in their occurrences, give place and occasion to self-showing phenomena. The life of thought and language is found in the ways they engage the manifest lives of things. Is the engagement defined by organizational structures? By the power of will? By categories of knowledge? By means of production? By patterns of trade? People's lives are normally carried out with such structures and activities. In addition to our normal ways of acting, however, we might also give attention to the self-showing dimension of anything that is present with us. We might learn to connect with things with a sense that their very happening addresses us, that our "hearing" is found in the ways we live with them. If our living provides ways to allow events prominence in their disclosive dimension, our thinking and speaking might well grant to them a dwelling place, not for their utility only, but also for their self-showing, for the essence of *their* lives. Although that manner of living is not forecast in language and thought dominated by the importance of subjectivity and objectivity, it does appear significantly in the work of some of the artists and poets Heidegger read. He explored and experimented for many years with possibilities for language and thought that are influenced by poetic rather than traditionally philosophical kinds of awareness. He intended to find areas of encounter between poetic and philosophical language, to enrich each in their engagements, and at best to occasion the emergence of a kind of mentation that finds its truth in allowing the truth—the self-disclosure—of whatever happens in its environment.

CONTRIBUTIONS TO PHILOSOPHY

Heidegger's sense of failure in the language of *Being and Time* to say what needed to be said of the questions of the meaning of being and truth figured a large part of the turning in his thought during the 1930s and 1940s. His *Contributions to Philosophy (From Enowning)* is a major work that emerged in that turning between 1936 and 1938. He wrote it in the impetus provided by his work on art and especially on poets, his rethinking of the inception and decline of European metaphysics, and his search for a new beginning for thinking. In its fuguelike formation, this series of meditations composes an effort to find ways to speak of what seems always to remain unsaid yet present in European philosophy. It is an effort to think in the obscure questions of being and truth, to speak in their modern wake, rather than to re-present them. Heidegger invites the reader to engage in strange and often wrenching movements of language as he attempts to let the questions emerge and turn thought and language from the tracks that move them inevitably away from what most threatens and yet impels the remarkable occurrence of European thinking. If he succeeds he reconceives *Being and Time* in a radical return to *Being and Time*'s issues and makes that return by the force of turning away from the book's structure and articulation. The lives and forces of the questions of truth and the meaning of being, not their resolution, guide this book's movements. It is a work that attempts to think the inconclusiveness of its major issues. Its success would be found in the emergence of a way of thinking that makes apparent what incited western thought and what western thought in its formation nonetheless virtually lost.

BEYOND HUMANISM

In 1946, Heidegger responded to questions raised by the French philosopher Jean Beaufret. Published in 1948 as "Letter on Humanism," Heidegger made explicit in his response not only his distance from "Existentialism," but also his reservations about "Humanism" as it is conceived in post-Enlightenment Europe and North America. Issues of human life and community are not best located in conceptions and images of human subjectivity. He developed his descriptive claim that humanistic values are often the source of destructive depreciations of human life. The essay comprises a sustained reflection on what is destructive and constructive for people and on basic assumptions regarding the essence of worldly life. It has had widespread influence on thinkers in the second half of the twentieth century who find in humanistic ideals elements that, contrary to their stated purposes, do harm to societies and individuals.

In the 1940s and 1950s, Heidegger also relentlessly pursued questions concerning the essence of technology. Technology for him constitutes a way of life that overrides the subtle and most important dimensions of the existence of things as well as of people. The word "technology" thus names the most dangerous form that European nihilism takes. Among the best known of his essays during this time are "The Question concerning Technology"

and "Building, Dwelling, Thinking." One of the most far reaching and profound of his works in this context is *The Way to Language*, in which Heidegger brings to bear in a cumulative way his preoccupations with thought, language, technology, and dwelling

"Engagement," "encounter," and "way" are important words to hold in mind as a person reads Heidegger's works. He often described thinking as made up of ways of letting things show themselves in the specificity of their contexts. Thinking composes engagements with all manner of manifest things—texts, behaviors, trees, bells, images, concepts. Manifest things are alive in their manifestness, and thinking properly allows their differences as they happen, engages them with alertness to *their* happening. At best, the engagement composes a dialogue, an encounter that no one individual produces. Thinking is what takes place in the dialogue. For Heidegger such an engagement is a social, historical, and communal event that cannot be reduced to the sum of its parts or generalized on the basis of universal meanings. Thinking is made up of engagements with living events in their happening, their eventuation. As he saw it, there are many ways, and the issue for thinking is not one of calculating the correctness of assertions but rather one of making evident or, in unfortunate instances, obscuring beings in their self-showing. Thinking, always opening to the differences of events, always coming to pass in its own life, always on the way to something else, remained at the center of Heidegger's preoccupation with the questions of being and truth until his death in 1976.

See also Existentialism; Hermeneutics; Phenomenology.

Bibliography

WORKS BY HEIDEGGER

Discourse on Thinking. Translated by John M. Andersen and E. Hans Freund. New York: Harper and Row, 1966.

On the Way to Language. Translated by Peter D. Hertz. New York: Harper and Row, 1971.

Poetry, Language, Thought. Translated by Albert Hofstaedter. New York: Harper and Row, 1971.

Early Greek Thinking. Translated by David Farrell Krell and Frank A. Capuzzi. New York: Harper and Row, 1975.

Gesamtausgabe, edited by Hermann Heidegger. Frankfurt am Main. Vol. 1, 1977. The volumes in this series continue to be published.

The Basic Problems of Phenomenology. Translated by Albert Hofstaedter. Bloomington: Indiana University Press, 1988.

Kant and the Problem of Metaphysics. Translated by Richard Taft. Bloomington: Indiana University Press, 1990.

History of the Concept of Time. Prolegomena. Translated by Theodore Kisiel. Bloomington: Indiana University Press, 1992.

Basic Writings, edited by D. F. Krell. San Francisco: Harper, 1993.

Aristotle's Metaphysics Theta 1–3. Translated by Walter Brogan and Peter Warnek. Bloomington: Indiana University Press, 1995.

Being and Time. Translated by Joan Stambaugh. Albany: State University of New York Press, 1996.

Hölderlin's Hymn "The Ister." Translated by William McNeill and Julia Davis. Bloomington: University of Indiana Press, 1996.

Pathmarks, edited by William McNeill. Cambridge, U.K.: Cambridge University Press, 1998.

Contributions to Philosophy (From Enowning). Translated by Parvis Emad and Kenneth Maly. Bloomington: Indiana University Press, 1999.

Introduction to Metaphysics. Translated by Gregory Fried and Richard Polt. New Haven, CT: Yale University Press, 2000.

WORKS ON HEIDEGGER

Brainard, Marcus, David Jacobs, and Rick Lee, eds. "Heidegger and the Political." Pts. 1 and 2. *Graduate Faculty Philosophy Journal* 14, no. 2; (1991): 15, no. 1 (1991):

Kisiel, Theodore. *The Genesis of Heidegger's* Being and Time. Berkeley: University of California Press, 1993.

Krell, David Farrell. *Daimon Life: Heidegger and Life Philosophy*. Bloomington: Indiana University Press, 1992.

Krell, David Farrell. *Intimations of Mortality: Time, Truth, and Finitude in Heidegger's Thinking of Being*. University Park: Pennsylvania University Press, 1986.

Sallis, John. *Delimitations: Phenomenology and the End of Metaphysics*. Bloomington: Indiana University Press, 1986.

Sallis, John, ed. *Reading Heidegger: Conversations*. Bloomington University Press, 1993.

Sefranski, Rüdiger. *Martin Heidegger: Between Good and Evil*. Translated by Ewald Osers. Cambridge, MA: Harvard University Press, 1998.

Scott, Charles E., Susan M. Schoenbohm, Daniela Vallega-Neu, and Alejandro Vallega, eds. *Companion to Heidegger's Contributions to Philosophy*. Bloomington: Indiana University Press, 2001.

Charles E. Scott (2005)

HEIDELBERG SCHOOL

See *Neo-Kantianism*

HEIM, KARL

(1874–1958)

Karl Heim, the German theologian, was born at Frauenzimmern in Württemberg. He studied at Tübingen and was professor of theology at Münster (1914) and at Tübingen from 1920 until his death.

Heim's work has philosophical interest insofar as he was concerned all of his life with the problem of restating

Christianity in a form that would be credible in the scientific age. His early work explored the epistemology of religious faith, and his developed account draws on the I–thou philosophy of Martin Buber and also on some of the concepts of modern science. Heim's fundamental point was that the experiencing subject cannot itself become an object and so cannot be brought under the objectifying categories of scientific thought. Thus we have a way of breaking out of, or transcending, the objective world of science, for there is open to us also a nonobjectifiable, interpersonal world. Heim spoke of this as a further dimension of experience, analogous to a dimension of space. This suggests a new way in which we may try to think about the transcendent God; and belief in such a God seemed to Heim the most important point calling for defense and restatement in the Christian tradition. Modern cosmology has made it senseless to talk of such a God as "up there" or "out there" or "beyond." But this God is not an object in the world any more than the experiencing subject is, and God too must be sought in the nonobjectifiable dimensions of experience, not in the realm of I–it. Just as the situation is revolutionized if we add a third dimension to a two-dimensional manifold, and what was hitherto impossible in two dimensions may be possible in three, so Heim believed that the conflicting attitudes of religion and science may be reconciled by admitting the multidimensional character of experience. We are familiar today with the notion of geometries of more than three dimensions, and we can think of an interpersonal space as having a different order and structure from physical space. Both kinds of space are embraced in an archetypal space, which is also a suprapolar space because it resolves the polarities of both the interpersonal and the physical spaces. This suprapolar space is the abode of God; it cannot be proved, but it is disclosed in the experiences of faith that may be likened to opening up new dimensions of a fuller life. Heim also taught a doctrine of panpsychism, which suggested that the further dimensions opened up in man's encounter with other persons and with God are at least potentially present at all levels of being.

See also Buber, Martin; Panpsychism; Religion and the Biological Sciences; Religion and the Physical Sciences; Space.

Bibliography

PRINCIPAL WORKS BY HEIM

Glaubensgewissheit. Leipzig, 1916.

Glaube und Denken. Berlin, 1931. Translated as *God Transcendent.* Edinburgh, 1935.

Jesus der Herr. Berlin, 1935. Translated as *Jesus the Lord.* Philadelphia: Muhlenberg Press, 1961.

Der christliche Gottesglaube und die Naturwissenschaft. Tübingen, 1949. Translated as *Christian Faith and Natural Science.* London, 1953; New York: Harper, 1957.

Die Wandlung im naturwissenschaftlichen Weltbild. Hamburg, 1952. Translated as *The Transformation of the Scientific World View.* London: SCM, 1953; New York: Harper, 1954.

Weltschöpfung und Weltende. Hamburg, 1958. Translated as *The World: Its Creation and Consummation.* Philadelphia, Muhlenberg Press, 1962.

WORKS ON HEIM

Eerikainen, Atso. *Two Dimensions of Time: The Dimensional Theory of Karl Heim.* New York: Peter Lang, 2002.

Eisenhuth, H. E. "Im Gedenken und Karl Heim." In *Theologische Literaturzeitung* 83 (Leipzig, 1958): 657–662.

Holmstrand, Ingemar. *Karl Heim on Philosophy, Science, and the Transcendence of God.* Uppsala, Sweden: Uppsala Universitet, 1980.

Spemann, F. *Karl Heim und die Theologie seiner Zeit.* Tübingen, 1932.

John Macquarrie (1967)

Bibliography updated by Christian B. Miller (2005)

HEISENBERG, WERNER

(1901–1976)

The German physicist Werner Heisenberg was born in Würzburg; he studied physics in Munich under Arnold Sommerfeld and received his doctorate from Munich in 1923. Heisenberg became a lecturer and assistant to Max Born at Göttingen in 1924. He continued his studies at the University of Copenhagen, where he collaborated with H. A. Kramers. He succeeded Kramers in 1926 as lecturer in physics there. Heisenberg was professor of physics at Leipzig from 1927 to 1941 and professor at Berlin and director of the Kaiser Wilhelm Institute for Physics from 1941 to 1945. He was named honorary professor and director of the Max Planck Institute for Physics at Göttingen in 1946 and served as honorary professor and administrative director of the Max Planck Institute for Physics and Astrophysics in Munich from 1958 to 1970. He was awarded the Nobel Prize for physics in 1932.

Heisenberg's contributions to physics are contained in more than 120 papers covering a great variety of topics. We shall here deal with two topics only, with the invention of matrix mechanics and with Heisenberg's more recent theory of elementary particles.

MATRIX MECHANICS

The older quantum theory of Niels Bohr and Sommerfeld had tried to combine classical physics with the new quantum laws and to use the predictive power of both. The resulting theory was a mixture of classical notions—some useful, others apparently redundant—of new ideas and of ad hoc adaptations. Thus, for example, transition probabilities and selection rules were calculated, or guessed at, by examining the Fourier coefficients of the motions

$$\Phi_i(t) = \sum_{n=-\infty}^{+\infty} X_i(n,\omega_i)\exp(in\omega_i t)$$

of the independently vibrating parts of the atom, while the motion Φ_i itself had to be denied any physical significance. In addition, the theory had failed in important respects. It clearly was but an intermediate step on the way to a satisfactory mechanics of the atom. The final theory is essentially due to the efforts and the very different philosophies of two men, Heisenberg and Erwin Schrödinger.

According to Heisenberg we must abandon all attempts to give a detailed description of the unobservable internal motions of the atom. Such motions are but the result of the continued use of classical ideas in a domain that is inaccessible to direct experimental examination. Considering that these ideas may be in need of revision it would seem to be wise to construct a theory that is expressed solely in terms of such "outer" magnitudes as frequencies and intensities of spectral lines. Speaking formally this means that we want to predict by using the X_i directly and without appeal to the Φ_i. Now Bohr's investigations had already gone a long way toward determining the required properties of the X. His idea of a rational generalization corresponds exactly to what Heisenberg had in mind. Heisenberg himself provided additional rules of calculation that were sufficient for solving some simple problems, such as the problem of the harmonic oscillator. It was not known to him at the time that the rules were those of an algebra of noncommuting matrices; this was soon recognized by Born, who, together with Pascual Jordan and Heisenberg, completed the formalism a few months after Heisenberg's first paper had appeared.

A new atomic mechanics was at last in sight. Its meaning, however, was far from clear. Macroscopic objects whose positions and momenta could be ascertained with a higher degree of precision were represented by infinite arrays of complex numbers, none of them corresponding in a simple way to visible properties. "Can you imagine," objected H. A. Lorentz at this stage, "me to be nothing but a matrix?" It was again Heisenberg who, after the theory had been completed in a somewhat unexpected fashion by Schrödinger, made an essential contribution here by showing, in his uncertainty relations, to what extent classical notions could still be used in the interpretation of microphysical theories.

Heisenberg was to use the principle to rebuild a theory by working "from the outside in" once more in 1943, in order to eliminate certain difficulties in the quantum theory of fields. Believing these difficulties to be due to the disappearance of the ordinary space-time relations below 10^{-13} centimeters, he tried to replace field theory by a formalism that for any interaction transforms asymptotic anterior states into asymptotic posterior states without dealing with the details of the interaction. This so-called S-matrix theory was taken up by Geoffrey Chew and others for the calculation of the properties of strongly interacting particles. This led to what some physicists regarded as the beginning of a "third revolution" of twentieth-century physics, to the idea that particles are composites and that the properties of all of them can be obtained in a step by step procedure, starting with the interaction of any small subset ("bootstrap hypothesis"). Spatiotemporal relations are alien to this scheme, which therefore cannot develop a theory of measurement. Nor does there seem to be any possibility of extending it to other types of interaction.

ELEMENTARY PARTICLE THEORY

Heisenberg, who had been the first to stress the nonexistence of a criterion for distinguishing "elementary" particles from composites, has in the meantime developed a different theory in which elementary particles are stationary states of a single physical system, "matter." The field operators refer no longer to particles but to this basic matter (which Heisenberg sometimes compares to Anaximander's *apeiron*). The masses of the particles arise wholly from the interactions due to the nonlinearity of the basic field equation. There are no "bare particles." Other properties are supposed to follow from the symmetries of the field equation. Strange particles of spin 0 and 1/2 have been dealt with, to a certain extent, on the basis of approximation methods (this refers to 1962). There are only programs, no exact predictions, for weak interactions.

Heisenberg's philosophical speculations were always intimately connected with his physics. They were original and exciting. The same cannot be said about his more

general observations on philosophical matters. However, he should not be blamed for this disparity, as it is at any rate only in close connection with reality that philosophy can be both interesting and fruitful.

See also Quantum Mechanics.

Bibliography

Heisenberg's early work and its relation to wave mechanics and to experiment is described in *Die physikalischen Prinzipien der Quantentheorie* (Leipzig: Hirzel, 1930), translated by Carl Eckart and Frank C. Hoyt as *The Physical Principles of the Quantum Theory* (Chicago: University of Chicago Press, 1930). The theory of the S-matrix is explained in "Die 'beobachtbaren Grössen' in die Theorie der Elementarteilchen," in *Zeitschrift für Physik* 120 (1943): 513–538 and 673–702. For a survey of Heisenberg's new field theory, see "Quantum Theory of Fields and Elementary Particles," in *Reviews of Modern Physics* 29 (1957): 269–278, and "Die Entwicklung der einheitlichen Feldtheorie der Elementarteilchen," in *Naturwissenschaften* 50 (1963): 3–7. *Die Physik der Atomkerne* (Brunswick, Germany: Vieweg, 1943) explains nuclear physics and Heisenberg's contributions to it. *Wandlungen in den Grundlagen der Naturwissenschaften* (Leipzig: Hirzel, 1935) is a survey of advances and discoveries from 1900 to 1930. *Physics and Philosophy* (New York: Harper, 1959), which contains the Gifford Lectures given in 1955/1956, deals with the same subject on a broader historical and philosophical basis and also discusses current objections to the Copenhagen interpretation of quantum theory.

For the atmosphere at Göttingen in the "golden twenties," see F. Hundt, "Göttingen, Kopenhagen, Leipzig im Rückblick," in *Werner Heisenberg und die Physik unserer Zeit*, edited by Fritz Bopp (Brunswick, Germany: Vieweg, 1961), pp. 1–7, and Robert Jungk, *Brighter Than a Thousand Suns* (New York: Harcourt Brace, 1958).

Paul K. Feyerabend (1967)

HELD, VIRGINIA

(1929–)

Virginia Held, American philosopher, received her PhD in philosophy from Columbia University and is a professor of philosophy at Hunter College and the Graduate School of the City University of New York. In addition to working as a reporter, she has also taught at Barnard College, Dartmouth College, the University of California at Los Angeles, and Yale University. She is the author of numerous scholarly books and journal articles in the areas of social and political philosophy, ethics, and feminist philosophy. In particular, she has contributed to our understanding of the moral importance of birth and mothering, to debates on limits on markets, to discussions on collective responsibility, and to the literature on moral methodology and metaethics.

According to Held, moral theorizing requires paying attention to actual moral experience. In *Rights and Goods*, Held argued for a view she calls "experimental morality," a version of John Rawls's method of reflective equilibrium, according to which actual moral agents ought to try out various moral approaches and see what it is like to live by them. An advocate of pluralism, she advances the view that different persons, in different roles or contexts, should develop and experiment with different approaches to morality. This link between experience and moral theory connects her work on moral theorizing in general (1984) and her work in feminist ethics (1993).

Focusing on the moral significance of experience, Held drew attention to how women's experiences have been left out or devalued by traditional moral theorizing. In particular, she has cast light on the experiences of women whose activities as mothers and caregivers have often been wrongly dismissed as mere biological reflexes. Critical of the tendency in some moral and political philosophy of depicting persons as rational, independent agents who make mutually disinterested agreements, Held encourages, instead, a conception of persons that appreciates our interdependence and the caring, rather than contractual, nature of our relationships. She rejects the notion that the impartial rule-follower is a paragon of moral virtue and recommends in its place the ideal of the empathetic caregiver, stressing that care ought to be valued for its own sake and not merely as a means of carrying out impartial rules.

Rather than construing care as a permissible nicety that comes into play only after justice and equality have been secured, Held has come to view care as one of the most basic moral values. Without care, she says, humans simply would not survive. Emphasizing the importance of care in the public realm, as well as in the private realm, is not, for Held, license for widespread paternalistic interference. Indeed, she thinks that appropriate care is often about cultivating a capacity for autonomy in the person cared for. She proposes a conception of care that can extend to distant others, predicting that a revaluation of care would inspire a more committed defense of others' rights and bolster political support for public health care and child care.

See also Ethics; Feminist Ethics; Feminist Philosophy; Feminist Social and Political Philosophy.

Bibliography

WORKS BY HELD

Public Interests and Individual Interests. New York: Basic Books, 1970.

Properties, Profits, and Economic Justice. Belmont, CA: Wadsworth, 1980.

Rights and Goods: Justifying Social Action. New York: Free Press, 1984.

Feminist Morality: Transforming Culture, Society, and Politics. Chicago: University of Chicago Press, 1993.

WORKS EDITED BY HELD

With Kai Nielson and Charles Parsons. *Philosophy and Political Action: Essays Edited for the New York Group of the Society for Philosophy and Public Affairs*. New York: Oxford University Press, 1972.

With Sidney Morgenbesser and Thomas Nagel. *Philosophy, Morality and International Affairs: Essays Edited for the Society for Philosophy and Public Affairs*. New York: Oxford University Press. New York, 1974.

Justice and Care: Essential Readings in Feminist Ethics. Boulder, CO: Westview Press, 1995.

Amanda Porter (2005)
Samantha Brennan (2005)

HELLENISTIC THOUGHT

The Hellenistic era extends from the death of Alexander the Great in 323 BCE to the conquest of Egypt by the Romans in 30 BCE. Though defined in terms of political events, it is also host to distinctive developments in Greek intellectual life. Chief among these are the foundation and consolidation of organized *schools* as the focus of philosophical life, especially in Athens; the growing independence of various special sciences from their original philosophical context; and a geographical expansion (in the wake of Alexander's conquests and the foundation of Greek-speaking kingdoms in the eastern Mediterranean) that had significant long-term consequences. The intellectual life of Hellenistic Greece changed again as Roman political authority gradually came to dominate in the region. Throughout this period, Greek intellectuals (both philosophers and scientists) became more prominent and important in civic life, often achieving political recognition even in foreign cities; in 155 BCE three prominent philosophers, none of them from Athens, were chosen to represent the city on an embassy to Rome. Prominent intellectuals were offered patronage by the new Hellenistic kingdoms.

The first major organized school at Athens was the Academy, founded by Plato. Aristotle's associate and successor, Theophrastus, and later Strato of Lampsacus, carried on the traditions of his work in the Lyceum. Other philosophical schools in the fourth century were of minor importance, although the hedonistic school based at Cyrene in North Africa was influential. Yet within the first few decades of the Hellenistic era two major new schools, representing significant philosophical directions with lasting influence, were established. Epicurus, the founder of Epicureanism, was an Athenian and established the Garden there, but his followers spread around the Aegean basin in a network of smaller institutions that remained connected to the original school. Zeno came from the town of Citium on the island of Cyprus to establish his school in the Painted Stoa in the Athenian agora, and throughout its history it continued to attract philosophers from all over the Hellenistic Greek world, especially Asia Minor.

Epicureanism and Stoicism quickly became successful and attracted adherents for centuries to come. Epicureanism revived the atomistic physics pioneered by Leucippus and Democritus and linked it tightly with a hedonistic ethics and quietistic political philosophy. Stoicism depended on the mainstream Socratic tradition; its cosmology and physics drew primarily on Plato and Aristotle and its ethical and political theory were heavily influenced by Socratic ideas colored by the Cynic tradition stemming from Diogenes of Sinope and Crates of Thebes.

Stoicism and Epicureanism were in some ways polar opposites. The former championed god's providence while Epicurus denied it. Stoic physics asserts the continuity of all matter (which is itself permeated by a divine cause giving it form), while for Epicurus all things, even the gods, are composed of atoms and void. Like Plato and Aristotle, Stoics believed that society and its institutions rest on deeply rooted features of human nature, but Epicureans held that societies are formed by agreements among people about mutual preservation and advantage. Stoicism (inspired in part by the dialectical school and Megarian philosophers) led the way in the development of logic and dialectic, while Epicurus rejected logic along with many other specialized intellectual endeavors as useless. For Epicurus even physics mattered only in so far as it was essential to achieving tranquility.

Despite these contrasts, the two schools shared a great deal. Both rested their philosophy on broadly empiricist epistemologies, according to which normal sensory experience was the ultimate source and criterion for knowledge, and both rejected the idea of causally efficacious incorporeal entities and emphasized the material foundations of all reality. Neither school could accept the

central role of *form*, either in the Platonic version in which forms were separate from material particulars, or in the immanentist version of Aristotle, for whom form and matter were the two components of all concrete objects; nor could they embrace the concepts of an incorporeal deity or an immortal and incorporeal soul animating the body.

As these new schools emerged, the Academy changed its intellectual course; under the leadership of Arcesilaus it adopted a skeptical practice, devoting its energies not to the development and refinement of positive theories but to the dialectical criticism of those philosophers who claimed certainty for their own views. Stoicism was its chief target, and it can be argued that the main inspiration for this skeptical turn was the desire to refute those who claimed that the physical world could yield certain knowledge. The Academy maintained its dialectical approach for nearly two centuries; its high point came under the intellectual leadership of Carneades in the second century BCE. His followers came to disagree about the nature of his commitment to skepticism and gradually reverted to *dogmatism*, the conviction that knowledge is achievable. The Lyceum (sometimes also called the Peripatos) did not long maintain its philosophical vigor after the death of Theophrastus and its leaders became better known for their achievements in the sciences than in philosophy. Only Critolaus, the contemporary of Carneades, achieved importance in philosophy proper. The renewal of Aristotelianism had to await the end of the Hellenistic era.

With each generation the Stoic school changed and developed, with most of its leaders making significant innovations. The third head, Chrysippus of Soli, systematized and reworked nearly every aspect of Stoic thought, developed the formal logic for which the Stoics remained famous until the end of antiquity, and exerted control of the school's trajectory for several generations after. In the late second century Panaetius of Rhodes and his student Posidonius of Apamea made a comparable mark, reintegrating Platonic and Aristotelian influences into the school's intellectual life. By contrast, in all but details the Epicurean school was marked by conservatism and doctrinal unity.

The interaction between philosophical schools and the special sciences is a topic of particular interest in this period. Except for the medical texts in the Hippocratic Corpus, there are few traces of specialized scientific writing before 300 BCE, although Aristotle makes frequent allusions to an optical and astronomical literature that was distinct from philosophy and had a mathematical character. Hellenistic optics, as represented by Euclid's *Optics*, was the physical science that engaged most actively with philosophy. Euclid uses a geometrical apparatus to model a selection of phenomena of visual perception that reflect not only Aristotle's analysis of the *objects of sense perception* but also contemporary Hellenistic epistemological concern with the reliability of the senses. The Euclidean model, invoking rectilinear "visual rays" that radiate from eye to object, could be reconciled with Stoic physics as well as with the more eclectic materialism of Theophrastus and his Peripatetic successors.

Astronomy, by contrast, seems to have disengaged from philosophy after Aristotle. Deeply impressed by the regularity of astronomical phenomena and by Eudoxus's ingenious hypotheses of rotating spheres that seemed to account for them, Aristotle posited a sharp discontinuity between the irregularly changeable globe of matter at the centre of the cosmos, in which we dwell, and the eternally unchanging outer shell, composed of a distinct kind of matter, that is the realm of the sun, moon, planets, and stars. The Stoics and Epicurus, by rejecting this discontinuity, made it harder to reconcile their physics with the mathematically abstract celestial models of the astronomers. Astronomical writers such as Aristarchus in the third century and Hipparchus in the second relied on geometry, arithmetic, and optical observation as criteria for their models, and sometimes put forward alternative models to explain the same phenomena. In the view of a physically oriented philosopher such as Posidonius, the astronomers' models did not constitute proper explanation, which only the philosophers could provide. Nonetheless such results of astronomical reasoning as estimates of the sizes and distances of the sun and moon and Eratosthenes's measurement of the earth's circumference became commonplaces of philosophical discourse.

Mesopotamian traditions of divination from celestial phenomena were known in the Greek world as early as the third century, and the Stoics in particular took a lively interest in them as they did in other forms of divination. It was only about the beginning of the first century BCE, however, that a distinctly Greek astrology endowed with sufficient complexity and rationale to claim scientific status took form. Astrology was founded on a physical cosmology that was loosely derived from Peripatetic and Stoic physics, though most of its literature concerned niceties of prognostication, not the analysis of cause and effect. The Stoic poet Aratus's versified description of the constellations achieved remarkable popularity in antiquity; but on the whole the Stoics tended to disregard technical astronomy, perhaps because they were uncom-

fortable with its mechanistic character. The Skeptical schools, on the other hand, found an easy target in astrology's pretensions to exact knowledge of the future derived from inexactly observed or calculated motions of the heavenly bodies.

The Hellenistic period was the heyday of Greek geometry. Euclid, Archimedes, Apollonius of Perge, and a host of lesser mathematicians published work of enduring value on difficult problems, typically involving the properties of curves and the areas and volumes bounded by geometrical figures. Much mathematical research was motivated by optics, mechanics, and astronomy, but Hellenistic mathematicians seem to have kept more aloof from the philosophers than their predecessors of Plato's and Aristotle's time.

Alone among the scientific disciplines, medicine was characterized in the Hellenistic period by a division into sects or schools, comparable to the contemporary emergence of the great philosophical schools. The Hellenistic medical sects took their start from the prolific early third century physicians Herophilus and Erasistratus, whose theoretical pronouncements on physiology and medical practice were founded on a level of anatomical research and experimentation (reportedly including human vivisections) that was unprecedented in Greek medicine. In their approaches to physical and biological explanation these men and their followers owed something to Aristotle and perhaps more to the later Peripatetics. The "Herophilean" and "Erasistratean" schools seem to have less direct engagement with Stoicism or Epicureanism, though in common with those philosophical sects they accepted that knowledge of hidden causes of phenomena was both possible and useful.

The medical sect of Empiricists, which rose in the third century, rejected hidden causes as both unknowable and unhelpful in medical practice, and advocated instead a strategy for progressing systematically from individual trial-and-error experience to generalized, teachable practical knowledge without recourse to anatomical or physiological theory. The debates between the Empiricists and the other sects, grouped under the heading of Rationalists or Dogmatists, centered on both epistemology and research ethics; Empiricist physicians found natural intellectual allies in the philosophical Skeptics, especially the Pyrrhonists. Few Hellenistic physicians, however, were themselves philosophers, and a broad, intellectually respectable effort to bring together the many threads of current medical and philosophical thought had to wait for Galen in the second century CE. Galen's contemporary Ptolemy had a comparable reintegrating role with respect to Hellenistic physical science and philosophy.

The relationship between philosophy and medicine was paralleled by that between philosophical analysis of language and the emerging disciplines of grammar and philology. While critical speculation about language began in the Presocratic period and developed dramatically in the fourth century BCE, in the Hellenistic era the study of language achieved greater independence from philosophy without fundamentally severing its ties. Pergamum and Alexandria became centers for the critical study of ancient texts, especially Homer, and for the analysis of linguistic phenomena. At the same time, Epicureanism promoted a naturalistic understanding of the origin and nature of language and the Stoics made enormous advances not just in the area of logic (Chrysippus developed propositional logic in contrast to Aristotelian term logic) but also in the analysis of the parts of speech and semantic theory. Philosophers and grammarians debated the roles of rule-driven morphological *analogy* and the variability of actual linguistic usage (*anomaly*) in the determination of linguistic norms. Here too Hellenistic developments laid the foundations for intellectual life in later antiquity.

At the end of the Hellenistic era, the dominance of Athens in Greek philosophical life came to an end. After the conquest of Athens by the Romans under Sulla during the Mithridatic wars (88–86 BCE), philosophy, like science, spread out around the Mediterranean world. Rome itself, as well as Alexandria and Rhodes, became an important locus of philosophical activity as the Hellenistic age, and with it the Roman Republic, came to an end. At the beginnings of the Roman Empire, philosophy changed its character and turned for inspiration to the close study of the classic texts of Plato and Aristotle written centuries before. The Hellenistic era in Greek thought came to an end appropriately with the rise of a productive form of scholasticism and the revival of the classical schools of thought which have remained central to our understanding of ancient philosophy ever since.

See also Ancient Skepticism; Arcesilaus; Aristotelianism; Carneades; Chrysippus; Cyrenaics; Epicurus; Greek Academy; Panaetius of Rhodes; Posidonis; Stoicism; Strato and Stratonism; Theophrastus; Zeno of Citium.

Bibliography

For philosophical developments in the period the starting point is *The Cambridge History of Hellenistic Philosophy* (ed.

K. Algra et al, Cambridge, U.K.: Cambridge University Press, 1999). Major texts and philosophical commentary are found in A. A. Long and D. N. Sedley, *The Hellenistic Philosophers* (2 vol.; Cambridge, U.K.: Cambridge University Press, 1987). For the sciences, G. E. R. Lloyd's *Greek Science after Aristotle* (London: Chatto and Windus, 1973) remains fundamental. For optics and its philosophical setting see Gérard Simon, *Le regard l'être et l'apparance dans l'optique de l'antiquité* (Paris: Seuil, 1988). H. von Staden, *Herophilus: the art of medicine in early Alexandria: edition, translation and essays* (Cambridge, U.K.: Cambridge University Press, 1989) and S. Cuomo, *Ancient Mathematics* (London: Routledlge, 2001) illuminate the intellectual world of Hellenistic medicine and mathematics.

Brad Inwood (2005)
Alexander Jones (2005)

HELMHOLTZ, HERMANN LUDWIG VON

(1821–1894)

Hermann Ludwig von Helmholtz, the German physiologist and physicist, was born in Potsdam and educated at the Potsdam Gymnasium, where his father taught philology and classical literature, and at the Royal Friedrich-Wilhelm Institute of Medicine and Surgery in Berlin, from which he graduated as a doctor of medicine at the age of twenty-one. Helmholtz's outstanding scientific talent led to the curtailment of his required ten-year service as a Prussian army physician and surgeon. After the presentation and publication of his famous paper *Über die Erhaltung der Kraft* (On the conservation of energy) in 1847, he held only academic posts. He was instructor in anatomy at the Academy of Arts in Berlin (1847–1848), professor of pathology and physiology at Königsberg (1848–1855), professor of physiology and anatomy at Bonn (1855–1858), professor of physiology at Heidelberg (1858–1871), professor of physics at Berlin (1871–1888), and the first president and director of the Physico-Technical Institute in Berlin from 1888 until his death.

Helmholtz contributed over two hundred papers and books of outstanding importance in medicine, anatomy, physiology, psychology, and physics. He also published papers in mathematics and in philosophy, and delivered many popular lectures to publicize significant scientific investigations and to point out their philosophical implications. He was the first to measure the speed of nerve impulses, and he invented the ophthalmoscope. His paper *Über die Erhaltung der Kraft* became the cornerstone of the science of thermodynamics and set the direction of much of physics for the next half century. His monumental three-volume *Handbuch der physiologischen Optik* (*Handbook of Physiological Optics,* 1856–1866), frequently called the principia in its field, was matched in 1863 by a work equally basic to physiological acoustics, *Die Lehre von dem Tonempfindung* (*On the Sensations of Tone as a Physiological Basis for the Theory of Music*). In mathematics he was a pioneer in the field of non-Euclidean geometry, arriving independently at conclusions similar to those of Bernhard Riemann and seeing more quickly than others the philosophical importance of these new developments. In physics he contributed substantially to the establishment of the Faraday-Maxwell conception of electrical phenomena, both by his own theoretical investigations and through his encouragement of his most famous student, Heinrich Hertz. Helmholtz greatly influenced the intellectual climate in many German universities, and he may rightly be considered one of the fathers of the philosophy of science.

EMPIRICISM AND OPPOSITION TO METAPHYSICS

Helmholtz wrote only one long essay, *Die Tatsachen in der Wahrnehmung* (The facts of perception; 1878), that he explicitly considered to be in the field of philosophy. Most of his philosophy is contained in a number of short, popular essays and in the body of his various scientific works. The scientific works, however, contain frequent passages of philosophical importance and always show a clear awareness of philosophical issues. Furthermore, many of his papers on science and mathematics, such as those on the foundations of physics and mathematics, would now be included in the philosophy of science.

Helmholtz's philosophy was at all times closely related with his scientific investigations. One of the motives for the work that led to the paper *Über die Erhaltung der Kraft* was his desire to discredit vitalism as a scientific hypothesis and as a metaphysical position. Indeed, from the beginning of his career he was opposed in general to metaphysical speculation, feeling that the idealists, Friedrich Schelling and G. W. F. Hegel in particular, and a number of materialists had perverted philosophy and turned it from its main function, which was the study of human knowledge. Helmholtz was close to Immanuel Kant in his philosophy; he believed that in *The Critique of Pure Reason* Kant had asked the right questions and had moved part of the way toward answering them. He was also close to the classical British empiricists, believing that a scientifically and mathematically sophisticated empiricism along the lines initiated by John Locke would

provide highly reliable answers to a number of the basic questions of philosophy.

KNOWLEDGE AND PERCEPTION

For Helmholtz the central questions in philosophy were "In what ways do our ideas correspond to reality?" and "What is true in our sense perception and thought?" Answering these questions was the common task of both philosophy and the sciences, the two disciplines approaching them from opposite directions. The task of philosophy is to study the formal aspects of knowledge, our forms of intuition and representation, and the general categories in terms of which we order knowledge. The task of the sciences is to study the world of reality and to find the laws of nature that cause or determine both objective sequences of events and the sensations we experience. The formal aspects of knowledge, our forms of intuition and representation, and our intellectual categories, condition the ways in which we should and do formulate scientific knowledge. Scientific investigations, specifically the findings of physiological optics and acoustics, help us to understand our forms of intuition and the mental operations involved in knowing.

Although Helmholtz's position was basically Kantian, it was markedly different from Kant's on certain important points because of Helmholtz's study of physiological optics, physiological acoustics, and non-Euclidean geometry. His answer to the question "In what ways do our ideas correspond to reality?" was based upon certain discoveries in the physiology of sensation and, in particular, upon the principle of specific nerve energies. This principle was implicit in the psychological theories of a number of British empiricists; it was made explicit by Johannes Müller and was extended significantly by Helmholtz. Fundamental to this view is the theory that all we know about the external world is brought to consciousness as the result of certain changes produced in our sense organs by external causes. These changes are transmitted by the nerves to the brain, where they first become conscious sensations. In the brain they are interpreted and combined to produce our perceptions of external objects by mental processes that Helmholtz called unconscious inferences—processes he considered to be the same as those that are operative when a child learns his native language. Thus, in the case of vision, excitations of the nerves of the retina are transmitted by the optic nerve to the brain, where they are experienced as sensations and where they are unconsciously interpreted and combined to form visual perceptions of objects and their properties.

According to the principle of specific nerve energies, there is no one-to-one correspondence between a sensation experienced and a specific property of the object causing that sensation. It is perfectly possible for similar or identical sensations to be the effects of diverse causes or for a single cause, because it affects more than one kind of nerve, to result in qualitatively distinct sensations. As a result, the most that can be claimed is that sensations are caused by external objects, that they are the subjective signs of these objects and their properties, but are in no way images of them. The relation is one of sign to object signified, and even so, as such it is not an invariant relation. The only exception—an important one—is the correspondence in temporal sequence between external events and subjective sensations. Indeed, it is this correspondence that enables the scientist to determine the order of external events—that is, to determine the invariant laws of nature.

Because, with the notable exception of temporal sequences, there are no invariant, but only fairly uniform, relations between the sensations we experience and the objective world, Helmholtz felt that we can speak of our ideas as true only in a practical sense. Sensations are signs that we learn to use in order to regulate our movements and actions. When we have learned to interpret these signs, we are able to control our actions and are able to bring about results we desire or to avoid dangers.

> [To ask, however,] whether the idea I have of a table, its form, strength, colour, weight, etc., is true *per se,* apart from any practical use I can make of this idea, and whether it corresponds with the real thing, or is false and due to an illusion, has just as much sense as to ask whether a certain musical note is red, yellow, or blue. Idea and thing conceived evidently belong to two entirely different worlds, which no more admit of being compared with each other than colours and musical tones or than the letters of a book and the sounds of the words they form. (*Handbook of Physiological Optics,* Vol. III, p. 19)

SPACE AND GEOMETRY

Helmholtz's study of perceptions of space and of spatially oriented objects led him into the field of non-Euclidean geometry. His interest in general problems of spatial perception led to the investigation of the analytic properties that any space must have in order to permit the establishment of congruence relations between bodies and surfaces. As he saw it, congruence can be established only if rigid bodies or systems of bodies can be moved toward

one another with unaltered form—that is, only if the congruence of geometrical figures is a relation independent of all movements in space. Thus, he took the actual fact of spatial measurements through the establishment of congruence as a starting point and investigated the most general analytical properties of any space in which the movements necessary for this measurement can occur. He found that such movements and measurements were possible not only in Euclidean space but also in the spaces investigated by Riemann and Nikolai Ivanovich Lobachevski or in any space with a constant measure of curvature. Helmholtz concluded that Kant was mistaken in claiming that the axioms of Euclidean geometry were synthetic a priori principles necessarily true of space. Spaces that are not Euclidean can be conceived; the geometries of these spaces can be formulated; models or interpretations of them can be given, and on the basis of experience, it is impossible to determine which of these geometries is that of real space. Kant was correct in considering space to be a form of intuition but wrong in claiming that space must necessarily possess Euclidean characteristics.

PHILOSOPHY OF PHYSICS

Helmholtz's philosophy of physics was a classic formulation of nineteenth-century mechanism. He felt that the primary function of the physical sciences was to search for laws that express observed particular processes in general terms, so that from these laws other particular processes could be deduced. The discovery of these laws is the task of experimental physical science.

> The theoretical part … endeavors to ascertain the unknown causes of processes from their visible effects, it seeks to comprehend them according to the law of causality… . Thus, the final goal of the theoretical natural sciences is to discover the ultimate invariable causes of natural processes. (*Über die Erhaltung der Kraft,* introduction)

According to Helmholtz, these ultimate causes are simple Newtonian forces, so that a causal explanation in physics is at the same time an explanation in terms of forces.

> Theoretical natural science must, therefore, if it is not to rest content with a partial view of the nature of things, take a position in harmony with the present conception of the nature of simple forces and the consequences of this conception. Its task will be completed when the reduction of phenomena to simple forces is completed, and when it can at the same time be proved that the reduction given is the only one possible which the phenomena will permit. This will then be established as the necessary conceptual form for comprehending nature, and we will then be able to ascribe objective truth to it. (Ibid.)

These statements always represented the ultimate aim of scientific explanation for Helmholtz. It was a goal that grew more distant as the nineteenth century advanced.

See also Geometry; Hegel, Georg Wilhelm Friedrich; Hertz, Heinrich Rudolf; Kant, Immanuel; Philosophy of Physics; Philosophy of Science; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

WORKS BY HELMHOLTZ

Über die Erhaltung der Kraft. Berlin: Reimer, 1847.

Handbuch der physiologischen Optik. 3 vols. Leipzig, 1856–1866. Edited and translated by J. P. C. Southall as *Handbook of Physiological Optics.* New York, 1924–1925.

Die Lehre von dem Tonempfindung. Leipzig, 1863. Translated by A. J. Ellis as *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* New York: P. Smith, 1948.

Populäre wissenschaftsliche Vorträge. 2 vols. Leipzig, 1865–1871. Translated as *Popular Lectures on Scientific Subjects, First and Second Series.* New York, 1873–1891.

Die Tatsachen in der Wahrnehmung (The facts of perception). Berlin, 1878.

Schriften zur Erkenntnistheorie. Edited by Paul Hertz and Moritz Schlick. Berlin: J. Springer, 1921.

WORKS ON HELMHOLTZ

Crombie, A. C. "Helmholtz." *Scientific American* 198 (3) (1958): 94–102.

Koenigsberger, Leo. *Hermann von Helmholtz.* 3 vols. Berlin, 1902–1903. Translated by F. A. Welby. Oxford, 1906.

McKendrick, John Gray. *Hermann Ludwig Ferdinand von Helmholtz.* New York: Longmans, Green, 1899.

Russell Kahl (1967)

HELVÉTIUS, CLAUDE-ADRIEN
(1715–1771)

Claude-Adrien Helvétius was born into a highly respected medical family; his father was first physician to the queen of France. After his education at the College Louis-le-Grand and at the age of only twenty-three, Helvétius obtained, through influence at court, the lucra-

tive post of *fermier-général*, in which he soon grew rich. He became known, however, for the philanthropic and enlightened uses he made of his great wealth, particularly as a patron of philosophers and men of letters. For a time Helvétius turned to poetry and, in a piece titled *Le bonheur*, extolled the supreme pleasures of the intellectual life. Taking his own advice, he resigned in 1751 from tax-farming, married, and retired to his country estate, thenceforth devoting himself primarily to philosophical and literary pursuits. The publication in 1758 of his principal work, *De l'esprit*, proved to be one of the ideological *causes célèbres* of the eighteenth century. Appearing at a moment of political reaction, *De l'esprit* was noisily condemned by the authorities, both ecclesiastical and ministerial, for its dangerously heretical and subversive opinions. Suppression of the book signaled a grave—but fortunately temporary—setback for the party of *philosophes* and Encyclopedists. Despite the recantations that Helvétius was forced to make regarding *De l'esprit*, he reaffirmed his ideas even more strongly in *De l'homme, de ses facultés intellectuelles, et de son éducation*, published posthumously in 1772.

The thought of Helvétius sprang mainly from the predominant current of sensualism in the Enlightenment, which he fashioned with marked originality into what may be described as a thoroughgoing doctrine of "environmental behaviorism." Like John Locke, he held that the primary function of the mind was the registering of sense impressions arising from the external world. Calling this faculty *sensibilité physique*, he held it to be the exclusive source not only of all ideas and judgments (which, he said, resulted from the mental comparison of sensations) and of memory (which is simply a weakened sensation), but of all the emotions as well; he described emotions as variations of the two root-sensations of pleasure and pain that usually accompany sensory experience. Diverging, however, from the mainstream of Lockeanism represented in France at the time chiefly by Étienne Bonnot de Condillac, Helvétius was concerned with deriving from sensationist premises a psychological, rather than epistemological, theory. His basic contention was that the intellectual and moral capabilities, no less than the entire complex of values and motivations peculiar to any individual, are to be explained solely as the product of education—that is, of the total cumulative environment from the moment of birth. The biological or hereditary influences on the individual are considered, by contrast, to be uniform in all who are "normally constituted"; moreover, it is deemed unnecessary to include such "constants" in the causal investigation of behavior. Whereas Locke had rejected innate ideas, Helvétius proceeded to the rejection of innate abilities, or, more exactly, of innate inequalities in the apparatus of natural talents and inclinations with which every normal person is endowed. In expounding this radically environmentalist psychology, he did not feel obliged to assume any metaphysical position concerning the substantive nature of mind. Nevertheless, it is clear enough, as the concept of *sensibilité physique* implies, that Helvétius, swayed perhaps by contemporary materialists such as Denis Diderot and Julien Offray de La Mettrie, had in fact adopted a naturalistic view of man as logically the most suitable context in which to develop his behaviorist thesis.

In the mass of empirical evidence, often valuable and sometimes penetrating, adduced in support of the absolute case for environment versus heredity, two arguments stand out. Aware that exceptional talents do not always result from equally exceptional educations, Helvétius attempted to justify environmentalism by introducing the notion of *hasard*. According to this, a trivial and chance occurrence—such as Isaac Newton's observation of an apple falling to the ground—could, under certain conditions, have the most far-reaching consequences—such as the discovery of universal gravitation. Helvétius reached the reassuring, if tenuous, conclusion that genius is common to all persons, but the special sequence of events needed to actualize it is exceedingly rare. His other and much more plausible line of reasoning affirms that the development of intelligence, talent, or any ability whatever is proportional to the degree of *passion*, or emotional motivation, felt by the individual. Having thus translated the problem of inborn inequalities of mental capability into one of inborn emotive potential, Helvétius ended by finding that anyone may, by the appropriate stimulation of his passions, be rendered superior in any field of endeavor.

Throughout his writings, Helvétius emphasized the ethical and social implications of his psychology. The conception of the human being as a sort of behavioral tabula rasa, uniformly malleable by external controls into whatever forms might be judged desirable, favors obviously the practical ideal of reforming man and society on a grand scale. It devolves upon the "legislator-philosopher" to achieve this goal by a system of reeducation based on the scientific knowledge of the mechanics of behavior in relation to environment—an ambition aptly summarized by the sanguine dictum *l'éducation peut tout*. The supreme ethical criterion is, in turn, described as "public utility" (or such equivalents as *intérêt général, bonheur général*), which Helvétius further defined, in keeping with the pleasure-pain principle, as the maxi-

mum of pleasure compatible with a minimum of pain in the whole of a given society. But individual conduct in any environment is already determined by the pleasure principle, or self-interest, and Helvétius tried to show that, in addition, the multiplicity of pleasurable ends that men automatically seek, however "nonmaterial" in appearance, are all ultimately reducible to *sensibilité physique*. It is futile, therefore, to try to inculcate the social virtues by mere moralizing and even more so by condemning pleasure. Men can be changed for the better only through the manipulation of their passions. The key idea, to which all of Helvétius's thinking leads, is that the moral improvement and happiness of humankind can result only from political reforms having as their object the establishment of a system of public education (in both the narrow and the broad sense), by means of which the closest possible linkage would be effected between any individual's socially beneficial acts and his rewards in the form of gratified *sensibilité physique*. For example, Helvétius suggested that society methodically offer the choicest sexual enjoyments to its most virtuous and useful members. More generally and especially in the long polemical sections of *De l'homme*, he inveighed against what he regarded as the two major obstacles to the triumph of a hedonistic ethics founded on the standard of public utility—namely, Christianity with its irrational dogmas and ascetic, otherworldly morality and the feudal structure, economic inequities, and autocratic practices of the *ancien régime*. An aura of agnosticism, no less than a revolutionary fervor, surrounds the writings of Helvétius; while he does not argue the philosophical case for atheism, he patently assigns no positive value to a belief in God and even finds it pernicious to the *bonheur général* that he envisions.

The principal weakness of Helvétius's philosophy is its one-sided, reductive use of the factor of environment, at the expense of physiological predisposition—a defect that provoked a solid refutation from Diderot. Stated in absolute form, environmental behaviorism is paradoxical. Because no two educations, as Helvétius himself admitted, can ever be even remotely similar, it is impossible either to prove or to disprove his basic supposition that the same environmental causes will invariably produce the same behavioral effects. The representation of the mind as essentially passive sets up, moreover, a false dichotomy between it and the natural as well as social environment in which, from a more realistic standpoint, the individual is perceived to be a peculiarly dynamic and, indeed, creative participant. Finally, the ideal of public utility, when considered positively, remains rather vague and elusive, despite the abuses of the *ancien régime* attacked in its name. But historically the many valid elements of Helvétius's thought exerted considerable influence in several directions; on the Encyclopedist movement, especially Baron d'Holbach; on Pierre Cabanis and the *idéologues*; on the British utilitarians, particularly Jeremy Bentham; and, in a general long-range sense, on the rise of both democratic and socialistic doctrines and on the growth in modern societies of the comprehensive role given to public education.

See also Agnosticism; Bentham, Jeremy; Cabanis, Pierre; Condillac, Étienne Bonnot de; Diderot, Denis; Encyclopédie; Enlightenment; Ethics, History of; French Philosophy; Holbach, Paul-Henri Thiry, Baron d'; Innate Ideas; La Mettrie, Julien Offray de; Locke, John; Newton, Isaac; Philosophy of Education, History of; Psychology; Utilitarianism.

Bibliography

PRINCIPAL WORKS BY HELVÉTIUS

De l'esprit; or Essays on the Mind and Its Several Faculties. London, 1759.

A Treatise on Man; His Intellectual Faculties and His Education. Translated by William Hooper. London, 1777.

Oeuvres complètes, 3 vols. Paris, 1818.

Notes de la main d'Helvétius. Edited by Albert Keim. Paris, 1907.

WORKS ON HELVÉTIUS

Bloch, Jean H. "Rousseau and Helvetius on Innate and Acquired Traits: The Final Stages of the Rousseau-Helvetius Controversy." *Journal of the History of Ideas* 40 (1979): 21–41.

Cumming, Ian. *Helvétius: His Life and Place in the History of Educational Thought*. London: Routledge and Kegan Paul, 1955.

Hardy, Henry, and Isaiah Berlin, eds. *Freedom and its Betrayal: Six Enemies of Human Liberty*. Princeton: Princeton University Press, 2002.

Horowitz, Irving. *Claude Helvétius: Philosopher of Democracy and Enlightenment*. New York: Paine-Whitman, 1954.

Keim, Albert. *Helvétius, sa vie et son oeuvre*. Paris: Alcan, 1907.

Momdjian, K. *La philosophie d'Helvétius*. Translated from the Russian by M. Katsovitch. Moscow, 1959.

O'Neal, John C. "The Authority of Experience: Sensationist Theory in the French Enlightenment." University Park: Pennsylvania University Press, 1996.

Smith, David. W. "Helvetius and the Problems of Utilitarianism." *Utilitas: A Journal of Utilitarian Studies* 5(2) (1993): 275–290.

Smith, G. W. "Freedom and Virtue in Politics: Some Aspects of Character, Circumstances and Utility from Helvetius to J. S. Mill." *Utilitas: A Journal of Utilitarian Studies* 1 (1989): 112–134.

Aram Vartanian (1967)
Bibliography updated by Tamra Frei (2005)

HEMPEL, CARL GUSTAV

(1905–1997)

Carl Gustav Hempel was born in Germany, immigrated to the United States, and became a naturalized citizen. He taught at Yale, Princeton, and Pittsburgh. Along with Sir Karl Popper and Thomas S. Kuhn, a former colleague, he would become one of the most important philosophers of science of the twentieth century. Popper exerted more influence upon natural scientists and Kuhn upon social scientists and the public alike, but Hempel's impact upon professional philosophers of science was unparalleled. His work, including the problems he addressed and the methods he employed, virtually defined the philosophy of science, not just for a few years, but for several decades.

Hempel sought solutions to philosophical problems that were not only well-supported by suitable arguments but which were also precisely formulated by means of symbolic logic. He proposed subtle and nuanced formulations of scientific philosophy and promoted the transition from *logical positivism* to what would become known as *logical empiricism*. Hempel was committed to extremely high standards of philosophical clarity and rigor, which enabled his explications to be subject to the most demanding inspection and critical examination. He cared more about finding the right solutions than whether his own solutions were right.

LOGICAL POSITIVISM

Thus, "logical positivism," the leading movement of the 1930s and 1940s, was based on three principles: the analytic/synthetic distinction; the observational/theoretical distinction; and the verifiability criterion of meaningfulness. Logical positivism thus affirmed that all *a priori* knowledge is analytic and that all synthetic knowledge is *a posteriori*, denying the existence of knowledge that is both synthetic and *a priori*. Sentences that are nonanalytic but also nonverifiable, including various theological and metaphysical assertions concerning the divine or the absolute, thereby qualify as cognitively meaningless.

The precise manner in which scientific theories are to be related to experience therefore became a crucial issue. *Observation language* is assumed to consist of names and predicates whose applicability or nonapplicability, under suitable conditions, could be ascertained by means of direct observation or relatively simple measurement. *Theoretical language*, which makes reference to nonobservables, such as malleability and conductivity as well as electrical fields and gravitational forces, must therefore either be reducible to observables or is empirically meaningless.

COGNITIVE SIGNIFICANCE

Hempel (1950, 1951) demonstrated that empirical knowledge was thereby restricted to observation sentences and their deductive consequences, which reduces scientific theories to mere logical constructions from observables. In articles on cognitive significance and empirical testability, he persuasively demonstrated that the verifiability criterion implies that *existential generalizations* are meaningful, but that *universal generalizations* are not, even though they include general laws, the principal objectives of scientific discovery.

Moreover, on the assumption that a sentence S is meaningful if and only if its negation is meaningful, Hempel demonstrated that implementing the verifiability criterion generates inconsistent consequences. The sentence, "At least one stork is red-legged," for example, is *meaningful* because it can be verified by observing one red-legged stork; yet its negation, "Not even one stork is red-legged," cannot be shown to be true by observing any finite number of red-legged storks and is therefore *meaningless*. Assertions about relative frequencies in finite classes and their negations are meaningful, but those about limits in infinite sequences are not.

SCIENTIFIC THEORIES

These realizations suggested that the logical relationship between scientific theories and empirical evidence cannot be exhausted by means of observation sentences and their *deductive consequences* alone, but needs to be expanded to include observation sentences and their *inductive consequences* (1958). The concepts of confirmation and disconfirmation (as forms of partial verification and partial falsification) warrant renewed attention, where the crucial feature of scientific hypotheses is their empirical testability rather than their verifiability.

Hempel (1960) argued further that the application of inductive logic supports certain logically impeccable, but psychologically surprising, consequences, such as that the observation of a white shoe confirms the hypothesis that all ravens are black because it is an instance of the hypothesis that everything is either not a raven or black, which, using extensional language, is logically equivalent to all ravens are black. And he proposed that cognitive significance should best be envisioned as a matter of degree that may only be evaluated relative to multiple criteria (1965a).

FIGURE 1

The Covering Law Model

PREMISES	General Law(s):	L1, L2, . . . , Ln	EXPLANANS
	Initial Conditions:	C1, C2, . . . , Cm	
CONCLUSION	Event Description:	E	EXPLANANDUM

FIGURE 2

Inductive-Probabilistic Explanation

The Statistical Law:	P(B/A) = r	
Initial Conditions:	Ax	
	═══════	[r]
The Explanandum:	Bx	

DISPOSITIONS AND DEFINITIONS

In *Fundamentals of Concept Formation in Empirical Science* (1952) he addresses the problem of definability in relation to dispositional predicates, such as "malleable," "soluble," and "magnetic," which designate, not directly observable properties, but rather tendencies on the part of some things to display specific reactions (say, attracting small iron objects) under specific circumstances (the presence of small iron objects in the vicinity). On first consideration, it might seem appropriate to define this predicate by means of a formulation employing a conditional: "x is magnetic at t" is taken to mean, "if, at t, a small iron object is close to x, then it moves toward x."

Interpreted as a *material conditional*, whose meaning is synonymous with "either not … or ____," however, the proposed definition would be satisfied by things not subject to the test condition at all—such as brown cows—when there are no small iron objects in their vicinity. This result threatened the integrity of the project of developing an adequate philosophical framework for understanding the language of science. Both Hempel and Rudolf Carnap displayed great ingenuity in employing the resources of formal logic to cope with it. Ultimately, Carnap would embrace intensional logic as the solution, but Hempel preferred extensional logic, which imposed methodological boundaries upon explications he found acceptable.

EXPLICATIONS OF EXPLANATION

Hempel's most important contribution to the philosophy of science, no doubt, was his masterful explication of the structure of *scientific explanations* as a refinement of the theory of explanation by means of subsumption by general laws, an approach whose precursors date from Aristotle. Hempel developed this conception by means of his "covering law" model, which he elaborated in several versions, understood as arguments whose premises ("the explanans") include at least one general law, Li, which explain why the event that is described by the conclusion ("the explanandum") occurred by showing it was to be expected relative to its initial conditions, C1-Cm (Hempel and Oppenheim 1948).

Thus Hempel presented a schema that has become familiar to generations of graduate students of the philosophy of science, which incorporated those conditions as follows in Figure 1. A simple example might explain why a small coin expanded when heated by invoking the law that copper expands when heated and noting it was copper. Hempel considered a vast variety of modes of explanation, contending that those which—implicitly or explicitly—conform to this conception are scientific.

INDUCTIVE-PROBABILISTIC EXPLANATIONS

Hempel included explanations of empirical generalizations by laws and of laws by theories within the scope of his approach, but devoted most of his attention to elaborating several precise and detailed accounts of the scientific explanation of singular events. And he advanced *deductive-nomological* and *inductive-probabilistic* versions to account for differences between subsumption by universal and by statistical covering laws. The differences between them, especially the peculiar difficulties generated by probabilistic explanations, would preoccupy much of his efforts for more than two decades, including Hempel (1948, 1965b, 1968).

The crucial problem turned out to be that of the rationale for the logical link between explanans and explanandum when the covering laws were not universal but statistical. Suppose, for example, that a statistical law of the form, P(B/A) = r, assigned probability of value r to the occurrence of an outcome of kind B, given conditions of kind A. Then an explanation of the form (see Figure 2),

invites the presumption that the bracketed variable [r] should be understood as a measure of evidential support. Hempel initially adopted such an approach, which reflects an epistemic interpretation of [r], but he would subsequently reject it on the grounds that the truth of the explanandum is already known: what we want to explain is why it occurred (Hempel 1968).

While the covering law approach dominated the philosophy of science in the 1950s and the 1960s, such difficulties, which were rooted in deep problems about the nature of explanatory relevance and of probabilistic laws, stimulated other investigations, the most important being the *statistical relevance* model of Wesley C. Salmon, which denied explanations were arguments and captivated the discipline in the 1970s. Salmon would later abandon the interpretation of nomic probabilities as relative frequencies for the Popperian alternative of propensities as probabilistic dispositions in the context of probabilistic explanation. During the 1980s and the 1990s, no approach would exert its grip upon the discipline as had Hempel's covering-law model, which made explanation a central function of science.

THE PROBLEM OF PROVISOES

One of the most remarkable features of Hempel's career is that he continued to publish original and innovative papers well into the eighth decade of his life. He authored a series of studies that moved away from the standard conception of scientific theories as *formal calculi* combined with *empirical interpretations* and, in *Philosophy of Natural Science* (1966), a widely used introduction to the philosophy of science that would be translated into ten other languages, he even advanced the novel explication of scientific theories as consisting of *internal principles* and *bridge principles*, where the general hypotheses that distinguish a theory are connected to observation and experiment by principles expressed in mixtures of ordinary and of technical language, where antecedent understanding replaces explicit definability.

More strikingly, Hempel (1988) noted that the application of scientific theories presupposes the absence of factors that might affect the internal principles of the theory, which goes beyond the content of the theory itself. Deriving predictions and explanations from classical mechanics, for example, presupposes that bodies are being acted upon exclusively by gravitational forces, where the presence of electromagnetic forces would invalidate those derivations. The function of these provisoes means that instrumentalist constructions of scientific theories as mere calculating devices and programs for the elimination of theoretical language by reduction to observational language alone are misguided and cannot be sustained.

See also Carnap, Rudolf; Confirmation: Qualitative Aspects; Confirmation Theory; Explanation in Science; Explanation, Theories of; Kuhn, Thomas; Logical Positivism; Popper, Karl Raimund; Salmon, Wesley; Scientific Theories; Verifiability Principle.

Bibliography

WORKS BY HEMPEL

"The Function of General Laws in History." *Journal of Philosophy* 39 (1942): 35–48.

"Studies in the Logic of Explanation." With P. Oppenheim. *Philosophy of Science* 15 (1948): 135–175.

"Problems and Changes in the Empiricist Criterion of Meaning." *Revue Internationale de Philosophie* 41 (11) (1950): 41–63.

"The Concept of Cognitive Significance: A Reconsideration." *Proceedings of the American Academy of Arts and Sciences* 80 (1) (1951): 61–77.

Fundamentals of Concept Formation in Empirical Science. Chicago: University of Chicago Press, 1952.

"The Theoretician's Dilemma." In *Minnesota Studies in the Philosophy of Science*. Vol. 2, edited by H. Feigl, M. Scriven, and G. Maxwell, 37–98. Minneapolis: University of Minnesota Press, 1958.

"Inductive Inconsistencies." *Synthese* 12 (1960): 439–469.

"Empiricist Criteria of Cognitive Significance: Problems and Changes." 1965a. In Hempel (1965c), 101–119.

"Aspects of Scientific Explanation." 1965b. In Hempel (1965c), 331–496.

Aspects of Scientific Explanation. New York: Free Press, 1965c.

Philosophy of Natural Science. Englewood Cliffs, NJ: Prentice-Hall, 1966.

"Maximal Specificity and Lawlikeness in Probabilistic Explanation." *Philosophy of Science* 35 (June 1968): 116–133.

"Valuation and Objectivity in Science." In *Physics, Philosophy, and Psychoanalysis: Essays in Honor of Adolf Grunbaum*, edited by Robert S. Cohen and Larry Laudan, 73–100. Dordrecht, Netherlands: Kluwer, 1983.

"Provisoes: A Problem concerning the Inferential Function of Scientific Theories." In *The Limitations of Deductivism*, edited by Adolf Grunbaum and Wesley C. Salmon, 19–36. Berkeley: University of California Press, 1988.

Selected Philosophical Essays: Early and Late, edited by Richard C. Jeffrey. Cambridge, U.K.: Cambridge University Press, 2000.

The Philosophy of Carl G. Hempel: Studies in Science, Explanation, and Rationality, edited by James H. Fetzer. Oxford: Oxford University Press, 2001.

WORKS ON HEMPEL

Davidson, Donald, et al. *Essays in Honor of Carl G. Hempel*.Edited by Nicholas Rescher. Dordrecht, Netherlands: D. Reidel, 1969.

Esler, Wilhelm, et al., eds. *Epistemology, Methodology, and Philosophy of Science*. Dordrecht, Netherlands: Kluwer, 1985.

Fetzer, James H., ed. *Science, Explanation, and Rationality: Aspects of the Philosophy of Carl G. Hempel*. Oxford: Oxford University Press, 2000.

Salmon, Wesley C. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press, 1990.

James H. Fetzer (2005)

HEMSTERHUIS, FRANS

(1721–1790)

The Dutch philosopher Frans Hemsterhuis was born at Franeker, the son of the famous Greek scholar Tiberius Hemsterhuis. Frans Hemsterhuis was a clerk of the State Council and devoted his free hours to his favorite studies—numismatics, fine arts, and philosophy. In his last years his philosophy was very much influenced by his friendship with the Princess von Gallitzin, the wife of the Russian ambassador at The Hague. Thus, his life and work may be divided into two periods.

In the first period Hemsterhuis's *Lettre sur l'homme et ses rapports* (1772) was his principal work, preceded by two small, closely connected treatises, *Lettre sur la sculpture* (1765) and *Lettre sur les désirs* (1769). In *Lettre sur la sculpture* Hemsterhuis argued that the essence of the aesthetic experience is a longing to unite oneself with the art object. This concept became part of his theory of ethics, which is set out in *Lettre sur les désirs.* The most perfect happiness for the soul is the union with the beloved object irrespective of whether it is an object of art, a person, or God. This Platonic Eros is for Hemsterhuis analogous to the power of attraction in the physical world. This theory is further developed in *Lettre sur l'homme,* on which the Platonic dialogues of his second period are based. On the subject of the nature of man Hemsterhuis thought in terms of a dualistic philosophy like René Descartes's, but Hemsterhuis's dualism was combined with an empiristic-sensationalistic theory that he probably derived from John Locke and Étienne Bonnot de Condillac. Through sensory perception man receives an image of what exists in reality. This image, however, is incomplete, and if man had other organs, he could perhaps see other aspects of reality. Through what Hemsterhuis calls the "moral organ" man is aware of an immediate feeling of his relationship with God. The moral organ is also responsible for the feeling of relation, rapport, that man has with thousands of other men, and the development of such relations is dependent on the perfection of the moral organ. This theory leads to an individualistic concept of man's moral duties, which is one of the reasons for Hemsterhuis's influence on the German philosophy of *Sturm und Drang* and romanticism.

In the second period of Hemsterhuis's life he wrote four Platonic dialogues the most important of which are *Aristée ou de la divinité* (1779) and *Alexis ou de l'âge d'or* (1783, but published in 1787). In *Aristée* Hemsterhuis, who originally believed in a personal God, is converted to a clear pantheism. God's omnipresence is the basis of man's relation to him, and it is mainly thanks to the moral principle, as the "moral organ" is called in later years, that man is able to come nearer to God. In *Alexis* Hemsterhuis, perhaps influenced by contemporary German philosophy, presented for the first time his concept of the golden age and the harmonious development of the individual. He also introduced the notion of the value of poetical truth (truth discovered by the poet in moments of enthusiasm). With these ideas Hemsterhuis had moved far from his earlier rationalism, and his thought was received with admiration and approval by representatives of the *Sturm und Drang* and romantic movements in philosophy.

In the first period F. H. Jacobi and J. G. Herder were among Hemsterhuis's admirers; in the second period he was very popular with and influenced the two Schlegels and Novalis.

See also Aesthetic Experience; Condillac, Étienne Bonnot de; Descartes, René; Herder, Johann Gottfried; Jacobi, Friedrich Heinrich; Locke, John; Novalis; Rationalism; Romanticism; Schlegel, Friedrich von.

Bibliography

WORKS BY HEMSTERHUIS

During his lifetime most of Hemsterhuis's works were printed for private circulation in small and anonymous editions. Some of his early treatises are still in manuscript form in public collections. The latest complete, but rather inaccurate, edition is François Hemsterhuis, *Oeuvres philosophiques,* edited by Louis Susan Pedro Meijboom (Leeuwarden, Netherlands: W. Eekhoff, 1846). A German translation edited by Julius Hilsz was published at Leipzig (1912). A more recent edition of one of his works is François Hemsterhuis, *Lettre sur l'homme et ses rapports. Avec le commentaire inédit de Diderot,* edited by George May (New Haven, CT: Yale University Press, 1964).

WORKS ON HEMSTERHUIS

Brachin, Pierre. *Le cercle de Munster (1779–1806) et la pensée religieuse de F. L. Stolberg.* Lyon and Paris, 1952.

Brachin, Pierre. "Hemsterhuis' Beziehungen zum Gallitzinkreis." In *Pariser Universitätswoche an der Ludwig-*

Maximilians-Universität zu München vom 14. bis 19. Februar 1955, 203–216. Munich, 1955.

Brummel, Leendert. "Frans Hemsterhuis." *Algemeen Nederlands Tijdschrift Voor Wijsbegeerte en Psychologie* 34 (1940): 17–26.

Brummel, Leendert. *Frans Hemsterhuis. Een philosophenleven.* Haarlem: H. D. Tjeenk Willink, 1925.

Loos, Waltrand. "Der Briefwechsel des Philosophen Hemsterhuis mit der Fürstin Gallitzin, mit einem Schlüssel zu seiner Geheimschrift." *Westfalen* 39 (1961): 119–127.

Leendert Brummel (1967)

HEN/POLLA

In one form or another the problem known as that of the One (*Hen*) and the Many (*Polla*) pervades the whole history of Greek philosophy. According to Aristotle (*Physics* I, 2–3), it arose first in the pre-Socratic inquiry into whether there is one first principle or source—for example, water alone, or air alone—for things, or whether there is more than one first principle. If we are to avoid "coming into being out of nothing," we must either deny with Parmenides that there is any multiplicity arising from the first principle, or else we must suppose that somehow or other multiplicity is already present within the unitary first principle. If we choose the second supposition, we are faced with a problem that is no longer purely physical, namely, how one thing can also be many.

In the period after Parmenides it became clear that this problem arose at a number of different levels of thought, though it usually seemed natural to suppose that a solution at one level would solve the problem at other levels as well. It arose in relation to the phenomenal world in three ways: How one thing can possess a number of different characteristics, how one thing can change into another, and how one thing can have many parts. It arose, above all for Plato, as a problem concerning metaphysical entities such as forms—how unitary forms can be split up among many particulars, and how one form can possess a number of attributes. It arose in the theory of predication as the question of how a number of predicates can be applied to a singular subject. It arose in logic especially as the problem of how classification of many things under one head can be justified. It was also discussed (by Aristotle in the *Physics*) in terms of the number of first principles, even when the question of their number was no longer seen from within a purely physical framework.

Only some of the more important treatments can be mentioned here. The earlier Pythagoreans subordinated One and Many to Limit and Unlimited in their table of opposites. It remained a standard problem for the Eleatics after Parmenides, and Zeno leveled some forty arguments against plurality, of which one or two survive. The Sophist Protagoras tackled the problem at the level of perception, and Gorgias's pupil Lycophron dealt with it as a problem of predication by banishing the word *is* in statements such as "Socrates is white"; Antisthenes, the Megarians, and the Eretrians, according to some, followed the same path. Plato discussed the question repeatedly, most explicitly in Part Two of the *Parmenides* and in the *Philebus,* but of vital importance is his approach to predication in the *Sophist.* His discussion was continued by Speusippus, and by Aristotle in the *Metaphysics* (I, 6) and in the *Physics,* where he propounded a solution in terms of "substrate," "privation," and "form." The Neoplatonists saw it as the problem of how the multiplicity of the world order can proceed from the ultimate absolute unity.

See also Antisthenes; Apeiron/Peras; Aristotle; Megarians; Parmenides of Elea; Plato; Protagoras of Abdera; Pythagoras and Pythagoreanism.

Bibliography

Virtually the whole of F. M. Cornford's *Plato and Parmenides* (London: K. Paul, Trench, Trubner, 1939) is concerned with the Hen/Polla problem in one way or another.

G. B. Kerferd (1967)

HENRY OF GHENT

(??–1293)

The Augustinian secular theologian Henry of Ghent, traditionally known as *Doctor Solemnis*, was born at Ghent or Tournai, probably in the second quarter of the thirteenth century. In addition to holding high ecclesiastical office at Bruges and Tournai, he taught both arts and theology at the University of Paris. In 1277 he served on the theological commission that prepared the condemnation issued by Stephen Tempier, bishop of Paris, against the Averroism of Siger of Brabant and Boethius of Dacia. He died in 1293.

Henry's principal writings are a *Summa Theologica* and fifteen *Quodlibeta* (occasional disputations). The extended criticism of his ideas by John Duns Scotus, William of Ockham, and others is a sign of his considerable influence in his own age. In the sixteenth century the

Servite friars chose him as their official theologian, although he had never belonged to their order.

As a philosopher, Henry of Ghent stood in the main line of development of medieval Platonism. The Augustinian tradition, already brilliantly represented in the thirteenth century by Bonaventure and Matthew of Acquasparta, was unmistakably the weightiest element in his thought, and the Platonic orientation thus established was further strengthened by the influence of Avicenna. At the same time, following Bonaventure and other earlier Augustinians, he incorporated a number of Aristotelian ideas into his synthesis. Furthermore, in adapting the Neoplatonic metaphysics of Avicenna to the requirements of the Christian view of God and creation, he anticipated certain critical tendencies of the fourteenth century, by which the whole structure of medieval realism was to be undermined. It is fair to say that Henry failed to blend these diverse elements into a fully coherent system. Nonetheless, in inspiration and aim he was the true precursor of Duns Scotus, the last great constructive philosopher of the Middle Ages.

BEING

For Henry of Ghent the starting point of metaphysical thinking was the idea of being (*ens* or *res* or *aliquid*), out of which the metaphysician draws the intelligible essences virtually contained in it. Analysis shows that being is an analogical idea. Taken in its widest sense, it includes both imaginary entities (*res secundum opinionem*), which exist only in the mind, and genuine beings (*res secundum veritatem*), which exist, or at any rate can exist, outside the mind. Genuine being, which is the proper object of metaphysics (*ens metaphysicum*), is further divided into Being Itself (*ipsum esse*), or God, and contingent beings, or creatures. Finally, creaturely being is divided into that which exists in itself (substance) and that which exists in another (accident).

Genuine beings, actual or possible, are distinguished from imaginary entities by their possession of "essential being" (*esse essentiae*). This essential being is not a rudimentary mode of existence. It is best described as an intrinsic possibility or intelligibility that pertains to definable essences as reflections of the divine ideas. It is to be contrasted with the intrinsic impossibility and incoherence of *res secundum opinionem*.

Actual beings are distinguished from merely possible beings by their possession of "existential being" (*esse existentiae*). This existential being is not a principle or act within actually existing things; Henry refused to accept the real distinction between essence and existence as formulated either by Giles of Rome or by Thomas Aquinas. The difference between essential and existential being is to be found not in things themselves but in the relation of essences to God. Essential being consists in being thought by God, while existential being consists in actually depending on God as creative Cause.

Apart from the fact that each actual being, as a product of divine creativity, is individually related to God, individual things, in which specific essences are multiplied, require no explanation of their individuality. Individuation involves no addition, whether of matter or of act of existence, to the intelligible essence. In analyzing the individual as such it is sufficient to say that it is internally undivided and is not identical in existence—that is, in its relation to the Creator—with any other individual.

DIVINE FREEDOM

The transition from essential to existential being, or the act of creation, is an act of divine freedom. Individual beings come into existence not from any intrinsic necessity but because God freely wills to create. Here the Christian and Augustinian conception of God's transcendent liberty excludes the Avicennian idea of the divine will as subject, equally with the divine intellect, to necessity. In his fear of compromising God's freedom in creation, Henry further minimized the intelligibility of individual beings. There are no divine ideas of singular things as such; God knows them only through their essences considered as multipliable in numerically distinct beings. Consequently, the existence of creatures can in no way be deduced from God's eternal ideas.

NECESSARY AND CONTINGENT BEING

The fundamental metaphysical notion of being is neither simply derived from sense experience nor strictly innate in the human mind. It is indeed formed by the mind from within, but on the occasion of sense experience. It would actually be more correct to say that two fundamental notions of being are formed by the mind, since the concepts of necessary or divine Being and contingent or creaturely being are radically distinct and cannot be deduced from a more general notion. When we conceive being unconfusedly, we always conceive either necessary Being or contingent being—never some undifferentiated, neutral being.

In thus asserting the irreducible duality of the notion of being, Henry was again trying to exclude any suggestion that God necessarily creates. If neither divine Being

nor creaturely being can be deduced from a universal concept of being, one argument for necessitarianism is effectively undermined. The further consequences of Henry's principle, however, were disastrous for his own metaphysical enterprise. His insistence that there is no positive content common to the two fundamental notions of being leaves a gap between divine Being and creaturely being that no mere affirmation of analogy between the two concepts can bridge. But if, as Henry claimed, there is some empirical factor in the formation of our notions of being, it is hard to see how necessary Being can be conceived, let alone demonstrated, as long as the gap remains. It is true that an a posteriori or "physical" proof of God's existence, based on experience of individual objects rather than on metaphysical principles, is possible, but such a proof can attain only to a supreme Being, not to a necessary Being. An appeal to divine illumination—the obvious remaining alternative—was excluded for Henry because he did not conceive of the divine light as a power impressing ideas upon the human mind. Although Henry refused to draw it, the conclusion seems inevitable that no firm basis can be found for a metaphysical theology.

With Henry of Ghent, medieval Platonism was clearly entering its final phase. In his thought, for all its predominantly Augustinian and Avicennian character, more or less novel concerns—a new stress on divine freedom, a fresh interest in sense experience, a new emphasis on sheer particularity—already modified the Platonic view of reality. Henry is thus a significant symbol of the transition from the constructive to the critical period of medieval thought.

See also Aristotelianism; Augustinianism; Avicenna; Being; Boetius of Dacia; Bonaventure, St.; Duns Scotus, John; Matthew of Acquasparta; Medieval Philosophy; Platonism and the Platonic Tradition; Siger of Brabant; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS

Series *Henrici de Gandavo Opera Omnia* (Ancient and Medieval Philosophy, Series 2), De Wulf-Mansion Centre.

Already Published

Macken, Raymond. *Bibliotheca manuscripta Henrici de Gandavo. I. Catalogue A-P*, (Henr. de Gand. *Opera Omnia*, I), Leuven/Leiden: Leuven University Press/E. J. Brill, 1979.

Macken, Raymond. *Bibliotheca manuscripta Henrici de Gandavo. II. Catalogue Q-Z*, (Henr. de Gand. *Opera Omnia*, II), Leuven/Leiden: Leuven University Press/E. J. Brill, 1979.

Quodlibeta: Quodlibet I. (edited by Raymond Macken); *Quodlibet* II (edited by Robert Wielockx); *Quodlibet* VI (edited by Gordon A. Wilson); *Quodlibet* IX (edited by Raymond Macken); *Quodlibet* X (edited by Raymond Macken); *Quodlibet* XII, q. 1–30 (edited by Jos Decorte); *Quodlibet* XII, q. 31, *Tractatus super facto praelatorum et fratrum* (edited by Ludwig Hödl and Marcel Haverals); *Quodlibet* XIII (edited by Jos Decorte).

Lectura ordinaria super S. Scripturam (edited by Raymond Macken).

Summa (*Quaestiones Ordinariae*): art. I–V (edited by Gordon A. Wilson); art. XXXV–XL (edited by Gordon A. Wilson); art. XLI–XLVI (edited by Ludwig Hödl).

STUDIES

Emery, Kent Jr. "The Image of God Deep in the Mind: The Continuity of Cognition according to Henry of Ghent." In *Nach der Verurteilung von 1277. Philosophie und Theologie an der Universität von Paris im letzten Viertel des 13. Jahrhunderts. Studien und Texte (After the Condemnation of 1277. Philosophy and Theology at the University of Paris in the Last Quarter of the Thirteenth Century. Studies and Texts)*, edited by Jan A. Aertsen, Kent Emery, Jr., and Andreas Speer. Berlin; New York: de Gruyter, 2001 (Miscellanea Mediaevalia, 28), 59–124.

Guldentops, Guy, and Carlos Steel, eds. *Henry of Ghent and the Transformation of Scholastic Thought. Studies in Memory of Jos Decorte*. Leuven: Leuven University Press, 2003. (Ancient and Medieval Philosophy, I/31), with Bibliography 1994–2002 [Studies by J. Decorte, C. Steel, J. Janssens, W. Goris, A. Speer, J. A. Aertsen, J.-L. Solère, C. Kann, R. Plevano, M. W. F. Stone, J.-M. Counet, S. P. Marrone, G. A. Wilson, J. C. Flores, G. Pini, M. S. de Carvalho, L. Hödl, P. Porro].

Macken, Raymond. *Essays on Henry of Ghent*. Leuven: Editions Medieval Philosophers of the Former Low Countries, Vol. 1, 1994; Vol. II, 1995; Vol. III, 1996; Vol. IV, 1998.

Macken, Raymond. "Henry of Ghent and Augustine." In *"Ad litteram": Authoritative Texts and Their Medieval Readers*, edited by Mark D. Jordan, and Kent Emery, Jr. Notre Dame; London: University of Notre Dame Press ("Notre-Dame Conferences in Medieval Studies,", 3), 1992, 251–274.

Macken, Raymond. "Synderesis and Conscience in the Philosophy of Henry of Ghent." *Franziskanische Studien* 70 (1988): 185–195.

Macken, Raymond. "The Superiority of Active Life to Contemplative Life in Henry of Ghent's Theology." *Medioevo* 20 (1994): 115–129.

Marrone, Steven P. *Truth and Scientific Knowledge in the Thought of Henry of Ghent*. Cambridge, MA: *The Medieval Academy of America* (Speculum Anniversary Monographs, 11), 1985.

Pasnau, Robert. "Henry of Ghent and the Twilight of Divine Illumination." *Review of Metaphysics* 49 (1995): 49–75.

Porro, Pasquale. "Metaphysics and Theology in the Last Quarter of the 13th Century: Henry of Ghent Reconsidered." In *Geistesleben im 13. Jahrhundert*, edited by Jan A. Aertsen and Andreas Speer. Berlin; New York: de Gruyter, 2000 (Miscellanea Mediaevalia, 27), 265–282.

Vanhamel, Willy, ed. *Henry of Ghent. Proceedings of the International Colloquium on the Occasion of the 700th Anniversary of His Death (1293)*. Leuven: Leuven University

Press, 1996 (Ancient and Medieval Philosophy, I/15), with Bibliography 1837–1993 [Studies by J. A. Aertsen, J. V. Brown, M. A. S. de Carvalho, J. Decorte, J. McEvoy, L. Hödl, J. Janssens, M. Laarmann, S. P. Marrone, P. Porro, B. B. Price, R. J. Teske, C. Trottmann, G. A. Wilson].

For further references, see Macken, Raymond, *Bibliographie d'Henri de Gand*, Leuven: Editions Medieval Philosophers of the Former Low Countries, 1994.

Eugene R. Rairweather (1967)
Bibliography updated by Maria Lucrezia Leone (2005)

HENRY OF GHENT [ADDENDUM]

Since the mid-1990s, remarkable progress has been made in the study of Henry of Ghent due to the ongoing publication of the critical edition of the *Opera Omnia.* Begun in 1979 at the De Wulf-Mansion Center of the Catholic University of Leuven, its labors have been complemented by two major international conferences held in 1993 and 2001 respectively.

The critical edition has not only made the texts of Henry much more accessible, but has directed scholarly attention to Henry's numerous (Greek, Latin, Arabic, and Hebrew) sources. Close attention to these sources has helped to settle several disputes concerning the status of Henry's works; for example the *Lectura ordinaria super sacram Scripturam,* the *Tractatus super facto praelatorum et fratrum,* and a number of sermons are now believed to be by Henry. However, the attribution of the *Syncategoremata* (ms. Brugge, Stadsbibl., 510, ff. 227ra-237va), and of a *Commentary* on the *Physics* (ms. Erfurt, Amplon. F. 349) and on the *Metaphysics* (ms. Escorial, h.ll.1) still remain matters of conjecture.

As a consequence of the published proceedings of the two conferences, Henry is no longer viewed solely as a "Platonic" and "neo-Augustinian" thinker, but as a thirteenth-century scholastic who possessed an astounding knowledge of the philosophy of Aristotle and Avicenna. While Henry's metaphysics and theology have been traditional areas of study, the conference proceedings began to focus attention on his ethics and economics, areas that await detailed investigation. Other recent studies have illuminated the facts of his biography. Henry was, most probably, born before 1240 (maybe between 1217 and 1223), was certainly in Paris by 1265, and from 1267 onward was registered as *magister* in the documents of Paris University.

Marialucrezia Leone (2005)

HENRY OF HARCLAY

(c. 1270–1317)

Henry of Harclay, the English scholastic theologian and philosopher, was born in the diocese of Carlisle. After studying at Oxford and Paris, he was ordained a priest in 1297 and obtained his master of theology about 1310. He taught at Oxford, becoming chancellor of the university in 1312. He wrote an unedited "Commentary on the *Sentences,*" and "Disputed Questions," most of which are unpublished. He died at Avignon.

Early in his career, while commenting on the *Sentences,* Henry defended the main theses of John Duns Scotus. Later, he criticized Scotism, teaching a doctrine of universals close to that of William of Ockham. He held that there are no common natures or essences in reality; there are only individuals, each of which has its own nature. Since there are no common natures, there is no need of the Scotist haecceity to render them individual. As Ockham later said, realities are individual not by an added "thisness" but by themselves.

Henry's doctrine of universals is based on this notion of reality. According to him, an individual can be conceived of either distinctly or indistinctly. When distinctly conceived, it is known through a particular concept; when indistinctly conceived, it is known through a universal concept. A universal is a confused concept by which the mind knows one individual without distinguishing it from others in the same genus or species. Inasmuch as an individual can be known through general concepts, Henry called it universal. For example, Socrates indistinctly conceived is man, animal, and body. Ockham criticized Henry's conceptualism because it ascribed some universality to things outside the mind.

Henry rejected the Scotist doctrine of the divine ideas as essences of creatures existing in God with cognitional being. He adopted a variation of the theory that the ideas are really the same as the divine essence itself known by God as imitable by creatures. God is known through concepts univocal to Him and creatures.

Henry stressed the omnipotence of God and the radical contingency of creatures. He claimed that no creature is naturally indestructible; the human soul is immortal not by nature but by divine grace. According to Henry, St. Thomas Aquinas betrayed Christianity by teaching the natural immortality of the soul.

See also Duns Scotus, John; Immortality; Scotism; Thomas Aquinas, St.; Universals, A Historical Survey; William of Ockham.

Bibliography

Balić, Carlum. "Henricus de Harcley et Ioannes Duns Scotus." In *Mélanges offerts à Étienne Gilson,* 93–121. Toronto: Pontifical Institute of Mediaeval Studies, 1959.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages,* 480–483, 779. New York: Random House, 1955.

Kraus, Johannes. "Die Universalienlehre des Oxforder Kanzlers Heinrich von Harclay." *Divus Thomas* 10 (1932): 36–58, 475–508; 11 (1933): 76–96, 288–314.

Maurer, Armand. "Henry of Harclay's Question on the Univocity of Being." *Mediaeval Studies* 16 (1954): 1–18.

Maurer, Armand. "Henry of Harclay's Questions on Immortality." *Mediaeval Studies* 19 (1957): 79–107.

Maurer, Armand. "Henry of Harclay's Questions on the Divine Ideas." *Mediaeval Studies* 23 (1961): 163–193.

Pelster, Franz. "Heinrich von Harclay, Kanzler von Oxford, und seine Quästionen." *Miscellanea Francesco Ehrle* 1 (1924): 307–356.

Armand A. Maurer (1967)

HERACLITUS OF EPHESUS

Heraclitus of Ephesus is an early Greek philosopher who lived around the end of the sixth century BCE. He was a native of Ephesus, an important Ionian city just north of Miletus on the western coast of Asia Minor, and his father's name was Bloson. If the story can be credited that he voluntarily surrendered to his brother a hereditary right to a ceremonial kingship, Heraclitus would be the oldest son of an old noble family. His birth and death dates are uncertain, but the evidence of our doubtful sources would place his floruit in the reign of Darius I of Persia. The authors Heraclitus names make it impossible for his single book to be dated much before the end of the sixth century, and since he is fond of naming his rivals, the lack of any reference or allusion in his surviving words to Parmenides of Elea argues for dating Heraclitus's book before the publication of Parmenides's poem.

Tradition tells us that Heraclitus deposited his book at the great temple of Artemis in Ephesus. His dedication of his book to the goddess may be tantamount to publishing it and to making his thoughts publicly available rather than hiding his thoughts away from the vulgar, as some have surmised. This publicity would be in keeping with Heraclitus's conviction that the truth is common and open to anyone and is not a private possession of the privileged few. From antiquity, Heraclitus is infamous for his obscurity, and he was dubbed early on "the dark." His obscurity has often been credited to his emulation of the Pythian Apollo, whose oracular deliverances Heraclitus analyzes insightfully: "The lord whose oracle is in Delphi neither tells nor conceals, but gives a sign" (frag. 93 Diels-Kranz). He highlights the indirection of the *lord* because of his conviction that the nature of things reveals itself indirectly, and he may mimic in his obscure writing what he takes to be the obscurity in reality itself. Instead of the hexameters of Apollo's priests and of the heroic poets, Heraclitus writes in prose, like most of the new intellectuals of the sixth century who were critical of the poetic tradition and undertook independent inquiry, or *historiê*, in a wide variety of areas.

The Milesian natural philosophers Anaximander and Anaximenes wrote on cosmology and cosmogony, while their fellow Milesian Hecataeus composed the first comprehensive geography of the Greeks, which in part he based upon what he learned from his own voyages. The fragments of Heraclitus's book, of which there are more than a hundred, provide the first substantial sample of Greek prose. Yet Heraclitus is also the most poetic of the early prose authors; he displays skillful use of traditional poetic devices, such as parallel and antithetical sentence constructions, chiasmus, alliteration, assonance, rhyme, and ring composition, as well as an adept use of wordplay that enhances his message. His book was probably not a continuous treatise of unbroken prose but a sequence of short passages, some of which are pithy enough in their moral import to look like a maxim of the Seven Wise Men: "It is hard to fight with anger; for whatever it should want it buys with the soul" (frag. 85 D-K). Despite his much-heralded obscurity, many of his sayings are as straightforward as this astute observation on moral psychology.

THE LOGOS AND THE UNITY OF OPPOSITES

Like his older contemporary Xenophanes of Colophon, Heraclitus is openly critical of the poets of the ancient past, but he also includes among his targets contemporary intellectuals. He is critical of "Hesiod and Pythagoras, and also Xenophanes and Hecataeus" for their "polymathy" that does not yield "understanding" (frag. 40 D-K). He finds "much learning" an impediment to understanding, and this puts him at odds with the new intellectuals who practice*historiê*, which depends upon polymathy. "Understanding" comes from heeding what Heraclitus calls "the Logos," by which "all things come to be," and whose message the common stock of humanity fail to appreciate, as well as those reputed to be wise. They live in a private world of their own making, comparable to dreams, but those who harken to the Logos live in the one public world of the wakeful (frag. 89 D-K). Along with Xenophanes, Heraclitus is among the first of the new

breed of intellectuals to make an issue of the human epistemic condition.

The nature of this Logos is contested. Some scholars understand it as the nature or essence of reality, as it shows itself in discourse, others as a universal principle or law that regulates the basic workings of reality, and a few render it as Heraclitus's true *account* of reality in the form of his own book, or *logos*. With his predilection for wordplay, Heraclitus could well allow Logos to stand for both his book and the subject of his book. He lays down a telling parallel when he urges "those speaking with understanding" to hold to what is "common to all things," presumably the Logos, just as a city holds to its "laws." The commonality of the Logos would be comparable to the way in which the laws of a city apply across the whole of its citizenry, as the rules that regulate their behavior and shape them into a single community, and not to the way the air of Anaximenes's cosmology is the common constitution of all things. What is comparable to "human laws" is also what they are "nourished by," "by one, the divine," which in his ambiguity Heraclitus may intend to be "the one divine law" (frag. 114 D-K). The importance of what sustains "human laws" devolves upon them, so that "The people must fight for the law as for a city wall" (frag. 44 D-K).

The one surviving explicit message of the Logos declares that "all things are one" (frag. 50 D-K). This unity is not the oneness of the monism Aristotle credits the earliest natural philosophers with advocating, but the unity of opposites. This "connection" lies "unseen" (frag. 54 D-K), beyond the patterns of ordinary ways of thinking, as well as the teachings of the old authorities and of the new intellectuals. A "strife" between opposing powers lies hidden within the nature of each thing, and without this strife, the cosmos and everything in it would perish. While contesting with one another, the opposing powers within the essence of each thing cooperate with one another and yield a unified object: "They do not comprehend how each thing quarreling with itself agrees; it is a connection turning back on itself, like that of the bow and the lyre" (frag. 51 D-K). There would be no bow or lyre unless there were a striving between the wood and string through their powers of pulling in opposing directions.

At the cosmic level, the unity of opposites displays itself in the strife between the great cosmic powers of the hot and cold, the dry and moist, since even as they strive with one another for dominion, in the form of fire, water, and earth, they are tightly linked. The destruction of one cosmic mass is the generation of another, "death for water is the birth of earth, from earth water is born" (frag. 36 D-K), where birth and death unite in a single event. The strife between opposing powers is beneficent and just, and "justice is strife" (frag. 80 D-K), contrary to the teaching of Anaximander, who describes the dominion of one opposite over another as "injustice." When people count some things as just and others as unjust, they divide justice from injustice, but from the objective position of god, all things are "fair and good and just" (frag. 102 D-K). The division between opposites is real enough, but so too is the unity, "it scatters and again brings together" (frag. 91b D-K). This divisive thought of the popular imagination leads to a false impression, which Homer fosters, that the positive of the pair is preferable, morally superior, and should dominate. Aristotle reports that Heraclitus criticizes Achilles's lament, "Would that strife might perish from among gods and men" (*Iliad* 18.107), since without strife there would be no peace, no coherent cosmos.

Heraclitus's originality lies most prominently in his efforts at establishing the integrity of each thing through the unity of opposing powers within each thing, and he revolutionizes thought about values through his insistence upon this unity. No opposite can be valued to the exclusion of its counterpart g powers, because of the various ways in which they are tied to one another for their presence in the world and their efficacy. Heraclitus goes beyond his predecessors in displaying the positive nature of those powers that ordinary ways of thinking deem to be purely negative.

EPISTEMOLOGY AND RATIONALISM

The truth is "hidden," yet "obvious." The blind poet Homer, who is the "wisest of all the Greeks," fails to appreciate the "obvious" (frag. 56 D-K). The truth is obscure, yet it remains open to anyone's inspection through simple means of comprehension. The unity of opposites is no mysterious dogma handed down from on high, and its confirmation may be achieved through observation and argumentation, linguistic analysis, and self-reflection. Heraclitus has confidence in the truth-yielding capacity of observation, "Those things that come from sight, hearing, learning from experience, these I esteem" (frag. 55 D-K), although observation must be evaluated carefully: "Eyes and ears are poor witnesses for men if they have barbarian souls" (frag. 107 D-K). Simple arguments premised on trivial empirical truths provide evidence for the unity: "Sea water is the purest and foulest of water, for fish it is drinkable and life-sustaining, for men it is undrinkable and deadly" (frag. 61 D-K).

In his exploitation of everyday language as a pathway to truth, Heraclitus puns on an uncommon word for "bow," which differs only in accent from the common word for "life," so that he may make manifest the connection between life and death: "The name of the bow is life, but its work is death" (frag. 48 D-K). Death is life, since, for example, the destruction of earth is the birth of water. When Heraclitus notes that "they would not know the name of Justice, if these things did not exist" (frag. 23 D-K), presumably "unjust things," he draws together opposites in the belief that a "name" like "justice" has no meaning in isolation, but only with its opposite, "injustice." Heraclitus will also appeal to a word's etymology for his evidence. The assistants of Justice, he reports, are the Furies (frag. 94 D-K), whose name meshes well with his identification of justice and strife (frag. 80 D-K), since it derives from "strife."

Heraclitus, unlike many of the new intellectuals, has no use for travel and the information it yields as a means for gaining "understanding." The only "journey" he ever mentions is into one's soul, in search of oneself (frag. 101 D-K), and "You would not find out the limits of the soul by going, even traveling over every road, so deep is its *logos* (frag. 45 D-K). This inward journey reveals the value of a measured existence for human well-being. The "measured man" learns from self-examination the proper limits of the great destructive forces of emotion and desire. Despite Heraclitus's revolutionary reassessment of values, he shows himself still bound to tradition when he pairs self-knowledge with measure (frag. 116 D-K), which are values highly esteemed by the Pythian Apollo, whose Alcmaeonid temple posted prominently the famous maxims of the Seven Wise Men: "Know yourself" and "Nothing too much." Like the "measured man," the world-order "lives" a measured existence; the cosmos is "fire ever living, kindled in measures and in measures going out" (frag. 30 D-K). When one cosmic mass changes into another, a *logos*, or proportion, holds between them, so that, for instance, the sea "measures up to the same *logos* it was before becoming earth" (frag. 31b D-K). The cosmos is a self-regulating system that keeps within spatial and temporal limits the great destructive forces of nature: "The Sun will not step over his measures" (frag. 94 D-K).

The Logos belongs to the soul as much as anything else; thus, self-knowledge may provide a path to cosmic knowledge and to "understanding." One need not go far afield or draw upon extraordinary powers to discover the truth. Heraclitus is no pessimist, in contrast with Homer and the poets who believe that humans left to themselves without the aid of the Muses have no knowledge of recondite topics and are the victims of "rumor" (*Il.* 2.485—486). Humankind has within its reach the truth of reality, "It belongs to all men to know themselves and to think in a measured way" (frag. 116 D-K), although Heraclitus thinks that few will ever exercise successfully these shared capacities. He tempers his optimism further when he maintains that a man hears from a divinity that he is "infantile" just as a child hears the same from a man (frag. 79 D-K), and some think that they detect poetic pessimism in his observation that "human character," in contrast with "divine," "has no judgment" (frag. 78 D-K).

Yet Heraclitus is also no mystic, if by "mysticism" is meant a private insight into the truth, vouchsafed to the few, which goes beyond the ordinary capacities of humankind. Instead of intuition, Heraclitus has recourse to argument and public verification. His rationalism holds, even if Aristotle is correct in charging him with contravening the principle of non-contradiction, when Heraclitus insists that the same thing displays opposing properties (*Metaphysics*, 1062A30-35). Aristotle's charge is plausible even if we may easily dispel the appearance of contradiction, when Heraclitus maintains, for example, that sea water is both pure and impure, by our pointing out that these properties are not contradictory since they are qualified in different ways by applying to different creatures: fish and humans. Heraclitus may have not been able to recognize that the ambiguity gives him only the appearance of contradiction. But he comes by his view of unity honestly, without mystery, through his appeal to argument and observation.

CHANGE AND FIRE

Heraclitus argues for the truth of the unity of opposites, by arguing that the contrary pairs, the living and the dead, the waking and the sleeping, and the young and the old, are the "same" because the contraries of each pair mutually replace one another (frag. 88 D-K). Living things die, but from the remains of the dead living things emerge. Day and night are "one," thinks Heraclitus, probably because of their mutual succession (frag. 57 D-K). Heraclitus ties together inferentially two of his important doctrines when he derives the unity of opposites from the fact of change, and he may see change at the foundation of his speculations. Despite the centrality of change, its nature has been subject to exaggeration. Plato, who may be more under the influence of Cratylus than of Heraclitus, finds change to be incomprehensible, and he credits Heraclitus with a doctrine of universal flux in which reality is likened to the flow of a river, where "you could not step twice into the same river" (*Cratylus* 402A). These

words indicate an extreme sort of change in which nothing retains its identity. The "river fragments" suggest something less extreme, that things constantly change but retain their identity: "As they step into the same rivers, other and still other waters flow upon them" (frag. 12 D-K). The rivers remain rivers; only the water that constitutes them is constantly changing.

Cosmic change is not chaotic, but occurs in an orderly way, as Heraclitus suggests when he speaks grandiloquently of the eternity of the world-order: "This cosmos, the same for all, no one of gods and men made, but always was and is and shall be fire ever living, kindling in measures and going out in measures" (frag. 30 D-K). The fire the cosmos is identified with changes, but in a measured way, changing in its extinction into the other great cosmic masses, and in an orderly pattern changing back again in its ignition. The flow of fire matches the flow of a river, but fire is more than an image when Heraclitus identifies it with the cosmos, and when he makes fire worth all else: "All things are an exchange for fire and fire for all things, just as gold for goods and goods for gold" (frag. 90 D-K). Heraclitus privileges fire, but not after the fashion of the monists, as Aristotle and Theophrastus would have us believe, as the stuff that constitutes all else. Theophrastus explains the "exchange" between fire and all things as fire's yielding everything else through its rarefaction and condensation, although he must admit that Heraclitus "sets out nothing clearly" (Diogenes Laertius 9.8).

It is not surprising that Theophrastus finds Heraclitus unclear, since the mercantile image of exchange, of "gold for goods and goods for gold," indicates that what is exchanged for fire is no more fire than the goods exchanged for gold are gold. In keeping with the mercantile image, the primacy of fire lies in its providing the standard that fixes the value of all else, as equivalent to so much fire, and Heraclitus may value fire above all else because it is psychic stuff: "For souls death is the birth of water, for water death is the birth of earth, from earth water is born, from water soul" (frag. 36 D-K). The sequential change back and forth between soul, water, earth suggests an exhaustive cosmic exchange, and thus the absence of the important cosmic mass of fire calls for its identification with soul. It is a "dry soul" that is "wisest and best" (frag. 118 D-K). Heraclitus has room for only three great cosmic masses, fire, water, earth, in his physics and no place for the air of Anaximenes.

THEOLOGY

The soul is the basis of life, but also of intelligence (frag. 107 D-K). Heraclitus links fiery stuff and intelligence when he says that "Thunderbolt steers all things" (frag. 64 D-K). "Thunderbolt," which stands for the guiding principle behind the cosmos, is the instrument of Zeus, the greatest god in the Greek pantheon, and Heraclitus intends for his ruling principle to be identified with the divine, but in a qualified way, when he says: "The one, the wise alone, is not willing and is willing to be spoken of by the name of Zeus" (frag. 32 D-K). Heraclitus appropriates a divine name from popular religion, but he warns against its literal application. His ruling principle, like Zeus, is the most powerful of deities, but, unlike Zeus, it should not be conceived in an anthropomorphic manner. The traditional language of divine anthropomorphism shows up in his praise of strife, "War is father of all and king of all" (frag. 53 D-K), which recalls the Homeric description of Zeus as "the father of men and gods," and Zeus is the "king" of the gods.

Heraclitus borrows freely from the conventional language and images of popular religion, but he applies them in unconventional ways, and his practice suggests that he is trying to formulate a new way of talking about the divine within the idiom of the old. Among the Greeks Xenophanes initiated the criticism of anthropomorphism, but, unlike Heraclitus, he purifies his language of the traditional anthropomorphic vocabulary. In one remarkable passage, Heraclitus draws together the divine, the opposites, and perhaps even fire: "The god, day-night, winter-summer, war-peace, satiety-hunger, and it undergoes change, as when mingled with perfumes, it is named according to the pleasure of each one" (frag. 67 D-K). The divine is actually identified with the opposites, as if their union was the divinity itself, and Heraclitus may treat the god like fire when referring to its changing in accord with the perfumes mixed into it. The significance of fire corresponds to the importance of Logos; war, like the Logos, is common; strife, like the Logos, is what all things come to be in accord with (frag. 80 D-K). The Logos, fire, strife, and divinity would seem to come together in Heraclitus's thought, although there is no evident formula for the expression of their convergence.

When he subjects particular cult practices to criticism, Heraclitus proves to be harsher than Xenophanes. The age-old practice of purging oneself of blood guilt through blood sacrifice Heraclitus ridicules as comparable to washing off mud with mud, praying to statues is like "chatting with houses" (frag. 5 D-K), and the "mysteries" men now believe in do no more than "initiate into

impiety" (frag. 14 D-K). The "procession" for Dionysus and the "chant for the phallus" would be shameful if they were not done for the sake of the god, and the participants in these practices do not even recognize that "Hades and Dionysus are the same" (frag. 15 D-K). Once again Heraclitus appropriates conventional divine names, but uses them in a shocking way by identifying traditional deities of widely contrasting natures, Hades and Dionysus, perhaps once more as a way of signifying the identity of life and death. The alcohol beloved of Dionysus Heraclitus condemns as turning a man into a boy by making his "soul moist" (frag. 117 D-K), and, despite the "joy" men take in moisture, it is death (frag. 77 D-K).

Heraclitus does not recommend any new practices to take the place of those he censures, in contrast with Pythagoras who recommends many new rituals to supplement those of tradition. When Heraclitus maintains that "Character for a man is fate" (frag. 119 D-K), he looks as if he were removing humankind from the tutelage of the gods. *Daemon*, the word for "fate," is [a] also the word for a guardian divinity, and thus in identifying a man's own character with his guardian, Heraclitus would be stressing that humans should take responsibility for their actions instead of laying blame upon the divine for their fortunes, both good and bad. Heraclitus is often thought to sanction immortality for at least some souls, perhaps of warriors: "Greater deaths are allotted greater portions" (frag. 25 D-K); "Those slain by Ares, gods and men honor" (frag. 24 D-K). Personal survival is not possible in a cosmos of universal destruction, and Heraclitus may mean no more than enlightenment when he speaks of those who "arise and become wakeful watchers of the living and the dead" (frag. 63 D-K). The wakeful may be those awakened from folly, since Heraclitus associates subjective misapprehension with sleep and objective comprehension with wakefulness (frag. 89 D-K).

COSMOLOGY

There is little reason to think that Heraclitus makes any advances in physics or astronomy. Unlike the Milesians, he does not seem to take much interest in the details of natural philosophy, and the fragments speak little to the issue. His words testify to his belief in an eternal cosmos (frag. 30 D-K), even though Aristotle and Theophrastus report otherwise. His rejection of cosmogony would mark him out significantly from the early natural philosophers, although Xenophanes may, too, have championed an eternal cosmos. Theophrastus and the doxographical tradition he founded report some astronomical and meteorological speculations. Bright and dark exhalations arise from earth and sea, and "bowls" in the heaven trap the bright exhalations and form the heavenly bodies. The rotations of these bowls account for the phases of the moon and eclipses. The preponderance of bright and dark exhalations contributes to the explanations of day, night, months, seasons, years, rains, and winds. In what may be Heraclitus's own words, he traces daylight back to the sun, "If there were no sun, it would be night" (frag. 99 D-K), and he believes, along with Xenophanes, that the sun is "new each day" (frag. 6 D-K). Theophrastus concludes his report by saying that Heraclitus offers no explanation of "what the earth is like, or even about the bowls." The Hellenistic grammarian Diodotus finds Heraclitus's book to be about "man's life in society," and its statements on nature to serve only as "illustrations." (D. L. 9.15) The doxographical tradition may be misguided in assimilating Heraclitus's work to the discipline of natural philosophy. Theophrastus indicates further difficulties he had with Heraclitus's book when he maintains that some things Heraclitus wrote were "half-finished," others "inconsistent," which Theophrastus puts down to Heraclitus's "melancholy." (D. L. 9.6)

INFLUENCE

Of the Presocratic thinkers, Parmenides had the most influence, but Heraclitus may have had the most influence upon him, in drawing his attention to the problematic nature of change. Heraclitus may have had a certain vogue in the fifth century. Some minor Hippocratic authors reflect something of his paradoxicality, and Plato jokes about the fidgeting Ephesians, who are not stable enough to carry on an argument. Cratylus of Athens pushed change to such an extreme that he finds it necessary to rebuke Heraclitus for thinking that one could not step twice into the same river, when one could not step even once. Aristotle reports that in his youth, Plato was under the influence of Cratylus and found the objects of perception unknowable because of their instability. Heraclitus's most profound philosophical influence was upon the Stoics, who credited him with anticipating them. Fire has a primacy for them, and they, too, adopt a Logos as a ruling principle that is eternal, divine, and common to all things. Unlike Parmenides, Heraclitus exercises his charm well beyond antiquity and beyond philosophy. Hegel finds a positive parallel between his logic and the doctrines of Heraclitus, and T. S. Eliot spins out much of *The Four Quartets* in imagery borrowed from Heraclitus. Heraclitus's poetic prose attracts many to this day.

See also Anaximander; Anaximenes; Cosmology: Cratylus; Epistemology; God, Concepts of; Homer; Meta-

physics; Parmenides of Elea; Philosophy of Religion; Philosophy of Religion, History of; Philosophy of Religion, Problems of; Pythagoras and Pythagoreanism; Xenophanes of Colophon.

Bibliography

WORKS

Text

Diels, Hermann, with Walther Kranz. *Die Fragmente der Vorsokratiker, griechisch und deutsch*. Vol. 1. 8th ed. Berlin: Weidmann, 1956–1959.

Texts and Commentaries

Kahn, Charles H. *The Art and Thought of Heraclitus: An Edition of the Fragments with Translation and Commentary*. Cambridge, U.K.: Cambridge University Press, 1979.

Marcovich, Miroslav. *Heraclitus: Greek Text with a Short Commentary*. Mérida: Los Andes University Press, 1967.

Robinson, Thomas M. *Fragments/Heraclitus: A Text and Translation with a Commentary*. Toronto: University of Toronto Press, 1987.

Books and Articles

Barnes, Jonathan. *The Presocratic Philosophers, Vol. 1: Thales to Zeno*. London: Routledge and Kegan Paul, 1979.

Guthrie, W. K. C. *A History of Greek Philosophy, Vol. 1: Earlier Presocratics and Pythagoreans*. Cambridge, U.K.: Cambridge University Press, 1962–1981.

Hussey, Edward. "Epistemology and Meaning in Heraclitus." In *Language and Logos: Studies in Ancient Greek Philosophy Presented to G. E. L. Owen*, edited by Malcolm S. Schofield and Martha C. Nussbaum. Cambridge, U.K.: Cambridge University Press, 1982.

Kirk, G. S., ed. *The Cosmic Fragments*. Cambridge, U.K.: Cambridge University Press, 1954.

Kirk, G. S., J. E. Raven, and M. Schofield. *The Presocratic Philosophers: A Critical History with a Selection of Texts*. 2nd ed., rev. Cambridge, U.K.: Cambridge University Press, 1983.

Nussbaum, Martha Craven. "Psuchê in Heraclitus." *Phronesis* 17 (1972): 1—16, 153–170.

Vlastos, Gregory. "On Heraclitus." *American Journal of Philology* 76 (1955): 337—368.

Wiggins, David. "Heraclitus' Conceptions of Flux, Fire, and Material Persistence." In *Language and Logos: Studies in Ancient Philosophy Presented to G. E. L. Owen*, edited by Malcolm S. Schofield and Martha C. Nussbaum. Cambridge, U.K.: Cambridge University Press, 1982.

Herbert Granger (2005)

HERBART, JOHANN FRIEDRICH

(1776–1841)

Johann Friedrich Herbart, the German philosopher, psychologist, and educational theorist, was born in Oldenburg; he entered the University of Jena in 1794. Although he studied under Johann Gottlieb Fichte, Herbart was unable to accept Fichte's view of the ego and its psychology, and in reaction he laid the basis for his own metaphysical and psychological views. In 1797 Herbart took a post in Switzerland as tutor. He held the position for three years and, during this period, worked out to a large extent the views that he was to refine and elaborate for the rest of his life.

After he took his doctorate at Göttingen in 1802, Herbart remained there for the next seven years. *Allgemeine Pädagogik* (General Theory of Education) and *Hauptpunkte der Metaphysik* (Main Points of Metaphysics), both of which appeared in Göttingen in 1806, and his *Allgemeine praktische Philosophie* (General Practical Philosophy; Göttingen, 1808) were major fruits of this period. In 1809 Herbart moved to Königsberg to occupy Immanuel Kant's former chair, and there he published his *Lehrbuch zur Psychologie* (Compendium of psychology; 1816), *Psychologie als Wissenschaft* (Psychology as a science; 1824–1825), and *Allgemeine Metaphysik* (General metaphysics; 1828–1829). When the political situation rendered Königsberg continually less attractive and Herbart failed to secure G. W. F. Hegel's chair at the University of Berlin, he returned to Göttingen in 1833 and remained there until his death.

THE PARTS OF PHILOSOPHY

Philosophy, according to Herbart, cannot be characterized by its subject matter but only by its method, which is the reworking (*Bearbeitung*) of concepts; and the possible kinds of such elaboration determine the major divisions of philosophy.

The first kind of reworking renders concepts clear and distinct. Distinct concepts can be formed into judgments and these judgments can be organized into inferences. This process of distinguishing and ordering concepts is logic.

But experience gives rise to many concepts which, the more distinct they become, the more contradictory they appear, a sure sign that we are missing both being and truth. Our ideas must, therefore, undergo an "enlargement" (*Ergänzung*), which will remove the contradiction. This second kind of reworking of concepts gives us the second great division of philosophy, metaphysics.

Concepts in the third class are like the metaphysical concepts in that they cannot remain merely at the level of clarity and distinctness, as do the logical concepts. But

while the metaphysical concepts involve only enlargement, this third class involves in addition an intuitive judgment of approval or disapproval. Thus we get the last great division, aesthetics. Aesthetics includes a series of doctrines of art or practical sciences. One of these, ethics, issues necessary (and not merely conditional) prescriptions because we continuously and necessarily concern ourselves with its object, ourselves.

METAPHYSICS

Metaphysics consists of four parts: (1) method, the general principles of the proper method and order of procedure; (2) ontology, the study of the real; (3) synechiology, the study of those forms of experience that have continuity (such as space, time, and motion), and (4) eidolology, the examination of the possibility of knowledge.

METHOD. The first task of metaphysics is to define "the given" in experience. Common sense says that it is "things with multiple and changing characteristics." But this concept violates the law of identity. For any single thing dissolves into a multiplicity of qualities when we describe it; it is at once both a unity and not a unity. Substance, cause, ego, time, and space are also contradictory. Yet they must be "given" in some sense, since we cannot change them at will. In moving to metaphysical concepts freed from contradiction, we use as our chief tool the " method of relations." A contradictory concept of experience, *A*, unites the contradictory terms, *M* and *N*. *M* is thus both identical and nonidentical with *N*. If one further divides *M* into *M'* and *M"*, one element will still be identical with *N*, the other not; this contradiction disappears only when we admit that, although each of the parts of *M* is not identical with *N*, they mutually modify each other so that together they become so. Thus, in the syllogism, the conclusion must be contained in the premises; and thus the premises must change into or cause the conclusion. But *M'* and *M"*, the premises or cause, are not individually the same as *N*, the conclusion or effect. Hence *M'* and *M"*, the premises, must mutually modify each other so as to become identical with the conclusion.

ONTOLOGY. Ontology deals with being. Since being is not directly given in experience, it is easy to say that there is no being. But it is hard to live with this judgment, for things continue to appear. What, then, is it that appears? Appearances or phenomena cannot be taken as the only reality, for the concept of absolute position (being in absolutely no relation whatever to anything) cannot be applied to them; they are always related to something else. Nor can phenomena be reduced to our sensations of them and then located in an ego, for our sensations are not just sensations of sensations, and the concept of the ego is itself contradictory.

We are thus led to posit "being" as a plurality of beings or reals (*Realen*), with the essence of each real a single quality, absolutely simple, without parts, degree, or negation, always immutably identical with itself. But how can this concept of being be reconciled with our experience, which is both the basis of metaphysics and its test? The absolute position of reals seems to contradict the multiple relations in which things appear to us. But being can be conceived by mind; and in mind a being is only an image (*Bild*). Mind also can simultaneously represent several beings, which, as images, can stand in many different relations to the first one. These relations are "contingent viewpoints" (*zufällige Ansichten*), which exist in thought, not in things-in-themselves. Just as in analytic geometry the same point can be part of an infinite number of curves, so a single real may enter many contingent viewpoints.

Experience presents us with complex aggregates that we call things. Yet we cannot say that the aggregate exists, for colors, sounds, and such exist only in the perceiving subject. Nor can we say that something having those qualities—that is, substance—exists, for substance cannot be being if being is simple, since substance appears as endowed with manifold and varying qualities. How, then, can attributes and modes inhere in substance? By the method of relations. If *A* is a substance and *a* an attribute, analysis of *A* into multiple elements (*A'*, *A"*, and so on) will not resolve the contradiction unless we say that *a* is not identical with any one of the *A*s but with the totality of them, the number of which remains undetermined but which must be at least two if mutual modification between them is to occur. Substance, then, is explainable as multiple beings in conjunction (*zusammen*) with each other, many reals grouped about one real. This conjunction explains that unity that we attribute to substance, although the essential unchanging quality of that central real is unknown to us; and the other reals in conjunction with it account for the varied and varying attributes we experience, although those attributes are not the essential qualities of those various reals.

The conjunction and the separation of the reals explain those sensible appearances which led to their postulation. Such mutual interaction would seem to lead to mutual destruction, but, just as in an equilibrium of forces both forces remain constantly what each is, though balanced, so in the concatenation of reals, the mutual perturbations (*Störungen*) that would lead to mutual

destruction are counterbalanced by individual acts of self-preservation through which each real strives to remain what it is.

These acts of perturbation and self-preservation constitute real phenomena (*das wirkliche Geschehen*), as distinct from sensible phenomena or appearances (*Erscheinung*), of which the real ones are the basis and explanation. To what or to whom do they appear? This is the problem of the ego, self, or soul-substance.

The ego, or self, poses the same problems as do the other substances, and the solution is the same. The unity and diversity of the ego are explained by the coming and going of reals. The soul, which is not the ego of consciousness, is a real, but one endowed with mind, which is the seismograph that records, in the form of presentations or ideas (*Vorstellungen*), the acts of self-preservation of the soul vis-à-vis the other reals. These presentations are the sensible phenomena given in experience. With metaphysics thus having shown the origin of our ideas, psychology will show their development and combination.

SYNECHIOLOGY. Synechiology concerns that which is continuous (*das Stetige*)—notably space, time, and motion. Continuity, as union in separation and separation in unity, is a contradictory concept (though undoubtedly given in experience), which must be explained by metaphysics. As far as being is concerned, space and time are "obviously nothing." They, like the continuity we attribute to them, are merely natural and necessary products of the psychic mechanism. What essentially characterizes space and time is the mutual exteriority (*Auseinander*) of the parts. But between points of space or time it is always possible to *conceive* additional parts, and this further functioning of the psychic mechanism makes space and time seem to flow uninterruptedly.

But the comings and goings of the reals imply some sort of space, time, and motion, even though these are distinct from their sensible counterparts, such as "intelligible space." Although two reals, *A* and *B*, are actually apart, we can conceive the possibility of *B*'s being with *A* and *A*'s being with *B*. Thus space is the simple possibility *in mind* for one real to be together with another from which it is separated in reality, an "image" without reality. Space thus being completely accidental for reals, we can, by putting *A* in the place of *B* and *B* in the place of *A* and further continuing to add more reals and more dimensions, generate lines infinite in all directions, even though each line is "fixed" (*starre*) with a determinate number of points rather than continuous (*stetige*) with mutual interpenetration (and hence indefiniteness in number) of the parts. The psychic mechanism then conceives of these "fixed" lines as continuous by interpolation. Intelligible space, as thus generated, corresponds to the sensible space of phenomena and shares its contradictions, but these need not trouble us, since they have nothing to do with being. Intelligible space is a purely conceptual artifact (*Gedankending*), not imposed by mind on things but generated by mind as a necessary aid to thought. Once intelligible space has been generated, the explanation of matter becomes possible. It is a question of asking what situation the images of reals should occupy with respect to intelligible space in order to account for matter. The answer is "incomplete interpenetration."

EIDOLOLOGY. Eidolology examines the possibility of knowledge and its limits. In all knowledge, matter and form can be distinguished. The matter is simple presentations or sensations. They do not enable us to know what is, but they do oblige us to believe that something is. But the given has form as well as matter. Sensations are not given us in isolation but formed into groups, which cannot be separated at will and which constitute things. Doubtless, in the form in which they appear to us, things exist only in and for mind. But the constancy and the modifications of these groups of sensations have their basis in the conjunction and separation of the reals. Thus mind, though it cannot know the qualities of the reals through sensation, does know their relations; and even were our sensations wholly different from what they are, their forms would be the same, arising as they do from the objective separations and connections of the reals. Knowledge through concepts is likewise valid knowledge, although it too is knowledge only of relations. "We live amid relations and need nothing more."

PSYCHOLOGY

Everything in mind arises in some fashion out of presentations. There are no faculties, no innate ideas, no concepts a priori. The soul is a real in which countless acts of self-preservation are provoked through its contacts with other reals, and these efforts in turn produce in mind the presentations, some of which oppose, some of which reinforce, each other. Although reals and presentations are not forces, they can best be understood by analogy with forces, and hence the synthetic part of psychology consists of the statics and mechanics of mind. Complex mathematical formulas, corresponding to those of the statics and mechanics of physics, describe the interplay of presentations.

Presentations of different sorts do not oppose each other (for instance, colors do not oppose sounds); but presentations of the same sort do (for instance, red opposes blue). In the latter case, what remains after arrest is an equilibrium, a weakening or obscuring of the original presentations that is reached progressively by a process of sinking (*Senkung*). A presentation, if it has not undergone arrest, is present in consciousness. Sinking under arrest, it may be forced below the threshold of consciousness. Yet a presentation below the threshold of consciousness is subject to recall or rising (*Hebung*) by the appearance of a new presentation similar to it, and the speed of this rising depends on the degree of similarity between the two presentations. This new presentation also produces a vaulting (*Wölbung*) or "arching" of all other arrested presentations similar to it. The coming of a new presentation, *B*, produces the rising (*Hebung*) of the similar older presentation *A*. But as *A* is pulled up, other older presentations similar to *A* but less similar to *B* are also pulled up in a *Wölbung*, or arching. The analogy to a beater being pulled out of stiff whipped cream is exact. The surface of the cream closest to the beater is pulled up most (*Hebung*), but the whole center surface arches somewhat (*Wölbung*).

The feelings, the desires, and the will have their origin in presentations. Some feelings arise out of the fusion of opposed presentations, and the pleasantness or unpleasantness depends on the amount of opposition. Other feelings originate in the strain that the rising, produced by a new presentation, puts on ties that an old presentation already has with one or more others. Thus the sight of an object belonging to a dead friend evokes the memory of him, but the thought of his death tends to repress the memory and thus to produce a painful feeling. Pleasant feelings arise in the contrary situation, when the other associated presentations all facilitate the recall of the original one. The desires are closely connected with the feelings. In a situation giving a painful feeling, where *A* is lifted toward consciousness by the appearance of *C* and is simultaneously depressed by its earlier relation to *B*, the feeling of effort by which the resistance is overcome will be a desire and *A* will appear as the object of desire. The will, in turn, is only a particular form of desire, the realization of which is seen as possible.

Concepts, also, have their origin in the fact that each new presentation produces the vaulting of the images of similar previous presentations already in mind, in which process the similarities are reinforced and the differences between them are repressed, as in a composite portrait.

ETHICS

Ethical judgments are aesthetic judgments involving pleasure and displeasure. Since the completely simple cannot be pleasing or displeasing, these judgments must be directed to something complex, to relations. Since Herbart, like Kant, sought the basis of ethics in the good will, the five possible relations of the will suggest five corresponding fundamental ethical ideas.

The idea of "inner freedom" is the correspondence between a single act of a single will and the judgment passed on it. Harmony between the "objective" will (the inclinations) and the "subjective" will (the intuitive ethical judgment) is absolutely pleasing; its contrary, displeasing.

"Perfection" relates the varied acts of a single will. To this multiplicity, three quantitative concepts may be applied: the strength of any single effort (intensity), the multiplicity of the objects encompassed by the will (extension), and the concentration of this manifold into a total power (a new intensity developing out of extension). There is no absolute standard, but the stronger and more concentrated will is more pleasing than the weaker.

The idea of "benevolence" arises when one will comes to terms with the will of another. Yet this relation is internal to the first will in that it takes the will of the other person as an object.

The idea of "law" concerns the relations between the wills of two persons who desire some one thing. The ensuing strife is not, however, merely the contrary of the idea of benevolence, since both wills are turned directly toward the object and only indirectly toward each other.

The idea of "equity" arises from the intentional doing of an act of ill or good, a displeasing imbalance between two wills that can be rectified only by some appropriate requital through reward and punishment.

These five basic ethical ideas cannot be isolated in estimating character or in organizing social or political life; each must be tempered by all the others. Together, they exhaust the possible relations of the will, since the addition of more wills repeats in more complicated fashion those already covered by law and equity. But extension is possible if the many wills of a group can be analogized to that of a single rational being. Then five analogous social ideas appear in the reverse order. Members of this society will seek to avoid strife through a "system of law." But transgressions lead to a "system of requital." The benevolent spectator would wish for the greatest possible sum of well-being attained through the rational distribution of the available goods according to a

"system of administration." Then increased well-being would produce an intensity and range of strivings reconciled under a "system of culture." With the obedience of each to the moral insight of all, the many would become one in an "ideal society."

EDUCATION

Education takes its aim from ethics; psychology then shows it the means and hindrances to this end. The aim is moral strength of character, a will with inner freedom whose volitions are always in accord with the moral law. The three major divisions of education are instruction (*Unterricht*), discipline (*Regierung*), and training (*Zucht*). Since psychology shows that the entire mental life (including the desires and the will) is built out of presentations, instruction (with its four steps of clarity, association, system, and method) is directed toward enlarging the child's circle of thought and developing in him a many-sided interest by efficiently introducing the proper presentations into his apperceptive mass. Discipline keeps the child obedient and attentive so that instruction and training can do their work before the child has developed a proper will of his own. Training works constantly with instruction and discipline to form the will directly through such means as environment, examples, and ideals. Under discipline, the child acts rightly because he must; under proper instruction and training, he acts rightly because he wills to do so.

See also Apperception.

Bibliography

The writings of Herbart were published in *Collected Works*, edited by G. Hartenstein, 2nd ed., 13 vols. (Hamburg: L. Voss, 1883–1893), and *Collected Works*, edited by K. Kehrbach and O. Fluegel, 19 vols. (Leipzig, 1887–1912; reprinted 1963).

An introduction to Herbart's philosophical psychology is G. F. Stout's essay (1888), reprinted in his *Studies in Philosophy and Psychology* (New York: Macmillan, 1930). Works on Herbart include G. Weiss, *Herbart und seine Schule* (Munich: Reinhardt, 1928); T. Fritzsch, *J. F. Herbarts Leben und Lehre* (Leipzig, 1921); and M. Mauxion, *La métaphysique de Herbart* (Paris, 1894).

OTHER RECOMMENDED WORKS

Crittenden, Brian S. "J. F. Herbart on the Moral Purpose of Education." *Educational Theory* 15 (1965): 112–120.

Dunkel, Harold B. "System Trouble in Herbart and the Herbartians." *Philosophy of Education* 23 (1967): 19–31.

Harold B. Dunkel (1967)
Bibliography updated by Tamra Frei (2005)

HERBERT OF CHERBURY

(c. 1582–1648)

Edward Herbert, the first Baron Herbert of Cherbury, courtier, soldier, diplomat, poet, historian, philosopher, and theologian, was the brother of George Herbert (1593–1633), the pastor and poet. He matriculated at University College, Oxford, in 1596. He moved to London in 1600, where he continued his studies and attracted the attention of the aging Queen Elizabeth. On his accession to the throne of England James I created him a knight. As a young man Herbert traveled on the Continent, on occasion being involved in warfare where he showed what some judged bravery and others foolhardiness. Visiting Rome, he called at the English College and showed his undogmatic spirit when he told a person whom he met there that while he was not a Roman Catholic, he judged that "the points agreed on both sides are greater bonds of amity betwixt us, than the points disagreed on" (*The Life of Edward Lord Herbert of Cherbury, Written by Himself*, p.105).

In 1619 he was appointed ambassador to the French court. In this post he showed himself to be a skillful diplomat who was prepared to use his own initiative and to give sensible, even if unpalatable, advice to his government. In 1624 he was recalled. The Crown failed to reimburse his debts as ambassador but sought to satisfy him with peerages, first the Irish barony of Castle Island, and later the English barony of Cherbury. In vain attempts to recover royal favor Herbert wrote two histories. The first, *Expeditio in Ream Insulam* (published posthumously in 1656), tries to defend the Duke of Buckingham's conduct in an English invasion of the Isle de Rhé in 1627 that was intended to support the Huguenots. Unsurprisingly in view of what happened on the expedition, it is not a convincing defense. The other was *The Life and Raigne of King Henry the Eighth* (1649), in preparing which he used official archives. It was long regarded as an authoritative study.

Although a member of Charles I's Council of War, Herbert sought as far as possible to avoid playing an active part in the English Civil War. He surrendered Montgomery Castle, where he was living, to Parliamentary forces when they augmented their challenge with the threat to sell the library that he possessed in London. Soon after, he moved to London and died there in 1648. He was buried, as he directed, without "shew of mourning" at midnight. In a lively and somewhat tongue-in-cheek autobiography, *The Life of Edward Lord Herbert of Cherbury, Written by Himself* (originally published by

Horace Walpole in 1764), he tells the story of his life and escapades up to his recall from France in 1624. His main claims to fame in the history of thought lie in his philosophical views (for which he is justifiably known as the first English author of a purely metaphysical study and for which he was respected as well as criticized by Hugo Grotius, Pierre Gassendi, and René Descartes) and in his religious thought (in which he produced a pioneer work on the study of other religions, and for which he has traditionally, but arguably inaccurately, been described as "the father of English deism").

PHILOSOPHICAL THOUGHT

During his ambassadorship Herbert completed his major philosophical work, *De Veritate, Prout Distinguitur a Revelatione, a Verisimili, a Possibili, et a Falso* (On truth, in distinction from revelation, probability, possibility, and error), and had it privately printed in Paris in 1624. In this work he seeks to show, contrary to the doubts of "imbeciles and sceptics," that "Truth exists" (*De Veritate*, p. 83).

Although he was in touch with contemporary scholars, although his works show wide knowledge of classical, scholastic, and Renaissance literature and of hermetic literature, and although his arguments are sometimes less than persuasive, Herbert should not be seen as an eclectic thinker who merely puts forward a collection of sometimes discordant ideas that happen to attract him. Conflicts between his ideas are rather due to a failure to be sufficiently thorough in developing his innovative position.

THE NATURE OF TRUTH. On the title page of the second and third editions of *De Veritate*, Herbert dedicated his work to "every reader of sound and unprejudiced judgement" (this was different from the first edition, which had been grander in its amusingly presumptuous dedication to "the whole human race without qualification"), regarding himself as an original thinker who thinks "freely" and recognizes only the authority of "right reason," and using what is at times rather infelicitous Latin, Herbert aims to determine the nature of truth and the way in which it is identified by "every normal human being." He regards such an investigation as necessary if people are to know how to avoid the errors of skepticism, dogmatism, and fideism that corrupt current thought and lead some to hold that "we can know nothing," and others that "we can know everything."

Although Herbert regards right reason as the final judge of what is true, he also puts forward a doctrine of universal consent as the criterion for truth. He defends this doctrine on the grounds that universal consent must be due to the work of Providence, and hence what receives it cannot be doubted. The doctrine is somewhat paradoxical, however, since the need for some such criterion only arises where people do not agree about what is true, and hence where there is no universal consent about the matter. Herbert's views on this issue reflect the question-begging nature of his appeal to the authority of right reason: those who disagree with him about what receives universal consent may be ignored because their disagreement shows that they are not people of "sound and unprejudiced judgement" who are clearly following the dictates of right reason. It also reflects his conviction that the overall providence of God prevents what is erroneous from receiving universal consent.

THE DOCTRINE OF THE 'FACULTIES.' Herbert rejects the notion of the mind as a passive blank sheet on which the objects of its knowledge make their impressions. Nevertheless, while he holds that what people truly know is determined by the structure and activity of their minds, he seeks to show that what is known is, as common sense maintains, what is actually the case. To do this he puts forward his doctrine of the faculties, using the term *faculty* to refer to an internal power of the mind that links a particular perception with a particular object. According to this doctrine an object, whether intellectual or physical, is perceivable as such, and only so perceivable, because there is a corresponding faculty preestablished in the mind. Within the mind of each person there are as many latent faculties as there are differentiable objects, and the existence of a faculty shows the existence of a corresponding object. Defining truth as "a matter of conformity between objects and faculties" (*De Veritate*, p. 78), Herbert maintains that true knowledge of an object occurs when the appropriate latent faculty is activated. He also claims that an inner sense of satisfaction shows when an object has been correctly perceived by its corresponding faculty.

Herbert distinguishes between four classes of truth and between four types of faculty. According to the former division, the four classes of truth are: the truth of things as they are in themselves, the truth of how things manifest themselves to people, the truth of concepts that differentiate between things, and the truth of judgments on the deliverances of the other faculties. The four types of faculty are natural instinct, internal apprehension, external apprehension, and discursive thought. The first of these, natural instinct, is described as "that mode of apprehension which springs from the faculties which

conform to the Common Notions" (*De Veritate*, p. 115). These common notions are implanted in people by God. They are therefore all present, even if in many cases only latent, in every sane and whole person. The common notions are characterized by the qualities of priority, independence, universality, certainty, necessity (for one's preservation), and "the way of conformation" [*De Veritate*, p. 139–141] (in the sense of being immediately recognized and not needing to be warranted by discursive thinking). When brought to consciousness through appropriate stimulation, the common notions are acknowledged by all reasonable people as the normative principles for discerning what is true and good and for exposing what is false and bad.

As Herbert himself admits, the forms of the second type of faculty, internal apprehension, are not easily distinguished from those of the natural instinct. What characterizes them is that they concern a person's active response to particular objects. They may be spiritual, bodily, excited by external objects, or mixed. Under the guidance of the natural instinct, they make judgments about what is good and what is evil. Conscience is the highest of them. It applies the common notions to individual cases and is only satisfied when the faculties are correctly adjusted to what is the case. It is also through this faculty that people sense what is erroneous. External apprehension, the third type of faculty, concerns the ways by which people become aware of the external characteristics of objects and of their relationships with each other. Although most of Herbert's discussion of this kind of apprehension is concerned with the conventional five senses, he denies that there are only these five external modes of apprehension. He maintains that each sense is a channel for many external forms of apprehension since there are as many forms of apprehension as there are differences between objects.

The final type of faculty, discursive thought, is peculiar to humankind. It draws inferences from what comes to be known through the other faculties. It is more liable to error than they are, and so its findings are not to be preferred to what they directly discover. To establish the proper limits and methods of discursive reasoning, Herbert presents what he calls *Zetetica* and *Euretica* (the terms seem to have been coined by him, presumably from the Greek words for "to seek for" and "to discover"). These rules of reasoning appear to be indebted to Aristotelian logic. According to Herbert right reasoning proceeds by asking the appropriate faculties ten questions about the object of enquiry, namely, whether it exists, what it is, what kind of object it is, what its size is, to what it is related, and how, when, where, whence, and why it exists. Herbert assures his readers that by using this method complete and true knowledge of an object will be obtained.

PROBABILITY, POSSIBILITY, AND ERROR. *De Veritate* closes by considering probability, possibility, and error. The first of these deals with knowledge of the past, in the course of which discussion Herbert indicates the insecurity of beliefs based on historical judgments, the second with knowledge of the future, and the last with the sources of wrong judgments. In 1645 Herbert published *De Causis Errorum: Una cum Tractatu de Religione Laici, et Appendice ad Sacerdotes, nec non quibusdam Poematibus* (Concerning the causes of errors, with a treatise concerning religion for the laity, and an appendix to priests, with certain poems), some copies of which were bound in an enlarged, third edition of the *De Veritate*. It does not add significantly to the epistemological discussions in the earlier work. Errors and fallacies are held to be the result of failing to satisfy the conditions for grasping the truth laid down in *De Veritate*.

RELIGIOUS THOUGHT

Herbert's thought is probably most widely known for his discussion of the common notions concerning religion. This was included in his treatment of common notions in *De Veritate* and was expanded in later editions. Herbert holds that the criterion for true religious belief is not found in some supposed revelation or ecclesiastical authority, but in five common notions. They are that (1) there is a God (Herbert's term for God is *Supremum Numen* [Highest divinity]); (2) this God ought to be worshipped; (3) the connection of virtue with piety is and always has been the most important part of religious practice; (4) while people are aware of their evils, they can and must expiate them by repentance; and (5) people face reward or punishment after this life. These common notions are, for Herbert, the foundation of the true catholic or universal church that covers all humanity and the only authentic source of salvation.

In *De Religione Laici* (On religion for the laity; 1645) Herbert argues that a layman, using both reason and prayer, can and should decide between competing claims to belief by choosing that religion whose doctrines and practices are closest to the five common notions of religion discerned by the natural instinct. Commands about what is to be believed and practiced that have supposedly been imparted to people by revelation and transmitted by tradition should only be regarded as credible if they are

consistent with the common notions. Religious persecution, the pretensions of priestcraft, and restricted schemes of salvation are to be rejected. Authentic religion is realized in a virtuous life that conforms to the common notions of religion. This is fundamentally, if sometimes obscurely, recognized by people of right reason everywhere and at all times. According to Herbert *De Religione Laici* is written, not "with a mind … hostile to the best religion," but to make clear what follows from holding that "universal divine providence" is "the highest attribute of God" (*De Religione Laici*, p. 125). In the attached *Appendix ad Sacerdotes* Herbert reaffirms his argument that God, as a universal providence, must have provided all people with the means of salvation and that these are found in the common notions. Nothing more is needed. Additions to them, whether proposed by priests or the Bible, are to be judged unnecessary.

Herbert wrote two works, both published posthumously, in which he attempts to justify his claim that people everywhere and all times recognize the truth and normative status of the common notions of religion. The first of these, *De Religione Gentilium* (On the religion of the Gentiles), first published in 1663 (with an English translation by William Lewis appearing in 1705 entitled *The Ancient Religions of the Gentiles and Causes of Their Errors Consider'd*), is a pioneering work in the English study of other faiths. In it Herbert seeks to show that the evidence about religious belief and practice, which he derives for the most part from classical authors although there are some references to more recent reports, confirms that the common notions of religion are acknowledged everywhere. Evidence that seems to contradict this conclusion is rejected, for the most part on the grounds that it either is due to the corrupting effects of priestcraft or arises from a hermeneutical failure to appreciate symbolic usage. Herbert's sympathetic approach to non-Christian faiths did not blind him, however, to evils present in them. The "sound, most ancient and universal parts of religion" have to be abstracted from a vast heap of "superstitious rubbish" (*De Religione Gentilium*, Lewis translation, p. 292), largely introduced to serve priestly self-interest.

While in *De Religione Gentilium* Herbert concentrates on the first two common notions of religion, he focuses on evidence about the third and fourth in *A Dialogue between A Tutor and His Pupil* (published in 1768 and whose text, at least on the whole, is now generally accepted to be correctly ascribed to Herbert). Here again he denounces the corruptions and perversions of priestcraft while defending his conviction that the "five catholick articles" have been universally "engraved" in human "souls by the hand of God" (*A Dialogue between a Tutor and His Pupil*, p. 105).

HERBERT AND DEISM

As has been mentioned, Herbert has commonly been dubbed "the father of English deism." When examined, the evidence of his thought and practice provides strong grounds for questioning the justification of this description. Apart from the case of Charles Blount (an eclectic and sometimes plagiarizing author), it is not clear that Herbert's views influenced the thought of those later writers commonly said to be deist (itself a designation whose vagueness renders it more misleading than useful as a description when applied to English writers in the seventeenth and eighteenth centuries).

In his epistemological attempt to find a reasonable way between the dogmatic errors of bigotry and skepticism, Herbert holds that God is a universal providence whose active benevolence influences people's lives, that salvation is available to all through repentance, that prayer is efficacious, that people are to live virtuous lives, and that people have a postmortem personal existence in which they are judged but also may expect to find fulfillment. He does not doubt that divine revelations are given to individuals and, indeed, claims that he only decided to publish *De Veritate* after praying for and receiving a sign from heaven. At the same time, he does limit the significance of appeals to revelation, gives rules for authenticating them, and points out that there is a crucial difference between what is actually revealed on some occasion and what is passed down as a tradition of historical faith, especially when priests claim to be the authorized bearers of the tradition.

As for the Bible, Herbert is aware that different faiths assert the authority of different sacred books and so holds that what is taught by any of them, including the Bible, is to be judged against, and interpreted in terms of, the common notions of religion, since these alone undoubtedly express "the undoubted pronouncements of God, transcribed in the conscience" (*De Religione Laici*, p. 101). Rather than having the dubious distinction of being called "the father of English deism," an ascription deriving from eighteenth-century works by Thomas Halyburton (1674–1712), Philip Skelton (1707–1787), and John Leland (1691–1766), it seems more accurate to regard Herbert of Cherbury as a thinker with liberal convictions who attempts to identify an understanding of theistic belief that avoids the gross errors of religious fanaticism and unbelieving skepticism.

See also Blount, Charles; British Philosophy; Deism; Descartes, René; Gassendi, Pierre; Grotius, Hugo; Skepticism; Truth.

Bibliography

WORKS BY HERBERT OF CHERBURY

De Religione Gentilium. Amsterdam: Faksimile-Neudruck der Ausg. Amsterdam, 1663. An English-language translation of this work by William Lewis called *The Ancient Religion of the Gentiles and Causes of Their Errors Consider'd* was published in London by publisher John Nutt in 1705.

The Life of Edward Lord Herbert of Cherbury, Written by Himself. London: J. Dodsley, 1770.

De Veritate. Translated by Meyrick H. Carré. Bristol, U.K.: J. W. Arrowsmith, 1937.

De Religione Laici. Translated and edited by Harold R. Hutcheson. New Haven, CT: Yale University Press, 1944.

A Dialogue between a Tutor and His Pupil, edited by Günter Gawlick. Stuggart–Bad Cannstatt, Germany: Friedrich Frommann Verlag, 1971.

WORKS ABOUT HERBERT OF CHERBURY

Bedford, R. D. *The Defence of Truth: Herbert of Cherbury and the Seventeenth Century*. Manchester, U.K.: Manchester University Press, 1979.

Pailin, David A. "Should Herbert of Cherbury Be Regarded as a 'Deist'?" *Journal of Theological Studies* 51 (1) (2000): 113–149.

Sorley, W. R. "The Philosophy of Herbert of Cherbury." *Mind* 3 (1894): 491–508.

Webb, Clement C. J. *Studies in the History of Natural Theology*. Oxford, U.K.: Clarendon Press, 1915.

David A. Pailin (2005)

HERDER, JOHANN GOTTFRIED

(1744–1803)

Johann Gottfried Herder, German philosopher and critic, was born in Mohrungen in East Prussia. His father was a schoolteacher and he grew up in humble circumstances. In 1762 he enrolled at the University of Königsberg, where he studied with Kant, who accorded him special privileges due to his unusual intellectual abilities. At this period he also began a lifelong friendship with the irrationalist philosopher Johann Georg Hamann. In 1764 he left Königsberg to take up a schoolteaching position in Riga. There he wrote the programmatic essay *How Philosophy Can Become More Universal and Useful for the Benefit of the People* (1765); published his first major work, on the philosophy of language and literature, the *Fragments on Recent German Literature* (1767–1768); and also an important work in aesthetics, the *Critical Forests* (1769). In 1769 he resigned his position and traveled—first to France, and then to Strasbourg, where he met, and had a powerful impact on, the young Goethe. In 1771 he won a prize from the Berlin Academy for his best-known work in the philosophy of language, the *Treatise on the Origin of Language* (1772). From 1771 to 1776 he served as court preacher to the ruling house in Bückeburg. The most important work from this period is his first major essay on the philosophy of history, *This Too a Philosophy of History for the Formation of Humanity* (1774).

In 1776, partly through Goethe's influence, he was appointed General Superintendant of the Lutheran clergy in Weimar, a post he retained for the rest of his life. During this period he published an important essay in the philosophy of mind, *On the Cognition and Sensation of the Human Soul* (1778); a seminal work about the Old Testament, *On the Spirit of Hebrew Poetry* (1782); his well-known longer work on the philosophy of history, the *Ideas for the Philosophy of History of Humanity* (1784–1791); an influential essay in the philosophy of religion, *God. Some Conversations* (1787); a work largely on political philosophy, written in response to the French Revolution, the *Letters for the Advancement of Humanity* (1793–1797); a series of *Christian Writings* (1794–1798) concerned with the New Testament; and two works written in opposition to Kant's critical philosophy, the *Metacritique* (1799) (against the theoretical philosophy of the *Critique of Pure Reason*) and the *Calligone* (1800) (against the aesthetics of the *Critique of Judgment*).

Already in the 1760s Herder developed certain distinctive general positions in philosophy that would endure for the rest of his career. Most of these were strongly influenced by Kant, but by the *pre*critical Kant of the early and middle 1760s (not the critical Kant, against whom Herder later engaged in the public polemics just mentioned). Among these positions were: an insistence that philosophy should be useful for people in general; a (Pyrrhonist-influenced) skepticism about metaphysics, and about apriorism in philosophy; a form of empiricism; a (Hume-influenced) noncognitivism in ethics; and a principled rejection of ambitious forms of systematicity in philosophy. The early essay *How Philosophy* is especially revealing in this connection.

LANGUAGE

On the Origin is Herder's best-known work in the philosophy of language, but it is in certain respects unrepresentative and inferior in comparison with other works such as the *Fragments* and should not monopolize attention.

On the Origin is primarily concerned with the question whether the origin of language can be explained in purely natural, human terms or (as Johann Peter Süßmilch had recently argued) only in terms of a divine source. Herder argues for the former position and against the latter. His argument is fairly persuasive. But this is unlikely to constitute a modern philosopher's main reason for interest in Herder's ideas about language (deriving its zest, as it does, from a religious background that is no longer ours).

Of far greater modern relevance are the following three theses already embraced by Herder as early as the 1760s, the first two of which founded the philosophy of language as we know it today: (1) Thought is essentially dependent on, and bounded in scope by, language—that is, one can only think if one has a language, and one can only think what one can express linguistically. (2) Meanings or concepts are not to be equated with the sorts of items, in principle autonomous of language, with which much of the philosophical tradition has equated them—for example, the referents involved, Platonic forms, or empiricist ideas. Instead, they consist in *usages of words*. (3) Conceptualization intimately involves (perceptual and affective) sensation. More specifically, sensation is the source and basis of all our concepts, though we are able to achieve nonempirical concepts by means of a sort of metaphorical extension from the empirical ones, so that all of our concepts ultimately depend on sensation in one way or the other.

Herder also develops original theories of interpretation and translation founded on these principles. Fundamental to these theories is also a further insight: (contra such eminent Enlightenment philosopher-historians as Hume and Voltaire) peoples from different historical periods and cultures often vary radically in their concepts, beliefs, sensations, and so forth; and similar, albeit usually less dramatic, variations occur even between individuals within a single period and culture. This situation makes accurate interpretation and translation extremely difficult. In particular, it entails that interpreters and translators constantly need to resist a temptation to erroneously assimilate the concepts and thoughts they interpret or translate to their own. Herder develops his theories of interpretation and translation largely in response to this challenge.

His theory of interpretation (which is scattered through several works) stresses, inter alia, the need to complement a focus on word usages with a focus on historical context, authorial psychology, and literary genre; to "feel one's way into '*sich hineinfühlen in*'" the author's meaning-internal sensations; to adopt a rigorously empirical approach to determining all of these things; to use "divination," in the sense of tentative hypothesis, when advancing beyond the available empirical evidence, for example, for an author's psychology; and to interpret the parts of a work in light of the whole work. This theory exercised a huge influence on subsequent theories of interpretation, in particular Friedrich Daniel Ernst Schleiermacher's (which has often been mistakenly credited with introducing the psychological component of interpretation and "divination," which were in fact Herder's innovations).

Herder's theory of translation (which is mainly developed in the *Fragments*) stresses the need to "bend" word usages in the target language in order to faithfully reproduce the alien word usages and hence meanings in the source language; and the need to reproduce not only the semantic content but also the musical "form" (e.g., the meter) of the source text, because this conveys nuances of the sensations internal to the author's meanings. With these two principles, Herder founded a new paradigm of foreignizing translation that came to dominate German translation theory and practice during and after his lifetime.

MIND

In *On the Cognition* and elsewhere Herder develops an interesting and influential philosophy of mind. The following are its main features.

(1) He argues for an uncompromisingly *naturalistic* and *anti-dualistic* conception of the mind. In particular, he tries to erase the division between the mental and the physical in two specific and suggestive ways: First, he advances a theory that minds and their conditions consist in *forces* (*Kräfte*) that manifest themselves in people's bodily behavior—just as physical nature contains forces that manifest themselves in the behavior of bodies. Second, he undertakes to explain the mind in terms of the phenomenon of *irritation* (*Reiz*), a phenomenon recently identified by Haller, and exemplified by muscle fibers contracting in response to direct physical stimuli and relaxing upon their removal—in other words, a phenomenon that, while basically physiological, also seems to exhibit a transition to mental characteristics.

(2) Herder also argues that the mind is a *unity*, that there is no sharp division between its faculties. This position contradicts theorists from the period such as Sulzer and Kant. It is not entirely original with Herder, having already been central to Rationalism, especially Wolff. But Herder's version of it is original in certain respects, e.g. in rejecting Rationalism's reduction of sensation and voli-

tion to cognition, and in establishing the unity thesis in an empirical rather than an apriorist way.

(3) Herder also argues that linguistic meaning is fundamentally social, so that thought and other aspects of human mental life (as essentially articulated in terms of meanings), and hence also the very self (as essentially dependent on thought and other aspects of human mental life, and as defined in its specific identity by theirs), are so too. Herder's version of this position is again empirically based (unlike later versions of it, e.g., Hegel's).

(4) In tension, though not contradiction, with the preceding position, Herder also holds that even within a single period and culture human minds are as a rule deeply *individual*, deeply different from each other—so that in addition to a generalizing psychology we also need a psychology oriented to individuality. (5) Finally (like predecessors in the Rationalist tradition and Kant), Herder rejects the Cartesian idea of the mind's self-transparency—instead insisting that much of what occurs in the mind is unconscious, so that self-knowledge is often deeply problematic. This whole theory of the mind exercised an enormous influence on successors such as Hegel, Schleiermacher, and Nietzsche.

AESTHETICS

As already noted, Herder's philosophy of language is committed to the two doctrines that thought is essentially dependent on and bounded by language, and that meaning is word-usage. These doctrines seem to stand in tension with the expression of thought and meaning by the nonlinguistic arts, however. In the *Critical Forests* (1769) Herder initially tried to cope with this problem by denying that such arts express thought or meaning, but that is an implausible position, and in later parts of the work he began to develop a much more plausible solution: they do so, but the thoughts and meanings in question depend on a prior linguistic articulation or articulability by the artist (so that the interpretation of "nonlinguistic" art requires the interpretation of language). This was henceforth Herder's normal position.

Another important Herderian contribution to aesthetics is his theory of *genre*. Herder believes, plausibly, that a work of art is always written or made to exemplify a certain genre, and that it is vitally important for an interpreter to identify the work's genre if he is to understand it correctly. Herder's basic conception of genre is that it consists in an overall purpose together with certain rules of composition dictated thereby. Genres are to a great extent socially pregiven, but they vary from society to society, and they always play their role via authorial intention, so that the individual artist is not inexorably locked into them but can and often does modify them.

Herder has two reasons for thinking it vitally important to define a work's genre correctly if one is to understand the work properly (both good reasons): first, because an author intends his work to exemplify a certain genre, there will normally be aspects of the work's meaning that are expressed, not explicitly in any of its parts, but rather through its intended exemplification of the genre; second, correctly identifying the genre is also required for correctly interpreting things that *are* expressed explicitly in the parts of a work. Just as Herder insists on a scrupulously empirical approach to interpretation generally, so he insists on it in connection with determining genres in particular; he sharply rejects apriorism here, including the relative apriorism of generalizing from certain familiar examples. Such relative apriorism is disastrous, in his view, because the superficial appearance of a similar genre shared by different historical periods or cultures, or even by different authors within one period and culture, or even by a single author in one work and in another, usually masks important differences.

Herder sees misguided relative apriorism in the definition of genres in many areas of interpretation. For example, his essay *Shakespeare* (1773) discerns it in the French critics' approach to tragedy, which assumes the universal validity in tragedy of Aristotelian genre-purposes and -rules originally derived exclusively from ancient tragedies, and consequently assumes that these provide an appropriate guide for interpreting Shakespearean tragedy, whose genre-conception is in fact different; and *This Too* discerns it in Johann Joachim Winckelmann's approach to Egyptian sculpture, which erroneously imports genre-purposes and -rules derived from Greek sculpture. Herder also stresses that determining the genre properly is vitally important not only for the correct *interpretation* of an artwork, but also for its correct *evaluation*. For example, the French critics not only make an interpretive mistake when they go to Shakespeare with a genre dogmatically in mind that was not his, but also an evaluative one: because they falsely assume that he must be aspiring to realize the genre-purpose and -rules that Aristotle found in ancient tragedy, they fault him for failing to realize these, while at the same time they overlook the different genre-purpose and -rules that he really does aspire to realize and his success in realizing these.

HISTORY

Herder's philosophy of history appears mainly in two works, *This Too* and the later *Ideas*. These works are famous for their teleological conception of history as the progressive realization of a divine purpose (unspecified in the former work, but specified in the latter as the realization of "humanity" or "reason"). This conception was influential on subsequent thinkers (especially Hegel). However, the philosophical interest of Herder's works today lies elsewhere.

Herder's main achievement here arguably consists in his insight into, and detailed empirical elaboration of, the thesis mentioned earlier that (contra such Enlightenment philosopher-historians as Hume and Voltaire) there exist radical mental differences between historical periods, that people's concepts, beliefs, sensations, and so on differ in deep ways from one period to another. This thesis is already prominent in *On the Change of Taste* (1766). It too exercised an enormous influence on successors (e.g., Hegel, Nietzsche, and Dilthey).

Herder indeed makes the empirical exploration of the realm of mental diversity posited by this thesis the core of the discipline of history. He takes relatively little interest in the supposedly "great" political and military deeds and events of history, focusing instead on this varying "innerness" of history's participants (consequently, for him *psychology* and *interpretation* take center-stage as methods in the discipline). This is a deliberate and self-conscious choice for which he has deep reasons: On the one hand, he is skeptical of the traditional justifications for a history that focuses on the "great" political and military deeds and events of the past, justifications in terms of their being morally edifying (his values rather incline him to find them morally repugnant), revealing an overall meaning in history (despite his own official teleology, he is skeptical about this), or affording efficient causal insights that will enable us to explain the past and predict or control the future (he considers the potential for such insights and benefits severely limited). On the other hand, he sees positive reasons for focusing on the "innerness" of human beings in history: His discovery of radical diversity in human mentality has shown there to be a much larger, less explored, and more intellectually challenging field for investigation here than previous generations of historians had realized.

Also, studying people's minds through their literature, visual art, and so on promises to contribute to our moral self-improvement, since, unlike political-military history, it exposes us to people at their moral best and hence is morally edifying, and it serves cosmopolitan and egalitarian moral ideals by enhancing our sympathies for peoples and indeed for peoples at all social levels, including lower ones. Finally, doing "inner" history is also valuable as an instrument for our *non*moral self-improvement: It advances our self-understanding, because contrasting our own outlook with the outlooks of other peoples enables us to recognize what is universal and invariant in it and what by contrast distinctive and variable, and because in order fully to understand our own outlook we need to identify its origins and how they developed into it (this is Herder's famous and influential "genetic method"); and additionally, by investigating the nonmoral ideals of past ages (e.g., their aesthetic ideals) it enables us to enrich our own nonmoral ideals and hence happiness. This whole position strongly influenced successors, especially Wilhelm Dilthey.

Herder is also impressive for having recognized, and if not solved then at least grappled with, a problem that flows from his picture of history (and intercultural comparison) as an arena of radical variations in human mentality: the threat of *skepticism*. Herder is determined to avoid skepticism. He vacillates between two main strategies for doing so that are inconsistent with each other: His first is to acknowledge the problem in its full force but to respond to it with relativism: especially in *This Too* he argues that (at least where questions of moral, aesthetic, and prudential value are concerned) the different positions taken by different periods and cultures are equally valid, namely for the periods and cultures to which they belong, and that there can be no question of any preferential ranking between them. His second strategy is to try to defuse the problem at its source by arguing that, on closer inspection, there is in fact much more common ground between different periods and cultures than it allows. This strategy plays a central role in the *Ideas*, where in particular "humanity" is presented as a common ethical value. The later *Letters* goes back and forth between these two strategies.

POLITICS

Herder's most developed statement of his political philosophy appears in a late work prompted by the French Revolution of 1789: the *Letters* (including the early draft of 1792). In domestic politics, the mature Herder is a liberal, a republican, a democrat, and an egalitarian. In international politics, he has often been classified as a "nationalist" or (even worse) a "German nationalist," but this is misleading and unjust. On the contrary, his fundamental position is a committed *cosmopolitanism*, an impartial concern for *all* human beings. This is a large

part of the force of his ideal of "humanity." Hence, for example, in the *Letters* he quotes with approval François de Salignad de la Mothe Fénelon's remark, "I love my family more than myself; more than my family my fatherland; more than my fatherland humankind." (2002, p. 389). Moreover, unlike the cosmopolitanism of his teacher Kant, Herder's is genuine: Whereas Kant's is vitiated by a set of empirically ignorant and morally inexcusable prejudices that he harbors—in particular, racism, antisemitism, and misogyny—Herder's is entirely free of these prejudices, which he indeed works tirelessly to combat. Herder does *also* insist on respecting, preserving, and advancing national groupings. But he has good reasons for doing so: (1) The deep diversity of values between nations entails that homogenization is ultimately impracticable, only a fantasy. (2) It also entails that, to the extent that it *is* practicable, it cannot occur voluntarily but only through external coercion. (3) In practice, attempts to achieve it—for example, by European colonialism—are moreover coercive from, and subserve, ulterior motives of domination and exploitation. (4) Furthermore, real national variety is positively valuable, both as affording individuals a vital sense of local belonging and in itself.

Moreover, Herder's insistence on respecting, preserving, and advancing national groupings is unalarming, for the following reasons: (1) For Herder, this is emphatically something that must be done for *all* national groupings *equally* (not just or especially Germany!). (2) The "nation" in question is not racial but linguistic and cultural. (3) Nor does it involve a centralized or militaristic state (Herder advocates the disappearance of such a state). (4) Herder's insistence on respecting national groupings is accompanied by the strongest denunciation of military conflict, colonial exploitation, and all other forms of harm between nations; a demand that nations instead peacefully cooperate and compete in trade and intellectual endeavors for their mutual benefit; and a plea that they should indeed actively work to help each other.

On the one hand, Herder's political philosophy can appear theoretically thin, but this is intentional and arguably a virtue, a salutary minimalism. There is certainly no grand metaphysical theory underpinning it (à la Fichte or Hegel). But that is deliberate, given his skepticism about metaphysics. Nor does he have an elaborate account purporting to justify the moral intuitions at work in it as a sort of theoretical insight (à la Kant or Rawls). But that is again deliberate, given his noncognitivism in ethics. Nor does he call on such tired staples of political theory as the state of nature, the social contract, natural rights, the general will, and utopias for the future. But again, he has good specific reasons for skepticism about these. On the other hand, he *does* have a "political theory" of another, and arguably more valuable, sort. For one thing, in accordance with his general empiricism, his political philosophy is deeply empirically informed (e.g., he argues that freedom of thought and expression is required for the advancement of truth and artistic creativity by appeal to historical examples, especially classical Athens). For another thing, conformably with his noncognitivism in ethics, he is acutely aware that his political position ultimately depends on moral sentiments—his own and, for its success, other people's as well—and this leads him to engage in moral theorizing of another sort, namely theorizing about how, and by what means, people's moral sentiments should be molded in order to realize his political ideals. *These* two sorts of theorizing *are* deeply developed in Herder's political philosophy.

RELIGION

Religion was a lifelong preoccupation of Herder's. He made important contributions to the theory of biblical interpretation and to the actual interpretation of the Bible (in particular, insisting on and applying the same sort of rigorous secular approach to interpretation that he advocates for profane texts). In addition, he played a major role in reviving a form of Spinozism, a position to which he was already attracted early in his career, but to which he gave fullest expression in his neo-Spinozistic *God. Some Conversations* of 1787. In this work he develops a version of Spinozism that consciously modifies the original in certain respects. He shares with Spinoza the basic thesis of *monism*, and like Spinoza equates the single, all-encompassing principle with God. But whereas Spinoza characterized it as *substance*, Herder characterizes it as *force*, or *primal force*.

Moreover, this modification involves further ones, including these: (1) Whereas Spinoza's theory rejected conceptions of God as a full-blooded person or mind, and a being who not only thinks but also has purposes, Herder's identification of God with force imports, thanks to his general identification of mind with force, a claim that God *is* in fact a mind, and and a being who not only thinksbut also has purposes. (2) Herder believes that Spinoza's original theory contained a residue of objectionable dualism, inherited from Descartes, in its conception of the relation between God's two known attributes, thought and extension (and similarly in its conception of the relation between finite minds and bodies); by contrast, Herder's conception of God as a force (and of finite

minds as likewise forces) overcomes this residual dualism, since forces are of their very nature expressed in extended bodies. Herder's neo-Spinozism, including these modifications, was largely responsible for a great wave of neo-Spinozism that swept German philosophy in this period (embracing Goethe, Schelling, Hegel, Schleiermacher, Hölderlin, Novalis, F. Schlegel, and others).

See also Descartes, René; Dilthey, Wilhelm; Empiricism; Fénelon, François de Salignac de la Mothe; German Philosophy; Goethe, Johann Wolfgang von; Haller, Albrecht von; Hamann, Johann Georg; Hegel, Georg Wilhelm Friedrich; Hölderlin, Johann Christian Friedrich; Hume, David; Kant, Immanuel; Nietzsche, Friedrich; Noncognitivism; Novalis; Philosophy of History; Philosophy of Language; Philosophy of Mind; Political Philosophy, History of; Pyrrho; Rationalism; Schelling, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von; Schleiermacher, Friedrich Daniel Ernst; Skepticism; Spinoza, Benedict (Baruch) de; Spinozism; Sulzer, Johann Georg; Voltaire, François-Marie Arouet de; Winckelmann, Johann Joachim; Wolff, Christian.

Bibliography

WORKS BY HERDER

Herders Sämtliche Werke, edited by B. Suphan et al. Berlin: Weidmannsche Buchhandlung, 1887–1913.

Johann Gottfried Herder Werke, edited by U. Gaier et al. Frankfurt am Main, Germany: Deutscher Klassiker Verlag, 1985–.

Outlines of a Philosophy of the History of Man. Translated by T. Churchill. London: Luke Hansard for J. Johnson, 1803. This is a translation of the *Ideas*.

God. Some Conversations. Translated by F. H. Burkhardt. New York: Veritas Press, 1940.

Forster, M. N. *Herder: Philosophical Writings*. Cambridge, U.K.: Cambridge University Press, 2002. Contains translations of *How Philosophy*, *On the Origin*, *On the Cognition*, and *This Too*, as well as various other pieces.

Barnard, F. M. *J. G. Herder on Social and Political Culture*. Cambridge, U.K.: Cambridge University Press, 1969. Contains partial translations of Herder's 1769 *Journal*, *On the Origin*, *This Too*, the *Dissertation on the Reciprocal Influence of Government and the Sciences*, and the *Ideas*.

Nisbet, H. B. *German Aesthetic and Literary Criticism: Winckelmann, Lessing, Hamann, Herder, Schiller, Goethe*. Cambridge, U.K.: Cambridge University Press, 1985. Contains two pieces by Herder on aesthetics, including his *Shakespeare*.

WORKS ABOUT HERDER

Barnard, F. M. *Herder's Social and Political Thought: From Enlightenment to Nationalism*. Oxford, U.K.: Clarendon Press, 1965. Chapters 3–5 on political philosophy are especially good.

Beiser, F. C. *Enlightenment, Revolution, and Romanticism: The Genesis of Modern German Political Thought, 1790–1800*. Cambridge, MA: Harvard University Press, 1992. See esp. chap. 8. An excellent treatment of Herder's political philosophy.

Beiser, F. C. *The Fate of Reason: German Philosopy from Kant to Fichte*. Cambridge, MA: Harvard University Press, 1987. Chap. 5. Covers several topics well, including Herder's philosophies of language, mind, and religion.

Berlin, I. *Vico and Herder: Two Studies in the History of Ideas*. New York: Viking Press, 1976. A concise and excellent overview.

Clark, R. T., Jr. *Herder: His Life and Thought*. Berkeley: University of California Press, 1955. Detailed and useful, though unimaginative.

Ergang, R. *Herder and the Foundations of German Nationalism*. New York: Columbia University Press, 1931. Helpful on Herder's political thought and on his intellectual influence.

Haym, R. *Herder nach seinem Leben und seinen Werken*. Berlin: Gaertner, 1880. A detailed and excellent intellectual biography. Still the best single work on Herder.

Irmischer, H. D. *Johann Gottfried Herder*. Stuttgart: Reclam, 2001. A short and helpful overview.

Lovejoy, A. O. "Herder and the Enlightenment Philosophy of History." In his *Essays on the History of Ideas*. Baltimore: Johns Hopkins Press, 1948. A good introduction to Herder's philosophy of history.

Meinecke, F. *Historism: The Rise of a New Historical Outlook*. London: Routledge, 1972. Chap. 9. Helpful on Herder's philosophy of history.

Nisbet, H. B. *Herder and the Philosophy and History of Science*. Cambridge, MA: Modern Humanities Research Association, 1970. A helpful account of Herder's stance towards science.

Norton, R. E. *Herder's Aesthetics and the European Enlightenment*. Ithaca, NY: Cornell University Press, 1991. Helpful both on certain aspects of Herder's aesthetic theory and on Herder's general relation to the Enlightenment.

Zammito, J. H. *Kant, Herder, and the Birth of Anthropology*. Chicago: University of Chicago Press, 2001. Especially helpful on Herder and the precritical Kant.

Michael N. Forster (2005)

HERMENEUTICS

While there is already a general sense of the word "hermeneutics" in ancient Greek thought, where it refers to the problems of interpretation and understanding, the first real consolidation of its meaning comes in the medieval world when the peculiar task of interpreting the Bible is theorized. The first systematic form of hermeneutic theory emerged out of the effort to supply methods and rules for biblical commentary. Hermeneutics as a theory of biblical exegesis was subsequently widened to include the concerns of interpreting juridical texts, where

the jurist faces the problem of applying a universal rule to particular cases. Over time, the domain of hermeneutic methodology was broadened to include any text the meaning of which could be disputed. Although a wide range of texts became the objects of hermeneutics, theological and legal texts long remained its preeminent concerns. In its primary concern for textual exegesis, hermeneutics tended to develop methods for interpretation and understanding based upon the rhetorical principles, and thus helped to define the difference between the humanities and modern natural science with its own emphasis on a method linked with mathematics.

The hermeneutic tradition underwent significant expansion and modification in the nineteenth century when first Friedrich Schleiermacher and Friedrich von Schlegel, and then Wilhelm Dilthey, expanded the scope of hermeneutical concerns while probing the character of the foundations of hermeneutic practices. Following in the wake of Immanuel Kant's critical philosophy, wherein both the conditions that render experience possible and the ineradicable limits of knowledge are exposed, the Romantics argued that all understanding—not simply the understanding of texts—is always already interpretive. In Romanticism, hermeneutics thus begins to take on the contours of the new meaning that it acquired in the twentieth century: it is no longer simply a matter of a strategy directed to the interpretation of a special domain of texts, rather it is now understood to be concerned with the character of any form of understanding that may emerge from human experience. In addition to this is the claim that all understanding takes place within language, making language thus one of the chief concerns of any hermeneutic theory. The plurality of languages, their history, and the problem of translation replace a concern with the word of God and the word of law dominating earlier conceptions of hermeneutics. Dilthey, for his part, elevated hermeneutics into a methodology for the entirety of the human sciences by insisting that the understanding of the historical life expressions, which encompass human experiencing, requires a methodology distinct from that of natural science. Dilthey maintained that whereas the natural sciences explain nature, the task of the human sciences is to understand historical life.

By the end of the nineteenth century, "hermeneutics" had ceased to designate simply a methodology or doctrine concerned with decoding the meaning and truth claims of texts. Instead, it had become the name for a broader methodology and a philosophical approach to experience that was sensitive to the limits of language and history. Hermeneutics at this stage of its development came to be especially attentive to those experiences directly challenging the possibility of understanding; for example, the translation of foreign languages, the comprehension of foreign cultures, and, in particular, the interpretation of other historical periods.

Martin Heidegger took the decisive steps in formulating the contemporary shape of philosophical hermeneutics as it is understood today, achieving this by gathering together and radicalizing the concerns dominating its prehistory while adding a new dimension whereby hermeneutics became the name for a full-fledged ontology. Heidegger does this under the rubric of a "hermeneutics of facticity." That notion, which Heidegger worked out in his lecture courses during the 1920s (above all in his courses dealing with Aristotle), is consolidated in his 1927 magnum opus, *Being and Time*. There Heidegger argues that understanding is not simply a cognitive task, but that it names one of the basic ways (existentialia) of being-in-the world. In short, understanding is now taken to be concerned with an experiencing that subtends methodological procedure. The form of such lived experience proper to human beings, for whom being is always a question and always defined by death and the ineluctability of nonbeing, is what Heidegger refers to as "factical life."

When Heidegger speaks of the hermeneutics of factical life it is a way of acknowledging both that the concerns of hermeneutics—language, history, and finitude—and the manner in which it takes truth to be a matter of interpretation rather than objectivity are especially well suited for the attempt to theorize factical life. The phenomenology of lived experience is now said to have the character of a hermeneutics. This means that lived experience is taken to be the working out of the factical conditions upon which any understanding whatsoever can be founded. The analysis of existence thus takes the form of a hermeneutics that traces the action of these conditions of understanding. The most important aspect of this new development is that now even self-understanding comes to be presented as a hermeneutic task. Hermeneutics is thus the manner in which existence discloses the truth of a world that is lived and it is the form by which self-understanding is achieved.

After *Being and Time*, Heidegger uses the word "hermeneutics" less frequently. It will be left to one of Heidegger's students from those lecture courses of the 1920s, Hans-Georg Gadamer, to systematically develop the notion of hermeneutics as a philosophical standpoint. Gadamer, whose name is most closely associated with the idea of contemporary philosophical hermeneutics, does

this most extensively in his magnum opus, *Truth and Method* (1982).

The title *Truth and Method* alludes to the early sense of hermeneutics, where it was understood as a method for getting at the truth of texts. But the argument of that book entails both a fundamental rethinking of the notion of truth and a powerful critique of the idea that a method can ever yield it. In *Truth and Method*, Gadamer identifies hermeneutics with the insight that the concept of method is inappropriate for the task of understanding in the domain of the human sciences. Other guideposts take the place of method in the effort to unfold a truth that is understood as belonging to the realm of a historical event rather than objective fact: Language, tradition, questioning, and conversation become the leading concerns of Gadamer's hermeneutics.

Gadamer creatively draws upon several sources for his formulation of a systematic philosophical hermeneutics. In addition to those figures already mentioned, Aristotle's notion of *phronesis* (prudence or practical wisdom) in his *Ethics*, the logic of question and answer as it is found in the Platonic dialogues, Kant's understanding of judgment as well as the relation of art and truth in his Third Critique, and G. W. F. Hegel's notion of the formation of traditions all play pivotal roles in Gadamer's hermeneutics. Without any significant departure from Heidegger's way of opening up the notion of philosophical hermeneutics, Gadamer places a greater emphasis on the relevance of three themes for hermeneutics: the role of *art* in the disclosure of truth, the force of the prejudices of *tradition* in any understanding, and the importance of the *question* in the opening of the restrictions of such prejudices and in the liberation of understanding to the new and the foreign.

Gadamer understands hermeneutics not as a method, but more as a sort of dialogue or conversation in which understanding increases insofar as one becomes aware of the formative roles of history and language in one's self-understanding. In such a genuine dialogue with others, one's self-understanding is challenged to reflect upon and reach beyond the limits that are inscribed in its own roots in tradition and language. With Gadamer, hermeneutics comes to refer to a philosophical sensibility that has a deep commitment to exposing the ways in which all forms of understanding, rooted in self-understanding, is finite and so remains always at best a task and an ideal.

One other key figure in the field of contemporary hermeneutics is Paul Ricoeur. Ricoeur's work has been marked both by its extension of hermeneutic concerns to include psychoanalysis, literary criticism, and linguistic analysis, as well as by the details of his treatment of issues such as problems in semantics, metaphor, narrative, and temporal structures. In his earlier work Ricoeur attempted to reintegrate the role of explanation into hermeneutics theory by relying on the insights from linguistic structuralism, while in his later writings Ricoeur was less prone to pursue methodological questions. The originality of Ricoeur's hermeneutics has taken shape primarily as a matter of practices and studies of special themes, rather than as a theory of hermeneutics proper. What one sees most in those studies is how the workings of language and time have come to dominate his sense of the task of hermeneutic reflection.

See also Dilthey, Wilhelm; Gadamer, Hans-Georg; Heidegger, Martin; Phenomenology; Ricoeur, Paul; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

Bruns, Gerald. *Hermeneutics: Ancient and Modern*. New Haven, CT: Yale University Press, 1992.

Gadamer, Hans-Georg. *Heidegger's Ways*. Translated by John Stanley. New York: SUNY Press, 1994.

Gadamer, Hans-Georg. *Philosophical Hermeneutics*. Translated by David Linge. Berkeley: University of California Press, 1976.

Gadamer, Hans-Georg. *Truth and Method*. Translated by Garrett Barden and John Cumming. New York: Crossroad, 1982.

Grondin, Jean. *The Sources of Hermeneutics*. New York: SUNY Press, 1995.

Heidegger, Martin. *Being and Time*. Translated by Joan Stambaugh. New York: SUNY Press, 1996.

Heidegger, Martin. *Ontology—The Hermeneutics of Facticity*. Translated by John van Buren. Bloomington: Indiana University Press, 1999.

Makkreel, Rudolf. *Dilthey: Philosopher of the Human Sciences*. Princeton, NJ: Princeton University Press, 1975.

Ricoeur, Paul. *From Text to Action: Essays on Hermeneutics*. Translated by Kathleen Blamey and John Thompson. Evanston, IL: Northwestern University Press, 1991.

Ricoeur, Paul. *Hermeneutics and the Human Sciences* edited by John Thompson. New York: Cambridge University Press, 1988.

Risser, James. *Hermeneutics and the Voice of the Other*. New York: SUNY Press, 1997.

Smith, Christopher. *Hermeneutics and Human Finitude*. New York: Fordham University Press, 1991.

Dennis J. Schmidt (2005)

HERMETICISM

"Hermeticism" is the outlook associated with the Hermetic writings, a literature in Greek that developed in the early centuries after Christ under the name "Hermes Trismegistus." Much of it is concerned with astrology, alchemy, and other occult sciences, but there is also a philosophical Hermetic literature. The treatise known as the *Asclepius* and the collection of treatises grouped as the *Corpus Hermeticum* are the most important of the philosophical Hermetica, though some other fragments are preserved in the anthology of Stobaeus. These writings are probably to be dated between 100 and 300 CE. They are an amalgam of Greek philosophy, particularly Platonic, with other elements from the heterogeneous late antique culture. The *Pimander*, the first treatise in the *Corpus Hermeticum*, has obvious affinities with Genesis, suggesting an influence of Hellenistic Judaism. There may also be Persian influences, and the possibility of some contact with Christianity cannot be excluded. The ascription of their authorship to "Hermes Trismegistus," supposed to be an Egyptian priest, encouraged the belief that these writings transmitted ancient Egyptian wisdom; the *Asclepius* in particular has a strong pseudo-Egyptian coloring.

There is much difference of opinion among scholars as to the various elements that make up the Hermetica, which are the work of an unknown number of unknown authors; even individual treatises may often be a fusion of fragments. They have a certain unity of tone, however, since they all exhibit a similar type of philosophical-religious approach to the cosmos, involving regenerative experiences and outbursts of religious ecstasy. It has been suggested that they may be the literature of a gnostic sect. The philosophical Hermetica, with their lofty aspirations, cannot be altogether isolated from the magical and occult type of literature which also goes under the name of "Hermes Trismegistus," for the experiences of the Hermeticist, as described in the philosophical-religious treatises, take place within an astrological framework and imply, particularly in the *Asclepius*, a religious use of magic.

HISTORY

Although much is in debate concerning the Hermetica themselves, we are on firmer ground when we come to the history of their legend. In the fourth century Lactantius taught that these writings were the work of an Egyptian seer who lived not long after the time of Moses, whose account of creation he confirmed and, indeed, improved and whose mentions of a "son of God" were prophetic of Christianity and to be compared with passages in the Gospel according to St. John. Augustine also believed in the extreme antiquity of "Hermes Trismegistus," but he disapproved of the magical cult described in the *Asclepius*. Nevertheless, there was ample authority in Christian writers for an attitude of respect for Hermes. Lactantius places him with the sibyls as a Gentile prophet of Christianity. The myth of "Hermes Trismegistus," the Egyptian sage who was the actual author of all the writings assigned to him and who lived long before the Incarnation, which he prophetically foresaw, was to give great authority to the Hermetica.

The *Asclepius* was known in the Middle Ages in the Latin translation wrongly attributed to Apuleius of Madaura; certain pseudo-Hermetic writings were also known. The collection of treatises grouped as the *Corpus Hermeticum* seems to have been already known in this form to Psellus in the eleventh century but did not reach the West until the Renaissance.

INFLUENCE ON RENAISSANCE

The Hermetica made an impact on the Renaissance the importance of which has begun to be realized only in recent years. About 1460 a manuscript containing an incomplete Greek text of the *Corpus Hermeticum* was brought to Florence. Cosimo de' Medici ordered Marsilio Ficino to translate this at once into Latin, before beginning his translation of the works of Plato. This illustrates the Renaissance attitude, which treated the Hermetica as texts much more ancient than the Platonic writings and as the "Egyptian wisdom" believed to be one of the founts of *prisca theologia* that descended in an unbroken line to Plato and the Neoplatonists. When Ficino found scraps of Platonic philosophy in the late antique Hermetica, he assumed that he was dealing with the ancient Egyptian source of Greek wisdom. Like the interpretation of "Hermes Trismegistus" as a Gentile prophet, in which Ficino also firmly believed, this view of the Hermetic writings as a source of Plato and the Platonists depended on the misdating of those writings. To this most influential error is due the fact that there is a Hermetic core to Renaissance Neoplatonism. Ficino's work on astral magic is based on the magical passages in the *Asclepius*. Giovanni Pico della Mirandola opened his *Oration on the Dignity of Man* with a quotation from the *Asclepius*.

Throughout the sixteenth century the Hermetic writings were eagerly read in the many editions of Ficino's translation, and new editions and commentaries were published by Jacques Lefèvre d'Étaples, Symphorien

Champier, F. Foix de Candale, Francesco Patrizi, and others. The first edition of the Greek text of the *Corpus Hermeticum* appeared in 1554.

The influence of this intensive study of the Hermetica can be traced throughout the Renaissance. It penetrated some types of Renaissance theology. Christian Hermeticists who wished to avoid the magic excluded the magical passages in the *Asclepius* from their canon. On the other hand, for Renaissance magicians and philosophers the animist and magical view of nature that they extracted from the Hermetic writings was the most attractive feature. A striking instance of Hermetic influence on a Renaissance philosopher is Giordano Bruno, who rejected the Christian interpretation of the Hermetica and regarded Hermeticism as a pure Egyptian religion and philosophy that he made the basis of his animist interpretation of nature.

In 1614 the great Greek scholar Isaac Casaubon dated the Hermetica as written in post-Christian times, thus shattering the myth of their ancient Egyptian authorship on which Renaissance Hermeticism had rested. With the rise of seventeenth-century thought the influence of Hermeticism receded, though there were many survivals of the Renaissance attitudes to the Hermetic writings. The part played in the immediately premodern period by Renaissance Hermeticism in the directing religious attention toward the cosmos and toward operating with cosmic powers has yet to be assessed.

See also Bruno, Giordano; Ficino, Marsilio; Literature, Philosophy of; Neoplatonism; Patrizi, Francesco; Pico della Mirandola, Count Giovanni; Plato; Platonism and the Platonic Tradition; Renaissance.

Bibliography

TEXTS AND TRANSLATIONS

Copenhaver, Brian P., tr. *Hermetica: The Greek Corpus Hermeticum and the Latin Asclepius*. New York: Cambridge University Press, 1992.

Dodd, C. H. *The Interpretation of the Fourth Gospel* (1963). Reprint, Cambridge, U.K.: Cambridge University Press, 1992.

Festugière, André Marie Jean. *La révélation d'Hermès Trismégiste*. 4 vols. Paris: Belles Lettres, 1990. Reprint of the 2nd edition by J. Gabalda, 1950.

Festugière, André Marie Jean, and Arthur Darby Nock, eds. *Corpus Hermeticum*. 4 vols. Text and French translations of *Corpus Hermeticum*. Paris: Belles Lettres, 2002. Reprint of the 1946–1954 editions.

Georgi, Dieter, and John Strugnell, eds. *Concordance to the Corpus Hermeticum*. Cambridge, MA: Boston Theological Institute, 1971.

Nock, Arthur Darby. *Conversion: The Old and the New in Religion from Alexander the Great to Augustine of Hippo*. Baltimore: Johns Hopkins University Press, 1998. First published by the Oxford University Press in 1933.

Reitzenstein, Richard. *Poimandres: Studien zur grieschisch-ägyptischen und frühchristlichen literatur*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1996. First published by B. G. Teubner in 1904.

Salaman, Clement, Dorine van Oyen, William D. Wharton, and Jean-Pierre Mahé, eds. *The Way of Hermes: New Translation of the Corpus Hermeticum and the Definition of Hermes Trismegistius to Asclepius*. Rochester, VT: Inner Traditions, 2000.

WORKS ON HERMETICISM

Debus, Allen G., and Ingrid Merkel, eds. *Hermeticism and Renaissance: Intellectual History and Occult in Early Modern Europe*. Cranbury, NJ: Associated University Press, 1988.

Dodd, C. H. *The Bible and the Greeks* (1935). Reprint, London: Hodder and Stoughton, 1964.

Fowden, Garth. *The Egyptian Hermes: A Historical Approach to the Late Pagan Mind*. Cambridge, U.K.: Cambridge University Press, 1986.

Kloppenborg, John S., and Stephen G. Wilson, eds. *Voluntary Associations in the Graeco-Roman World*. London: Routledge, 1996.

Kristeller, Paul Oskar, ed. *Supplementum ficinianum: opuscula inedita et dispersa* (1937). Reprint, Florence: L. S. Olschki, 1973.

Lazzarelli, Ludovico, Francesco D. Ana, Heinrich Cornelius Agrippa von Nettesheim, and Eugenio Garin, eds. *Testi umanistici su l'ermetismo*. Rome: Fratelli Bocca, 1955.

Valantasis, Richard. *Spiritual Guides of the Third Century: A Semiotic Study of the Guide-Disciple Relationship in Christianity, Neoplatonism, Hermetism, and Gnosticism*. Minneapolis: Fortress Press, 1991.

Walker, Daniel Pickering. *Spiritual and Demonic Magic: From Ficino to Campanella*. University Park: Pennsylvania State University Press, 2000. First published by the Warburg Institute, University of London, in 1958.

Wilson, Robert McLachlan. *The Gnostic Problem*. New York: AMS Press, 1980. Reprint of the 1958 edition by Mowbray.

Yates, Frances A. *Giordano Bruno and the Hermetic Tradition*. New York: Routledge, 2001. First published by the University of Chicago in 1964.

Frances A. Yates (1967)
Bibliography updated by Kevin Moore (2005)

HERMETICISM [ADDENDUM]

Hermeticism also flourished in the Jewish and Islamic world. Although only a fragment of the Greek *Corpus Hermeticum* has been discovered in Arabic, there are

numerous texts in Arabic and Hebrew that are attributed to the Greek god Hermes, purporting to provide ancient wisdom. Hermes was identified respectively with Idris; a mysterious prophet mentioned in the Qur'an) and Enoch (the grandfather of Noah) in the Jewish Bible. The routes by which this Hermes arrived in Arabic texts appear to be as much via Persia as directly from Greek sources, and Haran (ancient Carrhae) in northern Mesopotamia, as a cultural melting point and the home of the Sabaeans, who worshipped the planets, appears to have played a central role. Several interrelated cosmological texts that give instructions on the invocation of planetary spirits to empower talismans purport to be the wisdom of Hermes as conveyed by Aristotle, while the earliest doctrines and practices in alchemy, shoulder-blade divination, and several aspects of astrology (e.g., lunar mansions, Egyptian decans, and astrological lots) are attributed to him, sometimes in the company of Apollonius of Tyana (first century CE), and the legendary Agathodaimon, Asclepius, and Toz Graecus. These technical works were translated into Latin, Greek, and Hebrew and formed the basis of medieval Hermeticism.

See also Aristotle; Islamic Philosophy; Jewish Philosophy; Medieval Philosophy.

Bibliography

Burnett, Charles. "The Establishment of Medieval Hermeticism." In *The Medieval World*, edited by Peter Linehan and Janet L. Nelson, 111–130. New York: Routledge, 2001.

Copenhaver, Brian P. *Hermetica: The Greek* Corpus Hermeticum *and the Latin* Asclepius *in a New English Translation*. Cambridge, U.K.: Cambridge University Press, 1992.

Lucentini, Paolo. *Hermes Latinus*. 3 vols., ongoing. Turnhout, Belgium: Brepols, 1994–2001.

Lucentini, Paolo, ed. *Hermetism from Late Antiquity to Humanism: La tradizione ermetica dal mondo tardo-antico all'umanesimo*. Turnhout, Belgium: Brepols, 2003.

Charles Burnett (2005)

HERSCHEL, JOHN

(1792–1871)

John Herschel, the son of the astronomer William Herschel, worked in mathematics, chemistry, optics, and solid-state physics; pioneered in photography and instigated the regular use of photography in astronomy; invented blueprints; initiated simultaneous worldwide meteorological observations; introduced the theory of isostasy in geology; and was the world's leading observer-theoretician of double stars and nebulae. His *Treatise on Astronomy* (1833), continued as *Outlines of Astronomy* (1849), although deliberately common sense in treatment, was authoritative in content even for professionals until the 1860s. He was England's most famous scientist from 1830 to about 1860.

Herschel's *Preliminary Discourse on the Study of Natural Philosophy* (1831) was a starting point for his philosophic contemporaries, the more radical (post-Kantian) William Whewell and the more conservative (Humean) J. S. Mill; in fact, many errors were deleted from Mill's *Logic* in its second edition because Herschel supplied detailed criticisms of the scientific passages in the first edition. Herschel's full position was expressed later, in papers collected as *Essays from the Edinburgh and Quarterly Reviews* (1857) and *Familiar Lectures on Scientific Subjects* (1867) and in remarks in his scientific books. His best-known philosophic followers were William Stanley Jevons and James Clerk Maxwell.

In theory of knowledge, Herschel's basic concept was the law of continuity, which for him defined the rationality of a system. In his version of the law, he asserted that scientists *observe* not continuous phenomena (not even simple extension), but "dotted outlines which the mind … fills up." Thus "*we assume continuity where we find none.*" Herschel refused any philosophic solution of this disparity between observation and thought and accepted the harmony of mind with external nature as an ultimate fact, preestablished by God.

Next, he was a "decided disciple of old Boscovich"; matter is "a collection of mathematical points—mere *localization* of *forces*"; and therefore it is foolish to picture kinetic-molecular processes as "the 'clashing together' of 'atoms'" or as the "knocking about of billiard balls." Force as hitherto understood, he pointed out, was always associated with matter, that is, inertia; but in electricity and in the "quasi-undulatory propagation of qualities" we see noninertial agents. So the kind of force presented in theories of mechanics is not primary. More basic physical powers exist, he asserted, but are not yet (1840) understood.

Science should uncover not only laws (formal relations among parameters) but also causes. Causation is not Humean succession but (as in the Scottish common-sense school) is known from our consciousness of effort when we exert force. Causes are *not* will, however, but the physical intermediaries between will and muscular contraction. These may also exist in connection with inanimate bodies.

This general position is well beyond Roger Joseph Boscovich and Thomas Reid but is not idealistic. It points toward the theory of the conservation of energy, which, however, Herschel did not approve in its 1860 form. He felt that "potential energy" was not a physical reality, but a mere mathematical expression introduced into the theory "*to save the truth of its verbal enunciation.*"

In methodology Herschel was interested in discovery, not in a justification of the process of induction. (Mill's "methods" were derived directly from Herschel's *Discourse.*) Thus one Herschelian method was "at once to form a bold hypothesis," that is, to guess. Herschel emphasized the central importance of rigorous deduction to confirm hypotheses; it is this which makes science not a craft. One should at all costs avoid specialties of investigation (e.g., chemistry vs. physics), for no actual phenomenon is so divided. Herschel thought that contingency is the most obvious aspect of the universe. Science must grapple with the apparently arbitrary complexities of the actual world, such as sunspot changes, the shapes of nebulae, the variations in terrestrial magnetism, trade winds, and so on, and try to reduce them to scientific laws. It should not content itself with simple general laws concerning force and matter considered in abstraction.

Herschel's contemporary influence was perhaps greatest among working scientists. He gave a reasoned basis for the shift from a purely abstract treatment of physical parameters (as in Joseph-Louis Lagrange) to a belief in the actual existence of the entities used in scientific theories (e.g., the fields of force of his friend Michael Faraday and his admirer Maxwell, which were felt to be actually present in space, not merely mathematical symbols). He upheld the importance of the scientist's feeling for the reality of his constructs. Sheltered by his great authority, scientists pursued their intuitional ideas without worrying about attacks from Humean or other philosophers, or from Evangelical preachers. Herschel, for example, authoritatively established the naturalistic origin of species as a proper subject of investigation for Victorian Englishmen. Young scientists of the period, such as Charles Darwin and Thomas Andrews, admired him extravagantly.

See also Boscovich, Roger Joseph; Causation: Philosophy of Science; Darwin, Charles Robert; Epistemology; Epistemology, History of; Faraday, Michael; Jevons, William Stanley; Maxwell, James Clerk; Mill, John Stuart; Reid, Thomas; Whewell, William.

Bibliography

There are no modern editions of Herschel's works. Recent partial treatments are C. J. Ducasse, "John Herschel's Philosophy of Science," in the American Council of Learned Societies collection, *Studies in the History of Culture* (Menasha, WI, 1942), reprinted in *Theories of Scientific Method,* edited by E. H. Madden (Seattle: University of Washington Press, 1960); W. F. Cannon, "The Impact of Uniformitarianism," *Proceedings of the American Philosophical Society* 105 (1961): 301–314; and W. F. Cannon, "John Herschel and the Idea of Science," *Journal of the History of Ideas* 22 (1961): 215–239, containing an error on "will" that has been corrected in the present entry.

Walter F. Cannon (1967)

HERTZ, HEINRICH RUDOLF

(1857–1894)

The German physicist and philosopher of science Heinrich Rudolf Hertz was born in Hamburg. Early in his student days he showed an interest in engineering but soon took up the study of physics, to which he quickly made important contributions, mainly in the study of magnetism and electricity. He studied in Berlin under Hermann von Helmholtz and Gustav Kirchhoff and inherited their interest in the philosophy of science. In 1883 he began teaching in Kiel, where he worked on James Clerk Maxwell's electromagnetic theory. He became professor of physics at the technological institute at Karlsruhe in 1885, where he produced his most celebrated work on electromagnetic waves. In 1889 he was appointed professor of physics to succeed Clausius at Bonn. He was in failing health, however, and he died five years later.

Hertz's most important book for philosophy is his *Principles of Mechanics,* written during his last illness and published in 1894. This is an attempt to rewrite classical mechanics in such a way as to exhibit its systematic nature, increasing its rigor, reducing its assumptions to a minimum, and keeping it as empirical and nonmetaphysical as possible. His aims were firmly in the spirit of his teachers and of Ernst Mach, who expressed his admiration for Hertz's work. The preface to the *Principles of Mechanics* is a classic in the philosophy of science and deserves to be better known.

Hertz was prepared to admit that various logical categories of statement figure necessarily in the sciences; he even thought, unfashionably, that metaphysical statements could be of considerable value to the scientist. But, he held, it is of the utmost importance for anyone who

would understand the methods of the sciences, and for the scientist himself, to distinguish clearly the different categories of statement and not to suppose that, for example, a nonempirical statement is empirical. In his reconstruction of mechanics he wanted especially to ensure, among other things, that such distinctions are made.

Rather in the manner of Immanuel Kant, who greatly influenced the philosopher-scientists of this period, he begins by dividing mechanics into two parts, one depending upon the formal necessities of our thought and the other depending upon our experience. Moreover, as Jules Henri Poincaré was to argue later, certain features of mechanics depend upon our arbitrary choice. The structure of scientific theories in general exhibits these features, and understanding this structure involves disentangling them.

Further, very much in the modern manner, Hertz holds that a scientific theory is a deductive system that, according to whether it is correct or incorrect, corresponds or fails to correspond to the observable world. *The Principles of Mechanics* shows how one such theory can be set out as an axiom system in which we may deduce conclusions that are testable against reports of our observations.

However, Hertz's aim in this was not merely theoretical and academic, for he seems to have thought that the progress of science might be impeded if scientists do not fully and clearly understand the logic of the concepts they use. He holds, and regards it as generally held by scientists, that the laws of mechanics are fundamental in the solution of all problems in physics; yet there are concepts used in mechanics that are by no means clear and upon which physicists do not even agree. The outstanding example of such a concept—and here Hertz agrees with Mach—is "force." In fact, Hertz fiercely criticizes physicists for relying on this concept without having any very clear notion of what it entails.

The way to understand the concept of force is to see how it functions in the theories in which it is used. But when we look at classical mechanics, we find that force is not used in the way physicists think it is; the usual method of expounding mechanics obscures this and, in general, obscures the very nature of the concept of force. The understanding of scientific concepts is inextricably bound up with the understanding of the theories in which they figure.

Hertz's approach to mechanics is largely determined by his views about explanation in general; mechanics explains the motions of bodies by bringing them under laws, but these laws cannot be in terms only of what is directly observable. Hertz seemingly holds that it has been found that this is so, although he also shows signs of thinking that it must be so—that it is a necessity of explanation. At any rate, he points to many of the explanations accepted in the sciences and shows that they rely upon concealed mechanisms or, as he says, "'confederates' concealed beyond observed masses and motions."

REPRESENTATIONS OF MECHANICS

There are two existing interpretations of mechanics that rely upon force and energy, respectively, as the nonempirical concepts to be used in explanation. Here a further presupposition of Hertz's enters: We will understand our explanations best if all the concepts we use are as similar as possible to concepts of what we experience, that is, to empirical concepts. Force and energy are quite unlike anything we experience, so Hertz seeks to replace them, in his rewriting of mechanics, by motion and mass, which are exactly like observed motion and mass except that they are unobserved (concealed). Or, rather, force and energy are given minor and subordinate roles in his mechanics; all the important roles go to mass and motion. This, he believed, fitted in with the physical theorizing of his time. For example, Maxwell gave an account of electromagnetic forces in terms of concealed masses and motions.

Although force and energy are not empirical concepts, space, time, and mass are. Hertz therefore reconstructs mechanics using only space, time, and mass as primitive concepts. This means that they are not defined in any verbal or symbolic way, although we understand them through our experience of observable masses in motion. Force and energy must not figure in mechanics—as they have tended to do—as basic terms, *as if* they were empirical concepts. They may be introduced at some later stage, but only by defining them—ultimately, at least—in terms of the primitives.

Hertz outlines the two existing "representations" or "images" of mechanics, criticizes them, and then develops his own alternative, which forms the bulk of *The Principles of Mechanics.* He puts forward, as tools of criticism for theories or their representations, three conditions that they must fulfill. They must be *logically permissible* (sometimes abbreviated to *permissible*), that is, consistent with the laws of our thought; they must be *correct,* that is, their structure must not conflict with the structure of observable things; they must be *appropriate,* that is, they must be simple in the sense of containing the fewest pos-

sible superfluous or empty relations. Appropriateness is merely relative: we should, where there are alternatives, accept the more appropriate rather than the less appropriate. These requirements concern the three features of theories mentioned above—one depending upon the nature of our minds, one depending upon our external experiences, and one depending upon our conventional systems of notation.

The first representation Hertz considers is the one then current in most textbooks, taking space, time, force, and mass as its fundamental concepts. It is, among other things, too much influenced by the historical development of mechanics, the order of which may have little to do with its logical structure. It takes force as an independent concept and regards force as the cause of motion. However, the weakness of this representation is that the idea of force is not clear. This affects both the permissibility and the appropriateness of this version: Because our notion of force is vague, it cannot help us to reason precisely, and because we associate with it certain nonessential anthropomorphic ideas, it imports superfluous elements into mechanics. This latter point also seems, for Hertz, to include the idea that too much which is not directly perceptible is thus brought into mechanics. He looks askance at forces that "cancel out in the calculations" as robbing an explanation of its simplicity, or what Mach calls its economy. Apart from this, the first representation satisfies the condition of correctness; if we are merely considering alternative ways of expressing mechanics, we should indeed expect something to be satisfactory in each. What is satisfactory here is that the structure of this way corresponds to, or at least does not conflict with, the structure of observable phenomena.

The second representation is one that was favored in Hertz's day by the more advanced physicists, including Helmholtz. This representation attempts to sidestep the difficulties involved in the concept of force by taking as fundamental the concepts of space, time, mass, and energy. It is then possible to introduce force by definition and merely as an aid to calculation. The advantage of energy over force, it was claimed, was that energy depends upon direct experience because it depends only upon positions or velocities, both of which are directly experienced. This ensures that the second representation is more appropriate than the first. If we consider only motions that occur in nature, Hertz argues, it lacks nothing in correctness. Its weakness lies in its permissibility, as is seen when we try to define energy, as it is here used, in terms of "simple, direct experiences." A substantial view of energy tended to be associated with this representation, but it is difficult to treat potential energy as a substance, especially when, as is sometimes necessary, we must ascribe negative potential energy to a system or must regard the potential energy of a finite quantity of matter as infinite. This version is superior to the first, but it still contains serious difficulties.

HERTZ'S REPRESENTATION

Since force and energy, respectively, appear to be responsible for the problems arising over these two representations. Hertz attempts to do without them, at least as primitive concepts for his representation. He begins with space, time, and mass. That is, he begins with kinematics, the abstract study of motion, and sets out to derive the whole of mechanics from it without using force and energy except as devices for calculation. Kirchhoff had already asserted that three independent concepts are necessary and sufficient for mechanics.

Time, space, and mass are primitive terms for Hertz's system, but they are not mere abstract counters like the uninterpreted symbols of the logicians. They are understood through experience, and the particular experiences that are to count for the purposes of mechanics can be specified. Moreover, these concepts are, as we also discover in experience, permanently related in various ways. Hertz puts forward a "Fundamental Law" that has similarities to the law of inertia and that summarizes the connection between the three basic concepts taken together: "Every natural motion of an independent material system consists herein, that the system follows with uniform velocity one of its straightest paths." This law, together with the concepts of space, time, and mass and the hypothesis of concealed masses, allows us, by purely deductive reasoning, to derive the whole of mechanics and so to explain mechanical phenomena.

Other concepts, such as force and energy, are introduced into the system later by definition and so are regarded merely as aids to deduction. They are defined ultimately in terms of the primitive concepts.

The Principles of Mechanics is divided into two parts to emphasize the independence of the mathematical form and the physical content of mechanics. The equation "2 horses + 2 horses = 4 horses" has a mathematical form, expressed by "2 + 2 = 4," that is independent of its application to horses. In the same way, mechanics as a whole can be regarded as having these two aspects. Book I of the *Principles* draws out the implications of the fundamental ideas: space, time, and mass. At this stage these concepts are intuitive and independent of experience except insofar as all our intuitions and modes of reasoning depend

upon experience. Book II contains the application of these concepts to experience through the Fundamental Law and the derivation of testable assertions about observable phenomena. The apparently equivocal nature of Hertz's basic concepts can best be understood in relation to Kantian philosophy: Our intuitions, peculiarly adapted to fit the general form of what we experience, are analogous to colored spectacles which determine our seeing the world as colored. Nevertheless, the details of our pictures of the world have the nature of hypotheses and are open to empirical testing.

Hertz's account of mechanics is important from the standpoint of the philosophy of science because it represents an early attempt to see a scientific theory as a system and to bring out its logical structure accordingly. It was influential in connection with the conventionalism later championed by Poincaré and attempted to do justice to, on the one hand, the undoubted empirical nature of science and, on the other, the apparent claims of scientific laws to embody natural necessities. Hertz's view that mechanics is the foundation of all physical explanation was the most backward-looking element in his work and was strangely belied by both his scientific work, which was largely influential in breaking down that view, and his work in the philosophy of science, which contained the seeds of a far more flexible view of explanation.

See also Classical Mechanics, Philosophy of; Helmholtz, Hermann Ludwig von; Kant, Immanuel; Mach, Ernst; Maxwell, James Clerk; Philosophy of Science; Poincaré, Jules Henri.

Bibliography

WORKS BY HERTZ

Untersuchungen über die Ausbreitung der elektrischen Kraft. Leipzig: J.A. Barth, 1892. Translated by D. E. Jones as *Electric Waves.* London: Macmillan, 1893. This volume contains Hertz's most important papers on Maxwell's theory and a preface by Kelvin (in the English edition). Hertz approved the translation.

Gesammelte Werke (Collected works). Edited by Philipp Lenard. Leipzig: J.A. Barth, 1894–1895.

Die Prinzipien der Mechanik. Leipzig, 1894. Translated by D. E. Jones and T. E. Walley as *The Principles of Mechanics.* London: Macmillan, 1899. This is the most important source for Hertz's ideas on the philosophy of science.

Schriften Vermischten Inhalts. Leipzig, 1895. Translated by D. E. Jones and G. A. Schott as *Miscellaneous Papers.* London, 1896.

WORKS ON HERTZ

Braithwaite, R. B. *Scientific Explanation.* Cambridge, U.K.: Cambridge University Press, 1953. See especially Chs. 3 and 4.

Brill, Alexander von. "Über die Mechanik von Hertz." *Mathematisch-Naturwissenschaftliche Mitteilungen* 2 (1900): 1–16.

Fitzgerald, G. F. "The Foundations of Dynamics." *Nature* 51 (1895): 283–285.

Foulkes, P. "The Concept of Force." *Australasian Journal of Philosophy* 29 (1951): 175–180 and 30 (1952): 130–132. Reply to Smart.

Hempel, C. G. *Fundamentals of Concept Formation in Empirical Science.* Chicago: University of Chicago Press, 1952. This is Vol. II, No. 7, of *International Encyclopedia of Unified Science.*

Hesse, Mary B. *Forces and Fields.* London: Nelson, 1961. See especially pp. 212ff.

Jammer, Max. *Concepts of Force.* Cambridge, MA: Harvard University Press, 1957.

Lindsay, R. B., and Henry Margenau. *Foundations of Physics.* New York: Wiley, 1936. See especially Ch. 3.

Mach, Ernst. *Die Mechanik in ihrer Entwickelung historischekritisch dargestellt.* Leipzig: Brockhaus, 1883. Translated by T. J. McCormack as *The Science of Mechanics.* 4th ed., Chicago: Open Court, 1919; taken from the fourth German edition of 1901.

Maxwell, J. C. "Ether." In *Encyclopaedia Britannica,* 9th ed. Edinburgh, 1875–1889. Vol. VIII, pp. 568–572.

Poincaré, Henri. "Les idées de Hertz sur la mécanique." In *Oeuvres de Henri Poincaré,* Vol. VII, pp. 231–250. Paris: Gauthier-Villars, 1916–1956. This has been partially translated by A. Miller in *Philosophy of Science,* edited by Arthur Danto and Sidney Morgenbesser, 366–373. New York: Meridian, 1960. Both versions give an account of the reasons for dissatisfaction with the concept of force.

Russell, Bertrand. *The Principles of Mathematics.* Cambridge, U.K.: Cambridge University Press, 1903. Ch. 59.

Schilpp, P. A., ed. *Albert Einstein, Philosopher-Scientist.* Evanston, IL: Library of Living Philosophers, 1949. Contains an assessment of both the scientific and the philosophical works of Hertz. See especially Section I and Section II, Parts 2 and 19.

Smart, J. J. C. "The Concept of Force." *Australasian Journal of Philosophy* 30 (1952): 124–130.

Smart, J. J. C. "Heinrich Hertz and the Concept of Force." *Australasian Journal of Philosophy* 29 (1951): 36–54.

Watson, W. H. *On Understanding Physics.* Cambridge, U.K.: Cambridge University Press, 1938.

Peter Alexander (1967)

HERVAEUS NATALIS

(c. 1250–1323)

Hervaeus Natalis, or Harvey Nedellec (c. 1250–1323) was one of the first followers of Thomas Aquinas, but also an original thinker, especially in the areas of intentionality and the mental word. Hervaeus was born in Brittany in the mid-thirteenth century. He entered the Dominicans in 1276 and studied at the University of Paris, where he

commented on the *Sentences* of Peter Lombard, later received the degree of Master of Theology, and served as regent master in theology from 1307 to 1309. He was elected Provincial of France in 1309 and became Master General of the Dominicans in 1318. In the years following the condemnation by Étienne Tempier, the bishop of Paris, of 219 propositions—many of which touched upon the teaching of Thomas Aquinas—Hervaeus defended Aquinas's theological method in his *Defensio doctrinae fratris Thomae* and his theory of knowledge in his *Quodlibeta*. He actively promoted the canonization of Aquinas and died at Narbonne in 1323 on his way to it. Due largely to the work of Hervaeus, Aquinas became the official doctor of the Dominican Order, despite the conservative Augustinian atmosphere in the period after 1277.

Although he strongly promoted Aquinas's thought, Hervaeus did not follow Thomas on some of his most distinctive teachings, such as the real distinction between essence and existence in creatures and the five ways of proving the existence of God. Of the latter, Hervaeus retains only the ways of efficient causality and of degrees of perfection. It is noteworthy that in the conservative theological atmosphere following 1277 Hervaeus develops strictly philosophical proofs for the existence of God in his *De cognitione primi principii*. In his conflicts with Durandus of Saint Pourçain, a Dominican who leaned toward a more Augustinian position, Hervaeus upheld Thomism, but a Thomism that manifests the influence of Duns Scotus's thought. Hervaeus's still unpublished *Tractatus de secundis intentionibus* is the first treatise in the Middle Ages devoted to the topic of intentionality.

See also Durandus of Saint-Pourçain; Thomas Aquinas, St.

Bibliography

PRIMARY WORKS

In quattuor libros sententiarum commentaria and *De potestate papae*. Farnborough, U.K.: Gregg, 1966.

The Poverty of Christ and the Apostles. Translated by J. D. Jones. Toronto: Pontifical Institute of Mediaeval Studies, 1999.

Quolibeta with *De beatitudine*, *De verbo*, *De eternitate mundi*, *De materia*, *De relationibus*, *De unitate formarum*, and *De virtutibus*. Ridgewood, NJ: Gregg, 1966.

SECONDARY WORKS

Conforti, P. "Hervé de Nédellec et les questions ordinaires *De cognitione primi principii*." *Revue thomiste* 97 (1997): 63–82.

Iribarren, Isabel. "The Scotist Background in Hervaeus Natalis's Interpretation of Thomism." *The Thomist* 66 (2002): 607–627.

Lowe, Elizabeth. *The Contested Theological Authority of Thomas Aquinas: The Controversies between Hervaeus Natalis and Durandus of St. Pourçain*. New York: Routledge, 2003.

Mannath, Joseph T. "Harvey of Nedellec's Proofs for the Existence of God: 'De Cognitione Primi Principii, Qq. III-IV.'" *Salesianum* 31 (1969): 46–112.

Perler, Dominik. "Teil IV: Das Modell der intentionalen Präsenz: Petrus Aureoli und Hervaeus Natalis." In *Theorien der Intentionalität im Mittelalter*, pp. 253–317. Frankfurt am Main, Germany: Klostermann, 2002.

Trottmann, C. "Verbe mentale et noétique thomiste dans le *De verbo* d'Hervé de Nédellec." *Revue thomiste* 97 (1997): 47–62.

Roland J. Teske (2005)

HERZEN, ALEKSANDR IVANOVICH
(1812–1870)

Alexander Ivanovich Herzen, the Russian editor, essayist, and social philosopher, was the illegitimate son of I. A. Iakovlev. Herzen was graduated from the faculty of physics and mathematics of Moscow University in 1834 and was promptly exiled to the provinces for radicalism (1835–1840, 1841–1843). He emigrated from Russia in 1847 and spent the remainder of his life in western Europe. In London he founded the first "free Russian journal," *Kolokol* (The Bell), in 1852. There, during the 1850s he published, in eight parts, his *Byloe i dumy* (*My Past and Thoughts*), a brilliant personal memoir and history of nineteenth-century ideas.

Herzen's first essay in philosophy, "Diletantizm v Nauke" (Dilettantism in science), was published as a series of four articles in *Otechestvennye zapiski* (Notes of the fatherland) in 1843. Herzen used the term *science* (*nauka*) in the broad sense of G. W. F. Hegel's *Wissenschaft* and focused his critique on four kinds of scientific dilettantes: naive dabblers; romanticizers of the past without interest in the problems of today; pedantic specialists who in their ivory towers write erudite books about erudite books; and the "Buddhists of science," or right Hegelians, who offer a purely speculative account of historical reality and make no effort to change it. Herzen opposed to all four kinds of dilettantism a "committed" philosophy of the act (*filosofiia dela*) that seeks to reconcile abstract speculation with vital human needs. (Later, Herzen spoke of Hegel's dialectic as an "algebra of revolution.")

In "Pis'ma ob izuchenii prirody" (Letters on the study of nature), published as a series of eight articles in

Otechestvennye Zapiski in 1845–1846, Herzen attempted to reconcile the opposed interests of natural science, which tends toward empiricism, and philosophy, which tends toward idealism. But empiricism and idealism are Hegelian "moments," incomplete and one-sided dialectical phases, each requiring the other. In the end Herzen stressed the rights of empiricism as closer than those of idealism to the real needs of living individuals.

In works written after 1847 Herzen outlined his defense of the existing individual against the collective encroachments of society, history, and progress. Nonhuman individuals are constantly sacrificed to supraindividual ends, as in the slow formation of a coral reef from the skeletons of millions of tiny sea creatures. "The polyps die," Herzen wrote, "not suspecting that they have served the progress of the reef" (*Byloe i dumy*, in *Sobranie sochinenii* [Collected works], Vol. X, p. 123). But men are not polyps; human individuals should not be sacrificed to build any coral reef of historical progress, however grandiose.

Herzen developed, in *S togo berega* (*From the Other Shore*; Paris, 1850) and later works, a philosophy of contingency, emphasizing the "tousled improvisation" of history. Historical development does not exhibit the rational, purposive structure that Hegelians see in it; therefore, men are free to impose their own purposes on its "whirlwind of chances." "Outside us everything changes, everything vacillates. We are standing on the edge of a precipice and we see it crumbling. . . . We shall find no haven but in ourselves, in the consciousness of our unlimited freedom, of our autocratic independence" (*From the Other Shore*, p. 128).

Herzen stressed the lived sense of freedom and attempted to reconcile this psychological fact with a deterministic theory. In a letter (written in French in 1868) to his son, who had defended a fashionable physiological reductionism, he called the idea of freedom a "phenomenological necessity," adding, "the conscious self cannot move or act without positing itself as free, that is to say as to having within certain limits the power to do or not to do. Without this belief, individuality dissolves and is lost" (*Sobranie sochinenii*, Vol. XX, pt. 1, p. 436).

Anticipating Fëdor Dostoevsky's "Legend of the Grand Inquisitor" (1880), Herzen in 1847 stated the theme of escape from freedom and the burden of moral responsibility. The love of moral freedom, in Herzen's words, is "purely Platonic and ideal," whereas the love of intellectual authority is a solid *mariage de raison* in which "dreams and poetry are sacrificed for domestic comfort [and] order" (ibid., Vol. II, p. 90).

Herzen formulated not only an extreme moral relativism—"What was admirable behavior yesterday may be abominable today" (*From the Other Shore*, p. 141)—but also an embryonic emotivism in ethics. Moral judgments are expressions of taste or preference on the model of "I like lobster"; there is no point in my arguing with someone who does not like lobster. According to Herzen, "there are no general rules, but [only] an improvisation of conduct, … a tact, an aesthetics of human actions" (*Sobranie sochinenii*, Vol. XXIX, pt. 1, p. 148). On this, as on other points, Herzen anticipated Friedrich Nietzsche—who may have read some of Herzen's works in translation.

See also Dostoevsky, Fyodor Mikhailovich; Empiricism; Hegel, Georg Wilhelm Friedrich; Idealism; Nietzsche, Friedrich; Russian Philosophy.

Bibliography

WORKS BY HERZEN

My Past and Thoughts. 6 vols. Translated by Constance Garnett. New York: Knopf, 1924–1927.

Sobranie sochinenii v tridtsati tomakh (Collected works in thirty volumes). Moscow, 1954–1964.

From the Other Shore and The Russian People and Socialism. Edited by Isaiah Berlin. New York: Braziller, 1956. Introduction by the editor.

Selected Philosophical Works. Translated by L. Navrozov. Moscow: Foreign Languages Publishing House, 1956.

WORKS ON HERZEN

Malia, Martin. *Alexander Herzen and the Birth of Russian Socialism: 1812–1855*. Cambridge, MA: Harvard University Press, 1961.

Zenkovsky, V. V. *Istoriia Russkoi Filosofii*. 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*, 2 vols., New York: Columbia University Press, 1953. See pp. 271–298.

George L. Kline (1967)

HESS, MOSES

(1812–1877)

Moses Hess, the socialist journalist and organizer and intellectual precursor of Zionism, often called the father of German socialism, was born in Bonn of Jewish parents. A left-Hegelian, he was a mentor and coworker of Friedrich Engels, Karl Marx, and Ferdinand Lassalle. He led radical workers' groups in Paris and Belgium, edited the famous *Rheinische Zeitung,* and was the leader of the "true," or "philosophical," German socialists of the 1840s. Later he became Lassalle's chief organizer in the

Rhineland and a foreign correspondent for European and American newspapers. His published books and countless essays include works on the philosophy of history and on socialism, a famous call for a Jewish state, and a comprehensive theory of the laws of science, society, and socialism.

Hess used the principles of Benedict de Spinoza, G. W. F. Hegel, and Johann Gottlieb Fichte to demonstrate the inevitability and justice of a society lacking distinctions of class and wealth, without "contradictions" between private passion and public law, and without external compulsion. Hess took this to be both the social expression of pantheism and the inevitable result of the dialectical development of the self-realization of the Absolute Spirit in history. This was the theme of his early work, *Die Heilige Geschichte der Menschheit* (Stuttgart, 1837). Later, under the influence of Ludwig Feuerbach, Hess rejected Hegelian transcendentalism. He then created a "philosophy of the deed," based on a belief in the human spirit as the unconditioned ultimate reality. He stressed the creative power of man and man's historical "alienation" of that power to various mythical transcendent powers—God, the state, fate, or, in Hess's day, the laws of history and economics. Hess insisted that there are no objective limits to man's power to create a society free from exploitation and compulsion.

Marx and Engels attacked this kind of moralistic and philosophical socialism (and later Hess himself) as inadequate to the harsh realities of economic determinism and the class struggle. The influence of Marx and the failure of romantic idealism in the widespread revolutions of 1848 helped to convert Hess from idealism to materialism. He now spoke of ideas as the "reflex" of material conditions and the class struggle, and he predicted the inevitable termination of the economic "contradictions of capitalism" in overproduction, proletarian misery, depression, revolution, and finally, socialism.

In the end, however, Hess became pragmatic. He rejected dialectical materialism as he had rejected dialectical idealism. He worked with Lassalle to found German social democracy, and like Lassalle he hoped for radical social reform through universal suffrage and the nationalization of the means of production. And it was in this spirit, not Marx's, that the German Social Democratic Party started its career. In the 1860s, fearful of the future of the Jews of Europe, Hess worked for a Jewish and socialist state in Palestine.

See also Dialectical Materialism; Engels, Friedrich; Feuerbach, Ludwig Andreas; Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Lassalle, Ferdinand; Marx, Karl; Socialism; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY HESS

Die europäische Triarchie. Leipzig, 1841.

Jüdische Schriften. Edited by Theodor Zlocisti. Berlin: Lamm, 1905.

Rom und Jerusalem. Leipzig, 1862. Translated by M. Waxman as *Rome and Jerusalem,* New York: Bloch, 1918.

Sozialistische Aufsätze, 1841–1847. Edited by Theodor Zlocisti. Berlin: Welt-Verlag, 1920.

Moses Hess Briefwechsel. Edited by Edmund Silberner. The Hague: Mouton, 1959.

The Revival of Israel: Rome and Jerusalem, the Last Nationalist Question. Translated by M. Waxman. Lincoln: University of Nebraska Press, 1995.

WORKS ON HESS

Avineri, Shlomo. *Moses Hess: Prophet of Communism and Zionism.* New York: New York University Press, 1985.

Berlin, Isaiah. *The Life and Opinions of Moses Hess.* Cambridge, U.K.: Heffer, 1959.

Koltun-Fromm, Ken. *Moses Hess and Modern Jewish Identity.* Bloomington: Indiana University Press, 2001.

Lukács, Georg. "Moses Hess und die Probleme der idealistischen Dialectik." *Archiv für die Geschichte des Sozialismus und der Arbeiterwegung* 12 (1926): 105–155.

Weiss, John. *Moses Hess: Utopian Socialist.* Detroit: Wayne State University Press, 1960.

Zlocisti, Theodor. *Moses Hess, der Vorkämpfer des Sozialismus und Zionismus.* 2nd ed. Berlin: Welt-Verlag, 1921.

John Weiss (1967)
Bibliography updated by Philip Reed (2005)

HETEROSEXISM

Heterosexism may generally be understood as an attitude in which heterosexual relationships, social arrangements, and sexual activities are viewed as morally, culturally, religiously, biologically, and/or psychologically ideal, and are thus superior to and rightly privileged over any non-heterosexual option. Another less common usage describes heterosexism as an attitude in which the separation of sex, anatomy, gender, and gender roles into two discrete categories of male and female is assumed to be natural and required for coherent personal identity and social stability, and is influential in analyzing gendered social roles and identities (Butler 1990) and issues of transsexualism and transgendered identity.

In this first and predominant use, however, *heterosexism* is intended to parallel the concepts of sexism and racism and points toward characteristic prejudice and

discrimination against sexual minorities, mainly by heterosexuals, but also by self-disapproving homosexuals and bisexuals who have internalized a heterosexist attitude. The ways in which people who evince homosexual desire and behavior are discriminated against include legal inequity (military service prohibition, gay marriage and adoption prohibitions, sodomy laws, career restrictions), social treatment (housing discrimination, job discrimination, public denouncements), and cultural treatment (community invisibility, moral condemnation, stereotyping, greater risk of physical and verbal assault, pressure to stay "closeted," religious condemnation).

While these inequities are examples of discrimination, and in some cases obvious mistreatment, discrimination is not in and of itself unjust. Just discrimination occurs when some property of a person is contextually morally relevant to the decision to treat them differently (e.g., not allowing a blind person to drive), whereas unjust discrimination occurs when some property of a person is not contextually morally relevant to the decision to treat them differently (e.g., not allowing women to vote). The ongoing moral debate then, is about whether some heterosexist beliefs and consequent discriminatory practices are morally justified.

DISCRIMINATION AS MORALLY JUSTIFIED

Heterosexist laws and policies are typically held to be just based on the prior assumption that homosexual sex is immoral and/or that homosexual desire is defective. While heterosexist policies do not automatically result from heterosexist attitudes (one could be libertarian), heterosexist attitudes are a prerequisite for such policies. Though variations exist, there are three main classes of argument for the view that homosexuality is immoral or inferior.

Divine command theory arguments state that morality is determined by the edict of God and that those texts thought to be authoritative indicators of God's will have outlawed homosexual sex. In the historically dominant Judeo-Christian tradition of European and North American culture, the proof texts most often cited are the Hebrew Bible passages concerning creation (Gen. 1–2), Sodom and Gomorrah (Gen. 19), Gibeah (Judg. 19), the holiness code (Lev. 18:22 and 20:13), and three Christian New Testament passages (Rom. 1:26–27; 1 Cor. 6:9; 1 Tim. 1–10).

Natural law arguments state that there is an objective moral good and a path toward that good for all creatures depending on the kind of creature that they are. The kind of creatures that they are is typically understood (though not always—e.g., in Aristotle) as the result of God's eternal design, and so natural law is often closely connected to divine command theory. By attending to the visible organization of nature—including our bodies—we can determine what our proper functions and ends are and how we might best achieve those ends. Any actions that do not work toward achieving those aims or that actively block the fulfillment of those aims are immoral or at least not tending toward the moral. The obvious function of sex and sexual anatomy (it is claimed) is reproduction and family unity. Homosexual behavior interferes with these aims (as does other nonprocreative sex) and is therefore objectively unnatural and intrinsically immoral. Homosexual desire is not a behavior and thus not immoral, but by tending toward interference with the proper ends of our bodies is an objectively disordered state.

This approach to the moral condemnation of homosexuality is historically influential and frequent. The argument is present in Plato, who while praising homosexual love as a first step toward the realization of beauty and truth in the Symposium (209a–211c), nonetheless rejects homosexual sex as absent in animals and unnatural in Laws (836c–836e). Aristotle says very little on homosexuality but his emphasis in the Physics on explaining things by reference to their biological purposes is well-matched for the natural law argument that nonprocreative sex is essentially misdirected. He includes pederasty (the Greek tradition of a young man engaging in sex with an older male mentor) in a list of diseased states learned mostly through social custom, including chewing fingernails and eating coal (Ethics 1148a15–1148b30).

The true cultural power of the "against nature" argument, however, flourishes with the development of medieval Christian moral doctrine in St. Thomas Aquinas. Heavily influenced by Aristotle, he asks what the natural end of sex is and answers that it is procreation. As such, all sexual activity that has no chance of resulting in procreation is "contrary to right reason," contrary to the "natural order," and a lustful vice, a category which includes masturbation, bestiality, homosexuality, and any sort of sex even between men and women that occurs "by not observing the natural manner of copulation" (Aquinas, 1920 2.2.154.11). This view is currently reflected in the Catechism of the Catholic Church, which states that: "homosexual acts are intrinsically disordered" and "contrary to the natural law" (CCC n2357). The natural law approach has been influential in legal matters,

including Blackstone's famous and formative Commentaries on the Laws of England, which the U.S. Supreme Court appealed to in *Bowers v. Hardwick* (1986) in upholding the legality of sodomy laws. Justice Burger wrote: "Blackstone described 'the infamous crime against nature' as an offense of 'deeper malignity' than rape, an heinous act 'the very mention of which is a disgrace to human nature,' and 'a crime not fit to be named'" (*Bowers v. Hardwick*).

Not all natural law arguments say homosexual acts must always be immoral, but leave room open for the position that such sexual expression can be genuine goods for the homosexuals involved, even though same-sex relationships do not achieve the natural ideal. The analogy here is often to adoption. Ideally, parents raise and love their own biological children, but adoption may still be a genuine, lesser, good for the actual persons involved in actual situations (Cahill 1995). This view does not treat homosexuality as a neutral variation in human nature, like left-handedness, but rather as a deficit or disadvantage. Finally, there are secular variations of a natural law-style argument against homosexuality that avoid religion or concepts of intrinsic morality but retain the disease or disadvantage model. Sigmund Freud views a homosexual orientation as a stalled development resulting in sexual narcissism (Ruse 1988). Taking an evolutionary and Aristotelian approach, Michael Levin (1984) views homosexuality as abnormal and homosexual sex as a misuse of sexual organs, likely to result in endemic unhappiness, making it prudentially, but not intrinsically bad. As such, it is legitimate for a society to formally discourage and refuse to legitimize homosexuality—a position which crosses over into the social harm category of argument.

Social harm arguments state that homosexual activity damages individuals and society. As such, the state may legitimately discourage or criminalize homosexuality as a measure of public self-protection. While proponents of social harm arguments often begin with divine command or natural law convictions, the arguments themselves, as empirically based, stand independently. Many social harm arguments take as a starting point certain empirical claims about homosexual psychology, including that homosexuals are pedophiles (sometimes conflating the categories altogether), are unstable, are promiscuous, are disease-ridden, are depressed, and actively recruit others (particularly children) into their lifestyle.

Some arguments are more sociological than psychological, arguing that altering the historical social norms that have governed gender and family structures, such as permitting gay marriage or adoption, will confuse our notions of families, will traumatize children's understanding of parental roles, and will erode heterosexual families. For example, John Finnis argues that because homosexual relationships are immoral, incapable of actualizing mutual devotion, and harm the personalities of its participants, they are "deeply hostile to the self-understanding of those members of the community who are willing to commit themselves to real marriage" and are "an active threat to the stability of existing and future marriages" (1994, p. 515). As such, the political community has a "compelling interest" in discouraging and criminalizing homosexual sex. Finally, there is the slippery slope argument that once an institution as sacrosanct as heterosexual family and marriage is fundamentally altered, there will be no principled way to prevent the legalization of polygamy, bestiality, pedophilia, or incest.

DISCRIMINATION AS MORALLY UNJUSTIFIED

Defenders of the morality of homosexuality begin with the claim that individuals have a prima facie interest in pursuing their own goods as they interpret them and that such pursuit should not be denied without just reason. The first approach then to claiming that heterosexist beliefs and policies are unjust is to contend that all the arguments that defend the permissibility of heterosexism are unsound.

Against the divine command argument, frequent intrareligious criticism says that scriptural passages are often mistranslated (there being no term in either Hebrew or Greek for constitutional homosexuality as we know it), are misread out of context (the holiness code being instructions on ritual practices rather intrinsic morality or referring only to rape or prostitution), that the choice to enforce sexual regulations over other equally weighted ritual regulations are examples of prejudicial selectivity, or that biblical texts are not infallibly authoritative. The frequent extrareligious response is that divine command theory is false in the first place and thus such texts have no special moral authority, or at least that in a free society, religious beliefs are not state-enforced and thus have no special legal authority.

Against the natural law argument, critics argue that organs and sex may have more than one end (Mohr 1988), that *natural* and *unnatural* do not equate with *good* and *bad*, that there is no logical implication from the ways things are to the ways things ought to be (the is/ought distinction), that animals are no guides to our

behavior (Boswell 1980), that homosexuality is in fact natural and does occur among animals (Bagemihl 1999; Roughgarden 2004), and that homosexuality may be evolutionarily adaptive and antihomosexual discrimination may explain any unhappiness better than sociobiology (Murphy 1987).

Against the social harm argument, critics point out that numerous psychological studies have shown most of the heterosexist claims to be false—homosexuals are not more likely than heterosexuals to be pedophiles, they do not seek to convert heterosexuals, sexually transmitted diseases are as likely to occur in heterosexuals who engage in unsafe sex, and that any promiscuity is both exaggerated and as or more likely to be the result of social marginalization as an intrinsically homosexual trait. Slippery slope arguments, in addition, falsely assume that we will be unable in the future to distinguish the consensual, harmless sexual activity between rational adults (in both heterosexual and homosexual relationships) from the sorts of harmful abuses present in bestiality, incest, and pedophilia (Corvino 2005).

For all these reasons, defenders of the morality of homosexuality claim that proponents of heterosexism fail to meet the burden of proof that would justify heterosexist attitudes and policies. In a positive vein, they also argue that homosexuality is simply a normal, if minority, variation among human traits with widespread historical and cross-cultural representation, that homosexual relationships can actualize mutual devotion and other human goods in distinctive ways, and that citizens' interest in liberty and pursuit of their own good outweighs any state interest in promoting only procreative relationships. This last argument weighed heavily in *Lawrence v. Texas*, the 2003 U.S. Supreme Court decision overruling the tradition- and natural law-defense of *Bowers v. Hardwick*. The court stated: "The fact that a State's governing majority has traditionally viewed a particular practice as immoral is not a sufficient reason for upholding a law prohibiting the practice, and … individual decisions concerning the intimacies of physical relationships, even when not intended to produce offspring, are a form of 'liberty' protected by due process."

RELATION TO HOMOPHOBIA

Heterosexism is a relatively new term with a contested relationship to *homophobia*, a term popularized by George Weinberg (1972). The latter concept came to be used widely, and often interchangeably, with *heterosexism* to refer to any prejudice or morally negative attitude toward homosexuality, often with highly polemical intent. However, *homophobia* originated as a quasi-psychiatric term that emphasized the irrational hatred of homosexuals (Pharr 1988). This usage has been criticized as failing to adhere to the clinical symptomatology of true phobias (Richmond 1998) and for individualizing a problem that is better understood as a political and institutional phenomenon (Kitzinger 1989). It has been argued that homophobia is a phenomenon best understood as an extreme projection of background heterosexist attitudes that are themselves strongly shaped by background assumptions concerning gender identity (Hopkins 1996). In any case, *heterosexism* appears to have become the more dispassionate term of choice for describing an attitude that privileges heterosexuality.

See also Aristotle; Feminism and Continental Philosophy; Feminism and the History of Philosophy; Feminist Philosophy; Feminist Social and Political Philosophy; Philosophy of Sex; Racism; Sexism; Thomas Aquinas, St.

Bibliography

Aquinas, St. Thomas. *Summa Theologica*. 2nd ed. Revised and translated by the Fathers of the English Dominican Province, 1920. Available from http://www.newadvent.org/summa/.

Aristotle. *The Nicomachean Ethics*. Translated by David Ross. Oxford: Oxford University Press, 1998.

Bagemihl, Bruce. *Biological Exuberance: Animal Homosexuality and Natural Diversity*. New York: St. Martin's, 1999.

Boswell, John. *Christianity, Social Tolerance, and Homosexuality: Gay People in Western Europe from the Beginning of the Christian Era to the Fourteenth Century*. Chicago: University of Chicago Press, 1980.

Bowers v. Hardwick, 478 U.S. 186 (1986).

Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.

Cahill, Lisa Sowle. "Book Review of 'Heterosexism: An Ethical Challenge.'" *Theological Studies* 56 (1) (March 1995): 187–188.

Catechism of the Catholic Church (2357).

Corvino, John. "Homosexuality and the PIB Argument'" *Ethics* 115 (3) (April 2005): 501–534.

Finnis, John M. "Law, Morality, and 'Sexual Orientation.'" *Notre Dame Law Review* 69 (1994): 1049–1076.

Hopkins, Patrick D. "Gender Treachery: Homophobia, Masculinity, and Threatened Identities." In *Rethinking Masculinity: Philosophical Explorations in Light of Feminism*, edited by Larry May, Robert Strikwerda, and Patrick D. Hopkins. Lanham, MD: Rowman and Littlefield, 1996.

Kitzinger, Celia. "Heteropatriarchal Language: The Case Against 'Homophobia.'" *Gossip: A Journal of Lesbian Feminist Ethics* 5 (1989): 15–20.

Lawrence v. Texas, 539 U.S. 558 (2003).

Levin, Michael. "Why Homosexuality Is Abnormal." *Monist* 67 (2) (April 1984): 251–283.

Mohr, Richard D. *Gays/Justice: A Study of Ethics, Society, and Law*. New York: Columbia University Press, 1988.
Murphy, Timothy F. "Homosexuality and Nature: Happiness and the Law at Stake." *Journal of Applied Philosophy* 4 (2) (1987): 195–204.
Pharr, Suzanne. *Homophobia: A Weapon of Sexism*. Inverness, CA: Chardon Press, 1988.
Plato. *Laws*. Translated by Trevor Saunders. In *Plato: Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.
Plato. *Symposium*. Translated by Alexander Nehama and Paul Woodruff. In *Plato: Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.
Richmond, Janice P. "Homophobia: An Evolutionary Analysis of the Concept as Applied to Nursing." *Journal of Advanced Nursing* 28 (2) (1998): 362–369.
Roughgarden, Joan. *Evolution's Rainbow: Diversity, Gender, and Sexuality in Nature and People*. Berkeley: University of California Press, 2004.
Ruse, Michael. *Homosexuality: A Philosophical Inquiry*. Oxford: Blackwell, 1988.
Weinberg, G. *Society and the Healthy Homosexual*. New York: St. Martin's, 1972.

Patrick D. Hopkins (2005)

HEYTESBURY, WILLIAM

(before 1313–1372/3)

William Heytesbury, a fellow at Merton College in Oxford from 1330, belonged to the second generation of Mertonian "Calculators." His work depends on Richard Kilvington's *Sophismata* (1325) and Thomas Bradwardine's *Insolubilia* and *Tractatus de Proportionibus* (1328). His technique was to analyze sophismata—ambiguous problematic statements whose truth or falsehood is to be assessed under specified assumptions—and apply supposition theory, a form of semantic-logical analysis, to the explication of their underlying logical grammar. He is particularly noted for his work on motion and the continuum.

Heytesbury's most popular work was the *Rules for Solving Sophismata* (1335), which contains six treatises: "On Insoluble Sentences (Insolubilia)," dealing with self-referential paradoxes; "On Knowing and Doubting," concerning reference in intensional contexts; "On Relative Terms," considering the reference of relative pronouns; "On Beginning and Ceasing" and "On Maxima and Minima," about continua; and "On the Three Categories," on velocity and acceleration in changes of place, quantity, and quality.

In "On Beginning and Ceasing," Heytesbury considers the sophisma "some part of an object ceases to be seen by Socrates," given that the object is not now, but will, immediately after now, be partly occluded by an object passing in front of it. This statement may assert that there is a given part of the object that will, in every moment after this one, be entirely occluded, and if so, it is false. Or it may assert that at every moment after this present moment, there will be some part of the object entirely occluded at that moment (a different part for each moment), and then it is true.

The *Rules* became popular, and remained important on the European continent even after the Mertonians began to be ignored in Britain. It was taught at Padua and Paris through the early sixteenth century, influencing the Paduan school, fifteenth-century Italian logicians such as Paul of Venice (d. 1429), and the school of John Major at Paris. With the rest of medieval logic, Heytesbury's work sank into obscurity after that. In addition to *Rules* Heytesbury wrote two collections of sophismata, in one of which the (obviously false) statement, "you are a donkey," was repeatedly derived from seemingly harmless admissions. He also wrote some shorter works; for instance, "On the Compounded and Divided Senses," which deals with scope ambiguities similar to that involved in the preceding example.

In the sixth chapter of *Rules*, Heytesbury states the mean-speed theorem for uniformly accelerated motion: A uniformly accelerated body will, over a given period of time, traverse a distance equal to the distance it would traverse if it moved continuously in the same period at its mean velocity (one-half the sum of the initial and final velocities). Elsewhere, he points out, in a particular case, that a uniformly accelerated body will, in the second equal time interval, traverse three times the distance it does in the first. Domingo de Soto observed the applicability of the mean-speed theorem to free fall in 1555.

See also Bradwardine, Thomas; Kilvington, Richard; Medieval Philosophy; Paul of Venice.

Bibliography

WORKS BY WILLIAM HEYTESBURY

Hentisberi de sensu composito et diviso, Regulae solvendi sophismata. Venice: Bonetus Locatellus, 1494.
"William of Heytesbury on the Three Categories." Selections translated by E. A. Moody. In *The Science of Mechanics in the Middle Ages*, edited by Marshall Claggett. Madison: University of Wisconsin Press, 1959. Reprinted in *A Source Book in Medieval Science*, edited by Edward Grant. Cambridge, MA: Harvard University Press, 1974.
"William of Heytesbury on the 'Insoluble' Sentences." Translated with notes by Paul Spade. Toronto: Pontifical Institute of Medieval Studies, 1979.

On Maxima and Minima: Chapter 5 of "Rules for Solving Sophismata," with an anonymous fourteenth-century discussion. Translated with introduction and study by John Longeway. Dordrecht: D. Reidel Publishing Company, 1984.

"'The Compounded and Divided Senses,' and 'The Verbs Know and Doubt.'" Translated by Norman Kretzmann and Eleonore Stump. In *The Cambridge Translations of Medieval Philosophical Texts.* Vol. 1: *Logic and Philosophy of Language.* Cambridge, U.K.: Cambridge University Press, 1988.

WORKS ABOUT WILLIAM HEYTESBURY

Wilson, Curtis. *William Heytesbury: Medieval Logic and the Rise of Modern Physics*. Madison: University of Wisconsin Press, 1960.

John Longeway (2005)

HICKOK, LAURENS PERSEUS

(1798–1888)

Laurens Perseus Hickok was America's first systematic philosopher and also won distinction as a theologian and educational administrator. He was born in Bethel, Connecticut, and was educated at Union College. He trained for the ministry under William Andrews and Bennett Tyler, who was a major spokesman for "old school" Calvinism. Hickok served well as pastor at Kent, Connecticut (1823–1829), and Litchfield, Connecticut (1829–1836). He then became professor of theology at Western Reserve College (1836–1844) and Auburn Theological Seminary (1844–1855). His alma mater, Union College, called him to serve as vice-president and professor of mental and moral philosophy (1855–1866) and president (1866–1868). In 1868 he retired to Amherst, Massachusetts, where he wrote several books over the next twenty years.

The core of Hickok's philosophic enterprise was the attempt to allow adequate weight to the claims of reason and experience in all domains of intellectual life. Ultimately, he was convinced, the rational and the empirical modes of thinking could not lead to contradictory conclusions; human intelligence might begin with general principles and rationally deduce facts or might begin with observed facts and gradually uncover general principles. In either case the facts were the same, and the principles were the same. Rational science is science as known by God; empirical science is science as learned by men. Different criteria of validity are to be applied to man's ideas in these two types of scientific thinking. In the empirical area ideas are tested by their experimental consequences; in the rational area ideas are tested by their congruence within a systematic pattern. Each type of thinking has, however, its proper place; the speculative mode should not be used when the investigative mode is in order, but neither should men become so enamored of empirical investigation that they neglect rational speculation.

Despite this careful balancing of empirical and rational method, Hickok did not regard the discoveries of empirical science as part of philosophy. His own work *Empirical Psychology; or, The Human Mind as Given in Consciousness* (1854) was an introspective study of the workings of the human mind. Hickok thought of this study as prephilosophic. He also published a philosophic work in the same field—*Rational Psychology; or, The Subjective Idea and the Objective Law of All Intelligence* (1849). Here no attention was given to the data of introspection; hence, this work was properly "philosophy." In all the other fields to which he gave consideration, Hickok's work was completely dominated by rational, speculative system building.

Although there was some trace of the ideas of Immanuel Kant in American philosophy before Hickok, he was the first professor of philosophy in the United States to attempt to make systematic use of Kant and the post-Kantian German rationalists. Thus, he was an important figure in the transition from the orthodox academic teaching of Scottish realism in the first half of the nineteenth century to the dominance of idealism in the latter part of the century.

See also Empiricism; Experience; Kant, Immanuel; Rationalism; Reason.

Bibliography

WORKS BY HICKOK

Rational Psychology; or, The Subjective Idea and the Objective Law of All Intelligence. Auburn, NY: Derby Miller, 1849; rev. ed., New York, 1870.

A System of Moral Science. Schenectady, NY: G.Y. Van Debogert, 1853; rev. ed., with the cooperation of Julius H. Seelye, Boston: Ginn and Heath, 1880.

Empirical Psychology; or, The Human Mind as Given in Consciousness. Schenectady, NY: G.Y. Van Debogert, 1854; rev. ed., with the cooperation of Julius H. Seelye, Boston: Ginn and Heath, 1882.

Rational Cosmology; or, The Eternal Principles and the Necessary Laws of the Universe. New York: Appleton, 1858.

Creator and Creation; or, The Knowledge in the Reason of God and His Work. New York: Lee, Shepard and Dillingham, 1872.

Humanity Immortal; or, Man Tried, Fallen and Redeemed. Boston: Lee and Shepard, 1872.

The Logic of Reason, Universal and Eternal. Boston: Lee and Shepard, 1875.

WORKS ON HICKOK

Union Worthies, Number Two; Laurens Perseus Hickok, Class of 1820, published by Union College in 1947, contains three appreciations of Hickok by Harold A. Larrabee, Herbert W. Schneider, and Julius Seelye Bixler, as well as a bibliography of Hickok's chief writings. Other discussions are in H. W. Schneider, *A History of American Philosophy* (New York: Columbia University Press, 1946; 2nd ed., 1963), and in Joseph L. Blau, *Men and Movements in American Philosophy* (Englewood Cliffs, NJ: Prentice-Hall, 1952).

J. L. Blau (1967)

HIDDENNESS OF GOD

Many people are perplexed that God (if such there be) does not make His existence more evident. For many of them, the hiddenness of God puts their faith to the test. Others, however, claim that God's hiddenness is the basis of an argument against His existence. While this claim is no newcomer to religious reflection, it has been the focus of renewed debate since the early 1990s.

Two preliminary observations are in order. First, the God in question is the God of traditional theism, a personal God who is unsurpassably good. Second, *the hiddenness of God* is an inapt term to use in an argument for the conclusion that there is no God since God is hidden only if there is a God; the term *inculpable nonbelief* is better. At a first approximation, the argument is that there are people who, through no fault of their own, lack belief that God exists; thus, since there is a God only if there is no inculpable nonbelief, there is no God.

RELATIONSHIPS WITH THE ARGUMENT FROM EVIL

The argument from inculpable nonbelief is related in several ways to the more familiar argument from evil and suffering against the existence of God.

First, inculpable nonbelief is supposed to be evidence against the existence of God independent of evil and suffering. To see how this can be, imagine a society in a world much like our own but in which there is no evil or suffering. While no argument from evil could arise in such a society, some of its citizens might maintain that there is a God while others maintain that there is not since there are inculpable nonbelievers.

Second, evil and suffering are much more powerful evidence than inculpable nonbelief. It is difficult to view inculpable nonbelief as nearly as bad as the horrors of Auschwitz or the suffering caused by the tsunami of December 26, 2004. Perhaps this is due, in part, to the fact that, unlike evil and suffering, inculpable nonbelief is not bad in itself—indeed, it is bad only if there is a God.

Third, although inculpable nonbelief is weaker, independent evidence for atheism, it is arguably stronger precisely because of the suffering in the world. That is because suffering constitutes a context in which one's expectation increases that God would make Himself and His love sufficiently clear. For one is more in need of the assurance and comfort that God's manifest love would bring when one suffers. Thus, its absence in the suffering of many people, especially horrific and intense suffering, is more striking.

Fourth, formulations of the argument from inculpable nonbelief parallel those of the more familiar argument from evil. For example, one commonly distinguishes logical (deductive) arguments from evil from evidential (inductive or probabilistic) arguments from evil. A logical argument from evil affirms of some known fact about evil that it is incompatible with theism, while an evidential argument does not, either because it affirms that the fact in question is not known but only reasonably believed, or because it affirms that the fact in question is only improbable given theism, not incompatible with it. One can distinguish arguments from inculpable nonbelief along the same lines.

SCHELLENBERG'S VERSION OF THE ARGUMENT

More than anyone else, John L. Schellenberg is responsible for renewing the contemporary debate with his *Divine Hiddenness and Human Reason* (1993). The main argument is this:

(1) There are people who are capable of relating personally to God but who, through no fault of their own, fail to believe

(2) If there is a personal God who is unsurpassably great, then there are no such people

(3) So, there is no such God (from 1 and 2)

According to Schellenberg (1) is a generalization from two facts. First, there are honest seekers of the truth who are atheists and agnostics. Second, there are individuals who belong to cultures that lack the idea of a personal God altogether (e.g., the Chinese race in the period from the beginning of their history until the Christian Middle

Ages). In defense of (2), Schellenberg offers the following subargument:

(2a) If there is a personal God who is unsurpassably great, then there is a personal God who is unsurpassably loving

(2b) If there is a personal God who is unsurpassably loving, then for any human person H and any time t, if H is at t capable of relating personally to God, H has it within H's power at t do so (i.e., will do so, just by choosing), unless H is culpably in a contrary position at t

(2c) For any human person H and any time t, H has it within H's power at t to relate personally to God only if H at t believes that God exists

(2d) So, if there is a personal God who is unsurpassably great, then for any human person H and any time t, if H is at t capable of relating personally to God, H at t believes that God exists, unless H is culpably in a contrary position at t (from 2a through 2c)

In effect, (2d) is (2) of the main argument.

Schellenberg regards (2) as a necessary truth, reflecting part of the meaning of "there is a personal God who is unsurpassably great." (An evidential or probabilistic version of the argument would say that [2] is only likely to be true.) So Schellenberg regards each of (2a) through (2c) as necessary truths. He thinks (2a) is just obviously necessary. (2b), however, is not obvious. Indeed, what does it mean? In particular, what does Schellenberg mean by "relating personally to God" and being "capable" of such a relationship? He means this. To relate personally to God is to interact with God in the various ways that theistic religious traditions describe: on the divine side, God's guiding, supporting, and forgiving one, for example, and on the human side, one trusting Him, showing gratitude, and worshipping Him, among other things. Crucially, such a relationship would involve an explicit consciousness of God's presence and interaction with one. This relationship is to be conceived of developmentally, not as something that comes complete and mature. To be capable of a personal relationship with God is to have the cognitive and affective equipment required to be conscious of God's presence and interaction with one and to hold the attitudes and to perform the behavior involved in such a relationship; it also requires possession of the concept of God, or at least the materials from which it can be constructed.

So why should one suppose that (2b) is true? Schellenberg argues that it follows from the nature of unsurpassable love and can be supported by analogy with the best sorts of human love as well. An unsurpassable lover would seek a kind of close, explicit participation in the life of his or her beloved for its own sake, as well as for the beloved's sake, so that the beloved could draw from it what he or she needed to flourish. This would be especially true in the divine-human case. A close, explicit interaction with God would bestow moral benefits. For example, it would enable one to more easily overcome character flaws and it would provide one with a model for other relationships. Moreover, it would bestow experiential benefits, such as peace and joy, security and support in suffering, and the pleasure of companionship. Of course, God would not force Himself on one, as that would make the relationship a sham. He would leave it up to one to enter into and maintain it. Thus—one's own resistance, as well as the consequences of one's prior free choices—would be the only thing He would allow to prevent one from relating personally with Him. Otherwise, He would always be available just for the asking.

As for (2c), Schellenberg argues that it is absolutely impossible for one to have a personal relationship with another unless one believes that the other exists. Thus, as a matter of logical necessity, one has it within one's power to relate personally to another only if one believes that the other exists. The same goes for us and God.

Schellenberg's argument has enjoyed much critical scrutiny. To this attention is now directed.

NON-THEODICAL CRITICISMS OF THE ARGUMENT

Some critics say that the argument does not show that there is no God since it leaves open the possibility of an impersonal God, or a personal God that is not unsurpassably great. Others say that since God is so absolutely different, even incomprehensible, nothing could count as evidence against God's existence, including inculpable nonbelief. These responses are irrelevant, however, since Schellenberg's target is a God that is at least somewhat comprehensible insofar as He is said to be personal and unsurpassably great.

Another criticism holds that the argument is an occasion for observing that unsurpassable greatness does not imply unsurpassable love. In this connection think of the Stoic view of eudaemonia (happiness), according to which the sage—the person who has achieved moral and intellectual perfection—would possess benevolence but lack upsetting emotions like empathy, ecstasy, fear, and grief since these passions would upset the life of bliss characteristic of the sage. On such a view (2a) is false. A

personal God who was unsurpassably great would be a divine sage and, as such, would not possess the sort of attachment and passion characteristic of the love exhibited by parents for their children.

Some critics deny (2b). They note that one's view of the implications of unsurpassable love depends on what human analogies one takes to be most salient. They suggest that an emphasis on maternal love of children supports (2b), whereas an emphasis on familiar adult love or the love of a benevolent reconstructive surgeon is more apt and supports the denial of (2b). One might worry about this since, first, benevolence is not love, second, maternal love is offered as an apt analogy in lived theistic religions, and, third, a perfectly loving God would empathize with the plight of those who seek Him but who through no fault of their own come up empty-handed.

Many critics say that there are no inculpable nonbelievers who are capable of relating personally to God; as such, (1) is false. Chief among them are those who argue, first, that there is sufficient evidence to believe, in creation and history, or through the witness of one's conscience or *sensus divinitatis*. Next, they argue that nonbelief is best explained by the willful sinfulness of nonbelievers, in which case it is not inculpable. For example, Jonathan Edwards (1970) argues that God has endowed human beings with the faculties to discern, appreciate, and weigh the evidence for God's existence, but those faculties work properly only if they function in accordance with "true benevolence," which consists mainly in an intense desire for truth about God and for true holiness. So while there is plenty of evidence, some lack it because they lack true benevolence. In *Original Sin* Edwards denies that there are nonbelievers who possess true benevolence; after all, he says, the scriptures say (compare *Romans* 1:19–22) that there is "sufficient light for the knowledge of God," hence, nonbelievers must fail to believe "divine things" owing to "a dreadful stupidity of mind, occasioning a sottish insensibility of their truth and importance" (1970, p. 149, 157). This insensibility consists in a "proneness to idolatry" and a "disregard of eternal things"—dispositions to ignore familiar and obvious considerations, to be swayed by ridicule and deference to people in authority, to prejudice against religion, and so on—which impair the God-given ability to reason properly about God. People bring such impairments on themselves. One worry about this criticism is that, even though some nonbelievers lack true benevolence, the empirical evidence strongly suggests that others possess it since they really do earnestly seek the truth about God, love the Good, assess evidence judiciously, and, if anything, display a prejudice for God, not against Him.

Some critics appeal to implicit belief. The idea is that since God is the Good (or, God's moral goodness is His most salient feature), pursuit of the Good is, in fact, pursuit of God, even if one does not recognize it as such. This thought can be taken in different directions. On the one hand, one might infer that, since belief in the Good just is (or is one way to have) belief that God exists, one is a nonbeliever only if one fails to pursue the Good, a failure for which one would be culpable; so, there are no inculpable nonbelievers, and (1) is false. On the other hand, one might deny that belief in the Good is belief that God exists but still infer that one can relate personally with God (just by choosing) even if one does not believe that God exists since, after all, belief in the Good (moral goodness) is sufficient for the early stages of a developing personal relationship with God; as such, (2c) is false.

Another possibility is that one can begin to develop a personal relationship with God without belief that God exists. One option here is a kind of faith that God exists that has as its cognitive component acceptance rather than belief. Belief differs from acceptance in that, first, acceptance is a mental act, rather than a dispositional mental state (which is not to say that acceptance does not engender a complex behavioral disposition), and second, acceptance is under voluntary control while belief is not. Regarding the first point, although accepting a proposition is like believing it in that accepting it involves a positive stance toward it, accepting a proposition is unlike believing it in that accepting it involves one adopting it or taking it on for the purposes of theoretical and practical reasoning as well as behavior, even though one is not disposed to think "yes, that's how things are" on considering it, which is essential to belief. As for the second point, belief is a state one finds in oneself, the causal consequence of one's reasons, evidence, or grounds. However, when one's grounds for a proposition seem ambiguous, one can choose to accept it or choose to withhold acceptance. Now, if one's faith that God exists involves acceptance but not belief, one will nevertheless be disposed to act and feel in ways appropriate to God's existence (e.g., worshipping and feeling gratitude) and one will accept various experiences and sacraments as God's interacting with one (e.g., forgiving, guiding, and supporting one). In that case, one might argue that (2c) is false: one can have it within one's power to begin to relate personally to God even if one does not believe that God exists; faith is enough.

A SOUL-MAKING THEODICY

Many critics concur with (2c) but argue that (2b) is false. Toward that end, they offer *theodicies*, that is, reasons why God might lovingly permit inculpable nonbelief. Several general themes have emerged in the literature. First, God may well prefer temporary nonbelief to belief accompanied by a negative response. Second, God might have different reasons for different individuals depending on what attitudes and dispositions they possess; likewise, He might have different reasons for the same individual at different times. Third, God might have a combination of reasons, no one of which is enough but which, taken together, explain His permission of inculpable nonbelief. Fourth, on some versions of theism everyone will eventually have evidence sufficient for belief, if not in this life then in the next; so theodicies involve reasons for God to permit inculpable nonbelief for a time, not forever. Finally, evidence sufficient for belief that God exists need not involve arguments or spectacular "signs and wonders"; experiential awareness of God is enough. What follows are some representative theodicies.

According to one version of the *soul-making theodicy*, many people, at the dawn of their capacity to relate personally to God, are already ill disposed toward Him. Through no fault of their own, they have become inculpably ill-disposed nonbelievers. These include many of those who were raised to be hostile or indifferent toward religion, who were abused by excessively strict religious parents, or who had instilled in them an extreme self-centeredness or disrespect for proper authority. Perhaps God refrains from giving such people evidence sufficient for belief because they would not respond appropriately if they had it. Moreover, there's a grave risk in bringing inculpable ill-disposed nonbelievers to belief since there is a good chance they will confirm their defective disposition by an unfitting response; indeed, it might even be useless to give them evidence since they might be so ill disposed that they are more inclined to think they are institutionalizable ("hearing voices") than that God is communicating with them. Consequently, God waits, giving them the opportunity to become more receptive and apt to reciprocate His love, and influences them in subtle but respectful and loving ways toward this end.

As for inculpable nonbelievers who are well disposed toward God, this version of the soul-making theodicy considers separately those who were responsible for becoming well disposed and those who were not. Examples of the first group include those who have been virtually determined—say, by parental training—to become well disposed but who do not yet believe. In that case, as they become capable of relating personally to God, they are disposed to love God but they had little if any say in becoming so disposed. This is unfortunate because, all else being equal, a state of affairs in which one reciprocates God's love but had little if any say in being so disposed is not nearly as good as a state of affairs in which one reciprocates God's love and had a significant say about being so disposed. God prefers the better state of affairs, and so He does not bring to belief the well-disposed inculpable nonbeliever who is not responsible for being so disposed because He prefers them to confirm their disposition, on their own, in the face of contrary desires and competing allegiances, before bringing them to belief. In that way God allows them to make their involuntarily acquired good disposition toward Him genuinely theirs.

Now consider well-disposed inculpable nonbelievers who were responsible for becoming so disposed. Given the influences that shape childhood character, these will most likely be adults who have either reshaped their bad dispositions toward God for the good or confirmed their good dispositions over time. They constitute the most difficult case for the soul-making theodicy; nevertheless, it has some resources. For, as is well known, one can be disposed to love another for the wrong reasons. For example, sometimes one's love springs from a desire to extend one's power or influence, increase one's pleasure, or satisfy one's curiosity. Other times its source is insecurity or fear, for example, fear of being alone or unprotected. And there are other sources. Likewise with God. One may well be disposed to love God on coming to belief, but one might be so disposed for reasons that are not as fitting as they might be. For example, it is most fitting to love God mainly for His moral beauty, His holiness; relatedly, perhaps no disposition to love God is suitably motivated unless it is grounded in a strong desire to surrender wholly to His will. In that case the possibility arises that if God were to bring such people to belief, they would love Him, but their love would not be appropriately motivated. So He woos them, before bringing them to belief, influencing them behind the scenes in respectful and loving ways to change the source of their disposition to love Him and to confirm that change over time.

OTHER THEODICIES

A variety of other theodicies have been articulated. The *presumption theodicy* states that God does not bring some individuals to belief because if He did, they would relate to God in presumptuous and arrogant ways, not with due

contrition and humility, which are essential to a proper relationship with God. According to the *stimulus theodicy*, God does not produce belief in some individuals because if He did, they would be less apt to recognize the wretchedness of living life on their own, without God. Divine hiddenness stimulates such people to recognize this fact about the human condition, which is essential for entering into a proper relationship with God. A variation on the stimulus theodicy states that God does not provide evidence sufficient for belief in some individuals because if He did, the perceived risk required for an intense and passionate faith would be objectionably reduced, and without such a faith one cannot wholly enter into a proper relationship with God.

The *deception theodicy* asserts that some people are disposed in such a way that if God brought them to belief they would be deceived into thinking that they had arrived at a proper understanding of religious matters and would become complacent or relate to God at a superficial level. Proponents of the *intellectual virtue theodicy* say that God does not provide evidence sufficient for belief in some individuals because if He did, certain intellectual temptations would not be available to them and they would not have the opportunity to respond to those temptations virtuously. For example, if the evidence were too clear, sustained investigation and reflection and wrestling with doubt would be inhibited.

According to the *diversity theodicy* God does not produce belief in some people because if He did, diverse expression of religious imagination, creativity, and devotion would be greatly reduced, and religious variety of this sort is a great good. Finally, advocates of the *investigation theodicy* hold that it is a great good to pursue knowledge with others, all the more so when the knowledge is as important as knowledge of God. But people can pursue knowledge together only if some of them are ignorant. So God permits inculpable nonbelief so that human beings might help each other to learn about Him and to assist nonbelievers in starting personal relationships with Him.

Naturally enough, these theodicies have been criticized. Some critics claim that they provide no good reason for God to permit inculpable nonbelief at all, or at least not for every sort of inculpable nonbeliever. Others insist that if they are good reasons, then the problem of too much belief arises. Most importantly, it is claimed that the benefits of temporary inculpable nonbelief articulated in the theodicies can be accommodated within a developing, explicit personal relationship with God that involves evidence sufficient for belief that He exists.

One final critique of Schellenberg's argument should be mentioned. His argument invites one to affirm, at least tacitly, that there is no reason for God to permit inculpable nonbelief. Two themes have emerged on this score. First, one should accept the invitation only if the theodicies fail, individually but especially collectively, to account for why God might lovingly permit inculpable nonbelief. The worry here is that human beings are enormously complicated, and it is no easy task to tell whether any particular candidate for inculpable nonbelief possesses or fails to possess those motivations, attitudes, and dispositions that figure in the theodicies above. Second, even if there are inculpable nonbelievers whose nonbelief cannot be fully explained by any theodicy one knows of, the live possibility remains that there is some theodicy one does not know of. Indeed, would it really be all that surprising if God had some purpose for permitting inculpable nonbelief, as well as other bad things that happen, that one cannot understand?

See also Atheism; Common Consent Arguments for the Existence of God; Edwards, Jonathan; Eudaimonia; Evil; Evil, The Problem of; Happiness; Popular Arguments for the Existence of God.

Bibliography

Alston, William P. "Belief, Acceptance, and Religious Faith." In *Faith, Freedom, and Rationality: Philosophy of Religion Today*, edited by Jeff Jordan and Daniel Howard-Snyder. Lanham, MD: Rowman & Littlefield, 1996.

Drange, Theodore M. *Nonbelief and Evil: Two Arguments for the Nonexistence of God*. Amherst, NY: Prometheus Press, 1998.

Edwards, Jonathan. *Original Sin*, edited by Clyde A. Holbrook. New Haven, CT: Yale University Press, 1970.

Henry, Douglas V. "Does Reasonable Nonbelief Exist?" *Faith and Philosophy* 18 (1) (2001): 75–92.

Hick, John. *Faith and Knowledge*. 2nd ed. Ithaca, NY: Cornell University Press, 1966.

Howard-Snyder, Daniel. "The Argument from Divine Hiddenness." *Canadian Journal of Philosophy* 26 (3) (1996): 433–453.

Howard-Snyder, Daniel, and Paul K. Moser, eds. *Divine Hiddenness: New Essays*. New York: Cambridge University Press, 2002.

Lehe, Robert T. "A Response to the Argument from the Reasonableness of Nonbelief." *Faith and Philosophy* 21 (2) (2004): 159–174.

McKim, Robert. "The Hiddenness of God." *Religious Studies* 26 (1) (1990): 141–161.

McKim, Robert. *Religious Ambiguity and Religious Diversity*. New York: Oxford University Press, 2001.

Morris, Thomas V. "The Hidden God." *Philosophical Topics* 16 (2) (1988): 5–21.

Moser, Paul K. “Divine Hiddenness Does Not Justify Atheism” and “Reply to Schellenberg.” In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. VanArragon. Malden, MA: Blackwell, 2004.

Murray, Michael J. “Coercion and the Hiddenness of God.” *American Philosophical Quarterly* 30 (1993): 27–38.

Pascal, Blaise. *Pensées*. Translated by A. J. Krailsheimer. Harmondsworth, U.K.: Penguin, 1966.

Penelhum, Terence. *Reason and Religious Faith*. Boulder, CO: Westview Press, 1995.

Schellenberg, John L. *Divine Hiddenness and Human Reason*. Ithaca, NY: Cornell University Press, 1993.

Schellenberg, John L. “Divine Hiddenness Justifies Atheism” and “Reply to Moser.” In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. VanArragon. Malden, MA: Blackwell, 2004.

Schellenberg, John L. “The Hiddenness Argument Revisited (I)” and “The Hiddenness Argument Revisited (II).” *Religious Studies* 41 (2005).

Schellenberg, John L. *Prolegomena to a Philosophy of Religion*. Ithaca, NY: Cornell University Press, 2005.

Swinburne, Richard. *Providence and the Problem of Evil*. Oxford, U.K.: Clarendon Press, 1998.

Wainwright, William J. *Reason and the Heart: A Prolegomenon to a Critique of Passional Reason*. Ithaca, NY: Cornell University Press, 1995.

Daniel Howard-Snyder (2005)

HILBERT, DAVID

(1862–1943)

David Hilbert, the German mathematician, was born in Königsberg and, except for a semester at Heidelberg, did his university studies there. His dissertation, presented in 1884, was on a problem in the theory of algebraic invariants, and it was to this theory that Hilbert devoted his mathematical researches until 1892. Through these penetrating investigations Hilbert obtained many pregnant results, some of them (*Hilbertscher Nullstellensatz, Hilbertscher Irreduzibilitätssatz*) still known by his name. The methods he used in these investigations inaugurated a trend toward treating algebra more conceptually and have since become dominant in the field.

In 1886 Hilbert became a *Privatdozent* and in 1892 an extraordinary professor at the University of Königsberg. In 1893 he was named by the minister of culture Friedrich Althoff to succeed his teacher, Felix Lindemann, as an ordinary professor at Königsberg. In 1895 Hilbert accepted an invitation initiated by Felix Klein, to the University of Göttingen to succeed Heinrich Weber. Hilbert remained at Göttingen, despite many offers of other chairs, notably from Leipzig in 1898, Berlin in 1902, and Heidelberg in 1904. The invitation from Berlin led to Hilbert's obtaining, through the help of Althoff, a chair at Göttingen for Hermann Minkowski, whom Hilbert had known since they were students at Königsberg. The personal intercourse between the two investigators was highly stimulating to both men but was prematurely ended, to Hilbert's grief, by Minkowski's death in 1909.

Hilbert's most important mathematical investigations were carried out between 1892 and 1909. He simplified the existing transcendence proofs for the numbers e and π. His investigations in the theory of algebraic number fields, in particular his monumental report “Die Theorie der algebraischen Zahlkörper” (1897), greatly amplified existing theory and directed further research in the field. His famous *Grundlagen der Geometrie* is discussed below. He showed the possibility of directly supporting the Dirichlet principle, that the existence of a conformal mapping may be inferred from the presumed existence of a minimum of a certain integral (which Bernhard Riemann had taken as the basis for his general theorems concerning conformal mappings), by means of an existence proof. This method for giving an existence proof, when worked out by Richard Courant and Hermann Weyl, proved very successful. Hilbert's contributions to the calculus of variations, in particular his statement of the *Unabhängigkeitsatz* (“independence axiom”), constituted an illuminating commentary on Adolf Kneser's textbook in the field. He continued the theory of Ivar Fredholm concerning integral equations. In particular, he introduced the analysis of infinitely many variables and generalized the transformation to principal axes. The theory thus established has proved highly fruitful in topology and in physics, particularly in quantum mechanics. Utilizing a result of Adolf Hurwitz, Hilbert solved the Waring problem concerning the representation of natural numbers by sums of nth powers.

Hilbert's familiarity with the various domains of mathematics was impressively demonstrated by the address “Mathematische Probleme,” which he presented at the Second International Congress of Mathematicians in Paris in 1900. In this address Hilbert surveyed the situation then existing in mathematics, at the same time formulating twenty-three problems that have much occupied mathematicians since then. A great many of these problems have been solved in the meantime.

After Minkowski's death Hilbert turned to problems of theoretical physics. He first applied the theory of integral equations to the kinetic theory of gases and to the theory of radiation. Immediately after the appearance of Albert Einstein's general theory of relativity, Hilbert published “Die Grundlagen der Physik” (1915–1916), which

offered the first proposal of a way to unify gravitational theory and electrodynamics.

After 1916 Hilbert returned to the problems of the foundations of mathematics. These investigations led to the development of proof theory, which will be discussed below.

In his later years Hilbert gave lectures providing careful general surveys of mathematics, such as "Anschauliche Geometrie" (on intuitive geometry), as well as popular philosophical lectures. The spirit of these philosophical lectures can be seen in the speech "Naturerkennen und Logik," which he gave at the congress of natural scientists in Königsberg in 1930. At this congress his native city named him an honorary citizen.

Hilbert's character was not that of a specialized scientist. He took pleasure in the joy of life, especially in sociability, and also took a vivid interest in political events. He enjoyed the exchange of ideas both in science and in general thought; in discussions he had a predilection for pregnant, sometimes paradoxical, formulations.

Hilbert had a great many pupils, and he was the adviser on many famous dissertations whose themes were suggested by his investigations. He had the satisfaction of seeing his work highly appreciated in his own lifetime.

The memory of Hilbert's personality is vivid in all those who knew him, and the impulses he gave to science remained effective for decades after his death.

THE FOUNDATIONS OF GEOMETRY

In Hilbert's scientific work, his studies in the foundations of mathematics constitute an important part. These investigations fall into two stages separated by an interval of nearly thirteen years. The first period, which extends from about 1893 to 1904, embraces Hilbert's inquiries into geometric axiomatics and is highlighted by the publication of the *Grundlagen der Geometrie* (1899), the work that made Hilbert's name familiar to a wide public of scientists and philosophers. The second period, which began with the publication in 1917 of "Axiomatisches Denken," centers on the foundations of arithmetic and the development of Hilbert's program for proof theory.

ABSTRACT AXIOMATICS. A main feature of Hilbert's axiomatization of geometry is that the axiomatic method is presented and practiced in the spirit of the abstract conception of axiomatics that arose at the end of the nineteenth century and which has been generally adopted in modern mathematics. It consists in abstracting from the intuitive meaning of the terms for the kinds of primitive objects (individuals) and for the fundamental relations and in understanding the assertions (theorems) of the axiomatized theory in a hypothetical sense, that is, as holding true for any interpretation or determination of the kinds of individuals and of the fundamental relations for which the axioms are satisfied. Thus, an axiom system is regarded not as a system of statements about a subject matter but as a system of conditions for what might be called a relational structure. Such a relational structure is taken as the immediate object of the axiomatic theory; its application to a kind of intuitive object or to a domain of natural science is to be made by means of an interpretation of the individuals and relations in accordance with which the axioms are found to be satisfied.

This conception of axiomatics, of which Hilbert was one of the first advocates (and certainly the most influential), has its roots in Euclid's *Elements,* in which logical reasoning on the basis of axioms is used not merely as a means of assisting intuition in the study of spatial figures; rather, logical dependencies are considered for their own sake, and it is insisted that in reasoning we should rely only on those properties of a figure that either are explicitly assumed or follow logically from the assumptions and axioms. This program was not strictly adhered to in all parts of the *Elements,* nor could it have been, for its system of axioms was not sufficient for the purpose. The first axiom system meeting the requirements of the program was given by Moritz Pasch in his *Vorlesungen über neuere Geometrie* (Leipzig, 1882).

This abstract kind of axiomatics, which consists in separating out the purely mathematical aspects of a theory, is not the only possible one. Hilbert himself knew that it can be applied advantageously only in domains of science whose theoretical development is sufficiently advanced. But abstract axiomatics is useful wherever the logical dependence or independence of theoretical assumptions is under investigation.

The distinguishing property of Hilbert's axiomatics is frequently described by saying that in it the terms for the kinds of elements (points, straight lines) and for the relations (incidence, betweenness, congruence) are *implicitly defined* by the axioms. This expression, first introduced in 1818 by J. D. Gergonne (Hilbert did not employ it), is often used in a misleading way. The axioms generally impose conditions on the relations and on the kinds of elements of the system; some of these conditions are partial characterizations of the relations or the kinds of elements, others characterize the space with respect to the elements and relations. The entire axiom system—as Hilbert observed in a letter to Gottlob Frege—can be

regarded as a single definition. But this is an *explicit* definition of a term denoting the relational structure in question. The defined concept is a predicate of the second type (*zweiter Stufe,* as Frege called it), applying to domains of things and to certain relations between them.

NON-ARCHIMEDEAN SYSTEMS. Another main feature of Hilbert's *Grundlagen der Geometrie* is the development of geometry, and, in particular, of plane geometry, independently of the Archimedean axiom. This axiom states that given any two line segments, either may be exceeded by an entire multiple of the other. Thus, it partly compensates for the absence of general commensurability of line segments. It was with the help of this axiom that the theory of proportions was established in Book V of Euclid's *Elements* (attributed to Eudoxus). It is also a consequence of this axiom that, once a unit segment is chosen, there corresponds to any line segment a real number (in Richard Dedekind's sense of the term) that is its measure (ratio number); therefore, Hilbert also called the Archimedean axiom the axiom of measurement.

Recourse to the Archimedean axiom introduced an arithmetical element into reasoning, and hence avoidance of it in geometrical proofs amounts to an emancipation from a nongeometrical type of reasoning. The avoidance of nongeometrical reasoning does not preclude an analytic geometry. In fact, Hilbert was able to construct a calculus of line segments, independent of the Archimedean axiom, in two different ways.

One method operates within the framework of metric plane geometry. It is based on the axioms of incidence (for the plane), those of order, those of congruence, and the parallel axiom. Hilbert defines the sum of segments in the usual way and the product of segments, after establishing a unit segment, by a parallel construction; he then shows that by these definitions the usual computation laws for sum and product are satisfied.

By this segment calculus an elementary foundation of the theory of proportions and thereby also of analytic geometry is obtained. Hilbert further showed how with the aid of the segment calculus the theory of the areas of polygons can be set up without supposing, as is assumed in Euclid, that to any polygon there corresponds its area as a quantity, that is, in agreement with Euclid's axioms of quantities. Thus he showed that no accessory reliance on intuition is required for the theory of areas of polygons.

Hilbert conjectured that the theory of the volumes of polyhedrons is not fully analogous to the theory of the areas of polygons. He posed the problem of showing that tetrahedrons of equal volume cannot always be obtained from one another by a series of processes of pairwise additions and subtractions of congruent polyhedrons, a problem solved by Max Dehn ("Über raumgleiche Polyeder" and "Über den Rauminhalt"). Various investigations have derived from this problem.

Hilbert's second calculus of line segments independent of the Archimedean axiom is for *affine* geometry of the plane. A difficulty here is that the axioms of plane affine geometry do not suffice for the foundation of this geometry. The same holds for plane projective geometry.

Hans Wiener stated at the Naturforscherversammlung in Halle (1891) that it is impossible to give autonomous foundations to both plane projective geometry and plane affine geometry by adjoining to the axioms of incidence the Desargues theorem and a specialized form of the Pascal theorem on conic sections (with the conic section degenerated to a pair of straight lines). Hilbert was impressed by these statements and gave a proof of them for affine geometry by means of a calculus of segments. Here sum and product of segments are defined by elementary parallel constructions, and, with the aid of the Desargues theorem, the computation laws, with the exception of the commutative law for the product, are proved to be satisfied. These proofs were simplified by Arnold Schmidt in the seventh edition of the *Grundlagen.*

This calculus of segments leads to an analytic geometry over a skewfield—as it has come to be called—for the plane. This geometry can be extended, as Hilbert showed, to an analytic geometry of three-dimensional space satisfying the incidence axioms and the parallel axiom for the space. This is the extent of the role of the Desargues theorem. The specialized Pascal theorem is needed to prove that the segment calculus satisfies the commutative law for multiplication. This law, as Hilbert showed, can be inferred from the other computation laws and the laws of order with the aid of the Archimedean axiom, but not without it. (Gerhard Hessenberg proved, somewhat later, that the Desargues theorem is a consequence of the specialized Pascal theorem.)

Hilbert's positive treatment of the Archimedean axiom and, in particular, the question of its independence complemented his elimination of it from the foundations of geometry. The possibility of a non-Archimedean geometry was first considered in detail by Giuseppe Veronese in his *Fondamenti di Geometria* (Padua, 1891). This possibility can be inferred, by the methods of analytic geometry, from the existence of a (generalized) number system for which the operations of sum and product and their inverses, as well as the operation

$\sqrt{1 + a^2}$ and the relation "smaller than," can be defined in such a way that the familiar computation laws, but not the Archimedean axiom, are satisfied.

Hilbert gave as an instance of such a non-Archimedean system a system whose elements are algebraic functions of an argument *t*. But the instance he presented in "Über den Satz von der Gleichheit der Basiswinkel im gleichschenkligen Dreieck" is easier to operate with. (This essay is one of a series of studies closely connected with the *Grundlagen* and added to it as appendixes in the second and later editions; this essay forms Appendix II.) It deals with the possibility of restricting, in plane geometry, the last congruence axiom concerning triangles to the case of triangles assigned to one another in equal orientation. The effect of this restriction is to admit as congruences only those transformations obtained by translations and plane rotations, thus excluding symmetry from the notion of congruence. Two kinds of questions arise, those concerning the anomalies that can occur in a model of the restricted axiom system and those relating to the ways of compensating for the weakening of the triangle congruence axiom. Many anomalies are stated by Hilbert to occur in two models that he ingeniously constructed. Concerning different methods of compensating for the restriction of the triangle congruence axiom, see Supplement V2 of the ninth edition of the *Grundlagen* (pp. 264–268) and the literature mentioned there in the footnote on p. 265.

CHARACTERIZATION OF THE PLANE. As Blumenthal's biography indicates, Hilbert was led to the problems of Appendix II by investigations (reprinted in Appendix IV) in which he gave a very different foundation for plane geometry from that presented in the main part of the book. The problem here is to characterize the plane by means of the properties of the manifold of congruent motions. It was first treated by Hermann von Helmholtz and soon after by Sophus Lie, who emphasized its group-theoretic aspects. Both Helmholtz and Lie proceeded by the methods of the differential calculus. Hilbert sought to avoid any assumption concerning differentiability. His arguments in Appendix IV are within the framework of the theory of point sets. They rely especially on Camille Jordan's theorem concerning simply closed continuous curves (Jordan curves) in the "number plane," which generalizes the theorem on the decomposition of the plane by a simple polygon. Hilbert starts from a characterization of the geometric plane as a two-dimensional manifold by means of the concept of neighborhoods introduced in an axiomatic way—this is now a familiar method in topology, but at that time it was scarcely known at all.

Two characterizations of the "plane" are offered. According to the narrower definition the plane is topologically equivalent to a connected domain in the number plane; according to the wider definition it is locally equivalent (homeomorphic) to the interior of a Jordan curve and is globally connected. Hilbert chose the narrower characterization for simplicity.

The concept of continuous transformation can be defined by means of the representation of the geometric plane in the number plane. The motions are then taken as special continuous one-to-one transformations of the geometric plane onto itself such that in the representation each Jordan curve preserves its orientation. This provisional characterization of the geometric plane is then completed by three axioms on motions: (1) The motions constitute a group with respect to their composition; (2) given two different points, *A* and *B*, there are infinitely many points into which *B* can be transformed by a motion keeping *A* fixed; (3) if *A*, *B*, *C* and *A′*, *B′*, *C′* are triples of points in the geometric plane (the members of a triple not necessarily being different) and if in an arbitrary proximity of *A*, *B*, *C* there exist triples *P*, *Q*, *R* and in an arbitrary proximity of *A′*, *B′*, *C′* there exist triples *P′*, *Q′*, *R′* such that *P*, *Q*, *R* is transformed by a motion into *P′*, *Q′*, *R′*, then *A*, *B*, *C* is transformed by a motion into *A′*, *B′*, *C′*.

In a valuable discussion that made use of set-theoretic, topological, and group-theoretic arguments, Hilbert proved that from these axioms, with the obvious definition of congruence by means of the concept of motion and a suitable set-theoretic definition of straight line, it follows that the geometric plane under consideration satisfies the axioms of plane geometry as stated in the main part of the *Grundlagen,* with two exceptions: (1) the triangle congruence axiom is obtained only in the restricted form relating to motions, and (2) the parallel axiom does not result. Two possibilities then remain: the plane satisfies either Euclidean geometry or Bolyai-Lobachevski geometry.

Hilbert's handling of these problems disclosed a new direction of investigation, which is still being pursued. His results have been extended in three ways: (1) by weakening the topological assumptions through the adoption of the wider characterization, mentioned above, of a two-dimensional manifold, (2) by generalizing the discussion to higher dimensions, and (3) by modifying the axioms on the motions. (See the surveys of these researches in the introduction to Hans Freudenthal's "Neuere Fassung des

Riemann-Helmholtz-Lieschen Raumproblems" and his "Im Urkreis der sogenannten Raumprobleme.")

CONTINUITY. A final aspect of Hilbert's axiomatization of geometry in the *Grundlagen* is his treatment of continuity. The Archimedean axiom is stated as an axiom of continuity, yet it excludes only a particular kind of discontinuity. In fact, if this axiom alone is added to the Hilbert axioms of incidence, order, and congruence (including the parallel axiom), then the axiom system is satisfied by an analytic geometry constructed over a *restricted* number system consisting only of algebraic numbers and not including the square root of each positive number.

In this respect Hilbert's axioms differ from those of Euclid's *Elements.* Euclid explicitly postulated the construction of a circle around a given point with a given radius (and implicitly made assumptions about the intersection of circles and of circles with straight lines). However, in order to realize by constructions the existence statements of Hilbert's axioms, it is sufficient to have, in addition to a ruler, not a compass but an "*Eichmass*"—that is, an instrument for determining a given distance on a given straight line from a given point in a prescribed direction. Hilbert showed, in Chapter 7 of the *Grundlagen,* that the *Eichmass* and the ruler allow us to perform all the constructions corresponding to the existence axioms.

Chapter 7 also discusses the question of the analytical representation of the constructions with ruler and *Eichmass.* It turns out that the ratio numbers of line segments constructible from a given unit length with ruler and *Eichmass* are the real numbers obtainable by the elementary arithmetical operations together with the operation $\sqrt{1 + c^2}$. This domain of numbers is narrower than that obtained when the operation $\sqrt{1 + c^2}$ is replaced by that of extracting the square root of an arbitrary positive number. The latter domain is the one composed of the ratio numbers of the lengths constructible by ruler and compass, but by no means does it contain all algebraic numbers. Yet, whereas the set of all algebraic numbers is denumerable, the set of all ratio numbers has a higher infinity. Hence, in order to characterize the geometric continuum a further axiom is required. It then becomes apparent that geometric continuity is related to continuity in the theory of real numbers.

When Hilbert wrote the *Grundlagen* the question of conceptually formulating the continuity property of an ordered set had been settled by the Dedekind axiom of *Lückenlosigkeit* and its equivalent, the principle of the least upper bound. For a metrical set, each of these axioms implies the Archimedean property.

COMPLETENESS. In direct connection with his work on the foundations of geometry, Hilbert undertook an axiomatization of theory of real numbers. In the paper "Über den Zahlbegriff" (published in 1900 and reprinted as Appendix VI of the *Grundlagen*), he presents an axiom system characterizing the system of real numbers as an ordered Archimedean field that cannot be extended to a wider ordered Archimedean field. He thus replaced the continuity axiom by (1) the Archimedean axiom and (2) a condition of maximality which he called the axiom of completeness.

Hilbert introduced into geometry a corresponding axiom of completeness (which first appears in the second edition of the *Grundlagen*) stating that the space characterized by the axiom system including the axiom of completeness constitutes a maximal (that is, not extensible) model of the other axioms. The connection between the geometrical and the arithmetical completeness axiom is given by the circumstance that any model of the axioms of incidence, order, and congruence and of the parallel and the Archimedean axiom can be represented by an analytic geometry over an ordered Archimedean number field, which again is isomorphic with respect to sum, product, and order to a subfield of the field of all real numbers.

The statement of the completeness axiom is very suggestive, and it was with Hilbert's introduction of this axiom that the notion of a maximal model was first conceived. Yet, because of its reference to other axioms, the completeness axiom offers difficulties, particularly with respect to questions of independence. The possibility of decomposing the full continuity axiom into the Archimedean axiom and another axiom which does not entail it is given by Georg Cantor's continuity axiom. (See Federigo Enriques, "Prinzipien der Geometric," and Richard Baldus, "Zur Axiomatik der Geometrie III: Über das Archimedische und das Cantorsche Axiom.")

CONSISTENCY. In "Über den Zahlbegriff" Hilbert recommended substituting an axiomatic presentation of the theory of real numbers for the "genetic" method of treating them. Despite the great pedagogical value of the genetic method, he said, the axiomatic method is to be preferred for the definitive formulation and logical precision of the theory.

This point of view has decisive consequences for the problem of consistency. Hilbert proved the consistency of

the geometrical axiom system by using the arithmetical model provided by analytic geometry. But if arithmetic is set up as an axiomatic theory, then Hilbert's proof establishes only a relative consistency. This, of course, is a valuable result, since the structure described by the axioms for the arithmetical continuum is much simpler than that of Euclidean space. The reduction to arithmetic, however, cannot then be regarded as a kind of direct verification by intuitive evidence, for the task of proving the consistency of the axiomatic theory of real numbers remains. This problem was one of those Hilbert posed in "Mathematische Probleme" (*Gesammelte Abhandlungen,* Vol. III, pp. 290–329).

At that time Hilbert thought that a suitable modification of the methods of Dedekind and Karl Weierstrass in the theory of irrational numbers would suffice to obtain the desired proof of consistency. Not long after, however, in the address "Über die Grundlagen der Logik und der Arithmetik" to the Heidelberg Congress of Mathematicians, Hilbert presented an essentially altered view. This alteration was no doubt brought about through the discovery by Bertrand Russell and Ernst Zermelo of very significant forms of the logical paradoxes which gave a more fundamental aspect to the difficulties that Cantor had earlier found with respect to "inconsistent sets." These difficulties showed that in set theory we cannot in general assign to a predicate *P* "the set of all those things for which *P* holds" as an object belonging to the universe of discourse.

Hilbert stated that these paradoxes seemed to show that the views and methods of logic "conceived in the traditional sense" ("*im hergebrachten Sinne aufgefasst*") are not equal to the strong requirements of set theory. And, although he strongly opposed Leopold Kronecker's tendency to restrict mathematical methods, he nevertheless admitted that Kronecker's criticism of the usual way of dealing with the infinite was partly justified.

The resulting point of view was not yet explicitly developed in Hilbert's Heidelberg address. However, Hilbert presented there the following programmatic ideas: (1) One must include in the arithmetical theory whose consistency is to be demonstrated the methods of logical reasoning used in the theory; (2) the methods of symbolic logic for representing mathematical sentences by formulas are to be applied; (3) the sequences of formulas representing mathematical proofs can be made the object of intuitive elementary reasoning regarding their structural properties and relations, and in this way proofs of consistency can be carried out. Various devices for proving consistency were also exhibited.

Hilbert's investigations of the foundations of arithmetic remained in this provisional state for a long time. During the interval major developments took place in the foundations of mathematics and in mathematical logic. Zermelo proved the well-ordering theorem and published his axiom system for set theory in 1908. Two years later the first volume of Russell and Alfred North Whitehead's *Principia Mathematica* appeared. Julius König attempted to carry out Hilbert's plan, but his work was interrupted by his premature death and appeared only in fragmentary form, edited by his son, in 1914 (*Neue Grundlagen der Logik, Arithmetik und Mengenlehre,* Leipzig, 1914). In this work some steps of the later Hilbert proof theory are already carried out, but Hilbert did not know of it when he again took up his investigation of the foundations of arithmetic.

PROOF THEORY

Hilbert's return to the problem of the foundations of arithmetic was announced by his delivery at Zürich in 1917 of the lecture "Axiomatisches Denken." In the latter part of this lecture he pointed out several epistemological questions which, as he said, are connected with that of the consistency of number theory and set theory: the problem of the solubility in principle of every mathematical question; that of finding a standard of simplicity for mathematical proofs; that of the relation of contents and formalism in mathematics; and that of the decidability of a mathematical question by a finite procedure. Questions of this kind, he observed, seem to constitute a domain that should be investigated, and to carry out this investigation it will be necessary to inquire into the concept of mathematical proof. The general idea and the aims of proof theory were thus proclaimed, but the means of investigation were not thereby fixed, for indeed the theory was not to rely on the current mathematical methods.

At the time of his Zürich lecture Hilbert tended to restrict the methods of proof-theoretic reasoning to the most primitive evidence. The apparent needs of proof theory induced him to adopt successively those suppositions which constitute what he then called the "*finite Einstellung.*"

CONSISTENCY. In his first publication on proof theory, "Neubegründung der Mathematik, Erste Mitteilung," Hilbert explains how number theory can be treated in a finitist way, whereas mathematics in general transcends finitist methods. But, Hilbert argues, we can regain an elementary kind of mathematical objectivity by formalizing the statements and proofs, using the methods of sym-

bolic logic, and by taking the representing formulas and proofs directly as objects. In the same paper Hilbert also gives indications of the nature of formalization and presents an instance of a proof of consistency—as yet for only a very restricted system.

A more advanced stage is reached in Hilbert's lecture at the Leipzig congress of the Deutsche Naturforscher Gesellschaft in 1922, "Die logische Grundlagen der Mathematik."

In this speech the method is presented of dealing in proof theory with the logical forms of generality and existence (quantifiers) by means of a logical choice function which assigns to any predicate A an object τA for which A holds only if it is generally satisfied. This idea is formally expressed by the "transfinite axiom," $A(\tau A) \rightarrow A(a)$, in which a predicate expression can be substituted for A and any term representing an individual can be substituted for a. A slight modification, soon applied, replaced the function τA by the function ϵA, dual to it, whose axiom is $A(a) \rightarrow A(\epsilon A)$.

By means of the choice function the quantifiers can be eliminated from a formalized proof in such a way that the rules for the use of "all" and "exists" are reduced to applications of the transfinite axiom, so that the explicit logical structure of the proof becomes transformed into an elementary one, consisting only in applications of the propositional calculus and substitutions.

The task of proving the consistency of a formalized domain of arithmetic is thus essentially reduced. This task—in virtue of the law "ex falso quodlibet"—amounts to showing that the formula $0 \neq 0$ cannot be derived in the domain; in other words, to showing that in any formal derivation of the formalized domain having a numerical end formula, this end formula differs from the formula $0 \neq 0$. Consideration of formalized proofs can now be restricted to those obtained by the transformation using the function ϵA. The main problem is then to eliminate the formulas resulting from the transfinite axiom by substitution (the "critical formulas").

The method that Hilbert indicates for attacking this problem consists—after first removing the free variables, which is possible—of a sequence of steps. In each step the terms that occur are replaced by numerical values. Then, either all critical formulas turn into true numerical formulas, and the attempted elimination is effected, or the result of the step determines a next step. It must still be shown that the process has an end, and this, at least in the simple cases, can be seen to hold.

This method is not in principle restricted to cases where the predicates to which the logical choice function applies are number predicates and where the individuals are therefore natural numbers; it can also be used for individuals of higher types. The particular case in which number functions are taken as individuals is essential to the formalization of the theory of real numbers. In the Leipzig lecture, Hilbert gave several indications of how this formalization can be performed; in particular, he showed how some form of the Zermelo choice principle (used in the theory of functions of real numbers) can be derived from the transfinite axiom related to the type of real numbers (as individuals).

Thus, it seemed that carrying out proof theory was only a question of mathematical technique. Such an expectation, however, turned out to be illusory. An indication was that the first substantial consistency proof following Hilbert's scheme of reasoning by Wilhelm Ackermann (in his thesis, "Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit") required an essential restriction of the formal system not envisaged in the original plan. Similarly, in John von Neumann's inquiry "Zur Hilbertschen Beweistheorie," where a formal system for the logic of first and second order (including the first four Peano axioms) was set up and a consistency proof using Hilbert's method was given, the consistency proof did not apply to the full system but excluded the comprehension axiom, which provides the manipulation of substitutions for variables of second type. Thus, two highly able investigators did not succeed in obtaining a consistency proof for a formal system of the theory of real numbers by means of the above-mentioned Hilbert method (connected with the logical choice function) of eliminating the critical formulas.

A second method of eliminating the critical formulas, devised by Hilbert and elaborated by Ackermann, yields the proof of a general theorem which states that any axiomatic system, formalized within the frame of standard logic (that is, propositional logic and the rules governing quantifiers), whose axioms have a finitist interpretation is consistent (see Hilbert and Bernays, *Grundlagen der Mathematik,* Vol. II, Sec. I, esp. pp. 18–38). The method is one of the easiest for proving an important theorem of mathematical logic (first stated by Jacques Herbrand in his doctoral dissertation) that yields a kind of normal form for derivations in pure logic and which also can be applied to decision problems. But this method is not sufficient to demonstrate the consistency of the

proper formal system of number theory and therefore is the less sufficient for the systems of infinitesimal analysis.

COMPLETENESS. Ackermann revised and simplified the proof presented in his thesis. It was thought that by this modified proof and by that of von Neumann the consistency of formalized number theory, at least, had been proved. Such was the situation when Hilbert presented, at the International Congress of Mathematicians in Bologna in 1928, his "Probleme der Grundlegung der Mathematik." To the problem of proving consistency he here added two problems of completeness: the problem of showing that every universally valid logical schema is derivable by the rules of the predicate calculus and the problem of showing the completeness of formalized number theory, in the sense that the formal system of number theory contains no formula which, together with its negation, can be shown to be underivable in the system.

GÖDEL'S RESULTS. Kurt Gödel soon took up both these problems of completeness, but he stated completeness only for the case of the predicate calculus (first-order functional calculus), whereas he proved the incompleteness of formalized number theory even in the strong sense that no strictly formal system is possible in which each true number-theoretic proposition is derivable. At the same time Gödel proved a theorem from which it follows that a finitist proof of consistency for a formal system strong enough to formalize all finitist reasonings is impossible ("Über formale unentscheidbare Sätze der Principia Mathematica und verwandte Systeme I"). Von Neumann was convinced that this last condition holds for the formal system of number theory, and hence he inferred that Gödel's result implies the impossibility of a finitist consistency proof not only for the broader systems discussed by Gödel but even for the formal system of number theory.

To corroborate this inference he was able to show that in the proof of consistency of the formal system of number theory by the elimination of critical formulas, the demonstration that the process of elimination has an end did not apply in full generality (see Hilbert and Bernays, *Grundlagen,* Vol. II, pp. 123–125). It thus became clear that in two respects Hilbert's program demanded more than can be fulfilled: Mathematical theories cannot be formalized with full adequacy, and consistency proofs cannot be strictly finitist in the essential cases.

BROADENING OF PROOF THEORY. It soon became apparent that proof theory could be fruitfully developed without fully keeping to the original program. It was discovered that a proof of consistency for the formal system of number theory, although not a finitist one, is possible by methods of proof admitted by L. E. J. Brouwer's intuitionism.

Arend Heyting, in two papers of 1930, set up a formal system of intuitionistic number theory. And, as Gödel and Gerhard Gentzen independently observed, there is a relatively simple method of showing that any contradiction derivable in the formal system of classical number theory would entail a contradiction in Heyting's system. Hence, from the consistency of Heyting's system the consistency of the classical system follows (Kurt Gödel, "Zur intuitionistischen Arithmetik und Zahlentheorie"—Gentzen withdrew his own paper, already in print, because of the appearance of Gödel's paper).

In this way it appeared that intuitionistic reasoning is not identical with finitist reasoning, contrary to the prevailing views at that time. In particular, intuitionistic reasoning deals with concepts not admitted as methods in finitist proofs, such as the quite general concept of consequence when it is not delimited by any rules of proof. It thus became apparent that the "*finite Standpunkt*" is not the only alternative to classical ways of reasoning and is not necessarily implied by the idea of proof theory. An enlarging of the methods of proof theory was therefore suggested: Instead of a restriction to finitist methods of reasoning, it was required only that the arguments be of a constructive character, allowing us to deal with more general forms of inference.

By this modification of the program, various proofs of consistency for the formal system of number theory were obtained, the first by Gentzen ("Die Widerspruchsfreiheit der reinen Zahlentheorie," "Die gegenwärtige Lage in der mathematischen Grundlagenforschung," and "Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie"). Ackermann was then able to complete the consistency proof proceeding by the method of eliminating the critical formulas ("Zur Widerspruchsfreiheit der Zahlentheorie"). The broadened methods also permitted a loosening of the requirements of formalizing. One step in this direction, made by Hilbert himself, was to replace the schema of complete induction by the stronger rule later called infinite induction ("Die Grundlegung der elementaren Zahlenlehre" and "Beweis des Tertium non datur").

However, going beyond finitist methods is not generally required in proof theory; many important results have been obtained by finitist methods, results concerning the following topics: pure logic, the combinatorial

calculus, completeness (the completeness of a system of real algebra), undecidability, and relative consistency.

One main result concerning relative consistency is connected with Hilbert's attempt at a positive solution of Cantor's continuum problem in the paper "Über das Unendliche." The sketch of a proof given in this work contains many detailed arguments, and it stimulated the study of recursive definitions. As a whole, however, the sketch was fragmentary, and there were in principle hindrances to its completion. Twelve years later Gödel connected the ideas of Hilbert's paper with the concepts of axiomatic set theory and proved the consistency of Cantor's continuum hypothesis in its generalized form on the assumption that axiomatic set theory (without the axiom of choice) is consistent. (Nevertheless, this result, which is obtained by a powerful method of proof, does not settle the continuum problem. In fact, from results obtained by Paul Cohen it appears that axiomatic set theory, at least in its formal delimitation, leaves this problem fully undecided.)

On the whole, Hilbert's idea of making mathematical proof an object of mathematical research by means of formalization has proved to be very fruitful. And although Hilbert's work in the foundations of arithmetic has not had the effect he sought, "to remove once and for all the questions of foundations in mathematics" ("die Grundlagenfragen in der Mathematik als solche endgültig aus der Welt zu schaffen"—"Die Grundlagen der Mathematik," p. 65, and "Die Grundlagen der elementaren Zahlenlehre," p. 489), he did establish proof theory as a valuable domain of mathematical investigation, and thus Hilbert was a pioneer in the newer mathematical foundation theory, as he was in many other fields of mathematics.

See also Brouwer, Luitzen Egbertus Jan; Cantor, Georg; Continuity; Einstein, Albert; Frege, Gottlob; Geometry; Gödel, Kurt; Helmholtz, Hermann Ludwig von; Logic, History of: Modern Logic; Mathematics, Foundations of; Neumann, John von; Pascal, Blaise; Proof Theory; Russell, Bertrand Arthur William; Set Theory; Weyl, (Claus Hugo) Hermann; Alfred North Whitehead.

Bibliography

WORKS BY HILBERT

On the Foundations of Geometry

Die Grundlagen der Geometrie. Leipzig: Teubner, 1899; 9th ed., Stuttgart, 1961. The first edition was published as a *Festschrift* for the unveiling of the Gauss-Weber memorial. The second and later editions were published as textbooks. Most incorporate modifications and additions. Translated by E. J. Townsend as *Foundations of Geometry.* La Salle, IL: Open Court, 1902.

"Über den Satz von der Gleichheit der Basiswinkel im gleichschenkligen Dreieck." *Proceedings of the London Mathematical Society* 35 (1903): 50–68. Reprinted as Appendix II to the *Grundlagen,* 9th ed., 133–158.

"Über die Grundlagen der Geometrie." *Mathematische Annalen* 56 (1902): 381–422. Reprinted as Appendix IV to the *Grundlagen,* 9th ed., 178–230.

"Neue Begründung der Bolyai-Lobatschefskyschen Geometrie." *Mathematische Annalen* 57 (1903): 137–150. Reprinted as Appendix III to the *Grundlagen,* 9th ed., 159–177.

On the Foundations of Arithmetic

"Die Theorie der algebraischen Zahlkörper." *Jahresbericht der Deutschen Mathematikervereinigung* 4 (1897), I–XVIII, 175–546.

"Über den Zahlbegriff." *Jahresbericht der Deutschen Mathematikervereinigung* 8 (1900): 180–184. Reprinted as Appendix VI to the *Grundlagen,* 7th ed., 241–246. Leipzig, 1930.

"Über die Grundlagen der Logik und Arithmetik." In *Verhandlungen des Dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904,* 174–185. Leipzig, 1905. Reprinted as Appendix VII to the *Grundlagen,* 7th ed., 247–261.

"Axiomatisches Denken." *Mathematische Annalen* 78 (1918): 405–415.

"Neubegründung der Mathematik. Erste Mitteilung." *Abhandlungen aus dem mathematische Seminar der Hamburgischen Universität* 1 (1922): 157–177.

"Die logischen Grundlagen der Mathematik." *Mathematische Annalen* 88 (1923): 151–165.

"Über das Unendliche." *Mathematische Annalen* 95 (1926): 161–190.

"Die Grundlagen der Mathematik." *Abhandlungen aus dem mathematische Seminar der Hamburgischen Universität* 6 (1928): 65–85.

"Probleme der Grundlegung der Mathematik." *Mathematische Annalen* 103 (1929): 1–9.

"Die Grundlegung der elementaren Zahlenlehre." *Mathematische Annalen* 104 (1931): 485–494.

"Beweis des Tertium non datur." *Nachrichten des Gesellschaft der Wissenschaft zu Göttingen* (1931): 120–125.

Textbooks

Grundzüge der theoretischen Logik. With W. Ackermann. Berlin: Springer, 1928; 3rd ed., 1949. Translated by Lewis M. Hammond and others as *Principles of Mathematical Logic.* New York: Chelsea, 1950.

Grundlagen der Mathematik. 2 vols. With Paul Bernays. Berlin: Springer, 1934–1939.

Collected Works

Gesammelte Abhandlungen. 3 vols. Edited by Ernst Hellinger, Paul Bernays, and Otto Blumenthal. Berlin: Springer, 1932–1935.

STUDIES RELATED TO HILBERT'S WORK

Bernays, Paul. "Die Bedeutung Hilberts für die Philosophie der Mathematik." *Die Naturwissenschaften,* 10th year (1922),

No. 1. This number was devoted to Hilbert on his sixtieth birthday. It contains the following papers: Otto Blumenthal, "David Hilbert"; Otto Toeplitz, "Der Algebraiker Hilbert"; Max Dehn, "Hilberts geometrisches Werk"; Richard Courant, "Hilbert als Analytiker"; Max Born, "Hilbert und die Physik"; and a list of Hilbert's publications up to 1921 compiled by Karl Siegel.

Bernays, Paul. "Die Philosophie der Mathematik und die Hilbertsche Beweistheorie." *Blätter für deutsche Philosophie. Zeitschrift der Deutschen Philosophischen Gesellschaft* (Berlin) 4 (3–4) (1930): 326–367. This number also contains articles on the foundations of mathematics by Heinrich Scholz, A. Fraenkel, Rudolf Carnap, Carl Menger, and Walter Dubislav.

Bernays, Paul. "Sur le platonismc dans les mathématiques." *L'enseignement mathématique* 34 (1935): 52–60.

Bernays, Paul. "Über Hilberts Gedanken zur Grundlegung der Arithmetik." *Jahresbericht der Deutschen Mathematikervereinigung* 31 (1922): 10–19.

Bernays, Paul. "Hilberts Untersuchungen über die Grundlagen der Arithmetik." In *Gesammelte Abhandlungen,* by David Hilbert, Vol. III, 191–216. Berlin: Springer, 1935.

Blumenthal, Otto. "Lebensgeschichte," in *Gesammelte Abhandlungen,* by David Hilbert (see above), Vol. III, 388–429.

Kreisel, Georg. "Hilbert's Programme." *Dialectica* 88 (1923): 151–165. Reprinted in *Philosophy of Mathematics,* edited by Paul Benacerraf and Hilary Putnam, 157–180. Englewood Cliffs, NJ: Prentice-Hall, 1964.

On the Foundations of Geometry

Baldus, Richard. "Zur Axiomatik der Geometrie III. Über das Archimedische und das Cantorsche Axiom." In *Sitzungsbericht des Heidelberger Akademie des Wissenschaft,* 5. Abhandlung, 3–12. Leipzig, 1930.

Bernays, Paul. "Bemerkungen zu den Grundlagen der Geometrie." In *Studies and Essays Presented to R. Courant on His 60th Birthday,* 29–44. New York: Interscience, 1948.

Dehn, Max. "Über den Rauminhalt." *Mathematische Annalen* 55 (1902): 465–478.

Dehn, Max. "Über raumgleiche Polyeder." *Göttinger Nachrichten* (1900).

Enriques, Federigo. "Prinzipien der Geometrie." In *Enzyclopädie des mathematischen Wissenschaften,* Vol. III, Pt. I, 34ff. 1907–1910. See especially p. 36.

Freudenthal, Hans. "Die Grundlagen der Geometrie um die Wende des 19. Jahrhunderts." *Mathematisch-Physikalische Semesterberichte* (Göttingen) 7 (1) (1960): 2–25. A translated extract of this article appeared as "The Main Trends of the Foundations of Geometry in the 19th Century," in *Logic, Methodology and Philosophy of Science,* edited by Ernest Nagel et al. Stanford, CA: Stanford University Press, 1962.

Freudenthal, Hans. "Im Urkreis der sogennannten Raumprobeme." In *Essays on the Foundation of Mathematics,* edited by Yehoshua Bar-Hillel, 322–327. Jerusalem, 1961.

Freudenthal, Hans. "Neuere Fassung des Riemann-Helmholtz-Lieschen Raumproblems." *Mathematische Zeitschrift* 63 (1955–1956): 374–405. See in particular the introduction.

Freudenthal, Hans. "Zur Geschichte der Grundlagen der Geometrie." *Nieue Archief voor Wiskunde* 4 (1957): 105–142. Includes a review of the 8th edition of Hilbert's *Grundlagen der Geometrie.*

Hessenberg, Gerhard. "Beweis des Desarguesschen Satzes aus dem Pascalschen." *Mathematische Annalen* 61: 161–172.

Kagan, B. "Über die Transformation der Polyeder." *Mathematische Annalen* 57 (1903): 421–424.

Schatunowski, S. O. "Über den Rauminhalt der Polyeder." *Mathematische Annalen* 57 (1903): 496–508.

Schmidt, Arnold. "Die Herleitung der Spiegelung aus der ebenen Bewegung." *Mathematische Annalen* 109 (1934): 538–571.

Wiener, Hans. "Grundlagen und Aufbau der Geometrie." *Jahresberichte der Deutschen Mathematikervereinigung* 1 (1892): 45–48.

Important Papers on Proof Theory

Ackermann, Wilhelm. "Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit." *Mathematische Annalen* 93 (1924–1925): 1–36.

Ackermann, Wilhelm. "Zur Widerspruchsfreiheit der Zahlentheorie." *Mathematische Annalen* 117 (1940): 162–194.

Cohen, Paul J. "The Independence of the Continuum Hypothesis," I and II. *Proceedings of the National Academy of Sciences* 50 (1963): 1143–1148, and 51 (1964): 105–110.

Gentzen, Gerhard. "Die gegenwärtige Lage in der mathematischen Grundlagenforschung" and "Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie." *Forschungen zur Logik zur Grundlegung der exakten Wissenschaften,* n.s., (4) (1938): 5–18, 19–44.

Gentzen, Gerhard. "Untersuchungen über das logische Schliessen," I and II. *Mathematische Zeitschrift* 39 (1934): 176–210, 405–431.

Gentzen, Gerhard. "Die Widerspruchsfreiheit der reinen Zahlentheorie." *Mathematische Annalen* 112 (1936): 493–565.

Gödel, Kurt. *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory.* Princeton, NJ: Princeton University Press, 1940.

Gödel, Kurt. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I." *Monatschefte für Mathematik und Physik* 38 (1931): 173–198.

Gödel, Kurt. "Zur intuitionistischen Arithmetik und Zahlentheorie." *Ergebnisse eines mathematische Kolloquiums, 52 Kolloquium* (June 1932), 4 (1933): 34–38.

Herbrand, Jacques. "Recherches sur la théorie de la démonstration." *Travaux de la Société des Sciences et des Lettres de Varsovie,* Classe III, (33) (1930): 33–160. Ph.D. diss., University of Paris.

Heyting, Arend. "Die formalen Regeln der intuitionistischen Logik." *Sitzungsberichte der Preussischen Akademie der Wissenschaften,* Physical-mathematical section, 1930 II. Berlin, 1930.

Heyting, Arend. "Die formalen Regeln der intuitionistischen Mathematik." *Sitzungsberichte der Preussischen Akademie der Wissenschaften,* Physical-mathematical section, 1930 II. Berlin, 1930.

Neumann, John von. "Zur Hilbertschen Beweistheorie." *Mathematische Zeitschrift* 26 (1927): 1–46.

Paul Bernays (1967)

HILBERT, DAVID [ADDENDUM]

Bernays's entry on Hilbert still reads, after forty years, as a wonderful account of the essential contributions by Hilbert to the foundations of geometry and proof theory. However, recent developments have substantially increased the understanding of Hilbert's original investigations and pushed these investigations further. The following bibliography will help the reader navigate among the most important recent contributions. It is divided into four parts: (1) contributions to Hilbert's biography and mathematical work emerging from Hilbert's famous list of problems given in Paris in 1900; (2) historical work related to Hilbert's foundational views; (3) logico-foundational and philosophical developments related to Hilbert's program; and (4) ongoing work of publication of Hilbert's and Bernays's work.

See also Mathematics, Foundations of.

Bibliography

BIOGRAPHY AND MATHEMATICAL DEVELOPMENTS EMERGING FROM HILBERT'S PROBLEMS

The only book-length biography available is Constance Reid's *Hilbert* (New York: Springer, 1970). For mathematical developments arising from Hilbert's list of 23 problems in Paris, see Jeremy Gray's *The Hilbert Challenge* (Oxford: Oxford University Press, 2000) and Benjamin H. Yandell's *The Honors Class: Hilbert's Problems and Their Solvers* (Natick, MA: A. K. Peters, 2002).

HISTORICAL STUDIES OF HILBERT'S FOUNDATIONAL WORK

There has been a substantial amount of work in the study of the development of Hilbert's foundational views. A novelty in this area is the study of his work on the foundations of physics. Most of this scholarship is characterized by the extensive use of the Hilbert Nachlaß in Göttingen. The work has been divided below by periods.

Foundations of Geometry (1891–1899)

Toepell, Michael-Markus. *Über die Entstehung von Hilberts "Grundlagen der Geometrie."* Göttingen, Germany: Vandenhoeck und Ruprecht, 1986.

Axiomatizations of Physical Theories (1895–1917)

Corry, Leo. "David Hilbert and the Axiomatization of Physics." *Archive for History of Exact Sciences* 51 (1997): 83–198.

Corry, Leo. "Hilbert and Physics (1900–1915)." In *The Symbolic Universe: Geometry and Physics (1890–1930)*, edited by Jeremy Gray, 145–187. New York, Oxford University Press, 1999.

Majer, Ulrich. "The Axiomatic Method and the Foundations of Science: Historical Roots of Mathematical Physics in Göttingen (1900–1930)." In *John von Neumann and the Foundations of Quantum Physics*, edited by Miklos Redi and Michael Stoeltzner, 11–33. Dordrecht, Netherlands: Kluwer, 2001.

Sauer, Tilman. "The Relativity of Discovery: Hilbert's First Note on the Foundations of Physics." *Archive for History of Exact Sciences* 53 (1999): 529–575.

From Axiomatics to Proof Theory (1900–1922)

Peckhaus, Volker. *Hilbertprogramm und kritische Philosophie*. Göttingen, Germany: Vandenhoeck und Ruprecht, 1990.

Mancosu, Paolo. "Between Russell and Hilbert: Behmann on the Foundations of Mathematics." *The Bulletin of Symbolic Logic* 5 (3) (1999): 303–330.

Sieg, Wilfried. "Hilbert's Programs: 1917–1922." *The Bulletin of Symbolic Logic* 5 (1999): 1–44.

Mancosu, Paolo. "The Russellian Influence on Hilbert and His School." *Synthese* 137 (2003): 59–101.

Development of Logic in Hilbert's School (1905–1928)

Hallett, Michael. "Hilbert and Logic." In *Québec Studies in the Philosophy of Science*, pt. 1, vol. 177, edited by Mathieu Marion and Robert Cohen, 135–187. Dordrecht, Netherlands: Kluwer, 1995.

Peckhaus, Volker. "Hilberts Logik. Von der Axiomatik zur Beweistheorie." *Intern. Zs. f. Gesch. u. Ethik der Naturwiss. Tech. u. Med.* 3 (1995): 65–86.

Zach, Richard. "Completeness Before Post: Bernays, Hilbert, and the Development of Propositional Logic." *The Bulletin of Symbolic Logic* 5 (3) (1999): 331–366.

Moore, Gregory "Hilbert and the Emergence of Modern Mathematical Logic." *Theoria* 12 (1) (1997): 65–90.

Avigad, Jeremy, and Richard Zach. "The Epsilon Calculus." *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/epsilon-calculus/.

Proof Theory and Finitism Until Gödel's Theorems (1922–1931)

Majer, Ulrich. "Hilberts Methode der idealen Elemente und Kants regulativer Gebrauch der Ideen." *Kant-Studien* 84 (1993): 51–77.

Mancosu, Paolo. "Hilbert and Bernays on Metamathematics." In *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*, pp. 149–188. New York: Oxford University Press, 1988.

Zach, Richard. 2003 "The Practice of Finitism: Epsilon Calculus and Consistency Proofs in Hilbert's Program." *Synthese* 137 (2003): 211–259.

LOGICO-FOUNDATIONAL AND PHILOSOPHICAL CONTRIBUTIONS (1970–2003)

Logico-Foundational Developments

The best starting point are the articles from "A Symposium on Hilbert's Program":

Feferman, Solomon. "Hilbert's Program Relativized: Proof-Theoretical and Foundational Reductions." *Journal of Symbolic Logic* 53 (1988): 364–383.

Simpson, Steven. "Partial Realizations of Hilbert's Program." *Journal of Symbolic Logic* 53 (1988): 349–3631.

Sieg, Wilfried. "Hilbert's Program Sixty Years Later." *Journal of Symbolic Logic* 53 (1988): 338–348.

For further references see Richard Zach's "Hilbert's Program" in the *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/hilbert-program/.

Philosophical Contributions

Kitcher, Philip. “Hilbert’s Epistemology.” *Philosophy of Science* 43 (1976): 99–115.

Tait, William. “Finitism.” *Journal of Philosophy* 78 (1981): 524–546.

Detlefsen, Michael. *Hilbert’s Programme: An Essay on Mathematical Instrumentalism*. Dordrecht, Netherlands: Reidel, 1986.

Hallett, Michael. “Physicalism, Reductionism and Hilbert.” In *Physicalism in Mathematics*, edited by Andrew D. Irvine, 183–257. Dordrecht, Netherlands: Kluwer, 1990.

Parsons, Charles. “Finitism and Intuitive Knowledge.” In *The Philosophy of Mathematics Today*, edited by Matthias Schirn, 249–270. Oxford: Oxford University Press, 1998.

TRANSLATIONS AND EDITIONS OF HILBERT AND BERNAYS

Van Hejenoort 1967, Ewald 1996, and Mancosu 1998 provide an extensive coverage of translations into English of most of Hilbert’s and Bernays’s published articles on the foundations of mathematics for the period 1900–1931:

Ewald, William. *From Kant to Hilbert: Readings in the Philosophy of Mathematics*. Oxford: Oxford University Press, 1996.

Mancosu, Paolo., ed. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. New York: Oxford University Press, 1998.

van Heijenoort, Jean. 1967 *From Frege to Gödel*. Cambridge, MA: Harvard University Press, 1967.

In addition there are two scholarly editions in the making. The *Hibert Edition* (six volumes, Springer Verlag) includes a selection of the original unpublished lecture notes (in German) preserved at the University of Göttingen. The first volume, edited by Michael Hallett, appeared in 2004. There is also in preparation an edition of Bernays’s foundational writings in English, a description of which is available at http://www.phil.cmu.edu/projects/bernays/. This edition is scheduled to appear some time after 2004 for Open Court.

Paolo Mancosu (2005)

HILDEGARD OF BINGEN

(1098–1179)

Hildegard of Bingen, the first German mystic, wrote profusely as a prophet, poet, dramatist, musician, physician, and political moralist. She was an extraordinary woman who exerted a tremendous temporal and spiritual influence on her time and who has been rediscovered since the 1960s.

Hildegard was born in Bockelheim, the diocese of Mainz, on the Nahe River. Her father, Hildebert, was a knight in the service of Count Meginhard of Spanheim. At six, she began to have religious visions that continued the rest of her life. At eight, she was entrusted to the care of Jutta, sister of Count Meginhard. The two lived in a small cottage adjoining the church abbey at Disibodenberg. A sickly child, Hildegard continued her education under Jutta, learning to read and sing Latin. At fifteen, she was clothed in the habit of a nun in Jutta’s hermitage, a community following the Rule of St. Benedict. At thirty-eight, Hildegard became the abbess of the community.

Eventually, the archbishop of Mainz examined her visions with his theologians and ruled them divinely inspired, ordering Hildegard to record them in writing. From 1141 until 1151, she worked on her principal work, *Scivias* (*May You Know*, or *Know the Ways*). In 1147, Pope Eugenius III and his commission examined her visions and also authorized her to write whatever the Holy Spirit inspired her to write. Her growing fame then caused Hildegard to transfer her convent from Disibodenberg to Rupertsberg, near Bingen, between 1147 and 1150. She continued living there until her death on September 17, 1179. She was buried in her convent church, where her relics remained until 1632, when the convent was destroyed by the Swedes and her relics moved to Eibingen.

A woman of an extraordinarily energetic and independent mind, Hildegard wrote voluminously. *Scivias*, the first of her three mystical works, develops her view on the universe, on the theory of macrocosm and microcosm, the structure of humans, birth, death, and the nature of the soul. It also treats the relations between God and humans in creation, the redemption, and the church. The last of the twenty-six visions of *Scivias* contains *Ordo Virtutum*, the earliest liturgical morality play.

Liber Vitae Meritorum (*The Book of the Rewards of Life*, 1158–1163) studies the weaknesses separating us from God. It is one of the most subtle, psychologically fascinating, and intense works ever written on the relationship of various sins to their corresponding virtues.

Liber Divinorum Operum Simplicis Hominis (*The Book of the Divine Works of a Simple Man*, 1163–1173), the third of Hildegard’s mystical books, concerns itself with the unity of creation. Hildegard succeeds in synthesizing into one great whole her theological beliefs along with her knowledge of the elements of the universe and the structures within the human body. This work is often considered the epitome of science of her time.

Besides her three mystical books, Hildegard wrote a long physical treatise titled *Physica: Subtilitatum Diversarum Naturarum Creaturarum* (*Physical Things: Of the Simplicities of Various Natural Creatures*, 2001) and her book of medicine titled *Causae et Curae* (*Causes and Cures*, 1903). Although her theoretical knowledge of

medicine seems crude today, large numbers of sick and suffering persons were brought to her for cures. A thriving clinic in Konstanz, Germany, practices Hildegard's remedies today.

In addition, Hildegard wrote *Vita Sancti Disibodi* (*The Life of Saint Disibod*) and *Vita Sancti Ruperti* (*The Life of Saint Rupert*). Her *Solutiones Triginta Octo Quaestionum* (*Answers to Thirty-eight Questions*) comments on various theological and scriptural subjects. Her *Explanatio Symboli Athanasii* (*Explanation of the Symbol of Saint Athanasius*) and *Explanatio Regulae Sancti Benedicti* (*Explanation of the Rule of Saint Benedict*), written at the request of the Benedictine monastery of Huy in Belgium, are self-explanatory.

For the nuns of her convent, Hildegard wrote hymns and canticles—both words and music. She collected her songs into a cycle titled *Symphonia Armonie Celestium Revelationum* (*The Symphony of the Harmony of Heavenly Revelations*). These approximately seventy songs were written for a wide range of liturgical celebrations.

Finally, Hildegard wrote letters to popes, cardinals, bishops, abbots, kings and emperors, monks and nuns, men and women of various social levels both in Germany and abroad. Some of her letters are more personal, but the majority are mystical treatises, prophecies, sermons, and strong exhortations concerning various corruptions. Hildegard's clear intelligence foresaw that the ecclesiastical and political abuses of her time would ultimately burst into flames in some event such as the eventual Reformation or the Thirty Years' War. Hildegard represented a legacy to her own times, and now has been rediscovered in ours.

See also Macrocosm and Microcosm; Mysticism, History of; Women in the History of Philosophy.

Selected Bibliography

PRIMARY WORKS

Barth, Pudentiana, Immaculata M. Ritscher, and Joseph Schmidt-Gorg, eds. *Lieder: Nach den Handschrfiten herasugegeben*. Salzburg: 1969.

Fuhrkotter, Adelgundis, ed. and tr. *Briefwechsel: Nach den altesten Handschriften ubersetzt und nach den Quellen erlautert*. Salzburg: 1965.

Fuhrkotter, Adelgundis, and Angela Carlevaris, eds. *Scivias*. Corpus Christianorum, Continuatio Medievalis 43. 2 vols. Turnhout: 1978.

Heilkunde: Das Buch von dem Grund und Wesen und der Heilung der Krankheiten; nach den Quellen ubersetzt und erlautert. Translated by Heinrich Schipperges. Salzburg: 1957.

Kaiser, Paul, ed. *Hildegardis Causae et Curae*. Leipzig: 1903.

Der Mensch in der Verantwortung: Das Buck der Lebensverdienste (Liber vitae meritorum): nach den Quellen ubersetzt und erlautert. Translated by Heinrich Schipperges. Salzburg: 1972.

Migne, J. P., ed. *S. Hildegardis abbatissae Opera omnia*. Patrologiae cursus completus, Ser. Lat. 197. Paris: 1882.

Naturkunde: Das Buck von dem inneren Wesen der verschiedenen Naturren in der Schopfung: nach den Quellen ubersetzt und erlautert. Translated by Peter Riethe. Salzburg: 1959.

Pitra, Joannes Baptista, ed. *Analecta Sanctae Hildegardis Opera Spicilegio Solesmensi Parata*. Analecta Sacra 8. 1882. Reprint. Farnborough, England: 1966.

Van Acker, Lieven, and Klaes-Hachmoller. *Epistolarioum*. Corpus Christianorum, Continuatio Medievalis 91, 91a, 91b. Turnhout: 1991, 1993, 2001.

Welt un Mensch: Das Buch "De operatione Dei": aus dem Genter Kodex ubersetzt unde erlautert. Translated by Heinrich Schipperges. Salzburg: 1965.

Wisse die Wege: Nach dem Orginialtext des illuminierten Rupertsberger Kodex der Wiesbaden Landesbibliotheck in Deutsche ubertragen und bearbeitet. Translated by Maura Bockecler. Salzburg: 1954.

PRINTED BIBLIOGRAPHIES

Lauter, Werner. *Hildegard-Bibliographie: Wegweiser zur Hidegard-Literatur*. Alzey: 1970.

Lauter, Werner. *Hildegard-Bibliographie: Wegweiser zur Hildegard-Literatur. Band II 1970–82*. Alzey: 1984.

FURTHER REFERENCES

Beer, Frances. *Women and Mystical Experience in the Middle Ages*. Suffolk, U.K.: Boydell, 1992.

Burnett, Charles, and Peter Dronke, eds. *Hildegard of Bingen: The Context of Her Thought and Art*. London: Warburg Institute, 1998.

Craine, Renate. *Hildegard: Prophet of the Cosmic Christ*. New York: Crossroad, 1997.

Davidson, Audrey Ekdahl, ed. *The "Ordo Virtutum" of Hildegard of Bingen: Critical Studies*. Kalamazoo, MI: Medieval Institute Publications, 1992.

Dronke, Peter. "The Composition of Hildegard of Bingen's *Symphonia*." *Sacris Erudiri* 19 (1969–1970): 381–393.

Dronke, Peter. *Women Writers of the Middle Ages*. Cambridge, U.K.: Cambridge University Press, 1984.

Escot, Pozzi. "Gothic Cathedral and Hidden Geometry of St. Hildegard." *Sonus* 5 (1) (1984): 14–31.

Escot, Pozzi. "Universal Proportion in Hildegard von Bingen." *Sonus* 11 (1990): 33–40.

Flanagan, Sabina. *Hildegard of Bingen, 1098–1179: A Visionary Life*. 2nd ed. New York: Routledge, 1998.

Fox, Matthew. *Hildegard of Bingen's* Book of Divine Works *with Letters and Songs*. Sante Fe, NM: Bear, 1987.

Gies, Frances, and Joseph Gies. *Women in the Middle Ages*. New York: Perennial Library, 1986.

Grant, Barbara L. "A Feather on the Breath of God." *Parabola* 9 (2) (1984): 94–96.

Grant, Barbara L. "Five Liturgical Songs by Hildegard von Bingen." *Signs* 5 (3) (1980): 564–573.

Hozeski, Bruce W., tr. *Hildegard of Bingen: The Book of the Rewards of Life (Liber Vitae Meritorum)*. New York, Oxford University Press, 1994.
Hozeski, Bruce W., tr. *Hildegard's Healing Plants: From Her Medieval Classic* Physica. Boston: Beacon, 2001.
Hozeski, Bruce W., tr. *Scivias by Hildegard of Bingen*. Santa Fe, NM: Bear, 1986.
Jeskalian, Barbara J. "Hildegard of Bingen, Her Times and Her Music." *Anima* (Fall 1983): 7–13.
King-Lenzmeier, Anne. *Hildegard of Bingen: An Integrated Vision*. Collegeville, MN: Liturgical Press, 2001.
Kraft, Kent. "The German Visionary: Hildegard of Bingen." In *Medieval Women Writers*, edited by Katharina Wilson. Athens: University of Georgia Press, 1984.
McInerney, Maud Burnett, ed. *Hildegard of Bingen: A Book of Essays*. Garland Medieval Casebooks. New York: Garland, 1998.
Newman, Barbara. "Hildegard of Bingen: Visions and Validation." *Church History* 54 (2) (1985): 163–175.
Newman, Barbara. *Sister of Wisdom: St. Hildegard's Theology of the Feminine*. 2nd ed. Berkeley: University of California Press, 1997.
Newman, Barbara. *Voice of the Living Light: Hildegard of Bingen and Her World*. Berkeley: University of California Press, 1998.
Petroff, Elizabeth Alvilda. *Medieval Women's Visionary Literature*. Oxford: Oxford University Press, 1986.
Potter, Robert. "The *Ordo Virtutum*: Ancestor of the English Moralities?" *Comparative Drama* 20 (1986): 201–210.
Sacks, Oliver. "The Visions of Hildegard." In *The Man Who Mistook His Wife for a Hat*. Berkeley: University of California Press, 1986.
Schmitt, Miriam. "St. Hildegard of Bingen: Leven of God's Justice." *Cistercian Studies* 24 (1989): 69–88.
Scholz, Berhard W. "Hildegard von Bingen on the Nature of Woman." *The American Benedictine Review* 31 (1980): 361–383.
Steele, Francesca Maria. *The Life and Visions of St. Hildegarde*. St. Louis: B. Herder, 1915.

Bruce W. Hozeski (2005)

HINTIKKA, JAAKKO

(1929–)

The logician and philosopher Jaakko Hintikka was born in Vantaa, Finland. Receiving his doctorate from the University of Helsinki in 1956, he was a junior fellow at Harvard University from 1956 to 1959, a research professor at the Academy of Finland, and a professor of philosophy at the universities of Helsinki, Stanford, Florida State, and currently Boston University.

Hintikka developed semantical logical methods and uses them in philosophy. He advocates applying mathematical logic, especially model theory, in philosophy, most notably to questions in philosophy of language, but also to the study of Aristotle, Immanuel Kant, and Ludwig Wittgenstein. His main contributions in logic are those of model set, distributive normal form, possible-worlds semantics, and game-theoretic semantics.

A critical view of the Tarski truth definition led Hintikka to the concept of a model set as a more constructive approach to semantics. A model set has enough information to build a canonical term model in which sentences belonging to the set are true.

A model set is a set S of first-order formulas without identity (for simplicity), with negation in front of atomic formulas only, in a countable vocabulary, and containing possibly new individual constants, such that:

(1) No atomic sentence φ satisfies both $\varphi \in H$ and $\neg\varphi \in H$

(2) If $\varphi \wedge \psi \in H$, then $\varphi \in H$ and $\psi \in H$

(3) If $\varphi \vee \psi \in H$, then $\varphi \in H$ or $\psi \in H$

(4) If $\exists x\varphi(x) \in H$, then $\varphi(c) \in H$ for some constant c

(5) If $\forall x\varphi(x) \in H$, then $\varphi(c) \in H$ for all constants c occurring in H

A sentence has a model if and only if it is an element of a model set. Attempts to build a model set around the negation of a sentence form a tree, known as a semantic (or Beth) tableau. Infinite branches of this tree are model sets for $\neg\varphi$. If the tree has no infinite branches, it is finite and can be considered a proof of φ in the style of Jacques Herbrand and Gerhard Gentzen. Model sets came to play a central role in Hintikka's other work, such as distributive normal forms, possible-worlds semantics, and game-theoretic semantics.

Distributive normal forms, first introduced in monadic predicate logic by Georg Henrik von Wright, are defined as follows: Let $A^n_i(x_1, \ldots, x_n)$, $i \in K^n$ list all atomic formulas in a finite relational vocabulary (without identity, for simplicity), and the variables $x_1, \ldots, x_n$. If F is a formula, let $[F]^0 = F$ and $[F]^1 = \neg F$. Let $C^{0,n}_i(x_1, \ldots, x_n)$, $i \in I^{0,n}$ list all possible conjunctions $\bigwedge_j [A^n_j(x_1, \ldots, x_n)]^{\epsilon(j)}$ where ϵ runs through all functions $K^n \to \{0, 1\}$. Let $C^{m+1,n}_i(x_1, \ldots, x_n)$ $i \in I^{m+1,n}$ list all possible formulas

$$\bigwedge_{j \in J} \exists x_{n+1} C_j^{m,n+1}(x_1,\ldots,x_{n+1}) \wedge \forall x_{n+1}$$

$$\bigvee_{i \in J} C_j^{m,n+1}(x_1,\ldots,x_{n+1}),$$

where $J \subseteq I^{m,n+1}$.

If $a_1, \ldots, a_n$ satisfy $C_i^{m,n}(x_1, \ldots, x_n)$ in a model M and $b_1, \ldots, b_n$ satisfy $C_i^{m,n}(x_1, \ldots, x_n)$ in a model N, then $C_i^{m,n}(x_1, \ldots, x_n)$ codes a winning strategy for player 2 in the m-move Ehrenfeucht-Fraïssé game starting from the position $\{(a_1, b_1), \ldots, (a_n, b_n)\}$.

Every first-order sentence ϕ of quantifier rank m is logically equivalent to a unique disjunction of formulas of the form $C_i^{m,0}$. This disjunction is the distributive normal form of ϕ. The process of finding the distributive normal form of a given sentence cannot be made effective. Intuitively, one pushes quantifiers as deep into the formula as possible.

Distributive normal forms can be used to systematize definability theory, such as the Beth definability theorem, the Craig interpolation theorem, and the Svenonius theorem, and to systematize infinitary logic, emphasizing formal aspects more than the game-theoretic approach by Robert Vaught.

In the logic of induction Hintikka used distributive normal forms to give, in contrast to Rudolf Carnap, positive probabilities for universal generalizations. He developed a theory of surface information to support a thesis of the nontautological nature of logical inference, with applications to Kant's analytic-synthetic distinction.

Hintikka's formal definition of possible-worlds semantics, or model systems, for modal and epistemic logic is based on his concept of model set, unlike Saul Kripke's approach, which uses actual models as possible worlds.

A model system $(\mathfrak{C}, R)$ consists of a set $\mathfrak{C}$ of model sets and a binary alternativeness-relation R on $\mathfrak{C}$ such that:

(1) If $\Box\phi \in H \in \mathfrak{C}$, then $\phi \in H$.

(2) If $\Diamond\phi \in H \in \mathfrak{C}$, then there exists an alternative $H' \in \mathfrak{C}$ to H such that $\phi \in H'$.

(3) If $\Box\phi \in H \in \mathfrak{C}$ and $H' \in \mathfrak{C}$ is an alternative to H, then $\phi \in H'$.

A set S of formulas is defined to be satisfiable if there is a model system $(\mathfrak{C}, R)$ such that $S \subseteq H$ for some $H \in \mathfrak{C}$. A formula ϕ is valid if its negation is not satisfiable. Hintikka applied possible-worlds semantics to epistemic logic, deontic and modal logic, and the logic of perception and to the study of Aristotle and Kant. (See Hintikka [1969] for a summary of his theory of possible-worlds semantics. Hintikka's 1962 book is well-known outside of philosophy, most notably in the study of artificial intelligence and theoretical computer science.)

Game-theoretic semantics has its origin in Wittgenstein's language-games, Paul Lorenzen's dialogue games, Ehrenfeucht-Fraïssé games, and Leon Henkin's game theoretic interpretation of quantifiers. The semantic game of a sentence ϕ in a model M is a game between myself and nature about a formula ϕ and an assignment s. For $\phi = \phi_1 \wedge \phi_2$, nature chooses ϕ_i. For $\phi = \phi_1 \vee \phi_2$, I choose ϕ_i. Then we continue with ϕ_i and s. For $\phi = \forall x\psi(x)$, nature chooses s', which agrees with s outside x. For $\phi = \exists x\psi(x)$, I choose such s'. Then we continue with $\psi(x)$ and s'. For negation, we exchange roles. For ϕ atomic, the game ends. I win if s satisfies ϕ in M, otherwise nature wins.

Game-theoretic semantics became Hintikka's tool for analyzing natural language, particularly pronouns, conditionals, prepositions, definite descriptions, and the *de dicto* versus *de re* distinction and for challenging the approach of generative grammar. Sentences like "Every writer likes a book of his almost as much as every critic dislikes some book he has reviewed" led Hintikka to consider partially ordered quantifiers and eventually independence friendly (IF) logic (1996), with existential quantifiers $\exists x/y$, meaning that a value for x is chosen independently of what has been chosen for y. Thus, the semantic game of IF logic is a game of partial information.

IF logic is equal in expressive power to the existential fragment of second-order logic. The satisfiability of a sentence can still be analyzed in terms of model sets, but not provability. Wilfrid Hodges (1997) gave IF logic a compositional semantics in terms of sets of assignments, and Peter Cameron and Hodges (2001) proved it has no compositional semantics in terms of assignments only. Truth in various structures of mathematics can be reduced to logical consequence in IF logic, as in full second-order logic. IF logic has no negation and is not axiomatizable. This is countered by IF logic having a truth definition in IF logic.

See also Aristotle; Carnap, Rudolf; Model Theory; Philosophy of Language; Kant, Immanuel; Kripke, Saul; Logic, History of: Modern Logic; Modality, Philosophy and Metaphysics of; Modal Logic; Semantics; Semantics, History of; Tarski, Alfred; Wittgenstein, Ludwig Josef Johann; Wright, Georg Henrik von.

Bibliography

Cameron, Peter, and Wilfrid Hodges. "Some Combinatorics of Imperfect Information." *Journal of Symbolic Logic* 66 (2) (2001): 673–684.

Hintikka, Jaakko, and Merrill B. Hintikka. *Investigating Wittgenstein*. New York: Blackwell, 1986.

Hodges, Wilfrid. "Compositional Semantics for a Language of Imperfect Information." *Logic Journal of the IGPL* 5 (4) (1997): 539–563.

WORKS BY HINTIKKA

"Distributive Normal Forms in the Calculus of Predicates." *Acta Philosophica Fennica* 6 (1953).

Knowledge and Belief: An Introduction to the Logic of the Two Notions. Ithaca, NY: Cornell University Press, 1962.

Models for Modalities. Dordrecht, Netherlands: D. Reidel, 1969.

Logic, Language-Games, and Information: Kantian Themes in the Philosophy of Logic. Oxford, U.K.: Clarendon Press, 1973a.

Time and Necessity: Studies in Aristotle's Theory of Modality. New York: Oxford University Press, 1973b.

The Principles of Mathematics Revisited. New York: Cambridge University Press, 1996.

Selected Papers, Vols. 1–6. New York: Springer, 2005.

WORKS ABOUT HINTIKKA

Auxier, Randall E., and Lewis Hahn. *The Philosophy of Jaakko Hintikka*. Chicago: Open Court, 2005.

Jouko Väänänen (2005)

HIPPIAS OF ELIS

Hippias of Elis, the Greek Sophist and polymath, was probably born before 460 BCE. The date of his death is not known, but Plato speaks of him as one of the leading Sophists at the time of the death of Socrates in 399 BCE. On a number of occasions he acted as ambassador for his native city and also traveled widely, earning very large sums of money. He claimed to be a master of all the learning of his day, and his teaching and writings included elegies, tragedies, dithyrambs, historical works, literary discourses, epideictic speeches, discussions of astronomy, geometry, arithmetic, music, painting, sculpture, and ethics, and a technical system of mnemonics. None of his writings survives, but a reference in a papyrus book list of the third century suggests that at least one of his works survived until that date. Our knowledge of his teaching rests above all upon the picture of him given in Plato's dialogues, the *Hippias Major* (now generally accepted as written by Plato), the *Hippias Minor*, and the brief sketch in the *Protagoras*.

His polymathy invites comparison with Plato's more philosophic approach to reality, and Hippias has often been presented as standing for a superficial encyclopedic approach to knowledge, in contrast with the more profound penetration of the genuinely philosophic search for truth. This is the way Plato came to view all the Sophists, but it is probably unfair to Hippias, who in some ways anticipated Aristotle's approach to the whole range of human knowledge. Mathematics and astronomy in the sophistic period were certainly not studied for their practical application in everyday life but, rather, in the pursuit of knowledge for its own sake. Hippias made a really important contribution to mathematical development through his discovery of the curve known as the quadratrix, used for the trisection of an angle and later in attempts to square the circle. He was also used by Eudemus as a source for the early history of geometry, which would suggest that he himself may have written a history of mathematics. He was fairly certainly the source of Aristotle's information about the doctrines of Thales, and he may also have been responsible for the main lines of the schematized picture of the history of the pre-Socratics found in Plato's *Sophist* (242D).

Whether he had any general theory of the nature of reality is not certain, but it is probable that he did. In the *Hippias Major*, Plato attributes to him a "continuous doctrine of being," which implies that some particular doctrine was regularly attributed to him. This doctrine dealt with "continuous physical objects that spring from being" (301B), and was opposed to Socrates's attempt to distinguish "the beautiful" from "beautiful objects." While the details of the doctrine are not given, it seems clear that Hippias objected to attempts to explain phenomena in terms of qualities or entities whose existence does not lie wholly within the phenomena that exemplify them. If this is so, then he held to the standard sophistic rejection of the position of Parmenides—for Hippias, phenomenal reality was the whole of reality. If Plato presents the matter correctly, Hippias regarded reality as composed of concrete physical objects such that all qualities applicable to any group will also apply individually to each member of the group, and all qualities found in each of the individual objects will also apply to the group as a whole.

In ethics Hippias propounded an ideal of individual self-sufficiency. Plato's evidence in the *Protagoras*, together with that of Xenophon in the *Memorabilia* (Book IV, Ch. 4, Sec. 5), shows that Hippias made free use of the opposition between nature and convention and that he accepted the overriding claim of Nature in cases of conflict. That he originated this antithesis has often been asserted, but no ancient source suggests this; and there is good evidence that the origins of the doctrine are earlier than Hippias. In the *Protagoras*, Hippias declares that his listeners are kinsmen, friends, and fellow citizens by Nature because the friendship of like to like comes by Nature, not by convention. While this clearly contains the seeds of a doctrine of cosmopolitanism, it should be remembered that Hippias's listeners in the dialogue are

all Greeks and are all alike in their interest in sophistic discussion.

See also Sophists.

Bibliography

Fragments and testimonia are found in H. Diels and W. Kranz, *Fragmente der Vorokratiker,* 10th ed. (Berlin, 1961).

Guthrie, W. K. C. *The Sophists*. Cambridge, U.K.: Cambridge University Press, 1971.

Kerferd, G. B. *The Sophistic Movement*. Cambridge, U.K.: Cambridge University Press, 1981.

Sprague, Rosamond Kent, ed. *The Older Sophists; A Complete Translation by Several Hands of the Fragments in die Fragmente der Vorsokratiker, ed. by Diels-Kranz*. Columbia: University of South Carolina Press 1972.

Woodruff, Paul, ed. and trans. *Hippias Major* (Plato). Indianapolis: Hackett, 1982.

G. B. Kerferd (1967)
Bibliography updated by Paul Woodruff (2005)

HIPPOCRATES AND THE HIPPOCRATIC CORPUS

(b. 460 BCE)

Hippocrates, who came from the Aegean island of Cos, is said to have been born in 460 BCE. His legendary status as the father of medicine is secure—unfortunately, just about everything else about him is legendary too. Tradition records a number of entertaining stories, but it is plain that later writers in the notoriously unreliable Greek biographical tradition knew very little about him. Plato mentions him a couple of times, respectfully, and in *Phaedrus* (270b–d) ascribes approvingly to him the view that in order to know the parts of something one must know the whole. But there is no consensus even as to what the whole here refers to: The whole of the universe? Or simply the complete structure of the body? Different scholars, taking different lines, have consequently seen this remark as alluding to a variety of quite different treatises of the Hippocratic Corpus, and constructed accounts of the *authentic* Hippocrates accordingly. The author of the most comprehensive and learned recent account of Hippocrates and Hippocratic medicine, Jacques Jouanna, while noting the disagreements and the pitfalls, nonetheless tries to distil some spirit of fact from the mash of the biographical tradition and takes note of some relevant recent inscriptional evidence. But the picture is still obscure and speculative. There almost certainly was (though even here scholars contend) a school of medicine on the island of Cos from the fifth century onward, probably in rivalry with an alternative school at Cnidos on the Anatolian mainland. One of the texts of the Hippocratic Corpus refers to a lost treatise named *Cnidian Opinions*, and scholars have tried to reconstruct the methodological differences between the schools (the usual, although disputed, suggestion is that Cnidian medicine favored very precise disease classification and a reliance on purgative treatments, and certain texts in the surviving Corpus, notably *On Diseases* and *Internal Affections* have been classified as Cnidian on doctrinal grounds). Hippocrates himself was associated with the Coan school, and he may well have traveled elsewhere in Greece, perhaps to Thessaly and Macedonia (doctors of the time were often, although not invariably, itinerant). We need not credit the story, even though it is relatively well attested, that he was forced to leave after maliciously burning the archives of the Cnidian school.

So the pursuit of the historical Hippocrates is largely fruitless. However, there survives under his name a collection of some sixty texts (even this number is disputed since scholars cannot agree as to what constitutes separate treatises)—the Hippocratic Corpus. As has been realized since antiquity, they cannot all be ascribed to the same individual, much less to the historical Hippocrates. They exhibit wide divergences not just in subject matter but also in style and doctrine; and some cannot have been written earlier than the third century BCE (others, such as the fictitious correspondence between Hippocrates and Democritus, are later still). Many, however, clearly belong to the fifth century and as such are among the earliest surviving examples of Greek prose. Some (*On Art, On Breaths*) bear the unmistakable stamp of the Sophistic movement and, although containing much of methodological interest, are almost certainly not the work of practicing physicians. Others indubitably are: Some are severely practical and observational in tone (the *Epidemics, On Diseases, On Affections*), others are more theoretical (*Ancient Medicine, Nature of Man, On Prognosis, On Regimen*). Some address issues of medical ethics although in a fairly pedestrian way: *Decorum, The Oath, Precepts*. There are treatises on surgery (*On Joints, On Fractures, Wounds in the Head*), embryology (*On Seed, The Nature of the Child*), and several gynecological texts (*Diseases of Women, Sterile Women, Nature of Women*). The remainder of this article will consider, necessarily briefly, some of the more philosophically interesting texts and the topics they raise.

Greek medicine did not arise out of nowhere in the fifth century. The earliest surviving literary products of

Greek culture, the Homeric poems, mention both surgery and the administration of various treatments by human rather than divine agents; and there was a medical tradition of immense antiquity in Egypt although it is unclear how early it made any impact on the Hellenic world, if indeed it did at all. Various Presocratic thinkers were also renowned for their healing expertise—in particular, Alcmaeon and Empedocles. But the main innovation of the Hippocratic authors (as they shall now compendiously be referred to) seems to have lain in their desire for systematicity (although the Corpus contains several, evidently incompatible, such systems) and the related drives toward diagnostic precision, prognostic knowledge, and nosological explanation. For the Hippocratics (in general—for reasons by now obvious—no generalization across the entire Corpus, no matter how bland, is secure), medicine is about understanding: understanding the nature of health and disease and the measures needed to maintain the former and cure the latter. And they are, fundamentally, physical phenomena, to be approached from a physical point of view.

In a celebrated treatise on epilepsy and related seizure disorders, *The Sacred Disease*, the author opens with the following characteristic statement, which might serve as a motto for Hippocratic medicine in general: "Concerning the so-called 'sacred disease,' these are the facts. It seems to me to be in no way more divine or sacred than any other disease, but has a nature and a cause from which it arises, although men think it be something divine because of their inexperience and their wonderment at its dissimilarity with other illnesses" (*Sacred Disease* 1). The author goes on to castigate as charlatans those who propose religious or magical *cures* for it, declaring that in spite of its peculiar symptomology, it has a determinate physical cause (excess of phlegm in the brain), which may be countered by means both prophylactic and curative. At the end, he writes:

> The so-called sacred disease arises from the same type of cause as the others, from things that enter and leave the body, from cold and heat, and from the winds which constantly change and never rest. All these things are divine, so one should not distinguish this disease as being in any way more divine than the others: all are divine and all human. None is hopeless or untreatable; and most are cured by the same things which cause them.
>
> (SACRED DISEASE 21)

The latter claim is not to be understood as homeopathic: It is the *removal* (or counteraction) of the pathogenic substances that produces recovery, and such allopathy is a Hippocratic commonplace ("opposites cure opposites" occurs as a frequent slogan—see, for example, *Breaths* 1—although it was interpreted in widely different ways). It should also be noted that the author does not reject the claims of divinity altogether—all diseases have an aspect of the divine about them. But crucially, that does not mean that they are not amenable to rational understanding and cure.

Thus, the Hippocratic doctor positions himself in the Presocratic tradition of natural science. Moreover, for many of the authors of the Corpus, a thorough theoretical understanding of the nature of the universe is a prerequisite for understanding, and hence nurturing and curing, the human body. But different authors differ in how far they think such general knowledge should go. Perhaps the most extreme position is that of the author of *On Regimen*. This is, as the title suggests, a treatise about the ways in which lifestyle (diet, exercise, bathing, etc.) affects health. But it is much else besides (it is also perhaps the most traditionally *religious* text of the Corpus, advocating prayer as well as more typically Hippocratic types of therapy). But he begins by declaring that "someone who is to deal with human regimen correctly, must first understand and ascertain the general nature of man: understanding his primary constituents and understanding the parts from which he is composed" (*Regimen* 1.2). The primary constituents turn out to be fire and water, and everything in the universe is in some way an elaboration of these.

Moreover, their ratios of composition and degrees of purity account not only for the generation of other stuffs, but also for the phenomena of mental quickness and retentiveness. Fire is fundamentally motive, water fundamentally nutritive; whereas fire is basically hot and dry, water is cold and wet (although each contains some admixture of the other. The natural world consists of a perpetual fluid interaction between the elements and their properties, and there is no such thing as genuine generation or destruction, only rearrangement, mixture, and separation. So far, late Presocratic—and indeed the author's physical views—seem to be *a cento* of those of Parmenides, Empedocles, Anaxagoras, and (perhaps predominantly) Heraclitus. The ideal condition of the body is one of attunement of the elements whereas disease is disharmony; and the human body is a microcosm of the structure of the universe as a whole.

All this is obviously schematic and, as such, offers no practical clue as to what steps should be taken to combat illness and ill health beyond the bland injunction to cure

opposites with opposites or to suppress the pathogenic influences. There is no consensus in the Corpus as to what the basic elements are: *Nature of Man* (incidentally the one treatise in the Corpus of which authorship is relatively certain: It was composed, at least in part, by Hippocrates' son-in-law, Polybus) rejects the view that the doctors should offer accounts on the human constitution in terms of any of the so-called elements—air, fire, water, or earth—"or anything else that is not clearly a constituent of the human body" (*Nature of Man* 1). The author has, in fact, two distinct targets: one is monism, the view that a single underlying stuff could account for all that there is (plurality is needed for variation and change, he argues; and no unique stuff could suffer pain); but the other, as the quoted clause suggests, is excessive reductionism. One should describe the state of the body in terms of the balance or imbalance of the four humors (this treatise is the first in which that celebrated and long-lived doctrine appears in full), blood, phlegm, yellow bile, and black bile, which are (allegedly) observable constituents of the body. (Black bile is a problem—no one is really sure what this was supposed to answer to, and this fact in itself compromises the supposed empiricism of the theory).

Even more uncompromising is the attitude of the author of *Ancient Medicine*. He argues that "medicine has no need of novel hypotheses" (ch. 1), and rejects *philosophical* physiology in the manner of Empedocles (ch. 20). The *hypotheses* in question are that health and disease are the result of balance and imbalance among four fundamental qualities: hot, cold, wet, and dry. Such postulates are useless for medicine, he argues, since the terms either have their ordinary phenomenal senses, in which case changes and imbalances in them do not correlate with health and sickness, or they are arbitrarily specified *technical* terms, in which case they have no useful empirical content and are simply introduced after the fact to label what are—in the author's view, empirically observable correlations. Thus, it makes sense to categorize foodstuffs in terms of their phenomenal qualities (sweet, sour, salty, etc.) and to relate these to determinable physiological changes; such relations are to be discovered on the basis of long experience (hence the *ancient* of the title). But anything else is superfluous.

Needless to say, not all Hippocratics agreed. The author of *Breaths* is quite happy to describe his basic theoretical postulate (that different types of air are fundamentally responsible, along with food and drink, for health and disease) as a hypothesis and, moreover, one that his discourse has vindicated. But that *vindication* takes the form simply of supplying *explanations*, of a fairly far-fetched variety, for the incidence of particular illnesses (including apoplexy, epilepsy, and fever) in terms of his favored postulates.

Elsewhere, Hippocratic authors do show themselves to be aware, albeit dimly, of the need to support their explanations with empirical observation and sometimes even experiments of sorts. But these appeals to evidence are of widely varying quality and plausibility. Thus, the author of *Airs, Waters, Places*, a study of the generalized effects of climate and ambient environment on human health and character, holds that "water from snow and ice is always harmful, because once frozen it never recovers its previous quality" (ch. 8). The author thinks that "light, sweet" water is the most healthful, and that freezing drives off this part of it; in support of this claim, he says that if you measure water into a jar and leave it outside overnight to freeze, then melt the water in the morning, "you will find it considerably reduced in quantity." Here the hypothesis is plainly not entailed or, indeed, even supported, by the evidence.

Another strand of the Corpus is more observational and practical. The *Epidemics*, a disparate collection of general and particular observations of disease, illustrates this well. *Epidemics* I and III, which are almost certainly from the same pen, consist in general accounts (*Constitutions*) of prevailing epidemic diseases classified by season, place, and other general environmental features. Although apparently the products of disinterested observation, the types of general factors noted point to a particular theoretical account of the origin of disease, again involving the imbalance of climatological and environmental factors. Particular incidences of disease are to be explained in terms of the patients' specific conditions and of particular events that occur to them (excessive eating, drinking, sex, exercise, bathing, for example). The implicit idea, once again, is that the occurrence of disease (as well as the maintenance of health) can be given general, naturalistic explanations in terms of the patient's underlying physiological condition and external occasioning events. It is in this two-fold analysis of the structure of physical explanation, in terms of the interrelation between more or less permanent standing conditions and triggering events, that the Hippocratics made their greatest contribution to the development of the concept of physical explanation.

Much else of importance has been passed over—space permits only a passing mention of the development, in such texts as *On the Art* and *Regimen in Acute Diseases*, of concern with defending the scientific status of

medicine against its detractors: Doctors often fail to cure patients, and patients sometimes recover independently of treatment. These facts do not detract from the art itself: It is no condition of something being a genuine technical skill that it must yield 100 percent success; there are always other factors that can interfere, such as the failure of the patient to follow the prescription; the disease is already too deeply entrenched to be eradicated. Indeed, the author of the influential *On Prognosis* notes that one of the advantages conferred by prognostic ability is that of knowing which diseases are incurable and being able to leave well alone. Moreover, the fact that some practitioners are charlatans does not mean that they *all* are. The existence of such defenses as early as the fifth century BCE shows that the practitioners of the infant science of medicine were well aware of the seriousness of the challenge to their claims to expertise and that they were capable of considerable sophistication in rebutting them.

Finally, a number of texts, usually labeled *deontological*, deal with matters of professional conduct and ethics. The most famous of these, the *Oath*, still serves as a template for medical codes of conduct. Among its clauses are injunctions to protect the secrecy of medical knowledge, not to infringe on the turf of other professionals (in particular, surgeons), never knowingly to cause harm, and to resist the temptation to abuse one's professional position for sexual purposes. In spite of their pretensions to comprehensive theoretical and practical knowledge, the Hippocratics were aware of their own limitations, knowing that nature was the best hope for a cure in most cases—the job of the physician being to help nature in its healing course.

See also Alcmaeon of Croton; Anaxagoras of Clazomenae; Empedocles; Heraclitus of Ephesus; Leucippus and Democritus; Parmenides of Elea; Philosophy of Medicine; Philosophy of Science; Plato.

Bibliography

Hippocrates. Vol. 1, *Ancient Medicine, Airs, Waters, Places, Epidemics 1 & 3, The Oath, Precepts, Nutriment*. Vol. 2, *Prognostic Regimen in Acute Diseases, Sacred Disease, Art, Breaths, Law, Decorum, Physician (Ch. 1), Dentition*. Vol. 3, *On Wounds in the Head, In the Surgery, On Fractures, On Joints, Mochlicon*. Translated by E. T. Withington (1923). Vol. 4, *Nature of Man, Regimen in Health, Humours, Aphorisms, Regimen 1—3, Dreams; Heracleitus: On the Universe*. Translated by W. H. S. Jones (1923—1931). Vol. 5, *Affections. Diseases 1, Diseases 2*. Vol. 6, *Diseases 3, Internal Affections, Regimen in Acute Diseases*. Vol. 7, *Epidemics 2, 4—7*. Translated by Wesley D. Smith (1994). Vol. 8, *Places in Man, General Nature of Glands, Fleshes, Prorrhetic 1—2, Physician*. Translated by Paul Potter (1988—1995). Loeb Classical Library. Cambridge MA: Harvard University Press, 1923—1995.

Jouanna, Jacques. *Hippocrates*, translated by M. B. De Bevoise. Baltimore, MD: Johns Hopkins University Press, 1999.

Lloyd, G. E. R., ed. *Hippocratic Writings*. Harmondsworth, U.K.: Penguin, 1978.

R. J. Hankinson (2005)

HISTORICAL MATERIALISM

The materialist conception of history was put forward by Karl Marx and Friedrich Engels and subsequently adopted by their followers and incorporated in the doctrine of Marxism-Leninism. According to "historical materialism," the structure of society and its historical development are determined by "the material conditions of life" or "the mode of production of the material means of existence." These last two phrases are quoted from Marx's preface to his *Critique of Political Economy* (1859), in which he gave a brief presentation of the view. Marx and Engels had formulated it, however, in their *The German Ideology,* written in 1845–1846 but not published until 1932. Marx himself gave a brief account in his *Poverty of Philosophy* (1847) and more concisely perhaps in a letter to Paul Annenkov, written in December 1846, while Marx was working on the *Poverty of Philosophy.* A vigorous sketch is given in the *Communist Manifesto* of 1848. Marx's chief work, *Capital* (the first volume of which was published by Marx in 1867 and the other two by Engels after Marx's death) is an application of the historical-materialist view to the capitalist form of society.

ORIGIN OF THE THEORY

Marx wrote in the preface to the first edition of *Capital* that he conceived "the development of the economic structure of society to be a natural process." This is the main force of the adjective *materialist* in the phrase "materialist conception of history." Marx used the word *materialist* to make a contrast with what is obviously or implicitly supernatural, metaphysical, or speculative. He believed that a general science of human society could be worked out only by describing and explaining society in empirical terms. He admired those English and French writers who, by writing "histories of civil society, of commerce and industry," gave the writing of history "a materialist basis" (*The German Ideology,* p. 16). He and Engels regarded industry and commerce as "material" by contrast with religion and morals, and even by contrast with politics and law. Thus the materialist conception of history is intended to be a naturalistic, empirical, scientific

account and explanation of historical events, which takes industrial and economic factors as basic. It would seem that nothing could be more consonant with scientific common sense, nothing less metaphysical or speculative.

In some of Marx's writings of an earlier date than *The German Ideology,* however, it becomes evident that the later, would-be scientific view arose out of a metaphysical prototype, a sort of "Ur-Marxismus," which continued to exert an influence on all of Marx's systematic work. Prior to his collaboration with Engels, which began in 1844, Marx had justified his radical views by philosophical and moral, rather than economic, considerations. In 1844, however, Engels encouraged Marx to make an intensive study of economics, which resulted in an uncompleted and unpublished critique of political economy combined with a critique of the Hegelian philosophy. These so-called *Economic and Philosophical Manuscripts of 1844,* or "Paris manuscripts," are the first draft of the comprehensive treatise that Marx was engaged in writing all his life, and of which *The German Ideology,* the not published until 1953 *Grundrisse der Kritik der politischen Ökonomie* (Outline of a Critique of Political Economy; 1857–1858), the *Critique of Political Economy* itself (1859), and *Capital* are successive, but incomplete stages.

While writing the *Economic and Philosophical Manuscripts,* then, Marx was bringing his newly acquired economic knowledge to bear upon views he had reached in criticizing certain of G. W. F. Hegel's writings. Marx had noticed how Hegel described the development of the human mind as a process of externalizing its ideas in order to transform the material world and to "humanize" it. According to Hegel, the labor of men's hands was not, in general, an obstacle to human development but, rather, the very process by which it took place. Hegel recognized, of course, that when labor was greatly subdivided, some jobs became trivial and even degrading. But this, he thought, made possible, through the differentiation of society into orders or classes, the production of works of mind that would have been beyond the power of less differentiated societies. The word that Hegel had used for the process of externalizing ideas into the natural world was *alienation* (*Entäusserung*). Now Marx thought that in the capitalist social order the labor of individual men did not serve to develop the human mind and to humanize the natural world. Labor had become the production of commodities for sale and was itself a commodity bought and sold in the market, so that it served not to unfold the capacities of the laborer but to subject him to impersonal market forces over which he had no control. A worker's labor, and hence he himself, were alienated in the sense of being sold to someone else. His work resulted in the creation of a social system whose operations were hidden from him. The wage system perverted his labor so that the natural world was not transformed by that labor into a manifestation of human power but was rendered strange and even hostile to the workers.

Estrangement (*Entfremdung*) was another word used by Hegel that Marx took over in this context. A truly human existence would be possible only when money and private property, and hence wages too, had been abolished through the establishment of a communist social order. A communist society, Marx wrote, is "the solution to the riddle of history."

It is important to notice that in these early writings Marx was criticizing capitalism in metaphysical and moral terms. But for the perverting influence of capitalism, human labor would be what it ought to be, the self-development of the individual worker. It should be noted, too, that Marx, like Hegel, thought that the human mind could develop its powers only by working on, and transforming, the natural world. This conception is a metaphysical predecessor of the view that the "mode of production of the material means of existence" is what determines the development of society. Again, the view that capitalism distorts the efforts of the worker and is hence unnatural and impermanent is the metaphysical predecessor of the view that capitalism contains the seeds of its own destruction. Finally, the idea that communism would solve the riddle of history by releasing men from their own unwilled, unwanted productions is the metaphysical predecessor of the planned but noncoercive communism that Marx afterward believed must result from the dissolution of capitalism.

OUTLINE OF THE THEORY

Historical materialism consists, in the first place, of a sociological analysis thought to be applicable to all but the most primitive human societies. On the basis of this analysis an account is given of the rise and fall of various social systems. Marx's main work, of course, was his analysis of capitalism—indeed, the very use of the word *capitalism* for a form of society suggests that its characteristics depend upon its economy. Finally, on the basis of the sociological analysis, the prediction is made that capitalism will collapse and ultimately be succeeded by a communist society, in which there will be no wages, no money, no class distinctions, and no state.

Marx, who was greatly interested in the social structure of primitive societies, would doubtless have agreed

with Engels's description, in his *Origin of the Family, Private Property, and the State* (1884), of the most primitive societies as being without private property or political institutions. Within the more developed societies, with which he was principally concerned, Marx distinguished several elements: (1) "the productive forces," which consist of the tools, skills, and techniques by which men obtain the wherewithal for life; (2) "the relations of production," which are the ways in which the producers are related to one another in production and which form "the economic structure of society"; (3) the political and legal institutions of the society; and (4) the ideas, habits of thought, ideals, and systems of justification, in terms of which the members of the society think of themselves and of their relations to one another. Marx thought that these ideas were distorted pictures of, and relatively ineffective agents in, the social reality, and he therefore referred to them as "ideologies." Marx gave various lists of ideologies that, when combined, yield the following: religion, theology, speculative philosophy or metaphysics, philosophy, morality, ethics, art, and "political ideology," such as contrasting views on democracy, aristocracy, and the struggle for the franchise.

ANALYSIS OF SOCIAL STRUCTURE. Marx called the productive forces and the relations of production together "the material conditions of life." In the preface to the *Critique of Political Economy* he wrote that they are "the real basis on which a juridical and political superstructure arises and to which definite forms of social consciousness correspond." The primary social activity is production, which always involves relations with other men, both in the work itself and in the distribution of the product. It is upon these relationships that the political and legal superstructure and the ideological superstructure are formed. To understand the religion, morality, art, or philosophy of a society, and to understand its politics and law, it is necessary to ascertain the nature of its productive forces and economic structure. Whereas in the *Economic and Philosophical Manuscripts* Marx had deplored the way in which men's labor enslaves them to the production of commodities, in the *Critique of Political Economy* he explained or sought to explain, how the productive forces determine certain social structures into which men are forced to fit their activities. Thus Marx laid great stress upon the fact that the structure of society is something that individuals find waiting for them and are powerless to alter.

Division of labor. According to Marx, a vitally important connection between the productive forces and the productive relationships is the nature of the division of labor that has been achieved and the degree to which it has been developed. In *The German Ideology,* Marx and Engels wrote that "division of labor and private property are, moreover, identical expressions." This probably means that when products are made by specialists who do not themselves use them, then they must be exchanged by, or sold to, those who do and so must be owned by the original maker. An associated idea is that the division of labor fosters the production of goods for sale, thus encouraging the production of commodities and enhancing the power of money. Marx and Engels did not think, however, that property was all of one type, and in *The German Ideology* they distinguished four main types that play an important role in their theory of history and society: tribal property, which is characteristic of a low level of the division of labor; state property, such as the roads, public buildings, and stores of grain under the ancient forms of despotism; feudal property, consisting of lands and services controlled by military landowners whose needs are supplied by serfs; and capital, which rests on the separation between production and commerce and results in the employment of men who work for wages and produce goods that are sold in wider and wider markets to make profits for the capitalist.

Property and power. The next step in the Marxian analysis is the claim that the main power or influence in a society belongs to those who own and control the main type of property in it. In tribal society the property is jointly owned; hence power is diffused throughout the society and there is no dominant class. The other types of property involve a distinction between those who control property and those who do not. Those who control a predominant type of property are the predominant power in society and are able to make arrangements benefiting themselves at the expense of the rest of the population. In feudal society, for example, the feudal lords are the ruling class. They are able to get what they want from the serfs who work for them, and even from rich merchants, whose type of wealth is subordinated to the landed interests. The interests of serf, merchant, and lord are not the same; indeed, they necessarily conflict at certain points. But while the productive forces and type of property are predominantly feudal, the feudal lords are able to settle these conflicts in their own favor. While the feudal system operates, any frictions and tensions are dealt with within its terms. The political movements in a feudal society express, or "reflect," these conflicts of interest between classes.

Economics, politics, and culture. If the political activities of men are regarded as merely phenomenal in com-

parison to their productive and economic activities, then their moral beliefs, religious and artistic achievements, and philosophical theories must be regarded as even less real, as epiphenomenal. The writers of books on political philosophy, for example, are taking part, but in a rarefied or ghostly form, in the phenomenal political activities and the real industrial ones. The predominant mode of the material conditions of life will have the cultural forms appropriate to it, in which the religion, art, and philosophy are what they are because of the nature of the technology and economy. The controversies between "schools" of philosophy, the movements for the reform and renovation of religious belief, the revolutions of morality, and even changes in artistic style, are merely the shadows cast by the "real" business of human living, which is production and exchange.

HISTORICAL EPOCHS. Thus far an outline has been given of what, in Auguste Comte's language, might be called "the social statics" of historical materialism. It is now necessary to describe "the social dynamics" of the view: its account of historical change and development. In outline, this is the assertion that, just as "the material conditions of life" are fundamental in the structure of a society, important changes in the material conditions of life sooner or later bring with them important changes in the legal and political superstructure and in the ideological superstructure. It is also held that important changes in the superstructures can be brought about only by changes in the basis, that politics, law, and ideology are incapable in themselves of any fundamental influence on social development. All important social changes, it is held, must originate in productive activities and the organizations in which they take place. This is the central element of the theory of historical materialism.

This theory is also a theory of historical epochs. The original state of primitive communism was succeeded, according to the Marxist view, by the ancient forms of slave-owning society; these were succeeded by feudalism, and feudalism by capitalism. According to *The Origin of the Family,* the transition from primitive communism to the next phase was due to the introduction of private property. It is clear, of course, that the introduction of private property would bring with it very important social changes, but how is private property itself introduced? We have already seen that one idea is that it is brought about by the division of labor. In *The Origin of the Family* Engels also suggested that it was furthered by changes in the structure of the family and by the discovery of iron and bronze. The former would hardly be a technological invention, although the latter was. Engels's doubts on the matter may be seen from the fact that when he discussed the question of how the common ownership of herds was succeeded by private ownership he vaguely said that "the herds drifted into the hands of private individuals." However private property is held to have arisen, the division of labor brought with it the transformation of goods into commodities and their sale for money.

The next epoch after the period of primitive communism was that of ancient slave society. Marx and Engels held that it was the labor of slaves that made possible the art and science of ancient Greece and the cities, commerce, and bureaucracy of ancient Rome. The slave system broke down largely because of its wastefulness and was replaced by the feudal system, in which features borrowed from the social system of the barbarian invaders were utilized. The basis of the feudal system was the ownership of land by feudal lords, whose dependents had to render them services of various kinds.

The feudal system was fundamentally an agricultural society, but in the towns some men managed to become wealthy by means of trade and by organizing the production of goods in large workshops where they employed considerable numbers of men for wages. These bourgeois, as they were called, were the forerunners of the capitalist system. They attracted men from the countryside to work for them in producing goods sold in widely expanding markets. In this and other ways they acted in opposition to the predominant feudal arrangements that confined serfs to the areas of their birth. Finding themselves hampered by the feudal laws, the bourgeois endeavored to change them and thus entered upon a political struggle with the aristocracy. They justified their actions by appealing to a new ideology according to which aristocratic distinctions based on family connections, and control over the movements of men and over trade, were in opposition to the "natural" order of individual freedom and equality.

As the new methods of production and the new modes of life that went with them were extended, a new order of society was gradually formed within the old. New types of production and trade had been adopted that could come to fruition only if the laws and customs that hampered them were abolished. When, therefore, the bourgeoisie were strong enough, they took political action to achieve this and gained political power by a series of revolutions, of which the French Revolution of 1789 was the culmination. From being a progressive class they became the ruling class, and their landowning opponents declined from being the ruling class into being a reactionary class, which, however, could not return soci-

ety to its earlier state, because the new productive forces were superior to the old ones.

This interpretation of the change from feudalism to capitalism illustrates the Marxist analysis of political revolutions. Marx and Engels regarded such revolutions as the means by which a progressive class, that is, the class that controls some newly emerging productive force, brings about those changes in the productive relationships that enable the new productive forces to become effective. Feudal institutions and, in particular, feudal laws of property would have stifled the development of the capitalist modes of production. By their seizure of political power, the bourgeoisie made laws that enabled capitalism to become a going and growing concern.

SOCIAL PREDICTION. Historical materialism makes two main predictions. The first is that the capitalist system will break down as a result of its internal contradictions. The second is that, after a period of proletarian dictatorship, it will be succeeded by a communist society.

Breakdown of capitalism. In *Capital,* Marx was largely concerned with an analysis of the capitalist order, but he also briefly considered the future of capitalism. He held that the capitalist economy was so far out of human control that economic crises were inevitable features of it. He held, too, that in competing with one another to sell their goods at a profit the capitalists would find it necessary to push down the wages of their employees to the lowest level consistent with their being able to produce at all. Furthermore, the advantages of large-scale production would be such that the larger capitalists would drive their weaker rivals out of business and into the ranks of the proletariat. As a few capitalists became richer, the mass of workers would become poorer. At the same time the growth of scientific knowledge would enable the larger capitalist concerns to improve their technology, so that nature would be brought under human control as never before. Thus, the subdivision of labor is increased, and great numbers of men, organized in manifold ways, cooperate, often in ways unknown to one another, in the manufacture of a single article.

Although production is thus highly socialized, ownership of the means of production and of the commodities produced is still an individual matter. Engels expressed this by saying that there is a contradiction between capitalist appropriation and social production that must result in the elimination of the former. The conditions of life imposed on workers in capitalist production teach them how to cooperate against their employers. The capitalist mode of ownership stands in the way of the fullest development of planned production. "The centralization of the means of production and the socialization of labor reach a point where they prove incompatible with their capitalist husk. This bursts asunder. The knell of capitalist property sounds. The expropriators are expropriated" (*Capital,* Vol. I, Ch. 24).

Arrival of communism. Just as the bourgeoisie found it necessary to achieve control of the state in order to bring the feudal system to an end, so the proletariat will find it necessary to wrest the state from capitalist control in order to bring capitalism to an end. Thus while the proletariat, or their spokesmen, are criticizing the bourgeoisie, they constitute the rising, progressive class, and when they have overcome the bourgeoisie, they will become the ruling class. But once the bourgeoisie are ousted, there will be no other class for the proletariat to oppose. The proletariat will be the only class, or rather, the class that will bring class divisions to an end. In the absence of class conflicts, politics and the state will become redundant, and a social order will arise in which production will be carried out in accordance with plans devised without coercion for the good of all. According to *The German Ideology,* the outcome will be "the control and conscious mastery of those powers which … have till now overawed and governed men as powers completely alien to them." Twenty years later Marx wrote of "a process carried on by a free association of producers, under their conscious and purposive control," adding: "For this, however, an indispensable requisite is that there should exist a specific material groundwork (or a series of material conditions of existence) which can only come into being as the spontaneous outcome of a long and painful process of evolution" (*Capital,* Vol. I, Ch. 1).

PROBLEMS OF INTERPRETATION

In the course of the many discussions of historical materialism since Marx's day, among Marxists as well as between Marxists and their critics, various problems of interpretation have come to light. Questions arise about the nature and status of the theory itself. There is the question whether the theory is to be interpreted as asserting the primacy of technology both in the structure of society and in the promotion of social change or whether the prime element is wider in scope and is intended to embrace economic as well as technological relationships. A third problem concerns the connection or lack of connection between historical materialism as a value-free sociological theory and as an element in the socialist outlook and an ethical justification of socialist expectations.

NATURE AND STATUS OF THE THEORY. Is historical materialism the statement of an established sociological or historical law? Is it an extremely wide-ranging and complex hypothesis liable to refutation as research advances? Or is it, as some have suggested, not so much a hypothesis as a method, or recipe, or set of hints for framing one? The Marxist-Leninist tradition of the Russian and Chinese Communist parties undoubtedly adopted the view that it is an established law, as reference to Marxist-Leninist textbooks shows. It is sometimes said that Marx himself held the methodological view about his own theory. This is supported by a phrase in the preamble to his famous account of historical materialism in the preface to the *Critique of Political Economy*: "The general conclusion I arrived at—and once reached it served as the guiding thread in my studies." But in this passage Marx is describing how he came to adopt the view, so that the expression "guiding thread" relates to the use he made of the idea in its early stages rather than to the theory once it was established. It seems fair to say that historical materialism was a view that Marx was constantly trying to support but never to refute. Furthermore, as will be shown, the theory contains features suggesting that Marx held it to be a necessary truth. V. I. Lenin, in an early pamphlet titled *What the "Friends of the People" Are* (1894), said that historical materialism was "no longer an hypothesis, but a scientifically proven proposition," but he admitted at least the possibility of its being upset. In *Materialism and Empirio-Criticism* (1909), however, he considered that historical materialism was a consequence of dialectical materialism and thus to be proved in quite a different way.

THE PRIME SOCIAL DETERMINANT. Was the prime social determinant, in Marx's view, the productive forces, or was it the whole composed by the productive forces and the productive relationships? Was it, that is, technology alone, or technology plus economy? The Marxist-Leninist tradition favors the first interpretation, and there are many passages in Marx's writings to support it. For example, Marx wrote in *The Poverty of Philosophy*: "In acquiring new productive forces men change their mode of production, and in changing their mode of production, their manner of gaining a living, they change all their social relations. The windmill gives you society with the feudal lord; the steam mill, society with the industrial capitalist."

A similar point of view is indicated in the *Communist Manifesto,* in which Marx wrote: "The bourgeoisie cannot exist without constantly revolutionizing the instruments of production, and thereby the relations of production, and with them the whole relations of society." In a footnote to Chapter 13 of Volume I of *Capital* he said that "the only materialist method" is to show how technology "uncovers man's active dealings with nature, the direct productive process of his life, and, at the same time, of his social relationships (*seiner gesellschaftlichen Lebensverhältnisse*) and the mental conceptions that arise from them." In the same passage he talked about those who uncritically abstract from "this material basis," and he advocated tracing the development of "the celestial forms" of these real relationships (*wirklichen Lebensverhältnisse*) from the real relationships themselves. It is clear that Marx was here arguing that religious ideology should be explained in terms of real social relationships and that these, in their turn, should be explained by reference to technology. But the language he used does not suggest that he was making sharp distinctions. Indeed, what he criticized is the attempt to consider other forms of life in abstraction from technology, so that he could be regarded as upholding what Benedetto Croce in 1896 called the "realistic view of history."

Certainly Marx said a number of things that contradict a merely technological theory of history. Perhaps the most compelling evidence for the view that Marx regarded the basic social determinant as comprising more than technology is his account in *Capital* of the rise of modern capitalism. According to Marx, modern capitalism began with the setting up of large workshops in which men worked for wages in producing goods that the capitalist employer sold for profit. These workshops or factories were new forms of organization, not new methods of production. If they are to be regarded as productive forces, then organization is a productive force. How far is this to be taken? These early capitalists were trying to supply a wider market than had hitherto been possible, and thus considerations of demand and of economic efficiency enter into the notion of a productive force. This notion, indeed, can be extended to include commerce, piracy, and war, and Marx and Engels did so in the early pages of *The German Ideology.* But if commerce is a productive force, then the distinction between productive forces and productive relations is blurred, if not abolished altogether. And if war is a productive force, then it would seem that politics is also a productive force, and in this way the distinction between basis and superstructure disappears.

That Marx and Engels were not clear about all this may be seen in two letters from Marx to Engels on the subject of armies and armaments. In a letter to Engels dated September 25, 1857, Marx wrote: "The history of

the army brings out more clearly than anything else the correctness of our view about the connection of the productive forces and social relations. The army is particularly important for economic development, *e.g.* wage payments first fully developed in the army among the ancients. Thus the *peculium castrense* was among the Romans the first legal form in which the chattels of those who were not fathers of families were recognized.…" In a letter dated July 7, 1866, Marx referred to the new types of weapons that the manufacturers were trying to sell to Louis Napoleon and commented: "Where does our theory about the determination of the organization of labor by the means of production get more brilliant support than from the human slaughter industry?"

In the first of these letters the idea is that the waging and winning of war depend upon the refinements of armament manufacture, which, in their turn, depend upon the level of technology achieved in the society. Here the armaments industry seems to be regarded as a means of production, and the waging of war as the organization of labor. It should be noted, too, that in the first letter the distinction is between productive forces and social relations, where the social relations referred to are working for wages and owning chattels. In the second letter, however, the distinction is between the means of production and the organization of labor. It is possible that by "productive forces" and "means of production" Marx meant much the same thing, but "social relations" is clearly a much wider notion than "organization of labor." In the light of such examples, it can hardly be denied that Marx had no precise view of the theory that he was putting forward.

THE PLACE OF VALUES IN THE THEORY. The third problem of interpretation concerns the connection between historical materialism as an alleged scientific theory and the advocacy of an eventual classless society apparently involved in it. On the one hand, there is the claim that historical materialism is scientifically established and explains how things are and predicts what they will be. On the other hand, there is the promise that out of the contradictions of capitalism a superior form of society will arise in which there will be no more coercion or exploitation. By a happy conjunction a moral millennium is held to be predictable on scientific grounds. As was said at the beginning of this entry, the doctrine of historical materialism arose out of an earlier metaphysicomoral view in which scientific objectivity played no part. Some critics therefore take the view that Marx was at the same time a moralist and a sociologist and that he never succeeded in reconciling these roles. Others go still further and suggest that the scientific works are nothing but a vehicle for his moral aims.

Defenders of Marx argue that he rightly refused to make the distinction between fact and value that is implicit in the claim that social science should be "value-free." They argue that Marx considered that theory and practice are inextricably mingled, so that it is impossible to understand the working of social processes without at the same time obtaining control over them. Marx very probably believed that capitalist society develops in ways that are not intended by anyone and that it would be succeeded by a form of society in which men's aims and intentions would find scope for fulfillment. Thus, in his view, the processes of capitalist society can be observed and explained as if they were the workings of some alien, nonhuman entity in which individuals are caught up as in some monstrous mechanism. Nevertheless, he also held that the machine would break down and be destroyed and that the activities of men, thus released, would be explicable not in impersonal terms but in terms of their collective aims.

THE VALIDITY OF HISTORICAL MATERIALISM

It has already been pointed out that historical materialism has been supported on grounds of very different sorts. It has been regarded as a method of investigating the facts of history, as an established historical hypothesis of great generality, and as a deduction from materialism, or, more specifically, from dialectical materialism. It has also been said that Marx regarded his view as more than a method and that if he regarded it as a hypothesis, he hardly considered the possibility of its being upset. We shall consider the various reasons put forward in its support, so that we can get a clearer understanding of the theory.

DEDUCTION FROM DIALECTICAL MATERIALISM. The view that historical materialism is a deduction from dialectical materialism was apparently not put forward by Marx himself. Dialectical materialism may be implicit in Marx's writings but it is not explicit there, and when Marx wrote of materialism, he frequently meant nothing but a scientific, this-worldly view of things. In the Marxist-Leninist tradition, however, the argument has been used that if dialectical materialism is true, then historical materialism is true also. Thus in his *History of the Communist Party of the Soviet Union* (1938) Joseph Stalin wrote: "Further, if nature, being, the material world, is primary, and mind, thought, is secondary, deriv-

ative: if the material world represents objective reality, existing independently of the mind of men, while the mind is a reflection of this objective reality, it follows that the material life of society, its being, is also primary, and its spiritual life is secondary, derivative, and that the material life of society is an objective reality existing independently of the will of man, while the spiritual life of society is a reflection of this objective reality, a reflection of being."

A somewhat similar argument is to be found in section 2 of Chapter 6 of Lenin's *Materialism and Empirio-Criticism* (English translation, Moscow, 1939, p. 115). Both Lenin and Stalin supported this view by reference to Marx's statement in the *Critique of Political Economy* that "it is not the consciousness of men that determines their being but, on the contrary, their social being that determines their consciousness." But Marx, in this passage, was not referring to materialism as a philosophy of nature, but to the ideologies that are formed in specific social circumstances. Furthermore, it does not follow from the fact (if fact it be) that there is nothing but matter and its forms of being, that the productive and economic activities of man provide the key to his politics, law, religion, philosophy, art, and morals. The adjective *material* does not have the same meaning in Marx's usage as it has when used in the phrase "material world" or "material object." The general acceptance of materialism does not entail any particular view about which features of human life can be used to provide an explanation for the rest.

It might be argued, of course, that if materialism is true, all social facts are reducible to physical facts or that all social laws are reducible to laws of physics. Marx and Engels, however, did not believe this. In an interesting letter, one of the last to pass between them, Engels maintained that "labor" is a social term that cannot be reduced to "work" in its physical or mechanical sense.

HISTORICAL MATERIALISM AS OBVIOUSLY TRUE. It is an exaggeration to say, as some have, that Marx gave no reasons at all for the doctrine of historical materialism. It is clear, however, that both he and Engels regarded it as obviously true. Thus, in the *Communist Manifesto* occurs the following question: "Does it require deep intuition to comprehend that man's ideas, views and conceptions, in one word, man's consciousness, changes with every change in the conditions of his material existence, in his social relations and in his social life?" Engels, in his speech at the graveside of Marx, referred to Marx's "discovery" as the discovery of "a simple fact." This "simple fact" is clearly neither a deduction from dialectical materialism nor a complex hypothesis based on a mass of historical information. It would seem to be the fact that men could not engage in politics, religion, philosophy, and art unless they were alive, with the wherewithal to do so. No one could reasonably deny this, but is every reasonable man therefore an implicit upholder of historical materialism? For this to be so, it would be necessary to show that the theory that the material conditions of life must provide the explanation for all other human activities is deducible from the fact that men must get the wherewithal to live in order to be in a position to engage in political, religious, philosophical, and artistic pursuits. But from the fact that obtaining the wherewithal to live is a sine qua non of politics, religion, and philosophy, it does not follow that these latter activities can be explained only in terms of the former. It seems that a mistake has been made not unlike the failure to distinguish between necessary and sufficient conditions. From the fact that men could not engage in these activities unless they kept themselves alive, it does not follow that how they keep themselves alive explains or "determines" these activities. Engels's statement could be denied only by someone who held that politics, religion, and philosophy were the pursuits of disembodied spirits. His simple fact is too simple to be of any theoretical value.

ARGUMENT FROM THE ESSENCE OF MAN. Marx himself had another argument suggesting that there is something obvious in the view that the productive forces are the determining factors in human society and human history. He wrote in *Capital,* Volume I, that toolmaking is what distinguishes man from other animals. He and Engels had argued in a similar way in *The German Ideology* that men "begin to distinguish themselves from animals as soon as they begin to *produce* their means of subsistence.…" Of course, beavers and bees do this too, but their hives and dams (Marx and Engels would probably have argued) are never improved upon and never serve as the starting points for other devices. Whatever the difference, Marx and Engels held that what is peculiar to human beings is that they make (and presumably improve) their means of life and that, therefore, this fact must be the key fact in sustaining human society and in explaining the course of human history as distinct from natural history.

This is to adopt an Aristotelian method of explanation in terms of essences. What men do, it is supposed, depends upon what men essentially are. It is assumed that there is some central feature common to all human beings and to them alone upon which all their other specifically human activities depend and in terms of

which they must be explained. To this it may be objected, in the first place, that human beings are not the sort of beings to which essences may be attributed. Beings with essences are those that can be classified in some definite way in a well-defined system of classification. The Aristotelian scheme presupposed a world of things that can be so classified, and it was found necessary to abandon the scheme when it was realized that the world was too complex. Essences may be defined for artifacts with definite functions, such as chairs and knives. A knife is an instrument for cutting, a chair an article of furniture for seating one person. But human beings cannot be fitted into any single system of aims or functions.

The Aristotelian definition of man as a rational animal sums up a view of man's place and purpose in the cosmos. It is absurd to suppose that there is any single thing that constitutes the humanity of man, as cutting constitutes the nature of knives. The choice of a single word such as *reason* or *political* or *toolmaking* gives the appearance of such an essence, but it is an appearance only, since each of these words expresses a highly complex notion that cannot be caught up as a definition with a single classificatory scheme. It has already been noted that man is not the only animal that makes its means of life, but that bees and beavers—to mention only two—do so as well. What differentiates human productions is that they are constantly improved on and form the basis for new ones that become progressively less and less like those from which they originated. To say that toolmaking is the essence of man is to refer to his inventiveness in one of its most concrete forms. If man has an essence, it is that he has none.

Why did Marx and Engels pick on toolmaking as the feature that differentiates man from the other animals? There does not seem to be any single answer. Marx, at any rate, was influenced by the archeological classifications of the periods of prehistory into the Stone Age, Bronze Age, and Iron Age. But of course he was wrong if he supposed that because prehistory has to be reconstructed from the material things left behind, these material things are the basic explanatory factors in all human society. (In any case, some of the archaeological remains are not tools at all.) Insofar as archeologists adopt the hypothesis or method of historical materialism, they do so *faute de mieux,* for by the very nature of their business there is nothing else they can do.

A more fundamental reason for the view of Marx and Engels that toolmaking is the human essence is their acceptance, not perhaps altogether conscious in their later years, of the Hegelian view that men create their lives through labor. Technology is thus regarded as the concrete embodiment of the process by which nature is controlled and humanized.

Again, Marx and Engels lived at a time when people were becoming aware of the social effects of important industrial inventions. They saw that a new form of society was coming into being as a result of the invention of steam power and that a society with cotton mills and railroads required very different institutions from those of a society with cottage looms and stage coaches. In our own day the social influence of technological invention has become obvious, at any rate in a general way, even though the specific effects of particular inventions may sometimes be difficult to ascertain. But Marx and Engels noted this at a time when not everyone was aware of what was occurring. But it should be noted that this does not establish historical materialism. From the fact that important technological changes often make it necessary to change laws and to adopt new modes of life and thought, it does not follow that law and modes of life and thought can be decisively altered only as a result of technological change. Furthermore, from the great social importance of technological invention nothing follows as to the causes and conditions of technological invention itself.

LINKAGE OF PRODUCTIVE FORCES AND RELATIONS. In saying that Marx regarded historical materialism as obviously true we are saying that he regarded it as obvious that the productive forces "determine" the productive relationships. There is a sense in which productive relationships are necessarily linked with productive forces. For in inventing a new tool or machine it may well happen that the inventor is requiring so many men to work together such and such ways. A man might, for example, invent or design a sailing ship that required five men to sail it and each member of the crew to occupy a certain position in the vessel. Again, when it was discovered how to equip ships with steam or gasoline engines, the work demanded of seamen was altered and new relationships created among them. Controlling boilers and engines is quite different from handling lines and sails. The jobs are different, and the relationships of those who do the jobs are different too. The point therefore may be expressed by saying that sometimes the introduction of a new type of tool or machine necessarily involves the introduction of new job relationships. It would be natural enough to call these job relationships productive relationships in contrast with the tools or machines themselves, which might be called productive forces or means of production. With the terms understood in this way, then, it can happen that a change in productive

forces necessarily brings with it a change in productive relationships, since the productive forces and the productive relationships may be different aspects of the same thing.

How far does this sort of productive relationship extend? We may take the example of the invention of the airplane to elucidate this question. An airplane at first was flown by one man; later models require several operators. Hence there are certain job relationships for the actual operation of the machine. In addition, however, an airport is required and, if journeys are to be undertaken, other places for landing and refueling. If an airplane is regarded as a machine for flying considerable distances from its base, then the provision of airfields with men to supervise takeoffs and landings and to help in refueling is necessarily involved in the invention too. Thus there are rather extensive job relationships implicit in the invention of a machine for flying from one place to another.

Now there is a principle of Roman law according to which the owner of land owns the whole volume of earth and air below and above it, *de caelo usque ad inferas* (from heaven above to hell beneath). If this principle were insisted on, those who fly airplanes would find it necessary to obtain permission from, or even make payments to, the intervening landowners before they could fly from their own territory. Actually, a system of permissions and exclusions has arisen according to which landowners within a country generally cannot prevent airplanes from flying over their land, whereas governments have certain powers of control over flights crossing their boundaries. Someone might argue that in inventing a machine for flying considerable distances from its base, the inventor was providing not only for the piloting of the aircraft and for its landing and refueling but also for the rules by which it would be controlled as it went from place to place. But this would be to extend the notion of job relationships much too far. Whereas piloting and landing and refueling may be regarded as aspects of flying the machine, and hence as necessary features of the invention, the rules under which the flights may be allowed are a different matter. An injunction to prevent the flight might have been issued after arrangements had been made for it to take place. Thus the third set of relationships is connected with the invention in a contingent way. It might be convenient to call these last relationships productive relationships as distinct from job relationships, even though use of the adjective *productive* exaggerates the connection with the actual operation of the machine. Thus it is clear that whereas a given invention may necessitate certain job relationships, it will be inconsistent with certain wider relationships and consistent with a variety of others. Use of the word *determine* both for the job relationships and the wider ones obscures this difference and encourages the idea that technology sets bonds of necessity upon the social system.

ARGUMENT FROM THE HISTORY OF CAPITALISM. By far the greater part of Marx's historical work was concerned with the origins and development of capitalism, and it is therefore reasonable to regard this part of his work as an example and as a vindication of the doctrine of historical materialism. However, *Capital* deals mainly with the economic and industrial aspects of capitalism and all too briefly with political and ideological matters. It is not surprising that economic and industrial matters should play a large part in an analysis and history of economic and industrial developments. But *Capital* gives only minute and incidental support to the main thesis of historical materialism: the thesis of the dependence of other social institutions upon the technical and economic ones and the thesis of the primary historical influence of technology and economics. After Marx's death Max Weber put forward the view that the growth of capitalism in Europe was fostered by certain aspects of Protestant religious belief. Marx, of course, thought that religious belief is ideological and epiphenomenal, an ineffectual shadow of social reality. He would have found it necessary to reject Weber's view on grounds of principle, in spite of the concomitances and assimilations to which Weber called attention. This shows that Marx's view is not a hypothesis but part of a system of interpretation of very wide scope; part, indeed, of a philosophical outlook.

DIALECTICAL ASPECTS OF THE THEORY

The fundamental thesis of Marxist dialectics is that everything is in movement, and Marx and his followers have proclaimed the mutability of all existing social forms. This in itself, of course, would not distinguish historical materialism from, for example, Hegelianism or some types of liberalism. Another feature of Marxist dialectics, however, is the belief that although gradual changes are occurring all the time, there are also on occasion sudden changes of great scope in which existing types of being are succeeded by utterly new ones. This means that Marxists consider the emergence of new social forms to be as natural as evolutionary adaptation. One might say that their view of change is such as to make them expect the unexpected. A further tenet of Marxist dialectics is that development takes place through the clash of opposites.

Thus the doctrine of the class struggle is regarded by Marxists as a vital feature of historical materialism. Changes in the means of production provide the clue to class struggles and social revolutions out of which new forms of life and thought are born. Philosophers of the Marxist-Leninist tradition hold that in communist society contradictions and oppositions would continue but that, in the absence of class differences, they would be "nonantagonistic."

The foregoing might be called the metaphysics of Marxist dialectics. Marx himself, however, was much more concerned with dialectics as a method. Perhaps the most fundamental feature of the dialectical method as understood by Marx is its distrust of abstraction. This, too, is a Hegelian legacy, but whereas Hegel regarded the Absolute Spirit as the concrete reality, for Marx reality was the material world, along with embodied human beings organized together in various social orders. Philosophers who talk of spirit, or economists who talk of land, labor, and capital, according to Marx, obscure the physical basis of human life and action and substitute abstract categories for the concrete realities of human work and association. Abstraction, in this view of the matter, is a form of mystification. The only way to avoid mystification is to relate the things that people say and do to the material circumstances in which they live. But the abstract is contrasted not only with the concrete but also with what is whole or complete. Marx, like Hegel, thought that the parts of any whole were not indifferent to one another but were, on the contrary, linked closely together. This linkage was particularly close between the individuals and groups of human society. According to Marx, the institutions of work and production were the primary ones, but through their connection with these institutions, men's laws and politics, their philosophy, morals, art, and religion are interrelated and interdependent and cannot be understood in isolation from one another or from their material basis.

A further form of abstraction that Marx objected to was the claim that there are economic laws that apply to all human societies equally. Marx held (preface to *Capital,* Vol. I, 2nd ed.) that each main type of social order develops and functions in its own special ways, so that we cannot conclude from what happens in one type of society that anything similar will happen in another. Indeed, he said that to trace the laws of development of different types of society in this way, keeping the particular and peculiar in view, is the dialectical method. It should be noted, too, that Marx sometimes thought that the various social categories, such as productive forces and productive relations, could not be abstracted from one another, but collapsed one into the other, as Hegelian theories do. We have already seen that Marx treated forms of organization as means of production, thus blurring the distinction between productive forces and productive relationships. In the recently published *Outlines of a Critique of Political Economy* (1857) appears the following note: "Dialectic of the concepts productive force (means of production) and productive relationship, a dialectic to determine their limits, and which does not cancel their real distinction" (p. 29). It seems that Marx hoped to settle the problem by means of a dialectical *coup de main.*

RELATION TO OTHER ENDEAVORS

Marx was not the first to inquire into the history of technology and of industry and commerce, but undoubtedly his work greatly influenced the direction taken by historical research. Marxist historians have been particularly anxious to show how knowledge has been hindered or promoted by the prevailing productive forces and productive relationships. Thus, Benjamin Farrington, in his *Greek Science* (2 vols., London, 1944–1949), argued that the predominantly speculative and unpractical character of Greek science was due to the institution of slavery and the aristocratic contempt for manual work that went with it. George Thomson, in his *Studies in Ancient Greek Society,* 1: *The Prehistoric Aegean* (London, 1949), presented evidence in favor of Engels's views on primitive communism. In Volume II of the same work, subtitled *The First Philosophers* (London, 1955), Thomson linked the categories employed by the pre-Socratic philosophers with economic and class factors and with Marx's notion of a commodity as "the uniform socially recognized" incarnation of human labor, concluding that "the Parmenidean One, together with the later idea of 'substance,' may therefore be described as a reflex or projection of the substance of exchange value" (p. 103). B. Hessen, in an essay titled "The Social and Economic Roots of Newton's *Principia*" (*Science at the Crossroads,* 1931), argued that Isaac Newton was the typical representative of the rising bourgeoisie, and in his philosophy he embodies the characteristic features of his class" (p. 33). This type of view illustrates the more general inquiry into the connections between class and knowledge known as the sociology of knowledge. Karl Mannheim's *Ideology and Utopia* (*Ideologie und Utopie,* Bonn, 1929; translated by Louis Wirth and Edward Shils, London, 1936) shows how Marxism influenced this subject, but Max Scheler, who was not a Marxist, also helped develop it (*Die Wissenformen und die Gesellschaft,* Leipzig, 1926).

It should be emphasized that a materialist view of history is not necessarily linked with Marxist socialism, for it is possible to recognize the historical importance of the means of production and of economic and class interests without concluding that a classless, communist society must emerge. (This was done, for instance, by E. R. A. Seligman in *The Economic Interpretation of History,* New York, 1902). Furthermore, some historians and economists have adopted an economic interpretation of history without committing themselves to the Marxist views about the dominating influence of technology, of the means of production. Thus, Thorold Rogers, an undogmatic free trader, called attention to such influences as the shortage of labor created by the Black Death or the interference with trade routes by the Mongol invaders, but said: "You cannot, of course, separate, except in thought, and then only with no little risk of confusion, economical from social and political facts" (*The Economic Interpretation of History,* London, 1888, p. 281). Marxists have often gone to considerable lengths to distinguish the economic from the materialist conception of history. Thus, the Russian Marxist historian M. N. Pokrovsky has been criticized by orthodox Marxists for placing too much emphasis on market considerations and too little on the influence of the means of production.

See also Aristotelianism; Communism; Croce, Benedetto; Dialectical Materialism; Engels, Friedrich; Hegel, Georg Wilhelm Friedrich; Ideology; Lenin, Vladimir Il'ich; Mannheim, Karl; Marx, Karl; Plekhanov, Georgii Valentinovich; Scheler, Max; Socialism.

Bibliography

HISTORICAL MATERIALIST WORKS

Marx

Misère de la philosophie. Brussels and Paris, 1847. Translated by H. Quelch as *The Poverty of Philosophy.* Chicago: C. H. Kerr, 1910.

Zur Kritik der politischen Ökonomie. Berlin, 1859. Translated by N. I. Stone as *A Contribution to the Critique of Political Economy.* Chicago: C. H. Kerr, 1904.

Das Kapital, 3 vols. Hamburg, 1867–1894. Vols. II and III edited by Friedrich Engels; Vol. III in two parts. English translation, Chicago, 1915: C. H. Kerr. Vol. I translated by Samuel Moore and Edward Aveling, revised and amplified according to the fourth German edition by Ernest Untermann; Vol. II translated from the second German edition by Ernest Untermann; Vol. III translated from the first German edition by Ernest Untermann.

"Oekonomische-philosophische Manuskripte." In *Marx-Engels Gesamtausgabe,* edited by D. Riazanov and V. Adoratski. Berlin, 1932. Division 1, Vol. III. Translated by Martin Milligan as *Economic and Philosophical Manuscripts of 1844.* Moscow: Foreign Languages Publishing House; and London, 1959.

Grundrisse der Kritik der politischen Ökonomie (Rohentwurf). Berlin, 1953. One section has been translated by J. Cohen as *Pre-Capitalist Economic Formations,* edited by E. J. Hobsbawn. London: Lawrence and Wishart, 1964.

Karl Marx. Selected Writings in Sociology and Social Philosophy. Edited by T. B. Bottomore and Maximilien Rubel. London: Watts, 1956.

Marx and Engels

Manifest der kommunistischen Partei. London, 1848. Many English editions as *The Communist Manifesto.* See especially edition with an introduction by Harold Laski. London: Allen and Unwin, 1948.

Die deutsche Ideologie. Edited by V. Adoratski. Vienna, 1932. Translated as *The German Ideology,* edited by R. Pascal. London: Lawrence and Wishart, 1938.

Selected Correspondence (1846–1895). Edited by Dona Torr. New York: International, 1942. Contains Marx's 1846 letter to Paul Annenkov and important letters written by Engels after Marx's death.

Karl Marx, Frederick Engels: Collected Works, 49+ vols. Translated by Richard Dixon et al. New York: International, 1975–.

Engels

Umrisse zu Kritik der Nationalökonomie (1844). Translated by Martin Milligan as "Outlines of a Critique of Political Economy." In Karl Marx, *Economic and Philosophical Manuscripts of 1844* (op. cit.).

Herr Eugen Dührings Umwälzung der Wissenschaft. Leipzig, 1878. Translated by E. Burns as *Herr Eugen Dühring's Revolution in Science.* London: Lawrence and Wishart, 1934.

Die Entwicklung des Sozialismus von der Utopie zur Wissenschaft. Zürich, 1883. Translated by E. Aveling as *Socialism: Utopian and Scientific.* London: Sonnenschein, 1892.

Der Ursprung der Familie, des Privateigentums und des Staats. Zürich, 1884. Translated by E. Untermann as *The Origin of the Family, Private Property, and the State.* Chicago: C.H. Kerr, 1902.

Lenin

What the "Friends of the People" Are (1894). In his *Collected Works,* Moscow and London, 1960–. Vol. I, 1960.

"Karl Marx." In his *Collected Works.* Moscow and London, 1960–. Vol. 21, 1964.

Plekhanov

K Voprosu o Razvitii Monisticheskago Vzglyada na Istoriyu. St. Petersburg, 1895. Translated by Andrew Rothstein as *In Defence of Materialism: The Development of the Monist View of History.* London: Lawrence and Wishart, 1947.

Osnovnye Problemy Marksizma. 1908. Translated by E. Paul and C. Paul as *Fundamental Problems of Marxism.* London: Lawrence, 1929.

WORKS ON HISTORICAL MATERIALISM

Acton, H. B. *The Illusion of the Epoch: Marxism-Leninism as a Philosophical Creed.* London: Cohen and West, 1955.

Adler, Max. *Lehrbuch der materialistischen Geschichtsauffassung,* 2 vols. Vienna, 1930.

Barth, P. *Die Philosophie der Geschichte als Soziologie,* 3rd and 4th eds. Leipzig: Reisland, 1922.

Berlin, Isaiah. *Karl Marx.* New York: Oxford University Press, 1948.

Bober, M. M. *Marx's Interpretation of History,* 2nd ed. Cambridge, MA: Harvard University Press, 1950.

Clark, G. N. *Science and Social Welfare in the Age of Newton.* Oxford: Clarendon Press, 1937.

Croce, Benedetto. *Materialismo storico ed economica marxista.* Palermo, 1900. Translated by C. M. Meredith as *Historical Materialism and the Economics of Karl Marx.* London: Latimer, 1914.

Federn, Karl. *The Materialist Conception of History. A Critical Analysis.* London: Macmillan, 1939.

Goldmann, Lucien. *Recherches dialectiques.* Paris: Gallimard, 1959.

Hexter, J. H. *Reappraisals in History.* Evanston, IL: Northwestern University Press, 1961.

Hook, Sidney. *From Hegel to Marx.* New York: Reynal and Hitchcock, 1935.

Hook, Sidney. *Towards the Understanding of Karl Marx.* New York: John Day, 1933.

Masaryk, T. G. *Die philosophische und soziologische Grundlagen des Marxismus.* Vienna: Konegen, 1899.

Plamenatz, John. *German Marxism and Russian Communism.* London: Longman, 1954.

Pokrovsky, M. N. *A Brief History of Russia,* 2 vols. London, 1933.

Pokrovsky, M. N. *History of Russia from the Earliest Times to the Rise of Commercial Capitalism.* London: Lawrence, 1932.

Popper, K. R. *The Open Society and Its Enemies,* Vol. 2: *The High Tide of Prophecy.* London: Routledge, 1945.

Rubel, Maximilien. *Karl Marx. Essai de biographie intellectuelle.* Paris, 1957.

Tucker, Robert. *Philosophy and Myth in Karl Marx.* Cambridge, U.K.: Cambridge University Press, 1961.

Venable, Vernon. *Human Nature. The Marxian View.* London: Dobson, 1946.

ADDITIONAL SOURCES

Acton, H. B. *What Marx Really Said.* New York: Schocken Books, 1967.

Althusser, Louis. *Lenin and Philosophy.* New York: Monthly Review Press, 1972.

Armin, Samir. *The Law of Value and Historical Materialism.*New York: Monthly Review Press, 1978.

Ball, Terence, and James Farr, eds. *After Marx.* Cambridge, U.K.: Cambridge University Press, 1984.

Cohen, G. *Karl Marx's Theory of History,* expanded ed. Princeton, NJ: Princeton University Press, 2001.

Dupré, Louis K. *Marx's Social Critique of Culture.* New Haven, CT: Yale University Press, 1983.

Elster, Jon. *Making Sense of Marx.* Cambridge, U.K.: Cambridge University Press, 1985.

Giddens, Anthony. *A Contemporary Critique of Historical Materialism,* 2nd edition. Stanford, CA: Stanford University Press, 1995.

Habermas, Jürgen. *Communication and the Evolution of Society.* Translated by Thomas McCarthy. Boston: Beacon Press, 1979.

Little, Daniel. *The Scientific Marx.* Minneapolis: University of Minnesota Press, 1986.

Rader, Melvin Miller. *Marx's Interpretation of History.* New York: Oxford University Press, 1979.

Rigby, S. H. *Marxism and History: A Critical Introduction,* 2nd ed. Manchester, U.K.: Manchester University Press, 1998.

Shaw, William H. *Marx's Theory of History.* Stanford, CA: Stanford University Press, 1978.

Thompson, E. P. *The Making of the English Working Class.* New York: Vintage Books, 1966.

Wood, Allen. *Karl Marx.* London: Routledge and Kegan Paul, 1981.

H. B. Acton (1967)
Bibliography updated by Philip Reed (2005)

HISTORICAL SCHOOL OF JURISPRUDENCE

The historical school of jurists was founded by Friedrich Karl von Savigny (1779–1861). Its central idea was that a nation's customary law is its truly living law and that the task of jurisprudence is to uncover this law and describe in historical studies its social provenience. As in other schools of thought, acceptance of this approach did not necessarily mean agreement on its theoretical or practical consequences.

GERMANY

To followers of Savigny the identification of law with custom and tradition and the *Volksgeist,* or genius peculiar to a nation or folk, generally meant a rejection of rationalism and natural law; a rejection of the notion of law as the command of the state or sovereign, and therefore a disparagement of legislation and codification; and a denial of the possibility of universally valid rights and duties and of the individual's possession of nonderivable and inalienable rights. In positive terms, historical jurisprudence identified law with the consciousness, or spirit, of a specific people. Law is "found" by the jurist and not "made" by the state or its organs. Law is a national or folk and not a political phenomenon; it is a social and not an individual production; like language, it cannot be abstracted from a particular people and its genius; it is a historical necessity and not an expression of will or reason, and therefore it cannot be transplanted.

In addition to Savigny, the historical school was probably influenced by Johann Gottfried Herder (1744–1803) and the romantic notions of folk culture, by the emphasis on tradition in the work of Edmund Burke (1729–1797), by the stress on historical continuity in the work of Gustav Hugo (1764–1844), and by the Hegelian conception of Spirit. In Germany, the main proponents of

historical jurisprudence were G. F. Puchta, Karl Friedrich Eichhorn, Rudolph von Sohm, and Otto von Gierke.

ENGLAND

In England Henry Maine (1822–1888) was closely identified with the historical school, although there is no evidence that he was directly influenced by the German thinkers. Modern historical jurisprudence in England was born with the publication in London of Maine's *Ancient Law* in 1861, the year of Savigny's death. Until then historical research in law had been neglected, but from that time on, the field was assiduously cultivated. In reaction against natural law and under the influence of Thomas Hobbes, the tendency in England had been to regard law as the command of the state, and the task of the jurist was conceived as a concern with the analysis of positive law without regard to historical or ethical considerations. Maine broke with these traditional attitudes. Probably influenced by Rudolf von Ihering (*Der Geist des römischen Recht,* 3 vols., Leipzig, 1852–1865), Maine was stimulated to apply the historical method to jurisprudence. Charles Darwin's *Origin of Species,* published two years before *Ancient Law,* also probably influenced Maine.

Maine rejected the natural law, rationalistic, and a priori approaches to the nature of law. In his *Early History of Institutions* (London, 1875) he saw a people's law as compounded of opinions, beliefs, and superstitions produced by institutions and human nature as they affected one another. Indeed, English common law seemed better to exemplify Savigny's views than did the law of Germany, which drew heavily on Roman law. But as an Englishman, Maine saw in law more than a people's customs; he observed and took into account the creative and reforming work of Parliament, and so he was led to recognize legislation as an instrument of legal growth. And he found that equity and legal fictions played creative roles in the common law. In these respects he departed radically from Savigny's monistic approach to law and its sources.

Maine's comparative historical studies, which took into account diverse legal systems, kept him from a belief in the mystical uniqueness of a people and its genius and its law; he observed uniformities as well as differences in different legal orders, and so he was led to suggest that similar stages of social development may be correlated with similar stages of legal development in different nations. Maine differed from Savigny also in believing that custom might historically follow an act of judgment, so that the jurist could be seen to have had a creative role in making the law, even though he claimed only to have found it. Maine also noted the part played in early societies by the codification of customary law. In revealing the ideals operative in a society at a particular stage of its development and in relating them to social conditions, Maine stimulated the development of the use of the sociological method in jurisprudence. It thus became apparent that just as law cannot be divorced from history, so, too, it cannot be divorced from philosophy and sociology. Thus, if Savigny's historical jurisprudence was mainly conservative in import, Maine's work had a predominantly liberalizing effect. Then too, Maine's work influenced the development of comparative legal studies.

Other English scholars associated in varying degree with the historical school of jurisprudence are James Bryce (1838–1922), Frederic W. Maitland (1850–1906), Frederick Pollock (1845–1937), and Paul Vinogradoff (1854–1925).

Perhaps the greatness of historical jurisprudence lay in the fact that it provided its own seed of dissolution; for once it is admitted that law is historically conditioned, it is as impossible to limit the conception of law to a *Volksgeist* as to the commands of the sovereign; all forms of social control and all sources of law emerge as subjects for legitimate consideration and study.

See also Burke, Edmund; Darwin, Charles Robert; Herder, Johann Gottfried; Hobbes, Thomas; Legal Positivism; Philosophy of Law, History of; Philosophy of Law, Problems of; Savigny, Friedrich Karl von.

Bibliography

WORKS OF HISTORICAL JURISPRUDENCE

Eichhorn, Karl Friedrich. *Deutsche Staats- und Rechtsgeschichte,* 4 vols. Göttingen: Vandenhoeck and Ruprecht, 1808–1823.

Gierke, Otto von. *Das deutsche Genossenschaftsrecht,* 4 vols. Berlin: Weidmann, 1868–1913. Partially translated by Ernest Barker as *Natural Law and the Theory of Society.* Oxford: Cambridge University Press, 1934.

Puchta, G. F. *Das Gewohnheitsrecht,* 2 vols. Erlangen: Palm, 1828–1837.

Sohm, Rudolf von. *Institutionen des romischen Rechts.* Leipzig, 1883. Translated by J. C. Ledlie as *Institutes of Roman Law.* Oxford: Clarendon Press, 1892; 3rd ed., Oxford, 1907.

COMMENTARIES

Allen, C. K. *Law in the Making,* 7th ed. Oxford: Clarendon Press, 1964. Ch. 2.

Friedmann, Wolfgang. *Legal Theory,* 3rd ed. London: Stevens, 1953. Ch. 14, with bibliography.

Friedrich, C. J. *The Philosophy of Law in Historical Perspective,* 2nd ed. Chicago: University of Chicago Press, 1963. Ch. 15.

Jones, J. W. *Historical Introduction to Theory of Law.* Oxford, 1940. Ch. 2.

Pound, Roscoe. *Interpretations of Legal History.* New York: Macmillan, 1923.

Pound, Roscoe. *Jurisprudence.* St. Paul, MN: West, 1959. Vol. I.
Stammler, Rudolf. "Fundamental Tendencies in Modern Jurisprudence." *Michigan Law Review* 21 (1923): 623ff.
Stone, Julius. *The Province and Function of Law.* Cambridge, MA: Harvard University Press, 1950. Ch. 18.
Vinogradoff, Paul. *Outlines of Historical Jurisprudence.* London: Oxford University Press, 1920. Vol. I.
Walton, F. P. "Historical School of Jurisprudence and Transplantations of Law." *Journal of Comparative Legislation & International Law* 183 (3rd series, 1927).

ADDITIONAL SOURCES

Allison, J. W. F. *A Continental Distinction in the Common Law: A Historical and Comparative Perspective on English Public Law.* Oxford: Oxford University Press, 1996.
Diamond, Alan, ed. *The Victorian Achievement of Sir Henry Maine: A Centennial Reappraisal.* Cambridge, U.K.: Cambridge University Press, 1991.
Stein, Peter. *Legal Evolution: The Story of an Idea.* Cambridge, U.K.: Cambridge University Press, 1980.

Milton R. Konvitz (1967)
Bibliography updated by Philip Reed (2005)

HISTORICISM

The early history of the term "historicism" (*Historismus*) has not been sufficiently explored, as Erich Rothacker has pointed out. However, one clear case in which it was used in a sense closely allied to all of the senses which it has subsequently assumed is to be found in Carl Prantl's *Die gegenwärtige Aufgabe der Philosophie* (1852). Although the term was later employed as a means of characterizing the thought of Giambattista Vico, its first widespread use probably dates from methodological debates among German-speaking political economists. In these debates, Carl Menger criticized Gustav Schmoller and his school for making economic theory unduly dependent upon economic history; this he characterized as *Historismus.* Thus, the term took on a depreciatory sense; it suggested an inappropriate use of historical knowledge and a confusion regarding the sorts of questions that could be answered by means of such knowledge. One may conjecture that the extension of its use during the first decades of the twentieth century was fostered by the currency of its depreciatory analogue, "psychologism" (*Psychologismus*): Both terms were used in reference to attempts to extend the methods and results of a particular discipline into provinces in which that discipline was claimed to lack legitimate authority.

It was not until the period immediately following World War I, however, that *Historismus* came to be widely used. The impact of the war and the consequences of the German defeat led to attempts to reappraise the cultural and political traditions of the past, and in this reappraisal a central issue was whether a purely historical approach to human culture provided an adequate basis for the judgment of cultural values. This was not, of course, a new problem for theologians or for philosophers; it was one which had been forced upon their attention by dominant strains in nineteenth-century thought (for example, by Hegelianism, the results of historical biblical criticism, and evolutionism). Nevertheless, for those in Germany who had been reared in the tradition of historical studies and who were encountering the violent upheaval of the times, the question of the relations of cultural standards to historical change took on great immediacy. It was at this point that Ernst Troeltsch attempted to characterize historicism in a nonpolemical way, to examine its origins, and to assess its merits and limitations.

TROELTSCH

In *Der Historismus und Seine Probleme,* Troeltsch used "historicism" to mean a tendency to view all knowledge and all forms of experience in a context of historical change. He regarded this tendency as one of the two fundamental discoveries of the modern mind: The other, with which he compared it, was the generalizing, quantitative approach to nature that he termed *Naturalismus.* Thus, like Wilhelm Dilthey, Wilhelm Windelband, Heinrich Rickert, and others, Troeltsch drew a distinction between the forms of understanding characteristic of the natural sciences and those which are appropriate to what one may perhaps best term the "historical sciences" (*die Geisteswissenschaften*). What was of prime importance to him, however, was not the differences between the methodologies of the natural and the historical sciences, but the fact that each was a fundamentally different way of looking at the world, that is, each constituted a different *Weltanschauung.* Troeltsch documented the scope and the depth of historicism as a *Weltanschauung* by tracing its presence in the thought of a host of philosophers and sociologists of the nineteenth and twentieth centuries. He himself accepted the view that all knowledge and all forms of human experience are caught up in a process of change; however, he believed that this view tended to lead to an unmitigated moral and intellectual skepticism. It was this that constituted the crisis of historicism, and it was this that he sought to overcome. Unlike Rickert and others among his contemporaries, he believed that the skeptical consequences of historicism could be overcome only through history itself and could not be avoided by any appeal to transhistorical values. His own positive, religiously based views, however, received only partial

expression, for he died before he was able to complete the work that he had projected.

MANNHEIM

In 1924, almost immediately after the appearance of Troeltsch's work, Karl Mannheim wrote an essay, "Historismus," in which he too characterized historicism as a basic *Weltanschauung.* According to him, the static, theologically oriented conception of the world that characterized the Middle Ages had been retained in secularized form in the Enlightenment, because both cultures held to the doctrine of the atemporal character of the judgments of reason. According to Mannheim, this static conception had at last been abandoned, and all social and cultural reality was seen as being dominated by change. It was this radically temporalistic view of the world that he designated as historicism. Unlike Troeltsch, to whose work he devoted a portion of his essay, Mannheim did not recoil from the relativism of values that he saw that historicism entailed; rather, he was concerned to affirm it. However, on the basis of his own views regarding the intimate connections between theory and practice, he did not believe that either moral or intellectual skepticism was a necessary consequence of temporalistic relativism. Moral skepticism would not necessarily follow, since Mannheim believed that all values are rooted in the conditions of actual social existence and their discovery is not dependent upon our possession of some unchanging capacity for moral insight; furthermore, intellectual skepticism could be avoided through a recognition of the perspectival character of knowledge, and by means of the capacity of a sociology of knowledge to uncover the nature of divergent perspectives and reconcile them with one another. Thus, in Mannheim's use of "historicism," unlike Troeltsch's, there remained no vestige of the original depreciatory significance of the term.

MEINECKE

In 1936 Friedrich Meinecke published a historical study titled *Die Entstehung des Historismus* in which the term assumed a markedly different connotation. To be sure, Meinecke shared Troeltsch's view that historicism represented a break with those modes of thought which both characterized as naturalism. Furthermore, like Mannheim and others, he believed that there was a fundamental opposition between the modern historical sense and earlier political philosophies that had relied upon the conception of a universal and unchanging natural law as the basis for moral and political judgment. Thus Meinecke regarded historicism as opposed to a static view of the world, and in this he was in agreement with Troeltsch and Mannheim. However, he proceeded to characterize this new world view in terms of an interest in that which is concrete, unique, and individual; he found the clue to the new view expressed in Johann Wolfgang von Goethe's use of the dictum "*Individuum est ineffabile.*" This characterization of historicism was undoubtedly related to the fact that Troeltsch (among others) had viewed historical inquiry as concerned with the concrete, the unique, and the individual, and had contrasted this interest with the methods used in the natural sciences.

However, in translating this particular methodological doctrine into a worldview, Meinecke departed radically from the characterizations offered by Troeltsch and Mannheim. For them it was not the concept of individuality but the concepts of change and development that were fundamental to what they had termed "historicism." As a consequence of this difference in the meaning of the terms, some of the eighteenth-century historians who played dominant roles in Meinecke's account would not have been considered proponents of historicism by Troeltsch or by Mannheim. The difference emerges most strongly in the fact that Meinecke believed the culmination of modern historicism was to be found in the world views of Goethe and Leopold von Ranke, whereas one would expect such a high point to be identified with G. W. F. Hegel, with Karl Marx, or perhaps with later evolutionary thought, were one to take the term in the meaning ascribed to it by Troeltsch and Mannheim. As a consequence of this shift in the meaning of the term, Meinecke naturally did not regard historicism as a force that threatened human values or which could lead to a radical transvaluation of values; thus, for him there was no crisis of historicism as there had been for Troeltsch.

CROCE

The view with which Meinecke's characterization of historicism can best be compared is that of Benedetto Croce, even though Croce criticized Meinecke's work for its failure to emphasize nineteenth-century thought, and in particular because of its failure to appreciate Hegel's importance. Croce's own philosophic views had grown out of a reaction against positivism and materialism, in favor of idealism: in particular, he concerned himself with combating positivist and materialist philosophies of history. What he rejected in these views was not the historicism that Troeltsch and Mannheim correctly discerned in them, but the fact that they attempted to interpret history naturalistically, that is; in ways similar to those used by the sciences in dealing with the nonhuman world. Like Vico and Hegel, with whose thought his own

was directly affiliated, Croce regarded history as the self-development of the human spirit. Furthermore, since Croce, as an idealist, wished to deny that there was any realm of existence external to the human spirit, he interpreted the whole of reality as being encompassed within history: life and reality were nothing but the ever changing manifestations of the spirit.

It was primarily with reference to this radical metaphysical idealism, rather than with reference to any more general currents in Western intellectual history, that Croce used the term "historicism" (*storicismo*). While Croce's own emphasis on the pervasiveness of change did in fact provide an example of what Troeltsch and Mannheim considered to be the basic feature of historicism, it was not with their thought, but with that of Meinecke, that his views had the greater affinity. Like Meinecke, Croce held that the means by which a naturalistic worldview seeks to envision and grasp reality are totally inadequate because of the uniqueness and individuality of that which is historical. He therefore held—as did Meinecke—that genuine knowledge, as opposed to merely practical or pseudo-knowledge, comes only through an understanding of history. Croce endeavored to establish this antinaturalistic position throughout his philosophical writings; for Meinecke, the acceptance of this form of historicism was intimately connected with a religious sense of mystery.

ENGLAND AND THE UNITED STATES

The term "historicism" was adopted into the English language in the late 1930s and the 1940s both in the United States and in England. In neither country, however, was it used to refer primarily to a *Weltanschauung*; rather, what was of concern were questions regarding principles of explanation and of evaluation. In the United States, attention was directed to these issues through works by Morris R. Cohen, Maurice Mandelbaum, and Morton White, among others. In England, fuller discussions were to be found in articles by F. A. Hayek and Karl Popper.

One may plausibly infer from Hayek's discussion of historicism that the sense in which he and Popper conceived the notion probably derived from Menger's original contrast between scientific theory-construction and a primarily historical approach to problems in the social sciences. However, the specific form of historicism that both Hayek and Popper especially attacked was the nineteenth-century doctrine that there are laws of development that characterize social wholes and that it is possible, on the basis of a knowledge of such laws, to make scientific predictions about the future. Thus, the notion of "holism," which had not previously been directly associated with the definition of historicism, was injected into the discussion, and the chief protagonists of historicism were identified as Hegel, Auguste Comte, and Marx. When taken in this sense, three theses were common to historicist doctrines: (1) a rejection of "methodological individualism" in favor of the view that there are social wholes which are not reducible to the activities of individuals; (2) the doctrine that there are laws of development of these wholes, considered as wholes; (3) the belief that such laws permit predictions as to the course which the future will take. While these three theses were intimately connected with some of the doctrines previously characterized as examples of historicism, there seems to be no necessity for identifying historicism with holistic thought and with a belief in the possibility of prediction, as Popper and Hayek tend to do.

DEFINITION OF "HISTORICISM"

Considering the very great diversity in usage which we have now traced, one may ask whether there is any characterization of historicism which can serve to connect the various ways in which the term has been used and which at the same time can give it a relatively clear meaning. Without suggesting that all problems concerning the deviant meanings of historicism can be solved in this way, the following definition may be proposed as an approximation of that goal: Historicism is the belief that an adequate understanding of the nature of anything and an adequate assessment of its value are to be gained by considering it in terms of the place it occupied and the role it played within a process of development.

It will be noted that this definition does not characterize historicism as a particular *Weltanschauung* but as a methodological belief concerning explanation and evaluation. As Popper's discussion makes clear, in the late eighteenth and nineteenth centuries, forms of what has been termed "naturalism" have closely resembled antinaturalistic theories, with respect to their presuppositions about the relation of historical change to the explanation and evaluation of events. Since it is misleading to regard positions as divergent as those of, say, Hegel, Comte, Marx, and Herbert Spencer as representative of one and the same *Weltanschauung,* it is preferable to conceive of historicism as a methodological principle.

Troeltsch and Mannheim were in agreement with Meinecke and Croce in holding that this new methodological principle was based upon the rise of a new concept of change and of history. Its original challenge to older modes of thought lay partly in its tendency to link

evaluation with genetic explanation. It was this tendency that was fundamental to the so-called crisis of historicism, and it has also been against this tendency that Hayek and Popper, among others, subsequently rebelled. However, the most radical aspect of historicism as a methodological principle has been its conception of what is presupposed in all explanations and evaluations of past events: that each event is to be understood by viewing it in terms of a larger process of which it was a phase, or in which it played a part; and that only through understanding the nature of this process can one fully understand or evaluate concrete events. It is partly because of this emphasis upon relating each event to some larger developmental process that historicism has come to be identified with holism and a belief in historical prediction. Important as this connection has undoubtedly been, a definition in terms of it fails to stress the more fundamental fact that historicism involves a genetic model of explanation and an attempt to base all evaluation upon the nature of the historical process itself. Popper, in his characterization of the position, therefore tends to separate his own use of the term "historicism" from its other, more frequent uses. The definition suggested here constitutes an attempt to epitomize many of these uses and to connect them with one another even where they are found to diverge.

See also Cohen, Morris Raphael; Comte, Auguste; Croce, Benedetto; Dilthey, Wilhelm; Enlightenment; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hegelianism; Holism and Individualism in History and Social Science; Idealism; Mannheim, Karl; Marx, Karl; Meinecke, Friedrich; Moral Skepticism; Popper, Karl Raimund; Rickert, Heinrich; Troeltsch, Ernst; Vico, Giambattista; Windelband, Wilhelm.

Bibliography

The following bibliography is highly selective and deals only with those works in which historicism is the central topic of discussion and with works specifically mentioned in the article. The order is chronological, except that all works by the same author or concerned primarily with that author are placed together.

Prantl, Carl. *Die gegenwärtige Aufgabe der Philosophie.* Munich, 1852. A speech given before the (Bavarian) Akademie der Wissenschaften on March 27, 1852. In it, the term *Historismus* is used to designate a concrete, historically oriented method by means of which the impasses in the philosophy of the day could be avoided.

Menger, Carl. *Untersuchungen über die Methode der Socialwissenschaften.* Leipzig: Duncker and Humblot, 1883. Reprinted as Volume II (London, 1933) of the *Collected Works.* These four volumes constitute numbers 17–20 of "Reprints of Scarce Tracts in Economics and Political Science," published by the London School of Economics and Political Science.

Menger, Carl. *Die Irrthümer des Historismus in der deutschen Nationalökonomie.* Vienna, 1884. Reprinted in Volume III (London, 1935) of the *Collected Works,* 4 vols. A polemic against Schmoller's methodological views. This and the preceding item represent the first influential introduction of the term "historicism" into the literature of the social sciences.

Ritzel, Gerhard. *Schmoller versus Menger: Eine Analyse des Methodenstreits im Hinblick auf den Historismus in der Nationalökonomie.* Frankfurt am Main, 1950. A helpful doctoral dissertation concerned with the issues involved in the quarrel over method in political economy.

Troeltsch, Ernst. *Der Historismus und seine Probleme.* Tübingen: Mohr, 1922. Only the first volume of this work was completed. It takes its point of departure from the "crisis of historicism" in postwar German intellectual life and treats fundamental problems of both formal and material philosophies of history through detailed discussion of almost all important nineteenth- and twentieth-century figures in these fields. Unfortunately, the structure of the discussion is not clear, largely because Troeltsch used the concept of "value" in discussing both the formal problems of a philosophy of history and the problem of meaning in history. His own solution of the latter problem, which was to have been the theme of his second volume, is briefly adumbrated in posthumously published lectures that he was to have given at Oxford. They were originally published as *Christian Thought, Its History and Application* (London: University of London Press, 1923; reprinted, New York, 1957), but they have been better known under the title *Der Historismus und seine Überwindung* (Berlin: Teubner, 1924).

Troeltsch, Ernst. "Die Krisis des Historismus." *Die Neue Rundschau* 33 (1922): 572–590. An excellent, brief, untechnical statement of the central theme of Troeltsch's classic work.

Hintze, Otto. "Troeltsch und die Probleme des Historismus." *Historische Zeitschrift* 135 (1927): 188–239. Also to be found in Hintze's collected essays, *Zur Theorie der Geschichte* (Leipzig: Koehler and Amelang, 1942). A masterful analytical and critical exposition of Troeltsch's approach to historicism. It is also suggestive in the way in which it provides a rich historiographical background against which the work of Troeltsch and Meinecke can be viewed.

Mannheim, Karl. "Historismus." *Archiv für Sozialwissenschaft und Sozialpolitik* 52 (1924), translated and edited by Paul Kecskemeti and reprinted in *Essays on the Sociology of Knowledge* (London: Routledge, 1952). An important treatment of historicism as the dominant *Weltanschauung* of the post-Enlightenment period, and an attempt to work out its cultural and philosophic implications.

Croce, Benedetto. "Antistoricismo." In *Proceedings of the Seventh International Congress of Philosophy.* London: H. Milford, 1931. Translated into German by K. Vossler, and published in *Historische Zeitschrift* 143 (1931): 457–466. An attack upon what Croce regarded as two antihistorical tendencies in thought: "futurism" and the attempt to escape from historicity to fixed and absolute standards.

Croce, Benedetto. *La storia come pensiero e come axione.* Bari: Laterza, 1938. English translation by Sylvia Sprigge as

History as the Story of Liberty (New York: Norton, 1941). A series of studies representing Croce's most mature reflections on history, historiography, and historicism.

Cohen, Morris R. "History versus Value." In *Reason and Nature,* 369–385. New York: Harcourt Brace, 1931. This essay was originally published in *Journal of Philosophy* 11 (1914): 704–716, but its influence on Morton White and others probably dates from its inclusion in this volume.

Heussi, Karl. *Die Krisis des Historismus.* Tübingen: Mohr, 1932. Contains a useful summary of early uses of the term and helpful references. Heussi's analysis of the concept should be consulted, but lacks an appreciation of the connection between the main currents of nineteenth-century thought and the ways in which the term was used in the twentieth century.

Meinecke, Friedrich. *Die Entstehung des Historismus,* 2 vols. Munich: Oldenbourg, 1936. The classic account of the development of historicism interpreted as a new view of life and of historical experience. Its course is traced by means of a series of individual essays from seventeenth-century forerunners through Goethe and Ranke.

Meinecke, Friedrich. *Zur Theorie und Philosophie der Geschichte,* Vol. 4 of *Werke,* 8 vols. Stuttgart: Koehler, 1957–1979. Collected essays and reviews, including discussions of Heussi and Troeltsch, and additional materials concerning the theme of his own book.

Hofer, Walther. *Geschichtschreibung und Weltanschauung.* Munich: Oldenbourg, 1950. A study of Meinecke's thought, with special reference to historicism.

Rothacker, Erich. "Historismus." *Schmollers Jahrbuch* 62 (1938): 388–399. A noted theorist of the *Geisteswissenschaften* offers an interpretation of the characteristic nature of historicism. Rothacker's "Das Wort 'Historismus,'" *Zeitschrift für Deutsche Wortforschung* 16 (1960): 3–6, is a criticism of those who failed to trace early uses of the term.

Mandelbaum, Maurice. *The Problem of Historical Knowledge.* New York: Liveright, 1938. A brief discussion of historicism is included.

Hayek, F. A. "Scientism and the Study of Society: The Historicism of the Scientistic Approach." In *The Counter-revolution of Science.* Glencoe, IL: Free Press, 1952, Part I, Ch. 7. First published in *Economica* 10 (1943): 50–63. The above study first appeared as a whole in *Economica* 9 (1942), 10 (1943), 11 (1944).

Popper, Karl R. *The Poverty of Historicism.* Boston: Beacon Press, 1957. First published in *Economica* 11 (1944), 12 (1945). (For remarks on its earlier composition, see Popper's "Historical Note" prefixed to the printed volume.) The exposition and attack upon historicism are intimately connected with Popper's theories concerning methodology in the social sciences. With respect to all of these topics, see also Karl Popper, *The Open Society and Its Enemies* (London: Routledge, 1945).

Engel-Janosi, Friedrich. "The Growth of German Historicism." *Johns Hopkins University Studies in History and Political Science,* series 62 (1944), no. 2. Offers a brief survey of a number of leading German-speaking historians, emphasizing the historicist elements in their work.

White, Morton G. "The Attack on Historical Method." *Journal of Philosophy* 42 (1945): 314–331. A reply to some aspects of the attack on historicism.

White, Morton G. *Social Thought in America: The Revolt against Formalism.* New York: Viking, 1949. Includes a characterization of historicism and an estimate of its role in the thought of selected social theorists.

Lee, Dwight E., and Robert N. Beck. "The Meaning of 'Historicism.'" *American Historical Review* 59 (1953–1954): 568–577. A compilation and critical classification of the various ways in which the term "historicism" has been used; it proposes a pair of definitions to cover the most widely established usages.

Antoni, Carlo. *Lo storicismo.* Rome, 1957. Adds some information to that supplied by Heussi on the early uses of the term; offers a brief survey of the development of historicism (in its broadest sense) from the eighteenth century to Croce.

White, Hayden V. "On History and Historicisms." In *From History to Sociology,* by Carlo Antoni; English translation by Hayden V. White. Detroit, MI: Wayne State University Press, 1959. The translator's introduction is a brief, useful historical sketch of types of historicism.

Rand, Calvin G. "Two Meanings of Historicism in the Writings of Dilthey, Troeltsch, and Meinecke." *Journal of the History of Ideas* 25 (1964): 503–518. A further attempt to elucidate the varied meanings of the term.

Maurice Mandelbaum (1967)

HISTORICISM [ADDENDUM]

The debate over the nature of historicism has not yet been resolved, and perhaps never will be; for historicism is a broad, multifaceted phenomenon with highly diverse and often conflicting manifestations. Indeed, the only relationship between the myriad meanings of the term "historicism" seems to be that of family resemblance (i.e. constantly shifting patterns of similarities and differences across the multiple instantiations of the concept).

Nevertheless, it is possible to identify, within the historicist family, one strand that has been significant for the history of philosophy more than any of its counterparts: namely, that which reached its apex in the enormous impact of Martin Heidegger and Hans-Georg Gadamer on the human sciences over the second half of the twentieth century. In this light, historicism's most crucial and enduring legacy appears to be the insistence that the historical character of human existence sets it apart from the world of nature, both in its ontological features and in the epistemological exigencies deriving therefrom. At the ontological level, this worldview unpacks the historicity of the human world in terms of the finitude, temporality,

uniqueness, and the meaning-laden character of all human phenomena. At the epistemological level, these ontological features render erroneous and futile any attempt to understand historical existence through the prism of universal laws, timeless causal mechanisms, or theoretical abstractions.

Ultimately, this outlook not only contests the aspirations of positivist epistemology to universal applicability, but in fact poses a comprehensive challenge to the universalistic worldview of the Enlightenment and its nineteenth- and twentieth-century successors. This major strand of historicism developed chiefly (though not exclusively) within the German intellectual world, with four prominent figures towering above all others in their contribution to historicism's unabated and dramatic impact upon Western thought: Leopold von Ranke, Wilhelm Dilthey, Heidegger, and Gadamer.

Leopold von Ranke (1795–1886), one of the founding fathers of modern historiography, developed his historical thought to a large extent through critical engagement with the idealist, Hegelian philosophy of history. Whereas the latter viewed history as the march of reason, entailing the mediation of historical knowledge by conceptual abstraction and generalization (that is, the subordination of history to philosophy), Ranke on the contrary upheld the primacy of history over philosophy; that is, the irreducibility of the full concreteness and diversity characterizing the flow of human existence to abstract, general categories. Ranke still embraces the idea of the unity of world history, but only in the sense that all historical epochs and phenomena are of equal value before God by virtue of their concrete individuality. Their universal meaning—their essence—resides within their particularity rather than in some general concepts to which they are purportedly reducible. Accordingly, the causal sequences governing historical life are of a contingent nature that has nothing to do with the mechanistic causality of the natural sciences.

Wilhelm Dilthey (1833–1911) sought to place the distinction made by Ranke and others between the natural sciences and the human (*qua* historical) sciences upon firmer, more systematic ontological and epistemological ground. In Dilthey's ontology, human life is construed as a meaningfully structured flow of interconnected experiences through time. These experiences are in turn objectivated in all visible human phenomena: works of art, political and legal formations, religious rituals, and so on. Thus, history consists in meaning-laden expressions of life.

This ontology clears the way for harnessing hermeneutical principles, originally utilized for the interpretation of individual texts, as a key to systematizing the epistemology of the entire human sciences. The possibility of historical knowledge flows, according to Dilthey (1989), from two complementary observations, both of which are hermeneutical by nature. First, the objectivation of life expressions in the realms of art, religion, law, and so forth, renders those expressions visible to the historian. Second, those expressions are not only visible but also intelligible to us by virtue of our ability to reproduce, to re-experience the meaning of objectivated expressions of life within our own field of experience. Dilthey's hermeneutics thus functions both as the basis for a distinctly historical epistemology and as a methodological procedure for extracting meaning out of history.

Although, as we have seen, historicism has always contained ontological as well as epistemological and methodological precepts, Martin Heidegger (1889–1976) was the first to transform historicism into an all-encompassing ontology. This accomplishment was effected through a reversal of the entire Western metaphysical tradition. Instead of searching for the eternal essences of Being underlying external, changing appearances, Heidegger (1996) proposes to construct a universal ontology based on temporality as its most fundamental attribute. Human existence, for Heidegger, is fully contained within its concrete, temporal manifestations; nothing which belongs to Being resides outside of this concrete temporality (that is, outside of historicity), in some abstract universal reason or divine spirit. Concomitantly, there is no need for abstract concepts or scientific classifications in order to grasp truth. Rather, the latter is immediately accessible to us in the form of experience, thus pointing up the ontological status of understanding as a ubiquitous form of (historical) being rather than as a merely cognitive process.

Heidegger's most influential follower, Hans-Georg Gadamer (1900–2002), devoted his magnum opus *Truth and Method* to the systematic development of an ontological hermeneutics based upon Heidegger's insight concerning the historicity of understanding. For Gadamer (1989), we are always, as historical beings, situated within a tradition but at the same time constantly working through that tradition, by way of an interpretive engagement with the texts comprising it. This is the so-called "hermeneutical circle": While tradition is the universal ontological condition within which all understanding occurs, that same tradition is also constantly being reformed in the course of the interpretive process, which in turn entails a perpetual remaking of the interpreter's self-understanding. Thus, the understanding subject, the process of understanding, and that which is understood

(i.e., tradition) are but three facets of one and the same ontological condition: the temporal flux of being.

Gadamer's thought may be regarded as the culmination of a historical process whereby historicism persistently broadened its scope, from the rather impressionistic and quasi-theological observations of Ranke on the nature of history and historiography, to the comprehensive historicist ontology expounded by Heidegger and Gadamer. At the same time, the protracted debate with the universalistic project of the Enlightenment has been accompanying historicism throughout all of its permutations. From Ranke's debate with Hegelianism, through Heidegger's famous encounter with Ernst Cassirer at Davos in 1929, to the Habermas-Gadamer debate, the unresolved battle between historicism and Enlightenment embodies one of the leitmotifs of modern civilization: namely, the perpetual intellectual, cultural, and political tension between the particularizing and the universalizing moments of the modern condition.

See also Philosophy of History.

Bibliography

Ankersmit, Frank R. "Historicism: An Attempt at Synthesis." *History and Theory* 34 (3) (1995): 143–161. This attempt, undertaken by one of today's foremost historical thinkers, to grasp the meaning of historicism serves to point up the importance of understanding this phenomenon in juxtaposition with Enlightenment thought.

Bambach, Charles R. *Heidegger, Dilthey, and the Crisis of Historicism*. Ithaca, NY: Cornell University Press, 1995. Discusses the vicissitudes of historicism within the German philosophical discipline during the late-nineteenth and early-twentieth centuries.

Dilthey, Wilhelm. *Selected Works*. Vol. 1, *Introduction to the Human Sciences*. Translated by Rudolf A. Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 1989.

Dilthey, Wilhelm. *Selected Works*. Vol. 3, *The Formation of the Historical World in the Human Sciences*. Translated by Rudolf A. Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 2002. These two works of Dilthey together contain the bulk of his historical thought.

Ermarth, Michael. *Wilhelm Dilthey: The Critique of Historical Reason*. Chicago: University of Chicago Press, 1978. The most systematic discussion of Dilthey's thought.

Gadamer, Hans-Georg. *Truth and Method*. 2nd rev. ed. Translated by Joel Weinsheimer and Donald G. Marshall. New York: Continuum, 1989. In his magnum opus, Gadamer not only lays out his approach in great detail, but also discusses (and clarifies his own position vis-à-vis) many of his predecessors in the historicist and hermeneutical traditions, including Ranke, Droysen, Dilthey, and Heidegger.

Hamilton, Paul. *Historicism*. 2nd ed. London: Routledge, 2003. A useful introduction to the subject, extending the scope of historicism to include Freud, Lacan, Foucault, and Derrida, and placing it in the context of postcolonialism and globalization. Includes a short glossary and suggestions for further reading.

Heidegger, Martin. *Being and Time*. Translated by Joan Stambaugh. Albany: State University of New York Press, 1996. The definitive, though opaque, statement of Heidegger's philosophy.

Iggers, Georg G. *The German Conception of History: The National Tradition of Historical Thought from Herder to the Present*. Rev. ed. Middletown, CT: Wesleyan University Press, 1983. Still the standard account of the historicism that developed within the German historical and social scientific disciplines. Emphasizes the nationalist, illiberal ethical-political dimension of German historical thought through the middle of the twentieth century.

Iggers, Georg G. "Historicism: The History and Meaning of the Term." *Journal of the History of Ideas* 56 (1) (1995): 129–152. An extremely useful recapitulation both of recent trends in the study of historicism and of this authoritative scholar's own understanding of the phenomenon. Includes extensive references to primary and secondary sources.

Krieger, Leonard. *Ranke: The Meaning of History*. Chicago: University of Chicago Press, 1977. The most systematic exposition of Ranke's historical thought.

Reill, Peter Hanns. *The German Enlightenment and the Rise of Historicism*. Berkeley: University of California Press, 1975. A detailed analysis of the complex relationship between historicism and the German Enlightenment.

Asaf Kedar (2005)

HISTORY

See *Determinism in History*; *Historicism*; *History and Historiography of Philosophy*; *Holism and Individualism in History and Social Science*; *Philosophy of History*

HISTORY, PHILOSOPHY OF

See *Philosophy of History*

HISTORY AND HISTORIOGRAPHY OF PHILOSOPHY

The term *history of philosophy* is often used in two different senses. In one, it refers to past events (*res gestae*) and, in another, to accounts of those events (*historiae rerum gestarum*). "The history of ancient Greek philosophy" can be taken to indicate views entertained by Greek philosophers, but also the accounts that later historians give of

those views. The positions Aristotle takes in his *Metaphysics* are part of the first but not of the second, whereas those adopted by Joseph Owens in *The Doctrine of Being in the Aristotelian Metaphysics* (1951) are part of the second but not the first.

The term *historiography of philosophy* can also be taken in two senses. According to one, it refers to accounts of past events, and so it is interchangeable with *history* when this term is used in the second sense mentioned above. But *historiography of philosophy* can also be used to mean the discipline that studies and establishes the procedures to be followed in accounts of the views from past philosophers. Aquinas's statement, "whatever is moved is moved by another," is part of the history and historiography of philosophy in the first sense mentioned. But the claim, "A proper understanding of Aquinas's view, that whatever is moved is moved by another, presented in the *Summa theologiae*, requires that we look into what he says about movement elsewhere in his writings," is part of historiography when this is understood as a discipline.

In addition, both the history and the historiography of philosophy need to be distinguished from the philosophy of the history of philosophy. This last studies the history of philosophy understood as past events in order to make claims about its nature and how it develops in general. In doing so, it may refer to particular events of that history, but its primary aim is not to account for them. For example, philosophers of the history of philosophy might claim that philosophy develops according to certain stages, but when they identify the stages through which ancient philosophy passed in particular, they do so to illustrate or establish the first kind of claim.

Because the history of philosophy, the historiography of philosophy, and the philosophy of the history of philosophy are closely connected, their tasks are not often distinguished and philosophers engaged in the pursuit of one also frequently pursue the others. For the sake of clarity, however, this entry will keep them separate, concentrating only on the issues pertaining to the historiography of philosophy when this is understood as the study of the procedures to be followed in the investigation of the philosophical past and of the philosophical issues that this kind of study raises.

Six of these issues have been the focus of most discussions: (1) What kind of claim are historians of philosophy entitled to make? (2) What is the relation between philosophy and the study of its history? (3) What is the value of the study of the history of philosophy for philosophy? (4) What is the role of texts in the study of the history of philosophy? (5) What approach should historians of philosophy use? And (6) what are the main genres historians of philosophy employ?

1. CLAIMS

Disagreements concerning the kind of claim that historians of philosophy are supposed to make center on three possibilities: descriptive, interpretative, and evaluative. A descriptive claim consists of a proposition that accurately (1) presents what particular philosophers said or thought or (2) recounts contemporaneous and later views concerning the positions of the philosophers under study. These claims take forms such as "X stated that P," "X's stating that P is the reason that X gave for holding Q," "M, a contemporary of X, stated that X did not hold that P," "N, a later historian of philosophy, disagreed with M as to X's view," and so on.

In interpretative claims, historians of philosophy go beyond what particular philosophers and their historians said or thought, in order to establish nonexplicit relations between the stated or unstated views of a philosopher or a historian, or between the views of two or more philosophers or historians. They also formulate broad generalizations that purport to characterize the overall approach used by a philosopher or the philosophers from a particular period, and to translate the views of historical figures into the languages and conceptual frameworks of contemporary historians in order effectively to communicate their meaning. Interpretative claims can take various forms, such as: "X held that Q," "X held that Q because X held that P," "X held that Q because Y held that P," "X's view that P led to the abandonment of ~P by her contemporaries," and so on.

Evaluative claims make judgments about the value of philosophical views from the past. These judgments may concern truth, validity, coherence, adequacy, completeness, clarity, social relevance, and so on. Here are some forms that these claims may take: "X's view, that P, is true," "X's argument A is invalid," "X and Y were right in formulating problem P as they did," "X's view that P is a backward step in the history of philosophy," "X's position had an adverse effect on society S," and so on.

The question pertaining to descriptive, interpretative, and evaluative claims that concerns historiographers in particular is the following: Are historians of philosophy supposed to make claims that are descriptive, interpretative, evaluative, or some combination of these? At one extreme, positivist historiographers answer that historians should consign themselves to descriptive claims. Their job is to describe, and not to interpret or to evaluate, the philosophical past (Lafrance 1983). At the other

extreme, historicist historiographers maintain that historians should merely be concerned with interpretation and evaluation because description is impossible. Every historical event is unique and cannot be reproduced either in reality or thought. Therefore, the attempt to describe and understand the past as it was in itself, independently of how it appears to the present, is bound to fail. The job of historians is to present the past as it looks to them at present (Collingwood 1946).

Both positivist and historicist historiographers accuse each other of betraying the historical enterprise. According to the first, the second do so because they fail to account for the past by falling into Anachronism, that is, reading the present into the past. But historicists retort that positivists betray history because they misunderstand the past by falling into Antiquarianism, that is, by failing to grasp the significance of the past for the present.

In between these two extreme positions, various positions attempt to find a more sensible middle ground. Closer to positivism is the view that the history of philosophy needs to be disinterested, that is, it should refrain from any kind of value judgment or interpretation based on value judgments (Garber 1988). Closer to historicism is the position that the history of philosophy should not be conceived as a science at all, but rather, like all philosophy, as a process of edification. Accordingly, it is its current uses and meaning that matter, not what actually happened in the past (Rorty 1984). Closer to the middle, some historiographers argue that historians of philosophy need to engage in description, interpretation, and evaluation: Description, because their aim is to understand and account for the past; interpretation, because the understanding and account of the past requires interpretation; and evaluation, because a history of philosophy without evaluation has no use (Gracia 1992).

2. PHILOSOPHY VS. HISTORY OF PHILOSOPHY

But what is the relationship between philosophy and the history of philosophy? Are they compatible enterprises? And if compatible, how dependent are they on each other? The attempts to answer these questions are plagued with puzzles and difficulties (Powers 1986).

The positions adopted with respect to these questions generally follow those adopted in the previous one. On one side are those historians who draw a sharp distinction between the descriptive aim of the historian of philosophy and the interpretative and evaluative aims followed by the philosopher. According to them, philosophy and the history of philosophy are incompatible insofar as the philosopher seeks to establish truth in general, whereas the historian of philosophy is merely interested in historical truth, that is, in arriving at accurate descriptions of the philosophical past. The historian studies the history of philosophy in its own terms, not for the philosophical truth it may yield (Frede 1988).

On the other side are those who closely relate the task of description with those of interpretation and evaluation. For some, philosophy necessarily involves the study of its past, so it must be done historically (Cohen 1986); for others, studying the philosophical past requires doing philosophy (Kenny 1995, 1996); and for others still, the relation goes both ways (Taylor 1984). The reason, as given by philosophers with historicist leanings, is that philosophy is a rearticulation of a view about ourselves and the world, and this requires both the understanding of past articulations and a liberation from them. The study of the philosophical past, then, necessarily involves philosophical judgments, and philosophy must study its past to move beyond it; the history of philosophy must be done philosophically and philosophy must be done historically. Indeed, philosophy is a historical enterprise insofar as the thought or statement of a philosophical view is a historical event and thus part of the history of philosophy. So even contemporaneous philosophical discussions necessarily involve historical references and the understanding of the past, even if the history in question is recent (Popkin 1985).

These positions have been criticized in various ways. Some critics point out that they rely on an oversimplification of the issue (Janaway 1988; Alexander 1988), whereas others object that they fail to draw a distinction between objective and methodological necessity (Gracia 1992). Objective necessity holds between a discipline or study on the one hand and its object of study on the other. In this sense, the history of philosophy, considered as past philosophical views, is indeed necessary not just for the study of the history of philosophy but also for philosophy insofar as philosophy studies the world and all human experience of it and the history of philosophy is part of that object. Methodological necessity, however, holds between two studies or disciplines, insofar as there is a necessary dependence of the methods employed by them. This distinction opens the doors to an alternative position to the two mentioned. According to it, the study of the history of philosophy is not methodologically necessary for philosophy, although philosophy is methodologically necessary for the study of its history. One can philosophize without a historical aim or concern; but one cannot investigate the history of philosophy without a

philosophical understanding of the concepts and arguments it contains. The relation of necessity between philosophy and its history, then, is not reciprocal.

3. VALUE

Regardless of the position one takes with respect to the relation between philosophy and the study of its history, one may still ask whether the second is useful or detrimental for the first. Those who argue that the study of the history of philosophy is incompatible with philosophy see only negative influences on it: the study of the history of philosophy stultifies creativity, prevents discoveries, is irrelevant to present concerns, and wastes precious time (Descartes 1970). And for those who hold that doing the history of philosophy is necessary for doing philosophy, the question of the value of the first for the second is obviously irrelevant. However, for those who maintain that the study of the history of philosophy is neither incompatible with nor necessary for doing philosophy, it is pertinent. Some of these believe that the study of the history of philosophy is harmful, whereas others argue that it is beneficial and thus justify it in various ways. At least eight different justifications are common. They can roughly be divided into three groups: rhetorical, pragmatic, and essentialistic.

Rhetorical justifications in turn fall into two groups. According to one, the history of philosophy provides a source of inspiration: past philosophers function as role models whose lives, devoted to the pursuit of truth, inspire us to emulate them (Rée 1978). According to another, the history of philosophy can be a source of support and respectability, and in that way be used to validate the present (Gilson 1955).

Pragmatic justifications can be classified into four types. One argues that the consideration and analysis of a rich historical treasure of philosophical views and arguments can supply present-day philosophers with a fertile ground in which to train for the philosophical task (Yolton 1986). Another proposes that the history of philosophy is a source of solutions to important philosophical problems insofar as many great minds from the past have presented answers to questions still pertinent today and offer us alternatives to contemporary proposals (Curley 1986). A third maintains that the present state of philosophy is one of confusion and "ill health," and the study of the past can help us figure out how and where philosophy went wrong; the study of the history of philosophy can be therapeutic for the present (Bennett 1988). The fourth group combines all three of these justifications, arguing that philosophy can profit from both the failures and successes of the past (Mash 1987).

Essentialistic justifications are cashed out in terms of the nature of philosophy and the way it develops. At least four versions of them have been proposed. One, not explicit among historiographers of philosophy, although applicable to philosophy and used in some sciences, argues that the ontogeny of a discipline recapitulates its phylogeny. The acquisition of philosophical knowledge by an individual person goes through stages that mirror those that the human race as a whole has experienced in its philosophical understanding. The study of the history of philosophy, then, provides a shortcut to the level of understanding that individual philosophers seek. Another argues that the dialectical nature of philosophy requires that we study its past. Regardless of whether this dialectical nature is taken to apply to the dialogue thought to be fundamental to the philosophical enterprise (Veatch 1988) or to a set of stages of development that repeat themselves (Hegel 1974), it appears essential that philosophy engage its past. In the first case, this is because the variety of the past makes it an ideal interlocutor; and, in the second, it is because any stage in the development of philosophy relies on prior stages. A third justification argues that the understanding and management of science and technology is possible only on the basis of historical experience and the history of philosophy supplies it (Krüger 1984). A fourth argues that philosophy is a cultural enterprise that relies on historical elements such as language, values, presuppositions, and so on; to understand the philosophical present, then, we need to go back to the past, for it is from the past that the present has arisen (Gracia 2000).

4. TEXTS

The object studied by historians of philosophy consists of the views of past philosophers, but they have no way of establishing direct contact with those views except through texts. Their access to Kant's philosophy, for example, is only through the texts that express Kant's views, whether they were composed by the author himself or by subsequent historians. The study of the history of philosophy amounts, then, to the study of texts, and this poses a set of questions that fall within what is frequently called hermeneutics. They may be divided roughly into four categories, depending on whether they have to do with texts themselves, their interpretation, their authors, or their audiences.

With respect to texts, the most pertinent questions concern their nature and identity. For purposes of the

history of philosophy, the texts that matter most are written. Oral texts are relevant only insofar as they have survived either in written reports or have been taped. Historians who wish to give an account of William of Ockham's logic, then, begin by looking at copies of the pertinent texts from Ockham, say the *Logica*. But it turns out that the copies of the text they have are not the autograph Ockham wrote. Rather, they are reconstructions produced by editorial processes that took into account various manuscript versions of Ockham's text, and relied on the judgment of various editors as to the most historically accurate reading. This means that historians need to be aware of the distinction between the historical text—the one produced by Ockham—and the text they currently have, which may be called the contemporary text.

Even when historians have access to a philosopher's autograph, however, they may still ask themselves whether the script they have in front of them is the one intended by the author, for the philosopher may have written something he did not intend, or failed to write something he intended. So in addition to the historical text and the contemporary text, historians could take into account what they consider to be the intended text. But there is still more, for some historiographers argue that there is another text that is pertinent, namely the text the author should have written. Ockham may have written something that did not fit his view, because he was distracted or even failed to understand all the implications of his own position. Hence, in addition to the historical, contemporary, and intended, there is also what might be called the ideal text. These different ways of conceiving texts give rise to wide disagreement among historiographers concerning the kind of text that is most pertinent for the study of the history of philosophy.

Two questions in particular are pertinent concerning the interpretation of texts: "What is an interpretation?" and "What is its purpose?" According to a common conception, an interpretation of a text is the understanding that an interpreter has of the text; according to another, it is a text added to the text under interpretation. A example of the first sort is Thomas Aquinas's understanding of Aristotle's *Metaphysics*; an example of the second is Aquinas's *Commentary on Aristotle's "Metaphysics."* The purpose of the interpretation may vary in each case, and this has also been a subject of disagreement, which most frequently occurs along two lines: understanding the meaning of the text or relating the text to something else. The first, in turn, can be broken down depending on various ways of conceiving the meaning of a text: in terms of the author's understanding or intention, in terms of the understanding of a particular audience, or independently of either the author or any audience. The second purpose of interpretations has been prompted in part by questions raised about the nature and viability of meaning by such Analytic philosophers as W. V. Quine and such Continental philosophers as Jacques Derrida. These questions have undermined meaning-based conceptions of interpretation and have led some historiographers to favor relational ones instead (Daniel 1993). If the purpose of an interpretation is relational, then the interpretation depends on what the text is related to, such as another text, particular historical events, certain conceptual frameworks used in the interpretation, and so on.

Those who make interpretations dependent on authors need to establish the identity of the latter, but this again is contested, for at least three authors need to be considered: historical, pseudohistorical, and contemporary (the terms used to refer to them vary). The first is the person who produced the historical text—the person who wrote Ockham's *Logica* for instance. The pseudohistorical author is the person whom later historians think wrote the text. The pseudohistorical author of the *Logica*, for example, goes by the same proper name as the historical author, but it could in fact be different. The contemporary author is the author of the contemporary text. Recall that the contemporary text is a reconstruction of the autograph carried out by editors on the basis of various texts and readings, so it is likely that it is different from the historical text and, therefore, it would be incorrect to regard the historical author as having sole responsibility for it.

The audience also has frequently been thought pertinent for the interpretation of texts. And here, again, various understandings of it may be considered. For some, authors themselves may be conceived as audiences, whereas others refer to the audiences intended by authors, the audiences contemporary with the composition of the historical text under interpretation, or the audiences contemporary with interpreters. Naturally, these differences in audiences alter the character of an audience-based interpretation. It is one thing to grasp Aristotle's own understanding of his *Metaphysics* and another to comprehend what thirteenth-century scholastics thought of it.

5. APPROACHES

Different views concerning the interpretation of texts generate different approaches to them. Several taxonomies of these have been proposed, but most of them

include some of the following: ideological, sociological, biographical, scholarly, doxographic, apologetic, literary, idealistic, eschatological, dilettantish, and problematic. Ideological approaches use the history of philosophy for the justification of a chosen point of view and treat texts accordingly (Marxist historians). Sociological approaches break down into several, depending on whether they emphasize cultural (Gilson 1955), psychological (Kusch 2000), or generally contextual factors (Peckhaus 2000) in the study of the history of philosophy. Biographical approaches focus on personal histories (Rée 1978). Scholarly approaches seek to establish reliable texts, to produce accurate translations, to determine precise chronologies, and to reconstruct and expound the views of past authors and their relations without engaging in value judgments (Owens 1951). Doxography usually considers facts, figures, and ideas with a primarily informative aim (Diogenes 1925). Apologists see their goal as the defense of a particular author's point of view (John of St. Thomas 1931). The literary approach emphasizes form over content, stressing the need to take the former into consideration for the understanding of the latter (Danto 1985). Idealists consider the views they find in texts as imperfect renditions of what they think are perfect views, so they engage in speculative reconstruction (Russell 1937). Eschatologists view the history of philosophy as progressing toward some end, or as retreating from it (Aristotle 1984). Dilettantes focus on texts in isolation from historical contexts, being interested only in what they can find in them for their own philosophical purposes (Plantinga 1978). And those who adopt a problems approach look at the history of philosophy as a series of attempts to solve philosophical problems (Bennett). A recently proposed variant of the last is the framework approach. According to it, a proper historiographical method should make explicit the conceptual frameworks of problems and views used to study philosophical texts from the past in that such frameworks can be used to understand historical views better both in themselves and in relation to the views of the interpreters and their contexts (Gracia 1992).

A topic of occasional discussion in this context is the nature and value of what is frequently called the Principle of Charity. According to it, historians must attempt to develop the most favorable interpretations of the philosophical views they study. This applies whether the historians agree or disagree with them. If they agree, it is argued that this serves to support their own views, and if they disagree, that then they are presented with the best case against their own positions, forcing them to rethink those views or develop better arguments in their support.

6. GENRES

The genres used in the history of philosophy break down into at least two large categories: textual commentaries and systematic expositions. The first includes more or less literal commentaries. The second breaks down into general or particular histories. General histories of philosophy aim to provide accounts of the whole history of philosophy. Particular histories are concerned with the philosophy of particular periods, regions, nations, ethne, races, and authors, or with specific problems or ideas, and with their comparison. Here are some examples of particular histories: history of medieval philosophy, history of Latin American philosophy, history of French philosophy, history of Hispanic philosophy, history of Black philosophy, Hegel's philosophy, Gustav Bergmann's position on individuation, history of the problem of universals (in general or in a particular period), and the history of the idea of substance. Historiographers disagree on the comparative value of these genres, but they continue to use them.

See also Aristotle; Bergmann, Gustav; Continental Philosophy; Derrida, Jacques; Descartes, René; Feminism and the History of Philosophy; History and Value Judgments; John of St. Thomas; Kant, Immanuel; Ockhamism; Quine, Willard Van Orman; Thomas Aquinas, St.; William of Ockham; Women in the History of Philosophy.

Bibliography

The bibliography lists only sources from the last twenty-five years, except for some classical and medieval sources and a few earlier well-known historical accounts. However, these sources contain references to earlier historiographical literature.

Alexander, Peter. "History of Philosophy: The Analytical Ideal II." *Aristotelian Society* 62 (1988): S191–S208.

Aristotle. *Metaphysics*. Bk. A. In *The Complete Works of Aristotle: The Revised Oxford Translation*, edited by Jonathan Barnes. Vol. 2. Princeton, NJ: Princeton University Press, 1984.

Ayers, Michael. "Analytical Philosophy and the History of Philosophy." In *Philosophy and Its Past*, edited by Jonathan Reé, Michael Ayers, and Adam Westoby. Hassocks, U.K.: Harvester Press, 1978.

Bennett, Jonathan. "Response to Garber and Rée." In *Doing Philosophy Historically*, edited by Peter H. Hare. Buffalo, NY: Prometheus Books, 1988.

Cohen, Lesley. "Doing Philosophy Is Doing Its History." *Synthese* 67 (1986): 51–55.

Collingwood, R. G. *The Idea of History*. Oxford, U.K.: Clarendon Press, 1946.

Curley, Edwin. "Dialogues with the Dead." *Synthese* 67 (1986): 33–49.

Daniel, Steven. "Paramodern Strategies of Philosophical Historiography." *Epoché* 1 (1993): 42–61.

Danto, Arthur. *Narration and Knowledge: Including the Integral Text of Analytical Philosophy of History*. New York: Columbia University Press, 1985.

Descartes, René. *Descartes: Philosophical Letters*, edited and translated by Anthony Kenny. Oxford: Clarendon Press, 1970.

Diogenes Laertius. *The Lives and Opinions of Eminent Philosophers*. Translated by R. D. Hicks. 2 vols. The Loeb Classical Library. London: W. Heinemann, 1925.

Frede, Michael. "The History of Philosophy as a Discipline." *Journal of Philosophy* 85 (11) (1988): 666–672.

Garber, Daniel. "Does History Have a Future? Some Reflections on Bennett and Doing Philosophy Historically." In *Doing Philosophy Historically*, edited by Peter H. Hare. Buffalo, NY: Prometheus Books, 1988.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages*. New York: Random House, 1955.

Gracia, Jorge J. E. *Philosophy and Its History: Issues in Philosophical Historiography*. Albany: State University of New York Press, 1992.

Gracia, Jorge J. E. "Sociological Accounts and the History of Philosophy." In *The Sociology of Philosophical Knowledge*, edited by Martin Kusch. Dordrecht, Netherlands: Kluwer, 2000.

Graham, Gordon. "Can There Be a History of Philosophy?" *History and Theory* 21 (1982): 37–52.

Hare, Peter H., ed. *Doing Philosophy Historically*. Buffalo, NY: Prometheus Books, 1988.

Hegel, G. W. F. *Lectures on the History of Philosophy*. Translated by E. S. Haldane and Frances H. Simson. 3 vols. London: Routledge, 1974.

Irwin, William, ed. *The Death and Resurrection of the Author?* Westport, CT: Greenwood Press, 2002. Articles by Roland Barthes, Michel Foucault, Alexander Nehamas, Peter Lamarque, and others.

Janaway, Christopher. "History of Philosophy: The Analytical Ideal I." *Aristotelian Society* 62 (1988): S169–S189.

John of St. Thomas. *Cursus theologicus. Tractatus de approbatione et auctoritate doctrinae d. Thomae*. Disp. 2, vol. 1. Paris: Desclée, 1931.

Kenny, Anthony. "History of Philosophy: Historical and Rational Reconstruction." In *Methods of Philosophy and the History of Philosophy: Proceedings of the Entretiens of Institute International de Philosophie. Helsinki, August 27–30, 1995. Acta Philosophica Fennica* 61, edited by Simo Knuuttila and Ilkka Niiniluoto. Helsinki: Philosophical Society of Finland, 1996.

Kristeller, Paul Oskar. "Philosophy and Its Historiography." *Journal of Philosophy* 82 (11) (1985): 618–625.

Krüger, Lorenz. "Why Do We Study the History of Philosophy?" In *Philosophy in History: Essays on the Historiography of Philosophy*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge, U.K.: Cambridge University Press, 1984.

Kusch, Martin. *Psychologism: A Case Study in the Sociology of Philosophical Knowledge*. London: Routledge, 1995.

Kusch, Martin. "The Sociology of Philosophical Knowledge: A Case Study and a Defense." In *The Sociology of Philosophical Knowledge*, edited by Martin Kusch. Dordrecht, Netherlands: Kluwer, 2000.

Lafrance, Yvon. *Méthode et exégèse en histoire de la philosophie*. Montreal, Quebec: Les Éditions Bellarmin, 1983.

MacIntyre, Alasdair. "The Relationship of Philosophy to Its Past." In *Philosophy in History: Essays on the Historiography of Philosophy*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge, U.K.: Cambridge University Press, 1984.

Mandelbaum, Maurice H. "The History of Philosophy: Some Methodological Issues." In *Philosophy, History, and the Sciences: Selected Critical Essays*. Baltimore: Johns Hopkins University Press, 1984.

Mash, Roy. "How Important for Philosophers is the History of Philosophy?" *History and Theory* 26 (3) (1987): 287–299.

Morgan, Michael. "The Goals and Methods of the History of Philosophy." *Review of Metaphysics* 40 (1987): 717–732.

Owens, Joseph. *The Doctrine of Being in the Aristotelian Metaphysics: A Study in the Greek Background of Mediaeval Thought*. Toronto: Pontifical Institute of Mediaeval Studies, 1951.

Piaia, Gregorio. "Brucker versus Rorty? On the 'Models' of the Historiography of Philosophy." *British Journal for the History of Philosophy* 9 (1) (2001): 69–81.

Peckhaus, Volker. "The Contextualism of Philosophy." In *The Sociology of Philosophical Knowledge*, edited by Martin Kusch. Dordrecht, Netherlands: Kluwer, 2000.

Peperzak, Adriaan T. "On the Unity of Systematic Philosophy and History of Philosophy." In *History and Anti-History in Philosophy*, edited by T. Z. Lavine and V. Tejera. Dordrecht, Netherlands: Kluwer, 1989.

Plantinga, Alvin. "The Boethian Compromise." *American Philosophical Quarterly* 15 (1978): 129–138.

Popkin, Richard H. "Philosophy and the History of Philosophy." *Journal of Philosophy* 82 (11) (1985): 625–632.

Powers, Lawrence H. "On Philosophy and Its History." *Philosophical Studies* 50 (1) (1986): 1–38.

Rée, Jonathan. "Philosophy and the History of Philosophy." In *Philosophy and Its Past*, edited by Jonathan Reé, Michael Ayers, and Adam Westoby. Hassocks, U.K.: Harvester Press, 1978.

Rorty, Richard. "The Historiography of Philosophy: Four Genres." In *Philosophy in History: Essays on the Historiography of Philosophy*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge, U.K.: Cambridge University Press, 1984.

Russell, Bertrand. "Preface to the First Edition." In *The Philosophy of Leibniz*. 2nd ed. London: George Allen and Unwin, 1937.

Taylor, Charles. "Philosophy and Its History." In *Philosophy in History: Essays on the Historiography of Philosophy*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner. Cambridge, U.K.: Cambridge University Press, 1984.

Veatch, Henry B. "Response to Commentators." In *Doing Philosophy Historically*, edited by Peter H. Hare. Buffalo, NY: Prometheus Books, 1988.

Yolton, John W. "Is There a History of Philosophy? Some Difficulties and Suggestions." *Synthese* 67 (1) (1986): 3–21.

Jorge J. E. Gracia (2005)

HOBART, R. E.

See *Miller, Dickinson S.*

HOBBES, THOMAS

(1588–1679)

Thomas Hobbes, often called the father of modern analytic philosophy, was born in Malmesbury, Wiltshire, England. Hobbes later enjoyed jesting about the significance of his manner of entry into the world. (He was born prematurely when his mother heard of the approach of the Spanish Armada.) "Fear and I were born twins," he would say, adding color to his conviction that the fear of death and the need for security are the psychological foundations both of worldly prudence and of civilization itself. He died at the age of ninety-one in Hardwick, Derbyshire, after a life of travel, study, polemical controversy, and philosophical and literary activity that in his later years had virtually established him as an English institution.

EARLY YEARS

Hobbes's father, Thomas Hobbes, was vicar of Westport, an adjunct of Malmesbury, but his conduct reflected little credit on his cloth. After being involved in a brawl outside his own church, he had to flee to London, leaving Thomas to be brought up by a wealthy uncle, who took the matter of his education very seriously. When he was only fourteen, Hobbes was sent to Magdalen Hall, Oxford, where he remained for five years before taking his bachelor's degree. He seems to have been bored by his Aristotelian tutors, although he acquired considerable proficiency in logic. The strong Puritan tradition of his college impressed Hobbes, but the drunkenness, gaming, and other vices that were prevalent equally impressed him. On leaving Oxford in 1608, Hobbes had the good fortune to become tutor to the young son of William Cavendish, earl of Devonshire. This circumstance introduced him to influential people, to a first-class library, and to foreign travel.

In 1610, on the first of Hobbes's visits to the Continent, he discovered the disrepute into which the Aristotelian system of thought was beginning to fall. Johannes Kepler had recently published his *Astronomia Nova,* and Galileo Galilei had just discovered the satellites of Jupiter through his telescope. Hobbes returned to England determined to devote himself to the pursuit of learning, a resolve that was probably strengthened by his meetings with Francis Bacon. Hobbes, however, thought little of Bacon's so-called method of induction, with its stress on observation and experiment, which was later to become the inspiration of the Royal Society. Nevertheless, he agreed with Bacon in his contempt for Aristotelianism, in his conviction that knowledge means power to be used for the improvement of man's estate, and in his advocacy of clear and concrete speech instead of the vague abstractions of the schools.

At this period of his life Hobbes had turned to the classics to gain an understanding of life and of philosophy, which, he thought, could not be found in the schools. After a period of reading and reflection, he decided to translate Thucydides into English, a significant choice. Like Thucydides, Hobbes believed that history was written for instruction, and he wished to instruct his countrymen on the dangers of democracy. In 1628, when Hobbes published his translation, Charles I had been on the throne for three years and was already at loggerheads with Sir John Eliot and John Pym. Hobbes's translation was the first of his many attempts to bring his countrymen to their senses and to make them aware of the tragedy that they courted: that of civil war, from which proceed "slaughter, solitude, and the want of all things."

PHILOSOPHICAL AWAKENING

It was not until the time of his second journey to the Continent that Hobbes's career as a philosopher began. His patron had died, and as a temporary economy, Catherine, the countess of Devonshire, had dispensed with Hobbes's services. Hobbes took similar employment with Sir Gervase Clinton and, in 1629, accompanied Clinton's son on a journey to the Continent. There Hobbes developed a passionate interest in geometry, which impressed him as a method for reaching indubitable conclusions. Could not his convictions about the dangers of democracy be demonstrated? Could not his opinions about man, gleaned from his observation of the contemporary scene, from his insight into his own nature, and from his perusal of the pages of Thucydides and Niccolò Machiavelli, be postulated as axioms from which theorems about the conditions of a commonwealth might be generated?

Hobbes's discovery of geometry gave him a method of analysis and a conception of scientific method, but he still lacked a conceptual scheme to give content to his demonstrations about man and society. In Paris, during his third journey to the Continent (1634–1637), again in the service of the Devonshires as tutor to William, the succeeding earl, he became a member of the intellectual

circle of the Abbé Marin Mersenne, who patronized René Descartes and Pierre Gassendi. (Gassendi later became one of Hobbes's firmest friends.) Hobbes also made a pilgrimage to Italy in 1636 to visit Galileo, the leading exponent of the new natural philosophy. By the time of his return to England in 1637, he had conceived, perhaps at Galileo's suggestion, the main outlines of his philosophical system, in which the method of geometry and the concepts of the new science of motion were to be applied to man in society.

It is a mistake to think of Hobbes's interests as purely political. Hobbes claimed originality for his optics as well as for his civil philosophy, and at some point between his discovery of geometry and his return from his third journey to the Continent, he wrote his first philosophical work, the *Little Treatise,* in geometrical form, in which he sketched an explanation of sensation in terms of the new science of motion. His interest in sensation, according to his prose autobiography, arose from an encounter with some learned men who were discussing the cause of sensation. One of them asked derisively what sensation was, and Hobbes was astonished to find that none of them could say. From then on, he was haunted by the problem of the nature and cause of sense. He began to think he was near an explanation after it struck him that if bodies were always at rest or always moved at a constant rate, the ability to make discriminations would vanish, and with it all sensation. He concluded that the cause of everything, including that of sensation itself, must be in variations of motion.

In his verse autobiography, Hobbes graphically related how, on his third journey, he was obsessed by the omnipresence of motion. He was acclimating himself to Galileo's audacious suggestion that motion is the natural state of bodies and that they continue in motion to infinity unless they are impeded. This went against the crude evidence of the senses as well as against the established Aristotelian worldview, in which rest was regarded as the natural state. But if Galileo's supposition could be entertained, Hobbes thought, even apparition itself could be explained as a meeting place of motions, and from Galileo's law of inertia the phenomena of sense and imagination could be deduced.

The state of turmoil in England on his return drove Hobbes to make his first systematic attempt to employ his geometrical approach and mechanistic psychology to present the realities beneath the appearances of the contemporary issues. His *Elements of Law,* circulated in 1640 in manuscript form during the session of Parliament, was the result. This work, which demonstrated the need for undivided sovereignty, was published in 1650 in two parts, *Human Nature* and *De Corpore Politico.* However, its arguments were taken from general principles of psychology and ethics, rather than from appeals to divine right. Many regard Hobbes's *Human Nature* as one of his best works. It consists largely of traditional psychology coordinated and underpinned by the conceptual scheme he had learned from Galileo.

EXILE IN FRANCE

Hobbes claimed later that his life would have been in danger because of the views expressed in *Elements of Law,* had not the king dissolved Parliament in May 1640. Six months later, when the Long Parliament impeached Thomas Wentworth, earl of Strafford, Hobbes fled to the Continent in fear for his life, later priding himself on being "the first of all that fled." A warm welcome awaited him in Mersenne's circle, and he settled down in Paris to his most productive philosophical period.

His first work was the composition of some sixteen objections to Descartes's *Meditations,* which Mersenne submitted to Descartes in advance of its publication. This led to a rather acrimonious exchange between Descartes and Hobbes. In 1642 Hobbes published his *De Cive,* an expanded version in Latin of Part 2 of his *Elements of Law* (later to appear as *De Corpore Politico*). The additional sections dealt largely with a more detailed treatment of the relationship between the church and the civil power. During the period from 1642 to 1646, Hobbes published his *Minute or First Draught of the Optiques,* which he considered one of his most important and original works. He also started work on his most ambitious scheme—the construction of a trilogy on body, man, and citizen, in which everything in the world of nature and man was to be included in a conceptual scheme provided by the new science of mechanics. Hobbes made a beginning with *De Corpore,* which was to be the first work in the trilogy.

In 1646, however, political events again interfered with Hobbes's more abstract speculations. He was on the verge of accepting an invitation to retire in peace to a friend's house in Languedoc, in the south of France, when he was requested to act as tutor in mathematics to the future Charles II, who had just fled to Paris. Hobbes's tutorship, however, was interrupted, if not terminated, by a severe illness in 1647. He recovered after having consented to receive the sacrament on what he took to be his deathbed, and he was drawn again into political controversy by the presence of so many Royalist *émigrés.* A second edition of *De Cive* was published in 1647, but this was in Latin and had only a limited circulation. Hobbes

therefore decided to blazon abroad his views on man and citizen for all to read, in English, with the arresting title of *Leviathan.* With Mersenne's unfortunate death in 1648, Hobbes began to feel increasingly isolated, for he was suspected of atheism and was an outspoken enemy of the Catholic Church.

Political events in England provided a fitting prelude to the publication of *Leviathan.* Charles I was executed in 1649 and, until 1653, when Oliver Cromwell was made Protector, there was constant discussion and experimentation to find an appropriate form of government. *Leviathan,* published in 1651, was therefore very topical. It came out strongly in favor of absolute and undivided sovereignty, without the usual arguments from divine right. Indeed, Hobbes conceded popular representation but, by an ingenious twisting of the social contract theory, showed that it logically implied the acceptance of undivided sovereignty.

RETURN TO ENGLAND

Hobbes returned to England in 1651 after a severe illness and soon became embroiled in a heated debate with John Bramhall, bishop of Derry, Ulster, on the subject of free will. In 1645, in Paris, Hobbes had discussed the problem of free will with the bishop, and they both wrote their views on the matter soon afterward. A young disciple of Hobbes published his contribution in 1654, without Hobbes's consent, under the title *Of Liberty and Necessity.* Bramhall was understandably indignant and, in 1655, he published the whole controversy under the title *A Defence of True Liberty from Antecedent and Extrinsical Necessity.* In 1656 Hobbes replied by printing Bramhall's book, together with his own observations on it, which he called *The Questions concerning Liberty, Necessity, and Chance.* Bramhall replied in 1658 with *Castigations of Hobbes his Last Animadversions,* which carried an appendix called "The Catching of Leviathan the Great Whale." Bramhall died in 1663, and Hobbes had the last word a few years later.

There was another controversy in which Hobbes was caught up for the major part of the twenty years that were left to him. This one involved John Wallis, professor of geometry at Oxford, who mercilessly exposed Hobbes's attempt in *De Corpore* (1665) to square the circle—not then such a ridiculous enterprise as it now seems—and Seth Ward, professor of astronomy, who launched a polemic against Hobbes's general philosophy. These two men were members of the "invisible college" that the king had recognized as the Royal Society in 1663. They were Puritans in religion and Baconians in their approach to science. Hobbes had annoyed them not simply by his attack on their religion and his contempt for the method of induction, but also by his diatribes on the universities as hotbeds of vice and sedition. Hobbes replied to their published criticisms with an emended English version of *De Corpore* with "Six Lessons" appended for Wallis. This was in turn attacked by Wallis, and the controversy dragged on for many years, often descending into personal vituperation on both sides.

Not all of Hobbes's remaining years, however, were spent on this abortive controversy. *De Homine,* the second part of his trilogy, was published in 1657. This dealt with optics and human nature, matters on which Hobbes's opinions were already well known; accordingly, it attracted little attention and was not translated.

After the Restoration, Hobbes was granted a pension and "free accesse to his Majesty, who was always much delighted in his witt and smart repartees" (John Aubrey, *Brief Lives,* pp. 152–153). Only once again did he fear for his life. After the Great Plague (1665) and the Great Fire of London (1666), some reason was sought for God's displeasure, and a spasm of witch-hunting shook Parliament. A bill was passed by Parliament for the suppression of atheism, and a committee was set up to investigate *Leviathan.* The matter was eventually dropped, probably through the king's intervention, but Hobbes was forbidden to publish his opinions thereafter.

In 1668 Hobbes finished his *Behemoth*—a history of the period from 1640 to 1660, interpreted in the light of his beliefs about man and society. He submitted it to King Charles, who advised against its publication (it was published posthumously in 1682).

Even at this advanced age Hobbes was still capable of exerting himself both physically (he played tennis until he was seventy-five) and philosophically. John Aubrey, later his biographer, sent him Bacon's *Elements of Common Law* for his comments; and Hobbes, after protesting his age, managed to produce his unfinished *Dialogue between a Philosopher and a Student of the Common Laws of England* (published posthumously in 1681). This minor work was interesting in that Hobbes anticipated in it the analytical school of jurisprudence of the nineteenth century and came out unequivocally in favor of what has been called the command theory of law. At the age of eighty-four Hobbes wrote his autobiography in Latin verse after completing one in prose. At eighty-six, for want of something better to do, he published a verse translation of the *Iliad* and the *Odyssey.*

LOGIC AND METHODOLOGY

Hobbes lived during the emergence of men who challenged not only traditional tenets about political and religious authority but also the wisdom of the past, especially that of Aristotle. Men were exhorted to find out things for themselves, to consult their own consciences, and to communicate with God directly, instead of through the established religious hierarchy. It was widely believed that all men have the gift of reason but that they make poor use of it through lack of a proper method. Books such as Bacon's *Novum Organum,* Descartes's *Regulae* and *Discourse on Method,* and Benedict de Spinoza's *Ethics* were written to remedy this defect. Thus, Hobbes was not exceptional in believing that knowledge, which meant power, could be obtained only by adopting a certain kind of method.

According to Hobbes, the knowledge whereby most men live is the knowledge gleaned from experience, culminating in prudence and history—"the register of knowledge of fact." Hobbes described experience as "nothing but remembrance of what antecedents have been followed by what consequents." Bacon had tried to set out this sort of knowledge explicitly in his *Novum Organum,* and it was taken by the Royal Society to be the paradigm of science.

DOCTRINE OF NAMES. Hobbes, however, was very contemptuous of such grubbing around and peering at nature, not only in natural philosophy but also in civil philosophy. Had Galileo or William Harvey, the pioneers of the new philosophy, made a laborious summary of their experience? And in civil philosophy, what store is to be placed on the dreary saws of practical politicians or the ossified ignorance and superstitions of the common lawyers? Mere prudence, which is the product of experience, should not be mistaken for wisdom. Wisdom is the product of reason, which alone gives knowledge of "general, eternal, and immutable truths," as in geometry.

In geometry, definitions are of paramount importance. Therefore, claimed Hobbes: "The only way to know is by definition." Thus, science is "knowledge of all the consequences of names appertaining to the subject in hand." It gives knowledge not of the nature of things but of the names of things. We start with certain terms or names about whose definition we agree. We connect these into such statements as "A man is a rational, animated body," just as we add items in an account. We then find that if we follow certain methods of combining the statements so created, conclusions can be drawn that are contained in the premises but of which we were ignorant before we started reckoning. "For REASON, in this sense, is nothing but reckoning, that is adding and subtracting, of the consequences of general names agreed upon for the marking and signifying of our thoughts."

Obvious objections to such an account of scientific knowledge immediately come to mind. How, for instance, can we be sure that such a train of reasoning applies to anything? How are the meanings of Hobbes's names fixed, and how are the rules for their combinations determined?

Hobbes supposed that "names are signs not of things, but of our cogitations." Words are not the only things that can be signs; for instance, a heavy cloud can be a sign of rain. This means that from the cloud we can infer rain. This is an example of a natural sign; other examples are animal warnings of danger and summonses to food. These natural signs are to be distinguished from language proper, which consists of sounds, marks, and other such significations determined—as are the ruler of civil society—by decision. Animal noises come about by necessity, not by decision, as human speech does. That is why, on Hobbes's view, animals, though capable of imagery, cannot reason; for reasoning presupposes words with meanings fixed by decision.

Hobbes thought that every man has his own private world of phantasms or conceptions, for which words are signs that function for him like a private system of mnemonics. These words act as signs to others of what a man thinks and feels. Although some words signify conceptions, they are not names of conceptions; for Hobbes seemed to use the word *name* for the relation of reference between names and things, and words such as *signify* for the relationship between particular occurrences of a name and the idea in a person's mind. Some names are names of things themselves, such as "a man," "a tree," or "a stone," whereas others, such as "future," do not stand for or name things that as yet have any being. Such words signify the knitting together of things past and things present. In a similar way there are names, such as "impossible" and "nothing," that are not names of anything. Such names are signs of our conceptions, but they name or stand for "things" that do not exist.

Hobbes's doctrine was not altogether clear. He seemed to mean that all names serve as mnemonics to us of our conceptions and as signs to others of what we have in mind, but that only some names actually denote things in a strict sense. This leads to the distinctions that Hobbes introduced in relation to the logical function of names. Names can be either concrete or abstract. Concrete names can denote bodies, their accidents, or their names.

Abstract names come into being only with propositions and denote “the cause of concrete names.”

UNIVERSALS. There are two classes of concrete names: proper names and universal names. A proper name, such as “Peter,” is singular to one thing only; a universal name, such as “man,” denotes each member of a class of things. A universal name, “though but one name, is nevertheless the name of diverse particular things; in respect of which together, it is called a universal; there being nothing in the world universal but names; for the things named are every one of them individual and singular.”

Hobbes’s doctrine of universal names was crucial to his attack on the scholastic belief in essences. The world, Hobbes maintained, contains no such essences for universal names to designate. “Universal” is the name of a class of names, not of a diaphanous type of entity designated by a name. The error of those who believe in essences derives from their tendency to treat a universal name as if it were a peculiar kind of proper name. It is the *use* of a name that makes it universal, not the status of the thing that the name designates.

Hobbes’s doctrine of abstract names was more obscure but of cardinal importance in his account of scientific knowledge. Abstract names come into being when names are joined in propositions. A proposition is “a speech consisting of two names copulated, by which he that speaketh signifieth the latter name to be the name of the same thing whereof the former is the name.” For instance, in saying “man is a living creature,” the speaker conceives “living creature” and “man” to be names of the same thing, the name “man” being comprehended by the name “living creature.” This relation of “comprehension” can be brought out in some languages by the order of words without employing the verb “to be.” The copulation of the two names “makes us think of the cause for which these names were imposed on that thing,” and this search for the causes of names gives rise to such abstract names as “corporeity,” “motion,” “figure,” “quantity,” and “likeness.” But these denote only the causes of concrete names and not the things themselves. For instance, we see something that is extended and fills space, and we call it by the concrete name “body.” The cause of the concrete name is that the thing is extended, “or the extension or corporeity of it.” These causes are the same as the causes of our conceptions, “namely, some power of action, or affection of the thing conceived, which some call the manner by which anything works upon our senses, but by most men they are called accidents.” Accidents are neither the things themselves nor parts of them, but “do nevertheless accompany the things in such manner, that (saving extension) they may all perish, and be destroyed, but can never be abstracted.” Among such accidents some are of particular importance for science, those which Hobbes sometimes referred to as “universal things” or “such accidents as are common to all bodies.” These are the abstract concepts by means of which a theory is developed about the underlying structure of nature. The endeavor of the scientist is to understand, by means of the resoluto-compositive method of Galilean mechanics, the universal cause—motion—without knowledge of which such fundamental theories could not be developed.

MISUSES OF WORDS. Hobbes has often been called the precursor of modern analytical philosophy because he was particularly sensitive to the manner in which ridiculous (and dangerous) doctrines can be generated through confusion about how words have meaning. One class of absurdities is generated by failure to understand the different ways in which the copula “is” can function. Such terms as *essence, reality,* and *quiddity,* beloved by the schools, “could never have been heard among such nations as do not copulate their names by the verb ‘is,’ but by adjective verbs as runneth, readeth.” The word *is* in a proposition such as “Man is a living body” has the function of “comprehension” or class inclusion. Something of the form “If *x* is a man, then *x* is a living body” is being stated. There is no commitment to the existence of men that is implied when *is* occurs in such statements as “Here is Thomas Hobbes.”

Absurdities also arise if names of accidents are assimilated to names of bodies. For instance, those who say that faith is “infused” or “inspired” into a person treat faith as if it were the name of a body, for only bodies can be poured or breathed into anything. An accident is not in a body in the same sort of way that a body can be in a body—“as if, for example, redness were in blood, in the same manner, as blood is in a bloody cloth.” Hobbes was also eloquent on the subject of names that name nothing.

SCIENTIFIC TRUTH. Hobbes’s theory of scientific truth was not altogether consistent. He started with the important insight that “true” and “false” are attributes of speech, not of things. Truth, then, “consisteth in the right ordering of names in our affirmations.” It characterizes propositions in which names of limited generality are “comprehended” by those of wider generality: For example, “Charity is a virtue.” Hobbes held, it therefore seems, that all true propositions are analytically true, which is a plausible enough view if only geometrical truths are at issue. But Hobbes often spoke as if all truth must con-

form to this model. He saw that this raises the question of how the initial definitions are to be fixed, and about these definitions he often seemed to take a conventionist view by suggesting that "truth therefore depends upon the compacts and consents of men." He often linked the contract theory of the origin of civil society with a theory about agreement on definitions. When he was speaking about natural science, however, his position was not so clearly conventionist. The difference was caused by his assumption that men construct states just as they construct circles or triangles. But since they do not construct natural bodies in the same way, the problem therefore arises as to how Hobbes thought that propositions of natural science, which did not come into being through decisions of men, say what is true about the natural world.

Hobbes thought that all the propositions of natural science are deductions from the basic theory of motion, in which there are primary propositions containing such simple unanalyzable concepts as motion, extension, and straightness. These are "well enough defined, when, by speech as short as may be, we raise in the mind of the hearer perfect and clear ideas of the thing named" (*De Corpore*). Such conceptions are featured in Hobbes's account of evidence, which is "the concomitance of a man's *conception* with the *words* that signify such conception in the act of ratiocination" (*Human Nature*). A parrot could speak truth but could not know it, for it would lack the conceptions that accompany the speaking of truth by a man who knows truth. "Evidence is to truth, as the sap to the tree … for this evidence, which is meaning with our words, is the life of truth. Knowledge thereof, which we call *science,* I define to be *evidence of truth,* from some beginning or principle of *sense.*"

Conceptions, in Hobbes's view, are explained causally in terms of motions that arise in the head and persist after the stimulation of sense organs by external bodies. Names, which are joined together in true propositions, are signs of these conceptions in that they mark them for the individual and enable other people to make inferences about what he thinks. Thus, Hobbes must have thought that when a man knows (as distinct from when he merely speaks) what is true, his conceptions, as it were, keep pace with what he is saying. Some of these conceptions, those involved in understanding primary propositions, are clear and distinct ideas of things named. Thus, scientific systems are somehow anchored to the world of nature by means of names that refer to attributes of bodies of which we have a clear and distinct idea.

This theory resembles, in certain respects, the self-evidence theory of the Cartesians. However, it seems inconsistent with the conventionalism of Hobbes's other remarks about basic definitions and is a very confused account in itself, not very helpful in elucidating what makes scientific propositions true. In the empirical sciences the clarity of the ideas in the initial postulates is neither here nor there. What matters is whether statements deduced from them can be observationally confirmed.

SCIENTIFIC INQUIRY. The ambiguity in Hobbes's account of truth is paralleled by the ambiguity in his account of scientific method, which he equated with the search for causes. One of his most famous definitions of philosophy or scientific knowledge (he did not distinguish between the two) occurs at the start of *De Corpore* (Molesworth ed.): "philosophy is such knowledge of effects or appearances, as we acquire by true ratiocination from the knowledge we have first of their causes or generation: And again, of such causes or generations as may be from knowing first their effects." By "cause" Hobbes meant, of course, antecedent motion, and he was unusual in thinking that even geometrical figures are to be explained in terms of motion because of the movements involved in constructing them.

Hobbes's distinction between these two forms of philosophical knowledge is important. In the case of acquiring knowledge of effects from knowledge of causes or generation, his conventionist account of truth holds good. For instance, in the case of deciding that a figure must be a circle from our knowledge of the motions from which it was produced, "the truth of the first principles of our ratiocination, namely definitions, is made and constituted by ourselves, whilst we consent and agree upon the appellation of things." He used this method in *De Corpore* to explain parallel lines, refraction and reflection, circular and other forms of motion, angles, and similar concepts. It also seems that he had this model in mind when he thought about the generation of the artificial machine of the commonwealth.

When dealing with knowledge of causes from effects, however, Hobbes's account is far less clear-cut and conventionist. At the beginning of Part 4 of *De Corpore,* for instance, he said: "The principles, therefore, upon which the following discourse depends, are not such as we ourselves make and pronounce in general terms, as definitions: but such, as being placed in the things themselves by the Author of Nature, are by us observed in them." The explanations that we give in the natural sciences may be true, but it is impossible to demonstrate that they are necessarily true, for the phenomena are not generated by

human contrivance, as are the phenomena of geometry and politics.

The method on which Hobbes was relying in both these types of scientific inquiry was, of course, the resoluto-compositive method of Galilean mechanics. In this method a typical phenomenon, such as the rolling of a stone down a slope, was taken. Such properties as color and smell, which were regarded as scientifically irrelevant, were disregarded, and the situation was resolved into simple elements that could be quantified—the length and angle of the slope, the weight of the stone, the time the stone takes to fall. The mathematical relations disclosed were then manipulated until functional relations between the variables were established. The situation was then synthesized or "composed" in a rational structure of mathematical relations. This is what Hobbes called analysis—the search for causes, given the effects. "Synthesis" consisted in starting from the known causes and deducing effects from them. In Galileo's hands this method was highly successful because he tested such deductions by observation. In Hobbes's hands the method was not so fruitful because it always remained an imaginary experiment.

Similar ambiguities in Hobbes's methodology complicate our effort to understand his conception of his trilogy on body, man, and citizen. He thought of geometry as the science of simple motions that could demonstrate how figures are generated by varieties of motion. Second came the philosophy of motion, as usually understood in the Galilean system, in which the effects of the palpable motions of one body on another were considered. Third came physics, the investigation of the internal and invisible motions that explain why "things when they are the same, yet seem not to be the same, but changed." Sensible qualities, such as light, color, heat, and sound, were to be explained, together with the nature of sensation itself. After physics came moral philosophy, the study of the motions of the mind—appetites and aversions. Such motions of the mind had their causes in sense and imagination. Finally, there was civil philosophy, the study of how states are generated from the qualities of human nature.

It is probable that Hobbes did not view the hierarchy of sciences as a rigorous deductive system. To start with, he never worked out the deductions in any detail—for instance, in the transition from what he called physics to moral philosophy, or psychology. Furthermore, what he said about the possibility of a self-contained science of politics contradicts his suggestion that it must be deduced from the fundamental theory of motion and that it supports the conventionist account of truth in politics. Hobbes said that even those who are ignorant of the principles of physics and geometry might attain knowledge of the principles of politics by the analytical method. They could start, for instance, with the question of whether an action is just or unjust; "unjust" could be resolved into "fact against law," and "law" into "command of him or them that have coercive power"; "power" could in its turn be derived from the wills of men who established such power so that they might live in peace.

This line of argument, developed in *De Corpore* after admitting the possibility of using the synthetic method to start from the first principles of philosophy and deduce from them the causes and necessity of constituting commonwealths, is confirmed by Hobbes's injunction in the Introduction to *Leviathan* that a man who is to govern a whole nation must "read in himself, not this or that particular man; but mankind: which though it be hard to do, harder than to learn any language of science; yet when I shall have set down my own reading, orderly and perspicuously, the pains left another, will be only to consider, if he find not the same in himself. For this kind of doctrine admitteth no other demonstration." It appears that Hobbes envisaged a relatively self-contained doctrine of politics based on introspection. His trilogy was, therefore, probably not conceived as forming a strictly deductive system. Its various elements were to be more loosely bound together by the fact that all three were sciences of motion.

PHILOSOPHY OF NATURE

Hobbes's natural philosophy seems to have been stimulated largely by the problem of the nature and cause of sensation that had so long haunted him. His theory was that the cause of everything, including sensation itself, lies in the varieties of motion. His first sketches of such a theory were in his *Little Treatise* and his early optical treatises, and his *De Corpore* was an ambitious development of this fundamental idea. Geometry, physics, physiology, and animal psychology were all incorporated within the theory of motion. Sensation occupied a shadowy middle position between the gross motions of the external world and the minute motions of the bodily organs.

The strange thing about Hobbes's preoccupation with sensation is that he seems to have been little troubled by the problems that are almost the stock in trade of philosophers—the problems of epistemology. He assumed that things exist independently of our perceptions of them and was convinced that "conceptions and apparitions are nothing really but motions in some inter-

nal substance of the head." The "nothing but" is very hard to accept, for obviously when we speak of "thoughts" and "conceptions," we do not mean the same as when we talk of motions in the brain.

MOTION AND QUALITIES. On the status of the various sense qualities, Hobbes held, as did such natural philosophers as Kepler and Galileo, that secondary qualities—such as smells, colors, and sounds—are only appearances of bodies, whose real properties are those of extension, figure, and motion. Such secondary qualities are phantasms in the head, caused by the primary properties of external objects interacting with the sense organs, but the secondary qualities represent nothing outside. Hobbes argued that images and colors are "inherent in the sentient" because of illusions and because of images produced in other ways—for example, by blows on the optic nerve. But this proved too much, for representations of primary qualities are equally liable to deceive. Hobbes also proved too little, for he argued that secondary qualities represent no qualities of external objects because tastes, smells, and sounds seem different to different sentients. But there are standard tests for establishing the fact, for example, that a man is colorblind; and, as George Berkeley later showed, the perception of primary qualities is infected with a similar relativity owing to the point of view and peculiarities of the percipient. Hobbes, in fact, gave but a halting philosophical patter to justify a distinction deeply embedded in the thought and practice of the new natural philosophers, for the basic tenet of these thinkers was that bodies in motion exist independently of our perception of them and that mathematical thinking about them discloses their real properties.

Hobbes regarded sensation and apparition as a meeting place of motions. Sense organs, he thought, are agitated by external movements without which there would be no discrimination and, hence, no sensation. Therefore, to give the entire cause of sense, an analysis is required of all movements in external bodies, which are transmitted to the sense through a medium. But sensation is not simply the end product of external motions; it also functions as an efficient cause of actions of sentient beings. Actions, in Hobbes's view, are really reactions to stimuli that are passed on by means of the sense organs. Sensation acts as a bridge between movements in the external world and the behavior of animals and men.

Hobbes's mechanical theory was distinctive in that he extended the Galilean system in two directions: into geometry at one end, and into psychology and politics at the other. He thought that no one could understand the definitions of geometry without grasping how motion is involved in the construction of lines, superficies, and circles. Geometry is the science of simple motions. It paves the way for mechanics, which explains the effects of the motions of one body on another, and for physics, which deals with the generation of sensible qualities from the insensible parts of a body in contact with other moving bodies.

CAUSATION. All causation, in Hobbes's view, consists in motion. "There can be no cause of motion except in a body contiguous and moved." If bodies are not contiguous and yet influence one another, this influence has to be conveyed either by a medium or by emanations of minute bodies that impinge on others (the theory of effluxes). There can be no action at a distance. Hobbes combined this principle with his rendering of Galileo's law of inertia.

Hobbes extended this conception of causation to human actions: "A final cause has no place but in such things as have sense and will; and this also I shall prove hereafter to be an efficient cause." To bring about this transition from mechanics to physiology and psychology, Hobbes introduced the concept of "endeavour," which he defined as "motion made in less space and time than can be given … that is, motion made through the length of a point, and in an instant or point of time." In other words, he used the term to postulate infinitely small motions, and by means of this notion he tried to bridge the gap between mechanics and psychology. He thought that external objects, working on the sense organs, produce not only phantasms but also minute motions that proceed to the heart and make some alteration in the vital motions of the circulation of the blood. When these vital motions are thereby helped, we experience pleasure; when they are hindered, we experience pain. The body will be regulated in such a way that it will preserve the motions that help the vital motions and get rid of or shun those that hinder. This brings about animal motion. Even habits are nothing but motions made more easy by repeated endeavors; they are comparable to the bend of a crossbow.

Hobbes has often been called a materialist, but it is more appropriate to regard him as a great metaphysician of motion. He took concepts that have an obvious application to one realm of phenomena (mechanics) and developed a conceptual scheme that, he thought, could be applied to all phenomena. The plausibility of such a scheme derives from stressing tenuous similarities and ignoring palpable differences. There is a sense in which

social life is a matter of bodies moving toward and away from other bodies, just as there is a sense in which work is moving lumps of matter about. But such descriptions are either unilluminating truisms, or, if they carry the "nothing but" implication, they are misleading. Habits, for example, may be formed in part by a variety of movements, but to suggest that by "habit" we mean nothing but a buildup of movements is ridiculous. This either confuses a question of meaning with a question of genetic explanation or it demonstrates the length to which Hobbes was prepared to go in rigging appearances to suit his metaphysical redescription.

SUBSTANCE AND ACCIDENT. In his *De Corpore* Hobbes defined "body" as "that which having no dependence upon our thought, is coincident and coextended with some part of space." Bodies need not be visible. Indeed, "endeavours," which featured so widely in his system, are movements of minute unobservable bodies. Hobbes held that there is nothing else in the world but bodies, and he therefore did not flinch from the conclusion that "substance incorporeal" is a contradiction in terms. He argued that God cannot be such a substance. To Bishop Bramhall's question of what he took God to be, Hobbes replied, "I answer, I leave him to be a most pure, simple, invisible, spirit corporeal."

By "accident" Hobbes meant a property or characteristic that is not a part of a thing but "the manner by which any body is conceived." Most accidents, with the exception of figure and extension, can be absent without destruction of the body. But Hobbes was not altogether clear about the grounds for such an exception. If the grounds are the inconceivability of a body without figure and extension, why should not color be in the same category as figure? Hobbes regarded color as a subjective appearance brought about by the interaction of sense organs with the primary qualities of external objects; but if the criterion is one of conceivability, as Berkeley pointed out, it is as difficult to conceive of a body without color as it is to conceive of one without figure. Hobbes in fact defined "body" in terms of accidents that are mathematically tractable in mechanics and geometry. He tried to provide some kind of rationale for this basic assumption of the new natural philosophy by introducing the criterion of conceivability, which will not really do the work required of it.

Hobbes defined space as "the phantasm of a thing existing without the mind simply." By this he meant that what is called space is the appearance of externality. If the world were to be destroyed, and a man were left alone with his imagination and memories, some of these would appear external to him, or located in space, for the system of coordinates used to describe the relative position of bodies is a subjective framework. "Place is nothing out of the mind nor magnitude anything within it." A body always keeps the same magnitude, whether in motion or at rest, but it does not keep the same place when it moves. Place cannot, therefore, be an accident of bodies; place is feigned extension—an order of position constructed from experience of real extended things to provide a framework for their externality. Similarly, time is "the phantasm of before and after in motion." Time systems are constructed from the experience of succession.

Hobbes never made clear the relationship between any particular temporal or spatial system that an individual may devise and the system of coordinates adopted by the natural philosophers. Here again, Hobbes typically took for granted the system used by the scientists and tacked on a very brief philosophical story about its relation to the "phantasms" of the individual.

PSYCHOLOGY

Hobbes's psychology was not behavioristic, as it has sometimes been said to be, except insofar as behaviorism has often been associated with a materialistic metaphysical theory or with mechanical modes of explanation. Hobbes stressed the indispensability of introspection in the analysis and explanation of human behavior.

When Hobbes looked into himself he found, of course, motions that were in conformity with Galilean principles. He boldly proclaimed in *De Corpore* that "we have discovered the nature of sense, namely, that it is some internal motion in the sentient." The external body, either directly or via a medium, presses on the organ of sense, "which pressure, by the mediation of the nerves, and other strings and membranes of the body, continues inwards to the brain and heart, causeth there a resistance, or counterpressure, or endeavour of the heart to deliver itself, which endeavour, because outward, seemeth to be some matter without." Sensations are thus nothing but motions. They have the character of externality because of the "outward endeavor" of the heart.

PERCEPTION. Having provided a mechanical starting point for his psychology, Hobbes then tried to describe what was known about psychological phenomena in terms compatible with a mechanical theory. One of the most obvious features of perception is that it involves seeing something as something, some sort of discrimination or recognition. Hobbes's way of saying this was that sense

always has “some memory adhering to it.” This was to be explained by the sense organs’ property of acting as retainers of the movements of external bodies impinging on them. Without this retention of motions, what we call sense would be impossible, for “by sense we commonly understand the judgment we make of objects by their phantasms; namely, by comparing and distinguishing those phantasms; which we could never do, if that motion in the organ, by which the phantasm is made, did not remain there for some time, and make the same phantasm return.”

The selectivity of perception raised a further problem. Why is it that men do not see many things at once? Hobbes again suggested a mechanical explanation: “For seeing the nature of sense consists in motion; as long as the organs are employed about one object, they cannot be so moved by another at the same time, as to make by both their motions one sincere phantasm of each of them at once.” But this does nothing to explain why one object rather than another is selected. Hobbes’s ideomotor theory made it hard to give a plausible account of the influence of interests, attitudes, and sets on what is selected in perception.

Hobbes also attempted a mechanical explanation of the phenomena of attention and concentration. When a strong motion impinges on the sense organ, the motion from the root of the sense organ’s nerves to the heart persists contumaciously and makes the sense organ “stupid” to the registering of other motions.

IMAGINATION AND MEMORY. Hobbes’s account of imagination was explicitly a deduction from the law of inertia. “When a body is once in motion, it moveth, unless something else hinder it, eternally … so also it happeneth in that motion, which is made in the internal parts of a man when he sees, dreams, etc. For after the object is removed, or the eye shut, we still retain an image of the thing seen, though more obscure than when we see it.” Imagination, therefore, is “nothing but decaying sense.” This decay is not a decay in motion, for that would be contrary to the law of inertia. Rather, it comes about because the sense organs are moved by other objects, and subsequent movements obscure previous ones “in such manner as the light of the sun obscureth the light of the stars.”

Memory, Hobbes claimed, differs from imagination only in that the fading image is accompanied by a feeling of familiarity. “For he that perceives that he hath perceived remembers,” and memory “supposeth the time past.” Hobbes thus seems to have more or less equated what is past with what is familiar, which is most implausible even if familiarity is often a hallmark of what is past. It is also difficult to see how, in his view, remembering something could be distinguished from seeing it for a second time, if the second impression of the thing is not very vivid.

Hobbes’s fundamental mistake in all such descriptions and explanations was to attempt to distinguish performances, such as perceiving and remembering, by reference to subjective hallmarks vaguely consistent with his mechanical theory, rather than by reference to the epistemological criteria written into them. The fundamental difference between perception and imagination, for instance, is not one of vividness or any other such accidental property; it is an epistemological difference. To say that a person imagines a tree rather than perceives it is to say something about the status of what is claimed. To perceive is to see something that really is before one’s eyes; to imagine is to think one sees something that is not there. Similarly, to remember is to be right in a claim one makes about something in the past that one was in a position to witness, whereas to imagine is to be mistaken in what one claims. There are, of course, further questions about the mechanisms by means of which people perceive, imagine, and remember; and it could be that some such mechanical story as told by Hobbes might be true about such mechanisms. But in the language of such a story the basic epistemological differences between these mental performances could never be made, and although the mechanical story might give an account of some of the necessary conditions of such performances, it is difficult to see how it could ever serve as a sufficient explanation of them.

THOUGHT. The same general critique concerning neglect of epistemological criteria must be made of Hobbes’s treatment of thought, which he equated with movements of some substance in the head. There may be movements in the brain that are necessary conditions of thought, but no description of such conditions should be confused with what is meant by “thought.” We do speak of “the movement of thought,” but this is a description of transitions, as from premises to conclusions or from problems to solutions, not of movements explicable in terms of mechanical laws.

Even though Hobbes’s general account of thought was rather hamstrung by his obsession with mechanics, he nevertheless had some quite illuminating things to say about trains of thought, an account that owed more to Aristotle than to Galileo. Hobbes distinguished

"unguided" thought from that directed by a passionate thought or plan. Unguided thought followed principles that later came to be called principles of association—for example, spatiotemporal contiguity and similarity. Hobbes, however, made no attempt to formulate principles of this kind. He was much more interested in, and attached much more importance to, guided thought, in which desire for an end holds the train of thought together and determines the relevance of its content.

Hobbes distinguished two main types of regulated thinking. The first was the classic Aristotelian case of deliberation, where desire provides the end, and the means to this end are traced back until something is reached that is in a person's power to do. This faculty of invention is shared by the animals, but they do not share the other sort of guided thinking that Hobbes called prudence. In prudence the starting place is an action that is in a person's power to perform, and the store of past experience is used to speculate on its probable effects. In this case, deliberation leads forward to an end that is either desired or feared. Hobbes seemed to think that people's prudence is in proportion to the amount of past experience on which they can draw. This sounds improbable, for although children cannot be prudent, many old people miss the relevance of their past experience.

DREAMS. Dreams fascinated Hobbes. He attempted to determine what distinguishes them from waking thoughts and to develop a mechanical theory to explain them. He claimed that they lack coherence because they lack the thought of an end to guide them. Dreams consist of compounded phantasms of past sensations, for "in the silence of sense there is no new motion from the objects, and therefore no new phantasm." Dreams are clearer than the imaginations of waking men because of the predominance of internal motion in the absence of external stimulation. There is no sense of time in dreams, and nothing appears surprising or absurd.

There is an intimate connection between dreams and bodily states. Lying cold, for instance, produces dreams of fear and raises the image of a fearful object. The motions pass both from the brain to the inner parts and from the inner parts to the brain. So, just as anger causes overheating in some parts of the body, overheating of the same parts can cause anger and, with it, the picture of an enemy. Dreams are thus the reverse of waking imaginations. Motion begins at one end during waking life and at the other end during sleep. This tendency to project images produced by bodily states gives rise to belief in apparitions and visions. Hobbes's treatment of dreams typified his approach to such matters. He seemed uninterested in the epistemological questions to which they give rise, as, for instance, in the thought of his contemporary, Descartes.

PASSIONS. Hobbes's mechanical theory of human action hinged on his concept of "endeavour," by means of which he tried to show how the gross movements of the body in desire and aversion could be explained in terms of minute unobservable motions in the body. He postulated two sorts of motion in the body. The first is its vital motion, manifest in such functions as circulation of the blood, breathing, and nutrition, which proceeds without external stimulation or the help of the imagination. The second is animal motion, which is equivalent to such voluntary movements as walking and speaking. This is always "first fancied in our minds" and is produced by the impact of external stimuli on the sense organs, an impact that gives rise both to phantasms in the brain and to internal motions that impinge on the vital motions of the heart. If the motion of the blood is helped, this is felt as pleasure; if it is impeded, as pain. Pleasure, Hobbes said, is "nothing really but motion about the heart, as conception is nothing but motion in the head." In the case of pleasure, the spirits—which were thought of as vaporous substances flowing through the tubes of the nerves—are guided, by the help of the nerves, to preserve and augment the motion. When this endeavor tends toward things known by experience to be pleasant, it is called appetite; when it shuns what is painful, it is called aversion. Appetite and aversion are thus the first endeavors of animal motion. We talk about "will" when there is deliberation before acting, for will is "the last appetite in deliberating."

Hobbes's theory of the passions was an attempt to graft the traditional Aristotelian account of them onto his crude mechanical base. Love and hate are more or less the same as appetite and aversion, the only difference being that they require the actual presence of the object, whereas appetite and aversion presuppose its absence. These, together with joy and grief, which both involve foresight of an end rather than just an immediately perceived object, are the simple passions out of which others are compounded. Social life is a race for precedence that has no final termination save death. "So that in the first place, I put for a general inclination of all mankind, a perpetual and restless striving of power after power, that ceaseth only in death." To endure in the race requires foresight and scheming; to fail to compete is to die. A man who is convinced that his own power is greater than that of others is subject to what Hobbes called glory; its oppo-

site is humility or dejection. Pity is grief for the calamity of another, arising from imagination that a like calamity may befall ourselves. Laughter is the expression of sudden glory caused by something new and unexpected in which we discover some superiority to others in ourselves.

Hobbes also introduced motion into his theory of individual differences. He thought that such differences are derivative from differences in passions and in the ends to which men are led by appetite, as well as to the sluggishness or agility of the animal spirits involved in the vital motions of their respective bodies.

The basic difficulty in understanding Hobbes's theory of motivation arises from his attempt to underpin a psychology derived from introspection, from the shrewd observation of others, and from the tradition going back to Aristotle with a mechanical theory whose outline was only very briefly sketched. Perhaps the essential criticism of any such theory is that actions cannot be analyzed into mere movements because, in any action proper—as distinct from a nervous tic or a reflex—the movements take place because of an end that the person has in mind. This end is what makes the action one of a certain sort, and, provided that the movements are directed toward this end, an almost indefinite range of movements can form part of the same action. Similarly, the movements involved in raising one's hand can form part of quite different actions, depending on the purpose for which the hand is raised—for example, to signal, to test the direction of the wind, to stretch the muscles, and so on.

Having something in mind—which is part of the concept of "action"—is not a movement, still less a movement of some internal substance of the head, if this is what Hobbes really believed. But Hobbes was not at all clear on the relationship between movements, whether observable or unobservable, and the cognitive components of appetites, aversions, and the various passions. Indeed, he seems to have held an extremely paradoxical and overintellectualistic view about the cognitive component of the passions. For he saw that passions are to be distinguished by their objects and by the judgment of the possibility of attaining such objects, yet he injected into his account a bizarre kind of egocentricity. For Hobbes, in all cases of passions the notion of "self" was part of the content of cognition. He seemed to think that all such "phantasms" of objects, by reference to which the passions are to be distinguished, involve the thought of ourselves doing something or of our power to do something. Pity is thus seen as grief arising from our imagining ourselves in the same predicament as that of the one pitied. Hobbes's analysis of laughter palpably suffered from the same injection of egocentricity. Furthermore, how the highly sophisticated and narcissistic type of appraisal involved in the passions is to be reconciled with any attempt to represent them all as movements of the body and of some internal substance in the head is very difficult to determine.

For all its ambiguities, oversights, and obvious defects, Hobbes's psychology was remarkable, for he attempted to establish it as an objective study untrammeled by theological assumptions. To suggest that man is a machine was a great step forward in thought. Even though the hypothesis is probably untenable, it marked the beginning of the effort to use scientific methods and objective concepts in the sphere of human behavior. In the seventeenth century this was a novel undertaking, as well as a dangerous one.

ETHICS

Hobbes thought that, by employing the resolutive method, he could demonstrate the absolute necessity of leagues and covenants and the rudiments of moral and civil prudence from his two principles of human nature—"the one arising from the concupiscible part, which desires to appropriate to itself the use of those things in which all others have a joint interest; the other proceeding from the rational which teaches every man to fly a contranatural dissolution, as the greatest mischief that can arrive to nature." These two principles underlie Hobbes's account of the personal good, as well as his account of civil duty.

Hobbes was scornful of the notion that "good" and "evil" name any metaphysical essence. These words are "ever used with relation to the person that useth them: there being nothing simply and absolutely so; nor any common rule of good and evil, to be taken from the nature of the objects themselves." They name objects of our desires and aversions. We call a horse "good," for instance, because it is "gentle, strong, and carrieth a man easily." The desires of the individual determine what qualities are selected to furnish the ground for saying that an object is good.

Hobbes introduced a further refinement of this theory when he contrasted short-term goods with long-term goods. "Reason," he said, "declaring peace to be good, it follows by the same reason, that all the necessary means to peace be good also." This he contrasted with the sway of irrational appetite, whereby men "greedily prefer the present good." He thought that a man might not desire peace at a particular moment when influenced by some insistent desire; but when he sat down soberly in a cool

hour, he would see that peace is a necessary condition of satisfying most of his desires in the long run. Thus, peace is something that he must desire both because of his fear of death and because of the other things he desires to do that a state of war would make impossible.

Hobbes was a nominalist, and he thought that all words have meaning, as if they were some kind of name. He did not see, as Berkeley seems to have seen a little later, that words such as *good* have a prescriptive function and cannot be treated merely as if they were names. To say that something is good is to say that it is what it ought to be; it is to commend it. But also it implies that there are grounds for such commendation. It is to guide a person by suggesting grounds for his choice; it is not to order him or goad him. Hobbes saw that "good" is always thus connected with reasons, but he gave a very circumscribed account of what such reasons must be like, that is, characteristics of things desired. This was modified somewhat by what he said a man desires insofar as he uses his reason, that is, insofar as his "rational" as well as his "concupiscible" nature is involved. Hobbes's account of what a man desires would not be implausible if his account of human nature were acceptable, for then what men must desire could be predicted. But, if his account of human nature is rejected as oversimple, there cannot be quite such a tight connection as Hobbes suggested between "good" and what is, or will be, desired.

The connection is probably looser; given that words such as *good* have the practical function of guiding people's choices, it would be impossible to explain their effectiveness in this function if it were not generally the case that what was held up as good was something that people in general wanted. But it does not follow from this that any particular individual desires, or must desire, what is held up to him as good. Indeed, half the business of moral education consists in drawing people's attention to characteristics of things that they ought to desire but do not in fact desire.

STATE OF NATURE AND LAWS OF NATURE. Morality is not concerned simply with the pursuit of personal good; it is also concerned with the acceptance of rules that limit the pursuit of good when it affects that of others. A tradition going back to the Stoics held that there was a small corpus of such rules, called the law of nature; these rules, which were universal preconditions of social life, did not depend, as do custom and law, on local circumstances. The Dutch jurist Hugo Grotius regarded this law of nature as a self-evident set of principles binding on all men (on kings as well as on their subjects) that would provide a rational basis for a system of international law; it was, he claimed, fundamental in the sphere of social rules in the same sort of way that Galileo's postulates were fundamental in the realm of nature. Morals could be brought within the expanding empire of the mathematical sciences.

Hobbes, therefore, was not original in his claim that "the true doctrine of the laws of nature is the true moral philosophy," nor was he original in likening its precepts to axioms. What was original was his claim that its precepts were axioms of prudence, insofar as "prudence" implies considerations limited to those that affect only the agent. For Grotius, the maintenance of society was a major need of man as a social animal, irrespective of purely private benefits. Hobbes, however, maintained that more or less the same set of rules that Grotius regarded as binding (such as keeping faith and fair dealing) could be shown to be axioms that must be accepted by any man who is both rational and afraid of death. "All society, therefore, is either for gain or for glory; that is, not so much for love of our fellows as for love of ourselves."

Man, Hobbes argued, shuns death "by a certain impulsion of nature, no less than that whereby a stone moves downward." This is what saves man from anarchy and civilizes him, for if man were driven merely by his "concupiscible" part, there would be no society, and the life of man would be "solitary, poor, nasty, brutish, and short." Men are equal enough in body and mind to render negligible any palpable claims to superior benefits, and even the weakest is able to kill the strongest. But man's fear of death brings him up short in his pursuit of power and leads him to reflect upon the predicament of a state of nature. His reason tells him that peace is necessary for survival and also "suggesteth certain articles of peace, upon which men may be drawn to agreement. These articles are they, which otherwise are called the Laws of Nature." One of these laws is that "men perform their covenants made." In this way Hobbes claimed to demonstrate "the absolute necessity of leagues and covenants, and thence the rudiments both of moral and of civil prudence."

Hobbes's demonstration gave only the semblance of validity because he isolated the concupiscible and rational aspects of man's nature from each other and, as in a Galilean imaginary experiment, explored the consequences of each independently. Given only man's self-assertion, then there must be a state of nature; given only his overwhelming aversion to death, then he must accept the conditions necessary for avoiding death. These axioms of prudence are hypothetical in relation to man's

assumed fear of death. They are rules that a rational man must accept insofar as he wants to avoid death. But men are only partly rational and, although they have an overwhelming fear of death, they also want other things, such as power and glory. Presumably Hobbes, like Machiavelli, could also have laid down rules for obtaining power and glory that would have borne no resemblance to the laws of nature. Thus, Hobbes could not have been trying to show that virtue, as defined by adherence to the laws of nature, is natural to man or a deduction from his nature, as have many thinkers who have adopted a psychological starting point. Indeed, the general relationship between Hobbes's psychology and his ethics is too obscure for us to know quite what he was doing.

The key to Hobbes's "demonstration" really lies in what he did with it, for he went on to point out that the laws of nature are only theorems that any rational man would accept. Since these laws need the backing of the sword to ensure peace, men have need of a "common power to keep them in awe, and to direct their actions to the common benefit." The rationale of Hobbes's demonstration can now be seen, for at the time that Hobbes was writing, England was precariously poised between anarchy and civil disorder. Hobbes's analysis was a Galilean "resolution" of such a situation into the simple components of human nature that formed its basis. He pointed out that, insofar as men want peace and security (and all men do want this, although they want other things as well), then they must see that, human nature being what it is, there are certain means that they must accept if they are to have what they want. It is irrational to want something and yet to refuse to take the only means that will ensure that what is wanted is obtained. Since the acceptance of social rules is based only on the fear of death, it is only the fear of death that will ensure that these rules are obeyed. Men therefore cannot have the peace they all desire unless they accept the sword of the sovereign that will make death the consequence of breaking the rules that are a necessary condition of peace.

DETERMINISM AND FREE WILL. The indeterminate position of Hobbes's psychology in relation to his ethics was encouraged by his belief in determinism—or "necessitation," as he usually called it—which he outlined in his controversy with Bishop Bramhall. Hobbes denied that there is any power in men to which the term *will* refers; what is commonly called will is but the last desire in deliberating. Furthermore, he argued, only a man is properly called free, not his desires, will, or inclinations. The liberty of a man "consisteth in this, that he finds no stop, in doing what he has the will, desire, or inclination to do." Liberty is "the absence of all the impediments to action that are not contained in the nature and intrinsical quality of the agent." To speak of liberty is not to make any suggestions about the determinants or absence of determinants of man's deliberations or decisions; it is to suggest that man is not externally constrained in his actions. There is, therefore, no contradiction in saying that a man acts freely and that his actions are also determined. Since all actions have causes and thus are necessitated, it is pointless to use "free" in the sense of "free from necessitation," as distinct from "free from compulsion." There are no such actions, although we may think that there are because we are ignorant of the causes of actions.

There is much to be said for Hobbes's recommendation on the use of the word *free;* many others, such as John Locke and David Hume, have followed him in confining it to the absence of constraint on a man's actions. But Hobbes's claim that all actions are necessitated is not so straightforward. Certainly he was right in suggesting that all actions are explicable—if that is what is meant by saying that they have causes—but so many different things can count as causes, ranging from deliberation and understanding to a stab of pain or a crack on the skull. Since Hobbes thought of man as a natural machine, he therefore viewed all causes as mechanical pushes. His doctrine carried the suggestion that the behavior of men is not only explicable but also somehow unavoidable because men's decisions and choices are simply manifestations of internal pushes.

Significantly enough, Bramhall did not object to Hobbes's doctrine insofar as it related to actions shared with animals or to spontaneous actions. What he could not allow was that voluntary actions, which follow on election and deliberation, should also be "necessitated." Bramhall pointed out the difficulties of likening actions and the grasp of objects and of means of obtaining them, which are inseparable from the concept of "action," to processes in nature explicable in terms of antecedent motions. In this contention Bramhall was substantially right, for although actions may involve movements, they are not reducible to movements.

Hobbes also disagreed with Bramhall on the implications of his doctrine of "necessitation" for moral judgments and for the operation of the law. Bramhall argued that if human actions are necessitated, then praise and blame, reward and punishment, are both unjust and vain. To the charge that they are vain, Hobbes replied that they are to be viewed as further determinants of choice. Praise and blame, reward and punishment "do by example make and conform the will to good and evil." To the charge of

injustice, Hobbes argued that "the law regardeth the will and no other precedent causes of action"; also that punishments annexed to breaches of the law function as deterrents and necessitate justice. He went out of his way to distinguish punishment from acts of revenge or hostility and to stress its deterrent purpose, which is a sound position. Hobbes saw clearly that retribution is part of the meaning of punishment, but that it is the connection with authority that distinguishes it from other sorts of retributive acts. He also saw that, although retribution may be written into the meaning of punishment, its justification is not therefore necessarily retributive. Rather, it is to be justified for its preventive and deterrent function.

POLITICAL PHILOSOPHY

In his political philosophy Hobbes tried to conceptualize the relationship between the new nation-state, which had been emerging under the Tudors, and the individual citizen, who could no longer be regarded simply as having a set place in a divinely instituted order. In the old medieval society a man was bound by ties attaching to his status and by duties prescribed for him by the church. Tradition was the main form of social control, and traditions stretching back into the distant past assigned to a man his relatively fixed place in society. Aristotle's doctrine of natural kinds and natural places and his account of man as a social animal provided a fitting naturalistic foundation for the theological worldview that was accepted by rulers and ruled alike. But with the rise of individualism and the social mobility that accompanied the rise of commerce and capitalism, this old conception of man in society no longer applied. Men had shaken off the ties of their guilds and local communities, and the new natural philosophy was beginning to render the naturalistic foundations of the former worldview untenable.

Hobbes's picture of life as a race, in which we "must suppose to have no other good, nor other garland, but being foremost," was a gruesome caricature of an age of individualism, restless competition, and social mobility. But if the fetters of tradition were being cast away, what other form of social control could take its place to prevent the anarchy of a state of nature? The answer was to be found, of course, in the increasing executive power of the state and in the growth of statute law, together with the development of the individual conscience, whereby regulation from within replaced the external authority of the Catholic Church. Hobbes distrusted the anarchic tendencies of the individual conscience as much as he loathed the extramundane authority of the Church of Rome. Both were to be banished, along with traditional ties; civil society could be reconstructed as a simple mechanical system.

SOCIAL CONTRACT. Hobbes had a model ready at hand by means of which he might present his Galilean analysis of the rationale of civil society—the social contract theory. The social contract theory, despite its obvious flaws, was an attempt to rationalize political obligation, to substitute an intelligible bargain for mystifying appeals to tradition and divine right.

The contract theory was resorted to mainly by those who wanted to challenge the absolutist claims of monarchs, to uphold the claims of the common law, or to lay down some sort of moral limits on control and interference by the central executive. Hobbes's feat was to employ this model to demonstrate that absolutism is the only possible logical outcome of consistent concern for individual interests. Indeed, he prided himself on grounding the authority of sovereigns, as well as the liberty and duty of subjects, upon axioms of human nature rather than on tradition and supernatural authority. In his attitude toward tradition and divine right, he was at one with the defenders of government by consent. But because of his overriding concern for security, and because of his rather depressing estimate of human nature, he came to the somewhat gleeful conclusion—highly displeasing to those who believed in government by consent—that absolutism could be the only rationally defensible form of government.

Hobbes did not seriously consider the social contract, as some did, as a quasi-historical hypothesis on how civil society might have come into existence. In his account the contract was featured as a framework for a Galilean resolution of civil society into its simple elements. Hobbes imagined the individual in a state of nature as having an unlimited right to "protect his life and members" and "to use all the means, and do all the actions, without which he cannot preserve himself." But he also has a right to all things "to do what he would, and against whom he thought fit, and to possess, use, and enjoy all that he would, or could get." Hobbes here was employing a very strange concept of right, for usually, when we talk about a right, we are indicating a rule that protects or should protect a person from interference in the doing of something that he might want to do. Hobbes, however, used the term in this way to talk about both what a person is entitled to do (when it is correlative with duties of noninterference on the part of others) and what a person cannot be obliged to renounce. When Hobbes declared that men have a "right of self-

preservation," he meant not that an individual is entitled by some rule (of law, tradition, or morals) to life but that he cannot be obliged to renounce it because it is psychologically impossible for him to do so. "Natural rights" therefore have a quite different meaning in Hobbes's writing than in the works of Locke, Samuel von Pufendorf, and other such defenders of natural rights. In these classical theories, natural rights are interests protected by natural law against the interference of others. Hobbes's natural-law theory is not connected in this way with his rather bizarre concept of natural rights.

Hobbes's "rights" of nature are derivative from man's tendency to assert himself and to seek power. But, as already shown, Hobbes held that man would also be driven by his fear of death to accept certain laws of nature, the second of which prescribed that every man should lay down his right to all things and "be contented with so much liberty against other men, as he would allow other men against himself." This could be done either by not interfering with others' enjoyment of their rights or by transferring one's right to another, in which case the transferrer is obliged not to hinder the recipient. Injustice consists in hindering a person whom it is a duty not to hinder. The mutual transferring of such rights is called a contract, and the third law of nature is "that men perform their covenants made."

COMMONWEALTH. Hobbes deduced a mutual transfer of rights from his postulate of rational action under the impetus of fear. But men are not yet safe, for there may be danger in keeping covenants and it may be, on occasion, in people's interest to break them. "And covenants, without the sword, are but words, and of no strength to secure a man at all." Matters must be arranged so that it will never be in anyone's interest to break covenants, which cannot exist where there is no "common power" to enforce them. Thus, a social contract must be presumed in which it is as if every man should say to every other man, "I authorize and give up my right of governing myself, to this man, or to this assembly of men, on this condition, that thou give up thy right to him, and authorize all his actions in like manner." This contract unites the multitude into one people and marks the generation of "that great leviathan, or rather, to speak more reverently, of that mortal God, to which we owe under the immortal God, our peace and defence." The definition of commonwealth is, therefore, "one person, of whose acts a great multitude, by mutual covenants one with another, have made themselves every one the author, to the end he may use the strength and means of them all, as he shall think expedient, for their peace and common defence." The person that results is called sovereign, and everyone else is his subject. The sovereign is created by the contract but is not party to it. Thus, the people rule even in monarchies; a multitude becomes a people by having some device, such as that of representation, by means of which decisions binding on all are made on behalf of all. Some such "covenant" is implicit in speaking of a commonwealth as a people, as distinct from a multitude of men.

Up to this point there is much to be said for the sort of analysis that Hobbes gave, although some of its details are peculiar. He had considerable insight into the sort of thing we mean when we speak of a civil society, as distinct from a mere multitude of men. He saw clearly that societies are not natural wholes like toads, turnips, or colonies of termites. They exist because individuals act in accordance with rules that can be rejected, broken, or altered; they are artificial wholes. Therefore, if we are to speak of the "will" or "decision" of such an entity, there must be some higher-order rules of procedure, such as that of representation, by reference to which what is to count as a corporate decision is constituted. Individuals or groups of individuals are put in authority for such a purpose.

When Hobbes proceeded to the more concrete details of what must constitute the duties of rulers and subjects, however, he was not equally convincing, for this next step depended on his questionable account of human nature. The basic principle of human nature revealed by his Galilean resolution was "that the dispositions of men are naturally such that, except they be restrained through fear of some coercive power, every man will dread and distrust each other." No motive in human nature, except the fear of death, is strong enough to counteract the disruptive force of man's self-assertion. The fear of death must, therefore, be the explanation of the existence of civil society (insofar as there is a social order and not anarchy), and security must be the sole reason for the institution of the social order; there is simply no other reason for which men could be induced to give up their natural right to self-assertion. Since this is the sole reason for having a commonwealth, it follows logically that a commonwealth must be devised that will accomplish the end for which it exists. Sovereignty must be perpetual, undivided, and absolute, for to divide or limit sovereignty would be to risk anarchy; and such limitation would be illogical because it would be inconsistent with the raison d'être of sovereignty. *Salus populi suprema lex* (The safety of the people is the supreme law). Moreover, complete safety entails complete submission to an absolute sovereign. Thus, absolutism is the logical conse-

quence of government by consent, once the real interest of individuals, which is the presupposition of the institution of commonwealth, has been clearly understood.

There are two obvious flaws in this stage of Hobbes's argument. The first is the assumption that the desire for security, deriving from the fear of death, is the sole reason for the institution of commonwealth, a reason that Hobbes more or less wrote into the meaning of "commonwealth." It is obviously a very important reason, but that it should be the only reason is plausible only if Hobbes's psychology were to be accepted. Even so, Hobbes should not have written the reason for instituting a commonwealth into what is meant by "commonwealth." The second flaw was well brought out by Locke, who argued that, even if security were the sole reason for the institution of commonwealth, absolute authority is a dangerous expedient from the point of view of individual interest. For the hypothesis is that the timid individual would exchange the possible threat to life presented by 100,000 men, all of whom individually might attack him, for the threat to his life made possible by the arbitrary authority of one man who has 100,000 men under his command. "Are men so foolish that they take care to avoid what mischiefs may be done them by polecats or foxes, but are content, nay think it safety, to be devoured by lions?"

Hobbes was led to his advocacy of undivided sovereignty by his interest in constitutional and legal matters. When Hobbes was writing, there was a clash between the higher-order principles of common law and of statute law. The common-law principle that custom, as interpreted by the judges, is to be consulted in declaring what the law is, existed alongside the principle of statute law, that rules laid down by a determinate body or person (for example, Parliament or the king) determine what the courts must recognize as valid law. Statute law was on the increase during this period, and it was intolerable to any clearheaded man that these two principles should operate side by side. Hobbes advocated the unambiguous supremacy of the principle of statute law and the abolition of common law. The need to introduce clarity and coherence into the confused constitutional situation that prevailed in Hobbes's time was obvious enough. But for Hobbes to suggest that it was a logical truth that there must be an absolute sovereign in any commonwealth was to introduce dubious logical deductions into a field where a solution was more likely to be found by practical adjustments and compromises that reflected the strength of competing interests and were consonant with deep-seated traditions.

One of the traditions that Hobbes's geometric solution ignored was that of the liberty of the subject. In Hobbes's view, civil liberty lay "only in those things, which in regulating their actions, the sovereign hath praetermitted." It is unlikely, Hobbes suggested, that laws would be necessary to regulate buying and selling, and choice of abode, diet, a wife, a trade, and education. But whether such laws are necessary is entirely up to the sovereign. The liberty of the subject also consists in the lack of proscription of such acts that it would be vain to forbid because they are psychologically impossible for the subject to refrain from committing. These acts involve the right of the subject to preserve himself and to resist imprisonment. Hobbes also suggested that "in the act of submission consisteth both our obligation, and our liberty." Both the obligation and the liberty are to derive from the words "I authorize all his actions," which the subject is imagined to have expressed in instituting a commonwealth. The subject is released from his obligation only if the sovereign fails to do what he is there to do, namely, to guarantee security. This marks the extent of the subject's much-lauded "right to resist." Presumably Hobbes meant to stress that subjects submit voluntarily to authority. This is true enough, but what it has to do with the liberty of the subject, in any straightforward sense of "liberty," is difficult to grasp.

LAW. Hobbes's concept of the role of natural law, once the law of the state had been established, was not altogether clear. He maintained that the laws of nature were "but conclusions, or theorems concerning what conduceth to the conservation and defence of themselves; whereas law, properly, is the word of him that by right hath command over others. But yet, if we consider the same theorems, as delivered in the word of God, that by right commandeth all things; then they are properly called laws." These "laws" always obligate *in foro interno*—that is, in matters of private conscience—in prescribing a general readiness of mind; but *in foro externo,* that is, in actions, the laws may not be obligatory if certain conditions, such as peace and security, are absent. Such conditions, when present, will in fact render it to the interest of the subject that he follow the laws of nature. A law properly so called always obligates *in foro externo* because of its source in the command of the sovereign, as well as because civil society, by definition, provides the conditions of security and the sanction that will make it always against a man's interest to disobey it. But do the laws of nature oblige *in foro externo,* if not incorporated in the civil law, when the security of civil society prevails? This depends on how seriously Hobbes meant his reference to

theorems as authoritative edicts from God, for such derivation would give them a determinate source, as in the case of laws properly so called. Some take Hobbes seriously and claim that he really thought that the laws of nature oblige *in foro externo* as well as *in foro interno* whenever conditions of security prevail. Others hold that Hobbes never really thought that laws of nature oblige in a full sense *in foro externo* because his reference to their authoritative source is but a tactful concession to piety. He really thought of them merely as axioms of reason that oblige in a full sense only when they are issued by a temporal sovereign as commands and when conditions of security, together with sanctions, prevail in civil society.

Hobbes took this somewhat ambiguous view about the status of natural laws (or moral precepts) because of his extreme hardheadedness about laws properly so called. Law, he held, is the command of the sovereign, "the word of him that by right hath command over others." It is authority, not conformity with custom or reason, that makes a law. In this forthright view he was attacking the fiction of the common law that the law was there to be discovered, immanent in the customs of the people.

Whatever the merits of Hobbes's view—later adopted by the analytic school of John Austin—that laws are commands, Hobbes made a valuable contribution in helping to distinguish questions about law that are often confused. The question "What is a law?" should be distinguished from such other questions as "Is the law equitable or reasonable?" and "What makes a law valid?" Hobbes argued that a law is simply a rule issued by someone in authority. Whether it is reasonable or equitable is a further question, as are the questions of its validity, of its conformity with custom, and of the grounds on which a man could be obliged to obey it.

To claim that laws are commands was an oversimple and misleading way to bring out the prescriptive force of laws. But it was useful insofar as it connected law with authority, for laws, like commands, are utterances issuing from people in authority. In stressing the necessary connection between law and authority, Hobbes made an important contribution to political philosophy, for there is no necessary connection between authority and moral precepts or "laws of nature."

On the question of the person or body of men by whose authority laws should be made, Hobbes was more open-minded than is often realized. He thought that this was not a matter that could be demonstrated; it was a matter of factual argument. He believed that the relative advantages of each form of government had to be considered in the light of the sole end of security. It was a factual matter which type of government was most likely to promote such an end. On the whole, he argued, monarchy is preferable because it is more likely to be undivided, strong, and wise.

RELIGION

At the time Hobbes wrote, ethics and politics were inseparable from religion. Even the Royal Society was founded by men who believed that science would reveal more of the details of God's creation and thus enhance his worship. Hobbes was one of the pioneers in the process of distinguishing religious questions from other sorts. He rigorously excluded theology from philosophy and tried to map the proper domains of faith and knowledge. He outlined a theory of the causes of religion and superstition and discussed the grounds of religious belief, and he conducted an elaborate inquiry into the use of various terms in the Scriptures. But all this analysis and theorizing was subordinate to his main interest in religion as a possible source of civil discord. It is seldom realized that more than half of *Leviathan* is concerned with religious matters, with Hobbes trying to defend the "true religion" from both Catholicism and the priesthood of all believers. He saw clearly that these doctrines were two of the main obstacles in the way of the absolutism that he advocated.

Hobbes made some interesting speculations about the natural causes of religion, which he said were "these four things, opinion of ghosts, ignorance of second causes, devotion toward what men fear, and taking of things casual for prognostics." These seeds of religion could be cultivated according to natural invention, which leads to superstition and nature worship, or according to God's commandments. "Fear of power invisible, feigned by the mind, or imagined from tales publicly allowed, religion; not allowed, superstition. And when the power imagined is truly such as we imagine, true religion."

NOTION OF GOD. What, then, constituted true religion for Hobbes? To reasonable men, God's commands amounted to the laws of nature. God's nature, however, was a much more baffling matter, even for a rational man. Certainly God must have "existence," which Hobbes took to be an attribute of God, in spite of his remarks elsewhere about the ambiguities of the verb "to be." In *Leviathan* Hobbes held that God is the cause of the world, "that is, a first and an eternal cause of all things; which is that which men mean by the name of God." In his later *De Corpore,* however, he indicated the difficulties in the

notion of an unmoved mover. This was a difficult question for philosophers to determine and had better be handed over for decision to the lawful authorities. Hobbes also stressed God's irresistible power and maintained that the only solution to the problem of evil was to be found in this power. Did not God reply to Job: "Where wast thou, when I laid the foundations of the earth?" Job had not sinned; his suffering was an unfortunate consequence of God's manifestation of power.

The main function of reason, however, is to show what God cannot be—at ease, finite, figured, having parts, occupying a place, moved or at rest, plural, and having passions, rational appetite, sight, knowledge, and understanding. If we rely on natural reason, we must either qualify God in a negative way by adjectives, such as "infinite" and "incomprehensible," or by a superlative, such as "most high," and an indefinite, such as "holy," which are not really descriptions of his nature but expressions of our admiration. Thus, rational disputations about the nature of God are pointless and a dishonor to him, "for in the attributes which we give to God, we are not to consider the signification of philosophical truth; but the signification of pious intention, to do him the greatest honour we are able." The sovereign, therefore, must decide on God's attributes; and public, uniform worship must be instituted.

REASON AND REVELATION. Reason, however, should not be "folded up in the napkin of an implicit faith, but employed in the purchase of justice, peace, and true religion." There is nothing in God's word contrary to reason. We must, however, be prepared in this world "to captivate our understanding to the words; and not to labour in sifting out a philosophical truth by logic, of such mysteries as are not comprehensible, nor fall under any rule of natural science." Reason should be kept very much to the fore when one is confronted with those who claim revelation, for if a man says that God spoke to him in a dream, this "is no more than to say he dreamed that God spoke to him." There are psychological explanations of such phenomena that cast doubt on their reliability as valid communications with God.

Dreams, visions, and inspiration, however, should not be dismissed altogether, for it is by such means that prophets have been informed of the will of God. What is needed are criteria for detecting true prophets. Hobbes suggested two necessary criteria: the working of miracles and the teaching of doctrines not at variance with those already established. Since miracles had by then ceased, there was no sign left to single out true prophets. And, in any case, the Scriptures, since the time of Jesus, had taken the place of prophecy.

Reliance on the Scriptures, Hobbes realized, is not altogether straightforward. Even supposing that it could be decided which books are authentic, and that the sovereign, by his authority, could make their teaching law, there is still the problem of what many of the terms used in the Scriptures mean. Hobbes went through most of the key terms in the Scriptures, giving meaning to them in a way consistent with his mechanical theory. He argued, for instance, that God must have a body and that the proper signification of "spirit" in common speech is either a subtle, fluid, and invisible body or a ghost or other idol or phantasm of the imagination; it may also have a figurative use in such a phrase as "spirit of wisdom." "Angels" signify images raised in the mind to indicate the presence of God. Hobbes made acute remarks about the nature of miracles that mingled radical probing with subtle irony (indeed, one often wonders whether his whole treatment of "the true religion" is not a colossal piece of irony).

On the relationship between church and state, Hobbes of course adopted an uncompromising Erastian position. A church he defined as "a company of men professing Christian religion, united in the person of one sovereign, at whose command they ought to assemble, and without whose authority they ought not to assemble." There is, therefore, no universal church to which all Christians owe allegiance, for there is no supreme sovereign over all nations.

Hobbes concluded *Leviathan* with his famous section on the Kingdom of Darkness, in which he castigated superstition and Catholicism as enemies of the true religion. The papacy, he remarked "is no other than the ghost of the deceased Roman empire, sitting crowned upon the grave thereof." The papacy ruthlessly exploits the fears of ignorant men to perpetuate the power of unscrupulous priests as a rival to the secular power.

Hobbes held that there is only one article of faith necessary for salvation: that Jesus is the Christ. On what authority did such a belief rest? Hobbes had some interesting things to say about the difference between knowledge and faith. The object of both is propositions, but in the case of knowledge we consider the proposition and call to mind what its terms signify. Truth here is a matter largely of following the consequences of our definitions. But when reasons for assent derive "not from the proposition itself but from the person propounding, whom we esteem so learned that he is not deceived, and we see no reason why he should deceive us; our assent, because it grows not from any confidence of our own, but from

another man's knowledge, is called faith." Faith, therefore, depends on our trust in a man rather than on our grasp of truth. The faith that Jesus is the Christ must therefore come from the Scriptures and our trust in those who wrote them. But who is to interpret them? "Christian men do not know, but only believe the Scripture to be the word of God." St. Paul said, "Faith cometh by hearing," and that, according to Hobbes, means listening to our lawful pastors, who are appointed by the sovereign to interpret the Scriptures for us. Charles II and Cromwell must have been flattered by the magnitude of the problems on which they were required to issue authoritative edicts: the creation of the world, God's attributes, the authenticity of miracles, and the proper interpretation of the Scriptures. Hobbes regarded religion more as a matter of law than of truth.

Hobbes's treatment of religion leaves obscure exactly what he himself thought about such matters. His technique was always to push radical probing to the limit, and when the basis for the traditional doctrines seemed about to be cut away, the sovereign was summoned as a sort of deus ex machina to put everything in its orthodox place. Hobbes was obviously extremely skeptical about what could be demonstrated in the sphere of religion, but it is difficult to say whether his suggestion that the sovereign should pronounce on such matters as the creation of the world and the attributes of God was a subtle piece of irony, a pious protestation to protect himself against the charge of atheism, or yet another manifestation of his overwhelming conviction that there must be nothing touching the peace of the realm that the sovereign should not decide.

See also Aristotelianism; Aristotle; Authority; Bacon, Francis; Definition; Descartes, René; Determinism, A Historical Survey; Determinism and Freedom; Dreams; Galileo Galilei; Gassendi, Pierre; Geometry; Grotius, Hugo; Harvey, William; Human Nature; Hume, David; Images; Kepler, Johannes; Laws of Nature; Locke, John; Logic, History of; Machiavelli, Niccolò; Mersenne, Marin; Motion, A Historical Survey; Peace, War, and Philosophy; Sensa; Social Contract; Spinoza, Benedict (Baruch) de; Thucydides; Universals, A Historical Survey.

Bibliography

WORKS BY HOBBES

English Works of Thomas Hobbes, 11 vols., and *Opera Philosophica* (Latin works), 5 vols, edited by Sir William Molesworth. London: John Bohn, 1839–1845; reprinted Oxford, 1961. All quotations in the text of this entry are from this edition of the English works.

The Elements of Law, Natural and Political, edited by Ferdinand Tönnies. London: Cambridge University Press, 1928. This also includes extracts from the *Little Treatise, or A Short Tract on First Principles* and *Excerpta de Tractatu Optico.*

Leviathan, edited by Michael Oakeshott. Oxford, 1947; New York, 1962. Latter is paperback with introduction by Richard S. Peters.

Body, Mind, and Citizen, edited by Richard S. Peters. New York, 1962. Paperback, with introduction by Peters, that contains selections from the Molesworth and Tönnies editions.

WORKS ON HOBBES

Aubrey, John. *Brief Lives,* edited by E. O. Dick. London: Secker and Warburg, 1950. Contains biography of Hobbes.

Brandt, Frithiof. *Thomas Hobbes' Mechanical Conception of Nature.* Translated by Vaughan Maxwell and Annie I. Fausbøll. London: Hachette, 1928.

Laird, John. *Hobbes.* New York and London: Oxford University Press, 1934.

Peters, Richard S. *Hobbes.* Harmondsworth, U.K.: Penguin, 1956. Paperback.

Robertson, G. C. *Hobbes.* London and Edinburgh, 1886.

Stephen, Leslie. *Hobbes.* London: Macmillan, 1904.

Strauss, Leo. *The Political Philosophy of Thomas Hobbes.* London: Clarendon Press, 1936; Chicago, 1963. The 1963 edition is paperback.

Warrender, J. H. *The Political Philosophy of Thomas Hobbes.* Oxford: Clarendon Press, 1957.

Watkins, John W. N. *Hobbes's System of Ideas.* London, 1965.

R. S. Peters (1967)

HOBBES, THOMAS [ADDENDUM]

LIFE

Thomas Hobbes, still widely regarded as the greatest of English philosophers, was born on April 5, 1588, in Malmesbury, England, the son of a clergyman who later disappeared into London and left his sons to be raised by their uncle. Thomas died on December 3, 1679, at Hardwick Hall and was buried at Ault Hucknall, having lived a long and eventful life. After study at Magdalen Hall in Oxford, Hobbes was awarded the degree of BA in 1608 and was appointed tutor to William (1591–1628), the son of William Lord Cavendish (d. 1626); he lived much of the rest of his life as a member of the Cavendish household. This position gave him an opportunity, otherwise unlikely because of his relatively humble beginnings, for travel and to meet many of the leading intellectuals of the age, including Marin Mersenne, Pierre Gassendi, and Galileo Galilei. In England he was loaned by the

Cavendishes to Francis Bacon, to whom he acted as amanuensis; according to John Aubrey (1950), Bacon claimed that Hobbes was his best amanuensis because he understood what Bacon was talking about. Hobbes developed a great interest in mathematics and claimed geometry as the model for philosophical work; he took geometry to be descriptive of the properties of space and had little time for uninterpreted calculi, especially algebra. This mathematical interest led him to work for some time on ballistics with the Duke of Newcastle (1592–1676; a Royalist leader during the English civil wars and cousin to Hobbes's employer, who had become the first Earl of Devonshire) and to be appointed in 1646 as reader in mathematics to the Prince of Wales in Paris.

There is a lack of detailed knowledge of Hobbes's life. While Miriam M. Reik (1977), Arnold A. Rogow (1986), and A. P. Martinich (1999) discuss various aspects about Hobbes's life, the most interesting source to read on this subject remains Aubrey (1950).

SYSTEM AND SCIENCE

Though rarely mentioned in modern scientific literature, Hobbes was known in his own day mainly as a natural scientist. Despite increased interest in his work on morals and politics over the last century, no notice is taken of his work in the hard sciences, and little by any scientists except in political science (especially international relations) and economics. In his moral and political theory, Hobbes took his main contribution to be basing his enquiries on scientific principles and thus turning morals and politics into sciences. Hobbes addressed political issues of the day and his writings should be considered in that context, but he was not merely joining in political argument; in his own view, he was changing it by making it scientific and thus capable of producing definitive answers. The period was one of considerable argument about the appropriate methods for science, one account of Hobbes's contribution to this being Steven Shapin and Simon Schaffer's *Leviathan and the Air Pump: Hobbes, Boyle, and the Experimental Life* (1985), which deals with the dispute between Hobbes and Boyle about the value of the experimental method.

Hobbes claimed that he had created a system of philosophy with continuous argument from physics to politics. His grand plan encompassed a set of three books to carry this argument from the one point to the other, but events in England led him to publish the political work first rather than last: *De Cive* was published in Paris in 1642 with *De Corpore* not being published until 1655 and *De Homine* in 1658. By that time, Hobbes's other main political works had already appeared: *The Elements of Law* in two pirated parts in 1650 and *Leviathan* in 1651. There is much argument about the extent to which Hobbes's philosophy does form a system with the politics eventually following from the physics. Real debate on this issue springs from Leo Strauss's *The Political Philosophy of Hobbes: Its Basis and Its Genesis* (1936), in which he argues that Hobbes's political philosophy is independent of his natural philosophy. Major responses to Strauss can be found in John W. N. Watkins's *Hobbes's System of Ideas: A Study in the Political Significance of Philosophical Theories* (1965), M. M. Goldsmith's *Hobbes's Science of Politics* (1966), and Thomas A. Spragens Jr.'s *The Politics of Motion: The World of Thomas Hobbes* (1973). David Boonin-Vail's *Thomas Hobbes and the Science of Moral Virtue* (1994) includes discussion of the relationship between Hobbes's natural philosophy and his moral philosophy. Hobbes said that his political philosophy could be understood independently of his natural philosophy; many writers have taken him at his word and ignored the question of whether he has produced a system of philosophy of the sort he claimed.

RELIGION

In *The Political Philosophy of Hobbes: His Theory of Obligation* (1957) Howard Warrender argued that it was not science on which Hobbes's political philosophy was based, but God. This set the parameters for much discussion of Hobbes and provoked a renewed interest in Hobbes's religious beliefs. Most notable in this is Aloysius P. Martinich's *The Two Gods of Leviathan: Thomas Hobbes on Religion and Politics* (1992). S. A. Lloyd, in *Ideals as Interests in Hobbes's Leviathan: The Power of Mind over Matter* (1992) takes further the issue of the role of religious belief in Hobbes's moral and political philosophy.

METHOD

Hobbes officially espoused two methods of argument, which he claimed to have taken over from Galileo: the resolutive or analytical method, which involved taking things (at least in imagination) to find an explanation of them, and the compositive or synthetical method, which involved seeing how one had to put things together to construct what was to be explained. The compositive method, he says, is the only method that actually provides demonstration, and it is the method that should be used in teaching. Though he says that the analytical method will provide understanding of civil and moral philosophy to the less philosophically adept who lack a knowledge of geometry, one must conclude that it is the compositive

method he uses in his attempts to create a science of morals and politics.

The standard example of Hobbes's method, his discussion of the watch in the preface to *De Cive*, is a misleading example to take because it is an example of the resolutive method, taking the watch apart to understand how it works. More appropriate is the example of the circle in chapter 1 of *De Corpore*, since that is an example of the compositive method: Hobbes seeks real definitions and, specifically, definitions in terms of a method that is guaranteed to produce what is being defined. A circle, for example, is defined in terms of the movement of a point that remains equidistant from another point. Science, Hobbes says, seeks the causes of things, but care is needed in understanding what he means by "cause." He does not mean matter-of-fact causes in particular cases, as dropping a lighted match onto wood shavings might cause a house to burn down; the cause of something's being a circle, even if it was in fact stamped out by a die, is that it is a shape that could have been produced by moving a point so that it remained equidistant from another point. How a particular circle is, in fact, produced is not Hobbes's concern. His concern is with the compositive method, the setting out of a way of putting things together so as to guarantee that the outcome is whatever one is defining. His arguments about, for example, the state and the laws of nature should be read in the light of this method.

SCIENCE AND KNOWLEDGE

Hobbes distinguished science from "knowledge original," the latter being sense or remembrance of sense and the record of it in books being called history. Science is the knowledge of the truth of propositions and how things are called, the model Hobbes recommended being geometry. The truths of science, as Hobbes used the term and as is to be expected from the method he adopted, are conceptual or necessary truths on his account. One should take it, therefore, that his moral and political philosophy is not to be taken as history, but as applications of his compositive method.

HUMAN NATURE, MORALS, AND POLITICS

There was, in his own time, and still is, today, a common belief that Hobbes believed people to be essentially selfish. This idea of the role of self-interest in Hobbes's philosophy lies behind most of the interest that theoreticians of international affairs and economics show in his work. The most influential book taking this sort of interpretation of Hobbes is David P. Gauthier's *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes* (1969). This book led the reaction to Warrender's (1957) account. Gauthier introduced into discussion of Hobbes's moral theory the use of games theory, starting from the idea that Hobbes took all people to be naturally selfish and then, working from the further claim that, on Hobbes's account, both rights and obligations must have a prudential basis, constructed a Hobbesian argument that rational people must set up a sovereign if they are to achieve their own advantage. He concluded, furthermore, that the Hobbesian "moral" system is nothing more than universal prudence. Despite the work of Bernard Gert (1967) arguing persuasively that Hobbes did not regard people as naturally selfish, this general approach became for some time the received view of Hobbes. Jean Hampton employs the games-theoretical approach to interpreting Hobbes in its finest detail in *Hobbes and the Social Contract Tradition* (1986).

Unlike Hampton, some others taking a games-theoretic approach did not claim to be interpreting Hobbes's arguments but only to be using him as a starting point. Gregory S. Kavka's *Hobbesian Moral and Political Theory* (1986) makes that point explicit at the start. This is also true of many of the attempts to sort out a Hobbesian model for the study of international affairs, a move that has shared with the games-theoretic approach the assumption that the strong form of Hobbes's condition of mere nature (the relationship holding between people who have no common authority over them, so that each is bound only by his or her private judgment) is at least a coherent notion and can therefore constitute part of such a model.

Others followed Gert's (1967) interpretation of Hobbes, according to which Hobbes did not take everybody to be naturally self-interested. This alternative interpretation was present even when the games-theoretic approach was popular and has now come to dominate. Deborah Baumgold's *Hobbes's Political Theory* (1988) develops such an interpretation of Hobbes, as does R. E. Ewin's *Virtues and Rights: The Moral Philosophy of Thomas Hobbes* (1991), which also argues that Hobbes's condition of mere nature is not a coherent notion and was intended by Hobbes as part of a *reductio* argument, and Boonin-Vail's *Thomas Hobbes and the Science of Moral Virtue* (1994).

The different views of human nature led to different accounts of Hobbes's views of morals and politics. On the games-theoretic approach Hobbes was taken to be some sort of an ethical egoist as well as a psychological egoist. The main alternative to this in recent debate, mainly

sparked off by Gert's work (1967), when Hobbes is not taken to be a psychological egoist, has been that Hobbes is a virtues theorist (the view espoused by, for example, Ewin [1991] and Boonin-Vail [1994]), though interpretations of Hobbes's moral theory as depending on God (such as Lloyd's [1992] interpretation) still play a significant part in debate. Hobbes rejected Aristotle's moral theory because Aristotle took the relevant question to be whether a particular quality of character was a mean; the correct question was whether it tended to peace. The laws of nature, Hobbes says (among several different things he says about them), properly understood, are not laws, but are qualities fitting man for peace and for obedience. This suggests a virtues interpretation of Hobbes's moral theory, with his story of people agreeing on these laws of nature being the model produced by the compositive method to explain what these laws are.

Hobbes was clear that honest and intelligent people can disagree (though he often doubted the honesty or intelligence of those who disagreed with him), even in cases of morality. Where private judgments are in conflict and common action is necessary, disruption will occur. He makes the point firmly in, for example, chapter five of *Leviathan*. The only way of avoiding disruption, he believes, is to have a sovereign, a man or body of men, whose judgment is binding on those who disagree. Thus, the necessity of politics appears, in Hobbes's philosophy, as part of the working out of one's moral life.

THE PHILOSOPHERS' HOBBES AND THE HISTORIANS' HOBBES

Much work has been done, by Richard Tuck (1993), Noel Malcolm (2002), Quentin Skinner (1996), and others on the historical context of Hobbes's work and its significance in interpreting what Hobbes wrote. Philosophers, perhaps, had read Hobbes with too much concentration on what he took to be his universal science; historians, perhaps, had read him with too much emphasis on his concern with practical matters of the day. The different approaches became explicit with a disagreement between Warrender (1957) and Skinner (1996), but have largely fallen away as each group has come to pay more attention to the work of the other with a number of writers combining both sorts of interest in Hobbes.

See also Aristotle; Bacon, Francis; Boyle, Robert; Galileo Galilei; Game Theory; Gassendi, Pierre; Human Nature; Laws of Nature; Mersenne, Marin; Self-Interest.

Bibliography

Aubrey, John. *Aubrey's Brief Lives*. 2nd ed., edited by Oliver Lawson Dick. London: Secker and Warburg, 1950.

Baumgold, Deborah. *Hobbes's Political Theory*. New York: Cambridge University Press, 1988.

Boonin-Vail, David. *Thomas Hobbes and the Science of Moral Virtue*. New York: Cambridge University Press, 1994.

Chappell, Vere Clairborne, ed. *Essays on Early Modern Philosophers*. Vol. 5, *Thomas Hobbes*. New York: Garland, 1992.

Dietz, Mary G., ed. *Thomas Hobbes and Political Theory*. Lawrence: University Press of Kansas, 1990.

Ewin, R. E. *Virtues and Rights: The Moral Philosophy of Thomas Hobbes*. Boulder, CO: Westview Press, 1991.

Gauthier, David P. *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford, U.K.: Clarendon Press, 1969.

Gert, Bernard. "Hobbes and Psychological Egoism." *Journal of the History of Ideas* 27 (1967): 503–520.

Goldsmith, M. M. *Hobbes's Science of Politics*. New York: Columbia University Press, 1966.

Hampton, Jean. *Hobbes and the Social Contract Tradition*. New York: Cambridge University Press, 1986.

Kavka, Gregory S. *Hobbesian Moral and Political Theory*. Princeton, NJ: Princeton University Press, 1986.

Lloyd, S. A. *Ideals as Interests in Hobbes's Leviathan: The Power of Mind over Matter*. New York: Cambridge University Press, 1992.

Malcolm, Noel. *Aspects of Hobbes*. Oxford, U.K.: Clarendon Press, 2002.

Martinich, Aloysius. *Hobbes: A Biography*. Cambridge, U.K.; New York: Cambridge University Press, 1999.

Martinich, Aloysius P. *The Two Gods of Leviathan: Thomas Hobbes on Religion and Politics* New York: Cambridge University Press, 1992.

Reik, Miriam M. *The Golden Lands of Thomas Hobbes*. Detroit: Wayne State University Press, 1977.

Rogers, G. A. J., and Alan Ryan, eds. *Perspectives on Thomas Hobbes*. Oxford, U.K.: Clarendon Press, 1988.

Rogow, Arnold Rogow. *Thomas Hobbes: Radical in the Service of Reaction*. New York: W. W. Norton, 1986.

Shapin, Steven, and Simon Schaffer. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press, 1985.

Skinner, Quentin. "The Context of Hobbes's Theory of Political Obligation." *Historical Journal* 9 (1966): 286–317.

Skinner, Quentin. *Reason and Rhetoric in the Philosophy of Hobbes*. New York: Cambridge University Press, 1996.

Spragens, Thomas A., Jr. *The Politics of Motion: The World of Thomas Hobbes*. Lexington: University Press of Kentucky, 1973.

Strauss, Leo. *The Political Philosophy of Hobbes: Its Basis and Its Genesis*. Translated by Elsa M. Sinclair. Oxford, U.K.: Clarendon Press, 1936.

Tuck, Richard. *Philosophy and Government, 1572–1651*. New York: Cambridge University Press, 1993.

Warrender, Howard. *The Political Philosophy of Hobbes: His Theory of Obligation*. Oxford, U.K.: Clarendon Press, 1957.

Watkins, John W. N. *Hobbes's System of Ideas: A Study in the Political Significance of Philosophical Theories*. London: Hutchinson, 1965.

R. E. Ewin (1996, 2005)

HOBHOUSE, LEONARD TRELAWNEY

(1864–1929)

Leonard Trelawney Hobhouse, the British sociologist and philosopher, was born in a small village near Liskeard, in Cornwall. He was educated at Marlborough School and Corpus Christi College, Oxford, where he took firsts in classical moderations and "greats." During his undergraduate years he engaged in the study of current problems in politics and economics, along with other radically minded students such as Gilbert Murray and Arthur Acland. He was elected to a prize fellowship at Merton College in 1887 and to a fellowship at Corpus Christi in 1894.

Hobhouse's main interest was the study of the evolution of mind as the central factor in historical development. This, combined with an innate humanitarianism, made him dissatisfied with the passive role of an Oxford don, although even at Oxford he was active in the Labour movement, especially in such causes as trade unionism, the cooperative movement, and adult education. After leaving Oxford, Hobhouse became influential among the "New Liberals," who sought to combine Liberalism with a measure of organized collective action. He was very sympathetic to the Labour Party, although he never joined it. Toward the end of his life Hobhouse grew disillusioned with party politics, and by 1927 he had ceased to belong to any party.

On leaving Oxford in 1897, Hobhouse joined the staff of the *Manchester Guardian,* with which he was associated for most of the rest of his life in one capacity or another. Sociology and philosophy, however, were always his main interests. His *Mind in Evolution* (1901) and *Morals in Evolution* (1906)—a remarkable synthesis of anthropology, ethics, and the history of religious and social institutions—led to his appointment to the new Martin White part-time chair of sociology in the University of London, converted in 1925 to a full-time chair. Hobhouse first opposed Britain's entry into World War I, but he came to support the Allied cause wholeheartedly. He saw the war as the direct outcome of Hegelian teaching, and his own contribution to the war effort was *The Metaphysical Theory of the State,* an extreme attack on Hegelian political theory, especially as found in Bernard Bosanquet's *Philosophical Theory of the State.*

Hobhouse, besides being a philosopher of distinction, made important contributions to anthropological techniques and was a pioneer in comparative and social psychology and one of the founders of sociology as a synthesizing science. The encyclopedic scope of his work and the reluctance of English universities to accept the new subject of sociology contributed to an underestimation of his work in any one field. In philosophy his concern with the reconciliation of different schools meant that he did not himself belong to any one school, and this militated against his due recognition by philosophers.

It is impossible to separate Hobhouse's philosophy from the rest of his work, since his achievement lay in interpreting philosophically a wealth of general and detailed knowledge. There was, however, no question of fitting everything into a fixed scheme. His procedure was empirical and undogmatic, leaving a place for new facts from science and life. His comprehensive studies began with epistemology; went on to an evolutionary interpretation, first of mind in animals and humankind and then of moral and religious ideas; turned next to values in man and society; and ended with a grand synthesis of his philosophical and scientific theories.

The strongest influences on Hobhouse were Herbert Spencer's evolutionary philosophy, Auguste Comte's Positivism, and the social philosophy of John Stuart Mill and T. H. Green. He parted company with Spencer in regarding the appearance of minds as a turning point in the evolutionary process and in accepting the idealists' organic view of society. At the same time he rejected the idealists' reduction of all things to the spiritual. His theory of knowledge was realist and empirical; knowledge cannot make its own object, for it is based on experience and is of reality, not appearance. All knowledge is sociologically conditioned, but a positivist philosophy, applying our knowledge of these conditions, provides safeguards against error. The object of the physical sciences ("matter"), subject to mechanical laws, is only one aspect of reality; there is another aspect ("mind"), subject to teleological laws. Hobhouse traced the close relation of the two aspects in the developing world order.

See also Bosanquet, Bernard; Comte, Auguste; Epistemology; Green, Thomas Hill; Hegelianism; Idealism; Liberalism; Mill, John Stuart; Positivism.

Bibliography

WORKS BY HOBHOUSE

The Theory of Knowledge. London: Methuen, 1896.

Mind in Evolution. London: Macmillan, 1901.

Morals in Evolution. London: Chapman and Hall, 1906.

Social Evolution and Political Theory. New York: Columbia University Press, 1911.

Development and Purpose. London: Macmillan, 1913; rev. ed., 1927. The revised edition is largely rewritten.

The Metaphysical Theory of the State. London: Allen and Unwin, 1918.

The Rational Good. London: Allen and Unwin, 1921.

The Elements of Social Justice. London: Allen and Unwin, 1922.

Social Development. London: Allen and Unwin, 1924.

WORKS ON HOBHOUSE

Barker, Ernest. "Leonard Trelawney Hobhouse." *Proceedings of the British Academy* 14 (1929): 536–554.

Carter, Hugh. *The Social Theories of L. T. Hobhouse.* Chapel Hill: University of North Carolina Press, 1927.

Ginsberg, Morris. "The Contribution of Professor Hobhouse to Philosophy and Sociology." *Economica* 9 (1929): 251–266.

Ginsberg, Morris. "Leonard Trelawney Hobhouse." *Journal of Philosophical Studies* 4 (1929): 442–452.

Hobson, J. A., and Morris Ginsberg. *L. T. Hobhouse, His Life and Work.* London: Allen and Unwin, 1931.

Nicholson, J. A. *Some Aspects of the Philosophy of L. T. Hobhouse.* Urbana: University of Illinois, 1928.

A. K. Stout (1967)

HOCKING, WILLIAM ERNEST
(1873–1966)

William Ernest Hocking, the American idealist and philosopher of religion, was born in Cleveland, Ohio. He spent his early years in the Middle West and studied civil engineering at Iowa State University. Private reading stimulated an interest in philosophy and led him to study at Harvard, where he was influenced chiefly by William James and Josiah Royce. He completed his undergraduate and graduate studies at Harvard University and spent most of his long teaching career there, retiring in 1943.

Although his philosophical system embodies elements of pragmatism and realism, it is primarily an affirmation of Other Mind, or God, as ultimate reality known directly and intuitively. Hocking thus stands in the idealist tradition in modern philosophy and referred to his own position most commonly as "Objective Idealism." Primitive experience, involving the knowledge of other selves and the world, is conditioned by an immediate awareness of Other Mind, standing in an I–Thou relationship to the self. Both sensory and emotive experience have cognitive connections that point beyond self to Other Mind. Hocking's emphasis is on feeling linked inextricably with idea, so that the two are joined in immediate consciousness as an "idea–feeling couple." This concept of the union of idea and feeling is the source of the strong strain of mysticism in Hocking's philosophy, but it is a mysticism that does not abandon the role of intellect in clarifying and correcting intuition. He advances the "principle of alternation" between intuition and intellect as fundamental to the appropriation of metaphysical truth.

In his first book, *The Meaning of God in Human Experience* (1912), Hocking developed an empirical philosophy of religion, grounded in the tradition of classical idealism and at the same time drawing heavily on the mystical experience. In so doing, he sought primarily to defend idealism against arguments of the pragmatists and realists, and he has continued this defense over the years. His Gifford Lectures of 1938–1939 and other later works show a continuing concern with the problem of "meaning in experience," of "fact and destiny," which challenges man to go beyond his day-to-day existence and seek understanding in the wholeness of things. Thus, as a philosopher Hocking dealt primarily with metaphysical and epistemological questions in a manner in which religious sensitivity played a prominent part.

At no point in his long career did Hocking devote himself exclusively to intellectual issues. He played an active role in seeking United States acceptance of the League of Nations and in the 1920s and 1930s he was especially interested in social and political problems of the Middle East. After that time he participated in a study of freedom of the press in the United States and was active in support of the United Nations and other political and ethical causes. These active concerns found expression in at least ten books and scores of articles and extended his influence far beyond the realm of academic philosophy.

See also Idealism; James, William; Royce, Josiah; Other Minds; Pragmatism; Realism.

Bibliography

Hocking's major work is *The Meaning of God* (New Haven, CT: Yale University Press, 1912). Among other works that develop his philosophical and religious views are *Human Nature and Its Remaking* (New Haven, CT: Yale University Press, 1923); *The Self: Its Body and Freedom* (New Haven, CT: Yale University Press, 1928); *Types of Philosophy* (New York: Scribners, 1929; rev. eds., 1939, 1959); *Living Religions*

and a World Faith (New York: Macmillan, 1940); and *Science and the Idea of God* (Chapel Hill: University of North Carolina Press, 1944).

Rouner, LeRoy, ed. *Philosophy, Religion, and the Coming World Civilization: Essays in Honor of William Ernest Hocking.* The Hague: Martinus Nijhoff, 1966.

Rouner, LeRoy. *Within Human Experience: The Philosophy of William Earnest Hocking.* Cambridge, MA: Harvard University Press, 1969.

Richard C. Gilman (1967)
Bibliography updated by Christian B. Miller (2005)

HODGSON, SHADWORTH HOLLOWAY

(1832–1912)

Shadworth Holloway Hodgson, the English metaphysician and epistemologist, was educated at Rugby and Oxford. Although he worked outside the universities, Hodgson was widely respected among English philosophers; he was elected president of the Aristotelian Society at its founding in 1880 and was reelected for thirteen successive years. In the United States, William James recognized the similarity of many of his own doctrines to those of Hodgson, and acknowledged Hodgson's priority despite their profound differences on fundamental points of metaphysics.

Independent and workmanlike, Hodgson was remarkably free from the characteristics and fashions of late Victorian philosophy. He remained steadfast in a central position, attacking the superficial clarities of the associationists on the one side and the vague generalizations of the Germanizing idealists on the other. His primary achievement was to keep alive the firmness of intellectual analysis peculiar to the epoch of Sir William Hamilton and H. L. Mansel. In particular he carried out the line of investigation begun in J. F. Ferrier's *Institutes of Metaphysics.* J. C. Shairp, principal of St. Andrews University and Hodgson's friend and mentor, was his link with Ferrier. Hodgson got from Ferrier a sense of the importance of the relationship of being empirically distinguishable but inseparable, in the way, for example, that color is visually inseparable from shape but nevertheless distinguishable from it. As developed by Hodgson, this principle meant that the notion of logical independence is much more complex than most philosophers have realized. Color, for example, although it is not isolable from shape, does vary independently of shape. From this point of view, Hodgson was able to repudiate the crude logical atomism then prevalent among the associationists without running to the extreme of the sort of logical monism which denies outright the reality of independence.

At a deeper level still, Hodgson applied this same principle of distinguishability with inseparability to elucidate the relationship of consciousness to its objects, that is, of the subjective to the objective. This relationship was basic for Hodgson, and he felt it was disclosed by the kind of reflective analysis that René Descartes used in establishing his *cogito.* Indeed, one might say that Hodgson's starting point was the distinction between this reflective consciousness and a prereflective consciousness in which the distinction between subject and object has not yet emerged.

Although he lacked Ferrier's striking originality, Hodgson was a thinker of great intellectual honesty and thoroughness. What gives his work its special value is the modern manner in which his untiring examination of the fashionable nineteenth-century problems combined technical competence with clarity. The long discussion of G. W. F. Hegel in *The Philosophy of Reflection* is still of interest. So too is Hodgson's treatment of the relationships between particulars and universals and between perception and conception. His careful reconsideration of the problem of free will in *The Metaphysic of Experience* can also be profitably consulted. We are free in the sense that we determine our own actions, but that which does the determining in each case is a set of neurocerebral conditions that is not self-determined, accompanied by consciousness. In this way, he held, free will and determinism are compatible. He explained our awareness of being free as simply our awareness of the uncertainty of the outcome of our acts of volition.

Hodgson held that consciousness gives us knowledge of a reality which is independent of consciousness and which is its condition, even though consciousness is our only evidence for that reality. The material object revealed by consciousness causes sensations in consciousness. It is material, but it is composed of elements that apart from the object would not be material. Consciousness is an epiphenomenon. It is always conditioned by organic and interorganic interaction and never conditions such interactions. The proximate causes of all psychical events lie in the neurocerebral system. There might be immaterial causes of such events, but experience reveals none.

Hodgson resembles Edmund Husserl among later philosophers, rather than Bertrand Russell, G. E. Moore, and their followers. Hodgson's doctrine that things are what they are "known as" anticipated in a way Husserl's phenomenological reduction, and his technique of distinguishing between inseparables approximates to Husserl's reduction to essences. Hodgson's ethics, though

perhaps less interesting than his metaphysics, nevertheless shows the same conscientious struggle to clarify basic distinctions and can be as profitably studied as some other, better-known systems.

See also Analysis, Philosophical; Descartes, René; Ferrier, James Frederick; Hamilton, William; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; James, William; Mansel, Henry Longueville; Metaphysics; Moore, George Edward; Russell, Bertrand Arthur William.

Bibliography

WORKS BY HODGSON

Time and Space. London: Longmans, Green, 1865.

The Theory of Practice. 2 vols. London: Longmans, Green, 1870.

The Philosophy of Reflection. 2 vols. London: Longmans, Green, 1878.

The Metaphysic of Experience. 4 vols. London: Longmans, Green, 1898.

WORKS ON HODGSON

Carr, H. W. "Shadworth Holloway Hodgson." *Mind* n.s., 21 (1912).

Hicks, G. Dawes. "Shadworth Holloway Hodgson." *Proceedings of the British Academy* 6 (1913).

Stout, C. F. "The Philosophy of Mr. Shadworth Hodgson." *PAS* 1 (1892).

George E. Davie (1967)

HØFFDING, HARALD

(1843–1931)

Harald Høffding, the Danish philosopher and historian of philosophy, was born in Copenhagen and lived there throughout his life. From 1883 to 1915 he was professor of philosophy at the University of Copenhagen. Høffding received a degree in divinity in 1865, but he had already decided not to take orders. A study of Søren Kierkegaard's works, and especially of his views on Christianity, had led to an intense religious crisis ending in a radical break with Christianity. Høffding sought in philosophy a new personal orientation and gradually developed into an extraordinarily many-sided liberal humanist. His philosophical development was influenced during a stay in Paris (1868–1869) by the study of French and English positivism, especially that of Auguste Comte and Herbert Spencer. Høffding always worked hard, and his activity as a scholar ranged over every branch of philosophy, including psychology. His works display a vast knowledge, a keen eye for essentials, and a critically balanced judgment. They were translated into many languages and widely used as textbooks. By the turn of the twentieth century Høffding's reputation was worldwide and he knew personally many leading thinkers. He was the outstanding Danish philosopher of his day, and in 1914 the Royal Danish Academy of Sciences and Letters assigned him the honorary residence of Gammel Carlsberg, where he lived to the end of his life. The residence later passed to the physicist Niels Bohr, a younger friend of Høffding.

Of Høffding's many works only five can be discussed here. *Psykologi i Omrids på Grundlag af Erfaring* (Copenhagen, 1881; translated by M. E. Loundes as *Outlines of Psychology,* London) is based on the traditional tripartite division of the mind into knowledge, feeling, and will but puts primary stress on the will in the widest sense of the term. In this sense the will includes conation, urge, endeavor, need, demand, and desire. The will is seen as primary, knowledge as guiding the will, and feeling as a symptom of need or desire, which are themselves elements of the will. Høffding's view anticipated modern need and dynamic psychology.

In *Etik, en Frernstilling af de etiske Principper og deres Anvendelse på de vigtigste Livsforhold* (Ethics: an account of ethical principles and their application to the chief conditions of life; Copenhagen, 1887) Høffding associated himself with British utilitarianism, which he called welfare ethics. The greatest happiness of the greatest number is the fundamental value. In the conflict between individual and social ethics, Høffding took the liberal view. The psychological basis of ethical valuation is a sympathetic feeling that at its highest development takes on the character of a universal and disinterested sympathy.

Den nyere Filosofis Historie, en Fremstilling af Filosofiens Historie fra Renaissancens Slutning til vore Dage (2 vols., Copenhagen, 1894–1895; translated by B. E. Meyer as *History of Modern Philosophy,* 2 vols., London, 1900; reprinted, 2 vols., New York, 1955) is a concentrated account of the various modern philosophers and philosophical schools marked by a fine balance between exposition and criticism. It is of special interest as the first study of modern philosophy to put the primary stress on the mathematico-mechanical science and methods of Galileo Galilei and Isaac Newton in presenting the development of epistemology. Among the philosophers treated, Høffding found Benedict Spinoza, David Hume, and Immanuel Kant especially congenial.

Religionsfilosofi (Copenhagen, 1901; translated by B. E. Meyer as *Philosophy of Religion,* New York, 1906), in three parts, treats religious experience from the standpoints of epistemology, psychology, and ethics. Høffding

claimed that the basis of all religion is a desire for belief in the existence of values, and that the various religions may be characterized by the kinds of values that they claim exist. The presentation is distinguished by its reasoned objectivity and its respect for religion. Høffding himself was an agnostic.

In *Den menneskelige Tanke, dens Former og dens Opgave* (Human thought: its forms and its problems; Copenhagen, 1910) Høffding set forth his theory of knowledge, including an outline for a doctrine of categories whose usefulness has been reduced by the development of modern logic. Høffding's interest in epistemology was psychological rather than strictly logical, and his interest in the psychological basis of knowledge was constructive rather than phenomenological. In general, Høffding followed Hume and Kant in regarding the forms and principles of human knowledge as being peculiar to human beings and their absolute ontological validity as being incapable of proof. The result is a compromise between empiricism and the Kantian critical philosophy.

See also Bohr, Niels; Comte, Auguste; Epistemology; Galileo Galilei; Hume, David; Kant, Immanuel; Kierkegaard, Søren Aabye; Neo-Kantianism; Newton, Isaac; Psychology; Spinoza, Benedict (Baruch) de; Utilitarianism.

Bibliography

There is a brief autobiography containing an account of Høffding's fundamental views in *Die Philosophie der Gegenwart in Selbstdarstellungen,* Vol. IV, edited by Raymund Schmidt (Leipzig: F. Meiner, 1923). Kalle Sandelin, ed., *Harald Høffding in Memoriam* (Copenhagen: Gyldendal, 1932), contains an excellent bibliography that lists 393 publications by Høffding as well as 425 papers and reviews concerning his works.

Frithiof Brandt (1967)

HOLBACH, PAUL-HENRI THIRY, BARON D'
(1723–1789)

Paul-Henri Thiry, Baron d' Holbach, the foremost exponent of atheistic materialism and the most intransigent polemicist against religion in the Enlightenment, was born of honorable but obscure German parents in Edesheim, a small town in the Palatinate; his name was originally Paul Heinrich Dietrich. His upbringing and education were directed by his maternal uncle, Franciscus Adam d'Holbach, who had made a fortune in Paris and assumed French nationality. After studying at the University of Leiden, Holbach came to Paris, in 1749, married his second cousin Basile-Geneviève d'Aine, and soon became a French subject. On his uncle's death in 1753, he inherited the title of Baron d'Holbach, with properties yielding a handsome income of 60,000 livres. The following year his wife died, and in 1756 Holbach married her younger sister, Charlotte Suzanne d'Aine.

On settling in Paris, Holbach had associated with the younger *philosophes* who, with Denis Diderot, Jean Le Rond d'Alembert, and Jean-Jacques Rousseau, were grouping around the *Encyclopédie,* to which he also became a major contributor. His *salon* soon became the main social center, and a sort of intellectual headquarters, for the Encyclopedist movement. The gatherings on Thursdays and Sundays, during more than three decades, at Holbach's house in Rue Royale-Saint-Roch were famous not only for his excellent dinners but also as a unique "clearinghouse" for radical ideas of every type. The more intimate meetings at his country estate of Grandval, near Paris, have been described in fascinating detail in Diderot's letters. The members of Holbach's circle, besides the assiduous Diderot, included Melchior von Grimm, Claude-Adrien Helvétius, d'Alembert, Rousseau, Nicolas-Antoine Boulanger, Étienne Bonnot de Condillac, Jacques-André Naigeon, Baron de l'Aulne Turgot, and Marquis de Condorcet. Holbach also counted among his acquaintances many foreigners, notably David Hume, Edward Gibbon, Adam Smith, Joseph Priestley, Horace Walpole, David Garrick, Laurence Sterne, Cesare Beccaria, and Benjamin Franklin.

Because he left neither a body of correspondence nor personal papers, Holbach's character must be pieced together from contemporary accounts. The composite picture credits him with an impressive erudition, an extremely methodical mind, a collector's interest in art, and with the qualities of affability, discreet generosity, modesty, loyalty to friends, and a taste for virtuous simplicity. Diderot's more private remarks diverge somewhat from this public image, disclosing that the baron, at least with those nearest him, had moments of moodiness, petulance, and gruffness. But these traits just provide a touch of humanity without essentially altering the picture of him as the virtuous atheist. Even Rousseau, despite growing hostility, used him as the model for Monsieur de Wolmar, the altruistic unbeliever of *La nouvelle Héloïse.* Indeed, Holbach's comportment as a social being evidently conformed to his deep desire to illustrate, by his own life and personality, the truth of a most cherished

philosophical opinion, that atheism and morality are as plausibly bound together as religiosity and true virtue are not.

Although Holbach, until some years after his death, was publicly known merely as *le premier maître d'hôtel de la philosophie,* he had surreptitiously played a far greater role, known only to a few. Almost everything he wrote—whether because it expounded atheism and materialism, attacked Christianity, or castigated absolute monarchy, the state church, and feudal privilege—was highly subversive under the *ancien régime* and could have exposed him to the severest penalties. Consequently, his innumerable manuscripts were usually forwarded through secret channels to Holland for publication, after which the books were smuggled back into France. Owing to the strict anonymity that Holbach maintained, bibliographers have since been faced with insoluble problems of exact attribution concerning many texts linked to him.

PHILOSOPHICAL ORIENTATION

Holbach's literary career falls conveniently into three phases. A competent although uncreative student of chemistry, metallurgy, mineralogy, and geology, he translated into French, mainly during the 1750s, a number of works (mostly German) from these fields. He also contributed to the *Encyclopédie,* beginning in 1752, almost 400 articles dealing with the same sciences. These interests shaped Holbach's philosophical outlook, for his materialism corresponded to the methodology and scope of a rigorously scientific explanation of things. In particular, the new evidence offered by geology concerning Earth's history negated, in his view, the doctrine of creation, and with it the existence of God.

The second phase of Holbach's activity, coinciding with the 1760s, consisted of a relentless militancy against organized religion in general and the Catholic church in particular. Not content with the repeated broadsides of his own composition, he also translated anticlerical, deistic, or materialistic works by various British authors (among them Peter Annet, Anthony Collins, Thomas Woolston, John Toland, and Thomas Hobbes), and he published, with the collaboration of Naigeon, a number of French antireligious texts that had long been circulating clandestinely in manuscript copies. Among Holbach's own tracts, the most important were *Le Christianisme dévoilé, ou Examen des principes et des effets de la religion chrétienne* (1761); *Théologie portative, ou Dictionnaire abrégé de la religion chrétienne* (1767); *La contagion sacrée, ou Histoire naturelle de la superstition* (1768); *Lettres à Eugénie, ou Préservatif contre les préjugés* (1768); and *Histoire critique de Jésus-Christ, ou Analyse raisonnée des Évangiles* (1770).

The themes recurring throughout these and similar books represent a vehement restatement of almost all the arguments for unbelief current in eighteenth-century France. The most characteristic are the following: The idea and cult of God sprang from the ignorant terror of primitive man seeking to placate the destructive powers of nature, and they have survived ever since through superstition; religious history is a catalogue of senseless disputes, intolerance, prejudice, persecution, and crime; the clergy is ordinarily engaged in exploiting the gullibility of the people for its own profit; religions have invariably supported tyrannical governments to further their own ambitions of domination; Scriptural "proofs" of Christianity are worthless as objective historical evidence; theological dogmas are a maze of delusion and mystification on which no rational, just, or useful social institution can be built.

ATHEISTIC MATERIALISM

The third and properly philosophical stage of Holbach's output began in 1770 with the *Système de la nature, ou des Lois du monde physique et du monde moral.* This first—and only—example in the Enlightenment of a comprehensive, unmitigated defense of atheistic materialism was the culmination of a whole trend of ideas already expressed in varying degrees by Julien Offray de La Mettrie, Helvétius, Diderot, and others. It caused much consternation in France, not only among spokesmen for the official faith but among the deistic *philosophes* as well. It was suppressed by judicial decree, and among the flood of refutations it provoked were those of Voltaire (the article "God" in the *Philosophical Dictionary*) and Frederick the Great (*Examen critique du Système de la nature*).

The *Système de la nature* defines man as a product entirely of nature, subject to the laws governing the physical universe that, in turn, constitutes the whole of reality. The soul, or spiritual substance, is an illusion; the moral and intellectual attributes of man are simply his organic machine considered in certain of its special, less visible operations. Since sensibility is a primary function of the animal organism, all our higher faculties are derived ultimately from the different forms that sensation takes. The only means of knowing man in nature is through the empirical and rational investigation of matter.

Nature is the sum of matter and motion. All matter is actually or latently in motion, since energy, or force, is among its inherent properties. The material universe is self-created and eternal. All change in nature represents a

communication of motion, a redistribution of energy, which modifies the corresponding combination or disposition of material particles, elements, or aggregates. The totality of matter and motion are eternal and constant, but the specific forms they exhibit—rocks, plants, animals, oceans, heavenly bodies, and so forth—are forever changing. Each thing or being tends, by the laws of attraction and repulsion, to persist in its essence, until it is finally transformed into something else. Man is no exception: The ephemeral life of his species depends on the stability of the physical environment.

There is neither chance nor disorder in nature: All is necessity and order, an irreversible chain of causes and effects. Freedom is objectively meaningless when applied to human behavior, which, controlled by such factors as temperament, education, and environment, takes part in the universal determinism of nature. Virtue and vice, moreover, need not depend on free will; they simply describe actions favoring or hindering the mutual happiness of society and the individual.

ETHICS

Holbach's principal aim was to construct a system of ethical and political values on materialistic grounds. The supreme natural goal of human existence is happiness, but no one can be happy without the services of others. Ethics, therefore, is the science of human cooperation to promote the well-being of the individual through that of society, and it is based on the positive knowledge of men's reciprocal social needs. If humankind has always been morally corrupt and unhappy, religion has been mainly to blame. Supernatural theology, by falsifying man's nature and linking his salvation to the illusory notions of God and immortality, has entirely subverted ethical truth. Holbach takes pains to show that, all attempted definitions of God being hopelessly self-contradictory, "God" is logically a meaningless term. It is understandable, then, that belief in God should have been historically of no moral utility. For religious morality, founded on dogmatic obscurantism and ritualistic futilities rather than on natural and social realities, has prevented human beings from perceiving and correcting the actual conditions productive of vice and misery. Atheism is thus the prerequisite of all valid ethical teaching. In place of the condemnation of sin, Holbach's exposition of secular and utilitarian ethics is typically accompanied by vibrant appeals to humanitarianism and moving exhortations to civic virtue—all in the name of "nature" and "happiness."

POLITICAL THEORY

In *Le bon-sens, ou Idées naturelles opposées aux idées surnaturelles* (1772), the most widely read of his books, Holbach offered a popular, unsystematic version of his philosophy. Thereafter, with the growing troubles of the Bourbon regime, he focused his attention on national problems and developed at great length the ethical and political sections of the *Système de la nature* in a new series of works: *Politique naturelle, ou Discours sur les vrais principes du gouvernement* (1773); *Système social, ou Principes naturels de la morale et de la politique* (1773); *Éthocratie, ou le Gouvernement fondé sur la morale* (1776); and *La morale universelle, ou les Devoirs de l'homme fondés sur sa nature* (1776).

His own term *ethocracy* describes the gist of Holbachian political thought. The state, whose role is simply an extension of the social ethics of enlightened self-love, ought to nurture, in every possible way, the virtues of cooperation on which the good of society and the felicity of each of its members depend. The social pact itself is based on the useful services that the individual and society are able to render to one another, and it remains valid only to the extent that its mutually beneficial aims are fulfilled. Since, therefore, the legitimacy of any government varies directly with the happiness of one and all living under it, Holbach proclaimed with courageous logic the people's right, if there were no other hope of assuring their welfare, to overthrow and replace their rulers. Where the happiness of a society was at stake, it was the sovereign; governments, which were merely means to an end, had no absolute or divine authority.

More specifically, Holbach proposed radical political and economic reforms for France in keeping with the ethocratic ideal. He advocated, as against the extremes of republicanism and enlightened despotism, a limited, constitutional monarchy, in which intermediate parliamentary bodies would represent the interests of society and would maintain a balance between the opposing dangers of either popular or autocratic tyranny. He called for the abolition of hereditary class privileges and for their replacement by a hierarchy of status based on the degree of socially useful service actually rendered by its members. He defended the principle of progressive taxation according to wealth and wanted individual ownership of property to be as proportionate as possible to the value of work performed, thus eliminating the extremes of opulence and poverty. He insisted on the complete separation of church and state and on the toleration of all religious sects, with the government as a neutral preserving peace among them. Freedom of thought and of the press were

to be inviolable; and government had the duty of providing a system of secular public education, with its main objective the inculcation of the social and civic virtues.

SOURCES AND INFLUENCE

Among the sources of Holbach's philosophy were classical and modern Epicureanism, the Cartesian universe of matter and motion in perpetual flux, the logical and metaphysical materialism of Hobbes, the determinism and "atheism" of Benedict de Spinoza, the sensationalism of John Locke, and Leibnizian dynamics. Nearer in time, Holbach was indebted to Helvétius for the utilitarian conception; to La Mettrie for the physiological psychology of the *homme machine*; and to the experimentalist, evolutionary materialism of Diderot, with whom he had the closest personal and ideological ties.

Despite serious shortcomings, Holbach's ideas are still of considerable interest. Although the value of his critique of Christianity is today limited by the one-sidedness and unimaginativeness resulting from his polemical stance and propagandist aims, historically it led toward the objective and psychological study of religion as a distinctly human invention. The *Système de la nature* suffers, no doubt, from too much reliance on outmoded scientific theories; from an excessive generalization and simplification of the concrete complexities of nature; and from a tiresome combination of doctrinaire tone and humorless prolixity that were, unfortunately, peculiar to the author. Nonetheless, it remains a classic text in the development of atheistic materialism as the philosophical expression par excellence of modern science. The main weakness of Holbach's political thought is that it exaggerated a rationalist, moralistic, and prescriptive approach to the subject at the expense of the perhaps more important role of economic, sociological, and historical laws of development on which political institutions, and the changes to be made in them, must depend. Nevertheless, it served significantly to prepare for the French Revolution and contributed subsequently to the progress of democratic and utilitarian doctrines.

See also Alembert, Jean Le Rond d'; Annet, Peter; Atheism; Beccaria, Cesare Bonesana; Collins, Anthony; Condillac, Étienne Bonnot de; Condorcet, Marquis de; Determinism, A Historical Survey; Diderot, Denis; Encyclopédie; Enlightenment; Epicureanism and the Epicurean School; Ethics, History of; Franklin, Benjamin; Gibbon, Edward; Helvétius, Claude-Adrien; Hobbes, Thomas; Hume, David; La Mettrie, Julien Offray de; Locke, John; Materialism; Naigeon, Jacques-André; Priestley, Joseph; Rousseau, Jean-Jacques; Smith, Adam; Spinoza, Benedict (Baruch) de; Toland, John; Turgot, Anne Robert Jacques, Baron de L'Aulne; Voltaire, François-Marie Arouet de; Woolston, Thomas.

Bibliography

WORKS BY HOLBACH

Holbach's works, despite innumerable editions, have never been published in collected form, and few of them have been reprinted since the first half of the nineteenth century. The following are English translations:

Christianity Unveiled; Being an Examination of the Principles and Effects of the Christian Religion. New York: Columbian Press, 1795; London: R. Carlile, 1819.

Common Sense: or Natural Ideas Opposed to Supernatural. New York, 1795, 1833, 1836. Also published as *Superstition in All Ages.* New York, 1878, 1890, 1920; Chicago: De Laurence Scott, 1910.

The System of Nature; or, The Laws of the Moral and Physical World. London, 1795–1796; 1797, 1816, 1820, 1834, 1839, 1840, 1863, 1884; Philadelphia, 1808; New York, 1835; Boston, 1853.

Ecce Homo! or, A Critical Inquiry into the History of Jesus of Nazareth. Edinburgh, 1799; London, 1813, 1823; New York, 1827.

Letters to Eugenia, on the Absurd, Contradictory, and Demoralizing Dogmas and Mysteries of the Christian Religion. London, 1819; New York, 1833; Boston, 1857.

WORKS ON HOLBACH

Charbonnel, Paulette, ed. *Textes choisis; Préface, commentaire et notes.* Paris, 1957.

Cushing, Max Pearson. *Baron d'Holbach; A Study of Eighteenth-Century Radicalism in France.* Ph.D. diss., Columbia University, New York, 1914.

Hubert, René. *D'Holbach et ses amis.* Paris, 1928.

Lange, Friedrich Albert. *The History of Materialism.* Translated by E. C. Thomas. New York: Humanities Press, 1950. Bk. I, Sec. iv, Ch. 3.

Mauthner, Fritz. *Der Atheismus und seine Geschichte im Abendlande,* 4 vols. Stuttgart and Berlin, 1920–1923. Vol. III. Ch. 5.

Naville, Pierre. *Paul Thiry d'Holbach et la philosophie scientifique au XVIII[e] siècle.* Paris, 1943.

Topazio, Virgil. *D'Holbach's Moral Philosophy; Its Background and Development.* Geneva, 1956.

Wickwar, W. H. *Baron d'Holbach; A Prelude to the French Revolution.* London: Allen and Unwin, 1935.

Aram Vartanian (1967)

HÖLDERLIN, JOHANN CHRISTIAN FRIEDRICH
(1770–1843)

Johann Christian Friedrich Hölderlin, a German poet, novelist, philosophical essayist, and dramatist, was born

in Lauffen, Germany. His father died when he was two, leaving Hölderlin an inheritance administered by his mother, who demanded strict obedience to her plans for his future. His mother married Johann Christoph Gok, subsequently the mayor of Nürtingen, in 1774; and a half brother, Karl Gok, with whom Hölderlin maintained a significant correspondence, was born in 1776. His stepfather, whom Hölderlin admired, died in 1779, leaving Hölderlin in his mother's sole charge.

Hölderlin was educated first at the local school in Nürtingen, where he studied Latin, Greek, and rhetoric. He became friends there with Friedrich Wilhelm Joseph von Schelling. He then studied further in Lutheran monastery schools, first at Denkendorf (1784–1786) and then at Maulbronn (1786–1788). During this time he read Ferdinand Canning Scott Schiller, Friedrich Gottlieb Klopstock, and Pindar, and he began composing verses.

Hölderlin entered the Lutheran theological seminary in Tübingen in 1788, at the same time as Georg Wilhelm Friedrich Hegel. Schelling joined the seminary two years later, and Schelling, Hegel, and Hölderlin developed a close friendship. Together, they read Jean-Jacques Rousseau, Benedict (Baruch) de Spinoza, Gottfried Wilhelm Leibniz, Plato, and Immanuel Kant, and they shared enthusiasm for the French Revolution. Throughout his school years Hölderlin displayed intellectual ability, anxiety, emotional intensity, and a readiness to fall in love with intellectually inclined young women. His emotional and intellectual life made him chafe under the regimes and orthodoxies of the seminary, and he found himself pulled more toward poetry than toward a career in the ministry. He published his first poems in 1791, and he began work on the novel *Hyperion, or The Hermit in Greece*.

While continuing to accept the formal control of his future as a minister by the Lutheran consistory, Hölderlin left Tübingen in 1793 to become a private tutor in Waltershausen. From Waltershausen he traveled frequently to Jena in 1794, where he attended Johann Gottlieb Fichte's lectures, met Johann Wolfgang von Goethe, and visited regularly with Schiller, who published "Fragment of Hyperion" in his magazine. In 1795, with Wilhelmine Kirms, a married but separated lady's companion of his employers, he had a daughter, who died of smallpox at thirteen months. Following increasing difficulties in controlling his pupil, Hölderlin was dismissed from Waltershausen in 1795, but was provided with enough money to settle in Jena to study philosophy. There, he lived for a time with Isaac von Sinclair, a close friend and political radical. Most of his strictly philosophical essays date from the 1794–1795 period of his Jena visits and residence.

In January 1796 Hölderlin again became a private tutor, now in the home of Jakob Friedrich Gontard, a wealthy Frankfurt banker. He continued to work in philosophy, and the famous "Oldest System-Program for German Idealism" fragment, arguably by Hölderlin but only later discovered in Hegel's hand and published first in 1918, dates from this period. Here, Hölderlin also encountered the beautiful and talented twenty-seven-year-old Susette Gontard, the wife of his employer, with whom he began a passionate affair. She figures as the model for Diotima in *Hyperion* and as the addressee in some of his finest poems. Volume one of *Hyperion* was published in April 1797. While in Frankfurt, Hölderlin continued to correspond with Schiller, and he imagined a series of "New Letters on the Aesthetic Education of Man," planning both to explain and to overcome all divisions between subject and object and between theoretical and practical reason. Already his friends had begun to worry about his enthusiasms, anxieties, and depressions. Hölderlin completed some fifty-five poems in Frankfurt and began work on the verse drama *Der Tod des Empedokles*.

In September 1798 the affair with Gontard became evident, and Hölderlin was forced to leave Frankfurt for Bad Homburg. He remained in Bad Homburg, except for occasional visits to Nürtingen, until 1800. During this time he continued work on *Hyperion* and *Empedokles*, and he began translations of Pindar and of the tragedies of Sophocles. He produced his poetological essays during this period, as well as many new poems. Volume two of *Hyperion* was published in 1799.

Beginning in January 1801 Hölderlin worked as a private tutor in Hauptwyl, Switzerland. In April he was dismissed, and he returned to Nürtingen. Schiller broke off their correspondence. Throughout the year he completed a number of great poems, including "Bread and Wine," "Homecoming," and "Voice of the People." In December he left on foot to travel to Bordeaux, France, where he arrived in January and remained for three months. In June 1802 he reappeared in Nürtingen, pale, emaciated, and obviously deranged. Hölderlin was able to continue work on the translations of Sophocles (published 1803) and Pindar, as well as on a few poems. In 1804 Sinclair arranged for Hölderlin a position as a librarian in Bad Homburg, without duties. With Sinclair, Hölderlin met in Stuttgart with political radicals conspiring against the landgrave. Sinclair was tried for treason in

1805 but released for lack of evidence. Hölderlin avoided trial by being judged mentally incompetent.

Sinclair gave up his care of Hölderlin in September 1806, and Hölderlin's mother had him forcibly committed to a clinic in Tübingen for the mentally ill. In the summer of 1807 Hölderlin was released into the care of Ernst Zimmer, a Tübingen carpenter who admired his work. "Patmos," "The Rhein," and "Remembrance" were published. He remained in the Zimmer household for the next thirty-six years, where he spent much of his time playing the piano and flute. Wilhelm Waiblinger began to visit Hölderlin in Tübingen in 1822, and in 1830 Waiblinger published *Friedrich Hölderlin's Life, Poetry, and Madness*. Hölderlin's *Selected Poems* was published in 1826. Hölderlin died in June 1843.

It is difficult to locate Hölderlin's work—poetic, philosophical, or poetological—within standard literary and philosophical categories. Dieter Henrich (1992, 1997) established Hölderlin's continuing Kantianism, both in accepting the separation of discursive consciousness from immersion in and intuitive awareness of absolute being and in accepting independence of free, moral personality as an ideal. But Henrich also emphasized Hölderlin's commitment to love and to connectedness to nature and to other human beings. This commitment lends to his writing a sobriety or earnestness, different from the later Romantic irony of Friedrich von Schlegel and from other projects of purely cultural cultivation that are less freighted with ontology. Nor, given his Kantian antidogmatism, does Hölderlin offer any system of human life in relation to the absolute, in the manner of the absolute idealisms of Hegel or Schelling. In Henrich's terms Hölderlin is best characterized as articulating a *Vereinigungsphilosophie*: an account of human beings as always seeking both independence-moral sublimity and love-connectedness. In this continual seeking, moments of remembrance and of gratitude for one's course of life are possible, but without any lasting conclusiveness.

Hölderlin's poetry—while typically firmly metrically controlled by Greek models, especially ones taken from Pindar, and so is more classical than effusive—is also characteristically difficult syntactically, even hermetic. Argument over Hölderlin's significance has concerned whether Hölderlin is better understood as a confident prophet of an imminent transcendence of one's present cultural plights, as Martin Heidegger (1949a) urges, or rather principally as a paratactic writer, resistant to all formally closed plots of human experience, as Theodor W. Adorno (1992) urges. Here, Henrich's reading of Hölderlin's *Vereinigungsphilosophie* has the advantage of accepting the insights but avoiding the errors of these other, sharply opposed readings.

Hölderlin's sense of the continuing openness, but also provisional formability, of philosophico-poetic thinking is reflected in his *Wechseltonlehre* or theory of the proper modulation of fundamental moods, in poetry and in life, and this sense is enacted in his poetic practice. Together, his theory and poetic practice provide an image of nonfoundationalist seriousness in thinking that is likely to continue to attract substantial attention and interest.

See also Adorno, Theodor; Fichte, Johann Gottlieb; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hermeticism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Neo-Kantianism; Plato; Rousseau, Jean-Jacques; Schiller, Ferdinand Canning Scott; Schlegel, Friedrich von; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY HÖLDERLIN

Sämtliche Werke. 8 vols., edited by Friedrich Beissner. Stuttgart, Germany: J. G. Cottasche Buchhandlung Nachfolger, 1943–1985.

Werke und Briefe. 2 vols., edited by Friedrich Beissner and Jochen Schmidt. Frankfurt am Main, Germany: Insel-Verlag, 1969.

Sämtliche Werke: Frankfurter Ausgabe. 18 vols., edited by D. E. Sattler. Frankfurt am Main, Germany: Verlag Roter Stern, 1975–2000.

Essays and Letters on Theory. Translated and edited by Thomas Pfau. Albany: SUNY Press, 1988.

Hyperion and Selected Poems, edited by Eric L. Santner. New York: Continuum, 1990.

WORKS ABOUT HÖLDERLIN

Adorno, Theodor W. "Parataxis: On Hölderlin's Late Poetry." In *Notes to Literature*, 2:109–149. Translated by Shierry Weber Nicholsen; edited by Rolf Tiedemann. New York: Columbia University Press, 1992.

Constantine, David. *Hölderlin*. Oxford, U.K.: Clarendon Press, 1988.

Donelan, James H. "Hölderlin's Poetic Self-Consciousness." *Philosophy and Literature* 26 (1) (April 2002): 125–142.

Förster, Eckart. "'To Lend Wings to Physics Once Again': Hölderlin and the 'Oldest System Program of German Idealism.'" *European Journal of Philosophy* 3 (2) (August 1995): 174–198.

Hayden-Roy, Priscilla A. *A Foretaste of Heaven: Friedrich Hölderlin in the Context of Württemberg Pietism*. Amsterdam, Netherlands: Rodopi, 1994.

Heidegger, Martin. "Hölderlin and the Essence of Poetry." In *Existence and Being*, 270–291. Chicago: Henry Regnery, 1949.

Heidegger, Martin. "… Poetically Man Dwells …" In *Poetry, Language, Thought*, 211–229. Translated by Albert Hofstadter. New York: Harper and Row, 1971.

Heidegger, Martin. "Remembrance of the Poet." In *Existence and Being*, 233–269. Chicago: Henry Regnery, 1949.

Henrich, Dieter. *The Course of Remembrance and Other Essays on Hölderlin*, edited by Eckart Förster. Stanford, CA: Stanford University Press, 1997.

Henrich, Dieter. *Der Grund im Bewusssein: Untersuchungen zu Hölderlins Denken, 1794–1795*. Stuttgart, Germany: Klett-Cotta, 1992.

Ryan, Lawrence J. *Hölderlins Lehre vom Wechsel der Töne*. Stuttgart, Germany: W. Kohlhammer, 1960.

Santner, Eric L. "Introduction: Reading Hölderlin in the Age of Difference." In *Hyperion and Selected Poems*, by Friedrich Hölderlin, xxxii–xl, edited by Eric L. Santner. New York: Continuum, 1990.

Waibel, Violetta L. *Hölderlin und Fichte: 1794–1800*. Paderborn, Germany: Schöningh, 2000.

Richard Eldridge (2005)

HOLE ARGUMENT

The original "hole argument" (*lochbetrachtung*) was created by Albert Einstein. The point of the argument may be put as follows: If a physical theory's equations are *generally covariant* (that is, invariant under a wide group of continuous coordinate transformations) then the theory is in a certain specific sense *indeterministic*. Einstein put the argument to two different uses. First before the discovery of his final field equations for the General Theory of Relativity (GTR), the argument was put forward as a justification for accepting *non*-generally covariant field equations, namely those of the 1913 Einstein-Grossman *Entwurf* theory. Einstein was not fully satisfied with that theory, in part because he believed that general covariance was necessary if a theory were to capture a fully general relativity of motion, and so the hole argument served to help Einstein reconcile himself (temporarily and only partially) to the *Entwurf* theory. The second use of the hole argument came in 1915 when Einstein came to see the argument, taken in its first form, as a mistake. From his second point of view the argument rests on a mistaken interpretation of the mathematics of general covariance. The indeterminism allegedly shown by the hole argument is spurious, and the argument cuts no ice in favor of any particular theory or interpretation of the nature of space-time.

Seven decades later, after the rediscovery of Einstein's argument by John Stachel and John Norton, history repeated itself. A close cousin of Einstein's hole argument was put forth by John Earman and John Norton (1987) as an argument claiming to show that, if one embraces a *substantival* view of space-time, then in a generally covariant theory such as the GTR, one is committed to an unpleasant form of indeterminism. Earman and Norton argued that the problem is reason enough to justify rejecting a substantival view of space-time in GTR. But within a few years this view of the argument's significance was widely rejected. Instead most philosophers came to think that the hole argument's indeterminism is merely an artifact of a particular interpretation of the mathematical structure of GTR that we are not logically compelled to accept.

Regardless of which viewpoint is better supported, it is indisputable that Earman and Norton's hole argument led to a huge resurgence of interest in the interpretation of space-time in GTR, and lies at the core of much of the philosophy of space-time theories published since 1987. Subsequently philosophers have explored the status of general covariance, and therefore of the hole argument, in the domain of quantum gravity theories.

THE 1987 HOLE ARGUMENT

GTR describes the dynamical interaction of material substances in space-time with other material substances, as well as their interactions with the variably-curved structure of space-time itself. Einstein's field equations describe these interactions, and delimit the set of models, or physically possible worlds, corresponding to the theory.

A model of GTR is usually presented as a triple $<M, \mathbf{g}, \mathbf{T}>$ consisting of a four-dimensional, continuously differentiable manifold M, a metric-field tensor $\mathbf{g}$ (representing the geometry of space-time) defined everywhere on the manifold, and a stress-energy tensor $\mathbf{T}$ representing the material substances in space-time. Like $\mathbf{g}$, $\mathbf{T}$ is defined everywhere in the space-time, but unlike $\mathbf{g}$, $\mathbf{T}$ may be exactly equal to zero at some or even all points of space-time. (In the latter case we say the space-time is "empty," but it may still have an interesting structure as encoded in $\mathbf{g}$.) Notice that each of these objects is *four*-dimensional, representing not just how things are at a specific time but rather how things are over the entire history of the (model-) universe.

The manifold is a collection of points with a local and global topology built-in. For example some models of GTR have M structurally identical to $\Re^4$, which means that space-time can be coordinatized (all the points labeled) with four-dimensional Cartesian coordinates. The metric tensor defines the metric and geometric structure of the space-time: distances between points A

and *B*, whether points *A*, *B* and *C* are collinear, whether line *L* is a straight line (geodesic) or curved, and so on. Note that *M* by itself does not have such geometric structure; there are no distances between points in *M* alone, no straight lines, and so forth. Finally **T** represents the matter, the energy-momentum, existing in space-time.

Physicists and philosophers confront a set of interpretational issues regarding GTR and its model worlds, and one prominent issue is this: Should space-time be thought of as an object existing in its own right, that is as a *substantial entity*? To answer this question in the affirmative is to take GTR as vindicating space-time substantivalism, a close cousin of Newton's absolutism. But what exactly is space-time according to GTR? Earman and Norton (1987) argued that the manifold *M*, by itself, is what deserves the name of substantival space-time in GTR:

> We take all the geometric structure, such as the metric and derivative operator, as fields determined by partial differential equations. Thus we look upon the bare manifold—the "container" of these fields—as space-time …
>
> The advent of general relativity has made most compelling the identification of the bare manifold with space-time. For in that theory geometric structures, such as the metric tensor, are clearly physical fields in space-time. The metric tensor now incorporates the gravitational field and thus, like other physical fields, carries energy and momentum whose density is represented by the gravitational field stress-energy pseudo-tensor …
>
> If we do not classify such energy-bearing structures as the [gravitational] wave as contained within space-time, then we do not see how we can consistently divide between container and contained.
>
> (PP. 518–551)

If space-time substantivalism is understood as the thesis that (a) the manifold *M* by itself represents space-time and (b) its points are substantial entities themselves, then the ground is prepared for the hole argument. The general covariance of the Einstein field equations, interpreted in an active sense, allows one to take a given model $M_1 = \langle M, \mathbf{g}, \mathbf{T} \rangle$ and construct a second via an automorphism *h* on the manifold. The automorphism maps points of *M* to other points of *M* in a smooth fashion. The effect of this re-arranging of the points is the production of a new model: $M_2 = \langle M, h^*\mathbf{g}, h^*\mathbf{T} \rangle$ which also satisfies the field equations, and in which the "contents" of space-time, **g** and **T**, have been "slid around" on the manifold. The kind of automorphism employed in the hole argument is usually called a "hole diffeomorphism." Think of M_2 as obtained from M_1 by sliding the metric and matter fields around on the point-manifold in the region of *M* called "the Hole," leaving everything unchanged elsewhere. (Equivalently one can think of the hole diffeomorphism as a kind of shifting-around of the manifold points, moving the points around "underneath" the metric and material contents of the space-time.)

If M_2 and M_1 agree or match for all events before a certain time *t*, but differ for some events afterward (inside the Hole), then we have a form of indeterminism, at least on the most straightforward way of defining determinism in the context of GTR. Relative to our chosen substantial entities, space-time points considered as the elements of *M*, we can say: In GTR, what happens at what space-time locations is radically underdetermined. Earman and Norton (1987) presented this indeterminism as an argument against the kind of substantivalism (*manifold* substantivalism) they see as most natural in GTR:

> Our argument does not stem from a conviction that determinism is or ought to be true. There are many ways in which determinism can and may in fact fail … Rather, our point is this. If a metaphysics, which forces all our theories to be deterministic, is unacceptable, then equally a metaphysics, which automatically decides in favor of indeterminism, is also unacceptable. Determinism may fail, but if it fails it should fail for reasons of physics, not because of a commitment to substantival properties which can be eradicated without affecting the empirical consequences of the theory.
>
> (P. 524).

Substantivalism about space-time is thus, according to this argument, ruled out as an acceptable interpretive option for GTR. Before we consider responses to the hole argument, we need to note three points. First this indeterminism is unobservable: M_1 and M_2 are qualitatively indistinguishable. Second, a tacit assumption of the hole argument is that the identities of the manifold points may be taken as given or specified, in some sense, independently of the material/ observable processes occurring in space-time (represented by **g** and **T**). In fact one way of thinking of a hole automorphism is as a (continuous) *permutation* of the points underlying physical processes, or (equivalently?) as a *re-labeling* of the points. Third, a manifold is a collection of space*time* points, not space points. In other words the points do not have duration;

FIGURE 1

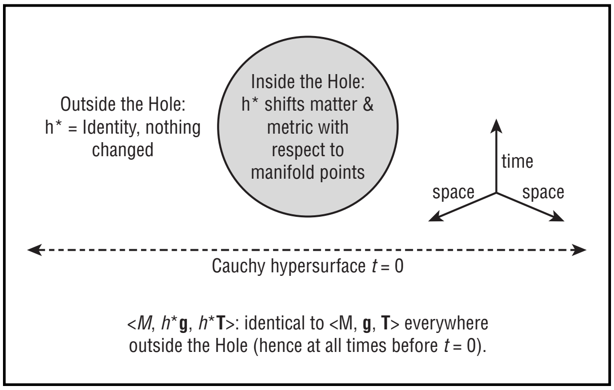

each one is an ideal point-event, a representative of a spatial location at a single instant of time. They do not exist *over* time and hence serve as a structure against which motion may be defined, as Newton's space points did. In light of the second point just above, the indeterminism at issue is not a failure of the determination of future events at pre-existing spatial locations, but rather a failure of the mathematics to specify which individual points would *pop into and out of existence* underneath specified physical events.

Not surprisingly most responses to Earman and Norton's hole argument have departed from these three points, arguing either that the indeterminism is innocuous, or that substantivalism can be reinterpreted in ways that do not lead to the apparent indeterminism.

RESPONSES TO THE HOLE ARGUMENT

Two authors, Cartwright and Hoefer, have responded to the hole argument by denying that it has any *prima facie* force at all. Their response attacks the logic of Earman and Norton's reasoning. Since the indeterminism is both unobservable and peculiarly metaphysical (involving as it does only questions of *which points*, considered as identity-bearing individuals, will underlie *which physical events*, it is not properly speaking a *physical* indeterminism at all and hence not something that ought to be ascribed physical/ ontological importance. Most other authors however have not questioned the logical force of Earman and Norton's argument, agreeing with them that determinism must "be given a fighting chance" (Earman 1989, p. 180). But most authors have also rejected the hole argument's anti-substantivalist conclusion. They argue either for a different understanding of substantivalism, a different definition of determinism, or both.

The first to respond to Earman and Norton's argument were Tim Maudlin (1988) and Jeremy Butterfield (1989). Both accepted the *prima facie* reasonableness of the hole argument but then argued that a metaphysical mistake was nevertheless being committed in the course of the hole argument. For Butterfield, the mistake lay in (a) taking the identities of manifold points between models as an unproblematic given and hence (b) defining indeterminism in too direct and unsubtle a fashion. Butterfield argued that we should avail ourselves of something like David Lewis's apparatus of counterpart theory in order to decide which points in a given model are identical with which points in a different model and accordingly revise the definition of determinism in terms of counterpart relationships. The technical details are too complicated to present here, but the upshot is that GTR turns out not to be indeterministic after all once both point trans-world identity and determinism are properly understood.

Maudlin rejected Earman and Norton's claim that the manifold by itself represents space-time. Instead he argued that the manifold plus metric is what represents space-time, and moreover that we should consider the spatio-temporal, geometric properties ascribed to points of *M* by **g** to be *essential properties* in a strong metaphysical sense. In support of the former point, Maudlin (1988) and Hoefer (1996) adduce the following points:

> 1) A manifold by itself has few of the paradigmatic spatiotemporal properties we would expect space-time to have: Distance relations between points, collinearity on a straight line, and so forth. In fact there is not always even a distinction to be found between space-like directions and time-like directions! So it is odd to think of *M* alone as representing space-time.
>
> 2) There is an easy way to separate between space-time (the "container") and its contents (the "contained"): it is the distinction between *M* + **g** and **T**. Mathematically the distinction is clear. Moreover, as was true for classical substances in the Newtonian tradition, **T** can vanish at some, or even all, space-time locations. **g** cannot vanish anywhere, in any genuine part of a GTR model space-time.
>
> 3) If it is accepted that **g** can carry genuine (stress-) energy content, that only makes it even more substantial than Newtonian space-time's structure was; it is hardly a reason for considering **g** not to be part of the characterization of space-time in GTR.

4)"[I]f the metric is classified as a physical field *in* space-time, rather than as representing part of space-time itself, the following odd situation emerges. Space-time itself is not appealed to in explaining the motions of material things; they are explained by relations to a different kind of physical field. Even distances and other geometric relations have nothing to do with space-time, but instead with the relations between two kinds of physical fields in space-time. When substantivalism starts to sound like relationism, something is wrong!" (Hoefer 1996, p. 13).

Most authors now seem to agree that Earman and Norton's identification of *M* as the sole representor of space-time is questionable. This alone does not block their hole argument, though it points the way toward various different versions of substantivalism, incorporating the metric as part-representor of spacetime, which may avoid the hole argument's indeterminism.

Maudlin's essentialism about the metrical properties and relations of space-time points blocks the hole argument by making the metrical properties of individual points be (metaphysically) essential properties: Thus, if model $M_1 = \langle M, \mathbf{g}, \mathbf{T} \rangle$ represents a genuine physically possible world, then model $M_2 = \langle M, h^*\mathbf{g}, h^*\mathbf{T} \rangle$ cannot in general do so, since it ascribes metaphysically impossible properties to the points of *M*. Thus, properly interpreted, GTR does not allow a determinism-violating plethora of indistinguishable space-times.

In only slightly different ways, Maidens (1993), Stachel (1993) and Hoefer (1996) diagnose the hole argument as resting on an interpretive mistake: The mistake of considering models such as M_1 and M_2 as representing (meta-) physically distinct space-times. As noted above, the differences between such models concern only *which substantial individuals* (manifold points) underlie *which material happenings and relations*. Without exaggeration one can put the distinction like this: While M_1 says that the point Larry underlies my fingertip at this moment, M_2 says that the (qualitatively identical, in all respects) point Fred does so instead. Earman and Norton's interpretation of substantivalism therefore ascribes *primitive identity* to the points of space-time, and models such as M_1 and M_2 differ only in what philosophers call *haecceitistic* ways, that is, in which properties are ascribed to which individuals, where the individuals are mere "bare particulars".

General relativists routinely deny the significance of such alleged differences, and say that diffeomorphic models like M_1 and M_2 represent just *one* physically possible world (thereby advocating *Leibniz Equivalence*). We should do the same, urge these authors; when we do, the hole argument evaporates, and we are nevertheless left with a strong form of substantivalism, one that takes M + **g** to represent space-time and considers any two diffeomorphic mathematical models as representing one and the same physically possible world. The disadvantage of taking this interpretive route is that one loses the ability to describe certain metaphysical possibilities that were accepted by Newton and Samuel Clarke, that is, the possibility that every event in the world's history could have taken place five meters to the East of its actual location. Some philosophers maintain that these metaphysical possibilities are an essential part of any substantivalist view.

Not all those inspired by the hole argument to work on spacetime issues try to shore up substantivalism. The hole argument inspired those with relationist leanings to revive the idea, advocated by Reichenbach earlier in the twentieth century but effectively killed by Earman (1989) and Friedman (1983), that GTR can be interpreted as fully compatible with relationism. Teller (1991), Huggett (1999), and Saunders (2003) are examples of this approach. What makes this position possible is the adoption of a liberal attitude toward the idea of *relations* between material things. If the manifold is viewed as only representing the continuity, dimensionality, and topology of spacetime (as some substantivalists would agree anyway), then what is really indispensable is the metric. Can it be interpreted relationally? Those philosophers who argue that it can are not espousing a Machian reduction of metrical structure to material relations. Instead they claim that the metric itself can be interpreted as merely giving the structure of actual *and possible* spatiotemporal relations between material things. **g** is not a thing or substance. Where matter is present, it is crucial to the definition of local standards of acceleration and non-acceleration; the Einstein field equations record just this relationship. In many ways the desires of traditional relationists (especially Leibniz, Huygens, and Mach) are—arguably—met by GTR when interpreted this way.

FURTHER DEVELOPMENTS

By the late 1990s a broad consensus was reached among philosophers of spacetime that there are acceptable interpretations of spacetime in the GTR context that do not run afoul of the hole argument indeterminism (both substantival and relational interpretations). In a series of papers, however, John Earman and Gordon Belot (2001) argued that consideration of the extension of GTR to a quantum theory vindicates the importance of the hole argument, and reveals the vacuity of certain philosophical

responses to it. The issues and arguments involved in this new broaching of the hole argument are complicated and technical; only a cursory review can be attempted here.

One approach to quantizing GTR begins by recasting the theory in the Hamiltonian formalism, wherein a three-dimensional configuration representing the state of the physical world evolves in accordance with the Hamiltonian equations of motion. This is a natural way of formulating GTR preparatory to attempting to quantize it, since there are established recipes for quantizing theories starting in the Hamiltonian framework. But the (active) general covariance of GTR makes for a resulting indeterminism in the Hamiltonian presentation of the theory, just as it did for the standard theory when interpreted as a theory about what happens at individual manifold points. Various ways of dealing with this indeterminism can be linked conceptually to respective philosophical responses to the hole argument, and they appear to lead to genuinely different *theories* after quantization is done. So Earman and Belot (2001) claim:

> There is a correspondence between interpretations of the general covariance of general relativity and approaches to—and interpretations of—quantum gravity … One demands that one's interpretation of general relativity should underwrite an approach to quantization which leads to a viable theory of quantum gravity, and that one's understanding of quantum gravity should lead to a way of viewing general relativity as an appropriate classical limit.
>
> (P. 249)

Different responses to the hole argument make for different interpretations of space-time in GTR, and according to Earman and Belot these correspond to different approaches to quantum gravity. The "relationist" response of Teller, Saunders, and others which rejects the idea that diffeomorphic models are distinct physical possibilities, corresponds to the "gauge invariant" approach to quantum gravity. But the quite similar interpretation of classical GTR that is offered by Maidens, Stachel, Hoefer, and others is pejoratively labeled "sophisticated substantivalism" by Earman and Belot, and they find it the one view unworthy of even entering the playing field of the interpretive game. They claim that these philosophers are obliged to produce a gauge invariant mathematical treatment of classical GTR in the Hamiltonian framework, which may or may not be possible:

> [W]e maintain that there is one sort of response to the hole argument which *is* clearly undesirable: the sort of sophisticated substantivalism which mimics relationalism's denial of the Leibniz-Clarke counterfactuals. It would require considerable ingenuity to construct an (intrinsic) gauge-invariant [LE-based] substantivalist interpretation of general relativity. And if one were to accomplish this, one's reward would be to occupy a conceptual space already occupied by relationalism. Meanwhile, one would forgo the most exciting aspect of substantivalism: its link to approaches to quantum gravity.
>
> (2001, P. 248)

Earman and Belot evidently still characterize substantivalism as essentially a matter of believing in space-time points as individuals with primitive identity (a key conceptual part of manifold substantivalism, as we saw above), and relationism as the denial of such points-as-primitives. Most philosophers would reject both viewpoints, in line with points 1–4 above that argue for including the metric field in our characterization of space-time. It is also unclear why the sophisticated substantivalist faces a technical difficulty of constructing a "gauge-invariant" interpretation of Hamiltonian GTR, unless the relationist that Earman and Belot cites approvingly does too—given that they occupy the same conceptual space.

Earman categorizes the loop quantum gravity approach of Rovelli and others as lined up with relationism. A crucial aspect of such a gauge invariant approach should be its evasion of hole argument-style indeterminism problems. But Rickles (2005) claims that a perfect analog of the hole argument can be constructed within the framework of loop quantum gravity. Rickles argues that indeterminism-via-surplus-structure can infect either relationist or substantivalist interpretations of GTR, whether in classical or quantized form, and therefore that the two issues should henceforth be kept apart.

As for the original hole argument itself, we should note that it was an argument that *used* general covariance to argue against a certain ontological view (manifold substantivalism), in the context of classical GTR. Regardless of what view of general covariance and determinism/indeterminism issues is eventually vindicated in the realm of quantum GTR (if one view is—there is no guarantee this will happen), it will not alter the dialectic of the hole argument itself, or the philosophical issue of whether GTR as a self-standing theory does or does not give us a picture of space-time deserving the label "substantival." Compare with the absolute/ relational debate in the context of classical Newtonian mechanics. Greater mathematical rigor and conceptual clarity in the founda-

tions of Newtonian mechanics did have a bearing on that philosophical debate, and led Earman (1989) and Michael Friedman (1983) to declare a hands-down victory for absolute space. But the nature and status of space-time in General Relativity has not been taken to be relevant to that earlier debate (or the correlate debate about space-time in Special Relativity), even though both of these earlier theories are "appropriate classical limits" of GTR. The moral would appear to be that philosophers should tackle interpretive issues one theory at a time. If and when a successful quantum gravity theory emerges, the substantival/ relational debate can be addressed anew in light of its particular mathematical structure.

See also Clarke, Samuel; Determinism, A Historical Survey; Earman, John; Einstein, Albert; Leibniz, Gottfried Wilhelm; Logic, History of; Mach, Ernst; Newton, Isaac; Reichenbach, Hans; Relativity Theory; Space; Time.

Bibliography

Belot, Gordon, and John Earman. "Pre-Socratic Quantum Gravity." In *Physics Meets Philosophy at the Planck Scale*, edited by Craig Callender and Nick Huggett. Cambridge, U.K.: Cambridge University Press, 2001.

Butterfield, Jeremy. "The Hole Truth." *British Journal for the Philosophy of Science* 40 (1989): 1–28.

Cartwright, Nancy, Carl Hoefer, and John Earman. "Who's Afraid of Absolute Space?" *Australasian Journal of Philosophy* 48 (1970): 297–319.

Earman, John. *World Enough and Space-Time*. Cambridge, MA: MIT Press, 1989.

Earman, John, and John Norton. "What Price Space-Time Substantivalism? The Hole Story." *British Journal for the Philosophy of Science* 38 (1987): 515–525.

Einstein, Albert, and Marcle Grossman. *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*. Teubner, 1913; CP4, doc. 13.

Friedman, Michael. *Foundations of Space-time Theories*. Princeton, NJ: Princeton University Press, 1983.

Hoefer, Carl. "Einstein's Struggle for a Machian Gravitation Theory." *Studies in History and Philosophy of Science* 25 (1994): 287–335.

Hoefer, Carl. "The Metaphysics of Space-time Substantivalism." *Journal of Philosophy* 93 (1996): 5–27.

Huggett, Nick. "Why Manifold Substantivalism Is Probably Not a Consequence of Classical Mechanics." *International Studies in the Philosophy of Science* 13 (1999): 17–34.

Maidens, Anna. "The Hole Argument: Substantivalism and Determinism in General Relativity." PhD diss. University of Cambridge, 1993.

Maudlin, Tim. "The Essence of Space-time." In *Proceedings of the Philosophy of Science Association 1988.* Vol. II, edited by Mickey Forbes and Arthur Fine, 82–91.

Norton, John. "How Einstein Found His Field Equations: 1912–1915." *Studies in History and Philosophy of Science* 14 (1984): 253–316.

Rickles, Dean. "A New Spin on the Hole Argument." *Studies In History and Philosophy of Science. Part B: Studies in History and Philosophy of Modern Physics* 35 (2005): 1–26.

Saunders, Simon. "Indiscernibles, General Covariance and Other Symmetries: The Case for Non-Reductive Relationalism." In *Revisiting the Foundations of Relativistic Physics: Festschrift in Honour of John Stachel*, edited by Abhay Ashtekar et al. Dordrecht: Kluwer, 2003.

Stachel, John. "The Meaning of General Covariance." In *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, edited by John Earman et al. Pittsburgh: University of Pittsburgh Press, 1993.

Teller, Paul. "Substance, Relations, and Arguments about the Nature of Space-Time." *Philosophical Review* (1991): 363–397.

Carl Hoefer (2005)

HOLISM AND INDIVIDUALISM IN HISTORY AND SOCIAL SCIENCE

In most recent philosophical discussion, the contrast between holism and individualism in history and the social sciences has been presented as a methodological issue. Stated generally, the question is whether we should treat large-scale social events and conditions as mere aggregates or configurations of the actions, attitudes, relations, and circumstances of the individual men and women who participated in, enjoyed, or suffered them. Methodological individualists say we should. Methodological holists (or collectivists, as some prefer to be called) claim, rather, that social phenomena may be studied at their own autonomous, macroscopic level of analysis. Social "wholes," they say, not their human elements, are the true historical individuals.

This issue obviously bears directly upon the way we are to conceive the relations between such social sciences as psychology and sociology, and between these and historical inquiry. But it is commonly thought also to involve us in wide-ranging metaphysical problems—those of historicism and organicism, for example—and to have grave ethical and political implications as well. Sir Isaiah Berlin, in *Historical Inevitability* (Oxford, 1954), moves quickly from methodological to metaphysical issues when he represents holists as believing in "invisible powers and dominions," conceived as "impersonal enti-

ties at once patterns and realities, in terms of which … men and institutions must behave as they do." And May Brodbeck, in "Methodological Individualisms: Definition and Reduction," expresses a common opinion when she writes: "Culturally, holism is intimately connected with hostility toward the liberal political individualism of the Western tradition." Individualists, in their turn, have been castigated by their opponents for encouraging laissez-faire in economics and anarchy in politics, the alleged natural consequences of adopting an "atomistic" view of social life. Indeed, the threat of appropriate social consequences seems to have been regarded by some as a reason for accepting one or the other of these methodological positions. F. A. Hayek and K. R. Popper are well-known champions of the principle of methodological individualism as a bulwark against the supposed horrors of the "planned society"—or at any rate, against anything worse than "piecemeal social engineering."

It is not, in fact, entirely accurate to say that the methodological, metaphysical, and political doctrines have invariably gone together. Thomas Hobbes, for example, was in effect a methodological individualist who advocated something close to political absolutism; and Maurice Mandelbaum, as will appear below, is a contemporary methodological holist who would certainly repudiate "invisible powers" and "impersonal entities." But political or ethical argument has, in any case, a dubious place in an examination of holism and individualism as methodological prescriptions for social and historical research. Even if metaphysical questions cannot ultimately be ignored, it is worthwhile, at least at the outset, to try to consider the contending methodological doctrines in their own terms. The discussion that follows makes no attempt to trace the considerable history of the problem in Western philosophy; rather, it is a report on what some contemporary philosophers have said by way of exposition and defense of the two positions. Since it has generally been the individualists who have taken the initiative in controversy, it will be convenient to set forth their position first.

METHODOLOGICAL INDIVIDUALISM

J. W. N. Watkins, one of the most prominent recent advocates of methodological individualism, has presented it as primarily a theory of sociological or historical explanation. In his "Ideal Types and Historical Explanation," Watkins stated its requirements thus: "Social processes and events should be explained by being deduced from (*a*) principles governing the behaviour of the participating individuals and (*b*) descriptions of their situations." The elaboration of criteria for acceptable explanation is, of course, an activity characteristic of philosophers. What has most often concerned them, however, has been the formal or structural features of explanation, that is, the logical relation that must hold between an explanans and explanandum. Watkins's criterion, by contrast, is a material one. It makes a stipulation about the content of a social or historical explanans, holding that it must be "psychological," at least in the sense of being, in Watkins's words, about "the situations, dispositions and beliefs of individuals."

In formulating their material requirement, individualists often have in mind successful patterns of explanation in other branches of science. According to Watkins, the principle of methodological individualism is a correlate of the principle of mechanism in physics, which held triumphant sway from the seventeenth to the nineteenth centuries. An especially prestigious example of the application of the mechanistic principle is the explanation of the solar system by reference to Isaac Newton's laws and the positions, masses, and momenta of its component "individuals." Another example, often cited, is the explanation of the macro properties of a gas—its temperature, for example—as a resultant of the micro properties of its molecules. The best illustration of the same explanatory procedure in social science is afforded by classical economics, which regards macro states of the market as resultants of the dispositions and consequent activities of individual producers and consumers. There are differences (some will be discussed later) between the way particles in a mechanistic system are linked with what they explain and the way psychological facts about individuals are linked with social events. Methodological individualists, however, regard the likenesses as more instructive than the differences.

METHODOLOGICAL HOLISM

The rival thesis of methodological holism is that explanations in history and social science may (some would say "must") employ holistic societal laws or dispositions. Social dispositions are envisaged as being holistic, not only in the sense of being macroscopic relative to individual behavior but as being irreducibly so. Except in extreme versions of the theory (usually framed by opponents for polemical purposes), psychological elements are not actually excluded from a social explanans; they are merely regarded as insufficient. Thus, in their most usual form, the two methodological doctrines are not contraries but contradictories.

In elaborating their position, holists often match paradigm cases with the individualists. In economics, for example, they point to the Keynesian theory, which relates such variables as national income and savings, as showing the need to supplement the classical approach with a macroscopic one. In physics they note the decline of mechanism with the development of wave and field notions. And methodological holists do not limit their claims to cases in which social phenomena are explained by other societal factors. The explanation of individual actions themselves, they insist, may often have to be given partly in societal terms, employing laws that link individual behavior with types of social conditions. They deny, however, that this commits them either to organicism or to historicism. For sui generis societal laws can be of various logical types. They need not be organic, in the sense of relating the parts of the social system in a way that makes society self-regulating or self-maintaining, nor need they be developmental. There is thus no necessary connection between methodological holism and the dismal conclusion that men are caught up in some inexorable process that possesses something like a life of its own.

REFINEMENTS OF INDIVIDUALISM

The basic response of methodological holism to the individualist claim is that the procedures of history and social science are in fact largely holistic, and that attempts to apply the principle of individualism do not work. The theory of the social sciences should accept the consequences. To methodological individualists, on the other hand, failures of application simply indicate a need for further analysis and research. Yet the discrepancy between fact and theory has induced individualists to make a few concessions, which are often represented as "refinements" or "clarifications" of the original thesis. A brief look at four of these may help to sharpen the issue.

LEVELS OF EXPLANATION. Individualists generally concede, first, that macro explanations may sometimes be both true and informative. The temperature of a gas, for example, may be explained by referring to a heat source that was applied to it, or to such simultaneous macro conditions as its volume and pressure; the outbreak of a revolution may be similarly explained by referring to economic or social trends in the society as a whole. According to Watkins, all the methodological individualist claims is that until we manage to reduce such explanations to terms of the molecular theory of gases or the psychology of individuals, we fail to achieve a full understanding of what has occurred. Thus, what the individualist seems to offer is not a criterion of being an explanation at all (for this, the satisfaction of formal criteria may be enough), but of being an ultimately satisfactory one. Yet the acceptability of "half-way explanations" (to use Watkins's term) is said to depend on the possibility of eventually reducing them to "rock-bottom explanations." The concession, in other words, is only with regard to "practice"; nothing is yielded at the level of "principle."

ANONYMOUS INDIVIDUALS. A second refinement arises out of the suspicion that what is actually possible in social science, even "in principle," is seldom an explanation in terms of the dispositions of the specific individuals involved. We might explain the rise in a stock's value, for example, by pointing out that the individual dispositions that most stockholders may be presumed to share lead them to be willing to pay a higher price under the circumstances; but we could hardly hope to ground our conclusion in knowledge of the detailed motives and beliefs of all the individuals involved. Methodological individualists consequently limit their prescription, even for "rock-bottom explanations," to typical dispositions of anonymous individuals. Such explanations, they will point out, still follow the model of mechanistic physics, in which information about specific particles is not required. Unlike physical particles, it cannot, of course, be presumed that human beings are all alike, or even that they are similar in all respects relevant to the social resultant that is being studied. This is particularly the case in historical inquiry, with its concern for unique rather than recurring circumstances and events. Thus, it will often be impossible to give adequate historical explanations without taking at least some named individuals into account. Even in the field of history, however, there is considerable scope for the anonymous.

UNINTENDED RESULTS. Advocates of individualism often emphasize that if explanation need not be in terms of the actions and dispositions of specific human beings, still less need it show that social phenomena are brought about deliberately, or even knowingly, by individuals. Methodological individualists do not question the contention, constantly reiterated by holists, that social phenomena are largely the *unintended* results of the behavior of hosts of interacting human beings. The individualist principle is thus to be distinguished from what K. R. Popper, in *The Open Society and Its Enemies* (London, 1945), has called the "conspiracy theory of society": the view that for every social effect there is a manipulator (hero or villain) to be found. Not that individualists doubt that public affairs are controllable through the knowledgeable

intervention of people; they hold, rather, that even when events are not so controlled, they can be explained individualistically. The individualist principle is also to be distinguished from a second doctrine with which Watkins felt it is sometimes confused, namely, the view that social phenomena "reflect" the dispositions of component individuals. Social characteristics are often, in fact, quite different qualitatively from the characteristics of the individuals referred to in explaining them. But there is likewise no qualitative similarity between the thermodynamic properties of a gas and the mechanical properties of its elements.

EXCEPTIONS. Some individualists are willing to make a concession that leads to still a fourth refinement of the original doctrine. They allow that there are some social phenomena, at least, that may not be open to individualistic explanation at all, although they usually add that these exceptions are not very important for the theory of the social sciences—certainly not important enough to justify the acceptance of methodological holism as a general principle for explanation in these fields. The exceptions fall into two classes. The first contains phenomena that can be treated only statistically. The second consists of occasional instances of what may be genuinely organic "social" behavior: Watkins mentioned the physical union of mating couples, the ecstatic singing of revival meetings, the rioting of panicking crowds. But, individualists argue, we cannot extrapolate from such cases to the nature of "higher-grade" forms of social organization. The latter, even when unplanned, are related by "ideas" and involve people widely separated in space and time.

ARGUMENTS FOR METHODOLOGICAL INDIVIDUALISM

Clarified and refined, then, the principle of methodological individualism asserts that *ultimate or final* explanation of the more *significant* social phenomena must be given in terms of at least *typical* dispositions (including beliefs, attitudes, and volitions) of *anonymous* individuals involved. Individualists often seem to present this principle as self-evident. Yet arguments for it have been offered, among the most characteristic in contemporary literature being the five that are considered below. No separate presentation of the case for methodological holism will be given, since holists are generally content to offer rebuttals of what their opponents claim.

METAPHYSICAL ARGUMENTS. One common argument appeals directly to ontological considerations. According to Watkins, "the ontological basis of methodological individualism is the assumption that society … really consists only of people." Social "things" may even be said to be "created" by individuals, by their attitudes as well as by their actions. "Remove the attitudes of food officials, shop-keepers, housewives, etc., towards ration books," Watkins observed, "and they shrivel into bits of cardboard." To a methodological individualist it seems paradoxical to suggest that social objects, thus constituted, could be explained other than individualistically. To try to explain individual actions in social terms seems to involve referring what really exists to a mere "construction." Yet although ontological individualism offers the methodological doctrine a "basis," Watkins conceded that the former does not actually entail the latter. It might still be true that what is constituted by individual actions and attitudes is governed by autonomous social law, although the ontology of individualism makes this difficult to believe.

Today, few holists would argue directly from a corresponding ontological thesis, which would rest upon some such principle as "a whole is not equal to the sum of its parts," the social whole thus being conceived as free to operate in accordance with laws which hold true at its own "level of existence." Typical of objections to this are Ernest Nagel's observation, in *The Structure of Science,* that wholes are recognized in physical science, too, apparently without presenting special problems for individualistic explanation; and Popper's jibe, in *The Poverty of Historicism,* that the metaphysical principle of holism, although "trivially true," applies even to three apples on a plate. However, most methodological holists (for example, Maurice Mandelbaum in "Societal Facts") prefer to argue that although social phenomena can be said to be ontologically dependent upon the actions and attitudes of individuals, the two are not simply identical. They point out, too, that their doctrine does not commit them to claiming that societies could exist without people, this being an absurdity eschewed even by full-blooded ontological holists like G. W. F. Hegel. The frequent use, in this connection, of the epiphenomenalist account of the mind-brain relation to show what might be meant by ontological dependence without identity is rather unfortunate. For, whereas a mind with no brain may be conceivable, few, if any, methodological holists would allow that society was conceivable without individuals. Many methodological holists, in fact, profess complete ontological individualism. What they demand of individualists is a willingness to try to *find out* whether there are any irreducible societal laws.

This takes us within range of a second metaphysical argument. According to Watkins, it is a "metaphysical commonplace that social events are brought about by people." He interpreted this "commonplace" to imply that individual men and women "together with their material resources" are the "only moving agents," indeed the "only causal factors," in history. Social wholes, whether or not they can be said really to exist, cannot *do* anything; in particular, they cannot affect the behavior of the concrete human beings who constitute them. Methodological individualists therefore disagree with economists who regard long-term cyclical waves in economic activity as, in Watkins's words, "self-propelling, uncontrollable and inexplicable in terms of human activities." They similarly oppose historical materialism, which, in its more uncompromising forms, at any rate, asserts a one-way causal relation between certain social conditions (the economic substructure) and the thoughts and actions of those who live under them. It is a "central assumption of the individualist position," Watkins declared, that "no social tendency is somehow imposed on human beings 'from above' (or 'from below')." Actually, even the more modest (and more usual) thesis of "interaction" between the social and the individual spheres is often deemed unacceptable by individualists.

The usual response of the holists to this line of argument is to ridicule the implied denial of social conditioning—as if people were not born into social situations in the first place. The "real oddity" of methodological individualism, wrote Ernest Gellner in "Explanations in History," is that "it seems to preclude *a priori* the possibility of human dispositions being the dependent variable in an historical explanation—when in fact this is what they often or always are." An associated peculiarity is that it precludes "the possibility of causes … being a complex fact which is not describable in terms of the characteristics of its constituent parts alone—which again seems often to be the case." Individualists, of course, would regard this charge as a misunderstanding of their doctrine. They would hold that the social conditioning of individuals, although real, is simply their conditioning by other individuals, referred to compendiously by holistic terms. And they would accept this claim that causes may be complex facts as long as the complexity of the cause is regarded as "resultant" from individual actions in the way indicated by the ontological argument. (Some individualists, however, would find it less easy to counter the argument that to speak of causes as "moving agents" at all is tacitly to accept an "activity" view of causation that has been suspect since David Hume.)

Alan Donagan, in *The Later Philosophy of R. G. Collingwood,* provided a version of the individualist's causal argument that turns on a conception of human action made familiar by idealist philosophers. The only way men's actions can be explained, Donagan maintained, is through their "thoughts"; it is not men's actual situations which explain what they do, but their conception of the situations (although it may be necessary to refer to the actual situation in explaining a man's success or failure in translating his intentions into action). Thus, if physical causes, like climate, operate in history, they must operate indirectly; and the same is true of such social events and conditions as an economic depression or a military victory. Unless we are to challenge the common assumption that the causal relation is transitive, however, methodological holists may well feel that such considerations, even if they are acceptable in themselves, do little to establish Watkins's original contention. For to say that social causes require the mediation of individual thoughts and responses is not to establish the latter as the only "moving forces" in history. On the contrary, to cause individuals to cause is still to cause.

EPISTEMOLOGICAL ARGUMENTS. The theory of action thus indicated has a bearing on a third general argument that is sometimes used by methodological individualists. This argument develops Watkins's contention that even if we learned to describe, predict, and control social events and conditions holistically, we still could not properly claim to understand them without treating them as a collection of individual responses. For "understanding," Watkins seemed to insist, requires the explanation of what happened in terms of *intelligible* human dispositions. What he appears to have had in mind is the discerning of the participants' reasons for doing what they did, which allows us the intellectual satisfaction of seeing why they thought their responses were appropriate. As Gellner has pointed out, there is a dual thesis here: first, that social or historical explanation must be couched in terms of the dispositions of individual human beings; second, that these dispositions must be of a special kind. For those who would claim, on general philosophical grounds, that explanation by reference to an agent's reason or motive is logically different from subsuming an occurrence under a law (or even under a "disposition" properly so called), the present claim opens up the possibility of giving individualistic explanations of social phenomena without reference even to psychological laws.

Many methodological holists would agree that to accept the additional thesis would make their position quite untenable, for it might be claimed that "intelligible"

dispositions could be sought at the level of social wholes only on the assumption either of an immanent group mind or of an external historical providence: in other words, methodological holism now *would* require ontological holism. Something just short of this nevertheless sometimes appears to be entertained. Thus Morris Ginsberg, in *On the Diversity of Morals* (London, 1956), while denying for ontological reasons that society is itself a mind, conceded that it has a "mental organization" or "inner side" that is not identical with the mentality of any of its component individuals. Most methodological holists, however, simply deny the necessity of the additional criterion of explanation. They would hold that the essential claim of methodological individualism could be achieved without reference to intelligible dispositions if appropriate psychological laws could be found. And they would similarly claim for their own position that subsumption under autonomous societal laws (if such laws could be found) would yield understanding in the only sense significant to "scientific" inquiry.

A fourth argument makes the even more basic epistemological claim that whereas we can observe human individuals, we cannot similarly obtain knowledge of the macro features of social groups. As Watkins put it: "The social scientist and historian have no 'direct access' to the overall structure and behaviour of a system of interacting individuals (in the sense that a chemist does have 'direct access' to such overall properties of a gas as its volume and pressure and temperature, which he can measure and relate without any knowledge of gas molecules)." Since reliable knowledge of the dispositions and situations of individuals is readily available, Watkins continued, and since these individuals constitute the group, "a theoretical understanding of an abstract social structure should be derived from more empirical beliefs about concrete individuals." How else can what is said about social wholes be verified? Such an appeal to "hardheaded" empiricism is a shrewd blow. For the contenders here are (with a caveat, perhaps, about history) two theories of "scientific" inquiry. It would be odd if they did not both claim to be "empiricist."

Faced with this argument, many methodological holists insist that *some* social phenomena, at least—for example, parades, trials at law, battles—are directly observable. It is true that no one will notice such things if he lacks certain interpretative ideas or concepts. But this is not a peculiarity of social observation. Individual human actions themselves will not be "observed" unless we are able to discern the intentions and motives of the actors; and it may be questioned whether these can be known "directly" in a sense in which group phenomena cannot. The epistemological criterion of the individualists, in other words, either allows some social phenomena to be counted as observable or excludes the most interesting individual phenomena. Many holists nevertheless concede that the social wholes of most significance for history and social science cannot, in any ordinary sense, be directly observed. They reject instead the implication that this puts them at odds with accepted procedures in natural science. Not all physical theorizing proceeds, as in chemistry, from observables to what explains them: astronomy, for example, "constructs" its wholes as surely as sociology does. They admit, too, that assertions about societies must be verified by discovery of what individuals do. But they deny that this undermines their claim to possess knowledge, not just of the individuals but of the social wholes of which they are elements.

LINGUISTIC ARGUMENT. The fifth argument for methodological individualism, although it obviously has some affinity with both ontological and epistemological ones, is presented as a point of logic or semantics. As L. J. Goldstein stated it, in his "The Two Theses of Methodological Individualism," individualists require, as a condition of their being meaningful, "that all of the concepts used in social science theory be exhaustively analyzable in terms of the interests, activities, volitions and so forth of individual human beings." If this condition were met, the apparent holism of explanations employing societal laws would be tolerable, because it would be eliminable "by translation." Watkins himself denied that this conceptual thesis actually belongs to the central position of methodological individualism. The latter, he maintained, is a theory of explanation, not of concept-formation or description. Yet in arguing for the explanatory thesis, he characteristically slipped into the conceptual one. He maintained, for example, that to an individualist, the statement "The Jewish race is cohesive," if it is to be empirically meaningful, must mean such things as "Jews usually marry Jews"—a statement about anonymous individuals. And he commended Max Weber for insisting that the only way to make the meaning of social terms precise is to *define* them individualistically—as if such concepts appear holistic only when they remain vague or undefined.

Methodological holists have denied both that such analysis, definition, or translation is possible and that the conceptual thesis, even if sound, would establish the explanatory one. In arguing for the first of these positions in "Methodological Individualisms: Definition and Reduction," Brodbeck allowed that there are no insuper-

able problems for conceptual individualism so long as we are dealing with group concepts that are basically statistical—as in "He got his votes from the poor." Less straightforwardly statistical locutions like "Boom follows slump" might present problems for individualistic translation only because their implied statistical reference is vague. The real difficulty arises over such terms as "renaissance" or "the government." There seems to be no finite list of individual actions and attitudes that would count as their exemplifications; yet the problem does not seem to be one simply of vagueness. Appropriate exemplifications, furthermore, seem to vary from culture to culture, without our being able to say that the relevant terms are ambiguous—which suggests, perhaps, an evaluative element in their meanings. Terms within this range of social description appear to be logically holistic. As Mandelbaum has observed, the problem here is analogous to one which phenomenalists have paid great attention to in contemporary theory of perception. The full meaning of a material object statement, it is generally agreed, cannot be given by any finite set of sensation reports alone; we always need reference to "conditions of appearance," which are stated in the material object language. Attempts to translate societal statements into psychological terms founder on the similar need to specify the social conditions under which an action must be performed in order for it to count as an exemplification.

Mandelbaum himself actually wavered on this point. He conceded that partial translatability, at least, is required; otherwise there would be no way of verifying societal statements (an echo of the epistemological argument). He was sufficiently moved, too, by metaphysical considerations (both ontological and causal) to believe that full translation may be possible "in principle," even if this cannot be made the basis for a "practical" methodology. Many methodological holists have claimed, however, that full translatability would still not warrant the acceptance of the individualist thesis as it is most commonly understood, namely, that explanations should be (or should be capable of being) limited to psychological terms, with psychological laws as the only permissible kind of connecting generalizations.

SOME OBSCURITIES

Consideration of the claim that full translatability would not warrant acceptance of the individualist thesis calls attention to three important obscurities that have dogged much contemporary discussion. The first concerns the sense of "explain" in which a methodological individualist asserts that ultimately all explanations must be individualistic. The conceptual claim, it should be noted, has been stated not only as an argument for the explanatory one but also as though it were itself a theory of explanation—and so it is, in one important sense of the term. According to Watkins, every complex social situation or event is "the result of a particular configuration of individuals, their dispositions, beliefs, and physical resources and environments." This is often, and plausibly, read as meaning that we must be able to explain large-scale social phenomena as configurations or resultants of individual ones. But to explain something *as* something else is to explicate its nature: It is to explain it in the sense of showing *what it really is.* Although this kind of explanation is common in history and social science, however, it does not seem to be what methodological individualists have usually had in mind when advancing their explanatory thesis. What they envisage is explanation in the sense of showing how or why something *came to be* what it is: explanation that goes on to give causes, for example. It remains to be seen whether the conceptual thesis has any bearing on individualist claims about such explanations.

It may be objected that this ambiguity underlying the notion of being a "resultant" need not destroy the support given by the conceptual thesis to the full explanatory thesis, since what is specified in the productive sense may itself be treated, in its turn, as a resultant in the constitutive sense, thus achieving full "reduction." But this directs attention to a second obscurity in the individualist thesis, the question whether a "why" or "how" explanation in which all societal terms were replaced by psychological ones would involve the reduction of societal to psychological *laws.* Holists such as Nagel and Brodbeck have contended that it would not, at any rate, achieve reduction in the sense most familiar to the philosophy of science. For the derivability of macro laws from micro laws, no matter what the field of inquiry, is at least partly an empirical matter. Even in such exemplary cases as the reduction of chemistry to physics, they have pointed out, composition laws, which specify the way individual behavior changes as groups increase in size, must be added to the ordinary laws of the micro discipline; and these, however "self-evident" they often seem, have an empirical status. If laws of individual psychology are to be related "reductively" to laws of group phenomena, empirical composition laws would similarly have to be found. The reduction could never be just a matter of definition.

Individualists may complain that this involves too restricted an interpretation of their demand that sociological and historical explanations be reduced to "psychological" terms. Thus, when Mandelbaum, a

methodological holist, attacked the conceptual thesis on the assumption that no statement will be counted as psychological if it employs any societal term, his argument was rebutted by Donagan, an individualist, on the ground that hardly any human dispositions would be psychological on such a criterion. Watkins made it clear that, for him, a psychological disposition is simply one which specifies a way of acting and thinking that is open to an individual. Thus, he argued against an anthropological holist that marriage rules are widespread dispositions of anonymous individuals in a society to behave in a certain way, and against a historical holist that the "Calvinistic outlook" of seventeenth-century Huguenot traders was similarly individualistic. Behind the uncertainty about what is to count as "psychological," there in fact appear to lie two different interpretations of the conceptual thesis itself. The first, which imposes a limitation upon the *ways of behaving* that may be cited in a "rock-bottom" explanans, is the translatability thesis. The second, to which many methodological individualists appear to retreat under pressure, is the much weaker demand that an acceptable explanation employ concepts which can be *attributed* to an individual, or jointly to a group of them.

Willingness to move in the latter direction suggests that contemporary methodological individualists and holists are not really as far apart as they often seem. In spite of their insistence that what they put forward is a *methodological* doctrine which is merely supported by metaphysical considerations, it seems clear that what interests methodological individualists most is the related ontological claim that human beings are the "ultimate constituents" of the social world. By contrast, although methodological holists find themselves continually under fire for allegedly flirting with dangerous metaphysical notions, it appears that what they are most concerned to uphold is the *logical* respectability of using holistic collective concepts and macroscopic laws, if need be. As was indicated, many methodological holists protest their allegiance to ontological individualism, and this appears to be a perfectly coherent claim. Some would even accord the corresponding methodological principle of individualism "regulative status" as formulating an ideal to be striven for. What they resist is the conclusion a priori that we can realize the ideal, and the associated temptation to refuse anything less.

See also Berlin, Isaiah; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Hume, David; Nagel, Ernest; Newton, Isaac; Philosophy of Social Sciences; Popper, Karl Raimund; Scientific Method.

Bibliography

Brodbeck, May. "Methodological Individualisms: Definition and Reduction." *Philosophy of Science* 25 (1) (1958): 1–22.

Danto, A. C. "Methodological Individualism and Methodological Socialism." *Filosofia* 13 (1) (1962): 3–24.

Donagan, Alan. *The Later Philosophy of R. G. Collingwood.* Oxford: Clarendon Press, 1962.

Gellner, Ernest. "Explanations in History." *PAS,* Supp., 30 (1956): 157–176. Reprinted under the title "Holism versus Individualism in History and Sociology," in *Theories of History,* edited by Patrick Gardiner, 489–503. Glencoe, IL: Free Press, 1959.

Goldstein, Leon J. "The Two Theses of Methodological Individualism." *British Journal for the Philosophy of Science* 9 (33) (1958): 1–11.

Hayek, F. A. *The Counter-Revolution of Science.* Glencoe, IL: Free Press, 1952.

Mandelbaum, Maurice. "Societal Facts." *British Journal of Sociology* 6 (4) (1955): 305–317. Reprinted in *Theories of History,* edited by Patrick Gardiner, 476–488. Glencoe, IL: Free Press, 1959.

Mandelbaum, Maurice. "Societal Laws." *British Journal for the Philosophy of Science* 8 (31) (1957): 211–224. Has full bibliographical reference to much of the current literature.

Nagel, Ernest. *The Structure of Science,* 336–397, 536–546. New York: Harcourt Brace, 1961. Like Donagan, he treats the problem of holism versus individualism in the context of a broader discussion of science.

Popper, K. R. *The Poverty of Historicism.* London: Routledge, 1957.

Watkins, J. W. N. "Historical Explanation in the Social Sciences." *British Journal for the Philosophy of Science* 8 (30) (1957): 104–117. Reprinted in *Theories of History,* edited by Patrick Gardiner, 503–514. Glencoe, IL: Free Press, 1959.

Watkins, J. W. N. "Ideal Types and Historical Explanations." In *Readings in the Philosophy of Science,* edited by Herbert Feigl and May Brodbeck, 723–743. New York, 1953.

W. H. Dray (1967)

HOLISM AND INDIVIDUALISM IN HISTORY AND SOCIAL SCIENCE [ADDENDUM]

The current philosophical discussion on holism and individualism can be considered on the basis of the notions of supervenience and intrinsic properties. A prominent conception of supervenience is David Lewis's thesis of Humean supervenience:

> It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. … We have geometry: a system of external relations of spatio-temporal distance between points. Maybe

> points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that. (1986, pp. ix–x)

This is a thesis of global supervenience, applying to the world as a whole. To put the matter in the formulation that tends to be preferred in today's discussion, "Any world which is a *minimal* physical duplicate of our world is a duplicate *simpliciter* of our world" (Jackson 1998, p. 12). Global supervenience thus conceived is not in dispute in the philosophy of social science. No serious holist denies that if you duplicate the whole domain of physical properties that are instantiated in the world, you thereby also duplicate all the social properties that are instantiated in the world.

The dispute is about what exactly has to be included in the supervenience base. According to Lewis (1986), apart from spatiotemporal relations, all the relations that are instantiated in the world supervene on the intrinsic properties of the individuals. If this idea is applied to social science, it is to say: Given the spatiotemporal positions of all the individual persons in the world (their worldlines), the intrinsic properties of all the individual persons in the world fix all the social and historical facts and institutions. The dispositions of individuals count among their intrinsic properties. While individualism implies this thesis, holism disputes it, holding that there are more nonsupervenient relations than the spatiotemporal ones. As regards social science, holists maintain that social relations do not supervene on the intrinsic properties of individual persons and their spatiotemporal arrangement.

The claim about intrinsic properties in Lewis's (1986) thesis of global supervenience is already disputed in fundamental physics. Quantum systems admit what is known as entangled states. Quantum entanglement can be taken to consist in certain relations among quantum systems; there are no intrinsic properties whatsoever that could constitute a supervenience basis for these relations. Quantum entanglement can therefore be considered as indicating the ultimate failure of the explanation paradigm of individualism, which is seen as being tied to classical, atomistic physics (Teller 1986).

The most serious challenge to individualism in the social sciences stems from the rule-following considerations as put forward by Saul A. Kripke (1982) in his interpretation of Ludwig Wittgenstein's *Philosophical Investigations* (1953/1993). According to Kripke, Wittgenstein maintains that the conceptual content of the belief states of people does not supervene on their intrinsic properties—neither mental intrinsic properties such as mental ideas, nor physical intrinsic properties such as dispositions to behavior. Following Kripke's Wittgenstein, any mental idea and any disposition to behavior is finite and therefore insufficient to fix a precise conceptual content, which implies determining an indefinite number of cases. Furthermore, it is thereby insufficient to determine what is the correct application of a rule (concept) in contrast to its incorrect application. The problem of rule-following is an important challenge to individualism in the social sciences, because it calls one central presupposition of any individualistic position in question, namely that the content of our belief states is ontologically independent of social interactions.

The most prominent individualist reply to the problem of rule-following is to develop a sophisticated dispositionalism to overcome the objections from the finitude and the non-normativity of dispositions. The main versions make use of computationalism (Miscevic 1996) or teleosemantics (Millikan 1990), or a combination of both. However, according to teleosemantics the content of a belief state is determined by a biological function, and biological functions depend on the history of the organism in a given environment. Consequently, biological functions are not intrinsic properties so that, according to this view, the content of the belief states of a person does not supervene on intrinsic properties of the person either.

The holist reply to the problem of rule-following takes social relations to be the decisive factor in the determination of the conceptual content of our belief states. It thus implies social holism: Having beliefs with a determinate conceptual content depends on social interactions. The basic idea is that social interactions (social practices) put at the disposal of people a distinction between what they take to be correct and what is correct in the light of others. On this basis social interactions drive, notably via sanctions, a process that determines a conceptual content for the belief states of the people who participate in them.

There are two versions of this position: The more radical one is the skeptical solution to the problem of rule-following that Kripke (1982) himself attributes to Wittgenstein (1953/1993) and according to which there is no standard of correctness beyond communal agreement (see also Kusch 2002). The more moderate position takes

social practices to fix conceptual content, but maintains that there are truth conditions for our beliefs beyond communal agreement (Brandom 1994; see also Esfeld 2001, chapters 3, 5).

The debate that is initiated by the rule-following considerations enables us to cast the discussion between individualism and holism in the social sciences in the following framework: The point at issue is to what extent social facts and institutions depend on social relations that do not supervene on the intrinsic properties of the individuals that stand in these relations (cf. Pettit 1993, part 2). To the extent that one rejects such nonsupervenient social relations, one subscribes to individualism. A social holist goes as far as maintaining that even the conceptual content of the belief states of the individual persons is fixed by such social relations.

See also Extrinsic and Intrinsic Properties; Kripke, Saul; Lewis, David; Rule Following; Supervenience; Wittgenstein, Ludwig Josef Johann.

Bibliography

Brandom, Robert B. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press, 1994.

Esfeld, Michael. *Holism in Philosophy of Mind and Philosophy of Physics*. Dordrecht, Netherlands: Kluwer Academic, 2001.

Jackson, Frank. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford, U.K.: Clarendon Press, 1998.

Kripke, Saul A. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge, MA: Harvard University Press, 1982.

Kusch, Martin. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford, U.K.: Clarendon Press, 2002.

Lewis, David. *Philosophical Papers*. Vol. 2. New York: Oxford University Press, 1986.

Millikan, Ruth Garrett. “Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox.” *Philosophical Review* 99 (3) (1990): 323–353.

Miscevic, Nenad. “Computationalism and Rule Following.” *Proceedings of the Aristotelian Society* 96 (1996): 215–229.

Pettit, Philip. *The Common Mind: An Essay on Psychology, Society, and Politics*. New York: Oxford University Press, 1993.

Teller, Paul. “Relational Holism and Quantum Mechanics.” *British Journal for the Philosophy of Science* 37 (1986): 71–81.

Wittgenstein, Ludwig. *Philosophical Investigations* (1953). Translated by G. E. M. Anscombe. Malden, MA: Blackwell, 1993.

Michael Esfeld (2005)

HOLKOT, ROBERT

(d. 1349)

Robert Holkot [Holcot] was the most significant Dominican theologian of the fourteenth and fifteenth centuries. He received his doctorate at Oxford, lecturing on Peter Lombard’s *Sentences*, the main theology textbook, in the years 1331–1333, and served as regent master there, most likely from 1336–1338. He spent time in London as a clerk for Richard of Bury, the bishop of Durham, and probably lectured on the biblical Book of Wisdom at Cambridge from 1340–1342. From 1343 to his death from the plague in 1349, he resided at the Dominican priory in Northampton.

The Condemnations of 1277 and the arguments of John Duns Scotus at the turn of the fourteenth century established the view that no absolute necessity governs creation: God has always had the power to do other than he does and to create a reality other than this one. The working out of the implications for philosophy and theology of such a contingent reality framed scholarly debate during Holkot’s time. The tools available to attack the problem had also undergone major changes. In the generation prior to Holkot, William Ockham had subjected thirteenth-century Aristotelianism to a severe critique. Holkot adopted Ockham’s philosophy as his starting point.

The most important and controversial of Holkot’s views involve his use of the distinction between God’s absolute and ordained power. Omnipotence means that God has the absolute power to do whatever does not involve a contradiction. Because no necessity attaches to the ethical precepts that govern the created order (God could without contradiction have created a world in which merit would accrue to doing the opposite of each of the Ten Commandments), human salvation depends upon a covenant between God and human beings established under the New Law of Christ. God’s ordained system, the system that instantiates one or another of the many possible creatable orders, displays his expressed power, but could have been, or in the future could still be, other than it is.

Because, in Holkot’s view, divine goodness owes nothing to creation, there would be no contradiction in God’s replacing the current order with another, even without fulfilling the promises or covenants integral to the current ordination. The principle of noncontradiction provides the ultimate security. If God were to change the ordained system, he would either inform people of the new conditions for salvation or not. If God did not,

then no one could be held accountable for the new conditions. It would involve a contradiction for God to hold people to account for what they can not know.

To analyze the contingent theological order, Holkot adapted the rules of "obligational" debate, a form of debate in which an "opponent" usually proposed some contingent possibility as the initial starting point, to be held true during the debate, and a "respondent" would admit or exclude further proposed propositions as they were consistent with or contradicted the initial proposition. For Holkot, God's revelations functioned like the initial proposals in such debates, and it was incumbent on the believer to hold them as true and to accept the consequences of supposing them true, all the while knowing that the contingent order of creation might mean they were false and never have been true. Holkot's development of this "obligational" theology was his most distinctive contribution.

See also Aristotelianism; Duns Scotus, John; Ockhamism; Peter Lombard; William of Ockham.

Bibliography

Gelber, Hester Goodenough. *It Could Have Been Otherwise: Contingency and Necessity in Dominican Theology at Oxford, 1300–1350*. Leiden: Brill, 2004.

Hoffmann, Fritz. *Die theologische Methode des Oxforder Dominicanerlehrers Robert Holcot*. Bieträge zur Geschichte der Philosophie und Theologie des Mittelalters: Texte und Untersuchungen, n.s. no. 5. Münster: Aschendorff, 1972.

Kennedy, Leonard A. *The Philosophy of Robert Holcot, Fourteenth-Century Skeptic. Studies in the History of Philosophy*, no. 27. Lewiston: The Edward Mellon Press, 1993.

"Robert Holkot [Holcot]." In *Stanford Encyclopedia of Philosophy*, July 2001. http://plato.stanford.edu/entries/holkot/

Hester Goodenough Gelber (2005)

HOLOCAUST

A number of philosophical issues arise relating to the destruction of most of the Jewish community in Europe in the twentieth century by the Nazis and their allies. This event has been labeled the Holocaust—or *Shoah* in Hebrew—in order to indicate its unique status.

UNIQUENESS

Was the Holocaust unique? There has been an extended debate on this issue. The obvious response is that it was not unique, because other ethnic groups have also been singled out for destruction—and have been destroyed—by powerful enemies. Hitler famously referred to the massacres of the Armenians in Turkey in the early part of the twentieth century when the question arose as to whether people would object to the Holocaust. Because few were interested in the fate of the Armenians—who had so recently been massacred—who would care about the Jews? Throughout human history groups of people who were in some way distinctive have been singled out for persecution and death, and the Jews are hardly the only target. Nor was the Holocaust the only large-scale act of genocide to occur; even in the same century there were several other instances of attempts to destroy an ethnic group.

It has been argued that the Holocaust is unique because never before, or since, have the entire technological resources of the state been directed in such a protracted manner against an indigenous community. The Holocaust took place over many years, against a group of people who could not be realistically regarded as any sort of threat to the state, and was in many ways carried out in opposition to the main aims of the war. For example, when the German army was short of railway stock in order to transport troops, the organizers of the Holocaust increased their efforts to direct stock away from the military in order to continue with the policy of annihilation. Even when the war was clearly lost the policy continued to be pursued until almost the last moment of practicability.

Why does the issue of uniqueness matter? It matters because if the Holocaust was unique, then it may call for new answers and directions. For example, it may give some validation to the creation of the State of Israel as a home for the Jewish people who survived. It may also call for new responses because it would then represent a break in history, and in particular in Jewish history. After all, Jewish history is replete with disasters of one kind and another, and the large-scale destruction of Jewish communities is a familiar feature of that history over the millennia. Is the Holocaust just another disaster among many similar—albeit more limited—disasters? Or does it represent a change in quality, not just quantity?

RADICAL RESPONSES

One of the most radical responses to the Holocaust is provided by Richard Rubenstein (1966), who argues that the events of the Holocaust rule out the traditional God of the Jewish Bible. The traditional God participated in Jewish history; were such a God to exist he would surely have participated in the Holocaust, and prevented it.

Because he did not, it follows that the concept of God has to change. What is needed is a concept of God that takes people closer to nature rather than away from it. Thus the traditional Jewish laws and rituals that emphasize the denial of nature are to be transcended and replaced with a far more hedonistic form of practice. For Rubenstein the return of the Jews to Israel typifies this, to a degree, because it represents a return to the land and to a more rooted and organic form of existence. By forging a new relationship with nature, Jews can transcend the negativity of history—whose paradigm is the Holocaust—and change Judaism itself.

Irving Greenberg (1981) also takes the Holocaust to compel radical steps—an end to the idea of a covenant between the Jews and God. Whereas to begin with the Jews were the junior partner in the covenant—and later on became equal partners—after the Holocaust the Jews are the senior partner, because God has to show himself prepared to act on behalf of the survivors if he is to play any role in their continuing lives at all. The whole idea of an agreement implies that both parties to it will uphold their side of the agreement, and God has clearly not held up his side because he allowed the Holocaust to occur. While Greenberg does see the hand of God in some events after the Holocaust—in particular in relationship to the State of Israel—he clearly holds God in dereliction of his duty, and calls therefore for a new relationship with him.

Arthur Cohen (2002) derives from the Holocaust the silence of God, and his distance from human affairs. To a degree this is not a new factor, because God has always been remote; he has to be if he is to allow people to be free and make their own decisions. Yet the God who emerges is clearly not the ordinary God of religion, but rather a deity who often hides his face and leaves his creatures to get on with their lives by themselves. Clearly such a God cannot be implicated readily in the State of Israel either, and Cohen is skeptical of the point of such a state, reflecting the doubts of Franz Rosenzweig on Zionism. It is the role of the Jewish people to typify a long and difficult relationship with God, not to live in a state of their own like everyone else.

Clearly these responses to the Holocaust call for a new definition of the relationship between God and the world. They also call for a new understanding of what constitutes religious practice, because the old prayers and rituals of Judaism may seem to be irrelevant given this new concept of God. What is worth noting is the crucial significance of the Holocaust to the propounders of these views. The Holocaust is not taken to be one disaster coming after many other disasters, but as an event with an existential meaning all its own. It is a unique event and so calls for unique responses. If those responses demand an entirely new understanding of Judaism, then it would be intellectually dishonest not to establish such an understanding.

ART AND THE UNIQUENESS DOCTRINE

There are important implications of the uniqueness doctrine for art. Adorno famously is supposed to have said that after Auschwitz there could be no art. As a factual claim this is problematic, because not only has there been art subsequently, there was even art during the Holocaust. However grim the conditions are under which artists work, they always manage to operate—some even believe that the harder the conditions the more important it is to respond aesthetically. What Adorno may be suggesting is that the whole context within which art takes place has changed irrevocably due to the Holocaust, and so art that does get done no longer has the character that it appears to have. For example, it may be that the German language has been so corrupted by its use in Nazi Germany that it can never be used again in a fresh and creative way. Although this may be plausible about German, it hardly would extend to other languages—and in fact does not even seem to describe German. Indeed, there has been no shortage of successful German prose and verse since the Holocaust, and in fact that event has often been its subject. It is difficult to make sweeping claims about art, of course, but it does not seem to have been noticeably altered by the Holocaust, nor has art changed much since the Holocaust.

Adorno probably means something a bit less obvious by his claim. Art rests on a whole range of human practices and expectations, and the Holocaust seriously threatened many of these. A defenseless and inoffensive minority were ruthlessly murdered by their fellow citizens, not as a random act of violence but through the machinery of the state and with little evidence of anyone outside of the minority disapproving. The scientific and rational forces of society were used for this purpose, occurring in what had until then been widely regarded as one of the most civilized and advanced societies in the world.

Adorno is pointing to the end of what is sometimes known as the Enlightenment Project, the idea that over time the world would progress as a result of the growing reliance on rationality and science. During the Holocaust, rationality and science were put entirely at the disposal of the murderers, and those techniques were revealed to be mere tools to be employed without reference to moral

restrictions. The optimism of the Enlightenment was thereby undone and should be replaced to by a thoroughgoing realism about the possibility of human progress.

The implications for art are clear. Whereas in the past it was thought that art has a civilizing impact, the Holocaust taught people that it may be enjoyed just as much by the morally corrupt as by anyone else. Thus its status changes from being an aspect of human nobility and cultural progress to becoming a morally neutral means of distraction. Hence Adorno's claim that art has irretrievably changed after the Holocaust.

OTHER RESPONSES TO THE UNIQUENESS THESIS

Emil Fackenheim (1982) presents a powerful defense of the uniqueness thesis, deriving from it what he calls the 614th commandment (there are traditionally held to be 613 commandments applicable to the Jews) that Hitler is awarded no posthumous victories. Such victories include assimilation and the destruction of the State of Israel, but Fackenheim does not see the Holocaust as calling for a radically new approach to Judaism itself or to the relationship between Jews and God.

Another important thinker is Elie Wiesel (1969), who wrote powerfully on his experiences and those of others during the Holocaust. He also sees it as not calling for a new understanding of faith. In particular, to the question of where was God at Auschwitz, he replies with the question where was humanity? The Holocaust represents an event carried out by human beings against other human beings and it is squarely on the shoulders of the murderers that the responsibility should be placed. One cannot expect God to rescue people from the evil decisions and actions of others, because were he to do so their capacity to act freely would be severely constrained.

This latter point is drawn on extensively by Eliezer Berkovits (1973) and Ignaz Maybaum (1965), different thinkers who agree that the Holocaust can be put within a normal Jewish theological context. The Holocaust does not represent a break in history, it is just one more disaster undergone by the Jewish people, and these disasters do have a point to them. God has a role in mind for the Jews, and this is to represent the divine role in history. That the Jews are never entirely destroyed reveals God's actions on behalf of the Jews. For Berkovits the Jews have to undergo suffering in order to sanctify the Holy Name, the traditional interpretation of Jewish suffering. For Maybaum the Holocaust represents an important stage in human history, and the sufferings of the Jews are supposed to lead the gentile world to reflect on the direction that their actions are leading them to pursue. Both thinkers discuss the difficult balancing act that God undertakes. He has to separate himself sufficiently from his creation in order to allow people to be free, whereas at the same time he has to enter into the human world in order to play a part in history.

THE HOLOCAUST FROM A CHRISTIAN POINT OF VIEW

A theme of many Christian views calls for some introspection into the responsibility of various churches for antiSemitism and its eventual outcome in the Holocaust. There have also been more positive analyses, in particular the argument that only Christianity can properly explain human suffering, because only Christianity has at its heart the notion of a suffering deity, in the person of Jesus Christ. The normal conception of God in Judaism is abstract, and many leading Jewish thinkers—such as Maimonides—have strenuously fought against any anthropomorphizing of the concept of the deity. This rather distant notion of a deity is said to be unhelpful during events such as the Holocaust.

It is certainly true that if people are in pain it is good to be comforted by someone who knows precisely what it is like to share in that condition. However, it may be argued that if such a person were in a position to relieve the pain, and does nothing, then the comfort is somewhat reduced. People may be more interested in pain relief than in sympathy, and indeed the latter may be valued largely as a stage on the route to the former. The Christian approach to the Holocaust does, however, raise the important question that runs through the debate—namely, what concept of God can survive the Holocaust experience? The more radical responses insist that a new concept of God is needed, whereas the less radical approaches defend the continuation of the traditional notion of God, and see the Holocaust as just another stage in Jewish history.

THE INDESCRIBABILITY THESIS

Fackenheim (1982) and others have declared that the Holocaust is indescribable. This follows to a degree from its uniqueness. If it is, as an event, really unique, then it could be argued that it escapes the normal categories of description. The indescribability of the Holocaust also explains to a degree why it has been little discussed by philosophers. The indescribability thesis does not appear to be plausible, because there have been many accounts of the Holocaust, and there seems to be little difficulty in

describing it. Around the world there are museums, memorials, and libraries designed to ensure that the world does not forget the Holocaust. In order not to forget the event itself must be described.

The indescribability thesis is rather like the uniqueness thesis in that it is intended not to literally make a claim about the Holocaust, but instead is a metaphorical indication of the extraordinary nature of the attempted extermination of the Jews in Europe as carried out by Germany and its allies. To say that this is easy to describe and that it is just one example of mass murder among many others has seemed to many commentators to diminish the enormity of the Holocaust.

ISRAEL

As noted earlier, one of the effects of the Holocaust is taken to be the creation of the State of Israel. Fackenheim suggests that the Holocaust represented the breakdown of Christian-Jewish relations, and the State of Israel is a *tikkun*, a repair of those relations. In the same way that the Holocaust is a break in history, so is the creation of Israel. It may well be that there is a factual link between the Holocaust and the State of Israel, but from a philosophical point of view it is difficult to see the logical link. Could God not have brought about a state for the Jews in a less costly manner? Was it really necessary for so many innocent people to die? And what about the rights of the people displaced from Palestine to make room for a Jewish state? These people played no part in the Holocaust, and yet were uprooted from the land by Zionism. The opposition to Zionism frequently compares it as a doctrine with Nazism in order to try to weaken the idea that Israel's existence is justified by the occurrence of the Holocaust.

RELATED MORAL ISSUES

An event of the stature of the Holocaust brings out sharply some interesting moral topics, such as the responsibility of the bystander for what goes on in his or her country, and the possibility of forgiveness for a crime of such enormity. During the Holocaust a large number of civilians apparently could have helped the victims, but did not, or could have expressed their views on what was happening, but declined to do so. Of course, there would have been a cost involved, yet the attitude of many was that the events of the Holocaust were not their responsibility because they were not actually the perpetrators and they had not troubled to find out precisely what was going on.

Since the Holocaust, the responsibility of the bystander—as opposed to the actual criminal actor—has become much more of an issue. For the criminal actor, the excuse of only following orders has become less defensible. Agents are expected to be able to consider the moral acceptability of the orders they are given and not carry them out if they are immoral. Finally, the issue of forgiveness and national responsibility arises. Who if anyone is entitled to forgive the agents of the Holocaust? Under what circumstances should they be forgiven—if they should be forgiven at all? What responsibility do their descendents have for the well-being of the survivors, or the Jews in general, or the State of Israel? Is it appropriate to blame a country—Germany—when most of its citizens were born after the Holocaust? How far can a country be held to be guilty at any time?

See also Jewish Philosophy.

Bibliography

Berkovits, Eliezer. *Faith after the Holocaust*. New York: Ktav, 1973.

Cohen, Arthur. *The Tremendum: A Theological Interpretation of the Holocaust*. New York: Crossroad, 1981.

Cohn-Sherbok, Dan, ed. *Holocaust Theology: A Reader*. New York: New York University Press, 2002.

Fackenheim, Emil. *To Mend the World*. New York: Schocken, 1982.

Garrard, Eve, and Geoffrey Scarre, eds. *Moral Philosophy and the Holocaust*. Aldershot: Ashgate, 2003.

Greenberg, Irving. *The Third Great Cycle in Jewish History.* New York: CLAL Resource Center, 1981.

Katz, Steven. "The Holocaust." In *History of Jewish Philosophy*, edited by Daniel Frank and Oliver Leaman, pp. 854–874. London: Routledge, 1997.

Leaman, Oliver. "The Holocaust." Chap. 10 in *Evil and Suffering in Jewish Philosophy*. Cambridge, U.K.: Cambridge University Press, 1985.

Maybaum, Ignaz. *The Face of God after Auschwitz*. Amsterdam: Polak & Van Gennep, 1965.

Rubenstein, Richard. *After Aushwitz: Radical Theology and Contemporary Judaism*. New York: Bobbs-Merill, 1966.

Wiesel, Elie. *Night*. Translated by Stella Rodway. New York: Avon Books, 1969.

Oliver Leaman (2005)

HOLT, EDWIN BISSELL

(1873–1946)

Edwin Bissell Holt, an American psychologist and philosopher, was noted for his innovations in philosophical psychology. His influence was greater in psychology than in philosophy. In his time he was the American psy-

chologist best known and most respected by the British. Holt completed his undergraduate and graduate work at Harvard and taught there from 1901 to 1918, first as an instructor and then as assistant professor of psychology. In 1926 he returned from retirement to become visiting professor of social psychology at Princeton, but he retired permanently in 1936.

NEW REALISM

Holt was one of the original six American New Realists who banded together in the first decade of the twentieth century in a polemic against idealism and representational realism. Holt was the only one, however, to attempt a systematic development of New Realism, first in a neutral monism, then, after giving that up, in a behaviorist theory of consciousness. In this attempt, Holt uncovered the fatal problems that were in New Realism from its beginning.

The New Realists took their start from the theory of consciousness of William James. James argued that consciousness was an external relation between a sentient organism and its objects, not a substance or entity. The latter view was the basis of the doctrine of the dualism of psychic and physical substances, and of an idealism that defined objects in terms of psychic or subjective substance, thereby giving them a mental or ideal status.

Holt replaced the dualism and psychic monism with a monism that was neutral, defining Being, or reality, neither in terms of mind (idealism) nor in terms of matter (materialism). The basic category of this neutral monism was "Being," which connoted nothing and denoted everything. This neutral Being could most readily be found in the concepts of logic and mathematics, the simplest known elements of Being. But in thus identifying Being with logical and mathematical terms and propositions, Holt gave it a distinctly mental or conceptual character. He admitted borrowing this approach from the idealist Josiah Royce, but he claimed that rather than arguing for idealism, his neutral monism reaffirmed the "sadly neglected truism" of New Realism: "everything is precisely what it is, and is not to be explained away as something else."

Yet Holt's analysis had an inescapable reductivist outcome. All things turn out to be "really" the same. That is, they turn out to be neutral entities (logical and mathematical terms) and the complexes made out of them (propositions), not the material things of common sense or the particles and elements of science. As one critic pointed out, this meant that it is the mathematical logician, not the physicist, who tells us what things are. By failing to keep clear the difference between the simplest elements of Being and the simplest known elements of Being, Holt threw doubt on the neutrality of his monism. The supposedly neutral logical and mathematical entities, he said, generate the further terms and propositions that make up all systems of being, or universes of discourse, through a "motion" of their own. Though Holt denied that this motion was a mental process, he did term it a "deduction," an intrinsic activity at work in the universe. Like any other object or aggregate, consciousness thus can be "deduced" from Being, and since Being is neutral, consciousness too is neutral; for all the complex constructions in experience are basically composed of neutral entities that maintain their identity despite the constructions they go into.

It is to these propositions, then, generated by the neutral entities of logic and mathematics through a "motion" of their own, that the nervous system responds. Although he admitted this might be considered fantastic, Holt stood by his position. James had said that the content of knowledge is the object of knowledge; content and object are not two separate things but are numerically one. Holt modified this only by noting that since our knowledge of an object is never complete, our ideas are never completely identical with their objects. When we say "My thought is of an object," we should say "My thought is a portion of the object; a portion of the object is my thought." Holt thought the representationalists had failed to see that an idea can represent an object only to the extent that it is identical with that object. An idea cannot represent space, then, without itself being spatial; the only adequate idea of a minute or an hour is just a minute or an hour. Holt thus passed from a partial qualitative identity of knowledge or consciousness and its objects to a numerical or existential identity. He had forgotten that he had begun with James's idea of consciousness as a relation.

No other New Realist developed monism in this thoroughgoing fashion. Holt carried it to its furthest conclusion. If consciousness and its object are numerically identical, do objects then have the character of consciousness (panpsychism), or does the content of consciousness have the character of objects (an inverted panpsychism, or "panobjectism")? Holt's anti-idealism ruled out panpsychism; consequently the elements of consciousness became objects themselves among all other objects, and the world for Holt is populated with all those entities usually placed in consciousness: error, hallucination, delusion, secondary qualities, even volitions. The objective world contains physical counterparts to the

errors of ordinary sense experience. Errors of thought, always cases of contradictory propositions, are equally objective. The "real," or objective, world is contradictory through and through; nature is a "seething chaos of contradiction."

Holt, in a later paper on the locus of concepts, confessed that his neutral monism had led him to write a mistaken book, an "absurd hocus-pocus" conjured up because he did not know at the time the true locus of these neutral "timeless and changeless entities." His failure to maintain their neutrality is admitted: These entities have no objective existence in nature. Although he promised to return to the subject, Holt never did. Nor did he produce the planned second volume that was to carry out the epistemological implications of his neutral monism. Instead, he turned to the development of a behaviorist theory of consciousness.

BEHAVIORIST THEORY OF CONSCIOUSNESS

Holt saw that an extreme behaviorism would make the materialist's mistake of denying the facts, as well as the theory, of consciousness. While he described his own behaviorism as part of the "objective tendency" to abolish the subjective and to interpret mental phenomena in an "objective relational manner," he consciously sought to avoid slurring over or repudiating the "facts" of consciousness, and he modified his behaviorism accordingly.

Increasingly, ideas suggestive of subjectivity, if not dualism, such as integration of behavior, capacity to respond, suppression, and split personality, appeared in Holt's writings. The result was an oscillation between his objectivist, behaviorist ideas and the subjectivist ideas that he needed in order to do justice to the facts of consciousness.

In *The Freudian Wish,* he described behavior by examining the way in which reflexes are combined and integrated to produce that organized "synthetic novelty" which is the specific response, or behavior, and which is also the point at which awareness is born. He identified this response with Sigmund Freud's "wish," including in it purpose, tendency, desire, impulse, and attitude. It was the replacement, Holt claimed, for sensation as the unifying factor of psychology. But he denied that this view meant he was falling back on the psychic or subjective; the basis for the view was objectively observable in what an organism does. While he did not deny that we have unobservable thoughts, he argued that they are often an "embroidery, a mere irrelevance to action," and eventually they too can be observed if one looks to behavior that is yet to come.

Holt thought the Freudian wish was the first key that psychology had discovered for an explanation of mind. It meant psychology "with a soul," not the "ghost-soul" but the "wishes" which are the soul. Like Aristotle, he identified the soul with the dynamic form of a body endowed with the capacity of life: it is what it can do. The behavior of such a body is distinguished from its random movements by its purposiveness, an objective reference that is found in every reflex. Behavior occurs when more than one reflex is set off by a stimulus. As the number of reflexes increases, the immediate stimulus "recedes" as the inciting and controlling factor. This recession of the stimulus is part of intelligence and deliberation. Holt also used it to give an account of consciousness and knowledge of spatially and temporally remote objects. Still, Holt could not avoid a basic monism. The "objective" world is the only world. What has been called the "subjective" world is the subtler workings of integrated objective mechanisms. It is the body that is the knower; the environing objects to which it responds are the known. And Holt revived his claim that the mind is the thing of which it is thinking.

By the end of his career, despite his lifelong objectivist-subjectivist oscillations, Holt was committed to an objectivist position. He described his last published book (on the learning process) as an essay toward radical empiricism, and it was supposed to complete James's work of ridding philosophy and psychology of the absurdities of subjectivism and any form of psychophysical parallelism. There is only a sketchy idea at the end as to what direction Holt's epistemology might have taken. He thought at that point that he was but one short step away from a definition of awareness and consciousness in physiological terms. His "objectism" was reaffirmed: he sought to formulate a wholly physical and physiological psychology as a basis for the solution of any psychological problem. But he admitted that such a psychology had not yet given the slightest clue to the problem of secondary qualities.

Holt's last published writing set forth a materialism without apologies. Mind and cognition are neither mental nor cognitive, but physical—a matter of nerves and muscles. The active self is the physical body, that and nothing else. An experience of "self" is an experience of parts of one's body. Anything other than that, whether a self, ego, soul, or knower, does not exist.

Still, Holt modified this objectivism. He admitted that our physiological apparatus of perception and

thought habitually distorts, mutilates, and disguises what it is perceiving. It subtracts from "the objective reality," and with the remainder it fuses inseparably "a vast amount of unreality of its own motor creating (subjective reality)." In a mistaken but significant interpretation of Immanuel Kant, Holt claimed that these distortions are strikingly analogous to the Kantian categories in their distortion of things-in-themselves in intuition and understanding.

The ghost of subjectivism remained. Holt and the New Realists may have exorcised its idealist form, but the need for its inclusion was a constant embarrassment to them and was eventually the reason for the failure of New Realism to be anything more than an anti-idealist polemic.

See also Aristotle; Behaviorism; Being; Freud, Sigmund; Idealism; James, William; Kant, Immanuel; Monism and Pluralism; New Realism; Panpsychism; Royce, Josiah.

Bibliography

Works by Holt are *The Concept of Consciousness* (London: George Allen, 1914; manuscript completed in 1908), which presents his doctrine of neutral monism; "The Place of Illusory Experience in a Realistic World," in E. B. Holt et al., *The New Realism* (New York: Macmillan, 1912), pp. 313–376—the book contains essays by the six New Realists elaborating their 1910 platform statement; "Response and Cognition," in *Journal of Philosophy* 12 (1915): 365–373, 393–409, reprinted in *The Freudian Wish; The Freudian Wish and Its Place in Ethics* (New York: Holt, 1915), his most popular and widely read book and one of the first introductions of Freud to a large American audience; "Prof. Henderson's 'Fitness' and the Locus of Concepts," in *Journal of Philosophy* 17 (1920): 365–381; *Animal Drive and the Learning Process: An Essay towards Radical Empiricism* (New York: Holt, 1931), Vol. I, probably Holt's most significant publication in psychology; "The Whimsical Condition of Social Psychology and of Mankind," in Horace M. Kallen and Sidney Hook, eds., *American Philosophy Today and Tomorrow* (New York: Furman, 1935); and "Materialism and the Criterion of the Psychic," in *Psychological Review* 44 (1937): 33–53.

A secondary source is Syed Zafarul Hasan, *Realism: An Attempt to Trace Its Origin and Development in Its Chief Representatives* (Cambridge, U.K.: Cambridge University Press, 1928), an extensive consideration of American philosophic realism in all its variations, including a critical exposition of Holt's philosophy.

Thomas Robischon (1967)

HOME, HENRY
(1696–1782)

Henry Home (Lord Kames), an aesthetician and moral philosopher, was born at Kames, Berwickshire, Scotland. He was educated at home and indentured to a writer of the signet in Edinburgh, but he resolved to become an advocate and was admitted to the Scottish bar in 1724. He became a judge of the Court of Session in 1752 and assumed the judicial title of Lord Kames. He was advanced to the High Court of Justiciary in 1763 and was still serving at the time of his death.

Kames wrote a number of books, several of them on legal subjects. His *Sketches of the History of Man* (2 vols., Edinburgh, 1774) bridged his interests in history and philosophy, and he frequently referred to the *Sketches* in his *Essays on the Principles of Morality and Natural Religion* (3rd ed., Edinburgh, 1779). His other philosophical work is *Elements of Criticism* (2nd ed., Edinburgh, 1763), a discussion of aesthetic principles.

Kames argued that the fundamental principles of the fine arts, or the elements of criticism, must be drawn from human nature. The fine arts are suited to human nature because humans, as sensitive beings, are capable of pleasure; and the fine arts are calculated to give pleasure to eye or ear. Kames devoted the opening chapters of the *Elements* to an account of human emotions and passions. These chapters form the psychological prolegomena that he believed aesthetics requires. Perceptions and ideas occur independently of our wills, though we can sometimes will the cessation of a train of ideas. Ideas follow our perceptions and each other in accordance with the laws of association (resemblance, contiguity in time or place, and cause and effect). Emotions and passions occur in relation to our train of perceptions and ideas. A passion is an emotion that is accompanied by a desire. The general rule for the occurrence of emotions is that we love what is agreeable and hate what is disagreeable. Kames's basic principle of criticism is that every work of art that is conformable to the natural course of our ideas is so far agreeable, and every work of art that reverses that course is so far disagreeable. On the one hand, Kames wanted to establish that the agreeableness or disagreeableness of things is prior to our love or hatred; but on the other, he accounted for our emotional reactions to certain things by saying that the nature of man is originally framed with a relish for regularity, uniformity, proportion, order, and simplicity.

The fine arts that Kames had in mind are painting, sculpture, music, poetry, gardening, and architecture; but

the first three are not discussed systematically in the *Elements.* Poetry is given the most extended criticism. Kames was especially interested in plays, and gardening and architecture share a chapter. He divided aesthetic qualities into two sorts: those that an object may possess in itself and those that it has in relation to other objects. Qualities of the first sort are grandeur, sublimity, motion, force, novelty, "laughableness," and beauty, which he conceded are both intrinsic and relational. The relational qualities that Kames discussed are resemblance and dissimilitude, uniformity and variety.

Kames argued that it should be possible to establish a standard of taste against which productions in the fine arts might be judged. We believe that things of a certain kind have a common nature, and individuals are perfect or right insofar as they conform to the common nature of their kind. Thus, it should be possible to determine the common nature that works of art of a certain kind ought to share and to assess the success with which a given work of art meets the ideal of its kind. Kames noted that every person is not fit to become a judge of the fine arts, since not everyone is capable of the refinement of taste that is required. This is no great hardship on the bulk of humankind. The fine arts only contribute to our pleasure and amusement, and it is not as necessary for everyone to have an authoritative sense of right and wrong in the fine arts as it is for everyone to have an authoritative moral sense.

In *Essays on the Principles of Morality and Natural Religion,* Kames discussed a wide range of philosophical topics, including liberty and necessity, personal identity, belief, external senses, and cause and effect. His thinking is influenced by David Hume, either in quiet concurrence or by vigorous reaction. The two longest and most important essays are Essay II, "The Foundation of Morality," and Essay VIII, "Knowledge of the Deity."

For Kames, the foundation of morality is to be found in human nature. Looking there, he finds the moral sense that approves certain natural principles, which are enforced by natural rewards or punishments of pleasure or pain. These principles bind us to refrain from harming others, to tell the truth, to keep our promises, to act faithfully toward those who rely on us, to be grateful, and to be benevolent. While the moral sense is rooted in the nature of man, it admits of great refinements by culture and education.

A hasty reader might conclude that whenever Kames needed to solve a new perplexity in the foundation of morals, he discovered a new sense in humankind. For instance, he resolved the long-standing dispute over the artificiality of justice by declaring that justice is natural because it is founded on a natural sense of property. He claimed that this sense is necessarily antecedent to any social agreement; and indeed that any agreement to organize a society presupposes the existence in men of a sense of property.

In natural religion, Kames believed that he had brought to light a new argument to prove the existence of a god. In D. Cranz's *The History of Greenland* (London, 1767) Kames found an account of a Greenlander who argued in the following way for the existence of an artisan superior in power to man: A kayak is a work of art that can be made only by the most skilled of men, but a bird is an even greater work of art than a kayak; thus there must be an artisan to make birds who is even greater than man. Kames was most impressed by the fact that this argument came from a savage and concluded that "the perception we have of Deity must proceed from an internal cause, which may be termed the Sense of Deity."

In the *Essays,* Kames generalized the Greenlander argument, contending, "We are so accustomed to human arts, that every work of design and use will be attributed to man, if it exceed not his known powers. *Nor do effects above the powers of man unhinge our notion of a cause*: They only lead the mind to a more powerful cause." The italicized words in the passage above are especially interesting, because in an addendum to the third edition of the *Essays* (1779), Kames complained that Hume ignored the Greenlander argument in his *Dialogues concerning Natural Religion*; and Kames believed that argument immune from any strictures on natural religion was found in Hume's *Dialogues.*

See also Aesthetic Judgment; Hume, David; Pleasure; Religion and Morality.

Bibliography

Lewis Amherst Selby-Bigge, ed. *British Moralists* (Oxford: Clarendon Press, 1897), Vol. II, contains a representative selection of passages from the essay on the foundation of morals (Essay II) in *Essays on the Principles of Morality and Natural Religion.*

Elmer Sprague (1967)

HOMER

The Homeric poems *Iliad* and *Odyssey* (probably eighth century BCE) are of interest to the historian of philosophy because they provide the background, in language

and to some extent in thought, from which Greek philosophy emerged. The hexameters of Parmenides and Empedocles follow the Homeric pattern closely, and they both use Homeric words and coin words for themselves after the Homeric model. They also sometimes use the same thought forms. For instance, a comparison may be drawn between Parmenides' journey (see Fr. 1) and Odysseus's journey to the underworld (*Odyssey,* Book 11). The Homeric simile is the forerunner of the natural philosopher's "working model," by which an unfamiliar process is explained by comparison with a more familiar one. For example, to illuminate his description of an evenly poised battle Homer introduced a "careful working woman" weighing wool in her scales; Empedocles compared the breathing process in animals with operations performed with a household instrument, the *clepsydra.*

Apart from these questions of language and style, the *Iliad* and the *Odyssey* influenced the content of later philosophical thought in various ways.

THE WORLD

The Homeric world picture was of a flat, disk-shaped earth, with the sky set over the top like an inverted metal bowl and Hades underneath the earth in a more or less symmetrical relation to the sky. The sun, moon, and stars were taken to move across the fixed heaven from east to west, but the manner of their return journey was not clear. The space between the earth and the sky contained *aer* (mist), and above that was *aether* (the bright air of the upper heavens). The earth was completely surrounded by the river of Ocean, personified and deified as Okeanos. In one exceptional passage (*Iliad,* Book 14, 200–248) Okeanos is called "the begetter of gods" and "the begetter of all things." Aristotle (*Metaphysics* A 3, 982b27) half seriously suggested that Homer's Okeanos was the forerunner of Thales' cosmogonical water. Plato, even less seriously, suggested (*Theaetetus* 152E) that Okeanos provided the origin of Heraclitus's flux theory. These are farfetched ideas; the cosmology of Homer, such as it was, can hardly be seen as anything but a contrast with Ionian theories (see G. S. Kirk in *The Presocratic Philosophers,* Ch 1). But connections can be traced between some details of Homer's descriptions of the natural world and the speculations of later Greek philosophers of nature (see Charles Mugler, *Les origines de la science grecque chez Homère*).

THE GODS

The historian Herodotus observed that Homer and Hesiod together had determined for all the Greeks what their gods were like, and this is probably the greatest significance of the Homeric poems for the history of philosophy (*History* II, 53). There is one general feature about the Homeric gods that is of much importance: They were not dark gods, accessible only to mystics and appeasable by magic, but on the whole very human and rational. They had powers over the world of human experience, and their powers were defined and hierarchical; in this we can see a hint of the orderly cosmos of later theory.

Some philosophers objected to the Homeric gods. Xenophanes launched the first attack: The gods behaved immorally; moreover, the conception of them was relative to the believer (see Fr. 16: "Ethiopians imagine their gods as black and snub-nosed"). Heraclitus's objections were not explicitly against gods but against Homer as the educator of Greece; the Olympian gods were, however, near the center of his target. Plato's onslaught in the *Republic* (376Eff.) is well known; he wished to censor everything in the Homeric poems that was discreditable to the gods before the poems could be used in the education of the "Guardians" (it was general practice in Greece both before and after Plato's time to use Homer as the basis for moral and religious education).

MAN

The Homeric view of man shows interesting differences from later theories. There was no unified soul, contrasted with body, as in the Pythagorean-Platonic tradition; instead, the psychic functions were distributed without much consistency over a number of entities. The *psyche,* which held the position of greatest importance from the time of Pythagoras, was merely a life-soul in Homer; it played no part in the thoughts, emotions, and actions of the living man. The *psyche* survived after death; it did not, however, retain the complete moral personality, as in the Platonic eschatology, but was a bloodless, helpless shadow. The thoughts and feelings of the living man were attributed to the *phrenes* (roughly speaking, the organs of the chest, although in later Greek the word means "diaphragm"), the heart, and the *thymos* (a mysterious entity probably connected, like *psyche,* with breath). *Nous* (mind), which became the most important part of the *psyche* in the psychology of Plato and Aristotle, was generally restricted in Homer to the intuitive understanding of a situation (like the English "to see" in its metaphorical sense); consequently, it was often connected with sense perception, not contrasted with it as in Plato. Unlike *phrenes, nous* was not a physical thing for Homer but a function.

HUMAN ACTION

The actions of the human characters in the *Iliad* and *Odyssey* are represented as being influenced or manipulated more or less constantly by the gods. Actions that might be otherwise difficult to explain, such as a sudden access of superhuman courage, are especially attributed to the intervention of a god. But it is not only the inexplicable or the uncharacteristic that is described thus; a successful shot with the spear or an unsuccessful one, a plan adopted, a fit of anger, a bad bargain, an untimely sleep—these and many other unremarkable events are described as caused by a god. The gods handle the heroes as arbitrarily as a mortal king might treat his subjects, although not, as a rule, with savagery.

The fact that so much of human action is attributed to the gods has led modern interpreters to say that Homeric man is "an open field," that Homer denies free will, and that he has no concept of the human personality. This is true in a sense, but it is misleading. Homer was not a philosopher who had confronted the free-will problem and decided upon determinism; apart from an occasional exception he offered no theories about motivation and responsibility. From the point of view of the responsibility of human characters, there is no opposition between "caused by a god" and "due to a human agent"; for example, one and the same attack by Sarpedon is described as due to Zeus and a few lines later as due to Sarpedon's *thymos* (*Iliad,* Book 11, 292 and 307). The moral relations between human beings are on the whole, although not entirely, unaffected by the interventions of the gods; a god may stir a man to excessive anger, but it is still felt appropriate to blame the man for his anger. The individual characters of the heroes remain fairly stable; the activity of the gods is not such as to make human beings unpredictable. But the Homeric poems generally show a limited sense of moral responsibility. They were composed at a time when shame still predominated over guilt as a motivating force, and the intention of the agent and his knowledge of the circumstances of his act (the two factors that of course played the chief part in later legal and philosophical theories of responsibility) receive little attention.

Homer provided the material for much of later Greek literature, which examined the relation of Homeric gods and men in a new way. The problem of individual human responsibility for actions in which gods were said to be involved, though hardly seen by Homer, was much discussed by the fifth-century tragedians and Sophists (see, for example, Aeschylus, *Agamemnon* 1497ff.; Euripides, *Troades* 914ff.; Gorgias, *Helen*).

See also Aristotle; Empedocles; Heraclitus of Ephesus; History and Historiography of Philosophy; Parmenides of Elea; Plato; Platonism and the Platonic Tradition; Pre-Socratic Philosophy; Pythagoras and Pythagoreanism; Sophists; Xenophanes of Colophon.

Bibliography

The best text of the *Iliad* is edited by Martin L. West (Munich and Leipzig: Saur, 1998–2000); for the *Odyssey* see the text edited by Peter von der Muehll (Stuttgart: Teubner, 1962). Translations include *The Iliad,* translated by Richmond Lattimore (Chicago: University of Chicago Press, 1962), and *The Odyssey,* translated by Robert Fitzgerald (New York: Doubleday, 1961). For commentaries see Geoffrey S. Kirk, ed., *The Iliad: A Commentary* (Cambridge, U.K.: Cambridge University Press, 1985–1993) and Alfred Heubeck, Stephanie West et al., *A Commentary on Homer's Odyssey* (Oxford: Oxford University Press, 1988–1992). Relevant texts with translation and commentary are included in Geoffrey S. Kirk, John E. Raven, and Malcolm Schofield, *The Presocratic Philosophers,* 2nd ed. (Cambridge, U.K.: Cambridge University Press, 1983), Ch. 1, sections 1 and 2.

Secondary sources include E. R. Dodds, *The Greeks and the Irrational* (Berkeley: University of California Press, 1951); Francis M. Cornford, *From Religion to Philosophy* (London, 1912); R. B. Onians, *The Origins of European Thought* (Cambridge, U.K.: Cambridge University Press, 1951); Bruno Snell, *The Discovery of the Mind,* translated by Thomas G. Rosenmeyer (Oxford: Oxford University Press, 1953), Chs. 1 and 2; Félix Buffière, *Les mythes d'Homère* (Paris: Belles Lettres, 1956); Charles Mugler, *Les origines de la science grecque chez Homère* (Paris: C. Klincksieck, 1963); Hermann Fränkel, *Early Greek Poetry and Philosophy,* translated by Moses Hadas and J. Willis (Oxford: Oxford University Press, 1975), Ch. 1; Jasper Griffin, *Homer on Life and Death* (Oxford: Oxford University Press, 1980), Chs. 5 and 6; Hugh Lloyd-Jones, *The Justice of Zeus* (Berkeley: University of California Press, 2nd ed., 1983), Chs. 1 and 2; Naoko Yamagata, *Homeric Morality* (Leiden: Brill, 1994).

David J. Furley (1967)
Bibliography updated by Richard Janko (2005)

HÖNIGSWALD, RICHARD
(1875–1947)

Richard Hönigswald, the German philosopher, was born in Magyarovar, a small Hungarian town near the Austrian border. He received a degree in medicine from the University of Vienna in 1902 and then studied philosophy under Alexius Meinong at Graz and Alois Riehl at Halle, receiving a doctorate in philosophy in 1904. He taught at the University of Breslau from 1906 until 1930, when he accepted a chair in philosophy at the University of Munich. Because he was a Jew, Hönigswald was deprived

of his academic position in 1933. After the pogrom of 1938 he was sent to Dachau, but in 1939 he managed to immigrate to the United States. He lived in New York and engaged in research and writing until his death.

Hönigswald remained closer to the original doctrine of Immanuel Kant, as exemplified in the Transcendental Aesthetic, the *Critique of Practical Reason,* and the *Critique of Judgment,* than did such Neo-Kantians as Hermann Cohen, Paul Natorp, and Heinrich Rickert. However, he emphasized the insufficient consideration Kant had devoted to the importance of the concrete subject as a historical and empirical entity. Out of this criticism of Kant, Hönigswald developed his own influential theory of concrete subjectivity, the psychology of thinking (*Denkpsychologie*). According to Hönigswald, the concrete subject, an individual monad, is both fact (*Tatsache*) and principle (*Prinzip*)—that is, it is both a constituent of the world and an entity that recognizes itself as the correlate of the world, confronting it in cognition, volition, and artistic productivity. In the concrete subject, ground and grounded, objectivity and object, coincide; in a natural object they are separated. This doctrine forms the basis of Hönigswald's cosmology. In its attempt to determine the concrete subject's position in the world and its specific temporal structure in terms of a regional ontology, Hönigswald's philosophy exhibits similarities to Edmund Husserl's *Konstitutionslehre,* Martin Heidegger's analysis of *Dasein,* and Nicolai Hartmann's theory of stratified being. Hönigswald's approach differs from these in that he adhered to classical principles of validity (*Geltungsprinzipien*) in epistemology, ethics, legal and political philosophy, aesthetics, and the philosophy of religion. He found the key to the differentiation of the corresponding judgments and cultural realms in the constitutive features of the subject (thereby departing from Marburg and southwest German Neo-Kantianism), which he classed as intentionality, self-determination, reference to nature, and unlimitedness.

Hönigswald's philosophy of language made a considerable impact on Continental linguistics. Just as fact and principle coincide in the individual monad, Hönigswald claimed, the intermonadic reference of language constitutes the one other instance of the coincidence of fact and principle. Hönigswald's educational thought influenced such philosophers as Moritz Löwi and Alfred Petzelt, who, like him, emphasized the notions of tradition, concentration, and projection into the future. A number of thinkers, including Bruno Bauch, Theodor Litt, Wolfgang Cramer, and Hans Wagner, have engaged in evaluating Hönigswald's teachings for the study of fundamental problems in philosophy.

See also Cohen, Hermann; Hartmann, Nicolai; Heidegger, Martin; Husserl, Edmund; Kant, Immanuel; Language, Philosophy of; Meinong, Alexius; Natorp, Paul; Neo-Kantianism; Rickert, Heinrich; Riehl, Alois; Subjectivity.

Bibliography

WORKS BY HÖNIGSWALD

Über die Lehre Hume's von der Realität der Aussendinge. Berlin, 1904.

Hobbes und die Staatsphilosophie. Munich: Reinhardt, 1924.

Immanuel Kant. Breslau, 1924.

Die Philosophie des Altertums, 2nd ed. Leipzig: Teubner, 1924.

Die Grundlagen der Denkpsychologie, 2nd ed. Leipzig: Teubner, 1925.

Über die Grundlagen der Pädagogik, 2nd ed. Munich: Reinhardt, 1927.

Grundfragen der Erkenntnistheorie. Tübingen: Mohr, 1931.

Geschichte der Erkenntnistheorie. Berlin, 1933.

Philosophie und Sprache. Basel: Haus zum Falken, 1937.

Schriften aus dem Nachlass. Vols. I–IV, Stuttgart, 1957–1961; Vol. V, Bonn, 1965; Vol. VI, 1967; Vol. VII–VIII, 1970.

A fuller but incomplete bibliography may be found in Werner Ziegenfuss, *Philosophen-Lexikon,* Vol. I, 553–554. Berlin: W. de Gruyter, 1949.

Posthumous papers are available in Hönigswald-Archiv, University of Bonn.

WORKS ON HÖNIGSWALD

Opahle, Oswald. "Richard Hönigswald und seine Stellung zum Ganzheitsproblem." *Die Ganzheitsschule* 10 (1961): 1–9.

Ueberweg, Friedrich. *Grundriss der Geschichte der Philosophie,* 13th ed., Vol. IV, 432. Tübingen, 1951.

Wolandt, Gerd. "Problemgeschichte, Weltentstehungsmythos und Glaube in der Philosophie Richard Hönigswalds." *Zeitschrift für philosophische Forschung* 12 (1958): 188–217.

Wolandt, Gerd. *Gegenständlichkeit und Gliederung. Untersuchungen zur Prinzipientheorie Richard Hönigswalds.* Cologne: University of Cologne Press, 1964.

Gerd Wolandt (1967)
Translated by Eva Schaper

HOOKER, RICHARD
(1553–1600)

Richard Hooker, the English theologian and social and political philosopher, was born at Heavitree, near Exeter. His family was poor but well connected, and in 1568 Bishop John Jewel secured for Hooker a clerk's place at Corpus Christi College, Oxford. He became a fellow in 1577 and upon his marriage in 1581 was presented with

the living of Drayton-Beauchamp and a few months later with the mastership of the Temple in London. At the Temple, Hooker came into violent conflict with William Travers, a Calvinist who lectured there in the evenings. Although Hooker always retained a high regard for Travers's intellect and integrity, he was forced by his own convictions to oppose the views of Travers. It was during this controversy that Hooker seems to have conceived the idea of writing a systematic treatise to uphold the establishment of church and state as represented by Queen Elizabeth's policies. In order to carry out this plan, he requested a transfer from the unquiet position in London to a country rectory. Thus he went to Boscombe near Salisbury, where he was able to write and complete the first four books of his projected treatise, *The Laws of Ecclesiastical Polity,* by 1593 or 1594. In 1595 he was promoted to the rectory of Bishopsbourne near Canterbury, where he completed the fifth, purely theological part of his treatise by 1597. During the following three years he wrote another three books for the *Laws,* but he did not live to see them published. He died toward the end of 1600.

HOOKER'S IMPORTANCE

Hooker was not an original thinker. His importance lies in the fact that he drew upon the various currents of medieval thought in order to explain the ecclesiastical and political institutions of Elizabethan England. Together with Francisco Suárez and Cardinal Robert Bellarmine he belonged to the first Counter-Reformation generation, and like the two Jesuits he elaborated the final implications of Aristotelianism and of Thomism in social and political philosophy. But unlike his two Jesuit contemporaries, he did not live in the orbit of the Roman Catholic revival. To both Suárez and Bellarmine the Renaissance, the Protestant Reformation, and the Erastian state were merely threats they had heard of—threats from the outside. But Hooker was an Englishman who had grown up and lived through the turmoil occasioned by the attempt of radical Protestantism to force Queen Elizabeth from her conciliatory path. As a result he had to parry the practical attack of the extreme Protestant wing, and he finally came face to face with the secular state's opposition to that wing. This confrontation lends Hooker's thought an air of real drama; and if he was less systematic in his exposition than Suárez, his writings have the advantage of revealing a genuine intellect at work, wrestling with problems, not merely teaching what is imagined to be the truth.

THE SOURCE OF AUTHORITY

Hooker's analysis of the Puritan attack on the Elizabethan settlement in church and state had revealed to him the essential similarity of that case with a line of argument that had a long and distinguished medieval ancestry and in some ways went back as far as St. Augustine. The attack the Puritans mounted against the Elizabethan settlement drew heavily on John Calvin and to a lesser extent on John Wyclyf, and was ultimately analogous to all those medieval arguments that had denied the validity of natural law and therefore of the justification of secular authority in terms of natural law. Lacking a justification in natural law, the secular state, if it was to have any legal and moral basis at all, had to be subject to divine authority. To medieval writers this divine authority was represented on Earth by the papacy; to the sixteenth-century Calvinists, it resided in the presbyteries of the godly and the elect. In order to combat the view that men have no natural reason with which to discover a natural law, and the view that any law discovered or made by men is incompatible with divine law, Hooker fell back upon the philosophy of St. Thomas Aquinas. That philosophy had been developed during the thirteenth century to establish a doctrine of natural law and natural reason and to show how the rules thus discovered were fully compatible with those supernaturally revealed by God. The first book of Hooker's treatise is therefore a readable sixteenth-century compendium of Thomistic philosophy.

NATURAL AND REVEALED LAW

Like St. Thomas, Hooker believed that man is by nature a social animal and that both the impulse to live in society and the need for some kind of government is inherent in human nature. Man is therefore created by God with the rational endowments necessary for the conduct of society and government. All social and political arrangements are hence subject to natural law, which is immutable. But since conditions of life vary from time to time and place to place, it is necessary to supplement the dictates of natural law with positive or "human" rules. All this was taken from Aristotle, but translated by both St. Thomas and Hooker into the context of Christian thought. Men desire not only to live, however, but also to live well. This further desire implies that they must find their ultimate happiness. Such ultimate happiness cannot be found in the attainment of a temporal, and therefore temporary, good, but only in the ultimate perfection that is God. Owing to the Fall, man cannot know by natural reason what he must do to obtain this final supernatural end. God has therefore revealed to man certain rules to supplement

natural law. Hence it becomes clear that in order to achieve full human stature, man needs both natural law, for social and political purposes, and revealed law, for everlasting felicity. Revealed law is contained in the Bible and the traditions of the church. Natural law and revealed law are jointly, and not separately, the correct guide.

THE "LEX AETERNA"

In order to establish his point that the two sets of laws must be brought into operation jointly, Hooker delved into cosmology. God, he wrote, is the author of everything. He is a law unto himself, and that law is the *lex aeterna,* which is both the source of all other law and itself manifested in all other laws. In the divinely revealed law, it is manifested directly, so to speak; in natural law, indirectly. For natural law is discovered by human reason, and human reason is created by God according to the *lex aeterna*; therefore the dictates of natural law, and even the positive rules of human law, spring from the *lex aeterna.* God has given reason to every man. He has "illuminated" him. Although there is no explicit reference to St. Thomas in Hooker's text, this argument is a transcription of one of the central tenets of Thomism: *signatum est super nos lumen vultus tui, domine* ("the light of thy countenance is signed upon us, Lord"; Psalms 4:6–7). Hence we learn the will of God by using our reason.

OTHER THOMIST DOCTRINES

Hooker identified himself with all the more salient doctrines of St. Thomas. He argued that God is pure act and that in him existence and essence coincide; that angels are immaterial and that they differ from all natural, not purely intellectual creatures in that they behold the face of God directly; that the soul is the form of man, and not a separate substance as St. Thomas's opponents had argued.

The will of man, Hooker wrote, is free. Everything good that reason sees as such has something unpleasant annexed to it. And everything evil that reason sees as such has something pleasant attached to it. For reason cannot see the absolutely good. Hence, although we always will the good, we can never will the absolutely good; as a result the will is always free to choose between several relative goods. Hooker believed that the two springs of human action are knowledge (reason) and will. The will always wills the good; and the good is apprehended by reason. Sin results from the imperfect operation of reason, which can never apprehend the absolutely good. Sin is therefore intimately linked with both the freedom of the will and the imperfection of reason. It is never committed as a positive action or desired for its own sake, but is the result of a loss. Evil, by implication, is a privation.

These subsidiary arguments were important to Hooker not only because they enable the reader to identify the main lines of Thomism but also because they help to lead to the goal of the main argument. To avoid evil, it is necessary to supplement the law of nature. And since the law of nature is embodied in secular government, the revealed law is embodied in the church. Thus Hooker arrived at his main objective, the proof that church and state are intimately connected.

ECCLESIASTICAL AND SECULAR SOCIETY

As long as the argument remained confined to a high level of generality, it was easy to take for granted that this philosophy amounted in fact to a defense of the Elizabethan establishment, in which church and state were closely identified. Such reforms as Henry VIII and Elizabeth introduced into the church never really severed the visible continuity of ecclesiastical institutions and of canon law in England. The Elizabethan settlement, like Henry's acts of law, had been made by Parliament; in a very general sense, Parliament appeared to Hooker not as a purely secular institution. The bishops were part of it; and the electors themselves, being members of the church as well as members of a secular society, could easily be deemed to constitute in fact an ecclesiastical polity.

Hooker was explicit on the importance to his argument of the identity of the people who were the church with the people who were the commonwealth. He admitted that in countries where no such identity could be presumed, the natural society (being hierarchically lower than the ecclesiastical society) could not be deemed capable of making laws for the church. But in England, he was confident, complete identity obtained.

Thus Hooker was able to establish his initial point that the Puritan attack upon the Elizabethan settlement and the Puritan demand for the establishment of presbyteries and congregations was based on a false estimate of human nature. For it assumed that there was no natural law to justify the existence of secular society and of secular government, that all authority would ultimately have to be vested in the congregations representing the godly and the elect who embodied the only law there was, the divine law.

NATURALISM AND ANTI-PLATONISM

When Hooker turned to writing about the more particular arrangements of the Elizabethan settlement, he had difficulty squaring his Thomist theory with political practice, which was Erastian and naturalistic in the extreme. In an attempt to do so he drew heavily upon the ideas of Marsilius of Padua, who had completely subjected the church to the state. Hooker had begun as a confident Thomist; with the discovery that Thomism did not suffice to account for the intricacies of late Tudor politics, he found himself in a tangle once he began drawing upon ideas from the naturalistic thought of Marsilius, which was completely incompatible with Thomism.

The crux of the tangle was Hooker's unflinching Aristotelianism, probably absorbed when he was a student at Oxford. It was his Aristotelianism that prompted the experiment of bringing together the two great Aristotelian strands, that of St. Thomas and that of Marsilius, and yoking them to the defense of the Tudor state as Tudor ecclesiastical polity. If Hooker had been more observant and less wedded to Aristotle, he would have found another growing tradition of thought—Platonism—ready to hand.

Basically, Hooker was a Christian humanist, tolerant and fairly latitudinarian in theology. In the fifth book, which was devoted entirely to theology, he went out of his way to provide theological formulations that embraced to the point of ambiguity all the most controversial issues of the sixteenth century, so that as many disputants as possible would feel at home in his ecclesiastical polity. He was convinced that man was not wholly depraved and that the judicious exercise of human reason was absolutely essential to a Christian life. He saw no great and insurmountable chasm between nature and the supernatural and held that the mark of the Divine Creator can be detected in every creature.

Christian humanism had in a way been the mainstay of medieval Thomism. But in the sixteenth century, with Marsilio Ficino and Desiderius Erasmus, it had severed its connections with Aristotle and had been poured instead into the mold of Plato. Hooker was not only completely unaware of this revolution in thought; he actually went out of his way to attack one of the most popular Platonist teachers of his day, Peter Ramus. It is true that Ramus's variety of Platonism was a vulgar one and that one cannot blame Hooker for taking up cudgels against him. But viewed in perspective, Hooker's stubborn Aristotelianism acquired an unnecessarily aggressive edge when it was led into the fray against the Ramists, who were conspicuously active at Cambridge at that time. Against their nimble handling of Ramus's theories of rhetoric, Hooker reiterated all the old stock in trade of Aristotelianism and thought that he had vanquished his opponents simply by his demonstration that they differed from Aristotle. In this respect Hooker showed himself to be much more medieval than one is led to expect from his high baroque prose style and his freely discursive and informal way of arguing.

Through his conviction that Aristotelianism was the only satisfactory vehicle of Christian humanism, Hooker weakened his own case. For it was this conviction that deprived him of the opportunity of becoming the link between the humanism of John Colet, Erasmus, and Thomas More at the beginning of the sixteenth century and the Platonism of the Cambridge Platonists of the early seventeenth century. Platonism was fashionable enough in the England of Hooker: Edmund Spenser, William Harvey, Roger Ascham, Sir Philip Sidney were all Platonists in one way or another. But their Platonism was purely literary and emotional. Hooker was perhaps the only Elizabethan who could have deepened it. His Aristotelianism kept him aloof from these currents of thought, and thus he missed the unique opportunity that his great learning and the lucidity of his thought afforded him: injecting systematic philosophy into the Platonist current.

NATURAL LAW AFTER HOOKER

Although it may seem that Hooker's grand vindication of natural law helped to prepare the way for the revival of natural law in the seventeenth century, his arguments bear no relation to those of Hugo Grotius or of John Locke. To Hooker natural law was the dictate of reason; and reason was a discursive power of sensibility, capable of intuiting the good. It can therefore provide premises as well as help to draw out conclusions and dictate right conduct. To Grotius, on the other hand, the dictates of right reason were mere calculations of enlightened self-interest. In his theory of natural law, reason merely provided the long-term views necessary for survival, and natural law ceased to be identified with the rules set down, indirectly, by God. They were, on the contrary, made out to be completely independent of God.

See also Aristotelianism; Aristotle; Augustine, St.; Bellarmine, St. Robert; Calvin, John; Cambridge Platonists; Colet, John; Erasmus, Desiderius; Ficino, Marsilio; Grotius, Hugo; Harvey, William; Locke, John; Marsilius of Padua; More, Thomas; Natural Law; Plato; Ramus,

Peter; Social and Political Philosophy; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Wyclyf, John.

Bibliography

WORKS BY HOOKER

The best text for Hooker's works is J. Keble, ed., *Works,* 3 vols., in the 7th ed. reedited by R. W. Church and F. Paget (Oxford, 1888). The most easily accessible edition is that of *The Folger Library Edition of the Works of Richard Hooker,* 7 vols., edited by W. S. Hill (Cambridge, MA: Harvard University Press; Binghamton, NY: Medieval and Renaissance Texts and Studies, 1977–1998). See also *Book I, Laws of Ecclesiastical Polity,* edited by R. W. Church (Oxford, 1882); *Book V, Laws of Ecclesiastical Polity,* edited by R. Bayne (London, 1902); *Hooker's Ecclesiastical Polity: Book VIII,* edited with an introduction by R. A. Houk (New York, 1931); and *Of the Laws of Ecclesiastical Polity,* edited by Arthur McGrade (Cambridge, U.K.: Cambridge University Press, 1989).

WORKS ON HOOKER

Allison, C. F. *The Rise of Moralism: The Proclamation of the Gospel from Hooker to Baxter.* London: SPCK, 1966.

Atkinson, Nigel. *Richard Hooker and the Authority of Scripture, Reason and Tradition: Reformed Theologian of the Church of England?* Carlisle, U.K.: Paternoster Press, 1997.

Beiser, Frederick C. *The Sovereignty of Reason: The Defense of Rationality in the Early English Enlightenment.* Princeton, NJ: Princeton University Press, 1996.

Entréves, A. P. d'. *The Medieval Contribution to Political Thought.* London: Oxford University Press, 1939. Ch. 6.

Faulkner, Robert K. *Richard Hooker and the Politics of a Christian England.* Berkeley: University of California Press, 1981.

Hill, W. Speed, ed. *Studies in Richard Hooker.* Cleveland: Case Western Reserve University, 1972.

Kirby, W. J. Torrance, ed. *Richard Hooker and the English Reformation.* Dordrecht: Kluwer Academic, 2003.

McGrade, Arthur Stephen, ed. *Richard Hooker and the Construction of Christian Community.* Tempe, AZ: Medieval and Renaissance Texts and Studies, 1997.

Munz, Peter. *The Place of Hooker in the History of Thought.* London: Routledge and Kegan Paul, 1952.

Paget, Francis. *Introduction to the Fifth Book of Hooker's Treatise of the Laws of Ecclesiastical Polity.* Oxford: Clarendon Press, 1899.

Secor, Philip B. *Richard Hooker: Prophet of Anglicanism.* Turnbridge Wells, U.K.: Burns and Oates, 1999.

Shirley, F. J. *Richard Hooker and Contemporary Political Ideas.* London: SPCK/Church Historical Society, 1949.

Sisson, C. J. *The Judicious Marriage of Mr. Hooker and the Birth of the Laws of Ecclesiastical Polity.* Cambridge, U.K.: Cambridge University Press, 1940.

Peter Munz (1967)

Bibliography updated by Philip Reed (2005)

HORKHEIMER, MAX

(1894–1972)

Max Horkheimer, a German-American philosopher and social theorist, was born in Stuttgart, Germany, to a wealthy industrialist. After receiving a PhD in philosophy at the university of Frankfurt in 1922 with a dissertation on Kant supervised by Hans Cornelius, Horkheimer joined the *Institut für Sozialforschung* (Institute for Social Research) that was established in Frankfurt in 1923 as the first Marxist-oriented research center affiliated with a major German university. Under its director, Carl Grunberg, the institute's work in the 1920s tended to be empirical, historical, and oriented towards problems of the European working-class movement.

Horkheimer became director of the institute in 1930 and gathered around him many talented theorists, including Erich Fromm, Franz Neumann, Leo Lowenthal, Herbert Marcuse, and T. W. Adorno. Under Horkheimer, the institute sought to develop an interdisciplinary social theory that could serve as an instrument of social transformation. The work of this era was a synthesis of philosophy and social theory, combining sociology, psychology, cultural studies, and political economy.

During the 1930s, Horkheimer wrote many articles in philosophy, validating progressive ideals of reason, democracy, justice, morality, and other traditional concepts, while criticizing assaults on these ideals in the contemporary era and in particular developing critical perspectives on German fascism and its ideology. Most members of the Institute were both Jews and Marxist radicals and were forced to flee Germany after Hitler's ascendancy to power. The majority emigrated to the United States and the Institute became affiliated with Columbia University from 1931 until 1949, when it returned to Frankfurt.

From the mid-1930s, the Institute referred to its work as the "critical theory of society." For many years, "critical theory" stood as a code for the Institute's Marxism and was distinguished by its attempt to found a radical interdisciplinary social theory rooted in Hegelian-Marxian dialectics, historical materialism, and the critique of political economy. Members argued that Marx's concepts of the commodity, money, value, exchange, and fetishism characterized not only the capitalist economy but also social relations under capitalism, where human relations and all forms of life are governed by commodity and exchange relations and values.

In "Traditional and Critical Theory" (1937 [trans. 1972 in *Critical Theory*]), Horkheimer argued that "tradi-

tional theory" (which included modern philosophy and science since Descartes) tended to be overly abstract, objectivistic, and cut off from social practice. "Critical theory," by contrast, was grounded in social theory and (Marxian) political economy, carried out systematic critique of existing society, and allied itself with efforts to produce alternatives to capitalism and bourgeois society (then in its fascist stage in much of Europe). The goal of critical theory is to transform these social conditions, and help produce "an association of free people in which each has the same possibility of self-development" ("Traditional and critical theory," p. 219).

Working collaboratively with T. W. Adorno, their *Dialectic of Enlightenment* (1947 [1972]) sketched out a vision of history from the Greeks to the present that discussed how reason and enlightenment became their opposite, transforming what promised to be instruments of truth and liberation into tools of domination. Under the pressure of societal systems of domination, reason became instrumental, reducing human beings to things and objects and nature to numbers. While such modes of abstraction enabled science and technology to develop apace, it also produced societal reification and domination, culminating in the concentration camps that generated an instrumentalization of death. In the "dialectic of Enlightenment," reason thus turned instrumental, science and technology had created horrific tools of destruction and death, culture was commodified into products of a mass-produced culture industry, and democracy terminated into fascism, in which masses chose despotic and demagogic rulers. Moreover, in their extremely pessimistic vision, individuals were repressing their own bodies and renouncing their own desires as they assimilated and made their own repressive beliefs and allowed themselves to be instruments of alienated labor and war.

Sharply criticizing enlightenment scientism and rationalism, as well as systems of social domination, Adorno and Horkheimer implicitly implicated Marxism within the "dialectic of enlightenment" because it too affirmed the primacy of labor, instrumentalized reason in its scientism and celebration of "socialist production," and participated in Western modernity and the domination of nature. After the Second World War, Adorno, Horkheimer, and Pollock returned to Frankfurt to reestablish the institute in Germany, while Lowenthal, Marcuse and others remained in the United States.

In Germany, Adorno, Horkheimer, and their associates published a series of books and became a dominant intellectual current. At this time, the term "Frankfurt School" became widespread as a characterization of their version of interdisciplinary social research and of the particular social philosophy developed by Adorno, Horkheimer, and their associates. They engaged in frequent methodological and substantive debates with other theories, most notably "the positivism dispute," where they criticized empirical and quantitative approaches to social theory and defended their own more speculative and critical brand of thought. The German group around Adorno and Horkheimer was also increasingly hostile toward orthodox Marxism and were in turn criticized by a variety of types of "Marxism-Leninism" and "scientific Marxists" for their alleged surrender of revolutionary and scientific Marxian perspectives.

Horkheimer's *Eclipse of Reason* (1947) presents a popularized version of *Dialectic of Enlightenment* for an English-speaking audience and *Critique of Instrumental Reason* (1974) brings together Horkheimer's key essays since the end of World War II. The late Horkheimer became increasingly pessimistic and combined Schopenhauer's stoicism with a quest for the "totally other," a religious desire for transcendence that entered his materialist philosophy in later years.

See also Adorno, Theodor; Critical Theory.

Bibliography

WORKS BY HORKHEIMER

Twenty volumes of Horkheimer's *Gesammelte Schriften* are available in a Suhrkamp (Frankfurt) edition edited by Alfred Schmidt and Gunzelin Schmid Noerr. Important translations include:

Between Philosophy and Social Science. Cambridge, MA: MIT Press, 1993.

Critical Theory. New York: Herder and Herder, 1972.

Critique of Instrumental Reason. New York: Seabury, 1974.

Dialectic of Enlightenment.With T.W. Adorno. New York: Herder and Herder, 1972.

Eclipse of Reason [1947]. New York: Seabury, 1974.

WORKS ON HORKHEIMER

Benhabib, S., W. Bonss, and J. McCole, eds. *On Max Horkheimer: New Perspectives*. Cambridge, MA: MIT Press, 1993.

Wiggershaus, R. *The Frankfurt School*. Cambridge, U.K.: Polity Press, 1994.

Douglas Kellner (2005)

HOW AND WHY

See *Why*

HOWISON, GEORGE HOLMES
(1834–1916)

George Holmes Howison, the American personalist philosopher and mathematician, was a graduate of Marietta College in Ohio and professor of mathematics at Washington University, where he became a member of the St. Louis Philosophical Society. He taught philosophy at the Massachusetts Institute of Technology, at the Harvard Divinity School, and at the Concord School of Philosophy before moving in 1884 to the University of California, where he organized what was to become an influential department of philosophy.

Howison, calling his system "Personal Idealism," maintained that both impersonal, monistic idealism and materialism run contrary to the moral freedom experienced by persons. To deny the freedom to pursue the ideals of truth, beauty, and "benignant love" is to undermine every profound human venture, including science, morality, and philosophy. Thus, even Personalistic Idealism (B. P. Bowne and E. S. Brightman) and Realistic Personal Theism (Thomas Aquinas) are inadequate, for they make finite persons dependent for their existence upon an infinite Person and support this view by an unintelligible doctrine of *creatio ex nihilo.*

Howison's Personal Idealism, therefore, is founded on what he believed to be an undeniable fact: The freedom crucial to human existence is untenable if the individual is dependent for his existence upon any other being, including a Creator-God or an Absolute One. Therefore, self-determining beings must be uncreated and eternal; yet the unique quality of human freedom presupposes that each person stands in an individual relationship to other persons, subpersonal beings, and God.

How, then, does this plurality of uncreated beings compose a universe and not a mere collection of beings, a pluriverse? Howison answers that it is the very nature of undeniable, self-active, unified, thinking beings to define themselves and to fulfill themselves as individuals. In this very act of self-definition and self-fulfillment they find themselves related to other beings. "Thus, in thinking itself as eternally real, each spirit thinks the reality of other spirits."

Is there a God to unify the many grades of self-active beings? Yes, but any unification must not infringe upon individual growth to moral perfection. Creation as efficient cause must give way to creation in accordance with an Ideal present in each being. The fulfillment of this Ideal calls for a world composed of "*all* the individual differences compatible with the mutual reality of all." Thus, basic harmony is possible because, as each individual defines himself, he finds the Ideal of self-definition by which to measure himself. And God, who is "defined as self-existent by every other self-defining being," is the indispensable standard for measuring reality.

In this Personal Idealism there is, then, no one Prime Mover or Creator. Reality is a republic of self-active, self-defining spirits, each moving toward the Ideal exemplified by God, "changelessly attentive to every other mind, rationally sympathetic with all experiences, and bent on its spiritual success." Nor are the vast number and the gradation of minds that compose the different levels of matter, life, and mind the product of evolution; what we know as nature and evolution is the product of the various kinds of self-active beings, moved ultimately by the final causes of their inner beings toward a common goal.

See also Bowne, Borden Parker; Brightman, Edgar Sheffield; Idealism; Materialism; Personalism; Thomas Aquinas, St.

Bibliography

Buckingham, J. W., and G. W. Stratton. *George Holmes Howison, Philosopher and Teacher.* Berkeley: University of California Press, 1934. Includes a bibliography and complete reprint of *The Limits of Evolution and Other Essays Illustrating the Metaphysical Theory of Personal Idealism* (London: Macmillan, 1901).

Cunningham, G. W. *The Idealistic Argument in Recent British and American Philosophy.* New York: Century, 1933. Ch. 12.

Peter A. Bertocci (1967)

HSÜN TZU

See *Xunzi*

HUANG ZONGXI
(1610–1695)

Huang Zongxi, also known as Huang Lizhou, was the most important figure during the transitional period from the late Ming to the early Qing dynasty. He was the last in line as a Sung-Ming neo-Confucian philosopher, and was also an intellectual historian who studied in depth the whole Sung-Ming neo-Confucian Movement. Huang was the disciple of Liu Zongzhou (1578–1645),

and compiled the influential *Mingru xue'an* (Cases in Ming Confucianism) according to the guidelines he learned from his teacher.

The dominant trend of philosophy in the Ming dynasty was Wang Yangming's (1472–1529) *xinxue* (learning of mind). It was in sharp contrast to Zhu Xi's (1130–1200) *lixue* (learning of principle), which had been the dominant neo-Confucian philosophy as well as state ideology since the Yüan dynasty, because Zhu's *Commentaries to the Four Books* had been adopted as the basis for civil service examinations since 1313. Zhu had taught a dualism of *li* (principle) and *qi* (material or vital force); *xin* (mind-heart), for Zhu, consisted of the subtlest kind of *qi* that encompasses *li* (principles). Wang felt that Zhu's dualism was detrimental to self-discipline. Instead Wang taught a monism that identified *xin* with *li*.

Liu was in sympathy with Wang, but when Wang put too much emphasis on *liangzhi* (innate knowledge of the good), some of his followers claimed that sages are all over the street. In order to remedy the situation, Liu shifted the emphasis to *chengyi* (sincerity of the will) and *shendu* (vigilance in solitude). Huang inherited his teacher's monistic outlook, and went further, claiming that *li* is but the *li* of *qi*, and that there is no *benti* (substance) aside from *gongfu* (discipline). Such a tendency inadvertently led to a radical naturalistic interpretation of monism, which abandons the transcendent aspect of neo-Confucian philosophy altogether, thus causing a paradigm shift in early Qing philosophy.

Although Huang had firm convictions of his own, he chose not to write on his philosophy; instead, he worked hard to compile case studies. Because Wang Yangming taught different things in different places and periods, Huang took pains to study the different branches of philosophy under the school, devising a scheme to cover them all (although he did not neglect the other schools of philosophy). With its breadth and depth, Huang's *Mingrue xue'an* was unprecedented. It became so dominant, in fact, that when it was published it was taken as the only doorway through which one should study Ming Confucianism. Huang had also planned to provide case studies in Sung-Yüan Confucianism, but he never completed the task; the study was finally put together by Quan Zuwang (1705–1755).

Huang was also an expert on textual studies of the Classics. A case in point was his study of *Yijing* (Book of changes). He and his brother argued that the diagrams attached to this classic, which had been around since the Sung dynasty, were spurious. Huang's influence was contagious; Yan Rouju (1636–1704), who claimed Huang as his mentor, produced a critical study that showed the *Book of History* in ancient script was spurious. It is well known that Zhu Xi had established the orthodox line of transmission of the Way by quoting from the alleged fabricated document. With Yan's study, the foundation of Zhu's claim was now apparently undermined. Again, inadvertently, Huang appeared to have helped Qing Confucianism undergo a paradigm shift from philosophy to philology.

When the Ming dynasty was overthrown by the Manchus, Huang reflected deeply on politics and wrote the *Mingyi daifang lu* (Waiting for the dawn: A plan for the prince). He felt that since the establishment of the dynasties the rulers had taken the country as their private property, thus causing much misfortune. Huang urged a return to the ancient time when sage-emperors served the country and the people without selfish desires. (Although the *Mingyi daifang lu* has nothing to do with democracy in the West. Huang's book was used as propaganda against the Qing regime, inadvertently propelling the intellectuals to hope for a republican government of, for, and by the people.) The last dynasty was overthrown in 1912.

See also Chinese Philosophy: Confucianism.

Bibliography

Elman, Benjamin, A. *From Philosophy to Philology*. Cambridge, MA: Harvard University Press, 1984.

Huang Zongxi. *Mingru xue'an* [Cases in Ming Confucianism]. Selected translation in *The Records of Ming Scholars*, edited by Julia Ching. Honolulu: University of Hawaii Press, 1987.

Huang Zongxi. *Mingyi daifang lu*. Translated with an introduction by William Theodore de Bary as *Waiting for the Dawn: A Plan for the Prince*. New York: Columbia University Press, 1993.

Liu Shu-hsien. *Essentials of Contemporary Neo-Confucian Philosophy*. Westport, CT: Praeger, 2003.

Liu Shu-hsien. *Understanding Confucian Philosophy: Classical and Sung-Ming*. Westport, CT: Greenwood, 1998.

Shu-hsien Liu (2005)

HUET, PIERRE-DANIEL
(1630–1721)

Pierre-Daniel Huet, the last Christian skeptic in the line of Michel Eyquem de Montaigne and Pierre Charron, was born in Caen, Normandy. His father had been converted from Calvinism. Young Huet studied with the Jesuits and, after taking a degree in mathematics, went in 1652 with the Protestant scholar Samuel Bochart to the court of

Queen Christina of Sweden. There, he discovered a manuscript of Origen, which later led him to write *Origenis Commentaria in Sacrum Scripturam* (Rouen, 1668). En route home in 1653 he stopped in The Netherlands, where he met many savants. A discussion with one of them, Rabbi Menasseh (Manasseh) ben Israel, led Huet to write *Demonstratio Evangelica* (1679).

From The Netherlands Huet returned to Caen, where he founded the Academy of Sciences, corresponded with learned men throughout the world, and worked on his studies on Origen. He often traveled to Paris and entered several of the learned literary salons. His reputation as a man of letters and science grew, and in 1670 Louis XIV appointed him to be Jacques Bénigne Bossuet's assistant as the dauphin's teacher. While holding this post, Huet started a famous set of editions of classical authors, *Ad Usum Delphini*.

After several years Huet decided to become a priest and was appointed abbot of Aunay and afterward bishop of Soissons. He did not like that post and exchanged it for the bishopric of Avranches. In 1699 he retired to a Jesuit institution in Paris, to which he had donated his enormous library (transferred after the suppression of the Jesuit order to the Bibliothèque Nationale, where it constitutes a basic part of the collection).

Huet wrote many works on history, philosophy, theology, and literature and was regarded by figures like Gottfried Wilhelm Leibniz as the most learned man of his age and as an excellent Latin poet. His most philosophically interesting works are the *Demonstratio Evangelica*, *Censura Philosophiae Cartesianae* (1689), *Nouveaux Mémoires pour servir à l'Histoire du cartésianisme* (1692), *Questiones Alnetae de Concordia Rationis et Fidei* (1692), and its notorious concluding section, the *Traité philosophique de la foiblesse de l'esprit humain*, published posthumously in 1723.

HUET'S SKEPTICISM

Demonstratio Evangelica shows signs of philosophical skepticism and empirical and liberal views. After arguing that no absolute certainty could be attained in mathematics or theology, Huet tries to establish religious truth inductively, by showing the common elements in all religions, ancient and modern. The privileged position of Christianity was primarily because of its expressing best the features of natural revelation. (Doctrinal differences within Christianity had little interest for Huet. Hence, he could join his friend Leibniz in trying to reunite all the churches.)

In Huet's papers there is some material that indicates the special flavor of his skepticism and religious approach. While he was bishop of Avranches, somebody noticed that a Jesuit from Normandy had received a doctorate for a dissertation claiming that there is no evidence that Christianity is true and that of all the religions in the world Christianity is the least probable. This raised a scandal, and the case was turned over to Huet to examine the Jesuit. Huet sent back a report saying that he agreed on everything. Since Christianity is a matter of faith, there should be no evidence and if it were at all probable, that would count as evidence. Further material about this may be found in the massive collection of his papers in the Medici library in Florence, Italy.

Huet's writings against Cartesianism show a much more developed epistemological skepticism. He utilizes all of Sextus Empiricus's weapons to attack René Descartes's claims that the *cogito* is the fundamental, indubitable truth and that whatever is clearly and distinctly conceived is true. Joining the previous critics Pierre Gassendi, Thomas Hobbes, and Simon Foucher, Huet, in an intensive examination of the Cartesian theory of knowledge, contends that "I think, therefore I am" is a dubious claim and that no certain knowledge about the world could be attained by Descartes's "way of ideas." In *Censura* and in an unpublished defense of it Huet argues not only that "I think, therefore I am" is an inference but also that it involves a time sequence from the moment when thinking is occurring to the moment when one realizes that one thought and that memory may be inaccurate. If one is immediately conscious of thinking, the realization about the existence is a possible future event. Hence, one cannot be simultaneously aware and certain of the ingredients of the cogito and, thus, of its indubitability.

Besides analyzing the Cartesian arguments, Huet ridicules both the theory and its founder. *Nouveaux Mémoires* is a spoof about Descartes's life after his supposed death in Stockholm, Sweden, in which Descartes tries to expound his philosophy in Lapland. Huet also joins the Jesuit anti-Cartesians in accusing Cartesianism of irreligion and incoherence, advocating, instead, a type of probabilistic nonmetaphysical view of the world.

THE 'TRAITÉ PHILOSOPHIQUE'

The full presentation of Huet's skepticism appears in the posthumous *Traité philosophique*, which the Jesuits denounced as a forgery written to embarrass the church. (The manuscript, which is in Huet's handwriting, and discussions in his correspondence eliminate any doubts

about Huet's authorship.) The traditional Pyrrhonian position is set forth, criticisms of skepticism are considered and refuted, and a modern skepticism are advocated in opposition to Cartesianism. Huet's skepticism consists of doubting that any genuine knowledge about reality can be attained by human means while offering experimental science and pure fideism as the means for finding out something about nature, God, and man. In *Traité philosophique*, in his correspondence, and in his marginalia, especially in his copy of Blaise Pascal's *Pensées*, an extreme fideism appears, in which it is denied that there can be any rational defense of religion. Huet thought Pascal too rationalistic because of his wager argument. Faith, and faith alone, could lead to any religious views. It is difficult to determine what or how much Huet himself actually believed. As a prelate and theologian, he was extremely latitudinarian and was in friendly contact with scholars everywhere, regardless of their religious or nonreligious affiliations.

Huet's *Traité philosophique*, first published in 1723, was quickly translated into English, Italian, German, and Latin and was studied throughout the eighteenth century. David Hume read it and, like many others, was amused that the author was a total skeptic and a learned clergyman. Huet's contribution to skeptical discussion of the period underlines his explanation of how the skeptic can deal with normal human situations. Huet states, "It is one thing to philosophize, another to live." He then points out that the skeptic, like everybody else, lives according to customs and habits while at the same time doubting that there can be any justification. At the end of book 1, part 4 of *A Treatise of Human Nature* (1737), Hume gives pretty much the same explanation as Huet. Hume also cites Huet in the *Dialogues on Natural Religion*, posthumously published in 1779.

In his day Huet was influential and was taken seriously by Leibniz, Pierre Bayle, and others (Benedict [Baruch] de Spinoza even feared that Huet was writing a refutation of his views). A major transitional figure, he helped to destroy Cartesianism and to further empirical science. His immense erudition provided some of the basic materials for the Enlightenment. His pioneering work in comparative religion was taken by later scholars and was used as ammunition against traditional religion. However, his skeptical argumentation was taken less seriously than that of Bayle or Hume.

Recent studies suggest that Huet had an overall theological and philosophical perspective that would be united in the total corpus of his works. So far, there are only piecemeal pictures of it, which show him to be an important scholar of the early Enlightenment. There is still an enormous amount of unpublished material in his correspondence, his markings and notes on books in his library, and in manuscripts of works not completed.

See also Bayle, Pierre; Bossuet, Jacques Bénigne; Cartesianism; Charron, Pierre; Descartes, René; Enlightenment; Foucher, Simon; French Philosophy; Gassendi, Pierre; Hobbes, Thomas; Hume, David; Leibniz, Gottfried Wilhelm; Menasseh (Manasseh) ben Israel; Montaigne, Michel Eyquem de; Origen; Pascal, Blaise; Sextus Empiricus; Skepticism; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY HUET

Demonstratio Evangelica. Paris: Apud Stephanum Michallet, 1679.

A Philosophical Treatise concerning the Weakness of the Human Understanding. London: Printed for G. Dommer, 1725. Originally published as *Traité philosophique de la foiblesse de l'esprit humain* (Amsterdam: Henri du sauzet, 1723).

Against Cartesian Philosophy. Translated by Thomas M. Lennon. Amherst, NY: Humanity Books, 2003. Originally published as *Censura Philosophiae Cartesianae* (Paris: Apud Joannem Anisson, 1689).

WORKS ABOUT HUET

Barthomèss, Christian. *Huet, évêque d'Avranches, ou, Le scepticisme théologique*. Paris: Imprimerie de Marc Ducloux et Comp., 1850.

Dupront, A. *Pierre-Daniel Huet et l'exégèse comparatiste au XVIIe siècle*. Paris: Librairie E. Leroux, 1930.

Lennon, Thomas M. "Huet, Descartes, and the Objection of Objections." In *Skepticism in Renaissance and Post-Renaissance Thought: New Interpretations*, edited by José R. Maia Neto and Richard H. Popkin. Amherst, NY: Humanity Books, 2004.

Maia Neto, José, and Richard H. Popkin. "Bishop Pierre-Daniel Huet's Remarks on Pascal." *British Journal of the History of Philosophy* 3 (1) (1995): 147–160.

Popkin, Richard H. "Bishop Pierre-Daniel Huet's Remarks on Malebranche." In *Nicolas Malebranche: His Philosophical Critics and Successors*, edited by Stuart Brown, 10–21. Assen, Netherlands: Van Gorcum, 1991.

Popkin, Richard H. "Pierre Bayle and Bishop Huet, the Master Sceptics." In *The Columbia History of Western Philosophy*, edited by Richard H. Popkin, 404–411. New York: Columbia University Press, 1999.

Popkin, Richard H. "The New Sceptics: Simon Foucher and Pierre-Daniel Huet." In *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Popkin, Richard H. "The Manuscript Papers of Bishop Pierre-Daniel Huet." In *Year Book of the American Philosophical Society*, 449–453. Philadelphia, PA: American Philosophical Society, 1959.

Rapetti, Elena. *Percorsi anticartesiani nelle lettere a Pierre-Daniel Huet*. Firenze, Italy: L. S. Olschki, 2003.
Tolmer, Léon. *Pierre-Daniel Huet, 1630–1721: Humaniste-physicien*. Bayeux, France: Colas, 1949.

Richard H. Popkin (1967, 2005)

HÜGEL, BARON FRIEDRICH VON

(1852–1925)

Baron Friedrich von Hügel, the Roman Catholic philosopher of religion and writer on mysticism, was born in Florence, Italy, and succeeded to his father's (Austrian) title in 1870. Most of his life was spent in England. His most important writings were *The Mystical Element of Religion as Studied in St. Catherine of Genoa and Her Friends* (London, 1908), *Essays and Addresses on the Philosophy of Religion* (London, 1921 and 1926), and *The Reality of God* (published posthumously; London, 1931).

Von Hügel's philosophical position was opposed both to idealism and to what he called positivism. By positivism he meant the doctrine that knowledge is exclusively confined to sense perceptions and to the laws that connect them. He rejected this position on the grounds that sense experience is accompanied by a strong "pressure on our minds" to credit it with "trans-subjective validity" (that is, to accept that it tells us something about an external world existing independently of our experience of it) and that refusal to assent to this pressure would mean that positivism collapses into skepticism, which is self-defeating. Moreover, since it is our own minds that we are immediately aware of, our apprehension of reality will be more certain if there is no phenomenal content. This idea paves the way for von Hügel's justification of the epistemological importance of mystical experience. He criticized idealism for a subjectivism similar to that implicit in positivism.

Von Hügel distinguished between knowledge of abstract ideas and of numerical and spatial relations, on the one hand, and knowledge of real existences on the other. The former is clear and readily intelligible; the latter is never totally clear, since any statement or set of statements about a real object will fail to exhaust what is to be discovered in it. The "higher," or more complex, the entity, the less clear is one's apprehension of it. Von Hügel was therefore concerned with opposing philosophical theories that claimed to give a clear and exhaustive analysis of types of existence (for example, he criticized David Hume's account of the individual). Reality, according to von Hügel, is indefinitely apprehensible, a fact that serves to explain both the revisionism of science and the gropings of religion. The obscurity involved in religion is an index of the richness of its subject matter. "Religion," he said, "can't be clear if it is worth anything."

The concept of the Infinite occupies a central position in von Hügel's philosophy. He held that there was no good reason for neglecting or doubting the validity of man's sense of the Infinite, which should be taken quite as seriously as sense experience; in this, he in effect conjoined a critique of religious experience and traditional Catholic natural theology. The critique of religious experience involved the examination of the claims of great religious figures of all ages. He was opposed to simply accepting the testimony of the individual; rather, he pointed to the errors and excesses of many individual interpretations of religion, some of which involved the denial of plain facts. At the same time, he was sympathetic to the insights claimed by non-Christian religions. His doctrine of religious knowledge was not exclusive to any one tradition, although he was opposed to relativism. Von Hügel also argued against various theories arising from religious experience, such as the extreme dualism of Søren Kierkegaard and the monism of some mystics (for instance, the doctrine of the identification of the soul with God). He maintained that Kierkegaard differentiated to such an extent between God and man that intercourse between the two became incomprehensible. A crucial argument of von Hügel's was that the impinging of the Infinite on man's experience, emotions, and will implies the spiritual nature of the Infinite; for otherwise it would be hard to account for its inspiring power.

According to von Hügel's theory of knowledge, it is artificial to divorce the cognitive aspects of experience from the affective and volitional aspects; and therefore the religious apprehension of the Infinite is not limited to grasping a theoretical concept but includes a vital response. For this and other reasons, von Hügel defended a sacramental and institutional faith—namely, that of Catholicism. But in his openness to and sympathy with the critical evaluation of the biblical tradition by the methods of scientific history, von Hügel belonged to the Catholic Modernist movement.

See also Hume, David; Idealism; Infinite, The; Kierkegaard, Søren Aabye; Modernism; Positivism.

Bibliography

Additional works by von Hügel are *Eternal Life* (Edinburgh: Clark, 1912); *The German Soul* (London: Dent, 1916);

Readings from von Hügel, edited by Algar Thorold (London: Dent, 1928); and *Letters to a Niece* (New York: Dutton, 1950).

A work on von Hügel is L. V. Lester-Garland, *The Religious Philosophy of Baron Friedrich von Hügel* (London, 1938).

Ninian Smart (1967)

HUGH OF SAINT VICTOR

See *Saint Victor, School of*

HUI SHI

(370–318 BCE)

Hui Shi (also Hui Shih) was an ancient Chinese logician and a major figure of the School of Names (*Ming-Jia*) whose members are also known as dialecticians or sophists for the sake of their emphasis on rational argumentations and their focus on deep structures of concepts. Hui Shi's philosophical thoughts are primarily delivered in his ten seemingly paradoxical propositions as recorded in the part "*Tian-Xia*" (ch. 33) of the *Zhuang-Zi*; these ten propositions are given as follows (with this author's brief explanation in brackets attached to each).

(1) "The greatest dimension [of the universe] has nothing beyond itself and is thus called 'the great unity,' while the smallest dimension [of the universe] has nothing within itself and is thus called 'the small unity.'" [The universe as a whole unity is both the greatest and the smallest in infinity; and the greatest and the smallest are intrinsically connected.]

(2) "That which has no thickness cannot be increased in thickness, and yet in extent it covers one thousand li [miles]." [This is one way to illustrate the point that being (extension) and non-being (non-thickness) come from each other.]

(3) "The heaven is as low as the earth; mountains are on the same level as marshes." [The high and the low in nature are not absolute but relative.]

(4) "The moment the sun reaches the zenith at noon, it is declining; the moment the creature is born, it is dying." [This characterizes the two features of changing and becoming process in nature: things will develop in the opposite direction when they become extreme; being and non-being interpenetrate each other.]

(5) "A great similarity differs from a little similarity; this is called 'the little similarity-and-difference.' All things are both similar/identical to one another and different from one another; this is called 'the great similarity/identity-and-difference.'" [Things have not only their more or less similar or identical aspects but also their distinct aspects that distinguish one from another.]

(6) "The South has no limit and has a limit." [Some things have both their finite aspects and their infinite aspects at the same time. For example, the South as a location has its limit in space but has no limit in regard to, say, its development in time.]

(7) "One goes to the State of Yüe today and arrives there yesterday." [This highlights temporal relativity.]

(8) "Connected rings can be in separation." [Connected rings themselves are separated from each other in regard to the identity of each ring; each ring is at the same in connection with and separation from the other rings. The point is that seemingly opposed and unrelated states or processes can be possessed by the same thing and thus be interpenetrating and complementary.]

(9) "I know where the center of the world is; it is in the north of the State of Yan and the south of the State of Yüe." [This stresses spatial relativity.]

(10) "Extend love to all things; Heaven and Earth are the one unity." [The fundamental unification-character of all things in the universe constitutes the metaphysical foundation for extending love to all things.]

Hui Shi puts more emphasis on common aspects, connections, and unification of things in the universe (as highlighted in propositions 1 and 10) and relativity of their distinctions (as illustrated in propositions 2, 3, 4, 6, 7 and 9). In contrast, Gongsun Long, another major figure of the School of Names, stresses distinct aspects of things. Nevertheless, though the two thinkers appear to have different orientations, their difference is rather in emphasis. Hui Shi also pays attention to, or even stresses, distinct aspects, as suggested in propositions 5 and 8. Indeed, one central point suggested in Hui Shi's ten propositions is that many seemingly paradoxical or opposed contraries turn out to be interdependent, interpenetrating, and complementary. This essentially reflects the crucial point of the fundamental *Yin-Yang* way of thinking in view of cosmological and ontological characters of the universe. Moreover, as suggested by the points of all the ten propositions, Hui Shi, as "logician," is primarily concerned with metaphysical foundation of logical discourse rather than with its purely formal character.

Bibliography

Fung, Yu-lan (Feng, Yu-lan). *A History of Chinese Philosophy*, 2 vols., translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1952–1953.

Graham, Angus Charles. "Three Studies on Gongsun Lung." In *Studies in Chinese Philosophy and Philosophical Literature*. Albany: State University of New York Press, 1990.

Hansen, Chad. *Language and Thought in Ancient China*. Ann Arbor: University of Michigan Press, 1983.

Hu Shi. *The Development of the Logical Method in Ancient China*, 2nd ed. New York: Paragon Books, 1969.

Bo Mou (2005)

HUMAN GENOME PROJECT

It was the Moravian monk Gregor Mendel, working in his monastery's garden in the 1860s, who cracked the secrets of heredity, but it was not until the beginning of the twentieth century that his insights were recognized and developed. Thanks particularly to the work of Thomas Hunt Morgan and his associates, working at Columbia University in New York, it was learned that the basic units of heredity, the genes, lie along thin, paired strings (chromosomes), in the centers of cells, and that these are not only the units of function—the things that carry the information used in building the finished organism—but also the units of heredity—the things passed on from one generation to the next.

In 1953, working in Cambridge, England, James Watson and Francis Crick confirmed the growing suspicion that the genes are long macromolecules of deoxyribonucleic acid (DNA), and famously they showed that the genes themselves consist of paired strings, twisted together in a double helix. The information carried by the genes comes not so much from the content of the DNA itself—a string or chain of four basic submolecules (nucleotides or bases)—but rather in the order of these submolecules along the chain. Information is read off from the DNA by another nucleic acid (RNA), and then this is used to pick up amino acids within the cell, which are then in turn strung together to make polypeptide chains, the building blocks (proteins) of new cells. Because there are twenty different amino acids used by the body, and because there are four different molecules in the DNA chain, in order take information from the DNA, there had to be (at a minimum) at least three nucleotides used in each transfer of information to catch without ambiguity a particular amino acid. Cracking this genetic code was the second great triumph of molecular biology of the 1950s.

With the basic theory now in place, biologists could turn their attentions to the discovering of the particular genes in particular organisms, a task made much easier, in the 1970s, by the discovery of powerful tools (recombinant DNA techniques) for dissecting the parts of the genome—the totality of a particular organism's genetic components. Thus was made realistically possible the idea of mapping the complete genetic content (the whole genome) of any particular animal or plant. Naturally, because people are humans, the idea of mapping human genes came to the fore, and thus the Human Genome Project (HGP) was conceived. Visionary credit is usually given to the biologist Robert Sinsheimer, who convened a crucial meeting of pertinent molecular scientists in California in 1985 to discuss and explore the feasibility of the project.

A task such as this is incredibly expensive, and governments of the world were soon asked for support and became involved both financially and organizationally. In the United States, the Department of Energy (DOE) was an early backer, a somewhat strange connection explained by the Department's involvement in the genetic effects of radiation on humans—a connection dating back to the Japanese atomic bombs and the testing of the subsequent Cold War. Less surprisingly, the National Institutes of Health (NIH) became involved and soon took the lead. By 1988, the codiscoverer of the structure of DNA, James Watson, had been appointed head of the project. (Watson resigned in 1992 and his post was then taken by the molecular biologist Francis Collins.)

The approach taken was basically one of brute force, squeezing out the information from the genome step by step, base by base. However, work in the 1990s was transformed not just by ever more powerful and rapid methods of getting results, but by the researcher J. Craig Venter. He argued for different strategies, initially aiming just to skim the genome for the really interesting results and leaving the rest until later, and then leaving governmentally supported institutions and going after private money, aiming to use even more powerful techniques to beat the NIH at its own game. In the end, both sides, public and private, announced a first draft of the human genome in 2000, and in 2003 a more detailed map was produced. The HGP was completed. Over 99 percent of the human genome has been mapped ("sequenced"), to an accuracy of 99.99 percent. Functionally speaking, there are about thirty thousand genes (a lot fewer than many would have estimated before the project began),

and a huge amount of the genome seems to be nonfunctional (it is estimated that "junk" DNA may be over 90%). The cost of this enterprise, factored in at 1991 U.S. dollars (when the estimates and projections were being done), was $2.7 billion. The genome sequenced was not taken from one individual person, but was a mixture of different people for the different chromosomes.

Focusing now on philosophical questions arising from the HGP, there are those more epistemological and those more ethical and social. This entry will take them in turn.

WORTHWHILE SCIENCE?

The big epistemological question that has haunted the HGP is whether it should have been done in the first place. As it was being discussed, many top-notch molecular biologists argued that it was unneeded and would divert resources and attentions from far more worthwhile projects. They felt that the project was motivated mainly by biologists who, feeling insecure around physical scientists, wanted their own equivalent of the moon-landing project—and it would have about as much scientific value.

Concerns such as this keep reappearing. Simply given a list of bases seems to have little or no scientific merit. The philosopher Alex Rosenberg gives a memorable metaphor. Imagine two stacks of phone books twisted around each other in a helix and reaching a mile and a half into the sky. The covers have been removed so each book leads at once to the next; the names have been removed so there are only lists of numbers—not in columns but one after another with no punctuation; area codes are assigned here, there, and everywhere at random, and no one has any idea what code corresponds to what geographical region. The only certainty is that 90 percent or more of the codes are fictitious and have no region to which they refer.

What, asks Rosenberg, is the point of having information on something such as this? Why bother to go to the effort of feeding all of these numbers into a computer? No one in their right mind would think it worth the effort. And yet this is equivalent to just what the HGP set out to do. And now that there is a string of numbers, so what? It is not that sequencing as such is worthless, but that unless there is some idea of what is being sequencing and why, it is misconceived effort. Far better to concentrate on specific problems—finding particular genes of interest and tracing their effects—than simply scooping up everything in one massive project.

Not surprisingly, from the beginning many biologists have disagreed. The junk DNA is a particularly sore issue because it does seem as if huge effort is being given over to mapping something—the vast portion of the genome—that is without function or purpose. The biologist and ethicist Frederick Grinnell argues that an attitude such as Rosenberg's is intellectual ludditism. It is precisely because scientists do not now see any function behind junk DNA that it should be explored and uncovered. Only then can its true nature be seen, and perhaps in fact they will find that far from being without purpose, perhaps junk DNA plays a crucial role in the living being.

Above the level of detail however—and more importantly—is that the project has never been simply one of listing bases on a line. From the start, biologists have been using the data to locate and to identify particular genes of interest, and to explore their functioning. As a tool, the results of the HGP have been made far more powerful because, from the beginning, the human genome was not the only genome being sequenced. Many others—bacteria such as *E. coli*, insects such as the fruit fly Drosophila, and mammals such as the mouse, as well as yeasts and plants—have been studied and mapped. The comparative results to which these have led have given rise to some of the most exciting and forward-looking branches of pure science active today. Particularly noteworthy is the field of so-called evolutionary development ("evo-devo"), where comparisons between gene sequences are a vital component of understanding and have led to insights as surprisingly unexpected as they are of far-reaching consequence. It is now known, for instance, that almost identical sequences of genes are to be found in organisms as separate as fruit flies and humans, pointing not only to shared ancestry but also to the fact that even today the ways in which these different organisms develop are identical, and for the same reasons. Results such as these are calling for significant rethinking of the workings of evolution. As Charles Darwin himself always suspected, so much change is a matter of making do with what one has rather than regrouping and starting again to go for an ideal solution.

No doubt more results will come in future years. While one may probably legitimately question the motives of all of the early backers of the HGP—the DOE certainly seemed to be looking for a project to justify its existence—it would be incorrect to say that no good science has emerged and churlish to say that the project has set molecular biology entirely down a misguided path.

EUGENICS

Even more than its potential scientific worth, the HGP was promoted as something that would have major social and ethical virtues, particularly in areas of medical care. As one moves toward these topics, what is first encountered is a cluster of issues around the topic of eugenics—however it may be called and described. Traditionally, "eugenics" referred to programs designed to improve the human race by interfering with or modifying its genetic constitutions. This could be thought of in two ways—positively, trying actively to perfect human genomes; and negatively, trying simply to eliminate the worst effects of the genes. The whole program fell into great disrepute because of the horrific activities of the Nazis, and today eugenics tends to be one of the topics from which people of all persuasions flee. Naturally, any suggestion that the HGP may in some way be an excuse for a new eugenical program is something that fills people with horror. It is something from which the promoters of the program have been at pains to divorce themselves.

Prima facie, one can see why there are worries. If the aim is simply to list the human genes, then the next move could well be trying to produce designer babies of some kind. But before one rushes to this as an inevitable conclusion, it is worth stopping and asking whether eugenics is necessarily always a bad thing. In particular, one may well reject positive eugenics—one may deplore the attempt to produce a race of superhumans—and yet endorse negative eugenics—attempting to eliminate horrendous genetic diseases. In fact, by another name, namely "genetic counseling," negative eugenics has never stopped. Paradoxically, although it was Jews who were the focus of the Nazi race laws, it is Jews who have been at the fore of genetic counseling as they try to discover carriers of the gene for Tay-Sachs disease, and through selective abortion prevent children being born with the affliction.

In this sense, therefore, one can see the completion of the HGP as a powerful tool in this direction, but at the same time argue that it is of positive moral and social worth rather than the other. Of course, one can see how this could be the thin end of the wedge to more draconian—perhaps even state-enforced—policies designed to eliminate even minor or idiosyncratic traits, or to eliminate those who may carry traits that could in certain circumstances prove wrong or disliked by some others in society. Perhaps Tay-Sachs disease yesterday, alcoholism (if it proves to have a genetic link) today, and homosexuality (again, if there is a link) tomorrow. This is clearly a danger, although whether the HGP as such should be faulted or whether any kind of genetic approach to humankind is more truly to blame is another matter.

In fact, one may argue that the HGP is a good thing in this respect, precisely because it takes the focus from individual genes and shows that the human being is a conglomeration of genetic factors. That (as the late Stephen Jay Gould pointed out) blunt "reductionistic" approaches to human nature—where each feature is supposedly controlled by one gene—are highly simplistic. Thirty thousand genes are not many to produce a complex entity such as a human being. It is clear that much about us is a product not just of one gene working in splendid isolation, but rather of many genes interacting to give rise to many complex traits. Hence, one starts to realize that, even with the best of intentions, thoughts of eliminating all genetic diseases are probably unrealistic. As one eliminates one gene, one affects the workings of others, which may in itself have negative effects. There is no simple route to perfection, and the HGP underlines this fact.

In any case, as has also been hinted at by the reference to evo-devo, the HGP is part of an ever-increasing awareness brought on by molecular biology. Simply thinking of genes leading to completed features is as naive as unqualified reductionism. There is no such thing as unadulterated genetic determinism, where people are the product of and only of their genes. It is always the genes in interaction with the environment, whether this be the physical environment or culture or whatever. Rather than argue that the HGP will at once lead to programs for genetic elimination, one could argue that the HGP will lead to programs to manipulate environments so that deleterious genes will not have full effect. Suppose that alcoholism is a function of the genes in some way. Rather than eliminating the genes and their carriers, one may rather try for programs that help people to learn to live adequately and happily with their biologically informed natures.

The point is that, as always in dealing with moral and social issues, scientific findings are never definitive or the last word. How one deals with these findings—including the findings of the HGP—is going to be a function of many things. While new knowledge can be dangerous, it is unfair to think that it will always be disastrous.

PARIAHS?

This leads at once to what is already an immediate and pressing concern. Now that there is the information from the HGP in the public domain, what is to stop institutions—public and private—from using this information

for their own ends, ends that may not necessarily be the ends of individuals? Take most obviously the question of insurance. The idea behind insurance is that people bet on ignorance, with a group of other people, to protect themselves if things go wrong. People buy health insurance knowing that 10 percent of the population will need a procedure; each individual in the group hopes he or she will not need the procedure, but knows that if so it will have cost (approximately) 10 percent of what it would have cost without the insurance. If people know they do not need the procedure, they will not buy the insurance—no man is going to buy insurance for gynecological problems. But, if a person is known to need the procedure, no one will sell that person the insurance—no one will sell cancer protection to someone with leukemia.

The worry is that the HGP is going to render some people pariahs and ineligible for insurance. Even worse, these will not necessarily be people who are or ever become physically sick. Thanks to the HGP, scientists are increasingly discovering genes that lead to various ailments or the predispositions to such ailments. This means that individuals can be tested against such standards to see if they are or are not possessors of the pertinent genes. Suppose that the knowledge obtained from the project enables a test for the disposition to a certain ailment or habit, alcoholism for instance. Normally, one may think that alcoholics should be discriminated against because of their habits; people with drunk driving convictions should pay more in car insurance. But is it fair to discriminate against the person who has a gene that simply sets up a predisposition? Does the state have the right to demand of a private insurance company that it not test for such a gene? Is this not unfair to others who buy insurance from that company? Do they not have the right to demand that their company obtain all pertinent information? The company surely has the right—the obligation—to discover if their would-be clients have drunk driving convictions.

The general problem has led some, notably the philosopher Philip Kitcher, to argue that the only viable solution to the new knowledge pouring forth from the HGP is some kind of state-supported, universal health care. Private insurers are simply going to be unwilling and unable to carry on with such plans unless there are massive, state-enforced rules preventing the obtaining of pertinent knowledge—which in itself is not necessarily a good thing (because, apart from anything else, one may need such knowledge), and probably not enforceable anyway. Society generally frowns upon fetal sex-determination, but that is little barrier for those determined to find and act on such information.

Even if Kitcher's proposal is followed through—after all, virtually every sophisticated society other than the United States already has such health care—this does not solve other related problems, such as life insurance. The moral ambiguity of such issues and the difficulties of working toward satisfactory solutions presents two notions of prejudice. On the one hand is the social prejudice that will be shown against those unable to buy such insurance. On the other hand is the potential for the prejudice abhorred by insurers, namely that society should not discriminate unfairly against the fortunate for the benefit of the unfortunate. A business deal is a business deal, not a policy of social welfare.

ASSESSMENT

One point is clear. The HGP is not some isolated phenomenon; it is part of a general move toward the understanding of organisms at the genetic level, and of technological applications that stem from such understanding. It is hardly a philosophical argument to point out that no amount of bemoaning the fact is going to stop this process. But philosophical argument does show that although there are great dangers and problems, there are also exciting opportunities and prospects, theoretical and practical. The philosopher's aim must be to guide in the right direction the bounties that stem from this new knowledge and its powers.

See also Philosophy of Biology; Science Policy.

Bibliography

FURTHER READING

The Internet is the place to start with the quest for information. Organizations such as the National Institutes of Health have detailed and clear web pages and links to other pertinent sites. Two books that capture the excitement of the new genetics are by the science writer Matt Ridley, *Genome* (New York: HarperCollins, 2000), and *Nature via Nurture* (New York: HarperCollins, 2003). A collection of philosophical articles on the HGP is included in *The Philosophy of Biology*, edited by David L. Hull and Michael Ruse (Oxford: Oxford University Press, 1998). This includes Rosenberg's article with his analogy. Frederick Grinnell replies in "Philosophy of Biology and the Human Genome Project" (*Biology and Philosophy* 15 [2000]: 595–601). A good starting guide to some of the social and moral issues is by Philip Kitcher, *The Lives to Come: The Genetic Revolution and Human Possibilities* (New York: Simon and Schuster, 1996).

Michael Ruse (2005)
Lucyle T. Werkmeister (2005)

HUMANISM

Humanism in the present era signifies an ideological doctrine that places human beings, as opposed to God, at the center of the universe. Although a focus on human nature and human life can be traced back ultimately to ancient Greek thought, humanism in the modern sense, with its anthropocentric belief in the boundless potentiality of unfettered human reason and its secular conviction that human destiny is entirely in human hands, has its roots in the Enlightenment of the eighteenth century. This philosophical orientation should not be confused with the intellectual movement known as Renaissance humanism. Unlike its contemporary namesake, Renaissance humanism was not specifically concerned with promoting and exalting human values. It was, instead, a hugely influential cultural and educational program dedicated to the revival of the classical ideal of cultivated and civilized learning, referred to in Latin as *humanitas* and in Greek as *paideia*. For humanists of the Renaissance and their successors, the only way to achieve this ideal was through the *studia humanitatis*, the study of Graeco-Roman civilization through its literature, history, philosophy, and surviving artifacts. The zeal for recovering and reviving antiquity reached its height from 1300 to 1650. Recent scholarship has, however, highlighted earlier periods in which brief bursts of enthusiasm for ancient learning can be identified.

HUMANISM IN THE MIDDLE AGES

An intensified interest in the classical legacy, leading to a general cultural revival, occurred in the Islamic world during the tenth century. In contrast to Western Europe, where the Latin heritage was always supreme, Arab scholars, both Christian and Muslim, were concerned exclusively with Greek erudition. Moreover, their interest was entirely in the scientific and philosophical patrimony of ancient Greece, leaving aside its literary and historical works. During this period a large number of texts by Plato, Aristotle, Euclid, Ptolemy, Hippocrates, and Galen were translated into Arabic, sometimes via Syriac intermediaries. The philological efforts that went into this enterprise bore fruit in the achievements of thinkers such as Avicenna (980–1037) who were able to build up their own philosophical, scientific, and medical systems by drawing on the Greek material newly available to them.

Two epochs during the Western European Middle Ages witnessed revivals of ancient learning that have been seen as foreshadowing the humanism of the Renaissance. The first, associated with the reign of Charlemagne, occurred in the eighth and ninth centuries. In this period a small group of scholars, most notably Lupus of Ferrières (c. 805–862), studied, edited, and copied texts by Cicero, Valerius Maximus, and Aulus Gellius, among other Latin authors. It was partly thanks to their philological interests and skills that these works survived and were transmitted to later generations. Another Carolingian intellectual, John Scotus Eriugena (c. 810–c. 877), used his knowledge of Greek, a rare accomplishment in the Middle Ages, to gain access to Neoplatonic sources, which played a significant role in his highly original philosophical and theological writings.

The second period of medieval humanist activity took place in the twelfth century. A coterie of scholars, mainly located in northern France, began to study writings from classical antiquity with a new intensity and sense of purpose. They explored a wider range of Latin texts than their Carolingian predecessors, including late ancient translations of Greek works, such as Chalcidius's partial version of and commentary on Plato's *Timaeus*, dating from the fourth century, and the Aristotelian translations and commentaries of Boethius (c. 480–c. 524). They pursued predominantly scientific and philosophical interests, though less single-mindedly than the Arab scholars of the tenth century, and with a considerable emphasis on the Roman as well as the Greek tradition. The outlook of John of Salisbury (c. 1115–1180), an Englishman educated in Chartres and Paris, reflects the characteristic strains of humanism in this era. His knowledge of Latin literature, much of it culled from medieval *florilegia*, or anthologies, rather than through direct acquaintance with the ancient texts, was impressively broad though often shallow and perforce patchy. He wrote in a fluent and accomplished Latin style though without any attempt to imitate classical authors; and though he peppered his treatises with quotations and anecdotes from ancient literature, these snippets were deployed solely for his own purposes with no concern for their original context and import. Other twelfth-century scholars engaged in cosmological speculation and provoked accusations of heresy by employing Platonic concepts to investigate the relationship between God and the created universe.

FROM MEDIEVAL TO RENAISSANCE HUMANISM

Another product of twelfth-century France was the development of a new approach to Latin grammar based on a philosophical and logical analysis of syntax rather than on the careful study of Roman authors. Codified in two

enormously influential manuals, Alexander of Villedieu's *Doctrinale* (1199) and Evrard of Béthune's *Graecismus* (1216), both written in verse for easy memorization, this unclassical method of teaching grammar quickly spread to Italy where it remained the staple of elementary education in Latin until the end of the fifteenth century. Equally successful was Geoffrey of Vinsauf's *Poetria nova*, a verse textbook on rhetoric written between 1208 and 1213. Devoid of classical examples, it provided rules for obtaining an abstract eloquence unrelated to the prose of ancient Roman authors. This technique fitted in well with the *ars dictaminis*, a simplified method of composing public letters that had been widely adopted in Italy for the training of notaries, lawyers, and chancery officials. Their need for a practical, efficient, and uncluttered form of Latin expression led to a rejection of Roman models.

The first glimmerings of humanism appeared against the background of—and most likely in reaction to—this neglect of the classical tradition in thirteenth-century Italy. The Paduan notary and judge Lovato dei Lovati (c. 1240–1309), usually described as a *pre-* or *proto-humanist*, broke new ground with his attempt to write Latin verse epistles in the style of Roman poets. Medieval scholars who took an interest in classical literature had not aspired to write Latin in an authentically ancient manner. Lovati, by contrast, made a deliberate (though far from successful) effort to imitate the vocabulary, meter, and tone of the Roman poetry he admired, including the tragedies of Seneca, whose metrics he explained in a brief treatise, and the lyric poems of Catullus, Tibullus, and Propertius, which were hardly known at the time. In his official capacity, however, Lovati continued to write the traditional Latin of the *ars dictaminis*. His disciple Albertino Mussato (1261–1329) not only composed a Senecan verse tragedy, *Ecerinis*, but also extended the classicizing reform of Latin to prose by modeling his history of Emperor Henry VII on Livy. Yet he, too, continued to use medieval Latin in public letters and speeches, as did other humanists throughout the fourteenth century.

It was with Petrarch (Francesco Petrarca, 1304–1374) that the nascent humanist movement came into its own. He became an exemplary figure whose predilections, interests, and activities set the agenda for later generations of humanists. In his own Latin compositions, he emulated both the prose and poetry of Roman authors as well as working out a sophisticated theory of imitation. He collated and edited manuscripts of Livy and applied his philological acumen to the correction of other classical texts. He recovered works that had been effectively lost since antiquity, including Cicero's letters to Atticus. He rejected medieval scholasticism, with its emphasis on Aristotelian logic, natural philosophy, and metaphysics, and favored instead the rhetoric of Cicero and Seneca, which had the power to move hearts and stir emotions. He felt that he was living at the dawn of a new era following a dark age of ignorance and barbarism; and he believed that the moving force behind this large-scale cultural transformation was the gradual recovery of the heritage of classical antiquity through his own efforts and those of like-minded scholars.

RENAISSANCE HUMANISM IN FIFTEENTH-CENTURY ITALY

There have been attempts by modern scholars to connect the rise of Renaissance humanism with the political circumstances of fifteenth-century Italy. Its origins have been linked to the struggle of republican Florence against the monarchical tyranny of Milan in the early years of the century, or it has been seen as both reflecting and fostering a new spirit of active engagement of the citizenry in communal affairs. This notion of *civic humanism* has not stood up well in the face of overwhelming evidence showing that humanism flourished in a wide variety of political and social contexts. Similarly, the multiplicity of mutually contradictory views held by humanists on any given subject has undermined persistent efforts to align the movement with a particular ideological bent or philosophical persuasion. The only conviction that humanists demonstrably held in common was their passionate dedication to study of classical antiquity.

Paul Oskar Kristeller's definition of a humanist as a professional teacher of the *studia humanitatis* therefore corresponds most closely to the historical facts and has consequently won widespread acceptance. The term humanist, or *umanista*, in fact derives from late fifteenth-century Italian university slang that denoted a teacher or student of the *studia humanitatis*, just as a *legista* was someone who taught or studied law. The expression *studia humanitatis* itself had even longer associations with the movement. Petrarch noted it in his manuscript of Cicero's *Pro Archia*, a speech he himself discovered in 1333, and his devoted follower Coluccio Salutati (1331–1406), the first in a long line of humanist chancellors of Florence, began using it in 1369 to describe the study of classical literature. It soon became a frequent refrain, indeed a battle cry, among humanists defending or promoting their own activities.

Although Kristeller's account of the *studia humanitatis* as consisting of five academic subjects—grammar, rhetoric, poetry, history, and moral philosophy—has also

been widely adopted, it needs further refinement. These subjects may have been at the center of the Italian humanist curriculum during the fifteenth century. Yet their own interests ranged far beyond these fields, extending into all disciplines that relied on the wisdom of the ancients. In the Renaissance this meant almost every branch of learning: medicine, law, science, political thought, music, architecture, and all branches of philosophy. It was on this basis that Angelo Poliziano (1454–1494), the most learned classical scholar of the Italian Renaissance, decided that he would no longer lecture at the University of Florence on Latin and Greek literature, the normal subjects for a humanist professor but, instead, give a course on Aristotelian logic. Such an audacious move predictably provoked cries of derision from the philosophers whose academic territory he was invading. He responded to this outrage by maintaining, in his inaugural lecture of 1492, that as an expert on antiquity he was qualified to interpret any ancient text, not just poetry, history, and rhetoric, but also medicine, law, and philosophy. He further demonstrated this point in a dazzling series of philological investigations ranging over a broad spectrum of texts, including the *Corpus iuris civilis* (the sixth-century codification of Roman law), the scientific writings of Pliny the Elder, and the Greek sources of Latin medical terminology.

Fifteenth-century Italian humanists contributed to the *studia humanitatis*, broadly construed, in a number of ways. In the first place, they uncovered manuscripts of classical Latin texts that had been virtually unknown throughout the Middle Ages. In the wake of Petrarch's discovery of Cicero's letters to Atticus and *Pro Archia*, Salutati turned up a copy of his familiar letters while a humanist of the next generation, Poggio Bracciolini (1380–1459), found more unknown speeches of Cicero. Poggio's energetic hunt through monastic libraries in northern Europe also produced a complete copy of Quintilian's *Education of the Orator*, previously circulating in fragmentary form, and a masterpiece of Roman poetry, Lucretius's *On the Nature of Things*. Other important works unearthed by Italian humanists include the histories of Tacitus and the *Brutus* of Cicero.

When humanists found new works, or more accurate copies of ones that were already known to them, they ensured their further survival and diffusion, before the invention of printing, by copying and circulating them in manuscript. One of the innovations introduced by Italian humanists was to replace the crabbed and illegible gothic handwriting used in the late Middle Ages with an elegant and readable script that they believed was modeled on ancient Roman letter forms but that, in reality, dated from the Carolingian era. They also devised a cursive script, which is the ancestor of our italic character. In addition to copying texts, humanists also attempted to correct the errors that had inevitably crept into those texts through centuries of scribal transmission. They did this by comparing readings in different manuscripts or by making conjectural emendations—techniques that classicists still use today though with far greater methodological sophistication.

The next stage in dealing with the text of an ancient author was to explain and interpret it, often for the benefit of students. A large number of humanist commentaries on classical works grew out of university lectures. At a lower level, humanists also made their living as schoolteachers, equipping youngsters with the basic tools of Latin literacy that would enable them to gain access to the literary monuments of antiquity. Humanists such as Pier Paolo Vergerio (1370–1444) and Battista Guarino (1434–1503) wrote treatises touting the novelty of their teaching methods and boasting that the classical education they provided would inculcate a love of virtue and nobility in their young charges. Such extravagant claims no doubt assisted humanists to corner the educational market in the Renaissance. As is now known, however, the textbooks used in elementary Latin training changed relatively little from the thirteenth to the end of the fifteenth century. Moreover, the pedagogical techniques employed by humanists, which emphasized rote memorization and focused on grammatical, historical, and literary minutiae, were ill-suited to produce moral improvement. At the later stages of schooling, humanists made a greater impact, giving Virgil and Cicero a more prominent place in the curriculum than they had previously enjoyed and downgrading late ancient authors such as Boethius.

Fifteenth-century Italian humanists continued the classicizing reform of Latin style initiated by Lovati, Mussato, and Petrarch. By the early decades Cicero had become the accepted model for prose writers though it was not until the end of the century that a slavish and exclusive imitation of Cicero came into fashion. As an aid to writing correct classical Latin, Lorenzo Valla (1407–1457) compiled his *Elegantiae*, a catalogue of subtle linguistic distinctions, fine shades of meaning, and nuances of usage, based on his exhaustive knowledge of the entire Roman literary canon. In a bravura display of humanist historical scholarship, Valla deployed this same knowledge to discredit the "Donation of Constantine," a document underwriting papal claims to temporal sovereignty, as a crude medieval forgery. He also believed, like

Poliziano, that his superior command of Latin permitted him to interpret the *Corpus iuris civilis* more accurately than the legal scholars of his day.

Valla not only had an unrivaled mastery of classical Latin, he also knew ancient Greek well enough to translate the historians Herodotus and Thucydides. His expertise in both languages also allowed him to point out errors in the Vulgate, the standard Latin version of the New Testament, by comparing it with the Greek original. The revival of the study of Greek, which was well on its way by Valla's time, was one of the most important achievements of humanism. Very few medieval scholars had any acquaintance with Greek; and even though the works of some authors, including Aristotle, had been translated into Latin, the bulk of Greek philosophy, science, history, and literature was unknown in Western Europe. Beginning with Petrarch, humanists recognized the importance of recovering the Greek as well as the Latin heritage of antiquity. By traveling to Greece or studying with Byzantine émigrés in Italy, they learned the language and started to apply the techniques of editing and interpretation that had been developed for Latin texts to Greek ones. Greek, nonetheless, remained the preserve of a minority of humanists who served the larger intellectual community by translating a large body of texts into Latin. The writings of the Greek Church Fathers, many of them translated by Ambrogio Traversari (1386–1439), general of the Camaldulensian Order, formed an important element in this corpus.

Humanists were not concerned solely with texts. The material remains of antiquity, which were especially plentiful in Italy, were also of great interest to them. They visited architectural ruins and avidly collected Roman coins, inscriptions, and sculptures. Humanist historians, such as Flavio Biondo (1392–1463), subjected these artifacts to critical scrutiny and used them to supplement written records. This aspect of humanism laid the groundwork on which the disciplines of archaeology, numismatics, and epigraphy were later constructed.

RENAISSANCE HUMANISM IN SIXTEENTH-CENTURY EUROPE

By the turn of the sixteenth century, humanism had begun to spread from Italy to other European countries. The movement took on new contours and colors, reflecting the different cultures into which it was transplanted. Nonetheless, the humanist program that had taken shape in fifteenth-century Italy did not undergo radical changes but continued to develop within the same broad outlines. This process is well illustrated in the writings of the most outstanding and influential humanist of the period, Erasmus (c. 1469–1536). In his educational works the Dutch scholar banished the last vestiges of the medieval tradition of learning Latin and presented a thoroughly humanist pedagogical method firmly based on the study of Roman and Greek authors. Erasmus also brought the humanist reform of Latin style to new heights. With the entire resources of classical Latin at his command, he adopted and promoted a flexible and eclectic approach to prose composition, rejecting the rigid Ciceronianism of his day. Carrying forward the achievements of Valla and Traversari, Erasmus demonstrated the relevance of humanism to Christian as well as pagan antiquity by applying philological techniques to the text of the New Testament and producing numerous critical editions and translations of the Church Fathers.

The inroads that fifteenth-century Italian humanists had made into disciplines such as medicine, philosophy, and law were extended during the sixteenth century by scholars from all over Europe. The Englishman Thomas Linacre (c. 1460–1524) helped to edit the Greek text of Galen and translated many of his treatises into Latin. The Flemish scholar Justus Lipsius (1547–1606) reconstructed the philosophical system of the ancient Stoics, relying on Greek as well as Latin sources, and gave impetus to a popular fad for Stoicism that lasted until the 1660s. The French humanist Guillaume Budé (c. 1467–1540) brought the weight of his vast classical erudition to bear on the elucidation of obscure passages and terms in Roman law. He also wrote learned treatises on Roman coinage and Greek grammar.

Though Latin remained the *lingua franca* of humanism, facilitating communication among scholars of different nations, a feature of the movement in the sixteenth century was the increase of humanist writings in the vernacular. This phenomenon was not unheard of in the fifteenth century: Leon Battista Alberti (1404–1472) wrote a humanist treatise on household management in Italian, partly in order to demonstrate that the language was a suitable vehicle for scholarly discourse. Now, however, it proliferated and attained a respectability that it had previously lacked so that even a hard-core humanist such as Budé was prepared to write his treatise on the education of the prince in French. *The Prince* of Niccolò Machiavelli (1469–1527) and the *Essays* of Michel de Montaigne (1533–1592) are just two examples of influential works in the vernacular that were steeped in humanist culture.

THE LEGACY OF RENAISSANCE HUMANISM

The humanists' aim of reviving and restoring the heritage of classical antiquity was largely achieved by the seventeenth century. Although a few discoveries were yet to be made, almost all ancient Greek and Latin writings known today were available to scholars who could consult them in printed editions, often accompanied by learned commentaries and, in the case of Greek works, Latin translations. It was at this stage, however, that the seismic changes in European culture brought about by the Scientific Revolution, and the rise of modern philosophy made this body of knowledge, so revered by the humanists, increasingly irrelevant to contemporary needs. They continued to develop ever more sophisticated methods of investigating the textual and material remains of antiquity, gradually transforming themselves into the classicists and archaeologists of the present day. By 1809, when the term *humanism* was first coined by a German philologist to defend the study of Greek and Latin, the movement had become synonymous with the profession of classical scholarship. Although greatly marginalized since its heyday in the Renaissance, humanism continued to exert a significant and widespread cultural influence until well into the twentieth century through the resilient ideal of the classical education.

See also Aristotle; Avicenna; Carolingian Renaissance; Cicero, Marcus Tullius; Enlightenment; Erasmus, Desiderius; Galen; Hippocrates and the Hippocratic Corpus; John of Salisbury; Petrarch; Plato; Stoicism.

Bibliography

Baron, Hans. *The Crisis of the Early Italian Renaissance: Civic Humanism and Republican Liberty in an Age of Classicism and Tyranny*. 2 vols. Princeton, NJ: Princeton University Press, 1955.

Black, Robert. "Humanism." In *The New Cambridge Medieval History*, Vol. VII: *c. 1415–c. 1500*, edited by Christopher Allmand. Cambridge, U.K.: Cambridge University Press, 1998.

Black, Robert. *Humanism and Education in Medieval and Renaissance Italy: Tradition and Innovation in Latin Schools from the Twelfth to the Fifteenth Century*. Cambridge, U.K.: Cambridge University Press, 2001.

Celenza, Christopher S. *The Lost Italian Renaissance: Humanists, Historians, and Latin's Legacy*. Baltimore, MD: Johns Hopkins University Press, 2004.

Garin, Eugenio. *Italian Humanism: Philosophy and Civic Life in the Renaissance*. Translated by Peter Munz. New York: Harper and Row, 1965.

Gersh, Stephen, and Bert Roest, eds. *Medieval and Renaissance Humanism: Rhetoric, Representation and Reform*. Leiden: Brill, 2003.

Goodman, Anthony, and Angus MacKay, eds. *The Impact of Humanism on Western Europe*. London: Longman, 1990.

Grafton, Anthony. *Defenders of the Text: The Traditions of Scholarship in an Age of Science, 1450–1800*. Cambridge, MA: Harvard University Press, 1991.

Grafton, Anthony, and Lisa Jardine. *From Humanism to the Humanities: Education and the Liberal Arts in Fifteenth- and Sixteenth-Century Europe*. Cambridge, MA: Harvard University Press, 1986.

Hankins, James, ed. *Renaissance Civic Humanism: Reappraisals and Reflections*. Cambridge, U.K.: Cambridge University Press, 2000.

Kallendorf, Craig, ed. and trans. *Humanist Educational Treatises*. Cambridge, MA: Harvard University Press, 2002.

Kraye, Jill, ed. *The Cambridge Companion to Renaissance Humanism*. Cambridge, U.K.: Cambridge University Press, 1996.

Kraye, Jill, and M. W. F. Stone, eds. *Humanism and Early Modern Philosophy*. London: Routledge, 2000.

Kraemer, Joel L. *Humanism in the Renaissance of Islam: The Cultural Revival during the Buyid Age*. Leiden: Brill, 1986.

Kristeller, Paul Oskar. *Renaissance Thought: The Classic, Scholastic and Humanist Strains*. New York: Harper and Row, 1961.

Kristeller, Paul Oskar. *Renaissance Thought II: Papers on Humanism and the Arts*. New York: Harper and Row, 1965.

Nauert, Charles G., Jr. *Humanism and the Culture of Renaissance Europe*. Cambridge, U.K.: Cambridge University Press, 1995.

Otten, Willemien. *From Paradise to Paradigm: A Study of Twelfth-Century Humanism*. Leiden: Brill, 2004.

Rabil, Albert, Jr. ed. *Renaissance Humanism: Foundations, Forms, and Legacy*. 3 vols. Philadelphia: University of Pennsylvania Press, 1988.

Reynolds, L. D. and N. G. Wilson. *Scribes and Scholars: A Guide to the Transmission of Greek and Latin Literature*. 3rd ed. Oxford: Clarendon Press, 1991.

Rummel, Erika. *The Humanist-Scholastic Debate in the Renaissance and Reformation*. Cambridge, MA: Harvard University Press, 1995.

Seigel, Jerrold E. *Rhetoric and Philosophy in Renaissance Humanism*. Princeton, NJ: Princeton University Press, 1968.

Trinkaus, Charles. *In Our Image and Likeness: Humanity and Divinity in Italian Renaissance Thought*. 2 vols. Chicago: University of Chicago Press, 1980.

Weiss, Roberto. *The Renaissance Discovery of Classical Antiquity*. 2nd ed. Oxford: Basil Blackwell, 1988.

Wilson, N. G. *From Byzantium to Italy: Greek Studies in the Italian Renaissance*. London: Duckworth, 1992.

Witt, Ronald G. *In the Footsteps of the Ancients: The Origins of Humanism from Lovato to Bruni*. Leiden: Brill, 2000.

Jill Kraye (2005)

HUMAN NATURE

The phrase "human nature" is multiply ambiguous. Some early modern thinkers such as Thomas Hobbes, John

Locke, and Jean-Jacques Rousseau tended to mean by it the supposed nature of human beings before the advent of organized human society. But there is every reason to believe that human beings have always been highly social creatures, and that the idea of individuals coming together to form society is a myth.

Another ambiguity, exemplified in the opposition between Mencius and Hsun-tzu in the ancient Confucian tradition in China, and between differing traditions within Christianity, is over whether human nature is basically good and in need only of appropriate sustenance and education, or whether we are inherently evil and stand in need of discipline or radical transformation.

A further difference is between a conception of human nature as what is in each individual at birth (or, given modern understanding of genetics, at conception), as opposed to the nature of the fully formed adult after maturation, socialization and education. This has given rise to endless nature versus nurture debates.

The distinction between a priori and a posteriori truths allows us place both for philosophical analysis of concepts of human nature, and for the discovery of empirical facts in physiology, psychology, anthropology, sociology, and history.

A PRIORI THEORY: RATIONALITY

What is most distinctive of human beings as opposed to other animals— rationality, language, consciousness, self-consciousness, freewill, moral responsibility, the ability to love? (And must all these go together?) How would we recognize beings from outer space as having any of these capacities? Perhaps the most obvious criterion they would have to meet to count as rational thinkers and agents is that they should be able to give reasons for their beliefs and their actions, in language of some sort that we could come to understand.

What makes such rationality possible in us? Plato and René Descartes believed that we are essentially immaterial souls, so our distinctively rational nature lies beyond scientific investigation. But must minds, consciousness and rationality involve something nonmaterial, or are we made of matter alone? Aristotle saw our rationality as superimposed on what we share with the animals (perception and self-movement), which is itself superimposed on the basic functions of all life including plants (metabolism and reproduction). According to this understanding, we are animals of a special rational kind.

But even if we reject a dualism of substances, and say that mind or soul is whatever the brain enables us to do, we find an unavoidable duality of aspects. There are mental descriptions of our beliefs, desires, hopes and fears, and there are physical descriptions of neuron firings and chemical changes. We thus use an irreducible duality of explanations—justifying our actions and beliefs in terms of reasons, and explaining brain events in terms of their physiological causes.

EMPIRICAL THEORY: HUMAN NATURE AND NURTURE

Into this a priori conceptual framework we can fit empirical discoveries about human nature (and perhaps one day, about other rational beings elsewhere). There are plenty of such facts about the structure and functioning of our bodies—it is surely the size and complexity of our brains that explains our linguistic and rational abilities. There are also facts about our mental capacities, for example our recognition of faces, our tendency toward pair-bonding, and the need of children for attachment to parents.

In the light of Charles Darwin's theory of evolution, we can offer a scientific account of how the basic physical and mental commonalities of humans have evolved on this planet. With the aid of genetics and the fossil record, scientists are now piecing together the complicated story of how the faculties of rational thought and agency have come to be embodied in the human species. But we have to be very careful in applying Darwinian theory to human phenomena. Sociobiologists and evolutionary psychologists have tended to exaggerate. It is highly disputable whether every detail of contemporary human behavior has an evolutionary explanation; for example, donations to charity, the pursuit of religious vocations, and politicians' decisions to go to war.

Our reasons for action involve our beliefs and values, expressed in terms of our culturally developed concepts. Culture is at least as crucial to the realities of our contemporary human nature as evolution. It is superimposed on basic human biology, of course. That there are *some* innate tendencies in human nature is indisputable—for example, our sexual behavior. But the forms sexuality takes vary considerably between societies, and over time, and in devotedly celibate individuals its expression may be suppressed. The details of our behavior depend on the particular culture we have been brought up in. In the high-tech capitalist economy that now dominates the world, much of the social influence is exerted through the power of money, advertising, and the media. But it should not be forgotten that much behavior

depends on individual choices, as existentialists and religious traditions have emphasized.

HUMAN NEEDS AND RIGHTS

Like Plato and Aristotle, Immanuel Kant offered an objective basis for ethics, appealing both to pure reason and to empirical facts about human nature. Though Kant seemed to want to derive morality from rationality alone, he can be seen as appealing to a fundamental moral principle of respect for all rational beings "as ends in themselves." In this, he was obviously inspired by the Judeo-Christian ideal of love for one's neighbor as oneself, whereas Plato and Aristotle were more selective in their bestowal of respect for others. Karl Marx's and Jean-Paul Sartre's sense of human possibilities and the injustice of their denial were surely also influenced by the Judeo-Christian tradition. The Confucian notion of benevolence, and the Hindu and Buddhist programs of detachment from self, seem to point in the same direction of universal compassion.

Respect for all rational beings implies recognition of the rights and needs of all human beings. Rights imply corresponding obligations on other people, and the most appropriate place for talk of rights is in the negative cases: the rights not to be killed, injured, tortured, enslaved, imprisoned without trial, or exploited for someone else's benefit.

In his "second ethics," Sartre thought of human needs as objective values which demand to be fulfilled, if human beings are to flourish. The notion of need applies at several levels. There are things we need to maintain life and health—air, water, carbohydrates, protein, vitamins, medicines. There are psychological needs—most fundamentally the need of children for loving care, and there are typical adult needs for friendship, for sexual fulfillment, and for children of one's own. There are also needs for education and group membership, and needs to work, or contribute in some way to society.

When a human need is not met, it does not follow that someone is to blame. But when a human right is abused, then some person or group or social agency has done or encouraged the killing or torture, the enslavement or exploitation. And why have they done it? The answer will typically involve their seeing some advantage to themselves. There may be sadistic individuals who find intrinsic pleasure in causing pain, and many more are prepared to inflict suffering in the name of a "greater" cause (nation, party, or church), but most people do what they see as best for themselves. As Kant said, there is a "radical evil" in human nature, which consists in the tendency to prefer one's own interests over those of everyone else. But this is consistent with saying that we also have a potential for goodness and love.

LOVE

In the light of scarcity of resources, and individual and social evils, can we still entertain any hope for ethical and social progress, like Kant and other Enlightenment thinkers? Scarcity may perhaps be alleviated by scientific discovery and technological ingenuity. But new affluence breeds new needs and demands. There is an inherently competitive streak in human nature; we constantly rank ourselves against others. Our competitive tendencies may be acceptable in business and sport and in scientific and artistic achievement, but they easily turn into ruthlessness, cheating or greed. They may help drive social and cultural progress, but they need limitation by higher ideals of compassion and the common good.

What remedies are there? The first step is surely to name the evils, to try to make people aware of what is wrong, in ourselves and in society. For we are adept at finding good names for what we do: there are many possibilities of self-justification, self-deception, Freudian repression or Sartrian bad faith, and what Marx called "ideology," which covers up exploitation.

The notion of rationality alone does not give us much guidance as to what is ultimately worth aiming at, which of our desires are to be encouraged and developed, and which should be suppressed or transformed. People can be superbly intelligent and energetically persistent in action, yet utterly selfish or perverse.

The notion of love is perhaps more promising, provided we distinguish it, as in the Christian tradition exemplified in C. S. Lewis, from erotic and parental love, and from merely human friendship (as in Aristotle). The New Testament presents us with the ideal of *agape* (divine love, traditionally translated as "charity"), as presented in I Corinthians 13, Galatians 5, and I John 4. Sigmund Freud thought it impossible to fulfill, but even that dour old pessimist Arthur Schopenhauer recognized the possibility of saintly renunciation of self (and also, aesthetic contemplation) as a way that human beings can sometimes escape the near-universal domination of biologically based "will."

See also Altruism; Egoism and Altruism; Evolutionary Ethics; Moral Psychology.

Bibliography

Anderson, Thomas C. *Sartre's Two Ethics; From Authenticity to Integral Humanity*. Peru, IL: Open Court, 1993.

Aristotle. *Nichomachean Ethics*. Books I and X. Translated by Sarah Broadie. Oxford: Oxford University Press, 2002.

Clark, Mary E. *In Search of Human Nature*. London: Routledge, 2002.

Freud, Sigmund. *Civilization and its Discontents* (1930). Translated by Joan Riviere. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. Vol 21. London: Hogarth Press, 1953–1974.

Kant, Immanuel. *Groundwork of the Metaphysic of Morals* (1785). Translated by Mary J. Gregor. In *Immanuel Kant: Practical Philosophy*. Cambridge, U.K.: Cambridge University Press, 1996.

Kant, Immanuel. *Religion within the Boundaries of Mere Reason* (1793). Part 1. Translated by George di Giovanni. In *Immanuel Kant: Religion and Rational Theology*. Cambridge, U.K.: Cambridge University Press, 1996.

Lewis, C. S. *The Four Loves*. London: Fontana, 1963.

Marx, Karl. *Selected Writings in Sociology and Social Philosophy*. Edited by T. B. Bottomore and Maximilien Rubel. Harmondsworth: Penguin, 1963.

The Oxford Study Bible. New York: Oxford University Press, 1992.

Pinker, Steven. *The Blank Slate: The Modern Denial of Human Nature*. London: Allen Lane, 2002.

Plato. *Phaedo; Republic*. Books IV and IX. Translated by G. M. A. Grube. Indianapolis, IL: Hackett, 1992.

Rose, Steven, R. C. Lewontin, and Leon J. Kamin. *Not in Our Genes: Biology, Ideology, and Human Nature*. Harmondsworth: Penguin, 1984.

Schopenhauer, Arthur. *The World as Will and Representation* Vol. I (1819). Translated by E. F. J. Payne. New York: Dover, 1966.

Stevenson, Leslie, and David Haberman. *Ten Theories of Human Nature*. 4th ed. New York: Oxford University Press, 2004.

Wilson, Edward O. *On Human Nature*. Cambridge, MA: Harvard University Press, 1978.

Leslie Stevenson (2005)

HUMBOLDT, WILHELM VON

(1767–1835)

Wilhelm von Humboldt, the Prussian statesman, humanist, and linguistic scholar, was born in Potsdam; a younger brother was the scientist and explorer Alexander von Humboldt. Wilhelm von Humboldt's early education was placed in the hands of private tutors and was augmented by private instruction in Greek, philosophy, natural law, and political economy from distinguished men of Germany's Enlightenment. From these youthful studies Plato's idea of the soul and Gottfried Wilhelm Leibniz's concept of force left lasting impressions on his thought.

During and after his university years at Frankfurt an der Oder (1787) and at Göttingen (1788–1789), Humboldt began to question the rationalistic presuppositions of the Enlightenment. Like Johann Gottfried Herder, he viewed human society as a manifold of organic forces, closer to nature than to reason, and came to believe that true knowledge of humanity depended on the cultivation not of pure analytical reason but of deep-lying intuitive faculties.

Humboldt's political philosophy was outlined in a long essay, *Ideen zu einem Versuch die Grenzen der Wirksamkeit des Staats zu bestimmen*, written in 1791. Focused on the central theme of his thought—the inalienable value of the individual—this work propounds the humanistic creed that man's goal is "the highest and most proportional development of his powers to a complete and consistent whole." Reason must guide this development, but reason for Humboldt was a formative rather than a generative faculty. He criticized state control of education and religion for inflicting an arbitrary framework on diverse, organically developing human forces, whose unity could not be imposed from without but sought only from within.

In the last decade of the eighteenth century Humboldt was occupied with various scholarly projects, none of which he completed; at the same time his growing friendship with Friedrich Schiller and Johann Wolfgang von Goethe brought him into contact with contemporary aesthetic problems. From 1802 to 1807 he was Prussian ambassador to the Vatican, and in 1808 he was appointed to the ministry of religious and educational affairs in Berlin, in which position he drafted several papers on education and was chiefly responsible for the foundation of the University of Berlin. Thereafter, he served as Prussian diplomatic representative in Vienna (1810–1813), at the peace negotiations before and after Napoleon Bonaparte's downfall (1814–1815), and in London (1817–1818). Defeated in his effort to achieve a constitutional monarchy for Prussia in 1819, he retired from public service and devoted the remainder of his life to study.

HISTORY

Humboldt's humanism was based on his idea of historical experience. "The broadening of our existence and of our knowledge," he wrote in a letter of 1823, "is possible historically only through the contemplation of previous existence." Searching for a discipline by which man's accumulated historical experience could become the

foundation for a philosophy of man, Humboldt had already written several essays and drafts outlining principles for the study of Greek antiquity (*Über das Studium des Altertums und des griechischen insbesondere*, 1793), for a comparative anthropology (*Plan einer vergleichenden Anthropologie*, 1795), and finally for the historian's profession (*Die Aufgabe des Geschichtschreibers*, 1821). Sharing his generation's enthusiasm for ancient Greece, Humboldt believed that the study of Greek culture in its broadest aspects would promote a true philosophical knowledge of men, including "the knowledge of the manifold intellectual, sentient, and moral human powers." For Humboldt the Hellenic world was a unity of diverse forces, a cultural unity that his own times lacked but might regain through a comprehensive study of the Greeks. His plan for a comparative anthropology was to study the moral character of different human types; a great variety of sources would provide the data for establishing an ideal norm, which was not adequately represented by any specific individuality. To comprehend the wholeness in the diversity of human types required aesthetic insight, which was fundamental to the art of the historian. In an essay on Goethe's *Hermann und Dorothea*, he concluded that epic poetry, of which Goethe's drama was an example, could be compared to history. "The condition of the soul which gives rise to the necessity of history (in the truest and highest sense of the word) is similar to that out of which an epic is produced with the help of imagination and art." In Humboldt's essay *Die Aufgabe des Geschichtschreibers*, in which the affinity of his thought to Friedrich von Schelling's philosophy is clearly manifested, the historian's imagination is likened to the poet's. It differs from the free fantasy of the poet's in that it is more strictly subordinated to the historian's experience and feeling for reality; it is actually a "divining faculty" (*Ahndungsvermögen*) and a "connecting ability" (*Verknüpfungsgabe*).

The most notable feature of this essay is Humboldt's attempt to elucidate the role of ideas in history. "Everything that is active in world history," he declared, "is also stirring in the inner being of man." The ideas in history have preserved human experience in the mind. "The eternal original ideas of everything conceivable provide existence and value, the beauty of all physical and spiritual forms, the truth in the unalterable working of every force according to its indwelling law, the justice in the inexorable course of events which are eternally regulated and meted their just reward." For Humboldt the goal of history is "the realization of the idea representing itself through humanity from all sides and in all forms in which the finite forms can be connected with the idea." The task of the historian is therefore to represent this process of ideas being actualized in history.

LANGUAGE

Humboldt's language studies represent his chief legacy to posterity and marked, according to Ernst Cassirer, a new epoch in the history of the philosophy of language. Humboldt saw in the origin of language that crucial moment when man emerged from nature and, thus, the moment of connection between nature and idea. Language is for Humboldt the faculty by which man is identified as man. Speech and understanding are only different products of the power of language. The formation of languages depends on the spiritual forces of humanity, and languages are thus not merely an intermediary between individuals but "the most radiant sign and certain proof that man does not possess intrinsically separate individuality." Languages delineate the cultural characteristics of nations, each of which has its own individuality and arouses a sense of unity in men.

Humboldt's chief contribution to the study of linguistics was his concept of the "inner form" of languages (*innere Sprachform*), which consists of more than just external grammatical principles; it implies a deep-rooted subjective view of the world, a spiritual attitude, that controls the formation of concepts. "Because of the mutual dependency of thought and word," he wrote, "it is evident that the languages are not really means of representing the truth that has already been ascertained, but far more, means of discovering a truth not previously known. Their diversity is not a diversity of sounds, but of world outlook."

Humboldt's idea that each language has its own characteristic outlook, or inner form, found support in the linguistic studies of A. F. Pott and Heymann Steinthal in the nineteenth century and was suggested anew in the twentieth in the works of Benjamin Lee Whorf and Edward Sapir. His influence can also be traced in other areas of nineteenth-century thought—a passage from his political treatise provided the motto for J. S. Mill's essay *On Liberty*; his notion of the idea in history is closely related to Leopold von Ranke's doctrine of ideas; and his notion of historical experience is basic to the philosophy of Wilhelm Dilthey. In the twentieth century Cassirer, in the first volume of *The Philosophy of Symbolic Forms*, provided a penetrating evaluation of Humboldt's linguistic insights and a general philosophical context for the unmethodical profusion of his thought.

See also Cassirer, Ernst; Dilthey, Wilhelm; Enlightenment; Goethe, Johann Wolfgang von; Herder, Johann Gottfried; Language, Philosophy of; Leibniz, Gottfried Wilhelm; Mill, John Stuart; Plato; Schelling, Friedrich Wilhelm Joseph von; Schiller, Friedrich; Semantics, History of.

Bibliography

WORKS BY HUMBOLDT

Gesammelte Werke, 7 vols. Edited by Karl Brandes. Berlin, 1841–1852.

Gesammelte Schriften, 17 vols. Edited by the Royal Prussian Academy. Berlin: B. Behr, 1903–1936.

Werke. Edited by A. Flitner and K. Giel. Darmstadt, 1960–. Four of the projected five volumes are completed.

Humanist without Portfolio: An Anthology. Edited by Marianne Cowan. Detroit: Wayne State University Press, 1963. Translated by the editor.

WORKS ON HUMBOLDT

Biographies

Haym, Rudolf. *Wilhelm von Humboldt. Lebensbild und Charakteristik*. Berlin, 1856.

Kaehler, Siegfried. *Humboldt und der Staat*. Munich and Berlin, 1927.

Leroux, Robert. *Guillaume de Humboldt: La Formation de sa pensée jusqu'en 1794*. Paris: Belles lettres, 1932.

Spranger, Eduard. *Wilhelm von Humboldt und die Humanitätsidee*. Berlin, 1909.

Studies

Adler, G. J. *Wilhelm von Humboldt's Linguistical Studies*. New York: Press of Wynkoop and Hallenbeck, 1866.

Arens, Hans. *Sprachwissenschaft*. Munich, 1955.

Goldfriedrich, J. *Die historische Ideenlehre in Deutschland*. Berlin, 1902.

Hegselmann, Rainer. "Humboldt's Argument against the Welfare State: A Reconstruction in Terms of Game Theory." *Vienna Circle Institute Yearbook* 2 (1994): 229–243.

Kittel, Otto. *Wilhelm von Humboldts geschichtliche Weltanschauung im Lichte des klassischen Subjektivismus der Denker und Dichter von Königsberg, Jena und Weimar*. Leipzig, 1901.

Leroux, Robert. *L'anthropologie comparée de Guillaume de Humboldt*. Publications de l'Université de Strasbourg, Fasc. 135. Strasbourg, 1958.

Love, Nancy S. "Rawlsian Harmonies: Overlapping Consensus Symphony Orchestra." *Theory, Culture and Society* 20 (6) (2003): 121–140.

Reill, Peter H. "Science and the Construction of the Cultural Sciences in Late Enlightenment Germany: The Case of Wilhelm Von Humboldt." *History and Theory* 33 (3) (1994): 345–366.

Robins, Robert H. "Leibniz and Wilhelm Von Humboldt and the History of Comparative Linguistics." In *Leibniz, Humboldt, and the Origins of Comparativism*, edited by Tullio DeMauro. Philadelphia: John Benjamins, 1990.

Sorkin, David. "Wilhelm Von Humboldt: The Theory and Practice of Self-Formation ('Bildung'), 1791–1810." *Journal of the History of Ideas* 44 (1983): 55–74.

Sweet, Paul R. "Young Wilhelm Von Humboldt's Writings (1789–93) Reconsidered." *Journal of the History of Ideas* 34 (1973): 469–482.

Vogel, Ursula. "Liberty Is Beautiful: Von Humboldt's Gift to Liberalism." *History of Political Thought* 3 (1982): 77–101.

Wach, Joachim. *Das Verstehen*, Vol. I. Tübingen, 1926.

Wulf, Christoph. "Perfecting the Individual: Wilhelm Von Humboldt's Concept of Anthropology, Bildung and Mimesis." *Educational Philosophy and Theory* 35 (2) (2003): 241–249.

Howard Isham (1967)
Bibliography updated by Tamra Frei (2005)

HUME, DAVID
(1711–1776)

David Hume, considered by many the finest Anglophone philosopher, one of the first fully modern secular minds, and, along with Adam Smith, the leading light of the Scottish Enlightenment, was the author of four major philosophical works and many essays.

Born on April 26, 1711, in Edinburgh, Scotland, Hume spent his childhood mostly at Ninewells, the family estate near Berwick. Though his family was of good social standing, they were not rich, and, as the second son, he had to be prepared to earn a living to supplement an inadequate inherited income. He attended Edinburgh University from the ages of eleven to fifteen, in which city he remained to study law. Finding this not to his taste, Hume returned to Ninewells and threw himself into an intensive program of intellectual self-development. He read widely in ancient and modern literature, improved his knowledge of science and languages, and devoted himself above all to philosophy. In this way, sometime before he turned eighteen, Hume achieved the breakthrough that, he reported, "open'd up to me a new Scene of thought, which transported me beyond Measure, & made me, with an Ardor natural to young men, throw up every other Pleasure or Business to apply entirely to it" (*The Letters of David Hume* 1932, vol. 1, pp. 13–14).

However, the strain eventually told on Hume's health, and he was obliged to curtail his studies and pursue a more active life. To this end, he secured employment with a Bristol merchant in 1734. Though this venture into the world of commerce was brief, his health was sufficiently restored to enable him to undertake the composition of the systematic philosophical treatise by which he hoped to make his literary mark. To stretch his meager

income further than was possible in Britain, Hume relocated to France—first to Reims, then to La Flèche in Anjou—where he was able to benefit from the outstanding library of the Jesuit college.

Hume returned to England in 1737 with the intention of publishing the first two books, *Of the Understanding* and *Of the Passions*, of the work he decided to call *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. After publishing them as volume 1 in 1739, he went home to Scotland to revise the third book, *Of Morals*, which he published as volume 2 the following year. Never before or since has anyone so young published a philosophical work so comprehensive, ambitious, original, or accomplished. Still, Hume's obvious aspiration to be acknowledged the Isaac Newton of philosophy did not sit well with contemporaries. Reviewers were mostly hostile and uncomprehending, so that the *Treatise* "fell *dead-born from the Press*; without reaching such distinction as even to excite a Murmur among the Zealots" (1987, p. xxxiv).

Having wisely taken the precaution to publish anonymously, Hume soon recovered from his failure and decided to apply his immense literary gifts to the more widely accessible medium of the essay. His *Essays, Moral and Political* of 1741 and 1742 duly succeeded where the *Treatise* failed. With a public won, together with a keen sense of its tastes, Hume presented a selection of the doctrines of the *Treatise* with some previously unpublished material in the form of *Philosophical Essays concerning Human Understanding* in 1748 (retitled *An Enquiry concerning Human Understanding* in 1758). With its companion published three years later, *An Enquiry concerning the Principles of Morals*, Hume firmly established his reputation as one of the leading philosophical thinkers of his day. Around the same time Hume composed his *Dialogues concerning Natural Religion*, but was prevailed on not to publish it during his lifetime. From that point on, Hume devoted himself to essays and wrote his most popularly successful work of all: the six-volume *History of England: From the Invasion of Julius Caesar to the Revolution of 1688* (1754–1762).

Hume held a number of posts during his life, though he never succeeded in securing an academic position. In 1745 he served as tutor to the mentally unbalanced Marquess of Annandale. From 1746 to 1749 he was secretary to Lieutenant-General James St. Clair (1720–1806), whom he accompanied on a military expedition to Brittany. He was keeper of the Advocates Library in Edinburgh from 1752 to 1757. In 1763 Hume became private secretary to Lord Hertford (1718–1794), the British ambassador to France, where he spent the next three years being continually fêted and forming friendships with several leading figures of the French Enlightenment, including Denis Diderot, Jean Le Rond d'Alembert, and Jean-Jacques Rousseau (though this last connection was to end in conflict). The last position he held was that of secretary of state in the Northern Department, from 1767 to 1768.

Physically, Hume was tall, somewhat ungainly, and, by the mid-1740s, corpulent. He never married, initially for lack of means to support a family, and afterward from a preference for bachelor life. Hume's most extraordinary quality was his personality. Warm, generous, even-tempered, and honorable in all matters, he gained and kept an enormous number of close, devoted friends. This included many prominent clergymen who time and again staunchly defended him against his persecutors. Hume was thus able to spend his final years in Scotland in tranquillity, surrounded by well-wishing friends and family. When death came on August 25, 1776, he took it in the best spirit imaginable, while also making sure that no tales could be spread that his religious skepticism had weakened in the end.

Hume's influence on philosophy during his lifetime was nothing like what it later became. His moral theory undoubtedly made an impact on Smith's *Theory of Moral Sentiments* (1759), while his theory of the understanding provided Thomas Reid with his principal foil in *Inquiry into the Human Mind, on the Principles of Common Sense* (1764). Reid and other less respectful philosophers of the Scottish "commonsense" school focused many of their severest criticisms on the *Treatise*. Their misunderstandings and misrepresentations of that work so infuriated Hume that he published an advertisement with the final edition of the *Enquiries* produced under his supervision (1777), desiring that these maturer efforts would "alone be regarded as containing his philosophical sentiments and principles."

A sea change in the reception of Hume's theory of understanding occurred in 1783, when Immanuel Kant declared that Hume's treatment of cause and effect was responsible for awakening him from his dogmatic slumber. Kant's own transcendent importance in the history of philosophy, and the scholarly attention devoted to almost his every word, led to a reappraisal of the worth and importance of the philosopher Kant credited with making his achievements possible, and it was not long till the *Treatise* came to be recognized as Hume's masterpiece.

Being cast as Kant's John the Baptist did, however, have its downside, and many have labored to bring

Hume's legacy out from under the shadow of Kant. Influenced by the latter, philosophers in the nineteenth century, and for much of the twentieth as well, tended to esteem Hume almost exclusively for the power of his skeptical arguments regarding reason, the natural world, and religion. Since then, the positive, constructivist aspects of his theory of understanding have come to be equally prized, as have his theories of passion, actions, morality, and aesthetics. Today, interest in Hume's philosophy is greater than ever and the wave shows no sign of cresting.

THE TREATISE AND THE ENQUIRIES

Most scholars accept the essential correctness of Hume's assertion that there are few substantive differences between the *Treatise* and the *Enquiries*, and none of great consequence. Instead, the earlier and later works differ primarily in inclusiveness and style. The *Treatise* was pitched at the highest level, to pass muster with the most learned, exigent readers. Questions left unraised in the *Enquiries* were pursued at considerable length, whole batteries of arguments were assembled in support of major theses, and every effort was made to be both systematic and comprehensive.

By contrast, the *Enquiries* were aimed at the same readers who enjoyed Hume's more philosophical essays. This seems to have been the principal reason for his decision to omit from the first *Enquiry* almost everything in parts 2, "Of the Ideas of Space and Time," and 4, "Of the Skeptical and Other Systems of Philosophy," of book 1 of the *Treatise*. Much of parts 1, "Of Ideas," and 3, "Of Knowledge and Probability," were also sacrificed, so that what remains seems less like a condensation of the *Treatise* than a greatly expanded and improved version of the abstract of the *Treatise* that Hume published in 1740. The second *Enquiry* drew on the moral philosophy of book 3 of the *Treatise*, while eschewing the theoretical framework of the latter in favor of a more strictly literary approach (which both explains why Hume thought it his finest work and why so few today agree). Neither *Enquiry* contains any considerable trace of book 2 of the *Treatise*, on the passions, and though occasional echoes of it are to be found in Hume's essays, they give no idea of the impressive, highly sophisticated theoretical framework one finds in book 2 of the *Treatise* (and the same is true of Hume's *A Dissertation on the Passions* [1757]). Thus, despite Hume's wish not to be judged by the *Treatise*, its unity, scope, and rigor make it the work that best represents what is most important and enduring in his philosophy.

HUME'S SCIENCE OF HUMAN NATURE

Hume believed human nature to be the proper focus of the philosopher because its first principles necessarily carry over to every human endeavor, cognitive and conative alike. A science of human nature affords fundamental insight not only into such domains as morals, aesthetics, and politics but "even *Mathematics, Natural Philosophy, and Natural Religion*," which "are in some measure dependent on the science of MAN; since they lie under the cognizance of men, and are judged of by their powers and faculties" (1978, p. xv). Situating himself in the line of British empiricist thinkers extending from Francis Bacon and John Locke, Hume restricted the investigation of human nature to evidence gleaned from "careful and exact experiments, and the observation of those particular effects, which result from its different circumstances and situations" (p. xvii). It constitutes a science insofar as one "must endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes." This may require one to revise initial determinations in the light of new experiments (Hume's evolving characterization of the difference between memory and imagination is a prime example), and obliges one to determine whether the fundamental principles of human nature have even wider scope (thus, Hume considered it a plus that much of his account of human nature extends to animals as well). Finally, the mandate for maximal simplicity means that the science of man should take the form of a system, deriving its principal authority from "the agreement of [its] parts, and the necessity of one to explain another" (p. 154).

THE ELEMENTS OF HUME'S SCIENCE OF HUMAN NATURE

OBJECTS. Hume considered human nature always and only in terms of perceptions. *Perception* is Hume's substitute for Locke's term *idea*, and it refers to all objects insofar as they are immediately present to one by consciousness, be it in sensation, reflexion, or thought (*reflexion* is Hume's catch-all term for the objects present to internal sense or inward sentiment, including passions, emotions, desires, volitions, and mental operations generally). For Hume, just as for Locke with *idea*, the indeterminacy of *perception*—the impossibility of contrasting it with anything that is not a perception because "[t]he mind never has anything present to it but the perceptions"—is its principal virtue. If things other than perceptions exist, then, as what never "can be present to the

mind, whether we employ our senses, or are actuated with passion, or exercise our thought and reflection" (1999, p. 202), they are no different from perfect nonentities so far as one's thoughts and actions are concerned. By contrast, even objects as fanciful as a billiard ball that transforms itself into wedding cake on being struck, though never present to the senses, are still objects of one's thought, and so too perceptions.

Perceptions come in two kinds: impressions and ideas. Impressions comprise sensations and reflexions, and ideas thoughts (the mental contents of thought, considered in themselves rather than in the capacity of signs used to signify other perceptions, whether by resemblance, linguistically, or in any other significative capacity). According to Hume, the difference between impressions and ideas consists in the greater "force and vivacity" of the former. This does not mean that impressions always make a forceful impression, for they can be so gentle as altogether to escape notice. Nor does it mean that they are vivid in the usual sense, since seeing a gray blur on an otherwise black night (visual sensation) is still more vivid than a brilliantly lit, detailed image in a daydream (visual idea).

The best indication of what Hume had in mind by "force and vivacity" is his subsequent equation of it with belief in the real existence of a content present to one in sensation, reflexion, or thought, all perceptions. According to Hume one believes in the reality of something that one merely thinks if one's conception of it exhibits force and vivacity, as when, on seeing smoke coming into the room, one not only thinks of a fire somewhere outside the room but believes that a fire really exists. Similarly, "the *belief* or *assent*, which always attends the … senses, is nothing but the vivacity of those perceptions they present" (1978, p. 86). More particularly, the vivacity of a perception seems to consist in a feeling distinctive of the manner in which an object in sensation or reflexion is apprehended, or an object in thought is conceived, in virtue of which it is regarded as really existent—actual rather than merely possible, fact rather than fiction.

If this reading is correct, then one needs to distinguish two senses of *exists* in Hume: an object, even if it is a mere fiction, exists simply in being present to consciousness (p. 66–67), but it is taken to be really existent if, in addition, it is perceived or conceived in a lively manner (pp. 84–123). Sensations and reflexions are impressions because human (and animal) nature is so constituted that these objects have only to appear to be believed really existent, whereas objects present to one only in thought are not believed really to exist unless circumstances intervene to induce one to conceive them with a high enough degree of force and vivacity. One of the principal occupations of Hume's theory of understanding was to determine what those circumstances are and to identify the underlying principles.

Finally, Hume distinguished perceptions according to whether they are complex or simple. In general, an impression or idea counts as simple if it cannot be distinguished into two or more components (different significative uses to which the same simple perception may be put do not compromise its intrinsic simplicity). But Hume also seems to allow that perceptions distinguishable in this way may still be simple if it is impossible for them to be derived by the combination or blending of perceptions already in one's possession (e.g., "The impressions of touch are simple impressions, except when consider'd with regard to their extension" [1978, pp. 230–231]).

THE COPY PRINCIPLE AND HUME'S THEORY OF ORIGINS. The "*full* examination" of the question of how impressions and ideas "stand with regard to their existence, and which of the impressions and ideas are causes and which effects" is "the subject of the present treatise" (1978, p. 4). To this end, Hume notes that one's simplest perceptions all seem to come in duplicate impressions and nearly exactly resembling ideas, and asks if there is any causal significance to this relation. He then formulates perhaps the most important principle of his science of human nature: because experience shows that simple impressions invariably precede their resembling ideas, "all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent" (p. 4). The causal dependence of ideas on impressions expressed in Hume's copy principle owes its importance to his preeminent methodological concern to find a better method of clarifying the ideas at the heart of traditional metaphysical disputes than definition can provide:

> Complex ideas may, perhaps, be well known by definition, which is nothing but an enumeration of those parts or simple ideas, that compose them. But when we have pushed up definitions to the most simple ideas, and find still some ambiguity and obscurity; what resource are we then possessed of? By what invention can we throw light upon these ideas, and render them altogether precise and determinate to our intellectual view? Produce the impressions or original sentiments, from which the ideas are copied. These impressions are all strong and sensible.

> They admit not of ambiguity. They are not only placed in a full light themselves, but may throw light on their correspondent ideas, which lie in obscurity. And by this means, we may perhaps attain a new microscope or species of optics, by which, in the moral sciences, the most minute, and most simple ideas may be so enlarged as to fall readily under our apprehension, and be equally known with the grossest and most sensible ideas, that can be the object of our enquiry.
>
> (1999, pp. 135–136)

Hume's science of human nature is, in the first instance, a critique of traditional philosophical definitions whereby they are supplemented or, more usually, supplanted, by psychological accounts tracing ideas to their originating impressions. These accounts inform everything else in the science, and it is often impossible to understand the positions Hume takes without returning to his explications of the relevant ideas in terms of their originating impressions.

RELATIONS. To understand the nature of relation for Hume, one first needs to consider the two ways in which relations may be affirmed. If one can affirm a relation independently of the senses, and so of all matters of fact and real existence, one's affirmation is a case of knowledge and the relation affirmed is a necessary one. For "the necessity, which makes two times two equal to four, or three angles of a triangle equal to two right ones, lies only in the act of the understanding, by which we consider and compare these ideas" (1978, p. 166). When immediate, the knowledge of a relation is intuition, when it consists of a continuous sequence of intuitions, it is a demonstration.

Knowledge of a relation of ideas is attainable (1) when one is sensible of the impossibility of forming one idea without including another as a constituent, as, for example, one cannot form the idea of a valley without incorporating into one's conception the idea of mountains (p. 32), or, (2) even if the ideas can be conceived separately, one is sensible of the impossibility of conceiving a change in their relation without conceiving a change in the ideas themselves (p. 69), as "the shortest distance between two points is sa straight line" is known to be necessary even though shortness (a quantity) and straightness (a quality) are conceivable independently (pp. 49–50). (The first type coincides with Kant's notion of an analytic judgment, the second with that of a synthetic *a priori* judgment; Hume did not, however, see fit to subdivide intuitive knowledge this way, that is, he either did not recognize or did not attribute to the question of the possibility of synthetic *a priori* judgments the same importance Kant would afterward accord it.) Either way, one's affirmation of a relation amounts to knowledge if and only if one is sensible of the impossibility of conceiving the ideas concerned in any other relation (pp. 652–653).

Where knowledge is lacking, and other relations between the ideas (or none at all) are conceivable, one can still affirm a relation between distinct perceptions with probability, that is, with a certainty extending anywhere from just above logical possibility all the way to a certainty so great as to be immune to doubt (termed *proofs* by Hume, e.g. "the sun will rise tomorrow" and "all men must die"). Such relations consist essentially in transitions of thought characterized by a quality Hume termed *facility* (1978, pp. 99, 204, 220, 260). There is considerable evidence that Hume conceived of facility as affective; that is, like the vivacity of impressions or ideas in virtue of which one believes them really to exist, the facility constitutive of probable relations is a content the mind does not conceive but feels. Facility and vivacity tend to go together in Hume's theorizing. When a relation between ideas is known, facility and vivacity affect are redundant to the relation and its affirmation since one is "necessarily determin'd to conceive them in that manner" (p. 95). Only when one remains free to conceive both sides of the question can assent be supposed to be a matter of feeling rather than an act of thought. In this regard, one of the most important principles of Hume's theory of understanding is that the more facile the transition from a lively perception to an idea in thought (= the stronger the relation), the more nearly the vivacity of one's conception of it (= belief in its real existence) approaches that of the lively perception itself (pp. 98–99).

Association. The effect of a facile transition between perceptions is to associate them in reflexion or thought, and it is in this association that their relation consists. With the precedent of Newtonian gravitation in mind, Hume saw fit to characterize association as "a kind of ATTRACTION, which in the mental world will be found to have as extraordinary effects as in the natural, and to shew itself in as many and as various forms" (1978, pp. 12–13). In the absence of the real connections falsely imputed to perceptions by the sophisticated and simple alike, the associative ties felt between perceptions are the source of all order and unity among them. Finally, in accordance with his scientific ideal of maximal generality and simplicity, Hume resolved all species of association into expressions of three fundamental associative princi-

ples: the contiguity of perceptions in space or time, their resemblance, and their connection as cause and effect—as "these are the only ties of our thoughts, they are really to us the cement of the universe, and all the operations of the mind must, in great measure, depend on them" (p. 662).

Natural and philosophical relations. Not all relations are constituted by facile transitions of thought. Hume designated those that are natural and those that are not philosophical relations. Since one can arbitrarily compare anything with anything else, and since no two objects admit comparison unless they have some degree of resemblance, resemblance counts not only as a natural but also as a philosophical relation; and philosophical resemblance is, in turn, the condition for other natural relations to assume a nonassociative philosophical dimension: identity, space and time, quantity (in number), quality (in degree), cause and effect, and contrariety. The crucial thing to remark here is that, except in cases of intuitive or demonstrative knowledge, philosophical relations seem to have no independent power to generate belief (vivacity), and so are parasitic on natural relations for their power to influence one's thoughts and actions. Hume made this explicit in the case of the cognitively preeminent relation, causation, for "tho' causation be a *philosophical* relation, as implying contiguity, succession, and constant conjunction, yet 'tis only so far as it is a *natural* relation, and produces an union among our ideas, that we are able to reason upon it, or draw any inference from it" (1978, p. 94).

HUME'S REJECTION OF ABSTRACT IDEAS. Hume expressed complete agreement with George Berkeley's exclusion of abstract ideas from the explanation of general ideas and terms. The keystone of this critique of abstraction is the separability principle that Hume, like Berkeley before him, made a centerpiece of his philosophizing. According to this principle, whatever objects (perceptions) are different are distinguishable, and so separable in thought; and vice versa (1978, p. 18). So far as abstraction is concerned, this means that one cannot abstract any X from any Y unless X can be perceived and conceived even in the absence of Y. For example, because the distinction between the shape and color of a visible object fails to satisfy the separability principle, the notion that these are distinct perceptions (different abstract ideas, as Locke supposed) has to be rejected as an illusion cast by language. For while there is indeed a significative distinction to be drawn in the use of the idea of a visible object to designate, on the one hand, things resembling it in shape and, on the other hand, things resembling it in color, when the idea is considered in itself, apart from any significative use to which it may be put, its shape and color are ineluctably one.

Accordingly, differences of aspect—that is, distinctions that fail to conform to the separability principle (sometimes called distinctions of reason)—are never intrinsic to the object to which they are ascribed, but are instead always the by-product of the relations in which it stands to other objects. Thus, a globe of white marble may be found to resemble a black globe of papier-mâché, a white cube of sugar, or an oblong piece of red marble; and since resemblance is an associative relation, the facile transition from a white globe to a black globe will set up an relational dynamic in which it becomes easier to make a transition next to the idea of a blue globe, red globe, or yellow globe, than to any nonspherical white or red object. In the same way, a transition from the white globe to a white cube will make it easier to transition next to the idea of a white oblong or any other white shape than to a black globe or red oblong. It is in these divergent axes of resemblance relations, ramifying in various directions from the same object, as it were, that aspects have their basis.

Resemblance association alone does not, however, suffice to explicate general representation. Custom is equally indispensable, "If ideas be particular in their nature, and at the same time finite in their number, 'tis only by custom they can become general in their representation, and contain an infinite number of other ideas under them" (1978, p. 24). The habits instilled by frequently encountered axes of resemblance association lie in readiness to be triggered by any of the infinitely many possible stimuli (determinate, nonabstract impressions or ideas) capable of triggering it (= representational generality); and which of the many habits it happens to trigger will determine to which species a given stimulus will be recognized as belonging (i.e., under which general sort it will be subsumed or classified). For example, a single, fully determinate (nonabstract) perception of an equilateral triangle one inch in circumference can serve as a general representation of figures, rectilinear figures, regular figures, triangles, or equilateral triangles, according to which custom one uses it to represent or which custom it triggers in a particular context (pp. 21–22). Finally, with the addition of words to overcome the confusion that would otherwise result either from the capacity of the same idea to trigger any of various customs, or from the same custom to be triggered by dissimilar ideas, one arrives at Berkeley's principle "that all general ideas are nothing but particular ones, annexed to a certain term, which gives them a more extensive signification, and

makes them recall upon occasion other individuals, which are similar to them" (p. 17).

SPACE AND TIME. Hume's treatment of abstract ideas exemplifies his general method of tracing ideas to their originating impressions; only here, where association and custom are indispensable, the experience of the operations of one's own mind (transitions of thought, the facility affect essential to associative relation, and the triggering of customs) proves to be the source of contents essential to these ideas. The abstract ideas of space and time are a case in point. Just as the shape and color of a visible object are one and indistinguishable, so, too, are extension and color. That is, the only idea one can derive from an impression of, say, uniform purple is the idea of uniform purple. To distinguish the extension from the color, one must compare the impression to others, associate them according to their resemblances, and, from the different axes of resemblances thus formed, arrive at last at an ineluctably relational conception of their difference.

Even so, to form a visual idea of space it is not enough simply to find what is resembling between purple, green, yellow, and other uniformly colored expanses, or between these and nonuniformly colored expanses. Visual space is the idea of something in which visible objects do or can appear and disappear, change their color and contour, grow, shrink, and alter their relative visible positions and situations inside, outside, alongside, adjacent, separated, above, below, right, left, in front, or behind one another. An idea with such limitless determinability is impossible except when visual perceptions are conceived of as an ordered manifold, or nexus, formed of coexistent loci (points) that preserve their relative positions to one another (their situation and relations) through any and all changes in respect of light and color ("co-existent parts dispos'd in a certain order, and capable of being at once present to the sight" [1978, p. 429]). That is, for Hume, the visual idea of space is the outcome of comparing visible objects, associating them according to their various resemblances, and forming habits when these associations are continuously reinforced, whether by frequent recurrence or some other cause. The key, as with aspects and distinctions of reason generally, is that visible space is never anything present to our eyes, prior to and independently of experience and habit, but rather something that exists only in and through the actions and affects of associative imagination (imagination in its associative capacity).

Unless this is appreciated, one cannot hope to understand how, on Hume's view, it is possible to form an idea of space common to vision and touch alike, notwithstanding the qualitative incommensurability of the objects of the two senses. For, lacking the ability to discriminate aspects immediately (nonrelationally), one can no more distinguish the extension of a tangible object from its other distinctively tactual qualities (hard or soft, smooth or rough, and wet or dry) than one can distinguish the extension of a visible object from its color. Consequently, to find visible and tangible space in any way resembling in appearance (sensible quality), one would have to find wet to be "like" yellow, red "like" softness, and so on, which of course is impossible. The locus of resemblance in virtue of which tangible and visible objects alike are supposed to instantiate the same general idea of space must instead lie in the operations the mind performs on these otherwise incommensurable appearances.

In particular, by contrast with data of the other senses, one is able to discern, and keep track of, distinctions of the finest, subtlest kind among visible and tangible appearances—distinctions sufficient in each case for association and custom to yield the abstract idea of an ordered manifold of coexistent loci (points) that preserve their relative positions to one another (their situation and relations) through any and all changes. To the imagination, then, producing and operating with two such similar manifolds feels so similar that, notwithstanding their radical qualitative disparity as appearances, it ranks them under a single, highly general idea of space. Moreover, thanks to the innumerable correlations (constant conjunctions) disclosed by experience between the objects situated in the respective imaginary spaces of each sense, one fancies that one is dealing not with distinct instances of the same general idea, but with a single, mulitsensory space, with its own, sense-divide transcending objects.

Hume's account of the origin of the idea of time differs from that of space in two principal regards: (1) whereas ideas of spatial features originate only in vision and touch, temporal ideas can be "deriv'd from the succession of our perceptions of every kind, ideas as well as impressions, and impressions of reflection as well as of sensation" (1978, pp. 34–35); and (2) whereas the manner of appearance of the spatial is defined by "that quality of the co-existence of parts," the temporal "is compos'd of parts that are not co-existent … and consequently that idea must be deriv'd from a succession of changeable objects" (p. 36). These differences aside, the psychological processes whereby ideas of the temporal are acquired are identical to those that give rise to ideas of the spatial.

From an unchanging object no idea of time can be derived "since it produces none but co-existent impres-

sions"; only "a succession of changeable objects" can yield the idea of something composed of noncoexistent parts. But since the successiveness of, say, five notes played on the flute cannot be perceived or conceived independently of the sounds—"The ideas of some objects it [the mind] certainly must have, nor is it possible for it without these ideas ever to arrive at any conception of time" (p. 37)—any supposition that the former, as the manner of appearance of these auditory objects, is something really distinct from these objects themselves falls foul of Hume's antiabstractionist separability principle. So, just like the idea of space, that of time can only be formed by comparing distinct perceptions and associating them in resemblance relations, until a custom is produced that stands in readiness to be triggered by all and only those stimuli to which ideas of succession and duration are applied. Time, understood as an ordered manifold of determinable positions composed of indivisible, noncoexistent instants, is thus, on Hume's account, as much an amalgam of the senses and associative imagination as space.

It is in connection with time that Hume formulated another of his principles, restricting the application of ideas according to the copy principle, "Ideas always represent the objects or impressions, from which they are deriv'd, and can never without a fiction represent or be apply'd to any other" (1978, p. 37). Like the copy, separability, and other principles of concern to Hume, this principle governs only one's perception of objects in sensation, reflexion, and thought, and does not imply any restriction on one's talk of objects. Nevertheless, since perceptions are the only objects that can ever be present to one's mind, the principle restricting the application of ideas according to the copy principle restricts one's discourse to the extent that objective meaning can attach to what one says only insofar as it cashes out ideationally. And temporal ideas are a case in point: While one is free to speak of unchanging objects, no objective meaning can attach to one's discourse since one has no ideas other than those copied from fleetingly existent perceptions.

Denial of infinite divisibility. Because one's abstract ideas of space and time "are really nothing but particular ones, consider'd in a certain light" (1978, p. 34), Hume concluded that infinitely divisible space and time are impossible even to conceive. For since particular ideas are one and all copied from particular impressions, and since experience shows that one's impressions admit being divided to the point where an indivisible temporal and/or spatial minimum is reached, it follows that the ideas one derives from these impressions can never serve to conceive an infinitely divisible spatial or temporal object. (For similar reasons, Hume denied the conceivability of a vacuum in space or time.) Thus, whatever mathematicians may pretend to the contrary, the first principles of mathematics "are founded on the imagination and senses: The conclusions, therefore, can never go beyond, much less contradict these faculties" (p. 638).

HUME'S THEORY OF UNDERSTANDING

Causal relations are the centerpiece of Hume's theory of understanding. Without them, "[i]nference and reasoning concerning the operations of nature would, from that moment, be at an end; and the memory and senses remain the only canals, by which the knowledge of any real existence could possibly have access to the mind" (1999, p. 149). This is because, of all relations linking ideas to impressions, none approaches cause and effect in its power to produce belief (enliven ideas). If I see smoke coming into the room, my belief in the reality of the unseen fire causing it is as great as in the smoke itself. If the hearing of voices on the other side of the fence brings persons to mind as their cause, I not only think there are people there, I believe them really to be there. Thus, whenever I infer a cause for a given effect or an effect for a given cause, I thereby expand the scope of what for me constitutes reality beyond the immediate evidence of my senses and memory.

Although the other principles of association, contiguity and resemblance, also have power to enliven the ideas they associate with impressions, without the support of causal relations "their influence is very feeble and uncertain" (1978, p. 109). For while I can think constant relations of time and place exist beyond the scope of my senses and memory, or think an identity based on the resemblance between nonsimultaneous resembling objects, it is only insofar as causal relations underlie them that I am able to believe these relations really to exist (pp. 73–74). Thus, when it comes to explaining reasoning in matters of fact and real existence, one has no choice but to focus on the relation of cause and effect, as "the only one, that can be trac'd beyond our senses, and informs us of existence and objects, which we do not see or feel" (p. 74).

ANALYSIS OF CAUSE AND EFFECT. Hume identified four constituents crucial to the idea of cause and effect: objects relatable as cause and effect must be distinct in the sense specified in the separability principle; they must be contiguous in time and (where the objects concerned are spatial) in place; the cause must precede the effect; and there must be a necessary connection between them.

Since the first three are fairly straightforward, Hume focused on necessary connection, with an eye to clarifying the idea by tracing it to its originating impression.

To understand why Hume proceeded as he did in this matter, the inherently paradoxical character of the idea of a necessary connection between distinct existents must first be taken into account. It stipulates a necessary connection between the existence of items presupposed as distinct. For example, one does not consider valleys and mountains candidates for terms of a causal relation because their necessary connection is merely conceptual, incorporated into the ideas themselves: Valleys cannot be conceived to exist in the absence of mountains and vice versa. By contrast, fire and smoke qualify as candidates for terms of a causal relation precisely because each can be conceived to exist without necessitating one to conceive the existence of the other. But there lies the rub: If to conceive them as distinct is to conceive the existence of the one to be possible even in the absence of the other, and to conceive them as necessarily connected is to conceive the existence of the one to be impossible in the absence of the other, then their combination in a single concept seems self-contradictory.

The general causal maxim. By far the most important illustration of the unintelligibility of the notion of necessary connection is Hume's analysis of the general causal maxim that everything that begins to exist must have a cause of its existence (1978, pp. 78–82). While recognition of the contingency of any determination in accordance with the maxim was a commonplace among pre-Humeans—that this specific thing causes that one—the truth of the maxim itself—that everything that comes into existence must have some cause—was taken to be an intuitively certain necessary truth, and so "one of those maxims, which tho' they may be deny'd with the lips, 'tis impossible for men in their hearts really to doubt of" (p. 79). Still, for Hume, the notion that the general maxim is a matter of knowledge rather than probability is easily refuted by a simple consideration of the concept of necessary connection itself. Its presupposition that the objects to be related in it are distinct already of itself implies the possibility that each of the objects can be conceived to exist in the absence of the other (pp. 79–80). Since even so much as a single conceivable exception is sufficient to show that a general proposition is not knowable intuitively or demonstrably, Hume concluded that the certainty of the general causal maxim is of a completely different nature, consisting not in any necessity of thought (relation of ideas) but in irresistible feeling (great force and vivacity), founded on experience and rooted in the nature of human (and much nonhuman-animal) associative psychology (pp. 82 and 172; Kant rightly recognized in this result a challenge to the possibility of metaphysics itself).

THE ORIGIN OF THE IDEA OF NECESSARY CONNECTION. A source of the idea of necessary connection in the objects present to one in sensation or reflexion is precluded by the fact that all perceptions as such conform to the separability principle, and so are "distinct" in the sense implying that it is always possible to conceive any one to exist in the absence of any other, or all others. Accordingly, Hume sought the origin of the idea in the experiencing subject and the ways it regards its objects, and, in particular, in the acts and affects incident to customary transitions from impressions to ideas (1978, pp. 165–166). When one object is found by experience to constantly succeed another, a habit is formed so that when one of them is present in sensation or reflexion, it straightaway brings to mind its constant concomitant, and one not only conceives it but believes it really to exist. The facility of this transition, with the force and vivacity felt in the conception of the idea when the transition to it is from an impression, constitutes the sole and entire content of the impression-of-reflexion original of the idea of necessary connection (1999, p. 145). To be sure, a projective illusion induces one to ascribe the impression of reflexion immanent to associative imagination to the objects it considers (1978, p. 167). Nevertheless, the necessity of causes is never anything but a subjective necessity felt in the mind that considers objects, and it is in this sense that the "necessary connexion betwixt causes and effects," and "the transition arising from the accustom'd union … are, therefore, the same" (p. 165).

Since Hume defined causal necessity both as a philosophical relation, in terms of constant precedence, and as a natural relation, in terms of customary association, many interpreters have supposed that the former has a meaning and scope of application unrestricted to associative imagination. Against this, one should note that, for Hume, (1) the idea of necessary connection is an essential element in all ideas of causal relations, (2) constant precedence as such does not include an idea of necessary connection, (3) the only source from which the idea of a necessary connection can be derived is customary association, and (4) ideas can never represent any objects other than those from which they are derived.

Accordingly, the only thing that can distinguish philosophical causation from constant precedence is the addition of the idea of necessity derived from customary

association, so that the necessity that "makes an essential part" of both definitions of causality is "at bottom the same" (1999, p. 160). This means that philosophical causation owes its influence on one's thoughts and actions entirely to its inclusion of a content no less bound up with conscious mind than pleasure, fear, or love; and to forget this by attempting to apply causal concepts directly to objects, apart from "that determination of the mind, which is acquir'd by custom," is to "either contradict ourselves, or talk without a meaning" (1978, p. 267).

EMPIRICAL RATIONALITY. In matters of fact and real existence, reasoning, as Hume understood it, is a transition in thought from a more vivid impression or idea to a less vivid idea in which the latter is conceived with more vivacity because of the relation the transition effects between them (where facility feeling is the essence of the relation). Since, in Hume's view, the enlivening of ideas primarily depends on their association with impressions, and since causal relations far exceed any other in their ability to enliven ideas to the point where they approach the vivacity of impressions, customary transitions from impressions to ideas are at once the source of the impression originals of ideas of necessary connection and the template of all empirical reasoning. This is just to say that the one indispensable item of evidence in any inferential matter of fact or real existence is an impression of necessary connection. For, in the absence of such an impression (maximally vivid perception), there could be no belief that an idea is connected to an impression in the manner requisite to enliven it, with the consequence that the impression would not then be regarded as a reason to affirm the idea. Thus, to explicate the nature of empirical reasoning, and to distinguish reasonable (factually justified) cases of reasoning from unreasonable ones, Hume undertook an investigation into the causes of such impressions.

The nonrational basis of empirical reasoning. The principal, and the most efficacious cause, of impressions of necessary connection is frequent experience of the items connected in them in an unvarying sequence—termed *constant conjunction* by Hume. As the evidence for this causal connection is itself a remembered constant conjunction (between relations of constant conjunction and subsequently felt impressions of necessary connection), Hume queried whether one infers the necessary connection from experience "by means of the understanding or of the imagination; whether we are determin'd by reason to make the transition, or by a certain association and relation of perceptions" (1978, pp. 88–89).

Nothing in Hume's philosophy has received more attention than his solution to this question (usually called the problem of induction). He began by premising that if reason were responsible for the conclusion that a necessary connection exists whenever a relation of constant conjunction is found, then the inference would be grounded on the "principle, *that instances, of which we have had no experience, must resemble those, of which we have had experience, and that the course of nature continues always uniformly the same*" (1978, p. 89). The question thus becomes whether one's belief in this uniformity principle is itself a product of rational argument, demonstrative or probable, or whether the implicit confidence one places in it derives from a different, nonrational source (associative imagination). Demonstrative reasoning (knowledge) is easily ruled out, since "[w]e can at least conceive a change in the course of nature" and "[t]o form a clear idea of any thing, is an undeniable argument for its possibility, and is alone a refutation of any pretended demonstration against it" (p. 89). Hume next excluded probable reasoning on the ground that it cannot be the source of a belief it presupposes:

> We have said, that all arguments concerning existence are founded on the relation of cause and effect; that our knowledge of that relation is derived entirely from experience; and that all our experimental conclusions proceed upon the supposition, that the future will be conformable to the past. To endeavour, therefore, the proof of this last supposition by probable arguments, or arguments regarding existence, must be evidently going in a circle, and taking that for granted, which is the very point in question.
>
> (1999, p. 115)

Since the past can only matter to one in forming of beliefs about the present or future in probable reasoning if one already believes the future is conformable to the past, one's belief in this uniformity must have a basis other than probable reasoning. According to Hume its basis is none other than customary association, which instills in one a belief in the uniformity of nature long before one has left one's cradle and determines the reasoning of brute beasts in the same way it does humans (1999, p. 118 and 1978, p. 178).

Philosophical and unphilosophical probability. When conjunctions of perceptions are remembered to be less than constant, one's evidence of necessary connection falls short of the certainty of proof. How much credence should one accord each of the competing causes and/or effects? That is, what constitutes reasonable belief here?

According to Hume the natural procedure is also the rational one: the accumulated belief (vivacity feeling) is distributed among the contrary causes or effects according to their relative constancy in past experience, subtract the lesser from the greater, and accord only so much credence (vivacity) to the latter as remains (1978, pp. 132–140). In other words, experience shows that one proportions belief in causal connections according to the constancy of the conjunction of the items concerned in them in the past and that this experience is so natural and universal that such proportioning has in all times and places been regarded as the hallmark, if not indeed the essence, of reasonable belief, or philosophical probability.

Of course, Hume was well aware that experience shows there to be many other causes of impressions of necessary connection than experienced conjunction and that these causes sometimes prevail over the evidence of experience: the ebb and flow of passions, calculations of interest and gain, laziness, hastiness, credulity, the persistence of tenets in education that have ceased to be proportioned to experience, and so on. One may be tempted to object that Hume's distinction between such unphilosophical (unreasonable or even irrational) reasoning and reasonable inferences proportioned to experience is arbitrary, since both alike are functions of feeling (vivacity transference effected by facile transitions of thought). Was he simply endeavoring to reflect linguistic practice? More likely, Hume's distinction derives from the account of the origin of impressions of necessary connection on which all causal inference depends. Experience is the natural and original cause of ideas of causal relations: It operates most constantly and steadily on the imagination and is most inseparable from the nature of that faculty (compare to 1978, p. 280). So, even in the absence of any objective or normative paradigm of rationality, nature itself, on Hume's account, sets experience at the foundation of empirical rationality.

A WORLD IN IMAGINATION. In denying that one has intuitive or demonstrative knowledge of the truth of the general causal maxim, Hume at the same time affirmed that one has another kind of certainty that everything must have a cause of its existence, arising from observation and experience (1978, p. 82) and consisting in the great vivacity of one's idea of the relation of any beginning of existence (thing, action, or state) to something precedent from which its existence follows by necessity (p. 172). The consequence is an unquestioning assumption, in any particular instance, that a cause inferred for a given effect is itself the effect of some other cause. For example, if the sight of smoke makes me think and believe that there is a fire in the hall outside, I at the same time take for granted a cause of this fire, a cause of this cause, and so on. If I reflect on this regress, I might attribute the fire to the frayed wiring I saw earlier, this to the gnawing of mice, the presence of mice in the building to the construction going on next door, the construction to the renovation plans of the new owner, the purchase of the building to the death of the old owner and the greed of the new one, and so on. But even if my theory should turn out to be mistaken (it was arson), I still remain absolutely certain of the existence of some chain of causes leading to the fire.

Since similar causal chains, with fewer or more of the blanks filled in, are taken for granted in respect of every beginning of existence, the space and time of real things demarcated by the purview of one's senses and memory comes to be dwarfed by the sphere comprised of the realities one infers to exist by means of customary association in relations of cause and effect:

> 'Tis this latter principle, which peoples the world, and brings us acquainted with such existences, as by their removal in time and place, lie beyond the reach of the senses and memory. By means of it I paint the universe in my imagination, and fix my attention on any part of it I please. I form an idea of ROME, which I neither see nor remember; but which is connected with such impressions as I remember to have received from the conversation and books of travellers and historians. This idea of Rome I place in a certain situation on the idea of an object, which I call the globe. I join to it the conception of a particular government, and religion, and manners. I look backward and consider its first foundation; its several revolutions, successes, and misfortunes. All this, and every thing else, which I believe, are nothing but ideas; tho' by their force and settled order, arising from custom and the relation of cause and effect, they distinguish themselves from the other ideas, which are merely the offspring of the imagination.
>
> (P. 108)

INDIVIDUALS

Hume explicated one's ideas of complex individuals (bodies and minds), both at a time (which he called simplicity) and over time (identity), as fictions resulting from failures to distinguish relations of genuine individuals from these individuals themselves. While granting that, in appearance, these fictitious individuals do not resemble

genuine ones, he insisted that their feeling to the imagination in contemplating its objects is so similar in the two cases, and the associative influence of the resemblance relation so strong, that one affirms their simplicity or identity even in the face of contrary appearances (1978, pp. 202–204 and 253–254).

Hume opted for associationist explications of these ideas because he could find no way to make sense of complex individuals objectively. The only kind of simplicity one is capable of conceiving in objects (impressions and ideas) is incompatible with complexity and manifestly different from it: Perceptions may be simple, in which case there must be only one, or complex, in which case there must be more than one, but since they cannot be both one and more than one at once, the notion of a complex individual is, strictly speaking, unintelligible. The predicament is even worse when it comes to the identity of an object over time. Since "all impressions are internal and perishing existences, and appear as such" (1978, p. 194), no idea can be copied from them that is not of existents "interrupted, and perishing, and different at every different return" (p. 211). Hume took this so far as to insist that duration is inconceivable apart from succession, and so can never be represented otherwise than as a multiplicity (p. 37). To be sure, one can represent something as the same as itself at one and the same time; but this is unity, not identity (pp. 200–201). Thus, unlike simplicity, the notion of identity seems to premise a combination of unity with number that, objectively at any rate, seems unintelligible.

PERFECT IDENTITY. While there may be nothing objectively to distinguish the presence to consciousness of a single continuing existent from a succession of distinct qualitatively identical fleeting existents, on the subjective side there is a feeling that suffices to mark a difference:

> The faculties of the mind repose themselves in a manner, and take no more exercise, than what is necessary to continue that idea, of which we were formerly possest, and which subsists without variation or interruption. The passage from one moment to another is scarce felt, and distinguishes not itself by a different perception or idea, which may require a different direction of the spirits, in order to its conception.
>
> (1978, p. 203)

Presumably, one's mind might have been so constituted that, instead of being all but effortless, the act of successively repeating the same idea might have required great exertion and a continuous redirection of the spirits to effect it. In that case, however, the change (succession of the distinct) would be as unmistakable here as with a kaleidoscopically varying flux. Alternatively, instead of being "scarce felt," contemplating a qualitatively invariant succession might involve no feeling at all. Still, in that case, there would be nothing to induce the imagination to confuse the observation of a continued, invariant sequence of perceptions with interrupted or variable ones and Hume's account of complex individuals could not even get off the ground. Thus, the original of the idea of what Hume called *perfect identity* lies not merely in the objects contemplated but also in the sustained affective disposition of the imagination in successively reproducing the same idea.

THE IMPERFECT IDENTITY OF BODY (CONTINUED AND DISTINCT EXISTENCE). Perfect identity is terminated by the first interruption or variation sufficient to necessitate a new direction of the spirits. However, "a succession of related objects places the mind in this disposition, and is consider'd with the same smooth and uninterrupted progress of the imagination, as attends the view of the same invariable object" (1978, p. 204). Since the very nature or essence of relation is facility, a succession of a single relation of ideas (facility feelings) produces the same continuity of affective disposition distinctive of a successive repetition of the same idea, and so leads one to confound them (= imperfect identity). In the case of bodies (continued and distinct existents) the principal relation is resemblance:

> We find by experience, that there is such a *constancy* in almost all the impressions of the senses, that their interruption produces no alteration on them, and hinders them not from returning the same in appearance and situation as at their first appearance. … This resemblance is observ'd in a thousand instances, and naturally connects together our ideas of these interrupted perceptions by the strongest relation, and conveys the mind with an easy transition from one to another. An easy transition or passage of the imagination, along the ideas of these different and interrupted perceptions, is almost the same disposition of mind with that in which we consider one constant and uninterrupted perception. 'Tis therefore very natural for us to mistake the one for the other.
>
> (p. 204)

To be sure, the identity the imagination wishes to ascribe to these appearances directly conflicts with the new direction of the spirits necessitated by their interrupted

appearances. Since these interruptions "are so long and frequent, that 'tis impossible to overlook them; and as the *appearance* of a perception in the mind and its *existence* seem at first sight entirely the same, it may be doubted, whether we can ever assent to so palpable a contradiction, and suppose a perception to exist without being present to the mind" (1978, p. 206). Given that one does so virtually every moment of one's life, the question for Hume was not whether but how one reckons with the contradiction. He found the answer in the associative nature of the idea of the mind to which perceptions appear. If the mind is not, as most of Hume's predecessors believed, a real substantial unity on which perceptions essentially depend, but something conceivable only associatively, as a "connected mass of perceptions," then "there is no absurdity in separating any particular perception from the mind" (p. 207). That is, if, in accordance with the separability principle, one can conceive any perception to exist in the absence of any other or even all others, then one can conceive any perception to exist in the absence of the mind if the mind is, indeed, just another perception (namely, a complex idea produced in associative imagination).

By calling such absences interruptions in its appearance, one can attribute to the perception a reality independent of the mind. Of course, since the separability principle holds of all perceptions without exception, this is something one is capable of doing with any perception whatsoever—smells, pains, fears, desires, volitions, and thoughts no less than spatial (visible and tangible) objects. That one only exercises this conceptual capacity in the case of spatial objects is due solely to the fact that they alone exhibit the constancy requisite to produce resemblances sufficiently strong between interrupted perceptions to generate an affective disposition liable to be mistaken for perfect identity.

Even so, the distinction between the appearance and reality of spatial objects employed here is merely external (relative). Consequently, it can only disguise, not eliminate, the feature that sets up the palpable contradiction in the first place: the appearance and reality of perceptions are one and indistinguishable. Given that "all impressions are internal and perishing existences, and appear as such," the distinct, continued existence one accords to visual and tactual impressions has nothing whatsoever to do with either the reality or the appearance of these perceptions, and everything to do with operations of the imagination that considers them. That is, the only idea one is capable of forming of the identity of bodies is inseparably bound up by content with the subjective acts and affects of association imagination, and so is fictitious through and through.

In designating body a fiction, it was by no means Hume's intent to imply that one does or even can doubt its reality. For not only is the fiction rooted in fundamental principles of human nature, it is in effect self-confirming. The memories whereof ideas of bodies consist are, in general, one's most vivid ideas. Since the effect of the fiction of a continued existence is to unite the scattered memories of resembling appearances in a single idea, their vivacity feelings are pooled together in that idea, thereby producing the strongest conviction in the real existence of the continued existent thereby conceived (1978, pp. 208–209). For this reason, "[w]e may well ask, *What causes induce us to believe in the existence of body?*, but 'tis vain to ask, *Whether there be body or not?* That is a point, which we must take for granted in all our reasonings" (p. 187).

THE SIMPLICITY OF BODY: THE IDEA OF SUBSTANCE. Hume explicated the idea of simplicity of bodies (their individuality at a time) by means of an associative fiction closely analogous to that responsible for one's idea of their identity. The appearance and reality of one's perceptions are ignored because of the powerful influence on the imagination of its own affective disposition when it contemplates coexistent perceptions bound together by customary associations of contiguity and causality:

> The connexion of parts in the compound object has almost the same effect, and so unites the object within itself, that the fancy feels not the transition in passing from one part to another. Hence the colour, taste, figure, solidity, and other qualities, combin'd in a peach or melon, are conceiv'd to form *one thing*; and on account of their close relation, which makes them affect the thought in the same manner, as if perfectly uncompounded.
>
> (1978, p. 221)

Here, too, the contradiction between one's feelings and the manifest difference in appearance between a genuinely simple object and a body—that is, the distinctness in the latter, according to the separability principle, of the color from the taste, these from the visible figure, these in turn from its tangible solidity, and so on—is too pronounced to ignore, and so must be palliated by some fiction, even if the contradiction can only be disguised thereby, not eliminated. Accordingly, we "feign an unknown something, or *original* substance and matter, as

a principle of union or cohesion among the qualities, as what may give the compound object a title to be call'd one thing, notwithstanding its diversity and composition" (p. 221).

THE IMPERFECT IDENTITY OF THE MIND (SELF AND PERSON). In the case of the mind, one is induced to attribute identity in the face of recalcitrant appearances more by causal relations than by resemblance:

> As to *causation*; we may observe, that the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. Our impressions give rise to their correspondent ideas; and these in turn produce other impressions. One thought chaces another, and draws after it a third, by which it is expell'd in its turn.
>
> (1978, P. 261)

One is a witness continuously, almost from the beginning of conscious life, to impressions causing idea copies of themselves to be formed, of these ideas being the occasion of further thoughts, passions, desires, and/or volitions, these in turn causing copies of them to be formed, and so on. One's perceptions may be subject to constant change, but never, even for a moment, is a causal relation between them of some kind absent from one's purview. Since "the very essence of these relations consists in their producing an easy transition of ideas" (1978, p. 260), the facility feelings incident to contemplating an unvarying, uninterrupted series of causal relations signify the presence in one of an unvarying, uninterrupted affective disposition. The strength of this disposition, with the strength of the feeling of its resemblance to the affective disposition incident to perfect identity, leads one to attribute an identity to this system of causal relations (pp. 253–254), notwithstanding that, on the side of the appearances, one's perceptions are "a perpetual flux and movement" and nothing "remains unalterably the same, perhaps for one moment" (pp. 252–253). (Hume's account of the simplicity of the self is essentially the same as that of body [p. 263].)

Second thoughts. Hume's explication of the idea one has of oneself thus shows it to be no less fictitious than that of the idea of external objects: nothing "really binds our several perceptions together," it merely "associates their ideas in the imagination"; one never observes any "real bond" among them, one "only feel one among the ideas we form of them" (1978, p. 259). Still, by excluding all real relations from the account of the self, Hume eventually came to realize that he had no way to "explain the principles, that unite our successive perceptions in our thought or consciousness" (appendix published with the second volume [book 3 1978, p. 636] of the *Treatise*). Hume saw no way out of this quandary, nor did he ever return to this topic in any subsequent work.

SKEPTICISM

Was Hume a skeptic? Though generally reputed to be among the most extreme of skeptics, the question is not so absurd as it may seem. If a skeptic is one who doubts or even rejects the use of reason as a means of arriving at truth, then Hume was no skeptic. So long as one is guided by intuition in one's inferences in mathematics and by experience in matters of fact, "Our reason must be consider'd as a kind of cause, of which truth is the natural effect" (1978, p. 180). Furthermore, Hume recognized that many beliefs are pointless to doubt because one is literally incapable of disbelieving them or not taking them for granted in all one's reasoning, including such philosophically contentious topics as the existence of external objects and the self, space and time, and the necessity of a cause to every beginning of existence.

Consequently, many commentators have come to regard Hume's skepticism as considerably more moderate and narrowly focused than traditionally supposed. For them, what makes Hume a skeptic is that he supposed one's ineliminable beliefs skeptically unassailable not because they are founded on reasons too strong to be undermined by skeptical argument but because they are not founded on reasons at all. It is nature, not reason, that has determined one to believe certain things. Nor is reason, when understood as Hume would have one do, capable of supplying these beliefs with a rational basis immune to skeptical assault.

The problem with this view is that it focuses almost exclusively on beliefs to the neglect of their ideational contents. If Hume did indeed deem belief in the existence of body skeptically unassailable, it must also be remembered that psychological processes—the actions and affects of associative imagination—are not merely essential to the formation of the idea in which this belief is reposed but also contribute elements essential to its content (i.e., apart from which bodies are inconceivable), and limit its application accordingly. Indeed, what is perhaps most distinctive of Humean skepticism is the conceptual dimension, in which association supplies subjective-psychological surrogates, as the only way around the "contradictions which adhere to the very ideas of matter,

cause and effect, extension, space, time, motion; and, in a word, quantity of all kinds" (1992, pp. 189–190).

For Hume, it is impossible even so much as to conceive these things without incorporating into one's ideas of them contents copied from impressions as irreducibly subjective as pain or disgust. What does it matter that the belief (vivacity) conferred on these ideas renders them skeptically unassailable if the ideas themselves are of such a nature that no skeptic would think to contend against them? One's reliance on associative imagination for the content of one's ideas comes at a price. If, for example, "we suppose necessity and power to lie in the objects we consider, not in our mind, that considers them," then, apart from this, "it is not possible for us to form the most distant idea of that quality" (1978, p. 167). This restriction on the scope of application of concepts so fundamental to human understanding as causation and body to the purview of a suitably constitutive experiencing mind unquestionably qualifies as a form of extreme skepticism.

VARIETIES OF HUME'S SKEPTICISM. When Hume himself characterized his philosophy as skeptical, he meant that it abounds with "discoveries concerning the weakness and narrow limits of human reason and capacity" (1999, p. 145). Although virtually everything in Hume's philosophy is directed to this end, among the arguments, analyses, and approaches to which he explicitly appended the term *skeptical*, three seem most deserving of being singled out.

Skepticism with regard to reason. After explicating empirical rationality as inferential belief proportioned to the evidence of past experience in *Treatise* I.iii, Hume advanced an argument in I.iv.1 to show that the result of adhering always and only to the canons of empirical rationality leads inexorably to the conclusion that "all is uncertain, and that our judgment is not in *any* thing possest of *any* measures of truth and falsity," so that "the understanding, when it acts alone, and according to its most general principles, entirely subverts itself, and leaves not the lowest degree of evidence in any proposition, either in philosophy or common life" (pp. 183 and 267–268). While most commentators consider his reasoning fallacious, Hume himself clearly deemed it impeccable and irresistible on any conception of empirical rationality, his own included (pp. 184–185). What interested him was why the argument nevertheless fails to convince. The reason he offered is that "[n]ature, by an absolute and uncontroulable necessity has determin'd us to judge as well as to breathe and feel" (p. 183).

More particularly, the argument lacks the affective force on which all relation (facility) and belief (vivacity) depend, "Where the mind reaches not its object with easiness and facility, the same principles have not the same effect as in a more natural conception of the ideas; nor does the imagination feel a sensation, which holds any proportion with that which arises from its common judgments and opinions" (p. 185). Vivacity (belief) follows facility (relation); so even if experience and custom support a certain inference, if for some reason, however trivial, facility feeling fails, vivacity will as well. And the circumstance in which understanding would subvert itself is a case in point:

> We save ourselves from this total scepticism only by means of that singular and seemingly trivial property of the fancy, by which we enter with difficulty into remote views of things, and are not able to accompany them with so sensible an impression, as we do those, which are more easy and natural.… We have, therefore, no choice left but betwixt a false reason and none at all.
>
> (P. 268)

Skepticism with regard to the senses. However impossible it may be for one in ordinary life not to believe in the distinct, continued existence of the bodies one sees and touches, only "a very little reflection and philosophy is sufficient for us to perceive the fallacy of that opinion" (1978, p. 210). Still, even if the more philosophical part of humankind recognizes this, they typically attempt to salvage the common opinion by arguing that unperceived objects correspond to perceptions that resemble them in various particulars but not their internal perishing existence. Many interpreters believe that Hume judged the philosophical view capable of sustaining skeptical scrutiny. This, however, is hard to credit in the face of his assertion that the philosophical view "contains all the difficulties of the vulgar system, with some others, that are peculiar to itself" (p. 211). If it contains all the difficulties, how can it withstand skeptical scrutiny any better? Hume's skepticism regarding the vulgar view centered on the content of the idea of a distinct, continued existence: the indispensability to it of something of the nature of an affective disposition (as is true of the idea of identity itself, this being the only means whereby the manifest differences between an interrupted or varying existence and a genuine identity can be overlooked and the two confounded).

Since the idea carries this content with it into all its applications, Hume cannot have exempted its philosophical employment from the same skeptical arguments to

which he subjected its vulgar. Indeed, because the philosophical view was erected in express opposition to the verdict of the most powerful, deep-seated natural human psychological propensity to believe in the distinct, continued existence of immediately perceived visible and tangible objects (sensations), only the weakest, most ephemeral conviction can be accorded to the philosophers' objects (p. 213). Finally, Hume contended that philosophers, having no means of conceiving their would-be objects except their own perceptions, in effect do no more than "arbitrarily invent a new set of perceptions" (p. 218). If, to avoid this implication, they suppose their objects to be specifically different from everything one can conceive, the result will be an "unknown, inexplicable *something* … a notion so imperfect, that no sceptic will think it worth while to contend against it" (1999, p. 203).

Academic, or mitigated, skepticism. Despite the extremity of the skepticism resulting from the "deficiency of our ideas" (1978, p. 267), Hume saw fit to describe his philosophy as an exercise in "mitigated scepticism" (1999, pp. 207–211). A skepticism qualifies as such if, instead of advocating the rejection of reason in all its forms, it counsels one to reject all abstract reasoning other than mathematics, and all reasoning regarding matters of fact and experience that is not carefully and precisely calibrated to accord with the deliverances of experience.

Does Hume's own philosophical reasoning meet these criteria? It was because the empirical investigation of human understanding turns up no evidence of any other faculties besides sense and imagination that he endeavored to account for all the phenomena of perception, judgment, and reasoning (mathematics included) in terms of their operations. And it was because the only empirical source to which ideas of causal connection, substance, real existence, space, time, and the mind could plausibly be ascribed as associative imagination that he was compelled to conclude that even one's most basic, indispensable concepts of objects incorporate an ineliminably subjective element of feeling into their content (facility and vivacity). To be sure, with the understanding thus transformed (in part) into an organ of feeling, Hume's philosophy became the first to set reason on a par with pleasure and pain, passions, desires, and everything else previous philosophers had denigrated as belonging to the baser, animal part of human nature; and this may seem skeptical indeed. But since his conclusions are fully consonant with the strictures of a mitigated skepticism, he could at least be confident that his books would not be incinerated by anyone answering his call to "commit to the flames" any volume that fails to respect them.

THE WILL

Will is "the internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind" (1978, p. 399). There is no implicit proposition the affirmation of which constitutes the act of volition. Volitions, for Hume, are not ideas or manners of conceiving, but feelings, felt excitations to mental or physical action. They are full-fledged perceptions (impressions of reflexion) in their own right, distinct from all others under the separability principle, capable of existing in complete isolation (p. 625). As such, they are completely indefinable: like flavors, to know volitions—to be able to form (copy) clear ideas of them—it is necessary to have the corresponding impressions; to lack the impressions is to be completely ignorant of will, to be unable to form even the most obscure idea of it.

With nothing more to be said of the will *per se*, Hume focused on the causes of its actuation. Nothing precludes reason from doing so since here, as always, "to consider the matter *a priori*, any thing may produce any thing" (1978, p. 247). Still, as a matter of fact, one finds "that reason alone can never be a motive to any action of the will" (p. 413). Convinced by reason that I am about to be devoured by a ravenous beast, for example, I would be completely indifferent to the fact, and not be provoked by this belief to any exercise of will, without the mediation of some passion in response to (caused by) the belief. Indeed, if human nature was such that being devoured by the beast was one of our fondest desires—because, say, passing through the digestive tract of a beast of that species was indispensable to reproduction—then this belief, with the passion, would excite actions to facilitate our capture. Alternately, our passionate response to the belief might be as tepid as that of a fifth grader to his or her belief regarding the result of the fifteenth of a series of long-division homework problems, so that we merely yawn at the imminent prospect of being devoured. Only passions actuate the will. Reason, according to Hume, is neither a necessary nor sufficient to do so.

For similar reasons, Hume argued that reason can never directly oppose, curb, or in any way act as a counterweight to the actuation of the will by passions. It can do so only indirectly, by giving rise to some new passion, as when it informs one that the object of one's desire is unattainable, or attainable only by a different course of action, whereupon it will produce an aversion to counter, or a desire to override, the existing passion. Conse-

quently, when one speaks of "sweet reason" prevailing over "brute passion," it is not passionless, volitionally impotent, reason that is being invoked, but other, calmer passions. Their gentleness should not, however, be confused with weakness:

> 'Tis evident passions influence not the will in proportion to their violence, or the disorder they occasion in the temper; but on the contrary, that when a passion has once become a settled principle of action, and is the predominant inclination of the soul, it commonly produces no longer any sensible agitation. … We must, therefore, distinguish betwixt a calm and a weak passion; betwixt a violent and a strong one.
>
> (1978, PP. 418–419)

Is there such a thing as a rational passion? According to Hume, no. For even though a belief can be the invariable cause of a certain passion, passions are one and all original existences: none of their features are copied from the ideas that cause them or in any way derivable from them (1978, p. 415); and even when a passion has an object—as pride takes the idea of oneself for its object and love the idea of someone else—the object remains distinct (by the separability principle) from the passion itself, and only becomes an object to it by the mediation of some feeling of pleasure, such as that given by the beauty of the beloved or the opulence of a house that has passed into one's ownership (p. 279). Passions are therefore never rational in and of themselves; and since experience shows that only passions can actuate the will, reason

> is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them. … 'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. 'Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an *Indian* or person wholly unknown to me. 'Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater, and have a more ardent affection for the former than the latter.
>
> (PP. 415–416)

GOOD/BAD AND PLEASANT/UNPLEASANT ARE INDISTINGUISHABLE. Since reason, considered apart from whichever passions its deliverances may provoke, leaves the will indifferent, it cannot be the source of any of one's ideas of good and bad. This means that nothing propositional in character (rule, maxim, principle) can be intrinsically good or bad: carnally, spiritually, aesthetically, or in any other way. Since the only place left to look for the impression originals of ideas of good and bad are pleasant and unpleasant feelings (sensations and passions), goods and ills must all be pleasures and pains of one sort or another (1978, p. 439). Thus, for Hume, the standards one applies in all one's value judgments have their origin exclusively in pleasant and unpleasant sensations or reflexions, and neither the goals of one's actions, the deeds themselves, one's volitions to perform them, nor the character of the person who wills can be supposed good or bad either intrinsically or in relation to any rule of conduct (maxim and principle) under which they fall; they are good or bad solely by virtue of the feelings that caused them and/or the feelings they arouse.

DENIAL OF FREE WILL. The question of freedom of the will takes on a different aspect according to how a philosopher analyzes volition. If one deems will and reason inseparable, as Berkeley did, and conceives of volition as the affirmation or denial of a proposition, like René Descartes, then any external cause that necessitates one to affirm or deny will be construed as a constraint on the freedom of one's will. But if, like Hume, one distinguishes reason from will and equates volition with a nonintellectual feeling of excitation to action (impression of reflexion), then a free will, unrestrained by any necessitating cause, would be one that acted blindly and randomly, unresponsive to one's desires and heedless of one's beliefs, and so is something rather to be dreaded. Thus, from his standpoint, it is fortunate that experience shows one will not to be free, but instead to act only when necessitated to do so by some passion, be it calm or violent, beneficial or destructive, responsive or unresponsive to the deliverances of reason.

Complementing Hume's denial of free will is his analysis of causal necessity in the operations of bodies as consisting of nothing more than facile transitions of thought from one perception to its customary conjunct. For this means that there is nothing "the mind can perceive, in the operations of matter, some farther connexion between cause and effect … that has not place in the voluntary actions of intelligent beings" (1999, p. 157). All there is to causal necessity is what one experiences in every facile transition from an impression to the idea of its usual antecedent or successor. Thus, Hume's necessitarianism does "not ascribe to the will that unintelligible necessity, which is suppos'd to lie in matter," but "ascribe[s] to matter, that intelligible quality, call it necessity or not, which does or must allow to belong to the will" (1978, p. 410).

Consequently, to prove that one is in practice a necessitarian, protestations to the contrary notwithstanding, he had only to assemble reminders that one naturally and inevitably draws on one's past experience of regularities in human voluntary behavior to predict the actions of minds in precisely the same way one does to predict the actions of physical objects (1999, p. 150). To object that one encounters contrariety in the human sphere and often finds the actions of minds puzzling and unpredictable is futile since the same is true in the physical sphere as well, nor does one infer the freedom of bodies from causal necessitation because of the contrariety one finds there.

Illusions of freedom. Hume identified several reasons why one nonetheless insists on supposing oneself to be free. First, by not distinguishing the will as effect from the will as cause, one confuses two different notions of freedom. The will is free as a cause to the extent the actions of one's body and mind are subject to its control, that is, causally necessitated by it. This is the freedom one would lose if one's body or mind became unresponsive to the will or responded only to some external control. By contrast, the will is free as an effect only if its action is not necessitated by any cause, including one's own passions and beliefs, and so acts at random. The latter is the kind of freedom no one wants and, on the evidence of experience, no one has. But it is precisely this sort that matters philosophically, since the other is not only compatible with universal causal necessitation but would not be worth having otherwise.

There is also a psychological illusion of freedom implicit in the idea of necessity itself. When one perceives two objects, one does not feel a causal connection between them unless and until one observes their similarity to past constantly conjoined objects between which such a connection is felt, and then transfer the idea copied from this feeling (the reflexive impression of necessary connection) to the objects presently before one. By contrast, when one is not an observer but a performer of actions, no such reflection occurs, and consequently no connection is felt between one's perceptions (1978, pp. 408–409). For example, if I believe someone has betrayed me, and I become enraged and smash a vase against the wall, I feel no causative forces necessitating my actions; it is only afterward, when I reflect on what happened, that I recognize the necessitation of my action by the passion and the passion by my belief. Even so, I am still apt to resist the claim that in so doing my will and action were no less necessitated than a body released from a height is necessitated to fall. But apart from the fact that "there is no known circumstance, that enters into the connexion and production of the actions of matter, that is not to be found in all the operations of the mind" (p. 404), this is simply to say I can reimagine the situation so that, instead of the vase, I hurled something else or nothing at all, or that I somehow stopped myself from becoming enraged in the first place. That is not the same as supposing my volition to have been unnecessitated. It only means that, given different antecedents, different causes would have necessitated something other than the action I performed under the circumstances that actually prevailed.

THE PASSIONS

Though Hume devoted as much of the *Treatise* to developing a theory of the passions as he did to the understanding, the former has never attracted as much attention as the latter has. This is regrettable. Hume's theory of the passions is the mirror image of his theory of understanding: just as he was able to show the understanding to be as much an organ of feeling as of thought by explaining its most basic and important operations in terms of principles of association, so, too, by showing how surprisingly far these same principles go toward explaining the operations of the passions, he was able to reveal a deeper, underlying affinity between reason and feeling that otherwise, apart from his associationist doctrine, must remain concealed. This fundamental unity of perceptions that, to all appearances, seem disparate, or even opposed, was surely prominent in Hume's mind when he compared the place of association in the science of man to that of universal gravitation in Newtonian science of nature. One may therefore hope that Hume's theory of passions will someday receive the same amount of careful study and attention that has hitherto been reserved for other topics in his philosophy.

DIRECT PASSIONS. Hume distinguished passions into two basic types: direct and indirect. Direct passions such as grief, joy, hope, fear, despair, and security arise immediately from some good or ill (pleasure or pain), or are themselves productive of good or ill (natural impulses such as punishing enemies and rewarding friends, as well as natural instincts such as hunger, lust, and other bodily appetites). Because their immediate cause or effect is some impression or idea of pleasure or pain, Hume could identify no role for the association of ideas in explaining their origin and only an occasional, incidental role for the association of impressions (where there is only association by resemblance). Nevertheless, he found a number of cases in which associative imagination proves crucial to

enable passions already present in the mind either to commingle (or not) or to oppose one another (or not).

INDIRECT PASSIONS. The passions of principal interest for Hume's associationist science are those he classified as indirect ideas and their associative relations are found to be causally essential to their production. The most fundamental indirect passions are pride/humility and love/hatred, but they also include ambition, vanity, envy, pity, and malice. These share a causation that takes the form of a "double relation of ideas and impressions" (1978, p. 286). Thus, an object causes a pleasure of some kind; if the object happens to be related to me by a strong enough relation, this relation of ideas (of the object to me), together with the pleasurable quality (impression) of the object, causes me to feel the resembling (because also pleasing) passion of pride (impression), whereas that same object, if productive of something unpleasant, will, given the same relation to me, cause the resembling (displeasing) passion of humility. Take away that object's relation to me, and I will feel neither pride nor humility in response to its pleasing or displeasing quality; take away its pleasing or displeasing quality and again I will feel neither passion. Consequently, pride and humility are found by experience to exist only in conjunction with an idea of myself, another object strongly related to (associated with) me, and some pleasing or displeasing quality related to (associated with) that object.

What differentiates love and hate from pride and humility is simply the object of the passion. For just as I take pride in my body or mind, or some object, insofar as it possesses some pleasing quality and has a strong relation to me—my looks, my brilliance, the imposing house I own, the beautiful painting I created, the coveted office to which I have been elected, and so on—so, too, I love or esteem someone else from precisely the same causes. Otherwise, these passions exhibit the same double relational structure.

Hume was well aware of the profusion of seeming counterexamples to this structure and spared no effort to rebut or deflect them. Still, to many, these efforts have something ad hoc about them, and Hume tends to be condemned for too rigid an adherence to theory in the face of recalcitrant phenomena. But much of this criticism may be due to a failure to appreciate the significance that double relations in question are associative in character, that is, their essence consists in facile transitions felt between impressions and ideas (1978, pp. 289, 309, 335–336, 378). This is never clearer than when, in the last three of Hume's "Experiments to Confirm This System" (pp. 332–347), he shows what seem to be counterexamples are really cases in which something interferes not with the relation considered abstractly (philosophically) but with the degree of facility felt in it, so that one or both of the relations requisite to produce an indirect passion are deprived of their associating quality, either by losing facility or because some opposing, even more facile transition prevails. Thus, when one factors in the affective dimension of Humean associationism, one can begin to appreciate Hume's evident excitement at the prospect of an explanatory principle that, for the first time, permits a systematic exposition of the human conative mind (pp. 346–347).

SYMPATHY. The compass of one's passions would be narrowly confined to those with whom one has close personal relations if sympathy did not overcome one's indifference by communicating to one the feelings of others and enabling these to arouse one's own feelings, whether they be strangers, those known to one only by reputation, persons long dead, members of far away societies, even characters in myth. Thus, sympathy plays a key role in the operation of the passions in the wider context of human society. Regarded from Hume's perspective, however, sympathy is simply an extension of the associationist principle into the societal sphere. For, in and of itself, it is just one among species of the general associationist operation of enlivening ideas related to impressions to the point where they approach or equal the vivacity of the impressions themselves; one can call it sympathy when it increases the vivacity of an idea related to the passion felt by another to the point where it equals or approaches the original impression (1978, p. 319).

MORALITY

Hume's approach to morality is of a piece with the rest of his philosophy. Are there specifically moral ideas, or does moral discourse have nothing in the only object ever present to one—one's perceptions—to confer objective meaning on its pronouncements? If there are ideas, then their content must be determined by tracing them back to their originating impressions: whether they have their source in the perception of some object in sensation or reflexion (impression) or in acts of associating ideas of these objects. With the origin of moral ideas determined, enough would become evident about their place in the cognitive and/or conative economy of the human mind to permit the discovery of the fundamental principles governing moral judgment and action.

The question whether causal discourse has a basis in the objects present to one's mind came down to the question whether one experiences nothing but constant conjunctions or whether there is something more—even if that something should turn out not to be the objectively real necessary connections one's discourse might lead one to expect. In the case of moral discourse the question that was decisive for Hume regarding its objective significance is whether one's experience of good and ill is limited to passions and desires, or whether there is, in addition, a source of distinctively moral ideas.

Hume's confidence that there is more to causal discourse than experienced constant conjunction stemmed from a conviction that, given only this, reality, for one, would be restricted to the narrow compass of the senses and memory. Where morality is concerned, his confidence in its ideational foundations seems to have derived from the abundant evidence of morally motivated actions: action undertaken not for selfish reasons, from partiality for those one loves, from dread of the consequences of not performing them, or for any identifiable purpose other than the sheer morality of it. Accordingly, in tracing ideas of moral good and ill to their origin, Hume's first task was to determine whether they derive from the features or relations of the objects immediately present to one in perception or, like ideas of necessary connection, from something felt in their contemplation.

MORAL IDEAS ARE COPIED NEITHER FROM OBJECTS NOR THEIR RELATIONS. For Hume, morality would count as objective if actions or things were moral or immoral prior to and independently of any course of reflection on them and, *a fortiori*, any feeling that arises only in the course of such reflection. For example, if willful murder were objectively immoral, then some impression embodying its immorality must exist to be copied in an idea. But what does one find when one considers such crimes objectively but a sequence of thoughts, passions, motives, volitions, and actions? The action itself is not immoral or else an avalanche would be immoral for taking the lives of skiers. That the action is voluntary does not of itself make it immoral or else lions would be guilty of immorality every time they killed. Nor does its immorality consist in the anger, greed, or other passion that determined the will, since these feelings are in themselves neither moral nor immoral. Finally, even if the course of reasoning that eventuated in the resolve to murder included an awareness that murder is wrong, its immorality, if objective, would derive not from this thought as such, but from the preexisting objective state of affairs recognized in it.

If not in the objects whereof willful murder consists, does its immorality reside in some relation of these objects discoverable by reason? Reason, as explicated by Hume, consists either in (intuitive or demonstrative) knowledge of the relations of ideas derived from objects or in belief (a vivid idea) regarding a matter of fact inferred from some other matter of fact. Against the former supposition, Hume argued that none of the knowable relations into which ideas can enter—resemblance, contrariety, degrees in quality, and proportions in quantity and number—seem capable even of distinguishing the moral from the nonmoral, much less the moral from the immoral.

If there is some other kind of knowable relation in which objective morality consists, Hume confessed to being ignorant of it. But even if there were, it would have to satisfy two conditions that seem impossible to meet. In the first place, to be a knowable yet genuinely moral relation, it could only relate two species of objects to the exclusion of all others: internal actions of the mind to external objects. Otherwise, internal actions of the mind that never eventuate in any deed could be moral or immoral, as could deeds with no mental components (thoughts and volitions). Still, so selective a relation of ideas seemed to Hume beyond the scope of what is intuitable or demonstrable by mere human minds. Second, even if such a relation did exist and were known, it would still remain for one actually to intuit or demonstrate its power to determine the will of every being possessed of a knowledge of it, divine no less than human. Since the components of the relation—knowledge and volition—are distinct perceptions, such determination could only take place via causal necessitation. Still, if Hume's analysis of causal connections shows anything at all, it is that no connection is ever intuitable or demonstrable "by the simple consideration of the objects," since "[a]ll beings in the universe, consider'd in themselves, appear entirely loose and independent of each other. 'Tis only by experience we learn their influence and connexion; and this influence we ought never to extend beyond experience" (1978, p. 466). Therefore, it seems that no moral relation can ever be knowable and vice versa.

Objective morality is also not discoverable by probable reason. Deeds objectively comprise thoughts, passions, volitions, and bodily actions. In which relation of these does its morality consist? Even if experiment revealed the existence of some hidden object, a neurochemical perhaps, that reliably tracked the distinctions one makes between the moral and nonmoral, and the moral and immoral, one's ideas of the moral and

immoral could still not be originally derived from such a source since, in and of itself, neurochemicals are just as nonmoral as any of the more obvious objects concerned in moral and immoral deeds. Thus, there is nothing rationally discoverable in the objects, and expressible by an "is" or "is not," that can lead one simply by reasoning to any properly moral recognition, expressible by an "ought" or "ought not" (1978, pp. 469–470).

THE SUBJECTIVE ORIGIN OF MORAL IDEAS IN INTERNAL SENTIMENT. With objects excluded as the source of moral ideas, Hume saw no alternative but to conclude that, like ideas of cause connections, they have their origin in something one feels in the act of contemplating objects. However, the exclusion of empirical reason as their source ipso facto precludes the facility and vivacity affects immanent to associative imagination. Instead, moral ideas originate in a species of impression of reflexion that is entirely independent of imagination. This, for Hume, is not to deny that experience shows that certain processes of thought are causally essential to moral impressions; it is only to say that these processes—by contrast with the impression originals of ideas of necessary connection and identity—contribute nothing to their content. As such, moral sentiments are distinct from these processes, and from every other perception, under the separability principle, and so might conceivably have arisen in total isolation from processes of thought, as hunger and sexual appetites do, or from causes different from those experience in fact reveals. The special status of the impression of reflexion source of moral ideas therefore derives not from any special authority intrinsic to these feelings themselves—they are simply one among many other varieties of pleasure and pain—but from the unique circumstances of their causation and the special place in one's life they derive therefrom.

THE CAUSATION OF MORAL SENTIMENTS. Experience reveals that moral sentiments are aroused only in the course of reflecting on the doings of human beings, specifically the mental characteristics responsible for their voluntary actions, and of these only those most firmly rooted in a person's character: the most efficacious and enduring characteristics of the identity that constitutes an individual human mind. This causation explains why moral feeling weakens or vanishes altogether when one contemplates actions not considered to be tests of character, because, say, their performance was prompted by an uncharacteristic whim, an excusable misjudgment regarding the facts, fever, disease, medicinal side effects, or involuntarily through some unavoidable external cause.

The causal structure of moral feeling resembles that of the indirect passions of pride/humility and love/hate in that it involves a double relation of impressions and ideas: an object (idea) related to a person (another idea) is the subject of some pleasant or unpleasant feeling (impression) that, because of the relation between the objects, gives rise to its resembling (pleasing or displeasing) moral feeling (another impression). Indeed, with the proviso that the causes of moral feelings are restricted to mental characteristics strongly related to the person, the pleasures and pains that arouse moral feelings prove to be precisely the same ones that arouse feelings of pride/humility in oneself and to love/hate toward others (1978, pp. 574–575), so that moral feelings may be regarded as "nothing but a fainter or more imperceptible" (p. 614) variety of these passions themselves.

There are, however, two further features of the causation of moral sentiments that distinguish them from indirect passions:

Moral feeling requires a general point of view. The indirect passions are invariably partial for or against their particular object (oneself or another). Moral sentiments, by contrast, tend to be felt only when "we fix on some *steady* and *general* point of view" in which one abstracts from "our situation of nearness or remoteness, with regard to the person blam'd or prais'd, and … the present disposition of our mind" (1978, pp. 581–582). Moral feelings are at their strongest (remembering that, for Hume, the strength of a sentiment is often inversely proportional to its violence) when the character of the person is viewed from the standpoint where it

> appears the same to every spectator. … And tho' such interests and pleasures touch us more faintly than our own, yet being more constant and universal, they counter-ballance the latter even in practice, and are alone admitted in speculation as the standard of virtue and morality. They alone produce that particular feeling or sentiment, on which moral distinctions depend.
>
> (P. 591)

From a personal perspective, one may be far more moved by the moral perfections of a best friend than by those of some moral giant of the past like Gandhi. Still, this delight is not moral sentiment. That feeling can arise only when one brackets out one's personal feelings for the person, whereon one cannot help feeling a far stronger feeling in contemplating Gandhi than one's friend (though

this is no guarantee that, when it comes to determining the will, one's moral sentiments will be strong enough to prevail over nonmoral ones).

Moral feeling requires sympathy. Since reason is impotent to determine the will and useless by itself to distinguish moral right from wrong, moral action is wholly at the mercy of moral sentiment. But if moral sentiments can arise only through their association with other pleasures or pains (in the context of a double relation of impressions and ideas), how is it possible for moral feeling to arise if it requires one to regard persons from a general point of view in which abstraction is made from everything determinative of one's present affective disposition? Hume's answer is that the capacity to remain affectively engaged depends on one's ability to sympathize with the persons one considers from a general point of view. Thanks to this societal variety of association, one continues to feel pleasure or displeasure from the consideration of the mental qualities rooted in the characters of persons one considers impartially. Since this permits the condition for the double relation of impressions and ideas requisite to produce moral sentiment is met, one then has only to contemplate the character from the general point of view requisite for moral sentiment for the pleasant or unpleasant feelings produced by sympathy to cause a corresponding pleasant or unpleasant moral sentiment.

VIRTUE AND VICE. Another way in which the impression of reflexion originals of moral ideas and those of ideas of necessary connection are alike is that, despite being subjective (felt only in contemplating objects), they are illusorily projected onto the objects contemplated and treated as though they were properties of the objects themselves (1978, pp. 167, 224–225). In the case of moral feelings, the objects that take on moral attributes are the mental characteristics whose agreeableness or disagreeableness cause moral feelings, whereon they count as virtues or vices: "*taste* … gives the sentiment of … vice and virtue … [and] has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises, in a manner, a new creation" (1998, p. 163).

Hume's typology of virtue of and vice. Hume distinguished four (nonexclusive) types of virtue:

(1) Mental qualities immediately agreeable to their possessors, such as skill, greatness of mind, cheer, equanimity in the face of adversity, and courage

(2) Qualities immediately agreeable to others, such as tact, delicacy, wit, and good manners

(3) Qualities useful to their possessors, such as intelligence, industriousness, skill, patience, and perseverance

(4) Qualities useful to others, such as gratitude, faithfulness, reliability, and charity

The pleasure one takes in these mental qualities in and of themselves is enhanced by the moral pleasure with which one responds to them, thereby adding a moral beauty to their original, nonmoral beauty. Similarly, the displeasure occasioned by their contraries is augmented by moral displeasure, and to their natural ugliness moral repugnancy is added. This, in turn, increases the effects these qualities have on other passions, above all the pride or love and humility or hatred felt on their account. Indeed, as mental qualities capable of stirring moral sentiments in one when considered with sympathy from a general point of view, pride/humility and love/hate now take on a moral value in their own right. Thus, if the pride another takes in his or her character is the effect of real virtues and proportionate to them, our contemplation of his or her pride (a pleasing quality) can only add to the pleasure we derive from contemplating the pleasing qualities in which he or she takes pride, whereas if his or her pride is a perverse pleasure deriving from morally repugnant mental qualities, his or her feelings about him- or herself can only increase the contempt we feel in contemplating those qualities.

Hume seems convinced that many of the qualities commonly deemed virtuous in his and other societies would not be considered virtues, or even be deemed vices, if people could overcome the distorting influences that prevent them from attaining a truly impartial, sympathetic perspective on human characters. Religious education, for example, can condition one to regard as virtuous the asceticism of monks, the fanaticism of zealots, or the credulity of the faithful—qualities of mind that would otherwise be certain to strike one as both repellant in themselves and harmful (1998, pp. 146–147). But, for Hume, the fact that miseducation, harsh conditions of life, and other factors can lead people to mistake virtues for vices and vices for virtues no more makes the one really the other than the fact that people are often influenced to discount or ignore past experience in their reasoning means that there is no real difference, rooted in human nature, between good and bad empirical reasoning. Nothing—interest, expediency, or serendipity—can make disagreeable or harmful mental qualities be, or appear to be, anything other than they really are. Nevertheless, outside influences may intervene to prevent one from attaining the constancy and universality of perspective, and/or the sympathetic engagement, requisite to

bring one's moral sense to bear on such disagreeable or harmful qualities and respond to them with the contempt they would otherwise naturally and universally inspire.

Of course, even if human nature ensures that universal agreement regarding virtue and vice is possible in the abstract, things are different when it comes to judging, in any particular instance, whether an action issued mainly from moral, immoral, or amoral motives, and in which proportions. Hume was keenly aware, in his capacity as philosopher no less than that of essayist or historian, that motives for particular actions can be complex and obscure, even to the agent, and that agreement in one's judgments regarding the morality may be impossible owing to differences in experience, education, access to information, and individual mental abilities. Matters are further complicated by the fact that moral sentiments must compete with other passions for influence on the wills of agents and the hearts of judges. Nonetheless, even if human nature cannot always reveal what one ought to do in each particular instance, Hume still deemed moral sentiment a universally valid standard accessible to anyone concerned to know what kind of person he or she ought to be; and, in this regard, moral sentiment serves as a dependable guide in moral decision making and judgment.

ARTIFICIAL VIRTUES. Institutions such as property, contracts, government, intergovernmental relations, and marriage must exist before the virtues of justice (the rightful possession of property), promise-keeping, allegiance, treaty-keeping, and chastity are even possible. A first precondition is that everyone, or nearly everyone, realize that they stand to benefit when every member of society, selves included, adheres to the rules requisite for these institutions to exist and flourish. Second, each person's recognition of their interest in everything that promotes universal adherence to these rules leads them to take pleasure in those mental qualities of persons that contribute most to making them just, faithful keepers of promises, loyal subjects, good treaty-makers and -keepers, and good husbands or wives. Only then, when reflecting on these pleasing qualities of persons from a general point of view, will each person's moral sense respond to these qualities with its own distinctive feeling, whereupon qualities originally prized only from self-interest at last come to elicit one's admiration as virtues.

What prompted Hume to classify these and other virtues as artificial rather than as natural, even though their origin in a recognition of the utility of certain mental qualities is no different from many natural virtues? Justice, for example, presupposes property, which, as an institution founded on a tacit convention, is, in Hume's view, thoroughly artificial, and in that sense unnatural. Although there is possession, property in the strict sense (as carrying an obligation not to hinder possession) does not yet exist in a state of nature, where something is mine if, by strength or wit, I can get it and keep anyone else who wants it from taking it. When goods are either too plentiful or too scarce, and generosity is confined to one's closest relations, there is no interest or intrinsic virtue to inhibit one from taking anything one wants from anyone else, even if one's need for it is not desperate. But when goods are neither too plentiful nor too scarce, a condition in which everyone takes whatever they want whenever they can prevents anyone from enjoying the benefit of secure possession of the goods they want or need for future use. The resulting dissatisfaction with the existing state of things thus creates an openness to change.

The problem is that it is not in my interest to leave anyone else in secure possession of my goods if I cannot be assured that the other person will do the same for me. This impasse is broken only with the establishment of a tacit convention, based on self-interest, of leaving others in possession of their goods provided they are prepared to leave one in possession of oneself. Moreover, since it is in the interest of all to be able to exchange some of the goods one has for others one needs or desires more, the convention of secure possession must also provide means whereby the goods of another can become one's own and vice versa, so that secure possession is transferred with them. Thus, through the artifice of tacit conventions, property in goods, over and above their mere possession, first comes into existence.

The reason that Hume classified justice in matters of property as an artificial virtue is that there is nothing about any good one desires to possess or retain, considered in and of itself, that can convey to one an idea of it as property. Property is unintelligible apart from established conventions, and conventions, however universal, tacit, and informal, are always artificial. For this reason, Hume denied that there is any natural interest or virtue in justice. Only after one has been inducted into the mysteries of the institution of property can one arrive at a recognition of one's interest in universal adherence to the rules requisite to maintaining it and so, *a fortiori*, come to prize as virtues the mental qualities most conducive to that interest. The same is true of every other virtue that presupposes human institutions founded on tacit conventions secured by a recognition of self-interest: contracts, laws, public offices, government, and so on. So, even though artificial virtues are no less genuine or powerful

expressions of moral sentiment than natural ones, Hume deemed them as unnatural to one's species as speaking English or paying in British currency.

RELIGION

One cannot be certain what Hume's actual views were with regard to belief in God. He was quite clear that he was not a Christian, and he seems to have regarded all religions as expressions of superstition, vestiges from less enlightened times that might (or might not) someday be superseded or wither away. However, Hume was also somewhat skeptical concerning contemporary atheistic conceptions. Matters are further complicated by the times in which he lived. Apart from legal sanctions (after a period of relative openness, new censorship laws began appearing in the late 1730s), a person's career prospects, social position, and tranquillity would be put in jeopardy by too open an expression of views liable to be construed as impious. For anyone unconcerned with mundane matters, zealous in the cause of atheism and enlightenment, desirous of being the focus of controversy, or sufficiently naive, these impediments might not matter. But Hume was not such a person. He was too worldly wise and fond of his place in society to bring down on himself the consequences of a frontal assault on the religious beliefs and institutions dear to the overwhelming majority of humankind. So, while many would agree with contemporary charges that his views on such matters as the general causal maxim and freedom of the will are implicative of atheism, Hume himself always professed the contrary (1978, pp. 409–10, 633n; 1999, pp. 160–164 1745/1967). And though his writings on religion seem to lead inexorably to the conclusion that a rational faith in God or revealed religion is an impossibility, he never ceased to proclaim that "the existence of a DEITY is plainly ascertained by reason" (1992, p. 280).

What is one to make of Hume's claims that his philosophy is consistent with, even supportive of, a rational belief in God? If these pretensions had been sincere, he would have had every reason to advertise the opinion, as other philosophers did who employed skepticism to humble reason to elevate faith. But one finds no evidence of this in his philosophizing beyond occasional brief asides, which seem too casually thrown out for one not to suspect that they are there merely to provide cover for his skeptical forays. It seems unquestionable that Philo, rightly regarded as Hume's principal mouthpiece in the *Dialogues concerning Natural Religion*, was not serving in that capacity when he declared that "[t]o be a philosophical skeptic is, in a man of letters, the first and most essential step towards being a sound, believing Christian" (1992, p. 292). Hume's actual skepticism points in a different direction, as a close examination of the arguments in his writings on religion reveals.

THE IDEA OF GOD. Hume professed agreement with Locke and other anti-innatists that the idea of "an infinitely intelligent, wise, and good Being" has its origin in one's "reflecting on the operations of our own mind, and augmenting, without limit, those qualities of goodness and wisdom" (1999, pp. 97–98). Nevertheless, he also maintained that the attempt to realize this definition in an idea is fraught with difficulty. Not only is "the capacity of the mind … limited, and can never attain a full and adequate conception of infinity" (1978, p. 26), even large numbers are representable only by means of the power of multiplying ideas, and, like all powers, rests ultimately on custom (pp. 22–23). The case of qualitative superlatives such as wisdom and goodness is even more problematic, for, finite or infinite, they "are not, like quantity or number, susceptible of any exact mensuration, which may be the standard" (1992, p. 281). In addition, Hume devoted the greater part of the *Dialogues* to showing that the empiricist definition of the divine founded on qualities of the human mind can never provide one with an idea remotely adequate to underwriting the conception of God featured in the discourse of philosophical theologians. Had he been bolder, he might also have applied to the case of God the implications of his associationist explications of the ideas of power and efficacy (necessary connection), substance, identity over time, the simplicity of complex beings, personhood, and reason. For their result is to show that these ideas are all inseparably bound up by content with the actions and affects of associative imagination, and so cannot be used to comprehend anything that exists prior to and independently of idea-enlivening, transition-facilitating. Therefore, it is ironic (no doubt intentionally so) that Hume ended up on the same side as the most pious monotheists (represented by Demea in the *Dialogues*) in insisting on the incomprehensibility of the nature of the divine.

A *PRIORI* ARGUMENTS FOR THE EXISTENCE OF GOD. The ontological argument for the existence of God advanced by many philosophers before Hume depends on treating existence as a property of God in the same sense in which goodness, wisdom, power, and other attributes are ascribed to the nature of divinity, and, moreover, like them, a necessary property. Hume argued against the first part of the thesis by denying that existence can ever be conceived of as a property, be it of God

or any other being. For to be able to do so, existence would have to be a distinct idea in its own right, capable of being combined with other ideas to form a complex idea, and there is no such idea in one's possession. Nor is the real existence attributed to God when, instead of merely conceiving him to exist, one believes him actually to exist, any new addition to the idea either, "When I think of God, when I think of him as existent, and when I believe him to be existent, my idea of him neither encreases nor diminishes" (1978, p. 94).

Even if there was an idea of real existence one could conjoin with one's idea of God, one still could not suppose it to apply necessarily, "Nothing that is distinctly conceivable implies a contradiction. Whatever we conceive as existent, we can also conceive as non-existent. There is no being, therefore, whose non-existence implies a contradiction. … The words, therefore, 'necessary existence' have no meaning; or, which is the same thing, none that is consistent" (1992, p. 251). If it is objected that God might in fact be a necessary existent even if existence does not attach to God of necessity in the idea one's feeble mind is able to form of divinity, the reply is that the same may be true of the unknown nature of any object, sensible objects included. The point is that one can never have reason to include existence in one's idea of God as a necessary attribute.

A *POSTERIORI* ARGUMENTS FOR THE EXISTENCE OF GOD. Insofar as Hume's explications of ideas such as cause and effect show them to be bound up by content with the actions and affects of associative imagination, the scope of their application is limited to the purview of appropriately constituted conscious minds. Consequently, in order to even to raise the question whether experience provides any justification for inferring the existence of God, Hume had first to set aside these explications. This should not be forgotten when trying to assess the true nature and scope of his critique of *a posteriori* theistic reasoning.

COSMOLOGICAL ARGUMENTS FOR THE EXISTENCE OF GOD. Many philosophical theists employ the general causal maxim to argue from the fact that something exists that some first cause must exist as well, since the supposition of an infinite regress of causes implies that the whole chain of causes and effects would lack a cause or reason for existing, and this is inconsistent with the maxim. Hume regarded such reasoning as fallacious:

> [T]he uniting of these parts into a whole, like the uniting of several distinct countries into one kingdom, or several distinct members into one body, is performed merely by an arbitrary act of the mind, and has no influence on the nature of things. Did I show you the particular causes of each individual in a collection of twenty particles of matter, I should think it very unreasonable should you afterwards ask me what was the cause of the whole twenty. This is sufficiently explained in explaining the cause of these parts.
>
> (1992, PP. 252–253)

ARGUMENTS FROM DESIGN. Though given a pass in the *Treatise* and elsewhere in Hume's corpus, Hume subjected the design argument for the existence of God to critical scrutiny in section 11 of *Enquiry concerning Human Understanding*, "Of a Particular Providence and of a Future State." The discussion takes the form of a dialogue between Hume and a paradox-loving skeptical friend who imagines what Epicurus might have said in his defense if brought before a tribunal on charges of impiety and endangering the state because of his denial that religion (the existence of God and of a providence and future) can be established "upon principles of reason" (1999, p. 189).

For the sake of argument, Epicurus grants that the order, beauty, and wise arrangement everywhere observed in the universe cannot have resulted from material causes alone, so that the point at issue is what kind of author(s) can be inferred from the work according to the canons of empirical reasoning. Since the cause is something that has never been observed by any mortal, and since the given effect (the totality of design in nature) is so singular as to afford no basis for determining the general characteristics (species) of its cause, Epicurus maintains that one has no choice here but to subject one's reasoning to the "maxim, that where any cause is known only by its particular effects, it must be impossible to infer any new effects from that cause, since the qualities, which are requisite to produce these new effects along with the former, must either be different, or superior, or of more extensive operation, than those which simply produced the effect, whence alone the cause is supposed to be known to us" (1999, p. 196n).

This means that one must incorporate into one's conception of the cause the abundant empirical evidence of disorder, ugliness, indifference to human welfare, and the unjust distribution of talents, goods, and fates. So, even with the concession that matter and motion are insufficient to account for the world, the cause one is warranted in inferring from the effect as one empirically finds it falls far short of the superlative, benevolent intel-

ligence proponents of the design argument claim to be able to infer.

In the *Dialogues* this line of argument is deepened and expanded, even while Hume maintains the pretense that the design argument suffices to prove the existence of a deity and fails only when it comes to providing insight into the nature of that deity (like Kant after him, Hume suggests, in the *Dialogues* [dialogue 5], that empirical reasoning would need to be supplemented by *a priori* if this want were to be made good). It is impossible here to do justice to this splendid work, possibly the finest philosophical dialogue since Plato. Suffice it to say that its conclusion is "that the causes or causes of order in the universe probably bear some remote analogy to human intelligence" (1992, p. 291).

What this means becomes clearer in the light of Philo's observation in dialogue 7 that intelligence is just one of four known causes of order in the world and that the same claim of a remote analogy with the cause(s) of order in the universe can, with equal reason, be made for instinct (a bird's design of its nest), generation (of offspring by animals), and vegetation (seeding). Since even an atheist can admit that, in this highly attenuated analogical sense, it is proper to think of the cause of order in the world as similar to intelligence—and possibly to many other, as yet unknown principles of order as well—nothing of any consequence seems to be warranted by the conclusion reached in the *Dialogues*. Indeed, it is no wonder that Hume has Philo argue that the difference between atheists and certain theists is merely verbal (1992, pp. 280–281).

Nor does Philo deny that, among the unknown principles of order in the world, some may be inherent in matter itself, such that over vast periods of time, a minute probability that the motions of particles will eventuate in the production and replication of stable, orderly forms must eventually be realized (1992, pp. 244–247). Since other principles of order, known and unknown, may themselves be explicable in terms of principles inherent in matter, even the modest conclusion reached at the end of the *Dialogues* is put in jeopardy by this concession—"So dangerous is it to introduce this idea of necessity into the present question! And so naturally does it afford an inference directly opposite the religious hypothesis!" (p. 1992, p. 253) Since Hume elsewhere made no secret that he embraced necessity in precisely this sense, one cannot help wondering if the neo-Epicurean excursus in *Dialogues* (dialogue 7) was not intended to remind his reader of Hume's own explication of cause and effect, to the end of rejecting all causal reasoning in matters of religion—as happens overtly in *Enquiry concerning Human Understanding*:

> It is only when two *species* of objects are found to be constantly conjoined, that we can infer the one from the other; and were an effect presented, which was entirely singular, and could not be comprehended under any known *species*, I do not see, that we could form any conjecture or inference at all concerning its cause. If experience and observation and analogy be, indeed, the only guides which we can reasonably follow in inferences of this nature, both the effect and cause must bear a similarity and resemblance to other effects and causes, which we know, and have found, in many instances, to be conjoined with each other. I leave it to your own reflections to pursue the consequences of this principle.
>
> (1999, p. 198)

REASON AND REVELATION. Is it ever rational to accept the truth of revealed religion? Those who answer affirmatively typically point to prophecies fulfilled and miracles performed. Since such evidence comes to nearly all of us by way of oral or scriptural testimony, Hume asked if conditions exist under which one could rationally credit reports of prophesies and miracles and, if so, whether any revelation has ever met these conditions. The key to his reasoning in this matter is the recognition that human testimony on any topic owes whatever authority it has in the eyes of reason to the same source causal inferences do: past experience. Finding there to be a fairly constant conjunction between the facts as reported by witnesses and as ascertained by other means, one has only to hear or read (have an impression of) a report for one's mind not only to think (form an idea) of the event reported but also to believe it to the extent (enliven the idea to the degree) warranted by experience. For, besides lending authority to testimony in general, experience also teaches one that particular reports are more or less credible depending on the reporter, the circumstances under which the report is given and received, and the event reported itself. If a report falls short of maximum credibility on any of these counts, then reasonable persons must refuse to give it the same credence they accord to empirical beliefs founded on a frequently encountered, perfectly constant conjunction, having the certainty of proofs.

Reports of miracles are intrinsically suspect because the events they report are, by their nature, the least creditable. As defined by Hume, an event is miraculous only if it meets two conditions: it contradicts a law of nature and

does so "by the particular volition of the Deity, or by the interposition of some invisible agent" (1999, p. 173n). A law of nature is a causal sequence found by constant experience to be invariable, and so has the highest authority empirical reason can confer. Accordingly, to determine whether one can rationally credit any report of a miracle, one must follow the procedure empirical reason prescribes whenever two beliefs regarding matters of fact are found to conflict: deduct from the empirical support of one of the beliefs the amount of support possessed by the other and, if any support remains, accord it only so much credence as that remainder warrants; otherwise, discount it or (if the beliefs have equal support) refrain from believing either way. However, when one does this, one finds that

> no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish: And even in that case, there is a mutual destruction of arguments, and the superior only gives us an assurance suitable to that degree of force, which remains, after deducting the inferior.
>
> (P. 174)

Since it is impossible that experience could ever give one reason to regard the falsehood of any report of miracles as more improbable than the falsehood of a law of nature, even the most credible testimony imaginable could not win one's acceptance if belief were always proportioned to experience. The same is true of prophecies, for these are simply a species of miracle ("If it did not exceed the capacity of human nature to foretel future events, it would be absurd to employ any prophecy as an argument for a divine mission or authority from heaven" [1999, p. 186]). Thus, one's acceptance of revealed religion can never possess the rational authority to which belief proportioned to the evidence of experience can alone lay claim.

RELIGIOUS BELIEF. Having established that one has no clear idea of God to underwrite religious discourse nor any rational basis for religious belief, Hume devoted the remainder of his discussion of miracles, as well as other writings ("The Natural History of Religion" [1757] most notably), to examining the nature and causes of religious belief. The upshot is that one believes in God and accepts the proofs of purported revelation from the same causes that lead one to form other beliefs not proportioned to experience (unphilosophical probabilities): failure to clarify one's ideas or to ascertain the existence of ideas corresponding to one's words; education; credulity; self-interest; the influence of the passions; eloquence and other appeals to imagination that detach reason from its moorings in experience; the errors and exaggerations that tend to creep in with each new telling of a story; and so on. The implication is that, however widespread a religious belief may be, it is not imposed on one by human nature, and so is not irresistible in the way that belief in causes, continued distinct existents, and the self are.

Hume did not deny that religious belief can ever be agreeable or useful, either for the individual or society, but he did seem to think that, in the forms it actually takes—especially when vitiated by superstition or enthusiasm—it is neither. For example, in two essays, "Of Suicide" and "Of the Immortality of the Soul" (written in 1755 but published posthumously in 1777 [though a French edition appeared in 1770]), he argued that there is no rational or moral basis for the prohibition of the former or for belief in the latter. Still, his single most important philosophical contribution to the effort of combating the deleterious influence of religion is the example set by his theory of morals: It illustrates how universally valid moral standards can be understood non-theologically, in terms exclusively of natural sentiment and artificial interest.

See also Aesthetics, History of; Alembert, Jean Le Rond d'; Bacon, Francis; Beauty; Berkeley, George; Causation: Metaphysical Issues; Causation: Philosophy of Science; Colors; Common Sense; Cosmological Argument for the Existence of God; Determinism, A Historical Survey; Determinism and Freedom; Diderot, Denis; Enlightenment; Human Nature; Induction; Kant, Immanuel; Locke, John; Newton, Isaac; Perception; Philo Judaeus; Reason; Reid, Thomas; Revelation; Rousseau, Jean-Jacques; Skepticism, History of; Smith, Adam; Space; Virtue and Vice; Volition.

Bibliography

WORKS BY HUME

Four Dissertations. London: Printed for A. Millar, 1757.

The Philosophical Works of David Hume. 4 vols. Edited by T. H. Green and T. H. Grose. London: Longmans, Green, 1875.

The Letters of David Hume. 2 vols. Edited by J. Y. T. Grieg. Oxford, U.K.: Clarendon Press, 1932.

New Letters of David Hume. Edited by Raymond Klibansky and Ernest Campbell Mossner. Oxford, U.K.: Clarendon Press, 1954.

A Letter from a Gentleman to His Friend in Edinburgh (1745). Edited by Ernest Campbell Mossner and John V. Price. Edinburgh, Scotland: Edinburgh University Press, 1967.

A Treatise of Human Nature. Edited by L. A. Selby-Bigge. Oxford, U.K.: Clarendon Press, 1978.

The History of England, edited by William B. Todd, Indianapolis: Liberty Classics, 1983.

Essays Moral, Political, and Literary. Edited by Eugene F. Miller. Indianapolis, IN: Liberty Classics, 1987.

Writings on Religion. Edited by Antony Flew. La Salle, IL: Open Court, 1992.

The Clarendon Edition of the Works of David Hume. Edited by Tom L. Beauchamp, David Fate Norton, and M. A. Stewart. New York: Oxford University Press, 1998–.

An Enquiry concerning the Principles of Morals: A Critical Edition. Edited by Tom L. Beauchamp. New York: Oxford University Press, 1998.

An Enquiry concerning Human Understanding. Edited by Tom L. Beauchamp. New York: Oxford University Press, 1999.

A Treatise of Human Nature. Edited by David Fate Norton and Mary J. Norton. New York: Oxford University Press, 2000.

WORKS ABOUT HUME

Àrdal, Páll S. *Passion and Value in Hume's "Treatise"*. Edinburgh, Scotland: Edinburgh University Press, 1966. A detailed textual analysis of Hume's theory of the passions.

Baier, Annette. *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, MA: Harvard University Press, 1991. An authoritative, highly persuasive vision of the *Treatise* as the onerous progress of reason toward self-comprehension, which it finds in society.

Baillie, James. *Hume on Morality*. London: Routledge, 2000.

Beauchamp, Tom L., and Alexander Rosenberg. *Hume and the Problem of Causation*. New York: Oxford University Press, 1981. An attempt to defend Hume's views on induction and causation against contemporary views.

Bricke, John. *Mind and Morality: An Examination of Hume's Moral Psychology*. New York: Oxford University Press, 1996.

Capaldi, Nicholas. *Hume's Place in Moral Philosophy*. New York: Peter Lang, 1989.

Flew, Antony. *David Hume: Philosopher of Moral Science*. Oxford, U.K.: Basil Blackwell, 1986.

Fogelin, Robert L. *Hume's Skepticism in the Treatise of Human Nature*. London: Routledge and Kegan Paul, 1985. Aspires to restore Hume's skepticism to a status equal, and complementary, to his naturalism, both in epistemology and in morals.

Forbes, Duncan. *Hume's Philosophical Politics*. New York: Cambridge University Press, 1975.

Garrett, Don. *Cognition and Commitment in Hume's Philosophy*. New York: Oxford University Press, 1997. One of the finest books on Hume to date.

Gaskin, J. C. A. *Hume's Philosophy of Religion*. 2nd ed. Atlantic Highlands, NJ: Humanities Press International, 1988.

Haakonssen, Knud. *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith*. New York: Cambridge University Press, 1981.

Harrison, Jonathan. *Hume's Theory of Justice*. Oxford, U.K.: Clarendon Press, 1981.

Laird, John. *Hume's Philosophy of Human Nature*. London: Methuen, 1932. A classic.

Livingston, Donald W. *Hume's Philosophy of Common Life*. Chicago: Chicago University Press, 1984. Maintains that Hume adopted the approach of the social historian even in epistemology.

Mackie, John L. *The Cement of the Universe: A Study of Causation*. Oxford, U.K.: Clarendon Press, 1974.

Mackie, John L. *Hume's Moral Theory*. London: Routledge and Kegan Paul, 1980. A projectivist reading of Hume's sentimentalism, with a keen analysis of his account of virtue.

Miller, David. *Philosophy and Ideology in Hume's Political Thought*. Oxford, U.K.: Clarendon Press, 1981.

Mossner, Ernest Campbell, ed. "Hume's Early Memoranda, 1729–1740." *Journal of the History of Ideas* 9 (1948): 492–518.

Mossner, Ernest Campbell. *The Life of David Hume*. 2nd ed. Oxford, U.K.: Clarendon Press, 1980.

Norton, David Fate, ed. *The Cambridge Companion to Hume*. New York: Cambridge University Press, 1993.

Norton, David Fate. *David Hume: Common-Sense Moralist, Sceptical Metaphysician*. Princeton, NJ: Princeton University Press, 1982. Contains an excellent account of Hume's relation to Frances Hutcheson and examines other influences on Hume.

Owen, David. *Hume's Reason*. New York: Oxford University Press, 1999. Corrects many misconceptions, especially regarding Hume's conception of deductive reason. An outstanding work.

Passmore, John. *Hume's Intentions*. 3rd ed. London: Duckworth, 1980.

Pears, David. *Hume's System: An Examination of the First Book of His Treatise*. New York: Oxford University Press, 1990. Effects a balance between Hume's empiricism and naturalism by correlating the first to a theory of meaning and the second to a theory of truth.

Philips, D. Z., and Timothy Tessin, eds. *Religion and Hume's Legacy*. New York: St. Martin's, 1999.

Read, Rupert, and Kenneth A. Richman, eds. *The New Hume Debate*. New York: Routledge, 2000.

Russell, Paul. *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*. New York: Oxford University Press, 1995.

Smith, Norman Kemp. *The Philosophy of David Hume*. London: Macmillan, 1941. Instigated the move away from skeptical readings of Hume.

Snare, Francis. *Morals, Motivation, and Convention: Hume's Influential Doctrines*. New York: Cambridge University Press, 1991.

Stewart, M.A., and John P. Wright, eds. *Hume and Hume's Connexions*. University Park: Pennsylvania State University Press, 1995.

Stove, D. C. *Probability and Hume's Inductive Scepticism*. Oxford, U.K.: Clarendon Press, 1973.

Strawson, Galen. *The Secret Connexion: Causation, Realism, and David Hume*. New York: Oxford University Press, 1989.

Stroud, Barry. *Hume*. London: Routledge and Kegan Paul, 1977. Presents Hume as advancing a theory of human nature in which the traditional conception of man as rational animal is overturned.

Townsend, Dabney. *Hume's Aesthetic Theory: Taste and Sentiment*. London: Routledge, 2001.

Tweyman, Stanley, ed. *David Hume: Critical Assessments*. 6 vols. London: Routledge, 1995.

Tweyman, Stanley, ed. *Hume on Miracles*. Bristol, U.K.: Thoemmes, 1996.

Waxman, Wayne. *Hume's Theory of Consciousness*. New York: Cambridge University Press, 1994.

Waxman, Wayne. *Kant and the Empiricists: Understanding Understanding*. New York: Oxford University Press, 2005. Emphasizes the affinity between Hume and Kant as exponents of a psychologistic method of explicating concepts.

Whelan, Frederick G. *Order and Artifice in Hume's Political Philosophy*. Princeton, NJ: Princeton University Press, 1985.

Wilson, Fred. *Laws and Other Worlds: A Humean Account of Laws and Counterfactuals*. Dordrecht, Netherlands: D. Reidel, 1986.

Wayne Waxman (2005)

HUMOR

Although the laughable is not usually thought of as a subclass of the beautiful (Aristotle, indeed, said that it was a subclass of the ugly), the problem of "humor" is a special case of the central problem of aesthetic theory. To find something laughable is to have a special kind of aesthetic emotion, but it is not at all easy to say just what features of the laughable situation evoke this emotion. Theories of humor attempt to answer this question.

TYPES OF HUMOR

The only way to evaluate theories of humor is to see how well they apply to different types of jokes or humorous situations. For this we need a list of the main types of humor. The attempt to provide one may, however, prejudge the issue, since the basis of classification may itself presuppose a theory of humor. Moreover, if any one theory is right, then in the final analysis jokes will be of only one type: They will all turn on release of inhibitions, or superiority to the misfortune of others, or whatever it may be.

With these reservations, the following may be regarded as the main types of humorous situations: (*a*) Any breach of the usual order of events, as wearing an unusual costume or eating with chopsticks when one is used to knife and fork (or with knife and fork when one is used to chopsticks). (*b*) Any breach of the usual order of events that is also felt to break a rule, whether of morality or etiquette. The drunkard, the glutton, the hypocrite, the miser are all stock figures of comedy, on the stage and elsewhere. (*c*) A special case of the second type is indecency, as in Restoration comedy or any smoking-room story. This has a different flavor from comic vice, just as comic vice has a different flavor from mere novelty and oddity. (*d*) Introduction into one situation of what is felt to belong to another, as George Bernard Shaw's reference to conventional sexual morality as "the trade unionism of married women" or Mark Twain's introduction of a Connecticut Yankee into the Court of King Arthur. Finding connections between things we usually keep in separate compartments of our minds is, according to one version of the incongruity theory, the ultimate source of all humor. Whether this is correct or not, it is certainly one source that needs to be noted. (*e*) Anything masquerading as something it is not. This has been a favorite stage device, from *Twelfth Night* to *Charley's Aunt,* and is common enough in other forms of comedy. (*f*) Wordplay, of which puns are the most obvious, but not of course the only, example. (*g*) Nonsense, especially of the Edward Lear or Lewis Carroll type, which often turns on wordplay but is distinct from it. (*h*) Small misfortunes, like those provided by the banana skin, the custard pie, the thumb beneath the hammer. (*i*) Want of knowledge and skill, as in the schoolboy howler or the circus clown clumsily attempting to imitate the acrobat. (*j*) Veiled insults, as in the catty remarks in *The School for Scandal.*

THEORIES OF HUMOR

Most theories find the essence of humor in one or another of the following: superiority, incongruity, and relief from restraint. It has also been suggested that humor derives from ambivalent feelings, in which attraction and repulsion are both present.

SUPERIORITY THEORIES. If we laugh at the miser, the drunkard, the glutton, the henpecked husband, the man who gets hit by the custard pie, the schoolboy howler, the person with faulty pronunciation, may it not be because we feel superior to all of these? This could account for our pleasure in humor. Accordingly, Thomas Hobbes regarded laughter as the result of a sudden access of self-esteem ("sudden glory") when we realize that our own situations compare favorably with the misfortunes or infirmities of others. We also laugh, he said, at our own past follies—provided we are conscious of having surmounted them—or at unexpected successes.

In support of Hobbes, or perhaps as a modification of his view, it may be said that in humor at its best we are conscious of surveying the whole human scene from some godlike level at which all men and women look pretty much alike: all weak, all lovable, all transparently obvious in their petty pretenses. If "superiority" is interpreted as this god's-eye view rather than as simply a

sneering contempt for some failing we do not have, it is possible to account for laughter not merely at comic vice but also at comic virtue, as in Mr. Pickwick or Don Quixote. It may even explain why we often laugh with comic vice rather than at it. No one feels superior to Falstaff, but we may feel pleasantly conscious of "seeing through" him, and perhaps, in sympathizing with him, we feel superior, if only for the time being, to the conventional morality he flouts.

By extending Hobbes's theory in this way, it is possible to account for many of our classes of humor: indecency and masquerade as well as comic vice, small misfortunes, and ignorance. Alexander Bain extended Hobbes in two directions. Sometimes, Bain suggested, our laughter may be a manifestation not of our own feeling of superiority but of our sympathy with someone else who has triumphed in some way. This would account for laughter at veiled insults. Second, the triumph need not be over a person; it can be over anything at all that is conventionally treated with respect. Mark Twain's debunking of feudal values was not directed at any individual, and Samuel Butler degraded a sunrise by comparing it to a boiled lobster. According to Bain, the essential feature of humor is degradation. Some writers have argued, not very plausible, that in wordplay we triumph over the degradation of words. More credibly, nonsense may be regarded as the degradation of what Arthur Schopenhauer called "that strict, untiring, troublesome governess, the reason." Even incongruity, it is argued, always involves degradation. Typically, the incongruous effect is obtained by the bringing of something exalted into contact with something trivial or disreputable. Shaw's phrase has its force because trade unionism is much lower on the conventional scale of values than is chastity: The pleasure in seeing them linked is, at least in part, malicious.

Henri Bergson maintained that the particular characteristic exciting derision is inflexibility, the inability to adapt oneself to the ever-changing demands of life. Laughter is always at "something mechanical encrusted upon the living." With Molière in mind, Bergson claimed that the comic character is usually a man with a fixed idea. This fits in with early stage comedy and with the etymology of the word *humor:* A humor was originally a quirk, a kink, a mental (and primarily a physiological) oddity that throws a man off balance and twists his view of life. Hence, the comic character is simply a man with an obsession. The joke is to see how this obsession crops up again and again in the most varied situations, so that he always behaves in a manner wildly inappropriate to the circumstances as others see them but entirely appropriate to his own ruling passion.

With more ingenuity than plausibility, Bergson attempted to apply his formula to wordplay, which consists, he claimed, in showing that language is too rigid to be an accurate mirror of an infinitely fluid universe. His main emphasis, however, was on the social function of laughter; it is leveled, according to him, at the eccentric or nonconformist. This seems an unduly restricted view: The most penetrating humor is often aimed at the social code itself. There is nothing in Bergson's theory of humor that need have prevented him from conceding this: The conventions of society may often enough be characterized as "something mechanical encrusted upon the living."

INCONGRUITY THEORIES. It can be doubted whether the concepts of "superiority" or "degradation" or even "inelasticity" do justice to the very large element of humor that consists in the intellectual and emotional pleasure of finding connections where none were thought to exist. It is true that if this were the whole of humor, humor would be indistinguishable from fancy or imagination; but then, if "degradation" were the whole of humor, humor would be indistinguishable from malice.

Immanuel Kant asserted that humor arises "from the sudden transformation of a strained expectation into nothing," and since his time incongruity has often been identified with "frustrated expectation." But there is more to incongruity than mere surprise, or even anticlimax; we must be, as it were, jolted out of one mental attitude into another completely and violently opposed to it. Usually this results from bringing together two things normally kept in separate compartments of our minds. Shaw's aphorism about the trade unionism of married women may once again serve as an example. Another is Butler's "God and the Devil are an effort after specialisation and division of labour." In Kant's view, the "degradation" of one of the two disparate ideas is quite incidental. What is important is that they normally evoke very different attitudes and that the connection between them appears to be genuine, not artificially contrived. It is on these two features that the neatness of a joke depends.

Kant's formula may be regarded as defective in that by putting the emphasis on surprise it ignores the logical connection between the two ideas that are linked. This is Schopenhauer's criticism. He claimed that all humor can be "traced to syllogism in the first figure with an undisputed major and an unexpected minor, which to a certain extent is only sophistically valid."

This formula applies most obviously to the mock-heroic or to certain types of satire. The point of Henry Fielding's *Jonathan Wild,* for example, might be summarized syllogistically as: All generals and those who behave like generals are heroes; highwaymen behave like generals; therefore, highwaymen are heroes. Here the major premise is, conventionally, undisputed. The minor is, no doubt, "only sophistically valid," but only "to a certain extent"; there is enough resemblance in behavior to give the satire sting.

The formula applies, however, to other types of humor as well. Oscar Wilde is reported to have said, when he was in prison, "If this is the way the Queen treats her convicts, she doesn't deserve to have any." Here the major premise is: "Those who ill-treat their dependents deserve to lose them." This generalization is then made to apply to a case in which losing them would be no hardship and deserving to lose them no demerit. What is sophistical about the minor premise is the assumption that a convict is, along with a servant, a child, and the like, the kind of dependent to whom the generalization applies.

The objection to Schopenhauer's analysis is that it stresses the formal side of a joke to the exclusion of the content. For him, humor was purely a matter of finding connections where (except in a "sophistical" sense) none exist. By this view, all humor is of the type of Richard Whately's *Historic Doubts Relative to Napoleon Bonaparte.* The essence of it lies in the ingenuity of the argument, underlined by the absurdity of the conclusion. If any derision creeps in, it is at the expense of the reasoning, or perhaps of the governess Reason herself.

What this overlooks is the part that the abrupt dissolution of an attitude plays in our emotional lives. Kant's phrase "strained expectation" hints at this but does not characterize it adequately. *Jonathan Wild* would not be funny if it were not for the whole complex of emotions that cluster round the concepts of patriotism and national glory. To take another example, Gerald Bullett's adaptation of Alfred, Lord Tennyson, "Wearing the white feather of a blameless life," is funny, not merely because of its close resemblance to the wording of the original ("the white *flower* of a blameless life") but because of the startling difference in attitude that results from the alteration of a single word.

So far as superiority theories call attention to the emotional element in humor, they do something to correct this inadequacy. It is doubtful, however, whether the emotion involved is either self-congratulation or malice. In any community certain attitudes are felt to be appropriate to some things and not to others, and there develop "stereotypes" of such figures as the typical politician, poet, businessman. The humorist drags into light the inconvenient facts that shatter these attitudes and puncture these stereotypes. Sometimes, as Bergson pointed out, the humor is at the expense of the person who is unable to live up to the conventional requirements, and here malice may creep in, but often enough the effect is to cast doubt on the conventional attitudes and values. Sometimes it is not clear which effect is intended. Wilde's witticism "Work is the curse of the drinking classes" may be taken either as a gibe at the working classes or as a questioning of the conventional Victorian attitudes to work and to drink. In either case one element in our enjoyment is certainly the sense of enlarged horizons that comes from seeing unexpected connections. This is in part an intellectual pleasure. So far as it is a conventional attitude that has been convicted of inadequacy, the accompanying emotion may be not malice or superiority but a feeling of liberation at the removal of intellectual blinkers.

RELIEF THEORIES. Liberation, or relief from restraint, is regarded in a third type of theory as the central element in humor.

It is well known that people who have been undergoing a strain will sometimes burst into laughter if the strain is suddenly removed. It has been argued that all laughter is of this type and that any joke will be found, in one way or another, to remove the restraints which society imposes on our natural impulses. It is the liberation of our impulses from social constraints, not of our intellects from too narrow a point of view, that is emphasized by this type of theory.

What are these impulses that need liberating? One obvious one is the sexual impulse. Since the mention of the (conventionally) unmentionable is in itself a sufficient cause of laughter, it seems reasonable to say that at least one important type of humor depends on our being able to give vent to forbidden thoughts and feelings.

But thoughts about sex are not the only ones that society calls on us to suppress. Our aggressive impulses are also repressed. Children are taught that it is "rude" both to expose their bodies and to speak insultingly to others. Consequently, the relief theory can account, plausibly enough, for the malicious element in humor and, in general, for most of the aspects of humor that have given rise to superiority theories. Even nonsense can be explained, if it is conceded that trying to be rational all the time is a strain for most of us.

Relief theories have been given considerable impetus by the rise of psychoanalysis. Sigmund Freud himself wrote a book on humor, in which he suggested that there is a basic resemblance between jokes and dreams. Both are essentially means of outwitting the "censor," the name by means of which Freud personified our internal inhibitions. In dreams forbidden thoughts are distorted and disguised; in humor insults are veiled, masquerading perhaps as compliments, and sexual references lurk behind apparently innocent remarks.

Freud did not, however, regard all humor as the release of inhibition. He distinguished between "harmless wit," indulged in for its own sake, and "tendency wit," which gives us the additional gratification of giving rein to repressed sexual or aggressive impulses. Harmless wit delights us because it provides us with "infantile play-pleasure." In learning to use words, Freud pointed out, children "experience pleasurable effects which originate from the repetition of similarities, the rediscovery of the familiar, sound-associations," and the like. In other words, the pleasure of playing with words and ideas, on which incongruity theories place so much stress, is admitted by Freud to be enjoyable for its own sake, not just as a means of seeking relief from restraint. It is, indeed, because this intellectual play is enjoyable in itself that we can use it to beguile the censor. When Wilde, for example, complained that "the youth of to-day are quite monstrous; they have absolutely no respect for dyed hair," we must suppose that the censor is so diverted by the discovery that this remark differs only in one word from the conventional headshaking of the stuffier kind of matron that the malice in the remark (its complete exposure of the matron's pretensions and its revelation of her envy of youth) is allowed to go unchecked.

Freud explained "infantile play-pleasure" by invoking the concept of "psychic economy." In this he was influenced by Herbert Spencer. Spencer thought that humor consists essentially in the abrupt transition of thought from a noble or elevated idea to a trivial or degrading one, leaving the psyche with an unexpended fund of nervous energy that overflows into laughter, which is, according to him, a physical release of energy. Freud adapted this notion for his own purposes, identifying "psychic economy" first with the line of least resistance and then with the brevity and neatness that is the soul of wit.

Neither Spencer's nor Freud's use of the concept is very satisfactory. It may be pointed out against Spencer that when, for example, an innocent remark is transformed into a sexual reference, the second might be expected to call forth more emotional energy than the first. Against Freud it may be said that the lazy pleasure of following the path of least resistance is very different from our appreciation of the skill with which a master of humor links disparate ideas. When writers such as François Rabelais, G. K. Chesterton, Christopher Fry, James Joyce, and even Laurence Sterne play with words and ideas, it is exuberance rather than economy that they display.

RELATION OF THE THEORIES TO THE TYPES

If the theories are evaluated by their ability to explain the main types of humor listed earlier, it would seem that none is completely adequate by itself. Each of them relies mainly on particular kinds of humor, either ignoring the rest or giving relatively lame accounts of them. Satire and laughter at small misfortunes are very well explained by superiority theories. Incongruity theories find difficulty in dealing with these but are much more satisfactory than superiority theories in dealing with wordplay, nonsense, and indecency. Relief theories can explain malice and indecency, and perhaps nonsense, but are driven to admit that wordplay and the finding of unexpected connections have an intrinsic appeal that cannot be reduced to relief from restraint.

See also Aristotle; Bain, Alexander; Bergson, Henri; Butler, Samuel; Carroll, Lewis; Freud, Sigmund; Hobbes, Thomas; Kant, Immanuel; Schopenhauer, Arthur.

Bibliography

GENERAL

Monro, D. H. *Argument of Laughter.* Melbourne: Melbourne University Press, 1951; Notre Dame, IN: University of Notre Dame Press, 1963.

Sully, J. *Essay on Laughter.* London: Longmans, 1902.

SUPERIORITY THEORIES

Bain, Alexander. *The Emotions and the Will,* 3rd ed. London: Longmans, 1880.

Bergson, Henri. *Le rire.* Paris, 1900. Translated by Cloudesley Brereton and Fred Rothwell as *Laughter.* London: Macmillan, 1911.

Leacock, Stephen. *Humour: Its Theory and Techniques.* London: John Lane, 1935.

INCONGRUITY THEORIES

Eastman, Max. *Enjoyment of Laughter.* New York: Simon and Schuster, 1936.

Schopenhauer, Arthur. *Die Welt als Wille und Vorstellung,* 2nd ed., 2 vols. Leipzig, 1844. Translated by E. F. J. Payne as *The World as Will and Representation,* 2 vols. Indian Hills, CO: Falcon's Wing Press, 1958.

RELIEF THEORIES

Freud, Sigmund. *Der Witz und seine Beziehungen zum Unbewussten.* Leipzig and Vienna, 1905. Translated by James Strachey as *Jokes and Their Relation to the Unconscious.* London: Routledge and Kegan Paul, 1960.

D. H. Monro (1967)

HUMOR [ADDENDUM]

Since 1980 a number of philosophers have explored the psychology, aesthetics, and ethics of humor. The incongruity theory dominates though, as in earlier theorizing, the precise relation of humor to incongruity is seldom made clear. Not just any experience of incongruity constitutes humor. Coming home to find one's family murdered would be incongruous but not funny. Even incongruity intended to be humorous may not be funny to everyone. What seems necessary for humorous amusement is that the incongruity be enjoyed.

The enjoyment of incongruity is not sufficient for amusement, however, for one may enjoy something bizarre or fantastic for its incongruity without being amused. Some people also enjoy grotesque and macabre works of art and horror movies and novels in part for their incongruity. In aesthetics there is no general agreement on what distinguishes amusement from such cases; one suggestion is that in amusement we tend to laugh.

Another psychological and aesthetic issue is the relation of humor to emotions. Since Plato most philosophers have treated amusement as an emotion, but there are significant differences between amusement and standard emotions. The practical orientation of standard emotions is lacking in amusement. Emotions evolved in early mammals as adaptive reactions to threats and opportunities. The bodily changes in fear and anger, for example, energize animals and humans for fighting or fleeing. Sexual love motivates reproductive activity and parental love motivates nurturing. But the bodily changes in amusement do not prepare us to take action; indeed, uncontrollable laughter is incapacitating.

A second difference is that amusement does not require belief in the reality of its object as emotions typically require belief in the reality of their objects. News that I have won the lottery might make me feel joy, but that joy evaporates when I discover that the news was false. Humor, by contrast, seems to work as well with playful, merely entertained thoughts as with beliefs. Indeed, those who produce jokes and other forms of comedy work mostly with intentional objects known by everyone to be fictional. A third difference is that in standard emotions, there is a positive or negative attitude toward the object of those emotions while in amusement there need be no positive attitude toward the amusing object. People value what they love and what brings them joy, but they need not value the things they find funny. If at a funeral one sees someone dressed in a garish yellow and pink outfit, one may be amused without having a positive attitude toward that person or the outfit. Indeed, Aristotle classified the humorous as a species of the ugly.

These differences between amusement and standard emotions suggest that humor involves a more sophisticated kind of mental processing than is found in at least the basic emotions, and a different relation between its mental and physical components. Those who want to continue the Platonic classification of amusement as an emotion, then, one should at least provide an explanation of why these differences should not push amusement out of the category of emotion.

Turning lastly to the ethics of humor, since the mid-1970s, philosophers have examined humor that seems to express morally objectionable beliefs and attitudes, such as racism and sexism. Joke telling is often based on stereotypes representing various groups as stupid, lazy, greedy, or promiscuous. Are those who tell such jokes asserting or presupposing the truth of those stereotypes? One strong position, called *moralism* by Berys Gaut, says that appreciating a joke involves subscribing to the beliefs and attitudes it expresses, and so joke tellers are fully answerable to ethical considerations. At the other extreme is *antimoralism*, which treats humor as a form of play in which ideas and attitudes are merely entertained and not subscribed to, so that joke telling is not bound by ethical constraints. Antimoralists often point out that one can laugh at an ethnic joke merely for its cleverness. A joke about stupid Frisians can be amusing even though one has no idea who Frisians are! However, moralists point out that racists tell racist jokes to express and spread their beliefs and attitudes and not simply to play with ideas.

Whereas ethical examinations of humor have focused on what can be wrong with it, a few have shown how humor, particularly about oneself, can foster virtues such as humility, patience, tolerance, and forgiveness. The person with a rich sense of humor also tends to think critically, which has made humor a natural accompaniment to philosophy since Socrates.

See also Plato; Socrates.

Bibliography

GENERAL

Monro, D. H. *Argument of Laughter*. Notre Dame: University of Notre Dame Press, 1963.

Morreall, John, ed. *The Philosophy of Laughter and Humor*. Albany: State University of New York Press, 1987.

Morreall, John. *Taking Laughter Seriously*. Albany: State University of New York Press, 1983.

SUPERIORITY THEORIES

Bergson, Henri. *Laughter: An Essay on the Meaning of the Comic*. Translated by Cloudesley Brereton and Fred Rothwell. Whitefish, MT: Kessinger, 2004.

Leacock, Stephen. *Humour: Its Theory and Techniques*. New York: Dodd, Mead, 1935.

INCONGRUITY THEORIES

Eastman, Max. *Enjoyment of Laughter*. New York: Simon and Schuster, 1936.

Schopenhauer, Arthur. *The World as Will and Representation*. 2nd ed. Translated by E. F. J. Payne. Mineola, NY: Dover, 1974.

RELIEF THEORIES

Freud, Sigmund. *Jokes and Their Relation to the Unconscious*. Translated and edited by James Strachey. Harmondsworth, NY: Penguin, 1976.

AESTHETICS AND ETHICS

Bergmann, Merrie. "How Many Feminists Does It Take to Make a Joke: Sexist Humor and What's Wrong with It." *Hypatia* 1(1986): 63–82.

Buckley, Francis H. *The Morality of Laughter*. Ann Arbor: University of Michigan Press, 2003.

Carroll, Nöel. "Horror and Humor." *Journal of Aesthetics and Art Criticism* 57 (1999): 145–160.

Gaut, Berys. "Just Joking: The Ethics and Aesthetics of Humor." *Philosophy and Literature* 22 (1998): 51–68.

Roberts, Robert C. "Humor and the Virtues." *Inquiry* 31 (1988): 127–149.

John Morreall (2005)

HUS, JOHN

(c. 1369–1415)

John Hus, the Czech church reformer and national hero, was born at Husinec in southern Bohemia. He made his way through the University of Prague, receiving his A.B. in 1393, his M.A. in 1396, and his B.D. in 1404. Some of the logical works of John Wyclyf were known in Prague in the early 1390s, and there is still extant a copy of a half dozen of Wyclyf's philosophical works in Hus's hand, made in 1398. Wyclyf's realism (*universalia ante rem*) found a warm welcome among Czech professors and students, not least because the German community at the university was strongly Ockhamist and Wyclyf's vigorous defense of universals (prior to individuals) fortified the Czechs' position. He was deeply influenced by the Augustinianism of the Victorine school of the twelfth century.

Hus became well known and popular, partly for his teaching and partly for his preaching in the vernacular. In 1402 Hus was named stated preacher in the Bethlehem Chapel, and his sermons in Czech were well attended by Czechs of all classes. In October 1401 Hus was elected dean of the arts faculty and in 1403 rector of the university (though there is some uncertainty as to this first rectorate). By this time disputes over Wyclyf's teachings had become acrimonious, and Hus with some of his friends undertook to defend Wyclyf from charges of heresy against a party largely of German professors, who demanded strict condemnation of Wyclyf's teachings. Hus continued his preaching and writing in the interest of reform, but in 1408 the Prague conservative hierarchy (mainly German) lodged specific charges of heresy against him. Soon thereafter the struggle for predominance in the university broke out between Czech and German. The Germans had three votes, the Czechs only one. Hus led the fight for a reversal of the proportion, and King Wenceslaus decided in the Kutná Hora decree of 1409 that the Czech professors and students should have three votes and all others combined, one vote. The Germans left in a body to form the University of Leipzig. Hus, as leader of the national Czech party, was elected rector of the university.

Opposition to Hus on the part of the conservative Czech clergy remained, and the serious charges of 1408 were renewed in 1409 and 1410. He disobeyed a summons to Rome and was excommunicated in 1411. Hus had formed his opinions clearly by then and was prepared to defend them under any conditions. He believed firmly in predestination and the unity of the church under the headship of Christ. He was deeply influenced by the teaching of Wyclyf but in one important matter he categorically disagreed. He rejected Wyclyf's teaching on the Eucharist, accepting completely the church's doctrine of transubstantiation. Realist philosophy was important in the formulation of his theological positions, and his competence in Scholastic exposition is evident in all his writings. From the excommunication of 1411 to his death four years later it was clear that his position and that of the established hierarchy were irreconcilable. In 1412 King Wenceslaus reluctantly had to withdraw his protection, and Hus went into exile to relieve the city of Prague from the interdict. It was during his exile that he finished his most important work, the *De Ecclesia*, very similar to a book under the same title by Wyclyf. He argued against

the authority of the pope and the cardinalate over the church and their control of the means of salvation, basing his conclusions on the doctrine of predestination. "The church is the body of the predestinate." Inasmuch as only God knows who is predestinate, the pope's function and power are readily dispensable. The hierarchy could not tolerate so basic an attack on its existence. Hus appealed to the general council called for November 1414 at Constance and, receiving a safe-conduct from Emperor Sigismund, arrived in Constance on November 3. However, the safe-conduct was soon disregarded; Hus was imprisoned and interrogated at length. He asked simply to be shown from Scriptures or the Fathers where he was in error. The council demanded that he make a blanket recantation. No compromise was possible. Hus's concept of the church as the body of the predestinate, regardless of the decision of the pope and the hierarchy, was declared pure heresy. He was "relaxed to the secular arm" on July 6, 1415, and burned at the stake that morning. His martyrdom set off the Hussite Wars (1419–1434), which in turn isolated Bohemia from the rest of Europe for several generations. Hussitism, as it developed, took forms that Hus might not have approved.

Hus may not have been one of the leading minds of his century. On the other hand his commentary on the *Sententiae* of Peter Lombard, composed in 1407–1409, is a very impressive work and shows complete familiarity with the dominant currents of philosophical thought in the fourteenth and fifteenth centuries and an easy ability in the handling of contradictory arguments. His realism is confident and precise.

See also Augustinianism; Peter Lombard; Realism; Universals, A Historical Survey; Wyclyf, John.

Bibliography

WORKS BY HUS

The early edition of his collected works (1558; reprinted in 1715) is still indispensable. Some of his works have been published in modern critical editions: *Opera,* edited by V. Flajšhans (Prague, 1903–1912), and *De Ecclesia,* edited by S. H. Thomson (Boulder, CO, 1956). Also indispensable is *Documenta … Mag. Joannis Hus,* edited by Francis Palacký (1869), containing his correspondence and salient documents of his career.

The Letters of John Hus. Manchester, U.K.: Manchester University Press, 1972.

WORKS ON HUS

There are several useful biographies in English, including D. S. Schaff, *John Huss* (New York: Scribners, 1915), and Francis Lützow, *Life and Times of Master John Hus* (London: Dent, 1909). Johann Loserth, *Hus and Wiclif* (London: Hodder and Stoughton, 1884), is bitterly anti-Hus, arguing that Hus borrowed all his ideas from Wyclyf. The classic study in Czech is by V. Novotný and V. Kybal, *M.J. Hus, Život a Učení,* 5 vols. (Prague: Laichter, 1919–1931). See also two more recent and important studies by P. de Vooght, *L'hérésie de Jean Huss* (Louvain, 1960) and *Hussiana* (Louvain: Bibliothèque de l'Université, Bureaux de la Revue, 1960).

Spinka, Matthew. *John Hus' Concept of the Church.* Princeton, NJ: Princeton University Press, 1966.

Spinka, Matthew. *John Hus: A Biography.* Princeton, NJ: Princeton University Press, 1968.

S. Harrison Thomson (1967)
Bibliography updated by Christian B. Miller (2005)

HU SHI
(1891–1962)

Hu Shi, the Chinese pragmatist, was educated in China, at Cornell University, and at Columbia University under John Dewey. He was successively professor, chancellor of Peking National University, ambassador to the United States, and president of Academia Sinica in Taipei, Taiwan.

In 1916 he inaugurated the Literary Revolution in China by advocating the use of the vernacular style for writing instead of the formal, classical style, which, radically different from the spoken language, had become rigid and decadent. He succeeded in spite of strong opposition and thus set Chinese literature free. Since freedom of expression means also freedom of thought, the new literature led to the Intellectual Renaissance in China in 1917.

Hu did not claim to be a philosopher, but his own credo represented a new philosophy in China at the time. According to Hu Shi the universe, infinite in space and time, was not supernaturally created but is naturalistic and is governed by natural laws. All things, including psychological phenomena, have a scientific basis and can therefore be scientifically understood. Immortality is not personal but the sum total of individual achievement living on in the Larger Self. Truth must be historically and scientifically tested and is best expressed in democracy, freedom, progress, and social action.

His contributions to Chinese philosophy are important. As the leading disciple of Dewey in China, in 1919 he introduced pragmatism, which exerted tremendous influence and became the first concerted philosophical movement in twentieth-century China. Although the philosophy declined in influence in the later 1920s, its spirit of practical application, emphasis on problems

instead of theories, the insistence on results, the critical approach, and the scientific method had become the generally accepted outlook in China.

In his writings on Chinese philosophy Hu Shi was the first to give it a clear outline, free from religious beliefs and legendary philosophy. He provided it with a historical and social environment. Laozi, for example, was presented as a rebel against oppressive government and hypocritical society. Hu Shi discovered the methodology in Chinese philosophy, notably the "rectification of names" in Confucianism, the "three standards" or "laws of reasoning" in Mohism, and the method of "names and actuality" in other philosophers. He removed the mysticism of Laozi and Zhuangzi, whom he regarded as realists championing the cause of complete individual freedom. While these views are extreme, he created an entirely new atmosphere in Chinese philosophy.

See also Chinese Philosophy; Dewey, John; Laozi; Pragmatism; Zhuangzi.

Bibliography

For a complete listing of his works in both Chinese and English, see *Bulletin of the Institute of History and Philology, Academia Sinica,* Vol. 28 (Taipei, Taiwan, 1957), 889–914. For literature on Hu Shi, see Wing-tsit Chan, "Hu Shih and Chinese Philosophy," in *Philosophy East and West,* Vol. 6, No. 1 (April 1956), 3–12.

Fan, Kuang-huan. "A Study of Hu Shih's Thought." Ph.D. New York University, 1963.

Hu, Shi. *The Chinese Renaissance.* Chicago: University of Chicago Press, 1934.

Hu, Shi. *The Development of the Logical Method in Ancient China.* Shanghai: The Oriental Book Co., 1922.

Wing-tsit Chan (1967)
Bibliography updated by Huichieh Loy (2005)

HUSSERL, EDMUND
(1859–1938)

Edmund Husserl (1859–1938), the founding figure of the philosophical movement known as phenomenology, was born in Prossnitz in Mähren, then part of the Hapsburg Austro-Hungarian Empire but now Prostějow in Moravia in the Czech Republic. Husserl studied astronomy at Leipzig from 1876 to 1878 and mathematics in Berlin from mid-1878 to 1881 under the eminent mathematicians Karl Weierstrass (1815–1897) and Leopold Kronecker (1823–1891). Husserl completed his mathematical training in Vienna, receiving the PhD in January of 1883, and while completing his degree, he attended the philosophy lectures of Franz Brentano (1838–1917). Husserl went back to Berlin briefly for further study with Weierstrass, but soon returned to Vienna to study philosophy again with Brentano from 1884 to 1886. At Brentano's suggestion, Husserl studied with Carl Stumpf (1849–1936) at the university at Halle, where in 1887 he submitted a *Habilitationsschrift* titled "Über den Begriff der Zahl. Psychologische Analysen." Husserl taught at Halle from 1887 to 1901, at Göttingen from 1901 to 1916, and at Freiburg from 1916 until his retirement in 1928.

Husserl published relatively little during his lifetime, and his publications were for the most part a series of introductions to phenomenology that were largely methodological and programmatic. However, these works were far from the total of his output. At his death he left more than forty-five thousand pages of unedited manuscripts written in shorthand, the continuing publication of which since 1950 has shed much light on the details and development of Husserl's philosophy.

PSYCHOLOGISM, PSYCHOLOGY, AND PHENOMENOLOGY

Husserl's *Philosophie der Arithmetik* (1891) attempts to realize Weierstrass's program of grounding mathematics in the cardinal numbers by describing those mental acts in which we are conscious of cardinal numbers. While Husserl was satisfied with his discussion of the intuitive presentation of the lower cardinals, he was dissatisfied with the psychologism in his analysis of the symbolic presentation of the higher cardinal numbers. Internal exigencies in Husserl's continued reflections on logic and mathematics—even by 1891—eventually turned him away from psychologism. By 1893 and 1894, Husserl clearly distinguished the subjective presentation, that is, the psychological act presenting an object, from both the logical content of the presentation and the object presented in the presentation, a threefold distinction much indebted to Bernard Bolzano (1781–1848) and Kasimir Twardowski (1866–1938). Husserl in the following years completed his critique of psychologism, culminating in his lectures on logic at Halle, lectures that form the basis for the *Prolegomena* to the *Logische Untersuchungen* (1900–1901), which is considered by many the locus classicus of the critique of psychologism.

Husserl criticizes psychologism for its reduction of the ideality and transcendence of logical objects (e.g., meanings, concepts, judgments, number, and so forth) to the reality and immanence of psychological contents, and for its reduction of the ideality and universality of logical laws to the factuality and generality of empirical, psycho-

logical laws. While rejecting psychologism, Husserl does not, however, altogether reject the descriptive-psychological approach of the *Philosophie der Arithmetik* or its results, for his anti-psychologism is united with the recognition that insofar as logical laws govern the ideal, objective content of acts of thinking, the relation between these ideal contents and the acts in which they are thought must be elucidated. Husserl's problematic in the main body of the *Logische Untersuchungen*, then, is to account for the relation between meaning and mind while preserving the objectivity and ideality of meaning. He typically poses this problem as a problem in epistemology, specifically, the problem concerning the relationship between the subjectivity of knowing and the objectivity of what is known. So, Husserl is committed to finding a new, nonpsychologistic epistemology to account for the relations among acts, ideal contents, and objects.

In the first edition of the fifth of the *Logische Untersuchungen* Husserl identifies phenomenological contents with psychological contents and distinguishes these from intentional contents. Ideal, intentional contents, in other words, are not properly included within the scope of a phenomenological description, and Husserl must account for meaning by appealing solely to the phenomenological-psychological contents. This is suspiciously close to a psychologism that accounts for meaning by focusing on the act. In the discussion of expressive acts in the first investigation Husserl avoids this conclusion by making the contents of the act that account for its intentional directedness the instantiation of an ideal essence, a meaning-species. The meaning itself remains objective and ideal, and the particular act's relation to this ideal meaning is one of instantiation such that the expressive act intends an object, whether or not that object exists, by means of conferring this meaning on a sensible sign.

Husserl contrasts these meaning-conferring intentions with fulfilling intentions that involve the actual presence of the object to consciousness and therefore involve some intuitive dimension. As Husserl later in the sixth of the *Logische Untersuchungen* turns to the discussion of these fulfilling acts, he recognizes that there are problems in his general account of meaning. Because fulfilling acts present the objects emptily intended in expressive acts, the sense of the fulfilling act is rooted in the object itself rather than in an ideal meaning-species. It is the sense of the object, the significance it has for us in its actual presence, that confirms or disconfirms what we intend as its sense in the expressive act that confers meaning on a sensible sign. Hence, Husserl recognizes that an account of meaning cannot focus exclusively on the subjective conditions of objective knowledge.

In the second edition of the *Logische Untersuchungen* and *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch* (both 1913), Husserl incorporates the intentional contents into the phenomenological contents on which he reflects. He thereby turns to the investigation of the correlation between what he in *Ideen* calls the *noesis*, that is, the intending act, and the *noema*, that is, the object just as intended. Some interpreters of Husserl's theory of intentionality as expressed in *Ideen* understand the noema to be an abstract, ideal intensional entity ontologically distinct from the intended object. This abstract entity can in turn be understood on the model of the *Logische Untersuchungen* as a type that is tokened in different acts having the same determinate object, or it can be understood as an abstract particular by means of which an object is intended. On both interpretations, the noema serves as a mediator between the act and its intended object. Other interpreters, however, claim that Husserl's continuing reflections on intentionality, especially those acts that can serve as fulfilling acts in which the object is intuitively grasped, made him aware of the philosophical difficulties in saying that the act's intentional relation to an object is mediated by an abstract entity. For these interpreters, the noema is the intended object just as intended (whether or not that object actually exists), and the object is the identity in the manifold of noematic presentations (whether veridical or not).

THE PHENOMENOLOGICAL REDUCTION AND TRANSCENDENTAL PHENOMENOLOGY

Husserl's goal was to develop a new philosophical science as the radical critique of the possibility of experience, a science that did not take the possibility of cognition for granted. However, because any science existing on the same plane as the natural and psychological sciences already presupposes both the possibility and the general validity of the experience of the world, this new science must exist on a different plane. This new plane—the plane of transcendental subjectivity—is disclosed by the methodological technique of the phenomenological reduction. Reminiscent of the universal Cartesian doubt, it is nevertheless different therefrom. Whereas the distinguishing characteristic of Cartesian doubt is that it annuls the positing of an object's existence or the validity of a judgment, the distinguishing characteristic of the phenomenological reduction is that it withholds participa-

tion in the positing of the existence of objects and the general validity of experience that characterizes one's natural experience—a positing Husserl characterizes as the general thesis of the natural attitude.

In suspending one's participation in the affirmation characteristic of ordinary experience, the objects given in experience are not lost to reflection but are instead considered only as presumed existents. They remain available for reflection just insofar as they are experienced; the index attaching to them, however, has changed, and their status as objects of experience has been modified so that they are now viewed exclusively in their being as objects of that experience in which they are originally posited. Concrete transcendental subjectivity includes its object as intended without reducing that object to an immanent, psychological content.

The reduction is a change in attitude that leads our attention back to the subjective achievements in which the object as experienced is disclosed in a determinate manner and to the achievements in which we realize the evidence appropriate to confirming or disconfirming their natural experiences. The reduction, in other words, leads our attention to the intentional correlation itself, and Husserl's discussions of intentionality and the reduction are inseparable. The subjective achievements, insofar as they are the medium of access to objects as experienced, have a certain kind of priority over the object that they disclose, but Husserl does not believe that all intelligibility derives from these achievements. The investigation of intentional achievements reveals: (1) how it is that we come to experience objects in determinate manners, including those objects that are always already there for us as transcendental subject before thinking becomes active in the world; (2) how our different experiences are related to one another, and, therefore (3) how the different kinds and levels of objectivity are related; and, finally (4) how our experience confirms or disconfirms in fulfilling intentions what was emptily or mistakenly intended.

Natural straightforward experience is directed to objects in their significance for us. However, it is possible to adjust the manner in which we attend to the object, and when doing so we focus attention not on the object as such but on its significance for us, its noematic sense. This is not to turn our attention to some different entity called a *sense* or *meaning*; it is simply to refocus attention from the significant object to the significance of the object as the object of an intending act. The methodological point picks out what the substantive analyses of meaning reveal as a way of proceeding; that is, we need to focus our attention on both the subjective and objective conditions of meaning by focusing on the essential features of the correlation between the noetic and noematic dimensions of the experiences in which objects are disclosed in determinate ways. To turn our attention to this correlation is to perform the phenomenological reduction.

TEMPORALITY AND PASSIVE SYNTHESIS

The revision of the theory of intentionality and the development of the methodological principle of the phenomenological reduction are two of the three major developments in Husserl's thought during the Göttingen years. The third is the development of his views on the nature of the consciousness of inner time, a development that leads to the disclosure of absolute consciousness. A phenomenological description of the awareness of experience as temporally extended—that is, as beginning in the past, enduring in the present, and aimed at the future—requires that Husserl distinguish two strata in consciousness: (1) the nontemporal, time-constituting absolute consciousness that makes possible the awareness of inner time by virtue of a compound intentionality directed at once to the now, the just elapsed, and the yet to come; and (2) the flow of temporally ordered experiences themselves. This distinction accounts at once for the temporality of lived experience, for the momentary, prereflective awareness of that experience as a temporal unity, and for the prereflective self-awareness of one's own temporally ordered and unified stream of experiences.

Whereas the revisions in the theory of intentionality and the methodological discussions centered around the phenomenological reduction find their way into *Ideen*, the reflections on the nature of inner time-consciousness and absolute consciousness, which had reached a mature form by 1911, do not. The implications of the reflections on time-consciousness point toward a less static and more genetic account of the origin of sense or meaning, an account whose development becomes a dominant aspect of Husserl's reflections in the 1920s. These analyses, which develop an approach known as genetic phenomenology, take the form of extensions of the theory of time-consciousness, and in them Husserl describes the intentionalities at work in what he calls passive syntheses. These syntheses occur on two levels: the primary passivities of near and distant association and the secondary passivities of history, tradition, and community. This project comes to fruition in *Formale und transzendentale Logik* (1929) and *Die Krisis der europäischen Wis-*

senschaften und die transzendentale Phänomenologie (1936).

TRANSCENDENTAL LOGIC

Formale und transzendentale Logik, in which Husserl returns to the issue of the grounding of logical and mathematical sciences, brings his career full circle. The nature of logic cannot be fully clarified without the phenomenological reduction, for the reduction enables us to see more clearly how sense—in a manner relevant for logic—arises in our experience. Acts of judging are directed to categorially formed, complex states of affairs. The logical or apophantic domain first emerges in a critical turn occasioned by a concern with the truth or falsity of judgments. The positing involved in the straightforward encounter of objects and states of affairs is neutralized. However, this positing is not denied or negated; nor does the original state of affairs disappear from view to be replaced by a new entity—the proposition—that was always there but an unnoticed mediator in our intentional relation to the state of affairs. Instead attention is turned to the objective sense of the state of affairs as intended in the judging, and this objective sense is considered simply as a supposition in order to seek confirmation or disconfirmation of the state of affairs as supposed.

The judged state of affairs and the proposition are properly distinguished, therefore, by means of a difference in the way the meant objectivity is apprehended. Straightforward experience apprehends the categorial objectivity or state of affairs as such; critical reflection apprehends the state of affairs as supposed—that is, the supposition or proposition and, more precisely, the noematic sense of the intended state of affairs. Such critical or propositional reflection is continuous with our natural concern with the way things are. The natural concern with the truth of things is addressed in the interplay between the critical and natural attitudes, between the judgment as such and the state of affairs as such, between propositional reflection and the categorial intuition of states of affairs. Although it is only phenomenological reflection that grasps clearly what occurs in our apprehension of the logical domain, the critical reflection that focuses on the sense or logical content of an experience is different from the phenomenological reflection that views the object as the correlate of an intending. In a phenomenological reflection, the proposition is considered not in relation to the state of affairs straightforwardly experienced, but in relation to the critical experience in which it is intended.

Formale und transzendentale Logik also explores the relation between the Aristotelian and Leibnizian traditions in logic. Husserl contrasts formal apophantics derived from Aristotle with formal ontology derived from Leibniz's notion of *mathesis universalis*, but he also views them as inseparably united. The ground of their unity is the intentional relation between acts and their objects. Formal ontology results from the articulation of the formal structures, relations, and combinations of objects. Formal logic arises from the articulation of these same formal structures, relations, and combinations considered as meanings, as objective states of affairs merely as supposed. The meaning-forms are teleologically ordered toward fulfillment in our intuitive apprehension of object-forms. If the meanings are confirmed in fulfilling experiences, then the identity obtaining between meaning-forms and object-forms is disclosed. The identity-in-correlation of the logical and the ontological, therefore, is properly and fully a *mathesis universalis* realized only at what Husserl calls the third level of logic, the logic of truth.

While both the *Méditations cartésiennes* (1931), which, like *Ideen*, present an overview of Husserl's transcendental philosophy, and *Formale und transzendentale Logik* incorporate the results of Husserl's reflections on time-consciousness and passive synthesis, they remain focused on the nature of theoretical knowledge and the objectivity appropriate to it. They point to the need for regressive inquiries into the constitution of sense, inquiries that reveal the layering of sense over time and its development in intersubjective communities of inquirers. However, they continue to neglect in large part the historicality of the experiences themselves. Husserl addresses this question most explicitly in his last work *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie* (1936), as well as in texts collected and published posthumously as *Erfahrung und Urteil* (1939) and *Analysen zur passiven Synthesis* (1966).

HISTORY AND ETHICS.

The *Krisis* emphasizes how experiences in both the natural sciences and philosophy itself are formed within the context of living traditions. In this context, Husserl identifies the important notion of the life-world. His account of the life-world is somewhat ambiguous. It means at different times: (1) an abstractly conceived world on which higher meanings of the sort belonging to science, philosophy, and culture in general are grounded; and (2) the concrete world that is already pregiven and taken for granted in our experience, a world that already includes

the sedimented deposits of the history of science, philosophy, and culture. The first sense captures Husserl's idea that different levels of experience are built on more fundamental levels, and this abstract notion of the life-world is the meaning-fundament on which higher levels of sense are built. The second sense captures the idea that experience of the world is already historically formed in secondary passivities before someone comes to think actively about that world. This world is already rich in emotional dimensions, functional and practical dimensions, theoretical dimensions, and cultural dimensions. New experiences—new ways of making sense—both depart from this world and contribute to it. Although experience has an appropriate historicality, this does not negate Husserl's view that the ideal meanings constituted in experience can, in certain cases such as logic and mathematics, be trans-temporal in character.

Another dimension in which the concrete historicality of experience plays an important role is in Husserl's ethical reflections. His move to Freiburg had occurred in a period of political turmoil that soon turned into personal tragedy and that affected his philosophy profoundly. Husserl, who lost one son and whose other son suffered serious injuries in World War I, saw both the war and its aftermath as a sign of a loss of faith in reason. Moreover, although Husserl's postretirement years were active with continued writing and speaking, he was, after the rise of the Nazis to power, no longer free to teach or lecture in Germany. What for Husserl had, early in his career, been a philosophical crisis regarding the proper grounding of knowledge now at Freiburg revealed itself as a cultural crisis, the loss of faith in reason itself. In hindsight it can be said that there was always a moral urgency at the center of Husserl's philosophy, a moral imperative to retrieve a proper sense of rationality and to develop a sense of self-responsibility in which each person seeks the truth and decides about it for himself or herself in the light of evidence.

Husserl's early ethical reflections are centered around two themes: (1) values are constituted in emotional experiences that are grounded in objectifying acts; and (2) there is a need for a formal axiology and a formal theory of practice—both analogous to formal logic—that will counter ethical empiricism, analogous to psychologism, and establish universal moral norms. These two themes are in some tension. After World War I, however, Husserl focuses on the first theme and speaks of vocations—that is, commitments to certain goods that order and give moral meaning to life—and of absolute values grounded in love. Such language makes the enunciation of universal moral principles more difficult, but Husserl never abandoned his commitment to rationality in ethics. But his notion of reason was an expanded one; it is not merely theoretical reason, but axiological reason and practical reason. Just as theoretical reason is teleologically ordered toward the fulfillment of empty cognitive intentions, both axiological reason and practical reason are teleologically ordered toward the evidential fulfillment of empty axiological and volitional intentions.

It is just this commitment to reason and to fulfilling evidences that characterize the moral urgency at the center of all of Husserl's reflections. All are born into moral communities, but all must decide for themselves about what is truly good and about what emotions and actions are appropriate for different circumstances. And if one's vocation is a theoretical or philosophical one, then the search for truth regarding the transcendental conditions for truthfully encountering a world that has intertwined cognitive, affective, axiological, practical, and cultural dimensions must be the unwavering goal of one's reflections.

HUSSERL'S SUCCESSORS

The continual publication of Husserl's unedited manuscripts not only provides a more complete view of Husserl's thought and its development, but also affects one's view of the relations between Husserl and his successors. Discussions of embodiment, intersubjectivity, passive synthesis, community, tradition, and the life-world were all present in Husserl's work before any of the major works of his successors appeared. This implies that the understanding of these differences between Husserl and his successors must be carefully nuanced. In particular, the idea of a pure consciousness or ego separate from both the empirical ego and the world must be rejected in the light of the discussions of embodiment, volition, and historical community. Nevertheless, Husserl could criticize Merleau-Ponty and Heidegger for failing properly to distinguish, respectively, psychology and transcendental phenomenology or anthropology and transcendental phenomenology. Moreover, given his views on temporality and intentionality, Husserl could criticize Sartre for an inadequate view of the ego and a too voluntaristic account of intentionality. Just as his successors' critiques of him must be nuanced by what is known of Husserl's unedited writings, so too Husserl's critiques or potential critiques must not rely on too sharp a distinction between himself and his successors. This is especially the case with those analyses undertaken by his successors that in their own way involve a transcendental reduction even as they

emphasize more than Husserl did the worldliness and existential condition of the subject.

See also Phenomenology.

Bibliography

WORKS BY HUSSERL

Edmund Husserl Collected Works. The Hague: Martinus Nijhoff, and Dordrecht: Kluwer, 1980–. Ten volumes to date of English translations of Husserl's works supervised by the Husserl Archives:

Vol. 1. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. Third Book: *Phenomenology and the Foundations of the Sciences*. Translated by Ted E. Klein and William E. Pohl, 1980.

Vol. 2. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. First Book: *General Introduction to a Pure Phenomenology*. Translated by Fred Kersten, 1983.

Vol. 3. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. Second Book: *Studies in the Phenomenology of Constitution*. Translated by Richard Rojcewicz and André Schuwer, 1989.

Vol. 4. *On the Phenomenology of the Consciousness of Internal Time (1893–1917)*. Translated by John B. Brough, 1991.

Vol. 5. *Early Writings in the Philosophy of Logic and Mathematics* Translated by Dallas Willard, 1994.

Vol. 6. *Psychological and Transcendental Phenomenology and the Confrontation with Heidegger (1927–1931)* Translated by Thomas Sheehan and Richard E. Palmer, 1997.

Vol. 7. *The Idea of Phenomenology*. Translated by Lee Hardy, 1999.

Vol. 8. *Thing and Space: Lectures of 1907*. Translated by Richard Rojcewicz, 1997.

Vol. 9. *Analyses concerning Passive and Active Synthesis: Lectures on Transcendental Logic*. Translated by Anthony J. Steinbock, 2001.

Vol. 10. *Philosophy of Arithmetic: Psychological and Logical Investigations, with Supplementary Texts from 1887–1901*. Translated by Dallas Willard, 2003.

Cartesian Meditations: An Introduction to Phenomenology. Translated by Dorion Cairns. The Hague: Martinus Nijhoff, 1960.

Formal and Transcendental Logic. Translated by Dorion Cairns. The Hague: Martinus Nijhoff, 1969.

Logical Investigations. Translated by J. N. Findlay. London: Routledge and Kegan Paul, 1970.

The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy. Translated by David Carr. Evanston, IL: Northwestern University Press, 1970.

Experience and Judgment: Investigations in a Genealogy of Logic. Translated by James S. Churchill and Karl Ameriks. Evanston, IL: Northwestern University Press, 1973.

WORKS ON HUSSERL

Bell, David. *Husserl*. London: Routledge, 1990.

Bernet, Rudolf. *La vie du sujet. Recherches sur l'interprétation de Husserl dans la phénoménologie*. Paris: Presses universitaires de France, 1994.

Bernet, Rudolf, Iso Kern, and Eduard Marbach. *An Introduction to Husserlian Phenomenology*. Evanston, IL: Northwestern University Press, 1993.

Bernet, Rudolf, Iso Kern, and Eduard Marbach. *Edmund Husserl. Darstellung seines Denkens*. Hamburg: F. Meiner, 1989.

Brainard, Marcus. *Belief and Its Neutralization: Husserl's System of Phenomenology in* Ideas I. Albany: State University of New York Press, 2002.

Brough, John B. "The Emergence of an Absolute Consciousness in Husserl's Early Writings on Time-consciousness." *Man and World* 5 (1972): 298–326.

Buckley, R. Philip. *Husserl, Heidegger and the Notion of Philosophical Responsibility*. Dordrecht: Kluwer, 1992.

Carr, David. *Phenomenology and the Problem of History: A Study of Husserl's Transcendental Phenomenology*. Evanston, IL: Northwestern University Press, 1974.

Cobb-Stevens, Richard. *Husserl and Analytic Philosophy*. Dordrecht: Kluwer, 1990.

Crowell, Steven. *Husserl, Heidegger, and the Space of Meaning: Paths Toward Transcendental Phenomenology*. Evanston, IL: Northwestern University Press, 2001.

Drummond, John J. *Husserlian Intentionality and Non-Foundational Realism: Noema and Object*. Dordrecht: Kluwer, 1990.

Hart, James G. *The Person and the Common Life: Studies in a Husserlian Social Ethics*. Dordrecht: Kluwer, 1992.

Kern, Iso. *Husserl und Kant. Eine Untersuchung über Husserls Verhältnis zu Kant und zum Neukantianismus*. The Hague: Martinus Nijhoff, 1964.

McIntyre, Ronald, and David Woodruff Smith. *Husserl and Intentionality: A Study of Mind, Meaning and Language*. Dordrecht: D. Reidel, 1982.

Mohanty, J. N. *Edmund Husserl's Theory of Meaning*. The Hague: Martinus Nijhoff, 1964.

Mohanty, J. N. *Husserl and Frege*. Bloomington: Indiana University Press, 1982.

Mohanty, J. N. *The Possibility of Transcendental Philosophy*. Dordrecht: Martinus Nijhoff, 1985.

Sokolowski, Robert. *Husserlian Meditations: How Words Present Things*. Evanston, IL: Northwestern University Press, 1974.

Sokolowski, Robert. *Moral Action: A Phenomenological Study*. Bloomington: Indiana University Press, 1985.

Sokolowski, Robert. *Presence and Absence: A Philosophical Investigation of Language and Being*. Bloomington: Indiana University Press, 1978.

Sokolowski, Robert. *The Formation of Husserl's Concept of Constitution*. The Hague: Martinus Nijhoff, 1964.

Waldenfels, Bernhard. *Das Zwischenreich des Dialogs. Sozialphilosophische Untersuchungen im Anschluss an Edmund Husserl*. The Hague: Martinus Nijhoff, 1971.

Welton, Donn. *The Other Husserl: The Horizons of Transcendental Phenomenology*. Bloomington: Indiana University Press, 2000.

Willard, Dallas. *Logic and the Objectivity of Knowledge: A Study in Husserl's Early Philosophy*. Athens: Ohio University Press, 1984.

Zahavi, Dan. *Husserl's Phenomenology*. Stanford, CA: Stanford University Press, 2003.

Zahavi, Dan. *Self-Awareness and Alterity: A Phenomenological Investigation*. Evanston, IL: Northwestern University Press, 1999.

John J. Drummond (2005)

HUTCHESON, FRANCIS

(1694–1746)

Francis Hutcheson, a moral-sense theorist, was born at Drumalig in County Down, Ulster. His father and grandfather were Presbyterian ministers. In 1711 he entered the University of Glasgow, taking both the arts and theological courses and probably finishing in 1717. He was licensed as a probationer preacher by the Ulster Presbyterians in 1719. Not long after, he was invited by the Presbyterians of Dublin to found a dissenting academy for their youth, and he remained in Dublin for the next ten years as head of the academy. His stay there was a turning point in the development of his thought, for he came under the influence of admirers of the Earl of Shaftesbury's philosophy. Hutcheson's first two, and perhaps most important, books were published during this period. The University of Glasgow elected Hutcheson to its professorship of moral philosophy in 1730, a position that he held until his death. In 1746, while visiting Dublin, he contracted a fever and died.

At Glasgow, Hutcheson devoted himself to enriching the culture and softening the Calvinism of his fellow Presbyterians. The Presbytery of Glasgow tried him for teaching, in contravention to the Westminster Confession, the following "false and dangerous" doctrines: (*a*) that the standard of moral goodness is the promotion of the happiness of others and (*b*) that it is possible to have a knowledge of good and evil without, and prior to, a knowledge of God. Afterward, Hutcheson was able to speak of the matter as the "whimsical buffoonery" about his heresy, but the fact that the charges were brought is doubtless a measure of the effectiveness of his teaching. David Hume sent a draft of Part III of *The Treatise of Human Nature,* "Of Morals," to Hutcheson for his comments prior to publication. Some indication of the spirit in which Hutcheson wrote his own work can be gathered from his rebuking Hume for a lack of warmth in the cause of virtue, which "all good men would relish, and could not displease among abstract enquiries."

THE MORAL SENSE

Hutcheson's contributions to philosophy lie in aesthetics and moral philosophy. In the one he offers a theory of an internal sense by which we perceive beauty, and in the other he offers a theory of a moral sense by which we perceive and approve virtue and perceive and condemn vice. Hutcheson meant his theory of the moral sense to be a contribution to the contemporary discussion of how to analyze man's moral knowledge. There were two sides in the discussion. Samuel Clarke and his followers held that moral distinctions are made by reason on the basis of our knowledge of the unchanging and unchangeable fitness of things. The other side, owing its original allegiance to Shaftesbury, held that moral distinctions are the deliverances of a moral sense.

Both sides held two points in common. First, moral knowledge must be accounted for by showing how it can be acquired by the exercise of some human faculty. In this respect they were all Lockeans: If something is knowable, you must show how it can be perceived. Second, moral knowledge cannot be simply a revelation from God, though of course God may enter the picture indirectly by having endowed us with our moral faculty. And when it came to picking out actual instances of virtue and vice, both sides were in agreement about the value of benevolence and the wrongness of acts of violence against other persons. Their debate, then, was over the character of the moral faculty.

PERCEPTION AND APPROVAL OF VIRTUE. Hutcheson plucked from Shaftesbury's rhapsodies the notion of a moral sense and endeavored to give a systematic account of it as the moral faculty of humankind. To see what Hutcheson's claim means, we must first of all consider what led him to make it. When you see someone doing something that is helpful to another, you say that his action is a virtuous one. But why is a helpful action counted as virtuous? It might be said that a helpful action is virtuous because it exhibits benevolence. But this does not take us very far, for we may still ask why benevolence is a criterion of virtue. Hutcheson knew the answer that some moral writers had given to this question: Helpfulness or benevolence is a possible relation between two human beings, and it is a fitting one. Therefore, it is virtuous. But how do you tell what is fitting and what is not? Your reason tells you. At this point, however, Hutcheson asked whether fittingness could be discovered by reason. After all, reason can tell us only that a certain relation does or does not exist; the moral quality of the relation, if any, remains to be apprehended. But by what?

Certainly not by reason, Hutcheson argued, because the moral quality is not a relation. And if not by reason, then the only thing left is a sense: the moral sense.

Hutcheson's task was to offer an account of how the moral sense works. He located the moral sense on the map of Lockean psychology. Its deliverances are ideas of reflection that arise from our original perceptions of human actions. As he first described the moral sense, it is a determination of our minds "to receive amiable or disagreeable ideas of actions." The "amiable idea" or, as he sometimes spoke of it, "our determination to be pleased," has two jobs. It is both our perception of the virtue of an action and our approval of it. It so happens that those actions which we discern to be virtuous are always benevolent actions, and we are necessarily determined to discern their virtue as soon as we observe them. Hutcheson attributed both the connection between virtue and benevolence and our necessary perception of the virtuousness of benevolence to arrangements superintended by God. Like sight, the moral sense is universal in humankind. But just as some men are born blind, and others have defective sight, some men have no moral sense at all and others have defective moral senses.

The strength of Hutcheson's theory of the moral sense lies in his having given an account of how we know that benevolent actions are virtuous: They are virtuous because they please. He was careful to point out that they please irrespective of any advantage they may have to ourselves. He noticed that we may indeed perceive as virtuous an action that displeases us because it goes against our selfish interest, and we may desire that someone act in a certain way even though we should call it vicious. He also argued that in the first instance the moral sense works independently of education, custom, and example. These factors may strengthen the moral sense but cannot create it, for they really presuppose a moral sense. In order for a person to be given an education in morality, he must already be able to discern moral qualities. Similarly, in order for customs to be moral customs and for examples to be moral examples, morality must already have been discovered in order to give these factors a moral character.

In saying that virtue is what it is because it pleases, Hutcheson thought that he had given a completely satisfactory account of the nature of virtue. By means of the moral sense, virtue is perceived for what it is. It is an end to be sought for itself, and no further characterization of it is required. Hutcheson's critics, however, found that he had paid a disastrous price for making virtue comprehensible by the human understanding. If virtue is that which pleases, then must any action that pleases be virtuous? Why are the actions that exhibit benevolence the only ones that are counted as virtuous? These questions seem bound to be asked despite the stipulations with which Hutcheson hedged his account of our knowledge of virtue.

Both his theory and its difficulties stem from Hutcheson's tacit assumption of the Lockean guide that a piece of knowledge must be accounted for through an appeal to the faculty by which it is known. It was not open to Hutcheson to try the gambit that it would be logically odd to call an act of highway robbery, for example, virtuous. His first line of defense was to insist that the deliverances of the moral sense with respect to virtue are a distinctive kind of pleasure. But in later editions of his *Inquiry concerning Moral Good and Evil,* Hutcheson played down the perceptual function and stressed approving and disapproving. Thus, the moral sense becomes a "determination of our minds to receive the simple ideas of approbation or condemnation, from actions observed … ." To call these ideas simple is to claim that they are not subject to further analysis and, hence, to further characterization. But this new position is not without its own difficulties. Approbation and condemnation are dispositions, not sensations; and only a most slavish allegiance to John Locke's model of the mind could lead one to construe all mental acts as perceptions.

MOTIVATION TO VIRTUOUS ACTION. Hutcheson's theory of the moral sense has yet a third part. As well as using it to account for the perception and approbation of virtue, he also used it to account for a person's motivation to behave in a virtuous way. A person pursues virtue because virtuous acts are pleasing to him and avoids vicious acts because they pain him. This account of moral motivation is perhaps the most convincing part of Hutcheson's theory. It enabled him to close the gap between someone's knowing what ways of acting are virtuous and his being inclined to act virtuously. Yet even here the theory gives us less than we might hope, for someone will be motivated to act benevolently only if benevolence pleases. And if other ways of acting please, even malevolence, perhaps, what then? Once more Hutcheson entered a stipulation too pat to be absolutely convincing: God has determined most people to be benevolent. Once again we must admit that he took this position for the best of reasons, for he was opposing those who would reduce all human motives to self-interest—and the many disinterested actions that people perform show the absurdity of this contention. But what Hutcheson's account of moral motivation requires is not the sensation of being pleased

with benevolence but a disposition to be benevolent. Unfortunately, the psychological theory on which Hutcheson relied required him to reduce all mental phenomena to some sort of perception. Thus, his account of motives lacks an effective analysis of dispositions.

AESTHETICS

Hutcheson's aesthetics closely parallels his theory of the moral sense. He found that we have an internal sense of beauty, a determination to be pleased by regular, harmonious, uniform objects, by grandeur, or by novelty. These perceptions occur necessarily and independently of our wills, but there is no corresponding "pain or disgust, any farther than what arises from disappointment." This limitation seems to have the curious consequence of leaving Hutcheson no room to account for perceptions of the ugly. The task of approving of the beautiful is not assigned to our sense of beauty. Presumably Hutcheson thought indifference to beauty allowable but indifference to virtue never so.

ROLE OF PASSIONS AND REASON IN MORAL LIFE

Hutcheson's *Essay on the Nature and Conduct of the Passions and Affections* and *Illustrations upon the Moral Sense* (published jointly in London, 1728) supplement the part of the *Inquiry* devoted to morals. In the essay on the passions, Hutcheson defined *sense* as every determination of the mind either to receive ideas independently of the will or to have perceptions of pleasure or pain. This definition led to the introduction of several new senses into Hutcheson's system. For instance, there is a public sense, which is our determination to be pleased by the happiness of others and to be uneasy at their misery. There is also the sense of honor, which makes the approbation or gratitude of others for any actions we have done the necessary occasion of pleasure.

In the *Illustrations upon the Moral Sense,* Hutcheson's general aim was to characterize the role of reason in the moral life. With regard to actions, Hutcheson said that we may reason either to account for what excites someone to act as he does or to account for what justifies our approbation of an act. For example, we give the "exciting" reason when we account for a luxury-loving man's pursuit of money by pointing out that money may be used to purchase pleasures. We give the "justifying" reason when we account for our approving of a man's risking his life in war by pointing out that his conduct evidences public spirit. But it is never true that reasons are to be found independently of feelings, for "exciting" reasons presuppose instincts and affections, and "justifying" reasons presuppose the moral sense.

Supposing that we get our ideas of virtue and vice through a moral sense, Hutcheson acknowledged that there are certain truths which might be proved by reason. These are (1) what actions or affections obtain the approbation of any observer, and what actions or affections obtain condemnation; (2) what quality of actions gains approbation; (3) what actions really evidence kind affections and tend to the greatest public good; and (4) what motives excite men to publicly useful actions.

See also Aesthetics, History of; Aesthetics, Problems of; Beauty; Clarke, Samuel; Hume, David; Locke, John; Moral Epistemology; Moral Sense; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Virtue and Vice.

Bibliography

Hutcheson first presented his philosophy in the *Inquiry into the Original of Our Ideas of Beauty and Virtue* (London, 1725). The *Inquiry* is divided into two parts: "Concerning Beauty, Order, Harmony, Design" and "concerning Moral Good and Evil." It is a much-revised work, the most notable changes occurring in the fourth edition (Glasgow, 1738). Lewis Amherst Selby-Bigge, ed., *The British Moralists* (Oxford: Clarendon Press, 1897), Vol. I, reproduces a substantial part of the second edition of the "Inquiry concerning Moral Good and Evil," which taken by itself is an incomplete representation of Hutcheson's thought. In *A Short Introduction to Moral Philosophy* (Glasgow: R. Foulis, 1747), an English version of the *Philosophiae Moralis Institutio Compendiaria* (Glasgow, 1742), Hutcheson uses "conscience" interchangeably with "moral sense," a possible sign of Butler's influence. Hutcheson is also the author of *A System of Moral Philosophy,* published posthumously in two volumes (London: A. Millar, 1755) by his son Francis. The *System* contains a biography by William Leechman. Other works by Hutcheson are *Metaphysical Synopsis* (Glasgow, 1742) and *Logical Compendium* (Glasgow, 1756).

For Hume's letters to Hutcheson, see J. Y. T. Grieg, ed., *The Letters of David Hume* (Oxford: Clarendon Press, 1932), Vol. I.

Contemporary criticism may be found in *Letters concerning the True Foundation of Virtue or Moral Goodness, wrote in Correspondence between Mr. Gilbert Burnet and Mr. Francis Hutcheson,* edited by Hutcheson (Glasgow, 1772), first published in the *London Journal* (1728).

Biographical and analytical material is contained in W. R. Scott, *Francis Hutcheson* (Cambridge, U.K.: Cambridge University Press, 1900); Ernest Albee, *History of English Utilitarianism* (London: Swan Sonnenschein, 1902); and T. Fowler, *Shaftesbury and Hutcheson* (London: Sampson, Low, Marston, Searle & Rivington, 1882).

For critical discussions, see James Bonar, *Moral Sense* (London: Allen and Unwin, 1930), and D. D. Raphael, *The Moral Sense* (London: Oxford University Press, 1947).

Elmer Sprague (1967)

HUTCHESON, FRANCIS [ADDENDUM]

Although Francis Hutcheson's name is frequently associated with the third Earl of Shaftesbury (1671–1713), the truth is that once he rejects John Locke's (1632–1704) contention that beauty is a complex idea, his aesthetic theory is thoroughly Lockean. That beauty is a simple idea, for Hutcheson, is made clear when he postulates a sense of beauty as necessary to perceive it. And such a move is fully sanctioned by Locke himself when he writes that "I have here followed the common opinion of men's having but five senses, though, perhaps there may justly be counted more" (John Locke, *Essay concerning Human Understanding*, II, ii, 3).

Hutcheson's most succinct and influential statement of his basic position goes as follows: "[T]he word *beauty* is taken for *the idea raised in us*, and a *sense* of beauty for *our power of receiving this idea*" (*Inquiry concerning Beauty*, I, ix, p. 34).

The idea of beauty, which is a simple idea, is caused to be excited in the sense of beauty, however, by a complex idea, namely any collection of ideas that possesses what Hutcheson calls "uniformity amidst variety." The sense of beauty perceives ideas, and not the external world directly, because it is what Hutcheson calls a "reflex" or "subsequent" sense, requiring the five "external" senses to provide its objects.

Few today will find this Lockean account of beauty and its perception at all plausible. Nonetheless, it is no exaggeration that it defined and drove aesthetic speculation through the whole of the eighteenth century in Great Britain. And because it was philosophy at the cutting edge, the influence was good, even though it turned out the doctrine was not.

See also Aesthetics, History of; Beauty; Locke, John; Shaftesbury, Third Earl of (Anthony Ashley Cooper).

Bibliography

Hutcheson, Francis. *An Inquiry concerning Beauty, Order, Harmony, Design*, edited by Peter Kivy. The Hague: Martinus Nijhoff, 1973.

Kivy, Peter. *The Seventh Sense: Francis Hutcheson and Eighteenth-Century British Aesthetics*. 2nd ed. Oxford, U.K.: Clarendon Press, 2003.

Peter Kivy (2005)

HUXLEY, THOMAS HENRY
(1825–1895)

Thomas Henry Huxley, the biologist and the most versatile man of science of nineteenth-century England, was born at Ealing, near London. Like many eminent Victorians, Huxley was self-educated. While still an adolescent he read extensively in history and philosophy, learned several foreign languages, and began a medical apprenticeship. In 1842 he entered Charing Cross Hospital, where he distinguished himself by winning prizes in chemistry, anatomy, and physiology, and by publishing his first research paper. From 1846 to 1850 he was assistant surgeon on H.M.S. *Rattlesnake* while it conducted surveying operations in Australian waters. Huxley made capital out of this voyage, as Charles Darwin had done on the voyage of the *Beagle,* and sent home a number of scientific papers dealing with marine animals. These papers established his reputation as a first-rate biologist, and in 1851 he was elected a fellow of the Royal Society. After leaving the navy he settled in London, where he eventually obtained several small appointments, the chief one being that of naturalist at the Government School of Mines. Here he began his paleontological investigations, which resulted in more than twenty memoirs on the anatomy and classification of fossils. During the next four decades Huxley became one of the intellectual leaders of England. His strong, skeptical, earnest mind was enlisted on behalf of a great variety of causes. He championed Darwin's theory of evolution, disputed with churchmen about the Bible, worked for educational reforms, served on eight royal commissions, and refused a professorship at Oxford. As a public lecturer he was brilliant at clarifying abstruse subjects and developing polemical arguments. He also wrote copiously in forceful, eloquent prose. Yet he produced no really seminal ideas or magnum opus, partly because his efforts were so dispersed. In the following discussion, attention will be limited his views on the nature of science, metaphysics, ethics, and religion.

THE NATURE OF SCIENCE

For Huxley, two aspects of the sciences were of special importance. One was their historical continuity with modes of thought used by men in the ordinary commerce of life. "Science," he once said, "is nothing but trained and organized common sense, differing from the latter only as a veteran may differ from a raw recruit … . The man of science, in fact, simply uses with scrupulous exactness the methods which we all habitually use carelessly" (*Collected Essays,* Vol. III, pp. 45–46). Hence there is a unity of pro-

cedure in all the sciences. This was the other aspect of the sciences that he deemed important, because it allowed a specification to be given of the steps that must be taken if the procedure is to be properly carried out. In an essay of 1854, "On the Educational Value of the Natural History Sciences" (*Collected Essays,* Vol. III), Huxley listed the steps as: (1) observation of the facts, including those elicited by experiment; (2) comparison and classification, leading to general propositions; (3) deduction from the general propositions to the facts again; and (4) verification. Later he came to see that hypotheses are essential to the procedure of science, especially as devices for "anticipating nature." But he did not sufficiently stress the connection between a hypothesis and the scientific problem that initiates an inquiry, or the role of the hypothesis in determining what facts are to be observed.

EVOLUTION

It was the effective use of scientific method in Darwin's *Origin of Species* that helped to convert Huxley to the doctrine of evolution by natural selection. As a young man he had held antievolutionary views, not because he believed in the special creation of species, but because he failed to find a scientific explanation of how their transmutation could have been effected. Darwin's book proved to be "a flash of light which, to a man who has lost himself in a dark night, suddenly reveals a road that, whether it takes him straight home or not, certainly goes his way" (*Life and Letters of T. H. Huxley,* Vol. I, pp. 245–246). His reflection on having mastered Darwin's central thesis was, "How exceedingly stupid not to have thought of that."

Huxley espoused Darwinism not as a dogma, however, but as a "most ingenious hypothesis" that offered a rational account of how the organic world came to be what it is. The hypothesis was not contradicted by any known evidence, nor was it seriously rivaled; yet it was not established beyond a shadow of doubt. For instance, certain physiological peculiarities of organisms, such as hybrid sterility, had still to be explained in terms of natural selection. To Huxley, some of Darwin's formulations seemed quite unsatisfactory. To speak of variations "arising spontaneously" was to employ "a conveniently erroneous phrase." To commit oneself to the principle *natura non facit saltum* ("nature makes no leap") was to invite needless trouble. For in fact, Huxley declared, "Nature does make jumps now and then, and a recognition of this is of no small importance in disposing of many minor objections to the doctrine of transmutation" (*Collected Essays,* Vol. II, p. 77). But even if it remained to be shown that natural selection sufficed for the production of species, "few can doubt that it is a very important factor in this operation." To that extent Darwinism was certainly here to stay.

Huxley was sensitive to a number of philosophical questions generated by the theory of evolution. The questions that particularly interested him arose when three considerations were taken seriously. First, like all scientific theories, Darwinism "starts with certain postulates … and the validity of these postulates is a problem of metaphysics." Second, the theory of evolution had to be extended to the cosmos as a whole, if its scope was not to be arbitrarily restricted. But at that point philosophical issues had to be faced. Did the cosmos evolve from some "epicurean chance-world," or had its order been eternally the same? Finally, the study of organisms pointed to the conclusion that they began as, and are now, physicochemical systems. It could therefore be assumed that molecular motions are the basis of all vital processes, including so-called conscious ones. But if this was so, metaphysical materialism gained strong support.

METAPHYSICS AND EPISTEMOLOGY

The philosophical standpoint most congenial to Huxley was derived from his reading of René Descartes, George Berkeley, and David Hume. Of prime importance was the contention "that our certain knowledge does not extend beyond states of consciousness, or the phenomena of mind. … Our sensations, our pleasures, our pains, and the relations of these, make up the sum total of the elements of positive, unquestionable knowledge" (*Collected Essays,* Vol. VI, pp. 317–318). Beyond this we have only uncertain inferences or beliefs. Hence, when we talk about "matter" and "the physical world," we are interpreting some mental phenomena, just as we are interpreting other phenomena when we talk about "mind" and "the self." For matter is only a postulated cause of certain conscious states, in the same way that mind is a postulated substratum of those same states. This is all that criticism leaves of "the idols set up by the spurious metaphysics of vulgar common sense."

Huxley expressed many of his philosophical ideas in a book on Hume that he wrote for the English Men of Letters series in 1878. He agreed with Hume's account of perception as a process that yields only sense impressions, but he held that Hume had failed "to recognize the elementary character of impressions of relation" and also had failed to make clear that having a sense impression is a case of knowing. Hume had correctly represented the order of nature as an unbroken succession of causes and effects, so that there can be no uncaused volitions such as

proponents of "free will" postulate. But determinism, Huxley affirmed, is entirely compatible with ascribing responsibility to human beings for their deliberate actions. As Hume had rightly understood, "the very idea of responsibility implies the belief in the necessary connexion of certain actions with certain states of mind" (*Collected Essays,* Vol. VI, p. 222).

METAPHYSICAL PRESUPPOSITIONS OF SCIENCE. From this philosophical standpoint, Huxley dealt with questions that fall under the three considerations mentioned above. Science, he affirmed, postulates a rational order of nature, the operation of material forces, the universality of causation, and the immutable necessity of laws. All these factors need to be properly interpreted. Thus, "nature" is simply the totality of phenomena, whose regular occurrence constitutes nature's "rational order." Material forces are at best hypothetical entities which Huxley said he could not conceive clearly. As Hume had insisted, "causation" refers to the relation of invariable succession among phenomena. "Necessity" is a term that should be limited to logic and has no warranted application to the physical world. For the laws that science formulates are records of observed regularities, not agents which "force" things to happen as they do. Hence, "our highest and surest generalizations remain on the level of justifiable expectations, that is, very high probabilities." The quest for certainty in science is an irrational pursuit.

AGNOSTICISM. It is also irrational to hope that we can ever know anything about the ultimate origin or ultimate nature of the universe. Speculation about such matters is fruitless, for they lie outside the limits of philosophical inquiry. To identify his position on this issue, Huxley coined the name "agnostic" about 1869. "It came into my head," he said, "as suggestively antithetical to the "agnostic" of Church history, who professed to know so much about the very things of which I was ignorant" (*Life and Letters of T. H. Huxley,* Vol. I, p. 462). Agnosticism, however, is not another creed; it is an outlook that results from the adoption of a principle, at once intellectual and moral, which states that a man ought not to assert that he knows a proposition to be true unless he can produce adequate evidence to support it. Conversely, an agnostic repudiates as immoral "the doctrine that there are propositions which men ought to believe without logically satisfactory evidence." The justification of this principle lies in the success which follows upon its application, whether in the field of natural or of human history.

Because of an agnostic's outlook, he cannot accept the tenets of metaphysical materialism, according to which nothing exists in the world save matter, force, and necessity. For these three concepts are intelligible only insofar as they are related to the phenomena of mind. Hence, "Materialism is as utterly devoid of justification as the most baseless of theological dogmas." Yet to reject materialism is by no means to espouse idealism or spiritualism. "Spiritualism is, after all, little better than Materialism turned upside down" (*Collected Essays,* Vol. IX, p. 133). Nor does it follow that the sciences must eschew materialistic language. On the contrary, such language is often useful in investigating the order of nature, as Huxley himself showed in more than one paper. But to use materialistic language for scientific purposes is quite different from accepting a metaphysics based on materialism.

EPIPHENOMENALISM. As a biologist, Huxley took the view that the bodies of animals, including humans are best regarded as mechanical systems. Yet the mind and states of consciousness undeniably exist, and their relation to the working of the physical body has to be explained. This was the question Huxley discussed in a well-known paper of 1874, "On the Hypothesis that Animals are Automata." States of consciousness are represented as being no more than effects of bodily processes—chiefly, the molecular changes in brain substance that has attained a certain degree of organization. Furthermore, no evidence can be found for supposing "that any state of consciousness is the cause of change in the motion of the matter of the organism." Animals, then, are conscious automata. The working of their bodily mechanism is unaffected by their mental activity. "The mind stands related to the body as the bell of the clock to the works, and consciousness answers to the sound which the bell gives out when it is struck." This is Huxley's version of epiphenomenalism. The doctrine did not purport to give an ultimate explanation of the mind-body relationship. It did not even purport to explain how the passage from molecular movement to conscious states is effected. Concerning the details of this passage, he declared, "I really know nothing and never hope to know anything."

RELIGION

The greatest impact of Huxley's agnosticism was on the religious dogmas of his time. As a young man he accepted a form of theism. In a paper of 1856, "On Natural History as Knowledge, Discipline, and Power" (*Royal Institution Proceedings,* London, Vol. II, 1854–1858, pp. 187–195), he

contended that the design revealed by nature pointed to the existence of an Infinite Mind as its author. But he discarded this view when he became a Darwinian, on the ground that the argument from design had received its deathblow. Thenceforth, he attacked those who claimed to prove that a supernatural God exists, or who affirmed that biblical and Christian doctrines are rationally credible. His most dramatic clash was with Bishop Wilberforce at the Oxford meeting of the British Association in June 1860, and his most protracted controversy was with W. E Gladstone in the pages of the *Nineteenth Century,* from 1885 to 1891. Neither of these defenders of the faith was a match for Huxley.

MIRACLES. Huxley hammered away at the inconsistencies in, and the lack of evidence for, the biblical cosmology, the creation stories, and the belief in demons, spirits, and miraculous occurrences that Christianity requires. The subject of miracles was of deep interest to him; miracles could not be rejected as impossible, he thought, because they are logically conceivable. Hence Hume's a priori reasoning against them was mistaken. Yet one can say that the occurrence of an alleged miracle, being antecedently a most improbable event, needs strong supporting evidence. But in each recorded case, evidence of this kind was lacking. In several essays Huxley discussed particular biblical reports of miracles and found them unconvincing.

GOD. Although he rejected supernaturalism, Huxley was prepared to accept a Spinozistic conception of God as being identical with nature in its infinite complexity. "The God so conceived is one that only a very great fool would deny, even in his heart. Physical science is as little Atheistic as it is Materialistic" (*Collected Essays,* Vol. IX, p. 140). Once, in a letter to Charles Kingsley, he said that he believed in "the Divine Government" of the universe. The phrase expressed his conviction that the cosmic process is rational rather than random, that the reign of law is universal, and that the order of nature has existed "throughout all duration." Yet the governing principles of the universe appear to be amoral, since what happens to men is "accompanied by pleasures and pains, the incidence of which, in the majority of cases, has not the slightest reference to moral desert" (*Collected Essays,* Vol. IX, p. 202).

ETHICS

Toward the close of his life Huxley thought a good deal about the foundations of morality. He was dissatisfied with the attempts of Darwin and Herbert Spencer to harmonize man's moral sentiments and the theory of evolution. It was not that he doubted the evolutionary origin of those sentiments; what he doubted was whether Darwin or Spencer had appreciated the extent to which morality and nature are at war with each other. This was the theme with which he startled the Victorian world in his famous Romanes lecture, "Evolution and Ethics" (*Collected Essays,* Vol. IX, pp. 46–116), on May 18, 1893.

Its central contention is that "ethical nature, while born of cosmic nature, is necessarily at enmity with its parent." For a dominant feature of the natural world is "the intense and unceasing competition of the struggle for existence." Thomas Hobbes's depiction of nature as the war of all against all is correct. In this world, ruthless and predatory action is "best" for the individual. But in human society, "ape and tiger methods" are precisely what man's moral sense condemns. Hence, the practice of that which is *morally* best involves a repudiation of "the gladiatorial theory of existence" portrayed by Darwinism: "Let us understand, once for all, that the ethical progress of society depends, not on imitating the cosmic process, still less in running away from it, but in combating it." Accordingly, man although himself a product of evolution, has an obligation to subjugate the amoral or immoral aspects of evolution to moral ends. Yet Huxley's grounds for this conclusion are by no means clear. His only recourse was to fall back on a kind of ethical intuitionism which is hardly compatible with his other views.

The philosophical garment that Huxley wove is coarsely textured and has a number of loose ends. Thus his radical phenomenalism is not carefully interwoven with his evolutionism, and his agnosticism seems unconnected with his Spinozistic affirmations. Yet his grasp of philosophical issues was remarkable for a man who was also a leading scientist, educator, and public figure of his time.

See also Agnosticism; Animal Mind; Berkeley, George; Consciousness; Darwin, Charles Robert; Darwinism; Descartes, René; Evolution; Evolutionary Ethics; Hobbes, Thomas; Hume, David; Materialism; Miracles; Natural Law; Phenomenalism.

Bibliography

WORKS BY HUXLEY

Man's Place in Nature. London: Williams and Norgate, 1863.

Hume. New York: Harper, 1879.

Collected Essays. 9 vols. London: Williams and Norgate, 1893–1895.

Scientific Memoirs. 5 vols. London: Macmillan, 1898–1903.

Life and Letters of Thomas Henry Huxley. 2nd ed. 3 vols. Edited by Leonard Huxley. London: Macmillan, 1903.

With Julian Huxley. *Evolution and Ethics, 1893–1943*. London: Pilot Press, 1947; 1969.

The Essence of T. H. Huxley: Selections from His Writings;. Edited by Cyril Bibby. London, Melbourne: Macmillan; New York, St. Martin's, 1967.

On the Origin of Species: Or, The Causes of the Phenomena of Organic Nature. Ann Arbor, University of Michigan Press, 1968.

Autobiography and Essays. Edited by Brander Matthews. New York and Chicago: Gregg, 1969.

Darwiniana. Hildesheim, New York: G. Olms, 1970.

T. H. Huxley on Education. Edited by Cyril Bibby. Cambridge, U.K.: Cambridge University Press, 1971.

Scientific Papers and Correspondence of Thomas Henry Huxley, c. 1843–1895 from the Imperial College of Science, Technology and Medicine, London. Reading, Berkshire, U.K.; Woodbridge, CT: Research Publications, 1990.

Lay Sermons, Addresses and Reviews. Boston: Elibron Classics, 2003.

WORKS ON HUXLEY

Bibby, Cyril. *T. H. Huxley: Scientist, Humanist, Educator.* London: Watts, 1959.

Collie, Michael. *Huxley at Work: With the Scientific Correspondence of T.H. Huxley and the Rev. Dr. George Gordon of Birnie, near Elgin*. Basingstoke, Hampshire, U.K.: Macmillan Press, 1991.

Darwin, Charles, and Thomas Henry Huxley. *Autobiographies [of] Charles Darwin, Thomas Henry Huxley*. Edited by Gavin De Beer. London: Oxford University Press, 1974.

Davis, J. R. A. *Thomas Henry Huxley.* London and New York, 1907.

Desmond, Adrian J. *Huxley: The Devil's Disciple*. London: M. Joseph; New York: Viking Penguin, 1994.

Di Gregorio, Mario A. *T. H. Huxley's Place in Natural Science*. New Haven, CT: Yale University Press, 1984.

Irvine, William. *Thomas Henry Huxley.* London: Longmans, Green, 1960.

Jensen, J. Vernon. *Thomas Henry Huxley: Communicating for Science*. Newark: University of Delaware Press; London: Associated University Presses, 1991.

Mitchell, P. Chalmers. *Thomas Henry Huxley.* New York: Putnam, 1900.

Paradis, James G. *T. H. Huxley: Man's Place in Nature*. Lincoln: University of Nebraska Press, 1978.

Peterson, Houston. *Huxley, Prophet of Science.* New York: Longmans, Green, 1932.

T. A. Goudge (1967)
Bibliography updated by Michael J. Farmer (2005)

HYLOZOISM

See *Panpsychism*

HYPATIA
(370/75–415 CE)

Hypatia was a philosopher, mathematician, and astronomer who, though female and pagan, achieved the honor of being named by the Christian Roman government to the position of philosopher at the museum of Alexandria. Students reading philosophy at the Alexandrian School would also study mathematics and astronomy as technical, applied disciplines of the more traditional studies of metaphysics and cosmology. Hypatia's father, Theon of Alexandria, was the museum's most famous mathematician-astronomer, and it is largely through Theon that we have a reliable source of Ptolemy's *Syntaxis Mathematica* (*Almagest*).

Hypatia likely assumed the directorship of the school of philosophy in about 400. The recently converted Christian, Synesius of Cyrene, later the bishop of Ptolemais, became her student in 393. From Synesius's works we surmise that Hypatia's early philosophical teachings concentrated on Plato's metaphysical works, especially the *Timeaus*. Her mathematical and astronomical writings can be understood primarily as applications of Neoplatonist metaphysical and cosmological theories to mathematical problems whose solution informed astronomical theories. These in turn were considered to illuminate Neoplatonist cosmological theories. Six of Hypatia's works have been tentatively identified. They include an edition of Diophantus's *Arithmetica* with new lemmas, a lost prototype based on Archimedes's *Sphere and Cylinder* surviving as John of Tynemouth's *De Curvis Superficibus*, a text on isoperimetric figures incorporated by a later author into *Introduction to the Almagest*, a commentary on Archimedes's *Dimension of the Circle*, and a commentary edition of Apollonius Pergaeus's *Conics* upon which later commentary editions were based. But her most important work appears to have been a revision of a work by her father Theon appearing in Book III of his *Commentary of Ptolemy's Syntaxis Mathematica*.

Hypatia was an eclectic philosopher with a Cynic's literary and personal style that may have had as much to do with her risky status as both woman and pagan as with her philosophical affiliation. Accounts of outrageous tactics to counter sexist male student behavior may be apocryphal (Lewis 1921, Toland 1720). Nevertheless, they provide insight into the personality of a defensive female professor in a brutally misogynist environment. A traditional middle Platonist, Hypatia was sympathetic to Porphyrian metaphysics and to Stoicism. She preferred Euclidean methodology to the Archimedean in formulat-

ing results of problems and as a pedagogical tool for teaching philosophical mathematics. In 415, she was savagely dismembered by a gang of monks. She appears to have been succeeded by Hierocles.

See also Alexandrian School; Cynics; Feminism and the History of Philosophy; Neoplatonism; Platonism and the Platonic Tradition; Porphyry; Sexism; Stoicism; Women in the History of Philosophy.

Bibliography

Bregman, Jay. *Synesius of Cyrene: Philosopher-Bishop*. Berkeley: University of California Press, 1982.

Evrard, Etiénne. "A Quel titre Hypatie enseigna-t-elle la philosophie?" *Revue des études grecques* 90 (1977): 69–74.

Heath, Thomas. *Diophantus of Alexandria*. Cambridge, U.K.: Cambridge University Press, 1910.

Heath, Thomas. *Greek Astronomy*. New York: E.P. Dutton, 1932.

Heath, Thomas. *A History of Greek Mathematics*. Oxford: Clarendon Press, 1921.

Index Bibliothecae Mediceae. Firenze: Tip dell' arte della stampa, 1882.

Knorr, Wilbur. "Observations on the Early History of the Conics." *Centaurus* 26 (1982): 1–24

Knorr, Wilbur. *Textual Studies in Ancient and Medieval Geometry*. Boston: Birkhäuser, 1989.

Knorr, Wilbur. "John of Tynemouth *alias* John of London: Emerging Portrait of a Singular Medieval Mathematician." *British Journal for the History of Sciences*, 1990.

Lampropoulou, S. "Hypatia philosophe Alexandrine." *Athenai, Bibliopholeion*. Athens, 1977.

Lewis, Thomas. *The History of Hypatia, a Most Impudent Schoolmistress of Alexandria*. London: T. Bickerton, 1721.

Ligier, Hermann. *De Hypatia philosophà et eclectismi alexandrini fine*. Dijon: Mercurri Curte, Chassel, 1879.

Meyer, Wolfgang Alexander. *Hypatia von Alexandria: Ein Beitrag zur Geschichte der Neuplautonismos*. Heidelberg: G. Weiss, 1886.

Montucla, Jean Etienne. *Histoire des mathématiques*. Paris: A. Blanchard, 1967. This important work was originally published in 1799.

Rome, Adolphe. *Commentaires de Pappus et de Théon d'Alexandrie sur l'almageste, tome III, Théon d'Alexandrie commentaire sur les livres 3 et 4 de l'almageste*. Studi e testi 106. Vatican City, 1943. This established Greek text with introduction and notes in French is substantially based on Halma, Nicolas. B. *Theon et Hypatie, Commentaires de Theon d'Alexandrie sur le troisième livre de l'Almageste de Ptolemée*. Paris: Merlin, 1820 and the introduction by Halma in his *Almageste de Ptolémée, Commentaires de Théon D'Alexandrie sur le premier livre de la composition mathématique de Ptolemée*. Paris: Merlin, 1821.

Rome, Adolphe. "Le Troisième livre des commentaires sur l'almageste par Théon et Hypatie." *Annales de la société scientifique de Bruxelles* 46 (1926), 1–14.

Rome, Adolphe. "Observations d'équinoxes et de solstices dans le chapitre 1 du livre 3 du commentaire sur l'almageste par Théon d'Alexandrie." *Annales de la société scientifique de Bruxelles* 57 (1937): 213–236 (première partie); 58 (1938): 6–26 (second partie).

Suidas. *Lexicon*. 5 vols. Translated by Ada Adler. Stuttgart: B.G. Teubner, 1967–1971. This work was reissued in 1989-1994.

Tannery, Paul. *Diophanti Alexandrini Opera Omnia*. 2 vols. Stuttgard: B.G. Teubner, 1974. This work was originally published in 1893–95.

Tannery, Paul. "L'Article de Suidas sur Hypatia." *Annales de la faculté des lettres de Bordeaux*. Bordeaux, 1880.

Tannery, Paul. *Mémoires scientifiques*. Paris: Editions Jacques Gabay, 1995. This work was originally published in Toulouse in 1912.

Toland, John. *Hypatia, or the History ….* In *Tetradymus, part III*. London: J. Brotherton and W. Meadows, 1720.

Usener, Hermann. "Fasti Theonis Alexandrini." In *Chronica Minora Sasc. IV. V, VI, VII*, edited by T. Mommsen. Berlin: Berolini, 1898.

Waithe, Mary Ellen. "Finding Bits and Pieces of Hypatia." In *Hypatia's Daughters: Fifteen Hundred Years of Women Philosophers*, edited by Linda Lopez McAlister, 4–15. Bloomington and Indianapolis: Indiana University Press, 1996.

Waithe, Mary Ellen. "Hypatia of Alexandria." In *A History of Women Philosophers*, Vol. 1, *Ancient Women Philosophers 600 BC–500 AD*. Dordrecht: Martinus Nijhoff/Kluwer Academic Press, 1987.

Mary Ellen Waithe (1996, 2005)

HYPPOLITE, JEAN

(1907–1968)

Born at Jonzac, France, Jean Hyppolite had an illustrious university career: professor at Université de Strasbourg in 1945; at the Sorbonne in Paris in 1949; director of the Ecole Normale Supérieure in 1954; and finally, the chair at the Collège de France in "Histoire des systèmes" from 1963 until his death. He belonged to the post–World War II generation of French philosophers that included Maurice Merleau-Ponty, Jean-Paul Sartre, and Jacques Lacan. However, Hyppolite's most enduring legacy is his students from the Sorbonne and the Ecole Normale Supérieure: Jacques Derrida, Gilles Deleuze, and Michel Foucault.

Hyppolite became famous as the French translator of Hegel's *The Phenomenology of Spirit* in 1941. He then produced a commentary, *Genesis and Structure of Hegel's Phenomenology of Spirit*, in 1947. In many essays, Hyppolite recounts the French reception of Hegel. The French reception had first been formed by Jean Wahl, but during the 1930s especially by the humanistic reading Kojève produced. Kojève's reading had oriented the philosophies of Sartre and the early Merleau-Ponty. Hyppolite, how-

ever, tried to show that Hegel goes beyond the human, an attempt that is obvious in his *Genesis and Structure*: Its chapter on the master-slave dialectic—the foundation for Kojève's humanistic reading—is its shortest, about three pages long.

Nevertheless, *Genesis and Structure* aims to be a comprehensive reading of Hegel's *Phenomenology of Spirit*. Indeed, it is a classic reading of a philosophical text, for it not only tracks all of the dialectical movement or procedures of the *Phenomenology*, but also shows how they are connected to the history of philosophy, how they are connected to other German idealists such as Friedrich Wilhelm Joseph von Schelling, and finally how the book itself grew out of Hegel's earlier reflections—in particular, his earlier Jena Logic. But Hyppolite's text is not only a classic reading, it is a masterful philosophical text in its own right. Unlike Kojève—for whom the leading question was how man take possession of himself as the purpose of all dialectical development—Hyppolite's leading question was: "How is the *Phenomenology* connected to the later *Logic*?" In other words, even if there is an ambiguity in Hegelianism between phenomenology (as the science of the appearances of the forms of consciousness, resulting in absolute knowledge) and ontology (as the science of all being), which one is the authentic mode of procedure in Hegelianism? Is Hegel's ontologic independent of all phenomenology? All of *Genesis and Structure* is directed at responding to this question. For Hyppolite, the intersection of knowledge and being is central to Hegelianism.

The centrality of this intersection becomes most evident in Hyppolite's 1952 *Logic and Existence*, a text that makes three basic claims. First, Hyppolite tries to show that Hegel's philosophy is a logic in the literal sense of the word, a *logos*: language. If we start from language, we can see that Hegel's philosophy attempts to reconstruct the genesis from sensible (experience) to sense (or essence). But second, again if we start from language, we can see that Hegel's thought "completes immanence," as Hyppolite says. This claim means that Hegel, like Nietzsche, is an anti-Platonist; there is no second world of ideas or essences behind the first sensible one; there is only sense. In this second claim, Hyppolite is returning to *Genesis and Structure*, in which he claimed that the most difficult idea in Hegel's thought was the difference between essence and appearance—that is, the *difference* within immanence itself. For Hyppolite, following Hegel, difference must be "pushed all the way up to contradiction" (Hyppolite 1997, p. 113). In other words, if we are to remain true to the thought of immanence, we must think totality. But to think totality, we must have opposites be internal to themselves. The infinite, for example, cannot be opposed externally to the finite; if it were, then the infinite would be finite because it would have the finite as its boundary. So, the infinite must include the finite inside of itself; it must be both finite and infinite and thus contradict itself. For Hyppolite, following Hegel, there can be difference within immanence only through self-contradiction. Is self-contradiction really difference? This question is explored further below.

The difference within immanence leads to the third and final claim made by Hyppolite in *Logic and Existence*. Hegel is not a humanist because sense (which has now replaced the old metaphysical concept of essence) is indeed different from man. Hegel therefore is trying to think not man but across man, and through this *antihumanism* Hyppolite's reading no longer shares any similarity with that of Kojève.

Logic and Existence sets up the philosophies of Derrida, Deleuze, and Foucault. Indeed, in his inaugural address to the Collège de France in 1970, Foucault, who was then assuming the chair vacated by Hyppolite's death, said that "*Logic and Existence* established all the problems that are ours" (Foucault 1972, p. 236). In other words, when Hyppolite discusses the problem of difference in Hegel, he is setting up the entire philosophy of difference that will arise in France in the 1960s. Yet as seen above, for Hyppolite difference must be pushed all the way up to contradiction. This occurs by indeterminate differences being converted into oppositions; each thing that is different must find *its* other, as Hyppolite says. Then, after having pushed all the indeterminate differences up into oppositions, one can see that each position makes sense only with or through its opposition. Nature, for instance, makes sense only through *its* opposite, which is culture. Thus each position includes its opposition in itself; each position is a self-contradiction. But, the philosophy of difference that arises in France during the 1960s consists of the attempt to push difference back down from self-contradiction to indeterminate difference. We can see this project already in Deleuze's 1954 review of *Logic and Existence*. But, the project is fulfilled in at least two different ways. On the one hand, one reconceives what looks to be a position and an opposition as two positivities or two positions; this is Deleuze (and Foucault). On the other hand, one reconceives what looks to be a position and an opposition as mutual contamination; this is Derrida. These two fulfillments are the legacy of Hyppolite's thought.

See also Derrida, Jacques; Deleuze, Gilles; Foucault, Michel; Hegel, Georg Wilhelm Friedrich; Hegelianism;

Lacan, Jacques; Merleau-Ponty, Maurice; Nietzsche, Freidrich; Ontology, History of; Phenomenology; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

PRIMARY WORKS

Figures de la pensée philosophique. 2 vols. Paris: Presses Universitaires de France, 1971. No English translation exists of these two volumes.

Genèse et structure de la Phénoménology de l'esprit de Hegel. 2 vols. Paris: Aubier, 1946–1947. Translated by Samiel Cherniak and John Heckman as *Genesis and Structure of Hegel's Phenomenology of Spirit* (Evanston, IL: Northwestern University Press, 1974).

Logique et existence. Paris: Presses Universitaires de France, 1952. Translated by Leonard Lawlor and Amit Sen as *Logic and Existence* (Albany, NY: SUNY Press, 1997).

SECONDARY WORKS

Baugh, Bruce, and French Hegel. *From Surrealism to Postmodernism*. London: Routledge, 2003.

Deleuze, Gilles. *Logique et existence*. Review of Jean Hyppolite. English translation found as an appendix to Jean Hyppolite, *Logic and Existence*, 191–196 (Albany, NY: SUNY Press, 1997).

Foucault, Michel. "The Discourse on Language." In *The Archeology of Knowledge*, translated by A. M. Sheridan Smith, 215–237. New York: Pantheon, 1972.

Roth, Michael S. *Knowing and History: Appropriations of Hegel in Twentieth Century France*. Ithaca, NY: Cornell University Press, 1983.

Leonard Lawlor (2005)

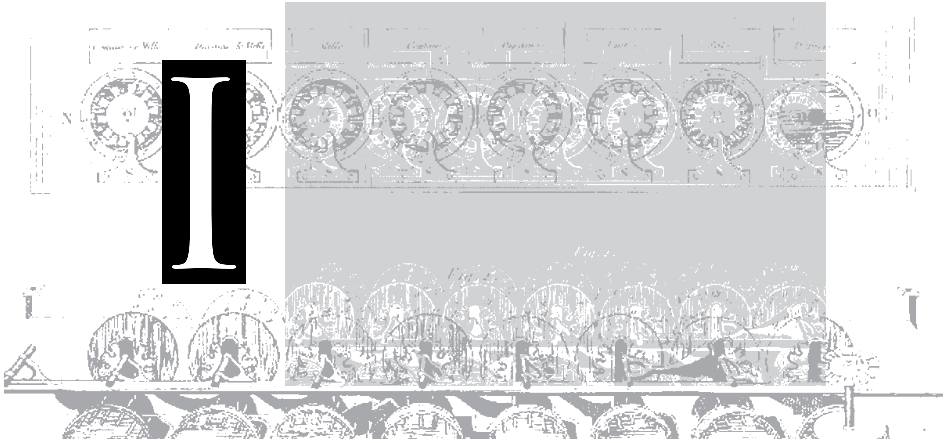

IAMBLICHUS

(c. 245–320 CE)

The sources available for our knowledge of Iamblichus's life are highly unsatisfactory, consisting as they do largely of a rather hagiographical and ill-informed *Life* by the sophist Eunapius, who was a pupil of Chrysanthius, who had been in turn a pupil of Iamblichus's pupil Aedesius, but enough evidence can be gathered to give a general view of his life-span and activities.

LIFE AND WORKS

The evidence points to a date of birth around 245 CE, in the town of Chalcis-ad-Belum (modern Qinnesrin) in Northern Syria. Iamblichus's family was prominent in the area, and the retention of an old Aramaic name (*yamliku*) in the family points to some relationship with the dynasts of Emesa in the previous centuries, one of whose names this was. This noble ancestry does seem to somewhat color Iamblichus's attitude to tradition—he likes to appeal on occasion for authority to "the most ancient of the priests" (*Iamblichus: De Anima*, §37).

As teachers, Eunapius provides (*VP* 457–8) us with two names: first, a certain Anatolius, described as second in command to the distinguished Neoplatonic philosopher Porphyry, the pupil of Plotinus, and then Porphyry himself. We are left quite uncertain as to where these contacts took place, but we must presume in Rome, at some time in the 270s or 280s, when Porphyry had reconstituted Plotinus's school. If that is so—and it is plain that Iamblichus knew Porphyry's work well, even though he was far from a faithful follower—then it seems probable that he left Porphyry's circle long before the latter's death, and returned to his native Syria (probably in the 290s) to set up his own school, not to his home town, but rather to the city of Apamea, already famous in philosophical circles as the home of the Neopythagorean Numenius. There he presided over a circle of pupils, including a local grandee, Sopater, who seems to have supported him materially, and as long as Licinius ruled in the East, the school flourished. After the triumph of Constantine, however, the writing had to be on the wall for such an overtly Hellenic and theurgically-inclined group, and on Iamblichus's death in the early 320s the school broke up; his senior pupil Aedesius moved to Pergamum, where the Iamblichean tradition was carried on quietly for another generation or so.

Iamblichus was a prolific author. Unfortunately, with the exception of *Reply to the Letter of Porphyry to Anebo* (popularly known, since the Renaissance, as *On the Mys-*

teries of the Egyptians), only his more elementary works survive intact. Chief among these is a sequence of nine, or possibly ten, works in which he presented a comprehensive introduction to Pythagorean philosophy. Of these, the first four are still in existence, beginning with a *Bios Pythagorikos*. This work is not simply a *life of Pythagoras*, but rather an account of the Pythagorean way of life, with a biography of Pythagoras woven into it. It was followed by *Exhortation to Philosophy*, the treatise *On the General Science of Mathematics*, and a commentary on the *Introduction to Arithmetic*, the second-century Neopythagorean Nicomachus of Gerasa. The doxographical portion of a treatise *On the Soul*, and extracts from a series of philosophical letters also survive in the *Anthology* of John of Stobi.

Other than those, however, we have considerable evidence of commentaries on works of both Plato and Aristotle, fragments of which survive (mainly) in the later commentaries of Proclus. Most notable among these are commentaries on the *Alcibiades*, *Phaedrus*, *Timaeus*, and *Parmenides* of Plato, and the *Categories* of Aristotle (this latter preserved extensively by Simplicius). He is also on record as having composed a commentary on the *Chaldaean Oracles*, and a Platonic Theology. The *Reply to the letter of Porphyry to Anebo* mentioned above is an odd production, in that it is a response to a polemical open letter by Porphyry attacking the practice and theory of theurgy, which Iamblichus, taking on the persona of a senior Egyptian priest, Abammon, elects to defend.

PHILOSOPHY

Iamblichus's system of philosophy is essentially an elaboration of Plotinus's Platonism, though strongly influenced by Neopythagorean writings and the *Chaldaean Oracles*. He accepts the triadic system of principles, the One, Intellect and Soul, but he introduces elaborations at every turn.

First of all, in an attempt to resolve the contradiction between a One which is utterly transcendent but which also constitutes the first principle of all creation, he postulates a totally ineffable first Principle above a more positive One, which itself presides over a dyad of Limit and Unlimitedness. These in turn generate a third principle, the One-Existent (*hen on*), which constitutes a link with the next hypostasis, that of Intellect (*Nous*), whose highest element it also is. Inhering in the One-Existent, we may also discern a multiplicity of *henads*, which serve as unitary prefigurations of the system of Forms which are the contents of Intellect.

Intellect, meanwhile, also suffers elaborate subdivision in Iamblichus's system, first into a triad of three *moments* or aspects, Being, Life, and Intellect proper, and then into a series of three triads (again, of Being, Life, Intellect) arising out of each of these. He thus becomes the ancestor of the elaborate system of the later Athenian School of Syrianus and Proclus. The impulse for such elaborations seem to stem from a consciousness of the complexity of the spiritual world, and of the many levels of divinity which inhabit it.

Soul, likewise, is distinguished into Pure, or Unparticipated Soul, and Participated Soul, which is in a way the sum-total of individual souls. Some individual souls, likewise, transcend any contact with body, while others are destined to be embodied, and even these descend into body on various different terms.

Iamblichus, in his treatise *On the Soul*, sought to differentiate himself from his predecessors Plotinus and Porphyry, on the issue of the relation of the soul with what is above it, postulating a less direct contact with Intellect and the One, and a corresponding need of theurgy, or *sacramental* ritual, to secure personal salvation. He may thus be reasonably accused of making Platonism much more of a religion, a characteristic which endeared him in particular to the Emperor Julian, a generation after his death.

See also Neoplatonism; Plotinus; Porphyry; Proclus; Simplicius.

Bibliography

Blumenthal, H. J., and E. G. Clark, eds. *The Divine Iamblichus, Philosopher and Man of Gods*. Proceedings of First International Conference on Iamblichus, Sept. 1990. London: Bristol Classical Press (Duckworth), 1993.

Blumenthal, H. J., and John F. Finamore, eds. *Iamblichus: The Philosopher*. *Syllecta Classica*, Vol. 8. Iowa City: University of Iowa, 1997.

Clarke, Emma C. *Iamblichus's De Mysteriis: A Manifesto of the Miraculous*. Aldershot: Ashgate, 2001.

Clarke, Emma C., John M. Dillon, and Jackson P. Hershbell, eds., trans., introductions, and notes. *Iamblichus: On the Mysteries*. Atlanta: Society for Biblical Literature, 2003.

Dillon, John. "Iamblichus of Chalcis." In *Aufstieg und Niedergang der römischen Welt* II 36: 2, edited by W. Haase and H. Temporini, 862–909. Berlin/New York: De Gruyter, 1987.

Dillon, John. *In Platonis Dialogos Commentariorum Fragmenta*. Leiden: Brill, 1973.

Dillon, John, and Jackson Hershbell, text, trans., and notes. *Iamblichus: On the Pythagorean Way of Life*. Atlanta: Scholars Press, 1991.

Finamore, John, and John Dillon. *Iamblichus: De Anima*. Leiden: Brill, 2002.

Shaw, Gregory. *Theurgy and the Soul: The Neoplatonism of Iamblichus*. University Park: Pennsylvania State University Press, 1995.

Wallis, Richard. *Neoplatonism*. London: Duckworth, 1972.

John Dillon (2005)

IBN AL-ʿARABĪ

(1165–1240)

Muhyī al-Dīn Muhammad ibn ʿAlī, known as Ibn al-ʿArabī (or Ibn ʿArabī, without the definite article), was arguably the most influential philosopher of the second half of Islamic history. Born in Murcia in Muslim Spain in the year 1165, he left the west permanently in 1200, settled in Damascus in 1223, and died there in 1240. He is not normally classified as a philosopher (*faylasūf*) because he made no attempt to fall into line with the schools of thought that adopted Greek methodologies. Nonetheless, "love of wisdom" was central to his project, and he praised the high aspiration of "the divine Plato" and others who engaged in the philosophical quest. With extraordinary faithfulness to the sources of the Islamic tradition and unprecedented originality, he offered diverse interpretations of the fundamental issues of philosophical and theological thought. He was enormously prolific, yet maintained a consistently high level of discourse without repeating himself. He has typically been classified as a "Sufi" or a "mystic," but this simply means that he supplemented rational investigation with suprarational intuition, not that he avoided philosophical issues. His pervasive influence was not eclipsed until the collapse of Islamic institutions under the pressure of colonialism and modernity. Current renewed interest in his legacy throughout the Muslim world stems largely from the realization that the Enlightenment project has reached a dead end.

Henry Corbin (1969) suggests that Ibn al-ʿArabī's meeting during his teenage years with Averroes (d. 1198), the last of the great peripatetics, symbolizes the parting of the ways between Islamic and Christian civilizations. According to Ibn al-ʿArabī's account, his father had told Averroes that his son had been opened up to the invisible realms, and Averroes requested a meeting. He asked the boy if philosophical theory (*nazar*) reached the same conclusions as divine unveiling (*kashf*). Ibn al-ʿArabī replied, "Yes and no: Between the yes and the no, spirits fly from their matter and heads from their bodies" (*al-Futûhât al-makkiyya*, p. 154). In Corbin's reading, Ibn al-ʿArabī and subsequent Muslim philosophers preserved the creative tension between the "yes," or the affirmation of the legitimacy of rational thought, and the "no," or the declaration of its inadequacy in face of the divine. In contrast, European thought, partly under the influence of Latin Averroism, lost the balance between reason and intuition and fell into deep dichotomies between philosophy and theology, science and religion, history and symbol, mind and heart.

Corbin considered Ibn al-ʿArabī's main theoretical contribution to lie in his stress on *khayâl*, "imagination" or "image," specifically the human faculty of creative imagination and the cosmic *mundus imaginalis*, the "imaginal"—not imaginary—world located between the sensible and intelligible realms. Though this world had been implicitly affirmed by Avicenna's cosmology, it was denied by Averroes. When Ibn al-ʿArabī reformulated Avicenna's thought in terms of imagination, the eventual result was a synthetic rather than analytic philosophical vision that stressed the essential unity of human beings and the cosmos.

Corbin, however, neglects a third and deeper meaning of the word *khayâl*. It also denotes the cosmos as a whole, the realm of contingency and becoming. While discussing the cosmos as image, Ibn al-ʿArabī offers unprecedented analyses of *wujûd*, "being" or "existence," the basic topic of the philosophers. He follows the Avicennan picture by classifying *wujûd* into two basic sorts—necessary (*wâjib*) and contingent (or possible, *mumkin*)—but he reminds us that the Arabic word *wujûd* also signifies consciousness, awareness, finding, ecstasy, and bliss; and he reformulates the whole discussion in terms of the Qurʾanic doctrine of divine names and attributes. The Necessary *Wujûd* is not only that which is and cannot not be, but also that which knows and cannot not know, lives and cannot not live, loves and cannot not love. To say that the cosmos is the realm of contingent *wujûd* means that it stands halfway between being and nothingness, awareness and unconsciousness, life and death, mercy and wrath. In the same way, the *mundus imaginalis* is the intermediate realm of cosmic becoming, situated between the luminosity of the angelic realm and the darkness of corporeality. As for the human self (*nafs*), it is the imagination of the microcosm, hanging between heaven and earth, spirit and body, intelligence and ignorance, virtue and vice.

In later times Ibn al-ʿArabī came to be known as the great expositor of *wahdat al-wujûd* or "the Oneness of Being," even though he and his immediate followers did not employ the expression (at least not in a technical sense). Ontology was unquestionably central to his proj-

ect, but so were epistemology, hermeneutics, theology, cosmology, spiritual psychology, ethics, and jurisprudence. He addressed all of the basic fields of human understanding, yet he was not attempting to provide an overarching system. Rather, he was adumbrating the major categories of human cognitive participation in the infinite and never-repeating disclosures of the Necessary *Wujûd*, disclosures that are none other than the imaginal realm known as the cosmos. He employs the mythic language of the Qurʾan to provide a broad framework for the stations of wisdom that designate the realms of human possibility, and he assures us that each of the 124,000 prophets sent by God embodies a distinct archetype of human perfection. His most famous book, *Fusûs al-hikam*, is arranged in terms of twenty-seven prophetic Logoi, each of which incarnates a specific divine attribute.

To provide some sense of the scope of his work, a few of the many themes that he discussed with unparalleled thoroughness and which then reverberated down through Islamic intellectual history will be outlined:

WISDOM AS REALIZATION

According to the Qurʾan, one of God's names is *haqq*, a word that designates reality, truth, appropriateness, rightness, right (as in "human rights"), and justice, along with the corresponding adjectives. From al-Kindī onwards, Muslim philosophers often defined their discipline in terms of this word. As al-Kindī put it at the beginning of his *Metaphysics*, "The philosophers' purpose in their knowledge is to hit upon the *haqq* and, in their practice, to practice according to the *haqq*" (*Rasâʾil al-Kindī al-falsafiyya*, p. 25) They took *haqq* in its purest sense as a designation for the Necessary *Wujûd* in itself, but they also recognized that it denotes the realm of contingency known as the cosmos, the truth that is to be grasped (right understanding), and the embodiment of truth in correct activity (ethics and virtue). To say that the cosmos is the realm of *haqq* means that human beings, like everything else, manifest truth, reality, and right in their essential nature.

In contrast to other things, however, human beings partake of enough freedom to affect the degree to which they understand and embody *haqq*, and it is this that necessitates praxis. The quest for wisdom is then called *tahqīq* (from the same root as *haqq*), that is, "realization"—literally "actualizing *haqq*." *Tahqīq* is contrasted with *taqlīd*, "imitation" of the beliefs and opinions of others. To be a sage demands far more than studying philosophy and memorizing the words of Aristotle and Avicenna. The real goal is to see *haqq* for oneself (*theôria* as vision) and to act in keeping with one's own impartial seeing. To the extent that people remain imitators, they are held back from their human substance. To the extent that they achieve realization, they participate in *haqq*, that is, the reality and consciousness of the Necessary *Wujûd*. The quest for wisdom remains intensely personal, for the cumulative theories of philosophers and scientists, not to speak of the conventional, imitative knowledge of society, are as nothing compared to the knowledge of self that only the self can achieve for itself.

THE COMPLEMENTARITY OF LOGIC AND MYSTICISM

Ibn al-ʿArabī's account of his encounter with Averroes highlights two terms, "theory" and "unveiling." The former designates the rational and discursive knowledge achieved by philosophers and theologians, the latter the suprarational intuition granted to mystics and visionaries. A good deal of Ibn al-ʿArabī's writing deals with the inadequacies of exclusive reliance on one or the other of these two modes of understanding. Theory he calls "the eye of reason," and unveiling "the eye of imagination." Both are located in the "heart" (*qalb*), which, in Islam as in China and other traditional contexts, is the seat of consciousness and selfhood. The quest for realization is the attempt to actualize the vision of both eyes and to achieve the harmonious marriage of logos and mythos, philosophy and poetry, science and art.

THE HARMONY OF THE ONE AND THE MANY

The eye of reason has the capacity to discern the individual reality—the essence or quiddity—of each thing. In contrast, the eye of imagination is able to perceive the actual presence of the Necessary *Wujûd* in all that exists. Reason acknowledges difference and recognizes its *haqq*—its reality, appropriateness, and rightness. It grasps that the Necessary *Wujûd* is utterly other than the cosmos and the existents (*mawjûdât*). In contrast, imagination perceives sameness and denies otherness. It finds the face of the Necessary *Wujûd* in every contingency. Theologically, this means that reason perceives God as distant and asserts his incomparability or transcendence (*tanzîh*); imagination sees God as present and asserts his similarity or immanence (*tashbîh*). Only the heart that sees simultaneously with both eyes can understand God, the universe, and the self as they truly are. The sage does not fall into the traps laid down by the principle of noncontradiction. Instead he grasps the exact manner in which all

things pertain to the imaginal realm of both/and, or neither/nor. According to Ibn al-ʿArabī's most succinct formulation of the actual state of affairs, the sage sees that everything in the cosmos is "He/not He" (*huwa lâ huwa*)—God/not God, being/nothingness, necessity/impossibility, consciousness/unawareness.

LANGUAGE AS THE DETERMINANT OF REALITY

We perceive only language, for at root nothing is accessible to us but the Logos, the self-expression of the Necessary *Wujûd* (called the "Active Intellect" by the peripatetic philosophers). According to the Qurʾan, "When God desires a thing, He says to it 'Be!', and it comes to be." In Ibn al-ʿArabī's terms, what "comes to be" is the infinite words of God articulated in "the Breath of the All-Merciful" (*nafas al-rahmân*). God's mercy, his all-embracing bounty and kindness, is the Necessary *Wujûd* itself, which demands the existence of every possibility. Within the Breath, God voices all things as letters, words, and sentences, arranging them in three grand books: the cosmos, the human self, and revelation. All language, whether divine, cosmic, or human, pertains to the imaginal realm, for words are neither the speaker nor other than the speaker, neither the spoken nor other than the spoken. The key to deciphering the message inscribed in cosmos and self lies in prophetic revelation, which explicates the Logos in the language most accessible to human understanding.

THE CORRELATION BETWEEN MACROCOSM AND MICROCOSM

The Breath of the All-Merciful deploys itself as a hierarchy of being and consciousness on three basic levels: intelligence or spirit, imagination or soul, and corporeality. Each of these ontological levels is inhabited by appropriate entities (e.g., angels, jinn or psychic beings, inanimate objects). The sum total of the three levels along with the infinite proliferation of inhabitants is the cosmos. The cosmos is then called the "macrocosm" when contrasted with the human being as microcosm. What differentiates the human microcosm from all other creatures is its all-comprehensive image of the Necessary *Wujûd*. The macrocosm embraces all contingent beings in their distinctiveness, all the individual words uttered by the All-Merciful Breath. The microcosm combines all the characteristics of macrocosmic reality in a unified and focused whole that opens up inwardly in the direction of the undifferentiated Logos.

ANTHROPOCOSMIC TELEOLOGY

The cosmos can be looked upon as a static hierarchy, but Ibn al-ʿArabī more typically describes it in terms of the dynamism of its unfolding. We and all things come into existence from the One in a quasi-neoplatonic manner, but our existential concerns are determined by the path we follow in retracing our steps to the Origin. Given that consciousness and self-awareness are centralized in the human microcosm, the way back to the One—which is the Necessary *Wujûd*, the unity of Being, Consciousness, and Bliss—goes by way of the full realization of the human self as the immanent Logos. The purpose of human life is the recovery of the original unity of being and intelligence by way of self-understanding. But self-understanding cannot be achieved without understanding the cosmos and revelation, and none of these three grand books can be deciphered unless they are read with both eyes of the heart. The cosmic subjectivity of the human state is so central that it provides the raison d'être for the existence of the world. Human beings—or rather, those whom Ibn al-ʿArabī calls "perfect human beings" (*al-insân al-kâmil*)—are the final cause of the contingent realm. In them alone are realized God's words as related by Muhammad in the famous saying, "I was a Hidden Treasure and I desired to be known, so I created the creatures that I might be known."

See also al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotle; Averroes; Averroism; Corbin, Henry; Enlightenment, Islamic; Imagination; Islamic Philosophy; Macrocosm and Microcosm; Mysticism, History of; Sufism.

Bibliography

Addas, Claude. *Quest for the Red Sulphur: The Life of Ibn ʿArabī*. Cambridge, U.K.: Islamic Texts Society, 1993.

al-Kindī. *Rasâʾil al-Kindî al-falsafiyya*, edited by Muhammad ʿAbd al-Hâdî Abû Rîda. Cairo: Dâr al-Fikr al-ʿArabî, 1978.

Almond, Ian. *Sufism and Deconstruction: A Comparative Study of Derrida and Ibn ʿArabi*. London: Routledge, 2004.

Arabī, Ibn al-ʿ. *al-Futûhât al-makkiyya*. Vol 1., Cairo: 1911.

Chittick, William C. *Ibn ʿArabi: Heir to the Prophets*. Oxford: Oneworld, 2005.

Chittick, William C. *The Sufi Path of Knowledge: Ibn al-ʿArabī's Metaphysics of Imagination*. Albany: State University of New York Press, 1989.

Coates, Peter. *Ibn ʿArabi and Modern Thought: The History of Taking Metaphysics Seriously*. Oxford: Anqa, 2002.

Chodkiewicz, Michel. *An Ocean without Shore: Ibn ʿArabī, the Book, and the Law*. Albany: State University of New York Press, 1993.

Corbin, Henry. *Creative Imagination in the Sûfism of Ibn ʿArabī* (re-issued as *Alone with the Alone*). Princeton, NJ: Princeton University Press, 1969.

Dagli, Caner K. *Ibn al-Arabi: The Ringstones of Wisdom (Fusus al-hikam)*. Chicago: Kazi Publications, 2004.

Izutsu, Toshihiko. *Sufism and Taoism: A Comparative Study of Key Philosophical Concepts*. Los Angeles: University of California Press, 1983.

Morris, James. *The Reflective Heart: Discovering Spiritual Intelligence in Ibn ʿArabī's Meccan Illuminations*. Louisville: Fons Vitae, forthcoming.

Sells, Michael A. *Mystical Languages of Unsaying*. Chicago: University of Chicago Press, 1994.

William C. Chittick (2005)

IBN BĀJJA

(d. 533 AH/1138 CE)

Abū-Bakr Muḥammad ibn Yaḥyā ibn al-Sāyigh ibn Bājja, the Islamic philosopher, was known to the medieval Scholastics as Avempace. He was born in Saragossa at the end of the fifth century AH, eleventh century CE, and died in Fez, Morocco, in 533 AH/1138 CE. During his brief life he endured the tribulations occasioned by the Christian "reconquest" of Andalusia. It is known that he wrote several commentaries on Aristotle's treatises and that he was very learned in medicine, mathematics, and astronomy. He was involved in the quarrel initiated by the Peripatetics, during which al-Bītrogī, whom the Scholastics called Alpetragius, distinguished himself. Ibn Bājja opposed his own hypotheses to Ptolemy's system.

Ibn Bājja's philosophical works have remained incomplete, notably the treatise that gained him his reputation, *Tadbīr al-motawaḥḥid* (The rule of the solitary). For a considerable length of time this treatise was known only through a detailed analysis of it in Hebrew by Moses of Narbonne (fourteenth century) in his commentary on the *Ḥayy ibn Yaqẓān* of Ibn Ṭufayl, the pupil of Ibn Bājja. Salomon Munk based his account of Ibn Bājja on this analysis. The Arabic original (now in the Bodleian Library at Oxford) was rediscovered by Miguel Asin Palacios.

The work's central theme is that of an *itinerarium* leading the man-spirit to unite itself with the Active Intellect (*ʿAql faʿʿal, Intellegentia agens*). He who speaks of a "rule" or "discipline" assumes a mode of life regulated by actions demanding reflection, and this can be found only in the solitary man. This is why the solitary man's discipline should be the model for a member of the perfect City and the ideal State. The ideal State, it must be noted, is not the result of a priori conceptions, nor can it come into being by a political coup d'état; much more than a mere "social" reform, it is the fruit of a reform of customs that seeks to realize the fullness of human existence in each individual. For the time being, the solitary individuals live in imperfect states, with neither judge nor doctor except God. Their task is to become members of the perfect City. In order to found the regime of these individuals it is necessary at first to analyze and classify human actions, using the forms that they strive to fulfill as the point of departure.

For this reason the treatise is presented essentially as a "theory of spiritual forms," a sketch of the phenomenology of the spirit. The spirit progressively evolves from forms engaged in matter to forms that have been abstracted from it. Having then become intelligible in act, these forms thereby attain the level of intellect in act, reaching the level of pure spiritual forms, those forms that, inasmuch as they exist for the Active Intellect, have not had to pass from power to act.

Ibn Bājja imposed upon Islamic philosophy in Spain a completely different orientation than did Mohammad al-Ghazali. The motives of the solitary individual, of the stranger, and of the *allogène,* however, merge with the motives typical of the mystical gnosis in Islam. The same type of spiritual man is realized in these individuals, although their perception of the common goal differs and thereby the choice that determines their course. One of these courses in Spain was that of Ibn Masarra, which was continued by Ibn al-Arabi. Another was that of Ibn Bājja, later continued by Averroes.

See also al-Ghazālī, Muhammad; Averroes; Ibn al-ʿArabī; Ibn Ṭufayl; Islamic Philosophy; Logic, History of: Logic in the Islamic World; Peripatetics.

Bibliography

Ibn Bājja's *Tadbīr al-motawaḥḥid* has been edited and translated into Spanish by Miguel Asin Palacios as *Avempace, El régimen del solitario* (Madrid and Granada, 1946); the first section has been translated into English by D. M. Dunlop in *Journal of the Royal Asiatic Society* (1945): 61–81.

For discussions of Ibn Bājja's philosophy, see Henry Corbin, *Histoire de la philosophie islamique,* Vol. I (paperback, Paris: Gallimard, 1964), pp. 317–325; S. H. Masumi, *Ibn Bajjah's ʿIlm al-Nafs* (Karachi: Pakistan Historical Society, 1961); and Salomon Munk, *Mélanges de philosophie juive et arabe* (Paris: A. Franck, 1859; new ed., Paris: J. Vrin, 1955).

Henry Corbin (1967)

IBN BĀJJA [ADDENDUM]

One of the unusual aspects of Ibn Bājja's political philosophy is the doctrine that in imperfect societies the philosopher has the status of weeds or *nawabit*. In a society governed by reason, the representatives of reason—the philosophers—find an important place. They are important people in the state because the state requires them to help it pursue the most rational course. But where the state is not governed by reason and instead by some less perfect rationale, philosophers will find themselves out of favor; they will be regarded as useless—like weeds—and possibly even dangerous to the state itself. The only happiness that philsophers will be able to establish for themselves is a private happiness because they will not be able to use publicly their knowledge of how the state ought to be run. Ibn Bājja's remarks on this topic are poignant, bringing out nicely the alienation experienced by intellectuals in a culture where their views are disregarded.

See also Islamic Philosophy.

Bibliography

Bājja, Ibn. *Aristotle's Physics and its Perception in the Arabic World with an Edition of the Unpublished Parts of Ibn Bajja's Commentary on the Physics*. Leiden, Netherlands: Brill, 1994.

Goodman, Lenn. "Ibn Bajjah." In *History of Islamic Philosophy*, edited by Seyyed Nasr and Oliver Leaman, 294–312 London: Routledge, 1996. Contains a useful bibliography.

Oliver Leaman (2005)

IBN GABIROL, SOLOMON BEN JUDAH

(c. 1021–1058 or 1070)

Solomon ben Judah ibn Gabirol was first mentioned by Sāʿd the Qadi of Toledo (c. 1029–1070), who claimed that ibn Gabirol lived in Saragossa, was a keen student of philosophy, especially logic, and died sometime around 1058 CE, after he had passed the age of thirty. The Andalusian Jewish poet Moses ibn Ezra (c. 1060–1139) claimed that ibn Gabirol was born in Malaga and reared in Saragossa and spent a short but fruitful life in the service of philosophy and poetry. The Jewish philosopher Abraham ibn Daud (c. 1110–1180) said that ibn Gabirol died in 1070, but 1058 is more generally accepted.

The tone of some of Ibn Gabirol's secular songs, gloomy and bitter, is sometimes considered an indication of his unhappy lot—orphaned at an early age, poor, and ostracized by many of his contemporaries because of his irascible disposition and unorthodox philosophy. He did find some favor with Yequtiel ben Ishaq ibn Hasan, a veritable Maecenas, at the court in Saragossa, and later with his patron, Samuel ibn Nagrella, at the court of Zirid in Granada. Most, if not all, of this patronage seems to have resulted from his reputation as the greatest Jewish poet of his time in the West.

There are some 400 extant secular and religious poems attributed to ibn Gabirol. One, *The Kingly Crown,* has become a part of the Sephardic Jewish liturgy for the Day of Atonement. Its rhythmical, rhymed simplicity gives it a distinct biblical flavor. In this poem of forty stanzas, ibn Gabirol celebrated the divine attributes, the last of which is Will, so prominent in his philosophical work *The Fountain of Life,* and the wonders of creation, reminiscent of an Aristotelian-Ptolemaic worldview. He concluded with an Augustinian self-analysis, marked by confession, penitence, and supplication.

In addition to his work in poetry, an anthology of ethical and sapiential sayings, *Choice of Pearls,* is attributed to him, but its authenticity is doubted by some.

Ibn Gabirol's *The Improvement of the Moral Qualities* exists in one known Arabic text and four Hebrew versions, as well as in translations into other languages. The ethical aspect of Ibn Gabirol's philosophy is interesting because it appears to be an early, if not the first, attempt to systematize the basic principles of medieval Jewish ethics independently of religious dogma, ritual, or belief. The impulses of the human soul and how they can be trained to virtue or permitted to fall into vice are explained in relation to the five external senses, which are in turn explained by the four-element, or simple-body, theory of Aristotle. Stephen S. Wise claimed that "Gabirol's object is to establish a system of purely physio-psychological ethics." Certainly it is true that his interest was mainly in the animal rather than the rational soul. He emphasized the virtuous order that can be achieved in the external senses under the direction of the rational soul. In his treatment of the virtues and vices, Ibn Gabirol did refer to biblical writings, but in a superficial and summary way, as a support of his own allegorical-poetic viewpoint.

It was not until nearly the end of the first half of the nineteenth century that Salomon Munk showed conclusively that ibn Gabirol, the great Jewish poet, was the same man as Avicebron, the recognized author of *The Fountain of Life.* An examination of the abstracts translated into Hebrew from the original Arabic by Shem Tob

Falaquera in the thirteenth century and attributed to ibn Gabirol showed substantial agreement with related passages of the *Fons Vitae* attributed to Avicebron by the twelfth-century translators John of Spain (Ibn Daud, Avendehut) and Dominic Gundissalin. In the text of *The Fountain of Life* are found references to two other works by ibn Gabirol, "The Treatise on *Esse*," Book 5.8, and the book of the Will, which is titled "*Origo Largitatis et Causa Essendi*," Book 5.40. Unfortunately, these works cannot be found or identified. They may constitute, with *The Fountain,* the three parts of Wisdom: knowledge of matter and form (*The Fountain*), knowledge of Will (*The Origo*), and knowledge of the First Essence (*De Esse*). *The Fountain* is like the ethical work in its purely rational approach but differs in its complete lack of references to the Bible, the Talmud, or the Midrash. It is a treatise in the strict philosophical area of Neoplatonism as related to an eleventh-century Jewish mind. In it we find a Neoplatonic universe dependent on the Will of the First Author, supreme and holy.

The Fountain of Life, though composed in a dialogue form involving master and pupil, has none of the beauty and charm of the dramatic dialogues of Plato, the only other person mentioned by name in the work. The pupil seems to be a fictitious straight man, asking the proper questions at the proper time and giving a verbal nod of the head when appropriate. The opening section tells us that the discussion concerns the first part of Wisdom, the science of universal matter and universal form. Because of the nature of the topics involved, the work falls neatly into five parts:

(1) What we must presuppose in order to assign universal matter and form and predicate them of composite substances.

(2) The substance upholding the corporeity of the world.

(3) The acceptance of simple substances, such as the separated intelligences (i.e., angels).

(4) The science of understanding matter and form in simple substances.

(5) Universal matter and universal form in themselves.

The general method followed in this dialectical investigation is a search for the nature and existence of certain properties, which when found reveal the existence of the being that has these properties. In things we find there is something that "exists in itself," "is of one nature," is the "vehicle of diversity," and "gives everything its essence and name." These are the properties of universal matter. If one abstracts every sensible and intelligible form from things, the remainder is the common denominator called universal matter. Universal form is found "to subsist in another," "to perfect the essence of that in which it is," and "to give it being." By inspecting universal and particular sensible things, one finds four grades of matter and form: artificial-particular matter and its appropriate form, natural-particular matter (the matter of art products) and its form, natural-universal matter and its form, and celestial matter (the matter of the simple intelligences) and its form. Hence, there are common denominators for both matter and form: universal matter and universal form.

Every reality, except the First Essence, when viewed with its form is called a substance; when one conceives of something as receptive of form, then it is called matter, or hyle. Sensible forms require an extended substrate or body. The corporeal body is formed out of matter (which is itself incorporeal) and the corporeity-form, quantity. The first and simplest form and the highest matter are those that when united constitute the Intelligence. The Intelligence is the highest existence next to the First Essence. Below this are the rest of the hylomorphically composed souls—rational, sensitive, and vegetative—and then nature, the foundation of all inorganic things. Nature serves as the matter for the corporeity-form, quantity; the resulting substance is the matter of sensible qualities, like color.

One might say that Ibn Gabirol's universal hylomorphism represents an intermediate between the universal formlessness of Augustine and the later Franciscan variations. There are many differences from, as well as similarities to, scholastic thought, but the influence of Jewish religious ideas provides a basis for creation in his Neoplatonic universe. It seems that the Neoplatonic element in his thinking led ibn Gabirol to consider the origin of all things as a necessary emanation from the First Author. But the Jewish element may have rebelled against this, as a necessary emanation would be in conflict with the absolute transcendence of God. The solution results in an intuitive view of the relation of all things to the First Author. This relation is necessary because matter is an expression of the essence of God, who is himself necessary. However, the dynamism of the Will of God leads to the need for a variety of forms that are initiated by God. Hence, the relation is voluntary and therefore free. In *The Fountain of Life* we have at times a strange mixture of Jewish religious ideas, Arabian Aristotelianism, and Alexandrine Neoplatonism, though we cannot be

absolutely sure of the source of the ingredients because of the absence of definite historical information.

Solomon ibn Gabirol's direct influence in philosophy seems to have been confined to certain Franciscans of the Augustinian tradition in the thirteenth and fourteenth centuries. They thought that ibn Gabirol's universal hylomorphism supplied them with a suitable philosophical way of expressing the difference between creatures and God. The universal principle of limitation—namely, matter—becomes spiritual matter in all other creatures.

See also Aristotle; Augustine, St.; Jewish Philosophy; Matter; Neoplatonism; Plato.

Bibliography

WORKS BY IBN GABIROL

Munk, S. *Mélanges de philosophie juive et arabe.* Paris: Gamber, 1859. Includes French translation of *Fons Vitae.*

"'Fons Vitae' ex Arabico in Latinum Translatus ab Johanne Hispano et Dominico Gundissalino. Ex Codicibus Parisinis, Amploniano, Columbino," edited by Clemens Bäumker, in *Beiträge zur Geschichte der Philosophie des Mittelalters,* Band I, Hefte 2–4. Münster, 1892–1895.

Choice of Pearls. Translated by A. Cohen. Library of Jewish Classics, IV. New York: Bloch, 1925.

The Improvement of the Moral Qualities. Translated by Stephen S. Wise. Columbia University Oriental Studies Publications, 1. New York: Columbia University Press, 1902.

The Fountain of Life. Translated by H. E. Wedeck, introduction by T. E. James. New York: Philosophical Library, 1962.

Tikkun Middot ha-Nefesh, edited and translated by S. Wise as *The Improvement of the Moral Qualities.* New York: Columbia University Press, 1966.

Keter Malkhut (*The Kingly Crown*). Translated by I. Zangwill in *Selected Religious Poems of Solomon ibn Gabirol,* edited by I. Davidson. New York: Arno Press, 1973.

Duties of the Heart, edited and translated by M. Hyamson. Jerusalem: Feldheim, 1986.

WORKS ON IBN GABIROL

Brunner, Fernand. *La source de vie.* Book III. Paris, 1950.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages.* New York: Random House, 1955.

Goodman, L. E., ed. *Neoplatonism and Jewish Thought.* Albany: State University of New York Press, 1992.

Guttmann, Jacob. *Die Philosophie des Salomon ibn Gabirol.* Göttingen: Vandenhoeck and Ruprecht, 1889.

Kaufmann, David. *Studien über Salomon ibn Gabirol.* Budapest, 1899.

Pessin, S. "Jewish Neoplatonism: Being above Being and Divine Emanation in Solomon ibn Gabirol and Isaac Israeli." In *Cambridge Companion to Medieval Jewish Philosophy,* edited by D. Frank and O. Leaman, 91–110. Cambridge, U.K.: Cambridge University Press, 2003.

Rudavsky, T. "Conflicting Motifs: Ibn Gabirol on Matter and Evil," *New Scholasticism* 52 (1978): 54–71.

Samuelson, N. (1997) "Medieval Jewish Aristotelianism: An Introduction" In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman. London: Routledge, 1997.

Sirat, C. *A History of Jewish Philosophy in the Middle Ages,* 68–81. Cambridge, U.K.: Cambridge University Press, 1985.

Ueberweg, Friedrich. *Grundriss der Geschichte der Philosophie.* Vol. II, 11th ed., edited by B. Geyer. Berlin, 1928.

Theodore E. James (1967)
Bibliography updated by Oliver Leaman (2005)

IBN KHALDŪN

(1332–1406)

Ab-Ar-Rahman ibn Khaldūn, the Muslim statesman, historian, philosopher of history, sociologist, and political thinker of the fourteenth century, is probably the greatest creative genius produced by Muslim civilization. To Arnold Toynbee, Ibn Khaldūn's philosophy of history "is undoubtedly the greatest work of its kind that has ever yet been created by any mind in any time or place."

Ibn Khaldūn was born in Tunis into a family of southern Arabian origin that had immigrated to Andalusia in the eighth century. With the decline of Muslim rule in Spain the family immigrated to northwest Africa, establishing itself first in Morocco and then in Tunisia. Muslim emigrants from Spain constituted an aristocracy in the Maghreb, and the Khaldūn family won fame in scholarship and statesmanship.

Ibn Khaldūn surpassed the achievements of all the members of his family. Brought up in the traditional religious sciences and the philosophical-rational sciences that formed the two major streams of Islamic culture, he studied the Qurʾan, Arabic, traditions, jurisprudence, logic, and philosophy under several of the best scholars of his time and studied, taught, and occupied high positions in Tunis, Algeria, Morocco, southern Spain, and Egypt.

Medieval Muslim rulers were eager to enlist scholars either for government service or for the prestige that goes with their presence in the court. Ibn Khaldūn enjoyed all the privileges of princely positions and suffered the odds of their fluctuations in medieval courts. He shared in the political maneuvers and conspiracies that accompanied the rise and fall of different rulers, and in trying periods, when he was in prison or was forced into exile, he devoted himself to the study of power and meditated on its historic laws and social dynamics.

Ibn Khaldūn's greatest work, *Al-Muqaddimah* (The prolegomena), was the first of seven volumes of his universal history of the Arabs and Berbers, *Kitab al-ʿIbar.*

Although the last two volumes are of special value to historians as the best source for the history of northwest Africa, especially for the history of the Berbers, the introduction that outlines ibn Khaldūn's philosophy of history overshadowed the narrative. The philosophic originality of this introduction was so great that ibn Khaldūn became known as the author of *Al-Muqaddimah.*

Prior to ibn Khaldūn, Muslim philosophers had concerned themselves with the reconciliation of Qurʾanic truth and rational truth, but this had led to an assimilation of Greek rationalism by Muslim theology rather than to the emergence of Muslim rationalism. The concern with religion and philosophy penetrated all Muslim disciplines—law, history, and the like. By the fourteenth century this method, which had its religious origins in the Qurʾan and the traditions of the Prophet and its philosophic origins in Greek rationalism, had reached its height.

Ibn Khaldūn was an accomplished student of Muslim learning, but witnessing the decline of Muslim power and metaphysics, he decided to seek the concrete causes of this decline. History rather than metaphysics gave the answers to his questions about the changes in Islam's fortunes.

Ibn Khaldūn sought not only historic truth but history as the way to truth. The Preface of his *Muqaddimah* reveals him to be a forerunner of all modern historicists. His Muslim predecessors had narrated the train of historic events; as he said, they saw history on the surface as no more "than information about political events. … They over-looked its inner meaning," which "involves speculation and an attempt to get at the truth, subtle explanation of the causes and origins of existing things, and deep knowledge of the how and why of events." For Ibn Khaldūn history is therefore firmly rooted in philosophy and deserves to be accounted a branch of philosophy.

Traditional Muslim theologians, who were best represented by Mohammad al-Ghazālī, rejected the Aristotelian notion of natural causality. They conceived of God as the first and only cause of all that is. Ibn Khaldūn, as a Muslim believer, agreed with their ontology but introduced natural causality into history. Reason can see historic causes, not ontological causes. God in revelation is the teacher of ontological causes. Reason can grasp the limited phenomenon, but revelation introduces the limitless.

Ibn Khaldūn's concern with historic methodology led him to historicism. For historic accuracy Muslim historiography had relied on the criticism of the sources. It elaborated on the method of *hadith,* the study of the traditions and sayings of the Prophet. Ibn Khaldūn criticized this method and called for philosophical and rational methodology. The test of the accuracy of an event is not the reliability of the source but its conformity to the natural character or the natural law that the event should manifest.

To attain accuracy, the historian should therefore be a student of sociological and political causes and laws. He ought to be a philosopher of history.

> [If the historian] trusts historical information in its plain transmitted form and has no clear knowledge of the principles resulting from custom, the fundamental facts of politics, the nature of civilization, or the conditions governing social organization, and if, further-more, he does not evaluate remote or ancient material through comparison with near or contemporary material, he often cannot avoid … deviating from the high road of truth.

Ibn Khaldūn called this introductory science to the study of history the science of *ʿumran,* or the science of civilization, and claimed to be its originator.

Civilization is the beginning and end of social development and political organization. Man is born naturally sociable. Society rises through man's ability to cooperate with other men for the satisfaction of his natural needs. Countrymen or nomads in primitive or tribal societies seek the satisfaction of their elementary need for food; townsmen in urban and more complex societies pursue higher economic, intellectual, and spiritual needs. Political organization, or the state, arises from individual and social needs for restraint, arbitrage, defense, and prosperity.

Asabyia, or group feeling, is the way to achieve leadership, enforce authority, and expand. Political organization or statehood leads to power and prosperity. The state is the form of civilization.

Arts and sciences can prosper only within a state. Resulting luxury is conducive to social and political disintegration. Like individual human beings, all societies, states, and civilizations go through cyclical states of emergence, growth, and decay. Civilizations, however, live longer than states, for the cultural faculties acquired by individuals and societies enable civilizations to survive political disintegration. The systematic formulation of this organistic theory of civilization is full of original observations about the influence of climate on social organization, the forms of society, the economic forces, the relation between labor and value, the psychological, social, and economic foundations of power, the forms of the state, the relation of state and religion, the role of edu-

cation in society, the interdependence of prosperity and culture, and many other subjects.

Because these observations were formulated as natural laws, Ibn Khaldūn has been studied not only as a philosopher of history but also as a sociologist, political thinker, economist, educator, epistemologist, and historian of Muslim sciences. Guided by their own disciplines or convictions, different scholars have proclaimed him a forerunner of Niccolo Machiavelli, Giambattista Vico, Baron de Montesquieu, G. W. F. Hegel, Charles Darwin, Herbert Spencer, Karl Marx, Toynbee, and others. In his methodology and style Ibn Khaldūn is more a modernist than a medievalist. This partially explains his limited influence in medieval times and growing influence in modern times.

Al-Muqaddimah was written at a time when translation from Arabic into Latin had waned. Rediscovered by modern scholars and orientalists in the nineteenth century, excerpts of the book have been translated into French, German, Italian, English, and Japanese.

Ibn Khaldūn has also been rediscovered by modern Muslim and Arab authors. More books in Arabic have been written about him than about any other medieval Muslim thinker. He has influenced historic, sociological, and political writings.

See also al-Ghazālī, Muhammad; Culture and Civilization; Darwin, Charles Robert; Hegel, Georg Wilhelm Friedrich; Historicism; Islamic Philosophy; Logic, History of: Logic in the Islamic World; Machiavelli, Niccolò; Marx, Karl; Montesquieu, Baron de; Philosophy of History; Toynbee, Arnold Joseph; Vico, Giambattista.

Bibliography

Ibn Khaldūn's *Kitab al-ʿIbar* was published in seven volumes (Būlāq, 1867) and translated into French by M'G. de Slane as *Histoire des Berbères* (Algiers: Algerian Government, 1851–1856). *Al-Muqaddimah* has been translated by Franz Rosenthal as *The Muqaddimah: An Introduction to History,* 3 vols. (New York: Pantheon, 1958); it has also been translated into French by M'G. de Slane, 3 vols. (Algiers, 1862–1868).

Most of the biographical data covering Ibn Khaldūn's long and interesting career can be found in his autobiography, *At-Tarif bi-ibn Khaldoun wa rihlatuhu gharban wa-Sharqan,* translated by Muḥammad Tawit at-Tanji as *Biography of ibn Khaldoun and Report on His Travels in the West and in the East* (Cairo, 1951).

Drawing from the autobiography and other sources, Walter J. Fischel wrote the story of Ibn Khaldūn's meeting and peace negotiations with Tamerlane in *Ibn Khaldūn and Tamerlane* (Berkeley: University of California Press, 1952) and *Ibn Khaldūn's Activities in Mameluk Egypt 1382–1406* (Berkeley, 1951).

For works on Ibn Khaldūn see Gaston Bouthoul, *Ibn Khaldoun, sa philosophic sociale* (Paris, 1930); Muḥammad Abdullah Enan, *Ibn Khaldoun: His Life and Work,* translated by Ashraf (Lahore: Muhammad Ashraf, 1941); Abd el-Aziz Ezzat, *Ibn-Khaldoun et sa science sociale* (Cairo: Tsoumas, 1947); Muhsin Mahdi, *Ibn Khaldūn's Philosophy of History: A Study in the Philosophic Foundation of the Science of Culture* (London: Allen and Unwin, 1957; paperback ed., Chicago: University of Chicago Press, 1964); Sobhi Mahmassani, *Les idées économiques d'ibn Khaldoun* (Lyon, 1932); and Hussein Taha, *La philosophie sociale d'ibn Khaldoun* (Paris, 1925).

Hassan Saab (1967)

IBN KHALDŪN [ADDENDUM]

Although most attention has been paid to Ibn Khaldūn's writings on history and society, he made an important contribution to other aspects of Islamic thought. He played a role in the extensive debate over Sufism, and defended its position in Islamic culture, provided that it adhered to the proprieties of religious law. The idea of coming to mystical knowledge by abandoning Islamic law and practices is a constant source of attack by Ibn Khaldūn; he emphasizes the role of the *shaykh*, or spiritual guide, and the place of mysticism within an orthodox understanding of Islam. He was also rather critical of many of the ambitious claims of the philosophers who thought that they could acquire knowledge of the most important features of reality by using reason alone.

Ibn Khaldūn is certainly no enemy of reason, but he argues that it operates within limits, and that religion is required for acquiring deeper knowledge than reason can provide. Similarly, when it comes to political philosophy, he criticizes the highly theoretical approaches of the philosophers who talk about the constitution of the ideal state as though this is something that could be established by reason and nothing else. The state cannot, he argues, be divorced from its history and social structure, and religion has to play a crucial role in its organization and goals. Although his ideas are heavily influenced by the philosophical concepts and arguments current at his time, Ibn Khaldūn was consistently skeptical of the ability of philosophy to reveal much of interest in either political or religious theory.

See also Islamic Philosophy.

Bibliography

Ahmed, Zaid. *The Epistemology of Ibn Khaldun*. London: Routledge, 2003.

al-Azmeh, Aziz. *Ibn Khaldun in Modern Scholarship*. London: Third World Centre, 1981. An extensive bibliography.

Ibn Khaldūn. *The Muqaddimah*. Translated by F. Rosenthal. 3 vols. Princeton, NJ: Princeton University Press, 1958. Ibn Khaldūn's most important work.

Issawi, Charles. *An Arab Philosophy of History*. Princeton, NJ: Princeton University Press, 1986.

Lakhsassi, Adbderrahmane. "Ibn Khaldun." In *History of Islamic Philosophy*, edited by Seyyed Nasr and Oliver Leaman. London: Routledge, 1996.

Mahdi, Muhsin. *Ibn Khaldun's Philosophy of History*. London: Allen and Unwin, 1957.

Oliver Leaman (2005)

IBN PAQUDA, BAḤYA BEN JOSEPH

See *Baḥya ben Joseph ibn Paquda*

IBN RUSHD

See *Averroes*

IBN SĪNĀ

See *Avicenna*

IBN ṬUFAYL

(d. 580 AH/1185 CE)

Abū Bakr Muḥammad ibn ʿAbd al-Malik ibn Ṭufayl, the Islamic philosopher, was known to medieval Scholastics as Abubacer. Few details are known about the life of Ibn Ṭufayl, who was born at Guadix in the province of Granada and died in Morocco. Like all his colleagues, he was a scholar whose knowledge was encyclopedic; he was a mathematician, astronomer, philosopher, and poet. He served as vizier for and was a friend of the Almohad sovereign Abū Yaʿqūb Yūsuf, and it was he who recommended that his friend Averroes be assigned the task of analyzing the works of Aristotle. Ibn Ṭufayl became known to medieval Scholastics (Abū Bakr having become Abubacer) through Averroes's translation of *De Anima*, which contained a brief criticism of Ibn Ṭufayl's doctrine identifying the possible (or passive) intellect with the imagination.

It was, however, because of his "philosophical novel," *Ḥayy ibn Yaqẓān*, a work that remained unknown to the Scholastics, that Ibn Ṭufayl later gained fame. It is worth noting that in the same era in the East Shihāb al-Dīn Yahyā Suhrawardī composed his own tales of symbolic initiations, in which he introduced, by extending the cycle of Avicennian tales, the "oriental philosophy" that Avicenna had already opposed to Peripatetic philosophy, but with only partial success. Ibn Ṭufayl referred to the Avicennian tales in the prologue to his philosophical novel, because he knew that the secret of Avicenna's "oriental philosophy" was partially contained therein.

Ibn Ṭufayl's work, however, is completely original and not in the least a mere amplification of an Avicennian tale. All it owes to Avicenna are the names of the *dramatis personae*: Ḥayy ibn Yaqẓān (*Vivens filius Vigilantis*), and Salamān and Absāl (a spelling certainly preferable to the mutilated form "Asāl," which figures in certain manuscripts).

In the works of Avicenna the name Ḥayy ibn Yaqẓān typified the Active Intellect, the central figure of Islamic Neoplatonism, simultaneously angel of knowledge and angel of revelation (the Holy Ghost and the angel Gabriel). For Ibn Ṭufayl this name is also that of the absolute hermit, mysteriously abandoned or spontaneously born on a desert island; in the absence of any human master and of all social falsification, the hermit becomes the perfect Sage. The superior pedagogy of the Active Intellect alone develops in him its natural faculties through a slow, rhythmic process evolving over the years. On a neighboring, inhabited island live two friends, Salamān, who typifies the practical and social spirit, and Absāl, contemplative and mystical, who lives like one in exile in his own country and finally decides to immigrate to the hermit's island, where he meets Ḥayy ibn Yaqẓān. In the course of their long conversations Absāl discovers that all that had been taught to him in matters of religion Ḥayy ibn Yaqẓān, the solitary, philosophical wise man, already knows, but in a purer form. Absāl discovers that religion is the *symbol* of a truth otherwise inaccessible to the common run of men. Together they attempt to deliver their spiritual message to the men on the island opposite them. Alas! in the face of the growing hostility that they encounter, they must accept an inescapable truth: The ordinary man is not able to understand.

Ibn Ṭufayl's novel is not an anticipation of Robinson Crusoe; each external episode must be understood on a spiritual level. On the other hand, in spite of its pes-

simistic ending it should not be concluded that the conflict Ibn Ṭufayl set forth (that between religion and philosophy) attained desperate proportions in the Muslim faith. In fact, another position and solution to the problem are sought in the "prophetic philosophy" of Shiʿism.

See also Aristotle; Averroes; Avicenna; Imagination; Islamic Philosophy; Logic, History of: Logic in the Islamic World; Peripatetics; Suhrawardī, Shihāb al-Dīn Yaḥyā.

Bibliography

Ibn Ṭufayl's *Ḥayy ibn Yaqẓān* has been translated by S. Ockley as *The Improvement of Reason* (London, 1708) and revised by A. S. Fulton (London, 1929).

For literature on Ibn Ṭufayl, see Henry Corbin, *Histoire de la philosophie islamique,* Vol. I (paperback, Paris: Gallimard, 1964), pp. 327–337 and the bibliography on p. 362.

See also Henry Corbin, *Avicenna and the Visionary Recital,* translated from the French by Willard R. Trask (New York: Pantheon, 1960).

Henry Corbin (1967)

IBN ṬUFAYL [ADDENDUM]

There is indeed a rather pessimistic strand in Ibn Ṭufayl's philosophical novel. When Ḥayy travels to Absāl's island the former is an instant celebrity, but he discovers that the inhabitants are not interested in really coming close to the truth; they are only prepared to adhere to the symbols of their religion and have no interest in peering behind those symbols at the deeper truth they represent. When Ḥayy investigates the rules of religion he finds a good deal of discussion on matters for which he has no time at all, issues about money and possessions and other material topics. Ḥayy and Absāl eventually give up and return to Ḥayy's island where they can live in seclusion and avoid the infelicities of social life among a population unconcerned about spiritual truth. The implication is that religions such as Islam are built on solid principles, but most of their adherents never appreciate the nature of these principles—they remain at a more superficial level of understanding and merely carry out the rituals of the religion without investigating their roots. Ibn Ṭufayl's version of the story of Ḥayy ibn Yaqzān is much more radical than that of his predecessor Ibn Sīnā, who hinted at the role of mystical knowledge but did not explicate its centrality in religion. Ibn Sīnā is guarded throughout his account, using allusion rather than direct argument to make his points. Ibn Ṭufayl writes with boldness and clarity and does not hesitate to present his highly critical analysis of traditional religion as it is normally understood.

See also Islamic Philosophy.

Bibliography

Conrad, Leonard, ed. *The World of Ibn Tufayl: Interdisciplinary Perspectives on Hayy Ibn Yaczan*. Leiden, Netherlands: Brill, 1996.

Goodman, Lenn. "Ibn Tufayl." In *History of Islamic Philosophy*, edited by Seyyed Nasr and Oliver Leaman. London: Routledge, 1996. See pp. 313–329 for a useful bibliography and guide to the secondary literature.

Hawi, Sami. *Islamic Naturalism and Mysticism: A Philosophical Study of Ibn Tufayl's Hayy Yaqzan*. Leiden, Netherlands: Brill, 1974.

Ibn Ṭufayl. *Ḥayy ibn Yaqzān*. Translated by Lenn Goodman. Los Angeles: Gee Tee Books, 1983.

Oliver Leaman (2005)

IBN ZADDIK, JOSEPH BEN JACOB

(d. 1149)

Joseph ben Jacob ibn Zaddik, like other Jewish philosophers of a Neoplatonic cast, such as Yehuda Halevi and Abraham ibn Ezra, was a poet as well as a philosopher and legist. Very few of his poems survive, and although he was highly praised as a Talmudist and served for the last eleven years of his life (1138–1149) as judge (*dayyan*) of the Jewish community of Córdoba, he does not seem to have written any systematic legal work. His philosophic work, on which his chief reputation rests, was originally written in Arabic, but the original no longer survives; a Hebrew translation, under the title *Olam Katon* (The Microcosm), was circulated in manuscript during the Middle Ages but was not printed until the mid-nineteenth century.

The general thesis of Joseph ibn Zaddik's work is that since man's nature duplicates in reduced form the nature of the universe, knowledge of the self provides a key to all knowledge. It is unnecessary to study the special sciences. The study of man, the microcosm, will lead to the understanding of the universe, the macrocosm.

Ibn Zaddik's *Olam Katon,* in fulfillment of this program, develops in its first part a metaphysical basis for the theory of man as the microcosm. Here the author showed acquaintance with both the Platonic and the Aristotelian traditions in the form in which they were maintained by

Muslim philosophers. The second part of the work discusses both the physical and the psychological natures of man; it asserts a point-for-point correspondence between human nature and the physics of the universe. In the third part Ibn Zaddik turned to theological questions, particularly the question of divine unity. His theological discussion includes a proof of creation from the finiteness of the world: Where there is creation, there must be a Creator; hence God exists. The philosopher was aware of the difficulties presented by a naive doctrine of divine attributes and resolved these difficulties by denying to the attributes a positive character. The fourth and final division of the work, continuing the pattern established by the Muslim philosophers, is devoted to God's justice and the implications of the divine government of the universe for man's duties. Ibn Zaddik was firmly committed to a belief in human free will; he believed that a man must use his freedom to imitate the goodness of God and to seek knowledge of him. Success or failure in so doing leads to reward or punishment in the future life, but apparent rewards and punishments in this world are merely natural happenings and should not be understood as indications of divine favor or disfavor.

See also Aristotelianism; Determinism and Freedom; Halevi, Yehuda; Jewish Philosophy; Macrocosm and Microcosm; Platonism and the Platonic Tradition.

Bibliography

WORKS BY IBN ZADDIK

The Microcosm. Translated by J. Haberman. London: Associated University Press, 1954.

WORKS ON IBN ZADDIK

Samuelson, N. "Medieval Jewish Aristotelianism: An Introduction." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman. London: Routledge, 1997.

J. L. Blau (1967)
Bibliography updated by Oliver Leaman (2005)

IDEALISM

"Idealism" in its philosophical sense, is the view that mind and spiritual values are fundamental in the world as a whole. Thus, idealism is opposed to naturalism, that is, to the view that mind and spiritual values have emerged from or are reducible to material things and processes. Philosophical idealism is also opposed to realism and is thus the denial of the commonsense realist view that material things exist independently of being perceived. Some philosophers who have held the idealist view in its antinaturalist form have not opposed commonsense realism, and thus it is possible to be a metaphysical idealist and an epistemological realist. More often, however, arguments against commonsense realism have been used in order to establish metaphysical idealism. The description "subjective idealism" is sometimes used for idealism based on antirealist epistemological arguments, and the description "objective idealism" for idealism that is antinaturalist without being antirealist.

In terms of these definitions, philosophical theism is an idealist view, for according to theism God is a perfect, uncreated spirit who has created everything else and is hence more fundamental in the world than any material things he has created. Marxist philosophers have therefore held that there are in principle only two main philosophical systems: idealism, according to which mind or spirit is primary in the universe, and materialism, according to which matter is primary in the universe. If "primary" is taken not to mean "earlier in time" but rather to mean "fundamental" or "basic," then these Marxist definitions agree with those given above. The only objection to them is that many philosophers who accept theism would be unwilling to be labeled idealists, since they would take the view that idealists belittle the material world and regard it as illusory by comparison with mind or even as less real than mind, whereas theists do not belittle matter or regard it as in any way less real than mind. Certainly this is a difference between theism and some forms of idealism, but there is force in the argument that theism and both subjective and objective idealism may be classed together as opposed to materialism. Pantheism may be regarded as a more thoroughly idealist view than theism, since pantheism is the view that nothing exists except God and his modes and attributes, so that the material world must be an aspect or appearance of God. Theism, in contrast, is the view that God has created a world beyond or outside himself so that the material world, although dependent on him, is not an aspect or appearance of him. What unites idealism both with theism and with pantheism is the rejection of materialism and the assertion of a metaphysic that is favorable to religious belief.

HISTORY AND ORIGIN OF THE TERM

The word *idealism* came to be used as a philosophical term in the eighteenth century. Gottfried Wilhelm Leibniz, in his *Réponse aux réflexions de Bayle* (written 1702; published in *Philosophischen Schriften*, edited by C. I. Gerhardt, 7 vols. Berlin, 1875–1890), criticized "those

who like Epicurus and [Thomas] Hobbes, believe that the soul is material" and held that in his own system "whatever of good there is in the hypotheses of Epicurus and of Plato, of the greatest materialists and the greatest idealists, is combined here" (Vol. IV, pp. 559–560). In this passage Leibniz clearly means by "idealists" philosophers who uphold an antimaterialist metaphysic like that of Plato and himself. When, later in the century, George Berkeley's views came to be discussed, the word *idealism* was applied, however, to the view that nothing could be known to exist or did exist except the ideas in the mind of the percipient. (Berkeley called his own view "immaterialism," not "idealism.") Thus, Christian Wolff (1679–1754), a follower of Leibniz, included idealists, along with materialists and skeptics, among "three bad sects" that he reprobated, and Denis Diderot (1713–1784) wrote in 1749: "We call *idealists* those philosophers who, knowing only their own existence and that of the sensations that follow one another within them, do not grant anything else" (*Lettre sur les aveugles*, London, 1749). The term *egoists* was also applied to holders of this view, as can be seen from the article titled "Égoistes" in the *Encyclopédie*, edited by Jean Le Rond d'Alembert and Diderot, which started publication in 1750. Today the word *solipsists* is applied to what were then called "egoists" or "idealists." In the *Critique of Pure Reason* (Riga, 1781) Immanuel Kant referred to his own view as "transcendental idealism," and in his *Prolegomena to any Future Metaphysics* (Riga, 1783) he called it "critical idealism." Thus, by this time the word *idealism* was beginning to lose the pejorative meaning that had linked it with extreme subjectivism.

The word *idealism* is derived from the Greek word *ἰδέα*, which simply means something seen, or the look of something. Plato used the word as a technical term of his philosophy to mean a universal (such as whiteness) in contrast to a particular (such as something white) or to mean an ideal limit or standard (such as absolute Beauty) in contrast to the things that approximate or conform to it (such as the more or less beautiful things). According to Plato an Idea, or Form, is apprehended by the intellect, does not exist in time, and cannot come into existence or cease to exist as temporal things do and is hence more real than they are. In medieval philosophy Ideas or Forms were regarded as the patterns in accordance with which God conceived of things and created them, and hence they were thought of as existing in the mind of God. René Descartes used the word *idea* for thoughts existing in the human mind, sometimes retaining, however, the intellectual and objective character of ideas as understood in the Platonic tradition. But he also used the word *idea* for the effects in embodied minds of external objects acting on the sense organs, and hence the word came to stand for changing sense perceptions as well as for unchanging objects of the intellect. Descartes also used the word *idea* for a shape or form stamped upon a soft material, as when he said in Section XII of his *Rules for the Direction of the Mind* (1628) that "shapes or ideas" are formed *in the brain* by things outside the body acting upon it. John Locke, in *An Essay concerning Human Understanding* (London, 1690), used the word *idea* for perceptions of "sensible qualities" conveyed into the mind by the senses and for "the perception of the operations of our own mind within us, as it is employed about the ideas it has got" (Bk. II, Ch. I, Sec. 4). The mind, he held, "stirs not one jot beyond those ideas which *sense* or *reflection* have offered for its contemplation" (ibid., Sec. 24). Berkeley adopted Locke's terminology and held that by our senses "we have the knowledge *only of our sensations*, ideas, or those things that are immediately perceived by sense" (*Principles of Human Knowledge*, Dublin, 1710, Sec. XVIII). Thus, Berkeley here repeats a view already held by Locke.

Thus, the word *idea* was used variously to mean a Form in the Platonic sense, a Form as apprehended in the mind of God or by the human mind, a shape impressed on soft, yielding material, and, apparently by analogy with this last sense, a modification produced in a mind by the influence on it of external things that affect the sense organs. Neither a Platonic Form nor a shape is a mental entity. "Operations of the mind" clearly are, and so would be the effects in minds of material objects that produce "impressions" in them. Ideas in this last sense would seem to be like mental images, but mental images produced not by imagining but by the operation of external objects. This variation in meanings can be seen in Berkeley's *A New Theory of Vision* (Dublin, 1709), where he writes (Sec. XLI): "a man born blind being made to see, would, at first, have no idea of distance by sight; the sun and stars, the remotest objects as well as the hearer, would all seem to be in his eye, or rather in his mind. The objects intromitted by sight, would seem to him (as in truth they are) no other than a new set of thoughts or sensations, each whereof is as near to him as the perceptions of pain or pleasure, or the most inward passions of his soul." It will be noticed that in his passage Berkeley comes close to assimilating "in his eye," a physical condition, to "in his mind," meant presumably to be a mental condition. Again, he puts "sensations" in apposition with "thoughts," although sensations and thoughts would seem to be as different as pains and concepts. There is also the suggestion that what is near to us is "in the mind," so that if colors and shapes are not, as they seem to be, at a distance

from us, they must be in our minds. The passage is an important one for indicating the conflicts and confusions involved in the word *idea* and carried over into some of the arguments for idealism.

IMMATERIALISM

Berkeley gave the name "immaterialism" to the central thesis of his philosophy, the thesis that there is no such thing as material substance. Immaterialism has been prominent in idealist theories just because to prove that there is no material substance would be the most effective and spectacular way of disproving materialism. If there is no material substance, then matter cannot be the basis of what is or all that there is. Immaterialism has been supported by two main lines of argument. Along one line it has been argued that it is impossible that matter could be independently real. The arguments to this effect may be called the metaphysical arguments for immaterialism. Along the other line it has been argued that the colors, shapes, and sounds that are naturally taken to belong to independently existing material objects are in fact sensible qualities that cannot exist apart from being perceived.

The arguments to establish this may be called the epistemological arguments for immaterialism. Although he did not call himself an immaterialist, Leibniz, on the evidence of the passage we have quoted, would have regarded himself as an idealist, and his arguments were metaphysical rather than epistemological. Berkeley, of course, is best known for his epistemological arguments, even though his argument that the very notion of something existing totally unperceived is self-contradictory may be classed as metaphysical. Arthur Collier, in his *Clavis Universalis* (London, 1713), used both epistemological and metaphysical arguments; the subtitle of his book, "a Demonstration of the Non-existence or Impossibility of an External World," allowed for both types of approach.

LEIBNIZ. Leibniz's metaphysical idealism consisted of two main theses: (1) that matter is necessarily composite and hence cannot be substantially or independently real, and (2) that simple (that is, noncomposite) substances must be perceiving and appetitive beings even though they are not necessarily conscious or self-conscious. He gave the name "monad" to these independently real and essentially active substances, and he argued that space and time cannot be real containers in which substances exist but must be the order in which monads are related to one another. Thus, he held that space and time are not absolute existences but relations of coexistence and succession among created monads. He did not conclude from this, however, that space and time and material objects are mere illusions or delusions; delusions and dreams, he held, are by their very nature inconsistent and unpredictable, whereas the material world in space and time is regular and in part predictable. Leibniz was not quite explicit on the matter, but he seems to have believed that space and time were a sort of mental construction or *ens rationis* and that material things are regular appearances rather than real substances. Sometimes, however, he used the expression *phenomena bene fundata* for space and time.

However this may be, Leibniz argued for an idealist system in which there is a series of realms of being with God as the supreme, uncreated spiritual substance. In the realm of created substances all the members are active and immaterial and some are self-conscious substances created in God's image. In the realm of appearances the elements are "well-founded" in the substantial realities, and in consequence they show a rational order even though, like the rainbow, they disappear when closely examined. Finally, there are isolated realms of mere illusion and delusion that, however, have their place in the total scheme of things. Leibniz believed that this metaphysical system could be proved by reason. He held, too, that sense experience is not an independent source of knowledge but is reason in a state of obscurity and indistinctness. Thus, he held that "we use the external senses as … a blind man does a stick" and that the world is revealed as it is by means of reason, not by means of the senses (*Letter to Queen Charlotte of Prussia*, 1702). Thus, he denied not only the substantial reality of matter but also the efficacy and even the possibility of mere sense experience. This is a theme that many later idealists have developed. It runs counter, however, to the empiricist immaterialism of Berkeley.

BERKELEY. Berkeley is the best-known exponent of immaterialism on epistemological grounds. His basic argument is that what we immediately perceive are sensations or ideas, that sensations or ideas are necessarily objects of perception (their *esse*, as he put it, *is percipi*, their essence is to be perceived), and that what we call physical things, such as trees and rocks and tables, are orderly groups or collections of sensations or ideas and are hence mind-dependent like the sensations or ideas that compose them. This argument proceeds on the assumption that sense experience is basic and reliable. Matter is rejected on the ground that the senses inform us of ideas but not of material substances to which these ideas belong. The very notion of a material substance dis-

tinct from sensible qualities or ideas is, according to Berkeley, unimaginable and inconceivable.

Berkeley made the surprising claim that this view is in full accordance with common sense. According to common sense, he argued, trees and rocks and tables are immediately perceived and have the characteristics they are immediately perceived to have. But according to those who believe in material substance, what is immediately perceived are the ideas produced in the mind by material substances of which we can only have mediate or indirect knowledge. Furthermore, these indirectly perceived material substances do not have the characteristics of color, hardness, etc., which common sense says they have. Hence, Berkeley thought that material substances, even if they were conceivable, would be problematic existents, so that the theory in which they figured would give rise to skepticism about the existence of familiar things like trees and rocks and tables. Immaterialism, in contrast, with its claim that such things, being ideas, are immediately perceived, does not lead to skepticism about them.

In its reliance on sense experience, then, and in its acceptance of the view that trees and rocks and tables are immediately perceived and are as they seem to be, Berkeley's immaterialism is very different from that of Leibniz. On the other hand, there is an important point of similarity between their views that is often overlooked. Leibniz held that substances, or monads, that is, the basically real things that make up the world, must be active, perceiving beings. Berkeley held this too, for he argued that sensible qualities or ideas are dependent and passive existences that depend on independent and active beings. These independent and active beings, according to Berkeley, are selves. The difference between Berkeley and Leibniz is that Berkeley held that only selves are active, whereas Leibniz held that activity is possible at a lower level than that of selves. However, this view that what is real is active is an element in a number of idealist theories.

Berkeley also supported immaterialism with the argument that it is not possible even to conceive of anything existing apart from being thought of, for it must be thought of in the very act of being conceived. This argument was not used by Leibniz, but it has played an important part in the arguments of many idealists since Berkeley.

COLLIER. Arthur Collier's *Clavis Universalis*, which appeared posthumously in 1713, was possibly written before Berkeley's *Principles of Human Knowledge* (1710), in which Berkeley's immaterialist philosophy was first published. Collier used epistemological arguments to prove immaterialism, but, unlike Berkeley, he made no attempt to reconcile immaterialism with common sense. On the contrary, he said that in denying the existence of the material world he meant that bodies are as delusory as the visions of lunatics. Collier also produced metaphysical arguments for immaterialism, maintaining, for example, that matter can be proved to be both infinite in extent and not infinite in extent, infinitely divisible and not infinitely divisible, and since nothing can in fact have contradictory characteristics, matter cannot exist.

Knowledge of immaterialism was spread in Germany by the publication of a book that contained German translations of Berkeley's *Three Dialogues between Hylas and Philonous* (London, 1713) and Collier's *Clavis Universalis* and whose title was *Sammlung der vornehmsten Schriftsteller die die Wirklichkeit ihren eigenen Körper und der ganzen Körperwelt leugnen* (Rostock, 1756). The translator and editor, Johann C. Eschenbach, set out to refute as well as to translate the two books.

TRANSCENDENTAL IDEALISM

Kant, in his *Critique of Pure Reason*, described his own view as formal, critical, or transcendental idealism. Nevertheless, a famous passage of that book (B 274) is headed "Refutation of Idealism." Kant called the types of idealism he claimed to be refuting problematic idealism and dogmatic idealism, respectively. By problematic idealism he meant the view, which he attributed to Descartes, that the existence of objects in space outside us is doubtful. By dogmatic idealism he meant the view, which he attributed to Berkeley, that "space and all the things to which it belongs as an inseparable condition" is "something impossible in itself and hence looks upon things in space as mere imaginations" (B 274). Kant's interpretation of Descartes is not quite adequate, but his interpretation of Berkeley is so completely at fault that it seems possible that he had made use of Eschenbach's book and confused Collier's arguments with those of Berkeley.

In any case, Kant's transcendental idealism is very different from the types of idealism we have so far considered. Kant held that it is not possible to gain knowledge of the world by rational thought alone, and thus he rejected all attempts such as those of Leibniz and Wolff to do so. Nonetheless, he also held that mere sense experience does not give knowledge of the world either, since in the absence of interpretation, sense experience is "blind." Thus, Kant argued that unless our perceptions were organized within what he called the pure a priori intuitions of space and time in terms of rational principles

such as the requirement that our perceptions refer to things in causal relation with one another, knowledge of an objective world would be impossible. Without the a priori intuitions of space and time and the categories of the understanding, there would be a manifold of fluctuating sensations but no knowledge of the natural world. When Kant refuted the two types of idealism mentioned above, he argued that no one could become aware of himself unless there were enduring material substances with which he could contrast his own fleeting experiences. We should not be aware of selves unless we were also aware of material things. This line of argument disposes of the view that we could be certain of our own existence but doubtful about the material world and also of the view that material things are "mere imaginations." Unless there were material things in space, we should not know of our own existence or of our own imaginations.

Kant's transcendental idealism, therefore, is his view that space and time and the categories are conditions of the possibility of experience rather than features of things as they are in themselves. Whether things-in-themselves are in space and time and whether they form a causally interacting system we do not know, but unless we were so constituted as to place everything in spatiotemporal contexts and to synthesize our sensations according to the categories of the understanding, we should not have knowledge of an objective world. Kant did not think that this synthesizing was carried out by the empirical selves we are aware of in ourselves and others. He thought, rather, that a transcendental self had to be postulated as doing this, but of this transcendental self nothing could be known, since it was a condition of knowledge and not an object of knowledge. The natural world, or the world of appearances, as he calls it, somehow depends on a transcendental self of which we can know nothing except that it is. Whereas at the empirical level selves and material things are equally real, the knowledge we have at this level presupposes the synthesizing activities of a transcendental self of which we can know nothing.

Kant was regarded in his own day as a destroyer not only because he maintained that there was no basis for the rationalist, metaphysical constructions of Leibniz and Wolff but also because he held that no single one of the traditionally accepted arguments for the existence of God was valid and that it is impossible to prove the immateriality and immortality of the soul. Idealists such as Leibniz and Berkeley and Collier had considered that they had framed philosophical arguments that favored religious belief. Berkeley, for example, emphasized that his conclusions made atheism and skepticism untenable. He also claimed to have provided a new and cogent argument for the existence of God. According to Kant, however, sense experience cannot lead us beyond the natural world, and the categories of the understanding can be validly applied only where there are sense experiences and if applied beyond them can lead only to insoluble antinomies. For example, if the category of cause is used to transcend sense experience, then equally valid proofs can be made to show that there must be a first cause and that there cannot be a first cause. In the appendix to the *Prolegomena* Kant says that "idealism proper always has a mystical tendency" but that his form of idealism was not intended for such purposes but only as a solution of certain problems of philosophy. All this seems to place Kant outside the main idealist tradition and to indicate that he was developing a positivistic view. Nevertheless, at the end of the eighteenth century a group of philosophers who are known as Absolute idealists claimed to have been inspired by him. What, then, are the features of Kant's idealism that gave rise to views so different from his?

One is that Kant called specific attention to the elements of activity and spontaneity in knowledge. His view that knowledge of nature would be impossible apart from the activity of the understanding in synthesizing sensations in accordance with the categories led some of his successors to regard knowledge as analogous to construction or making. Another feature of Kant's philosophy that pointed in the direction of Absolute idealism was the thesis that synthesizing in terms of the categories presupposed a unitary transcendental self. It is true that Kant himself said that as a presupposition of experience the transcendental self could not be an object of knowledge, but some of his successors claimed to be rather more familiar with it.

Some of Kant's views on morality and on freedom of the will also gave scope for development in an idealist direction. Kant held that the free will problem is insoluble by metaphysical argumentation, for it can be proved both that there must be a freedom of spontaneity and that there is no freedom and everything takes place according to laws of nature. But in his ethical writings that followed the *Critique of Pure Reason*, Kant argued that our knowledge of and respect for the moral law presupposed freedom of the will. He emphasized that this was not a metaphysical or speculative proof; his point was that metaphysics could not disprove freedom of the will, so that we are justified in accepting what morality presupposes. He argued, furthermore, that the existence of God and the immortality of the soul might also be accepted as practical concomitants of morality, as long as the fundamental impossibility of their being theoretically

proved was recognized. Again, Kant introduced into his account of knowledge a faculty of reason (*Vernunft*), which, remaining dissatisfied with the understanding's confinement to the ordering of sense experiences, constantly strove for completeness and totality. Kant thought that the reason might in practice advance our knowledge by seeking for a completeness that is not in fact to be found—Kant used the expression *focus imaginarius* in this connection. Some of his successors transformed this suggestion into the claim that reason reveals a real, not an imaginary or merely methodological, totality.

ABSOLUTE IDEALISM

FICHTE. The development from Kant's idealism to Absolute idealism can be most readily seen in the writings of Johann Gottlieb Fichte (1762–1814). Like Kant, Fichte believed that strict determinism is incompatible with morality and that our knowledge of the moral law presupposes the freedom of the will. Therefore, the philosopher is faced with choosing between two systems of thought, the deterministic system that Fichte called "dogmatism," of which Benedict de Spinoza is the chief representative, and "critical idealism." Fichte recognized that the philosophy a man chooses depends on the sort of man he is, but he also thought that reasons could be given for preferring the idealist course. A reason on which Fichte placed great weight is that thought and intelligence cannot be accounted for within a system of causes and effects, for, in comprehending causal determination, they necessarily go beyond it. If, therefore, there is to be a fundamental account of things, it must start from the intellect. Fichte was here developing a suggestion by Kant in the *Groundwork of the Metaphysic of Morals* that the operations of the intellect transcend the phenomenal series of causes and effects. Thus, according to Fichte a free, intelligent ego (*Ich*) must be the starting point of philosophy, and everything else must somehow be "deduced" from this ego. Fichte, therefore, endeavored to go beyond Kant by showing that space and time and the categories are not just facts that must be accepted as they are but necessary conditions of intelligence. Even the material world is not merely matter of fact but is presented as a series of obstacles that must be overcome in the performance of our duties.

SCHELLING. Friedrich Wilhelm Joseph von Schelling (1775–1854) began his philosophical career as a supporter of Fichte—as the titles of two of his early works show: *Vom Ich als Prinzip der Philosophie* (Tübingen, 1795) and "Philosophische Briefe über Dogmatismus und Kritizismus" (in *Philosophische Journal*, 1796). Schelling's first account of his distinctive views was titled *System des transzendentalen Idealismus* (Tübingen, 1800), but he later described his view as "absolute idealism," explaining that things are always conditioned by other things, whereas mind is undetermined and absolute. Fichte's idealism has sometimes been called a "moral idealism," since its basis is a system of active moral beings. Schelling's has sometimes been called an "aesthetic idealism," since Schelling argued that it is the artist who makes men aware of the Absolute. Although, like Fichte, he believed that free activity is basic in the world, he placed less emphasis on the distinction between individuals and came nearer to pantheism.

HEGEL. Georg Wilhelm Friedrich Hegel (1770–1831) is too individual a philosopher to be readily classifiable, but he was undoubtedly the most comprehensive and the most influential of the Absolute idealists. In his *Encyclopedia* (Sec. 95) he writes of "the ideality of the finite," which he says is "the main principle of philosophy," and says that "every genuine philosophy is on that account idealism." Like much that Hegel wrote, this is somewhat cryptic, but it appears to mean that what is finite is not real and that the true philosophy, idealism, recognizes this. The matter is more fully discussed in the *Science of Logic* (Bk. 1, Sec. 1, Ch. 2), where Hegel says that philosophical idealism is the view that "the finite is not genuinely real." Here he also contrasts his form of idealism with subjective idealism and says that in denying the reality of the finite, idealist philosophy is at one with religion, "for religion no more admits finitude to be a genuine reality, than it admits finitude to be ultimate, absolute, or as basic (*ein Nicht-Gesetztes*), uncreated, eternal."

We need not linger over Hegel's rejection of subjective idealism, except to refer to what he says about Berkeley's immaterialism in the *Lectures on the History of Philosophy*. Hegel there argues that Berkeley says very little when he says that things are ideas, for this only amounts to recommending a change of nomenclature and calling things ideas, and this throws no new light on the status of things and ideas. Hegel's arguments are metaphysical rather than epistemological. He thought that Fichte was right when he tried to deduce or give reasons for the categories, and Hegel's *Science of Logic* may be regarded as his view of how the deduction should be carried out. Insofar as such a compact work can be summarized, its argument is that we say very little about the world when we say that it is, rather more when we say that it is measurable, or that it is a series of interacting things, more again when we think of it in terms of chemical

combinations, still more when we apply the categories of life, more again when we apply the categories of theoretical reason, and most of all when we come to the categories of will and the pursuit of the good. What remains of the older metaphysical arguments is his view that the incomplete and inadequate categories lead to contradictions. These contradictions, Hegel held, are resolved as the higher categories are reached, in particular the category of the Absolute idea.

Hegel also tried to show that rudimentary mind operates in the natural world. But what most concerned him was the working of mind in human society. He set out a series of stages of human achievement proceeding from the family organization to "civil society" (what today we call the market economy), from civil society to the state, and then, at the highest levels, to art, religion, and philosophy. The idealist character of this construction may be seen from the fact that when Marx wished to set out a materialist view of society he took the economy as basic, the state as dependent on it, and regarded art, religion, and philosophy as ideologies that had no real influence.

Hegel's philosophy was elaborated after his death by a series of able successors and criticized from many points of view. It came to be known in England about the middle of the century, and Benjamin Jowett translated some passages (which he never published) for the use of his students. Absolute idealism was made known to a larger British public by James Hutchinson Stirling's *The Secret of Hegel* (2 vols., London, 1865), (Fichte's moral idealism had earlier influenced Thomas Carlyle, and Samuel Taylor Coleridge had been influenced by his reading of Schelling, although he had not accepted all of Schelling's views. William Wordsworth's definition of poetry as "emotion recollected in tranquillity" seems to be a translation of a phrase of Schelling's that Coleridge noticed and copied into his notebook).

NEO-HEGELIANISM

About the time when German Absolute idealism was becoming known in England through the writings of Coleridge and Carlyle, it was also becoming known in the United States through a group of writers (mostly Unitarians) who came to be called the transcendentalists. Later, in the 1860s, idealist philosophy received more detailed and professional attention on both sides of the Atlantic. In 1867 at St. Louis, William Torrey Harris founded the *Journal of Speculative Philosophy*, in the first issue of which he referred disparagingly to the prevailing "brittle individualism" that he considered should be replaced by a philosophy in which the state was properly comprehended as a support for freedom. In the same period Thomas Hill Green was teaching philosophy at Oxford with the support of Jowett. The nature of Green's influence may be seen from a letter sent to Green in 1872 asking him to speak to an essay society whose members felt the need for "earnest effort to bring speculation into relation with modern life instead of making it an intellectual luxury, and to deal with various branches of science, physical, social, political, metaphysical, theological, aesthetic, as part of a whole instead of in abstract separation," and sought for "co-operation instead of the present suspicious isolation." This letter was signed by, among others, F. H. Bradley, who had recently become a fellow of Merton College (Melvin Richter, *The Politics of Conscience. T. H. Green and His Age*, London, 1964, pp. 159–160). Both Harris and his circle and Green and his were critical of social individualism as well as of positivism and materialism. They aimed to provide an alternative to utilitarianism, which they thought was based on an inadequate pluralistic metaphysics.

Green's form of idealism was rather closer to that of Kant than to that of Hegel. It was built around two main themes, that the natural world cannot be self-contained and ultimate, and that there is no merely given experience. The first theme is an extension of Kant's theory of the transcendental ego, which Green held implied that nature presupposes "a principle which is not natural," a "spiritual principle" (*Prolegomena to Ethics*, Oxford, 1883, Sec. 54). The second theme, on the other hand, goes well beyond Kant, who believed that there was a "manifold of sense" which the understanding synthesized. Green's view that there is no merely given sense experience, and that all experience implies some sort of intelligent organization, was a central theme of subsequent idealist argument. It has a certain kinship with Leibniz's theory that ideas of sense are confused ideas of reason.

Green died in 1882, and the leading English idealist philosophers after that were F. H. Bradley and Bernard Bosanquet. In Scotland, where idealism very soon prevailed in the universities, Edward Caird's *A Critical Account of the Philosophy of Kant* (Glasgow, 1877) and Andrew Seth's (later Pringle-Pattison's) *Hegelianism and Personality* (London and Edinburgh, 1887) were notable contributions. But from the 1880s to the 1920s Bradley and Bosanquet dominated the philosophical scene in Great Britain. Bradley attempted to discredit the commonsense view of the world by bringing to bear a multitude of arguments to show that it involved self-contradictions, and he argued that these contradictions could

be eliminated only if the world is shown to be a single, harmonious experience. The central theme of Bosanquet's idealism was that every finite existence necessarily transcends itself and points toward other existences and finally to the whole. Thus, he advocated a system very close to that in which Hegel had argued for the ideality of the finite. Bradley and Bosanquet influenced one another a great deal. For example, Bradley's *Ethical Studies* (London and Edinburgh, 1876) influenced Bosanquet's account of society, and Bosanquet's *Knowledge and Reality* (London, 1885) led Bradley to modify very considerably the views he had set out in his *Principles of Logic* (London, 1883).

In the United States the most impressive contribution to the philosophy of idealism is Josiah Royce's *The World and the Individual* (first series, New York, 1900; second series, New York, 1902). Royce was extremely learned in the literature of idealism, both German and British, and *The World and the Individual* was written in the light of his study of Kant, Fichte, Schelling, and Hegel and of his reading of Bradley's *Appearance and Reality* (London, 1893). Furthermore, Royce was acquainted with the empiricism and pragmatism of C. S. Peirce and William James and with Peirce's work in formal logic. Like Pringle-Pattison, Royce considered that Bradley went too far in regarding the individual mind as "fused" or "transformed" in the Absolute. The mystic who regards finite experience as mere illusion, Royce held, is an improvement on the realist who uncritically accepts it just as it is, but nevertheless the very point of idealism would be lost if the individual self is deprived of all cosmic significance. Royce believed he could show that the "world … is a realm of individuals, self-possessed, morally free, and sufficiently independent of one another to make their freedom of action possible and finally significant" (*The World and the Individual*, first series, p. 395). Like Fichte, Royce endeavored to support this view by an analysis of the moral, rational will.

By the beginning of the twentieth century idealism had become a powerful force in the universities of the English-speaking world. Empiricism and realism were held to have been finally discredited, along with the utilitarianism and individualism that had so often accompanied them. Philosophical truth was thought to be a unity, so that similar principles animated idealist works on aesthetics, ethics, religion, and politics. Such leading British statesmen as Arthur J. Balfour and Richard B. Haldane and the South African prime minister Jan C. Smuts wrote books defending the idealist point of view. When the new provincial universities were being founded in Great Britain at that time, Haldane used his influence to foster the study of philosophy in them, as a central, unifying subject.

At the same time, however, points of view opposed to idealism were being vigorously developed. An example is G. E. Moore's "The Refutation of Idealism," which appeared in *Mind* (n.s. 12 [1903]: 433–453). Another example is *The New Realism*, a collection of articles by American philosophers critical of idealism that was published in New York in 1912. Bertrand Russell urged that idealists were ignorant of new developments in logic and that this rendered their theories untenable. Ludwig Wittgenstein's *Tractatus Logico-Philosophicus* (London, 1922) was symptomatic of a new, pluralist, antispeculative approach to philosophical problems. Moore's "The Conception of Reality" (*PAS*, 1913–1914) attempted to show that one of Bradley's theses was nothing but a consequence of his not realizing that the proposition "Unicorns are thought of" is of quite a different logical form from the proposition "Lions are hunted." In the 1920s the very possibility of speculative metaphysics was denied on the basis of the allegedly empiricist principle of verifiability. Furthermore, the idealist theses about the "unreality" of finite individuals and the "reality" of society or the state were held to be evil as well as meaningless.

But during this period when the idealist movement was under increasing attack, three important treatises appeared in which comprehensive idealist theories were developed. John M. E. McTaggart's *The Nature of Existence* (Cambridge, U.K., 2 vols., 1921–1927) defended a pluralistic idealism by means of metaphysical arguments designed to show that space, time, and matter cannot possibly be real. Michael Oakeshott's *Experience and Its Modes* (Cambridge, U.K., 1933), unlike McTaggart's work, seems to have been completely unaffected by the realist and empiricist arguments so widely accepted at that time. Brand Blanshard's *The Nature of Thought* (2 vols., London, 1939), on the other hand, maintains a constant and detailed criticism of behaviorist and empiricist arguments. It is noteworthy that in none of these elegantly written idealist works is there any attempt to defend a theistic position. Indeed, *The Nature of Existence* concludes its discussion of God by saying that "there can be no being who is a God, or who is anything so resembling a God that the name would not be very deceptive" (Sec. 500).

IDEALIST SOCIAL THEORY

Most nineteenth-century and twentieth-century idealist philosophers were agreed that utilitarians and individual-

ists had a false view of what constitutes an individual person. They believed that since individuals are constituted by their relations to one another, the idea that society is an association of independently existing individuals is absurd. They thought, too, that it follows from this that freedom is something more positive than just being left alone by the government. Insofar as government is concerned with the common aims of individuals, it is not merely a constraint on them but a manifestation of their most rational purposes. Some idealist writers, therefore, saw no serious harm in Rousseau's claim that men can be forced to be free. T. H. Green was thus able to support temperance legislation on the ground that it enabled those protected by it to fulfill their abiding aims rather than their passing whims.

Even so, Green had no doubts about the ultimate reality of individual persons, whereas Bosanquet, in his *Philosophical Theory of the State* (London, 1899) argued that the state is the real individual and that individual persons are unreal by comparison with it. But Bosanquet did not think that this justified socialist control. On the contrary, he believed that if society is organic and individual, then its elements can cooperate apart from a centralized organ of control, the need for which presupposes that harmony has to be imposed upon something that is naturally unharmonious.

McTaggart was the one leading idealist who denied the relevance of metaphysics to social and political action. He was a Hegelian scholar who was in general agreement with Hegel's views, but he thought that Hegel was wrong in supposing that metaphysics could show that the state is more than a means to the good of the individuals who compose it. McTaggart concluded that "philosophy can give us very little, if any guidance in action.… Why should a Hegelian citizen be surprised that his belief as to the organic nature of the Absolute does not help him in deciding how to vote? Would a Hegelian engineer be reasonable in expecting that his belief that all matter is spirit should help him in planning a bridge?" (*Studies in Hegelian Cosmology*, Cambridge, U.K., 1901, p. 196).

SOME COMMENTS ON IDEALISM

ACT AND OBJECT. Moore, Russell, and other realist philosophers at the beginning of the twentieth century objected to idealism that its exponents failed to distinguish between the act of perception and the object of the act. It was rightly argued that the words *idea* and *sensation* were used vaguely and thus encouraged the confusion. According to the realist argument, colors and shapes are objects of the mind, whereas pains and feelings are states of mind, and what the idealists do is to say of the former that they are essentially mental, when this is true only of the latter. It may be questioned, however, whether the idealists were thus confused. Certainly Berkeley was not, since in the first of the *Three Dialogues between Hylas and Philonous* he himself made this objection only to reject it on the ground that the only acts of mind are acts of will, and in perceiving we are passive and do not exert acts of will.

In any case it is not easy to be sure that we can recognize or identify acts of perception. William James, for example, said he could distinguish no such thing (*Essays in Radical Empiricism*, New York, 1912), and Russell later took this view as well (*The Analysis of Mind*, London, 1921). Furthermore, even if the distinction is acceptable, what the object of perception is still remains to be determined. It is hard to maintain that what is immediately perceived is a physical object, since this seems to be inconsistent with the physiology of perception. If the immediate object is a sense datum, as Moore and Russell argued, then this suggests a representative theory of perception. But representative theories of perception are liable to the objection that they make our knowledge of physical objects problematical. If, on the other hand, sense data are not intended to play their part in a representative theory of perception but are meant to be all that can be perceived, then commonsense realism has been abandoned and Berkeley is vindicated. Apart from this, the very notion of a sense datum is dubious, since it is impossible to specify what a sense datum is without reference to physical objects. The distinction between act and object does not, therefore, lead to any effective arguments against idealism.

EXISTENCE APART FROM MIND. We have seen that Berkeley supported his immaterialist theory with the argument that nothing could exist apart from mind, since if we try to think of something existing unthought of we have to think of it, so that there is a contradiction in the very notion of thinking of something unthought of. Berkeley was by no means the only idealist who used this argument. It seems to have been accepted by Bradley, for example, when he wrote in Chapter 14 of *Appearance and Reality*:

> We perceive, on reflection, that to be real, or even barely to exist, must be to fall within sentience …. Find any piece of existence, take up anything that any one could possibly call a fact, or could in any sense assert to have being, and then judge if it does not consist in sentient expe-

> rience. Try to discover any sense in which you can still continue to speak of it, when all perception and feeling have been removed; or point out any fragment of its matter, any aspect of its being, which is not derived from and is not still relative to this source. When the experiment is made strictly, I can myself conceive of nothing else than the experienced.

This general line of argument came under attack in *The New Realism*, where the objection to it was that it falsely concludes that whatever is must be experienced from the evident tautology that whatever is experienced is experienced. From the fact that nothing can be experienced without being experienced it does not follow that everything must be experienced. Another way of stating this objection is to distinguish (*a*) it is impossible to-think-of-something-existing-unthought-of and (*b*) it is impossible to think of something-existing-unthought-of. Berkeley and Bradley are accused of denying the possibility of (*b*) because of the obvious impossibility of (*a*) (G. Dawes Hicks, *Berkeley*, London, 1932).

IDEALIST METAPHYSICS. Idealism involves the existence of some ultimate spiritual reality beyond what appears to common sense and ordinary sense experience. If it could be proved, therefore, that it does not make sense to speak of something that transcends sense experience, then idealism, like all other metaphysical systems, would be meaningless, as is claimed by logical positivism. Logical positivism, however, has been subjected to serious criticism and is by no means the chief alternative to idealism. It is linguistic philosophy, the philosophy that seeks to solve or to dissolve philosophical problems by showing that they arise out of linguistic misunderstandings, that today is the strongest opponent of idealism.

Moore's insistence on the act-object distinction was not, as we have seen, a successful mode of attack on idealism. But when he criticized Bradley for misunderstanding the logic of propositions in which something is said to be real, he was starting a sort of philosophizing that has proved most inhospitable to idealist theories. Moore saw that when Bradley said that time is unreal he had no wish to deny such things as that people are sometimes late for their trains. Yet if there were no temporal facts, there would be no trains and no people to catch or to lose them. Moore felt that something had gone wrong with Bradley's argument, and he tried to locate the fault. He thought that Bradley believed that even though time is unreal, if it can be thought of then it must have some sort of existence. Moore thought he could show that this belief is groundless and arises from a misunderstanding of what is being said when something is said to be real. But Moore also came to believe that we know for certain such things as that there are trains and people and that in consequence we are justified in denying out of hand those philosophical views that would require trains and people and space and time and matter to be mere appearances or not to be real at all. It was through his attempts to understand the prevailing idealist metaphysics that Moore came to adopt his philosophy of common sense. This philosophy and the linguistic philosophy that grew out of it regard our prephilosophical beliefs and concepts as in a certain sense unassailable. If this view is correct, then idealism is based on misunderstandings. If it is not correct, then the idealist criticisms of our prephilosophical beliefs have to be taken seriously.

IDEALISM AND THE NATURE OF THOUGHT. The idealist movement is important in the history of philosophy quite apart from the success or failure of idealist metaphysics. Idealists have insisted from Kant onward that thinking is an activity. This view of thinking was Kant's particular contribution to philosophy and is opposed to the Cartesian theory of knowledge. According to Descartes knowledge consists in the intuition of clear and distinct natures. What keeps us from obtaining knowledge, Descartes held, is the existence of prejudices that keep us from getting face to face with the ultimate clarities; once the prejudices are removed, the world shows itself as it really is. On this view the human mind is like a mirror that reflects what is there when it has been wiped clean. According to Kant, however, the mind approaches the world with concepts and presuppositions of its own. It does not reflect the world but tries to understand and interpret it. The activity of synthesizing is an activity of interpreting, and this can be done only by means of concepts that we already possess. According to Descartes we must wipe the mirror clean to be ready for undistorted visions; inquiry ends in revelation. According to Kant we gain knowledge as we improve and test our theories. Apart from natural science, nature is nothing but what men have to contend with in their daily concerns. This view was metaphysically elaborated by Kant's idealist successors, but they did not lose sight of an important implication of it that Kant had seen, the implication that the pursuit of knowledge was a spontaneous activity. They argued that knowledge and freedom go together and that therefore determinism and reductive materialism cannot be true. This would appear to be the essence of the idealist argument.

See also Absolute, The; Coherence Theory of Truth; Dialectical Materialism; Hegelianism; Ideas; Neo-

Kantianism; New England Transcendentalism; Panpsychism; Personalism; Realism; Relations, Internal and External; Solipsism.

Bibliography

A number of idealist texts are available in *The Idealist Tradition: From Berkeley to Blanshard*, edited by A. C. Ewing (Glencoe, IL: Free Press, 1957).

The meaning of the term *idealism* is discussed in Norman Kemp Smith, *Prolegomena to an Idealist Theory of Knowledge* (London: Macmillan, 1924).

On the history of idealism, see G. Watts Cunningham, *The Idealistic Argument in Recent British and American Philosophy* (New York: Century, 1933); Nikolai Hartmann, *Die Philosophie des Deutschen Idealismus*, 2 vols., Vol. I: *Fichte, Schelling und die Romantik* (Berlin and Leipzig, 1923), Vol. II: *Hegel* (Berlin and Leipzig, 1929), 2nd ed., 1 vol. (Berlin, 1960); J. H. Muirhead, *The Platonic Tradition in Anglo-Saxon Philosophy. Studies in the History of Idealism in England and America* (London, 1931); and Josiah Royce, *Lectures on Modern Idealism* (New Haven, CT: Yale University Press, 1919).

On idealist social theory, see A. J. M. Milne, *The Social Philosophy of English Idealism* (London: Allen and Unwin, 1962), which, in spite of its title, has a chapter on Royce.

For an analysis and criticism of idealism, see A. C. Ewing, *Idealism: A Critical Survey* (London: Methuen, 1934).

OTHER RECOMMENDED TITLES

Ayers, Michael. "What Is Realism?" *Proceedings of the Aristotelian Society*, supplement 75 (2001): 91–110.

Bradley, F. H. *Appearance and Reality*. Oxford: Oxford University Press, 1930.

Foster, J. *The Case for Idealism*. London: Routledge and Kegan Paul, 1982.

Franks, Paul. "From Kant to Post-Kantian Idealism: German Idealism." *Proceedings of the Aristotelian Society*, supplement 76 (2002): 229–246.

Kulp, Christopher, ed. *Realism/Antirealism and Epistemology*. Lanham, MD: Rowman & Littlefield, 1997.

Lovejoy, Arthur O. *The Revolt against Dualism*. La Salle, IL: Open Court, 1930.

McDermid, Douglas James. "The World as Representation: Schopenhauer's Arguments for Transcendental Idealism." *British Journal for the History of Philosophy* 11 (2003): 57–87.

Neill, E. "Evolutionary Theory and British Idealism: The Case of David George Ritchie." *History of European Ideas* 29 (2003): 313–338.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Royce, J. *The Religious Aspect of Philosophy*. Gloucester, MA: Peter Smith, 1965.

Sprigge, T. L. S. *The Vindication of Absolute Idealism*. Edinburgh: Edinburgh University Press, 1983.

Van Cleve, James. "Time, Idealism, and the Identity of Indiscernibles." *Nous*, supplement 16 (2002): 379–393.

H. B. Acton (1967)
Bibliography updated by Benjamin Fiedor (2005)

IDEAL OBSERVER THEORIES OF ETHICS

The ideal observer theory (IOT) offers an account of the truth/objectivity of moral judgments in terms of the approval or disapproval of an ideal observer. The theory receives explicit treatment by Adam Smith and Henry Sidgwick; Roderick Firth is the most well-known proponent of the theory in the twentieth century.

There are two versions of the theory. On one account the IOT is an analysis of what (some, all) moral judgments mean: A judgment that some act (or event or state of character) is good may be analyzed in terms of that act being approved of by an ideal observer (IO); some act is wrong if it would be disapproved of by an IO. Most, but not all, such accounts conceive of the IO in hypothetical terms, leaving open the question of whether there actually is an IO. The traits of the IO vary between ethical theories, but they often include impartiality, knowledge of all of what may be called nonmoral facts (facts that may be conceived of and known without ipso facto knowing the moral status of the fact), and an affective awareness of the points of view of all involved parties. The reason for employing a term such as *nonmoral facts* is to avoid an explicit circularity, for the theory must be more informative than claiming that some act is morally right if and only if it would be approved of by a being who is omniscient with respect to all moral facts.

While strong versions of the IO theory offer an analysis of what moral rightness and wrongness mean, moderate proposals hold that the IO point of view amounts to an analysis of the moral point of view; that is, the point of view from which ideal moral judgments are made. On this account, what it means for persons to carry out an inquiry into the moral status of some act is to engage in an inquiry aimed at achieving impartiality, knowledge of the relevant nonmoral facts, and an affective awareness of the points of view of all involved parties. Arguably, these conditions might be both necessary and sufficient for moral inquiry, and yet the IOT would not amount to an analysis of what it means for some act to be right.

Both versions are subject to objections. Against both accounts philosophers have questioned the feasibility and desirability of impartiality. IO accounts that appeal to the hypothetical responses of an IO face a problem in terms of moral psychology; someone making a moral judgment about some act need have no interest in the responses of some other, hypothetical observer. There is also the recurrent charge against both versions that the theory is circu-

lar. They build into the concept of an IO the notion that the observer is in fact ideal; being impartial, for example, is a positive moral ideal. If so, the theories presuppose a moral ideal and so cannot be used to analyze what it is to be morally ideal. Some argue that neither account is able to avoid conflict between IOs or those seeking the moral point of view. There is also the charge that both accounts fail because it is coherent to claim that an IO or one who achieves the ideal moral point of view may get matters wrong.

The first two objections may be played against each other. Evidence that philosophers disagree about the moral desirability of impartiality is evidence that impartiality is not an obvious moral ideal. If the case against the moral desirability of impartiality is successful, the IO theory will need amending to allow for specific, partial duties and goods. Some versions of the IO theory have been articulated that accommodate the thesis that IOs disagree (Thomas Carson 1984), while others argue that there is no reason to suppose that there would be disagreement (Charles Taliaferro 1988).

There have been several replies to the charge that IOs or those taking the ideal moral point of view may be wrong. Some link the IO theory with a divine command theory according to which moral rightness and wrongness is constitutive of an actual IO's God's approval and disapproval. The apparent coherence of there being something approved of by God that is morally wrong is accommodated the way in which some philosophers accommodate the apparent possibility that one might have water without H20.

On behalf of using the second, modest form of the IO theory, it has also been charged that the following state of affairs is incoherent. A person morally disapproves of some act but she believes that if she were an IO she would approve of the act. According to the moderate version of the theory, the person disapproves of the act but simultaneously believes that she would reverse her view if she were actually impartial, knew more of the relevant nonmoral facts, and had an affective understanding of the points of view of those affected. Arguably, when we disapprove of some act morally, we often allow the *possibility* that we may not be impartial, we may be ignorant of the nonmoral facts, and we may lack an awareness of the feelings of those involved, but it would be peculiar for a person to disapprove of an act while believing that one is *actually* mistaken about the nonmoral facts, and so on.

See also Ethical Subjectivism; Metaethics; Objectivity in Ethics; Sidgwick, Henry; Smith, Adam.

Bibliography

Anderson, Pamela. "What's Wrong with the God's Eye Point of View: A Constructive Feminist Critique of the Ideal Observer Theory." In *Faith and Philosophical Analysis*, edited by H. A. Harris and C. Insole. Burlington, VT: Ashgate, 2005: 85-99.

Carson, Thomas. *The Status of Morality*. Boston, MA: Reidel, 1984.

Carson, Thomas. *Value and the Good Life*. Notre Dame: University of Notre Dame Press, 2000.

Firth, Roderick. "Ethical Absolutism and the Ideal Observer." *Philosophy and Phenomenological Research* 12 (1952): 317–345.

Frankena, William. *Ethics*. 2nd ed. Englewood Cliffs: Prentice-Hall, 1973.

Hare, R. M. *The Language of Morals*. Oxford: Oxford University Press, 1952.

Smith, Adam. *The Theory of Moral Sentiments*, edited by K. Haakonssen. Cambridge, U.K.: Cambridge University Press, 2002.

Taliaferro, Charles. "Relativizing the Ideal Observer Theory." *Philosophy and Phenomenlogical Research* 49 (1988) 123–138.

Zagzebski, Linda. Divine Motivation Theory. Cambridge, U.K.: Cambridge University Press, 2004.

Charles Taliaferro (2005)

IDEAS

The word "idea" is a transliteration of a Greek word of which the root meaning is "see." In classical Greek it never lost the possible meaning "visual aspect"; thus Plato writes of a person as being "very beautiful in idea," meaning "beautiful in visual aspect" or "good-looking" (*Protagoras* 315E). Very often visual shape is primarily involved, as when Plato refers to the "idea of the earth," meaning "the visible shape of the earth" (*Phaedo* 108D). The transferred sense of "type" or "kind" springs quite naturally from this use. Thus Thucydides writes of "many ideas [kinds] of warfare" (*Histories* I, 109).

In Plato's more technical use, the Ideas or Forms are always spoken of as (1) the objects of intelligence, in contrast with the objects of perception; (2) things that truly *are,* in contrast with changing objects of perception, which are in a state of *becoming*; (3) eternal, in contrast with the perishable world of change. But there are at least two irreconcilable strands in Plato's thought about Ideas. Sometimes he seems to have thought of Ideas much as later philosophers have thought of universals, as when he says that "we are accustomed to posit a single form for each group of many things to which we give the same name" (*Republic* 596A); consistent with this he speaks sometimes of the *presence* of the form in the particular or

of the particulars as participating in the form (*Phaedo* 100D). But sometimes Plato writes as if his Forms were, rather, perfect exemplars or paradigms of which the sensible world is an imperfect copy or imitation; thus in the *Parmenides* Socrates says that the Forms are "as it were paradigms" and that "other things are like them and are copies of them" (132D). When the Forms are thus described, we also find Plato insisting that they are "separate," a doctrine in conflict with the language of "presence" and "participation" noted above. It is plausible to suggest that there is here a tension between the theory of universals and the theory of resemblance to standard objects as explanations of common names.

But it is the theory of Ideas as separate and eternal paradigms that appears in the *Timaeus,* the dialogue that had incomparably the greatest influence on later antiquity and the Middle Ages; there the divine demiurge is depicted as forming the world on the pattern of the eternal Forms. It will therefore be the aspect of Forms as paradigms, perfect exemplars, blueprints, particularly as patterns used by a divine agent in creation, which will be important in the development of the philosophical notion of an idea.

In the *Timaeus* the Forms, or Ideas, are eternal and independent objects to which the demiurge looks as patterns. But one of the most important and early modifications of this Platonic view is the religious conception of the Ideas as the thoughts of God. This is the view of Plotinus (*Ennead* III, 9, i), of Philo (*De Opificio Mundi* 4) and of Augustine (*De Diversis Quaestionibus LXXXIII,* Question 46). Clement of Alexandria simply defines an idea as a "thought of God" (*Stromateis* V, iii, 16.3). The ideas are still perfect and eternal exemplars, but now they are in the mind of God.

It is not a very long step to extend the term *idea* to cover patterns, blueprints, or plans in anybody's mind, not only in God's. Thus we find Thomas Aquinas saying that "the word 'idea' signifies a certain form thought of by an agent in the likeness of which he intends to produce an external work" (*Quaestiones Quodlibetales* IV, I, lc); similarly Goclenius says that "in general an idea is a form or exemplar of a thing with an eye on which a workman makes what he has planned in his mind" (*Lexicon Philosophicum* 208A).

When the word *idea* was taken over into the French and English vernacular by learned men in the sixteenth century, there were thus two elements in the concept of an idea—that it was an exemplar or pattern and that it was a thought in a mind. Using the pattern element alone, François Rabelais could speak of Pantagruel as being the "idea and exemplar of every joyous perfection" (*Pantagruel,* Book III, Ch. 51); but a pattern and its copy could be easily muddled so that Rabelais also could say, "En leur mariage semble reluire quelque idée et représentation des joyes de paradis" ("In their marriage some idea and representation of the joys of paradise seems to be reflected"; *Pantagruel,* Book III, Ch. 10). When the other, mental element is introduced, the meaning of "idea" quickly becomes "mental representation"; this is a very common meaning in sixteenth-century French and English, and the phrase of Michel Eyquem de Montaigne, "Ayant par longue conversation planté vivement dans son âme unc générale idée de celle de Plutarque" ("Having by long communion vividly emplanted in his own a general idea of the mind of Plutarch"; *Essays,* II, 4), could be paralleled many times.

DESCARTES

Thus when René Descartes first began to write, the meaning "image or representation," often but not necessarily "in the mind," was already well known in the vernacular. In spite of the fact that Descartes is usually credited with the invention of the non-Platonic use of the term, we find him at first following this vernacular use. In his first Latin work, the *Regulae,* the word *idea* appears infrequently, but Descartes always uses it to mean an image or representation; when he first introduces it in the *Meditations,* he at once says, "Quelques-unes [de mes pensées] sont comme les images des choses, et c'est à celles-là seules que convient proprement le nom d'idée" ("Some [of my thoughts] are like images of things, and it is to these alone that the name 'idea' properly belongs"). It is only under the pressure of philosophical difficulties that he extends the term *idea* to cover the unimaginable, for which Thomas Hobbes duly reprimanded him: "When I think of a man I represent to myself an idea or image composed of colour and shape. … of God we have no image or idea" (*The Third Set of Objections,* Objection 5). There is therefore no need for any explanation why the word *idea* tends to mean "mental image" to seventeenth-century philosophers; this is what the word ordinarily meant in their time.

What does need explanation is why, if Descartes found the word *idea* to mean "properly" only "an image of a thing," he and other philosophers could use "having an idea" as a proper designation of all thought and could define an idea as the object of a mind when it thinks, in a liberal sense of "think" that includes sense perception. Part of the explanation is to be found in the representative theory of perception, held in some form by all the

philosophers of the period; there was no extension of meaning in using "idea" of sense perception because it was believed that what was directly perceived was not things, but images of things—the images caused by and more or less resembling the things themselves. Another part of the explanation is the "image theory" of thinking: To think of something is or includes having either a mental image of that thing or, as some believed, a physical image on that part of the brain termed the "corporeal phantasy." Such a view was in the air at the beginning of the seventeenth century and was accepted by Pierre Gassendi and Hobbes without reservation. Descartes never doubted that many of our thoughts are images of things; his extension of the term arises from his gradual realization of the inadequacy of the image theory to account for all our thought even while he persevered in the use of its terminology. His use of the term to denote any object of thought became the standard one in philosophy, via such influential writings as the *Port-Royal Logic* and John Locke's *Essay.* Only a few scholastically trained philosophers, such as Immanuel Kant, have stood out for a more Platonic usage; thus Kant in the *Critique of Pure Reason* holds to the terminology of the transcendental ideas of reason to which no corresponding object can be perceptually given, as distinct from the concepts of the understanding ("Transcendental Dialectic," I, 2).

Most of the confusions in the "way of ideas" arise at least in part from the use of the term *idea* to cover both the representative percept and the object of conceptual thought. This can be illustrated in terms of the doctrines of innate ideas, concrete and abstract ideas, and simple and complex ideas.

INNATE IDEAS

The mature Gottfried Wilhelm Leibniz always maintained, and Descartes sometimes maintained, that all our ideas are innate. Thus Leibniz said that "all the thoughts and acts of the soul come from its own depths, with no possibility of their being given to it by the senses" (*New Essays concerning Human Understanding,* Book I, Ch. i, 1). But this is a theory of *perception,* as is made clear by Descartes in his *Notes Directed against a Certain Program,* his defense of the view:

> Nothing comes from external objects to our mind through the organs of sense save certain corporeal motions … but not even these motions, and the configurations to which they give rise, are conceived by us as they occur in the sense-organs. … Whence it follows that the very ideas of motions and configurations are innate in us. So much more must the ideas of pain, colours, sounds, and the like be innate, so that our mind can, on the occasion of certain corporeal motions, display them to itself; for they have no similarity to the corporeal motions.

There is nothing here from which Locke would dissent, except verbally; no wonder that Leibniz said, in the preface to his *New Essays,* "I am led to believe that at bottom his [Locke's] view upon this point is not different from mine." The true controversy with Locke is, rather, exhibited by Descartes's view of concepts, expressed in the same terms as and never distinguished from the perceptual theory by philosophers of the time. According to this theory, some ideas are innate—for example, those of God, mind, body; others are adventitious—one's ordinary idea of the sun; still others are made (*factae*) or factitious—the ideas of the sun astronomers construct by reasoning. It is those innate and factitious ideas—which Descartes could as little say were *occasioned* by "corporeal motions" as Locke could say they were *caused* by "corporeal motions"—which raised a still-pressing difficulty.

ABSTRACT AND CONCRETE IDEAS

The distinction between abstract and concrete ideas is virtually the distinction, misleadingly put, of concepts and percepts. The doctrine of abstract ideas was held by the Cartesians, and the best statement of it is to be found in *Port-Royal Logic,* Book I, Ch. 6. To have an abstract idea is to think of some feature or features of the perceptible without attending to other features that it has and that are as inseparable from it (except in thought) as are the length and breadth of a road. Locke took over the Port-Royal account of what abstraction was without change, even echoing its language, but tried to give a more thorough account of what it involved. He tried to give an account of abstraction in terms of a doctrine of simple and complex ideas, but by failing to distinguish thought and perception, he gives two incompatible accounts of this distinction. In Book III of the *Essay* he tells us that all ideas save those denoted by proper nouns are abstract. Of these some are indefinable; they are simple ideas. Others are definable; these are complex ideas. "The ideas first in the mind, it is evident, are those of particular things" (*Essay,* Book IV, Ch. vii, Sec. 9)—that is, we first perceive particular things; in thinking about them, we may come to form some very general ideas by omitting less interesting features and concentrating on those common to a whole group, which taken together form a complex idea; by further abstraction we can get to less and less complex ideas. It is clear that according to this view simple ideas

involve the highest degree of abstraction. But in Book II we are told that the simple ideas enter the mind in perception simple and unmixed; they are objects of perception. Thus a theoretical analysis of the construction of concepts is inextricably confused with an atomistic doctrine of perception. If simple ideas are objects of perception and complex ideas are formed from them, then all abstract ideas ought to be imaginable, and George Berkeley's famous sneers about the abstract idea of a triangle have some justification. But neither Berkeley nor David Hume could emancipate himself from the basic confusion; this is true of Hume in spite of his famous distinction between ideas and impressions.

Thus the classical theory of ideas, which had held virtually undisputed sway in the seventeenth and eighteenth centuries among rationalists and empiricists, was based on the theories of representative perception and image-thinking. To continue to use the terminology after these theories had been abandoned as inadequate could lead only to confusion and a skepticism which, consistently developed, would be even more extreme than Hume's.

Reasonably, therefore, outside the empiricist tradition the term *idea,* as employed in the seventeenth and eighteenth centuries, soon ceased to appear in philosophical writings. Kant's representations have, indeed, some resemblance to ideas of sensation, and the thing-in-itself plays a part somewhat analogous to Locke's substratum. But there are important differences, and his concepts of the understanding are very far from being copies of representations. He does, indeed, use the term *idea* technically, but with a yet further removed significance. In the *Critique of Pure Reason,* he says: "I understand by 'Idea' a necessary concept of reason to which no corresponding object can be given in sensation" ("Transcendental Dialectic," I, 2). These ideas, such as that of the absolute unity of the subject, have, Kant holds, a valid regulative employment, but if we try to apply them to experience we become involved in metaphysical paralogisms. Insofar as the term continued to be used in Continental philosophy it was used, as by G. W. F. Hegel, in senses far removed from that in pre-Kantian philosophy.

But in British philosophy the terminology did not die an easy death. The empiricists could not abandon it, especially in their philosophical psychology in which the doctrine of the association of ideas continued to play the dominant role given to it by Hume. It was largely F. H. Bradley's polemic against psychologistic logic that finally led to the abandonment of the "way of ideas." But even Bradley, in the first chapter of his *Logic,* which is a locus classicus for the attack on psychologism, showed that he had not completely emancipated himself. He could still write that "the idea, in the sense of mental-image, is a sign of the idea in the sense of meaning," and added "without ideas no judgment," though in a note of 1922 he rejected these statements. By 1922 his own work and that of G. E. Moore had led to the elimination of the term *idea* from British philosophy, except as a part of nontechnical idiom.

In the United States, also, the term *idea* continued to have considerable currency. It was a key term in the pragmatism of Charles Sanders Peirce, William James, and John Dewey, reflecting the fact that they, too, were heirs to the empiricist tradition though not to Humean skepticism. They avoided this skepticism in part by wholly abandoning the image theory of thinking with which the terminology of ideas was traditionally linked. In Dewey's instrumentalism, ideas became tools for directing our activities, responses to sensation rather than sensations. They were tied to practical transactions. In calling the idea a law of action, Dewey reminds us rather of the definition given by Thomas Aquinas quoted earlier in this article than of the traditional empiricist position. But Peirce could still think, like Bradley, of ideas as psychological entities, as well as in terms of pragmatic epistemology; and in James also the pragmatic doctrine that our ideas of an object have to be explained in terms of the sensations we expect from it and the reactions we make toward it had still not been completely disentangled from a more traditional empiricism.

See also Augustine, St.; Berkeley, George; Bradley, Francis Herbert; Clement of Alexandria; Concepts; Descartes, René; Dewey, John; Empiricism; Gassendi, Pierre; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Hume, David; Innate Ideas; James, William; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Montaigne, Michel Eyquem de; Moore, George Edward; Peirce, Charles Sanders; Philo Judaeus; Plato; Plotinus; Plutarch of Chaeronea; Psychologism; Rabelais, François; Socrates; Thinking; Thomas Aquinas, St.; Thucydides; Universals, A Historical Survey.

Bibliography

For critical and historical discussions, see Albert G. A. Balz, *Idea and Essence in the Philosophies of Hobbes and Spinoza* (New York: Columbia University Press, 1918); M. H. Carré, *Realists and Nominalists* (Oxford: Oxford University Press, 1946); E. Garin, *La théorie de l'idée suivant l'école thomiste* (Paris: Desclèe, de Brouwer, 1932); Paul Natorp, *Platons*

Ideenlehre (Leipzig, 1930); W. D. Ross, *Plato's Theory of Ideas* (Oxford: Clarendon Press, 1951); G. J. Warnock, *Berkeley* (Harmondsworth, U.K., 1953); John W. Yolton, *John Locke and the Way of Ideas* (London: Oxford University Press, 1956).

For the nontechnical use of the term *idea* (*idée*) in the sixteenth and seventeenth centuries, see *The New English Dictionary* and Huguet's *Dictionnaire de la langue française du seizième siècle.*

See also the bibliographies in these entries: Concept; Innate Ideas; Rationalism; Thinking.

J. O. Urmson (1967)

IDENTITY

The word *is* is multiply ambiguous. When it can be expanded to read "is the same thing as," or "is identical with," or (in numerical contexts) "is equal to," it expresses the relation of identity. The simplest identity statements contain the "is" of identity flanked by singular terms, either names or definite descriptions: "Samuel Clemens is Mark Twain"; "The U.S. president in 1996 was Bill Clinton"; "Four is the sum of two and two." A more complex identity statement might, for example, combine the "is" of identity with quantifiers: "Every even number is the sum of two primes."

Identity, on its face, is simple and unproblematic: It is that relation that everything bears to itself and to nothing else. Yet discussions of identity in contemporary philosophical logic and metaphysics are brimming with controversy. From where does this controversy arise? Some of it is not genuine, being based on confusion; and some of it, though genuine, is not genuinely about identity. However, a residue of controversy survives, owing to the view, perpetrated by Peter Geach, that identity statements are meaningless unless relativized, that there is no absolute relation of identity.

SOURCES OF CONFUSION

One source of confusion is the ambiguity of "identical" in English. We do sometimes say that two things are identical, as when we speak of identical twins, or say that some coat is identical with some other. This is qualitative identity: Things are qualitatively identical if they resemble one another sufficiently in relevant qualitative respects. Numerical identity is different: Two things, no matter how closely they resemble one another, are never numerically identical. Numerical identity is the topic of this article.

A second source of confusion is English grammar, which allows, for example, "Clemens is identical with Twain" to be rewritten equivalently as "Clemens and Twain are identical" or as "they are identical." But then it seems that two persons (or two *somethings*) are being said to be identical, which is absurd. A general response is familiar from other cases: Surface grammar often misrepresents the underlying logic: One must beware inferring logical from grammatical form. More specifically, it can be verified that plural noun phrases in English do not, in all contexts, entail or presuppose reference to a plurality.

A third source of confusion is Gottlob Frege's puzzle of informative identity statements, sometimes introduced by the following argument. To say of something that it is identical with itself is trivial, to say of something that it is identical with something else is false; therefore, identity statements are all either trivial or false, and there can be no point in asserting them. This conclusion is manifestly incorrect: Identity statements are often both true and informative, as witness, "the capital of Honduras is Tegucigalpa." The puzzle is to say where the argument goes wrong.

One response rejects the second premise by taking identity to be a relation between names or descriptions rather than between the objects named or described: Identity is then the relation of codesignation, the relation that holds between singular terms whenever those terms designate the same object. That would indeed allow identity statements to be both true and informative. But the response is not viable, for many reasons. For one, it fails to account for uses of identity that do not involve singular terms, such as: "Everything is identical with itself." For another, it fails to allow identity statements between different singular terms to be uninformative, as they are when the singular terms are synonymous. For another, it fails to provide a unified solution to analogous puzzles of informativeness, such as how "the capital of Honduras is in Honduras" and "Tegucigalpa is in Honduras" can differ in informativeness, even though both ascribe the same property to the same thing.

A better response is due to Frege. Identity is a relation between objects; a simple identity statement is true just in case the objects referred to by the singular terms stand in that relation. But singular terms have sense in addition to reference; a true identity statement is informative just in case its singular terms differ in sense. (Just what is included in the sense of a singular term varies from theory to theory; but note that senses must be rich enough to allow codesignative proper names—such as "Mark Twain" and "Samuel Clemens"—to differ in

sense.) Now the puzzle may be solved by rejecting the argument's first premise: One can say informatively of an object that it is identical with itself by referring to the object twice over, using singular terms that differ in sense. That is how "The capital of Honduras is Tegucigalpa" manages to be both true and informative. Identity statements are useful in ordinary language because we often refer to the same object from different points of view, using terms with different senses. (Frege's statement of the puzzle, and his solution, is in Frege 1892; see also Kripke 1980, Salmon 1986.)

THE LOGIC OF IDENTITY: LEIBNIZ'S LAW

Relations may be classified according to their general, logical characteristics. The logical characteristics of the identity relation are easily enumerated. First, as already noted, identity is reflexive: Every object is identical with itself. Second, identity is symmetric: If an object *x* is identical with an object *y*, then *y* is identical with *x*. Third, identity is transitive: If an object *x* is identical with an object *y*, and *y* is identical with an object *z*, then *x* is identical with *z*. A relation that is reflexive, symmetric, and transitive is called an equivalence relation. Finally, identity is the strongest equivalence relation, entailing all other equivalence relations: If an object *x* is identical with an object *y*, then *x* bears *R* to *y*, for every equivalence relation *R*. Since being the strongest equivalence relation (or, equivalently, being the strongest reflexive relation) uniquely characterizes identity in purely logical terms, identity may properly be classified as a logical relation and the theory of identity as a branch of logic.

All of the logical characteristics of identity can be derived from a single principle, sometimes called Leibniz's law: An object *x* is identical with an object *y* if and only if every property of *x* is a property of *y* and vice versa. Leibniz's law is a biconditional and thus the conjunction of two conditionals, one giving a necessary, the other a sufficient, condition for identity to hold. Say that an object *x* is indiscernible from an object *y* just in case every property of *x* is a property of *y* and vice versa. The half of Leibniz's law that gives a necessary condition proclaims the indiscernibility of identicals: If *x* is identical with *y*, then *x* is indiscernible from *y*. This principle is useful for establishing nonidentity: To show that *x* is not identical with *y*, it suffices to find a property had by *x* but not by *y* or vice versa. Most famously, perhaps, the principle has been used to argue that persons are not identical with their bodies. The half of Leibniz's law that gives a sufficient condition proclaims the identity of indiscernibles: If *x* is indiscernible from *y*, then *x* is identical with *y* (more on this below).

(Note that Leibniz's law is stated within second-order logic: It involves quantification over properties. The first-order theory of identity substitutes for Leibniz's law an axiom schema containing, for each [monadic] predicate of the language, an axiom stating: If *x* is identical with *y*, then *x* satisfies the predicate if and only if *y* satisfies the predicate. This schema, together with an axiom of reflexivity, entails the entire first-order theory of identity. The first-order theory is weaker than the full second-order theory; in particular, no logically sufficient condition for identity is expressible within first-order logic.)

The indiscernibility of identicals is beyond dispute: If *x* and *y* are identical, then there is only one thing; how can that one thing both have and not have some property? Nonetheless, the principle has been disputed. Consider the following attempt at a counterexample (discussed in Quine, 1953). It is true that Giorgione was so called because of his size, let us suppose, and that Giorgione is identical with Barbarelli; yet, apparently contrary to the principle, it is not true that Barbarelli was so called because of his size. But to see this as a violation of the indiscernibility of identicals, one would have to hold that the predicate "is so called because of his size" expresses some genuine property of objects and expresses the same property when applied to "Giorgione" as when applied to "Barbarelli." On the contrary, when considered in isolation the predicate expresses no property at all but rather a relation between objects and names. When applied to "Giorgione" it expresses the property was-called-Giorgione-because-of-his-size; and that property is true of Barbarelli, in accord with the indiscernibility of identicals. Other attempts at counterexamples are more subtle than this: But all seem to involve naively reading subject–predicate sentences as simple property-to-object attributions. (For examples involving modality see Cartwright 1971, Quine 1953).

IDENTITY OF INDISCERNIBLES

The other half of Leibniz's law proclaims the identity of indiscernibles; but now one must be careful just what "indiscernible" means. If indiscernibles have *all* of their properties in common, where properties are conceived abundantly, then the identity of indiscernibles is trivially true. For, on an abundant conception of property, for any object *y* there is the property is-identical-with-*y*. Now suppose that *x* is indiscernible from *y*. Then, since *y* has the property is-identical-with-*y*, *x* must have this property too; that is, *x* is identical with *y*, as was to be shown.

If we interpret "indiscernible" instead in terms of properties more sparsely conceived, for example, as "indiscernible in all qualitative respects," then we arrive at a substantial metaphysical principle, the identity of qualitative indiscernibles; the trivial "proof" above is blocked because properties such as is-identical-with-*a* (where "*a*" names some object) are not (or, at any rate, are not trivially) qualitative. There are different versions of the principle, however, corresponding to different interpretations of "qualitatively indiscernible"; and for each version one might ask whether the principle is logically necessary, is contingently true, or neither. Let us consider three versions.

According to the strongest (and least plausible) version, objects that share all of their intrinsic qualitative properties—intrinsic duplicates—are identical. This principle seems to be false even at the actual world: According to current physics, distinct elementary particles of the same kind—for example, distinct electrons—have all of their intrinsic properties (charge, mass, etc.) in common.

According to the second (and most familiar) version, objects that share all of their intrinsic and extrinsic qualitative properties—absolute indiscernibles—are identical. Absolute indiscernibles must not only be intrinsic duplicates, they must be exactly similarly situated with respect to all of their surroundings. But, surely it is at least possible that there be distinct yet absolutely indiscernible objects; that is, the principle is not necessarily true. For, to take the standard counterexample (from Black 1952), it is logically possible that the world contains nothing but two perfectly round globes, exactly similar down to their smallest parts and separated, say, by one meter. The globes share all of their intrinsic qualitative properties, having the same mass, shape, and so on. And the globes share all of their extrinsic qualitative properties—for example, each is one meter from a globe of a certain mass, shape, and so on. (Note that properties that would only be expressible using names for the globes, such as is-one-meter-from-globe$_1$, are not qualitative). In short, the globes are absolutely indiscernible; yet they are two, not one.

A defender of the identity of absolute indiscernibles might simply deny that there is any such possibility; but there is a substantial cost. The claim that it is logically possible that there be nothing but two absolutely indiscernible globes can be backed up by a subsidiary argument (Adams 1979). Surely, there could be nothing but two almost indiscernible globes, differing, say, only in the placement of a single atom. To hold that that atom could not have been shifted in a certain way (because, if it had, there would have been two absolutely indiscernible globes), but that any other atom could have been shifted in that way, would amount to an implausibly inegalitarian approach to what is and is not possible.

Perhaps an even weaker version of the principle should be considered: Objects that share all of their qualitative properties, and stand in the same qualitative relations to any given object—relative indiscernibles—are identical. (On absolute vs. relative indiscernibility, see Quine 1960.) The possibility just considered of the two globes is not a counterexample to the necessity of this version: The globes are discerned by spatial relations; each globe is one meter from the other globe but not one meter from itself. A counterexample, however, is not far to seek. Consider the possibility that there be nothing but two absolutely indiscernible globes standing in no spatial relation (or other qualitative external relation) to one another, two absolutely indiscernible "island universes." (This possibility can be motivated, too, by first considering "almost" island universes, connected, say, by a single "wormhole.") Such globes would be relatively, as well as absolutely, indiscernible; they stand in no relations that could serve to discern them. So even this weakest version of the identity of qualitative indiscernibles seems not to be a necessary truth. (Indeed, it may not be contingently true: So-called identical particles in quantum mechanics are arguably distinct but absolutely and relatively indiscernible.)

IS IDENTITY DEFINABLE?

Identity has been characterized many times over. Do any of these characterizations provide a (noncircular) definition of the identity relation? Can identity be understood in terms not involving identity? Our initial characterization—that everything is identical with itself and with nothing else—clearly will not do as a definition: To be "else" is to be other, that is, nonidentical. Moreover, the characterization of identity as the strongest equivalence relation fares no better: Identity characterized by quantifying over all relations, identity included.

Leibniz's law gives a necessary and sufficient condition for identity by quantifying instead over properties. But among the quantified properties are haecceities, properties of being identical with some given object. The question whether an object *x* shares with an object *y* the property of being identical with *y* is just the question whether *x* is identical with *y*; the purported definition takes one around in a circle. Similarly defective is the oft-heard definition "*x* is identical with *y* if and only if *x* and

y belong to the same classes." The question whether *x,* like *y,* belongs to the class whose only member is *y* is just the question whether *x* is identical with *y.*

What if some version of the identity of qualitative indiscernibles were necessarily true (contrary to what was argued above)? That would indeed provide a noncircular criterion for the identity of *objects.* But the identity or distinctness of qualitative properties (and relations) would remain undefined. Indeed, any purported definition of identity would have to quantify over some sort of entity; the definition could not be understood without a prior understanding of the identity and distinctness of the entities quantified over. We must conclude, then, that identity, at least as applied to the most basic entities, must be taken as primitive and unanalyzable; there is no fully general (noncircular) definition of identity.

Questions remain, some of which might seem to pose problems for the classical conception of identity. We shall see, however, that in each case replies exist that leave classical identity unscathed. (Each of the issues raised below is discussed in Lewis 1993.)

PARTIAL IDENTITY

Classical identity is all or nothing; it never comes in degrees. Yet, when objects overlap, we may say they are "partially identical, partially distinct": And when objects extensively overlap, we may say they are "almost identical." Do we have here a challenge to classical identity? No, we have an ambiguity: Identity, in the sense that admits of degrees, is simply overlap; identity, in the classical sense, is equivalent to the extreme case of total overlap. The two notions of identity are not in conflict; they fit together as well as you please.

VAGUE IDENTITY

Classical identity is determinate and admits of no borderline cases. That is not to say that identity statements cannot be vague or indeterminate in truth-value. If I say "that cloud in the sky is identical with *A*," where "*A*" names some precisely specified aggregate of water molecules, what I say may be neither determinately true nor false. But such vagueness resides in the reference of singular terms—in this case, "that cloud in the sky"—not in the identity relation itself.

Some philosophers, however, hold that there is vagueness, not only in our reference to objects, but in the objects themselves; not only in our language and thought, but in the world. Let us suppose, charitably, that such a view makes sense. Might not these vague objects be vaguely identical? That depends. If vague identity is understood so that vaguely identical objects are neither determinately identical nor determinately not identical, then the answer is no, as the following argument shows. (Versions are in Evans 1978, Salmon 1981). Suppose *a* and *b* are vaguely identical; then they differ in some property, namely, being vaguely identical with *b.* For although *a* has the property, *b* does not: Nothing is vaguely identical with itself. By the indiscernibility of identicals, then, *a* is (determinately) not identical with *b.* So, vaguely identical objects are (determinately) not identical! That sounds odd; but there is no contradiction if vague identity is understood in some way that detaches it from indeterminacy of truth-value. So understood, vague identity poses no challenge to classical identity.

TEMPORARY IDENTITY

The Greek philosopher Heraclitus argued that one cannot bathe in the same river twice, something as follows. Rivers flow. The stretch of water that comprises the river on Monday is not the same as the stretch of water that comprises the river on Tuesday. But a river is not something separate and distinct from the stretch of water that comprises it; be it on Monday or on Tuesday, the river and the stretch of water are one and the same. It follows, by a double application of the indiscernibility of identicals, that the river on Monday is not the same as the river on Tuesday. If one bathes in the river on Monday, and returns to bathe at the same place on Tuesday, one has not bathed in the same river twice.

One wants to say: On Monday, the river is identical with a certain stretch of water; on Tuesday, the same river is identical with a different stretch of water. More generally, identity can be temporary, holding at some times but not at others. Temporary identity, however, is disallowed by the above argument, not just for rivers, but for all entities whatsoever. Should we abandon the classical notion of identity that the argument presupposes?

There are at least two responses to Heraclitus's problem compatible with classical identity. According to the first response (inspired by Aristotle, when we say that a river *is* just a certain stretch of water, we are using not the "is" of identity but the "is" of constitution; and constitution is never identity (see Lowe 1989). On this view there are two fundamentally different kinds of entities that occupy space and persist through time. There are ordinary material objects, such as rivers, trees, statues, and tables; and there are portions of matter that may temporarily constitute the ordinary objects. At any time an ordinary object is constituted by some portion of matter

or other; but at no time is it identical with that portion of matter, either wholly or in part. In particular, the very same river is constituted by one stretch of water on Monday and by a different stretch of water on Tuesday. No conflict arises with the laws of classical identity, and Heraclitus's problem is solved.

This response, however, is not without problems. A dualism of ordinary objects and the portions of matter that constitute them is neither necessary nor sufficient to solve the general problem of temporary identity. It is not sufficient, because some cases of temporary identity have nothing to do with constitution. Consider a tree that, at some bleak stage of its career, consists of nothing but a trunk. Later, however, the tree sprouts new branches and leaves. Then we have another prima facie case of temporary identity: The tree is identical with the trunk at the bleak time but not identical with the trunk at the happier time. In this case, however, invoking constitution is of no avail: Neither the trunk nor the tree constitutes the other, in the relevant sense. (This example is from Hirsch 1982.)

Nor is such a dualism necessary to solve the problem of temporary identity, because another response is available, one (arguably) more economical in its ontological commitments (see Hirsch 1982, Quine 1950). On this second response objects that persist through time are composed of (more-or-less) momentary stages, of temporal parts. A persisting river is a sum of stages unified in a way appropriate for rivers; a persisting aggregate of water molecules is a sum of stages unified in a way appropriate for portions of matter. A persisting river and a persisting aggregate of water molecules may overlap by having a stage in common; in that case a stage of the river and a contemporaneous stage of the aggregate of water molecules are identical. But the persisting river is not identical with the persisting aggregate of water molecules: Later stages of the river are in about the same place as earlier stages and are no less spatially continuous; later stages of the aggregate of water molecules are downstream of earlier stages and are spatially scattered. When we say that, at any time, a river is nothing separate and distinct from the water that comprises it, this must be understood as asserting not an identity between persisting objects but an identity between stages. Identity between stages, however, is all one needs to avoid the uneconomical dualism of the constitution view. All objects that occupy space and persist through time are composed of a single kind of entity: Stages of portions of matter. (The stage view of persistence is argued for in Lewis 1986.)

Heraclitus's problem is now easily solved. One cannot bathe in the same river stage twice; but one can bathe in the same river twice by bathing successively in two river stages belonging to a single persisting river. That these two stages are not stages of a single persisting aggregate of water molecules is irrelevant. There is no conflict with classical identity.

CONTINGENT IDENTITY

A change in example, however, makes trouble for the stage view of persistence. Consider a statue called *Goliath* that consists entirely of a lump of clay called Lumpl; and suppose that the statue and the lump came into being, and ceased to exist, at exactly the same times. Then, on the stage view, every stage of *Goliath* is identical with a stage of Lumpl and vice versa; *Goliath* and Lumpl are the same sum of stages and so are identical. But, surely, they are not necessarily identical. *Goliath* could have been destroyed without destroying Lumpl—say, by being squashed—in which case *Goliath* would have lacked Lumpl's final stages and would have been a distinct sum from Lumpl. So, *Goliath* and Lumpl are identical, but only contingently identical. (The example is from Gibbard 1975.)

Trouble arises because contingent identity, no less than temporary identity, is incompatible with identity, classically conceived—or so the following argument seems to show. Consider the property is-necessarily-identical-with-*y*, for some object *y*. Surely *y* has it: Everything is necessarily identical with itself. Now suppose an object *x* is identical with *y*. Then, by the indiscernibility of identicals, *x* has the property as well; that is, *x* is necessarily identical with *y*. Thus, objects are necessarily identical if identical at all; objects are never contingently identical.

Whether this argument is unassailable will depend upon one's interpretation of modal properties, of modality *de re*. If objects have their modal properties absolutely, in and of themselves, then the argument is sound. Since *Goliath* and Lumpl are not necessarily identical, they are not identical at all. *Goliath* and Lumpl are numerically distinct objects that occupy the same place at all times that they exist. *Goliath* is not identical with any sum of matter-stages, contradicting the stage view of persistence.

The stage view can be preserved, however, if one takes the view that modal predicates do not apply to objects absolutely, in and of themselves; their application is relative to how the objects are conceived, classified, or referred to. For example, could the lump of clay—that is, the statue—have survived a squashing? *Qua* lump of clay, it could; *qua* statue, it could not. There is no violation of the indiscernibility of identicals because the modal predicate "could survive a squashing" expresses no property

when considered out of context and expresses different properties when attached to "the lump of clay" (or "Lumpl") and to "the statue" (or *Goliath*). In this way the stage view can accept the contingent identity of Lumpl and *Goliath,* without forfeiting classical identity. (For versions of this strategy, see Gibbard 1975, Lewis 1971.)

RELATIVE IDENTITY

Classical identity is absolute: Whether identity holds between objects does not depend upon how those objects are conceived, classified, or referred to. In ordinary language we often say "*a* is the same *F* as *b,*" for some general term "*F*"; but this is naturally analyzed as a restriction of absolute identity: *a* is *F,* and *b* is *F,* and *a* is (absolutely) identical with *b.*

Geach has argued, on the contrary, that all identity statements are relative: "*a* is the same *F* as *b*" cannot be analyzed as restricted absolute identity, because there is no absolute identity; when we say simply "*a* is the same as *b,*" some general term "*F*" must be supplied by context, or what we say is meaningless (Geach 1970). To support his claim, Geach has presented examples in which we would say: *a* and *b* are the same *F,* and *a* and *b* are *G*'s, but *a* and *b* are not the same *G.* Consider the word *tot.* It contains three letter tokens, two letter types. The first letter token and the last letter token are not the same letter token, but they are the same letter type. That contradicts the claim that "the same *F*" is to be analyzed as restricted absolute identity.

The defender of classical identity has a simple and natural reply: Sometimes the relation is-the-same-*F*-as is not restricted identity but rather some weaker equivalence relation; that is, sometimes it is a species of qualitative, rather than numerical, identity (see Perry 1970). For example: If I say that you are wearing the same coat as I am, I (probably) do not mean the numerically same coat. Similarly, letter tokens of the same type are qualitatively similar—equiform—not numerically identical. To the extent that Geach's point is just that "the same *F*" cannot always be analyzed as restricted identity, it is a point no one should deny.

Any rejection of absolute identity, it seems, must be based upon arguments of a more abstract sort. Indeed. Geach explicitly rejects the standard characterization of identity through Leibniz's law on the grounds that second-order quantification over properties leads to paradox. And he rightly points out that, within first-order logic, characterizations of identity are inevitably relative to the predicates of the language. But how does this impugn the meaningfulness of absolute identity? Does Geach's argument simply amount to the demand, Define absolute identity, or count it as meaningless? That demand, certainly, is too strong. No fundamental notion of logic or metaphysics could meet it.

See also Aristotle; Frege, Gottlob; Heraclitus of Ephesus; Kripke, Saul; Leibniz, Gottfried Wilhelm; Modality, Philosophy and Metaphysics of; Personal Identity; Properties; Quine, Willard Van Orman; Vagueness.

Bibliography

Adams, R. M. "Primitive Thisness and Primitive Identity." *Journal of Philosophy* 76 (1979): 5–26.

Black, M. "The Identity of Indiscernibles." In *Problems of Analysis.* Ithaca, NY: Cornell University Press, 1954.

Cartwright, R. "Identity and Substitutivity." In *Identity and Individuation,* edited by M. Munitz. New York: New York University Press, 1971.

Evans, G. "Can There Be Vague Objects?" *Analysis* 38 (1978): 208.

Frege, G. "Uber Sinn und Bedeutung." *Zeitschrift fur Philosophie und philosophische Kritik* 100 (1892): 25–50. Translated by P. Geach and M. Black as "On Sense and Reference." In *Philosophical Writings of Gottlob Frege.* Oxford: Blackwell, 1952.

Geach, P. T. "Identity." *Review of Metaphysics* 21 (1967): 3–12.

Gibbard, A. "Contingent Identity." *Journal of Philosophical Logic* 4 (1975): 187–221.

Hirsch, E. *The Concept of Identity.* Oxford: Oxford University Press, 1982.

Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.

Lewis, D. "Counterparts of Persons and Their Bodies." *Journal of Philosophy* 68 (1971): 203–211.

Lewis, D. "Many, but Almost One." In *Ontology, Causality, and Mind: Essays in Honour of D. M. Armstrong,* edited by K. Campbell, J. Bacon, and L. Reinhardt. Cambridge, U.K.: Cambridge University Press, 1993.

Lewis, D. *On the Plurality of Worlds.* Oxford: Blackwell, 1986.

Lowe, E. J. *Kinds of Being.* Oxford: Blackwell, 1989.

Perry, J. "The Same *F.*" *Philosophical Review* (1970): 181–200.

Quine, W. V. O. "Identity, Ostension, and Hypostasis." *Journal of Philosophy* (1950).

Quine, W. V. O. "Reference and Modality." In *From a Logical Point of View.* Cambridge, MA: Harvard University Press, 1953.

Quine, W. V. O. *Word and Object.* Cambridge, MA: MIT Press, 1960.

Salmon, N. *Frege's Puzzle.* Cambridge, MA: MIT Press, 1986.

Salmon, N. *Reference and Essence.* Princeton, NJ: Princeton University Press, 1981.

Phillip Bricker (1996)

IDENTITY OF INDISCERNIBLES

See *Identity*

IDENTITY OF MENTAL AND BODILY STATES

See *Mind-Body Problem*

IDEOLOGY

Though often employed as a catchall term in contemporary usage, including some philosophers' usage, *ideology* has a clearly identifiable historical origin and since its invention has borne some clear though disparate meanings (as well as, to be sure, some unclear ones) in several traditions of thought, most notably in the Marxian tradition.

It was Antoine Destutt de Tracy who, toward the end of the eighteenth century, conceived the notion of developing a science of ideas that would trace them back to their supposed material elements. The group around him became known as the *Idéologues* and at first found favor with Napoleon Bonaparte, whose coup ending the period of the French Revolution and its immediate aftermath they at first applauded. But they soon became his vehement critics (concerning, for example, his policy-driven revival of religion), and Napoleon returned the compliment by denouncing them for, among other things, allegedly indulging in wild ideas rather than respecting the exigencies of the concrete political situation. Thus did "ideologists" become an epithet, an expression of contempt.

As such, the term was picked up and used by Karl Marx and Friedrich Engels some four decades later. In *Die Deutsche Ideologie* (*The German Ideology*, 1976), a lengthy work, they lampoon their neo-Hegelian near-contemporaries, notably Ludwig Feuerbach, Bruno Bauer, and Max Stirner, on the ground that the supposedly weighty disputes of the latter are pseudobattles among merely abstract, primarily theologically based ideas, lacking any influence on, or even much connection with, the actual sociohistorical world. Here, "ideology" is equated with religion, metaphysics, moral theory, and similar products of pure consciousness and is given roughly the same highly pejorative valence, though affixed to an entirely different object, as that formerly given by Napoleon to the objects of his wrath.

But, unlike some of Marx's criticisms of the idealist philosophy of G. W. F. Hegel himself, which were printed during Marx's lifetime, *The German Ideology* was not actually published, and hence its textual details were not generally known, until 1932. Marx does, however, mention it, in a brief autobiographical sketch that appeared in 1859, as having been the early outcome—one left to the "gnawing criticism of the mice" when the original arrangement to have it published fell through—of his and Engels's newly elaborated systematic opposition to the "ideological" standpoint of German philosophy. At the same time, in their widely circulated *Manifesto of the Communist Party* of 1848 they at one point allude to the anticipated defection from class solidarity of a section of the bourgeois class, notably some (though by implication just a few) of the bourgeois ideologists, by virtue of the latter's having achieved a comprehensive overview of the process of history. While this passage is revealing as a veiled self-reference, it is equally interesting for its suggestion that "ideology" can have a positive connotation, as well. Hence the later ambivalence of the term in Marxist and non-Marxist contexts alike is already to be found in the classical writings of the Marxian tradition.

CONNOTATIONS IN LATER MARXISMS

The pejorative understanding of "ideology," linked as it is to the idea that *most* philosophers and other intellectuals typically engage in mystificatory, distortive justifications, or legitimizations, of the existing social order with the effect of reinforcing the dominant insitutions of the ruling class of which they are a part, continued to predominate especially in so-called "Western Marxism." By this is meant those strands of neo-Marxist thought that preserved their independence from the Communist Party, based primarily in the Soviet Union from the time of the Russian Revolution until the final decade of the twentieth century, the successive leaders of which stipulated the terms of what they considered to be orthodox Marxist theory.

Western Marxists, accused of "revisionism" by these leaders and their followers, tended rather to consider the wooden, dogmatic style and content of "orthodox Marxist" writings to constitute a serious distortion of Marx's ideas. For example, Herbert Marcuse, a leading figure in the early Frankfurt School in Germany before migrating to the United States, retained, analyzed, and applied the pejorative sense of the term "ideology" both in his 1958 critique of the Soviet Union, *Soviet Marxism*, and in his

early 1960s indictment of Western society as tranquilized, democratic, but profoundly unfree, *One-Dimensional Man* (1966). Similarly, his erstwhile Frankfurt colleague, Theodor Adorno, another strong social critic, equated "ideology" with "false consciousness" and regarded it as being characteristic of those who are obsessed with enforcing identity and conformity and who fail to respect differences. The best-known member of the "later" Frankfurt School, Jürgen Habermas, while he has diverged from Marxism in a great many respects, has continued to equate ideology with systematically distorted communication, to be combated through what he calls emancipatory critique.

The self-styled "orthodox" Marxists, however, took their cue on the question of the meaning of ideology above all from the Russian revolutionary leader, V. I. Lenin, who in his early call to arms, *What Is to Be Done?*, insisted that a clear-cut choice had to be made between bourgeois ideology, which he contemned, and "socialist" ideology, which he espoused and thought it necessary for professional revolutionaries to inculcate in the minds of the masses. Most of the subtler thinkers within the orbit of the Communist Party, such as the Hungarian György (Georg) Lukács (1971) and the Italian Antonio Gramsci, also saw ideology as a potentially and at least partially positive phenomenon, with Lukács depicting Marxism as the ideological expression of the proletarian class. One of the most complex and idiosyncratic conceptions of ideology to be developed by someone who was at the time a Communist Party member was that of the once influential French philosopher Louis Althusser (1969): He contrasted ideology with science, of which he saw Marx's theory of history as a leading instance, but at the same time he took ideology to be a pervasive and ineliminable part of human experience, regardless of a given historical society's class configurations.

NON-MARXIST CONCEPTIONS

In *Ideology and Utopia* Karl Mannheim (1986), the German sociologist of knowledge who was himself influenced by the early Lukács, distinguished between what he called the "particular" and the "total" concepts of the term, the former being linked with suspicion concerning the motives of others as interested and biased—in other words, "ideology" as more purely pejorative—and the latter characterizing the comprehensive views of many large groups, such as classes, in the modern world. "Ideology" in the latter sense is to be seen as a pervasive historical phenomenon. Espousing a nonevaluative approach to the understanding of diverse worldviews (*Weltan-schauungen*) that he denominated "relationism," Mannheim in effect paved the way for the much broader, more all-encompassing, less critical usage of the term "ideology" that has become common.

No treatment of the meandering evolution of this term could pretend to adequacy without noting the curious recurrence of announcements that its supposed referent has, or may have, ceased to exist. Political scientists, such as Seymour Lipset, and other philosophically oriented sociologists, such as Raymond Aron, have evoked this as at least a possibility, but no doubt its most famous assertion occurred in a lengthy tome by Daniel Bell, an American sociologist strongly influenced by the Marx scholar and philosopher turned fervent anticommunist, Sidney Hook. The title of Bell's book, especially its less well-known subtitle, accurately captures its principal claim; it is *The End of Ideology: On the Exhaustion of Political Ideas in the Fifties* (1960). Understandably, but rather unfortunately, Bell's main title led to oversimplified interpretations of what he actually intended, which was *not* an umbrella thesis supposedly applicable to all future times and places.

To some (e.g., Hannah Arendt), ideology means totalitarianism, of which Communist ideology is a salient example; to others (e.g., Edward Shils attempting to define the term in the *International Encyclopedia of the Social Sciences*), it means, above all, intolerant belief systems that are by and large inimical to science. It is in any case evident that the confusion and even contradictoriness of meanings of the term that are traceable to its historical origins have continued to characterize its deployment, which remains widespread in the literatures of philosophy, political science, sociology, psychology, literary theory, and even popular journalism despite its alleged demise as a phenomenon at the end of the 1950s.

EVALUATION

Apart from those, if there are any, who still share Destutt de Tracy's youthful confidence in the possibility of generating a materialist science of ideas based on an analysis of the origins of their components, it would seem that "ideology" would indeed lose its purpose as a part of our vocabulary if all negative, critical connotations were to be excised from it. At least one important meaning of it remains, and should continue to remain, that of suspect generalized claims, often entire theories, which purport to be true but are in fact intellectual constructions designed to reinforce particular interests, especially the interests of those in power. Although many of those who believe in particular "ideologies" as so understood may do so unreflectively—in an important sense, after all, it is

the aim of skilful ideologists to maximize the number of such believers—ideologies in this sense of the word should in the last analysis be capable of being unmasked as sophistic and in bad faith.

But there are a number of philosophical problems involved in elaborating a coherent conception of ideology. The first of these concerns the question of one's basis for designating another's set of propositions or beliefs as ideological: How can one be sure that one's own supposedly critical standpoint is not itself ideological? May it not also be, in the last analysis, merely an elaborate apologia for an alternative special interest aspiring to social dominance? Marx and Engels thought to evade this difficulty by painting the proletariat as the class, the coming to power of which would usher in a classless society, without particular interests or internal relationships of dominance and subordination; hence, they believed, the class consciousness of the proletariat, history's first truly "universal class," should be regarded as radically different in kind from the bourgeois, feudal, and other ideological standpoints of the past. But may this not be just one more intellectual sleight of hand?

Another problem inherent in the conception of ideology as critical and "suspicious" is that of its explicit or implicit tendency to relegate philosophy itself, in its various branches, to the realm of ideology. At times Marx and Engels wrote as if philosophical and other ideas were in fact just epiphenomena, ghostly by-products of a real world of which the "base" consisted of the dominant forces of production; that is, the crafts, industries, and technologies of any given historical period, and the "superstructure" consisted of the political, legal, and other institutions developed in conformity with those forces. This conception, taken to its extreme, would deny that ideological phenomena have any autonomy, any force—in other words, that ideas as such can ever have real consequences. But such a claim runs counter to much of human experience. (Engels himself lived long enough to express regret over this misinterpretation of his views. He located its origin in the long past *Zeitgeist* of the era when he and Marx had begun formulating their own ideas, an era when the Hegelian and neo-Hegelian idealist philosophies, to which they were so opposed, were in the ascendancy.)

Ultimately, of course, the fundamental problem concerning ideology is the fundamental problem of virtually all of philosophy; that is, the problem of truth itself. How could we ever succeed in assuring ourselves and others, beyond all doubt, that a claim or set of claims that we assert to be an ideological distortion of the "true" state of affairs actually is such? For to do so would presuppose, contrary to *all* past experience, a complete and comprehensive grasp, on our part, of the true state of affairs.

See also Cosmopolitanism; Postcolonialism; Republicanism.

Bibliography

Althusser, Louis. *For Marx*. Translated by Ben Brewster. London: Allen Lane, 1969.

Barrett, Michèle. *The Politics of Truth: From Marx to Foucault*. Stanford: Stanford University Press, 1991.

Bell, Daniel. *The End of Ideology: On the Exhaustion of Political Ideas in the Fifties*. New York: Free Press, 1960.

Eagleton, Terry. *Ideology: An Introduction*. London: Verso, 1991.

Larrain, Jorge. *The Concept of Ideology*. Athens: University of Georgia Press, 1979.

Lukács, Georg. *History and Class Consciousness: Studies in Marxist Dialectics*. Translated by Rodney Livingstone. Cambridge, MA: MIT Press, 1971.

Mannheim, Karl. *Ideology and Utopia*, 1936. Translated by Louis Worth and Edward Shils. San Diego: Harcourt Brace Jovanovich, 1986.

Marcuse, Herbert. *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Boston: Beacon Press, 1966.

Marx, Karl, and Friedrich Engels. *The German Ideology*. Moscow: Progress Publishers, 1976.

McLellan, David. *Ideology*. Minneapolis: University of Minnesota Press, 1986.

Mészáros, Istvan. *The Power of Ideology*. New York: New York University Press, 1989.

Nielsen, Kai. *Marxism and the Moral Point of View: Morality, Ideology, and Historical Materialism*. Boulder, CO: Westview Press, 1989.

Parekh, Bhikku. *Marx's Theory of Ideology*. Baltimore: Johns Hopkins University Press, 1982.

Rosen, Michael. *On Voluntary Servitude: False Consciousness and the Theory of Ideology*. Cambridge, MA: Harvard University Press, 1996.

William L. McBride (2005)

IKHWĀN AL-ṢAFĀʾ

The Ikhwān al-Ṣafāʾ (The Brethren of Purity or The Sincere Brethren) were the anonymous adepts of an esoteric fraternity of lettered urbanites that was principally based in the Mesopotamian cities of Basra and Baghdad in the second half of the tenth century CE. This learned brotherhood occupied a prominent station in the history of science and philosophy in Islam due to the wide intellectual reception of their famed tracts the *Rasāʾil Ikhwān al-Ṣafāʾ* (The epistles of the brethren of purity). The exact

dating and authorship of this encyclopedic compendium remain unsettled polemical questions, and it is widely assumed that the provenance of the Ikhwān's ideas is primarily ascribable to Ismāʿīlī sources. Nonetheless, this is controversial, and it is rather more circumspect to attribute their outlook to a broadly Shīʿī lineage.

The Ikhwān's corpus displays a tolerance for multifarious pagan and monotheistic traditions. Besides their filial observance of the teachings of the Qurʾān, the Ikhwān also reverently appealed to the Torah of Judaism and to the Gospels of Christianity, primarily in their accounts of prophetology. In addition, they heeded the legacies of the Stoics and of Pythagoras, Hermes, Socrates, Plato, Aristotle, Plotinus, Euclid, Ptolemy, Porphyry, and Iamblichus. They moreover strived to establish some form of harmony between faith and reason, in a manner that is partly reminiscent of the practices of al-Fārābī (d. 950), Avicenna (Ibn Sīnā) (d. 1037) and Averroes (Ibn Rushd) (d. 1198).

Motivated by an active soteriological pursuit of happiness, the Ikhwān promoted a convivial and earnest companionship of virtue. Their eschatological views were furthermore articulated by way of an intricate cyclical view of history and an uncanny hermeneutic interpretation of the microcosm and macrocosm analogy. The multiplicity of the voices that were expressed in their tracts reflect a genuine quest for wisdom that is driven by an impetus that is not reducible to a mere eclecticism. Their ecumenical syncretism, which may have been partly influenced by the outlooks of the Sabaeans of Harran, grounded their aspiration to establish a spiritual refuge that would transcend the sectarian divisions that beset their era.

Customarily enumerated as fifty-two epistles, the Ikhwān's *Rasāʾil* offer synoptic explications of the classical sciences of the ancients and the moderns of the age. Divided into four classificatory parts, these treatises treated themes in mathematics, logic, physics, psychology, and theology. This series was also accompanied by a concise tract titled: *al-Risāla al-Jāmiʿa* (*The Comprehensive Epistle*), which acted as the summary of their corpus, and was supplemented by an abridged appendage known as *Risālat Jāmiʿat al-Jāmiʿa* (*The Condensed Comprehensive Epistle*).

The eloquent literary style of the *Rasāʾil* covers the technicalities of mathematics, logic, physics, and medicine, together with religious speculations, occultist incantations, along with the poetic elaborateness of fables, odes, and didactic parables. Although the influence of the *Rasāʾil* in Ismāʿīlī circles was prominent, and in spite of being partially manifest in various doctrinal citations in Islam, the impact that these epistles may have had on the philosophers and the dialectical theologians has been rather exaggerated. In spite of the extensive thematic scope of the *Rasāʾil*, which may have occasionally been plagued by repetitions, these epistles do not establish a convincing intellectual relationship with the achievements of the classical authorities of Arabic sciences and philosophy. Despite being usually classed as philosophers, the Ikhwān would more fairly rank as learned compilers of knowledge when compared with the philosophical luminaries of the period.

Although the Ikhwān's erudite reflections on spirituality show signs of originality, this does not make the *Rasāʾil* the principal reference for all the disciplines that they endeavored to tackle. Their investigations in geometry, arithmetic, logic, and physics remained diluted in essence, and although these inquiries were complemented by oral instructions in seminars, they nevertheless represented a minor aspect of the disciplines they addressed. Even though it is usually claimed that Avicenna may have been implicitly influenced in his intellectual formation by their teachings, his philosophical acumen remained superior to that of the Ikhwān. A similar observation may be made concerning the Ikhwān's impact on the unfolding of the Illuminationist and emanationist tenets in Islam, or the implicit influence that they may have exercised on the dialectical doctrines of the exponents of *kalām*. Moreover, although some Shīʿī illuminati and Sunnī literati may have professedly cited the tutelage of the Ikhwān, this did not entail that they were affiliates of this line in thinking. Despite some of these scholarly shortcomings, the *Rasāʾil* represent a commendable populist adaptation of science and philosophy that merits the privilege of being ranked among the high literature of Islam.

See also al-Fārābī; Aristotle; Averroes; Avicenna; Islamic Philosophy; Neoplatonism; Plato; Plotinus; Porphyry; Stoicism.

Bibliography

ARABIC EDITIONS

Rasāʾil Ikhwān al-Ṣafā. Vols. 1–4. With introduction by Butrus Bustani. Beirut, Lebanon: Dar Sadir, 1957.

REFERENCES

ʿAwa, ʿAdel. *L'esprit critique des Frères de la Pureté: Encyclopédistes Arabes du IVe/Xe siècle*. Beirut: Imprimerie Catholique, 1948.

Baffioni, Carmela. *L'epistola degli Ikhwan al-Safa', sulle opinioni e le religioni*. Naples: Istituto Universitario Orientale, 1989.

Bausani, Alessandro. *L'enciclopedia dei Fratelli della Purità*. Naples: Istituto Universitario Orientale, 1978.

De Callatay, Godefroid. *Ikhwan al-Safa': Les révolutions et les cycles (Épîtres des Frères de la Pureté, XXXVI)*. Beirut: al-Buraq, 1996.

Diwald, Susanne. *Arabische Philosophie und Wissenschaft in der Enzyklopädie Kitab as-Safa' (III): Die Lehre von Seele und Intellekt*. Wiesbaden, Germany: Otto Harrassowitz, 1975.

Goodman, Lenn E. *The Case of the Animals versus Man before the King of the Jinn: A Tenth-Century Ecological Fable of the Pure Brethren of Basra*. Boston: Twayne, 1978.

Hamdani, Abbas. "A Critique of Paul Casanova's Dating of the *Rasa'il Ikhwan al-Safa'*." In *Mediaeval Isma'ili History and Thought*, edited by Farhad Daftary. Cambridge, U.K.: Cambridge University Press, 1996.

Marquet, Yves. "Ikhwan al-Safa'." *The Encyclopaedia of Islam*. Leiden, Netherlands: E. J. Brill, 1960.

Marquet, Yves. *La Philosophie des Ikhwan al-Safa', nouvelle édition augmentée*. Milan: Archè, 1999.

Netton, Ian Richard. *Muslim Neoplatonists: An Introduction to the Thought of the Brethren of Purity*. London: Allen & Unwin, 1982.

Netton, Ian Richard. "The Brethren of Purity (Ikhwan al-Safa')" In *History of Islamic Philosophy*, edited by Seyyed Nasr and Oliver Leaman, 222–230. London; New York: Routledge, 1996.

Tibawi, Abdul-Latif. "Ikhwan as-Safa' and their *Rasa'il*: A Critical Review of a Century and a Half of Research." *Islamic Quarterly* 2 (1955): 28–46.

Nader El-Bizri (2005)

IL'IN, IVAN ALEKSANDROVICH

(1883–1954)

Ivan Aleksandrovich Il'in was a Russian religious philosopher, legal and political theorist, philosopher of Russian national identity, orator, and publicist. Born in Moscow of a noble family, he studied law at Moscow University, from which he graduated in 1906. Deemed an extraordinary scholarly achievement, his dissertation on Georg Hegel (1770–1831) earned him at once master's and doctoral degrees in 1918. The two volumes of his published dissertation, *Filosofiia Gegelia kak uchenie o konkretnosti Boga i cheloveka* (The philosophy of Hegel as a doctrine of the concreteness of God and man, 1918), have been described as one of the more significant commentaries on Hegel published in the twentieth century in any language (Grier 1997). As a resolute foe of the Bolsheviks, he was exiled from Soviet Russia under threat of execution. Il'in and his wife left Russia for Germany in 1922. He was a renowned anti-Bolshevik and anticommunist, which suited the Nazis, but his critique of totalitarianism was not at all appreciated by the Nazi regime. Moreover, in 1934 he refused to accept their orders to spread Nazi propaganda in the Russian Academic Institute and was subsequently removed from his post by them and banned from all further employment. Finally, in 1938 he and his wife escaped from Nazi Germany and found refuge in Switzerland.

In his book *Religioznyi smysl filosofii* (The religious meaning of philosophy, 1925) Il'in proposed the following task to his readers: "To find authentic, spiritually objective environments." The spiritual labor of such a philosopher approximates that of the monk. They differ only in that the former occurs in the realm of cognition, while the latter seeks redemption. Both the philosopher and the monk comprehend, Il'in believed, that the world is "really a school where reasonable souls exercise themselves, a training ground where they learn to know God" (see St. Basil the Great's [c. 329–379] *Hexaëmeron* 1.6).

POLITICAL THEORY OF LEGAL CONSCIOUSNESS

Convinced that Hegel's doctrine of right and the state was ultimately a failure, Il'in began to elaborate his own theory in 1919, while still in Soviet Russia. It was published posthumously as *O sushchnosti pravosoznaniia* (On the essence of legal consciousness) in 1956. The central notion in it is *pravosoznanie*, from *pravo* (right or law) and *soznanie* (consciousness), which can perhaps best be translated as "legal consciousness." He formulated three axioms of legal consciousness—"a feeling of one's own spirit dignity; an ability to self-obligation and self-government; a mutual respect and trust of people to each other" (1956b, chapters 15, 17, 19)—and added nine axioms of authority, which included its *pravovoe polnomochie* (legal plenary powers), that the state authority must be *edina* (singular) and that it must be realized by "the best" people who meet ethical and political qualification (chapter 14). Then he joined natural law and positive law (i.e., legal norms including agreement with morality and justice, and legal norms established by legal authority and being subject to application) in an appropriate way: "A rational system of positive law would reflect the structure of natural law" (chapters 5, 6). Unfortunately, our political reality does not give a good example of even an approximate realization of such a doctrine.

ETHICS DOCTRINE

On the Essence of Legal Consciousness contains a number of uncommon but profound observations that are useful

to any national leader who wishes to have "a deeper religious and moral motivation" for ruling, for example, to follow Il'in's conception in which "the ultimate justification of state authority would be the development in the citizenry of a moral, legal and spiritual culture in which the requirements of natural rights would be so widely exemplified in human conduct as to make genuine self-government a reality" (Grier 1998, p. 693)

In 1925 Il'in published his polemical book *O soprotivlenii zlu siloiu* (On resistance to evil by force), concerning an important dual ethical problem: "May a human being who is trying to achieve ethical perfection *resist* evil by force, using the sword?" and "May a human being who believes in God and accepts His creation, and who knows his place in this created world *not resist* evil by force, using the sword?" (1925, chapter 19). Il'in gave a single direct answer to both: that one not only *may* but *must* also resist evil by force. He wrote:

> Physical intervention and coercion may become the direct religious and patriotic duty of a human being; and once this happens one must not evade it. To fulfill this duty is to become a participant in the great historical battle between God's servants and the forces of the underworld; and this battle will force him not only to draw his sword but to take upon himself the burden of homicide. (chapter 19)

This book evoked a strong response not only in the Russian émigré community but in the Soviet Union as well (Poltoratzky 1975, Lisitsa 1996).

Il'in lived and worked with the single-minded purpose of reconstructing Russia in an authentic way in the aftermath of the Bolshevik regime. He wrote *Osnovy gosudarstvennogo ustroistva: Proekt Osnovnogo Zakona Rossii* (The foundations of government: A proposal for the fundamental law of Russia, 1996) as a post-Bolshevik constitution. Between 1940 and 1954 he produced 215 anonymous bulletins for a restricted list of readers only, and these *Nashi zadachi: Stat'i 1948–1954* (Our tasks: Articles 1948–1954) were published in 1956 in two volumes only after his death. This two-volume work is nothing other than an "Axiomatics of Political Life," analogous to his *Aksiomy religioznogo opyta* (Axioms of religious experience, 1953); it is clearly intended to treat the disease of "political nihilism."

RELIGIOUS THOUGHTS

Between 1938 and 1945 Il'in created in German a wonderful literary triptych—*Ich schaue in Leben: Ein Buch der Besinnung* (I am peering into life: A book of thoughts, 1938–1939), *Das verloschollene Herz: Ein Buch stiller Betrachtungen* (The singing heart: A book of quiet contemplations, 1943), and *Blick in die Ferne: Ein Buch der Einsichten und Hoffnungen* (A look in the distance: A book of reflections and hopes, 1945)—and described it as "devoted not to theology, but to a quiet, philosophical praising of God." Despite all the striving of humanity to unveil the mystery of world creation, it has been losing access to this mystery on the path that it has selected. "For the world remains as before," wrote Il'in, "i.e., a great mysterious wonder, created by a rational inner Authority, carried by a rational inner force, and moving toward a certain inner goal" (*Put' k ochevidnosti* [The path to self-evidence], 1957, chapter 18). And this "lost mystery" might be returned to humanity through a contemplative heart, but only if the heart is open, loving, and marveling.

In 1953 Il'in published in Paris *Axioms of Religious Experience*, his main work, in two volumes on which he had been working for thirty-three years. It was a profound and original investigation of the personal "religious act." One of its axioms, "The autonomy of religious experience," and the motifs connected with it, such as "loneliness" and "tragedy in the world," were received somewhat critically by the Russian theologians Archimandrite Kostantin Zaitsev and Father Vladislav Sveshnikov. And yet Il'in derived this axiom from one of the Church Fathers, Petrus Chrysologus, who described God as *solus, sed non solitarius* (God alone is, but He is not lonely).

When Il'in died on December 21, 1954, and was buried in a cemetery in the village of Zollikon, near Zürich, there appeared on his monument an epitaph composed by the philosopher himself:

Alles empfunden	(Felt it all
So viel gelitten	Suffered so much
In Liebe geschauet	Had it revealed through love
Manches verschuldet	Guilty of some things
Und wenig verstanden	Understood very little
Danke Dir, ewige Güte!	Thank you, eternal Goodness!)
	(Il'in 1993).

Il'in's legacy is enormous and well preserved. It contains more than 40 books and brochures, 600 articles, 100 lectures, a large collection of letters, humorous poems, several "naive political fairy tales," memoirs, and documents that are in several archives in various countries. Il'in's largest archive was organized in 1963 by Professor Nikolai Poltoratsky at Michigan State University

Libraries. This collection and other materials are being published as a project of the publisher Russkaia Kniga (Russian Book) in Moscow, and are expected to contain forty volumes in the series *I. A. Il'in: Sobranie sochinenii v desiati tomakh* (Il'in, I. A., Collected Works), Moscow: Russkaia Kniga, 1993–1999.

See also Authority; Consciousness; Hegel, Georg Wilhelm Friedrich; Natural Law; Philosophy of Religion, History of; Russian Philosophy.

Bibliography

PRIMARY WORKS

Aksiomy religioznogo opyta (Axioms of religious experience). 2 vols. Paris: n.p., 1953.

Blick in die Ferne: Ein Buch der Einsichten und Hoffnungen (A look in the distance: A book of reflections and hopes). Affoltern a.A.: Aehren-Verlag, 1945.

Das verloschollene Herz: Ein Buch stiller Betrachtungen (The singing heart: A book of quiet contemplations). Bern: Verlag Paul Haupt, 1943.

Filosofiia Gegelia kak uchenie o konkretnosti Boga i cheloveka (The philosophy of hegel as a doctrine of the concreteness of God and man). Moscow: Russkaia Kniga, 2002. Originally published in 1918. 2 vols. (Moscow: Izd. G.A. Lemana i S.I. Sakharova, 1918).

I. A. Il'in: Sobranie sochinenii v desiati tomakh. Vol. 1. Moscow: Russkaia Kniga, 1993.

Ich schaue in Leben: Ein Buch der Besinnung (I am peering into life: A book of thoughts), Berlin: Furche-Verlag, 1938–1939.

Nashi zadachi: Stat'i 1948–1954 (Our tasks: Articles 1948–1954), 2 Vols. Paris: Izd-vo Russkago Obshche-Voennago Soìuza, 1956a.

O soprotivlenii zlu siloiu (On resistance to evil by force). (1925). London, Canada: Zaria, 1975.

O sushchnosti pravosoznaniia (On the essence of legal consciousness). Münich: n.p., 1956b.

Osnovy gosudarstvennogo ustroistva: Proekt Osnovnogo Zakona Rossii (The foundations of government: A proposal for the fundamental law of Russia). Moscow: Rarog, 1996.

Put' dukhovnogo obnovleniia (The path of spiritual renewal). (1935). Münich: n.p., 1962.

Put' k ochevidnosti (The path to manifest truth). Münich: n.p., 1957.

Religioznyi smysl filosofii (The religious meaning of philosophy). Paris: YMCA Press, 1925.

SECONDARY WORKS

Grier, Philip. "Il'in, Ivan Aleksandrovich." In *Encyclopedia of Philosophy*. Vol. 4. London and New York: Routledge, 1998.

Grier, Philip. "The Speculative Concrete: I. A. Il'in's Interpretation of Hegel." In *Hegel, History, and Interpretation*, edited by Shaun Gallagher, 169–193. Albany: State University of New York Press, 1997.

Lisitsa, Iurii. "O soprotivlenii zlu siloiu pro et contra. Polemika vokrug idei I.A. Il'ina" (On opposing evil with force pro et contra: The controversy over I. A. Il'in's ideas), an appendix to *I. A. Il'in: Sobranie sochinenii v desiati tomakh*. Vol. 5. Moscow: Russkaia Kniga, 1996.

Poltoratzky, N. *I. A. Il'in i polemika vokrug ego idei o soprotivlenii zlu siloiu* (I. A. Il'in and the controversy surrounding his idea of opposing evil with force). London, Canada: Zaria, 1975.

Poltoratzky, N. *Ivan Aleksandrovich Il'in: Zhizn', trudy, mirovozzrenie: Sbornik statei* (Ivan Aleksandrovich Il'in: Life, works, worldview: A collection of articles). Tenafly, NJ: Ermitazh, 1989.

Robinson, P. "On Resistance to Evil by Force: Ivan Il'in and Necessity of War." *Journal of Military Ethics* (2003) 2 (2): 145–159.

St. Basil the Great. *Hexaëmeron*. The Master Christian Library, Version 5, 1997, AGES Software,www.ageslibrary.com

Iurii Lisitsa (2005)

ILLUMINATION

The idea of a divine "illumination" in the mind occurs in both philosophical and religious contexts. Often it forms one of the links between the two types of thought, and sometimes it bears distinctly religious overtones even in its more philosophical applications. This is one of the characteristic features of the theory of illumination in the thought of Plato, where it played, for the first time in its long history, a major part. Plato, like many other thinkers, creative artists, prophets, and mystics, spoke readily of the sudden flash of understanding or insight in the mind as a flood of light (see, for example, his *Seventh Letter,* 341C, 344B). The image is, indeed, one that occurs naturally in many languages and is especially apt for the description of insight thought to have been achieved as a result of external aid of some kind, of an "inspiration." The language of inspiration is based on the entry of breath, and that of illumination on the entry of light into the mind. The Stoic tradition can be said to have developed the former analogy in its metaphysics; Plato was undoubtedly the father of the philosophical tradition to which the analogy of light is fundamental.

In his *Republic,* Plato employed the analogy of light and vision to describe the process of understanding or of knowledge in general (Books V–VIII). The mind's knowledge of the world of intelligible reality, of the forms or ideas, was held to be analogous to the awareness of material objects accessible to the eye's vision when illuminated by the light of the sun. Plato developed a detailed correspondence between physical and intellectual sight (*Republic* 507f.), according to which the mind corresponds to the eye and the form to the physical object seen; an "intellectual light" emanating from the supreme

form, the Good, and pervasive of the whole intelligible world as well as the mind, corresponds to the sun. Understanding, in terms of this analogy, depends on the intellectual illumination of the mind and its objects, just as vision depends on a physical illumination of the eye and its objects.

A theory of this type, in one or another of many variant forms, became an essential part of a vast body of thought cast in Platonic molds. During the Hellenistic and Roman periods it was widely diffused and incorporated into Jewish and Christian thought. In the Hellenized Judaic milieu of Alexandria the divine wisdom was sometimes spoken of in terms of light, for instance, by the author of the book of Wisdom, who referred to it as "an effulgence of eternal light," which he interpreted as an image of God's goodness (7, 26). Thoughts of this kind found a place in the work of Philo and in the prologue to the Fourth Gospel. Middle-Platonist thinkers, such as Albinus, took the step—perhaps already hinted at by Plato in some passages—of placing the forms within a divine mind and, in effect, identifying the "intelligible world" with the mind of God. In this way a long and rich future was prepared for the theory of illumination within the body of Christian thought.

In Christian thought it is in the work of St. Augustine of Hippo that the theory of illumination is found in its most highly developed form. Like Plato, Augustine thought of understanding as analogous to seeing. Understanding, or intellectual sight, was therefore, he held, conditional on illumination, just as physical sight was; only here the light was the intelligible light that emanated from the divine mind and in illuminating the human mind endowed it with understanding. Understanding, in the last resort, was an inward participation of the human mind in the divine. The scope of illumination was further extended, at the cost of precision, in the work of the pseudo-Dionysius. His favorite designation for God, the absolutely transcendent One, was in terms of light. God is the intelligible light beyond all light and the inexhaustibly rich source of brightness that extends to all intelligence. His illuminating activity gathers and reunites all that it touches; it perfects creatures endowed with reason and understanding by uniting them with the one all-pervading light (*De Divinus Nominibus,* IV, 6). In true Neoplatonic fashion, the pseudo-Dionysius conceived of the cosmos as a hierarchically ordered system, descending in order of reality and value from its source, the One. Illumination, in general terms, is the means by which intellectual creatures ascend and return to unity, and the "hierarchy" (understood as extending through both the cosmos and the church) is defined as the divine arrangement whereby all things, participating in their measure in the divine light, are brought back to as close a union with the source of this light as is possible for them (*De Coelestia Hierarchia,* III, 1). In a more special sense, illumination is the second of three phases—namely purification, illumination, and perfection—of man's return to the One. In this more specialized sense the church's sacramental system and the grades in the ecclesiastical hierarchy concerned with its administration are agencies of divine illumination. Illumination is the intermediate stage of approach to God, between initial purification and final perfection (*De Ecclesiastica Hierarchia,* V, 1, 3). In the most restricted sacramental contexts "illumination" thus becomes synonymous, in accordance with an old Christian usage, with "baptism." In the work of the pseudo-Dionysius the theory of illumination was merged with an inclusive conception of the spiritual life formulated in the language of light and illumination.

The reputation enjoyed by Augustine and by the writings of the pseudo-Dionysius in the Middle Ages assured their views a long future. In the thirteenth century the rise of Christian Aristotelianism provided the first serious alternative theory of knowledge. In this there was no place for the intervention of a divine illumination as an essential constituent of knowledge. Knowledge was accounted for entirely in terms of mental activity and its objects, and no reference to God was necessary to explain it. Nevertheless, the *lumen intellectuale* of the mind was held to be a participation in the *lumen divinum* of the divine mind, since God was present everywhere, in the mind no less than in other things. In this way Christian Aristotelians, such as St. Thomas Aquinas, were able to endorse some characteristically Augustinian statements in spite of the fact that their theories of knowledge were built on a radically different structure. The Augustinian version of the theory of illumination continued to have a vogue among some thinkers of the thirteenth century, such as St. Bonaventure, and even later. It found echoes in the thought of some modern philosophers, such as Nicolas Malebranche. Increasingly, however, in the later Middle Ages and after, the language of illumination, especially as elaborated by the pseudo-Dionysius, became the special property of mystical writers and writers on the spiritual life.

See also Alcinous; Augustine, St.; Bonaventure, St.; Malebranche, Nicolas; Plato; Pseudo-Dionysius; Thomas Aquinas, St.

Bibliography

Allers, R. "St. Augustine's Doctrine on Illumination." *Franciscan Studies* 12 (1952): 27–46.

Geach, P. *Mental Acts.* London: Routledge and Paul, 1957. Section 11 and the appendix include corrections of the standard account of Thomas Aquinas's theory of concept formation.

Gersh, Stephen. *From Iamblichus to Erigena: An Investigation of the Prehistory and Evolution of the Pseudo-Dionysian Tradition.* Leiden: Brill, 1978.

Jolivet, R. *Dieu soleil des esprits.* Paris: Desclée de Brouwer, 1934. A study of the theory of illumination.

Markus, R. A. "St. Augustine on Signs." *Phronesis* 2 (1957): 60–83. Includes a discussion of illumination in Augustine's theory of knowledge.

Marrone, Steven. *The Light of Thy Countenance: Science and Knowledge of God in the Thirteenth Century.* Leiden: Brill, 2001.

Pasnau, Robert. "Henry of Ghent and the Twilight of Divine Illumination." *Review of Metaphysics* 49 (1995): 49–75.

Pasnau, Robert. *Thomas Aquinas on Human Nature: A Philosophical Study of Summa Theologiae 1a 75–89.* Cambridge, U.K.: Cambridge University Press, 2002.

R. A. Markus (1967)
Bibliography updated by Christian B. Miller (2005)

ILLUMINATIONISM

Illuminationism (also, Illuminationist philosophy) is the name given to a school of philosophy founded in the twelfth century by the innovative Persian philosopher, Shihāb al-Dīn Suhrawardī (d. 1191), who is well-known by the honorific epithet, "Master of Illumination" (Shaykh al-Ishrāq).

BACKGROUND

The philosophy of Illumination is a holistically constructed system that aims to refine the period's peripatetic philosophy, which was known predominantly in the corpus of philosophical writings by the acclaimed Persian philosopher and scientist, Abū ʿAlī Sīnā, well-known in European traditions as Avicenna, the latinized version of his name. The intense Greek-inspired scientific and philosophical activity from the ninth to the eleventh centuries, centered mainly in Baghdad (the Abbasid Caliphate's political, cultural, and scientific capital), but also in the emerging centers of learning in Iran (such as the cities Rayy, Hamadan, Isfahān, and Nayshpur) as well as central Asian centers of Pesianate.

Linguistic and cultural influence produced remarkable results manifest in many texts covering the range of pure and applied sciences, including medicine, astronomy, mathematics, logic and philosophy, and so on. In this, the creative period of Islamic philosophy, two domains of intellectual endeavor, political philosophy and holistic theoretical philosophy, are defined and creatively expressed in texts that together constitute the dominant side of Islamic philosophy to this day. In practical philosophy the Persian thinker Abu Naṣr Fārābī (875–950)—Abunaser, or Alfarabius in medieval Latin texts, also called "The Second Teacher"—creates seminal works of political philosophy, such as *Opinions of the Inhabitants of the Virtuous City*, where he redefined Greek political philosophy and theorized that human beings could gain access to "prophetic" yet objective knowledge through conjunction with the Active Intellect, not restricted to Divine Will. His political order, legislated by a founding prophet-lawgiver and "scientifically" reformed by learned (*ʿulamā*) guardians, ensured just rule necessary for the universal pursuit of earthly and eternal happiness.

In theoretical philosophy Avicenna's texts, *Healing* (*al-Shifā*); *Directives and Remarks* (*al-Ishārāt wa al-Tanbīhāt*); and *Deliverance* (*al-Najāt*) define Islamic Peripatetic philosophy, which has had the greatest impact on all subsequent philosophical works to this day. This highly creative rationalist philosophical endeavor was, however, seriously curtailed by the antirationalist movement of Ashʾarite theology augmented by the antiphilosophical polemics of the state-sponsored theologian, Abu Ḥāmid Ghazzālī (d. 1111). This is where Illuminationism is critical, for had it not been for the definition and construction of the philosophy of Illumination by Suhrawardī the unbound and creative philosophical endeavor could have died out altogether in the history of Islam. As is, in part due to antirational polemics and fundamentalist religious zeal, much of Islam's intellectual life became confined by structures defined and dictated by Juridical creed.

The impact of such polemics is seen in the philosophical sphere where scholastic philosophical compositions after Avicenna are reduced to the production of "textbooks" (e.g., Athīr al-Dīn Abharī's Guide to Philosophy [*Ḥidāyat al-ḥikma*]) limited by theological presuppositions, whereas philosophy, if allowed, is employed solely as the handmaiden of theology. For awhile the Mongol rule of eastern Islam did allow for a properly free and creative scientific endeavor, which in the philosophical domain is exemplified by noted thinkers who, starting in the thirteenth century, wrote commentaries on Suhrawardī's texts and also composed independent works, some distinctly inspired by the Illuminationist system.

It is in this respect that one can witness an Illuminationist-inspired analytical trend that helped rescue genuine philosophy from deteriorating altogether to dogmatic theology or to ideological mysticism. In part the origins of Illuminationism may be viewed as attempts to respond to antiphilosophical polemics. The Illuminationists' daring philosophical position, however, was that peripatetic philosophy itself needs to be refined and reconstructed to remove a set of presumed logical gaps, and to provide epistemological and other theories to better explain being, knowing, and cosmology. For the most part—specifically in philosophical circles—Aristotle's authority was unquestioned, and Avicenna's work was considered the perfect and consistent Arabic and Persian expression of Aristotelian philosophy. Suhrawardī is among the first philosophers to raise well-reasoned, nonpolemic, and nonideologically driven objections against Aristotelian philosophy. His aim—to refine philosophical arguments by rethinking the set of questions that constitute holistic systems—does lead to novel analysis covering the principles of knowledge, ways of examining being, and of new cosmological constructs. The Illuminationist legacy exemplifies refined rational process, and must not be confused with polemics to refute reason, nor to change reason to subjective, social, and ethical mysticism.

II. ORIGINS AND CONSTRUCTION OF ILLUMINATIONIST PHILOSOPHY

The most important and clearly stipulated aim of the philosophy of Illumination is the construction of a holistic system to define a new method of science, named "Science of Lights"(*ʿilm al-anwār*), a refinement of Aristotelian method, and capable of describing an inclusive range of phenomena where peripatetic theory has been thought to have failed. Suhrawardī's novel ideas are expressed in four major texts that together constitute the new system and form an integral and ordered syllabus on the philosophy of Illumination. They are: the first text, *the Intimations* (*al-Talwīḥāt*); and second its addendum, *the Apposites* (*al-Muqāwamāt*), composed in standard peripatetic structure and language with the aim to present a working synopsis of Avicenna's philosophical system, but also to point out the elements where the Illuminationist position differs from that of the peripatetic and to introduce arguments to prove the former. The third text is *the Paths and Havens* (*al-Mashāriʿ wa al-Muʔāraḥāt*), the longest of Suhrawardī's compositions, in which he presents detailed arguments concerning Illuminationist principles in every domain of philosophical inquiry set against those of the peripatetics, mainly the strictly Avicennan.

The fourth text of the corpus is the text eponymous with the system itself, *the Philosophy of Illumination* (*Ḥikmat al-Ishrāq*), and is the most well-known of all of Suhrawardī's works. This text is the final expression of the new analysis and its systematic construction; it is structured differently than the standard three-part logic, physics, and metaphysics of peripatetic texts, and it employs a constructed symbolic metalanguage named the "Language of Illumination" (*lisān al-ishrāq*). All things pertaining to the domains knowing, being, and cosmology are depicted as lights, where distinction is determined by equivocation—that is, in terms of degrees of the intensity of luminosity. The One origin of the system is the most luminous, hence most self-conscious light, named the Light of Lights, and all other entities are propagated from it in accordance with the increasing sequence 2^n—where n is the rank of the propagated light starting with the First Light—and together they form the continuum luminous whole of reality.

The foundations of the new philosophy commence in logic, where Suhrawardī draws on an earlier twelfth-century Persian thinker, ʿUmar ibn Sahlān Sāvī, and his perhaps Stoic-inspired views in semantics and other parts of logic, and restructures the Peripatetic nine books of the *Organon*. The restructuring of peripatetic work becomes the most apparent distinguishing characteristic of Illuminationist texts since the twelfth century. For example, topics pertaining to semantics, and formal and material logic, plus a novel set of questions on fallacies, are placed together—this for the first time in the history of logic—and given the title "Rules of Thought." There are technical innovations in Illuminationist formal logic, such as reduction of terms; formal redefinitions of the Second and Third Figures of Syllogism as simple inferences based on the First Figure; and the critical reevaluation of negation in simple and compound propositions, where negation is defined as an independent operator that distributes.

The traditional nine books of the Arabic *Organon* are rearranged according to a more well-defined concept of logic as a whole, where expository propositions (the Stoic *logos apophantikos*) are distinguished from proof theory, and indicate a clear view of three-part logic: semantics, formal logic, and material logic. The *Philosophy of Illumination*'s restructured logic is seen as follows: Book Two of the *Organon* (the *Categories*) is removed from logic and a reformulated theory that reduces the number of Aristotelian categories to five—the Stoic four, substance, quality, quantity, relation, plus the fifth, motion, which is

the common continuous category in all existent things—is introduced in physics. Selected subjects introduced in Book One (the *Isagoge*), Book Three (*De Interpretatione*), and Book Six (the *Topics*) are brought together in Section One, titled "On Things Known and On Definitions." Other selected subjects from Book Four (*Prior Analytics*), Book Five (*Posterior Analytics*), and Book Six (the *Topics*) are brought together in the Section Two, which is titled "On Proofs and Their Principles."

Finally, selections from the remaining three Books of the *Organon* (*Sophistical Refutations*, *Rhetoric*, and *Poetics*)—but mostly from Book Seven (*Sophistical Refutations*)—are brought together in the Third Section, which is titled "On Sophistical Refutations and Disputations On the Validity of Illuminationist Principles Vs. the Peripatetic Principles," which further includes subjects traditionally treated in other Aristotelian texts (e.g., selections from *De Anima*; questions on the physics of sight and sound; a critique of the Aristotelian dyad Prime Matter-Form; discussion of Platonic Forms; plus a novel discussion of subjects best described as foundations of mathematics). The Three Sections together are placed in Part One, titled "On Rules of Thought," from the book *Philosophy of Illumination*.

THE PARAMOUNT PROBLEM IN ILLUMINATIONIST PHILOSOPHY. The most important philosophical problem in which Illuminationist philosophy diverges from the peripatetics concerns the epistemology of obtaining primary principles and the first step taken in the construction of scientific systems. The Illuminationist position argues that: (1) the first step in science cannot be demonstrated based on the construction of essentialist definitions (*al-Ḥadd al-tāmm*); (2) laws of science cannot be formulated as universal affirmative propositions (because of future contingency there may be always elements discovered that negate universality); and (3) the peripatetic conjunction with the Active Intellect is a false position or law.

Suhrawardī argues in his "destruction" (*hadam*) of the peripatetic formula that the essentialist definition is based on (1) an elaborate critique of predication aimed at rejecting it as tautological; and (2) the impossibility of counting each and every member of the constituents of the thing to be defined, a condition that must be met for the peripatetic essentialist definition to indicate the essence of the *definiendum*, which is similar the impossibility of a definition by extension. The alternative, as stipulated by Suhrawardī, is that primary principles must be known by "other" ways, which is then stipulated to be an immediate intuitive mode. Suhrawardī's Illuminationist critique of predication may be summed this way: to say "x is y" without knowing the essence of x prior to the predication does not inform of anything other than a change in terms x to y without added signification. Moreover, x includes $\{x_i\}$, then for the predicative definition to inform of the essence, y must be identical to Σx_i, which is not possible as {xi} may be uncountible, or unbound.

The peripatetic position, based on the Stagirites's own view stipulated in many of his texts, was that primary principles may be known through the cognitive mode named "immediate knowledge" but Suhrawardī argues that Aristotle's position on immediate knowledge had not been fully explained and was left ambiguous. This point is best exemplified in early passages of the *Posterior Analytics*, I.2: 71b.20–72a.25, which may be summed up as follows: Science rests on necessary, true, primary, and most prior premises, which are known not through syllogistic demonstration, but by an "immediate," intuitive way. The Illuminationist position, however, is that Aristotle does not systematically present what is the intuitive, immediate cognitive mode; that he does not discuss an epistemological well-structured process that could describe primary intuition; and that he leaves this question in an ambiguous state—because Aristotle refers to immediate knowledge as "opinion" (*doxa*) in his works. Suhrawardī's Illuminationist construction of a unified epistemological theory, named "Knowledge by Presence," is claimed to resolve the ambiguity in Aristotle's position, and Suhrawardī is acclaimed for having, for the first time in Islamic philosophy, described intuitive knowledge in a systematic, "scientific" way.

The Illuminationist ontological position, called "primacy of quiddity," distinguishes philosophical schools in the development of Islamic philosophy in Iran up to the present day. It is also a matter of considerable controversy. Those who believe in the primacy of being, or existence (*wujūd*), consider essence (*māhiyya*) to be a derived, mental concept (*amr iʿtibārī*, a term of secondary intention), whereas those who believe in the primacy of quiddity consider existence to be a derived, mental concept. The Illuminationist position is this: if existence is real outside the mind (*mutaḥaqqaq fī khārij al-dhihn*), then the real must consist of two things—the principle of the reality of existence, and the being of existence, which requires a referent outside the mind (*miṣdāq fī khārij al-dhihn*). And its referent outside the mind must also consist of two things, which are subdivided, and so on, ad infinitum. This is clearly absurd. Therefore existence must be considered an abstract, derived, mental concept.

III. SUMMARY OF THE MAIN TOPICS OF THE ILLUMINATIONIST HOLISTIC SYSTEM

(1) Principles of knowledge, and the first step in science, rest on the primary and immediate intuitive cognitive mode. This knowledge is of essence, is prepropositional, and rests on the atemporal Illuminationist relation between the self-conscious knower (*mudrik*) and the essentially knowable thing—the object of knowledge—the known (*mudrak*). This "relation" between the knower and the known, or knowing and being, is an identity preserving "sameness" and replaces the peripatetic principle of "conjunction" between the elevated human intellect and the Active Intellect.

(2) Reality is a continuum of monad-like "light" entities that are distinguished only by equivocation in terms of degrees of "luminousity" (*nuriyya, istināra*). Self-consciousness is an essential specific aspect of all lights determining rank of each and every entity propagated from the One source, the Light of Lights. All entities are propagated according to the sequence 2^n, where n is the ordered rank. Consciousness and degrees of abstraction from material extension decrease as n increases, and are associated with each and every member of the Whole (*al-kull*), which is also conscious of self.

(3) There is a two-fold process, "vision-illumination" (*mushāhada-ishrāq*), that acts on all levels of reality. In the corporeal realm of sense-perception, the process acts as sight (*ibṣār*). The eye (*al-baṣar*; or the seeing subject, *al-bāṣir*), when capable of seeing, sees an object (*al-mubṣar*) when the object itself is illuminated (*mustanīr*). In the incorporeal realm every "abstract light" "sees" the "lights" that are above it in rank, whereas the higher illuminates it instantaneously, at the moment of vision. The Light of Lights (*Nur al-anwār*) illuminates everything. Knowledge is obtained through this "coupled" activity of vision-illumination, and the impetus underlying the operation of this principle is self-consciousness. Thus every being comes to know its own degree of perfection, an act of self-knowledge that induces a desire (*shawq*) to "see" the being just above it in perfection, and this act of "seeing" triggers the process of illumination. By means of the process of illumination, "light " is propagated from its highest origin to the lowest elements.

(4) The Illuminationist cosmos adds a fourth realm of being to the standard three—Intellect, Soul, and Matter—of the peripatetic named "Mundus Imaginalis" (al-ʿālam al-khayāl), and is a boundary between the intellect and the soul. This realm, described as the "essence of wonders" (dhāt al-ʿajāʾib), is the veritable wonderland of visionary experience as described in the Illuminationist allegorical recitals, where time and space are different from time as measure and euclidean space. Movement into this realm brings about qualitative change described in amazing allegorical tales.

IV. ILLUMINATIONIST PHILOSOPHY AFTER SUHRAWARDĪ

The Illuminationist system continues after Suhrawardī's execution in 1191 through the composition of scholastic commentaries on his texts, and also by the gradual creation of independent work in the Illuminationist tradition by a number of leading philosophers, and thus gains widespread acceptance in scholastic centers of learning in Iran. There are two ways in which Illuminationist philosophy continues. Firstly, the thirteenth-century Persian philosopher and historian of philosophy, Shams al-Dīn Shahrazurī in his commentaries on Suhrawardī's texts, *Commentary on the Philosophy of Illumination* (*Sharḥ Ḥikmat al-Ishrāq*); *Commentary on the Intimations* (*Sharḥ al-Talwīḥāt*); and in his independent magnum opus encyclopedic text *The Metaphysical Tree* (*al-Shajara al-Ilāhiyya*), emphasizes the symbolic, and the distinctly nonperipatetic components of Illuminationist philosophy. He also further extends and greatly embellishes the inspirational, allegorical, and fantastic side of Illuminationist texts. Secondly, later in the thirteenth century the well known Jewish philosopher and occulist of Baghdad, Saʿd ibn Manṣur Ibn Kammunam in his *Commentry on the Intimations* (*al-Tanqīḥāt fī Sharḥ al-Talwīḥāt*), and in his major independent philosophical work, *The New Philosophy* (*al-Jadīd fī al-Ḥikma*), as well as in his shorter works, such as *Treatise on the Soul* (*Risāla fi al-Nafs*) emphasizes the purely discursive and systematically philosophical side of Illuminationist Philosophy.

The most philosophically important impact of the Illuminationist system is seen in the latest, creative, and holistic work in Islamic philosophy. This is the constructed system named "Metaphysical Philosophy" (*al-Ḥikma al-Mutaʿāliya*) by the famous Persian thinker Ṣadr al-Dīn al-Shīrāzī, best known as Mullā Ṣadrā (d. 1640). The most widely studied text by Mullā Ṣadrā is his *The Four Intellectual Journeys* (*al-Asfār al-Arbaʿa al-ʿAqliyya*), where in almost the entire range of philosophical investigation the author draws heavily from the Illumina-

tionist tradition—this after the texts of Avicenna and Suhrawardī are first carefully analyzed and then problems and arguments are reconstructed usually along the systematic principles of the Illuminationist system. The most enduring impact of Illuminationist systematic philosophy is in the domain of epistemology, where Mullā Ṣadrā adopts and refines Suhrawardī's unified theory Knowledge by Presence to discuss, among other things, God's knowledge, and the "scientific" validity of inspirational knowledge as well as of revelation. Mullā Ṣadrā's discussion of the proposition "sameness of knowing and being," or "unity of the knower and the known" (*ittiḥād al-ʿāqil wa al-maʿqūl*), is distinctly Illuminationist, whereas sameness, or "unity," is nonpredicative.

See also Mullā Ṣadrā; Suhrawardī, Shihāb al-Dīn Yahyā; School of Qom.

Bibliography

Corbin, Henry. *Archange empourpré, Quinze traités et récits mystiques traduits du persan et de l'arabe, présentés et annotés par Henry Corbin*. Paris: Fayard, 1976.

Corbin, Henry. *Livre de la Sagesse Orientale, Kitāb Ḥikmat al-Ishrāq*. Translation and notes by Henry Corbin, with an introduction by Christian Jambet. Paris: Verdier, 1986.

Haʾiri Yazdi, M. *The Principles of Epistemology in Islamic Philosophy: Knowledge by Presence*. New York: SUNY Press, 1992.

Ibn Kammuna. *al-Tanqīḥāt fī Sharḥ al-Talwīḥat*, edited, with introduction and analysis by Hossein Ziai and Ahmed Alwishah. Costa Mesa, CA: Mazda, 2003.

Shahrazurī, Shams al-Dīn. *Sharḥ Ḥikmat al-Ishrāq* (Commentary on the philosophy of illumination), edited with introduction and notes by Hossein Ziai. Tehran, Iran: Institute for Cultural Studies, no. 736, 1993. 2nd ed., 2001.

Suhrawardī, Shihāb al-Dīn *The Book of Radiance: A Parallel English-Persian Text*, edited and translated, with introduction by Hossein Ziai. Costa Mesa, CA: Mazda, 1998.

Suhrawardī, Shihāb al-Dīn. *The Philosophy of Illumination: A New Critical Edition of the Text of Ḥikmat al-ishrāq*. Translation, Notes, Commentary, and Introduction by John Walbridge and Hossein Ziai. Provo, UT: BYU Press, 1999.

Thackston, W. M., Jr. *The Philosophical and Mystical and Visionary Treatises of Shihabuddin Yahya Suhrawardi*. Costa Mesa, CA: Mazda Publishers, 2001.

Walbridge, John. *The Science of Mystic Lights: Qu?b al-Dīn Shīrāzī and the Illuminationist Tradition in Islamic Philosophy*. Cambridge, MA: Harvard Middle Eastern Monographs 26, Harvard University Press, 1992.

Ziai, Hossein. "al-Suhrawardī, Shihāb al-Dīn." In *Encyclopedia of Islam*, 2nd ed., 781–784. Leiden: Brill.

Ziai, Hossein. "Beyond Philosophy: Suhrawardī's Illuminationist Path to Wisdom." In *Myth and Philosophy*, edited by Frank E. Reynolds and David Tracy, chap. 8. New York: State University of New York Press, 1990.

Ziai, Hossein. *Knowledge and Illumination: A Study of Sohravardī's Ḥikmat al-Isahrāq*. Brown Judaic Series, no. 96. Atlanta: Scholars Press, 1990.

Ziai, Hossein. "Shihāb al-Dīn Yaḥyā Suhrawardī." In *The Routledge History of Islamic Philosophy*, edited by Sayyed Nasr and Oliver Leaman, chap. 28. London: Routledge, 1995.

Ziai, Hossein. "Source and Nature of Authority: A Study of Suhrawardī's Illuminationist Political Doctrine." In *The Political Aspects of Islamic Philosophy*, edited by Charles Butterworth, 304–344. Cambridge, MA: Harvard University Press, 1992.

Ziai, Hossein. "The Illuminationist Tradition." In *The Routledge History of Islamic Philosophy*, edited by Sayyed Nasr and Oliver Leaman, chap. 29. London: Routledge, 1995.

Hossein Ziai (2005)

ILLUSIONS

Most of the major philosophical problems of perception derive from the fact of "illusions." These problems center on the question whether perception can give us true and direct knowledge of the world, and thus they are basic to epistemology. This entry will describe illusions and set forth and examine the argument from illusion that perception cannot be trusted as a source of knowledge of the external world but affords direct awareness only of appearances or sensa.

THREE KINDS OF ILLUSORY EXPERIENCE

The term *illusion* is used by philosophers to cover a range of phenomena approximately classifiable as follows.

ILLUSIONS PROPER. Illusions proper occur when the percipient is deceived or is liable to be deceived in identifying the object perceived or its properties. Psychologists have produced a number of optical illusions, such as equal lines that appear to be of unequal length; a stationary balloon that when inflated and then deflated seems to advance and then recede; and a specially constructed Distorted Room, in which a man looks smaller than a boy. Diseases or drugs, including alcohol, may produce other illusions, such as double images or the unearthly colors and multiple shapes an object may assume for one who has taken mescaline. Other examples are mirages, mirror effects, and conjurers' tricks. The perception of motion introduces many more: At the cinema a rapid succession of slightly different stills on a flat screen makes us see a scene with a three-dimensional perspective in which people move about; the wheels of a coach may seem to be

going backward when really they are moving rapidly forward (stroboscopic effect).

RELATIVITY OF PERCEPTIONS. A round plate that looks elliptical when seen from an angle and a square table that looks diamond shaped illustrate the relativity of perception. The same water may feel cool to one person and warm to another; the same wine may taste sweet or dry, depending on what one has just been eating; green hills may look blue in the distance; and as a train rushes past, the pitch of its whistle may seem to vary. Further examples are color blindness, shortsightedness, and other physical defects that alter the appearance of things. In all these cases the apparent properties of an object vary relative to the position of the percipient, the distance and media between him and the object, the lighting, the state of his health, body, or sense organs, etc. These are not strictly illusions (they usually do not deceive), and they vary around a norm in which the objects are perceived accurately.

HALLUCINATIONS. In pure hallucinations—for example, the pink elephant a drunkard sees, the apparitions of delirium, Macbeth's dagger—some physical object is "perceived" when neither it nor anything at all like it is present. In contrast are illusions where the mistake is about the properties, position, or identity of some object actually in view.

Some, perhaps even most, hallucinations are triggered by some perceived feature of a very different character; for instance, a beam of light may be taken to be a person. Many hallucinations are integrated; they fit well with the real background, cast shadows, and vary in size and perspective as they move. One may also class phantom limbs as hallucinations. Pain or other sensations are felt "in the toes," for example, of a leg that has been amputated—the victim still feels he possesses the missing limb.

ARGUMENT FROM ILLUSION

The main aim of the argument from illusion is to show by means of illusions that the senses are not to be trusted and that perception is not direct and certain awareness of the real properties of material objects but awareness of appearances only. In fact, this argument involves three subarguments.

(A) A SKEPTICAL CLAIM. However sure we are about our perceiving, it is always possible that we are being deceived by one of the many kinds of illusion or hallucination, since it is characteristic of such states that we cannot tell that we are suffering from them. This may in practice be a negligible possibility, but philosophy is concerned with the highest standard of exactitude, and from this strict position perceiving is not absolutely certain because there is always some theoretical possibility of error. Various conclusions can then be drawn. One is that for certain knowledge we must rely not on the senses but on some other faculty, such as intellectual intuition (as in René Descartes); another is that we must abandon common-sense realism.

(B) NATURE OF APPEARANCES. In all these illusions there is some thing or quality that does not coincide with the object or object-properties that are in fact present—for example, the apparitions of hallucinations, the elliptical appearance we see when we look at a round plate, the black shape the color-blind person sees when looking at a red box, the oasis of a mirage, and the second bottle in double vision. All these are merely appearances and cannot be identified with real objects or properties. What then are these appearances? In some cases, and probably in all, they must be *sensa*, private, probably mental, objects of awareness quite distinct from external material objects, although no doubt they are caused by or resemble material objects.

(C) SIGNIFICANCE OF CONTINUITY. If one were to change from seeing an appearance, a private and transitory sensum, to seeing a public, enduring physical object ("public" meaning observable by several persons at one time), one would expect a sudden change in the character of one's sensory experience. But no such jump occurs: There is normally an unbroken continuity between situations where we cannot actually be seeing the material object but are aware only of appearances and situations where we think we see the material object. As we move from where the plate looks elliptical to where it looks round, or as the drunkard looks first at the pink rat and then at the real bed on which it sits, there is a smooth transition. Consequently, even in these seemingly genuine or veridical perceptions we must also be aware of appearances or sensa and not directly of the object itself.

We may note three things concerning our subarguments: (1) Argument (*b*), unlike (*a*), does not depend on there being error; even if one is not deceived by perspectival distortion, double vision, and so on, the argument that what is really perceived must be sensa is unaffected. (2) The claim in (*b*), that the appearances are private and mental existents, depends to some extent on considerations of continuity. Almost all hallucinations, the dark

shapes a color-blind person sees or the results of diseases and drugs, are plausibly private to the percipient. But simple perspectival distortions will be private only to the viewpoint. For instance, the elliptical appearance of the plate is as public as the round one in that many may see it at once; this holds similarly for mirages and reflections. Unless causal considerations are introduced, the supposition that each person is then seeing a numerically different but qualitatively similar elliptical appearance or sensum must rely partly on similarity with cases where the content of illusion is undeniably private and partly on the assumption that if the plate is round, then the elliptical appearance must be something other than the plate; but these are hardly compelling grounds. (3) The charge may be made, How do we know that the plate is round or what its real color is? These points would normally be settled by measurement or by reference to standard lighting conditions, but the argument does not rely on this. To take the plate example, it may be put thus: The plate looks elliptical to *A* and round to *B*; it cannot be both round and elliptical, for that would be a self-contradiction; therefore, one of these appearances at least must be quite distinct from the plate—and perhaps (by continuity) both are.

CRITICISM OF THE ARGUMENT FROM ILLUSION

The argument from illusion can be countered in various ways.

CERTAINTY. The skeptical claim is often met by stressing the comparative rarity of illusion and the efficacy of the various tests that can be made to remove doubt. We can use one sense to help another. For example, wax fruit may look like real fruit, but touch and taste reveal it; sight, memory, and testimony can show that a phantom limb does not exist; measurement can settle the real shape of an object; confirmation from others can show up many hallucinations, though there are some group hallucinations; we soon learn to discount alcohol and drugs and may generally argue from known causal factors present. But although these tests reduce the possibility of error in a tested perception to extremely slight and in practice negligible proportions, the critic will still say that it is not absolutely certain and that only absolute certainty will satisfy the philosopher. To this there are two replies. (1) It is logically impossible that we suffer from hallucinations all the time; if no perception were ever certain, then there would be no way of distinguishing hallucinations and illusions from normal perception. (2) The skeptic is misusing the word *certain*; well-tested perceptions are just the things we refer to as certain. If we say they are only probable we destroy the normal useful distinction between *certain* and *probable*. If nothing is certain, the word has no meaning, and we shall just have to invent a new term for that ordinary distinction.

We may comment on these replies. Reply (1) is of no help in deciding whether any particular perception is certain or not—which is one of the main points—and anyhow, a merely approximate certainty would serve to distinguish perceptions from hallucinations. Reply (2) seems to depend on confusing meaning and reference. It is true that perceptions are things we refer to as certain, but that may only be due to our ignorance of the possibility of illusions. The normal meaning of "certain"—without any possibility of doubt—is correctly adopted by the skeptic; he merely argues that it may only be used of the results of intuition or of mathematical demonstration, not of perception; that is he differs only as to the referents of the word. Also, he can still distinguish between "probable" and "practically certain" in perceptual statements. However, a modified reply to the skeptic may be made (3) that he is in fact limiting the word *certain* to cases of logical necessity, to those that it is self-contradictory to deny. This limitation not only has the practical disadvantage of destroying the ordinary certainty-probability distinction but also rules out a priori the possibility of any perceptual statement's being certain; thus the lack of certainty in perception is due not to any defect in perceiving but simply to its not being something quite different from what it is, namely intuition or entailment. It would therefore be much more appropriate to use a relaxed standard in dealing with perception and to allow a perceptual statement to be regarded as certain if it has passed all conceivable or all recognized tests. At any rate, there is no reason to suppose that ordinary perceptions are uncertain in the way that the result of a horse race or the nature of next year's weather is uncertain.

HALLUCINATIONS. The argument from the significance of continuity claims (1) that hallucinations are private sensa, not public material objects, and (2) that since they are indistinguishable from the objects of perceptual consciousness, especially when integrated with them, the latter must also be groups of sensa—representations, perhaps of external objects.

(*a*) One answer, based on the usual psychological account of hallucinations, would be that they are not sensa but mental images of an unusually vivid type that are confused with normal perception. To

meet point (2), the unusual vividness and the lack of normal discrimination may be stressed and explained by the special circumstances in which almost all hallucinations occur, as when the victim is suffering from fever, drunkenness, drugs, starvation, religious ecstasy, or madness or is influenced by lesser factors, such as fear, acute anxiety, or drowsiness. (In the hallucinations of mescaline the person's mental powers are unimpaired, but he usually recognizes the hallucinations as such and is not deceived into thinking they are real.) It is questionable whether these factors, especially the lesser ones, can account for the integration and triggering of hallucinations—cases in which the continuity argument is strong and imagery would seem to merge with genuine perceptions. Also, to be complete this answer would need to offer an explanation of the nature of mental imagery and of why it resembles perceiving. Probably imagery depends on reactivation of the kinds of brain and nervous activity that occur in perception (or in action, if it is motor imagery), and the occurrence of such activity can be detected during the imagery. However, this involves the causal processes, study of which leads by a different route to the abandonment of commonsense theories.

(*b*) It has been pointed out, by J. L. Austin, for example, that the argument from illusion as applied to hallucinations relies on certain dubious assumptions, namely, that if two things (i.e., an object of genuine perception and an object of hallucination) are not generically the same they cannot look alike, and that they cannot be distinguishable if we in fact fail to distinguish them. The special circumstances cited in point (*a*) may come in here as providing reasons for the victim's failing to distinguish what are in fact distinguishable and quite different experiences. This criticism certainly undermines the argument from illusion as a demonstration; for it to be that, these assumptions would have to be accepted as universally true. But it can be replied that an explanation is still required for the general similarity between the two things (sufficiently close a similarity for people suffering only from anxiety to confuse them); also, we need some general theory of the nature of hallucinations and of their integration and triggering.

Phantom limbs are not covered by these points: One can hardly say that the pains and sensations involved are images of genuine ones—they are genuine enough—nor are the victims suffering from drugs or delirium. The usual physiological explanation is that the nerves from the toes, for example, remain in the untouched part of the limb and, being irritated at the stump, send impulses to the brain similar to those they would send if the toes were being crushed or the pain and other receptors in the toes were being otherwise stimulated. This seems to confirm that pain and somatic sensations are private sensa and accords with the general causal theory of representative realism. But it is still arguable that such sensations are very different from sight and hearing, so that nothing follows about the nature of the latter.

ILLUSIONS AND RELATIVITY. The argument from the nature of appearances relies on the odd assumption that things cannot look other than they are, that when one apparently sees as elliptical a plate that is actually round, then one cannot really be seeing the plate; one is seeing something, an appearance or sensum, which, being elliptical, cannot be the round plate. But one can simply deny the assumption and say that one is in fact seeing a round plate from such a position that it looks elliptical; its elliptical appearance is not some entity different from it. To treat appearances as entities, as though they were things, is a quite unjustified reification; when we speak of the appearance of something we speak of how it, the original object, appears, not of some other object distinct from it. One may confirm this point by noting that the elliptical shape will appear on a photograph too, so that it cannot be subjective or mental.

The same answer may be applied to the various examples of relativity and illusion. The distant green mountains are actually seen but look blue and may be so photographed; in the Doppler effect we still hear the whistle (a "public" noise), and its apparent variation in pitch may be recorded on tape. In the optical illusions we are still seeing lines on paper, balloons, or a man and a boy, and cameras will photograph them with their deceptive appearances. Again, as Austin has shown in detail, in refraction and reflection we still see the object—the face in the mirror or the stick in the water; even in a mirage we see a real oasis, though it appears many miles nearer than it actually is. It may also be claimed that the color-blind man sees the red box, even if it looks black to him, and that the man with double vision sees the one bottle, but it looks double to him. (This last point is more dubious: It may be said that looking double is not like looking blue for it involves an extra apparent object and not a dif-

fering quality of the one object. On the other hand, the percipient is not seeing two bottles in the same way normal people see one: The bottle and the background have a doubled, slightly defocused appearance that perhaps makes it reasonable to say they look double.)

For some people this general answer is immediately convincing, and it seems incredible that the argument from illusion was ever taken seriously. But others protest that it is inadequate and neglects the immediacy of perception; in the various situations mentioned they seem clearly and directly to be aware of an elliptical shape or a blue expanse of mountain, an advancing balloon, two bottles, or, if color-blind, a black box-shaped expanse. Thus to be told that they are aware only of a round plate, green mountain, stationary balloon, one bottle, and so on, is to them unconvincing and fails to do justice to the facts of experience. This feeling for the immediacy of sensory awareness, and the belief that confrontation is so direct that its apparent object must exist as perceived, is at the bottom of the sense-datum theory. The alternative is to dismiss as illusory this apparent direct and mistake-proof confrontation in perception; perceiving is variable in quality, is affected in its accuracy by position, distance, and many other factors, and may thus be inefficient. It is more plausible to suppose that position, distance, and media distort perception of a round or green object or that color blindness and shortsightedness prevent one from seeing it properly than to suppose that these factors give one excellent and perfect awareness of some elliptical or blue sensum different from the object.

But this is not a final answer, for if one then seeks to discover how these factors affect the quality of perception, one has to go into scientific details. Angle of sight varies the pattern of light striking the eye, refraction or reflection bends the light rays, dust scatters them or absorbs some frequencies rather than others, drugs affect the activity of the nervous system, lack of certain retinal pigments alters the eye's response to light, the Distorted Room and other optical illusions rely in their effects on misleading cues. In short, the effects of illusion point beyond themselves to the causal and psychological processes that underlie perception and constitute its most serious theoretical problem.

See also Perception; Realism; Sensa.

Bibliography

The argument from illusion is a stock feature of introductions to philosophy and is dealt with in most of the books in the bibliographies for the entries Sensa, Realism, and Perception. The books listed below give a fuller treatment of the subject.

Henry Habberley Price's *Perception* (London: Methuen, 1932) and Charlie Dunbar Broad's *Scientific Thought* (London: K. Paul, Trench, Trubner, 1923) and *The Mind and Its Place in Nature* (London: Kegan Paul, 1923) develop the argument to support sense data. Alfred Jules Ayer's *Foundations of Empirical Knowledge* (London: Macmillan, 1940) develops it to support the introduction of a "sense-datum language." Price has some second thoughts on perspectival distortion in his article "Illusions," in Volume III of *Contemporary British Philosophy*, edited by Hywel David Lewis (London, 1956), and Ayer's views are presented in his clear introductory work *The Problem of Knowledge* (London: Macmillan, 1956).

John Langshaw Austin, *Sense and Sensibilia* (London: Oxford University Press, 1962), D. M. Armstrong, *Perception and the Physical World* (London: Routledge and Kegan Paul, 1961), and Rodney Julian Hirst, *The Problems of Perception* (London: Allen and Unwin, 1959), are critical of the argument from illusion. See also George McCreath Wyburn, Ralph William Pickford, and R. J. Hirst, *Human Senses and Perception* (Toronto: University of Toronto Press, 1964), which, besides criticism, gives some physiological and psychological details. For other attempts to meet the facts of illusion without postulating sensa, see the works listed under "Perspective Realism and the Like" in the bibliography for Realism and under "Sense Data" in the bibliography for Sensa.

The neurologists John Raymond Smythies, in *Analysis of Perception* (London: Routledge and Paul, 1956), and Russell Brain, in *The Nature of Experience* (London: Oxford University Press, 1959), give some useful references and nontechnical information on hallucinations. The psychological details of illusion in general can be found in any good introductory textbook on psychology, such as David Krech and Richard S. Crutchfield, *Elements of Psychology* (New York: Knopf, 1958).

OTHER RECOMMENDED TITLES

Alston, William P. *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press, 1993.

Chisholm, Roderick. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.

Crane, Tim. "The Waterfall Illusion (1988)." In *Essays on Nonconceptual Content*, edited by York H. Gunther. Cambridge, MA: MIT Press, 2003.

Dancy, Jonathan. "Arguments from Illusion." *Philosophical Quarterly* 45 (1995): 421–438.

Dretske, Fred. *Seeing and Knowing*. London: Routledge and Kegan Paul, 1969.

Firth, Roderick. "Austin and the Argument from Illusion." *Philosophical Review* 73 (1964): 372–382.

Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.

Kenny, A. "The Argument from Illusion in Aristotle's Metaphysics, 1009–10." *Mind* 76 (1967): 184–197.

Reynolds, Steven. "The Argument from Illusion." *Nous* 34 (2000): 604–621.

Robinson, Howard. "The Objects of Perceptual Experience—II." *Proceedings of the Aristotelian Society*, supplement 64 (1990): 151–166.

R. J. Hirst (1967)
Bibliography updated by Benjamin Fiedor (2005)

IMAGERY, MENTAL

In many ways, mental imagery has been a fundamental issue in the history of philosophy. At least since Aristotle, philosophers have argued that knowledge is often represented in the form of mental images, taken to be inner pictures of some sort. However, questions have frequently been raised about the capacity of such images to play roles in thinking, remembering, and imagining; for instance, in George Berkeley's well-known doubts about the possibility of general or abstract images. Debates about mental imagery have been important in the history of psychology as well. Because the images in question are the bearers of conscious experience, claims about them have often been made on the basis of introspection, and the rejection of introspection in favor of behavioral studies was central to the emergence of psychology as a science. However, with the rise of cognitive science, quantified behavioral research has put mental images back on the map.

For example, Roger Shepard and his colleagues (1982) asked subjects to determine whether one geometrical figure matched another, the overall orientation of which was tilted relative to that of the original figure. Reaction times were a linear function of the angle of the tilt: The greater the displacement between the two otherwise identical figures, the longer it took subjects to respond. The implication is that reaction time depends on an operation such as rotating one of the perceived figures through space. Assuming a constant rotation rate, time to respond will depend on the distance through which the figure is rotated. One conclusion that can be drawn is that imaging is like perceiving, because matching rotating objects or figures in perception is similarly governed by a time-to-distance law.

This *perceptual similitude* thesis is an important part of pictorialist theories, according to which images are like mental pictures. It is particularly important on Stephen Kosslyn's account, the most fully developed version of pictorialism. In one well-known experiment, Kosslyn asked subjects to visualize a map they had previously studied and to focus on one of several items represented on the map (e.g. a hut, a pond, a tree). Subjects were then asked to say whether various items were located on the map. Reaction times were a linear function of the distance between the original focal point and the identified item. This suggests that they were scanning a mental map and not simply accessing a description or list. While reaction times might be due to the position of terms on a list, given the initial conditions of free study, there is no reason to think that the locations would be listed systematically by their proximity to the focal point, with the nearest first, the farthest last, and so on.

In Kosslyn's *Image and Mind* (1980), scanning and other operations, such as panning and zooming, are defined as functions that could be performed by a digital computer. This use of a computer model illustrates why theories of mental imagery do not need to treat the mind as an immaterial substance or entail a homunculus to view the inner pictures. Forming and accessing images can be explained in terms of more basic level operations, which can themselves be further decomposed into fundamental processes that a machine could perform.

The same is true for the mental sentences posited by descriptionalist theories of imagery, which constitute the opposing camp. The best-known of these has been developed by Zenon Pylyshyn in *Computation and Cognition* (1984). His argument has two parts. First, he claims that evidence shows that imaging is cognitively penetrable. It is influenced by background knowledge and belief. Therefore, he argues, there cannot be perception-like processes of the sort that pictorialism requires; that is, generic operations such as scanning or rotating at a standard rate that are part of a fixed functional architecture employed similarly across imaging tasks. Second, he maintains that what is a vice for the pictorialist is a virtue on the descriptionalist account. The data on reaction times can be explained, he argues, precisely in terms of the effects of tacit knowledge in the face of experimental task demands. This tacit knowledge is expressed in language-like representations and operates through the production of the descriptions in which imaging consists.

For instance, when four-year-old children were shown an inclined beaker containing colored liquid and then later asked to draw it, they typically drew the fluid level as perpendicular to the sides of the beaker. The implication is that the children's memory images, upon which the drawings are based, are not simply pictures that reproduce the perceived object, and the images reflect that young children do not possess an understanding of geocentric level. Extending this analysis to rotation and scanning studies, Pylyshyn argues that the results can be explained in terms of task demands: Subjects are led to

believe that, in visualizing objects, they are to replicate the process of perceiving the objects. Knowing that perceived object rotations must obey a time-to-distance law, they reproduce the relevant reaction times, although not necessarily with conscious intent.

However, a number of objections have been made to these claims. First, it is sometimes argued that Pylyshyn's descriptionalist view makes images epiphenomenal, giving them no role in causal explanations of behavior. Of course, if such images can be identified with the underlying data structures that take a descriptive form, the charge is not strictly correct. Nonetheless, such a construal will not explain the phenomenal properties of conscious imagery, which are thus excluded from scientific accounts.

Second, not all reaction time studies can be explained in terms of task demands, a point that Pylyshyn now concedes. Moreover, because imaging is affected by background knowledge, it need not be taken to undercut explanations in terms of a basic set of perceptual operations. Kosslyn agrees that imaging is cognitively penetrable. He notes, for example, that the rate of scanning may vary across individuals or tasks. However, that does not mean that scanning cannot be defined in terms of standard operations—such as shifting attention incrementally—or that the employment of those operations is not governed by law-like generalizations. Scanning might be one of a fixed set of operations available to everyone—even if it is not always used—and it can exhibit regularities, despite the effects of knowledge and belief. For example, it can be assumed to occur at a constant rate within individual subjects on a given task.

Nonetheless, a positive account must be given of the knowledge effects that imaging does display, and this requires more than an appeal to perceptual similitude. Thus Kosslyn argues that imaging occurs in a *visual buffer*, a distinctively spatial medium analogous to an internal computer monitor. Although the representations on such a screen will be composed of distinct elements—such as cells in a matrix that can be labeled—the images are said to be pictorial, in the sense that spatial properties of objects are represented by the spatial properties of the medium.

Originally posited as part of Kosslyn's computational model, this visual buffer is identified in his *Image and Brain* (1994) with topographically organized areas of visual cortex. In topographic representations, the features of an object can be distorted. Nonetheless, spatiallydefined regions of the medium will correspond systematically to spatial regions of the object. Moreover, unlike descriptions, such images have the property that the farther apart two points appear to be on an object, the more representational elements there will be between representations of the points. Although these elements need not be closely contiguous, they cannot be just anywhere. If two points appear to be adjacent in a represented object, then the elements that represent them must be—at least in an extended sense—adjacent as well. Several types of evidence from brain research can now be cited in support of the pictorialist view; for instance, lesions to the visual system cause subjects to be unaware of one side of the visual field in imaging, just as they do in perception.

One objection often made to pictorialism is that mental pictures lack the syntactical regularities that would allow them to express thoughts precisely. Sentences can be used to single out certain types of information while ignoring others, but pictures will inevitably represent features that are irrelevant to the task at hand. Thus Daniel Dennett (1981) has argued that imagining cannot be mental picturing, because the former can be more indeterminate than the latter. On the one hand, it is possible to imagine a striped tiger without envisioning it as having a definite number of stripes. On the other hand, it is impossible to depict a striped tiger without showing the number of stripes that it has. However, this line of argument commits what Ned Block (1983) has called the "photographic fallacy." It assumes that pictures cannot employ selective devices; there are actually several ways in which pictures can omit details, Block argues (e.g., by virtue of viewpoint, occlusion, atmospheric blurriness, or schematization). Moreover, the argument from indeterminacy can be turned around. In *The Imagery Debate* (1991), Michael Tye has argued that there are certain kinds of corollary or implicit information that both pictures and images inevitably carry. For instance, any perceptual or imagistic representation of two objects, A and B, will necessarily represent an apparent direction of one to the other. Descriptions of A and B need not contain information of that sort. Thus, Tye argues, images cannot be construed as descriptions alone.

However, one way to capture picture-like properties in descriptionist terms has been proposed by Geoffrey Hinton (1979). According to him, imagery does not occur in a special medium, a visual buffer of the sort that Kosslyn describes. However, neither does it depend on the same format and processes as higher-order thought—that is, descriptions in Pylyshyn's sense. Rather, it involves a distinctive format and set of operations, albeit defined over descriptions of a more elaborate kind. Attached to object-centered descriptions of shapes are egocentric

coordinates for objects in a scene, which add spatial information and allow for operations of a special sort (e.g., a gradual alteration of the coordinates, in terms of which rotation and scanning can be described). This account explains why subjects find it hard to identify figures embedded in complex geometrical shapes (a triangle in a star of David) or to reinterpret ambiguous figures, once an original interpretation has been made (e.g., to see the rabbit in the duck-rabbit image if it was initially seen as a duck). Interpretations not included in or derivable from the original descriptions will be hard to come by, and this may be particularly so if they require a revision of the coordinate reference frame. The problem is that Kosslyn's evidence shows that subjects are able to reinterpret images even when the new interpretation is incompatible with the original reference frame. This would not be predicted on Hinton's account.

Tye has proposed a hybrid theory, according to which an image consists in an array and an interpretive description combined. The descriptive components are limited, consisting primarily of part descriptions that are produced whenever the part is scrutinized. Thus the array is not rendered irrelevant by a complex description that could simply take its place. However, it is unclear exactly how arrays and interpretations are combined on this account; that is, why certain descriptions are generated for an array on certain tasks and precisely how the properties of the array are used to perform the task. The question is why, on the one hand, basic part descriptions are not simply activated directly on a visual memory task, thus making the array unnecessary. On the other hand, if the array functions to support the discovery of previously unnoticed features, then there is no guarantee that ambiguities will not appear in descriptions of basic shapes themselves.

One promising avenue for research is suggested by Tye's argument that part descriptions are generated only as needed. That claim is consistent with Kosslyn's current view that imagery and perception are governed by a principle of *opportunistic processing*: Representational resources can be deployed in diverse and sometimes limited ways, as required by a task. In his 1994 book and subsequent research, Kosslyn argues that imaging is not a single capacity, but comprises a set of subsystems that are distributed in the brain. Although these subsystems are functionally specialized, they can interact, and they can be employed strategically in various combinations. This approach implies that the interpretation of images in the visual buffer need not always consist in inferences over language-like representations. Instead, assignment of content in an image is constrained by the particular operations and strategies in which aspects of the image are incorporated. For instance, image scanning consists in enhancing activity in various parts of the visual buffer, thus priming specific features, making them easier to encode. In that sense, scanning patterns constitute interpretations, because they bias the content that can be ascribed to the image by the visual system. This account has the potential to explain individual differences in image interpretation in terms of variations in perception-like strategies. Born out of Kosslyn's turn to neural networks and connectionist modeling, this emphasis on imaging strategies tracks the ongoing development of cognitive science and philosophy of mind.

See also Images.

Bibliography

Aristotle. *De Anima*.

Berkeley, George. *A Treatise concerning the Principles of Human Knowledge*. Chicago: Open Court Press, 1907.

Block, Ned, ed. *Imagery*. Cambridge, MA: MIT Press, 1981.

Block, Ned. "The Photographic Fallacy in the Debate about Mental Imagery." *Nous* 17 (1983): 654–664.

Dennett, Daniel. "The Nature of Images and The Introspective Trap." In *Imagery*, edited by Ned Block. Cambridge, MA: MIT Press, 1981.

Hinton, Geoffrey. "Some Demonstrations of the Effects of Structural Descriptions in Mental Imagery." *Cognitive Science* 3 (1979): 231–250.

Kosslyn, Stephen. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press, 1994.

Kosslyn, Stephen. *Image and Mind*. Cambridge, MA: Harvard University Press, 1980.

Pylyshyn, Zenon. *Computation and Cognition*. Cambridge, MA: MIT Press, 1984.

Rollins, Mark. *Mental Imagery: On the Limits of Cognitive Science*. New Haven, CT: Yale University Press, 1986.

Rollins, Mark. "The Strategic Eye: Kosslyn's Theory of Perception and Imagery." *Mind and Machines* 11 (2001): 267–286.

Shepard, Roger, and Lynn Cooper. *Mental Images and Their Transformations*. Cambridge, MA: MIT Press, 1982.

Tye, Michael. *The Imagery Debate*. Cambridge, MA: MIT Press, 1991.

Mark Rollins (2005)

IMAGES

Aristotle's claim that "it is impossible even to think without a mental picture" (*On Memory and Recollection* 450a) has frequently been echoed by subsequent philosophers. David Hume equated thinking with having mental

images, since he appears to have considered ideas and images to be the same; for of any sense impression "there is a copy taken by the mind, which remains after the impression ceases; and this we call an idea" (*Treatise of Human Nature,* Book I, Part I, Sec. ii). The sole contents of the human mind are original impressions and these copies of them. Thomas Hobbes was stating much the same view when he said, "Imagination therefore is nothing but *decaying sense*" (*Leviathan,* Ch. 2).

Many other philosophers have also accepted the existence of such mental contents without examining their nature; they had assumed that images are things whose nature or existence is obvious to all human beings and that can most simply be described as "copies" or "pictures" of the external world. Views denying the existence of such objects have been rare; the chapter in Gilbert Ryle's *Concept of Mind* that seems to attack the commonly held view of the imagination as the power of producing mental images is felt by many to be contrary to normal experience. The images that Hobbes and Hume were talking of, and that Ryle attacks, are mental existents, depending on our prior experience of the physical world, though they may have objective counterparts in the brain. In this they differ from the Epicurean *eidola* or *simulacra,* which Lucretius defined as "images of things, a sort of outer skin perpetually peeled off the surfaces of objects and flying about this way and that through the air" (*De Rerum Natura,* Book IV, 11. 29ff.). These images Lucretius thought of as physical objects, albeit rather ethereal ones, whose function is to explain perception as well as images and dreams. When actual existence is attributed to them they are made to resemble the physicists' "real images," which are the representations of objects formed on screens or in space by lenses, or on the retina of the eye by the same mechanism. Physicists also talk of a "virtual image," a visual appearance that cannot be detected by physical means in the place in which it seems to be (for example, the appearance of objects behind the mirror's surface). This usage, which implies that there is something unreal about the image, is nearer to the normal philosophical or psychological use than is that of the term *real image.* The connection between "image" and "imaginary" is preserved in ordinary usage.

IMAGES AS THE MEANINGS OF WORDS

Undoubtedly the strongest desire to maintain the existence of mental images has come from the need to provide something to serve as the bearer of meaning for words of our language. George Berkeley's attack on John Locke in the introduction to *The Principles of Human Knowledge* is mainly concerned with this question. Against what he took to be Locke's view of the existence of "abstract general ideas," or the meanings of general terms, Berkeley argued that images must be particular. It is, he claimed, impossible for anyone to form a general idea (by which he clearly meant "image") of a triangle, for it would have to be "neither oblique nor rectangle, neither equilateral, equicrural, nor scalenon, but all and none of these at once." Whether Locke had meant this by his argument for abstract general ideas will not be discussed here; the important point is that Berkeley may be said to have shown that in some cases thought may proceed without images, because there could be no image or "mental picture" to correspond with some terms of our vocabulary.

Nevertheless, it may still be claimed that imagery is an important part of our mental life; this is argued by H. H. Price in his *Thinking and Experience* and elsewhere. Empirical evidence would appear to show that there is considerable divergence in the amount of mental imagery experienced by different individuals; Sir Francis Galton (in *Inquiries into Human Faculty and Its Development*) stated that imagery tended to be lacking in "scientific" minds and to be common in those of artistic bent. The Würzburg school of psychologists in the early twentieth century maintained that their experiments proved the existence of "imageless" thought. The difficulty here lies at least partially in determining what is to be called a "mental image." Although most people, as has been said, understand the instruction "Picture to yourself a familiar building" and claim to be able to do so, it is obvious that what they do in such a case is not the same thing as looking at a picture or photograph of the object, and it is not clear what connection this ability has with that of using the words of a language.

WITTGENSTEIN'S CRITICISM

Ludwig Wittgenstein argued that if some form of mental picture is needed to "give meaning" to a word, then an actual picture can be used instead; for example, asked to get a red apple, a man could use a color chart that gave a specimen of red opposite the word *red.* He could then compare apples with this sample until he found one that matched. Those who think of images as being essential to the use of language are talking as if each person carried such charts "in his head" and proceeded in the same way in the absence of an actual sample. The difficulty with this view, in Wittgenstein's opinion, is that the command "Imagine a red patch" can be given and obeyed; here it is obvious that the "mental sample" will be of no use or will

lead to an infinite regress. The image can itself be recognized as red without the use of any intermediary, so there is no reason why a specimen of red should not also be recognized. Most people do, in fact, immediately recognize specimens of the common colors, though they may need a chart for the rarer ones. Wittgenstein summarized his attack on the false picture of recognition as follows:

> It is as if I carried a picture of an object with me and used it to perform an identification of an object as the one represented by the picture. Our memory seems to us to be the agent of such a comparison, by preserving a picture of what has been seen before, or by allowing us to look into the past (as if down a spy-glass). (*Philosophical Investigations,* Sec. 604)

There are two further difficulties about this view of the image as the bearer of meaning. First, it is not clear how an actual picture functions, and second, the comparison of the image with a picture itself gives rise to difficulties.

FUNCTIONING OF ACTUAL PICTURES

Price has stated that "both words and images are used as symbols. They symbolise in quite different ways, and neither sort of symbolisation is reducible to or dependent on the other. Images symbolise by resemblance" (*Thinking and Experience,* p. 299). Price's arguments for his weakened version of the imagist theory rest, as the quotation shows, on the assumption that images, like other pictures, are related to their objects by resemblance. Such a view assumes that there is no problem in recognizing a picture of, say, a man as a man. Just as anyone who could pick out a real man could identify a mirror image of a man, so, it is thought, could he pick out a pictorial representation of a man.

But what is to count as a picture of a man here? A child's matchstick man consisting of five lines and a circle? A rough sketch? A "lifelike" portrait by a Royal Academician? A life-size photograph? As the art historian E. H. Gombrich has shown in his *Art and Illusion,* the representation and the recognition of three-dimensional objects on a two-dimensional surface is a sophisticated activity. Our children are taught something of the appropriate techniques at about the same time as they learn their native language. There is no basis for feeling that the procedure of representing objects in these ways is more "natural" than describing them by means of words. It has been said that some primitive peoples find it impossible to recognize a photograph of one of their number because they have not learned to interpret the pattern of black and white in the appropriate way. Yet it would seem that a photograph is the most "natural" representation because it is the product of a purely objective projection of the object; drawings and paintings depend on a variety of learned techniques of representation.

It is necessary to distinguish between the way in which a picture is produced and the use that is made of it. There may be a method of projection, but it is not because of that method that we accept the picture as a likeness. Furthermore, it is not clear from the picture itself, though it may be from the title, what it is meant to be the likeness of. A picture of an oak may be that of a particular historic tree (King Charles's Oak, for example), an example of an oak tree for purposes of identifying the species, an illustration for a general article on trees, a sign for a forest, or a composition to hang on the wall for its "artistic" quality. Without some rule it is impossible to tell what the picture is for and hence what its subject is; its meaning, what it symbolizes, lies in the use we make of it. In the context of a botany class it may be quite clear that the picture of an oak is being shown to enable students to identify specimens of that tree; here the rule is given by the situation in which the picture is used. Similarly, it is clear that the man who carries a photograph of his sweetheart does it to remind himself of her, uses it as a kind of substitute for her presence. Real pictures have a variety of uses.

IMAGES AS PICTURES: OBJECTIONS

A picture may be used to give information; from a picture of the Pantheon it is possible to discover the number of columns in the facade. But as Jean-Paul Sartre points out in *L'imaginaire* (p. 117), an image of the Pantheon may not be sufficiently detailed to enable this, even though before the question was asked the agent thought his image was perfectly clear. If he does not already know the number, then he cannot count the columns in his image. In this the image differs radically from the picture. Furthermore, it is usually known what the image is an image of without the need to inspect it for clues. Even when an image arises in the mind and cannot be recognized, no closer examination will provide clues to its identity; we have to wait until the name comes to us. In the extreme case of dreaming, we may "recognize" a person even though his characteristics are entirely different from those possessed in real life. A picture, on the other hand, may be identified gradually by the collection of clues. Thus, "having an image" of an object differs from contemplating either the object or a picture of it. The image

is not a picture in a special private gallery (cf. Ryle, op. cit., p. 247).

Part of the difficulty, as Ryle stresses, is due to an excessive concentration on the sense of sight; we naturally talk of "picturing" or of "visualizing," but there are also aural, tactual, and olfactory imagery. (A blind man's imagery, presumably, would be entirely of these kinds.) But in these cases there is no recognized means of representing the sound, touch, or smell—what would such a process be like?—and hence no temptation to talk of such images in terms that are drawn from the inspection of physical representations.

We do find it very natural to talk of mental images, and because external objects are normally described in visual terms, these terms are also applied to images.

Images are not always under our control; a person may find he is "haunted" by the image of a street accident or by the cries of the victim. Images do occur and must be accounted for. But to say this need not lead us to think of them as "decaying sense." Such a description would apply to afterimages, caused by staring at a bright light and then looking away. But these are actually perceived and can be physically located, on or just in front of whatever is looked at. Mental images have no location and are not related to public visual space; it is useless to ask a subject, as some psychologists have done, to project his mental image onto a screen, for it is impossible to look at the physical world and contemplate an image at the same time. But the "seeing" of a visual image or the "hearing" of an auditory one is only, in Sartre's terminology, a "quasi observation"; as Ryle puts it, "an imagined shriek is neither louder nor fainter than a heard murmur. It neither drowns it nor is drowned by it" (op. cit., p. 250; despite differences in terminology, there is a measure of agreement between Ryle and Sartre on this topic). The "quasi-observational" nature of our apprehension of images is marked by the device, naturally adopted, of putting quotation marks around "see" and "hear" in this context.

Nevertheless, the question "What is a mental image?" is wrongly posed, for it implies that there is some definite mental content to which the words can be applied. As has been shown above, the similar question "What is a picture?" equally has no definite answer. A picture may be regarded as a pattern of pigment on a piece of canvas, and much can be said about it in this respect. But such a description leaves out of account its function *as a picture,* which may be to recall the face of an absent friend. When it is being used for this purpose its characteristics as a physical object are ignored; the person is seen "through" the painted representation. It is he in whom we are interested. Similarly, when a mental image is being used it is the object that is of interest to us, not the image itself. "When we are thinking, although we must know what our images are of, it is not necessary for us to know what our images are like—even whether they are clear and distinct, or fuzzy and shifting" (D. W. Hamlyn, "The Stream of Thought," 71). Indeed, it is hard to see how it is possible to know "what they are like," for they are described only in terms of their objects. In the case of the portrait there is a public object that can be described in physical terms and serves as the "analogue" of the absent friend.

It has been suggested that there are similar analogues in the case of mental imagery—for example, movements of the eyeballs. These may well occur, but their occurrence is not part of what is meant by having an image. In the case both of the picture and of such movements it is the way in which an existing but absent object is indicated or referred to that constitutes the essence of the representation. Sartre has suggested that we can set up a series of representations, starting with a photograph, continuing with a full portrait, a drawing, a caricature (which may be a few lines on paper or a piece of behavior on the part of an actor). All these are ways of indicating a particular person. The series can be continued with a mental image, and finally with the person's name. These different ways of thinking of him depend on a relation of meaning. "For the contents (images for instance) which accompany or illustrate them are not the meaning or intending. … If God had looked into our minds He would not have been able to see there whom we were speaking of" (Wittgenstein, op. cit., p. 217). Or whom we were thinking of. So far from being the vehicles of meaning, images are dependent on a prior ability to mean or intend particular objects for their very existence. In this they are like pictures, but this fact must not lead us into talking of our apprehension of images as if it were the inspection of private pictures.

See also Aristotle; Berkeley, George; Hobbes, Thomas; Hume, David; Imagery, Mental; Imagination; Locke, John; Lucretius; Ryle, Gilbert; Sartre, Jean-Paul; Thinking; Wittgenstein, Ludwig Josef Johann.

Bibliography

CLASSIC WORKS

Aristotle. *On Memory and Recollection.*
Aristotle. *On the Soul.*
Berkeley, George. *The Principles of Human Knowledge.*
Hobbes, Thomas. *Leviathan.*
Hume, David. *Treatise of Human Nature,* Bk. I.
Locke, John. *Essay concerning Human Understanding.*

Lucretius. *De Rerum Natura* (*On the Nature of Things*).

MODERN WORKS

Flew, Annis. "Images, Supposing and Imagining." *Philosophy* 28 (106) (July 1953): 246–254.

Furlong, E. J. *Imagination.* New York: Macmillan, 1961.

Furlong, E. J., C. A. Mace, and D. J. O'Connor. "Abstract Ideas and Images." *PAS,* Supp., 27 (1953): 121–158. A symposium.

Galton, Francis. *Inquiries into Human Faculty and Its Development.* London: Macmillan, 1883.

Gombrich, E. H. *Art and Illusion.* London: Phaidon, 1960.

Hamlyn, D. W. "The Stream of Thought." *PAS,* n.s., 56 (1955–1956): 63–82.

Harrison, Bernard. "Meaning and Mental Images." *PAS,* n.s., 63 (1962–1963): 237–250.

Humphrey, George. *Thinking.* London and New York, 1951.

Price, H. H. "Image Thinking." *PAS,* n.s., 52 (1951–1952): 135–166.

Price, H. H. *Thinking and Experience.* London, 1953.

Ryle, Gilbert. *The Concept of Mind.* London: Hutchinson, 1949.

Sartre, Jean-Paul. *L'imaginaire.* Paris, 1940. Translated by Bernard Frechtman as *The Psychology of the Imagination.* London and New York, 1949.

Shorter, J. M. "Imagination." *Mind* 61 (244) (October 1952): 528–542.

Wittgenstein, Ludwig. *Philosophical Investigations.* Translated by G. E. M. Anscombe. Oxford: Blackwell, 1953.

A. R. Manser (1967)

IMAGINATION

Imagination is generally held to be the power of forming mental images or other concepts not directly derived from sensation. In spite of the popular usage of the term, the majority of philosophers from Aristotle to Immanuel Kant considered it in relation to knowledge or opinion. They conceived it either as an element in knowledge or as an obstacle to it—as in Plato's attack on art—or as both an obstacle and an element. David Hume is a representative of the last view: "Nothing is more dangerous to reason than flights of the imagination, and nothing has been the occasion of more mistakes among philosophers." Yet in the same place he wrote of the understanding as "the general and more established properties of the imagination" (*Treatise of Human Nature,* Book I, Part IV, Sec. vii). The fancy, the power of the imagination to combine ideas in fantastical ways, is to be avoided, but nevertheless imagination is vital to knowledge.

This latter element in Hume's view had its greatest development in Kant's *Critique of Pure Reason,* where the imagination is described as a "blind but indispensable function of the soul, without which we should have no knowledge whatsoever, but of which we are scarcely ever conscious." Kant thought that the imagination has two tasks to perform in giving rise to knowledge, though it is not always easy to separate them. First, it completes the necessarily fragmentary data of the senses: it is impossible to perceive the *whole* of an object at once, yet we are seldom aware of the partial nature of our perception. For example, we cannot see more than three sides of a cube at one time, but we think of it as having all six sides. This completion of perception is the work of the "reproductive" imagination (called reproductive because it depends on prior experience for its operation). Kant contrasted this with the "productive" imagination, which has an even more important role to play.

The two names mark different functions of the imagination, rather than imply that it is twofold. The productive imagination gives rise to the transcendental synthesis of imagination, which combines our experience into a single connected whole. Kant called this operation "transcendental" because it is prior to experience, not subsequent to it; without such a synthesis no coherent experience of a world would be possible. So central is the work of the imagination to the first *Critique* that it is sometimes hard to separate from the understanding; Kant even said in one passage: "*The unity of apperception in relation to the synthesis of the imagination* is the *understanding*; and this same unity, with reference to the *transcendental synthesis* of the imagination, the *pure understanding*" (A 119).

ARTISTIC IMAGINATION

In spite of Kant's emphasis on the productive nature of the imagination and the importance he gave to it, his view of it in the first *Critique* is still as a faculty for forming images, images that are at the service of the cognitive powers of the mind. It is our normal apprehension of the world that is mainly at issue in that work. Consequently, it is hard to see how this use of the term is related to that by which we talk of writers and artists as "imaginative." Many critics and philosophers have written as if the artist or writer were a person especially good at imagining, in the sense of visualizing, scenes or events that had not occurred, which he then transmitted to the public by means of his art. The mental operations were of the "fancy" in Hume's sense of the term, the imagination recombining materials it had previously received from the senses into new forms that were not reproductions of previous experiences. The degree to which an artist could do this was the measure of his imaginative powers, while the reader or viewer reproduced in his own mind what

the artist had had in his. Two contemporary literary critics have attacked this view:

> But much great literature does not evoke sensuous images, or, if it does, it does so only incidentally, occasionally and intermittently. In the depiction even of a fictional character the writer may not suggest visual images at all. … If we had to visualise every metaphor in poetry we would become completely bewildered and confused. (Wellek and Warren, *Theory of Literature,* pp. 26–27)

It has even been suggested that the term *imaginative* has now come to fill the place in the critical vocabulary left by the general abandonment of the term *beautiful* in aesthetics; a "work of imaginative power" would previously have been called "beautiful." Clearly it is inadequate to equate "imagination" with the power of the mind to produce images. Interestingly enough, the germ of a better theory of the imagination might be seen in Kant's discussion of teleological judgment in his *Critique of Judgment*: to think of nature *as if* it had a purpose is an imaginative activity, though there do not seem to be any actual images involved in the process.

COLERIDGE. One of the most important contributions to the theory of the imagination in the nineteenth century was that of Samuel Taylor Coleridge, put forward in *Biographia Literaria* and elsewhere. He strongly contrasted the Fancy and the Imagination; the former he defined as "no other than a mode of Memory emancipated from the order of time and place." It operates almost mechanically and is responsible for the production of verse, whereas the Imagination is the source of true poetry. This he divided into two: the Primary Imagination, which is the equivalent of Kant's productive imagination and is responsible for all human perception, and the Secondary Imagination, which is the source of art. Coleridge described the operation of the Secondary Imagination as follows: "It dissolves, diffuses, dissipates, in order to re-create … it struggles to idealise and to unify. It is essentially *vital,* even as all objects (*as* objects) are essentially fixed and dead." This vital nature of the imagination meant for Coleridge that it is a way of discovering a deeper truth about the world; he would have agreed with John Keats's "What the imagination seizes as beauty must be Truth," and thus he went beyond the Kantian original of this theory. In this he sided with the romantics, for whom art and science were alternative ways of reaching the real world; previous writers had tended to think of science and philosophy as superior to art in this respect.

RYLE ON IMAGINATION

Coleridge and those who followed him, including both Benedetto Croce and R. G. Collingwood, still thought of the Imagination as a single faculty or power of the mind. Gilbert Ryle, in his chapter on imagination in *The Concept of Mind,* stresses that there is no one thing that can be called "imagination" but rather a variety of activities that are imaginative, among which are pretending, acting, impersonating, fancying, and so-called imaging. His arguments clearly establish his central thesis, though his subsidiary denial of mental images, which is not essential to the main point, is open to doubt. A child shows his imaginative ability, Ryle maintains, not by what goes on in his head but rather by the way in which he plays—for instance, the manner in which he pretends to be a bear. An actor, again, demonstrates his ability by the way he performs on the stage, his public appearance, to which mental accompaniments are largely, if not entirely, irrelevant.

Many of the activities called "imaginative," Ryle says, are "mock-performances"; he talks of boxers sparring as "making these movements in a hypothetical and not a categorical manner" (p. 261). This is closely connected with supposal, the running over in the mind of a future possibility. Indeed, in ordinary speech the word *imagine* is often synonymous with "suppose" or "think"; the instruction "imagine what it would be like if" is equivalent to "think what it would be like if." In both cases the evidence that the instruction had been carried out would be a report in words; even the operation itself might have been purely verbal, without any "images" passing through the mind. Hence, Ryle can argue that there is no need for an artist or writer—or, indeed, for anybody at all—to have "mental imagery."

IMAGINATION AND TRUTH

Because there is such a close connection between "imagining" and "supposing" or "fancying," it is easy to see why what is imagined is often thought to be unreal or false. In fact, "I must have imagined it" is a common form for the admission of a mistake of some kind. Hence, it is natural for epistemologically minded philosophers to assume that all imaginative activity is false or unreal. Ryle, in spite of the overall excellence of his account, may be criticized on this score: Such forms of expression as "mock-performance" and the use of quotation marks stress this element. However, the falsity of the imagination may, by philosophers of other persuasions and interests, be welcomed as a sign of the mind's freedom. Jean-Paul Sartre would appear to be of this number. E. J. Furlong, in his

book *Imagination,* agrees with Sartre on this point: "to act 'with imagination' is to act with freedom, with spontaneity; it is to break with the trammels of the orthodox, of the accepted; it is to be original, constructive" (p. 25). But, as has already been mentioned, artists and writers about art often want to go further than this, to stress the "truth" of imaginative works. Collingwood, for example, in a section of *The Principles of Art* titled "Imagination and Truth," has said, "Art is not indifferent to truth; it is essentially the pursuit of truth" (p. 288). It is clear that the truth in question is one somehow connected with the imagination rather than with the ordinary cognitive powers of the mind.

The difficulty of assessing this claim is increased by the fact that the idealist theory of art, of which Collingwood and Croce are the chief representatives, places the locus of the work of art not in its physical manifestation, the painting or poem, but in the imagination of the artist and spectator. The real work of art is an experience in the mind of the artist, and the spectator is moved to re-create the experience of the artist in his own imagination when he contemplates the picture. The picture is thus connected with the work of art but is not the work itself. The main difficulty here lies in the fact that it is an imaginative experience, not a statement, which is said to be true. A subsidiary problem is that such a view leads to the undervaluing of the actual product of the artist, the picture, novel, or poem. But the stress on the part played by the imagination in appreciating art is shared by some writers not normally thought of as idealists. For instance, Sartre says, "In a word, reading is directed creation" (*Situations II,* p. 96). The writer, he argues, has only provided a series of clues that the reader has to "solve" and complete by his own activity. Sartre even goes so far as to talk of reading as a "dream under our own control" (ibid., p. 100), which assimilates the appreciation of art even more closely with activities normally thought of as imaginative—for example, daydreaming.

One aspect of the idealist account of art clearly fits in with our normal thinking on the subject, for a person said to be "imaginative" is frequently one who is capable of appreciating works of art or of fiction. A man who could not read novels because "they are not factual" would be unimaginative. But the antithesis imaginative-factual that is here employed would seem to contradict the idealist claim that art is connected with truth. In ordinary conversation a novel may be described as "true to life" or "realistic." A child pretending to be a bear may also be praised for the realism of his performance, as may a young actor playing the part of an old man. In these and similar instances no one need be deceived by the novel or the performance; the readers or spectators can be fully aware that they are not reading a factual account or seeing a genuinely old man. Indeed, if they were not so aware their reactions would be different. The spectator who responds to the stage performance as to an actual event has made a serious mistake; many events on the stage would be too painful to contemplate if they took place in real life. This kind of awareness has sometimes been described as "aesthetic distance," but it is the same feature that was above described as the "unreality" of the imagination. Sartre expresses this fact by saying that the image "contains a certain nothingness." He continues: "However lively, however affecting or strong an image may be, it is clear that its object is non-existent" (*L'imaginaire,* p. 26). For Sartre, when someone imagines the face of an absent friend he is supposing that the friend is present to him, which ex hypothesi he is not. A person who forgets that he is imagining, that his thought is supposal, not fact, has made the same mistake as the spectator who thinks a real murder has been committed on the stage. The sense in which imagination may provide, in works of fiction, for example, a "truth" that is not conformity to actual fact can thus only be that the world which is supposed is a possible one, in the sense that it is self-consistent. Those who claim that the imagination gives another "truth" must be extending the meaning of the word in a way that requires justification, or at least explanation.

What has just been said also serves to point to a solution of the difficulty of the idealist account, that of the actual mode of existence of the work of art, whether it is in the mind or is the physical object it is ordinarily taken to be. Against the idealist view it is normally asserted that what is criticized in a work of art is the work itself, not its effects on the imagination, which would be private to each person; the critic thinks he is talking about a public object. The solution lies in the ambiguous nature of the work of art, as Sartre stresses, in that a picture, for example, can be viewed either as paint on canvas or as a picture of an absent friend. The picture does not produce an image of the absent person, but, as Sartre says, we respond to the picture in some of the ways in which we would respond to the friend himself, albeit we are aware that he is not present. The ability to respond in this way is the imagination, but the response does not require a flow of imagery in the mind. To have established this is one of the merits of Ryle's account.

MIMESIS

It is now possible to see the connection between many of the various, apparently disparate uses of "imagination." The man who is thoroughly immersed in reading a story, who is almost dreaming it, is very like the child who is fully occupied with pretending to be a bear. These are in a position similar to that of the man who is taking the behavior of a young actor on the stage for that of an old man. There is a common element in the behavior of all three, which is shared by the man who is supposing that something is the case, though his activity is less full. This man, again, is not dissimilar to the person having a mental image, who is fancying or supposing that he is seeing or hearing something he is not seeing or hearing, although aware that he is not.

All of these notions are related to an earlier account of art, the Greek *mimesis,* or imitation, although it has often been thought that there was a radical difference between them. Aristotle's idea of an "instinct of imitation" in the *Poetics* (IV, 1) is not entirely unlike Ryle's account of the imagination. In both cases there is something unreal about the activity, as Sartre has tried to indicate by his talk of "nothingness" as a feature of imagination; in these areas the implications of normal life do not hold. Thus, in spite of the apparent diversity of usage, there is a "family likeness," in Ludwig Wittgenstein's phrase, between the various terms, which makes talk of "the Imagination" legitimate.

See also Aristotle; Coleridge, Samuel Taylor; Collingwood, Robin George; Croce, Benedetto; Hume, David; Imagery, Mental; Images; Kant, Immanuel; Plato; Sartre, Jean-Paul; Wittgenstein, Ludwig Josef Johann.

Bibliography

CLASSIC WORKS

Aristotle. *On the Soul.*

Aristotle. *Poetics.*

Coleridge, Samuel Taylor. *Biographia Literaria.* Edited by J. Shawcross. Oxford: Clarendon Press, 1907.

Hume, David. *Treatise of Human Nature,* Bk. I.

Kant, Immanuel. *Critique of Judgment.* Translated by J. H. Bernard. New York, 1951.

Kant, Immanuel. *Critique of Pure Reason.* Translated by N. Kemp Smith. London, 1929.

MODERN WORKS

Cameron, J. M. "Poetic Imagination." *PAS,* n.s., 62 (1961–1962): 219–240.

Carritt, E. F. *The Theory of Beauty.* London, 1914.

Collingwood, R. G. *The Principles of Art.* Oxford: Clarendon Press, 1938.

Croce, Benedetto.*Estetica come scienze dell'espressione e linguistica generale.* Bari, Italy: Laterza, 1902. Translated by D. Ainslie as *Aesthetic.* London: Macmillan, 1922.

Furlong, E. J. *Imagination.* New York: Macmillan, 1961.

Gombrich, E. H. "Meditations on a Hobby Horse." In *Aspects of Form,* edited by L. L. Whyte. London: Lund Humphries, 1951.

James, D. G. *Scepticism and Poetry.* London: Allen and Unwin, 1937.

Koestler, Arthur. *Act of Creation.* New York: Macmillan, 1964.

Langer, S. K. *Feeling and Form.* New York: Scribners, 1953.

Levi, A. W. *Literature, Philosophy and the Imagination.* Bloomington: Indiana University Press, 1962.

MacDonald, M. "Art and Imagination." *PAS,* n.s., 53 (1952–1953): 205–226.

Mischel, T. "Collingwood on Art as 'Imaginative Expression.'" *Australasian Journal of Philosophy* 39 (1961): 241–250.

Price, H. H. *Thinking and Experience.* London: Hutchinson, 1953.

Ryle, Gilbert. *The Concept of Mind.* London: Hutchinson, 1949.

Sartre, Jean-Paul. *L'imaginaire.* Paris, 1940. Translated by Bernard Frechtman as *The Psychology of the Imagination.* London and New York, 1949.

Sartre, Jean-Paul. "Qu'est-ce que la littérature." In *Situations II.* Paris: Gallimard, 1948. Translated by Bernard Frechtman as *What Is Literature?* New York: Philosophical Library, 1949; London: Methuen, 1950.

Shorter, J. M. "Imagination." *Mind,* n.s., 61 (244) (October 1952): 528–542.

Wellek, R., and A. Warren. *Theory of Literature.* New York: Harcourt Brace, 1949.

Wittgenstein, Ludwig. *Philosophical Investigations.* Translated by G. E. M. Anscombe. Oxford: Blackwell, 1953.

Ziff, P. "Art and the 'Object of Art.'" *Mind,* n.s., 60 (1951): 466–480.

A. R. Manser (1967)

IMAGINATION [ADDENDUM]

Late-twentieth-century discussions of imagination have tended to focus on three sorts of issues. Discussions in the philosophy of mind have focused on the cognitive architecture underlying imagination, and on the ways that imagination differs from and resembles belief, perception, and supposition. Discussions in modal epistemology have focused on the extent to which imaginability—and its cousin conceivability—can serve as guides to possibility. And discussions in aesthetics have focused on a cluster of issues concerning our imaginative engagement with fictional characters and events.

PHILOSOPHY OF MIND

Within the philosophy of mind, three distinct notions of imagination have been discussed: sensory imagination (quasi-perceptual experience in the absence of appropriate stimuli); recreative imagination (mental simulation); and creative imagination (combining ideas in unexpected and unconventional ways), with the great bulk of discussion devoted to the former two.

SENSORY IMAGINATION. Drawing on work by cognitive psychologists (e.g., Shepard 1982, Farah 1999), philosophers have explored the extent to which sensory imagination in general, and visual mental imagery in particular, employs the same systems as those involved in corresponding perceptual experience, and the related question of whether mental images are encoded in analogue form (as mental pictures) or propositionally (as descriptions). While the mainstream view holds that entertaining a visual mental image involves inspecting some sort of picture-like object (Kosslyn 1994), critics—most notably Zenon Pylyshyn (2003)—maintain that reasoning with mental images need not involve any sort of quasi-sensory representation. (A selection of early philosophical papers on the topic of visual imagery can be found in Block 1981; an overview of the debate is presented in Tye [1991/2000]; discussion of related issues in the phenomenological tradition can be found in Casey 2000; these themes are also examined in McGinn [2004] and Williams [1973].) Discussions of other sensory modalities have been explored less thoroughly, though some philosophical attention has been paid to the question of how motor imagery ought to be understood—whether as imagined action or imagined perception of action (Jeannerod 1997, Currie and Ravenscroft 2002); and there is a small literature on auditory imagery (primarily in the phenomenological tradition; cf. also Reisberg 1992).

A related discussion concerns the intentional status of mental images: do they derive their content through resemblance alone (an image of a maple leaf resembles and thus represents a maple leaf), or through some other mode of representation (an image of a maple leaf represents a maple leaf only as a result of being "labeled" as such)? Many analytic philosophers, following Jerry Fodor (1975) and Hilary Putnam (1981) have endorsed the latter view, though there has been some dissent.

RECREATIVE AND CREATIVE IMAGINATION. The other main area of exploration in the philosophy of mind has concerned what Gregory Currie and Ian Ravenscroft (2003) have dubbed recreative imagination: the capacity that underpins one's ability to take perspectives other than one's own. This capacity, sometimes called off-line simulation, seems to play a central role in the understanding of other minds, in the contemplation of counterfactual scenarios, in the planning of behavior, and in engaging in explicit games of pretense. (Autistic children, for example, show marked deficits both in pretend play, and in understanding the mental states of others.) Beginning in the early 1990s—prompted in part by work of philosophers Alvin Goldman, Robert Gordon, and Jane Heal, and by empirical work by psychologists such as Simon Baron-Cohen, Paul Harris, Alan Leslie, and Uta Frith—investigation of these connections began in earnest, resulting in a number of important anthologies (Carruthers and Smith 1996, Davies and Stone 1995a, 1995b), collections (Currie 2004, Heal 2003) and book-length treatments (Currie and Ravenscroft 2003, Nichols and Stich 2003).

In a related vein, philosophers concerned with simulation and pretense have offered hypotheses concerning the cognitive architecture that supports such a capacity (Nichols and Stich 2003) and offered discussions of the features which it shares with and which distinguish it from related attitudes such as belief and supposition (Currie and Ravenscroft 2002, Velleman 2001; essays in Lopes and Kieran 2003; Nichols 2006).

Work on the creative imagination has been primarily in the context of empirical psychology (Boden 2003, Csikszentmihalyi 1996, Sternberg 1998; see also Gaut and Livingston 2003.)

MODAL EPISTEMOLOGY AND AESTHETICS

Within the context of modal epistemology, a great deal of scholarly attention has been focused on the question of whether there is some representational capacity, imagining or conceiving, that can serve as a reliable guide to possibility, and, if so, whether particular conclusions about, for example, mind-body dualism can be established by exploring what one can or cannot conceive or imagine. Advocates of conceivability-possibility theses, most notably David Chalmers, have contended that under certain conditions, conceivability is a fully reliable guide to certain sorts of possibility (Chalmers 1996); critics have maintained that the relevant notions of conceivability or possibility are of doubtful philosophical utility (Gendler and Hawthorne 2002.)

Within the general purview of aesthetics, philosophers interested in the imagination have discussed both

general questions about the nature of the imagination as well as a number of specific and general puzzles, among them the Paradox of Fictional Emotions (how and why one feels apparent emotional reactions toward characters and events that one explicitly recognizes as fictional), the Paradox of Tragedy (how and why one enjoys artworks that induce feelings such as sadness and fear) and the Puzzle of Imaginative Resistance (why one is reluctant to imagine fictional worlds that differ morally from the actual world). (Influential anthologies on these topics include Bermúdez and Gardner 2003, Hjort and Laver 1997, Lopes and Kieran 2003, Nichols 2006; collections of essays include Currie 2004, Levinson 2002; book-length treatments include Carroll 1990, Currie 1990, 1995, Feagin 1996, Robinson 2005, Scruton 1997, Walton 1990; see also Brann 1992.)

See also Aesthetics, History of; Art, Expression in.

Bibliography

Baron-Cohen, Simon. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press, 1997.

Bermúdez, José Luis, and Sebastian Gardner, eds. *Art and Morality*. London: Routledge, 2003.

Block, Ned. *Imagery.* Cambridge, MA: MIT Press, A Bradford Book, 1981.

Boden, Margaret. *The Creative Mind: Myths and Mechanisms*. London: Routledge, 2003.

Brann, Eva T. H. *The World of the Imagination: Sum and Substance*. Savage, MD: Rowman & Littlefield, 1992.

Carroll, Noel. *The Philosophy of Horror*. New York: Routledge, 1990.

Carruthers, Peter, and Peter K. Smith, eds. *Theories of Theories of Mind*. Cambridge, U.K.: Cambridge University Press, 1996.

Casey, Edmund. *Imagining: A Phenomenological Study*. 2nd ed. Bloomington: Indiana University Press, 2000.

Chalmers, David. *The Conscious Mind*. New York: Oxford University Press, 1996.

Csikszentmihalyi, Mihaly. *Creativity: Flow and the Psychology of Discovery and Invention*. New York: Harper Collins, 1996.

Currie, Gregory. *Arts and Minds*. New York: Oxford University Press, 2004.

Currie, Gregory. *Image and Mind: Film, Philosophy and Cognitive Science*. Cambridge, U.K.: Cambridge University Press, 1995.

Currie, Gregory. *The Nature of Fiction*. Cambridge, U.K.: Cambridge University Press, 1990.

Currie, Gregory, and Ian Ravenscroft. *Recreative Minds: Imagination in Philosophy and Psychology*. New York: Oxford University Press, 2003.

Davies, Martin, and Tony Stone, eds. *Folk Psychology*. Oxford: Basil Blackwell, 1995a.

Davies, Martin, and Tony Stone, eds. *Mental Simulation*. Oxford: Basil Blackwell, 1995b.

Farah, Martha. *The Cognitive Neuroscience of Vision*. Oxford: Basil Blackwell, 1999.

Feagin, Susan. *Reading With Feeling: The Aesthetics of Appreciation*. Ithaca, NY: Cornell University Press, 1996.

Fodor, Jerry. *The Language of Thought*. Cambridge, MA: Harvard University Press, 1975.

Frith, Uta. *Autism: Explaining the Enigma*. Oxford: Basil Blackwell, 1992.

Gaut, Berys, and Paisley Livingston, eds. *The Creation of Art*. Cambridge, U.K.: Cambridge University Press, 2003.

Gendler, Tamar Szabó, and John Hawthorne, eds. *Conceivability and Possibility*. New York: Oxford University Press, 2002.

Harris, Paul. *The Work of the Imagination*. Oxford: Blackwell, 2000.

Heal, Jane. *Mind, Reason and Imagination: Selected Essays in Philosophy of Mind and Language*. Cambridge, U.K.: Cambridge University Press, 2003.

Hjort, Mette, and Sue Laver, eds. *Emotion and the Arts*. New York: Oxford University Press, 1997.

Jeannerod, Marc. *The Cognitive Neuroscience of Action*. Oxford: Basil Blackwell, 1997.

Kosslyn, Stephen M. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press, 1994.

Lamarque, Peter, and Stein Haugom Olsen. *Truth, Fiction, and Literature: A Philosophical Perspective*. New York: Oxford University Press, 1996.

Levinson, Jerrold. *Aesthetics and Ethics: Essays at the Intersection*. Cambridge, U.K.: Cambridge University Press, 2002.

Lopes, Dominic, and Matthew Kieran, eds. *Imagination, Philosophy, and the Arts*. London: Routledge, 2003.

McGinn, Colin. *Mindsight: Image, Dream, Meaning*. Cambridge, MA: Harvard University Press, 2004.

Nichols, Shaun, ed. *The Architecture of the Imagination: New Essays on Pretense, Possibility, and Fiction*. New York: Oxford University Press, 2006.

Nichols, Shaun, and Stephen Stich. *Mindreading: An Integrated Account of Pretense, Self-Awareness and Understanding Other Minds*. New York: Oxford University Press, 2003.

Putnam, Hilary. *Reason, Truth, and History*. Cambridge, U.K.: Cambridge University Press, 1981.

Pylyshyn, Zenon. *Seeing and Visualizing*. Cambridge, MA: MIT Press, 2003.

Reisberg, Daniel, ed. *Auditory Imagery*. Hillsdale, NJ: Erlbaum Associates, 1992.

Robinson, Jenefer. *Deeper Than Reason: Emotion and Its Role in Literature, Music, and Art*. New York: Oxford University Press, 2005.

Scruton, Roger. *Art and Imagination: A Study in the Philosophy of Mind*. South Bend, IN: St. Augustine's Press, reissued 1997.

Shepard, Roger. *Mental Images and Their Transformations*. Cambridge, MA: MIT Books, 1982.

Sternberg, Robert. *The Handbook of Creativity*. Cambridge, U.K.: Cambridge University Press, 1998.

Tye, Michael. *The Imagery Debate*. Cambridge, MA: MIT Press, 2000.

Velleman, David. *The Possibility of Practical Reason*. New York: Oxford University Press, 2001.

Walton, Kendall. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press, 1990.

Williams, Bernard. *Problems of the Self: Philosophical Papers 1956–1972*. Cambridge, U.K.: Cambridge University Press, 1973.

Tamar Szabó Gendler (2005)

IMMORTALITY

The literature on the philosophical problems involved in the question of a future life begins with Plato. We cannot therefore do better than start with a quotation from *Phaedo,* the dialogue in which Plato deployed what is put forward as a demonstration of the immortality of the soul. Having heard and apparently accepted the supposed proof put into the mouth of Socrates, Crito asks:

> "But how shall we bury you?" "However you please," Socrates replied, "if you can catch me and I do not get away from you." And he laughed gently, and looking towards us, said: "I cannot persuade Crito, my friends, that the Socrates who is now conversing and arranging the details of his argument is really I: he thinks I am the one whom he will presently see as a corpse, and he asks how to bury me. And though I have been saying at great length that after I drink the poison I shall no longer be with you, but shall go away to the joys of the blessed, he seems to think that was idle talk uttered to encourage you and myself."
>
> (115CD)

This passage can be employed to fix two fundamental points by reference to which the main problems can be mapped.

The first point is that the essence of doctrines of personal survival (or immortality)—and precisely this and this only is what gives them their great human interest—is that they should assert that after our deaths we shall continue to exist (forever). Only in this way can they provide the basis for what John Wisdom has called "these logically unique expectations"—that we shall, to put it as noncommittally as possible, have "experiences" after death, that death will be not our terminus but the beginning of a new journey. This has to be underlined both because there have been some famous philosophical doctrines of immortality which have not been of this sort and because there have been other doctrines which either were from the beginning substitutes for, or have been so hedged and interpreted that they have now ceased to be, the genuine article, personal immortality. Thus, whatever one makes of Aristotle on the immortality of the intellect (*De Anima* 429a–431a) or of Benedict Spinoza on the eternal element in the mind (*Ethics,* V, xxiii ff.), it is certain that these views do not, nor were they intended to, provide any ground for such "logically unique expectations." Again, the man who urges that we all of us live forever because the ill (and sometimes even the good) men do lives after them indicates by the very irrelevance of his supporting reason that this sort of immortality is not the authentic personal brand.

The second basic reference point is that any doctrine of survival or immortality has one enormous and immediate obstacle to surmount before it can begin to qualify for any further consideration. This obstacle consists simply in our manifest universal human mortality. It is due to this ineluctably familiar fact that "All men are mortal" has become a trite truistic example in logic and that we so use the word *survive* that it is logically impossible for one and the same passenger to be both dead and a survivor after a crash. One can recognize and respect the longing that may lie behind the epitaph "Not dead, but sleeping," yet no one can deny that, literally interpreted, it is false. Indeed, even to consider the contrary possibility would be to enter the ghoulish world of Edgar Allan Poe. This second reference point is, of course, as obvious as the first though the inducement to disregard it is perhaps greater. And both points have to be kept constantly in view.

THREE DOCTRINES OF IMMORTALITY

There seem to be three ways of trying to circumvent the massive initial difficulty confronting any doctrine of personal immortality. If once the initial obstacle could be overcome, the remaining problems would turn out to be not philosophical but factual and practical.

IMMORTAL-SOUL DOCTRINE. The first of the three doctrines is the Platonic attempt to demonstrate two points. One point is that we are essentially composite beings. Besides the more familiar corporeal element, the body, there is also something else, different in kind—the incorporeal soul. For the duration of a life the soul is somehow attached to—incorporated into or imprisoned in—its body. Although the soul is incorporeal, it is nevertheless a substance, something that could significantly be said to exist independently of anything else. This is precisely the point of Plato's arguments to the conclusion that the soul is not a harmony, for it would not even make sense to suggest that a harmony might survive or, for that matter, precede the elements of which it is a harmony.

This is no more possible than that there could be a grin without a face to grin it (*Phaedo* 85E–86D and 91C–95A).

The other point is that the soul must be the person or, at any rate, the real, true, or essential person. This, though it is sometimes neglected, is crucial. Unless it is established that I am my soul, the demonstration of the survival of my soul will not demonstrate my survival, and the news that my soul will last forever could provide me with no more justification for harboring "these logically unique expectations" than the rather less elevated assurance that my appendix is to be preserved eternally in a bottle.

RECONSTITUTION DOCTRINE. The second doctrine is, in its purest form, extremely simple and direct. It consists in urging the resurrection of the body or, more accurately, the reconstitution of the person. Whereas the first doctrine insists that I am the sort of thing that could perfectly well escape unharmed and unnoticed at death ("if you can catch me and I do not get away"); the second recognizes that to be truly a human person, I have to have the corporeal human form. It then relies on an act of sheer omnipotence to produce the immortal me.

SHADOW-MAN DOCTRINE. Considering that the reconstitution doctrine or something approaching it has been part of traditional Christianity, it is surprising to find that Thomas Aquinas seems to be the only philosopher of the first rank to have discussed a version of it at any length. (See especially "Treatise of the Resurrection" in the supplement to Part III of the *Summa Theologiae.*) But the shadow-man doctrine seems to have received even less attention. It can perhaps best be regarded as an attempt to combine the strong points of the other two doctrines. It is the claim that a person is a kind of shadow man, sufficiently human and corporeal to overcome the problem of identification with the familiar flesh and blood person and at the same time sufficiently ethereal and elusive to have no difficulty in escaping unnoticed from the ordinary earthy body which is destined to be burned or buried. This view is found in some of the Christian Fathers (for instance, Tertullian in *De Anima*). A similar view is also held by some modern spiritualists—"the astral body" detaches itself at death to proceed on its "journey to the summerland." Perhaps the best way of conveying the idea to the modern reader unfamiliar with either patristic or spiritualist literature is to refer to the many films in which a "spirit" is shown as a tenuous shadowy replica of a man that detaches itself from him at death and is thereafter visible to the entire audience but only to favored characters in the film.

DIFFICULTIES IN THESE DOCTRINES. At first sight the third doctrine might appear to be the most promising way to avoid the initial difficulty. However, the doctrine is bold precisely where the other two are discreet. By insisting upon the essential incorporeality of the soul, the first neutralizes all the ordinary weapons of empirical inquiry; the second, by deferring the corporeal resurrection to an unspecified time and place, indefinitely postpones any occasion for their deployment. But astral bodies, detaching themselves at death from the other sort, should be empirically detectable here and now. The crucial and probably insoluble dilemma for the shadow-man doctrine is to provide a specification of the nature of an astral body in which an astral body remains sufficiently like an ordinary flesh and blood person to avoid difficulties of identification and at the same time to ensure that the claim that there are such things would be verified or, at least, not immediately falsified by the appropriate factual investigation.

In the time of Tertullian and even in the early days of modern psychical research there may have been some slight basis for believing that this might possibly be done. Tertullian himself appealed not only to purely theological considerations but also to such cases as that of the woman who claimed to have seen "a transparent and lucid figure in the perfect form of a man." The systematic investigation of such phantasms has shown, however, that though they do undoubtedly occur, they belong to the category of purely subjective and hallucinatory experience. (See, for instance, G. N. M. Tyrrell, *Apparitions,* London, 1953.) The third way must therefore be dismissed as a blind alley.

Any reconstitution doctrine confronted with the question "How is the reconstituted person on the last day to be identified as the original me, as opposed to a mere replica, an appropriately brilliant forgery?" There seems to be no satisfactory answer to this question, at least for a pure reconstitution theory. This question is, however, logically prior to all questions about the reasons, if any, that might be brought forward in support of such a doctrine.

This decisive objection seems rarely to have been raised, and when it has been, its force has not usually been felt. No doubt, the explanation lies largely in the fact that the doctrine is scarcely ever found in the pure form, unalloyed with any Platonic elements. There are two defenses that might be offered. It could be urged that God will infallibly ensure that the unending torments and the eternal ecstasies are allocated to the right people—that is, in the one case to the very same people who had incurred his disapproval and in the other case to precisely those who had won his favor. It might also be suggested that

though there might indeed be cases in which all other merely human observers must be entirely at a loss, the person himself could not fail to know whether he was the original person rather than a changeling or a replica. These are, in substance, the arguments presented by John Locke in developing and defending his own analysis of personal identity; however, that analysis itself was intended to meet difficulties about the identification of the future victims of divine judgment with people now living on Earth.

Both these arguments, though they possess a strong appeal, seem to miss the point. Notwithstanding the form of the original question, the difficulty is not one of "How do you know?" but of "What do you know?" The objection is that the reconstituted people could only be mere replicas of and surrogates for their earthly predecessors. Neither the appeal to the cognitive and executive resources of Omnipotence nor the appeal to the supposed special status of the person in question does anything at all to meet this contention.

The point can be brought out better in the case of the argument that the person cannot fail to know who he is. This argument depends on the premises that if I remember doing something, I must have done it and that normal people in normal situations are usually able to remember the most important features of their lives. Both premises are true, but it does not follow from them that if someone on the last day or any other day claims in all honesty to remember doing something, he must in fact have done it. From "He remembers doing that" it follows necessarily that "He did that," just as from "He knows that that is true" it follows necessarily that "That is true." But from "He claims to remember doing that, and he is not lying" it does not follow that "He did that," any more than from "He claims to know that that is true, and he is not lying" it follows that "That is true."

The crux in both cases is the possibility of honest error. However honest and however convinced, claims that you did something do not guarantee that you actually did it. In normal circumstances most such claims are no doubt entirely reliable, and the memories involved can properly be said to constitute knowledge. But the circumstances envisaged by the reconstitutionists are conspicuously not normal. In these circumstances the question of whether any of the ostensible memories enjoyed by the reconstituted people can properly be counted as memories at all must wait on the resolution of the logically prior issue of whether they have any past to remember, of whether, in particular, they are indeed the people they apparently think they are.

Thomas Aquinas seems to have appreciated that the immediate objection to any pure reconstitutionist view is decisive. This insight is no doubt one of the reasons that his own view incorporated important Platonic elements. In answer to the objection "that it will not be identically the same man that shall rise again. … After the change wrought by death the selfsame man cannot be repeated," Thomas replied:

> The form of other things subject to generation and corruption is not subsistent of itself, so as to be able to remain after the corruption of the composite, as it is with the rational soul. For the soul, even after separation from the body, retains the being which accrues to it when in the body. … Consequently there has been no interruption in the substantial being of a man, as would make it impossible for the selfsame man to return on account of an interruption in his being. (Op. cit., IIIa, Supp. 79, 2, *ad* 1)

IMMORTAL-SOUL DOCTRINE

It thus seems that if any headway is to be made toward overcoming the first gigantic obstacle in the way of doctrines of survival or immortality, it will have to be, at least in its first stages, Platonic in an extremely broad sense. In this sense René Descartes's views on the nature of the soul and its relations to the bodily machine can be characterized as Platonic (see, for instance, *The Passions of the Soul* I, xxx ff.), and even Thomas must count as at least Platonizing. For everyone who maintains that the mind or the soul is a substance, in the sense that it could significantly be said to exist alone and disembodied, is thereby Platonizing, and everyone who identifies this putative substantial mind or soul as the real or true person is adopting a fully Platonic position.

Platonizing concerning the nature of the soul seems to be the essential condition of the possibility of any defensible doctrine of personal immortality or even survival. Philosophy that is Platonizing or Platonic in this sense constitutes an enormous field, so this article will concentrate on the views of Plato, Aristotle, and Descartes, making their positions serve as focuses. All three are well suited to serve in this way both because their work has been and is enormously influential and because they can each be seen as representative of a different approach to the problem. Since all three believed in some sort of immortality and since it is scarcely possible to treat their arguments about the nature of the soul without touching on their ideas of immortality, it will be convenient to consider the two together.

DIFFICULTIES IN THE NOTION OF THE SOUL. One great and rather peculiar difficulty of our present subject is that it is very hard and, hence, very uncommon for anyone really to begin from the beginning. The key terms such as *body, mind,* and *soul* and their equivalents or near equivalents in other languages are all in a way familiar to all those acquainted with the languages concerned. Though there may be much unclarity about their meanings and perhaps even a certain indeterminacy in their usage, they are not the recognizably fresh coinages that both demand definitions and enable meanings to be prescribed with some confidence that the appropriate usage can be followed consistently by the prescriber and by others. (The point here is simply that speech habits, like other habits, can be very hard to break.)

Again, in the air there always are and always have been vulgar or not so vulgar theories about bodies, minds, and souls and their several natures, destinies, and relations, with no guarantee that these theories can be harmonized with the meanings of such terms as *mind* or *soul* as actually determined by whatever is the accepted correct usage of the culture concerned. For instance, correct usage—and this surely is and must be the only possible standard of meaning—may very well determine that the soul, in the relevant sense of *soul,* is not a substance, even though the people concerned entertain fantasies which presuppose that it is. This thesis has been fully developed in this connection by Gilbert Ryle, chiefly in *The Concept of Mind.* The consequence of this is that the issues tend to be presented not as a matter of first giving some suitable sense to the word *soul* or *mind* and then asking whether in that sense the term would in fact have any application but, rather, as an inquiry which presupposing that we have or are souls or minds, asks what is the nature of the soul or the mind. Such a presentation is bound to constitute a temptation to prejudge the question whether, in the sense eventually chosen, we have or are such things. By itself this might give rise to no more and no less trouble than the insistence of the modern analytic philosopher that he is not concerned with the reality of matter, time, or whatever but is merely searching for an adequate analysis of the notions of matter, time, and so on. In the present case, however, there are also the various theories in the background, and these theories happen to be of very different sorts and suited to answering very different questions.

The fundamental distinction needed is that between theories that might serve as answers to philosophical questions about the meanings of the terms *mind, soul,* and the like and theories that offer some sort of explanation of why the creatures that are said to have minds, souls, and the like behave and suffer as they do. Thus, for instance, in reading Aristotle's criticisms of his predecessors (*De Anima* I), we must distinguish—even if they, or he, did not—between those thinkers who said that the soul was a vapor, blood, or something of that order and must to that extent be interpreted as embryonic scientists and those thinkers who urged that the essence of the soul is motion, sensation, or some combination of the two and are therefore to be counted as philosophers.

Another crucial distinction must be made between explanatory and descriptive concepts. Suppose that for some person the meaning of the words *mind* and *soul* is to be given entirely in terms of certain capacities or incapacities and that in his use having a first-class mind is just being able to compass certain sorts of achievement. In his sense of *mind,* then, it can be no explanation to say that someone can do these things because he is endowed with a first-class mind, for to say this is, for him, only to redescribe the phenomenon. If his concept of mind is to be explanatory, he must give the word *mind* a meaning such that minds would be, in the terminology already explained, substances.

PLATO. There seems to be only one, very brief passage in Plato (*Alcibiades* I 129B–130C) where the argument is explicitly directed toward the justification of the Platonic presupposition that the soul is the person. Even in this passage the first presupposition, that we are composite beings, is very much taken for granted.

Soul as the person. Socrates is talking with Alcibiades, and the question is raised, "What are we, and what is talking with what?" The conclusion is that we are our souls. The argument runs in this way. In speaking, we use words. The user and the thing used are always different. We use our hands, our eyes, our whole bodies. Thus, I cannot be my body. Yet it is agreed that I must be my soul, my body, or a combination of both. However, because the user and the thing used are always different and because I use my body, I cannot be either my body or my body and soul combined. Thus, I must be my soul.

Considering how vital the conclusion is, Plato's argument may seem inadequate. But if, sympathetically, we call in the rest of Plato's writings to provide supplementary evidence, it becomes clear that it has to be taken as the epitome of many arguments. For Plato constantly talked of the phenomenon of self-control and the lack of it and of all those times when we are inclined to speak of being let down or dragged down by the weaknesses or by the excessive strength of the body or some part of it. Few

concepts are, in the ordinary narrow sense, more typically Platonic.

In all the innumerable cases of bodily control or lack of it, it is possible to produce arguments of basically the same form, and it is not at all necessary to appeal to the assumption, which may not seem as obvious to everyone today as it did to Plato's Alcibiades, that I must be my soul, my body, or both. Thus, starting from the known fact that our eyes sometimes play tricks on us, we may go on to argue that this shows that we see not with, but through, them, that they are, as it were, built-in optical instruments (compare *Theaetetus* 184C ff.). Or, again, noting how natural and entirely proper it is to describe someone on some desperate occasion as "flogging on his protesting body," we may infer that we drive our bodies as we drive our cars. In every case the Platonic conclusion, expressed in a more modern idiom, would be that the personal pronouns, personal names, and all other person words, words which clearly must refer to something and which, it seems, equally clearly cannot refer to bodies, the only available corporeal objects, must therefore refer to some incorporeal objects, the conclusion is that these are the objects to which we apply the term *souls.*

This conclusion is wrong, however, and all arguments of this kind are misguided. It is not true that person words are words for any sort of incorporeal objects. People are what you meet. We do not meet only the sinewy containers in which other people are kept, and they do not encounter only the fleshy houses that we ourselves inhabit. It is therefore wrong to suggest that the word *person* is equivalent to the word *soul* in this sense of *soul* and, hence, to imply that it is contradictory to deny that people are incorporeal objects and that it is absurd to say that you can see a person. This basic fact about the meanings of person words is central and fundamental to the entire problem.

To deal in detail here with all the variations on the present argument would be impossible. The mistake involved in all such arguments seems to be that of insisting that because expressions such as "that person" are, for one reason or another, not synonymous with "that human body" and because we use all sorts of idioms in which I and my body are spoken of as if they were two substances, there must therefore be a special class of incorporeal objects for person words to refer to. This false conclusion seems to be one more product of the perennially disastrous *unum nomen, unum nominatum* theory of meaning—the misconception that every different class of word must refer to a different class of object. The truth seems to be that in this area we have a vastly rich and idiomatic vocabulary that provides us with all manner of subtle linguistic instruments, all of which we employ to say things about one sort of inordinately complicated but essentially corporeal creature, ourselves.

Argument from reminiscence. Plato's second argument to the conclusion that we are incorporeal souls was his doctrine of reminiscence. This has two forms, each proceeding to the same conclusion from rather different premises. In one form the premise is that we can all be shown to possess some knowledge, which we have not acquired in this life, of a priori truths (*Meno* 81B–86B). In the other it is that we all have certain ideal concepts, such as the ideal of perfect equality, which we cannot have acquired in our lives because they are never fully instantiated in this world (*Phaedo* 73A–77A). In both forms the conclusion is that these facts can be accounted for only in terms of memory. We, or "our souls," must have acquired our knowledge of the conclusion of the theorem of Pythagoras or have been acquainted with the Platonic Idea of equality before this life began.

This argument has never been very popular, partly because of a well-grounded mistrust of both premises and partly because of the fact that the notion of preexistence involved in the conclusion does not square with the demands of Western orthodoxy. For its force the argument depends on the existence of an important logical link between (true) memory and personal identity. If I really do remember certain truths or being acquainted with certain objects, then it follows that at some time in the past I must have learned them or made that acquaintance. Plato's argument is sound, but he draws the wrong conclusion. The correct conclusion is not that we must be remembering from a former existence but that memory cannot be involved. It cannot be memory for the simple and basic reason that we were not available to acquire knowledge or anything else before we existed, because we are not, what this argument in fact assumes that we are, the sort of incorporeal things which could preexist our conception and growth.

It is worth remarking that although to products of Western cultural conditioning preexistence appears much less credible than immortality, Plato, in insisting on both, was adopting a much less arbitrary position than that of those who assert immortality alone. It was not without reason that in the ancient world Lucretius and other spokesmen for human mortality made much of the comparison between our nothingness before birth and our annihilation in death (see, for instance, Lucretius, *De Rerum Natura* III. 11. 830–842 and 973–977). As George

Santayana once remarked, "The fact of being born is a poor augury of immortality."

Argument from rationality. Another argument was developed from distinctions embedded in Plato's account of Socrates' intellectual history (*Phaedo* 96A ff.). The Platonic Socrates here tells of his dissatisfaction with Anaxagoras, who apparently wanted to explain how the universe works rather than justify why everything is for the best. Socrates then goes on to contrast the physiological conditions of human behavior with the reasons the agent has for acting as he does. These categorial distinctions could serve as the foundations of an argument to the conclusion that since there are no necessary connections between the concepts of physiology, on the one hand, and the concepts that are peculiar to the distinctively human business of giving reasons for actions, on the other, it must therefore follow that rational agents are of their very nature incorporeal.

It would probably be going too far to attribute the argument in this form to Plato, although the conclusion and all the ideas involved are thoroughly Platonic. It is nevertheless one that needs to be noted here. C. S. Lewis and other contemporary apologists have tried to use these ideas to show that rationality is somehow essentially supernatural and that the bodily occurrences involved in rational behavior cannot be completely compassed in any scientific explanation.

There are two crucial points to be made in reply to this argument. The first is that precisely because the justification and appraisal of actions is so totally different from the causal explanation of physiological events, questions and answers belonging to the one universe of discourse cannot rival those belonging to the other. It is thus entirely possible to be confronted by a series of corporeal events—those, for instance, involved in what would normally be described as the oral development of an argument—and to ask and to answer both logical questions about the rationality of the whole performance considered as an argument and physiological questions about the causes of all the various glottal, oral, and nervous happenings considered as subject matter for the physiologist.

The second point is that to show that the concepts involved in the rational assessment of conduct are not logically reducible to purely physiological terms is not the same as to establish that agents must be essentially incorporeal. It would be equally impossible by purely logical analysis to translate the statement "Italy declared war" into a series of assertions about individual Italians, but this is no reason for thinking that "Italy" is the word for some incorporeal substance. Nor is the impossibility of a logical reduction any reason for thinking that it could even make sense to talk of incorporeal rational agents' or of Italy's taking part in international affairs if there were no individual Italians.

Life as a substantial soul. The Platonic approaches thus far considered have all involved thinking of the soul as the person; at the same time the person was wrongly thought of as an incorporeal substance. Another approach starts from the notion of a soul as a principle of life. It helps to note some peculiarities of the Greek language. The word *ψυχή*, translated "soul," is etymologically related to such words as *ἔμψυχος*, meaning "alive" (literally, "ensouled"), and *λιποψυχία*, meaning "swooning" or "death" (literally, "abandonment" by the soul). A popular idea to which Plato makes gently contemptuous reference was that death was a matter of the soul's permanently leaving its body; the soul was thought of as a puff of air, an invisible vapor, that would be dispersed in the breeze (see *Phaedo* 77D; compare, for instance, Euripides, *Supplices* 553–554). In this connection we might therefore distinguish two senses of "soul." In the first "to have a soul" means merely "to be alive"; in the second "soul" is the word for a class of supposititious entities, corporeal but elusive.

In the first sense one might speak, rather pretentiously, of the soul as the principle of life. In this sense we do have souls, for to say that a creature possesses a soul in this sense is just a misleadingly substantival way of describing it as alive. At this point Plato took another step, apparently without recognizing that any step was involved and, therefore, without providing the slightest warrant for taking it. He simply assumed what is manifestly false—that the word *soul* in this sense is equivalent to the term *soul* construed as a synonym for *person* (albeit for persons recognized to be incorporeal objects). He unjustifiably equated this "soul as the principle of life" with what the older commentators call "the soul as the bearer of moral values" (see, for instance, *Republic* 353D, a passage that is no less revealing for being found in an argument a little removed from our present concern).

Incompatibility of life and death. The false equation of two senses of "soul" is crucial in the most considerable of Plato's arguments for immortality (*Phaedo* 100B–107A). Of his other arguments the only one that retains more than antiquarian interest is the contention that the soul is something that moves itself and that whatever moves itself must be ingenerable and incorruptible (*Phaedrus* 245C–246A). And the interest of this argument lies mainly in its later theological development. It was the

germ of some of the theology of Plato's *Laws.* This theology led to Aristotle's notions of God as the Unmoved Mover. And Thomas later quarried Aristotelian materials for the first of his five ways—the argument to the First Mover.

This most interesting of the Platonic arguments presupposes Plato's general theory of Ideas, or Forms, especially as expounded in *Phaedo,* in the *Republic,* and elsewhere; in fact, Plato's other arguments derive what plausibility they may have from the theory of Ideas as a background assumption. Plato believed that for every significant word, such as *justice* or *equality,* there is a corresponding abstract Idea, or Form. These Ideas are eternal and incorporeal substances, intelligible to the intellect as material things are sensible to the senses. All the many particular instances of some general class of things "participate" in the appropriate unique Idea, and this Idea serves as an ideal standard, itself apparently preeminently possessing the characteristic concerned. These Ideas are thought of as providing answers both to questions about criteria—What makes an *X* count as an *X*?—and to questions of a more causal character—What is ultimately responsible for the existence of *X*'s?

The argument to show that life is incompatible with death starts from the notion of the soul as the principle of life, and this is equated, in terms of the theory of Ideas, with the Form of Life. Now, life in the abstract is as incompatible with death as equality is with inequality. Life can never be overcome by Death. Thus, the conclusion is that the soul as the very Idea of Life is essentially deathless and eternal and, hence, "the immortal part" of us is not destroyed by death, for "our souls will exist somewhere in another world" (*Phaedo* 106E and 107A).

The answer to this argument is that since Life and Soul are convertible terms in this context, there is as much or as little reason for saying that the Idea of Soul is eternal as there is for maintaining the eternal reality of any other Form. But, as Plato himself always insisted, the abstract Form is entirely different from the particular individual, whereas the nerve of the entire argument lies precisely in the equation of the Form of Soul with the particular soul. This identification is impossible not merely because, as we have seen, there is no reason to equate souls in the present sense with the souls that are people but, more fundamentally, because in the present sense no meaning has been given to the expression "an individual soul." This vital fact is one of the many that are obscured by the confusion of explanatory and descriptive concepts and by the failure to separate philosophical questions about criteria from factual questions about causes. Once these distinctions are made, it becomes clear that to say that someone has a soul (is alive) is not to say that he is alive only thanks to the presence of some mysterious extra substance, whether corporeal or incorporeal.

ARISTOTLE. Aristotle's *De Anima* is perhaps best approached as a philosophical treatise on life. *Anima* is Latin for "soul" and is the word from which our *animate* and *inanimate* are ultimately derived; the declared aim of *De Anima* is "to ascertain the nature and essence of soul" as "the principle of life" (402a–403b). The fundamental thesis is that life or the soul is "the form of the particular living body." The Aristotelian notion of form is complex and is to be distinguished from the Platonic. R. D. Hicks, the editor of the classic English language edition, stated that by the thesis that the soul is the form of the body Aristotle "so far from favouring materialism, secures once and for all the soul's absolute immateriality" (R. D. Hicks, ed., *Aristotle: De Anima,* Cambridge, U.K., 1907, p. xliii). Aristotle does dispose of all ideas that the soul is a lump of stuff. However, Aristotle's basic thesis is quite un-Platonic and leaves no room at all for any doctrine of immortality. An Aristotelian form is no more a corporeal thing than a Platonic Form would be, but it is not an incorporeal one either. In our sense it is not a substance at all. The soul as the form stands to the stuff of the particular body—and the examples are all Aristotle's—as the configuration of the statue to the materials of which it is made, as vision to the eye capable of seeing, as cutting power to the serviceable ax. Whatever else may be obscure, here it is obvious, as Aristotle himself said, that in this view the soul is not separable from the body (413a) and, furthermore, that this inseparability is a matter not of physical but of logical impossibility.

Had this been all that Aristotle said, Aristotle's views could not have been used in support of immortality. He also maintained, however, certain Platonic views that have given rise to much discussion and development, particularly among the Scholastics and others committed to a belief in personal immortality. These views concern the intellectual aspects of man and the corresponding intellectual (functions of the) soul. Despite the enormous labors of the commentators, precisely what Aristotle thought on these points is far from clear, possibly because Aristotle was not very clear in his own mind. There are nevertheless some relevant points that may usefully be made.

Immortal Abstract Intellect. The tradition descending from Alexander of Aphrodisias through Averroes

attributes to Aristotle a belief in some sort of Eternal Intellect. This is, however, a doctrine of personal immortality offering "prospect of rewards and punishments," a point emphasized by St. Thomas (*De Unitate Intellectus Contra Averroistes*). Furthermore, since the Abstract Intellect as opposed to the intellects of particular men is necessarily unique, it is not at all the right material to serve Thomas's own vital theoretical need for bridges between us and our successors in the next life.

Two senses of "eternal." The one kind of reason that might be proffered for saying that the Abstract Intellect (or any other putative Abstract Reality) is essentially eternal is really no reason for saying that anything at all actually goes on forever. In a way it is correct to say that such things as necessary truths and the logical relations between abstract mathematical concepts are somehow timeless and eternal. Yet this is not a matter of the existence of anything imperishable but, rather, of its not making sense to ask temporal questions about the periods during which any of these truths and these relations obtain. From *eternity* in this sense we can have nothing either to hope or to fear.

This distinction between two senses of *eternity* is fundamental in discussing personal immortality. We are concerned only with a life that would live on forever; the eternity of mere abstractions is not to the point. In the light of this distinction we can better appreciate the significance and the error of Plato's contention that the soul belongs to the same category as the abstract Ideas and, hence, that it is the sort of thing that may be presumed immortal (*Phaedo* 78B–80B).

The presumption that an incorporeal substance would be naturally incorruptible has always been the philosophers' favorite argument for the immortality of the soul. Thomas appealed to it, for instance, although he, of course, would not have included Platonic abstract Ideas in this category and although he was also careful to insist that souls, like everything else, are sustained by God and would be at once annihilated if he chose to withdraw his support (*Summa Theologiae* Ia, 75, 6; *ad* 2). It was perhaps with the same idea in mind that in the *Republic* Plato offered, as if it were a proof, the following unconvincing argument: Because every sort of thing has its one congenital evil, because nothing can be destroyed by anything but its own congenital evil, because the congenital evil of the soul is wickedness, and because wickedness as such is never directly lethal to the wicked man, our souls must be immortal (*Republic* 608E–611A).

Reason not localized. Assuming that Aristotle had really wanted to suggest that it could make sense to talk of an individual intellect's existing separately, then, presumably, a large part of his reason would have lain in his belief that ratiocination, unlike sight or hearing, is not localized in any organ (*De Anima* 402a, 408a, 429a). This belief has, of course, turned out to be erroneous. But even if it had not, the absence of any special corporeal organ provides no justification for assuming that our intellectual attributes must, or even might, be those of special incorporeal substances. The lack of special organs of melancholy or of volition is surely not to be construed as grounds for seeking invisible subjects to which to attribute Eric's feeling glum or Katrina's wanting to go to sleep. These are simply and obviously attributes of the people concerned.

Aristotle himself never employed any argument of this sort. On the contrary, he urged that the intellect, "since it thinks all things must needs, in the words of Anaxagoras, be unmixed with any, if it is to rule, that is, to know" (429a). This dark saying has been construed as an expression of a belief that our intellects are both incorporeal and substances, a belief that might seem to mesh well with the conviction, which Aristotle undoubtedly did have, that pure intellectual activity, abstract cognition, is something rather grand, almost divine, an occupation only for the highest sort of person (see, for instance, *Nicomachean Ethics* 1177a–1179b). Aristotle immediately went on to insist, however, that this pure intellect "has no other nature than this, that it is a capacity," and a capacity is not at all the sort of thing that can significantly be said to exist separately. Again, he seems elsewhere to have dismissed the idea of individual immortality with contempt (ibid., 1111b). And in the whole range of his works there is neither a positive treatment of the subject of a future life nor any promise of such treatment. The most plausible interpretation of Aristotle's view is surely that defended by Pietro Pomponazzi in Chapter 9 of his great polemic *De Immortalite Animae.* He concluded that the soul, including the intellect, "is in no way truly itself an individual. And so it is truly a form beginning with, and ceasing to be with, the body; nor can it in any way operate or exist without the body."

THOMAS AQUINAS. Thomas, as was mentioned, had urgent theoretical reasons for wanting to show that the soul is a substance, that it is, as he put it, "something subsistent." He was therefore inclined as far as possible to read Aristotle as holding the same view. In his *Commentary on Aristotle's De Anima* he explained the passage considered above (429a) in this way:

> But our intellect … must itself lack all those things which of its nature it understands. Since then it naturally understands all sensible and

> bodily things, it must be lacking in every bodily nature; just as the sense of sight, being able to know colour, lacks all colour. If sight itself had any particular colour, this colour would prevent it from seeing other colours, just as the tongue of a feverish man, being coated with a bitter moisture, cannot taste anything sweet.

If this were indeed what Aristotle meant, he really was confused. For if intellect is, reasonably enough, to be compared with the sense of sight, it is because they are both (cognitive) capacities. But we need no particular argument to show why a capacity, as opposed to the subject that may possess that capacity, cannot itself have any material characteristics. The reason that the sense of sight is not yellow is not that being yellow must render it or its possessor incapable of seeing yellow things but that, generally, it is nonsense to attribute sensible characteristics to a capacity. (It is hard not to regard all this as the product, at least in part, of the bad habit of making nouns out of verbs and then succumbing to the temptation to presume that a substantive must be a word for a substance.)

It might seem that it is upon precisely this argument that Thomas himself relied in the *Summa Theologiae* to establish "that the principle of intellectual operation which we call the soul is … both incorporeal and subsistent." He even employed the same example of "a sick man's tongue being vitiated by a feverish and bitter humour," but here he was comparing the soul as "the principle of intellectual operation" with the organ, not the sense, of sight. Having thus supposedly established that "it is impossible for the intellectual principle to be a body," he proceeded:

> It is likewise impossible for it to understand by means of a bodily organ; since the particular nature of the organ would prevent its knowing all bodies; compare the way in which liquid put into a glass vase seems to be of the same colour, not only when some particular colour is in the pupil of the eye but even when it is in the vase. (Ia, 75, 2)

In this version the argument escapes the previous criticism. But it escapes only at the price of removing what was in the commentary on *De Anima* offered as the proof of its major premise, here formulated as the proposition that "whatever knows certain things cannot have any of them in its own nature; because that which is in it naturally would impede the knowledge of anything else." The question arises, "How is this premise known?" The answer seems to be that it is not known, perhaps even that it is known to be false. Take Thomas's own example of the eye as the organ of sight. The eyes are admittedly material, yet that does not prevent us from using them for seeing material things, including other people's eyes and even—in mirrors—our own. Furthermore, even if the Thomist proposition did fit all the facts about our present sense organs, this would at most suggest that it was a contingent truth about them. But to serve Thomas's purpose, it must be known to be, if not actually necessary, at least sufficiently universal to apply not only to sense organs but also to "that principle of intellectual operation which we call the soul"—something which he himself is here trying to show to be radically different from anything corporeal.

DESCARTES. Plato and Aristotle can be regarded as the archetypical protagonists of two opposing views of man. Plato is the original spokesman for a dualistic view, and it seems that it is upon dualism that a doctrine of personal immortality must be grounded if it is to possess any initial plausibility. As a defender of a monistic view, Aristotle was neither so consistent nor so wholehearted. Yet it is still fair to see him at his most characteristic as the philosophical founding father of the view that the person is the living human organism, a view that apparently leaves no room whatsoever for belief in personal immortality. Thomas, who generally followed Aristotle on this point, characteristically attempted a synthesis that would have opened, had it been successful, the doors to heaven and to hell. In the present perspective Descartes must be placed squarely in the Platonic tradition. Thus, in the final paragraph of Part V of the *Discourse on Method,* after remarking that "next to the error of those who deny God … there is none which is more effectual in leading feeble spirits from the straight path of virtue, than to imagine that … after this life we have nothing to fear or to hope for, any more than the flies or the ants," Descartes concluded that "our soul is in its nature entirely independent of the body, and in consequence that it is not liable to die with it. And then, inasmuch as we observe no other causes capable of destroying it, we are naturally inclined to judge that it is immortal."

Soul as a thinking substance. Although his conclusions were thoroughly traditional, Descartes was nevertheless a revolutionary thinker. Unlike Plato, his chief intellectual interests were science, in particular physiology. Like Thomas Hobbes, the other great metaphysician of his period, Descartes quickly grasped the wider significance of the work of William Harvey and Galileo Galilei. Harvey's discovery of the circulation of the blood suggested to Descartes that both animals and human bodies might be regarded as machines. Descartes then asked

himself how the creatures that we know might be distinguished from living machines. His answer was that with respect to animals there simply was no distinction in principle but that an automaton in human shape, however brilliantly constructed, could always be distinguished from a true human being in two ways. There were two sorts of test which were bound to reveal the absence of the vital rational soul: without a rational soul such an automaton would not be able "to reply appropriately to everything … said in its presence, as even the lowest type of man can do," and their lack of versatility would always reveal that the automata "did not act from knowledge, but only from the disposition of their organs" (*Discourse on Method,* Part V).

One fundamental distinction, often overlooked in discussing these questions, is that between logical and technical impossibility. In Part V of the *Discourse,* his first published treatment, Descartes seems to have been making a purely factual claim "that it is morally impossible that there should be sufficient diversity in any machine to allow it to act in all the events of life in the same way as our reason causes us to act." To make any such would-be factual claim must be both rashly premature and scientifically defeatist. Elsewhere and later, it becomes clear that what Descartes, like so many successors, really wanted to say is that it is inconceivable that any material mechanism could be responsible for certain sorts of things. Thus, in the *Passions of the Soul* he laid down the principle "that all that which is in us and which we cannot in any way conceive as possibly pertaining to a body, must be attributed to our soul" (I, iv). And in his view what has to be thus attributed is thought, in his own rather broad sense of "thought," which seems to include all actions and passions considered to involve consciousness (ibid., I, xvii ff.). "By the word thought I understand all that of which we are conscious as operating in us. And that is why not only understanding, willing, imaging, but feeling also here count as thought" (*Principles of Philosophy,* I, ix).

Descartes was thus insisting that it is inconceivable that matter, however disposed, could in this sense think. This is a notion of the same sort as the idea that purposive and rational beings could not, without benefit of control by some Higher Purpose, have evolved first from creatures of a lower order and, ultimately, from inanimate matter, an idea found in both some objections to evolutionary theory and some versions of the Argument to Design. Presumably, Descartes would have accepted both contentions and many others like them because they fall under the generic principle, which he formulated as the fourth of his "axioms or common notions"; "All the reality of perfection which is in a thing is found formally or eminently in its first and total cause" (Addendum to the Replies to the Second Set of Objections to the *Meditations*).

It has since Immanuel Kant become the custom to dignify such principles with the title "synthetic a priori propositions." But the one with which we are here concerned, though certainly synthetic, can be described as a priori only in the quite artificial sense that it is wholly arbitrary and unwarranted. Descartes's more specific idea had been forcibly challenged long before by the Epicureans (see, for instance, Lucretius, *De Rerum Natura* II. 865–870 and 875–882). The challenge was later repeated by both Spinoza and Locke even before David Hume launched his decisive onslaught on the generic notion that it is possible to know a priori that some thing or sort of thing must be or cannot be the cause of some other thing or sort of thing. The points made, in their different ways, by both Spinoza and Locke were that there is no contradiction in the idea of something material being endowed with thought and that we are in no position to deny dogmatically that there are material creatures so endowed.

Subjectivism. Thus far, Descartes's originality, as against the Platonic tradition, has chiefly been in his positive scientific interests and in his mechanistic ideas about the body. His achievement was to form a new framework of discussion and to provide a metaphysical foundation for the further development of physiology. He was also revolutionary on a second count, for it was he who developed with compelling dramatic power a new approach to questions of mind and matter. For three full centuries this remained part of the accepted philosophical orthodoxy, an orthodoxy that even Hume seems never to have thought to question. This approach can be characterized, though with no intended moral overtones, as self-centered.

Whereas Plato generally—and Descartes, too, when he suggested tests of humanity—approached people from our common public world, Descartes at his most characteristic tried to approach the world from inside the closed circle of his logically private consciousness. Thus, in Part IV of the *Discourse,* having reached his rock-bottom certainty in the proposition *cogito ergo sum,* he asked what he was. "I saw that I could conceive that I had no body, and that there was no world nor place where I might be; but yet that I could not for all that conceive that I was not." He concluded that he "was a substance the whole essence or nature of which is to think, and that for its existence there is no need of any place, nor does it depend

on any material thing; so that this 'me,' that is to say, the soul by which I am what I am, is entirely distinct from body, … and even if the body were not, the soul would not cease to be what it is" (compare, especially, *Meditations,* II).

Much of the power of the Cartesian argument lies in the use of the first-person personal pronoun and in the idiosyncratic choice of tenses and moods. For there is surely no difficulty at all, even for Descartes, in supposing that Descartes may one day be annihilated or that Descartes might never have been born. The most fundamental objections are founded upon a rejection of his unstated general assumption that (his) words obtain their meaning by reference to (his) logically private experiences. In particular, Descartes mistakenly assumed that all the words for all the things that he comprehended under the term *thinking* are words for such private experiences. Only on this assumption is it possible to assert that there could be—much less that we are—essentially incorporeal beings and, as such, fully capable of every sort of thinking. To insist that this assumption is wrong is not necessarily to adopt either a complete logical behaviorism—saying that all terms of this type refer only to public performances—or Ludwig Wittgenstein's extreme later position—apparently denying the very possibility of a language's containing words defined in terms of one man's logically private experience. It is sufficient to commit oneself only to the more modest claim that most thinking words refer wholly or partly to various actual or possible proceedings that are necessarily corporeal. To recognize that this is true and could scarcely be otherwise, it is sufficient to reflect for a moment upon the whole context in which we learn to use these terms; consider how we should teach the meaning of "He argued with her" or "She drew her own conclusions." In this perspective it becomes no wonder that, as Wittgenstein said, "The human face is the best picture of the human soul."

Personal identity and parapsychology. The appeal of the Cartesian approach and its influence can be appreciated by considering two examples, both relevant to the question of immortality—first, the discussion of personal identity initiated by Locke and continued by Joseph Butler, Hume, and Thomas Reid and, second, the investigation of the question of human survival by modern parapsychologists through the study of the possible relevance of the evidence furnished by all types of mediumistic performances.

Both investigations have started from the self-centered Cartesian standpoint and have taken for granted that, essentially, people are bodiless. Thus, the problem of personal identity was generally taken to be one of the identity of an incorporeal thinking thing. Locke tried to provide an analysis of personal identity, so construed, in terms of consciousness (memory). The decisive objection to any such analysis was sharply put by Butler: "And one should really think it self-evident that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity" (Dissertation I, "Personal Identity," appended to the *Analogy of Religion*).

But most of Locke's critics, Butler included, seem to have failed to appreciate just how difficult—even, perhaps, impossible—the problems of the nature of the identity and of the principle of the individuation of such putative incorporeal beings must be. If people are thought of as incorporeal substances having sorts of thinking, in the wide Cartesian sense, as their qualities (the substance, or "pure ego," theory of the self), then the question is how such substances are to be identified, what sense can be given to the expression "pure ego." If, with Hume, one is unable to provide any satisfactory answer to this question, the only alternative seems to be thinking of people as collections of experiences (the serial, or "bundle," theory of the self). Theories of this sort face two difficulties. First, it does not seem to make sense to speak of thoughts or experiences as "loose and separate" without anyone's having them, and, second, there seems to be no string capable of tying the bundles of experiences together while keeping one bundle distinct from another. The first difficulty may or may not be merely grammatical. The second, once the impossibility of using memory as the string is fully realized, appears very formidable. It was the second difficulty in a slightly different form that Hume had to confess to be "too hard for my understanding" (Appendix to *Treatise of Human Nature*).

In parapsychology it seems to have been almost universally assumed that mediumistic material, insofar as it cannot be either satisfactorily explained away in terms of fraud and delusion or conveniently redescribed in terms of telepathic and clairvoyant transactions among the living, can and must be interpreted as evidence for human survival. Yet to interpret such material in this way is not to provide support for, but rather to presuppose, a Platonic-Cartesian view of man. For it is only insofar as a person is essentially incorporeal that it can even make sense to suggest that someone years ago dead, buried, and dissolved is even now communicating with us through a medium.

OTHER ARGUMENTS CONCERNING IMMORTALITY

This article has thus far concentrated on philosophers who have adopted, more or less consciously, a Platonic or Platonizing view of man and who, if they have argued philosophically for any sort of immortality, have urged that the nature of the soul is such that it must be or must be presumed to be imperishable. None of these arguments requires any reference to a deity, and none appeals to any moral premises. This may perhaps be surprising, for most people—at least those in the European cultural tradition—are likely to think that beliefs in God and in immortality must go together. They are inclined to take it for granted that the main if not the only point of immortality—and sometimes perhaps of God, too—is to provide inordinate rewards and punishments. Yet there is no obvious inconsistency in believing in a Creator while denying that he has established a new world in a future life to redress the moral unbalances of the old. Nor does it appear that to assert our immortality is logically either to presuppose or to imply the existence of any sort of god. It may seem odd, but it is not manifestly inconsistent, for such avowedly atheist philosophers as J. M. E. McTaggart and C. J. Ducasse to affirm immortality, McTaggart offering exclusively metaphysical reasons and Ducasse appealing mainly to the evidence of parapsychology.

MORAL ARGUMENTS. The most considerable philosopher to rest his case for immortality on morality was Kant. Unfortunately, this is one of the many cases in which it is difficult to give an account of Kant's position and reasons that is clear, consistent, persuasive, precise, and acceptable to Kant scholars. Kant himself may be at fault here not merely, as usual, because he obscured his thought with cumbrous and idiosyncratic expression but also because he presented imprecise and uncompelling arguments.

But with these warnings it can be said that in the *Critique of Practical Reason* and in the *Critique of Judgment* Kant offered freedom, immortality, and God as the three postulates of practical reason. Practical reason is for Kant the source of the universal imperatives of morality. A "postulate of pure practical reason" is defined in the *Critique of Practical Reason* as "a theoretical proposition which is not as such demonstrable, but which is an inseparable corollary of an a priori unconditionally valid moral law" (translated by L. W. Beck, Chicago, 1949, pp. 225–228). The form of the argument, in the cases of immortality and God, appears to be that something is said to be commanded by the moral law but could be obeyed only on a certain condition; therefore, the conclusion is drawn not that that condition must obtain but that it must be a postulate of practical reason. The first difficulty is to see how Kant, who in and after the *Critique of Pure Reason* regularly denied the possibility of proofs of immortality or of the existence of God, proposed to reconcile this denial with insistence on the validity of the present deductions. The most promising response to this is to suggest that they cannot be rated as proofs of the doctrines that Kant maintained to be unprovable because it has not been and cannot be shown that the moral ideas are indeed soundly based but that they do prove that to act in accordance with moral ideas is to act as if, or to act on the assumption that, these doctrines are true.

The second difficulty lies in the supposed derivations themselves. In the case of the postulate of immortality the conclusion is to be drawn from the premise that the moral law commands us to achieve a perfect correspondence between our will and that law. This is taken to be out of the question in this life. Thus, what the law really requires is an endless progress toward the ideal, which is possible "only under the presupposition of an infinitely enduring existence and personality of the same rational being." If this is what Kant meant—as it certainly is what he said—then the moral law includes one very strange command. For to reach the proposed conclusion, we have to construe that law not as stating that we should approach as near to perfection as is humanly possible, or, as Kant seemed at first inclined to say, that we must actually achieve perfection, but, rather, that we must forever approach asymptotically this eternally unattainable ideal.

In the case of the third postulate of God the moral premise is that the law requires us to promote the highest good, which involves a perfect correspondence between the morality and the happiness of every individual. But the only guarantee of the possibility of this correspondence would be the existence of God, presumably because God alone would possess the power necessary to achieve it. Consequently, practical reason demands this postulate.

There seems to be a crucial disharmony between the premises of the second and the third arguments. Only at first in the second but throughout the third, Kant apparently wanted to insist that the ideals prescribed by practical reason must be practically and not just theoretically possible. Surely, it is merely the contingent weakness of the flesh that makes holiness something "of which no rational being in the world of sense is at any time capable," whereas if the theoretical possibility of achieving the necessary correspondence was all that was at stake, there would be no call "to assume the actual existence of God."

Yet Kant had urged in the second argument that the true imperative is to press ever closer to an ideal which is, it seems, not even theoretically attainable. This is fatal. If it is once allowed that an imperative can be to get as near as is humanly possible to an ideal that may be practically—even theoretically—unattainable, then the whole foundation of both arguments collapses. For such more modest demands could be satisfied in an earthly lifetime and without benefit of God.

Three general points about the Kantian arguments should be particularly noted. First, that the cases for the second and for the third postulate are separate. Second, Kant scrupulously avoided any suggestion that the authority of the moral law is at all dependent on the availability, here or hereafter, of rewards and punishments. Third, Kant was careful not to make the mistake of trying to deduce what is the case from premises affirming only what ought to be. It is not often that any of these things can be said for some more popular arguments for immortality.

For instance, it is often urged—most commonly, perhaps, in Roman Catholic textbook apologetic but elsewhere, too—that the lack of appropriate rewards and punishments would make nonsense of the claims of morality. Thus, the Jesuit M. Maher wrote:

> But in the judgement that conduct entailing a sacrifice *ought* to be pursued, there is implied a further judgement that it cannot be ultimately *worse for the agent himself* to do that which is right. … The supposition that virtue can finally result in … misery for the agent; or that wickedness may effect an increase in the total quantity of his personal happiness is seen to be in conflict with reason, and to be destructive of all morality. (*Rational Psychology,* London, 1940, p. 530)

Maher proceeded to argue that God could not permit this and, therefore, that there must be immortality, with penalties and compensations. He himself believed that the existence of God is independently established by natural theology (p. 533), but "some of the proofs of Immortality are amongst the most forcible arguments for the existence of a Deity" (ibid., pp. 525–526). It is interesting to compare the distress of Henry Sidgwick, who saw the moral situation similarly but was unable to share the supposedly saving religious convictions (*Methods of Ethics,* London, 1874, especially Part IV, Ch. 7).

Even if it were to be allowed that some such view is correct, it certainly does not warrant the suggested conclusions. Suppose we allow that rewards and penalties are indeed morally necessary; at most, this could support a demand not for immortality but for a temporary survival. Nothing has been said in the premises to explain why these necessary rewards and penalties have to be eternal. Indeed, to the secular moralist, to whom no revelation has been vouchsafed, it might seem that to provide eternal penalties for temporal offenses would be to make the universe infinitely worse. More generally, it is essential to insist that no argument from purely gerundive premises—stating only what ought to be or what is in some other way desirable—can by itself either establish or make probable any conclusion about what is actually the case. (Compare J. S. Mill, *Three Essays on Religion,* London, 1874, especially the essay "Theism.")

In any case it is certainly no part of the meaning of moral obligation that the obligation must always accord with the eventual self-interest or the person obliged. The sense in which categorical imperatives can be characterized as essentially rational refers to their universality and impartiality rather than to any implication that obedience must always be ultimately the best-paying policy. If anything, surely, it is part of the very idea of morality that sacrifices are sometimes required.

ARGUMENT FROM DESIRE. Many other arguments have been and are put forward. It is urged, for instance, that the allegedly almost universal belief in survival is somehow evidence of its own truth, a contention rejected by J. S. Mill for the decisive reason that to urge this is not to offer a good ground but, rather, if anything, to concede tacitly that there is none. Again, attempts have been made (by Dugald Stewart, for instance) to make something of the allegedly almost universal desire for immortality or of the existence of human potentialities that cannot be realized in a mere three score years and ten. If the existence of a desire really were a reason for affirming not merely, as perhaps it is, that this desire has some describable object but also, as it manifestly is not, that this object must actually be realized, then the argument could still be refuted by the consideration—pressed by Hume in his essay "Immortality"—that the certainly no less nearly universal fear of annihilation equally demands its real object. Of course, the existence of a desire for immortality, where it is found, does call for—and can easily be given—a naturalistic explanation. Such a desire, however, begins to be useful to the advocate of immortality only insofar as it can be used in conjunction with some idea of a God who may be relied on to arrange for the ultimate fulfillment of (some of?) the desires and (some more of?) the potentialities that he has arranged for us to have. (The qualification "some of" has, presumably, to be put in to allow for

the existence of ambivalent and evil desires; the "some more of" is needed if we are to have an argument that even appears to hold.)

To consider the possibility of establishing the existence of such a God is beyond the scope of this article. Yet it is perhaps worth suggesting that the existence of these ostensibly frustrated wholesome desires and these apparently unfulfilled splendid potentialities must, by itself, count as evidence against, rather than for, the existence of this kind of God. It is, as Hume insisted both in his *Dialogues* and in the first *Enquiry,* very odd—notwithstanding that it is very common—to argue from what in themselves would have to be rated as defects of the familiar world to the conclusion that this world is the work of a being without defect, who will in the future make good all present deficiencies.

ARGUMENTS AGAINST IMMORTALITY

Philosophers opposed to the belief in immortality have generally confined their case to attacking weaknesses in the arguments thought up by immortalists. But some have also advanced arguments intended to show that human beings do not survive the death of their bodies. Thus, in the essay mentioned above Hume was not satisfied with pointing out the flaws in the metaphysical and the moral arguments for immortality but urged a number of considerations "from the analogy of nature" in favor of "the mortality of the soul." Similarly, many other writers who are not materialists and who are not committed to the view about the meaning of person words presented in this article maintain that there are powerful empirical grounds supporting a negative position on immortality. The most popular and impressive of these is what may be called the "body-mind dependence argument," an argument that, according to its more recent exponents, has received powerful confirmation from modern brain research. Bertrand Russell wrote,

> We know that the brain is not immortal, and that the organized energy of a living body becomes, as it were, demobilized at death and therefore not available for collective action. All the evidence goes to show that what we regard as our mental life is bound up with brain structure and organized bodily energy. Therefore it is rational to suppose that mental life ceases when bodily life ceases. The argument is only one of probability, but it is as strong as those upon which most scientific conclusions are based. (*Why I Am Not a Christian,* New York, 1957, p. 51)

Philosophers in the Hume-Russell tradition have also generally insisted that in the case of immortality the onus of proof must lie entirely with the believers. In their view it is quite wrong that we start with an open question. As this article urged at the beginning, the familiar facts of life and death establish an overwhelming presumption of mortality. Given these facts and the fact that person words mean what they do mean, there are massive philosophical obstacles to be overcome before the question of a future life can be shown to be sufficiently open to leave any room at all for appeals to evidence or even to faith.

Of course, there is nothing to stop anyone from giving what sense he likes to the expression "disembodied person." The difficulty is to attach enough sense to the expression so that some discovery about disembodied people could provide us with grounds for believing that we survive death. In their present senses person words have logical liaisons of the very greatest human importance. Personal identity in the present sense is the necessary condition of both accountability and expectation. This is only to say that it is unjust to reward or punish someone for something unless, as a minimum condition, he is the same person who did the deed and that it is absurd to expect things to happen to me in 2014 unless, as a minimum condition, there is going to be a person in existence in 2014 who will be the same person as I. The difficulty is to change the use of person words so radically that it becomes significant to talk of people's surviving dissolution without changing it in such a way that these crucial logical liaisons must be broken.

If this difficulty cannot be overcome—and there seems little reason to think that it can—then the apocalyptic words of the early Wittgenstein are to the point: "Our life is endless as the visual field is without limit. Death is not an event in life. Death is not lived through" (*Tractatus Logico-philosophicus,* London and New York, 1922, Secs. 6.431 and 6.1411).

See also Aristotle; Butler, Joseph; Death; Descartes, René; Ducasse, Curt John; Eternity; Galileo Galilei; Harvey, William; Hobbes, Thomas; Hume, David; Kant, Immanuel; Lewis, C. S. (Clive Staples); Locke, John; McTaggart, John McTaggart Ellis; Mill, John Stuart; Mind-Body Problem; Moral Arguments for the Existence of God; Parapsychology; Personal Identity; Plato; Pomponazzi, Pietro; Punishment; Reid, Thomas; Russell, Bertrand Arthur William; Ryle, Gilbert; Santayana, George; Socrates; Spinoza, Benedict (Baruch) de; Tertullian, Quintus Septimius Florens; Thomas Aquinas, St.; Wisdom, (Arthur) John Terence Dibben.

Bibliography

For a massive survey of the history of the doctrine of immortality, see W. R. Alger, *A Critical History of the Doctrine of the Future Life* (New York, 1871). For surveys of particular aspects of the doctrine, consult the bibliography of Antony Flew, ed., *Body, Mind, and Death* (New York: Macmillan, 1964), an anthology that includes a number of important pieces on immortality. The sources for the views given in the text are cited there.

St. Augustine's *De Immortalite Animae* has been translated by George G. Leckie in St. Aurelius Augustine, *Concerning the Teacher and On the Immortality of the Soul* (New York and London: Appleton-Century, 1938). Pietro Pomponazzi's *De Immortalite Animae* is available in an English translation by W. H. Hay II, revised by John H. Randall Jr., in *The Renaissance Philosophy of Man,* edited by Ernst Cassirer, Paul Oskar Kristeller, and John H. Randall Jr. (Chicago: University of Chicago Press, 1948).

Moses Mendelssohn's *Phaedon* (Berlin, 1767) is a famous dialogue on immortality that consciously imitates Plato's *Phaedo.* Hume's "Of the Immortality of the Soul" first appeared in *Two Essays* (London, 1777); it is reprinted in *A Modern Introduction to Philosophy,* 2nd ed., edited by Paul Edwards and Arthur Pap (New York: Free Press, 1965). The translation of Immanuel Kant's *Critique of Practical Reason* quoted in the text is that of L. W. Beck (Chicago: University of Chicago Press, 1949). Dugald Stewart's *Philosophy of the Active and Moral Powers of Man* is contained in *The Works of Dugald Stewart,* Vol. V (Cambridge, MA: Hilliard and Brown, 1829), pp. 366–423. J. M. E. McTaggart's views can be found in *Some Dogmas of Religion* (London: Arnold, 1906). For C. J. Ducasse's views see *Nature, Mind, and Death* (La Salle, IL: Open Court, 1951) and *The Belief in a Life after Death* (Springfield, IL, 1960). Joseph Butler, *The Analogy of Religion* (London, 1736), Part I, Ch. 1, is a detailed argument for immortality with replies to objections. John Fiske, *The Destiny of Man* (Boston and New York: Houghton Mifflin, 1884) and *Life Everlasting* (Boston: Houghton Mifflin, 1901); William James, *Human Immortality* (Boston: Houghton Mifflin, 1898; 2nd ed., with reply to critics, 1917); Harry Emerson Fosdick, *The Assurance of Immortality* (New York, 1926); John Baillie, *And the Life Everlasting* (London, 1934); A. E. Taylor, *The Christian Hope of Immortality* (London: Unicorn Press, 1938); and Jacques Maritain, *The Range of Reason* (New York: Scribners, 1952) all present arguments for immortality, as does "Ten Reasons for Believing in Immortality," a sermon by John Haynes Holmes, published in *The Community Pulpit* (1929–1930) and reprinted in Edwards and Pap, op. cit. G. T. Fechner, *Büchlein vom Leben nach dem Tode,* translated by John Erskine as *Life after Death* (Chicago, 1906), is a panpsychist defense of immortality. A. S. Pringle-Pattison, *The Idea of Immortality* (Oxford: Clarendon Press, 1922), holds that if there is personal immortality, it is not the inherent possession of every human soul but must be won by continual effort. E. S. P. Haynes, *The Belief in Personal Immortality* (Londo: Watts, 1913); Chapman Cohen, *The Other Side of Death* (London, 1922); Clarence Darrow, *The Story of My Life* (New York: Scribners, 1932); and Corliss Lamont, *The Illusion of Immortality* (London and New York, 1932) argue against immortality. Bertrand Russell, *Why I Am Not a Christian and Other Essays* (London: Allen and Unwin, 1957), should be consulted for a representative sample of Russell's arguments against immortality. A negative position similar to that presented in this article is found in C. B. Martin, *Religious Belief* (Ithaca, NY: Cornell University Press, 1959).

C. D. Broad's *The Mind and Its Place in Nature* (London: Kegan Paul, 1925), though somewhat dated, is perhaps the most valuable work on the problem of mind and body so crucial to the question of immortality. In Section D, Broad argues against the moral argument for immortality.

The source of the discussion of personal identity is Locke's *Essay concerning Human Understanding,* edited by A. C. Frazer, 2 vols. (Oxford: Clarendon Press, 1894), Vol. II, Ch. 27. For the contributions of Butler, Hume, and Reid and for later works on this topic, see the bibliography of the Personal Identity entry.

The classic parapsychological study of immortality is F. W. H. Myers, *Human Personality and Its Survival of Bodily Death* (London and New York: Longmans, Green, 1903). See also Ducasse, op. cit.; C. D. Broad, op. cit., *Religion, Philosophy, and Psychical Research* (London: Routledge, 1953), and *Lectures on Psychical Research* (London: Routledge, 1962); Antony Flew, *A New Approach to Psychical Research* (London: Watts, 1953); and the bibliography of the Parapsychology entry. For the investigations of psychical researchers into the nature of the experiences of spirits, see, for instance, G. N. M. Tyrrell, *Apparitions* (London: Society for Psychical Research by G. Duckworth, 1953). J. R. Hick, "Theology and Verification IV," in *Theology Today,* 17 (April 1960), reprinted in *The Existence of God,* edited by J. R. Hick (New York: Macmillan, 1964), presents a view opposite to the thesis of this article on the difficulty of identifying the reconstituted person with the original person.

F. C. Copleston, *Aquinas* (Harmondsworth, U.K.: Penguin, 1955), gives a sympathetic and illuminating account of Thomas's views on immortality.

For the most important recent criticism of the Cartesian doctrine of the soul, see Gilbert Ryle, *The Concept of Mind* (London and New York: Hutchinson, 1949). For a Wittgensteinian criticism see P. T. Geach, *Mental Acts* (London: Routledge & Paul; New York: Humanities, 1957).

For the view that the bodily occurrences involved in rational behavior can be completely explained only supernaturally, see, for example, C. S. Lewis, *Miracles* (London and New York, 1957), especially Ch. 2. For a fuller development of the view that such occurrences can be the subject matter of the physiologist while logical questions as to the rationality of the performance considered as an argument can be raised and answered, see Antony Flew, *Hume's Philosophy of Belief* (New York: Humanities Press, 1961), Ch. 5.

Antony Flew (1967)

IMMORTALITY [ADDENDUM]

The arguments against immortality on offer at the beginning of the twenty-first century are essentially those dis-

cussed by Professor Flew in the first edition of this reference work and will only be summarized here, following which we will consider non-traditional conceptions of immortality, and arguments in support of immortality. Four principal arguments have been presented against immortality: (1) The notion of a disembodied person is incoherent: to be a person just is to be embodied. (2) Advances in neuroscience have led to increased understanding of the brain and support physicalist theories in philosophy of mind. Physicalism (and *a fortiori* philosophical naturalism) denies the existence of an immaterial soul. The dependence of the mind on the functioning brain then grounds a formidable argument against the possibility of surviving the dissolution of the brain. (3) Determining identity conditions that would enable us meaningfully to assert that a particular disembodied soul was identical to a particular deceased person is a problem with no clear answers. Plausibly, souls would be individuated by mental contents. But it is logically possible for two minds to have identical contents, and in the absence of some bodily criterion of identity, it is not possible to tell which one has the genuine contents. Hence even if Smith somehow survived death, there would be no way to know that a particular soul was Smith. (4) The theory of evolution challenges the belief that there is anything special about humans in virtue of which we, and not other living things, would have immortal immaterial souls. But the conclusion that all living things are immortal—bacteria to baboons, barnacles to bananas—is absurd.

NONTRADITIONAL CONCEPTIONS OF IMMORTALITY

The difficulties inherent in the traditional concept of immortality, coupled with a pervasive human fear of death, have led to nontraditional conceptions that affirm immortality even while denying the survival of the person after death.

(1) CURING DEATH. Advances in biomedical technologies suggest that death is a disease to be cured. Immortality thus is conceived as prolongation of human life. Whereas it may be theoretically possible to postpone death indefinitely (through technologies such as cryogenics or genetic engineering), the desirability of prolonging life indefinitely is a significant ethical question that raises troubling questions about human nature and the purpose of life. Leon R. Kass (2001) argues that deferring death would radically change social institutions and undercut the universal human drive to procreation and protection of offspring, even to sacrifice on their behalf. Kass urges us to "resist the siren song of the conquest of aging and death" (p. 24). Regardless of the merits of Kass's argument, it is doubtful that many of us would embrace immortality if it were just "more of the same." Would this not be what Bernard Williams called "the tedium of immortality"?

(2) CYBERNETIC IMMORALITY. Assuming that the mind—thoughts, memories, feelings, dispositions—is not an immaterial substance but is reducible to patterns of neural activity, advances in computer science (nanotechnology, quantum computing) offer the prospect of transferring the entire contents of one's brain to a computer chip—far less likely to wear out than an organic body, and easily transferred again if chip failure becomes imminent. In *The Age of Spiritual Machines* (1999), Ray Kurzweil predicts such immortality will be achieved by uploading the contents of our brains into ever-better computers. Frank J. Tipler (1994) proposes a similar cybernetic immortality.

Granting for sake of argument that a computer program could instantiate Jones's mind, would the software offer immortality to Jones? Not only do familiar problems of personal identity plague this scenario, but it seems that (at least normally) humans are relational creatures who interact with their environment and with other persons by means of their bodies. In what sense then would the silicone-chip Jones be human? And if not human, how could it be Jones?

(3) ANTIREALISM. An antirealist view of immortality sees "eternal life" as a matter of quality, not quantity, of life. There seems to be a rather widespread sense that death is somehow transcended through one's projects or progeny. But this surely is not immortality traditionally conceived. Nor is the antirealist theological view of D. Z. Phillips (1970), according to which eternal life is participation in the life of God—that is, personal transformation in this temporal life.

(4) REINCARNATION. John Hick melds Christian and Eastern ideas to reinterpret immortality, resurrection, and reincarnation as "the divine creation in another space of an exact psycho-physical 'replica' of the deceased person" (1976, p. 279). Hick sees such divine recreations as occurring multiple times and in different "spaces" within a multiverse. Critics of Hick's view grant that God would have the power to create an exact replica, but deny that the replica would in fact be the person who died. The right sorts of connections that would preserve identity through recreation do not obtain. Further, the possibility

of multiple replicas raises both metaphysical and epistemological reasons to deny that the replica is identical to the original in any interesting sense.

DEFENSES OF IMMORTALITY

Clearly, if God exists, the prospects for an afterlife increase dramatically. A remarkable development in analytical philosophy in the last four decades has been the resurgence of philosophy of religion and natural theological arguments for God's existence. Indeed, Professor Flew (2004) himself, while still denying the possibility of immortality, recently acknowledged a significant change in his beliefs prompted by the Argument from Design. It is no surprise then that arguments for immortality, generally in the form of bodily resurrection, have also seen a renewal. Defenses of the immortality of the disembodied soul have been largely absent from recent literature.

CHRISTIAN VIEWS. Christian analytic philosophers have offered a variety of philosophical arguments defending the possibility of immortality through resurrection. There are two views on the nature of such immortality. Dualists affirm that the mind/soul is an immaterial substance distinct from the body, whereas monists adopt a physicalist philosophy of mind that denies the existence of an immaterial soul.

Christian dualists generally reject a strong Cartesian form of dualism. John W. Cooper (2000) argues that "holistic dualism" is the proper Christian view. Richard Swinburne (1997) offers a modal argument in defense of substance dualism, and William Hasker (1999) presents an argument from the unity of consciousness in defense of his emergent substance dualism.

J. P. Moreland and Scott B. Rae (2000) argue that free will and agency entail an immaterial mind. Moreland, with Gary R. Habermas (1992), also considers the empirical support offered by near-death experiences. A significant body of recent literature, beginning with Raymond A. Moody's *Life After Life* (1975), documents perimortem experiences. Habermas and Moreland point to features such as perceptions of physical facts that the clinically dead patient would have been physically unable to perceive which, if veridical, strongly support the possibility of conscious existence apart from the body.

Stephen T. Davis (1989) defends the traditional Christian understanding of the soul existing temporarily in a disembodied state, followed by the reunion of the soul with a resurrected body. Because personal identity is grounded in the soul rather than in the physical body, whether any or all of the atoms constituting the body of a person at death are incorporated in the resurrected body is not a significant issue. Davis further argues that the problems concerning individuating disembodied souls confuse criteria of identity with evidence for identity; epistemological uncertainty about identity does not undercut the possibility of genuine metaphysical individuation.

Christian monists reject the possibility of disembodied existence, yet hold to immortality gained through a resurrected body. Because the resurrection is eschatological, monists generally hold that personal identity is compatible with a temporal gap. Peter van Inwagen (1978) suggests that at death God miraculously preserves the essential core physical component(s) of the body, from which God reconstructs the resurrection body, thus preserving personal identity by means of the right sort of causal connections. Nancey Murphy (2002) accepts a physicalist anthropology, and contends that all the personal attributes that constitute personal identity supervene on the body, so if God creates a perfect replica of the body, that body will subserve all the necessary attributes constituting identity.

JEWISH VIEWS. Two poles of Jewish thought are seen as early as the Middle Ages. Maimonides (1135–1204) believed that the soul is immortal only because God sustains it. Although he wrote his *Essay on the Resurrection* to silence critics' claims that he denied the doctrine, he believed that the resurrected dead will eventually die again. Nachmanides (1194–1270) defended resurrection as the teaching of the Torah, while speculating on the ethereal nature of the resurrection body.

The tension in Jewish philosophy between immortality and resurrection may be traced into the present. Moses Mendelssohn (1729–1786), the first modern Jewish philosopher, was also the last to argue that the existence of the afterlife was rationally demonstrable (because the soul is by nature indestructible). The existentialist thinkers Martin Buber (1878–1965) and Franz Rosenzweig (1886–1929) tended to place discussion of the afterlife in a mythical category. Neil Gillman (1997) moves from the theological argument that God is more powerful than death, to philosophical arguments that a person is both body and soul, a psychophysical unity, and concludes that body and soul will be resurrected. In general, contemporary Reformed and Conservative Jewish thinkers reject bodily resurrection in favor of spiritual immortality, whereas Orthodox Jews retain belief in resurrection.

MUSLIM VIEWS. The nature of the soul occupies a prominent place in classical Islamic philosophy. Both the Neoplatonist Al-Kindī (d. 870) and the great Aristotelian Avicenna (Ibn Sīnā, 980–1037) asserted that the rational soul, being simple, is naturally indestructible. Al-Fārābī (875–950) held that only the rational soul that has knowledge of universals—eternal aspects of the universe—is indestructible. Averroes (Ibn Rushd, 1126–1198), departing from Aristotle, argued that the acquired intellect was indestructible because it was one with the divine mind; for this he was condemned by fellow Muslims as well as Christians and Jews.

Contemporary Islamic philosophers familiar in Europe and North America tend to focus on politico-ethical thinking (e.g., Rachid Ghannoushi, Mohamed Arkoun) rather than metaphysics or philosophy of religion. However, belief in immortality, and more particularly in resurrection, the Day of Judgment, and heaven and hell (the specific characters of which vary widely among different Islamic sects), is deeply embedded in all branches of Islam.

See also al-Fārābī; al-Kindī Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotle; Averroes; Avicenna; Buber, Martin; Cartesianism; Cybernetics; Dualism in the Philosophy of Mind; Islamic Philosophy; Jewish Philosophy; Maimonides; Mendelssohn, Moses; Philosophy of Mind; Physicalism; Reincarnation; Rosenzweig, Franz.

Bibliography

Cooper, John W. *Body, Soul, and Life Everlasting: Biblical Anthropology and the Monism-Dualism Debate*. 2nd ed. Grand Rapids, MI: Eerdmans, 2000.

Davis, Stephen T. "The Resurrection of the Dead." In *Death and Afterlife*, edited by Stephen T. Davis. New York: St. Martin's, 1989.

Flew, Antony. "My Pilgrimage from Atheism to Theism." *Philosophia Christi*. 6 (2004): 197–211.

Gillman, Neil. *The Death of Death: Resurrection and Immortality in Jewish Thought*. Woodstock, VT: Jewish Lights, 1997.

Hasker, William. *The Emergent Self*. Ithaca, NY: Cornell University Press, 1999.

Hick, John. *Death and Eternal Life*. London: Macmillan, 1976.

Kass, Leon R. "L'Chaim and Its Limits: Why Not Immortality?" *First Things* 113 (2001): 17–24.

Kurzweil, Ray. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Viking, 1999.

Moody, Raymond A. *Life after Life*. Atlanta: Mockingbird Books, 1975.

Moreland, J. P., and Gary R. Habermas. *Beyond Death: Exploring the Evidence for Immortality*. Eugene, OR: Wipf & Stock, 2004.

Moreland, J. P., and Scott B. Rae. *Body and Soul: Human Nature and the Crisis in Ethics*. Downers Grove, IL: InterVarsity, 2000.

Murphy, Nancey C. "The Resurrection Body and Personal Identity." In *Resurrection: Theological and Scientific Assessments*, edited by Ted Peters, Robert John Russell, and Michael Welker. Grand Rapids, MI: Eerdmans. 2002.

Phillips, D. Z. *Death and Immortality*. London: Macmillan, 1970.

Swinburne, Richard. *The Evolution of the Soul*. Rev. ed. Oxford: Clarendon Press, 1997.

Tipler, Frank J. *The Physics of Immortality: Modern Cosmology, God, and the Resurrection of the Dead*. New York: Doubleday, 1994.

Van Inwagen, Peter. "The Possibility of Resurrection." *International Journal for Philosophy of Religion* 9 (1978): 114–121

Garrett J. DeWeese (2005)

IMPARTIALITY

Impartiality is a more complex concept than is generally recognized. Judging a person to be impartial is not as straightforward as judging a person to have some moral virtue such as kindness or trustworthiness. People do not even understand what it means to claim that one is impartial unless they know both the group toward which that person is impartial and the respect in which one is impartial with regard to that group. The impartiality required by morality also requires a specification of the group toward which morality requires impartiality and the respect in which it requires impartiality with regard to that group.

The most common characterization of general impartiality is that it requires that like cases be treated alike. Almost all philosophers take this characterization as trivially true, but it is mistaken. Consider a baseball umpire who is upset because he believes that umpires are not appreciated. While staying within the accepted interpretations of the rule, he changes the strike zone every three innings; he starts with a widest zone, goes to the narrowest one, and then returns to a widest one. If he changes without regard to which team benefits or is harmed by this change, then he is impartial with regard to the two teams in calling balls and strikes. Because he does not treat like cases alike—that is, he calls balls and strikes differently in the first and fifth innings—he is a bad umpire, but he is still completely impartial with regard to the two teams with respect to calling balls and strikes. He is inconsistent, but inconsistency should not be confused with impartiality. A good umpire must be consistent as well as impartial. An inconsistent umpire will be sus-

pected of not being impartial, but when the disgruntled umpire is not influenced at all by who is benefited or harmed, he remains impartial with respect to calling balls and strikes with regard to the two teams.

A person is impartial with regard to a group in a specified respect insofar as that person acts impartially in that respect with regard to that group. The basic concept of impartiality is defined as follows:

> A is impartial in respect R with regard to group G if and only if A's actions in respect R are not influenced at all by which member(s) of G are benefited or harmed by these actions.

A teacher can be impartial with regard to a group G—for example, the students in her class in respect R; or, for example, grading their exams—but not impartial in a different respect, such as calling on them in class, for she may favor boys over girls in this respect. Two umpires, both consistent and impartial with regard to two teams, need not be impartial with regard to pitchers and batters. If one prefers a higher scoring game and the other a lower scoring one, they may, within the accepted interpretations, call some pitches differently. Both show partiality toward pitchers or toward batters, but both are still impartial with regard to the two teams.

Some contemporary consequentialists claim that morality requires impartiality whenever any sentient being's interests are involved. However, not only is there disagreement about whether all sentient beings are included in the group toward which morality requires impartiality, it is generally recognized that even with agreement about the group, morality does not require impartiality with respect to all actions affecting people's interests. It is generally agreed that morality does not even require impartiality when following moral ideals—for example, relieving or preventing pain, or helping the needy. Unless one does not act on these ideals at all, it is impossible to act on them impartially even with regard to all moral agents; no one can relieve or prevent pain impartially with regard to all moral agents. The only respect in which morality requires impartiality is with respect to violating moral rules—for example, those rules prohibiting killing, causing pain, deceiving, and breaking promises. It is only with regard to these kinds of moral rules—those that can be formulated as prohibitions—that it is humanly possible to act impartially with regard to a group large enough to be an appropriate group.

The examples of the teacher and umpire show that the group with regard toward which impartiality is usually required is often small and usually does not include the agent. The impartiality required by morality differs from this kind of impartiality in that it requires impartiality with respect to violating a moral rule toward a group composed of at least all moral agents, including the person violating the rule. Morality requires impartiality with regard to those moral agents affected by a violation of a moral rule—for example, being partial toward friends is not morally allowed. It also requires impartiality with respect to whether one can violate a moral rule; that is, it is not morally allowed to violate a rule in circumstances if it would be irrational to be willing for everyone to know that they are allowed to violate the rule in those same circumstances.

Sometimes all impartial rational persons favor violating a moral rule—for example, deceiving a hired killer in order to save an innocent person's life. Because morality always requires impartiality with respect to violating moral rules, it must be possible to violate a moral rule and still be acting impartially in this respect. This kind of impartiality can be achieved by violating a moral rule only when one would be willing for everyone to know that they are allowed to break the rule in the same circumstances. This achieves Kant's point about morality not allowing a person to make special exceptions for herself without creating the kinds of problems caused by the claim that morality requires acting on the categorical imperative.

Kant claims that morality requires that the group with regard to which one must be impartial with respect to violating a moral rule include only moral agents, that is, those persons who are required to act morally. Jeremy Bentham claims that the group includes all sentient beings. Most people, including most philosophers, do not agree with either Kant or Bentham; almost all want to include infants and children in the group toward which morality requires impartiality, but there is considerable disagreement about whether morality requires impartiality with regard to fetuses or to nonhuman animals. However, many who think that morality does not require impartiality with regard to nonhuman animals hold that morality does provide some protection to sentient nonhuman animals.

Because the concept of impartiality presupposes that there be some group with regard to which one is impartial, it does not make sense to claim that there is an impartial method for picking the group with regard to which morality requires impartiality. Recognizing that rational persons can differ about the composition of the group with regard to which morality requires them to be impartial helps explain the moral disputes concerning abortion and the treatment of animals. Morality limits the freedom of moral agents, so that the larger the group

with regard toward which morality requires impartiality with respect to violating a moral rule, the greater the limitation on the freedom of moral agents. A rational person can rank this freedom of moral agents higher than the welfare of nonmoral agents or vice versa. The former is more likely to hold that morality does not require impartiality with regard to nonmoral agents, whereas the latter may hold that it does.

Even when there is agreement about the composition of the group with regard to which morality requires impartiality with respect to violating a moral rule, rational persons who are impartial with regard to all members of this group can still disagree. Because rational persons can rank the various evils—for example, death, pain, and disability—differently, one impartial rational person can favor everyone knowing that they are allowed to break a rule in circumstances in which another impartial rational person would not favor this. Impartiality does not require unanimity, as some philosophers such as Kant and Rawls seem to claim. If it did, then assuming that all Supreme Court justices know all of the relevant information, and do not suffer from any other mental dysfunction, one would be forced to hold that whenever the United States Supreme Court issues a split decision, at least one justice is not impartial.

See also Bentham, Jeremy; Kant, Immanuel.

Bibliography

Baier, Kurt. *The Moral Point of View*. Ithaca, NY: Cornell University Press, 1958. Abridged edition, New York: Random House, 1965.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation* 1789. Many editions.

Firth, Roderick. "Ethical Absolutism and the Ideal Observer." *Philosophy and Phenomenological Research* 12 (1952): 317–345.

Gert, Bernard. *Common Morality: Deciding What to Do*. New York: Oxford University Press, 2004.

Gert, Bernard. *Morality: Its Nature and Justification*, rev. ed. New York: Oxford University Press, 2005.

Hare, R. M. *Moral Thinking*. Oxford: Oxford University Press, 1981.

Kant, Immanuel. *Grounding for the Metaphysics of Morals.* 1785. Many editions.

Mill, John Stuart. *Utilitarianism* 1863. Many editions.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Singer, Marcus George. *Generalization in Ethics*. New York: Alfred A. Knopf, 1961.

Singer, Peter. *Practical Ethics*. Cambridge, U.K.: Cambridge University Press, 1993.

Bernard Gert (1996, 2005)

IMPETUS

Aristotle distinguished between two sorts of motion: natural and unnatural. Natural motions were those induced by the elemental constituents of things to seek their natural places—the earth at the center of the cosmos, fire under the periphery, water and air in their intermediate locations. Anything not in its natural place (i.e., not at that point in the stratification of things appropriate to its elemental composition) has an internal inclination to reach its natural place that will be exercised so long as nothing impedes it. Moreover, the speed of any body in natural motion (i.e., approaching its natural place) is a function of its heaviness in the case of downward motion, lightness in the case of upward, and an inverse function of the resistance of the medium through which it moves.

Not all motions, however, are natural; heavy objects can be hurled upwards, buoyant ones forcibly submerged. These unnatural motions are the result of force, yet Aristotle also notoriously held that in any change (including change of position) there must be a continuously acting agent of change; in the case of projectile motion, he supposed that the original action of the thrower endowed successive enveloping portions of the medium (air or water) with the ability both to receive and to transmit motive force; and so the projectile continues to move after it leaves the thrower's arm as a result of the continuing—albeit diminishing—power successively induced in the surrounding elastic medium. This explanation was often felt to be less than adequate; John Philoponus explicitly rejected it, supposing rather that the thrower imparted a certain quantum of force into the projectile, which it gradually exhausted in the course of its flight until it fell to earth.

Thus Philoponus crucially rejects the Aristotelian assumption that there must be continuous contact between a thing in motion and some external mover of it, speaking of an "induced power" (*endotheisa dunamis*) possessed—albeit temporarily and in this sense nonnaturally—by the moving body. This was to become the *vis impressa* of medieval theorists such as Jean Buridan, who attacked the reemerging Aristotelian orthodoxy with vigor (although perhaps without fully understanding it).

Impetus theory (the term is owed to Pierre Maruice Marie Duhem) thus maintains that all motion relies ultimately on the transmission of force from a mover, although not necessarily simultaneously with that motion, as Aristotle required. The theory also maintained that force gradually diminishes, although different versions of the impetus theory disagreed as to whether it

does so simply because all movement requires force, and that force is like a fuel to be consumed, or (as some writers in the Arabic tradition such as Avicenna apparently held), only as a result of contact with a retarding medium (and hence that in a vacuum the impetus would continue for ever). Whereas both versions retain the ancient commitment to the view that all action requires a continuous active cause, the latter was a step in the direction of the inertial notions that would revolutionize physics. Galileo Galilei, in his early *De Motu* (1590) is still an impetus theorist, and gratefully acknowledges Philoponus and other predecessors for showing the way beyond Aristotle. By the time of the mature physics of the *Discorsi* (1638), impetus theory itself has been left decisively behind.

See also Aristotle; Buridan, Jean; Galileo Galilei; Philoponus, John.

Bibliography

Wolff, Michael. "Philoponus and the Rise of Preclassical Dynamics." In *Philoponus and the Rejection of Aristotelian Science*, edited by R. Sorabji. London: Duckworth, 1987.

R. J. Hankinson (2005)

INDETERMINISM

See *Chance; Determinism, A Historical Survey; Determinism and Freedom; Determinism and Indeterminism*

INDEXICALS

Suppose that Natasha says "I am right and you are wrong" to Joey. Natasha's utterance of "I" designates Natasha and her utterance of "you" designates Joey. The truth-conditions of her statement are that Natasha is right and Joey is wrong.

Now suppose that Joey responds by uttering the exact same words back to Natasha: "I am right and you are wrong." He has said the same words, with the same meaning, but he has not said the same thing. Joey's utterance of "I" designates Joey and his utterance of "you" designates Natasha. The truth-conditions of his statement are that Joey is right and Natasha is wrong. Joey has directly disagreed with Natasha.

In this article, "meaning" refers to the rules or conventions that are associated by a language with the expressions in it, the rules that one learns when one learns the language. Given this, the meanings of Natasha's words and of Joey's are the same. What differs is the objects the particular expressions designate and the truth conditions of the statements. This aspect of utterances will be called "content."

The crucial differences between the first and second utterances were the speakers and the addressees. Such facts about an utterance can be called its "context." Differences in the contexts of the utterances account for the differences in their contents.

(The role of context in this case differs from that in a case of homonymity or ambiguity. With homonymity the context helps us determine which word is being used; with ambiguity, which meaning of a word or phrase is being used. But in this case context still has a role to play after questions of words and meanings have been settled. The meanings of "I" and "you" direct us to features of the context, to determine who is designated.)

The content of an utterance using "I" or "you" is determined by contextual facts about the utterance in accord with their meaning. Such expressions we call "indexicals."

In addition to "I" and "you," the standard list of indexicals includes the personal pronouns "my," "he," "his," "she," "it," the demonstrative pronouns "that" and "this," the adverbs "here," "now," "today," "yesterday," and "tomorrow," and the adjectives "actual" and "present" (Kaplan 1989). The words and aspects of words that indicate tense are also indexicals. And many other words—for instance, "local"—seem to have an indexical element.

According to David Kaplan's account, each indexical, and each sentence containing an indexical, has a meaning or character that is a function from contexts to content. The character of "I" is a function whose value, for each context, is the speaker or agent of that context. The character of "now" is a function whose value, for each context, is the time of that context. The character of "you" is a function whose value, for each context, is the person addressed by the speaker in that context. The character of the sentence spoken by Natasha and Joey is a function whose value, for a context with a speaker x and an addressee y, is the proposition that x is right and y is wrong. Natasha and Joey's words have the same characters, but their utterances have different contents.

In the formal development of his theory, Kaplan equates content with the intensions of intensional semantics. He criticizes earlier attempts to provide a formal theory within this framework for treating contexts on a par with "circumstances of evaluation" (Kaplan 1989, pp. 507ff.). The context determines which proposition is expressed by Joey's utterance of "I am right and you are

wrong"; the circumstance of evaluation determines whether or not the proposition is true. The necessity for such a distinction was seen by Hans Kamp (1971).

(Kaplan notes that at the level of character it makes sense to talk about the logic of indexicals. "I am here now" is a truth in the logic of indexicals, because, given its character, this sentence will have a true content at each context. The content will be contingent and can be expressed by a sentence that is not a logical truth.)

Kaplan's concept of content corresponds to "what is said" by an utterance (let us call this "official" content). This is what someone who knows the meaning and the context grasps. Other philosophers have thought it important also to bring in the concept of token-reflexive or diagonal content. This is what someone who knows the meaning but does not know the context grasps (Burks 1949, Perry 1993, Stalnaker 1981).

Consider an utterance *u* of "Je ne comprends pas l'anglais" made by Erin during a cocktail party. Suppose that Natasha hears the words and understands French but does not see who said them. Joey hears the words, understands French, and also sees that Erin said them. Based on her knowledge of French, Natasha can assign utterance-reflexive truth conditions to *u:* Natasha knows that *u* is true iff (if and only if) (1) the speaker of *u* does not understand English. Joey, since he knows who is talking, can assign nonreflexive truth conditions to *u:* Joey knows that *u* is true iff (2) Erin does not understand English. Natasha knows what the world has to be like for *u* to be true, given the meaning of the words in *u.* Joey knows what the world has to be like, given the meaning of the words in *u* and the relevant facts about context. What Joey knows, (2), is the official content of Erin's remark. It is what we would ordinarily say Erin said. Erin did not say (1): She did not make a remark about her own utterance. Nevertheless, (1) corresponds to an important level of understanding that we must take account of to explain the cognitive significance of sentences containing indexicals. (When Erin said what she did, she probably wanted her listeners to grasp that the person in front of them, at whom they were looking and with whom perhaps trying to converse in English, did not understand that language. This would be an easy inference from the proposition expressed by (1)—that the person who was producing the utterance they were hearing did not understand English. To understand Erin's plan, we seem to need the reflexive content of Erin's remark, and not only its official content.)

See also Kaplan, David; Philosophy of Language.

Bibliography

Burks, A. "Icon, Index, and Symbol." *Philosophical and Phenomenological Research* 9 (1949): 673–689.

Castañeda, H.-N. "'He': A Study in the Logic of Self-Consciousness." *Ratio* 8 (1966): 130–157.

Evans, G. "Understanding Demonstratives." In *Meaning and Understanding,* edited by H. Parret and J. Bouveresse. Berlin: W. de Gruyter, 1981.

Kamp, H. "Formal Properties of 'Now.'" *Theoria* 37 (1971): 237–273.

Kaplan, D. "Demonstratives." In *Themes from Kaplan,* edited by J. Almog, J. Perry, and H. Wettstein. New York: Oxford University Press, 1989.

Nunberg, G. "Indexicality and Deixis." *Linguistics and Philosophy* 16 (1993): 1–43.

Perry, J. *The Problem of the Essential Indexical and Other Essays.* New York: Oxford University Press, 1993.

Recanati, F. *Direct Reference: From Language to Thought.* Oxford: Blackwell, 1993.

Stalnaker, R. "Indexical Belief." *Synthese* 49 (1981): 129–151.

Yourgrau, P., ed. *Demonstratives.* Oxford: Oxford University Press, 1990.

John Perry (1996)

INDIAN PHILOSOPHY

The "India" in question is the Indian subcontinent—the land constituting present-day India, Pakistan, Bangladesh, and surrounding countries such as Srī Lanka to the south and Bhutan, Sikkim, Afghanistan, and Nepal to the north. And although *philosophy* in the sense in question covers much of what is covered by the term *philosophy* in its contemporary usage in English-speaking countries, it also has a specific use in the Indian context, in which it refers to the thoughts expressed in the literature relating to liberation (*mokṣa; nirvāṇa*). In this usage, philosophy, and the philosophical literature of India, is contrasted in Indian thinking with the literature pertaining to other matters, notably the literature concerned with political and social concerns (*arthaśāstra*), with interpersonal relations such as the sexual and aesthetic dimensions of love (*kāmaśāstra*), and with morals (*dharmaśāstra*), each of which has a pertinent literature of its own. The "philosophical" literature of India, then, relates to ultimate concerns, especially how to achieve liberation from rebirths and the nature of a universe in which liberation is possible and available. It is a literature that does not primarily include such Western fields of philosophy as political and social philosophy (for that is *artha*), aesthetics (for that is *kāma*) and ethics (for that is *dharma*). It also does not include the literature concerning the natural and social sciences (although it is arguable that parts

of Indian philosophy are offshoots of aspects of early Indian protoscience) or the applied sciences (agriculture, astronomy, and so on); nor does it include the domain of poetry and prose literature.

Whether Indian philosophy overlaps religion or not is a matter of how one thinks of "religion." The majority of the early Indian philosophical systems (*darśana*) do not acknowledge, and in some cases explicitly deny, the existence of a supreme being or lord (*īśvara*). All classical Indian thinking accepts gods (*deva*). They are viewed as unliberated, like humans; they are beings who inhabit other realms and occasionally visit ours. They eventually live out their lengthy period as gods and are reborn into lower realms as humans or even animals. This process is part of the Indian theory of karma—accepted until modern times—according to which selves are beginningless and are caused by their past actions to inhabit a series of bodies ranging from insects (or even plants) up to gods, depending on the particular portion of the stored-up results of past actions (*karman*) that becomes activated (*prārabdha*) as one enters the next birth.

In what has been dubbed the "*bhakti* period"—the period beginning around the turn of the second millenium CE, many philosophical viewpoints became inspired by and wedded into one or another religious movement. These movements typically recognize and worship one or more of the Hindu deities such as Śiva or Viṣṇu, and their literature is a mixture of devotional and philosophical concerns. However, the systems that originated in the previous centuries—Sāṁkhya and Yoga, Nyāya and Vaiśeṣika, Pūrvamīmāṁsā and Advaita Vedānta, Buddhism and Jainism—have persisted up to the present. Their literature continues to expand, and for most of them (except Buddhist systems) there are still *maṭha*s and *āśrama*s in India where followers devote their lives to the study of one of these systems.

In the "modern" period—from the nineteenth century to the present—the application of the term "Indian philosophy" has become more complex because the British-founded system of higher education has bred a group of philosophers (in the Western sense) who are native South Asians. Because of the broad Western connotation of the terms *philosophy* and *philosophers*, among modern Indian philosophers one finds not only academic philosophers but also profound and influential political and social thinkers such as Ramakrishna, Vivekananda, Aurobindo, and even Gandhi.

THE INDIAN PHILOSOPHICAL LITERATURE

This article covers, under a variety of topical headings, the philosophical views (*darśana*) discussed in the classical literature. The writers focus chiefly on the path that can lead to liberation from karmic bondage. These writers also defend the very possibility of gaining liberation against doubts. In mounting this defense, they explore the nature of the kind of universe that would allow for the working out of karma. They examine the very possibility of liberation and what it takes to confront and overcome the causes of bondage. These writings have yielded a rich variety of profound metaphysical, logical, and epistemic theories.

The language of this Indian philosophical literature is mainly Sanskritic. The broad designation "Sanskritic" includes not only Sanskrit itself but also vernaculars such as Pāli (the language of early Buddhist philosophical treatises) and Prakrit, the language of some early Jain works. In the case of a few of the "*bhakti*-period" movements, some of the philosophical literature comes to us in Tamil and occasionally other modern Indian languages. But for the most part, classical Sanskrit is the language of Indian philosophy.

Great foundational works of Sanskrit literature are frequently included within the literature of Indian philosophy, specifically works such as the Vedas, the Upaniṣads, the epics, the *Bhagavadgītā*, the Pāli canon, the canonical Jain works, and the *Bhāgavatapurāṇa*. These works certainly include matters that pertain to "Indian philosophy" as characterized above, but it is commentaries on these foundational works that are the locus of a significant portion of India's philosophical literature.

Of the many philosophical systems that have grown up in India, there is one basic text that has come to be viewed as the basic scripture for each system. Sometimes such works have a title that ends in "-*sūtra*s." Thus the *Nyāyasūtras, Vaiśeṣikasūtras, Mīmāṁsāsūtras*, and *Brahmasūtra*s play this basic role, respectively, for the Nyāya, Vaiśeṣika, Pūrva-mīmāṁsā, and Vedāntic schools of thought. For other systems a basic text (if there is one) has a different kind of title. From these basic texts—whether called "*sūtra*" or not—an interpretive literature grew up over the centuries that is a significant part of the Indian philosophical corpus. For example, in the case of Nyāya, there is a series of commentaries upon commentaries stemming from the basic *Nyāyasūtra*s. The number of commentaries and subcommentaries on the *Brahmasūtra*s is vast, because there are many Vedānta systems,

and each has its own interpretive literature based on that text.

THE SYSTEMS OF INDIAN PHILOSOPHY

Indian scholars traditionally speak of "six systems of Indian philosophy." But there are many more than six. A "system" or "school" (*darśana*, literally a "view") in this context constitutes a set of theories about liberation and the means to it based on a certain ontology, logic, and epistemology. This definition has to be understood in a loose way. The system known as Cārvāka has a theory about liberation—it denies that it is possible to be liberated—so its position is largely an extended polemic against the rest of the systems, whose theories are adopted specifically to account for and illuminate the possibility of liberation. Some of the Vedāntic systems place limitations on who is and who is not capable of liberation, and some of them elevate devotion to God to a position that equals or even surpasses liberation itself. The later Buddhist notion of the Bodhisattva who declines his own liberation until all beings have been liberated implies another exception to the general view that it is one's own liberation from karmic bondage that is the defining concern of Indian philosophy.

It is, indeed, impossible to give a finite list of Indian systems of philosophy. For one thing, new schools are being founded even now; their durability might be far from certain, but some recent ones have their adherents. For another thing, it is not always clear how to differentiate one system from another—it is not obvious, for example, whether those called "Buddhist logicians" are to be counted as a separate school of Buddhist thought or not; and there are clearly several disparate branches of Mīmāṁsā; there are an indefinite number of schools that call themselves "Vedānta."

ASSUMPTIONS COMMON TO ALL SYSTEMS (EXCEPT CĀRVĀKA)

As noted earlier, the Cārvākas do not accept liberation as a feasible goal. Their outlook has been culled from references in polemical passages by others attempting to refute their views and a limited number of literary works such as Jayarāśi's *Tattvopaplavasiṁha* and Kṛṣṇa Miśra's *Prabodhacandrodaya*.

All the other systems in this survey accept at least two relevant theses: The first is that there was no absolute beginning of things, that the series of lives each of us has lived is without beginning. This doctrine of beginninglessness (*anāditva*) entails, of course, that there can be no God who created us *ab initio* or who functions as the first cause of the universe. As we shall see, this does not necessarily stop Indian philosophers talking about God (*īśvara*); various roles are assigned to Him aside from that of ultimate creator.

The other thesis generally accepted by all systems except Cārvāka is what is often referred to as the "karma theory" or the "law of karma." Although many details about how karma works can be gleaned from the pages of the Indian philosophical literature, karma remains an assumption underlying all philosophical theories rather than a theory itself. It is infrequently defended, merely assumed.

What is this "karma theory"? First, given the assumption of the beginninglessness of selves, each person has always existed; each is always performing actions ("action" being the basic meaning of *karman*), at least some of which lay down "karmic traces" (*saṁskāra; vāsanā*) that are stored up in the agent until each is eventually "worked off" through performance of another action at some later date. These traces, which constitute each self's "karmic baggage," are carried through life and over into the next birth, where a certain portion of that baggage is identified or "ticketed" as requiring working off during that coming lifetime. In working off the ticketed portion, one performs more actions that in turn breed more traces, so that one must be born again and again in order to work off both stored-up karma from previous lives and un-worked-off karma from one's present life.

A commonly cited passage in the *Yogasūtra*s specifies three aspects of one's life that are determined by the traces stored up from previous lives. One is the kind of life one gets at each birth. Some part of one's karmic store, perhaps the traces of the latest or perhaps of the most virtuous or vicious actions, results in the coming birth's occurring at an appropriate place in the great chain of being—perhaps as an animal if the aspects are vicious, perhaps as a god if they are virtuous. A second aspect of one's life said to be governed by one's activated karma is the length of the life one is about to lead. And a third is the kind of experiences one is likely to have as one goes through this coming lifetime—relatively pleasant if good karma predominates, relatively unpleasant if bad karma predominates.

The philosophical literature often denies that karma implies fatalism. Although karmic traces are powerful, they are not indestructible. It is possible, although not easy, to resist the force of one's karma. That is why phi-

losophy has an important role to play in indicating the modes of thought and action that can avoid laying down new traces and thereby destroy the power of one's ticketed karma. Traveling the path to liberation usually requires the personal attention and advice of a teacher (*guru*) who can give advice about how to meditate and behave so as to bring the powers of past actions to heel.

EIGHT TYPES OF PHILOSOPHICAL SYSTEMS

Indian philosophical systems are divided into eight groups (nine if one includes the Cārvākas). It is traditional to distinguish a basic three systems: Jain, Buddhist, and Hindu. The Hindu systems are here distinguished into six groups: (1) Sāṁkhya and Yoga, which share a common metaphysics; (2) Nyāya and Vaiśeṣika, which share an ontology; (3) "Mīmāṁsā," more properly "Pūrvamīmāṁsā," whose members share a common approach to the interpretation of the authority of the Vedas, which they view mainly as a source of prescriptions about behavior; (4) "Vedānta," which treats the "closing sections of the Vedas" (*vedānta*)—the Upaniṣads—as authoritative; (5) a group of philosophical systems whose common ground is that their proponents are worshipers of Śiva; and (6) the Grammarians (*vaiyākaraṇa*), who view the study of language as providing the key to liberation. Many of these approaches claim ancient authority for their standpoints.

SĀṀKHYA AND YOGA

The characteristic terminology of these two systems, featuring terms such as *prakṛti* and *puruṣa*, is found in the earliest Indian literature, the Vedas, which date roughly from the end of the second millennium BCE. Sāṁkhya terms are also prominent in the great Indian epics, especially the *Mahābhārata*, for example in the portion that constitutes the *Bhagavadgītā*. There is reason to believe that several Sāṁkhya authors lived prior to the fourth century CE, at which time the basic Sāṁkhya text, the *Sāṁkhyakārikās*, ascribed to one Īśvarakṛṣṇa, appears to have been composed. This work plays the role of the system's basic *sūtra*s. A much later work by Kapila (1375) is named the *Sāṁkhyasūtra*s, but its claims to antiquity are usually disputed (though the name "Kapila" is ascribed to one of the otherwise unknown earlier sages cited in the later literature). One or two commentaries on the *Sāṁkhyakārikās* are regularly studied to elucidate that text. The most frequently cited is by Vācaspati Miśra (940), the *Tattvakaumudī*, but there are numerous others. And a few independent works on Sāṁkhya were written over the centuries: the *Yuktidīpikā*, of unknown authorship and date; and several works by a relatively late writer, Vijñānabhikṣu (1575), an interpreter.

The fundamental text of Yoga philosophy is the very popular, widely studied and quoted *Yogasūtra*s of Patañjali (300). An ancient (475) commentary, the *Yogabhāṣya*, is ascribed to someone named Vyāsa, and Vācaspati Miśra (940) has written a commentary on that named *Tattvavaiśāradī*. The term *yoga* is of course now standard throughout the world and no longer merely a Sanskrit term. In general usage it connotes techniques of breath control, bodily postures, and meditation, among the topics addressed in the texts just cited. The underlying metaphysics of Yoga, however, is the same as that of the Sāṁkhya ontology; in effect, Yoga provides the account of how to go about achieving liberation through meditation, whereas Sāṁkhya lays out the account of the ontological, logical, epistemogical, and psychological truths that form the basis of what is to be meditated upon.

Sāṁkhya (and thus Yoga as well) postulates two fundamental kinds of real entities: *puruṣas* and *prakṛti*. Each self in the universe is termed a *puruṣa* in this system, and there are as many *puruṣa*s in the world as there are embodied minds, perhaps even more if bugs and plants are included. However, a peculiarity of the system is that a self has only one function: to be a seat of consciousness. All other features ascribed to humans belong to *prakṛti*.

One might think that *prakṛti* thus corresponds to the "body" side of the mind/body dichotomy. But this is not true for Sāṁkhya. The consciousness that is *puruṣa* is merely "pure" consciousness, not any particular awareness or mental state. Particular modes of awarenesses, such as sensations, emotions, and mental events of all sorts, along with particular physical features such as the particular bodies, sense-organs, and activities of what is ordinarily called a "self," are features of a being's *prakṛti*. (Remember that selves are not limited to human beings—each center of consciousness from gods down to insects and perhaps plants are selves.)

Karmic bondage pertains only to one's *prakṛti*, since it is that being's *prakṛti* alone that changes from time to time, birth to birth. The true self, *puruṣa*, remains unaffected in reality; it appears to be affected only by karma. To use an analogy that is constantly appealed to, a *puruṣa* is like a lamp that lights up things that are themselves inert; some of those things—thoughts and modes of awareness, for example—may seem to us to be conscious, others in turn to be objects of those modes of awareness, but for Sāṁkhya fundamentally all *prakṛti*—whether psy-

chological activity or physical objects—is unconscious, inactive.

One must not, however, misunderstand Sāṁkhya as contending that all *prakṛti* is unreal. Unmanifest (*avyakta*) *prakṛti* is real, permanent, and is the real cause of real effects. Appealing to language of great antiquity we are told that *prakṛti* is made up of three *guṇa*s, called *sattva, rajas* and *tamas*. Every actual element in manifest *prakṛti*, that is, in the world that we experience, is different from every other, and the difference is due to which *guṇa*s are dominant in each thing and to what extent.

So for Sāṁkhya both causes and effects are real. Unmanifest *prakṛti*, made up of the three *guṇa*s, actually causes real effects (the term used is *pariṇāma*, usually translated as "evolves" or "transforms"), just as milk really produces real curds. And so there is no question about the reality of bondage either, since a state of bondage is the real product of causal forces that are themselves constituted by the balance of the three *guṇa*s that constitute that state.

For this system bondage is due to a failure to discriminate between what is permanent—the pure self and unmanifest *prakṛti*—and what is temporary—all the evolutionary states of *prakṛti* that we normally construe as our nature and the nature of the world and the universe. To learn to discriminate the permanent from the temporary is the purpose of meditation, of yoga. When properly discriminated, the self will no longer be subject to the limitations of *prakṛti* and will have achieved liberation.

NYĀYA AND VAIŚEṢIKA

These two systems merged by the end of the first millennium CE. At the start, though, there were two distinct sets of *sūtra*s, the *Vaiśeṣikasūtra*s ascribed to Kaṇāda (whose date is unknown but who lived probably during the first couple of centuries CE) and the *Nyāyasūtra*s ascribed to some Gautama (not the Buddha), who most likely flourished in the second century CE.

The *Vaiśeṣikasūtra*s lay out six categories of actual entities that the aspirant needs to thoroughly understand and that are intended to cover all the things that exist in the world. These six are substances (*dravya*), qualities (*guṇa*), motions (*kriyā*), univeral properties (*sāmānya*), individuators (*viśeṣa*), and inherence (*samavāya*). The *Vaiśeṣikasūtra*s, like most *sūtra*s, are extremely laconic and require a commentary to be understood. The earliest extant commentary is by Praśtapāda (530) entitled *Padārthadharmasaṁgraha*. Commentaries on Praśastapāda's work by Vyomaśiva (950), Śrīdhara (991), and Udayana (1054) develop accounts of the six categories and add a seventh, the category of negative entities or "absences" (*abhāva*). Most commentaries on the *Vaiśeṣikasūtra*s prior to the fifteenth century are known only from references by others, but there is a full commentary by Śaṁkara Miśra (1440).

The author of the *Nyāyasūtra*s works from a broad conceptual base that suggests its origins in a worldview with an approach broadly similar to that of Kaṇāda. Gautama starts from a list of sixteen topics, some epistemic, some ontological, but most of them harking back to the ancient practice of holding debates. The metaphysics, though not identical with Vaiśeṣika's categories, clearly presupposes them. The epistemology proposes four ways of gaining knowledge: through perception (*pratyakṣa*), inference (*anumāna*), comparison (*upamāna*), and authoritative language (*śabda*). The sixteen categories are as follows: instruments of knowledge, objects of knowledge, doubt, purpose, example, tenets, members of an inference, reductio ad absurdum, ascertainment, discussion, sophistry, cavil, fallacies, quibble, futile rejoinder, and ways of losing a debate. The last two "debate categories" by themselves take up the the entire fifth and last book of the *Nyāyasūtra*s.

A series of commentaries on commentaries is developed through the ages stemming from these *Nyāyasūtra*s, each commentary explaining the previous one in the list. This list comprises the *Bhāṣya* by Vātsyāyana (450), the *Vārttika* by Uddyotakara (610), the *Tātparyaṭīkā* of Vācaspati Miśra (960), and the *Pariśuddhi* by Udayana (1054). Important independent expositions of the Nyāya system are found in the *Nyāyamañjarī* by Jayanta Bhaṭṭa (870) and in several works by Udayana (1054).

Udayana is likewise the author of a widely studied book that has become the standard work concerning arguments for the existence of God: *Nyāyakusumāñjali*. In Nyāya-Vaiśeṣika God is another self, though one unlike us in never having been involved in the round of transmigration and rebirth. At the beginning of each era (*kalpa*) His function, according to Nyāya, is to cause the atoms, which were unmoving during the period (*pralaya*) between creations, to come into contact, thus starting the production of physical bodies that the selves (who have persisted throughout *pralaya*) inhabit through their karma. Since the *Vaiśeṣikasūtra*s themselves make no mention of God, and the *Nyāyasūtra*s mention Him only in reporting the opinion of an objector, it appears that the development of the role of God in Nyāya-Vaiśeṣika may have been influenced by the growing tendency toward

devotionalism in India toward the close of the first millennium CE.

The method of gaining liberation as described by Nyāya-Vaiśeṣika appears to be mainly intellectual, requiring study of the tenets of the system, which eventually removes the ignorance (*avidyā*) that occasions defects (*doṣa*) defects that occasion the desires and other mental attitudes that conduce to bondage.

With Gaṅgeśa (1320) an important new phase of Nyāya-Vaiśeṣika begins, known appropriately enough as *Navyanyāya*, "new Nyāya." Gaṅgeśa's sole work was the seminal *Tattvacintāmaṇi*, which takes up the four ways of gaining knowledge in a fresh way and employs a style of explanation that involves a host of new technical terms to indicate the various relations among the things contained in the Nyāya categories. These new relations make it possible for Navyanyāya to develop an "artificial" or "ideal" language in a way that resembles the methods of the logical positivists in modern analytic philosophy in the West. It also, not surprisingly, makes reading Navyanyāya texts especially difficult for the reader uninitiated into the technical terminology.

Gaṅgeśa's text is the basis for a flowering of hundreds of commentaries composed over the following centuries down to the present. The best known of these is the *Dīdhiti* by Raghunātha Śiromaṇi (1510), the subject, in turn, of myriads of commentaries, the most influential of which are those by Jagadīśa Tarkālaṁkāra (1620), Mathurānātha Tarkavagīśa (1650), and Gadādhara (1660). Also during this post-Gaṅgeśa period two works were composed that introduce the student to the terminology, categories, and logic of Nyāya-Vaiśeṣika: the widely studied introductory work by Annambhaṭṭa (1500) *(Tarkasaṁgraha)* and that by Viśvanātha Nyāyācārya Bhaṭṭācārya (1640) (*Bhāṣāpariccheda*, with its autocommentary *Siddhāntamuktāvalī*).

PŪRVAMĪMĀṀSĀ

The term *mīmāṁsā* connotes a method of textual interpretation, especially of the Vedas. Rules determining the proper way to read and interpret the passages of the Vedic corpus developed early. The earlier portions of the Vedas did not speak of liberation; the subject discussed was taken to be *dharma*, featuring prescriptions on sacrifice, how to live, and what actions to perform and not to perform.

These Mīmāṁsakas viewed the Vedas as without any author—not even the gods were the authors of scripture. The authority of the Vedas is based on their being beginningless and thus authorless and so not subject to the foils of any human or even divine creator. However, because the Upaniṣads, the later part of Vedic scripture, allude to liberation, later Mīmāṁsā philosophers beginning with Kumārila and Prabhākara recognized the possibility of attaining liberation.

While the Vedas themselves constitute the basic literature of Mīmāṁsā, a particular set of aphorisms—the *Mīmāṁsāsūtra*s ascribed to Jaimini (25 CE?)—is regularly cited as Mīmāṁsā's basic text, with the commentary (*Bhāṣya*) by Śabara (400) on it appealed to for explanations. These two works are largely devoted to matters that concern the proper interpretation of Vedic maxims about how to sacrifice and act in appropriate ways. But the literature of the Mīmāṁsā philosophical systems about to be discussed include interpretive and other works that, although largely concerned with Vedic interpretation, develop categorial frameworks that are comparable to those found in the other systems of Indian thought, and attempt to controvert the views of those other schools.

The Bhāṭṭa school of Mīmāṁsā looks to the interpretation of Kumārila (660) as found in that writer's commentary on Śabara's *Bhāṣya*, particularly in that portion of it titled *Ślokavārttika*, in which Kumārila makes a trenchant attack on other views known to him and provides reasons for preferring his own interpretation. The Bhāṭṭa literature develops through the works of Maṇḍana Miśra (690) (*Brahmaviveka, Vidhiviveka*), Pārthasārathi Miśra (1075) (*Nyāyaratnamālā, Śāstradīpikā, Nyāyaratnākara*), Āpadeva (1610) (*Mīmāṁsānyāyaprakāśa*), Laugākṣi Bhāskara (1660) (*Arthasaṁgraha*), and Kṛṣṇa Yajvan (1750) *(Mīmāṁsāparibhāṣā)*.

The Prābhākara school is named after Prabhākara (700), author of the commentary *Bṛhati* on Śabara's *Mīmāṁsābhāṣya*. Among later Prābhākara Mīmāṁsakas, far fewer in number to the Bhāṭṭas, one may note Śālikānātha Miśra (825), author of the *Prakaraṇapañcikā.*

There is also said to be a third Mīmāṁsaka school known as the "Miśras" after Murāri Miśra, reputed author of several works most of that are now lost.

ADVAITA VEDĀNTA

The term *vedānta* literally means "the end or final portions of the Vedas." Those final portions are the Upaniṣads. The various systems that are called "Vedānta" take at least the older Upaniṣads as authoritative, bolstered by the *Bhagavadgītā* and the *Brahmasūtra*s. Vedānta commentaries and independent works expressing the views of these schools claim to represent the cor-

rect interpretations of these scriptural materials. But the philosophical positions they take vary widely.

The best-known Vedānta system is Advaita, "nondualism." It takes the position that there is only one real entity: the Brahman identified in the Upaniṣads as the true Self (*ātman*). Perhaps the oldest completely extant text expounding Advaita views is a commentary on the Māṇḍūkya Upaniṣad by Gauḍapāda (600). But the writer acknowledged as the authoritative source of Advaita is Śaṁkarācārya (710), who wrote a number of works and who is assumed to be the author of many others of later origin. The most important of Śaṁkara's works is his commentary (*Bhāṣya*) on Bāḍarāyana's (50 CE) *Brahmasūtra*s. Śaṁkara is also probably the author of commentaries on several of the oldest Upaniṣads and on the *Bhagavadgītā*, and of at least part of an independent work titled *Upadeśasāhasrī*.

Śaṁkara refers to at least two other Vedānta systems. One is regularly termed "Bhedābhedavāda." In contrast to Śaṁkara's austere nondualistic position that there is only one real entity, Brahman, and that all difference and thus all plurality is illusory, the Bhedābheda view is that Brahman is both different (*bhedā*) and nondifferent (*abheda*) from the world. Śaṁkara also clearly has in mind for refutation a contemporary named Maṇḍana Miśra, who, in a work entitled *Brahmasiddhi*, defends an interpretation of Advaita according to which one does not (contrary to Śaṁkara's interpretation) achieve complete liberation prior to death. Maṇḍana holds that an enlightened person must still continue to practice meditation after achieving emancipation.

Śaṁkara's own position is that the Upaniṣadic texts are of one or the other of two types, in effect comprising two distinct portions of scripture referred to as the *karmakāṇḍa* and the *jñānakāṇḍa*. The *karmakāṇḍa*, as its name implies, consists of those portions of scripture that are governed by injunctions about how one should act. Because it prescribes actions, proper attention to it should lead one to perform appropriate kinds of action, as the Pūrvamīmāṁsakas correctly suppose. The viewpoint required by one who appeals to the *karmakāṇḍa* for advice is a view that assumes differences (*bheda*) between things. One could hardly act if one did not assume differences—between what is and what should be, between what is done and what ought to be done, between action and agent, between you and me.

In contrast, Śaṁkara claims that the other part of the Upaniṣadic texts, the *jñānakāṇḍa*, deals not with what is to be done but with what one should know. Instead of injunctions to act, the contents of this part refer solely to what is actually the case. Instead of enjoining us to act, this part provides us with knowledge; instead of dealing with differences among the many things and beings of the world, including ourselves, the *jñānakāṇḍa* speaks merely to the one Reality in which no distinctions or differences can ACTUALLY abide. That Reality is called Brahman, and the "great sentences" (*mahāvākya*) of the Upaniṣads—sentences such as "that art thou" (*tattvamasi*)—can provide us with enlightenment concerning the ultimate unity of Brahman and one's true Self if and when we are ready to appreciate it.

Once realization has dawned, one is completely liberated from bondage to actions, for that bondage requires recognition of differences and the liberated person no longer recognizes any differences as real. True, one remains alive and appears to act because of the *prārabdhakarman* that constituted the rationale for his present life, but there will be no future lives for such a one, no rebirth. The contrary view of the Bhedābhedavādins and of Maṇḍana Miśra, that one must still meditate even after liberation, is claimed by Śaṁkara to be incorrect, for meditation is an act, and the liberated self is incapable of performing any action because that would require recognition of that reality of differences among things and people that, in his liberated state, he no longer recognizes.

Important Advaita treatises are ascribed to Śaṁkara's pupils Padmapāda (740) (*Pañcapādikā*) and Sureśvara (740) (*Naiṣkarmyasiddhi* and commentaries on at least two of Śaṁkara's commentaries on the Upaniṣads). The standard account of post-Śaṁkara Advaita, which has been subjected to serious dispute, distinguishes two or three schools of Advaita, one stemming from Padmapāda, though named after a commentary on that work, the *Vivaraṇa* by Prakāśātman (975); another from Vācaspati Miśra's (940) commentary *Bhāmatī* on Śaṁkara's *Brahmasūtrabhāṣya*; and (sometimes) a third, unnamed, school stemming from Sureśvara. The vast majority of interpretations of Advaita defend the Vivaraṇa position. Among the writers who represent this school are Jñānaghana (975), author of *Tattvaśuddhi*, Vimuktātman (975) (*Iṣṭasiddhi*), Sarvajñātman (1027) (*Saṁkṣepaśārīraka*), Citsukha (1200) (*Citsukhī*), Vidyāraṇya or Mādhava (1350) (*Pañcadaśī*), Sadānanda (1500) (*Vedāntasāra*), Prakāśānanda (1505) (*Vedāntasiddhāntamuktāvalī*), Madhusūdana Sarasvatī (1570) (*Advaitasiddhi*), and Dharmarājadhvarīndra (1615) *(Vedāntaparibhāṣā*). The *Khaṇḍanakhaṇḍakhādya* of Śrīharṣa (1180) is a polemical treatise attacking the Nyāya-Vaiśeṣikas on behalf of Advaita. The far less numerous Bhāmatī-school authors include Amalānanda (1255)

(*Vedāntakalpataru*) and Akhaṇḍānanda Sarasvatī (1670) (*Ṛjuprakāśikā*).

VIŚIṢṬĀDVAITA VEDĀNTA

In contrast to Advaita's monistic interpretation of the relation of Brahman to the world, which says that only Brahman is real and that the world is illusory, the system of Viśiṣṭādvaita views Brahman and the world as real and takes Brahman and the world to be the same thing. The earliest author of this persuasion whose works are available is Yāmuna (1010), the author of *Āgamaprāmāṇya*, *Bhagavadgītārthasaṁgraha* and *Siddhitraya*. The real founder of Viśiṣṭādvaita is, however, Rāmānuja (1120), the author of a *Bhagavadgītābhāṣya*, the *Śrībhāṣya* on the *Brahmasūtra*s, and several independent works (*Vedāntadīpa*, *Vedāntasāra*,and *Vedārthasaṁgraha*). The most important writers in the ensuing centuries include Lokācārya Pillai (1300), who wrote in Tamil; Vedānta Deśika (1330), who is believed to have written more than thirty Sanskrit works; and Śrīnivāsa (1625), whose *Yatīndramatadīpikā* provides a useful summary of the major tenets of the school.

Viśiṣṭādvaita, frequently rendered as "qualified nondualism," is a kind of pantheism in which the unity of Brahman is gained not (as in Advaita) by denying Brahman's relation to anything else but rather by construing Brahman's unity as an "organic" unity of everything. Rāmānuja postulates three distinct real types of entities—selves, matter, and God—and construes "Brahman" as referring to the organic whole that they constitute. That is, Brahman for Viśiṣṭādvaita is *saguṇa*—it really has qualities—whereas for Śaṁkara it is *nirguṇa*, without any qualities whatsoever.

DVAITA PHILOSOPHY

Taking a straightforwardly pluralistic attitude toward the relations between Brahman, God, humans, and the things in the world was Madhva or Ānandatīrtha (1250), who wrote commentaries on the usual body of Vedānta texts (the *Brahmasūtra*s, *Bhagavadgītā*, and Upaniṣads) along with a number (usually reckoned as ten) of independent treatises. His system is known as Dvaita or "dualistic," or, in this case, more aptly, "pluralistic." Jayatīrtha (1370) comments on most of Madhva's works. The *Nyāyāmṛta* of Vyāsatīrtha (or Vyāsarāya) is a polemical treatise in which the author uses Navyanyāya methods to defend Dvaita and to criticize the views of others. Where Rāmānuja divided reality into three aspects of the organic unity of Brahman, Madhva's position features a basic distinction between Brahman, the Lord, who is deemed independently Real (*svatantra*), and the other Reals (including selves and things in the world), which are classed as real but dependent (*paratantra*) on God—dependent not for their being, which is beginningless, but for their being allowed to live, act, and gain release.

OTHER VEDĀNTIC SYSTEMS

The Bhedābhedavāda position, mentioned previously, which Śaṁkara criticizes, was apparently propounded in works, now lost, by writers who preceded Śaṁkara, notably Bhartṛprapañca (550). Later, Bhāskara (750) wrote a *Brahmasūtrabhāṣya* defending the position. A similar standpoint, called "Dvaitādvaita," is defended by Nimbārka (1250) in his commentary on the *Brahmasūtra*s, *Vedāntaparijātasaurabha*, and in other works.

A position known as Acintyabhedābheda has achieved some importance outside of India through the influence of the Hare Krishna movement. The founder of this school is Caitanya (1520), who wrote no works but whose views are capably expounded and defended by Sanātana Gosvāmin, Rūpa Gosvāmin, and Jīva Gosvāmin, all of whom seem to have lived in the sixteenth century. The writings of these and other exponents of this religious philosophy construe liberation as devotion to God culminating, according to A. C. Bhaktivedanta Swami (*The Nectar of Devotion*, p. 38), in "five liberated stages, which are 1) to be one with Me, 2) to achieve residence on My planet, 3) to have My opulences, 4) to possess bodily features similar to Me, and 5) to gain personal association with Me."

The Śuddhādvaita Vedānta school's literature starts with the numerous works of Vallabha (1525), which include an *Anubhāṣya* on the *Brahmasūtra*s and a commentary on the *Bhagavata Purāṇa*, along with more than thirty independent treatises. A series of commentators during the sixteenth and seventeenth centuries is capped by the prolific Puruṣottama Pītamabara Sarasvatī, who is credited with more than eighty works. The term *Śuddhādvaita*, "pure monism," is based on this school's theory that Brahman (called the highest Self [*puruṣottama*] or Śrī Kṛṣṇa) by nature emanates existence, intelligence, and joy like sparks from a fire. A spark where the joy portion becomes concealed by the existence portion constitutes an individual self. When, through devotion generated by God's grace (*puṣṭibhakti*), the lost joy is regained, one rejects liberation and chooses eternal service of Lord Kṛṣṇa, enjoying the boundless joy experienced in eternal play.

ŚAIVA SYSTEMS

Of the philosophical systems devoted to Śiva, perhaps the oldest literature is that of Kashmir Śaivism. Vasugupta (840) is the reputed author of *Śivasūtra*s and the *Spandakārikās*, and Somānanda (850) of *Śivadṛṣṭi*. Utpala's (925) *Īśvarapratyabhijñā* is another important work. Abhinavagupta (1014) contributes copiously and significantly to this tradition in a variety of works of which the *Tantrāloka* is the most enterprising. Kṣemarāja (1040) is the author of a popular exposition of the system in his *Pratyabhijñāhṛdaya*.

Kashmir Śaiva philosophy is also called the Pratyabhijñā system because it considers the ultimate aim to be self-realization (*pratyabhijñā*). Ontologically it teaches that the world is made to appear by Śiva's power (*śakti*) of consciousness. An individual self (*paśu*) is a center of consciousness, which is different from Śiva's in that a self's consciousness is limited by impurities (*mala*). Through following the spiritual path, one removes the impurities and gains realization. Three basic kinds of means (*upāya*) constituting the spiritual path are distinguished: external (*āṇava*), consisting of yogic postures, control of breathing, and so on; mental (*śākta*), voluntary meditation involving conceptual construction (*vikalpa*); and spontaneously viewing the entire world including oneself as a reflection of Śiva, a view that is effortless or construction-free (*nirvikalpaka*). In this final state the mind is dissolved into consciousness because of the removal of obstructions.

Śaiva Siddhānta has its home in the South of India. Its literature was written entirely in Tamil. Among the important authors and works in this tradition are those by Sadyojyoti (890), Meykanta Tevar (1221) (*Śivajñānabodha*), Aruṇanti Śivacariyār (1253) (*Śivajñānasiddhiyār*), Umāpati Śivacariyār (1310) (*Tiruvaruṭpayan*), and Śivāgra Yogi (1600) (*Śaivaparibhāṣā*).

Other Śaiva systems that have a literature in Sanskrit include those of Vīraśaiva and Śivādvaita. The most widely known work of the latter system is Śrīkaṇṭha's (1400) *Śrīkāra-bhāṣya* on the *Brahmasūtra*s, along with Appayya Dīkṣita's (1585) *Śivakaraṇidīpikā*. Appayya Dīkṣita has also written two independent treatises on this system.

GRAMMARIAN PHILOSOPHY

One of the remarkable achievements of early Indian science was in linguistics or grammar. Pāṇini (perhaps fifth century BCE) anticipates the linguistic analysts of the twentieth century in having managed, in his *Aṣṭādhyāyī*, to have shown how to generate the entire Sanskrit language from a series of rules, including rules about how to apply the rules. Grammar (*vyākaraṇa*), which formed a distinct science with a sizable literature, also caught the attention of philosophers of several of the systems discussed previously. But Grammarian philosophy is largely a product of Bhartṛhari (450) who, in his work *Vākyapadīya*,wedded grammar, epistemology, and ontology into a full-fledged philosophical system.

Bhartṛhari's innovations pertain to the proper account of language, of epistemology and ontology. The characteristic and unique idea of Bhartṛhari's view of language is the notion of the *sphoṭa*, which is conceived to be a unitary and permanent entity underlying the significance found in syllables, words, and sentences. A cognition is likewise construed as a unitary mental event that appears as having distinctions of subject and object, of time and space. But ultimately such distinctions are transcended: only language itself existse—even physical objects are no longer discriminated.

Maṇḍana Miśra (see above) has contributed an important work on Grammarian philosophy, the *Sphoṭasiddhi*. Several of the best grammatical works contain considerable philosophical material: for example, Kauṇḍa Bhaṭṭa's (1650) *Vaiyākaraṇa-bhūṣaṇa* and several works by Nāgeśa (or Nāgojī) Bhaṭṭa (1700).

Jain Philosophy. Jainism and Buddhism (along with Cārvāka) are sometimes referred to as "heterodox" schools in that they, unlike the systems listed above, do not view the Vedas as authoritative. (In fact, among the foregoing schools Nyāya-Vaiśeṣika and Sāṁkhya-Yoga do not cite the Vedas as authority either, although some texts of those schools very occasionally appeal to Vedic passages when suitably supportive.)

The Jain *sūtra*s represent early (the precise dating is unclear) representations of a variety of concerns only occasionally philosophical. They are written in Prakrit, a Sanskritic vernacular. Some of the Jain interpretive literature is also in Prakrit, though for the most part it exists in Sanskrit.

Perhaps the most notable feature of Jain thought is its refusal to accept a single account of reality. This is reflected in the Jain theory of *anekāntavāda*, that there are several equally true aspects of any given thing or topic. This is spelled out in their theories of *syādvāda* and *saptabhaṅgī*, which emphasize that everything we cognize can be viewed in several different ways each of which is acceptable given its particular orientation. The earliest writers of texts on Jain philosophy appear to have been

Kundakunda (200?) (*Pañcāstikāyasāra, Pravacanasāra, Samayasāra*) and Umāsvāti (200?) (*Tattvārthasūtra*). The interpretive literature on the latter work is voluminous, starting with Pujyapāda (500) (*Sarvārthasiddhi*), who also wrote the *Samādhitantra*. Bhadrabāhu (550) is the author of authoritative commentaries on the Jain *sūtra*s. Bhaṭṭa Akalaṅka (680) contributed a number of important works to the extensive literature on logic, along with Vidyānanda (850). A Siddhasena Divākara, whose date is not entirely certain, is the author of a small work, *Nyāyāvatāra*, which is perhaps the usual beginning point for students beginning the study of Jain philosophy. Other important contributors to the Jain philosophical literature are Haribhadra Sūri (750) (at least twenty-five works), Maṇikyanandin (950) (*Parīkṣāmukha*), Nemicandra Siddhāntacakravartin (1080) (*Gomatasāra, Dravyaviveka*), Vādideva or Devasūri (1143) (*Pramāṇanayatattvāloka*), Hemacandra (1150) (*Pramāṇamīmāṁsā, Anuyogavyavacchedadvātriṁśikā*) and Yaśovijaya (1680), to whom over thirty works are attributed.

BUDDHIST PHILOSOPHY

Although the number of Buddhist philosophical schools is still unsettled six distinct positions will here be distinguished. The standard division into Hīnayāna and Mahāyāna is not really relevant to these distinctions, although Theravāda, Sarvāstivāda and Sautrāntika are usually classed as Abhidharma (an expression we prefer instead of the pejorative "Hīnayāna"), the other three discussed here as Mahāyāna. Since Buddhism is discussed extensively elsewhere the most important authors and Indian texts on Buddhist philosophy are here merely listed.

Both the Theravāda and Sarvāstivāda acknowledge a list of seven Abhidharma texts, but a different seven for each of the two. The chronology of this literature is unclear, although we may conjecture that the seven texts in both lists existed prior to the beginning of the Christian era. The seven Sarvāstivāda or "northern" Abhidharma texts are *Dharmaskandha, Saṅgītiparyāya, Prajñaptibhāṣya, Dhātukāya, Vijñānakāya, Prakaraṇapāda* and *Jñānaprasthāna*, ascribed to various authors. The Sarvāstivāda reached its maturity, however, in the (*Mahā)Vibhāṣā*, compiled by committee in the first half of the second century CE, in which are recorded the opinions of many Abhidharma teachers concerning the proper understanding of Sarvāstivāda tenets. The names of a large number of Buddhist schools are also given in this and in various later works; the relations of some of these schools to the ones discussed in our brief overview are still being determined. In any case, the term "Vaibhāṣika" is used synonymously with "Sarvāstivāda" in recognition of the importance of the *Vibhāṣā*.

The fourth-century (?) author Vasubandhu (the dating is still controversial) is the author of the best-known exposition of Vaibhāṣika theses, entitled *Abhidharmakoṣa*, together with his own commentary (*Bhāṣya*) in which Vasubandhu criticizes those very Vaibhāṣika views from the standpoint of the interpretation labelled "Sautrāntika" (derived from "*sūtra*"). The Sautrāntikas urged going back to the Buddha's own words as found in the Buddhist canon. Vasubandhu provides us with a detailed acount of the Sautrāntika's opinions, and perhaps goes on to criticize both Vaibhāṣika and Sautrāntika from a Yogācāra (see below) perspective. This defection from Vaibhāṣika tenets produced a violent reaction by Vasubandhu's contemporary Saṁghabhadra in his *Nyāyānusāra*.

Turning to the Theravāda, their list of seven Abhidharma works comprises the *Dhammasaṅgaṇī, Vibhaṅga, Dhātukathā, Puggalapaññatti, Kathāvatthu, Yamaka* and *Paṭṭhāna*. The place of the *Kathāvatthu* is somewhat similar to that of the *Vibhāṣā* in the Sarvāstivāda tradition, in that it records many different opinions about a variety of Buddhist concerns, some doctrinal, others practical, which appear to have caught the notice of the author, who wrote in about the third century BCE A noncanonical work that appears to date from the pre-Christian era as well is the popular *Milindapañha*, a literary treatise presenting itself as recording a discussion between King Milinda (Menander?) and the monk Nāgasena.

Buddhism appears to have moved to what is now Śrī Laṅka several centuries after the Buddha, although the earliest literary remnants of Buddhism there are now lost. Around 425 two Indians, BuddhagHosa and Buddhadatta, appear to have visited Śrī Laṅka—there is a tradition about their meeting on the seas between island and mainland—and each wrote works in Pāli (a vernacular of Sanskrit) recording the philosophical position of the Theravādins. Buddhadatta's work, titled *Abhidhammāvatāra*, is not studied frequently nowadays, but the works of Buddhaghosa (who remained in Lanka), especially the mammoth *Visuddhimagga*, are seminal to the philosophical theses of the Buddhism that has flourished since, initially in Ceylon and eventually throughout Southeast Asia. Buddhaghosa is said to have also written commentaries on all seven of the works of the Theravāda or "Pāli canon" Abhidharma.

Beginning at least by the first century CE a type of literary production began to appear, the importance of

which for understanding the subsequent development of Buddhism both in India as well as throughout Asia is very apparent from the vast interpretive literature that has grown up around it. The texts in question are often referred to as "Mahāyāna *sūtra*s." While the exact connection between these works and the coming of Mahāyāna Buddhism (not to mention the use of the self-laudatory word "Mahāyāna") is not well understood, some of these works are among those most familiar and dear to the heart of millions of Buddhists throughout the world under names such as the "Lotus Sutra," "Heart Sutra," "Diamond Sutra," etc.

It is still being argued by scholars what is the correct account of the rise of what has come to be called "Mahāyāna." Some accounts connect it with Nāgārjuna (150 CE), author of a number of philosophical works such as the *(Mūla)Madhyamaka-kārikās* and *Vigrahavyāvartanī*. The connection between that Nāgārjuna and Mahāyāna is not at all clear, however. It is other Nāgārjunas, the apparent authors of works of probably later vintage, that show affinities with what are taken to be particularly Mahāyānic topics and theories.

In any case, the Madhyamaka tradition persists in India, and eventually in Tibet and East Asia, in a literature of which we mention here only the Indian portion. The works attributed to the second-century Nāgārjuna show their author to be a masterly critic of all philosophical positions, so much so as to have earned him the charge of being merely a skeptic or—worse still—a nihilist. A vast secondary as well as a lively interpretive literature concerns itself with the proper interpretation of his position. Although it is still somewhat controversial to say so, the major lines of interpretation seem to be two, terms for which have been borrowed from subsequent Tibetan commentators. One line insists that when Nāgārjuna says that everything is "empty" or "void" (*śūnya*) he means what he says, i.e., that (as he himself says) he has no thesis whatsoever, that he uses language solely to refute those who do take positions. This interpretation is known as "Prāsaṅgika," and its earliest Indian protagonists are Buddhapālita (480) and Candrakīrti (600). An alternative, "Svātantrika" line of thinking is particularly defended by Bhavya (or Bhāvaviveka) (550), who wishes to allow for the positive use of inferential arguments to establish the interdependence of all things, termed "emptiness" in Madhyamaka. All these writers composed commentaries on Nāgārjuna's *Mādhyamikasūtra*s.

Madhyamaka is one of the philosophical positions regularly identified as Mahāyāna. The other is known as Yogācāra or Vijñānavāda Buddhism. Whereas Madhyamaka's position (if it has one) is that nothing is real (not even emptiness), the Yogācāras exempt consciousness itself from this denial. What exist are many streams of consciousness (or perhaps only one stream); what appears as an independent world is merely a construction based on our karma. The earliest proponents of this position are regularly held to be the brothers Asaṅga and Vasubandhu (fourth century), although scholars are fairly certain that these two did not found the system, that it was already in place and expounded in works slightly earlier, e.g., in the *Saṁdhinirmocanasūtra* and *Laṅkāvatārasūtra* (both about 325 CE). Furthermore, there is some reason to believe that at least parts of the works ascribed to Asaṅga are of an earlier vintage and that the Vasubandhu who wrote the influential Yogācāra works titled *Triṁśikā* and *Viṁśkā* as well as the *Karmasiddhiprakaraṇa* was not the same person as the one who wrote the *Abhidharmakośa*. In any case, a healthy literature soon grew up around these works, eventually leading to attempts to find *rapprochement* between Yogācāra and Madhyamaka, as in Śāntarakṣita (750) (*Tattvasaṁgraha*) and his commentator Kamalaśīla (770), and perhaps to the development of the Buddhist Logic tradition (see below) which is itself viewed by East Asian interpreters as merely a branch of Yogācāra. Eventually Yogācāra becomes very important in Tibet, expounded there by the influential monk Atīśa (or Dīpaṁkara Śrījñāna, 1035) and others.

Attention is given to the theory of inference or logic by Vasubandhu and others prior to Dignāga (510) (*Pramāṇasa-muccaya*), but it has become customary (influenced by Th. Stcherbatsky's book *Buddhist Logic*, available in a popular publication in English early in the twentieth century) to refer to Dignāga's philosophical position as that of the "Buddhist Logicians." The works by Dignāga and others that constitute the literature of this school concentrate on the methodology of inferential reasoning, but also speculate on many of the same metaphysical and epistemological questions that other Buddhist systems were addressing, for example, whether consciousness alone is real or whether nothing at all is real, and whether it is even possible to speak or think truly. Dignāga is ambiguous on those questions, but nevertheless he develops an identifiable position stemming from an epistemic distinction between "perception"—which grasps what Westerners might call "sense-data," which are taken to be actual entities—and inference which, since it deals with universal properties, is deemed unreal since it can only concern conceptual constructions. Dignāga's approach is taken up and critically clarified with zeal in subsequent periods, notably by Dharmakīrti (610), Dhar-

mottara (770), Jitāri (990), Jñānaśrīmitra (1015), Ratnakīrti (1070), and eventually elegantly summarized by Mokṣākara Gupta (1100) in his *Tarkabhāṣā.*

See also Atomic Theory in Indian Philosophy; Brahman; Buddhism; Causation in Indian Philosophy; Karma; Knowledge in Indian Philosophy; Logic, History of: Logic and Inference in Indian Philosophy; Meditation in Indian Philosophy; Mind and Mental States in Buddhist Philosophy; Negation in Indian Philosophy; Philosophy of Language in India; Self in Indian Philosophy; Truth and Falsity in Indian Philosophy; Universal Properties in Indian Philosophical Traditions.

Bibliography

Bhaktivedanta Svami, A. C. *The Nectar of Devotion*. Boston, MA: ISKCON Press, 1970.

Bhatt, Govardhan P. *Epistemology of the Bhāṭṭa School of Pūrva Mīmāṁsā.* Chowkhamba Sanskrit Studies Vol. XVII. Varanasi: Chowkhamba Sanskrit Series Office, 1962.

Chari, S. M. Srinivasa. *Fundamentals of Viśiṣṭādvaita Vedānta. A Study Based on Vedānta Deśika's Tattva-muktā-kalāpa*. Delhi: Motilal Banarsidass, 1987.

Coward, Harold G., and K. Kunjunni Raja, eds. *The Philosophy of the Grammarians*. Vol. V of *Encyclopedia of Indian Philosophies*. Princeton, NJ: Princeton University Press, 1990; Delhi: Motilal Banarsidass, 1990.

Dasgupta, Surendranath. *A History of Indian Philosophy*. Cambridge, U.K.: Cambridge University Press. 5 vols., 1922, 1932, 1951, 1957, 1969, 1973; Delhi: Motilal Banarsidass 1975.

Franco, Eli. *Perception, Knowledge and Disbelief. A Study of Jayarāśi's Scepticism*. Alt- und Neu-Indische Studien herausgegeben vom Seminar für Kultur und Geschichte Indiens an der Universität Hamburg 35. Franz Steiner Verlag Wiesbaden Gmbh, Suttgart 1987.

Ganeri, Jonardon. *Philosophy in Classical India. The Proper Work of Reason*. London: Routledge, 2001.

Jaini, Padmanabh S. *The Jaina Path of Purification*. Berkeley: University of California Press, 1979.

Jha, Ganganatha. *The Prābhākara School of Pūrva Mīmāṁsā.* Allahabad 1911; Delhi: Motilal Banarsidass, 1978.

Larson, Gerald J., and Ramakrishna Bhattacharya, eds. *Sāṁkhya. A Dualist Tradition in Indian Philosophy*. Vol. IV of *Encyclopedia of Indian Philosophies*. Princeton, NJ: Princeton University Press, 1987; Delhi, Motilal Banarsidass, 1987.

Matilal, Bimal Krishna. *The Character of Logic in India*, edited by Jonardon Ganeri and Heeraman Tiwari. Albany, N.Y.: SUNY Press 1998.

Matilal, Bimal Krishna. *Perception. An Essay on Classical Indian Theories of Knowledge*. Oxford: Clarendon Press, 1986.

Mishra, Kamalakar. *Kashmir Śaivism. The Central Philosophy of Tantrism*. Delhi: Sri Satguru Publications, 1999.

Mohanty, Jitendra Nath. *Reason and Tradition in Indian Thought. An Essay on the Nature of Indian Philosophical Thinking*. Oxford: Clarendon Press, 1992.

Mohanty, J. N. *Classical Indian Philosophy*. Lanham, MD: Rowman and Littlefield Publishers, Inc. 2000.

Potter, Karl H., ed. *Advaita Vedānta up to Śaṁkara and His Pupils*. Vol. III of *Encylopedia of Indian Philosophies*. Princeton, NJ: Princeton University Press, 1981; Delhi, Motilal Banarsidass, 1981.

Potter, Karl H. *Bibliography*. Vol. I of the *Encyclopedia of Indian Philosophy*. 1st ed., 1970; 2nd ed., 1983. Most recently published in two volumes as the 3rd ed., Delhi: Motilal Banarsidass, revised 1995. Updated semiannually, it is available on the Internet at: http://faculty.washington.edu/kpotter.

Potter, Karl H., ed. *Buddhist Philosophy from 100 to 350 A.D.* Vol. VIII of *Encyclopedia of Indian Philosophies*. Delhi: Motilal Banarsidass, 1999.

Potter, Karl H., ed. *Buddhist Philosophy from 350 to 600 A.D.* Vol. IX of *Encyclopedia of Indian Philosophies*. Delhi: Motilal Banarsidass, 2003.

Potter, Karl H., ed. *Indian Metaphysics and Epistemology. The Tradition of Nyāya-Vaiśeṣika up to Gaṅgeśa*. Vol. II of *Encyclopedia of Indian Philosophy*. Princeton, NJ: Princeton University Press, 1977; Delhi: Motilal Banarsidass, 1977, 1995.

Potter, Karl H. *Presuppositions of India's Philosophies*. Englewood Cliffs, NJ: Prentice-Hall, 1963; Westport, CT: Greenwood Press, 1963; Delhi: Motilal Banarsidass, 1991, 1999, 2002.

Potter, Karl H., Robert E. Buswell Jr., Padmanabh S. Jaini, and Noble Ross Reat, eds. *Abhidharma Buddhism to 150 A.D.* Vol. VII of *Encyclopedia of Indian Philosophies*. Delhi: Motilal Banarsidass, 1996, 1998.

Potter, Karl H., and Sibajiban Bhattacharyya, eds. *Indian Philosophical Analysis. Nyāya-Vaiśeṣika from Gaṅgeśa to Raghunātha Śiromaṇi*. Vol. VI of *Encyclopedia of Indian Philosophies*. Princeton, NJ: Princeton University Press, 1993; Delhi, Motilal Banarsidass, 1993.

Radhakrishnan, Sarvepalli, and Charles A. Moore, eds. *A Sourcebook in Indian Philosophy*. Princeton, NJ: Princeton University Press, 1957, 1973, etc.

Sharma, B. N. K. *Philosophy of Śrī Madhvācārya*. Bombay: Bharatiya Vidya Bhavan, 1962.

Sivaraman, K. *Śaivism in Philosophical Perspective. A Study of the Formative Concepts, Problems and Methods of Śaiva Siddhānta*. Delhi: Motilal Banarsidass, 1973.

Stcherbatsky, Th. *Buddhist Logic*. Two Vols. Bibliotheca Buddhica 26. Leningrad 1930; The Hague 1958; New York 1962; Osnabruck 1970; Tokyo 1977; Delhi 1972.

Williams, Paul, with Anthony Tribe. *Buddhist Thought. A Complete Introduction to the Indian Tradition*. London: Routledge, 2000.

Karl Potter (2005)

INDIRECT REALISM

See *Realism*

INDISCERNIBLES, IDENTITY OF

See *Identity*

INDIVIDUALISM

See *Holism and Individualism in History and Social Science*

INDUCTION

The name "induction," derived from the Latin translation of Aristotle's *epagoge*, will be used here to cover all cases of nondemonstrative argument, in which the truth of the premises, while not entailing the truth of the conclusion, purports to be a good reason for belief in it. Such arguments may also be called "ampliative," as C. S. Peirce called them, because the conclusion may presuppose the existence of individuals whose existence is not presupposed by the premises.

Thus, the conclusion "All *A* are *B*" of an induction by simple enumeration may apply to *A*'s not already mentioned in the finite number of premises having the form "A_i is *B*." Similarly, in eduction (or arguments from particulars to particulars) the conclusion "Any *A* is *B*" is intended to apply to any *A* not yet observed as being a *B*.

It would be convenient to have some such term as *adduction* to refer to the sense of induction here adopted, which is broader than the classical conception of induction as generalization from particular instances. Most philosophical issues concerning induction in the classical sense arise in connection with the more general case of nondemonstrative argument.

In what follows it will be convenient to use Jean Nicod's expression "primary inductions" to refer to those nondemonstrative arguments "whose premises do not derive their certainty or probability from any induction." Problems of philosophical justification are most acute in connection with such primary inductions.

It may be added that "mathematical induction" is a misnomer because the useful types of reasoning so labeled are rigorously demonstrative. Given that the first integer has a certain property and also that if any integer *n* has that property then so does $n + 1$, the next, it follows demonstratively that all the integers have the property in question. Inductive arguments, as here conceived, do not constitute mathematical or logical proofs; by definition induction is not a species of deduction.

TYPES OF INDUCTIVE ARGUMENTS

In addition to the types of arguments already mentioned, the following are most frequently discussed:

(1) Elaborated induction (as it might be called) consists of more or less sophisticated variations of induction by simple enumeration, typically including supplementary information concerning the mode of selection of the individuals named in the premises and perhaps including reference to negative instances.

(2) Proportional induction is inference from the frequency of occurrence of some character in a sample to the frequency of occurrence of the same character in the parent population—that is, from "m_1/n_1 *A*'s selected by a stated procedure *P* are *B*" to "m_2/n_2 *A*'s are *B*." Here the ratio stated in the conclusion may be other than the one stated in the premise; it is often advantageous to locate the final ratio within a certain designated interval.

(3) Proportional eduction is argument from sample to sample. From the same premises as in proportional induction a conclusion is drawn concerning approximate frequency of occurrence in a further sample obtained by the same procedure or by another one.

(4) Proportional deduction (commonly called "statistical syllogism") is inference from "m/n *C*'s are *B*" (where m/n is greater than 1/2) and "*A* is a *C*" to "*A* is a *B*."

In all the above cases modern writers usually insist upon inserting some more or less precise indication of probability or likelihood, either within the conclusion itself or as an index of reliability attached to the mark of inference ("therefore," "hence," or the like). Careful attention to the probability or likelihood attributed to a given inductive conclusion is a distinct merit of modern treatments of the subject.

The foregoing list cannot claim to be exhaustive, nor are its items to be regarded as mutually irreducible. There is no general agreement concerning the basic forms of inductive argument, although many writers regard simple enumeration as in some sense the most fundamental.

HISTORY OF INDUCTIVE METHODS

Interest in the philosophy and methodology of induction was excited by the extraordinary successes of natural science, which tended to discredit the rationalistic conception of knowledge about matters of fact. The classical writers on the subject, from Francis Bacon on, have lamented the powerlessness of deduction to do more than render explicit the logical consequences of generalizations derived from some external source. If recourse to intellectual intuition or to self-evidence is repudiated as a source of factual knowledge, nothing better seems to remain than reliance upon the empiricist principle that all knowledge concerning matters of fact ultimately derives from experience. However, experience, whether conceived as sporadic and undirected observation or as the systematic search for specific answers extorted by planned experiment, seems to supply knowledge only of particular truths. Empiricists are therefore faced with the problem of accounting for the crucial step from knowledge of experiential particulars to reasoned acceptance of empirical generalizations sufficiently powerful to serve as the major premises of subsequent logical and mathematical deduction.

The aspiration of early writers was, characteristically, to demonstrate the conclusions of acceptable inductive arguments as true; not until the end of the nineteenth century did a more modest conception of inductive argument and scientific method, directed toward acquiring probability rather than certainty, begin to prevail.

PROBLEM OF INDUCTION

The celebrated problem of induction, which still lacks any generally accepted solution, includes under a single heading a variety of distinct, if related, problems. It is useful to distinguish the following:

(1) The general problem of justification: Why, if at all, is it reasonable to accept the conclusions of certain inductive arguments as true—or at least probably true? Why, if at all, is it reasonable to employ certain rules of inductive inference?

(2) The comparative problem: Why is one inductive conclusion preferable to another as better supported? Why is one rule of inductive inference preferable to another as more reliable or more deserving of rational trust?

(3) The analytical problem: What is it that renders some inductive arguments rationally acceptable? What are the criteria for deciding that one rule of inductive inference is superior to another?

These problems may be briefly labeled "justification," "differential appraisal," and "analysis." Many writers on induction have also occupied themselves with the task of codification, the formulation of a coherent, consistent, and comprehensive set of canons for the proper conduct of inductive inference. Important as it is, this task is not distinctively philosophical, except insofar as it requires in advance answers to the questions listed above.

In practice the three problems here distinguished cannot be pursued separately; a comprehensive general defense of inductive procedures involves specification, inter alia, of legitimate forms of inductive argument, and selection between alternative inductive rules or methods must rely, explicitly or not, upon determination of what, if anything, makes an inductive argument "sound." The *why* of inductive argument cannot profitably be isolated from the *how*.

It is characteristic of much recent investigation of the subject to concentrate on the last two of the problems listed, often in the hope of formulating precise canons of inductive inference (an inductive logic). These comparative and analytical versions of the problem of induction are thought worth pursuing even by writers who reject the general problem of justification as insoluble.

HUME'S VIEW OF CAUSATION. For better or worse, all modern discussion of the philosophy of induction takes off from David Hume's celebrated analysis of causation, whose connection with the philosophical problems of induction (a word that Hume never used) arises from his view that all reasoning concerning matters of fact is founded on the relation between cause and effect. Although Hume may be held to have given undue prominence to causation (his skeptical conclusions do, in fact, challenge every kind of nondemonstrative argument, whether or not grounded in causal imputation), it is easy to overlook and to be misled by the special form in which he conceived the problem of justification.

Hume, unlike such later writers as J. S. Mill, was not satisfied to analyze the notion of cause and effect into the notions of spatial contiguity, temporal succession, and joint occurrence; he fatefully added to these the criterion of "necessary connexion." That objects of certain kinds have been conjoined or associated in past experience might be no more than an extended coincidence. Something more is needed before one event can properly be recognized as the cause of the other; we must be able to pass from *post hoc* to *propter hoc*. In predicting a putative effect of a given event we can ensure contiguity and succession by choosing to look only for a spatiotemporally

proximate event, and memory (if that can be relied on) will furnish knowledge of constant conjunction in the past. Whether we are truly justified in predicting the occurrence of the putative effect will therefore turn entirely upon whether there is good reason to assert that it is necessarily connected with its neighbor. Hume, in effect, challenged his reader to find anything in the observation of a single case of supposed causal action (for instance, in the favorite example of a collision between two billiard balls) that answers to the required "necessary connexion" between two events. No observation, however attentive, will discover more than contiguity and an internal habit of expecting association. Nor will examination of a series of cases, all exactly alike, help at all: A sum of zeroes is still zero.

But what did Hume mean by "necessary connexion"? Although he did not tell us in so many words, his main proof that we can "never demonstrate the necessity of a cause" rests simply upon the conceivability, and hence the logical possibility, of an event's being bereft of its putative cause. He seems, therefore, to have implied that our notion of a cause and its effect requires the existence of the one to be entailed by the existence of the other. If so, it does not need much argument to show that we can have no impression (direct sensory experience) of such entailment. Hume concluded that necessity cannot reside in the external world but must arise, as an idea, from an internal impression of the mind, a "determination to carry our thoughts from one object to another."

Repeated observation of the association of events leads us to the habit of expecting the association to continue "by means of an operation of the soul … as unavoidable as to feel the passion of love, when we receive the benefits" (*Enquiry concerning the Human Understanding*, Sec. 5, Part 1). Our idea of necessary connection is nothing more than an internal response to the habit of expecting effects: "Upon the whole, necessity is something in the mind, not in objects." At this point skepticism is just around the corner; we are on the verge of such famous conclusions as that "all probable reasoning is nothing but a species of sensation" (*Treatise*, Book I, Part 3, Sec. 8).

The reference to habit or custom explains nothing, of course, and is at best only a concise reference to the truism, which according to Hume's view simply has to be accepted, that men do in fact expect events to be accompanied by effects. Without such habits of causal expectation men could hardly have survived—but this reflection, itself based on induction, cannot be a reason for belief in causation. For a philosopher so critical of such allegedly occult entities as power and energy, Hume was strangely carefree in his reliance upon habit or custom as a *vera causa*. In keeping with his own principles he ought to have turned as skeptical an eye on habit as on cause and ought to have concluded that our idea of habit is derived from nothing more than a habit of expecting that a man who acts in a certain way will continue to do so. But now the account looks circular. Have we any better reason to believe in the existence of habits—even if construed, in as reductionist a fashion as possible, as mere constant conjunctions—than we have to believe in causes? And would not everything that tended to show we have no sufficient basis in external experience for belief in the objective reality of causal connection also tend to show, by parity of reasoning, that we have no basis for believing in the existence of those habits that are invoked at least to explain, if not to justify, our ordinary causal beliefs?

It has seemed to nearly all of Hume's readers that his method must lead to a skepticism more sweeping than he himself was perhaps willing to recognize or to accept. If Hume had been correct about the origin of the idea of necessity, he would have been committed to a totally skeptical answer to the general problem of justification. Whether or not we can escape from the bondage of causal expectation, we are at any rate free to see that such a habit can provide no reason, in Hume's sense, for the belief in causal connection. And once we see this, wholesale skepticism concerning inductive inference seems inescapable.

Hume's skeptical conclusions cannot be dismissed on the ground that they originated in an oversimplified psychology of ideas and impressions, for his argument can, with little difficulty, be made independent of any psychological assumptions. Cause and effect are logically independent, not because repeated search fails to find any logical connection, as Hume's own account misleadingly suggests, but because it is a part of what we mean by cause and effect that the two shall be logically separable. It is tempting to say, then, that there is no reason why the separable consequent should follow its antecedent in any particular instance. We can very well imagine or conceive the cause's occurring without its usual consequent, and, in Hume's words, "nothing of which we can form a clear and distinct idea is absurd or impossible" (*A Treatise of Human Nature*, Book I, Part 1, Sec. 7).

NEO-HUMEAN ARGUMENTS. Even if Hume was wrong in including logical necessity in the idea of causal connection, a neo-Humean can correct his argument without weakening its skeptical force. It is reasonable to say that what distinguishes a causal connection from a

merely accidental association is that empirical rather than logical necessity obtains between the two events. This, in turn, may be rephrased by saying that the observed conjunction is a case of lawful and not merely accidental association. But then Hume's challenge to discover such lawfulness in experience remains as formidable as ever; no matter how many instances of joint occurrence we encounter, we will never observe more than the de facto association and will never have ultimate, noninductive grounds for believing in a de jure connection.

Thus, Hume's problem can be put into modern dress, without restriction to causal inference, as follows: An inductive inference from an observed association of attributes (A_n—B_n) can justify inference to another case (A_{n+1}—B_{n+1}) or inference to the corresponding generalization ("All *A* are *B*") only if the association is somehow known to be lawlike, not merely accidental. Yet how can this be known in primary inductions that do not themselves rest upon the assumed truth of other laws? Certainly not by immediate experience, nor a priori, nor, without begging the question, by appeal to induction.

The sharpest form of this version of the problem (called by its author the "new riddle of induction") is that of Nelson Goodman. Suppose all emeralds examined before a certain time *t* have been green; use the label "grue" for the property of being green up to the time *t* and being blue thereafter. Then all the evidence supports equally well the competing laws "All emeralds are green" and "All emeralds are grue." Here an instance of the comparative problem is raised in a particularly pointed and instructive way.

Goodman's challenge awaits an answer. Some writers have hoped to defend the received or standard modes of inductive argument by invoking criteria of relative simplicity. But apart from the yet unsolved problem of clarifying what simplicity is to mean in this connection, there seems no good reason why nature should obligingly make correct inference simple; often enough the best-confirmed law is less simple than others that would accord with the given evidence. Goodman's own suggestion to restrict defensible inductions to "entrenched" predicates (roughly speaking, those that have been frequently employed in previous inductive judgments) seems less than satisfying.

From the standpoint of the philosophy of induction the chief significance of Hume's memorable discussion (apart from its tonic effect in disturbing "dogmatic slumber") is that it brought into full daylight the problem of distinguishing between a merely accidental series of associations and the genuine laws that we seek by means of inductions.

DEDUCTIVE STANDARD OF JUSTIFICATION. A demand that induction be justified arises, of course, from some supposed deficiency or imperfection. If all were obviously well with inductive argument, there would be no point in asking for any defense or justification. It is therefore of the first importance to be clear about the alleged weakness or precariousness of induction and the corresponding standard of justification to which appeal is covertly made. We need to know what is supposed to be the trouble with induction, for only when the disease is understood will the search for a remedy have much prospect of success.

The root of the trouble is plain enough in the writings of a hundred writers who have trodden in Hume's footsteps. All have been haunted by the supposedly superior certainty of demonstrative reasoning. If valid deduction from premises known to be true transmits certainty to the conclusion, even the best induction will seem inferior by comparison. (John Locke said that induction from experience "may provide us convenience, not science"—*Essay concerning Human Understanding*, Book IV, Ch. 12, Sec. 10.) The nagging conviction that induction somehow falls short of the ideals of rationality perfectly exemplified in valid deductive argument has made the problem of induction needlessly intractable.

If Hume, for instance, did not require that induction be shown as somehow satisfying the criteria of valid deduction, an answer to his question about how "children and peasants" learn from experience would be easy. The method employed, as he himself stated, is that of arguing from similarity of causes to similarity of effects. However, such an answer would obviously not have satisfied him, because this method will not guarantee the truth of the conclusion drawn; that is, it is not the kind of method that would be acceptable as justifying a valid deduction. Hume would have liked an inductive conclusion to follow from (be entailed by) premises known to be true, for anything less would not have seemed genuinely reasonable. Having shown, in effect, that no reason of this kind can be produced for primary inductions, he was forced to regard the question of justification as demonstrably insoluble. This conclusion has the notable inconvenience of leaving the comparative problem also insoluble (while the analytical task vanishes for lack of an object).

Hume's conclusion must be granted if his is the only sense of "reason" in point. If we never have a reason for an inductive conclusion unless we know the conclusion to

follow strictly from premises known to be true, then we can have no reason for believing in primary inductive conclusions; it is as reasonable to expect that thistles will bear figs, or something equally absurd, as it is to expect anything else extending beyond past experience. (Whether we can in fact bring ourselves to believe anything so absurd is beside the point.) Only in recent times have serious efforts been made to escape from the spell of the deductive model, used by Hume and his innumerable followers, by inquiring whether there may not be other proper and relevant senses of "reasonable." It will be argued later that belief in induction is reasonable in principle and that belief in one kind of inductive conclusion is more reasonable than belief in another.

The lasting attraction of the deductive model is not hard to understand. The raison d'être of deductive argument seems enticingly plain: Valid deductions are truth-transmitting and truth-preserving—which, given an interest in obtaining novel truth, seems enough to show the point of deductive reasoning. (That this cannot be the whole story is obvious from the uses of deductive reasoning in exhibiting the consequences of propositions hypothetically entertained—not to mention *reductio* arguments and other uses.) By contrast the raison d'être of induction seems unclear and mysterious. It would be easy, although unsatisfying to the genuinely perplexed, to say that sound inductive arguments are "likelihood-transmitting," for likelihood is as unclear a concept as inductive correctness. Thus, it is natural to ask for and to expect a detailed answer to the question "Why should a reasonable man rely upon likelihood in default of truth?" Even if the power of sound induction to confer likelihood upon conclusions is regarded as sufficient to make inductive argument reasonable beyond further cavil, the question how such likelihood is conferred will remain. Attention thus shifts to the analytical task.

It may be added that an enduring source of disquiet concerning inductive argument is its disorderliness and formlessness by contrast with deductive argument. In deductive argument we flatter ourselves upon readily perceiving the underlying principles and their necessary connection with logical form. By contrast with such classic simplicity, and order the realm of inductive argument seems disconcertingly complex, confused, and debatable: An inductive argument accepted by one judge may be rejected, on good grounds, by another, equally competent judge; supposedly sound arguments from different sets of true premises may yield opposed conclusions; the very soundness of induction seems not to be clear-cut but to admit of gradations of relative strength and reliability. Given all this, it is not surprising that although many students have labored to introduce order into the field, others, abandoning any hope of so doing, have turned away from induction as a tissue of confusions.

TYPES OF SOLUTION

The answers given in the literature to Hume's problem can be briefly summarized as follows:

(1) Hume's challenge cannot be met; consequently, induction is indefensible and ought to be expunged from any reasoning purporting to be rational.

(2) In the light of Hume's criticisms, inductive arguments as normally presented need improvement, either (*a*) by adding further premises or (*b*) by changing the conclusions into statements of probability. In either case a conclusion's validity is expected to follow demonstratively from the premises, and inductive logic will be reconstructed as a branch of applied deductive logic.

(3) Although inductive argument cannot be justified as satisfying deductive standards of correctness, it may be proved that inductive policies (rather than rules or principles) are, in a novel sense to be explained later, reasonable. Induction can be vindicated if not validated.

(4) Hume's problem is generated by conceptual and linguistic confusions; it must therefore be dissolved, rather than solved, by exposing these confusions and their roots.

These approaches are not all mutually exclusive. Thus (3), the pragmatic approach, is usually combined with (1), repudiation of induction as an acceptable mode of reasoning. Apart from (4) all the approaches accept or make substantial concessions to Hume's major assumption—namely, that the only wholly acceptable mode of reasoning is deductive. This is true even of those who hold (3), the "practicalists," who might be supposed, at first glance, to be relaxing the criteria of rationality.

REJECTION OF INDUCTION. The rejection of induction as a proper mode of scientific reasoning is sometimes found in the guise of advocacy of the so-called hypothetico-deductive method. According to such a view, the essence of genuinely scientific reasoning about matters of fact is the framing of hypotheses not established by given empirical data but merely suggested by them. Inference enters only in the control of hypotheses by the verification of their observable consequences: Negative

instances strictly falsify a hypothesis, whereas positive instances permit its use, pending further experimental tests, as a plausible, if unproved, conjecture. Science, as well as all reasoning about matters of fact aspiring to the reliability of scientific method, needs only the kind of reasoning to be found in deductive logic and in mathematics. Some such position was already adumbrated in the writings of William Whewell. It has at least the merit of drawing attention to the role of hypotheses in scientific method, a welcome corrective to the excessive claims of early partisans of inductive logic.

The most influential, and possibly the most extreme, of contemporary writers following this line is Karl Popper, who often maintained that what is called induction is a myth, inasmuch as what passes under that title "is always invalid and therefore clearly not justifiable." In his own conception of scientific method such repudiation of induction is linked with the thesis that the purpose of scientific theorizing is falsification (demonstration of error) rather than verification or confirmation (provisional support of an approximation to the truth). Those who agree would rewrite putatively inductive inferences to make them appear explicitly as hypothetical explanations of given facts. (Thus, instead of inferring "All *A* are *B*" from premises of the form "A_n is *B*," the first statement is offered as a more or less plausible explanation of why all the A_n should have been found to be *B*.)

In spite of its enthusiastic advocacy, it is hard to see where this proposal accomplishes more than a superficial change in the form in which inductive arguments are written and a corresponding alteration in the metalanguage in which they are appraised. Any hypothetical explanation of given empirical data is intended to reach beyond them by having empirical consequences amenable to subsequent tests. If all explanations consonant with the known facts (always an infinite set) were treated as equally unjustified by the evidence, Hume's problem would certainly be set aside, but only at the cost of ignoring what provoked it—namely, the apparent existence of rationally acceptable nondemonstrative arguments. It can hardly be denied that there are nondemonstrative arguments lending reasonable support to their conclusions; otherwise it would be as reasonable to expect manna from heaven as rain from a cloud. Anti-inductivists have seldom been hardy enough to brand all inductive arguments as equally invalid, but as soon as they discriminate between alternative hypotheses as more or less corroborated, more or less in accord with available facts, they are faced, in a new terminology, with substantially the original problems of justification and differential appraisal.

INDUCTIVE SUPPORT FOR INDUCTION. To the layperson the most natural way of defending belief in induction is that it has worked in the past. Concealed in this reply, of course, is the assumption that what has already worked will continue to do so, an assumption that has seemed objectionably circular to nearly all philosophers of induction. A stubborn minority (including R. B. Braithwaite and Max Black), however, insists that the appearance of circularity arises only from overhasty application of criteria applicable to deduction. Even in the limiting case, where the rule governing the supporting argument from previous efficacy is the very rule that is to be defended, it can be plausibly argued that no formal circularity is present. Nor is there the more subtle circularity that would obtain if knowledge of the conclusion's truth were needed to justify use of the self-supporting argument. In spite of spirited objections, this line of reasoning has not yet, in the writer's opinion, been shown to be mistaken.

The point that inductive support of induction is not necessarily circular has some importance as illustrating the interesting self-applying and self-correcting features of inductive rules; in virtue of these features, scrutiny of the consequences of the adoption of such rules can, in favorable cases, be used to refine the proper scope of inductive rules and the appropriate judgments of their strength.

A more serious weakness of this kind of defense, if it deserves to be called that, is lack of clarity about what counts as success in using the rule, which is connected in turn with the insufficiently discussed question of the raison d'être of induction considered as an autonomous mode of reasoning.

But even if this controversial type of inductive support of inductive rules ultimately survives criticism, it will not dispose of the metaphysical problems of induction. Those satisfied with Hume's conception of the problem are at bottom objecting to any use of inductive concepts and of the language in which they are expressed unless there is deductive justification for such use. They will therefore reject any reliance upon induction by way of defense, however free from formal defect, as essentially irrelevant to the primary task of philosophical justification. It must be admitted that inductive support of induction, however congenial to the layman, does not go to the roots of the philosophical perplexity.

A PRIORI DEFENSES. A few twentieth-century writers (notably D. C. Williams and R. F. Harrod) maintained that certain inductive arguments, unimproved by the addition of supplementary premises or by modification of the form of the conclusion, can be proved to be valid. Williams argued, with surprising plausibility, that the probable truth of the conclusion of a statistical syllogism can be shown to be necessitated by the truth of the premises, solely by reference to accredited principles of the mathematical theory of chances. While admiring the ingenuity displayed in this approach, critics have generally agreed in finding it fallacious. That some modes of inductive argument are certified as sound or acceptable on broadly a priori (perhaps ultimately linguistic) grounds is, however, a contention of some versions of the linguistic approach.

DEDUCTIVE RECONSTRUCTION. The effort to provide justification for induction through a reconstruction of inductive arguments so as to make them deductively valid has chiefly taken two forms.

Search for supreme inductive principles. If a given nondemonstrative argument, say from the amalgamated premise *P* to a conclusion *K* (where *K*, for the present, is regarded as a categorical statement of fact containing no reference to probability), is looked at through deductive spectacles, it is bound to seem invalid and so to be regarded as at best an enthymeme, needing extra premises to become respectable. It is easy, of course, to render the original argument deductively valid by supplying the additional premise "If *P* then *K*" (this premise will be called *Q*). In order for induction to be defended in the classical way, however, the premises have to be true and known to be true. Since *P* was supposed not to entail *K*, the new premise, *Q*, will be a contingent statement of fact, knowledge of whose truth is presumably to be derived either by deduction from more general principles or by induction from empirical data. In either case, if the deductive standard of justification is to be respected, the process must continue until we obtain general factual principles, neither capable of further empirical support nor needing such support.

The line of thought is the following: Since *K* does not follow strictly from *P*, the fact that the truth of propositions resembling *P* in assignable ways is regularly associated with the truth of propositions resembling *K* is a contingent fact about the actual universe. Looked at in another way, if events occurred purely at random, it would be impossible to make successful inductions; conversely, if inductions of a certain sort do systematically produce true conclusions, there must be a contingent regularity in the universe that should be capable of expression in the form of supreme principles or postulates of induction. Only if such postulates are true can inductions be sound; they must therefore be the assumed but unexpressed premises of all sound inductive arguments.

Favored candidates for the role of such enabling postulates have been the principle that the future resembles the past (Hume), a general principle of causation to the effect that every event has a sufficient cause (Mill), a principle of spatiotemporal homogeneity, which makes locations and dates causally irrelevant (Mill again), and a principle of limited independent variety ensuring that the attributes of individuals cluster together in a finite number of groups (J. M. Keynes, C. D. Broad; Keynes's principle, however, was intended to ensure only the probability of inductive conclusions). Any of these, if true, records the presence in the universe of a certain global regularity or order that permits inductive procedures to produce the desired true conclusions. For example, if we somehow knew in advance that a given attribute *C* of an observed event must have some other attribute invariably associated with it, and if we further knew that the associated attribute must be included in a finite list of known attributes, say E_1, E_2, …, E_n, then there would be a good prospect that repeated observations of similar events would eliminate all but one of the possible associations, E_1—E_i. Refinements aside, this is how Mill, for instance, conceived of inductive method; his celebrated "methods" (which have received attention out of all proportion to their merits) reduce, in the end, to deductive procedures for eliminating unfit candidates for the title of necessary or sufficient conditions. (Later attempts to develop eliminative induction follow substantially the same path.)

It is clear that the whole interest of this program rests upon the considerations that can be advanced in favor of the supreme premises. If the supreme premises can be known to be true, the remaining processes of inference become trivial (so that there is no need for an autonomous logic of induction); if not, the entire project floats in the void.

The task of formulating plausible principles of the sort envisaged by this program has proved harder than Mill supposed. However, it may be argued that the search for them is pointless and misguided. For one thing, they would accomplish too much: If known to be true, they would allow the conclusions of selected primary inductions to be demonstrated as true, which is too much to expect. It is generally agreed (and rightly so) that the conclusion of even the best inductive argument may without

contradiction turn out to be false—if only through bad luck.

Still more serious is the problem of how, from the standpoint of this program, the desired supreme premises could ever be known to be true. Since appeal to induction is excluded at this point on the score of circularity, and since the principles themselves cannot be analytic if they are to serve their desired purpose, there seems no recourse at all. At this point those who search for supreme inductive principles find themselves with empty hands. Mill, for instance, was compelled to let his whole program rest upon the supposed reliability of simple enumeration (the method he regarded as the weakest), in whose defense he had nothing better to say than that it is "universally applicable" (which, on his principles, delightfully begs the question); Keynes, forsaking his empiricist principles for a half-hearted flirtation with Immanuel Kant, could do no better than to suggest that the ultimate principles rest upon "some direct synthetic knowledge" of the general regularity of the universe. Induction may indeed beg to be spared such defenders as these; better the robust skepticism of Hume or Popper than the lame evasions of Mill or Keynes. The conclusion seems inescapable that any attempt to show (as Bacon and many others have hoped) that there are general ontological guarantees for induction is doomed to failure from the outset.

Recourse to probability. A more promising way, at least at first sight, of hewing to the deductive line is to modify the conclusion of an inductive argument by including some explicit reference to probability. This approach, influential since Keynes's spirited exposition of it, still has many adherents. If there is no prospect of plugging the deductive gap between *P* and *K* by adding further premises known to be true, then perhaps the same end can be achieved by weakening the conclusion. If *K* does not follow from *P*, why not be satisfied with a more modest conclusion of the form "Probably, *K*" or perhaps "*K* has such and such a probability relative to *P*"?

The most impressive projects of this sort so far available have encountered severe technical difficulties. It is essential to Keynes's program, for instance, that the probability of a generalization relative to an unbroken series of confirmatory instances steadily approach unity. The conditions necessary for this to be possible in his program are at least that the generalization have an initial nonzero probability and that infinitely many of the confirmatory instances be independent, in the sense of having less than maximal probability of occurrence given the already accumulated evidence. The supreme ontological principles to which Keynes was ultimately driven to appeal (see the preceding section) hardly suffice to satisfy these conditions; subsequent criticism—for example, by Nicod and G. H. von Wright—has shown that even more rigorous conditions are needed. (Von Wright has argued that the desired asymptotic convergence will result only if in the long run every instance of the generalization is scrutinized—which would certainly render the theory somewhat less than useful in practical applications.) For all his importance as a founder of confirmation theory, the theory advocated by Keynes must be judged a failure.

Carnap's construction. The merits of Rudolf Carnap's impressively sustained construction of inductive logic, following in the tradition of Laplace and Keynes but surpassing the work of both in elaboration and sophistication, are still in dispute. Taking probability to express a logical relation between propositions, Carnap has shown how, in certain simplified languages, it is possible to define the breadth or logical width of a given proposition. (Roughly speaking, the degree of confirmation given by a proposition x to a proposition y is the ratio of the width of $x \cdot y$ to the width of x.) The definition of logical width depends on the class of possible universes expressible in the language in question. In order to assign a definite measure of logical width it is necessary to adopt some method of weighting the various possible universes ("state descriptions," in Carnap's terminology) compatible with a given proposition.

One of the merits of Carnap's analysis is to have shown that there is an entire continuum of alternative weighting procedures and associated inductive methods, each of which is internally coherent. The arbitrariness thereby recognized in inductive procedure has worried even the most sympathetic of Carnap's readers; still more disturbing is the emergence of what might be called the paradox of the unconfirmable generalization—the impossibility of ensuring, by Carnap's principles, that an unbroken series of positive instances will raise the probability of a generalization above zero. (Carnap retorts that an instance confirmation—that is, the conclusion of an eduction—does acquire progressively increasing probability, but this is insufficient to satisfy those critics who still hope to find a place for authentic generalization within inductive method.) It is too soon to decide whether such problems as these are more than the teething pains of a new subject. The ingenious modifications of Carnap's program suggested by, among others, J. G. Kemeny and Jaakko Hintikka offer some hope for their elimination.

More serious is the fundamental difficulty that flows from Carnap's conception of confirmation statements as

analytic. If it is a truth of logic (broadly speaking) that given the selected definition of confirmation, presented evidence confirms a given hypothesis to such-and-such a degree, then how could such an a priori truth justify any rational belief in the hypothesis? Or, again, if someone were to adopt a different definition of confirmation and thereby be led to a contrary belief, then how could he be shown to be in error?

Carnap's answer is based on the notion that the bridge between confirmation, as defined by him, and rational belief is to be found in some principle for the maximization of expected utility (due allowance being made, however—in his sophisticated rendering of that principle—for subjective estimates of probabilities and utilities). Yet it seems that because considerations of probability also enter into the calculation of probabilities and expected utilities, a logical circle is involved here. Since Carnap's discussions of this fundamental point are still comparatively rough and provisional, it would be premature to reach any final judgment on the success that he and those who agree with him are likely to achieve in coping with this basic difficulty. (It might be said that difficulty with the connection between probability judgments and practice is not peculiar to Carnap's work, since it arises in one form or another for all theorists of induction who take the trouble to work out in detail the consequences of their principles and assumptions.) It may be held, however, that Carnap's relatively cursory judgments about the justification of induction belong to the least satisfactory parts of his work on inductive logic.

How much the recourse to probability will accomplish depends, of course, upon how the reference to probability is construed. With empirical interpretations of probability, such as those favored by "frequentists," the probability conclusion still extends beyond the premises by covert reference to finite or infinite sets of events not covered by the given premises. The inductive leap remaining in the reconstructed argument will thus still leave the problem of induction unsolved. If, however, probability is construed in some logical way (as by Keynes or Carnap), the amended conclusion will say less than the premises and will therefore be untouched by subsequent empirical test; the deductive validity of the reconstructed argument will be saved only at the cost of rendering problematic its relevance to prediction and empirical control. In converting a purportedly inductive argument into a valid deductive one, the very point of the original argument—that is, to risk a prediction concerning the yet unknown—seems to be destroyed.

PRAGMATIC DEFENSES. Answers of the pragmatic type, originally offered by Peirce but independently elaborated with great resourcefulness by Hans Reichenbach, are among the most original modern contributions to the subject. To many they still offer the best hope of avoiding what seems to be the inevitable failure of the attempts so far discussed. The germ of the pragmatic strategy is the reflection that in ordinary life, situations sometimes arise where, in default of reliable knowledge of consequences, problematic choices can still be justified by a "nothing to lose" argument. Faced with a choice between an operation for cancer and a sure death, a patient may choose surgery, not because of any assurance of cure but on the rational ground that nothing is lost by taking the chance.

Reichenbach's "vindication." According to Reichenbach, the case is similar in what he takes to be the paradigmatic inductive situation. Given an antecedent interest in determining the probability of occurrence of a designated character (construed, by him, as the limit, in an infinitely long run of events, of the relative frequency of occurrence of that character), Reichenbach argues that the only rationally defensible policy is to use the already ascertained relative frequency of occurrence as a provisional estimate of the ultimate limiting value. A man who proceeds in this way can have no guarantee or assurance that his estimates, constantly revised as information about the series gradually accumulates, will bring him into the neighborhood of a limiting value of the frequency, for the provisional values of the relative frequencies may, in fact, diverge. In that case no predictive policy at all will work, and successful induction is impossible.

However, if this should not be the case and the series really does have a limiting value for the relative frequency in question, we can know in advance, and with certainty, that the policy is bound eventually to lead the reasoner to estimates that will remain as close to the limit as desired. There is therefore nothing to lose by adopting the inductive policy: If the series of events under scrutiny is sufficiently regular to make induction possible, the recommended policy is bound to yield the desired result ultimately (and we know before we start that it will do so), whereas if the series is irregular enough to defeat the standard inductive policy, nothing will avail, and we are no worse off than if a contrary decision had been made.

This type of justification is often called "vindication," as Herbert Feigl termed it. It is claimed that in a sense the type of vindication sketched above resolves Hume's problem by bypassing it. We know for certain that what Hume desired—namely, certification of the soundness of inductive argument by the standard of demonstrative reason-

ing—cannot be supplied. But it would be fainthearted to leave the matter there. By conceiving the practice of induction as the adoption of certain policies, applied in stoic-acceptance of the impossibility of assured success in obtaining reliable knowledge concerning matters of fact, we are able to see that such policies are, in a clear sense, preferable to any of their competitors. Standard induction is preferable to soothsaying because we know that it will work (will approach limiting values in the long run) if anything will.

To these plausible claims it has been objected that the analogy with genuinely practical decisions to act upon insufficient evidence is misconceived, for in the state of perfect ignorance postulated by defenders of the pragmatic approach no method at all can be regarded as superior to any other. Vindicationists have been relatively undisturbed by such general criticism; they have, however, felt obliged to seek remedies for a grave technical flaw that threatens to wreck their entire program. Given the assumption that the best to be achieved by an inductive policy is asymptotic convergence to a limiting relative frequency, it is obvious that no policy for inductive estimation in the short run is excluded as unreasonable. Thus, from the standpoint of pragmatic vindication an unbroken run of *A*'s found to be *B* would not make it unreasonable to predict the subsequent occurrence *in the short run* of *A*'s that are not *B*, provided only that the adopted estimates are chosen so as to converge eventually to the limit (if it exists). But since the long run is in fact never attained, even by immortal beings, it follows that the pragmatic defense yields no criteria for inductive decisions in short-run cases, to which inductive prediction is confined, and offers no differential reasons for preferring one inductive policy to another.

In spite of strenuous attempts (notably by Wesley Salmon) to improve Reichenbach's original conception by providing supplementary reasons for rejecting unwanted nonstandard policies, the prospects for vindicationism remain dubious. Even if some plausible way could be found of assigning, on vindicationist principles, a special status to the standard policy of induction, the approach would be vulnerable to the objection that it conceives inductive method in an eccentrically restricted fashion. The determination of limiting values of relative frequencies is at best a special problem of inductive method and by no means the most fundamental.

Peirce's views. Peirce, whose views on induction have exerted a lasting influence on the subject since the posthumous appearance of his *Collected Papers*, had a more complex conception of scientific method than latter-day vindicationists. Induction, conceived by him as a process of testing statistical hypotheses by examining random samples, has to be understood in its relations to two other procedures, statistical deduction and abduction.

Statistical deduction consists of inference from the frequency of occurrence of an attribute in a population to the probable and approximate occurrence of that attribute in a sample randomly drawn from it. Given Peirce's definition of probability as limiting frequency and his conception of randomness, it follows demonstratively that most of the samples drawn will have nearly the same composition as the parent population; statistical deduction is thus "valid" in the sense that it generates conclusions that are true most of the time.

Abduction, the creative formulation of statistical hypotheses and the only mode of scientific inference introducing new ideas, is a kind of inversion of statistical deduction. It has almost no probative force, its value being rather that it provides new generalizations needing independent verification and having "some chance of being true."

When the three procedures are used in combination, induction is seen to be a self-correcting method that if indefinitely followed must in the long run lead the scientific community, although not the individual reasoner, indefinitely close to the truth. In such asymptotic convergence to the truth lies the peculiar validity of induction.

Peirce cannot be held to have succeeded in his effort to defend the rationality of inductive policies in terms of long-range efficacy in generating conclusions approximately and for the most part true. Since the intended justification of induction depends essentially upon the randomness of the samples used, it must be objected that there is normally no way of guaranteeing in advance the presence of such randomness. (To this objection Peirce had only the lame and unsupported rejoinder that inductive inference retains some probative force even in the absence of the desired randomness.) The following are among the most obvious weaknesses of Peirce's views about induction.

The self-corrective tendency of induction, which Peirce, in his last writings on the subject, came to view as the heart and essence of inductive method, remains obscure, in spite of his eulogies. That inductive estimates will need, on Peirce's principles, repeated adjustments as further evidence accumulates is clear enough, but that this process will show any convergence toward a limiting value cannot be guaranteed a priori. If the samples to be

examined were random in Peirce's severe sense of that term, we could at least count upon an overall predominance of approximately correct estimates, but even then we should have no reason, in the absence of additional guarantees, to expect the better estimates to come near the end of the testing process. In any case, supposing realistic conditions for the testing of hypotheses (such as our necessary reliance on cases that we are in a position to examine), it seems clear that the conditions for the kind of sampling demanded by Peirce cannot be fulfilled.

Peirce's references to the long run seem on the whole incoherent. Much of the time he seems to have been thinking of what would prove to be the case in an actual but infinitely extended series of trials. Toward the end of his life, however, he appears to have recognized that his definitions of probability and of the validity of induction needed to be construed more broadly, by reference to the "would be" of events, conceived as real general characters or habits. How such general features of events can in fact be disclosed, even by very lengthy series of trials, Peirce never made plain. Yet the need for clarification is great for anybody attracted by his approach. The infinitely long run is a chimera, and to be told that a certain method, if consistently pursued, would in such a long run eventually lead as close as we pleased to the truth is to be told nothing that can be useful for the actual process of verification. All verification is necessarily performed in the finite run, however extended in length, and what would happen if *per impossibile* the "run" were infinite is not relevant to the relative appraisal of given hypotheses. We need a method for adjudicating between rival hypotheses, if not now then in the foreseeable future, and this Peirce's conception cannot provide. Because of his reliance upon the infinitely long run Peirce's pragmaticism, which initially seems so hardheaded in its emphasis upon success and practical consequences, ends by being as utopian as any of the metaphysical conceptions that he derided.

JUSTIFICATION AS A PSEUDO PROBLEM. In view of the quandaries that beset all known attempts to answer Hume's challenge, it is reasonable to consider whether the problem itself may not have been misconceived. Indeed, it appears upon examination that the task of logical justification of induction, as classically conceived, is framed so as to be a priori impossible of solution. If induction is by definition nondeductive and if the demand for justification is, at bottom, that induction be shown to satisfy conditions of correctness appropriate only to deduction, then the task is certainly hopeless. But to conclude, for this reason, that induction is basically invalid or that a belief based upon inductive grounds can never be reasonable is to transfer, in a manner all too enticing, criteria of evaluation from one domain to another domain, in which they are inappropriate. Sound inductive conclusions do not follow (in the deductive sense of "follow") from even the best and strongest set of premises (in the inductive sense of "strongest"); there is no good reason why they should. Those who still seek a classical defense of induction may be challenged to show why deductive standards of justification should be appropriate. Perhaps the retort will be that there is no clear sense in which assertion of a conclusion is justified except the sense in which it is known to follow strictly from premises known to be true, so the burden of argument rests upon anybody who claims the existence of some other sense.

Linguistic approach to the problem. The challenge to the claim that inductive arguments cannot be said to be justified might be met in the following way: Suppose a man has learned, partly from his own experience and partly from the testimony of others, that in a vast variety of circumstances, when stones are released they fall toward the ground. Let him consider the proposition *K*, that any stone chosen at random and released will do likewise. This is, in the writer's opinion, a paradigm case for saying that the man in question (any of us) has a good reason for asserting *K* and is therefore justified in asserting *K* rather than not-*K*. Similarly, this is a paradigm case for saying that the man in question is reasonable in asserting *K* and would be unreasonable in asserting not-*K*, on the evidence at hand. Anybody who claimed otherwise would not be extraordinarily and admirably scrupulous but would be abusing language by violating some of the implicit criteria for the uses of "good reason," "justified," and "reasonable," to which he, like the interlocutor with whom he succeeds in communicating, is in fact committed.

Any man—say, one from Mars—who used these words according to criteria that would really make it improper for him to apply them in the kind of situation envisaged would not, in the end, be understood by us. Worse still, he would be trying, if he were consistent, to change our actual concepts of reason and reasonableness so that it would be logically impossible to have reasons for assertions concerning the unknown or to be reasonable in expecting one matter of fact rather than another on the basis of empirical evidence. (He would be behaving like a man who insisted that only stallions deserved to be called horses.) Nor would such distortion achieve anything significant, for the man who proposed to make "empirical reason" as impossible of application as "being in two places at once" would find himself forced to rein-

troduce essentially the same concept under some such label as “generally accepted as a reason” or “what commonly passes for a reason.” The distinction between what ordinary men and what scientists call “good reasons” and “bad reasons” is made for a good purpose, has practical consequences, and is indispensable in practice. Thus, the dispute between the advocate of the linguistic approach and his opponent seems to reduce to a verbal one, ripe for oblivion.

Given the intertwined complexity of the concepts entering into alternative formulations of the problem of induction and the seductive plausibility of the distortions to which such concepts are subject, no brief reply such as the above can be expected to clarify and to expose the conceptual confusions upon which traditional formulations of the problem rest. A full discussion would at least also have to consider the relevant senses of “knowledge” and “possibility” and related epistemological notions. The outline of the strategy is perhaps sufficiently plain; the line to be taken is that close and detailed examination of how the key words in the statement of the problem occur will show that criteria for the correct uses of such terms are violated in subtle and plausible ways. If this can be established, the celebrated problem of justifying induction will dissolve, and the confused supposition that induction needs philosophical justification or remains precarious in its absence will disappear.

The comparative problem and the analytical problem do not dissolve under this attack. Advocates of the linguistic approach can be fairly reproached for having been too often content to show to their own satisfaction that the general problem of justification is rooted in confusion, while neglecting the constructive tasks of rendering clearer the criteria for preferential appraisal of inductive arguments.

To those unsympathetic with the linguistic approach such an attack upon the traditional problem has sometimes seemed to be operating with dubious and insufficiently elaborated theories of meaning or use and to be altogether too glib in its attribution of semantical confusions. Moreover, a number of critics have thought that an appeal to ordinary language cannot be ultimately decisive from a philosophical standpoint. Even if it were established that it is a violation of ordinary language to describe the conclusion of some inductive arguments as supported by less than good reasons, the critics ask, what is there in the nature of things that requires us to continue talking in the ordinary way or to be bound by the encapsulated metaphysical prejudices of those originally responsible for establishing the rules of use to which appeal is now made? The linguistic philosopher necessarily uses such key words as *reasonable* in his polemic against the traditional approaches to the problem. But to use the crucial terms in a discussion of the nature of the inductive problem, it might be urged, is to beg the very question at issue. A lunatic or an eccentric philosopher might well use the expression “good reason” in a way that would be blatantly improper, yet he might be able to prove, by appeal to his own criteria, that he had “good reasons” to use the phrase in the way he did. But are we ourselves in any better position? Are we not obligated to break through the linguistic barrier and at least to show why the alleged criteria for good reasons to which appeal is made should continue to receive our allegiance?

There is no short way of dealing with this type of objection. It may be helpful, however, to sketch the general view upon which the present writer, as a defender of the linguistic approach, would rely.

Defense of the linguistic approach. All normal adult human beings follow the same broad and systematic patterns for drawing inferences concerning the unobserved and apply the same general principles for appraising such nondemonstrative inferences. For instance, all normal persons expect observed cases of association of attributes to be confirmed in further experience unless there are countervailing factors (the principle of simple enumeration), all count increase in the number of independent confirmatory instances of a law as strengthening (or at least not weakening) the probability of the law’s truth, and all alike share the inductive beliefs that underpin causal notions. It is, therefore, not fanciful to conceive of all sane adult human beings as participating in a complex system of ways of learning from experience that might be called the inductive institution. Like other institutions (warfare, the law, and so on), it has a relatively fixed, though not immutable, structure, transmitted from one generation to the next and crystallized in the form of prohibitions and licenses, maxims of conduct, and informal precepts of performance. Like other institutions, the inductive institution requires that its participants have mastered a system of distinctive concepts (among them the concepts of good reason, sound argument, and relative likelihood) having both descriptive and normative aspects.

Such mastery is shown in capacity to use the corresponding language correctly—which, in turn, implies recognition of, though not invariable obedience to, associated rules for assertion, for evaluation, and for the appraisal of actions. Understanding what people mean by reasons for empirical conclusions requires acceptance of

certain types of situations as paradigmatic of empirical evidence; to call given facts sound reasons for some conclusion is to imply the acceptability of certain criteria for judging one reason to be better than another; asserting that some belief about the hitherto unobserved is reasonable commits the speaker to holding that other things being equal, action based on such belief should be approved.

The philosophical problem of justifying induction can arise only for somebody who is a member of the inductive institution and is therefore already bound by its constitutive rules. A spectator can understand bridge without being a player, but all of us are necessarily players of the "inductive game" before we achieve the reflective self-consciousness characteristic of philosophical criticism.

The constitutive rules of the inductive institution (whose precise delineation remains a still unfinished task for philosophers of induction) are highly abstract, schematic, and limited in their practical usefulness. Indeed, the general principles of inductive inference are about as relevant to practice as the abstract principles of justice are to decisions on concrete legal issues. In particular situations concerning the soundness of empirical hypotheses the reasoner is compelled to fall back upon his specific knowledge of relevant facts and theories. In this way the conduct of concrete inductive inference resembles the exercise of a craft or skill more than it does the automatic application of a decision procedure. Yet the constitutive rules provide important general constraints that cannot be violated without generating nonsense. To be in command of inductive language, whether as a master of advanced techniques of statistical inference or as a layperson constantly and more or less skillfully anticipating future experience, is necessarily to be subject to the implicit norms of belief and conduct imposed by the institution.

The inductive concepts that we acquire by example and formal education and modify through our own experiences are not exempt even from drastic revision. The norms may be usefully thought of as formal crystallizations into linguistic rules of general modes of response to the universe that our ancestors have, on the whole, found advantageous to survival, but the earlier experience of the race never has absolute authority. Piecemeal reform of the inductive institution can be observed in the history of modern science.

What is clearly impossible, however, is the sort of wholesale revolution that would be involved in wiping the inductive slate clean and trying to revert to the condition of some hypothetical Adam setting out to learn from experience without previous indoctrination in relevant rules of inductive procedure. This would be tantamount to attempting to destroy the language we now use to talk about the world and about ourselves and thereby to destroy the concepts embodied in that language. The idea of ceasing to be an inductive reasoner is a monstrosity. The task is not impossibly difficult; rather, its very formulation fails to make sense. Yet it remains important to insist that the inductive institution, precisely because its raison d'être is learning from experience, is intrinsically self-critical. Induction, like the Sabbath, was made for humankind, not vice versa. Thus, constantly renewed experience of the successes and failures of the specific inductive procedures permitted within the general framework of the inductive institution provides a sound basis for gradual reform of the institution itself, without objectionable circularity.

Yet even if no feature of the institution is exempt, in principle, from criticism and reconstruction, the entire institution cannot be called into question all at once without destroying the very meaning of the words in which the philosophical problems of induction are stated. Wholesale philosophical skepticism about matters of fact is senseless and must be shown to be so. If this is the "linguocentric predicament," we must make the best of it.

The view here outlined must be carefully distinguished from what is commonly called conventionalism. The argument is not that the constitutive inductive rules hold by convention but rather that the sweeping question "Why should we accept *any* inductive rules?" can be shown to make no sense.

Our sketch may be usefully compared with Hume's view of induction as a habit or custom. Both views agree in regarding inductive practices as being, on the whole, social and contingent facts obtaining at given periods in human history. It is, after all, a contingent fact that there have existed animals sufficiently rational to be able to speak and hence to have inductive concepts. The present conception differs significantly from Hume's, however, in regarding the inductive institution as partly constituted by normative inductive rules to which the philosopher, like every reasoning individual, finds himself already committed. Thus, the encompassing social fact of the existence of the inductive institution includes within itself the means for appraisal and criticism of inductive procedures; we cannot regard inductive inference as something merely "given," as a natural fact, like the Milky Way, that it would be absurd to criticize. To understand induction is necessarily to accept its authority. However

(to repeat), questions about the general or ultimate justification of induction *as such*, questions of the form "Why should any induction be trusted?" must be recognized as senseless. If we persist in trying to raise them, we come, as Wittgenstein expressed it, to the "limits of language," and we can see that we have done so by perceiving that what we had hoped were important and fundamental questions are no better than nonsense masquerading as sense. The foregoing will undoubtedly strike critics of the linguistic approach as too facile, for the tangle of philosophical problems that have been dubbed "the problem of induction" constitute, in their depth, their importance, their elusiveness, and their capacity to bewilder and confuse, a very paradigm of philosophical perplexity.

The preceding survey indicates that no wholly satisfactory philosophy of induction is yet available. The work still to be done may be summarized as follows: For those who recognize the crucial role of probability in inductive inference, to develop a consistent, systematic, and relevant reconstruction of the concept of probability; for those who reject induction as an outmoded myth, to elaborate a detailed and comprehensive account of scientific practice that will be reasonably close to the best actual procedures used in reasoning about matters of fact; for those who pin their hopes on the construction of an inductive logic, to remove the constraints imposed by the study of artificially simplified languages and to show in detail how analytical statements of probability can be relevant to the practice of inductive prediction; for vindicationists, to solve the comparative problem of selecting competing hypotheses and to show how eventual convergence in the long run can bear upon short-run judgment; for those who regard induction as a pseudo problem, to articulate the theory of language presupposed and to demonstrate in convincing detail the origins and the character of the stubborn confusions that have infested the subject.

See also Laws of Nature; Probability and Chance.

Bibliography

COMPREHENSIVE STUDIES

The best introduction to the whole subject of induction is still William Kneale, *Probability and Induction* (Oxford: Clarendon Press, 1949). See especially Part II, "The Traditional Problem of Induction." A shorter and more up-to-date discussion is Stephen F. Barker, *Induction and Hypothesis* (Ithaca, NY: Cornell University Press, 1957), which is especially good on the role of simplicity in inductive inference. Georg Henrik von Wright, *The Logical Problem of Induction* (Oxford, 1941; 2nd ed., 1957), is invaluable for its ample discussion of the history of the subject and also contains penetrating criticism. John Patrick Day, *Inductive Probability* (London: Routledge and Paul, 1961), uses somewhat opaque symbolism but is very comprehensive. *Induction: Some Current Issues*, edited by Henry E. Kyburg Jr. and Ernest Nagel (Middletown, CT: Wesleyan University Press, 1963), is a conference report containing edited versions of discussions by Black, Braithwaite, Nagel, Salmon, and others.

TERMINOLOGY

For a good brief treatment of terminology, see Kneale, op. cit., pp. 24–48, which includes discussion of Aristotle's uses of *epagoge*. Kneale is good also in his treatment of intuitive induction and mathematical induction.

S. F. Barker, "Must Every Inference Be Either Deductive or Inductive?," in *Philosophy in America*, edited by Max Black (Ithaca, NY: Cornell University Press, 1965), pp. 58–73, answers in the negative; it is an illuminating effort to distinguish criteria for nondemonstrative inference.

HISTORY OF INDUCTIVE METHODS

André Lalande, *Les théories de l'induction et de l'expérimentation* (Paris: Boivin, 1929), contains brief discussions of Isaac Newton, John Herschel, Claude Bernard, and other scientists, as well as of Bacon and Mill. There is no satisfactory general history of inductive methods.

Most students of Bacon confine themselves to Book I of the *Novum Organum* (available in many editions). The best philosophical commentary on Bacon remains R. L. Ellis's general preface to *The Works of Francis Bacon*, edited by James Spedding, R. L. Ellis, and D. D. Heath (London: Longman, 1857–1854). For a lucid and sympathetic appreciation of Bacon's general program—but with comparatively little discussion of Bacon's views on induction—see the commemorative address by C. D. Broad, *The Philosophy of Francis Bacon* (Cambridge, U.K.: Cambridge University Press, 1926).

For Mill, see *A System of Logic*, 8th ed. (London, 1872), especially Book III, Ch. 3, "On the Ground of Induction," and Ch. 21, "Of the Evidence of the Law of Universal Causation." Ernest Nagel, ed., *John Stuart Mill's Philosophy of Scientific Method* (New York: Hafner, 1950), is a useful selection from the *Logic*, including part of Mill's *Examination of Sir William Hamilton's Philosophy*. The editor's introduction is helpful.

A serious student of Mill will wish to consult his unjustly neglected critic, William Whewell, whose *The Philosophy of the Inductive Sciences* (London: J.W. Parker, 1847)—especially Part II, titled "Novum Organum Renovatum"—is still instructive. A useful summary of Whewell's views is C. J. Ducasse's "William Whewell's Philosophy of Scientific Discovery," pp. 183–217 in *Theories of Scientific Method: The Renaissance through the Nineteenth Century*, edited by Edward H. Madden (Seattle: University of Washington Press, 1960).

For Peirce's views, see the *Collected Papers of Charles Sanders Peirce*, edited by Charles Hartshorne, Paul Weiss, and Arthur Burks (8 vols., Cambridge, MA: Harvard University Press, 1931–1958), especially "Ampliative Reasoning," Vol. II, pp. 272–607. Unfortunately, Peirce's writings on induction are scattered throughout the volumes, out of chronological

order. A concise critical examination is Thomas A. Goudge, "Peirce's Treatment of Induction," in *Philosophy of Science* 7 (1940): 56–68. Part VI of the same author's *The Thought of C. S. Peirce* (Toronto: University of Toronto Press, 1950) has a fuller treatment, for which the rest of the book provides good background.

For Keynes, see *A Treatise on Probability* (London: Macmillan, 1921). Chs. 18–22 are the most relevant; they include, inter alia, Keynes's original ideas concerning analogy. Jean Nicod, "The Logical Problem of Induction," Part II of *Foundations of Geometry and Induction* (London: K. Paul, Trench, Trubner, 1930), contains lucid and trenchant criticisms of Keynes's position. For more recent appraisal, see von Wright, op. cit., pp. 127–132.

Rudolf Carnap, *Logical Foundations of Probability* (Chicago: University of Chicago Press, 1950), is an elaborate treatise that, although still unfinished and in part superseded by Carnap's later studies, remains the most important original source for his views.

The Philosophy of Rudolf Carnap, edited by P. A. Schilpp (La Salle, IL: Open Court, 1963), is now indispensable to the serious student because of its important essays by Carnap's defenders and critics and the detailed comments and replies by Carnap. John G. Kemeny, "Carnap's Theory of Probability and Induction," pp. 711–738, is an outstandingly successful effort to convey the gist of Carnap's position sympathetically; Ernest Nagel, "Carnap's Theory of Induction," pp. 785–825 (vigorously attacked in Carnap's rejoinder, pp. 989–995), expresses at length the misgivings of those who see in Carnap's view a retreat from empiricism; Hilary Putnam, "'Degree of Confirmation' and Inductive Logic," pp. 761–783, is a highly ingenious attempt to demonstrate that a logic based on confirmation must violate accepted canons of scientific method.

Carnap's "The Aim of Inductive Logic," pp. 303–318 in *Logic, Methodology, and Philosophy of Science*, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski (Stanford, CA: Stanford University Press, 1962), is an important pioneering discussion of the application of inductive logic. Carnap concludes, "Induction, if properly reformulated, can be shown to be valid by rational criteria."

PROBLEM OF JUSTIFICATION

There is no substitute for reading Hume. His own summary, *An Abstract of a Treatise of Human Nature* (London, 1740; reprinted, with introduction by J. M. Keynes and Piero Sraffa, Cambridge, U.K.: Cambridge University Press, 1938), should not be overlooked. Pp. 11–20 express the essence of Hume's position. The serious student must, of course, read *A Treatise of Human Nature*, edited by L. A. Selby-Bigge (Oxford: Clarendon Press, 1896). See especially Book I, Part 3. Sec. 6 contains the famous skeptical attack on the objectivity of causal connection.

See also Hume's *Enquiry Concerning the Human Understanding*, edited by L. A. Selby-Bigge, 2nd ed. (Oxford: Clarendon Press, 1902), especially Sec. 4, "Sceptical Doubts concerning the Operations of the Understanding."

Bertrand Russell, "On Induction," Ch. 6 of *The Problems of Philosophy* (London: Williams and Norgate, 1912), is a lucid and concise version of what is basically Hume's skeptical position.

The best statement of the "new riddle of induction" is in Nelson Goodman, *Fact, Fiction and Forecast* (Cambridge, MA: Harvard University Press, 1955), Chs. 3 and 4. Goodman's position, with some unpublished improvements, is sympathetically discussed in Israel Scheffler, *The Anatomy of Inquiry* (New York: Knopf, 1963), pp. 291–326. The same book is valuable for its lucid appraisal of various theories of inductive confirmation.

REJECTION OF INDUCTION

See Karl Popper, *The Logic of Scientific Discovery* (London: Hutchinson, 1959), especially Ch. 1. John Oulton Wisdom, *Foundations of Inference in Natural Science* (London: Methuen, 1952), is written from a Popperian standpoint; Parts II–IV, pp. 85–232, deal with induction. For criticism of Popper's general position, see John Arthur Passmore, "Popper's Account of Scientific Method," in *Philosophy* 35 (1960): 326–331.

A PRIORI DEFENSES

Donald C. Williams, *The Ground of Induction* (Cambridge, MA: Harvard University Press, 1947), claims to have answered Hume on a priori grounds. For a dissenting verdict, see the extended review by Max Black in *Journal of Symbolic Logic* 12 (1947): 141–144.

Roy F. Harrod, *Foundations of Inductive Logic* (London: Macmillan, 1956), is still another ingenious attempt to show that induction must be successful. For criticism, see Popper's review in *British Journal for the Philosophy of Science* 9 (1958–1959): 221–224; Harrod's reply appears in 10 (1959–1960): 309–312.

INDUCTIVE SUPPORT FOR INDUCTION

Consult Richard B. Braithwaite, *Scientific Explanation* (Cambridge, U.K.: Cambridge University Press, 1953), Ch. 8, "The Justification of Induction," pp. 255–292.

Abner Shimony's severely critical "Braithwaite on Scientific Method," in *Review of Metaphysics* 7 (1953–1954): 644–660, has not been rebutted. On this topic see also H. E. Kyburg Jr., "R. B. Braithwaite on Probability and Induction," in *British Journal for the Philosophy of Science* 9 (1958–1959): 203–220.

A defense of the soundness of self-supporting inductions can be found in Max Black, "Inductive Support of Inductive Rules," pp. 191–208 in *Problems of Analysis* (Ithaca, NY: Cornell University Press, 1954), and "Self-Supporting Inductive Arguments," pp. 209–218 in *Models and Metaphors* (Ithaca, NY: Cornell University Press, 1962). For criticism of Black's views, see Peter Achinstein, "The Circularity of a Self-Supporting Argument," in *Analysis* 22 (1961–1962): 138–141. See also Black's reply in the same journal, 23 (1962–1963): 43–44, and Achinstein's rejoinder, ibid., 123–127.

DEDUCTIVE RECONSTRUCTION

See Bertrand Russell, *Human Knowledge: Its Scope and Limits* (New York: Simon and Schuster, 1948). Russell states and defends postulates of scientific inference in the concluding part, pp. 421–507. Arthur W. Burks, "On the Presuppositions of Induction," in *Review of Metaphysics* 8 (1955): 574–611, is a persuasive exposition by the ablest contemporary defender of deductive reconstruction. See also Burks's article in *The Philosophy of Rudolf Carnap*,

edited by Schilpp (above), and his *Chance, Cause, Reason: An Inquiry into the Nature of Scientific Evidence* (Chicago: University of Chicago Press, 1977). Deductive reconstruction is criticized in most of the articles listed below in the section on induction as pseudo problem.

C. D. Broad, "On the Relation between Induction and Probability," in *Mind* 27 (1918): 389–404, and 29 (1920): 11–45, is an early and unjustly neglected discussion whose second part is a suggestive development of Keynes's ideas.

PRAGMATIC DEFENSES

Herbert Feigl's views are contained in two of his articles, "De Principiis Non Est Disputandum …?," in *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950), and "Validation and Vindication," in *Readings in Ethical Theory*, edited by Wilfrid Sellars and John Hospers (New York: Appleton-Century-Crofts, 1952), pp. 667–680. The first paper argues the case for vindication as a special mode of justification, and the second enlarges upon the idea.

Hans Reichenbach, *The Theory of Probability* (Berkeley: University of California Press, 1949), is the locus classicus of recent pragmatic defenses; see especially "The Justification of Induction," pp. 469–482. The same author's *Experience and Prediction* (Chicago, 1938) contains a more popular presentation.

Isabel P. Creed, "The Justification of the Habit of Induction," in *Journal of Philosophy* 37 (1940): 85–97, and Everett J. Nelson, "Professor Reichenbach on Induction," in *Journal of Philosophy* 33 (1936): 577–580, are two thorough appraisals and criticisms of Reichenbach's position.

Max Black, "'Pragmatic' Justifications of Induction," in his *Problems of Analysis* (Ithaca, NY: Cornell University Press, 1954), pp. 157–190, is a full-scale attack upon "pragmatic vindication."

Wesley C. Salmon, "The Short Run," in *Philosophy of Science* 22 (1955): 214–221, is an admirable, if ultimately unsuccessful, attempt to avoid reliance upon the infinitely long run. In "Should We Attempt to Justify Induction?," *Philosophical Studies* 8 (1957): 33–48, Salmon elaborately defends the pragmatic approach; Black's "Can Induction Be Vindicated?," in his *Models and Metaphors* (Ithaca, NY: Cornell University Press, 1962), pp. 194–208, is a detailed reply to Salmon.

Salmon's "Vindication of Induction," in *Current Issues in the Philosophy of Science*, edited by H. Feigl and G. E. Maxwell (New York: Holt, Rinehart and Winston, 1961), pp. 245–256, is the latest version of Reichenbach's approach by its ablest contemporary advocate. See also Salmon's contributions to Nagel and Kyburg, eds., *Induction*, above.

INDUCTION AS A PSEUDO PROBLEM

For an introduction to the growing number of studies written from this standpoint, see Alice Ambrose, "The Problem of Justifying Inductive Inference," in *Journal of Philosophy* 44 (1947): 253–272; Max Black, "The Justification of Induction," in his *Language and Philosophy* (Ithaca, NY: Cornell University Press, 1949), pp. 59–88; Frederick L. Will, "Will the Future Be Like the Past?," in *Mind* 56 (1947): 332–347; and Paul Edwards, "Russell's Doubts about Induction," in *Mind* 58 (1949): 141–163. Perhaps the most accessible brief account is Peter F. Strawson, "The 'Justification' of Induction," in *Introduction to Logical Theory* (London: Methuen, 1952), pp. 248–263. Brief criticism of the "linguistic approach is to be found in many of the works listed above. No satisfactorily broad statement of the position of the "linguists" or their critics is yet available.

BIBLIOGRAPHIES AND SURVEYS

Lengthy bibliographies are supplied in the above-mentioned books by Keynes (to 1921), Carnap (to 1951), and von Wright (to 1955).

Black's "Induction and Probability," in *Philosophy in the Mid-century*, edited by Raymond Klibansky (Florence: Nuova Italia, 1958), pp. 154–163, is a critical survey of the work of the decade prior to its publication.

H. E. Kyburg Jr., "Recent Work in Inductive Logic," in *American Philosophical Quarterly* 1 (1964): 249–287, is a highly useful analysis of the main trends, with criticism and an extensive bibliography.

OTHER RECOMMENDED TITLES

Hacking, Ian. *Logic of Statistical Inference*. Cambridge, U.K.: Cambridge University Press, 1965.

Horwich, P. *Probability and Evidence*. Cambridge, U.K.: Cambridge University Press, 1982.

Jeffrey, R. C. *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press, 1983.

Jeffrey, R. C. *Probability and the Art of Judgment*. New York: Cambridge University Press, 1992.

Jeffrey, R. C. "The Valuation and Acceptance of Scientific Hypotheses." *Philosophy of Science* 23 (1956): 237–246.

Kaplan, M. *Decision Theory as Philosophy*. New York: Cambridge University Press, 1996.

Kyburg, H. *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel, 1974.

Levi, I. *The Fixation of Belief and Its Undoing*. Cambridge, U.K.: Cambridge University Press, 1991.

Lewis, David. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*. Vol. 2, edited by Richard C. Jeffrey. Berkeley: University of California Press, 1980.

Salmon, Wesley. *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press, 1966.

Salmon, Wesley. "Inductive Inference." In *Philosophy of Science: The Delaware Seminar*, edited by B. Baumrin. New York: Interscience, 1963.

Skyrms, B. *Choice and Chance*. 4th ed. Encino, CA: Wadsworth, 2000.

Skyrms, B. *Pragmatics and Empiricism*. New Haven, CT: Yale University Press, 1984.

Sober, Elliot. "Bayesianism—Its Scope and Limits." In *Bayes' Theorem*, edited by Richard Swinburne. Oxford: Oxford University Press, 2002.

Swinburne, Richard, ed. *The Justification of Induction*. Oxford: Oxford University Press, 1974.

Van Fraassen, Bas. *Laws and Symmetry*. Oxford: Clarendon Press, 1989.

Max Black (1967)

Bibliography updated by Benjamin Fiedor (2005)

INDUCTION, CANONS OF

See *Mill's Methods of Induction*

INEQUALITY

See *Equality, Moral and Social*

INERTIA

See *Mass*

INFERENCE TO THE BEST EXPLANATION

In an inductive inference, we acquire a belief on the basis of evidence that is less than conclusive. The new belief is compatible with the evidence, but so are (possibly many) competing hypotheses that we are unwilling to infer. Such is the situation for a great number of the inferences we make, and this raises a question of description and a question of justification. What principles lead us to infer one hypothesis rather than another? And do we have any reason to believe that these principles are good ones, leading us to accept hypotheses that are true and to reject those that are false? Inference to the Best Explanation offers partial answers to both questions.

According to this model, explanatory considerations are a guide to inductive inference. We decide which of the competing hypotheses the evidence best supports by determining how well each of the competitors would explain that evidence. Many inferences are naturally described in this way. Seeing the ball next to the broken vase, I infer that my children have been playing catch in the house because this is the best explanation of what I see. Charles Darwin inferred the hypothesis of natural selection because, although it was not entailed by his diverse biological evidence, natural selection would provide the best explanation of it. When astronomers infer that a galaxy is receding from the Earth with a specified velocity, they do this because the supposition of such a recession would provide the best explanation of the observed red-shift of the galaxy's characteristic spectrum. When the detectives infer that it was Moriarty who committed the crime, they does so because this hypothesis would best explain the fingerprints, blood stains, and other forensic evidence. Sherlock Holmes to the contrary, this is not a matter of *de*duction. The evidence will not *entail* that Moriarty is to blame, since it always remains possible that someone else was the perpetrator. Nevertheless, Holmes is right to make his inference, since the supposition of Moriarty's guilt provides a better explanation of the evidence than does the supposition of anyone else's guilt.

Inference to the Best Explanation can be seen as an extension of the idea of "self-evidencing" explanations, where the phenomenon that is explained in turn provides an essential part of the reason for believing the explanation is correct. In the example above, the speed of recession explains the red-shift, but the observed red-shift may at the same time be an essential part of the reason astronomers have for believing that the galaxy is receding at that speed. Self-evidencing explanations exhibit a curious circularity, but this circularity is apparently benign. The recession is used to explain the red-shift and the red-shift is used to determine the recession, yet the recession hypothesis may be both explanatory and well supported. According to Inference to the Best Explanation, this is a common situation: Hypotheses are supported by the very observations they are supposed to explain. Moreover, on this model, the observations support the hypothesis precisely because it would explain them.

Inference to the Best Explanation thus partially inverts an otherwise natural view of the relationship between inference and explanation. According to that natural view, inference is prior to explanation. First we must decide which hypotheses to accept; then, when called upon to explain some observation, we will draw from our pool of accepted hypotheses. According to Inference to the Best Explanation, by contrast, it is only by asking how well various hypotheses would explain the available evidence that we can determine which hypotheses merit acceptance. In this sense, Inference to the Best Explanation has it that explanation is prior to inference.

Although it gives a natural account of many inferences in both science and ordinary life, the model needs further development. What, for example, do we mean by "best?" It is sometimes taken to mean "likeliest" or "most plausible," but Inference to the Likeliest Explanation would be a disappointingly uninformative model, since the main point of an account of inference is to say what leads one hypothesis to be judged likelier than another, that is, to give the symptoms of likeliness. A more promising approach construes *best* as "loveliest." In this view, we infer the hypothesis that would, if correct, provide the greatest understanding.

The model should thus be construed as "Inference to the Loveliest Explanation." Its central claim is that loveliness is a guide to likeliness, that the explanation that

would, if correct, provide the most understanding, is the explanation that is judged likeliest to be correct. This at least is not a trivial claim, but it faces at least three challenges. The first is to identify the explanatory virtues, the features of explanations that contribute to the degree of understanding they provide. There are a number of plausible candidates for these virtues, including scope, precision, mechanism, unification, and simplicity. Better explanations explain more types of phenomena, explain them with greater precision, provide more information about underlying mechanisms, unify apparently disparate phenomena, or simplify our overall picture of the world. But analyzing these and other explanatory virtues is not easy, and it also leaves the other two challenges. One of these is to show that these aspects of loveliness do indeed match judgments of likeliness, that the loveliest explanations tend also to be those that are judged likeliest to be correct. The remaining challenge is to show that, granting the match between loveliness and judgments of likeliness, the former is in fact our guide to the latter.

In addition to offering a description of our inductive practices, Inference to the Best Explanation has been used to justify them, to show that those hypotheses we judge likely to be correct really are so. For example, it has been argued that we have good reason to believe that our best scientific theories are true, since the truth of those theories is the best explanation of their wide-ranging predictive success. Indeed, it has been claimed that the successes of a theory would be inexplicable unless it were at least approximately true. This argument has considerable plausibility, but it faces serious objections. If scientific theories are themselves accepted on the basis of Inferences to the Best Explanation, then to appeal to an argument of the same form to show that those inferences lead to the truth seems to beg the question. Moreover, it is not clear that the truth of a theory really is the best explanation of its predictive success. For one thing, it seems no better an explanation than would be the truth of any other competing theory that happens to share those particular predictions. For another, to explain why our current theories have so far been successful may not require an appeal to truth if scientists have a policy of weeding out unsuccessful theories.

See also Epistemology; Naturalized Epistemology; Realism.

Bibliography

Ben-Menahem, Yemima. "The Inference to the Best Explanation." *Erkenntnis* 33 (1990): 319–344.

Day, Timothy, and Harold Kincaid. "Putting Inference to the Best Explanation in Its Place." *Synthese* 98 (1994): 271–295.

Harman, Gilbert. "The Inference to the Best Explanation." *The Philosophical Review* 74 (1965): 88–95.

Lipton, Peter. *Inference to the Best Explanation*, 2nd ed. London: Routledge, 2004.

Okasha, Samir. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in the History and Philosophy of Science* 31 (2000): 691–710.

Psillos, Stathis. *Scientific Realism: How Science Tracks Truth*. London: Routledge, 1999.

Salmon, Wesley. "Explanation and Confirmation: A Bayesian Critique of *Inference to the Best Explanation*." In *Explanation: Theoretical Approaches and Applications*, edited by Giora Hon and Sam Rakover. Dordrecht: Kluwer, 2001.

Thagard, Paul. "The Best Explanation: Criteria for Theory Choice." *The Journal of Philosophy* 75 (1978): 76–92.

Peter Lipton (1996, 2005)

INFINITESIMALS

The ubiquitous use of infinitely small quantities in mathematics dates back at least to the seventeenth century. Despite continuing qualms as to their legitimacy and their supposed elimination as a result of the thoroughgoing reform movement of the nineteenth century, "infinitesimals" have continued to be used, especially in applied mathematics. The logician Adolf Fraenkel gave what was no doubt the widely accepted view when he stated, "The infinitely small is only to be understood as a manner of speaking based on the limit concept, hence a *potential* infinite; it is a matter of variable … [positive] numbers or quantities that can ultimately decrease below any arbitrarily small positive value. A fixed [positive] number different from zero that can serve as a lower bound to all finite positive values is not possible" (1928, p. 114, my translation, emphasis in original). In 1960 Fraenkel's one-time student Abraham Robinson showed how to obtain just such a "fixed number" and thereby vindicated the discredited infinitesimal methods.

The benefits of the free use of infinitesimal methods were amply demonstrated by the success of Gottfried Wilhelm Leibniz's version of the differential and integral calculus and the continued use of these methods by the Bernoulis and especially by Leonhard Euler. Working mathematicians had no difficulty in knowing just which properties of ordinary numbers infinitesimals could be assumed to possess and just when it was legitimate to equate such quantities to zero. But the lack of any clear justification for these methods provided an opening for scathing attacks such as that of George Berkeley. The need for rigorous methods was felt by mathematicians them-

selves and eventually supplied (Edwards 1979, Robinson 1974).

Robinson's key insight was that the methods of model theory could be used to construct a powerful rigorous theory of infinitesimals. Thus, for example, we may consider a first-order language in which a constant symbol is provided as a "name" for each real number, a function symbol is provided as a "name" for each real-valued function defined on the real numbers, and the only relation symbols are = and <. Let T be the set of all true sentences of this language when each symbol is understood to have its intended interpretation. Let δ be a new constant symbol, and let W consist of the sentences of T together with the infinite set of sentences:

$$\delta > 0$$

$$\delta < 1, \delta < \tfrac{1}{2}, \delta < \tfrac{1}{3}, \delta < \tfrac{1}{4}, \ldots$$

Since any finite subset of W can be satisfied in the ordinary real numbers by interpreting δ as a sufficiently small positive number, the compactness theorem for first-order logic guarantees that W has a model. But in that model, the element serving to interpret δ must be positive and less than every positive real number (i.e., infinitesimal). The structure with which we began of real numbers and real-valued functions can readily be embedded in the new model. Thus if r is a real number and c_r is the constant of the language that names r, we may regard the element of the new model that serves to interpret c_r as simply r itself. Functions can be embedded in the same way. One speaks of the new model as an enlargement.

Moreover, because $T \subseteq W$, all true statements about the real numbers that can be expressed in our language are also true in this enlargement. A false statement about the real numbers is likewise false in the enlargement: If the statement S is false, then $\neg S$ is a true statement about the reals and hence is also true in the enlargement. It is this transfer principle, the fact that statements are true about the real numbers if and only if they are true in the new enlarged structure, that makes precise just when an assertion about ordinary numbers can be extended to apply to infinitesimals as well.

The enlargement will contain infinitely large as well as infinitesimal elements. This is readily seen by applying the transfer principle to the statement that every nonzero real number has a reciprocal. One may even speak of infinite integers; their existence follows on applying the transfer principle to the statement that for any given real number there is a positive integer that exceeds it.

The basic facts of real analysis can be established on this basis using modes of argument that would earlier have been quite correctly regarded as illegitimate. For example, the basic theorem that a continuous function on a closed interval assumes a maximum value can be proved by dividing the interval into infinitely many subintervals, each of infinitesimal length, and selecting an endpoint of such a subinterval at which the function's value is greatest (Davis 1977, Robinson 1974). By beginning with a more extensive language, it is possible to apply infinitesimal methods to branches of mathematics requiring a more substantial set-theoretic basis (e.g., topology, functional analysis, probability theory). It has even proved possible to use these "nonstandard" methods to settle certain open questions in mathematics.

For those with qualms concerning nonconstructive methods in mathematics, these infinitesimal methods are bound to seem unsatisfactory. Because the underlying language is built on an uncountable "alphabet," the use of the compactness theorem hides an application of some form of the axiom of choice. This in turn is reflected in a basic indeterminacy; we can establish the existence of enlargements but cannot specify any particular enlargement. Robinson himself has emphasized that although nonstandard analysis "appears to affirm the existence of all sorts of infinitary entities," one always has the option of taking the "formalist point of view" from which "we may consider that what we have done is to introduce *new deductive procedures* rather than new mathematical entities" (1974, p. 282, emphasis in original).

See also Berkeley, George; Leibniz, Gottfried Wilhelm; Logic, History of; Model Theory; Number.

Bibliography

Davis, M. *Applied Nonstandard Analysis.* New York: Wiley, 1977.

Edwards, C. H., Jr. *The Historical Development of the Calculus.* New York: Springer-Verlag, 1979.

Fraenkel, A. *Einleitung in die Mengenlehre,* 3rd ed. Berlin: Springer, 1928; reprint, New York: Dover, 1946.

Robinson, A. *Nonstandard Analysis,* 2nd ed. Amsterdam, 1974.

Martin Davis (1996)

INFINITY, AXIOM OF

See *Russell, Bertrand Arthur William,* section on logic and mathematics

INFINITY IN MATHEMATICS AND LOGIC

The notion of infinity, and the problems, both philosophical and mathematical, that arise from it have been a central concern for over two millennia. Any serious thought about the nature of space, time, God (or gods), mathematics, and motion quickly leads to more general concerns regarding the notion, or notions, of infinity intimately tied up with such issues. As a result, it is unsurprising that philosophers throughout history have thought deeply about what infinity is, whether the notion is coherent, whether there are infinite entities (or infinitely many entities), and how we can know about such entities if they exist.

This entry focuses on two aspects of the infinite. The first is infinite divisibility, the idea that an object can, in some sense (and perhaps only ideally), be divided into an infinite collection of smaller and smaller parts. The puzzles that arise from such division are central both to philosophical thinking about notions such as part and whole and to the mathematical analysis of lines, surfaces, and other continuous objects. The second aspect to be addressed is already implicit in the first—the idea that there can be infinitely large collections at all. Much of the history of mathematics and philosophy can be seen as an (often indirect) inquiry into the coherence of such collections and how they differ from finite collections.

As a result, this entry will for the most part ignore other interesting, but less central, issues within the literature on infinity, including infinitesimals, mereological theories containing gunk, unrestrictedly general quantification, nonstandard set theories, and nonmathematical uses of the term "infinite" (e.g., theological understandings of the infinite). The discussion below, then, is not meant to be a comprehensive survey of all aspects of infinity (or even all aspects of this notion as it appears within mathematics, logic, and metaphysics), but is instead intended to provide a basic understanding of two important themes underlying hundreds of years of thought on the topic.

INFINITE DIVISIBILITY: ARISTOTLE AND ZENO

Given the long pedigree of thought regarding infinity, it seems apropos to begin at (or near) the beginning, with the ancient Greeks. While mathematics in general, and geometry in particular, was central to Plato's philosophy, he has little to say regarding the nature of the infinite. His few comments on the topic occur in the *Philebus*, where he equates the infinite with the unlimited, unbounded, excessive, and indefinite.

Identifying the infinite with such notions is, in hindsight, less helpful than one might hope: The surface of a sphere, which in a certain sense has no boundaries or limits, nevertheless has finite area. Extremely large collections can be thought of as excessive yet finite. Indefiniteness has, recently, been associated more with vague phenomenon (such as the boundary between colors) than with the infinite. Thus, Plato's discussion, while interesting, fails to provide clear criteria for distinguishing between finite and infinite collections, relying rather on the idea that we can tell the difference when we see it.

Plato's star pupil, Aristotle, follows his teacher in neglecting to provide a clear definition of the infinite (and in equating it with suspect notions such as unboundedness). Even so, he made an influential contribution by distinguishing between two different types of infinity, or two different ways of conceiving infinite collections. This contribution was due to Aristotle's need to respond to Zeno of Elea's paradoxes.

Zeno presented four paradoxes that, through clever uses of infinity, demonstrated (so it was claimed) that motion was impossible. Here we consider only two: the paradox of Achilles and the paradox of the runner (an adequate response to either can likely be generalized to the others).

The paradox of Achilles is perhaps the best known of Zeno's puzzles. Swift Achilles is to run a race against a tortoise, and the tortoise is given a head start. Zeno argues that, no matter how fast Achilles runs, he can never overtake the tortoise. First Achilles must reach the point at which the tortoise started, call it P_1. By the time he does so, however, the tortoise will have traveled some short distance further, to a point we can call P_2. So Achilles's next task is to run from P_1 to P_2. By the time he achieves this, the tortoise will have traveled a bit further, to P_3. So Achilles's next task is to run to P_3. But by then the tortoise will have reached P_4, and so on.

Thus, according to Zeno, Achilles can never pass the tortoise and win the race, because, no matter how fast he runs, each time he reaches a point where the tortoise was, the tortoise will have moved a bit farther on. Stating the conclusion more carefully, Zeno's argument does not (and was most likely not intended to) show that Achilles cannot overtake the tortoise; rather, it demonstrates that there is a conceptual puzzle regarding how he does so.

Zeno's paradox of the runner is similar, slightly less well known, but mathematically a bit more elegant. Imagine a runner who must run from point 0 to point 1. Before reaching 1, he must reach the midpoint between 0 and 1 (i.e., point 1/2). Once he has reached 1/2, however he must, before he can reach 1, run to the midpoint between 1/2 and 1, (i.e., 3/4). Then he must run to the midpoint between 3/4 and 1 (i.e., 7/8), and then run to 15/16, and 31/32, and so on.

0 1/2 3/4 7/8 ... 1

Zeno concluded that, in traveling from 0 to 1, the runner traverses an infinite number of distinct distances.

It is worth noting that this paradox depends on a by now well-known mathematical fact: Some infinitely long lists of numbers (or *infinite series*) have a finite sum. In particular, the construction of the paradox demonstrates an at least implicit awareness that:

$$1 = 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \dots$$

or, in more modern notation:

$$\sum_{i=1}^{\infty} 1/2^n = 1$$

We can provide an intuitively compelling (although not mathematically rigorous) argument demonstrating this as follows. Set the infinite sum equal to x:

$$x = 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \dots$$

Multiply both sides by 2:

$$2x = 1 + 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \dots$$

Subtract the first line from the second, and we obtain the desired result:

$$x = 1$$

As Aristotle realized, there are two puzzles lurking within Zeno's paradoxes, and only one of them is difficult. First, the worry might be that Zeno's arguments suggest that we can accomplish infinitely many tasks in a finite amount of time (for example, traveling through the infinite sequence of distances 0 to 1/2, 1/2 to 1/4, 1/4 to 1/8 … or inhabiting the infinite sequence of points 1/2, 1/4, 1/8 …). This, however, as Aristotle noted, is a misleading way of characterizing the situation, because if we can divide distances in the way envisioned by Zeno, then we can divide time in the same way (so our minute of time can be divided into the first half of a minute, then the next quarter of a minute, then the next eighth of a minute, …). Rather, according to Aristotle, the puzzle concerns the idea that we can ever complete infinitely many tasks (no matter how they are described). In other words, is the infinite division of either space or time that Zeno envisioned legitimate?

The answer for Aristotle had to be "no," because, as already noted, he equated the infinite with the unbounded or unlimited. As a result it should be impossible to complete infinitely many tasks, because it is impossible to reach the bound or limit of something that, by its nature, is unlimited. Aristotle had little choice but to conclude that there is a mistake lurking within Zeno's argument, and it is in his explanation of this mistake that his potentially/actually infinite distinction makes its appearance.

Zeno's paradoxes result from the fact that line segments are (or at least seem to be) infinitely divisible. Consider an arbitrary (finite) line segment. We can easily divide the line into two halves, producing two parts. Additional divisions, of course, are also possible. The crucial question now arises: How many distinct parts does the line segment contain? In some sense, at least, the correct answer is infinitely many, since for any part we can further subdivide it into two halves, obtaining two more parts. Thus, for any finite number of parts we divide the line into, we can further subdivide those segments to obtain more parts.

Aristotle distinguished this sort of unboundedness—the potentially infinite—from actually infinite collections. On the one hand, a collection is potentially infinite if we can continue to add to it without limit. On the other hand, the actual infinite is, for Aristotle, a completed totality, that is, an unbounded collection that is nevertheless present all at once. In considering the division of a line segment Aristotle writes that "It is always possible to think of a larger number: for the number of times a magnitude can be bisected is infinite. Hence the infinite is potential, never actual; the number of parts that can be taken always surpasses any assigned number" (Physics 207b8).

In understanding Aristotle's potential/actual distinction it is useful to distinguish between three distinct theses:

(1) Any part of a line can be divided into distinct subparts.

(2) A line contains infinitely many distinct parts.

(3) A line contains infinitely small parts.

The first claim, according to Aristotle (and most thinkers since) is undeniable. The third, which asserts the existence of so-called infinitesimals, would be of great importance in the development of the calculus during the seventeenth century. The second claim is, for our purposes, the crucial one. According to Aristotle, claim (2) is ambiguous, having both a true and a false reading.

Aristotle understood the first of these three claims as something such as: given a time t_1, any part of a line that exists at t_1 can be divided into distinct subparts at some future time t_2. Once we recognize the temporal ingredient of claim (1), two distinct readings of (2) are then apparent:

> (2a) For any number n, there is a time t such that a line has been (or can be) divided into (at least) n parts at t.
>
> (2b) There is a time t such that a line has infinitely many parts at time t.

Although (2b) implies (2a), Aristotle argued, in effect, that (2a) does not imply (2b): Imagine that we have a line segment in front of us, and each hour we divide up each of its parts into two subparts. Then (2a) is true (assuming we live forever and never forget to carry out the divisions) yet (2b) fails, because there is never a particular time at which we have finished dividing. The collection of parts of a line segment is potentially infinite if (2a) is true, and is actually infinite if (2b) holds.

It is important to note that Aristotle's distinction contains both a constructive aspect and a temporal aspect. First, lines (and other objects) are not presented to us already divided into their parts, rather, the division of an object into its components is somehow a construction that we perform on it. Second, these constructions cannot be carried out all at once, but must be carried out one by one in time. This aspect of Aristotle's view is what prevents the potential infinite from collapsing into the actual infinite, because it allows us to distinguish between a series of ever increasing finite collections spread out through time and an infinite collection existing at a particular time.

Aristotle applied the potential versus actual distinction, not just to parts of line segments, but to the natural (counting) numbers 0, 1, 2, 3, … as well (we include 0 and 1 here because they are now considered to be the first two natural numbers, although Aristotle would not have recognized 0, and perhaps not even 1, as a number at all). The natural numbers, for Aristotle, are potentially infinite, because for any number we have counted to, we can always count one (or ten, or one hundred) numbers further. Nevertheless, we will never reach a time when we will have counted out all the numbers.

Aristotle's final contribution lies in his insistence that not only some but all infinities are potential, not actual. Thus, there are no infinitely long line segments, only a potential series of ever longer finite line segments, and no infinitely large collections, only series of larger and larger finite ones.

The denial of absolute infinity provided Aristotle with a solution to Zeno's puzzles: If space and time are only potentially infinite, then the sort of division of space and time necessary for Zeno's construction is illicit. In the paradox of the runner, the runner does not pass through infinitely many different distances (nor does he occupy infinitely many different points, because Aristotle thought that points only exist if a line has in fact been divided into two parts that meet at that point). Rather, given any particular time, there are only finitely many parts into which the distance between 0 and 1 will have been divided into. Of course, at a later time we might further subdivide the distance into a larger (but still finite) collection of parts. It does not follow, however, that we have therefore traveled over infinitely many distances, because the parts of the path from 0 to 1 are merely potentially infinite.

For approximately two thousand years Aristotle's view remained unchallenged—philosophers and mathematicians both (for the most part) denied the existence of actual, completed infinities, arguing that the notion of potential infinity could, within both philosophy and mathematics, fulfill any role that the infinite might need to play.

RIGOR AT LAST: DEDEKIND INFINITY

During the development and rigorization of the calculus (a long, torturous period stretching from the beginning of the seventeenth century to the end of the nineteenth century) it became evident that the further development of classical mathematics required actual infinities. (The details need not detain us here. Suffice it to say at least some infinite collections needed to be viewed as complete, because it became necessary to study arbitrary subcollections of these infinite structures). If Aristotle and his followers were correct, however, and the only viable understanding of the infinite was the potential one, then classical mathematics was facing a crisis.

Fortunately, a number of philosophers and mathematicians stepped into the breach. They were faced with two main mathematical tasks and one philosophical one:

first, to provide a rigorous definition of "infinite"; second, to provide a mathematical theory of the infinite that clarified how the behavior of infinite collections differed from the well-known behavior of finite collections; and third, and perhaps most importantly, to provide a philosophical account of the actual infinite that defended its intelligibility. The first two tasks were successfully carried out during the nineteenth century, by Richard Dedekind (1901) and Georg Cantor (1955) respectively.

Before considering definitions of infinity, it is useful to introduce some terminology. The existence of an actually infinite collection involves the idea that such a collection can be presented all at once, as a completed totality—that is, in some sense, as a single thing. Such totalities, considered as single objects, are called *sets*. The central idea behind *set theory* is that, given any collection of objects (or almost any, see the discussion of *Russell's paradox* below), there exists another object—the set containing exactly the original objects. Thus, if we start out with three distinct persons, Alan, Bob, and Carl, we obtain the following sets:

{Alan}, {Bob}, {Carl}, {Alan, Bob}, {Alan, Carl},
{Bob, Carl}, {Alan, Bob, Carl}

Note that the one-membered set {Alan}, called the *singleton* of Alan, is not the same thing as Alan himself, because {Alan} is a set, yet Alan, who is a person, is not. At this point it should be noted, as well, that the collection containing no objects, the so-called *empty set* { } or ∅, although somewhat puzzling (how can there exist a collection formed out of nothing?) is nevertheless accepted as a set, and thus an object, in most accounts of set theory.

Sets are objects, so there is nothing to prevent us from forming collections of these (i.e., sets of sets), obtaining, for example:

{∅, {Alan}, {Bob}, {Alan, Bob}}

This expression names a single object, a set collecting together four other sets.

There are two important relations that can hold between sets and other objects: *membership* and *subsethood*. The members of a set are those objects that were collected together to form the set. Using "$x \in y$" to express the claim that x is a member of y (and "∉" for nonmembership), we have:

Alan ∈ {Alan}

{Alan}∈ {∅, {Alan}, {Bob}, {Alan, Bob}}

and:

{Alan} ∉ {Alan}

Alan ∉ {∅, {Alan}, {Bob}, {Alan, Bob}}

Our second notion, subsethood, can be defined in terms of membership. Given two sets A and B, A is a subset of B if and only if every member of A is a member of B. This implies that every set is a subset of itself, and, using "$x \subseteq y$" to express the claim that x is a subset of y, we have:

{Alan, Bob} ⊆ {Alan, Bob, Carl}

{Alan, Bob} ⊈ {∅, {Alan}, {Bob}, {Alan, Bob}}

Finally, a *function f* from a set A to a set B, symbolized by:

f: A → B

is a mapping that assigns to each member of A exactly one member of B. For example, there is a function:

f_1: {Alan, Bob} → {Alan, Bob, Carl}

which maps Alan to Bob and Bob to Carl (f_1 does not map anyone to Alan). We can express this symbolically as:

f_1(Alan) = Bob

f_1(Bob) = Carl

Another example is:

f_2: {Alan, Bob, Carl} → {Alan, Bob}

mapping Alan to Bob, Bob to Alan, and Carl to Bob:

f_2 (Alan) = Bob

f_2 (Bob) = Alan

f_2 (Carl) = Bob

There are two conditions on functions that will be crucial in what follows. First, a function is from A to B is *injective* (or *one-to-one*) if and only if no two distinct members of A get mapped onto the same member of B. In the examples above, f_1 is injective, but f_2 is not, because both Alan and Carl both get mapped onto Bob. Second, a function from A to B is *surjective* (or *onto*) if and only if every member of B gets mapped onto by some member of A. f_2 above is surjective, whereas f_1 is not, because neither of Alan or Bob gets mapped onto Alan. A function that is both one-to-one and onto is *bijective*.

We can now consider possible definitions of infinite. One obvious approach suggests itself: To define finite in terms of counting, that is, a set is finite if and only if we can assign 0 the first member of the set, 1 to the second, 2 to the third, … and at some point we reach the last member of the set, to which we assign some natural number. Using the terminology introduced above:

A set A is finite if, and only if, there is a natural number n and a bijective function:

f: A → {0, 1, 2, …, n–1}

Thus, on this definition a set is finite if there is a number n and a bijective function from A to the set of natural numbers less than n. A set is infinite if there is no such function.

For example, the set {Alan, Bob, Carl} is finite because:

f_3: {Alan, Bob, Carl} → {0, 1, 2}

f_3 (Alan) = 0

f_3 (Bob) = 1

f_3 (Carl) = 2

is bijective, whereas the set of natural numbers {0, 1, 2, 3, …} is, on this definition, infinite.

While this definition provides the desired results, there is a problem. The definition works by determining whether a set is finite or infinite in terms of whether it can be mapped bijectively onto the natural numbers less than n, for some n. The reason this works is that we know, for any number n, that the set of numbers less than n is finite. What could be more obvious? The problem, however, is that the definition lacks the sort of generality required in both mathematics and philosophy. What we want is a criterion that tells us which sets are infinite and which sets are finite. What we have is a criterion that tells us this, assuming that we already know that certain sets of natural numbers are finite.

An appropriately general definition of infinite set was produced by Richard Dedekind in 1888. The definition is based on an insight into infinite collections that traces back to Galileo Galilei (and probably to the ancient Greeks). Galileo noticed that there is a bijective mapping between the natural numbers and the even natural numbers:

f_4: {0, 1, 2, 3, …} → {0, 2, 4, 6, …}

provided by mapping each number onto its double:

$f_4(0) = 0$

$f_4(1) = 2$

$f_4(2) = 4$

$f_4(3) = 6$

etc.

Galileo argued that this provided further evidence against the notion of actual infinity, because if both the natural numbers and the even numbers were completed infinities, then the latter would be a part of the former. The existence of the bijective mapping, however, seemed to imply that there was just as much "stuff" in each infinite collection, violating the (at the time sacrosanct) dictum that the part must be less than the whole.

Dedekind, however, embraced the puzzling nature of this discovery, and proposed an alternative definition that does not rely on prior knowledge that certain sets of natural numbers are finite. Instead, a set is infinite, according to Dedekind, if it can be mapped bijectively onto a proper subset of itself (a subset of a set A is *proper* if it is not identical to A itself):

A set A is *Dedekind infinite* if, and only if, there is a function:

f: A → A

(i.e., a function from a set A to itself) that is injective but not surjective.

A set is (*Dedekind*-) *finite* if there is no such mapping. With this definition in place, actual, completed infinities are at least partially vindicated, insofar as mathematicians and philosophers now have a general, formal criterion for distinguishing between finite and infinite sets.

At roughly the same time Bernard Bolzano (2004) produced a competing definition of infinite set: A set A is infinite if there is a nonterminating series of sets B_1, B_2, B_3, … such that each set in the series is a subset of A, each set in the series is a subset of all sets that follow it in the series, and no set in the series is a subset of any set in the series that precedes it. In other words, a set is infinite if it contains a series of subsets that get "bigger" and "bigger" without end. While Bolzano's definition is especially interesting in light of its obvious connection to Aristotle's ideas regarding the infinite, there are legitimate worries regarding whether the definition is of any help, since "nonterminating series" seems synonymous with "infinite series." Thus, this definition, like the one we began with (but unlike Dedekind's), assumes an understanding of the concept that we are trying to define.

With a rigorous definition of infinite set in hand, the next step in securing the notion of actually infinite from Aristotelian worries would be a well worked out theory of the existence and "behavior" of infinite sets. Much of this theory was worked out by Cantor, and will be the subject of the next section. Before examining Cantor's work, however, it should be noted that, even after the work of

Dedekind and Cantor, not everyone accepted that infinite totalities existed. A number of influential thinkers denied that Dedekind's definition, interesting or not, applied to any complete collections, returning to Aristotle's distinction between potential and actual infinity (although most took Aristotle's idea of the potentially infinite as being potentially extendable in time to be little more than a metaphor). Post-Dedekind/Cantor views of this sort include Kronecker's finitism, Brouwer's intuitionism, Weyl's constructivism, and the later Wittgenstein.

SIZES OF INFINITY: CANTOR'S INFINITE NUMBERS

Georg Cantor began with the idea that infinite sets, like finite sets, have a corresponding number. Such numbers, which measure how many things are contained in a set, are called *cardinal numbers* (or *cardinalities*). We represent the cardinal number of a set A as Card(A).

The idea that infinite collections have cardinal numbers, or, equivalently, that there are "infinite" numbers coming after the natural numbers 0, 1, 2, 3 … , was rejected prior to the nineteenth century on the grounds that infinite numbers required actual infinite totalities, and this in turn (so the story went) implied a contradiction. For example, the great seventeenth-century philosopher, logician, and mathematician Gottfried Leibniz wrote that "I proved beyond any doubt that the number or multitude of all numbers implies a contradiction, if taken as a unitary whole. I think that the same is true of the largest number" (1849, p. 535). One notable aspect of this line of thought is the assumption that the existence of infinite numbers implies the existence of a largest infinite number, or a number of all numbers.

This assumption traces to a rather basic intuition: won't any two "unbounded" or "unlimited" collections, whether complete or potential, contain the same number of elements (or members), because both just keep going and going? Cantor's great contribution to both mathematics and philosophy was his discovery that the answer to this question is "no."

Cantor begins, not by asking which number should be attached to a particular set, but by asking what criteria can be given for deciding when two sets have the same number (independently of which particular number or numbers are involved). To illustrate the difference between the two approaches, consider the following situation: you have two baskets of fruit, one containing apples, the other containing oranges, and you need to determine whether the number of apples in the first basket is the same as the number of oranges in the second basket.

There are two strategies. First, you could count the apples, count the oranges, and then determine if the two numbers are the same. This strategy corresponds to the first approach, where we first assign particular numbers to sets and then compare them. On the second approach, you repeatedly remove one apple from the first basket and one orange from the second, until you run out of either apples or oranges. When you reach a stage where you cannot remove a pair consisting of one apple and one orange, there are three options: Either there are apples left in the first basket, in which case there are more apples than oranges; or there are oranges left in the second basket, in which case there are more oranges than apples; or both baskets are empty, in which case the number of apples is the same as the number of oranges. (The second strategy, while perhaps more efficient, is less informative, because you do not, in the end, know how many apples or oranges were in the baskets, but only whether or not there was more of one than the other. A similar difficulty arises with regard to the *continuum hypothesis*, which will be discussed below.) The second "pairing" strategy amounts to nothing more or less than attempting to construct a bijective mapping between the set of apples and the set of oranges.

Cantor's insight was in noticing that, although extending the first strategy (counting and comparing numbers obtained) to infinite sets is not viable until we have a well-worked-out theory of infinite cardinal number (which was exactly what he was attempting to formulate), the second strategy can be applied to infinite sets (almost) as easily as to finite ones.

The following principle, a version of what has come to be called *Hume's Principle* and which we can call *Hume's Principle for Sets*, sums up Cantor's approach:

> *HPS*: For any sets A and B:
>
> $$\text{Card}(A) = \text{Card}(B)$$
>
> if, and only if, there is a bijective function $f\colon A \rightarrow B$.

We supplement this with the following definition of "less than" for cardinal numbers:

> $Def_<$: For any sets A and B:
>
> $$\text{Card}(A) < \text{Card}(B)$$
>
> if, and only if, there is an injective function $f\colon A \rightarrow B$ but no surjective function $g\colon A \rightarrow B$.

This definition agrees with the intuitive one for finite sets. For example, we can verify that Card({Alan, Bob}) <

Card({Alan, Bob, Carl}) (i.e., two is less than three) by noting that there is an injective mapping from the first to the second (f_1 above provides one of the six possible injective mappings) but there is no surjective mapping, because one of the three members of the latter set will always be "missed."

Consider Galileo's puzzle again. Galileo showed that there was a bijective mapping between the set of natural numbers {0, 1, 2, 3 …} and the set of even natural numbers {0, 2, 4, 6 …}. Thus, Card({0, 1, 2, 3 …}) = Card({0, 2, 4, 6 …}). More surprisingly, Cantor also proved that the set of rational numbers (i.e., all numbers that can be written as a fraction a/b where a and b are both natural numbers) has the same cardinal number as the set of natural numbers. Cantor called this number $\aleph_0$, and a set is *countable* if, and only if, it is either finite or its cardinal number is $\aleph_0$ (reflecting the fact that such sets are the same size as some set of natural, or "counting" numbers). Dedekind's definition of infinite set implies that $\aleph_0$ is the smallest infinite cardinal number; in other words, there is no infinite set that cannot be mapped surjectively onto the natural numbers.

Cantor used $\aleph_1$, $\aleph_2$, $\aleph_3$ … as names for the second, third, fourth … infinite numbers, and sets with these cardinal numbers are called *uncountable*. Providing names for infinite cardinal numbers is one thing, however; showing that there are sets that have those numbers is something else. Thus, Cantor's next task was to demonstrate that there were infinite sets that do not receive $\aleph_0$ as their number, that is, that there are infinite sets "bigger than" the set of natural numbers. He did this in two ways.

The first strategy depends on a second species of number, the *ordinal numbers*. Whereas cardinal numbers measure how many members a set has, ordinal numbers are a measure of particular orderings on that set (compare one, two, three … with first, second, third …). In other words, ordinal numbers attach, not to sets by themselves, but to a set plus an ordering on that set, and the same (infinite) set can correspond to different ordinal numbers if we consider different orderings on it.

More carefully, an ordinal number attaches to a pair consisting of a set A and an ordering ≤ on A. If a set/ordering pair ⟨A, ≤⟩ is to receive an ordinal number, then the relation ≤ must be a well-ordering (we call such pairs well-ordered sets, and represent the ordinal number of a well-ordered set ⟨A, ≤⟩ as Ord(A, ≤)).

We begin with the notion of a *totally ordered set*. An ordering ≤ is a total ordering on A (i.e., ⟨A, ≤⟩ is a totally ordered set) if, and only if, it satisfies the following three conditions (here and below I assume familiarity with the notation of first-order logic):

Antisymmetry: $(\forall x)(\forall y)((x \leq y \wedge y \leq x) \rightarrow x = y)$

Transitivity: $(\forall x)(\forall y)(\forall z)((x \leq y \wedge y \leq z) \rightarrow x \leq z)$

Comparability: $(\forall x)(\forall y)(x \leq y \vee y \leq x)$

More intuitively, a relation on a set A is a total ordering if, and only if: (i) given two distinct objects in A, it cannot be the case that the first is less than or equal to the second and the second is less than or equal to the first (if so, then they would be the same object); (ii) for any three objects in A, if the first is less than or equal to the second, and the second is less than or equal to the third, then the first is less than or equal to the third; and (iii) given any two objects in A, either the first is less than or equal to the second, or the second is less than or equal to the first (this implies that any object in the ordering is less than or equal to itself). Two examples of total orderings are the natural numbers {0, 1, 2, 3 …} on their standard ordering (i.e., $0 \leq 1 \leq 2 \leq 3 \leq \ldots$), and the integers {… –3, –2, –1, 0, 1, 2, 3 …} on their standard ordering (i.e., $\ldots \leq -3 \leq -2 \leq -1 \leq 0 \leq 1 \leq 2 \leq 3 \leq \ldots$).

A totally ordered set ⟨A, ≤⟩ is a *well-ordered set* if and only if the following additional condition holds:

Well-foundedness: $(\forall B \subseteq A)(\exists x \in B)(\forall y \in B)(x \leq y)$

Loosely put, if ≤ is a well ordering on a set A, then there is no "infinitely descending chain" in ⟨A, ≤⟩, that is, there is no infinite sequence $x_1 \geq x_2 \geq x_3 \geq x_4 \geq \ldots$ (although there can be infinitely ascending chains $x_1 \leq x_2 \leq x_3 \leq x_4 \leq \ldots$). The standard ordering on the natural numbers is a well ordering, whereas the standard ordering on the integers is not, because the negative integers form an infinitely descending chain.

Finally, we need the notion of an *order-preserving* function from one ordered set to another:

Given two ordered sets $\langle A, \leq_1\rangle$ and $\langle B, \leq_2\rangle$, and a function $f: A \rightarrow B$, f is *order preserving* if and only if, for any two members x and y of A, $x \leq_1 y$ if, and only if $f(x) \leq_2 f(y)$.

In other words, if we take two members of A, where the first is, according to the ordering on A, less than or equal to the second, then an order preserving function will map the first object onto a member of B that is, according to the ordering on B, less than or equal to the member of B onto which the second object is mapped.

We can now provide an analogue of *Hume's Principle* for ordinals, which we can call the *Order Type Principle For Sets* (OTP):

OTP: For any well-ordered sets $\langle A, \leq_1\rangle$ and $\langle B, \leq_2\rangle$:

$$\mathrm{Ord}(A, \leq_1) = \mathrm{Ord}(B, \leq_2)$$

if, and only if, there is an order preserving bijective function f: A → B.

and provide an analogous definition of *less than* for ordinal numbers:

Def: For any well-ordered sets $\langle A, \leq_1\rangle$ and $\langle B, \leq_2\rangle$:

$$\mathrm{Ord}(A, \leq_1) < \mathrm{Ord}(B, \leq_2)$$

if, and only if, there is an order preserving injective function *f*: A → B, but no order preserving surjective function *f*: A → B.

Given this definition, Cantor was able to prove that the ordinal numbers, ordered by < as defined above, are themselves well-ordered.

Cantor also proved that the ordinal number of the natural numbers on their standard ordering is the smallest infinite ordinal number, which he called ω (An ordinal is infinite if it is the ordinal of some well-ordered set $\langle A, \leq\rangle$ where A is infinite, an ordinal is countable if it is the ordinal of some well-ordered set $\langle A, \leq\rangle$ where A is countable, and so on).

The standard ordering on the natural numbers, however, is not the only way to order them. Instead, we might move zero from the beginning to the end, so that (on this ordering) 1 is the least natural number, and zero is greater than any other natural number (i.e., 1, 2, 3, 4, 5 … 0). The ordinal corresponding to this ordering is greater than ω, and there is no ordinal less than this one but greater than ω. Thus, this ordinal, called $\omega + 1$ (because it consists of a "copy" of ω followed by a single element) is the second infinite ordinal. Continuing in this way, we can then move 1 from the beginning to the end (i.e., 2, 3, 4, 5 … 0, 1) obtaining $\omega + 2$, and then $\omega + 3$, $\omega + 4$, … We can then consider the ordering consisting of all of the even natural numbers followed by all of the odd natural numbers (i.e., 0, 2, 4, 6, … 1, 3, 5, 7 …), whose ordinal number is $\omega + \omega$, and so on.

Because there are many different ways of well-ordering the natural numbers, resulting in different countable ordinals, a natural question to ask is how many different countable ordinals there are (or, equivalently, how many different types of well-ordering can be constructed from the natural numbers). Cantor proved that the cardinal number of the set of natural numbers (what Cantor called the *first number class*) is less than the cardinal number of the set of countable ordinals (the *second number class*), and that there is no set whose cardinal number is greater than the former and less than the latter. In other words, $\aleph_1$, the second infinite cardinal number, is the cardinal number of the set of countable ordinals.

We can then go on to ask about the cardinal number of the set of ordinals of well-ordered sets of size $\aleph_1$ (i.e., how many different types of well-ordering are there on a set of objects of size $\aleph_1$?). The answer is $\aleph_2$. How many ordinals of size $\aleph_2$? Surprise, its $\aleph_3$! And so on. In this way, Cantor managed to use ordinal numbers to prove the existence of a series of sets of cardinality $\aleph_1$, $\aleph_2$, $\aleph_3$, and even $\aleph_\omega$ (the first infinite cardinal number that is larger than infinitely many other infinite cardinal numbers). In fact, for any ordinal number α, there is a set whose cardinal number is $\aleph_\alpha$.

Cantor had a second means by which to prove that there are uncountably infinite sets, a method that relies on the notion of *powerset*. The powerset of a set A (or $\wp(A)$) is the set that contains exactly the subsets of A. For example:

$\wp(\{\text{Alan, Bob}\}) = \{\emptyset, \{\text{Alan}\}, \{\text{Bob}\}, \{\text{Alan, Bob}\}\}$.

Notice that whereas {Alan, Bob} has two members, its powerset has $2^2 = 4$ members. This holds more generally: if κ is the cardinal number of a set A, then 2^κ is the cardinal number of $\wp(A)$ (even for infinite sets).

Cantor proved that, for any set A, $\mathrm{Card}(A) < \mathrm{Card}(\wp(A))$. The method of proof is known as the method of *diagonalization*, and generalizations of it have become immensely important in mathematics. For more technically interested readers, a proof follows (readers not interested in purely mathematical matters may skip the next paragraph).

We can provide an injective function from A to $\wp(A)$ by mapping each member of A onto its singleton (because, for each member of A, its singleton is a subset of A and thus a member of $\wp(A)$). No mapping from A to $\wp(A)$ can be surjective, however. Let f: $A \to \wp(A)$ be an arbitrary function from A to its powerset. Define the set B as follows: for any object x, $x \in B$ if, and only if, $x \in A$ and $x \notin f(x)$. Assume, for *reductio*, that there is $c \in A$ such that $f(c) = B$. By the definition of B, we have $c \in B$ if, and only if, $c \in A$ and $c \notin f(c)$, which implies that $c \in B$ if, and only if, $c \notin B$. Contradiction, so f cannot be surjective.

Thus, the cardinality of any set is less than the cardinality of its powerset, and, as a result, for any cardinal number $\aleph_\kappa$, $\aleph_\kappa < 2^{\aleph_\kappa}$. We might wonder why anyone would make all this fuss over powersets, however. Haven't we already seen that we can construct larger and larger sets, and thus obtain larger and larger cardinal numbers, using the ordinal numbers?

There are two aspects of Cantor's diagonalization result that are noteworthy. The first concerns the import of cardinality results for ordinary mathematics. We might wonder why everyday mathematicians (and ordinary nonmathematicians) should worry about different "sizes" of infinity, because neither ordinary folk nor most professional mathematicians run across infinite sets of ordinal numbers in their everyday business. Cantor's result, however, connects the theory of cardinal numbers to more intuitive, everyday mathematical concerns, because the cardinal number of the set of real numbers (i.e., the cardinal number of the set containing all the numbers on the continuous number line) is $2^{\aleph_0}$. As a result, there are more points on a continuous line than there are natural numbers! Because real numbers and lines are commonly used within basic mathematics, this result provides a direct connection between Cantor's theory and the practice of everyday measurement and mathematics.

The second reason that Cantor's result regarding powersets is interesting is that it introduced one of the great unsolved problems of mathematics. One might wonder, because we have the notation $\aleph_0$, $\aleph_1$, $\aleph_2$, … $\aleph_\omega$ … for the series of cardinal numbers, why we use a different notation ($2^{\aleph_0}$) for the cardinal number of the powerset of the natural numbers. The reason is simple: Although we know that $2^{\aleph_0}$ is larger than $\aleph_0$, we do not know how much larger. In particular, we do not know whether Cantor's *continuum hypothesis*:

$$2^{\aleph_0} = \aleph_1$$

is true or false. The truth of the continuum hypothesis amounts to the claim that there are no sets that are strictly larger than the set of natural numbers yet smaller than the set of real numbers. (Our account of cardinal numbers failed to settle this problem because our bijection strategy tells us whether one number is larger without necessarily telling us how much larger.)

More generally, for any ordinal number α, we can ask whether:

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}$$

is true. The claim that the above is true for all ordinals α is know as the *generalized continuum hypothesis*.

Despite great effort, Cantor (and others that followed him) failed to settle the issue one way or another. In retrospect, this is not surprising. During the 1940s Kurt Gödel (1986, 1989) proved that if the standard principles of set theory are consistent, then they do not allow one to refute the continuum hypothesis, in other words, adding the continuum hypothesis to standard set theory does not lead to inconsistency. Roughly two decades later Paul Cohen (1963, 1964) proved that the same basic axioms fail to prove the continuum hypothesis as well (again, assuming that standard set theory is consistent). As a result, we can add either the continuum hypothesis or its negation to standard set theory, and either way no contradiction results.

THE INFINITELY LARGE TODAY: ZERMELO FRAENKEL SET THEORY

Of course, the Gödel/Cohen result is not all that interesting until one knows what the standard axioms of set theory are. The theory in question is called *Zermelo Fraenkel set theory* (or ZFC) and consists of the following axioms and axioms schemes (assume here that the quantifiers range only over sets).

First, we have the *axiom of extensionality*:

Extensionality: $(\forall x)(\forall y)(x = y \leftrightarrow (\forall z)(z \in x \leftrightarrow z \in y))$

which says that there cannot be two distinct sets with exactly the same members (i.e., sets are individuated by their members). Next, we have two purely existential axioms, that is, axioms that assert outright the existence of objects. The *empty set axiom*:

Empty Set: $(\exists x)(\forall y)(\neg\, y \in x)$

states that there is a set that contains no members (i.e., $\emptyset$). The *axiom of (Zermelo) infinity*:

Infinity: $(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \rightarrow \{y\} \in x))$

states that there is a set that contains the empty set and the singleton of every set which it contains (i.e., the set contains $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}$ …).

Following these we have what we can call conditional existence axioms, which tell us which sets can be "built up" from previously existing objects. The first of these is the *pairing axiom*:

Pairing: $(\forall x)(\forall y)(\exists z)(\forall w)(w \in z \leftrightarrow (w = x \vee w = y))$

The pairing axiom asserts that, given any two objects, there is a set that contains exactly those two objects and nothing else. The pairing axiom guarantees that the sin-

gleton of any object exists (because we can just take the pair of an object and itself). Next we have the *union axiom*:

Union: $(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\exists w)(w \in x \wedge z \in w))$

which states that, given any set, there is second set that contains exactly the members of the members of the first. Unsurprisingly, we also have an axiom asserting the existence of the powerset of any set:

Powerset: $(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\forall w)(w \in z \rightarrow w \in y))$

Our next principle, the *axiom of choice*, states that, given any set of nonempty pairwise disjoint sets, there exists a second set that contains exactly one member from each of the sets contained in the original set. The axiom of choice is often replaced with a more easily understood, but provably equivalent, principle called the *well-ordering principle*:

Well-Order: $(\forall x)(\exists \mathrm{R})(<x, \mathrm{R}>$ is a well-ordering)

The well-ordering principle guarantees that for any set, no matter how large, there is a relation that well-orders its members. During the first half of the twentieth century the status of the axiom of choice was highly controversial; since then it has become a standard part of the everyday mathematician's toolkit.

Our final two conditional existential principles take the form, not of single axioms, but axiom schemes, which have infinitely many instances. The first is the *axiom(s) of separation*. Given any condition Φ expressible in the language of set theory:

Separation: $(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (z \in x \wedge \Phi z))$

This axiom states that, given any set and any condition on objects, there exists a set that contains exactly the members of the original set that satisfy the condition in question. The second schematic principle is the *axiom(s) of replacement*, which consists of all instances of:

Replacement: $(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\exists w)(w \in x \wedge f(w) = z))$

where f is any function definable in the language of set theory. Put loosely, given any set and any function, replacement insures that there is a set containing exactly the objects obtained by applying the function to the members of the original set.

The final axiom of *Zermelo Fraenkel set theory* does not assert the existence of any sets, but instead imposes a restriction on what sorts of sets exist. The *axiom of foundation*:

Foundation: $(\forall x)((\exists y)(y \in x) \rightarrow (\exists z)(z \in x \leftrightarrow \neg (\exists w)(w \in x \leftrightarrow w \in z)))$

asserts that any nonempty set (i.e., any set other than $\emptyset$) contains as a member some second set that has no members in common with the original set. Although it is difficult to sum up the consequences of this axiom in simple terms, its main purpose it to rule out the existence of sets that contain themselves, such as the (potential) set that has itself as its only member, that is, $\Omega = \{\Omega\}$.

When confronted with such a list, a number of questions naturally arise, including: (1) Might there be a simpler set of principles that does the same job?; (2) Why have we chosen these principles?; (3) Might there be additional principles that we have overlooked? Complete answers to all of these questions are beyond the scope of the present article, but partial answers can at least be given.

Regarding the first question, we can rule out one initially promising simplified theory, *Naive Set Theory*, which has one axiom schema. The principle in question is the *naive comprehension principle*:

Naive Comp: $(\exists y)(\forall z)(z \in y \leftrightarrow \Phi(z))$

which states that, for any condition whatsoever (as long as we can express it in our set theoretic language), there is a set containing exactly the objects that satisfy that condition (note the similarity to separation above). The naive comprehension principle entails all of the axioms given above. Unfortunately, however, it also entails a contradiction, as was famously proved by Bertrand Russell (1996).

The reasoning, which has come to be known as Russell's paradox, proceeds as follows. Given naive comprehension, some sets will be members of themselves (such as the set of all sets) and some will not (such as the empty set). Because a set's not being a member of itself is a condition expressible in the language of set theory (i.e., $x \notin x$), if the following instance of naive comprehension:

$$(\exists y)(\forall z)(z \in y \leftrightarrow x \notin x)$$

were true then there would be a set that contains exactly those sets that are not members of themselves. Call this (supposed) set the *Russell Class*, or R (collections too ill-behaved to form a set are called *proper classes*, although this terminology obscures the fact that such "classes" cannot be objects at all). To obtain the contradiction, we need only ask whether R is a member of itself. By the criteria just stated, we can conclude that R is a member of itself if, and only if, it is not a member of itself. The contradiction is evident. (Similar arguments demonstrate

that there cannot be a set of all ordinals, a set of all cardinals, or a set of all sets.)

Thus, we must reject naive comprehension, and with it any conception of set that allows for the existence of the Russell Class. ZFC provides us with one means to achieve this. The next question, however, is why we have chosen these axioms, because presumably there are other collections of principles that would do the same job. The answer is that the axioms were not chosen at random, nor was their selection based merely on their individual plausibility. Instead, these axioms were selected because they correspond to natural thoughts regarding how one should separate the legitimate sets from those "collections," such as the Russell Class, that are not well-behaved enough to correspond to sets.

There are, roughly speaking, two intuitive pictures of the universe of sets that have motivated the formulation of ZFC. The first, and more influential, is founded on the idea that each set is built up from other sets or objects that are simpler, or at least prior to, the set in question. This notion of set is known as the *iterative conception of set*, and is summarized by George Boolos:

> According to the iterative, or cumulative, conception of sets, sets are formed at stages; indeed, every set is formed at some stage of the following "process": at stage 0 all possible collections of individuals are formed … The sets formed at stage 1 are all possible collections of sets formed at stage 0, … The sets formed at stage 2 are all possible collections of sets formed at stages 0 and 1. The sets formed at stage 3 are all possible collections of sets formed at stages 0, 1, and 2 … The sets formed at stage 4 … In general, for any natural number n, the sets formed at stage n are all possible collections of sets formed at stages earlier than n, i.e. stages 0, 1, … , $n - 1$. Immediately after all stages 0, 1, 2, … there is a stage, stage ω. The sets formed at stage ω are, similarly, all possible collections of collections of sets formed at stages earlier ω, i.e., stages 0, 1, 2, … After stage ω comes stage ω+1: at which … In general, for each α, the sets formed at stage α are all possible collections of sets formed at stages earlier than α. There is no last stage: each stage is immediately followed by another. Thus there are stages ω+2, ω+3, … and so it goes. (1989, p. 88)

Another notion that has been used to justify the particular choice of axioms constituting ZFC (and which is closer to what Cantor originally had in mind) is the *limitation of size conception of set*. The underlying thought is that problematic collections, such as the Russell Class, are not sets because they are, in some sense, too big. Boolos sums up the limitation of size conception as the view that "objects form a set if and only if they are not in one-one correspondence with all the objects there are" (1989, p. 90), in other words, a collection forms a set if, and only if, it is smaller than the universe of all sets or all objects.

The exact role that such conceptions of set can, and should, play within the philosophy and practice of set theory is something of an open question. One particular worry surrounding such intuitive pictures is that the axioms of *ZFC* do not all seem to follow from a single conception. For example, the powerset axiom seems basic on the iterative conception of set but less obvious on a limitation of size understanding, whereas the axiom of replacement seems straightforward on the latter approach but somewhat questionable on the iterative conception. Nevertheless, we can at least recognize that these conceptions played a significant role in the actual choice of axioms included within ZFC.

The final question to ask is whether there might be additional axioms that can be added to the basic theory. The answer, of course, is "yes." We have already seen one such principle, the continuum hypothesis. As we noted, we can add either the continuum hypothesis or its negation to ZFC, and in either case we obtain a consistent theory. If one thinks there is a unique universe of sets (and debate rages over this question) then at most one of these theories can be correct. Even so, the Gödel/Cohen results are enough to guarantee that ZFC alone leaves at least some set theoretic questions unanswered.

There is another type of question to which ZFC provides only a partial answer, namely determining how many sets there are. The axioms of ZFC imply that the universe of sets is larger than any particular set, otherwise the axiom(s) of replacement would provide us with a set of all sets that, using the axiom(s) of separation, would provide us with the Russell Class and a contradiction. Thus, we can ask: How big must the universe be in order to satisfy the axioms of ZFC?

Before providing the answer, we need a few more definitions. First, the *supremum* of a set of cardinal numbers A (i.e., sup(A)) is the smallest cardinal greater than or equal to each of the cardinal numbers in the set. Second, a cardinal number κ is *regular* if and only if, given any set of cardinals A, if card(A) $< \kappa$ and, for every cardinal $\gamma \in$ A, $\gamma < \kappa$, then sup(A) $< \kappa$ (intuitively, a cardinal κ is regular if we cannot reach it by summing up smaller cardinals unless we add up at least κ-many smaller cardinals). Finally, a cardinal κ is *strongly inaccessible* if and

only if it is uncountable, it is regular and, for each cardinal γ less than κ, 2^γ is also less than κ.

We can now answer the question just raised: any collection of objects that satisfies the axioms of ZFC has a cardinal number at least as big as the first strongly inaccessible cardinal number. (We are fudging here a bit, because we have assumed that only sets have cardinal numbers. One should consult a good textbook to see the formal details.)

If there must be (according to the axioms of set theory) at least strongly inaccessibly many objects, our next question might be: Why should there not be a set that is that big? The answer, according to set theorists, is that there is (or at least, it is worth exploring the assumption that there is). Thus, one of the most common axioms added to ZFC is the *strong inaccessible cardinal axiom*, which is equivalent to the claim that there is a set whose cardinal number is strongly inaccessible.

We need not stop here. Once we have determined how big the universe needs to be in order to satisfy all of the axioms of ZFC plus the strong inaccessible cardinal axiom, we can add axioms that assert the existence of even larger sets (and thus larger cardinal numbers). This process can continue indefinitely.

Thus, recognition of the fact that there is no set of all sets and no largest set leads us to posit the existence of larger and larger sets, resulting in stronger and stronger theories. Principles that assert the existence of extremely large sets are called *large cardinal axioms* (technically, a *large cardinal* is any cardinal number whose existence cannot be proved from the axioms of ZFC alone). Surprisingly, large cardinal axioms often have consequences for less esoteric areas of mathematics, such as the theory of the real numbers, although for the most part they fail to impact the status of the continuum hypothesis. In addition, adopting stronger theories including large cardinal axioms allows us to prove the consistency of weaker theories. As a result, the study of large cardinal axioms is one of the most fruitful areas of research within set theory.

The introduction of large cardinal axioms is one instance of a more general method for generating new set theoretic principles. The method, known as *reflection*, assumes that, for any sentence true of the entire set theoretic universe, there will be a set such that the sentence will also be true when restricted to that set. Thus, letting *SI* be the claim that the universe contains strongly inaccessibly many sets, reflection tells us that, because *SI* is true, *SI* must be true when restricted to some set (i.e., there is a set that contains inaccessible many other sets as members, or, equivalently, there is a set whose cardinal number is a strong inaccessible cardinal).

Interestingly, reflection principles imply (and in some cases are equivalent to) strong versions of the axiom of choice. For example, let Φ be the claim that there is no well-ordering on the entire universe of sets. Reflection implies that if Φ is true, then there is some set whose contents cannot be well ordered. The axiom of choice is equivalent to the claim that every set can be well-ordered, however, so, if we accept both the axiom of choice and reflection, then Φ must be false; in other words, there must be a relation that well-orders the entire universe of sets. This thesis, known as *global choice*, is stronger than the axiom of choice.

We can now observe the great irony of the theory of the infinite. The revolution in the mathematics and philosophy of the infinite occurred in late nineteenth century when mathematicians such as Cantor and Dedekind abandoned the Aristotelian view that infinite collections could not, and should not, be thought of as completed totalities (i.e., sets). Their account of infinite totalities, however, and subsequent work on large cardinals, suggests a universe of sets that in some sense is ever expanding, with no upper limit to the size or variety of sets themselves. As a result, Aristotle was at least partially right, because there are collections (such as the collection of all objects, or all sets) that cannot, and should not, be thought of as completed, definite totalities.

Instead, the universe of sets, considered as a whole, is what has come to be called *indefinitely extensible*. Russell, reflecting on the paradoxes that arise when some intuitive "collections" are treated as sets (such as the paradox that bears his name) described the situation as follows:

> A concept is indefinitely extensible if, for any definite characterization of it, there is a natural extension of this characterization, which yields a more inclusive concept; this extension will be made according to some general principle for generating such extensions, and, typically, the extended characterization will be formulated by reference to the previous, unextended characterization. (1963, pp. 195–196)

Understanding Russell's "definite characterization" as the contents of a set, any time we think we have collected together all the objects, or all of the sets, together into a single set, there will turn out to be more objects, or sets, which we somehow missed. To put the point loosely, Cantor's embrace of the actually infinite has led us, in the

end, to the recognition that the set theoretic hierarchy itself is (or is something very much like) potentially infinite.

ACTUAL INFINITY AND INFINITE DIVISIBILITY

As we saw at the beginning of this entry, Aristotle's solution to Zeno's paradoxes relied on the fact that we do not accomplish infinitely many things when we run from point 0 to point 1, because it is never the case that all of the parts of this journey (including the infinite sequence of distances 0 to 1/2, 1/2 to 1/4, 1/4 to 1/8, …) are present at one time. This solution relied, in turn, on the distinction between potential and actual infinity and on Aristotle's insistence that the actual infinite is illusory. If, however, our post-Cantorian mathematical view allows, and even embraces, actually infinite collections, then we are left with the possibility that the problematic parts of the runner's path from 0 to 1 are actually present as a completed totality. If so, then, assuming that motion is possible, it follows that we can accomplish infinitely many tasks (and we do so every time we wiggle our little finger!)

The most tempting response to this line of thought is "So what?" We might be surprised to learn that any action (no matter how slight) involves infinitely many tasks, but this astonishing fact is little more than the result of a clever (and perhaps misleading) description of the event. The movement from 0 to 1 might be, on one description, composed of infinitely many smaller motions, but it can also be viewed as a single continuous action, that of moving from 0 to 1. It is this latter fact that explains how we can accomplish the movement, and furthermore do so in a finite amount of time.

A task that consists of infinitely many subtasks carried out in a finite amount of time (such as Zeno's description of the runner traveling from 0 to 1) is called a *supertask*. Once we admit that Zeno was right, and we can sometimes carry out supertasks, problems emerge. There are other easily describable supertasks that seem to lead to paradoxes. Two of the most famous are *Thompson's Lamp* and *Bernardete's Paradox*.

Imagine an ideal lamp (indestructible, and able to be switched on or off instantaneously) that, at exactly 12:00, is turned on. 1/2 minute later, it is switched off, then 1/4 minute later it is switched back on, then 1/8 minute later it is switched off, and then 1/16th minute later it is switched on, … The infinite series of switchings will be completed at exactly 12:01. The puzzling question that arises is: Will the lamp be on or off once the supertask is completed? There seems to be no good reason to answer one way rather than another. Nevertheless, a lamp, even an ideal one, must be either on or off.

Thompson's Lamp is similar to Zeno's paradox of the runner in that it divides a unit of time into the first half, the first half of the second half, the first half of the last fourth, the first half of the last eighth, … Bernardete's paradox, however, forces us to bizarre (if not outright contradictory) conclusions using the mirror image of this division, namely dividing a unit of time into the last half, the last half of the first half, the last half of the first fourth, the last half of the first eighth, … (more intuitively, we can see it as having the structure … + 1/32 + 1/16 + 1/8 + 1/4 + 1/2 instead of 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + …).

José Bernardete constructs a number of different versions of his paradox, here we will consider a particularly striking formulation. Imagine an object that exists up until 12:00, when an infinite series of gods take notice of it. The gods are omnipotent and, additionally, they always carry out the actions they decide upon. The first god decides that if the object still exists at 1/2 minute after 12:00 then he will annihilate it (he will do nothing otherwise). The second god decides to annihilate the object if it still exists at 1/4 minute past 12:00 (and again, do nothing otherwise). The third god decides to annihilate the object at 1/8 minute after 12:00 if it still exists, and so on. There is no threat to the existence of the object other than the intentions of each of the infinite series of gods. We can conclude that the object will suddenly cease to exist at exactly 12:00, yet nothing (and in particular, no god) will have caused its destruction.

First of all, assume that the object exists past 12:00. The there must be some fraction of a minute $1/x$ such that the object existed for (at least) that long after 12:00. If so, however, then one of the gods failed to live up to his intentions, because there will be some god in the list who decided to destroy the object if it still existed at $1/y$ minutes past 12:00 where $1/y$ is less than $1/x$ (because the series 1/2, 1/4, 1/8 1/16… tends to zero). But this contradicts the fact that the gods always carry out their intentions. Thus, the object will cease to exist at 12:00.

The natural assumption to make is that one or more of the gods must have destroyed it, but we can see this is incorrect as well. Each of the gods decided to destroy the object at a time after 12:00 (and to do nothing otherwise). Because the object did not survive past 12:00, the gods did nothing. Thus, the object blinks out of existence at exactly 12:00 yet nothing acted in such a way as to cause its disappearance. Even if not exactly paradoxical, Bernardete's paradox is deeply puzzling.

Of course, unlike Zeno's runner, both of these puzzles begin with an absurd situation. On the one hand we have a lamp that can be turned on and off arbitrarily fast and infinitely many times without malfunction, on the other we have an infinite collection of gods (most of us nowadays still puzzle over whether or not there is even one god!). It is tempting to conclude that the proper reaction to these puzzles should be mere amusement but not worry, because they concern situations that are at such a distance from our everyday experience of the world.

This would be a mistake. Both Thomson's Lamp and Bernardete's paradox are intended to challenge our understanding of the infinite and infinite divisibility. To dodge such challenges by noting that they require physically impossible situations or events is to miss the point. Surely neither an indestructible lamp nor an infinite pantheon of gods is logically impossible, and if they are logically possible then their behavior is relevant to our understanding of the infinite (which is, after all, a logical, or at least mathematical, concept). Thus, these puzzles, and others like them, are not mere curiosities but instead represent important unsolved problems confronting the coherence of, and our understanding of, the notion of infinity at the heart of mathematics and philosophy.

This entry ends in roughly the same way in which it began, with a discussion of infinite divisibility. This is the topic Aristotle focused on, and where his notion of the potentially infinite gained its greatest (albeit temporary) success. Even as we abandoned Aristotle's proscription on the actually infinite and embraced modern set theory, puzzles such as Russell's paradox and the nonexistence of a set of all sets prevented us from rejecting the potentially but not actually infinite altogether. Finally, given our acceptance of at least some actually infinite collections, paradoxes similar in structure to those that Aristotle first considered continue to plague our understanding of collections that keep going and going and going and going …

See also Aristotle; Bolzano, Bernard; Brouwer, Luitzen Egbertus Jan; Cantor, Georg; Constructivism and Conventionalism; Continuity; Galileo Galilei; Gödel, Kurt; Intuitionism; Leibniz, Gottfried Wilhelm; Plato; Russell, Bertrand Arthur William; Set Theory; Weyl, (Claus Hugo) Hermann; Wittgenstein, Ludwig Josef Johann; Zeno of Elea.

Bibliography

HISTORICAL SOURCES

Aristotle. *The Physics*. Book 8. Translated by Daniel W. Graham. Oxford, U.K.: Clarendon Press, 1999.

Bolzano, Bernhard. *The Mathematical Works of Bernhard Bolzano*. Translated by Steve Russ. Oxford, U.K.: Oxford University Press, 2004.

Cantor, Georg. *Contributions to the Founding of the Theory of Transfinite Numbers*. New York: Dover, 1955.

Dedekind, Richard. *Essays on the Theory of Numbers*. New York: Dover, 1901.

Leibniz, Gottfried. *Mathematische Schriften*. 7 vols, edited by C. I. Gerhardt. Berlin: A. Asher, 1849–1863.

Leibniz, Gottfried. *Philosophical Papers and Letters*. 2nd ed. Translated by Leroy Loemker. Dordrecht, Netherlands: Kluwer, 1976.

Plato. *Philebus*. Translated by R. Hackforth. Cambridge, U.K.: Cambridge University Press, 1972.

Russell, Bertrand. *The Principles of Mathematics*. New York: Norton, 1996.

CONTEMPORARY VIEWS ON INFINITY AND SET THEORY

Boolos, George. "Iteration Again." *Philosophical Topics* 42 (1989): 5–21. Reprinted in *Logic, Logic, and Logic*, edited by Richard Jeffrey. Cambridge, MA: Harvard University Press, 1999, 88–104.

Cohen, Paul. "The Independence of the Continuum Hypothesis." *Proceedings of the National Academy of Sciences USA* 50 (1963): 1143–1148.

Cohen, Paul. "The Independence of the Continuum Hypothesis II." *Proceedings of the National Academy of Sciences USA* 51 (1964): 105–110.

Enderton, Herbert. *Elements of Set Theory*. New York: Academic Press, 1977.

Gödel, Kurt. *Collected Papers*. Vol. 1, *1929–1936*. Oxford, U.K.: Oxford University Press, 1986.

Gödel, Kurt. *Collected Papers*. Vol. 2, *1938–1974*. Oxford, U.K.: Oxford University Press, 1989.

Kunen, Kenneth. *Set Theory: An Introduction to Independence Proofs*. Amsterdam: North Holland, 1998.

Moore, A. *The Infinite*. New York: Routledge, 2001.

ZENO'S PARADOXES AND SUPERTASKS

Bernardete, J. *Infinity: An Essay in Metaphysics*. Oxford, U.K.: Clarendon Press, 1964.

Earman, John, and John Norton. "Infinite Pains: The Trouble with Supertasks." In *Benacerraf and his Critics*, edited by A. Morton and Steven Stich. Cambridge, MA: Blackwell, 1996.

Kirk, Geoffrey, John Raven, and Malcolm Schofield. *The Presocratic Philosophers*. Cambridge, U.K.: Cambridge University Press, 1983.

Roy T. Cook (2005)

INFINITY IN THEOLOGY AND METAPHYSICS

It would be profitless (even if it were possible) to catalog every nuance that the word *infinity* possesses in minor, as well as major, thinkers. Fortunately, the dominant strands are clear. Among these the theistic one is the most important both historically and in terms of contemporary debate.

GREEK PHILOSOPHY

ANAXIMANDER. The first Western philosopher to speculate on infinity was the pre-Socratic Anaximander. By the infinite (*to apeiron*) he meant a limitless substance from which the limited things that constitute the world have come. This substance is limitless in three respects: It is eternal, not having a beginning or an end; it is inexhaustible; and it lacks internal boundaries and distinctions. But it is not spatially unlimited, for Anaximander (almost certainly) conceived it as a sphere. Also, it is not qualitatively indeterminate, like Aristotle's unformed matter, for it contains nature's basic elements in a fused, nonseparated state.

PYTHAGORAS. The Pythagoreans adopted Anaximander's concept. Some of them identified it with air (which Anaximenes considered to be the basic constituent of the universe). But their main contribution was to posit a limit (*peras*) as a principle that gives structure to the limitless or infinite. This limit was mathematical; the limitless once limited gives the point, twice limited the line, thrice limited the plane, and four times limited the solid. Later writers interpreted Pythagoras theologically. Thus in the *Placita* we are told that he believed in two principles—the monad (God, the Good, the essential nature of the One, *Nous* alone and by itself) and the indefinite dyad (or evil, which is bound up with materiality and multitude).

PLATO. Plato's speculations on infinity are contained in his *Philebus.* He gives a fourfold classification of "all that now exists in the universe." The whole world can be viewed in terms of the unlimited, limit, mixture, and the cause of the mixture. This theory is an application of the axiom that the nature (and therefore the good) of anything consists in an intelligible order or proportion. The cosmic cause mixes limit with the unlimited and so imposes structure on the world. In 15D–17A Plato interprets the *peras-apeiron* contrast logically. The unlimited stands for particulars, and the limited for the species into which they can be put. But in 23C–26D the contrast has an ontological significance of a Pythagorean kind. The limitless consists in a collection of opposites (for example, hot and cold, dry and moist). Limit consists in "all that puts an end to the conflict of opposites with one another, making them well proportioned and harmonious by the introduction of number" (25E). This principle of limitation is essential also in the moral realm. Plato affirms that human pleasures (which, in themselves, tend to unlimited excess) ought to be rationally controlled by a law and order that are marked by limit.

Thus, in classical Greek philosophy *infinity* represents a substratum that is formless, characterless, indeterminate. It is a pejorative word. An entity is good to the extent that it is limited by form. The Pythagoreans identified this form with numerical ratios. But, as the *Philebus* shows, it can be nonnumerical (such as a universal essence or the personal activity of reason).

An important fact emerges from this survey. Plato could not envisage God (or the divine) as infinite. If God is perfect, he must represent the principle of limit. The cause of cosmic mixture in the *Philebus* is equivalent to the Demiurge in the *Timaeus.* The latter's task is to impose intelligible form on preexistent matter and thereby make an ordered whole. Otherwise the world would be a vast *apeiron*—a formless, unintelligible chaos. Hence, to say that he is *apeiros,* or that the Forms that he copies are *apeira,* would have seemed self-contradictory.

PLOTINUS. Plotinus occupies a place between Plato and Christian theologians who, if they are orthodox, regard infinity as the first among God's attributes. Plotinus applied the concept of the infinite, or unbounded (*apeiron* or *aoriston*), to two categories of being. First, he applied it to matter, which is evil because it tends intrinsically to formlessness. In this he developed philosophical tradition. But second, he applied it to the divine hypostasis. Thus, he called Mind infinite because of its endless power, its complete unity, and its self-sufficiency. Yet while he says that the One is formless, he does not say that it is infinite. The history of *apeiron* prevented him from predicating it of the Absolute. He expressed the infinite nature of the One by denying that any positive idea abstracted from finite experience is applicable to it.

MEDIEVAL AND MODERN PHILOSOPHY AND THEOLOGY

Throughout the postclassical period of Western thought it has been widely assumed that God, or the Absolute, is infinite, or limitless. The division lies between those philosophers (such as Giordano Bruno, Benedict de Spin-

oza, and G. W. F. Hegel) who interpret God pantheistically and those (especially Christian theists) who hold that he wholly transcends the world. According to the first group of thinkers, the world, being divine, is also infinite (even if particular things and persons reflect its infinity in limited degrees). According to the second group, the whole world is finite (as created), and only God (as the Creator) is infinite.

PANTHEISM. The clearest example of the pantheistic group is Spinoza. Having posited one substance (God or nature), he affirmed that it must be infinite both in its essence and in the number of its attributes. God must be infinite in his essence because if he were finite, we could suppose the existence of something else by which he is limited, so that he would not be the sole reality. His attributes must be infinite because if his essence is infinite, there must be an infinite number of ways in which it can be conceived.

Hegel's theory is more dynamic and complex. It was based on the conviction that finite and infinite are correlative terms within a single system of thought and reality. The Absolute Spirit (God) is infinite. But it does not exist outside the finite spirits through whom it manifests itself. Since the world is the manifestation of the Absolute, and since the Absolute requires the world for its development, we can predicate infinity either of the Absolute (considered as an identity-in-differences) or of the world (considered as a rational totality). Hegel considered Christianity to be the highest form of religion because it represents a perfect reconciliation between man and God, the finite and the infinite.

Any theory that views the finite as, in some sense, the self-expression of the infinite is exposed to two basic objections.

(1) The world (so the theist claims) is not limitless. It is limited in two main ways. First, it is morally imperfect. The premise of Immanuel Kant's moral argument for immortality is irrefutable. We cannot in this life bring our wills into complete accordance with the moral law, and even if we could do so, the spatiotemporal order could not fulfill our deepest longings (as A. E. Taylor argued in his Gifford Lectures). Second (and this is the core of theism), the world in all its aspects bears the marks of radical contingency, so that its existence cannot be explained unless we suppose it to be derived from a transcendent being who is infinite or absolute.

(2) In any case, the world is full of differences and discordances. How can these be reconciled within a unitary Absolute? How can a set of finite (that is, limited and mutually exclusive) entities constitute a nonlimited and all-inclusive whole? In particular, how can this whole, if it is complete and perfect (as it must be if it is infinite), contain within itself both good and evil? There is no satisfactory answer to these questions. Nicholas of Cusa, in his pantheistic moments, affirmed that in God there is a "synthesis of opposites" (*coincidentia oppositorum*). Similarly, Friedrich von Schelling affirmed that the Absolute is a self-identity in which all differences vanish. But these affirmations are metaphysically vacuous, as Nicholas admitted when, using mystical terminology to conceal a contradiction, he called our knowledge of the all-inclusive Maximum a *docta ignorantia.*

THEISM. Theists do not have to face the above problems. Certainly they hold that all perfections preexist in God eminently. But they also hold that the mode of this existence is determined by the infinity that God does not share with any creature. God's infinity means that he is "not-finite." He is free from the limitations that affect every other being. There are two fundamental limitations.

First, every other being is a mode of existence (or existing). A man exists in one way, a dog in another. But God is existence *simpliciter.* He does not suffer from the determinations that are reflected in genera and species. We can express this (with deliberate paradox) by saying that he is his own genus.

Second, if God is existence "in itself," he must be self-existent in the sense that he does not derive his being from any other source. Every other being is dependent or derived. It does not contain within itself the cause of its existence. It depends continuously on the creative act of God who alone exists *a se* (that is, by his own intrinsic power).

Both these aspects of God's infinity are affirmed by the Scholastic dictum that in him essence and existence are identical. The finitude of any being consists in the lack of this identity at both the points mentioned above. Its essence limits its existential act (or pattern of activity), and this limitation follows from its dependent character. It exists as "this" or "that" by its derivation from a being who is existence in a necessary and perfect form.

This view of God's infinity must be safeguarded by the following assertions.

(1) God's infinity is not to be interpreted as formlessness (as if it were equivalent to Plato's *apeiron*). It is the nature of finite being (at any rate in the subangelic realm) to be a *compositum* of form and matter. The form limits matter. Without some degree of limitation there would be no difference (either generically or individually) between one finite being and another. But since God's essence and existence are identical, his form cannot be a principle of limitation. "Matter," Thomas Aquinas wrote, "is perfected and made definite by form. Infiniteness attributable to matter is imperfect and amorphous. On the other hand, form as such is not perfected by matter, but contracted rather; hence infiniteness attributable to form is perfection" (*Summa Theologiae* I, 7, 1).

(2) God's infinity is incomprehensible. We cannot imagine or conceive it. We can know that God is self-existent. But how he is self-existent is utterly unknowable by us in our present state. As soon as we try to represent his infinity through a univocal use of concepts, we commit three errors. We fall into anthropomorphism; we confuse infinity with formlessness; and, finally, we reach a self-contradiction, for the essence of a finite entity (however high it may be on the scale of being) is to possess a form that acts as a limit that excludes other forms.

However, the various attributes that constitute God's character are all deducible from his self-existence. He must be absolutely simple, for if in him essence and existence are identical, his qualities must be coinherent through the whole range of his activity. He must be spiritual and nontemporal, for corporeality entails spatial limitation and temporal successiveness implies divisibility. He must be omniscient and omnipotent, for there cannot be any externally imposed limit to his knowledge or his power. Finally, he must be absolutely good.

Two of these characteristics, spirituality and eternity, call for comment. Since God is nonspatial and nontemporal, the concept of his infinity is unaffected by the views we hold concerning space and time. Whether space and time are limited or unlimited makes no difference to the claims of the theist concerning God's infinity and his relation to the world. Thus, even if the world has existed for an endless length of time, it would still (according to the Cosmological Argument) be endlessly incomplete, so that we should still have grounds for positing a nontemporal act of divine creativity.

Yet the theistic view of God's infinity raises problems of its own. Four are especially urgent. First, if God is infinite and we are finite, how can we speak of him positively (as the biblical writers and doctrinal theologians do)? Second, Christians affirm that God is personal. But does not the idea of personality conflict with the idea of infinity? (This objection was first urged by Carneades and later elaborated by David Hume.) Third, is it not contradictory to say that all God's attributes (for example, justice and mercy) can coexist in a limitless degree? Are not even theists forced to posit a *coincidentia oppositorum* in the Godhead? Fourth, if God is infinite both in goodness and in power, how can we explain the presence of evil in the world?

The answers that theists normally give to these objections are as follows.

(1) While we cannot speak of God univocally, we can do so analogically. But in applying any analogue to God, we must distinguish between the manner of predication and the object signified. The only positive meaning that we can attach to a term we predicate of God is the one which it has when predicated of finite beings. Yet since God and the creature are ontologically related by an analogy of attribution, we can *affirm that* (although we cannot *know how*) the divine analogate possesses the analogue, according to the analogy of proportionality, in a manner appropriate to his infinite existence.

(2) The basic answer to the second question is that we need not equate the essence, or norm, of personality with its human mode. On the contrary, the latter (according to the Bible) is a created image of an infinite archetype. The theist would claim that while we cannot *see how* God can be both infinite and personal, we can *understand that* an infinite existence, so far from being incompatible with personality, would represent it in its most perfect form. At any rate (so the theist would maintain), it is not contradictory to assert that individuality can exist without individuation and that God therefore can have a positive character without possessing characteristics of the kind that differentiate a member of one created genus from a member of another.

(3) If God's attributes were essentially incompatible, they could not be predicated of him infinitely and simultaneously without a logical contradiction that could be solved (as Nicholas and Schelling found) only by an *asylum ignorantiae.* But theists

claim that any contradiction is only apparent. Everything depends on how we define our terms. Thus, if we take justice to mean retribution, it is bound to be incompatible with mercy, if both are infinitely conceived. But if we take it to mean the vindication of the moral order, mercy becomes (as St. Paul saw) the primary form of its expression.

(4) Most Christian theists would admit that the fact of evil seems to be incompatible with belief in a God who is infinite both in goodness and in power. But they would also claim that the apparent incompatibility disappears once we recognize first, that since God's power and goodness are inconceivable, his purposes are bound to be largely inscrutable and second, that in Christ he has shown that he not only can but also does bring the greatest good out of the greatest evil.

See also Absolute, The; Analogy in Theology; Anaximander; Bruno, Giordano; Carneades; Cosmological Argument for the Existence of God; Eternity; Hegel, Georg Wilhelm Friedrich; Hume, David; Nicholas of Cusa; Plato; Plotinus; Pythagoras and Pythagoreanism; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de; Taylor, Alfred Edward; Theism.

Bibliography

For literature on Anaximander, see F. M. Cornford, "Anaximander's System," in *Principium Sapientiae* (Cambridge, U.K.: Cambridge University Press, 1952), pp. 159–186. Plato's views in the *Philebus* are discussed by R. Hackforth in his commentary in *Plato's Examination of Pleasure* (Cambridge, U.K.: Cambridge University Press, 1945). A. H. Armstrong surveys Plotinus thoroughly in "Plotinus" Doctrine of the Infinite and Christian Thought," in *Downside Review* (Winter 1954–1955): 47–58. The teaching of Thomas Aquinas on God's infinity (together with associated attributes) is collected by Thomas Gilby in his anthology *Philosophical Texts* (Oxford: Oxford University Press, 1956). Nicholas of Cusa, *De Docta Ignorantia,* is translated by G. Heron, with an introduction by D. J. B. Hawkins (New Haven, CT: Yale University Press, 1954). Bruno's theory of infinity can be found in *On the Infinite Universe and Worlds,* which has been translated by D. W. Singer in *Giordano Bruno: His Life and Thought* (New York: Schuman, 1950). A lucid summary of Hegel and Schelling (with ample references) is given by Frederick Copleston in *A History of Philosophy,* Vol. VII (London, 1963). For a discussion of infinity in traditional Christian theism (and the problems that it raises), see A. M. Farrer, *Finite and Infinite* (London: Dacre Press, 1943), and E. L. Mascall, *Existence* and Analogy (London: Longmans Green, 1949).

Adams, Robert. *Finite and Infinite Goods: A Framework for Ethics.* Oxford: Oxford University Press, 1999.

Benardete, José. *Infinity: An Essay in Metaphysics.* Oxford: Oxford University Press, 1964.

Bolzano, Bernard. *Paradoxes of the Infinite.* Translated by F. Prihonsky. London: Routledge and Kegan Paul, 1950.

Duhem, Pierre. *Medieval Cosmology: Theories of Infinity, Place, Time, Void, and the Plurality of Worlds.* Chicago: University of Chicago Press, 1985.

Ganssle, Gregory, and David Woodruff, eds. *God and Time: Essays on the Divine Nature.* Oxford: Oxford University Press, 2001.

Kretzmann, Norman, ed. *Infinity and Continuity in Ancient and Medieval Thought.* Ithaca, NY: Cornell University Press, 1982.

Le Poidevin, Robin, ed. *Questions of Time and Tense.* Oxford: Clarendon Press, 1998.

Moore, A.W. *The Infinite.* London: Routledge, 1990.

Moore, A.W., ed. *Infinity.* Aldershot, U.K.: Dartmouth, 1993.

Sweeney, Leo. *Divine Infinity in Greek and Medieval Thought.* New York: Peter Lang, 1992.

H. P. Owen (1967)

Bibliography updated by Christian B. Miller (2005)

INFORMATION THEORY

Among the more interesting trends of the past half-century has been the consolidation of probability, statistics, combinatorial optimization, information theory, and computer science into a single imposing discipline of *infomatics*.

MINIMAL BELIEF CHANGE

Of special philosophical interest is the enrichment of Bayesian inference by a rule of belief change that goes by minimizing the distance between two probability distributions, $P = (p_1, \ldots , p_k$ and $Q = (q_1, \ldots , q_2)$, as measured by the expected log-likelihood ratio:

$$H(P,Q) = \sum_{i=1}^{k} p_i \ln(p_i / q_i) \quad (1)$$

The likelihood ratio, $P(e|h):P(e|k)$, is a fundamental index of the support that e accords h over k (see the entry "Foundations of Statistics").

Using the visually transparent Gibbs inequality,

$$\ln x \leq x - 1 \quad (2)$$

with equality if and only if (iff) $x = 1$, in the equivalent form $\ln x \geq 1 - 1/x$, it follows that $H(P, Q) \geq 0$ with equality iff $P = Q$. Notice, however, that $H(P, Q) \neq H(Q, P)$.

Alan Turing and his wartime assistant, Irving John Good, used $H(P, Q)$ in their code-breaking work, but it

was not until 1959 that another wartime code breaker, Solomon Kullback, developed its properties systematically in his book *Information Theory and Statistics* (1959), unleashing a floodtide of applications to classification, contingency tables, pattern recognition, and other topics.

A second line of development began with Claude Shannon's creation of information and coding theory (Shannon and Weaver 1949), whose central concept is a measure of uncertainty,

$$H(P) = -\sum_{i=1}^{k} p_i \ln p_i \tag{3}$$

that Shannon dubbed the entropy function. One sees that minimizing $H(P, U)$ against a uniform distribution, $U = (k^{-1}, \ldots, k^{-1})$, namely,

$$H(P,U) = \sum_{i=1}^{k} p_i \ln(p_i / k^{-1}) = \sum_{i=1}^{k} p_i \ln p_i + \ln k$$

is equivalent to maximizing the entropy (MAXENT). In view of this relation $H(P, Q)$ is often called the cross (or relative) entropy of P with respect to Q, and the rule of minimizing it MINXENT.

Shannon characterized $H(P)$ axiomatically as continuous, strictly increasing in k when $P = (k^{-1}, \ldots, k^{-1})$ is uniform, and by the general form of a consistency requirement exemplified by:

$$A(9) = \frac{3}{9}A(3) + \frac{4}{9}A(4) + \frac{2}{9}A(2) + H\left(\frac{3}{9}, \frac{4}{9}, \frac{2}{9}\right)$$

where $A(3) = H(⅓, ⅓, ⅓)$, and so on. Roughly speaking, one's uncertainty about a lottery is the same for an equivalent two-stage form of it, where in the example $H(3/9, 4/9, 2/9)$, is one's uncertainty about the first stage.

Exploiting the conceptual link Ludwig Boltzmann established between thermodynamic entropy and Shannon entropy through his celebrated formula, $S = k \ln W$, Edwin T. Jaynes showed how to derive the probability distributions (over microstates) of statistical mechanics by maximizing the (Shannon) entropy subject to mean value constraints given by measured values of macro variables like pressure, volume, or internal energy. Intuitively, the maxent distribution is the most spread out of all those satisfying the constraints, the one that is maximally noncommittal with regard to missing information. Statistical mechanics could thus be seen as a branch of statistical inference operating on physically given constraints. Jaynes quickly realized that MAXENT admits of much wider application. A collection of his papers (Jaynes 1983) sparked an explosion of applications whose subsequent progress can be followed in the published proceedings of workshops on maximum entropy and Bayesian methods held annually since 1981 (e.g., see Erickson and Smith 1988). Brian Buck and Vincent A. Macaulay's *Maximum Entropy in Action* (1991) is a sampler of physical applications to such fields as spectroscopy, X-ray crystallography, the structure of macromolecules, magnetic resonance, thermodynamics, and plasma physics. The method is universal, however, and applies equally well to image reconstruction or time series.

TABLE 1

k	1	2	3	4	5	6	entropy
n_k	3246	3449	2897	2841	3635	3932	
g_k	0.16230	0.17245	0.14485	0.14205	0.18175	0.19660	1.784 990
p_k	0.15294	0.15818	0.16361	0.16922	0.17502	0.18103	1.790 103
q_k	0.16433	0.16963	0.14117	0.14573	0.18656	0.19258	1.785 225
r_k	0.16139	0.17361	0.14434	0.14256	0.18215	0.19594	1.784 993

INFERRING HIDDEN CAUSES

Irving John Good (1983) viewed $H(P, Q)$ as a natural measure of deviation suitable for testing multinomial hypotheses, and like Karl Pearson's better-known chi-squared measure of deviation

$$X^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = n \sum_{i=1}^{k} \frac{(f_i - p_i)^2}{p_i}$$

the analogous informational measure

$$\psi = 2nH(F,P) = 2n \sum_{i=1}^{k} f_i \ln(f_i / p_i)$$

is asymptotically distributed as χ^2_{k-1}, the chi-square distribution with $k - 1$ degrees of freedom. By comparing the current model to the ideally best-fitting model, the resulting psi test sets limits to how much room there is for improving support by moving to a new (possibly more complicated) model (Jaynes 2003). When such a move is indicated, MINXENT helps guide one to plausible better-fitting models, as the following example illustrates:

Seen in the first row of Table 1 are the frequencies with which the six faces of a white die turned up in $N =$

20,000 tosses recorded by the Swiss astronomer Rudolf Wolf. One sees at a glance that Wolf's die was biased; the question of interest is how to find physically plausible hypotheses that best account for these data. The distributions shown in rows 3 to 5 of the table afford successively better approximations to the empirical distribution, $\{g_j\}$, of row 2. Jaynes (1983) found them by moving the initial uniform distribution just enough to satisfy mean value constraints given by the data that correspond to physical imperfections of a die. Among the most plausible are:

(1) A shift in the center-of-mass due to the different amounts of ivory removed from opposite faces

(2) Oblateness, due to the difficulty of machining a perfect cube (Jaynes 1983)

The first implies a mean higher than the 3.5 of an "honest" die. Maximizing the entropy subject to the observed mean, $\langle i \rangle = 3.5983$, yields the distribution *P* of row 3. Notice that it assigns higher probability to the 6 spot than the 1 spot, the 5 spot than the 2 spot, and the 4 spot than the 3 spot, with decreasing margins, which is just what the posited physical cause would lead one to expect. But while *P* improves on the uniform distribution, its fit to the data is still poor, with $X^2 = 33.4$. This points to the existence of another constraint. The lower than expected frequencies of faces 3 and 4 are best explained by oblateness of the die, the 3-4 dimension being longer than the other two. This is reflected in the nonzero mean value, $\bar{f}_2 = 0.1393$, of the oblateness function, $f_2(j) = 1$ for j = 1, 2, 5, 6 and $f_2(j) = -2$ for j = 3, 4. Adding this constraint to the first leads to the maxent distribution *Q* for both constraints (row 4). The fit to Wolf's data is now fairly good, but there is some slight evidence for a third constraint, a slightly chipped 2-3-5 corner. Maximizing entropy subject to all three constraints then yields the distribution of the last row, whose fit to the data is preternaturally good. Note: It makes no difference in which order the constraints are applied, provided each one is retained in applying the next one.

If Wolf's die still exists, one could actually put the posited physical imperfections to the test of careful measurements. In any case, one would expect other dice to exhibit the first two imperfections. As it happens, Wolf also tossed a red die 20,000 times, and this expectation is realized (Jaynes 1983). To permit such inferences to physical causes, one must assume, of course, that Wolf tossed his dice in a manner that precludes skill.

MAXENT first enters, then, as a technique of mathematical modeling, as a means of generating plausible hypotheses for explaining old data or testing against new data (Good 1983, pp. xvi–xvii, 41, 99–100). Indeed, most of the probability distributions that figure prominently in pure and applied probability are maxent distributions for a suitable set of constraints. For example, the exponential distribution is maxent for a positive random variable whose (finite) first moment is given. (This result depends on extending the cross entropy function to continuous distributions, sums giving way to integrals.) Jaynes (2003, §7.6) even makes a case for thinking the ubiquitous use of the normal law of errors is owing to the fact that it is the maxent distribution having specified values of the first two moments.

When predictions based on MAXENT are verified, one's belief in the completeness of the given constraints is borne out, but one learns most when one's predictions fail. For then one infers the existence of a previously unsuspected cause to which the observed deviations point one. However, for this kind of "learning from error" to be effective requires that those inferences be one's best inferences (Jaynes 2003, p. 326). In what sense, then, are MINXENT or MAXENT inferences best possible?

One can look, first, at how MAXENT operates on a mean value constraint:

$$\sum_{i=1}^{n} p_i f(x_i) = \bar{f}$$

With $\sum p_i = 1$, the method of Lagrange for constrained extrema yields a solution,

(4) $$p_i = Z(\lambda)^{-1} \exp(\lambda f(x_i)) \qquad i = 1, 2, \ldots, n$$

with $Z(\lambda) = \sum_{i=1}^{n} \exp(\lambda f(x_i))$. More generally, given $m < n$ such constraints, $\sum_{i=1}^{n} p_i f_k(x_i) = \bar{f}_k$, a solution is

(5) $$p_i = Z(\lambda_1, \ldots, \lambda_m)^{-1} \exp[\lambda_1 f_1(x_i) + \ldots + \lambda_m f_m(x_i)]$$

for $i = 1, \ldots, n$, where the partition function, *Z*, is defined by:

(5a) $$Z(\lambda, \ldots, \lambda_m) = \sum_{i=1}^{n} \exp[\lambda_1 f_1(x_i) + \ldots + \lambda_m f_m(x_i)]$$

Moreover, the moments of the resulting maxent distribution are given by the derivatives of ln *Z*, so that, in particular, the mean values are:

(5b) $$\frac{\partial}{\partial \lambda_i} \ln Z = \bar{f}_i$$

as one can verify. Thus, Z is a kind of (exponential) generating function and is called the partition function for reasons given by Jaynes (2003, pp. 280–282). The parallel minxent distribution is given by:

(6)
$$p_i = \frac{p_i^o \exp[\lambda_1 f_1(x_i) + \ldots + \lambda_m f_m(x_i)]}{\sum_{i=1}^{n} p_i^o \exp[\lambda_1 f_1(x_i) + \ldots + \lambda_m f_m(x_i)]}$$

where $P^0 = (p_z^0, \ldots, p_n^0)$ is the initial "pre-distribution." (6) reduces to (5) when $p_i^0 = n^{-1}$ is constant.

For a simple discrete example, let the mean for a die be $\langle i \rangle = 4$. Here, $x_i = i$ and $f(x) = x$, so the partition function becomes (with $x = e^\lambda$):

$$Z(\lambda) = \sum_{i=1}^{6} e^{\lambda i} = x(1 + x \ldots + x^5) = x\frac{1 - x^6}{1 - x}$$

Then since $dx/d\lambda = e^\lambda = x$, the chain rule yields:

$$\frac{d\ln Z}{d\lambda} = x\frac{d\ln Z(x)}{dx} = x\left[\frac{1}{x} - \frac{6x^5}{1-x^6} + \frac{x}{1-x}\right]$$

$$= \frac{(1-x)(1-x^6) - 6x^6(1-x) + x(1-x^6)}{(1-x)(1-x^6)}$$

whereupon, setting $d\ln Z/d\lambda = 4$, $3(1 - x)(1 - x^6) = 5x^7 - 6x^6 + x$, which is solved numerically for x to yield, $x = 1.190804264$, or $\lambda = \ln x = 0.1746289309$. Hence, the maxent distribution of mean 4 is:

$$(p_1, \ldots, p_6) = (.103, .123, .146, .174, .207, .247)$$

The general formula for mean $\bar{f}$ works out to $(6 - \bar{f})x^7 - (7 - \bar{f})x^6 + \bar{f}x - \bar{f} + 1 = 0$, and one seeks a root other than $x = 1$. Verify that for mean 3.5, the only root is $x = 1$, so the maxent distribution is uniform. Thus, MINXENT gives back the pre-distribution if it already satisfies the constraint—a redundancy property that should hold generally. Computer programs are available for finding maxent distributions subject to mean value constraints.

The method of Lagrange for constrained extrema is not guaranteed to yield a global maximum, and so to settle this point, one invokes the Gibbs inequality (2) to obtain:

$$\sum p_i \ln(u_i / p_i) \leq \sum p_i \left(\frac{u_i}{p_i} - 1\right) = 0$$

when $\sum u_i = \sum p_i = 1$, whence

$$-\sum p_i \ln p_i \leq -\sum p_i \ln u_i$$

with equality iff $p_i = u_i$. Knowing the form (5) of the maxent distribution, one sets

$$u_i = Z(\lambda_1, \ldots, \lambda_m)^{-1} \exp[\lambda_1 f_1(x_i) + \ldots + \lambda_m f_m(x_i)]$$

and the last inequality then specializes to:

$$H(p_i, \ldots, p_n) \leq \sum_{i=1}^{n} p_i[\ln Z - \sum_{k=1}^{m} \lambda_k f_k(x_i)]$$

$$= \ln Z - \sum_{j=1}^{m} \lambda_j \bar{f}_j$$

with equality iff the p_i are given by (5). This not only shows that (5) is the (one and only) distribution of maximum entropy satisfying the given constraints, but that the right side,

(7)
$$H_{max} = \ln Z(\lambda_1, \ldots, \lambda_m) - \sum_{j=1}^{m} \lambda_j \bar{f}_j$$

is the maximum entropy permitted by those constraints. Any distribution of lower entropy must be importing additional information. Or, as in the example of Wolf's die, if the data distribution is of lower entropy than a hypothesized distribution, an additional condition must be constraining those data to lie in a proper subset of those allowed by the hypothesis in question.

AXIOMATIC CHARACTERIZATION

Every property of MINXENT (MAXENT) rules out potential rivals. Write $P = P^0 \circ C$ for the distribution, P, nearest the pre-distribution, P^0, among all those satisfying a constraint, C (of class C), where for equality or inequality linear constraints the class C is closed and convex and there is a nearest P to P^0. The following properties of MINXENT are then easily proved (Shore and Johnson 1980, §4; Williams 1980):

Uniqueness: The minxent distribution is unique

Redundancy $P \circ C = P$ when P satisfies the constraint C [since $H(Q, P) = 0$ only if $Q = P$]

Chain consistency: The order in which constraints are applied is immaterial, provided each is retained in applying the next one

System independence: The "post-distribution" should not import a dependence between two vari-

ates that is neither implied by the constraints nor the pre-distribution

Invariance: $P = P^0 \circ C$ should not depend on one's choice of a coordinate system

Subset independence: It should not matter whether one obtains a conditional distribution for disjoint subsets of outcomes by finding the post-distribution nearest the conditional prior for each subset or by obtaining the post-distribution for the whole outcome space and conditioning it on each subset

To illustrate the last property, the even and odd faces in tossing a die may be reported separately. Let the prior be uniform and the mean for both the odd and even faces be 4. By redundancy, the post-distribution, *Q*, is also uniform on the even faces: $Q(2) = Q(4) = Q(6) = 1/3$, while for the odd faces, one finds (as above) that $Q(1) = 0.1162$, $Q(3) = 0.2676$, and $Q(5) = 0.6162$. One can also solve the problem by finding the post-distribution for all six states by applying the mean value constraints conjointly obtaining:

$Q'(1) = 0.0524$	$Q'(2) = 0.1831$
$Q'(3) = 0.1206$	$Q'(4) = 0.1831$
$Q'(5) = 0.2778$	$Q'(6) = 0.1831$

To condition on the subsets of odd and even faces, divide each column by its sum and obtain the same conditional distributions found earlier. If these two ways of solving the problem did not agree, MINXENT could be justly deemed inconsistent.

In 1980 John E. Shore and Rodney W. Johnson vindicated Jaynes's conjecture that "deductions made from any other information measure … will eventually lead to contradictions" (1983, p. 9) by deriving MINXENT for mean value constraints from uniqueness and the last three mentioned properties. All these ring changes on the consistency requirement that two ways of doing a calculation must agree, the very condition from which Richard T. Cox derived the basic rules of probability (see the entry "Foundations of Statistics"). Rather than attempt a sketch of their proof, one can illustrate how rival rules violate the Shore-Johnson axioms and lead to contradictions.

INCONSISTENCY OF OTHER RULES

An alternative measure of the spread of a distribution, $\sum p_i^2$, dubbed the repeat rate by Turing, is, like entropy, continuous and assumes its extreme values of $1/n$ and 1 at the extremes of uniformity and concentration. The closely related Gini diversity, $1 - \sum_{i=1}^{n} p_i^2$, has a well-established place in statistics as a measure of the qualitative diversity of a population, with widespread applications to such fields as genetics, ecology, and linguistics. Table 2 shows how closely the distribution of a die of mean 4 obtained by minimizing $\sum p_i^2$, approximates the maxent distribution for this constraint found earlier.

On a superficial examination, one might easily suppose that the repeat rate (RR) rule performs about as well as MAXENT. Still, one knows it must violate one of Shore and Johnson's (1980) consistency postulates, and inconsistency always brings a degradation of performance in its train.

In casting a red and a white die with mean values, $E(R) = 4$ and $E(W) = 3$, and a uniform prior, the method of Lagrange for the RR-rule leads in this case to three linear equations in three unknowns (the Lagrange multipliers), which are easily solved to yield the joint post-distribution:

$$P(R=i, W=j) = \pi_{ij} = \frac{1}{36} + \frac{i-j}{210}$$

from which the marginal distributions for *R* and *W* are computed to be:

$$P(R=i) = p_i = \sum_j \pi_{ij} = \frac{3i+7}{105}$$

and

$$P(W=j) = q_j = \sum_i \pi_{ij} = \frac{-3j+28}{105}$$

One easily checks that $p_i q_i \neq 1/36 = \pi(i,i)$. Hence, the RR-rule violates system independence. Its inconsistency shows up even more clearly when one enlarges the problem to the case of symmetric mean values, $E(R) = 3.5 + \Delta$ and $E(W) = 3.5 - \Delta$. From the joint post-distribution, one can compute a value of Δ, namely, $\Delta = 7/12$ at which the RR-rule makes the smallest of the joint probabilities zero, even though no outcome is excluded by the pre-distribution or the constraints. For values of Δ greater than $7/12$, the RR-rule breaks down completely, making some of the joint probabilities negative.

Consider the more general family of rules that minimize a Csiszar divergence:

$$H_f(P, Q) = \sum q_i f(p_i q_i)$$

with *f* a convex function. This family includes MINXENT as the special case $f(x) = x \ln x$, as well as the chi-squared

rule that minimizes $\sum_{i=1}^{k}\frac{(pi-qi)^2}{qi}$, given by $f(x) = (x - 1)^2$. Applied to a uniform prior, $q_i = 1/n$, this is minimized by minimizing $\sum p_i^2$, hence, it inherits the inconsistencies of the RR-rule.

RELATION TO BAYESIAN CONDITIONING

The upshot of Shore and Johnson's (1980) axiomatic derivation of MINXENT is to place it on a par with Bayes's rule for revising a probability assignment—the most basic rule of belief change. Hence, it is of interest to see what MINXENT delivers in this case (Williams 1980). For any distribution P satisfying $P(B) = 1$:

(8) $$H(P, P^0) = H(P_B, P_B^0) - \ln P^0(B)$$

writing $P_B(A) = P(AB)/P(B)$ for the conditional measure. For then using $\sum_{xi\in B} P(x_i) = P(B) = 1$,

$$H(P_B, P_B^0) = \sum_{x_i\in B} P(x_i)1\text{n}[P(x_i)/\frac{P^0(x_i)}{P^0(B)}]$$

$$= \sum_{x_i\in B} P(x_i)1\text{n}[P(x_i)/P^0(x_i)] + 1\text{n}\,P^0(B)$$

since $P_B(x_i) = P(x_i)/P(B) = P(x_i)$, which proves (8). Since (8) is clearly a minimum when $H(P_B, P_B^0) = 0$, hence, when $(P_B = P_B^0)$, one sees that the distribution P nearest P^0 among those for which $P(B) = 1$ is the Bayesian posterior distribution, P_B^0.

What if the constraint $P(B) = 1$ is weakened to $P(B) = q$, $0 < q < 1$? For any such P one has the following straightforward generalization of (8):

(8a) $$H(P,P^0) = qH(P_B, P_B^0) + \bar{q}H(P_{\bar{B}}, P_{\bar{B}}^0) - q\,1\text{n}\left[\frac{P^0(B)}{q}\right] - \bar{q}\,1\text{n}\left[\frac{P^0(\bar{B})}{\bar{q}}\right]$$

with $\bar{q} = 1 - q$. The right side is minimized by making both $H(P_B, P_B^0)$ and $H(P_{\bar{B}}, P_{\bar{B}}^0)$ zero, which means that the nearest P to P^0 is given by:

$$P = qP_B^0 + \bar{q}P_{\bar{B}}^0$$

a q-weighted average of the conditional distributions for B and its negation. There is an obvious generalization to a partition $B_1, \ldots, B_n$ and the constraint $P(B_i) = q_i$, with $\sum q_i = 1$. This special case of MINXENT is known as Jeffrey conditioning. Indeed, Williams (1980) generalizes further to the case where the B's need not be mutually exclusive. The validity of this rule requires that the sole affect of the datum or sensory input is to raise the probability of B to a value $q < 1$. For conditions under which Jeffrey conditioning is not valid, see Jaynes (2003, §5.6).

TABLE 2

i	1	2	3	4	5	6
Maxent	.103	.123	.146	.174	.207	.247
RR-rule	.095	.124	.152	.181	.209	.238

FREQUENCY CONNECTIONS

MAXENT also has frequency connections (Jaynes 1983). Of the k^N outcomes (i.e., outcome sequences) of N trials, the number that yields category counts $(n_1, \ldots, n_k)$ with $\sum n_i = N$ is given by the multinomial coefficient:

$$W = \frac{N!}{n_1!\ldots n_k!}$$

Using Stirling's approximation to the factorial function, one easily proves that

(9) $$N^{-1}\ln W \to H(f_1, \ldots, f_k)$$

Hence, the maxent distribution is W-maximizing, hence realized by the most outcomes. Moreover, given two sets of relative frequencies, $\{f_i\}$and $\{f_i'\}$,

(10) $$\frac{W}{W'} \sim Ae^{N(H-H')}\left[1 + \frac{B}{12N} + O(N^{-2})\right]$$

gives the ratio of the number of ways each can be realized, where $H = H(f_1, \ldots, f_k)$ and $H' = H(f_1', \ldots, f_k')$, $A = \prod_{i=1}^{k}(f_i / f_i')$ and $B = \sum_{i=1}^{k}\frac{(fi-f_i')}{fif_i'}$. For example, for the two distributions of Table 2, $H = 1.7485056$ and $H' = 1.7470082$, and so at $N = 20{,}000$ trials, the maxent distribution is realized by $W/W' = 9.86 \times 10^{12}$ more outcomes than the similar looking distribution of the RR-rule. The peak is thus enormously sharp. Just how sharp is quantified by Jaynes's concentration theorem (1983), which allows one to compute the fraction of possible outcome sequences whose category counts, f_i, have entropy in the range $H_{max} - \Delta H \leq H(f_1, \ldots, f_k) \leq H_{max}$. This gives, in effect, a new kind of significance test for detecting when additional constraints are hidden in one's data.

Not unrelated results concern typical outcomes of a stationary Markov process (Khinchin 1957). Namely, if H is the entropy of a regular Markov chain and s sufficiently large, then almost all s-step outcomes C satisfy:

$$\left| \frac{\log \frac{1}{c}}{s} - H \right| < \eta$$

with η arbitrarily small. That is, almost all s-step sequences of the process deliver information arbitrarily close to the average, $H^{(s)} = sH$. The entropy is defined in the obvious way as the expectation of

$$H_i = -\sum p_{ik} \log p_{ik}$$

the one-step uncertainty, so that $H = \sum_i P_i H_i$ where $P = (P_l, \ldots, P_k)$ is the stationary distribution and (p_{ik}) is the one-step transition matrix.

Aleksandr Khinchin's prophetic remark, that "the study of entropy will become a permanent part of probability theory" (1957, p. 2), has been borne out not only by the flowering of the maxent method but also by D. S. Ornstein's proof that entropy is a complete invariant of an ergodic Markov chain (for an informal treatment and references to the mathematical literature, see Suppes 2002, §4.5). That is, two ergodic chains (in which any state is reachable from any other) are isomorphic iff they have the same period and the same entropy.

INFORMATION THEORY AND STATISTICAL MECHANICS

The frequency implications of MAXENT play their most important role, however, in statistical mechanics. Thus, Ludwig Boltzmann found the famous Maxwell-Boltzmann energy distribution for molecules in a conservative force field by partitioning the 6n-dimensional phase space of position and velocity (or momentum) coordinates into cells, R_k, small enough for the energy to be a constant, E_k, and large enough to contain a sizeable number, N_k, of molecules. Then the total number, N, and the total energy, E, are constants of the molecular motion. Boltzmann argues that the most probable distribution, $(N_l, \ldots, N_s)$, is the one that is realized by the most microstates among those compatible with the constraints:

$$\sum N_i = N \text{ and } \sum N_i E_i = E$$

By virtue of (5), this most probable distribution,

(11) $$\hat{N}_i = \frac{N}{Z(\beta)} \exp(-\beta E_i)$$

is none other than the maxent distribution for the given constraints.

In this derivation, Boltzmann may be said to have launched MAXENT and the information theoretical approach to statistical mechanics. All the canonical distributions J. Willard Gibbs (1902) later derived are simply maxent distributions for other sets of constraints, for example, that for fixed values of the total energy and angular momentum is Gibbs's rotational ensemble. Unfortunately, Gibbs slipped in the logical basis of his derivation so unobtrusively that most readers missed it. Moreover, he provided no clear or compelling rationale, so that in their famous review article of statistical mechanics published a decade later, Paul Ehrenfest and Tatiana Ehrenfest (1912) dismissed Gibbs's method of derivation as "a mere analytic trick" (for the relevant history, see Jaynes 1983, pp. 98ff). Thus began a long siege of confusion and controversy over the justification of Gibbs's formalism that continues to this day, notwithstanding Jaynes's rediscovery of the MAXENT method of Gibbs in 1957 and his clear rationale for using it (1983, chapter 1). In particular, the information theoretical approach dispenses with the ergodic hypothesis.

Jaynes's second great contribution was to extend the Gibbsian (MAXENT) formalism to irreversible processes and nonequilibrium thermodynamics (1983, chapter 10, §D). He writes:

> The final breakthrough came in the Christmas vacation period of 1962 when, after all else had failed, I finally had the courage to sit down and work out all the details of the calculations that result from using the Maximum Entropy Principle; *and nothing else*. Within three days the new formalism was in hand, masses of the known correct results of Onsager, Wiener, Kirkwood, Callen, Kubo, Mori, MacLennon, were pouring out as special cases, just as fast as I could write them down; and it was clear that this was it. (p. 239)

The unbelievably short derivation of (11) as the equilibrium distribution of the energies has seemed too short to many physicists and philosophers. This initial impression is only reinforced when it is seen that (11) implies the familiar barometric formula,

(11a) $$\rho(z) = \rho(0) \exp(-\beta mgz)$$

for the density of the atmosphere at height z, as well as Maxwell's velocity distribution in the form:

(12) $$f(v_b) = B(\beta)\exp(-\beta m v_b^2/2)$$

where $f(v_b)dv_b$ is now the probability that a molecule has a velocity in a tiny neighborhood of v_b. Indeed, even more follows, for (12) does not depend on the position of the molecule (an assumption Boltzmann was forced to make in his derivation of the Maxwell distribution from his collision equation). This implies, in turn, the dynamic stability of the distribution. However, the MAXENT derivation seems to ignore the dynamics altogether.

Jaynes responded in several ways. First, the derivation does not ignore the dynamics; it uses conservation of energy as well as the preservation of the volumes of the cells, R_k, under evolution of the system (Liouville's theorem). Jaynes emphasizes that, in addition, one is trying to predict reproducible macrostates. These are *ipso facto* under the experimenter's control, and so the myriads of details concerning microstates not under his control must needs be irrelevant for prediction. Moreover, reproducible macroscopic properties must be characteristic of the overwhelming preponderance of microstates in the allowed region of the phase space. Given the large number of degrees of freedom entailed, the maxent distribution will be enormously peaked. Hence, predictions of other macro quantities based on their mean values will be correct (within experimental error) with probability close to one (Jaynes 1983). As far as Jaynes is concerned, these considerations fully explain the predictive success of equilibrium thermodynamics as "inferences from the available information."

In particular, there is no need to appeal, as Maxwell and Boltzmann did (but Gibbs did not) to the equality of infinite time averages and averages with respect to the canonical distribution. Even if this could be established from some other easily verified assumption—the program of ergodic theory—it would be nothing to the purpose. For one would have to show, in addition, that the averages over finite time intervals involved in measuring macro quantities closely approximate their infinite time averages, and there are positive reasons to doubt this (Jaynes 1983).

Apart from the clarity, unity, and simplicity the information theoretical approach brings to the foundation of statistical mechanics, David Hestenes (1993) considers Jaynes's greatest merit to lie in his recognition that "in the evolution of statistical mechanics the principles of physics had gotten confused with principles of statistical inference" (1993, p. 153).

Jaynes regards the formalism of quantum mechanics as a similar "nasty omelet" scrambling together properties of physical systems and our information about them in ways that are difficult to unscramble. See the cited article by David Hestenes for one noteworthy attempt to disentangle subjective and objective aspects of the electron wave function, namely, a probability factor and a kinematic factor, using a powerful "universal geometric calculus" based on Hermann Grassmann's Ausdehnumgslehre. Hestenes purports to show, in particular, that the complex "probability amplitudes" of the formalism have nothing to do with probability per se but have, instead, a physical origin in the "Zitterbewegung", or circular dance, of the electrons generating their spin and magnetic moment.

Jaynes's views on quantum mechanics are outlined (with references) in the same volume—a festschrift in his honor—in which the article by Hestenes appears. For other views, see Philosophy of Statistical Mechanics and Quantum Mechanics.

See also Bayes, Bayes's Theorem, Bayesian Approach to Philosophy of Science; Bell, John, and Bell's Theorem; Bohr, Niels; Boltzmann, Ludwig; Copenhagen Interpretation; Einstein, Albert; Gibbs, Josiah; Maxwell, James Clerk; Philosophy of Statistical Mechanics; Probability; Quantum Mechanics; Semantics; Turing, Alan M.

Bibliography

Buck, Brian, and Vincent A. Macaulay, eds. *Maximum Entropy in Action*. Oxford, U.K.: Clarendon Press, 1991.

Cover, Thomas M., and Joy A. Thomas. *Elements of Information Theory*. New York: Wiley, 1991.

Erickson, Gary J., and C. Ray Smith, eds. *Maximum-Entropy and Bayesian Methods in Science and Engineering*. 2 vols. Dordrecht, Netherlands: Kluwer Academic, 1988.

Gibbs, J. Willard. *Elementary Principles of Statistical Mechanics*. New Haven, CT: Yale University Press, 1902.

Good, Irving John. *Good Thinking*. Minneapolis: University of Minnesota Press, 1983.

Hestenes, David. "The Kinematic Origin of Complex Wave Functions." In *Physics and Probability: Essays in Honor of Edwin T. Jaynes*, edited by W. T. Grandy Jr. and P. W. Milonni, 153–160. New York: Cambridge University Press, 1993.

Jaynes, Edwin T. "A Backward Look to the Future." In *Physics and Probability: Essays in Honor of Edwin T. Jaynes*, edited by W. T. Grandy and P. W. Milonni, 261–275. New York: Cambridge University Press, 1993.

Jaynes, Edwin T. "Clearing up Mysteries—The Original Goal." In *Maximum Entropy and Bayesian Methods*, edited by John Skilling. Dordrecht, Netherlands: Kluwer Academic, 1989.

Jaynes, Edwin T. *Papers on Probability, Statistics, and Statistical Physics*, edited by R. D. Rosenkrantz. Dordrecht, Netherlands: D. Reidel, 1983.

Jaynes, Edwin T. *Probability Theory: The Logic of Science*. New York: Cambridge University Press, 2003.

Khinchin, Aleksandr. *Mathematical Foundations of Information Theory*. New York: Dover, 1957.

Kullback, Solomon. *Information Theory and Statistics*. New York: Wiley, 1959.

Shannon, Claude, and Warren Weaver. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1949.

Shore, John E., and Rodney W. Johnson. "Axiomatic Derivation of Maximum Entropy and Minimum Cross Entropy." *IEEE Transactions on Information* IT-26 (1980): 26–37.

Suppes, Patrick. *Representation and Invariance of Scientific Structures*. Stanford, CA: CSLI Publications, 2002.

Williams, Paul M. "Bayesian Conditionalization and the Principle of Minimum Information." *British Journal for the Philosophy of Science* 31 (1980): 131–144.

Roger D. Rosenkrantz (2005)

INFORMED CONSENT

A fundamental requirement of both ethics and the law is that medical treatment cannot be given to competent patients without their "informed consent." This represents a rejection of more traditional authoritarian or paternalistic accounts of the physician/patient relationship in which the physician had decision-making authority in favor of a process of shared decision making between physicians and patients. In this respect informed consent helps shape the nature of nearly all health-care treatment decision making. Informed consent also has special importance in a narrower class of cases in which patients and their physicians are unable to agree on a course of treatment. In these cases a competent patient is given the right to refuse any recommended treatment, even including life-sustaining treatment, no matter how strongly the physician or others believe that the treatment should be undertaken.

There are two principal moral values that are served by and justify the informed-consent requirement in health care. The first is patient well-being—arguably the fundamental goal of all health care. The concept of patient well-being, as opposed to the apparently more objective goals of protecting and promoting patients' health and lives, signals the important respect in which what will best serve a particular patient's well-being is often to a significant degree a subjective determination that depends on the particular aims and values of the patient in question. Increasingly, there are medically acceptable alternative treatments (including the alternative of no treatment), no one of which is best for all patients with a particular medical condition. The patient's participation in decision making is therefore necessary in order to select the treatment that best fits his or her aims and values. The other fundamental moral value that undergirds the informed-consent requirement is individual self-determination or autonomy. Self-determination in this context is the moral right of ordinary persons to make significant decisions about their lives for themselves and according to their own aims and values. Requiring that health care not be rendered without a competent patient's informed consent respects this right of self-determination. The informed-consent requirement reflects the fundamental moral point that it is the patient and the patient's body that undergo the treatment, and so it should be the patient who is morally entitled to authorize or refuse the treatment.

Three conditions are necessary for ethically valid informed consent—that the patient's decision be informed, voluntary, and competent. The requirement that the decision be informed places a responsibility on the patient's physician to provide the patient with information, in an understandable form, about the patient's condition or diagnosis and the prognosis if no treatment is provided, together with the alternative treatments that would improve that prognosis, along with their risks and benefits. This typically does not require that the physician provide, or that the patient understand, complex medical and scientific information, but rather information about how the various alternatives would likely affect the patient's pursuit of his or her plan of life. Legal requirements regarding how much and which information must be provided vary, but the ethical ideal is to provide the information that the particular patient would reasonably want to know in order to make his or her decision.

The requirement that the consent be voluntary means that treatment must not be rendered against the patient's will, either by force or by coercing the patient's choice. More important, it also forbids physicians from manipulating the patient's choice through selective provision of information, playing on the patient's fears, and other means. Ethically objectionable manipulation, as opposed to appropriate informing and persuasion, aims to produce a different choice from what a competent patient would have made if fully informed and freely choosing.

The third requirement of competence is the most complex. Usually, patients are either clearly competent, with their normal decision-making capacities intact, or

clearly incompetent, unable to make any decision. In borderline cases in which there is significant, but not total, impairment of the patient's decision-making capacities, the competence determination is often controversial. The competence evaluation should address the process of the patient's decision making in order to determine whether there are significant impairments, limitations, or mistakes in that process that have resulted in a choice different from what the patient would have wanted in the absence of those impairments, limitations, or mistakes. The proper standard of competence in borderline cases is controversial but increasingly understood to be a variable standard, requiring a higher level of understanding and reasoning when the patient's well-being would be seriously affected by the decision in question and a lower level when there would be only limited impact on the patient's well-being. While treatment refusal may reasonably trigger an evaluation of the patient's competence, it should not serve as any evidence of the patient's incompetence—that evidence must come from impairments or limitations that cannot be remedied in the process of the patient's reasoning. When the requirements for ethically valid informed consent (that is, informed, voluntary, and competent) are met, the patient's choice should be reasonably in accord with his or her well-being, and his or her self-determination will have been respected.

When the patient has been determined to be incompetent to make his or her own treatment choices, a surrogate or proxy, typically a close family member, should substitute for the patient in the decision-making and consent process. The patient's informed consent is also not required in emergency conditions, when taking time to obtain consent would involve serious risks to the patient's well-being, or when the patient has waived his or her right to give consent and has authorized another to make the treatment decision.

See also Applied Ethics; Bioethics; Biomedical Ethics; Euthanasia; Medical Ethics.

Bibliography

Appelbaum, P. S., C. W. Lidz, and A. Meisel. *Informed Consent: Legal Theory and Clinical Practice.* New York: Oxford University Press, 1987.

Faden, R., and T. L. Beauchamp. *A History and Theory of Informed Consent.* New York: Oxford University Press, 1986.

Dan W. Brock (1996)

INFORMED CONSENT IN THE PRACTICE OF LAW

The doctrine of informed consent has deeper historical roots in the practice of law than in the practice of medicine. The modern rule that medical treatment cannot be given without the informed consent of competent patients did not arise until the late twentieth century, whereas a century earlier, courts already recognized that, like other agents, lawyers may breach their fiduciary duty to client-principals when they fail to provide them with sufficient information. In some respects, however, the legal profession may have lagged behind the medical profession. Thus lawyers, who played a critical role in developing the informed consent model in medicine, have been criticized for failing to adequately develop such a model for their own practice.

The legal profession has long recognized that lawyers may not reveal confidential information or represent conflicting interests unless the client consents after full disclosure. Because it has not always been clear precisely what information the lawyer must disclose to the client before acting, the American Bar Association (ABA) amended its Model Rules of Professional Conduct in 2002 to require the lawyer to obtain the client's *informed consent*, taking advantage of lawyers' familiarity with that term in the context of medical practice. The Model Rules also draw on language familiar in the medical context in defining informed consent to require the communication of "adequate information and explanation about the material risks of and reasonably available alternatives to the proposed course of conduct" (ABA *Model Rules*, Rule 1.0[e]).

Aside from confidentiality and conflicts of interest, the use of informed consent doctrine in legal practice is unclear. The difficulty lies in determining which decisions are for the client (in which case lawyers are required to provide sufficient explanation for the client to make informed decisions) and which decisions are for the lawyer (with or without consulting the client). Certain decisions are clearly reserved for the client—whether to accept a settlement offer and, in criminal cases, what plea to enter, whether to waive a jury trial, and whether the client will testify. In addition, there is agreement that the client defines the objectives of the representation. What remains unclear is how to distinguish between a client's objectives and the means of achieving those objectives (for example, when the client wants to win a lawsuit by asserting a particular right), how properly to allocate decision-making as to the means of the representation

(for example, when the client does not want to pay for additional depositions), and whether there are some decisions lawyers may make without first consulting the client (for example, deciding what questions to ask a witness on direct or cross-examination).

As initially promulgated, the ABA *Model Rules* appeared to suggest that all means decisions were for the lawyer to make after consultation with the client. Aside from the difficulty of distinguishing between objectives and means, many believe that there are some means decisions that ought to be for the client to decide, particularly when they involve expenses or concern for third persons, as opposed to technical, legal and tactical matters. On the other hand, it would be inefficient to require the lawyer to consult the client prior to taking any action, especially when the matter is in trial. Case law has not resolved these issues; some courts state that a lawyer must obey all lawful instructions of a client, while others adhere to the traditional distinction between ends and means.

The amended *Model Rules* provide that lawyers may act without prior consultation when the action is "impliedly authorized" to carry out the representation (Rule 1.2[a]), as when the lawyer reasonably assumes that the client would not want to be consulted because the matter is highly technical and does not involve significant risk to the client. As to other means decisions, the *Model Rules* require the lawyer to reasonably consult the client, but do not prescribe how to resolve disagreements between lawyer and client, other than suggesting that the lawyer may withdraw from the representation when the lawyer has a fundamental disagreement with the client and that, conversely, the client may resolve the disagreement by discharging the lawyer.

A similar problem exists in medical practice. Historically, the failure to obtain the patient's informed consent to surgery or some other invasive procedure was treated in law as a battery. In the 2000s, most courts agree that the cause of action is better understood as deriving from the patient's right to self-determination rather than the right to reject a nonconsensual touching; therefore, they base the informed consent action in negligence rather than battery. The question arises, however, whether the physician must obtain the patient's consent to all treatment (and perhaps diagnostic) options, even when the proposed treatment is noninvasive, such as bed rest, and one or more of the options is not one that the physician would recommend. Some courts hold that a physician may not subject a patient to a course of treatment, whether invasive or noninvasive, without disclosing information that will enable the patient to intelligibly evaluate the available options and risks of each.

These courts also recognize, however, that it would be unduly burdensome to require physicians to explain in detail all treatment options in every case, such as when a physician proscribes one of several potentially appropriate antibiotics in treating a respiratory infection. Other courts refuse to apply informed consent doctrine to procedures or diagnostic options not recommended by the physician or to situations where the patient refuses the recommended treatment, fearing that recognizing such a duty would in effect require physicians to give a minicourse in medical science and would further suggest that physicians should defer their medical judgment to the patient's wishes.

The difficulty of determining when to use informed consent doctrine stems from the need to balance the individual's right to self-determination with other concerns. Like physicians, lawyers resist being forced to give detailed explanations of every exercise of professional judgment, because some explanations entail either excessive costs or unwarranted invasions of professional autonomy. Obviously there must be some limits to the reach of informed consent doctrine. One approach is to draw admittedly arbitrary lines between the objectives and means of a representation, or between invasive and noninvasive procedures (or recommended and nonrecommended treatment or diagnostic options). Another approach is to reject arbitrary line-drawing in favor of fact-intensive, case-by-case determinations of the proper allocation of decision-making between professionals and consumers, guided by the decisions reasonable consumers would presumably want to make. In any event, the precise boundaries of informed consent doctrine, whether in legal or medical practice, continue to be debated.

See also Contractualism; Discourse Ethics; Medical Ethics; Philosophy of Law, History of; Philosophy of Law, Problems of.

Bibliography

American Bar Association. *Model Rules of Professional Conduct.* Chicago: Author, 2004.

Burns, Robert P., and Steven Lubet. "Division of Authority Between Attorney and Client: The Case of the Benevolent Otolaryngologist." *University of Illinois Law Review* (2003): 1,275–1,297.

Martyn, Susan R., and Lawrence J. Fox. *Traversing the Ethical Minefield: Problems, Law, and Professional Responsibility.* New York: Aspen, 2004.

Morries, Grant H. "Dissing Disclosure: Just What the Doctor Ordered." *Arizona Law Review* 44 (2002): 313–371.

Strauss, Marcy. "Toward a Revised Model of Attorney-Client: The Argument for Autonomy." *North Carolina Law Review* 65 (1987): 315–349.

Nancy J. Moore (2005)

INGARDEN, ROMAN

(1893–1970)

Roman Ingarden, the Polish phenomenologist, was born in Kraków. He studied philosophy under Kazimierz Twardowski at Lvov and under Edmund Husserl at Göttingen. At Göttingen he also studied mathematics under David Hilbert and psychology under G. E. Müller. Ingarden followed Husserl to Freiburg, where he received his PhD in 1918 with the dissertation "Intuition und Intellekt bei Henri Bergson." The same year Ingarden returned to Poland, where he taught mathematics in high schools. After his habilitation in 1921 he was named Privatdozent in philosophy at the University of Lvov. During the German occupation Ingarden was basically preoccupied with writing "Controversy over the Existence of the World"; universities in Poland were closed at that time. In 1945 he accepted the chair of philosophy at the Jagellonian University at Kraków. During the early 1950s the Polish government barred him from teaching philosophy because of his adherence to "idealism"; during this period he translated Immanuel Kant's *Critique of Pure Reason* into Polish. Ingarden regained his chair in 1956 and retired in 1963, but he continued to be philosophically active.

Ingarden was one of the ablest pupils of Husserl. He accepted Husserl's main analytical results and the phenomenological method, but he rejected Husserl's transcendental idealism, showing instead how phenomenology could lead to realism. Max Scheler, Jean Hering, and, in her earlier works, Hedwig Conrad Martius, also exerted some influence on Ingarden. Traces of Ingarden's ideas can be found in the work of Nicolai Hartmann, Herbert Spiegelberg, and Michel Dufrennes, as well as in that of such American aestheticians as René Wellek.

Ingarden's philosophy is a fusion of two traditions: the variety of German speculative metaphysics as represented by Franz Brentano and the restrained and painstaking Polish analytical philosophy. Ingarden wove grand philosophical designs, but he wove them with great care and clarity. He opposed what he regarded as the narrowness and one-sidedness of the analytical trend, and he was probably the first to argue (in 1934) that the logical positivist verification principle of meaning, since it is a metalanguage statement, is itself unverifiable; and since it is not analytic, it is therefore meaningless. Ingarden followed this criticism with many others, but he nevertheless acquired and used the skills and techniques of the analytical philosophers. His phenomenology is therefore marked by an intelligibility and clarity rare among metaphysicians and ontologists.

AESTHETICS

Ingarden's earliest work was in epistemology, which he conceived of as an independent discipline able to show the certainty of its own conclusions. The center of his investigations later shifted to ontology, which he regarded as a science of pure possibilities. Ontology determines and describes these possibilities in order to provide us with conceptual apparatuses by which we can express various existential situations.

Ingarden also conducted significant work in aesthetics. His fully elaborated and original theory of art is perhaps the best-known part of his philosophy. He arrived at this theory through the ontological investigations that were central to his thought, and the theory itself was a preparation for his realistic ontology. One of the possible ways of settling the controversy between idealism and realism is through examining the nature of objects that exist. There seems to be a necessary connection between a mode of being and its formal structure. Ingarden first attempted to investigate this problem through examining works of art, which, in contrast with spatiotemporal objects, are dependent for their existence on the conscious act of the creator but which nevertheless transcend this act and continue to exist in their material shape afterward. What makes them works of art is the intention of the creator to endow them with significance, and it requires another intentional act on the part of the receiver to decipher this significance expressed by physically perceptible signs. Thus, the work of art possesses many strata. In a literary work of art, for example, the following can be distinguished: (1) the visual or phonic stratum; (2) the stratum of the meanings of words and sentences; (3) the stratum of objects described; (4) the stratum of the appearances of these objects. All these strata are polyphonically orchestrated to compose one work of art. In a poem it is not the printed marks in the shape of letters, nor even the actual meanings of particular words, that matter; rather, it is the "poetic significance" achieved through these printed marks and through the meanings of particular words. The inten-

tional act of the creator and another intentional act of the receiver are indispensable for the existence of the work of art. And because of this, works of art are called purely intentional objects.

ONTOLOGY

It is customary to link phenomenology with existentialism, as if Husserl, Martin Heidegger, and Jean-Paul Sartre were three links in the development of one trend and as if existentialism were an inevitable development of phenomenology. But the linking of phenomenology and existentialism in this manner blurs the fact that for Husserl phenomenology was primarily a cognitive philosophy, seeking to acquire knowledge, whereas for Sartre the main function of philosophy was consolatory, to explain the mystery of man and justify his tragic existence. Ingarden's philosophy was a continuation, development, and restatement of the cognitive core of Husserl's philosophy, and perhaps was closer to its cognitive spirit than any other development of Husserl's doctrine by his numerous pupils. Ingarden perhaps succeeded better than Husserl himself in making his phenomenological inquiries consistent and coherent. Husserl, as Ingarden observed, was entangled in a vicious circle of phenomenology: In order to conduct the phenomenological reductions that are to yield self-evident knowledge, Husserl had to assume that our consciousness is transcendental, whereas it is precisely through application of the phenomenological reductions that consciousness is revealed to be transcendental.

Ingarden attempted to break away from this circle by what he called eidetic analyses, the penetration of the nature of essences in an "objective" way, as opposed to the transcendental approach Husserl used in his later work. Ingarden's objective approach was to clear the ground for philosophy as an independent and self-sufficient discipline. He contended that any reconstruction of our knowledge must start from thorough analyses of the nature of the objects of our knowledge, both existing and possible. Ontology is basic to other philosophical endeavors because the manner of our cognizing is determined by the objects of cognition. It follows that there are as many types of immediate experience as there are types of objects and types of relationships occurring among objects.

Ingarden devoted his principal work, *Spór o Istnienie Świata* (The controversy over the existence of the world), to the analysis of these various objects and relationships. According to Ingarden, existence is not that which exists but that by means of which something exists. Not everything that can be distinguished in an object belongs to its attributes: Existence is not an attribute of an object. Ingarden attempted to account for the specific role of existence in whatever is, by distinguishing between modes of being (*modus existentiae*) and existential moments (*momentum existentiale*). The real existence (reality) of something, the possibility of something, and the ideal existence of something are examples of modes of being (modes of existence). Nonexistence, however, is not a mode but the absence of any being. An existing object can never be experienced by us without its mode of being. In every mode of being we can distinguish existential moments. The existential moments are the elemental units of the modes and thus are the key to understanding them. Many different existential moments can be distinguished intuitively in each mode of being of something. What we grasp in the object is not existence as such, which is a certain universal idea, but particular existential moments.

Ingarden divided moments of being into mutually exclusive pairs. There are four basic pairs. The first pair comprises existential autonomy and existential heteronomy. "Something is self-existent (is existentially autonomous) if it has its existential foundation *in itself*. It has such a foundation if it is *immanently determined* in itself" (*Time and Modes of Being*, p. 43). Otherwise it is existentially heteronomous. "An object is *existentially original* if, in its essence, it cannot be produced by any other object" (ibid., p. 52). If it can be so produced, it is existentially derivative. "An object is existentially separate if, *for its existence, it does not in its essence require the existence of any other object with which it would have to coexist, because of its essence, within the compass of one and the same whole*" (ibid., p. 82). If it does require such another object, it is "inseparate." The fourth pair of existential moments are existential self-dependence and existential contingency. Existential contingency involves separate objects that, in spite of being separate, require for their existence some other existentially separate object. An existentially self-dependent object, which is also an existentially separate object, does not require such another object.

Ingarden discussed at length both time and causality. In the analysis of time he distinguished further pairs of existential moments, including actuality and nonactuality, persistence and fragility, fissuration and nonfissuration. His original interpretation of the causal relation arose out of his analysis of the moments of existential originality and existential derivation. For Ingarden a causal relation occurs between *C* and *E* if: (1) *C* and *E* are

diverse; (2) *C* actually conditions *E* but *E* does not condition *C* in the same way; (3) both *C* and *E* are events or processes (as far as their form is concerned); (4) the occurrence of *E* is simultaneous with that of *C*; (5) both *C* and *E* are real actual).

Modes of being consist of noncontradictory combinations of existential moments. Ingarden distinguished four basic modes, or regions, of being: absolute being, temporal (or real) being, ideal (or extratemporal) being, and purely intentional being. Absolute being is characterized by the existential moments of autonomy, originality, separateness, and self-dependence. The other modes have many subtypes, each of which is characterized by a number of existential moments.

Each of Ingarden's analyses of pairs of existential moments is a small monograph on traditional ontological problems usually rooted in Aristotle and scholastic philosophy. On one level they may appear to be analyses of language, as one linguistic philosopher has pointed out, but they are of a scope not generally undertaken by linguistic philosophers, and Ingarden regarded linguistic analysis as an inadequate tool for the systematic analysis of philosophical problems. The analyses contained in this work were to pave the way for the eventual solution of the controversy between idealism and realism over the nature of the world and our relation to it. They follow in many instances the spirit of Aristotle's analysis of categories, but to be fully comprehended they presuppose familiarity with medieval discussion of pure possibilities.

See also Aesthetics, History of; Aesthetics, Problems of; Aristotle; Brentano, Franz; Existence; Hartmann, Nicolai; Hilbert, David; Husserl, Edmund; Idealism; Kant, Immanuel; Ontology; Scheler, Max; Time; Twardowski, Kazimierz.

Bibliography

WORKS BY INGARDEN

Gesammelte Werke, Tübingen: Max Niemeyer, 1992 ff.

Epistemology

"Über die Gefahr einer Petitio Principii in der Erkenntnistheorie." *Jahrbuch für Philosophie und phänomenologische Forschung* 4 (1921): 545–568.

"Intuition und Intellect bei Henri Bergson." *Jahrbuch für Philosophie und phänomenologische Forschung* 5 (1922): 286–461.

"Essentiale Fragen. Ein Beitrag zum Problem des Wesens." *Jahrbuch für Philosophie und phänomenologische Forschung* 7 (1925): 125–304.

Ontology

Spór o Istnienie Świata (The controversy over the existence of the world). 2 vols. Kraków: Polish Academy of Sciences, 1947–1948. Vol. I partially translated by Helen R. Michejda as *Time and Modes of Being*. Springfield, IL: Charles C. Thomas, 1964. A German version appeared as *Der Streit um der Existenz der Welt*. 3 vols. Tübingen, 1964–1966. The third volume is not published in Polish.

Aesthetics

Das literarische Kunstwerk. Eine Untersuchung aus dem Granzgebiet der Ontologie, Logik und Literaturwissenschaft. Halle: Niemeyer, 1931.

O Poznawaniu Dzieta Literackiego (On comprehending the work of literature). Lvov: Ossolineum, 1937.

Studia z Estetyki (Studies in aesthetics). 2 vols. Warsaw: Polish Scientific, 1957–1958.

"Prace Filozoficzne Romana Ingardena." In *Szkice Filozoficzne Romanowi Ingardenowi w Darze*. Warsaw: Polish Scientific, 1964. Lists 150 works by Ingarden written between 1915 and 1963.

TRANSLATIONS OF WORKS BY INGARDEN

The Literary Work of Art: An Investigation on the Borderlines of Ontology, Logic and Theory of Literature. With an Appendix on the Functions of Language in the Theatre. Translated and with an introduction by George G. Grabowicz. Evanston, IL: Northwestern University Press, 1973.

The Cognition of the Literary Work of Art. Translated by Ruth Ann Crowley and Kenneth R. Olson. Evanston, IL: Northwestern University Press, 1973.

On the Motives which Led Husserl to Transcendental Idealism. Translated from Polish by Arnòr Hannibalsson. The Hague: Nijhoff, 1975.

Man and Value. Translated from Polish and German by Arthur Szylewicz. Washington, DC: Catholic University of America Press, 1983.

The Work of Music and the Problem of Its Identity. London: Macmillan, 1986.

The Ontology of the Work of Art: The Musical Work, the Picture, the Architectural Work, the Film. Translated by Raymond Meyer with John T. Goldthwait. Athens: Ohio University Press, 1989.

WORKS ON INGARDEN

Ameriks, Karl. "Husserl's Realism." *The Philosophical Review* 86 (1977): 498–519.

Bostar, Leo. "Reading Ingarden Read Husserl: Metaphysics, Ontology, and Phenomenological Method." *Husserl Studies* 10 (1994): 211–236.

Dziemidok, Bohdan, and Peter McCormick, eds. *On the Aesthetics of Roman Ingarden: Interpretations and Assessments*. Dordrecht: Kluwer Academic Publishers, 1989.

Gierulanka, D., and A. Polatowski. "Kirunki Badan Filozoficznych Romana Ingardena." In *Szkice Filozoficzne Romanowi Ingardenowi w Darze*. Warsaw: Polish Scientific, 1964.

Küng, Guido. "Roman Ingarden (1893–1970): Ontological Phenomenology." In *The Phenomenological Movement*. Edited by Herbert Spiegelberg. The Hague: Martinus Nijhoff, 1982, pp. 223–233.

Rudnick, Hans H., ed. "Ingardeniana II: New Studies in the Philosophy of Roman Ingarden, with a New International Ingarden Bibliography." *Analecta Husserliana* 30 (1990).

Smith, Barry. "Roman Ingarden: Ontological Foundations for Literary Theory." In *Language, Literature and Meaning*. Edited by John Odmark. Amsterdam: Benjamins Press, 1979, pp. 373–390.

Stepien, A. "O Filozofii Romana Ingardena." *Ruch Filozoficzny* 22 (1963–1964): 153–159.

Tymieniecka, A. T. "Le dessin de la philosophic de Roman Ingarden." *Revue de métaphysique et de morale* 60 (1955): 32–57.

Tymieniecka, A. T. *Essence et existence: Étude à propos de la philosophie de Roman Ingarden et Nicolai Hartmann*. Paris, 1957.

Tymieniecka, A. T., et al. *For Roman Ingarden. Nine Essays in Phenomenology*. The Hague: Nijhoff, 1959.

Henryk Skolimowski (1967)
Bibliography updated by Thomas Nenon (2005)

INGE, WILLIAM RALPH

(1860–1954)

William Ralph Inge, the English ecclesiastic and religious thinker, was born at Crayke, Yorkshire. Educated at Eton and Cambridge, he was fellow of Hertford College, Oxford, from 1889 to 1905, vicar of All Saints Church, Knightsbridge, from 1905 to 1907, and professor at Cambridge from 1907 to 1911, when he was appointed dean of St. Paul's Cathedral, London. During his long tenure of this high office, he became one of the best-known Englishmen of his generation. He continued his lifelong studies in philosophy and mysticism, and his penetrating comments on the events of his time, especially on the foibles of contemporary civilization, earned him the sobriquet of "the gloomy dean." He retired in 1934 to Brightwell Manor, Berkshire, where he spent twenty years more of thought and activity before his death.

What provoked Inge's criticism of contemporary culture was its preoccupation with material progress; against this, he pleaded for an end to the separation of fact and value. He maintained that Plato taught an abiding truth when he instructed us to seek reality beyond what is present to the senses; and only a culture that is based on the invisible but eternal values of truth, beauty, and goodness is securely founded. These values are in turn grounded in God, the ultimate spiritual reality, so that Inge's plea was for a religious attitude toward life. The model for such an attitude is provided by the mystic, who penetrates the phenomena of the sensible world to the realm of values and whose soul ascends toward union with God. However, this advocacy of mysticism is not to be understood as escapism or as a denial of the reality of the world of the senses. Inge considered himself in some ways more of a realist than an idealist, and he insisted that any adequate philosophy must take account of the findings of the natural sciences. Mysticism, as he understood it, does not imply emotionalism or irrationalism. Mysticism is itself a kind of spiritual philosophy, a quest for knowledge of the real. If today there is a conflict between the rational and the religious approaches to reality, this is because modern rationalism has become too narrow in its understanding of reason. A genuine rationalism takes account of values as well as of facts; this is the kind of rationalism that flourished in the earlier tradition of Western philosophy, and such a broadly based rational philosophy conduces to the same results as the mystical insights of religion. Both lead, Inge claimed, to "perfect knowledge of the Perfect."

Inge steeped himself in the history of mystical and religious thought, but there was one particular school that seemed to him to approach his ideal of combining genuine rationalism with mystical insight and that therefore strongly attracted him: the Neoplatonism of Plotinus. Inge spoke of Plotinus in terms of almost exaggerated respect as not merely an intellectual teacher but also a spiritual director, and he studied his philosophy not just as a historical phenomenon but also as the classic statement of the insights that have guided Western culture—and thus as a message for our time. Platonism, Christianity, and Western civilization, Inge believed, are inseparable and interdependent; and a restatement of the philosophy of Plotinus can provide an intellectual basis that, when combined with the spirit of Christianity, can lead to the rejuvenation of the West.

See also Beauty; Mysticism, History of; Neoplatonism; Plato; Platonism and the Platonic Tradition; Plotinus; Rationalism; Truth.

Bibliography

WORKS BY INGE

Christian Mysticism. London: Methuen, 1899. Given as the Bampton Lectures.

Faith and Its Psychology. London: Duckworth, 1909.

The Philosophy of Plotinus, 2 vols. London: Longman, 1918. Given as the Gifford Lectures.

Outspoken Essays, 2 series. London: Putnam, 1919–1922.

Lay Thoughts of a Dean. London: Putnam, 1926.

The Platonic Tradition in English Religious Thought. London: Longman, 1926.

Christian Ethics and Modern Problems. London: Hodder and Stoughton, 1930.

God and the Astronomers. London: Longman, 1933.

Diary of a Dean. London: Hutchinson, 1934.
Mysticism in Religion. London: Hutchinson, 1947.

WORKS ON INGE

Crook, Paul. "W. R. Inge and Cultural Crisis, 1899–1920." *Journal of Religious History* 16 (1991): 410–432.

Fox, Adam. *Dean Inge.* London; Murray, 1960.

Helm, Robert M. *The Gloomy Dean.* Winston-Salem, NC: John F. Blair, 1962.

John Macquarrie (1967)
Bibliography updated by Christian B. Miller (2005)

INGENIEROS, JOSÉ
(1877–1925)

José Ingenieros, the Argentine positivist metaphysician and ethical philosopher, was born in Buenos Aires. He studied, successively, medicine, psychiatry, axiology, and metaphysics and held appointments on the faculties of medicine and of philosophy and letters in Buenos Aires; he also founded the *Revista de Filosofía.* Ingenieros lived for some years in Germany and Switzerland. He had great influence in Latin America, and some of his works were translated into several European languages.

In *Proposiciones relativas al porvenir de la filosofía* (Buenos Aires, 1918), Ingenieros set forth a prospectus for a metaphysics of the "inexperiential." By the "inexperiential" he did not mean a transcendent object but those parts of the natural world that the limitations of the senses and instruments exclude from present experience. He rejected the "classical" problems of the existence and nature of God, immortality, and freedom, finding them to be not so much meaningless as falsely stated under the influence of theological and ethical orthodoxy. The legitimate problems of metaphysics are those of metacosmology, metabiology, and metapsychology; in metabiology, for example, some legitimate problems are the origin of life, the possibility of life beyond this planet, and the final purpose of life. Because its objects lie beyond experience, metaphysics cannot achieve certainty. Its statements are hypotheses, which must be logically consistent and compatible with experience. Like the sciences, the metaphysics of the future will be antidogmatic, tentative and indefinitely perfectible, and impersonal in the sense that it will be the work of many collaborators.

The ethics of Ingenieros, discussed with visionary enthusiasm in *El hombre mediocre* (Madrid, 1913), is naturalistic, evolutionary, and deterministic. Values or ideals are hypotheses for the perfecting of human life. They arise out of experience, are formulated by the imagination, are tested in the evolutionary process, and are at once relative and a challenge to strenuous philosophy. They are created by exceptional men, or idealists, and are often thwarted, at best conserved, by the mass of mediocre men. For these reasons *El hombre mediocre* is critical of democracy, although it calls for equality under law while asserting an aristocracy of merit.

See also Ethical Naturalism; Evolutionary Ethics; Metaphysics; Positivism.

Bibliography

WORKS BY INGENIEROS

Psicología genética. Buenos Aires, 1911.

Hacia una moral sin dogmas. Buenos Aires: Talleres gráficos de L.J. Rosso, 1917.

La evolución de las ideas argentinas, 2 vols. Buenos Aires: Talleres gráficos de L.J. Rosso, 1918–1920.

Obras completas, 24 vols. Buenos Aires, 1930–1940.

WORKS ON INGENIEROS

Torchia Estrada, Juan Carlos. *La filosofía en la Argentina.* Washington, DC: Unión Panamericana, 1961. Ch. 9.

Arthur Berndtson (1967)

INNATE IDEAS

The theory of "innate ideas," in any of its philosophically significant forms, claims that all morally right judgment or all science, or both, rest upon or consist in a knowledge a priori either of (*a*) universal principles governing reality or (*b*) objects transcending sensory experience. Representative of such universal principles are "From nothing, nothing comes" (*Ex nihilo, nihil fit*); "Equals added to equals give equals"; "It is wrong to murder." Illustrations of transcendent objects are Platonic Forms and God. Concomitantly, the theory attempts to explain the genesis and epistemological status of the conception of such principles and objects. For this purpose it introduces the notion of innate ideas.

Proponents of the theory of innate ideas (henceforth "innatists") would typically agree with empiricists that sensory experience consists of particulars. They would claim, however, that scientific knowledge is knowledge that holds good everywhere and at every time, that such knowledge in fact exists, and that the abstracting and compounding of sensory particulars in empiricist inductions cannot possibly provide us with such knowledge, but at most only with opinion. Innatists would also maintain, in agreement with some empiricists, that the

abstractions and compoundings of sensory particulars described by empiricism as the basis of conception cannot possibly provide us with the conception of such universal principles or transcendent objects as are referred to above. At the same time, innatists would typically disagree with those realists who claim that such conceptions and knowledge are attained through direct perceptions or intuitions of nonsensory reality, or if they did join the theory of innate ideas with a theory of such intuitions, as Plato seems to have done, they would hold that scientific knowledge, though it may conclude in such intuitions, does not commence with them. (To maintain that scientific knowledge commences with such intuitions would be to make the notion of innate ideas methodologically and epistemologically superfluous.) The notion of innate ideas rests, for its philosophical significance, on the assumption that knowledge of reality is not given directly—at least, not in its chronologically first premises—but through representations. Where reality is viewed as something distinct from sensory particulars, innatists are thus representative realists.

Since proponents of the theory of innate ideas deny that such conceptions of universal principles or of transcendent objects as are described above are derived either from sensory experience or from intuitions of nonsensory reality, they are left with the problem of explaining their genesis. This they solve by holding these conceptions to be innate or inborn—to be, in short, innate ideas. But in speaking of innate ideas, proponents of the theory seem to mean two things. By "idea" they sometimes mean an object of awareness, like a mental image. When speaking in this way, innatists must maintain that conceptions of universal principles or of transcendent objects are present in the mind from birth or even prior to it. Innatists then typically explain why children and savages do not seem to be cognizant of the principles or objects in question by holding that these conceptions or representations, though present in the mind, are obscured by the presence of other conceptions or ideas—in particular, sensory ideas or percepts—much as the sound of a flute might be present in the air but be inaudible because of other sounds or noises. Again, innatists sometimes mean by "idea" not an object of awareness but, rather, a disposition of the mind or reason to form a determinate conception under certain conditions or stimuli. In René Descartes, for instance, whenever consciousness occurs, there also occurs the conception that something is conscious—namely, oneself—and this is innate in the dispositional sense.

An equally crucial problem for proponents of the theory of innate ideas is to explain the epistemological status of innate conceptions. Since these conceptions are held to constitute the foundation for all science and since science is conceived of as depicting reality, the question arises: How can we know that these conceptions apply to reality?

Again, two answers are traditionally given to this question. One answer, originating in Plato, holds that innate ideas are actually memories. These memories are the representations of direct intuitions of reality experienced before birth. Innate ideas express knowledge, then, in the way that memories do. A second answer, exemplified in Descartes, holds that the truth of innate ideas can be internally validated. Thus, in Descartes we find upon reflection that two innate ideas, the idea that I am and the idea that from nothing nothing comes, possess a special property—they not only involve the immediate assent of reason (their denial being a contradiction of sorts) but they cannot be subjected to doubt, since any possible argument of doubt, as, for instance, appeal to an evil demon as the source of these ideas, must implicitly affirm the ideas in question. Thus, in arguing that an evil demon might be deceiving me, I at the same time affirm that I am and employ the principle that from nothing nothing comes. Taking a stand on these two innate ideas, Descartes then purports to prove the existence of God and God's goodness; by so doing, he thinks to establish clarity and distinctness as both the necessary and the sufficient condition of an idea's being true and thus validate all other innate ideas.

In summary, then, the theory of innate ideas states that certain conceptions of universal principles and nonsensory objects are innate, in the sense of being either images present in the mind at or before birth or inborn dispositions of the mind to form conceptions under certain circumstances. Since these conceptions, taken as either images or dispositions, exist chronologically before sensory experience, they are a priori in the literal, temporal sense of the term. Since they are not composed from or testable in sensory experience but since they provide the basis for all scientific knowledge, they are also a priori in the logical and epistemological senses.

HISTORY OF THE THEORY

The notion of innate ideas patently lends itself to theological speculation and to systems of metaphysics that locate reality in realms transcending sensory experience. Plato employed the notion as the bridge to the realm of Forms, and similar metaphysical and theological uses of

the doctrine occur in the works of the Neoplatonists (Plotinus, for example), as well as in the works of later philosophers and theologians belonging to the Platonic and the Neoplatonic tradition, including St. Augustine in the early period of Christianity, and Marsilio Ficino in the Italian Renaissance. Outside the strictly Platonic and Neoplatonic line, Descartes, as already noted, employed the doctrine in his proof of God's existence, and it was used in a similar fashion by the ancient Stoics, Herbert of Cherbury, and many other philosophers.

The doctrine of innate ideas also has an intimate relationship with the philosophy of science. Historically, this relationship has minifested itself in the fact that philosophical controversy over the doctrine has been greatest just when philosophers have been most concerned to establish foundations and methods for science. Thus, the existence of innate ideas was especially debated in the fourth century BCE by the philosophers of the Academy and the Lyceum and in the seventeenth century by the Continental rationalists and the British empiricists.

The question of whether innate ideas exist is not without consequences in the establishment of science. The doctrine of innate ideas favors certain scientific procedures and discourages others. In particular, it favors meditation as opposed to laboratory experimentation and mathematical methods as opposed to inductive methods. It might seem, however, that philosophical theories concerning the origins and foundations of scientific knowledge could have had, and therefore have had, no actual influence on the establishment of science, just as it has seemed to many philosophers that philosophical theories concerning ethics could have had and have had no actual influence on men's moral behavior. But this view overlooks the failures of would-be science in the seventeenth and eighteenth centuries: On the one hand, the imposing but vacuous systems spun out of the doctrine of innate ideas, and, on the other, the aimless experimentation and observation that Jonathan Swift, for example, caricatures in parts of *Gulliver's Travels*. The truth would appear to be, not that the doctrine of innate ideas could have had no real influence upon the development of science, but that if it had been strictly and universally adhered to in the seventeenth century and afterward, science would not have been established; but, then, universal adherence to the stricter forms of empiricism would also have been a sterile cause, and so, too, it would seem, would have been an intellectual climate in which neither philosophical empiricism or philosophical rationalism played any part in men's thinking.

EVALUATION

The classic attack upon the doctrine of innate ideas is made by John Locke in the first book of *An Essay concerning Human Understanding*. Locke argues that if the doctrine of innate ideas were true, one would expect to find certain ideas, such as the idea of God or the idea that whatever is, is, possessed by everyone and consciously employed in all their reasonings. This is not the case, however. Small children and savages do not possess these ideas, nor do persons consciously employ them in all their reasonings.

Commentators on the theory of innate ideas have sometimes complained that Locke's criticism of the theory sets up a straw man that no responsible innatist has ever cared to defend. In particular it has been claimed that responsible innatists have not held, and have not pretended to hold, that universal recognition and acceptance are corollaries to the existence of innate ideas. But this complaint is beside Locke's point. It is clear, for instance, that if small children everywhere, at the commencement of their discourse with others, appealed explicitly to the idea that whatever is, is, or to the idea of God, there would be good empirical grounds for supposing that innate ideas existed. For these would be at least some of the crucial empirical consequences one would want to deduce from the theory. Since these crucial consequences are not observed but might theoretically be observed, the theory is an empirical theory and, as such, it stands refuted by experience.

It has been argued against the theory of innate ideas that whatever transcendent principles or conceptions the theory pretends to account for can be accounted for more plausibly by supposing them to be constructed from givens of experience or acquired through transcendent intuitions. This argument, however, is not very convincing. It is, for example, impossible to conceive how the concept of infinity could be constructed from givens of experience or acquired through the contemplation of some transcendent realm of entities. But it is not clear, either, how possession of the concept can be accounted for through the theory of innate ideas.

See also A Priori and A Posteriori; Augustine, St.; Descartes, René; Ficino, Marsilio; Greek Academy; Herbert of Cherbury; Hume, David; Knowledge, A Priori; Leibniz, Gottfried Wilhelm; Locke, John; Nativism, Innatism; Neoplatonism; Plato; Plotinus; Rationalism; Renaissance; Stoicism; Swift, Jonathan.

Bibliography

PRIMARY SOURCES

Plato connects the theory of innate ideas with the doctrines of reminiscence and metempsychosis in *Meno* and *Phaedo*. (See *Dialogues*, edited by B. Jowett, 2 vols., 4th rev. ed., Oxford, 1953). He supports this by the case of the slave boy who carries out geometrical demonstrations in a manner suggesting recollection.

Herbert of Cherbury bases his doctrine of natural religion upon a theory of innate ideas. In his *De Veritate*, translated as *On Truth* by M. H. Carré (Bristol, U.K., 1937), universal recognition and acceptance are treated as criteria distinguishing innate ideas from other ideas. Pragmatic overtones are introduced: Our common or innate notions are also those that conduce to our preservation, and conversely.

Descartes lays the foundation for most subsequent discussion, pro and con, of the theory of innate ideas in his account of the wax tablet and our judgments of other minds in the "Second Meditation" and in his threefold division of ideas in the "Third Meditation." See *Meditations*, in *The Philosophical Works of Descartes*, translated by E. S. Haldane and C. T. Ross. 2 vols. (Cambridge, U.K.: Cambridge University Press, 1911, 1931; New York, 1955).

Locke's classic attack upon the theory of innate ideas is given in Book I of the *Essay concerning Human Understanding* (edited by J. W. Yolton, 2 vols., London, 1961). Locke argues that the theory that all knowledge is acquired from experience can be substantiated in experience; the doctrine of innate ideas is disconfirmed by experience.

Gottfried Wilhelm Leibniz's *New Essays concerning Human Understanding* contains an exhaustive examination and critique of Locke's attack upon the theory of innate ideas. See "Specimens of Thoughts upon the First Book of the Essay on Human Understanding, 1698," the preface, and "Book I" in the *New Essays* as translated by A. G. Langley. 3rd ed. (La Salle, IL, 1949). In opposition to Locke, Leibniz argues that insensible perceptions exist and that thus Locke's arguments concerning children and savages being unaware of the concept of God or such principles as whatever is, is, do not refute the theory of innate ideas. Contains Leibniz's own version of the theory of innate ideas.

David Hume in his *Enquiry concerning Human Understanding* offers a very brief but penetrating and informative discussion of the theory of innate ideas and its relation to his principle that every simple idea is the copy of a precedent simple impression. See Sec. 2, footnote 1, in the *Enquiry* edited by C. W. Hendel (New York, 1955).

N. O. Losskii connects the historical appearance of the doctrine of innate ideas in the seventeenth century with the failure of empiricists of that era to account for transcendent knowledge, necessary truths, and knowledge of an external world. See Ch. 2 of his *The Intuitive Basis of Knowledge*, translated by N. A. Duddington (London: Macmillan, 1919), where he argues that although the pre-Kantian rationalists seemed able to resolve the difficulties in question by the doctrine of innate ideas, the cost of this resolution was prohibitive. They "had to assume that *the whole of knowledge is innate*."

R. I. Aaron in *The Nature of Knowing* (London: Williams and Norgate, 1930) discusses innate ideas in connection with his intuitionist theory of knowledge. Although he maintains that all discursive knowledge rests upon a priori knowledge of indubitable or self-evident principles, he denies that the latter are innate ideas.

H. H. Price maintains in *Perception* (London: Methuen, 1932) that there are a priori innate ideas, but that "we only come to clear consciousness of such concepts … when we have already applied them many times."

Lewis E. Hahn's *A Contextualistic Theory of Perception* (Berkeley and Los Angeles: University of California Press, 1942) contains a detailed discussion of the interrelationship of strict sense-data theories of perception, pragmatic theories of perception, and the doctrine of innate ideas. Hahn argues that a strict sense-data theory of perception forces one to accept the doctrine of innate ideas in order to account for one's knowledge and conception of material things, whereas a pragmatic theory of perception does not. He takes this consequence to count in favor of pragmatic theories of perception and against sense-data theories.

John Wild maintains in his *Introduction to Realistic Philosophy* (New York: Harper, 1948) that what evidence we possess strongly indicates that rational knowledge does not rest upon innate ideas, but is acquired. This evidence consists in the fact that "first we do not know. Then we know."

SECONDARY SOURCES

Bealer, George. "The Incoherence of Empiricism." *Proceedings of the Aristotelian Society*, supplement 66 (1992): 99–138. Reprinted in *Rationality and Naturalism*, edited by S. Wagner and R. Warner. Notre Dame, IN: University of Notre Dame Press, 1993.

Carruthers, Peter. *Human Knowledge and Human Nature: An Introduction to an Ancient Debate*. New York: Oxford University Press, 1992.

Carruthers, Peter. "What Is Empiricism?—I." *Proceedings of the Aristotelian Society*, supplement 64 (1990): 63–79.

Flage, Daniel, and Clarence Bonnen. "Descartes and the Epistemology of Innate Ideas." *History of Philosophy Quarterly* 9 (1992): 19–33.

Fodor, Jerry. "The Current Status of the Innateness Controversy." In *Representations*. Cambridge, MA: MIT Press, 1981.

Jolley, Nicholas. "Leibniz and Malebranche on Innate Ideas." *Philosophical Review* 97 (1988): 71–91.

Newman, Lex. "Descartes on Unknown Faculties and Our Knowledge of the External World." *Philosophical Review* 103 (1994): 489–531.

Rose, F. O. *Die Lehre von den angeborenen Ideen bei Descartes und Locke*. Berne, 1901.

Schmaltz, Tad. "Descartes on Innate Ideas, Sensation, and Scholasticism: The Response to Regius." In *Studies in Seventeenth-Century European Philosophy*, edited by M. A. Stewart. New York: Clarendon Press, 1997.

Sigall, E. *Platon und Leibniz über die angeborenen Ideen.* 2 vols. Chernovtsy, Russia, 1897–1898.

Stich, S. *Innate Ideas*. Berkeley: University of California Press, 1975.

John O. Nelson (1967)
Bibliography updated by Benjamin Fiedor (2005)

INNATE IDEAS, NATIVISM

Nativism, or the doctrine of innate ideas, is the thesis that human beings possess at least some substantive knowledge innately. The doctrine has long been the subject of intense controversy among philosophers, and since the late 1950s, among cognitive scientists as well. It is generally understood in opposition to the doctrine of empiricism, according to which all substantive human knowledge derives from sense experience.

Proponents of nativism argue that experience alone cannot account, either *de facto*, or in principle, for the extent and specific content of human knowledge. Arguments of the first type focus on the type and amount of information actually available to a human being during a given period of empirical experience, and purport to show that the information contained in that experience is insufficient to account for a person's manifest cognitive achievements at that point. Such arguments have become known by the term introduced by linguist Noam Chomsky as "poverty-of-the-stimulus" (POS) arguments, and constitute the most common way of defending nativist hypotheses (Chomsky 1957, 1965, 1975). Arguments of the second type are transcendental in character. They purport to show that certain kinds of knowledge are preconditions of empirical learning, and could not, therefore, be the products of such learning. Philosopher Jerry Fodor, who posits an innate system of representation, is probably the most well-known contemporary author of an argument of this type (Fodor 1975). In both kinds of argument, nativists tend to focus on the particular mechanisms of knowledge acquisition presumed or posited by their opponents.

Opponents of nativism also employ a variety of strategies. Some attempt to rebut POS arguments, by showing that the relevant body of empirical experience does, in fact, contain the right types and amount of information to account for human knowledge, and that actual human learning mechanisms can, in fact, extract it. Others critics attack the doctrine directly, arguing either that the notion of innate knowledge is hopelessly obscure, or that it cannot be made scientifically respectable. Still other opponents contend that the doctrine properly understood is trivial—that it posits nothing an empiricist need reject.

Indeed, philosophical discussion of nativism has tended to focus on the meaning of the doctrine, and in particular, the operative terms, "knowledge" and "innate." Philosophers have traditionally made a provisional distinction between declarative, or theoretical knowledge "knowing-*that*"—and practical knowledge—"knowing-*how*." It is one thing to attribute to someone practical knowledge of, for example, logic, and quite a different thing to attribute to her declarative knowledge of logic. Normal human beings all have the ability to reason in accordance with the laws of logic, whereas few of us can articulate those laws, or even recognize them if they were laid out explicitly before us. Accordingly, the claim that some practical ability is innate is, on its face, different from the claim that some bit of declarative knowledge is.

The term "innate" has both a categorical and a dispositional sense. In the categorical sense, a trait is innate only if it is manifestly present at the creature's beginning; in the dispositional sense, a trait can be innate if the creature is disposed to manifest it under normal conditions. Now it may well be that a newborn human infant lacks the ability to reason—the relevant neuronal connections may not yet have formed—so that practical knowledge of logic is not innate in the categorical sense. Still, a normal infant is disposed to acquire the ability to reason, in normal circumstances. That is, the infant is so constituted that the necessary neurological connections will be formed, provided the infant gets enough to eat, suffers no head injuries. Practical knowledge of logic should therefore be counted as innate in the dispositional sense. The same point can be made with respect to declarative knowledge: It could be innate in the dispositional sense even if it is not innate in the categorical sense.

There are further complications. Knowledge can be either explicit or implicit. While it could be argued that few individuals have explicit declarative knowledge of the laws of logic, it is clear that everyone has—in virtue of being a competent reasoner—implicit knowledge of those laws. Contemporary cognitive science introduces another wrinkle: Knowledge can reside at the "personal" or at the "sub-personal" level. Personal-level knowledge is knowledge that can be consciously accessed, and that can play a role in conscious deliberation. Sub-personal knowledge is posited to reside in cognitive "modules"—subsystems that perform cognitive operations that are inaccessible to consciousness, and that are isolated from personal-level beliefs and desires. Many philosophers find the notion of sub-personal "knowledge" oxymoronic, and yet the notion has come to play an increasingly important role in contemporary cognitive science.

Contemporary cognitive science and cognitive ethology (the study of nonhuman animal cognition and behavior) have generated a great deal of support for the view that much apparently intelligent animal behavior is to be explained by natively structured cognitive "mod-

ules" containing domain-specific information and algorithms. Explanations of this sort have been posited for phenomena as diverse as bird navigation systems and bee foraging strategies (Gallistel et al. 1991). Psychologist Steven Pinker suggests that this conception of "instinct" provides the best way to understand the species-specific human capacity for language (Pinker 1994).

Nativism as it is understood in philosophy and contemporary cognitive science is a theory of human universals, of species-wide characteristics. It should therefore be sharply distinguished from "biological determinist" theories—theories that purport to explain individual human differences with respect to such traits as intelligence and criminality in terms of innate and supposedly immutable genotypic differences.

HISTORY OF THE DEBATE

The basic outlines of the debate were set in ancient Greece, and were further elaborated in the early modern period. Plato employed a POS argument for a version of nativism, his doctrine of "recollection" (*anamnesis*). In *Meno*, Plato's character Socrates elicits from an uneducated slave boy the solutions to a series of problems in geometry, culminating in the identification of an irrational number, the square root of two. Since the boy could not have learned it, Socrates argues, the relevant information must have already been present within his soul, needing only the stimulus of Socrates's questioning for it to be "recollected." Socrates goes on to argue that the origins of this knowledge, which every human being could "find … within himself" (Plato 1981, p. 75), are prenatal, the result of the soul's early encounter in a previous life with the objects of eternal truths.

Aristotle rejected Plato's doctrine of innate ideas, averring that the mind is initially blank. Universals, or "intelligible forms" can only be grasped through experience, by abstraction from the "sensible forms" delivered to the mind through sense perception (*Posterior Analytics*, Bk. II). This Aristotelian empiricism was embraced and elaborated by the Scholastics in the Middle Ages. According to Aquinas, apparently *a priori* knowledge such as that of Plato's slave boy could be accounted for in terms of the mind's detection of its own operations. Although the mind natively has the propensity to perform these operations, it is not until the mind is stimulated by the provision of the sensible forms imparted through perception that such operations actually occur. In this sense, then, even the apprehension of one's own inner mental life is indirectly dependent upon sensory experience (Adams 1975).

René Descartes revived nativism. He rejected the Aristotelian/Scholastic doctrine that our ideas of perceptual qualities resemble or share a form with the qualities of the bodies that occasioned them. On Descartes's view, perception, like all causal processes, involved the movement of tiny bodies or corpuscles, the properties of which bore no resemblance to the ideas of sensory qualities in which the process culminated. In that case, Descartes argued, our ideas of color, shape, movement could not have been imparted by sensation, but must rather be innate, and merely occasioned by sensation. In calling such ideas innate, Descartes was not claiming that they were manifestly present in the mind since birth. Rather, he meant that our minds are innately constituted in such a way that sensory experiences of certain sorts reliably give rise to ideas of certain sorts (Descartes 1648/1969).

Although Descartes did not believe that ideas must share forms with the objects they represented, he did accept a principle that constrained the relation between the two. This principle, sometimes called the "principle of proportionate reality," formed the basis of a type of POS argument, albeit an a priori one. The principle states that a cause must contain at least as much reality as its effect. Applied to the special case of the relation between ideas and their causes, Descartes claimed that the principle entails that the cause of an idea must possess at least as much "formal" reality as the "objective" (or representative) reality contained in the idea.

There are three possibilities, Descartes says, as to the causes of those of our ideas that are involved in judgments, that is, that admit of truth or falsity: they can be (a) caused by something outside of ourselves ("adventitious" ideas), (b) fabricated by ourselves ("factitious"), or (c) innate. Now our idea of God is an idea of an infinite, eternal, omniscient, omnipotent and perfectly benevolent being. Since the amount of objective reality in such an idea is infinite, the cause of this idea itself must be something infinite and perfect. No finite substance could be the cause, and hence it could not have been caused by a substance exeeeee e external to myself, nor could I have caused it myself. Only an infinite, perfect being could have caused it, i.e., only God Himself. Furthermore, since this idea does not come upon me "unexpectedly" as do ideas that come through the senses, and since it is not within my power to change it as is the case with ideas that I originate, my idea of God, Descartes concludes, must be innate, implanted by God at the beginning of my existence (Descartes 1641/1969).

Descartes also made two empirical arguments for the domain-specificity of the human language capacity, argu-

ments explicitly cited by Noam Chomsky in his twentieth-century defense of the same conclusion. In *Discourse on Method* (1637/1969) Descartes sought to establish the existence of a "special faculty" in human beings that accounted for our capacity, distinctive within the animal kingdom, for creative language use. He considers the objection, offered by a hypothetical interlocuter, that human linguistic competence might be due to a merely quantitative difference in some general ability between human beings and other animals, rather than, as Descartes's view had it, a difference in kind.

Descartes's first response is to point out that even the "stupidest" human children acquire language without difficulty, whereas not even the most intelligent infrahuman animals are able to acquire it at all. (It is noteworthy that Descartes's judgment about the linguistic incapacity of infrahuman animals was confirmed during the 1970s, when a variety of researchers attempted, without success, to teach American Sign Language to apes and chimpanzees. Despite assiduous training, not one of the otherwise highly intelligent and creative animals even approached the linguistic achievements of the average deaf human three-year-old (Pinker 1995). This response, however, is vulnerable to the following rejoinder: It may be that there simply is no overlap between humans and other animals in the ranges of variation in the relevant ability, that the "stupidest" human has a far greater amount of this hypothetical general ability than does the "most perfect parrot or monkey." Descartes's second argument speaks to that possibility: if human linguistic ability is a manifestation of a more general ability, then we'd expect that human beings would excel above animals in all activities to the same extent we exceed them in the practice of communication. But this is not what is observed (Descartes 1637/1969).

John Locke, like Aristotle, held that the mind is initially only a "white paper void of all characters, without any ideas" (1689/1979, p. 2). He explicitly rejected Descartes's doctrine of innate ideas. It was unnecessary, he argued, to posit "innate principles" in order to account for any feature of human knowledge, when the humanly "natural faculties" of sensation and reflection could be shown to be quite sufficient. He particularly objected to the view that ideas of sense were innate, arguing that "it would be impertinent to suppose the ideas of colors innate in a creature to whom God hath given sight, and a power to receive them by the eyes from external objects" (1689/1979, p. 1). But he was equally adamant that truths of reason became known only through experience. The *faculty* of reflection was sufficient to ensure that we recognize such propositions as true as soon as we apprehend them.

Locke's *Essay* was meant to provide a systematic explanation of the origins of all human knowledge from these raw materials of human faculties and simple ideas of sense, and thus was addressed to Descartes's POS arguments. But Locke also had principled reasons for rejecting the doctrine of innate ideas, reasons that echo today in contemporary empiricists' objections to empirical nativist theories. To begin with, he argued, these allegedly innate principles are not in fact universally known—"children and idiots have not the least apprehension … of them" (p. 4)—and so could not have been imprinted upon the human soul at birth. If it is replied that the principles could have been imprinted without the subjects' being aware of them, Locke argues, then the doctrine of innate ideas is trivialized. If we can make sense of the notion of an unperceived idea, Locke argues, it can only be as an idea that we have the *capability* of acquiring. If the doctrine is understood in this way, however, it becomes trivial, because *any* truth a person can come to know in the course of a lifetime would count as "innate" in this sense. And in any case, even if "universal consent" were established, nativism would not provide the only explanation.

Locke's positive account of the acquisition of ideas and the formation of knowledge relied on posited mechanisms of association and abstraction for the accumulation and manipulation of sensory impressions into general ideas, abstract ideas, and judgments. Subsequent empiricist models of concept acquisition and learning have followed Locke's model in essentials.

Gottfried Leibniz, in his *New Essays*, attempted a systematic rebuttal of Locke's critique of the doctrine of innate ideas. Leibniz first addressed Locke's in-principle objection to the notion of an innate, but unperceived idea. Leibniz argues that Locke is operating with too restrictive a notion of "knowledge" if he does not acknowledge the existence of implicit or unconscious knowledge. And to Locke's objection that allowing implicit knowledge would trivialize the doctrine of innate ideas, Leibniz responds that a distinction can and must be made between sciences like arithmetic and geometry, that we can "construct for ourselves … in our private room … without learning through sight or even touch the truths which we need," and those which require sensory experience (Leibniz 1704/1975). Furthermore, Leibniz argues, there is a difference between the mind's actually possessing structure, of a sort that permits the generation, *a priori*, of knowledge, and the mind's simply having the

potential for acquiring truths. This disagreement between Locke and Leibniz about the nature and significance of innate mental structure has echoes in contemporary philosophical debates.

Leibniz then turned his attention to Locke's positive account of the development of knowledge, and here makes an argument of the type this entry has termed "transcendental." Leibniz focuses on two concepts repeatedly cited by Locke as examples of ideas that could not possibly have been innate, but must be acquired through experience, namely impossibility and identity. According to Locke, these ideas can only be the product of the comparison of various specific sensory experiences, such as the taste of a normal nipple, versus one rubbed with wormwood. But this, according to Leibniz makes no sense: abstraction cannot explain the acquisition of these concepts, because abstraction presupposes the ability in the subject to classify experiences as similar or different, and thus as possessing the concepts same and different. Possession of such concepts is a precondition of, and thus cannot be the result of empirical learning (Leibniz 1704/1975). Ironically, a similar argument was made by the empiricist David Hume for the innateness of the principle of induction. The notion that the future tends to resemble the past is not one, Hume argued, that could be acquired through experience, because we need to presume that principle in order to take experience as evidence for anything at all (Hume 1748/1977).

Immanuel Kant, from whom my use of the term "transcendental" is borrowed, argued similarly that empirical experience as we know it—"intuitions"—are only possible because of innate forms and structures that characterize our perceptual and intellectual capacities. Space and time, he argued, are not features of reality considered "in itself" but are rather the "a priori" forms of perception. Perceptual experiences are only cognitively available to us because of the "pure categories of the understanding," highly general concepts like object and cause, that allow us to utilize experience to form judgments (Kant 1781/1787/1929).

CURRENT CONTROVERSIES

At the beginning of the twentieth century, under the influence of logical positivism, many philosophers rejected not only the doctrine of innate ideas, but the very notion of the mind. Hypotheses about mental structures and processes were held to be at best unverifiable, and at worst, unintelligible; the only possible "science of the mind" would be, ironically, a science of behavior. J. B. Watson and B. F. Skinner developed what they claimed was a fully general account of learning, applicable to all behavior, whether "intelligent" or reflexive, human or infrahuman. The basic mechanisms posited by behaviorists—classical and operant conditioning—involved the evocation, shaping and reinforcement of patterns of behavior. Although behaviorists eschewed any reference to the mental, they did tacitly accept Leibniz's point against Locke. Accordingly, they posited what philosopher W. v. O. Quine called an "innate similarity space"—a disposition to treat stimuli as falling into similarity classes, manifested by "stimulus generalization," the transfer of a reinforced response to similar but novel circumstances, a precondition of the learning of complex behavior (Skinner 1953).

Skinner thought that behaviorist principles could and would account for even the most complex behaviors acquired by human beings in their lifetimes, including the mastery of language. His book *Verbal Behavior* was the first, and to date, the last, effort by an anti-nativist, anti-mentalist to provide a systematic and relatively detailed explanation of the acquisition of human language (Skinner 1957). But in 1959, Noam Chomsky published a devastating review of the book in which he showed that the theory faced a fatal flaw: It was either grossly empirically inadequate or else devoid of empirical content (Chomsky 1959).

Chomsky had outlined his own positive account of language acquisition in 1957, an account that was iconoclastic at the time in reviving not only mentalism, but nativism. Unlike Skinner, who derived his model of language learning top-down, Chomsky urged a naturalistic approach to the study of language, one that focused on the actual conditions in which children acquired their linguistic competence. This focus quickly revealed several important points: (1) children attain language without any explicit instruction; (2) language is almost universally acquired, even for blind children (and deaf children if they are given access to signed language) within the first five years of life; and (3) the body of evidence available to children during acquisition—the body of "primary linguistic data"—is badly "impoverished" relative to the body of information eventually mastered. In particular, the data do not contain "negative evidence"—information that certain constructions are *not* licensed. This is highly significant, as there seem to be certain kinds of ungrammatical constructions that never appear among children's early mistakes. Given all this, Chomsky concluded that language acquisition was not the result of some general learning mechanism operating on sensory data, but that it rather involved, a domain-specific cogni-

tive mechanism that embodied constraints on the forms human linguistic systems can take—the "language acquisition device" (LAD). This natively specified set of constraints, "universal grammar" (UG), greatly simplifies the acquisition task by sharply constraining the set of candidate grammars a child must consider in response to data and by filling in gaps left by experience.

The POS paradigm has since been applied in many other ways within cognitive science. Psychologists Barbara Landau and Lila Gleitman have shown that blind children acquire verbs of sight (for example, "look" "see") in much the same way as sighted children, though, obviously, in the absence of any visual experience (Landau and Gleitman 1985). Gleitman has argued that children's mastery of semantics (as opposed to syntax, which is Chomsky's focus) cannot be explained on the basis of experience, showing that the most common verbs in human language have no distinctive profiles of contingency of usage from which the child could infer their meanings (Gleitman 1990). These findings directly contradict the empiricist accounts of meaning acquisition found in Skinner and Quine (Skinner 1957, Quine 1973).

Elizabeth Spelke, pioneering an experimental paradigm for studying infant cognition, has provided empirical support for the views of Leibniz and Kant that our basic conceptual organization of the external world is natively specified, as a precondition of all other empirical learning (Spelke 1995). Evidence from the study of autistic persons has led many researchers to posit an innate "theory of mind" that enables non-autistic persons to interpret the facial expressions, gestures, intonation patterns of their fellow human beings in a reflexive way, and to effortlessly generate appropriate hypotheses about their likely intentions, desires, and reactions (Frith 1992, Baron-Cohen 1995). Empiricist critics of this so-called "theory-theory" approach to our capacity for psychological understanding argue that we do not need innate psychological knowledge; that rather, we utilize our capacity to mimic our conspecifics to run "simulations" of other individuals' psychologies, and thus to learn empirically what is going on in their heads (Gordon 1986, Heal 1994, Goldman 2006).

Philosophical critics of Chomsky have challenged the intelligibility of his model of mind and his conception of innate knowledge, raising many of the points raised by Locke and other empiricists. Gilbert Harman and Hilary Putnam each have argued that Chomsky's arguments demonstrate, at best, the existence of native principles of induction governing the learning of language, a thesis any empiricist could accept. Putnam also argued that nativism is not the only or the best explanation for the existence of linguistic universals, or the ubiquity of acquisition (Putnam1967, Harman 1967). Chomsky replies that the real issue is whether there are specialized mechanisms that in some sense embody domain-specific information, not how that information is embodied (See also Katz 1966). He insists that the notion of "innate knowledge" should be understood as an abductive posit in a scientific theory, with the details to be worked out as part of the relevant scientific investigation, as is usual in other branches of science (Chomsky 1969). Pinker and others have stressed, in response to Putnam's second point, that it is the poverty of the stimulus, rather than universality, that provides the strongest support for nativism. Alvin Goldman considers and responds to the objection, first found in Locke, that mere innateness cannot secure the justification required for knowledge, arguing that an externalist epistemological framework that takes account of selectional processes that account for native beliefs or biases can provide the systematic warrant needed for knowledge (Goldman 2006).

On the empirical side, critics have focused, appropriately, on the POS argument. Some have attacked the logic of the POS argument, contending that many cognitive achievements are made in the absence of apparently needed evidence, with there being no reason to think the relevant knowledge is innate. Fiona Cowie, for example, argues that she acquired the concept "curry" without explicit instruction, without negative evidence, but that that is no reason to think the concept is innate (Cowie 1998). Others contend that the primary linguistic data are richer than Chomsky has supposed. For example, it has been widely documented that parents and other speakers interacting with infants tend to produce a simplified and pedagogically friendly version of human language, dubbed "Motherese," that may provide the child with a salutarily biased sample of the language (Snow 1972). Other critics argue that our cognitive resources for extracting information from the environment are more powerful than Chomsky supposed. These latter critics include advocates of a new empiricist model of mind called "connectionism." Connectionists reject the view that cognition involves the manipulation of structured representations, arguing instead that the mind is a vast network of neural nodes, capable, given suitable "training," of detecting and responding, at levels not accessible to consciousness, to very subtle regularities in the data stream (Elman et al. 1996).

Chomsky's theory is computationalist—that is, it presupposes that the mind is, *inter alia*, an information

processing device. In 1975, philosopher Jerry Fodor argued that not only Chomsky's, but all successful and fruitful psychological theories were tacitly committed to this model, and that the model itself carried a heavy ontological commitment. Computations, Fodor argues, require a *medium of computation*. Since the acquisition of natural language is one of the processes that is, *ex hypothesi*, computational, natural language cannot itself be the medium in which the relevant computations take place. There must be an antecedently existing medium, an innate "language of thought" (LOT), with at least as much syntactic complexity and expressive power as needed to fully represent natural languages (Fodor 1975).

Fodor's view on the latter point has developed. Initially, he claimed that the argument showed that all concepts were innate, but he has recently modified this to the claim that no concept is *learned*, leaving open the possibility that there are non-computational, and hence non-psychological means for acquiring concepts (Fodor 1997). The argument that no concepts are learned is transcendental in character, and simple to state: (1) Concept learning (if there is such a thing) would involve the formulation, projection and confirmation of hypotheses as to the extension of the concept to be learned; (2) Such processes presuppose the means to represent the extension of the concept to be learned; (3) But any system that has the means to represent the extension of a concept *ipso facto* possesses the concept. Therefore, (4), one cannot learn a concept without already possessing that concept; hence (5) concept *learning* is impossible.

Fodor's critics include many of the philosophers who have challenged Chomsky. Critics have charged that the LOT leads to an infinite regress of languages and interpreters (Dennett 1975, Harman 1975). Connectionists have argued that Fodor's fundamental assumption—that mental operations involve the manipulation of structured symbols—is mistaken, and that there is therefore no need to posit an innate medium of computation (Clark 1993). Other critics have argued that Fodor has constructed a false dilemma about concept learning: that there are plausible models that do involve a rational extraction of information, and that should thus count as learning, but that do not employ the hypothesis-confirmation model that Fodor presumes to be the only alternative to non-psychological triggering (Margolis 1998, Cowie 1998).

Fodor is also responsible for revitalizing the theory of mental modularity (Fodor 1983). According to Fodor, processing in each sensory modality, as well as linguistic processing, takes place in specialized functional regions of the brain, equipped with proprietary algorithms, memory, and computational vocabulary. Such processing is characteristically fast, automatic, and "informationally encapsulated"—insensitive to information from outside the module, typically information stored in central systems. The character of these modules is specified in the genome, although their development may require experiential inputs at crucial stages. Contemporary anthropologists and social psychologists, along with some cognitive psychologists, have become increasingly interested in developing modular explanations for a wide variety of human cognitive and psychological traits, from reasoning about social contracts to men's alleged aesthetic preference for firm breasts in women (Barkow et al. 1992; see Pinker 1997, 2002 for an overview). Ironically, Fodor has emerged as the most vocal critic of this line of thought, objecting both to the strict adaptationist presumptions of the modularists' methodology, and to what he regards as the emptiness of the conception of modularity they employ (Fodor 2000). The debate beginning with Plato will no doubt continue as philosophers, psychologists, and cognitive scientists try to bring fresh insight to the issue of innate ideas.

See also Aristotle; Artificial and Natural Languages; Chomsky, Noam; Connectionism; Dennett, Daniel C.; Descartes, René; Fodor, Jerry A.; Harman, Gilbert; Hume, David; Innate Ideas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Plato; Putnam, Hilary; Quine, Willard Van Orman; Skinner, B. F.; Thomas Aquinas, St.

Bibliography

Adams, Robert Merrihew. "Where Do Our Ideas Come From?" In *Innate Ideas*, edited by Stephen P. Stitch. Los Angeles: University of California Press, 1975.

Aristotle. *Posterior Analytics* (350 BCE). In *The Basic Works of Aristotle*, edited by Richard McKeon. Oxford: Oxford University Press, 1941.

Barkow, J. H., L. Cosmides, and J. Tooby. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press, 1992.

Baron-Cohen, Simon. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press, 1995.

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.

Chomsky, Noam. *Reflections on Language*. New York: Pantheon Books, 1975.

Chomsky, Noam. "A Review of B. F. Skinner's *Verbal Behavior*." *Language* 35 (1) (1959): 26–58.

Chomsky, Noam. *Syntactic Structures*. The Hague: Mouton, 1957.

Clark, Andy. *Connectionism, Concepts, and Representational Change*. Cambridge, MA: MIT Press, 1993.

Cowie, Fiona. *What's Within? Nativism Reconsidered*. Oxford: Oxford University Press, 1998.

Descartes, René. *Discourse on the Method of Rightly Conducting the Reason* (1637). Translated by E. S. Haldane and G. T. R. Ross. In *The Essential Descartes*, edited by Margaret Wilson. New York: New American Library, 1969.

Descartes, René. *Meditations on First Philosophy* (1641). In *The Essential Descartes*, edited by Margaret Wilson. New York: New American Library, 1969.

Descartes, René. *Notes Directed Against a Certain Programme* (1648). In *The Essential Descartes*, edited by Margaret Wilson. New York: New American Library, 1969.

Dennett, Daniel. "Brain Writing and Mind Reading." *Minnesota Studies in the Philosophy of Science* 7 (1975): 403–415.

Elman, J. L., E. A. Bates, M. H. Johnson, and A. Karmiloff-Smith et al. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: Bradford Books/MIT Press, 1996.

Fodor, Jerry. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press, 1997.

Fodor, Jerry. *The Language of Thought*. New York: Crowell, 1975.

Fodor, Jerry. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press, 2000.

Fodor, Jerry. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

Frith, Uta. *Autism: Explaining the Enigma*. Oxford: Blackwell, 1992.

Gallistel, C. R., A. L. Brown, S. Carey, R. Gelman, and F. C. Keil. "Lessons from Animal Learning for the Study of Cognitive Development." In *The Epigenesis of Mind: Essays on Biology and Cognition*, 3–36. Hillsdale, NJ: Lawrence Erlbaum Associates, 1991.

Gleitman, Lila R. "The Structural Sources of Verb Meanings." *Language Acquisition* 1 (1) (1990): 3–55.

Goldman, A. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press, 2006.

Gordon, Robert. "Folk Psychology as Simulation." *Mind and Language* 7 (1986): 11–34.

Harman, Gilbert. "Language, Thought, and Communication." *Minnesota Studies in the Philosophy of Science* 7 (1975): 270–298.

Harman, Gilbert. "Psychological Aspects of the Theory of Syntax." *Journal of Philosophy* LXIV (1967): 75–87. In *Innate Ideas*, edited by Stephen P. Stitch. Los Angeles: University of California Press, 1975.

Heal, Jane. "Simulation vs. Theory-theory: What Is at Issue?" In *Objectivity, Simulation, and the Unity of Consciousness*, edited by C. Peacocke. Oxford: Oxford University Press, 1994.

Hume, David. *An Enquiry concerning Human Understanding* (1748). Indianapolis, IN: Hacket Publishing Co., 1977.

Katz, Jerrold J. *The Philosophy of Language*. New York and London: Harper and Row, 1966. Excerpted in *Innate Ideas*, edited by Stephen P. Stitch. Los Angeles: University of California Press, 1975.

Kant, Immanuel. *The Critique of Pure Reason* (1781a/1787b). Translated by Norman Kemp Smith. New York: Macmillan, 1929.

Landau, Barbara, and Lila R. Gleitman. *Language and Experience: Evidence from the Blind Child*. Cambridge, MA: Harvard University Press, 1985.

Leibniz, G. W. F. *New Essays on Human Understanding* (1704). Excerpted in *Innate Ideas*, edited by Stephen P. Stitch. Los Angeles: University of California Press, 1975.

Locke, John. *An Essay concerning Human Understanding* (1689). New York: Oxford University Press, 1979.

Margolis, Eric. "How to Acquire a Concept." *Mind and Language* 13 (1998): 347–369.

Pinker, Steven. *The Blank Slate*. New York: Viking/The Penguin Group, 2002.

Pinker, Steven. *How the Mind Works*. New York: Norton, 1997.

Pinker, Steven. *The Language Instinct*. New York: Harper Collins, 1994.

Plato. *Meno*. Translated by G. M. A. Grube. Indianapolis: Hackett Publishing Co., 1981.

Putnam, Hilary. "The 'Innateness Hypothesis' and Explanatory Models in Linguistics." *Synthese* 17 (1967): 12–22. In *Innate Ideas*, edited by Stephen P. Stitch. Los Angeles: University of California Press, 1975.

Quine, W. v. O. *The Roots of Reference*. La Salle, IL: Open Court, 1973.

Skinner, B. F. *Science and Human Behavior*. New York: The Free Press, 1953.

Skinner, B. F. *Verbal Behavior*. Acton, MA: Copley, 1957.

Snow, C. E. "Mothers' Speech to Children Learning Language." *Child Development* 43 (1972): 459–465.

Spelke, Elizabeth. "Initial Knowledge: Six Suggestions." *Cognition* 50 (1995): 433–447.

Stich, Stephen P., ed. *Innate Ideas*. Los Angeles: University of California Press, 1975.

Wilson, Margaret, ed. *The Essential Descartes*. New York: New American Library, 1969.

Louise M. Antony (2005)

INNER SENSES

The scholastic theory of the inner senses can be viewed as an attempt to explain and classify cognitive abilities shared by human beings and nonrational animals, abilities that go beyond pure sensation and require a certain level of abstraction. Given that capacities such as reason or belief were generally denied to animals beginning in the classical period of Greek philosophy, these powers or faculties of the sensible soul were thought to account for goal-directed or intentional animal behavior as well as memory and dreaming in humans and animals.

Historically, the concept of the inner senses is rooted in Aristotle's (384–322 BCE) remarks on postsensory faculties of the soul in the second and third books of *De Anima* and in *De memoria et reminiscentia*. A model list-

ing three "inner" psychic faculties, assigned to three cerebral ventricles (imagination/front ventricle, intellective faculty/middle ventricle, memory/rear ventricle), stems from the writings of Galen (129–c. 199) and was handed down to medieval thinkers via Nemesius's (fourth century AD) *De natura hominis* (chapter 5) and John Damascene's (c. 675–749) *De fide orthodoxa* (chapters 32–34). St. Augustine (354–430) was the first to use the Latin term *sensus interior*, meaning Aristotle's common sense (*Confessions*, book 1, chapter 17; *vis interior* in book 7, chapter 27).

However, the notion of the inner sense only appears there in its singular form. The tendency to posit a plurality of inner senses was probably most influenced by Avicenna's (980–1037) *Liber de anima sextus de naturalibus* (part 1, chapter 5). The Islamic philosopher lists five inner senses as powers of the apprehensive part of the sensible soul: (1) the common sense (*sensus communis*) combines the forms it receives from the five external senses; (2) the imagination (*imaginatio*) keeps these forms stored; (3) the imaginative power (*vis imaginativa*) combines and separates forms kept in the imagination; (4) the estimative power (*vis aestimativa*) judges perceived salient or of interest (e.g., the sheep that apprehends the perceived wolf as something it should flee from); and (5) the memory (*vis memorialis et reminiscibilis*) keeps these prerational estimations. Although Avicenna gives three- and fourfold classifications as well, this fivefold classification came to be frequently cited in medieval texts.

Albert the Great (Albertus Magnus, c. 1200–1280) uses Avicenna's classification and combines it with a description of the brain and the functions of animal spirits taken from Costa ben Luca's (c. 864–923) *De differentia animae et spiritus* to localize the inner senses. According to Albert the classification reflects different levels of abstraction and corresponds to the grades of subtlety of the animal spirits (*Summa de homine*). The common sense belongs to the same level of abstraction as the five external senses because its function depends on the immediate presence of a perceived object. Nonetheless, it is not counted as an external sense because it does not receive its forms directly from the external object, but from the external senses.

St. Thomas Aquinas (1225–1274) calls the common sense the "root and principle" of the external senses because it joins the different impressions of the external senses and thus combines the raw sense-data to form a unified episode of perceiving an object. Besides the common sense Aquinas's fourfold list of inner senses (*Summa Theologiae*, pars Ia, quaestio 78, articulus 4) includes the imagination (*imaginatio sive phantasia*), the functioning as storage for sensible forms, the estimative power (*vis aestimativa*), and the memory (*vis memorativa*, *memoria sive reminiscentia*). In contrast to Albert and Avicenna, Aquinas—following Averroes (1126–1198)—stresses that in human beings the animal estimative power is replaced by the cogitative power (*vis cogitativa sive ratio particularis*) that accounts for quasi-propositional perception. In modern philosophy the term *inner sense* is used to signify the mind's ability to reflect on its own operations (Locke 1975, Kant 1998).

See also Aristotle; Augustine, St.; Avicenna; Thomas Aquinas, St.

Bibliography

Kant, Immanuel. *Critique of Pure Reason*. Translated and edited by P. Guyer and A. W. Wood. Cambridge, U.K.: Cambridge University Press, 1998.

Kemp, Simon. *Medieval Psychology*. New York: Greenwood, 1990.

Locke, John. *An Essay concerning Human Understanding*, edited by P. H. Nidditch. Oxford: Clarendon, 1975.

Steneck, Nicholas H. "Albert the Great on the Classification and Localization of the Internal Senses." *Isis* 65 (227) (1974): 193–211.

Tellkamp, Jörg Alejandro. *Sinne, Gegenstände und Sensibilia: Zur Wahrnehmungslehre des Thomas von Aquin*. Boston: Brill, 1999.

Wolfson, Harry Austryn. "The Internal Senses in Latin, Arabic, and Hebrew Philosophic Texts." *Harvard Theological Review* 28 (2) (April 1935): 69–133.

Alexander Brungs (2005)

INSTRUMENTALISM

See *Dewey, John; Pragmatism*

INTENSIONAL TRANSITIVE VERBS

A verb is transitive if it takes a direct object and intensional if it exhibits one or more intensionality effects in its direct object. The three main such effects are (i) resistance to interchange of coextensive expressions, such as coreferential names or common nouns that happen to apply to exactly the same objects; (ii) lack of existence entailments even when the direct object is existentially

quantified; and (iii) a relational-notional ambiguity if the direct object is quantified.

Verbs of search, desire, and expectation, exhibit all three intensionality effects. Thus, for (i), "Lois seeks Superman" and "Lois seeks Clark" might differ in truth-value even though Superman is Clark. For (ii), "Perseus seeks a gorgon" can be true even if there are no gorgons (contrast the extensional transitive *found*). For (iii), "Richard III seeks a horse" is normally understood to mean that his search may be concluded successfully by finding any one of a range of horses: There need be no particular horse he must find. This is the notional reading. The relational reading is that there is some horse such that he is looking for that horse: Finding other horses will not do. Extensional transitives only allow relational readings: If Richard III rode a horse, there is a particular horse he rode. The relational/notional distinction was named and explored in Quine (1956).

Other groups of verbs exhibit various effects in various ways, providing much for an account of intensional transitives to explain. For example, depiction verbs generate a relational-notional ambiguity only with certain quantifiers in the direct object. "Guercino drew a dog" has both the relational reading—some specific dog—and the notional reading—no specific dog. But "Guercino drew every dog" seems to advert to some antecedent domain on which "every dog" is interpreted, requiring him to have drawn particular dogs (similarly with "most dogs" and "the dog"). By contrast, "Aldrovandi seeks every dog on his property" has a notional reading, according to which he simply has a general intention to find all the dogs there may be in the area. So depiction verbs are a special case.

Verbs of evaluation, such as *despise, fear, respect* and *admire*, resist interchange of coreferential expressions (for example, "Lex Luthor fears Superman but not Clark") but it is not so clear that they give rise to relational/notional ambiguities, at least with existential direct objects. The sentence "Churchill scorned a pedant" (pedantry was something up with which he was not prepared to put) can be understood in two ways: There is a relational reading, according to which there was a particular pedant who was the object of his scorn, and there is a generic reading, which attributes a ceteris paribus response-disposition to him and allows for exceptions. Generic readings are not notional ones, since they are just as common with extensional verbs—for instance, "Corporations overcompensate their CEO's" (see Ariel 1999 for more on generics). So "scorned a pedant" lacks a notional reading.

The verb *need* (and transaction verbs such as *wager* and *owe*) displays the opposite behavior: notional readings are unproblematically available, but substitutions are permitted that fail with evaluation verbs and desire verbs. For example, if Richard III is dehydrated and therefore needs some water, it follows, since water is H_2O, that he needs some H_2O, whether or not he has such concepts as hydrogen and regardless of his other beliefs. But if he thinks H_2O is a kind of rat poison, he may want some water without wanting some H_2O and fear H_2O without fearing water.

An obvious question is whether the three types of intensionality effect have a common ground or whether two or even three distinct mechanisms are involved. The contrast between *need* and *scorn* suggests that one mechanism is involved in substitution resistance and another in generating notional readings: The former is not available to *needs*, the latter not to *scorn*. And since one may need more flu vaccine even if there is no more left, existence neutrality apparently goes with the availability of notional readings.

It might be objected that we get existence neutrality with evaluation verbs as well: The Ancient Greeks worshipped Zeus even though there is no such entity. However, serious use of names for fictional or mythical items requires an ontology of abstract fictional or mythical entities that exist contingently (they would not have existed if the corresponding fictions or myths had not been created; see Salmon [2002]). We should also note that *worship* is peculiar among evaluative verbs as regards notional readings. For instance, if a priest sacrifices with the words "To whichever god is out there," this is, arguably, a case of worshipping a god, but no particular one.

The next step is to describe the mechanisms accounting for the intensionality effects. One possibility is that in all three cases, the problems with intensional transitive verbs simply duplicate those encountered with propositional attitude verbs. This position seems quite plausible for substitution-resistance. If we have a good account of why substitution fails in "Lex Luthor fears that Superman is nearby" we would surely expect it to transfer straightforwardly to "Lex Luthor fears Superman"—unless, that is, the account for propositional attitude verbs depends on the presence of a clausal complement, as appears to be the case with Davidson's "paratactic" analysis of propositional attitude ascriptions (Davidson 1969). However, other accounts of substitution-failure in propositional attitude ascriptions transfer more smoothly—for instance, any account on which a name is associated with

a way of thinking of the referent and that way of thinking somehow enters into the truth conditions of the ascription. (See, for example, the "hidden indexical" mechanisms explored in Crimmins [1992] and Forbes [2000]).

There is also the view that substitution resistance is an illusion (locus classicus Salmon [1986]), which, if correct for propositional attitude verbs, should be equally correct for intensional transitives. On this view a name contributes only its referent to the meaning of a sentence, so no semantic distinction is to be drawn between "Lex fears Superman" and "Lex fears Clark."

The idea that no new problem is presented by intensional transitives runs into trouble, however, when we consider the relational-notional distinction. With propositional attitude verbs, the difference between (a) ascribing some cognitive relation between a subject and a specific item that the attitude is about and (b) not making such an ascription is captured by a scope distinction: "Lex fears that an extraterrestrial is nearby" gets the relational meaning when the quantifier "an extraterrestrial" is moved out of the attitude-content specification so that it has scope over "fears," as in "An extraterrestrial is such that Lex fears it is nearby"; the notional reading corresponds to unambiguous restriction of the quantifier to the attitude-content specification, as in "Lex fears-true the proposition that an extraterrestrial is nearby." But when we turn to intensional transitives, we find that, at least within a first-order framework, notional readings cannot be represented as ones in which the intensional verb has scope over the quantifier. For the verb to have wide scope, the quantifier must be one of its arguments: "seeks (Lois, an extraterrestrial)." But in first-order language a quantifier cannot be an argument to a relation: it must take scope over a sentence, open or closed, hence "without an inner sentential context, distinctions of scope disappear" (Kaplan 1986, p. 266).

According to propositionalism, the inner sentential context is there but partly hidden. Quine (1956) advances this view in the thesis that search-verb sentences can be paraphrased in terms of trying to find. So "Perseus seeks a gorgon" would be paraphrased as "Perseus is trying to find a gorgon." Partee (1974, p. 97) notes that search verbs cannot all be paraphrased using "trying to find," since they are not all synonymous (*cf.* "hunt" and "rummage about"), but in defense of propositionalism, both Parsons (1977) and Larson (2001) suggest using the search verb itself along with "to find." So we get "Perseus seeks to find a gorgon," or, in a more explicitly propositionalist formulation, "Perseus seeks (in order) to make it true that he himself finds a gorgon."

Evidence for an implicit inner sentential context varies with different kinds of verbs. For example, "Richard III needs a horse quickly" barely makes sense if *quickly* is understood to modify "needs" or any other explicit material. It seems instead to modify an implicit *get*. Along with other evidence (see Den Dikken et al. [1996] for more) this makes it quite plausible that desire verbs and *needs* are not really transitive but take infinitival *to get* clauses as their true complements. However, comparable evidence for search verbs is hard to find, and whether converting the direct object into a purpose clause is meaning-preserving can be doubted. Depiction verbs and evaluative verbs present even more of a challenge. For instance, to fear x is not to fear encountering x, since one may not fear x but may fear encountering x because x has a dangerous communicable disease. Nor is fearing x the same as fearing that x will hurt you, since you may fear that your accident-prone dentist will hurt you without fearing your dentist. It is therefore conceivable that intensional transitives are not a unified semantic group: for some, such as desire verbs, *need*, and maybe verbs of expectation, propositionalism is workable, but not for others.

The main alternative to propositionalism is developed in Montague (1973) as part of a higher-order, type-theoretic semantics for natural language. In this framework, quantifiers can be arguments to verbs, so "seeks (Lois, an extraterrestrial)" is allowed as the semantics of the notional reading of "Lois seeks an extraterrestrial." Montague's ideas are refined, revised, and developed in Zimmerman (1993), Moltmann (1997), and Richard (2001), although in all these accounts, notional readings of search-verb sentences still put the searcher into a search relation to an abstract entity, the meaning of the quantifier (in standard Montague grammar, this is something rather complicated, a function from possible worlds to sets of intensional entities; see Dowty et al. [1981] for an accessible account). It is unclear that such a semantics is compatible with the evident univocality of *seeks* in "seeks an extraterrestrial, but no particular one" and "seeks a particular extraterrestrial."

The approach of Forbes (2000) avoids this problem by employing a Davidsonian event-semantics (Davidson 1967) in which verbs are treated as predicates of events and the same predicate *search* appears in both relational and notional readings. In relational readings, the syntactic object signals a theme of the event, but in notional readings, it simply classifies the search as being one of a certain kind, for instance, as being a search of the at-least-one-extraterrestrial kind.

Intensional transitives raise interesting logical problems. It may be argued that propositional attitude ascriptions have no logic at all: even "x believes that p and q" does not logically entail that "x believes that p": at best we may endorse a psychological principle that persons aware that they accept a conjunction will also accept each conjunct individually. But for intensional transitives, there are substantial questions about the validity of certain inference-patterns. For example, if Richard III needs a warhorse, does it follow that he needs a horse? If notional readings are glossed in terms of indifference ("any would do," as in Lewis [1972, p. 199]) it does not follow: Even if Richard III needs a warhorse, and any one will do, it does not follow that he needs a horse, and any one will do—in the mayhem of the Battle of Bosworth, a cart horse would not do. On the other hand, the standard glossing of notional readings using "no particular one" seems to leave open the logical status of the inference rather than settling it one way or the other. These and other issues about the validity of specific inference patterns are pursued in Richard (2000) and Forbes (2003).

See also Davidson, Donald; Language; Language, Philosophy of; Propositions; Quine, Willard Van Orman; Semantics; Sense.

Bibliography

Cohen, Ariel. *Think Generic!* Stanford: CSLI Publications, 1999.

Crimmins, Mark. *Talk About Belief*. Cambridge, MA: The MIT Press, 1992.

Davidson, Donald. "The Logical Form of Action Sentences." In *The Logic of Decision and Action*, ed. Nicholas Rescher. Pittsburgh: University of Pittsburgh Press. 1967.

Davidson, Donald. "On Saying That." In *Words and Objections: Essays on the Work of W. V. Quine*, edited by Donald Davidson and Jaakko Hintikka. Dordrecht: Reidel, 1969.

Den-Dikken, Marcel, Richard Larson, and Peter Ludlow. "Intensional 'Transitive' Verbs and Concealed Complement Clauses." *Rivista di Linguistica* 8 (2) (1996): 331–348.

Dowty, David, Robert Wall, and Stanley Peters. *Introduction to Montague Semantics*. Dordrecht: Kluwer, 1981.

Forbes, Graeme. "Meaning-Postulates, Inference, and the Relational/Notional Ambiguity." *Facta Philosophica* 5 (2003): 49–74.

Forbes, Graeme. "Objectual Attitudes." *Linguistics and Philosophy* 23 (2000): 141–183.

Kaplan, David. "Opacity." In *The Philosophy of W. V. Quine*, edited by Lewis Hahn and Paul Schilpp. LaSalle: Open Court, 1986.

Larson, Richard. "The Grammar of Intensionality." In *Logical Form and Natural Language*, edited by Gerhard Preyer and Georg Peter. Oxford and New York: Oxford University Press, 2001.

Lewis, David. "General Semantics." In *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman. Dordrecht: Reidel, 1972.

Moltmann, Friederike. "Intensional Verbs and Quantifiers." *Natural Language Semantics* 5 (1997): 1–52.

Montague, Richard. "The Proper Treatment of Quantification in Ordinary English." In *Approaches to Natural Language*, edited by Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes. Dordrecht: Reidel Publishing Company, 1973.

Parsons, Terence. "Meaning Sensitivity and Grammatical Structure." In *Structures and Norms in Science*, edited by Maria Luisa Dalla-Chiara, Kees Doets, Danieli Mundici, and Johan van Benthem. Dordrecht: Kluwer, 1997.

Partee, Barbara. "Opacity and Scope." In *Semantics and Philosophy*, edited by Milton Munitz and Peter Unger. New York: NYU Press, 1974.

Quine, W. V. O. "Quantifiers and Propositional Attitudes." *The Philosophical Review* 53 (1956), pp. 177–187.

Richard, Mark. "Seeking a Centaur, Adoring Adonis: Intensional Transitives and Empty Terms." In *Midwest Studies in Philosophy*. Vol. 25, *Figurative Language*, edited by Peter French and Howard Wettstein. Oxford: Basil Blackwell, 2001.

Salmon, Nathan. *Frege's Puzzle*. Cambridge, MA: The MIT Press, 1986.

Salmon, Nathan. "Mythical Objects." In *Meaning and Truth*, edited by J. Campbell, M. O'Rourke, and D. Shier. New York: Seven Bridges Press, 2002.

Zimmerman, Thomas Ede. "On the Proper Treatment of Opacity in Certain Verbs." *Natural Language Semantics* 1 (1993): 149–179.

Graeme Forbes (2005)

INTENTION

Philosophical work on intention is motivated by three general concerns. First, philosophers of action want to understand what it is for an event to be an intentional action and how intentional actions are produced by their agents. They have good reason to think hard about what intentions are and how they may be involved in the production of intentional actions, because, even if it is unclear exactly how intentional actions and intentions are related to each other, it is clear that they are intimately related. Second, moral philosophers and others in the business of developing theories of the evaluation of actions and their agents need an account of intentional action, and such an account is likely to involve intention in an important way. Moral evaluations of actions have intentional actions as their primary subject matter, even if people sometimes are proper targets of moral blame for some unintentional actions (e.g., when a drunk driver accidentally injures or kills someone). Third, some philosophers have the goal of crafting analyses of philo-

sophically interesting concepts as they are reflected in ordinary language.

INTENTIONS AND RELATED STATES OF MIND

It is generally agreed that intentions are closely linked to desires—especially *action-desires*, desires to do things—and beliefs. An intention to do something *A* has a motivational dimension, as does a desire to *A*. Having an intention also is widely regarded as requiring the satisfaction of a belief condition of some sort. Few philosophers of action would maintain that people who believe that their chance of winning today's lottery is about one in a million intend to win the lottery, no matter how strongly they desire to win. A relatively popular claim is that having an intention to *A* requires believing that one (probably) will *A*. The proposal is designed to capture, among other things, the confidence in one's success that intending allegedly involves. A less demanding claim is that having an intention to *A* requires that one lack the belief that one (probably) will not *A*. (The agent may have no belief on the matter.) Other alternatives include the requirement that the agent believe to some nonzero degree (even a degree associated with a subjective probability well below 0.5) that he or she will *A* and the requirement that the agent believe that there is a chance that he or she can *A*.

Philosophers are divided on how tight the connection is between intentions, on the one hand, and desires and beliefs, on the other. In particular, they disagree about whether intentions are reducible to combinations of action-desires and beliefs. The central point of contention is whether the settledness that intention encompasses can be captured in terms of beliefs and desires. One who desires to *A*—even someone who desires this more strongly than he or she desires not to *A* and who believes on inductive grounds that he or she probably will *A*—may still be deliberating about whether to *A*, in which case the person is not settled on *A*-ing. Ed wants more strongly to respond in kind to a recent insult than to refrain from doing so, but, owing to moral qualms, he is deliberating about whether to do so. He is unsettled about whether to retaliate, despite the relative strength of his desires and despite his inference from his past behavior in similar situations that he is more likely to retaliate than not to do so (Mele 1992). In acquiring an intention to retaliate—or an intention to refrain from retaliating—Ed becomes settled (but not necessarily irrevocably) on a course of action.

Two ways of coming to intend to *A* should be distinguished. Many philosophers claim or argue that to decide to *A* is to perform a mental action of a certain kind—an action of forming an intention to *A*. According to one version of this view, deciding to *A* is a momentary mental action of intention formation, and it resolves uncertainty about what to do (Mele 2003). The assertion that deciding to *A* is momentary is meant to distinguish it from, for example, a combination of deliberating and deciding. Students who are speaking loosely may say, "I was up all night deciding to major in English," when what they mean is that they were deliberating or fretting all night about what major to declare and eventually decided to major in English. Not all intentions are actively formed. For example, "When I intentionally unlocked my office door this morning, I intended to unlock it. But since I am in the habit of unlocking my door in the morning and conditions … were normal, nothing called for a *decision* to unlock it" (Mele 1992, p. 231). If I had heard a fight in my office, I may have paused to consider whether to unlock the door or walk away, and I may have decided to unlock it. But given the routine nature of my conduct, there is no need to posit an act of intention formation in this case. My intention to unlock the door may have been acquired without having been actively formed.

Some intentions are for the nonimmediate future and others are not. Ann may decide on Tuesday to attend a meeting on Friday, and she may decide now to phone her mother now. The intention formed in the former decision is aimed at action three days in the future. The intention Ann forms when she decides to phone her mother now is about what to do now. Intentions of these kinds are, respectively, *distal* and *proximal* intentions. Proximal intentions also include intentions to continue doing something that one is doing and intentions to start *A*-ing (e.g., start running a mile) straightaway. *Temporally mixed* intentions have both proximal and distal aspects. Consider an intention to watch the movie *Dangerous Intentions* in one sitting, beginning now. Executing it requires doing something now and continued activity for some time.

INTENTION'S FUNCTIONS AND CONSTITUTION

What work do intentions do? And how are they likely to be constituted given that they do this work? Functions plausibly attributed to intentions include initiating and motivationally sustaining intentional actions, guiding intentional action, helping to coordinate agents' behavior over time and their interaction with other agents, and

prompting and appropriately terminating practical reasoning (see Brand 1984, Bratman 1987, McCann 1998, Mele 1992, and Searle 1983).

Intentions, like many psychological states, have both a representational and an attitudinal dimension. The representational content of an intention may be understood as a *plan*. The intending attitude toward plans may be termed an *executive* attitude. Plans, on one conception, are purely representational and have no motivational power of their own. People have many different attitudes toward plans, in this sense. They may believe that a plan is too complicated, admire it, hope that it is never executed, and so on. To understand the executive dimension of intention—something at work in the initiation of action—recall that intending to *A*, unlike desiring to *A*, is partially constituted by being settled on *A*-ing. To have the intending attitude toward a plan is to be settled (but not necessarily irrevocably) on executing it. In virtue of this motivational feature of intentions, acquisitions of proximal intentions are well suited to the task of initiating actions and the persistence of intentions that initiate actions is well suited to sustain them. (In the case of an intention for a not-doing—for example, an intention not to vote in tomorrow's election—the agent may instead be settled on not violating the simple plan embedded in it, the plan not to vote.)

Why do acquisitions of proximal intentions initiate and sustain the actions that they do? Why, for example, does acquiring a proximal intention to order a hamburger and fries initiate and sustain one's ordering a hamburger and fries rather than one's ordering a salad or one's singing a song? Attention to the representational side of intentions provides an answer. An intention to *A* incorporates a plan for *A*-ing, and which intentional action(s) an intention generates is a partial function of the intention-embedded plan. In the limiting case, the plan in an intention has a single node. It is, for example, a prospective representation of one's pushing a window closed. Often, intention-embedded plans are more complex. The proximal intention to check his bank account online that Bob is executing incorporates a plan that includes clicking on his bank's link, then typing his ID and password in a certain pair of boxes, and so on. Agents who successfully execute an intention are guided by the intention-embedded plan. The guidance depends on agents monitoring progress toward their goals. The information (or misinformation) that Bob has entered his ID, for example, helps to produce his continued execution of his plan.

Although the content of an intention is a plan, such expressions as "Bob's intention to check his bank account now" and "Ann intends to shoot pool tonight" are common. It should not be inferred from such expressions that the agent's intention-embedded plan is structurally simple. Often, ordinary expressions of an agent's motivational attitudes do not identify the full content of the attitude and are not meant to. Bob says, without intending to mislead, "Ann wants to shoot pool tonight," even though he knows that what she wants is to play eight-ball with him at Pockets tonight for a dollar a game until the place closes, as they normally do.

Intention's coordinative capacities lie both in its executive aspect, which includes settledness, and in its plan component. Comprehensive plans for extended activity can be constructed out of plans embedded in less inclusive intentions, and developments in plans will be influenced and constrained by what one is already settled on doing. (This is not to deny the possibility of revising earlier intentions.) Moreover, knowledge of what others are settled on doing assists one in forming intentions and plans for cooperative ventures. To the extent to which coordination depends on practical reasoning, intention promotes coordination by providing motivation for required reasoning—motivation deriving from the settledness intention encompasses. Michael Bratman argues that the coordinating roles of distal intentions rest on several features of these intentions: they have the capacity to control behavior, they "resist (to some extent) revision and reconsideration," and they involve dispositions to reason with a view to intention-satisfaction and "to constrain one's intentions in the direction of consistency" (1987, pp. 108–109). All of these features are tied to the settledness intentions encompass.

Intention is an appropriate terminator of practical reasoning precisely because in forming or acquiring an intention one becomes settled on a course of action. Practical reasoning is aimed at action; and, if all goes well, one does what one has become settled on doing on the basis of one's practical reasoning. Intention's capacity to prompt such reasoning, as just noted, also derives from the settledness it involves.

INTENTIONS AND REASONS

Are people's reasons for intending to *A* limited to their reasons for *A*-ing? Gregory Kavka's (1983) "toxin puzzle" suggests that they are not. In this puzzle, a trustworthy billionaire offers you a million dollars for intending tonight to drink a certain toxin tomorrow afternoon. You are convinced that he can tell what you intend independ-

ently of what you do. Although drinking the toxin would make you ill for a day, you do not need to drink it to get the money. Constraints on prize-winning intentions include prohibitions against creating special incentives for yourself to drink the toxin, various tricks, and forgetting relevant details of the offer. For example, you will not receive the money if you hire a hit man to kill you should you not drink the toxin or persuade a hypnotist to implant the intention in you. If, by midnight tonight, without violating any rules, you intend to drink the toxin tomorrow afternoon, you will find a million dollars in your bank account when you awake tomorrow morning. Because you are well aware of this point and would love to be a millionaire, you seemingly have a great reason to form the intention. Now, you probably would drink the toxin for a million dollars. But can you, without violating the rules of the offer, intend tonight to drink it tomorrow? Apparently, you have no reason to drink the toxin and an excellent reason not to drink it. Seemingly, you will infer from this that you will not drink the toxin. Indeed, it seems that you will be confident that you will not drink it, and your confidence in that seems inconsistent with your having an intention to drink it.

Kavka draws the moral that intentions are "dispositions to act that are based on *reasons to act*—features of the act itself or its (possible) consequences that are valued by the agent" (1983, p. 35). However, because not all the work in Kavka's puzzle is done by truths about intention, reasons, and the like, his perfectly general claim about intentions cannot be established by reflection on the puzzle. Were it not for the rule against forgetting, for example, you could become a millionaire. If, tonight, you can so arrange things that at midnight you will be confident that the toxin will be in your favorite afternoon drink tomorrow and confident, as well, that by tomorrow you will have forgotten about the toxin, then at midnight you can intend to drink the toxin tomorrow. The content of your intention may be described roughly as follows: "Tomorrow afternoon, I drink the toxin unintentionally while sipping my customary afternoon tea." Even though you will have a reason tomorrow to drink tea, you will have no reason at all to drink the toxin; and that is clear to you at midnight. This scenario falsifies the idea that all possible intentions to *A* are based on reasons to *A*. A more cautious diagnosis of your apparent inability to intend to drink the toxin given the constraints Kavka imposes is that having an intention to *A* is inconsistent with being convinced that one will not *A*.

The preceding scenario leaves open a more modest version of Kavka's moral. Perhaps all possible intentions to *A* such that in executing them one would *intentionally A* are based on reasons to *A*. Although one cannot find in reasons for *A*-ing a necessary basis for all possible intentions to *A*, one may find in them a necessary basis for all intentions of the sort just identified—*orthodox intentions*. The relatively cautious diagnosis previously mentioned provides a hint about how to test this hypothesis. Might there be agents who know that they have no reason to drink the toxin, have not forgotten anything relevant, and nevertheless believe that they will drink it?

Consider the following story. An evil genius tricks Ted into drinking nonlethal liquid toxins whenever such toxins happen to be nearby, and Ted is well aware of this. Ted also has—as he knows—a condition called *intention perseverance*: once he forms an intention, he will not abandon it unless he has a good reason to abandon it. Finally, Ted is indifferent between drinking toxins unintentionally and drinking them intentionally: only the subsequent illness bothers him.

Seemingly, Ted can get the big prize in Kavka's scenario. Although normal folks are confident that they will not drink the toxin, Ted is confident that he will drink it. He also has an excellent reason to decide to drink it: in so deciding he would form an intention that will make him a millionaire. And he can count on the intention formed in his decision to persist and to result in intentional toxin drinking, given that he lacks a good reason to abandon the intention after he forms it. Ted's intention to drink the toxin is such that, in executing it, he intentionally drinks the toxin. So he undermines even the more modest version of Kavka's moral. His intention to drink the toxin is based on his reasons for forming that intention, and it is not based at all on reasons for drinking the toxin. This leads back, then, to the relatively cautious diagnosis of one's apparent inability to intend to drink the toxin. The diagnosis is about a completely general connection between intention and belief, not a completely general connection between intention and reasons: having an intention to *A* is inconsistent with being convinced that one will not *A*.

Sometimes people consider reasons for and against taking a prospective course of action. Gilbert Harman (1986) and Michael Bratman (1987) argue that the concept of intentional action is sensitive to reasons agents have for not doing what they do in a way in which the concept of intention is not. The upshot is that agents sometimes intentionally do things that they lack an intention to do. For example, Bill knows that his vacuuming his carpets today will cause Beth to sneeze, and he counts that as a reason not to vacuum them today. Even so,

because he believes that it is important to vacuum today, he does so, and he notices Beth sneezing as he works. Harman and Bratman would say that even though making Beth sneeze is no part of what Bill intends, he intentionally makes her sneeze. This judgment may be in line with ordinary usage of the terms at issue, and it may be a judgment that a majority of nonspecialists would make. However, granting the existence of intentionally produced side effects that the agent does not intend to produce would complicate the task of philosophers of action who say that they are in the business of explaining how intentional actions are produced by their agents. They would need a theory that explains intentional actions of two different kinds: actions the agent is trying to perform and actions the agent is not trying to perform. Such philosophers may do well to seek—and to set up as the target of their explanatory efforts—a more circumscribed notion of intentional action that is no more sensitive to reasons against doing what one does than Harman and Bratman say the concept of intention is.

See also Belief; Content, Mental; Propositional Attitudes.

Bibliography

Anscombe, G. E. M. *Intention*. 2nd ed. Ithaca, NY: Cornell University Press, 1957.

Audi, Robert. *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press, 1993.

Brand, Myles. *Intending and Acting*. Cambridge, MA: MIT Press, 1984.

Bratman, Michael. *Faces of Intention*. Cambridge, U.K.: Cambridge University Press, 1999.

Bratman, Michael. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Harman, Gilbert. *Change in View*. Cambridge, MA: MIT Press, 1986.

Kavka, Gregory. "The Toxin Puzzle." *Analysis* 43 (1983): 33–36.

McCann, Hugh. *The Works of Agency*. Ithaca, NY: Cornell University Press, 1998.

Mele, Alfred. *Motivation and Agency*. New York: Oxford University Press, 2003.

Mele, Alfred. *Springs of Action*. New York. Oxford University Press, 1992.

Searle, John. *Intentionality*. Cambridge, U.K.: Cambridge University Press, 1983.

Alfred R. Mele (2005)

INTENTIONALITY

The term "intentionality" was used by Jeremy Bentham to distinguish between actions that are intentional and those that are not. It was reintroduced by Edmund Husserl in connection with certain doctrines set forth in Franz Brentano's *Psychologie vom empirischen Standpunkt* (1874). The word is now used primarily in this second sense.

Brentano wrote:

> Every mental phenomenon is characterized by what the scholastics of the Middle Ages called the intentional (and also mental) inexistence of an object, and what we would call, although not in entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality …), or an immanent objectivity. Each one includes something as an object within itself, although not always in the same way. In presentation something is presented, in judgment something is affirmed or denied, in love [something is] loved, in hate [something] is hated, in desire something is desired, etc.
>
> This intentional inexistence is exclusively characteristic of mental phenomena. No physical phenomenon manifests anything similar. Consequently, we can define mental phenomena by saying that they are such phenomena as include an object intentionally within themselves. (*Op. cit.,* Vol. I, Book II, Ch. 1)

This passage contains two different theses: one, an ontological thesis about the nature of certain objects of thought and of other psychological attitudes; the other, a psychological thesis, implying that reference to an object is what distinguishes the mental or psychological from the physical. These two theses are the subject matter of the present article. It should be noted, however, that "intentionality" is also used in connection with certain other related theses of phenomenology and existentialism.

INTENTIONAL INEXISTENCE

The problem that gave rise to the ontological thesis of intentional inexistence may be suggested by asking what is involved in having thoughts, beliefs, desires, purposes, or other intentional attitudes, which are directed upon objects that do not exist. There is a distinction between a man who is thinking about a unicorn and a man who is thinking about nothing; in the former case, the man is

intentionally related to an object, but in the latter case he is not. What, then, is the status of this object? It cannot be an actual unicorn, since there are no unicorns. According to the doctrine of intentional inexistence, the object of the thought about a unicorn *is* a unicorn, but a unicorn with a mode of being (intentional inexistence, immanent objectivity, or existence in the understanding) that is short of actuality but more than nothingness and that, according to most versions of the doctrine, lasts for just the length of time that the unicorn is thought about.

EARLY THEORIES. St. Anselm's ontological argument was thus based upon the assumption that, if God is thought about, he thereby "exists in the understanding." Anselm then proceeded to contrast the perfections of that which "exists in the understanding alone" with that which "exists in reality." Peter Aureol and William of Ockham contrasted the intentional existence of the objects of thought with the subjective existence of the thoughts themselves. The term "objective existence," referring to the existence of something *as* an object of thought, was used by medieval philosophers and by René Descartes as a synonym for "intentional existence"; Descartes thus contrasted the formal, or subjective, existence of actual objects with the objective existence in the mind of objects that are merely thought about. The terms *objective* and *subjective,* in these uses, had connotations quite different from those that they have now; that which was said to have objective existence (for instance, a unicorn as an object of thought), unlike that which had subjective existence (the idea of a unicorn, for instance), need not exist in fact.

Advantage of the doctrine. The doctrine of intentional existence, or, as Brentano called it, intentional inexistence, had at least the advantage of providing a literal interpretation for the dictum that truth consists in a kind of correspondence between mind and thing: an affirmative judgment is true if the properties of the intentional object are the same as those of the actual object. The very statement of this advantage, however, betrays the fact that the judgment is directed, not upon the intentional object, but upon the actual object, in which case, as Pierre Gassendi pointed out, the intentional object would seem to be superfluous.

Intentional reference. The difficulty of the apparent superfluity of the intentional object may be traced, in part, to the fact that the phenomenon of intentionality has two sides. Our intentional attitudes may be directed upon objects that do not exist (Diogenes looked for an honest man), but they may also be directed upon objects that do exist (there is a certain dishonest man whom the police happen to be looking for). The object of the latter quest, obviously, is not a thing having only immanent or intentional existence. But this is also true, as Brentano was later to point out, of the object of the former quest: Diogenes was not looking for an immanent object (for, if the doctrine of intentional inexistence were true, he already had one in his mind); he was looking for an actual, existing honest man, despite the fact that, as we may suppose, no such man exists. Thus, Brentano said, "If we think about a horse, the object of our contemplation is a horse and not a contemplated horse."

In the expression of the ontological thesis of intentionality, "intentional" may be said to refer to a mode of being *within* the mind; but in the expression of the psychological thesis of intentionality, "intentional" is used to describe the direction upon objects that may exist *outside* the mind. It is not inaccurate to say that intentional entities were posited in the attempt to account for intentional reference, but precisely because they were intentional, the attempt did not succeed. Husserl said, in the fifth of his *Logische Untersuchungen,* that the objects of our intentional experiences are never immanent—never intentional objects—but are always transcendent.

BRENTANO'S LATER VIEWS. Thus, for various reasons Brentano abandoned the ontological part of his doctrine of intentionality. In his later writings, he said that "unicorn" in the sentence "John is thinking about a unicorn" has no referential function; a contemplated unicorn is not a type of unicorn. "Unicorn," in such sentences, is used syncate-gorematically to contribute to the description of the person who is said to have a unicorn as the object of his thought. But this conclusion seems to leave us with our problem. The statement "John is thinking about a unicorn" does not describe *John* as a unicorn; how, then, does "unicorn" serve to contribute to his description?

The ontological problem, therefore, may be said to survive in the question, "How are we using 'unicorn' in 'John believes that there are unicorns'?" There is a temptation to say that the use of "unicorn" in such sentences has no connection at all with the use it would have in "There are unicorns." That this would be false, however, may be seen by noting that "John believes that there are unicorns" and "All of John's beliefs are true" together imply "There are unicorns." Thus, Ludwig Wittgenstein remarked:

> One may have the feeling that in the sentence "I expect he is coming" one is using the words "he is coming" in a different sense from the one they

> have in the assertion "He is coming." But if it were so, how could I say that my expectation had been fulfilled? If I wanted to explain the words "he" and "is coming," say by means of ostensive definitions, the same definitions of these words would go for both sentences. (*Philosophical Investigations,* p. 130e)

CARNAP'S THEORY. In the *Logical Syntax of Language* (London, 1937), Rudolf Carnap suggested that linguistic entities are the objects of our intentional attitudes. "Charles thinks (asserts, believes, wonders about) *A,*" he said, might be translated as "Charles thinks '*A.*'" Taken literally, this suggestion would imply, falsely, that a man who wonders whether there are unicorns is a man who wonders whether there is the word *unicorns.*

INSCRIPTIONAL THEORY. A closely related view has been developed by W. V. Quine and Israel Scheffler. These authors, however, instead of saying that our intentional attitudes have linguistic entities as their objects, suggest instead that certain sentences, which relate people to words or to other linguistic entities, might be used to perform all of the functions of intentional sentences; if this view were adequate, the problem of the status of the intentional object might be avoided. Thus, "John believes-true a Socrates-is-mortal inscription" may be interpreted as a sentence affirming a certain relation to hold between John and a linguistic entity or "inscription," but a relation that is true only under the conditions under which "John believes that Socrates is mortal" is true; hence, if we use the former sentence instead of the latter, we relate John only to inscriptions.

However, it may be held (1) that the plausibility of this approach depends upon the assumption that there are certain semantic sentences (for instance, "The German sentence '*Sokrates ist sterblich*' means that Socrates is mortal") that are true of certain inscriptions and (2) that these semantic sentences are abbreviations for intentional sentences that leave us with our original problem (for instance, "German-speaking people use '*Sokrates ist sterblich*' to express and convey the belief that Socrates is mortal").

This inscriptional approach, moreover, fails to distinguish between such sentences as "Someone is looking for a horse" and "There is a horse that someone is looking for"; these two types of sentence, as noted above, reflect the two different sides of the phenomenon of intentionality. It has been suggested that sentences of the latter sort may be illegitimate, on the ground that they quantify, in effect, into contexts that are referentially opaque, in a sense explained below. To say that such intentional sentences are illegitimate is to imply that the mind is incapable of referring to objects that exist and, hence, that we cannot "get outside the circle of our own ideas."

RESPONSE THEORY. There have been still other approaches to the problem of the intentional object. Some of the American New Realists proposed, behavioristically, that to think about a unicorn might merely be to "put one's unicorn responses in readiness." The thinker, instead of relating himself to unicorns, disposes himself to behave in just those ways in which he would behave if there *were* unicorns. Other psychological attitudes were treated analogously. It would seem, however, that "unicorn responses" cannot be adequately specified except by reference to beliefs and desires that are directed upon unicorns, since the ways in which a man would respond to a unicorn would be, in part, a function of what he otherwise perceives, desires, and believes. More recent revivals of the specific-response theory seem to be subject to similar difficulties.

CHURCH'S VIEW. Alonzo Church, in his *Introduction to Mathematical Logic* (Princeton, NJ, 1956), proposed that the sentence "Schliemann sought the site of Troy" asserts that a certain relation holds between Schliemann and the *concept* of the site of Troy; Church said, negatively, that the relation is "not quite like that of having sought," but he did not say more positively what it is. This view suggests a return to the medieval doctrine, at least to the extent of viewing the objects of our intentional attitudes as beings of reason.

ANALOGICAL THEORY. Thomas Aquinas seems to have held that "unicorn," in such sentences as "John is thinking about a unicorn," is used *analogically* (*De Potentia* 7c; *Summa Theologiae* 1, 13, 10). There is ground for questioning whether the doctrine of analogical predication is itself sufficiently illuminating to throw light upon the problem of intentionality, but the fact that we can understand the use of "unicorn" and cannot say just what function the word there performs may, on the other hand, throw some light upon the doctrine.

The most plausible defense of the doctrine of intentional inexistence, therefore, would seem to be that this doctrine, unlike most of its alternatives, *does* provide us with a straightforward account of the use of "unicorn" in "John is thinking about a unicorn": The word is being used simply to designate a unicorn.

PSYCHOLOGICAL THESIS OF INTENTIONALITY

According to Brentano's second thesis, intentionality is peculiar to psychological phenomena and thus provides a criterion by means of which the mental may be distinguished from the nonmental. The problem for the proponent of this second thesis is not so much that of showing that mental phenomena *are* intentional as it is that of showing that physical phenomena are *not* intentional. Some now believe that the thesis can be defended by reference to the language we use in describing psychological phenomena—that the sentences we must use in describing psychological phenomena have certain logical properties that are not shared by any of the sentences we must use in describing nonpsychological phenomena, and that these properties are correctly called intentional. If this view is true, then the basic thesis of physicalism and the unity of science is false. Can we find, then, a logical criterion of the intentional, one that we may then use to distinguish the mental from the physical?

UNSATISFACTORY CRITERIA OF INTENTIONALITY. It has been suggested that failure of existential generalization yields a logical criterion of the intentional. The intentional "John is thinking about a horse," unlike the nonintentional "John is riding on a horse," does not imply that there are horses. However, existential generalization also fails in application to some of the terms in the following statements that describe physical phenomena: "New Zealand is devoid of unicorns," "That lady resembles a mermaid," and "The dam is high enough to prevent any future floods."

Nonextensional occurrence has also been proposed as a possible criterion of the intentional. A phrase, *p,* may be said to occur nonextensionally in a sentence, *s,* provided that the result of replacing *p* in *s* by any phrase having the same truth value as *p* will be a sentence having the same truth value as *s.* Thus, "Johnson is Kennedy's successor" may replace "Socrates was a philosopher" in "Either Socrates was a god or Socrates was a philosopher," without altering the truth value of the whole, whereas similar replacement is not possible in "Plato believed that Socrates was a philosopher." Nonextensional occurrence, however, is not peculiar to sentences that are intentional; compare "It is necessarily true that if Socrates was a member of the class of philosophers, then Socrates was a philosopher."

Referential opacity has also been proposed as a criterion of the intentional. The occurrence of a substantival expression in a sentence, *s* (for instance, "Truman's successor" in "Joe Martin believed that Dewey would be Truman's successor"), is referentially opaque if its replacement in *s* by another substantival expression (such as "Eisenhower") designating the same individual may result in a sentence having a truth value different from that of *s.* However, referential opacity is not peculiar to the intentional; we may assert "It is necessarily true that if Dewey was Truman's successor, then Dewey was Truman's successor," but not "It is necessarily true that if Dewey was Truman's successor, then Dewey was Eisenhower."

SATISFACTORY CRITERIA. The failure of nonextensional occurrence and referential opacity has led some to believe that there are no logical characteristics peculiar to intentional statements. However, there are other criteria that do seem to be satisfied only by intentional statements. We may mention two.

Let us refine upon ordinary English in the following way: instead of writing propositional clauses as "that" clauses, we will eliminate the "that" and put the remainder of the clause in parentheses; for example, instead of writing "John believes that there are men," we will write "John believes (there are men)." A simple sentence prefix may be said to be an expression that contains no proper part that is logically equivalent to a sentence or to a sentence function and that is such that the result of prefixing it to a sentence in parentheses is another sentence. We may say that a simple sentence prefix, *M,* is *intentional* if, for every sentence *p,* *M*(*p*) is logically contingent. Thus, "it is impossible" is not intentional, since when prefixed to "(some squares are circles)," it yields a sentence that is necessary and therefore not contingent; "it is right" is not intentional since, when prefixed to "(there is not anything of which it can be truly said that it is right)," it yields a sentence that is contradictory and therefore not logically contingent.

However, *every* sentence, whether it is itself contingent or not, is such that the result of thus prefixing it by "John believes" is contingent. Similar observations apply to "John questions," "John desires," and to other prefixes referring to intentional attitudes. Thus, we might say that the psychological differs from the nonpsychological in this respect: an adequate description of the psychological requires the use of intentional prefixes.

It may also be argued that some intentional prefixes (for instance, "John believes") are such that the possible ways of inserting them into a universally quantified sentence (for instance, into "For every *x, x* is material") and into the corresponding existentially quantified sentence

("There exists an *x* such that *x* is material") yield four statements ("John believes that, for every *x, x* is material"; "For every *x,* John believes that *x* is material"; "John believes that there exists an *x* such that *x* is material"; and "There exists an *x* such that John believes that *x* is material") that are logically related in ways in which no corresponding sentences with nonintentional prefixes are related. Thus, it may be said of the four sentences just cited: Neither the first nor the third implies any of the others; the second implies all but the first; the fourth implies the third but does not imply either the first or the second; and there is no nonintentional prefix that will yield four sentences that are similarly related. This contention, to the extent that it applies to "John believes," is based upon the assumptions that in believing a thing to have certain properties, one thereby believes that the thing exists; that one may believe falsely, of some nonuniversal set of things (some set comprising less than everything there is), that it comprises everything there is; and that one may believe falsely, of a universal set of things, that it does not comprise everything there is.

There are other psychological sentences—for instance, "He is in pain" and "He is thinking about Jupiter"—that may not satisfy the above criteria of intentionality. The first of these sentences, however, might be said to be intentional if, as some believe, one cannot be in pain if one is not *aware* that one is in pain; or if one does not *believe* that one is in pain; or if, at any one instant, one does not *remember* the pain of previous instants; and analogously for the second quoted sentence. Another possible view, however, is to say that intentionality is at least a sufficient if not a necessary condition of the psychological.

See also Anselm, St.; Bentham, Jeremy; Brentano, Franz; Carnap, Rudolf; Church, Alonzo; Existentialism; Gassendi, Pierre; Husserl, Edmund; Language and Thought; Meaning; New Realism; Nonexistent Object, Nonbeing; Ontological Argument for the Existence of God; Peter Aureol; Phenomenology; Philosophy of Mind; Physicalism; Quine, Willard Van Orman; Reference; Thomas Aquinas, St.; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

Brentano, Franz. "The Distinction between Mental and Physical Phenomena." Translated by D. B. Terrell, in *Realism and the Background of Phenomenology,* edited by R. M. Chisholm, 39–61. Glencoe, IL: Free Press, 1960. Translation of part of Vol. I, Book II, Ch. 1 of Brentano's *Psychologie.*

Brentano, Franz. *Kategorienlehre.* Leipzig: F. Meiner, 1933.

Brentano, Franz. *Psychologie vom empirischen Standpunkt.* Vienna, 1874; 3rd ed., Leipzig, 1925.

Chisholm, R. M. "Notes on the Logic of Believing." *Philosophy and Phenomenological Research* 24 (1963): 195–201.

Chisholm, R. M. *Perceiving: A Philosophical Study.* Ithaca, NY: Cornell University Press, 1957.

Hayen, André. *L'intentionnel selon saint Thomas,* 2nd ed. Paris, 1954.

Husserl, Edmund. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie.* Halle, Germany: Niemeyer, 1913. Translated by W. R. Boyce Gibson as *Ideas: General Introduction to Pure Phenomenology.* New York: Macmillan, 1931.

Husserl, Edmund. *Logische Untersuchungen,* 4th ed. Halle, Germany, 1928.

Quine, W. V. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53 (1956): 177–187.

Quine, W. V. *Word and Object.* Cambridge, MA: Massachusetts Institute of Technology Press, 1960.

Scheffler, Israel. *The Anatomy of Inquiry.* New York: Knopf, 1963.

Sellars, Wilfrid, and R. M. Chisholm. "Intentionality and the Mental." In *Concepts, Theories, and the Mind-Body Problem,* edited by H. Feigl, M. Scriven, and G. Maxwell, 507–539. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, 1958.

Spiegelberg, Herbert. "Der Begriff der Intentionalität in der Scholastik, bei Brentano, und bei Husserl." *Philosophische Hefte* 5 (1936): 75–91.

Roderick M. Chisholm (1967)

INTENTIONALITY [ADDENDUM]

The medieval word *intentionality* was reintroduced into late nineteenth-century philosophy by Franz Clemens Brentano. Intentionality is the power of minds to represent, stand for, or be about things, properties, and states of affairs. The English word *intentionality* stems from the Latin verb *intendere*, which can be used to denote the act of stretching a bow string with the aim of propelling an arrow into its target. In Brentano's sense intentionality is the mental tension whereby the human mind aims at objects. The nature of intentionality has been a central topic in the philosophy of mind and language in the twentieth century in both the phenomenological tradition (founded by Edmund Husserl, a student of Brentano) and the analytic tradition.

In several well-known paragraphs from his 1874 classical work, *Psychology from an Empirical Standpoint*, quoted by Roderick M. Chisholm at the beginning of his entry, Brentano did two things: he provided a puzzling definition of intentionality and he put forward the thesis

that intentionality is the mark of the mental. This entry briefly considers some of the logical and ontological puzzles raised by Brentano's definition of intentionality. It then turns toward issues raised by Brentano's thesis that intentionality is the mark of the mental, considering whether only mental (or psychological) phenomena exhibit intentionality and whether all do, in this order.

LOGICAL AND ONTOLOGICAL PUZZLES

WHAT INTENTIONALITY DOES NOT MEAN. To grasp the philosophical significance of the concept expressed by the word *intentionality* for contemporary philosophy, it is important not to confuse it with two of its cognates: *intention* and *intension*. *Intention* and *intending* refer to specific mental states, events, or processes that, unlike others (e.g., beliefs, judgments, expectations, perceptions, fears, desires, and hopes), play a distinctive role in the etiology, the preparation, and the execution of actions. By contrast, in Brentano's philosophical tradition, intentionality is a property of many—if not all—such mental states and events as beliefs, judgments, desires, and perceptions. Nor should *intentionality* be confused with *intensionality* (with an *s*). *Intensional* and *intensionality*, which mean respectively "nonextensional" and "nonextensionality," refer to logical features of words and sentences. To take Willard Van Orman Quine's famous example, "creature with a heart" and "creature with a kidney" have the same extension because they apply to the same things: all the creatures with a heart have a kidney and conversely. But the two expressions have different intensions because *heart* and *kidney* have different meanings and extensions.

THE IMPACT OF BRENTANO'S DEFINITION. Brentano defined *intentionality* as what enables a psychological state or act to be directed on what he called an intentional object. He further ascribed to intentional objects the puzzling property that he called intentional inexistence or immanent objectivity. Finally, Brentano's quote suggests that a satisfactory explanation of intentionality must be able to account for the inferences illustrated by the following sentences, which contain a psychological verb that refers to a mental intentional state (e.g., "to watch," "to admire," or "to fear"), in which (b) is a conclusion that follows from premise (a), and (c) represents the logical form of (b):

(1a) George W. Bush watches Laura Bush.

(1b) George W. Bush watches something.

(1c) (∃x) (George W. Bush watches x).

(2a) George W. Bush admires Sherlock Holmes.

(2b) George W. Bush admires something.

(2c) (∃x) (George W. Bush admires x).

(3a) George W. Bush fears God.

(3b) George W. Bush fears something.

(3c) (∃x) (George W. Bush fears x).

As illustrated by sentences (1) to (3), all sorts of things can be the target of intentionality and, therefore, qualify as intentional objects: concrete physical objects in space and time (e.g., stars, stones, plants, animals, and persons), abstract objects (e.g., numbers), and objects of fiction (e.g., Sherlock Holmes) that are neither in space nor in time, and even objects whose properties are inconsistent with known natural or geometrical laws (e.g., unicorns and squared circles). As Chisholm elegantly puts it in his entry, although unicorns do not exist, thinking about a unicorn is not the same thing as thinking about nothing. Arguably, by ascribing to intentional objects the property he called intentional inexistence or immanent objectivity, what Brentano presumably meant was that intentional objects need not exist in space and time: it is enough that they exist within the mind.

ARE THERE THINGS THAT FAIL TO EXIST? Brentano's definition gave rise to a fundamental gap in twentieth-century philosophical logic between intentional objects theorists, who claimed that there are things that do not exist, and their opponents, who denied it. On the one hand, Alexius Meinong (who was Brentano's student) and his followers took it as a genuine condition of adequacy on an account of intentionality that it ought to provide a uniform account for the validity of all inferences from premises of type (a) to conclusions of type (b), in sentences of type (1) to (3), notwithstanding the fundamental ontological differences between the kinds of entities over which the bound variable ranges in the logical form of conclusions of type (c).

For example, Meinong (1904/1960) supposes that for any set of properties, there is an intentional object that instantiates them, but only some, not all, such objects exist. On his view, it is one thing to say that there are unicorns. It is another thing to say that unicorns exist. On the contrary, following Bertrand Arthur William Russell, Quine (1960) and most twentieth-century analytic philosophers denied that there are things that do not exist, on the threefold grounds that (1) existence is not a property (or the English verb "to exist" should be treated as a quantifier, not as a predicate); (2) the theory of inten-

tional objects incurs intolerably heavy ontological commitments; and (3) an ontology of nonexistent objects can hardly be reconciled with the ontology of the contemporary natural sciences. The rest of this entry is devoted to issues that arise only if one denies that there are things that do not exist.

PSYCHOLOGICAL AND LINGUISTIC INTENTIONALITY

What Quine (1960) calls semantic ascent (and others the linguistic turn) enables one to rise from talk about things to talk about talk about things (i.e., words and sentences). As Chisholm explains in the second half of his entry, the ascent from talk about intentionality to talk about talk about intentionality raised the prospect of a linguistic doctrine of intentionality and it also gave rise to an objection to Brentano's thesis that only mental phenomena exhibit intentionality.

IS INTENSIONALITY A CRITERION OF INTENTIONALITY? Sentences about intentionality exhibit the two logical features of intensionality (or referential opacity). First, sentences about intentionality do not always license substitution of coreferential terms *salva veritate*. Thus, even though sentence (4a) expresses a true proposition, it does not automatically follow that sentence (4b) does, even though "Cicero" and "Tully" name one and the same individual:

(4a) George W. Bush believes that Cicero was bald.

(4b) George W. Bush believes that Tully was bald.

If George W. Bush fails to know that "Cicero" and "Tully" are coreferential, then the truth of (4a) does not entail the truth of (4b). Second, sentences about intentionality do not always license the rule of existential generalization. A speaker may truly believe that Bush holds the belief ascribed by (5a) without committing him- or herself to the truth of (5b):

(5a) George W. Bush believes that Zeus transformed himself into a bull.

(5b) (∃x) (George W. Bush believes that x transformed himself into a bull).

On this basis, Chisholm (1957) argues, on the one hand, that linguistic descriptions of intentional phenomena, which fail the criteria of extensionality, cannot be replaced by descriptions of observable behavior (or bodily movements), which satisfy the criteria of extensionality. On the other hand, he contemplates the prospect that the intensionality of sentences describing intentional phenomena might constitute a logical criterion for the intentionality of the described phenomenon. This criterion fails for two reasons. First, reports of visual perception may be extensional. For example, if "Ralph saw Laura" is true, then "Ralph saw something" is true. Furthermore, if "Laura" and "the president's wife" are coreferential and if "Ralph saw Laura" is true, then "Ralph saw the president's wife" cannot fail to be true. So if "Ralph saw Laura" truly reports an intentional phenomenon, then the intensionality of a report cannot be a necessary condition for the intentionality of the reported phenomenon. Second, many sentences that do not describe intentionality (e.g., sentences about natural laws) exhibit intensionality. So, the intensionality of a report cannot be a sufficient condition for the intentionality of the reported phenomenon.

DERIVED VERSUS UNDERIVED INTENTIONALITY. Brentano's thesis that only mental phenomena exhibit intentionality can be questioned on the grounds that the utterances of sentences, which have meaning but are nonmental things, exhibit intentionality. Most contemporary philosophers of mind, such as Jerry A. Fodor (1987), John Haugeland (1981), and John R. Searle (1983), would respond to this objection to Brentano's thesis by offering a distinction between the underived (or primitive) intentionality of a speaker's mental states and the derived intentionality of the utterances of the sentences whereby he or she expresses his or her mental states. On their view, sentences of natural languages have no intrinsic meaning and would be deprived of meaning if humans did not use them for the purpose of making their private thoughts known to others. Thus, according to the revised version of Brentano's thesis, only mental phenomena have underived intentionality. Daniel C. Dennett (1987), however, disagrees because, according to him, the distinction between derived and underived intentionality is just an illusion.

HOW TO NATURALIZE INTENTIONALITY

QUINE'S DILEMMA. The thesis that only mental phenomena possess underived intentionality lead both Chisholm and Brentano to embrace a version of the ontological dualist distinction between mental and physical things. Quine (1960) agrees with Chisholm's thesis that sentences describing an agent's intentional phenomena cannot be successfully paraphrased into sentences about the agent's behavior or bodily movements: intentionality, therefore, cannot be naturalized. He does not, however,

endorse ontological dualism. Instead, he embraces a physicalist ontology according to which everything, including allegedly mental things, are physical, and he puts forward an influential dilemma: since intentionality cannot be naturalized, one must choose between a physicalist ontology and intentional realism (i.e., the view that intentionality is real). Some of the physicalists who accept Quine's dilemma (e.g., Churchland 1989) embrace eliminative materialism and deny purely and simply the reality of beliefs and desires. It is, however, difficult to make sense of the belief that there are no beliefs. Others (such as Dennett 1987) take the instrumentalist view that although the intentional idiom is a useful stance for predicting a complex physical system's behavior, it lacks an explanatory value. But the question arises how the intentional idiom could make useful predictions if it fails to describe and explain anything.

As a result of the difficulties inherent to both eliminative materialism and interpretive instrumentalism, several physicalists choose to deny both Quine's challenge and Brentano's thesis that only nonphysical things can exhibit intentionality. Their project is to naturalize intentionality. Since they are physicalists, they assume that all so-called mental things are physical things, and their goal is to show that uncontroversially physical things exemplify, if not full-blown intentionality, at least the seeds of intentionality.

INFORMATIONAL SEMANTICS. One influential strategy for doing so has been Fred I. Dretske's (1981) proposal that a device that carries information exhibits some degree of intentionality. According to Paul Grice's (1957/1989) distinction between natural and nonnatural meaning, whereas the English word *fire* nonnaturally means fire, *smoke* naturally means fire. Dretske extends the Gricean notion of natural meaning into an information-theoretic approach according to which the informational relation is the converse of a nomic relation. If the length of a metal bar is nomically correlated with variations in temperature, then the former carries information about the latter. The direction of the needle of a compass carries information about the location of the North Pole because it nomically co-varies with the location of the North Pole. If it is not a law that polar bears live at the North Pole, then a compass will fail to indicate where polar bears live even though it indicates the location of the North Pole and this is where polar bears happen to live. If so, then the linguistic report of the information carried by a signal exhibits some of the intensionality (with an *s*) of reports of intentionality.

It is widely recognized, however, that pure informational semantics fails to generate the full intentionality of human mental representations. It is part and parcel of human mental representations that they can misrepresent what they are about. Since the informational relation is the converse of a nomic correlation, it cannot account either for misrepresentation or for the normativity of the contents of representations.

TELEOSEMANTICS. A second major proposal for naturalizing intentionality is the teleosemantic approach championed by Ruth Garrett Millikan (1984). It starts by offering a direct solution to the problems that plague informational semantics, namely the problems of misrepresentation and the normativity of mental content. Its basic insight is that what Brentano calls intentional inexistence is exemplified by biological (so-called teleological) functions and is, therefore, a particular case of a more general biological phenomenon. For example, if it is the function of a mammal's heart to pump blood, then a mammal's heart ought to pump blood even though it might fail to do so: it may fail to fulfill its function. Of course, a heart has no semantic properties: it represents nothing and does not exhibit intentionality in Brentano's sense. Millikan's teleosemantic claim is not that having a function is a sufficient condition for semantic aboutness, but that it is necessary. Something cannot be a representation unless it can misrepresent what it is about. It could not misrepresent anything unless it could malfunction and it could not malfunction unless it had a function. Arguably, nothing can have a function unless it results from some historical selection process. Selection processes are design processes. According to teleosemantic theories, design is the main source of function, which in turn generates intentionality.

Such theories are called teleosemantic in virtue of the intimate connection between design or teleology and content (or intentionality). Selection processes can be intentional or nonintentional. The fundamental goal of the teleosemantic approach is to derive the intentionality of human mental states from the nonintentional process of natural selection that gave rise to human cognition in the course of human phylogeny.

IS INTENTIONALITY CONSTITUTIVE OF MENTALITY?

THE SEVERAL DIMENSIONS OF INTENTIONALITY. Ever since Brentano put forward his thesis that intentionality is constitutive of the mental, philosophers have identified two major dimensions along which intentional

phenomena vary. The first psychological distinction arises from the ontological distinction between objects and states of affairs (such that the former are constituents of the latter). Some mental states are intentional in virtue of being directed toward an object under a particular psychological mode. Thus, perceiving, remembering, loving, or hating someone or something may be instances of such object-directed psychological states. Other (arguably more complex) psychological states are intentional in virtue of the fact that they represent full states of affairs. True beliefs and judgments represent actual states of affairs (or facts). False beliefs, false judgments, desires, intentions, and hopes represent possible or even impossible states of affairs.

One important issue in the philosophy of mind has been whether object-directedness can be reduced to propositional content. Philosophers of perception, in particular, sharply disagree about the intentionality of visual perception: does visual perception always represent full states of affairs, as David Malet Armstrong (1968) and Dennett (1991) argue? Or can visual perception be merely object-directed or nonepistemic, as Dretske (1969) argues? These questions are closely linked to the question whether perceptual experiences have nonconceptual content, as argued by, for example, Gareth Evans (1982) and Christopher Peacocke (1992), and as denied by John McDowell (1994).

The second distinction is internal to the category of psychological states that represent full states of affairs. Russell calls them propositional attitudes, because they are ascribed by complex sentences containing a singular term standing for a person, a main verb expressing his or her attitude (e.g., believing, judging, desiring, and hoping) followed by a that clause expressing the propositional content that is the object of the attitude (as in "John believes that Pegasus flies"). Following Gertrude Elizabeth Margaret Anscombe (1957) and Searle (1983), beliefs and judgments are said to have a "mind-to-world direction of fit" because their function is to record facts. Intentions and desires are said to have a "world-to-mind direction of fit" because their function is not to represent facts but nonactual states of affairs that are goals for actions.

THE VARIETIES OF CONSCIOUS AWARENESS. What do a pain, a visual experience of blue, the taste of a glass of burgundy, the smell of a rose, an intention to act, a thrust of anger, a feeling of depression, and the belief that 2 is a prime number have in common? If they have nothing in common, then presumably the English word *mental* fails to express any well-defined property. If so, then as Richard Rorty (1979) argues, the word *mental* may just be part of an academic language game with no realistic explanatory, let alone scientific, import. According to Brentano's thesis, all these psychological phenomena have something in common: they exhibit intentionality (in the form of either object-directedness or propositional content). Is Brentano's thesis true? Much twentieth-century philosophy of mind has addressed this issue by asking the question: Can one be consciously aware without being aware of something or other?

As many philosophers would recognize, however, the concept expressed by conscious awareness needs some clarification. A first clarification is provided by David M. Rosenthal's (1991) distinction between creature consciousness and state consciousness. A creature can be said to be conscious and a creature's mental state (e.g., one of his or her beliefs) can be said to be conscious. Second, as Rosenthal also points out, if conscious, a creature can either be intransitively conscious or transitively conscious of something. A creature is intransitively conscious if he or she is not unconscious (e.g., as in a coma). He or she can be transitively conscious of things, properties, and states of affairs by either perceiving them or thinking of them. Whereas a creature can be either intransitively conscious or transitively conscious of something, a creature's mental state (or representation) can only be intransitively conscious. Third, one of the things a creature can be transitively conscious of is him- or herself. Arguably, a creature could not be self-conscious or self-aware unless he or she had some concept of him- or herself. But it does not follow that a creature could not be transitively conscious of things in his or her environment unless he or she was self-conscious and unless he or she had a concept of the self. Finally, as Ned Block (1997) argues, there are two distinct ways a mental state can be conscious: It is access conscious if it is poised for free use in reasoning and for the direct rational control of action and speech. It is phenomenally conscious if, as Thomas Nagel (1974) famously puts it, "there is something it is like" to be in that state, that is, if the state has a phenomenal character.

THE DISPUTES BETWEEN INTENTIONALISTS AND NONINTENTIONALISTS. Philosophers who accept Brentano's thesis that all psychological phenomena exhibit intentionality are intentionalists. Philosophers who do not are nonintentionalists. The nonintentionalists argue that the phenomenal character (phenomenal consciousness) of qualitative mental states (*qualia*) such as pains cannot be accounted by its intentionality if it has any. Radical nonintentionalists, such as Searle (1992) and

Galen Strawson (1994), reject Brentano's view that intentionality is the mark of the mental and argue that either accessibility to consciousness or phenomenal consciousness is the true criterion of the mental. If accessibility to consciousness were the criterion of the mental, then psychological states and processes investigated by cognitive science that are unavailable to consciousness would fail to qualify as mental (or psychological). If phenomenal consciousness were the criterion of the mental and if the belief that 5 is a prime number lacks phenomenal consciousness, then this belief would fail to be mental. Many nonintentionalists, such as Block (1997), David J. Chalmers (1996), and Charles P. Siewert (1998) hold a more moderate dual view according to which many, if not all, mental states have both intentional and phenomenal properties.

If a psychological state has some phenomenal character, then it is incumbent on an intentionalist either to show that its phenomenal character is derivable from its intentionality or to argue that its having a phenomenal character is merely an illusion. Radical intentionalists, such as Dennett (1991), who are *qualophobes*, choose the last option and argue that the mysteries of phenomenal consciousness should be explained away or dissolved. Philosophers who subscribe to moderate forms of intentionalism take the former option, which in turn can be divided into two distinct strategies. According to the higher-order thought (HOT) theory of conscious states (defended by Rosenthal [1991]), what makes a person's mental state phenomenally conscious is that the person is transitively conscious of it by virtue of forming a HOT about it. If, however, a creature (e.g., a human baby) lacks the ability to form a HOT about his or her own perceptions, then his or her perceptions will lack phenomenal consciousness—a consequence that many will find implausible. Finally, pure intentionalists, such as Michael Tye (1995) and Tim Crane (2001), argue that *qualia* are mental representations with nonconceptual content. Tye, for example, argues that pains are mental representations of damaged bodily parts and that the phenomenal features of pains arise from the nonconceptual content of the bodily representation. Arguably, the intentionalist account of pain may derive some empirical support from the phenomenon of phantom pains, whereby people who have had limbs amputated may still experience pain in their phantom limb. A bodily part need not exist for someone to feel pain in it: it is enough that the bodily part be mentally represented.

See also Belief; Content, Mental; Propositional Attitudes: Issues in Philosophy of Mind and Psychology; Reference.

Bibliography

Anscombe, G. E. M. *Intention*. Ithaca, NY: Cornel University Press, 1957.

Armstrong, David Malet. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul, 1968.

Block, Ned. "On a Confusion about a Function of Consciousness." In *The Nature of Consciousness: Philosophical Debates*, edited by Ned Block, Owen Flanagan, and Güven Güzeldere, 375–415. Cambridge, MA: MIT Press, 1997.

Brentano, Franz Clemens. *Psychology from an Empirical Standpoint*. Translated by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister; edited by Linda L. McAlister. London: Routledge and Kegan Paul, 1973.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.

Chisholm, Roderick M. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.

Churchland, Paul M. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press, 1989.

Crane Tim. *Elements of Mind: An Introduction to the Philosophy of Mind*. New York: Oxford University Press, 2001.

Dennett, Daniel C. *Consciousness Explained*. Boston: Little, Brown, 1991.

Dennett, Daniel C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Dretske, Fred I. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

Dretske, Fred I. *Seeing and Knowing*. London: Routledge and Kegan Paul, 1969.

Evans, Gareth. *The Varieties of Reference*. New York: Oxford University Press, 1982.

Fodor, Jerry A. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press, 1987.

Grice, Paul. "Meaning" (1957). In *Studies in the Way of Words*, 213–223. Cambridge, MA: Harvard University Press, 1989.

Haugeland, John. "Semantic Engines: An Introduction to Mind Design." In *Mind Design: Philosophy, Psychology, Artificial Intelligence*, edited by John Haugeland, 95–128. Cambridge, MA: MIT Press, 1981.

McDowell, John. *Mind and World*. Cambridge, MA: Harvard University Press, 1994.

Meinong, Alexius. "The Theory of Objects" (1904). In *Realism and the Background of Phenomenology*, edited by Roderick M. Chisholm, 76–117. Glencoe, IL: Free Press, 1960.

Millikan, Ruth Garrett. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press, 1984.

Nagel, Thomas. "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435–450.

Peacocke, Christopher. *A Study of Concepts*. Cambridge, MA: MIT Press, 1992.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.
Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.
Rosenthal, David M. "Two Concepts of Consciousness." In *Nature of Mind*, edited by David M. Rosenthal, 462–477. New York: Oxford University Press, 1991.
Searle, John R. *Intentionality: An Essay in the Philosophy of Mind*. New York: Cambridge University Press, 1983.
Searle, John R. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press, 1992.
Siewert, Charles P. *The Significance of Consciousness*. Princeton, NJ: Princeton University Press, 1998.
Strawson, Galen. *Mental Reality*. Cambridge, MA: MIT Press, 1994.
Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press, 1995.

Pierre Jacob (2005)

INTERNAL RELATIONS

See *Relations, Internal and External*

INTERNALISM AND EXTERNALISM IN ETHICS

Among the various uses of the term *internalism* in ethics, there are two that are central and importantly different. In the following entry, these two uses will be distinguished as judgment internalism and reason internalism.

JUDGMENT INTERNALISM

Judgment internalism is the view that moral judgments can be sufficient to motivate actions. Motivation is internal to morality. Externalists, by contrast, hold that the motivation to act morally is supplied by motives that are only contingently related to moral judgments. Internalism is thus opposed to the view that we need to appeal to special motives in order to explain compliance with moral demands, such as sympathy, as well as to a Hobbesian outlook claiming that the motivation to act is always self-interested, and that the motivation to act morally must therefore be self-interested, too. Internalism in this sense has been defended by Thomas Nagel (1970), John McDowell (1978), Christine Korsgaard (1986), and possibly by Immanuel Kant (1785).

One of the first to introduce the term in this sense was William Frankena (1958) who is critical of internalism. Externalism—the view that moral judgments as such cannot motivate moral actions—has few explicit defenders. However, John Stuart Mill (1861) claimed that we should distinguish sharply between the 'proof' of the moral principle (the principle of utility, as he sees it) and its 'sanctions.' While it can be demonstrated to anyone that an action is morally wrong if it violates the principle of utility, the motivation to act in accordance with the principle will be present only in those who received an appropriate education.

One response to judgment internalism is error theory (Mackie 1977). On the level of semantics, internalists have it right: Moral judgments involve an attempt to refer to properties that exist independently of a person's desires, but that are capable of motivating him or her. Thus, on the one hand, those properties must be features of the world as it is independently of our responses to it. But, on the other hand, we necessarily respond to them in certain ways. This combination of claims is, according to Mackie, ontologically speaking, 'queer:' It requires that moral properties be primary and secondary properties at the same time. But there can be no such properties. Therefore, all our moral judgments are false (for a critical discussion of J.L. Mackie's argument, see McDowell [1985]).

REASON INTERNALISM

Yet there is a different use of the terms *internalism* and *externalism* that in effect reverses the one sketched above. Bernard Williams in his influential essay *Internal and External Reasons* (1980) defends the view that all practical reasons are internal reasons. By *internal* he means that they are related to a person's given desires—to the elements of his or her subjective motivational set. This Hume-inspired view is based on an explanation of motivation in terms of desires as a distinct kind of psychological state.

Practical reasons are potentially both explanatory and justificatory: They determine what a person should do, but also explain his or her actions (if he or she acts for those reasons). But as explanation must appeal to an agent's motives (or desires), reasons have to be suitably linked to those. Desires, in turn, are not (ultimately) the product of reasons. Therefore, in order to be explanatorily relevant, a person's reasons must be based on his or her given desires. A person has a reason to ϕ, if he or she can reach the conclusion to ϕ by a sound deliberative route starting from his or her given desires (Williams 1989).

Desires, Williams explains, need not be conceived narrowly. The term applies to a whole array of states of a

very different kind comprising a person's projects, commitments, and loyalties. Desire is simply a term of art that can be used to refer to all motivationally relevant attitudes. It follows that a person has reason to act in a certain way only if he or she happens to have an appropriate desire: a desire that will be satisfied if he or she acts accordingly, provided the desire is not based on false belief and formed on the basis of correct information about the relevant facts. Therefore a person's reasons do not exist independently of his or her psychological states.

This view is at odds with the normal understanding of moral reasons, and of practical reasons more generally. We tend to interpret at least some reason statements as referring to how things are in the world (independently of the agent's attitude toward them). They are thus external reasons, according to Williams's terminology: reasons that are independent of a person's psychological states. In interpreting reason statements as referring to external reasons, Williams claims, we are mistaken because external reasons are incapable of explaining a person's actions (for an earlier, yet different defense of a similar view, see Davidson [1963]).

Williams's defense of internalism led to an intense and continuing debate (see Hooker [1987], Smith [1995], Millgram [1996], FitzPatrick [2004]). McDowell (1995) replied that Williams may well be right thinking that if reasons can be external, then not everyone is capable of being motivated by practical reasons that apply to him or her. But the externalist is not committed to thinking that they can. The externalist's crucial claim is that reasons exist independently of motives—not that they can motivate anyone independently of what his or her motives happen to be. Is the externalist committed to denying Williams's claim that practical reasons are both justificatory and explanatory then? According to McDowell, he or she is not. Those who are motivated by reasons may not be so motivated by a desire whose existence is independent of the reason.

McDowell suggests an Aristotelian alternative to Williams's Humean view: The capacity of being motivated in the right way is a matter of moral upbringing. But moral upbringing is (in part) the ability to be motivated by moral reasons. The moral person is one who responds to his or her perception of the morally salient features of the his or her situation. Thus McDowell can agree with Williams that practical reasons are both justificatory and explanatory, but denies that explanation must appeal to desires that exist independently of reasons. Reasons exist independently of desires, and they can motivate independently of them—at least those who have been brought up in the right way.

Various versions of reason externalism have been proposed in recent years (see Dancy [2000], Parfit [1997], Raz [1999], Scanlon [1998]). Korsgaard (1986 and 1996) defends a version that is stronger than the one proposed by McDowell: She claims that a reason can motivate a person insofar as he or she is rational (independently of given motives). According to her, being rational is the ability to respond to reason, and we all have that ability (perhaps to a lesser or higher degree). Thus there is no emphasis on moral upbringing in Korsgaard's account of motivation.

Against this, Michael Smith (1987) provides an a priori argument for internalism, or—as he puts it—the Humean theory of motivation. Smith develops Hume's view that beliefs and desires are distinct psychological states, distinguishing them by their different direction of fit. Beliefs aim to represent the world as it is, whereas desires are an agent's dispositions to change the world in such a way that it fits with the desire. Beliefs have a mind-to-world and desires a world-to-mind direction of fit. Only states with the right direction of fit (i.e., desires) can motivate. Beliefs as such cannot. If we are to understand value judgments as beliefs (as Smith thinks we should) they will not be sufficient to explain actions (Smith 1994). This argument gave rise to an ongoing discussion (Wallace 1990; Velleman 1992).

The two uses of internalism can be seen as related: The older tradition of judgment internalism identifies internalism with the claim that moral judgments as such are capable of explaining actions. This claim, however, bears some similarity to Williams's claim that practical reasons are both justificatory and motivating. The main difference is that judgment internalism is confined to moral judgments, whereas Williams is concerned with practical reasons more generally (a further difference is that reason internalists are not committed to accepting that practical reasons are, at least in part, judgments; for the significance of this difference see Dancy [2000]).

Yet, according to Williams, put together with some version of the Humean theory of motivation, the claim that practical reasons are both explanatory and justificatory leads to the conclusion that reasons must be based on desires, which is the view that he calls internalism: reason internalism. Thus, roughly, judgment internalism labels one of the premises of Williams's argument internalism, whereas Williams himself uses the term to refer to its conclusion. The focus of disagreement is then on the

Humean theory of motivation, which divides the two approaches.

See also Error Theory of Ethics; Hume, David; Kant, Immanuel; Mackie, John Leslie; McDowell, John; Metaethics; Mill, John Stuart; Moral Motivation; Nagel, Thomas; Normativity; Response-Dependence Theories; Williams, Bernard.

Bibliography

Dancy, J. *Practical Reality*. Oxford: Oxford University Press, 2000.

Davidson, D. "Actions, Reasons, and Causes" (1963). In *Actions and Events*. Oxford: Oxford University Press, 1980.

FitzPatrick, W. J. "Reasons, Value, and Particular Agents: Normative Relevance without Motivational Internalism" *Mind* 113 (2004): 285–318.

Frankena, W. K."Obligation and Motivation in Recent Moral Philosophy" (1958). In *Perspectives on Morality: Essays by William K. Frankena*, edited by K.E. Goodpaster. Notre Dame, IN; London: University of Notre Dame Press, 1976.

Gibbard, A. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Oxford University Press, 1990.

Heuer, U. "Reasons for Action and Desires." *Philosophical Studies* 121 (2004): 43–63.

Hobbes, T. *The Leviathan* (1651), edited by J.C.A. Gaskin. Oxford: Oxford University Press, 1998.

Hooker, B. "Williams's Argument Against External Reasons." *Analysis* 47 (1987): 42–44.

Hume, D. *A Treatise of Human Nature* (1739–1740). 2nd ed., edited by L. A. Selby-Bigge. Oxford: Clarendon Press, 1978.

Kant, I. *Grundlegung zur metaphysik der sitten, akademie-ausgabe*. Riga: J. F. Hartknock, 1785.

Korsgaard, C. "Skepticism about Practical Reason." *Journal of Philosophy* 83 (1) (1986): 5–25.

Korsgaard, C. *The Sources of Normativity*. Cambridge, U.K.: Cambridge University Press, 1996.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin, 1977.

McDowell, J. "Are the Requirements of Morality Hypothetical Imperatives?" *Proceedings of the Aristotelian Society* 52 (1978): 13–29.

McDowell, J. "Virtue and Reason." *The Monist* 62 (1979): 331–350.

McDowell, J. "Values and Secondary Qualities." In *Morality and Objectivity*, edited by T. Honderich. London: Routledge and Kegan Paul, 1985.

McDowell, J. "Might There Be External Reasons?" In *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, edited by J.E.J. Altham and R. Harrison. Cambridge, U.K.: Cambridge University Press, 1995.

Mill, J. S. *Utilitarianism* (1861), edited by G. Sher. Indianapolis, IN: Hacking Publishing Company, 1979.

Millgram, E. "Williams's Argument Against External Reasons." *Nous* 30 (2) (1996): 197–220.

Nagel, T. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press, 1970.

Parfit, D. "Reasons and Motivation." *Proceedings of the Aristotelian Society* 71 (1997): 99–130.

Raz, J. *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press, 1999.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Smith, M. "The Humean Theory of Motivation." *Mind* 96 (1987): 36–61.

Smith, M. *The Moral Problem*. Oxford: Blackwell, 1994.

Smith, M. "Internal Reasons." *Philosophy and Phenomenological Research* 55 (1) (1995): 109–131.

Velleman, D. "The Guise of the Good." *Nous* 26 (1992): 3–26.

Wallace, R. J. "How to Argue about Practical Reason." *Mind* 99 (1990): 355–385.

Williams, B."Internal and External Reasons" (1980). In *Moral Luck: Philosophical Papers, 1973–1980*. Cambridge, U.K.: Cambridge University Press 1981.

Williams, B. "Internal Reasons and the Obscurity of Blame" (1989). In *Making Sense of Humanity and Other Philosophical Papers, 1982–1993*. Cambridge, U.K.: Cambridge University Press, 1995.

Williams, B. "Replies." In *World, Mind and Ethics. Essays on the Ethical Philosophy of Bernard Williams*, edited by J. E. J Altham and R. Harrison. Cambridge, U.K.: Cambridge University Press, 1995.

Ulrike Heuer (2005)

INTERNALISM VERSUS EXTERNALISM

Internalism in epistemology is a thesis about the nature of epistemic normativity, or the sort of normativity that is involved in the evaluation of cognition. Specifically, internalists claim that the (epistemically) normative status of a belief is entirely determined by factors that are relevantly "internal" to the believer's perspective on things. By contrast, externalists in epistemology deny this. The externalist says that the epistemic status of a belief is not entirely determined by factors that are internal to the believer's perspective.

When internalism and externalism are characterized in this way, several things become apparent. First, internalism is a rather strong thesis, in the sense that it says that epistemic status is entirely a function of internal factors. By contrast, the denial of internalism is a relatively weak thesis. Externalism in epistemology holds that some factors that are relevant to epistemic status are not internal to the believer's perspective. A second point to note is that there are several kinds of epistemically normative status, corresponding to several kinds of epistemic evaluation. We can say that a belief is justified, rational, reasonable, or intellectually responsible, and these need not

mean the same thing. It is possible, then, to be an internalist about some kinds of epistemic status and an externalist about others. Hence, there are a variety of internalisms and a corresponding variety of externalisms.

Third, we get different understandings of internalism (and externalism) depending on different ways that we may understand the phrase "internal to the believer's perspective." The most common way to understand the phrase is that something is internal to a believer's perspective just in case the person has some sort of privileged access to the thing in question. For example, some fact F is relevantly internal to some person S's perspective if S can know by reflection alone whether F obtains. A related, though not equivalent, understanding of "internal to S's perspective" is as follows: Some factor F is internal to S's perspective just in case F constitutes part of S's mental life. For example, a person's perceptual experience counts as internal on this understanding, since how things appear perceptually to S is part of S's mental life in the relevant sense. Also, any belief or representation that S has about how things are would be internal on this understanding, since one's beliefs and other representations are also part of one's mental life. These two understandings are related because it is plausible to think that one has privileged access to what goes on in one's mental life, and perhaps only to what goes on in one's mental life. In that case the two understandings would amount to the same thing for practical purposes. Internalism would then be the thesis that epistemic status (of some specified sort) is entirely a function of factors that are part of one's mental life, and to which one therefore has privileged access.

Finally, it is apparent that some varieties of internalism are initially more plausible than others. That is, some sorts of epistemic evaluation are obviously externalist on the previous understandings. Most importantly, and perhaps most obviously, whether a belief counts as knowledge is an external matter, if only because a belief counts as knowledge only if it is true, and whether a belief is true is typically an external matter.

OBJECTIVE VERSUS SUBJECTIVE EVALUATIONS

There is another reason knowledge and many other sorts of epistemic evaluation must be understood as externalist, however. Consider that we can evaluate both persons and their beliefs in two different ways. Broadly speaking, we can evaluate them either from an objective point of view or from a subjective point of view. From the objective point of view we can ask whether there is a good fit between the person's cognitive powers and the world. For example, we can ask whether the person has a good memory or an accurate vision. Also from this point of view we can ask whether a person's methods of investigation are reliable, in the sense that they are likely to produce accurate results. By contrast, there is a second broad category of epistemic evaluation. This sort does not concern whether a belief is objectively well formed, but whether it is subjectively well formed. It asks not about objective fitness, but about subjective appropriateness. Internalism is pretty much a nonstarter with respect to evaluations of the first category. Evaluations from an objective point of view involve factors such as accuracy, reliability, and appropriate causal relations to one's environment, and these are paradigmatically external factors. Therefore, internalism is best understood as a thesis about the second broad category of epistemic evaluation: It is a thesis about what factors determine subjective appropriateness. Let us use the term *epistemic justification* to signify this second sort of epistemic status. In that case internalism is the thesis that epistemic justification is entirely a function of factors that are within the believer's perspective.

THREE CONSIDERATIONS IN FAVOR OF INTERNALISM

Why would someone be an internalist? Three considerations have been stressed in the literature. The first begins with an assumption about the nature of epistemic justification (where *epistemic justification* refers to the sort of subjective appropriateness required for knowledge or some other important epistemic status). The assumption is this: A belief is epistemically justified just in case it is epistemically responsible. However, the argument continues, epistemic responsibility is entirely a matter of factors that are internal to S's perspective. Therefore, epistemic justification is entirely a matter of factors that are internal to S's perspective.

A second consideration put forward in favor of internalism invokes a strong intuition about epistemic justification. Namely, in many cases it seems that believers who are alike in terms of internal perspective must also be alike in terms of epistemic justification. The point is often illustrated by considering René Descartes's victim of an evil deceiver. Suppose that the victim is exactly like you in terms of internal perspective. Even if the victim lacks knowledge, the argument goes, surely his beliefs are as well justified as yours are. If you are justified in believing that there is a table before you, and if the victim's perspective is exactly as yours, then he must be justified in believing that there is a table before him.

A third consideration invoked in favor of internalism is that externalism makes an answer to skepticism too easy. Philosophical problems are supposed to be difficult. If the externalist has an easy answer to the problem of skepticism, this argument goes, then that is good reason to think that externalism is false. At the least it is good reason to think that the externalist has changed the subject—that he is no longer talking about our traditional notions of justification and knowledge.

How does externalism make an answer to skepticism too easy? The idea is roughly as follows: According to the skeptic one can know via sense perception only if one knows that sense perception is reliable. Similarly, one can know by inductive reasoning only if one knows that inductive reasoning is reliable. This creates problems for the internalist, because it is hard to understand how one can mount a noncircular argument to the desired conclusions about the reliability of one's cognitive powers. There is, however, no such problem for the externalist since the externalist can deny the initial assumption of the skeptical argument. For example, an externalist can insist that sense perception gives rise to knowledge so long as sense perception is in fact reliable. There need be no requirement, on an externalist account, that one know that one's perception is reliable. What is more, on an externalist account one seemingly can know that one's cognitive powers are reliable, and easily so. For example, one can use reliable perception to check up on perception, and then reason from there that perception is reliable. Similarly, one can use reliable induction to check up on induction, and then reason from there that induction is reliable.

EXTERNALIST REPLIES

Externalists reply that none of these considerations adequately motivate internalism. First, externalists argue, even if epistemic justification is to be understood in terms of epistemic responsibility, it is false that epistemic responsibility is entirely a matter of factors that are internal to S's perspective. This is because whether a belief is epistemically responsible is partly a function of the belief's etiology, or how S came to have the belief in the first place. For example, whether a person is epistemically responsible in holding some belief is partly a function of the person's prior behavior: If S's reasons for believing b are the result of prior negligence, then S is not now blameless in believing b. Similarly, we can make a distinction between (1) merely having good reasons for a belief, and (2) believing on the basis of those reasons. Plausibly, a belief is epistemically praiseworthy only if it is believed on the basis of good reasons—merely having good reasons, if one does not use them, is not enough. But etiological considerations such as these involve external factors; that is, factors that are not typically internal to S's perspective.

The same line of argument has been used to counter the second consideration in favor of internalism. The problem is that two believers might be alike internally, and yet different regarding the causal genesis of their beliefs. Suppose that two persons arrive at the same internal perspective, but that one does so in a way that is epistemically responsible, whereas the other does so in a way that is careless and thick-headed. The two persons will not be alike in epistemic justification, although they share the same internal perspective.

Finally, externalists argue that the third consideration in favor of internalism is self-defeating. In effect, internalists claim that only they can give a satisfactory answer to traditional skeptical concerns. On the contrary, externalists argue, internalism makes it impossible to answer the skeptic. This is because traditional skeptical arguments assume internalism about epistemic justification. Moreover, if one concedes that assumption, externalists argue, then the skeptic has all he or she needs to construct skeptical arguments that are otherwise sound. Therefore, externalists conclude, internalism about epistemic justification guarantees skepticism about epistemic justification.

See also Epistemology.

Bibliography

Alston, William. "Concepts of Epistemic Justification." *Monist* 68 (1985): 57–89.

Alston, William. "Internalism and Externalism in Epistemology." *Philosophical Topics* 14 (1986): 179–221.

BonJour, Laurence. "Externalist Theories of Empirical Knowledge." *Midwest Studies in Philosophy* 5 (1980): 53–73.

Conee, Earl, and Richard Feldman. "Internalism Defended." In *Epistemology: Internalism and Externalism*, edited by Hilary Kornblith. Malden, MA: Blackwell Publishers, 2001.

Fumerton, Richard. *Metaepistemology and Skepticism*. Lanham, MD: Rowman & Littlefield, 1995.

Goldman, Alvin. "The Internalist Conception of Justification." *Midwest Studies in Philosophy* 5 (1980): 27–51.

Greco, John. "Agent Reliabilism." In *Philosophical Perspectives, 13, Epistemology*, edited by James E. Tomberlin. Atascadero, CA: Ridgeview Press, 1999.

Greco, John. "Internalism and Epistemically Responsible Belief." *Synthese* 85 (1990): 245–277.

Greco, John. "Justification Is Not Internal." In *Contemporary Debates in Epistemology*, edited by Matthias Steup and Ernest Sosa. Oxford, U.K.: Blackwell Publishers, 2005.

Kornblith, Hilary, ed. *Epistemology: Internalism and Externalism*. Malden, MA: Blackwell Publishers, 2001.

Plantinga, Alvin. *Warrant: The Current Debate*. New York: Oxford University Press, 1993.

Sosa, Ernest. "Beyond Internal Foundations to External Virtues." In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, edited by Laurence BonJour and Ernest Sosa. Malden, MA: Blackwell Publishers, 2003.

John Greco (1996, 2005)

INTRINSIC VALUE

A thing's intrinsic value is the value it has in itself as opposed to the instrumental value it derives from causally producing something else. Such value is important for the theory of the right, since on most views at least one moral duty is to promote intrinsic goods and prevent intrinsic evils. But it also matters in itself. If an earthquake causes suffering no one could have prevented, the suffering is still intrinsically bad. In fact, it is distinctive of the concept of value that, unlike that of ought or right, it is not restricted to what is under our voluntary control. And intrinsic value is the central kind of value. Something is instrumentally good if it produces something else good, but on pain of infinite regress, what it produces must eventually be intrinsically good.

There are two types of questions about intrinsic value: conceptual questions about what it is and substantive questions about what things have it. One conceptual issue is whether judgments of intrinsic value are objectively true or false or merely express emotions, but this question arises equally for all normative concepts. A more distinctive question is how intrinsic goodness relates to other normative concepts. One view, defended by G. E. Moore (1903), says the concept of goodness is simple and unanalyzable, even in normative terms; a rival view analyzes the good as what one ought to desire (Sidgwick 1907) or as the desire for which is correct or appropriate (Brentano 1969, Ewing 1948). Like the first view, the second allows judgments of goodness to extend beyond ones about right action, since one can desire what one cannot produce. But it cannot say, as the first can, that one ought to desire something *because* it is good, and it faces the further difficulty that, since one often cannot produce desires at will, its *ought* cannot be the simple deontic one and is therefore not clearly distinguishable from *good*.

A second conceptual question is what type of entities have intrinsic value. The most common view is that they are states of affairs, so to say that pleasure is good is really to say that the state of affairs in which someone enjoys pleasure is good. But some attach value primarily to objects. Thus, Immanuel Kant (1998) held that the prime bearers of intrinsic value are persons (and not the states of affairs in which persons exist) while others find such value in, say, beautiful works of art. A related question is what the morally required response to intrinsic value is. Many who locate such value in states of affairs say this response is only to desire and promote the good; others say there is a separate and stronger requirement to respect good states by not directly destroying them. In fact the Kantian view centers on respect. Since one cannot promote values that already exist, as the Kantian value of persons does, the required response is not to treat them in a way that disregards their value.

A final conceptual question is what kind of properties intrinsic value can depend on. Moore held the strict view that a thing's intrinsic value can depend only on its intrinsic properties, those it has independently of relations to other things. He therefore tested for intrinsic value by a *method of isolation*, which involves imagining a world where only a given thing exists and asking whether that world is good. Others hold, more liberally, that intrinsic value can depend on relations, so a belief can be intrinsically better if it corresponds to reality and therefore is true, or a pleasure can be better if it is that of a virtuous person and so deserved. (Some restrict the term *intrinsic value* to value that depends only on intrinsic properties and use *final value* for what can vary with relations; they still differ from Moore in holding that some value worth promoting for itself depends on relations.)

The substantive issues about intrinsic value arise most clearly if it is located in states of affairs; the question then is which states are worth desiring and promoting for their own sakes. Here, the simplest view is hedonism, which holds that only pleasure is intrinsically good and only pain intrinsically evil. Though defended by Jeremy Bentham, Sidgwick, and other utilitarians, hedonism faces numerous objections. It implies that a world of intense mindless pleasures like those of *Brave New World*, where people are systematically deluded and exercise no serious skills, can be supremely good; it makes morally vicious pleasures, such as pleasure in another's pain, purely good; and it also makes undeserved pleasure good.

In response, many philosophers develop pluralistic views according to which the intrinsic goods include not just pleasure but also, for example, other states of persons such as knowledge, the achievement of difficult goals, and moral virtue; patterns of distribution of goods across persons, such as equal distributions or ones proportioned

to people's deserts; and even states outside persons, such as the flourishing of complex ecosystems. These pluralistic views require comparisons between their different goods whereby some may be higher or greater than others. Thus, scientific knowledge may be better than the pleasure of eating chocolate and the virtue of compassion better than both. But in so far as there is a duty to promote intrinsic value, one should pursue the best overall outcome, weighing all goods appropriately against each other.

Many nonhedonic goods involve what Moore called the "principle of organic unities (1903, p. 27-36) according to which the intrinsic value of a whole need not equal the sum of the values its parts would have on their own, so states with little value when apart can make for significant value when combined. The exact formulation of this principle depends on whether intrinsic value can depend only on intrinsic or also on relational properties. Either way, the greatest value can be found not in simple states such as pleasures but in complexes combining several elements in a specified way.

A final substantive question is whether all intrinsic values are agent-neutral, so everyone has equal reason to pursue them, or some are agent-relative, with greater value from some people's point of view than from others'. Given Moore's view that goodness is an unanalyzable property, agent-relativity is impossible: A state either has the simple property, in which case everyone has a duty to promote it, or it does not. But analyses of good in terms of ought allow relativity: We can say that each person's pleasure is something only that person ought to desire, or that parents should care more about their children's pleasure than about that of strangers. On the agent-neutral view, intrinsic value specifies a common goal that everyone is to pursue together; given agent-relativity, there can be different goods and different required goals for different people.

See also Bentham, Jeremy; Intuitionism and Intuitionistic Logic; Kant, Immanuel; Objectivity in Ethics; Sidgwick, Henry; Value and Valuation.

Bibliography

Anderson, Elizabeth. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press, 1993.

Brentano, Franz. *The Origin of Our Knowledge of Right and Wrong*. Edited by Roderick M. Chisholm. Translated by Roderick M. Chisholm and Elizabeth Schneewind. London: Routledge & Kegan Paul, 1969.

Ewing, A. C. *The Definition of Good*. New York: Macmillan, 1948.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Edited and translated by Mary Gregor. Cambridge, U.K.: Cambridge University Press, 1998.

Korsgaard, Christine M. "Two Distinctions in Goodness." *Philosophical Review* 92 1983: 169–195.

Lemos, Noah M. *Intrinsic Value: Concept and Warrant*. Cambridge, U.K.: Cambridge University Press, 1994.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Ross, W. D *The Right and the Good*. Oxford: Clarendon Press, 1930.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1907.

Thomas Hurka (2005)

INTROSPECTION

The term *introspection* might be defined as the direct, conscious examination or observation by a subject of his or her own mental processes. The term is derived from two Latin words, *spicere* ("to look") and *intra* ("within").

From at least the time of René Descartes up to the early twentieth century, it would have been considered unproblematic that the mind can reflect (or bend its attention back) upon itself. In the eighteenth and nineteenth centuries, if not earlier, self-reflection began to be interpreted, in the main, as introspection. In turn, to introspect one's own mental processes was explained in terms of the capacities (1) to focus the full glare of one's conscious attention upon the task of observing some particular, first-level, conscious process (or mental act), which was an item in one's stream of consciousness, and (2) to report in a privileged and incorrigible way upon the results of such observation. This introspective act was considered to be a form of inner, though nonsensuous, perception, and deliberate parallels were frequently drawn between it and ordinary outer perception by means of our senses, such as those of vision or hearing.

In the nineteenth century, Franz Brentano and other philosophical psychologists were at pains to distinguish introspection (sometimes called inner observation) from its close relative, self-consciousness (sometimes called inner perception). Introspection was a deliberate act of focusing a subject's attention on some inhabitant in his stream of consciousness. Self-consciousness was an indeliberate but inescapable, though partial, concomitant awareness on the part of a subject of at least some features of some of his first-level conscious mental acts. To put it metaphorically, introspection was a deliberate ogling with the inner mental eye; self-consciousness was

unavoidably catching sight of something out of the corner of one's mental eye.

However, even as this canonical version of introspection was being formulated, doubts were being voiced about the possibility of splitting consciousness into two processes that operated at two different levels at the same time. Pushing aside these doubts, the early psychological introspectionists—such as Wilhelm Wundt, Edward B. Titchener, Narziss Ach, Karl Bühler, and William James—believed that either introspection proper or some version of self-consciousness was nevertheless the only possible method for inaugurating a truly empirical, that is, scientific, psychology. For only the subject of mental acts or processes can have "eye witness," knowledge by acquaintance of the denizens of his or her stream of consciousness. So, the very first psychological laboratories were devoted to introspection (for this term came to be used for both introspection proper and for scientific versions of self-consciousness). In carefully designed laboratories bristling with chronograph and tachistocope, subjects were asked to produce detailed introspective reports on various aspects of the inner conscious effects of carefully controlled stimuli applied to their senses.

These experiments resulted in some of the most tedious literature that psychology has ever produced. Also, there could be found little or no agreement about results across schools or from one laboratory to the next. Yet another consequence, which Wundt, for example, readily admitted, was that introspection experiments seemed confined to a study of comparatively trivial mental episodes.

Surprisingly, the failure of introspectionism did not lead many people to question the inherent model of introspection. As psychology and philosophy wound their way through behaviorism and versions of the mind-brain identity theory to contemporary forms of physicalism, such as functionalism, both were faithful to the original, classical model of introspection. They abandoned the Cartesianism of the psychological introspectionists and questioned the privileged status of introspection reports, but they did not question the basic two-level picture—that introspection was a second-level monitoring, observing, registering, or tracking of some first-level process or processes.

Thus, classical psychological behaviorists such as John Broadus Watson or B. F. Skinner gave, as at least one account of one employment of introspection, that it was a literal monitoring by the subject of his thinking (which for a classical behaviorist was to be analyzed as inner truncated movements in the muscles of speech, or "stopped short" speech). Only the repeated failure of experiments seeking to verify this theory led to the abandonment of that particular, and now notorious, explanation.

The philosophers, or most of them, also championed some version of the two-level account of introspection, and still do. Even the most tough-minded of the physicalists, such as David M. Armstrong or Daniel Dennett, stick resolutely to a two-level monitoring account of introspection. Thus, in *A Materialist Theory of the Mind* Armstrong describes introspection as one part of the brain scanning another part of the brain such that the subject, whose brain it is, generates (in entirely causal fashion) a belief about the nature of the first-level, scanned, brain process. In *Content and Consciousness* and again in *Brainstorms* and *Consciousness Explained*, in an uncompromising functionalist account of mind, Dennett describes introspection in terms of one part of the brain "accessing" another (like one part of a computer accessing another) and then, via the speech center, "printing out" the results.

In philosophy and psychology since the 1950s, there has been a minority view that this two-level account of introspection is simply mistaken. Humans have no such second-level inspecting or scanning or monitoring capacity. Earlier, Gilbert Ryle (1949) argued convincingly that this two-level account did not make theoretical sense. Unfortunately, he substituted for it an unconvincing behaviorist account (in terms of the ordinary perceptual "retrospection" of ordinary behavior). More recently, psychologists and philosophers (such as Wilson and Nisbett 1977 and Lyons 1986) have suggested that, besides those theoretical grounds for rejecting the two-level account of introspection, there are also empirical grounds for rejection drawn from contemporary experimental psychology, anthropology, and the brain sciences. In contemporary introspective experiments subjects produced reports that were more like stereotyped and predictable "folk" interpretations than detailed eyewitness accounts of inner events. Besides, it seems that in cultures more or less uninfluenced by European culture people do not claim to have powers of introspection. More important, there does not seem to be any part of the brain that functions as a monitor of those neurophysiological states that maintain and control conscious states. Finally, it seems both possible and more plausible to give an account of what humans are doing, when they claim to be introspecting, in terms of the exercise of the internal but quite ordinary capacities of memory and imagination. This opposition of views has not yet been resolved, and,

because of this, introspection (like consciousness itself) is likely to receive more direct and sustained treatment in the future.

See also Armstrong, David M.; Behaviorism; Brentano, Franz; Descartes, René; Functionalism; James, William; Philosophy of Mind; Physicalism; Ryle, Gilbert; Skinner, B. F.; Wundt, Wilhelm.

Bibliography

IN PHILOSOPHY

Armstrong, D. M. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul, 1968.

Brentano, F. *Psychology from an Empirical Standpoint* (1874), edited by O. Kraus and L. McAlister. Translated by A. C. Rancurello, D. B. Terrell, and L. McAlister. New York: Humanities Press, 1973.

Churchland, P. M. *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*. Cambridge, MA: MIT Press, 1984.

Dennett, D. *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, VT: Bradford Books, 1978.

Dennett, D. *Consciousness Explained*. Boston: Little Brown, 1991.

Dennett, D. *Content and Consciousness*. London: Routledge and Kegan Paul, 1969.

Dretske, F. "Introspection." *Proceedings of the Aristotelian Society* 94 (1994): 263–278.

Hamlyn, D. W. "Self-Knowledge." In *Perception, Learning, and the Self: Essays on the Philosophy of Psychology*. London: Routledge and Kegan Paul, 1983.

Lyons, W. *The Disappearance of Introspection*. Cambridge, MA: MIT Press, 1986.

Ryle, G. *The Concept of Mind*. London: Hutchinson, 1949.

IN PSYCHOLOGY

Boring, E. G. "A History of Introspection." *Psychological Bulletin* 50 (1953): 169–189.

Danziger, K. "The History of Introspection Reconsidered." *Journal of the History of the Behavioral Sciences* 16 (1980): 241–262.

Hebb, D. O. "The Mind's Eye." *Psychology Today* 2 (1969): 55–68.

Humphrey, N. *Consciousness Regained: Chapters in the Development of Mind*. New York: Oxford University Press, 1983.

James, W. *The Principles of Psychology* (1890). 2 vols. New York: Dover Publications, 1950.

Miller, G. *Psychology: The Science of Mental Life*. Harmondsworth, U.K.: Penguin, 1966.

Nisbett, R. E., and T. D. Wilson. "Telling More than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84 (1977): 231–259.

Titchener, E. B. *A Primer of Psychology*. New York: Macmillan, 1898.

Wundt, W. *An Introduction to Psychology*. Translated by R. Pintner. London, 1912.

OTHER RECOMMENDED TITLES

Cassam, Quassim, ed. *Self-Knowledge*. Oxford: Oxford University Press, 1994.

Hill, Christopher. *Sensations: A Defense of Type Materialism*. Cambridge, U.K.: Cambridge University Press, 1991.

Kornblith, Hilary. "Introspection and Misdirection." *Australasian Journal of Philosophy* 67 (4) (1989): 410–422.

Lycan, William. "Consciousness as Internal Monitoring." In *Consciousness and Experience*. Cambridge, MA: MIT Press, 1986.

Rosenthal, D. M. "Thinking that One Thinks." In *Consciousness: Psychological and Philosophical Essays*, edited by Martin Davies. Oxford: Blackwell, 1993.

Shoemaker, S. "Self-Knowledge and 'Inner Sense.'" *Philosophy and Phenomenological Research* 54 (2) (1994): 249–315.

William Lyons (1996)
Bibliography updated by Benjamin Fiedor (2005)

INTUITION

The broadest definition of the term *intuition* is "immediate apprehension." *Apprehension* is used to cover such disparate states as sensation, knowledge, and mystical rapport. *Immediate* has as many senses as there are kinds of mediation: It may be used to signify the absence of inference, the absence of causes, the absence of the ability to define a term, the absence of justification, the absence of symbols, or the absence of thought. Given this range of uses, nothing can be said about intuition in general. Instead, it is necessary to pick out those principal meanings of the term that have played the most important roles in philosophical controversy and to discuss each of these individually.

Four principal meanings of *intuition* may be distinguished: (1) Intuition as unjustified true belief not preceded by inference; in this (the commonest) sense "an intuition" means "a hunch." The existence of hunches is uncontroversial and not of philosophical interest. (2) Intuition as immediate knowledge of the truth of a proposition, where *immediate* means "not preceded by inference." This is a philosophically important sense, since philosophers have found it puzzling that one can have knowledge, and thus justified belief, without having made oneself aware through the process of inference of any justification for this belief. (3) Intuition as immediate knowledge of a concept. "Immediate knowledge" here means, roughly, "knowledge that does not entail ability to define the concept." (4) Intuition as nonpropositional knowledge of an entity—knowledge that may be a necessary condition for, but is not identical with, intuitive knowledge of the truth of propositions about the entity. This sense of *intuition* is exemplified by (*a*) sense percep-

tions, considered as products of a cognitive faculty distinct from the faculty of forming judgments concerning the entity sensed; (*b*) intuitions of universals, or (as in Immanuel Kant) of such insensible particulars as time and space—intuitions that are necessary conditions of our intuitive knowledge of a priori truths; (*c*) mystical or inexpressible intuitions that, unlike sense perceptions and intuitions of universals, do not make possible knowledge of the truth of propositions about the entities intuited—such intuitions as Henri Bergson's inexpressible intuition of duration, Johann Gottlieb Fichte's intuition of the Transcendental Ego, and the mystic's intuition of God.

FACULTY AND LINGUISTIC EXPLANATIONS OF INTUITIVE KNOWLEDGE

INTUITIVE AND NONINFERENTIAL KNOWLEDGE. There is both a strong and a weak sense of "intuitive knowledge that *p*." In the weak sense of this term, *S* knows that *p* intuitively if (*a*) *p* is true, (*b*) he is justified in believing that *p*, and (*c*) his knowledge that *p* is not based upon his inferring *p* from other propositions. The criterion for its not being so based is simply that *S* would deny, for any set of propositions *p** from which *p* follows, that he believes that *p because* he believes that *p** (although he might in fact believe *p** and be willing to adduce *p** to satisfy someone else's doubts about *p*). In this sense of "intuitive," we may know intuitively that we have two legs or two children, but we cannot know intuitively that the Civil War was caused by slavery, or that nothing can move faster than the speed of light. In this sense, the existence of intuitive knowledge is unquestionable; and "intuitive" in this sense is synonymous with "noninferential."

In the stronger sense of *intuition*, however, only a certain species of noninferential knowledge is intuitive: *S* knows that *p* intuitively only if (*a*) *p* is true, (*b*) he is justified in believing that *p*, and (*c*) there are no accepted procedures for resolving doubts about the truth of *p*, given *S*'s belief that *p*. Thus we may be justified in believing without inference that we have two legs, but if we have doubts we can undertake such tests as looking and seeing, asking others, and checking the possibility of collective hallucination. Given these tests, so much evidence may appear to show that one leg is missing that it would be irrational to maintain our previous belief. But in certain cases—for example, our belief that we are in pain, or that every event has a cause—there are (at present) no procedures available for resolving doubt. It is never irrational to continue to believe that *S* has a pain once one knows that he believes he does, despite, for instance, the failure of physiologists to find a concomitant neural process. Again, if someone thinks that some events are uncaused, we have no way of testing his hypothesis. Yet we are not willing to give up our claim to know that he is wrong. In both sets of cases—so-called rock-bottom data of perception and introspection, and so-called unquestionable first principles—justified belief is accompanied by the lack of procedures to settle doubt. These are the two paradigm cases of "intuitive knowledge," in the strong sense of the term—first-person statements about those psychological states to which one has "privileged access" and underived a priori truths.

In this stronger sense, too, the existence of intuitive knowledge is unquestionable. Two points should, however, be noted. First, if in formulating the conditions for the application of this sense of intuitive knowledge we had simply said "*p* is indubitable" rather than "there are no accepted procedures for resolving doubts about *p*," then it would have been questionable whether any such knowledge existed. It can plausibly be argued that, under sufficiently peculiar circumstances, it may be rational to doubt one's belief that one is in pain, or that every event has a cause. In general, it can plausibly be argued that there are no intrinsically indubitable propositions, for rational doubt may outstrip the possibility of rationally settling doubt. Second, it is possible for procedures to come into existence for settling doubt in areas where none existed before. Thus we now take *S*'s belief that he was in pain as the best possible evidence for his having been in pain, but advances in physiology may bring about a practice of withdrawing claims to have been in pain when the relevant neural processes have failed to occur. Under these conditions, *S*'s belief that he was in pain would be intuitive in the weak sense, but no longer in the strong sense. Again, some philosophers would argue that, with the rise of quantum theory, we are now in a position to treat "every event has a cause" as an empirical hypothesis, even though it was once the paradigm of an unquestionable first principle. In general, whether a proposition can count as the object of intuitive knowledge (in the strong sense) is a function of the availability of accepted procedures for settling doubt, and it is doubtful that we can know a priori in what areas such procedures will and will not be developed.

Noninferential will here be used in place of the weak sense of intuitive, and *intuitive* in place of the strong sense. Both noninferential and intuitive knowledge seemed to philosophers to require explanation because the paradigm of knowledge has, since Aristotle, fre-

quently been taken to be inferential knowledge—the case in which one knows not only that *p* is true but also why *p* is true, and believes that *p* is true because one believes certain other propositions from which *p* may validly be inferred (see Aristotle's *Posterior Analytics* I, 2). Noninferential knowledge has often been explained by being assimilated to this paradigm through the use of the notion of implicit or unconscious inference. Cases of nonintuitive knowledge have been treated as cases in which an inference from intuitively known premises was performed, and cases of nonintuitive, noninferential knowledge as cases in which the knower is not aware of having performed the appropriate inference.

FACULTY THEORY. Various explanations have been given of the existence of intuitive knowledge. As was noted, the objects of intuitive knowledge seem to fall into two quite different groups—such very particular facts as "This looks white" or "This hurts," and such very general facts as "Every event has a cause" or "If *p* implies *q*, and *p*, then *q*." Our knowledge of the particular has often been referred to as sensory intuition, and of the very general as nonsensory intuition. The simplest, most familiar, and least helpful explanation of our possession of these two sorts of intuition is that we possess faculties which produce such knowledge. Accepting this explanation amounts to granting that the presence in our mind of the original starting points of knowledge is inexplicable and must be accepted as a brute fact. Aristotle was content with this solution, and so was René Descartes. In Cartesianism this inexplicability was woven into the fabric of a metaphysical dualism, according to which no mental event (such as a coming-to-know) could be caused by any sequence of physical events, and in which the only mental relation that could bring about a coming-to-know was the relation of being inferred from. This picture of the mind required that comings-to-know which were not preceded by inference be treated as uncaused causes, incapable of explanation.

Descartes's extreme rationalism led him to insist that sensory intuitions are not really cases of knowledge at all, and this in turn led him to hold that they are not really mental events but merely physical ones. Thus he did not recognize two intuitive faculties (one sensory and one nonsensory) but only one, the nonsensory. In his view, sense perception is in principle nonessential to attaining complete knowledge, although it is mysteriously necessary in practice. This paradoxical position was criticized by John Locke and others. Under the impact of such criticisms, a more moderate rationalistic position was developed, according to which both sense perception and the intellect are sources of genuine knowledge and enjoy equal status as intuitive faculties.

LINGUISTIC THEORY. The new moderate rationalism was attacked by the immoderate empiricism of David Hume, according to which our only intuitive faculty is that of sensory intuition. Hume, however, and such later empiricists as Bertrand Russell, continued to accept the Cartesian metaphysical framework, thus admitting that no explanation can be given of the fact that a physical event *p* (the modification of one of *S*'s sense organs) is frequently followed by the mental event *M* (*S*'s coming-to-know that *p*). They insisted, however, that an explanation can be given of the acquisition of our nonsensory intuitive knowledge and that consequently it is not necessary to postulate a special faculty that provides us with knowledge of first principles. The alternative explanation (in the form it was given by the logical positivists) was that all such knowledge is knowledge of analytic truths and that the process of acquiring such knowledge is identical with the process of learning the conventions of one's language. This view—sometimes called the linguistic theory of a priori knowledge—held that to know, for example, that all events are caused is simply to know something about the meanings of the words *event* and *cause*, and that this knowledge is acquired by easily understandable processes of psychological conditioning. To this suggestion, rationalists objected, first, that the process of learning the meaning of *cause* cannot be accounted for except by invoking a special faculty of intuitive acquaintance with universals; and, second, that the linguistic theory represents a confusion of acquiring knowledge with acquiring the ability to express this knowledge.

PRELINGUISTIC KNOWLEDGE. The rationalists held, concerning the linguistic theory, that even granted that it would be a violation of linguistic conventions to speak of "uncaused events," the real question is: How do we know that this is the right convention to adopt? Is not this latter piece of knowledge, knowledge of nonlinguistic fact? Are not linguistic conventions adopted on the basis of such prelinguistic knowledge? Such questions, many philosophers thought, show that the linguistic theory does not enable us to dispense with a faculty of nonsensory intuition. As long as the central presupposition of these questions—that *S* can properly be said to know that *p* prior to his ability to express *p* in language—was granted, this rationalist rebuttal created a new deadlock.

The influence of Cartesianism, and particularly of the Cartesian notion of sense perception as a special,

unanalyzable mental act correlated with certain modifications of sense organs, made it difficult to question this presupposition. Sense perception was, it seemed to most philosophers, a clear example of our ability to know facts without having the ability to express them. If a child, by virtue of his faculty of sensory intuition, can see that a physical object *O* has the sensory quality *Q* by a simple, uncaused act, prior to acquiring the ability to express this fact in language, then why cannot the same child see with his mind's eye that every event has a cause and, on the basis of this prelinguistic intuitive knowledge, check the correctness of conventions concerning the words *cause* and *event*?

BEHAVIORIST ANALYSIS. The notion of prelinguistic knowledge, and with it the whole Cartesian conception of comings-to-know as mental occurrences, was questioned by Gilbert Ryle, Ludwig Wittgenstein, and their followers. Under the influence of these writers, many philosophers have come to treat "*S* knows that *p*" not as a statement about *S*'s mind but as a statement that, besides presupposing the truth of *p*, asserts that *S* is disposed to assert *p* on appropriate occasions, and also either that *S* is prepared to give good reasons for believing that *p* or that *S* is justified in believing *p* even though he is unable to give reasons for believing that *p*. The last case covers all noninferential knowledge, both intuitive and nonintuitive. In the case of *S*'s nonintuitive, noninferential knowledge—that, for example, he has two children, or that there is a house in front of him—the criteria that establish that *S* is entitled to assert these propositions are of two sorts: those that determine whether he knows the meanings of the terms he uses and those that determine whether his situation and abilities are normal (where *normal* means, roughly, that the sincere reports of persons with these abilities in these situations are usually confirmed when checked by independent means). For example, *S* would be justified in believing that there is a house in front of him if he knew what a house is (that is, knew what *house* means), had his eyes open, and had normal vision. He would be justified in believing this even if, when asked, "How do you know that that's a house?" he was too unsophisticated to make any reply except "I just know." Whether *S* satisfies these criteria can be determined by public procedures—testing his grasp of the language, his vision, and his position vis-à-vis the house in straightforward and unmysterious ways.

According to the Cartesian view, what justifies *S* in believing *p* in the absence of an ability to produce good reasons for believing *p* is a special, private, introspectable mental state. *S* introspects to see whether or not he knows that *p*, and thus he knows intuitively that he knows that *p* and has better ground for the belief that he knows that *p* than anyone else can have. The behaviorist alternative asserts, on the contrary, that the fulfillment of public criteria is not just an external symptom of the presence of an occult mental state called knowledge, but that the statement of such criteria gives a full account of the meaning of "to know." This treatment of such cases of nonintuitive, noninferential knowledge as "I see that *O* is *Q*" is designed to replace the Cartesian notion of sense perception as a simple, unanalyzable act with the view that to see that *O* is *Q* cannot happen prior to the ability to use correctly the terms *O* and *Q* (or some equivalent expressions). Infants and animals, confronted by *O*, have sensations but do not have perceptions. They begin to perceive that *O* is *Q* when these sensations, and only these sensations, are accompanied by a disposition to assert or assent to "*O* is *Q*." Thus, they begin to perceive that *O* is *Q* only when their belief that *O* is *Q* becomes a reliable indicator of the truth of "*O* is *Q*."

This behaviorist analysis of nonintuitive, noninferential knowledge can be used to explain the difference between this case and the case of intuitive knowledge. The difference is that in the case of intuitive knowledge the only criterion that *S* must satisfy in order to be entitled to believe *p* without being able to offer good reasons for *p* is that he knows his language. The paradigms of intuitive knowledge—knowledge of "private" psychological states and knowledge of underived a priori truths—are such that if a person claims knowledge of this sort, the only way in which his claim can be refuted is to show that he does not know his language. For example, if someone sincerely believes that he is in pain, we cannot show that his belief is mistaken unless (as in the case of a young child) we can show that his use of *pain* is idiosyncratic. Again, if someone claims to know that every event has (or does not have) a cause, we cannot show that his belief is unjustified unless we discover that he does not understand what he is saying (and we discover this by discovering that his use of *event* or of *cause* is idiosyncratic). To know what one is saying is, in certain cases, to know that what one says is true.

Behaviorist analysis also permits an explanation of our possession of intuitive knowledge that dispenses with the notion of intuitive faculties. In the case of sensory intuition, the process of acquiring intuitive knowledge is simply the occurrence of certain sensations in a person who knows a language that contains ways of describing these sensations (that is, contains expressions whose utterance speakers of the language are conditioned to

correlate with occurrences of these sensations). In the case of nonsensory intuition, we acquire intuitive knowledge simply by reflecting upon our own linguistic behavior (where *reflecting* means, roughly, "asking ourselves questions about what we would say if …"). In both cases, the crucial precondition is knowledge of a language, and the process of acquiring this knowledge is taken to be a matter of psychological conditioning—conditioning whose operations are explicable entirely in terms of a stimulus-response model. Whereas according to the traditional Cartesian faculty view the difference between men and animals is a matter of man's possession of a special sui generis power (variously called awareness, consciousness, spirit, reason, and the like), this difference is regarded by many contemporary philosophers as a matter of the ability (due, presumably, to a more complex central nervous system) to respond in more diverse ways to a wider variety of stimuli—as a matter of degree rather than of kind.

CARTESIAN AND WITTGENSTEINIAN ATTITUDES. The difference between Cartesian and Wittgensteinian attitudes toward the fact that intuitive knowledge that *p*, such that belief in *p* is justified yet there is no way to settle doubt about *p*, exists may be summed up by saying that for a Cartesian the claim that belief in *p* is justified must reflect a natural fact—for example, some intrinsic feature of that belief (considered as a mental state), such as self-evidence. For the Wittgensteinian, this claim need reflect only a social convention. On the Cartesian view, it is only contingently true that we possess intuitive knowledge, a fact that is to be explained (if at all) by reference to the makeup of our minds. On the Wittgensteinian view, our possession of intuitive knowledge is a necessary truth, built into the use of the word *know*. The Cartesian reasons that since there cannot be an infinite regress—and thus justification of beliefs must stop somewhere—there must be certain kinds of belief that are intrinsically of a special sort, such that to have them is to know that they do not require justification. Followers of Wittgenstein reason that since there can be no infinite regress—and thus justification of belief must stop somewhere—one would expect, given our use of the word *know* to mean "justified belief," that there would be certain conventions dictating that certain beliefs are justified even in the absence of good reasons. For the Cartesian, these conventions reflect introspectable facts about the mind or about entities (such as universals) visible to the eye of the mind; for the Wittgensteinian, they do not reflect anything. To ask why we have procedures for settling doubt about *S*'s claim that he sees a house, although we do not have procedures for settling doubt about his claim that he has a pain, is, according to Wittgensteinians, to ask why we use the words *pain, house*, and *see* as we do. To such questions there is no answer. Nor is there any answer to the question why we use *event* or *cause* in such a way that it does not make sense to ask whether or not a given event was uncaused. We just do. That in certain cases it does not make sense to ask certain questions—for example, the question "How do you know?"—is, on this view, as much a matter of convention as the fact that one normally says "I am in pain" when being tortured but not when being caressed.

OBJECTIONS TO THE LINGUISTIC EXPLANATION. Much contemporary epistemological controversy consists of arguments for and against the behaviorist analysis of knowledge and the linguistic explanation of intuitive knowledge. The principal objections to the linguistic explanation are three: (1) It has been claimed that no behavioristic analysis of *believes* (and thus a fortiori of *knows*) can be achieved without recourse to terms that, like *believes* itself, exhibit intentionality. (2) It has been argued that the view that there is no awareness, perception, consciousness, or knowledge prior to the acquisition of linguistic ability makes it impossible to understand how we can learn language in the first place. (In rebuttal, it has been argued that to suppose that we learn how words are used by associating certain awarenesses with certain utterances is a misleading backward projection of the way in which an adult learns new words into the original learning of language by the child.) (3) It has been argued that the stimulus-response model is inadequate for explaining the learning of languages, on the ground that one who knows a language is able to produce grammatical sentences he has never heard. This fact has suggested to some theorists that we must postulate innate knowledge in order to explain language-learning.

This entry will not attempt to resolve these issues, but will only describe how the linguistic explanation has been brought to bear upon (*a*) the notion of unconscious inference, (*b*) the notion of intuitive awareness of universals, and (*c*) the notion of nonpropositional knowledge.

NONINFERENTIAL KNOWLEDGE AND UNCONSCIOUS INFERENCE

It has traditionally been held that all knowledge that is not intuitive is inferential, and thus that the cases of nonintuitive, noninferential knowledge should properly be regarded as the products of unconscious inference. This view is most familiar in the form of the phenomenalist

claim that *S*'s knowledge that, for instance, there is a white house before him is always the result of an inference from propositions concerning the sense data that *S* is currently having or concerning the appearances that the house is presenting to him. Proponents of this view regard *S*'s denial that he made such an inference or believed such propositions simply as evidence of a lack of philosophical sophistication. Such a view results from the assumption that only certain special propositions are suited, by virtue of their intrinsic properties, to be objects of noninferential knowledge. Thus, phenomenalists hold that "That is a white house" is inherently unsuited to be noninferentially known, whereas "I am now having a white sense datum" or "There now seems to me to be something white in my visual field" is inherently suited to be so known. The occurrence of an unconscious inference in *S*, they hold, is guaranteed by the fact that his belief is unsuited to be an expression of direct sensory awareness. No empirical evidence is allowed to disconfirm that such an unconscious inference was performed.

The criterion for being an expression of direct sensory awareness used by sense-datum theorists usually takes one or the other of the following forms:

(1) *p* expresses *S*'s direct sensory awareness if and only if *S* has intuitive knowledge that *p* (if, in other words, there are no procedures available that would provide better evidence against *p* than the fact of *S*'s belief that *p* provides for *p*), and if *S*'s coming to know that *p* is correlated with *S*'s having a certain sensation.

(2) *p* expresses *S*'s direct sensory awareness if and only if a sufficient condition of the acquisition of knowledge that *p* by *S* is that *S* has a certain sensation (so that none of *S*'s antecedent knowledge interferes to provide an interpretation of what his senses give).

These two criteria are often taken as interchangeable by philosophers who have gone in quest of the "given" elements in experience—for, at first blush, such intuitively knowable propositions as "I am in pain" or "I seem to be seeing something white" seem the most promising candidates for satisfying the second form.

The linguistic explanation of sensory intuition attempts to dispense with both the given and unconscious inference. According to the linguistic theory nothing could possibly satisfy the second form, since a sensation is never a sufficient condition for the acquisition of a bit of knowledge. Also, there is nothing paradoxical in saying that a man may simultaneously come to know, without performing any inferences, that this is an airplane, a Boeing airplane, and a B-29 as a result of a single modification of the eyes—the same modification that, in a child, would produce only the knowledge that this is something silver. According to this theory, the man's belief in all these propositions is justified because, roughly speaking, he has been conditioned to utter statements expressing each of them when certain sensory stimuli are received. Some men, as we say, just know a B-29 when they see one, and others do not. An aircraft spotter trained to respond to the appearance of a B-29 by saying "There is a B-29" would have a justified belief in this proposition even if he were unable to list any criteria for B-29-hood (and thus were unable to provide any reasons for his believing the plane to be a B-29).

For those who accept a linguistic explanation of intuitive knowledge, the traditional attempt to identify noninferential and intuitive knowledge by means of the notion of unconscious inference results from a confusion of the context of *S*'s acquisition of the knowledge that *p* with the context of his justifying his belief that *p* to one who doubts *p*. If an argument between *S* and a doubter of an empirical proposition *p* were carried to its ideal limit, *S* might eventually have to retreat to such intuitively known statements as "It seems to me that I remember that *q*" and the like. The ideal empiricist would be the man who never believes an empirical proposition *p* unless he has previously performed an inference embodying the argument that he would give in defense of *p* when challenged by a die-hard doubter. (The ideal empiricist, in other words, is the ideal Cartesian doubter; he always doubts every proposition he knows how to doubt.) The notion that we are all unconsciously ideal empiricists is a confusion of "*S* would not be able to justify his belief that *p* to a die-hard doubter without appealing to certain propositions that he knows intuitively to be true" with "*S* is not justified in believing *p* if he has not previously so justified his belief to himself."

Once we adopt the linguistic explanation of intuitive knowledge, its defenders argue, we see that whereas noninferential knowledge is a matter of one's disposition to make certain statements being a sufficient ground for one's belief that they are true, intuitive knowledge is a matter of that disposition serving as the best possible evidence for their truth. The propositions that can be noninferentially known by *S*, like those that can be intuitively known by him, are determined by *S*'s training, circumstances, and abilities, together with the conventions in force within his linguistic community. The fact that certain propositions are usually known noninferentially, and

others usually known intuitively, by normal adults has misled philosophers into thinking that certain special intrinsic properties belong to all those propositions, and only to those propositions, properties detectable by our mental eye. The linguistic theory, freeing us from the "mental eye" model, directs our attention to the factual criteria that we use in deciding whether a certain belief, held by a certain person, is justified.

INTUITIVE ACQUAINTANCE WITH CONCEPTS

A person is said to have intuitive acquaintance with a concept if he is able to understand a large range of propositions that employ a term signifying this concept and is unable to explain the significance of this term. Thus (confining ourselves, for the sake of simplicity, to descriptive concepts) we may say that *S* grasps F-ness intuitively if and only if he can use the expression "F" correctly, and he does not know any noncircular definition of "F," where a "definition of 'F'" is any true statement of the form "*X* is called 'F' (or 'an F') if and only if it is ______," and "noncircular" means that the blank is filled by some expression that neither contains "F" nor contains any word whose definition itself contains "F," nor any word whose definition contains words whose definition contains "F," and so on.

ACT OF ABSTRACTION THEORIES. As in the case of intuitive knowledge that *p*, there is no dispute among philosophers about the existence of intuitive acquaintance with concepts. Rather, as in the former case, controversy arises concerning the explanation of this fact. In this case also, philosophers working within a Cartesian tradition accept a "simple act" theory. On this traditional view, we possess a faculty called abstraction that, for example, peels the whiteness of white objects from these objects and holds the whiteness up before our mental eye; once we have whiteness clearly in focus, we can label it with the term *white* and thus can acquire a knowledge of how to use this term. This act of abstraction, like the act of intuiting that *p*, is specifically mental, simple, and unanalyzable. Within this Cartesian framework, the principal issue is that between rationalists and empiricists: Whether such a simple act of abstraction must be postulated to explain only our knowledge of apparently indefinable sensory concepts (like "white"), or whether it is also needed to explain our knowledge of apparently indefinable nonsensory concepts, such as "being," "cause," "necessity," or "good." Empiricists have traditionally held that these latter concepts are not grasped intuitively. They have claimed either that our knowledge of how to use terms signifying them is a result of our implicitly or unconsciously possessing noncircular definitions of them, or that these terms do not refer to concepts at all but are without meaning. Consequently, they have devoted themselves to proposing such definitions, or to developing theories of meaningfulness that would permit the conclusion that these terms have no meaning. Rationalists, on the other hand, have insisted that certain terms signify a priori concepts, and that none of the definitions of these terms proposed by empiricists (such as Hume's definition of *causation* as "constant conjunction") are adequate.

LINGUISTIC THEORY OF CONCEPTUAL INTUITION. The traditional account of our intuitive grasp of concepts contains many of the same elements as the traditional view of intuitive knowledge that *p*. It is again assumed that we need to account for a difference between humans and animals (the fact that we can use concepts, whereas animals can merely respond to stimuli) by postulating a simple sui generis mental act and that this simple act does not occur in all the cases that, prima facie, are cases of immediate knowledge, but that some such cases are cases of unconscious mediation. Just as recent philosophical thought has turned away from the notion that intuitive knowledge that *p* is to be regarded as such a simple act, and has offered an account of the acquisition of such knowledge in terms of a theory according to which the use of language is a necessary condition of the possession of any piece of knowledge, recent thought has likewise asserted that the ability to use "F" correctly is all that is signified by the phrase "acquaintance with F-ness," and thus that the notion of a prelinguistic grasp of F-ness is incoherent. According to this newer view, no object of acquaintance (such as a concept, conceived of as a sort of mental particular) need be postulated as that with which language learners correlate utterances of general terms. We learn such terms as *white* not by correlating utterances of them with anything but by being subjected to a conditioning process that leads us, after some trial and error, to utter these words in appropriate contexts in appropriate situations. This process need not, at any stage, involve our knowing the truth of any proposition of the form "*X* is called 'F' only when it is an instance of F-ness."

The older view, in insisting on the necessity of such knowledge, assumes that the process of learning the use of an indefinable word such as *white* must parallel the process of learning the use of a word by learning its definition. Just as we might correlate utterances of "bachelor" with situations in which we would be inclined to say

"unmarried male," and thus learn the meaning of "bachelor," so (the older theory holds) we correlate utterances of "white" with situations in which we are aware of whiteness. But, proponents of the newer view object, the only test we have for knowing whether we are aware of whiteness is whether or not we are inclined to utter "white." Nothing is added to an explanation of learning words ostensively by a reference to acquaintance with concepts, save the unverifiable claim to possess a piece of prelinguistic knowledge. If this newer view (largely due to Wittgenstein and his followers) is accepted, then what distinguishes us from the animals is not that they cannot perform the mysterious operation of intuiting concepts but simply that we can respond in much more various ways to a much greater variety of stimuli than they can (and, specifically, we can develop patterns of linguistic behavior). Once again, the difference between humans and animals reduces to the possession of language.

One advantage claimed by defenders of this newer view is that, if it is accepted, the old controversy about the existence of a priori concepts that divided rationalists from empiricists is rendered moot. The question of whether we must postulate a sort of nonsensory ostention of such concepts as causality, or an innate grasp of them, no longer arises if the same sort of process that enables us to learn the use of *white* enables us to learn the use of *cause*. To acquire the concept of causality is, on this view, to learn the use of the word *cause*; this can be done without correlating utterances of *cause* with anything, but simply by trial and error: Sometimes when we say "This caused that," we are rebuked, and sometimes praised, until gradually we get it right. (Before we got it right, we were said not to know the meaning of *cause*, just as we were said not to know the meaning of *white* as long as we called "white" what our parents called "gray.") The question of whether *cause* (and other terms that have been held to signify a priori concepts) is definable without circularity now loses its philosophical interest.

INTUITION AS NONPROPOSITIONAL KNOWLEDGE

The final sense of intuition comes primarily from Kant, who defined "intuition" as "knowledge that is in immediate relation to objects" (see *Critique of Pure Reason*, A19–B34, A320–B377). By *immediate* he here meant "without the mediation of concepts," and he took sense perception as the paradigm of intuition (although he also argued for the existence of pure intuitions of space and time). Kant sharply distinguished immediate knowledge from knowledge of the truth of judgments concerning the objects sensed, since he held that the formation of judgments requires the addition of concepts to intuitions. The former sort of knowledge is a necessary condition of the latter. The knowledge gained in sense perception is expressed by judgments concerning the objects sensed but exists prior to the formation of these judgments. Perceptual knowledge of *O* is, on this view, not reducible to knowledge that *O* has certain properties.

This distinction between immediate knowledge of objects and mediate knowledge of facts about these objects was formulated by Russell, in *The Problems of Philosophy*, as the distinction between "knowledge by acquaintance" and "knowledge by description." He proceeded to explain a priori knowledge by postulating a faculty, analogous to sensation, that acquaints us with universals and with the relations between universals. The assertion of the existence of universals has, traditionally, gone hand in hand with the faculty explanation of our intuitive knowledge of a priori truths and of our grasp of nonsensory concepts. It is still current among contemporary philosophers who resist the linguistic explanation of this knowledge. These philosophers include both such traditional rationalists as Brand Blanshard and phenomenologists who adopt Edmund Husserl's notion of intuition of essences.

The Kantian notion of sense perception as a kind of nonjudgmental knowledge has had the effect of opening the door to the suggestion that we possess a certain sort of knowledge that is like sense perception, or Russellian acquaintance with universals, in being immediate but unlike either in being inexpressible. In other words, it is suggested that we have an intuition of a certain object *O* even though we do not know the truth of *any* proposition of the form "*O* is *Q*." The reason usually given for our failure to have the latter sort of knowledge is that conceptual thought (or language) is inadequate to capture the essence of *X*. For example, Bergson argued that duration cannot be captured by concepts (nor, a fortiori, expressed in language) because concepts (and thus language) are designed precisely to freeze and stabilize (and thus to distort) the flux of experience, whose essence is duration. Again, God's perfect simplicity—his identity with his own attributes—is held to make it impossible truly to apply any predicate to him, and thus to know any true propositions about him.

Philosophers who adopt the view that there is no knowledge prior to the possession of language, and who construe knowledge in the behavioristic manner, naturally object to the notion of nonpropositional knowledge. On their view, the original Kantian notion of sense per-

ception as a kind of knowledge is based upon a confusion. Once this confusion is dissipated, the analogy to sense perception that is the basis of Russellian accounts of a priori knowledge and of theories of inexpressible intuition will no longer be available, and the notion of knowledge of *O* that is irreducible to the knowledge that *O* has certain features will appear as paradoxical as it really is. The original confusion, these philosophers argue, is that of the cause of the belief that some sensed object *O* has the feature *Q* with the justification of this belief. Specifically, the fact that knowledge that *O* is *Q* is caused by a sensation of *O* is combined with the assumption that nothing can serve to justify *S*'s claim to know about *O* except another piece of knowledge about *O* by *S*. This produces the conclusion that the mere sensing of *O* is itself a case of knowing—distinct from, because giving a ground for, the knowledge that *O* is *Q*. Since sensing *O* is construed as a direct relation between the knower and *O*, whereas knowing that *O* is *Q* is construed as a relation between the knower and something distinct from *O* (a fact or a proposition), it is inferred that there are two sorts of knowing, one of which is primitive and direct and the other derivative and indirect. A causal condition for knowledge is thus confused with a special type of knowledge—knowledge by acquaintance.

Philosophers who deny the existence of such nonpropositional knowledge by acquaintance argue that the notion of knowledge of *O* that is not knowledge that *O* has some feature is neither present in ordinary usage nor part of a useful explanatory theory. On their view, all knowledge of objects is knowledge of the truth of propositions about these objects. This anti-Kantian position is supported by, and supports, the anti-Cartesian behaviorist position, according to which knowledge cannot occur prior to the ability to learn language. Although it is logically possible to hold both that there can be prelinguistic knowledge of facts and that there is no such thing as knowledge of particulars as distinct from knowledge of facts, this position is not popular. Contemporary epistemological thought is, by and large, split between those who adopt both a Cartesian "simple act" explanation of the intuitive knowledge that *p* and a Kantian notion of nonpropositional knowledge as a necessary condition for intuitive propositional knowledge, and those who reject both of these views in favor of a radically behavioristic approach.

INTUITION OF THE INEXPRESSIBLE

Even philosophers who have remained faithful to the traditional Cartesian and Kantian positions tend to criticize the use of the notion of intuition as nonpropositional awareness made by such philosophers as Fichte, Bergson, and contemporary Thomists. Their criticism is based on the view that the only criterion for knowing whether *S* has nonpropositional knowledge of *O* is his knowledge of the truth of propositions about *O*. Thus both groups reject claims to have knowledge that one is unable to express (except, perhaps, in analogies and metaphors). Anti-Cartesian philosophers, however, argue that it is precisely the Kantian view that sensing is a kind of knowing that opens the gates to claims to intuit the inexpressible. This view leads naturally to the conclusion that even the objects of ordinary sensory acquaintance are incommunicable and inexpressible. No amount of talk by Jones (who has seen *O*) will suffice to reproduce in Smith (who has not) the sensation Jones had when he was in the presence of *O*. This failure to reproduce an experience is, given the view that sensing is a kind of knowing, taken as a failure to convey knowledge of *O*, even though Smith may learn, from Jones's reports, every fact about *O* that Jones knows. We thus find ourselves adopting a novel, and peculiarly philosophical, sense of "express"—a sense in which an experience would be expressed only if it were reproduced. Whereas in the normal sense of the term, my seeing a white house is completely and adequately expressed by some finite set of such propositions as "That's a white house," in this new sense such propositions are inherently unsatisfactory surrogates. This line of thought, opponents of nonpropositional knowledge argue, plays into the hands of those who, like Bergson, hold that language is inadequate to reality.

The claim that language is inadequate to express one's intuitive knowledge of reality would, in itself, be harmless. However, the danger of adopting this new meaning of "inexpressible" is that we may find ourselves claiming private justification for our moral, philosophical, religious, aesthetic, or other beliefs by saying, "Although I cannot, of course, express (or communicate or put into words) the experience that I had, and hence cannot supply you with reasons for believing that *p*, I am nevertheless entitled to believe that *p* solely on the strength of that experience." The plausibility of this sort of reasoning stems from the fact that, in the case of noninferential belief about physical objects, we sometimes say things like "Since you haven't seen a flying saucer, you have no reason to believe that there are flying saucers; but I have seen one, and so I do believe in them." Here we seem to be justifying a belief solely on the basis of private experience. The difference is that "I saw a flying saucer" is a complete and adequate expression of this experience, in the ordinary sense of "express." The justification is sufficient because the state-

ment that an experience *E* was had analytically implies *p*. (If *S* saw a flying saucer, then there are flying saucers to be seen.) In the former case, however, the statement that an experience *E* was had cannot entail any statement about the object of the experience because the nature of the experience is, ex hypothesi, inexpressible.

This obvious disanalogy is veiled by the fact that in the second, philosophical sense of "inexpressible," our experiences of seeing houses or flying saucers are just as inexpressible as the Thomists' intuition of Being, or Bergson's intuition of duration. In other words, a tacit shift to a new sense of "express" creates the sophistical argument "Since your sensory experiences are inexpressible, and yet sufficient to justify your beliefs, it is unfair of you not to let my inexpressible nonsensory experiences justify my beliefs." Of course, in the ordinary sense of "express," sensory experiences are as expressible as experiences can be.

In addition to this criticism of the ambiguity contained in the philosophers' use of "inexpressible," a further criticism of such claims to private justification is available if the behaviorist view of the nature of justification of claims to noninferential knowledge is adopted. If this justification is viewed not as a matter of an intrinsic, introspectable property (self-evidence) of certain beliefs but rather as a matter of social convention, then one will hold that we know which of our noninferential beliefs are justified only by knowing which ones our peers would agree are justified. In the flying saucer example, we rightly think that our belief in flying saucers is justified if we think we have seen flying saucers, because we are confident that anyone who had had the sensations we have had would have been disposed to utter "I see a flying saucer." We know that our belief is justified because our peers admit that if they should ever have an experience of the sort we claim to have had, they would share our belief. The only element of privacy lies in the fact that they can have doubts about, for example, whether we are being truthful in claiming that we had this experience, or whether we were sober, or attentive, whereas we cannot.

In the "inexpressible intuition" case, however, we cannot tell whether our peers would share our belief if they shared our experience, for we do not know what our experience was. Here we could speak of a private justification only if we had a private language in which we could express to ourselves, although to no one else, what we experienced, and private criteria of justification formulated (in part, at least) in this private language. But, aside from the general difficulties in the notion of a private language pointed out by Wittgenstein, "private criteria of justification" is an intrinsically paradoxical notion. One can no more have private rules for justifying beliefs than one can have private rules for justifying actions. A criminal has no greater claim on our sympathy if he proclaims that his private ethical code differs from ours, and a believer in untestable beliefs has no greater claim on our attention when he says that his epistemological code is not ours.

See also Intentionality; Objectivity in Ethics.

Bibliography

The medievals used "intuitive cognition" as we would use "sensory intuition" to refer to knowledge about objects present to the senses. The term was opposed to "abstractive cognition," which included memory and imagination. They also, however, used "intuition" to refer to a vision of God. This use of "intuition" for any sort of knowledge that has the same noninferential character as knowledge of the apparent features of an object present to the senses was continued by Descartes (*Regulae* XII), Benedict de Spinoza (*Ethics* II, Prop. 40, Note 2), and Locke (*Essay concerning Human Understanding*, Book II, Ch. 2, Sec. 1). These philosophers used the term as we would use "nonsensory intuition"—to refer to our noninferential knowledge of, for instance, mathematical axioms and analytic truths, and of the validity of valid inferences. Between Descartes and Kant, "intuition" was rarely used in reference to perceptual knowledge, nor was a clear distinction made between propositional and nonpropositional knowledge. Since Kant, however, it has been usual to speak of both nonpropositional perceptual knowledge of a particular, and of the propositional knowledge derived from this nonpropositional knowledge, as cases of intuition.

Whereas Kant had denied the existence of intellectual intuition (nonpropositional knowledge of insensible objects), Fichte asserted it in his *Werke*, edited by I. H. Fichte (Berlin, 1845), Vol. I, pp. 463ff. However, Fichte argued that he did not really disagree with Kant because the object of this intuition, the Transcendental Ego, was an act rather than a thing. The same strategy is adopted by contemporary neo-Thomists, who speak of an intuition of Being; what is intuited, they say, is an act rather than a thing or an essence. See Jacques Maritain, *Existence and the Existent* (New York: Pantheon, 1948), Ch. 1, and Étienne Gilson, *Being and Some Philosophers* (Toronto: Pontifical Institute of Mediaeval Studies, 1949), Ch. 6. The most influential recent proponent of a faculty of nonpropositional knowledge other than sense perception is Henri Bergson; see his *Introduction to Metaphysics* (New York: Putnam, 1913). For a criticism of Bergson's notion of intuition, consult G. Watts Cunningham, *A Study in the Philosophy of Bergson* (New York: Longmans Green, 1916), Ch. 3. For a discussion of the philosophical importance of the ineffable intuitions claimed by mystics, see W. T. Stace, *Mysticism and Philosophy* (New York, 1960), Chs. 1 and 3.

W. H. Walsh, *Reason and Experience* (Oxford: Clarendon Press, 1947), contains an account of traditional controversies between rationalists and empiricists concerning intuitive knowledge. For the traditional view that intuitive knowledge of facts about objects sensed is based on a nonpropositional

acquaintance with these objects, see Bertrand Russell, *The Problems of Philosophy* (London: Williams and Norgate, 1912), Ch. 5. Criticism of the notion of knowledge by acquaintance, which Russell develops in *The Problems of Philosophy*, is found in H. L. A. Hart, "Is There Knowledge by Acquaintance?" in *PAS*, Supp. 23 (1949): 69–90; also see the essays by G. E. Hughes and J. N. Findlay on the same topic in the same volume, 91–128, and Wilfrid Sellars, *Science, Perception, and Reality* (London, 1963), pp. 127–196. Additional criticisms of the view that sensing is a form of knowing occur in H. A. Prichard, "The Sense-Datum Fallacy," in *PAS*, Supp. 17 (1938): 1–18; Wilfrid Sellars, "Physical Realism," in *Philosophy and Phenomenological Research* 15 (1954–1955): 13–32; and Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), Ch. 7.

The notion of unconscious inference is presented in Bertrand Russell, *The Problems of Philosophy*, Ch. 13, and a defense of this notion in Gilbert Harman, "How Belief Is Based on Inference," in *Journal of Philosophy* 61 (1964): 353–359. For the criteria of direct sensory awareness, see Bertrand Russell, "On Verification," in *PAS* 38 (1937–1938): 1–20. Russell's *Analysis of Mind* (London: Macmillan, 1921), Ch. 12, states the view that beliefs are introspectable mental occurrences. For criticism of this view, see Gilbert Ryle, *The Concept of Mind*, Chs. 2 and 5. The view that we can introspectively differentiate knowledge from mere belief is found in H. A. Prichard, *Knowledge and Perception* (Oxford, 1950), p. 88. Prichard is criticized on this point by Norman Malcolm in his *Knowledge and Certainty* (Englewood Cliffs, NJ: Prentice-Hall, 1963), p. 58.

For the contemporary reaction to Cartesianism, see Gilbert Ryle, *The Concept of Mind*, and Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1953). An earlier reaction against the Cartesian account of intuitive knowledge is C. S. Peirce's *Collected Papers* (Cambridge, MA: Harvard University Press, 1933–1958), Vol. V, pp. 135–189. For the linguistic account of intuitive knowledge of the truth of propositions, see Wilfrid Sellars, *Science, Perception, and Reality*, pp. 164–170. The view that intuitive knowledge of a priori truths is founded upon a nonpropositional knowledge of universals or essences is found in Bertrand Russell, *Problems of Philosophy*, Ch. 10; Brand Blanshard, *Reason and Analysis* (La Salle, IL: Open Court, 1962), Chs. 6, 9, 10; and Edmund Husserl, *Ideas* (New York: Macmillan, 1931). For critical discussion of this view and of the linguistic account of a priori knowledge, see Arthur Pap, *Semantics and Necessary Truth* (New Haven, CT: Yale University Press, 1958).

OTHER RECOMMENDED TITLES

BonJour, Laurence. *In Defense of Pure Reason*. New York: Cambridge University Press, 1998.

Casullo, Albert. *A Priori Justification*. New York: Oxford University Press, 2003.

DePaul, Michael, and William Ramsey, eds. *Proceedings of the Notre Dame Intuition Conference*. Totowa, NJ: Rowman & Littlefield, 1997.

Weatherson, Brian. "What Good Are Counterexamples?" *Philosophical Studies* 115 (2003): 1–31.

Richard Rorty (1967)
Bibliography updated by Benjamin Fiedor (2005)

INTUITION [ADDENDUM]

In the history of philosophy "intuition" has been used primarily as a term for an intellectual, or rational, episode intimately tied to a priori knowledge. The term has sometimes been used in a broader way to include certain sensory episodes (appearances) and certain introspective episodes (e.g., inner awareness of the passage of time). In contemporary philosophy this broader use has fallen out of fashion (except among Kantians), and the narrower use prevails.

An intuition in this sense is simply a certain kind of seeming: For one to have an intuition that *P* is just for it to seem to one that *P*. This kind of seeming is intellectual, not sensory or introspective, in the following sense: Typically, if it is possible for someone to have the intuition that *P*, then it is possible for someone to have the intuition that *P* in the absence of any particular sensory or introspective experiences relevant to the truth or falsity of the proposition that *P*. For this reason, intuitions are counted as "data of reason" not "data of experience." In this connection, intuitions are sometimes called "a priori intuitions" or "rational intuitions."

Intuition must be distinguished from belief: Belief is not a seeming; intuition is. For example, I have an intuition—it still seems to me—that the naive set-abstraction axiom from set theory is true despite the fact that I do not believe that it is true (because I know of the set-theoretical paradoxes). There is a rather similar phenomenon in sense perception. In the Müller-Lyer illusion, it still seems to me that one of the two arrows is longer than the other, despite the fact that I do not believe that one of the two arrows is longer (because I have measured them). In each case, the seeming persists in spite of the countervailing belief. Similar considerations show that intuitions must likewise be distinguished from guesses, hunches, and common sense.

Many philosophers identify intuitions with linguistic intuitions. But this is mistaken if by "linguistic intuition" they mean intuitions about words, for most of our intuitions simply do not have any linguistic content. Other philosophers think of intuitions as conceptual intuitions. Nothing is wrong with this if "conceptual intuition" is understood broadly enough. But there is a common construal—originating in David Hume's notion of relations of ideas and popular with logical positivists—according to which conceptual intuitions are all analytic. The problem is that countless intuitions are not analytic on the traditional construal of that term (convertibility into a logical truth by substitution of synonyms). For example, the intu-

ition that, if region r_1 is part of region r_2 and r_2 is part of region r_3, then r_1 is part of r_3. Possibility intuitions are also not analytic (e.g., in epistemology the intuition that the Gettier situations are possible). In response, some philosophers have countered that possibility intuitions are just intuitions of consistency, but this view is mistaken on several counts. For example, it is consistent to hold that region r_1 is part of r_2, r_2 is part of r_3, but that r_1 is not part of r_3, despite the fact that such a thing is not possible.

Standard practice in logic, mathematics, linguistics, and philosophy is to use intuitions as evidence. (For example, in epistemology Roderick Chisholm uses intuitions to show that traditional phenomenalism is mistaken, and Edmund Gettier uses intuitions to show that the traditional identification of knowledge with justified true belief is mistaken. In metaphysics Saul Kripke uses intuitions to show that, if water is H_2O, then it is necessary that water is H_2O. In philosophy of mind, Hilary Putnam uses intuitions to show that logical behaviorism is mistaken, and so forth.) A great many philosophers believe that use of intuitions is essential to the indicated disciplines.

Radical empiricists, who doubt that intuitions have evidential weight, usually defend their view by pointing to the fact that intuitions can be unreliable. They cite, for example, the fact that our intuitions about naive set theory are in conflict with our intuitions about classical logic. But this shows only that traditional infallibilism is mistaken, not that intuitions lack evidential weight. After all, sense perceptions have evidential weight even though they can be unreliable. (Incidentally, although various cognitive psychologists—Peter C. Wason, Philip Johnson-Laird, Eleanor Rosch, Richard E. Nisbett, D. Kahneman, A. Tversky, and others—have examined human rationality with a critical eye, their studies have not attempted to test empirically the reliability of intuitions, and it will be quite difficult to do so.)

Why should intuitions have evidential weight? A plausible answer is that intuitions have an appropriate tie to the truth: As a noncontingent fact, if a subject's cognitive conditions (intelligence, attentiveness, and so forth) were suitably close to ideal, the subject's intuitions would be sufficiently reliable to permit the subject to arrive at a mostly true theory regarding the subject matter of those intuitions. This is a consequence of an analysis of what it is to possess concepts determinately: A necessary and sufficient condition for determinately possessing one's concepts is that one's intuitions have this kind of tie to the truth; if the subject's intuitions lacked this sort of tie to the truth, that would only show that the subject did not determinately possess those concepts (or that the subject's cognitive conditions were not sufficiently good). In contemporary philosophy, many have come to accept (some form of) this moderate rationalist theory of intuitions and concept possession.

See also A Priori and A Posteriori; Belief; Chisholm, Roderick; Empiricism; Hume, David; Kripke, Saul; Philosophy of Mind; Putnam, Hilary; Truth.

Bibliography

Bealer, G. "A Priori Knowledge and the Scope of Philosophy." *Philosophical Studies* (1966), with replies by E. Sosa and W. Lycan in the same volume.

BonJour, L. "Against Naturalized Epistemology." *Midwest Studies in Philosophy* 19 (1994).

Kahneman, D., and A. Tversky. "On the Study of Statistical Intuitions." In *Judgment under Uncertainty,* edited by D. Kahneman, P. Slovic, and A. Tversky. Cambridge, U.K.: Cambridge University Press, 1982.

George Bealer (1996)

INTUITION [ADDENDUM 2]

An intuition is a noninferential awareness of something: a concept, a proposition, space or time, a physical object, our own existence, or God. While sometimes people talk of sensory intuitions of perceptual objects, by which they mean an immediate awareness of how they appear, this use of "intuition" is becoming more rare. Nowadays philosophers use the term primarily to mean a nonsensory and nonintrospective awareness of a proposition or concept. Some philosophers hold that an intuition must be of a proposition that seems necessarily, or possibly, true. But people who lack the concepts of necessity and possibility are able to have something very like what philosophers call intuitions. So a more plausible view is that a person has an intuition that P if and only if P seems true, or possibly or necessarily true, where that appearance is intellectual—that is, based on the understanding, not on perception or introspection. George Bealer thinks that intuitions are not beliefs because we can disbelieve something that still appears true. Perhaps some argument has convinced us that in a lottery with seventy-six million tickets we know before the drawing that we hold a losing ticket if the ticket is in fact a loser, but it may still seem that at that time we do not know that it will lose.

Thus, by "intuition" most philosophers mean a rational intuition—or a rational insight—that is based solely on understanding the proposition that is its object. The intu-

ition that a person who commits suicide kills himself, that nothing can be red and green all over at the same time and in the same respects, and that it is prima facie wrong to flog or torture an infant to death are based solely on understanding the relevant propositions. These examples show that the propositions that are the objects of intuitions can be either analytic or synthetic, where analytic propositions are ones that can be reduced to logical truths by the substitution of relevant synonyms, and synthetic ones are nonlogical truths that cannot be so reduced. The first proposition about suicide is analytic but the other two are not.

The propositions that are the objects of intuitions need not be true, though they must at least seem true. People often have the intuition that an unmarried male of marriageable age is a bachelor but come to see that this is false once they consider the example of the pope, who is such an unmarried male of marriageable age but not a bachelor. This shows that intuitions are fallible.

It is also possible to distinguish between immediate and mediate intuitions. For some propositions, such as "bachelors are unmarried," the proposition immediately seems true upon considering it. For others, such as "knowledge requires true belief," or "personal identity requires continuity of memory," reflection may be required before a sufficient understanding of the concepts involved develops and enables one to "see" that these things are true. A person might have to consider examples where it is intuitively obvious that a person lacks knowledge because he has a false belief before the general claim that knowledge requires true belief seems true to him. Similarly, a person might have to consider a case where someone has completely and irretrievably lost her memory in an auto accident to see that personal identity requires continuity of memory.

Intuitions are mental events that occur and then end, and so are unlike beliefs that are dispositions that endure even when nothing is presently before the mind (as happens when one is in a dreamless sleep). However, it is the understanding, which is the basis of rational intuitions, that provides a priori justification and sometimes knowledge. We are justified in believing that if A is greater than B, and B is greater than C, then A is greater than C solely on the basis of understanding what that proposition asserts. We will know that proposition, and not simply be justified in believing it true, provided that it is true and we believe it on the basis of the corresponding intuition.

There are some necessary truths that we cannot know, or even be justified in believing, on the basis of intuition. No one can be justified in believing that water is H_20 solely on the basis of understanding that proposition. Empirical investigation is needed. However, we can know on the basis of intuition that some clear, odorless, colorless, drinkable liquid that does not have the same underlying physical structure as water is not water.

One challenge to the justificatory force of intuitions is that they lack such force when directed to propositions that are expressed by sentences that contain natural kind terms, terms such as water, aluminum, horse, topaz, and ruby. Further, some argue that the concepts that philosophers are interested in are natural kind concepts. Some hold, for instance, that causation, knowledge, and justice are natural kind concepts whose essence can only be discovered empirically. Note that some terms such as "ruby" might express what might be called hybrid concepts because we can know via intuition that a ruby must be red, but can learn only through empirical investigation what chemical structure rubies must have. Rubies differ from topaz in this respect because topaz can be any color, and the essence of topaz can be discovered only by empirical investigation. If concepts of interest to philosophers are either not natural kind concepts at all or hybrid concepts such as the concept of a ruby, then intuitions can be used to discover, and justify, conceptual truths involving those concepts.

Another challenge to the justificatory force of intuitions is that people with different backgrounds have conflicting intuitions about, say, the nature of knowledge. Gettier examples show that having a justified true belief is not sufficient for knowledge. For instance, if you are driving down a road where ninety-nine out of a hundred structures that look like barns are really indistinguishable facades, you do not know, though you are justified in believing, that you are looking at a barn if you happen to stop in front of the one barn on that road. Here it seems intuitively obvious to many that you have a justified true belief but lack knowledge. However, it is reported that people from other cultures do not have this intuition. So how can intuitions provide justification if there is this sort of disagreement? Surely no one would think that, say, clairvoyance had justificatory force if people "saw" events happening at a distance but did not agree on what they "saw."

A third challenge says that if intuitions provide justification then there must be a faculty of intuition similar to one of our five senses or to memory. But because there is no such faculty, intuitions cannot provide justification. One response to this objection is to maintain that intuition is a faculty, similar to the language faculty and similar to what Jerry Fodor calls a module, and then ask why

that is a problem. Another response denies that intuitions must come from some faculty if they are to provide justification.

A fourth challenge says that we can never be justified in believing that intuitions sometimes provide justification because we would have to rely on intuitions that say they do, and such a circular argument could never provide real justification. Of course, this is a problem for all basic sources of evidence. For instance, to be justified in believing that memory provides justification, we would have to remember instances where our memories have been correct. But then we would be relying on memory to justify reliance on memory.

A plausible view to hold is that intuitions can provide a priori justification when their objects are certain sorts of propositions, say, mathematical, logical, or what at least appear to be conceptual truths, though this justification is fallible and can be defeated by widespread disagreement among people who understand the relevant propositions equally well. An argument for the view that intuitions sometimes justify can be constructed from the notion of what it is to possess a concept: To possess the concept, C, we must be able to apply C and not-C correctly to most hypothetical situations. Hence our possession of the concept guarantees reliability in its application, though reliability does not require infallibility.

See also A Priori and A Posteriori; Fodor, Jerry A.; Perception.

Bibliography

Audi, R. *Moral Knowledge and Ethical Character*. Oxford, U.K.: Oxford University Press, 1997.

Boghossian, Paul, and Christopher Peacocke, eds. *New Essays on the A Priori*. Oxford, U.K.: Oxford University Press, 2000.

BonJour, L. *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge, U.K.: Cambridge University Press, 1998.

DePaul, Michael, and William Ramsey, eds. *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield, 1998.

Bruce Russell (2005)

INTUITIONISM, ETHICAL

Philosophers thought of as intuitionists include Henry Sidgwick, H. A. Prichard, W. D. Ross, C. D. Broad, and A. C. Ewing. More recent intuitionists include Derek Parfit, John McDowell, and Thomas Scanlon. Though all but one of these philosophers are British, the expression "British intuitionism" standardly refers only to work done in the first half of the twentieth century by Prichard, Ross, Broad, and Ewing.

WHAT IS INTUITIONISM?

To be an ethical intuitionist is to hold a combination of five views in metaethics, only one of which says anything about intuitions. The first view is the pluralist view that there are many different ways in which an action can get to be right or wrong, good or bad. This is opposed to monism, the view that all moral requirements can be captured in one basic principle. The classic example of monism is John Stuart Mill's utilitarianism, which holds that an action is right if and only if, of all available alternatives, it has the best (or least bad) effect on the general happiness.

Note, however, that there are two sorts of monism: monism about the right and monism about the good. Mill's position combines both. He held that there is only one way for an action to get to be right, and only one sort of thing that is good; the only sort of thing that is good is happiness, and the only way for an action to be right is to produce as much of that good as possible. G. E. Moore, by contrast, was a monist about the right and a pluralist about the good; he agreed that a right action is one that produces the most good, but held that there are several different goods, mainly social intercourse and aesthetic appreciation, which cannot be reduced to one. So Moore is not an intuitionist, because he is a monist about the right, though he agreed with the intuitionists on everything else. Intuitionists, by contrast, combine both forms of pluralism. W. D. Ross, for instance, provided a list of goods, as Moore did, but also argued that there are many different moral duties and that they cannot all be reduced to the duty to produce the most good.

The second view is the realist (or "objectivist") view that some ethical or moral judgments are true and others are false; there is truth and there are facts of the matter in ethics. This is opposed to the noncognitivist claim that moral assertions are expressions of attitude (pro or con) rather than of belief; in saying that an action is wrong, one is not trying to characterize correctly some slice of moral reality, but expressing one's opposition to the action, the stance one takes toward it.

The third view is the nonnaturalist view that moral facts are not natural facts; the opposite view is naturalism. A natural fact is one that can be discovered using the methods of the natural sciences. (At least, this is one way

of trying to say what it is for a fact to be natural.) There are naturalist and nonnaturalist forms of utilitarianism. The naturalist form holds that rightness *is* the natural property of increasing general happiness; sociologists are presumably the people best equipped to tell us which actions have that property. Mill, however, held that actions that increase the general happiness have the *further* property of being right, and he showed no signs of thinking that this further property is natural. (One could however think that to show a moral property to be natural we do not need to find a second, "natural" way of picking it out.) Intuitionists, anyway, think that moral properties and moral facts are irreducibly nonnatural. They tend to argue for this position by pointing out that moral properties and facts are *normative*, and insisting that nothing normative can also be natural. The most famous way of making this point is Moore's "open question argument": since the question "Is increasing the general happiness being right?" is not the same question as "Is increasing the general happiness increasing the general happiness?", to increase the general happiness cannot be the same thing as to be right.

The fourth view is that the normative status of moral or ethical facts cannot be explained, but also requires no explanation. Normativity is not a mystery. When we say that it matters whether people are free or oppressed, we are saying that something is morally important, or makes a moral difference. Intuitionists think that this "making a moral difference" is a feature that some things have and others do not, but they do not think that the special nature of such features is one that calls for elaborate explanations. In this they are opposed to those nonnaturalists who feel called on to provide such explanations—in particular, they oppose Kantians, whose explanations of the morally important run in terms of some relation to the will of a rational being. Compared to the Kantians, intuitionists are quietists; they do not feel the need to say any more.

The fifth view is the one that talks about intuitions. If there are these distinct, nonnatural, normative facts and features, how do we find out about them? By which aspect of our intellectual or sensory equipment are we rendered capable of discerning when an action is right and when it is not? Intuitionists maintain that we are capable of coming to know *basic* moral facts directly, in ways that involve no inference from other nonmoral facts. They have to say this, since it seems impossible simply to move by inference from a natural belief to a normative one, given the great difference intuitionists discern between the two. And they have to say that there are basic moral facts, since even if some of our moral beliefs are defended by appeal to others, this process has to stop somewhere. There must, then, be some basic moral beliefs, which if true are beliefs in basic moral facts. How then do we come to recognize those basic moral facts? The intuitionist answer is (supposedly): by a special faculty of moral intuition. Hence the name.

COMMENTS ON THESE VIEWS

The form of monism that intuitionists targeted most eagerly would now be called consequentialism: the view that actions are made right or wrong by the value of their consequences. Intuitionists tended to argue against this view by appeal to example. Ross (1930, 1939), for instance, said that when I keep my promises, my thoughts are not normally on the future, on the good that I will achieve by doing what I once promised to do, but on the past, on the fact that I did promise to do it. That is my reason for doing it, and it is my having promised to do it that makes the action my duty. More generally, Ross thought that actions can become duties in various ways, only some of which have anything to do with consequences. One can have duties that derive from one's role as a neighbor, or as a teacher, or as a friend, and these duties are not necessarily related to making things go better. In fact, they are not grounded in thoughts about what is better or worse, in considerations of value, at all. At this point Ross is expressing a deontological stance; and indeed intuitionists are standardly deontologists.

Historically speaking, intuitionists tended not to argue for the view that there is truth in ethics; they more or less assumed it. They were familiar with the idea that to express a moral opinion involves expressing an attitude or feeling, but Ross wrote, "What we *express* when we call an object good is our attitude towards it, but what we *mean* is something about the object itself and not about our attitude towards it" (1939, p. 255).

Nonnaturalism arouses extreme passions in some quarters. The real pressure behind naturalism is the sense that there cannot be two distinct realms: the familiar natural realm investigated by physics and a much less familiar normative realm learned about in quite other ways. If there really are such properties as right and wrong, good and bad, they must be properties that natural objects such as people and actions can have (and of course physical objects too can be good or bad, of course, though not morally so). But surely natural objects can have only natural properties, and so the normative properties must somehow be natural. To claim anything else is to commit oneself to the existence of some most peculiar features,

quite unlike, and utterly unconnectable to, any natural features. To this the intuitionists reply that there is a connection; it is *because of* their natural features that objects come to have, or to lack, normative ones. This "because of" is not a causal "because," of course. Actions are not caused to be wrong by their natural features; they are made wrong by them. Put another way, the natural features are the reasons why an action is wrong, the reasons for not doing it. The question then focuses on this notion of a reason, which is itself not a natural notion but a normative one. Intuitionists would say that nobody can do without this notion, and that it is not possible to think of "being a reason for" as a natural relation.

This focus on what it is for one thing to be a reason for an action gives the intuitionists something to say against the Kantians. Is it possible to think that the relation of "being a reason for" is one that can be, and needs to be explained (whether in natural terms or, as the Kantians would urge, in nonnatural terms)? Intuitionists have more recently tended to say that the basic notion here is that of favoring; a fact is a reason for an action if it favors doing that action. But there is, they say, very little that can be said about this notion of favoring, even though it is one with which we are all familiar.

Finally, with respect to intuitions, we need first to decide what sorts of facts are basic moral facts. Henry Sidgwick (1874) thought that the utilitarian principle was a basic moral fact. For him, a basic moral fact is a universal principle, then, and he held that such principles are self-evident, meaning that one has only to consider them in order to recognize their truth. W. D. Ross held that what we know first is that a certain feature counts in favor of (or is a reason for doing) a particular action, and that we work from this to the recognition of a general "prima facie" duty by a process of "intuitive induction." In Ross's view, then, what we are capable of directly recognizing is that one thing is a reason for another, and that is not a matter of drawing inferences. This seems to commit him to the view that we know these things by intuition (though he never explicitly said as much). The question is, What else could one say about our ability to recognize reasons?

See also Metaethics; Moore, G. E.; Objectivity in Ethics; Ross, William David.

Bibliography

Dancy, Jonathan. "From Intuitionism to Emotivism." In *The Cambridge History of Philosophy, 1870–1945*, edited by Thomas Baldwin. Cambridge, U.K.: Cambridge University Press, 2003.

Ewing, A. C. *Ethics*. London: English Universities Press, 1953.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Prichard, H. A. *Moral Writings*. Oxford: Clarendon Press, 2002.

Ross, W. D. *The Foundations of Ethics*. Oxford: Clarendon Press, 1939.

Ross, W. D. *The Right and the Good*. Oxford: Clarendon Press, 1930.

Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Sidgwick, Henry. *The Methods of Ethics*. London: Macmillan, 1874.

Jonathan Dancy (2005)

INTUITIONISM AND INTUITIONISTIC LOGIC

Logic, in the modern preponderantly mathematical sense, deals with concepts like truth and consequence. The main task of logic is to discover the properties of these concepts. Ever since Aristotle it had been assumed that there is one ultimate logic for the case of descriptive statements, which lent logic a sort of immutable, eternal appearance. Only in the beginning of the twentieth century were certain principles of traditional logic submitted to a critical revision. It was L. E. J. Brouwer, who, in a radical constructive framework of mathematics, discovered that traditional logic could not be upheld in its full extent.

This entry sketches the basic ideas of Brouwer's constructivism, which goes by the name of intuitionism, and then discusses the fundamental principles. Next, an exposition of the familiar notions, such as proof system, semantics, and the like, is provided. In particular, this entry will show how the Brouwerian mathematical universe takes a special place in terms of its logical properties.

INTUITIONISTIC TRUTH

For all practical purposes it suffices to consider in mathematical logic only a few logical constants, or connectives. The traditional conjunction (*and*, $\wedge$), disjunction (*or*, $\vee$, and *not*, $\neg$), and implication (*if* …, *then* … , $\rightarrow$) will do for propositional logic. By adding the two quantifiers, the universal quantifier (*for all*, $\forall$) and the existential quantifier (*there is*, $\exists$), one gets the required connectives for predicate logic. The notion of intuitionistic truth is laid down for composite statements by an inductive procedure. Hence, one has to start by explaining the truth of the basic statements, the so-called atomic ones.

In intuitionism the objects and relations of mathematics are mental creations of the individual, called sub-

ject here. And logic has its domain of action in the self-constructed universe. The fundamental construction is that of the natural numbers. The move of time causes the subject to experience a sensation, which is succeeded by another sensation, while the first one is stored in memory. Brouwer called this the ur-intuition. Through a process of abstraction (identification of experienced two-ities), the empty two-ity, that is, the number 2 is obtained. A process of iteration then yields the natural numbers, and reflection on the process yields the principle of complete induction. Now, one can establish simple identities like $2 + 3 = 5$. This is done by a construction of the following kind: (1) construct 2, (2) construct 3, (3) construct the sum of 2 and 3, (4) construct 5, (5) carry out the comparison construction for the outcome of (3) and (4). The success of (5) tells one that "$2 + 3 = 5$." The basic statements are thus verified, or proved, by constructions. One can denote "the construction a proves φ" by "$a{:}\varphi$." It is easy to see how a conjunction is proved; one has a proof of $\varphi\wedge\psi$ if one has a proof of φ and a proof of ψ. So the proof of $\varphi\wedge\psi$ is an ordered pair (a, b), such that $a{:}\varphi$ and $b{:}\psi$.

The disjunction is slightly more problematic; $\varphi\vee\psi$ is a statement that contains two possibilities φ and ψ, and at least one of these is to be the case. Constructively, one wants therefore to select one of the two and provide a proof for it. We put this as $(a, b){:}\varphi\vee\psi$ if $a = 0$ and $b{:}\varphi$, or $a = 1$ and $b{:}\psi$. So a distinguishes between φ and ψ and b is the required proof. The classical notion of disjunction (the word *classical* is used here for "nonintuitionistic") is more liberal, something like "it is φ or ψ but I don't know which." This notion is not constructively acceptable. The implication is by far the most interesting; p is a proof of $\varphi\to\psi$ means that p is an operation that carries proofs into proofs such that for a proof a of φ, $p(a)$ is a proof of ψ. In symbols: $p : \varphi \to \psi \Leftrightarrow$ for any a $a{:}\varphi \Rightarrow p(\mathrm{a}){:}\psi$. This notion is in fact the natural interpretation of implication. The interpretation of negation has also to be rendered constructively, that is, we have to explain what $a{:}\neg\varphi$ means, a mere "not" will not do. Brouwer's solution was: $\neg\varphi$ is true if φ can be reduced to a contradiction (absurdity); in terms of constructions: $p{:}\neg\varphi$ if for any proof $a{:}\varphi$, $p(a){:}\bot$, where $\bot$ stands for a contradiction. That is, $\bot$ is a statement that has no proof. As a consequence, φ cannot have a proof. In Brouwer's conception the paradigm of a contradiction is $0 = 1$.

For the quantifiers, one has the following proof clauses, where one considers some fixed domains D of objects d:

$$a{:}\forall x\ \varphi(x) \text{ if for each } d \in D\ a(d){:}\varphi(d).$$

$$(a, b){:}\exists x\ \varphi(x) \text{ if } b{:}\varphi(a).$$

The constructive character of existence is thus made explicit in the sense that the proof contains a witness a and a proof of $\varphi(a)$.

The above proof interpretation defines *truth* as "having a proof," where a proof is a construction of an open-ended, unspecified nature. Hence, there is no preferred class of constructions specified—one recognizes a construction when one sees one. The precise form was published by Arend Heyting in 1934, and a similar interpretation was provided by Kolmogorov in his "problem-interpretation" (1932). In an informal manner the proof interpretation was conceived and used by Brouwer. He demonstrated the use of the interpretation, for example, in his proof of the historically first theorem of intuitionistic logic: $\neg\varphi \leftrightarrow \neg\neg\neg\varphi$. On the proof interpretation the principle of the excluded middle (PEM), $\varphi\vee\neg\varphi$, is not always provable (i.e., true). For it would demand that one can give for each proposition φ in advance either a proof of φ, or a construction that reduces any proof of φ into a proof of $\bot$. This is not a matter of abstract logic, but a matter of a general superalgorithm that decides any φ, a so-called omniscience principle. Hence, on the proof interpretation, PEM is equivalent to Hilbert's dogma, which states that all mathematical problems are solvable (1900). However, on the same interpretation, there are no absolutely unsolvable problems, for if $\varphi\vee\neg\varphi$ has no proof, φ cannot have a proof, and this is equivalent to "$\neg\varphi$ has a proof." So in this respect, David Hilbert was right, "there is no ignorabimus."

FORMAL SYSTEMS

The first formalization of (a fragment of) intuitionistic logic was presented by Kolmogorov (1925), and Heyting followed in 1930 with a full system. Both used Hilbert-type formalizations. The axioms were chosen in such a way that the classical system is obtained by adding just one extra axiom (schema): the principle of the excluded middle. The following is an axiom system:

$\varphi\to(\psi\to\varphi)$	$(\varphi\to\psi)\to((\varphi\to(\psi\to\sigma))\to(\varphi\to\sigma))$
$\varphi\to(\psi\to(\varphi\wedge\psi))$	$(\varphi\to\sigma)\to((\psi\to\sigma)\to((\varphi\vee\psi)\to\sigma))$
$(\varphi\wedge\psi)\to\varphi$	$\varphi\to(\varphi\vee\psi)$
$(\varphi\wedge\psi)\to\psi$	$\psi\to(\varphi\vee\psi)$
$(\varphi\to\psi)\to((\varphi\to\neg\psi)\to\neg\varphi)$	$\neg\varphi\to(\varphi\to\psi)$
$\varphi(t)\to\exists\ x\varphi(x)$	$\forall x\varphi(x)\to\varphi(t)$

There are three derivation rules:

From φ and $\varphi\rightarrow\psi$ derive ψ, in symbols $\varphi, \varphi\rightarrow\psi / \psi$ (modus ponens); $\sigma\rightarrow\varphi(x) / \sigma\rightarrow\forall x\varphi(x)$; $\varphi(x)\rightarrow\sigma / \exists x\varphi(x)\rightarrow\sigma$.

An alternative formalization was presented by Gerhard Gentzen (1909–1945); his system of natural deduction differed from the Hilbert-type formalization in having no axioms and only rules. The rules are particularly perspicuous, in the sense that they tell one how to introduce a connective and how to eliminate it. In other words, what one has to assume to conclude a composite statement, and what are the consequences of a composite statement. For example, from φ and ψ infer $\varphi\wedge\psi$; from $\varphi\wedge\psi$ infer φ and likewise ψ. In Gentzen's notation:

$$\frac{\varphi\ \psi}{\varphi\wedge\psi} \qquad \frac{\varphi\wedge\psi}{\varphi} \qquad \frac{\varphi\wedge\psi}{\psi}$$

Thus, the rules appear in the form of *introduction rules* and *elimination rules*. Some rules are purely local in the sense that they only concern the immediate premises and the consequence, and some are more complicated in that they concern the whole history of the derivation. The remaining rules can be found in the literature.

The Gentzen rules can be viewed as a means of providing the meaning of the connectives. This was the underlying idea of Per Martin-Löf's (1984) type theory, and it also played a role in Michael A. E. Dummett's (2000) meaning theory.

A question of immediate interest concerns the relation between intuitionistic and classical logic. The first is a subsystem of the second, that is, every intuitionistic theorem is a classical one. There is also a natural way of interpreting classical logic in intuitionistic logic. Kolmogorov (1925), Kurt Gödel (1934) and Gentzen (1934) translated classical logic so that derivability was preserved. If one denotes classical derivability by $\vdash_c$ and intuitionistic derivability by $\vdash$, then $\Gamma \vdash_c \varphi \Leftrightarrow \Gamma^o \vdash \varphi^o$ where φ^o is the translation of φ. One can see here the Gödel translation:

$\varphi^o = \neg\neg\varphi$ for atomic φ	$(\varphi\wedge\psi)^o = \varphi^o\wedge\psi^o$
$(\varphi\vee\psi)^o = \neg\varphi^o\wedge\neg\psi^o$	$(\varphi\rightarrow\psi)^o = \varphi^o\rightarrow\psi^o$
$(\forall x\varphi(x))^o = \forall x\varphi^o(x)$	$(\exists x\varphi(x))^o = \neg\forall x\neg\varphi^o(x)$

The translation clearly indicates from an intuitionistic point of view a weakening of the meaning of the connectives. The translation also works for concrete theories, for example, arithmetic. Let **PA** be the standard formalization of Peano's arithmetic, and HA the same for Heyting's arithmetic, where **PA** = **HA** + PEM. Now **PA** $\vdash_c$ A $\Leftrightarrow$ **HA** $\vdash$ A^o. This immediately shows that **PA** and **HA** are equally consistent. Historically, this came as a surprise. **HA** was supposed to be more consistent on the ground of its constructive nature. Translations, such as Gödel's, can also be carried out for certain higher order systems.

Gödel (1934) also observed that the extra strength of intuitionistic logic (from a constructive point of view) had an implicit modal character. He translated propositional logic into a modal logic, where $\Box\varphi$ had the heuristic meaning "I know φ" or "I have established φ." The translation is given by:

$\varphi^m = \Box\varphi$ for atomic φ	$(\varphi\wedge\psi)^m = \varphi^m \wedge \psi^m$
$(\varphi \vee \psi)^m = \neg(\neg\varphi^m \wedge \neg\psi^m)$	$(\varphi \rightarrow \psi)^m = \Box(\varphi^m \rightarrow \psi^m)$

Intuitionistic logic thus translates into modal logic, and derivability is preserved in the sense that $\vdash \varphi \Rightarrow$ *S4* $\vdash$ $\varphi\psi)^m$. Gödel's idea was recently adopted and vigorously extended by Artemov (2002), who combined the proof interpretation and the modal interpretation for arithmetic and extensions.

There are many classically derivable propositions that are not derivable in intuitionistic logic, the best known being PEM and the double negation principle $\neg\neg\varphi\rightarrow\varphi$. The latter obviously follows from the first, but conversely the schema $\neg\neg\varphi\rightarrow\varphi$ implies the schema $\varphi\vee\neg\varphi$ (Bernays). One simply applies the double negation principle to $\varphi\vee\neg\varphi$ (and uses some intuitionistic logic).

It is convenient to remember the following rule: $\exists$ and $\vee$ are strong connectives. So, as to be expected, for example, $\neg(\varphi\wedge\psi)\rightarrow(\neg\varphi\vee\neg\psi)$, $(\varphi\rightarrow\psi)\rightarrow(\neg\varphi\vee\psi)$, $\neg\forall x\varphi(x)\rightarrow\exists x\neg\varphi(x)$, $\forall x(\varphi\vee\psi(x))\rightarrow(\varphi\vee\forall x\psi(x))$ are not derivable.

In an obvious way intuitionistic logic is weaker than classical logic, that is to say, as a subsystem. However, it is stronger on the ground of certain metalogical properties. For example, if $\vdash\varphi\vee\psi$, then $\vdash\varphi$ or $\vdash\psi$ (disjunction property). This testifies to the strength of the disjunction in intuitionistic logic. Similarly, one has $\vdash\exists x\varphi(x) \Rightarrow \vdash\varphi(c)$ for a constant c (in a language without function symbols) (existence property).

SEMANTICS

Where classical logic basically has one natural semantics, the truth table semantics of true, false, intuitionistic logic has many semantics, each with its specific features. The intended intuitionistic semantics is the proof interpretation, but for model theoretic applications it is rather too open ended. There are semantic interpretations of intuitionistic logic that have the flexibility one needs for

uncovering logical properties. There are roughly two classes of semantics: those based on (classical) set theory and those of an algorithmic nature. The set-theoretical ones will be considered first.

THE TOPOLOGICAL INTERPRETATION (TARSKI). In classical logic one can interpret propositional logic by means of Venn diagrams (or rather interpret monadic logic), that is, one assigns subsets of a set to propositions and interprets $\wedge$, $\vee$, $\neg$ as $\cap$, $\cup$, ^{c}. The topological interpretation refines this by considering open sets in a topological space and reinterpreting $\neg$ by the interior of the complement (where the interior of a set is the largest open subset of that set). More precisely, given a topological space X and a domain D, assign an open set $[\![\varphi]\!]$ in X to each atom φ (and $\emptyset$ to $\bot$) then

$[\![\varphi\wedge\psi]\!] = [\![\varphi]\!] \cap [\![\psi]\!]$, $\quad [\![\varphi\vee\psi]\!] = [\![\varphi]\!] \cup [\![\psi]\!]$
$[\![\varphi\rightarrow\psi]\!] = \mathrm{Int}([\![\varphi]\!]^c \cup [\![\psi]\!])$, $\quad [\![\neg\varphi]\!] = \mathrm{Int}([\![\varphi]\!]^c)$
$[\![\exists x\varphi(x)]\!] = \cup\{[\![\varphi(d)]\!] \mid d\in D\}$, $\quad [\![\forall x\varphi(x)]\!] = \mathrm{Int}(\cap\{[\![\varphi(d)]\!] \mid d\in D\})$

Under this interpretation one obtains completeness:

$\vdash\varphi \Leftrightarrow [\![\varphi]\!] = X$ for all topological spaces X.

Brouwer's theorem, $\neg\neg\neg\varphi\rightarrow\neg\varphi$, is topologically explained by the fact that the interior of the interior is the interior (also "the closure of the closure if the closure"). One can easily see that $\varphi\vee\neg\varphi$ and $\neg\neg\varphi\rightarrow\varphi$ are not valid, and hence not derivable: consider the real line R and put $\varphi = R-\{0\}$, then $[\![\neg\varphi]\!] = \emptyset$ and $[\![\neg\neg\varphi]\!] = R$.

HEYTING VALUED INTERPRETATIONS (TARSKI AND MCKINSEY). Classical propositional logic, from an algebraic point of view, is a Boolean algebra if one identifies provably equivalent propositions. A similar fact holds for intuitionistic logic, only the algebra fails to be Boolean, for in general $\varphi\leftrightarrow\neg\neg\varphi$ is not valid. The algebras for intuitionistic logic are called Heyting algebras. Since Heyting (and Boolean) algebras have a partial ordering (given by $[\![\varphi]\!] \leq [\![\psi]\!] \Rightarrow \vdash\varphi\rightarrow\psi$) with suitable properties, they turn out to be lattices. The algebraic and lattice theoretic approach are closely related. Jaskowski (1936), in his pioneering work, constructed certain lattices that in a specific sense capture intuitionistic propositional logic. The applications of Heyting algebras and Heyting valued models can be found in Peter T. Johnstone's book, *Stone Spaces* (1982).

BETH AND KRIPKE MODELS. Where the previous interpretations are more technical than foundational in character, Beth and Saul Kripke formulated in the 1950s and 1960s interpretations that intended to do justice to Brouwer's philosophical motivations, in particular his conception of mathematics, and thus logic, as a creation of the subject. Beth introduced his models in 1956 and Kripke in 1963. Kripke showed how the two interpretations are in a precise sense equivalent. The underlying idea is to take the notions "*the subject knows* (or 'has established') φ *at time t*" seriously. So Beth and Kripke consider stages of research (or knowledge) in time for the subject. These states are partially ordered, as the subject has at each stage a number of options of how to pursue his or her research further. The subject is assumed to have perfect memory, so what is known at a certain stage is preserved in future stages. The subject directly becomes aware of and establishes basic facts, so at each stage k a number of atomic statements are accepted; one can denote "φ is known at stage k" by $k \Vdash \varphi$, and we can say "φ is forced at stage k." The interpretation for composite statements is defined inductively. Some decisions are made on the spot, for example, $\varphi \wedge \psi$ is forced at k if both φ and ψ are forced at k; some, however, involve the future. The paradigmatic case is the implication: $\varphi \rightarrow \psi$ may be accepted as correct although no information is available about the correctness of φ and ψ. Here is a Brouwerian example: Let φ stand for "there are 100 consecutive zeros in the decimal expansion of π," and ψ for "there are 99 such decimals." Obviously $\varphi \rightarrow \psi$ is correct, even for an intuitionist. So what does $k \Vdash \varphi\rightarrow\psi$ mean? The natural solution is: If at a future stage k' one has $k' \Vdash \varphi$, then $k' \Vdash \psi$. In fact, this is the natural interpretation of implication, which is only obscured by the truth table definition.

The universal quantifier also involves the future. The subject not only establishes atomic facts but also constructs objects of the domain. So at a later stage, in general, more objects will be available (exist) than at an earlier one. So the universal quantifier has also to take future elements into account; one can denote the domain at stage k by $D(k)$. One can now give Kripke's definition of $\Vdash$:

$$k \Vdash \varphi\wedge\psi \Leftrightarrow k \Vdash \varphi \text{ and } k \Vdash\psi;\ k \Vdash \varphi\vee\psi \Leftrightarrow k \Vdash \varphi \text{ or } k \Vdash \psi;$$
$$k \Vdash \varphi \rightarrow \psi \Leftrightarrow \forall\, k'\geq k(k' \Vdash\varphi \Rightarrow k' \Vdash\psi);$$
$$k \Vdash\exists x\varphi(x)\Leftrightarrow\exists d\in D(k)(k \Vdash\varphi(d));$$
$$k \Vdash \forall x\varphi(x)\Leftrightarrow\forall\, k'\geq k\forall d\in D(k')(\, k' \Vdash \varphi(d)).$$

Furthermore $\bot$ is never forced. So for negation one has $k \Vdash\neg\varphi \Leftrightarrow \forall\, k'\geq k$ not $k' \Vdash\varphi$. This semantics is sound, and even complete for intuitionistic logic. A particular partially ordered set with prescribed forcing for atoms, and a given domain assignment $D(k)$ for all k, is called a Kripke model. It is a simple exercise to construct coun-

termodels for nonderivable statements. For example, the model with two stages, k_o and k_1, such that $k_o < k_1$ and k_o forces no atoms, and

$k_1 \Vdash \varphi$, has the property that not $k_o \Vdash \varphi \vee \neg\varphi$.

The completeness theorem (Kripke 1963) states that

$\vdash\varphi \Leftrightarrow$ for all Kripke models K and all stages $k \in K$ $k \Vdash \varphi$.

For propositional intuitionistic logic it even suffices to consider only finite Kripke models. In that case the so-called finite model property holds: if not $\vdash \varphi$, then there is a finite Kripke model that refutes φ.

Kripke models have been extensively used in metalogical research. Unfortunately, the proofs of the completeness theorem use a classical metatheory and are based on PEM. Hence, not all the applications of Kripke semantics are intuitionistically correct. W. Veldman (1976) generalized the notion of the Kripke model, so that the use of PEM can be avoided.

Beth models are similar to Kripke models, be it that the clauses rely strongly on the future. For example, $\varphi \vee \psi$ is forced at stage k if any (infinite) sequence of future stages $k \leq k_o \leq k_1 \leq \ldots$ eventually leads to a stage in which φ or ψ is forced.

Beth models are less flexible than Kripke models, but they have their advantages for metalogical applications and for modeling second-order theories.

ALGORITHMIC INTERPRETATIONS. In 1945 Stephen Cole Kleene introduced an algorithmic interpretation of intuitionistic arithmetic called realizability. One may think of the constructions in the proof interpretation as partial recursive functions, so that the proof evidence can be given as the index n of such a function. We can say that "n realizes φ" (n **r** φ). For closed atoms the correctness can immediately be verified, so one defines n **r** $t_1 = t_2$ if $t_1 = t_2$ is true. The interesting case is, again, the implication: n **r** $\varphi{\rightarrow}\sigma$ if $\forall m(m \mathbf{r} \varphi \Rightarrow \{n\}m \mathbf{r} \sigma)$, where $\{n\}m$ is assumed to converge. The class of all realizable sentences is a theory that extends **HA**; it is axiomatized by some natural axioms. The interesting and surprising fact is that Church's thesis, CT_0 $\forall x \exists y \varphi(x,y) \rightarrow \exists e \forall x\, \varphi(x\{e\}x)$, is realized. Church's thesis claims that all algorithms (on natural numbers) are recursive functions. Hence, in this extended arithmetic (which is inconsistent with PA) all functions are given by Turing machines.

Gödel introduced an interpretation of arithmetic by means of (effective) functionals of higher types, the so-called Dialectica interpretation (1958). Another such interpretation was introduced by Georg Kreisel (b. 1923): modified realizability. For a survey of the previously discussed text and more, see the bibliography.

SECOND-ORDER SYSTEMS

On Brouwer's view, there are two basic kind of objects, given by the first and second act of intuitionism. The first act gives one the discrete objects, and the second infinite sequences. These sequences, say of natural numbers, need not be given by a law; they are chosen more or less arbitrarily by the subject, hence the name choice sequences. The codification takes place in second-order logic with variables for natural numbers $x, y, z, \ldots$ and variables for infinite sequences of natural numbers, $\alpha, \beta, \gamma, \ldots$ (see Troelstra and Dalen 1988). There are a number of systems treating this intuitionistic analysis. Besides the obvious axioms they usually contain an axiom of countable choice, $\forall x \exists y \varphi(x,y) \rightarrow \exists \alpha \forall x \varphi(x, \alpha(x))$, and continuity axioms of various strengths. The motivation for the continuity principle that comes to mind first runs as follows: suppose one has a function F from choice sequences to natural numbers; since the output has to be computed in a finite time, only a finite part of the input choice function α can be used. Hence F is a continuous function from Baire space to N. This argument is an oversimplification (see Atten and Dalen 2002). A formulation of the (weak) principle of continuity, WC, runs as follows: $\forall \alpha \exists x \varphi(\alpha,x) \rightarrow \forall \alpha \exists xy \forall \beta\, [\alpha(0)=\beta(0) \wedge \ldots \wedge \alpha(y-1) = \beta(y-1) \rightarrow \varphi(\beta,x)]$. For functions F one gets the $\forall\exists!$ prefix.

This continuity principle evidently fails for the following $\varphi(\alpha,x)$: $[\forall z(\alpha(z)=0 \wedge x=0) \vee \exists z(\alpha(z)>0 \wedge x=1)]$. Hence, intuitionistic systems with continuity principles are inconsistent with PEM. Furthermore, Brouwer introduced a certain principle of transfinite induction, the principle of bar induction, which lends the system a considerable proof theoretic strength (enough to prove consistency of **HA** and **PA**). A corollary is the so-called fan theorem, which basically says that the intuitionistic Cantor space is compact. A consequence of the fan theorem is the following: "every real function on a closed interval is uniformly continuous" (Brouwer 1975–1976); Brouwer also showed that the continuum is indecomposable, that is, R cannot be split into two nonempty parts.

In the 1920s Brouwer started to exploit the idealized choice activity of the subject. This was formalized and further analyzed by Kreisel, Kripke, and John Myhill in the 1960s. Kreisel laid down axioms for the theory of the creating subject in a kind of modal formulation. Let $\Box_n \varphi$ stand for "the subject knows (has evidence for) φ at time n." The axioms are $\Box_n \varphi \rightarrow \Box_{n+m} \varphi$; $\Box_n \varphi \vee \neg\Box_n \varphi$; $\exists x \Box_x \varphi \leftrightarrow \varphi$.

Kripke simplified the theory by condensing the total action of the creating subject in one schema, $\exists\alpha(\varphi\leftrightarrow\exists(\alpha(x) \neq 0))$, Kripke's schema (KS). This schema allowed a formal reconstruction of Brouwer's later results. A recent application of KS showed that, for example, the set of irrationals is indecomposable.

Of course, there is also a second-order arithmetic with set variables, HAS. In general, sets are less palpable then functions. In constructive mathematics one cannot reduce sets to characteristic functions, since only decidable sets have characteristic functions.

METAMATHEMATICAL ASPECTS

There are striking similarities between function spaces and logic, which were first observed by Haskell Curry (1900–1982), who indicated the relation between certain combinators and propositional axioms. In the Curry-Howard isomorphism, this is extended to full logic. The present typed lambda calculus is based on these ideas. Henk Barendregt (b. 1947) systematized the various logical- and typed λ-calculus systems and arranged them in Barendregt's cube. Martin-Löf had already in the early 1970s introduced type systems that treat in a uniform way intuitionistic logic and higher-order type systems. These systems yield a most perspicuous presentation of mathematics in a constructive setting and also allow a thorough proof theoretical analysis. It may be remarked that in Martin-Löf's systems (suitable forms of) the axiom of choice became provable, thus confirming the strength of a constructive approach to mathematics. The type theoretic approach to mathematics has born fruit in computer science; there are, for example, implementations of type theory that allow automatic theorem proving. Nicolaas Govert de Bruijn (b. 1918) was the first to develop modern type systems in his AUTOMATH system.

The various semantics that were introduced over the years have found a generalized formulation in category theory. There is a special class of categories, called toposes, that in a most flexible way yields the existing older semantics. A topos can be viewed as a higher-order intuitionistic universe. Hyland (1982) introduced the effective topos, in which Church's thesis holds. Thus nowadays there are many models of all kinds of constructive universes around that allow metamathematical analysis. In particular, there are generalizations of the topological interpretation, the so-called sheaf interpretations, that perfectly model intuitionistic analysis and topology.

One particular line of research by McCarthy shows that in a universe with Church's thesis, intuitionistic truth (for predicate logic) is not arithmetical (hence, the logic is incomplete), and moreover that HA has up to isomorphism, only the standard models. This extends earlier work by Kreisel and refines insights into incompleteness of intuitionistic logic.

A great deal of metamathematical research deals with basic principles, such as transfinite inductions, axioms of choice, and continuity principles. Some of these principles are far from neutral with respect to logic, for example, the full axiom of choice implies PEM.

See also Brouwer, Luitzen Egbertus Jan; Logic, Non-Classical; Mathematics, Foundations of.

Bibliography

Atten, Mark van. *On Brouwer*. Belmont, CA: Wadsworth, 2004.

Atten, Mark van, and Dirk van Dalen. "Arguments for the Continuity Principle." *Bulletin of Symbolic Logic* 8 (3) (2002): 329–347.

Brouwer, L. E. J. *Collected Works*. Vol 2, edited by A. Heyting. Amsterdam, Netherlands: North-Holland, 1975.

Buss, Samuel R. *Handbook of Proof Theory*. New York: Elsevier, 1998.

Dalen, Dirk van. "From Brouwerian Counter Examples to the Creating Subject." *Studia Logica* 62 (2) (1999): 305–314.

Dalen, Dirk van. "Intuitionistic Logic." In *Handbook of Philosophical Logic*. Vol. 5. 2nd ed., edited by D. M. Gabbay and F. Guenthner, 1–114. Dordrecht, Netherlands: Kluwer Academic, 2002.

Dalen, Dirk van. *Logic and Structure*. 4th ed. Berlin. Springer-Verlag, 2004.

Dalen, Dirk van. *Mystic, Geometer, and Intuitionist: The Life of L. E. J. Brouwer*. Vols. 1 and 2. New York: Oxford University Press, 1999, 2005.

Dalen, Dirk van, and M. van Atten. "Intuitionism." In *A Companion to Philosophical Logic*, edited by Dale Jacquette, 513–530. Malden, MA: Blackwell, 2002.

Dragalin, A. G. *Mathematical Intuitionism: Introduction to Proof Theory*. Providence, RI: American Mathematical Society, 1988.

Dummett, Michael A. E. *Elements of Intuitionism*. 2nd ed. New York: Oxford University Press, 2000.

Heyting, A. *Mathematische Grundlagenforschung: Intuitionismus, Beweistheorie* Berlin: J. Springer, 1934.

Johnstone, Peter T. *Stone Spaces*. New York: Cambridge University Press, 1982.

Mancosu, Paolo. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. New York: Oxford University Press, 1998.

Martin-Löf, Per. *Intuitionistic Type Theory*. Naples, Italy: Bibliopolis, 1984.

Mints, Grigori. *A Short Introduction to Intuitionistic Logic*. New York: Kluwer Academic, 2000.

Rasiowa, Helena. *The Mathematics of Metamathematics*. Warsaw, Poland: Pánstowe Wydawnictwo Naukowe, 1963.

Troelstra, A. S., and D. van Dalen. *Constructivism in Mathematics: An Introduction*. 2 vols. Amsterdam, Netherlands: North-Holland, 1988.

Troelstra, A. S., and H. Schwichtenberg. *Basic Proof Theory*. New York: Cambridge University Press, 1996.

Van Heijenoort, Jean. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press, 1967.

Dirk van Dalen (2005)

IONESCU, NAE

(1890–1940)

Nae Ionescu, the Romanian logician, metaphysician, and religious philosopher, studied at the University of Bucharest and received his doctorate from the University of Munich in 1919 with the thesis *Die Logistik als Versuch einer neuen Begründung der Mathematik*. From 1920 on he was professor of logic, history of logic, and metaphysics at the University of Bucharest. He was also the editor in chief (1924–1928) and director (1928–1934) of the newspaper *Cuvântul*, in which he published more than 1,000 articles on religious, political, and economic problems.

Ionescu's scholarly publications were few—some articles on logic, a few prefaces, and a series of articles in the theological journal *Predania* (1937–1938). Nevertheless, his influence from 1922 to 1940 was enormous. His teachings and writings inspired a new interest in metaphysics and religious philosophy in Romania. Although he was primarily a logician, he strove to understand all forms of human activity. According to Ionescu, the philosopher must take into consideration not only the theoretical expression of historical life—from religion to logic and science—but also its meaningful creations: crafts, arts, biographies, political events, and all others. He approached the history of logic, as well as the history of metaphysics and of religion, as a typology of the human spirit. Such a typology he regarded as always a creation of history and ultimately of life. This seems to imply a radical historicization of the mind's activities, but God, for Nae Ionescu, is present in history through the Incarnation. On the other hand, man's mode of being is completely fulfilled only through death, and death is above all transcendent.

See also Logic, History of; Philosophy of Religion.

Bibliography

WORKS ON IONESCU

A complete posthumous edition of Ionescu's philosophical works in twelve volumes was announced during World War II, but only three volumes were published: *Istoria Logicei* (Bucharest, 1941) and *Metafisica*, 2 vols. (Bucharest, 1942–1943). New editions of many of Ionescu's works appeared during the 1990s, all edited by Marin Diaconu and published by Humanitas Press, Bucharest, most notably: *Curs de metafizica* (2 vols., 1991 and 1995; Lectures in metaphysics), *Curs de istorie a logicii* (1993; Lectures in the history of logic), and *Curs de logica* (1993; Lectures in logic).

WORKS ON IONESCU

In English

Calinescu, Matei. "The 1927 Generation in Romania: Friendship and Ideological Choices (Mihail Sebastian, Mircea Eliade, Nae Ionescu, Eugene Ionesco, E. M. Cioran)." *East European Politics and Societies* 15 (3) (2001): 649–677.

In German

Müller, Hannah. "Mircea Eliade und Nae Ionescu. Der Schüler und sein Meister." *Zeitschrift für Religionswissenschaft* 12 (1) (2004): 79–98.

In Romanian

Bradatan, Costica. *O Introducere la istoria filosofiei românesti în secolul XX* (An Introduction to the History of Romanian Philosophy in the Twentieth Century). Bucharest: Romanian Cultural Foundation Press, 2000.

Mircea Eliade (1967)
Bibliography updated by Vicki Harper (2005)

IONIAN SCHOOL

See *Pre-Socratic Philosophy*

IQBAL, MUHAMMAD

(1877–1938)

Muhammad Iqbal, an Islamic poet and metaphysician, was born in Sialkot, Pakistan. He studied philosophy at Cambridge for three years under J. M. E. McTaggart and James Ward. He received his Ph.D. from Munich University in 1908 for his thesis *The Development of Metaphysics in Persia.*

Inheriting the classical tradition of Muslim mystic poets, both Persian and Urdu, Iqbal was for a long time an admirer of the Spanish Sufi philosopher Ibn al-Arabi (1165–1240), the most consistent advocate of pantheism among Muslim thinkers. Very soon, however, he realized that this philosophy was foreign to the simple and invigorating message of Islam as embodied in the Qurʾan and as represented in the dynamic life of Muḥammad and his

early followers. Under the influence of Jalāl al-Din Rūmī (1207–1273), the great mystic poet, whose philosophical outlook was allied in several important respects with post-Kantian voluntaristic thought in the West, as represented by Friedrich Nietzsche and Henri Bergson, he evolved a new system of thought that was meant to revitalize the faith of the Muslims of the Indo-Pakistan subcontinent. At first his message, written in verse in the *Secret of the Self* (1915), raised a storm of opposition, but very soon this opposition died its natural death, and the whole subcontinent reverberated with his inspiring melodies. He exerted great influence in molding the pattern of political, social, and intellectual life of the Muslims in the early decades of the twentieth century, an influence that is visible everywhere even now. In 1930, as president of the Muslim League, he proposed the creation of a "Muslim India within India." Pakistan, Iqbal's dream, came into being in 1947, nine years after his death. As a tribute to his memory, the government of Pakistan established in 1951 a statutory body known as Iqbal Academy, in order "to promote the study and understanding of the works of Iqbal."

The system of thought that he evolved may be called theistic pluralism in contradistinction to Ibn al-Arabi's pantheistic doctrine of the unity of being, which denied not only the unique personality of the Divine Being and his existence as distinct from the universe but also the existence of human individuals and their partnership with God in constituting the commonwealth of ends.

Immanuel Kant's negative answer to the possibility of metaphysics provided Iqbal with a basis on which to construct his thought. Human thought, Kant asserted, is circumscribed by the categories of space and time; therefore, the Ultimate Reality, which, by definition, is beyond these categories, cannot be comprehended by pure thought, which is intimately related to and based on the normal level of experience. According to Iqbal, however, time and space are not fixed and unvarying modes, as Kant had thought; their significance may vary with the beings of higher or lower grade, the degree of being determined by greater or lesser psychic powers. Moreover, this normal level is not the only level of knowledge-yielding experience. The level above spatiotemporal experience is revealed by intuition, a form of perception that is allied to ordinary experience in giving objective knowledge but which is quite distinct from it in not being solely dependent upon sense perception; intuitive experience is individual and incommunicable. It is not simple Bergsonian "intellectual sympathy," which implies negation of the perceiver; intuition, according to Iqbal, by bringing the perceiver into contact with the Most Real, has the power to vitally transform his character and to endow him with a new personality, which reveals to him the higher consciousness of his manifold relations with God and the universe. Through his contact with Reality, the individual discovers his uniqueness, his metaphysical status, and the possibility of improvement in that status. The experience of intuition not only serves to confirm his reality and deepen his whole being but also sharpens his will with the creative assurance that the universe is not something to be really seen and known through concepts but rather something to be made and remade by continuous action, by interpreting the intuition of reality as a stimulus to ideal ends and purposes. Conceptual knowledge gives us knowledge of relations, not of reality; it is only through intuition that we can grasp the Real and give a fresh direction to the course of human history.

To Iqbal, ego is the basic reality revealed by intuition as the center of all efforts—a revelation that is vouchsafed not in the barren contemplation of the recluse but in moments of great decision and action, which are expressive of a firm faith in the ultimate purposiveness of the universe. The life of the ego consists in meeting obstruction in its contact with matter and overcoming it. This gives the ego the power to act freely. It is partly determined and partly free, and it reaches fuller freedom by approaching the individual who is most free—God. In other words, the ego is continually moving from a state of lesser freedom to that of greater freedom.

The ego is also immortal. According to Averroes immortality means transindividual eternity of intellect; according to Nietzsche immortality is synonymous with what he calls eternal recurrence, a most "intolerable" conception, as Iqbal put it. Immortality, according to Iqbal, must be individual and personal. He repudiated the pantheistic belief that the self, as a differentiation of the Absolute, will in the end be submerged and lose its identity in the Whole. It was to save man from this fate that Iqbal advocated that immortality is not a gift that every ego will enjoy; rather, it is a hope, an aspiration, depending, of course, upon a particular philosophy of life and a particular ethic that tends to maintain the state of tension in the ego and develop self-reliance, self-respect, self-confidence, self-preservation—even self-assertion, when such a thing is necessary in the interest of life—and the power to stick to the cause of truth, justice, and duty, even in the face of death. Such behavior helps in the integration of the forces of ego, thus hardening it against the forces of disintegration and dissolution. Because the ego, which exists only in the state of tension, is the most valu-

able achievement of man, he should exert all efforts not to revert to a state of relaxation. We are mortal insofar as we keep ourselves fettered to spatialized time; as soon as we rise above it and immerse ourselves in what Bergson called duration, we become timeless. It is possible, Iqbal held, to realize this timelessness even in this life, although it be but for a moment. It is the moral duty of man to keep the state of tension intact by repudiating life-negating philosophies and to attain immortality by his ego-sustaining behavior. It is in this sense that attaining immortality, according to Iqbal, becomes a moral duty.

How is the ego related to the world of matter? Iqbal viewed matter, as did Albert Einstein, as "a system of interrelated events" and the universe as an "organism," as did Alfred North Whitehead. Every atom, however low in the scale of being, is an ego. Mind, with its capacity for self-consciousness, is a higher ego, and body is a combination of subegos. Thus, on this principle the universe is of the nature of life—free, creative, and original. The universe is constantly growing and progressing toward an end—a rationally directed creative life.

How is the Ultimate Ego (God) related to the universe and to the human ego? To the Absolute Self the universe is not a reality confronting him as an "other"; it is only a passing phase of his consciousness, a fleeting moment of its infinite life. Iqbal began with Einstein's view that the universe is finite but boundless and added that it is finite because it is a passing phase of God's extensively infinite consciousness and boundless because the creative power of God is intensively infinite. But the human self is the exception; it is not a mere passing phase in God's consciousness, for it is self-centered and exclusive. It is distinct but not isolated from God. The Ultimate Ego is characterized by the most beautiful names and attributes; he is transcendent and yet immanent, and above all he is a Person who responds to man's inner yearning in "the awful silence of the universe."

See also Absolute, The; Averroes; Bergson, Henri; Einstein, Albert; Ibn al-ʿArabī; Kant, Immanuel; Intuition; McTaggart, John McTaggart Ellis; Metaphysics; Nietzsche, Friedrich; Pantheism; Religious Pluralism; Ward, James; Whitehead, Alfred North.

Bibliography

Most of Iqbal's works are in verse, Persian and Urdu. Only books available in English are mentioned here. *The Secrets of the Self,* translated by R. E. Nicholson (London: Macmillan, 1920); *The Mysteries of Selflessness,* translated by A. J. Arberry (London: J. Murray, 1953); *Tulips of Sinai,* translated by A. J. Arberry (London, 1947); *The Persian Psalms,* translated by A. J. Arberry (Lahore: M. Ashraf, 1948); *New Garden of Mystery,* translated by B. A. Dar (Lahore: Institute of Islamic Culture, 1964); *Pilgrimage of Eternity,* translated into English verse by S. M. Ahmed (Lahore: Institute of Islamic Culture, 1961); *Development of Metaphysics in Persia* (London, 1908); *Reconstruction of Religious Thought in Islam* (Lahore: J. Iqbal, 1951); and *Poems from Iqbal,* translated by V. G. Kiernam (London: Murray, 1955).

For literature on Iqbal, see B. A. Dar, *A Study of Iqbal's Philosophy* (Lahore, 1944) and *Iqbal and Post-Kantian Voluntarism* (Lahore: Bazm-i-Iqbal, 1956); I. H. Enver, *Metaphysics of Iqbal* (Lahore: Muhammad Ashraf, 1944); K. G. Saiyadain, *Iqbal's Educational Philosophy* (Lahore: Arafat, 1938); and S. A. Vahid, *Iqbal, His Art and Thought* (Lahore: Muhammad Ashraf, 1944).

B. A. Dar (1967)

IRIGARAY, LUCE

(1930–)

Luce Irigaray is a Belgian-born French feminist philosopher whose work draws on her multiple doctorates in the areas of linguistics, philosophy, and psychoanalysis. Her main contributions are her concept of sexual difference and the methodology she developed for a feminist interpretation of the history of philosophy. Like many feminist philosophers, Irigaray argues that women have always been defined in relation to men. She would agree with the mid-twentieth century French feminist philosopher Simone de Beauvoir, who argued that "the relation between the sexes is not quite like that of two electrical poles, for man represents both the positive and the neutral … whereas woman represents only the negative, defined by limiting criteria." Irigaray agrees that the feminine tends to be described "in terms of deficiency or atrophy, as the other side of the sex that alone holds a monopoly on value: the male sex" (1985b, p. 69). Irigaray demonstrates this idea with examples from literature, philosophy, everyday life, and economic and social history.

Irigaray stresses that texts from the history of philosophy have been inconsistent in their discussions of women. They have included conflicting, often overlooked hypotheses about the sexes. These may be explicit contradictions in canonical literature or implicit alternatives. In this sense, the most sex-biased historical text may be a rich resource for a feminist rereading. Rather than dismissing sex-biased caricatures of women as false or irrelevant, Irigaray recommends critiques of their their incoherence as part of a project of imaginative literary elaboration. She therefore argues that feminists should not forgo the close study of historical texts about the

sexes, particularly those of the history of philosophy, and her interpretations have focused on such figures as Plato, Aristotle, Plotinus, Hegel, Nietzsche, Heidegger, Sartre, Levinas, Marx, Lévi-Strauss, Freud, and Lacan.

Irigaray's methodology involves extensive citation, parody, and whimsical or ironic diagnoses of what a thinker does not "want" to say about women. Such diagnoses are often accompanied by writing experiments in which Irigaray attempts to describe women or write as a woman in ways that she claims would be deemed undesirable by the authors of the texts she is analyzing.

"Sexual difference," as an Irigarayan concept, does not refer to historical depictions of men and women counterposed in terms of such unsatisfactory hierarchies as reason/sensibility, wisdom/ignorance, culture/nature, and public/private. Instead, Irigarayan "sexual difference" refers to a hypothetical, alternative means of envisaging the sexes, according to which they would be considered neither like men nor their opposites or complements, but genuinely different. The concept is not generated through empirical description nor utopian imagination. Instead, it is primarily grounded in Irigaray's notion that such a prospect seems to have been "excluded" historically.

Irigaray has argued, controversially, that equality often means "equal *to*" a default individual (for example, male or white or able-bodied). She has proposed the alternative notion of equivalent rights for men and women and has devised a short bill of "sexuate" rights (1993). Such initiatives embody her view that legal reform should include a concern with the quality of representation of sexual identity. She has also directed collective research on empirical differences in the speech habits of contemporary European men and women, and she has formulated linguistic reforms corresponding to a hypothetical culture that would affirm sexual difference.

See also Continental Philosophy; Feminism and Continental Philosophy; Feminist Philosophy.

Bibliography

WORKS BY LUCE IRIGARAY

Speculum of the Other Woman. Translated by Gillian C. Gill. Ithaca, NY: Cornell University Press, 1985a.

This Sex Which Is Not One. Translated by Catherine Porter. Ithaca, NY: Cornell University Press, 1985b.

Marine Lover of Friedrich Nietzsche. Translated by Gillian C. Gill. Ithaca, NY: Cornell University Press, 1991.

Je, tu, nous: Toward a Culture of Difference. Translated by Alison Martin. New York: Routledge, 1993.

Penelope Deutscher (2005)

IRRATIONALISM

Like other words in current philosophical use, such as *historicism* and *subjectivism, irrationalism* is an exceedingly imprecise term that is employed with a wide variety of meanings and implications. Consequently, any attempt to elucidate its sense within the confines of a clear-cut and tidy formula quickly runs into difficulties. It might be said, for instance, that to describe a writer as an irrationalist is to speak of him as denying the authority of reason. But how is the notion of "reason" itself to be understood, and in what respects is its authority supposed to be flouted? It would scarcely be sufficient to reply that denial of reason consists in illogicality or confusion of thought, or that it manifests itself in a tendency to arrive at unacceptable conclusions; for this would apply to the work of many thinkers to whom the label "irrationalist" is clearly inapplicable. In addition, the suggestion fails to identify the primary point of calling a writer an irrationalist. A man may be accused of irrationality if he is prone to make mistakes of a particular kind or to indulge in invalid reasoning; but it is only insofar as he maintains some specific doctrine concerning such things as the status and role of reason or the relevance of rational standards within various domains of experience or inquiry that he can be called an irrationalist. In other words, attention is focused not on an unwitting failure to conform to norms of generally recognized validity, but on the explicit repudiation, or putting into question, of such norms in the light of certain considerations or in relation to certain contexts.

ENLIGHTENMENT RATIONALISM

A more promising approach to the understanding of irrationalism is the historical. One might try to understand irrationalism by contrasting it with that "belief in reason," that faith in the application of mathematical and scientific procedures, which was so prominent in the thought and speculation of seventeenth-century and eighteenth-century Europe and which provided the inspiration for the Enlightenment. Such a proposal, however, runs the risk of invoking generalities as vague as they are misleading. Seventeenth-century and eighteenth-century theorists interpreted the ideal of rationality in widely differing ways, and they assumed it or sought to realize it at various levels of inquiry—metaphysical, epistemological, ethical, and political. René Descartes, John Locke, Thomas Hobbes, Gottfried Wilhelm Leibniz, Benedict de Spinoza, and David Hume shared the conviction that in their speculations concerning the nature of the world and our knowledge of it they were conforming

to a course acceptable to reason and were applying methods that reason prescribed. But they differed fundamentally concerning both what constituted rational procedure and what types of discovery such procedure was capable of achieving. Similar disagreements may also be discerned at the other levels of investigation mentioned.

NINETEENTH-CENTURY IRRATIONALISM

The diversity of opinion attributable to thinkers who all held a general belief in rationality puts in doubt the notion that irrationalism can be neatly and unambiguously identified by reference to its rejection of a single set of assumptions allegedly shared by philosophers associated with the Enlightenment. Nevertheless, the ideas of those thinkers most typically classified as irrationalists did develop to a large extent in reaction to the ambitious claims made on behalf of reason by Enlightenment theorists and their nineteenth-century successors, however widely such claims may have varied in actual content and formulation. That the world is in some sense a rational or harmonious whole, that the human mind is capable of comprehending it, and that there exist certain communicable and teachable methods by means of which its inner workings can be revealed; that this knowledge can be systematically utilized in a manner that will insure the continuous improvement of human society in the foreseeable future; that man is by nature a reasonable and progressive being whose potentialities can be realized through the removal of ignorance and the creation of institutions based upon principles of justice—it has been against views like these that irrationalist philosophies have, in different ways, characteristically protested. Vociferous insistence upon the limitations and weaknesses of reason followed an equally vociferous insistence upon its possibilities and powers.

ONTOLOGICAL IRRATIONALISM

The belief that reality, at least in its innermost nature, represents an intelligible, ordered system whose fundamental character is accessible to the human intellect, is an ancient one; in philosophy, it dates at least from Plato. During its long history it has admittedly been subjected to a number of widely differing interpretations, ranging from the animistic or religious to the mathematical or scientific. Yet the notion of some kind of comprehensible pattern or rational structure to which all that exists or happens can finally be shown to conform retained its hold. From this point of view the world we belong to is not an alien world; on the contrary, it is one in which, by virtue of our own rationality, we can feel at home.

PASCAL'S SKEPTICISM. There have, however, been thinkers to whom the consoling idea of an intelligible world has seemed less acceptable. Thus even in the seventeenth century, the heyday of Cartesian rationalism, Blaise Pascal was questioning the conception of reality as a logically coherent whole, transparent to human reason and in which everything, including man himself, can be seen to have its necessary place: "Too much clarity darkens," he wrote with reference to Descartes's famous "clear and distinct ideas." Forcibly impressed both by the contingent character of human existence in an unfathomable universe and by the inadequacies of human reason, Pascal had little use for rational theology with its pretended proofs of God; he eschewed all such forms of ratiocination in favor of an inward religious faith that transcended ordinary methods of argument and justification and that was beyond demonstration.

THE WORLD AS WILL. The intense dissatisfaction and disquiet Pascal experienced when he contemplated the world and our situation within it has been echoed in the works of many subsequent writers, although they have not always shared the religious convictions that ultimately sustained him. For some it has appeared necessary simply to acquiesce in the realization that reality, far from representing an intellectually satisfying or morally acceptable system, is in truth devoid of all rational meaning or purpose and that salvation can only be reached through a complete liberation from its trammels. Such an attitude found perhaps its most eloquent and forceful exponent in Arthur Schopenhauer. In Schopenhauer's conception of existence there was an explicit and uncompromising reversal of the traditional approach. He made it his object to show, not that the world is governed according to some beneficent teleological principle or that it is the embodiment of certain fundamental rational categories, but that, on the contrary, what lies at its center is something antithetical to all reason and value, namely, a blind unconscious force or striving he termed "will." It is this that constitutes the metaphysical essence of the world, and not (as G. W. F. Hegel and his followers had taught) Absolute Spirit or Mind manifesting itself according to the inner laws of its own rational development. For Schopenhauer, in fact, all forms of rationalism—metaphysical and scientific alike—involve an illicit projection into the ultimate nature of reality of principles whose actual source and spring is the human intellect alone.

THE DOCTRINE OF ABSURDITY. Schopenhauer's theory rested, in the last analysis, upon a professed knowledge of what "really" lies beneath the phenomenal (and finally illusory) surface of things. Yet there have also been thinkers whose skepticism, although quite as profound as Schopenhauer's, did not derive from claims of this kind but instead took as its point of departure the concrete facts of ordinary experience. Such is the doctrine of *absurdité* in the work of twentieth-century French existentialists like Jean-Paul Sartre and Albert Camus. In some respects Sartre remained firmly within the Cartesian tradition, founding his epistemology upon the conception of man as a thinking consciousness confronted by an external world of unthinking substance. But the world that we are aware of is not, for Sartre, an intrinsically intelligible world whose nature conforms to a determinate logical order and whose existence is guaranteed by a benevolent deity. Sartre's view of material existence is perhaps most succinctly expressed in his first novel, *La nausée*, a book that contains in embryo many of the cardinal themes that later figure in his impressive philosophical treatise *L'être et le néant*. The hero of *La nausée*, Roquentin, is described as experiencing in a peculiarly vivid and horrifying way the brute "contingency" of things, their palpable failure to measure up to the standards of logical rigor and necessity, of clarity and distinctness, that reason of its nature seeks to impose upon or find realized within the world. Roquentin is impressed by the loose and arbitrary character of our modes of classifying objects and by the manner in which existence, in all its rich and pointless superfluity, seems inevitably to elude the network of interpretative concepts and schemes that we try to throw over it. When so perceived, the world can strike us as divested of all significance or value. "The world of explanations and reasons," Roquentin remarks, "is not the world of existence."

The impossibility of trying to reduce experienced reality to a system, whether Cartesian, Hegelian, or some other, had already been accepted by Søren Kierkegaard, who is often regarded as the originator of modern existentialism. But in Sartre's work one is conscious of a more positive and explicit insistence upon the opacity and ultimate unintelligibility of the world and its resistance to the abstract categories of thought. For Kierkegaard there was something eccentric, some element of radical misunderstanding, in the entire project of attempting to explain or justify existence as a whole in rational terms. By contrast, both Sartre in his philosophical works and Camus in *Le mythe de Sisyphe* are plainly sympathetic to those who demand intellectually or morally satisfying systematic accounts of existence; it is felt to be in some sense an imperfection of our condition as human beings in the world that such demands are necessarily incapable of being satisfied. The essence of what they call *absurdité* lies precisely in the contrast between the contingent amorphous character of reality, on the one hand, and the understandable requirements of reason that reality so patently fails to meet, on the other.

EPISTEMOLOGICAL IRRATIONALISM

Irrationalism sometimes finds expression, not in the claim that reality itself is devoid of ultimate senses or purpose, but in the distinguishable idea that the customary or scientific methods by means of which we are accustomed to explore its nature and to which we accord the honorific title of "rational," are inherently defective or suspect. There are clearly close connections between this view and the conceptions of ontological irrationalism. For if the world really is irrational in the ways it is sometimes declared to be, this presumably implies that, at some level at least, it is not amenable to those modes of investigation typically regarded as rational. But some philosophers, while agreeing that such methods are incapable of leading us to any finally acceptable and satisfying explanation of the nature of things, have not supposed themselves to be thereby committed to holding that all comprehension of the desired kind is in principle impossible. They have suggested, in other words, that alternative modes of apprehending and understanding the world, free from the limitations that beset standard procedures, remain open. The object of their strictures has been the distortions inherent in these procedures, rather than the world itself.

THE LIMITS OF RATIONAL INQUIRY. The belief that there exist determinate limits to what we can discover by the resources of ordinary sense and understanding received precise and systematic exposition in the works of Immanuel Kant. To prescribe limits to what rational inquiry can accomplish is not, as such, to impugn such inquiry, and much of the argument in the *Critique of Pure Reason* is, in fact, expressly concerned with establishing and explaining the validity of mathematical and scientific forms of reasoning within the empirical realm. But there were, nevertheless, two strands in Kant's philosophy that led to doctrines far removed in spirit from those Kant himself propounded. One of these was the claim that the fundamental principles in terms of which phenomenal reality is intelligible derive from the human mind and understanding; the other was the claim that there is a "noumenal" realm of things-in-themselves that is necessarily inaccessible to rational investigation.

SUBJECTIVITY OF CRITERIA OF RATIONAL INQUIRY. Kant's description of the means by which phenomenal reality is intelligible gave rise to the suggestion that the criteria of rational judgment and inference we normally accept are not the stable, objectively grounded things we take them to be but are, on the contrary, essentially subjective and even susceptible to change and variation. Thus Johann Gottlieb Fichte, at any rate in his earlier writings, often gave the impression of having thought that the basic principles in terms of which human beings interpret their experience ultimately fall within the sphere of individual choice or commitment; as prerational posits they cannot be themselves subject to rational assessment and must, instead, be evaluated by reference to the needs and demands of human beings conceived as volitional agents in the world. Fichte ended by taking refuge in the notion of an Absolute Spirit or rational ego that transcended all particular human selves.

Other nineteenth-century thinkers, however, reinterpreted Fichte's initial postulates in a fashion that implied a definite skepticism regarding the claims of rationality. This was true above all of Friedrich Nietzsche, who—at least in certain aspects of his complex and not always consistent thinking—exhibited a profound suspicion of accredited concepts and procedures. Possibly more sensitive than any previous philosopher to the emotional drives and attitudes that operate beneath the surface of human life and unconsciously influence thought and behavior, he was at times prepared to speak as if the entire manner in which we approach the world were founded upon pervasive myths and fictions. The "lies and frauds" that permeate our cherished forms of scientific investigation and description are not devoid of all value; on the contrary, from a "life-furthering, life-preserving, species-preserving" point of view they are actually indispensable. But insofar as we take them to embody or reveal the truth, we are the victims of deception.

NONRATIONAL COGNITION. Although his own confident affirmations concerning the limitations of common sense and science might seem to have required it, Nietzsche did not, in fact, postulate a superior form of cognition capable of circumventing the delusive schemes of ordinary thought and experience and of arriving at some clear, unsullied understanding of the world as it is in itself: in the last analysis there could be no escape from particular interpretations and perspectives. But to other thinkers this has not seemed so evident. Friedrich von Schelling, the contemporary of Fichte and Hegel, evolved an elaborate system in which intuition of a mystical or quasi-religious character was accorded a central place and was held to provide access to the ultimate nature of reality. "The nature of the Absolute *itself*," Schelling wrote, "which as ideal is also immediately real, cannot be known through explanations, but only through intuition" (*Philosophie und Religion*, p. 15). Later Henri Bergson also drew a sharp distinction between the intellect, regarded as having a basically practical function and as rationalizing experience through the construction of mechanistic models and hypotheses, and intuition, whereby an inner sympathetic consciousness of the creative flow that underlies and pervades the universe was attainable.

The division between rational and nonrational or suprarational modes of apprehending the world, which these and other writers have stressed, often merges into further, related contrasts; for example, between conventional perception and artistic perception, between scientific and historical understanding, or between technical know-how, which is communicable in words, and a sense of, or feel for, the inward direction and meaning of things, which is not. Rationalists have tended to point out in return that such contentions are open to serious objections. Emphasis is laid upon the "privacy" of the alleged "insight" or "intuition"; but how can such insight aspire to the status of knowledge if no public criteria are available whereby its findings may be tested or confirmed? Again, in what sense can one speak of knowledge or understanding if—as often seems to be assumed—the intuition is of a kind that precludes conceptualization? Nevertheless, whatever difficulties irrationalist epistemology may present, these have not prevented its adherents from claiming that there are modes of awareness of the deepest significance to which rationalistic theorists have remained perennially blind.

ETHICAL IRRATIONALISM

Questions have also been raised with regard to our claims to moral knowledge and certainty. For instance, a number of writers of an empiricist persuasion (including Rudolf Carnap, A. J. Ayer, and C. L. Stevenson) adopted views concerning the meaning and function of moral judgments that would seem to deny, or at least put in doubt, the possibility of treating these as the proper subjects of rational argument. Yet such writers would certainly reject the suggestion that they are irrationalists in any of the senses so far distinguished. If they owe a historical debt, it is to David Hume (himself a skeptic concerning the rationality of morals) rather than to Continental sources, and they would in any case claim that their theories are grounded upon purely logical considerations related to the analysis of moral concepts and terms rather than

upon alleged discoveries about the nature of the world or the status of human beings within it. Nor would they be likely to admit that what they say entails any dramatic consequences so far as the realm of practical choice and action is concerned; on the contrary, they have tended to contend that their theories, being of a wholly conceptual character, are neutral between particular moral standpoints and outlooks.

ABSENCE OF A MORAL ORDER IN THE WORLD. Not every challenge to the rationality of morals has, however, been characterized by a comparable detachment. One of the strongest motives in recent times for belief that moral convictions are without basis or justification has been precisely the decay of all-encompassing theological and philosophical interpretations of reality; for these were thought of as providing the moral consciousness with the kind of backing it logically required. Along with the religious beliefs to which it was sometimes allied, the conception of a moral order at the heart of existence, either revealing itself directly to the eye of reason or manifesting itself empirically in the course of human life and history, was already in decline during the nineteenth century. Schopenhauer's theory of all-pervasive metaphysical will was directly expressive of this development, but it was Nietzsche, not Schopenhauer, who drew the radical consequences. According to Nietzsche, it was necessary to recognize, once and for all, that there is no moral order, no system of ready-made values, objectively subsisting "out there" in the world—"there are no moral phenomena, only moralistic interpretations of phenomena," he wrote in *Beyond Good and Evil*. The notion of moral facts is a philosopher's delusion. With such ideas in mind Nietzsche, in effect, did two things. First, he embarked upon a devastating analysis intended to show how traditional moral codes, far from resulting from the operations of contemplative reason, derive instead from deep-lying nonrational forces in the human psyche, from motives like resentment and sadism and fear. Second, he urged that it is now possible for us—since, in his famous phrase, "God is dead"—to create new values, more fitted to preserving the dignity of humanity and to realizing those human energies and capacities that still await their true fulfillment.

ALTERNATIVE THEORIES. The claim that it is now possible to determine new values along these lines drew attention to a difficulty that has beset theorists who have denied the possibility of appealing to rational canons within the moral sphere. Nietzsche was a moralist who wished to insist that certain forms of character and behavior were evidently superior to others; at the same time, he was committed to the opinion that, objectively considered, there was nothing to justify preference for one way of life, one system of values, rather than another—*nichts ist wahr, alles ist erlaubt* ("nothing is true, everything permitted"). If traditional Christian morality is without foundation in fact or reason, then so, likewise, is any alternative ethics with which we may seek to replace it.

Similar tensions and ambiguities underlie other varieties of individualist or existentialist teaching, from Max Stirner and Kierkegaard on. Sometimes it seems to be maintained that sheer intensity and sincerity of commitment is all that ultimately counts from a moral point of view. What is chosen is not a matter for argument, since in the last resort there is no yardstick, no privileged set of criteria, against which rival possibilities may be assessed and evaluated. The vital thing is for a man to assert his essential freedom by refusing to conform his will to forces and agencies external to himself, including the falsely substantialized standards of conventional religion and ethics.

Sometimes, on the other hand, an attempt is made to give the notion of an acceptable mode of living more positive content, the implication being that certain forms of behavior are more appropriate to our situation in the world than others. For beings who find themselves in an alien and meaningless world, which is bereft of purpose or value, there may be virtue, or at any rate fittingness, in conduct that reflects the inescapable absurdity of their condition. Suggestions as to how conduct might be said to do so have for the most part been as vague as they have been various. Living in the present or for the moment, giving spontaneous vent to instincts or passions (as opposed to trying to heed the reasonable dictates of conscience or prudence), indulging in anarchical or incongruous behavior for its own sake, undertaking certain types of useless artistic activity—these are among the proposals that may be extracted from works purporting to show what is meant. Such works often seem to be inspired by a curious form of inverted rationalism; the rational response to an irrational world is to act irrationally. Yet it would be incorrect to imply that this is the only consideration that has been used to justify such behavior. Instead, the recommendation appears to be held by some proponents to follow from a realization of what constitutes our true innermost nature as human beings; and this claim introduces a further dimension of irrationalist thought.

PSYCHOLOGICAL AND SOCIAL IRRATIONALISM

The claim that it is not the human situation that is intrinsically absurd, but that human nature itself is in some fundamental sense irrational, is not confined to philosophers of a metaphysical or speculative persuasion; its adherents also include psychologists, political scientists, social theorists, historians, literary artists, and even statesmen. In this area, above all others, a pervasive departure from certain dominant Enlightenment conceptions may be discerned, involving a shift of outlook that has led to drastic changes in the approach adopted by many writers to problems concerning man and society.

It is difficult neatly to summarize the complex and sometimes conflicting ideas involved here. One underlying theme, however, has been that the *idéologues* of the eighteenth century, together with the utilitarians and progressive radicals who followed them in the nineteenth century, grossly exaggerated the extent to which human behavior is motivated, or is capable of being modified, by rational consideration. It has further been suggested that such overvaluation of reason or intellect caused liberal and democratic thinkers to adopt absurdly optimistic, unrealistic, and naive views concerning the capacity of men to improve themselves and the conditions under which they live.

INDIVIDUAL PSYCHOLOGY. At the level of individual psychology it is held to be false that people usually or consistently are activated by calculations regarding their best interests or that they can confidently be expected to respond to considerations of abstract moral principle or general advantage once these are clearly apprehended and understood. Such doctrines are the fictions of philosophical theory and ignore three essential points. First, vast areas of human behavior are, in fact, governed by overriding antisocial passions like pride and cruelty. The indulgence of these is in general detrimental to the agent's long-term advantage, frequently causing as much harm to him as to those against whom his actions may be directed. Second, it is a mistake to write off as mere eradicable superstition the various myths, religious and otherwise, in terms of which men are prone to conduct their lives. These are often attuned to powerful nonrational forces in the psyche that demand expression and that, if frustrated, are likely to seek outlet in other, possibly more dangerous forms. Third, it is important to appreciate how often people are totally unaware of the true motives and drives that determine their actions; human beings are adept at rationalization and self-deception, and their conduct may appear to be guided by reason when, in reality, it is directed by quite different factors. Intimations of these notions occurred in the writings of the Marquis de Sade and Joseph de Maistre at the close of the eighteenth century; and they were subsequently given forceful expression in the works of romantic and postromantic thinkers like Schopenhauer and Nietzsche. More recently, they have been regarded as receiving impressive and detailed corroboration from the advances in psychoanalysis initiated by Sigmund Freud and C. G. Jung.

POLITICAL AND SOCIAL THOUGHT. In the sphere of political and social theory, insistence upon the irrationality of human nature has tended to be combined with traditionalist, authoritarian, or reactionary conceptions of government. To some, it has seemed obvious that the only enduring way of preserving the integrity of society against the disruptive forces of violence and passion lurking beneath the thin surface of civilized life consists in the use of coercion and suppression. De Maistre, for instance, considered the executioner to be the most significant figure in the state. Stress is laid on the importance of instilling habits of obedience to authority by appeals to supernatural or providential powers and by safeguarding the atmosphere of reverence and awe that surrounds the person of the ruler in established societies—a principal objection to proposals for the reorganization of social life according to egalitarian or consciously utilitarian general principles has been the belief that they can only lead to a loosening of the mysterious ties that hold a political community together. Likewise, attempts to displace unreasoned acceptance of the existing order of things by the propagation of scientifically inspired ideas and policies strike at the root of all that makes for social cohesion.

Edmund Burke was, for these reasons, deeply distrustful of revolutionary theories and plans. He thought that the true sources of political harmony lay below the level of rational reflection and showed considerable prescience concerning the consequences likely to ensue if the checks upon men's passions provided by traditional arrangements were challenged or removed. He did not, however, share the curiously ambivalent attitude toward violent or sadistic human propensities discernible in certain later social thinkers, who saw these as something to be systematically exploited rather than inhibited and for whom the ideas of force and brutality seem to have possessed a powerful emotional appeal. In the case of Vilfredo Pareto, for instance, the approach adopted toward the role of the irrational in human life was not as detached or objective as he tried to present it. Such writers did not merely dismiss humanitarian schemes for

social amelioration and improvement as ultimately unrealistic, impracticable, or utopian; it was also strongly suggested in their works that if these schemes were to be realized, this would constitute an intrinsically undesirable state of affairs. It is for pressing the second claim, as well as the first, that fascism is often described as an irrationalist ideology.

Major currents of thought do not originate in a vacuum, and the various components of modern irrationalism have many diverse sources. Among them are the void left by the decay of institutionalized religion, the recurrent failure of large-scale reformist movements (like the French and Russian revolutions) to fulfill the hopes that originally inspired them, and the inability of contemporary industrial society to provide scope for individual self-expression. But it would be a mistake to regard irrationalist trends as purely pathological symptoms or to suppose that they have contributed nothing of value to the development of thought. It is common for Anglo-Saxon critics to denounce some irrationalist claims as having played a pernicious role in the formation of extremist political ideologies and to dismiss others as representing no more than inflated or misleading formulations of familiar logical doctrines—for instance, it has been suggested that the existentialist conception of the world as irrational is (partly at least) a bombastic restatement of the Humean insight that there exist no necessary connections between matters of fact. Up to a point such objections may be justified. However, it is worth remembering that there are important areas of human consciousness and behavior that theorists of a rationalistic temper have been characteristically prone to overlook and that it has been largely left to theorists of a different outlook to explore and define these areas. To say that the task has sometimes been perversely performed is not to say that it should not have been undertaken at all.

See also Ayer, Alfred Jules; Bergson, Henri; Burke, Edmund; Camus, Albert; Carnap, Rudolf; Descartes, René; Enlightenment; Existentialism; Fascism; Fichte, Johann Gottlieb; Freud, Sigmund; Jung, Carl Gustav; Hegel, Georg Wilhelm Friedrich; Historicism; Hobbes, Thomas; Hume, David; Kant, Immanuel; Kierkegaard, Søren Aabye; Leibniz, Gottfried Wilhelm; Locke, John; Maistre, Comte Joseph de; Myth; Nietzsche, Friedrich; Pareto, Vilfredo; Pascal, Blaise; Sartre, Jean-Paul; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Stevenson, Charles L.; Stirner, Max.

Bibliography

MAJOR IRRATIONALIST WORKS

Bergson, Henri. *Introduction to Metaphysics*. Translated by T. E. Hulme. New York: Putnam, 1913.

Bergson, Henri. *Time and Free Will*. Translated by P. L. Pogson. London: Sonnenschein, 1910.

Bergson, Henri. *The Two Sources of Morality and Religion*. Translated by R. A Audra and Cloudesley Brereton. Garden City, NY: Doubleday, 1954.

Camus, Albert. *Le mythe de Sisyphe*. Paris: Gallimard, 1942.

Dostoevsky, Fëdor. *Notes from Underground*. Translated by Constance Garnett, in his *Works.* 12 vols. New York, 1912–1920.

Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson. New York: Harper, 1962.

Kierkegaard, Søren. *Concluding Unscientific Postscript*. Translated by D. F. Swenson and Walter Lowrie. Princeton, NJ: Princeton University Press, 1941.

Kierkegaard, Søren. *Philosophical Fragments*. Translated by D. F. Swenson, with introduction and commentary by Niels Tholstrup. Princeton, NJ: Princeton University Press, 1962.

Maistre, Joseph de. *Les soirées de Saint-Pétersbourg.* 6th ed. Paris, 1850.

Müller-Freienfels, Richard. *Irrationalismus*. Berlin, 1923.

Müller-Freienfels, Richard. *Metaphysik des Irrationalen*. Leipzig: Meiner, 1927.

Nietzsche, Friedrich. *Beyond Good and Evil*. Translated by Helen Zimmern. New York: Macmillan, 1924.

Nietzsche, Friedrich. *The Genealogy of Morals*. Translated by H. B. Samuel. New York: Macmillan, 1924.

Nietzsche, Friedrich. *The Portable Nietzsche,* edited and translated by Walter Kaufmann. New York: Viking Press, 1954.

Nietzsche, Friedrich. *The Will to Power*. Translated by A. M. Ludovice. New York: Macmillan, 1924.

Pareto, Vilfredo. *Mind and Society*. Translated by Arthur Livingstone. London: Cape, 1935.

Pascal, Blaise. *Pensées*. Translated by W. F. Trotter. New York, 1954.

Sade, Marquis de. *Selected Writings*. Translated by Paul Dinnage. London, 1962. Introductory essay by Simone de Beauvoir.

Sartre, Jean-Paul, *L'être et le néant*. Translated by Hazel Barnes as *Being and Nothingness*. New York: Philosophical Library, 1956.

Sartre, Jean-Paul. *La nausée*. Paris: Gallimard, 1938.

Schelling, F. W. J. von. *Philosophie und Religion*. In his *Werke*, edited by M. Schröter. Munich, 1927–1928. Vol. IV.

Schopenhauer, Arthur. *Die Welt als Wille und Vorstellung*. Translated by E. F. J. Payne as *The World as Will and Representation.* 2 vols. Indian Hills, CO: Falcon's Wing Press, 1958.

Sorel, Georges. *Reflections on Violence*. Translated by T. E. Hulme. New York: Huebsch, 1914.

Spengler, Oswald. *The Decline of the West*. Translated by C. F. Atkinson. New York, 1947.

Tolstoy, Lev. *Death of Ivan Ilyitch* and *A Confession*, in his *Works*, translated by Louise Maude and Aylmer Maude. 21 vols. Oxford, 1928–1937.

GENERAL COMMENTARIES AND SURVEYS

Agassi, Joseph, and I. C. Jarvie. "The Rationality of Irrationalism." *Metaphilosophy* 11 (1980): 127–133.

Aiken, H. D. *The Age of Ideology*. New York: New American Library, 1956. Selections from nineteenth-century philosophers, with introduction and interpretive commentary.

Almond, Brenda. "Philosophy and the Cult of Irrationalism." *Philosophy*, supplement 33 (1992): 201–217.

Ayer, A. J. "Some Aspects of Existentialism." *Rationalist Annual* (1948).

Barrett, William. *Irrational Man*. Garden City, NY: Doubleday, 1958.

Berlin, Isaiah. *The Hedgehog and the Fox*. London: Weidenfeld and Nicolson, 1953. This study of Tolstoy's view of history includes a discussion of the antirationalist strains in Tolstoy's thought, and explores some of the parallels between his ideas and those of Joseph de Maistre.

Blackham, H. J. *Six Existentialist Thinkers*. New York, 1952.

Brown, Harold. *Rationality*. London: Routledge, 1988.

Copleston, Frederick. *History of Philosophy.* Vol. VII: *Fichte to Nietzsche*. London, 1963.

Mele, Alfred. "Motivated Irrationality." In *The Oxford Handbook of Rationality*, edited by Alfred Mele. Oxford: Oxford University Press, 2004.

Rescher, Nicholas. *Pluralism: Against the Demand for Consensus*. New York: Oxford University Press, 1993.

Rogers, W. Kim. "Existentialism Is not Irrationalism: A Challenge to the Common Interpretation of Existentialism." *Journal of the British Society for Phenomenology* 20 (1989): 77–83.

Rorty, Richard. "Pragmatism, Relativism, and Irrationalism." *Proceedings of the Addresses of the American Philosophical Association* 53 (1980): 719–738.

Russell, Bertrand. "The Ancestry of Fascism." In *In Praise of Idleness*. London, 1960.

Russell, Bertrand. *Power, a New Social Analysis*. New York: Norton, 1938. Ch. 16.

Simon, Herbert. *Reason in Human Affairs*. Stanford, CA: Stanford University Press, 1981.

Smart, J. J. C. "Critical Notice: David Stove, *On Enlightenment and Scientific Irrationalism*." *Australastian Journal of Philosophy* 81 (2003): 429–433.

Stich, Stephen. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Young, Julian. "Is Schopenhauer an Irrationalist?" *Schopenhauer Jahrbuch* 69 (1988): 85–100.

Watkins, John. "On Stove's Book, by a Fifth 'Irrationalist.'" *Australasian Journal of Philosophy* 63 (1985): 259–268.

Wolf, L. H. de. *The Religious Revolt against Reason*. New York: Harper, 1949.

Wollheim, Richard. "The Political Philosophy of Existentialism." *Cambridge Journal* 7 (1953): 3–19.

Patrick Gardiner (1967)

Bibliography updated by Benjamin Fiedor (2005)

ISAAC OF STELLA

(c. 1100–c. 1169)

Isaac of Stella, one of the great monastic thinkers of the Middle Ages, was born in England about 1100. He apparently studied in both England and France before entering the monastery of Citeaux. After several years at Stella (L'Étoile) in Poitou, where he became abbot, Isaac attempted to found a monastery on the lonely island of Ré, near La Rochelle, but soon returned to Stella, where he died about 1169. His writings include a treatise on human nature (*De Anima*), an exposition of the liturgy (*De Officio Missae*), and fifty-four sermons, preached either at Ré or at Stella. Through the *De Spiritu et Anima* of Alcher of Clairvaux his psychological theories became widely influential, notably in the Franciscan school of the thirteenth century.

Isaac's mind, schooled in the biblical spirituality of the Cistercians, was steeped in Scripture, and his writings are full of biblical allusions. In contrast to many of his contemporaries, however, he was careful and systematic in his use of Scripture. Moreover, although most monastic interpreters were content with the moral lessons derivable from the biblical text, Isaac was deeply interested in its doctrinal content. Thus, his biblical exegesis reflects his metaphysical concerns.

As a philosophical theologian, Isaac stood in the tradition of Christian Neoplatonism at the point where it first felt the impact of the Aristotelian renaissance. Both the Greek Fathers and Augustine were extensively studied by the Cistercians, but Isaac's grasp of their teaching was exceptional. Indeed, apart from Erigena no earlier medieval thinker could equal his knowledge of Eastern and Western Neoplatonism. On the one hand, as both his doctrinal tendencies and his extensive use of a Dionysian vocabulary, including at least a dozen Greek terms, indicate, he was well acquainted with the works of the pseudo-Dionysius. On the other hand, he was thoroughly familiar with the philosophical, theological, and mystical thought of Augustine.

Isaac's ambition to reconcile Neoplatonism and Aristotelianism is apparent in his account of human knowledge, which combines the Augustinian doctrine of illumination with the theory of abstraction. In his synthesis reason forms universal concepts by abstraction from sense experience of corporeal objects. Intelligence, however, must be aided by divine illumination in its effort to apprehend incorporeal beings.

The influence of the pseudo-Dionysius can be seen in Isaac's insistence on the negative approach (*via nega-*

tiva) to the knowledge of God. It appears also in his emphasis on the hierarchical structure of reality, in his exemplarist doctrine of creaturely participation in the divine perfections, and in his strong interest in liturgical symbolism.

The influence of Augustine's theology is most conspicuous in Isaac's discussion of predestination. With frequent echoes of Augustine's own style, he fully develops the theme of God's initiative in the process of human salvation. Augustinian influences are obvious also in Isaac's teaching on many points, including the Trinity, the virtue of charity, and the church as Christ's mystical body.

See also Aristotelianism; Augustine, St.; Determinism, A Historical Survey; Erigena, John Scotus; Illumination; Neoplatonism; Pseudo-Dionysius.

Bibliography

See Louis Bouyer, *The Cistercian Heritage* (London: Mombnay, 1958); George Bosworth Burch, *Early Medieval Philosophy* (New York: King's Crown Press, 1951); and W. Meuser, *Die Erkenntnislehre des Isaak von Stella* (Bottrop, Germany, 1934).

Eugene R. Fairweather (1967)

ISLAMIC PHILOSOPHY

In Islam the development of philosophical thought, properly speaking, succeeded earlier schools of dialectical theology (*kalām*) that began to arise in the eighth century (second century AH in the Islamic calendar) through the action of foreign ideas—particularly Greco-Christian—on certain fundamental moral issues raised within the Islamic community. These moral issues clustered particularly around the problems of the freedom of the human will, God's omnipotence and justice, and God's relationship to the world. Although these early schools do not properly belong within the scope of this article, since they are theological rather than philosophical, a very brief characterization of the main groups and their tenets will serve to elucidate the content of the philosophical movement itself. Broadly speaking, there were two theological schools. The so-called rationalist, or Muʿtazila, school maintained the freedom of the will; insisted that right and wrong are knowable through reason independently of, but confirmed by, revelation; and claimed that God's attributes are identical with his essence and that God cannot do what is unreasonable or unjust. However, the Muʿtazilites posed and solved all these problems theologically, not philosophically; their entire thought was theocentric. For example, they did not pose the problem of the will absolutely but discussed it mainly insofar as it is relevant to the concept of a just God. However, their opponents (the Ahl al-Sunnah waʾl-Jamāʿah), who came to constitute the orthodoxy, accused them of stark humanism and opposed them on all these major questions. The orthodoxy, after a long, hard struggle, completely routed the Muʿtazilites as a theological school, but the spark of the Muʿtazilites kindled the purely rationalist movement in philosophic thought.

The work of the original philosophers in Islam was preceded by feverish translation that began around 800 and lasted for about two hundred years; its climax was reached in the time of Caliph al-Maʾmūn al-Rashid (reigned 813–833). Al-Maʾmūn set up the first official seat of liberal learning in Islam, called the House of Wisdom, whose main function was to translate the works of the Greek masters of science and philosophy. The translations, however, were mostly from Syriac versions and not directly from the Greek. These translations, which were made almost invariably by Arab Christians, covered the entire range of Greek civilization—that is, its thought content—but excluded such specifically cultural aspects as mythology, drama, and literature, which were foreign to the Arabs and to Islam. The Arabs were able to develop a highly technical philosophical diction with astonishing rapidity and to integrate it into the Arabic language so successfully that a philosopher like al-Fārābī (c. 873–950), who was a Turk and not an Arab, was able to express himself philosophically in Arabic with remarkable facility. All this happened within a span of about 150 years in a language that had previously known no technical philosophical literature whatsoever.

The main character of Islamic philosophy was set by the combination of Aristotle and Neoplatonism that had constituted an important tradition in the late stages of Hellenistic philosophy and that was represented particularly by the Neoplatonic commentators on Aristotle in Athens and Alexandria, such as Simplicius and John Philoponus (sixth century). The Muslim philosophers introduced into this tradition other fundamental concepts in order to adapt it to an Islamic milieu; the most important were the ideas of contingent and necessary being and of prophethood. Despite these fundamental changes, the Muslim philosophers accepted the general cosmological scheme they had inherited from the Greek traditions. Thus, an important place in their cosmology and metaphysics is occupied by the role of the stars and the heavenly bodies, a role that has no place in the scheme of reality of the Qurʾan. This must be attributed to the

Greek beliefs about the status of stars and the heavenly bodies and their creative influence on the sublunary sphere, although such a picture of the universe was also quite in harmony with other traditions existing in the Middle East, for instance, Sabaeans and Babylonians.

AL-KINDĪ

The first important Muslim philosopher was the Arab prince Abū-Yūsuf Yaʿqūb ibn Isḥāq al-Kindī (d. after 870). Al-Kindī's philosophic thought is directly connected with, on the one hand, Greek philosophical doctrines transmitted to him through translations and, on the other, with the rationalist theological movement of the Muʿtazilites. He seems to have espoused the Muʿtazilite doctrines in toto and to have sought to create a philosophical substructure for them. Thus, the Muʿtazilite dogma of the attributeless transcendence of God must have led him to the somewhat parallel idea of God as absolute and transcendent being, a combination of the Aristotelian concept of God and the Neoplatonic concept of the One. It is this affinity that must have led him further to formulate the doctrine, common to all the great Muslim philosophers, that philosophy and religion, or the rational truth and the revealed truth, not only do not conflict with each other but, in fact, lend support to each other and are basically identical. This recalls the Muʿtazilite doctrine that the source of our knowledge of values is reason confirmed by revelation.

In his philosophy, al-Kindī was more of a Neoplatonist than an Aristotelian. (The Arabs attributed certain Neoplatonic works, such as *De Causis* and *Theologia Aristotelis,* to Aristotle.) He adopted the Neoplatonic doctrine of emanation in his metaphysics and cosmology. Also, in his theory of intellectual knowledge he adopted the doctrine of the active intellect and the passive intellect, originally formulated by Aristotle, later elaborated by the commentator Alexander of Aphrodisias, and subsequently reworked and essentially modified by Neoplatonists. Al-Kindī introduced into the Greek framework of ideas some fundamental doctrines of Islam. Thus, although he accepted the theory of emanation, he asserted that the first being was created by the sheer act of God's will and out of nothing, an antithesis to the general Greek doctrine that nothing comes out of nothing. Aristotle had postulated two ultimates—one was God, the form of forms; the other, the prime matter—each of which had "existed" independently of the other. Similarly, although the Neoplatonic doctrine of emanation differs vitally from Aristotle's theory of the cosmic movement, it still seeks to avoid having to accept creation ex nihilo by postulating the emanatory process. However, it is difficult to see how, in the last analysis, the emanation theory can overcome the difficulties of creation ex nihilo. Al-Kindī, however, simply asserted emanationism and creationism side by side without reconciling the contradiction between the two. It was Avicenna (Ibn Sīnā) who later attempted the reconciliation, but it was important to the development of Islamic philosophy that al-Kindī, far from giving up the Islamic requirements of the relationship of God and the world, juxtaposed both the Islamic and the Greek doctrines. In his theory of intellection, al-Kindī was attracted by the ideal of a form of knowledge that would do justice to the demands of reason and revelation, although in his extant works we do not find an elaborated theory of prophethood. This, again, was taken up later by al-Farabi and Avicenna, but it was al-Kindī who initiated development of the theory of intellection in Islamic philosophy.

AL-FĀRĀBĪ

With al-Fārābī, philosophy reached maturity in Islam. Not many of his works have come down to us, but his writings that we do possess reveal an unusually incisive and clear mind. In his cosmology, as well as in his psychology, al-Fārābī was almost entirely Aristotelian, except for the doctrine of emanation. In political theory, which seems to have preoccupied him considerably more than it did other Muslim philosophers, he based himself on Plato's *Republic* and *Laws,* but he adapted the Platonic system to his contemporary political situation with a remarkable ingenuity. He developed the doctrine of the intellect from the point at which al-Kindī had left off, and he constructed a theory of divine inspiration that was to serve as a model for Avicenna. But apart from his original theories, the importance of al-Fārābī lies in his attempt to elevate philosophy to the place of highest value and to subordinate the revelation and the *sharīʿa,* or religious law, to it. In this also he served as a model for both Avicenna and Averroes (Ibn Rushd), but it was precisely this doctrine, in which the *sharīʿa* took an inferior place as a symbolic expression of a higher intellectual truth, that was also ultimately responsible for the fatal attacks on the philosophical movement by representatives of the orthodoxy.

In his religious attitudes, al-Fārābī was a genuinely universalistic spirit who believed that the entire world should have one religion, of which all particular religions would be considered symbolic expressions. However, it would be a mistake to regard al-Fārābī as a relativist. He tells us in no uncertain terms that not all religions are

equal either as adequate symbols of truth or as the effective harnessing of men's minds and hearts. Indeed, he believed that there are religious symbolisms that are positively harmful and must be discarded. He did affirm, however, that there are religions which are equivalent in their religious value; and any one of these symbolic systems may be applied in a given milieu, depending upon circumstances. Although al-Fārābī gave no concrete examples of religions or names of prophets, there is little doubt that the prophet Muḥammad was fixed in his mind as a paradigm par excellence of a prophet and a lawgiver. This becomes clear in his insistence that the teachings of a prophet should not only be universal but should also be successful in history.

Al-Fārābī's writings give us a full-scale picture of the basic world view of Muslim philosophy. At the apex of his scheme of reality stands God, who is both the One of Plotinus and the First Cause of Aristotle. From him proceeds the first intelligence, which is also the archangel. The first intelligence has a dual nature and gives rise to two further beings: the highest sphere on the physical side and the second intelligence on the spiritual side. This process of emanation continues until we reach the tenth sphere and the last intelligence, identified as the angel of revelation, Gabriel, on the one hand, and as the sphere of the moon on the other. The entire process of the world below the moon is an interaction between the materials emanating from the sphere of the moon and the spiritual influence generated by the tenth intelligence, called the Active Intellect. This interaction generates the world process, and its culminating product is man, with his fully organized body and rational soul.

The goal of man, wherein lies his ultimate bliss, is to develop his rational faculty by his will. The rational faculty is developed by the action of the active intelligence upon it, through which actual thought arises. The end of man, therefore, is to reach philosophic contemplation, and al-Fārābī categorically states that men whose rational faculty remains undeveloped cannot attain immortality but perish with their physical death. The actual activation of man's rational power, however, demands certain practical virtues as well, and this makes it necessary for man to live in organized societies rather than in isolation. People who are ultimately responsible for organizing and directing human societies are those possessed of philosophical wisdom, for it is not possible to enunciate practical laws for humankind without having theoretical wisdom. Therefore, for al-Fārābī the philosopher and the prophet are identical. It is the philosopher-prophet who can formulate the practical principles and laws that will lead men to their final goal of philosophic bliss. Societies governed by such laws are "good societies"; others are "ignorant societies," "misguided societies," or "retarded societies."

At the final stage of the intellective development, the philosophical mind becomes like matter to the Active Intellect, which becomes its form. This is the absolute apogee of human bliss. The prophet is a person who, having attained this philosophical illumination, transforms the philosophic truth into an imaginative myth that moves people to action and can influence societies toward greater morality. It is because of his imaginative power, the power to represent the intellectual truth in the form of a figure or a symbol, that the prophet is able to make laws and to bring revelation. Revelation, therefore, is not philosophic truth but imaginative truth. Only a few gifted philosophical spirits can pierce the imaginative shell and reach the philosophic truth. In al-Fārābī 's theory of prophethood, there seems to be no place for miracles; the accommodation of miracles on a philosophical basis was the work of Avicenna.

Al-Fārābī likened the ruler to the head in the human organism and, like Plato, developed the idea of a hierarchy in which each stratum receives orders from above and issues commands to those below. Just as at the top there is a ruler who is not ruled, so at the bottom there are those who are ruled but do not rule. It is a fully authoritarian view of government, and some scholars have suggested that al-Fārābī was influenced by Shiʿite doctrine. The fact that al-Fārābī was at the court of the Shiʿite ruler is supposed to lend some support to this view. We do not have sufficient historical evidence for such a judgment, but it should be noted that the ultimate ruler of the Farabian state does resemble the Shiʿite Imam, the repository of divine wisdom.

BRETHERN OF PURITY

During the tenth century, a secret coterie of popular philosophers known as the Brethren of Purity (*Ikhwan al-Safa*) was formed, and they wrote a series of "epistles," or treatises, titled Rasāʾil Ikhwān as Safāʾ, to propagate their views. The epistles exhibit a thoroughly Neoplatonic character. They seek to formulate a worldview culminating in a universalistic religion transcending all organized religions, which, at best, serve as so many different ladders to the ultimate truth. The philosophy preached by the Brethren of Purity is also esoteric, and there are strong reasons to believe that this group was either formed by members of or was connected with the Ismāʿīli movement, a religious sect; it is very likely that it was

through such channels that Ismāʿīlism absorbed those Greek philosophic elements which were rejected by the Muslim orthodoxy but were akin to certain patent Oriental theories and to attitudes about religion and the nature of the ultimate truth. The view of the Brethren of Purity does not constitute philosophy in the strict sense but is a kind of vague and romantic idealism; nevertheless, it is important to note it because its ideas have also influenced the development of another powerful spiritual movement in Islam, Sufism.

AVICENNA

The most important and original of Muslim philosophers was Abū ʿAlī ibn Sīnā, known to the West as Avicenna (980–1037). The philosophic movement in eastern Islam comes to its fullest fruition in the thought of Avicenna, who elaborated one of the most cohesive, subtle, and all-embracing systems of medieval history. In the West his ideas had a profound influence on medieval scholastic philosophy, and in the Muslim world his system is still taught in the traditional centers of Islamic learning. The central thesis of Avicenna's metaphysics is the division of reality into contingent being and Necessary Being. In order to formulate this doctrine, whose influence has been so palpable and enduring in both Eastern and Western thought, Avicenna devised his theory of the distinction between essence and existence. In this theory, he refined the implications of the Islamic doctrine of creation, which al-Kindī had crudely asserted, into an integrated philosophic system.

The bases of this theory of essence and existence are set in Aristotle's doctrine of movement and in the Neoplatonic doctrine of emanation, but in order to achieve the desired results, Avicenna had to effect basic changes both in the doctrine of emanation and in the Aristotelian doctrine of matter and form. Briefly, Aristotle had taught that matter is the principle of potentiality and form the principle of actuality, and that through the interaction of the two the actual movement of the universe takes place, in which potentialities are progressively actualized. Thus, the analysis of any given thing—with the exception of God and prime matter—falls into matter and form. There are, however, grave objections to this view. How can an actual thing come into existence through the interaction of a matter that, according to Aristotle, does not exist and a form that also does not exist? Why should things not remain unactualized in their potentialities, and where is the necessity of movement? Emanation seems to simplify this problem by asserting a single, universal process of outward movement, but it gives no rationale of this movement.

Closer examination led Avicenna to posit three factors—matter, form, and existence—and to postulate a Necessary Being as the basis for the world process. There is little doubt, however, that it was not merely these philosophic reasons that led him to formulate this doctrine but also the fact that Islam demanded a fundamental distinction between God and the world. Since Avicenna could not accept the creationism of the Muslim theologians because it implied temporal priority of God over the world, he affirmed that God is distinguished from the world by the fact that his being is necessary and simple; God cannot be composed of matter and form but must be pure existence. From God emanate the intelligences, which, although they have no matter, are nevertheless composites of essence and existence; the material beings are composed of matter and form, which constitute their essence, and the fact of their existence—all existence flowing from God.

Avicenna was thus able to solve, to his own satisfaction, the contradiction that seemed to exist between the Greek philosophic world view and the Islamic doctrine of creationism: in accord with the philosophers he affirmed the eternity of the world and rejected temporal creation, but with the Islamists he made the world entirely and eternally dependent upon God. This solution led him to establish the relationship between religion and philosophy. Since the findings of religion and of philosophy do not contradict one another on this crucial point but are not identical either, they run parallel to one another. From this, Avicenna expounded his further view that religion is a kind of philosophy for the masses: It does not tell the naked philosophical truth but is an endeavor to make the masses come as near to the philosophical truth as possible. The prophets are, then, mass psychologists who launch religious movements as pragmatic endeavors to make people virtuous. Thus, Avicenna reaffirms al-Farabi's position that revelation is not philosophic truth but symbolic truth.

The possibility of prophethood in Avicenna's system is intimately connected with his theory of knowledge, particularly with his theory of the creative knowledge and of the "internal sense," which appears to be his own contribution to the history of thought. According to Avicenna, all genuine intellectual discovery implies an intuitive act of knowledge, and our ratiocination merely prepares for us this intuitive act. However, there can be—and there are—people who possess a tremendous native intuitive power even without any ratiocination and

process of learning. The ultimate limit of such a gifted mind is the prophetic mind, which does not receive knowledge through learning but creates knowledge. This constitutes the prophetic revelation at the intellectual level. But this intellectual power, in a genuine prophet, flows into the imagination or the "internal sense" as well, thus enabling the imaginative faculty to transform the intellectual truth into images and symbols capable of moving people's minds and bodies. It was on the basis of this power of imagination and suggestion that Avicenna explained the possibility of miracles attributed to prophets. He was thus able to accept even the miracle doctrine of the orthodoxy, although he rejected certain miracles as being "impossible."

AL-GHAZĀLĪ

Avicenna's system went furthest in integrating the traditional demands of the orthodox religion with the purely Greek rationalism, which explains why his works continue to be studied in the traditional Islamic schools even today. However, his system was made the object of denunciatory criticism by the orthodoxy on certain points: the eternity of the world, the inferior status of the *sharīʿa* (religious law) as a mere symbol of the higher truth, and the rejection of the resurrection of the body. The classical criticism was carried out by al-Ghazālī (1058–1111) in his famous work *Tahāfut al-Falāsifa* (Incoherence of the philosophers), which was also rendered into Latin in the thirteenth century under the title *Destructio Philosophorum.*

AVERROES

The unrelenting criticism of philosophy as it appeared in Avicenna's system by al-Ghazālī and others led Ibn Rushd, known in the West as Averroes (c. 1126–c. 1198), to defend the claims of philosophy. In the process of doing this, Averroes sought to resurrect the original Aristotelian doctrines from the later Neoplatonic and Muslim accretion as much as possible. He wrote many commentaries on the works of Aristotle, whom he believed to be the philosopher par excellence. He accused both Avicenna and al-Ghazālī of having mutilated philosophical theses and of having confused them with religious doctrines. Averroes, however, did not advocate a theory of two truths, although this may be a logical conclusion of what he said in his work titled *Faṣl al-Maqāl* (The decisive statement) on the relationship between philosophy and religion.

Averroes rejected Avicenna's distinction between essence and existence. He insisted that existence is, in a way, part of the essence of a thing. The one conspicuous doctrine on which Averroes does not appear to be a faithful follower of Aristotle is that concerning intellect. He declared the passive human intellect also to be eternal and incorruptible and, indeed, to be universal to all humankind, like the Active Intellect. This doctrine of the unity of intellect, besides being apparently unfaithful to Aristotle, was also unacceptable to the followers of the revealed religions. He was thus attacked both by Muslims and, in the West, by Thomas Aquinas, who wrote a special treatise, titled *De Unitate Intellectus,* against the Averroistic doctrine. It must, however, be pointed out that the common objection raised against Averroes' doctrine of the universality of the intellect ever since Thomas's classic formulation of it as *ego intellego* is very superficial. Averroes not only never held that the act of cognition is universal but was, in fact, at pains to prove its individual character. What he seems to be concerned to show is that all thinking, although it occurs individually, becomes in a real sense universal, and that this universal aspect is more intrinsic to human cognition than is the fact that it is the product of such-and-such an individual or individuals. In any case, it is certain that Averroes never denied the individuality of the act of cognition.

Although Averroes believed that religion and philosophy are in two different orbits, he nevertheless felt the necessity of reconciling the two and of so stating the philosophic doctrines as not to offend religion and of so conceiving the religious dogmas that they would not conflict with philosophy. We are, therefore, back at the position of Avicenna. On the question of the eternity of the world, Averroes taught the doctrine of eternal creation. Although he did not reject the religious dogmas of the resurrection of the body, as Avicenna had done, he taught that the numerically same body cannot be resurrected. There was, however, bitter opposition to the doctrines of Averroes, who was also the *qadī* (judge) of Seville, and today very few of his works survive in the original Arabic; they are to be found mostly in Hebrew and Latin translations.

ABUʾL-BARAKĀT IBN MALKĀ

In the East we find another important attempt at the *rapprochement* of the content of religion and philosophy in the works of Abuʾl-Barakāt ibn Malkā (also known as Abuʾl-Barakāt al-Baghdādī, d. c. 1174/1175). A Jew converted to Islam, Abuʾl-Barakāt's doctrines show a decisive trend toward Islamic orthodox beliefs. Thus, on the question of the attributes of God, he affirmed all the attributes of the Deity in the positive sense and not as pure nega-

tions, as his predecessors had done. His doctrine that the eternal essence of God can be the subject of changing accidents is palpable proof of his conscious orthodoxy. The doctrine is so obviously removed from the teaching of the early great Muslim philosophers and of Aristotle himself that, while it did not seem to have much appeal for the philosophic tradition in Islam, it evoked enthusiastic approval from such orthodox ʿUlamāʾ (the "learned") as Ibn Taymiya (thirteenth and fourteenth centuries). Similarly, Abuʾl-Barakāt taught that the intellectual and the perceptual faculties are not different but are one and the same. He rejected the teachings of the Aristotelians that God does not know the particulars but only the universals, and he obviously did not accept Avicenna's formulation of the doctrine that God knows every particular but "in a universal way" rather than through perception. According to Abuʾl-Barakāt, both sense perception and intellective perception belong to the soul and do not intrinsically involve the body. Then he concludes that God knows the particulars just as he knows the universals.

Although further progress of philosophy was cut off by the blows of the orthodoxy, philosophical developments, especially the system of Avicenna, had exerted a rejuvenating influence on orthodox theology (*kalâm*). After al-Ghazālī's refutation of philosophy, the scope of theology was expanded to include all the epistemological and metaphysical questions the philosophers had dealt with but to which theological answers were now provided. The first person to attempt this and who is, in fact, the forerunner of all Islamic theologians is Fakhr ad-Dīn ar-Rāzī (1149–1209). Logic was simply taken over by *kalam* as a necessary instrumental science. Thus, the official theology set itself up as "the crown of the religious sciences" and began to function as a sufficient substitute for philosophic thought. Rational thought was thus banished from the schools as being redundant; only Avicenna's works (and commentaries and compendia based upon them) were taught, but more in order to be refuted than to instigate independent thought.

Under the attacks of orthodoxy, philosophy went underground, as it were, and lived on in the form of now one theosophy, now another. Instead of continuing as a purely rational expression of the human mind, it emptied its contents into intellectual Sufism. Henceforth, we do not get pure philosophy in Islam but a mystical philosophy. After the activity of the pantheist Sufi theosoph Ibn al-Arabī (1165–1240), the new philosophic mysticism developed into a closely argued and elaborate system in the works of Ṣadr al-Dīn al Shīrazī, commonly known as Mullā Ṣadrā (1571/1572–1640). Mullā Ṣadrā represents a conjunction of the Shiʿite doctrine, the philosophic tradition of Avicenna, the mystical intellectualism of Shihāb al-Dīn Yahyā Suhrawardī (executed at Aleppo in 1192), and of Ibn al-Arabī. He is a typical representative of the intellectual-spiritual tradition of late medieval Islam. A monist, Mullā Ṣadrā believed in a doctrine of mystic "return" to the First Principle of being. The reality as given is constituted by three levels of "worlds"—the spiritual, the imaginative, and the physical. The "imaginative" world (*ʿālam al-mithāl*) is the world of symbols or images that relates the spiritual and the physical realms to one another, and it is the realm essentially relevant to the genesis and interpretation of symbols given in religious experience. This doctrine exercised a very considerable influence on subsequent developments in Islamic thought until the dawn of modern times. The centrality of "the world of symbols," with its religious implications and with its escapism from the external world, is symptomatic of the refined spiritual and intellectual culture of Islam in the later Middle Ages until the impact of Western influence upon it.

The story of philosophic thought in Islam after Averroes still remains to be written. Modern Western students of Islamic philosophy generally stop short at Averroes because the Muslim philosophic movement exerted an influence on medieval Western philosophy until his time. It is a pity that Muslim philosophy has been studied not as an internal whole but essentially from the point of view of its impact upon and relationship to Western philosophy. However, even a thorough account of the influence of Islamic ideas on Western thought is still lacking.

See also Alexander of Aphrodisias; al-Fārābī; al-Ghazālī, Muhammad; al-Kindī Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotelianism; Aristotle; Averroes; Averroism in Modern Islamic Philosophy; Avicenna; Determinism and Freedom; Dialectic in Islamic and Jewish Philosophy; Enlightenment, Islamic; God, Concepts of; Ibn al-ʿArabī; Mullā Ṣadrā; Mysticism, History of; Neoplatonism; Philoponus, John; Rationalism; Simplicius; Sufism; Suhrawardī Shihāb al-Dīn Yaḥyā Thomas Aquinas, St.

Bibliography

GENERAL WORKS

Boer, T. J. de. *The History of Philosophy in Islam.* London, 1903.

Corbin, Henry. *Histoire de la philosophie islamique,* Vol. I. Paris: Gallimard, 1964. Paperback.

Menasce, P. J. de. *Arabisches Philosophie.* No. 6 in Bibliographische Einführungen in das Studium der Philosophie. Bern, 1948.

Munk, Salomon. *Mélanges de philosophie juive et arabe.* Paris: A. Franck, 1859; 2nd ed., Paris: J. Gamber, 1927.

Walzer, Richard. "Islamic Philosophy." In *History of Philosophy, Eastern and Western,* Vol. II, edited by Sarvepalli Radhakrishnan et al. London: Allen and Unwin, 1952–1953. Contains a selected bibliography on individual Muslim philosophers.

Watt, W. Montgomery. *Islamic Surveys. Vol. I, Islamic Philosophy and Theology.* Edinburgh: Edinburgh University Press, 1962.

SUPPLEMENTARY WORKS

In addition to the works listed in the bibliography of the Walzer work, above, the following should be consulted:

"Al-Kindi's Treatise on Intellect." Edited and translated by R. J. McCarthy. *Islamic Studies* 3 (4) (1964).

Averroes' Tahafut al-Tahafut (The Incoherence of the Incoherence). Translated with an introduction and notes by Simon van den Bergh. 2 vols. London: Luzac, 1954.

Corbin, Henry. *Oeuvres philosophiques et mystiques de Sohrawardī.* Teheran: Institut Franco-Iranien, 1952. Especially *Prolégomènes.*

Masumi, M. S. H. *Ibn Bajjah's ʿIlm al-Nafs.* Karachi: Pakistan Historical Society, 1961.

Nasr, Seyyed Hossein. *An Introduction to Islamic Cosmological Doctrines: Conceptions of Nature and Methods Used for Its Study by Ikhwān al Safā, al-Bīrūnī, and Ibn Sīnā.* Cambridge, MA: Belknap Press of Harvard University Press, 1964.

Nasr, Seyyed Hossein. *Three Muslim Sages.* Cambridge, MA: Harvard University Press, 1963. Discusses Avicenna, Suhrawardī, and Ibn al-Arabī.

Rahman, Fazlur. "Dream, Imagination and *ʿĀlam al-Mithāl.*" Islamic Studies 3 (2) (1964).

Rahman, Fazlur. *Prophecy in Islam.* London: Allen & Unwin, 1958.

Sharif, M. M., ed. *A History of Islamic Philosophy.* Vol. I. Wiesbaden, 1963. Published by the Pakistan Philosophical Congress. Essays on all aspects of the history of Islamic thought. Contains much useful information but has scholarship of uneven quality.

Walzer, Richard. *Greek into Arabic: Essays on Islamic Philosophy.* Cambridge, MA: Harvard University Press, 1962. Reprints selected essays of the author, including "Islamic Philosophy."

Fazlur Rahman (1967)

ISLAMIC PHILOSOPHY [ADDENDUM]

MYSTICISM (SUFISM)

Mysticism is of enormous significance in Islamic philosophy. Few Islamic philosophers were not committed to some form or another of mysticism. Ibn Sabʿīn (1217–1270) led an important school of thought that argued that Aristotelian philosophy and logic were of no use in understanding the way things really are. Logical thought is analytical, a process of dividing a concept into its parts. This method fails to represent the basic unity and wholeness that exists in reality, a unity that reflects the unity of God and everything as part of God, and any system of thought that is accurate is one that is based on unity, not division. The most influential mystical thinker Ibn al-ʿArabī (1165–1240) established this line of thought and represented himself as burying the old Peripatetic form of thought when he transported the bones of Ibn Rushd (Averroes) back to al-Andalus from North Africa.

This antagonism to analytical philosophy is certainly not the only position adopted by those committed to mysticism, though. Many philosophers managed to combine mysticism with peripatetic philosophy, arguing that they were just alternative philosophical methodologies, with different objects of thought. Peripatetic thought deals with the natural world and science, while mysticism goes deeper, and investigates the inner and the secret. The philosophers talked about a science of mysticism, and adopted a systematic attitude to this approach to understanding the nature of reality so that it is seen to represent something different from the exercise of subjective feelings. Sufism takes the searcher after knowledge further, it allows the individual to develop and understand significant experiences as well as valid concepts.

ILLUMINATIONIST (ISHRĀQĪ) PHILOSOPHY

Illuminationist (ishrāqīi) thought comes from the term *ishrāq*, a term linked with the idea of the east, and like Sufism was often adopted by thinkers who combined it with peripatetic thought. There is a more extreme form of illuminationist thought, though, that opposes peripatetic thought by attacking the crucial notion of definition, and using in its place immediate or intuitive knowledge. The peripatetic approach to reasoning has at its core the idea of definition in terms of genus and differentia, a process of explaining something by breaking it down into its smaller parts. Illuminationist thinkers such as al-Suhrawardī (1154–1191) criticize this approach as an attempt at explaining the unknown in terms of something even less known than itself, because the parts of the definition themselves will require definition, and so on ad infinitum. These criticisms also apply to deductive knowledge itself because this sort of knowledge is based on analysis and definition, and so operates on principles that are not philosophically respectable. Aristotelian demonstration is supposed to be the gold standard of argument and proof, and thus attacking it is an effective

way of suggesting that an alternative approach is necessary. The ishrāqī thinkers replace demonstration with knowledge by presence, knowledge that is so immediate that it cannot be doubted. If this knowledge is completely trustworthy, then it is highly appropriate as the starting point of one's reasoning process.

According to many ishrāqī thinkers, there is a type of knowledge that is so self-evident that it cannot be doubted. What counts as self-evident knowledge for the ishrāqī thinkers is that level of knowledge that is so intimately tied in with peoples' perception of themselves that in doubting it they would doubt themselves, and that implies doubting what makes the doubting possible in the first place. The conclusion is taken to be that such doubt is impossible. The truth that is presupposed by any perception is that the subject of perception exists. It is perhaps Suhrawardi who explores this notion of immediate knowledge, *ʿilm al-ḥuḍūrī*, most precisely; he argues that immediate knowledge is so immediate and incontrovertible that it is known in far more than an intellectual sense. That is, there are propositions that are known through reason and that are known perfectly, in the sense that humans grasp all aspects of them and can hold them in their minds all at once perfectly. The sorts of knowledge that are called *ʿilm al-ḥuḍūrī* are not only indubitable, but people experience their indubitability. The light of knowledge that shines on them makes the truth they possess evident to people in more than merely an intellectual sense. Of course, another advantage that perception of the self has over discursive knowledge is that the assumption is made that the self is basically a simple thing, so the use of human intelligence implies the activity of a simple self, a self that is pure agency.

The key term in illuminationist thought, as its name suggests, is light, and the idea is that immediate knowledge is lit up or illuminated in such a way to make it impossible to doubt. Light is commonly experienced as pervasive throughout the universe, and because people notice things clearly if they are brightly illuminated, this concept has familiarity as its basis. Yet people also tend to think that the things that are illuminated already exist before the light strikes them. For most of the ishrāqīs, what things are depends on degrees of their luminosity or light, not on their essences. God is often identified with the Light of Lights, the light that is the source of all other light and that does not itself receive light.

Illuminationism is often combined with both Sufism and Peripateticism, although more often with the former than with the latter. Although it sounds like a mystical approach, it is actually often carried out with analytical exactitude, and there is a good deal of controversy in the literature as to whether ishrāqī thought represents a long-standing esoteric tradition in philosophy, or whether it is basically logical in the widest sense of that term. Ishrāqī thought is largely limited to the Persian cultural world, and so there is a temptation to see in it something linked with early Iranian ideas such as Zoroastrianism because light was also important there. However, the argument for such influence is difficult to make plausible.

It is difficult to overemphasize the significance of Persians in Islamic philosophy, although they generally wrote in Arabic. Most Islamic philosophers were Persians—albeit often living in other parts of the Islamic world—and Persia has a long history of absorbing philosophical ideas from a wide variety of sources. Mullā Ṣadrā is with little doubt the most outstanding thinker to have emerged from Persia, and his thought has defined the Persian philosophical curriculum since his day. Unlike some ishrāqī thinkers, such as al-Suhrawardī, Mullā Ṣadrā suggested that existence precedes essence, and so the first question in ontology deals with the characteristics of what exists, not with what they would need to be like to exist. He argued that existence is equivalent to God and so when people talk of ordinary things existing, they are really describing a relationship that they have with God. This is always going to be difficult to describe because ordinary language is based on a form of existence that it itself cannot explain. What is required for people to understand existence is first of all to comprehend all the different ways in which things exist, how those different forms of existence are linked, and the nature of the intermediaries between the different levels of existence. This concern for grasping the unity of being links Mullā Ṣadrā with the mysticism of Ibn al-ʿArabī, while the desire to understand the nature of the different levels of existence involves the sort of analysis found in the illuminationist tradition.

CONTEMPORARY ISLAMIC PHILOSOPHY

There are a variety of different approaches to philosophy in the contemporary Islamic world—as is hardly surprising—but some themes do recur. One of these themes comes from the notion of Islamic philosophy and deals with how Islamic it should be and what links it has to have with Islam itself, if any. How does Islamic philosophy's taking place within a particular cultural context shape it? How far should it continue to shape it? Although many Islamic philosophers continue to use techniques and ideas from outside of the region, the links that should

be established between Islam and the rest of the world is frequently a contentious issue. It is worth adding that the Islamic world can no longer, if it ever could, be identified with a specific geographical area because Islamic thinkers are today to be found virtually everywhere in the world.

The issue of the links between Islam and the rest of the world has persisted for some time and was highlighted in the nineteenth century as a result of orientalism and colonialism. In the subsequent centuries this has become an even hotter topic, because globalization and Zionism have been seen as yet another assault on the Islamic world and its distinct ethos. In the past the Islamic world was far in advance of the rest of the world, and yet for many centuries this has been entirely reversed and has led to many debates about the sources and significance of this apparent relative decline. It has been taken to be more than just a social or economic issue; it is a cultural one also, and clearly philosophy is then relevant in trying to resolve it.

THE RENAISSANCE IN ISLAMIC PHILOSOPHY. The *Nahḍa* or Islamic renaissance that started in the nineteenth century and became significant in Egypt in particular played a large role in shaping the modern debate in the Arab world. The *Nahḍa* involved a strategy of maintaining a distinctive Islamic identity while also incorporating within Islamic society those aspects of modernity that are not incompatible with religion. The idea was to combine the traditions of Islam with what was desirable from the rest of the world, in particular science but also aspects of culture such as philosophy.

It is often said that Islam never went through an enlightenment, yet the *Nahḍa* movement was clearly a sort of Enlightenment, albeit one that unlike some advocates of the European Enlightenment did not involve hostility to traditional religion. Jamāl al-Dīn al-Afghānī (1838/9–1897) and Muḥammad ʿAbduh both argued that Islam is perfectly rational and in no way opposed by European and North American scientific and cultural ideas, so there is no problem in not using those ideas. The Egyptian philosopher Muṣṭafā ʿAbd al-Rāziq went so far as to argue that all the main Islamic schools of thought are inherently rational and in no way inimical to European and North American science and rationality. By contrast, Muḥammad ʿAbd al-Jābrī is critical of much traditional Islamic thought, arguing that the reasons for the decline of the Arab world need to be analyzed clearly. He calls for a reexamination of the argument between those who emphasize the glory of the Islamic past and those who praise European and North American modernity. What is required is a liberation of the Arab consciousness from its traditional ties to its Islamic past, and yet also a cautious attitude to the ideas that have come from Europe and North America and are aspects of foreign domination.

Fuʾād Zakariyya agrees that Arab failure is linked with the failure to criticize tradition, while Fazlur Rahman outlines the links between Islam and social progress. He argues that Islamic traditionalism is opposed to Islam itself because the religion is in favor of economic and social development and change. The attempt to fix a rigid and stultified version of Islam as the ideal is to fail to understand how science and technology can improve the lifestyle and moral welfare of the community. Hasan Hanafi uses the methodology of phenomenology to describe the concept of *tawḥīd* or unity. He suggests that Islam is dynamic enough to extend this notion so that it may provide a generally acceptable principle of unity and equality for everyone, something we can observe by examining how unity is actually used within contemporary and prior Islamic culture. He is also critical of blind faith in European and North American progress, suggesting that Europe and North America are now entering into a period of decadence that will require an infusion of ideas from elsewhere and in particular the Middle East. The idea that Islam is based on fixed rules he finds unrealistic, it is based on a revelation appropriate at its own time and place, but now other interpretations of the message should be adopted to match present conditions and represent more accurately the dynamism of Islam.

PHILOSOPHY'S PRESENCE IN THE ISLAMIC WORLD. It is often said that philosophy declined in the Islamic world after the death of Ibn Rushd in the twelfth century, but this is far from the truth. Today there is a lively philosophical presence in most of the Islamic world, often with the infusion into Islamic philosophy of ideas such as logical positivism, hermeneutics, pragmatism, Hegelianism, deconstructionism, and so on. Philosophy continued vigorously in the Persian cultural world, especially the philosophy of Ibn Sīnā and the ishrāqī (illuminationist) thinkers developing and commenting on al-Suhrawardī and Mullā Ṣadrā. In Iran philosophy has now moved away from the theological school, the *madrasa*, into the university. A good example of this is represented by the thought of Mehdi Haʾiri Yazdi. He develops a complex theory of knowledge that is based on knowledge by presence, a form of knowledge that is immediate and incorrigible and that serves as the foundation of other knowledge claims. He uses ideas from both ishrāqī thinkers like al-Suhrawardī, and the modern philosopher Wittgenstein.

Rather similar to Hanafi's phenomenology, ʿAlī Shariati uses the ishrāqī school's intermediary position between mysticism and Peripateticism to develop a view of humanity having God at its essence while maintaining the scope to determine its own form of existence. The notion of unity (*tawḥīd*) is seen as therapeutic—it links both personal and political justice and harmony. He interprets the main figures of Shiʿite Islam as models for people not only in a personal sense but also to bring about more progressive social ideals, and he sees them as representing archetypes that have always been regarded as desirable. Over time these archetypes themselves have not changed in essence, but they have changed in appearance to make them more appropriate to the particular audiences for whom they are designed.

This kind of link of the personal and the political is significant in modern Persian thought. An excellent example is Ayatollah Khomeini (Khumaynī), who led the Islamic revolution in Iran and combined the roles of spiritual and temporal ruler of the Islamic Republic of Iran. He wrote about and practiced a political philosophy that has become much discussed today in the Islamic world and beyond. The arguments for theocracy have become familiar again in political philosophy as a result of the Islamic revolution and its theoretical context.

CONTEXT FOR THE ISLAMIC REVOLUTION. What is the Islamic revolution's theoretical context? In his account of the Islamic state, Khomeini follows the familiar strategy of reflecting on the past and what took place then. He points out that the Prophet instituted a practical way of life as well as transmitting God's message, and also made provision for his successors. (It is on this issue of successors that the big divide in Islam between the Sunni and the Shiʿa occurred, the latter believing that the Prophet's son-in-law ʿAlī was his legitimate successor as head of the polity.) Islam is not only a religion in the sense of a system of belief about spiritual issues, but it also includes specific rules about how people ought to live, and if they can live in those ways then the legal and material provisions of the state must support that form of life.

There are three strong arguments for the continuing relevance of Islamic government. As Khomeini points out, the claim that Islamic government was only appropriate in earlier times might be taken as equivalent to the thesis that Islam itself is only valid at earlier times. Also, if Islam were not supposed to be a comprehensive and constant legislative system, why would it in fact consist of such detailed prescriptions? Finally, had there existed a unified Islamic polity, the constant humiliations of the Islamic world at the hands of its enemies would not have taken place. This is a reference not only to the creation of the State of Israel and its continuing dominance in the Middle East, but also to the repugnant actions of groups such as the Jews, according to Khomeini.

From an Islamic perspective, the state is not neutral. States are either Islamic or founded on unbelief and corruption. The Muslim cannot live in the latter kind of state without being irretrievably affected by it, unless he actively opposes it. It is the duty of all Muslims to struggle against the state unless the state is Islamic. Khomeini denounces the division of the international Islamic community into individual states, one of the effects of imperialism to weaken and divide Islam, in his view. Everything in Islam, he argues, is opposed to injustice, and yet, he argues, we see injustice in what is called the Islamic world. What is needed is the overturning of the corrupt regimes and their replacement by real Islamic governments.

A theme of Khomeini's thought is that religion does not just apply to private morality but must also be applied to the state as a whole, and the religious authorities should be in charge of the state because only then will the community be rightly guided. The school of Qom, of which he was a member, contained also Muhammad Hossein Tabatabaʾi, Murtaza Mutahheri, and Muhammad Taqi Misbah Yazdi, all important religious Shiʿite thinkers who nonetheless did not reject ideas just because they came from Europe. They argued that traditional Islamic philosophy could only gain by opening itself to some of the important philosophical achievements created outside of the Islamic world. All of the main religious thinkers in Iran disapproved of the work of Abdul Soroush, who took a rather distanced view of religion when he applied what he took to be the arguments of Popper, Moore and Wittgenstein to them. Soroush was opposed by Sadiq Larijani, the chief representative of the School of Qom, who suggested that Soroush had misapplied the theories of Popper, Stalnaker, Watkins, and Hempel. It is interesting that the debate took the form not of the clash between religion and reason, but rather of the correct understanding of philosophical theories, although it is fairly clear that there are serious issues of the role of religion in philosophy implicitly in the debate. Soroush managed to infuriate both the school of Qom and also the supporters of Heidegger, and that left him thoroughly isolated intellectually in Iran.

THE FUTURE OF ISLAMIC PHILOSOPHY Perhaps the best-known Iranian thinker outside the country in the early twenty-first century is Seyyed Hossein Nasr. He enters the debate on modernity by being critical of European and North American science—he praises some of its material achievements but points to the ecological consequences of a worldview that does not base itself on the presence of God. Science without spirituality is blind to moral issues, Nasr believes, because there is nothing that it holds sacred; it bases itself entirely on measurements of quantities, not on the quality of existence. More spiritual philosophies such as those based on Islam are holistic and integrative; they embed spiritual values in the technological agenda and so make ecological disasters less likely. For Nasr, the main question is not what the Middle East should take from Europe and North America, but vice versa.

Along with this view, Nasr has established in some detail the theoretical presuppositions of Sufism, the school of mysticism in Islamic thought. His historical accounts of this doctrine have played a large role in its increasing domestication outside of the traditional Islamic world. Indeed, as the Islamic world spreads out ever more widely, it is likely to involve itself much more in the ideas that it finds in an originally non-Islamic source. In this way Islamic philosophy is returning to its roots, in a sense, because it was the meeting of Islam with Greek philosophy in the early years of Islam that led to the subject coming into existence in the first place.

See also Aristotelianism; Averroes; Avicenna; Ibn al-ʿArabī; Illuminationism; Enlightenment, Islamic; Hanafi, Hassan; Moore, George Edward; Mullā Ṣadrā; Mysticism, History of; Nasr, Seyyed Hossein; Peripatetics; Popper, Karl Raimund; Sufism; Suhrawardī, Shihāb al-Dīn Yaḥyā; Wittgenstein, Ludwig Josef Johann; Zoroastrianism.

Bibliography

A source of general bibliographical information may be found in Oliver Leaman's "A Guide to Bibliographical Resources," in Seyyed H. Nasr's and Oliver Leaman's *History of Islamic Philosophy*, 1173–1179 (London: Routledge, 1996).

Al-Suhrawardī. *The Philosophy of Illumination*. Translated by John Walbridge and Hossein Ziai. Provo, UT: Brigham Young University Press, 1999.

Fakhry, Majid. *A History of Islamic Philosophy*. 2nd ed. New York: Columbia University Press, 1987.

Fakhry, Majid. *A Short Introduction to Islamic Philosophy, Theology, and Mysticism*. Oxford: One World, 1997.

Hahn, Louis, et al., eds. *The Philosophy of Seyyed Hossein Nasr*. Chicago: Open Court, 2001.

Hourani, George. *Reason and Tradition in Islamic Ethics*. Cambridge, U.K.: Cambridge University Press, 1985.

Leaman, Oliver. *Brief Introduction to Islamic Philosophy*. Oxford: Polity Press, 1999.

Leaman, Oliver. *Introduction to Classical Islamic Philosophy*. Cambridge, U.K.: Cambridge University Press, 2002.

Nasr, Seyyed H., and Oliver Leaman, eds. *History of Islamic Philosophy*. Chapters 61, 63–65, 69, 71. London: Routledge, 1996.

Netton, Ian. *Allah Transcendent* London: Routledge, 1989.

Rahman, Fazlur. *Islam and Modernity: Transformation of an Intellectual Tradition*. Chicago: University of Chicago Press, 1982.

Yazdi, M. Ha'iri. *The Principles of Epistemology in Islamic Philosophy: Knowledge by Presence*. Albany: State University of New York Press, 1992.

Oliver Leaman (2005)

ISRAELI, ISAAC BEN SOLOMON
(c. 855–955)

Isaac ben Solomon Israeli, the first Jewish Neoplatonist, was one of the most distinguished Jewish physicians of the Middle Ages. He was so renowned for his medical competence, both in theory and in practice, that his works were widely circulated in manuscript, translated into Latin, and printed in the early years of the sixteenth century, as *Omnia Opera Ysaac* (Lyons, 1515). This printed edition and the manuscripts on which it was based contained some of Israeli's philosophic writings as well as his scientific treatises. As a result, his name became well-known, beyond his philosophic deserts; indeed, his fame among Christian scholars was second only to that of Moses Maimonides. Yet Maimonides held Israeli's philosophy in no great esteem, referring to him as "merely a physician."

Isaac Israeli was a native of Egypt. He left his native land to study medicine in the intellectual center of Kairouan, in north Africa, under the tutelage of Ishaq ibn Imram, a Muslim. Later Israeli served as court physician to Ubaydullah al-Mahdi, founder of the Fatimid dynasty in north Africa.

In addition to the philosophic materials in his "Book of Elements" (a medical work), Israeli has long been known as the writer of a "Book of Definitions." Recent studies have added also a "Book of Substances," a "Book on Spirit and Soul," and, probably, a short "Chapter on the Elements," found in a unique manuscript in the Bibliotheca Communale of Mantua and ascribed to Isaac Israeli by A. Altmann. On the basis of these works, Israeli

can be confidently classified as a Neoplatonist whose work is akin to that of other Neoplatonists among the Muslim philosophers of his age.

His surviving works do not include any significant discussion of the existence and nature of God but they do describe God as a perpetually active Creator. God's original creative act is a creation out of nothing; later acts of creativity in nature are not of the same order but are "the passing of corporeal substances from privation to existence" in accordance with God's will. Along with this account, however, Israeli also maintained a doctrine of emanation. Thus, on the one hand God creates because of his goodness, while on the other his creativity is a perpetual overflowing. These two accounts of creation are never reconciled in Israeli's thought.

The process of emanation terminates with the emergence of the visible sphere. From this point, Israeli's explanation of the universe is physical and more closely akin to the views of Aristotle. Retaining the classical Greek theory of the four elements, he accounted for everything in the world of our experience by the combination of the elements earth, air, fire, and water. Once again, however, we are confronted with an uncertainty. In the "Book of Definitions," Israeli asserted that the four elements came into being through the movement of the sphere of heaven, but in the "Book of Elements" they are attributed to the power of God. Except by straining the language, these two views cannot be reconciled.

A similar double view emerges in Israeli's doctrine of the soul. Here he spoke of a cosmic soul, which exists independently of body, appearing in three successive stages of emanation—rational, animal, and vegetable—and also of a divine spark within the individual, striving ever upward toward the cosmic soul. Perhaps in this double account of soul we have a reflection of the Neoplatonic doctrine of man as the microcosm. If so, we can understand the emphasis Israeli put on self-knowledge, the road to the knowledge of the universe. Self-knowledge is knowledge of both body and soul; one who knows himself in both soul and body knows everything, and he alone is worthy of the name of philosopher.

See also Aristotle; Jewish Philosophy; Maimonides; Neoplatonism; Self-Knowledge.

Bibliography

WORKS BY ISRAELI

Kitab al-jawahir (Book of Substances), edited by S. M. Stern, "The Fragments of Isaac Israeli's 'Book of Substances.'" *Journal of Jewish Studies* 7 (1956): 13–29.

Collected Works, edited by A. Altmann and S. M. Stern. *Isaac Israeli. A Neoplatonic Philosopher of the Early Tenth Century: His Works Translated with Comments and an Outline of his Philosophy.* London: Oxford University Press, 1958.

WORKS ON ISRAELI

Altmann, A. "Creation and Emanation in Isaac Israeli: A Reappraisal." In *Studies in Medieval Jewish History and Literature,* edited by I. Twersky, 1–15. Cambridge, MA: Harvard University Press, 1979.

Pessin, S. "Jewish Neoplatonism: Being above Being and Divine Emanation in Solomon ibn Gabirol and Isaac Israeli." In *Cambridge Companion to Medieval Jewish Philosophy,* edited by D. Frank and O. Leaman, 91–110. Cambridge, U.K.: Cambridge University Press, 2003.

Samuelson, N. "Medieval Jewish Aristotelianism: An Introduction." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman. London: Routledge, 1997.

Wolfson, H. "The Meaning of Ex Nihilo in Isaac Israeli." *Jewish Quarterly Review* 50 (1959): 1–12.

J. L. Blau (1967)
Bibliography updated by Oliver Leaman (2005)

ITŌ JINSAI

(1627–1705)

Itō Jinsai, a Japanese Confucianist of the *kogakuha* ("school of ancient learning"), was born in Kyoto, the son of a poor merchant, and spent his life there as an educator. After studying the official Zhu Xi Confucianist doctrine, he rediscovered ancient Confucianism and became its systematizer and, through the Kogidō, a school he founded in 1680, its propagator. The novelty of his teaching aroused the suspicion of the central government in Edo (Tokyo). However, it was not suppressed although his *kogigaku,* or "learning-of-the-ancient-meaning," was gaining a large following. Through the able guidance of his scholarly son, Tōgai, and of his grandson the school was operated until 1871, when all Confucianist schools were abolished in favor of the new Western system.

Itō's philosophy, stemming from a great admiration for Confucius and Mencius, is quite contrary to the Neo-Confucianism of Zhu Xi. Itō is clearly a monist in the sense that he does not admit any priority of *ri,* the principle (reason), over *ki,* the material force, which for him is material energy. A primordial material energy (*ichi genki*), having neither beginning nor end, is the root of everything. *Ri* is but a pattern of *ki; ki,* through the motion of the yin-yang, or passive-active, elements, forms the great living organism (*dai-katsubutsu*), the universe itself.

Itō holds with Mencius that human nature is originally good, and he does not make the usual Zhu Xi distinction between physical and original nature, which he treats as a spurious Daoist influence. Evil in physical nature need not be explained as if it arose from lack of cultivation of the potentialities of human nature. The four sources of virtue (in Chinese, *ssu tuan*; in Japanese, *shitan*) according to Itō are righteousness, humaneness, ritual or propriety, and wisdom. Righteousness is the pivotal virtue of Itō's ethics. Humaneness is benevolent love, or condescension from the superior to the inferior, for in Confucianism universal equalitarian love is practically nonexistent. Morality, the natural Way of things, has a cosmological meaning in addition to the ethical one. The material energy of the universe is manifested in humankind through humaneness or love. Itō's principles of education centered on forming moral character rather than on imparting knowledge; will is above the intellect.

Itō did not make much of astronomy and mathematics, but he was very fond of history. However, unlike most other Confucianists of the "ancient learning" school, he did not become a nationalist through the study of history. For him China remained the fountainhead of culture. Itō's outstanding merits as a Sinologist were the result of painstaking research in ancient texts, yet he patiently bore the faultfinding of his gifted son and the criticisms of his best pupil, Namikawa Temmin (1679–1718).

See also Chinese Philosophy; Confucius; Human Nature; Mencius; Wisdom; Zhu Xi (Chu Hsi).

Bibliography

Itō's chief works can be found in collections in Japanese, including Inoue Tetsujirō, ed., *Nihon rinri ihen* (Library on Japanese ethics; Tokyo: Ikuseikai, 1901), Vol. V, pp. 11–181, and *Dai Nihon shisō zenshū* (Collected works on the thought of great Japan; Tokyo, 1934), Vol. XLI, pp. 7–249. See also Ishida Ichirō, *Itō Jinsai* (Tokyo, 1964), which is in Japanese.

For works in English see J. J. Spae, *Itō Jinsai: A Philosopher, Educator* and *Sinologist of the Tokugawa Period* (Beijing: Catholic University of Beijing, 1948), and W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 410–422. The second work contains selections in translation with introductions.

Gino K. Piovesana, S.J. (1967)

IVANOV, VIACHESLAV IVANOVICH
(1866–1949)

A major poet and theorist of the symbolist literary movement in Russia, Viacheslav Ivanov left an elaborate and influential body of work on art, culture, and religion. To the chagrin of his contemporaries, he never formalized his protean and wide-ranging ideas as a philosophical system. However, Ivanov's writings can be divided into several core areas, which succeeded each other at the center of his attention: the ritual roots of tragedy; the artwork as symbol of the transcendent; the role of art in creating historical myth; and the prospects for a religious revival in modernity. Despite his protean views, Ivanov can be seen as a philosopher in the hermeneutic tradition for whom the world reveals itself as an historical continuum of discrete acts of expression and understanding.

BIOGRAPHY

Born in Moscow, Ivanov was educated as a classical historian and philologist at the universities of Moscow (1884–1886) and Berlin (1886–1891). In 1895 he abandoned his academic career and devoted himself to poetry. This awakening was instigated by his move to Italy, his adulterous affair with Lidiia Zinov'eva-Annibal (who became his second wife in 1899), and his discovery of Friedrich Nietzsche. Ivanov entered into contact with the philosopher Vladimir Solov'ëv, who approved some of Ivanov's poems and the title of his first book of poetry *Kormchie zvezdy* (Pilot Stars, 1902). Ivanov followed up on his poetic debut with a series of lectures for Russians in Paris called *Ellinskaia religiia stradaiushchego boga* (The Hellenic Religion of the Suffering God, published 1904; continued in 1905 as *Religiia Dionisa* [The Religion of Dionysus]).

After completing another book of poetry, *Prozrachnost'* (Transparency, 1904), Ivanov moved to St. Petersburg in the revolutionary year 1905. Between 1905 and 1912, Ivanov hosted weekly symposia at his "Tower" apartment which attracted many major writers, artists, and thinkers. The first session, for example, was devoted to the question of "Eros" and chaired by philosopher Nikolai Berdyaev. Ivanov was also active at the St Petersburg Religious-Philosophical Society. His essays from 1904 to 1909 were gathered into the volume *Po zvezdam* (By the Stars, 1909). After his wife's sudden death in October 1907, Ivanov commemorated her in an elaborate poetic cult, the highpoint of which is marked by the two-volume book of poetry *Cor Ardens* (1911–1912).

In 1912 Ivanov married Lidiia's daughter from a previous marriage Vera Shvarsalon, who gave birth to a son Dmitrii; these events inspired Ivanov's 1912 collection *Nezhnaia taina* (The Tender Mystery). Moving to Moscow in 1913, Ivanov became active in the Moscow Religious-Philosophical Society and became close to such "neo-Slavophile" philosophers as Pavel Florenskii, Sergei Bulgakov, and Vladimir Ern. Ivanov identified himself closely with the memory of Fëdor Dostoevsky, on whom he wrote extensively. He also became a friend of the mystical composer Aleksandr Scriabin, on whom Ivanov wrote a series of poems and essays.

During World War I Ivanov published two books of essays: *Borozdy i mezhi* (Furrows and Boundaries, 1916) and *Rodnoe i vselenskoe* (Matters Native and Universal, 1917). In the latter Ivanov's strident political tone reflects his enthusiastic embrace of the February 1917 revolution and his initial opposition to the Bolshevik revolution in October 1917. However by late 1918 Ivanov had assumed an important position in the cultural organs of the fledgling Soviet state and published two earlier works: the autobiographical narrative poem *Mladenchestvo* (Infancy, 1918) and the drama *Prometei* (Prometheus, 1919).

After Vera's death in August 1920 Ivanov moved to Baku, where he taught at the new Azerbaijan State University for four years. In 1921 Ivanov defended his doctoral dissertation, *Dionis i pradionisiistvo* (Dionysus and Pre-Dionysianism, 1923), a more rigorous elaboration of his earlier ideas about Greek religion and tragedy. The Baku period was comparatively barren of original work, especially poetry; Ivanov wrote only a satirical drama *Liubov'—mirazh?* (Is Love a Mirage?, 1924). In 1924 he emigrated to Italy, where in 1926 he became a Roman Catholic. From 1926 to 1935 Ivanov taught at Collegio Borromeo in Pavia.

Ivanov achieved some renown in European intellectual circles between the wars. Most notably, *Perepiska iz dvukh uglov* (The Correspondence from two corners), which Ivanov coauthored in 1920 with cultural historian Mikhail Gershenzon, was translated into numerous languages beginning in 1926. In 1932 he reworked and translated his essays on Fë Dostoevsky as *Dostojewkskij: Tragödie—Myth—Mystik* (Dostoevsky: tragedy, myth, mysticism; translated as *Freedom and the Tragic Life* [1952]), the single best introduction to Ivanov's thought. In 1939 he published *Chelovek* (Man), a philosophical poem written mainly between 1915 and 1919. In 1944 he kept a lyric diary which was included in his posthumous book of poetry *Svet vechernii* (The fading light, 1962).

IVANOV'S THOUGHT

Influenced by Arthur Schopenhauer and by Nietzsche, in his early metaphysics Ivanov viewed the physical world as a veil of *Maya* or nonbeing, which can be overcome only in cathartic ritual or ritual-like tragedy. In Ivanov's initial aesthetic statements, the ineffable transcendent event resisted concrete expression, and so the emphasis fell squarely on the psychological transformation of artist and beholder in mimetic performance. Following Richard Wagner, Ivanov projected the renewal of tragedy as a synthesis of the existing arts that would lead to a religious revival. In particular Ivanov equated the rebirth of the tragic chorus in art to the achievement of *sobornost'*, the spiritual unity of the believing community or nation. Ivanov christianized Nietzsche, identifying the suffering god Dionysus with Christ and ancient tragedy with the Christian liturgy.

In politics Ivanov elaborated a theory of mystical anarchism, predicated on the expectation that the community would be unified inwardly by common ritual practice in symphonic unanimity, or *sobornost'*, instead of by formal legal and political structures. He constructed a cyclical theory of cultural history in which periods of critical or classical culture alternate with organic periods of barbarian (i.e., Scythian, Anglo-Saxon, and Slavic) energy. Ivanov counted the impending rebirth of tragedy, the synthesis of the arts, and the revolutionary tumult of 1905 among the symptoms of a new organic era of mystical activity that would lead to a just society and a reinvigorated church.

By 1908 Ivanov had shifted his attention from tragic and ritual performance to its concretization in aesthetic symbols and religious dogma. His aesthetics became increasingly metaphysical, drawing especially on Plato and Vladimir Solov'ëv, from whom he took the term *theurgy* to denote the artist's transfiguration of phenomena into an ontologically higher reality (the symbol) that approximates the divine prototype. Ivanov described artistic creation and reception as an ascent *a realibus ad realiora* (from the real to the more real). Adopting linguistic terminology, and referring to neo-Kantian philosopher Heinrich Rickert, Ivanov claimed that any proposition is based on the verb "to be" and is a normative projection of being as value. In religious terms, any proposition imbues reality with an assertion of divine being. Therefore Ivanov posited that the statement "Thou art" actually elevates the being of both speaker and addressee through the energies of God contained in language itself.

Ivanov's concepts of catharsis and the linguistic symbol led to a communicative philosophy that took its final shape around the time of his move to Moscow in 1913. In

"On the Limits of Art" (1914), a major restatement of his aesthetics, Ivanov described the aesthetic process as a continuum of expressive and receptive acts, in which art stimulates individuals to further creativity, leading to the gradual transfiguration of reality through human agency (instead of through semimagical theurgy). Ivanov still attributed a central role to the artist in guiding the individual's transformation; however, he was now eager to describe how the transcendent revelation of the art work is transcribed into a narrative myth that communicates memory and projects future human action.

At this time Ivanov also integrated his aesthetics with an account of history. In a series of articles on the history of literature, Ivanov described how Byron's texts contributed to the development of individual consciousness in Russia, which in turn allowed Pushkin and then Dostoevsky to re-appropriate Russian history and spirituality from within modernity. Defining Dostoevsky's major works as "novel-tragedies" allowed Ivanov to explain both their cathartic grip on readers and their ideological influence. In his philosophical and artistic prose, Ivanov elaborated a new myth of Russian history, which he hoped would result in the country's transformation into a truly Christian empire.

In a 1909 essay "On the Russian Idea" Ivanov described the complex interaction of understanding and action in terms of an Aristotelian triad: catharsis (cleansing), mathesis (learning), and praxis (action). This hermeneutic standpoint received its most accomplished expression in *The Correspondence from Two Corners* (1920), an epistolary exchange between Ivanov and Mikhail Gershenzon. Surrounded by the ruins of their former world, both authors grappled with their own lives by inscribing their projected identities into a text, which is immediately read and answered by the other.

In his Italian exile Ivanov adjusted earlier ideas and constructs in the light of his Roman Catholicism. Like Solov'ëv, Ivanov explained his conversion to Catholicism as an affirmation of the Roman Catholic Church as a historical symbol of divine unity. He adopted some of Jacques Maritain's neo-Thomist vocabulary, for example defining art as *transparentia formae*.

In his heyday, Ivanov's intellectual constructs enjoyed broad renown and were key influences on such thinkers as Nikolai Berdyaev, Pavel Florenskii, Sergei Bulgakov, Aleksei Losev, and Mikhail Bakhtin. His views on tragic performance as a social panacea influenced the public celebrations in the early Soviet Union. His impact has also been felt in Orthodox theology, which has sometimes adopted his formulations of the symbol, the idea of aesthetic ascent and descent, and the primacy of ritual experience in generating *sobornost'*. His conception of culture as a historical continuum of creative acts remains an underappreciated aspect of his work.

See also Hermeneutics; Russian Philosophy.

Bibliography

PRIMARY WORKS

Sobranie sochinenii. 4 vols. Brussels: Foyer oriental chrétien, 1971–1987.

With Mikhail Gershenzon. *Correspondence across a Room*. Translated by Lisa Sergio. Marlboro, VT: Marlboro Press, 1984.

Freedom and the Tragic Life: A Study in Dostoevsky. Introduction by Robert Louis Jackson. Wolfeboro, NH: Longwood Academic, 1989.

Selected Essays. Translated and with notes by Robert Bird, edited with an introduction by Michael Wachtel. Evanston, IL: Northwestern University Press, 2001.

SECONDARY WORKS

Bird, Robert. "Martin Heidegger and Russian Symbolist Philosophy." *Studies in East European Thought* 51 (3) (June 1999): 85–108.

Bird, Robert. *The Russian Prospero: The Creative Universe of Viacheslav Ivanov*. Madison: University of Wisconsin Press, forthcoming.

Davidson, Pamela. *Viacheslav Ivanov: A Reference Guide*. New York: G. K. Hall, 1996.

Gogotishvili, L. A. "Mezhdu imenem i predikatom (simvolizm Viach Ivanova na fone imiaslaviia)." In *Viacheslav Ivanov: Arkhivnye materialy i issledovaniia*, edited by A. A. Gogotishvili and A. T. Kazarian, 303–382. Moscow: Russkie slovari, 1999.

Jackson, Robert Louis, and Lowry Nelson Jr., eds. *Vyacheslav Ivanov: Poet, Critic, and Philosopher*. New Haven, CT: Yale Center for International and Area Studies, 1986.

Marchenkov, Vladimir. "Vladimir Solov'ëv and Vjačeslav Ivanov: Two Theurgic Mythosophies. " In *Vladimir Solov'ëv: Reconciler and Polemicist*, edited by Wil van den Bercken, Manon de Courten, and Evert van der Zweerde, 211–221. Leuven, Paris; Sterling, VA: Peeters, 2000.

Meerson, Michael. *The Trinity of Love in Modern Russian Theology: The Love Paradigm and the Retrieval of Western Medieval Love Mysticism (from Solovyov to Bulgakov)*. Quincy, IL: Franciscan Press, 1998.

Rizzi, Daniela, and Andrej Shishkin, eds. *Archivio russo-italiano III: Vjačeslav Ivanov – Testi inediti*. Salerno: Europa Orientalis, 2001.

Stepun, Fedor. *Mystische Weltschau: Fünf Gestalten des russischen Symbolismus*. Munich: Carl Hanser, 1964.

Wachtel, Michael, ed. *Vjačeslav Ivanov: Dichtung und Briefwechsel aus dem deutschsprachigen Nachlass*. Mainz, Germany: Liber-Verlag, 1995.

West, James. *Russian Symbolism: A Study of Vyacheslav Ivanov and the Russian symbolist aesthetic*. London: Methuen, 1970.

Robert Bird (2005)

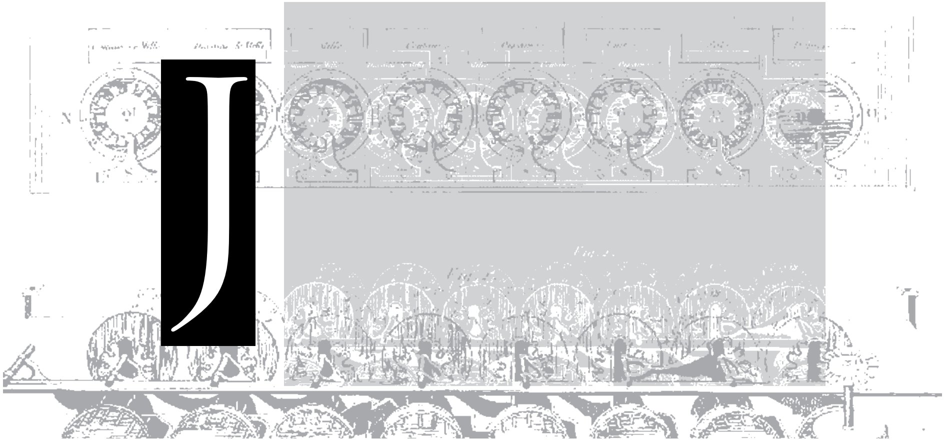

J

JACOBI, FRIEDRICH HEINRICH

(1743–1819)

Friedrich Heinrich Jacobi was a leading representative, with Johann Georg Hamann, of the philosophy of feeling and a major critic of Immanuel Kant. He was born in Düsseldorf on the Rhine. Jacobi received an education preparing him for a business career, but an inner urge drove him to the pursuit of philosophical studies. He studied the works of Claude-Adrien Helvétius, Jean-Jacques Rousseau, Ferguson, and Benedict de Spinoza, the last of which had a negative influence on him, provoking opposition and criticism; he was also influenced by the English philosophers of feeling—the earl of Shaftesbury and others. His friend Hamann, a kindred spirit, lived in his home for a long period, and his influence on Jacobi cannot be overestimated. In 1804, Jacobi was appointed president of the Academy of Sciences in Munich. He was in literary contact with the prominent thinkers of his time—Moses Mendelssohn, Karl Leonhard Reinhold, Jakob Friedrich Fries, and Johann Wolfgang von Goethe. His discussions with his contemporaries are as important for the understanding of his philosophy as are his original works.

Jacobi developed a philosophy of feeling and faith. He was critical of speculations leading to the concept of the prevalence of necessary laws above freedom, hence Jacobi's rejection of Spinoza's pantheism and of the philosophy of Johann Gottlieb Fichte, Friedrich von Schelling, and G. W. F. Hegel, in which there are manifest pantheistic tendencies. Because of Jacobi's concept of the primacy of freedom, he found that the actions of man are not to be deduced from his thinking, for thinking is not the primary force in man. The history of man is not the result of his mode of thought; rather, the former determines the latter. Herein is anticipated the method of the historical school of law as it was later developed by Friedrich Karl von Savigny. For Jacobi the immediately given is the determining factor in our cognition of cultural phenomena. Objects have to be given to us through immediate feeling or faith before thought comes into play. The task of discursive thinking is to observe, analyze, compare, and order perceptions by reducing them to their fundamental principles. But unless something real is previously given through feeling, discursive thinking cannot take place.

Jacobi was a master of criticism. His strength lay in grasping a system of thought as a whole and detecting those elements in it that are incompatible. This capacity

of critical analysis is manifest in his appraisal of dogmatic rationalism and the critical philosophy. Jacobi subjected both Spinoza and Kant to severe criticism. He pointed to hidden contradictions and inconsistencies in both their systems. The dogmatic rationalism of Spinoza employs the mathematical method in the realm of metaphysics; it accepts as real only what can be proven and deduced mathematically. By this method, however, neither God nor freedom can be maintained. These ideas cannot be deduced by an absolute system of causality, which is the essence of Spinozism. Absolute necessity leads to atheism, and the denial of freedom leads to fatalism. To Jacobi, Spinozism and pantheism were synonymous terms, and pantheism was identical with dogmatic rational atheism. (He ignored the possibility of interpreting Spinoza's system as acosmism instead of as atheism—an interpretation that was first suggested by Salomon Maimon and then by Hegel.)

Jacobi's ethicoreligious worldview is the background of this criticism of Spinoza. While recognizing the dangers implied in Spinozism, Jacobi and Gotthold Ephraim Lessing were the first to acknowledge the philosophical genius of Spinoza. Through Jacobi's discussions with Mendelssohn about Spinozism and Lessing's relation to it, in the course of which the arguments for and against Spinoza were brought forth, Spinoza's philosophy became a force in the intellectual life of the time; it acquired a universal significance. Spinoza and Kant were two opposing poles of thought for Jacobi. For the former all being, including man, is determined by necessary laws; for the latter freedom and creativity are the essence of man. The whole period of the development of post-Kantian speculative idealism was determined by the two intellectual forces: the dogmatic rationalism of Spinoza and the critical philosophy of Kant. Jacobi was critical of the philosophy of speculative idealism (Fichte, Schelling, and Hegel) for its manifestation of Spinozistic tendencies.

JACOBI ON KANT

Jacobi's enthusiasm for Kant's precritical essay *Der einzig möglicher Beweisgrund zu einer Demonstration des Daseins Gottes* (The only possible ground for a demonstration of God's existence) is indicative of his conception of the method by which we can attain knowledge of reality. Kant had shown in this work that the absolute and unconditioned being must be grasped as existing in and through itself, not as a predicate or as a consequence of something else. The attainment of some reality that is simple, insoluble, and immediately given is the ultimate aim in our striving for certainty. Cognition by way of discursive thought cannot attain certainty. A method of deduction of consequences from premises is an endless process that can never attain the original unconditional and primary being. Certainty is acquired only in an immediate perception of a reality not requiring any deduction.

Jacobi admired Spinoza because he had reversed the whole process of philosophizing as it was known since Aristotle. Instead of proceeding from the phenomena of experience, leading gradually to being as such, Spinoza started with a definition of substance as something that is conceived in itself and through itself—that is, a simple and immediately given reality. This simple and indissoluble datum is, however, according to Jacobi, not free from contradiction. Spinoza's substance is not a free, independent, self-sufficient being, but a necessary and causally bound being. The God of Spinoza is nothing else but a manifestation of the logical-mathematical determination of being.

The critical philosophy can be maintained only if it consistently removes all traces of a dogmatic, realistic nature. The concept of a thing-in-itself has to be completely eliminated because it is incompatible with the system as a whole. The Kantian position is, according to Jacobi, pure idealism. As such it cannot retain the concept of things in themselves. The *Critique of Pure Reason* deduces the objects from the constitution of our cognitive capacity. It has therefore to deny objective reality existing independently of and beyond the conditions of cognition. The object has to be completely resolved in subjective presentations of our mind. Kantian philosophy is thus interpreted by Jacobi as pure subjective idealism. Since we perceive the objects through forms of sensibility (space and time) and concepts of understanding, constituting the human capacity of cognition, the "external" objects cannot be beyond us. According to Jacobi, René Descartes intended by the principle *cogito ergo sum* to deduce the totality of the inner subjective world from the consciousness of the self as a thinking subject. Self-consciousness of oneself as a thinking being is the primary condition of man's knowledge of the inner world. In a similar manner Kant tried to prove that external objects are likewise conditioned by and dependent on the subject with its forms of sensibility and understanding. Hence, the subjective idealism of Descartes was extended by Kant to encompass the world of objects, too. The Kantian position is thus, according to Jacobi, universal idealism, but since he took Kant to mean that the cognition of things is determined by the individual ego and not by the objective mind as it is presented in scientific thought, universal ide-

alism, according to Jacobi, is a system of absolute subjectivity, which implies a "nihilism" with reference to the objects. This system recognizes only the ego as real; it is thus speculative egoism. Jacobi found this position self-contradictory. Sensibility is a receptive function, according to Kant. But a consistent idealism excludes a receptive capacity in the process of cognition. It is incongruous with idealism to assume the reality of things-in-themselves existing independently of our mind, yet these things are supposed by Kant to supply the material of experience that affects our senses. The first part of the Kantian *Critique* deals with the forms of sensibility as a receptive capacity. Thus, things-in-themselves are assumed, by which our sensibility is affected. "Hence we cannot enter into the *Critique* without assuming things-in-themselves, but we cannot retain this assumption upon leaving the *Critique*."

Since Jacobi understood the Kantian position as subjective idealism, he did not consider the second edition of the *Critique of Pure Reason* an improvement on the first. The Kantian philosophy that Jacobi took to be a form of pure subjective idealism is presented in the first edition of the *Critique*, and this he took to be its genuine and adequate presentation. To Jacobi belongs the priority of recognizing the difference between the two editions, but he was wrong in its evaluation. He failed to grasp the essential characteristic of critical idealism, which is grounded in analysis of objective scientific cognition and not in analysis of the process of cognition of the individual subject. The problem posed by Kant was How are synthetic propositions a priori in mathematics and natural science possible?, not How is cognition of the individual subject as a psychological phenomenon possible? Whereas the Kantian inquiry constitutes the essence of the transcendental method, leading to objective idealism, the investigation of the individual process of cognition appertains to the psychological method, resulting in subjective idealism. The second edition of the *Critique*, which tries to eliminate the psychological sections of the first edition, is the preeminent presentation of the transcendental method.

FAITH—THE SENSE OF REALITY

In opposition to the critical philosophy, which is, according to Jacobi, absolute subjectivity, he proposed a thesis of absolute objectivity. The objective reality of things-in-themselves existing beyond man and independently of the human cognition is based for Jacobi on an original, immediate certainty that does not require any proof or demonstration. The certainty of the existence of things in themselves is based on faith.

Our consciousness presupposes the reality of things as a necessary correlate of cognition. The idealistic position contradicts an assumption that is inherent in every act of cognition of an object of experience. To be sure, the reality of the things cognized cannot be conclusively derived from the process of cognition as such, which is a subjective phenomenon, but only from the immediate sense of reality accompanying every act of cognition of an object. This sense of reality, which cannot be accounted for logically but is nonetheless present in our mind, is designated by Jacobi by such terms as *faith*, *feeling*, and, later, *revelation*.

With Kant Jacobi recognized that analysis of cognition cannot lead to things-in-themselves, since the validity of the categories is confined to the realm of experience and does not extend beyond it. But the Kantian *Critique* had also shown that reason leads to a realm of faith in addition to mere cognition. Hence, an object that cannot be proven as real on the basis of cognition may still be real on the basis of faith. Kant employed the concept of faith only with reference to the moral and religious realm, but Jacobi extended the scope of faith to include the knowledge of things-in-themselves. In recognizing the validity of faith for the theoretical realm, Jacobi followed David Hume, who designated the feeling of reality of the natural human consciousness as faith. The skepticism of Hume showed that the reality of things cannot be derived from sense perception. Analysis of perception cannot lead to cognition of substance and causality; only through faith can we know the reality of things. Hume thus ascribed to faith a positive theoretical function inasmuch as it is a source of knowledge of the reality of the things of experience. The belief in the reality of things, which accompanies our sensuous experience throughout our lives, is incomprehensible, but, according to Jacobi, it commands certainty just as if it were an act of revelation. He understood by revelation a certainty that we are aware of but which we cannot explain rationally. This conception of the belief in the reality of things is radically different from naive realism and commonsense philosophy. The latter does not realize the extraordinary nature and the problematic character of the concept of reality, of things-in-themselves. Naive realism takes for granted that we perceive things as they are. Jacobi, however, realized the miraculous nature of such a belief. The possibility of transition from consciousness to things, from the subject to objects, cannot be comprehended by our understanding. Jacobi was right to affirm the position of critical ide-

alism that we can know of things only what we ourselves put into them. While we cannot cognize things-in-themselves, our belief in their reality can be accounted for as something irrational that is an indispensable ingredient of human consciousness. Our rational thinking cannot lead us to cognition of reality of things-in-themselves. However, the necessary condition of the existence of man as a conscious being is grounded in an incomprehensible and irrational act of faith commanding certainty that is not subject to any doubt. In face of this belief as a necessary condition of human consciousness, the arguments of rationalism, of critical philosophy, and of skepticism are powerless.

RELIGION

Jacobi's philosophy of religion is grounded in the same principle on which his theory of cognition of reality is founded. The concept of faith as having a theoretical function is the ground of the certainty of real objects beyond us and of a supersensuous reality. This immediate certainty of reality is present in our consciousness of God as it is present in our perception of objects. Through belief man has the capacity of intuiting God. Dogmatic religionists maintain that through an act of faith God reveals himself to man by grace. For Jacobi faith is a mode of cognition or a form of intellectual intuition. And this is not an exclusively religious phenomenon, for through belief man likewise perceives the reality of things of experience. The distinction between the reality of the things and the transcendent, supersensuous reality is that the former reveals itself through an external perception, whereas the latter is intuited through an internal revelation. Both forms of revelation constitute the very essence of human existence as a conscious being.

For Kant, it is impossible through faith to transcend the sphere of the subject, but for Jacobi we are aware through faith of a reality that is not subjective, since in the act of faith the nature of the real thing reveals itself to us. Faith thus commands not only ethical certainty, as Kant held, but also theoretical certainty. To be sure, the transcendent reality cannot be known by the forms of understanding that are confined to the realm of experience. But faith as a function of reason (*Vernunft*) is capable of transcending experience and thus can perceive the supersensuous by an act of intellectual intuition.

Intellectual intuition, which is attained through faith, overcomes the Kantian dualism of sensibility and understanding, which is a necessary condition of cognition of objects of experience. Kant considered intellectual intuition an idea of knowledge of the infinite mind (*intellectus archetypus*), which is not attainable by the finite, human mind. But for Jacobi intellectual intuition is attained through faith, or immediate feeling. Jacobi thus prepared the way for the post-Kantian speculative metaphysicians to consider intellectual intuition a capacity of human reason.

CRITICISM

By ascribing to belief the function of knowledge of things-in-themselves and of the existence of God and of freedom, Jacobi disregarded the essential difference between the theoretical and the ethical realms. Kant's concept of faith is a new principle of validity but not a mode of knowledge. In the *Critique of Practical Reason* Kant discovered an "unconditioned" in opposition to the conditioned reality of experience. God and freedom as ideas of practical reason are not metaphysical things but principles of ethical conduct. It is the unconditioned of freedom and the "ought to be," not the existence of transcendent reality, that is discovered through faith.

Jacobi is rightly critical of the dogmatic rationalism of the Enlightenment; he realized the limitations of rational thought in face of the endlessness of that which is problematic. But he was wrong in subordinating the realm of science, which is grounded in discursive thinking, to that of feeling and faith. He did not realize the problem involved in his concept of belief and immediate feeling as the highest means of attaining knowledge of reality. The appeal to feeling, belief, and immediate evidence opens up possibilities for abuse and willful arbitrariness. Feeling and immediate sense of reality are subjective, and whenever a capacity of the subject is elevated to a principle of knowledge, objective truth is in jeopardy. The rightful place of faith is therefore the ethical and the religious realm, which is concerned with the "ought," not with being as it is. Theoretical knowledge of reality can be attained only by discursive thinking, which is the scientific method.

Jacobi said of himself that he was a pagan in his mind but a Christian in his heart. He thus recognized the conflict between reason and faith that he caused by the extended role he ascribed to faith. His belief in the reality of things-in-themselves, of a supersensuous being, and of freedom not only claims ethical and religious validity but also pretends to possess the rank of theoretical knowledge; it is therefore in conscious disagreement with reason. The price we pay for extending the scope of faith is its clash with reason.

See also Pantheismusstreit.

Bibliography

WORKS BY JACOBI

Über die Lehre des Spinozas, in Briefen an Herrn Moses Mendelssohn. Breslau: G. Löwe, 1785.

David Hume über den Glauben, oder Idealismus und Realismus. Breslau: G. Löwe, 1787.

Sendschreiben an Fichte. Hamburg, 1799.

Über das Unternehmen des Kritizismus, die Vernunft zu Verstande zu bringen und der Philosophic überhaupt eine neue Absicht zu geben. Hamburg, 1802.

Von den göttlichen Dingen und ihrer Offenbarung. Leipzig, 1811.

Jacobi's Werke, 6 vols. Edited by F. Roth. Leipzig: G. Fleischer, 1812–1825.

Hauptschriften zum Pantheismusstreit zwischen Mendelssohn und Jacobi. Edited by M. Scholz. Berlin: Reuther and Reichard, 1916.

WORKS ON JACOBI

Ahlers, Rolf. "Vitalism and System: Jacobi and Fichte on Philosophy and Life." *Idealistic Studies* 33 (1) (2003): 83–113.

Behler, Ernst, ed. *Philosophy of German Idealism: Fichte, Jacobi, and Schelling*. New York: Continuum, 1987.

Bowie, Andrew. "Rethinking the History of the Subject: Jacobi, Schelling, and Heidegger." In *Deconstructive Subjectivities*, edited by Simon Critchley. Albany: State University of New York Press, 1996.

Di Giovanni, George. "The Early Fichte as Disciple of Jacobi." *Fichte Studien* 9 (1997): 257–273.

Di Giovanni, George. "Fichte's Rhetoric of Deception: Reflections on the Early Fichte in the Spirit of Jacobi." *Revue Internationale de Philosophie* 49 (191) (1995): 9–78.

Di Giovanni, George. "From Jacobi's Philosophical Novel to Fichte's Idealism: Some Comments on the 1798–99 'Atheism Dispute.'" *Journal of the History of Philosophy* 27 (1989): 75–100.

Di Giovanni, George. "The Jacobi-Fichte-Reinhold Dialogue and Analytical Philosophy." *Fichte Studien* 14 (1998): 63–86.

Franks, Paul. "All or Nothing: Systematicity and Nihilism in Jacobi, Reinhold, and Maimon." In *The Cambridge Companion to German Idealism*, edited by Karl Ameriks. New York: Cambridge University Press, 2000.

Janssens, David. "The Problem of the Enlightenment: Strauss, Jacobi, and the Pantheism Controversy." *Review of Metaphysics* 56 (3) (2003): 605–631.

Lévy-Bruhl, L. *La philosophie de F. H. Jacobi*. Paris: Alcan, 1894.

Milbank, John. "Knowledge: The Theological Critique of Philosophy in Hamann and Jacobi." In *Radical Orthodoxy: A New Theology*, edited by John Milbank, Catherine Pickstock, and Graham Ward. New York: Routledge, 1999.

Schmidt, F. U. *F. H. Jacobi*. Heidelberg, 1908.

Schmidt, James. "Liberalism and Enlightenment in Eighteenth-Century Germany." *Critical Review* 13 (1–2) (1999): 31–53.

Snow, Dale E. "F. H. Jacobi and the Development of German Idealism." *Journal of the History of Philosophy* 25 (1987): 397–415.

Solomon, Robert C. *Introducing the German Idealists: Mock Interviews with Kant, Hegel, Fichte, Schelling, Reinhold, Jacobi, Schlegel, and a Letter from Schopenhauer*. Indianapolis: Hackett, 1981.

Westphal, Kenneth R. "Hegel's Attitude toward Jacobi in the 'Third Attitude of Thought toward Objectivity.'" *Southern Journal of Philosophy* 27 (1989): 135–156.

Williams, Robert R. "The Other: F. H. Jacobi and German Idealism." In *Hegel on the Modern World*, edited by Ardis Collins. Albany: State University of New York Press, 1995.

Wurzer, Wilhelm S. "Between Idealism and *Dasein*: Friedrich Heinrich Jacobi." *Dialogos* 11 (1977): 123–136.

Zirngiebl, E. *F. H. Jacobi*. Vienna, 1867.

Samuel Atlas (1967)
Bibliography updated by Tamra Frei (2005)

JAMES, HENRY
(1811–1882)

Henry James, an American philosophical theologian in the Swedenborgian tradition, is perhaps best known as the father of the novelist Henry James and the philosopher-psychologist William James. Although the elder James was physically handicapped from his early teens, an inheritance from his father, a dominant figure in upper New York State real estate provided him with a lifelong income. Henry James graduated from Union College in 1830 and studied for the Presbyterian ministry at Princeton Theological Seminary from 1835 to 1837. Dissatisfied with the ritual formality and absence of spirituality in what he called "professional religion," he left the seminary and traveled to England, where he came under the influence of the idiosyncratic theology of Robert Sandeman, author of *Letters on Theron and Aspasio,* which James edited for American publication in 1838. Soon afterward, through J. J. Garth Wilkinson, James discovered Emanuel Swedenborg. During the remainder of his life, he developed his own insights in the language of, and within the broad framework of, Swedenborgian ideas.

Central to James's view was the belief that selfhood (Swedenborg's *proprium*) is the sin of sins. Since the movement of creation is a move away from God, it is during this phase that selfhood flourishes. Religion and morality form, as it were, a reflecting surface from which the individual is "bounced back" toward God, thus initiating the movement of redemption, in which selfhood is replaced by "sociality" as a dominant motivation. Thus, as one of James's titles indicates, society is the redeemed form of man. Selfhood is destructive of the Divine intention with regard to created nature, whereas sociality is reconstructive. The ideal of redemptive society that James envisioned was largely derived from the social theories of

Charles Fourier and emphasized social solidarity and democracy.

Because of this double allegiance to Swedenborg and Fourier (an allegiance James shared with many of his contemporaries, including Parke Godwin, Horace Greeley, and Albert Brisbane), James was able to assert that the highest points of European life were reached in Protestantism and constitutional liberty, and that both of these had been raised to still higher levels in America. Beneath the sometimes crude externals of American democracy, he saw "the soul of fellowship that animates and redeems it." Thus, he conceived of democracy as the herald of moral perfection and the means of "preparing the way for the reign of infinite Love." In this way James linked his theology of redemptive society to American democratic practice and to its ideal theory.

See also Democracy; Fourier, François Marie Charles; James, William; Liberty; Swedenborg, Emanuel.

Bibliography

WORKS BY JAMES

Selected Works of Henry James, Sr. (1846). New York: AMS Press, 1983.

Moralism and Christianity. New York: Redfield, 1850.

Lectures and Miscellanies. New York: Redfield, 1852.

The Nature of Evil. New York: Appleton, 1855.

Christianity, the Logic of Creation. London and New York: Appleton, 1857.

Substance and Shadow. Boston: Ticknor and Fields, 1863.

The Secret of Swedenborg. Boston: Fields, Osgood, 1869.

Society the Redeemed Form of Man. Boston: Houghton, Osgood, 1879.

The Literary Remains of the Late Henry James. Edited by William James. Boston: Houghton Mifflin, 1884. Contains a lengthy introductory essay.

WORKS ON JAMES

Duban, James. *The Nature of True Virtue: Theology, Psychology, and Politics in the Writings of Henry James, Sr., Henry James, Jr., and William James.* Madison, NJ: Fairleigh Dickinson University Press; London: Associated University Presses, 2001.

Habegger, Alfred. *The Father: A Life of Henry James, Sr.* New York: Farrar, Straus, and Giroux, 1994.

Hoover, Dwight W. *Henry James, Sr., and the Religion of the Community.* Grand Rapids, MI: Eerdmans, 1969.

Perry, Ralph Barton. *The Thought and Character of William James.* Boston: Little Brown, 1935. Contains discussions of the elder James.

Warren, Austin. *The Elder Henry James.* New York: Macmillan, 1934. Includes much biographical information.

Young, Frederic H. *The Philosophy of Henry James Sr.* New York: Bookman Associates, 1951.

J. L. Blau (1967)

JAMES, WILLIAM
(1842–1910)

William James, the American philosopher and psychologist, was born in New York City to Mary Robertson Walsh James and Henry James Sr., the eccentric Swedenborgian theologian. James's paternal grandfather and namesake was an Irishman of Calvinist persuasion who immigrated to the United States in 1798 and became very rich through felicitous investment in the Erie Canal. James had three brothers and a sister; one of them, the novelist Henry James, achieved equal fame.

James's early environment was propitious; his father's enthusiastic and unconventional scholarship, his personal and unorthodox religion, his literary association with men like Oliver Wendell Holmes Sr. and Ralph Waldo Emerson all stimulated free intellectual growth. Even more important was the rather extraordinary respect that the elder James lavished upon the youthful spontaneities of his children; each, he thought, must go his own way and become that most valuable of creatures, himself. There was no straitlaced dogmatism in the James household, and William James was free to accept or reject the ideas of his father and his father's friends. The thought and sympathies of these transcendentalists and romantic humanitarians of the New England tradition never seemed to James the ultimate answers to his own philosophical and personal problems, but they dealt with genuine issues that he did not evade in his later work.

James's primary education took place at his father's table; its main constituents were the spirited discourse that the family held on every topic and the example of the parents, loving and unworldly. Formal education took place irregularly in various private establishments. From 1855 to 1860 James (often in the company of his younger brother Henry) attended schools in England, France, Switzerland, and Germany. There, as his father said, he and his brother were able "to absorb French and German and get a better sensuous education than they are likely to get here" (Ralph Barton Perry, *The Thought and Character of William James,* p. 59). During this European sojourn James's interest was divided between natural science and art, especially painting.

In spite of his continuing enthusiasm and talent for scientific inquiry, James's interest in painting became so strong by 1860 that he resolved to spend a trial period learning to paint. The elder James was not anxious for his son to become a painter, thereby prematurely cutting himself off from the rest of life's possibilities; any definite vocation, according to the father, was sadly "narrowing"

(ibid., p. 171). It was nevertheless arranged that James should begin study with William M. Hunt in Newport. This experiment convinced James that he lacked the ability to be anything more than a mediocre artist, than which there was, he thought, nothing worse. The lesson at Newport permanently discouraged James's pursuit of an artistic vocation, but throughout his scientific and philosophical career he retained the artist's eye, his predilection for concrete sensuous detail, and his concern for style.

In 1861, James entered the Lawrence Scientific School, Harvard, studying first in the chemistry department under Charles W. Eliot, later in the department of comparative anatomy and physiology under Jeffries Wyman and Louis Agassiz. From Wyman he learned the importance of evolution; from Agassiz, an appreciation of "the world's concrete fulness" (William James, *Memories and Studies,* p. 14) and of acquaintance with empirical facts as against abstraction. In 1864, James transferred to the medical school, though without the intention of ever practicing medicine. His medical studies, although fruitful, were attenuated and sporadic.

While at medical school James joined Agassiz as an assistant on the Thayer expedition to Brazil during 1865/1866. In Brazil he contracted smallpox and suffered from sensitivity of the eyes. This was the first serious manifestation of that constitutional failure which was to recur throughout James's life, imposing upon it a pattern of interrupted work and of periodic flights to Europe which were always, at least in part, searches for health.

In 1867 ill health and the desire to study experimental physiology led James to Europe, to Germany in particular. While little formal study of physiology proved to be possible, James read widely and thoughtfully. His first professional literary effort, a revision of Herman Grimm's *Unüberwindliche Mächte,* published in the *Nation* (Vol. 5, 1867), dates from this period.

James returned to Cambridge in November 1868 and received his medical degree in June 1869. After a period of illness and retirement, he began teaching anatomy and physiology at Harvard in 1873, psychology in 1875, and philosophy in 1879. This order is very nearly accidental and gives no adequate indication of James's development. Philosophy was an early interest which grew with his scientific studies; for James the more narrowly scientific questions could never be separated, even theoretically, from the more general questions which philosophy considers.

It was indeed a specifically philosophical concern which precipitated James's profound emotional crisis of 1870. He had been suffering from a sense of moral impotence that only a philosophical justification of the belief in the freedom of the will could cure. In the *Essais de critique générale* of Charles Renouvier, James found the basis of the justification he sought. And throughout his life the problem of maintaining free will and the moral attitude in the face of either religious monism or scientific determinism, as well as the problem of legitimating belief despite various intellectual skepticisms, continued to engage James's attention and to influence his mature philosophy. That philosophy, growing out of personal need and agitation, has a strong eschatological flavor. It cannot, however, be reduced either to a scheme of personal salvation or to an apology for some special way of life. James offered a philosophical, not an emotional, defense of free will, moralism, and belief. These topics became important test cases for a general metaphysics that James sought to elaborate not for its own sake but to satisfy interests which were distinctly rational and theoretical.

Having settled into the career of philosopher and teacher, if one may speak of James's settling into anything, he maintained close but not constant association with Harvard until his final resignation in 1907. He married Alice Howe Gibbens in 1878; the marriage seems to have increased his sense of purpose and coincided with a noticeable improvement in James's health. Thenceforth, he led an intensely active life, teaching at Harvard, lecturing widely, and publishing a series of books which became undeniable classics of American philosophy. Three series of James's lectures deserve special mention. He gave the Gifford Lectures at Edinburgh in 1901/1902, published as *The Varieties of Religious Experience* (1902); lectures on pragmatism at the Lowell Institute and Columbia in 1906 and 1907, published as *Pragmatism* (1907); and the Hibbert Lectures at Oxford in 1908/1909, published as *A Pluralistic Universe* (1909).

CHARACTER OF JAMES'S PHILOSOPHY

This brief biography gives no indication of that range and richness of James's experience which so struck those who knew him and which entered into everything he wrote. James was a highly social man whose friends formed an intellectual community of great distinction. Chauncey Wright, C. S. Peirce, Shadworth Hodgson, Charles Renouvier, Josiah Royce, George Santayana, John Dewey, Henri Bergson, and F. H. Bradley were a few of those whom James knew as friends and fellow laborers.

From all of these men and others James drew philosophical nourishment, and the very number of sources and influences renders the search for antecedents otiose. James was essentially an original thinker, and he borrowed only what fitted his own design. This must be maintained in spite of James's habitual humility and his characteristic generosity of acknowledgment.

James impressed his friends with his vitality and strength of character, with his open-mindedness and sympathy. His spirit and attitude were admired even by those whose philosophical conclusions differed radically from his own. Santayana, for example, in his witty and condescending memoir *Character and Opinion in the United States* is forced to praise James, at least as an enthusiastic and explosive force. Because James wrote as he talked, much of his vividness and personal style is retained in his works. The majority of James's books are simply transcriptions of lectures; they have all the virtues and vices of spoken discourse, and the circumstances of their presentation must help to determine the kind of analysis to which they can be fruitfully subjected.

James addressed himself to the people, not especially to other philosophers, and he listened to the people to find out what life meant to them. He respected not so much their common sense as their common feelings and hopes and would not allow his philosophy to dismiss cavalierly that which figured largely in the experiences of men. The people listened to James, and his books sold well. By the end of his life he was nearly a legendary figure, and he was generally regarded as the chief representative of American philosophy. Nevertheless, professional philosophers, when they have discussed James at all, have tended to concentrate on those of his ideas that, separated from the body of his work and often distorted, have achieved currency. To this general picture there are important exceptions, such as Ralph Barton Perry, who has done more for James scholarship than anyone else.

To provide a proper perspective for the study of James, three corrective measures must be taken. First, attention must be diverted from his life, however interesting, to his published philosophy. For all its validity the biographical motive can be, and has been, pressed to the point where it precludes philosophical clarity. Second, James must be seen within the general philosophical tradition, in relation to the fundamental philosophical problems that he attempted to solve and not in relation to his position as a distinctly American thinker. To attempt to evaluate James's philosophy in terms of his American background is neither more nor less rewarding than to attempt to evaluate Immanuel Kant, say, in terms of his German background. Third, the objective aspect of James's philosophy must be stressed. James himself thought that philosophy involved the subjective factors of temperament and personal vision. In the first chapter of *Pragmatism,* he drew a very plausible correlation between tough-minded and tender-minded temperaments and empirical and rationalist philosophical positions. Again, in the essay "The Sentiment of Rationality" James argued that there can be no adequate definition of reason which ignores the feeling of rationality, the ultimate sense of logical fit. James believed that the subjective (or what might better be called the aesthetic) dimension was a feature of philosophy as such. James's philosophy is subjective, therefore, because it is philosophy, not because it is James's philosophy. Objectivity, like truth and reality, was redefined, not abandoned, by James.

The remainder of this entry is divided into sections on James's psychology, philosophy of religion, pragmatism, and metaphysics. This arrangement is simply an expository device. If pragmatism is a theory of all belief, then religious philosophy is a subdivision of pragmatism. If pragmatism is a description of what actually happens when men seek truth, then it is part of psychology. If the dualism between human and natural processes is finally inadmissible, then psychology is a chapter of general metaphysics. The interdependence of the various parts of James's philosophy, suggested here, will be exhibited below.

PSYCHOLOGY

The *Principles of Psychology* (1890) is, according to James himself, "mainly a mass of descriptive details"; certainly, this work more than any other justifies Alfred North Whitehead's remark that James's primary task was philosophical assemblage. The *Principles* "assembles" in two senses. First, there is a brilliant gathering, through extensive quotation and reference as well as careful documentation, of relevant material from the Scottish, English, French, and German schools. Second, there is the exhibition of facts which may never have appeared prominently in any system, either of psychology or of philosophy.

It has become customary, and it is certainly legitimate, to praise the *Principles* for its sensitive evocation of the evanescent inner life. It is indeed a kind of generalized psychic autobiography by a master of introspection, but it is much more than a document of literary psychology. The concrete rendering of experience is an essential element in the development of James's mature philosophy, for when he spoke of the world as "a world of pure experience," he referred to experience as it is described in the

Principles. If experience had not the ramifications and possibilities so lovingly and exuberantly detailed by James in his "psychological" writings, it could never have become, as it did for James, the central image of complete reality. Moreover, James was not in his early days merely "collecting" facts whose subsequent careers happened to include the incident of being generalized into a total world view. James, as he said himself, "hated collecting" (Perry, op. cit., p. 225). The material of the *Principles* is already thrown into philosophical form, is already illuminated or stained (however one decides the matter) by the foundational metaphysical categories that recur, with greater generality, in the later works.

DESCRIPTION. If the *Principles* is to be regarded as primarily a descriptive work, one must be clear about what is involved in description as James understood it. He was convinced that pure description in the manner of phenomenology is impossible. Description cannot be other than conceptual; concepts, in turn, are tools of classification that have inexpugnable conventional and theoretical elements. Concepts do not passively mirror; they select according to human interests and purposes. Assumptions, James maintained, have a way of establishing themselves "in our very descriptions of the phenomenal facts" (*Principles,* Vol. I, p. 145). Naive phenomenology attempts to eliminate assumptions from descriptive statements. This is an impossible task if for no other reason than that every allegedly assumption-free phenomenology must itself make doubtful assumptions, including the assumption that there can be description without classification. James's own approach was to examine the assumptions involved in all descriptions, making those assumptions "give an articulate account of themselves before letting them pass" (loc. cit.). Pragmatism as it appears in the *Principles* consists simply in spelling out what claims our theories and assumptions make for us and in eliminating elements which are superfluous, elements, that is, which can be eliminated without changing the tenor of what we really want to say. Pragmatism here can be fruitfully regarded as a general theory of theory criticism, as an attempt to make clear what we are actually committed to by the theories we entertain. The chapters that criticize the conscious automaton theory (Vol. I, Ch. 5) and the mind-stuff theory (ibid., Ch. 6), respectively, are indeed the first extended exercises in pragmatic criticism.

SCIENCE AND METAPHYSICS. Purely phenomenological description being considered impossible by James, the question arises as to what is scientific about the *Principles.* The standard interpretation—the interpretation upon which the judgment of its great historical importance is based—finds the work very nearly the first attempt to treat psychology from the standpoint of a natural science—that is, descriptively and apart from metaphysical theories. The sharp distinction which we are likely to draw between scientific theories and metaphysical theories is difficult to sustain from James's own point of view and therefore cannot be used to differentiate the *Principles* from metaphysical treatments of the same subject matter.

A much more pregnant distinction is that between a priori and a posteriori metaphysics. A priori metaphysics was, for James, a totally illegitimate enterprise consisting of vacuous abstractions excogitated apart from any experience of the world. Throughout his work James often referred to a priori metaphysics simply as "metaphysics," and his frequent criticisms of metaphysics must therefore be carefully interpreted in their contexts. The *Principles* is antimetaphysical where metaphysics means "scholastic rational psychology" or "philosophical pyschology."

The more interesting problem is defining the relation between the *Principles* and the kind of metaphysics of which James did approve, a posteriori metaphysics, which is continuous with science and, like science, is both descriptive and theoretical. Here the differentiation must be emphatic rather than absolute. The *Principles* may be regarded as a deliberate (and artificial) restriction of general metaphysical scope. Science, as James saw it, must grow into metaphysics. Explanation must become more complete and more comprehensive even if, as James certainly believed, it cannot become total and absolute. But science must *be* science before it can become metaphysics. In the *Principles* metaphysics is, necessarily, postponed; its positivism is provisional rather than dogmatic and final.

The relative autonomy that science is given in the *Principles* is "for the sake of practical effectiveness exclusively," as James said in his essay "A Plea for Psychology as a 'Natural Science'" (*Collected Essays and Reviews,* p. 317). Science left to itself, with its "convenient assumptions" unquestioned, is best able to accumulate a mass of factual details which lead to the subsequent enrichment and "thickening" of the content of metaphysics. The danger of premature metaphysical reconstruction is thinness, impoverishment of content, and abstraction.

MENTAL STATES. The basic assumption of the *Principles* and its "convenient" point of departure is the existence of mental states. The first task of psychology is to describe

the conditions of these mental states with as much detail and completeness as possible. Chapter 2 of the *Principles* is an extended examination of the ways in which various brain states condition various mental states. The search for conditions among bodily experiences generally and brain experiences particularly is the only alternative to treating mental states as frankly miraculous. James, the evolutionary naturalist, had to maintain that mental states grow out of physical states, in spite of whatever difficulties this view entails. Since mental states, in addition to arising from physical antecedents, themselves give rise in all cases to changes in the physical world, it seems utterly impossible to create any kind of dualistic ontological chasm between the two types of process, mental and physical. There is indeed a discriminable subject matter of psychology, which James referred to both as "mental states" and as "mental life." This subject matter must be treated autonomously, which means, in practice, guarding against the reduction of mental phenomena to nothing but physical phenomena in the interest of some schematic monism. In this context James at times spoke of "irreducible dualisms," but what he meant to emphasize might perhaps better be called "irreducible dualities," discriminations that remain what they are no matter what supervenient integrations may also be pointed out.

The whole question of the dualism between the physical and the mental is complicated by the fact that James was, even in the *Principles,* developing a view of physical nature at large which departed radically from the familiar deterministic, mechanical model. It is often maintained, for example, that James's treatment of the will as irreducible to antecedent mechanical factors creates a dualistic chasm between natural processes and characteristically human processes. This would be true only if James had retained the customary deterministic model of nature. James, however, did not retain this model; he would sooner have conceived of all nature as willful than of man's will as an exception to nature.

James believed that the borderline of the mental is vague. Mentality, as James defined it, exists wherever we find the choice of means for the attainment of future ends. Mental life is purposive in a way that involves the overcoming, through suitable invention and appropriation, of any obstacles lying in the way of its purpose. The mind is a tactical power that reveals itself in the struggle with its environment. The only kind of world in which minds can conceivably develop and be found is one in which success is neither automatic nor impossible. An interesting consequence of James's view is that an omnipotent God could not have a mind; neither could a purely contemplative deity. The notion of mind as an instrument within the general economy of purpose and resistance to purpose, a notion which has justly been called "biological" and "Darwinian," is simply an ungeneralized expression of pragmatism.

Although it is necessary to consider mental states as "temporal events arising in the ordinary course of nature" (ibid., p. 319), with emphasis on their natural antecedents and results, it is also necessary to consider mental states in themselves as realities to be described as they are found with their generic particularity and variety intact. Here again, it must be emphasized that James was not attempting a phenomenology of mental life or consciousness. What he was attempting was the provision of adequate description that would not be guilty of gross oversimplification or distortion.

INTROSPECTION. Adequate description must, of course, be based somehow upon observation, and, James maintained, the principal method of psychology is introspective observation. Introspection, as an observational process, is similar to other kinds of observation. James could find in introspection no peculiar epistemological characteristic; it is neither more nor less fallible than other kinds of observation. Its frequently alleged infallibility, based on some notion of the immediate relation obtaining between a mind and its contents, is simply contradicted by experience. Even if feeling is unmistakably what it is, our "naming, classing, and knowing" (*Principles,* pp. 189–190) of every feeling share in the notorious general human fallibility. The truth of any observation, introspective or otherwise, is not to be found in the character of the source of observation but in the consequent service, especially theoretical service, which the observation and its correlative preservation in description can be made to render. There is therefore no simple and immediate verification of observations, no once and for all validation of descriptions. For James "the only safeguard [of truth] is the final *consensus* of our further knowledge about the thing in question, later views correcting earlier ones until at last the harmony of a consistent system is reached" (ibid., p. 192). James's own descriptions in the *Principles* must lend themselves to this kind of pragmatic corroboration.

THOUGHT. The famous "descriptive" chapter, "The Stream of Thought" (perhaps the heart of the *Principles*), cannot be evaluated from a simply empirical point of view. What is described and how it is described are determined by markedly theoretical affinities and avoidances. James singled out five traits of thought in that chapter: (1) Thought tends to be part of personal consciousness—

that is, thought is not experienced as simply *a* thought but as *my* thought; (2) thought is always changing; (3) within each personal consciousness thought is sensibly continuous; (4) thought deals with objects independent of itself; and (5) thought is selective and has interests. The metaphysical model that James had in mind here is of a process that is partially determined and partially self-determining—that is, centered or focused and essentially temporal. Although the analysis in the *Principles* is limited to one kind of process, consciousness, the structure of the analysis is similar to that Whitehead offers of all actual occasions. James himself came to believe that all of reality must be describable in terms like those used for human experience. This belief is elaborated in *Essays in Radical Empiricism* as the notion of a world of "pure experience."

Each of the five traits of thought which James distinguishes repudiates some important philosophical position. One dimension of James's work clearly apparent in the *Principles* is a sustained criticism of the "classic-academic" version of mind. No easy summary of what this meant to James is available, but its main features would seem to be the marshaling of instances of mental phenomena according to a priori canons of clarity and rationality, the overwhelming influence of the assertive paradigm (as opposed to the judgments implicit in making and doing) in construing the problems of belief and judgment, and allegiance to the spectator theory of knowledge with whatever passivity is therein involved. These attitudes James attacked in the name of a richer experience, encompassing all the concrete information we possess about the functions of mind. This is the information, so carefully assembled and considered in the *Principles,* which James urged the epistemologist to work into his official model and the philosopher generally to consider in making his pronouncements.

EXPERIENCE. The appeal to experience is not new in philosophy; James was solidly in the venerable tradition of empiricism. But empiricism in its classic British form is essentially an epistemological position that regards experience as an exclusive witness before a cognitive tribunal in which other sources of evidence are ruled out of court as uncertain or unreliable. The genius of James's empiricism lies precisely in ruling nothing out of court. His theory of experience, the object of so much of James's later labor, is perhaps the first such theory which is cosmological, rather than strictly epistemological, in intention and logical form. This shift of the total frame of reference within which experience is considered has, for better or worse, influenced a subsequent movement in philosophy typified by Whitehead and his disciples. It is this influence which points to the main philosophical significance of the *Principles.*

PHILOSOPHY OF RELIGION

Even in the introduction to *The Literary Remains of the Late Henry James* (1885), a relatively early work that might be thought no more than an act of filial devotion, James's own ideas about religion were quite clear. There is, of course, a sympathetic exposition of his father's superpersonal theological monism, for William James could honestly admire his father's "instinct and attitude" even if he could not condone the "cold accounts" and abstract formulations of the elder James's system. It is religious experience, rather than religious doctrine, that matters. Unless it is a part of vital experience, religion becomes "fossil conventionalism." Here James shared his father's attitude; his father wrote so much, according to James, because he was dissatisfied with every verbal encapsulation. Writing was a necessary evil and, like the labor of Sisyphus, self-stultifying.

That James could not accept in any unqualified way the religious vision of his father is evident. The difficulty is simply this: "Any absolute moralism is a pluralism; any absolute religion is a monism" (*The Literary Remains,* p. 118). The recognition of the essential opposition of morality and religion was clearly made by the elder James. The logic of his system required him to reject the finite moral agent with his frantic moral efforts. It is certain that James benefited from his father's insight even though he aligned himself with morality and pluralism. The working attitude of the healthy mind must always be, for James, a moral one which takes seriously the difference between good and evil and which commits itself to struggle for the first and against the second. To adopt the religious attitude is to step out of life's fight and to justify that withdrawal by some belief about the character of the world and either the ineffectiveness or superfluity of action within it. For James the character of the world, the nature of reality, does not justify, as a general attitude, the quietism that religion counsels. On the contrary, the world is the kind of place in which moral endeavor is, as a rule, supremely worthy. James neither denied the satisfaction that religion gives to many nor declared that satisfaction illusory. The very fact of pluralism allowed him to suppose at least some aspect, however fragmentary, of reality that justifies the religious option. Religious belief gives us, in James's famous phrase, a "moral holiday." Like any holiday it may be enjoyed for its own sake; more

important to James, however, holidays indirectly affect the work week.

EVOLUTIONARY THEORY. James was strongly influenced by the Darwinian theory of evolution and was therefore predisposed to find in all feelings, including religious feelings, clues about what the world is like. Feelings that evolved in the world must somehow reflect the world. The most eccentric fancy, for example, tells us that we have the kind of world in which such a fancy is possible.

Evolutionary theory, as James saw it, begins with the presupposition that each part of reality has a function, that each part is in some way or other good for something or other. The strictly useless, according to such a theory, cannot endure, and all flourishing realities command a certain minimal respect. Religious experience is not especially justified by evolution because nothing is *especially* justified. Religion and irreligion, insofar as they both exist, are exactly equal before the evolutionary tribunal. Belief in evolution, at least as James interpreted that belief, makes simple dismissal impossible; even that which is evil cannot be negligible. The questions must be asked of religion as it must be asked of everything. How is it that it came to be what it is? What is it for?

ANTECEDENTS AND VALUE OF RELIGION. In his major work on religion, *The Varieties of Religious Experience,* James attempted to account for the antecedents and value of religion. The question of how it came to be what it is, is a matter of classifying religious feelings and religious propensities with other kinds of human experience which are found to be similar to them. The initial task, therefore, of *The Varieties* is the provision of a "descriptive survey" beginning with as many and as varied examples of typically religious experience as possible. The emphasis, here as elsewhere, is on spontaneous religious emotions rather than theological interpretations or institutional prolongation and regularization.

James was scrupulously careful to explain religious phenomena by ordinary scientific laws and principles, if at all possible. Accordingly, religious visitations of all kinds are classed as sudden influxions from the subject's own subconsciousness. Conversion is seen as the radical rearrangement of psychic energy around some new center of interest. Examples of this kind of felicitous theorizing could be multiplied.

James, however, was equally concerned with promoting the thesis that nothing said about the history or genesis of religious phenomena can shed the slightest light on the spiritual worth and significance of those phenomena. The older dogmatists attempted to justify religion once and for all by pointing to its privileged origin in some kind of revelation; newer dogmatists—the "medical materialists"—attempted to discredit religion once and for all by pointing to its disreputable origin in some curious bodily state. Neither approach is acceptable. Religion must be judged in the same way that everything else is judged, by proving itself useful (in specifiable ways) in some possible future. Religion must "run the gauntlet of confrontation with the total context of experience" (*The Varieties,* p. 426). This context includes the collection of all our established truths as well as all the exigencies of our affective and intellectual natures. Therefore, the defense of religion that can be found in James is not based on appeals to either mere social utility or subjective feeling. The question of the truth of religion arises only when religion makes some concrete, specific prediction about the world's future. Religion having framed its hypotheses, these hypotheses are supported or refuted in terms set out by James's general theory of belief, known as pragmatism.

BELIEF. James's notorious defense of the right to believe in the widely read essay "The Will to Believe" and elsewhere, though generally given a limited religious interpretation, is, in fact, not primarily a defense of religious belief but of moral belief, belief in the efficacy of action, including, as an important instance, the active experimentalism of modern science. The point of James's doctrine is its repudiation of the methodological caution epitomized by the Baconian injunction not to "suffer the understanding to jump and fly from particulars to remote and most general axioms" or by the Cartesian rule "that the understanding should always know before the will makes a decision." James was making a general statement in support of the method of empirical science, with special emphasis upon the initially unwarranted character of every scientific hypothesis. We must at least believe our hypotheses sufficiently to bestir ourselves to test them; without our active interest in and partisanship of belief the enterprise of science would come to a silent, ghostly end. It is the theoretical daring of science that inspired James. His doctrine on the will to believe is no fuzzy ad hoc concession to self-indulgent piety but an integral part of his general theory of belief.

The doctrine of the will to believe, with all its genial encouragement of risking belief, is balanced, in James, by an unremitting fallibilism. Belief, however justified originally, is always conditional. Belief must continue to justify itself; there is no possibility of a definitive, once and for all certification. Both the options of practical life and the

tenets of religion may be justified as peculiar kinds of scientific hypotheses, the first sort peculiar because of their limitation to some particular matter or situation, the second because of their elusive generality. There seems to be no difficulty in interpreting the practical decisions of life, with their inherent predictions about relevant future events, as closely analogous to the predictive, if not to the explanatory, activity of science. Religious belief, on the contrary, may seem intrinsically isolated from the arena of confirmation and disconfirmation and, therefore, alien to the scientific pattern. For James all genuine belief, including religious belief, must address itself to the tribunal of experiment. If all possible procedures of verification are irrelevant to some religious doctrine, then that doctrine cannot rightly be the object of any belief; such a doctrine, having no positive content, would be meaningless.

James did, in fact, think that at least a few religious hypotheses were truly empirical, that they made a difference which somewhere could be noticed. James was careful not to prejudice the case against religion by adopting some single restrictive paradigm of verification. If religious belief makes a difference, it is not altogether surprising that we should have to look for that difference with greater sympathy, imagination, and patience than we are used to exercising in more straightforward cases.

JAMES'S RELIGIOUS BELIEF. James's own religious belief, expressed without dogmatism in the last chapter and the postscript of *The Varieties* and again in the last chapter of *Pragmatism,* consists essentially in the affirmation that the world is richer in realities than conventional science is willing to recognize. Religious experience at least suggests that there is what James called a "higher part of the universe" (ibid., p. 516) which, though beyond the immediate deliverance of the senses, is nevertheless effective in the world in a way that makes a noticeable difference. This assertion that the higher part makes concrete and local differences constitutes James's famous "piecemeal super-naturalism" (ibid., p. 520), really only a name for an enlarged and tolerant naturalism. The higher part is perhaps impossible to define given the present state of our knowledge. Certainly, for James it cannot be the infinite and omnipotent God of traditional theism who guarantees the successful outcome of the universe. The higher part is better conceived as a finite power (or perhaps even a polytheistic medley of powers) that, like men, works toward the good and helps achieve it. This is a theological notion compatible with the significance of moral choice in a way that the conventional notion is not.

The vagueness of much of James's treatment, a vagueness he frequently admitted to, has been amply noted by his critics. What must also be noted, however, is the forceful way in which a fundamental idea of our tradition, the idea of God, has been radically reconstructed by James in a manner that makes the idea more consonant with religious experience and that frees it from the congeries of paradoxes associated with the problems, among others, of free will and of evil.

PRAGMATISM

The chief locus for James's pragmatism is, of course, his immensely popular and influential work *Pragmatism: A New Name for Some Old Ways of Thinking.* The origin of pragmatism, however, as James always acknowledged, is found in C. S. Peirce's essay "How to Make Our Ideas Clear," published in 1878. This essay remained generally unnoticed until James's 1898 lecture on pragmatism, "Philosophical Conceptions and Practical Results," at the University of California (in *Collected Essays and Reviews*). This lecture may be taken as the beginning of pragmatism as an explicit, although never a unified, movement, but the essentials of the doctrine as developed by James are found earlier in the *Principles of Psychology* and even in the introduction to *The Literary Remains of the Late Henry James.* Indeed, James rarely wrote anything, early or late, which did not at least imply pragmatism.

Pragmatism may be approached as a mere method, an eristic device which vouchsafes hints as to either the meaning or truth of propositions or to both together; it may be taken as a theory of meaning or a theory of truth or, once again, as a theory of both meaning and truth. A. O. Lovejoy in "The Thirteen Pragmatisms" insisted upon distinctions such as these and chided James for neglecting them. In fact, though James erred in emphasizing the autonomy of the various aspects of pragmatism, he wished to persuade his readers of the truth of whichever part he was recommending at the moment, and he therefore tended to stress the self-contained plausibility of elements which, if plausible at all, are so only when taken together in the total view.

It is the contention of James's sympathetic commentators that his pragmatism is plausible as nothing less than a theory of reality. It is the descriptive naturalism central to James that saves pragmatism from being merely a convenient device for settling philosophical disputes. The fundamental assumption that generates pragmatism is the assumption that "knowledge," "truth," and "meaning," as well as any other possible object of discourse or any other possible subject matter for philosophical dis-

cussion, must be explicable as a natural process or as a functional medley or competition of natural processes. The world, for James, is a plurality of temporal processes related in so many specifiable and concrete ways that it cannot be accounted for by abstract speculation alone.

James believed that an individual's personal, peculiar vision counts most in philosophy; not surprisingly, it is vision, not method, which is primary in James. Reality dictates the method by which it may be known. The gross encounter with the world is primary in the determination of what character the world will have for us. Theories of knowledge and of method, existing at a high level of abstraction, are second to the ineluctable fact of experience breaking in upon us.

TRUTH AND MEANING. James's pragmatism is an attempt to formulate a metaphysics of truth and of meaning. Logically, such an attempt is exactly on a par with the metaphysical treatment of any discriminable subject matter. By metaphysics James meant the quest for adequate general descriptions either of reality as a whole or of some distinguishable part of it. The descriptions offered by metaphysics are, in principle, continuous with those offered by science, although their range and focus may differ. The distinction between science and metaphysics was not crucial for James; he saw the possibility of unrestricted intercourse and cooperation exactly where later thinkers are likely to see division and competition for cognitive respectability. It is therefore helpful from James's own point of view to regard pragmatism's description of truth in the same light, say, as geology's description of continental drift. Both are characterizations of natural processes, and both attempt to portray what actually happens.

The metaphysical perspective of *Pragmatism* itself (even apart from the context of James's total work) is so unmistakable that the prevailing interpretation of pragmatism as a set of newly devised rules that serve a certain practical purpose seems totally unjustified. If James was right, people have always unwittingly followed the "pragmatic method." A purely theoretical illumination like pragmatism will indeed clarify practice and improve it; for James no process—least of all, a process where human influence intervenes—is so canalized that modifications are utterly beyond hope. Metaphysics must recognize the plasticity of its subject matter as well as the limits of plasticity.

Pragmatism discusses truth without falling into the epistemological frame of mind habitually assumed by professional philosophers. James's description of actual processes rejects the usual question of what we ought to believe. If there is something we "ought" to believe, the authority of the "ought" itself must be explained concretely. There is no authority which is merely formal. Pragmatism therefore becomes the justification of truth's prestige in terms of the world's exigencies.

One factor discernible in the complex process called "believing truly" is the compulsion of fact or the unavoidability of a residual nonplastic pole in determination of what is true. It is here that we find truth's authority and importance.

Truth, for James, is what we must somehow take account of if we are not to perish. Men cannot in the long run believe what is false not because truth extracts from them a categorical imperative in its own behalf but because reality compels men in spite of themselves, and it is from this that the authority of truth is derived. "Agreement with reality" as a criterion of truth cannot be taken to indicate any fixed structural relation (such as the "copying" relation). The truth relation is characterized not by stasis but by the fluid resourcefulness of functional harmony. The character of the harmony itself may be anything that is compatible with survival. Even in the Darwinian world that James pictured, there is more than one way to survive as truth.

Raw compulsion may account for the authority of truth, but truth is hardly a mute registry of bruises received from the world. Indeed, people create truth, and truth is so exclusively the result of human activity that James's own view has been called "humanism." Central to this humanism is the distinction (so often insisted on by James, so often neglected by his critics) between ideas and objects, between what takes account and what is taken account of. The objects constitute what James referred to as the "unhumanized fringe," the yet to be conceptualized. What must be taken account of is presumably just what it is. Truth and falsity, however, apply not to objects but only to our ideas of objects. Our ideas of objects are mutable in the sense that we can modify ideas or replace one idea by another. In such a situation ideas are to be judged better or worse; such judgments fall between the ideal limits of complete good and complete bad. These are the same limits usually called "truth" and "falsity." Truth is viewed by James as one species of the good. The good is itself interpreted as a plurality of "good fors." In this view ideas are instruments for taking account theoretically, practically, aesthetically, and so on, of reality.

The point of James's view of truth, as Bergson suggests in *The Creative Mind* (p. 256), is that truth is to be described as an invention rather than a discovery. Truth,

or propositions which are true, might be compared to cleverly made maps or apt predictions. If they serve us as we expect them to serve us, we have no legitimate complaint. There are, of course, ontological relations between such inventions as maps, predictions, and propositions (as well as inventions such as light bulbs and cotton gins) and what, in summary fashion, is referred to as reality. Inventions are conventional but not arbitrary. They are not arbitrary because they must somehow take account of reality; they are conventional because they embody one way (among alternatives) for that taking account.

The relationship between two processes within experience constitutes truth—(1) the inventive process or activity of proposing, of framing propositions, and (2) the particular chain of natural processes with which the proposition in question is concerned. The emphasis on the truth relation as a relation within experience and totally construable in terms of "positive experienceable operation" (*Meaning of Truth,* p. x) is one instance of James's general metaphysical position that all relations are within experience. Experience, as it were, forms a cohesive, self-explanatory whole; it hangs together, as James liked to say, and needs no transcendental connectives or supports.

Since the truth relation was taken by James's contemporaries as transcending experience, the strategic function of pragmatism is apparent. It is an extension of radical empiricism, an attempt to place the particularly troublesome truth relation within the total perspective of metaphysical naturalism.

James spoke of true ideas as those which "work," which "lead" propitiously, which give various kinds of satisfaction, and which bring about various kinds of success. He also spoke approvingly of the "cash value" of ideas and thought that meaningful ideas are those which make "practical differences." These highly (and obviously) metaphorical expressions have confused many commentators.

There are those who have found James vague. He intended, however, that all these metaphors should be functionally specific and indeterminate only in respect to instances. "Working," "leading," "satisfying," and "succeeding" are generic terms as respectable and as precise as terms such as "copying" and "agreeing." They are, however, functional rather than static. For those who see functions as inherently insubstantial, shadowy, and vague, any functional definition of truth will be unacceptable, but this hardly seems to be an insurmountable objection.

Other commentators have seized upon the prominence of the "practical" in James's account of meaning and truth. Surely, this is a difficult term in James, if for no other reason than that he used it as it is used in ordinary language—that is, variously. His prevailing usage, however, cannot be equated with some narrow notion of commercial efficiency. Pragmatism is not a philosophical vindication of the businessperson's common sense or acumen. It was James, after all, who saw the tendency to worship "the bitch goddess, success," as the principal weakness in the American character. It is especially in our theoretical and moral practice that meaningful ideas, according to James, are to make a difference. Belief divorced from action may well be morally effete, and James set forth this point, though not in its crudely athletic form; his main thesis, however, was that belief divorced from action is *theoretically* inexplicable. James's quest was not for a formula that would rouse his fellows to civic virtue or efficiency of some peculiarly American sort but for criteria which would be descriptively adequate to belief. His philosophical purpose was to find out what it means to believe, what it means to entertain ideas which may be meaningful and true.

METAPHYSICS

Although frequently attempted, it is not possible to isolate a final "metaphysical" period in James. The theory of the various kinds of belief, which formed his philosophy of religion and of pragmatism, has as a conspicuous feature the assumption that anything which can be meaningfully said about belief must take into account the grounding of belief in natural processes, particularly human processes. It is possible to formulate a theory of belief apart from a general metaphysics only by adopting an assumption that James explicitly rejected. This is the epistemological assumption that an existentially neutral *logic* of belief can be constructed. In fact, on this assumption existential reference is regarded as the indication of a certain categorial confusion frequently labeled "psychologism." James insisted, even in his least metaphysical passages, that "knowledge," "belief," "truth," and "meaning" indicate discriminable natural existences in the same way that all terms do, or at least all terms that figure as possible subjects of philosophical discussion. This is simply the corrective application of the basic postulate of metaphysical naturalism to the recalcitrant subject matter of epistemology. James regarded the prominence accorded this subject matter since the time of Kant as a distortion of perspective that his own philosophy was intended to correct.

But for the development of James's metaphysics, the psychology—or the treatment of characteristically human processes—was even more important than the theory of belief. His metaphysics was simply the attempt to apply to all reality categories originally framed for human experience. The radical generalization of the concept of experience, so central in James, is necessitated by two ideas. First, James believed that metaphysical dualism is always unacceptable. Whatever dualities or pluralities are distinguished for certain purposes, ultimately the philosopher cannot operate with irreducible categories. Second, if one categorial set or one metaphysical model must be adopted, James believed that this categorial set or metaphysical model must arise from the consideration of our own experience. It is only of human experience that we have anything like "complete concrete data." Anthropocentrism is therefore thought to be a consequence of any genuine empiricism. For James even panpsychism is at least a possible and interesting empirical hypothesis.

In the seminal essay "Does Consciousness Exist?" (1904), James asks us to assume that there is just one "primal stuff" of which everything in the world is made. This stuff, called "pure experience," is not a single entity, like Thales' water; "pure experience" is a collective name for all sensible natures, for all the "that's" which anywhere appear. The monism implied in this concept of the *one* primal stuff is therefore merely formal. Explanatory monism must be accepted before specific metaphysical descriptions may be attempted. In the same essay James provided a sample of metaphysical description. Consciousness is there described as a certain relation of parts of experience to one another. Consciousness is not an unanalyzable substance but simply the name that is given to a certain discriminable function within experience, the knowing function. All other functions are to be explained in the same way as consciousness. Functional explanations in terms of related strands of experience allow the abrogation of traditional dualisms because the same isolable part of experience may enter into many and various relations. What is subject may also be object; what is object may be subject. The knower may also be the known and vice versa, depending on the "context of associates" within which the part of experience so labeled is considered.

James's frequent use of the expression "part of experience" was not meant to suggest that experience has an atomistic constitution. Indeed, James constantly argued against the "pulverization" of experience in British empiricism. We experience not isolated parts but continuities of indeterminate extension. Parts and the relations between parts, both directly experienced, form new functional wholes. The use of the word *part* indicates nothing more than the theoretical and practical need for emphatic focus.

James regarded the "concrete" data appealed to by British empiricism as abstract, intellectual products; he accused that empiricism of committing what Whitehead later called "the fallacy of misplaced concreteness." If James's philosophy is to be classified historically as a criticism of British empiricism, it must also be emphasized that it is self-consciously offered as an alternative to the criticism of empiricism by idealists from Kant to Royce.

If the facts pointed to by the usual empiricism are abstract in the sense of being incomplete, inadequate, or partial, it still cannot be said, as it is said by absolute idealism, that there are no facts at all or that there is just one fact, the immovable "block-universe," as James referred to this notion that he always found slightly ridiculous. There are no general grounds, according to James, for the rejection of the obvious particularity and individuality that characterize the plural parts of experience. James certainly held that any allegedly self-sufficient fact may turn out from some point of view or for some purpose, intellectual or practical, to be partial or abstract. But there are many points of view and many purposes with equal titles to rationality. There are therefore many levels of fact, and words such as "part," "whole," "unity," "concrete," "abstract," "particular," and "individual" do not qualify any reality simply or always. These words are definable only within purposive contexts. Absolute idealism, in contrast, sets up a single standard of rationality and develops a characteristic vocabulary which it applies *simpliciter.* This procedure yields a certain clarity and neatness but suffers from "vicious intellectualism" or "*the treating of a name as excluding from the fact named what the name's definition fails positively to include*" (*A Pluralistic Universe,* p. 60).

The notion of self-sufficient centers within experience emphasized by James as particulars or individuals is a generalization of that first trait of the stream of thought referred to in *Principles of Psychology.* Although made familiar by Whitehead, it was James who first used the concept of personal order to replace the traditional concept of some fundamental and thinglike substance.

Other traits of existence that impressed themselves on James are first annunciated in the *Principles* as traits of the stream of thought or of the central human process. So, for example, the doctrine that thought is always changing becomes the doctrine that reality is always changing. Again, human freedom is eventually inter-

preted as a special case of universal indeterminism. My future, though continuous with my past, is not determined by it. Just so the future of the world; although it grows out of the total past, it is not a mere result of that past. If I am creative—that is, if human freedom is effectual—then the world is creative, if for no other reason than that I am part of the world. What is constant in my behavior is the result of habits that never entirely lose their flexibility. In the same way the constancies charted by the laws of science are only more inveterate habits.

Objections can be raised against all these contentions, especially in the enthusiastic, unguarded form in which James made them. They do, however, add up to a serious philosophical position which has, in fact, borne fruit in the subsequent history of philosophy and is worthy of continuing serious study.

See also Bergson, Henri; Bradley, Francis Herbert; Determinism and Indeterminism; Dewey, John; Emerson, Ralph Waldo; Empiricism; Evolutionary Theory; Hodgson, Shadworth Holloway; Introspection; James, Henry; Kant, Immanuel; Panpsychism; Peirce, Charles Sanders; Perry, Ralph Barton; Philosophy of Religion; Pragmatism; Psychology; Rationality; Renouvier, Charles Bernard; Royce, Josiah; Santayana, George; Whitehead, Alfred North; Wright, Chauncey.

Bibliography

WORKS BY JAMES

Only James's chief works are listed here. For a complete list see Perry's bibliography, below. Secondary literature is copious, and only the most important works are listed here. Those who wish to sample the periodical literature are advised to consult the fifty-year (1904–1953) index of the *Journal of Philosophy* (New York, 1962), under the heading "James."

The Literary Remains of the Late Henry James. Edited and with an introduction by William James. Boston: Houghton Mifflin, 1885.

Principles of Psychology, 2 vols. New York: Holt, 1890. Regarded by many as James's major work, it is a prime source not only for his psychology but also for his metaphysics.

The Will to Believe and Other Essays in Popular Philosophy. New York: Longman, 1897. In addition to the title essay, the essay "The Sentiment of Rationality" is, if interpreted in the context of James's total thought, an important source of his basic convictions. The book also contains the famous essay "The Dilemma of Determinism," James's fullest statement of his views on free will.

The Varieties of Religious Experience: A Study in Human Nature. New York: Longman, 1902.

Pragmatism: A New Name for Some Old Ways of Thinking. New York: Longman, 1907.

The Meaning of Truth: A Sequel to "Pragmatism." New York: Longman, 1909. A collection of polemical essays; the preface is especially important, for it answers certain criticisms of pragmatism and states James's conception of the relation between pragmatism and radical empiricism.

A Pluralistic Universe. New York: Longman, 1909. A sustained criticism of absolute idealism and intellectualism; contains chapters on G. W. F. Hegel, Gustav Theodor Fechner, and Bergson.

Some Problems of Philosophy: A Beginning of an Introduction to Philosophy. New York: Longman, 1911. James's last project; it is incomplete. Valuable for its many very clear formulations; three chapters outline his theory of perception.

Essays in Radical Empiricism. New York: Longman, 1912. A related series of essays expounding James's mature philosophy; the essays "Does Consciousness Exist?" and "A World of Pure Experience" are especially important.

Memories and Studies. New York: Longman, 1912. Fifteen popular essays and addresses selected by James's son Henry James; includes commemorative addresses on Agassiz and Emerson, an essay on Spencer, and several essays on psychical research and academic life.

Collected Essays and Reviews. New York: Longman, 1920. Thirty-nine articles, selected by Ralph Barton Perry, extending from 1869 to 1910; many historically important works, including the California lecture "Philosophical Conceptions and Practical Results."

Letters. Edited by Henry James. Boston: Atlantic Monthly Press, 1920. Charming letters, primarily of biographical and historical significance, edited by James's son.

WORKS ON JAMES

Aiken, H. D. "American Pragmatism Reconsidered: William James." *Commentary* (September 1962): 238–266.

Bergson, Henri. "On the Pragmatism of William James: Truth and Reality." In his *The Creative Mind,* translated by Mabelle L. Andison. New York: Philosophical Library, 1946. Brief but provocative development of the thesis that to understand James's pragmatism, we must modify our general conception of reality.

Dewey, John. *Characters and Events,* Vol. I. New York: Holt, 1929. Book I, Ch. 12. "William James" consists of three occasional pieces that together provide an informal but penetrating analysis of James's contribution to and place in American philosophy.

Lovejoy, Arthur O. *The Thirteen Pragmatisms and Other Essays.* Baltimore: Johns Hopkins University Press, 1963. In addition to the very influential title essay, there are eight other essays on James or pragmatism.

Moore, G. E. *Philosophical Studies.* London: Routledge, 1922. "William James' 'Pragmatism'" attempts to refute James from the commonsense point of view.

Perry, Ralph Barton. *Annotated Bibliography of the Writings of William James.* New York: Longmans, Green, 1920. A listing of 312 items from 1867 to 1920, with helpful indications of each item's content and value.

Perry, Ralph Barton. *In the Spirit of William James.* New Haven, CT: Yale University Press, 1938. A solid and sympathetic philosophical interpretation of James. The chapters "An Empirical Theory of Knowledge" (Ch. 2) and "The Metaphysics of Experience" (Ch. 3) are especially noteworthy.

Perry, Ralph Barton. *The Thought and Character of William James.* 2 vols. Boston: Little Brown, 1935. A massive, richly documented study; the single most important work on James.

Royce, Josiah. *William James and Other Essays on the Philosophy of Life.* New York: Macmillan, 1911. An attempt to show James's place in American social history.

Russell, Bertrand. *Philosophical Essays.* New York: Longman, 1910. See "William James' Conception of Truth" and "Pragmatism."

Santayana, George. *Character and Opinion in the United States, with Reminiscences of William James and Josiah Royce and Academic Life in America.* New York: Scribners, 1920. Unsympathetic but extremely interesting in its own right.

William James Earle (1967)

JAMES, WILLIAM [ADDENDUM]

William James is to classical American philosophy as Plato was to Greek and Roman philosophy: an originating and inspirational fountainhead. Thinkers as diverse as C. S. Peirce, Josiah Royce, John Dewey and the late work of A. N. Whitehead took their point of departure from William James, especially his monumental *Principles of Psychology*. Influential philosophers elsewhere were also deeply influenced by James, for instance Henri Bergson, Edmund Husserl, Miguel de Unamuno, and Ludwig Wittgenstein.

With the completed publication of all of James's writings, including his manuscripts and notebooks, the full range and philosophical virtuosity of his work comes into focus. For too long the thought of William James was taken to be novel and intriguing but lacking in technical sophistication. In reading James the first response is one of elation at the apparent simplicity and obvious elegance of the literary style. After several careful and close readings, however, the philosophical depth and complexity emerge. The consequence of these more mature readings of James's thought are now found in many areas of contemporary philosophy—for example, the philosophy of mind, ethics, and the philosophy of religion. More significant still is that James represents a helpful philosophical stance, one that is wary of narrowness and rigid conceptual schematisms and affirms the messages of human experience no matter the source. William James believes that philosophy itself is "the habit of always seeing an alternative" ("Essays in Philosophy," *Works*, 1978, p. 4). He was convinced as well that no matter how recondite the issue in question—for example, the meaning of consciousness or his innovative doctrine of radically empirical relations—the kernel of the position taken could be articulated in prose accessible to the intelligent reader as well as to the philosopher.

The most salutary result of recent commentaries on the philosophy of William James has been the rescue of two of his most beleaguered positions, that of the pragmatic theory of truth and his doctrine of "The Will to Believe." In both areas James's thought was often subject to mocking dismissal and shallow interpretations. With the completion of James's *Works*, the girth and sophistication of his philosophy is now apparent. Witness, for example, the sterling introductory essays by H. Standish Thayer on James's theory of truth as found in "Pragmatism" (*Works*, 1975) and "The Meaning of Truth" (*Works*, 1975). Similarly, one finds an equivalently clarifying essay by Edward H. Madden in his introductory essay to "The Will to Believe" (*Works*, 1979).

The divide that has existed between mainstream analytic philosophy and pragmatism is no longer purposeful. Transformations of this conflict are now at hand. Hilary Putnam, for decades a major figure in contemporary philosophical thought, writes in his *Pragmatism* (1995):

> I believe that James was a powerful thinker, as powerful as any in the last century, and that his way of philosophizing contains possibilities which have been too long neglected, that it points to ways out of old philosophical "binds" that continue to afflict us. In short, I believe that it is high time we paid attention to Pragmatism, the movement of which James was arguably the greatest exponent.

Although in no way gainsaying the importance of specific philosophical contentions held by James, nonetheless it can be said that the most signal reason for paying serious attention to this work is found in his philosophical attitude, his approach to philosophical inquiry. William James was no stranger to philosophical debate or argument, as one finds in his brilliant and jousting correspondence with F. H. Bradley. Yet James was uneasy about closure, answers, and finality of any kind. In a "Notebook" entry of 1903 James writes of "bad taste," by which he means:

> All neat schematisms with permanent and absolute distinctions, classifications with absolute pretensions, systems with pigeon-holes, etc., have this character. All 'classic,' clean, cut and dried, 'noble,' fixed, 'eternal,' *Weltanschauungen* seem to me to violate the character with which life concretely comes and the expression

> which it bears of being, or at least of involving, a muddle and a struggle, with an 'ever not quite' to all our formulas, and novelty and possibility forever leaking in.

For the thought and person of William James, the novel call of experience inevitably trumps categories of explanation. Consequently, possibility rather than solution becomes the philosophical watchword, especially in matters of profound human importance.

See also Bergson, Henri; Bradley, Francis Herbert; Dewey, John; Husserl, Edmund; Peirce, Charles Sanders; Philosophy of Mind; Philosophy of Religion; Pragmatism; Pragmatist Epistemology; Royce, Josiah; Unamuno y Jugo, Miguel de; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

Since the original publication of *The Encyclopedia of Philosophy*, virtually all of William James's writings have been published in a critical edition. Under the general editorship of Frederick Burkhardt, see *The Works of William James*, 19 volumes (Cambridge, MA: Harvard University Press, 1975–88). William James is widely admired for his brilliant style of writing, and nowhere is that more apparent than in his letters. To that end, with John J. McDermott as general editor and edited by Ignas Skrupskelis and Elizabeth Berkeley, a critical edition of *The Correspondence of William James* has been published in 12 volumes (Charlottesville, NC: University of Virginia Press, 1992–2004). Of commentaries, the finest and most thorough is that of Gerald Myers, *William James: His Life and Thought* (New Haven, CT: Yale University Press, 1986). Other recent studies of note include: Samuel Henry Levinson, *The Religious Investigations of William James* (Chapel Hill, NC: North Carolina University Press, 1981); George Cotkin, *William James: Public Philosopher* (Baltimore: Johns Hopkins University Press, 1990); Charlene Haddock Seigfried, *William James's Radical Reconstruction of Philosophy* (Albany, NY: State University of New York Press 1990); Ruth Anna Putnam, ed., *The Cambridge Companion to William James* (Cambridge, U.K.: Cambridge University Press, 1997); David C. Lamberth, *William James—The Metaphysics of Experience* (Cambridge, U.K.: Cambridge University Press, 1999); Eugene Fontinell, *Self, God, and Immortality: A Jamesian Investigation* (New York: Fordham University Press, 2000); and Phil Oliver, *William James's Springs of Delight* (Nashville: Vanderbilt University Press, 2001). For an "Annotated Bibliography of the Writings of William James" and complete selections from his major works, see John J. McDermott, ed., *The Writings of William James* (Chicago: University of Chicago Press, 1977). The entire family is chronicled by R. W. B. Lewis in *The Jameses: A Family Narrative* (New York: Farrar, Straus and Giroux, 1991). Biographically, see Linda Simon, *Genuine Reality—A Life of William James* (New York: Harcourt Brace and Co., 1998) and Robert Richardson, *William James* (Boston: Houghton Mifflin, 2006).

John J. McDermott (1996, 2005)

JANKÉLÉVITCH, VLADIMIR
(1903–1985)

Vladimir Jankélévitch, the French moral philosopher, was born in Bourges. He was educated at the Lycée Louis-le-Grand and the École Normale Supérieure. Having become an agrégé in philosophy in 1926, he took his doctorate in 1933. After teaching at the French Institute in Prague and at various *lycées*, he served as lecturer at Toulouse from 1936 to 1937 and at Lille from 1938 to 1939. He was dismissed by the Vichy government in 1940 but returned to academic life in 1945 as professor at Lille, going from there to the University of Paris as professor of morals and moral philosophy.

Jankélévitch's philosophy is highly individual, though it displays a sympathetic understanding of widely divergent philosophical traditions. In content it has affinities with Christian morality and with the philosophy of Søren Kierkegaard. In expression it is idiosyncratic and always lively.

Jankélévitch's first notable work was *Henri Bergson*. On its first appearance, in 1931, it bore a prefatory note by Henri Bergson himself, praising its "intellectual sympathy." Jankélévitch's own philosophy made its first appearance in his main doctoral thesis, on Friedrich von Schelling's later philosophy, and even more clearly in his secondary thesis, on bad conscience (*La mauvaise conscience*). Bad conscience is consciousness directed not unreflectingly forward but regretfully backward toward its own past, which is irremediable because time is irreversible. The problem posed is how to restore the flow of living that tends to be halted by retrospective brooding. How is consciousness freed and time unfrozen? Jankélévitch did not favor the detachment from one's predicament effected by irony, precisely because it intellectualizes and detemporalizes that predicament. Time alone, in its flow, frees us.

In two of his postwar works, *Philosophie première* and *Traité des vertus*, Jankélévitch was perhaps at his best. Just as he rejected intellectual recourse to irony or conceptualization as consolations for the discontent attendant upon self-consciousness, so he showed, in *Philosophie première*, that the concern of metaphysics is not with the world of ideas, eternal truths, or transcendent models, which are ultimately as contingent as the

reality that they rationalize, but with the "entirely other Order" of radical contingency. Here, in effect, Jankélévitch suggested that "sufficient reason" is never really sufficient. The instant always brings novelty over and above the schemata that demonstrate its "necessity."

The real importance of this fact is moral and leads to the treatise on the virtues. In this work virtues are classified according to either their intellectual quality of equity or their "non-natural" quality of goodness, to use the language of G. E. Moore. For Jankélévitch the virtues of consistent conduct—the "virtues of the interval," fidelity and justice—are inferior to the creative "virtues of the instant," courage and charity.

See also Bergson, Henri; Conscience; Ethics, History of; Kierkegaard, Søren Aabye; Moore, George Edward; Schelling, Friedrich Wilhelm Joseph von; Virtue and Vice.

Bibliography

WORKS BY JANKÉLÉVITCH

Henri Bergson. Paris: Alcan, 1931; rev. ed., Paris, 1959.
La mauvaise conscience. Paris: Alcan, 1933.
L'odyssée de la conscience dans la derniére philosophie de Schelling. Paris: Alcan, 1933.
L'ironie ou la bonne conscience. Paris: Alcan, 1936; 2nd rev. ed., Paris, 1950.
L'alternative. Paris: Alcan, 1938.
Du mensonge. Paris, 1943.
Le mal. Paris, 1947.
Traité des vertus. Paris: Bordas, 1949.
Philosophic première. Paris, 1954.
L'austérité et la vie morale. Paris: Flammarion, 1956.
Le je-ne-sais-quoi et le presque-rien. Paris, 1957; Paris: Editions du Seuil, 1980.
La mort. Paris: Flammarion, 1966.
Le pardon. Paris: Aubier-Montaigne, 1967.
Traité des vertus. Paris: Bordas, Mouton, 1968.
La vie et la mort dans la musique de Debussy. Neuchâtel: A la Baconnière, 1968.
Ravel. Paris: Éditions du Seuil, 1969.
L'ironie. Paris: Flammarion, 1972.
De la musique au silence. Paris: Plon, 1974.
L'irréversible et la nostalgie. Paris: Flammarion, 1974.
L'aventure, l'ennui, le sérieux. Paris: Aubier, 1976.
Le pur et l'impur. Paris: Flammarion, 1978.
Le paradoxe de la morale. Paris: Seuil, 1981.
Le sérieux de l'intention. Paris: Flammarion, 1983.
La musique et l'ineffable. Paris: Éditions du Seuil, 1983.
La présence lointaine: Albeniz, Séverac, Mompou. Paris: Editions du Seuil, 1983.
Philosophie première: Introduction à une philosophie du "presque". 2nd ed. Paris: Presses universitaires de France, 1986.
Les vertus et l'amour. Paris: Flammarion, 1986.
L'innocence et la méchanceté. Paris Flammarion, 1986.
L'imprescriptible: Pardonner?; Dans l'honneur et la dignité. Paris: Seuil, 1986.
La musique et les heures. Paris: Seuil, 1988.
Penser la mort?. Paris: Liana Levi, 1994.
Premières et dernières pages. Paris: Editions du Seuil, 1994.
Liszt: Rhapsodie et improvisation. Paris: Flammarion, 1998.
Music and the Ineffable. Princeton, NJ: Princeton University Press, 2003.

WORKS ON JANKÉLÉVITCH

Barthélemy-Madaule, Madeleine. "Autour du *Bergson* de M. V. Jankélévitch." *Revue de Métaphysique et de Morale* 65 (1960): 510–524.
Smith, Colin. *Contemporary French Philosophy*. London: Methuen, 1964. See pp. 181–201.
Suarès, Guy. *Vladimir Jankélévitch*. Lyon: La Manufacture, 1986.
Vax, L. "Du Bergsonisme à la philosophie première." *Critique* 11 (1955): 36–52.

Colin Smith (1967)
Bibliography updated by Michael J. Farmer (2005)

JANSENISM

Jansenism is a polemical term introduced by Jesuit critics to label those sympathetic to the theological views of Cornelius Jansen, the Louvain theologian and later bishop of Ypres. Supporters of Jansen protested that Jansenism is merely a "phantom" of the Jesuits and preferred to be called Augustinians. Jean Orcibal (1953) draws attention to the considerable difficulties in providing a precise definition of the term *jansénisme*. Even so, Jansenism can be understood in contrast to Jesuit theology, and the Jansenist movement did play a particularly significant political role in pre-revolutionary modern France. Moreover, Jansenism is of philosophical interest given its connections, both real and perceived, to Cartesianism.

JANSENIST THEOLOGY

Jansen's main theological work is his *Augustinus*, posthumously published in 1640. He called for a return to the emphasis in Augustine on the importance of the workings of grace in the salvation of the elect. He was explicit in rejecting the view of the Jesuits, defended in the sixteenth century by the Spanish Jesuit Luis de Molina, that though grace is needed for salvation, it is also necessary that the will freely cooperate with the working of grace. For Molina, such freedom requires an "indifference" that makes it possible for the will to reject divine assistance. In

response Jansen insisted that since the grace that heals the will is fully efficacious, it determines the will to meritorious action in a way that excludes indifference.

The Jesuit charge was that the theology of the *Augustinus* gives aid and comfort to a heretical Calvinist view that God determines one's salvation in complete independence of the works of one's free will. Jansen's defenders countered that the standard Molinist line supports the heretical doctrine of Pelagius, which Augustine and the early church had condemned as heretical, that one's salvation is due to one's free actions rather than to the workings of grace.

Pope Urban VIII initially condemned the *Augustinus* in the 1643 bull *In eminenti* on the technical grounds that it violated an earlier decree in Rome prohibiting inflammatory remarks concerning free will and grace. This decree was connected to a bitter dispute at the end of the sixteenth century that pitted Molina and the Jesuits against Domingo Bañez and other members of the Dominican order (of which St. Thomas Aquinas had been a member). To resolve the dispute, Pope Clement VIII established the Congregatio de Auxiliis (Congregation on Grace) in 1597 to determine whether Molina's views were heretical. At certain points there were rumors that the decision was about to go against Molina and the Jesuits. However, the congregation ended in a stalemate, and Pope Paul V closed it in 1607 with a decree forbidding either side of the dispute to charge the other with heresy. He further promised a resolution of the issue at "an opportune time," but the issue remains unresolved to this day.

In the case of Jansenism, however, the church took action. In 1653 Innocent X issued the bull *Cum occasione*, which condemned as heretical or temerarious the following five propositions that anti-Jansenist theologians in the Sorbonne had claimed to find in the *Augustinus*:

(1) Some of God's commandments are impossible for the just despite their desire and their effort [to keep them], given the forces that they have presently and also the lack of the grace that makes them possible.

(2) In the state of fallen nature, no one [can] ever resist interior grace.

(3) For merit or demerit in the state of fallen nature, it is not required that man be free from the necessity of willing and acting; it is sufficient for him to be free from constraint.

(4) The Semi-pelagians admitted the necessity of interior prevenient grace for all good works, even for the beginning of faith; but they were heretical in claiming that this grace is such that the human will may either resist or obey it.

(5) To say that Jesus Christ died and shed his blood for all men, without a single exception, is to speak as a Semi-pelagian (Denzinger 1963, p. 445f).

In 1656 Alexander VII closed a loophole created by *Cum occasione*, which failed to mention the *Augustinus* explicitly, and issued *Ad sacram*, which claimed that the propositions are to be found in Jansen's text in their condemned sense. Thus did Jansenism become a formally defined heresy within the Catholic Church.

FRENCH JANSENISM

Jansen was a friend of the French religious figure Jean Duvergier de Hauranne, abbé de Saint-Cyran, who served as the spiritual advisor to the reformist convent of Port-Royal des Champs. Though Jansenism began in Louvain, several theologians associated with Saint-Cyran and attached to Port-Royal as solitaires came to be identified as the leaders of the Jansenist faction. These individuals were known for their opposition to a moral laxism that they found in the work of the Jesuits. This aspect of Jansenism is most evident in the *Lettres provincials* (1656–1657) of the Port-Royalist solitaire Blaise Pascal (the brother of Jacqueline Pascal, a member of Port-Royal). Another famous solitaire, the Sorbonne theologian Antoine Arnauld (the brother of Jacqueline-Marie-Angélique Arnauld, the prioress of Port-Royal), wrote a defense of Saint-Cyran's penitential theology in the 1643 *De la fréquente communion*, and later was an active supporter of Jansen and the *Augustinus*.

Saint-Cyran had been an opponent of Cardinal Richelieu, the French first minister, and Cardinal Mazarin, Richelieu's successor, inherited a suspicion of those associated with Port-Royal. As part of a campaign against Port-Royal, Mazarin lobbied for an official condemnation of Jansenism in Rome. The effort resulted in *Cum occasione*, but Arnauld argued that the five propositions have both heretical and nonheretical senses and that the *Augustinus* endorses them only in their nonheretical senses. When Rome answered with the explicit condemnation of Jansen's text in *Ad sacram*, Arnauld then insisted on the distinction between questions of faith (*questions de droit*), on which the pope's word is authoritative, and questions of fact (*questions de fait*), on which the pope has no special authority. He appealed to this distinction in claiming that whereas Catholics must accept the pope's claim that the propositions are heretical, they are entitled to retain a "respectful silence" with respect to

the claim that the propositions are to be found in the *Augustinus*. Arnauld was expelled from the Sorbonne in 1656 for writing in support of the Duc de Liancourt (Roger du Plessis-Liancourt), who was refused absolution due to his failure to affirm the presence of the condemned propositions in the *Augustinus*.

After Mazarin's death in 1661, Louis XIV followed his former first minister's policy of suppressing Jansenism and started to put considerable pressure on the nuns and solitaires at Port-Royal. His attempts during the 1660s to impose an anti-Jansenist formulary both on those associated with Port-Royal and on the clergy brought the French church to the brink of schism. However, with the help of Pope Clement IX, Louis was able to institute a Peace of the Church in 1669 that allowed for the respectful silence that Arnauld had advocated. This brought to an end the predominance of what Louis Cognet (1968) calls "First Jansenism," for which theological issues were most crucial. Cognet contrasts this sort of Jansenism with a "Second Jansenism" that started to emerge after the end of the Peace of the Church in 1679 and that was more concerned with political issues. During the 1680s Jansenists such as Arnauld, then in exile in the Spanish Netherlands, took the side of the pope in political disputes between Paris and Rome. After 1700 Louis attempted to ally himself with anti-Jansenist elements of the Roman curia to bring an end to what he took to be a politically subversive form of Jansenism in the work of Pasquier Quesnel. His efforts led Pope Clement XI to issue the bull *Unigenitus* (1713), which condemned Quesnel's views. The Parlement de Paris initially refused to register the bull due to Gallican concerns that the sanctioning there of unjust excommunication compromised French sovereignty. Though Louis succeeded in having the bull registered and approved through intimidation, there was significant parlementary opposition to *Unigenitus* throughout the eighteenth century. Port-Royal could no longer be a source of support for this opposition, however, since Louis, with the encouragement of his confessor, the Jesuit père de la Chaize (Jean Chastain), disbanded the convent in 1709 and had it destroyed the following year.

During the 1730s Jansenism became associated with the *convulsionnaires*, so called because they experienced uncontrollable convulsions at the grave of François de Paris in the cemetery of Saint-Médard. This group of individuals claimed that the grave was the site of miracles that confirmed divine support for Jansenism. In response the French government closed the cemetery and countered Jansenist political opposition by making *Unigenitus* a law of state. After the 1650s, however, the Jansenist journal *Nouvelles Ecclésiastiques* played a prominent role in the parlementary campaign against the Jesuits, and this campaign ultimately resulted in the suppression of the Jesuit order in France in 1764 (see Van Kley 1975). Jansenist resistance to the French political establishment also arguably helped to set the stage for the French Revolution at the end of the eighteenth century, though the secularism that dominated the revolution was far removed from the Augustinian spiritualism that pervaded the Jansenist movement.

JANSENISM AND CARTESIANISM

Francisque Bouillier, the nineteenth-century historian of Cartesianism, claims that there is "a natural alliance of the doctrine of Jansenius with that of Descartes." His specific proposal is that this alliance derives from the fact that the Cartesians "make God the unique efficient cause, the only actor who acts in us," whereas the Jansenists "give everything to the grace that operates in us without us" (1868, vol. 1, p. 432f).

The association of Jansenism with Cartesianism goes back to the seventeenth century, as indicated by the remark in the 1690 *Voyage du Monde* of the Jesuit Gabriel Daniel that "there are very few Jansenists who are not Cartesians" (1690, p. 285). This association was due in no small part to the interest in Cartesianism at Port-Royal. Cartesianism infuses *La logique ou l'art de penser* (1662) of Arnauld and Pierre Nicole, a text that reflects the teaching in the *petite écoles* of the convent before their suppression in 1660. Moreover, issues involving Cartesianism were prominent at discussions among Port-Royalist sympathizers held at the hôtel of the Duc de Liancourt from 1669 to 1674.

Even so, it is important to keep in mind that René Descartes himself was never associated with Jansenism during his lifetime, although the *Augustinus* was published a decade before his death. The connection between Jansenism and Cartesianism is a genuinely post-Descartes phenomenon. It is also noteworthy that there was significant opposition to Cartesianism from within Port-Royal, as indicated in the record of the Liancourt discussions. Some of this opposition was due to a fear of the heretical implications of Cartesian natural philosophy, particularly with regard to the theology of the Eucharist. Another source of opposition was the belief, widespread among the Jansenists and arguably present in the *Augustinus* itself, that human reason can accomplish little on its own given the corrupted state of fallen human nature. Even Arnauld, who among the Port-Royalists was

the most supportive of Cartesianism and the use of reason in philosophy, criticized Descartes at one point for offering an account of human freedom in his correspondence that is "full of Pelagianism."

The historical realities are thus more complex than Bouillier's thesis suggests. There is need for caution even with respect to his proposal that the Cartesian claim that God is the only cause is naturally connected to the Jansenist emphasis that grace brings about one's meritorious action. Nicolas Malebranche was most responsible for the perception that Cartesianism leads to the occasionalist conclusion that God alone has causal power. But Malebranche insisted, against the Jansenists, that one has the freedom to reject divine grace. Arnauld in fact found such a view to be Pelagian, and he campaigned to have Malebranche's works placed on the *Index of Forbidden Books* in Rome, which they were in 1690. For his part, Malebranche noted the irony that he was being condemned because "I have refuted the opinions that the Church has condemned in Jansenius" (1958–1984, vol. 19, p. 548).

See also Arnauld, Antoine; Augustine, St.; Cartesianism; Descartes, René; Logic, History of; Malebranche, Nicolas; Molina, Luis de; Nicole, Pierre; Pascal, Blaise; Pelagius and Pelagianism; Thomas Aquinas, St.

Bibliography

Armogathe, Jean-Robert, and M. Dupuy. "Jansénisme." In *Dictionnaire de spiritualité*, 8:102–148. Paris: Beauchesne, 1974.

Arnauld, Antoine. *Oeuvres de Messire Arnauld*. 42 vols. Brussels: Culture et Civilisation, 1964–1967.

Bouillier, Francisque. *Histoire de la philosophie cartésienne*. 3rd ed., 2 vols. Paris: Delagrave, 1868.

Cognet, Louis. *Le jansénisme*. 3rd ed. Paris: Presses Universitaires de France, 1968.

Daniel, Gabriel. *Voyage du monde de M. Descartes*. Paris: Bernard, 1690.

Denzinger, Heinrich. *Enchiridion Symbolorum Definitionum et Declarationum de Rebus Fidei et Morum*. 32nd ed. Barcelona: Herder, 1963.

Jansen, Cornelius. *Cornelii Iansenii, Episcopi Iprensis, Augustinus*. 3 vols. in 1. Frankfurt: Minerva, 1964.

Malebranche, Nicolas. *Oeuvres Complètes de Malebranche*, 20 vols. Edited by André Robinet. Paris: J. Vrin, 1958–1984.

Molina, Luis de. *On Divine Foreknowledge: Part IV of the Concordia*. Translated by Alfred F. Freddoso. Ithaca, NY: Cornell University Press, 1988,

Orcibal, Jean. *Jansénius d'Ypres (1585–1638)*. Paris: Etudes augustiniennes, 1989.

Orcibal, Jean. "Qu'est-ce que le jansénisme?" *Cahiers de l'Association internationale des etudes françaises* 3–5 (1953): 39–53.

Orcibal, Jean. *Saint-Cyran et le jansénisme*. Paris: Éditions du Seuil, 1961.

Sainte-Beuve, Charles-Augustin. *Port-Royal*. 5 vols. Paris: N.p., 1840–1859.

Schmaltz, Tad M. "What Has Cartesianism to Do with Jansenism." *Journal of the History of Ideas* 60 (1) (1999): 37–56.

Sedgwick, Alexander. *Jansenism in Seventeenth-Century France: Voices from the Wilderness*. Charlottesville: University Press of Virginia, 1977.

Taveneaux, René. *Jansénisme et politique*. Paris: A. Colin, 1965.

Van Kley, Dale. *The Jansenists and the Expulsion of the Jesuits from France, 1757–1765*. New Haven, CT: Yale University Press, 1975.

Tad M. Schmaltz (2005)

JAPANESE PHILOSOPHY

The first, and perhaps the most interesting, question regarding Japanese philosophy is whether there is such a thing. Or, to be more precise, whether there was any Japanese philosophy before Nishida Kitarō's 1911 *An Inquiry into the Good* (*Zen no kenkyō*). Some Japanese scholars today, such as Sakamoto Hyakudai, deny that there has ever been any Japanese philosophy. Others, like Nakamura Yōjirō, argue that there was none before Nishida. This is somewhat surprising in that since 1920 much of the same literature originating from China has been called Chinese philosophy by the Chinese, while a little later a Korean version was labeled Korean philosophy by the Koreans.

To understand why the Japanese have not followed suit, we need to examine how the notion of "philosophy," as it is known in the West, first took shape in the Japanese intellectual world during the Meiji period (1868–1911). At that time the Japanese government was encouraging the wholesale importation of Western intellectual culture, including something called "philosophy." To designate this newly introduced Western study, Nishi Amane introduced in 1874 a new word, *tetsugaku* (a shortened form of "*kitetsugaku,*" " *kitetsu*" itself abbreviated from *kikyō tetsuchi*), which he formed using two Chinese characters, or *kanji*, meaning the "science of seeking wisdom." The first philosophy instructors were foreigners, who began to arrive three years later, and it was not until 1893 that they began to be replaced by Western-trained Japanese professors of philosophy. This development fostered the idea that that thing called "philosophy" was a solely Western product standing alongside other Western disciplines such as chemistry, physics, and biology.

Since *tetsugaku* was formed of two Chinese characters, the Chinese themselves adopted the Japanese convention and began toward the end of the nineteenth century to refer to Western philosophy using these same two characters (pronounced in Chinese, zhu-shway, or, in pinyin, zhe xue). As in Japan, most Chinese scholars initially thought that *zhe xue* was one of the Western sciences and was therefore something previously nonexistent in either China or Japan except in very rudimentary form. However, as it gradually became clearer that Western philosophy was *not* a science but a metaphysical and speculative world view based largely on a sense of cultural values (partially through the efforts of John Dewey and Bertrand Russell, who visited China just after the World War), Chinese scholars began to see greater similarities between Western philosophy and ancient Confucianism, Mohism, Daoism, and Legalism.

The final shift in definition was achieved following the great debates on this issue in China during 1922–1923, led by Liang Shuming and Chang Chunmai (Carson Chang). Chinese intellectuals now reached the consensus that much ancient Chinese writing (Confucian, Daoist, and some Buddhist texts) should be considered *zhe xue* and that *zhe xue* must be divided into Western, Indian, and Chinese, each representing different value orientations or *Weltanschauungen* of these different cultures. Since philosophy was now deemed not a science but rather the expression of cultural values, Liang and his group successfully argued that the Chinese should embrace Western science but continue to espouse Chinese philosophy.

Shortly afterward Koreans began referring to their ancient literature derived from Chinese sources as "Korean philosophy," but the Japanese disagreed, refusing to designate Japanese versions of this same Chinese literature as Japanese philosophy. It is true, as Japanese intellectuals such as Nakae Tokusuke (pen name Nakae Chōmin 1901) argued, that anything one might want to designate as Japanese philosophy was borrowed and evolved from Chinese sources. But this is no more true of Japanese borrowing from external sources than of British, German, or French philosophy borrowing from Greek sources. A transplanted tradition often becomes culturally identified with its adopted country if and when it takes deep root and permanently transforms the original product into its own image. And this seems no less true of Japanese versions of Chinese philosophy than it is of Korean renditions of Chinese philosophy or of British transformations of Greek philosophy. The important issue, therefore, is not the origins of Japanese philosophy but how Japanese philosophers interpreted, criticized, modified, developed, and used imported Chinese philosophical ideas and methods in accordance with Japanese predilections and needs, and how their writings contributed to a continuing, distinctively Japanese tradition of thought. Exactly the same criteria should be used to distinguish twentieth-century Japanese philosophy of a Western or international style from the earlier study of European philosophy in Japanese universities (in the late nineteenth and early twentieth centuries).

DISTINGUISHING JAPANESE FROM CHINESE PHILOSOPHY

One difference between Japanese and Chinese philosophy arises from the fact that Japanese philosophy is highly selective about the much larger range of philosophical schools that arose in China. This is partly due to the historical accident that by the time Chinese philosophy was imported to Japan in the seventh century, many earlier Chinese schools had already become obsolete or absorbed into other philosophical schools.

Part of the selection process, however, reflected Japanese political priorities and cultural preferences. In China philosophy had developed independently of government. In Japan, by contrast, philosophy was admitted by the government for the aid it could provide the government in the service of the state. Hence there never developed until quite late an independent class of literary specialists from which scholars could be selected for government service, as was the case in China with its famous meritocratic examination system. In Japan government positions tended to be hereditary.

For all these reasons Japanese tended to select only those aspects of Chinese philosophy best suited to the perceived needs of Japanese government leaders and advisers. So, for example, Japanese never developed (until the late Tokugawa era—eighteenth and nineteenth centuries) the idea, so prominent in China, of the role of philosophy as a tool for self-cultivation. Also, the Japanese were never very interested (again until late Tokugawa) in China's second most important and popular philosophy, philosophical Daoism (*Dao Jia*), which the Japanese government leaders thought encouraged anarchy, rebellion, and lack of loyalty to the government and devotion to the state.

For similar reasons, the Japanese tended to exclude Kongzi's (Confucius's) and, more so, Mengzi's (Mencius's) theory of the "mandate of heaven," the view that to be successful, governments must be acceptable to a moral order of heaven, without which they could be legitimately

overthrown (not a popular idea among government leaders anywhere). Where Mengzi and the Confucian tradition offered advice to governments on how they ought to rule in order to fulfill their moral obligations to their people and to heaven, this advisory function was largely excluded from Japanese Confucianism, at least until very late in the Tokugawa period. Also, Japanese Confucians emphasized loyalty to the state government over filial piety (family loyalty), whereas for the Chinese it was just the reverse.

Japanese Buddhism, in its early centuries, was similarly politically enmeshed, having been introduced into Japan by government leaders as a way to protect and bring good fortune to the state and not as a popular movement of personal faith among ordinary Japanese people. Whereas Chinese Buddhism spread among all classes of people in all parts of the country, the Japan variant was limited for several centuries to aristocratic families living in the capital. Moreover, whereas Chinese Buddhists tried and largely succeeded in staying out of government service, Japanese Buddhists were from the beginning heavily involved in the affairs of state.

The particular selection of texts the Japanese made from the Chinese and Korean traditions and the interpretations these texts received were also much affected by Japanese cultural predispositions. In the later neo-Confucianism of the Tokugawa period, for example, Japanese philosophers rejected the more abstract, transcendental, and rationalist elements of the philosophy of Zhu Xi (Shushi) in favor of material, phenomenal, sensual, immediate, intuitive principles. Japanese philosophers often explicitly criticized Chinese philosophers for being too intellectual, abstract, logical, and otherworldly.

JAPANESE BUDDHIST PHILOSOPHY

Although Confucianism and Buddhism arrived more or less simultaneously in Japan as part of a "package deal" of Chinese culture, for various reasons Buddhism played by far the greater role before Tokugawa (seventeenth century). One reason for this was the rising power of Buddhism over Confucianism in China at the time of significant contact with Japan. In the Sui dynasty (seventh century) Buddhism was at its peak in China and was strongly supported by the Sui rulers. A more practical reason inhibiting the spread of Confucianism was the enormous difficulty Japanese people had in reading Chinese. Although the characters are the same, the grammar of the two languages is completely different. Not until the Tokugawa period a thousand years later were these problems sorted out, affording Japanese greater access to Chinese sources.

Indeed, for nearly a thousand years Buddhism played much the same educational role in Japan as Confucianism had in China. Throughout most of this long period, the Buddhists ran the schools and educated most of the ruling and military elites. Ironically, it was the Zen Buddhists who introduced neo-Confucianism (Zhu Xi and Wang Yangming) to Japanese in the seventeenth century. Even in the early twenty-first century the most successful and distinctively Japanese philosophy is that of the Kyōto school, which combines Zen Buddhism with the European philosophies of Hegel and Heidegger.

In light of the Japanese traditional preference for the aesthetic surface of the world as it directly appears to us, the Japanese in general rejected any transcendent, otherworldly, metaphysical reality "behind" appearances and embraced instead the "here-now" phenomenal world sanctified and glorified as aesthetic ritual.

The philosophically most sophisticated Buddhism to emerge in China (Tien Tai, Hua Yen, and Chan Na [Japanese, Zen]) endorsed the profound and paradoxical idea that the changing, dependent phenomenal world is simply a false way of seeing the eternal, ultimate reality. This striking theory results from carrying to its logical conclusion the idea that there is nothing in the world but this one Buddha reality. There is therefore no dualism by which we might contrast the Buddha reality with the ordinary space-time physical world. What we experience as ordinary mundane existence is simply the one Buddha reality misunderstood. Because this outlook fit in very well with Japanese predispositions, Japanese Buddhist philosophers (Saichō [767–822], Kōkai [774–835], and Dōgen [1200–1253]) developed this aspect of Buddhism to the fullest.

JAPANESE CONFUCIAN PHILOSOPHY

As previously indicated, Confucianism did not have much immediate impact and was for a time hardly studied in any detail. In this early period, as can be seen in Prince Shōtoku's "Constitution" (604 CE), the main Japanese interest in Confucianism was its support for the ancient Chinese customs that Kongzi defended and systematized and that therefore became attached to his name. Many of these customs were similar to ancient Japanese practices—they were understood, that is, as a justification and theoretical support for hierarchy in society and cohesion within the family and more generally within society. The governing principle in both cultures was loyalty to the superior and loyalty to the group.

Neo-Confucianism was introduced at the beginning of the Tokugawa era through Zen Buddhists in whose monastery schools (Song dynasty [960–1279]) neo-Confucianism was studied as a sideline. What Western scholars call neo-Confucianism Chinese call *Dao Xue Jia*, the School of the Study of Dao. And this name indicates the new metaphysical and spiritual direction of Chinese Buddhism and late Daoism, beginning in the Tang dynasty (618–907) but coming to maturity in the Song dynasty. Although neo-Confucians rejected Buddhism because it was not Chinese either in origin or in tradition, they absorbed into Confucianism many elements of both Buddhism and Daoism. Neo-Confucians also selected those Confucian texts more in line with Song dynasty Buddhist-Daoist spiritualism and then interpreted those texts in the new way. Mencius is selected over Xunzi and interpreted spiritually and idealistically, emphasizing the idea in Mengzi that everything lies within us, that we share the goodness of human nature with heaven, that the direct, spontaneous feeling or intuitive thought is the best insight into reality. The key virtues during this period were not so much the social ones of propriety and benevolence but rather the self-cultivation of an inner quality of Buddhistlike mental tranquility and sincerity.

Cosmologically, the *ba gua*, or trigrams (and the sixty-four hexagrams of paired trigrams), were added to the older Daoist cosmology of *qi* in an effort to explain the evolution of the natural world from a single element into the multifaceted world we are familiar with. The original *qi* ether is said to divide into the *yin* and *yang* ethers (representing the passive and active forces in nature), which in turn evolve into the five elements (*wu xing*: earth, wood, metal, fire, and water), which finally produce the "ten thousand things." Philosophically the most important element added during this period is the notion of *li* in opposition to *qi*. *Qi* is the material stuff of the world, and *li* is the formative principle that shapes it into stable and predictable forms.

This idea probably comes from Buddhist Tien Tai and Hua Yen metaphysics (which may, in turn, have been influenced by Daoism, suggestions of which are found still earlier in the *I Jing*), where the root idea is that the inner nature of everything is the same, namely the Buddha nature. In neo-Confucianism the emphasis is more specific and somewhat more secular, each *kind* of thing being governed by its own principle, or *li*. The *li* of chickens makes their eggs hatch into chicks which then grow into chickens, and so on. But as in *Yogō cōra* Buddhism, an understanding of all the *li* lies innate within each person's mind. By quietly reflecting within our own minds, we can come to realize the inner *li* of all things.

Zhou Tunyi, Shao Yung, and Chang Cai (eleventh century), are all "fathers" of Chinese neo-Confucianism, but the tradition really begins with the Cheng brothers, Cheng Hao and Cheng Yi (eleventh century). Cheng Yi and Zhu Xi (late eleventh century) form the Cheng-Zhu *Li Xue* school (also called the Rationalist school), whereas Cheng Hao, along with Lu Jiuyuan (better known under his literary name Lu Xiangshan [twelfth century]) and Wang Yangming (fifteenth to sixteenth century) form the Lu-Wang *Xin Xue* school (also called the Idealist school).

Li Xue held that *li* exist independently of particular things and also independently of human consciousness (or minds). As Fung Yulan points out, this view is akin to Plato's theory of Forms. Xin Xue held that *li* do not exist independently of human consciousness (or particular things). So, for the Li Xue we discover *li* by examining things in the world, whereas for the Xin Xue we discover *li* by examining our own minds. Also, for the Li Xue human nature is *li*, whereas for the Xin Xue human nature is mind (human consciousness). That is, for Li Xue human consciousness is part of the *qi*, the material stuff, or body, whereas for Xin Xue it is the essential characteristic of human beings.

TOKUGAWA CONFUCIANISM

The first phase of Tokugawa Confucianism in Japan was basically a variation of the neo-Confucianism of Zhu Xi. Zhu Xi is clearly the most important neo-Confucian and the one who had the greatest influence outside of China (he is the central figure in Korean and Vietnamese Confucianism as well). Zhu Xi interprets the Supreme Ultimate (*tai qi*) as a metaprinciple, the superprinciple that governs the other principles just as they govern the formation of individual things. So Zhu Xi argues not only that every distinct kind of thing has its own principle but also that everything in the world has the *same* nature or principle (the *Tai Qi*, or Supreme Ultimate, sometimes referred to as the *Dao*), the superprinciple of principles, which governs other principles as they govern particular things in the world.

Unlike Buddhism, however, this inner nature of everything, according to Zhu Xi, is not Buddhahood but the central Confucian virtue of *ren* or human-heartedness. The difference between human beings and plants, rocks, or other animals, each of which has its own *li*, is that this nature (Supreme Ultimate, *tai qi*) is more clearly displayed, more prominent, and more accessible in human beings. Zhu Xi reasoned that in a profound sense

the Supreme Ultimate meta- *li* of *ren* (human-heartedness) was the controlling force of the world. No longer, then, is *ren* merely *one* of the human virtues; it has now become a metaphysical principle governing the entire universe.

Because neo-Confucianism was first presented to the Japanese by Buddhists within the context of Buddhism, the antagonism between Confucianism and Buddhism was not apparent at first, and the two coexisted peacefully for centuries. But the new political regime of Tokugawa shōguns, in their attempt to unite the many feudal principalities of Japan into one nation under the nominal head of the Emperor but controlled by the Shōgun, found the differences between Confucianism and Buddhism politically useful and therefore encouraged the development of a new Confucianism that was not only different from Buddhism but also antagonistic to it. Whereas Buddhism was perceived as otherworldly, spiritual, personal, and metaphysical, Confucianism came to be perceived as being this-worldly, humanistic, rational, and focused on social and political concerns. As a result, Buddhism declined as Confucianism rose, though not to such a great extent as in China. Buddhism was disparaged as superstitious, emotional, and socially useless, whereas Confucianism was praised as humanistic, rationalistic, and pragmatic.

Nonetheless, imported Chinese culture was always adjusted to Japanese sensibilities and needs, and neo-Confucianism was no exception. Almost immediately Japanese intellectuals, including Fujiwara Seika (1561–1617), Hayashi Razan (1583–1657), Gahō Razan (1618–1680), Hōkō Razan (1644–1732), Nakae Tōju (1608–1648), Yamazaki Ansai (1618–1682), Kumazawa (1619–1691), and Itō Jinsai (1627–1705) accepted that part of neo-Confucianism that suited their needs and rejected those parts they considered un-Japanese. Basically, they accepted the humanism and rejected the rationalism. The main criticism of Zhu Xi was his stress on rationality at the expense of emotion.

What is most interesting about Japanese followers of Zhu Xi (Shushi) is their complete rejection of his notion that the ultimate reality of the world is the abstract, immaterial, eternal, and unchanging *li*. Korean Confucians, by contrast, took this Platonic element in Zhu quite seriously, actively debating for centuries whether both *li* and *qi* exist (that is, whether the abstract *li* can exist independently of the material *qi*) and, if so, which of the two is primary.

Some 300 years after Zhu Xi, Wang Yangming (*ō yōmei*) rejected Zhu's "*li xue*" (the philosophy of principle) in favor of " *xin xue*" (the philosophy of mind). And this, too, had its important counterpart in Japan, especially in the work of Oshio Heihachirō (1793–1837). Wang Yangming identified the ultimate nature or essence of things with mind, adopting a position similar to Western idealism; the ultimate reality is mind and ideas entertained by mind. For Wang human nature is mind, not *li*, and the Supreme Ultimate (the overarching *Dao* of everything) is Mind (*xin*), not *li*.

Other differences follow from Zhu's privileging of *li* and Wang's preference for *xin*. Whereas for Zhu we follow the *Da Xue* (the Han dynasty [third century BCE to the third century CE] Confucian classic, *The Great Learning*) in "extending learning by investigating things," Wang contends that, following the other Han dynasty Confucian classic, *Zhong Yong* (*Doctrine of the Mean*), one can best learn the ultimate principles of reality by simply reflecting within oneself. The ultimate *Dao* is Mind, and where better to study Mind than one's own mind? The other major difference between these leading neo-Confucians is that whereas Zhu (somewhat like Aristotle) sees a gap between knowledge and action (that one can know the right thing to do and not do it [Aristotle's "weakness of will"]), Wang argues (somewhat like Socrates and Plato) that if one truly understands what is right, one will do it. Of course, part of the disagreement between Zhu and Wang on this point has to do with different notions they have of knowledge, Zhu stressing something akin to ordinary common sense knowledge, and Wang, something closer to meditative quasi-Buddhist enlightenment. While it seems clear that Zhu Xi borrows from Hua Yen Buddhism the *li-qi* (in Hua Yen *li-ji*) distinction, Wang Yangming's indebtedness to Yogōcōra Buddhism is equally clear.

A major contribution to Chinese Mahōyōna Buddhism was Yogōcōra idealism, which held that everything is the Buddha Mind, that the phenomenal world is a mentally produced illusion. Yogōcōra joins with Nōgōrjuna's *Mōdhyamika,* or "middle way" (which holds that reality is empty, not mental) to form most of the leading schools of Chinese Buddhism, especially, when further combined with Daoism, Chan (Japanese Zen). For that reason Wang is often called a closet Chan Buddhist. Most Japanese Buddhist philosophers rejected Yogōcōra idealism as too remote from common-sense realism and too alien from the peculiarly Japanese celebration of the infinite aesthetic richness of the phenomenal world of everyday sense experience. For this reason most Japanese neo-Confucians rejected Wang's idealism, though many Japanese found great sympathy for the spiritual sincerity

of Wang's emphasis on inner reflection and self-cultivation. Nonetheless, Japanese followers of Wang rejected his main idea, just as followers of Zhu rejected *his* main idea. In both cases Japanese adopted Chinese neo-Confucianism in their own, peculiarly Japanese way.

MODERN NATIONALISTIC PHILOSOPHY

In the Qing dynasty (seventeenth through twentieth centuries), Chinese Confucians rejected all the Daoist and especially Buddhist elements with which Song (Zhu Xi) and Ming (Wang Yangming) dynasty neo-Confucianism had become embedded and urged a return to the original Confucianism of the Han and pre-Han period. In reaction to this Buddhistic (and hence Indian, ergo, non-Chinese) Confucianism, the Qing dynasty Confucians led a movement "back to the original (thoroughly Chinese) Confucianism." And this too was closely followed by Japanese Confucians.

More interesting was the Japanese adaptation of this "back to the (nationalistic) origins" as a "return" to *Japanese*, not Chinese, ancient writing. Of course, there is no Japanese writing of comparable antiquity to that of China, but there were the early "histories," such as the *Kojiki*, commissioned in the seventh and eighth centuries by Japanese rulers. These accounts were mostly collections of mythological prehistories of what later became known as Japanese Shintō. Like its Chinese counterpart, this trend represents the first dawning in Japan of a kind of "intellectual nationalism" that became increasingly important all over the world in the early twentieth century, especially in the period from 1920 to 1940.

Whereas Japanese Confucians rejected Confucian rationalism in favor of humanism, their embrace of Confucian humanism was itself qualified. On the whole, it was rejected politically but accepted morally; that is, Confucian humanism was rejected, at least initially (in the seventeenth century), as part of the political philosophy supporting the new Japanese Shōgunate "*bakufu*" government, whereas it was accepted as the foundation for a more general and widespread moral code throughout the country. In its military guise, Japanese government was less paternalistic and more rigidly duty-bound. Military leaders demanded and expected absolute obedience from their citizens. Here again Chinese thought was used to support, justify, and defend Japanese traditions rather than to modify them. On the other hand, Confucianism was a very important factor in the development of Japan's early modern (seventeenth and eighteenth century) moral consciousness, especially among the rising middle class of wealthy, educated merchants in the cities.

At first Japanese Confucians sought to find this more humanistic side of Confucianism in the earlier Han and pre-Han Confucianism of the *Analects* and the *Mengzi*. But eventually Japanese Confucians turned away from Chinese sources altogether for this missing ingredient and began to look instead within their own ancient Japanese traditions. This opened the door for the focus on the more religious, nonrational ancient *Japanese* (i.e., Shinto) learning, which rejected both the humanism and the rationalism of neo-Confucianism. Japanese ancient learning portrayed the secular humanism of neo-Confucianism as a weakness, not a strength; it was bad because it ignored the ancient Japanese belief in the *kami*, a mysterious power that cannot be discovered by logical analysis or empirical investigation but only by the authority of ancient texts.

The most important of these "National Learning" philosophers was Motoori Norinaga (1730–1801). For more than thirty years Motoori struggled to have the *Kojiki* made the basis of accepted Shintō scripture. The problem was that the *Kojiki* is mainly a loose collection of ancient myths, legends, and genealogical records of the imperial family. It contains little abstract or profound philosophical thought. Motoori, nonetheless, tried to show that this was a strength and not a weakness. He argued that, like other sacred texts, the religious truths in the *Kojiki* are beyond ordinary sense perception, common sense, or reason. He also interpreted certain elements in the *Kojiki* as a purely Japanese sensibility of spontaneous sentiment privileging the emotional and aesthetic side of human nature over its more rational and moral side as favored by the Chinese.

Not surprisingly, Motoori was severely criticized by the neo-Confucian philosophers of his day for his naive and irrational theories. One objection Motoori tackles head-on is the criticism that instead of appealing to a universal human reason that could be appreciated by all people everywhere, as most philosophers try to do, Motoori *isolates* the Japanese people from everyone else in the world. According to Motoori's explanations, only the Japanese who follow the *Kojiki* know the truth and follow the true Way; only they are the chosen people. But that is just the way it is, Motoori responds. The gods favored Japan and more clearly revealed the Way of the gods to them, and the Japanese people have preserved this ancient, sacred tradition better than other people, who have abandoned what religious understanding they once

had in favor of new, man-made philosophical explanations.

Motoori tends to be theistically fatalistic: everything is decreed by the gods, whether for good or evil. How can we explain the existence of evil? Instead of looking for rational reasons to justify the fact of undeserved evil, Motoori simply says that we know from the *Kojiki* that this is what the gods decided and the way they acted. If you go on to ask *why* the gods did things in this way, you are asking a question that cannot be answered. The Way of the gods is not the Way of man.

Instead of constantly trying to control or restrain our emotions, as the rationalistic philosophers are always telling us to do, Motoori insists on a more frank acknowledgment of the power of emotion in our lives. Sometimes, it is true, emotion leads us into indiscretions which we later regret. But we cannot help ourselves. We should not be so judgmentally harsh on ourselves or on other people, Motoori urges us, but rather sympathetically recognize (with fatalistic resignation) the power of emotion to occasionally lead us astray.

Motoori forms the transition to the late-nineteenth-century Meiji rejection of Chinese in favor of Western learning as the only way to compete with the West and avoid Western domination. The central paradox of this early adjustment to the modern ways of the West is that the Japanese suffered no sense of inferiority in their need to emulate at least some aspects of the West. The general consensus among Japanese intellectuals of this period followed Motoori's conviction that the gods arose in Japan and therefore favored the Japanese people, giving them a natural and undeniable edge over all other peoples. By the early twentieth century, Japan had thoroughly mastered Western science and technology but was torn politically between liberal and more conservative Western thought—specifically between British empirical and neo-Kantian ideas on the one hand and German Hegelian and Heideggerian doctrines on the other. In the end the conservatives won the day.

Perhaps the most important thinker in this regard was Watsuji Tetsurō (1889–1960). His 1935 work *Fōdo* (*Climate and Culture*) replaces Heidegger's emphasis on time (in *Being and Time*) with a focus on space. Like other Japanese philosophers, Watsuji did not accept Heidegger's thought without reserve. Indeed, it was deemed by most Japanese to be too concerned with the individual at the expense of the social. In Watsuji's view, Heidegger's neglect of space precluded a description of human existence concrete enough to allow for a true depiction of history and (to cite Watsuji's chosen focus) of the role of climate within it. Again, we see the Japanese predilection for the empirical and phenomenal over the abstract and transcendental.

In *Rinrigaku* (*Ethics,* 1937–1949) as in *Fōdo*, Watsuji again taxes Heidegger with having scanted space relative to time and thereby the social relative to the individual. The German philosopher, he complains, "stuck fast to an atomistic individuality"*(Rinrigaku*, 224). This Cartesian, Hobbesian notion of self, Watsuji declares, is artificial and must be replaced with a more communitarian view of authenticity. Watsuji's terminological lynchpin for this idea is *ningen,* the Japanese word for "human being," where *nin* means "person" and *gen* signifies "between or together," thus implying a communal relationship. Inseparable from their cultural and social context, humans are fixed in a tensed, contradictory relationship to society; each person is at once an individual and a member of a social order and never wholly one or the other.

In place of Heidegger's "nothingness," Watsuji substitutes the Buddhist notion of emptiness (*sunyata*), an "authentic" surrender of selfish ego out of which compassion may arise. The problem with this idea, of course, is that his "authentic individual" can easily be submerged in totalitarianism because the community of which he or she is a *selfless* member is, in practical terms, inseparable from the state.

THE TWENTIETH-CENTURY KYŌTO SCHOOL

Using our criteria for what constitutes "Japanese philosophy," the first clearly Japanese philosophy of the post-Meiji period is the 1911 publication of *Zen no kenkyō* (*An Inquiry into the Good*), by Nishida Kitarō (1870–1945). Nishida was the pioneer of the *Kyōto-ha*, the Kyōto School of philosophers, which included other notable thinkers such as Tanabe Hajime and Nishitani Keiji. On the one hand, all the school's major members were condemned in some quarters for having collaborated with, or at least having endorsed, ultranationalist objectives; on the other hand, there has been widespread admiration for the quality of their purely philosophical activity, the best of which has been deemed of worldwide significance.

The originality of *Zen no kenkyō* lay in its author's attempt to express the Zen ideal of "unity of thought" within a densely argued philosophical system applying Western methods and concepts. Like his contemporary readers, Nishida faced the problem of reconciling Japan's traditional values with those implicit in the Western-inspired technological revolution. Nishida saw Zen insight as a possible solution for the crisis facing Japan.

Nishida's solution centered on the Western notion of "pure experience." Western writers had used the term "pure experience" in a way that seemed to him fundamentally flawed. What dissatisfied him in Mach, James, and others was their dualistic analysis of pure experience. Nishida held that any theoretical representation of pure experience inevitably introduces falsification. Sense-datum psychologists and philosophers had tried to describe our experience prior to perceptual syntheses and conceptual classifications but had nonetheless presupposed a subject-object dichotomy in which the perceiver is aware of himself looking at a world beyond himself. What Nishida sought to describe was a still more elementary, "pure" experience prior to any subject-object distinction—the experience of a newborn child.

From the standpoint of Zen Buddhism, in which the self is perceived as an artificial construction that inhibits the Buddha vision and is therefore best "dissolved," such a position does not surprise. But expressed philosophically, it took on the power of subversive dialogue with established Western beliefs. If self precedes experience, universal principles posited on the basis of individual experience are suspect. To avoid solipsism, Western thinkers have traditionally had to make assertions *beyond*experience. "Higher" realms and "hidden" essences have become chimerical foci in the hopeful quest of a human commonality. If, however, pure experience precedes self, then such experience itself can be declared a universal principle. The problem for modern Western philosophy—How do I get my private individual self to a reality beyond, understood in terms of universal principles accessible to everyone?—is not a problem in Nishida's notion of "pure experience," for he presupposes no division between "me" and "reality" nor between me and others.

Nishida's "logic of place (*basho*)" or "logic of nothingness" was quite unlike the "objective" logic of Western rationalism. Just as he had criticized the dualistic opposition in the Western representation of pure experience, so he calls here for a regress from the standpoint of reason (ensconced in the constructed subject) to the very starting point of our awareness, prior to the construction of self with its constructions of categories of determinate being. Only after pure experience has been differentiated into self and world—and after world, in turn, has been classified into categories of conceptual thought—can reason and cognition begin their work. Nishida developed a logic prior to the confrontation of a knower confronting an object.

CONCLUSION

Whether there is a Japanese philosophy or not depends on how one defines the word *philosophy* and how one judges the efforts to indigenize, or nationalize, borrowed, alien philosophical sources. Since the early 1920s there has been a general if not unanimous consensus in the philosophical community that there are three independent (literate) philosophical traditions: Greek, Indian, and Chinese (of which Japanese philosophy is an offshoot). Throughout Japan's 1,300-year history, that nation's philosophers have borrowed freely from outside sources—first from Chinese and later from Western philosophical traditions. But the Japanese have always interpreted and used these foreign sources in their own distinctive way. What is perhaps most peculiar about Japanese philosophy is that whereas nearly every other literate (and nonliterate) non-Western tradition is eager to claim for some of its thought the honorific title of "philosophy," the Japanese have mostly been reluctant to do so—or indeed to identify themselves with other cultures in any way that might appear to detract from their cultural uniqueness.

See also Aristotle; Buddhism; Cheng Hao; Cheng Yi; Chinese Philosophy; Confucius; Hayashi Razan; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Itō Jinsai; James, William; Korean Philosophy; Kumazawa Banzan; Lu Xiangshan; Mach, Ernst; Mencius; Nāgārjuna; Nakae Tōju; Nishi Amane; Nishida, Kitarō; Plato; Self; Socrates; Wang Yang-ming; Watsuji Tetsurō; Yamazaki Ansai; Zhu Xi (Chu Hsi).

Bibliography

Arima, Tetsuo. *The Failure of Freedom: A Portrait of Modern Japanese Intellectuals*. Cambridge, MA: Harvard University Press, 1969.

Barshay, Andrew E. *State and Intellectual in Imperial Japan: The Public Man in Crisis*. Berkeley, CA: University of California Press, 1988.

Blocker, H. Gen, and Christopher Starling. *Japanese Philosophy*. Albany, NY: State University of New York Press, 2000.

Kashiwahara, Yōsen, and Kōyō Sonoda, eds. *Shapers of Japanese Buddhism* Translated by Gaynor Sekimori. Tokyo: Kosei, 1994.

Maruyama, Masao. *Studies in the Intellectual History of Tokugawa Japan*. Translated by Mikiso Hane. Tokyo: University of Tokyo Press, 1974.

Matsunaga, Alicia, and Daigan Matsunaga. *Foundation of Japanese Buddhism*, 2 vols. Los Angeles and Tokyo: Buddhist Books International, 1976.

Miyoshi, Masao, and H. D. Harootunian, eds. *Postmodernism and Japan*. Durham, NC: Duke University Press, 1989.

Nakamura, Hajime. *Ways of Thinking of Eastern Peoples,* edited by Philip Wiener. Honolulu: University Press of Hawaii, 1964.
Nishida, Kitarō. *An Inquiry into the Good* (*Zen no kenkyō*. Translated by Abe Masao and Christopher Ives. New Haven, CT: Yale University Press, 1990.
Nishida, Kitarō. *Intelligibility and the Philosophy of Nothingness*. Translated by Robert Schinzinger. Westport, CT: Greenwood Press, 1973.
Nishida, Kitarō. *Intuition and Reflection in Self-Consciousness*. Translated by Valdo Viglielmo. Albany, NY: State University of New York Press, 1987.
Ōe, Kenzaburō. "Japan's Dual Identity." In *Postmodernism and Japan*, edited by Miyoshi and Harootunian. Durham, NC: Duke University Press, 1989.
Parkes, Graham, ed. *Heidegger and Asian Thought*. Honolulu: University of Hawaii Press, 1987.
Piovesana, Gino K. "Contemporary Japanese Philosophy." In *Asian Philosophy Today*, edited by Dale Riepe. New York: Gordon and Breach, 1981.
Reischauer, Edwin O., and Albert M. Craig. *Japan: Tradition and Transformation*. New York: Houghton Mifflin Co., 1973.
Tanabe, Hajime. *Philosophy as Metanoetics* (*Zangedō to shite no tetsugaku*). Translated by Takeuchi Yoshinori. Berkeley, CA: University of California Press, 1986.
Tsunoda, Ryōsaku, William, Theodore de Bary, and Donald Keene, eds. *Sources of Japanese Tradition.* 2 vols. New York: Columbia University Press, 1958.
Watsuji, Tetsurō. *Climate and Culture: A Philosophical Study*. Translated by Geoffrey Bownas. Westport, CT: Greenwood Press, 1988.
Watsuji, Tetsurō. *Rinrigaku*. Translated by Yamamoto Seisaku and Robert E. Carter. Albany, NY: State University of New York Press, 1996.

H. Gene Blocker (2005)

JASPERS, KARL
(1883–1969)

Karl Jaspers was one of the architects of contemporary existentialism and one of the first philosophers to use the term *existentialist*. He was a prolific writer with a prolix style that is often inelegant, superficial, sentimental, and unclear and that over the years showed itself to be repetitious. Yet careful and extensive reading of his works shows him to be a rigorous and responsible thinker. Appearances notwithstanding, he was perhaps the most systematic of all existentialist philosophers. His philosophy is neither linguistic analysis nor metaphysics. It can be best characterized as a disciplined and organized description of the critical fringes of human existence, such as impenetrable limits, unmitigated freedom, and the experienced indefinite expanse of space, time, and consciousness. Jaspers fulfilled the commonsense image of the philosopher through his vital concern with the contemporary political situation and his trenchant reflections on the threats to man's integrity and fulfillment posed by twentieth-century social, economic, and political institutions. He spoke with authority to the nonphilosophic mind because of his deep and successful roots in medicine and psychology. He was suspicious of contemporary overconfidence in science and, as an antidote, stresses the irrational in man. As Jaspers saw it, philosophy begins where reason has suffered shipwreck. Philosophy is an activity, a becoming, not a state of being or a body of facts. Philosophy is philosophizing. To appreciate philosophic insights we must—as Socrates and Sigmund Freud saw—arrive at them ourselves. We must live philosophy, since we cannot meaningfully paraphrase its conclusions. Genuine philosophy arises directly out of the problems confronting the individual philosopher in his existential, or historical, situation. General problems are mere derivatives. Philosophy need not be metaphysics; it can only illuminate some of the potentialities of an individual existence, an existence that is ineffable, unique, and free.

Jaspers was influenced especially by Immanuel Kant, but also by Søren Kierkegaard and Friedrich Nietzsche, whom he admired because they were prophets who articulated the structure of their existence, because they were not academic philosophers, because their thinking welled up directly from their personal existence, and because they illustrated the axiom that philosophic thinking begins in the attempt to communicate to another the nature of one's *Existenz*. The influence of Edmund Husserl is also apparent, although it is perhaps unconscious, since it is mostly unacknowledged. Jaspers used Husserl's method of descriptive phenomenology and adopted Husserl's concept of intentionality as a central function of the self. Furthermore, Husserl's ideas of the transcendental ego and transcendental consciousness conform to Jaspers's descriptions of the inner self (*Existenz*) and the outermost boundaries of the world (*das Umgreifende*). Jaspers's religious thought, although it ignored Aristotelianism and Scholasticism, was deeply influenced by Plotinus, Giordano Bruno, Benedict de Spinoza, and Friedrich von Schelling and gives a modern phenomenological restatement of many of the classical religious intuitions of humankind.

LIFE AND WORKS

Jaspers was born in 1883 in the East Frisian city of Oldenburg. His father was a banker, constable, and jurist. Jaspers studied law at the universities of Heidelberg and Munich, and medicine at Berlin, Göttingen, and Heidelberg. He received his MD from Heidelberg in 1909, upon

completion of his dissertation on *Heimweh und Verbrechen* (Nostalgia and crime). Immediately upon graduation he became a volunteer assistant in psychiatry at Heidelberg. His first major work, *Allgemeine Psychopathologie* (*General Psychopathology*, 1913), is a book on methodology showing the merits and limits of various psychological procedures and descriptions. In 1916 he became professor of psychology at Heidelberg. Shortly after World War I he published his *Psychologie der Weltanschauungen* (Psychology of world views; 1919), which consists of descriptions of many different attitudes toward life. It is based on Wilhelm Dilthey's *Typologie der Weltanschauungen* and marks Jaspers's transition from psychology to philosophy. He later called it the first genuinely existentialist work. Both of these early works were based on his medical experience.

He received a professorship in philosophy at Heidelberg in 1921, after declining similar offers from the universities of Kiel and Greifswald. In 1932 he published his magnum opus, the three-volume *Philosophie*, which is a detailed development of the notions of transcendence and *Existenz*. In 1937 he was relieved of his duties by the National Socialist regime, but was reinstated in 1945. In 1946 he was named honorary senator of Heidelberg University, and from 1948 on he taught at the University of Basel in Switzerland. In 1958 he was awarded the German Peace Prize at the Frankfurt Book Fair. The first volume of his *Philosophische Logik* appeared in 1947. Throughout his life, Jaspers was greatly concerned with communication. Personal relationships had great philosophic significance to him. In addition to his parents, particularly significant persons in his life were his teacher Max Weber, his friend Ernst Meyer, and Meyer's sister Gertrud, who became Jaspers's wife. Since she was Jewish, Jaspers lived, through her, the agony of the Jewish people during World War II, and this led him to publish in 1946 his reflections on the question of German guilt, *Die Schuldfrage, ein Beitrag zur deutschen Frage*.

Any classification of Jaspers's views into traditional philosophic disciplines is artificial. For purposes of exposition, however, such an expedient is necessary.

EPISTEMOLOGY

Jaspers's method is generally skeptical. It consists of the exploration, description, and analysis of first-person experiences. These form the basic data for philosophical generalizations and are for any person the sole source of his information about reality. Jaspers goes far beyond René Descartes in emphasizing the epistemological primacy of subjectivity: My thinking begins and ends with subjectivity, since awareness, as Kant saw, always consists partially of interpretations. Although the results of these descriptions do not form a universal ontology—they apply, strictly speaking, to my own self exclusively—they are nonetheless verifiable inasmuch as egos may compare experiences. Jaspers follows Kierkegaard in describing immediate experiences (which consist not only of sense data but also of love and anxiety, hope and despair) and examining their ontological import. Since he describes fringe states of consciousness, areas of experience that are difficult, perhaps even impossible, to focus sharply, his language necessarily becomes ambiguous.

There is no certainty either in philosophy or in science. I am forced to depend ultimately on the intuitions and decisions of my own ego. Science is not an ultimate form of knowledge because it excludes the observer, because it is replete with unexamined and often erroneous assumptions, and because one method of inquiry is insufficient for a complete world picture. Although the spirit of scientific inquiry is an antidote to dogma in religion, politics, and philosophy, it gives us only surface knowledge, which is, at best, a workable mythology.

PSYCHOLOGY

The nature of the self is discovered through illumination of existence (*Existenzerhellung*), which discloses the possibilities of man, that is, the possibilities of an entity seeking understanding of self and of being. *Existenzerhellung* yields access to the questioner himself. Ordinary modes of perception and cognition, which imply a subject apprehending an object, always bypass the real self (the *Ursprung*). The real and valuable, that is, the authentic, in man is called *Existenz. Existenz*, the genuine self, is nonobjective and unique. It is infinitely open to new possibilities and inaccessible to traditional philosophical investigations. Although *Existenz* is that crucial aspect of human existence that cannot be conceptually delimited, it is nonetheless clearly experienced: It can be lived; it is illuminated through philosophical reflection; it can be communicated. *Existenz* is the experience of the total freedom that defines man; it is the experience of the infinity of possibilities for styles of life; it is, finally, the experience of loneliness that cries in the wilderness. *Existenz* is the eternal in man, while *Dasein* (not to be confused with that which Martin Heidegger designates by the same word) is his temporal dimension. *Dasein* is that aspect of man that has describable characteristics and is accessible to theoretical reflection. To confuse mere *Dasein* with the authentic ground of my being, *Existenz*, is crass materialism and leads to shipwreck, while to

ignore *Dasein* altogether leads to nihilism. A tension (*Spannung*) between the two is the golden mean.

Man is alienated from his world. He comes from a dim past and goes into an indefinite future. Life is a flux in which he seeks anchor. Existence is rich in mysterious paradoxes and antinomies, such as those of freedom coexisting with dependence, communication with solitude, good with evil, truth with falsehood, happiness with grief, life with death, and progress with destruction. Authentic *Existenz* is disclosed through reason (*Vernunft*), while intellect (*Verstand*) concerns itself with the pragmatic management of existence. *Verstand* is satisfied with practical results, while *Vernunft* engages in endless searching. Man is both *Vernunft* and *Existenz.*

Existenz is limited by impenetrable boundaries (*Grenzsituationen*). To experience these and to exist are one and the same, since despair can be, in the last analysis, a cognitive and elevating emotion. A defining characteristic of man is his finitude, which he experiences as the limits to his existence. Jaspers's analysis of these boundary situations is the existential formulation of the problem of evil and has been most influential. Authentic existence will push back these limits as far as possible and then accept and bear them. Death is one of the most dramatic of these barriers. It is the source of anxiety, but it also elevates the spirit because it emphasizes the urgency of living authentically without postponement. Consciousness of the inevitable presence of death gives man courage and integrity: It gives him an authentic perspective on the things that matter most. Guilt is another important boundary situation. Man not only *feels* guilty but, because of his total freedom, is guilty. He always could have chosen otherwise. Ultimate guilt cannot be removed: It must be accepted and can thereby become constructive. Our guilt demonstrates the power that our freedom has over our destiny. The boundary of "situationality" is the fact that we are partially thrust and partially choose ourselves into a particular human condition. We can be inauthentic and inevitably fall into these situations or be authentic and make them happen. Other important boundaries are chance, suffering, and conflict.

Freedom is central to man; it leads to the overriding importance of choice, which becomes the problem of moral responsibility.

ETHICS

For Jaspers, ethics is the exploration of the experience and the potential of free will. Freedom is identified with choice, awareness, and selfhood. To choose means to be free, and man's freedom is his being. I am only to the extent that I choose freely. To be is to be conscious that one is free. I do not choose life's meanings; I do not "define" man, as Jean-Paul Sartre contends, since I am limited by my historicity—my past choices bind me. But within these confines my freedom is total. Freedom is experienced as both spontaneity and action; it is thus more important to act and be an *homme engagé* than to observe and be a theoretician. To know and use my freedom is the raison d'être of *Existenzerhellung.* Whenever I choose, I act, I am conscious of my action, I am aware of the values involved, I take chances (since the consequences of my choice are often uncertain), and I realize that commitment to some values is unavoidable.

The presence of anguish adumbrates the sacred nature of my freedom. Since each choice carries with it the accumulated weight of previous decisions, the first choice overshadows all others. Consequently, guilt is the inevitable concomitant of my freedom. My original choice (*Urentschluss*) bears down on my subsequent existence and assumes the role of original sin. I am accountable for that first choice, so that to be responsible means to have accepted that guilt. In addition, I am ceaselessly confronted with the choice between sacrificing my integrity for the sake of a longer life or surrendering myself to my authentic existential possibilities. The inherent difficulty of these choices leads to further guilt, which I may alleviate by imagining absolute standards and then approximating them. But in my heart I know there are no fixed standards and that absolutism is therefore a rationalization: the boundary of guilt is indeed impenetrable.

Anguish also appears when I realize I may lose the promise of my possibilities. But that same anguish gives me the urgency and courage to choose with my full being to implement the authentic potential of my *Existenz.* I reach this pedagogically expedient brink caused by anguish when 1 recognize the limits of scientific thought or when I am faced with critical decisions. Confronted with the abyss, I may accept a philosophic or religious orientation, I may act as if I did not recognize the existence of the abyss before which I stand, or I may adopt the nihilistic position that judges these problems to be meaningless.

Subjectivity is essentially intersubjective. I am only to the extent that another *Existenz* reflects me. Jaspers describes true communication as the feeling that men have known each other since eternity. My own freedom is in essence the search for the "loving strife" of communication with another *Existenz.* In fact, the search for *Existenz* cannot be accomplished in the abyss of absolute estrangement. Existential philosophy is self-disclosure

through communication, even being itself, although it can be represented only in ciphers as symbols, is made transparent solely through authentic communication (*Existenzursprung*). Existential communication is neither friendship nor psychotherapy; it is not fusion, esteem, or unanimity; it is, strictly speaking, as with *Existenz* itself, ineffable.

But in the end, human existence is a failure. There is no escape from man's limits (the limit of death in particular), yet man is condemned to endless striving. In this dreadful paradox between finite existence and striving for infinity, man finds the ultimate symbol of his salvation, which is transcendence.

METAPHYSICS AND THEOLOGY

Jaspers maintained that just as ethical considerations grow out of philosophical psychology, so religious answers emerge from metaphysical descriptions of being.

He follows Kant in criticizing the usual arguments for the existence of God. He rejects theism, pantheism, revealed religion, and atheism alike. All these are but symbols (ciphers), and we are in danger of taking them literally. Phenomenological descriptions of the fringes of inward and outer experiences give us the only accurate understanding of the intuitions that metaphysics and theology have traditionally attempted to articulate.

When man reflects on his freedom, he experiences it as a gift; he dimly knows that he does not stand alone. That gift, in turn, points vaguely to an ultimate horizon as its source and foundation. Awareness of transcendence also originates in the consciousness of our finitude: Through our boundaries we recognize the infinite possibilities within us. In general, the world itself points to a region beyond. Transcendence is thus experienced as the intimation of a power by virtue of which man himself exists. Confronted with these clues, man is free to pursue or to ignore them.

Jaspers uses the term *encompassing* (*das Umgreifende*) to designate the ultimate and indefinite limits of being as we experience it in all its fullness and richness, limits that surround, envelop, and suffuse all there is. It is the ultimate experienceable horizon. He uses the expression "being-as-such" to mean the encompassing or the totality of being as it is thought, conceived, or conceptualized, while he reserves the term *transcendence* to mean man's personal, devoted, and committed effort to reach the encompassing. In other words, the encompassing manifests itself in at least three modes: the total encompassing of the world, the encompassing that is the empirical world of ordinary and scientific experience, and the encompassing that is one's own self. Although we are at a loss to describe its essence, we can say of the encompassing that it is. In a sense, I and the world are identical with the encompassing. In it, the severance between subject and object disappears, since both are manifestations of the same encompassing. On similar grounds, the encompassing (and this then applies to Jaspers's reinterpretation of God) can never be viewed as one object among many. It is all of being as well as all the differentiations within being. It is likewise beyond idealism, materialism, positivism, and naturalism, since all metaphysical positions are events within the encompassing but do not in any way delimit it. Therefore, in Jaspers's view, God, the unthinkable (*das Undenkbare*), becomes Rudolf Otto's "wholly other." I cannot grasp conceptually the encompassing that I am; similarly, the world is not exactly an illusion, since it is the only language through which the encompassing can reach me. The ultimate encompassing envelops both the I-pole and the object-pole of experience.

Man can search for transcendence by various means. He can explore the world, as science does. In that way he achieves a worldview. Or he can search for it by examining the relation between himself and the world, as we find it in epistemology, ethics, and psychology. He thereby achieves illumination of *Existenz*. Finally, he can search for God, in which case he deals directly with the problem of penetrating being itself. But he must never succumb to the error of identifying the encompassing with a particular substance or substratum of the world.

The encompassing manifests itself through the "footsteps of God," through analogical predication, through symbols, or, in Jaspers's own words, through ciphers (*Chiffren*), a notion borrowed from Pascal. The encompassing is like the horizon that is the perennial goal of the sailor: It always shows itself and yet is forever inaccessible. The major purpose of metaphysics is the disclosure of the ciphers that manifest encompassing, but in the end, metaphysical elucidation of the ciphers is a highly personal undertaking. Ciphers may appear suddenly and spontaneously in the presence of empirical facts, for example, an overwhelming mountain. They may appear in art forms, in religious myths and dogma, and in theological disputations; they may become manifest in the symbolism of the history of philosophy and its metaphysical systems; and finally, they may appear through reflection on the mystery of being as well as on the death that awaits every man.

Jaspers's religious prescription is called philosophic faith (*philosophische Glaube*). It consists of the convictions that man is open to transcendence and consequently wills infinity; that there is in fact a transcendence to the ordinary world; that personal freedom is to be maintained and respected; that man, as he finds himself, is inadequate; that man can rely on help from transcendence; and that the world is grounded and supported. To reject faith means to hold that the immediate world is all there is, that man's destiny is fully determined, that man is perfectible and alone, and that the world is self-supporting. Although there are significant similarities between Jaspers's philosophical faith and that of traditional Christianity, he rigidly opposed the absolutism of the latter to the openness and toleration of his philosophic faith. The Bible, for example, is a highly suggestive instrument for his philosophic faith, especially through its ciphers of one God and its emphasis on love, on choosing between good and evil, on the eternal in man, on the ordered and yet contingent universe, and on the image of God as the refuge. Nevertheless, transcendence is discovered through doubt, not reassurance: There can be no rational justification for the final leap of faith, even for a philosopher.

See also Bruno, Giordano; Descartes, René; Dilthey, Wilhelm; Doubt; Existentialism; Heidegger, Martin; Husserl, Edmund; Intentionality; Kant, Immanuel; Kierkegaard, Søren Aabye; Nietzsche, Friedrich; Phenomenology; Plotinus; Psychology; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de; Subjectivity; Weber, Max.

Bibliography

WORKS BY JASPERS

Allgemeine Psychopathologie. Berlin: Springer, 1913. Translated from the 7th German ed. by J. Hoenig and Marian W. Hamilton as *General Psychopathology*. Chicago: University of Chicago Press, 1963.

Psychologie der Weltanschauungen. Berlin: Springer, 1919.

Die geistige Situation der Zeit. Berlin: de Gruyter, 1932. Translated by Eden Paul and Cedar Paul as *Man in the Modern Age*. London: Routledge, 1933.

Max Weber. Oldenburg: Stalling, 1932.

Philosophie, 3 vols. Berlin: Springer, 1932.

Vernunft und Existenz. Groningen: Wolters, 1935. Translated by William Earle as *Reason and Existenz*. New York: Noonday Press, 1955.

Nietzsche. Berlin: de Gruyter, 1936. Translated by C. F. Wallraff and F. J. Schmitz. Tucson: University of Arizona Press, 1965.

Descartes und die Philosophie. Berlin: de Gruyter, 1937.

Existenzphilosophie. Berlin: de Gruyter, 1938.

Die Idee der Universität. Berlin: Springer, 1946. Translated by H. A. T. Reiche as *The Idea of the University*. Boston: Beacon Press, 1959.

Nietzsche und das Christentum. Hameln: F. Seifert, 1946. Translated by E. B. Ashton as *Nietzsche and Christianity*. Chicago: Regnery, 1961.

Die Schuldfrage, ein Beitrag zur deutschen Frage. Zürich: Artemis, 1946. Translated by E. B. Ashton as *The Question of German Guilt*. New York: Dial Press, 1947.

Philosophische Logik, Vol. I: *Von der Wahrheit*. Munich: Piper, 1947. Partially translated by Jean T. Wilde, William Kluback, and William Kimmel as *Truth and Symbol*. New York: Twayne, 1959.

Der philosophische Glaube. Zürich: Artemis, 1948. Translated by Ralph Manheim as *The Perennial Scope of Philosophy*. New York: Philosophical Library, 1949.

Vom Ursprung und Ziel der Geschichte. Zürich: Artemis, 1949. Translated by Michael Bullock as *The Origin and Goal of History*. New Haven, CT: Yale University Press, 1953.

Vernunft und Widervernunft in unserer Zeit. Munich: Piper, 1950. Translated by Stanley Fine-Godman as *Reason and Anti-reason in Our Time*, New Haven, CT: Yale University Press, 1952.

Einführung in die Philosophie. Zürich: Artemis, 1950. Translated by Ralph Manheim as *The Way to Wisdom*. New Haven, CT: Yale University Press, 1951.

Rechenschaft und Ausblick. Munich: Piper, 1951.

Über das Tragische. Munich: Piper, 1952. Translated by H. A. T. Reiche and others as *Tragedy Is Not Enough*. Boston: Beacon Press, 1952.

Die Frage der Entmythologisierung. Written with Rudolf Bultmann. Munich: Piper, 1954. Translated as *Myth and Christianity*. New York: Noonday Press, 1958.

Wesen und Kritik der Psychotherapie. Munich: Piper, 1955.

Über Bedingungen und Möglichkeiten eines neuen Humanismus. Munich, 1956.

Die Atombombe und die Zukunft des Menschen. Munich: Piper, 1957. Translated by E. B. Ashton as *The Future of Mankind*. Chicago: University of Chicago Press, 1961.

Die grossen Philosophen. Munich: Piper, 1957. Translated by Ralph Manheim as *The Great Philosophers*. New York: Harcourt Brace, 1962.

Plato, Augustin, Kant. Munich, 1957.

Philosophie und Welt. Munich: Piper, 1958.

Freiheit und Wiedervereinigung. Munich: Piper, 1960.

Wahrheit und Wissenschaft. Munich, 1960.

WORKS ON JASPERS

Bollnow, O. F., *Existenzphilosophie und Pädagogik*. Stuttgart: Kohlhammer, 1958.

Ehrlich, L., and R. Wisser, eds. *Karl Jaspers Today: Philosophy at the Threshhold of the Future*. Lanham, MD: University Press of America, 1988.

Olson, A. M. *Transcendence and Hermeneutics: An Interpretation of Karl Jaspers*. The Hague: Nijhoff, 1979.

Samay, S. *Reason Revisited: The Philosophy of Karl Jaspers*. Notre Dame, IN: Notre Dame University Press, 1971.

Saner, H. *Karl Jaspers in Selbstzeugnissen und Bilddokumenten*. Reinbeck: Rowohlt, 1970.

Schilpp, P. A., ed. *The Philosophy of Karl Jaspers*. New York: Tudor, 1957.

Schrag, O. O. *Existence, Existenz, and Transcendence*. Pittsburgh, PA: Duquesne University Press, 1971.

Tilliette, X. *Karl Jaspers: Théorie de la vérité, Métaphysique des chiffres. Foi philosophique*. Paris: Aubier, 1959.

Walraff, C. *Karl Jaspers: An Introduction to His Philosophy*. Princeton, NJ: Princeton University Press, 1970.

Young-Bruehl, E. *Freedom and Karl Jasper's Philosophy*. New Haven, CT: Yale University Press, 1981).

Peter Koestenbaum (1967)
Bibliography updated by Thomas Nenon (2005)

JEANS, JAMES HOPWOOD
(1877–1946)

James Hopwood Jeans, an English physicist and astronomer was educated at Merchant Taylor's School and Trinity College, Cambridge, where he received high honors in mathematics in 1898. He taught mathematics at Cambridge as university lecturer from 1904 to 1905, at Princeton as professor of applied mathematics from 1905 to 1909, and again at Cambridge as Stokes lecturer from 1909 to 1912. In 1912 he resigned all regular offices to live on a private income and later also on the sale of several popular books. He was honorary secretary of the Royal Society, president of the British Association for the Advancement of Science, and professor of astronomy at the Royal Institution.

Jeans was a man of undoubted ability and originality and early won a deservedly high reputation, being elected a fellow of the Royal Society at the age of twenty-eight. His main contributions to science were in two fields: the kinetic theory of gases, in particular the equipartition of energy and radiation; and cosmogony, in particular the forms of equilibrium of rotating gravitational masses and the kinetic theories of aggregates of stars. The last constitute perhaps his best and most enduring work.

During the early 1930s Jeans wrote a number of highly successful books popularizing science, and these, together with *Physics and Philosophy* (1942), contain his philosophical writings. His popular expositions of scientific theories are marked by their simplicity of expression and by the striking and illuminating examples and analogies they contain.

Although Jeans contributed nothing substantial to philosophy, his views gained attention because of his eminence in the scientific field and because of their being presented together with expositions of abstruse scientific theories widely agreed to be of philosophical interest. Jeans's writings on philosophy were slight in quantity as well as in quality; even *Physics and Philosophy* contained only about fifty pages of his own views.

His position was never consistently developed and is therefore unclear. Indeed, he seems almost to have felt that it would be against the spirit of philosophy to argue with rigor, clarity, and decent caution. His work is certainly characterized by loose reasoning, and not infrequently by plainly false or confused premises. Broadly, however, his views were that science must connect observables with observables by means of chains of mathematical equations. He held that mathematical formalization is the prime part of physical knowledge and that interpretative models of this formalism are outdated and confusing crutches in coming to know about the world. This was not because Jeans believed that only propositions about observables have a meaning. He was no positivist, despite his claim to be one. On the ground that physical measurement reveals only relations between instruments (including one's eyes and ears) and reality, he believed in a Lockean substratum that is forever hidden from us. He also held that modern science suggests that there is some room for the operation of free will, but it is unclear why he adopted this opinion. His attitude to the common fallacy that the uncertainty relations of quantum physics establish the possibility of free will is quite ambiguous.

The most striking and most widely discussed of Jeans's conclusions is that reality, the Lockean substratum, is mental, not material. This conclusion reaches its most startling form in the final chapter of *The Mysterious Universe,* where Jeans argued that the universe consists of the thoughts of a Pure Mathematician, God.

Jeans asserted—it is hardly an argument—that the universe is shown to be rational by the very fact that a mathematical description of it is possible. He argued that as physics has progressed it has discarded models as an aid to explanation and discovery. Post-Galilean physics discarded the biological model of Aristotle, and modern physics has now discarded mechanical theories and models, being content to present its theories as pieces of mathematical formalism. Jeans put the matter this way: We cannot interpret the multidimensional configuration spaces of quantum physics as material space because material space has but three dimensions. Nor can we interpret the axioms of non-Euclidean geometry, especially the geometry of finite spaces, in terms of the congruences of material rods in material space. (This last claim is simply unwarranted.) Consequently, argued Jeans, the formalism of modern physical theory must be given a pure mathematical interpretation. (However,

there is no sense in which we can speak of a pure mathematical *interpretation,* since "pure" here means "*un*interpreted.") Since the subject of pure mathematics is just thoughts, we may conclude, according to Jeans, that the stuff of the universe is mental. It is thought in the mind of God, the Pure Mathematician.

See also Aristotle; Cosmology; Geometry.

Bibliography

WORKS BY JEANS

Eos; or Wider Aspects of Cosmogony. London: Kegan Paul, 1929.

The Universe around Us. Cambridge, U.K.: Cambridge University Press, 1929.

The Mysterious Universe. Cambridge, U.K.: Cambridge University Press, 1930.

The Stars in Their Courses. Cambridge, U.K.: Cambridge University Press, 1931.

New Background of Science. Cambridge, U.K.: Cambridge University Press, 1933.

Through Space and Time. Cambridge, U.K.: Cambridge University Press, 1934.

An Introduction to the Kinetic Theory of Gases. Cambridge, U.K.: Cambridge University Press, 1940.

Physics and Philosophy. Cambridge, U.K.: Cambridge University Press, 1942.

WORKS ON JEANS

Milne, E. A. *Sir James Jeans: A Biography.* Cambridge, U.K.: Cambridge University Press, 1952.

Stebbing, Susan. *Philosophy and the Physicists.* London: Methuen, 1937. Ably criticizes Jeans's views.

G. C. Nerlich (1967)

JEFFERSON, THOMAS

(1743–1826)

"Here was Buried Thomas Jefferson
April 2, 1743 O.S. –July 4, 1826
Author of the Declaration of Independence
And of the Virginia Statue for Religious Freedom
and Father of the University of Virginia."

These are the words that Thomas Jefferson wrote for his epitaph. They indicate what he thought were his life achievements. What is notable here is that he does not mention that he was secretary of state, vice president, and president. These political accomplishments were not at the top of his list.

Jefferson was born in Shadwell, Virginia. His father, Peter Jefferson, was a self-taught surveyor and mapmaker. Thomas was sent to William and Mary College in the colony's capital, Williamsburg. Jefferson took to the law under the tutelage of George Wythe. It should be noted that Thomas Jefferson had wide-ranging intellectual interests (as evidenced by the personal libraries he assembled that included natural philosophy, history, and the fine arts). Jefferson was also a man of action and in 1769 was elected to the House of Burgesses. He became active in politics and published "A Summary View of the Rights of British America." This tract took a Whig-oriented republican view. In 1776 he wrote the Declaration of Independence from a Lockean standpoint. And in 1787 he completed his only full-length book, *Notes on the State of Virginia* (part encyclopedia and reflections on the same). Most of his other writings consist of speeches, legislation, and letters. He was the architect for his famous house, Monticello, and he founded the University of Virginia.

HIS PHILOSOPHY

Most commentators cite the influence of John Locke on Jefferson. Locke's *Second Treatise on Government* depicted the strong individual within a state of nature. This individual possessed natural rights that came into play in establishing the social contract. Government was created by the people and could be dissolved if it did not serve popular purposes. Locke's approach is so greatly in evidence within the Declaration of Independence that Carl Becker has said that "Jefferson copied Locke" (1945, p. 79).

This is probably most true of the Declaration but is less true of Jefferson's other works, which show a broad influence from the liberal Enlightenment. The argument for this can be illustrated by Jefferson's repudiation of the Church of England (which set out a default religious position modified by toleration of other religions—a position accepted by Locke). Jefferson, however, insisted upon the absolute separation of church and state. The result of this is the Virginia Statute for Religious Freedom, which Jefferson wrote.

J. G. A. Pocock (1969) has argued that Jefferson's thought reflects the thinking of nonliberal republican thinkers such as Cicero, Machiavelli, and James Harrington. This argument follows from humankind's nature as political beings and the political heritage from ancient Rome onwards. The argument for this interpretation lies largely in discursive passages of the *Notes* and in his correspondence.

The empiricism of Francis Bacon is also present in Jefferson's work on agriculture. In the end, it seems safe to conclude that though Jefferson was greatly influenced by

John Locke, there are many philosophical lights that guided him. What Jefferson did was to assimilate these various influences and apply them to practical problems that confronted him in his role as a prominent man of action.

A MAN OF CONTRADICTIONS

In the end, any evaluation of Thomas Jefferson must come to terms with his many contradictions. On the one hand, he was an agrarian, individualist, advocate for limited government, and yet on the other hand he served in three national offices (including the presidency) and expanded the country greatly—particularly through the Louisiana Purchase. He also stated in the Declaration that "all men are created equal." Slavery certainly flies in the face of equality. At times in both "The Rights of British America" and in his correspondence, Jefferson calls for various versions of ending slavery. And yet Jefferson continued to own slaves himself. This can probably be explained by the fact that Jefferson did not completely believe that nonwhite individuals were fully human. For example, Jefferson says in the *Notes* (query 14) that "Never yet could I find that a black had uttered a thought above the level of plain narration …" If blacks and Native Americans were not fully human, then they have no place in the new Republic. They must be exiled so that the "pure" fully human European Americans might appropriate the wilderness—viewed as the state of nature. In the Lockean state of nature, if one could work the land and make productive use of it, then it was his. Because the native peoples were not fully human, the fact that they were using the land first would be irrelevant.

However, once we set out the above position, we are again faced with a contradiction. In November 1998 the magazine *Nature* published an article that strongly suggested that Thomas Jefferson was the father of his slave Sally Hemmings's last son, Madison Hemmings (1805–1877). It was also possible that Jefferson and Hemmings had five other children: four daughters and one other son. Though the evidence for this is not conclusive (because the DNA tests could have also been the same if another Jefferson relative were the father), still these results raise questions of the relationship between the races in Colonial times. Why would Thomas Jefferson have children with someone he believed to be subhuman? Two possibilities present themselves: (a) either Jefferson thought that interracial sexuality was merely a way to satisfy desire without thought of outcome; or (b) Jefferson's private actions did not match his public writings. The first alternative makes Jefferson into the sort of animal brute he publicly eschewed. The second alternative humanizes Jefferson and shows that he might fall in love with and honor a woman of color. Under this second hypothesis, he might personally believe in a realm of equality that he could never publicly express (even though it matched his words in the Declaration of Independence). Racism against native peoples and African slaves was the public dogma. Yet, perhaps he found a human with whom he could share and cherish true human love? The real truth may be a combination of (a) and (b). Such tortured reasoning is reminiscent of William Faulkner's *Absalom, Absalom!* The worldviews of the public and private are so riddled with contradictions that they often lead to bizarre and brutal results.

In the context of all these contradictions stands Thomas Jefferson, the third president of the United States—though (in his own mind) most to be honored as the author of the Declaration of Independence, the Virginia Statue for Religious Freedom, and founder of the University of Virginia.

See also Bacon, Francis; Cicero, Marcus Tullius; Deism; Enlightenment; Harrington, James; Locke, John; Machiavelli, Niccolò; Rights.

Bibliography

WORKS BY JEFFERSON

The Writings of Thomas Jefferson, edited by Paul L. Ford. 10 vols. New York: G. P. Putnam's Sons, 1892–1899.

The Writings of Thomas Jefferson, edited by Andrew Lipscomb. 20 vols. Washington, DC: Thomas Jefferson Memorial Association, 1903–1904.

The Papers of Thomas Jefferson, edited by Julian Boyd, Charles Cullen, and John Catanzariti. 27 vols. (to date). Princeton, NJ: Princeton University Press, 1950–.

Jefferson: Political Writings, edited by Joyce Appleby and Terrance Ball. Cambridge, U.K.: Cambridge University Press, 1999.

SECONDARY TEXTS

Adams, William Howard. *The Paris Years of Thomas Jefferson*. New Haven, CT: Yale University Press, 1997.

Becker, Carl L. *The Declaration of Independence: A Study in the History of Political Ideas*. Rev. ed. New York: Knopf, 1945.

Bernstein, R. B. *Thomas Jefferson*. Oxford: Oxford University Press, 2003.

Brown, Stuart Gerry. "The Mind of Thomas Jefferson." *Ethics* 63 (1963): 79–99.

Cooke, J. W. "Jefferson on Liberty" *Journal of the History of Ideas* 34 (1973): 563–576.

Dumbauld, Edward. *Thomas Jefferson and the Law*. Norman: University of Oklahoma Press, 1978.

Gordon-Reed, Annette. *Thomas Jefferson and Sally Hemings: An American Controversy*. Charlottesville: University Press of Virginia, 1997.

Kennedy, Roger G. *Mr. Jefferson's Lost Cause: Land, Farmers, Slavery, and the Louisiana Purchase*. New York: Oxford University Press, 2003.

Lewis, Jan Ellen, and Peter S. Onuf, eds. *Sally Hemings & Thomas Jefferson: History, Memory, and Civic Culture*. Charlottesville: University Press of Virginia, 1999.

Maier, Pauline. *American Scripture: Making the Declaration of Independence*. New York: Knopf, 1997.

Malone, Dumas. *Jefferson and His Time*. 6 vols. Boston: Little Brown, 1948–1981.

Mayer, David N. *The Constitutional Thought of Thomas Jefferson*. Charlottesville: University Press of Virginia, 1994.

McColley, Robert. *Slavery and Jeffersonian Virginia*. 2nd ed. Urbana: University of Illinois Press, 1973.

Miller, John C. *The Wolf by the Ears: Thomas Jefferson and Slavery*. New York: Free Press, 1977.

Peterson, Merrill. *Thomas Jefferson and the New Nation: A Biography*. New York: Oxford University Press, 1970.

Pocock, J. G. A. *The Machiavellian Moment: Florentine Political Thought and the Atlantic Republican Tradition*. Princeton, NJ: Princeton University Press, 1969.

Reck, Andrew J. "The Declaration of Independence as an 'Expression of the American Mind'" *Revue Internationale de Philosophie* 31 (1977): 401–437.

Sheldon, Garrett Ward. *The Political Philosophy of Thomas Jefferson*. Baltimore, MD: The Johns Hopkins University Press, 1991.

Michael Boylan (2005)

JEVONS, WILLIAM STANLEY

(1835–1882)

The British economist and logician William Stanley Jevons was the son of Thomas Jevons, a Liverpool iron merchant, and Mary Anne Roscoe, a lady of some literary note. After early schooling at Liverpool, he attended University College School and University College, London, where he sat under Augustus De Morgan. In 1854 he left London to take up the post of assayer at the mint in Sydney, Australia, but returned five years later to complete his studies. Soon after, in 1863, he secured a junior teaching position at Owens College, Manchester. By this time he had already published various minor papers on meteorology and economics, a statistical study of commercial fluctuations, and a small work, titled *Pure Logic* (London, 1864, reprinted 1890), reflecting the influence of George Boole. His book on *The Coal Question* (London, 1865) attracted the attention of William Gladstone and was the first to make him known as an economist. In 1866 he was appointed professor of logic and political economy at Manchester, and in the following year married Harriet Taylor, daughter of the proprietor of the *Manchester Guardian.*

Jevons was a conscientious lecturer, but he neither enjoyed nor excelled at the work; and his laborious habits of study led to recurrent breakdowns of health, which had to be repaired by Continental travel, generally to Norway. In spite of this, he wrote prolifically, publishing *The Substitution of Similars* (London, 1869, reprinted 1890); *Elementary Lessons in Logic* (London, 1870), a widely used textbook introductory to J. S. Mill; and *The Principles of Science* (London, 1874; 2nd ed., 1877), his most important contribution to scientific methodology, containing, among much else, an account of his celebrated logical machine. *The Theory of Political Economy* (London, 1871; 2nd ed., 1879) was an equally important landmark in the development of mathematical economics and the theory of utility, followed soon after by a no less influential work of applied analysis and description, *Money and the Mechanism of Exchange* (London, 1875). The once-famous speculations on the relation of sunspot cycles to financial crises, posthumously published in *Investigations in Currency and Finance* (London, 1884), exhibit, more curiously, the range of his interests and the originality of his mind.

Wearying of his duties, Jevons resigned his chair at Manchester in 1876 to take up a similar but more congenial post as professor of political economy at University College, London. This he also resigned, however, in 1880. The main works of this later period were *Studies and Exercises in Deductive Logic* (London, 1880) and *The State in Relation to Labour* (London, 1882). In 1882 he accidentally drowned, probably as the result of a heart attack, while bathing off the coast of Kent. His *Letters and Journal,* edited by his wife (London, 1886), gives an interesting portrait of him. His last work, *The Principles of Economics,* appeared, unfinished, in 1905.

LOGIC

Although marked by no special distinction of style, the writings of Jevons are still worth reading, both for their logical penetration and for their wealth of factual information drawn from many sources of knowledge. His logic owes something to De Morgan and a good deal more to Boole. It represents in the main an attempt to simplify Boole's system by eliminating the more complex and uninterpretable of its mathematical operations and by reducing its procedures of calculation to a mechanical routine. Jevons's own claim to independence in developing his logic as a calculus of qualities, rather than of classes or propositions, is of no great significance; and his method of treating propositions as identities and inferring from them by substitution, though simple enough in

its way, is too lacking in subtlety to have become the "logic of the future" that he once hoped it would be. The most successful of his reforms of the Boolean algebra have been the removal of its inverse operations of subtraction and division and the proposal to read the disjunctive symbol ("either … or") as including the possibility "both"—a practice now universal and resisted at the time only by the conservative John Venn.

Jevons's most interesting adaptation of Boole is to be seen in his method of indirect inference—the principle underlying his "logical piano" and other mechanical aids to calculation—whereby premises are used to eliminate inconsistent combinations of terms from a matrix listing all the possibilities under which a given set of terms and their negatives can be associated. The machine itself, exhibited at the Royal Society in 1870 and described in the *Philosophical Transactions* for the same year, anticipates in its design a number of the features of modern logical computers, while its mode of operation has some fairly obvious affinities with the use of a truth table, though it can hardly be said that Jevons had much grasp of its applications in that respect.

INDUCTION

The logical machine gave its answers only by displaying the combinations compatible with the information fed to it, leaving to the operator the task of finding a compendious formula to express them. The difficulties of this "inverse process" resist mechanical solution and are comparable, in Jevons's view, to those of induction, which he represents accordingly as the inverse operation of deduction—the problem, that is, of deciphering from a given set of phenomena the hidden laws they obey. The treatment of this problem in *The Principles of Science* is in line with the work of William Whewell and De Morgan and in somewhat embittered opposition to the views of Francis Bacon and Mill. Jevons, in short, is an apostle of the hypothetico-deductive method in science, although, unlike Whewell at least, he does not believe it to be a demonstrative procedure or capable of extending knowledge beyond the range of present or past observation. We are necessarily ignorant of the long-term behavior of the universe at large, and when to this ignorance are added the inevitable deficiencies of observation and measurement, it is evident that inductive conclusions can never be more than probable.

PROBABILITY

Jevons was led by the above considerations to give detailed attention to the theories of measurement, approximation, and error and also to bring the whole conception of inductive inference into closer association with the theory of probabilities than was usual with the writers who preceded him. Probability he holds, with De Morgan, to be essentially subjective, though it is a measure of appropriate, rather than of mere actual, belief. It determines "rational expectation, by measuring the comparative amounts of knowledge and ignorance," as represented by the evidence available. That evidence, as nature presents it in the inductive situation, consists of sets of phenomena in combination. Having previously ascertained them (and presumably selected them, somehow, for relevance), we proceed, by more or less intuitive methods (of which Jevons gives no satisfactory account), to erect a hypothesis to explain them. From this in turn we deduce the direct probability of various sets of possible consequences. We then compare these supposed consequences with the known facts in order to determine the probability of their having occurred under the hypothesis in question. This process being repeated for every conceivable hypothesis, we are thereby in a position to assign a probability to each of them by use of the inversion theorem derived, via De Morgan, from Pierre Simon de Laplace. There is no guarantee that by this method the right answers will be forthcoming; but it justifies the adoption of the most probable hypothesis as a matter of practical policy, and that is the best we can expect.

The mathematical theory of inverse probability is, unfortunately, not equal to the weight that Jevons here put upon it, and his conclusions are accordingly unsound. There is no means of knowing that the a priori probabilities of the rival hypotheses are equal, as the theory requires; and there is still less warrant for its extension, by the "rule of succession," to the prediction of new instances or for the employment, where ignorance is total, of the "principle of indifference" to confer a probability of ½ on a proposition merely because knowledge of its truth or falsity is the same (namely, nil) in either case. The fallacies that Jevons committed under this head have since become notorious; the measurement of ignorance is less simple—and nature less like a ballot box—than he was apt to suppose. Errors of conception apart, however, his general view of scientific method has in recent years met with increasing support and is probably his most enduring legacy to the history of thought.

See also Bacon, Francis; Boole, George; De Morgan, Augustus; Induction; Logic, History of; Logic Machines; Mill, John Stuart; Probability and Chance; Venn, John; Whewell, William.

Bibliography

Jevons had no commentators. For a discussion of his logic, see references in the bibliography to the Logic, History of, entry. There is an early criticism of his theory of induction in Thomas Fowler, *Elements of Inductive Logic,* 3rd ed. (London: Clarendon Press, 1876), and scattered remarks of value occur in such general treatises as J. M. Keynes, *A Treatise of Probability* (London, 1921); William Kneale, *Probability and Induction* (Oxford: Clarendon Press, 1949); and G. H. von Wright, *The Logical Problem of Induction,* 2nd ed. (Oxford, 1949), and *A Treatise on Induction and Probability* (London: Routledge and Paul, 1951). See also the brief sketch in J. A. Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1957), pp. 132–136.

P. L. Heath (1967)

JEWISH AVERROISM

The thought of Averroes (Ibn Rushd) was popular in Jewish circles in the Middle Ages, as can be seen by the number of translations made into Hebrew. Some of his books survive only in Hebrew. Not all of these readers could be called Averroists, but some certainly did adhere to what they took to be the central ideas of Averroes himself. Jewish Averroism often included some degree of allegiance to Maimonides, who also developed a complex theory of how to link religion and philosophy. The major Averroists were Isaac Albalag, Joseph ibn Caspi, Moshe Narboni, Elijah Delmedigo, and many other more minor figures extending throughout the South of France and Italy.

One of the main features of Jewish Averroism was its way of distinguishing between rational and religious truths. Proving that religion is true by using reason is a mistake because religion and reason involve entirely different forms of argument. The Jewish Averroists nonetheless argued for the rational superiority of Judaism over against Christianity because the former, unlike the latter, does not call for the acceptance of logically self-contradictory beliefs such as those of transubstantiation, the Trinity, and the Virgin Birth.

THE MAJOR JEWISH AVERROISTS

Isaac Albalag came from the Pyrenees region during the second half of the thirteenth century. Albalag, like Averroes, regarded demonstrative argument to be the paradigmatic method of philosophy. Only philosophers can really use this sort of thought, Albalag claimed, and so only philosophers can really be allowed to say that they know what is true. He argues that when the literal sense of a religious text cannot be reconciled with its philosophical sense, both the literal sense and the philosophical understanding have to be accepted, but in different ways. The literal sense is accepted as something that one would understand completely if one were in the same position of the prophets who had originally transmitted the text. This takes him close to the so-called doctrine of double truth often ascribed to the Christian Averroists in their more radical moments.

Joseph ibn Caspi, born in 1279 in Provence, defended the literal sense of many passages in Scripture as accurate accounts of past events. He gives a naturalistic account of miracles and prophecy; the former are ill-understoood natural events, while prophets, according to Ibn Caspi, are people who understand the links between the present and the future.

Moses Narboni was born in Perpignan around 1300 and was critical of Maimonides' use of arguments drawn from Averroes. Narboni recognized that Averroes sought to challenge the Neoplatonic metaphysics of Ibn Sina (Avicenna), which formed an important part of Maimonides's thinking. Narboni also used Averroes's theory of the active intellect to provide an interest account of philosophical psychology. As human thinking becomes gradually perfected it moves from being largely imaginative to becoming more abstract and intellectual, and the material side comes under the control of thought. This is how religion and prophecy work: In themselves they are abstract but come to affect the material by inspiring and moving people to action. The prophets, according to Narboni, are provided for the majority who do not have the ability to use abstract thought because prophecy represents philosophical truths in imaginative language. There is one truth that is expressed in at least two different ways, one intellectually rigorous and the other practical and effective.

See also Averroes; Jewish Philosophy.

Bibliography

Hayoun, Moise-Ruben. *Moshe Narboni*. Tubingen, Germany: Mohr, 1986.

Leaman, Oliver. "Jewish Averroism." In *History of Islamic Philosophy*, edited by Seyyed Nasr and Oliver Leaman, 760–780. London: Routledge, 1996.

Sirat, Colette. *A History of Jewish Philosophy in the Middle Ages*. Cambridge, U.K.: Cambridge University Press, 1985.

Vajda, Georges. *Isaac Albalag—Averroiste juif, traducteur et commentateur d'Al-Ghazali*. Paris: Vrin, 1960.

Oliver Leaman (2005)

JEWISH PHILOSOPHY

The works constituting the Old Testament touch upon various problems that are discussed in philosophical texts, and the literary forms of some of these works, for instance that of the dialogue in the book of Job and that of Ecclesiastes, bear some similarity to those found in certain philosophical writings. However, a conception of philosophy that included biblical wisdom would lose in clarity and definiteness of outline what it would gain in comprehensiveness. Accordingly, there seems to be a certain amount of justification for considering, as is often done, that the history of Jewish philosophy commences in Alexandria around the beginning of the Christian era, when the first noteworthy attempt was made to use Greek philosophical concepts and methods to come to terms with facts that in the philosophical view are most peculiar, namely, Jewish history as interpreted in religious tradition and biblical revelation.

HELLENISTIC PERIOD

PHILO. The attempt to apply Greek philosophical concepts to Jewish doctrines was made by Philo Judaeus (fl. 20 BCE–40 CE), a prominent member of the Jewish community of Alexandria—he was a member of a delegation sent by this community in the year 39/40, when he was in his own view an old man, to the Roman Emperor Caligula to complain of persecution. Philo, a scholar who combined Greek and Jewish learning, was a most elusive thinker. The immense difficulties that beset any inquiry into Philo's basic conception of the world spring from a variety of sources. Some of the difficulties result from our ignorance of the Greek philosophical authors belonging to Philo's time, for we have only secondhand knowledge of them. Also insufficient is our information about postbiblical Jewish beliefs and speculations, which may be supposed to have shaped Philo's outlook—at least in part and perhaps decisively. However, Philo seems to have had some acquaintance with the oral law, which was being evolved in his time, mainly by the Pharisees in Palestine, and which much later was set down in writing in the Mishnah and in other works belonging to the Talmudic literature. He also knew of the Essenes, whom he praised highly. Some of the sect's theological doctrines, its ethical lore, and its pseudepigraphic literature may have been adapted by Philo to his own purposes.

In a sense Philo's main life's work was hermeneutic. On the one hand, he provided Jewish conceptions with the hallmark of intellectual (or cultural) respectability by stating them in Greek philosophical terms; on the other, he showed that from the point of view of Judaism many Greek notions were unexceptionable—they could be regarded as consonant with Philo's own Jewish doctrine and with the allegorical sense of biblical texts. The homiletic character of most of his writings gave him full scope for his labor of interpretation. He had two schemes of reference—Jewish religious tradition and Greek philosophy—and the fact that he took care to stress the primacy of the former may have been more than mere lip service. In many of Philo's religious speculations the Jewish tradition in the particular form he adopted was not interpreted and explained away—as it was by most of the medieval Aristotelians—as being a mere rehash of philosophical doctrines in a language suited to the limited intellectual capacity of most people. It may be argued with a certain amount of plausibility that in central points of his thought, such as his conception of the Logos, Philo used philosophical notions as trappings for an originally nonphilosophical belief.

A main function of the Logos as conceived by Philo is to serve as an intermediary between the transcendent, unknowable God and the world, a view that probably has a close connection with the view of his Jewish contemporaries concerning the Word (Logos) of God, by means of which he accomplishes his designs. It is significant that the Logos of God is said by Philo to be the place occupied by the world of Ideas: This world is also called by Philo the intelligible world (*kosmos noetos*). The conception of Idea intended here is clearly the Platonic one, conceived of as having been "thought out" by God. The expression used by Philo may indicate that in his time Platonistic philosophers already tended, as the Middle Platonists and the Neoplatonists later did—to place Ideas in the mind of God.

Above philosophical and theological speculations Philo placed mystic ecstasy, of which he may have had a personal experience, "when, … as at noon-tide God shines around the soul, and the light of the mind fills it through and through and the shadows are driven from it by the rays which pour all around it" ("On Abraham," in *Philo,* 10 vols., translated by F. H. Colson and G. H. Whitaker, Cambridge, MA, and London, 1929–1937; Vol. VI, p. 63).

Philo's approach, his method of interpretation, and his way of thinking, as well as some of his conceptions, primarily that of the Logos, exerted a considerable influence on early Christian thought, but not to any comparable extent upon Jewish thought. Later, in the Middle Ages, knowledge of Philo among Jews was either very slight or, in the majority of cases, nonexistent.

TALMUDIC LITERATURE. Most Hellenized Jews were no doubt absorbed into the Christian communities. On the other hand, such historical catastrophes as the destruction of the Temple and the crushing of the various Jewish insurrections by the Romans may have brought about a spiritual withdrawal of the Jews from the circumambient Greco-Roman civilization, a stressing of their separateness. Moreover, as a result of these disasters the spiritual center of Jewry shifted to Iraq, a country that was part of the Persian Empire and less permeated by Greek culture than the regions belonging to the Imperium Romanum.

Some traces of a knowledge of popular, mainly Stoic, philosophy may be found in the Mishnah, a codification of the oral law composed in Palestine in the second century of the Christian era, and in the subsequent Talmudic literature set down in writing in Palestine and Iraq. On the whole, these traces are rather slight. Nevertheless, some scholars believe that the influence of Greek philosophy on Palestinian Jewry was far-reaching, but the case, to say the least, is not proven.

Jewish theological and cosmological speculations occur in the Midrashim, which, under the guise of interpreting biblical verses, propound allegorical interpretations, legends, and myths, and in the *Book of Creation* (*Sefer ha-Yeṣīra*), a work attributed to Abraham, which is a combination of a cosmogony and a grammar. There is no clear evidence of the period in which it was written; both the third century and the sixth or seventh century have been suggested.

MIDDLE AGES

MEDIEVAL LITERATURE. Hayuye (usually called Hivi) al-Balkhī, who appears to have lived in the ninth century in Muslim central Asia, seems to have been a Jewish representative of a brand of free thought also known in Islam, one that under dualistic influence criticized the God of the Bible, who, in view of the prevalence of evil and the fact of his omnipotence, cannot be just. Al-Balkhī seems to have favored Manichaeism—which at that time had a number of adepts—or at least to have been suspected of this heresy; this inference can be made from a preserved fragment of a polemical work directed against him by Saadya in the tenth century. According to Saadya, “the Lord” of al-Balkhī is being eaten, drunk, burnt, and commingled (v. 54 of *Sa’adia Refutatum*), a description that fits the primeval man of Manichaean mythology and the elements belonging to him.

In the ninth and tenth centuries, after a very long hiatus, systematic philosophy and ideology reappeared among Jews, a phenomenon indicative of their accession to Islamic civilization. There is undoubtedly a correlation between this rebirth of philosophy and theology and the social trends of that period, which produced Jewish financiers—some of whom were patrons of learning and who in fact, although perhaps not in theory, were members of the ruling class of the Islamic state—and Jewish physicians who associated on equal terms with Muslim and Christian intellectuals. The evolution of Islam in the ninth and tenth centuries showed that Greek scientific and philosophic lore could be separated at least to some extent from its pagan associations, could be transposed into another language and another culture; it also tended to show—and many Jewish thinkers learned the lesson—that a culture of which the sciences and philosophy and/or theology were an indispensable part could be based upon a monotheistic, prophetic religion that in all relevant essentials was closely akin to Judaism. The question whether philosophy is compatible with religious law (the answer being sometimes negative) constituted the main theme of the foremost medieval Jewish thinkers.

Approximately from the ninth to the thirteenth centuries, Jewish philosophical and theological thought participated in the evolution of Islamic philosophy and theology and manifested only in a limited sense a continuity of its own. Jewish philosophers showed no particular preference for philosophic texts written by Jewish authors over those composed by Muslims, and in many cases the significant works of Jewish thinkers constitute a reply or a reaction to the ideas of non-Jewish predecessors. Arabic was the main language of Jewish philosophic and scientific writings.

There was little regular teaching of philosophy in the religious universities of Islam (though some taught a brand of Kalām approved by the government) and none in the Jewish schools. Many Jewish philosophers seem to have earned their living or a part of it by practicing medicine, a fact that sometimes influenced their thought. A certain number (among them some physicians) were teachers of and authorities in religious law and active in community matters.

Iraq, a very important center of Jewish thought in the ninth and tenth centuries, counted several Jews among its intellectuals steeped in Greek philosophy. However, by far the most productive and influential Jewish thinkers of this period represented a very different tendency, that of the Muʿtazilite Kalām. Kalām (literally, “speech”) is an Arabic term used both in Islamic and in Jewish vocabulary to designate several theological schools that were ostensibly opposed to Greek, particularly to

Aristotelian, philosophy; the Aristotelians, both Islamic and Jewish, regarded Kalām theologians (called the Mutukallimūn) with a certain contempt, holding them to be mere apologists, watchdogs of religion, and indifferent to truth. Herein they did not do justice to their adversaries.

SAADYA. The Muʿtazilite school formed in the eighth century appears to have had, at certain periods, representatives actuated by a genuine theoretical impulse. Its theology, forged in disputes with the Zoroastrians, the Manichaeans, and the Christians, claimed to be based on reason. This belief in reason, as well as most of the tenets of Muʿtazilite theology, were taken over by Saadya ben Joseph (882–942). He prepared an Arabic translation of the books of the Bible provided with commentaries and composed a number of legal and polemical treatises.

Saadya's main theological work, whose Arabic title, *Kitāb al-Amānāt waʿl'i tiqādāt,* may be translated "The Book of Beliefs and Creeds," is modeled to a considerable extent on similar Muʿtazilite treatises and on a Muʿtazilite classification of theological subject matter known as the "Five Principles." Like many Muʿtazilite authors, Saadya starts out by setting forth in his introduction a list and theory of the various sources of knowledge. It may be noted that in beginning systematic theological treatises in this way the Jewish and Islamic adherents of Kalām approximated not Greek philosophical practice but the custom of Indian philosophical writings, which also normally begin by propounding a doctrine of the sources of knowledge (*pramānāh*). The Organon and the expositions of logical disciplines stemming from it that in the Corpus Aristotelicum and in the treatises of the medieval Aristotelians precede the disquisitions on the natural sciences and metaphysics are very different from these analyses of the sources of knowledge.

Knowledge. Saadya distinguished four sources of knowledge: (1) The five senses, (2) the intellect, or reason, (3) necessary inferences, and (4) reliable information given by trustworthy persons. Concerning the first source, he was aware of the doubts expressed by skeptics about the truth of the sense data but rejected these doubts. He held that as a rule a healthy man, one without disabilities, may trust his senses. Exceptional cases do not carry the weight attributed to them by the Skeptics. In Saadya's sense of the word, intellect or reason (*al-ʿaql*) means first and foremost an immediate a priori cognition. In "The Book of Beliefs and Creeds" the intellect is characterized as having immediate ethical cognitions, that is, as discerning what is good and what is evil. However, in his commentary on the book of Proverbs, Saadya also attributes to it the cognition of simple mathematical truths. The third source of knowledge concerns inferences that, if we may judge by the examples given by Saadya, are of the type "if there is smoke, there is fire." These inferences are based on data furnished by the first two sources of knowledge. The fourth source of knowledge is meant to validate the teachings of Scripture and of the religious tradition. Teachings of Scripture must be held to be true because of the trustworthiness of the men who propounded them. One of the main purposes of the work is to show that the knowledge deriving from the fourth source concords with that discovered by means of the other three, or, in other words, that religion and human reason agree.

Saadya's "intellect," postulated as the second source of knowledge, has a function quite different from that of the intellect of the medieval Aristotelians, who did not regard even the most general ethical rules as being a priori cognitions. According to them these rules are accepted as true in virtue of a universal consensus; because of this, validity, unlike that of a priori intellectual truths, can be questioned.

In discussing the third source of knowledge, Saadya does not refer to the Aristotelian theory of the syllogism, but this may be because of ignorance; such knowledge of Greek thought as he possessed was derived mainly from compendiums of doxographers translated into Arabic or adapted by Arabic authors. However, unlike the Muʿtazilites and the Karaites, who were atomists, Saadya adopted a number of doctrines resembling Aristotle's physical views. Nevertheless, he had no use for the conception of an eternal order of nature. This position does not necessarily deny all validity to the theory of genera and species, which is a main concern of the Aristotelian syllogistic, but it certainly tends to limit, or in some cases to negate, the relevance of this theory to the actually existing world.

Theology. Saadya did not merely deny the eternity of the world but held, in common with other, less eclectic partisans of the Muʿtazilite Kalām that the demonstration of the temporal creation of the world must precede and pave the way for the proof of the existence of God the Creator. Of the four arguments which he brought forward in favor of temporal creation, the last is the most noteworthy: Creation in time is an inference from the impossibility of supposing that the past (the whole of time which has elapsed up to the present moment) is of infinite duration—for its infinitude would preclude its coming to an end; the present would never arrive.

Given the demonstrated truth that the world has a beginning in time, it can be proved that it could have been produced only through the action of a Creator. It can further be proved that there can have been only one Creator. God's unity means that he is not a body. It also means, according to a conception taken over from the Muʿtazilites, that he has no attributes superadded to his essence. This applies also to the three attributes that Saadya singled out, perhaps rather inconsistently, as belonging to the Creator: He must be held to be living, possessed of power, and possessed of knowledge.

Justice and free will. The theology of Saadya, like that of the Muʿtazilites, hinges on two principles, of which the unity of God is one; the other is the principle of justice, whose formulation in Islam may have been influenced by attacks of dualists similar to those of Hayuye al-Balkhī (see above), who contended that in view of the existence of evil, an omnipotent God cannot be regarded as just.

This principle takes issue with the view (widespread in Islam and present also in Judaism) that the definition of what is just and what is good depends solely on God's will, to which none of the moral criteria found among men is applicable; according to this view a revelation from God can convert an action now generally recognized as evil into a good action. Against this way of thinking, Saadya and the Muʿtazilites believed that being good and just or evil and unjust are intrinsic characteristics of human actions and cannot be changed by divine decree. The notions of justice and of the good as conceived by man are binding on God himself. In the words of a later thinker, Gottfried Wilhelm Leibniz, he can act only *sub specie boni.* Since, according to Saadya, man has a priori knowledge of good and evil, just and unjust, the fact that human ethical judgments are valid for God means that man's ethical cognitions are also those of the Deity.

This point of view cannot be accorded with strict determinism if one believes, as Saadya professed to do, that men are rewarded for good and punished for evil deeds. It would be contrary to divine justice to condemn or to recompense them for something they cannot help doing; hence, man must be a free agent. For sharing this doctrine with Saadya, the Muʿtazilites were accused of being the dualists of Islam; because of it, they could not regard God as the sole Doer. In Judaism the doctrine of man's free will and free action had very respectable antecedents, and Saadya's position on this point does not seem to have aroused antagonism.

Saadya's simple solution to the problem of reconciling free will with divine prescience seems to be in accord with traditional religious formulas. God has foreknowledge of all the actions that men will perform in the future, but this knowledge does not interfere with human freedom, which enables men to do whatever they wish, both good and evil.

Religious law. The function of religious law is to impose on man the accomplishment of good actions and to prohibit bad ones. Because Saadya believed that man has a priori knowledge of good and evil and that this knowledge coincides with the principles underlying the most important portions of the revealed law, he was forced to ask the question whether this law is not supererogatory. He could, however, point out that whereas the human intellect recognizes that certain actions—for instance, murder or theft—are evil, it cannot by itself discover the best possible definition of what constitutes a particular transgression; nor can it, if it has no other guidance than its own reflections, determine the punishment appropriate for a transgression. On both points the commandments of religious law give the best possible answers.

The commandments of religious law that accord with the behests of the human intellect were designated by Saadya as the "intellectual," or "rational," commandments, According to him they include the duty of manifesting gratitude to the Creator for the benefits he has bestowed upon man. Saadya recognized that a considerable number of commandments, for instance those dealing with the prohibition of work on the Sabbath, do not belong to this category. He held, however, that the obligation to obey them may be derived from the "rational" commandment that makes it incumbent upon man to be grateful to God, for such gratitude entails obedience to his orders.

THE KARAITES. Saadya's adoption of the "rational" Muʿtazilite theology was a part of his overall activity, directed toward the consolidation of rabbinical Judaism, which was being attacked by the Karaites. This Jewish sect, which was founded by Anan ben David in the eighth century and which seems to have had some connection or some affinity with earlier Jewish sects of the period of the Second Temple, rejected the authority of the oral Law, that is, of the Mishnah and the Talmud. In the tenth century and after, the Karaites accepted as their guides the Bible and human reason in the Muʿtazilite sense of the word. Their professed freedom from any involvement with postbiblical Jewish religious tradition obviously facilitated a "rational" approach to theological doctrine. This approach led the Karaite authors to criticize their opponents, the rabbinical Jews, for holding anthropo-

morphic beliefs based, in part at least, on texts of the Talmudic period. In formulating his theology Saadya had in mind the need to disprove this enlightened criticism.

The Karaites themselves adopted wholesale Muʿtazilite Kalām, including its atomism. The atomism of the Karaite theologians has only a very slight similarity to what is known of the theories of Democritus and Epicurus, although Epicurus's hypothesis concerning *minima,* about which we are ill informed, does bear some resemblance to an important point in Islamic and Jewish doctrines. These doctrines appear to have a certain similarity to a Greek mathematical atomism, about which we possess very scanty information. It may derive from the theories of the Pythagoreans and of Xenocrates. Furthermore—and it is a significant point—Muʿtazilite and Karaite atomism in important points are reminiscent of Indian atomistic theories, those of Buddhism and that of the Nyāya-Vaiśeṣika; a historical connection is not wholly impossible.

The Muʿtazilite atomists, followed by the Karaites, held that everything that exists consists of discrete parts. This applies not only to bodies but also to space, to time, to motion, and to the "accidents"—that is, qualities—which the Islamic and Jewish atomists regarded as being joined to the corporeal atoms (but not determined by them, as had been believed by the Greek atomists). An instant of time or a unit of motion does not continue the preceding instant or unit. All apparent processes are discontinuous, and there is no causal connection between their successive units of change. The fact that cotton put into fire generally burns does not mean that fire is a cause of burning; rather, it may be explained as a "habit," signifying that this sequence of what is often wrongly held to be cause and effect has no character of necessity. God's free will, which is not bound by the nonexistent laws of nature, is the only agent of everything that occurs, with the exception of one category. Man's actions are causes that produce effects—for instance, a man who throws a stone at another man, who is then killed, directly brings about the latter's death. This inconsistency on the part of the theologians was necessitated by the principle of justice, for it would be unjust to punish a man for a murder that was a result not of his action but of God's. This grudging admission that causality exists in certain strictly defined and circumscribed cases was occasioned by moral, not physical, considerations. It may be added that because of the opposition it aroused, the Kalām's denial of the existence of a necessary succession of events seems to have strengthened the conviction of the Muslim and Jewish Aristotelians that such order exists and that it is immutable.

ISRAELI. Outside Iraq, philosophical studies were pursued by Jews in the ninth and tenth centuries in Egypt and in the Maghreb. Here the outstanding figure is Isaac ben Solomon Israeli, who died in the beginning of the second half of the tenth century—when he was over a hundred years old, if we are to believe his biographers.

Israeli, a famed physician, was the propagator of a type of philosophy that did not satisfy the exigencies of the strict Aristotelians of a later period; Maimonides denied his being a philosopher, saying that "he was only a physician."

In his philosophical works, such as the "Book of Elements" and the "Book of Five Substances," he drew largely upon the Muslim popularizer of Greek philosophy Abū-Yūsuf Yaʿqūb ibn al-Kindī and also in all probability upon a lost pseudo-Aristotelian text. The peculiar form of Neoplatonic doctrine that seems to have been set forth in this text had, directly and indirectly, a considerable influence on medieval Jewish philosophy.

According to Israeli, God creates through his will and power. This reference to two aspects of the Deity has been compared to certain passages in Plotinus and in Arabic texts that in a considerable measure derive from Plotinus. It may be noted in addition that power and will are singled out for mention as attributes of God in some Christian texts (see, for instance, Ignatius's *Epistle to the Smyrnaeans,* in *The Apostolic Fathers,* edited by Kirsopp Lake, Vol. I, London, 1959, p. 253, and "Isaac ex Judaeo, *Liber Fidei,* in *Patrologia Graeca,* edited by J. P. Migne, Paris, 1857–1866, Vol. XXXIII, Col. 1543). The two things that were created first are form, identified with wisdom, and matter, which is designated as the genus of genera and which is the substratum of everything, not only of bodies, as was the opinion of the Aristotelians, but also of incorporeal substances. This conception of matter seems to derive from the Greek Neoplatonists Plotinus and Proclus, particularly from the latter. In Proclus's opinion, generality was one of the main criteria for determining the ontological priority of an entity. Matter, because of its indeterminacy, obviously has a high degree of generality; consequently, it figures among the entities having ontological priority. According to the Neoplatonic view, which Israeli seems to have adopted, the conjunction of matter and form gives rise to the intellect. A light sent forth from the intellect produces the rational soul. The animal soul is an emanation of the rational soul, and in its turn it gives rise to the vegetative soul.

As far as Jewish philosophy is concerned, Israeli's doctrine of prophecy seems to be the earliest theory attributing prophecy to the influence of the intellect on the imaginative faculty. According to Israeli this faculty receives from the intellect spiritual forms that are intermediate between corporeality and spirituality. This explanation implies that these forms "with which the prophets armed themselves" are inferior to purely intellectual cognitions.

IBN GABIROL. In essentials the schema of creation and emanation propounded by Isaac Israeli and his Neoplatonic source or sources was taken over by Solomon Ben Judah ibn Gabirol, a celebrated Hebrew poet of the eleventh century, who seems to have been the earliest Jewish philosopher of Spain.

Ibn Gabirol's chief philosophical work, "The Source of Life" (or *The Fountain of Life*), written in Arabic, has been preserved in full only in a twelfth-century Latin translation titled *Fons Vitae.*

Fons Vitae makes no reference to Judaism or to specifically Jewish doctrines; it is a nonironical dialogue between a disciple and the master who teaches him true philosophical knowledge. In the Middle Ages it was criticized with some reason for its prolixity; it is also full of contradictions. Nevertheless, it is a strangely impressive work. Few medieval texts so effectively communicate the Neoplatonic conception of the existence of a number of planes of being that differ according to their ontological priority, the derivative and inferior ones constituting a reflection in a grosser mode of existence of those which are prior and superior.

A central conception in Ibn Gabirol's philosophy is concerned with the divine will, which appears to be both part of and separate from the divine essence. Infinite according to its essence, the will is finite in its action. It is described as pervading everything that exists and as being the intermediary between the divine essence and matter and form. Will was one of a number of traditional appellations applied in various, mainly negative, theologies to the entity intermediate between the transcendent Deity and the world or, according to another, not necessarily incompatible interpretation, to the aspect of the Deity involved in creation. According to a statement in *Fons Vitae,* matter derives from the divine essence, whereas form derives from the divine will. This suggests that the difference between matter and form has some counterpart in the godhead and also that universal matter is superior to universal form. Some of Ibn Gabirol's statements seem to bear out the latter impression; other passages, however, appear to imply a superiority of universal form. The apparent contradiction seems to result from two conflicting approaches: the Aristotelian, which assumes that form (which is held to be *in actu*) is superior to matter (which per se exists only potentially), and the Neoplatonic, which in at least one of its manifestations consistently professed the superiority of matter, which, being indeterminate, could be held to be of a more universal, all-encompassing nature than form.

Form and matter, whether they be universal or particular, exist only in conjunction. All things, with the sole exception of God, are constituted through the union of the two; the intellect no less than the corporeal substance. In fact, the intellect is the first being in which universal matter and form are conjoined. In other words, Ibn Gabirol considered—in accord with Israeli—that the intellect is not one simple substance, as was thought by the faithful disciples of Aristotle; in his view its unity proceeds from a duality. The intellect contains and encompasses all things. It is through the grasp of the various planes of being, through ascending in knowledge to the world of the intellect and cognizing what is above it—the divine will and the world of the Deity—that man may "escape death" and reach "the source of life."

In the twelfth century Ibn Gabirol's system seems to have enjoyed a certain vogue among Jewish intellectuals living in Spain. Thus, Joseph Ben Jacob ibn Zaddik (d. 1149) and Abraham ibn Ezra (c. 1092–1167) were at least to some extent disciples of his. Ibn Zaddik was the author of the *Microcosm,* a work written in Arabic but extant only in a Hebrew translation, which draws a parallel between man and the microcosm. However, Abraham ibn Da'ud (see below) criticized Ibn Gabirol at length, denouncing the feebleness of his argumentation and the incorrect (that is, non-Aristotelian) conception of matter.

HALEVI. Yehuda Halevi (c. 1075–1141), also of Spain, who, like Ibn Gabirol, was a Hebrew poet, has the distinction of being the earliest and the most outstanding medieval Jewish thinker whose theology or philosophy (he would have repudiated the latter term) does not merely take Judaism in its stride, as was largely true of Saadya and the Karaites, to mention only two, but is consciously and consistently based upon arguments drawn from Jewish history.

His views are set forth in an Arabic dialogue whose full title is translated as "The Book of Proof and Demonstration in Aid of the Despised Religion." According to a custom that finds some justification in one of Halevi's letters, this work is usually referred to as the "Kuzari," the

Hebrew name of the king of the Khazars who is one of the two protagonists of the dialogue.

Basing his narrative on the historical fact that the Khazars were converted to Judaism, Halevi relates that their king, a pious man who did not belong to any of the great monotheistic religions, dreamed of an angel who said to him, "Your intentions are pleasing to the Creator, but your works are not." To find the correct way of pleasing God, the king seeks the guidance of a philosopher, of a Christian, of a Muslim, and, finally, after hesitating to have recourse to a representative of a people degraded by its historical misfortune, of a Jewish scholar, who converts him to Judaism.

The words of the angel heard in a dream may, in accordance with both religious and philosophical doctrine, be regarded as an (inferior) species of revelation. The use of this element of the story enabled Halevi to suggest that it is not the spontaneous activity of human reason that impels man to undertake the quest for the true religion; for this one needs the gift of prophecy, or at least a touch of the prophetic faculty (or a knowledge of the revelations of the past).

The argument of the philosopher whose advice is sought by the king brings this point home. This disquisition is a brilliant piece of writing, for it lays bare the essential differences—which the medieval philosophers often endeavored to dissimulate by means of circumlocution and double talk—between the Aristotelian God, who is totally ignorant of and consequently wholly indifferent to human individuals, and the God of religion.

Within the framework of philosophical doctrine, the angel's words are quite meaningless. Not only is the God of the philosophers, who is a pure intellect, not concerned with man's works, but the (cultural) activities, involving both mind and body, to which the angel clearly referred, cannot from the philosophical point of view either help or hinder man in the pursuance of the philosophers' supreme goal, the attainment of union with the Active Intellect. This union was supposed to confer knowledge of all the intelligibles. Thus, man's supreme goal was supposed to be of a purely intellectual nature.

In opposition to the philosopher's faith, the religion of Halevi's Jewish scholar is based upon the fact that God may have a close, direct relationship with man, who is not conceived primarily as a being endowed with intellect. The postulate that God can have intercourse with a creature made of the disgusting materials that go into the composition of the human body is scandalous to the king and prevents his acceptance of the doctrine concerning prophecy expounded by the Muslim sage (just as the extraordinary nature of the Christological dogmas deters him from adopting Christianity). It may be noted that the opposition on this point between the king and the philosopher on the one hand and the Jew and the Muslim on the other reflects one of the main points of controversy between pagan authors and the Church Fathers (and some Gnostics in the first centuries of the Christian era). The moot point is whether a superior kind of man or, as many pagans believed, the souls or spirits ruling the heavenly bodies are the proper intermediaries between God, humankind, and the teachers of the arts and sciences. An echo of this controversy is found in Arabic literature, and Halevi, in developing his point of view, had probably adapted to some extent an older, non-Jewish source, at the same time making extensive use of Jewish religious tradition.

His position is that it is contemplation not of the cosmos but of Jewish history that procures knowledge of God. Halevi was aware of the odium attaching to the doctrine of the superiority of one particular nation; he held, however, that only this doctrine explains God's dealing with humankind, which like many other things, reason is unable to grasp. The controversies of the philosophers serve as proof of the failure of human intelligence to find valid solutions to the most important problems. Halevi's description of the specific Jewish position has also exercised a certain fascination upon several modern Jewish philosophers, such as Franz Rosenzweig.

HALEVI'S CONTEMPORARIES. As a speculative author Halevi was by no means an isolated phenomenon. During the period comprising the second half of the eleventh century and the first half of the twelfth century a number of Jewish thinkers appeared in Spain.

Bahya. In this period Bahya ben Joseph ibn Paquda (second half of the eleventh century) wrote one of the most popular books of Jewish "spiritual" literature, the "Commandments of the Heart," which combines a theology influenced by although not identical with that of Saadya with a moderate mysticism inspired by the teachings of the Muslim Sufis. The commandments of the heart—that is, those relating to men's thoughts and sentiments—are contrasted with the commandments of the limbs—that is, the Mosaic commandments enjoining or prohibiting certain actions. Bahya maintained that both sets of commandments should be observed, thus rejecting the antinomistic position. However, he made clear that first and foremost he was interested in the commandments of the heart.

Bar Hiyya. Abraham bar Hiyya (first half of the twelfth century), an outstanding mathematician, an astrologer, and a philosopher, outlined in *Megillat ha-Megalleh* a view of Jewish history which in some particulars is rather reminiscent of that of Yehuda Halevi but which does not emphasize to the same degree the uniqueness of that history and is set forth in much less impressive fashion. Living in Barcelona under Christian rule, bar Hiyya wrote his scientific and philosophical treatises not in Arabic but in Hebrew.

Ibn Ezra. Hebrew was also used by Abraham ibn Ezra, a native of Spain, who traveled extensively in Christian Europe. His commentaries on the Bible contributed to the diffusion among the Jews of Greek philosophical thought, to which Ibn Ezra made many, although as a rule disjointed, references.

Abuʾl-Barakāt al-Baghdādī. The last outstanding Jewish philosopher of the Islamic East, Abuʾl-Barakāt al-Baghdādī (died as a very old man after 1164), sometimes called Abuʾl-Barakāt ibn Malkā, also belongs to this period. Being a borderline case he illustrates a certain indeterminacy in the definition of a Jewish thinker.

Abuʾl-Barakāt al-Baghdādī, an inhabitant of Iraq, was converted to Islam in his old age (for reasons of expediency, according to his biographers). His philosophy appears to have had a very strong impact on Islamic thought, whereas its influence upon Jewish philosophy and theology is very hard to pin down and may be practically nonexistent. His chief philosophical work, *Kitāb al-Muʿtabar,* a title that according to Abuʾl-Barakāt's own interpretation means "The Book of That Which Has Been Established by Personal Reflection," has very few references to Jewish texts or topics. His theory appears mainly to represent a kind of dialectic development of Avicenna's doctrine concerning the existence of the soul; it is a radicalization that plays havoc with the greater part of Avicenna's psychology and theology. On the other hand, another important work of his, a philosophical commentary on Ecclesiastes, attests his knowledge of and interest in Jewish tradition.

Ibn Kammūna. Ibn Kammūna, who lived in the second half of the thirteenth century, may be regarded as the last Jewish philosopher of the Islamic East. There is a possibility that he too was converted to Islam. He wrote a curious treatise, *Tanqīh al-abhāth biʾl-mabhath ʿan al-milal al-thalāth,* dealing, ostensibly impartially, with the three monotheistic religions—Judaism, Islam, and Christianity. His philosophical doctrine seems to derive from Avicenna and his thirteenth-century disciple Naṣīr al-Dīn al-Ṭūsī.

IBN DAʾUD. With regard to the adoption of Aristotelianism (including such systems as that of Avicenna, which in many essentials stems from, but profoundly modifies, the pure Peripatetic doctrine) there is a considerable time lag between the Islamic East in the one hand and Muslim Spain and the Maghreb on the other.

Abraham ibn Daʾud (died in the second half of the twelfth century), who is regarded as the first Jewish Aristotelian of Spain, was primarily a disciple of Avicenna. According to a not unlikely hypothesis, he may have translated or helped to translate some of Avicenna's works into Latin, for Ibn Daʾud lived under Christian rule in Toledo, a town that in the twelfth century was a center for translators. His historical treatises, written in Hebrew, manifest his desire to familiarize his coreligionists with the historical tradition of the Latin world, which at that time was alien to most of them. But his philosophical work, *Sefer ha-Emunah ha-Ramah* (The book of sublime religion), written in 1161 in Arabic, shows few, if any, signs of Christian influence.

The doctrine of emanation set forth in *Sefer ha-Emunah ha-Ramah* describes in the manner of Avicenna the procession of the ten incorporeal intellects, the first of which derives from God. This intellect produces the second intellect, and so on. Ibn Daʾud questioned in a fairly explicit manner Avicenna's views on the way the second intellect is produced; his discipleship did not by any means spell total adherence.

Ibn Daʾud's psychology was also, and more distinctively, derived from Avicenna. The argumentation leading to a proof that the rational faculty is not corporeal attempts to derive the nature of the soul from the fact of immediate self-awareness. Like Avicenna, Ibn Daʾud tended to found psychology on a theory of consciousness.

Concerning "practical" philosophy, that is, ethics and political theory, Ibn Daʾud was of the opinion that all that Aristotle discovered in this field of inquiry can be found in the Torah in a more perfect manner.

Sefer ha-Emunah ha-Ramah was said by its author to have been written in response to a question concerning free will and determinism. Obviously, this problem is closely bound up with the problem of God's knowledge. According to Ibn Daʾud events in this world are in part predetermined by necessity and in part contingent. Insofar as they are contingent, their occurrence or failure to occur may depend on man's actions. The necessary events are known by God as necessary, and the contingent as contingent. With regard to contingent events, he has no

certain knowledge of whether they will come about in the future.

Ibn Da'ud often referred to the accord that, in his view, existed between philosophy and religious tradition. As he remarked, *Sefer ha-Emunah ha-Ramah* was not meant to be read either by readers who, in their simplicity, are satisfied with what they know of religious tradition or by those who have a thorough knowledge of philosophy. It was intended for readers of one type only, those who, being on the one hand acquainted with the religious tradition and having on the other some rudiments of philosophy, are "perplexed." It was for the same kind of people that Maimonides wrote his *Guide of the Perplexed.*

MAIMONIDES. Maimonides (Moses ben Maimon, 1135–1204), a native of Spain, is incontestably the greatest name in Jewish medieval philosophy, but it is not because of outstanding originality in philosophical thought, in the proper sense of the term, that his reputation is deserved. Rather, the distinction of Maimonides, who is also the most eminent codifier of Jewish religious law, is to be found in the vast scope of his attempt in the *Guide of the Perplexed* to safeguard both the religious law and philosophy (whose divulgation is, as he was aware, destructive of the law), without suppressing the issues and without trying to impose, on the theoretical plane, a final, universally binding solution of the conflict.

As Maimonides made clear in his introduction to the *Guide,* he regarded his self-imposed task as perilous, and he therefore had recourse to a whole system of precautions destined to conceal his true meaning from the people who, lacking the necessary qualifications, were liable to misread the book and abandon observance of the law. According to Maimonides' explicit statement, these precautions include deliberately contradictory statements meant to mislead the undiscerning reader. It clearly follows that there is no possibility of propounding an interpretation of Maimonides' doctrine which would not be disproved or seem to be disproved by some passage or other of the *Guide.* Nevertheless, a consideration of the system as a coherent whole and of certain indications found in this work appears to suggest that Maimonides' true opinions on certain capital points are not beyond conjecture.

Conception of God. The apparent or real contradictions that may be encountered in the *Guide* are perhaps most flagrant in Maimonides' doctrine concerning God. There seems to be no plausible hypothesis capable of explaining away the differences between the following three views:

(1) God has an eternal will that is not bound by natural laws. Through an act of his will he created the world in time and imposed on it the order of nature. This creation is the greatest of miracles; if and only if it is admitted can other miracles, such as God's interventions, which interfere with the causally determined concatenation of events, be regarded as possible. The philosophers' God who is not free to cut the wings of a fly is to be rejected. This conception is in keeping with the traditional religious view of God and is adopted by Maimonides, if a statement of his is to be taken at its face value, because failure to do so would undermine religion.

(2) Man is incapable of having any positive knowledge concerning God. The ascription to God of the so-called divine attributes—wisdom or life, for instance—should not be regarded as an assertion that God is endowed with a positive quality designated as wisdom or life because it is similar to the corresponding quality found in created beings, for the fact is that their being homonyms is the only resemblance between human and divine wisdom or, for that matter, between man's and God's existence. Contrary to the attributes predicated of created beings, the divine attributes are strictly negative; they state what God is *not*; for instance, he is not *not-wise,* which, as Maimonides believed, is not a positive assertion.

Negative theology of a similar kind may be found in the writings of Islamic philosophers, such as Avicenna, who are known to have had some influence on Maimonides, but they put much less emphasis on this aspect of their doctrine concerning God. Maimonides used it, inter alia, to justify the statement that the only positive knowledge of God possible is that which is known through his acts, identified in the *Guide* as the sometimes beneficent and sometimes destructive operations of the natural order. In other words, human knowledge of God is assimilated into the knowledge of the two sciences that treat of this order, physics and metaphysics.

(3) In accordance with the doctrine of Aristotle, God is an intellect. The formula current among medieval philosophers which maintains that in him the cognizing subject, the cognized object, and the act of intellectual cognition are identical derives from Aristotle's thesis that God cognizes only himself. Maimonides, however, in adopting the formula interpreted it in the light of human psychology and epistemology, pointing out that according to a theory of Aristotle the act of human (not only of divine) cognition brings about an identity of the cogniz-

ing subject and cognized object. The parallel drawn by Maimonides between the human and the divine intellect quite evidently implies a certain similarity between the two; in other words, it is incompatible with the negative theology of other passages of the *Guide.* Maimonides' interpretation also implies that God knows not only himself (if the reflexive pronoun is taken to refer to his transcendent essence only) but also objects of cognition, that is, intelligibles held to be outside himself; however, in virtue of the eternal act of cognition, the objects of cognition—which should perhaps be assimilated into the intelligible structure of the world—are identical with God himself.

In view of the relation that it implies between God and the world, the conception of God as an intellect can scarcely be reconciled with Maimonides' negative theology; nor can it be reconciled with his theological doctrine, which is centered on God's will and which asserts that the structure of the world (created in time) came into being through the action of his will.

Prophecy. The enigma of the *Guide* would be nonexistent if Maimonides could be held to have believed that truth can be discovered in a suprarational way, through revelations vouchsafed to the prophets. This, however, is not the case. Maimonides held that the prophets (with the exception of Moses) combine great intellectual abilities, which qualify them to be philosophers, with a powerful imagination. As he put it, the intellectual faculty of the philosophers and the prophets receives an "overflow" from the Active Intellect. In the case of the prophets this "overflow" not only brings about intellectual activity but also passes over into the imaginative faculty, giving rise to visions and dreams. The fact that prophets have a strong imagination gives them no superiority in knowledge over philosophers, who do not have it. Moses, who belonged to a higher category than the other prophets, did not have recourse to imagination. According to another text of Maimonides, his commentary on the Mishnah, the prophets achieve union with the Active Intellect; hence, they are the supreme philosophers.

The laws and religion as instituted by Moses are intended not only to ensure the bodily welfare and safety of the members of the community but also to facilitate the attainment of intellectual truths by individuals gifted enough to uncover the various hints embodied in religious laws and practices. This does not mean that all the beliefs inculcated by Judaism are true. Some indeed express philosophical truths, although in an inaccurate way, in a language suited to the intellectual capacity of the common people, who in general cannot grasp the import of the dogmas they are required to profess. Other beliefs, however, are false, but "necessary" for the preservation of a public order upholding justice. Such is the belief that God is angry with wrongdoers.

Religious law. As far as the law—that is, the religious commandments—is concerned, two aspects of Maimonides' position may be distinguished. On the one hand, he had to maintain that it is unique in its excellence; there is no basis of comparison between Moses, who promulgated this law, and any other prophet (or any other man) who existed in the past or who may appear in the future and, consequently, the law is valid for all time. This profession of faith, at least with regard to its assumptions about the future, lacked philosophical justification; however, in view of the Muslim polemics and perhaps also in view of incipient tendencies among the "perplexed" to neglect the observance of the commandments, it could be regarded as necessary for the survival of Judaism.

In its second aspect Maimonides' position is characterized by his awareness of the role of historical contingencies in the institution of the commandments. He insisted time and again that Moses had to fulfill two requirements: His law had to be different, but it could not be too different from the customs and ordinances of the pagans among whom the children of Israel lived. The people could not have borne too sharp a break with the way of life to which they were accustomed. For instance, the commandments concerning sacrifice arise from this awareness of the necessities of a specific historic situation. Like nature, which uses many complicated devices in forming a viable organism, the political leader, who must fashion his community, is sometimes compelled to have recourse to a "ruse" or a roundabout method.

Like Aristotle, Maimonides held that the "theoretical life" constitutes the highest perfection possible to man. But he believed (partly under the influence of the Platonic political doctrine adopted by al-Farabi and others) that certain individuals, for example, the Patriarchs and Moses, are capable of combining contemplation with a life of action. In its supreme manifestation the activity of the prophet-lawgiver imitates that of God or nature.

THE THIRTEENTH CENTURY. For four or five centuries (and in certain regions for an even longer time), the *Guide of the Perplexed* exercised a very strong influence in the European centers of Jewish thought; in the thirteenth century, when the *Guide* was twice translated into Hebrew, these centers were Spain, the south of France, and Italy. Rather paradoxically, in view of the unsystem-

atic character of Maimonides' exposition, it was used as a standard textbook of philosophy—and condemned as such when the teaching of philosophy came under attack. The performance of this function by the *Guide* was rendered possible or at least facilitated by the fact that from the thirteenth century onward the history of Jewish philosophy in European countries acquired a continuity it had never had before. First and foremost, this development seems to have resulted from a linguistic factor: In Spain, where the Christian reconquest was destroying piecemeal the power of Islam, Jewish philosophers abandoned the use of Arabic as the language of philosophical exposition. The Jews did not, however, switch to Latin, the language of Christian philosophy. They and their coreligionists in other European countries wrote in Hebrew, and they read original and translated texts extant in Hebrew, which were much less numerous and less diverse than those found in Arabic philosophical literature. Owing to the existence of a common and relatively homogeneous philosophical background and to the fact that Jewish philosophers reading and writing in Hebrew naturally read the works of their contemporaries and immediate predecessors, something like a dialogue can be discerned. In striking contrast to the immediately preceding period, European Jewish philosophers in the thirteenth century and after frequently devoted a very considerable part of their treatises to discussions of the opinions of other Jewish philosophers. That many of the Jewish philosophers in question wrote commentaries on the *Guide* undoubtedly furthered this tendency.

The influence of Maimonides' contemporary Averroes, many of whose commentaries and treatises were translated into Hebrew, was second only to that of Maimonides. Indeed, it may be argued that for philosophers, as distinct from the general reading public, it often came first. In certain cases, commentators on the *Guide* tend, in spite of the frequent divergences between the two philosophers, to quote Averroes's opinions in order to clarify those of Maimonides.

The influence of Christian scholastic thought on Jewish philosophy was in very many cases not openly acknowledged in the period beginning with the thirteenth century, but it seems to have been of great significance. Samuel ibn Tibbon, one of the translators of the *Guide* into Hebrew and a philosopher in his own right, remarked on the fact that the philosophical sciences were more widely known among Christians than among Muslims. Somewhat later, at the end of the thirteenth century and after, Jewish scholars in Italy (Hillel of Verona and others) translated into Hebrew texts of Thomas Aquinas and other Scholastics; not infrequently, although by no means always, some of them acknowledged the debt they owed their Christian masters.

In Spain and in the south of France a different convention seems to have prevailed up to the second half of the fifteenth century. Whereas Jewish philosophers of these countries felt no reluctance about referring by name to Greek, Arabic, and of course other Jewish philosophers, as a rule they refrained from citing Christian thinkers whose views had, in all probability, influenced them. In the case of certain Jewish thinkers this absence of reference to the Christian Scholastics served to disguise the fact that in many essentials they were representative of the philosophical trends, such as Latin Averroism, that were current among the Christian Scholastics of their time.

Albalag. Quite evident is the resemblance between certain views professed by the Latin Averroists and the parallel opinions of Isaac Albalag, a Jewish philosopher who lived in the second half of the thirteenth century, probably in Catalonia, Spain, and who wrote a commentary in Hebrew on the "Intentions of the Philosophers," an exposition of Avicenna's doctrine written by the Muslim philosopher Mohammad al-Ghazali. No serious attempt at interpreting Albalag's assertion that both the teachings of the Bible and the truths demonstrated by reason must be believed even if they are contradictory can fail to pose the question whether some historical connections exists between this view and the Latin Averroist doctrine that there are two sets of truths, the religious and the philosophical, and that these are not necessarily in accord.

In most other points Albalag was a consistent follower of the system of Averroes himself (although a few of Albalag's doctrines appear to be in closer accord with Ibn Sīnā). This philosophical position may be exemplified by his rejection of the view that the world was created in time. He professed, it is true, to believe in what he called "absolute creation in time." However, this expression merely signifies that at any given moment the continued existence of the world depends on God's existence, an opinion which is essentially in harmony with Averroes.

Bedersi. Yedaʿya Hapnini Bedersi, of Béziers in the south of France, who lived from the end of the thirteenth century to the beginning of the fourteenth century, appears to have been influenced by the teaching of John Duns Scotus, for he believed in the existence of what he called individual forms, which seem by and large to correspond to the *haecceitas* of the Scotists.

Kaspi. One of Bedersi's contemporaries, Joseph Kaspi, a prolific philosopher and exegetical commentator, maintained a somewhat unsystematic philosophical position that seems to have been influenced by Averroes. He expressed the opinion that knowledge of the future, with that of God himself, is like that possessed by experienced people concerning the way in which business transactions or marriages may be expected to turn out—that is, such knowledge is of a probabilistic nature. The prescience of the prophets is of the same nature. It is more than likely that Kaspi's interest in this problem had some connection with the debate about future contingents in which Christian Scholastics were engaged at that time.

Kaspi also held that in view of the vicissitudes of history the return of the Jews to Palestine may on probabilistic grounds be considered likely. As a result he rejected the distinction—which for Yehuda Halevi, for instance, had been a basic one—between sacred and profane history, the first being the history of the people of Israel and the second that of other nations.

LATE MEDIEVAL PERIOD

KABBALAH. One of the most urgent problems with which Jewish philosophers were faced in the fourteenth and fifteenth centuries was that of the attitude to be adopted toward the Kabbalah (literally, "tradition"), a body of mystic and Gnostic doctrines, part of which was being elaborated in that period in the countries in which the philosophers lived.

Many of the Kabbalists incorporated philosophical doctrines in their writings and claimed Maimonides as a Kabbalist but at the same time regarded philosophy as such as an inferior kind of science. This disdainful attitude was reciprocated by some philosophers. Nonetheless, attempts were made to effect a reconciliation between philosophy and Kabbalah. Such an attempt was made by Joseph ibn Wāqār, a fourteenth-century philosopher of Toledo, who wrote "The Treatise Which Reconciles Philosophy and Religious Law" in Arabic (which in that period and country was atypical for a Jewish philosopher). According to Ibn Wāqār, the opinions of the philosophers are founded on reason, whereas those of the Kabbalists owe their validity solely to their having been transmitted by a tradition whose authority guarantees their truth. Although recognizing in theory the superiority of the Kabbalistic doctrine to the teachings of philosophy, Ibn Wāqār endeavored to show the basic similarity, masked by a difference of terminology, of the two systems of thought. He also affirmed that knowledge of philosophy increases the aptitude to apprehend the mystic doctrine of the Kabbalah.

NARBONI. Moses of Narbonne, or Moses Narboni, who lived in the south of France in the fourteenth century, was, like many other Jewish writers of this period, mainly a writer of commentaries. He wrote commentaries on biblical books, on treatises of Averroes, apocryphal treatises, and on Maimonides' *Guide.* In his commentary on the *Guide,* Narboni often interprets the earlier Jewish philosopher's opinions by recourse to Averroes's views. Narboni also expounded and gave radical interpretations to certain conceptions that he understood as implied in the *Guide.*

According to Narboni, God participates in all things because he is the measure of all substances. From another point of view all things exist in God, "the Agent being the essence of the patient." God is the form of the world. In Narboni's interpretation (which, not quite correctly, he opposed to that of Maimonides) this formula means that God is a form which, although it is not in a body, is "with a body": God's existence appears to be bound up with that of the world, to which he has a relation analogous to that existing between a soul and its body (a comparison already made in the *Guide*). As the form of the world, God also determines the fact that the extension of the world is limited. It may be added that, according to a conception of Narboni that runs counter to the views of many Aristotelian philosophers, prime matter has its place in the thought of God. Narboni seems to have been a consistent (and on the whole unusually outspoken) adherent of the Aristotelian tradition that crystallized in the Arabic period of Jewish philosophy.

GERSONIDES. Gersonides (Levi ben Gerson, 1288–1344), another fourteenth-century Jewish philosopher born in the south of France, wrote the systematic philosophical work *Milhamot Adonai* (The wars of the Lord) as well as many philosophical commentaries. As an astronomer he enjoyed a certain fame among Christian scholars. Gersonides apparently never explicitly mentioned Christian scholastic philosophers; he cited Greek, Arabic, and Jewish thinkers only, and in many ways his system appears to have stemmed from the doctrines of Maimonides or Averroes, regardless of whether he agreed with them. For example, he explicitly rejected Maimonides' doctrine of negative theology. However, a comparison of his opinions and of the particular problems that engaged his attention with the views and debates found in scholastic writings of his period suggest that he was also influenced by the Latins on certain points.

Creation. Gersonides disagreed both with the Aristotelian philosophers who maintained the eternity of the world and with the partisans of the religious who believed in the creation of the world in time out of nothing. He maintained that God created the world in time out of a preexistent body lacking all form. As conceived by Gersonides this body seems to be similar to primal matter. According to the Aristotelian conception, the "now" separates the past from the future; because of this function its existence at any moment of time entails the existence of a past. Hence, an absolute beginning is impossible, which means that the world is eternal. Gersonides rejected this argument because he believed that it is possible for a "now" to be restricted to the function of beginning or terminating an interval of time. Hence, there is no difficulty in supposing that the existence of a "now" at the instant of the creation of the world in time did not entail the existence of a past. This argument was discussed prior to Gersonides in a Latin Averroistic treatise whose author is unknown, and Gersonides may have been influenced by Latin Scholasticism on this point.

Free will and divine omniscience. The problem of human freedom of action and a particular version of the problem of God's knowledge of future contingents form an important part of Gersonides' doctrine. Gersonides—who, unlike the great Jewish and Muslim Aristotelians, believed in astrology—held that all happenings in the world except human actions are governed by a strict determinism. God's knowledge does not, however, extend to the individual human acts that actually occur. It embraces the general order of things that exist; it grasps the laws of the universal determinism but is incapable of apprehending events resulting from man's freedom. Thus, the object of God's knowledge is an ideal world order, which differs from the "real" world insofar as the latter is in some measure formed according to man's free will.

Political philosophy. In political and social doctrine there is a fundamental difference between Maimonides and Gersonides. Gersonides does not appear to have assigned to the prophets any political function; according to him their role consists in the prediction of future events. The providence exercised by the heavenly bodies ensures the existence in a given political society of men having an aptitude for and exercising the handicrafts and professions necessary for the survival of the community. He remarked that in this way the various human activities are distributed in a manner superior to that outlined in Plato's *Republic.* Thus, he rejected explicitly Plato's political philosophy, which, having been adapted to a society ruled through the laws promulgated by a prophet, had been an important element of Jewish philosophy in the Arabic period.

Gersonides' deviations from this philosophical tradition may have involved various factors, such as the influence of Thomas Aquinas (whose conception of human freedom, to mention but this example, resembles that of Gersonides) or Gersonides' belief in astrology or his pronounced predilection for personal speculation. These deviations did not, however, affect his fundamental allegiance to medieval Aristotelianism.

CRESCAS. Both Hasdai Crescas (1340–1410), a Spanish Jewish thinker, and Gersonides had thorough knowledge of Jewish philosophy and partial knowledge of Islamic philosophy, and both seem to have been influenced by scholastic thought; moreover, in certain important respects Crescas was influenced by Gersonides himself. However, in Crescas's main work, *Or Adonai* (The light of the Lord), one of his objectives, quite contrary to Gersonides, was to expose the weakness and insufficiency of Aristotelian philosophy. This attitude may be placed in the wider context of the return to religion itself as opposed to the Aristotelian rationalization of religion and the vogue of Kabbalah, characteristic features of Spanish Jewry in Crescas's time. This change in attitude has been regarded as a reaction to the increasing precariousness of the position of the Jewish community in Spain.

The low estimation of the certainties and the rationalistic arrogance of the medieval Aristotelians coincided chronologically with a certain disintegration of and disaffection toward what may be called the classical Aristotelian Scholasticism. Relevant to this decline were the so-called voluntarism of Duns Scotus, the nominalism of William of Ockham and other Scholastics, and the development, in the fourteenth century and after, of the anti-Aristotelian terminist physics at the University of Paris and elsewhere. It is significant that there is a pronounced resemblance between Crescas's views and two of these trends, Scotism and the "new" physics.

Divine attributes. Crescas accepted Gersonides' view that divine attributes cannot be negative, but unlike his predecessor he centered his explanation of the difference between the attributes of God and those of created existents on the antithesis between an infinite being and finite beings. It is through infinitude that God's essential attributes—wisdom, for instance—differ from the corresponding and otherwise similar attributes found in created beings. In Crescas's as in Benedict Spinoza's

doctrine, God's attributes are also infinite in number. The central place assigned to the thesis of God's infinity in Crescas's system suggests the influence of Duns Scotus's theology, which is similarly founded upon the concept of divine infinity.

Physics. The problem of the infinite approached from an altogether different angle was one of the main themes of Crescas's critique of Maimonides' twenty-five propositions; these propositions, concerned mainly with Aristotelian physical doctrines, had been set forth in the *Guide* as the basis of Maimonides' proofs for the existence of God. Crescas's declared purpose in criticizing and rejecting several of these propositions was to show that the traditional Aristotelian proofs (founded in the first place on physical doctrines) were not valid.

In the course of his critique Crescas attempted to disprove the Aristotelian thesis that the existence of an actual infinite is impossible. He held that space is not a limit but a tridimensional extension, that it is infinite, and that, contrary to Aristotle, the existence of a vacuum and of more worlds than one is possible. He also criticized as being impossible the thesis of the Aristotelian philosophers that there exists an infinite number of causes and effects, which have order and gradation. This thesis refers not to a temporal succession of causes and effects which have a similar ontological status but to a vertical series, descending from God to the lowest rung in creation. His attacks were likewise directed against the Aristotelians' conception of time and of matter.

The physical doctrines that emerged in Crescas's critique resemble the "new," mainly "terminist" physics, which was being worked out in the fourteenth century at the University of Paris and other Christian seats of learning and which had a considerable influence on the classical physical theories of Galileo Galilei and others. There is no difficulty in supposing that Crescas was acquainted with some of the terminist theses. Crescas may on the whole be regarded as an outstanding representative of the medieval "new physics."

Ascendancy of soul over intellect. Crescas's fundamental opposition to Aristotelianism is perhaps most evident in his rejection of the conception of intellectual activity as the supreme state of being for man and for God. Crescas's God is not first and foremost an intellect, and the supreme goal to which man can aspire is to love God with a love corresponding as far as possible to the infinite greatness of its object and to rejoice in the observance of his commandments. God too loves man, and his love, in spite of the lowliness of its object, is proportionate to his infinity.

Crescas attacked the separation of the intellect from the soul as conceived by the Aristotelians and attempted, perhaps under the influence of Yehuda Halevi, to refute the Peripatetic doctrine that the actualized intellect, in contradistinction to the soul, survives the death of the body. According to Crescas the soul is a substance in its own right and can be separated from the body; it continues to subsist after the body's death.

Crescas's depreciation of the intellect did not lead to an emphasis on man's freedom of action. Crescas's view concords with that of Avicenna: There is no such freedom; everything in the world is subject to a strict determinism. Man's actions are as predetermined as all other happenings; they depend on his makeup and conditioning and on his reactions to stimuli from the external world. Crescas did not deny man's freedom only with regard to the domain of external action, for he pointed out that a man's beliefs and knowledge are not within his power.

ALBO. Whereas Crescas unmistakably regarded the Aristotelian philosophers as adversaries to be criticized or combated, the attitude of Joseph Albo (c. 1380–1444), who regarded Crescas as his teacher, is much less clearly defined. Albo did not eschew self-contradiction, apparently considering it a legitimate precaution on the part of a philosophical or theological author; indeed, he indulged in it in a much more obvious way than did Maimonides. But whereas the latter's fundamental philosophical position is fairly clear, the problem being how far he was prepared to deviate from Aristotelian doctrine in the interests of religion, there may be valid doubt whether Crescas and the Jewish religious tradition or Maimonides and Averroes were Albo's true masters. Mainly because of this perhaps deliberate failure to explain to the reader where he really stood, Albo has often been dismissed as an eclectic. He was strongly influenced not only by the authors just mentioned but also by Saadya. He seems to have had a considerable knowledge of Christian theology, even adopting for his own purposes certain scholastic doctrines. He differs from Crescas and to some extent resembles Maimonides in having had a marked interest in political theory.

The proclaimed theme of Albo's magnum opus, *Sefer ha-Ikkarim* (*The Book of Principles*), is the investigation of the theory of Jewish religious dogmas, whose number Maimonides, in a nonphilosophical work, had set at thirteen, whereas Albo, following a doctrine that in the last analysis seems to go back to Averroes, would limit them to three: existence of God, providence in reward and pun-

ishment, and the Torah as a divine revelation. One section, usually including the philosophical and the traditional religious interpretations side by side, is devoted to each of these dogmas. However, as far as Jewish philosophy is concerned, Albo's principal relatively novel (although in view of the likelihood of a Christian influence, probably not original) contribution to doctrinal evolution is the classification, in his introduction, of natural, conventional, and divine law. Natural law is necessary because man, being political by nature, must belong to a community, which may be restricted in size to one town or may extend over the whole earth. Natural law preserves society by promoting right and repressing injustice; thus, it restrains men from stealing, robbing, and murdering. The concept of "natural law" may have been taken over by Albo from the Christian Scholastics; the term is rarely used in philosophical works written in Arabic, and when it occurs it has an altogether different meaning. Albo did not mention whether natural law accords with human nature; he accounted for the need for and acceptance of natural law on purely utilitarian grounds. He did, however, believe that natural law is the same among all people, at all times, and in all places.

The positive laws instituted by wise men take into account the particular nature of the people for whose benefit they are instituted, as well as other circumstances. This means that they differ from the natural law in not being universally applicable. However, neither natural law nor the more elaborate conventional laws lead men toward true spiritual happiness; this is the function of divine laws instituted by a prophet, which teach men true theoretical opinions.

Contrary to Maimonides, but in agreement with the Scholastics and to some extent with Saadya, Albo believed that men are capable of establishing an orderly society by their own efforts, without the help of prophets.

Whereas Maimonides maintained that Judaism was the only divine law promulgated by a true prophet, Albo considered that the commandments given to Noah also constitute divine law, which ensures, although to a lesser degree than does Judaism, the happiness of its adherents. This position justifies a certain universalism; in accordance with a Talmudic saying, Albo believed that the pious among the non-Jews—that is, those who observe Noah's laws—have a share in the world to come. But he rejected the pretensions of Christianity and Islam to be divine laws.

RENAISSANCE. In the last few decades before their expulsion (1492), the Spanish Jews seem to have freely acknowledged the influence of the Christian Scholastics. A tribute to Christian thought was made not only by Habilla, a translator of several scholastic texts into Hebrew, but also by Isaac Arama and Isaac Abravanel, both of whom immigrated to Italy after the expulsion. Both are critical in various degrees of Aristotelians. Some of the views of Arama (who seems to have influenced Abravanel) mark a return to Yehuda Halevi. Abravanel's political doctrine is of some interest because it refers to and bestows praise on the regimes of the Italian republics of the period.

The son of Isaac Abravanel, Judah Abravanel, better known as Leone Ebreo (1460–c. 1521), was the author of *Dialoghi d'amore* and as such is one of the outstanding representatives of Platonism in Italy. He is perhaps the first example in postmedieval times of an important Jewish thinker who does not belong primarily to the history of Jewish philosophy (for the conception of Jewish philosophy presupposed in this assertion, see below).

Elijah del Medigo (c. 1460–1493), who was born in Crete, was a Jewish Averroist and a companion of Giovanni Pico della Mirandola. In his Hebrew treatise *Beḥinat Hadat* (The testing of religion) he opposed the trend among Jewish philosophers to read philosophical meanings into biblical texts by means of allegorical interpretations. Del Medigo, like Averroes, did not countenance any attempt to amalgamate religious law and philosophy. The Jewish philosophers who had such an amalgam in mind—it is pretty clear that Maimonides is the foremost object of these strictures—are neither (true) philosophers nor (true) professors of religious law.

MODERN AND CONTEMPORARY PERIODS

The expulsion of the Jews from Spain and Portugal produced a new center of Jewish thought, Holland, where many of the exiled Jews found a new and safer domicile; the tolerance of the regime seemed to provide guarantees against external persecution. This did not prevent, and indeed may have furthered, the establishment of an oppressive orthodoxy that was prepared to chastise rebellious members of its community.

DA COSTA. Both Uriel da Costa, or Acosta (1585–1640), and Spinoza (1632–1677) rebelled against Jewish orthodoxy. Uriel da Costa came to Amsterdam from Portugal, where, belonging to a family of Marranos (Jews who had converted to escape the Spanish Inquisition), he had been brought up in the Catholic faith; his philosophical position was to a great extent determined by his antagonism

to the orthodox Judaism (the Judaism of "the Pharisees," to use his term) that he encountered in Amsterdam. He was struck by the fact that the commandments as interpreted by his contemporaries did not conform to the text of the Torah, and he formulated a number of theses to prove his point. His growing estrangement from generally accepted Jewish doctrine is attested by his Portuguese treatise *Sobre a Mortalidade da Alma* (On the mortality of the soul). Apparently under the influence of Michael Servetus, he came to the conclusion that the soul is the vital spirit located in the blood and that it dies with the death of the body, there being no difference in this respect between the human and the animal soul. He considered that the belief in the immortality of the soul has had many evil effects, for it impels men to choose an ascetic way of life and even to seek death. According to him nothing has tormented men more than the belief in an eternal good and evil. God tolerates this opinion merely to torture the conscience of those who have abandoned his truth. At this stage da Costa affirmed the authority of the Bible from which, according to him, the mortality of the soul can be proved.

In his autobiography, written in Latin and titled *Exemplar Humanae Vitae* (An example of human life), he takes a more radical position. He proclaims the supreme excellency of the natural moral law (which, when arguing before the Jews, he seems to identify with the divine commandments to Noah—a comparison may be made with the view of Albo). Accordingly, he denies the validity of the argument that natural law is inferior to Judaism and Christianity, because he believes that both these religions teach the love of one's enemies, a precept which is not a part of natural law. According to da Costa, no good can come of demanding a manifest impossibility.

SPINOZA. Although medieval philosophers of Jewish origin for whom Judaism does not constitute a primary philosophical theme are thought of as belonging to the history of Jewish philosophy, a classification of this kind applied to such modern philosophers of Jewish origin as Salomon Maimon, Henri Bergson, Edmund Husserl, and L. I. Shestov might lead to some significant conclusions but would nevertheless seem inappropriate. It would certainly not be in keeping with the intentions of the philosophers themselves, and their views would be taken out of their natural contexts.

These considerations, however, may not, for the following reasons, be quite so valid with respect to Spinoza: (1) It was through the study of Jewish philosophical texts that Spinoza was first initiated into philosophy. (2) It may be argued with some reason that at least in part (if one abstracts the influence of René Descartes and of seventeenth-century physics and certain other constitutive elements), Spinoza's system is a radicalization or perhaps a logical corollary to medieval doctrines; although its importance may be contested, the impact of Maimonides and of Crescas is evident. (3) A considerable portion of Spinoza's *Tractatus Theologico-politicus* deals with problems related to Judaism. Reference to some of the views set forth in the *Tractatus* and to their connection with medieval Jewish doctrines may not be out of place here.

Prophecy. As the first chapters of the *Tractatus* show, the doctrine of prophecy is of central importance to Spinoza's explanation of Judaism. These chapters can also provide proof that, as far as this subject is concerned, Spinoza to a large extent used Maimonides' categories, although he applied them to different people or groups of people. In fact, the relationship of Spinoza to Maimonides—although antagonistic—is much closer than that of most of the fifteenth-century Jewish philosophers who did not break with Judaism.

Maimonides held that the prophets combined intellectual perfection, which made them philosophers, with perfection of the imaginative faculty. He also referred to a category of people endowed with a strong imagination but possessing no extraordinary intellectual gifts; this category includes, for example, lawgivers and statesmen. Spinoza took over this last category but applied it to the prophets, whom he described as possessing vivid imaginations but as not necessarily having outstanding intellectual capacities. He denied that the biblical prophets were philosophers and used a philosophical and historical approach to the Scriptures to show that the contrary assertion is not borne out by the texts.

Spinoza also denied Maimonides' assertion that the prophecy of Moses was essentially different from that of the other prophets and that this was largely because Moses, in prophesying, had no recourse to the imaginative faculty. According to Spinoza the distinctive fact about Moses' prophecy was that he heard the voice of God in a prophetic vision—that is, in a state in which his imagination was active. In this assertion Spinoza employed one of Maimonides' categories of prophecy, differentiated in the *Guide* according to certain characteristics of prophetic dreams and visions. However, Maimonides thought it improbable that the voice of God was ever heard in prophetic vision; he held that this category is purely hypothetical. It seems evident that in his classification of Moses, Spinoza was concerned not with what

really happened in history but with pigeonholing the evidence culled from the Bible into Maimonides' theoretical framework in such a way that it fit in with his theologicopolitical purpose.

This purpose made it imperative to propound in the *Tractatus* a theory concerning Jesus, whom Spinoza designates as Christus. The category and the status assigned to Jesus are by and large similar to those that Maimonides attributed to Moses. Thus, Jesus is referred to in the *Tractatus* as a religious teacher who makes recourse not to the imaginative faculty but solely to the intellect. However, in following up this hypothesis Spinoza was guilty of an inconsistency. Whereas in the case of the Old Testament prophets he rejected allegorical interpretations predicated on the supposition that the prophets adapted their discourses to the understanding of the general public, in the case of Jesus he adopted this interpretation because he wished to explain away those of Jesus' sayings that he regarded as incompatible with true philosophical doctrine. Both Maimonides' Moses and Spinoza's Jesus are absolutely unique personalities; there is, however, an important difference between them. In the opinion both of Maimonides and of Spinoza, Moses' legislation created the Jewish community and state, whereas Jesus as conceived by Spinoza was not a lawgiver and, as far as his direct activity was concerned, not a statesman, though within Spinoza's blueprint for an ideal State, he is assigned a political function: His authority may be used to institute and strengthen the religion Spinoza called *religio catholica,* which has little or nothing in common with any of the major manifestations of historic Christianity.

Critique of Judaism. The difference between Judaism and Spinoza's *religio catholica* corresponds to the difference between Moses and Jesus. After leaving Egypt the Jews found themselves, in Spinoza's view, in the position of people who had no allegiance to any positive law; they had, as it were, reverted to a state of nature and were faced with the need to enter into a social pact. They were also an ignorant people and very prone to superstition. Moses, a man of outstanding ability, made use of the situation and characteristics of the people in order to make them accept a social pact and a state founded upon it, which, contrary to Spinoza's schema for his ideal communities, were not based first and foremost upon utilitarian—that is, reasonable—consideration of the advantages of life in society over the state of nature.

The social pact concluded by the children of Israel in the desert was based upon a superstitious view of God as "King" and "Judge," to whom the children of Israel owed whatever political and military successes they obtained. It was to God rather than to the representatives of the popular will that the children of Israel transferred political sovereignty. In due course political sovereignty was vested in Moses, God's representative, and in his successors. It should be added that in spite of Spinoza's insistence on the superstitious foundations of the state of the children of Israel in ancient times, his account of its regime was not wholly unsympathetic. He did, however, believe that it contained the seeds of its own destruction and that with the extinction of this state the social pact devised by Moses had lapsed and all the political and religious obligations incumbent upon the Jews become null and void.

"Religio catholica." It could be argued that because the state conceived by Spinoza is based not on superstitious faith but on a social contract originating in rational, utilitarian considerations it does not necessarily need to have its authority safeguarded and stabilized by means of religion. However, Spinoza appears to have held the view—perhaps derived from a purely empirical knowledge of the behavior of the common run of men—that there is a need for religion. In order to fulfill the need for some religion and to obviate the danger of harmful religions, he devised his *religio catholica,* the universal religion, which has the following distinctive traits: (1) Its main purpose, a practical one (which is furthered by recourse to the authority of Jesus), is to impel men to act in accordance with justice and charity. Such conduct is tantamount to obedience to the laws of the state and to the orders of the magistrates, in whom sovereignty is vested. For disobedience—even if it springs from compassionate motives—weakens the social pact, which safeguards the welfare of all the members of the community; in consequence, its evil effects outweigh whatever good it may produce. (2) Although religion, according to Spinoza, is not concerned with theoretical truth, in order to be effective the *religio catholica* requires dogmas, which he set forth in the *Tractatus.* These dogmas are formulated there in terms that can be interpreted in accordance both with the philosophical conception of God that Spinoza regarded as true and with the superstitious ideas of ordinary people. It follows that if they are accepted as constituting by themselves the only creed that everybody is obliged to profess, people cannot be persecuted on account of their beliefs; Spinoza held that such a persecution is liable to lead to civil war and may thus destroy the state. Philosophers are free to engage in the pursuit of truth and to attain, if they can, the supreme goal of man, freedom grounded in knowledge. There can be little doubt that the furtherance of the cause of tolerance for

philosophical opinions was one of Spinoza's main objects in writing the *Tractatus.*

MENDELSSOHN. Moses Mendelssohn (1729–1786) opens what may be called the German period of Jewish philosophy. This period, in which a considerable number of works on Jewish philosophy were written in German and often under the influence of German philosophy, is also marked by the emancipation of the Jews (that is, by the abrogation of discriminatory laws directed against them) and by their partial or complete assimilation. In this period in particular, it appears indicated to apply the term "Jewish philosophy" first and foremost to works whose main purpose or one of whose main purposes consists in proposing a definition of Judaism and a justification of its existence. The second task is often conceived as necessitating a confrontation of Judaism with Christianity rather than with philosophy, which served as a point of comparison for many medieval philosophers. This change seems to have been a result of the demarcation of the sphere of religion in such a way that, at least in the opinion of the philosophers, possible points of collision no longer existed between it and philosophy.

This demarcation was largely furthered by the doctrine of Spinoza—from whom Mendelssohn and others took over and adopted for their own purposes certain fundamental ideas concerning Judaism. Like Spinoza, Mendelssohn held (according to his treatise *Jerusalem* and other writings) that it is not the task of Judaism to teach rational truths, although they may be referred to in the Bible. Contrary to what he called Athanasian Christianity (that is, the doctrine set forth in the Athanasian creed), Judaism has no binding dogmas; it is centered on inculcating belief in certain historical events and on action—that is, observance of religious law (including the ceremonial commandments). Such observance is supposed to lead to happiness in this world and in the afterlife. Mendelssohn did not reject this view offhand, as Spinoza would have done; indeed, he seems to have been prepared to accept it—God's mysteries being inscrutable, and the radicalism and what may be called the consistency of Spinoza being the complete antithesis of Mendelssohn's apologetics. Non-Jews were supposed by Mendelssohn to owe allegiance to the law of nature. He did not affirm the superiority of Judaism over this law and was prepared to regard Jesus as a great prophet. He declared his belief that the differences between the various religions are not eternal and that when the whole earth is united in the knowledge of God, the Jews will be permitted to abandon their peculiar rites and ceremonies. But that time has not yet arrived.

Mendelssohn was well grounded in medieval Jewish philosophy and referred quite frequently to the writings of Maimonides and of other Jewish thinkers of the Middle Ages. The three principles on which he held Judaism to be based call to mind those propounded by Albo. They are God, providence, and the divine law.

FORMSTECHER. Whereas Mendelssohn continued the medieval tradition, at least to some extent, or adopted Spinoza's doctrine for his purposes, the Jewish philosophers of the first half of the nineteenth century (except, at least in a certain measure, Solomon Steinheim) may be regarded as disciples of the philosophers of their own time. In *Die Religion des Geistes* (The religion of the spirit), Solomon Formstecher (1808–1889) was greatly influenced by Friedrich Schelling in his conception of nature and spirit as manifestations of the divine. There are types of religions that correspond to these manifestations: (1) the religion of nature in which God is conceived as the principle of nature or as the world soul, and (2) the religion of the Spirit which conceives of God as an ethical being. According to the religion of the Spirit, God has produced the world as his manifestation in full freedom and not, as the religion of nature tends to profess, because the world was necessary for his own existence.

The religion of the Spirit, which corresponds to absolute religious truth, was first manifested in the Jewish people. The religious history of the world may be understood as a process of universalization of the Jewish religion. Thus, Christianity propagated Jewish conceptions among the nations; however, it combined them with pagan ideas. The pagan element is gradually being eliminated—Protestantism, for instance, in this respect marks considerable progress. When at long last the Jewish element in Christianity is victorious, the Jews will be right to give up their isolation. The process that will bring about this final religious union is already under way.

HIRSCH. The main philosophical work of Samuel Hirsch (1815–1889), titled *Die Religionsphilosophie der Juden* (The philosophy of religion of the Jews) was decisively influenced by G. W. F. Hegel. This influence is most evident in Hirsch's method and in the task that he assigned to the philosophy of religion—the transformation of religious consciousness into conceptual truth. However, contrary to Hegel, he did not consider religious truth to be inadequate as compared to philosophical truth.

Hirsch believed that man's awareness of himself as an ego is identical with his awareness of his freedom. This freedom is, however, abstract; it became concrete in the

various historical religions. Man may renounce this freedom and believe that he is dominated by his senses. This means recognizing the absolute sovereignty of nature regarded as a divine principle, which is the point of departure of the pagan passive religions. If, however, he subordinates his nature to his freedom, his freedom becomes concrete. God is conceived not only as the giver of abstract freedom but as willing man's concrete freedom. This is the principle of Judaism. Christianity, was conceived by Hirsch as it was by Formstecher, as being intermediate between Judaism and paganism.

God revealed himself in the first stages of Jewish history by means of miracles and of prophecy. At present he manifests himself in the miracle that is constituted by the existence of the Jewish people. At its beginning in the time of Jesus, Christianity was identical with Judaism. The decisive break between the two religions was caused by Paul. According to Hirsch, when the Pauline elements are eliminated from Christianity, it will be in all essentials in agreement with Judaism, which, however, will preserve its separate existence.

KROCHMAL. Nachman Krochmal (1785–1840), a native of Galicia (at that time part of Austria), was the author of *Moreh Nebukhei ha-Zman* (Guide of the perplexed for our time), a treatise in Hebrew on philosophy of history and on Jewish history, which had a considerable influence.

Krochmal, like Hirsch, was influenced by Hegel and perhaps also by other German philosophers approximately of Hegel's period, such as Johann Gottlieb Fichte. Krochmal's philosophical thought was centered on the notion of "spirit," Krochmal being mainly concerned with the "national spirit," the particular "spirit" that is proper to each people and that accounts for the peculiar characteristics differentiating one people from another in every domain of human activity.

The national "spirits" of all peoples except the Jewish are, according to Krochmal, essentially particular. Hence, the national spirit either becomes extinct with the extinction of the nation or, if it is a powerful spirit, is assimilated by some other nation. The Jewish people has a special relation to the Universal Spirit, who is the God of Israel. This relation accounts for the perpetuity of the Jewish people.

STEINHEIM. Solomon Ludwig Steinheim (1789–1866), the author of *Die Offenbarung nach dem Lehrbegriff der Synagoge,* (Revelation according to the doctrine of the synagogue), was apparently influenced by the antirationalism of Friedrich Jacobi.

His criticism of science is based on Jacobi's criticism, but he did not agree with Jacobi in opposing discursive reason to our intuitive knowledge of God—Steinheim contrasted human reason to divine revelation. The main point on which the revelation vouchsafed to the prophets of Israel is opposed to reason is to be found in the fact that the God posited by reason is subject to necessity, that he can act only in accordance with laws. Moreover, reason affirms that nothing can come from nothing. Accordingly, God is free to create not a good world but only the best possible world. Revealed religion, on the other hand, affirms the freedom of God and the creation of the world out of nothing.

COHEN. There seems to be little connection between the Jewish philosophers of the first half or two-thirds of the nineteenth century and Hermann Cohen (1842–1918), the head of the Marburg Neo-Kantian school. In a certain sense Cohen may be regarded as a rather unusual case among the philosophers of Judaism of his and the preceding generations, because of the two aspects of his philosophical thought—the general and the Jewish—and the uneasy equilibrium between them. Judaism was by no means the only important theme of his philosophical system; it was one of several and not even his point of departure. There is no doubt that for most of his life Cohen was wholly committed to his brand of Kantianism, in the elaboration of which he displayed considerable originality—it has been maintained with some justification that his doctrine manifests a certain (unintentional) kinship with Hegel's. However, Cohen's idea of God derives from an analysis and a development of certain conceptions of Immanuel Kant.

In Cohen's view, reason requires that nature be conceived of as conforming to *one* rational plan and that harmony exist between the domains of natural and of moral teleology. These two requirements in turn necessitate the adoption of the idea of God—the word *idea* being used in the Kantian sense, which means that no assertion is made about the metaphysical reality of God. Cohen's theory of ethics stemmed to a considerable extent from Kant's, but he held that the most important ethical principles were discovered by the prophets of Israel, who freed religion from its entanglement with mythology. A harmony also exists between the Messianic notion of the Jewish prophets and the exigency of ethics that the task of coming ever closer to moral perfection be pursued unceas-

ingly. This goal will never be wholly attained. Messianism, too, is an idea in the Kantian sense of the word.

Cohen seems to have changed his attitude in the last years of his life; at least, although he did not explicitly renounce his previous positions, a considerable shift of emphasis can be discerned in his doctrines. The notion of the human individual—an individual who is weak and full of sin—comes to the fore, as well as the conception of a *correlation,* a relationship between God and the individual. This relationship is one of love, the love of God for man and the love of man for God. It is difficult to reconcile the conception of God expounded in Cohen's works of his last period with his Kantian or Neo-Kantian attitude toward metaphysics.

The conceptions of God and the individual and cognate conceptions are set forth in Cohen's posthumously published book *Die Religion der Vernunft aus den Quellen des Judentums* (Religion of reason from the sources of Judaism) and in a series of articles reprinted in his *Jüdische Schriften.*

ROSENZWEIG. Franz Rosenzweig (1886–1929) published his main philosophical work, *Der Stern der Erlösung* (The star of redemption), in 1921. This work begins with a rejection of the traditional philosophical attitude that denies the fear of death, maintaining, instead, that this fear is the beginning of the cognition of the All. Man should continue to fear death, despite the indifference of philosophy and its predilection for accepting death. Traditional philosophy is interested exclusively in the universal, and it is monistic—its aim is to discover one principle from which everything can be derived. However, this tendency of philosophy denatures human experience, which knows not one but three separate domains (which Kant had referred to in a different context), namely, God, the world, and man.

According to Rosenzweig, God (like the world and like man) is known through experience (the experience of revelation). In Greek paganism, the most perfect manifestation of paganism in general, every one of these domains subsists by itself: the gods, the cosmos, and man as the tragic, solitary, silent hero. The biblical religion is concerned with the relation between the three: the relation between God and the world, which is creation; the relation between God and man, which is revelation; and the relation between man and the world, which leads to salvation. The philosophy that renounces the ambition to find one principle for everything that exists and that follows biblical religion in centering on the connections between the three domains and between the words and acts that bring about and develop these connections Rosenzweig termed the "narrative" philosophy; the term and the concept were taken over from Schelling, whose influence Rosenzweig repeatedly emphasized.

The biblical faith brought forth two valid religions, Christianity and Judaism. The first is described by Rosenzweig as the eternal way: The Christian peoples seek in the vicissitudes of time and history the way to salvation. In contradistinction to them the existence of the stateless Jewish people is not concerned with time and history; it is—notwithstanding the hope for final salvation—*already* an eternal life, renewed again and again according to the rhythm of the liturgical Jewish year. Thus, Rosenzweig did not, like Yehuda Halevi (many of whose poems he had translated and who was very much in his thoughts) oppose the sacred history of the Jewish people to the profane history of the rest of the world but rather to what he considered as the historical existence of the Jews, their involvement in the history of the other nations.

BUBER. Since the early years of the twentieth century, Martin Buber (1878–1965) has exercised a powerful influence on both Jews and non-Jews. His theology, centered on the I and Thou relationship, on the conception of a dialogical life, and on the primal importance of the category of "encounter" are discussed in the entry on Buber.

In recent years new works dealing with the history of Jewish thought in one of its aspects or in one of its periods appear on the whole to have been more significant than purely philosophical or purely theological Jewish works; in certain cases scholarly works give expression to a personal attitude toward Judaism or toward religion in general. These remarks apply to the two main centers of Jewish philosophical, theological, and scholarly activities, the United States and Israel, as well as to such other countries as France and England. However, it may be too early to attempt to give a definitive summing-up of the tendencies and achievements of a period that verges upon the present.

See also Albo, Joseph; al-Ghazālī, Muhammad; al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥā; Aristotelianism; Averroes; Avicenna; Baḥya ben Joseph ibn Paquda; Bergson, Henri; Buber, Martin; Cohen, Hermann; Costa, Uriel da; Crescas, Hasdai; Descartes, René; Duns Scotus, John; Enlightenment, Jewish; Epicurus; Galileo Galilei; Gersonides; Halevi, Yehuda; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; Ibn Gabirol, Solomon ben Judah; Ibn Zaddik, Joseph ben Jacob; Israeli, Isaac ben

Solomon; Jacobi, Friedrich Heinrich; Jewish Averroism; Kabbalah; Kant, Immanuel; Logic, Traditional; Maimon, Salomon; Maimonides; Mendelssohn, Moses; Naṣīr al-Dīn al-Ṭūsī; Neo-Kantianism; Philo Judaeus; Pico della Mirandola, Count Giovanni; Renaissance; Rosenzweig, Franz; Saadya; Servetus, Michael; Shestov, Lev Isaakovich; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.

Bibliography

GENERAL WORKS

Guttman, Julius. *Philosophies of Judaism.* New York: Holt, Rinehart and Winston, 1964. Contains extensive bibliography.

Husik, Isaac. *A History of Mediaeval Jewish Philosophy.* Philadelphia: Macmillan, 1916.

Munk, Salomon. *Mélanges de philosophie juive et arabe.* Paris: J. Gamber, 1859.

Strauss, Leo. *Persecution and the Art of Writing.* Glencoe, IL: Free Press, 1952.

Strauss, Leo. *Philosophie und Gesetz.* Berlin: Schocken, 1935.

Vajda, G. *Jüdische Philosophie.* Bern: A. Francke, 1950.

TEXTS

For the works of individual philosophers other than those listed here, see the bibliographies to the articles devoted to them.

Abravanel, Isaac. *Perush l'Moreh Nebukhim.* Vilna, Poland, 1904.

Bar Hiyya, Abraham. *Hegyon ha-Nefesh o Sefer ha-Musar.* Edited by Y. I. Freiman. Leipzig, 1860.

Bar Hiyya, Abraham. *Megillath ha-Megalleh.* Edited by A. S. Poznanski. Berlin, 1924.

Cohen, Hermann. *Die Religion der Vernunft aus den Quellen des Judentums.* Leipzig, 1919; 2nd ed., Frankfurt am Main: J. Kauffmann, 1929.

Cohen, Hermann. *Jüdische Schriften.* Edited by B. Strauss, 3 vols. Berlin: C.A. Schwetschke, 1924. Introduction by Franz Rosenzweig.

Del Medigo, Elijah. *Sefer Behinat ha-Dat.* Vienna, 1833. Contains commentary and notes by I. S. Reggio.

Ebreo, Leone. *Dialoghi d'amore.* Edited by Carl Gebhardt. Heidelberg: C. Winter, 1929.

Formstecher, Solomon. *Die Religion des Geistes.* Frankfurt am Main: J.C. Hermann, 1841.

Heller-Vilensky, S. *Yitzhaq ʿArama ve-Mishnato.* Jerusalem and Tel Aviv, 1956.

Hirsch, Samuel. *Das System der religiösen Anschauung der Juden und sein Verhältnis zum Heidentum, Christentum und zur absoluten Philosophie. Vol. I, Die Religionsphilosophie der Juden.* Leipzig, 1842.

Ibn Da'ud, Abraham. *Sefer ha-Emunah ha-Ramah.* Edited with a German translation by Samson Weill. Frankfurt, 1852.

Ibn Ezra, Abraham. *Yesod Morah v'Sod Torah.* Edited by Samuel Waxman. Jerusalem, 1931.

Kaspi, Joseph. *Amuday Kesef u-Maskiyot Kesef.* Edited by S. Werbluner. Frankfurt, 1848. Contains two commentaries on Maimonides' *Guide.*

Krochmal, Nachman. *Writings.* Edited by Simon Rawidowicz. Waltham, MA, 1961.

Narboni, Moses. *Commentary to the Guide of the Perplexed.* Edited by Goldenthal. Vienna, 1852.

Pines, Shlomo. *Nouvelles Études sur Abu'l Barakāt al-Baghdādī.* Paris, 1955.

Saadya ben Joseph. *Polemic against Hiwi al-Balkhi.* Edited by Israel Davidson. New York, 1915.

Steinheim, Solomon. *Die Offenbarung nach dem Lehrbegriff der Synagoge.* 4 vols. Frankfurt am Main: S. Schmerber, 1835–1865.

Strauss, Leo. *Die Religionskritik Spinozas als Grundlage seiner Bibelwissenschaft.* Berlin: Akademie-Verlag, 1930.

Vajda, G. *Isaac Albalag, averroïste juif.* Paris: J. Vrin, 1960.

Shlomo Pines (1967)

JEWISH PHILOSOPHY [ADDENDUM]

Levinas's earlier work seemed to have nothing to do with religion, whereas after the Holocaust his writings often touched on religious and specifically Jewish themes. These latter works are often described as his translations of Judaism into Greek, his explanations of the universal significance of the Torah and all that goes with it in the living experience of the Jewish people. The central problem here is that philosophy is a universal and theoretical enterprise, whereas religion is often practical and its obligations are limited to a particular group of people.

It might be thought even harder to identify Judaism with philosophy because most Jews are Jews because of birth, not through a choice to accept particular doctrines or practices. As human beings we all share certain characteristics, and as members of smaller and more limited groups we share features with others in the same group that we do not share with everyone else. The universal role of philosophy is stressed throughout Levinas's thought. He starts by replacing the usual starting point in metaphysics of ontology with ethics. The obvious place to start philosophically is to ask the question: What is a human being? Once we know the answer to that question, which includes other questions such as what can we know, we can then start to work out our duties and responsibilities for others.

Yet Levinas turns this starting point entirely the other way around. The first question we need to ask is not who we are, but what our responsibilities to others are. Only once we can answer the question who is the other are we in a position to know who we are. Again, the normal way to start constructing a metaphysics is to build up

a model from the self to the other, and then suggest that I have certain responsibilities for others, as they have for me. But these responsibilities for others are limited, because it is unreasonable for them to be otherwise. Levinas transforms this to argue that our responsibilities for others are not limited, but are infinite. The idea is that if one starts with the self and then moves to the concept of the other one is always in the position of trying to demarcate between the roles of different selves in ways that limit, perhaps too severely, the links between these selves. In saying that we have infinite responsibility for the other, Levinas means that we must always be able to respond to the other, because unless we do, we lose our identity as a subject. As he says: "The word *I* means *here I am*, having to answer for everything and everyone" (1981 [1974], p. 145).

This radical turn in philosophy has interesting implications for all the familiar philosophical problems, but one implication that one might assume it would have is to direct attention away from the individual and toward groups of individuals, or indeed the whole of humanity, because it is to the whole of humanity for which the self is answerable. This implication is certainly there, and Levinas seems to have a horror of the solitary life, in marked contrast to his mentor Heidegger who appears to value above all retreating from the world into oneself and being alone with nature. For Levinas the more one tries to retreat into oneself, the less self there is to retreat to. Yet there is a problem in accounting for the treatment of groups larger than the individual self and less than the whole of humanity. If one gives one's attention primarily to a limited group, is this not to take it away from the whole of humanity, and thus deny one's full selfhood? Hence on the one hand Levinas's criticisms of love between people, which he sometimes sees as limiting the links of responsibility that we establish between each other. On the other hand, the way in which love can be embodied in an institution that gives rise both to the possibility of new life and also to new structured links between individuals is positive in the sense that it cements and furthers the ethical links between people. We need to point out here an important feature of such links, because for Levinas they are not reciprocal. We have responsibilities for the other, but they have none for us. The idea is that our responsibility is not based on their appropriate behavior or response; rather, we are responsible, and never stop being responsible however much others ignore, deny, or are unworthy of our concern. Our responsibility even extends further than our lives, into the infinite.

Here we have a system of philosophy that came to be greatly elaborated, and that appears to apply to every human being, if it applies to anyone. What is Jewish about it? Here we get into difficult territory. Judaism is profoundly ethical; it emphasizes the practical as against the theoretical, the group as compared to the individual; it compares speech with prophecy, has a messianic future in prospect for humanity, and is generally suspicious of the capacity of the state to embody ethical life appropriately. In his Jewish writings Levinas examines a wide variety of biblical and other literature that he interprets as bringing out some of the key aspects of his philosophy. What Levinas is doing in these works is using his philosophy to bring out what he thinks are the meanings of Jewish texts, and in this he is following in a long and distinguished tradition in Jewish philosophy.

Levinas is interested in those parts of the Bible that refer to the individual as responsible for others. When God asks Cain "Where is Abel your brother?" (Gen. 4:8) the response that Cain is not his brother's keeper is entirely inappropriate. Nor is the question only directed at Cain, but at everyone. Everyone is responsible for those who are murdered even if they are not directly involved in the crime. Levinas picks out those parts of the Torah in which individuals respond to God by saying "*hineni*" or "I am here." What is meant by that is that the individual recognizes the claim of God to be heard, he accepts that he is the person to whom the divine question is to be put and acknowledges responsibility for the task that God may have in mind. But is this not just the individual responding to God? What is specifically ethical about it? For Levinas references to God are not references only to a being. The point of religion is to bring out to the individual the significance of her links with others, so that when she is involved with others God is available to her, but when she is thinking only of how to come close to God, he is distant. There are plenty of passages in the Torah and other Jewish writings that emphasize the presence of God as contingent on a certain way of human acting. Buber, for example, was impressed with Hasidic stories of ordinary actions being imbued with spirituality when those actions were carried out for the sake of others.

Is Judaism then specifically directed toward practice, toward the ethical? It certainly looks like it, and its legal structure, the integration of the life of the Jew within a ritualistic system, is evidence of the significance of behavior, in particular behavior that is linked with the behavior of others. Levinas describes the move from the particularity of Judaism to the universalism of philosophy as part of a process of liberation that needs to be continued.

This means that we have not yet finished translating the Bible, in the sense that we have only just started bringing out its universalist message. We tend to get it the wrong way around, he suggests, in that we often try to invest a biblical idea with a universalist notion already specified in European or Greek thought, whereas what we should do is bring out the universalist idea that is already there in the Hebrew, and explain it in Greek. So the apparent clash between Athens and Jerusalem, between reason and revelation, comes out as a pseudo-conflict after all, because revelation contains everything that reason does, albeit in a different form.

This is also a familiar solution to the apparent conflict between the Torah and philosophical thought within the tradition of Jewish philosophy. Levinas is thus moving through familiar territory here, although the ways in which he characterizes his theoretical environment differs from his predecessors. One difference lies in the sort of philosophy that Levinas argues runs through the Bible and Talmud, and that is of course his philosophy in a much wider sense than one would expect given the above definition.

When we examine the detailed defense of *halakhah* (Jewish law) that Levinas produces, a number of questions arise. He tends not to defend the practices of halakhah as a whole, but the principle of there being a halakhah. We are supposed to accept that once the principle of there being a halakhah is understood, the details of what to do will be in accordance with the legal texts. Perhaps it does not matter so much exactly which understanding of the precise nature of the law one follows, so long as one adheres to an appropriate legal rule that determines the nature of the law. The general point that Levinas makes is certainly appropriate, that Jews seeking to find a way of embodying their ethical behavior in a practice need look no further than halakhah, which is a form of practice specifically designed to replicate the truths of faith. Yet because Levinas often says that when he talks about Jews he means humanity at large, we are still left not knowing how most people ought to behave. He relishes the prospect of Jews being obliged to follow 613 commandments, whereas the rest of humanity only needs follow the seven Noahide laws (the basic laws of humanity by tradition given by God to Noah), because this exemplifies the idea of owing more to the other than one asks in return. But is the only route to such an ethical life for the Jew the halakhic route? It is one thing to defend the acceptability of such a route, but another to demonstrate its inevitability. Because Levinas is firmly part of the demythological tendency in modern theology, he cannot argue that one should follow halakhah because God had commanded us to act thus in the Torah. Such appeals to what Levinas rather scathingly calls the numinous are ruled out from the start.

Then we have the problem of knowing how gentiles are to live. Are they to follow the laws of their faith? Are they to consider conversion to Judaism? Are they limited to the seven Noahide laws? They too may wish to be in a position to say *me voici* even if they cannot say *hineni*. Their route to reaching that position is presumably through Greek, because they cannot go through Hebrew, but if they can attain their end along the Greek route, there seems no reason to deny such a route to assimilated Jews. It would be invidious to suggest that a particular ethnic group is especially advantaged in knowing how to live as compared with others, unless one can appeal to some supernatural rationale, which Levinas rejects. Even if my ancestors opened themselves to God in a way different from other ethnic groups, it is difficult to see how that elevates my consciousness above those of members of other ethnic groups.

In a celebrated discussion Levinas argues that what made the Jews unique was that they undertook to obey God's law even before they heard what that law was, and in this way acknowledged the priority of the ethical over the ontological. That is, they accepted that the first question to be answered is where our responsibilities lie, and the second question follows from that, and concerns who we are. But what is the link between that event and the nature of halakhah as it has come down to us today? There is obviously some sort of link in terms of tradition, yet it is the case that for many Jews the ties of tradition no longer have any emotional resonance. The only way for them to find that resonance again is through the Greek, and that is why Levinas seeks to reinvigorate tradition by exploring its universal values in language that is accessible to those who have lost the ability to understand the Hebrew. When he talks of them no longer understanding Hebrew he does not refer to technical mastery of the language, but the ability to link the Hebrew to present-day ethical and political issues. What needs to be brought out is the universalist message of the Hebrew, so that it is not seen as the repository of a small and remote community, but as implicit within the rules of behavior everywhere and at all times.

Bibliography

Abravanel, Isaac. *Principles of Faith*. Translated by M. Kellner. Oxford: Littmann Library, 1982.

Buber, Martin. *Mamre: Essays in Religion*. Translated by G. Hort. Melbourne: Melbourne University Press, 1946.

Cohen, Hermann. *Religion of Reason out of the Sources of Judaism*. Translated by S. Kaplan. Atlanta: Scholars Press, 1995.

Fackenheim, Emil. *The Jewish Bible after the Holocaust: A Re-Reading*. Manchester, U.K.: Manchester University Press, 1990.

Frank, Daniel, and Oliver Leaman, eds. *History of Jewish Philosophy*. London: Routledge, 1997.

Frank, Daniel, Oliver Leaman, and Charles Manekin, eds. *The Jewish Philosophy Reader*. London: Routledge, 2000.

Leaman, Oliver. *Evil and Suffering in Jewish Philosophy*. Cambridge, U.K.: Cambridge University Press, 1995.

Leaman, Oliver. "The Future of Jewish Philosophy." In *The History of Jewish Philosophy*, edited by D. Frank and Oliver Leaman. London: Routledge, 1997.

Leaman, Oliver. *Moses Maimonides*. Richmond: Curzon, 1997.

Leaman, Oliver. "The Philosophy of Religion." In *The Future of Philosophy*, edited by Leaman. London: Routledge, 1998.

Levinas, Emmanuel. *Autrement que d'être*. The Hague: Nijhoff, 1974. Translated by A Lingis as *Otherwise than Being or Beyond Essence* (The Hague: Nijhoff, 1981).

Levinas, Emmanuel. *Beyond the Verse: Talmudic Readings and Lectures*. Translated by G. Mole. London: Athlone, 1994.

Levinas, Emmanuel. *Difficult Freedom: Essays on Judaism*. Translated by S. Hand. London: Athlone, 1990.

Levinas, Emmanuel. *Nine Talmudic Readings*. Translated by A. Aronowicz. Bloomington: Indiana University Press, 1990.

Maimonides, Moses. *Guide for the Perplexed*. Translated by S. Pines. Chicago: University of Chicago Press, 1963.

Mendelssohn, Moses. *Jerusalem*. Translated by A. Arkush. Hanover, NH: Brandeis University Press, 1983.

Philo. *On the Creation of the World*. Vol. 1 of *Works: English and Greek*. Translated by F. Colson and G.Whitaker. Cambridge, MA: Harvard University Press, 1929.

Rosenzweig, Franz. *The Star of Redemption*. Translated by W. Hallo. New York: Holt, Rinehart and Winston, 1970.

Rubenstein, Richard. *After Auschwitz: Radical Theology and Contemporary Judaism*. New York: Bobbs-Merrill, 1992.

Saadya Gaon. *The Book of Doctrines and Beliefs*. Translated by A. Altmann. Philadelphia: Jewish Publications Society, 1960.

Spinoza, Baruch. *A Spinoza Reader: The Ethics and Other Works*. Translated by E. Curley. Princeton, NJ: Princeton University Press, 1994.

Oliver Leaman (2005)

JINUL

(1158–1210)

Jinul (whose name is spelled Chinul under the McCune Reischauer Romanization system), a Korean Buddhist monk of the Goryeo period, is considered by many scholars to be the most influential figure in the formation of Korean Seon Buddhism. Deeply disturbed at the degree of corruption that had crept into the Buddhist monastic system, he sought to establish a new movement that he called the *samādhi* and *prajñā* society. The goal of this organization was the establishment of a new community of disciplined, pure-minded practitioners deep in the mountains. Jinul eventually brought this mission to fruition with the founding of the Seonggwangsa monastery at Mount Jogye, which still serves as an center for Korean Seon practice.

A major issue that received special attention from Jinul was the relationship between so-called gradual and sudden approaches to Buddhist practice and enlightenment. Drawing on various Chinese treatments of this topic, most importantly those established by Zongmi (780–841) and Dahui (1089–1163), Jinul came up with his "sudden enlightenment followed by gradual practice" approach. Jinul believed that for religious practice—especially meditative practice—to have efficacy, the practitioner must first have a deep and transformative experience of insight into the emptiness of things, to see their nature of innate enlightenment. He believed that if one tries to practice without such an experience, all of one's practice will be based on the dualistic thinking habits that are the causes of delusion, and thus, no matter how hard one might try, progress cannot be made. One metaphor that Jinul used to express this idea was that of the morning dew and the sunshine. Before the sun rises, the cool morning grass is wet with dew. Try as one may to wipe away the dew, it will continue to reappear. Once the sun rises, however, the dew can be wiped away and will be less apt to return. In the same way, once one has had an awakening experience, efforts toward the eradication of bad cognitive and emotive habits will have enhanced efficacy.

Jinul's approach to Buddhist practice ended up becoming an interesting blend of *gongan* (in Japanese *kōan*) meditation, coupled with scriptural study, incorporating a Hwaeom (in Chinese *Huayan*) approach that tended to see the mutual containment of ostensive opposites. While incorporating the *gongan* method into his system of practice, Jinul also believed that scriptural study was a vitally important component of Buddhist cultivation. This approach is enunciated in the oft-repeated story that Jinul did not undergo his enlightenment experiences as the result of the classical so-called personal mind-to-mind transmission between teacher and student as characterized in the Seon school. Rather, each of his three enlightenment experiences came in connection with the contemplation of a passage in a Buddhist text.

Jinul's philosophical resolution of this issue brought a deep and lasting impact on Korean Buddhism and can be seen as a repeated theme in the works of many subsequent Seon masters, including such famous figures as Gihwa (1376–1433) and Hyujeong (1520–1604), who followed Jinul's way of thinking in addressing the issue of practice and study in their own writings. Jinul produced a number of important disciples who passed on his teaching and continued to work within his discourse.

See also Buddhism—Schools: Chan and Zen.

Bibliography

Buswell, Robert E., Jr., tr. *Tracing Back the Radiance: Chinul's Korean Way of Zen*. Honolulu: University of Hawaii Press, 1991.

Keel, Hee-Sung. *Chinul: The Founder of the Korean Son Tradition*. Seoul, South Korea: Pojinjae, 1984.

Ko, Ik-chin. "Chinul's Explanation of Emptiness in the Meditation School." In *Buddhism in Koryo: A Royal Religion*, edited by Lewis R. Lancaster, Kikun Suh, and Chai-shin Yu, 103–138. Berkeley: Institute of East Asian Studies, University of California, 1996.

Shim, Jae Ryong. "The Philosophical Foundation of Korean Zen Buddhism: The Integration of Sŏn and Kyo by Chinul (1158–1210)." PhD diss., University of Hawaii, 1979.

Charles Muller (2005)

JOACHIM OF FIORE
(c. 1135–1202)

Joachim of Fiore, the Christian mystical philosopher of history, lived in Calabria, Italy, a region characterized by the remote hermit life, yet close to Sicily, the hub of the Mediterranean. This combination of withdrawal from and encounter with the world also characterized Joachim's life. Becoming a Cistercian, by 1177 he was abbot of Curazzo, but he obtained papal permission to retire from monastic administration to a more remote mountainous region, where he founded the order of San Giovanni in Fiore about 1192. Yet he descended to dramatic encounters—in which he prophesied on contemporary events and the advent of Antichrist—with Pope Lucius III (1184), King Richard I of England (1190–1191), and the Holy Roman Emperor Henry VI (1191), and he meditated deeply on contemporary history, especially the two great menaces to Christianity: the infidel and the heretic.

Joachim recorded two mystical experiences: one at Easter, when he was given understanding of the inner concords between the two testaments, and one at Whitsuntide, when he received illumination on the doctrine of the Trinity. Disclaiming the title of prophet, he believed that through the gift of spiritual intelligence he understood the inner spiritual meaning of history.

With papal encouragement, Joachim set out to expound this belief in his three main works, the *Liber Concordiae,* the *Expositio in Apocalypsim,* and the *Psalterium Decem Chordarum.* His exposition turns chiefly on an interwoven double pattern of twos and threes. The two testaments represent history in two eras, culminating, respectively, in the First and Second Advents and marked continually by concords—for example, twelve Tribes and twelve Churches, seven Seals and seven Openings. History is also trinitarian, growing treelike from the Age (*status*) of the Father (Law) to that of the Son (Grace) to that of the Spirit (Spiritual Understanding), yet in a double "procession" of the third *status* from both the first and the second. This third *status* represents an apotheosis of history, which Joachim equated with the Seventh, Sabbath Age of the traditional Seven Ages, placing it between the worst Antichrist and the end of history. He saw himself on the threshold of the last two generations of the Sixth Age, into which will be crowded the greatest tribulations before the church "crosses Jordan" into the Sabbath of the third *status.*

His strong visual imagination led him to embody this philosophy of history in the remarkable *Liber Figurarum,* through which it was widely disseminated. This doctrine contained revolutionary seeds. Joachim avoided dangerous implications by using his pattern of twos to proclaim that the authority of the Scriptures and church would endure until history ended. His pattern of threes culminated in a spiritual state rather than a historic era. Nonetheless, he almost gave it a starting date—1260—and expected its *Ecclesia Spiritualis,* symbolized in John, to "outlast" the church designated in Peter. This inspired fanatical groups to proclaim the Third Age, the overthrow of existing ecclesiastical institutions, and the transfer of authority to the Eternal Evangel. Joachim's prophecies of two new spiritual orders to lead the church into the Third Age were claimed first by Franciscans and then by Dominicans, Augustinian friars, and even Jesuits.

Condemned as a heretic, revered as a saint, Joachim seldom met with indifference to his views. From the thirteenth to the sixteenth century, when an optimistic expectation of history was proclaimed, it usually drew inspiration from Joachimism.

See also Mysticism, History of; Philosophy of History.

Bibliography

WORKS BY JOACHIM

Liber Concordiae Novi ac Veteris Testamenti. Venice, 1519.

Expositio in Apocalypsim. Venice, 1527.

Psalterium Decem Chordarum. Venice, 1527.

Tractatus Super Quatuor Evangelia. Edited by E. Buonaiuti. Rome: Tipografia del Senato, 1930.

Il libro delle figure, 2nd ed., Vol. I, edited by L. Tondelli; Vol. II, edited by L. Tondelli, M. Reeves, and B. Hirsch-Reich. Turin: Società Editrice Internazionale, 1953.

WORKS ON JOACHIM

Bloomfield, M. "Joachim of Flora. A Critical Survey of His Canon, Sources, Biography and Influence." *Traditio* 13 (1957): 249–311.

Grundmann, H. *Neue Forschungen über Joachim von Fiore.* Marburg: Simons, 1950.

Grundmann, H. *Studien über Joachim von Floris.* Leipzig and Berlin, 1927.

La Piana, G. "Joachim of Flora. A Critical Survey." *Speculum* 7 (1932): 257–282.

Reeves, M. "The *Liber Figurarum* of Joachim of Fiore." *Mediaeval and Renaissance Studies* 2 (1950): 57–81.

Reeves, M., and B. Hirsch-Reich. "The Seven Seals in the Writings of Joachim of Fiore." *Recherches de théologie ancienne et mediévale* 21 (1954): 211–247.

Marjorie E. Reeves (1967)

JODL, FRIEDRICH

(1849–1914)

Friedrich Jodl ranks as one of the most significant representatives of German positivism, although this designation by no means adequately characterizes the full scope of his ideas. Jodl was born in Munich, where in 1880 he qualified as a *Privatdozent* in philosophy. Five years later he was named professor of philosophy at the German University in Prague. In 1896 he accepted a call to the University of Vienna. His many publications ranged over the fields of philosophy and the history of philosophy and ethics, as well as psychology and aesthetics.

Jodl categorically rejected metaphysical speculation. For him, the boundaries of experience were at the same time the boundaries of knowledge; hence, there could be no a priori knowledge, nor any metaphysical cognition of the transcendental. The task of philosophy, he maintained, is to order scientific knowledge systematically and to comprehend it in a unified view of the world. The basis of philosophy, like that of science, can only be experience.

As a consistent empiricist, Jodl criticized phenomenalism, preferring critical realism. The factual existence of a transsubjective reality is guaranteed by the thou-experience, by the existence of one's fellow men. Moreover, without the assumption of an objective external world and without the assurance that we know it as such, natural science would be impossible. Hence, the forms of our intuition and of our thought are not subjective in the sense meant by extreme epistemological idealism; rather, they are also conditioned by the relationships of things. Our knowledge of the world is not subject to a theoretical limit beyond which our consciousness is unable to grasp reality; there is only a frontier that can always be pushed further back, with the result that the world in its totality constitutes an endless problem, a task for knowledge that can never be definitively solved.

Jodl sought a naturalistic conception of the world, free of religion and metaphysics, such as that of the monistic movement, which he energetically promoted. "We need no other mediator between us and nature except our understanding and a courageous will, nor any mystery behind nature to console us for her; we are alone with nature, and we feel secure because we possess intellect and she behaves according to laws" (*Vom wahren und vom falschen Idealismus,* p. 40).

Jodl treated the problem of God on the basis of this naturalistic monism. Somewhat like John Dewey after him, Jodl, while denying the existence of God in any traditional sense, retained the term *God* as a designation for the highest ideals to which human beings aspire.

In his psychology too Jodl confined himself to the clearest possible presentation of the empirically given facts of mental life, renouncing all metaphysical assumptions. His psychological investigations are unusually rich in acute analyses and genetic explanations. Consciousness is not a substance but an act; it is the inwardness of a living creature. The bearer of consciousness is not an immaterial soul but the living organism; the soul is nothing other than the unified coherence of experience. "Mental" and "physical" are simply two expressions in different languages for one and the same occurrence. Body and consciousness are one; the psychical is the internal, subjective experiencing of neurological processes. An individual experiences as subject the whole complex of his brain processes in internal perception.

In ethics, Jodl was a convinced evolutionist. Ethical values have been subject to continuous transformation; morality is an evolutionary product of the interaction between the individual and society. Jodl made a sharp distinction between the subjective, psychological basis of morality and the objective, axiological criterion for it, although the two, in his view, were most intimately connected. The basis of morality is the will, which rests on

social instincts, is influenced by reason, and is aimed at the welfare of the whole. A different question is the establishment of moral norms by which to measure the worth of human attributes and deeds. This requires that one take into account both the motivation and the utilitarian value of an action. In his penetrating studies in the history of ethics, Jodl showed that this discipline has, in the course of its development, increasingly freed itself from metaphysics and has replaced the theocentric foundation with an anthropocentric one.

Jodl, characteristically, was not content with theoretical (historical and systematic) studies in ethics, but sought beyond that to carry out in life a practical, ethical idealism. Imbued with a faith in the value of life and a vigorous optimism in regard to culture and progress, he was an "enlightener" advocating the humanization of culture; an ethically based social life in the spirit of a purely secular, humane morality and freedom of thought. He strongly supported and promoted the system of free popular education and the Ethical Culture movement.

See also Dewey, John; Evolutionary Ethics; History and Historiography of Philosophy; Positivism.

Bibliography

WORKS BY JODL

Leben und Philosophie David Humes. Halle: C.E.M. Pfeffer, 1872.

Die Kulturgeschichtsschreibung. Halle, 1878.

Geschichte der Ethik als philosophischer Wissenschaft, 2 vols. Stuttgart: Cotta, 1882–1889.

Lehrbuch der Psychologie, 2 vols. Stuttgart and Berlin: Cotta, 1897.

Ludwig Feuerbach. Stuttgart: F. Frommans, 1904.

Der Monismus und die Kulturprobleme der Gegenwart. Leipzig: A. Kröner, 1911.

Vom wahren und vom falschen Idealismus. Leipzig: A. Kröner, 1914.

Vom Lebenswege. Edited by Wilhelm Börner, 2 vols. Stuttgart and Berlin: Cotta, 1916–1917. Collected essays and lectures.

Aesthetik der bildenden Künste. Edited by Wilhelm Börner. Stuttgart and Berlin, 1917; 2nd ed., 1922.

Allgemeine Ethik. Edited by Wilhelm Börner. Stuttgart and Berlin: Cotta, 1918.

Kritik des Idealismus. Edited by Carl Siegel and Walther Schmied-Kowarzik. Leipzig: Akademische Verlagsgesellschaft, 1920.

Geschichte der neueren Philosophie. Edited by Karl Roretz. Vienna, Leipzig, and Munich, 1924.

WORKS ON JODL

Börner, Wilhelm. *Friedrich Jodl.* Stuttgart and Berlin: Cotta, 1911.

Jodl, Margarete. *Friedrich Jodl, sein Leben und Wirken, dargestellt nach Tagebüchern und Briefen.* Stuttgart and Berlin: Cotta, 1920.

Franz Austeda (1967)
Translated by Albert E. Blumberg

JOHN DUNS SCOTUS

See *Duns Scotus, John*

JOHN OF DAMASCUS

(c. 675–c. 750)

John, whose secular name was Mansur, was born in Damascus probably in the third quarter of the seventh century. His father and grandfather had been prominent in the fiscal administration of Syria, and it is believed that his father was in charge of the fiscal administration of the Umayyad Empire, with its capital in Damascus, in the latter decades of the seventh century. John received a good Hellenistic education and probably entered the service of the Caliph at Damascus in his father's footsteps. At some point—probably at the beginning of the eighth century, when the administration of the Umayyad Empire was put in the hands of Muslim officials—John resigned his post and became a monk in the Holy Land, according to a late tradition at the monastery of Mar Saba in the Judaean Desert (though it is more likely that he was associated with the church of the Anastasis in Jerusalem itself). In any event, he was close to John V, patriarch of Jerusalem from 706 to 735, who ordained John of Damascus to the priesthood. It is believed that he died around 750, because at the Iconoclast Synod of Hiereia (754) he was anathematized under his name Mansur as if he was already dead.

John possessed genuine literary gifts, knowledge both theological and philosophical, and considerable intellectual acumen. In his lifetime he achieved fame as a preacher (evident from the references to him in the *Chronicle* of Theophanes); his liturgical poetry still forms the core of the Byzantine liturgical office; through his works of theology he came to exercise an unparalleled influence, not only throughout the Byzantine world, but on theology in the West from the period of Scholasticism (for which he provided the principal access to the developed theology of the Byzantine East) up to at least the time of Friedrich Daniel Ernst Schleiermacher. Many of his theological works are polemical; he wrote a treatise against Manichaeism, as well as treatises against the

Christological heresies of monophysitism, monothelitism, and Nestorianism and is the first Christian theologian explicitly to attack the new religion of Islam, on which he was impressively informed.

John was the most notable defender of icons against the iconoclasm of the Emperor Leo III. His best-known work is a three-part treatise known as *The Fountain Head of Knowledge* (in Greek: *Pēgē gnōseōs*), consisting of an introduction to logic (*Dialectica*), a summary account of heresies (*De haeresibus*) and an epitome of the principal themes of the Christian faith (*Expositio fidei*, or *De orthodoxa fide*). The critical edition published by the Benedictine Dom Boniface Kotter in the last decades of the twentieth century reveals that each of these sections was intended to be a century (or half-century)—that is, a collection of one hundred chapters or paragraphs. The century was a genre of monastic literature, popular in Byzantine circles, and John's choice of this genre makes clear that his purpose in this work was essentially monastic: the intellectual training and learning it provides was ultimately to help monks in their life of prayer. His principal purpose was not to provide a systematic theology, as was suggested by the division of the *Expositio fidei* into four books, corresponding to the four books of Peter Lombard's *Sentences*, a division introduced into the Latin translations in the thirteenth century (and thence into the older editions and translations), but unknown in the Greek manuscript tradition. In fact, although John intended a final version of the work in three parts as indicated above, the most popular form in the Byzantine world was a combination of the *Dialectica* and the *Expositio*, usually known as *Philosophical and Theological Chapters*, usually consisting of 150 chapters.

The *Dialectica* is a compilation—belonging to a tradition of Christian introductions to logic, popular in the seventh and eighth centuries—that provided an introduction to basic philosophical terminology as an aid to understanding the issues raised by Christological controversy in the East, which had raged since the fifth century, and concerned concepts such as being, nature, person, and latterly activity and will. The *Dialectica* in the earlier form (the only one that survives complete; John probably died while revising it) seems to lead up to the notion of *hypostasis* or person, and "hypostatic union," key terms in the Christological orthodoxy to which John belonged. Those who contributed to this tradition of Christian introductions to logic drew their material from the sixth-century Alexandrian commentaries on Aristotle and Porphyry; unlike the sixth-century commentators, however, the compilers of these textbooks (including John himself) were not concerned to advance an understanding of logic, but simply to provide the basic tools for engaging in the theological arguments of the day.

The *Expositio Fidei* is also a work of compilation, drawing, often word for word, on earlier theologians in its presentation of the fundamental concepts of the Christian faith. It concerns the doctrine of God and the Trinity; creation and the nature of the created order, especially human nature; the doctrine of Christ (to which most space is devoted); and various questions of religious practice, especially those that marked off Christians from Jews and Muslims (though there is no explicit reference to the latter). John expressly sets aside any claim to originality; even the selection of authorities is probably not original to John, but represents an established tradition, much influenced by Maximos the Confessor (580–662). The only doctrines where some originality could be claimed for John are the doctrines of the will and its freedom, which had become central to the controversy over the heresy that Christ had only one will (monothelitism), and possibly his treatment of the infinity of God. In both cases, however, John's contribution is not much more than a refinement of the tradition that had reached him. Some aspects of the tradition he had received are ignored, possibly felt to be too daring: for example, his dependence on Maximos (and through him on the fourth-century Nemesios of Emesa) for his understanding of creation and human nature is particularly marked, but he ignores completely Maximos's developed doctrine of the principles (or *logoi*) of creation. It is doubtless the clarity of John's exposition that is the reason for his immense influence.

See also Aristotle; Byzantine Philosophy; Mani and Manichaeism; Medieval Philosophy; Porphyry; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

WORKS BY JOHN DAMASCENE

His works my be found in *Patrologia Graeca*, edited by J.-P Migne, vols. 94–96 (Paris: J.-P. Migne, 1860). A critical edition of (most of) his authentic prose works is *Die Schriften des Johannes von Damaskos*, 5 vols., edited by Bonifatius Kotter (Patristische Texte und Studien 7, 12, 17, 22, 29. Berlin: de Gruyter, 1969–1988). An English translation by Frederic H. Chase Jr. (from Migne) of *The Fountain Head of Knowledge* is *Writings* (Fathers of the Church 37. New York: Fathers of the Church, 1958). An English translation by Andrew Louth (from Kotter) of works against the iconoclasts is *Three Treatises on the Divine Images* (Crestwood, N.Y.: St. Vladimir's Seminary Press, 2003).

WORKS ABOUT JOHN DAMASCENE

Frede, Michael. "John of Damascus on Human Action, the Will, and Human Freedom." In *Byzantine Philosophy and its Ancient Sources*, edited by Katerina Ierodiakonou, 63–95. Oxford, U.K.: Clarendon Press, 2002.

Louth, Andrew. *St. John Damascene: Tradition and Originality in Byzantine Theology*. Oxford, U.K.: Oxford University Press, 2002.

Roueché, Mossman. "Byzantine Philosophical Texts of the Seventh Century." *Jahrbuch der Österreichischen Byzantinistik* 23 (1974): 61–76.

Sweeney, Leo. "John Damascene and Divine Infinity." *New Scholasticism* 35 (1961): 85–86.

Sweeney, Leo. "John Damascene's 'Infinite Sea of Essence.'" *Studia Patristica* 6 (1962): 248–263.

Andrew Louth (2005)

JOHN OF JANDUN

(c. 1286–c. 1328)

John of Jandun, also known as Jean de Jandun and Johannes de Janduno, was foremost among the Averroists at Paris in the fourteenth century. He was born in the village of Jandun in the French province of Champagne. The estimate of his date of birth is based on the year 1310, the earliest date found on any writing definitely attributable to him; at the time of this first or very early publication, John would have been a recent master of arts, and reckoning by the age and curricular requirements in effect at the University of Paris in the early fourteenth century, he could not have been much more than twenty-four years of age. John was active throughout the next decade and a half as master of arts at the Collège de Navarre, in Paris, although he was nominally canon at Senlis—the kind of preferment awarded a practicing teacher and scholar during the Middle Ages. At Paris he lectured on the standard curriculum of the Faculty of Arts: Aristotle's *Physics, De Coelo et Mundo, De Anima, Metaphysics, Parva Naturalia,* and *Rhetoric* and Averroes's *De Substantia Orbis.* John's commentaries on these works date from 1310 to 1323. Additional writings from this period attest to his interest in particular problems arising in his lectures and commentaries; there still survive many independent *quaestiones* and *disputationes,* which, in the medieval tradition, supplemented the normal course of studies with special studies and advanced seminars.

By 1324 he was closely associated with Marsilius of Padua, also a master of arts at Paris, in connection with Marsilius's famous and controversial *Defensor Pacis,* published that year. Although John does not seem to have shared in the actual composition of the work, he was apparently an intellectual intimate of Marsilius. The *Defensor Pacis,* a powerful affirmation of the temporal and civil authority over the spiritual and papal, occasioned enough ecclesiastical outrage for John and Marsilius to deem it prudent to leave Paris and seek the protection of Louis IV of Bavaria. Louis was himself embroiled with Pope John XXII on matters of political and spiritual authority and was soon to harbor another intellectual fugitive from Paris, William of Ockham. In 1326 and 1327 a series of papal bulls appeared specifically attacking John and Marsilius, and the final one, dated October 23, 1327, excommunicated them as "heretics and heresiarchs."

The remainder of John's life was brief. He followed Louis in the invasion of Italy and was rewarded with the episcopate of Ferrara. Probably en route to assume his new duties, he died at Todi, not later than August 31, 1328.

THOUGHT

To treat John of Jandun's philosophy, as many historians have done, as a blind recapitulation of the Commentator (the title by which Averroes was referred to throughout the Middle Ages and the Renaissance) and his special views on Aristotle would be an oversimplification. It is true that John did, at one point in his commentaries, call himself the "ape of Averroes," but this was in the context of a particular passage of Aristotle's *Metaphysics,* where John considered Averroes's remarks perfectly adequate. It is also true that John preferred, generally, Averroes's rendering of Aristotle, but it is not illuminating to call him "Averroist" without severe qualifications. Other medieval philosophers (for example, Siger of Brabant) can be termed Averroist, but their speculative positions were sometimes methodologically quite distinct from those of John. It is probably most accurate to place him in the philosophical tradition and method sometimes exemplified in Christian Augustinianism, always recognizing, however, that he was oriented intellectually within the traditions of the Faculty of Arts rather than those of the Faculty of Theology.

John's espousal of a *sensus agens* (active principle in the process of sensation), of a plurality of substantial forms in the individual (one for each of the three functions of living: vegetating, sensing, and thinking), of the soul's ability to grasp separate substances (that is, forms) directly, of form as the immanent and essential cause of natural activity, and of other kindred doctrines can be found in the thinking of many Augustinian theologians close to his time, such as Bonaventure, Peter John Olivi,

Roger Marston, John Duns Scotus, and Peter Aureol. John's own advocacy of these views arose, however, out of the use of Averroes in the analysis of Aristotle for the Faculty of Arts curriculum. Although John's version of Averroism and the tradition called Christian Augustinianism had much in common methodologically, they sprang from different institutional contexts.

John's interpretations of Averroes's commentaries on Aristotle were both acute and influential (as late as the seventeenth century his writings were still used alongside those of Averroes by the Paduan pedagogue Cesare Cremonini), but his place in intellectual history is due less to the conspicuous originality of his thought than to his unusually explicit delineation of the respective domains of faith and reason. Whenever confronted, in his analysis of Aristotle, with a conclusion severely at variance with some doctrine of Christian faith, John appended an apologia of the following kind: "It must be noted that, although the *dicta* are … according to the principles of Aristotle and the Commentator, it must be replied firmly according to faith and truth that the world is not eternal." Similar passages abound in John's commentaries; whenever conclusions of reason arrived at in the logic of Aristotle and Averroes differed from the dictates of Christian dogma, John introduced statements proclaiming the consistency of the reasoning but immediately ceding truth itself to the preeminent demands of faith.

Such remarks have had two interpretations. First, John has been indicted, with other so-called Averroists; as holding a theory of "double truth"—that is, that statements of faith, on the one hand, and conclusions of reason, on the other, can be simultaneously true, yet contradictory. This charge has been discounted effectively by Étienne Gilson; no medieval writings maintaining such a self-inconsistent view have yet been found. Medieval thinkers never stated more than the position that although reason can systematically reach certain conclusions, Christian faith is nevertheless the final arbiter of truth when such conclusions conflict with matters of doctrine.

Second, certain of John's disclamatory passages have been interpreted as actually revealing a fundamental religious insincerity. For example, he said:

> This is not known *per se,* nor is it demonstrable by any human proof, but we believe this to be so solely by divine authority and by the Sacred Scriptures. And to the credulity toward things of this kind and similar things, the habit of listening to this sort of thing from childhood adds a good deal.

Or again:

> If anyone knows how to prove this and to make it accord with the principles of philosophy, let him rejoice in this possession, and I will not grudge him, but declare that he surpasses my ability.

And finally:

> although every form inherent in matter is corruptible I say, however, that God can perpetuate it and preserve it eternally from corruption. I do not know the manner of this; God knows it.

Such statements have been interpreted as indicating a radical insincerity in John's thinking, a covert mocking of Christian faith. Thus, some historians have suggested that John was not merely maintaining a "double truth" but actually affirming the superior reliability of the conclusions of unaided reasoning in the mode of Aristotle and could therefore stand as an early precursor of seventeenth-century rationalism and libertarianism.

On close examination, however, John's position on the relation between the claims of faith and claims of reason does not seem to have been distinctively more radical than the thinking of many other medievals. (Similar disclaimers of reason in favor of faith can be found in many commentaries on Aristotle, including those of Thomas Aquinas.) Such discrepancies and apparent conflicts reflect, in small part, a strong institutional rivalry between the faculties of arts and theology and, in large part, a fundamental intellectual crisis occasioned by the confrontation between Greek rationalism and Christian dogma.

See also Aristotle; Augustinianism; Averroes; Averroism; Duns Scotus, John; Marsilius of Padua; Marston, Roger; Olivi, Peter John; Peter Aureol; Sensa; Siger of Brabant; Thomas Aquinas, St.

Bibliography

WORKS BY JOHN OF JANDUN

Quaestiones de Anima. Fourteen printings in Venice between 1473 and 1587.

Quaestiones Super Libros Physicorum. Twelve printings in Venice, Paris, and Lyons between 1488 and 1598.

Quaestiones in Duodecim Libros Metaphysicae. Four printings in Venice between 1525 and 1586.

John's most influential writings, the commentaries on Aristotle—including those listed above—can be found only in Renaissance editions. Some lesser writings exist in printed editions published in the fifteenth and sixteenth centuries.

All of John's works can be found in manuscript form in various European libraries.

WORKS ON JOHN OF JANDUN

Gilson, Étienne. "La doctrine de la double vérité." In *Études de philosophie médiévale,* 51–75. Strasbourg: University of Strasbourg, 1921.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages.* New York: Random House, 1955.

MacClintock, Stuart. *Perversity and Error.* Bloomington: Indiana University Press, 1956.

Valois, N. "Jean de Jandun et Marsile de Padoue, auteurs du *Defensor Pacis.*" In *Histoire littéraire de la France,* Vol. XXXIII, 28–623. Paris, 1903.

Stuart MacClintock (1967)

JOHN OF LA ROCHELLE

(c. 1190–1245)

John of La Rochelle, or de Rupella, was a Franciscan philosopher, theologian, and preacher at the University of Paris. The first clear reference to him (in Thomas of Cantimpré) indicates that in 1238 he was already a friar and a master in theology. From John's own writings, as well as from his knowledge of and interest in philosophy, we may deduce that he had studied and perhaps taught in the faculty of arts before becoming a theologian. His *Summa de Vitiis* (Summa on vices and sins), which manifests his penchant for ethical questions, is directly dependent on William of Auxerre, Prevostinus, and Stephen Langton, who apparently were John's teachers in the faculty of theology. It seems that only after Alexander of Hales entered the order, in 1236, did John become acquainted with that famous theologian; thereafter John was Alexander's faithful companion and collaborator. Both seem to have taught at Paris until their deaths in 1245.

Though a famous preacher and biblical commentator, John is known primarily as a "summist" interested in questions of psychology and morals. Both topics are combined in his early *Tractatus de Anima et de Virtutibus* (Tract on the soul and the virtues), a kind of rambling compilation of definitions of the soul, the divisions of the soul's powers according to the philosophers, and the division of the virtues according to Plotinus, Cicero, Aristotle, and Augustine. Ethical questions predominate in John's proposed *Summa Theologicae Disciplinae* (Summa of theological learning). As set forth in the prologue to the *Summa De Articulis Fidei* (Summa on the articles of faith), the larger summa of theology was to include both doctrines and morals. "Morals is divided into two parts: on sins and the remedies of sins. These remedies are four in number: commandments, virtues, the gifts of the Holy Spirit, and the sacraments" (ms. Milan, *Brera A.D. IX. 7,* fol. 75a). Only parts of such a summa seem to have been completed: the "Summa on Vices and Sins," *De Praeceptis et Consiliis* (On precepts and counsels; a tract), *De Virtutibus* (On virtues), and the *Summa de Donis* (On the gifts of the holy spirit). The same interest is reflected in the lengthy and influential tract *De Legibus et Praeceptis* (On laws and precepts), which is probably John's work, and in such Disputed Questions as "The Fall of Human Nature," "On Negligence, Hypocrisy, the Seven Capital Sins," "On Usury," and "On the Just War," all as yet unpublished.

His early "Tract on the Soul" was developed into the more mature *Summa de Anima* (Summa on the soul), which is rightly regarded as the first scholastic textbook of psychology. Beginning with proofs (from Avicenna and Augustine) for the existence of the soul, the first part examines the essence, causes, and properties of the soul and its union with the body (giving a none too clear, yet basically Aristotelian, solution of this latter problem), with a final section on immortality and the status of the soul after death. In the second half John considered at length the problem of the powers of the soul: their relation to the essence and their division according to Pseudo-Augustine (*De Spiritu et Anima*), John of Damascus, and Avicenna. The classification of the external and internal senses, the cognitive and motive powers, follows closely the *De Anima* of Avicenna, with some slight additional material. A comparison with the earlier "Tract" leads us to conclude that the *Summa de Anima* is incomplete. It ends abruptly in the midst of a discussion on the will. The large number of extant manuscripts attests to its popularity in the Middle Ages. If not strikingly original, the summa is of interest also for its use of the philosophers at a time when theologians were inclined to reject their help. John pointedly rejected such an attitude in a university sermon:

> If philosophy is neglected, one may fear lest "there be found no smiths in Israel" (I Samuel 13:19), that is, philosophers who will sharpen our wits like "swords" and with shining "lances" attack the enemy at a distance. The devil himself seeks to stamp out the study of philosophy because he does not want Christians to have sharp minds. (*Collectanea Franciscana,* Vol. 28 [1958], 50)

Last, John is to be considered the primary author or compiler of the first and third books of the so-called "Summa of Alexander of Hales."

See also Alexander of Hales; Aristotle; Augustine, St.; Avicenna; Cicero, Marcus Tullius; John of Damascus; Plotinus; Psychology; Virtue and Vice.

Bibliography

WORKS BY JOHN OF LA ROCHELLE

Tractatus de Anima et Virtutibus. Edited by P. Michaud-Quantin. Paris, 1964.

Summa de Anima. Edited by J. G. Bougerol. Paris: J. Vrim, 1995.

Eleven Marian Sermons. Edited by Kilian F. Lynch. St. Bonaventure, NY: Franciscan Institute, 1961.

WORKS ON JOHN OF LA ROCHELLE

Brady, I. "Law in the Summa Fratris Alexandri." *Proceedings of the American Catholic Philosophical Association* 24 (1950): 133–147.

Doucet, V. *Prolegomena in Librum III Necnon in Libros I et II "Summa Fratris Alexandri,"* pp. 211–227 and passim. Quaracchi, Italy, 1948.

Lottin, Odon. *Psychologie et morale aux XII*[e] *et XIII*[e] *siècles,* 6 vols. Gembloux, Belgium, 1942–1960.

Michaud-Quantin, P. "Les puissances de l'âme chez Jean de la Rochelle." *Antonianum* 24 (1949): 489–505.

Ignatius Brady, O.F.M. (1967)

JOHN OF MIRECOURT
(fourteenth century)

John of Mirecourt belongs to a generation of philosopher-theologians discussing the nature of knowledge and especially the varieties of evidence for human knowledge. Biographical information on him is scarce, but he lectured on the *Sentences* of Peter Lombard (c. 1095–1160) at the University of Paris in 1344–1345. Propositions taken from his work were condemned by the chancellor in 1347. His two apologies are the best-known writings by him, although his commentary on Lombard's *Sentences* has also survived. Traditionally, Mirecourt has been described as a skeptic associated with Nicholas of Autrecourt (c. 1300–after 1350). Research in the last decades of the twentieth century gave a somewhat more accurate picture of his epistemology, but other areas of his philosophy, like his theory of ethics and the will, have not been examined.

In his epistemology Mirecourt distinguishes between abstractive and intuitive cognition, following William of Ockham (c. 1285–?1349), John Duns Scotus (1266?–1308), and some earlier scholars. *Abstractive cognition* can be defined as a cognition that can be had without its object being present, while *intuitive cognition* is dependent of the presence of its object and allows (or even produces) evident knowledge that the object exists. For example, when one sees Peter, one gains an intuitive cognition with Peter as the object of the cognition and is able to give evident assent to the proposition "Peter exists." In abstractive cognition the object need not be present, and thus the examples are often of a conceptual or mathematical nature.

According to Mirecourt's classification, suspicion, opinion, and conjecture constitute *inevident assent*. In contrast, *evident assent* is firm belief, which, as he sees it, can be either supranatural or grounded in natural causes. Mirecourt concentrates more on the latter. There are cases where it is impossible that evident assent is wrong, and Mirecourt talks in this context of *special evidence*. Evidence reducible to the certainty of first principles is like this. Mirecourt's examples include: "If it is a man, it is an animal" and "God is God." The latter is a logical truth, but the former shows that he has in mind also conceptual necessities. Surpassing the border of what may be called analytic truth, there is also special evidence that something exists. Mirecourt proves this by saying, "If one doubts whether something exists or whether one exists, one has to concede that it follows evidently: One doubts whether something exists, therefore one exists, since if one did not exist, one would not doubt. So it follows: one exists, therefore something exists" (*In I librum Sententiarum*, q. 6). Here, knowledge of one's own existence is shown to have special evidence. Mirecourt continues by pointing out that nothing else is known to exist with infallible evidence.

Not all natural evidence is special. Mirecourt faces the skeptical challenge that any belief without special evidence could be false by God's absolute power. Our ordinary experiences could thus be like dreams or hallucinations. Just like René Descartes (1596–1649) some centuries later, Mirecourt accepts that this is in some sense possible. Nevertheless, Mirecourt thinks that ordinary experiences can constitute genuine knowledge and thus he is not really a skeptic in the classical sense.

See also Descartes, René; Duns Scotus, John; Nicolas of Autrecourt; Peter Lombard; William of Ockham.

Bibliography

WORKS BY MIRECOURT

In I librum Sententiarum (Commentary on Book I of the *Sentences*). Preliminary edition at Massimo Parodi's web site at http://filosofia.dipafilo.unimi.it/~mparodi/.

Apologiae, edited by F. Stegmüller in his article "Die zwei Apologien des Jean de Mirecourt." *Recherches de théologie ancienne et médièvale* 5 (1933): 40–78 and 192–204.

WORKS ON MIRECOURT

Courtenay, William. "John of Mirecourt and Gregory of Rimini on Whether God Can Undo the Past." *Recherches de théologie ancienne et médièvale* 39 (1972): 224–256 and 40 (1973): 147–174.

Parodi, Massimo. "Questioni inedite tratte dal I libro del Commento alle Sentenze di Giovanni di Mirecourt (qq. 13–16)." *Medioevo* 3 (1977): 237–284.

Mikko Yrjönsuuri (2005)

JOHN OF PARIS

(c. 1255–1306)

John of Paris, or John Quidort, also known as Surdus or Monoculus, was a Dominican scholastic philosopher and theologian, priest, and author. A native of Paris, John studied and taught philosophy at the University of Paris before entering the Dominican order at St. Jacques prior to 1279. As bachelor in theology he lectured on the *Sentences* of Peter Lombard (1284–1286) and energetically defended the then suspect doctrines of Thomas Aquinas in a famous refutation, *Correctorium* "*Circa,*" of the *Correctorium* of William de la Mare, which had been officially adopted by the Franciscans. Certain unknown adversaries managed to twist or misinterpret sixteen statements delivered in class, and in 1286 they had John denounced to the authorities. Although he ably explained the true meaning of his innocent statements, his academic career was temporarily suspended. In his defense of Thomas, John showed a clear understanding of the Thomistic distinction between essence and existence in creatures, the unicity of substantial form in material substance, the individuation of material substances by matter alone, and the pure potentiality of first matter.

From 1300 on John was again active in Paris, teaching, preaching, and writing. His sermons and treatises testify to the political and social unrest of the times. During the struggle between Pope Boniface VIII and Philip the Fair, John wrote the important treatise *De Potestate Regia et Papali* (On royal and papal power; 1302), in which, following Thomas Aquinas, he defended a middle position between the papalist and imperialist extremes. He clearly distinguished between two autonomous societies in Christendom—church and state—each of which has its independent, legitimate source of authority and its rightful area of concern. For him the source of royal power was not delegation from the pope but the nature of humankind acting reasonably and freely for the common good of society.

In 1304, John was given license to incept in theology, succeeding Raymond Romani as master. In 1305, John presided over a solemn disputation before the bishop of Paris and the faculty of theology, in which he maintained that since the church had not yet defined the doctrine of transsubstantiation, one could hold as equally probable the doctrine that later became known as "impanation"—that is, the continued existence of bread after consecration, now assumed in Christ. This novel view was examined by a number of bishops and theologians, who considered it heretical. John was suspended from all teaching and preaching, perpetual silence being imposed upon him under pain of excommunication. John appealed his case to the papal *curia* at Bordeaux, where he died on September 22, 1306, while awaiting a decision. At the very beginning of the Eucharistic controversy he had publicly expressed his willingness to retract his view should it prove contrary to the teaching of the church.

John was a gifted speculative thinker who, while accepting the basic principles of Thomas Aquinas, was eager to deal with new problems in philosophy and theology.

See also Essence and Existence; Medieval Philosophy; Peter Lombard; Thomas Aquinas, St.; Thomism.

Bibliography

Glorieux, Palémon. *Répertoire des maîtres en théologie* 1 (60) (1933): 189–193.

Grabmann, Martin. "Studien zu Johannes Quidort von Paris, O.P." In *Sitzungsberichte der Bayerischen Akademie der Wissenschaften zu München.* Munich, 1922.

Käppeli, Thomas. "Praedicator Monoculus." *Archivum Fratrum Praedicatorum* 27 (1957): 120–167.

Leclercq, Jean. *Jean de Paris et l'ecclésiologie du XIII[e] siècle.* Paris: J. Vrin, 1942.

Müller, Jean P. "Les reportations des deux premiers livres du Commentaire sur les Sentences de Jean Quidort de Paris, O.P." *Angelicum* 33 (1956): 361–414.

Müller, Jean P. "La date de la lecture sur les Sentences de Jean Quidort." *Angelicum* 36 (1959): 129–162.

Quétif, Jacques, and Jacques Échard. *Scriptores Ordinis Praedicatorum,* Vol. I, 500–502. Paris: J.-B.-C. Ballard and N. Simart, 1719.

Roensch, Frederick J. *Early Thomistic School.* Dubuque, IA: Priory Press, 1964.

James A. Weisheipl, O.P. (1967)

JOHN OF ST. THOMAS
(1589–1644)

John of St. Thomas, the Spanish theologian and philosopher, was born John Poinsot, the son of an Austrian, at Lisbon, Portugal, and died at Fraga, Spain. When he entered the Dominican order he took his name from St. Thomas Aquinas. John studied philosophy at Coimbra, Portugal, and theology at Louvain, taught philosophy and theology in Dominican houses of study, at Alcalá de Henares (1613–1630), and from 1630 to 1643 was a professor at the University of Alcalá. Apart from certain Latin and vernacular works of devotion, his writings consist of two series of textbooks, one in philosophy, the *Cursus Philosophicus* (which comprises "Ars Logica," covering logic, and "Philosophia Naturalis," on natural philosophy), the other in theology, the *Cursus Theologicus* (a systematic commentary on Thomas's *Summa of Theology*).

The "Ars Logica" is fundamentally Aristotelian logic, but John developed the content of the course in two directions: toward a formal theory of correct reasoning and toward a material logic that attends to the meaning of the actual terms of a proposition and thus anticipates some of the problems of epistemology and semantics. John's terminology differs from that of modern logic (*propositio copulativa* is the modern conjunctive proposition; *propositio disiunctiva* the alternative proposition; *bona consequentia* means *implication*). However, it has been claimed, by J. J. Doyle, that the "Ars Logica" and Alfred North Whitehead and Bertrand Russell's *Principia Mathematica* are fundamentally similar as formal systems. Concerning material implication, John taught that one may infer from the particular proposition ("Some man is rational") to the universal proposition ("Every man is rational") in cases where the matter is necessary. To some extent he anticipated problems in the philosophy of science and the metasciences and also the theory of induction.

His philosophy of nature is a systematic exposition of a type of Thomism much influenced by the commentaries of Cajetan. Nature is the world of bodies, of being that is subject to change (*ens mobile*), explained in terms of the four Aristotelian causes, substance and accidents, act and potency, matter and form.

John treated certain questions in a novel way—for example, immanent action, the sort of activity that begins and ends within one agent and is typical of psychic functions (see *Cursus Philosophicus*, "Philosophia Naturalis," I, q. 14, a. 3). John had no separate treatise on metaphysics, but his views on the ultimate character of reality were frequently presented in his explanation of parallel problems (substance, causality, potency) in the "Philosophia Naturalis." The "Theological Course" also contains explanations of problems in speculative philosophy. Cognition, on the sensory and intellectual levels, is explained in terms of a metaphysics of causality (I, q. 1, disp. II, a. 12, n. 4). John was one source of the theory of the distinction between three degrees of knowledge—physical, mathematical, and metaphysical—popularized in the twentieth century by Jacques Maritain.

In his discussion of the gifts of the Holy Ghost (*Cursus Theologicus*, IV, disp. XVII), John had much to say on the relation of knowledge to wisdom. He viewed ethics and political philosophy as speculative sciences and did not write much on practical philosophy. On moral questions he adopted the position called "probabilism"; that is, in moral situations where a person is really in doubt about what he should do, he may solve his doubt by adopting any judgment that has been made by a prudent moralist concerning the proposed action (*Cursus Theologicus*, IV, disp. XII, a. 3, n. 4).

John's writings are useful for their historical information on later scholasticism. He influenced many recent Thomists, notably Maritain, J. M. Ramírez, Joseph Gredt, and Yves Simon.

See also Aristotelianism; Cajetan, Cardinal; Induction; Logic, History of; Maritain, Jacques; Philosophy of Science, History of; Russell, Bertrand Arthur William; Thomas Aquinas, St.; Whitehead, Alfred North.

Bibliography

WORKS BY JOHN OF ST. THOMAS

Cursus Philosophicus, 3 vols. Edited by Beatus Reiser. Turin, 1930–1937; 2nd ed., 1948.

Cursus Theologicus. Opere et studio Monachorum Solesmensium. Paris, Rome, and Tournai, Belgium, 1933–.

The Material Logic of John of St. Thomas: Basic Treatises. Translated by Yves Simon, J. J. Glanville, and G. D. Hollenhorst. Chicago: University of Chicago Press, 1955.

Outlines of Formal Logic. Translated by F. C. Wade. Milwaukee: Marquette University Press, 1955.

WORKS ON JOHN OF ST. THOMAS

Bondi, Eugene. "Predication: A Study Based in the 'Ars Logica' of John of St Thomas." *Thomist* 30 (1966): 260–294.

Doyle, J. J. "John of St. Thomas and Mathematical Logic." *New Scholasticism* 17 (1) (1953): 3–38.

Furton, Edward J. *A Medieval Semiotic: Reference and Representation in John of St. Thomas' Theory of Signs*. New York: Peter Lang, 1995.

Gracia, Jorge J. E., and Kronen, John. "John of Saint Thomas (B. 1589; D. 1644)." In *Individuation in Scholasticism*, ed.

Jorge Gracia. Albany: State University of New York Press, 1994.

Maritain, Jacques. *Philosophy of Nature*. New York: Philosophical Library, 1951. See for a reworking of details of John's theory.

Wolicka, Elzbieta. "The Notion of Truth in the Epistemology of John of St Thomas." *New Scholasticism* 53 (1979): 96–106.

Vernon J. Bourke (1967)
Bibliography updated by Tamra Frei (2005)

JOHN OF SALISBURY
(c. 1115–1180)

John of Salisbury, the scholar, humanist, and bishop, was born at Old Sarum (Wiltshire), England. After primary instruction from a rural priest he went to France to study in 1136. He read dialectic first under Peter Abelard, during the latter's last period at Paris, then under Alberic and Robert of Melun. In 1138 he began the study of grammar under Richard of Arranches, probably at Chartres, where he also studied under William of Conches; at Chartres too he studied rhetoric and part of the quadrivium. In 1141 he took up theology at Paris under Gilbert of Poitiers and Robert Pullen and made the acquaintance of other masters. He was then probably secretary for a short time to Abbot Peter of Celle (1147–1148). He was a member of the Roman Curia, and in 1148 attended the Council of Rheims, where he knew well both Bernard of Clairvaux and Gilbert of Poitiers. That year he was introduced by St. Bernard to Theobald, archbishop of Canterbury, with whom he spent a short time. Between 1149 and 1153 John was a member of the Roman Curia in Apulia and elsewhere and was on terms of intimacy with Pope Adrian IV (Nicholas Breakspear). From 1153/1154 to 1161 he was the trusted secretary of Archbishop Theobald and was one of a distinguished household that included Thomas Becket, Roger of Pont l'Évêque, later archbishop of York, and the Italian lawyer Vacarius. He advised and represented the archbishop and wrote his letters, many of which dealt with business of the Curia.

After Theobald's death, John entered the service of Thomas Becket, to whom he remained a loyal, although not blind, supporter during Thomas's later controversy with King Henry. Accused by King Henry II of encouraging appeals to Rome, John preceded his patron into exile in 1163 and spent some years in Rheims living with Peter of Celle, then abbot of St. Rémy, and working in Thomas's interest with King Louis VII of France. He rejoined Thomas shortly before the latter's return to England in December 1170 and preceded him to Canterbury. John was at dinner with the archbishop when the knights arrived and was present, although perhaps in concealment, at Thomas's murder in the cathedral. He subsequently worked for Thomas's canonization and, in return, was invited by King Louis in July 1176 to become bishop of Chartres. He attended the third Lateran Council in 1179 and died the following year at Chartres, where he was buried.

John was author of a multitude of letters as well as short lives of Anselm and Thomas Becket, the latter a jejune work that is doubly disappointing in view of the writer's literary skill and intimate knowledge of his subject. His *Historia Pontificalis* is a continuation of the *Chronicle* of Sigebert of Gembloux and covers the years 1148–1152. As a scholar he composed the versified *Entheticus de Dog-mate Philosophorum* (1155), a rehearsal of his knowledge of ancient philosophy, as well as the two works on which his medieval reputation rested: the *Policraticus* (The statesman) and *Metalogicon.*

The *Policraticus,* subtitled *De Nugis Curialium et Vestigiis Philosophorum* (Concerning the vain purposes of courtiers and the traditions of philosophers), is a disorderly, rambling work without detailed plan. Dealing in part with such faults and follies of the great as hunting, gaming, dreams, and astrology and with witchcraft, it contains a variety of anecdotes and personal experiences. Books 6–10 deal with the character and duties of a prince, and the work has consequently been called—somewhat misleadingly—the first medieval treatise on political thought. It is, in fact, a sociological study, but it contains a well-known passage on the ministerial function of the prince, who holds the sword in order to perform duties beneath the dignity of the priesthood, which John always considers the superior power, even when emphasizing the virtue of patriotism. The passage shows no clear indication of acquaintance with the almost contemporary teaching of St. Bernard on the possession of two swords by the papacy. In the last book John proclaims the right and duty of citizens to kill a tyrant. The passage has often been quoted in later centuries as authoritative, but it is probably merely an echo of Roman republican rhetoric without any practical application to the world of the twelfth century.

The *Metalogicon* (1159–1160) was written at almost the same time as the *Policraticus.* It is an apologia for true logic, or rather for philosophical training as an introduction to a civilized way of life, contrasted with the technical logic of the schools, which was fit only for sciolists or such careerists as Cornificius, whose name recurs as an unidentified opponent of humane learning. John

recounts his own educational experiences (*Metalogicon* II, 10), with a tribute to Bernard of Chartres and a sketch of his methods, and sets out several current opinions on the nature of universals (ibid., II, 17–20). There are the pure nominalists, such as Roscelin, who held universals to be mere words (*voces*), and Abelard (as John understood—or misunderstood—him), who substituted the term *sermones.* There are those who, like Bernard of Chartres, regarded universals as the Ideas of Plato, and with these may be reckoned Gilbert of Poitiers with his "original forms" (*nativae formae*), while others regarded them merely as a group (*collectio*). John himself adopts an Aristotelian position: Universals are not independent realities, but mental images (*figmenta rationis*) of real kinds (*genera, species*) into which things can be grouped and from which the intellect can abstract those qualities that resemble those of other members of the group. John wrote no systematic philosophical work and declared himself a tolerant skeptic of the Academy, or of what is now known as Late Platonism. Nevertheless, he had great admiration for Aristotle, whom he called "the Philosopher" par excellence; without his New Logic, John maintained, and especially without the *Topics,* dialectic is doomed to be a hit-or-miss affair (*sine eo non disputatur arte sed casu*). Both these considerable works, the *Policraticus* and the *Metalogicon,* and probably also the *Entheticus,* were dedicated to Thomas Becket.

The *Historia Pontificalis,* written over a period of years and finally revised in 1164, is not professedly concerned with thought, although the controversy between St. Bernard and Gilbert of Poitiers is the most important episode contained in it. John gives the theological position of the bishop of Poitiers at considerable length, although he does not clarify the real point at issue.

John was neither a theologian nor an original thinker. He was rather, in the words of Bishop William Stubbs, "the central figure of English learning," or, perhaps more accurately, the writer of the twelfth century who came nearest to the modern critical attitude toward men and their ideas. His celebrated comparison of St. Bernard and Gilbert is the keenest analysis of character and style to appear between the days of Augustine and those of Petrarch, and this is not the only section of his writing that attains such a high level. Moreover, he is the only writer of the twelfth century to pass in review the schools of the day. He read widely, and his style was perhaps the most classical and idiomatic of all medieval attempts to write in imitation of classical models. He lacks the virtuosity and the emotional appeal of Bernard, and his vocabulary and constructions are at times difficult. He is unable to plan or to discard. But his cool judgment and unemphatic language always satisfy the reader. Similarly, his letters, especially those written during the denouement of the Becket affair, display a caustic wit that does not appear in his longer works. He knew on terms of equality almost all the distinguished men of his day. This width of acquaintance he shared with St. Bernard, but whereas the abbot of Clairvaux saw them as figures in black and white, the objects of his emotion and rhetoric, John saw them with a detached, slightly cynical eye that could observe their foibles as well as their gifts. He has won his reward: We speak of his age and society as the age of John of Salisbury.

See also Abelard, Peter; Anselm, St.; Aristotle; Augustine, St.; Bernard of Chartres; Bernard of Clairvaux, St.; Gilbert of Poitiers; Greek Academy; Medieval Philosophy; Petrarch; Plato; Political Philosophy, History of; Roscelin; William of Conches.

Bibliography

The collected works of John of Salisbury were published in Latin as *Joannis Saresberiensis Opera Omnia,* 5 vols., edited by J. A. Giles (Oxford, 1848). A reprint, edited by J. P. Migne, appears in *Patrologia Latina,* Vol. CXCIX. Individual works include the *Entheticus de Dogmate Philosophorum,* edited by C. Petersen (Hamburg, 1843); *Policraticus sive de Nugis Curialium,* 2 vols., edited by C. C. J. Webb (Oxford: Clarendon Press, 1909); *Metalogicon,* edited by C. C. J. Webb (Oxford, 1929); and *Historia Pontificalis,* edited by R. L. Poole (Oxford, 1927).

An English translation of Books 1–4 and parts of Books 7 and 8 of the *Policraticus* may be found in J. B. Pike, *Frivolities of Courtiers and Footprints of Philosophers* (Minneapolis: University of Minnesota Press, 1938); a translation of Books 4–6 and parts of Books 7 and 8 appears in J. Dickinson, *The Statesman's Book* (New York: Knopf, 1927). Other translations are the *Metalogicon,* translated by D. D. McGarry (Berkeley: University of California Press, 1955); and the *Historia Pontificalis,* excellently translated and annotated by M. Chibnall from the Poole edition (Edinburgh and London: T. Nelson, 1956).

For works on John of Salisbury, see C. S. Schaarschmidt, *Johannes Saresberiensis* (Leipzig: Teubner, 1862); and R. W. Carlyle and A. J. Carlyle, *A History of Medieval Humanism in the Life and Writings of John of Salisbury* (London, 1950). A valuable introduction appears in W. J. Miller, H. E. Butler, and C. N. L. Brooke, *The Letters of John of Salisbury,* Vol. I (1153–1161) (Edinburgh and London: T. Nelson, 1955).

David Knowles (1967)

JOHN OF THE CROSS, ST.
(1542–1591)

St. John of the Cross, or Juan de la Cruz, the Spanish mystic and poet, was born at Fontiveros, near Ávila. His family was poor, and as a child he worked in a hospital at Medina del Campo in return for training at the Jesuit school. In 1563 he entered the Carmelite order in Medina, and in the following years he studied at the University of Salamanca. In 1567, the year he was ordained priest, he met St. Teresa of Ávila and planned to start a monastic community in line with the kind of reform she had effected among nuns. Such a community was started in 1568, under the original Carmelite rule and in conditions of great poverty and austerity. This was a prelude to energetic reforming work by St. John and growing opposition on the part of his superiors. In 1577 he was imprisoned at Toledo for eight months and was maltreated. In 1591 he was banished to a lonely monastic house at Úbeda, where he died near the end of that year. His chief prose writings were *The Ascent of Mount Carmel, The Dark Night of the Soul, The Spiritual Canticle*, and *The Living Flame of Love*. His poems have given him a secure place in the history of Spanish literature.

The best-known feature of St. John's mystical writings is his description of the dark night of the soul (or spirit—*noche oscura del espíritu*). The imagery of night is indeed very prominent in his works and was used by him in a variety of senses. By "the dark night" he principally meant the extreme sense of desolation and despair that overcomes the soul after its first illumination by God. This illumination is not the highest state, for eventually the soul will achieve a perfect, lasting union with God—the Spiritual Marriage. The earlier illumination, which St. John called the Spiritual Betrothal, is a "high state of union and love." It thus appears that the dark night is brought on by the deprivation felt when the mystical state of illumination ceases.

St. John saw this dark night in relation to what he called the dark night of sense. This is the purgation of the body and of sense experience, in which the contemplative turns inward from the world. This self-discipline, which involves great asceticism and which constitutes the preliminary training needed for contemplation, culminates in the emptying of the mind of discursive thought and mental images. It is in this state that the Spiritual Betrothal can take place. The dark night of the soul that follows this was explained by St. John as follows.

The soul, despite the Betrothal, still has to endure further purgation, which is psychologically rather than physically painful. This is not due to a change of attitude on the part of God but results from the continued impurity of the soul, which is not able to withstand the glory of the divine illumination. In this situation the theological virtues of faith, hope, and love are essential. Faith enables the contemplative to continue undismayed through the "night"; hope turns the soul toward the future rather than to the memory of deprivation; love turns the soul toward God and men. Ultimately, then, the soul will gain the full union of the Spiritual Marriage. This is described as a complete transformation of the soul in God; and St. John tended to use language identifying the soul with God at this stage, which is contrary to theistic orthodoxy. It is interesting that in his commentary on the poem *The Living Flame of Love* he expressed great unwillingness to write about this, the loftiest state he had experienced. He also said, like other mystics, that the communication of God to the soul is ineffable. However, his use of the imagery of marriage and love indicated that he affirmed the essential distinction between the soul and its Lover.

The attainment of the highest state, according to St. John, is limited to very few persons. Such mystics long for death, after which they may enjoy the Beatific Vision in perpetuity in the next life.

St. John of the Cross and St. Teresa influenced each other, and they are the two most important figures in the history of Christian mysticism in Spain.

See also Asceticism; Illumination; Mysticism, History of; Teresa of Ávila, St.

Bibliography

St. John's works may be found in *Obras de San Juan de la Cruz*, edited with notes by Padre Silverio de Santa Teresa (Burgos: Tip de "El Monte Carmelo," 1929). The works are translated from the above edition by E. Allison Peers as *The Complete Works of St. John of the Cross*, 3 vols. (London: Burns and Oates, 1934–1935; reprinted 1953).

Also see E. Allison Peers, *Studies of the Spanish Mystics* (London: Sheldon Press, 1927), Vol. I.

OTHER RECOMMENDED WORKS

Herrera, Robert A. *Silent Music: The Life, Work, and Thought of St. John of the Cross*. Grand Rapids, MI: W. B. Eerdmans, 2004.

Nieto, José C. *Mystic, Rebel, Saint: A Study of St. John of the Cross*. Genève: Droz, 1979.

Pax, Clyde. "Companion Thinkers: Martin Heidegger and Saint John of the Cross." *Philosophy Today* 29 (1985): 230–244.

Payne, Steven. *John of the Cross and the Cognitive Value of Mysticism*. Boston: Kluwer, 1990.

Pike, Nelson. "John of the Cross: Epistemic Value of Mystic Visions." In *Rationality, Religious Belief, and Moral*

Commitment, edited by Robert Audi. Ithaca, NY: Cornell University Press, 1986.

Pike, Nelson. "St. John of the Cross on Mystic Apprehensions." In *Religious Experience and Religious Belief*, edited by Joseph Runzo. Lanham, MD: University Press of America, 1986.

Sanderlin, David. "Faith and Ethical Reasoning in the Mystical Theology of St. John of the Cross." *Religious Studies* 25 (3) (1989): 317–333.

Tillyer, Desmond B. *Union with God: The Teaching of St. John of the Cross*. London: Mowbray, 1984.

Ninian Smart (1967)
Bibliography updated by Tamra Frei (2005)

JOHN SCOTUS ERIGENA

See *Erigena, John Scotus*

JOHNSON, ALEXANDER BRYAN

(1786–1867)

Alexander Bryan Johnson, an American philosopher and semanticist, was born in Gosport, England, of Dutch-Jewish ancestry. He immigrated to the United States in 1801 and settled in Utica, New York, where he achieved wealth and prominence as a banker. His main interests were intellectual, primarily in theory of knowledge and the problem of linguistic meaning. He published works on the politics of his day, on economics and banking, and moralistic tales for the young, as well as a series of philosophical works.

LANGUAGE AND NATURE

Johnson's preoccupation with language derived from his view that "our misapprehension of the nature of language has occasioned a greater waste of time, effort, and genius than all the other mistakes and delusions with which humanity has been afflicted" (*A Treatise on Language,* p. 300; except where otherwise noted, page references are to the 1959 edition of this work). He found its source in our tendency to interpret nature by language. "My lectures," he wrote, "will endeavor to subordinate language to nature—to make nature the expositor of words, instead of making words the expositor of nature. If I succeed, the success will ultimately accomplish a great revolution in every branch of learning" (p. 40). A rich harvest of philosophically important insights arose from the detailed application of this principle to a wide variety of topics.

Nature, or reality as it appears to us in objects apprehended, is divisible, according to Johnson, into three irreducible classes—the physical (that is, the sensible), the emotional, and the intellectual (thoughts and concepts, which Johnson called "intellections"). Each class includes several subclasses. Sights, sounds, tastes, (tactile) feels, and smells constitute the physical class; the emotions of joy, pain, fear, awe fall into the second class; and concepts (intellections) such as cause, identity, and infinity fall into the third. Words occurring in discourse constitute a subclass of the physical; insofar as they occur in thinking, they are intellectual in nature. The inevitable discrepancy between the practical infinity of natural existences and the necessarily limited number of words of a language results in a one-many relation between words and things (objects of reference). This ambiguity, along with carelessness and ignorance, accounts for the intellectual confusions whose elimination, or at least marking, was the aim of Johnson's lessons on the nature of language.

The terms *physical, emotional,*and *intellectual* throw no light on the nature of the realities they name, but simply refer to them. Only sensing, feeling, and conceiving can inform us what is so referred to. And as the objects, even within each category, are themselves different, acquaintance with some objects of a given kind will not give knowledge of others not confronted. This is not to deny that distinct elements within a given domain resemble one another sufficiently to justify referring to them by a common term. But we err if we suppose that the word *resembles* refers unambiguously to a unique relation. To know that *A* resembles *B* is not to know how it resembles *B*; this can be learned only by specific experience. The elements—the sights, emotions, intellections—that constitute the ultimate referents of significant words are not thought of as mental in the sense of, say, René Descartes or George Berkeley. They are precisely what we find when we confront them, and no words or theories can enlighten us as to their natures. Ultimate meanings can only be *shown* or *had,* never *said.* To understand language we must pass beyond it to the world. Language does not explain the world; the world explains language.

WORDS AND THE MULTIPLICITY OF NATURE

Johnson used his theory to throw light on practically the whole body of traditional philosophical puzzles, most of which are the result of projecting upon nature our misunderstandings of our language about it. For example, we impute to nature a oneness corresponding to the unitary words used to refer to it. Finding nature not always in

agreement with our verbal predications or imputations, we deem this to be ground for impugning our knowledge of its character. The term *gravity* is a verbal unit, but its referents constitute a multiplicity of diverse phenomena. The discrepancy between verbal unity and phenomenal multiplicity leads us to distinguish between gravity and its appearances or manifestations and finally to the view that what gravity is in itself is a mystery, or unknowable. Similar considerations apply to truth, magnetism, cholera, death, the self, and other concepts. "The word gravity names many interesting and important phenomena; but if, in addition to these, we look for gravity itself, we act as ignorantly as the child at the opera, who, after listening with impatience to the musick, singing, and dancing, said, 'I am tired of these; I want the opera'" (p. 77).

In the same vein Johnson criticized Berkeley's view that distance is invisible, by pointing out the obvious fact that "distance" names feels as well as sights. The theory that we cannot see distance derives from our often unconscious restriction of the term to the feel.

Similarly, the question "whether seeing can or not inform us of an external universe, depends on the meaning which we attach to the word external. The question relates to language, not to nature" (p. 63). If *external* is used to refer to what can be tactually felt only, then seeing cannot inform us of an external universe. A sight is not a feel. If we use *external* as referring to a sight, as we frequently and properly do, then seeing can inform us of such a universe.

The origin of theories, according to Johnson, is frequently simply our desire to reconcile these incongruities between what we suppose our language implies and what in fact nature discloses. We invent theories to reconcile the multiplicity of nature to the oneness of language, to supply the unit we suppose must exist but which we fail to find in nature.

KINDS OF MEANING

In his early writings Johnson assumed that if a word had no sensible meaning (referent) it must refer to some inner feeling, or to some other word; otherwise it would be void of meaning, "an empty salvo." Such words as *love* and *hope,* insofar as they do not refer to anything accessible to our senses, would mean other words, their synonyms or definitions, except insofar as they referred to inner feelings. For a person lacking these feelings the word *love* would have only verbal meaning. However, such a person could engage in meaningful discourse involving the word *love* by virtue of being able to explain it by means of other words. He could even have verbal knowledge about love, in the sense that he could make correct verbal deductions from statements containing the term to others entailed by them. In this sense a blind man might have much knowledge of optics, making correct deductions from given premises, even though the sensible meanings, if any, would be beyond his comprehension.

SENSIBLE AND VERBAL SPACE

The distinction between sensible and verbal meaning led Johnson to the difference between physical (sensible) and mathematical (verbal) space, and to the distinction between pure and applied mathematics. The infinite divisibility of space (or matter), not being ascertainable by any of our senses that are cognizant of sensible space, must therefore, he argued, be verbal in nature, since the theory obviously does not refer to any of our inner feelings. Verbal or mathematical space *is* infinitely divisible, our common notion of space entailing such divisibility. The paradox of Achilles and the tortoise is to be explained in terms of this distinction between sensible and mathematical space. In the visual space in which the race is run, Achilles overtakes the tortoise at precisely the moment no light is visible between the two by an observer standing on a line at right angles to the just-touching racers. In mathematical space the process of increasing the denominator of the fraction expressing the "distance" separating them can go on forever. The puzzle is due to our failure to understand that the one-to-one correspondence between the sensible distance and the mathematical distance separating Achilles and the tortoise during the early moments of the race no longer exists at the later stages. When calculation shows that Achilles is one yard behind the tortoise there exists a sensible gap separating them, but when calculation tells us that Achilles is behind the tortoise a distance of one-billionth of an inch nothing in visible space corresponds to this quantity. Hence while still separated in mathematical space they are no longer so in sensible space. The calculations are not faulty. We err in supposing that there must always be a correspondence between the calculated and the observed distance separating them simply because there once was. What is true of mathematical space need not be true of sensible space.

SENSIBLE SPACES

Johnson was aware that there are many different sensible spaces having different properties. Visual space is not identical with tactile space. This fact is important in dealing with certain epistemological puzzles, such as the dis-

crepancy between seen distance and felt distance, seen and felt size or shape, seen location and felt location. The well-known skeptical conclusions derive largely if not entirely from a failure to realize or draw the correct conclusions from the fact that what kind of correlations are found to hold between the diverse referents of such ambiguous terms as *size, shape,* and *location* is a matter purely of experiences—experiences a sensible man will adjust his theories to, but which do not require that he invoke the two-world theory of appearance and reality.

QUALITIES

The question whether secondary qualities are located in things in the external, or physical, world or are subjective representations of objective primary qualities is, according to Johnson, the unhappy result of our failure to realize the ambiguity of spatial prepositions. When we ask for the location of something—whether, for example, the green we see is in the leaf, in our minds, or in the brain—we fail to appreciate that there are several different sensible spaces and that visual, tactile, and olfactory space have each their peculiar properties. In the sense appropriate to visual space the term *in* is correctly used when we say that the (seen) color is in the visual leaf. If we speak of the tangible leaf, the color is neither in the leaf nor not in it. All that can sensibly be said to be in it or not in it is a feel. Colors not being feels, there is no sense to the question if it is based on the presupposition that a sight is a feel or can be felt.

MEANING

Johnson thus understood that in some cases it makes no sense either to assert or deny that a certain object has a certain property, and hence that the law of excluded middle breaks down in certain ways. He made this insight the key to his treatment of many philosophical puzzles.

Since our questions and answers involve sentences, not isolated words or phrases, the meaning of such expressions is of fundamental importance. Declarative sentences, possibly expressing theories, such as "Air has weight," invoked to explain the phenomenon of water rising in a vacuum, gain their referential meaning from the facts, if any, to which they refer. To determine which facts these are, we must ascertain to what phenomena the sentences are attached by a given speaker.

> Pressure, like every other word, possesses no invariable signification, nor any inherent signification. Its signification is governed by the existence to which we attach it. When it refers to the effort of my hand against this table, it names a feel; and when applied to the ascent of water in a vacuum, it names the ascent. If we suppose it names also some insensible operation of the air on the water, this is merely our theory, which signifies nothing; or rather it signifies all to which we refer in proof of the pressure. (p. 227)

The last clause expresses Johnson's view of statement or propositional meaning. A statement means, for a speaker, whatever evidence he adduces or can adduce in support of it. Speaking of Earth's sphericity, Johnson advises us to pay attention to the evidence given in support of it by an astronomer, such as Earth's shadow in an eclipse of the moon or various calculations, and concludes: "After hearing all that he can adduce in proof of the earth's sphericity, consider the proposition significant of these proofs. If you deem it significant beyond them, you are deceived by the forms of language" (p. 129).

This principle of the meaning of propositions is of the type now called the "operational" theory of meaning. Johnson's version, by virtue of his concentration on the referential function of language, implies that propositions change their meaning with every accretion of evidence in support of them. Propositions purportedly about the future must in fact refer to what has already occurred, since one cannot refer to what is not, nor can a speaker refer to what he has not experienced. False propositions must be devoid of (sensible) meaning, since they are false precisely because what they purport to refer to does not exist. Since, however, one is rarely—if ever—unable to adduce some kind of evidence in support of one's assertions, genuinely meaningless or false propositions are extremely rare. In fact, Johnson held that "nearly every proposition is true when interpreted as the speaker interprets it" (p. 133).

Despite the obvious difficulties of this conception of propositional meaning, which needs emendation to allow for what is called "sense" as well as "reference," Johnson was able to suggest some very interesting interpretations of statements that anticipate views now in the center of philosophical controversy.

He held that a theory is a tool whose value is determined by its utility in correlating phenomena already known and enabling us to make true predictions.

He claimed that psychological statements, especially those about other minds, feelings, and thoughts, exhibit duality of meaning. They refer in one interpretation to expressive, that is, external, manifestations; in another to what is supposedly expressed or manifested. This, he

held, explains the dispute concerning the possibility of knowing other minds.

He said that true unrestricted universal propositions are such not because they hold in an infinite number of cases, but because the evidence offered in their support is our failure to find an exception. The statement asserting an exception refers to nothing and lacks sensible meaning—hence the unrestricted scope of the universal.

Typical religious or theological propositions have meaning by virtue of their reference to sacred texts or to inner feelings. It is sensibly, but not verbally, meaningless to assert that the universe either had or did not have a creator.

In his later writings Johnson allowed conceptual as well as verbal meaning to propositions. He came to believe that there are certain "predestinate ideas," concepts or intellections, that express man's intellectual nature. For example, in certain senses of the term, causal connections cannot be sensed, but all men nevertheless think causally. Men likewise impute personal identity to themselves and others, although what is sensibly or emotionally given does not exhibit the implied unity or connection. The verbal meaning that remains when sensible and emotional meanings are eliminated seemed no longer adequate in such cases, and he invoked intellectual meanings, which however are not objective or external; the intellectual words standing in relation to their referents, according to Johnson, as imprecations do to the feelings that give rise to them: "as therefore, the internal organic feeling which prompts an imprecation is the unverbal meaning of the imprecation; so the organism of the intellect that conceives any given words is the unverbal meaning of the verbal conception" (*The Meaning of Words,* p. 202). He thus treated them as expressing certain tendencies of our intellectual nature, though using the language of referential meaning.

THE MEANING OF QUESTIONS

Johnson's anticipation of a form of the operational theory of propositional meaning was accompanied by a detailed discussion of the topic of the meaning of questions. Like Ludwig Wittgenstein he arrived at the view that "the riddle does not exist," that there are no unanswerable questions. Corresponding to the verbal meanings of statements are the verbal questions to which the statements are answers. In every interrogation we must make clear the nature of the answer desired, whether verbal, emotional, sensible, or intellectual. For example, the question, "What is life?" may be answered by a definition or a theory, by an inner experience, or by indicating certain forms of overt observable behavior.

NECESSARY TRUTHS

Concerning necessary, analytical, logical truths, Johnson held to a twofold doctrine. These truths express verbal necessities based on meanings assigned to their constituent words, but these definitions or verbal necessities are themselves based on physical, nonverbal necessities.

> Why cannot the same spot be, at the same time, both white and black? Because the word white implies that the spot is not black. But how came white by this implication? Was it arbitrarily imposed by the framers of language? No. The incompatibility of the two colours is a result of experience. If I assert that the same spot cannot be both white and hard, the proposition will be untrue. Why? Because my senses can discover such a coincidence. No other reason exists. (*A Treatise on Language,* p. 195)

The same reasoning applies to the axioms of geometry. For instance, the transitivity of the relations of equality is ultimately based not on verbal but on physical facts. Nothing will explain why two sticks equal in length to a third are necessarily equal in length to each other except what one finds when one tries to construct two sticks equal in length to a third but not to each other.

APHORISMS

Johnson's works are studded with striking and revealing aphorisms:

> The heathen make graven images—we make verbal ones; and the heathen worship not more ardently the work of their hands, than we the work of our pens. (*A Treatise on Language,* p. 205)
>
> Though we deem any mental phenomenon inexplicable unless we can show it to be analogous to physical operations, we deem the operations of Deity well explained when we can show them to be analogous to mental operations. (p. 263)
>
> We employ words as though they possess, like specie, an intrinsick and natural value; rather than as though they possess, like banknotes, a merely conventional, artificial, and representative value; … We must convert our words into the natural realities which the words represent, if

> we would understand accurately their value. (p. 174)
>
> We can no more exemplify with words that there is a limit to their applicability, than a painter can demonstrate with colours, that there are phenomena that colours cannot delineate. (p. 246)

See also Language; Semantics.

Bibliography

WORKS BY JOHNSON

The Philosophy of Human Knowledge, or a Treatise on Language. New York: G. and C. Carvill, 1828.

A Treatise on Language: or the Relation Which Words Bear to Things. New York: Harper, 1836; edited, with introduction and critical essay, by D. Rynin, Berkeley, CA, 1947; 1947 edition, reprinted 1959.

Religion in Its Relation to the Present Life. New York: n.p., 1841; 2nd ed., retitled *Morality and Manners,* 1862.

The Meaning of Words: Analysed into Words and Unverbal Things, and Unverbal Things Classified into Intellections, Sensations, and Emotions. New York: Appleton, 1854; 2nd ed., 1862.

The Physiology of the Senses: or How and What We See, Hear, Taste, Feel and Smell. New York: Derby and Jackson, 1856.

Deep Sea Soundings and Explorations of the Bottom; or the Ultimate Analysis of Human Knowledge. Boston, 1861, privately printed.

WORKS ON JOHNSON

Johnson's views were ignored in his lifetime and were lost sight of for nearly a century. Despite the recent publication of some of his writings, he remains almost completely neglected. For biographical and bibliographical information and evaluations see the introduction and the critical essay in D. Rynin's 1947 edition of *A Treatise on Language* and M. M. Bagg's *The Pioneers of Utica* (Utica, NY: Curtiss and Childs, 1877). The *Dictionary of American Biography* contains a brief entry on Johnson. For an account of Johnson as an economist see Joseph Dorfman's *The Economic Mind in American Civilization* (New York: Viking, 1946).

David Rynin (1967)

JOHNSON, SAMUEL

(1696–1772)

Samuel Johnson, the American philosopher, was born in Guilford, Connecticut. He studied and taught at the college at New Haven, later called Yale. One of the first colonials to read Francis Bacon, John Locke, and Isaac Newton, he introduced their thought into the college program. In 1722, having abandoned the Calvinism in which he had been raised, he went to England to receive orders in the Anglican Church. On George Berkeley's arrival in Rhode Island in 1729, Johnson paid him several visits, corresponded with him, and became one of his disciples. At the invitation of Benjamin Franklin, Johnson collaborated in the founding of the University of Pennsylvania. In 1754 he helped found King's College, later called Columbia University; he was its first president (until 1763).

Johnson wrote an autobiography and numerous letters, including correspondence with Cadwallader Colden as well as with Berkeley. His philosophical works include *Synopsis Philosophiae Naturalis*, written about 1714; *Logic*, written in 1714; *Encyclopedia of Philosophy*, written in 1714 and revised in 1716; and *Elementa Philosophica*, published by Benjamin Franklin. The *Elementa* was the first textbook in philosophy published in America. It has two parts, "Noetica" and "Ethica"; the "Ethica" had been published alone under the title *A New System of Morality* (Philadelphia, 1746).

Johnson's early works reflect the scholastic Platonism and Calvinistic theology in vogue in the New England colonies during the seventeenth century. The *Encyclopedia*, also called *Technologia sive Technometria*, was a product of his school days and shows the influence of the method and ideas of Peter Ramus. While using Aristotle's physics, it criticizes his metaphysics and ethics as secular and irreligious. Johnson held that there should be no secular science but that all learning should enter into religion and foster it.

Johnson's reading of Bacon, Locke, and Newton broadened and liberalized his thinking. He became an enthusiastic follower of Berkeley's immaterialism, blending with it elements of Puritan Platonism. The English divines, especially Samuel Clarke, influenced him to give up Calvinism and to join the Church of England.

His mature philosophy is contained in his *Elementa Philosophica*. The first part, "Noetica," contains his views on reality and mind; the second, "Ethica," concerns moral behavior. Mind or spirit is defined as intelligent, active being. The objects of mind are ideas or notions. There are no material substances corresponding to our ideas; sensible reality is a system of ideas communicated to us by God as copies of the archetypal ideas in the divine mind.

Arguing against Cadwallader Colden, Johnson maintained that minds are the only agents or active causes; matter is purely passive. Bodies, which are a set of ideas impressed on our minds by God, are entirely inactive and powerless. In Johnson's view, the *vis inertiae*, which Newton attributed to matter, is not a power at all; it is simply

resistance, and resistance is the direct action of God on the human mind.

The existence of God is proved by the presence of eternal truths in our minds. Since these truths do not depend on our minds or on the actual existence of things, they must be communications of an eternally existing and necessary mind or God. We know these truths when our minds are illuminated by the divine mind. God is the fullness of being, and consequently he has the positive perfection of infinity.

Johnson defended the freedom of the will on moral grounds. If human actions are not free, then moral laws, rewards, and punishments are meaningless. God is not the only active cause; human minds are also genuine agents, endowed with freedom to choose or to reject, to act or not to act. Johnson accepted Newton's laws as regulating the movement of inanimate nature, but he insisted that the human spirit is not bound by necessary laws. In opposition to his former pupil Jonathan Edwards, he upheld the freedom of the human will and rejected the Calvinist doctrine of predestination as incompatible with genuine human freedom and as destructive of morality.

Johnson's writings are an important source for the condition of philosophy in pre-Revolution America and for the changes it underwent owing to the impact of eighteenth-century English thought.

See also Aristotle; Bacon, Francis; Berkeley, George; Clarke, Samuel; Determinism, A Historical Survey; Edwards, Jonathan; Franklin, Benjamin; Locke, John; Newton, Isaac; Platonism and the Platonic Tradition; Ramus, Peter.

Bibliography

WORKS BY JOHNSON

Encyclopedia of Philosophy. London, 1731.

Elementa Philosophica. Philadelphia: B. Franklin and D. Hall, 1752.

Synopsis Philosophiae Naturalis. New York, 1929.

Logic. New York, 1929.

WORKS ON JOHNSON

Beardsley, E. E. *Life and Correspondence of Samuel Johnson*. New York: Hurd and Houghton, 1874.

Blau, J. L. *Men and Movements in American Philosophy*, 15–17, 39–40. New York: Prentice-Hall, 1952.

Carroll, Peter N. *The Other Samuel Johnson: A Psychohistory of Early New England*. Rutherford, NJ: Fairleigh Dickinson University Press, 1978.

Ellis, Joseph J. *The New England Mind in Transition: Samuel Johnson of Connecticut*. New Haven, CT: Yale University Press, 1973.

Potkay, Adam. *The Passion for Happiness: Samuel Johnson and David Hume*. Ithaca, NY: Cornell University Press, 2000.

Ryder, John. "Cadwallader Colden, Samuel Johnson, and the Activity of Matter: Materialism and Idealism in Colonial America." *Transactions of the Charles S. Peirce Society* 32 (2) (1996): 248–272.

Schneider, H. *A History of American Philosophy*, 7–11, 21–26. New York: Columbia University Press, 1946.

Schneider, H., and C. Schneider, eds. *Samuel Johnson, President of King's College: His Career and Writings*, 4 vols. New York: Columbia University Press, 1929. Vol. I contains Johnson's autobiography; see Vol. II for his philosophical writings and "The Mind of Samuel Johnson" by H. Schneider.

Armand A. Maurer (1967)

Bibliography updated by Tamra Frei (2005)

JOHNSON, SAMUEL

(1709–1784)

Samuel Johnson, the English man of letters, poet, lexicographer, moralist, and humanist, was born in Lichfield, the son of an indigent bookseller. After his early education at Lichfield Grammar School, he tried schoolmastering for a brief period. In 1728 he entered Pembroke College, Oxford, but was compelled to leave the following year because of lack of funds. As a child he had suffered from scrofula and later from melancholia, a mental illness that plagued him throughout life, at times pushing him to the brink of insanity. In 1735 he married Mrs. Henry Porter, a widow who was twenty years his senior. After more futile attempts at schoolmastering, Johnson set out for London on horseback in 1737, taking with him one of his pupils, David Garrick. A journalist and hack writer par excellence, Johnson wrote for the *Gentleman's Magazine* and in addition produced poetry, essays, biographies, translations, a play, a proposal for a new edition of William Shakespeare, and a proposal for a new dictionary. As a "harmless drudge" he labored from 1746 to 1755 on the *Dictionary of the English Language*, a work that established the practice of elucidating definition of words by quotations from leading authors. Its appearance brought him fame and belated honorary doctorates from Dublin (1765) and Oxford (1775), but little money. Johnson's famous letter of 1755 to Lord Chesterfield repudiated the system of personal patronage. In 1762, however, despite the fact that he had defined "pension" as "pay given to a state hireling for treason to his country," he set aside his scruples to accept a pension from George III.

The Rambler (1750–1752) and *The Idler* (1758–1760) essays, although acclaimed as literature and as statements on morality, were hardly successful financially. The novel *Rasselas* (1759) was well received, as were the edition of Shakespeare (1765), *A Journey to the Western Islands of Scotland* (1775), and finally, *The Lives of the English Poets* (1779–1781). Johnson's political publications, *The False Alarm* (1770), *Thoughts on The Late Transactions Respecting The Falkland Islands* (1771), and *Taxation No Tyranny* (1775), were, on the contrary, mere diatribes and did him no credit. Yet the charge that they were written as repayment for his pension has no foundation in fact. His general theory of politics was close to that of Edmund Burke: conservative, traditional, and distrustful of all popular upheavals.

With a royal pension of £300 a year, poverty and Grubstreeting were over, and Johnson was able to indulge more freely his social proclivities and his desire to travel. The meetings with James Boswell in 1763, and with the wealthy Mr. and Mrs. Henry Thrale in 1764, and the founding of "The Club" in the same year, were happy omens of the new life. Charter members of "The Club" included Joshua Reynolds (who originated the idea), Edmund Burke, and Oliver Goldsmith. Later members of note included Boswell, Garrick, Thomas Warton, Bishop Percy, Sheridan, Fox, Edward Gibbon, and Adam Smith.

Johnson has been immortalized by his great biographer Boswell in *The Journal of a Tour to the Hebrides with Samuel Johnson* (1791), *The Life of Samuel Johnson* (1785), and in present times in the ever increasing number of volumes based upon Boswell's private journals and papers now in the archives of Yale University. Boswell's ability to draw Johnson out in conversation has presented posterity with a wide panorama of the latter's opinions and beliefs. Indeed, it is no exaggeration to say that more intimate details are known about both Johnson and Boswell than about any other persons of that or any previous age. As he grew older Johnson mellowed considerably; he was no longer the irascible, bitter, and not infrequently rude man of earlier years. Although he loved life, he feared death—despite (or perhaps because of) a deep religious faith. As he once put it, life is everywhere "supported with impatience and quitted with reluctance." He died in 1784 after a prolonged and painful siege of the dropsy. His last words are said to have been, *Iam moriturus*, "I who am about to die." He was buried in Westminster Abbey.

RELIGION AND MORALITY

Johnson acknowledged an early predilection for becoming a metaphysician, but instead he became a philosopher, in the wider sense of a thinking man struggling with the problems of life, death, and immortality. A notable excursion into the realm of metaphysics, however, is his 10,000-word critical review of Soame Jenyns's *Free Inquiry into the Nature and Origin of Evil* (1757). The rationalistic optimism inherent in the Great Chain of Being—an optimism wherein whatever is conceivable must exist (a concept justifying the necessity of evil)—was to Johnson morally monstrous as well as metaphysically illogical. It is illogical because however many links there may be in the Chain, from the Godhead at the one extreme to the lowliest atom at the other, it is always possible to conceive of gaps between the links ad infinitum. The morality of justifying poverty and pain as cosmologically necessary was monstrous to a humanist who had personally suffered both poverty and pain. Although God may move in a mysterious way his wonders to perform, it is idle to be told by a metaphysician that in some mysterious way evil in reality is good. It is small comfort to be complacently informed that poverty is merely the want of riches, and that, just as man has animals for food and diversion, so beings superior to man may be privileged to deceive, torment, or destroy man simply for the sake of utility or pleasure. In short, it was Johnson's belief that "life must be seen, before it can be known." His philosophical novel *Rasselas*, a fictional assault on metaphysical optimism, again exemplifies Johnson's favorite admonition, "Clear your mind of cant."

Johnson never systematized his thinking on morality and religion and consequently exhibits many inconsistencies. An ardent Christian and Anglican high-churchman, although not a regular churchgoer, he was forever seeking further evidence and reasons that would bolster his will to believe. He held that every man is entitled to liberty of conscience, but not necessarily the liberty of talking, preaching, or publishing. It is the prerogative of the magistrate to prohibit what he deems politically injurious to the society over which he presides. If the magistrate is morally or theologically wrong in his prohibitions, then truth may suffer. Consequently, the only way in which religious truth can be established is by martyrdom. In the persecution of a martyr, the magistrate is right politically and the martyr is right morally and religiously.

Johnson was afraid of death not only because he was fond of life (even though he held a tragic sense of life), but also because he was acutely aware of the wages of sin. The occasional sermons that Johnson composed for cler-

ical friends and acquaintances (frequently for a fee) are revealing as expressions of his views on specific theological issues. On a deeply intimate level, the "Prayers and Meditations" (begun in 1729 while he was still at Oxford and continued until a few days before his death) provide poignant evidence of repeated resolutions to reform his mode of living (that is, his habitual indolence), to steel himself against religious doubts, scruples, and fear of damnation, and to purge his mind of morbidity and the dread of recurring insanity.

Johnson claimed that we know the distinction between right and wrong by reason; from experience he also knew the difficulties that man encounters in trying to live the life of virtue. Accordingly, he felt the necessity of a mandate from Christian revelation—but never, to be sure, in the sense of the personal "enthusiasm" of seventeenth-century Puritans or eighteenth-century evangelists. He was thus both a rationalist and fideist, but the former tempered by a healthy empiricism and the latter by the requirement of "works." On the one hand, he had unbounded admiration for the Anglican rationalist theologian, Samuel Clarke (1675–1729), and on the other, for the nonjuring pietist and mystic, William Law (1686–1761), neither of whom qualify as orthodox. The sermons of Clarke provided Johnson with rational treatment of thorny theological problems; for example, Law's *Serious Call to a Devout and Holy Life* (1728) augmented faith through a reason that provides spiritual light. Johnson was sufficiently the ethical rationalist (with the qualifications mentioned above) to oppose the nonrational moralists of sentiment or moral sense, such as the earl of Shaftesbury, Francis Hutcheson, Joseph Butler, David Hume, and Adam Smith. Johnson, like Thomas Hobbes, did not consider benevolence or the will to do good to others a natural instinct. Charity, however, as a requisite Christian virtue, Johnson practiced religiously throughout his life. The desire for fame, he maintained in a *Rambler* essay, is basically the desire of "filling" the minds of others. Johnson achieved fame as a didactic writer and moralist who regarded the end and the rites of religion as divinely instituted for "the perpetual renovation of the motives to virtue." This concept of religious need and Christian stoicism received its most memorable poetical statement in one of Johnson's earliest works, *The Vanity of Human Wishes* (1749); it was a statement that was to be reaffirmed countless times throughout his life.

See also Burke, Edmund; Butler, Joseph; Clarke, Samuel; Fideism; Gibbon, Edward; Hobbes, Thomas; Hume, David; Hutcheson, Francis; Law, William; Rationalism; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Smith, Adam.

Bibliography

WORKS BY JOHNSON

The Letters of Samuel Johnson, 3 vols. Edited by R. W. Chapman. Oxford: Clarendon Press, 1952.

The Poems of Samuel Johnson. Edited by D. N. Smith and E. L. McAdam. Oxford: Clarendon Press, 1941.

The Works of Samuel Johnson, 9 vols. Oxford: Talboys and Wheeler, 1825.

Yale Edition of the Works of Samuel Johnson. Edited by Allan T. Hazen et al. 13 vols. to date. New Haven, CT: Yale University Press, 1958–.

WORKS BY BOSWELL

Boswell's Journal of a Tour to the Hebrides with Samuel Johnson. Edited by F. A. Pottle and C. H. Bennet. New York: Viking Press, 1936.

Boswell's Life of Johnson, 6 vols. Edited by G. B. Hill; revised by L. F. Powell. Oxford, 1934–1950.

MODERN STUDIES

Bate, Walter J. *The Achievement of Samuel Johnson*. New York: Oxford University Press, 1955.

Cairns, W. T. *The Religion of Dr. Johnson and Other Essays*. London: Oxford University Press, 1946.

Clifford, James L. *Young Sam Johnson*. New York: McGraw-Hill, 1955.

Hinnant, Charles H. *Samuel Johnson: An Analysis*. New York: St. Martin's Press, 1988.

Hodgart, M. J. C. *Samuel Johnson and His Times*. London: B. T. Batsford, 1962.

Hudson, Nicholas. *Samuel Johnson and Eighteenth-Century Thought*. New York: Oxford University Press, 1988.

Krutch, Joseph W. *Samuel Johnson*. New York: Holt, 1944.

Quinlan, Maurice J. *Samuel Johnson: A Layman's Religion*. Madison: University of Wisconsin Press, 1964.

Schwartz, Richard B. *Samuel Johnson and the New Science*. Madison: University of Wisconsin Press, 1971.

Schwartz, Richard B. *Samuel Johnson and the Problem of Evil*. Madison: University of Wisconsin Press, 1975.

Sushko, S A. "Samuel Johnson as Moralist." *Soviet Studies in Philosophy* 25 (1986): 87–104.

Temmer, Mark J. *Samuel Johnson and Three Infidels: Rousseau, Voltaire, Diderot*. Athens: University of Georgia Press, 1988.

Voitle, Robert. *Samuel Johnson the Moralist*. Cambridge, MA: Harvard University Press, 1961.

Wahba, Magdi, ed. *Johnsonian Studies*, 37–113. Cairo: Socíeté orientale de publicité, 1962. Six essays treat Johnson's religion and morality.

Wilcox, Lance. "The Religious Psychology of Samuel Johnson." *Ultimate Reality and Meaning* 21 (3) (1998): 160–176.

Ernest Campbell Mossner (1967)
Bibliography updated by Tamra Frei (2005)

JOUFFROY, THÉODORE SIMON
(1796–1842)

Théodore Simon Jouffroy, a French commonsense and spiritualist philosopher, was born at Pontets, near Pontarlier, in the department of Doubs. After his preliminary schooling he entered the École Normale in Paris in 1814 and began teaching there three years later. He was attracted to the study of philosophy by Pierre Paul Royer-Collard and Victor Cousin, who were lecturing on the Scottish school. In 1826, Jouffroy published a translation of Dugald Stewart's *Outlines of Moral Philosophy,* and in 1828 he prepared a six-volume translation of the works of Thomas Reid. Jouffroy's rise in the academic hierarchy was rapid; by 1828 he was lecturing at both the École Normale and the Collège de France, where he was appointed professor of Greek and Roman philosophy in 1833. In the same year he was made a member of the Academy of Science.

Jouffroy's interests were varied, covering psychology, aesthetics, legal philosophy, and epistemology, yet he published very little. He is best known for two volumes of miscellaneous essays, *Mélanges philosophiques,* published in 1833, and *Nouveaux Mélanges philosophiques,* which appeared the year of his death.

Jouffroy's ambition was to found a science of psychology based on Scottish philosophy. A survey of the soul's activity revealed to him six different faculties; basic to each of these is a fusion of love of power, curiosity, and sympathy. Upon this foundation rest sensitivity to pleasure and pain, intelligence, "expression," movement, and volition. The soul is thus a community of faculties, all of which must cooperate if the truth is ever to be discovered. It reproduces in the individual that fusion of human souls which is known in the Scottish philosophy as common sense.

It is common sense that alone possesses absolute truth, access to which is denied individuals. Each of us, Jouffroy believed, should attempt to reach the truth by the use of reasoning, but we must accept its conclusions by "a blind act of faith." For none of our faculties is capable of acting in the name of the collective wisdom of the race. Jouffroy held so strongly to this idea that he regarded individual philosophers as mere mouthpieces for the societies and cultures in which they live. As early as 1827 he showed an interest in society as a being having its peculiar influence on the individuals who compose it, but he was never clear about the nature of this being. Jouffroy maintained that if people understood their dependence on the totality of individuals, they would cease to fight with one another and would form a unified fraternal community. This community would be the explicit embodiment of common sense, which already exists implicitly in all human beings.

Common sense expresses itself in self-evident principles that appear in logic and in the dictates of the moral conscience. They are the source of an all-inclusive philosophy illustrated in natural law, which is that system of moral and political principles that underlies the statutes of all nations. Since this system is always consistent, it can act as a test for all truths. What William James, in his *Varieties of Religious Experience,* called Jouffroy's conversion to skepticism stemmed from this idea. For what man other than a mystic could transcend the limits of his individuality to grasp ideas that were overindividual?

In spite of this, Jouffroy maintained that intelligence can apprehend these self-evident principles, just as conscience can apprehend the difference between right and wrong. Here he departed from his theory that men express the ideas of periods and societies and insisted instead that each man's conscience is his sole guide to the good. For the good turns out to be the accomplishment of a man's destiny and evil the failure to accomplish it. A man's destiny is incorporated in his individuality, no two men having precisely the same goals. In general, however, pleasure and pain indicate to a man whether he is fulfilling his destiny, which is apparently the reason men are pleased by different experiences. Unfortunately, Jouffroy's conclusions on this point are lost. And, indeed, he may not have drawn any conclusions, for he was more given to preparatory analyses than to inferences.

Aesthetics, according to Jouffroy, deals exclusively with the nature of beauty. Just as truth is not the possession of any individual, neither is beauty. Beauty does not reflect the character of our life; it is the sublime that takes beauty's place in experience. "The ideas of our present life," Jouffroy said in his *Cours d'esthétique,* "are more familiar to us than the ideas of a more perfect life, and we are consequently less sensitive to beauty than to the sublime." Though the *Cours d'esthétique* consists of notes taken by his pupils and hence cannot be regarded as wholly his, it is clear that the metaphor of the whole of which we know but limited parts dominated Jouffroy's thought. Whether the problem was that of truth, goodness, or beauty, he believed it is the nature of the whole that contains the answer and men are condemned never to possess the answer.

See also Beauty; Common Sense; Cousin, Victor; James, William; Psychology; Reid, Thomas; Royer-Collard, Pierre Paul; Skepticism, History of; Stewart, Dugald.

Bibliography

WORKS BY JOUFFROY

Mélanges philosophiques. Paris: Paulin, 1833.

Cours de droit naturel. 1st ed., 2 vols. Paris: Prévost-Crocius, 1834–1842. 2nd ed., with additional notes by Philibert Damiron, ed., 3 vols. Paris: Hachette, 1843.

Nouveaux Mélanges philosophiques. Paris, 1842.

Cours d'esthétique. Paris: Hachette, 1845.

WORKS ON JOUFFROY

Boas, George. *French Philosophies of the Romantic Period,* pp. 239–250. Baltimore: Johns Hopkins Press, 1925.

Mignet, F. A. "Notes sur Jouffroy." *Memoires de l'Académie des Sciences Morales et Politiques* 25 (1853): 197.

Ollé-Laprune, Léon. *Théodore Jouffroy.* Paris: Perrin, 1899.

Taine, Hippolyte. *Les philosophes classiques du XIX^e siècle en France.* Paris: Hachette, 1868.

George Boas (1967)

JOURNALS, PHILOSOPHICAL

See "Philosophy Journals" in Vol. 10

JUNG, CARL GUSTAV
(1875–1961)

Carl Gustav Jung, the originator of analytical psychology, was born in Kesswil, Switzerland, studied medicine in Basel, and then became an assistant in psychiatry at Zürich, interrupting his stay there to visit and study under Pierre Janet in Paris. He was a pupil of Eugen Bleuler, and he became Sigmund Freud's friend and collaborator for a few years, after having been influenced by his writings. He became the first president of the International Psychoanalytic Society in 1911. In 1914 he broke with Freud, founding his own school of analytical psychology. His earlier studies of association tests and of dementia praecox were followed by an attempt to classify types of personality and by the gradual development not only of a theory of the collective unconscious but also of the implications of that theory for the study of culture and especially for the study of mythology and religion.

Jung traveled widely in Africa, America, and India and collaborated with Richard Wilhelm in Chinese studies and with Kárly Kerényi in the study of mythology. In June 1933 the German Society for Psychotherapy came under Nazi control. Ernst Kretschmer at once resigned from the office of president, and it is regrettable and noteworthy that Jung took his place. Among many other distinctions, he received honorary degrees from Harvard (1936), Oxford (1938), and Geneva (1945).

THEORY OF PSYCHOLOGICAL TYPES

Jung, like Kretschmer, distinguished initially between the extraverted type of personality—sociable, outgoing, and optimistic—and the introverted type—more apt to withdraw from external reality, less sociable, more absorbed in his own inner life. This initial distinction was accompanied by a distinction between four functions of personality—sensation, thinking, feeling, and intuition. By "sensation" Jung meant all that we acquire through sense perception. "Thinking" was used in its familiar meanings. "Feeling" was the capacity for making evaluations of oneself and of others. "Intuition" was the perception of realities that are not consciously perceived; it worked spontaneously for the solution of problems that cannot be grasped rationally.

Types of personality were discriminated in terms of which function is dominant and whether the person is extraverted or introverted. For example, the extravert in whom thinking is dominant will be fascinated by facts and concerned to order them rationally, will tend to underrate the emotions and thus be subject from time to time to uncontrolled and perhaps unrecognized bursts of emotion. The introverted thinking type is one in which facts are never of value for their own sake but only in relation to the creative inner theorizing of the thinker. Both types of thinking are accompanied by an undeveloped feeling function, for thinking and feeling are essentially opposite and even inimical. Sensation and intuition are paired in the same way.

On Jung's view one very rarely finds a person who is a pure example of one of these categories. Most often one function is dominant, although modified by the presence of one of the others. In more complex personalities two functions may coexist in dominance, and very occasionally three, but there will always be at least one function neglected and unacknowledged. Jung's classification into types is, of course, a classification in terms of types of conscious response to the world; however, the notion of parts of the self that are unacknowledged requires some reference to the unconscious.

PERSONAL AND COLLECTIVE UNCONSCIOUS

The personal unconscious consists of those associated webs of ideas and emotions that Jung named complexes, which have been repressed from consciousness because it found them too painful to acknowledge, and also of those perceptions of reality that have never forced their way into consciousness. Each individual's personal unconscious is thus to some extent explicable in terms of his own life history. Even the personal unconscious, however, has features that are common to every individual and do not derive from his personal history.

Consider the contrast between what Jung termed the "persona" and the "shadow." The persona is the socially accepted and socially imposed mask behind which dwells the true ego. The existence of such a mask is an unavoidable necessity, but the ego can fail to achieve self-realization either by identifying itself too strongly with its persona or by not developing an adequate persona at all. The counterpoint to this accepted and exposed part of the personality is the shadow, the rejected and usually imprisoned set of desires, emotions, and attitudes that we personify in dreams as an unpleasant or hostile figure. The shadow is essentially infantile, for it is untouched by the process of maturation or education. The inability to acknowledge one's shadow is always a potential danger to the personality, for the shadow unacknowledged and unrecognized is stronger and more wayward than the shadow recognized and accepted.

Although every individual has a shadow, since the shadow is the product of what his particular consciousness has repressed, it belongs to the personal unconscious. However, beside it in the personal unconscious is found another major force, the image, the image that constitutes the feminine in a man or the masculine in a woman, termed by Jung "anima" and "animus," respectively. The character of the anima is not determined by a man's private history in the way the character of the shadow is; rather, the anima determines how the opposite sex is perceived or misperceived. The anima is an inherited collective image of woman as such. Thus, what matters to the child is not merely how his mother treats him; his experience of the mother is produced both by the mother's actual behavior and by the way his anima determines his view of and feelings about her. Jung connected the anima especially with the function of feeling, the animus with that of thinking, supposing that thinking is more likely to be dominant in the man, feeling in the woman.

The animus and anima belong to the collective unconscious of humankind, along with persona and shadow. They are among the "archetypes," inherited tendencies of psychic functioning contained in the collective unconscious. Other key archetypes are those of the old wise man, the earth mother, and the self. An archetype plays a variety of roles: Not only does it condition the ways in which our conscious experience is formed but also it can appear directly in a number of guises in dreams and fantasies, and the individual may even unconsciously come to be so dominated by one of these images that he might be said to be possessed by it or to identify himself with it. When this happens the personality is itself in danger; it has been taken over and magnified into something that expresses not the individual person but the collective image. This Jung called inflation.

Jung contrasted the self with the ego. The ego is the actual center of consciousness; the self is spoken of by Jung as the center of the unconscious, but clearly it is potentially rather than actually so. Religious visions, dreams, and the magic diagram that Buddhists call the mandala are all images of a possible unity in which the self is at the center. The achievement of this unity by any given individual is a task that belongs especially to the second half of life. In the first half of life the individual is necessarily largely preoccupied with work, marriage, and the bringing up of children; it is when these tasks are mostly accomplished that the individual has to come to terms with himself. Hence the psychological crisis period that occurs in the late forties. At this point the nature of Jungian psychotherapy becomes important.

PSYCHOTHERAPY

According to Jung a neurotic symptom is never to be explained solely in terms of the patient's past. It always represents something positive in the present, an attempt to solve the problems that confront the patient. Jung was prepared to accept that Freud was correct in ascribing many neuroses to the problems arising out of repressed sexuality and that Alfred Adler was correct in ascribing many others to an unrecognized will to power. However, he felt that behind sexuality and the will to power lie other more fundamental causes. Sexuality, for example, is important because it represents the chthonic element in man, an element represented in pre-Olympian Greek religion and in other mythologies. Moreover, the type of neurosis that can be understood correctly, within limits, in Freudian or Adlerian terms belongs characteristically

to the earlier part of life. It arises from the inability to carry through the practical tasks of life.

In psychotherapy the patient comes to acknowledge hitherto unrecognized parts of his personality. Jung believed that free association, as practiced by Freudian analysts, leads not toward but away from the complexes of which we need to become aware. However, more is involved in the therapeutic process that ridding oneself of symptoms, as the patient discovers when he brings what was repressed into view, for example through a new awareness of the significance of his dreams, which function, according to Jung, as compensations for deficiencies in the dreamer's waking life. To rid oneself of symptoms, one has to become aware of the process of individuation, of the need for the creation of a harmonious synthesis of the functions in which the nature of the shadow and the power of the archetypes of the collective unconscious have been reconciled with the demands of the conscious personality.

MYTHOLOGY AND RELIGION

Jung used his central theoretical concept, that of the collective unconscious, to explain not only the occurrence in dreams and the awareness in analysis of contents of the unconscious that could not have been repressed into it by the individual psyche but also the widespread recurrence of the same symbols and themes in widely different times and places in mythologies and religions. Thus, Jung found in the dreams and paintings of patients material that closely resembles that in Eastern religious writings, and in literature and art the archetypal images continually recur. Modern man stands, however, in a peculiar relationship to the contents of the collective unconscious.

Jung held that the increase in scientific understanding has led to a dehumanization of the natural and social worlds. A former unconscious acceptance of natural phenomena, which involved endowing them with symbolic power, has disappeared. To treat thunder, for example, not as the voice of a god but as an explicable phenomenon is to have become alienated from external nature. A loss of belief in gods and demons has produced a lack of awareness of the powers within human nature. Modern man is thus specially a prey to psychological disorders.

It follows that men have a strong need for religious beliefs and experiences, since in religious form they are able to encounter and accept the contents of the collective unconscious. Religious beliefs, Jung conceded, cannot be shown to be true; but he held that they cannot be shown to be false, either. Whether to believe or not is thus a matter of choice, on purely pragmatic grounds. Jung regarded with deep suspicion, as essentially one-sided and distorting, the rationalist traditions of scientific thought. Indeed, he dated the disorientation of modern man partly from the original Christian break with paganism, but more importantly from the Enlightenment.

CRITICISM

Of all Jung's work his classification of types of personality as extravert or introvert has won the widest acceptance. H. J. Eysenck has developed this distinction for use in experimental psychology, and it may well be that other Jungian concepts and theories can also be tested experimentally. However, the linchpin of Jung's theorizing, the concept of the collective unconscious, is so formed that it appears that whereas the existence of the collective unconscious was advanced as an explanatory hypothesis, the question of whether the collective unconscious exists cannot be answered by any possible observation or experiment. That the existence of the collective unconscious is intended as a hypothesis seems clear from the fact that it is avowedly introduced to explain why the same symbols keep recurring in dreams, mythologies, and works of art. However, there are no predictions that we can deduce from this hypothesis other than the vague generalization that such symbols do and will recur—and this, after all, is what the hypothesis was originally intended to explain. Moreover, Jung is open to criticism for treating the collective unconscious not as a theoretical entity to which reference is made in an as yet untested hypothesis but as something whose existence is an established fact. Jung actually asserted that although the facts about personality and the unconscious are undeniable, they cannot, by their very nature, be formulated in such a way as to satisfy the demands of either science or logic.

At the root of the problem lies an ambiguous set of ontological claims. Jung insisted that the contents of the psyche are as real as what exists in the external world. He clearly meant by this more than the obvious, which nobody would be disposed to deny, for example, that there are recurrent patterns of symbolism. But what he meant beyond this remains unclear. Sometimes he seems to have treated the archetypal images as autonomous agents and the collective unconscious as a realm where they dwell. However, his insistence on the inapplicability of the ordinary canons of logic in these matters makes it difficult to press the questions that this seems to raise.

Finally, it is worth noting that we possess no statistical evidence of a worthwhile kind about the efficacy of Jungian psychotherapy. Lacking this evidence, we are forced to conclude that although Jung established a psy-

chological system of some complexity, there are as yet no grounds for believing any of its propositions that go beyond recording empirical data, either as to the nature of personality or as to the process of cure.

See also Adler, Alfred; Freud, Sigmund; Myth; Psychoanalytic Theories, Logical Status of; Psychology; Religion, Psychological Explanations of; Unconscious.

Bibliography

Jung's collected works were published in an English edition, in 19 volumes (London and New York, 1953–1979). See also *Psychological Types,* translated by H. G. Baynes (London: K. Paul, Trench, Trubner, 1923); *Psychological Reflections,* selected and edited by Jolande Jacobi (New York: Pantheon, 1953); *Psyche and Symbol: A Selection from the Writings of C. G. Jung,* edited by Violet S. de Laszlo (New York: Doubleday, 1958); and *The Basic Writings of C. G. Jung,* edited by Violet S. de Laszlo (New York: Random House, 1959).

More recently published editions of Jung's work include *Portable Jung (Viking Portable Library),* edited by Joseph Campbell (New York: Viking, 1971); *Letters of C. G. Jung,* edited by Gerhard Adler and Aniela Jaffé (*Volume 1, 1906–1950,* New York: Routledge, 1973; *Volume I, 1951–1961,* New York: Routledge, 1976); *The Freud/Jung Letters: The Correspondence between Sigmund Freud and C. G. Jung,* edited by William McGuire, translated by R. F. C. Hull and Ralph Manheim (Cambridge, MA: Harvard University Press, 1988; Princeton, NJ: Princeton University Press, 1974); *Collected Works of C.G. Jung,* 21-volume hardcover set, edited by Gerhard Adler, Michael Fordham, Herbert Read, and William McGuire (Princeton, NJ: Princeton University Press, 2000); *Atom and Archetype: The Pauli/Jung Letters, 1932–1958,* edited by C. A. Meier, translated by David Roscoe (Princeton, NJ: Princeton University Press, 2001); *The Earth Has a Soul: The Nature Writings of C.G. Jung,* edited by Meredith Sabini (Berkeley, CA: North Atlantic Books, 2002).

For literature on Jung, see Frieda Fordham, *An Introduction to Jung's Psychology* (London, 1953); E. Glover, *Freud or Jung* (London: Allen and Unwin, 1950); and Jolande Jacobi, *The Psychology of C. G. Jung* (London, 1942; revised 6th ed., New Haven, CT: Yale University Press, 1962), and *Complex/Archetype/Symbol in the Psychology of C. G. Jung* (New York: Pantheon, 1959).

More recently published works on Jung include Deirdre Bair, *Jung: A Biography* (Boston: Little, Brown, 2003); Paul Bishop, editor, *Jung in Contexts: A Reader* (London and New York: Routledge, 1999); Vincent Brome, *Jung* (London: Macmillan, 1978); Northrop Frye, *Anatomy of Criticism: Four Essays* (Princeton, NJ: Princeton University Press, 2000); Ronald Hayman, *A Life of Jung* (New York: W.W. Norton, 2001); Peter Homans, *Jung in Context: Modernity and the Making of a Psychology,*2nd edition (Chicago: University of Chicago Press, 1995); Frank McLynn, *Carl Gustav Jung: A Biography* (London: Bantam, 1996); Mary Ann Mattoon, *Jungian Psychology in Perspective* (New York: Free Press, 1985); Richard Noll, *The Aryan Christ: The Secret Life of Carl Jung* (New York: Random House, 1997); Marian L. Pauson, ed., *Jung the Philosopher: Essays in Jungian Thought (New Studies in Aesthetics, Vol 3)* (New York: P. Lang, 1988); Andrew Samuels, *Jung and the Post-Jungians* (Boston: Routledge & K. Paul, 1985); Andrew Samuels, *A Critical Dictionary of Jungian Analysis* (New York: Routledge & K. Paul, 1986); Sonu Shamdasani, *Cult Fictions: C.G. Jung and the Founding of Analytical Psychology* (New York: Routledge, 1998); Robert Steele, *Freud and Jung: Conflicts of Interpretation* (Boston: Routledge & K. Paul, 1982); Murray Stein, *Jungian Analysis,* 2nd ed. (Chicago: Open Court, 1995); Anthony Stevens, *On Jung* (Princeton, NJ: Princeton University Press, 1999); Richard Sugg, *Jungian Literary Criticism* (Evanston, IL: Northwestern University Press, 1992); J. H. Van Der Hoop, *Character and the Unconscious (International Library of Psychology),* 2nd ed. (Routledge, 1999); Jos Van Meurs and John Kidd, eds., *Jungian Literary Criticism, 1920–1980* (Rowman & Littlefield, 1988); Joseph F Vincie, *C. G. Jung and Analytical Psychology: A Comprehensive Bibliography,* Garland Reference Library of Social Science, v. 38 (New York: Garland, 1977); Steven F. Walker, *Jung and the Jungians on Myth: An Introduction (Theorists of Myth)* (New York: Garland, 1994).

On Jung's concept of extraversion-introversion, see H. J. Eysenck, *Dimensions of Personality* (London: K. Paul, Trench, Trubner, 1947).

Alasdair MacIntyre (1967)
Bibliography updated by Alyssa Ney (2005)

JÜNGER, ERNST
(1895–1998)

Ernst Jünger was a German novelist and cultural critic who, by embracing total war as an exemplary pattern of life, helped to prepare the ideology of the National Socialist revolution of 1933. He was born in Heidelberg and educated in Hanover. In 1913 he joined the French Foreign Legion in north Africa in search of "the extraordinary beyond the social and moral sphere … a zone in which the war of the forces of nature found its pure and aimless expression." This quest for an exotic life in artificially heightened experience revealed Jünger's metaphysical attitudes and anticipated his later pattern of life. Jünger joined the German army at the outbreak of World War I. He fought on the western front and was commissioned, repeatedly wounded, and highly decorated. To him the war appeared "a means for self-realization, a wild upsurge of life … a splendid bloody play which makes the gods rejoice" that offered the key to all essential experience: "ecstasy, sleep and death." After the war Jünger developed his views in a series of brilliant war descriptions: *In Stahlgewittern* (1920); *Der Kampf als inneres Erlebnis* (1922); *Das Wäldchen 125* (1925); *Feuer und Blut* (1925; Adolf Hitler annotated his gift copy); culminating in *Totale Mobilmachung* (1930) and *Der Arbeiter* (1932).

Jünger was also fascinated by modern technology, which had transformed the character of warfare and was creating a new form of industrial society. He envisioned the emergence of a new type of technical elite: the worker-soldier in the nationalized, socialist, militarist-imperialist, and dictatorial state of the future. He also discerned a "new consciousness of reality," nihilist in its relations to traditional values. But although he welcomed the rise of technology as a triumph of man, Jünger deplored its mechanization and dehumanization of life. In the Marxian solution of this problem, the common existential experience of the proletariat leads to class solidarity; its mastery of the tools of production leads to the liberation and human autonomy of the proletariat, which represents humankind. Similarly, Jünger's worker-soldier, simultaneously savior and saved, was to achieve the collective salvation of the rotting democratic-humanist society.

Technology, however, was inseparably bound up with war, "a fiery marriage between the spirit of chivalry and the severe coldness of our forms of work." The world of factories and calculated organization, of production, and of transport finds its true measure in battle. "The battle is a tremendous touchstone of industry, and victory marks the success of a competitive effort which knows how to work more quickly and ruthlessly." The individual worker-soldier finds his liberty in accepting the necessity to be part of "the greater force. Here one can only drift and be formed under the grip of the *Weltgeist.*" The worker-soldier type thus replaced the individualist personality of the nineteenth century. Technology became both the means and the end of human endeavor—the means because it procured mastery over others, the end because the old values were dead, and collective power, the product of technology, was equated with value: "Technology and ethos have become synonymous."

Jünger's "national-Bolshevist" conception of technology provided a scintillating and heady approach to totalitarianism, an approach based also on his belief in inexorable historical trends and his romantic conviction that the individual finds fulfillment only by sacrificial immersion of himself in the whole. Jünger promised redemption for the sacrifice of the obedient soldier but showed scant sympathy for that of the Socratic nonconformist. His *Der Arbeiter* is thus less a sociological interpretation of his times than the revelation of a political myth, a clarion call that exerted a wide influence in Germany among the bewildered generation of the 1920s.

Jünger's misinterpretation and rejection of liberalism prevented his playing a constructive part as a citizen and caused him to be a destructive intellectual force. An anarchic pride in his own independence, however, saved him from effective collaboration with National Socialism. Jünger first parted ways with the Nazi Party in 1929, when he backed a terrorist peasant movement opposed by Hitler. Between the lines of his novel *Auf den Marmorklippen* (1939) he criticized the prevailing tyranny, but he took no part in active resistance to the regime. He again fought in the German army in 1940, although he suffered misgivings as a member of the army of occupation in France and Russia. These feelings found expression in *Strahlungen* (1949), Jünger's journals from 1939 to 1949, in which he corrected certain of his former tenets and, in a fashion, held out a hand to Western values and to the Christian religion. In his novel *Heliopolis* (1949) he took up once more the problems raised in *Auf den Marmorklippen. Heliopolis* contained an indictment of a closely knit totalitarian order but, at the same time, preserved Jünger's distance from Western rationalism and liberalism. The same theme recurred in *Der Waldgang* (1951); *Gläserne Bienen* (1957), which again expressed Jünger's fascination with technology; and *Der Weltstaat* (1960), which called for international political unity as a historically determined necessity.

Jünger conceived of the writer as a seer and pathfinder. His diagnosis of his times was, however, based on an untrained and intuitive sociological and economic knowledge, poetical and pretentious rather than scholarly. His widely acclaimed concept of the Gestalt, or Typus, of the worker offered no methodological advance and in substance was merely ideological. Jünger's significance was as a spokesman of the powerful romantic strand in the German intellectual tradition that unites elements of *Naturphilosophie,* Neoplatonic mysticism, and a Protagorean theory of knowledge with the negative aspects of Jean-Jacques Rousseau's and Edmund Burke's critiques of society and the Enlightenment. In its modern representatives, such as Jünger and Oswald Spengler, such thinking leads to a rejection of the rational, abstract, and mechanical achievements of civilization, the "high-treason of the intellect against life," and to the extolling of the instinctive, oceanic "night side" of life. Although not original, Jünger's philosophy was presented in a highly personal manner and in an evocative style, drawn from military language and a minute observation of nature. As a novelist, however, he did not succeed in creating concrete character.

See also Burke, Edmund; Enlightenment; Fascism; Gestalt Theory; Liberalism; Philosophy of Technology; Rousseau, Jean-Jacques; Spengler, Oswald.

Bibliography

The collected works of Ernst Jünger were published in a definitive and partly revised edition as *Sämtliche Schriften,* 10 vols. (Stuttgart, 1960–1964).

English translations of Jünger's writings include *The Storm of Steel,* translated by Basil Creighton (Garden City, NY: Doubleday, 1929); *Copse 125,* translated by Basil Creighton (London: Chatto and Windus, 1930); *On the Marble Cliffs,* translated by Stuart Hood (Norwalk, CT: New Directions, 1947); *Peace,* translated by Stuart Hood (Chicago: H. Regnery, 1948); and *African Diversions,* translated by Stuart Hood (London, 1954).

For writings on Jünger, see G. Loose, *Ernst Jünger, Gestalt und Werk* (Frankfurt am Main: V. Klostermann, 1957), pp. 371–380; Karl O. Paetel, *Ernst Jünger in Selbstzeugnissen,* Vol. 72 in Rowohlt's Monographien (Hamburg: Rowohlt, 1962), pp. 168–175; Hans Peter Schwarz, *Der konservative Anarchist: Politik und Zeitkritik Ernst Jüngers* (Freiburg: Rombach, 1962), pp. 309–315; and J. P. Stern, *Ernst Jünger* (New Haven, CT: Yale University Press, 1953).

H. O. Pappé (1967)

JUNGIUS, JOACHIM
(1587–1656)

Joachim Jungius, of Lübeck, represents the German counterpart to Galileo Galilei in Italy, René Descartes in France, and Francis Bacon in England as an innovator in science and philosophy. Unlike these men, Jungius did not achieve an international reputation; even among scholars, interest in him has been largely confined to Germans, whose curiosity has been whetted by Gottfried Wilhelm Leibniz's enthusiastic praise of his merits as a philosopher. But Jungius exercised a wide personal influence in Germany as an active teacher. Furthermore, like Bacon, he envisaged a scientific society that would promote the welfare of humankind: Jungius actually organized a group called the Societas Ereunetica, whose stated objective was to promote sound science and combat false opinions. This group, with its stress on mathematics and logic as an antidote to metaphysical and mystical speculation, invites comparison with the Vienna circle of the twentieth century as well as with the Royal Society. Although Jungius has been linked by legend with the Rosicrucians, there is no evidence whatsoever to support this conjecture, according to G. C. Guhrauer.

Jungius studied at Rostock and Giessen before traveling to Italy to take a medical degree from Padua in 1618. During the early seventeenth century, philosophy in the German schools relied to a large extent on Aristotelian compendia drawn up by Philipp Melanchthon or by Peter Ramus, supplemented by metaphysics of the Suarezian type. Both traditions were diligently studied by Jungius before he rejected them. Jungius had taught mathematics at Rostock; hence, he must have found the atmosphere of Padua congenial, because of the school's emphasis on a research-oriented natural philosophy, medical training, and mathematics.

On his return to Germany, Jungius resumed his teaching duties, presiding over disputations in which Aristotelian views in physics were mercilessly criticized. He was dissatisfied with the doctrine of the four elements and wished to substitute for it an atomism that, he believed, would be confirmed by future research but which, in any event, offered a more promising hypothesis. Jungius considered atomism more sound from the methodological point of view since it did not require the postulating of entities ("forms") to explain the rise of all sorts of new qualities in things. "Democritus was an Ockhamist," he remarked.

In 1625 Jungius began teaching medicine at Helmstedt, stressing the value of Galen, whose logical empiricism he found congenial. In 1628 Jungius took an unusual step—he left university teaching to assume charge of a secondary school in Hamburg. Jungius rescued the school from the decline into which it had fallen, sending out from it students trained to a high level of critical analysis. For them Jungius composed the famous "Hamburg Logic" (1638), called by Heinrich Scholz "the most significant logic of the seventeenth century," eclipsing the better-known Port-Royal logic. Jungius's critical presentation of traditional logic shows what the more sophisticated neo-Aristotelian contemporaries of Descartes were thinking about causation, induction, and the nature of scientific demonstration. Jungius was also interested in natural history; he and his students collected plants, minerals, and fossils. His botanical views attracted the attention of Johann Wolfgang von Goethe, who planned a monograph about him.

Most of Jungius's writings in manuscript were destroyed by fire in 1691. The works posthumously published under his name, such as the *Doxoscopiae Physicae Minores* (Hamburg, 1662), were compilations made by students. Such writings as we do have bear the stamp of an active and critical mind, free from any mystical leanings and directed toward a scientific reconstruction of philosophy.

See also Atomism; Bacon, Francis; Descartes, René; Galen; Galileo Galilei; Goethe, Johann Wolfgang von; Leibniz, Gottfried Wilhelm; Logic, History of; Logical

Positivism; Melanchthon, Philipp; Ramus, Peter; Suárez, Francisco.

Bibliography

WORKS BY JUNGIUS

Logica Hamburgensis. Edited by R. W. Meyer. Hamburg: Augustin, 1957.

WORKS ON JUNGIUS

Ashworth, E. J. "Joachim Jungius, 1587–1657, and the Logic of Relations." *Archiv Fuer Geschichte der Philosophie* 49 (1967): 72–85.

Guhrauer, G. C. *Joachim Jungius und sein Zeitalter*. Stuttgart: Cotta, 1850. The definitive biography.

Wohlwill, E. "Joachim Jungius und die Erneurung atomistischer Lehren im 17. Jahrhundert." *Abhandlungen aus dem Gebiete der Naturwissenschaften*. Hamburg, 1887. Vol. X, pp. 3–66.

Neal W. Gilbert (1967)
Bibliography updated by Tamra Frei (2005)

JURISPRUDENCE

See *Analytical Jurisprudence*; *Historical School of Jurisprudence*; *Legal Positivism*; *Philosophy of Law, History of*; *Philosophy of Law, Problems of*

JUSTICE

Justice names not a thing, but a property of things. It makes sense therefore to focus the explication on the adjective "just"—or, better still, "unjust." Doing so facilitates clarification of how justice judgments are distinctive within the larger realm of moral judgments, and the even larger universe of evaluative judgments.

The application of ordinary empirical predicates, such as "tree" or "hard," is two-tiered: based on a definition and empirical facts. Any dispute about whether such a predicate applies thus reduces to linguistic and empirical differences. Such a dispute can be resolved by agreeing on a definition and settling the empirical disagreement.

Evaluative predicates, by contrast, have this special feature that their application is only conditioned, not determined, by their definition and the empirical facts. Thus, people can disagree about whether a painting is beautiful, even if they use this predicate in exactly the same sense and also agree about all empirical features of the painting. In such cases it may be said that they have different conceptions of beauty.

The same holds for moral predicates. Despite agreement on all relevant empirical facts, people disagree about whether something is praiseworthy or not. Such a disagreement could stem from one party's failure to understand the meaning of the word; but more typically the disputants know what the word means, and their disagreement shows then that the empirical facts and the meaning of the word together do not determine its correct application. Those who judge it praiseworthy to teach children through beatings are morally mistaken; they need not linguistic instruction to improve their understanding of "praiseworthy," but a good discussion about how children should be educated.

A dispute over the application of an evaluative predicate such as "just" may thus be due to differences of three kinds: *linguistic* differences about its meaning; *theoretical* differences about which substantive conception or criterion of justice should guide its application; and *empirical* differences about the evaluated object.

The boundaries between these three kinds of differences is not sharp and may shift over time, as Wittgenstein memorably describes:

> It might be imagined that some propositions, of the form of empirical propositions, were hardened and functioned as channels for such empirical propositions as were not hardened but fluid; and that this relation altered with time, in that fluid propositions hardened, and hard ones became fluid. The mythology may change back into a state of flux, the riverbed of thoughts may shift. But I distinguish between the movement of the waters on the riverbed and the shift of the bed itself; though there is not a sharp division of the one from the other.… And the bank of that river consists partly of hard rock, subject to no alteration or only to an imperceptible one, partly of sand, which now in one place now in another gets washed away, or deposited. (Wittgenstein 1969, §§ 96, 97, 99)

Not discussed by Wittgenstein, the boundary between linguistic and theoretical differences (in the case of evaluative predicates) is fuzzy and fluid in part because the controversies among competing conceptions reflect back upon the concept. Each conception seeks to introduce a certain order and unity. And sometimes elements of such a conception are widely found to be so convincing that they harden into an element of the concept.

CONDUCT-GUIDING FUNCTION

The predicates "just" and "unjust" have not merely an evaluative but also a normative—hence conduct-guiding—function. Calling a possible action or law unjust is to oppose its implementation. To be sure, people do make justice judgments about the distant past, about hypotheticals and fiction. But even these judgments imply oughts—for example: that the Athenians ought not to have attacked neutral Melos; that it would have been all right for them to attack a Melos allied with Sparta; or that Angelo (in Shakespeare's *Measure for Measure*) ought not to have demanded Isabella's virginity for the life of her brother.

The words "rational" and "irrational" are also conduct-guiding in this sense. Yet there is a difference. Whether an action is rational or irrational depends on the ends of the actor. Whether it is just or unjust is independent of these ends. Insofar as even an agent's ultimate ends can be criticized as irrational (Parfit), this difference becomes less deep. Like morality, rationality can then deliver unconditional judgments: an action is irrational if it was performed in pursuit of an irrational end.

Generally, claims of injustice are meant to evoke emotional rejection: condemnation, outrage, resentment. But this may be part not of the meaning of the word but of the usual pragmatic context of its employment.

JUDICANDA

The meaning of predicates is partly determined by their domain of application. Anyone who understands the meaning of "beautiful" and "just" knows how the domains of these predicates differ—knows, for instance, that a painting can be (un)beautiful but not (un)just whereas a patent regime can be (un)just but not (un)beautiful.

Things to which evaluative predicates are applicable can thus be called judicanda, from the Latin *judicandum*: that which is to be judged. The judicanda of justice may be categorized under four headings:

(a) individual and collective actors; that is, individuals as well as organized and unorganized groups such as a family, firm, state, or mob;

(b) the conduct of such actors, their actions and omissions;

(c) social rules, such as laws, social institutions, and conventions;

(d) states of affairs and events, such as the fact that some are much worse off than others or that some good persons suffer while some bad ones enjoy good fortune.

Because of the normativity of justice assessments, judicanda of the first two categories have a certain primacy. It is ultimately actors—and their conduct—who bear responsibility for the justice of social rules and, in part through these rules, for the justice of states of affairs. To be sure, sometimes people complain of injustice in states of affairs beyond human control. Such complaints may have a religious context; but they may also, where religion has faded, constitute a purely evaluative (non-normative) use of "unjust."

One may think that this list is underinclusive, that human feelings constitute a fifth judicandum as exemplified by the anger of Achilles and its critique as unjust. But such locutions are better understood as meaning to assess the person and her or his conduct. It is Achilles whose justice is in question—on account of his tendency to get angry allegedly without good reason. Thus, talk of unjust feelings is imprecise, meant to call into question the justice of the person who feels this way.

One may think that the first category reduces to the second in a similar way: that justice assessments of actors really mean to judge these actors' conduct, and that the list is therefore overinclusive. But the nature of such a reduction would be controversial: Do assessments of actors refer to their actual conduct, to their conduct dispositions, to their intentions (toward conduct)? Such controversy shows that the proposed reduction does not hold as a matter of meaning. It can be a substantive element in a conception of justice. But, to state it as such, actors must be shall be a separate category of judicanda.

The same holds for the possible reduction of the third category to the second. One may think, for instance, that social rules are unjust if and only if it would be unjust to (help) impose them. But here, too, there are various other ways of formulating the reduction; and it is a substantive moral question which of these formulations, if any, is correct.

In view of the diversity of judicanda, one may dispair of Plato's grand ambition in the *Republic* and conclude that it is impossible to give a general characterization of the concept—let alone a general conception—of justice that plausibly covers all four categories of judicanda. Such doubt can only be dispelled by setting forth a plausible general structure for the concept (building on the general points already made) and some main hypotheses toward a unified conception of justice. This attempt is made in what follows.

A UNIFIED CONCEPT OF JUSTICE

In contrast to many one-place evaluative predicates (beautiful, good, conscientious, modest), "unjust" has an essential second place. It is indeed often used as a one-place predicate (as in "this man is unjust"). But reference to a second place is always implicit. In this respect, "unjust" is like the predicate "mother." To say that she is a mother is to say that she is the mother of someone. One cannot understand what it is to be a mother without understanding what it is to be the mother of someone. Likewise here: To call a man unjust is to say that he is unjust to others. One cannot understand what it is for a judicandum to be unjust without understanding what it is for it to be unjust to someone. Injustice conceptually requires *recipients*: those who receive unjust treatment from the judicandum.

To be sure, one may call a proposed law unjust even if it fails to pass and thus never treats anyone unjustly, and one may call a woman unjust, even if she treats no one unjustly, on account of her intentions or dispositions. But even in such cases an implicit reference to recipients is essential: one implies that the proposed law, if adopted, would treat some persons unjustly, and that the woman intends, or is disposed, to treat others unjustly.

Recipients need not necessarily be victims. It is possible that the injustice of an action or rule entails that some are treated better than they should be. In such cases, there may be other victims—people who, because of the action or rule in question, are unjustly treated worse than those who were treated too well. But when there are no such victims, can one then still speak of (victimless) injustice? To use an example from Kant's *Metaphysics of Morals* (p. 333): Would it be unjust if a society about to dissolve itself were to set free a convicted murderer from its jail? (Kant suggests that this would be unjust—but not because his execution is owed to the murderer himself, but because it is owed to his victim(s) and to all of humankind.) The concept of justice does not settle this question; different conceptions of justice will answer it differently.

The concept of justice involves an essential third place in that the notion of recipients, of those who receive just or unjust treatment, presupposes benefits and burdens that these recipients either should but do not have or do but should not have. What sorts of benefits and burdens these are, and whether they are understood in absolute terms or relative to what other recipients have, varies with judicandum and context.

Various aspects of justice have traditionally been distinguished. These distinctions can be displayed in three dimensions.

DIMENSION ONE: FIRST-ORDER AND PROCEDURAL (IN)JUSTICE. In a first dimension, one can distinguish assessments of a particular allocation of benefits and burdens (*first-order* justice) from assessments of the way in which such an allocation comes about (*higher-order* or *procedural* justice, sometimes also called *fairness*). Thus a judicial divorce may be unjust on account of the ordered division of marital property, and it may also, and independently, be unjust because avoidably only one of the two parties was allowed to speak. The latter injustice—a violation of the classical precept *audiatur et altera pars*—involves (relative) second-order benefits and burdens: the advantaging of one party and the disadvantaging of the other in the decision-making process. In other cases, violations of procedural justice may involve absolute higher-order benefits or burdens, as when exculpatory or incriminating evidence or witnesses are arbitrarily excluded from a criminal trial.

Procedural justice plays an important role even outside jurisprudence. Rules and decisions about the awarding of honors, contracts, jobs, promotions, and university admissions may violate procedural justice. There are blatant cases, as when a coveted job goes to an insider's spouse without advertisement or search. In other cases it is controversial what the requirements of procedural justice are, exactly. The most important such controversies in recent decades have centered around the question of whether certain selection processes may or should favor people of a particular color, gender, or ethnicity so as to compensate for, or to help overcome, group disadvantages due to past or present discrimination.

In the realm of politics, especially, there may be third-order and fourth-order judgments of justice. This is the case when the rules or the participants for some procedure are selected through a metaprocedure. Thus, the electoral law of a country may be unjust independently of whether this has any effect on the composition of its legislature and, thereby, on its legislative output.

Even in private life one may find violations of procedural justice. It may be unjust if on the basis of a nasty rumor one discriminates against a person, or perhaps even repeats the rumor to third parties, without giving this person an opportunity to respond to the allegation. Insofar as the injustice of such conduct is independent of the truth of the rumor in question, it involves once more a higher-order burden.

DIMENSION TWO: FORMAL AND MATERIAL (IN)JUSTICE. Perpendicular to this first dimension is the common distinction between formal and material (in)justice. *Formal* justice requires that relevantly similar cases be treated similarly. There is much room for controversy about whether particular cases are relevantly similar and about what is to count as similar treatment. The requirement of formal justice is therefore primarily a demand for justification. The complaint that relevantly similar cases are being treated in a dissimilar way demands a justification, showing that the treatment was not dissimilar or else was appropriately responsive to dissimilarity among cases. This requirement of justification holds across the first dimension: The procedures through which benefits and burdens are allocated to recipients as well as these allocations themselves must satisfy formal justice. It must be justifiable that one accused was acquitted and another convicted (first-order justice), say, and also that one accused but not another was provided an attorney at public expense (procedural justice).

Judicanda can be gravely unjust even when they clearly do treat similar cases alike. Thus, parents who beat all their children without cause violate material justice, as do judicial systems that deny all accused rights to speak, to appeal, and to consult legal counsel.

One might think that justice reduces to material justice: When dissimilar treatment of cases cannot be justified, this merely shows that the treatment of some of these cases is materially unjust, independently of how the others are treated. But this view is surely not implicit in the concept of justice. And it is substantively implausible. In many cases, justice is comparative. To illustrate, there are indefinitely many schedules according to which individuals and businesses might be taxed. While the demands of material justice disqualify a wide range of these options, they do not mandate one uniquely just tax code. Still, it would clearly be unjust to impose diverse materially just tax codes—some more advantageous than others—upon the various households or businesses of a single jurisdiction. Unjust is here the unequal treatment of taxpayers, which is a violation of formal justice.

DIMENSION THREE: VARIOUS DOMAINS OF MATERIAL (IN)JUSTICE. *Material* justice imposes various requirements whose content is controversial. This multiplicity can be ordered in a third dimension by distinguishing various domains. Thus *distributive* justice deals with access to scarce resources—from the division of a pie to the structure of an economic order that regulates access to raw materials and the distribution of the jointly created social product. *Commutative* justice governs exchanges, which may be faulted for first-order flaws, as when the items exchanged are not equivalent, or for higher-order (procedural) flaws, such as excessive inequalities in information or bargaining power. *Corrective* or *restitutive* justice is concerned with how to make up for violations of social and moral rules and how to deal with the costs such violations cause. *Retributive* justice, finally, deals with the ascertainment and punishment of such violations.

In each of these four domains there are procedures, which may be formally or materially unjust, as well as particular allocations of benefits and burdens, which also may be formally or materially unjust. The traditional three-dimensional schema, as here reconstructed, thus contains sixteen boxes for sorting particular justice assessments. In the first dimension, the assessment is either first-order or higher-order. In the second dimension, it concerns either formal or material justice. And in the third dimension, it falls into the domain of distributive, commutative, corrective, or retributive justice.

This conceptual structure leaves open the possibility that some of these boxes may contain no substantive constraints. For example, there may be no first-order material demands of commutative justice if any exchange, no matter how lopsided, is morally acceptable so long as it has been performed freely by the exchanging parties without unfair inequalities in bargaining power or information. But this is not a conceptual point, but rather a substantive claim, affirmed by some (especially modern) and denied by other (especially medieval) conceptions of justice.

This complexity can be increased to sixty-four because—perpendicular to the three dimensions in which the character of diverse justice assessments can be differentiated—there is still the distinction between four categories of judicanda (introduced earlier) to which any such justice assessments can be applied: namely, individual and collective actors, their actions and omissions, social rules and institutions, and states of affairs and events.

In some accounts, *international* justice is given as a separate domain that, traditionally, is heavily focused on the use of force and, in particular, on just and unjust causes for going to war (*ius ad bellum*), and on just and unjust ways of fighting a war (*ius in bello*). (In his *Metaphysics of Morals* [pp. 343, 347–349], Kant adds a third theme: the just way of concluding a war so as to lay the foundation for a stable peace [*ius post bellum*]) It seems more appropriate, however, to think of states as one class

of collective actors among others. It may indeed be claimed that states are a special class of actors and that their conduct and relations are subject to distinctive justice requirements. But this is a substantive moral claim that should not be prejudged in the conceptual explication.

A state is understood not merely as a collective actor, but also as a comprehensive system of social rules and institutions enforced in a particular territory. A state can therefore be criticized as unjust not merely on account of its conduct toward or institutional relations with outsiders, but also on account of its internal institutional order—for instance, on account of how it distributes rights and duties, regulates and taxes economic cooperation, and enforces its laws against its members. So understood, states exemplify one kind of institutional scheme among others, whose special status once again should not be prejudged in the conceptual explication.

TOWARD A UNIFIED CONCEPTION OF JUSTICE

The preceding sections were meant to sketch the structure of the contemporary concept of justice and thereby to characterize the linguistic consensus that underlies current debates about questions of justice. Even these general thoughts may not be wholly uncontroversial. But they are much less controversial than the thoughts to follow, which are meant to describe plausible elements of a conception of justice. Should these elements be found convincing, they might gradually become elements of a clearer and more unified concept of justice. But they do not now fully accord with the various understandings of this concept that are dominant in contemporary public and academic discussions.

It is possible to begin once more from the question, briefly raised in (1): What is specific about justice, how justice judgments are distinctive within the larger realm of moral judgments; what is being said with the complaint that some judicandum is unjust, over and above the claim that this judicandum is immoral (morally flawed)? What follows are four hypotheses toward answering this question.

FIRST HYPOTHESIS: INJUSTICE INVOLVES ABUSE OF MORALITY ITSELF. Morally flawed judicanda are unjust only if they involve an abuse of morality itself—that is, only if they appear with a moral pretension they do not live up to. Unjust is someone who is prepared to violate moral principles she herself likes to appeal to. Unjust is someone who allows his official conduct to be influenced by bribes even while he pretends to be an impartial judge, umpire, or mediator. Unjust is a beating falsely presented as deserved punishment. And unjust is legislation designed for the benefit of a small minority and yet claimed to impose moral obligations on the oppressed population. In all these cases, the injustice comes about through the false moral claim: Without the pretense of deserved punishment, the beating is wrong but not unjust. Without the claim of moral authority (that compliance is morally required), coercive rules can be wrong, but not unjust.

According to this first hypothesis, then, justice is a part of morality that defends the authority and dignity of morality itself. This would explain the central place justice is thought to occupy within morality. Unjust are only those judicanda that do not merely—openly or covertly—violate morality, but also appear under color of morality.

SECOND HYPOTHESIS: JUSTICE MAY IN PRINCIPLE BE ENFORCED. A second essential mark of injustice may be that it involves a violation of a right and, more precisely, of a moral right. (Injustice need not violate positive law. To the contrary, because positive law may be unjust or otherwise immoral, conduct it permits or even requires may still be unjust.) Violations of rights are those moral infractions that may in principle be averted through the use or threat of coercive force. The main implication of the second hypothesis therefore is that any injustice may in principle be forcibly averted. Justice may be enforced.

The expression "in principle" flags two qualifications. First, it is not meant that anyone is permitted forcibly to avert any rights violations. In a well-ordered society, for instance, only the police and the courts are permitted to use force in response to rights violations—ordinary citizens may use or threaten force only in urgent emergencies. Second, there is no permission to use force regardless of the morally significant costs of doing so. Defensive force may be grossly disproportionate and therefore impermissible. And efforts forcibly to avert a rights violation may also be impermissible when they carry the risk of triggering much graver violations of the rights of third parties (a point well illustrated in Heinrich von Kleist's story "Michael Kohlhaas"). That any rights violation may in principle be forcibly averted thus means only that those responsible for it and those profiting from it have no moral right against having this violation blocked by force. Here justice contrasts with other moral qualities. No moral rights are violated by judicanda lack-

ing in generosity, civility, charity, humanity, decency, kindness, courtesy, mercy, or beneficence. Enforcing these virtues is wrong in principle and violates moral rights of those against whom they are enforced.

THIRD HYPOTHESIS: INJUSTICE HARMS IN VIOLATION OF NEGATIVE DUTIES. Related to the preceding, another essential mark of injustice may be that it always involves a violation of negative duties on the part of those who are responsible for the moral quality of the judicandum in question. What is at issue here, roughly speaking, is the traditional distinction between moral infractions that harm recipients and those that merely fail to help them. However this distinction may be made precise, the third hypothesis is simply that a claim of injustice always involves a claim of undue harm.

This hypothesis is related to the second hypothesis insofar as the two distinctions may wholly or largely coincide. That they coincide precisely would mean that people in every case have a right not to be harmed but in no case a right to be helped. Therefore, only harmings may be forcibly averted, refusals of help must not be.

It is doubtful that the two distinctions coincide precisely. Many civil-law countries have "Good Samaritan" laws proscribing failure to render assistance. These statutes presuppose, as seems plausible, that there are cases in which morally mandatory assistance may be compelled through threat of criminal sanctions. Conversely, there may also be harmings that are so trivial that in principle they must not be forcibly averted.

Still, the two distinctions largely coincide, and some moral predicates fall entirely on one side or the other. Judicanda lacking in morally required or recommended generosity, civility, charity, humanity, decency, kindness, courtesy, mercy, or beneficence fail to help people without harming them. Such infractions may not be forcibly averted. Any injustice in a judicandum, however, does harm people and thus may in principle be forcibly averted.

The last two hypotheses differ in their informative value for the investigation of justice. The second hypothesis is illuminating: something important is learned about justice, if one finds that in principle every injustice may be forcibly averted. The third hypothesis would be similarly informative if it started out with a clear justice-independent notion of harm. The claim would then be that a judicandum is unjust only if it harms persons (in this independent sense). But the connection postulated by the third hypothesis can also be taken in the opposite way: Any unjust treatment is to count as a harming. So understood, this claim would be using a harm-independent notion of injustice to illuminate the notion of harm—rather than the other way around. Still, although the latter definitional sequencing seems more adequate, it is moderated by the need to preserve the core meaning of harm. Coming back to Kant's example, failure to give a criminal his deserved punishment cannot plausibly be presented as an instance of harming him. If one wants to call such an act of grace unjust (while maintaining the third hypothesis), then one must identify another harmed party—other criminals arbitrarily excluded from the amnesty, perhaps, or future victims of crimes that would not have occurred but for this failure to punish. Given this constraint imposed by the core meaning of harm, the third hypothesis is not then entirely uninformative.

FOURTH HYPOTHESIS: JUSTICE IS NOT PURELY RECIPIENT-ORIENTED. The last hypothesis is negative in character. It merely rejects a hypothesis that has come to dominate Anglophone academic discussions of justice. The refutation of this approach eliminates only one of many possibilities and may thus seem to make little progress. But because the rejected approach is so elegant and influential, it is interesting nonetheless to explore how it fails.

The approach to be criticized has developed out of utilitarianism, which holds that the moral assessment of actors, conduct, and social rules should be based solely on each such judicandum's relative impact on the world, and on human happiness in particular.

If such a view is accepted, the ordinary distinctions among moral predicates lose much of their significance. Whatever moral predicate may be used to criticize a rule, action, or person—cowardly, unjust, evil, indecent, and so on—the complaint always boils down to the judicandum's failure to be optimally happiness-promoting. Utilitarians have little use for the received panoply of moral predicates because it ultimately does not matter how a suboptimal judicandum is squandering potential human happiness.

Those who reduce the traditional multiplicity of moral defects in this way may simply want to do away with the surplus predicates. Alternatively, they may prefer to redeploy these predicates—using them not (as traditionally) for different defects of the same judicanda, but instead for defects in different kinds of judicanda. In this way, justice has come to name the specific moral virtue of social rules.

In his famous *A Theory of Justice* (1999 [1971]), John Rawls adopts this redeployment. His theory is focused exclusively on social institutions and, even more narrowly, on a society's basic institutional design. And, presenting justice as the first virtue of social institutions, he recognizes no further moral predicates by means of which such institutions might be subjected to potentially competing moral judgments.

Rawls further follows the utilitarians by adopting a broadly consequentialist mode of moral assessment, and one that focuses specifically on the well-being of individuals. Thus he shares with utilitarianism the purely recipient-oriented mode of moral assessment. The unique characteristic of this approach can be expressed as the assumption that the only information needed in the moral assessment of any judicandum is information about its relative impact on persons, that is, information about how persons fare with the judicandum as it is versus how they would fare if the judicandum were different.

Unlike the utilitarians, Rawls employs the purely recipient-oriented approach on two distinct levels and moreover rejects, on both levels, utilitarian conceptions of well-being as happiness or desire fulfillment. Central to his theory is a criterion ("the two principles") of justice that assesses alternative institutional designs on the basis of the distribution of social primary goods each would produce. The just design is the one that would produce the best feasible distribution among citizens of basic rights and liberties, income and wealth, powers and responsibilities of office, and other social bases of self-respect. This proposed criterion of justice is justified, again in a purely recipient-oriented manner, on the ground that its public adoption would lead to a better fulfillment of citizens' three higher-order interests than the public adoption of any alternative criterion of justice. The argument on this second level is presented in terms of a contractualist thought experiment ("original position") in which representatives of citizens, informed only that their clients have the three higher-order interests but given no further specific information about them, come to agree on a particular public criterion of justice.

Rawls's theory, with its particular assessment standards on the two levels, has received much lively critique. Under the *Equality of What?* label, academics debate whether human well-being should be conceived in terms of happiness, welfare, desire fulfillment, Rawlsian social primary goods, Dworkinian resources, or capabilities à la Sen or Nussbaum. And academics also debate how such well-being information about individuals should be aggregated: Should well-being simply be averaged or should special weight be given to equality, sufficiency, or the worse off? Yet, beneath all this disagreement, the purely recipient-oriented approach is largely taken for granted in the Anglophone countries.

The fourth hypothesis is that any such purely recipient-oriented conceptions of justice are untenable.

DEFENSE OF THE FOURTH HYPOTHESIS. When it comes to actors and their conduct, the fourth hypothesis is—at least outside academic philosophy—hardly controversial. Here any conception of the form "to be just is to promote a good distribution among one's recipients" is bound to be unacceptable because it disregards morally relevant information about how a judicandum has its effects. It is widely taken to make a great moral difference whether some conduct brings about the death of an innocent person or merely fails to prevent such a death. The world at large may be worse, perhaps even less just, if two good persons die prematurely than if a single less deserving person dies in their stead. And yet, it is not morally required, nor even permitted, to save the former by killing the latter.

An analogous point applies to social institutions. This is clearest when one reflects on the criminal law. Here Rawls's first priority rule (*A Theory of Justice* 1999 [1971], p. 266) holds that basic liberties may be restricted whenever such restriction, by making the remaining basic liberties more secure, is a gain on balance for the basic liberties of the representative citizen. This claim fits well with his purely recipient-oriented approach as enshrined in the contractualist thought experiment of the original position. But it is inconsistent with the considered judgments his theory was meant to accommodate and to unify. One of the examples Rawls gives (*A Theory of Justice* 1999 [1971], pp. 212–213) involves a strict-liability criminal statute that permits conviction without a showing of mens rea. Although it violates citizens' basic liberties, this law is nonetheless said to be permissible as "the lesser of two evils" if it is necessary to block even greater dangers to citizens' basic liberties (a danger of civil war, in his example).

Similar arguments could plausibly be made in regard to other aspects of the criminal law. Constraints on searches, seizures, and interrogations should be relaxed if this would entail, through more effective crime fighting, a net gain for citizen's basic liberties. Standards of evidence ("beyond a reasonable doubt") should be lowered, and less than unanimity be required for a jury conviction, if this would, through increased deterrence and through disabling more repeat offenders, produce a net gain for

citizen's basic liberties. And draconian punishments (e.g., execution) should be imposed for high-elasticity crimes (e.g., drunk driving) if citizens' overall risk of a premature violent death can thereby be reduced.

These theory-produced judgments collide with people's intuitive sense of justice, which is responsive not merely to the magnitudes, probabilities, and distribution of morally significant goods and ills, but also to the attitude social rules manifest and to the causal pathways on which they affect their recipients. Even if these recipients themselves have no reason to care whether burdens falling on them (such as hunger or risks of premature death) are imposed by social rules or merely not averted by them, this difference is nonetheless morally significant. The purely recipient-oriented approach cannot then fulfill Rawls's stated ambition of accommodating and unifying his compatriots' considered judgments.

The reason is that judgments of justice take account not only of the passive perspective of recipients, of citizens as governed by social rules, but also of the active perspective of authors, of citizens as co-responsible for these rules. For citizens as recipients, all threats to their basic liberties are indeed on a par—whether they arise from criminal or crime-fighting activities. But for citizens as co-legislators it makes a considerable difference whether people are roughed up by criminals or by police interrogators, are killed prematurely by a drunk driver or through execution for drunk driving. As co-legislators, people take greater moral responsibility for harms they mandate or authorize through their institutional order than for equal harms they could prevent through it. That some institutional arrangements would be better for citizens as recipients is not sufficient to show that they are morally permitted—let alone required. Purely recipient-oriented conceptions of justice are bound to fail because they systematically ignore the active perspective of those who bear responsibility for a particular judicandum.

Plausible justice assessments of social rules must be sensitive to both citizen perspectives. What matters in the moral assessment of social rules is not merely how these rules affect, but how they treat their recipients. The importance of this point extends well beyond the criminal law. A rule under which those suffering from some genetic defect are not entitled to life-saving treatment may be exactly as bad for them as a rule mandating execution for those with certain physical features is for the executed. Though comparable in their relative impact on recipients, these two rules are nonetheless worlds apart in how they treat such recipients.

Once such differences in attitude and causal pathways have been restored their proper moral significance, people may well find that justice is not the only moral virtue of social institutions. If a society's institutional order provides little funding for public health services, more citizens will avoidably die or suffer. And yet, such an order is surely less unjust than one under which equal suffering and deaths are explicitly imposed by the rules, perhaps on members of a certain ethnic or religious minority. In fact, in a poor society with other urgent needs, low funding for public health services may not be unjust at all. And even in an affluent society, such low funding may not be unjust, so long as society helps treat the medical conditions it causes (e.g., through pollution). Still, if the institutional order of an affluent society ignores the plight of citizens with congentital medical conditions, say, it may well be morally flawed in other ways: be ungenerous, mean, or inhumane.

This result supports the fourth hypothesis by showing it to be a step toward a conception of justice that covers all judicanda in a unified way: In regard to actors and their conduct people are already familiar with the possibility of actors being beyond reproach in terms of justice and yet morally flawed in other ways.

See also Affirmative Action; Civil Disobedience; Distant Peoples and Future Generations; Feminist Social and Political Philosophy; Just War Theory; Natural Law; Punishment; Racism; Rawls, John; Reflective Equilibrium; Rights; Social Contract; Terrorism; Utilitarianism.

Bibliography

Ackerman, Bruce. *Social Justice and the Liberal State*. New Haven, CT: Yale University Press, 1980.

Aristotle. *Nicomachean Ethics*. [c. 330 BCE]. Translated by Terence Irwin. Indianapolis: Hackett, 1985.

Barry, Brian. *Theories of Justice*. Berkeley: University of California Press, 1989.

Beitz, Charles. *Political Equality*. Princeton, NJ: Princeton University Press, 1989.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation* [1789], edited by J. H. Burns and H. L. A. Hart. New York: Methuen, 1982.

Berlin, Isaiah. *Four Essays on Liberty*. Oxford: Oxford University Press, 1969.

Cohen, G. A. *If You're an Egalitarian, How Come You Are So Rich?* Cambridge, MA: Harvard University Press, 2001.

Daniels, Norman. *Just Health Care*. Cambridge, U.K.: Cambridge University Press, 1985.

Dworkin, Ronald. *Sovereign Virtue*. Cambridge, MA: Harvard University Press, 2000.

Dworkin, Ronald. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press, 1978.
Elster, Jon. *Local Justice*. New York: Russell Sage Foundation, 1992.
Galston, William. *Justice and the Human Good*. Chicago: University of Chicago Press, 1980.
Gauthier, David. *Morals by Agreement*. Cambridge, U.K.: Cambridge University Press, 1987.
Hart, H. L. A. *The Concept of Law*. Oxford: Oxford University Press 1961.
Hayek, Friedrich A. *The Mirage of Social Justice*. Chicago: University of Chicago Press 1976.
Hegel, G. F. W. *Elements of the Philosophy of Right* [1821]. Translated by H. B. Nisbet. Cambridge, U.K.: Cambridge University Press, 1991.
Hobbes, Thomas. *Leviathan* [1651], edited by C. B. MacPherson. Harmondsworth, U.K.: Penguin, 1968.
Hume, David. *An Enquiry concerning the Principles of Morals* [1751], edited by Jerome B. Schneewind. Indianapolis: Hackett, 1983.
Hurley, Susan. *Justice, Luck and Knowledge*. Cambridge, MA: Harvard University Press, 2003.
Kant, Immanuel. *The Metaphysics of Morals* [1797]. Translated by Mary Gregor, cited by Academy pagination. Cambridge, U.K.: Cambridge University Press, 1996.
Locke, John. *Two Treatises of Government* [1689], edited by Peter Laslett. Cambridge, U.K.: Cambridge University Press, 1960.
Marx, Karl. *Selected Writings,* edited by David McLennan. Oxford: Oxford University Press, 2000.
Mill, John Stuart: *The Basic Writings of John Stuart Mill: On Liberty, The Subjection of Women and Utilitarianism* [1859, 1869, 1861], edited by Jerome B. Schneewind. New York: Random House, 2002.
Miller, David. *Principles of Social Justice*. Cambridge, MA: Harvard University Press, 1999.
Nagel, Thomas. *Partiality and Equality*. Oxford: Oxford University Press, 1991.
Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
Nussbaum, Martha. *Sex and Social Justice*. Oxford: Oxford University Press, 1999.
Okin, Susan. *Justice, Gender, and the Family*. New York: Basic Books, 1989.
O'Neill, Onora. *Toward Justice and Virtue*. Cambridge, U.K.: Cambridge University Press, 1996.
Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
Plato. *The Republic.* [ca. 380 BCE]. Translated by G. M. A. Grube. Indianapolis: Hackett, 1974.
Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.
Rawls, John. *A Theory of Justice* [1971]. Cambridge, MA: Harvard University Press, 1999.
Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.
Roemer, John. *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press, 1996.
Rousseau, Jean-Jacques. *The Social Contract* [1762]. Translated by M. W. Cranston. Harmondsworth, U.K.: Penguin, 1968.
Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1999.
Sen, Amartya. "Equality of What?" In *Choice, Welfare and Measurement*, edited by Amartya Sen. Cambridge, U.K.: Cambridge University Press, 1982.
Sen, Amartya. *Inequality Reexamined*. Cambridge, MA: Harvard University Press, 1992.
Sen, Amartya, and Bernard Williams, eds. *Utilitarianism and Beyond*. Cambridge, U.K.: Cambridge University Press, 1982.
Sidgwick, Henry. *The Methods of Ethics* [1874]. Indianapolis: Hackett, 1981.
Smith, Adam. *The Theory of Moral Sentiments* [1759], edited by D. D. Raphael and A. L. Macfie. Oxford: Oxford University Press, 1979.
Tuck, Richard. *Natural Rights Theories*. Cambridge, U.K.: Cambridge University Press, 1979.
van Parijs, Philippe. *Real Freedom for All: What (if Anything) Can Justify Capitalism?* Oxford: Oxford University Press 1995.
von Humboldt, Wilhelm. *The Limits of State Action* [1852]. Cambridge, U.K.: Cambridge University Press, 1969.
Walzer, Michael. *Just and Unjust Wars*. New York: Basic Books, 1977.
Walzer, Michael. *Spheres of Justice*. New York: Basic Books, 1984
Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.
Wittgenstein, Ludwig. *On Certainty,* edited by G. E. M. Anscombe and G. H. von Wright. Oxford: Blackwell 1969.

Thomas Pogge (2005)

JUSTIFICATION OF MORAL PRINCIPLES

See *Moral Principles: Their Justification*

JUST WAR THEORY

In traditional just war theory there are two basic elements: an account of just cause and an account of just means. Just cause is usually specified as follows:

(1) There must be substantial aggression.

(2) Nonbelligerent correctives must be either hopeless or too costly.

(3) Belligerent correctives must be neither hopeless nor too costly.

Needless to say, the notion of substantial aggression is a bit fuzzy, but it is generally understood to be the type of aggression that violates people's most fundamental rights. To suggest some specific examples of what is and is not substantial aggression, usually the taking of

hostages is regarded as substantial aggression while the nationalization of particular firms owned by foreigners is not so regarded. But even when substantial aggression occurs, frequently nonbelligerent correctives are neither hopeless nor too costly. And even when nonbelligerent correctives are either hopeless or too costly, in order for there to be a just cause, belligerent correctives must be neither hopeless nor too costly.

Traditional just war theory assumes, however, that there are just causes and goes on to specify just means as imposing two requirements:

(1) Harm to innocents should not be directly intended as an end or a means.

(2) The harm resulting from the belligerent means should not be disproportionate to the particular defensive objective to be attained.

While the just means conditions apply to each defensive action, the just cause conditions must be met by the conflict as a whole.

It is important to note that these requirements of just cause and just means are not necessarily about war at all. Essentially, they constitute a theory of just defense that can apply to war but can also apply to a wide range of defensive actions short of war.

THE INTENDED/FORESEEN DISTINCTION

Just war theory presupposes that we can, in practice, distinguish between what is foreseen and what is intended, and some have challenged whether this can be done. So first one needs to address this challenge.

The practical test that is frequently appealed to in order to distinguish between foreseen and intended elements of an action is the Counterfactual Test, according to which two questions are relevant:

(1) Would you have performed the action if only the good consequences would have resulted and not the evil consequences?

(2) Would you have performed the action if only the evil consequences resulted and not the good consequences?

If an agent answers "yes" to the first question and "no" to the second, some would conclude that (1) the action is an intended means to the good consequences, (2) the good consequences are an intended end, and (3) the evil consequences are merely foreseen.

But how well does this Counterfactual Test work? Douglas P. Lackey argues that the test gives the wrong result in any case where the "act that produces an evil effect produces a larger good effect" (1987, p. 260). He cites the bombing of Hiroshima, Japan, as an example. That bombing is generally thought to have had two effects: the killing of Japanese civilians and the shortening of World War II. Now suppose we were to asked:

(1) Would Harry S. Truman have dropped the bomb if only the shortening of the war would have resulted but not the killing of the Japanese civilians?

(2) Would Truman have dropped the bomb if only the Japanese civilians would have been killed and the war not shortened?

And suppose that the answer to the first question is that Truman would have dropped the bomb if only the shortening of the war would have resulted but not the killing of Japanese civilians, and that the answer to the second question is that Truman would not have dropped the bomb if only the Japanese civilians would have been killed and the war not shortened. Lackey concludes from this that the killing of civilians at Hiroshima, self-evidently a means for shortening the war, is by the Counterfactual Test classified not as a means but as a mere foreseen consequence. On these grounds Lackey rejects the Counterfactual Test as an effective device for distinguishing between the foreseen and the intended consequences of an action.

Unfortunately, this is to reject the Counterfactual Test only because one expects too much from it. It is to expect the test to determine all the following:

(1) Whether the action is an intended means to the good consequences.

(2) Whether the good consequences are an intended end of the action.

(3) Whether the evil consequences are simply foreseen consequences.

In fact, this test is capable of meeting only the first two of these expectations. And the test clearly succeeds in doing this for Lackey's own example, where the test shows the bombing of Hiroshima to be an intended means to shortening the war, and shortening the war an intended consequence of the action.

To determine whether the evil consequences are simply foreseen, however, an additional test is needed, which can be called the Nonexplanation Test. According to this test the relevant question is: Does the bringing about of

the evil consequences help explain why the agent undertook the action as a means to the good consequences? If the answer is "no," that is, if the bringing about of the evil consequences does not help explain why the agent undertook the action as a means to the good consequences, the evil consequences are merely foreseen. But if the answer is "yes," the evil consequences are an intended means to the good consequences.

Of course, there is no guaranteed procedure for arriving at an answer to the Nonexplanation Test. Nevertheless, when we are in doubt concerning whether the evil consequences of an act are simply foreseen, seeking an answer to the Nonexplanation Test will tend to be the best way of reasonably resolving that doubt. For example, when applied to Lackey's example, the Nonexplanation Test comes up with a "yes," since the evil consequences in this example do help explain why the bombing was undertaken to shorten the war. For, according to the usual account, Truman ordered the bombing to bring about the civilian deaths, which by their impact on Japanese morale were expected to shorten the war. So, by the Nonexplanation Test, the civilian deaths were an intended means to the good consequences of shortening the war.

Just war theory has been challenged in various ways. Three of the most important are a conventionalist challenge to just means, a collectivist challenge to just means, and a feminist objection to just cause and just means.

A CONVENTIONALIST CHALLENGE TO JUST MEANS

The criteria of just means have been incorporated to some degree into the military codes of different nations and adopted as international law. George Mavrodes (1984) contends that the criteria of just means ought to be met simply because they have been incorporated into military codes or adopted as international law. Mavrodes arrives at this conclusion largely because he finds the standard attempts to specify the convention-independent basis for condition (2) of just means to be so totally unsuccessful. All such attempts, Mavrodes claims, are based on an identification of innocents with noncombatants. But by any plausible standard of guilt and innocence with moral content, Mavrodes contends, noncombatants can be guilty and combatants innocent. For example, noncombatants who are doing everything in their power to support an unjust war financially would be morally guilty, and combatants who were forced into military service and intended never to fire their weapons at anyone would be morally innocent. Consequently, the guilt-innocence distinction will not support the combatant-noncombatant distinction.

Hoping to support the combatant-noncombatant distinction, Mavrodes suggests that the distinction might be grounded on a convention to observe it. This would mean that our moral obligation to abide by condition (2) of just means would be a convention-dependent obligation. Nevertheless, Mavrodes does not deny that we have some convention-independent obligations. Our obligation to refrain from wantonly murdering our neighbors is given as an example of a convention-independent obligation, as is our obligation to reduce the pain and death involved in combat. But to refrain from harming noncombatants when harming them would be the most effective way of pursuing a just cause is not included among our convention-independent obligations.

Still, Mavrodes does not claim that our obligation to refrain from harming noncombatants is purely convention-dependent. He allows that, in circumstances in which the convention of refraining from harming noncombatants does not exist, we might still have an obligation to unilaterally refrain from harming noncombatants, provided that our action will help give rise to a convention prohibiting such harm with its associated good consequences. According to Mavrodes, our primary obligation is to maximize good consequences, and this obligation requires that we refrain from harming noncombatants when that will help bring about a convention prohibiting such harm. By contrast, someone who held that our obligation to refrain from harming noncombatants was purely convention-dependent would never recognize an obligation to unilaterally refrain from harming noncombatants. On a purely convention-dependent account, obligations can only be derived from existing conventions; the expected consequences from establishing a particular convention could never ground a purely convention-dependent obligation. But while Mavrodes does not claim that our obligation to refrain from harming noncombatants is purely convention-dependent, he does claim that this obligation generally arises only when there exists a convention prohibiting such harm. According to Mavrodes, the reason for this is that, generally, only when there exists a convention prohibiting harm to noncombatants will our refraining from harming them, while pursuing a just cause, actually maximize good consequences.

But is there no other way to support our obligation to refrain from harming noncombatants? Mavrodes would deny that there is. Consider, however, Mavrodes's own example of the convention-independent obligation

not to wantonly kill our neighbors. There are at least two ways to understand how this obligation is supported. Some would claim that we ought not to wantonly kill our neighbors because this would not maximize good consequences. This appears to be Mavrodes's view. Others would claim that we ought not to wantonly kill our neighbors, even if doing so would maximize good consequences, simply because it is not reasonable to believe that our neighbors are engaged in an attempt on our lives. Both of these ways of understanding how the obligation is supported account for the convention-independent character of the obligation, but the second approach can also be used to show how our obligation to refrain from harming noncombatants is convention-independent. According to this approach, since it is not reasonable to believe that noncombatants are engaged in an attempt on our lives, we have an obligation to refrain from harming them. So interpreted, our obligation to refrain from harming noncombatants is itself convention-independent, although it will give rise to conventions.

Of course, some may argue that whenever it is not reasonable to believe that persons are engaged in an attempt on our lives, an obligation to refrain from harming such persons will also be supported by the maximization of good consequences. Still, even if this were true, which seems doubtful, all it would show is that there exists a utilitarian or forward-looking justification for a convention-independent obligation to refrain from harming noncombatants; it would not show that such an obligation is a convention-dependent obligation, as Mavrodes claims.

A COLLECTIVIST CHALLENGE TO JUST MEANS

According to a collectivist challenge to just means, more people should be included under the category of combatants than the standard interpretation of just means allows. Just means, as noted earlier, imposes two requirements:

> (1) Harm to innocents should not be directly intended as an end or a means.
>
> (2) The harm resulting from the belligerent means should not be disproportionate to the particular defensive objective to be attained.

According to advocates of this challenge to just means, the problem is that the standard interpretation of (1) does not assume that the members of a society are collectively responsible for the actions of their leaders unless they have taken radical steps to oppose or disassociate themselves from those actions, for example, by engaging in civil disobedience or by emigrating. Of course, those who are unable to take such steps, particularly children, would not be responsible in any case; but for the rest, advocates of this collectivist challenge contend that failure to take the necessary radical steps, when one's leaders are acting aggressively, has the consequence that one is no longer entitled to full protection as a noncombatant. Some of those who press this objection against the just means criteria of just war theory, like Gregory Kavka (1985), contend that the members of a society can be directly threatened with nuclear attack to secure deterrence, and so reject noncombatant immunity, but then deny that carrying out such an attack could ever be morally justified. Others, like James W. Child (1986), reject both noncombatant immunity and proportionality by contending that the members of a society who fail to take the necessary radical steps can be both indirectly threatened and indirectly attacked with what would otherwise be a disproportionate attack.

In response to this collectivist challenge the first thing to note is that people are more responsible for disassociating themselves from the unjust acts of their leaders than they are for opposing those same acts. For there is no general obligation to oppose all unjust acts, even all unjust acts of one's leaders, because this would impose an unreasonable demand on individuals, and we are not morally required to be saints. Nevertheless, there is a general obligation to disassociate oneself from unjust acts and to minimize one's contribution to them, because this is not an unreasonable demand to impose on each of us. Of course, how far one is required to disassociate oneself from the unjust acts of one's leaders depends on how much one is contributing to those actions. If one's contribution is insignificant, as presumably a farmer's or a teacher's would be, only a minimal effort to disassociate oneself would be required, unless one's action could somehow be reasonably expected, in cooperation with the actions of others, to put a stop to the unjust actions of one's leaders. However, if one's contribution is significant, as presumably a soldier's or a munitions worker's would be, a maximal effort at disassociating oneself would be required immediately, unless by delaying one could reasonably expect to put a stop to the unjust actions of one's leaders.

In support of this collectivist challenge to just war theory Child (1986) offers the example of a member of a board of directors of a company that is engaging in the immoral and illegal activity of pouring large quantities of arsenic into the public water supply as a matter of ongo-

ing operations. When the policy is before the board, she votes no but does nothing else. Later, when sued in tort (or charged in crime) with these transgressions of duty, she pleads that she voted no. Child argues that mere formal dissent in this case does almost nothing to relieve her liability, legal or moral.

But while one might agree with Child that in this case the member of the board of directors had at least the responsibility to disassociate herself from the actions of the board by resigning, this does not show that farmers and teachers are similarly responsible for disassociating themselves from the unjust action of their leaders either by engaging in civil disobedience or by emigrating. This is because neither their contributions to the unjust actions of their leaders nor the effect of their disassociation on those unjust actions would typically be significant enough to require such a response.

This is not to deny that some other response (e.g., political protest or remunerations at the end of the war) would not be morally required. However, to meet this collectivist challenge, it suffices to show that not just any contribution to the unjust actions of one's leaders renders the contributor subject to attack or threat of attack; one's contribution must be significant enough to morally justify such a response.

A FEMINIST CHALLENGE TO JUST CAUSE AND JUST MEANS

A formidable challenge to both the just cause and just means criteria of just war theory comes from feminism. According to the feminist challenge to just war theory, sexism and militarism are inextricably linked in society. They are linked, according to Betty Reardon (1985), because sexism is essentially a prejudice against all manifestations of the feminine, and militarism is a policy of excessive military preparedness and eagerness to go to war that is rooted in a view of human nature as limited to masculine characteristics. Seen from a militarist perspective, other nations are competitive, aggressive, and averse to cooperation, the same traits that tend to be fostered primarily in men in a sexist society. By contrast, the traits of openness, cooperativeness, and nurturance that promote peaceful solutions to conflicts tend to be fostered primarily in women, who are then effectively excluded from positions of power and decision making in a sexist society. Consequently, if people are to rid society of militarism, Reardon argues, they need to rid society of sexism as well.

But even granting that sexism and militarism are inextricably linked in society in just the way Reardon maintains, how does this effect the validity of just war theory? As just war theory expresses the values of proportionality and respect for the rights of innocents, how can it be linked to militarism and sexism? The answer is that the linkage is practical rather than theoretical. If the leaders in militarist-sexist society have been socialized to be competitive, aggressive, and averse to cooperation, then they will tend to misapply just war theory when making military decisions. This represents an important practical challenge to just war theory. And the only way of meeting this challenge seems to be to rid society of its sexist and militarist attitudes and practices so as to increase the chances that just war theory will be correctly applied in the future.

Of course, still other challenges could be raised to just war theory but, in large measure, just war theory has stood the test of time. Moreover, if the theory can be reconciled with the most morally defensible form of pacifism, such that the only wars and large-scale conflicts that definitely satisfy the requirements of just war theory are the only wars and large-scale conflicts to which antiwar pacifists cannot reasonably object, then it is really hard to see how the theory could be displaced.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism.

Bibliography

Cady, Duane L. *From Warism to Pacifism: A Moral Continuum*. Philadelphia: Temple University Press, 1989.

Child, James W. *Nuclear War: The Moral Dimension*. New Brunswick, NJ: Transaction, 1986.

Holmes, Robert L. *On War and Morality*. Princeton, NJ: Princeton University Press, 1989.

Kavka, Gregory. "Nuclear Deterrence: Some Moral Perplexities." In *The Ethics of War and Nuclear Deterrence*, edited by James P. Sterba, 127–138. Belmont, CA: Wadsworth, 1985 .

Lackey, Douglas P. "The Moral Irrelevance of the Counterforce/Countervalue Distinction." *The Monist* 70 (3) (July 1987): 255–276.

Mavrodes, George. "Conventions and the Morality of War." In *Morality in Practice*, edited by James P. Sterba, 302–310. Belmont, CA: Wadsworth, 1984.

Reardon, Betty. *Sexism and the War System*. New York: Teachers College Press, 1985.

Sterba, James P. "Reconciling Pacifists and Just War Theorists." *Social Theory and Practice* 18 (1) (Spring 1992): 21–38.

Walzer, Michael. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. 2nd ed. New York: Basic, 1992.

James Sterba (2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

KABBALAH – MARXIST PHILOSOPHY

5
volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON
GALE™

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON
GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

K

KABBALAH

Kabbalah (literally "tradition") is used both as a general name for Jewish mysticism and as the specific designation for its major medieval variety. Mystical awareness is to be found in the biblical and rabbinic tradition and had literary expression in some of the prophetic writings, psalms, and apocalypses. More characteristically, however, what is referred to as Kabbalah is a type of occult theosophical formulation of the doctrines of the Jewish religion, particularly those concerned with creation, revelation, and redemption. This occult system structures and, in part, fossilizes individual intuitions of divine reality in terms of the culture in which it arose. Typically, the purpose of the complicated structuring of these formulated intuitions is to supply a focus in contemplation by which the Kabbalist can recover the untarnished brightness of direct mystical awareness.

Besides the sources of Kabbalah in the doctrines and literature of the Jewish tradition, a wide variety of other sources has been noted, which have introduced elements from the various cultures with which the Jewish people have come in contact in their dispersion. Among these influences should be included some Persian elements, both Parsi and Zoroastrian, and Neo-Platonic and Neo-Pythagorean elements which entered during the Hellenistic period; Christian influences and Gnostic themes added at a somewhat later time; and borrowings from Muslim sectarianism after the emergence of Islam. This mixture of elements explains the difficulty that scholars have found in disentangling the sources of Kabbalah. It should be said, however, that the pursuit of sources has less relevance here than it may have for other subjects, because what is essential is not the materials out of which the Kabbalistic theosophical system was created, but rather the use that was made of the materials.

MAJOR DOCTRINES

CREATION. All Jewish mysticism has seen the need for reinterpretation of the literal account of creation given in the book of Genesis. As it stands, the account does not sufficiently emphasize the transcendence of God. God is too close to humankind and the world to be the Supreme Mystery that the mystical temper insists He must be. The reinterpretation has generally taken form as a demiurgic theory. In such a theory, God Himself, the Boundless, the Infinite, the Transcendent, did not perform the material act of creating the world. This was the work of a lesser spirit, or demiurge, who was brought into existence by God in order to do this specific job. As the conception of

God's transcendence developed, one demiurge seemed insufficient to express the sense of awesome distance between divinity and the material world. The remoteness of God from the world was heightened, therefore, by adding other intermediaries and thus forming a chain from God to matter whose links were of increasing materiality.

A second problem in the biblical account of creation concerns matter. If we accept God as infinite, all must be contained in Him. Where, then, is there a place for matter outside of God? This issue was finally resolved by a theory that combined the idea of God's voluntary self-contraction with the concept of emanation. In this account, God, prior to creation, was actually infinite. To make room for creation, however, He voluntarily contracted or limited Himself. Some excess of spiritual substance overflowed into the space from which God had removed Himself, and this excess, or emanation, provided both the demiurgic intermediaries described above and the matter out of which the world was created. Because all substance is thus ultimately an overflowing of God's substance, Kabbalah is a pantheistic doctrine. The completed series of emanations served the additional purpose of providing the road by which humanity's aspiring spirit might reach the heights of divinity; thus, it served both as the mechanism of creation and as the "itinerary of the mind to God" (to borrow an expression from St. Bonaventure).

REVELATION. After the first destruction of the Temple at Jerusalem, and even more after its second destruction, the Scriptures served as a focus for the religious devotion of the Jews. Their state was no more; their cultus was no more; all that was left to them was their belief in God and His Word. For the continuance of the Jewish religion, it came to seem necessary that not only the content of revelation, but even its physical form, should be considered sacrosanct and unchangeable. In all types of Judaism this regard for the letter of Scripture made necessary the development of exegetic techniques for raising the level of significance of much that is trivial in the Scriptures. For the mystics the problem was particularly difficult, because the level on which they had to interpret revelation to make it serve their purpose was highly symbolical. To make this reinterpretation possible, the Kabbalists developed letter and number symbolisms of great variety and complexity.

REDEMPTION. The Kabbalists maintained and even intensified the traditional Jewish view of redemption. In the Kabbalistic view salvation of the individual was little considered; it entered only as a means to the greater end of the salvation of humankind. This would come about through the agency of a Messiah of the Davidic line, who would lead the Jews in triumph to the Holy Land and inaugurate a reign of truth, justice, and mercy. The ideal of salvation is thus the establishment of an earthly paradise of human life, raised to its highest humanity. Other elements clouded this doctrine at various times in the history of mystical Messianism. For example, in the sixteenth century Isaac Luria introduced the idea that this regeneration could not take place until all preexisting souls had satisfactorily completed their earthly existence and that, since some souls were too weak to go unaided through life to perfection, other superior souls might coexist with them in one body to ensure their success. Although Luria's doctrine of transmigration found followers, it was exceptional rather than typical; in general, the Kabbalistic view of redemption was an extreme form of traditional Messianism. Attempts to calculate the exact date of the coming of the Messiah were widespread; the coincidence of various calculations in fixing on dates close to each other was sufficient to start a wave of Messianic movements and even to touch off a major explosion like the widespread impassioned support of Sabbatai Zevi, the so-called Messiah of Ismir (1626–1676).

HISTORICAL EXPRESSIONS

While a number of smaller groups, such as the Essenes of Palestine, the Therapeutae of whom Philo wrote, and the eighth-century Persian "Men of the Caves" whom the tenth-century Karaite historian Joseph ben Jacob al-Kirkisani described, maintained views similar in part to those that have been presented, these groups do not lie in the mainstream of Jewish mysticism. The main development is rather to be traced from the Jewish Gnosticism of the first millennium of the common era, with its concentration on the glory of God as manifested in His throne, supposedly located in the innermost of seven heavenly mansions, into the parallel forms of the medieval European developments of the Kabbalah—the practical, ethical, and sometimes magical mysticism of the German Jews and the speculative mysticism of the French and Spanish Jews. Thence the movement became enmeshed in the morbidity of seventeenth-century Messianism, before the two strains of mystical speculation and socioethical piety were reunited, in eastern Europe, in the still-flourishing movement of Hasidism.

The German pietist movement developed during the century between 1150 and 1250. Its chief formulators were Samuel the Hasid (fl. 1150), his son Judah the Hasid

(d. 1217), and a relative, Eleazar of Worms (fl. 1220). The chief literary expression of the movement is the *Book of the Pious* (Hebrew, *Sefer Hasidim*), a collection of the literary remains of the three founders, with special emphasis on Judah the Hasid, whose character and influence recall those of his Christian contemporary, St. Francis of Assisi, and, perhaps, remind one also of Paracelsus, who lived in the sixteenth century and who also combined genuine piety with magic. In addition to its concern with the doctrinal elements that have already been discussed as characteristic of all forms of Jewish mysticism, German Hasidism defined an ideal human type and a way of life-devoutness, rather than learning or traditionalism. The three chief elements in this devoutness were mental serenity, ascetic renunciation, and extreme altruism, leading to heights of devotion in which true fear of God and love of God became one. At these heights, the Hasid was thought to achieve a creative power of a magical nature.

In southern France, at the beginning of the thirteenth century, a more speculative Kabbalistic development began, under the sponsorship of Isaac the Blind (fl. 1200) and his disciples Ezra and Azariel. Their chief concern was the elaboration of emanation theory; they also suggested a doctrine of metempsychosis, although they did not develop it fully. In Spain, Abraham ben Samuel Abulafia (1240–c. 1292) combined this speculation with the development of number and letter symbolism and thus became one of the central figures in the development of Kabbalah. His disciple, Joseph ben Abraham Gikatilia (c. 1247–1305), presented both the techniques for symbolic interpretation and the doctrine of the ten emanations (Hebrew, *sephiroth*) in systematically interrelated form. About 1290 the Spanish Kabbalist Moses ben Shemtob de Leon (d. 1305) produced the work that, for many, represents the Kabbalah in its entirety: the lush compendium of esoteric doctrines in the form of a commentary on the Pentateuch known as *The Book of Splendor* (Hebrew, *Sefer Ha-Zohar*). From the time of its composition, this work has been the chief source of inspiration for later Jewish mystics and for Jewish mysticism. Of later Kabbalistic leaders, two in particular should be mentioned: Moses ben Jacob Cordovero (1522–1570), whose book, *A Garden of Pomegranates* (Hebrew, *Pardes Rimmonim*), is the most systematic and philosophical exposition of the doctrines of the Kabbalah up to his time; and his pupil, Isaac Luria (1534–1572), who left no written legacy, but whose disciples have made it clear that he developed the theosophic doctrines of creation and redemption far beyond his predecessors.

There are still Kabbalistic groups in existence, chiefly in Israel, but they are for the most part outgrowths of eighteenth-century Polish Hasidism, a movement akin to, though by no means identical with, earlier German pietism. Among major Jewish thinkers of the twentieth century, the chief rabbi of Jerusalem, Abraham Isaac Kook (1865–1935), approached most closely the spirit of the Kabbalah in his mystical awareness of the Messianic role of the Jewish people and in his Lurianic and Hasidic stress on the spark of holiness that is veiled by the material shell of things perceived by the senses. Martin Buber, whose reinterpretations of the Hasidic view of life are profound and suggestive, may also be named here and, among younger thinkers, Abraham Joshua Heschel, whose thought has clear kinship with Hasidic social ethics.

See also Bonaventure, St.; Buber, Martin; Cordovero, Moses ben Jacob; Creation and Conservation, Religious Doctrine of; Gnosticism; Jewish Philosophy; Mysticism, History of; Mysticism, Nature and Assessment of; Paracelsus; Philo Judaeus; Revelation.

Bibliography

TEXTS

Of the primary Kabbalistic literature, only the chief sections of the Zohar are available, in an English translation by Harry Sperling, Maurice Simon, and Paul P. Levertoff. 5 vols. (London: Soncino Press, 1931–1934). Other segments of the Zohar in inferior translations are included in S. L. MacGregor Mathers, *The Kabbalah Unveiled* (London, 1887). A theosophized version of *Sefer Yetzirah* is Knut Stenring, *The Book of Formation* (London, 1923). The Hebrew texts have not been critically edited.

Among recent writers of a mystical bent, the works of Martin Buber are readily available in English translations. None of Abraham Isaac Kook's works have been translated; however, there are good discussions of his life and thought in Jacob Agus, *Banner of Jerusalem* (New York: Bloch, 1946) and Isidore Epstein, *Abraham Kook, His Life and Works* (London, 1951). A. J. Heschel is best represented by *God in Search of Man: A Philosophy of Judaism.* (New York: Farrar Straus, 1955).

HISTORY OF KABBALAH

See Joshua Abelson, *Jewish Mysticism* (London, 1913); Christian D. Ginsburg, *The Kabbalah: Its Doctrine, Development, and Literature* (London, 1920); Adolph Franck, *The Kabbalah: or, the Religious Philosophy of the Hebrews* (New York: Kabbalah, 1926); Abba Hillel Silver, *A History of Messianic Speculation in Israel from the First through the Seventeenth Centuries* (Boston: Beacon Press, 1959); Gershom G. Scholem, *Major Trends in Jewish Mysticism* (New York: Schocken, 1946; and Joseph L. Blau,

The Story of Jewish Philosophy (New York: Random House, 1962).

J. L. Blau (1967)

KABBALAH [ADDENDUM]

Medieval Jewish philosophy contributed considerably to the mystical branch of Judaism known as Kabbalah. This movement is generally regarded as having its origins in twelfth and thirteenth-century Provence in the midrashically styled *Bahir* (*Book of Enlightenment*). Some, however, consider the much earlier *Sefer Yetsirah* (*Book of Formation*)—from the third through the seventh centuries—to be the earliest work of Kabbalah.

Chief among the philosophers who influenced concepts within Kabbalah were those who thrived in the Muslim cultures of Babylon (Iraq) and Spain. An example is Saadya Gaon (882—942), head of the Babylonian Yeshivah (religious academy) of Pumbedita. Although Saadya was a rationalist philosopher, he nevertheless published a detailed commentary on *Sefer Yetsirah*. In addition, he posited an intermediary between God and creation, known as the *kavod* or "glory." It is possible that this concept was influenced by the Karaite thinker, Benjamin al-Nahawandi (830–860), and that both were influenced by the Muslim *kalamic* (theological) view of the "creative word" of God. Contextually, the idea of the *kavod* is less likely to have been influenced by Christian ideas of the *logos*. The concept of an intermediary between God and creation influenced the seminal idea of the *sefirot* (emanations from the Divine), as developed in all major kabbalistic works.

Abraham ibn Ezra (1089–1164) was born in Muslim Tudela, northern Spain, but lived to see both his own birthplace and other major Spanish cities taken by Christian forces before he was thirty. At fifty he left Spain and traveled through northern Christian Europe, dying in a pogrom in London in 1164. Through his travels, he influenced kabbalistic thought in Ashkenazi and Christian domiciles at both a theoretical and practical level. For example, Ibn Ezra's complex attitude to the preexistence of "matter" impacted on circles in Provence, out of which the foundations for the *Bahir* emerged. The problem of "matter," which had not been widely discussed in works of popular Jewish biblical exegesis before Ibn Ezra, played a seminal part in kabbalistic thinking, both in relation to the *sefirot* and also in discussions about the origins and role of evil in the universe. This is particularly true of the sixteenth-century Lurianic Kabbalah of Sfat, northern Israel.

In some ways an even bolder innovation on Ibn Ezra's part was his emphasis on the importance of the *mitzvot* (religious commandments) that, when practiced correctly, could affect the deity. This theory influenced theurgical Kabbalah. It was instrumental in lending a psychological dimension to the practice of Kabbalah, in which human beings could be regarded as influencing the deity by means of the *sefirot*.

It is therefore not completely accurate to view Kabbalah solely as a movement (or series of movements) that emerges during certain tragic times of Jewish history. It is more accurate to see it as being embedded at the heart of the Jewish religion, with biblical and rabbinic antecedents. Kabbalah has also been compared to mystical traditions in other religions, notably Sufism, in which emphasis is placed on experience of the Divine. This approach has paralleled neuroscientific interest in the field of consciousness studies. Lastly, developments in the study of language and linguistics have led to emphasis on the importance of the "text" and letter mysticism in Kabbalah. Interest in Kabbalah may thus be summarized as historical, philosophical, psychological, linguistic, and experiential, but as being grounded in the same intellectual milieu as more conventional Jewish genres.

See also Chinese Philosophy: Buddhism; Consciousness; Experimentation and Instrumentation; Islamic Philosophy; Jewish Philosophy; Mysticism, History of; Philosophy of Language; Postmodernism; Sufism.

Bibliography

TRANSLATIONS INTO ENGLISH OF MAJOR KABBALISTIC WORKS

Kaplan, Aryeh. *The Bahir*. New York: Samuel Weiser, 1979.

Kaplan, Aryeh. *Sefer Yetzirah*. Rev. ed.York Beach, ME: Samuel Weiser, 1997.

Matt, Daniel. *The Zohar*. Vols. 1–2. Stanford, CA: Stanford University Press, 2004.

Tishby, Isaiah. *The Wisdom of the Zohar*. 3 vols. Translated by David Goldstein. Oxford: Oxford University Press, 1991 (1949).

SECONDARY SOURCES

Idel, Moshe. *Absorbing Perfections*. New Haven, CT: Yale University Press, 2002.

Idel, Moshe. *Kabbalah: New Perspectives*. New Haven, CT: Yale University Press, 1988.

Lancaster, Brian, L. *Approaches to Consciousness: The Marriage of Science and Mysticism*. Houndsmill, U.K.: Palgrave Macmillan, 2004.

Lancaster, Brian, L. *The Essence of Kabbalah*. London: Arcturus, 2005.

Lancaster, Irene. *Deconstructing the Bible*. London: Routledge Curzon, 2003.

Scholem, Gershom. *Major Trends in Jewish Mysticism*. New York: Schocken, 1961.

Scholem, Gershom. *On the Kabbalah and Its Symbolism*. New York: Schocken, 1965.

Wolfson, Elliot R. *Through a Speculum That Shines*. Princeton NJ: Princeton University Press, 1994.

Irene Lancaster (2005)

KAFKA, FRANZ

(1883–1924)

Franz Kafka, the German author, was the son of a Jewish businessman who had been a peddler in southern Bohemia. The family was German-speaking. Kafka studied law at the German University of Prague and at Munich and became an official of a workers' accident insurance company. He began writing in 1907 but by his own choice published little. About that time he contracted tuberculosis and for some years lived in various sanatoriums. His two engagements ended unhappily. In 1923 he moved to Berlin, where, living with a girl who was in charge of a Jewish orphanage, he achieved what happiness he was to know. He died of a tubercular infection of the larynx in a nursing home at Kierling, near Vienna.

The central experience of Kafka's life, it seems, was a manifold alienation—as a speaker of German in a Czech city, as a Jew among German and Czech Gentiles in a period of ardent nationalism, as a man full of doubts and an unquenched thirst for faith among conventional "liberal" Jews, as a born writer among people with business interests, as a sick man among the healthy, and as a timid and neurasthenic lover in exacting erotic relationships.

Kafka's narrative art is at once immensely original, prophetic, and fragmentary—hence the large number of mutually exclusive interpretations it has received. Several elements of his prose were the stock in trade of the minor literature of his day. His language is unemphatic and prosy and occasionally contains Prague-German provincialisms; some of the subjects of his stories belong to the horror literature of the turn of the twentieth century; he shared the modern interest in psychological motivation; and he often used the smaller prose genres cultivated by his contemporaries in Prague and Vienna. But the use Kafka made of these elements is startlingly original, and the compelling gnostic vision of the world that is fashioned from them has become one of the major literary and intellectual influences of our age. In Kafka's work the existentialists' conceptions of absurdity and dread are fully explored. Unlike the later existentialists, he did not derive a positive value from these modes of experience; the value of his writings lies in the intense lucidity of the exploration.

It is obvious from the very titles of many of Kafka's stories—*The Trial,* "The Judgment," "Before the Law," "The Penal Settlement"—that his work is informed by a strong legalistic strain, possibly derived from his Jewish heritage but then secularized. In the famous "Letter to His Father" (1919) he recounted a certain childhood episode that violated his sense of justice. Characteristically, its terror for him lay in his inability to connect the trivial "crime" with the monstrous punishment he received.

The novel *The Trial,* begun in 1914 and published by Kafka's friend Max Brod in 1925, at once challenges and refines our conventional ways of connecting causes and effects through the story of a young man, Josef K, who one day wakes up in his lodgings to find himself arrested without knowing what wrong he has done. He makes various attempts to justify himself against the enigmatic accusation and to influence a number of people who he believes may effect his acquittal. Although offered a chance of repudiating the jurisdiction of the court that is concerned with his case, he ends up by being marched off to his execution, to die "like a dog."

The question What has Josef K done? receives a number of detailed answers, the total effect of which is to undermine the reader's notion of guilt. Josef K has lived the unremarkable life of an average young man, a bank clerk. Since in his "ordinary" life he always based his relations with other people on asserting what he believed were his "rights" in this or that situation, it is consistent with his character that he should seek to justify himself before the Law. The only thing he knows about that Law (and the all but unattainable authority behind it) is that it is powerful, whereas he is weak. According to the "inescapable logic" of the world, he must therefore be outside the Law and thus, in some sense, guilty. With his every move the not wholly irrational sense of guilt drags more violently at his soul. At first, this sense is no more than an uneasy "They are sure to have something on me," but gradually it is magnified by all the actions, in themselves trivial, which constitute "normal" behavior in our world, coupled with Josef K's inability to live "outside the Law," which for Kafka amounted to consciousness itself. Simplifying the subtly involuted and complex texture of the novel, we may conclude that "minor guilt + situation

of weakness + self-justification = major sense of guilt," which is tantamount to saying that Kafka's dialectical ingenuity is expended on making convincing the equation "[subjective] sense of guilt = [objective] guilt."

Similar dialectical devices are used in the second major work, the unfinished novel *The Castle* (1921–1922, published 1926). K, a land surveyor, has been called to a village that is governed by an authority that resides in a nearby castle. The village and its inhabitants are described only as they are related to K and to his attempts to justify his presence there. His commission, the authority on whose behalf he is to perform it, its relation to himself and to the villagers, the extent of its power, and the morality of its commands—all these are not so much vague as complexly contradictory. (Kafka was prophetically describing the anonymous, muffled workings of a totalitarian ministry as they affect the helpless victim, but since his style is that of an "objective" report, he allowed himself no expressions of pity.) Every assurance that K receives is thrown into doubt either by an oblique contradiction or by K's own unnerved (and, to the reader, unnerving) insistence on exploring its possible ambiguities.

Again, the novel elaborates a vicious circle. K uses the people he meets in order to wrest from them hints or indications about his task and status but because he lacks the assurance of a clearly defined status and task, he is an outsider and thus in a position of weakness. He is therefore bound to construe all these hints as hostile and thus distrust them. K does not have enough strength to break the spell that the Castle (like the court in *The Trial*) seems to be casting over him, for he looks to it as the place that, in justifying him, will give him strength. And, to keep alive K's torments of uncertainty, the Castle need do little more than send an occasional hint of a possible way of deliverance.

Leaving aside the various Freudian, Marxist, and Christian interpretations that Kafka's work has received, its fragmentary nature points to a fundamental hiatus. His heroes' desolate quests for justice, recognition, and acceptance by the world are meaningful to us because they invoke our sense of pity and justice, whereas the matter-of-fact ways in which these quests are presented invite us to accept cruelty and injustice as though they were necessary and self-evident modes of life. Thus, the meaningfulness of the quests is impaired. Kafka's writings are indeed prophetic intimations of the logic of the concentration camps; the monstrous insinuation inherent in his prophecies is that the exterminator is not wholly in the wrong, that his hold over his victim is something more than a matter of superior might, for the victim cooperates in his own destruction.

See also Alienation; Consciousness; Existentialism; Guilt; Metaphor.

Bibliography

WORKS BY KAFKA

Most of Kafka's writings were published posthumously and against his express wishes by his friend Max Brod. The complete edition is *Werke* (Frankfurt, 1952–). The "definitive" English edition, published in London, includes *The Trial,* translated by Willa Muir and Edwin Muir (1945); Kafka's *Diaries,* 2 vols., translated by J. Kresh, M. Greenberg, and H. Arendt (1948–1949); *America,* translated by Willa Muir and Edwin Muir (1949); *In the Penal Settlement: Tales and Short Prose Works,* translated by Willa Muir and Edwin Muir (1949); *The Castle,* translated by Willa Muir and Edwin Muir (1953); *Wedding Preparations in the Country and Other Posthumous Prose Writings,* translated by Ernst Kaiser and Eithne Wilkins (1954); and *Description of a Struggle and The Great Wall of China,* translated by Willa Muir and Edwin Muir and Tania Stern and James Stern (1960). See also Kafka's *Letters to Milena* (Jesenská), translated by Tania Stern and James Stern (London, 1953), and G. Janouch's *Conversations with Kafka* (New York, 1953).

WORKS ON KAFKA

Three biographical studies are available: Max Brod, *Franz Kafka: Eine Biographie* (Frankfurt, 1937), translated as *The Biography of Franz Kafka* (London: Secker and Warburg, 1947); K. Wagenbach, *Franz Kafka: Eine Biographie seiner Jugend, 1883–1912* (Bern: Francke, 1958); and P. Eisner, *Franz Kafka and Prague* (New York: Arts, 1950).

For critical works on Kafka see G. Anders, *Franz Kafka* (London: Bowes and Bowes, 1960); Ronald D. Gray, ed., *Kafka: A Collection of Critical Essays* (Englewood Cliffs, NJ: Prentice-Hall, 1962), which has important contributions by Albert Camus and E. Heller; and Heinz Politzer, *Franz Kafka: Parable and Paradox* (Ithaca, NY: Cornell University Press, 1962).

OTHER RECOMMENDED SOURCES

Biemel, Walter. "Franz Kafka: The Necessity for a Philosophical Interpretation of His Work." In *Explanation and Value in the Arts*, edited by Salim Kemal and Ivan Gaskell. New York: Cambridge University Press, 1993.

Cooper, Gabriele von Natzmer. *Kafka and Language: In the Stream of Thoughts and Life*. Riverside: Ariadne Press, 1991.

Heidsieck, Arnold. *The Intellectual Contexts of Kafka's Fiction: Philosophy, Law, Religion*. Columbia: Camden House, 1994.

Wade, Geoff. "Marxism and Modernist Aesthetics: Reading Kafka and Beckett." In *The Politics of Pleasure: Aesthetics and Cultural Theory*, edited by Stephen Regan. Buckingham; Philadelphia: Open University Press, 1992.

J. P. Stern (1967)
Bibliography updated by Desiree Matherly Martin (2005)

KAIBARA EKKEN
(1630–1714)

Kaibara Ekken, or Ekiken, a Japanese Confucianist influential in popularizing Confucian ethics among ordinary people, was born in Fukuoka. The son of a physician, he became a doctor himself, then left medicine to become a Zhu Xi neo-Confucianist. His teachers in Kyoto were Kinoshita Junan (1621–1698) and Yamazaki Ansai. At thirty-nine Kaibara returned to Fukuoka, where he spent the rest of his life in the service of the Kuroda fief. Blessed with an extraordinary capacity for work but little originality, he wrote on many subjects. He became an important botanist with the issuing of separate books on the vegetables, the flora, and the medicinal herbs of Japan. His books on education were pioneering works in pedagogy; *Onna daigaku* (The great learning for women), the standard book on women's ethics in the Tokugawa era, is attributed variously to him and to his well-educated wife. His books were a great success. Unlike most Confucianists, who wrote in Chinese, he wrote in Japanese; furthermore, his teaching was highly practical, applying Confucian morality to everyday life. His pedagogical ideas were not equalitarian (he assigned to women the role of mere submissiveness and obedience to their husbands), and his botanical studies were not at all scientific in the modern sense, but he played an important role in spreading education.

Kaibara's philosophical importance today rests on his *Taigiroku* (The great doubt), in which he aired his dissent with the official doctrine of the Zhu Xi school. Kaibara was also critical of the "ancient learning" school of Confucianism and its scholars Itō Jinsai and Ogyū Sorai, and of the Wang Yangming school, the rival of Zhu Xi. Kaibara disagreed with Zhu Xi Confucianism in his elevation of *ki,* the material force, over *ri,* the principle immanent in all things. For him *ki* is the "great limit" or the "ultimate" and is an all-pervading life force. Kaibara does not distinguish the original form of human nature from its acquired form; contrary to Zhu Xi, he is an optimist in his view of man and of the natural world. His cosmology is characterized by cosmic love that embraces all men, born as they are of heaven and earth. Man's indebtedness to nature is limitless, and for him the Confucian virtue of *jen,* "humaneness," comes close to being a religious benevolence, first toward nature and then toward men. His practical bent, however, makes it difficult to clarify his position, which seems to be one of eclectic doubt rather than critical inquiry. In administrative matters Kaibara opposed imitating Chinese ways; rather he was an ardent patriot, loyal in support of the emperor.

See also Chinese Philosophy; Itō Jinsai; Japanese Philosophy; Ogyū Sorai; Wang Yang-ming; Yamazaki Ansai; Zhu Xi (Chu Hsi).

Bibliography

Kaibara's works are available in Japanese in *Ekken zenshū* (Complete works of Kaibara Ekken), edited by Ekkenkai, 8 vols. (Tokyo, 1911). A secondary source in Japanese is Inoue Tadashi, *Kaibara Ekken* (Tokyo, 1963).

See also O. Graf, *Kaibara Ekiken* (Leiden: Brill, 1942); S. Atsuharu, "Kaibara E. and *Onna daigaku,*" in *Cultural Nippon* 7 (4) (1939): 43–56; and W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 374–377.

Gino K. Piovesana, S.J. (1967)

KALON

Kalon: the neuter of the Greek adjective *kalos*, beautiful, fine, also admirable, noble; accompanied by the definite article (*to kalon*), for example, *the* beautiful (or beauty). In Greek culture, what is *kalon* is typically the object of *erôs*, passionate or romantic love, and in (male-dominated) literature (and art), the term is predominantly applied to males around the age of puberty. Plato appropriates the *kalon* (along with the good and the just) as a key object for human striving and understanding in general, discovering in it, along with the good, one of the properties of the universe and of existence; *erôs* itself, in Plato, is transformed from a species of love into love or desire *tout court*, for whatever is truly desirable—and good (for the human agent). See especially his *Symposium*, *Phaedrus* (*Hippias Major*, possibly not by Plato, represents an unsuccessful attempt to define the *kalon*). The truly beautiful, or fine, is identical with the truly good, and also with the truly pleasant, as it is for Aristotle (*Eudemian Ethics* I.1, 1214a1–8). The Aristotelian good man acts "for the sake of the fine (*to kalon*)" (*Nicomachean Ethics* IV.2, 1122b6–7), an idea which is sometimes used as a basis for attributing to Aristotle a quasi-Kantian view of the ideal agent as acting *morally*, even—if occasion arises—*altruistically*, as opposed to acting out of a concern for his or her own good or pleasure. Against this, we need to take account of Aristotle's treatment of his good person as a *self-lover*, someone who seeks a disproportionate share of *the fine* for himself or herself (*NE* IX.8, 1169a35–b1), though he or she may willingly concede his or her share to a friend (*NE* IX.8, 1169a32–34). This is consistent with Aristotle's wanting

to treat the fine (or the admirable) as itself part—the most important part—of the human good; and indeed, he ultimately seems to recognize only two objects of desire, the good and the pleasant (*NE* VIII.2, 1155b18–21; cf. e.g. *EE* VII.2, 1235b18–23). In this context *the pleasant* will include only those pleasures that are not *fine* and good. For this move we may compare Plato's *Gorgias* (474C–475D), where Socrates actually reduces *fine* to *good*, *pleasant*, or *both*. Later Greek philosophy trades on, while sometimes modifying, this complex of ideas, which also forms the basis for the analysis of beauty in literature or in the visual arts.

See also Aristotle; Beauty; Good, The; Plato; Pleasure; Socrates.

Bibliography

Aristotle. *Eudemian Ethics*. In *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Aristotle. *Nicomachean Ethics*. Translated by Christopher Rowe. In *Aristotle: Nicomachean Ethics*, edited by Sarah Broadie and Christopher Rowe. Oxford: Oxford University Press, 2002.

Plato. *Gorgias*. In *Plato, Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.

Plato. *Hippias Major*. In *Plato, Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.

Plato. *Lysis*. Translated by Christopher Rowe. In *Plato's Lysis*, edited by Terry Penner and Christopher Rowe. Cambridge Studies in the Dialogues of Plato. Cambridge, U.K.: Cambridge University Press, 2005.

Plato. *Phaedrus*. Translated by Christopher Rowe. Harmondsworth: Penguin Books, 2005.

Plato. *Symposium*. Translated by Christopher Rowe. In his *Plato: Symposium*. Warminster/Oxford: Aris & Phillips/ Oxbow Books, 1998.

Christopher Rowe (2005)

KAMES, LORD

See *Home, Henry*

KANT, IMMANUEL

(1724–1804)

Immanuel Kant, the propounder of the critical philosophy, was born at Königsberg in East Prussia; he was the son of a saddler and, according to his own account, the grandson of an emigrant from Scotland. He was educated at the local high school, the Collegium Fridericianum, and then at the University of Königsberg, where he had the good fortune to encounter a first-class teacher in the philosopher Martin Knutzen. After leaving the university, about 1746, Kant was employed for a few years as a tutor in a number of families in different parts of East Prussia. He kept up his studies during this period and in 1755 was able to take his master's degree at Königsberg and to begin teaching in the university as a *Privatdozent*. He taught a wide variety of subjects, including physics, mathematics, and physical geography as well as philosophy, but nevertheless remained poor for many years. It was not until 1770, when he was appointed to the chair of logic and metaphysics at Königsberg, that his financial stringencies were eased.

Kant's first book, *Gedanken von der wahren Schätzung der lebendigen Kräfte* (Thoughts on the True Estimation of Living Forces), was published as early as 1747 (Königsberg), and between 1754 and 1770 he produced an impressive stream of essays and treatises. His earlier works are primarily contributions to natural science or natural philosophy, the most notable being his *General History of Nature and Theory of the Heavens* of 1755; it was not until after 1760 that philosophical interests in the modern sense became dominant in his mind. Kant's publications had already won him a considerable reputation in German learned circles by the time he obtained his professorship. The ten years following his appointment form a period of literary silence during which Kant was engaged in preparing his magnum opus, the *Critique of Pure Reason*. The appearance of the *Critique* was eagerly awaited by Kant's friends and philosophical colleagues, but when it at last came out in 1781 the general reaction was more bewilderment than admiration. Kant tried to remove misunderstandings by restating the main argument in the *Prolegomena to Every Future Metaphysics* of 1783 and by rewriting some of the central sections of the *Critique* for a second edition in 1787. At the same time he continued, with most remarkable energy for a man of his years, the elaboration of the rest of his system. By 1790 the *Critique of Practical Reason* and the *Critique of Judgment* were in print, and of the major treatises only *Religion within the Bounds of Mere Reason* (1793) and *Metaphysic of Morals* (1797) had still to appear. Kant then enjoyed a tremendous reputation throughout Germany and was beginning to be known, though scarcely to be understood, in other European countries. In his declining years, however, he suffered the mortification of seeing some of the ablest young philosophers in his own country, among them Johann Gottlieb Fichte, Friedrich von Schelling, and J. S. Beck, proclaim that he had not really understood his own philosophy and

propose to remedy the deficiency by producing "transcendental" systems of their own. There is reason to believe that the work on which Kant was engaged in the last years of his life was intended as a counterblast to such critics. But Kant was not able to complete it before his death, and all that remains of it are the fragments gathered together under the title *Opus Postumum*.

Kant's outer life was almost entirely uneventful. He never married. The one occasion on which he might have become politically prominent was in 1794 when, after the appearance of his book on religion, the Prussian king asked him not to publish further on a topic on which his views were causing alarm to the orthodox. But Kant duly promised, and no scandal ensued. For the rest, he fulfilled the duties of his professorship and took his turn as rector of the university; dined regularly with his friends; admired Jean-Jacques Rousseau and the French Revolution from afar; conversed eagerly with travelers who brought him news of a wider world he never saw himself. Never very robust in body, he carefully conserved his physical resources and was in good health until a relatively short time before his death. He was nearly eighty when he died.

CHARACTER OF KANT'S PHILOSOPHICAL WORK

Kant was the first of the major philosophers of modern times to spend his life as a professional teacher of the subject. He was required by university regulation to base his philosophy lectures on particular texts, and he used for this purpose not the works of such major thinkers as René Descartes and John Locke, but the handbooks of his professorial predecessors, notably Christian Wolff, Alexander Gottlieb Baumgarten, and G. F. Meier. Wolff and Baumgarten had dressed out the philosophy of Gottfried Wilhelm Leibniz in what they took to be decent academic garb, presenting Leibniz' thoughts in the form of a system and with an air of finality foreign to the original; Meier did the same for the doctrines of formal logic. Their example had a near-fatal effect on Kant, for he too thought that philosophy must be thorough if it is to be academically respectable—meaning, among other things, technical and schematic.

In the *Critique of Pure Reason* Kant set out his theories in what he later called progressive order, starting from what was logically first and working forward to familiar facts; in that work he also employed an elaborate terminology of his own and an apparatus of "parts," "divisions," and "books" whose titles are alarming and whose appropriateness to the subject matter is not immediately obvious. It is not surprising that his first readers were unable to discover what the work as a whole was about. The *Critique of Practical Reason* and the *Critique of Judgment* were still more pedantic in form, since in them Kant persisted with much of the formal framework already used in the *Critique of Pure Reason*, in each case proceeding from a part labeled "Analytic" to another labeled "Dialectic," uncovering one or more "antinomies" in dealing with the dialectic, and ending with an untidy appendix irrelevantly titled "Doctrine of Method." The fact that Kant was already an old man when he composed these works doubtless explains his attachment to what some commentators have called his architectonic; it is a major obstacle to the proper grasp and unprejudiced evaluation of his ideas. Yet, as passages in his ethical writings in particular show, Kant was capable of expounding his thoughts with clarity, even with eloquence. He was not by nature a bad writer, but he accepted uncritically the scholastic manner cultivated by his fellow professors.

The first task in reading Kant is thus to cut through the formal academic dress in which he clothes his opinions. When this is done, what emerges is not a provincial pedant like Wolff or Baumgarten, but a person of remarkable intellectual and moral stature. Kant's knowledge of the major European philosophers was often no more than superficial, and his estimate of the work of some of his own contemporaries was certainly overgenerous. But he had, for all that, a sure sense of what was intellectually important at the time; he alone among the eighteenth-century philosophers at once appreciated the greatness of Isaac Newton and was fully aware of the challenge for ethics Newton's work presented once its seemingly deterministic implications were understood. To sum up Kant's mature philosophy in a single formula: He wished to insist on the authority of science and yet preserve the autonomy of morals. To achieve this result was a gigantic task, involving consideration of the whole question of the possibility of metaphysics as well as the construction of a theory of scientific knowledge and the elaboration of an ethical system.

Nor was Kant one to be content with mere generalities; he sought to work out his position in detail, with many specific arguments, as well as to state a general case. But the obscurities of his language combine with the extent of his intellectual ambitions to prevent the average reader from grasping precisely what Kant was after; individual points are picked up, but the shape of the whole is not discerned. Yet to be fair to Kant the reader must see the individual views in the wide setting in which Kant saw them himself. To estimate their philosophical value with-

out taking account of their position in the Kantian system, as many critics have tried to do, is quite indefensible.

PRECRITICAL WRITINGS

Kant's philosophical career is commonly divided into two periods, that before 1770, usually referred to as "precritical," and that after 1770, usually referred to as "critical." The word *critical* comes from Kant's own description of his mature philosophy as a form of "critical idealism," an idealism, that is to say, built on the basis of a critique of the powers of reason. The precritical period of Kant's thought is interesting primarily, though not exclusively, for its anticipations of his later ideas. Kant was educated by Knutzen in the Wolff-Baumgarten version of Leibniz, and he was, like his master, an independent Leibnizian from the first, although it was many years before he made a decisive break with the Leibnizian way of thinking. The main influence operating against Leibniz in Kant's early thought was Newton, to whose work he had also been introduced by Knutzen. In the more narrowly philosophical field another independent Leibnizian, Christian August Crusius, proved an important subsidiary influence. Just when David Hume awakened Kant from his "dogmatic slumber" is uncertain, but it seems likely that Kant had moved some way in the direction of empiricism before that event took place.

CAUSATION. How little the early Kant had learned from Hume can be seen from some of his first metaphysical essays. In the *Principium Primorum Cognitionis Metaphysicae Nova Dilucidatio* (Königsberg, 1755) he discoursed in effect on the subject of causality, discussing at length the relationship of the Leibnizian principle of sufficient reason to the logical principles of identity and contradiction. Kant knew at this stage, as Crusius did, that Wolff's attempt to subordinate the real to the logical was a mistake, but he had only a hazy idea of what he was later to call the synthetic nature of propositions asserting real connections. He moved a step nearer his mature view in the 1763 essay on negative quantities (*Versuch, den Begriff der negativen Grössen in die Weltweisheit einzuführen*, Königsberg) when he pointed out that opposition in nature is quite different from opposition in logic: Two forces acting against one another are quite unlike a proposition in which the same predicate is simultaneously affirmed and denied. But in none of his writings of the time did Kant explicitly raise the question of the sphere of application of the causal principle, as Hume did.

EXISTENCE. Kant's failure to press home his questions on causation is paralleled in his otherwise striking treatment of existence in another work published in 1763, "The Only Possible Ground of Proof of God's Existence." He began this work by declaring that even if the proposition that existence is no predicate or determination of anything seems "strange and contradictory," it is nevertheless indubitable and certain. "It is not a fully correct expression to say: 'A sea unicorn is an existent animal'; we should put it the other way round and say: 'To a certain existing sea animal there belong the predicates that I think of as collectively constituting a sea unicorn.'" On these grounds Kant rejected the Cartesian version of the Ontological Argument. But he held, even so, that an alternative conceptual proof of God's existence could be found: Nothing could be conceived as possible unless (as the point had already been put in the *Nova Dilucidatio*) "whatever of reality there is in every possible notion do exist, and indeed, absolutely necessarily. … Further, this complete reality must be united in a single being." There must, in other words, be a perfect being if there are to be any possibilities. Kant was to recall this proof in his derivation of the idea of the *ens realissimum* in the *Critique of Pure Reason*, but he then no longer believed that it had constitutive force. His treatment of attempts to produce causal proofs of God's existence in the *Critique* was also altogether more trenchant than in the precritical works, for though he saw there that the ordinary First Cause Argument was unsatisfactory, he regarded the Argument from Design as generally acceptable, even if not logically compulsive.

METAPHYSICAL PROPOSITIONS. Kant was more successful in another treatise written at the same period, "Untersuchungen über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral" (On the Distinctness of the Principles of Natural Theology and Morals; 1764). The Berlin Academy had proposed the question, Are metaphysical truths generally, and the fundamental principles of natural theology and morals in particular, capable of proofs as distinct as those of geometry? If not, what is the true nature of their certainty? Kant answered by drawing a series of radical distinctions between argument in philosophy and argument in mathematics. The mathematician starts from definitions that are in effect arbitrary combinations of concepts; the philosopher must work toward definitions, not argue from them, since his business is to "analyze concepts which are given as confused." Mathematics contains few unanalyzable concepts and indemonstrable propositions; philosophy is full of them. Then too, the relationship

between mathematical ideas can always be observed *in concreto*, whereas the philosopher, having nothing to correspond to mathematical diagrams or symbolism, necessarily works on a more abstract level. The lesson of all this might seem to be that philosophical truths are incapable of strict demonstration, but Kant did not draw this conclusion in the case of natural theology, where he held to his attempted conceptual proof, though he inclined toward it in respect to "the primary grounds of morals." In general, Kant's tendency was to say that metaphysics must be an analytic activity that should follow a method that is fundamentally Newtonian: "It is far from the time for proceeding synthetically in metaphysics; only when analysis will have helped us to distinct concepts understood in their details will synthesis be able to subsume compounded cognitions under the simplest cognitions, as in mathematics" (*Critique of Practical Reason and Other Writings*, Beck translation, 1949, p. 275).

Kant viewed the prospects of attaining genuine metaphysical knowledge with increasing skepticism as the 1760s went on. In the enigmatic *Dreams of a Spirit-Seer* of 1766 he compared the thought constructions of metaphysics to the fantasies of Swedenborg, in a manner that is scarcely flattering to either. Metaphysical contentions are groundless, since metaphysical concepts such as spirit cannot be characterized in positive terms. To survive, metaphysics must change its nature and become a science of the limits of human knowledge. Kant's skepticism about metaphysics was increased by his discovery of the antinomies, which is often dated 1769 although something like the third antinomy is to be found in the *Nova Dilucidatio*. Astonishingly, however, in his inaugural dissertation in 1770 he reverted in some degree to the old dogmatic conception of the subject and argued for the possibility of genuine knowledge of an intelligible world. But the main interest of the dissertation lies in its account of sensory knowledge, which prepared the way for the fundamental criticisms of metaphysical pretensions in the *Critique of Pure Reason*.

THE INAUGURAL DISSERTATION

Kant's Latin dissertation, "On the Form and Principles of the Sensible and Intelligible Worlds," publicly defended on August 21, 1770, was his inaugural lecture as professor of logic and metaphysics at Königsberg. At least one of the themes of the dissertation, the status of the concept of space, represented a long-standing interest. As early as 1747 Kant had argued that the proposition that space has three dimensions is contingent; given a different law of the effects of different substances on one another, "an extension with other properties and dimensions would have arisen. A science of all these possible kinds of space would undoubtedly be the highest enterprise which a finite understanding could undertake in the field of geometry" ("Living Forces," Handyside translation, in *Kant's Inaugural Dissertation and Early Writings on Space*, p. 12). Later, however, he regarded three-dimensionality as a necessary property of space, and used its necessity as a ground for rejecting Leibniz' account of the concept. In a short essay on space published in 1768 Kant had seemed to suggest that Newton's view of space as an absolute reality was the only alternative to Leibniz, but in the dissertation he rejected both theories and widened his treatment of the question so that it covered time as well as space. Despite this extension the dissertation is best viewed as directed mainly against Leibniz.

SPACE AND TIME. In general, Leibniz had followed the other great rationalists in interpreting perception as a confused form of thinking. Like Descartes, he had treated the deliverances of the senses as sometimes clear but never distinct. In the dissertation Kant developed two main arguments against this position. He maintained in the first place that it could not do justice to the special character of space and time, which are not, as Leibniz supposed, systems of relations abstracted from particular situations and confusedly apprehended, but rather unique individuals of which clear knowledge is presupposed in all perceptual description. The ideas of space and time are intuitive rather than conceptual in character; moreover, they are "pure" intuitions insofar as the essential nature of their referents is known in advance of experience and not as a result of it.

SPACE AND GEOMETRY. To reinforce this point Kant brought forward his second argument, that Leibniz' theory could not account for the apodictic character of geometry. There was, Kant supposed, an essential relation between geometry and space, for geometry "contemplates the relations of space" and "does not demonstrate its universal propositions by apprehending the object through a universal concept, as is done in matters of reason, but by submitting it to the eyes as a singular intuition, as is done in matters of sense" ("Dissertation," in *Kant's Inaugural Discussion and Early Writings on Space,* Sec. 15 C). But if space is what Leibniz said it was and if, as Kant added, "all properties of space are borrowed only from external relations through experience," then:

> geometrical axioms do not possess universality, but only that comparative universality which is acquired through induction and holds only so

> widely as it is observed; nor do they possess necessity, except such as depends on fixed laws of nature; nor have they any precision save such as is matter of arbitrary convention; and we might hope, as in empirical matters, some day to discover a space endowed with other primary affections, and perhaps even a rectilinear figure enclosed by two straight lines. (Sec. 15 D)

Kant's own account of space at this stage was that it "*is not something objective and real*, neither substance, nor accident, nor relation, but [something] *subjective and ideal*; it is, as it were, a schema, issuing by a constant law from the nature of the mind, for the co-ordinating of all outer sensa whatever" (Sec. 15D). One major advantage of this subjectivist view, in Kant's eyes, was that it explains the possibility of applying geometry to the physical world. Space being a universal form of sensibility, "nothing whatsoever … can be given to the senses save in conformity with the primary axioms of space and the other consequences of its nature, as expounded by geometry" (Sec. 15 E).

APPEARANCE AND REALITY. Kant's view had another, more startling implication, namely that we cannot know things as they really are through sense perception. If space and time are contributed by the knowing mind, spatial and temporal objects will be altered in the very act of being apprehended. It follows that the world known through the senses—the world investigated by the physical sciences and familiar in everyday experience—can be no more than a phenomenal world. Kant was prepared to accept this conclusion in the dissertation, but he balanced it by saying that over and above this phenomenal world is another world of real objects, knowable not by the senses but by reason. Reason lacks intuitive powers—we cannot be acquainted with things as they are. But (and in this the contrast with the *Dreams* is at its strongest) reason possesses certain concepts of its own, among them "possibility, existence, necessity, substance, cause," by means of which it can arrive at a "symbolic cognition" of such things; that is, know some true propositions about them. The intellect, in its real as opposed to its logical use, can form the concept of a perfect being and use this both to measure the reality of other things and for moral purposes.

ACHIEVEMENTS. The doctrine of pure intellectual concepts in the dissertation is at best impressionistic and had to be completely rethought in the ten years that followed. But against this may be set Kant's positive achievements in the dissertation, seen from the point of view of his future work. First, Kant had convinced himself that there is an absolute difference between sensing and thinking, and that sense experience need not be in any way confused. Second, he had worked out the main lines, though by no means all the details, of what was to be his mature theory of space and time. Third, he had revived the old antithesis of things real and things apparent, objects of the intellect and objects of the senses, to cope with the consequences of his views about space and time; in this way he was able to show (or so he thought) that physics gives us genuine knowledge, though only of appearances, and that the task of telling us about things as they really are is reserved for metaphysics. Fourth and last, he had recognized the existence of a special class of concepts, "given through the very nature of the intellect," and had seen that these have an important bearing on the question of the possibility of metaphysics.

What Kant had not done was to pose the problem of metaphysics with all its wider implications. As in the *Dreams*, he treated the question whether we have any knowledge of a world of pure spirit as one that is asked primarily for its theoretical interest. It was intellectual curiosity, that is to say, which at this stage prompted Kant to inquire whether physics and metaphysics could coexist, and, if they could, what should be said of their respective objects. He retained this curiosity when he wrote the *Critique of Pure Reason*, but it was not by then his only motive. For he had seen by 1781 that the question of the possibility of metaphysics was important not only to the academic philosopher, but because of its bearing on the universally interesting topics of God, freedom, and immortality, to the plain man as well; that it was a matter not just of intellectual, but also of moral, concern.

CRITIQUE OF PURE REASON: THEME AND PRELIMINARIES

Kant's principal task in the *Critique of Pure Reason* was to determine the cognitive powers of reason, to find out what it could and could not achieve in the way of knowledge. The term *reason* in the title was intended in its generic sense, to cover the intellect as a whole; Kant was not exclusively interested in the reason that he himself distinguished from and opposed to understanding. He was, however, particularly concerned with the capacities of "pure" reason, that is, with what reason could know when operating by itself and not in association with another faculty. Kant believed it important to answer this question for two reasons. He saw that there are spheres (mathematics, for instance) in which it is plausible to claim that pure reason is a source of important truths. He

also saw that in another field, that of metaphysics, remarkable claims were advanced on reason's behalf: It was alleged that, by simply thinking, we could arrive at ultimate truth about the world, establishing thus a series of propositions whose certainty was unassailable and whose subject matter was of supreme importance. Kant, who had himself made this sort of claim in the dissertation, never doubted that what the metaphysician wants to say matters, but he did question his competence to say it. The fact that reason "precipitates itself into darkness and contradictions" once it enters this field struck him as deeply significant; the "intestine wars," the interminable disputes, of metaphysicians could only mean that their claims were pitched too high.

Nor was the scandal of metaphysics—the fact that nothing in metaphysics could be regarded as settled—of concern only to metaphysicians. By failing to make good his proofs, the metaphysician brought doubt on the acceptability of his conclusions, including such fundamental articles of belief as that God exists and that the will is free. In proposing a radical reexamination of the capacities of pure reason, Kant's ultimate motive was to safeguard such convictions by making clear that although they cannot be matters of knowledge, they can all the same be held to as matters of what he called pure rational faith.

TYPES OF JUDGMENT. In the preface to the *Critique*, Kant formulates his main question as "how much can understanding and reason know apart from all experience?" (A xvii). (The first edition is customarily referred to as A, the second edition as B.) In the introduction, he takes his first step toward an answer by substituting the formula "How are synthetic *a priori* judgments possible?" Two closely connected sets of distinctions lie behind these celebrated words. First, Kant distinguishes propositions that are a priori from all others; an a priori judgment "in being thought is thought as *necessary*" and is also thought "with strict universality, that is, in such a manner that no exception is allowed as possible" (B 3–4). A priori judgments have the twin characteristics of necessity and universality, neither of which can be found in conclusions from experience.

In holding that experience can present us with no more than contingent truths Kant echoes the views of many of his predecessors. But in his other distinction, between synthetic and analytic judgments, he shows greater originality. A judgment is analytic, he explains, if what is thought in the predicate-concept has already been thought in the subject-concept; a judgment is synthetic if this condition does not obtain. Thus, "All bodies are extended" is analytic because our idea of a body is of something that is extended or occupies space; "All bodies have weight" is synthetic because the notion of weight is not comprised in the notion of body (we learn by experience that bodies have weight). In analytic judgments, again, the connection of subject and predicate is "thought through identity"; or, as Kant puts it elsewhere in the *Critique*, the highest principle of all analytic judgments is the principle of contradiction. It follows from this that every analytic judgment is a priori in that it is true or false without regard to experience; every analytic judgment is either necessarily true or necessarily false, and we establish its truth or falsity by reference only to definitions of the terms it contains and to the principle of contradiction. Synthetic judgments, by contrast, require for their authentication a different sort of reference, since in their case the connection of subject and predicate terms is "thought without identity." In the case of everyday judgments of fact, for example, we need to consult experience to see whether the connection asserted actually holds.

So far Kant's distinction is simply a more elaborate version of Hume's division of propositions into those that assert relations of ideas and those that express matters of fact and existence, a version inferior to Hume's in that it is formally tied to statements of the subject-predicate form. But at this point Kant gives the distinction a fresh twist by asserting that there are judgments that are both synthetic and a priori, thus cutting across the usual classifications. Nearly all the propositions of mathematics answer this description, according to Kant; he also thinks it obvious that "*natural science (physics) contains* a priori *synthetic judgments as principles*." He gives two examples: "in all changes of the material world the quantity of matter remains unchanged; and … in all communication of motion action and reaction must always be equal" (B 17). The very existence of these judgments shows that reason has special cognitive powers of its own, and so lends plausibility to the claims of metaphysicians. But before accepting the claims of metaphysicians, Kant suggests, we need to ask ourselves how (under what conditions) it is possible to assert judgments of this type in the two fields concerned. Only when this question is answered can we decide whether metaphysicians can draw support from the example of mathematics and "pure" physics. This inquiry is what Kant is concerned with in the first half of the *Critique*.

ANALYTIC AND SYNTHETIC. The terms in which Kant states his problem seem at first sight clear, but the clarity diminishes on closer inspection. There is the criticism

that he offers a dual account of the analytic-synthetic distinction, once in psychological and once in logical terms, and the criticism that reference to the principle of contradiction alone is inadequate for the logical formulation of the distinction (he should have referred to logical laws generally). Apart from these two matters, Kant's treatment is marred by a failure to offer any discussion of his key idea, "what is thought in a concept." This omission is the more remarkable because Kant in fact had views on the subject of definition, views that are hard to reconcile with his apparent assumption that every judgment is unequivocally analytic or synthetic. Elsewhere in the *Critique* he states that, according to the real meaning of "definition," an empirical concept "cannot be defined at all, but only made explicit" (B 755). He means that we cannot give the "real essence" (in Locke's terminology) of such a concept, but only its "nominal essence," or conventional signification, which is liable to change as knowledge increases or interests shift. If this is correct, it seems to be only by convention, or provisionally, that the judgment "All bodies are extended" is analytic and the judgment "All bodies have weight" synthetic.

Nor is Kant's other distinction, between a priori and a posteriori, as simple as he pretends. He tries to clarify it by explaining that the first class of judgments have the characteristics of necessity and universality, which serve as criteria that are "inseparable from one another." He fails to notice, however, that the necessity that belongs to synthetic a priori judgments must on his own account differ from that which characterizes analytic judgments. Analytic judgments are, or rather claim to be, logically necessary—to deny a true analytic judgment would be, if Kant is correct, to dispute the validity of the law of contradiction. But though no synthetic judgment can contravene the laws of logic, none can be true in virtue of these laws and of meanings alone. Accordingly, if any synthetic judgment is to be described as necessary, it must be necessary in some further sense.

Kant recognizes in practice that the synthetic a priori judgments he takes to be valid have their own special kind of necessity. In his own terminology, they are "transcendentally" necessary; necessary, that is to say, if we are to have the knowledge and experience we actually have. But he would have done better to acknowledge the ambiguity in his term *a priori* from the outset. It would also have been helpful had he given some elucidation of his statement that, when a judgment is thought with strict universality, "no exception is allowed as possible." He cannot mean that no exception is logically possible, or every a priori judgment would be analytic. But he does not, at least at this early stage, make clear what other sort of possibility he has in mind.

TRANSCENDENTAL AESTHETIC

Kant's next step in the solution of the problem of how synthetic a priori judgments are possible is to examine the two types of case in which, in his view, we undoubtedly can make synthetic a priori judgments, and then to exhibit the bearing of his results on the possibility of metaphysical knowledge. In his short but important *Prolegomena to Every Future Metaphysics* he approaches these tasks directly. In the *Critique* itself his method is more roundabout, since he proposes there to delineate the entire cognitive powers of the mind and so to clarify the background against which synthetic a priori judgments are made. This leads him to undertake an inquiry first into the a priori elements involved in sensory knowledge (the "Transcendental Aesthetic") and then into the corresponding elements involved in thought (the "Transcendental Logic"). The sharp distinction between the senses and the intellect argued for in the dissertation is the obvious basis of this division.

A PRIORI INTUITIONS. It seems at first sight contradictory to say that there might be a priori elements involved in sensory knowledge. According to an old philosophical and psychological tradition, sensation is an essentially passive affair; the senses present us with data and we have no choice but to accept. Kant was quite ready to agree to this as a general account of sensation. But he was persuaded that there are some features of sensory experience that cannot be accepted as empirically given.

Kant identifies these features by a process similar to that in the dissertation: an examination of our ideas of space and time. These ideas, he argues, represent the form of experience rather than its matter; through them we structure the sensory given in the very act of sensing it. To establish this position Kant appeals to a variety of considerations.

First, he insists on the fundamental and ubiquitous character of space and time, as opposed to features like color and sound. Spatial predicates apply to whatever we know through the five senses, temporal predicates both to these and to the immediately experienced flow of our inner lives. Second, he argues that we cannot acquire the ideas of space and time by reflecting on what is empirically given. Some philosophers had said that we come by the idea of space by noticing such things as that one object is adjacent to another, and that we come by the idea of time by observing the way in which events suc-

ceed, are simultaneous with, or precede one another. Kant points out that the very description of such situations presupposes familiarity with space and time as such. For to know what is meant by saying that one thing is "next to" or "on top of" another we need to appreciate how the things in question are situated in a wider spatial framework, which in turn falls within a yet wider spatial system, until we come to the thought of space as a whole. Particular spaces are not instances of space, but limitations of it, and space is accordingly a special sort of particular. The same argument applies to time. Adding to these two points the fact that we know certain things to be necessarily true of space and time (space has only three dimensions, different times are not simultaneous but successive), Kant infers that the ideas of space and time are not only "intuitions," but "*a priori* intuitions."

MATHEMATICS. Kant finds confirmation for his view of space and time exactly as he had in the dissertation: in the thought that this view alone can explain the possibility of pure and applied mathematics. Pure geometry is possible because we are able to "construct," or show the real possibility of, its concepts in pure intuition. An experiment conducted in imagination shows at once that a triangle is a real spatial possibility, whereas a figure bounded by two straight lines is not. Applied geometry is possible because whatever is apprehended by the senses must necessarily accord with the forms of sensibility. Kant attempts at various points in his writings to extend his doctrine of the importance of pure intuition for mathematical thinking from geometry to the other parts of mathematics, but it cannot be said that he is ever convincing on this point. His reasons for saying that "seven and five are twelve" is a synthetic proposition were sharply and properly criticized by Gottlob Frege. His account of algebra (B 745, 762) is so sketchy as to be virtually unintelligible. Kant tries to say that in algebra there is a "symbolic construction" corresponding to the "ostensive construction" of the concepts of geometry, but it is not in the least clear what this has to do with the pure intuition of either space or time.

Some critics speak as if Kant's failure to produce a satisfactory philosophy of mathematics invalidated the whole "Aesthetic," and it is true that the central point of this part of his work is destroyed if his main contentions about mathematics are rejected. Kant's explanations fall to the ground if it turns out that there is no intrinsic connection between mathematics and space and time, or if it is held that mathematical propositions are analytic, not synthetic a priori. But it does not immediately follow that the whole Kantian doctrine of space and time must be rejected, for many of his arguments on this matter are independent of his philosophy of mathematics. Nor is it decisive against him that the treatment of space and time in modern physics is very different from his; he claims to be dealing with the space and time of immediate perception.

SIGNIFICANCE. Apart from the questions about truth, however, it is vital to appreciate the importance of the conclusions of the "Aesthetic" in the economy of the *Critique of Pure Reason* as a whole. The "transcendental ideality" of space and time carries with it, for Kant, the proposition that whatever we know through the senses (including "inner sense") is phenomenal; Kant's celebrated distinction between appearances and things-in-themselves has its origin, if not its justification, at this point. And the view that space and time are a priori forms of intuition is not only the model on which Kant constructed his theory of categories as concepts embodying the pure thought of an object in general; the view is carried over intact into the "Transcendental Analytic," and plays a crucial part there. To treat the theories of the "Aesthetic" as if they merely embodied a series of views that Kant had outgrown by the time he completed the *Critique*, as some commentators have proposed to do, is not in accord with Kant's own intentions. It is also to ignore a series of arguments that are of independent philosophical interest, and that demand careful notice from anyone writing on the philosophy of perception.

PURE CONCEPTS OF THE UNDERSTANDING

The main contentions of the aesthetic are to be found in the dissertation. Of the doctrine of pure intellectual concepts put forward in that inaugural lecture, on the other hand, almost nothing survives in the *Critique of Pure Reason*.

OBJECTIVE REFERENCE. In the dissertation Kant argues along two lines: First, that pure intellectual concepts are not derived from sense experience (they could not be described as "pure" if they were); and second, that they serve to give us information about things as they really are. Soon after writing this work, however, Kant realized that there was a fundamental difficulty in this position, a difficulty he stated at length in a letter to his friend Marcus Herz dated February 21, 1772. It was that of knowing how "pure" concepts could be said to determine an object of any kind. To elucidate the difficulty, Kant isolated two contrasting types of intelligence, *intellectus ectypus*, "which derives the data of its logical procedure from the

sensuous intuition of things," and *intellectus archetypus*, "on whose intuition the things themselves are grounded." The concepts of the first type of intelligence, deriving as they do from objects, have a guaranteed relationship to objects. The concepts of the second type determine objects, because, in this sort of case, thinking itself brings objects into existence in the same way in which "the ideas in the Divine Mind are the archetypes of things." But the human intelligence, as described in the dissertation, answers to neither description, for some of its concepts are not empirically derived and yet none of its thinking is creative in the sense specified. The problem then arises, How can these concepts be said to have objective reference; how can we know that in using them we are thinking about anything actual? It is this problem that Kant professes to have solved in the *Critique of Pure Reason*. Roughly speaking, his solution is that pure concepts can be shown to determine an object if the object is phenomenal. By contrast, when an attempt is made to use them to specify characteristics of "things in general," there is no guarantee that anything significant is being said.

ANALYTIC AND DIALECTIC. The details of Kant's explanation of how pure concepts can be said to have objective reference is to be found in the lengthy section of the *Critique* labeled "Transcendental Logic" and divided into two main parts, "Transcendental Analytic" and "Transcendental Dialectic."

The first part contains an inventory of what at this point Kant calls pure concepts of the understanding, or categories, with an account of the function they perform in human knowledge and a series of arguments purporting to show that, in the absence of such pure concepts, objective knowledge would be impossible for human beings. In addition, the "Analytic" lists the principles that rest on these pure concepts and offers independent proofs of these principles. Transcendental analytic is said by Kant to be a "logic of truth," insofar as "no knowledge can contradict it without at once losing all content, that is, all relation to an object, and therefore all truth" (B 87). It deals, in short, with the proper use of a priori concepts, which is the use they have when they provide a framework for empirical inquiries.

Transcendental dialectic is introduced as if it were merely the negative counterpart of analytic—as if its sole purpose were to expose the illusions generated when dogmatic philosophers, unaware of the sensuous conditions under which alone we can make successful use of a priori concepts, attempt to apply them outside the sphere of possible experience. In fact a large part of the section titled "Dialectic" is devoted to the exposure of metaphysical sophistries. But insofar as Kant recognizes in this part of his work the existence of a further set of intellectual operations involved in scientific inquiry, he seeks to show that the faculty of theoretical reason as well as that of the understanding has its appropriate pure employment.

JUDGMENT OR BELIEF. A good way to approach the central doctrines of the analytic is to see them as an intended answer to Hume. Kant's knowledge of Hume was limited—he had no firsthand acquaintance with the *Treatise of Human Nature*—but he grasped the importance of many of Hume's most challenging points. For instance, Hume had argued that "*belief is more properly an act of the sensitive, than of the cogitative part of our natures*" (*Treatise*, edited by L. A. Selby-Bigge, 1888, Book I, Part IV, Sec. 1, p. 183); in the last resort it is a matter of subjective conviction. It is one of Kant's main objects in the analytic to demonstrate that such a view cannot do justice to an all-important feature of what Hume calls belief and he calls judgment, namely, its claim to be true. When I judge that something is the case I do not merely commit myself to a certain assertion; there is a sense in which I commit all rational persons too, for I purport to state what holds objectively, that is to say for everyone. To make judgment primarily a matter of feeling, something private to an individual person, is to leave out what is most characteristic of it. Similarly, to explain thinking about matters of fact and existence in terms of the association of ideas, as Hume did, is to confuse the objective with the subjective, to put science on the level of idle reverie. Empirical thinking, to deserve its name, must proceed according to rules, and there is all the difference in the world between a rule, which cannot of its nature be private, and association, which is the connecting of ideas on a purely personal plane.

THE UNITY OF EXPERIENCE. There are many philosophers who would accept this criticism of Hume but would deny that empirical thinking involves not only rules, but rules that are a priori or necessary rules. To understand why Kant asserts that thinking must proceed according to necessary rules, we must explain his attitude to another of Hume's doctrines, the famous contention that "all our experimental conclusions proceed upon the supposition that the future will be conformable to the past" (*Enquiry concerning Human Understanding*, Sec. IV, Part II). Kant agrees with Hume that empirical knowledge involves connecting one part or element of experience with another; he agrees too that connection of this sort ("synthesis") proceeds on a principle that is neither

analytically true nor empirically probable. But he refuses to follow Hume in deriving the principle from "Custom or Habit," for he sees more clearly than Hume the consequences of adopting this "sceptical solution." If it were really the case that events were as "loose and separate" as Hume supposed, not only should we be deprived of any insight into the connections of things, but we should have no unitary consciousness of any sort. For it is a necessary condition of having a unitary consciousness that we be able to relate what is happening here and now to things and events that lie outside our immediate purview; if the ability to relate is not a real possibility, then neither is unitary consciousness. What Kant calls in one place (A 113) "the thoroughgoing affinity of appearances" (the fact that appearances are capable of being connected in a single experience) thus relates closely to the ability of the observer to recognize himself as a single person with diverse experiences. In fact the relation is one of mutual implication.

It may be useful to cite Kant's explanation as he gave it in the first edition of the *Critique*, in a passage in which all the most characteristic ideas of the "Analytic" appear and which also illustrates Kant's persistent but nonetheless questionable tendency to move from saying that unity of consciousness means that appearances must be capable of connection to the conclusion that they must be capable of connection according to universal and necessary laws.

> There can be in us no items of knowledge, no connection or unity of one item of knowledge with another, without that unity of consciousness which precedes all data of intuitions, and by relation to which representation of objects is alone possible. This pure original unchangeable consciousness I shall name *transcendental apperception*. … This transcendental unity of apperception forms out of all possible appearances, which can stand alongside one another in one experience, a connection of all these representations according to laws. For this unity of consciousness would be impossible if the mind in knowledge of the manifold could not become conscious of the identity of function whereby it synthetically combines it in one knowledge. The original and necessary consciousness of the identity of the self is thus at the same time a consciousness of an equally necessary unity of the synthesis of all appearances according to concepts, that is, according to rules, which not only make them necessarily reproducible but also in so doing determine an object for their intuition, that is, the concept of something wherein they are necessarily interconnected. (A 107–108)

ROLE OF CATEGORIES. If the synthesis of appearances is to proceed in accordance with necessary laws, we must clearly operate not just with empirical but also with a priori concepts. But this must not be taken to mean that some items or features of fact can be known apart from all experience. For the role of an a priori concept is fundamentally different from that of its empirical counterpart. Categories are concepts of a higher order than empirical concepts; like the ideas of space and time, they have to do with the form of experience rather than its matter. Our possession of categories accordingly supplies no knowledge of particular things; categories are fertile only when brought to bear on empirical data. Thus, because we hold to the a priori concept of cause, we interrogate nature in a certain way; thanks to it, we refuse to believe that there could be an uncaused event. But the answers we get to our interrogation depend primarily not on the form of our questions, but on what turns up in experience. Those who accuse Kant of having believed in the material a priori have failed to understand his theory.

To summarize this part of Kant's argument: If we are to have knowledge (and it is Kant's assumption that we do), various conditions must be fulfilled. The different items that fall within our experience must be capable of being connected in a single consciousness; there can be no happenings that are genuinely loose and separate. But the connections thus demanded must be objective connections—they must hold not just for my consciousness, but for "consciousness in general," for everyone's. An objective connection for Kant is a connection determined by a rule, and a rule is of its nature something that claims intersubjective validity. Finally, if we are to establish the operation of empirical rules we must proceed in accordance with nonempirical rules of a higher order, rules that ensure that our different experiences are capable of connection within a single experience.

JUDGMENTS. In view of the close relation Kant sees between the making of judgments and the use of a priori concepts, it is perhaps not surprising that he tries to arrive at a full list of such concepts by scrutinizing the formal properties of judgments. In this connection he invokes the doctrines of general or formal logic, a science he believed had been brought to completion at a single stroke by Aristotle. Few scholars have been convinced by this section of his argument, for it seems clear that Kant adapted the list of judgment forms to suit his list of cate-

gories, rather than deriving the categories from the judgment forms. In any case, it is not obvious how formal logic, which is a logic of consistency, can supply a clue to the content of what professes to be a logic of truth.

IMAGINATION AND UNDERSTANDING. In the first part of the "Analytic" Kant has much to say not only about concepts, judgments, and the understanding but also about the imagination. For example, he remarks in a cryptic passage:

> Synthesis in general is the mere result of the power of imagination, a blind but indispensable function in the soul, without which we should have no knowledge whatsoever, but of which we are scarcely ever conscious. To bring this synthesis to concepts is a function which belongs to understanding, and it is through this function of the understanding that we first obtain knowledge properly so called. (B 103)

The contrasting and, in places, overlapping roles of understanding and imagination are among the most puzzling features of Kant's exposition. The reason why they are both introduced is related to the fact that, in the second edition of the *Critique of Pure Reason* in particular, Kant was concerned with two quite distinct questions. He first asked himself what conditions have to be fulfilled if any sort of discursive consciousness is to have objective knowledge; he then went on to put the question as it relates to the human discursive consciousness, which not only intuits data passively, but does so under the particular forms of space and time. When the first question is uppermost Kant tends to speak of the understanding; when the second is to the fore, he brings in the imagination as well. The passage quoted above, typical of many, suggests that it is the business of the imagination to connect, whereas that of the understanding is to make explicit the principles on which the connecting proceeds. But in one chapter, "Schematism of the Pure Concepts of Understanding," a more satisfying account of the relationship is offered.

SCHEMATA. The problem of the chapter on what Kant called "schematism" is the central problem of the analytic: How can concepts that do not originate in experience find application in experience? At first Kant speaks as if there were no comparable difficulty in the case of concepts originating in experience, although he later makes clear that there are schemata corresponding both to empirical and to mathematical concepts. To possess the concept triangle is to know its formal definition, to be able to frame intelligible sentences containing the word *triangle*, and so on; to possess the schema corresponding to the concept triangle is to be able to envisage the variety of things to which the word *triangle* applies. Thus for Kant a schema is not an image, but a capacity to form images or (perhaps) to construct models. Pure concepts of the understanding are such that they "can never be brought into any image whatsoever" (B 181); the thought they embody, springing from the pure intellect, cannot be pictured or imagined. Yet there must be some connection between the abstract idea and the experienced world to which that idea is expected to apply; it must be possible to specify the empirical circumstances in which pure concepts of the understanding can find application. Kant thinks that for the categories this requirement is met by the fact that we can find for each of them a "transcendental schema," which is, he explains, a "transcendental determination of time." Without such a schema the categories would be devoid of "sense and significance," except in a logical (verbal) way. With it, use of the categories is clearly restricted to the range of things that fall within time—meaning, for Kant, restricted to phenomena.

The meaning of this baffling doctrine can perhaps best be grasped through Kant's examples of schemata:

> The schema of substance is permanence of the real in time, that is, the representation of the real as a substrate of empirical determination of time in general. … The schema of cause… is the real upon which, whenever posited, something else always follows. It consists, therefore, in the succession of the manifold, in so far as that succession is subject to a rule. … The schema of necessity is existence of an object at all times. (B 183–184)

It emerges from these cryptic sentences that the transcendental schema is something like an empirical counterpart of the pure category. It is what the latter means when translated into phenomenal terms. In Kant's own words, the schema is "properly, only the phenomenon, or sensible concept, of an object in agreement with the category" (B 186). A category without its corresponding "sensible concept" would be a bare abstraction, virtually without significance. Insofar as he argues that schematization is the work of the imagination, Kant has found a genuine function for the imagination to perform.

ANALYTIC OF PRINCIPLES: PURE PHYSICS. In the first half of the "Analytic" Kant undertook to produce a "transcendental deduction," that is, a general proof of validity, of the categories. In the second half of the "Analytic" he

gives a series of demonstrations of the synthetic a priori principles that rest on individual categories.

The categories are divided, for this and other purposes, into four groups: quantity, quality, relation, and modality. The four sets of corresponding principles are labeled axioms of intuition, anticipations of perception, analogies of experience, and postulates of empirical thought in general. Only one principle falls under each of the first two classes; the third contains a general principle and three more specific principles; the fourth contains three separate though closely connected principles. The first two classes are grouped together as "mathematical" principles; the third and fourth are described as "dynamical." Mathematical principles are said to be "immediately evident" and again to be "constitutive of their objects"; they apply directly to appearances. Dynamical principles are concerned with "the existence of such appearances and their relation to one another in respect of their existence." They are no less necessary than mathematical principles, but must be distinguished from them "in the nature of their evidence" and in that they are not "constitutive" but "regulative."

Behind this formidable façade some interesting ideas are hidden. In the first place, Kant makes stimulating though not altogether convincing remarks on the subject of proving principles of the understanding. The statement that every event has a cause carries strict necessity with it and therefore cannot be grounded on an inductive survey of empirical evidence. But equally it is not analytic, and so not open to straightforward conceptual proof. To be assured of its authenticity we consequently require a different type of argument altogether, which Kant calls a "transcendental" argument "from the possibility of experience." His idea is that only if the principles of the understanding are taken to be operative and in order can we have the type of experience we in fact have. Kant perhaps supposes that this type of proof is logically compulsive, but if so he overlooks the difficulty of setting up the original premise, of being sure that only if such-and-such were true should we have the experiences we have. But even with this defect his procedure has an immediate appeal, and is not without modern imitators.

AXIOMS OF INTUITION. The details of the particular arguments for the principles corresponding to the categories also deserve careful attention. The principle of axioms of intuition, that "all intuitions are extended magnitudes," is perhaps the most difficult to take seriously, since what it purports to prove has apparently already been dealt with in the "Aesthetic." Kant is once more asking questions about the application of mathematics to the world; in this section of the *Critique* the problem that apparently troubles him is how we know that inquiries about sizes or areas are always appropriate when we are dealing with things that occupy space. His solution is that they must be appropriate, since every such thing can be regarded as an aggregate of parts produced by the observer as he synthesizes his experiences. "I cannot represent to myself a line, however short, without drawing it in thought, that is, generating from a point all its parts one after another" (B 203).

ANTICIPATIONS OF PERCEPTION. Under the term "anticipations of perception" Kant is concerned with the question of the applicability of mathematics to sensations. What guarantee have we, he asks, that every sensation will turn out to have a determinate degree, in principle quantifiable? Might we not find, for instance, that an object is colored but with no precise depth of saturation, or a smell present in a room but with no specific magnitude? Kant attempts to rule out such possibilities by attention to the formal properties of sensations. We cannot anticipate the matter of sensation, but we can say in advance of experience that every sensation will have intensive magnitude, that is, a determinate degree, because it is possible to think of any given sensation as fading away until it is imperceptible, and conversely as being built up by continuous transitions on a scale from zero to the magnitude it has. Whatever may be the merits of this solution, there can be no doubt of the importance, and for that matter the novelty, of the question Kant asks here.

ANALOGIES OF EXPERIENCE. The section on the analogies of experience contains ideas as significant as any in Kant's writings.

The permanence of substance. The principle of the first analogy is that of the permanence of substance: "in all change of appearances substance is permanent; its quantum in nature is neither increased nor diminished." To believe in the permanence of substance is to believe that, whatever happens, nothing goes completely out of existence and nothing totally new is created: All change is transformation. Kant justifies the acceptance of this presupposition (which in his view, it should be remembered, applies only to things phenomenal) by arguing that without it we could not have a unitary temporal system. Coexistence and succession make sense only against a background that abides, and since time itself cannot be perceived, that background has got to be one of permanent things. This does not mean that we can determine a

priori what form the permanent will take; empirical scientists are to pronounce on that question, and their answers may obviously change from time to time. All that Kant seeks to rule out is the possibility that there might be no permanent at all. His argument is defective at a vital point here, but presumably he is saying that if things could go completely out of existence, so that it would make no sense to ask what became of them, the establishment of connections between one part of experience and another would be impossible. Experience would be (or at least might be) full of unbridgeable gaps, with the result that no one set of happenings could be integrated with another, and the unity of time would be totally destroyed.

Causation. Kant carries his argument further in his discussion of the second and third analogies, in which he argues for the necessary operation of the concepts of cause and reciprocity (causal interaction). But just as the notion of substance he justifies is very different from that held by metaphysicians, so is the Kantian concept of cause different from that of, say, Leibniz; it seems at first sight much closer to Hume's idea of a cause as an invariable antecedent. Causality for Kant as for Hume is a relation between successive events; a cause is an event that regularly precedes its effect. But whereas Hume is content to treat the occurrence of regular sequences as an ultimate and entirely contingent fact, Kant believes that without the presumption of sequences that are regular (determined by a rule) there could be no knowledge of objective succession. His reason is that we have to distinguish successions that happen only in ourselves, successions merely in our apprehension, from those that occur in the objective world and are independent of us. We can do this only if an objective sequence is defined as a sequence happening according to a rule. The objective world is a world of events the occurrence of each of which determines the precise place in time of some other event. But though events are necessarily connected in this way, we must not conclude that causal connections can be established a priori; for Kant as for Hume causal propositions are one and all synthetic and empirical. All we can know a priori is that there are such connections to be found, provided we have the skill or good fortune to discover them.

POSTULATES OF EMPIRICAL THOUGHT. One way of expressing Kant's attitude to substance and causality is to say that he thinks the principle of substance licenses us to ask the question, What became of that? Whenever something happens, and that the principle of causality licenses the parallel question, What brought that about? If someone tried to say that things might go out of existence altogether, or happen for no reason at all, Kant would say that these were logical but not real possibilities. The contrast between real and logical possibility is explored by Kant in the section "The Postulates of Empirical Thought." This section contains an explanation of the notions of possibility, actuality, and necessity from the critical point of view. By "really possible" Kant means "that which agrees with the formal conditions of experience, that is, with the conditions of intuition and of concepts" (B 265). A two-sided figure enclosing a space is not really possible, though its concept is not self-contradictory, because such a figure does not accord with the formal conditions of intuition. Telepathy and precognition are not real possibilities; they "cannot be based on experience and its known laws" (B 270), presumably because their actuality would violate some principle of the understanding, although Kant fails to make the point clear. The notion of real possibility is for Kant intermediate between logical and empirical possibility. We need it and can use it only because the world we have to deal with is a world that is not independently existent, but has its being in essential relation to consciousness.

PHENOMENA AND THINGS-IN-THEMSELVES. The distinction between phenomena and things-in-themselves, insisted on in the "Aesthetic" to explain our having a priori knowledge of the properties of space and time, is invoked again in the "Analytic" to account for "pure physics." If the world we confronted were one of things-in-themselves, a priori knowledge of it, even of the very restricted sort for which Kant argues, would be quite impossible. The fact that we have such knowledge—that we possess the principles discussed above—is taken by Kant as proof that the objects of our knowledge are phenomena or appearances. He does not mean by this, however, that they are private objects, at least insofar as they are spatial. The world we know in everyday and scientific experience is common to many observers; if not independent of consciousness as such, it is independent of particular consciousnesses. Parts of it are known only to particular experiencers—my inner life, for example, is accessible only to me—but that does not affect the general point.

Kant's acceptance of the distinction between phenomena and things-in-themselves has met with much criticism. Without the idea of the thing-in-itself, said his contemporary F. H. Jacobi, we cannot enter the world of the *Critique of Pure Reason*; with it we cannot remain inside. At the end of the "Analytic" Kant tries to defend himself against criticism of this sort by arguing that though he says that the objects of experience are phe-

nomena and is prepared to admit that the obverse of a phenomenon is a noumenon or intelligible object, he is committed to noumena only in a negative sense. Having said that the categories, one of which is existence, apply only to phenomena, he cannot with consistency hold any other view. Nor is his position at this stage as devoid of logic as some have tried to make out. After all, to describe things as phenomena he does not need to assert that there actually are things of a different kind; he needs only the idea of such things. To talk about things as they might be in themselves is no more objectionable than to speak of an *intellectus archetypus*, as Kant did in the letter to Herz, or of an intuitive understanding, as he constantly does in both the *Critique of Pure Reason* and the *Critique of Judgment.*

THE ELIMINATION OF DOGMATIC METAPHYSICS

At the end of the section of the *Critique of Pure Reason* devoted to the transcendental analytic, there is a passage that can be taken as summarizing the second stage in Kant's emancipation from Leibnizian rationalism:

> The Transcendental Analytic leads to this important conclusion, that the most the understanding can achieve *a priori* is to anticipate the form of a possible experience in general. And since that which is not appearance cannot be an object of experience, the understanding can never transcend those limits of sensibility within which alone objects can be given to us. Its principles are merely rules for the exposition of appearances; and the proud name of an Ontology that presumptuously claims to supply, in systematic doctrinal form, synthetic *a priori* knowledge of things in general … must, therefore, give place to the modest title of a mere Analytic of pure understanding. (B 303)

Kant thus repudiates the possibility of knowledge through pure concepts of things as they really are; in 1770 he had still clung to it. Having disposed of ontology, Kant needed to consider, to complete the negative side of his work, the tenability of the remaining parts of metaphysics (rational psychology, rational cosmology, and natural theology in Baumgarten's classification), and this he did in the section titled "Transcendental Dialectic." To complete his own alternative to rationalism he needed to clarify the status of the propositions involved in "pure practical faith." His attempt to meet this requirement is made at the very end of the *Critique*, especially in the chapter "The Canon of Pure Reason" (B 823ff.).

REASON. Most of the conclusions of the "Dialectic" follow directly from those of the "Analytic," though there are new points of interest. As in the "Analytic," Kant's views are expressed inside a framework that is heavily scholastic. Kant claimed that human beings have an intellectual faculty in addition to the understanding. This additional faculty is reason, and it is equipped with a set of a priori concepts of its own, technically known as ideas of reason. An idea of reason can have no object corresponding to it in sense experience, for the ambition of reason is to arrive at absolute totality in the series of conditions for the empirically given, and in this way to grasp the unconditioned that falls outside experience altogether. However, this ambition can never be realized, and the only proper function for reason in its theoretical capacity is to regulate the operations of the understanding by encouraging it to pursue the search for conditions to the maximum extent that is empirically possible.

THE KNOWING SUBJECT. Kant's handling of the "psychological idea" at the beginning of the main part of the "Dialectic" is exceptionally brilliant. He maintains in the "Analytic" that what he there calls the "I think," or the unity of apperception, is the ultimate condition of experience, in the sense of being the logical subject of experience or the point to which all experience relates. All experience is experience for a subject; whatever thoughts or feelings I have I must be capable of recognizing as my thoughts or feelings. But the subject here referred to is not something substantial; it is merely a logical requirement, in that nothing follows about the nature of my soul or self from the fact that I say "I think." So far from being "an abiding and continuing intuition" (the sort of thing Hume vainly sought in the flow of his inner consciousness), for Kant the "representation 'I' … [is] simple, and in itself completely empty … we cannot even say that this is a concept, but only that it is a bare consciousness which accompanies all concepts. Through this I or he or it (the thing) which thinks, nothing further is represented than a transcendental subject of thoughts = *X*" (B 404). The same view is expressed in an earlier passage in the *Critique*, where Kant says that "in the synthetic original unity of apperception, I am conscious of, myself, not as I appear to myself, nor as I am in myself, but [I am conscious] only that I am. This *representation* is a *thought*, not an *intuition*" (B 157).

REFUTATION OF RATIONAL PSYCHOLOGY. These subtleties are unknown to the exponents of rational psychology, who develop the whole of their teaching around a "single text," which is "I think." From the fact that I am

the subject of all my thoughts they infer that I am a thinking substance; from the fact that the "I" of apperception is logically simple they conclude that I am, in substance, simple and not composite. The proposition that "in all the manifold of which I am conscious I am identical with myself" is taken by them as implying that I am possessed of continuing personal identity. Finally, my distinguishing my own existence as a thinking being from that of other things, including my own body, is put forward as proof that I am really distinct from such things and so could in principle exist in complete independence of them. None of these inferences is justified, for in each case a move is attempted from an analytically true premise to a synthetic conclusion. As Kant remarks, "it would, indeed, be surprising if what in other cases requires so much labour to determine—namely, what, of all that is presented in intuition, is substance, and further, whether this substance can be simple …—should be thus given me directly, as if by revelation, in the poorest of all representations" (B 408).

MIND AND BODY. Kant presents the doctrines of rational psychology in his own idiosyncratic way, but anyone who reflects on the theories of Descartes will see that Kant was by no means attacking men of straw. Kant's treatment of the fourth paralogism, "of Ideality," is of special interest in this connection. Descartes inferred from his *cogito* argument that mind and body were separate in substance, which meant that the first could exist apart from the second. Bound up with this was the view that I am immediately aware of myself as a mind, but need to infer the existence of material things, which is in principle open to doubt. A great many philosophers have subscribed to this opinion, but Kant thought he could show it to be definitively false. In order to say that my inner experiences come one before another I need to observe them against a permanent background, and this can only be a background of external objects, for there is nothing permanent in the flow of inner experience. As Kant put it in the second edition, in which he transposed the argument to the discussion of existence in connection with the postulates of empirical thought), "*The mere, but empirically determined, consciousness of my own existence proves the existence of objects in space outside me*" (B 275). Kant is in no sense a behaviorist; he thinks that empirical self-knowledge is to be achieved through inner sense and declares in one passage that, for empirical purposes, dualism of soul and body must be taken as correct. Yet his commitment to "empirical realism" is quite unambiguous.

THE ANTINOMIES. Of the remaining parts of the "Dialectic," only the sections on the antinomies and on the existence of God can be discussed here. In the "Antinomy of Pure Reason," Kant first sets out a series of pairs of metaphysical doctrines (which he says have to do with cosmology but which are in fact of wider interest). The two doctrines in each pair seem to contradict one another directly. He then produces for each pair what he regards as watertight proofs of both sides of the case, maintaining that if we adopt the dogmatic standpoint assumed without question by the parties to the dispute, we can prove, for example, both that the world has a beginning in time and that it has no beginning in time, both that "causality in accordance with laws of nature is not the only causality" and that "everything in the world takes place solely in accordance with laws of nature." Thus Kant exhibits in systematic form the famous contradictions into which, as he notes, reason precipitates itself when it asks metaphysical questions. Kant is enormously impressed by the discovery of these contradictions, and it is regrettable only that he does not sufficiently discuss their formal character or illustrate them with genuine examples.

The only way to avoid these antinomies, in Kant's opinion, is to adopt his own (critical) point of view and recognize that the world that is the object of our knowledge is a world of appearances, existing only insofar as it is constructed; this solution enables us to dismiss both parties to the dispute in the case of the first two antinomies, and to accept the contentions of both parties in the case of the other two. If the world exists only insofar as it is constructed, it is neither finite nor infinite but indefinitely extensible and so neither has nor lacks a limit in space and time. Equally, if the world is phenomenal we have at least the idea of a world that is not phenomenal; and natural causality can apply without restriction to the first without precluding the application of a different type of causality to the second. This is admittedly only an empty hypothesis so far as theoretical reason is concerned, but Kant argues that it can be converted into something more satisfactory if we take account of the activities of practical (moral) reason.

THE EXISTENCE OF GOD. The fourth antinomy is concerned with God's existence. Kant's full treatment of the subject is not in the section on the antinomies but in that headed "The Ideal of Pure Reason," the locus classicus for Kant's criticisms of speculative theology. These criticisms have proved as devastating as those he brought against rational psychology.

Speculative proofs. There are, Kant argues, only three ways of proving God's existence on the speculative plane. First, we can proceed entirely a priori and maintain that the very idea of God is such that God could not *not* exist; this is the method of the Ontological Argument. Second, we can move from the bare fact that the world exists to the position that God is its ultimate cause, as in the First Cause, or Cosmological, Argument. Finally, we can base our contention on the particular constitution of the world, as in the "physicotheological proof" (the Argument from Design).

Kant argues that all three types of proof are fallacious. The Ontological Argument fails because it treats existence as if it were a "real predicate," whereas "it is not a concept of something which could be added to the concept of a thing. It is merely the positing of a thing, or of certain determinations, as existing in themselves" (B 626). The First Cause Argument fails on several counts: because it uses the category of cause without realizing that only in the schematized form is the category significant; because it assumes that the only way to avoid an actually infinite causal series in the world is to posit a first cause; finally and most important, because it presupposes the validity of the Ontological Proof, in the step which identifies the "necessary being" or First Cause with God. The Argument from Design makes all these mistakes and some of its own, for even on its own terms it proves only the existence of an architect of the universe, not of a creator, and such an architect would possess remarkable but not infinite powers.

The moral proof. In spite of Kant's criticisms of the classical arguments for God's existence, he is neither an atheist nor even a believer in the principle of *credo quia impossibile*. He both believes in God and holds that the belief can be rationally justified. For although speculative theology is, broadly, a tissue of errors, moral theology is perfectly possible. But the moral proof of God's existence differs from the attempted speculative proofs in at least two significant respects. First, it begins neither from a concept nor from a fact about the world, but from an immediately experienced moral situation. The moral agent feels called upon to achieve certain results, in particular to bring about a state of affairs in which happiness is proportioned to virtue, and knows that he cannot do it by his own unaided efforts; insofar as he commits himself to action he shows his belief in a moral author of the universe. Affirmation of God's existence is intimately linked with practice; it is most definitely not the result of mere speculation. Again, a proof like the First Cause Argument claims universal validity; standing as it does on purely intellectual grounds it ought, if cogent, to persuade saint and sinner alike. But the moral proof as Kant states it would not even have meaning to a man who is unconscious of moral obligations; the very word *God*, removed from the moral context that gives it life, is almost or quite without significance. Accordingly Kant states that the result of this proof is not objective knowledge but a species of personal conviction, embodying not logical but moral certainty. He adds that "I must not even say '*It is* morally certain that there is a God …,' but '*I am* morally certain'" (B 857). In other words, the belief or faith Kant proposes as a replacement for discredited metaphysical knowledge can be neither strictly communicated nor learned from another. It is something that has to be achieved by every man for himself.

ETHICS

Kant perhaps intended originally to make the *Critique of Pure Reason* the vehicle of his entire philosophy, but it was clear before he completed it that some of his views, especially those on ethics, could be only touched on there. In the years immediately following its publication he displayed exceptional energy in defending and restating the theories he had already put forth and in extending his philosophy to cover topics he had hitherto not treated, or not treated in detail. By 1788 he had not only published the second, substantially revised edition of the *Critique of Pure Reason*, but had laid the foundations for his ethics in his short but influential *Groundwork of the Metaphysic of Morals* (1785) and had undertaken a more elaborate survey of moral concepts and assumptions in the *Critique of Practical Reason* (1788). He had also, in passing, written his essay *Metaphysical Foundations of Natural Science* (1786), intended as a first step toward a projected but never completed metaphysics of nature. Two years after the *Critique of Practical Reason* he produced yet another substantial work, the *Critique of Judgment*, in which he expressed his views on, among other topics, aesthetics and teleology.

MORAL ACTIONS. If he had published nothing else but the *Groundwork of the Metaphysic of Morals* Kant would be assured a place in the history of philosophy. Difficult as it is to interpret in some of its details, this work is written with an eloquence, depth of insight, and strength of feeling that make an immediate impact on the reader and put it among the classics of the subject. Kant says that his "sole aim" in the book is "to seek out and establish *the supreme principle of morality*." He wishes to delineate the basic features of the situation in which moral decisions

are made, and so to clarify the special character of such decisions.

The situation as he sees it is roughly as follows. Man is a creature who is half sensual, half rational. Sensuous impulses are the determining factor in many of his actions, and the role of reason in these cases is that assigned to it by Hume; it is the slave or servant of the passions. But there is an identifiable class of actions in which reason plays a different part, leading rather than following. This is the class of moral actions. Such actions have the distinguishing feature that they are undertaken not for some ulterior end, but simply because of the principle they embody.

INTENTIONS AND MORAL JUDGMENTS. The moral worth of an action, as Kant puts it (*Grundlegung*, 2nd ed., p. 13), lies "not in the purpose to be attained by it, but in the maxim in accordance with which it is decided upon." Whether or not I attain my ends does not depend on me alone, and my actions cannot be pronounced good or bad according to the effects they actually bring about. But I can be praised or blamed for my intentions, and I can, if I choose, make sure that the maxim or subjective principle of my action accords with the requirements of morality. To do this I have only to ask myself the simple question whether I could will that the maxim should become a universal law, governing not merely this particular action of mine, but the actions of all agents similarly circumstanced. For it is a formal property of moral as of scientific judgments, recognized in practice even by the unsophisticated, that they hold without distinction of persons; the result is that an action can be permissible for me only if it is permissible for anyone in my situation.

PRACTICAL REASON. There are difficulties in this position of which Kant seems to have been unaware. In particular, he never asks how I am to decide what is the correct description, and hence the maxim, of my act or proposed act. Nor is it obvious how the theory shows the falsity of Hume's view that "reason alone can never be a motive to any action of the will"—how it can be shown, in Kant's language, that pure reason really is practical. The practical effectiveness of reason is manifested not in the capacity to reflect, which both Kant and Hume allow, but in the power to originate or inhibit action. Kant obviously thinks that the facts of temptation and resistance to temptation, which he sees as ubiquitous in the moral life, have a clear bearing on the question whether reason really has such a power. Recognition that I ought to follow a certain course of action, whether I want to or not, and that anything that is morally obligatory must also be practically possible, is enough in his view to show that I am not necessarily at the mercy of my desires. In favorable cases, at any rate (Kant pays too little attention to the factors that diminish and sometimes demolish responsibility), I am free to resist my sensuous impulses and to determine my actions by rational considerations alone.

CONSEQUENCES OF THE MORAL LAW. Some commentators have seen Kant as an ethical intuitionist, but this view is clearly mistaken. His "practical reason" is not the faculty of insight into the content of the moral law; it is rather the capacity to act. In determining what the moral law commands, I have initially no other resources at my disposal than the reflection that it must be applied impartially. But in practice this criterion carries others with it. If the moral law applies without distinction of persons, Kant believes it follows that I must treat all human beings as equally entitled to rights under it, and that therefore I must regard them as ends in themselves and never as merely means to my own ends. Further, once I recognize that other people are morally in the same position as I am myself, and that we belong to the same moral community, I recognize both that I can legitimately pursue those of my purposes that do not conflict with the moral law and that I also have a duty to facilitate the like pursuit on the part of my fellows. So though Kant is a formalist in his view of moral reason (as in his view of the theoretical intellect), he sees his ethics as having practical consequences of the first importance. He sets these consequences out in his lectures on ethics and develops them in detail later in his 1797 *Metaphysic of Morals*. To judge him by the *Groundwork* alone, or even by the *Groundwork* and the *Critique of Practical Reason* taken together, is to do less than justice to the scope of his ethical reflection.

MORAL IMPERATIVES. Previous moral philosophies, Kant writes, whether they put their stress on moral sense or on moral reason, have all been vitiated by a failure to recognize the principle of the autonomy of the will. Utilitarianism, for instance, is a heteronomous ethical theory because, according to its supporters, the point of a moral action is to promote an end or purpose beyond the action, the greatest happiness of the greatest number. Kant is not unaware of the importance of ends and purposes in actions: In the *Critique of Practical Reason* he corrects the one-sidedness of the *Groundwork* by discoursing at length on the concept of "good" as well as on that of "duty." But he holds, even so, that consideration of ends cannot be of primary importance for the moral agent, since a moral action is one that is commanded for its own sake, not with a view to some purpose it is

expected to bring about. The imperatives of morality command categorically, unlike those of skill or prudence, which have only hypothetical force. A rule of skill or a counsel of prudence bids us take certain steps if we wish to attain a certain end—good health or overall happiness, for example. There is no "if" about a command of morality; it bids me act in a certain way whether I want to or not, and without regard to any result the action may bring about. It represents a course of conduct as unconditionally necessary, not just necessary because it conduces to a certain end.

FREEDOM AND NECESSITY. The concepts of duty, the categorical imperative, the moral law, and the realm of ends (in which we are all at once subjects and lawgivers) are intended by Kant to illuminate the moral situation. But even when we know what that situation is, there are many features of it that remain mysterious. Morality as Kant expounds it involves autonomy of the will, and such autonomy clearly makes no sense except on the supposition of freedom. But how we can think of the will as free and at the same time regard ourselves as subject to the moral law, that is, as under obligation, has still to be explained. To throw light on this question, Kant invokes the concept of the two worlds, the sensible and the intelligible, to which he made appeal in the *Critique of Pure Reason*. Insofar as I exercise the faculty of reason I have to regard myself as belonging to the intelligible world; insofar as I exercise my "lower" faculties I am part of the world of nature, which is known through the senses. Were I a purely rational being, possessed of what Kant sometimes calls a "holy will," all my actions would be in perfect conformity with the principle of autonomy, and the notions of obligation and the moral law would have no meaning for me. They would similarly have no meaning if I were a purely sensuous being, for then everything I did would occur according to natural necessity, and there would be no sense in thinking that things ought to be otherwise. The peculiarities of the human moral situation arise from the fact that men are, or rather must think of themselves as being, at once intelligible and sensible. Because I regard myself as belonging to the intelligible order, I see myself as "under laws which, being independent of nature, are not empirical but have their ground in reason alone" (*Critique of Practical Reason,* p. 109). But I am also a natural being, and those laws therefore present themselves to me in the form of commands that I acknowledge as absolute because I recognize that the intelligible world is the ground of the sensible. We can thus see "how a categorical imperative is possible."

What we cannot see, if Kant is to be believed, is how freedom is possible. "All men think of themselves as having a free will. … Moreover, for *purposes of action* the footpath of freedom is the only one on which we can make use of reason in our conduct. Hence to argue freedom away is as impossible for the most abstruse philosophy as it is for the most ordinary human reason" (*Critique of Practical Reason,* p. 113–115). Yet freedom remains what it is in the *Critique of Pure Reason*, "only an idea whose objective reality is in itself questionable," and there is a prima facie clash between the claim to freedom and the knowledge that everything in nature is determined by natural necessity. Kant seeks to dissolve the antinomy of freedom and necessity by means of two expedients. First, he insists that the idea of freedom required for morals is not a theoretical but a practical idea. Freedom does not need to be established as a metaphysical fact; it is enough that we find it necessary to act on the assumption that freedom is real, since "every being who cannot act except under the idea of freedom is by this alone—from the practical point of view—really free" (p. 100). The status of the proposition that the will is free is identical with that of the proposition that there is a God. Both are postulates of practical reason—beliefs that we "inevitably" accept; but they are emphatically not items of knowledge in the strict sense of that term. Second, Kant sees no difficulty in our accepting the postulate of freedom, because there is no contradiction in thinking of the will as free. As an object of theoretical scrutiny I must regard myself as a phenomenon; as a moral agent possessed of a will I transfer myself to the intelligible world of noumena. I can be at once under necessity qua phenomenon and free qua noumenon. But the question of how I can be free leads to the extreme limits of practical philosophy. Freedom cannot be explained, for we lack all insight into the intelligible world; the most we can do is make clear why it cannot be explained. The critical philosophy purports to have performed this task.

EPISTEMOLOGY AND ETHICS. Kant advocates a form of nonnaturalist theory in ethics. But neither his ethics nor his theory of knowledge can be fully understood in isolation one from the other. The two together constitute an overall theory that is not so much a metaphysics as a substitute for a metaphysics: A theory that argues that human insight is strictly limited, but urges that, so far from being regrettable, this testifies to "the wise adaptation of man's cognitive faculties to his practical vocation" (*Critique of Practical Reason and Other Writings*, Beck translation, 1949, p. 247). If we knew more, we might indeed do as we ought, for "God and eternity in their

awful majesty would stand unceasingly before our eyes," but we should not then do things as a matter of duty, but rather out of fear or hope. And thus the world would be poorer, for we should lose the opportunity to manifest "good will," the only thing in the world, "or even out of it, which can be taken as good without qualification."

THE CRITIQUE OF JUDGMENT

None of Kant's other writings is as forceful or original as the first two *Critiques* and the *Groundwork*. The *Critique of Judgment* contains some fresh ideas of remarkable power, but it constitutes a series of appendixes or addenda to Kant's earlier work rather than something wholly new. It should really be seen as three or four separate essays whose connecting link is the concept of purpose.

SYSTEM OF SCIENCE. The first essay, the introduction, begins with a pedantic discussion of the status of the power of judgment. It then takes up a problem aired in the appendix to the "Dialectic" in the *Critique of Pure Reason*—the problem of the special assumptions involved in the belief that we can construct a system of scientific laws. If we are to have such a system, Kant argues, we must proceed on the principle that nature is "formally purposive" in respect of empirical laws; that nature is such that we can make sense of it not merely in general, but also in detail. Kant's object is to show that this principle is not a constitutive principle of things, but simply a subjective maxim of judgment.

In the *Critique of Pure Reason* (B 670ff.) Kant argues for what he calls the regulative employment of the ideas of reason: the use of ideas to order empirical inquiries in such a way that we try at once to find greater and greater diversity of form in the material before us and to group different species and subspecies together under ever higher genera. In actual practice we assume that nature will display the unity-in-diversity required for this program to be carried out, but we cannot prove that it will do so as we can prove that whatever falls within experience will conform to the categories. Hence we are concerned not with objective rules, but only with maxims, defined in this connection as "subjective principles which are derived, not from the constitution of an object but from the interest of reason in respect of a certain possible perfection of the knowledge of the object" (B 694).

In the *Critique of Pure Reason* Kant ascribes these maxims to reason. In the *Critique of Judgment*, he assigns them to judgment, in effect the identical doctrine. The difference is accounted for by two facts. First, by the time Kant wrote the *Critique of Judgment*, the term *reason* suggested to him nothing but practical reason. Second, he had come to think that if the power of judgment is genuinely separate from understanding on the one hand and reason on the other it must have a priori principles of its own. A division within the power of judgment itself, into determinant and reflective activities, had helped to make this last point plausible, at least in the eyes of its author.

AESTHETICS. The "Critique of Aesthetic Judgment," the first major division of the *Critique of Judgment*, uses the term *aesthetic* in what has become its modern sense. The discussion is Kant's contribution to the controversies initiated by Lord Shaftesbury and Francis Hutcheson when they made both moral and aesthetic judgments matters of feeling; Kant rejects this view and also explains why he yet cannot approve of Baumgarten's attempt to "bring the critical treatment of the beautiful under rational principles, and so to raise its rules to the rank of a science" (B 35, note *a*). Kant needs to show, for the purposes of his general philosophy, that aesthetic judgments are essentially different from moral judgments on the one hand and scientific judgments on the other. This need apart, he had a long-standing independent interest in the subject; in 1764, thirty years before the *Critique of Judgment*, he published an essay on the beautiful and the sublime (*Beobachtung über das Gefühl des Schönen und Erhabenen*, Königsberg). Such an interest may seem surprising in view of the obvious limitations of Kant's own aesthetic experience; he had some feeling for literature, especially for satire, but little or no real knowledge of either painting or music. But what he has in mind in discussing the beautiful is the beauty of nature as much as anything, and his main interest is not in making aesthetic judgments, but in deciding on their logical status.

Judgments of taste, as Kant calls them, are peculiar in that they not only rest on feeling but also claim universal validity. That they rest on feeling seems to him obvious: When I ascribe beauty to an object or scene I do so not because I have observed some special character in it, but because contemplation of its form gives me immediate delight. But it is an entirely disinterested form of delight, quite different from that we feel concerning things that are agreeable, or even things that are good. When we take pleasure in something beautiful we are not desiring to possess it, or indeed taking up any attitude toward its existence. The fact that aesthetic delight is disinterested allows us to think of it as universally shared:

> Since the delight is not based on any inclination of the subject (or any other deliberate interest), but the Subject feels himself completely *free* in

> respect to the liking which he accords to the object, he can find as reason for his delight no personal conditions to which his own subjective self might alone be party. Hence he must regard it as resting on what he may also presuppose in every other person; and therefore he must believe that he has reason for demanding a similar delight from every one. (*Critique of Judgment*, Meredith translation, Sec. 6)

Because they claim universal validity, judgments of taste appear to rest on concepts, but to think that they do is a mistake. The universality attaching to judgments of taste is not objective but subjective; to explain it we must refer to "nothing else than the mental state present in the free play of imagination and understanding (so far as these are in mutual accord, as is requisite for *cognition in general*)" (Sec. 9). As in the *Critique of Pure Reason*, Kant argues that both imagination and understanding are involved in the apprehension of any spatiotemporal object but that when we simply contemplate any such object aesthetically, no definite concept is adduced; and so the two faculties are in free play. It is the harmony between the faculties in any act of aesthetic contemplation that Kant takes to be universally communicable, and believes to be the basis for the pleasure we feel.

In addition to analyzing judgments about the beautiful, Kant devoted considerable attention in the *Critique of Judgment* to another concept which figured prominently in the aesthetics of his day, that of the sublime. Burke and others had given what was in effect a psychological description of the conditions in which we judge, say, the sight of a mountain range or a storm at sea to be sublime. Kant was all the more anxious to specify more exactly the meaning of such judgments and to establish their transcendental conditions because he was convinced that we here also have to do with a feeling that is held to be universally communicable. The feeling for the sublime, as he explained it, is connected not with the understanding, as is that for the beautiful, but with reason. To put his view somewhat crudely, we are at first abashed by the formlessness of some parts of nature, only to be elevated when we reflect on the utter inadequacy of these objects to measure up to our own ideas, and in particular to our moral ideas. Thus the sublime is not, as might at first sight be supposed, a quality which inheres in natural objects, but a feeling which the contemplation of natural objects provokes in us. It could have no existence for a being totally lacking in culture (a savage might feel fear on observing "thunderclouds piled up the vault of heaven," to use one of Kant's own examples, but could not recognize their sublimity), yet it is not a mere product of culture or social convention. "Rather is it in human nature that its foundations are laid, and, in fact, in that which, at once with common understanding, we may expect everyone to possess and may require of him, namely, a native capacity for the feeling for (practical) ideas, that is, for moral feeling" (Sec. 29).

TELEOLOGY. One of Kant's motives for wanting to avoid making beauty an objective characteristic was that he thought such a view would lend force to the Argument from Design, and so encourage the revival of speculative theology. If things could be said to possess beauty in the same sort of way in which they possess weight, it would be a short step to talking about the Great Artificer who made them to delight us. Arguments of the same general kind were still more vividly present to his mind when he came to write the second main section of the *Critique of Judgment*, the "Critique of Teleological Judgment." Indeed, he ended the book with a lengthy section that underlines yet again the shortcomings of "physicotheology" and points up the merits of "ethicotheology."

Before confronting theology directly, Kant embarked on a detailed and penetrating discussion of the nature and use of teleological concepts. The existence of organic bodies, he argues, is something for which we cannot account satisfactorily by the mechanical principles sanctioned by the physical sciences; to deal with organic bodies we must employ a distinct principle, the principle of teleology, which can do justice to the fact that "*an organized natural product is one in which every part is reciprocally both means and end*" (Sec. 66). Such a principle cannot be used for cognitive purposes in the strict sense; it can be employed only by reflective judgment to guide "our investigation of … [organic bodies] by a remote analogy with our own causality according to ends generally, and as a basis for reflection upon their supreme source" (Sec. 65). Teleology is a concept that occupies an uneasy intermediate position between natural science and theology. We cannot help using it to describe the world about us, yet we cannot assign to it full scientific status. Kant mitigates the austerities of this position by suggesting in his section "The Antinomy of Judgment" that in the end the mechanical and teleological principles stand on the same level, both belonging to reflective judgment. But it is hard to see how this can be made consistent with the doctrines of the *Critique of Pure Reason*, which ascribes constitutive force to the concepts of "pure physics," or even with the distinction in the *Critique of Judgment* itself between explaining something and merely "making an estimate" of it. We use the categories to

explain, but can employ teleological concepts only for the purpose of making an estimate. Kant's underlying attitude to the whole question is revealed most clearly in the passage at the end of Sec. 68 of the *Critique of Judgment*, where he asks why teleology "does not … form a special part of theoretical natural science, but is relegated to theology by way of a propaedeutic or transition." He answers:

> This is done in order to keep the study of the mechanical aspect of nature in close adherence to what we are able so to subject to our observation or experiment that we could ourselves produce it like nature, or at least produce it according to similar laws. For we have complete insight only into what we can make and accomplish according to our conceptions. But to effect by means of art a presentation similar to organization, as an intrinsic end of nature, infinitely surpasses all our powers. (Meredith translation)

It would be interesting to know if Kant would say the same were he alive today.

OTHER PHILOSOPHICAL WRITINGS

After publishing the three *Critiques*—Kant was sixty-six when the *Critique of Judgment* appeared—he continued to publish essays and treatises on a wide variety of philosophical subjects. Most of these are in fact contributions to applied philosophy, for he took the view that scientific inquiries and practical activities alike stand in need of philosophical foundations. In many cases he attempts to supply these foundations by means of the principles established in his main works—hence the general shape of his philosophies of science and religion, and of his political philosophy. It would, however, be wrong to see these as no more than mechanical applications of general Kantian conclusions. For although Kant was deeply and indeed unduly devoted to system, he also had a wide and in some cases penetrating knowledge of many different branches of learning and human activity, and there are few philosophical topics that he touches without illuminating; in fact, Kant gave the names still in use to most of the branches of applied philosophy he took up.

PHILOSOPHY OF NATURE. In the preface to his *Metaphysical Foundations of Natural Science*, Kant argues that the very concept of scientific knowledge is such that we can use the term properly only when dealing with truths that are both apodictically certain and systematically connected. A discipline that is thoroughly and entirely empirical cannot comply with these requirements; hence Kant pronounces chemistry to be no better than "systematic art or experimental doctrine." But the situation is different in physics. Although Kant was as firmly persuaded as any empiricist that detailed knowledge of the physical world could be arrived at only by observation and experiment, he was also sure that physics has an unshakable a priori basis that makes it worthy of the name of science. It owes this, in Kant's judgment, to the fact that its fundamental concepts are capable of mathematical expression, as those of chemistry are not, and to the close connection of these concepts with the categories, the basic concepts of rational thought.

The main object of the *Metaphysical Foundations* is to demonstrate the second of these points by means of an examination of the idea of matter. Starting from what professes to be an empirically derived definition of matter, "that which is capable of movement in space," Kant proceeds to a deduction of its main properties in the light of the table of categories. The result is, in effect, a rereading or reinterpretation of then-current physical theory in which all the main doctrines of Newton find their place, but which is distinctive in that the atomism professed by many physicists of the day is rejected in favor of a dynamical theory of matter resembling that of Leibniz. Kant argues in the *Critique of Pure Reason* that only mistaken metaphysics leads scientists to think they must accept the notions of absolutely homogeneous matter and absolutely empty space. In the *Metaphysical Foundations* he works out an alternative conception of matter in terms of moving forces, omnipresent but varying in degree, and puts it forward as both theoretically satisfactory and consistent with the empirical findings.

It is difficult not to see in these views the beginnings of *Naturphilosophie* as it was to be practiced by Schelling and G. W. F. Hegel, the more so if we read the *Metaphysical Foundations* in the light of Kant's further treatment of the subject in the notes published as *Opus Postumum*. But in 1786 at any rate Kant was still far from committing the extravagances of the speculative philosophers of nature. For one thing, he was both more knowledgeable about and more respectful of the actual achievements of physical scientists than were his romantic successors, doubtless because, unlike them, he was something of a physical scientist himself. For another, the lesson he drew from his 1786 inquiries was not how much physical knowledge we can arrive at by the use of pure reason, but how little. To establish the metaphysical foundations of natural science was a useful task, but it was in no sense a substitute for empirical investigation. Despite these differences from *Naturphilosophie*, it must be allowed that *Metaphysical*

Foundations testifies, in name as well as in content, to the extent of Kant's commitment to rationalism (his theory of science could scarcely be further from Hume's) and to the way in which he was at least tempted by the constructivism favored by some of his younger contemporaries.

PHILOSOPHY OF HISTORY. Although Kant was quite unaware of the problems about historical knowledge and explanation with which philosophers since Wilhelm Dilthey have dealt, he made an important and characteristic contribution to speculative philosophy of history in his essay "Idee zu einer allgemeinen Geschichte in Weltbürgerlicher Absicht" (Idea of a Universal History from a Cosmopolitan Point of View; *Berliner Monatsschrift*, November 1784, 386–410). Observing that the actions of men, when looked at individually, add up to nothing significant, he suggests that nature or providence may be pursuing through these actions a long-term plan of which the agents are unaware. To see what the plan may be we have to reflect on two points: First, that nature would scarcely have implanted capacities in human beings if she had not meant them to be developed, and second, that many human intellectual capacities (for example, the talent for invention) are such that they cannot be satisfactorily developed in the lifetime of a single individual.

The development of such capacities belongs to the history of the species as a whole. Kant suggests that the hidden plan of nature in history may well be to provide conditions in which such capacities are more and more developed, so that men move from barbarism to culture and thus convert "a social union originating in pathological needs into a moral whole." The mechanism of the process lies in what Kant calls the "unsocial sociability" of human beings—the fact that they need each other's society and help and are nevertheless by nature individualists and egotists—which ensures that men develop their talents to the maximum extent, if only to get the better of their fellows, and at the same time necessitates man's eventually arriving at a form of civil society that allows for peaceful rivalry under a strict rule of law. But such a "republican" constitution would be of no value unless it had its counterpart in the international sphere, for the struggles of individuals against one another are paralleled by the struggles of states. We must accordingly conclude that the final purpose of nature in history is to produce an international society consisting of a league of nations, in which war is outlawed and the way is finally clear for peaceful competition between individuals and nations.

The difficulty with this as with other lines of Kant's thought is to understand its relation to empirical inquiries. From what Kant says it seems clear that he intended "philosophical" history to be an alternative to history of the everyday kind, not a substitute for it. Nor did he pretend to be writing philosophical history himself; his essay merely puts forward the idea of or offers a "clue" to, such a history, leaving it to nature to produce someone really capable of making sense of the historical facts as Johannes Kepler and Newton made sense of physical facts. It is difficult to see, even so, how Kant could have possessed the idea of history as meaningful without knowing the facts, or alternatively how he could know that the idea throws light on the facts when it was discovered without any reference to them.

PHILOSOPHY OF LAW AND POLITICS. Kant's views about law and politics, like his philosophy of history, are obviously tied up with his ethics. Kant holds that legal obligations are a subspecies of moral obligation; thus the rational will, and neither force nor the commands of God, is the basis of the law. His standpoint in philosophy of law is thus broadly liberal, though his attitude on many particular legal issues is far from liberal as the term is now understood. He holds, for instance, that if one of the partners to a marriage runs away or takes another partner, "the other is entitled, at any time, and incontestably, to bring such a one back to the former relation, as if that person were a thing" (*Metaphysic of Morals*, Sec. 25). He is notorious as a strong supporter of the retributive theory of punishment and an uncompromising advocate of the death penalty for murder. The explanation of his harshness in these matters is to be found in his legalistic approach to ethics, which leaves little room for sympathy or forgiveness.

In politics also Kant combines a fundamentally liberal attitude with specific views that are conservative, if not reactionary. Following Rousseau, he attempts to explain political authority partly in terms of the general will and partly in terms of the original contract. Insofar as he insists on the contract, which he interprets not as a historical fact but as a regulative idea, he is advocating a version of political liberalism which lays particular emphasis on the rule of law; insofar as he grounds supreme political authority in the will of the people as a whole, he is obviously flirting with more radical doctrines—from whose consequences he is quick to draw back. An admirer of the French Revolution, he nevertheless denies that the subjects of the most ill-governed states have any right of rebellion against their rulers. And though the mixed constitution he favors is one in which citizens can make their

voices heard through their representatives, he is for confining the franchise to persons who possess "independence or self-sufficiency," thus excluding from "active" citizenship (according to Sec. 46 of the *Metaphysic of Morals*) apprentices, servants, woodcutters, plowmen, and, surprisingly, resident tutors, as well as "all women." The truth is, however, that Kant's political theorizing was done in a vacuum; in his day there was no real chance for a Prussian professor of philosophy to influence political events.

PHILOSOPHY OF RELIGION. In the sphere of religion the views of a professor of philosophy could be influential, and Kant's views on this subject were certainly provocative. He treats religion as essentially, if not quite exclusively, a matter of purity of heart—thus dispensing with speculative theology altogether and assigning a meager importance to the institutional side of religion. To adopt the religious attitude, as Kant sees it, is to look on duties as if they were divine commands. But this, he explains, is only to insist on the unconditioned character, the ineluctability, of moral obligation; it is a way of representing morality, not a way of going beyond it. Knowledge of the supersensible, as Kant thought he had shown in the *Critique of Pure Reason*, is impossible; and although moral practice carries with it belief in God and a future life, the whole meaning and force of that belief is to be found in a persistence in moral endeavor and a determination to repair moral shortcomings. The pure religion of morality needs no dogma apart from these two fundamental articles of belief, which are accessible immediately to the simplest intelligence. Still less has it any need of the external trappings of religion—priests, ceremonies, and the like—although the body of believers must think of themselves as belonging to a church, universal but invisible, and the practices of visible churches sometimes serve to stimulate or strengthen moral effort, in a way which is useful but not indispensable.

The religion of morality is on this account a religion of all good men. Despite this, Kant took a particular interest in Christianity, which he saw as at least approximating true religion though corrupted by the presence of extraneous elements derived from Judaism. His book *Religion within the Bounds of Mere Reason* (1793) is in effect a commentary on and a reinterpretation of Christian doctrine and practice, written with the object of making this conclusion clear. In this reinterpretation the doctrine of original sin is transformed into a doctrine of the radical evil in human nature, which is the positive source of moral failing; and that of the Incarnation is replaced by an account of the triumph of the good principle over the bad, the part of the historical Jesus being taken by an idea of reason, that of man in his moral perfection. Kant sets aside the historical elements in Christianity as having no importance in themselves: Whatever is true in the religion must be derivable from moral reason. To think of the uttering of religious formulas or the performance of formal services to God as having a value of their own is to fall into the grossest superstition. It is perhaps scarcely surprising that these sentiments, whose attraction for youth can be seen in Hegel's *Jugendschriften*, should have struck the Prussian authorities as subversive and led the orthodox King Frederick William II to demand that Kant refrain from further pronouncements on religion. Though Kant, in his letter acceding to this demand, protested that he had no thought of criticizing Christianity in writing his book, it is hard to take his protest quite seriously, for he had certainly meant to suggest that many of the beliefs and actions of practicing Christians were without value, if not positively immoral. Indeed, the originality and continuing interest of his work on religion connect directly with that fact.

THE *OPUS POSTUMUM*. In the last years of his life—from about 1795 on—Kant was engaged in the composition of what would have been a substantial philosophical work; the preparatory notes for it have been published as *Opus Postumum*. Its original title was "Transition from Metaphysical Foundations of Natural Science to Physics," and in its original form its object was to carry further the process, begun in 1786 in the *Metaphysical Foundations of Natural Science*, of finding an a priori basis for physics. No longer content with the formal structure for which he had argued earlier, Kant thought he had to show that some of the particular laws of nature could be known in advance of experience. The broadest types of physical possibility were determined by the constitution of the human mind; it was this, for example, which explained the presence in nature of just so many fundamental forces, and even of an omnipresent ether.

These speculations about the foundations of physics led Kant to epistemological considerations of a wider kind. The whole subject of the relation of the form of experience to its matter, with the question how far the form shapes the matter, arose in his mind anew, doubtless because of the criticisms directed against the formalist position of the *Critique of Pure Reason* by self-professed disciples such as Fichte. In 1799 Kant dissociated himself publicly from the views expressed in Fichte's *Wissenschaftslehre*, according to which the subject of knowledge "posits" the objective world and so, in a way, creates nature. Yet the evidence of the *Opus Postumum* is that at

this time, or shortly thereafter, Kant was toying with similar ideas and was even using some of the same vocabulary. It is perhaps fortunate for Kant's reputation that he was not able to get his final philosophical thoughts into publishable form.

See also Aesthetic Judgment; Appearance and Reality; Aristotle; Baumgarten, Alexander Gottlieb; Beck, Jakob Sigismund; Burke, Edmund; Causation; Cosmological Argument for the Existence of God; Crusius, Christian August; Descartes, René; Determinism and Freedom; Dilthey, Wilhelm; Ethics, History of; Fichte, Johann Gottlieb; Geometry; Hegel, Georg Wilhelm Friedrich; History and Historiography of Philosophy; Hume, David; Intuition; Jacobi, Friedrich Heinrich; Kepler, Johannes; Knutzen, Martin; Leibniz, Gottfried Wilhelm; Locke, John; Logic, History of; Meier, Georg Friedrich; Newton, Isaac; Ontological Argument for the Existence of God; Perception; Propositions; Reason; Rousseau, Jean-Jacques; Schelling, Friedrich Wilhelm Joseph von; Space; Teleology; Time; Wolff, Christian.

Bibliography

WORKS BY KANT

Collected Works

Gesammelte Schriften. 23 vols., edited under the supervision of the Berlin Academy of Sciences. Berlin: Reimer, 1902–1955. The standard collected edition of Kant's writings; contains his correspondence and hitherto unpublished notes (including those for the *Opus Postumum*) as well as everything he published. Further volumes covering Kant's lectures are in preparation.

Kants Werke. 10 vols, edited by Ernst Cassirer. Berlin: Cassirer, 1912–1922. Contains the published works, with full indications of the contents of the original editions.

Philosophische Bibliothek series. Leipzig and Hamburg, 1904–. Separately bound editions of all the treatises listed below under the head "Main Treatises," with useful introductions; also includes *Kleinere Schriften zur Geschichtsphilosophie, Ethik, und Politik*.

Main Treatises

Allgemeine Naturgeschichte und Theorie des Himmels. Königsberg and Leipzig, 1755. English translation by W. Hastie in *Kant's Cosmogony* (Glasgow, 1900).

Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes. Königsberg, 1763. No serviceable translation.

Träume eines Geistersehers, erläutert durch Träume der Metaphysik. Königsberg, 1766. Translated by E. F. Goerwitz into English as *Dreams of Spirit-Seer* (New York, 1900).

De Mundi Sensibilis atque Intelligibilis Forma et Principiis Dissertatio. Königsberg, 1770. Translated into English in Handyside's *Kant's Inaugural Dissertation and Early Writings on Space* (Chicago: Open Court, 1929).

Kritik der reinen Vernunft. Riga, 1781; 2nd ed., Riga, 1787. The first edition is customarily referred to as A, the second edition as B. The most useful modern edition is by Raymond Schmidt (Leipzig, 1926). There are English translations by Francis Heywood (London, 1838); J. M. D. Meiklejohn (London, 1854), F. Max Müller (2 vols., London, 1881), and N. Kemp Smith (London Macmillan, 1929). Kemp Smith's version is the fullest and most reliable.

Prolegomena zu einer jeden künftigen Metaphysik die als Wissenschaft auftreten können. Riga, 1783. English translations are available by John Richardson (*Prolegomena to Future Metaphysics*, in *Metaphysical Works of the Celebrated Immanual Kant*, London, 1836), J. P. Mahaffy and J. H. Bernard (*The Prolegomena*, London, 1872), E. B. Bax (*Kant's Prolegomena and Metaphysical Foundations of Natural Science*, London, 1883), Paul Carus (*Prolegomena to Any Future Metaphysics*, Chicago, 1902), L. W. Beck (same title as Carus's, New York: Liberal Arts Press, 1951), and P. G. Lucas (same title as Carus's, Manchester, U.K., 1953). The Beck and Lucas translations are much the best.

Grundlegung zur Metaphysik der Sitten. Riga, 1785; 2nd ed., Riga, 1786. Translated into English by T. K. Abbott as *Fundamental Principles of the Metaphysic of Morals* (in *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*, London, 1873), by L. W. Beck as *Foundations of the Metaphysics of Morals* (in *Critique of Practical Reason and Other Writings in Moral Philosophy*, Chicago: University of Chicago Press, 1949), by H. J. Paton (from the 2nd edition) as *The Moral Law, or Kant's Groundwork of the Metaphysic of Morals* (London: Hutchinson, 1948). All three versions are good; Paton's is the most elegant. Quotations in the text of this article are from Paton's translation; the citations to page numbers in the 2nd edition are taken from Paton's marginal notation.

Metaphysische Anfangsgründe der Naturwissenschaft. Riga, 1786. Translated into English by E. B. Bax (in *Kant's Prolegomena and Metaphysical Foundations of Natural Science*, London, 1883).

Kritik der praktischen Vernunft. Riga, 1788. Translated into English by T. K. Abbott, and also by L. W. Beck, as *Critique of Practical Reason* in the books cited for the *Grundlegung zur Metaphysik der Sitten*. The Beck translation is accurate but somewhat clumsy; it has been published separately (New York, 1956).

Kritik der Urteilskraft. Berlin and Liebau, 1790. Translated into English by J. H. Bernard as *Kritik of Judgement* (London, 1892; later reprinted as *Critique of Judgement*). Also translated by J. C. Meredith in two parts, *Critique of Aesthetic Judgement* (Oxford, 1911) and *Critique of Teleological Judgement* (Oxford, 1928), which were reissued together as *The Critique of Judgement* (Oxford: Clarendon Press, 1952). Both the Bernard and the Meredith versions are of poor quality; Meredith's is slightly the better. Kant wrote an introduction to the *Kritik* that he discarded; it is available as *Erste Einleitung in die Kritik der Urteilskraft*, Vol. V, *Kants Werke* (Berlin, 1922), and in English translation by H. Kabir as *Immanuel Kant on Philosophy in General* (Calcutta: Calcutta University Press, 1935).

Die Religion innerhalb der Grenzen der blossen Vernunft. First part published separately, Berlin, 1792; published complete at Königsberg, 1793. Translated into English by John Richardson as *Religion within the Sphere of Naked Reason* (in

Essays and Treatises, London, 1798), by J. W. Semple as *Religion within the Boundary of Pure Reason* (Edinburgh, 1838), T. M. Greene and H. H. Hudson as *Religion within the Limits of Reason Alone* (Chicago: Open Court, 1934; 2nd ed., with new matter by J. R. Silber, New York: Harper, 1960). The 1960 edition (Greene-Hudson-Silber) is the best.

Zum ewigen Frieden: ein philosophischer Entwurf. Königsberg, 1795. The two best English translations are L. W. Beck's *Perpetual Peace* (in *Critique of Practical Reason and Other Writings in Moral Philosophy*, Chicago, 1949; separately bound, New York, 1957) and C. J. Friedrich's *Inevitable Peace* (New Haven, CT, 1948).

Metaphysik der Sitten, Part I, *Metaphysische Anfangsgründe der Rechtslehre*; Part II, *Metaphysische Anfangsgründe der Tugendlehre*. Separately bound, Königsberg, 1797. Part I translated into English by W. Hastie as *Kant's Philosophy of Law* (Edinburgh, 1887). Part II translated into English by James Ellington as *Metaphysical Principles of Virtue* (Indianapolis, 1964) and by Mary Gregor as *The Doctrine of Virtue* (New York: Harper and Row, 1964).

Anthropologie in pragmatischer Hinsicht. Königsberg, 1798. No English translation.

Eine Vorlesung über Ethik, edited from student's notes by P. Menzer. Berlin, 1924. Translated into English by Louis Infield as *Lectures on Ethics* (London: Methuen, 1930).

Kant on History. Indianapolis, 1963. Translations by L. W. Beck, R. E. Anchor, and E L. Fackenheim; contains Kant's minor essays on philosophy of history.

WORKS ON KANT

Life

The main sources for Kant's life, apart from his letters, are three memoirs published in Königsberg in 1804: L. E. Borowski's *Darstellung des Lebens und Characters Immanuel Kants*; R. B. Jachmann's *Immanuel Kant, geschildert in Briefen an einen Freund*; and E. A. C. Wasianski's *Immanuel Kant in seinen letzten Lebensjahren*. Wasianski's memoir is extensively used in Thomas De Quincey's "The Last Days of Kant" (*Works*, Vol. XII). See also Ernst Cassirer's *Kants Leben und Lehre* (Berlin, 1921), and for a useful short life Karl Vorländer's *Immanuel Kants Leben* (Leipzig, 1911).

Commentaries

Rudolf Eisler's *Kantlexicon* (Berlin, 1930) and Heinrich Ratke's *Systematisches Handlexicon zu Kants Kritik der reinen Vernunft* (Leipzig, 1929) are valuable aids to the Kantian student. The periodical *Kantstudien* has published many important contributions to Kantian scholarship and discussion.

For commentaries on the *Critique of Pure Reason*, see Hans Vaihinger, *Kommentar zur Kritik der reinen Vernunft*, 2 vols. (Stuttgart: Spemann, 1881–1892), which covers the opening sections only; N. Kemp Smith, *A Commentary to Kant's Critique of Pure Reason* (London: Macmillan, 1918; rev. ed., 1923), which is strongly influenced by Vaihinger's "patchwork" theory; H. J. Paton, *Kant's Metaphysic of Experience*, 2 vols. (London: Allen and Unwin, 1936), which covers the first half only and is sharply critical of Kemp Smith; A. C. Ewing, *A Short Commentary on Kant's Critique of Pure Reason* (London: Methuen, 1938). T. D. Weldon, *Kant's Critique of Pure Reason* (Oxford, 1945; 2nd ed., Oxford: Clarendon Press, 1958).

On the theory of knowledge see also H. A. Prichard, *Kant's Theory of Knowledge* (Oxford: Clarendon Press, 1909); A. C. Ewing, *Kant's Treatment of Causality* (London: K. Paul, Trench, Trubner, 1924); W. H. Walsh, *Reason and Experience* (Oxford: Clarendon Press, 1947); Graham Bird, *Kant's Theory of Knowledge* (London: Routledge and Paul, 1962); R. P Wolff, *Kant's Theory of Mental Activity* (Cambridge, MA: Harvard University Press, 1963).

For commentaries on the *Critique of Practical Reason* and other works on ethics, see L. W. Beck, *A Commentary on Kant's Critique of Practical Reason* (Chicago, 1960), a good source; H. J. Paton, *The Categorical Imperative* (London: Hutchinson, 1947), a detailed commentary on the *Grundlegung*; W. D. Ross, *Kant's Ethical Theory* (Oxford, 1954); A. R. C. Duncan, *Practical Reason and Morality* (Edinburgh, 1957); M. J. Gregor, *Laws of Freedom* (Oxford: Blackwell, 1963), which expounds the *Metaphysic of Morals*; P A. Schilpp, *Kant's Pre-Critical Ethics* (Evanston, IL: Northwestern University, 1938), A. E. Teale, *Kantian Ethics* (London: Oxford University Press, 1951).

For commentaries on the *Critique of Judgment*, see Konrad Marc-Wogau, *Vier Studien zu Kants Kritik der Urteilskraft* (Uppsala, Sweden, 1938); H. W. Cassirer, *A Commentary on Kant's Critique of Judgment* (London: Methuen, 1938).

Commentaries on other aspects of Kant's thought are also available. On the precritical writings, see Giorgio Tonelli, *Elementi metafisici e metodologici in Kant precritico* (Turin, 1959). On Kant's philosophy of religion, see C. C. J. Webb, *Kant's Philosophy of Religion* (Oxford: Clarendon Press, 1926) and F. E. England, *Kant's Conception of God* (London: Allen and Unwin, 1929). On Kant's philosophy of history, see Klaus Weyand, *Kants Geschichtsphilosophie* (Cologne: Cologne University Press, 1964). On Kant as a scientist, see Erich Adickes, *Kant als Naturforscher*, 2 vols. (Berlin, 1924–1925). On the *Opus Postumum*, see Erich Adickes, *Kants Opus Postumum dargestelt und beurteilt* (Berlin, 1920).

General Studies

S. Körner, *Kant* (Harmondsworth, U.K., 1955) is the best general introduction; see also G. J. Warnock's chapter in *A Critical History of Western Philosophy*, edited by D. J. O'Connor (New York: Free Press of Glencoe, 1964).

Edward Caird, *The Critical Philosophy of Kant*, 2 vols. (Glasgow, 1889), criticizes Kant from the Hegelian position. Martin Heidegger, *Kant und das Problem der Metaphysik* (Bonn: Cohen, 1929) examines Kant as an "ontologist." H. J. de Vleeschauwer, *La déduction transcendentale dans l'oeuvre de Kant*, 3 vols. (Antwerp, 1934–1937), presents an exhaustive survey of Kant's writings on the problem; de Vleeschauwer offers a one-volume summary in *L'évolution de la pensée kantienne* (Paris, 1939), which has been translated into English by A. R. C. Duncan as *The Development of Kantian Thought* (London: Nelson, 1962). Gottfried Martin, *Immanuel Kant, Ontologie und Wissenschaftstheorie* (Cologne, 1951), translated into English by P. G. Lucas as *Kant's Metaphysics and Theory of Science* (Manchester, U.K.: Manchester University Press, 1955), is influenced but not dominated by Heidegger. Richard Kroner, *Kants Weltanschauung* (Tübingen, 1914), translated into English by J. E. Smith as *Kant's Weltanschauung* (Chicago: University of Chicago Press, 1956), stresses Kant's emphasis on the practical. R. Daval, *La métaphysique de*

Kant (Paris: Presses Universitaires de France, 1951), presents "schematism" as the key idea in Kant's thought.

RECENT TRANSLATIONS

The Cambridge Edition of the Works of Immanuel Kant: Lectures on Metaphysics. Edited and translated by Karl Ameriks and Steve Naragon. Cambridge, U.K.: Cambridge University Press, 1997.

Critique of Pure Reason. Edited and translated by Paul Guyer and Allen W. Wood. Cambridge, U.K.: Cambridge University Press, 1998.

Ethical Philosophy. Translated by James W. Ellington. Indianapolis: Hackett, 1983.

Immanuel Kant: Critique of Judgment. Translated by Werner S. Pluhar. Indianpolis: Hackett, 1987.

Immanuel Kant: Critique of Pure Reason: Unified Edition (with all variants from the 1781 and 1787 editions). Edited and translated by Werner S. Pluhar. Indianapolis: Hackett, 1996.

The Metaphysics of Morals. Edited and translated by Mary Gregor. Cambridge, U.K.: Cambridge University Press, 1996.

Opus Postumum/Immanuel Kant. Edited by Eckart Forster. Translated by Eckhart Forster and Michael Rosen. Cambridge, U.K.: Cambridge University Press, 1993.

Perpetual Peace and Other Essays on Politics, History, and Morals. Translated by Ted Humphrey. Indianapolis: Hackett, 1983.

Prolegomena to Any Future Metaphysics: That Will Be Able to Come Forward as Science with Selections from the Critique of Pure Reason. Rev. ed. Edited and translated by Gary Hatfield. Cambridge, U.K.: Cambridge University Press, 2004.

RECENT SECONDARY LITERATURE

Allison, Henry E. *Kant's Theory of Freedom*. New York: Cambridge University Press, 1990.

Ameriks, Karl. *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy*. New York: Cambridge University Press, 2000.

Ameriks, Karl. *Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason*. 2nd ed. New York: Oxford University Press, 2000.

Beck, Lewis White. *Early German Philosophy; Kant and His Predecessors*. Cambridge, MA: Belknap Press of Harvard University Press, 1969.

Cohen, Ted, and Paul Guyer, eds. *Essays in Kant's Aesthetics*. Chicago: University of Chicago Press, 1982.

Di Giovanni, George. *Freedom and Religion in Kant and His Immediate Successors: The Vocation of Humankind, 1774–1800*. New York: Cambridge University Press, 2005.

Findlay, John N. *Kant and the Transcendental Object: A Hermeneutic Study*. New York: Oxford University Press, 1981.

Forster, Eckhart, ed. *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*. Stanford, CA: Stanford University Press, 1989.

Friedman, Michael. *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press, 1992.

Guyer, Paul, ed. *The Cambridge Companion to Kant*. New York: Cambridge University Press, 1992.

Guyer, Paul. *Kant and the Claims of Knowledge*. New York: Cambridge University Press, 1987.

Guyer, Paul. *Kant and the Claims of Taste*. 2nd ed. New York: Cambridge University Press, 1997.

Guyer, Paul, ed. *Kant and the Experience of Freedom: Essays on Aesthetics and Morality*. New York: Cambridge University Press, 1993.

Guyer, Paul. *Kant on Freedom, Law, and Happiness*. New York: Cambridge University Press, 2000.

Henrich, Dieter. *Aesthetic Judgment and the Moral Image of the World: Studies in Kant*. Stanford, CA: Stanford University Press, 1992.

Keller, Pierre. *Kant and the Demands of Self-Consciousness*. Cambridge, U.K.: Cambridge University Press, 1998.

Kitcher, Patricia. *Kant's Critique of Pure Reason: Critical Essays*. Lanham: Rowman & Littlefield, 1998.

Kuehn, Manfred. *Kant: A Biography*. New York: Cambridge University Press, 2001.

Longuenesse, Beatrice. *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Translated by Charles T. Wolfe. Princeton, NJ: Princeton University Press, 1998.

Neiman, Susan. *The Unity of Reason: Rereading Kant*. New York: Oxford University Press, 1994.

Stratton-Lake, Philip. *Kant, Duty, and Moral Worth*. London; New York: Routledge, 2000.

Van Cleve, James. *Problems from Kant*. New York: Oxford University Press, 1999.

Walker, Ralph C. S. *Kant*. New York: Routledge, 1999.

Walker, Ralph C. S. *Kant: The Arguments of the Philosophers*. Boston: Routledge & Kegan Paul, 1978.

Watkins, Eric. *Kant and the Metaphysics of Causality*. New York: Cambridge University Press, 2005.

Wood, Allen W. *Kant*. Malden, MA: Blackwell, 2005.

Wood, Allen W. *Self and Nature in Kant's Philosophy*. Ithaca, NY: Cornell University Press, 1984.

Yovel, Yirmiahu. *Kant and the Philosophy of History*. Princeton, NJ: Princeton University Press, 1980.

W. H. Walsh (1967)
Bibliography updated by Tamra Frei (2005)

KANT, IMMANUEL [ADDENDUM]

Immanuel Kant's philosophy continues to exercise significant influence on philosophical developments and generates an ever-growing body of scholarly literature. Work on Kant has progressed in two main directions. Central doctrines of the *Critique of Pure Reason* have been reconstructed, examined, and revised in the light of current philosophical concerns and standards; and the focus of scholarship has widened to include aspects and parts of Kant's work hitherto neglected, especially in the areas of ethics, aesthetics, philosophy of history, political philosophy, anthropology, and philosophy of science.

THE CRITIQUE OF PURE REASON

Further advances in interpreting the first *Critique* have occurred in three related areas: the nature and validity of Kant's overall argumentative procedure, with special emphasis on the deduction of the categories; the meaning and function of transcendental idealism and the associated distinction between things in themselves and appearances; and the role of mental activity in Kant's theory of experience.

The deduction of the categories, in which Kant sought to identify and justify the basic concepts underlying all experience and its objects, has become the center of major interpretive efforts. Stimulated by the neo-Kantian analytic metaphysics of Peter F. Strawson, philosophers have attempted to distill a type of argument from Kant's text that refutes skeptical doubts about the reality of the external world and other minds by showing how the skeptical challenge tacitly and unavoidably assumes the truth of the very assumptions it sets out to deny, namely, the reality of external objects and other minds.

While the force of such *transcendental arguments* remains controversial, the analytic–reconstructive approach to the deduction of the categories has also resulted in more textually based interpretations that reflect the whole spectrum of Kant scholarship. Readings of the deduction start either from the assumption of experience and proceed from there analytically to the necessary conditions of experience (the categories and the principles based on them), or take as their starting point some conception of self-consciousness or self-knowledge, either understood in Cartesian purity (a priori unity of apperception) or in phenomenological embeddedness (empirical self-consciousness), and argue from there to the synthetic conditions for the very possibility of such self-awareness. A key insight shared by many interpreters is the mutual requirement of object-knowledge and self-knowledge in Kant.

In interpretations of Kant's transcendental idealism, a major alternative has opened up between those scholars who see things-in-themselves and appearances as different aspects of one and the same things (*two-aspect view*) and those who regard the two as so many different sets of objects (*two-object view*). On the former view appearances are genuine objects. On the latter view they are representations. While the textual evidence is not conclusive for either view, the two-aspect theory has found many adherents because of its ontological economy and its avoidance of a phenomenalist reduction of things to representations.

The central role of human subjectivity in the deduction of the categories and in the defense of transcendental idealism has led to a renewed interest in Kant's philosophy of mind. Kant's theory of subjectivity is more and more seen as an integral part of his theoretical philosophy. Special areas of interest are the essential role of imagination in perception and experience, the distinction between inner sense and apperception, the relation between subjective or psychological and objective or logical grounds of knowledge, and the functional unity of sensibility and understanding. While no one advocates the derivation of the logical from the psychological in the manner of a reductive psychologism, the exact function of specifically psychological considerations in transcendental philosophy remains controversial. There is a minimal consensus that the self involved in the grounding of experience is distinct from the transcendent, noumenal self of the metaphysics of the soul, so forcefully rejected by Kant in the Transcendental Dialectic of the first *Critique*, and equally to be distinguished from the empirical self known through inner experience. Interpreters typically stress the formal and functional rather than the material and substantial sense of this *third*, transcendental self in Kant.

OTHER WORKS

Important new work on other parts of Kant's philosophy has occurred in three main areas: his practical philosophy, especially ethics; the *Critique of Judgment*, especially its aesthetics; and his philosophy of science. Scholarship on Kant's ethics has widened beyond the limited concern with the principle of morality (categorical imperative) to include other aspects of Kant's ethics as well as the position of Kant's moral theory within his social philosophy in its entirety and within the wider architectonic of the critical philosophy. A main inspiration of the work on Kant's ethics has been the neo-Kantian political philosophy of John Rawls, who sought to extract from Kant's formal approach to morality procedural guidelines for the ideal construction of the principles of social conduct. Increased attention has been paid to Kant's account of agency, the possible grounding of the categorical imperative in a generic conception of practical rationality, and the key features of Kant's moral psychology—including the theory of motivation, the role of moral judgment, and the function of subjective principles of action (maxims).

The move beyond the confines of Kant's foundational writings in moral philosophy has extended not only to his philosophy of law and theory of moral duties contained in the *Metaphysics of Morals* but also to his work in the phi-

losophy of religion, political philosophy, philosophy of history, and anthropology to be found in a number of his smaller works, often written in a more popular vein. The picture of Kant's practical philosophy that emerges from these reconstructions, revisions, and rediscoveries is that of a highly complex theory that is sensitive to the social dimension of human existence and well being able to respond to the charges and challenges posed by utilitarianism and communitarianism as well as virtue ethics.

In work on the *Critique of Judgment*, the standard emphasis on Kant's theory of aesthetic judgments has been widened considerably in recognition of the role of the third *Critique* as a synthesis of theoretical and practical philosophy in a comprehensive philosophy of human cultural development. A main focus of the scholarship on Kant's philosophy of science has been the *Opus postumum* and its attempts to specify the transition from an a priori theory of material nature to physics proper.

See also Cartesianism; Communitarianism; Neo-Kantianism; Psychologism; Rawls, John; Strawson, Peter Frederick; Utilitarianism.

Bibliography

WORKS BY KANT

The Cambridge Edition of the Works of Immanuel Kant, general editors Paul Guyer and Allen W. Wood. 15 vols. This set is the first comprehensive edition of Kant's published works, literary remains, correspondence, and lecture transcripts in English. The individual volumes of this edition are as follows: *Theoretical Philosophy 1755–1770*, translated and edited by Michael Walford with Ralf Meerbote. Cambridge, U.K.: Cambridge University Press, 1992. *Lectures on Logic*, translated and edited by Michael Young. Cambridge, U.K.: Cambridge University Press, 1992. *Opus postumum*. Translated by Eckart Förster and M. Rosen; edited by Eckart Förster. Cambridge, U.K.: Cambridge University Press, 1993. *Practical Philosophy*, translated and edited by Mary J. Gregor. Cambridge, U.K.: Cambridge University Press, 1996. *Religion and Rational Theology*, translated and edited by Allen W. Wood and George di Giovanni. Cambridge, U.K.: Cambridge University Press, 1996. *Lectures on Ethics*, translated by Peter Heath, edited by Peter Heath and Jerome B. Schneewind. Cambridge, U.K.: Cambridge University Press, 1997.

Lectures on Metaphysics, translated and edited by Karl Ameriks and Steve Narragon. Cambridge, U.K.: Cambridge University Press, 1997. *Critique of Pure Reason*, translated and edited by Paul Guyer and Allen W. Wood. Cambridge, U.K.: Cambridge University Press, 1998. *Correspondence*, translated and edited by Arnulf Zweig. Cambridge, U.K.: Cambridge University Press, 1999. *Critique of the Power of Judgment*, translated by Paul Guyer and E. Matthews, edited by Paul Guyer. Cambridge, U.K.: Cambridge University Press, 2000. *Theoretical Philosophy after 1781*, translated by Henry E. Allison, M. Friedman, G. Hatfield, and Peter Heath, edited by Henry E. Allison with Peter Heath. Cambridge, U.K.: Cambridge University Press, 2002. *Notes and Fragments*, edited by Paul Guyer. Cambridge, U.K.: Cambridge University Press, 2005. *Anthropology, History, and Education*, translated and edited by Günter Zöller and Robert Louden. Cambridge, U.K.: Cambridge University Press, 2006. *Natural Science*, translated and edited by Eric Watkins. Cambridge, U.K.: Cambridge University Press, forthcoming. *Lectures on Anthropology*, translated and edited by Robert Louden. Cambridge, U.K.: Cambridge University Press, forthcoming.

WORKS ON KANT

Allison, Henry E. *Kant's Theory of Freedom*. Cambridge, U.K.: Cambridge University Press, 1990.

Allison, Henry E. *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment*. Cambridge, U.K.: Cambridge University Press, 2001.

Allison, Henry E. *Kant's Transcendental Idealism. An Interpretation and Defense*. New Haven, CT: Yale University Press, 1983. Rev. ed., 2004.

Ameriks, Karl. *Kant's Theory of Mind. An Analysis of the Paralogisms of Pure Reason*. Oxford: Clarendon Press, 1982. Rev. ed. New York: Oxford University Press, 2000.

Förster, Eckart. *Kant's Final Synthesis. An Essay on the Opus postumum*. Cambridge, MA: Harvard University Press, 2000.

Friedman, Michael. *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press, 1992.

Guyer, Paul, ed. *The Cambridge Companion to Kant*. Cambridge, U.K.: Cambridge University Press, 1992.

Guyer, Paul, ed. *The Cambridge Companion to Kant and Modern Philosophy*. Cambridge, U.K.: Cambridge University Press, 2006.

Guyer, Paul. *Kant and the Claims of Taste*. Cambridge, MA: Harvard University Press 1979.

Henrich, Dieter. *The Unity of Reason*. Cambridge, MA: Harvard University Press, 1994.

Jacobs, Brian, and Patrick Kain. *Essays on Kant's Anthropology*. Cambridge, U.K.: Cambridge University Press, 2003.

Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge, U.K.: Cambridge University Press, 1996.

Kuehn, Manfred. *Kant. A Biography*. Cambridge, U.K.: Cambridge University Press, 2001.

Longuenesse, Béatrice. *Kant and the Capacity to Judge. Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Princeton, NJ: Princeton University Press, 1997.

O'Neill, Onora. *Constructions of Reason*. Cambridge, U.K.: Cambridge University Press, 1989.

Strawson, Peter F. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. London: Methuen, 1966.

Timmons, Mark, ed. *Kant's Metaphysics of Morals. Interpretative Essays*. Oxford: Oxford University Press, 2002.

Van Cleve, James. *Problems from Kant*. New York: Oxford University Press, 1999.

Watkins, Eric, ed. *Kant and the Sciences*. New York: Oxford University Press, 2001.

Wood, Allen W. *Kant's Ethical Thought*. Cambridge, U.K.: Cambridge University Press, 1999.

Günter Zöller (2005)

KANTIAN ETHICS

Ethical theories may be said to be "Kantian" if they take their inspiration or focus from themes in the ethical theory of Immanuel Kant, while attempting something other than interpretation, development, or defense of Kant's own ethical theory. This is not a hard and fast distinction: What appears the right way to defend some thesis of Kant's to one may appear to another to be a complete departure from the crucial components of Kant's critical ethics. Moreover, some, like scholars Onora O'Neill (1975), Marcia Baron (1996), and Barbara Herman (1993), may see their work as exploring and defending the essential elements of Kant's moral theory, rather than developing an alternative theory inspired by him, even though they do not accept the metaphysical picture Kant thought crucial to his account. Many defenders of Kant's own account see the austere picture sometimes drawn of his ethics—as based on a rigoristic and formalistic obligation to *duty*—as mistaken, and argue that Kant's conception of what people are like as moral agents, and of what morality requires of people, is far richer and more satisfying than is often supposed. Still, it is useful to see Kantian theorists as holding that Kant had some crucial or seminal ethical matters right, while at the same time committing himself to claims or views that are from their perspective unacceptable. Thus, Kantian ethicists may be understood as attempting to rework cherished Kantian insights within the bounds of an overall more acceptable framework.

METAPHYSICS

For many Kantian theorists, the point of departure from Kant is Kant's metaphysics and the role his metaphysical commitments play in his ethical theory. Kant struggled for a solution to the problem of how moral agents could be held responsible for their actions in a world governed by natural laws of cause and effect. If every event has a cause, which is itself caused, how could one see human action as anything but determined by the causes antecedent to it? And if human action is caused by natural law, in what sense can individuals see themselves as morally responsible?

Kant's solution to the problem drew on the metaphysical view developed in his *Critique of Pure Reason*, where he distinguished two worlds, one the world of sense—natural, physical, and empirical—and the other rational or "intelligible." The empirical world is governed by natural law, and effects do follow causes in ways determined by natural law. However, human beings are not merely natural but rational, and as members of the rational order are capable of "spontaneity": of producing effects based on determinations of reason, not causes. Because we have these two-fold natures, people occupy both worlds at once, and their actions are simultaneously subject to natural law and (as rational agents) to moral law.

Many Kantian ethicists find the proposal that people are citizens of some nonnatural world of reason implausible and unattractive. They aim to reconstruct the crucial elements of Kant's ethical theory without Kant's reliance on these metaphysical speculations. Most Kantian ethics are intended to develop Kantian ethical ideas while drawing on people's understanding of themselves as simply members of the natural world.

UNIVERSALIZABILITY

The strain in Kant's ethics that has found broadest employment is his idea that a practical principle (or "maxim") suitable for morally worthy action must be one which can hold universally, or, as Kant puts it, can be willed as a universal law; this is the first formulation of his "Categorical Imperative." Kant thought that when one acts immorally, one makes an exception of oneself, or makes exceptions for "just this one time," from laws one would will that everyone obey. Morality is thus best understood as the apprehension of principles that are universalizable in their scope and application.

This element of Kant's thought has echoes in numerous later thinkers. Marcus Singer (1961), for example, focuses on the general logic of what he calls the "generalization principle": What is right for one person must be right for anyone in the same or similar circumstances. One accepts the force of the question, "What would happen if everyone did that?" and Singer's theory is a study in the conditions of its legitimate application. Singer maintains that this principle is presupposed by any genuine moral judgment, and is the key to the moral principles that ground any plausible moral theory. However, Singer departs from Kant both in the metaphysical commitments previously described, and in his departure from considering what one could *will* to be universal, to assessing the desirability of the consequences of a principle with universal application.

Alan Gewirth's moral theory takes on the principle of generic consistency as its supreme moral principle. Like Kant, Gewirth (1978) begins with the premise that people are agents who act for ends; unlike Kant, Gewirth holds that, as agents, one must see the ends one is acting to realize as good. One sees them as good, however, only in light

of certain properties, or "generic features," of those ends. For example, one might have the end of getting adequate nutrition in virtue of its natural role in healthy life and agency. But then, Gewirth argues, consistency requires that one sees anything else with those "generic features" as good as well; thus, to be consistent, one must see as good adequate nutrition for anyone. Moreover, people are committed to seeing as good not only their capacity for action but also the freedom and well-being that make it possible, and consistency requires that they see these as good for others as well. They must thus see themselves as having claims against others that they respect their "generic rights": rights to freedom and well-being. But the principle of universalizability requires that, if people see themselves as having claims against others, they must likewise see others as having the same claims against *them*. Thus, as in Kant, the bare idea of agency, coupled with the rational requirement of universalizability, leads to the fundamental moral principle, in this case the principle of generic consistency, "Act in accord with the generic rights of your recipients as well as of yourself" (Gewirth 1978, p. 135).

For many theorists drawing on elements of Kant's view (Singer is an example), the Kantian approach is attractive as a way to oppose consequentialism in ethics. However, not all consequentialists agree. R. M. Hare (1981) argues that the focus on universalizability can be taken to ground a form of consequentialism. Hare argues that Kant's insights into the logical properties of moral terms lead, not to Kant's own ethical conclusions, but to a form of utilitarianism. This is because people must recognize that moral principles are *prescriptions* of a certain sort, namely universal prescriptions. But such prescriptions are in turn best understood as a sort of *preference*, and when one considers one's preferences as being constrained by the requirement that they hold universally, one sees that one's prescriptions must take the familiar consequentialist form of maximizing utility.

RESPECT FOR PERSONS

The second formulation of Kant's categorical imperative stipulates that persons are not to treat other persons as means only, but always at the same time as ends. Kant is often thought to have identified something crucial to a proper understanding of morality in this principle, and this way of understanding our obligations of respect for other persons has been widely influential.

Alan Donagan's work begins with some of the essential elements recognized in the notion of universalizability, but develops them in a direction more congruent with this feature of Kant's theory. Donagan (1977) sees Kant as an exemplar of a moral theory based on a common core that reaches back to the Stoics, the Hebrews, and the Christian tradition. This core is based on the thought that morality is addressed to rational creatures as such, in virtue of their rationality, and that its precepts, or moral law, must somehow be accessible to moral agents in virtue of that rationality. In Donagan's view, what emerges from scrutiny of this common core is the requirement that every human being be treated with the respect due a rational creature. This is closely related to Kant's second formulation of the categorical imperative, which Donagan finds superior to the "universal law" formulation of that imperative. Thus Donagan is an example of a Kantian theorist who takes Kant's starting point in a shared capacity for rationality and ends with a focus on respect for human nature.

Others similarly have found this element of Kant's work central to their own ethical conceptions. Thomas Hill (1991) interprets the metaphysically hoary elements of Kant's theory as an examination of what it is for a deliberating agent to choose how to act, what ends to pursue, and so on. From this perspective, one's "autonomy"—one's capacity to see oneself as capable of more than simply the pursuit of self-interest or satisfaction of preferences—is crucial, as it presupposes that one's status as a rational agent must be essential in one's deliberating about how to act.

However, David Cummiskey (1996) argues that the focus on respect for persons as valuable in virtue of their status as rational agents can ground a consequentialist approach as well. Cummiskey maintains that Kant's attention to the value of persons as ends-in-themselves is appropriate, but is incapable of justifying the sorts of claims often made against consequentialist accounts, which by their nature require that value be maximized. Rather, Cummiskey argues, Kant's view that rational agents are ends in themselves is itself a view with a form of value at its core, and there is nothing in the balance of Kant's theory to block the inference that such value ought to be maximized as a matter of moral obligation.

CONSTRUCTIVISM

Without question the greatest single influence in Kantian ethics has been the work of John Rawls (1971, 1999). Rawls's best-known work is in political theory, not ethics, and it draws more from Kant's *method* than from the content of Kant's views. Rawls took Kant's singular contribution to moral theory to be the notion that moral truth is not constituted independently of human reasoning and

rationality—independent of individuals in such a way that moral truth can be treated as an object of investigation, as scientific truth is; instead, moral truth is something that instead people constitute or bring into being ("construct") through the very process of deliberating about it. In Kant's own theory, this idea is represented in the argument that people understand moral obligation by way of reflection on what principles could be willed as universal law. This approach brings to the foreground the *procedures* by which individuals deliberate about and attempt to determine fundamental moral principles. Rawls's political theory consists in large part of the characterization of such a procedure to arrive at principles of justice, which, he argues, are best understood not as something individuals discover, but as something they would arrive at on deliberation under certain carefully crafted conditions. The conditions Rawls specifies for this deliberation are also intended to capture important features of Kant's conception of what people are like as moral and political agents, in particular the distinction between individual persons, deserving of the sort of respect Rawls believes his theory of justice provides.

Rawls's influence can be seen not only in political theory, but in a resurgence of interest in Kantian foundations for moral and political theorizing generally. Christine Korsgaard (1996) has adapted the constructivist approach in developing her Kantian ethical theory. On her view people recognize that, as reasoning agents, they need reasons to act, and as they assess where such reasons can come from—as they consider possible "sources of normativity"—they realize in the end that they must come from their own rational natures. People take their reasons, Korsgaard argues, from their "identities," and fundamental to any and all of these identities is their moral identity—their identity as agents acting on reasons. Reasons, Korsgaard argues, are inherently public, in the sense that they must be shareable among agents, so the enterprise of reflecting on how to act itself gives rise to the principles governing one's conduct.

See also Categorical Imperative; Constructivism, Moral; Deontological Ethics; Kant, Immanuel; Rationalism in Ethics.

Bibliography

Baron, Marcia, *Kantian Ethics Almost without Apology*. Ithaca, NY: Cornell University Press, 1995.

Cummiskey, David, *Kantian Consequentialism*. New York: Oxford University Press, 1996.

Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1977.

Gewirth, Alan. *Reason and Morality*. Chicago: University of Chicago Press, 1978.

Hare, R. M. *Moral Thinking*. Oxford: Clarendon Press, 1981.

Herman, Barbara. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Hill, Thomas. *Autonomy and Self-Respect*. Cambridge, U.K.: Cambridge University Press, 1991.

Hill, Thomas. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca, NY: Cornell University Press, 1992.

Kant, Immanuel, *Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: St. Martin's Press, 1965.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Translated by Allen W. Wood. New Haven, CT: Yale University Press, 2002.

Korsgaard, Christine. *Creating the Kingdom of Ends*. New York: Cambridge University Press, 1996.

Korsgaard, Christine. *The Sources of Normativity*. New York: Cambridge University Press, 1996.

O'Neill, Onora. *Acting on Principle*. New York: Columbia University Press, 1975.

Rawls, John. *Collected Papers*. Edited by Samuel Freeman. Cambridge, MA: Harvard University Press, 1999.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971, 1999.

Singer, Marcus George. *Generalization in Ethics*. New York: Atheneum Press, 1961.

Mark LeBar (2005)

KANTIANISM

See *Neo-Kantianism*

KAPLAN, DAVID

(1933–)

An American philosopher and logician, David Benjamin Kaplan was born in Los Angeles in 1933 and has spent his career mainly at the University of California, Los Angeles: first as an undergraduate student (AB in Philosophy, 1956; AB in Mathematics, 1957); then as graduate student (PhD in Philosophy, 1964), where he wrote the last dissertation Rudolf Carnap supervised; later as a faculty member, where he became Hans Reichenbach Professor of Scientific Philosophy in 1994.

Kaplan is best known for his work in formal semantics, particularly on the semantics of demonstratives and other indexicals: expressions such as *this*, *that purple Mercedes convertible*, *I*, *you*, *here*, *now*, and *actually*. In *Demonstratives*, Kaplan developed a theoretical framework in which sentences express propositions relative to contexts. The content of an expression (relative to a context *C*) is what it contributes to the propositions

expressed (relative to *C*) by sentences that contain it. The content of an expression determines an intension: a function from circumstances of evaluation to extensions (truth-values for sentences, individuals for singular terms, sets of individuals for predicates). Circumstances include at least possible worlds and perhaps also times. The character of an expression determines a function from contexts to contents.

In this framework, indexicals have variable contents but stable characters. For example, relative to a context *c* whose agent is McX, *I* has a content *x* (which determines a function that maps every circumstance onto McX himself); whereas, relative to a context *c** whose agent is Wyman, *I* has a different content *y* (which determines a function that maps every circumstance onto Wyman himself). But, relative to either context, *I* has the same character (which determines a function that maps *c* onto *x* and *c** onto *y*). Kaplan proposed that the character of an expression is its linguistic meaning and that it is an expression's character that is responsible for its cognitive value: The difference in cognitive value between "His pants are on fire!" and "My pants are on fire!," for example, lies in the difference between the characters of the indexicals *his* and *my*.

Indexicals are directly referential: For any context *C*, the content *o* of an indexical relative to *C* is the entity that the function determined by *o* maps every circumstance onto. For example, relative to *c*, whose agent is McX, the content of *I* is McX himself. Because indexicals are directly referential, a sentence that contains an indexical expresses a singular proposition (relative to a context *C*): a proposition that contains the entity that is the content of that indexical (relative to *C*). For example, relative to *c*, whose agent is McX, "I'm right" expresses a proposition that contains McX himself. This proposition can be represented as the ordered pair ⟨McX, the property being right⟩.

One surprising feature of this framework is that it allows one to distinguish logical truth and necessity. For example, "I am here now" is a logical truth in something like the following sense: Relative to any context *C*, it expresses a proposition that is true relative to the circumstance of *C* (at least provided that the agent of *C* is located at the time and place of *C* at the circumstance of *C*). But, at least relative to most contexts, the proposition expressed by "I am here now" is not necessary: It is not true relative to every circumstance (likewise for "I exist" and "ϕ if and only if actually ϕ").

Kaplan's philosophical thought has moved from Fregeanism to Russellianism. In his 1964 dissertation, *Foundations of Intensional Logic*, Kaplan developed a Carnapian model-theoretic semantics for Alonzo Church's Fregean logic of sense and denotation. In "*Quantifying In*" (1968–1969), Kaplan developed a Fregean account of belief ascriptions and of belief, one that allows quantification into belief ascriptions (as in "There is an *x* such that Ralph believes that *x* is a spy") under certain circumstances. By *Dthat* (1978) Kaplan had turned away from his early Fregeanism toward a Russellian view on which "John is suspicious," for example, expresses a singular proposition, one that contains John himself and that can be represented as the ordered pair ⟨John, the property being suspicious⟩.

Kaplan went on to become a major proponent of the previously moribund theory of singular propositions. His Russellianism reached its apogee in *Demonstratives* (1989a), where he argued that indexicals are directly referential and, hence, that sentences containing indexicals express singular propositions. Although, in his 1989 *Afterthoughts*, Kaplan admitted to feeling "a resurgence of atavistic Fregeanism," he continued to treat indexicals as directly referential.

After *Demonstratives* and *Afterthoughts*, Kaplan has worked on a number of further topics. In *Words*, he argued that the relation between a word and its occurrences should be thought of as the relation, not between a type and its tokens, but rather between a perduring entity and its temporal parts. He also suggested that it is a word itself that is responsible for its cognitive value: The difference in cognitive value between "Hesperus equals Hesperus" and "Hesperus equals Phosphorus," for example, lies in the difference between the words *Hesperus* and *Phosphorus*. In work on expressives (expressions such as *ouch* and *oops*), Kaplan suggested that one should shift from a semantics that pairs expressions with entities (*meanings*) to a semantics that pairs expressions with rules for their correct use. Kaplan also suggested that characters might best be understood, not as entities, but rather as such rules.

See also Logic, History of; Philosophy of Language.

Bibliography

PRIMARY WORKS

"Quantifying In." *Synthese* 19 (1–2) (1968–1969): 178–214.

"Dthat." In *Syntax and Semantics*. Vol. 9, edited by Peter Cole. New York: Academic, 1978.

"Demonstratives: An Essay on The Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals" [1977]. In *Themes from Kaplan*, edited by

Joseph Almog, John Perry, and Howard Wettstein. Oxford: Oxford University Press, 1989a.

"Afterthoughts." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. Oxford: Oxford University Press, 1989b.

"Words." *Supplement to the Aristotelian Society* 64 (1990): 93–119.

Ben Caplan (2005)

KAREEV, NIKOLAI IVANOVICH

(1850–1931)

Nikolai Ivanovich Kareev, the Russian historian and philosopher, was educated at Moscow University, where he took his doctorate in history (1884). During the late 1870s and early 1880s he spent several years studying abroad. Kareev taught modern European history, first at Warsaw University and then at St. Petersburg University. He became a corresponding member of the St. Petersburg Academy of Sciences in 1910 and an honorary member of the Soviet Academy of Sciences in 1929. His main historical studies were devoted to eighteenth-century France, especially the Revolution of 1789.

Although a moderate in politics, Kareev was deeply influenced by such radical Russian thinkers as Aleksandr Herzen, Dimitrii Pisarev, Pëtr Lavrov, and N. K. Mikhailovskii. Like Lavrov and Mikhailovskii, Kareev was a "semipositivist," but he was less influenced by either G. W. F. Hegel or Karl Marx than Lavrov had been. His views of history echo Herzen's "philosophy of chance." "History," Kareev declared, "is not a straight line, not a regular design traced out on a mathematical plane, but a living fabric of irregular and sinuous lines, which are intertwined in the most varied and unexpected ways" (*Osnovnye voprosy* [Fundamental problems], Part I, p. 153).

Kareev's position in ethics, which he called ethical individualism, was even more Kantian than that of Lavrov's early works. He defended individual autonomy against three dominant anti-individualist tendencies: that which breaks down the self into a series of psychic events (David Hume); that which turns the individual into an expression of the *Zeitgeist* or *Volksgeist* (Hegel); and that which reduces the individual to a product of socioeconomic relations (Marx). From the point of view of the "human dignity and worth of the individual person," Kareev insisted, "external [sociopolitical] freedom is a necessary condition for the spiritual growth and happiness of all the members of society" (*Mysli*, 2nd ed., 1896, p. 135).

Kareev rejected the "utilitarian attitude toward the person, which treats her as an object," adding that the "principle of individuality" guarantees the individual's right "not to be an instrument or means for another" or reduced to the status of an organ of a "social organism" (ibid., p. 138). In attributing absolute value to individuals as such, Kareev said, we take account of both their natural rights and—as Lavrov had stressed—their present potentiality for future moral and intellectual growth. In the name of this absolute value, Kareev condemned not only political assassination and capital punishment but also euthanasia. On this point he came close not only to Immanuel Kant but also to Lev Tolstoy, whose philosophy of history, like those of Hegel and Marx, he had criticized perceptively and in detail.

See also Ethics, History of; Hegel, Georg Wilhelm Friedrich; Herzen, Aleksandr Ivanovich; Kant, Immanuel; Lavrov, Pëtr Lavrovich; Marx, Karl; Mikhailovskii, Nikolai Konstantinovich; Philosophy of History; Pisarev, Dmitri Ivanovich; Russian Philosophy; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

WORKS BY KAREEV

"K voprosu o svobode voli s tochki zreniya teorii istoricheskovo protsessa" (On the question of freedom of the will from the standpoint of the theory of the historical process) appeared in *Voprosy filosofii i psikhologii* 1 (4) (1889–1890): 113–142. It was reprinted in *Istoriko-filosofskiye i sotsiologicheskiye etyudy* (Studies in sociology and the philosophy of history; St. Petersburg, 1895; 2nd ed., St. Petersburg, 1899), pp. 279–304. *Mysli ob osnovakh Nravstvennosti* (Thoughts on the foundations of morality) was published in St. Petersburg in 1895; 2nd ed., St. Petersburg, 1896.

Osnovnye voprosy filosofii istorii (Fundamental problems in the philosophy of history). Moscow: L. F. Panteleev, 1883.

Filosofiia kul'turnoi i sotsial'noi istorii novago vremeni (The philosophy of modern cultural and social history). St. Petersburg: M. M. Stasiulevich, 1893.

Obshchii khod vsemirnoi istorii: ocherki glavneishikh istoricheskikh epokh (The general course of world history: Essays on main historical epochs). St. Petersburg: Brokgauz-Efron, 1903.

Osnovy russkoi sotsiologii (Foundations of Russian sociology). St. Petersburg: Izdatel'stvo Ivana Limbakha, 1996.

WORKS ON KAREEV

Pogodin, S. *"Russkaia shkola" istorikov: N. I. Kareev, I. V. Luchitskii, M. M. Kovalevskii* (The "Russian school" of historians: N. I. Kareev, I. V. Luchitskii, M. M. Kovalevskii).

St. Petersburg: Sankt-Peterburgskii gosudarstvennyi tekhnicheskii universitet, 1997, pp. 139–202.

Zenkovsky, V. V. *Istoriya Russkoi filosofii*. 2 vols. Paris, 1948 and 1950. Translated by G. L. Kline as *A History of Russian Philosophy*, 2 vols. (London and New York, 1953), pp. 374–375.

Zolotarev, V. P. *Istoricheskaia kontseptsiia N. I. Kareeva: soderzhanie i evoliutsiia* (N. I. Kareev's conception of history: Its content and evolution). Leningrad: Izdatel'stvo Leningradskogo universiteta, 1988.

George L. Kline (1967)
Bibliography updated by Vladimir Marchenkov (2005)

KAREYEV, NICHOLAS IVANOVICH

See *Kareev, Nikolai Ivanovich*

KARMA

Karma (Sanskrit, *karman*; literally, "deed," "action") is an adjunct in Indian religious thought to the doctrine of Reincarnation. In one form or another, it is part of the beliefs of Buddhism, Jainism, and Hinduism. The actions of a living being are regarded as having a special class of causal effects that determine his future spiritual condition, both in this life and in succeeding ones. These effects are known as the "fruits" of the action. Good deeds lead to progress toward liberation (*mokṣa,* nirvana); bad ones, to regress from this goal. Usually caste status, disease, prosperity, and so forth are thought to be the consequences of actions in previous lives. Thus, karma is an ethically oriented causal law; and although some Hindus regard karma as the work of God, the concept does not necessitate this interpretation, and the award of deserts is as often regarded as an automatic process in nature.

The archaic notion of karma seems to have been that action as such binds men to the world (and thereby to suffering and ignorance); hence, liberation must involve suspension of all activity. Thus, in Jainism, which represents a very ancient strand in Indian religion, even a good action, although inducing an influx of meritorious karma, ties the person to matter. Indeed, karma, as the force determining rebirth, is itself regarded as a subtle form of matter. Also—and hence the emphasis on "non-injury" (*ahiṃsā*)—especially evil effects follow from a person's destroying life, even microorganisms. Such ideas lay behind the heroically quietistic Jain ideal of suicide by self-starvation. Moreover, the concept of karma in Vedic literature had the meaning of ritual act, so that combined with the need to refrain from activity there runs through much Indian ascetic thought the notion that even religious acts, although they may bring heavenly rewards, bind men to the cosmos and to rebirth: heaven is part of the cosmos and itself must be transcended.

These ideas presented a number of problems to speculative and religious thinkers: (1) How can liberation ever be achieved if even the effort to be inactive, and inactivity itself, may be forms of binding action? (2) How can the ordinary man, involved in his worldly duties and concerns, have any hope of escaping rebirth? (3) By what mechanism does karma operate on future births? (4) Why, if karma is what keeps empirical life going, does the saint (*jīvanmukta*), who has attained serenity and release in this life, keep on living? (5) How can there be any human initiative or free will if our present state is inexorably determined by past karma?

Various answers to these questions were given, among them the following: (1) The Jains hold that karmic matter can be annihilated by austerities, so that gradually it can be totally removed from an individual. On the other hand, Buddhism transformed the notion of karma by holding that motives, rather than the acts themselves, are what count and that karma needs craving (*taṇhā*) as a necessary condition of its effectiveness. Hence, by removing craving through the purification of one's motives, one can find release from rebirth. For the Hindu theologian Śankara, the power of karma depends on ignorance, so that the contemplative knowledge that the Self is the sole reality brings liberation from the continuing effects of karma.

(2) On the one hand, the ordinary man can hope to become a recluse, monk, or holy man in a future life. On the other hand, theistic ideas introduced grace as a countervailing means of liberation. Thus, in the *Bhagavad Gītā* it is stressed that a man, in performing his duties without regard to their fruits and in sole reliance upon the Lord, can escape the bonds of karma. Likewise, in Mahāyāna Buddhism the theory of the transfer of merit involves the belief that the otherwise unworthy individual can be given merit by a bodhisattva (Buddha-to-be) out of the latter's infinite store, acquired through many lives of heroic self-sacrifice on behalf of living beings; thereby the individual qualifies for rebirth in paradise (where the conditions for attaining nirvana are peculiarly favorable). Thus the operation of karma is short-circuited by grace and faith.

(3) It is commonly held that karma is *adṛṣṭa,* an invisible force, so that the need to postulate an observable

mechanism is evaded. However, among some schools the doctrine that the soul is all-pervasive (and not localized) helps to explain the concept of karmic action-at-a-distance. Traditional medical writings (first or second century) affirm that a person's characteristics are not derived solely from his parents (in this, there is an incipient conflict between modern genetics and the theory of karma).

(4) It is generally held that there is a limited continuance of karmic effects, like the running on of a potter's wheel after the potter has stopped turning it—but when the saint's death occurs, there will be no further rebirth for him.

(5) Various positions are adopted concerning the question of free will. The Buddha, for instance, was clearly impressed by the principle that knowledge of causes gives one the opportunity to determine the future, so that a proper understanding of karma and its causality should in no way involve fatalistic conclusions. He attacked Makkhali Gosāla, a contemporary teacher, for holding a fatalistic predestinationism, allied to extreme asceticism (which was in no sense a cause of final release, but merely symptomatic of one's progress). The Jains held that theoretically, in its pure state, the life monad or soul is capable of any kind of effort: Because of this "omnipotence" it never needs to be subservient to karma.

Although some schools argued that, since the effects of karma are morally regulated, one must presuppose a conscious regulator, namely God, atheistic and agnostic proponents of karma theory held that the difficulties of belief in God are as great as, or greater than, those inherent in assuming the automatic operation of karma. Moreover, belief in God generally involves the notion that unworthy people can short-circuit karma through calling on God in faith, and this cuts against the concepts of moral responsibility and self-help.

See also Indian Philosophy; Reincarnation; Responsibility, Moral and Legal.

Bibliography

Dasgupta, S. N. *A History of Indian Philosophy.* Cambridge, U.K.: Cambridge University Press, 1922. Vol. I, Ch. 4.

Paranjoti, Violet. *Saiva Siddhānta,* 2nd ed. London: Luzac, 1954. Pp. 38ff.

Parrinder, Geoffrey. *Upanishads, Gītā and Bible.* London, 1962. Ch. 9.

Tucci, Giuseppe. *Storia della filosofia indiana.* Bari, Italy, 1957. Part II, Ch. 10.

Zimmer, Heinrich. *Philosophies of India.* Edited by Joseph Campbell. New York: Meridian, 1956. Ch. 4, Sec. 7.

Ninian Smart (1967)

KARSAVIN, LEV PLATONOVICH

(1882–1952)

Russian historian-medievalist and religious philosopher Lev Platonovich Karsavin was born in St. Petersburg, the son of a ballet dancer and master, and the brother of the famous ballerina Tamara Karsavina. He graduated from the Department of History of Petersburg University in 1906 and stayed there as a teacher, doing studies in medieval spirituality and culture. Being a disciple of the prominent medievalist Ivan Grews, he soon started to develop his own approach, which can be considered in retrospect as an early prototype of the method of the French Annales school. His first big monograph (1912) was devoted to the early history of the Franciscan Order and the heretical sects of the Waldenses and Cathars. His next monograph, *Foundations of Medieval Spirituality in the Twelfth and Thirteenth Centuries, Mainly in Italy* (1915), is an important theoretical work of a type close to future studies in historical and cultural anthropology. Here Karsavin developed a methodology for historical studies based on the formation of general concepts like "an average religious person," "basic religious fund," and so forth, and tried to perform a reconstruction of the personality of the medieval individual in all its dimensions. The long-forgotten historical work of Karsavin, which includes also *Introduction to History: The Theory of History* (1920) and *Philosophy of History* (1923), was rediscovered in the 1970s and 1980s (chiefly in influential works by Aron Gurevich) and won recognition as a pioneering effort.

During the period of the Russian Revolution (1917–1922) Karsavin's thought shifted gradually to philosophy. This transition was stimulated by his interest in methodological and philosophical problems of history and Christian doctrine. Like a medieval scholastic thinker, he came to general metaphysical problems from reflection on Christian dogmas. In the same period, important changes in his life took place. Karsavin was opposed to the Bolshevik regime, not politically (he even considered the Bolsheviks to be the only force capable of ruling Russia), but ideologically and spiritually. Having a provocative style, he demonstrated his Christian convictions much more than he had before the revolution, lectured in a theological institute, and became the target of a vicious campaign in the official press. In the summer of 1922 he was arrested and then expelled to Germany together with a large group of noncommunist public figures, including leading religious philosophers (Nikolai

Berdyaev, Sergey Bulgakov, Nikolai Lossky, Semen Frank). In exile, he lived in Berlin (1922–1926), then in Clamart, next to Paris (1926–1928), and finally settled in Lithuania, where he was invited to hold the chair of general history at Kaunas University. Between 1925 and 1929 he took an active part in the Eurasian movement, becoming the leading theoretician of its left wing characterized by pro-Soviet views. During the twenties he wrote all his principal philosophical works, creating an original system of religious metaphysics.

Karsavin's system is the last big system of the so-called metaphysics of All-Unity. This philosophical school founded by Vladimir Solov'ëv took the central place in Russian religious philosophy and included leading figures of the Russian religious-philosophical renaissance of the twentieth century. By definition, its systems are based on the fundamental concept of All-Unity that represents a specific transrational principle of inner form describing perfect unity of a manifold such that any part of this manifold is identical to the whole of it. Karsavin gives this concept a new treatment, describing All-Unity as a sophisticated hierarchical system, structured vertically (into components or "moments" of higher and lower order, the latter being subsystems of the former) and horizontally (into a variety of moments of the same order). Vertical connections in this structure are described by the notion of *contractio* borrowed from Nicolas of Cusa, while horizontal ones are characterized by means of *conglomeratio et exglomeratio centri* found in Giordano Bruno and meaning that any two moments of the same order are connected not in a direct (i.e., causal) way, but only via the center of the whole system.

Drawing upon ancient doctrines and using their concepts in a constructivist and systematic way close to the theory of systems, this treatment is both archaizing and modernist. In Karsavin's system, the principle of All-Unity is subordinate to another fundamental principle, that of Tri-Unity, modeled on the Holy Trinity as it is presented in Christian dogma. Karsavin follows here the paradigm of dynamic ontology: Like many metaphysical doctrines, from Plotinus to Hegel, he treats being as a process governed by a triadic principle of development, where All-Unity represents the static aspect of Tri-Unity, its "stopping and rest."

Three ontological notions are identical in Karsavin's system: (perfect) Tri-Unity, God, and (perfect) Person. This trilateral identification also serves as the definition of Person. Human being is interpreted as an imperfect person that strives to perfection, that is, to God; all kinds of collective units, social and religious groups, nations, and classes are also considered as imperfect, embryonic persons and called *symphonic persons*. Karsavin's personalistic turn was new for the metaphysics of All-Unity, which, starting with the Greeks, had traditionally developed in an impersonal symbolist vein. The personalistic trend is further enhanced in Karsavin's description of the world process. The three stages of ontological dynamics are primal unity, disjoining, reunification; the central stage is interpreted as nonbeing or death. In the act of creation God endows with being the reality that he creates, thus depriving himself of being (kenosis) and voluntarily choosing sacrificial death. This voluntary passing of one's own being to somebody, identical to voluntary sacrificial death for somebody, is the definition of (perfect) love—whence it follows that the creature, striving to God, advances to pass, in its turn, its own being to God and thereby ascends to its own sacrificial death out of love.

Thus Karsavin's philosophy presents itself as an ontological drama of death, sacrifice, and love. These principles of his thought turned out to be perfectly realized in the final years of his life. When in 1944 it was clear that the Soviet Union was about to recapture Lithuania, Karsavin refused to leave and move to the West. In 1946 he was dismissed from the university for his deliberately defiant attitude toward Soviet authorities. In 1949 he was arrested and sent to a concentration camp in Abez, near the polar circle. In the gulag he wrote about ten texts of spiritual poetry and metaphysics and until his final days (he died there from tuberculosis) was a spiritual guide and teacher for his fellow prisoners. After the fall of communism, all of Karsavin's principal works were republished in Russia and have been actively studied.

See also Philosophy of History; Philosophy of Religion; Russian Philosophy.

Bibliography

WORKS BY KARSAVIN

Vostok, Zapad i russkaia ideia. St. Petersburg: Academia, 1922. 2nd ed., Moscow, 1991.

Noctes Peropolitanae. St. Petersburg: A. S. Kagan, 1922.

O nachalakh. Berlin: Obelisk, 1925.

O lichnosti. Kaunas, Lithuania: 1929.

"Poema o smerti." *Eranus* 2 (1931): 231–310.

Religiosno-filosofskie sochineniia. Podgotovka tekstov, predislovie, kommentarii Sergey Horujy. Moscow: Renaissance, 1992.

Malye sochineniia. Podgotovka tekstov, predislovie, kommentarii Sergey Horujy. St. Peterburg: Aletheia, 1994.

WORKS ON KARSAVIN

Wetter, Gustav A. "L. P. Karsawins Ontologie der Dreieinheit." *Orientalia Christiana Periodica* 9 (1943): 366–405.

Wetter, Gustav A. "Zum Zeitproblem in der Philosophie des Ostens: Die Theorie der 'Allzeitlichkeit' bei L.P.Karsawin." *Scholastik* 3 (1949): 345–366.

While there are no English-language works devotedly solely to Karsavin, the following two English-language works contain sections on Karsavin: Zenkovsky, V. V., *A History of Russian Philosophy*, vol. 2 (London: Routledge & Kegan Paul Ltd., 1953); Meerson, Michael Aksionov, *The Trinity of Love in Modern Russian Theology* (Franciscan Press: 1998).

Sergey Horujy (2005)

KATHARSIS

Katharsis is a beneficial transformation of painful emotions through absorbed contemplation of a powerfully moving work of art. The root meaning of "katharsis" in Greek is cleansing. The word can indicate the removal of impurities from, hence the amelioration of, any kind of substance. Before Aristotle, some philosophers had spoken (metaphorically) of psychological katharsis. Aristotle's student Aristoxenus claimed that Pythagoreans "achieved katharsis of the body through medicine, katharsis of the soul through music" (frag. 26). Plato sometimes employs the terminology of "katharsis" for philosophically extricating the soul or intellect from bodily concerns (e.g., *Phaedo* 67c; compare *Sophist* 226d–231b). But Aristotle was the first person to apply the term "katharsis" to the experience of tragedy.

The last clause of Aristotle's definition of tragedy in *Poetics* 6 describes tragedy as "accomplishing through pity and fear the *katharsis* of such emotions." No further reference to katharsis as the effect of tragedy occurs in the *Poetics*. Controversy over the "katharsis" clause remains acute, with no solution commanding great confidence. At issue are questions like the following: Did Aristotle mean occurrent emotions or underlying dispositions? Are pity and fear the only emotions involved? Is emotion the object or only the agency of katharsis? Does the term "katharsis" carry medical and/or religious overtones? Are the minds of tragedy's spectators purged, purified, clarified, or refined?

Our best aid to interpreting tragic katharsis is the account of musical katharsis in Aristotle's *Politics* 8.6–7, where Aristotle posits both pathological and normal cases of the phenomenon. As pity and fear are specifically cited in this context and further elucidation is promised in a discussion of poetry, there is a clear link with the *Poetics*. While *Politics* 8, focusing on educational needs, distinguishes various uses of music, it adopts a fundamentally character-centered view of music's capacity to "change the soul" through the passions (1340a4–b19). Though Aristotle regards both tragedy and music as mimetic (representational and expressive) art forms that arouse intense emotional states in their audiences, in his general moral psychology, ethical judgment, while cognitive, is influenced by feeling (*Nicomachean Ethics* 2.2–5, *Rhetoric* 2.1–11). Hence, we should not drive a wedge between the emotional and cognitive implications of katharsis.

Aristotle partially compares the mental effects of musical katharsis to both medical and ritual katharsis, but he nonetheless keeps musical katharsis independent of those spheres. *Politics* 8 encourages a model of tragic katharsis that integrates cognitive, affective, and ethical reactions into the special pleasure of tragedy. Since these reactions stem from emotional engagement with a mimetic plot structure (*Poetics* 14), and since all experience of mimesis is guided by cognitive awareness (*Poetics* 4), Aristotle's larger theory of tragedy supports the view that katharsis operates together with cognition and pleasure. Even so, katharsis should be viewed not as tragic pleasure per se but as a beneficial transformation of painful emotions, through the absorbed contemplation of a powerfully moving artwork, into a key component of a satisfyingly unified experience.

Because katharsis requires an uninhibited flow of emotion, it may bring a sense of "relief" (*Politics* 1342a14) and reduce any excess. But the popular modern association of katharsis with mere draining of blocked emotion oversimplifies Aristotle's perspective. The combined evidence of the *Poetics* and *Politics* suggests that Aristotle addressed Plato's concerns about emotional responses to art (*Republic* 606) by maintaining that such heightened emotion could channel an ethically valuable alignment of feeling and understanding. If so, it is plausible that his concept of katharsis had application to several art forms, perhaps including comedy.

See also Aristotle; Emotion; Plato; Pythagoras and Pythagoreanism; Tragedy.

Bibliography

Aristotle. *Nicomachean Ethics*. Translated by Christopher Rowe. Oxford, U.K.: Oxford University Press, 2002.

Aristotle. *On Rhetoric*. Translated by George A. Kennedy. New York: Oxford University Press, 1991.

Aristotle. *Poetics*. Translated by Stephen Halliwell. Cambridge, MA: Harvard University Press, 1995.

Aristotle. *Politics. Books VII and VIII*. Translated by Richard Kraut. Oxford: Clarendon Press, 1997.

Halliwell, Stephen. *Aristotle's Poetics*. London: Duckworth, 1986.

Halliwell, Stephen. "La psychologie morale de la catharsis: Un essai de reconstruction." *Études philosophiques* 4 (2003): 499–517.

Lear, Jonathan. "*Katharsis*." *Phronesis* 33 (1988): 297–326.

Plato. *Complete Works*, edited by J. M. Cooper. Indianapolis: Hackett, 1997.

Stephen Halliwell (2005)

KAUFMANN, WALTER ARNOLD

(1921–1980)

Walter Kaufmann was born in Freiburg, Germany, on July 1, 1921. He emigrated to the United States in 1939, as conditions in Germany became ominous for those of Jewish descent (Kaufmann's father—although not his mother—had converted to Protestantism, with the consequence that Kaufmann had been raised in that faith; but he converted to Judaism in 1933, in an early display of the sensitivity to religious questions that became one of the central features of his intellectual life). He attended Williams College, from which he graduated in 1941, and then went to Harvard, from which he received an MA degree in Philosophy in 1942. After military service in Europe during the Second World War (in capacities that took advantage of his equal facility in German and in English), he returned to Harvard, receiving his PhD in 1947. He joined the Philosophy Department at Princeton University in the fall of that year, which remained his academic base until his untimely death on September 4, 1980, at the age of only 59, from a mysterious illness he apparently contracted while traveling in Egypt and Africa.

Kaufmann played a major role in the introduction of existential philosophy (of Jean-Paul Sartre in particular) and the rehabilitation of G.W.F. Hegel and Friedrich Nietzsche (who had come to be all too closely associated with the Germany of the kaiser and of Adolf Hitler) in the English-speaking world in the decades following the Second World War. As one of the few members of major philosophy departments in those years who had a strong interest in developments in post-Kantian European philosophy, and as a prolific translator as well as interpreter of the writings of some of the most important figures in that tradition, he emerged as its most prominent, visible, and articulate champion, during the very decades in which the new Britain-based import of analytic philosophy became dominant in the philosophy departments at most major American universities. Much of Kaufmann's career was spent in often heated conflict as an advocate of the continental tradition (as it came to be called) against the newly dominant analytical paradigm that he regarded as a disaster for philosophy, and also as an advocate of those within that tradition (Hegel, Nietzsche, Sartre, and Martin Buber in particular) against the influence and popularity of others within it of whom he had a very low opinion (such as Karl Marx and Martin Heidegger).

Because Kaufmann had a Jewish identity and made no secret of it (even though he also made much of his rejection of Jewish theology), he was ideally positioned to be able to reject the charge of anti-Semitism that had contributed to the widespread hostility to Nietzsche before, during, and after the war years, and to defuse the imputation to Nietzsche of other proto-Nazi sentiments along with it. His association of Nietzsche with Sartrean existentialism was another of his strategies in pursuit of this objective; for, unlike Heidegger, Sartre's anti-Nazi credentials were impeccable, and Sartre himself sought to portray his existentialism as a kind of radical humanism. Kaufmann further presented Nietzsche as a kindred spirit of the heroes of the Enlightenment, and even of Emersonian individualism and later American pragmatism. This interpretation of Nietzsche found a ready reception in a wide and growing audience in the years following the publication of Kaufmann's classic *Nietzsche: Philosopher, Psychologist, Antichrist* in 1950, which remains one of the best general introductions to Nietzsche's thought written for English-speaking readers.

Moreover, while Kaufmann never published another book-length study of Nietzsche, he exerted an even greater influence upon the reception of Nietzsche in the English-speaking world through his much-needed new translations of (and introductions and notes to) most of Nietzsche's major works over a period of two decades, beginning with his phenomenally popular anthology *The Portable Nietzsche* in 1954, culminating with Nietzsche's *The Gay Science* in 1974, and including the controversial collection of selections from Nietzsche's notebooks from the 1880s published after his death under the title *The Will to Power*, thereby giving that volume a prominence and appearance of legitimacy that many feel it does not deserve. And by passing over the various works Nietzsche published between *The Birth of Tragedy* and *Thus Spoke Zarathustra*, Kaufmann influenced what English-

speaking readers ever since have come to regard as Nietzsche's most important works.

Kaufmann simultaneously attempted to renew interest in Hegel, in a manner intended to liberate Hegel from the moribund tradition of interpretation that had flourished in Britain and America under the banner of idealism in the late nineteenth and early twentieth centuries. Kaufmann's Hegel was closer to existentialism than he was to that metaphysical idealism, as he tried to show in his *Hegel: A Reinterpretation* (1966); and his Hegel championed a political philosophy that was a major, but sadly forgotten and neglected, alternative to the options upon which attention was focusing in both analytical and Marxist circles at that time. So Kaufmann first published a study of *Hegel's Political Philosophy* (1970), and then a volume of his own essays in this area reflecting his own mix of Hegelian and Nietzschean elements, *Without Guilt and Justice* (1973). He aspired to be taken seriously as a moral, social, and political philosopher; but the failure of these volumes to attract significant attention led him to turn his efforts in other directions.

Kaufmann had followed his early study of Nietzsche and anthology of Nietzsche's writings with two very popular volumes attempting to do the same thing for existential philosophy—his anthology *Existentialism from Dostoevsky to Sartre* (1956), which was everyone's introduction to existentialism for many years, and his collection of essays *From Shakespeare to Existentialism* (1959), which sought to situate existentialism in intimate if not entirely harmonious relation to an intellectual tradition that included the greatest contributions to Western literature and thought. The relationship between existential and tragic thought, literature, and experience held a particular fascination for him, which he explored in his *Tragedy and Philosophy* (1968).

These interests led Kaufmann to attempt to position himself in relation to traditional forms of philosophical and religious thought, first in his combative early *Critique of Religion and Philosophy* (1958), and then in his impassioned attempt to formulate and articulate his own post-traditional secularly religious credo *The Faith of A Heretic* (1960). His attempts to come to terms with religion continued in two volumes published in 1976, a volume of essays on *Existentialism, Religion, and Death*, and a book intended for a wider audience and marking the beginning of his attempt to integrate philosophy and photography, *Religions in Four Dimensions: Existential and Aesthetic, Historical and Comparative.*

This experiment continued in a trilogy published three years later (1979), under the general title *Man's Lot.* In this three-volume study of the human condition—*Life at the Limits*, *Time Is an Artist*, and *What Is Man?*—Kaufmann revealed himself as a truly gifted photographer with a powerful ability to employ that gift in the service of his attempt to plumb the heights and depths of human reality. That trilogy was followed by another, *Discovering the Mind* (1980–1981), with which his life abruptly ended, and the third volume of which was published following his death.

In each of these three last volumes Kaufmann considered the contributions of three major figures to this discovery: J.W. Von Goethe, Immanuel Kant, and Hegel; Nietzsche, Heidegger, and Buber; and Sigmund Freud, Alfred Adler and Carl Jung. This, he believed, was the real philosophy of mind; and it was his hope, through these volumes, to enrich philosophical thinking with respect to the mind by connecting it with this tradition—as he had sought to enrich philosophical thinking with respect to the human condition in the previous trilogy, and to enrich moral, social, and political thought by an infusion into them of Hegelian and Nietzschean ways of thinking.

Kaufmann found it at first frustrating and then deeply distressing that he was not taken seriously by the new analytic-philosophical establishment of his day, other than (by some) as Nietzsche's best translator and most appealing reinterpreter. This made him increasingly estranged from and critical of that establishment and philosophical orientation, and may have prompted his involvement in his last years with the EST human potential movement and his willingness to be associated with the Moon Unification Church's International Conference on the Unity of Sciences in the 1970s.

His later work itself was of a character that could hardly have been more at odds with the aims and paradigms of analytic-philosophical inquiry. Yet he considered himself to be true to the real heart and soul of the Socratic philosophical tradition, and to be its advocate and defender in a time in which he felt academic philosophy had lost its way. He welcomed the opportunity to enter the fray of popular debate as a public intellectual who was more than willing to continue Nietzsche's effort to fight the good fight of disillusioned enlightenment that was neither religious, scientistic, nor historically optimistic. He thought that philosophy could and should make a difference in human life, and that that difference should be in the direction of an uncompromisingly secular, post-metaphysical, strongly individualistic, but intensely interpersonal, existential humanism. Had he lived to develop and make a case for that vision of authentic humanity, he might well have attained the

recognition in the philosophical community that escaped him.

See also Continental Philosophy.

Bibliography

WORKS BY KAUFMANN

Nietzsche: Philosopher, Psychologist, Antichrist. Princeton, NJ: Princeton University Press, 1950.

Critique of Religion and Philosophy. Princeton, NJ: Princeton University Press, 1958

From Shakespeare to Existentialism. Boston: Beacon Press, 1959.

The Faith of A Heretic. Garden City, NY: Doubleday, 1961.

Cain and Other Poems. Garden City, NY: Doubleday, 1962.

Hegel: A Reinterpretation. Notre Dame, IN: University of Notre Dame Press, 1966.

Hegel: Texts and Commentary. Notre Dame, IN: University of Notre Dame Press, 1966.

Tragedy and Philosophy. Garden City, NY: Doubleday, 1968.

Hegel's Political Philosophy. New York: Atherton Press, 1970.

Without Guilt and Justice: from Decidophobia to Autonomy. New York: P. H. Wyden, 1973.

Existentialism, Religion, and Death: Thirteen Essays. New York: New American Library, 1976.

Religions in Four Dimensions: Existential and Aesthetic, Historical and Comparative. New York: Reader's Digest Press, 1976.

The Future of the Humanities. New York: Reader's Digest Press, 1977.

Man's Lot. New York: Reader's Digest Press, 1979.

Discovering the Mind. New York: McGraw-Hill, 1980–1981.

Richard Schacht (2005)

KAUTSKY, KARL
(1854–1939)

Karl Kautsky was, with the exception of Karl Marx and Friedrich Engels, the leading theorist of orthodox Marxism before World War I. Born in Prague of Czech and German parentage, Kautsky studied at Vienna and showed much interest in social Darwinism and socialism. As an evolutionist and materialist, he found Marx's combination of dialectical materialism and economic determinism irresistible, and he worked with Engels himself during the 1880s. From 1883 to 1917 Kautsky was the editor of *Die neue Zeit,* the official organ of the German Social Democratic Party and the most influential socialist journal of the day. He edited and published the literary remains of Marx after Engels's death. In 1891 Kautsky wrote the famous first, or theoretical, part of the *Erfurter Programm,* the official policy statement of the German party. This document established that the greatest socialist party in history should be orthodox Marxist.

Kautsky, more than any other theorist of repute, accepted Marx's method and conclusions as he found them. The natural laws of economic development resulted in certain inevitable contradictions in capitalism that must necessarily lead to its destruction and replacement by socialism. This would occur, Marx and Kautsky held, because competition and technical improvements, together with the availability of surplus labor, would lead to the concentration of capital and the progressive immiserization of the proletariat, as well as the polarization of society into a few monopolists opposed by vast masses of starving workers. Recurrent depressions and economic catastrophes would finally destroy capitalism. Such crises would be caused mainly by the inability of the workers to purchase the products of their labor. The united proletariat, trained by its socialist leaders, would see that only social ownership of the means of production could end the contradiction between capitalism's ability to produce wealth and its inability to distribute that wealth through private ownership. Like Marx and Engels, Kautsky held that religion, philosophy, and ethics are reflections of the substructure of class interest and position and that the state is the puppet of the dominant social class.

Kautsky, the "defender of the faith," fought attempts of fellow socialists to make basic alterations in their Marxian heritage. He led the German Social Democratic Party in its struggle against Eduard Bernstein and the revisonists, who believed that the facts of European capitalism no longer supported his orthodox views and that parliamentary action and pragmatic flexibility could bring extensive and permanent reform. Kautsky was able to maintain the preeminence of orthodox Marxism in party theory, although the revisionists increasingly dominated party tactics and action. In the early years of the twentieth century, Kautsky and the orthodox centrists had increasingly to contend with the radical left wing of the party under Rosa Luxemburg and Karl Liebknecht. This group held strictly to Marx's economic teachings but rejected orthodox political tactics in favor of more immediately revolutionary doctrines. They hoped for more radical positions on questions before parliament and for greater encouragement of spontaneous revolutionary and general strike activity. Kautsky did not believe that the contradictions of capitalism or the class consciousness of the workers were advanced enough for such tactics. He did join the Left in parliament on various crucial ques-

tions, notably in its refusal to sanction the continuance of World War I as a war of conquest.

During the Weimar Republic, Kautsky lost his preeminent position as the reformists dominated the party and Leninism captured the Left. He was attacked by V. I. Lenin and Leon Trotsky for his castigation of their dictatorial and terroristic methods and their conquest of Georgia, then an independent socialist-controlled state. Forced into exile by the Nazis, Kautsky died in Amsterdam.

See also Darwinism; Dialectical Materialism; Engels, Friedrich; Lenin, Vladimir Il'ich; Marxist Philosophy; Marx, Karl; Socialism.

Bibliography

WORKS BY KAUTSKY

Das Erfurter Programm. Stuttgart, 1899.

The Social Revolution. London: Twentieth Century Press, 1909.

Bernstein und das sozialdemokratische Programm. Stuttgart, 1919.

Ethik und materialistische Geschichtsauffassung. Berlin, 1922.

Die Materialistische Geschichtsauffassung. Berlin, 1927. Translated by Raymond Meyer with John H. Kautsky as *The Materialist Conception of History.* New Haven, CT: Yale University Press, 1988.

Die Geschichte des Sozialismus. Berlin, 1947.

WORKS ON KAUTSKY

Lenin, Nikolai. *The Proletarian Revolution and Kautsky the Renegade.* London: British Socialist Party, 1920.

Renner, Karl. *Karl Kautsky.* Berlin, 1925.

Salvadori, Massimo L. *Karl Kautsky and the Socialist Revolution 1880–1938.* Translated by J. Rothschild. London: New Left Books, 1979.

Steenson, Gary P. *Karl Kautsky 1854–1938: Marxism in the Classical Years.* Pittsburgh: University of Pittsburgh Press, 1978.

Trotsky, Leon. *The Defense of Terrorism: A Reply to Kautsky.* New York, 1921.

John Weiss (1967)
Bibliography updated by Philip Reed (2005)

KAVELIN, KONSTANTIN DMITRIEVICH

(1818–1885)

Konstantin Dmitrievich Kavelin, the Russian historian and philosopher, was educated at Moscow University, where he was later professor of history. Kavelin also taught at St. Petersburg University and was for a time tutor to the royal family. In addition to numerous historical works, he wrote essays in psychology, sociology, and ethics. During the 1870s he carried on an active polemic with Vladimir Solov'ëv, defending a positivist (or "semipositivist") position against Solov'ëv's criticisms. In politics Kavelin was a moderate liberal; in religion he remained devoutly Russian Orthodox.

Kavelin's main work in ethical theory, *Zadachi etiki* (Tasks [or problems] of ethics), appeared in 1844. In it he criticized the then fashionable one-sided "objectivism," which, he charged, blurred the distinction between inner intention and outward behavior, leading to the conclusion that intentions may be "unlawful" or volitions "criminal." From the neo-Kantian viewpoint that Kavelin adopted in this book, such a conclusion is absurd. Intentions and volitions, he insisted, are to be judged only "by their relationship to consciousness, to the understanding and inner conviction of the person in whom they occur" (*Sobranie sochinenii* [Collected works], Vol. III, col. 907).

When utilitarians equate virtue with utility and vice with social harm they are taking an "outsider's" view of moral experience, the view of a spectator rather than that of a moral agent. In fact, moral virtue may or may not be useful; this depends on the particular social system involved, and the latter is a nonmoral factor. Hence, social utility cannot provide a sound criterion of morality.

It is human individuality as a unique locus of value, Kavelin asserted, which provides such a criterion. However, this assertion raised serious problems for Kavelin's "scientific ethics," since, as he admitted, concrete individuality systematically eludes the abstract generalities of science. In the end, the "scientific ethics" that Kavelin had been laboring to construct coincided with Christian ethics—the "last word in ethical wisdom" and "an incontrovertible truth of individual spiritual life" (*Sobranie sochinenii* [Collected works], Vol. III, Cols. 940–941).

Kavelin's attempt to provide a scientific foundation for ethics, like the attempts of other nineteenth-century thinkers, must be judged a failure. However, Kavelin eloquently restated ideas derived from Vissarion Belinskii, Aleksandr Herzen, and the Russian Populists concerning the individual person and his sense of freedom and the role of convictions in morality. His was a genuine, if modest, philosophical contribution.

See also Belinskii, Vissarion Grigor'evich; Ethics, History of; Herzen, Aleksandr Ivanovich; Metaethics; Neo-Kantianism; Philosophy of History; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich; Utilitarianism.

Bibliography

Vol. III of Kavelin's four-volume *Sobraniye sochinenii* (Collected works; St. Petersburg: Stasiulevicha, 1898–1900) contains Kavelin's philosophical works, including *Zadachi etiki*, cols. 897–1018.

For literature on Kavelin, see V. V. Zenkovsky, *Istoriia Russkoi filosofii*. 2 vols. (Paris: YMCA Press, 1948–1950). Translated by George L. Kline as *A History of Russian Philosophy* (New York: Columbia University Press, 1953), pp. 345–348.

See also V. I. Prilenskii, *Opyt issledovaniia mirovozzreniia rannikh russkikh liberalov* (A study of early Russian liberals' world view). Moscow: Rossiiskaia akademiia nauk, Institut filosofii, 1995, pp. 149–-205.

George L. Kline (1967)
Bibliography updated by Vladimir Marchenkov (2005)

KELSEN, HANS
(1881–1973)

Born in Prague on October 11, 1881, Hans Kelsen grew up in Vienna. He studied law at the University of Vienna and completed, in 1911, the *Habilitation* (major dissertation required for the *venia legendi* or state license to hold university lectures). After military service in World War I, he worked up a number of drafts of what became the Austrian Federal Constitution of October 1920. Here Kelsen's most distinctive contribution was centralized constitutional review, an entirely new institutional practice. During the 1920s, Kelsen served as professor of law at the University of Vienna and also as Constitutional Court judge. Ousted from the latter position in 1930 by Austria's right-of-center Christian-Social Party, Kelsen took up a professorship in Cologne. Ousted from this position in the spring of 1933, on the basis of the notorious Nazi statute for the "Restoration of the Professional Civil Service" (authorizing the dismissal of those seen as politically unreliable and also those of Jewish ancestry), Kelsen spent the period from 1933 to 1940 in Geneva. He left in May 1940 for the United States, where he eventually secured a position at the University of California at Berkeley. He died in Berkeley on April 19, 1973.

Kelsen's juridico-philosophical work breaks down into three phases, although there is no bright line between the first two. Kelsen's first phase, critical constructivism, runs from 1911 to approximately 1920. His primary concern is to show that naturalism in legal science is mistaken, and he goes on to construct the basic concepts of the law in nonnaturalisticnon-naturalistic terms. Kelsen's second phase, his classical or Neoneo-Kantian period, picks up at the end of the first phase and runs up to 1960. It is marked by two major developments. The first of these is Kelsen's attempt to provide a foundation for the concepts he constructed in the first phase. His "purity postulate" precludes any appeal either to natural law or moral theory on the one hand, or to empirical data on the other. What remains? Kelsen answers with a transcendental argument, proceeding in standard Neoneo-Kantian fashion from the *Faktum der Wissenschaft* (here, the fact of legal science) to the necessity of the basic norm *qua* normative category. Without the normative category, legal science would not be possible, but since legal science is given, it must be the case that the normative category is presupposed.

A rather different development in the early years of the second phase is represented by Kelsen's adoption of the *Stufenbaulehre* (doctrine of hierarchical structure) from his gifted Vienna colleague, Adolf Julius Merkl. This doctrine calls for ever-greater concretization as the law moves from the general norms of the constitution, at the apex of the hierarchy, to individual legal acts of law—implementation at its base. Accommodating norms that represent every species of law (constitutional rule, statutory provision, administrative regulation, official's legal act), the doctrine gives the lie to later nineteenth-century *Gesetzespositivismus* (statutory positivism), which held that the statute alone was characteristic of the modern legal system. In a juridico-philosophical vein, the doctrine of hierarchical structure marks the introduction, into Kelsen's theory, of empowering norms, which, as he argues at a later point, represent the most fundamental normative modality.

In his third and last phase, beginning in 1960, Kelsen throws overboard the Neoneo-Kantian edifice of the classical phase and defends a will theory of law—a remarkable development in the case of a philosopher who, for literally half a century, had criticized the will theory as well-nigh wrong-headed. Kelsen's skepticism in this last phase is reflected, for example, in his rejection of any role for logic in the law.

Kelsen's significance stems not least of all from his work on the philosophically difficult concept of normativity. A "strong normativity thesis," defended as an interpretation of Kelsen by Joseph Raz, speaks to the classical question in legal philosophy, namely, whether—and, if so, how—the obligation to obey the law is to be justified. A "weak normativity thesis," which reflects Kelsen's abiding interest in preserving the autonomy of the law and, by the same token, the "purity" of legal science, looks to normativity in the name of noncausal change as Kelsen's juridico-philosophical alternative to naturalism.

See also Constructivism and Conventionalism; Legal Positivism; Natural Law; Neo-Kantianism; Philosophy of Law, History of.

Bibliography

Ebenstein, William. *The Pure Theory of Law*. Madison: University of Wisconsin Press, 1945.

Raz, Joseph., "Kelsen's Theory of the Basic Norm." In *Normativity and Norms: Critical Perspectives on Kelsenian Themes*, edited by Stanley L. Paulson and Bonnie Litschewski Paulson, pp. 47–67. Oxford: Clarendon Press, 1998.

WORKS BY KELSEN

Die philosophischen Grundlagen der Naturrechtslehre und des Rechtspositivismus. Charlottenburg, Germany: Rolf Heise, 1928. Translated by Wolfgang Herbert Kraus under the titleas "Natural Law Doctrine and Legal Positivism" as an appendix to Kelsen, *General Theory of Law and State*, 389–446.

Reine Rechtslehre. 1st ed. Leipzig, Germany: Deuticke, 1934. Translated by Bonnie Litschewski Paulson and Stanley L. Paulson as *Introduction to the Problems of Legal Theory*. Oxford: Clarendon Press, 1992.

General Theory of Law and State. Translated by Anders Wedberg. Cambridge, MA: Harvard University Press, 1945.

Reine Rechtslehre. 2nd ed. Vienna: Deuticke, 1960. Translated by Max Knight as *The Pure Theory of Law*. Berkeley: University of California Press, 1967.

Essays in Legal and Moral Philosophy, edited by Ota Weinberger; translated by Peter Heath. Dordrecht, Netherlands: Reidel, 1973.

Allgemeine Theorie der Normen, edited by Kurt Ringhofer and Robert Walter. Vienna: Manz, 1979. Translated by Michael Hartney as *General Theory of Norms*. Oxford: Clarendon Press, 1991.

Stanley L. Paulson (2005)

KEPLER, JOHANNES

(1571–1630)

Johannes Kepler, the founder of modern astronomy, was born in Weil der Stadt, near Stuttgart. During his life he was a student of theology, teacher of mathematics and astronomy, assistant to Tycho Brahe, imperial mathematicus to the emperors Rudolf II and Matthias, and astrologer to the duke of Wallenstein. His principal scientific discoveries were the three planetary laws named after him, the principle of continuity in geometry, and the Keplerian telescope. He was also responsible for decisive advances in the theory of optics and in work that led to the development of the infinitesimal calculus, and incidentally he coined a number of terms whose paternity has been forgotten, including *satellite* (for the moons of Jupiter), *dioptrics, focus* (of a conic section), and *camera obscura*.

SIGNIFICANCE OF KEPLER'S LAWS

Kepler's three laws of planetary motion postulate that the planets travel in elliptical orbits, one focus of each ellipse being occupied by the sun; that the radius vector connecting sun and planet sweeps over equal areas in equal times; and that the squares of the periods of revolution of any two planets are in the same ratio as the cubes of their mean distances from the sun.

The promulgation of the three laws was in several respects a turning point in the history of thought. They were the first "laws of nature" in the modern sense: precise, verifiable statements, expressed in mathematical terms, about universal relations governing particular phenomena. They put an end to the Aristotelian dogma of uniform motion in perfect circles, which had bedeviled cosmology for two millennia, and substituted for the Ptolemaic universe—a fictitious clockwork of wheels turning on wheels—a vision of material bodies not unlike Earth freely floating in space, moved by physical forces acting on them. Kepler's laws severed the ties between astronomy and theology and replaced the moving spirits of medieval cosmology by physical causation.

What has come to be called the Copernican revolution was in fact mainly the work of Kepler and Galileo Galilei. Kepler's laws and Galileo's studies on the motion of projectiles were the basic ingredients of the Newtonian synthesis. Nicolas Copernicus's *De Revolutionibus* was published in 1543, nearly thirty years before Kepler was born. Its first edition of a thousand copies never sold out, and it had altogether four reprintings in 400 years. By way of comparison, Christopher Clavius's textbook *The Treatise on the Sphere* had nineteen reprintings within fifty years; Copernicus's book had one. This curiosity is mentioned because it illustrates the fact that the Copernican theory attracted very little attention on the continent of Europe for more than fifty years—that is, for the next two generations. *De Revolutionibus* was an unreadable book describing an unworkable system. It revived the Pythagorean idea of a heliocentric universe, first proposed by Aristarchus of Samos in the third century BCE, but it adhered to the dogma of circular motion. As a result, Copernicus was forced to let the planets run on no less than forty-eight epicycles and eccentrics. He was in fact, as Kepler remarked, "interpreting Ptolemy rather than nature."

Kepler was the first astronomer to raise his voice in public in favor of the Copernican system. His *Mysterium*

Cosmographicum, published in 1597, fifty-four years after Copernicus's death, initiated the controversy; Galileo only entered the scene fifteen years later. At that time Kepler—aged twenty-six—knew little of astronomy. He had started as a theologian, but a chance opportunity made him accept the post of teacher of mathematics and astronomy at the provincial school of Gratz in Styria. Three years later, however, he became assistant to Tycho Brahe, whose observational data, of a hitherto unparalleled richness and precision, provided the empirical foundation for Kepler's efforts to determine the orbit of Mars. It took Kepler eight years of nerve-racking labor to succeed. The result was his magnum opus, published in 1609, which contains the first and second laws (the third came nine years later). It bears a provocative title:

A NEW ASTRONOMY *Based on Causation*
or A PHYSICS OF THE SKY
derived from Investigations of the
MOTIONS OF THE STAR MARS
Founded on Observations of
THE NOBLE TYCHO BRAHE.

The title is indeed symbolic of the work's revolutionary intent and achievement. Astronomy before Kepler had been a purely descriptive geometry of the skies, divorced from physical reality. Since the observed motions of the planets did not conform to the demands of circularity and uniformity, an increasing number of auxiliary wheels had to be added to the fictitious clockwork to save the phenomena. These wheels were thought to be somehow connected with the eight crystal spheres of medieval cosmology, which were kept in motion by a hierarchy of angels, but any pretense to regard them as a physically workable model had to be abandoned. The situation was summed up in a famous remark by Alfonso X of Castile, called the Wise, when he was initiated into the Ptolemaic system: "If the Lord Almighty had consulted me before embarking on the Creation, I should have recommended something simpler."

Copernicus upset the cosmic hierarchy by placing the sun in its center, but his universe was still cluttered (in John Milton's words) "with centric and eccentric scribbled o'er, Cycle and epicycle, orb in orb." It was Kepler who, by banishing epicycles and eccentrics "to the lumber-room" (as he wrote), finally demolished the very scaffolding, as it were, on which the medieval universe rested and replaced its hierarchy of spirit forces with the interplay of physical forces. The tortuous way in which he achieved this may serve as a cautionary tale to scientists and philosophers and represents a significant episode in the history of thought.

MYSTICISM AND EMPIRICISM

In Kepler all the contradictions of his age seem to have become incarnate—the age of transition from the medieval to the "new philosophy," as the scientific revolution was called by its founders. One half of his divided personality belonged to the past; he was a mystic, given to theological speculation, astrology, and number lore. However, he was also an empiricist with a scrupulous respect for observational data, who unhesitatingly threw out his earlier theory of planetary motions, the product of five years of dogged labor, because certain observed positions of Mars deviated from those that the theory demanded by a paltry eight-minute arc. He later wrote that Ptolemy and Copernicus had been able to shrug away such minor blemishes in their theories because their observations were accurate only within a margin of ten minutes, anyway, but those who, "by divine kindness," were in possession of the accurate observations of Brahe could no longer do so. "If I had believed that we could ignore those eight minutes," he wrote in the *Astronomia Nova* (II, Ch. 19), "I would have patched up my hypothesis accordingly. But since it was not permissible to ignore them, those eight minutes point the road to a complete reformation of astronomy."

This newfound respect for hard, obstinate facts was to transform what used to be called "natural philosophy" into the "exact" (or "experimental") sciences and to determine, to a large extent, the climate of European thought during the next three centuries. It provided Kepler with the necessary discipline and put a restraint on his exuberant fantasy, but the primary motivation of his researches was mysticism of a Pythagorean brand. Throughout his life he was obsessed by certain mystic convictions, each of which had the power of an idée fixe. The first was the belief that the solar system was patterned on the perfect, or "Pythagorean," solids (Saturn's orbit circumscribed a cube into which was inscribed the orbit of Jupiter; into this was inscribed the tetrahedron that circumscribed the orbit of Mars; and so on down to the octahedron inscribed into the orbit of Mercury). The second was the equally Pythagorean belief that the planetary motions were governed by musical harmonies (the book containing the third law is called *Harmonice Mundi*). Fortunately, both lent themselves to mathematical juggling almost ad lib, until they fitted the data. Far from interfering with his reasoning powers, these irrational obsessions were harnessed to his rational pursuits and provided the drive for his tireless labors. From a subjective point of view, Kepler's fundamental discoveries were in fact merely by-products of his chimerical quest. Toward the end of his

life he proudly mentioned in retrospect some of his minor achievements, but there is no mention whatsoever of his epoch-making first and second laws.

EMERGENCE OF THE CONCEPT OF FORCE

The apparent paradox of a mystically inspired prejudice acting as a spur to scientific achievement is most clearly exemplified in the circumstances that led Kepler to introduce into astronomy the concept of physical forces. As has already been stated, he started his career as a student of theology (at the Lutheran University of Tübingen). The reason the concept of a heliocentric universe attracted the young theologian was later stated by him repeatedly. Thus, in the "Preface to the Reader" of his *Mysterium Cosmographicum* he explained that he had often defended the opinions of Copernicus in the discussions of the candidates at the seminary and had also written "a careful disputation on the first motion which consists in the rotation of the earth around the sun *for physical, or if you prefer, metaphysical reasons*." (The last phrase is emphasized because it is repeated verbatim in various passages in Kepler's works.)

He then proceeded to explain the nature of these "metaphysical reasons." They were originally based on a supposed analogy between the stationary sun, the stars, and interstellar space, on the one hand, and God the Father, the Son, and the Holy Ghost, on the other. In his first book the young Kepler promised the reader to pursue this analogy in his future cosmographical work; twenty-five years later, when he was over fifty, he reaffirmed his belief in it. "It is by no means permissible to treat this analogy as an empty comparison; it must be considered by its Platonic form and archetypal quality as one of the primary causes" (*Mysterium Cosmographicum*, note to 2nd ed.).

He stuck to this belief to the end of his life, as he stuck to the Pythagorean solids and the harmony of the spheres. But gradually his cherished analogy underwent a significant change. The fixed stars were replaced by the moving stars—the planets. The sun in the center of the planets, "himself at rest and yet the source of motion," continued to represent God the Father, and "even as the Father creates through the Holy Ghost" so the sun "distributes his motive force through a medium which contains the moving bodies" (letter to Maestlin, March 10, 1595).

Thus, the Holy Ghost no longer merely fills the space between the motionless sun and the fixed stars. It has become an active agent, a *vis motrix* that drives the planets. Nobody before had suspected the existence of such a force emanating from the sun. Astronomy had been concerned not with the causes of the heavenly motions but with their description. The passages just quoted are the first intimation of the forthcoming synthesis of cosmology and physics. Once he conceived the idea, derived from his analogy, that the sun was the source of the power that makes the planets go round, Kepler hit upon a question no one else had asked before him: Why do the planets closer to the sun go round faster than those farther away? His first answer to it, in the *Mysterium Cosmographicum*, was that there exists only one "moving soul" in the center of all the orbits—that is, the sun—which drives the planets "the more vigorously" the closer they are, but by the time it reaches the outer planets the force is quasi exhausted "because of the long distance and the weakening of the force which it entails."

Twenty-five years later, in the notes to the second edition, he commented that if we substitute for the word *soul* the word *force*, "then we get just the principle which underlies my physics of the skies." He continued to explain that he had once firmly believed the motive force was a soul; yet as he reflected that the force diminishes in proportion to distance, just as light diminishes in proportion to distance, he came to the conclusion "that this force must be something substantial—'substantial' not in the literal sense but … in the same manner as we say that light is something substantial, meaning by this an unsubstantial entity emanating from a substantial body."

The twenty-five years that separate these two quotations mark the transition from *anima motrix* to *vis motrix*, from a universe animated by purposeful intelligences to one moved by inanimate, "blind" forces devoid of purpose. For the rest of his life Kepler struggled with this new concept emerging from the womb of animism (its very name, *virtus*, or *vis*, betrays its origin) without ever coming to terms with it. At first he was not aware of the difficulties inherent in it. In a letter to a friend, which he wrote when the *Astronomia Nova* was nearing completion, he outlined his program:

> My aim is to show that the heavenly machine is not a kind of divine, live being, but a kind of clockwork (and he who believes that a clock has a soul, attributes the maker's glory to the work), insofar as nearly all the manifold motions are caused by a most simple, magnetic, and material force, just as all motions of the clock are caused by a simple weight. And I also show how these physical causes are to be given numerical and

> geometrical expression. (Letter to Herwart, February 10, 1605)

Kepler had defined the essence of the scientific revolution. But it turned out to be easier to talk about a "most simple, magnetic, material force" than to form a concrete idea of its working. Kepler's efforts to visualize the nature of the "moving force" emanating from the sun are not only of exceptional interest from the historian's point of view; they also illuminate the philosophical difficulties that were inherent in the concept of "force" from its very beginning. Since no English translation of the *Astronomia Nova* was published by the time this article was written, a few quotations may be found in order. First, Kepler compared the "moving force" of the sun with the light emitted by it:

> Though the light of the sun cannot itself be the moving force … it may perhaps represent a kind of vehicle, or tool, that the moving force uses. But the following considerations seem to contradict this. First, the light is arrested in regions that lie in shade. If, then, the moving force were to use light as a vehicle, darkness would bring the planets to a standstill. …
>
> This kind of force, just like the kind of force that is light, … can be regarded not as something that expands into the space between its source and the movable body but as something that the movable body receives out of the space it occupies. … It is propagated through the universe … but it is nowhere received except where there is a movable body, such as a planet. The answer to this is: although the moving force has no substance, it is aimed at substance, i.e., at the planet-body to be moved. …
>
> Who, I ask, will pretend that light has substance? Yet nevertheless it acts and is acted upon in space, it is refracted and reflected, and it has quality, so that it may be dense or sparse and can be regarded as a plane where it is received by something capable of being lit up. For, as I said in my *Optics*, the same thing applies to light as to our moving force: it has no present existence in the space between the source and the object it lights up, although it has passed through that space in the past; it "is" not, it "was," so to speak. (*Astronomia Nova*, III, Ch. 33)

Thus, Kepler's gropings brought him closer to the modern concept of the field than to the Newtonian concept of force, and the modern scientist grappling with the paradoxes of quantum theory will find here an echo of his own perplexities. This may be the reason Kepler, having hit on the concept of universal gravity, subsequently discarded it—as Galileo and René Descartes were to discard it.

GRAVITY AND ANIMISM

The most precise pre-Newtonian formulations of gravity are to be found in the preface to the *Astronomia Nova*. Kepler started by refuting the Aristotelian doctrine according to which all "earthy" matter is heavy because it is its nature to strive toward the center of the world—that is, Earth. But all "fiery" matter strives by its nature toward the periphery of the universe and is therefore light. Kepler explained that there is no such thing as lightness, but, rather, the

> matter that is less dense, either by nature or through heat, is relatively lighter … and therefore less attracted [to the earth] than heavier matter. … Supposing the earth *were* in the center of the world, heavy bodies would be attracted to it, not because it is in the center, but because it is a material body. It follows that regardless of where we place the earth, heavenly bodies will always seek it. …
>
> Gravity is the mutual bodily tendency between cognate [i.e., material] bodies toward unity or contact (of which kind the magnetic force also is), so that the earth draws a stone much more than the stone draws the earth. …
>
> If the earth and the moon were not kept in their respective orbits by a spiritual or some equivalent force, the earth would ascend toward the moon 1/54 of the distance, and the moon would descend the remaining 53 parts of the interval, and thus they would unite. But this calculation presupposes that both bodies are of the same density.
>
> If the earth ceased to attract the waters of the sea, the seas would rise and flow into the moon. …
>
> If the attractive force of the moon reaches down to the earth, it follows that the attractive force of the earth, all the more, extends to the moon and even farther. …
>
> If two stones were placed anywhere in space near to each other, and outside the reach of force of a third cognate body, then they would come together, after the manner of magnetic bodies, at an intermediate point, each approaching the other in proportion to the other's mass.

In the same passage is to be found the first approximation to a correct theory of the tides, which Kepler explained as "a motion of the waters toward the regions where the moon stands in the zenith." In a work written at the same time—"Somnium—A Dream of the Moon" (an early exercise in science fiction)—he furthermore postulated that the sun's attraction, too, influences the tides—that is, that the gravitational force of the sun reaches as far as Earth.

But here we are faced with another paradox. In the preface to the *Astronomia Nova*, Kepler, as we have seen, had grasped the essence of gravity and even the idea that its force is proportionate to its mass; yet in the text of *Somnium*, and all subsequent works, he seems to have completely forgotten it. The force that emanates from the sun in the Keplerian universe is not a force of attraction but a tangential force, a kind of vortex or "raging current which tears all the planets, and perhaps all the celestial ether, from West to East."

To the question of what made Kepler drop gravity no answer is found anywhere in his profuse writings. Everything points to some unconscious psychological blockage, and we may gather hints about its nature in the writings of the other pioneers of the scientific revolution. Kepler's suggestion that the tides were caused by the moon's attraction Galileo indignantly rejected as an "occult fancy" (*Dialogue concerning the Two Chief World Systems*). Descartes was equally repelled by the idea of a nonmechanical force acting at a distance and, like Kepler, substituted for it vortices in the ether. As for Isaac Newton, his attitude is summed up in his famous third letter to Richard Bentley, in which he said it is inconceivable that "inanimate brute matter" should, without some mediating material substance, act upon other bodies.

> That gravity should be innate, inherent, and essential to matter, so that one body may act upon another, at a distance through a vacuum, without the mediation of anything else, by and through which their action and force may be conveyed from one to another, is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking, can ever fall into it.

Kepler, Galileo, and Descartes did not fall into the philosophical abyss; their thinking was much too "modern"—that is, mechanistic—for that. The notion of a "force" that acts without an intermediary agent and pulls at immense stellar objects with ubiquitous ghost fingers appeared to them mystical and unscientific, a lapse into that Aristotelian animism from which they had just broken loose. Universal gravity, *gravitatio mundi*, smacked of the *anima mundi* of the ancients. Newton overcame the obstacle and made the concept of gravity respectable by invoking a ubiquitous ether, whose attributes were equally paradoxical, and by refusing to speculate on the manner in which gravity worked (his *hypothesis non fingo* refers to this problem, and to this problem only, though it is often quoted out of context). But above all, he provided a precise mathematical formula for the mysterious agency to which gravity referred. That formula Newton deduced from the laws of Kepler, who had intuitively glimpsed universal gravity and shied away from it. In such crooked ways does the tree of science grow.

SYNTHESIS OF ASTRONOMY AND PHYSICS

In the Aristotelian cosmos, physical forces operated only among the four elements in the sublunary sphere; the motions of the celestial bodies, made of a fifth element, were due to spiritual agencies and governed by the demands of geometrical perfection. Kepler and Galileo broke down this dualism by postulating that physical causality permeates the entire universe. Kepler's "physics of the sky" we know to have been all wrong. He had no notion of inertial momentum, and he had dropped gravity. In Kepler's universe the sun exerted a tangential force (diminishing in direct ratio with increasing distance), which the "lazy" planets resisted, and the eccentricity of the orbits was accounted for by magnetic forces. (Since the planets' magnetic poles always pointed in the same direction, they would be drawn closer to the sun in the aphelion and repelled in the perihelion.)

But though the model was wrong in every detail, his basic assumption, that there were several antagonistic forces acting on the planets, guided him in the right direction. A single force, as previously assumed—the Prime Mover and the allied hierarchy of angels—would never produce elliptical orbits and periodic changes of velocity. These could only be the result of some tug of war going on in the sky, and this dynamic concept, supported by a series of wild ad hoc hypotheses, led him in the end, after countless detours, to his three laws.

Kepler's determination of the orbit of Mars became the unifying link between two hitherto separate universes of discourse, celestial geometry and earthly physics. His was the first serious attempt to explain the mechanism of the solar system in terms of physical forces. Once the example was set, astronomy and physics could never again be divorced.

See also Aristotelianism; Copernicus, Nicolas; Descartes, René; Force; Galileo Galilei; Geometry; Laws of Nature; Mass; Matter; Milton, John; Nature, Philosophical Ideas of; Newton, Isaac; Philosophy of Physics; Pythagoras and Pythagoreanism; Scientific Revolutions.

Bibliography

WORKS BY KEPLER

Joannis Kepleri Astronomi Opera Omnia. 8 vols., edited by C. Frisch. Frankfurt and Erlangen: Heyder and Zimmer, 1858–1871.

Johannes Kepler, Gesammelte Werke, edited by W. van Dyck, Max Caspar, and Franz Hammer. Munich: C.H. Beck, 1938–.

"Somnium, Sive Astronomia Lunaris" (Somnium—A Dream of the Moon). Translated by Patricia Frueh Kirkwood, in John Lear, *Kepler's Dream*. Berkeley: University of California Press, 1965.

WORKS ON KEPLER

Barker, Peter. "Constructing Copernicus." *Perspectives on Science* 10(2) (2002): 208–227.

Caspar, Max. *Johannes Kepler*. Stuttgart: Kohlhammer, 1948. Translated by C. D. Hellman. New York, 1962 (paperbound).

Chen Morris, Raz D. "Optics, Imagination, and the Construction of Scientific Observation in Kepler's New Science." *Monist* 84(4) (2001): 453–486.

Dear, Peter. *Revolutionizing the Sciences: European Knowledge and its Ambitions, 1500–1700*. Princeton, NJ: Princeton University Press, 2001.

Diederich, Werner. "The Structure of the Copernican Revolution." *Dialogos* 36(77) (2001): 7–24.

Franklin, Allan, and Colin Howson. "Newton and Kepler, a Bayesian Approach." *Studies in History and Philosophy of Science* 16 (1985): 379–385.

Holton, Gerald. *Thematic Origins of Scientific Thought: Kepler to Einstein, Revised Edition*. Cambridge, MA: Harvard University Press, 1988.

Jarine, N. *The Birth of History and Philosophy of Science: Kepler's 'A Defence of Tycho Against Ursus', with Essays on Its Provenance and Significance*. New York: Cambridge University Press, 1984.

Kiikeri, Mika. "Interrogative Reasoning and Discovery: A New Perspective on Kepler's Inquiry." *Philosophica* 63(1) (1999): 51–87.

Kleiner, Scott A. "A New Look at Kepler and Abductive Argument." *Studies in History and Philosophy of Science* 14 (1983): 279–314.

Koestler, A. "Kepler and the Psychology of Discovery." In *The Logic of Personal Knowledge, Essays Presented to Michael Polanyi*, edited by Marjorie Grene. London: Routledge and Paul, 1961.

Koestler, A. *The Sleepwalkers*. London: Hutchinson, 1959.

Kozhamthadam, Job. *The Discovery of Kepler's Laws: The Interaction of Science, Philosophy, and Religion*. Notre Dame: University of Notre Dame Press, 1994.

Kozhamthadam, Job. "Kepler and the Sacredness of Natural Science." *Philosophy in Science* 7 (1997): 9–36.

Martens, Rhonda. "Kepler's Solution to the Problem of a Realist Celestial Mechanics." *Studies in History and Philosophy of Science* 30a(3) (1999): 377–394.

Phalet, A. "On the Logic of Kepler's Evolving Models" In *Logic of Discovery and Logic of Discourse*, edited by Jaakko Hintikka. New York: Plenum Press, 1985.

A. Koestler (1967)
Bibliography updated by Tamra Frei (2005)

KEYNES, JOHN MAYNARD
(1883–1946)

The English economist John Maynard Keynes, the son of a distinguished Cambridge logician and economist, was one of the most brilliant and influential men of the twentieth century. His role as the architect and chief negotiator of Britain's external economic policies in two world wars was only one side of his public life. During his own lifetime, his economic views, contained primarily in two great works, *A Treatise on Money* (London, 1930) and *The General Theory of Employment, Interest and Money* (London, 1936), revolutionized the economic practice, and to a lesser extent, the economic theory, of Western governments.

Keynes wrote only one philosophical work, *A Treatise on Probability* (London, 1921), but it is a philosophical classic. The following account of the book's leading ideas adheres to its own main divisions.

PHILOSOPHY OF PROBABILITY

Keynes's philosophy of probability is contained chiefly in Parts I and II. For Keynes, only a proposition can be probable or improbable. A proposition has probability only in relation to some other proposition(s) taken as premise(s). Hence a proposition may have different probabilities on different premises. Nevertheless, the probability that *p* does have, given *q* (which Keynes writes as p/q, is perfectly objective. Some probabilities are known to us indirectly—for example, as a result of applying the theorems of the probability calculus; but first, of course, some probabilities must be known directly. Where a probability is known to us directly, it is known to us in the way that the validity of a syllogistic argument is known, whatever that way is. The probability relation is not an empirical one. If it is true that $p/q > r/s$, or that $p/q > 1/3$, or that $r/s = 1/2$, then it is true a priori, and not in virtue of any matter of fact. In particular, the truth of such statements is independent of the factual truth of *p, q, r*, and *s*. Finally, $p/q = 0$ if *p* is inconsistent with *q*, and $p/q = 1$ if *q* entails *p*.

Keynes's fundamental thesis, of which the above statements are developments, is that there are inferences in which the premises do not entail the conclusion but are nevertheless, just by themselves, objectively more or less good reason for believing it. This thesis seems to require the existence of different degrees of implication. Such degrees are Keynes's probabilities. Thus, for Keynes the study of probability coincides exactly with the study of inference, demonstrative and nondemonstrative. He developed, though somewhat obscurely, a general theory of inference in Chapter X. However, from the axioms and definitions from which he derived the accepted theorems of the probability calculus, he also derived many theorems of demonstrative inference, for example, "if $a/h = 0$ then $ab/h = 0$."

It would be hard to exaggerate the importance of Keynes's fundamental thesis. Classical probability theory of the eighteenth and nineteenth centuries must have presupposed some such thesis. Recent theory on degrees of confirmation presupposes it. To Keynes, as to Pierre Simon de Laplace and Rudolf Carnap, this thesis appeared to be necessary as a means of avoiding skepticism about induction. But David Hume would presumably have rejected it outright, and it is by no means free from difficulty.

There are two negative theses that distinguish Keynes's philosophy of probability from most earlier or later formulations. One is that probabilities simply do not have a numerical value, except in certain exceptional circumstances, and never in normal inductive contexts. The other is that there are noncomparable probabilities, that is, probabilities that are neither equal to nor greater nor less than one another. For obvious reasons, these theses have contributed to the neglect of Keynes by statistical writers.

INDUCTION

In Part III, Keynes discussed induction. The most important arguments of those that are rational but not conclusive belong to the class of inductions whose conclusions are universal generalizations and whose premises are about instances of the generalization.

Keynes, like John Stuart Mill, regarded all scientific induction as essentially eliminative induction. His account of the circumstances in which we regard an inductive argument as strong is, in essentials (although not otherwise), a development in detail of Mill's method of agreement.

The mere number of confirmations of a hypothesis in itself is of no evidential weight. The important thing is the variety of the instances, in respects other than those that constitute the instances' confirming ones. We regard inductions as being of greatest weight when the evidence approaches the ideal case in which the confirming instances are known to be not all alike in every respect. Various ways in which our evidence can fall short of this ideal are discussed in Chapter XIX. Keynes thought that the extent to which the evidence, by its variety, eliminates alternative hypotheses is the only important factor—not only when our hypothesis is empirical, but when it is, for instance, mathematical or metaphysical.

Keynes very clearly distinguished between the task of analyzing those inductive arguments that we regard as strong and the task of justifying the fact that we regard them as strong.

The latter task, he appears to have assumed, requires a proof of the proposition that relative to instantial evidence, the probability of a universal hypothesis can approach certainty as a limit. It will do so, he purported to prove, if (and one must assume only if) the probability of the instantial evidence supposing the hypothesis to be false can be made small in comparison with the probability of the hypothesis prior to the instantial evidence (its "a priori" probability). To reduce the former probability is the object of "varying the circumstances." The required disparity between the two probabilities will exist, Keynes argued, if (and one must assume only if), inter alia, the hypothesis has finite a priori probability. This requires that it be a member of a finite disjunction of exhaustive alternatives.

When the universal hypothesis is an empirical one, this amounts to the assumption that there exists in nature the materials for only a finite number of generalizations linking empirical properties. In other words, the number of the logically independent properties of empirical objects, which a priori might have been constantly conjoined, is finite. This is the famous principle of limited independent variety (Chapter XX). Hence, the fact that the probability of any empirical universal generalization should approach certainty as a limit requires the assumption of this principle. Or rather, Keynes thought, all that is required for this principle is finite a priori probability, since experience can and does noncircularly support the principle, provided it does have this initial probability.

It does so, Keynes appears to have argued, because we have a direct apprehension of the truth of the principle, just as, he thought, we have an apprehension (not independent of experience, yet not inductively inferred) of

the truth of the statement, "Color cannot exist without extension."

STATISTICAL INFERENCE

The main subject of Part V is those inductive inferences whose premises include a statement of the frequency of a property *B* in an observed series of *A*'s, and whose conclusions concern *B*'s frequency in the population of *A*'s as a whole, or in a further series of *A*'s, or the probability of the next *A* being a *B*.

The theory of statistical inference had been dominated by two methods of making such inferences, both due to Laplace. One is the "rule of succession," according to which the probability of the next *A* being *B* is

$$\frac{m+1}{m+n+2}$$

if m out of $m + n$ observed *A*'s have been *B*. The other is the "inversion" of the great-numbers theorem of Bernoulli. This theorem permits us—under an important restriction—to infer what frequency of *B* is most probable among observed *A*'s, given its frequency among *A*'s as a whole. Laplace purported to supply a theorem that would guide our inferences in the reverse, inductive direction, that is, from observed *A*'s to *A*'s as a whole.

Keynes regarded both methods as "mathematical charlatanry." His many criticisms of them cannot be weighed here. Apart from these criticisms, however, he considered it absurd to imagine that we could have exact measures of the probability of statistical conclusions. Statistical induction is subject to all the difficulties that beset inductions with universal conclusions, and to others beside. Moreover, the only evidence taken into account by all methods like Laplace's is numerical. The vital requirement of variety in the instances is neglected. In statistical contexts, the variety of the positive "instances" takes the form of the stability of the observed frequency when the observed series is considered as divided into subseries according to many different principles of division.

Keynes did think that, under a number of extremely stringent conditions, an inversion of Bernoulli's theorem is legitimate. But even to license these inductive inferences, as Keynes interpreted them, the principle of limited independent variety is required.

Bibliography

Keynes's only philosophical work is *A Treatise on Probability* (London: Macmillan, 1921). On Keynes's life, see R. F. Harrod, *The Life of John Maynard Keynes* (London: Macmillan, 1951).

Valuable critical material on Keynes's theories of probability and induction may be found in the following: Jean Nicod, *Foundations of Geometry and Induction* (London: Kegan Paul Trench and Trubner, 1930); F. P. Ramsey, *Foundations of Mathematics* (London: Kegan Paul Trench and Trubner, 1931); G. H. von Wright, *The Logical Problem of Induction* (Helsingfors, 1941); G. H. von Wright, *A Treatise on Induction and Probability* (London: Routledge and Kegan Paul, 1951); Bertrand Russell, *Human Knowledge, Its Scope and Limits* (London: Allen and Unwin, 1948); and Arthur Pap, *An Introduction to the Philosophy of Science* (New York: Free Press of Glencoe, 1962). Joan Robinson, *Economic Philosophy* (London: Watts, 1962), Ch. 4, is a highly readable brief account of Keynes's place in economic theory.

For another facet of Keynes's many-sided career, see *Essays and Sketches in Biography* (New York, 1956), a varied collection of Keynes's writings on economists, politicians, acquaintances, and himself.

OTHER RECOMMENDED WORKS

Davis, John Bryan. *Keynes's Philosophical Development*. New York: Cambridge University Press, 1994.

Keynes, John Maynard. *The Collected Writings of John Maynard Keynes*. London: Macmillan; New York: St. Martin's Press, for the Royal Economic Society, 1971–1989.

Keynes, John Maynard. *The John Maynard Keynes Papers, King's College, Cambridge*. Cambridge, U.K.: Chadwyck-Healey, 1993.

Moggridge, D. E. *Maynard Keynes: An Economist's Biography*. New York: Routledge, 1992.

O'Donnell, R. M. *Keynes: Philosophy, Economics, and Politics: The Philosophical Foundations of Keynes's Thought and Their Influence on His Economics and Politics*. New York: St. Martin's, 1989.

Skidelsky, Robert Jacob Alexander. *John Maynard Keynes: Vol. 1, Hopes Betrayed 1883–1920: A Biography*. London: Macmillan, 1983.

Wittgenstein, Ludwig. *Letters to Russell, Keynes, and Moore*. Oxford: Basil Blackwell, 1974.

D. Stove (1967)

Bibliography updated by Michael J. Farmer (2005)

KEYSERLING, HERMANN ALEXANDER, GRAF VON

(1880–1946)

Hermann Alexander, Graf von Keyserling, a German philosopher of life and man, was born in Könno, Estonia. He studied geology and other natural sciences at the universities of Dorpat, Geneva, Heidelberg, and Vienna. In 1902 Keyserling received his doctorate at Vienna, where, under the influence of Houston Stewart Chamberlain, he turned to philosophy. He spent the next few years in Paris, interrupting his stay, however, by several trips to England. In 1908, after two years in Berlin, Keyserling returned to Estonia to take over his ancestral estate at

Rayküll. He traveled frequently and in 1911 and 1912 took a trip around the world. The loss of his property after the Russian Revolution led to Keyserling's immigration to Germany. In 1920 he founded the School of Wisdom in Darmstadt. Further journeys to North and South America followed. The last years of his life were spent in the Austrian Tyrol.

Keyserling was not a systematic philosopher; instead, he presented brilliant observations, suggestive generalizations, and in vague outline, an image of man. To measure his work by traditional philosophy is to reject his view of the philosophic enterprise. Keyserling wanted to replace the traditional philosopher with the sage, to replace critical examination with immediate appreciation, and to replace the university with his School of Wisdom. He held that, instead of criticizing another position, one should try to empathize with it. His own *Travel Diary* furnishes an example of this approach. Keyserling reduced philosophy to an exercise with the thoughts of other ages and cultures in the hope that such play would lead the reader to an awareness of the spirit that underlies these thoughts. Truth, in the sense of adequacy to fact, was of little concern to Keyserling; intuitive appreciation alone counted. Keyserling used the word *polyphonic* to distinguish his thinking from "homophonic," traditional philosophy. Polyphonic thinking has no definite point of view and presents no definite theses. It is essentially rootless, an exercise with possibilities, designed to reveal a meaning that escapes all philosophic systems.

Keyserling's approach to philosophy bears witness to his understanding of man. Following Arthur Schopenhauer, Friedrich Nietzsche, Wilhelm Dilthey, Henri Bergson, and Eastern thought, he asserted the rights of life in the face of the modern overemphasis on the intellect. His insistence on the protean nature of man anticipated the existentialists' claim that existence precedes essence. Keyserling asked us to intuit, amid cultural and natural diversity, the spirit that finds only inadequate expression in each definite form. Those matters that are truly important cannot be thought clearly but can only be intuited. Critical philosophy was renounced; the philosopher had become an artist. The success of Keyserling's works, particularly of the *Travel Diary,* was symptomatic of the spiritual situation following World War I. Keyserling lent expression to the feeling that many of the traditional answers had become meaningless. But instead of deploring this spiritual homelessness, Keyserling made it a necessary condition of the full life: Ideally, man is a traveler.

See also Bergson, Henri; Chamberlain, Houston Stewart; Dilthey, Wilhelm; Essence and Existence; Nietzsche, Friedrich; Schopenhauer, Arthur.

Bibliography

WORKS BY KEYSERLING

Unsterblichkeit. Munich: J. F. Lehmanns, 1907. Translated as *Immortality.* London: Oxford University Press, 1938.

Das Reisetagebuch eines Philosophen, 2 vols. Darmstadt: O. Reichl, 1920. Translated as *The Travel Diary of a Philosopher,* 2 vols. New York: Harcourt, Brace, 1925.

Schöpferische Erkenntnis. Darmstadt: O. Reichl, 1922. Translated as *Creative Understanding.* New York: Harper, 1929.

Wiedergeburt. Darmstadt: O. Reichl, 1927. Translated as *The Recovery of Truth.* New York: Harper, 1929.

Das Buch vom Ursprung. Baden-Baden: Roland, 1947.

Reise durch die Zeit. Vaduz, Liechtenstein, 1948.

Kritik des Denkens. Innsbruck: Palme, 1948.

Die Gesammelten Werke. Darmstadt, 1956–.

WORKS ON KEYSERLING

Feldkeller, Paul. *Graf Keyserlings Erkenntnisweg zum Übersinnlichen.* Darmstadt: O. Reichl, 1922.

Noack, Hermann. "Sinn und Geist." *Zeitschrift für philosophische Forschung* (1953): 592–597.

Parks, Mercedes G. *Introduction to Keyserling.* London: J. Cape, 1934.

Röhr, Rudolf. *Graf Keyserlings magische Geschichtsphilosophie.* Leipzig: S. Hirzel, 1939.

Karsten Harries (1967)

KHOMIAKOV, ALEKSEI STEPANOVICH
(1804–1860)

Aleksei Stepanovich Khomiakov (1804–1860), was a Russian philosopher, theologian, poet, and writer, a founder of Slavophilism. Born into a wealthy Muscovite family of landed nobility, Khomiakov was educated in Moscow University. In his youth he took part in the Russo-Turkish War of 1828–1829. In his mature years, he preferred to live as a "private" gentleman in Moscow and on the family. He traveled abroad on two occasions: in 1825–1826 to Paris to study painting, and in 1847 to Germany and England. In the Russian social order he preferred the niche of an independent writer, poet, and playwright. Before his death, he revived The Society of the Lovers of Russian Literature (first founded at the beginning of the nineteenth century) at Moscow University, and served as its head. He died when he contracted cholera while treating peasants on his estate.

Khomiakov was a man of encyclopedic knowledge and diverse talents who brought his polemical style to bear on discussion in several fields in the humanities. Perhaps of greatest significance is his contribution to the philosophy of history. In his *Semiramida*, a three-volume work in the genre of universal history that he began writing in 1837 and continued writing to the end of his life, Khomiakov's goal is to explore the prehistory of nations. His conclusion is that culture as a whole is an expression of a higher spiritual principle—that is, religion. The vista of universal history represents the action upon humanity of cultural-religious archetypes, combined with ideas of freedom and necessity. There initially existed, according to Khomiakov, two types of nations: "conquering nations" and "agricultural nations":

> In accordance with their original character, conquering nations permanently preserve the sense of personal pride and contempt not only for those who are conquered but also for all those who are foreign … When they are victorious, they repress those they have enslaved and do not mix with them; when they are defeated, they stubbornly resist the influence of the victors and preserve in their souls instincts engendered in them by epochs of former glory … [By contrast] agricultural nations are closer to universally human principles. They have not been affected by the proud magic of victory … Because of this they are more receptive to all things that are foreign. They do not experience aristocratic contempt for other nations; instead, they feel sympathy for all that is human. (1900)

Universal history, Khomiakov believes, unfolds according to the laws of the conflict between two opposite spiritual principles. Khomiakov calls the "agricultural" principle "Iranism," and its opposite "Kushitism." The spiritual history of humanity is viewed as the battle between Iranism and Kushitism. Such a conception was not entirely novel: Friedrich Schlegel had divided humanity into two opposed races—the Cainites and the Sethites—and in Hegel's *Philosophy of History* the Iranian "principle of light" is opposed to the Egyptian "principle of mystery." What was new was that Khomiakov did not base this antinomy on the principle of "good-bad"; instead, he viewed Iranism and Kushitism as two equally necessary forces in history.

Further, Kushitism consists in analysis and rationalism, whereas Iranism tends toward a synthetic and integral reception of the world. Therefore, these two types of national psychology are equally natural. Based on necessity, Kushitism engenders the state as a community based on convention. All of the civilizations of Kushitism were remarkable for being based on powerful state structures: Egypt, Babylon, China, Southern India. In contrast, Iranism proclaims the natural union of people and therefore rarely takes the form of a powerful political state. Thus, Khomiakov affirms that the historical process tends toward "the inevitable triumph of the Kushite principle" and to a "gradual decline of Iranism." "Iranism … has always been reestablished," writes Khomiakov, "by the particular efforts of great minds, whereas Kushitism has crept into the historical process by the unceasing action of time and of the national masses." If it happens that in Iranism there is an admixture of Kushitism, the latter is inevitably victorious (we find this, for example, in the history of ancient Greece and ancient Rome): Spiritual freedom must be absolute, and any concession to necessity leads to the death of freedom. The appearance of Christianity was the critical point of history: Christ represented a heroic effort to oppose the Kushitism of the world. But Christ's victory did not signify the victory of Iranism: Kushitism "closed itself up into the logic of the philosophical schools" (1900). And Hegelianism, which Khomiakov rejected, became the triumph of Kushitism in the nineteenth century. The Slavs belong to the Iranian type; that is what defines their place in history.

In Khomiakov's opinion, humans possesses the ability to strive toward being, toward God; but to preserve this striving, a special state is necessary: "true faith," where the diversity of a person's spiritual powers are gathered into a living, ordered wholeness. From this point of view, faith—which is simultaneously knowledge and life ("life-knowledge")—plays a special role in one's life.

Khomiakov's central conception is *sobornost'* ("catholicity," integrity, inner fullness), which characterizes not only the Christian church but also the nature of humans, society, and the processes of cognition and creativity. *Sobornost* is the organizing metaphysical principle of all being; by the power of love it gathers diversity into a "free organic unity" (in this it is distinct from "collectivity"). It was Khomiakov who introduced the principle of *sobornost* into the Russian thought of the nineteenth century. He defines *sobornost* as "a free and organic unity, whose vital principle is the Divine grace of mutual love." (1900) The foundation of *sobornost* is *grace*, a notion Khomiakov derives from Metropolita Ilarion's eleventh-century "Sermon on Law and Grace." Khomiakov also insists that divine grace is likewise the foundation of the real church, which can only be known from within, through one's lived experience.

Khomiakov based his theological conception on personal *experience*; and therefore affiliation with the church essentially became a prerequisite for knowing reality in general. Thus, Khomiakov extends the doctrine of *sobornost* beyond theology to the entire domain of Russian culture. Khomiakov wrote that "Christianity—even with all its purity, with all its elevatedness over all human individuals—takes different forms for the Slav, for the Roman, and for the German" (1900). It often happens that the aggregate of national beliefs and convictions is reflected neither in "verbal monuments" nor in "monuments of stone," and can be understood "only by looking at the entire life of a people, at its total historical development." Khomiakov elaborated this broad conception in his theological works, which, for reasons of censorship, in his lifetime could only be published abroad.

Despite their apparently paradoxical nature, Khomiakov's theological ideas were expressed at times with astonishing simplicity: "The Church is one, for two Churches do not exist"; "For there is one God and one Church, and there is no conflict or disharmony in her"; "The Church is not an institution"; "To assert that the Church is an authority is blasphemy." One does not "belong" to the church the way one belongs to an organization. In the church, people live the way they live at home, in the bosom of their family, "humbly conscious of their weakness and subordinating the latter to the unanimous decision of the conscience of all in sobornost" (1900). And only this life in the church gives people freedom, which is the greatest good. In his letter "To the Serbians" (written just before his death), Khomiakov expressed his view on "the meaning and virtue of faith" as follows:

> They are in great error, those who think that it [faith] is limited to the mere fulfillment of rituals or even to the relations of man to God. No: faith permeates the entire being of a man and all of his relations to his neighbor. As if with invisible threads and roots, faith grasps and is intertwined into all of a man's feelings, convictions, and aspirations. Faith is like a better air, transforming the earthly principle in a man; or it is like a most perfect light, illuminating all the moral notions of a man and all of his opinions of other people and of the inner laws connecting him with them. Thus, faith is also a supreme social principle … (1900)

Taking as his point of departure artistic intuition and "life-knowledge," which he strove to reconcile with scientific knowledge, Khomiakov attempted to unite two apparently incompatible sources: early patristics and ideas of Western romanticism and Western nature-philosophy. The organic principle of the interpretation of spiritual phenomena is evident not only in his ecclesiology, but also in his secular philosophy, as well as in his political and economic essays. The organic principle served as the foundation of his preference for gradual social development and conservatism. With the help of this principle Khomiakov sought to harmonize the Slavophile worldview with philosophical romanticism, bringing together such distinct categories as "the integrity of spirit," "the fullness of perception," and "the "organic character of social development" (1900). This principle was also the source of his doctrine of *sobornost* and of the view of the church as the regulator of the entire life of the Orthodox Christian.

In Khomiakov's social philosophy the opposition between *sobornost* and collectivity appears as the antithesis between *obshchina* (organic peasant community) and *druzhina* (organized "commune"), between "true brotherhood" and "conventional agreement." In Khomiakov's opinion, Russian history and Orthodox spirituality have manifested instances of true brotherhood, exhibited in the Russian peasant *obshchina*, which Khomiakov clearly idealizes, seeing in it the closest approximation to the social ideal. Petrine reforms, Khomiakov believes, led to the assimilation of "alien" principles by the Russian nobility and this, in turn, resulted in a split between the educated society and common people. Thus, in Khomiakov's opinion, genuine folk culture in Russia could be created only by returning to original folk principles. Khomiakov devoted to this subject numerous articles that provoked a polemic both in Russia and in Europe in the 1840s. Russian thought began to assimilate Khomiakov's heritage only many years after his death; his true stature became clear only at the end of the nineteenth century, when his major works were published (although not fully), and a Russian religious philosophy began to take shape.

See also Philosophy of History.

Bibliography

WORKS BY KHOMIAKOV

Polnoye sobranie sochinenii [Complete works]. 8 vols., edited by P. I. Bartenev and D. A. Khomyakov. Moscow, 1900–1911.

Stikhotvoreniia i dramy [Poems and plays], edited by B. F. Yegorov. Leningrad, 1969.

Sochineniia [Works]. 2 vols. Moscow, 1994.

On Spiritual Unity: A Slavophile Reader. Translated and edited by Boris Jakim and Robert Bird. Hudson, NY: Lindisfarne, 1998. This volume includes a translation of Khomiakov's

"The Church Is One," translations of three essays with "Remarks by an Orthodox Christian concerning the Western Communions," and the third and fifth letters to William Palmer. This volume also contains an extensive bibliography of works by and about Khomiakov, as well as about the Slavophiles in general.

WORKS ABOUT KHOMIAKOV

Berdiaev, N. A. *A. S. Khomiakov*. Moscow, 1912. Excerpts from Chapters 3, 4, and 8 of this work have been translated in Jakim and Bird, *On Spiritual Unity*. Hudson, NY: Lindisfarne, 1998.

Gratieux, A. A. *A. S. Khomiakov et le movement slavophile*. 2 vols. Paris, 1939. Translated from the French as *A. S. Khomiakov and the Slavophile Movement*. 2 vols. Belmont, MA: 1982.

Zenkovsky, V. V. *Istoriia russkoi filosofii*. 2 vols. Paris, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*. 2 vols. (New York: Columbia University Press, 1953). See volume one of Kline, 180–205.

Christoff, P. K. *An Introduction to Nineteenth-Century Russian Slavophilism: A Study in Ideas*. Vol. 1, *A. S. Xomjakov*. Paris: Mouton, 1961.

O'Leary, P. P. *The Triune Church: A Study in the Ecclesiology of A. S. Khomiakov*. Freiburg: Universitätsverlag Freiburg Schweiz, 1982.

Khomiakovskii sbornik [Collection of essays on Khomiakov]. Vol. 1. Tomsk, 1998.

Koshelev, V. *Aleksei Stepanovich Khomiakov: Zhizneopisanie v dokumentakh, rassuzhdeniiakh i razyskaniiakh* [Aleksei Stepanovich Khomiakov: Biography in the form of documents, reflections, and inquiries]. Moscow: Novoe Literaturnoe Obozrenie, 2000.

Viacheslav Koshelev (2005)
Translated by Boris Jakim (2005)

KIERKEGAARD, SØREN AABYE
(1813–1855)

Søren Aabye Kierkegaard, the Danish philosopher and religious thinker, frequently considered the first important existentialist, was the youngest son of Mikaël Pederson Kierkegaard and Anne Sørensdatter Lund, born when his father was fifty-six years old and his mother was forty-four. His early childhood was spent in the close company of his father, who insisted on high standards of performance in Latin and Greek, inculcated an anxiety-ridden pietist devotion of a deeply emotional kind, and awakened his son's imagination by continually acting out stories and scenes. Kierkegaard thus felt early the demand that life should be at once intellectually satisfying, dramatic, and an arena for devotion. Confronted with the Hegelian system at the University of Copenhagen, he reacted strongly against it. It could not supply what he needed—"a truth which is true *for me,* to find *the idea for which I can live and die*" (*Journal,* August 1, 1835). Nor could contemporary Danish Lutheranism provide this. He ceased to practice his religion and embarked on a life of pleasure, spending heavily on food, drink, and clothes.

The melancholy that originated in his childhood continued to haunt him, however, and was increased by his father's confiding in him his own sense of guilt for having somehow sinned deeply against God. For Kierkegaard, the question of how a man can be rescued from despair was consequently intensified. He resolved to return to his studies and become a pastor. He finished his thesis *On the Concept of Irony* (1841) and preached his first sermon. He became engaged to the seventeen-year-old Regine Olsen. But as he became aware of the uniqueness of the vocation that he felt within himself, he found himself unable either to share his life with anyone else or to live out the conventional role of a Lutheran pastor. For him, breaking off his engagement was a decisive step in implementing his vocation. (This cosmic view of the breach does not appear to have been shared by his young fiancée, whose natural hurt pride and rejected affection led to her marriage to Fritz Schlegel, afterward governor of the Danish West Indies.) From then on Kierkegaard lived a withdrawn life as an author, although he did involve himself in two major public controversies. The first followed his denunciation of the low standards of the popular Copenhagen satirical paper *The Corsair. The Corsair* in turn caricatured Kierkegaard unmercifully. The second sprang from his contempt for the established Danish Lutheran Church, and especially for its primate, Bishop Mynster, who died in early 1854. When Mynster's about-to-be-appointed successor, Professor Hans Martensen, declared that Mynster had been "a witness to the truth," Kierkegaard delivered a series of bitter attacks on the church in the name of the incompatibility he saw between established ecclesiastical conformism and the inward and personal character of Christian faith. He died shortly after refusing to receive the sacrament from a pastor. "Pastors are royal officials; royal officials have nothing to do with Christianity."

Kierkegaard's biography is necessarily more relevant to his thought than is the case with most philosophers, for he himself saw philosophical inquiry neither as the construction of systems nor as the analysis of concepts, but as the expression of an individual existence. The epitaph that he composed for himself was simply, "That individual." From his own point of view, any verdict on his thought can only be the expression of the critic's own existence, not a critical assessment which could stand or

fall according to some objective, impersonal standard. Hence all attempts at an objective evaluation of his thought were condemned by him in advance. He predicted and feared that he would fall into the hands of the professors. Moreover, the initial difficulty created by Kierkegaard's subjectivism is compounded by his style and manner of composition. Although he attacked G. W. F. Hegel, he inherited a large part of Hegel's vocabulary. Passages of great and glittering brilliance tend to alternate with paragraphs of turgid jargon. Both types of writing often prove inimical to clarity of expression. A great many of his books were written for highly specific purposes, and there is no clear thread of development in them. One device of Kierkegaard's must be given special mention: He issued several of his books under pseudonyms and used different pseudonyms so that he could, under one name, ostensibly attack his own work already published under some other name. His reason for doing this was precisely to avoid giving the appearance of attempting to construct a single, consistent, systematic edifice of thought. Systematic thought, especially the Hegelian system, was one of his principal targets.

THE SYSTEM, THE INDIVIDUAL, AND CHOICE

In Hegel's philosophical system, or rather in his successive construction of systems, the linked development of freedom and of reason is a logical one. Out of the most basic and abstract of concepts, Being and Nothing, there is developed first the concept of Becoming and the various phases of Becoming in which the Absolute Idea realizes itself during the course of human history. Each phase of history is the expression of a conceptual scheme, in which the gradual articulation of the concepts leads to a realization of their inadequacies and contradictions, so that the scheme is replaced by another higher and more adequate one, until finally Absolute Knowledge emerges and the whole historical process is comprehended as a single logical unfolding. It is this comprehension itself that is the culmination of the process, and this point was effectively reached for Hegel in his own philosophy. Thus, in *The Science of Logic* he was able to write that he was setting out not merely his own thoughts, but the thoughts of God—the idea of God being simply an anticipation of the Hegelian conception of the Absolute.

In the Hegelian view, both moral and religious development are simply phases in this total process. In *The Phenomenology of Mind,* Hegel described the moral individualism of the eighteenth century, for example, in terms of a logical progress from the hedonistic project of a universal pursuit of private pleasure, through the romantic idealization of "the noble soul," to the Kantian scheme of duty and the categorical imperative, trying to show how each was brought into being by the contradiction developed by its predecessor. In terms of the Hegelian view, an individual is essentially a representative of his age. His personal and religious views must give expression to his role in the total moral and religious development of humankind—a role that is imposed upon him by his place in the historical scheme. He can at best express, but not transcend, his age.

For Kierkegaard, Hegel dissolved the concreteness of individual existence into abstractions characteristic of the realm of concepts. Any particular conceptual scheme represents not an actuality but a possibility. Whether a given individual realizes this possibility, and so endows it with existence, depends upon the individual and not upon the concepts. What the individual does depends not upon what he understands, but upon what he wills. Kierkegaard invokes both Aristotle and Immanuel Kant in support of his contention that Hegel illegitimately assimilated concepts to individual existence; he praises in particular the manner of Kant's refutation of the Ontological Argument. But Kierkegaard, in his doctrine of the primacy of the will, is, in fact, more reminiscent of Quintus Septimius Florens Tertullian or Blaise Pascal.

Kierkegaard buttressed his doctrine of the will with his view of the ultimacy of undetermined choice. He maintained that the individual constitutes himself as the individual he is through his choice of one mode of existence rather than another. Christianity is not a phase in the total development of man's religious and moral ideas; it is a matter of choosing to accept or to reject God's Word. But choice is not restricted to this supreme decision; it is the core of all human existence. The Hegelian view that human existence develops logically within and through conceptual schemes is not merely an intellectual error. It is an attempt to disguise the true facts, to cast off the responsibility for choice, and to find an alibi for one's choices. Moreover, speculative system building falsifies human existence in another way, for it suggests that although those who lived prior to the construction of the system may have had to make do with a partial and inadequate view of reality, the arrival of the final system provides an absolute viewpoint. But according to Kierkegaard, such a viewpoint must be an illusion. Human existence is irremediably finite; its standpoint is incorrigibly partial and limited. To suppose otherwise is to yield to a temptation to pride; it is to attempt to put oneself in the place of God.

This conclusion is only a special case of Kierkegaard's general doctrine that his intellectual opponents are guilty fundamentally not of fallacies and mistakes, but of moral inadequacy. That Kierkegaard should have thought this not only reflects his unfortunate personality; it was a necessary consequence of his doctrine of choice. Another necessary consequence was his mode of authorship. On his own grounds, he cannot hope to produce pure intellectual conviction in his readers; all that he can do is to confront them with choices. Hence he should not try to present a single position. This explains Kierkegaard's method of expounding incompatible points of view in different books and using different pseudonyms for works with different standpoints. The author must conceal himself; his approach must be indirect. As an individual, he must testify to his chosen truth. Yet, as an author he cannot conceal the act of choice. From these views, it is apparent that Kierkegaard used a special concept of choice.

The essence of the Kierkegaardian concept of choice is that it is criterionless. On Kierkegaard's view, if criteria determine what I choose, it is not I who make the choice; hence the choice must be undetermined. Suppose, however, that I do invoke criteria in order to make my choice. Then all that has happened is that I have chosen the criteria. And if in turn I try to justify my selection of criteria by an appeal to logically cogent considerations, then I have in turn chosen the criteria in the light of which these considerations appear logically cogent. First principles at least must be chosen without the aid of criteria, simply in virtue of the fact that they are first. Thus, logical principles, or relationships between concepts, can in no sense determine a person's intellectual positions; for it is his choices that determine the authority such principles have for him. Is man then not even limited by such principles as those that enjoin consistency and prohibit contradiction? Apparently not. For even paradox challenges the intellect in such a way as to be a possible object of choice. The paradoxes that Kierkegaard has in mind at this point in his argument are those posed by the demands of ethics and religion. He is prepared to concede that in fields such as mathematics the ordinary procedures of reason are legitimate. But there are no objective standards where human existence is involved.

THE AESTHETIC AND THE ETHICAL

In *Either/Or: A Fragment of Life* (1843), the doctrine of choice is put to work in relation to a distinction between two ways of life, the ethical and the aesthetic. The aesthetic point of view is that of a sophisticated and romantic hedonism. The enemies of the aesthetic standpoint are not only pain but also, and above all, boredom. As Kierkegaard wrote of the protagonist of aestheticism in *Purify Your Hearts!,* "See him in his season of pleasure: did he not crave for one pleasure after another, variety his watchword?" The protagonist tried to realize every possibility, and no possibility furnishes him with more than a momentary actuality. "Every mood, every thought, good or bad, cheerful or sad, you pursue to its utmost limit, yet in such a way that this comes to pass *in abstracto* rather than *in concreto*; in such a way that the pursuit itself is little more than a mood…." But just because boredom is always to be guarded against, so its threat is perpetual. In the end, the search for novelty leads to the threshold of despair.

By contrast, the ethical constitutes the sphere of duty, of universal rules, of unconditional demands and tasks. For the man in the ethical stage "the chief thing is, not whether one can count on one's fingers how many duties one has, but that a man has once felt the intensity of duty in such a way that the consciousness of it is for him the assurance of the eternal validity of his being" (*Either/Or,* II, p. 223). It is important to note how intensity of feeling enters into Kierkegaard's definition of the ethical stage. He thought that what his own age most notably lacked was passion; hence one must not be deceived by the Kantian overtones of his discussions of duty. Kierkegaard's categorical imperative is felt rather than reasoned. He is an heir of such romantics as the Schlegel brothers in his attitude toward feeling, just as he is the heir of Hegel in his mode of argument. Kierkegaard is a constant reminder of the fact that those who most loudly proclaim their own uniqueness are most likely to have derived their ideas from authors whom they consciously reject.

In *Either/Or* the argument between the ethical and the aesthetic is presented by two rival characters: an older man puts the case for the ethical, a younger for the aesthetic. The reader, as we should expect, is allegedly left to make his own choice. But is he? The description of the two alternatives seems heavily weighted in favor of the ethical. The difficulty is that Kierkegaard wished *both* to maintain that there could be no objective criterion for the decision between the two alternatives, *and* to show that the ethical was superior to the aesthetic. Indeed, one difference between the ethical and the aesthetic is that in the ethical stage the role of choice is acknowledged. Kierkegaard frames this criticism of the man who adheres to the aesthetic: "He has not chosen himself; like Narcissus he has fallen in love with himself. Such a situation has certainly ended not infrequently in suicide." Remarks like

this suggest that in fact Kierkegaard thinks that the aesthetic fails on its own terms; but if he were to admit this, his concept of interested choice would no longer apply at this critical point. In one passage Kierkegaard asserts that if one chooses with sufficient passion, the passion will correct whatever was wrong with the choice. Here his inconsistency is explicit. According to his doctrine of choice, there can be no criterion of "correct" or "incorrect," but according to the values of his submerged romanticism, the criterion of both choice and truth is intensity of feeling.

This inconsistency is not resolved; rather it is canonized in the thesis that truth is subjectivity. On the one hand Kierkegaard wants to define truth in terms of the way in which it is apprehended; on the other he wants to define it in terms of what it is that is apprehended. When inconsistency results, he is all too apt to christen this inconsistency "paradox" and treat its appearance as the crowning glory of his argument.

Kierkegaard is not consistent, however, even in his treatment of inconsistency. For he sometimes seems to imply that if the ethical is forced to its limits, contradiction results, and one is therefore forced to pass from the ethical to the religious. "As soon as sin enters the discussion, ethics fails … for repentance is the supreme expression of ethics, but as such contains the most profound ethical contradiction" (*Fear and Trembling,* p. 147, footnote). What is this but Hegelianism of the purest kind?

Kierkegaard describes the transition from the ethical to the religious differently at different periods. In *Either/Or* the ethical sometimes seems to include the religious. By the time the *Concluding Unscientific Postscript* (1846) was written, the religious seems to have absorbed the ethical. In *Fear and Trembling* (1843), the passage from the ethical to the religious is even more striking than that from the aesthetic to the ethical. One of the heroes of this transition is Abraham. In demanding from Abraham the sacrifice of Isaac, God demands something that, from the standpoint of the ethical, is absolutely forbidden, a transgression of duty. Abraham must make the leap to faith, accept the absurd. He must concur in a "suspension of the ethical." At such a point the individual has to make a criterionless choice. General and universal rules cannot aid him here; it is as an individual that he has to choose. According to Kierkegaard, however, there are certain key experiences on the margins of the ethical and the religious through which one may come to censure oneself as an individual. One such experience is the despair that Kierkegaard describes in *The Sickness unto Death*; another is the generalized fear and anxiety that is characterized in *The Concept of Dread* (1844). Despair and dread point in the same direction. The experience of each forces the individual to realize that he confronts a void and that he is, in fact, responsible for his own sick and sinful condition. In the state of despair he is brought to recognize that what he despairs of are not the contingent facts (such as the loss of a loved one) that he claims to be the objects of his despair; the individual despairs of himself, and to despair of oneself is to see oneself confronting an emptiness that cannot be filled by aesthetic pleasure or ethical rule-following. Moreover, it is in order to become conscious that one has brought oneself to this point. In analyzing despair, we recognize guilt; so too with dread. Kierkegaard contrasts the fear that has a specific and identifiable object with the dread that is objectless; or rather he identifies the fear that is a fear of nothing in particular as a fear of Nothing. (The reification of negatives into noun phrases is typically Hegelian.) In the experience of dread I become conscious of my bad will as something for which I am responsible, and yet which I did not originate. Original sin is seen as a doctrine deduced from the analysis of experience.

In these works of Kierkegaard it is plain that the existentialist philosophy of choice is in some danger of being submerged in the romantic philosophy of feeling. But the testimony of feeling serves as a propaedeutic to the encounter with Christianity.

CHRISTIANITY

Kierkegaard regarded his own central task as the explanation of what is involved in being a Christian. Apart from Christianity, the only religions he discusses are those of the Greeks and the Jews, and those only as a foil to Christianity. At first sight, Kierkegaard's doctrines of choice and of truth stand in an uneasy relationship to his allegiance to Christianity. For surely Christianity has always claimed to be objectively true, independently of anyone's subjective commitment, and Kierkegaard recognized this. "Not only does it [Christian revelation] express something which man has not given to himself, but something which would never have entered any man's mind even as a wisp or an idea, or under any other name one likes to give to it" (*Journal,* 1839).

If what we believe depends on the believer's own ultimate choice of rational criteria, then surely all beliefs have an equal moment, or rather equal lack of moment, for claiming objective truth. Kierkegaard, however, tried to evade this conclusion and continued to argue both that ultimate choice is criterionless and that one choice can be more correct than another.

Unfortunately, Kierkegaard never considered the issues raised by religions other than Christianity; for it would clarify our view of his position considerably if we could know what he would have said about an account of Islam or Buddhism that was logically parallel to his account of Christianity, in that it made their claims rest on a doctrine of ultimate choice. But the choices that Kierkegaard discusses are always those that might arise for an educated Dane of the nineteenth century. The foil to Christianity is not another religion, but secular philosophy.

This particular contrast is most fully elucidated in the *Philosophical Fragments* (1844), in which Kierkegaard begins from the paradox posed by Socrates in Plato's *Meno.* How can one come to know anything? For either one already knows what one is to come to know, or one does not. But in the former case, since one already knows, one cannot come to know; and in the latter case, how can one possibly recognize what one discovers as being the object of one's quest for knowledge? Plato's answer to this paradox is that in coming to know, we do not discover truths of which we had hitherto been totally ignorant, but truths of which we were once aware (when the soul preexisted the body), but which we had forgotten. These truths lie dormant within us, and to teach is to elicit such truths. So Socrates makes the slave boy in the *Meno* aware that he knows geometrical truths which he did not know that he knew.

Suppose, however, Kierkegaard asks, that the truth is not within us already. It will then be the case that we are strangers to the truth, to whom the truth must be brought from outside. It will follow that the moment at which we learn the truth and the teacher from whom we learn the truth will not stand in a merely accidental relationship to us. On the Socratic view, one may learn geometry from this teacher or that, but the question of the truth of a geometric theorem is independent of the question from whom we learned it. Not so, on Kierkegaard's view. There are two possible conceptions of the truth that we must choose between, and the Socratic view represents only one alternative. It is important to note that in the *Philosophical Fragments* (1844) Kierkegaard does not say, as he says elsewhere, that one view of the truth is appropriate in matters of geometrical truth, but another is appropriate in matters concerning moral and religious truth. He speaks of two alternative views of the truth, which apparently cover every kind of subject matter, although for the rest of the book he discusses only religion.

Following Kierkegaard's preferred view of the truth, if the truth is not within us, it must be brought to us by a teacher. The teacher must transform us from beings who do not know the truth to beings who are acquainted with it. It is impossible to conceive any greater transformation, and only God could bring it about. But how could God become the teacher of man? If He appeared as He is, the effect on man would be to overawe him so that he could not possibly learn what God has to teach. (Kierkegaard cites the story of the prince in the fairy tale who could not appear to the swine girl as a prince because she would not have come to love him for himself.) Thus, Kierkegaard argues that if God is to be the teacher of man, He must appear in the form of a man, and more specifically, in the form of a servant. From the standpoint of human reason, the idea that God should come as a teacher in human form is an impossible paradox that reason cannot hope to comprehend within its own categories. But according to Kierkegaard, it is in encountering this paradox that reason becomes aware of the objective character of what it encounters.

To be a Christian is thus to subordinate one's reason to the authority of a revelation that is given in paradoxical form. The Christian lives before God by faith alone. His awareness of God is always an awareness of his own infinite distance from God. Christianity initially manifests itself in outward forms, and Kierkegaard reproaches Martin Luther for having tried to reduce Christianity to a pure inwardness—a project that has ended in its opposite, the replacement of inwardness by an ecclesiastical worldliness. Nonetheless, an inward suffering before God is the heart of Christianity.

As previously mentioned, Kierkegaard saw his own age as lacking in passion. The Greeks and the medieval monastics had true passion. The modern age lacks it, and because of this, it lacks a capacity for paradox, which is the passion of thought.

CRITICISMS OF KIERKEGAARD

Kierkegaard used Friedrich Trendelenburg's exposition of Aristotle's logic to criticize Hegel. But he never took the question of the nature of contradiction seriously, and hence he never explained the difference, if any, between paradox (in his sense of the word) and mere inconsistency. But without such a clarification, the notion is fatally unclear. The lack of clarity is increased by Kierkegaard's failure at times to distinguish between philosophy, as such, and Hegelianism. Kierkegaard sometimes seems to have thought that any philosophy that

claims objectivity must consist solely of tautologies (*Papirer* III, B, 177).

His doctrine of choice raises at least two fundamental questions: Are there criterionless choices? And is it by such choices that we either can or do arrive at our criteria of true belief? Actual cases of criterionless choice usually seem in some way to be special cases. Either they are trivial, random selections (as of a ticket in a lottery) or they arise from conflicts of duties in which each alternative seems equally weighted. But none of these are choices of criteria. Such choices arise precisely at the point at which we are not presented with objective criteria. How do we arrive at such criteria? They appear to be internally connected with the subject matter of the relevant beliefs and judgment. Therefore we cannot choose our ultimate criteria in mathematics or physics. But what about morals and religion? Can one choose to consider the gratuitous infliction of pain a morally neutral activity? We are strongly inclined to say that an affirmative answer would indicate that the word *morally* had not been understood. But what is certain is that Kierkegaard's fundamental positions must remain doubtful until some series of questions such as this has been systematically considered. Kierkegaard himself never tried to ask them.

See also Absolute, The; Aristotle; Being; Existentialism; Hegel, Georg Wilhelm Friedrich; Hegelianism; Kant, Immanuel; Luther, Martin; Ontological Argument for the Existence of God; Pascal, Blaise; Schlegel, Friedrich von; Tertullian, Quintus Septimius Florens.

Bibliography

WORKS BY KIERKEGAARD

Texts

Papirer, 20 vols. Edited by P. A. Heiberg, V. Kuhr, and E. Torsting. Copenhagen: Gyldendal, 1909–1948.

Samlede Vaerker, 2nd ed., 14 vols. Edited by A. B. Drachmann, J. L. Heiberg, and H. O. Lange. Copenhagen, 1920–1931.

Texts in English Translation

The following listing is in order of original date of publication.

Either/Or, 2 vols. Vol. I, translated by D. F. Swenson and L. M. Swenson. Princeton, NJ, 1941. Vol. II, translated by W. Lowrie. Princeton, NJ, 1944.

Fear and Trembling. Translated by R. Payne. London: Oxford University Press, 1939. Also translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

Repetition: An Essay in Experimental Psychology. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

Philosophical Fragments: Or, A Fragment of Philosophy. Translated by D. F. Swenson. Princeton, NJ: Princeton University Press, 1936.

The Concept of Dread. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1944.

Stages on Life's Way. Translated by W. Lowrie. Princeton, NJ, 1940.

Concluding Unscientific Postscript. Translated by D. F. Swenson and W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

The Sickness unto Death. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

The Point of View. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

Training in Christianity. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1944.

Purify Your Hearts! Translated by A. S. Aldworth and W. S. Fine. London: C. W. Daniel, 1937.

For Self-Examination. Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1941.

The Present Age. Translated by A. Dru and W. Lowrie. London: Oxford University Press, 1940.

Christian Discourses. Translated by W. Lowrie. London: Oxford University Press, 1939.

Works of Love. Translated by D. F. Swenson. Princeton, NJ: Princeton University Press, 1946.

The Attack upon "Christendom." Translated by W. Lowrie. Princeton, NJ: Princeton University Press, 1944.

The Journals of Søren Kierkegaard: A Selection. Edited and translated by A. Dru. Oxford: Oxford University Press, 1938.

WORKS ON KIERKEGAARD

Anthologies, Biography, and Critical Studies

Bretall, R. *A Kierkegaard Anthology.* Princeton, NJ: Princeton University Press, 1946.

Geismar, E. O. *Lectures on the Religious Thought of Søren Kierkegaard.* Minneapolis: Augsburg, 1937.

Hohlenberg, J. E. *Søren Kierkegaard.* London, 1954.

Jolivet, R. *Introduction to Kierkegaard.* London, 1950.

Lowrie, W. *Kierkegaard.* New York: Oxford University Press, 1938.

Lowrie, W. *A Short Life of Kierkegaard.* Princeton, NJ: Princeton University Press, 1942.

Swenson, D. F. *Something about Kierkegaard.* Minneapolis: Augsburg, 1941.

Wahl, J. *Études Kierkegaardiennes.* Paris, 1938.

Additional Background

Barrett, W. *Irrational Man.* Garden City, NY: Doubleday, 1958.

Blackham, H. J. *Six Existentialist Thinkers.* New York, 1952.

Collins, J. *The Existentialists: A Critical Study.* Chicago: Regnery, 1952.

Grene, M. *Introduction to Existentialism.* Chicago: University of Chicago Press, 1959.

Shestov, L. *Athènes et Jérusalem.* Paris: J. Vrin, 1938.

Alasdair MacIntyre (1967)

KIERKEGAARD, SØREN AABYE [ADDENDUM]

Søren Aabye Kierkegaard has been the subject of sharply rising scholarly interest since the mid-twentieth century.

In addition to several important works devoted to reexamining Kierkegaard's relation to G. W. F. Hegel, and numerous specialized treatments of key themes and problems in the authorship, newer studies have explored the significance of Kierkegaard's thought from literary, political, and historical viewpoints.

Niels Thulstrup (1967) traces the development of Kierkegaard's critical engagement with Hegel from 1835 to the conclusion of the pseudonymous authorship in 1846. Thulstrup carefully delineates the main sources of Kierkegaard's knowledge of Hegelian philosophy. This is an invaluable service, considering that much of what Kierkegaard knew about the German philosopher was actually gleaned from secondary sources. Of special interest are the Danish Hegelians, Johan Ludvig Heiberg and Hans Lassen Martensen, and the anti-Hegelians, Frederik Christian Sibbern and Poul Martin Møller. Thulstrup also examines the influence of important German writers such as Johann Erdmann, Johann Gottlieb Fichte, Friedrich von Schelling, Adolf Trendelenburg, Marheinecke, and Werder. The notable tendency in this work to read Hegel through a Kierkegaardian lens leads the author to conclude that the two "have nothing in common as thinkers." This conclusion, however, has been challenged by other commentators who claim to find deeper parallels in their thought.

Several such parallels are noted by Mark C. Taylor (1980). Taylor points out, for instance, that both thinkers see the spiritlessness of modernity as the chief obstacle to selfhood and that both attempt to recover spirit through a process of "aesthetic education." For Hegel, however, spiritlessness represents a form of self-alienation that can be overcome only by a reconciliation of self and other, a mediation of the individual's personal and social life; while for Kierkegaard, the threat to spirit lies in the modern tendency to objectify and systematize, to dissolve the distinction between the individual and "the crowd." Taylor argues that Kierkegaard's exclusive emphasis on the individual is ultimately self-negating, since the self is never merely the self but bears a necessary and internal relation to the other. Hegel's relational conception of selfhood is thus shown to be more adequate and more comprehensive than Kierkegaard's, which "necessarily passes over into its opposite—Hegelian spirit" (p. 272). There remains a genuine question, however, about whether Kierkegaard's critique of "the crowd" precludes the possibility of a genuine human community in which individual responsibility is preserved.

Stephen N. Dunning (1985) goes even further than Taylor, suggesting that a relational conception of selfhood is implicit in the dialectical structure of Kierkegaard's writings. Dunning argues that the solitude of the self is "always a moment in a development that embraces interpersonal relations that can be contradictory (the aesthetic stage), reciprocal (the ethical stage), or paradoxically both incommunicable and reciprocal (the religious stage)" (pp. 248–49). According to this reading the *Postscript* describes a religious dialectic that culminates in a paradoxical unity of the self as both "other to itself (in sin) and restored to itself by God" (p. 249), and at the same time related to the entire community of Christians by a deep bond of sympathy. In this way, the theory of stages confirms the Hegelian insight that the solitary self is incomprehensible apart from the relational structures that give it meaning. It has been noted, however, that the formal similarities between Kierkegaardian and Hegelian dialectic may mask important conceptual differences noted by Thulstrup and Taylor.

Three studies of Kierkegaard's moral and religious philosophy deserve special mention. The first is Gregor Malantschuk's excellent study (1968). Working mainly from the journals, Malantschuk shows that the authorship is governed by a qualitative dialectic, which is aimed at illuminating the subjective dimensions of human existence, while the later polemical writings make use of a quantitative dialectic, which invokes the visible degradation of Christ as a judgment on Christendom. The dialectical method is thus seen to be the golden thread that runs through all of Kierkegaard's writings and places the individual works in the larger context of his avowed purpose as a religious author.

C. Stephen Evans's study of the *Fragments* and *Postscript* (1983) is widely recognized as one of the best general introductions to the Climacus writings available in any language. Though the book is written for the "ordinary" reader rather than the specialist—there is no critical engagement with the secondary literature—students and scholars alike have found it immensely useful for its coherent presentation of the main themes in Kierkegaard's religious philosophy, including his complex use of irony and humor in connection with the theory of indirect communication. The clarity of Evans's exposition is unsurpassed, even by his 1992 book, which returns to many of the issues addressed in the earlier work.

M. Jamie Ferreira (1991) explores one of the most difficult conceptual problems in the authorship: the nature of religious conversion. Challenging volitionalist and antivolitionalist accounts of the Kierkegaardian leap, Ferreira reconceptualizes the transition to faith as a "reorienting, transforming, shift in perspective" (p. 57).

Central to this account is the concept of surrender, which is explicated in terms of the imaginative activities of suspension and engagement. Based on this analysis, Ferreira offers a compelling refutation of the popular but mistaken assumption that Kierkegaard viewed ethical and religious choice as criterionless and hence immune to critical appraisal. Her analysis suggests rather that the more wholeheartedly one chooses, the more likely one is to discover whether one has made the wrong choice. On this reading passionate engagement is not meant to guarantee that one will continue in a choice no matter what, but it does ensure that one will experience more fully what is implied by a choice. In this way passionate engagement is seen to facilitate the possibility of critical appraisal.

Louis Mackey (1971) uses the tools of literary criticism to explore the complex relation between the literary and philosophical dimensions of Kierkegaard's authorship. Mackey argues that even the most philosophical of Kierkegaard's books, the *Fragments* and *Postscript*, call into question the very nature of the philosophical enterprise. His use of literary devices, intended to create a poetic indirection, always leave the reader somewhere between assertion and irony. Mackey goes on to make a more general point about the relation between philosophy and poetry, observing that "all humane philosophy is a poetic and for that reason an indirect communication" (p. 295). Indeed, the philosophers of Western tradition have in this sense, he claims, "always been poetic philosophers" (p. 295). This theme is developed further in Mackey (1986), which attempts to situate Kierkegaard in relation to current trends in deconstructionist thought and literary practice.

Bruce Kirmmse (1990) traces the political, economic, and social history of Denmark from 1780 to 1850, giving us a detailed picture of the cultural milieu in which Kierkegaard lived and wrote. Focusing on the boundaries between the public and the private, between politics and religion, Kirmmse lays a foundation for understanding the connection between Kierkegaard's critique of society and his attack on the established church. The exposition is facilitated by a discussion of Kierkegaard's important religious writings, which are frequently overlooked in major surveys of his thought. Until recently Kierkegaard's social and political views had received scant attention in the secondary literature. Other notable discussions can be found in chapters 8 and 9 of Alastair Hannay (1982) and in Merold Westphal (1987).

See also Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

Adorno, T. *Kierkegaard: Construction of the Aesthetic.* Minneapolis: University of Minneapolis Press, 1989.

Collins, J. *The Mind of Kierkegaard*. Chicago: Regnery, 1953.

Dunning, S. N. *Kierkegaard's Dialectic of Inwardness: A Structural Analysis of the Theory of Stages*. Princeton, NJ: Princeton University Press, 1985.

Evans, C. S. *Kierkegaard's "Fragments" and "Postscript": The Religious Philosophy of Johannes Climacus*. Atlantic Highlands, NJ: Humanities Press, 1983.

Evans, C. S. *Passionate Reason: Making Sense of Kierkegaard's Philosophical Fragments*. Bloomington: Indiana University Press, 1992.

Ferreira, M. J. *Transforming Vision: Imagination and Will in Kierkegaardian Faith*. Oxford: Clarendon Press, 1991.

Hannay, A. *Kierkegaard*. London: Routledge and Kegan Paul, 1982.

Kirmmse, B. H. *Kierkegaard in Golden-Age Denmark*. Bloomington: Indiana University Press, 1990. Contains an extensive scholarly bibliography.

Mackey, L. *Kierkegaard: A Kind of Poet*. Philadelphia: University of Pennsylvania Press, 1971.

Mackey, L. *Points of View: Readings of Kierkegaard*. Tallahassee: Florida State University Press, 1986.

Malantschuk, G. *Dialektik og Eksistens hos Søren Kierkegaard* (1968). Translated by H. V. Hong and E. H. Hong as *Kierkegaard's Thought*. Princeton, NJ: Princeton University Press, 1971.

Taylor, M. C. *Journeys to Selfhood: Hegel and Kierkegaard*. Berkeley: University of California Press, 1980.

Taylor, M. C. *Kierkegaard's Pseudonymous Authorship: A Study of Time and the Self*. Princeton, NJ: Princeton University Press, 1975.

Thulstrup, N. *Kierkegaards forhold til Hegel*. Copenhagen: Gyldendal, 1967. Translated by G. L. Stengren as *Kierkegaard's Relation to Hegel*. Princeton, NJ: Princeton University Press, 1980.

Westphal, K. *Becoming a Self: A Reading of Kierkegaard's "Concluding Unscientific Postscript."* West Lafayette, IN: Purdue University Press, 1996.

Westphal, K. *Kierkegaard's Critique of Reason and Society*. Macon, GA: Mercer University Press, 1987.

Steven M. Emmanuel (1996)
Bibliography updated by Thomas Nenon (2005)

KILVINGTON, RICHARD
(c. 1302–1361)

Richard Kilvington, Master of Arts (c. 1325) and Doctor of Theology (c. 1335) at Oxford, was a member of Richard de Bury's household, later becoming archdeacon and finally dean of Saint Paul's Cathedral in London. Along with Thomas Bradwardine, Kilvington formed the first academic generation of the school known as the "Oxford Calculators." All of Kilvington's philosophical works—*Sophismata* and *Quaestiones super De generatione*

et corruptione (written before 1325), *Quaestiones super Physicam* (c. 1326) and *Quaestiones super libros Ethicorum* (before 1332)—and his theological questions on Lombard's *Sentences* (c. 1334) stem from lectures at Oxford. In his *Physics*, Kilvington found an original way to apply the Euclidean theory of ratios to a new formula relating speeds, forces, and resistances in motions. Because the new rule avoided a serious weakness in Aristotle's theory of motion, nearly everyone adopted it, including the most famous Oxford Calculator, Thomas Bradwardine, in his renowned treatise on velocities in motions, written in 1328.

Following William of Ockham, Kilvington refuted the Aristotelian prohibition against *metabasis* and was convinced that mathematics is useful in all branches of scientific inquiry. He made broad use of the most popular fourteenth-century calculative techniques to solve physical, ethical, and theological problems. Four types of measurement are present in his works: by limits, that is, by the first and last instants of continuous processes, and by the intrinsic and extrinsic limits of capacities of passive and active potencies; by latitude or degree of forms, to measure intensive changes; by a calculus of compounding ratios, to determine speed of motion; and by one to one correspondence, to compare different infinities. Having adopted Ockham's position of ontological minimalism, Kilvington claimed that absolutes—that is, substances and qualities—are the only subjects that change and therefore all other terms, such as "motion," "time," "latitude," or "degree," are modes of speech. Accordingly, he contrasted things that are really distinct with things that are merely distinct rationally or in imagination. Because imaginable means possible—that is, not self-contradictory—in physics Kilvington discussed *secundum imaginationem* (according to imagination) counterfactual cases, such as the rectilinear motion of the earth or motion in a vacuum, and pondered questions that would never arise from direct observation, because the structure of nature can only be uncovered by highly abstract analysis.

Like many Oxford Calculators, Kilvington refrained from including God in the speculations of natural science. However, like almost everyone in the fourteenth-century, he distinguished between God's absolute power (*potentia Dei absoluta*) and ordained power (*potentia Dei ordinata*). The laws of nature reflect God's ordained power. Thanks to his absolute power and will, a presently active power, God might intervene to change or contradict the order of things that he had established. Therefore, it is possible for the past to have been otherwise, because all past events are contingent. Kilvington's teaching on logic, natural philosophy, and theology was markedly influential both in England and elsewhere in Europe. He inspired both the next generation of Oxford Calculators and important Parisian masters such as Nicolas Oresme.

See also Bradwardine, Thomas; Buridan, John; Oresme, Nicholas; William of Ockham.

Bibliography

WORKS BY KILVINGTON

The Sophismata of Richard Kilvington, edited by Norman Kretzmann and Barbara E. Kretzmann. Oxford: Oxford University Press, 1990.

The Sophismata of Richard Kilvington. Introduction, translation, and commentary by Norman Kretzmann and Barbara E. Kretzmann. Cambridge, U.K.: Cambridge University Press, 1990.

WORKS ON KILVINGTON

Jung-Palczewska, Elżbieta. "Works by Richard Kilvington." *Archives d'Histoire Doctrinale et Littéraire du Moyen Age* 67 (2000): 181–223.

Katz, Bernard D. "On a *Sophisma* of Richard Kilvington and a Problem of Analysis." *Medieval Philosophy and Theology* 5 (1996): 31–38.

Elżbieta Jung (2005)

KILWARDBY, ROBERT

(c. 1210–1279)

Robert Kilwardby was an English Dominican. He was a master of arts at the University of Paris between 1237 and 1245 and a student and master of theology at Blackfriars, Oxford, between 1248 and 1261. He then became prior provincial of the English Dominicans and in 1273 he was consecrated archbishop of Canterbury. In 1278 he entered the papal service as cardinal-bishop of Porto and Santa Rufina; he died in Viterbo in 1279.

Kilwardby had a profound influence on thirteenth- and fourteenth-century Scholasticism. In general he tried to promote the philosophical views of Augustine in a time when Aristotle's influence was becoming more and more important. As archbishop of Canterbury he even tried to suppress Aristotelian views by condemning thirty errors in philosophy in the so-called Oxford condemnation of 1277.

His most important and long-lasting influence, however, was in logic. During his Paris years he commented on the whole *Organon* of Aristotle, wrote two *Sophismata*

(*Sophismata grammaticalia* and *Sophismata logicalia*) and also several books on grammar. His commentary on *Priscianus minor* is the most important. During this incredibly productive time of his life he also wrote a commentary on Porphyry's *Isagoge*, and perhaps the earliest commentary on Aristotle's *Nicomachean Ethics*.

Very few of these works have been studied, and most of them still remain in manuscripts. The logical work that in recent years has received most attention is his commentary on Aristotle's *Prior Analytics*. As an exposition of Aristotle's theory of the syllogism, the commentary maintains an extraordinarily high degree of fidelity to Aristotle's text. As part of his overall project of constructing faithful interpretations of Aristotle, Kilwardby aims in his commentary to produce an accurate interpretation of Aristotle's modal syllogistic. The commentary is significant because it appears to be the origin, in the Latin world, of a tradition in which Aristotle's essentialist metaphysics is deployed in the interpretation of his syllogistic.

Kilwardby's work makes use of a number of technical concepts in a very disciplined way. These include notions of a *per se* term and a *per se* necessity and two concepts of *simpliciter* predication. The analysis of these concepts requires both the notion of an essential property and the notion of a necessary proposition. For example, a term is *per se* provided that it is necessary that, whatever it is, it is essentially that. *Per se* terms are contrasted with *per accidens* terms like *walking* for which nothing that is walking is essentially walking. Hence, a sentence like "Every *B* is necessarily *A*" expresses a *per se* necessity provided that (i) "*B*" is a *per se* term and (ii) "*A*" is a *per se* term and (iii) "Every *B* is *A*" is a necessary proposition.

The most important works from Kilwardby's tenure in Oxford are the *De ortu scientiarum* (1250), which is a classification of the sciences and was intended to be an introduction to philosophy, and his questions on the *Sentences* of Peter Lombard from around 1256. The *Sentence*-commentary is influenced by Richard Rufus of Cornwall. Kilwardby also produced smaller but very interesting treatises on relation, on time, and on imagination during this period.

See also Aristotelianism; Augustinianism; Logic, History of: Medieval (European) Logic; Rufus, Richard.

Bibliography

PRIMARY WORKS

De natura relationis (On relation, c. 1256–1261), edited by L. Schmücker, Brixen: L. Schmücker, 1980. A late work on relations.

De ortu scientiarum (On the origin of science, c. 1250), edited by A. G. Judy. Auctores Britannici Medii Aevi 4, London: British Academy, 1976. Kilwardby on natural philosophy.

De spiritu fantastico (On imagination, c. 1256–1261). *On Time and Imagination: De tempore, De spiritu fantastico*, edited by P. O. Lewry. Auctores Britannici Medii Aevi 9. Oxford: Oxford University Press for the British Academy, 1987. A late work on imagination.

De tempore (On time, c. 1256–1261). *On Time and Imagination: De tempore, De spiritu fantastico*, edited by P. O. Lewry. Auctores Britannici Medii Aevi 9. Oxford: Oxford University Press for the British Academy, 1987. A late work on time.

In donati artem maiorem III (Commentary on Donatus, c. 1237–1245), edited by L. Schmücker. Brixen: Typographia A. Weger Fund, 1984. Commentary on the grammatical work of Donatus.

In libros Priorum Analyticorum expositio (Exposition on the books of the *Prior Analytics*, c. 1240). Printed under the name Aegidius Romanus, Venice 1516. Reprinted Frankfurt 1968. Commentary on the *Prior Analytics*.

Notule libri prisciani de accidentibus (Commentary on *De accidentibus*, c. 1237–1245), edited by P. O. Lewry, "Thirteenth-Century Teaching on Speech and Accentuation: Robert Kilwardby's Commentary on *De accidentibus* of Pseudo-Priscian." *Medieval Studies* 50 (1988): 96–185. Commentary on Priscian.

Quaestiones in librum [primum/secundum/tertium/quartum] Sententiarum (Questions of the [first/second/third/fourth] book of the *Sentences*) [c. 1256]. *primum*, edited by J. Schneider; *secundum*, edited by G. Leibold; *tertium*, part 1, *Christologie* (Christology), edited by E. Gössmann, and part 2, *Tugendlehre* [Virtue], edited by G. Leibold; *quartum*, edited by R. Schenk. Veröffentlichungen der Kommission für die Herausgabe ungedruckter Texte aus der mittelalterlichen Geisteswelt. Munich: Verlag der Bayerischen Akademie der Wissenschaften, *primum*, 1986; *secundum*, 1992; *tertium*, 1982 (part 1), 1985 (part 2); *quartum*, 1993. Questions on the *Sentences* of Peter Lombard.

SECONDARY WORKS

Braakhuis, H. A. G. "Kilwardby verus Bacon? The Contribution to the Discussion on Univocal Signification of Beings and Non-Beings found in a Sophism attributed to Robert Kilwardby." In *Medieval Semantics and Metaphysics*, edited by E. P. Bos, 111–142. Artistarium supplementa 2, Nijmegen: Ingenium, 1985. An introduction to Kilwardby's theory of meaning.

Celano, A. J. "Robert Kilwardby and the Limits of Moral Science." In *Philosophy and the God of Abraham*, edited by R. James Long, 31–40. Toronto: Pontifical Institute of Mediaeval Studies, 1991. Study of Kilwardby's moral philosophy.

Ebbesen, S. "Albert (the Great?)'s Companion to the Organon." In *Albert der Grosse. Seine Zeit, sein Werk, seine Wirkung*, edited by A. Zimmerman, 89–103. Berlin: Walter de Gruyter, 1981. Ebbesen shows that Albert the Great copies Kilwardby's *Prior Analytics* commentary.

Gál, G. "Robert Kilwardby's Questions on the Metaphysics and Physics of Aristotle." *Franciscan Studies* 13 (1953): 7–28. On the attribution of a set of questions on metaphysics and physics.

Lagerlund, Henrik. *Modal Syllogistics in the Middle Ages*. Leiden: Brill, 2000. Chapter 2 contains a detailed study on Kilwardby's modal syllogistics.

Lewry, P. Osmund. "The Oxford Condemnation of 1277 in Grammar and Logic." In *English Logic and Semantics from the End of the Twelfth Century to the Time of Ockham and Burleigh*, edited by H. A. G. Braakhuis and L. M. de Rijk, 235–278. Artistarium supplementa 1, Nijmegen: Ingenium, 1981. An account of the implications for logic of the 1277 condemnation.

Lewry, P. Osmund. "Robert Kilwardby on Imagination: the Reconciliation of Aristotle and Augustine." *Medioevo* 9 (1983):1–42. Discussion of Kilwardby's view of the soul.

Lewry, P. Osmund. "Robert Kilwarby's Commentary on the Ethica nova and Vetus." In *L'homme et son univers au moyen-âge*, edited by C. Wedin, 799–807. Philosophes médiévaux 27, Louvain-la-Neuve: Editions de l'Institut Superieur de Philosophie, 1986. A summary of Kilwardby's views on ethics.

Lewry, P. Osmund. "Robert Kilwardby's Writings on the Logica Vetus Studied with Regards to Their Teaching and Method." Ph.D. diss. University of Oxford, 1978. Still the standard text for a discussion of Kilwardby's logic.

Henrik Lagerlund (2005)

KIM, JAEGWON

(1934–)

Jaegwon Kim is a Korean American philosopher born in Taegu (Korea) and educated at Seoul National University, Dartmouth College, and Princeton University. He has taught at Cornell University, University of Michigan, and Brown University, among other institutions. Kim's decisive contributions to philosophy range mainly over many central topics in the philosophy of mind and metaphysics but extend to philosophy of science and epistemology as well. Kim's most influential views in metaphysics and his early stance about the mind were defended in essays published from the early 1970s to the early 1990s and collected in the book *Supervenience and Mind* (1993). His later views on the mind are defended in two books: *Mind in a Physical World* (1998) and *Physicalism, or Something Near Enough* (2005).

In metaphysics, Kim's most crucial influence has been in event theory and the nature of dependence relations, including causation and supervenience. Kim's *property exemplification* account of events is regarded, together with Donald Davidson's account, as one of the two main contenders in the field. According to Kim (1993, essays 1 and 3) an event is not a basic component of ontology; it is a complex entity constituted by a property P (or a relation) exemplified by an object O (or n-tuple of objects) at a time t. If events are the *relata* of causal relations and causal relations require nomological connections (two widespread assumptions that Kim supports), Kim's fine-grained account of events has the advantage of indicating, in a causal relation, which feature of the *cause event* (its constitutive property) is nomologically connected with which feature of the *effect event* (its constitutive property).

Kim argues that just as causation (about which he is a regularist and a realist) constitutes the diachronic connection among phenomena, there are other metaphysically significant cementing relations that are noncausal (1993, essay 2). One of those relations is particularly important: supervenience, a synchronic dependence relation that connects properties in a given *supervenient* level with properties of a more basic level so that the most basic ones fully determine the supervenient ones. Kim is widely regarded as the leading theorist on supervenience, having carefully distinguished between several types of supervenience relations (e.g., weak, strong, and global) their consequences for reduction and for naturalist ontologies, having applied the notion to a general ontological stance he calls the *layered view of reality* and having used the concept to analyze perennial issues in the mind-body problem (1993, essays 4 to 10).

In the philosophy of mind Kim's work can be divided in three phases. In the early 1980s he defended a nonreductive naturalist/physicalist model of mental causation called supervenient causation. Given two mental properties M and M^* that supervene, respectively, upon physical properties P and P^*, if P causes P^*, M superveniently causes M^*. And if M supervenes on P and P causes P^*, M superveniently causes P^*. Supervenient causation is not outright causation but Kim claimed it was sufficient to endow mental properties with causal efficacy since these properties supervene on properties involved directly in causal processes. Supervenience plays here the double role of articulating the naturalist commitment and accounting for an acceptable (yet somewhat deflationary) approach to mental causation. The model is nonreductive because despite the causal powers of mental properties being reduced to those of their bases, the properties themselves are not reduced since supervenience does not imply identity.

In the late 1980s Kim produced several famous attacks against different forms of nonreductive physicalism (1993, essays 13 to 17; 1998, chapters 2 and 3). Against Davidson's anomalism, Kim argues that the view implies that the fact that an event falls under a mental kind is a causally irrelevant fact. Against functionalism, he claims that its multiple realizability thesis implies *local*

reductions and as such does not have the intended nonreductive force. More generally, he develops an argument against all forms of nonreductive physicalism called the causal/explanatory exclusion argument. For a physicalist every physical event has to have a complete physical cause. Kim shows by analyzing and ruling out scenarios that go from partial causes to causal overdetermination that within that framework, we cannot attribute a causal role to the mental unless it is identified with the physical, turning nonreducible mental properties into epiphenomena. Since he also defends the principle that without causal efficacy an entity cannot be real, every form of nonreductive physicalism turns into an eliminativist view. It soon became evident to Kim as much as to his critics that his supervenient causation model is also an easy target of the exclusion argument. Additionally, Kim has lost faith on the explanatory power of the supervenience relation in general, and in particular as a tool for analyzing mental causation. If supervenience is only a superficial relation of property covariation between the mental and the physical and it is itself in need of explanation, it cannot articulate a deep explanatory relation between the mental and the physical.

With this background Kim developed in the 1990s an approach to the mental that can be called functional reductionism (1998, 2005). The proposal consists of grounding the mind-body supervenience relation on the realization relation proposed by functionalism. Mental properties are second-order properties defined over a set of first-order properties that satisfy a given causal/functional condition and thus are eligible as realizers of such second-order properties. Given a mental property *M* we attempt to construct a functionalization of it in which *M* is characterized in terms of its typical causes and effects. This functionalization of a property is, Kim argues, sufficient for reduction (under a non-Nagelian, functional account of reduction). Reductive functionalization explains why there are the dependence relations there are and provides ontological simplification by identifying the second-order property with an exhaustive disjunction of all its realizers or else, according to Kim, we may decide to recognize only second-order *concepts* or *predicates* but not second-order properties. Still, Kim thinks that the qualitative properties of experience, unlike the rest of mental states, cannot be functionalized. Since, according to Kim, they are not reducible through type identification with neural-biological properties either, we have to accept them as a *mental residue* that prevents us from embracing a fully generalized physicalism.

Within philosophy of science, Kim's most significant contribution is a sophisticated view of what he calls the *metaphysics of explanation* that combines explanatory realism and pluralism (1989, 1994). According to realism, explanations are grounded in structural, *world-cementing* objective relations between the events referred to by the *explanandum* and the *explanans*. According to pluralism there are, in addition to causal explanations, explanations tied to noncausal, structural dependence relations (such as supervenience). This view can be seen to accord well with Kim's views regarding causal realism and the importance of noncausal relations, and explicitly includes the claim that pluralist realism explains via unification the cognitive value of explanations. In epistemology, Kim has produced an influential critique of Willard Van Orman Quine's naturalized epistemology (1993, essay 12). While defending epistemological naturalism in the sense that epistemic properties supervene upon factual, nonepistemic properties, he criticizes Quine's purely nomological, nonnormative approach to studying how evidence relates to beliefs. The gist of Kim's argument is that the very concept of knowledge disappears if we abandon the normative notion of justification.

See also Davidson, Donald; Epistemology; Metaphysics; Ontology; Philosophy of Mind; Philosophy of Science, Problems in; Quine, Willard Van Orman.

Bibliography

WORKS BY KIM

"Psychophysical Supervenience as a Mind-Body Theory." *Cognition and Brain Theory* 5 (1982): 129–147.

"Explanatory Realism, Causal Realism and Explanatory Exclusion." *Midwest Studies in Philosophy* 12 (1988): 225–239.

Supervenience and Mind. Cambridge, U.K.: Cambridge University Press, 1993.

"Explanatory Knowledge and Metaphysical Dependence." *Philosophical Issues* 5 (1994): 51–69.

Mind in a Physical World. Cambridge, MA: MIT Press, 1998.

Physicalism, or Something Near Enough. Princeton, NJ: Princeton University Press, 2005.

WORKS BY OTHERS

Block, Ned. "Antireductionism Slaps Back." *Philosophical Perspectives* 11 (1997): 107–132.

Fodor, Jerry. "Special Sciences: Still Autonomous After All These Years." *Philosophical Perspectives* 11 (1997): 149–163.

Horgan, Terence. "Kim on Mental Causation and Causal Exclusion." *Philosophical Perspectives* 11 (1997): 165–184.

Loewer, Barry. "Comments on Jaegwon Kim's *Mind in a Physical World*." *Philosophy and Phenomenological Research* 65 (2002): 655–662.

Sabatés, Marcelo. "*Mind* in a Physical World?" *Philosophy and Phenomenological Research* 65 (2002): 663–670

Marcelo H. Sabatés (2005)

KINDĪ, ABŪ-YŪSUF YAʿQŪB IBN ISḤĀQ AL-

See *al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq*

KING, MARTIN LUTHER

(1929–1968)

Martin Luther King Jr. was born in 1929 in Atlanta, Georgia. He attended Morehouse College, Crozer Theological Seminary, and Boston University, where he earned a doctorate in philosophical theology. In 1964, he was awarded the Nobel Peace Prize. He was assassinated in Memphis, Tennessee, in 1968.

King first gained international attention when, after completing his doctoral studies and becoming pastor of the Dexter Avenue Baptist Church in Montgomery, Alabama, he led the fight to desegregate public transportation in Montgomery. His strategy was nonviolent passive resistance. The faith that underlay that strategy was that white Americans could be persuaded by black suffering and moral argument to agree on the injustice of laws requiring the segregation of the races. The essentials of that faith are eloquently summarized in his frequently reprinted "Letter from Birmingham City Jail," and in his arguably most famous speech, "I Have a Dream." In that letter and speech King stressed his vision of the "beloved community," his vision of the "color-blind society," his conviction that injustice could be cured if exposed to the light of human conscience, and his conviction that every person has a duty to love one's enemies, and to avoid violence.

However, even in these works, King was not as optimistic or as completely reliant on white conscience as many have apparently thought him to be. For example, as his essay on civil disobedience reveals, his strategy of civil disobedience was designed not only to appeal to white conscience, but also to bring economic pressure on merchants. It is therefore a mistake to identify his theory with that of John Rawls, although Rawls himself stated that the two theories are similar.

King's more pessimistic or at least realistic views emerged more clearly in later speeches. Probably he was influenced by nationalists like Malcolm X and Stokely Carmichael (later known as Kwame Ture.) Certainly he admitted that he had started seeing his dream turning into a "nightmare," and that most Americans were "unconscious racists." Like Frederick Douglass before him, King concluded that moral suasion alone would not succeed in moving the white political establishment to implement the needed reforms, and that black people and their allies should therefore seek political power, though unlike Douglass he never advocated violence. In King's mature philosophy this new turn coincided with a greater emphasis on the poverty of many black Americans, and the relation of their plight to America's behavior in the international arena. King believed that the injustice of that behavior was being then revealed dramatically by the war in Vietnam and his criticisms of that war, together with his evidently growing sympathies for socialism lost him many allies. King's last speech, "I See the Promised Land," seems to contain premonitions of his assassination on the next day.

Unfortunately, as scholars of King's philosophy have noted, conservatives of the late twentieth and early twenty-first centuries have skillfully misused King's vision of a future color-blind society, especially his longing for a nation in which his four little children "will not be judged by the color of their skin but by the content of their character," to oppose color-conscious means like affirmative action for achieving such a nation.

See also Civil Disobedience; Justice; Pacifism; Racism; Rawls, John; Rights; Violence.

Bibliography

WORKS BY MARTIN LUTHER KING

Stride toward Freedom. New York: Harper and Row, 1958.

Strength to Love. New York: Harper and Row, 1963.

Why We Can't Wait. New York: Harper and Row, 1963.

Where Do We Go from Here: Chaos or Community? New York: Harper and Row, 1967.

The Trumpet of Conscience. New York: Harper and Row, 1968.

A Testament of Hope, edited by James Melvin Washington. New York: Harper and Row, 1986. An anthology of sermons and essays.

WORKS ABOUT MARTIN LUTHER KING

Ansbro, John J. *Martin Luther King, Jr.: The Making of a Mind*. New York: Orbis, 1982.

Cone, James H. *Martin and Malcolm and America: A Dream or a Nightmare*. New York: Orbis Books 1991.

Dyson, Michael Eric. *I May Not Get There with You: The True Martin Luther King Jr*. New York: Touchstone Books, 2000.

Garrow, David J. *Bearing the Cross: Martin Luther King Jr. and the Southern Christian Leadership Conference*. New York: Random House, 1986.

Lewis, David L. *King: A Critical Biography*. Baltimore, MD: Penguin, 1970.

Lincoln, C. Eric, ed. *Martin Luther King Jr.: A Profile*. New York: Hill and Wang, 1984.

Moses, Greg. *Revolution of Conscience: Martin Luther King, Jr., and the Philosophy of Nonviolence*. Foreword by Leonard Harris. New York: Guilford Press, 1997.

Bernard Boxill (2005)

KIREEVSKII, IVAN VASIL'EVICH

(1806–1856)

Ivan Vasil'evich Kireevskii, Russian literary critic and religious philosopher, was born in Moscow in a family of the old nobility related to the important poet Vasilii Zhukovskii (1783–1852). Kireevskii's father died in 1812 after contracting typhus in a hospital he founded for wounded soldiers. After his father's death, the young boy's education was largely guided by Zhukovskii, who did much for the development of Kireevskii's literary talent. Zhukovskii repeatedly affirmed, with total sincerity, that his young relative could become a fine writer. In 1823 Kireevskii became a member of the *Obshchestvo liubomudrov* (Society of the lovers of wisdom), organized by Dmitrii Venevitinov and Prince Vladimir Odoevskii for the study of German philosophy, especially Schelling. To complete his education, Kireevskii went abroad in 1830. In Germany he attended the lectures of Hegel, Schelling, and Schleiermacher. When he returned to Russia, he began to publish the journal *Evropeets* (The European, 1832), which was soon prohibited by the government. The orientation of the journal was somewhat "pro-Western": Kireevskii had set himself the task of synthesizing Western-European and Russian thought.

Kireevskii's further evolution was closely connected with Slavophilism. In 1845, for a period of time he was the editor of the Slavophile journal *Moskovitianin* (The Muscovite); and later he expounded his religio-philosophical ideas in the collection of articles *Moskovsii Sbornik* (Moscow collection, 1852), published by the Slavophile circle. In the final years of his life, Kireevskii was working on a course of philosophy in which his intent was to clearly display the distinguishing characteristics of the Russian philosophical tradition. The course was not completed. Kireevskii's collected works were first published in 1861, in two volumes.

The central idea of Kireevskii's philosophy was the integrity of the spirit: A human being can remain a person as long as he preserves in himself the unity of his "mind and heart," the "integrity" of his consciousness, of his "inner organization." Meanwhile, Kireevskii's epistemological theories were closely connected with his sociohistorical views. Only by attaining a harmonious "integral thinking" can the person and society avoid the two extremes: the ignorance that separates a nation from the "living communion of minds" and "abstract logical thinking" (rationalism) that fragments the integrity of the spirit into its separate elements (Kireevskii 1984, pp. 221–222).

Kireevskii tended to associate what he perceived as the limitations of Western society primarily with the one-sidedness of rationalism. He viewed Hegel as the final and supreme peak of Western rationalistic thought, continuing the tradition of Aristotle. In assessing various attempts to overcome rationalism in Europe (Schelling), Kireevskii considered that their failure was predetermined: Philosophy depends on the "character of the dominant faith," but in the Catholic-Protestant West the two dominant Christian faiths are, according to the Slavophile assessment, profoundly rationalistic. Kireevskii's own allegiance was to Orthodox theism, and he viewed the future "new" philosophy as a harmony of reason and Orthodox faith based on feeling.

Kireevskii thought that Western culture had already passed the highest point of its development and exhausted its potential. In his article "On the Nature of European Culture and on its Relationship to Russian Culture," Kireevskii writes that contemporary Western man "fragments his life into separate strivings or tendencies"; in "one corner of his heart there lives the religious sense; in another corner, separately, there live the powers of the intelligence and exertions related to everyday occupations; in a third corner there lives the desire for sensuous pleasures; in a fourth there lives moral feeling related to family life; and in a fifth there lives the desire for personal gain" (Kireevskii 1984, pp. 203, 229). That is, the souls of contemporary Westerners is mosaiclike, fragmented.

According to Kireevskii, such a transformation of human consciousness into a "calculating machine" will lead, in the final analysis, to the triumph of the lower desires, the instincts, where people will shut themselves up in their physical persons and desire only material comfort. It is precisely for this reason that Kireevskii began to seek the sources of a "new" and "young" philosophy, which was destined to supplant rationalism, overcome the fragmentedness of man's being, and lead to the "integral spirit." Kireevskii turned his glance toward the Russo-Slavic culture, in which, in his opinion, Orthodoxy was the principle that unified all spheres of life, combining spirit, reason, conscience, will, and feeling into a "thinking that believes." This thesis of Kireevskii's was, not without justification, called "epistemological utopi-

anism" by Vasilii Zenkovsky, the well-known historian of Russian philosophy.

Kireevskii attempted to answer the question of why the European and Russian cultures were separated as it were by an invisible wall. In doing so, he defined the sources of the European culture of his day. He identified three such principles or "elements": (1) the influence of the Christian religion; (2) the spirit of the barbarian nations that destroyed the Roman Empire; and (3) the remnants of the ancient world, of classical scholarship. Kireevskii analyzed these principles of Western civilization and arrived at the conclusion that the development of Russia lacked the classical heritage of the ancient world.

This "lack of the classical world" (Kireevskii 1984, p. 72) was, in his opinion, the reason why the influence of the Orthodox Church on Russia was not as strong as the influence of the Roman Catholic Church on the Western European countries, the Roman Church having experienced the enormous influence of the Roman government and Roman law on its organization. As a result, the Christian church in Russia could not become a force that would unite spiritually and politically fragmented Russia, which because of this fragmentation fell subject to the Tatar Yoke for several centuries. On the one hand, without a spiritual center (the kind of center that the Vatican was for fragmented Europe), Russia could be unified not spiritually but only materially (in other words, not in a spiritual but in a material *sobornost*), and this material unification took many centuries. On the other hand, the peculiar character of the development of Russia led to a situation where the Russo-Slavic world found itself separated and protected from Europe's deadening rationalism, the external and formal character of Europe's juridical law (inherited from Rome), and the coercive character of European governmental power, which was formed as a result of military conquests; moreover, in these circumstances, the Church in Russia had preserved its "purity," remaining independent of the governmental authority and secular goals.

Eastern Christianity, leading (as Kireevskii believed) from discursive rationalistic thinking to a free moral intuitive understanding, was assimilated by Russia in a form undistorted by the classical heritage. According to Kireevskii, the purity and undistorted character of its Christian principles are what give Russian culture a right to claim that it has a special role to play in the history of humankind. The "seed" (which is how he figuratively referred to the religious idea) has fallen onto a special "soil"—the Slavic national soul, which is characterized "both by dignity and by humility, attesting to equilibrium of spirit" (Kireevskii 1984, p. 224). But the main thing is that the Slavic "soil" is characterized by an original native principle of the organization of social life—the *obshchina* (or Russian commune). Not the personal right to property (as in the West) but the communal ownership of land is the foundation of the "relations of social life" in Russian society, for which individualism is a foreign principle. This is precisely why Kireevskii believed that the "new" philosophy and culture, so indispensable for humankind, could arise in his country. He associated the birth of this new thinking not with the construction of systems but with a radical transformation of the social consciousness, with the "education of society" as a result of common efforts rooted in *sobornost*. In this way, society will experience the infusion of a new philosophy that will overcome rationalism. This new philosophy will reorient humankind's spiritual life and produce in both society and in the individual an inner integrity of consciousness, a harmony of the social life.

By no means did this opposition between the Western fragmentedness and individualism and the Russian integrity and sobornost lead Kireevskii to reject the Western tradition. He dreamt of "integrity"; and here his ideal was the synthesis of what he considered the best features of the spiritual life of the West and of the East in such a manner that the "Russian principles," without nullifying European culture, would bestow upon the latter "higher meaning and definitive development" (Kireevskii 1984, p. 238). In the light of this, for Kireevskii the task of an original Russian philosophy would be the reworking of contemporary Western philosophy in the spirit of the teachings of Eastern patristics.

Kireevskii's views influenced a number of twentieth-century Russian philosophers, including Nikolai Berdiaev, Sergei Bulgakov, and Dmitrii Merezhkovskii. Kireevskii's central ideas—for example, about Orthodoxy as the foundation of Russian culture; the "conciliar" (*soborny*) nature of knowledge; and the fundamental difference between the European and Russian cultural traditions—have had a great impact and become the subject of close study by philosophers both in Russia and in Europe and North America.

See also Aristotle; Berdyaev, Nikolai Aleksandrovich; Bulgakov, Sergei Nikolaevich; Hegel, Georg Wilhelm Friedrich; Rationalism; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Schleiermacher, Friedrich Daniel Ernst; Zen'kovskii, Vasilii Vasil'evich.

Bibilography

WORKS BY KIREEVSKII

Izbrannye stat'i [Selected articles], edited by V. Kotelnikov, Moscow: Sovremennik, 1984.

On Spiritual Unity: A Slavophile Reader. Translated and edited by Boris Jakim and Robert Bird. Hudson, NY: Lindisfarne, 1998.

"On the Nature of European Culture and Its Relation to the Culture of Russia." In *Russian Intellectual History: An Anthology*, edited by Marc Raeff. Atlantic Highlands, NJ: Humanities Press, 1978.

Polnoe Sobranie Sochinenii [Complete works]. Edited by Mikhail Gershenzon. 2 vols. Moscow, 1911.

WORKS ABOUT KIREEVSKII

Christoff, P. K. *An Introduction to Nineteenth-Century Russian Slavophilism: A Study in Ideas*. Vol. 2, *I. V. Kireevskij*. The Hague: Mouton, 1972.

Gleason, Abbott. *European and Muscovite: Ivan Kireevsky and the Origin of Slavophilism*. Cambridge, MA: Harvard University Press, 1972.

Zenkovsky, V. V. *Istoriia russkoi filosofii*. 2 vols. Paris, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*. 2 vols. (New York: Columbia University Press, 1953). See volume 1, 207–227.

Olga Volkogonova (2005)
Translated by Boris Jakim

KITCHER, PATRICIA
(1948–)

Patricia Kitcher is widely known for her work on Kant and on philosophy of psychology. Born Patricia Williams, she attended Wellesley College and then graduate school in philosophy at Princeton where she studied with George Pitcher. Kitcher's interest in cognition manifested early and has continued to shape and inform her work throughout her career. Her doctoral dissertation defended a psychological continuity criterion for personal identity but extended the scope of the psychological criterion beyond that traditionally posited to include broader and more abstract cognitive characteristics, such as cognitive approach or cognitive style. Since then her work has ranged widely from traditional philosophy of psychology, to Freud, and ultimately to her greatest philosophical passion: Kant scholarship.

In her early work Kitcher wrote a number of papers in philosophy of psychology, philosophy of mind, and philosophy of science. She argued for the viability of intentional psychology and the autonomy of functionalist psychology from neurophysiology. Later work predominantly concentrated on analysis of problems stemming from the interpretation of Kant's first *Critique*. Kitcher has written numerous articles on the forms of intuition, Kant's epistemology, self-consciousness, and on how transcendental arguments work.

Kitcher has written two books that also pursue psychological themes. *Kant's Transcendental Psychology* was a radical departure from most Kant exegesis. The book makes two main claims about the *Critique of Pure Reason*. First, contra Peter Frederick Strawson, Kitcher argues that to understand synthetic a priori knowledge, it is essential to consider transcendental psychology. Second, she explicates a Kantian argument for the necessity of an integrated thinking subject, which serves as a reply to David Hume's denial of the unity of the self. An expanded and amended version of this position is being fleshed out more fully in a book she is currently writing, *Kant's Thinker*, which also explores the question of how we are to understand the faculties, and how the *Critique* contributes to debates about conscious and unconscious ideas.

In *Freud's Dream* Kitcher argued that Freud was the first cognitive scientist: Psychoanalysis should be thought of as an exercise in interdisciplinary theory construction, and as such, it illuminates the pitfalls to which such interdisciplinary approaches are subject. (Kitcher jokes that her arguments managed to alienate all readers: Freudians, because she exposes the mistaken foundation of psychoanalysis, and anti-Freudians, because she portrays his program as scientifically legitimate.)

Around the turn of the new century, Kitcher's interests turned toward Kantian ethics. Her works from this period provide an account of Kantian *maxims* and an interpretation of Kant's argument for the Formulation of the Universal Law for the Categorical Imperative, a task that has led many other Kant experts to throw up their hands in perplexity.

Kitcher's prodigious published contributions to philosophy are matched by her contributions to the philosophical community. She has served as department chair in three different universities, on numerous academic committees (including being a founding chair of the UC committee on the status of women), as president of the Society for Philosophy and Psychology, as president of the North American Kant Society, and on the editorial board of *Journal of Philosophy*. Her philosophical integrity, her fiery lectures, and her incisive comments on student papers make her an inspiring teacher and mentor.

Patricia Kitcher has held faculty positions at the University of Vermont, the University of Minnesota, and University of California, San Diego, and a visiting position at

University of Michigan. In 1998 she went to Columbia University where she became the Mark van Doren Professor of the Humanities and chair of the philosophy department. She lives in New York City with her husband, Philip, also a philosopher, with whom she has two sons, Andrew and Charles.

See also Ethics; Hume, David; Kant, Immanuel; Philosophy of Mind; Philosophy of Science, Problems of.

Selected Bibliography

WORKS

"The Crucial Relation in Personal Identity." *Canadian Journal of Philosophy* 7 (1) (1978): 131–145.

"Kant on Self-Identity." *Philosophical Review* 91 (1) (1982): 41–72.

"Kant's Paralogisms." *Philosophical Review* 91 (4) (1982): 515–547.

"In Defense of Intentional Psychology." *Journal of Philosophy* 81 (2) (1984): 89–106.

"Narrow Taxonomy and Wide Functionalism." *Philosophy of Science* 52 (1) (1985): 78–97.

"Discovering the Forms of Intuition." *Philosophical Review* 96 (1987): 205–248.

Kant's Transcendental Psychology. New York: Oxford University Press, 1990.

Freud's Dream: A Complete Interdisciplinary Science of Mind. Cambridge, MA: Bradford Books/M.I.T. Press, 1992.

"Revisiting Kant's Epistemology: Skepticism, Apriority, and Psychologism." *Noûs* 29 (3) (1995): 285–315.

"Kant on Self-Consciousness." *Philosophical Review* 108 (1999): 345–386.

"On Interpreting Kant's Thinker as Wittgenstein's 'I.'" *Philosophy and Phenomenological Research* 61 (2000): 33–63.

"Kant's Argument for the Categorical Imperative." *Noûs* 38 (4) (2004) 555–584.

Adina L. Roskies (2005)

KLAGES, LUDWIG

(1872–1956)

Ludwig Klages, a German psychologist and philosopher, was the leading figure in the field of characterology. Born in Hanover, Klages studied chemistry, physics, and philosophy at Munich, receiving his doctorate in chemistry in 1900. As a member of the Stefan George circle, he collaborated with George in the editing of the *Blättern für die Kunst.* In 1905 Klages founded at the University of Munich a *Seminar für Ausdruckskunde,* which soon became Germany's main center of characterological psychology. In 1919 the seminar was moved to Kilchberg, near Zürich, where Klages remained until his death.

Klages was the principal representative in psychology of the vitalist movement that swept Germany from 1895 to 1915. His most important work was directed toward the formulation of a science of character that would reestablish the undifferentiated union of the life forms that had been ruptured by the emergence of ego in the human species. To this end he explored some of the more bizarre pseudo sciences, such as graphology, and attempted to use their insights as the bases for auxiliary disciplines in his study of character types.

In addition to the literary influences of the romantic poets, of Johann Wolfgang von Goethe, and of Stefan George, Klages was also influenced by the physiologist E. G. Carus and the psychologist Theodore Lipps and, most important, by the philosopher Friedrich Nietzsche. All of these strands of thought converged in Klages to make of him a major spokesman of a generation of intellectuals consciously dedicated to the repudiation of reason in the name of instinct, and of civilization in the name of life. In short, his work was similar in content and general effect to that of Ernst Jünger, Oswald Spengler, and Martin Heidegger in providing—however unintentionally—an intellectual basis for Nazism.

According to Klages, Nietzsche had perceived correctly that man was distinguished from the rest of animal nature only by his ability to clothe in images the reality given by the senses. But Nietzsche had been wrong, Klages maintained, to regard this image-making ability as necessarily acting in the service of vital forces. In fact, he argued, man's ability to conceive a world in the imagination and to present this imagined world as a project or possible attainment against lived experience was unnatural and, in the end, profoundly hostile to life itself. Human life, for Klages, differed from animal life in general by virtue of the emergence in man of spirit (*Geist*); man's capacity to think and to will provided the source of his estrangement from the world and the cause of his peculiar psychic illnesses.

Animal life is possessed of both body (*Leib*) and soul (*Seele*), whose functions constitute "genuine processes." "The Body finds expression in the process of sensation and in the impulse towards movement, the Soul in the process of contemplation and in the impulse to formation (that is, to the magical *or* mechanical realization of images) … ." The processes of body and soul express the "eternal" life force, which is characterized by spontaneous creativity and flows beneath individual duration. In man, however, spirit appears, characterized by the "act of apprehension and the act of willing," which are in turn the origin of ego, utterly lacking in animals and impelling

man to the "unnatural" desire for immortality "or, more briefly, the urge to self-preservation."

This unnatural urge to self-preservation in man creates the tensions of human life. Man is a field whereon animal consciousness and human consciousness vie for supremacy. The former promotes the impulse to return to nature, expressed in the quest for "eternal life," while the latter promotes the life-destructive impulse to transcend the animal condition, reflected in science, religion, philosophy, and even art. The different quanta of soul and spirit present within an individual account for differences in character. Characterology, which is the study of these differences, constructs a typology of attitudes and structural forms as manifested in different egos. Most men live in the middle range of a spectrum of characterological types that runs from an almost total repression of spirit, as in primitive peoples, to an almost total repression of bodily forces, as in the asceticism of the redemptive religions. But in the science of character, Klages hoped, the true nature of the struggle between life and spirit raging in the individual would be clarified, the disastrous consequences of the triumph of spirit over life would be revealed, and science, art, and religion would be turned upon the spirit, destroy it, and lead to the dissolution of the individual ego in the undifferentiated nature out of which it had unnaturally emerged.

See also Carus, Carl Gustav; Goethe, Johann Wolfgang von; Heidegger, Martin; Jünger, Ernst; Lipps, Theodor; Nietzsche, Friedrich; Psychology; Spengler, Oswald; Vitalism.

Bibliography

Works by Klages include *Prinzipien der Charakterologie* (Leipzig, 1910), of which the 4th and subsequent editions are titled *Grundlagen der Charakterkunde* (11th ed., Bonn, 1951). *The Science of Character,* a translation of the 5th and 6th editions of this work, was prepared by W. H. Johnson (London: Allen and Unwin, 1929). Other writings are *Handschrift und Charakter* (Leipzig, 1917); *Die psychologischen Errungenschaften Nietzsches* (Leipzig: J. A. Barth, 1926); *Der Geist als Widersacher der Seele,* 3 vols. (Leipzig: J. A. Barth, 1929–1932); *Graphologie* (Heidelberg, 1931); *Geist und Leben* (Berlin, 1935); *Ursprünge der Seelenforschung* (Leipzig: P. Reclam, 1942) and *Die Sprache als Quelle der Seelen-Kunde* (Zürich, 1948).

For works on Klages see Max Bense, *Anti-Klages* (Berlin, 1937); K. Haeberlein, *Einführung in die Forschungsergebnisse von Klages* (Kampen, 1934); Herbert Hönel, ed., *Ludwig Klages: Erforscher und Künder. Festschrift zum 75. Geburtstage* (Linz: Österreichischer, 1947); Hans Kasdorff, *Um Seele und Geist: Ein Wegweiser zum Hauptwerk von Ludwig Klages* (Munich, 1954); Hans Prinzhorn, ed., *Die Wissenschaft am Scheidewege von Leben und Geist: Ludwig Klages zum 60. Geburtstag* (Leipzig, 1932); Ernest Sellière, *De la déesse nature à la déesse vie* (Paris, 1931); and Jean Toulemonde, *La caractérologie* (Paris: Payot, 1951).

Hayden V. White (1967)

KLEIST, HEINRICH VON

(1777–1811)

Heinrich von Kleist, a German dramatist, poet, and novelist, was born in Frankfurt on the Oder. Following a family tradition, Kleist entered the Prussian military service at fourteen, but he left, dissatisfied, in 1799. Uncertain what profession to adopt, Kleist prepared himself for the university by studying privately philosophy, mathematics, and classical languages. An intensive study of Immanuel Kant, or perhaps of Johann Gottlieb Fichte, led to a spiritual crisis in March 1801. The relativity of all knowledge seemed to Kleist to render life, especially a life dedicated to the pursuit of knowledge, pointless. In disgust he discontinued his studies and journeyed to Paris and Switzerland. His decision to pursue a literary career led to a second crisis: Afraid that he had no talent, he burned his tragedy *Robert Guiskard* in 1803. A period of restless activity followed. In 1805 he obtained a minor civil service position in Königsberg, which relieved him of his immediate worries. His two comedies, *Amphitryon* and *Der zerbrochene Krug,* were written at this time. Eager to aid the anti-Napoleonic cause he left Königsberg for Berlin, where in 1807 he was seized as a spy and sent to prison in France. After his sister had obtained his release, Kleist made an attempt to establish himself in Dresden from 1807 to 1809. With Adam Müller he founded the literary magazine *Phöbus,* which, however, soon failed. Attempts to help the patriotic cause with his literary efforts (*Hermannsschlacht,* 1808) met with little response. He returned to Berlin, where for a time he published the *Abendblätter.* When this project also failed, partly because of political pressure, Kleist was left without means. On November 21, 1811, Kleist committed suicide with Henriette Vogel near Berlin.

Kleist's reading of Kant taught him that all attempts to penetrate the veil of phenomena were futile, that the world possesses no higher meaning. In his first play, *Die Familie Schroffenstein* (1803), love, the only value, is destroyed by the force of illusion and circumstance—a theme that was to recur in such stories as *Die Verlobung in St. Domingo* and *Das Erdbeben in Chile.* Like G. W. F. Hegel, Kleist saw life as essentially tragic, but unlike Hegel, he saw tragedy in absurdity, in the indifference of the world to man's demands for love and meaning.

Kleist's heroes confront this absurdity with demonic defiance. Thus Michael Kohlhaas, in the novella of the same name (1810), becomes inhuman in his pursuit of justice; and the heroines of Kleist's plays *Penthesilea* (1808) and *Das Käthchen von Heilbronn* (1810) become inhuman in their pursuit of love—one by being totally aggressive, the other by being totally submissive. In his last play, *Der Prinz von Homburg* (1810), Kleist attempted to oppose the order provided by the state to the uncertainties of the human situation. The prince disobeys orders, wins a battle, and yet is condemned to death. At first incapable of understanding this judgment and driven only by his fear of death, he regains control of himself when made judge of his own actions, and freely accepts the verdict.

See also Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Love.

Bibliography

WORKS BY KLEIST

Werke, 5 vols. Edited by E. Schmidt. Leipzig, 1905; 2nd ed., 7 vols., edited by G. Minde-Pouet, Leipzig: Bibliographisches institut, 1936–.

Werke, 2 vols. Edited by H. Sembdner. Munich, 1961.

The Marquise of O, and Other Stories. Translated by Martin Greenberg. New York: Criterion, 1960.

TRANSLATIONS OF PLAYS

"The Feud of the Schroffensteins." Translated by M. J. and L. M. Price. *Poet Lore* (Boston) 27 (5) (1916): 457–576.

The Prince of Homburg. Translated by C. E. Passage. New York: Liberal Arts Press, 1956.

Katie of Heilbronn. Translated by A. H. Hughes. Hartford, CT: Trinity College, 1960.

The Broken Pitcher. Translated by B. Q. Morgan. Chapel Hill: University of North Carolina Press, 1961.

WORKS ON KLEIST

Blankenagel, J. C. *The Dramas of Heinrich von Kleist.* Chapel Hill: University of North Carolina Press, 1931.

Blöcker, Günter. *Heinrich von Kleist oder Das absolute Ich.* Berlin: Argon, 1960.

Cassirer, Ernst. *Idee und Gestalt.* Berlin: Cassirer, 1921.

Fricke, Gerhard. *Gefühl und Schicksal bei Heinrich von Kleist.* Berlin, 1929.

March, Richard. *Heinrich von Kleist.* Cambridge, U.K.: Bowes and Bowes, 1954.

Muth, Ludwig. *Kleist und Kant.* Cologne, 1954.

Silz, W. *Heinrich von Kleist.* Philadelphia: University of Pennsylvania Press, 1962.

Stahl, E. L. *Heinrich von Kleist's Dramas.* Oxford: Blackwell, 1948.

Witkop, Philipp. *Heinrich von Kleist.* Leipzig: Haessel, 1922.

Karsten Harries (1967)

KNOWLEDGE, A PRIORI

The prominence of the a priori within traditional epistemology is largely due to the influence of Immanuel Kant's *Critique of Pure Reason* (1965), where he introduces a conceptual framework that involves three distinctions: the epistemic distinction between a priori and empirical (or a posteriori) knowledge; the metaphysical distinction between necessary and contingent propositions; and the semantic distinction between analytic and synthetic propositions. Within this framework, Kant poses four questions:

1. What is a priori knowledge?
2. Is there a priori knowledge?
3. What is the relationship between the a priori and the necessary?
4. Is there synthetic a priori knowledge?

These questions remain at the center of the contemporary debate.

Kant maintains that a priori knowledge is "independent of experience," contrasting it with a posteriori knowledge, which has its "sources" in experience (1965, p. 43). He offers two criteria for a priori knowledge, necessity and strict universality, which he claims are inseparable from one another. Invoking the first, he argues that mathematical knowledge is a priori. Kant's claim that necessity is a criterion of the a priori entails:

(K1) All knowledge of necessary propositions is a priori.

He also appears to endorse

(K2) All propositions known a priori are necessary.

Kant maintains that all propositions of the form "All A are B" are either analytic or synthetic: analytic if the predicate is contained in the subject; synthetic if it is not. Utilizing this distinction, he argues that

(K3) All knowledge of analytic propositions is a priori; and

(K4) Some propositions known a priori are synthetic.

In support of (K4), Kant claims that the predicate terms of "7 + 5 = 12" and "The straight line between two points is the shortest" are not contained in their respective subjects.

THE CONCEPT

Kant provides the core of the traditional conception of the a priori. When he speaks of the source of knowledge, he does not mean the source of the belief in question, but the source of its justification. Hence, according to Kant,

(APK) S knows a priori that p if and only if S's belief that p is justified a priori and the other conditions on knowledge are satisfied; and

(APJ) S's belief that p is justified a priori if and only if S's justification for the belief that p does not depend on experience.

(APJ) has been criticized from two directions. First, some maintain that it is not sufficiently informative; it tells one what a priori justification is not, but not what it is. Hence, Laurence BonJour (1985) rejects (APJ) in favor of

(AP1) S's belief that p is justified a priori just in case S intuitively "sees" or apprehends that p is necessarily true.

Alvin Plantinga (1993) and BonJour (1998) offer variants of (AP1). Second, others maintain that the sense of *dependence* relevant to a priori justification requires articulation and offer two competing accounts. Albert Casullo (2003) endorses

(AP2) S's belief that p is justified a priori if and only if S's belief that p is nonexperientially justified (i.e., justified by some nonexperiential source).

Hilary Putnam (1983) and Philip Kitcher (1983) favor

(AP3) S's belief that p is justified a priori if and only if S's belief that p is nonexperientially justified and cannot be defeated by experience.

(AP1) and (AP3) face serious objections.

The term *see* is used metaphorically in (AP1). Let us assume that it shares with the literal use of *see* one basic feature: "S sees that p" entails "S believes that p." Hence, (AP1) has the consequence that if S's belief that p is justified a priori then S believes that p is necessarily true. This consequence faces two problems. Suppose that Sam is a mathematician who believes some generally accepted theorem T on the basis of a valid proof. Presumably, Sam's belief is justified. But suppose that Sam is also a serious student of philosophy who has come to doubt the cogency of the distinction between necessary and contingent propositions and, as a consequence, refrains from modal beliefs. It is implausible to maintain that Sam's belief that T is not justified a priori merely because of his views about a controversial metaphysical thesis. (AP1) is also threatened with a regress. It entails that if S's belief that p is justified a priori then S believes that necessarily p. Must S's belief that necessarily p be justified? If not, it is hard to see why it is a necessary condition of having an a priori justified belief that p. If so, then presumably it is justified a priori. But for S's belief that necessarily p to be justified a priori, S must believe that necessarily necessarily p, and the same question arises with respect to the latter belief. Must it be justified or not? Hence, (AP1) must either maintain that having an unjustified belief that necessarily p is a necessary condition of having a justified belief that p, or face an infinite regress of justified modal beliefs.

(AP3) is also open to serious objection. Saul Kripke (1980) and Kitcher (1983) maintain that an adequate conception of a priori knowledge should allow for the possibility that a person knows empirically some proposition that he or she can know a priori. (AP3) precludes this possibility. Assume that

(A) S knows empirically that p and S can know a priori that p.

From the left conjunct of (A), it follows that

(1) S's belief that p is justified$_k$ empirically,

where "justified$_k$" abbreviates "justified to the degree minimally sufficient for knowledge." Consider now the empirical sources that have been alleged to justify mathematical propositions empirically: counting objects, reading a textbook, consulting a mathematician, and computer results. (Tyler Burge [1993] discusses the relationship between testimony and a priori knowledge.) Each of these sources is fallible in an important respect. The justification each confers on a belief that p is defeasible by an empirically justified overriding defeater; that is, by an empirically justified belief that not-p. If S's belief that p is justified by counting a collection of objects and arriving at a particular result, then it is possible that S recounts the collection and arrives at a different result. If S's belief that p is justified by a textbook (or mathematician or computer result) that states that p, then it is possible that S encounters a different textbook (or mathematician or computer result) that states that not-p. In each case, the latter result is an empirically justified overriding defeater for S's original justification. Hence, given the fallible character of empirical justification, it follows that

(2) S's empirical justification for the belief that p is defeasible by an empirically justified belief that not-p.

(2), however, entails that

(3) S's belief that not-p is justifiable_d empirically,

where "justifiable_d" abbreviates "justifiable to the degree minimally sufficient to defeat S's justified_k belief that p." Furthermore, the conjunction of (AP3) and the right conjunct of (A) entails

(4) It is not the case that S's nonexperiential justification_k for the belief that p is defeasible by S's empirically justified belief that not-p.

(4), however, entails that

(5) It is not the case that S's belief that not-p is justifiable_d empirically.

The conjunction of (3) and (5) is a contradiction. Hence, (AP3) is incompatible with (A). (AP2), however, is compatible with (A) since the conjunction of (AP2) and the right conjunct of (A) does not entail (4).

SUPPORTING ARGUMENTS

Kant offers the most influential traditional argument for the existence of a priori knowledge. He holds that necessity is a criterion of the a priori: "[I]f we have a proposition which in being thought is thought as *necessary*, it is an a priori judgment" (1965, p. 43). He then argues that "mathematical propositions, strictly so called, are always judgments a priori, not empirical; because they carry with them necessity, which cannot be derived from experience" (p. 52). Kant's argument can be presented as follows:

(K1) All knowledge of necessary propositions is a priori.

(K2) Mathematical propositions are necessary.

(K3) Therefore, knowledge of mathematical propositions is a priori.

Premise (K1) is ambiguous. There are two ways of reading it:

(K1T) All knowledge of the truth value of necessary propositions is a priori; or

(K1G) All knowledge of the general modal status of necessary propositions is a priori.

Kant supports (K1) with the observation that "[e]xperience teaches us that a thing is so and so, but not that it cannot be otherwise" (1965, p. 52). This observation supports (K1G) but not (K1T), since Kant allows that experience can provide evidence that something is the case, but denies that it can provide evidence that something must be the case. The conclusion of the argument, however, is that knowledge of the truth value of mathematical propositions, such as that 7 + 5 = 12, is a priori.

Kant's argument can now be articulated as follows:

(K1G) All knowledge of the general modal status of necessary propositions is a priori.

(K2) Mathematical propositions are necessary.

(K3T) Therefore, knowledge of the truth value of mathematical propositions is a priori.

The argument involves this assumption:

(KA) If the general modal status of p is knowable only a priori, then the truth value of p is knowable only a priori.

(KA), however, is false. If one can know only a priori that a proposition is necessary, then one can know only a priori that a proposition is contingent. The evidence relevant to determining the latter is the same as that relevant to determining the former. For example, if I determine that "2 + 2 = 4" is necessary by trying to conceive of its falsehood and failing, I determine that "Kant is a philosopher" is contingent by trying to conceive of its falsehood and succeeding. However, if my knowledge that "Kant is a philosopher" is contingent is a priori, it does not follow that my knowledge that "Kant is a philosopher" is true is a priori. Clearly, it is a posteriori.

Roderick Chisholm (1977) suggests the following reformulation of Kant's argument:

(K1G) All knowledge of the general modal status of necessary propositions is a priori.

(K2) Mathematical propositions are necessary.

(K3G) Therefore, knowledge of the general modal status of mathematical propositions is a priori.

This argument faces a different problem. Why accept Kant's claim that experience can teach one only what is the case? A good deal of one's ordinary practical knowledge and the bulk of one's scientific knowledge provide clear counterexamples to the claim. My knowledge that my pen will fall if I drop it does not provide information about what is the case for the antecedent is contrary to fact. Scientific laws are not mere descriptions of the actual world. They support counterfactual conditionals and, hence, provide information beyond what is true of the actual world. In the absence of further support, Kant's claim should be rejected.

A second strategy for defending the existence of a priori knowledge is offered by proponents of logical empiricism, such as Alfred Jules Ayer (1952) and Carl Hempel (1972), who reject John Stuart Mill's contention that knowledge of basic mathematical propositions, such as that $2 \times 5 = 10$, is based on induction from observed cases. Both draw attention to the fact that if one is justified in believing that some general proposition is true on the basis of experience, then contrary experiences should justify one in believing that the proposition is false. But no experiences would justify one in believing that a mathematical proposition, such as that $2 \times 5 = 10$, is false. Suppose, for example, that I count what appear to be five pairs of shoes and arrive at the result that there are only nine shoes. Ayer contends that

> [o]ne would say that I was wrong in supposing that there were five pairs of objects to start with, or that one of the objects had been taken away while I was counting, or that two of them had coalesced, or that I had counted wrongly. One would adopt as an explanation whatever empirical hypothesis fitted in best with the accredited facts. The one explanation which would in no circumstances be adopted is that ten is not always the product of two and five. (1952, pp. 75–76)

Since Ayer maintains that one would not regard any experiences as evidence that a mathematical proposition is false, he concludes that no experiences provide evidence that they are true.

Ayer's argument can be stated as follows:

(A1) No experiences provide evidence that mathematical propositions are false.

(A2) If no experiences provide evidence that mathematical propositions are false, then no experiences provide evidence that they are true.

(A3) Therefore, no experiences provide evidence that mathematical propositions are true.

Ayer's defense of (A1) is weak in several respects. First, it does not take into account the number of apparent confirming instances of the proposition in question. Second, it involves only a single disconfirming instance of the proposition. Third, the hypotheses that are invoked to explain away the apparent disconfirming instance are not subjected to an independent empirical test. In a situation where there is a strong background of supporting evidence for an inductive generalization and an isolated disconfirming instance, it is reasonable to discount the disconfirming instance as apparent and to explain it away on whatever empirical grounds are most plausible.

The case against premise (A1) can be considerably strengthened by revising Ayer's scenario as follows: Increase the number of disconfirming instances of the proposition so that it is large relative to the number of confirming instances; and subject the hypotheses invoked to explain away the apparent disconfirming instances to independent tests that fail to support them. Let us now suppose that one has experienced a large number of apparent disconfirming instances of the proposition that $2 \times 5 = 10$ and, furthermore, that empirical investigations of the hypotheses invoked to explain away these disconfirming instances produce little, if any, support for the hypotheses. Given these revisions, Ayer can continue to endorse premise (A1) only at the expense of holding empirical beliefs that are at odds with the available evidence.

OPPOSING ARGUMENTS

Radical empiricism is the view that denies the existence of a priori knowledge. Its most famous proponents are John Stuart Mill and Willard Van Orman Quine. One common strategy that radical empiricists employ in arguing against the existence of a priori knowledge is to consider the most prominent examples of propositions alleged to be knowable only a priori and to maintain that such propositions are known empirically. Since mathematical knowledge has received the most attention, this entry will focus on it.

Mill's (1973) account of mathematical knowledge is a version of inductive empiricism. Inductive empiricism with respect to a domain of knowledge involves two theses. First, some propositions within that domain are epistemically more basic than the others, in the sense that the nonbasic propositions derive their justification from the basic propositions via inference. Second, the basic propositions are known by a process of inductive inference from observed cases. Mill's focus is on the basic propositions of arithmetic and geometry, the axioms and definitions of each domain. His primary thesis is that they are known by induction from observed cases.

Mill's position faces formidable objections, such as those offered by Gottlob Frege (1974). Let us assume, however, that these objections can be deflected and that Mill offers a plausible inductive empiricist account of mathematical knowledge to assess how this concession bears on the existence of a priori knowledge. If Mill is right, then all epistemically basic propositions of arithmetic and geometry are justified on the basis of observa-

tion and inductive generalization. It follows that Kant's claim that mathematical knowledge cannot be derived from experience is wrong. It does not follow, however, that the claim that such knowledge is a priori is wrong. From the fact that mathematical knowledge is or can be derived from experience, it does not immediately follow that such knowledge is not or cannot be derived from some nonexperiential source. Mill is aware of the gap in his argument and attempts to close it with the following observations:

> They cannot, however, but allow that the truth of the axiom, Two straight lines cannot inclose a space, even if evident independently of experience, is also evident from experience. … Where then is the necessity for assuming that our recognition of these truths has a different origin from the rest of our knowledge, when its existence is perfectly accounted for by supposing its origin to be the same? … The burden of proof lies on the advocates of the contrary opinion: it is for them to point out some fact, inconsistent with the supposition that this part of our knowledge of nature is derived from the same sources as every other part. (1973, pp. 231–232)

Mill moves from the premise that inductive empiricism provides an account of knowledge of mathematical axioms to the stronger conclusion that knowledge of such axioms is not a priori by appealing to a version of the explanatory simplicity principle: If a putative source of knowledge is not necessary to explain knowledge of the propositions within some domain, then it is not a source of knowledge of the propositions within that domain. Mill's argument can be articulated as follows:

(M1) Inductive empiricism provides an account of mathematical knowledge based on inductive generalization from observed cases.

(M2) φ is a source of knowledge for some domain D only if φ is necessary to explain knowledge of some propositions within D.

(M3) Therefore, mathematical knowledge is not a priori.

The burden of the argument is carried by (M2), the explanatory simplicity principle.

Casullo (forthcoming) maintains that the explanatory simplicity principle conflicts with a familiar fact of one's epistemic life. The justification of some of one's beliefs is overdetermined by different sources. There are some beliefs for which one has more than one justification, each of those justifications derives from a different source, and each, in the absence of the others, is sufficient to justify the belief in question. For example, I have misplaced my wallet again and wonder where I might have left it. I suddenly recall having left it on the kitchen table when I came in from the garage last night. My recollection justifies my belief that my wallet is on the kitchen table. However, just to be sure, I walk out to the kitchen to check. To my relief, I see my wallet on the table. My seeing my wallet on the table also justifies my belief that my wallet is on the table. So here my justification is overdetermined by different sources. If the justification of my belief is overdetermined by two different sources, it follows that my belief is justified by two different sources. Hence, in the absence of an argument against the possibility of epistemic overdetermination, Mill's appeal to the explanatory simplicity principle simply begs the question.

Quine rejects inductive empiricism. He rejects the idea that there are basic mathematical propositions that, taken in isolation, are directly justified by observation and inductive generalization. Quine's account of mathematical knowledge is a version of holistic empiricism. Mathematical propositions are components of scientific theories. They are not tested directly against observation, but only indirectly via their observational consequences. Moreover, they do not have observational consequences in isolation, but only in conjunction with the other propositions of the theory. Hence, according to holistic empiricism, entire scientific theories, including their mathematical components, are indirectly confirmed or disconfirmed by experience via their observational consequences.

The main concern in this entry is not to assess the cogency of Quine's account of mathematical knowledge, but to determine whether it provides an argument against the existence of a priori knowledge. The argument of Quine's classic paper "Two Dogmas of Empiricism" (1963) remains controversial (for further discussion, see Boghossian 1996). The stated target of his attack is a conception of analyticity inspired by Frege: A statement is analytic if it can be turned into a logical truth by replacing synonyms with synonyms. Quine's contentions can be summarized as follows:

(1) Definition presupposes synonymy rather than explaining it.

(2) Interchangeability *salva veritate* is not a sufficient condition of cognitive synonymy in an extensional language.

(3) Semantic rules do not explain "Statement S is analytic for language L," with variable "S" and "L."

(4) The verification theory of meaning provides an account of statement synonymy that presupposes reductionism, but reductionism fails.

(5) Any statement can be held to be true come what may. No statement is immune to revision.

Quine's contentions appear to be directed at the concept of synonymy and the doctrine of reductionism. They are not explicitly directed at a priori knowledge. Hence, if "Two Dogmas" does indeed present a challenge to the existence of a priori knowledge, then some additional premise is necessary that connects those contentions to the a priori.

According to the traditional reading of his argument, Quine's contentions constitute an extended attack on the cogency of the analytic-synthetic distinction. Quine's ultimate goal is to undermine the central claim of the logical empiricist tradition:

(LE) All a priori knowledge is of analytic truths.

On this reading, (LE) provides the connection between his contentions and the rejection of the a priori. Let us grant that Quine's goal is to undermine (LE) and that he successfully challenges the cogency of the analytic-synthetic distinction. Does it follow that there is no a priori knowledge? No. (LE) is a thesis about the nature of the propositions alleged to be known a priori. If Quine is right, then (LE) itself is incoherent. But from the fact that a thesis about the nature of propositions known a priori is incoherent, it does not follow that there is no a priori knowledge.

An alternative response is to take (LE) as a conceptual claim; that is, to take it as claiming that the concept of a priori knowledge involves the concept of analytic truth. On this reading, the incoherence of the concept of analytic truth entails the incoherence of the concept of a priori knowledge. This response, however, rests on a false conceptual claim. The concept of a priori knowledge does not explicitly involve the concept of analytic truth. One might argue that it implicitly involves the concept of analytic truth by maintaining that all a priori knowledge is of necessary truths; and endorsing some version of the so-called linguistic theory of necessary truth. There are, however, two problems with this argument. First, the concept of a priori knowledge does not involve, either explicitly or implicitly, the concept of necessary truth. Second, there is no plausible analysis of the concept of necessary truth in terms of the concept of analytic truth.

Some champions of "Two Dogmas" propose an alternative connection between Quine's contentions and the rejection of the a priori. Putnam (1983) maintains that Quine's contentions are directed toward two different targets. The initial contentions are directed toward the semantic concept of analyticity. Contention (5), however, is directed toward the concept of a statement that is confirmed no matter what, which is not a semantic concept. The concept of a statement that is confirmed no matter what is an epistemic concept. It is a concept of apriority. Kitcher endorses Putnam's reading of Quine's argument, "If we can know a priori that *p* then no experience could deprive us of our warrant to believe that *p*" (1983, p. 80). But, according to Quine, no statement is immune from revision. Hence, the Putnam-Kitcher version of Quine's argument can be stated as follows:

(Q1) No statement is immune to revision in light of recalcitrant experience.

(Q2) If S's belief that p is justified a priori, then S's belief that p is not rationally revisable in light of any experiential evidence.

(Q3) Therefore, no knowledge is a priori.

The argument fails. Premise (Q2) is open to the objection presented against (AP3) in the first section.

THE EXPLANATORY CHALLENGE

A more recent challenge to the a priori derives from Quine's influential "Epistemology Naturalized" (1969). Epistemic naturalism comes in many different forms. The most radical form advocates the replacement of philosophical investigations into the nature of human knowledge with scientific investigations. More moderate forms advocate that philosophical theories concerning human knowledge cohere with scientific theories. Paul Benacerraf (1973), for example, argues that the truth conditions for mathematical statements make reference to abstract entities and that knowing a statement requires that one be causally related to the entities referred to by its truth conditions. Since abstract entities cannot stand in causal relations, one cannot know mathematical statements. The argument raises a more general challenge to the possibility of a priori knowledge since proponents of the a priori (apriorists) generally hold that most, if not all, a priori knowledge, is of necessary truths; and that the truth conditions of necessary truths make reference to abstract entities. Although some reject the argument on

the grounds that its epistemic premise appears to presuppose the generally rejected causal theory of knowledge, others, such as Hartry Field (1989), maintain that it points to a deeper problem. In the absence of an explanation of how it is possible to have knowledge of abstract entities, a priori knowledge remains mysterious.

The explanatory challenge goes beyond a commitment to epistemic naturalism. It derives support from broader epistemological considerations. To appreciate the full import of the challenge, two issues regarding the existence of a priori knowledge must be distinguished. Apriorists typically maintain that one knows certain logical, mathematical, and conceptual truths and that such knowledge is a priori. Radical skeptics deny that one has knowledge of the truths in question. Radical empiricists, however, are not radical skeptics. They do not deny that one knows the truths in question. Radical empiricists only deny that one's knowledge of these truths is a priori. Therefore, the primary dispute between apriorists and radical empiricists is over the source of the knowledge in question. They offer two competing theories of the source of the knowledge in question, and each maintains that its theory offers the better explanation of the knowledge in question. Therefore, to support their primary contention, apriorists must provide supporting evidence for the claim that there exist nonexperiential sources of justification and that such sources explain how one knows the truths in question.

BonJour (1998) and Ernest Sosa (2000) offer philosophical supporting evidence, a mix of phenomenological and a priori considerations. Casullo (2003) argues that a more promising approach is to supplement the philosophical evidence with evidence based on empirical investigations. Before empirical evidence can be enlisted to support the case for the a priori, however, additional philosophical work is necessary. The first step is to provide (1) a generally accepted phenomenological description of the cognitive states that noninferentially justify beliefs a priori, (2) the type of beliefs they justify, and (3) the conditions under which they justify the beliefs in question. Apriorists typically defend the claim that there are nonexperiential sources of justification by reflecting on their own cognitive situations and identifying phenomenologically distinct states, which they claim justify certain beliefs a priori. A cursory survey of the descriptions of these states offered by different theorists reveals wide variation. George Bealer (1996) and Sosa (1996) both maintain that the cognitive states that justify a priori are aptly described as *seemings*, but they offer different phenomenological descriptions of seemings. Plantinga (1993) and BonJour (1998) maintain that the states in question are more aptly described as *seeings*, but they offer different phenomenological descriptions of seeings. Bealer agrees with BonJour that the cognitive states that justify a priori are irreducible, but disagrees with him over the character of the states. On the contrary, Sosa agrees with Plantinga that the states are reducible to more familiar cognitive states, but disagrees with him over the character of the reducing states.

There is also wide variation among apriorists over the scope of beliefs justified a priori. Within the context of arguing against radical empiricism, the focus is on stock examples such as elementary logical or mathematical propositions and some familiar examples of alleged synthetic a priori truths. Few apriorists, however, believe that a priori justification is limited to those cases. Consequently, they must provide a more complete specification of the range of beliefs alleged to be justified by such cognitive states. One issue requires particular attention. The examples of a priori knowledge typically cited by apriorists are necessary truths. But here it is important to distinguish between knowledge of the truth value and knowledge of the general modal status of necessary propositions. A critical question now emerges: What is the target of a priori justification? Is it the general modal status of a proposition, its truth value, or both? If it is both, two further questions arise. Are beliefs about the truth value of a necessary proposition and beliefs about its general modal status justified by the same cognitive state or different cognitive states? Are some beliefs about the truth value of contingent propositions justified a priori?

Once the philosophical work is complete, the project of providing empirical supporting evidence for the a priori can be pursued. This involves providing (1) evidence that the cognitive states identified at the phenomenological level are associated with processes of a single type or relevantly similar types; (2) evidence that the associated processes play a role in producing or sustaining the beliefs they are alleged to justify; (3) evidence that the associated processes are truth-conducive; and (4) an explanation of how the associated processes produce the beliefs they are alleged to justify. The third area of empirical investigation offers the prospect of supporting the claim that there are nonexperiential sources of justification. Many prominent apriorists, including Bealer, BonJour, Plantinga, and Sosa, maintain that truth conduciveness is a necessary condition for epistemic justification. Moreover, even those who deny this concede that evidence that a source of beliefs is error conducive defeats whatever justification that the

source confers on the beliefs that it justifies. The claim that a source of beliefs is truth conducive or, more minimally, that it is not error conducive is a contingent empirical claim that can be supported only by empirical investigation.

The fourth area of empirical investigation offers the prospect of addressing the explanatory challenge. First, causal-perceptual models appear to be of limited utility in explaining how nonexperiential sources of justification provide cognitive access to necessary truths. Empirical investigation into human cognition offers the prospect of uncovering alternative models of cognitive access that can be utilized in the case of nonexperiential sources. Second, investigation of the specific cognitive processes associated with the cognitive states alleged to justify a priori may provide a better understanding of how the processes in question produce true beliefs about their subject matter. This understanding, in turn, is the key to providing a noncausal explanation of how the states in question provide cognitive access to the subject matter of the beliefs they produce. Third, although apriorists deny that epistemology is a chapter of science, they acknowledge that both epistemology and science contribute to the overall understanding of human knowledge. Establishing that the cognitive processes invoked by their epistemological theory are underwritten by their scientific commitments strengthens the apriorist's overall theory by demonstrating the coherence of its components.

See also Analyticity; A Priori and A Posteriori; Ayer, Alfred Jules; Chisholm, Roderick; Field, Hartry; Frege, Gottlob; Hempel, Carl Gustav; Kant, Immanuel; Knowledge and Modality; Kripke, Saul; Mathematics, Foundations of; Mill, John Stuart; Plantinga, Alvin; Putnam, Hilary; Quine, Willard Van Orman; Sosa, Ernest.

Bibliography

Ayer, A. J. *Language, Truth and Logic*. New York: Dover, 1952.

Bealer, George. "A Priori Knowledge and the Scope of Philosophy." *Philosophical Studies* 81 (1996): 121–142.

Benacerraf, Paul. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–679.

Boghossian, Paul A. "Analyticity Reconsidered." *Noûs* 30 (1996): 360–391.

BonJour, Laurence. *In Defense of Pure Reason*. New York: Cambridge University Press, 1998.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Burge, Tyler. "Content Preservation." *Philosophical Review* 102 (1993): 457–488.

Casullo, Albert. *A Priori Justification*. New York: Oxford University Press, 2003.

Casullo, Albert. "Epistemic Overdetermination and A Priori Justification." *Philosophical Perspectives* 19 (2005): 41–58.

Chisholm, Roderick M. *Theory of Knowledge*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1977.

Field, Hartry. *Realism, Mathematics, and Modality*. Oxford, U.K.: Blackwell, 1989.

Frege, Gottlob. *The Foundations of Arithmetic*. 2nd ed. Translated by J. L. Austin. Evanston, IL: Northwestern University Press, 1974.

Hempel, Carl. "On the Nature of Mathematical Truth." In *Necessary Truth*, edited by R. C. Sleigh. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Kant, Immanuel. *Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: St Martin's Press, 1965.

Kitcher, Philip. *The Nature of Mathematical Knowledge*. New York: Oxford University Press, 1983.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Mill, John Stuart. *A System of Logic*, edited by J. M. Robson. Toronto, Canada: University of Toronto Press, 1973.

Plantinga, Alvin. *Warrant and Proper Function*. New York: Oxford University Press, 1993.

Putnam, Hilary. "'Two Dogmas' Revisited." In *Realism and Reason: Philosophical Papers*. Vol. 3. New York: Cambridge University Press, 1983.

Quine, W. V. O. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Quine, W. V. O. "Two Dogmas of Empiricism." In *From a Logical Point of View*. 2nd ed. New York: Harper and Row, 1963.

Sosa, Ernest. "Modal and Other A Priori Epistemology: How Can We Know What Is Possible and What Impossible?" *Southern Journal of Philosophy* 38 (Supplement) (2000): 1–16.

Sosa, Ernest. "Rational Intuition: Bealer on its Nature and Epistemic Status." *Philosophical Studies* 81 (1996): 151–162.

Albert Casullo (2005)

KNOWLEDGE, THE PRIORITY OF

One fairly specific understanding of the priority of knowledge is the idea that instead of trying to explain knowledge in terms of belief plus truth, justification, and something, we should explain belief in terms of knowledge. This is to reverse the usual explanatory priority of knowledge and belief. This fairly specific idea generalizes in two directions. (1) Perhaps we should explain other notions in terms of knowledge as well. Some possibilities include assertion, justification or evidence, mental content, and intentional action. (2) Perhaps we could explain other relatively internal states like intentions, attempts, and appearances in terms of their more obviously external counterparts: intentional action and perception.

That knowledge is prior to belief has historically been a minority opinion. The idea that a belief, and the mind more generally, is what it is regardless of any actual connection to the external world is still widely accepted. Accepting the priority of knowledge constitutes a rejection of the picture of the mind as a self-contained, inner realm.

UNDERSTANDING BELIEF

Bernard Williams (1973) tries to explain the impossibility of believing at will in terms of the idea that belief aims at the truth. Suppose you are anxious about tomorrow's weather but have no access to a weather forecast or any other evidence. If you want to reduce your anxiety, then you might, if it was in your power to do so, simply decide to believe that it will be sunny tomorrow. But if you knew that this attempted belief was based not on evidence or any apparent connection to the facts, but on a decision, then it would be hard for you to see your attempted belief as aiming at the truth. It would also be hard for you to see it as a belief. So, perhaps, it could not be a belief.

Let us agree that in this particular case seeing your attempted belief as the result of a decision seriously casts doubt on the possibility of its being a belief. Is this best explained by the idea that belief aims at the truth? Since you have no evidence about the weather, the problem cannot be that you have reason to think the attempted belief will fail to achieve this aim. On the contrary, you have every reason to think it is at least possible that it will achieve this aim. So what keeps you from aiming at it?

If you merely guess that a flipped coin will come up heads, then you probably do not believe that it will. Guesses are not beliefs. But guessing aims at the truth. In guessing you are trying to get it right, and if you succeed, this is as good as a guess can get. When you see your attempted belief as the result of a decision, you may still be hoping, trying, or aiming to get things right. But you know believing is not epistemically justified for you in that instance. Whatever practical reasons you may have for believing that p, you have no evidence that p. It is seeing your state as unjustified while remaining in it that seriously casts doubt on the possibility of its being a belief. To understand belief, we need a connection between belief and justification, not just between belief and truth.

Suppose that someone has an unjustified, true belief. If belief aims at the truth, then this belief has achieved its aim. Perhaps justification is a good guide or a means to the truth. But if truth is the aim, and this belief has achieved that aim by other means, then epistemic justification or lack thereof is irrelevant to the evaluation of this belief. So if belief merely aimed at the truth, then it would not be automatically subject to evaluation from the epistemic point of view. If belief aims at knowledge, however, instead of mere truth, then it is clear why it is subject to this kind of evaluation. Unjustified beliefs may be true, but they cannot constitute knowledge.

Perhaps this does not capture what is meant in saying that belief aims at the truth. When you believe that p, you do not merely hope or try to get things right. In some sense it seems to you as though you already have gotten things right. We do not want to say that if you believe that p, you believe that your belief that p is true. This leads to an infinite number of beliefs. You do not need beliefs about beliefs to have beliefs about the world. But if you do have a view about your views, it must cohere with those views, where coherence involves more than just logical consistency. You can think that your belief that p is true, but you cannot think that your belief that p is false. You cannot assert, "I believe that p, but not p," and you cannot believe it either. This "cannot" is probably a normative "cannot," rather than an expression of logical impossibility.

What goes for error goes for ignorance as well. There is something wrong with assertions of the form "p, but I do not believe that p." Whatever is wrong with these assertions, they would be just as bad in the privacy of your own mind. If you think about Moore-paradoxical statements from the normative perspective, then the same kind of incoherence that is involved in the standard cases also seems to infect the following: p, but I have no reason to believe that p; I believe that p, but I should not believe it; and p, but I am completely unreliable about these things. The belief that p not only conflicts with the belief that you are wrong or that you do not believe that p. It also rules out the belief that you are unjustified or not in a position to know. These first-person facts about belief can be explained by the idea that belief aims at knowledge, but not by the idea that belief aims at truth.

The idea that belief aims at knowledge is a normative claim. From the point of view of belief there is something wrong with false beliefs, but there is also something wrong with unjustified beliefs. There is something wrong with accidentally true beliefs, even when they are justified. But from the point of view of belief, there is nothing wrong with knowledge. For a belief, knowledge is as good as it gets.

ASSERTION AND EVIDENCE

Moore's Paradox tells us not only about the nature of belief but also about the nature of assertion. Peter Unger

(1975) and Timothy Williamson (2000) are both defenders of the priority of knowledge. Both agree that when you assert that p, you not only represent p as being true—you not only represent yourself as believing that p—you also represent yourself as knowing that p. Propositions of the form "p, but I do not know that p" are unassertable because they violate the rule of assertion: assert only what you know. Unlike Williamson, Unger is a radical skeptic. When he tells you not to assert what you do not know, he is basically telling you to keep quiet. The consequences you draw from the priority of knowledge will depend on your general views about knowledge. But the basic idea does not discriminate against skeptics.

Unger and Williamson also agree that there is an important connection between knowledge and justification, though they articulate the connection in different ways. Unger's general idea is that if you are justified in believing that p, then you must know something. More specifically, he believes that if your reason for believing that p is that q, then you must know that q. According to Williamson evidence is knowledge. If your body of evidence consists of a set of propositions, then you must know each member of this set. Both of these views open up the possibility of merely apparent evidence. This is not a problem for Unger, since he thinks that all evidence is merely apparent.

Is there a problem for Williamson? Suppose you have a justified, false belief that p; you infer that q on the basis of this belief; and you think that p is your evidence that q. If evidence is knowledge, then you are simply mistaken in thinking that p is your evidence that q. You may even be mistaken in thinking that you have evidence that q. This can seem problematic if you think that evidence is such that, if you have some, then you are at least in a position to know that you have some; and if you do not have any, then you are in a position to know that you do not have any; and if p is or is not evidence for you to believe that q, then you are in a position to know whether or not this is so.

According to Williamson evidence is not this kind of thing, but neither is anything else. In Williamson's terminology a condition is "luminous" just in case one is in a position to know that the condition obtains, if it does. For example, you could easily be sleeping in a cold room without being able to tell that the room is cold. So the condition of one's being in a cold room is not luminous. But you might have thought that, if you feel cold, seem to see a red wall, or believe that there is life on Mars, then you are in a position to know that you feel cold, seem to see a red wall, or believe that there is life on Mars. In other words you might have thought that these conditions are luminous.

Williamson has a general argument designed to show that there are no nontrivial luminous conditions. Not even the condition that one feels cold is luminous. There is always a potential gap between the facts and your ability to know the facts, even when the facts are about your own present state of mind. So the idea that evidence would not be luminous if only knowledge were or could be evidence is no objection to the view. Evidence would not be luminous regardless what it was. If we do have some other form of privileged access to evidence or the justification of our own beliefs, and if our having that access is incompatible with the idea that evidence is knowledge, then it must be shown.

MENTAL CONTENT

Gilbert Harman (1999) believes that the basic mental notions are knowledge and intentional action. Belief and intention are generalizations of these that allow for error and failure. Harman therefore clearly endorses the priority of knowledge. He also believes that the content of a concept is determined by its functional or conceptual role: its typical or normal connections to perception, its role in practical and theoretical reasoning, and its connection to intentional action. Finally, he accepts content externalism: the view that it is possible for intrinsic duplicates to differ in the contents of their thoughts.

The first two of Harman's views explain why he holds the third. My concept of water is typically caused by perceptions of and hearing about water, and the concept is causally involved in my intentional interactions with water. Suppose that I have an intrinsic duplicate on Hilary Putnam's (1975) Twin Earth. On Twin Earth there is something that looks, smells, tastes, and feels like water but is not water. Call it XYZ. When I interact with water, my twin interacts with XYZ. This difference in our interactions does not influence our intrinsic natures. But it does influence the contents of our thoughts. Unlike me, my twin never perceives or interacts with water. Even if you dragged my twin into my kitchen, he would not intentionally interact with water, nor would he perceive that the water was running. My concept differs in content from my duplicate's concept because the functional roles of the concepts are different. The functional roles of the concepts are different because these roles must be understood in terms of knowledge, perception, and intentional action.

Harman is not the only philosopher to combine the priority of knowledge with a conceptual role account of

content. Christopher Peacocke (1999) has a sophisticated version of this view. According to Peacocke epistemically individuated concepts can be individuated, at least in part, in terms of the conditions under which certain judgments involving those concepts would constitute knowledge. Furthermore, every concept is either epistemically individuated, or individuated in part in terms of its relations to epistemically individuated concepts. If epistemically individuated concepts do in fact play this central role in our system of concepts, then the priority of knowledge may provide an explanation of this fact.

The conceptual role theory of content is or is a descendant of the idea that the content of a thought or concept is determined by what Wilfrid Sellars (1956) calls its place in a space of reasons. John McDowell (1996) argues, among other things, that if you take this idea seriously, then thinking of the space of reasons broadly enough to encompass not only beliefs but also knowledge is not an optional extra. There is no purely internal space of reasons. To understand how experience can be part of the logical space of reasons, and so how our thoughts can have any content at all, we need to understand how a subject can be open to the way things manifestly are, where this involves knowing about what is going on around you.

Setting aside conceptual roles, the priority of knowledge may provide an adequacy condition for the theory of content. Suppose there was a kind of content or a kind of representation that could not distinguish a situation in which water is wet from a situation in which something that merely looks, smells, feels, and tastes like water is wet. A picture in the head, qualitatively conceived, may be such a representation. Accepting this kind of representation or content could never constitute knowledge, since there would not be the right kind of distinction between justification and knowledge. If one of your beliefs about barns constitutes knowledge, then it matters whether or not there are fake barns in your neighborhood. If believing something about barns were a matter of accepting one of these phony propositions that cannot distinguish between real and fake barns, then it would not matter whether the barns were real or fake. According to the priority of knowledge, if these representations are not even candidates for knowledge, then they are not to be believed.

CONTACT

What justifies this preoccupation with knowledge? Each account of something in terms of knowledge must of course be judged on its own merits, but is there anything special about knowledge that holds them all together? A true belief will match or accurately represent the world, but knowledge seems to involve a kind of contact with the world. The recognition of the importance of this kind of contact is one of the underlying ideas that unifies these various approaches.

Edmund L. Gettier (1963) shows that justified, true belief alone is not sufficient for knowledge. If a justified, true belief is inferred from a false premise, then it will not constitute knowledge, even if that premise was justified. Not all cases of justified, true belief without knowledge involve inference from a false belief. Alvin I. Goldman (1992) imagines a case in which you look at a barn that is surrounded by realistic barn facades and form the justified, true belief that it is a barn. You do not know even though you are right because you just got lucky. Though you do get lucky, and it is just an accident, we cannot deny that your belief about the barn is causally connected to the barn. If we were trying to understand knowledge in terms of being in contact with the world, then we would need to specify the right kind of contact. But if you are using the notion of knowledge to explain other things, then it is easy to say what kind of contact you have in mind: you are connected in the right way to p if you know that p.

The presence or absence of this kind of contact matters in a variety of areas in philosophy. For example, as Unger argues, a factive propositional attitude either entails knowledge or the absence of knowledge. If you are happy that it is raining, or you notice that it is raining, then it follows that it is raining. These propositional attitudes are factive. Moreover, if you are in one of these mental states, then it also follows that you know that it is raining. By contrast, if you forget that it is raining, then it still follows that it is raining. Forgetting that p is just as factive as being surprised or embarrassed that p, but if you forget or are unaware that p, then it does not follow that you know that p. It follows that you do not know that p. Not all factive attitudes entail knowledge. But they do not leave the question of knowledge open. As Robert Gordon (1969) points out the propositional emotions, even the nonfactive ones, do not leave open the question of knowledge either. If you fear, hope, or are worried that p, then it does not follow that p. But it does follow that you do not know whether or not that p.

What matters in all these cases is genuine contact with the world, rather than merely a match between what is inside and what is outside the mind. You might be happy when it rains without being happy that it rains. You need the right kind of connection between the rain and the happiness for the happiness to be about the rain.

If the disturbing sight of the rain leads to your taking certain kinds of medication, then the rain, and your knowledge thereof, may cause the happiness, but you will not necessarily be happy that it is raining. The rain is causally related to the happiness, but not in the right way. What is the right way? It looks like the happiness has to be connected to the rain in the same way that a belief has to be connected to a fact for the belief to constitute knowledge.

Whenever something interesting requires contact between the mind and the world, a causal theory of that thing will at least look plausible. But any such theory will be faced with deviant causal chains: cases where there is a causal connection, but not the right kind of causal connection. You might intend to run over your uncle, and this may lead you to back your car out of your driveway to drive to his house. But if, unknown to you, your uncle is napping behind the wheels of your car, you will run him over; your intention to run him over will cause you to run him over; but you will not, in this case, run him over on purpose. Your intention to A is causally related to your A-ing, but not in the right way, so you do not intentionally A. This is a deviant causal chain.

Here is one thing to notice about the case. You correctly believe that backing out of your driveway will lead to running over your uncle. Given your plan, the belief, we may say, is justified. But the belief does not constitute knowledge. If it is just an accident that your belief is true, and you act on that belief, then it will just be an accident that your attempts are successful, if they are successful at all. To get intentional action, your means-ends beliefs must constitute knowledge. This is one suggestion for ruling out causal deviance in action theory. If this is right, then it not only follows that we can explain particular actions in terms of particular states of knowledge. It at least suggests that we understand intentional action, one kind of contact between the mind and the world, in terms of knowledge. Unless we also have to understand knowledge in terms of action, it looks as though knowledge is the more fundamental notion.

See also Belief.

Bibliography

Adler, Jonathan E. *Belief's Own Ethics*. Cambridge: Massachusetts Institute of Technology Press, 2002.

Burge, Tyler. "Individualism and the Mental." In *Midwest Studies in Philosophy*. Vol. 4, edited by Peter A. French, Theodore E. Uehling, and Howard K. Wettstein. Minneapolis: University of Minnesota Press, 1979.

Gettier, Edmund L. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121–123.

Gibbons, John. "Knowledge in Action." *Philosophy and Phenomenological Research* 62 (2001): 579–600.

Goldman, Alvin I. "Discrimination and Perceptual Knowledge." In *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge: Massachusetts Institute of Technology Press, 1992.

Gordon, Robert. "Emotions and Knowledge." *Journal of Philosophy* 66 (1969): 408–413.

Griffiths, A. Phillips, ed. *Knowledge and Belief*. London: Oxford University Press, 1967.

Harman, Gilbert. "(Nonsolisistic) Conceptual Role Semantics." In *Reasoning, Meaning, and Mind*. Oxford, U.K.: Clarendon Press, 1999.

McDowell, John. "Knowledge and the Internal." In *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press, 1998.

McDowell, John. *Mind and World*. Cambridge, MA: Harvard University Press, 1996.

Moore, G. E. *Commonplace Book, 1919–1953*, edited by Casimir Lewy. London: Allen and Unwin, 1962.

Peacocke, Christopher. *Being Known*. Oxford, U.K.: Clarendon Press, 1999.

Putnam, Hilary. "The Meaning of 'Meaning.'" In *Mind, Language, and Reality*. Cambridge, U.K.: Cambridge University Press, 1975.

Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." In *Minnesota Studies in the Philosophy of Science*. Vol. 1, edited by Herbert Feigl and Michael Scriven. Minneapolis: University of Minnesota Press, 1956.

Shoemaker, Sydney. "Moore's Paradox and Self-Knowledge." In *The First-Person Perspective and Other Essays*. Cambridge, U.K.: Cambridge University Press, 1996.

Sorensen, Roy A. *Blindspots*. Oxford, U.K.: Clarendon Press, 1988.

Unger, Peter. *Ignorance: A Case for Scepticism*. Oxford, U.K.: Clarendon Press, 1975.

Velleman, J. David. "On the Aim of Belief." In *The Possibility of Practical Reason*. Oxford, U.K.: Clarendon Press, 2000.

Williams, Bernard. "Deciding to Believe." In *Problems of the Self: Philosophical Papers, 1956–1972*. Cambridge, U.K.: Cambridge University Press, 1973.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford, U.K.: Oxford University Press, 2000.

John Gibbons (2005)

KNOWLEDGE, SOCIOLOGY OF

See *Sociology of Knowledge*

KNOWLEDGE, SPECIAL SORTS OF

See *Knowledge, A Priori; Knowledge and Modality; Moral Epistemology*

KNOWLEDGE, SPECIAL SOURCES OF

See *Inference to the Best Explanation; Introspection; Intuition; Memory; Perception; Perception, Contemporary Views; Testimony*

KNOWLEDGE, THEORY OF

See *Epistemology; Epistemology, History of*

KNOWLEDGE AND BELIEF

The nature of knowledge has been a central problem in philosophy from the earliest times. One of Plato's most brilliant dialogues, the *Theaetetus*, is an attempt to arrive at a satisfactory definition of the concept, and Plato's dualistic ontology—a real world of eternal Forms contrasted with a less real world of changing sensible particulars—rests on epistemological foundations.

The problem of knowledge occupies an important place in most major philosophical systems. If philosophy is conceived as an ontological undertaking, as an endeavor to describe the ultimate nature of reality or to say what there really is, it requires a preliminary investigation of the scope and validity of knowledge. Only that can reasonably be said to exist which can be known to exist. If, on the other hand, philosophy is conceived as a critical inquiry, as a second-order discipline concerned with the claims of various concrete forms of intellectual activity, it must consider the extent to which these activities issue in knowledge.

In modern philosophy in the widest sense of the phrase—that is, philosophy since the Renaissance—theory of knowledge has usually been the primary field of philosophical investigation. René Descartes and John Locke, David Hume and Immanuel Kant, were all, in the first instance, epistemologists. Epistemological considerations played an important part in the work of Arthur Schopenhauer, but they were less central in G. W. F. Hegel and Friedrich Nietzsche, who were more occupied with the nature of the human mind in general and with the institutions within which it is exercised than with its more narrowly cognitive aspects. With Søren Kierkegaard and his existentialist descendants the focus of interest was man's will rather than his intellect. Anglo-Saxon philosophy, however, has remained epistemological. J. S. Mill, Bertrand Russell, and the analytic philosophers of the twentieth century continued to work in the area marked out by Locke and Hume. Even the British Hegelians of the late nineteenth century, the school of Thomas Hill Green and F. H. Bradley, were led into far-reaching epistemological studies by the character of the native tradition they were seeking to overthrow.

Belief has had less attention from philosophers. It has generally been taken to be a more or less unproblematic inner state, accessible to introspection. But there has been disagreement about whether it is active or passive, Descartes having contended that assent is a matter of will, Hume that it is an emotional condition in which one finds oneself. Alexander Bain urged that belief should be interpreted in terms of the tendencies to action with which it is associated, and Charles Sanders Peirce took the view that it is an unobstructed habit of action that, like health, comes to our notice only when we have lost it. Faith, especially religious faith, and probability, the logic of rational belief, have been thoroughly examined, but belief itself has received surprisingly cursory treatment.

THE DEFINITION OF KNOWLEDGE

According to the most widely accepted definition, knowledge is justified true belief. That it is a kind of belief is supported by the fact that both knowledge and belief can have the same objects (thus, half an hour ago I believed I had left my raincoat in the garage; now I know that I have) and that what is true of someone who believes something to be the case is also true, among other things, of one who knows it. One who comes to know what he formerly believed does not lose the conviction he formerly had.

It is obvious and generally admitted that we can have knowledge only of what is true. If I admit that *p* is false, I must admit that I did not know it and that no one else did, although I may have thought and said so. It is urged, on the ground that beliefs that merely happen to be true cannot be regarded as knowledge, that knowledge must be justified. I may draw a true conclusion by invalid means from false premises or believe a truth on the strength of a dream or the misremembered testimony of a notorious liar. In such cases as these I do not really know the things I believe, although what I believe is true. There are, however, objections to all three parts of the definition of knowledge as justified true belief.

TRUTH. It has been suggested that the requirement that what is known be true is excessively stringent. Complete certainty of a statement's truth is not to be had; the best we can achieve is very strong grounds for thinking it true.

Thus, if knowledge entails truth, we can never attain knowledge or, at any rate, never know that we have done so. This objection is misconceived. If I firmly believe that something is true on what I take to be sufficient grounds, I am right to say that I know it. It may be that the grounds are, in fact, insufficient and that what I claim to know is false. In that case my claim is mistaken, but it does not follow that I was wrong to make it in the sense that I had no justification for doing so.

It has also been argued, with a view to showing that knowledge and belief are quite distinct and unrelated, that whereas beliefs can be true or false, knowledge is neither. This argument exploits the fact that we speak of a belief but not of a knowledge, only of a piece or item of knowledge. Furthermore, since all items or pieces of knowledge are by definition true, we never need to speak of them as true items or pieces in order to distinguish them from false ones.

BELIEF. It is often objected that knowledge cannot be a kind of belief, even though they can have the same objects, because they exclude each other. If I know that *p*, it would be wrong for me to say that I believe it, since this would suggest that I do *not* know it. If, knowing *p*, I am asked "Do you believe that *p*?," I should reply "No, 1 know it." This is hardly a serious argument. I should mislead people if I described my wife as the woman I live with, and I might say, "No, she's my wife," if I were asked whether she is the woman I live with. Nevertheless, my wife is the woman I live with. What is true is that I do not *merely* live with her. Likewise, if I know that *p*, I do not merely believe it, but I do believe it all the same. It is often wrong or misleading in certain circumstances to say something that is unquestionably true. The boy who, having taken two jam tarts, answers the question "How many have you had?" by saying "One" has told the truth but not the whole truth.

A more powerful argument against the definition of knowledge in terms of belief is that people can, it seems, know something to be the case and yet refuse, or be unable to bring themselves, to believe it. A woman told by wholly reliable witnesses with a wealth of circumstantial detail that her husband has been killed in an accident might be in this position. One way of getting around this objection is to say that she believes both that her husband is dead and that he is not. It is possible and not uncommon to believe something and its contradictory. It is not possible both to believe something and to not believe it at the same time, and what she will say is, "I don't believe it," although what she means is that she believes it is false. Another possibility is to say that although she has conclusive grounds for believing that her husband is dead, she does not, in fact, believe it and does not know it either. To have conclusive grounds is one thing; to recognize that they are conclusive is another.

It should be noted that where knowledge and belief overlap, the kind of knowledge involved is propositional knowledge, or what Gilbert Ryle called "knowing that." There is also "knowing how" (to skate, tie a reef knot, do long division), where there are no propositions to be true or false and where knowledge can vary in degree. The two kinds of knowledge are connected in that both are the outcome of learning. Belief is always propositional or believing that; there is no believing how that serves as a defective version of knowing how to do something.

JUSTIFICATION. We often express unreasonable hunches or intuitions by saying, "I know," and if they turn out, to our gratified amazement, to be correct, we rejoice by saying, "I knew it." Does this show that true belief can be knowledge even without justification? The emphasis we put on the verb when we use it in such a case suggests that it is an abnormal or marginal use. It is generally accepted that lucky guesses should not count as knowledge.

An important difficulty arises from the requirement that true belief must be justified if it is to be knowledge. What is it for a belief to be justified? One obvious answer is that my belief in *q* is justified if there is some other belief *p* that entails or supports it. It is clearly not enough that this further belief *p* should merely exist. It must also be a belief of mine; I must know it to be true, and I must know that it justifies *q*. But if this is a definition of justification, the original definition of knowledge is rendered circular and generates a regress. It has the consequence that before any belief can be justified, an infinite series of justifications must already have taken place.

How can such a regress be halted? A natural step is to ask whether all justification has to be of this propositional or inferential kind. As Russell has observed, we can define derivative knowledge in this way but must add an account of intuitive or uninferred knowledge. Philosophers have fastened on two forms of intuitive knowledge that, by standing as the uninferred first premises of all inference, can terminate the regress of justification. First, there are self-evident necessary truths, and, second, there are basic contingent statements, immediately justified by the experiences they report and not dependent on the support of any further statable items of knowledge.

In the first group are the axioms of logic and mathematics, such as the law of excluded middle and the principle of the commutativity of addition ($a + b = b + a$), and statements that correspond to familiar verbal definitions, such as that kittens are young cats. Some philosophers hold that such intuitive, necessary truths record the results of intellectual intuition, the direct inspection of the relations of timeless universals; others, that their truth is essentially verbal in character, that one must accept them in order to be regarded as understanding the ordinary meaning of the words they contain. To accept an intuitive, necessary truth is to be ready to draw inferences in accordance with it. If I understand and accept the truth of "If (if *p*, then *q*), then (if not-*q*, then not-*p*)," I must regard the deduction of "If he's not over twenty-one, he's not eligible" from "If he's eligible, he's over twenty-one" as valid. By applying such rules of inference to intuitive necessary premises, further demonstrative necessary truths are arrived at.

Intuitive contingent truths have been held to be those that describe the immediate objects of perceptual or introspective experience—for example, "There is a green patch in the middle of my visual field" or "There appears to me to be a green flag here" and "I am in pain" or "I want to go to sleep." Basic statements like these are said to be incorrigible in the sense that they are wholly certified by the experiences they report and are logically immune from falsification by the results of any further experience. There may be no green flag here, but whatever may happen, there does now appear to be one. I may find it impossible to go to sleep once I get into bed, but I still want to go to sleep now. A statement is incorrigible if its truth follows from the fact that it is believed by the person to whom it refers. Thus, although I can make such a statement falsely, I must know that the statement is false when I do so. I cannot be honestly mistaken about my pains or the contents of my visual field.

It has sometimes been denied that there are any contingent, empirical statements that are basic and incorrigible in this sense. Coherence theories of knowledge have been propounded by the absolute idealists of the late nineteenth century and by C. S. Peirce, Karl R. Popper, and W. V. Quine in more empiricist forms in which beliefs are seen as justifying one another but none as in any sense self-justifying. To overcome the apparent circularity of the doctrine, it has been argued that some beliefs are relatively basic in that they can be accepted as true by some kind of convention or posited for the time being but that the element of dogmatism involved is only provisional and is open to revision.

PLATO'S *THEAETETUS*. Several of the points raised concerning truth, belief, and justification were first made in the *Theaetetus*, that most modern in spirit of Plato's dialogues. In it three definitions of knowledge are examined, and in the end all are rejected. The three are that knowledge is (1) perception or sensation, (2) true belief, and (3) true belief *meta logou*, translated by John Burnet as "accompanied by a rational account of itself or ground." Against the view that knowledge is true belief Plato made the point that lawyers can persuade juries to accept beliefs that are, in fact, true by using rhetorical devices but cannot be said to provide them with knowledge by doing so. Against the third definition, which, in effect, takes knowledge to be justified true belief, he pointed out that it is circular and regressive.

There is an obvious objection to the definition of knowledge as perception. Perception itself must be defined in terms of knowledge—namely, as the acquisition of knowledge about the external world by means of the senses. Plato's meaning here is perhaps better rendered by understanding his first definition to equate knowledge and sensation. Certainly this makes more plausible Plato's identification of this definition with Protagoras's thesis that man is the measure of all things (or that the truth for each man is simply what appears to him to be the case). In fact, Protagoras's thesis would be more accurately interpreted as the view that knowledge and belief are one and the same. This contention has obviously contradictory implications, as Plato pointed out. We all believe some beliefs of others to be truer than our own, and most people believe that Protagoras's theory is false. Something like that theory persists, however, in the view, to which we shall later return, that the foundations of empirical knowledge consist of incorrigible statements about immediate experience. According to this view, what we believe about our current sensations or experiences, whatever we may choose to say about them, is true. If it is also correct that such sensations are self-intimating, in the sense that they cannot occur without our knowing them to occur, it follows that every sensation is an item of knowledge though not that every item of knowledge is a sensation.

In his discussion of knowledge as true belief Plato raised the problem of false belief. How can we believe falsely that *X* is *Y* since if the belief is false, there is no *X* that is *Y* to form a belief, true or false, about? A false belief, it seems, is no belief at all. A perhaps oversimple solution to the problem is that we can know a thing *X* well enough to be able to identify it as a subject of discourse without knowing everything about it (whether, for

instance, it is *Y* or not-*Y*). This draws attention to the point that the objects of knowledge are not always propositional, that not all knowledge is knowledge that. In addition to the knowledge how emphasized by Ryle, there is knowledge with a direct object, or knowledge of, claimed in such remarks as "I know Jones" or "I know Paris."

A claim to know a person can be intended and understood in two main ways. In saying that I know Jones, I may mean that I have met him and that I could not recognize him (and, usually, that we have had enough to do with one another for him to remember me). On the other hand, I may mean that I know what his character is like, what sort of things he is likely to do. According to the first interpretation, very little knowing that is involved, although I should be expected to be capable of giving some description of Jones's appearance; according to the second, some knowledge that relating to his character is implied, but none about his past history, health, occupation, and so on is.

A claim to know a place is ordinarily a claim to knowledge how, to an ability to find one's way about in it. It is not enough simply to have been there. Among other individual objects of knowledge are games, languages, and works of art. The last of these kinds of knowledge can be treated in much the same way as knowledge of persons; the others, as cases of knowing how, as claims to the possession of a skill. In general, knowledge of can be reduced to varying mixtures of knowing how and knowing that, though by no single recipe. It never involves a claim to knowledge that of all the facts involving the individual in question. A further point against Plato is that I can know enough about an individual or a thing to be able to refer significantly and successfully to him or it without being in a position to say that I know him or it *simpliciter*. I know enough about Samarqand to refer to it as a city in Uzbekistan and to ascribe to it a degree of beauty, historical interest, and size, but I do not know Samarqand at all, for I have never been there and could not find my way about in it.

IS KNOWLEDGE DEFINABLE? The English philosopher John Cook Wilson (1849–1915), closely followed in this by his disciple H. A. Prichard (1871–1947), strenuously maintained that the concept of knowledge is primitive and indefinable. Against such idealist logicians as F. H. Bradley and Bernard Bosanquet, they argued that judgment is not a genus of which knowledge, belief, and opinion are species. A judgment, said Wilson, is the conclusion of an inference, but some knowledge must be uninferred. Nor is knowledge a kind or species of thinking or a species of belief, for belief rests on knowledge in that it requires that there should be both some known evidence for it and the knowledge that this evidence is insufficient. No doubt, belief usually does rest on evidence or what is taken to be evidence, but it is not, as Wilson supposed, necessary that it should do so. I may believe a woman to be married because I take her to be wearing a wedding ring. The fact that it is not a wedding ring that she is wearing does not in the least imply that I do not really believe what I infer from my mistake.

According to Prichard, knowledge is completely sui generis and cannot, as he put it, "be explained." We cannot, he said, derive knowledge from what is not knowledge. This observation, if it is relevant at all, is simply a dogmatic assertion of the indefinability of knowledge. We can certainly define some things in terms of what they are not; for instance, not all cats are kittens, and not all young things are kittens, but a kitten is by definition a young cat. Knowledge and belief, Prichard held, are utterly distinct and cannot be mistaken for each other. We know directly and infallibly whether our state of mind is one of knowledge or belief. If so, knowledge and belief could not be related as genus and species, although they could still be different species of the same genus, another possibility that Prichard ruled out. His view that the two cannot be mistaken for each other seems clearly mistaken. We often claim with complete sincerity to know things that turn out to be false in the end. In so doing, we have taken a belief, mistakenly, to be knowledge.

Is the opposite possibility ever realized? Do we ever take to be mere belief something that, in fact, we really know? Is there a difference between knowing something and knowing that we know it? Benedict de Spinoza held that there is not. "He who has a true idea, knows at that same time that he has a true idea, nor can he doubt concerning the truth of the thing" (*Ethics*, Part 2, Proposition 43). As Spinoza expressed it, the doctrine is plainly false. I can perfectly well have very little confidence in a belief that is really true if, for example, it has been communicated to me by a notoriously unreliable informant. In other words, I can have a belief that is really true without knowing that it is true. But can I know that something is the case without knowing that I know it? I can certainly have a justified true belief without knowing that that is what it is, for I may not realize that the grounds I have for believing it really do justify it. The question deserves a more thorough investigation than it can be given here.

RATIONALIST THEORY OF KNOWLEDGE. Plato's distinction between knowledge and belief has had a greater influence on the subsequent course of philosophy than his penetrating but unsuccessful attempts to find a definition of the concept. His essential point was that knowledge and belief are not only distinct attitudes but that they also have distinct and proprietary objects. Knowledge can be only of what is eternal and unchanging, of Forms, Ideas or universals; belief has for its objects the changing sensible particulars that make up the temporal world. Plato's reflections on mathematics seem to have led him to this conclusion. The propositions of geometry are preeminently objects of knowledge in that they can be established as conclusively true, once and for all, by demonstrative reasoning. Our beliefs about matters of temporal fact, on the other hand, are much more liable to illusion and error. The sensible objects of perceptual belief are infected with contradiction; they undergo change and have contrary properties at different times. But the objects of mathematical knowledge are wholly different. The circles and triangles studied by geometers are exact and perfect; they are ideals that the circular and triangular things we perceive with the senses approximate but always fall short of.

There are three ways in which a circular concrete thing may not be really circular. It may be circular at one time and elliptical at another; it may be other things (for example, green, cold, and sweet) as well as circular; and as concrete and sensible, it may not be strictly or perfectly circular. From these facts Plato concluded that such a thing is not wholly real in the way that the ideal circle of the geometer is. The ideal circle is a genuine object of knowledge, and only such wholly knowable things can be wholly real. From the distinction between knowledge and belief, then, Plato derived a distinction between two sorts of object, each sort constituting a separate world of its own—the abstract world of eternal Forms, which is the knowable reality, and the concrete world of changing particulars, which is only appearance, not nonexistent but not wholly real either, and of which one can have not knowledge but only belief.

Plato's arguments for the unknowability and unreality of concrete, sensible things are not very persuasive. If this once circular mat is now elliptical, it does not follow that it was not really circular before. If this circular object is also green and cold, that does not in any way detract from its circularity. Finally, even if it is not perfectly circular, it may be quite definitely green. In general, there would seem to be many propositions that are known by some people but only believed by others; a mathematician will know the truth of a proposition he has proved, whereas another person will simply believe it on his authority. Some things I now know I used only to believe—for instance, that I should be writing this here today; some things I now only believe I once used to know—for instance, where I bought my raincoat. These considerations show that the objects of knowledge and belief are not wholly mutually exclusive. But it may still be true that there are some things that can be only believed, whereas others can be both believed and known.

At the center of Plato's thinking about this subject is a principle that defines one important sense of the word *rationalism*—the principle that only necessary truths, established by a priori reasoning, can really be known. Something like this principle was accepted by Aristotle, although he rejected Plato's doctrine that Forms or universals occupied a separate abstract world of their own beyond time and space. Aristotle agreed that only the form of things could be known and that the matter that individuated or particularized them was beyond the reach of knowledge. For him true knowledge was to be attained by a process of intuitive induction that discerned the necessary connections between the forms present in concrete things. A science or ordered body of knowledge must consist of propositions deduced from self-evident first principles of this kind.

Descartes's rationalism was inspired by the reflection that ordinary claims to knowledge often prove mistaken. True knowledge, he insisted, must be objectively certain and impossible to doubt. His methodical endeavors to doubt everything were brought up short by the celebrated "I think, therefore I exist." I cannot doubt that I doubt, for in the act of doubting it I prove it to be true; if I doubt, I think; and if I think, I exist. What, he then inquired, is so special about *cogito* and *sum*? What makes them so indubitably certain? His unhelpful conclusion is that they are clearly and distinctly perceived to be true. What he meant by this weakly formulated criterion of certainty can best be discovered by seeing what, in practice, he took it to certify. It appears that two sorts of proposition are clearly and distinctly perceived to be true: (1) necessary truths whose denial is self-evidently contradictory and (2) the immediate deliverances of sensation and introspection about one's own current mental state. Premises of both kinds figure in his first proof of God's existence:

> Every event must have an adequate cause.
> I have a clear and distinct idea of God.
> God alone is an adequate cause for my idea of him.
> Therefore, God exists.

In fact, *cogito*, I think, is not a clear instance of either, let alone both, of these two kinds of knowable, and even if it were, it would not follow from its being, on one hand, necessary and immediate and, on the other, certain that anything else that was necessary and immediate was also certain. Descartes's primary certainty was perhaps first thought of on a Thursday, but it does not follow that anything first thought of on a Thursday either by him or by anyone else is certain, too. It is not a necessary truth that I think or exist, for I might not be awake and might never have existed. If this is the case, the facts in question could not, of course, have been expressed in the first person singular.

Locke, despite his justly recognized position as a founding father of empiricism, reached much the same rationalist conclusion as Descartes, although by a very different route. He defined knowledge as "the perception of the agreement or disagreement of two ideas" (*Essay concerning Human Understanding*, Book 4, Ch. 1, Sec. 2). He went on to distinguish three kinds of knowledge: (1) intuitive knowledge of such things as the fact that red is not green and the fact of one's own existence; (2) demonstrative knowledge, which includes mathematics, morality, and the existence of God; and (3) sensitive knowledge, which is concerned with "the particular existence of finite beings without us." The third type of knowledge does not conform to his general definition, as he admitted. To become aware of a finite being outside us, we have to infer the existence of something that is not an idea from the ideas of sensation we take it to cause, and in part, to resemble. Locke's definition, as he understood it, restricts knowledge to the domain of a priori necessary truths. In intuition and demonstration there is a direct or indirect awareness of the connection between ideas present to the mind. But in the third case a connection is asserted between an idea of sensation and a physical thing that is not and cannot be directly present to the mind.

Locke did not introduce a special category to accommodate our knowledge of the ideas we passively experience but remitted them in passing to the category of intuitive knowledge. This sort of knowledge is quite unlike his exemplary cases of intuition, being contingent and empirical where the exemplary cases are necessary and a priori, and he might well have introduced a special category of reflective knowledge to accommodate it. It would comprise assertions of the connection of particular ideas, whereas intuition and demonstration would cover the connections of abstract, general ideas. Thus, although Locke's official definition of knowledge confines its application to necessary truths, it could, with a little modification, have been extended to cover a person's awareness of the present contents of his mind. But it could not, by any contortions, have been made to cover sensitive knowledge of real existence, that empirical knowledge par excellence which it was Locke's avowed purpose to justify and explain.

CERTAINTY. The indestructible vitality of the rationalist theory that necessary truths alone or necessary truths and reports of immediate experience are really knowledge was proved by its wide acceptance among empirically minded philosophers of the twentieth century—for example, Russell, C. I. Lewis, and A. J. Ayer. In support of it a powerful battery of arguments was produced, designed to show that despite the subjective certainty we feel in many kinds of belief, they cannot count as knowledge because they are not objectively certain.

Russell contended that all the sources of what we ordinarily regard as common knowledge of fact are in some degree untrustworthy. Perception is tainted by illusions, hallucinations, and dreams. Memory is notoriously fallible. Testimony, which plays such a large part in building up the social fabric of belief, presupposes an inference to other minds that is inevitably shaky and conjectural. Induction never certifies its conclusions, imparting at best only a measure of probability to them. Even introspection, if it is held to convey information about the self as a continuing personality, goes beyond what is directly present to the mind. Only what is directly present to it—currently occurring thoughts and feelings—is the object of certain, infallible, and indubitable belief.

Lewis generalized Russell's position by distinguishing expressive judgments that report current states of mind from all other empirical propositions on the ground that they alone are wholly nonpredictive and have no implications about future observable happenings by whose failure to occur they might be refuted. Ayer, at one time, went even further. He held that all contingent, empirical propositions whatsoever, including reports of immediate experience, are uncertain on the ground that every such proposition involves the application of a general predicative term to its subject and thus makes a comparison with previous and perhaps faultily remembered instances of the term's application.

This kind of fallibilism about empirical belief was doggedly resisted by G. E. Moore and, after him, by Ludwig Wittgenstein, J. L. Austin, and Norman Malcolm. Moore's main point was that the word *certain* is learned and thus acquires its meaning from such situations as that in which a man holds up his hand and makes the

perceptual judgment "I know for certain that this is a hand." Some rather subtler arguments are sketched in his book *Philosophical Papers*. Their general upshot is that the rationalists and fallibilists have been working with an unconsidered and excessively stringent concept of certainty. They have simply taken it for granted that for a belief to be certain, it must be impossible to doubt it. Russell, for example, began his search for certain knowledge with the question "Is there any knowledge in the world which is so certain that no reasonable man could doubt it?"

There are at least four senses in which it may be held that a belief cannot be doubted. The first is psychological; a man cannot doubt a belief if he cannot, in fact, bring himself to suspend judgment about it. This kind of certainty will vary from person to person and is of no direct philosophical interest. The second sense is logical. Here "doubt" is taken to mean "suppose false" and "can" to mean "can without logical inconsistency." This yields the strict rationalist view, since only necessary truths cannot be supposed false without inconsistency. A third sense identifies certainty with incorrigibility. According to it, a belief cannot be doubted if its truth follows from the fact that it is believed. Anyone who doubts an incorrigible belief shows that he does not understand the words that express it. The favorite examples of incorrigible beliefs are reports of immediate experience, such as "I am in pain" or "It seems to me now that there is a table here." But the notion would also apply to the more elementary and intuitive kind of necessary truth, such as the law of contradiction. Finally, there is the concept of certainty that, say Moore and his adherents, we actually employ in common speech where it means what cannot reasonably be doubted or supposed false. That people make all sorts of mistakes is not, according to this view, a reason for doubting the truth of a particular proposition. What is required to justify doubt is that propositions just like this, made in circumstances just like these and resting on just this kind of evidence, have in the past turned out to be mistaken. In this sense of certainty many beliefs based on perception, memory, testimony, and induction are objectively certain and thus properly regarded as items of knowledge. This view has the merit of allowing that many propositions that are, in fact, necessary truths are or once were less than certain, and it does not require the theory that there are any incorrigible propositions to be accepted. A further point in its favor is that such surprising theses as the one that no factual belief is certain can surprise us and escape triviality only if they are taken in this sense.

SOME MODERN VIEWS. In the mid-twentieth century, philosophical discussions of knowledge were much concerned with three distinctions drawn by Russell, Ryle, and Austin that must be briefly mentioned.

Acquaintance and description. In Russell's early writings he drew a distinction between knowledge of things and knowledge of truths, between knowledge of and knowledge that, a distinction marked in French by the verbs *connaître* and *savoir*. Within each kind he also discerned a distinction between an immediate and a derived form. Immediate knowledge of truths is conveyed in intuitive statements—for example, basic judgments of perception and the axioms of logic and mathematics; derivative knowledge of truths, in demonstrable necessary propositions and inferred empirical statements. Parallel to this on the side of knowledge of things is the distinction between knowledge by acquaintance and knowledge by description.

Acquaintance, as Russell defined it, is the converse of presentation; it is the direct and infallible apprehension of some sort of object. But objects of description, unlike those of acquaintance, can fail to exist. Russell held that we are acquainted with present and past particulars and also with universals. This doctrine has led to a good deal of confusion. Certainly we do know things, persons, and places by acquaintance, but to do so is generally to know that something is true of them and is at least to know how to recognize them. The words with which we refer to things we are not acquainted with can be defined or explained in terms of those connected with objects of acquaintance. But this produces understanding rather than knowledge, understanding of singular terms (whether what they purport to refer to exists or not) and of general terms (whether or not there is anything they apply to). Russell's principle of acquaintance ("Every proposition which we can understand must be composed wholly of constituents with which we are acquainted") is really a version of the empiricist theory of meaning. Asserted without qualification, it is highly unplausible. We are not acquainted with anything corresponding to the "if" that occurs in the verbal expression of a hypothetical proposition although we understand the word. In general, to become acquainted with things is to acquire some intuitive knowledge of truths in which they figure, particular objects of acquaintance being the subjects of such truths and universal objects of acquaintance their predicates. In other words, knowledge of things cannot be separated from and regarded as prior to knowledge of truths in the way Russell supposed.

Knowing how and knowing that. Ryle's distinction between knowing how and knowing that has already been mentioned. There is a parallel distinction between remembering how and remembering that (there is also memory *of* past events). Ryle is anxious to correct the intellectualist bias of theorists of knowledge and to draw attention to the dispositional nature of all kinds of knowledge and belief; we speak, after all, of the knowledge and beliefs of those who are fast asleep. He tends to suggest that knowing that is a special, verbal form of knowing how, that it consists in having learned how to answer certain questions and now being ready to answer them.

Performative and descriptive verbs. John Austin's work on performative utterances has interested many philosophers in that class of verbs that are used in the first person present to do things rather than to describe what is being done. Examples of such performative verbs are "promise," "swear," "take thee, *X*, to be my wedded wife," and "name this ship *Y*." A verb ϕ is performative if it follows that I ϕ from the fact that I say, "I ϕ." Austin appears to have thought, wrongly, that "know" is a verb of this kind and that its function is to guarantee or authorize the acceptance of the piece of information that followed it. It is true that to prefix "I know" to a statement of fact does not add much to its content. But *p* and "I know that *p*" are not equivalent, since the former may be true when the latter is false. Austin was right in denying that knowledge is a state of assurance stronger than the most assured belief, though it is not clear that anyone ever supposed that it was. But the correctness of this denial, although it entails that it is not some describable psychological feature of the knower's state of mind that differentiates knowledge from belief, does not entail that the difference is not at all describable and lies, rather, in some nondescriptive function that the word performs.

THE NATURE OF BELIEF

Most philosophers who have in any way adverted to the nature of belief have assumed that belief is an inner state of mind, directly accessible to introspection and distinct from, though causally related to, the believer's behavior. In *The Emotions and the Will* (1859) the Scottish philosopher Alexander Bain proposed that belief should be defined in terms of behavior: "Belief has no meaning except in reference to our actions … no mere conception that does not directly or indirectly implicate our voluntary exertions can ever amount to the state in question." In support of Bain's theory is the fact that not only can others check our claims to believe by considering whether we behave appropriately but we ourselves may also take the results of such a test to overrule claims to believe that we have sincerely made.

Careful statements of the opposing doctrines were given by H. H. Price and R. B. Braithwaite. Price's mentalist definition of belief equates it with entertainment of a proposition together with assent. To entertain a proposition is to understand and attend to its meaning; when it occurs by itself, it is neutral and uncommitted as regards the proposition's truth or falsehood. Price breaks assent down into a volitional and an emotional part. He describes the volitional element as a mental act of preferring a proposition to any incompatible alternatives that have occurred to one; the emotional element is a feeling of conviction or assurance and may vary in degree. Braithwaite identifies belief in a proposition with its entertainment together with a dispositional readiness to act as if it were true. "Being ready to act as if *p* were true" has at first sight a suggestion of circularity, for it seems to mean being ready to act as if one believed *p*. But this can be avoided. I act as if *p* were true if I act in a way that would satisfy my desires if *p* were in fact true.

Against both theories it should be said that "entertainment" is dispensable if the normal sense of "believe" is in question, for we attend consciously to the propositions we believe only at rare intervals. As regards Price, what is to be understood by an act of preferring as opposed to an emotion of preference? It looks very like the silent assertion of the proposition itself, an inner rehearsal of a piece of outward verbal behavior. Second, feelings of conviction do not always attend even the beliefs we consciously entertain. Unless our confident beliefs are actually challenged, our state would seem to be one of easy and unemotional taking for granted.

Against the view of Bain and Braithwaite it has been urged by Mill, Franz Brentano, and Russell that if a belief has behavioral effects different from mere entertainment, it must differ in its intrinsic mental character. This is a misunderstanding. For a behaviorist there is a difference in the dispositions of one who believes and of one who merely entertains a proposition. A more serious difficulty is presented by beliefs that have negligible practical consequences, such as those about remote historical or astronomical events. But even here there is a disposition to verbal behavior, and, again, a disposition can exist without being actualized. There is also the difficulty that my claims about what I believe become, according to this theory, inductive conjectures about what I should do if certain circumstances arose. One reply is that not all inductive conjectures are conjectural to that degree. I

need not, for example, feel very hesitant about what would happen if this iron table were dropped on that china teapot. Braithwaite adds that his theory has the merit of making possible rather precise measurements of subjective probability or degree of belief. The numerical probability I attach to a belief can be regarded as the least favorable odds I should accept on its turning out to be correct. Thus, unless I accept an odds-on bet, I do not believe something more than I believe its denial.

There is an interesting and extreme opposition in the history of philosophy between Descartes, who held that assent is a matter of will that can be freely given or withheld, and Hume, who represented us as largely passive in belief, which he conceived as a feeling that we find ourselves with and must put up with whether we like it or not, much as we find ourselves equipped with desires and aversions. Descartes's activism is shown first in his proposal that the philosopher should undertake a course of methodical doubt, suspending judgment about all the beliefs he has hitherto taken for granted. It reaches its fullest development in his attempt to solve the theological problem of error or intellectual evil, to reconcile the fact, on which his whole philosophy depends, that many of our beliefs are false with the goodness of God. The solution he offered is that God has fitted us out with limited intellects, appropriate to our earthly needs, but in his own image, with unrestricted freedom of will. When we make mistakes it is because we have culpably given free assent to propositions beyond the effective reach of our limited intellects.

In Descartes's favor is the fact that we do assess beliefs as more or less reasonable, a practice whose theory is logic and methodology. And the ethics of belief has not always been confined to distinguishing logically reasonable beliefs from others. It has often been held that some beliefs—in the existence of God, for example—are morally obligatory, and some beliefs are often recommended as prudent or useful. Hume himself propounded rules for judging causes and effects whose acceptance, he maintained, will enable us to advance science and avoid superstition. On Hume's side is the fact that it seems no more possible to resolve to believe something one actually does not believe than it is to increase one's height or eradicate one's distaste for endives by a simple effort of will. What one can do is to fortify or undermine one's belief in a proposition indirectly by voluntarily concentrating one's attention on the evidence for or against it.

It is quite commonly said that belief must rest on evidence and sometimes, especially by those who hold knowledge to be indefinable, that it must rest on knowledge. It is certainly usual for belief to rest on something the believer regards as evidence, whether or not it is true and whether or not it lends any support to the belief in question. But a wildly dogmatic or superstitious belief, maintained in the teeth of all the evidence, is still a belief, however unreasonable it may be.

FAITH. There is some point to the malicious definition of faith as firm belief in something for which there is no evidence, for faith does involve a measure of risk, a voluntary decision to repose more confidence in a proposition, person, or institution than the statable grounds for doing so would, if neutrally considered, justify. Locke defined faith as resting on authoritative testimony, "the assent to any proposition, not thus made out by the deductions of reason, but upon the credit of the proposer." This applies well enough to the religious faith of traditional Christianity, but it is too narrow to cover the general use of the concept. It is often said that science rests on faith in the uniformity and intelligibility of nature as much as religion does on an undemonstrable conviction that the world is under the direction of a wise and benevolent intelligence. Certainly, science would be wholly sterilized if men were not prepared to consider adventurous and unjustified hypotheses. But it is not obvious that these adventurous conjectures have to be believed by their propounders. The austere maxim of W. K. Clifford—"It is wrong, everywhere and for anyone, to believe anything upon insufficient evidence"—is not strictly incompatible with intellectual enterprise. Yet even Popper, who of all theorists of knowledge is most insistent on the conjectural and fallible nature of science, admits that "our guesses are guided by the unscientific, the metaphysical (though biologically explicable) faith in laws, in regularities which we can uncover."

See also A Priori and A Posteriori; Evans, Gareth; Kant, Immanuel; Knowledge, A Priori; Kripke, Saul; Meaning; Plantinga, Alvin; Propositions; Putnam, Hilary; Reference.

Bibliography

CLASSIC DISCUSSIONS

Bain, Alexander. *The Emotions and the Will*. London: J.W. Parker, 1859.

Descartes, René. *Meditations*.

Locke, John. *Essay concerning Human Understanding*. Book 4.

Plato. *Republic*. Books 5–7.

Plato. *Theaetetus*.

MODERN ACCOUNTS

Armstrong, D. M. *Belief, Truth, and Knowledge*. Cambridge, U.K.: Cambridge University Press, 1973.

Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knowledge.* 2nd ed. New York: Routledge, 2003.

Austin, John. "Other Minds." In his *Philosophical Papers*. London: Oxford University Press, 1962.

Ayer, A. J. *The Foundations of Empirical Knowledge*. London: Macmillan, 1940.

Ayer, A. J. *The Problem of Knowledge*. London: Macmillan, 1956. Chs. 1–2.

BonJour, Lawrence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Brewer, Bill. *Perception and Reason*. Oxford: Clarendon Press, 2002.

Chisholm, Roderick. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982.

Chisholm, Roderick M. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957. Part I.

Dretske, Fred. *Seeing and Knowing*. London: Routledge and Kegan Paul, 1969.

Foley, Richard. *Working without a Net*. New York: Oxford University Press, 1993.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Harman, Gil. *Thought*. Princeton, NJ: Princeton University Press, 1973.

Hintikka, Jaakko. *Knowledge and Belief*. Ithaca, NY: Cornell University Press, 1962.

Klein, Peter. *Certainty: A Refutation of Scepticism*. Minneapolis: University of Minnesota Press, 1981.

Lehrer, Keith. *Theory of Knowledge*. Boulder, CO: Westview Press, 1990.

Lewis, C. I. *Mind and the World-Order*. New York: Scribners, 1929. Ch. 9.

Malcolm, Norman. *Knowledge and Certainty*. Englewood Cliffs, NJ: Prentice Hall, 1963.

Moore, G. E. *Philosophical Papers*. London: Allen and Unwin, 1959.

Plantinga, Alvin. *Warrant and Proper Function*. Oxford: Oxford University Press, 1993.

Popper, Karl R. *Conjectures and Refutations*. London: Routledge, 1963. See the introduction.

Prichard, H. A. *Knowledge and Perception*. Oxford: Clarendon Press, 1950.

Russell, Bertrand. *Problems of Philosophy*. London: Williams and Norgate, 1912. Chs. 5, 13.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949. Ch. 2.

Sosa, Ernest. *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Wilson, John Cook. *Statement and Inference*. Oxford: Clarendon Press, 1926. Part 1, Chs. 2, 4; Part II, Chs. 1–3.

Woozley, A. D. *Theory of Knowledge: An Introduction*. London: Hutchinson, 1949. Ch. 8.

Anthony Quinton (1967)
Bibliography updated by Benjamin Fiedor (2005)

KNOWLEDGE AND MODALITY

The prominence of the modalities (i.e., necessity and contingency) in epistemological discussions is due to the influence of Immanuel Kant (1965), who maintained that:

(1) All knowledge of necessary propositions is a priori; and

(2) All propositions known a priori are necessary.

Saul Kripke (1971, 1980) renewed interest in Kant's account of the relationship between the a priori and the necessary by arguing that some necessary propositions are known a posteriori and some contingent propositions are known a priori. A cogent assessment of the controversy requires some preliminary clarification.

The distinction between necessary and contingent propositions is metaphysical. A necessarily true (false) proposition is one that is true (false) and cannot be false (true). The distinction between a priori and a posteriori knowledge is epistemic. S knows a priori that p just in case: (a) S knows that p; and (b) S's justification for believing that p does not depend on experience. Condition (b) is controversial. On the traditional reading, (b) is equivalent to (c): S's belief that p is nonexperientially justified. Hilary Putnam (1983) and Philip Kitcher (1983), however, argue that (b) is equivalent to (d): S's belief that p is nonexperientially justified and cannot be defeated by experience. Albert Casullo (2003) rejects the Putnam-Kitcher reading on the grounds that it yields an analysis of a priori knowledge that excludes the possibility that someone knows a posteriori a proposition that can be known a priori.

The expression "knowledge of necessary propositions" in (1) is ambiguous. The following definitions remove the ambiguity:

(A) S knows the *general modal status* of p just in case S knows that p is a necessary proposition (i.e., either necessarily true or necessarily false) or S knows that p is a contingent proposition (i.e., either contingently true or contingently false);

(B) S knows the *truth value* of p just in case S knows that p is true or S knows that p is false (assuming truth is always bivalent);

(C) S knows the *specific modal status* of p just in case S knows that p is necessarily true or S knows that p is necessarily false or S knows that p is contingently true or S knows that p is contingently false.

(A) and (B) are logically independent. One can know that Goldbach's Conjecture is a necessary proposition but not know whether it is true or false. Alternatively, one can know that some mathematical proposition is true but not know whether it is a necessary proposition or a contingent proposition. (C), however, is not independent of (A) and (B). One cannot know the specific modal status of a proposition unless one knows both its general modal status and its truth value.

(1) is crucial for Kant, because it is the leading premise of his only argument in support of the existence of a priori knowledge:

(1) All knowledge of necessary propositions is a priori.

(3) Mathematical propositions, such as that 7 + 5 = 12, are necessary.

(4) Therefore, knowledge of mathematical propositions, such as that 7 + 5 = 12, is a priori.

(1), however, is ambiguous. There are two ways of reading it:

(1T) All knowledge of the *truth value* of necessary propositions is a priori, or

(1G) All knowledge of the *general modal status* of necessary propositions is a priori.

The argument is valid only if (1) is read as (1T). Kant, however, supports (1) with the observation that although experience teaches that something is so and so, it does not teach us that it cannot be otherwise. Taken at face value, this observation states that experience teaches us that a proposition is true and that experience does not teach us that it is necessary. This supports (1G), not (1T).

Kripke rejects (1) by offering examples of necessary truths that are alleged to be known a posteriori. First, he maintains that if P is an identity statement between names, such as "Hesperus = Phosphorus," or a statement asserting that an object has an essential property, such as "This table is made of wood," then one knows a priori that:

(5) If P then necessarily P.

Second, he argues that because one knows by empirical investigation that Hesperus = Phosphorus and that this table is made of wood, one knows a posteriori that:

(6) P.

Kripke concludes that one knows by *modus ponens* that:

(7) Necessarily P.

(7) is known a posteriori because it is based on (6), which is known a posteriori.

How do Kripke's examples bear on (1)? Once again, a distinction must be made between (1G) and (1T). Kripke's examples, if cogent, establish that (1T) is false: They establish that one knows a posteriori that some necessary propositions are true. They do not, however, establish that (1G) is false: They do not establish that one knows a posteriori that some necessary propositions are necessary. It may appear that Kripke's conclusion that one has a posteriori knowledge that necessarily P entails that (1G) is false. Here a distinction must be made between (1G) and:

(1S) All knowledge of the *specific modal status* of necessary propositions is a priori.

Kripke's examples establish that (1S) is false: They establish that one knows a posteriori that some necessary propositions are necessarily true. Because knowledge of the specific modal status of a proposition is the conjunction of knowledge of its general modal status and knowledge of its truth value, it follows from the fact that one's knowledge of the truth value of P is a posteriori that one's knowledge of its specific modal status is also a posteriori. However, from the fact that one's knowledge of the specific modal status of P is a posteriori, it does not follow that one's knowledge of its general modal status is also a posteriori.

(1G) has not gone unchallenged. Kitcher (1983) argues that even if knowledge of the general modal status of propositions is justified by nonexperiential evidence, such as the results of abstract reasoning or thought experiments, it does not follow that such knowledge is a priori because the nonexperiential justification in question can be defeated by experience. Casullo (2003) rejects (1G) on the grounds that the Kantian contention that experience can provide knowledge of only the actual world overlooks the fact that much practical and scientific knowledge involves counterfactual conditionals, which provide information that goes beyond what is true of the actual world.

Kripke also argues that some contingent truths are known a priori. His examples are based on the observation that a definite description can be employed to fix the reference—as opposed to give the meaning—of a term. Consider someone who employs the definition description "the length of S at t_0" to fix the reference of the expression "one meter." Kripke maintains that this person knows, without further empirical investigation, that S is one meter long at t_0. Yet the statement is contingent

because "one meter" rigidly designates the length that is in fact the length of S at t_0 but, under different conditions, S would have had a different length at t_0. In reply, Alvin Plantinga (1974) and Keith Donnellan (1979) contend that, without empirical investigation, the reference fixer knows that the sentence "S is one meter long at t_0" expresses a truth, though not the truth that it expresses. Gareth Evans (1979) disputes this contention.

Bibliography

Casullo, Albert. *A Priori Justification*. New York: Oxford University Press, 2003.

Donnellan, Keith S. "The Contingent *A Priori* and Rigid Designators." In *Contemporary Perspectives on the Philosophy of Language*, edited by P. French et al. Minneapolis: University of Minnesota Press, 1979.

Evans, Gareth. "Reference and Contingency." *Monist* 62 (1979): 161–189.

Kant, Immanuel. *Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: St. Martin's Press, 1965.

Kitcher, Philip. *The Nature of Mathematical Knowledge*. New York: Oxford University Press, 1983.

Kripke, Saul. "Identity and Necessity." In *Identity and Individuation*, edited by M. K. Munitz. New York: New York University Press, 1971.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Plantinga, Alvin. *The Nature of Necessity*. Oxford: Oxford University Press, 1974.

Putnam, Hilary. "'Two Dogmas' Revisited." In *Realism and Reason: Philosophical Papers*. Vol. 3. Cambridge, U.K.: Cambridge University Press, 1983.

Albert Casullo (1996, 2005)

KNOWLEDGE AND TRUTH, THE VALUE OF

Questions concerning the value of knowledge and truth range from those that suggest complete skepticism about such value to those that reflect more discriminating concerns about the precise nature of the value in question and the comparative judgment that one of the two is more valuable than the other.

THE COMPARATIVE QUESTION AND THE PRAGMATIC ACCOUNT

The history of epistemology has its conceptual roots in the dialogues of Plato, and the question of the value of knowledge and truth arises there as well. In Plato's *Meno*, Socrates and Meno discuss a number of issues, including the issue of the nature and value of knowledge. Socrates raises the question of the value of knowledge, and Meno answers by proposing a pragmatic theory: knowledge is valuable because it gets us what we want. Socrates immediately proposes a counterexample, to the effect that true opinion would work just as well: If you want to get to Larissa, hiring a guide who has a true opinion of how to get there will have the same practical results as hiring a guide who knows the way. Meno then voices a philosophically deep perplexity, wondering aloud why knowledge should be more prized than true opinion and whether there is any difference between the two. Meno thus questions two assumptions, the first being the assumption that knowledge is more valuable than true opinion, and the second that knowledge is something more than true opinion.

Socrates's counterexample suggests another: If you want to get to Larissa, it matters not whether your guide has true opinion or merely empirically adequate views on the matter. To see the counterexample, we need to understand that an empirically adequate theory is one that "saves the appearances," in other words, one that would never be refuted by any sensory experience. The simplest way to see that such a theory is not the same thing as a true theory is to consider skeptical scenarios such as René Descartes's evil demon world. The denizens of such a world will have roughly the same views as we do, and their views will be as empirically adequate as ours. Since the demon is so skillful at carrying out his intentions, however, their views will be false even if ours are true. In such a world, there are no guides with true opinions about how to get to Larissa. Instead, the best one could hope for is a guide who has an empirically adequate view of the matter. Yet, if we compare the two situations, the one in the actual world where the hired guide has a true opinion, and the one in the demon world where the hired guide has only an empirically adequate opinion, no suffering accrues to the traveler in the demon world that does not also accrue to the traveler in the actual world, and no benefits are experienced by the traveler in the demon world that are not also experienced by the traveler in the actual world. That is to say, their experiences are indistinguishable, leaving us to wonder what practical advantage truth has over empirical adequacy.

SKEPTICISM ABOUT THE VALUE OF KNOWLEDGE AND TRUTH

Besides this Platonic threat to the value of knowledge and truth, there are other threats. One arises from the specter of skepticism. If we grant that there is no adequate answer to the skeptic, we might have the experience of philo-

sophical sour grapes, denying the value of what we cannot have.

More respectable threats to the value of knowledge and truth come from positions that question the ordinary thinking that knowledge and truth contribute to well-being. Pyrrhonian skepticism maintains that such ordinary thinking is mistaken, and that the path to happiness requires abandoning a search for knowledge and truth, ridding oneself of beliefs and instead "acquiescing to the appearances." Arguments for skepticism play an important role in this process insofar as they can play a role in eliminating the dogmatism purportedly inherent in belief, but the Pyrrhonian appeal to skepticism is not simply that of philosophical sour grapes: it is motivated instead by a conception of what human well-being involves and requires.

There is no question that the Pyrrhonian school was sensitive to a real threat to human happiness, for dogmatism has caused immense suffering (for one monumental example, think of the suffering caused by religious wars). It is philosophical overkill, however, to move from such obvious points to skepticism and a denigration of the value of knowledge and truth. For one thing, dogmatism is compatible with a full appreciation of the rights of other human beings and so need not lead to massive human rights violations. Moreover, even if dogmatism has practical consequences that are troubling, a defender of the value of knowledge and truth has a counterargument here. The typical epistemological approach involves abstracting away from the causal consequences of holding the beliefs in question, concerning itself more with intrinsic features of cognition, the kind reflected in talk of inquiry for its own sake. When we engage in inquiry for its own sake, successful results will partake of a kind of success that is independent of any causal contribution to well-being or other practical concerns. When epistemologists reflect on the nature of successful cognition and the extent to which an organism achieves it, the predominant approach has been to reflect on a kind of success that abstracts from the consequences of cognition, whether those consequences are practical, moral, religious, political, or social.

Given such an abstraction, a defender of the value of knowledge and truth can argue that even if Pyrrhonism is correct as a general approach to cognition, it fails to show that, from the abstract point of view of what is involved in inquiry for its own sake, knowledge and truth are not valuable. One of the factors to be considered in evaluating the plausibility of any view regarding the all-things-considered value of knowledge and truth is the perspectival value of these things, such as the value they (appear to) have from the perspective of inquiry for its own sake.

Moreover, the argument for Pyrrhonism as the best view of the all-things-considered value of knowledge and belief is weak. To the extent that dogmatism itself has untoward consequences, the proper remedy is a sense of human fallibility, and only a highly questionable theory in which knowledge must be infallible could view skepticism as the only antidote to dogmatism.

Another threat arose in the latter half of the twentieth century, from those whom Bernard Williams in his last major philosophical work (2002) labeled "deniers" of the value of truth. Some of these deniers claim, in postmodernist spirit, that the ideals of truth and objectivity in inquiry are pretensions in service of other, baser motives. Problems for such denials of the value of truth arise when attempts are made to delineate accurately the nature of the pretensions in question and the lessons to be learned about the human condition from such investigation. Some, such as Richard Rorty (1989), have sought to espouse views while at the same time denying their accuracy, but such a position is not intellectually stable. The instability of the view is masked by the false dilemma involved in always capitalizing terms like "Truth" and "Reality" to gain purchase for the view that these concepts always and everywhere posit a metaphysical space hidden behind the pale of language or experience, yielding the claim that inquiry should aim at something weaker than truth, such as widest possible agreement (see Rorty 1998). As Williams points out, however, it makes little sense to value the number of converts to a view unless convincing them of the view has something to do with convincing them that the view is true. Put more generally, among the regulating ideas concerning truth is that there is an obvious logical equivalence between *p* and *it is true that p*, so that to assert a claim is to represent that claim as being true, and no philosophical sleight of hand involving capitalization of terms or scare-quotes, to which such deniers are prone, undermines this central point about truth. The deniers may have useful and important critiques of pretensions to objectivity, but it is a fundamental principle of inquiry that claims and arguments that are self-refuting should be avoided.

THE NATURE OF THE VALUE IN QUESTION

So there are three primary questions regarding the value of knowledge and truth. The first is whether knowledge and truth are valuable, all things considered. The second

question is whether they are valuable from the abstract point of view of what is involved in inquiry for its own sake. And the third question pertains to the issue of explanation, asking whether it is really knowledge that is valuable from this purely cognitive point of view, or something else instead.

The first question is a very large one, but a proper answer to it depends on answers to the second two questions, for if knowledge and truth do not pass scrutiny when considered from the purely cognitive point of view, then they have little to be said in their favor from an all-things-considered point of view. Furthermore, a negative answer to the third question would threaten the significance of a positive answer to the second question.

THE VALUE OF TRUTH. The major concerns involved in the third question are whether knowledge is more valuable than its parts and whether truth has anything to be said on its behalf over mere empirical adequacy. From a purely cognitive point of view, as William James (1956) noted, human beings are motivated by two primary concerns, a concern for not being duped and a concern for not missing out on something important. The first concern is relevant to the issue regarding whether truth has anything to be said on its behalf over mere empirical adequacy. If we adopt the literary device of a narrator commenting on various scenarios, we find something of an answer to this question. If one of the scenarios is the evil demon world and the other the actual world (as we suppose it to be), with the narrator being the very same person in each of these scenarios, the narrative will almost certainly treat the evil demon scenario as disturbing in comparison to the actual scenario, precisely because the narrator is being duped in the former but not in the latter. The most straightforward explanation of this response is that we find getting to the truth intrinsically valuable in virtue of our concern for not being duped.

The second concern above, the concern for not missing out on something important, raises a further problem, the problem of whether all truth is intrinsically valuable or only the important truths (see Ernest Sosa 2003). It is certainly true that we view some truths as simply unimportant, but that fact need not be taken to undermine the intrinsic value of truth, for it may be that our practical needs, goals, and interests interact with the intrinsic value of truth so that some truths are simply unimportant, all things considered, even though truth is still intrinsically valuable from a purely cognitive point of view, or from the point of view of inquiry for its own sake.

THE VALUE OF KNOWLEDGE. The value of truth raises the question of whether knowledge is more valuable than the sum of its parts; an affirmative answer to this question faces serious obstacles. Note first the variety of ways in which one might defend the value of knowledge. After seeing the above defense of the value of truth, an obvious response would be to argue that knowledge is intrinsically valuable, valuable independently of any value possessed by its parts, and more valuable intrinsically than any collection of its parts. It is instructive to note that such a maneuver is not as promising here as it is in the case of truth. On the one hand, when asked, "Why, from a purely cognitive point of view, do you value truth?" we are hard pressed to say anything informative at all, and this difficulty is an indication that we do not value truth on the basis of our valuing something else, but rather that we value it intrinsically. On the other hand, when asked, "Why, from a purely cognitive point of view, do you value knowledge?" we are inclined to answer. Our answer might be that we want to be correct, but not merely by accident, as happens when one has merely a true belief. The inclination to answer in ways such as this suggests that we value knowledge in a way that is different from the way in which we value truth, that even if truth is intrinsically valuable, knowledge is valuable because of the features that distinguish it from true belief.

What are these features? The traditional view is that knowledge is true belief that is justified, but the literature deriving from Edmund Gettier's seminal paper of 1963 shows that no fallibilist view about justification can accept this account of knowledge. Fallibilism about justification is the view that justified false beliefs are possible, perhaps clarified in terms of the claim that no matter how good our evidence is for what we believe, we might still be wrong. Given this view, it turns out to be unavoidable that there could be cases of justified true belief that are not cases of knowledge. Hence another condition—a fourth condition—must be added.

Justification and knowledge. We should expect to find the value of knowledge, then, by examining the value of the additional elements of knowledge—justification and whatever fourth condition is needed. The standard conception of justification makes it difficult to use in a defense of the value of knowledge, however. The standard conception of justification is teleological: holding justified beliefs is the proper means to adopt when one's goal is to get to the truth (and avoid error). If we think of means to a goal in terms of that which makes achieving the goal likely, the standard conception of justification amounts to the idea that justification is a property of a

belief in virtue of which that belief is objectively likely to be true.

A theory will need to say something different from the simple claim that justification is to be understood in terms of objective likelihood of truth, however, if it is to have any hope of providing a basis for explaining the value of knowledge over the value of its parts. Recall that the task is to explain the value of knowledge over that of true belief, so if an appeal to justification is to aid in this task, the theory of justification provided must support the idea that justified true belief is more valuable than mere true belief. It is not enough simply that justification is a valuable property for a belief to have, for that result would only show that justified belief is more valuable than unjustified belief, not that justified true belief is more valuable than true belief. Another way to put this point is as follows: It is necessary for justification to be valuable for it to play a role in explaining the value of knowledge, but its having such value is not by itself sufficient for it to play such a role.

The reason the value of justification is not sufficient is because of the swamping problem, as explained by Linda Zagzebski (1996), Richard Swinburne (2001), and Jonathan Kvanvig (2003). To see the problem, consider the following analogy. Suppose one wants to visit a nearby bookstore with a good philosophy section while visiting an unfamiliar city, and one searches the Internet to find a store. Two sites are generated, one titled "Bookstores with a good philosophy section" and another titled "Bookstores likely to have a good philosophy section." Presumably, one will be more interested in the first than in the second, but the relevant point to note in our context is something different. Suppose one takes the time to construct the intersection of the two lists, resulting in a list of bookstores that both have and are likely to have a good philosophy section.

The point of the analogy is that it may be true that the first list is analogous to true belief, the second to justified belief, and the third to justified true belief. The swamping problem occurs in the bookstore example because the third list is no more valuable than the first when one's interests are simply to visit a bookstore with a good philosophy section. The swamping problem in epistemology is simply that the value of justification is swamped by the value of truth when justification is conceived solely in terms of objective likelihood of truth, for the same reasons that a list of bookstores that both have and are likely to have a good philosophy section is no more valuable than a list of bookstores that have a good philosophy section.

There are two ways to develop a theory of justification that addresses the swamping problem and thereby provides an account of justification that is helpful in an attempt to explain the value of knowledge. The first is to deny that the means-ends relationship needs to be one of objective likelihood. According to this approach, sometimes the means we adopt are nothing more than wishes or hopes or prayers for achieving the goals we have, but they are means to the goal in question nonetheless. For examples of such, think of the plight of the hopeless suitor, flailing away in the dark trying to find some way of winning the heart of his beloved. He knows he has no clue how to succeed and he knows that everything he tries may not even increase his chances of success. Even if his efforts are not successful, however, they still constitute the means he has adopted to achieve the goal in question.

Just so, justification may be a means to the goal of having true beliefs without being conceived to yield objective likelihood of having such. According to such subjective approaches, there is value in pursuing the truth by whatever means or methods are best by one's own lights, in full knowledge that these means or methods might having nothing more in their favor than hopes and wishes. Moreover, the value added by this property is not obviously swamped by the value of truth in the way that the property proposed by objective likelihood theorists is swamped, just as we value honesty and sincerity even when restricting our considerations to accurate reports. So one way of developing a theory of justification useful in the project of explaining the value of knowledge is to develop a subjective theory of justification.

The other way is to add further elements to the objective approaches so that the swamping problem is eliminated. One way to do so appeals to virtue epistemology, according to which knowledge is the product of the application of one's intellectual virtues (see Greco 2003, Riggs 2002, and Sosa 2003). On a standard account of the intellectual virtues, a virtue is a stable trait of character that makes the beliefs it produces likely to be true. In this way, standard virtue theories adopt objective likelihood accounts of justification. They do not stop, however, with the idea that justification is simply objective likelihood of truth. They add that this objective likelihood of truth must also arise from the display of some laudable intellectual character. The true beliefs that result are not merely likely to be true, they also constitute *accomplishments* of the believer, so that having the true belief is something for which the believer is responsible. As a result, the cognizer deserves credit for having a true belief, and this credit is valuable in a way not explained by

the likelihood that the belief is true. For this reason, virtue approaches to justification have some hope of avoiding the swamping problem of providing an account of justification that is useful in the project of explaining the value of justification in terms greater than the value of its parts.

The fourth condition for knowledge. Were knowledge nothing more than justified true belief, these approaches to justification would give significant hope to the idea that knowledge is more valuable than its parts. Knowledge, however, is more than justified true belief; it is justified true belief where the connection between justification and truth is, in an appropriate way, nonaccidental. Various theories have been proposed regarding the appropriate kind of nonaccidentality that is required for knowledge, with the two most popular being the defeasibility theory and the relevant alternatives theory. There are serious worries that any approach to the fourth condition undermines the idea that knowledge is more valuable than its parts, and we can use these two theories to illustrate the difficulties.

The fundamental problem faced by all theories of the fourth condition is an insensitivity to the problem of the value of knowledge. In the *Meno*, Meno's response to Socrates's counterexample was to question why we prize knowledge more than true opinion and, indeed, whether there is any difference between the two. Meno's response reveals an important constraint on a theory of knowledge. To the extent that the theory focuses on the nature of knowledge at the expense of being able to account for the value of knowledge, it is suspect; and to the extent that a theory focuses on the issue of the value of knowledge at the expense of being able to account for the nature of knowledge, it is suspect as well.

The two major approaches to the fourth condition cited above provide excellent illustrations of how to err in each of these directions. Take first the relevant alternatives theory. On a relevant alternatives approach, the difference between knowledge and justified true belief is determined by whether one would be immune from error in alternatives to the actual situation. In perceptual cases, for example, suppose the surrounding area is littered with fake barns, but one happens to be looking at the only real barn in the area. Then in alternatives to the actual situation, one is not immune from error, for had one been looking at a fake barn, one would still have believed of it that it is a (real) barn.

This theory handles the fake barn case quite well, but it also risks implying global skepticism, if we consider the alternative situation in which Descartes's evil demon is operative. In order to avoid this skeptical consequence, this approach introduces the qualifier "relevant," and holds that the evil demon scenario is not a relevant alternative to the actual situation. The pressing issue for this approach is to specify what makes a situation relevant, and here relevant alternatives theorists have had little to say. The most simplistic version of the view would simply rely on our intuitive understanding of the concept of relevance, claiming that no more precise theoretical specification is needed.

Such a theory is well suited to addressing the issue of the value of knowledge. Immunity from error is itself a good thing, and it would be hard to argue that one should prefer such immunity in *irrelevant* alternatives to immunity in *relevant* alternatives. Whether this value could withstand the scrutiny needed to provide a full and complete answer to the question of the value of knowledge would remain to be seen, but the theory provides some hope of such. It provides such hope by identifying a property with obvious evaluative dimensions, and in this way follows the strategy of addressing questions regarding the value of knowledge by identifying evaluative features of knowledge not present in mere true belief or even in justified true belief.

What this theory gains through the use of the concept of relevance in addressing the problem of the value of knowledge, however, it sacrifices in addressing the problem of the nature of knowledge. For without some clarification of the concept of relevance, this approach is a nonstarter for addressing the problem of the nature of knowledge. It is important to recognize explicitly the significance of the intuitive concept of relevance, however. For the evaluative nature of this concept gives one precisely what one would wish for when focusing on the question of the value of knowledge. It is unfortunate that the simplistic version of this approach has no similar hope of adequately addressing questions regarding the nature of knowledge.

The defeasibility approach begins from a starting point that appears attractive in the search for a solution to the problem of the value of knowledge as well. The starting point for such theories is that what distinguishes knowledge from mere justified true belief is the absence of defeaters—information that, if acquired, would undermine the justification in question. In the fake barn case above, the further (unknown) information is that the landscape is littered with fake barns that cannot be distinguished from real ones.

This starting point is attractive from the point of view of the problem of the value of knowledge, for it cites

a valuable property for a belief to have. It is valuable to have a belief whose justification cannot be undermined by learning any new information. The problem is that this starting point is inadequate, and to the credit of defeasibility theorists, they move beyond the simple relevant alternatives theory above by providing detailed and sophisticated accounts of precisely what unknown information undermines knowledge.

These accounts thus provide the detail needed in a serious effort to uncover the nature of knowledge, but the details of these accounts are completely insensitive to questions regarding the value of knowledge. The standard approach to developing the needed detail is to assemble a stable of examples, some of which involve knowledge and some of which do not, and attempt to find some distinguishing feature of the defeaters in cases of knowledge to use in refining the initial insight of the defeasibility theory. The result of this strategy is an approach that has little hope of providing a defeasibility condition that tracks any difference in value, and thus provides little hope in the attempt to explain the value of knowledge over that of its parts.

For example, consider one of the ways in which the simple defeasibility account is inadequate. Testimony by reliable persons often provides a defeater for what we would otherwise be justified in believing. Suppose we have visual evidence that a friend, Tom, left the library at 11 p.m. Our justification can be defeated if Tom's mother says that Tom has an identical twin that we did not know about who was in the library while Tom was at home fixing his mother's dishwasher. Whether it undermines our knowledge, however, depends on other factors such as who she reports this information to and what they know about her. It will not undermine our knowledge, for instance, if she fabricates the testimony to the police who are checking out a crime that occurred in the library, and the police have a large file of made-up stories from this woman in defense of Tom, who has a long criminal record, especially if the file contains precisely this concocted story, which the police have already checked in prior cases, discovering that Tom is an only child.

The simple defeasibility approach was attractive in the search for an explanation of the value of knowledge because it is valuable to have opinions that no further learning can undermine. Once we see cases such as the above, however, the defeasibility approach loses this attractive feature, for one can have knowledge even when further learning would rationally undermine one's opinion. In such cases, it is true that even more learning would restore one's original opinion, but there is little comfort to be found there, for the same will be true of any true belief, since if one knows all there is to know about a given claim, one will believe it if and only if it is true.

Defeasibility theories have had considerable difficulty in finding a condition that properly distinguishes when defeaters undermine knowledge and when they do not. The problem created by such approaches for the problem of the value of knowledge, however, is the tortured and ad hoc way in which various complex conditions are proposed to do the job. In light of the labyrinthine complexity that such accounts of knowledge display, no optimism is justified that such conditions will track any value difference between satisfying those complex conditions and not satisfying them. It appears that the most warranted conclusion to draw is that the task of distinguishing cases of knowledge from cases of nonknowledge has been revealed to be so difficult that epistemologists make progress on the question of the nature of knowledge only by proposing conditions that undermine any explanation of the value of knowledge by appeal to those conditions.

CONCLUSION

So the idea that truth is valuable on intrinsic grounds from a purely cognitive point of view may be defensible, but the same kind of defense of the value of knowledge is implausible. Instead, the more plausible approach tries to show that knowledge is valuable in virtue of its parts, but attempts along these lines founder on the admission that knowledge can be fallible. Such a result is compatible with truth and knowledge being valuable both from a purely cognitive point of view and from an all-things-considered point of view, but then knowledge will not have the type of value it is ordinarily assumed to have.

See also Truth.

Bibliography

Craig, Edward. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press, 1990.

Fogelin, Robert. *Pyrrhonian Reflections on Knowledge and Justification*. Oxford: Oxford University Press, 1994.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121–123.

Greco, John. "Knowledge as Credit for True Belief." In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski. Oxford: Oxford University Press, 2003.

James, William. "The Will to Believe." In his *The Will to Believe and Other Essays in Popular Philosophy*. New York: Dover Publications, 1956.

Kvanvig, Jonathan. *The Value of Knowledge and the Pursuit of Understanding*. New York: Cambridge University Press, 2003.

Riggs, Wayne D. "Beyond Truth and Falsehood: The Real Value of Knowing that P." *Philosophical Studies*107 (2002): 87–108.

Rorty, Richard. *Contingency, Irony, and Solidarity*. Cambridge, U.K.: Cambridge University Press, 1989.

Rorty, Richard. *Truth and Progress*. Vol. 3 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1998.

Sosa, Ernest. "The Place of Truth in Epistemology." In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski. New York: Oxford University Press, 2003.

Swinburne, Richard. *Epistemic Justification*. Oxford: Oxford University Press, 2001.

Williams, Bernard. *Truth and Truthfulness*. Princeton, NJ: Princeton University Press, 2002.

Zagzebski, Linda Trinkaus. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundation of Knowledge*. New York: Cambridge University Press, 1996.

Jonathan L. Kvanvig (2005)

KNOWLEDGE AND VAGUENESS

When anthropologists painstakingly identified the taxon of the skeleton that later became known as "Lucy's child,"

> There was no eureka. There was no grand turning point. The evidence kept dribbling in, and through hard labor and some dogged thinking we *did* solve the puzzle, not through revelation but through a sort of absorption, just below the level of explicit consciousness. It was as if the truth had slowly seeped through our pores, until we had come know it without knowing that we did. So when the final, indisputable confirmation came, we hardly noticed the event. What had once been a mystery had become—in hindsight, mind you—obvious from the start (Johanson and James Shreeve 1989, p. 203).

Instead of there being a clear point at which the anthropologists knew that the specimen was *Homo habilis*, there was stratification: The researchers began from obvious ignorance, inched up to being borderline knowers, and eventually emerged as clear knowers.

The vagueness of knowledge has substantial implications. When skeptics took over Plato's Academy, they tried to prove that there can be no knowledge. Such a proof would ensure that everything is a clear negative case of "knowledge." *Knowledge* would be a perfectly precise term; a skeptic should think twice before complaining about the vagueness of knowledge! Typically, borderline cases are flanked by clear cases (Figure 1), so

FIGURE 1

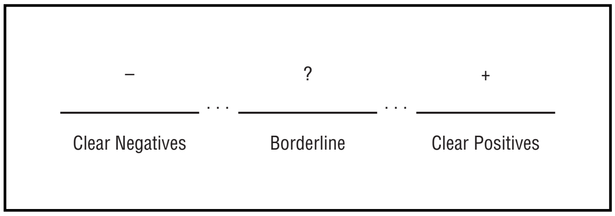

the vagueness of "know" positively invites the inference that there is at least some knowledge.

The vagueness of knowledge also affects principles of epistemic logic such as the "KK thesis": If you know, then you know you know. If the KK thesis were true, the anthropologists would have known that they knew from the moment they knew the taxon of Lucy's child.

Given a naturalistic perspective on knowers, the vagueness of "know" should be expected. Human perceptual capacities and memory trail off in the patterns made famous by evolutionary iconography (Figure 2).

1. THE SORITES PARADOX

Only a vague term (e.g., *human*) can serve as the inductive predicate of a sorites argument:

> Base Step: There are now humans.
>
> Induction Step: If there were humans *n* years ago, then there were humans *n* - 1 years ago.
>
> Conclusion: There were humans five billion years ago.

Because the earth is only 4.6 billion years old, the conclusion is false. The base step is clearly true and the argument is classically valid. Therefore, people naturally suspect the induction step. However, they are unable to specify a value for *n* at which the generalization is false.

If vagueness is merely a kind of ignorance, there is no need to find a counterexample to the induction step. One can know a generalization is false even if one cannot pinpoint where it breaks down. Consider an anthropologist who doubts that all of the skeletal fragments in a bag belong to a single individual but cannot identify any pair of fragments as belonging to distinct specimens. When the anthropologist weighs the bag and learns there are more than enough fragments to constitute one skeleton, that is all that is needed to refute the generalization that all of the fragments come from a single individual.

In common usage, a borderline case is often simply one that cannot be settled at a given stage of inquiry. When an archeologist sorts stones, a few are obviously tools and most others are clearly just rocks. There will be another

FIGURE 2

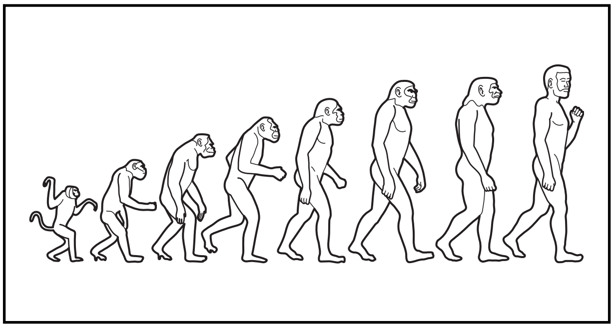

group of stones whose status cannot be determined by unaided observation. These borderline cases are put under a field microscope. The three-way sorting begins afresh. Borderline cases that survive this second stage of inquiry may eventually wind up under an electron microscope.

Philosophers focus on the minority of borderline cases in which there is no prospect of resolution. How many years did the Middle Ages last? Is Israel a new state or an ancient state? Philosophers are at sea with these questions, and because people are unsure what would count as correctly answering these questions, their ignorance cannot be relativized to a set of resources.

Epistemicists insist there remains a crucial resemblance between these absolute borderline cases and relative borderline cases; they take all vagueness to be a form of ignorance. Epistemicists solve the sorites paradox by claiming that there is a hidden counterexample to the induction step. After all, they know the base step is true and the conclusion is false; classical logic then licenses the deduction that the induction step is false.

In classical logic, denying the induction step of the above sorites argument is equivalent to asserting there is a number *n* such that *n* years ago there was at least one human being but the year before that there were no human beings. So belief that there is a counterexample to the induction step is equivalent to the belief that there was a first human!

Incredulous anthropologists counter that nature does not draw a sharp line between humans and nonhumans. Speakers have not made up for the absence of sharp boundary by supplying an artificial one. Consequently, anyone who searches for the exact year humans appear on the evolutionary timeline is conceptually confused.

2. INFINITE REGRESSES

David Sanford (1975a) points out that if finite sequences do not need beginnings or endings, there are neglected solutions to infinite regress problems. Consider the infinite regress of justification: A belief can only be justified by another justified belief. Justification cannot be achieved by reasoning in a circle. Nor can chains of justification be infinitely long. The skeptic concludes that no beliefs are justified. The foundationalist responds by conferring axiomatic status on some beliefs; axioms justify other beliefs without needing justification from other beliefs. The vagueness of "justified" suggests another solution to this infinite regress: Admit that the chain of justification is finite but deny it must terminate in an axiomatic belief.

Compare justification to motherhood. Each woman must have a mother. Her family tree cannot go back infinitely and cannot circle back on itself. Is one to conclude that some woman lacks a mother? Sanford instead appeals to the vagueness of "mother." As one moves down her ancestral line, what counts as a mother eventually becomes less and less clear. After passing through a stretch of borderline cases, one arrives at ancestors who clearly lack a gender and therefore are clear nonmothers. Sanford says that an insistence that finite sequences have terminal points is an incarnation of the sorites paradox.

3. THE LOGICAL PREDICAMENT

Because the sorites argument is classically valid, David Sanford must espouse a deviant logic. Supplemental logics (modal logic, deontic logic, etc.) merely add theorems to the standard stock; they cannot subtract the sorites from the list of valid arguments. So Sanford must target classical logic, weakening it just enough to stop its validation of the sorites—without causing too much collateral damage. In standard fuzzy logic, almost all classical theorems are rejected—except for the special case in which the truth-values equal full truth or full falsehood (Machina 1976). Sanford (1975b) accepts degrees of truth but prefers to keep all classical theorems by rejecting the truth-functionality of the logical connectives. Other deviant logicians reject some classical inference rules. For instance, intuitionists closely associate proof with truth and so try to derail the sorites paradox by rejecting the validity of double-negation (Putnam 1983). Supervaluationists either reject inference rules such as contraposition and reductio ad absurdum or reject core semantic principles such as Tarski's convention T (McGee and McLaughlin 1995).

These changes occur at the center of the human web of belief and so reverberate widely. Because knowledge implies truth, new questions are raised by the fuzzy logician's talk of degrees of truth. For instance, can one know

a proposition that has a degree of truth less than one? The fuzzy logician wants to explain human ignorance of typical borderline cases and so is committed to saying that people are ignorant of propositions that are as close to being false as to being true. But what about propositions that are nearly true? Fuzzy logicians say that many propositions that appear to be clear truths merely have a high degree of truth. So if knowledge implies full truth, people know less than they seem to know.

4. THE CREDIBILITY GAP

Knowledge does seem to imply full truth because knowledge implies belief and one can only believe what one considers to be fully true. "It is not fully true that the Black Skull is an australopithecine but I believe it is an australopithecine" is as hard on the ear as G. E. Moore's paradoxical sentence "It is raining but I do not believe it" (Moore 1942, 543).

This credibility gap hinders efforts to moderate epistemicism. Intuitively, people's wishy-washy attitude toward borderline cases seems like a reaction to the vagueness of these cases. But a subjectivist may reverse the relationship and say that the wishy-washy attitudes are what make propositions vague. If indeterminacy is a *projection* of human ambivalence, then people may hope to avoid the metaphysical burden of epistemicism. The epistemicist would be right in basing vagueness in the subject's limitations but wrong in postulating sharp thresholds.

Crispin Wright (2001) says that *x* is a borderline case of F-ness if two parties can disagree about whether *x* is F without either party being guilty of a cognitive shortcoming. Each party knows all the relevant facts, each is a competent speaker, and each has reasoned well. Wright compares this faultless stalemate with the cultural variation that makes relativism popular among ethnographers.

Critics of Wright object that anyone who takes a position on a borderline statement is guilty of a cognitive shortcoming; they ought to be agnostic. If one thinks that same-sex civil unions are borderline cases of marriages, then one cannot believe that they are marriages.

Stephen Schiffer (1998) has suggested that people have a special attitude toward cases that they take to be borderline. "Vague partial belief" differs from the belief humans extend to precise propositions. It also differs from the degrees of belief that people associate with probability theory. The probability calculus instructs people to assign a higher probability to a disjunction than either of its contingent disjuncts. But when the disjuncts are borderline cases, Schiffer only assigns the disjunction as much vague partial belief as he assigns the strongest disjunct.

This result (which echoes the fuzzy logician's rule for calculating disjunctions) grates against the observation that hedging a claim can make it more assertible. One can know that Blaise Pascal died at thirty-nine but not be sure whether this counts as dying as a young man. However, one can confidently say that *either* Pascal died as a young man or as a man in middle age.

Supervaluationists have a simple explanation of why people do not believe borderline statements: they lack truth-values. Belief aims at truth, thus people cannot believe a statement that they believe to be borderline. However, this explanation overgeneralizes, for it does seem possible to have weak propositional attitudes (guessing, doubting, and suspecting) toward statements that one acknowledges to be borderline.

Supervaluationists also have trouble explaining why one can make a statement more credible by adding an *epistemic* hedge. If one believes linguistic indecision prevents "ten is a small number" from having a truth-value, then one cannot believe it may be true. Yet if ten clearly is a borderline case of a small number, then it is appropriate to shrug one's shoulders and conclude "ten might be a small number and ten might not be a small number." Indeed, prefixing any statement that is clearly borderline with "maybe" seems to make it clearly true.

Supervaluationists use truth-value gaps and the principle that knowledge implies truth to explain why humans are absolutely ignorant of borderline statements. God cannot know when a fetus becomes a human being because there is nothing to know.

Supervaluationists pride themselves on the modesty of their revision of classical logic. The workhorse of their adjustment is the notion of super-truth: A statement is super-true if and only if it comes out true under all admissible precisifications of the statement. For instance, "Either the specimen is a *Homo erectus* or an archaic *Homo Sapien*" is super-true because it comes out true regardless of how one precisifies *Homo erectus* and *archaic Homo sapien*.

Any statement that has the form of a classical tautology will be super-true even if it contains vague terms. So supervaluationists claim to preserve all the theorems of classical logic.

But can one believe a statement by virtue of its super-truth? Truth under all *disambiguations* is not enough. Suppose a person says "bachelors are mammals"

and it is not clear whether that person is referring to unmarried men or to college graduates or to just any young male mammal. One knows the statement expresses a truth but does not know which truth it expresses. Ambiguous statements are not objects of knowledge.

But vague statements are objects of knowledge. People know "the number of men is either an even number or an odd number" even though the vagueness of "man" makes it impossible to count the number of men. Supervaluationists have trouble accepting asymmetries between vagueness and ambiguity. They characterize vagueness in semantic terms rather than epistemic terms, so supervaluationism looks more like a logic of ambiguity (Lewis 1982).

5. HIGHER ORDER VAGUENESS

In *Purity and Danger* Mary Douglas conjectures that the bearers of taboos are borderline cases (moles, eels, twilight, and so on). She interprets rituals of purification as attempts to reclassify doubtful cases (as when hermaphrodites are declared men through a rite of passage). Assessment of Douglas's hypothesis is hindered by the vagueness of "borderline case."

Borderline cases of "borderline case" are normal with vague terms. In addition to there being borderline cases of "human," there are borderline cases of "borderline human." So in addition to first order vagueness there is second order vagueness, third order vagueness, and so on, apparently ad infinitum.

Higher order vagueness is a problem for deviant logicians because they employ classical logic and set theory in the metalanguages they use to describe vague terms. This classical medium forces them to represent the transition from clear to borderline cases as a sharp threshold. For instance, supervaluationist semantics implies that there is a first point at which "*x* is a human" is true. So instead of having the epistemicist's sharp threshold between truth and falsehood, the supervaluationist has a sharp threshold between truth and absence of truth. Similarly, the fuzzy logician has sharp thresholds between each degree of truth, and can only approximate vagueness by using a large quantity of discrete micro-transitions. The fuzzy logician's representation of vagueness is like a dot matrix printer's representation of gray—a black and white affair when examined close up.

What originally bothered philosophers were *sharp thresholds*, not sharp thresholds *between truth and falsehood*. Thus epistemicists advertise themselves as just self-consciously biting a bullet that others gnaw absentmindedly.

6. EXPLAINING THE IGNORANCE

Recent epistemicists are careful to endorse the principle that inquiry into borderline cases is futile. That is why they stress that borderline statements are unknowable. But if these statements have truth-values, why can't they be known? One response is to challenge the presumption in favor of knowability—to portray ignorance as a natural state in need of no explanation.

However, Timothy Williamson (1994) directly answers the question of why borderline statements cannot be known. He traces the unknowability of borderline statements to the knower's need for a margin for error. When at a stadium, one can know there are about ten thousand people. But one cannot know there are exactly ten thousand, for a person cannot reliably discriminate between there being ten thousand and there being ten thousand and one. Given that "human" has the sort of precise threshold epistemicists allege, anyone who happened to correctly believe that humans originated *n* years ago, would have to be right by luck. For all this person knows, the origin could have been a year earlier.

Williamson believes that thresholds for vague predicates are determined by the psychology, social conditions, and environment of the speech community. These conditions are too complicated to allow humans to ascertain the threshold for vague terms.

The margin for error principle yields different limits for different kinds of knowers. For much of the history of *Homo sapiens* there were other hominids who had different cognitive capacities. Williamson's theory does not preclude these hominids from knowing the threshold of some vague terms. Some of what is chaotic to humans may be predictable to these homonids. Williamson is committed to the relativity of all borderline cases. Supervaluationists claim an advantage over Williamson insofar as they neatly model absolute borderline cases.

Roy Sorensen (2001) has speculated that an epistemicist can match the neatness of the supervaluationists by using truth-maker gaps instead of truth-value gaps. A truth-maker is what makes a proposition true. For instance, "Humans and chimpanzees had a common ancestor seven million years ago" is made true by a Miocene primate who had as descendants both Noam Chomsky and Nim Chimpsky. One learns the truth-value of propositions only by becoming appropriately related to their truth-makers. Propositions that lack truth-makers have truth-values that are not anchored to any piece of reality. This objective indeterminacy makes the propositions absolutely unknowable.

7. VAGUENESS AND EPISTEMIC LOGIC

If the relationship between knowledge and borderline cases is orderly, epistemicists can offer a logic of vagueness as a branch of epistemic logic. For instance, Timothy Williamson elaborates his "logic of clarity" in a way that makes it isomorphic to supervaluationism. The basic idea is that a statement is definitely true if it comes out true "under all sharp interpretations of the language indiscriminable from the right one" (Williamson 1999, p. 128). This mirrors the supervaluationist's principle that a statement is definitely true if it comes out true under all admissible completions of the language.

Epistemicists are divided on how closely vagueness is bound up with borderline cases. Everybody agrees that a vague term need not have actual borderline cases. Possible borderline cases are sufficient. But what about borderline cases that are merely epistemically possible? Perhaps the mere threat of an objective borderline case can be enough to make a predicate vague (Sorensen 2001). After all, if the threat cannot be exposed as false, then there will be irremediable linguistic ignorance without borderline cases. One would be able to embed the predicate in a sorites argument and bedevil people with doubts about termination points.

See also Agnosticism; Classical Foundationalism; Contextualism; Doubt; Laws of Thought; Relevant Alternatives.

Bibliography

Johanson, Donald, and James Shreeve. *Lucy's Child*. New York: Avon Books, 1989.

Lewis, David. "Logic for Equivocators." *Noûs* 16 (1982): 431–441.

Machina, Kenton. "Truth, Belief and Vagueness." *Journal of Philosophical Logic* 5 (1976): 47–78. Reprinted in *Vagueness: A Reader*, edited by Rosanna Keefe and Peter Smith, 174–203. Cambridge, MA: MIT Press.

McGee, Vann, and Brian McLaughlin. "Distinctions Without a Difference." *Southern Journal of Philosophy* 33 (Supplement) (1995): 203–251.

Moore, G. E. "A Reply to My Critics" in *The Philosophy of G. E. Moore*, edited by P. A. Schlipp. Evantston, IL: Northwestern University, 1942.

Putnam, Hilary. "Vagueness and Alternative Logic." In his *Realism and Reason*. Cambridge, U.K.: Cambridge University Press, 1983.

Sanford, David. "Borderline Logic." *American Philosophical Quarterly* 12/1 (1975b): 29–39.

Sanford, David. "Infinity and Vagueness." *Philosophical Review* 84 (1975b): 520–535.

Schiffer, Stephen. "Two Issues of Vagueness." *Monist* 81 (2) (1998): 193–214.

Sorensen, Roy. *Vagueness and Contradiction*. Oxford: Clarendon Press, 2001.

Sorensen, Roy. "The Vagueness of Knowledge." *Canadian Journal of Philosophy* 17 (4) (1987): 767–804.

Williamson, Timothy. "On the Structure of Higher-Order Vagueness." *Mind* 108 (1999): 127–143.

Williamson, Timothy. *Vagueness*. London: Routledge, 1994.

Wright, Crispin. "On Being in a Quandary." *Mind* 110 (437) (2001): 45–98.

Roy Sorensen (2005)

KNOWLEDGE ARGUMENT

The definitive statement of the knowledge argument was formulated by Frank Jackson in a paper titled "Epiphenomenal Qualia" that appeared in the *Philosophical Quarterly* in 1982. Arguments in the same spirit had appeared earlier (Broad 1925, Robinson 1982), but Jackson's argument is most often compared with Thomas Nagel's argument in "What Is It Like to Be a Bat?" (1974). Jackson, however, takes pains to distinguish his argument from Nagel's. This entry will follow standard practice in focusing on Jackson's argument, though it also describes the main points of alleged similarity and dissimilarity between these two arguments.

The knowledge argument targets physicalism about the mind, which claims that, as Jackson puts it in a follow-up article, "the actual world … is entirely physical" (1986, p. 281). The argument provided one of the chief sources of doubt about physicalism in the late twentieth century, and continues to shape discussion of the mind-body problem into the twenty-first. It is unclear whether the argument converted many to dualism; still, most readers found the argument's core thought experiment highly compelling. Physicalists thus faced the challenge of identifying an error in the argument. The potency of the knowledge argument is clear because while all materialists reject its conclusion, there is little agreement among them as to how, precisely, its reasoning is flawed.

THE ARGUMENT

Jackson's original argument is disarmingly brief. He invites the reader to imagine the following scenario: Mary, a brilliant neuroscientist, has spent her entire life in a room in which the only visible colors are black and white. Partly through the use of a black-and-white television monitor, Mary comes to know all of the physical facts about color vision. These facts include the nature of causal interactions between the surface reflectance properties of objects, wavelengths of light, and retinal stimulation. Jackson then asks: "What will happen when Mary is released from her black and white room or is given a

color television monitor? Will she *learn* anything or not?" He answers: "It seems just obvious that she will learn something about the world and our visual experience of it" (Jackson 1982, p. 130). He thinks that when Mary finally leaves the room and, for the first time, gazes upon an object that is red (and that she knows to be red), she learns *what it's like* to see red. Jackson concludes that, because physicalism requires that all facts are physical facts, physicalism is false.

Jackson's conclusion is a dualism of properties, rather than of substances; and this is all that his argument warrants. For a difference in properties—between the property *instantiating neurophysiological state N*, and the property *instantiating qualitative state Q*, say—suffices for a difference in corresponding facts.

A formalization of the argument will be useful.

(1) While in the black-and-white room, Mary knows all of the physical facts about color experience.

(2) Mary learns something about color experience upon her release.

(3) If Mary learns something about color experience upon her release, she does not know all of the facts about color experience while in the room.

(4) Mary does not know all of the facts about color experience while in the room (from 2 and 3).

(5) There are facts about color experience that are not physical facts (from 1 and 4).

(6) If physicalism is true, then all facts are physical facts.

Therefore,

(7) Physicalism is false (from 5 and 6).

As mentioned above, Jackson distinguishes this argument from Nagel's 1974 argument. Nagel had argued that no amount of physical information about bats—including knowledge of their neurophysiological, behavioral, and evolutionary features—could allow us to grasp the experiential aspect of using echolocation; that is, to know what it's like to be a bat. According to Jackson, these arguments differ in two ways. First, he claims that his argument concerns knowledge of a general property of experience, what it's like to see red, whereas Nagel's argument concerns knowledge of a property specific to an individual; that is, what it's like to be a (particular) bat. But to some, this difference has seemed at most a quirk of exposition: for Nagel's argument does draw into question whether we can know a general property of experience, namely, what it's like to use echolocation. However, others—including Jackson himself—have claimed that whereas Jackson's argument specifically targets the contrast between the phenomenal and the physical, Nagel's argument instead targets the contrast between the subjective and the objective.

The second point of contrast that Jackson draws is this: Nagel's argument simply shows that humans cannot imagine what it's like to use echolocation, and this limit to our imaginative powers is irrelevant to the issue of physicalism. Whether Nagel's argument rests on this issue about imaginability, or whether it would remain intact when using an experience that is within the normal course of human experience (as Jackson's does), is largely a question of interpretation. But the point about imaginability brings out an important and sometimes overlooked feature of Jackson's argument: that nothing in the argument excludes the possibility that Mary, perhaps through an exercise of imagination or as the result of taking a hallucinogen, undergoes an experience while in the room that is, in fact, a *seeing red* experience. Jackson's point remains so long as Mary is unable to determine that the experience is a *seeing red* experience as opposed to, say, a *seeing green* experience. This brings out the epistemic character of the argument. Jackson's argument requires only that Mary cannot *deduce* that a certain experience is the sort of experience her subjects undergo when seeing a ripe tomato (say). Upon leaving the room, Mary has the opportunity to *correlate* these, by gazing at a tomato herself. (She could, of course, correlate them while inside the room, by scanning her own brain while she is undergoing the *seeing red* experience. In the context of the argument, having the opportunity to make this correlation is tantamount to leaving the room.)

OBJECTIONS TO THE ARGUMENT

This entry now turns to the four most influential types of objection to the argument. The first is simply to deny the conjunction of premises (1) and (2). On this view, Mary does not know all of the physical facts unless she knows what it's like to see red. Daniel Dennett (1991) takes this approach, arguing that we cannot truly conceive knowing all of the relevant physical facts. This limitation explains why it seems that Mary learns something upon her release; but, Dennett maintains, if (1) is true, then (2) is false. In response, defenders of the knowledge argument have pointed out that the argument requires only that we understand the basic kind of knowledge that Mary has while in the room, not that we can mentally rehearse each bit of information she possesses. Because we do have a grasp of the sort of physical facts she knows, our powers

of conceiving are strong enough to evaluate the possibility that (1) and (2) are true simultaneously.

The next two objections deny premise (3). Churchland (1985) describes what Mary gains upon her release as a kind of *knowledge by acquaintance* of what it's like to see red; that is, as nonpropositional knowledge of this fact. Using this analysis of the case, he argues that a parallel argument would condemn dualism as well, because Mary would lack such knowledge by acquaintance even if she had exhaustive propositional knowledge about the nonphysical. Jackson (1986) responds that these are not on a par, for one could know all of the physical facts about seeing red without knowing what it's like, but one could not know all of the facts (physical and nonphysical) about seeing red without knowing what it's like. This may seem question-begging, but it has seemed highly intuitive to many philosophers, and hence this second avenue of objection has attracted relatively few proponents. (But see Earl Conee 1994 for a more developed version of the acquaintance analysis.)

Another objection that denies premise (3) claims that what Mary gains upon leaving the room is an ability, rather than knowledge of a fact. This objection originated in Laurence Nemirow's review of Nagel's argument (1980), and is defended by David Lewis (1988). On this *ability* approach to defusing the argument, when she finally sees something red, Mary learns how to remember, recognize, and/or imagine a *seeing red* experience. The fact that experience is required for such abilities carries no antiphysicalist consequences; after all, exhaustive propositional knowledge does not generally guarantee that one possesses the relevant ability. If it did, professional baseball teams would be staffed by physicists, who can master all of the relevant facts about how to hit a curve ball.

While the ability approach remains influential, it does face difficult challenges. One challenge is to specify an ability that is gained when, and only when, Mary learns what it's like to see red. At the moment of her grasping this, she is not yet able to remember what it's like, for the moment has not passed; and if she has a poor imagination, experience may not enable her to imagine what it's like. (For responses along these lines, see Conee 1994 and Torin Alter 1998.) Arguably, the best candidate for what Mary gains is the ability to recognize *seeing red* experiences. But the ability analysis may be mistaken even if this recognitional ability is perfectly correlated with knowing what it's like to see red. For, as Brie Gertler (1999) argues, it seems plausible that Mary is able to recognize a *seeing red* experience *because* she knows what it's like, where *because* is used in an explanatory sense. If knowing what it's like explains the recognitional ability, then it does not reduce to that ability.

The fourth and most widely accepted type of objection to the knowledge argument rejects premise (6). It claims that our ways of representing reality may be more fine-grained than the reality we represent, and what Mary gains is simply a new way to represent a portion of reality that was already known to her. (There are two competing ways to use *fact* in this context. One is to read *fact* as inheriting the fineness of grain that our representations possess; it is this reading that has been used in saying that this objection targets premise (6). The second reading uses *fact* as less fine-grained than our representations. On that reading, the current objection would instead reject premise (3), claiming that Mary didn't learn any new facts but only encountered old facts under a new guise or mode of presentation. The difference here is purely verbal, and this entry will continue to use *fact* in the former sense.)

This sort of objection was present in earlier papers (including Terence Horgan 1984 and Michael Tye 1986), but is usually associated with Brian Loar, who provided a nuanced version of it in 1990. Loar argues that a single property may be the referent of distinct concepts. In particular, a property that Mary knew as *instantiating neurophysiological state N* may be identical to the property *instantiating qualitative state Q*, even if knowledge that a state falls under the former concept does not generate knowledge that it falls under the latter. Thus, Mary's ignorance can be attributed to a distinction in concepts that does not imply any distinction in properties.

More generally, this line of response to the knowledge argument construes the change in Mary as purely epistemic, and denies that her epistemic advance, upon leaving the room, reflects any grasp of a hitherto unknown ontological feature of the world. As such, it represents a more general, highly influential position about the mind-body problem: The apparent disparity between physical and phenomenal features of the world (called *the explanatory gap* after Joseph Levine "1983") is purely epistemic, and not ontological.

This position belongs to a more general outlook known as *a posteriori physicalism*. According to *a posteriori* physicalists, antiphysicalist arguments that are based on thought experiments show, at most, that physicalism is not an a priori truth; but as Saul Kripke (1980) demonstrated, some identities are *a posteriori* (yet necessarily true). Strikingly, Kripke himself rejects *a posteriori* physicalism and claims that the distinctive way in which phenomenal concepts operate rules out the possibility of *a*

posteriori identities between phenomenal and physical (or functional) properties. In a paper co-written with David Chalmers, Jackson also objects to *a posteriori* physicalism. According to Chalmers and Jackson (2001), the approach used by *a posteriori* physicalists presumes that there is a deep schism between concepts and ontology, a schism that would undercut the justification for uncontroversial identity statements.

Despite his continuing opposition to *a posteriori* physicalism, Jackson now rejects the knowledge argument (Jackson 2003). He contends that phenomenal knowledge is deducible, in principle, from physical knowledge, even if we may be unable to perform the deduction. Jackson's turnabout is based on his acceptance of representationalism, which claims that the phenomenal character of a state is exhausted by its representational content. For instance, suppose that one of Mary's subjects, Joe, gazes at a ripe tomato. Representationalists maintain that the visual phenomenal quality of Joe's experience is fully captured by the fact that his state represents *there is something round and red before me*. (Specific representational contents will be much more detailed, of course.) Because Mary can, in principle, know the representational contents of Joe's states before her release, she can in principle know all that there is to know about what it is like to see red.

OVERALL ASSESSMENT

The knowledge argument is an argument against physicalism. Yet its importance stems as much from the richness and variety of the responses inspired by its provocative reasoning as from its conclusion. Discussion of the argument has profoundly affected debate on a range of issues, including: differences between propositional knowledge and ability, the relation between identity and deducibility, and the special features of phenomenal knowledge. While the majority of philosophers ultimately reject the argument, a vocal minority accepts it as sound.

See also Functionalism; Mind-Body Problem; Physicalism; Qualia.

Bibliography

Alter, Torin. "A Limited Defense of the Knowledge Argument." *Philosophical Studies* 90 (1998): 35–56.

Broad, C. D. *Mind and its Place in Nature*. London: Routledge and Kegan Paul, 1925.

Chalmers, David, and Frank Jackson. "Conceptual Analysis and Reductive Explanation." *Philosophical Review* 110 (2001): 315–360.

Churchland, Paul. "Reduction, Qualia, and the Direct Introspection of Brain States." *Journal of Philosophy* 82 (1985): 8–28.

Conee, Earl. "Phenomenal Knowledge." *Australasian Journal of Philosophy* 72 (1994): 136–150.

Dennett, Daniel. *Consciousness Explained*. Boston: Little, Brown, 1991.

Gertler, Brie. "A Defense of the Knowledge Argument." *Philosophical Studies* 93 (1999): 317–336.

Horgan, Terence. "Jackson on Physical Information and Qualia." *Philosophical Quarterly* 34 (1984): 147–52.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Jackson, Frank. "Mind and Illusion." In *Minds and Persons*, edited by Anthony O'Hear. Cambridge, U.K.: Cambridge University Press, 2003.

Jackson, Frank. "What Mary Didn't Know." *Journal of Philosophy* 83 (1986): 291–295.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Levine, Joseph. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (1983): 354–361.

Lewis, David. "What Experience Teaches." *Proceedings of the Russellian Society* 13 (1988): 29–57. Reprinted in Ludlow et al., *There's Something About Mary*.

Loar, Brian. "Phenomenal States." In *Philosophical Perspectives*, vol. 4: *Action Theory and Philosophy of Mind*, edited by James Tomberlin. Atascadero, CA: Ridgeview Press, 1990.

Ludlow, Peter, Yujin Nagasawa, and Daniel Stoljar, eds. *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. Cambridge, MA: MIT Press, 2004.

Nagel, Thomas. "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435–50. Reprinted in his *Mortal Questions* (Cambridge, U.K.: Cambridge University Press, 1979).

Nemirow, Laurence. "Review of Thomas Nagel's *Mortal Questions*" *Philosophical Review* 89 (1980): 473–477.

Robinson, Howard. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge, U.K.: Cambridge University Press, 1982.

Tye, M. "The Subjective Qualities of Experience." *Mind* 95 (1986): 1–17.

Brie Gertler (2005)

KNOWLEDGE IN INDIAN PHILOSOPHY

Almost all the philosophical texts in classical India were written in Sanskrit. How does one say *knowledge* in Sanskrit? And what do the Sanskrit terms that may be translated by the English word *knowledge* mean exactly? There are no simple answers to these questions.

In Western philosophy truth and falsity are usually ascribed to statements, propositions, or beliefs. In the Indian tradition truth and falsity are ascribed to a cogni-

tion or an awareness (the most common term is *jñāna*, but there are a relatively large number of synonyms, or quasi synonyms, such as *vijñāna*, *buddhi*, *dhī* and *citta*). The word *jñāna* is derived from the root *jñā*, which is etymologically related to the English word *know*. Nevertheless, the rendering of *jñāna* as knowledge is generally avoided because *jñāna* can be true or false, whereas false knowledge or wrong knowledge seems like a contradiction in terms (at least in English). Furthermore, *jñāna* is a particular and momentary event, whereas *knowledge* often refers to a general and lasting acquaintance with facts. Furthermore, knowledge is, or may be, an abstract entity that is shared by many persons; *jñāna* is always individual and belongs to a single person. Finally, *knowledge*, unlike *jñāna*, is a collective term and can only be used in the singular. A person has many *jñānas*, but not many *knowledges*.

The different ontologies of the various traditions of Indian philosophy necessitate different notions of *jñāna*. According to some Brahminical schools, *jñāna* is a momentary property of the eternal individual soul (*ātman*). The relationship between *jñāna* and *soul* is the relationship between quality and substance. It is the same relation that occurs between a color and the material substance like a pot in which it inheres. In contrast, the Buddhists reject the idea of substance in general and of a permanent soul or self in particular. According to them an awareness (*jñāna*) is a primitive (nonderivative) element of existence (*dharma*) that depends only on its causes and conditions (e.g., sense, object, and previous mental factors), not on any substrate such as a permanent soul. The Sāṃkhya and Yoga are unique in the Brahminical tradition in claiming that the cognitive and psychological processes occur in the realm of matter and have no direct contact with the conscious soul, which is distinct from them and completely passive (for more details, see Chakravarti 1975, pp. 171–196). Finally, according to the materialists (Cārvāka or Lokāyata), an awareness, or consciousness, arises from the combination of the material elements earth, water, fire, and wind when they evolve into body, sense, and object, just as the power of intoxication arises when certain substances ferment (Namai 1976, Franco 1997, pp. 98–99).

Knowledge in general as referring to an organized body of knowledge, or even a science, is usually called *veda* or *vidyā* (words that are cognate with Latin *videre* and the English *to wit*). When the word *veda* is mentioned without further qualification, it always refers to the four collections of texts known as Ṛgveda, Yajurveda, Sāmaveda, and Atharvaveda. These contain *the* knowledge, the knowledge par excellence. The Vedas are the primary scriptures of Brahmanism and Hinduism. According to Brahminical orthodoxy they are neither of human nor Godly origin, for they are eternal and infallible. The text of the Vedas was revealed (not created) by omniscient Gods such as Brahma, or directly heard by inspired seers (Rishis) of old. Various enumerations and classifications of systematic knowledge, or sciences, have been transmitted; perhaps the most common ones refer to fourteen or eighteen locations of knowledge (*vidyāshtāna*): the four Vedas and the six auxiliary sciences to the Vedic texts (the sciences of articulation or phonology, prosody, grammar, etymology, astronomy/astrology, and ritual/ceremony), religious and social law (*dharmaśāstra*), collections of ancient myths (*purāṇa*), hermeneutics (*mīmāṃsā*), and dialectics (*tarka*); the eighteenfold enumeration adds medicine (*āyurveda*), archery or the science of weapons in general (*dhanurveda*), and *arthaśāstra*, which includes politics and economy.

These lists do not exhaust all the sciences known in ancient India, but they point to an attempt at an exhaustive classification of human cultural practices (Pollock 1985, p. 502). Sheldon Pollock, who examined the notion of *sāstra* in classical India, points out that virtually every human activity had been codified into a science (or a theory, as he renders the word *śāstra*), for instance, cookery, erotics (*kāmaśāstra*), thievery (*cauraśāstra*), agriculture, mathematics, logic, ascetic renunciation, and spiritual liberation. As a rule (there are notable exceptions), the various sciences have not been discovered by their practitioners. Rather, all practice is said to be derived from previously existing knowledge. Science itself is primordial; it is not accumulative, and can only decrease with time.

In Buddhist texts (both in India and Tibet) one encounters a list of five places or locations of knowledge (*vidyāsthānas*) that are to be cultivated by the Bodhisattva on his way to enlightenment. The first of these, the inner science or the own science (*adhyātmavidyā*), is specific to Buddhists, the other four—the science of logical reasons, grammar, medicine, and arts and crafts—are external and considered common to Buddhist and non-Buddhists (Seyfort Ruegg 1995, pp. 9–10). However, the status of the science of reasons, that is, philosophy/dialectics/logic, was ambiguous. Although its position following the inner science clearly implies that it is an external (or non-Buddhist) science, it was sometimes considered to be part of the Buddhist teachings. The science of logical reasons could be assimilated either to *tarka*, dialectics, which have nothing particularly Buddhist about them, or it could be understood as the science of the means of knowledge (*pramāṇa*), as expounded by Dharmakirti (seventh century) and his followers that was closely associated to the

understanding and interpretation of the Buddhist teachings (Seyfort Ruegg 1995, p. 105). Deliberation and reasoning on the Buddha's teaching were widely perceived to be necessary steps before meditation. Traditionally, the study of the Buddhist scriptures was divided into three steps: listening to the Buddha's words, reflecting on them, and meditating on them.

However, another term that is often used to convey the idea of knowledge is *kalā*, sometimes translated by "art and craft," refers to both "knowledge that" and "knowledge how." There are long lists of the various *kalās* (also called *śilpas*), some of them enumerating sixty-four, some seventy-eight, some more than ninety types. A typical list would include the knowledges of writing, calculation, sculpting, painting, dancing, singing, playing on musical instruments, gambling, speaking courteously, various games, preparing drinks, preparing perfumes, composing poems in various meters, divination, poisons and antidotes, the movement of heavenly bodies, training horses and elephants, archery, and various forms of fighting.

However, these terms for knowledge are not extensively treated in Indian philosophical texts, and except for the four Vedas, do not play an important role in Indian theories of knowledge. For Indian philosophers are not so much concerned with the nature of knowledge as such, but with the means of knowledge (*pramāṇa*).

PRAMĀṆA

To the question "how can one know something?" all Indian philosophers would answer unanimously: by having a means of knowledge. This answer may sound almost tautological and no two significant philosophers would understand the term in exactly the same manner. Nevertheless, the term *pramāṇa* played a crucial role in structuring the Indian epistemologies. It is around this concept, its definitions, and its varieties that Indian philosophy developed in its most dynamic period (roughly from the fifth to the twelfth century). The most important means of knowledge are sense perception (*pratyakṣa*), inference (*anumāna*), and verbal communication (*śabda*), under which sacred writings such as the Vedas or the teaching of the Buddha are subsumed.

What are the means of knowledge (*pramāṇa*)?

The number of means of knowledge that are accepted by the different schools of thought varies strongly. Madhyamaka Buddhists like Nāgārjuna, skeptics like Jayarāśi (Franco 1994), and monists of the Advaita-Vedānta tradition like Śrīharṣa, all of whom deny the possibility of knowledge, obviously accept no means of knowledge to be reliable (Matilal 1977). All other schools admit that sense perception is a means of knowledge. The materialist school (Lokāyata) is distinguished from other schools by its claim that only sense perception is valid. The Vaiśeṣikas and the Buddhists after Dignāga (fifth century) admit two means of knowledge, namely, perception and inference. The Sāṃkhyas admit verbal communication by a trustworthy person (*āptavacana*) besides these two; Buddhist philosophers before Dignāga, for example, Vasubandhu, also admit verbal communication to be a means of knowledge. Philosophers of the Nyāya tradition, with the notable exception of Bhāsarvajña (ninth century), also admit analogy (*upamāna*) as a fourth means of knowledge. The same position was held by certain Buddhists (Franco 2001). The Prābhākara Mīmāṃsakas accept five means of knowledge: the previously mentioned four and presumption (*arthāpatti*). The Bhāṭṭa Mīmāṃsakas and Advaita-Vedāntins admit six means of knowledge: the previously mentioned five and absence (*abhāva*) or nonperception (*anupalabdhi*). In nonphilosophical texts one also encounters inclusion (*sambhava*) and tradition (*aitihya*) as means of knowledge. Since inference and verbal communication are dealt with in separate entries, this entry will focus mainly on a discussion of perception.

PERCEPTION AND SENSES

Perception here refers primarily to sense perception. Indeed, the Sanskrit word that is usually rendered by *perception* is *pratyakṣa*; it contains the semantic element—*akṣa*—which means "eye." However, in some cases such as mental perception of feelings or the extrasensory perception of Yogis, the senses play no role in its arising. Perception is usually said to arise from sense and object. In this connection one has to emphasize the distinction between sense (or sense-faculty) and sense organ. The senses are not identical with the bodily organs to which they are associated. It is an extremely common mistake in Western publications to refer to the senses of seeing, hearing, smelling, touching, and tasting as eyes, ears, nose, skin, and tongue. Indian philosophers, however, clearly distinguish between them.

Thus, according to Nyāya the sense of sight is not the eye, but an invisible ray of light that rests on the pupil of the eye and goes out to reach the object. The sense of hearing is not the ear, but a part of space-ether (*ākāśa*) that is enclosed in the ear. The sense of taste is not the tongue, but a watery substance in the form of half-moon that is spread at the front of the tongue. The sense of smell is a substance made of earth and is found inside the nose; its base is usually called *nāsā*—a cognate of nose—

but sometimes also *tripuñikā*, that is, "the three cavities," or "the triple cavity," which seems to indicate that its base is the root of the nose. The sense of touch, which is sometimes interpreted as a sense of temperature, is also found inside and throughout the body, not only on the skin.

Already in the early philosophy of nature, the senses were considered to be material. Each sense—except for the auditory—was composed of the four material elements (earth, water, fire, and wind). Their special ability to grasp a certain elemental quality was explained as being due to their composition. The gustatory sense consists mainly of water, and it possesses the quality to be grasped, namely, flavor (*VS* 8.16–17). Although the element earth also possesses flavor, this quality is not predominant in it. The elemental constitution of the senses is based on the principle that "similar perceives similar." The Nyāya, Vaiśeṣika, and the Mīmāṃsā accepted the so-called accumulation theory of qualities in elements.

Except for hearing, the senses are made of special invisible atoms. Therefore, they cannot perceive themselves and can only be inferred: From the fact that one has a visual awareness, one infers that one has a sense of sight. According to the Buddhists the senses are made of a special subtle and transparent matter (*bhūtaprasāda*); the transparency of this matter is used to explain both its invisibility and its receptivity to other forms. Unlike normal matter, the subtle matter of which the senses are made does not obstruct other matter. When Indian philosophers write about the senses, they think above all about sight. The sense of sight is often used as a model for all other senses; hearing is treated cursorily, the other senses are hardly ever discussed.

PERCEPTION AND CONTACT

There was a strong debate that lasted for centuries between Buddhists and Naiyāyikas on the question of whether the sense and the object must be in contact to produce sense perception. The debate concerned only the senses of seeing and hearing (for everyone agreed that the other senses must be in contact with their objects). The Naiyāyikas and the Mīmāṃsakas maintained that all senses must be in contact with their objects to perceive them. In response to the Buddhist objection that sight perceives objects at a distance and objects that are larger in size than the sense itself, the Naiyāyikas postulated an invisible ray of light that goes from the eye and enters in contact with the object. This ray of light has a broad tip so that it can be in contact with large objects. It is in this context that certain optical theories were developed (Preisendanz 1989).

PERCEPTION AND THE CRITERION OF TRUTH

For a general discussion of truth and error, notably of false inferences, see the entry "Truth and Falsity in Indian Philosophy." The problem of truth is addressed here only in respect to perception. The earliest discussion on the criterion of truth can be found in a short passage of an anonymous Mīmāṃsā commentary that is now lost except in quotations and references in later sources that refer to its author simply as "The Commentator" (*vṛttikāra*) (Frauwallner 1968, pp. 107–111). It may seem odd that a Mīmāṃsā commentary that deals with Vedic exegesis should contain digressions on perception and related epistemological problems. Indeed, the rationale for the treatment of perception in Mīmāṃsā writings was originally a negative one: the rejection of sense perception as a means for the apprehension of the *dharma*, understood here as Vedic injunctions (*MS* 1.1.4).

According to the Commentator, "true perception is the arising of awareness when the senses of a man are in contact with precisely that which the awareness has for its object" (*ŚBh* 26.3–4). In other words, when the internal object that appears in the awareness and the external object that is in contact with the senses are identical, the resulting awareness is perception. This is, however, only a general definition. How can one know whether a specific awareness has arisen when the senses are in contact with the same object that appears in the awareness, or whether they were in contact with a different object? One may have an awareness of silver, but how is one to know whether the senses are in contact with silver, or with a glittering conch shell that produces an illusion of silver? The Commentator answers that a sublating awareness (*bādhaka-jñāna*) arises in respect to a false awareness and asserts its falsity, "That was not silver, the awareness was false." However, the problem with sublation (*bādhā*) as a criterion of truth is that the sublating awareness arises later, sometimes much later, than the false awareness. How does one know when an awareness is true or false at the time it arises? At that moment there is no difference whatsoever between true and false awarenesses, for the person who mistakes a conch shell for a piece of silver also thinks, "My sense of sight is in contact with silver."

The Commentator suggests that when the causal complex that produces the awareness is disturbed, the awareness is false; otherwise it is true. For instance, when the mind is disturbed by hunger, when the sense of sight is disturbed by an illness, or when the external object is too subtle, the awareness is false; when the causal complex is not disturbed, the awareness that arises from it is

true. By this assertion the Commentator makes the true awarenesses the normal ones, those people usually have, and errors are considered to be an exception. In other words, there is nothing inherently wrong in the cognitive process itself. However, the assertion that a true awareness is produced by undisturbed causes tells one what happens, but not when it happens. That a particular awareness has arisen from undisturbed causes remains to be proved for every single case. The Commentator maintains that if one earnestly searches and does not find any fault with the causal complex, then, because there is no proof to the contrary, we should think (*manyemahi*) that the awareness is true.

Later Mīmāṃsakas like Kumārila (seventh century) had to deal with problems that the Commentator had left open. For instance, in certain cases one is not in a position to rectify an erroneous awareness (*ŚV, Vṛttikāragrantha* 23). A certain illness of the eye distorts vision in such a way that one sees a double moon. In such cases the mistaken person learns in his or her communication with other people that there is only one moon in the sky. Kumārila also had to deal also with errors that are immanent to the cognitive process. Such errors would render all everyday awarenesses, even those that are usually considered true, essentially erroneous. For instance, according to the Buddhists, every empirical awareness involves a conceptual construction. Empirical awarenesses have wholes (*avayavin*) and universals (*jāti*) as their objects, but these have no correspondence in reality. Even a simple awareness such as "this is a cow" contains at least two parts. The part *this* refers to some concrete individual, the part *cow* to a universal "bovinity" that, at least according to the Indian realists, is a single eternal entity present in all cows and is responsible for the fact that a great number of different individuals are all called cow.

The Buddhists have adduced powerful arguments against the existence of such universals. For instance, the universal bovinity cannot be present entirely in one individual cow, because if this were the case, it would not be able to reside in other cows. Nor can it be partly present in one cow, because it has no parts. Thus, all empirical awarenesses are false because they involve conceptual constructions, and conceptual constructions are faulty because they involve incoherent notions such as that of a universal. Kumārila's response to such objections was to refuse a philosophical engagement. No matter what arguments the Buddhists may raise: If everybody invariably has the awareness in respect to a certain individual, "this is a cow," then such awareness cannot be sublated, for it is more powerful than the other awareness that has found fault in it.

The concept of sublation may seem to presuppose a coherence theory of truth, in which truth is defined by relations between statements (or in the Indian case, between awarenesses), not in terms of relations between statements and reality, as is the case in a correspondence theory of truth. However, in general Indian philosophers always seem to presuppose a correspondence theory of truth. Even though only an awareness can sublate, or assert the falsity of another awareness, this is possible only because the sublating awareness corresponds to reality and the sublated awareness does not. The direct relationship between the two awarenesses remained problematic, and in the final analysis unexplained. To the question of how an awareness that arises later can apprehend the inexistence of an object of an earlier awareness Jayanta, a Nyāya philosopher of the ninth century, simply replies, "What [can] we do, since this is the way the awareness arises?" (*NM* I 171.12)

The correspondence theory of truth is clearly presupposed by the Nyāya criterion of truth called efficiency of activity (*pravṛttisāmarthya*). The Naiyāyikas argued in favor of a pragmatic principle of confirmation. When one has an awareness of water, one goes toward the perceived water, and if this endeavor is efficient, that is, if one obtains water, then the awareness is true. Otherwise it is false (*NBh*, Introduction). The discussions of the efficiency of activity seem to presuppose a difference in the reliability of the senses. The awareness that has to be confirmed is usually a visual one, and the confirming awareness is of touch or taste (as in the case of water). The expression "efficiency of activity" is often interchangeable with the expression "obtainment of an object/purpose" (*arthaprāpti*). The Naiyāyikas argue that when the awareness is true the object is obtained, and when it is false the object is not obtained.

Another similar but different criterion of truth is used by Dharmakirti and his followers. Dharmakirti argues that the production of efficient action (*arthakriyākāritva*) indicates whether an awareness is valid or not. The difference between this and the Nyāya criterion is that the former is not used to prove that the object of the awareness is real. According to Dharmakirti a false awareness can nevertheless be valid. Although all awarenesses that involve conceptual constructions are false, some such awarenesses (notably inferential awarenesses that always involve universals) lead to successful activity. Dharmakirti likens their case to someone who mistakes diamond rays for the diamond itself (*PV*, 3.57). Although such a person acts on a false awareness, he or she is nevertheless successful in obtaining the diamond. Another

important aspect that distinguishes the Buddhist criterion from that of the Nyāya is that the object seen and the object obtained can never be the same. According to the Buddhists everything is momentary. Thus, the water seen and the water obtained are not the same water. Another difference between the two criteria is due to the rejection of the substance. The Buddhists denied that there is a certain substance such as water that has properties such a color and flavor. Thus, the seen water and the tasted water are in fact entirely different kinds of atoms that are only loosely connected by a causal relationship (*PVSV* 70.14f)

The preceding discussion treats the realistic schools. The topic of the criterion of truth in idealistic and illusionistic schools, which consider all empirical awarenesses to be false, arises from a different set of problems and specific metaphysical doctrines. For instance, certain Buddhist Yogācāras consider only those awarenesses to be true that have a correspondence in an unconscious awareness called *ālayavijñāna*. Vedāntins like Śankara (700?–750?) consider empirical awarenesses to be provisionally true until one attains the realization of the identity between ātman and brahman. Everyday awarenesses are like a dream. As long as the dream lasts, the awarenesses of the dream are considered true; when one wakes up they are realized to have been false. These positions, however, are usually ignored in the philosophical debates in classical India.

A SKEPTICAL RESPONSE TO THE CRITERION OF TRUTH

Jayarāśi Bhaṭṭa (fl. c. 800), a skeptic philosopher loosely affiliated to the materialist Lokāyata school, raised a devastating critique of the various criteria of truth. The production by undisturbed causes, he says, cannot be used as a criterion, because it cannot be known whether the causes are undisturbed. The senses do not apprehend themselves, and therefore, cannot apprehend whether their functioning is disturbed or not. Nor can their proper functioning be inferred, because there is no inferential sign on which the inference can rest. If the correct awareness itself is considered to be such a sign, then the argument results in mutual dependence. The awareness is correct because the causes are undisturbed, and the causes are undisturbed because the awareness is correct.

Also, the absence of sublation cannot be used as a criterion of truth. At most one can say that those awarenesses that are sublated are false, but not that those that are not sublated are true. It is possible that sublations do not arise because some causal factor is missing. A person may have an illusion of water in respect to sun rays and not go toward the place of the sun rays. Thus, the causal factor that could produce the sublation (the proximity) is absent and the sublation does not arise. Besides, one may simply die before the sublation is produced. It is impossible to know at any given moment which awarenesses are true and which are going to be sublated in the future. Jayarāśi's argument bears an obvious similarity to Karl Popper's assertion that the scientific doctrines one holds to be true are only those that are not yet refuted, but they are liable to be so in the future. Of course, the basic concerns of Jayarāśi and Popper are entirely different.

The efficiency of activity based on an awareness also cannot be used as a criterion of an awareness' truth because the claim of efficiency also has to be confirmed: it has to be apprehended and its apprehension has to be ascertained as nonerroneous by another efficiency of activity. It is not true that an awareness will give satisfaction if and only if it is true. To repeat James's example, the pragmatist claims that if one believes that there are tigers in India, and one goes to India and finds tigers there, then, to use the Nyāya terms, the activity is efficient and the awareness is true. However, as critics of pragmatism point out, one may go to Syria, find some tigers there and think that one is in India, or one may go to India and mistake some big cats for tigers, or one can even go to India find tigers and mistake them for cats. Thus, a confirming awareness must be confirmed in its turn, and this would lead to an infinite regress. The arguments against Nyāya apply to the Buddhist criterion of production of efficient action, except that the Buddhist faces some additional difficulties due to the doctrine of momentariness and the rejection of universals.

VERBAL COMMUNICATION

The two main questions with which Indian philosophers who deal with verbal communication are concerned are: (1) What is the process by which one understands the meaning of words? (2) How does one know that words, once understood, are truthful? Concerning the first question see the entry "Philosophy of Language in India." This entry will focus only on the second question.

The veracity of words is crucial to Indian philosophers because knowledge derived from the sacred writings depends on it. Clearly, most religious doctrines could not be established by other means of knowledge such as perception or inference. Furthermore, when one is faced with a plurality of religious traditions, the question invariably arises as to which tradition can be trusted, for all of them cannot be true. Thus, each tradition had to adduce some arguments to justify the teachings it considered to be true. According to the Nyāya-Vaiśeṣika the

Veda was revealed to normal human beings by the Rishis who have direct knowledge of it, and consequently the truthfulness of the Veda, at least as known to one, depends on the truthfulness of the Rishis. Vātsyāyana (fifth century) enumerates three characteristics that must be present if one is to be considered a trustworthy or authoritative person: One has to have direct knowledge of things, compassion toward living beings, and the desire to teach things as they are.

There are basically two ways to prove the validity of a statement made by a reliable person. Either the reliability of the person making the statement is established, or the truthfulness of the statement is directly perceived or inferred. Ideally, the statement should be directly confirmed, but in the case of the Veda this is not always possible, for the truthfulness of a Vedic statement is often beyond the realm of examination by normal human beings, for example, statements concerning heaven. Vātsyāyana's proof is based on the assumption that the different parts of the Veda have the same authors. The statements of the Āyurveda and magical spells (*mantra*), which according to Vātsyāyana form a part of the Veda, have visible results. When certain spells that are intended to remove poison are uttered, the poison is actually removed. Furthermore, certain parts of the Veda proper also have visible results, for example, "One desirous of a village should perform a sacrifice" (*grāmakāmo yajeta*). Vātsyāyana's inference of the validity of the Veda runs as follows: From the parts of the Veda that have visible results one infers the trustworthiness of its authors (qualified by the three characteristics mentioned earlier), and because these are the same trustworthy authors as those of the rest of the Veda, the validity of the latter can be inferred.

The proof of reliability of a person was further developed by Dharmakirti, who was concerned with the trustworthiness of the Buddha. It was clear to Dharmakirti, who was conscious of the problem of induction, that the argument as it appears in *Nyāyabhāṣya* and *Nyāyavārttika* is not valid: Just because someone is trustworthy in matter x (e.g., medicine) does not necessarily mean he or she is trustworthy in matter y (e.g., rituals and sacrifices). Consequently, Dharmakirti modifies the argument in two points. First, he does not simply draw an inference from trustworthiness in any part x to trustworthiness in any part y; he allows such an inference only when one moves from the main part of a teaching to its secondary part. Second, the logical reason used in Dharmakirti's inference is not just the sameness of the author, but includes the motivations of the speaker in his reasoning, for example, one should consider whether the speaker may have a motivation to lie.

More specifically the proof runs as follows: The main part of the Buddha's teaching are the four noble truths. These truths can be established independently of the Buddha's authority through perception and inference. Once the four noble truths are established, one can conclude that the Buddha was knowledgeable at least in matters of salvation. From such knowledge one infers that the Buddha has practiced various means for salvation for a long time (i.e., during many lives). However, he need not have practiced for such a long time had he been interested only in his own salvation. Therefore, his efforts were for the sake of other people. His engagement for the benefit of other (in fact, all) living beings in this manner presupposes compassion. Furthermore, the Buddha does not lie, because he has nothing to gain by lying. Therefore, the Buddha is trustworthy. Consequently, one can infer the truth in secondary matters in his teachings that are not open to an examination by normal human beings. As an example for such a domain Dharmakirti mentions the law of karma. Later Tibetan commentators also mention certain monastic rules that cannot be established independently of the Buddha's word (Tillemans 1993).

Interestingly, the reliability of the Gods must also be established. The Śaiva commentator Sadyajyotis (ninth century) says: Why is the word of Śiva authoritative? Because he is a pure, infallible, gracious lord endowed with knowledge that extends to everything. And his words whose objects are seen can be perceived as fruitful. Therefore, it can be inferred that his words whose objects are not seen are fruitful in exactly the same manner (Franco 1997, pp. 41–42).

THE OTHER PRAMĀṆAS

It is unfortunate that the other means of knowledge receive little attention in the Indian tradition. The Naiyāyikas and the Mīmāṃsakas have accepted analogy or comparison (*upamāna*) as a separate means of knowledge, but discussions about it remain rudimentary. It is defined as "proof of what has to be proved from similarity to something well known" (*NS* 1.1.6). The stock example for the use of analogy is: Someone does not know what a gayal is and is told "a gayal is like a cow." He or she then goes to the forest and is able to recognize a gayal on seeing it. Another example concerns the recognition of something from its name. For instance, knowing that the herb called bean leaf is like a bean, a person who finds this herb realizes that this is the thing to which the name applies. The Naiyāyikas were not unanimous as to what exactly constitutes the means of

knowledge in this case. The older Naiyāyikas argued that the statement of the instructing person is the means of knowledge; the later Naiyāyikas maintained that it is the cognition of similarity that brings about the understanding. Means of knowledge, by definition, must lead to an awareness of an object previously unknown, for if the object is already known, its awareness will be nothing but recollection, and, except for the Jainas, no school of thought accepted memory as a means of knowledge.

There was some uncertainty as to what exactly is new about the object of the awareness resulting from comparison. To repeat the stock example, when one recognizes that a certain animal is a gayal, it is not the animal as such that is the object of the comparison, because it is apprehended by sense perception. It is also not that there is a similarity between the cow and the gayal, because the similarity was already conveyed by verbal communication. Nor can the resulting awareness consist in the conclusion that the particular animal observed for the first time is a gayal, because in this case comparison would not be different from inference. Indeed, some Mīmāṃsakas who professed this opinion were criticized by the Naiyāyikas for reducing comparison to inference (Bhatt 1962, pp. 290ff). The Naiyāyikas (*NBh* 1.1.6) as well as some Buddhists of the Kushana period (Franco 2001, pp. 11–12) maintained that the result of comparison is the awareness of the designation, that is, that the animal seen in the forest is called gayal. Nevertheless, it remained controversial what distinguishes analogy from inference on the one hand and from verbal testimony on the other, and different opinions were put forward on this issue. The Buddhists, the Vaiśeṣikas, and the Sāṃkhyas did not consider analogy to be a separate means of knowledge (Bhatt 1962, pp. 289–307).

Another potentially interesting means of knowledge that remained underdeveloped is *arthāpatti*. There is no agreed translation for this means of knowledge, and it is rendered by presumption, supposition, implication, negative implication, circumstantial evidence, and so on. The two most common examples for *arthāpatti* are: (1) Knowing that someone is alive and not finding him or her at home, one concludes that he or she is outside. (2) One is told that fat Devadatta does not eat during the day, and one concludes that he eats at night. The two examples are distinguished as presumption based on something seen (*dṛṣṭārthāpatti*) and presumption based on something heard (*śrutārthāpatti*). In later texts one distinguishes six types of presumption according to the six means of knowledge on which a presumption can be based.

The examples mentioned in this connection seem construed and artificial and are not taken from an actual philosophical discourse or from everyday life. For instance, presumption based on inference is illustrated as follows: One knows by inference that the sun moves (its movement cannot be perceived, but is inferred because it changes its place in the sky). However, things that move usually possess limbs such as legs. Thus, a conflict between two means of knowledge arises, and this conflict is resolved by the presumption that the sun has a moving power. Conflict or apparent contradiction (*anupapatti*) between two means of knowledge is the essential ingredient of *arthāpatti*, and the resulting presumption resolves the conflict. The contradiction must be apparent. If the contradiction is real, for example, two awarenesses about the same object, one perceiving it as silver and the other as mother-of-pearl, the way of resolving it is by rejecting one of the alternatives as false, not by making a new supposition. Among the important philosophical schools, only the Mīmāṃsā and Vedānta accepted presumption as an independent means of knowledge (Bhatt 1962, pp. 313–340).

The Bhāṭṭa Mīmāṃsakas accepted absence (*abhāva*) as a sixth means of knowledge. A discussion as to how mere absence or nonexistence can be an object of valid cognition appears already in *NS* 2.2.7–12. An objector argues that a negating cognition cannot be valid because it cannot refer to an object in reality. The objection is rebuked by reference to common experience. When some pieces of cloth are marked and some are unmarked, one can be told "Fetch the unmarked pieces," and one is able to do so. The Naiyāyikas, however, just like the Vaiśeṣikas, the Sāṃkhyas, the Buddhists, and the Prābhākara Mīmāṃsakas, considered absence or nonperception to be included in inference. Praśastapāda identified absence with inference from absence of effect to absence of cause.

CIRCULARITY OF PRAMĀṆAS

A general objection to the *pramāṇas* as such has been raised from the earliest times. If everything is established by means of knowledge, how are the means of knowledge themselves established? If they are established by other means of knowledge, these other means also have to be established by yet other means of knowledge and thus an infinite regress results. If the means of knowledge were to establish one another, a circularity would result. If one claims that the means of knowledge need not be established, the initial position that everything has to be established by means of knowledge has been abandoned. Some claimed that the means of knowledge establish both their objects and themselves, just as a lamp illuminates itself

and its surroundings. However, it remained unclear how this metaphor should actually apply to the *pramāṇas*, and some, like Nāgārjuna (*VV*, verses 30ff) even argued that actually a lamp cannot illuminate itself.

See also Atomic Theory in Indian Philosophy; Causation in Indian Philosophy; Liberation in Indian Philosophy; Logic, History of: Logic and Inference in Indian Philosophy; Meditation in Indian Philosophy; Mind and Mental States in Buddhist Philosophy; Philosophy of Language in India; Self in Indian Philosophy; Truth and Falsity in Indian Philosophy; Universal Properties in Indian Philosophical Traditions.

Bibliography

PRIMARY SOURCES

NBh: Nyāyabhāṣya of Vātsyāyāna in Taranatha Nyayatarkatirtha and Amarendramohan Tarkatirtha (ed.). *Nyāyadarśanam with Vātsyāyana's Bhāṣya, Uddyotakara's Vārttika, Vācaspati Miśra's Tātparyaṭīkā and Viśvanātha's Vṛtti*. Calcutta, India: 1936, 1944.

NM: Nyāyamañjarī in V. K. S. Varadacharya (ed.) *Nyāyamañjarī of Jayanta Bhaṭṭa with Ṭippaṇi—Nyāyasaurabha by the Editor*, Mysore: Prācyavidyāsamśodhanālayaḥ, Maisūruviśvavidyālayaḥ, 1969–1983.

NV: Nyāyavārttika of Uddyotakara. See *NBh*.

NS: Nyāyasūtra in Ruben, Walter, *Die Nyāyasūtra's. Text, Übersetzung, Erläuterung und Glossar*. Leipzig, Germany: 1928.

PV: Pramāṇavārttika of Dharmakīrti. Ed. Y Miyasaka. *Acta Indologica*, Vol. 2. Narita, Japan: Naritasan Shinshoji, 1971–1972.

PVSV: Prāmāṇavārttikasvavṛtti in R. Gnoli (ed.), *The Pramāṇavārttikam of Dharmakīrti*. Rome, Italy: 1960.

MS: Mīmāṃsāsūtra in Frauwallner 1968.

VV: Vigrahavyāvatanī of Nāgārjuna in K. Bhattacharya E.H. Johnston and A. Kunst, *The Dialectical Method of Nāgārjuna*. Delhi, India: 1978.

VS: Vaiśeṣikasūtra in Muni Jambuvijayaji (ed.), *Vaiśeṣikasūtra of Kaṇāda with the Commentary of Candrānanda*. Baroda, India: Oriental Institute, 1961.

ŚBh: Śābarabhāṣya in Frauwallner 1968.

ŚV: Ślokavārttika of Kumārila in Svāmī Dvārikadāsa Śāstrī (ed.), *Ślokavārttika of Śrī Kumārila Bhaṭṭa. With the Commentary Nyāyaratnākara of Śrī Pārthasārathi Miśra*. Varanasi, India: Tara Publications, 1978.

SECONDARY SOURCES

Bhatt, Govardhan P. *Epistemology of the Bhāṭṭa School of Pūrvamīmānsā*. Varanasi, India: Chowkhamba Sanskrit Series Office, 1962.

Chakravarty, Pulinbihari. *Origin and Development of Sāṃkhya System of Thought*. 2nd ed. Delhi, India: 1975.

Franco, Eli. *Perception, Knowledge, and Disbelief: A Study of Jayarāśi's Scepticism*. 2nd ed. Delhi, India: 1994.

Franco, Eli. *Dharmakīrti on Compassion and Rebirth*. Vienna, Austria: Arbeitskreis für Tibetische und Buddhistische Studien, Universität Wien, 1997.

Franco, Eli. "Fragments of a Buddhist *Pramāṇa*-Theory from the Kuṣāṇa Period." *Bukkyō Dendō Kyōkai Fellowship Newsletter* 4 (2001): 2–12.

Frauwallner, Erich. *Materialien zur Ältesten Erkenntnislehre der Karmamīṃāṃsā*. Vienna, Austria: Graz, Köln, Böhlau in Kommission, 1968.

Matilal, Bimal Krishna. *The Logical Illumination of Indian Mysticism: An Inaugural Lecture Delivered before the University of Oxford on 5 May 1977*. Oxford, U.K.: Clarendon Press, 1977.

Namai, Mamoru. "A Survey of Bārhaspatya Philosophy." *Indological Review* (Kyoto) 2 (1976): 29–74.

Pollock, Sheldon. "The Theory of Practice and the Practice of Theory." *Journal of the American Oriental Society* 105 (3) (1985): 499–519.

Preisendanz, Karin. "On *Ātmendriyamanorthasannikarṣa* and the Nyāya-Vaiśeṣika Theory of Vision." *Berliner Indologische Studien* 4–5 (1989): 141–213.

Seyfort Ruegg, David. *Ordre spirituel et ordre temporel dans la pensée bouddhisque de l'Inde et du Tibet: quatre conférences au collège de France*. Paris: Collège de France, Institut de civilisation indienne, Diffusion De Boccard, 1995.

Tillemans, Tom J. F. *Persons of Authority*. Stuttgart, Germany: F. Steiner, 1993.

Eli Franco (2005)

KNUTZEN, MARTIN

(1713–1751)

Martin Knutzen, the German Wolffian philosopher, studied at the University of Königsberg and became an extraordinary professor there in 1734. Because he was a Wolffian, even though an unorthodox one, he never attained a full professorship in that Pietist-dominated school. However, because he was also a Pietist, Knutzen could never attain such a position in other German universities where Wolffians held the power of appointment.

Knutzen disagreed with Christian Wolff on several significant points. His *Commentatio Philosophica de commercio Mentis et Corporis* (Philosophical Commentary on the Relation between Mind and Body; Königsberg, 1735) was an attempt to reconcile Wolff's theory of preestablished harmony with the Pietist doctrine of physical influence. He extended the problem beyond Wolff, from the relation of soul and body to the interrelations of simple substances in general. In this and in a panpsychistic metaphysics, he was closer to Gottfried Wilhelm Leibniz than to Wolff. Knutzen, in his cosmological work *Vernünftige Gedanken von den Cometen* (Rational thought concerning comets; Königsberg, 1744), was one of the first philosophers in Germany to accept, at least partially, the Newtonian theory of gravitational attraction. His theological work was derivative and of little significance.

Knutzen's reputation is due more to his having been the teacher of Immanuel Kant than to his own significance. His influence on Kant has been much overrated. Recent research has shown that his influence was confined to the solution given by Kant in his first essay, *Gedanken von den wahren Schätzung der lebendigen Kräfte* (Thoughts on the true estimation of living forces; Königsberg, 1747), to the problem of the interrelation of substances, and to Kant's acceptance of Newtonian attraction. On the second point, Kant was also strongly influenced by the Berlin circle around Pierre-Louis Moreau de Maupertuis, even though Maupertuis himself was reluctant to accept attraction; and in accepting attraction as a real force and in trying to give a metaphysical explanation for it, Kant went beyond the Berlin circle, Knutzen, and Isaac Newton himself in his published statements.

Both Kant's "Wolffianism" and his "Pietism" have been attributed by some historians to Knutzen's influence; but although Kant received a Pietist education, he was never either a Pietist or a Wolffian. Kant always opposed Wolff's doctrines, and any Pietist influence came through the general philosophical influence of C. A. Crusius. Even an alleged influence of Knutzen's theology on Kant's religious philosophy has been disproven.

See also Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY KNUTZEN

Dissertatio metaphysica de Aeternitate Mundi Impossibili. Königsberg, 1733.

Philosophischer Beweis von der Wahrheit der christlichen Religion. Königsberg, 1740.

Commentatio Philosophica de Humanae Mentis Individua Natura sive Immortalitate. Königsberg, 1741.

WORKS ON KNUTZEN

Biéma, M. van. *Martin Knutzen, La critique de l'harmonie préétablie*. Paris, 1908.

Bohatec, J. *Die Religionsphilosophie Kants*. Hamburg, 1938.

Erdmann, Benno. *Martin Knutzen und seine Zeit*. Leipzig: L. Voss, 1876.

Tonelli, Giorgio. *Elementi metodologici e metafisici in Kant dal 1745 al 1768*. Vol. 1. Turin, 1959. Chs. 1 and 2.

Watkins, Eric. "The Development of Physical Influx in Early Eighteenth-Century Germany: Gottsched, Knutzen, and Crusius." *Review of Metaphysics* 49 (2) (1995): 295–339.

Watkins, Eric. "Forces and Causes in Kant's Early Pre-Critical Writings." *Studies in History and Philosophy of Science* 34a (1) (2003): 5–27.

Giorgio Tonelli (1967)
Bibliography updated by Tamra Frei (2005)

KOCHEN-SPECKER THEOREM

See *Quantum Mechanics*

KOFFKA, KURT

(1886–1941)

Kurt Koffka, one of the three founders of the Gestalt movement in psychology, was born in Berlin. In 1903 he went to the university there to study philosophy, and he is said to have had a special interest in Immanuel Kant and Friedrich Nietzsche at that time. In 1904 he moved to Edinburgh, and in the next few years his interest in psychology became increasingly strong. Soon after receiving his doctorate at Berlin in 1908, he moved to Würzburg, where he served as an assistant to Oswald Külpe and Karl Marbe. In 1910–1911 he taught at the Academy at Frankfurt am Main, and it was during this period, as a result of the joint deliberations of Max Wertheimer, Wolfgang Köhler, and himself, that the central notions of Gestalt theory began to emerge. In 1911 Koffka became a lecturer at the University of Giessen, and from 1919 to about 1927 he was assistant professor.

The early 1920s saw the founding of *Psychologische Forschung,* a periodical in which several of the original articles on Gestalt theory were originally published, and of which Koffka was for many years the editor. During this decade he traveled extensively: A visit to Oxford for the International Congress of Psychology in 1923 resulted in much wider recognition of Gestalt theory than had hitherto been possible, and in succeeding years he was visiting professor at Cornell, Chicago, and Wisconsin. In 1927 he took up permanent residence in the United States, having accepted a professorship at Smith College, Northampton, Massachusetts. In 1932, at the invitation of the USSR State Institute, he joined an expedition to Uzbekistan to carry out ethno-psychological research, but at an early stage he was forced to return because of illness. He remained intellectually active until his death. He is said to have been a person of considerable kindness and charm, with wide interests that included music, art, and travel. His friendship with Wertheimer and Köhler was lifelong.

To separate Koffka's distinctive contributions from those of Wertheimer and Köhler is not easy, since each was influenced considerably by the other two. Koffka's *The Growth of the Mind* was an attempt to apply Gestalt principles to child psychology, while *Principles of Gestalt Psy-*

chology was a comprehensive account of a wide range of psychological work up to 1935, with detailed theoretical discussion. One of his central claims was that it is possible to take seriously the advances of science while still finding a place for the concepts of meaning and value; indeed, scientific inquiries themselves suffer if one does not do so. An aggressive materialism or behaviorism was quite foreign to him, but the alternative to this for Koffka was a new approach, using the concept of Gestalt, rather than a return to vitalism or Cartesian dualism. In an interesting passage in *Principles of Gestalt Psychology* he called attention to the difference in intellectual climate between Germany and America. The more abstract and speculative ideas, in which many German scholars were interested, had to be kept in the background when Gestalt theory was presented to the Americans, whose "high regard for science, accurate and earthbound" was accompanied by "an aversion, sometimes bordering on contempt, for metaphysics that tries to escape from the welter of mere facts into a loftier realm of ideas and ideals" (p. 18).

Philosophically interesting contributions found in *Principles of Gestalt Psychology* include the distinction between the geographical and behavioral environments, a discussion of the criteria by means of which "things" in the behavioral environment are distinguished from "not-things," and an attempt to reinstate the concept of ego. The behavioral environment is, in effect, the perceived world, the world of commonsense experience, whereas the geographical environment is the world as studied by the physical scientist. There are features in the geographical environment (such as infrared rays) that in ordinary circumstances are not present in the behavioral environment, whereas there are features in the behavioral environment (for example, the fact that two lines are grouped together when someone looks at them) that have no direct counterpart in the geographical environment. Examples of "things" are sticks, stones, clouds, and some types of fog; marginal cases are waves, words, and noises, while "a fog that makes our ocean liner reduce speed and sound its piercing horn is not thing-like at all, as little as the mist from which we emerge when we climb a mountain" (ibid., p. 70). The three characteristics of things are "shaped boundedness, dynamic properties, and constancy." As for the ego, "it has a very definite place in that [the behavioral] world, and well-defined, if variable boundaries…. 'In front,' 'to the left and right,' 'behind,' and 'above and below' are characteristics of space which it possesses with regard to an object which serves as the origin of the system of spatial co-ordinates" (ibid., p. 322). In this case science itself is seriously impoverished if the concept of the ego is simply ignored. The study (sometimes called phenomenology) of how the world appears at the commonsense level is logically independent, according to Koffka's view, of any new discovery in physics about what is "really" happening.

Many of the problems that Koffka raised are of current philosophical interest, and as a psychologist he ranks among the greatest of his generation.

See also Dualism in the Philosophy of Mind; Gestalt Theory; Kant, Immanuel; Köhler, Wolfgang; Külpe, Oswald; Nietzsche, Friedrich; Vitalism.

Bibliography

WORKS BY KOFFKA

Besides numerous articles published in journals, Koffka's works include *Principles of Gestalt Psychology* (New York: Harcourt Brace, 1935); *The Growth of the Mind,* translated by R. M. Ogden (London: Routledge and Kegan Paul, 1924); "Mental Development," in *Psychologies of 1925,* edited by Carl Murchison (Worcester, MA: Clark University, 1928), pp. 129–143; and "Some Problems of Space Perception," in *Psychologies of 1930,* edited by Carl Murchison (Worcester, MA: Clark University, 1930), pp. 161–187. See also "Perception: an Introduction to the *Gestalt-theorie.*" *Psychological Bulletin* 19 (1922): 531–585; "Problems in the Psychology of Art," in *Art: a Bryn Mawr Symposium* (Bryn Mawr, PA: Bryn Mawr College, 1940).

SECONDARY WORKS

For discussions of Koffka's work, see G. W. Hartmann, *Gestalt Psychology; A Survey of Facts and Principles* (New York: Ronald Press, 1935). See bibliography to the Gestalt Theory entry for works discussing Gestalt psychology as a whole.

ADDITIONAL SOURCES

Arnheim, R. "The Two Faces of Gestalt Psychology." *American Psychologist* 41 (1986): 820–824.

Ash, Mitchell G. *Gestalt Psychology in German Culture 1890–1967: Holism and the Quest for Objectivity.* Cambridge, U.K.: Cambridge University Press, 1995.

Harrower, Molly. *Kurt Koffka, an Unwitting Self-Portait.* Gainesville: University Presses of Florida, 1983.

Petermann, Bruno. *The Gestalt Theory and the Problem of Configuration.* Translated by Meyer Fortes. New York: Harcourt Brace, 1932.

Smith, Barry, ed. *Foundations of Gestalt Theory.* Munich: Philosophia, 1988.

T. R. Miles (1967)

Bibliography updated by Alyssa Ney (2005)

KÖHLER, WOLFGANG
(1887–1967)

Wolfgang Köhler, the German Gestalt psychologist, was born in Tallinn, Estonia. He studied first at the University of Tübingen and then at Bonn. He next studied physics under Max Planck and psychology under Carl Stumpf at the University of Berlin, and received his PhD from that school in 1909 for investigations on hearing. In 1911 he became Privatdozent at Frankfurt. Max Wertheimer came to Frankfurt in 1912, and in the same year Köhler and Kurt Koffka served as the subjects for Wertheimer's famous experiments on stroboscopic motion that are widely regarded as the beginning of Gestalt psychology.

In 1913 Köhler became director of the anthropoid experiment station operated by the Prussian Academy of Sciences at Tenerife in the Canary Islands, and he remained there, throughout World War I, until 1920. The pioneering studies in the psychology of chimpanzees that he carried out there were published in several papers and in the monograph *Intelligenzprüfungen an Anthropoiden* (*The Mentality of Apes,* 1917).

Köhler's next major work, *Die physischen Gestalten in Ruhe und im stationären Zustand* (Physical Gestalten in rest and in the stationary state), was published at Brunswick in 1920. It is primarily a work in physics and reveals Köhler's indebtedness to Planck, but its major themes played important roles in his more strictly psychological writings.

In 1921, with Wertheimer, Koffka, Kurt Goldstein, and Hans Gruhle, Köhler founded the journal *Psychologische Forschung,* which served as the leading organ of the Gestalt psychologists until Köhler was forced to suspend publication because of the difficulties of editing it from the United States. In 1922 Köhler succeeded Stumpf as director of the Psychological Institute and professor of philosophy at the University of Berlin. He held a visiting professorship at Clark University in the academic year 1925–1926 and returned to America for another visit in 1929. In the same year his *Gestalt Psychology* was published in English.

Köhler was the only leading member of the Gestalt school who was not Jewish, but he was strongly opposed to the Nazis. He published a letter against them in a Berlin newspaper after they took power and a bit later left Germany. Köhler gave the William James Lectures at Harvard in 1934 and published them as *The Place of Value in a World of Fact* in 1938. In 1935 he was appointed professor of psychology at Swarthmore College. His Page-Barbour Lectures given at the University of Virginia in 1938 were published in an expanded version in 1940 as *Dynamics in Psychology.* Köhler became professor emeritus at Swarthmore in 1957. In 1959 the school awarded him an honorary doctorate and he became visiting research professor at Dartmouth, a position he retained until his death.

Köhler is correctly thought of primarily as a psychologist. Nevertheless, throughout his career he never hesitated to interpret the results and methodology of the physical sciences and to apply his interpretations to the delineation of the proper task of psychology and to the elucidation of its problems. He admitted a debt to the phenomenology of Edmund Husserl, and his own work was broadly in the phenomenological stream. Both phenomenology and physics influenced his vocabulary, his methods of research, and his theoretical conclusions. Köhler was an ardent controversialist, and he engaged in a continuing polemical defense of the Gestalt theory. He believed that the theory offered a new resolution of the controversy between those who believe in innate ideas or tendencies and those who stress the importance of ideas acquired by learning. He thought that his Gestalt physics could resolve the biological controversy between mechanism and vitalism. He claimed to have dissolved the philosophical controversies between idealism and realism and between monism and dualism, and he advocated a form of epiphenomenalism or even an identity theory of mind and body. Köhler believed that by phenomenological analysis he could demonstrate both the existence and something of the nature of value, and that value, or "requiredness," was more general than moral philosophers and aestheticians believed; thus, he held, the psychologist's investigation of value was of prime importance to the philosopher.

Köhler, then, not only advanced psychological theories and views about the proper subject matter of this science but also presented well-reasoned opinions on speculative problems in biology, physiology, physics, and chemistry, and suggested possibly fruitful lines of research for these sciences to undertake. He also presented theories belonging to such central philosophical disciplines as epistemology, metaphysics, and value theory. This entry will discuss some of the philosophically interesting issues raised by Köhler in the physical sciences and psychology, as well as some of his general philosophical positions. It will not attempt to discuss his contributions to Gestalt psychology proper, except for his discussion of isomorphism.

PHYSICS AND PHYSIOLOGY

Köhler discussed physical concepts and discoveries for at least three main purposes: to demonstrate the existence of physical structures analogous to perceptual gestalten; to provide a physicochemical theory of perception and other mental functions; and to delineate the proper task of psychology by comparing its present status with the status of physics at various times in its history.

PHYSICAL GESTALTEN. Köhler, like the other Gestalt psychologists, claimed that a central subject of psychology is the investigation of certain kinds of structures in which "the whole is more than the sum of its parts." An analysis of these gestalten would explain many puzzling facts of vision, touch, hearing, memory, and understanding. The existence of such structures was denied on the ground that the whole can never be more than the sum of its parts. Köhler sought to show that there are a variety of recognized physical systems in which the whole is more than the sum of its parts. Machines are structures whose movements are strictly determined. From a knowledge of the parts of a machine and their interrelationships, we can know the motions of the whole. Thus a machine, according to Köhler, is no more than the sum of its parts. But in many physical systems it is the state of the whole that determines the state of the parts. Examples of such systems are the distribution of an electrical charge over the surface of a conductor, which varies with the shape of the conductor; the distribution of a current of electricity or fluid in a network of wires or pipes; the distribution of particles of a fluid body whose only constraint is the walls of the container; and a planetary system. The common characteristic of these systems is that the parts interact dynamically rather than mechanically. And in these systems, he claimed, the whole is greater than the parts.

These physical systems all exhibit another characteristic, which Köhler thinks is strikingly analogous to a characteristic of phenomenal gestalten. When the physical systems are disturbed, the interaction of their parts tends more or less rapidly to restore the systems to a state of equilibrium. They are thus dynamically self-regulating systems. Phenomenal gestalten are also dynamically self-regulating. The parts of the gestalten interact with one another to produce, or reproduce, systematic wholes within the perceptual field. Köhler recognizes, following Wertheimer, a set of five factors involved in the recognition of gestalten. If any of these factors are present, then we tend to perceive a gestalt, unless inhibiting factors are also present or the factors are so present as to cancel out one another. The five factors are (1) *proximity*: Objects that appear close together are more likely to be classed as part of the same gestalt than those which are far apart; (2) *similarity*: Objects that resemble each other tend to be classed as belonging together; (3) "*common destiny*": If objects move or change together, they tend to be perceived as part of the same thing or as belonging together; (4) "*good gestalt*": Forms that are not quite regular tend to be perceived as more regular than they are; (5) *closure*: Forms that are in some way incomplete tend to be perceived as complete—for example, a circle with a small arc missing will be perceived as a full circle.

The resemblance between dynamically self-regulating physical systems and phenomenal gestalten suggested to Köhler that it might be more fruitful to attempt to understand mental phenomena by means of a dynamic rather than a mechanical model, and in fact this model continued to serve Köhler throughout his career as a fruitful explanatory hypothesis in psychology. He was particularly successful in applying it to problems of perception, of memory, and of intelligence or insight—of coming to understand a situation or a problem.

Despite Köhler's apparent success in applying the two notions that in certain physical and phenomenal structures the whole is greater than the sum of its parts and that psychological phenomena should be interpreted dynamically rather than mechanically, they have been widely criticized. Both notions, it is said, are enormously vague. It is not surprising that they seem to "work," for by their very vagueness they can be made to fit almost any body of facts. Surely in some generally accepted sense of "whole" and "part" almost any whole can be shown to be greater than the sum of its parts. But it is not clear that Köhler was applying the two terms univocally in the phenomenal cases he adduced as examples, and it is even less clear that he was using them in the same sense when speaking of the parts of phenomenal gestalten and of the parts of physical systems. Similarly, although the dynamic model may have aided Köhler in the design of new experiments and the interpretation of many phenomenal facts, it has been claimed that, outside of a certain limited range of cases, the apparent use of a dynamic model can mean no more than a recognition that phenomena change. The substance of the theory is probably Wertheimer's set of dynamic factors, which had in large part been anticipated by earlier psychologists, and there seems no reason to connect them with any specific physical theory.

ISOMORPHISM. Probably the most central concept in all of Köhler's thought is isomorphism, or similarity of form. He used this notion for two major and several

minor purposes. The two major functions combine into a theory of knowledge that is partly conceptual and partly physicochemical and physiological. Köhler distinguished between (1) phenomena, or percepts; (2) their cortical correlates, or brain-states; and (3) nature, or the physical world. He was perfectly willing to believe that percepts and brain-states may eventually be shown to be identical and in this sense does not exclude the possibility of a metaphysical monism. He holds, in opposition to both phenomenalists and new realists, that the phenomenal world and the physical world are not identical, and thus is an epistemological dualist. (These points are discussed below.) It is the theory of isomorphism that serves as the connecting link among these three elements. Percepts, it is claimed, are related to one another within the phenomenal field as their cortical correlates are related to one another in the cortex and as the corresponding physical objects are related to one another in physical space. The structural relations within any of the three realms are reproduced in the others. If a man-percept appears in phenomenal space atop a horse-percept, then in physical space there is a man atop a horse, and in the brain there are two brain processes dynamically related to each other in the cortical correlative of the relation "on top of."

What concerns us here is the isomorphy between the phenomenal world and brain-states. In this connection Köhler formulated the principle of isomorphism for spatial relations (it can be formulated for any type of phenomenal ordering) as: "*Experienced order in space is always structurally identical with a functional order in the distribution of underlying brain processes*" (*Gestalt Psychology,* Mentor edition, New York, 1959, p. 39). The parts of the visual field are not independent of one another; they exhibit structural relationships. If, for example, there is in my visual field a white square on a black ground, then in my brain there are processes corresponding to the white square, the black ground, and the boundary between the two. The topological relations between the brain processes are functionally identical with the corresponding visual relations. Metrical relationships are not preserved, but such relationships as betweenness are. In memory, these relationships are preserved in memory-traces. Thus it is form or structure rather than exact pictorial images that are preserved.

Köhler holds that the physiological processes in the brain that are involved in perception and memory are very probably electrochemical in nature. In the case of the white square, the brain process corresponding to the square-percept contains a higher concentration of ions than the brain process corresponding to the black ground. The two processes are functionally connected at a boundary corresponding to the edge of the square. There is a potential difference across this boundary; an electric flow of ions therefore takes place, and the square is perceived. Changes in the solution leave memory traces, which are subject to alteration in the course of time. These traces are superimposed on one another and thus functionally mirror the order of time of the percepts themselves.

The theory of isomorphism, both in its conceptual outline and in its physiological accompaniment, has been only inadequately outlined here. The physiological element, despite the important role it plays in Köhler's claim that functionally an identity theory of mind and body is at least feasible, is a matter for empirical investigation. Much of what Köhler says sounds rather plausible, but there are difficulties in stating the theory with the proper degree of precision. Although he speaks of a cortical retina, Köhler does not mean that perception involves the reproduction of a (two-dimensional or three-dimensional) image of the object within the cortex. This would be complete isomorphism. On the other hand, almost any set of relationships can represent any other by some form of correspondence, and the correspondences, if any, actually involved in perception might be very complex or in some other way not what we would intuitively grasp as a correspondence.

There are other issues involved that can only be raised and not explored here. Suppose it were established that when a certain macroscopic brain-state is observed in people, they generally claim to perceive a certain object. For instance, take any of the reversible figures that appear to an observer now in one way and now in another, such as a Maltese cross, composed of alternating black and white rays, which can be seen in two different ways. In one way of looking at it certain parts appear as the figure and the others as ground, while in the other way what was ground appears as figure and what was figure appears as ground. According to Köhler, each way of seeing the figure corresponds to a different electrochemical state in the brain. Now suppose that one person's descriptions of the cross fail to correspond, in either a regular or irregular manner, to the descriptions that we have generally found associated with his brain-states. We may wish to claim that he is misdescribing what he is seeing. But how we choose to regard the situation is not merely a matter of fact; it involves at least one conceptual matter, a choice between conflicting criteria of what the person is seeing—the person's description (which is, of course, the only criterion we now have) and our knowl-

edge of his brain-states. And empirical investigation alone cannot settle this conflict.

The same point applies to another example, in which a further factor becomes apparent. There is experimental evidence that when people see two parallel lines close to each other, one of which extends beyond the other at each end, they claim to see shadowy lines connecting the ends of the two lines to complete a trapezoid. Köhler suggests that the shadowy lines are caused by potential barriers in the cortex created by the cortical correlates of the lines actually drawn. Again, if it could be shown that such potential barriers are present in a person's brain although he claims not to see such lines, we might put it down to misdescription. But surely here we are inclined to take him at his word. In the first case we can describe what it means to see the cross in one way rather than another. But in this case we can only point out where the shadowy lines ought to be seen. The achieving aspect of perception is perhaps more obvious here. It is not simply a matter of what is seen but also of how we learn to describe what we see. In most descriptions it is clear what the standards of an accurate description are, and we can understand a proposal for a change in standards. In the present case it is not even clear what the standards are, if there are any. It is this element of conventional standards, which Köhler has omitted from his discussion, that makes his problems of the relationship among percepts, objects, and brain-states not merely a matter of physiological and psychological experimentation but of conceptual analysis.

Isomorphism and language. Köhler developed an interesting linguistic theory as a corollary of his theory of isomorphism. This corollary, except for Köhler's added complexity, resembles the picture theory of meaning advanced by Ludwig Wittgenstein in his *Tractatus Logico-Philosophicus* and seems to have been developed out of similar considerations. If the only way one thing can represent another is by having the same form, then the only way language can represent a situation is through a common form. Since, according to the theory of isomorphism, a phenomenal event has a physiological correlate possessing a similar form, then language represents both the event and the physiological correlate indifferently. A statement ostensibly about an observed phenomenon can be interpreted as a statement about brain-states and vice versa: ". . . language . . . is the peripheral outcome of antecedent physiological processes, among others of those upon which my experience depends. According to our general hypothesis, the concrete order of this experience pictures the dynamic order of such processes. Thus, if to me my words represent a description of my experiences, they are at the same time objective representations of the processes that underlie these experiences. Consequently, it does not matter very much whether my words are taken as messages about experience or about these physiological facts. For, so far as the order of events is concerned, the message is the same in both cases" (*Gestalt Psychology,* p. 40).

PHYSICS AND PSYCHOLOGY. The third way in which Köhler has used physics is to elucidate what he regards as the proper program for psychology. Physics, in his view, is an old, established discipline whose techniques have been developed and refined over a long period of time. Quantitative methods and pointer readings are appropriate in physics because there are thoroughgoing and widely accepted theories that give meaning to the numbers arrived at. Even in the early days of physics, in the time of Galileo Galilei, many of the problems could be investigated quantitatively, because the phenomena investigated had long been known from everyday life and this knowledge provided the necessary qualitative meaning. Where everyday life did not supply the necessary qualitative background, as in the study of electricity, physics had to proceed by qualitative investigations before quantitative ones could be undertaken profitably. The problems of psychology, Köhler claims, are more often like those of electricity than those of Galilean mechanics. In general, in psychology the necessary meaning-giving theory is absent. Intelligence quotients are notoriously hard to interpret. The difficulty in assessing their significance arises out of a lack of any clear notion of what intelligence consists in. Psychology should first try to develop a theory of intelligence before it tries to measure intelligence. Until a satisfactory theory is arrived at, it can hardly be determined whether or not intelligence quotients do measure intelligence and how well they do it.

GESTALT PSYCHOLOGY

CRITIQUE OF BEHAVIORISM. Köhler's attempt to show that qualitative methods are the most appropriate in the present state of psychology arose in the context of his repudiation of behaviorism. His phenomenological view of the nature of the subject matter of psychology was radically different from the notion that psychology is the study of behavior, with its related stimulus-response physiological theory. The behaviorists, according to Köhler, have taken too much to heart one epistemological teaching but ignored its wider context. They seek to limit psychology to the observation of the response of human beings in scientifically controlled situations because they

have become aware of the truth that one person cannot directly observe another person's experience. However, the behaviorist cannot avoid the study of direct experience by limiting himself to the observation of human reactions in controlled situations, for the only evidence he has of such reactions is his own experience. The behaviorist seeks to be objective, but he confuses two pairs of meanings of the terms *subjective* and *objective.* In one sense, observations of another person's reactions are no less subjective than my hearing his statements about what he is experiencing: Both are part of my experience. But in the primary sense *subjective* and *objective* refer to differently characterized phenomena within my experience. In this sense there is no reason why I cannot examine both subjective and objective experience; in the first sense I cannot help but investigate subjective phenomena.

CRITIQUE OF INTROSPECTIONISM. Whereas Köhler criticized behaviorism for misunderstanding the nature of direct experience, he criticized introspectionism for distorting the facts of experience to fit a preconceived theory. By "introspectionism" Köhler does not mean the gathering of information from an inspection of one's own experience in general; he has criticized the behaviorists for their refusal to accept information so gathered as unscientific. When he attacks introspectionism, Köhler has in mind certain characteristic theories and procedures of the psychologists of his own and the previous generation who relied on introspection. Philosophers and psychologists long believed, under the influence of geometrical optics, that, for example, a round penny must appear elliptical in most positions or that a white surface under a very low degree of illumination must appear gray, and a darker gray than a black surface under a very high degree of illumination. Experimentation has shown, however, that a "naive" observer tends to describe the penny as round no matter what shape strikes the retina and the white surface as white in almost any circumstances. The naive observer, it was held, could not be seeing what he claimed to be seeing. Introspectionists devised elaborate techniques by which a "trained" observer could be made to claim to see what by the laws of optics he should be seeing. In essence, these techniques consisted in excluding from the visual field of the observer all of the surroundings of the object to be observed. In this way, the introspectionists claimed, all the factors of learning are excluded and the object is seen as it "really" appears, before the process of education has distorted our pristine perceptions.

Köhler rightly points out that by employing this technique of exclusion in the interests of a theory, all other factors that might explain why the round penny looks round have been barred. The Gestalt theory offers an alternative explanation of this fact that does not involve the notion of an elaborate hoax played upon the naive observer, an explanation that cannot even be tested by the exclusionary techniques of introspectionism. The defects of introspectionism were further evidenced, Köhler claims, by the fact that introspective psychology had degenerated into an investigation of minute and trivial facts of interest only to specialists.

ASSOCIATIONISM AND ATOMISM. Köhler criticized both the introspectionists and the behaviorists for their psychological atomism or, as he also called it, their mosaic theory. Closely related to psychological atomism is the theory of associationism, which Köhler likewise regarded as inadequate. Psychological atomism is the view that what we perceive is a mosaic of bits and pieces, each independent and essentially unconnected with any other. The parts of the visual and other sensory fields thus lack any sort of relatedness. Yet we do recognize this brown patch and that white patch as belonging together and both as being parts of a dog, rather than one belonging with the ground underneath the dog and the other to the wall behind the dog.

Psychological atomism, according to Köhler, is a theory about the nature of the objects of perception. The theory of association is a theory as to how the experience of order arises out of the unordered psychological atoms postulated by psychological atomism. I have seen white patches associated with dogs in the past, and thus I come to expect that when I see a white patch of a particular kind in the future, it will belong to a dog.

Köhler's answer to psychological atomism is that we do not experience the parts of the visual field, for example, as separate from and unrelated to one another, but that we experience relationships among its parts. Certain wholes separate themselves from other parts of the field, and these wholes are composed of parts related to each other by means of the Wertheimer factors mentioned earlier. If we are in fact led to see things as belonging together by the very structure of experience, then the theory of association is unnecessary. Köhler went on to show that it is also inadequate, in that it cannot fully explain all that it was intended to explain.

Many of Köhler's criticisms of atomism and associationism as psychological theories are justified. But he apparently thought that in arguing against psychological atomism he was also arguing against any epistemological atomism as well. Part of his theory of isomorphism is the

claim that the world as experienced contains experienced relationships among its constituents and that the observer does not add this structure to the world. But here, as earlier, conceptual matters are involved: It is not only a matter of experienced relationships but also of learning what it is to experience a relationship. We must learn the established criteria of what is to count as a relationship before we can know that what we are experiencing is a relationship.

Köhler also believed that the theory of associationism led to a hidden limitation in methods of investigation. According to the associationist, he holds, organization arises out of previous association, whereas, in his view, association depends on previous organization. Sensory gestalten, melodies, and meaningful sentences are organized wholes, and their parts are readily associated. Totally unrelated visual or auditory objects or nonsense syllables, on the other hand, have first to be organized into some kind of order before they can be recognized or be later remembered as having been associated. Köhler does not deny the facts of association but, rather, that association is a fundamental explanatory category. If it were recognized that order is more easily found than made, then it would be seen that organization should play a role in the design of experiments. As it is, far too many experiments fail. For instance, in experiments designed to test an animal's intelligence the apparatus may be too complex for the animal to grasp the relations of the parts and thus be beyond his capacity, whereas by a slight revision the apparatus could serve adequately in carrying out the experiments.

PHILOSOPHICAL PROBLEMS

EPISTEMOLOGY. Köhler's epistemological views are difficult to organize and apparently are not altogether consistent. Probably the most careful and accurate presentation of his views is found in *The Place of Value in a World of Fact.* His theory is, as he claims, a form of epistemological dualism, here couched in the form of a refutation of both phenomenalism and the new realism and aimed at showing that the body-mind problem is a pseudo problem. Köhler's theory, both in content and in terminology, is strikingly similar to that developed by Bertrand Russell in *The Analysis of Matter* and *The Outline of Philosophy.*

The body-mind problem, Köhler claims, concerns the location of percepts. Physiology tells us that they are in our interior, in our brains, yet they appear to be outside ourselves. The resolution is that percepts are inside our bodies in one sense and outside our bodies in quite a different sense. We should distinguish between the body as a physical organism and the body as a percept. Percepts depend on processes within the physical organism; without such processes they would not take place. They appear as located outside the body, which is itself a percept. This perceptual body has a definite place in perceptual space, and other percepts have a definite relation to it within perceptual space. There is no more need to wonder why a perceptual dog appears outside of my perceptual body than to wonder why it appears outside of a perceptual house. Relationships in perceptual space say nothing about the location of percepts in physical space.

In some way what Köhler was saying has been recognized at least since Immanuel Kant's distinction between phenomena and noumena, and Köhler's position seems open to much the same objections as Kant's. What is needed is an account of the relationships between physical space and perceptual space, or between physical object and percept, and this is not what Köhler has given. In physical space percepts are inside the observer's body; in perceptual space they are outside. Here is a radical disparity between spatial relations in the phenomenal and the transphenomenal realms. But Köhler wants to hold that relationships in the phenomenal and the physical worlds are isomorphic. The phenomenal house is between two phenomenal trees; the physical house is likewise between two physical trees. Phenomenal relationships are thus supposed to give us knowledge of physical relationships. And our knowledge of phenomenal relations is the only basis for any knowledge we may have of physical relations. But how do we get from percepts in the physical world to physical objects? And how can we avoid solipsism? Köhler claims that two scientists do not observe the same galvanometer. It is self-evident for him that neither can observe the other's phenomenal world. But physically the percept of each is different, for each is in his own brain. Köhler has not shown how we get from the two percepts to a common physical object.

That Kant spoke of things-in-themselves and Köhler of a physical world, or of nature, should not mask the fundamental similarities of their views. Despite Köhler's belief that the phenomenal world itself gives evidence of a nonphenomenal world, his physical world stands in exactly the same position as Kant's things-in-themselves. They are both unknowable.

CAUSATION. With his emphasis on experienced relationships between the parts of perceived entities, it is not surprising that Köhler denies David Hume's claim that we do not experience causal relations. Causation is only a

special case of a general characteristic of experienced phenomena that Köhler terms *requiredness,* other cases of which are discussed in the section on value. In any of various ways one experience "demands" another for its completion. What Köhler calls insight is the coming to see what is demanded, what is needed to complete a set of factors. Men, and animals to a more limited degree, can have insight into, among other things, what caused a particular event or what will be the probable outcome of a particular line of action. The insight is the experiencing of a causal relation between cause and effect. Köhler concedes that the Humean theory of regular sequence accounts for our practice in various situations of subjecting causal theories to experimental testing after they have occurred to us, but it cannot by itself account for our first recognition of a cause.

Köhler has been criticized by defenders of the regularity theory for confusing psychological issues with logical ones. It may well be the case that in human (as well as in purely physical) situations we frequently arrive at the true answer to a causal problem without any elaborate examination of classes of sequences. From this, however, it does not follow that causation is a "simple" relation like, for example, coexistence that can be given in a single experience. Granting that I may truly judge that A_1 is the cause of B_1 without having performed elaborate controlled experiments, Hume's regularity theory has nevertheless been vindicated as an analysis of the concept of causation if I am prepared to admit that A_1 was not really the cause of B_1 were I to discover that other instances of *A* are or were not followed by instances of *B.*

VALUE. Köhler's epistemological views are developed most fully in *The Place of Value in a World of Fact.* This volume is a contribution to the discussion of axiology that played such a prominent role in American philosophy during the 1920s and 1930s. The argument of the work is long, digressive, and difficult to summarize. The views on isomorphism and on epistemology mentioned above form an integral part of the argument. At the cost of oversimplifying Köhler's views to the point of distortion, it can be said that he holds that we can have direct perceptual knowledge of value. Value is an objective fact of the phenomenal, and hence also of the physical, world. Both phenomenal gestalten and physical gestalten spontaneously change in a certain direction. Melodies and visual shapes require completion in certain ways. Very often when we are attempting to remember something, the context in our mind shows us not only the sort of thing we seek to remember but also whether we are getting close to remembering it. Whatever the proper interpretation of these phenomena may be, Köhler believes that they all demonstrate the factor that he terms *requiredness* and that in the case of memory, the requiredness is a characteristic of something outside the present phenomenal situation. Valuation, an assessment of what ought to be, is not a unique phenomenon but another special case of the recognition of requiredness. Köhler does not directly undertake an analysis of valuation but only of requiredness in general. He hoped that his analysis would be of use to philosophers in their own analyses of ethical and aesthetic requiredness.

MECHANISM AND VITALISM. Toward the end of *The Place of Value in a World of Fact,* Köhler returns to two topics that had engaged him earlier, the dispute between mechanism and vitalism and the question of the precise metaphysical classification of his own theory. In the first case, as in many other situations, Köhler argues that the apparent alternatives are not exhaustive. Mechanists, in their treatment of living processes, take the same shortsighted view that they take of the nature of physical processes mentioned earlier. Mechanical systems are not the only kind of physical systems; there are also the dynamically self-regulating systems. The premise that man must be a machine because physics finds only mechanical systems in the world is thus undermined. On the other hand, one does not have to hold to vitalism just because men are obviously different from machines. Living organisms, including man, can quite easily be physical systems without being machines. And in fact, Köhler held, living organisms can be explained quite satisfactorily as dynamically self-regulating systems without postulating some mysterious nonphysical vital force.

BODY-MIND PROBLEM. Köhler seems to advocate an epistemological dualism. He was not, however, a dualist in the sense in which the term is used in connection with the body-mind problem. Other psychologists have labeled him a physicalist, and he did not totally reject the terms *materialist* and *monist* as used to describe his metaphysical views. He found the label "materialist" misleading because he accepted the modern physicists' account of the world, and this account is very different from any traditional account of matter as composed of solid impenetrable particles. He believed that eventually it may be shown that phenomenal colors are identical with chemical states in the brain and that in this way the physicists' account of reality would be complete. In this sense he did not reject the possibility that monism is true, but in the meantime phenomenal qualities appear so different from any physical correlates that the possibility of the false-

hood of monism likewise cannot be ruled out. There is some similarity between Köhler's views on this subject and the theory of J. J. C. Smart and U. T. Place that sensations and brain processes are identical. Like Smart and Place, Köhler argues that the undeniable phenomenological differences between colors and chemical states of the brain do not rule out the possibility that, in an important sense, they may nevertheless be identical. However, unlike Smart and Place, Köhler does not claim that such an identity has in fact been established.

See also Atomism; Behaviorism; Causation: Philosophy of Science; Epistemology; Galileo Galilei; Gestalt Theory; Hume, David; Husserl, Edmund; Introspection; Kant, Immanuel; Koffka, Kurt; Mind-Body Problem; Planck, Max; Realism; Psychology; Russell, Bertrand Arthur William; Smart, John Jamieson Carswell; Stumpf, Karl; Value and Valuation; Vitalism; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY KÖHLER

"Über unbemerkte Empfindungen und Urteilstäuschungen." *Zeitschrift für Psychologie* 63 (1913): 51–80.

Intelligenzprüfungen an Anthropoiden. Berlin, 1917. Revised 2nd ed. as *Intelligenzprüfungen an Menschenaffen.* Berlin: Springer, 1921. Translated from the 2nd ed. by Ella Winter as *The Mentality of Apes.* London: Routledge and Kegan Paul, 1925.

Die physischen Gestalten in Ruhe und im stationären Zustand. Brunswick, Germany, 1920; 2nd ed., Erlangen: Philosophische Akademie, 1924.

"An Aspect of Gestalt Psychology." In *Psychologies of 1925,* edited by Carl Murchison. Worcester, MA: Clark University Press, 1925. Ch. 8.

"Gestaltprobleme und Anfänge einer Gestalttheorie." *Jahresbericht über das gesamte Physiologie und experimental Pharmakologie* 3 (1925): 512–539.

"Komplextheorie und Gestalttheorie." *Psychologische Forschung* 6 (1925): 358–416.

Gestalt Psychology. New York: Liveright, 1929. German edition published as *Psychologische Probleme,* Berlin, 1933.

"Some Tasks of Gestalt Psychology." In *Psychologies of 1930,* edited by Carl Murchison. Worcester, MA: Clark University Press, 1930. Ch. 8.

"Zur Psychophysik des Vergleichs und des Raumes." *Psychologische Forschung* 18 (1933): 343–360.

The Place of Value in a World of Fact. New York: Liveright, 1938.

Dynamics in Psychology. New York: Liveright, 1940.

"On the Nature of Associations." *Proceedings of the American Philosophical Society* 84 (1941): 489–502.

"Figural After-Effects: An Investigation of Visual Processes." *Proceedings of the American Philosophical Society* 88 (1944): 269–357.

"Gestalt Psychology Today." *American Psychologist* 14 (1959): 727–734.

The Task of Gestalt Psychology. Princeton, NJ: Princeton University Press, 1969.

The Selected Papers of Wolfgang Koehler. Edited by Mary Henle. New York: Liveright, 1971.

In Anthologies

Ellis, W. D. *A Source Book of Gestalt Psychology.* London: Kegan Paul, 1938. Contains translation of portions of *Die physikalischen Gestalten in Ruhe und im stationären Zustand.*

Henle, Mary. *Documents of Gestalt Psychology.* Berkeley: University of California Press, 1961. Contains Köhler's presidential address to the American Psychological Association (1959) and other later papers.

WORKS ON KÖHLER

Arnheim, R. "The Two Faces of Gestalt Psychology." *American Psychologist* 41 (1986): 820–824.

Ash, Mitchell G. *Gestalt Psychology in German Culture 1890–1967: Holism and the Quest for Objectivity.* Cambridge, U.K.: Cambridge University Press, 1995.

Ayer, A. J. *The Foundations of Empirical Knowledge,* 113–135. London, 1947.

Ayer, A. J. *Language, Truth and Logic,* 2nd ed., 56–59. New York, 1946.

Boring, E. G. "The Gestalt Psychology and the Gestalt Movement." *American Journal of Psychology* 42 (1930): 308–315.

Driesch, Hans. "Physische Gestalten und Organismen." *Annalen der Philosophie* 5 (1925).

Grelling, Kurt, and Paul Oppenheim. "Der Gestaltbegriff in Lichte der neuen Logik." *Erkenntnis* 7 (1938): 211–225.

Hamlyn, D. W. "Psychological Explanation and the Gestalt Hypothesis." *Mind* 60 (1951).

Hamlyn, D. W. *The Psychology of Perception.* London: Routledge and Paul, 1957.

Hartmann, George W. *Gestalt Psychology: A Survey of Facts and Principles.* New York: Ronald Press, 1935.

Henle, Mary. "One Man against the Nazis: Wolfgang Köhler." *American Psychologist* 33 (1978): 939–944.

Hobart, R. E. "Hume without Scepticism." *Mind* 39 (1930).

Katz, David. *Gestalt Psychology: Its Nature and Significance.* Translated by Robert Tyson. New York: Ronald Press, 1950.

Ley, Ronald. *A Whisper of Espionage: Wolfgang Kohler and the Apes of Tenerife.* Garden City Park, NY: Avery, 1990.

Müller, G. E. *Komplextheorie und Gestalttheorie.* Göttingen: Vandenhoeck and Ruprecht, 1923.

Nagel, Ernest. *The Structure of Science,* 380–397. New York: Harcourt Brace, 1961.

Petermann, Bruno. *Die Wertheimer-Koffka-Köhlersche Gestalttheorie.* Leipzig: Barth, 1929. Translated by Meyer Fortes as *The Gestalt Theory and the Problem of Configuration.* London: K. Paul, Trench, Trubner, 1932.

Reiser, O. L. "The Logic of Gestalt Psychology." *Psychological Review* 38 (1931): 359–368.

Rignano, Eugenio. "The Psychological Theory of Form." *Psychological Review* 35 (1928): 118–135.

Sherrill, R., Jr. "Natural Wholes: Wolfgang Kohler and Gestalt Theory." In *Portraits of Pioneers in Psychology,* Vol. 1, edited by G. A. Kimble, M. Wertheimer, and C. White, 257–273.

Washington, DC: American Psychological Association; Hillsdale, NJ: L. Erlbaum Associates, 1991.

Smith, Barry, ed. *Foundations of Gestalt Theory.* Munich: Philosophia, 1988.

Philip W. Cummings (1967)
Bibliography updated by Alyssa Ney (2005)

KOREAN PHILOSOPHY

Scholars hold diverse opinions on the identity and origin of Korean philosophy. Although some trace the origin back to antiquity when the mythical figure Dangun supposedly founded the country in 2333 BCE, there is little historical evidence to support it. It is more plausible to estimate that philosophy began in Korea during the Three Kingdom era (second century CE) when people unfettered themselves from myths, legends, and shamanist beliefs of the tribes, and began to think in more general and philosophical terms. During this period Buddhism, a systematic and conceptually advanced religion, was introduced into the Three Kingdoms (Shilla, Baekje, and Koguryo), all of which embraced it to serve as a social and spiritual foundation for a trans-tribal ethical system. After its introduction, Korean Buddhism went through diverse phases of changes and developments, sometimes as a result of adaptations to changing social and political environments and sometimes as a result of theoretical debates. Neo-Confucianism and Western thought that were later introduced to Korea underwent similar turns and twists.

Korean philosophy, largely formed on the basis of external thought and influences, is notable not for the uniqueness of thoughts per se, but for the special manner in which it internalized the established and widely disseminated thought systems of Asia and the West and developed them into identifiably Korean forms. Korea's geographical and historical circumstances exposed the country to sudden and often torrential influxes of mature and powerful foreign culture and thought systems. Thus, the development of Korean philosophy has consisted in selecting an appropriate trend of thought carefully and reinterpreting it to meet the challenges of the society.

Because Korean philosophy had to concentrate on the selected trend, its characteristic is fundamentalist in that there was a tendency to select a specific trend or interpretation and adhere to it as the only source of truth to the exclusion of other trends. Because Korean philosophy attempted to synthesize diverse thought within the selected trend in order to meet the challenges of the society, the ability to weave divergent thoughts into a coherent whole was crucial. Even today when Western philosophy prevails, the two characteristics of fundamentalism and integrationism are still valid as a description of Korean Philosophy.

THE BEGINNING OF PHILOSOPHICAL THINKING—THE INTRODUCTION OF BUDDHISM AND THE DEVELOPMENT OF KOREAN BUDDHIST PHILOSOPHY

As the Three Kingdoms expanded to constitute sovereign states, politics began to separate from religion. Tribal federations were gradually transformed into monarchies, and the mythologies of clans and the associated religious rituals that had so far dominated the spiritual world of people were no longer adequate to serve as the basis of a state. This created a need for a unified belief system that would reconcile diverse native religious thought and practice, and provide a political rationale for the monarch-centered sovereign state. Such an ideology was also needed to counteract the aristocrats who resented the increasing concentration of political power in the monarch. The introduction of Buddhism from China at this time filled just this need, and it was welcomed by the royal authority.

From its inception Buddhism was allied with the royal authority, so it was advocated not only as a higher, more sophisticated religion, but also as a theoretical ground for strengthening the sovereignty. For example, the Buddhist notion of cause and effect, together with its karmic associations, were helpful in promoting the belief that their king was not a ruler arbitrarily chosen by Heaven, and that his status was a necessary consequence of the good deeds done in his past lives. Buddhist doctrines were also invoked to justify the authority and legitimacy of the royal rule. For that reason the Three Kingdoms endorsed at first the School of Precepts (the Vinaya School), which stressed the importance of rule abidance, in order to solidify the ethical norms and regulations of the newly established nations. As the number of Buddhist monks increased, their mission extended beyond the performance of ceremonies and rituals; they started to study the Buddhist doctrines and texts from a scholarly point of view.

Koguryo, in the north of the Korean peninsula, adopted a branch of Buddhism that interpreted Buddhism in terms of the Daoist concept of nothingness, a concept that was familiar in the local shamanist beliefs. It was succeeded by the Three Treatise School (the Madhyamika School), which upheld the doctrine of emptiness

(*Sunyata*) with the motto "What can be said, cannot be real." Whereas Buddhism, with an emphasis on nothingness or emptiness, was popular in Koguryo, a different perspective on Buddhism was embraced in Shilla. It was called the Consciousness-Only School (the Yogacara School). As the name suggests, their main claim was that the external world is nothing more than the objectification of inner cognitive activities and that only consciousness and cognition exist. It was popularized by Shilla monk Woncheuk (613–696), who studied and practiced his theory in China. His theory was influential not only within Shilla, but also in Tibet.

After the seventh century, more monks returned after studying abroad and brought with them Buddhist doctrines of numerous schools, adding diversity to the early Korean Buddhism. It also improved the quality of Buddhist studies, but at the same time it caused deep confusion. All the teachings were from one Buddha. So how could one make sense of all these diverse interpretations, some of them in conflict with others? The perplexity was especially acute in Shilla, which had an alliance with Tang China and sent many monks there to study Buddhist doctrines. This created fierce debates and disputes among the monks, each group arguing that what it had learned was the exclusive truth. Through this process, conflicting theoretical stances adjusted themselves to accommodate each other, which led to the unique characteristic of Korean Buddhism called *integrationism*.

Shilla monk Wonhyo (617–686) was the first Buddhist scholar who established his own unique theory. He meticulously analyzed three core concepts of Buddhism—mind (*citra*), enlightenment (*bodhi*), and ignorance (*avidyā*)—and attempted to illuminate their mutual relationship. According to Wonhyo, Buddha's mind and people's minds are one and the same and people born with the mind of Buddha lost track of the true facet of human existence because they are blinded by ignorance (i.e., self-centeredness and greed). Thus, being in the state of Buddha's mind (enlightenment) is nothing above and beyond being in the state of freedom from ignorance and thus returning to the original state of the human mind. On this basis, he argued that the Three Treatise School's method that tried to reach Buddha's mind by removing ignorance and the Consciousness-Only School's converse method of removing ignorance by reaching Buddha's mind were just two different paths to the same goal. This illustrates the way in which Wonhyo attempted to harmonize doctrinal differences among diverse schools. Because of Wonhyo's influence, the Buddhist schools in Korea henceforth sought in a single-minded way to reach an all-encompassing interpretation of Buddhism.

Whereas Wonhyo laid the philosophical foundation of Korean Buddhism, Uisang (625–702) focused his work on unifying numerous Buddhist schools active in all parts of the nation. Upon his return from Tang China shortly after Shilla absorbed and consolidated the other two kingdoms into the United Shilla (676), Uisang reorganized the Buddhist temples with divergent doctrinal allegiances by embracing the Flower Garland School (the Avatamsaka School). On the basis of the claim that particulars and universals, many and the one, were all different aspects of dharma (the principle, law, or a universal norm that orders both the natural world and human conduct), he advocated the holistic view that all things in the universe, causally interconnected under dharma, represented the same supreme mind. This holistic doctrine of the Flower Garland School provided a spiritual background for the harmony that must exist between individuals and the state, and between individuals and the universe. Thus it helped support the political consolidation of the Unified Shilla dynasty.

THE ACCEPTANCE OF ZEN BUDDHISM AND ITS DEVELOPMENT

In the eighth century the Unified Shilla made great strides in doctrinal studies, particularly in the areas of the Flower Garland and Consciousness-Only Schools. During the latter half of the eighth century, however, the role of king shrank to that of a protector of his own clan, and powerful clans in the provinces rose to supersede the royal authority. Accordingly, the Flower Garland School that provided the spiritual basis for unification was succeeded by Zen Buddhism backed by regional aristocrats. Zen Buddhism emphasized that enlightenment was attained not through laborious doctrinal studies, but through discovering the Buddha mind within oneself. Even though Korean Zen Buddhism prospered as diverse branches of Chinese Zen Buddhism were introduced, the philosophical message was no different from what had been taught by Flower Garland School or Wonhyo—that ignorance is the beginning of enlightment and that everything is dependent on one's mind. It should be noted, however, that practice-oriented characteristics of Zen Buddhism paved the way for Korean Buddhism to become a popular religion without being trapped in theoretical intricacies.

In 936 the Koryo dynasty emerged, leaving behind the chaotic ruins of the Shilla dynasty. While the Koryo dynasty was developing into a state, it exploited Confucianism for practical purposes. Confucianism was intro-

duced into Korea around the second century BCE and Koreans were familiar with its major teachings for more than 1,000 years. Although Confucian education was gradually strengthened mainly for the purpose of building a bureaucratic system, Koryo Confucianism at this time had yet to reach a level of philosophical significance. Spiritually, the primary concern of Koryo was integrating diversified schools of thought, and it was still Buddhism that undertook the role. Thus, one can witness the strong integrationist tendency in Buddhism throughout the Koryo dynasty.

Chinul (1158–1210) invigorated and established Zen Buddhism as a strong tradition in Koryo by providing it a firm philosophical basis. Thinking that Zen Buddhism of his time had dwindled in popularity mainly because of its inherent subjectivity and excessive aversion to doctrinal studies, he argued that both the doctrinal component and the meditative component must be incorporated into a correct version of Buddhism. This led to the creation of his own unique program of "sudden awakening and gradual cultivation." According to this program, one can clear oneself of secular concerns and arrive at Buddha's mind only if one comes to be enlightened by meditative insights and at the same time carries out self-cultivation to verify whether what one has understood by enlightenment corresponds to the general truth of Buddhism. This unique theory within the meditation camp became one of the most representative views of Korean Buddhism, influential up to the early twenty-first century.

After Chinul, there emerged a variety of Buddhist philosophies such as purely meditative Buddhism, a Confucian Buddhism, and so on. Still the unique characteristic of Korean Buddhism lies in the fact that it has constantly sought a synthesis of two major traditions of Buddhism, doctrinal tradition and Zen tradition, and it is often argued that Korean Buddhism has been most successful at that. With the formation of the Chosun dynasty, however, Buddhism came to be regarded as something to be overcome and was by and large excluded from ideological pursuits.

THE ACCEPTANCE OF NEO-CONFUCIANISM

Although it is hard to trace exactly when Confucianism was first introduced to Korea, it is estimated that its introduction accompanied the import of the Chinese writing system roughly around the second century BCE. Koreans began to accept Confucianism as the Three Kingdoms transformed themselves into ancient states and this created a need for Confucian bureaucrats who were versed in the Chinese writing system well enough to fulfill practical purposes of composing diplomatic documents. Each of the Three Kingdoms had Confucian educational institutions, which produced Confucian scholars and students. From the fact that Confucian virtues such as loyalty and filial piety were prized in the Three Kingdoms, it can be inferred that Confucianism was held in high esteem, even though the scholarship was not up to the level of philosophical analysis.

Confucianism during the Koryo period, as in the Shilla period, was chiefly used as a useful political and practical complement to Buddhism. After the eleventh century, however, as the sovereignty and its administrative structure became stabilized, Confucianism began to distinguish itself from Buddhism. Confucianism that had been only an object of a practical interest began to be the object of serious theoretical research as well. Koryo's Confucian scholars, represented by Choi Chung (984–1068) and his twelve disciples, considerably advanced the level of Confucian studies as they participated in public administration from the time of King Seong (who ruled from 981 to 997) to King Mun (1046–1083). The private Confucian educational institution Choi founded taught major Confucian Classics. Still, because the program of study was largely oriented toward preparing students for national examinations, it seems that more time was spent on literary exercises than on philosophical investigations.

The later Koryo period was an important time for Confucianism in Korea: This was when Korean Confucian scholars started distancing themselves from Buddhism. Scholars returning from Yuan China brought home with them the Confucianism that was already Yuan's political ideology, and this transformed Koryo's Confucianism in a novel way. The Neo-Confucian master Zhu Xi's writings were introduced in 1289 and numerous Confucian scholars from then on gradually extended the understanding of Confucianism and Neo-Confucianism. A truly novel phenomenon occurring was that these scholars began to mount an attack on Buddhism with philosophical arguments. Yi Saek (1328–1396), one of the last scholars to return from Yuan China, exerted an extensive influence on later Korean Confucians. Even though his own understanding of Neo-Confucianism remained still at a comparatively naive stage in that it simple-mindedly identified Confucian benevolence with Buddhist compassion, and Confucian repose with Buddhist calmness, Yi Saek produced prominent and influential disciples.

They were trained at the national Confucian educational institution, called Sungkeunguan, which was founded by the government in 1289. They became major figures during the transition period from Koryo to Chosun, which succeeded the Koryo dynasty in 1392. With philosophical explanations of why Buddhism was fundamentally a heresy, they decisively broke with the previous generations of scholars who were largely tolerant of Buddhism. They also played a crucial role in constructing, for the new state, an ideological framework based on Confucianism.

The Chosun dynasty, which replaced the Buddhist Koryo dynasty, adopted Confucian ideology, custom, and order as the political and social foundation of the new state. Those who framed the political philosophical framework for the new dynasty were a group of scholars led by Chung Dojeon. Chung had a leading role in laying the foundation of Chosun's Neo-Confucianism and enabled Confucian ideology to prevail. Because his interpretation of Neo-Confucianism was constructed with a deliberate intention to buttress the new society with a philosophical basis, his philosophy went beyond the personal realm of self-cultivation and moral improvement.

What Chung stressed the most as he propounded Neo-Confucianism was the criticism of Buddhism. He methodically compared the Buddhist worldview with that of Neo-Confucianism, arguing that whereas the basis of the Buddhist worldview was nihilism based on emptiness (*Sunyata*), a robust realism based on *li* and *qi* was the foundation of Neo-Confucianism. *Li* and *qi* are the two most important concepts in Neo-Confucianism. In Zhu Xi's philosophical system, *li*, which is similar to the Platonic idea or the Aristotelian notion of form, is an abstract being. *Li*, like the Buddhist dharma, is often appealed to in the explanation of universal truths governing the natural world and human conduct. *Qi*, on the other hand, corresponds roughly to matter in Western philosophy and it is often invoked to explain the changes in spatiotemporal objects including human bodies and minds. However, *qi* differs from matter as conceived in the West in two important respects. First, Neo-Confucianism locates mind in the domain of *qi*, whereas the Western tradition has tended to regard mind to be distinct from matter. Second, *qi* was construed to be animate, whereas matter is usually construed to be inert and inanimate.

Chung, following the Neo-Confucian tradition, explained the generation and decay of man and nature in terms of *qi* and, on its basis, attacked the Buddhist theory that argued for the illusory nature of the world, the unreality of things, and the transmigration and eternity of the soul. He also attacked the Buddhist doctrine of Karma by claiming that people's differences were not because of what they had done in the past, but because of the *qi* that each person possessed from birth. Chung distinguished Neo-Confucianism from Buddhism in the domain of morality as well. He contended that although the Buddhist notion of compassion had some similarities with the Confucian notion of benevolence, they fundamentally differed in that compassion required treating all beings with indiscriminate equality, whereas benevolence allowed for unequal treatments based on the type of relationship between the benefactor and the recipient. Confucian benevolence, thus construed, served as the fundamental value to sustain the order of the new hierarchical society. Chung's denunciation of Buddhism as a heresy successfully derailed the attempts to revive Buddhism during the early Chosun period and paved the way for other scholars of the upcoming generations to develop and systematize Korean Neo-Confucianism.

The groundwork laid by Chung, however, did not lead immediately to fruitful Confucian research. During the first years of Chosun, a period marked by intense conflicts among the major political factions, Neo-Confucianism as a national ideology lost its initial momentum and was bogged down in exegetical studies. It was during the years of King Sung (1457–1494) that Neo-Confucian scholars returned to hold positions of great influence in the government. Neo-Confucianism began to serve as a practical guide to governance, going beyond its role as a mere ideology. Cho Kwangjo (1482–1519) was the scholar who was most influential in this transition. He claimed that the ruler's moral cultivation was especially important because his moral commitments would exert great influence on the whole nation. Cho urged the view that an ideal Confucian state could be realized through the internalization of Confucian moral values on a national scale and he subsequently led a movement to actualize the view. Views like these were commonly held by the Confucian literati of the time, and it led Neo-Confucians to delve into the nature of human mind and explore the ground and the method of moral practice.

THE THEORETICAL DEVELOPMENT OF NEO-CONFUCIANISM

Although Neo-Confucianism during the early Chosun period put more emphasis on the practical side, the theoretical side was not completely ignored. For example, the concept of *qi* was exploited to explicate problems such as man and nature, life and death, and the existence of

souls and spirits. The scholar who added depth to the philosophy of *qi* was Seo Kyeongdeok (1489–1546). Seo, classified as a *qi*-philosopher during the early to middle Chosun period, constructed a highly complex and sophisticated theory of cosmology and human nature on the basis of *qi*.

Drawing on the views of Chinese *qi* philosophers during the Song dynasty, in particular Zhang Hengqi and Shao Kangjie, Seo attempted to explain the macroscopic movements and changes in nature in terms of the diverse phases of *qi* and transitions between them. For example, he discriminated between *qi* as a root of everything (pre-celestial *qi*) and *qi* as a changing phenomenon (post-celestial *qi*). Pre-celestial *qi* is the ultimate basis of existing entities, whose movement and change determine variance in post-celestial phenomena. The phenomenal world, which is generated through *qi*'s movements and changes, disappears as *qi* disperses, yet the dispersed *qi* returns again to the pre-celestial realm, which in turn becomes a causal basis of the regeneration of another phenomenal world. Seo associated this cosmology with the principle of Great Change as manifested in the Book of Changes, and applied his theory to the problems of life and death, and even to the question of life after death. His theory of *qi* enabled people to overcome the Daoist concept of nothingness and the Buddhist notion of eternity of the soul; most importantly, it helped the Neo-Confucianism of the Chosun dynasty to gain a unique perspective on man and nature.

The philosophers who completed the framework of Neo-Confucian moral philosophy were Yi Hwang (1501–1570) and Yi I (1536–1584). Yi Hwang, better known by his pen name Toegye, researched in depth the Chinese Neo-Confucian master Zhu Xi, whom he regarded as the ultimate source and authority for Neo-Confucianism. In contrast to Seo before him, he argued that *li* was the ultimate and essential being that determined the movement of *qi*. What particularly concerned Toegye, however, was not the ontology of *li* and *qi* per se, but their roles in grounding morality. He believed that if *li* did not act upon the external world, there would be no ontological ground for morality. In other words, he thought that moral intuition or wisdom would be useless if all human emotions are vulnerable to physical intemperance and overindulgence. It seemed obvious to Toegye, however, that humans had an intellectual control over the mind. From this, he concluded that there must be a domain of emotions that are distinctively moral, and that these must be distinguished from mundane nonmoral emotions. He went on to construct the unique view that everyday nonmoral emotions were manifestations of *qi*, whereas moral emotions were manifestations of *li*. In placing morality within the domain of emotions, Toegye put a greater emphasis on the cultivation of the emotions rather than on purely rational and intellectual training.

Another philosopher who elevated the Chosun dynasty's Neo-Confucianism to another level of sophistication was Yi I (1536–1584), better known by his pen name Yulgok. While revering Toegye's scholarship, he thought that Toegye's dualistic interpretation of Zhu Xi's philosophy had a fundamental problem. Placing a higher value on the aforementioned metaphysical system devised by Seo, Yulgok claimed that although *li* and *qi* were differentiated conceptually, they were not two independent beings. Applying this view to morality, Yulgok maintained that there was no separate source or domain of moral emotions; everyday emotions that conformed to the moral standard were themselves moral emotions. All the emotions including moral emotions were manifestations of *qi*, but they were regulated by *li*. A moral action was not a natural emanation from a separate moral emotion, but the outcome of the recognition of the universal norms and a personal decision to make that recognition bear on the mundane emotions. Because Yulgok considered reason, rather than emotion, to play a central role in living a moral life, he concluded that the enhancement of our rational capacity for right judgments should be emphasized over emotional enrichment.

Weighing between emotion and reason, and between *qi* and *li*, the philosophies of Toegye and Yulgok manifested subtle but significant differences in all respects, leading to two lineages of Neo-Confucianism during the Chosun period. One was *li*-centered and the other *qi*-centered. As the two schools contended for the title of Neo-Confucian orthodoxy, the Chosun dynasty's Neo-Confucianism became increasingly more dogmatic and doctrinaire, leading scholars to the rigid position that all social and individual conduct should conform to the Confucian code of behavior. Leaving behind the metaphysical basis of a moral mind, the debate now moved to another issue over how to apply abstract morality to the real world. Thus, the theory of rites and rituals came to replace the theory of mind, and formed the mainstream philosophy of the seventeenth-century Chosun dynasty.

As the Chosun dynasty's Neo-Confucianism became increasingly more doctrinaire and ritualistic, the chasm between theory and reality, and between philosophy and social development, widened. Scholars, convinced that a blind adherence to Zhu Xi's texts had led them into a dead end, began to search for a breakthrough outside Zhu

Xi. Two trends are notable as consequences of this movement; one was the acceptance of the Chinese Yangming philosophy that recognized the significance of the individual will and freedom. The other was the emergence of exegetical studies that focused on a positivistic interpretation of Confucian Classics free from political ideologies. Scholars involved in these studies hoped to overcome Zhu Xi's philosophy by an appeal to a superior authority (i.e., revered ancient Confucian Classics). The rejection of Zhu Xi's philosophy was significant and it exerted a strong influence on later philosophers, particularly on those belonging to the Practical Study School.

Meanwhile, the scholars from the Yulgok's lineage went on to articulate their philosophical system. In their attempt to refine Yulgok's philosophy, a discordance within his system was discovered, which led to the biggest philosophical debate of the eighteenth century and subsequently caused a split of the school into the *Ho* line and the *Rak* line (*Ho* and *Rak* are names of the regions where their advocates resided). The *Ho-Rak* debate was over the question whether there existed a nature common to both humans and other creatures in the world. The debate that initially started between two scholars gradually widened and came to involve almost all the scholars of the Yulgok school. The debate evolved to cover a wide range of topics such as the relationship between mind and nature, the distinction between the sage and the commoner, and the sameness or difference between human nature and animal nature. In debating over whether there was a general nature common to all things in nature, they came to address the relationship between *li* and *qi* and consequently it provided an opportunity to rethink the status and meaning of *li*. This in turn gave rise to a wide spectrum of thoughts such as the *qi*-only theory and the *li*-only theory.

THE RISE OF MODERN THOUGHT—THE INTRODUCTION AND RECEPTION OF WESTERN THOUGHT AND PRACTICAL STUDY

As Korea opened its door to Western thought in the eighteenth century, a notable change in the trend of Korean philosophy took place, and this was the emergence of the Practical Study School. From the early eighteenth century on, the inadequacy of Neo-Confucianism as a political ideology became increasingly more evident. In order to go beyond the limit of Neo-Confucianism and to go along with new social environments, a group of scholars turned their attention from morality and self-cultivation to more practical questions such as economy and the land system. This trend came to be called Practical Study. Scholars belonging to this movement tried to attain new philosophical insights by blending traditional Neo-Confucianism with newly introduced Western thought, especially Catholicism and Western sciences.

Yi Ik (1681–1763), deeply impressed by the astronomy and the solar calendar brought to Korea by the Christian missionaries, took an active part in introducing Western thought to Korea. He created an atmosphere that enabled his disciples to play leading roles in spreading and promoting Western thought. On the issue of accepting the Catholic doctrine, however, they diverged into a receptive group called the Accept-West Party and a critical group called the Reject-West Party. The latter criticized the fundamental premises of Catholicism including the theory of anima from a Neo-Confucian perspective on the nature of mind. They claimed that Catholicism and Confucianism differed in fundamental assumptions and could not be harmonized with each other.

The Accept-West Party maintained a more open attitude toward Western thought. Among the more influential members of this group was Chong Yakyong (1762–1836), better known by the pen name Dasan, who constructed a comprehensive and influential theory of the Practical Study School, incorporating Catholic theories in his philosophical system. Through a novel reinterpretation of Confucian Classics, not only did he attempt to recover the practical spirit of early Confucianism, but he also tried to synthesize Confucianism and Catholicism. For example, he argued that God in Christianity and Heaven in ancient Confucianism were one and the same; according to him, Heaven in the Confucian tradition was essentially a subject with volitions, desires, and perceptions, and also an agent who used those faculties to rule the universe. Thus, the Confucian Heaven was not to be explicated in terms of metaphysical and abstract principles such as *li* or yin and yang. According to him, then, the term *high-emperor* as employed by ancient Confucians portrayed the meaning of Heaven in the most adequate way, and Heaven, thus construed, was no different from the Christian God.

Dasan also drew on Christian ideas in his explications of morality. Criticizing the Neo-Confucian view that morality was a part of inherent human nature, he maintained that human nature was so constituted as to follow self-regarding desires and preferences and thus it was fundamentally egotistic and hedonistic. He advocated, on this basis, the Christian idea that moral perfection was possible only through recognizing God's will and acting accordingly. Then he attempted to graft Con-

fucianism onto Christianity by adopting the Confucian theory of cultivation as a way of internalizing God's orders. However, such attempts by Dasan and other Accept-West Party scholars caused, among the mainstream scholars who were still committed to Neo-Confucianism as their philosophical idea, a deep sense of insecurity. This played a part in bringing about an official oppression of the Catholic church later, which started in 1785 and lasted on and off for eighty years.

Unlike Yi Ik's disciples who attempted to overcome the limits of Neo-Confucianism by adopting Catholicism, other mainstream scholars in powerful positions embraced Western sciences to improve their Neo-Confucian system. They were called Study-North Scholars, and Hong Daeyong (1731–1783) and Choi Hangi (1803–1877) were the leading figures. Hong, keenly interested in Western sciences, turned his attention from a value-laden Confucian worldview to a morally neutral, positivistic, and scientific worldview. Believing that human existence was on the same level as the existence of any other natural beings, he attempted to explicate everything in terms of *qi*'s movement. Hong's notion of *qi* was similar to today's concept of matter, more so than that of any other *qi*-scholars. *Qi* was, for Hong, a concept suited to cosmology and useful in explaining natural phenomena; he explained the rotation of the earth, tides, and climatic changes by using the concepts such as shrouding *qi*, flowing *qi*, and great *qi*. Thus, in Hong's theory, the dynamic transformations of *qi* were more salient than the ultimate nature of *qi* itself. The significance of Hong's philosophy of *qi* was that it went beyond the Confucian moralist view of the natural world and gave Korean philosophy a modern naturalistic outlook by combining traditional philosophy with the newly introduced Western sciences.

In the case of Choi Hangi, the influence of Western science is even more evident. In Choi's theory, the traditional concept of *qi* played a critical mediating role in assimilating Western scientific theories into his own system. Choi believed that human conduct and natural phenomena were all manifestations of *qi*, and therefore that both Confucian ethics and Western science could be proven to be truths on the same level. Rejecting the Neo-Confucian perspective on morality, he claimed that ethical norms were based on, and derivable from, laws of nature. His *qi*-centered theory not only encompassed existing Confucianism, Buddhism, and Daoism, but could also be harmonized with Western scientific theories.

Two main characteristics of *qi* in Choi's theory were quantifiability and perceivability. Because everything was a manifestation of *qi* and *qi* was perceivable, one could accumulate knowledge only through empirical investigations. According to Choi, the knowledge thus obtained should be able to reach, through verifications and repeated corrections, a level where the fundamental principles common to humans and the natural world could be discovered and natural phenomena scientifically understood. He was also convinced that *qi* could be quantified by numbers. Because the numerical system could reveal changes of *qi* in an objective and general way, scientific studies such as menology (calendar studies), calculus, and physics could reveal the nature of the world most accurately. He even thought that the movement of *qi* could be proven mathematically. What is especially notable in Choi's theory is that Choi had unfettered himself completely from the value-centered, intuition-dependent philosophy of Neo-Confucianism and paved a way to a modern naturalistic way of thinking.

Dasan's Catholic Confucianism, Hong Daeyong's scientific Neo-Confucianism, and Choi Hangi's empirical epistemology were just a few representative attempts, during the eighteenth and nineteenth centuries, to embrace the newly introduced Western thought within traditional philosophy. They had the potential for launching a vital and original philosophical movement. With the fall of the Chosun dynasty, however, these philosophical endeavors did not lead to the formation of modern Korean philosophy. They remained only as one dead-end strand in the history of Korean philosophy.

MODERN KOREAN PHILOSOPHY

The period from the end of the nineteenth century to the beginning of the twentieth century was a critical turning point for Korea and for Korean philosophy. The Japanese colonialism backed by Western culture and technology began to encroach on Korea. Korea was forced to sign an unequal treaty with Japan in 1876. That provoked other imperialistic countries to coerce similar forms of agreements with Korea. As a result, Korea was defenseless against the tidal influx of Western culture, new languages, and new modes of thinking. Although Korean intellectuals at the time attempted to save Korea from colonization by westernizing Korea itself, it was too little and too late as Korea was annexed by Japan in 1910.

A notable phenomenon that followed was the shift of Korea's and Japan's roles in the transfer of cultures. Traditionally China was the dominant cultural force in the region, and Korea used to import Chinese culture and

incorporate it into its own, and then export the outcome to Japan. By the turn of the century, this pattern of cultural exchanges underwent a dramatic change; the West replaced China and Japan became the conduit of the Western culture to Korea. Even though Korea had earlier contact with Western religion and science, it was only after the Japanese colonization that Korea made its first encounter with Western philosophy. The word *chulhak* was also first introduced to Korea. The word, made up of two Chinese characters, was coined in Japan as a translation of the term *philosophy*, and it is now the standard term for philosophy in the Asian countries in which Chinese characters are used for academic purposes, including China, Korea, and Japan.

That Western philosophy was introduced to Korea through Japanese colonialism, combined with the prevalent picture of Western power and wealth, defined the early perception of Western philosophy in Korea. Philosophy was regarded as something indigenous to the West and completely alien to Korea, having nothing in common with the traditional thought of Korea. In the minds of Korean intellectuals at the turn of the century, the historical dominance of Confucianism was the main reason for Korea's falling behind in the process of modernization. Traditional ways of thinking and Confucianism, in particular, were what had to be overcome, whereas Western culture and philosophy were to be welcomed and assimilated. The introduction of a neologism, *chulhak*, to signify Western philosophy might have reinforced this frame of mind. For example, *Philosophy*, the first academic journal of philosophy published in 1933, contained no article on traditional Korean thought. It took many years to recognize the common features between Western philosophy and Asian thought and to apply the term *chulhak* to both.

Western philosophy was mostly German philosophy. Japan and Germany were allies and the Western philosophy in Japan was for the most part German philosophy. In consequence, Western philosophy introduced to Korea via Japan was also mostly German. Even though Bertrand Russell and John Dewey visited China and Japan respectively in 1910 and 1919 and that these visits aroused the interest of philosophers in Korea, their impact was limited. The dominance of German philosophy in Korea lasted for some time even after Korea's liberation from Japan in 1945, and this continued during the post-World War II years when the influence of German philosophy was diminishing in the rest of the world. Scholars specializing in German philosophy filled the philosophy faculties of the major universities, and they determined the overall shape and course of the profession until philosophers of a new generation began replacing them.

Writings of Korean philosophers in the early twentieth century were oriented toward practice. Korean philosophers, like any other Korean intellectual at the time, thought of themselves as pioneers of modernization and westernization. Philosophy was supposed to enlighten people and build a new way of thinking. The tendency to highlight the importance of doing philosophy with practical minds, rather than to introduce Western philosophy for its own sake, was manifest in the first issue of the above-mentioned journal, *Philosophy*. The articles published in the first issue included *One Question concerning the Starting Point of Philosophizing*, *What Is Philosophy?: On the Eternity of Philosophy*, *The Idea of an Ethical Evaluation*, and *The Structure of Concrete Existence*. In these articles, the nature of philosophy was defined with an emphasis on its relevance to practice. However, as the Japanese control over academia became ever more strict and rigid, emphases on practice grew weaker.

Korea was liberated from Japan in 1945. However, the country was divided into two Koreas with conflicting ideologies. This led, in 1950, to a calamitous national tragedy, the Korean War (1950–1953). This series of major events left significant marks on the contour of philosophy in Korea. Marxism, which was experimented with and advocated by a scant few philosophers during the Japanese colonial period, blossomed in the midst of the ideological conflicts that followed the liberation. Even though Marxism was soon officially suppressed in South Korea and many influential Marxist philosophers fled to the more hospitable North, Marxism left an indelible impression. Along with Marxism, existentialism emerged as a major player in Korean philosophy. This was mainly due to the Korean War; in particular, French existentialism, born in the ruins and despair of World War II (1939–1945), strongly resonated with Koreans with similar experiences during the Korean war.

A long-standing bias toward German philosophy began to change in the early 1950s. The Korean Philosophical Association was formed in 1953, and its official journal was founded. More important was that Korea started having direct contacts with Western philosophies. Philosophers came to visit Korea from the United States, Great Britain, and Germany. Students went to various parts of the world for studies. By having direct contacts, Korean philosophers gained firsthand access to Western philosophy, helping them to overcome the distortions inflicted by Japanese translations and interpretations.

Another outcome of this direct and broad exposure to Western philosophy has been the revival of interest in traditional philosophy. Ever-expanding contacts with diverse cultures and philosophies made Korean philosophers rethink the roots and identity of Korean thought. Traditional Korean philosophy, which had been ignored as useless and retrogressive during the Japanese colonial period, began to receive fresh scrutiny and assessment. In the late 1950s, Korean traditional thought came to be accommodated under the umbrella of philosophy.

As a result of interaction with diverse parts of the world, different trends of philosophy are evenly reflected in Korean philosophy today. Anglo-American analytic philosophy is one of the strongest trends. German philosophy is still going strong even though it is not as prominent as it once was. Many philosophers in Korea specialize in traditional Korean philosophy and other Asian philosophies. The world of philosophy in Korea is a melting pot. A large variety of traditions and trends are actively and vigorously represented—from phenomenology and existentialism to analytic philosophy, Buddhism, and Confucianism. Now philosophers pursuing diverse perspectives are starting to hold dialogues with each other. It is exciting to wait and see whether and how the world of philosophy in Korea will continue its tradition of integrationism and what the outcome will be.

See also Buddhism; Chinese Philosophy: Overview; Chinese Philosophy: Daosim; Confucius; Japanese Philosophy; Zhu Xi (Chu Hsi).

Bibliography

Buswell Jr., Robert E. *The Formation of the Ch'an Ideology in China and Korea*. Princeton NJ: Princeton University Press, 1989.

Chung, Edward. *The Korean Neo-Confucianism of Yi Toegye and Yi Yulgok—An Reappraisal of the "Four-Seven Thesis" and its Practical Implication for Self-Cultivation*. New York: SUNY Press, 1995.

Haboush, Jahyun Kim, and Wm. Theodore de Bary, eds. *The Rise of Neo-Confucianism in Korea*. New York: Columbia University, 1985.

Kalton, Michael. *The Four-Seven Debate—An Annotation of Translation of the Most Famous Controversy in Korean Neo-Confucianism*. New York: SUNY Press, 1994.

Keel, Hee-Sung. *Chinul, the Founder of the Korean Son Tradition*. Seoul: Pojinjae, 1984.

Kendall, Laurel, and Griffin Dix, eds. *Religion and Ritual in Korean Society*. Berkeley, CA.: Institute of East Asian Studies, 1987.

The Korean National Commission for UNESCO. *Main Currents of Korean Thought*. Seoul: Si-sa-young-o-sa Publishers, 1978.

Lancaster, Lewis. *Buddhism in Chosŏn*. Berkeley, CA.: Institute of East Asian Studies, 1996.

Lancaster, Lewis. *Buddhism in Koryŏ: A Royal Religion*. Berkeley, CA: Institute of East Asian Studies, 1996.

Lee, Peter H. *Sourcebook of Korean Civilization*. Vols. 1 and 2. New York: Columbia University Press, 1993.

Lee, Peter H. *Sources of Korean Tradition*. New York: Columbia University Press, 1997.

Park, Sung Bae. *Buddhist Faith and Sudden Enlightenment*. Albany, NY: SUNY Press, 1983.

Setton, Mark. *Chong Yag-yong—Korea's Challenge to Orthodox Neo-Confucianism*. New York: SUNY Press, 1997.

Shim, Jae-ryong. *Korean Buddhism: Tradition and Transformation*. Seoul: Jimoondang, 1999.

Namjin Huh and Kihyeon Kim (2005)

KORN, ALEJANDRO

(1860–1936)

Alejandro Korn, an Argentine metaphysician and ethical philosopher, was born in San Vicente. He took his doctorate in medicine and directed a hospital for the mentally ill. In 1906 he joined the faculty of philosophy and letters at Buenos Aires. Although he wrote little, he had immense personal influence on Argentine philosophy. His philosophical writing came late in his life: *La libertad creadora* (La Plata, 1930), his major work, is a compilation of five essays dating from 1918 to 1930.

Korn is sometimes called a positivist, a label suggested by his scientific training, his empiricism, the skeptical note in his metaphysics, and his ethical relativism. However, his "Incipit Vita Nova" (1918) set the stage for his own criticism of positivism. In this essay, he maintained that despite the scientific and technological progress of preceding decades, contemporary man is dissatisfied and disillusioned. The cause is the impairment of ethics by the spread of the positivistic doctrine that man is a machine without liberty; the remedy is a libertarian philosophy that subordinates science to ethics. Korn's sources were not Auguste Comte or Herbert Spencer, but Henri Bergson, Arthur Schopenhauer, and Immanuel Kant.

Korn's methodology rests on an experiential intuition whose objects are concrete particulars of ordinary experience. This common intuition is not passive and its content is not simple. Reason supplies concepts that are merely formal and symbolic but that penetrate intuition; the latter always has discursive elements. There is also a more intimate intuition or vision, which has intellectual, mystical, and aesthetic forms corresponding to metaphysics, religion, and art. Intuition as vision suggests pro-

found convictions and has an important place in the spiritual life of man, but it carries no assurance of truth. For comparative certainty we must turn to the two disciplines of ordinary intuition: science, which has a measurable object in the external world of fact, and axiology, which has an unmeasurable object in the internal world of evaluation. The third great intellectual enterprise, metaphysics, attempts to describe reality through concepts that transcend all possible experience. Metaphysical systems are dialectical poems. We cannot live without metaphysics, but we cannot convert it into a science; it should contain sincere convictions, free from dogmatism.

The external world of science, of the not-self, known through sensations, is spatial, measurable, and governed by strict causal law. The internal world of axiology, of the self, constituted of emotions, volitions, and judgments, is nonspatial, immeasurable, purposive, and free. These are the two halves of one encompassing domain of consciousness, which comprises all that we know and, it seems, all that is real. Common to both halves of consciousness are three further characters: activity or perpetual becoming, which shows that stable things and rigid names are false; relativity, which expresses the fact that every particular act has its reason in another; and time. Most significant in distinguishing the subjective from the objective order is freedom: economic freedom, or mastery of the external world, and ethical freedom, or mastery of self.

The search for an ultimate reality beyond consciousness led Korn to deny monistic realism, dualistic realism, and solipsism, and to affirm a type of absolute idealism. Experienced things, space, and time depend on consciousness, evidently because they involve organizing concepts or forms. A thing lying beyond consciousness and implied as cause of the experienced thing is denied: causality is a creature of our thought. The known object thus depends on consciousness and has its being there. But that does not entail the dependence of objects on my self. The self, or subjective order, is only a part of consciousness; it is not the source of the known world. The further definition of this idealism is through the theory of the *acción consciente*: consciousness as an everlasting, dynamic, and creative process, unknown in itself but manifested as aspiration toward absolute liberty.

This ontological goal is the key to Korn's theory of values. A value is the created object of an affirmative valuation, and valuation is the reaction of the human will to an event. Values therefore are subjective. There are instinctive, erotic, vital, economic, social, religious, ethical, logical, and aesthetic values, none of which can be reduced to any other. Values achieve unity through their common source in human personality and through their common goal in the liberty of man. Creative liberty is the recurring motif of Korn's philosophy.

See also Bergson, Henri; Comte, Auguste; Idealism; Intuition; Kant, Immanuel; Latin American Philosophy; Metaphysics; Positivism; Schopenhauer, Arthur.

Bibliography

Obras, 3 vols. La Plata: Tomás Palumbo, 1933–1940.

"Influencias filosóficas en la evolución nacional." *Revista de la universidad de Buenos Aires* (1912). Reprinted in *Obras,* Vol. III.

Apuntes filosóficos. Buenos Aires, 1935.

Arthur Berndtson (1967)

KOTARBIŃSKI, TADEUSZ
(1886–1981)

Tadeusz Kotarbiński, a Polish philosopher and logician, was born in Warsaw in 1886. He studied philosophy and the classics at the University of Lvov, where he obtained his doctorate in 1912. He began teaching at the University of Warsaw in 1918 and soon became perhaps the most influential philosophy teacher in Poland. His enlightened views, integrity, public spirit, and social zeal frequently brought him into conflict with established opinions and with the government, both before and after World War II. Admired by many and respected by all, Kotarbiński commanded a unique position of moral and intellectual prestige in his country. He was a member of the Polish Academy of Science and of the International Institute of Philosophy, and he was for a long time chairman of both bodies. He held an honorary doctorate from the Université Libre in Brussels and was a corresponding fellow of the British Academy and an honorary member of the Academy of Sciences of the U.S.S.R. and of other foreign scientific organizations.

CONCRETISM

Kotarbiński began his philosophical career as a minimalist. He advocated the abandonment of such terms as *philosophy* and *philosopher* because of their ambiguity and vagueness. The miscellaneous collection of subjects traditionally known as philosophy lacks any factual or logical coherence. These various subjects should be reconstructed as specialized fields of study and thus acquire some recognized criteria of professional competence. "The philosopher" should mean "the teacher of philosophy," and "philosophy" should be used restrictively to

denote moral philosophy and logic in the broad sense, which comprises formal logic, the philosophy of language, the methodology of science, and the theory of knowledge. Kotarbiński himself chose logic in this broad sense as the chief subject of his own concern. He wished to transform logic into a science as exact as mathematical logic and he applied himself to the construction of the conceptual apparatus necessary for this task. However, the results of this analytical work, accomplished between 1920 and 1935, exceeded the original design and produced a system known as reism or concretism. Kotarbiński regarded it as a program rather than a set doctrine and for linguistic reasons prefered *concretism* to *reism*.

Concretism arose from the puzzle about how qualities can belong to or inhere in the things of which they are characteristics. Kotarbiński believed that the puzzle can be resolved if we recognize that whereas things may be hard or soft, black or white, and so forth, nothing is hardness or softness, blackness or whiteness. Thus, the insight underlying concretism can be expressed in the proposition "only concrete individual objects exist." The expression "*a* exists" has the same meaning as "something is an *a*" (ex $a =_{\mathrm{Df}} (\exists x)$ *x* is *a*) and the meaning of *is* can be explicated as follows:

$$(a,b) :: a \in b . \equiv \therefore (\exists x) . x \in a \therefore (x) : x \in a . \supset . x \in b \therefore (x,y) :$$
$$x \in a .\ y \in a . \supset . x \in y .$$

This theorem is an early formulation of the single axiom of Leśniewski's ontology and should be read as an implicit definition of the functor "is" in expressions of the type "*a* is *b*," in which "is" has its main existential meaning.

SEMANTIC REISM. Concretism is both a metaphysical and a semantic doctrine; as metaphysics its basic characteristic is materialism and as semantics it is nominalism. Nominalism is an essential part of concretism, but materialism is not. For instance, Franz Brentano, although a concretist, was a Cartesian dualist.

If the dyadic functor "is" in expressions of the type "*a* is *b*" has the meaning defined above, then only genuine, empty or nonempty, shared or unshared names are admissible values for *a*. This should be clear in view of the fact that if *a* is *b*, then for some *x, x* is *a*, that is, *a* exists (therefore, if an empty name is substituted for *a*, "*a* is *b*" always becomes a false sentence). Semantic reism is a set of linguistic and logical rules that allow us to test the meaningfulness and truth of the expressions of language *L* as determined by their syntactic structure and semantic function.

According to semantic reism, names of concrete objects only, either corporeal or sentient, are genuine names. The names of properties, relations, events, facts, propositions, or classes are objectless and apparent names. Literally understood, sentences involving such fictitious names and implying the existence of properties, relations, events, facts, propositions, or classes are grammatically meaningful expressions, but reistically they are nonsense in disguise or falsehood. Only if, by a suitable transformation, such sentences can be reduced to equivalent expressions involving no apparent names can they become reistically meaningful and either true or false. For instance, in its literal meaning the sentence "the relation *being part of* is transitive" is either false or nonsensical. But if it is regarded as a shorthand statement of the fact that for all *x, y*, and *z*, if *x* is part of *y* and *y* is part of *z*, then *x* is part of *z*, the expanded version of this abbreviated sentence expresses a genuine and true proposition.

ONTOLOGICAL REISM. Nominalism is the view that the only admissible values for bound variables are entities of the lowest type as understood in the simplified theory of types. To apply this assumption outside logic and mathematics we need operational rules specifying the entities of the lowest type, that is, the referents of genuine names. For this purpose semantic reism must be supplemented by ontological reism; in other words, one's metaphysical commitments must be explicitly stated.

The basic proposition of ontological reism states that every object is a thing. *Object* is the most general ontological term, synonymous with *something*, the name of an arbitrarily chosen thing and thus extensionally equivalent to *thing. Thing* is a defined term and means a physical or a sentient body, in the nonexclusive meaning of *or. Physical* means spatial, temporal, and resistant, and *sentient* is defined by the Socratic definition as a term appropriately qualifying such bodies as animals or human beings (and probably also plants). Kotarbiński described ontological reism as somatism rather than as materialism, because for a reist "matter" is an apparent, quasi name, unless it is defined as a metatheoretical concept, in terms of which we speak about material or physical objects identified by the attributes of spatiality, temporality, and resistance and not by material substance. But somatism entails pansomatism, the proposition that every soul or mind (sentient entity) is a body. Therefore, a concretist who accepts pansomatism and asserts that there are only bodies in the universe is a materialist in the sense that he subscribes, speaking loosely, to the identity theory of mind and body. He leaves it to science to discover how it came about that there are sentient as well as physical bodies in the world.

In the theory of knowledge concretism implies the abandonment of the epistemological dualism of the theory of representative perception and the adherence to some form of sensational realism. Since there are no mental images or elements or sense data distinct from the object perceived, a concretist believes that all that is known is apprehended directly and that the so-called perceptual content is part of the physical object.

IMITATIONISM. If reality consists exclusively of bodies, and if the soul or mind is identical with part or the entire organism of a human individual, assertions about mental states and processes are not semantically well-formed sentences; they are objectionable on ontological grounds and consequently false. To be reistically acceptable they must be regarded as assertions of special sorts about persons, reducible, when fully stated, to descriptions of human individuals acting upon their environment and being affected by the external world. This view of the nature of psychological statements, together with the procedure by means of which they can be reduced to statements about persons doing and undergoing things, Kotarbiński called "imitationism." This name is intended to indicate that we come to understand the experiences of other people by imitating their behavior and, in general, that psychological knowledge is acquired not from introspection but by imitation or self-imitation.

Imitationism assumes that every singular psychological statement is a substitution of the schema "*A* experiences this: *P*," where *A* is a proper-name variable and *P* is a variable admitting all kinds of enunciations referring to the physical environment of the person whose name is substituted for *A*. The first part of the schema is the announcement by the experiencing person, *EP*, or the observer, *O*, of what its second part expresses by describing the environment in the same way that *EP* describes or would describe it. If *EP* and *O* are two different persons, the announcement refers to the imitation of *EP* by *O* and mentions the respect in which *EP* will be imitated. If *EP* and *O* are the same person, imitation becomes self-imitation and the description of the environment, including *EP*'s own body, is self-description.

PRACTICAL PHILOSOPHY

Kotarbiński had a lasting interest in practical philosophy. He saw its main task as the formulation of precepts and recommendations concerning the three questions of how to achieve happiness, how to live a good life, and how to act effectively. It is the second and third set of questions to which he devoted most attention. He was a staunch defender of the autonomy of ethics and approached its problems deontologically. Inspired both by a theoretical interest and by the desire to help his fellow men, he produced a general theory of efficient action known as praxeology. Although he had some predecessors, in particular A. A. Bogdanov (1873–1928) and Georges Hostelet (1875–1960), he accomplished pioneer work and opened a new field of study.

See also Brentano, Franz; Cartesianism; Logic, History of: Modern Logic: From Frege to Gödel; Materialism; Nominalism, Modern.

Bibliography

WORKS BY KOTARBIŃSKI

Wybór Pism (Selected works). 2 vols. Warsaw: Panstwowe Wydawn, 1957–1958. Includes all important essays and articles published between 1913 and 1954.

Elementy Teorii Poznania, Logiki Formalnej i Metodologii Nauk (Elements of the theory of knowledge, formal logic and methodology of science). Lvov, 1929; 2nd ed., Wrocław, Warsaw, and Kraków: Ossolinskich, 1961.

Kurs Logiki dla Prawników (A course of logic for lawyers). Warsaw: Gebethner and Wolff, 1951; 6th ed., 1963.

Traktat o Dobrej Robocie (Treatise on good work). Lodz: Zaklad im Ossolinskich we Wroclawiu, 1955; 2nd ed., Wrocław and Warsaw, 1958. Translated as *Praxiology: An Introduction to the Sciences of Efficient Action*. Oxford and Warsaw: Pergamon Press, 1965.

"The Fundamental Ideas of Pansomatism." *Mind* 64 (1955): 488–500.

Wyklady z Dziejów Logiki (Lectures on the history of logic). Lodz: Zaklad Narodowy im Ossolinskich we Wroclawiu, 1957. Translated as *Leçons sur l'histoire de la logique*. Paris: Presses Universitaires de France, 1964.

"Filozof" (The philosopher). *Studia Filozoficzne* (1) (1957): 4–16.

"Essai de réduire la connaissance psychologique à l'extraspection." In *Atti del XII Congresso Internazionale di Filosofia* (*Venezia, 12–18 Settembre 1958*)*—Proceedings of the XIIth International Congress of Philosophy*, 12 vols. Florence, 1958–1961. Vol. V (Florence, 1960), pp. 295–299.

"Zasady Etyki Niezależnej" (Principles of autonomous ethics). *Studia Filozoficzne* (1/4) (1958): 3–13.

"Fazy Rozwojowe Konkretyzmu" (The Stages of the development of concretism). *Studia Filozoficzne* (4/7) (1958): 3–13.

"La Philosophie dans la Pologne contemporaine." In *Philosophy in the Mid-Century: A Survey*. 4 vols., edited by Raymond Klibansky. Florence: Nuova Italia, 1958–1959. Vol. I, *Logic and the Philosophy of Science* (Florence, 1958), pp. 224–235.

"Psychological Propositions." In *Scientific Psychology: Principles and Approaches*, edited by Benjamin B. Wolman and Ernest Nagel, 44–49. New York, 1965.

WORKS ON KOTARBIŃSKI

Ajdukiewicz, Kazimierz. Review of Kotarbiński's *Elementy Teorii Poznania, Logiki Formalnej i Metodologii Nauk. Przegląd Filozoficzny* 33 (1930): 140–160. Reprinted in *Jezyk i Poznanie*. 2 vols. Warsaw: Panstwowe Wydawn, 1960, 1965, Vol. I, pp. 79–101. Available in Polish only.

Ajdukiewicz, Kazimierz. "Der logistische Antiirrationalismus in Polen." *Erkenntnis* 5 (1935): 151–161.

Grzegorczyk, Andrzej. "O Pewnych Formalnych Konsekwencjach Reizmu" (On certain formal implications of reism). In *Fragmenty Filozoficzne, Seria Druga*, 7–14. Warsaw: Panstwowe Wydawn, 1959. Available in Polish only.

Jordan, Z. A. *The Development of Mathematical Logic and Logical Positivism in Poland between the Two Wars*. London, 1945.

Jordan, Z. A. *Philosophy and Ideology: The Development of Philosophy and Marxism-Leninism in Poland since the Second World War*, 34–38, 195–197. Dordrecht, Netherlands: Reidel, 1963.

Jordan Z. A. "Próba Analizy Teorii Zdań Psychologicznych Prof. T. Kotarbińskiego" (An analysis of Professor Kotarbiński's theory of psychological statements). *Psychometria* 2 (1935): 347–375. Available in Polish only.

Lejewski, Czeslaw. "On Leśniewski's Ontology." *Ratio* 1 (1958): 150–176.

Rand, R. "Kotarbińskis Philosophie auf Grund seines Hauptwerkes: 'Elemente der Erkenntnistheorie, der Logik und der Methodologie der Wissenschaften.'" *Erkenntnis* 7 (1937–1938): 92–120.

Z. A. Jordan (1967)

KOZLOV, ALEKSEI ALEKSANDROVICH

(1831–1901)

Aleksei Aleksandrovich Kozlov, the Russian personalist philosopher, was the first major Russian exponent of a pluralistic idealism derived from Gottfried Wilhelm Leibniz. In his youth Kozlov studied the social sciences and was attracted to the ideas of Ludwig Feuerbach and François Marie Charles Fourier. His socialist views led to a short prison term in 1866 and the loss of his teaching position in a Moscow secondary school. He began to study philosophy seriously only in the 1870s, when, after an initial interest in materialism, he came successively under the influence of Arthur Schopenhauer, Eduard von Hartmann, and Immanuel Kant. In 1876 he became professor of philosophy at Kiev University, where he published the first Russian philosophical journal, *Filosofskii trekhmesiachnik* (Philosophical quarterly), and began to formulate his own mature position under the influence of Leibniz and his followers—notably Gustav Teichmüller. When illness forced Kozlov to retire in 1887, he moved to St. Petersburg and expounded his views systematically in a private journal, *Svoe slovo* (A personal word), published occasionally from 1888 to 1898.

In Kozlov's metaphysics, which he called panpsychism, there is a plurality of conscious spiritual substances, or monads. Each is an agent whose being consists not only in its substantiality, but also in its (psychic) activities and the contents of these activities. (Thus, Parmenides erred by considering substance alone, Johann Gottlieb Fichte by considering activity alone, and other philosophers erred similarly.) Together, these spiritual substances form a closed totality which is grounded in a Supreme Substance, God, and within which these substances (unlike Leibniz's monads) interact. The human body is a collection of less conscious spiritual substances with which our ego interacts until death. Kozlov suggested that after death the ego is reincarnated by interacting with other spiritual substances to form a new body.

The "material" aspect of the body, as of all supposed "material" entities, is produced by thought in our interaction with other spiritual substances, and is symbolic of these substances. Space and time (to which Kozlov devoted much attention) are likewise products of the thinking subject. Neither is objectively real, but each is symbolic of reality: Space is symbolic of the fact that real substances exist in connection, and time of the fact that within this connection there is variety and activity. Thus sense perception, which purports to show us objects in space and time, does not penetrate to the essentially timeless and spiritual reality. Kozlov developed an intuitivist epistemology, in which knowledge is based upon "primitive consciousness"—primarily consciousness of one's own ego. Primitive consciousness, however, being simple and immediate, is nonconceptual and ineffable. Knowledge, on the other hand, is complex and mediated; the mind constructs it by relating the elements of primitive consciousness. Thus we are directly conscious of God. Acquiring conceptual knowledge of God, however, is a difficult intellectual enterprise.

Kozlov did not develop his views fully in other areas, but his metaphysics and epistemology influenced many Russian philosophers, including his son, Sergei A. Askol'dov, and Nikolai Losskii.

See also Losskii, Nikolai Onufrievich.

Bibliography

ADDITIONAL WORKS BY KOZLOV

Filosofskie etiudy (Philosophical studies). 2 vols. St. Petersburg, 1876–1880.

Filosofiia kak nauka (Philosophy as a science). Kiev, 1877.

WORKS ON KOZLOV

Askol'dov, S. A. *A. A. Kozlov*. Moscow, 1912.

Lossky, N. O. "Kozlov: ego panpsikhizm" (Kozlov: His Panpsychism). *Voprosy filosofii i psikhologii* (Questions of philosophy and psychology) (58) (1901): 198–202.

Zenkovsky, V. V. *Istoriia Russkoi Filosofii*. 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*. 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

KRAUSE, KARL CHRISTIAN FRIEDRICH
(1781–1832)

Karl Christian Friedrich Krause, a German pantheistic philosopher, was born at Eisenberg in Thuringia. He studied at Jena, where he came under the influence of Johann Gottlieb Fichte and Friedrich Schelling. In 1812 he became Privatdozent, but his many efforts to secure a professorship were all unsuccessful. For a time he taught music in Dresden. In 1805 he joined the Freemasons, to further his ideal of a world society. His internationalist leanings were responsible for his failure to be appointed professor in Göttingen, and in Munich his chances were spoiled by the opposition of Schelling. Just as he finally obtained a position, Krause died of a heart attack.

Like several of his contemporaries, Krause claimed to be developing the true Kantian position. His orientation, however, was mystical and spiritualistic. The obscurity of his style is awesome; he expressed himself in an artificial and often unfathomable vocabulary which included such monstrous neologisms as *Or-om-wesenlebverhaltheit* and *Vereinselbganzweseninnesein*—words that are untranslatable into German, let alone into English. He called his system the theory of essence (*Wesenlehre*) and presented an elaborate set of categories, including Unity, Selfhood, Propositionality (*Satzheit*), "Graspness" (*Fassheit*), Unification-in-propositionality (*Satzheitvereinheit*), and so forth. The system was intended to mediate between pantheism and theism; hence Krause called his position "Panentheism," to suggest the idea that God or Absolute Being is one with the world, though not exhausted by it. From this central doctrine Krause derived a theory of man and of history. He regarded all men as part of a spiritual whole, an ideal League of Humanity (*Menschheitsbund*), the actualization of which is the goal of history.

Like Fichte, Krause took self-consciousness as his starting point in the belief that it provides a key to the essence of all things. The ego discovers itself to be both mind and body, enduring and changing; it is an organic, self-sustaining whole. According to Krause, this is the clue to the nature of other beings and of God. Considering its own finitude and that of other beings it encounters, the ego is led to the idea of an absolute, unconditioned principle upon which it and all other creatures and organizations are dependent. This principle is God, or Essence, whose nature is grasped in a spiritual intuition (*geistigen Schauen*), an immediately certain vision that is the foundation for all subsequent knowledge. God is primordial being (*Orwesen*), the being without contrareity; he is the unity of all that exists. Though he contains the world, he is nevertheless other than and superior to it. The distinction between God and the world is that of whole and part. Krause expressed this by speaking of God as *in himself* Contrabeing (*Gegenwesen*) and Unified Being (*Vereinwesen*), while *as himself,* or qua Primordial Being, he is absolute identity.

The existence of the world follows from an inner opposition in God's actuality (*Wesenheit*). Reason and Nature are two subordinate beings distinguished from, and yet lying within, God. Humanity is a synthesis of these. Humanity and the world, along with numerous basic human institutions, are organisms through which the divine life expresses itself. Thus, every being or group of beings is godlike in essence. Mind and body are integrated in the particular unified being that is man, reflecting the compresence of Reason and Nature in all things. Nature composes all individuals into a single whole. It is a mistake to view nature as a blind, mechanical system without consciousness; for its infinite perpetual activity, which is a pure self-determination, is free. Nature is a divine work of art; at the same time it is itself the artist, fashioning itself. The recognition of this divine character gives meaning and value to life.

Individual human minds together constitute the realm of Reason throughout which mind is organically distributed. But mind does not exist only in man and his institutions. Nature and Reason interpenetrate so fully that even animals are a unification of the two. Among animals, however, the career of each is fixed inexorably, according to the hierarchy of living forms. Man is the supreme unification of Reason and Nature, for he possesses the highest sort of mind joined to the highest sort of body. The individual souls that make up humanity are eternal, uncreated, immortal. Their number can neither be increased nor diminished. Humanity is thus complete at every moment.

What men should strive for is the imitation of the divine life in their own inner lives and in their social organizations. God is good, and men should participate in this goodness. The inner union with God (*Gottesinnigkeit*), or fervor for the divine, is the foundation of ethics, and ethics is the heart of religion. But individuals cannot achieve the moral life alone, since they are what they are only as parts of the whole. The community and its various institutions are thus indispensable.

Ideally, the community is governed by Right, which Krause defined as the organic whole of all of the internal and external conditions necessary for the completion of life that are dependent on freedom. This supernational law is grounded in the nature of the divine; it expresses the right of Humanity, not simply the right of individual human beings. The rights of individuals, groups, and nations can be recognized, but only as subordinate to the right of Humanity as a whole. Humanity is divided into a series of social organisms. There are, Krause speculated, human inhabitants in many cosmic systems. These human beings are subdivided into nations, races, communities, families, and so forth. There is an aesthetic community, a scientific community, a religious community, and a moral community. Each community has rights, although the right of Humanity takes precedence.

Men are all citizens of the universe, which is an infinite divine government. Because he revered the individual as a partial embodiment of the divine, Krause argued against the death penalty and maintained that punishment can be justified only as educative and reformatory. Only a republican form of government, he believed, is entirely compatible with the ideal of justice.

According to Krause's philosophy of history, the development of humanity is the temporal unfolding of a moral ideal. History follows a three-stage pattern, which is mirrored in every individual life as well. The development is not, however, purely progressive. There are two orders, one "ascending" and one "descending," so that the divine life may be presented again and again in the infinitely repeated epochs of history. The three steps in the ascending order are Wholeness, Selfhood, and Wholly-unified-selfhood. In the stage of Wholeness, each individual or higher organism exists germinally in the larger whole to which it belongs. In Selfhood, it enters into a free opposition to that whole and strives to develop its unique character. Evil appears as the individual organism tears itself loose from the harmony of the whole. Finally, the organism achieves a loving reunion with other beings (man, for example, becomes reunited with Nature, Reason, Humanity, and God), and with this rediscovery of harmony, all evil is negated. Afterward, however, the historical path leads downward, to a final involution that is both the ending of a career and the birth of a new life. Since the transition is gradual, an older age may survive for a time in a newer age. Each development, nevertheless, exhibits genuine, unforeseeable novelty.

Following this order, the individual man enters the world, proceeds through the stages of embryonic life, boyhood, and youth, and becomes increasingly independent, until he finally achieves the maturity of manhood, from which point he descends in a reverse series. Every human institution and organization pursues the same course of evolution, reflecting the basic laws of the divine organic life. In history, the first stage is marked by polytheism, slavery, caste systems, despotic governments, and a state of war between peoples. In the second period, the age of growth, men recognize the divine as an infinite being standing above all that is finite. This is monotheism, which Krause accuses of fostering theocracy, religious censorship of science and art, and contempt for the world. Finally, in the third stage (to which Krause's own philosophy is supposed to inspire men), humanity comes of age, the finite is reunited with the infinite, and world citizenship, philanthropy, and tolerance become the rule. According to Krause, the transition to this stage began with Benedict de Spinoza's discovery of the nature of being, and his own system was to be the development of that theory. He envisaged humanity as arriving at an organic completeness that represents the maturity of the race, and with visionary eloquence he depicted the unification of all humankind, as all men and all associations of men enter into a common life.

Krause's philosophy, while not very influential in Germany, found considerable support in Spain, where, for a time, "Krausism" flourished. This was largely due to the efforts of Julian Sanz del Rio, the minister of culture, who visited Germany and Belgium in 1844 and came into contact with a number of Krause's disciples, notably Heinrich Ahrens in Brussels and Hermann von Leonhardi in Heidelberg.

See also Consciousness; Fichte, Johann Gottlieb; Pantheism; Philosophy of History; Reason; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de.

Bibliography

Krause's most important work is *Das Urbild der Menschheit* (Dresden, 1812), translated into Spanish by Sanz del Rio as *El ideal de la humanidad* (Madrid, 1860). Included among his other works are *System der Sittenlehre* (Leipzig, 1810);

Vorlesungen über das System der Philosophie (Göttingen, 1828); and the short *Abriss des Systems der Rechtsphilosophie* (Göttingen, 1828).

For Krause's influence in Spain, see Sanz del Rio, *K. C. F. Krause: lecciones sobre el sistema de la filosofia analitica* (Madrid, 1850) and Juan Lopez Morillas, *El Krausismo español* (Mexico City, 1956). See also Hans Flasche, "Studie zu K. C. F. Krauses Philosophie in Spanien," in *Deutsche Vierteljarsschrift für Literaturwissenschaft und Geistesgeschichte* 14 (1936): 382–397. Flasche mentions the "left" and "right" wing of Krausism in Spain and, to account for Krause's success in Spain, tries to show (with a tenuous argument) the compatibility of Krause's views with Catholicism. Sharply critical of Krause is Eduard von Hartmann, "Krause's Aesthetik," in *Zeitschrift für Philosophie und philosophische Kritik* 86 (1) (1885): 112–130. Sympathetic accounts of Krause are to be found in Paul Hohfeld, *Die Krause'sche Philosophie* (Jena: H. Costenoble, 1879) and Rudolf Eucken, *Zur Erinnerung an K. Ch. Krause* (Leipzig, 1881). (Hohfeld edited a number of Krause's works, and Eucken studied with a student of Krause.) Clay Macauley, *K. C. F. Krause, Heroic Pioneer for Thought and Life* (Berkeley, CA, 1925) is a eulogistic pamphlet.

Arnulf Zweig (1967)

KRIPKE, SAUL

(1940–)

Saul Kripke is an American logician and philosopher born in New York in 1940. After earning a BA from Harvard University in 1962, he held positions at Harvard, Rockefeller, Princeton, New York Universities, and elsewhere.

MODAL LOGIC

Saul Kripke has worked in many branches of logic (higher recursion theory, set theory, models of arithmetic, and relevance logic), but the work best known to philosophers, and much cited in the literature of linguistic semantics, computer science, and other disciplines, is his development of Kripke models for modal and related logics. At the level of sentential logic such a model consists of a set *X* (of "states of the world," often misleadingly called "worlds"), a binary relation *R* (of "relative possibility") thereon, plus an assignment to each atomic formula *p* of the set of those *x* in *X* at which *p* is true. The assignment extends to all formulas, taking "Necessarily *A*" to be true at *x* if *A* is true at every *y* with *xRy*.

Kripke was the first to publish proofs of completeness theorems to the effect that truth at all *x* in all models with *R* reflexive (and transitive) (and symmetric) coincides with provability in the modal logic **T** (respectively **S4**) (respectively **S5**), and he obtained similar results for other modal logics. Announced in "Semantic Considerations on Modal Logic" (1963), and presented in detail in a subsequent series of technical papers, Kripke's work covers modal and intuitionistic sentential and predicate logic, and includes besides completeness theorems results on decidability and undecidability.

SEMANTIC PARADOXES

Also well known is Kripke's work on semantic paradoxes in "Outline of a Theory of Truth" (1975). A truth-predicate in a language *L* permitting quotation or equivalent means of self-reference would be a predicate *T* such that the following biconditional holds with any sentence of *L* in the blanks:

"*T*('__________')" is true if and only if "__________" is true.

The liar paradox shows there cannot be a truth-predicate in *L* if *L* has no truth-value gaps. Given a partial interpretation *I* of a predicate *U* (under which *U* is declared true of some items, declared false of others, or not declared either of the rest), any treatment of truth-value gaps, such as Stephen Cole Kleene's three-valued or Bas van Fraassen's supervaluational approach, will dictate which sentences containing *U* are to be declared true, declared false, or not declared either. If *U* is being thought of as "is true," this amounts to dictating a new partial interpretation I^* of *U*. For a fixed point, or partial interpretation having $I = I^*$, the biconditional displayed earlier holds.

Kripke's work, besides more purely philosophical contributions, shows how to obtain a minimal fixed point (contained in any other, and explicating an intuitive notion of *groundedness*), a maximal intrinsic fixed point (not declaring true anything declared false by any other fixed point), and many others, for any reasonable treatments of gaps.

WITTGENSTEIN AND SKEPTICISM

Turning from logic to philosophy of language and its applications to analytic metaphysics, Kripke has written two much-discussed books that are almost entirely independent of each other. In *Wittgenstein on Rules and Private Language* (1982) he advances as noteworthy, though not as sound, an argument inspired by his reading of Ludwig Wittgenstein's *Philosophical Investigations* (1953/1993) that is not unqualifiedly attributed to Wittgenstein. On Kripke's reading the target of the argument is any theory (such as that of the *Tractatus Logico-Philosophicus*) that conceives of meaning as given by

conditions for truth, conceived as correspondence with facts. Kripke compares Wittgenstein as he reads him to David Hume (more specifically, to a version of Hume that takes seriously his protestations that he is only a mitigated, not an extreme skeptic). So read, Wittgenstein's attack on correspondence theories of meaning consists, like Hume's attack on rationalist theories of inference, of two phases.

First there is a "skeptical paradox." Consider an ascription of meaning, say that according to which by "plus" I mean *plus*, so that 125 is the right answer to the question "what is 68 plus 57?" as I mean it. To what fact does this correspond? Not the record of how I have worked sums in the past. (Perhaps I have never worked this one before, and many rules are compatible with all the ones I have worked so far.) Not my ability to state general rules for doing sums, since this only raises the question what fact corresponds to my meaning what I do by the words in these rules. Not my behavioral dispositions (nor anything in the structure or functioning of my brain causally underlying them) since what answer I am *disposed* to give is one question, and what answer would be the *right* one for me to give is another question; and I am disposed to give wrong answers fairly often even for medium-sized numbers, and to give no answer at all for really big ones. Further considerations rule out also introspectable feelings accompanying calculation. No candidates seem to remain, so it seems that there is *no* fact to which an ascription of meaning corresponds. The conclusion is that *if* meaning consists in conditions for truth and truth of correspondence with facts, *then* ascriptions of meaning like "What I mean by 'plus' is *plus*" are neither true nor meaningful, and no one ever means anything by anything.

Second, there is a "skeptical solution," defying short summary. This solution identifies the meaningfulness of a sentence with the possession *not* of truth-conditions but of a potential for use within a speech community. The aspects of use—of usage and utility—that are emphasized are on the one hand the conditions under which assertion of a sentence is warranted, and on the other hand the applications warranted when a sentence is accepted.

One objection, anticipated by Kripke, is that Wittgenstein does accept talk of "truth" and "facts" in a deflated sense, in which sense to say, "It is true or a fact that by 'plus' I mean *plus*," amounts to no more than saying, "By 'plus' I mean *plus*," which on Kripke's reading Wittgenstein never denies. So a straightforward statement of Wittgenstein's view as the thesis that there are no "facts" corresponding to meaning ascriptions will not do. But as Kripke notes, one of the tasks of a reading of Wittgenstein is precisely to explain why he does *not* state his view in straightforward philosophical theses. Other objections to Kripke's interpretation, which has Wittgenstein opposing one theory of meaning to another, have been advanced by those who interpret Wittgenstein as a "therapist" who aims to treat philosophical questions not by developing philosophical theories (of meaning or of anything else) to answer these questions, but by developing methods to cure one of wanting to ask such questions. But such a reading may be less utterly irreconcilable with the reading of Wittgenstein as skeptic than its proponents generally recognize, since after all historical skepticism was itself a form of psychotherapy, aiming to achieve philosophic *ataraxia* by cultivating indifference to unanswerable questions.

REFERENCE AND METAPHYSICS

Kripke's most famous work is *Naming and Necessity* (1980), which consists of a transcription (with addenda and a preface written a decade later) of lectures given at Princeton in 1970. Only a rough, brief treatment will be possible here, leaving entirely to one side the influential ancillary papers "Speaker's Reference and Semantic Reference" and "A Puzzle about Belief," related work of Keith Donnellan (on proper names) and Hilary Putnam (on natural kind terms), and Kripke's provocative discussion of several side topics (among them the contingent *a priori* and the identity theory in philosophy of mind).

Kripke maintains the following doctrines about naming, illustrating them with examples, many of which have become famous. The reference of a proper name (e.g. "Phosphorus," "Feynman," "Newton") is *not* determined by some associated definite description (or cluster of descriptions, which is to say, description of the form "the object of which *most* of the following is true …"). The description a speaker associates with a name may be incorrect. (The speaker may describe Isaac Newton as "the man who was hit on the head by an apple and thereby struck with the idea of a force of gravity.") Even if correct, it may fail to be uniquely identifying. (The speaker may be able to describe Richard Feynman [1918–1988] only as "a famous physicist," which does not distinguish him from Murray Gell-Man [1929–].) Even if correct and uniquely identifying, it may be so only contingently, so that in speaking of certain counterfactual situations the description may denote something else or nothing at all. (Phosphorus, though it *is* the brightest object regularly visible in the eastern sky before sunrise,

might have only been second brightest, in which case "the brightest …" would have denoted something else; while if it had been tied for brightest, "*the* brightest …" would denote nothing.) By contrast, names designate rigidly, continuing to designate the same thing even when discussing counterfactual hypotheses. (If I say, "If there had been a brighter object, Phosphorus would have been only second brightest," I am still speaking of *Phosphorus*.)

A better picture than the description theory of how a name comes to denote its bearer would be this: The first user of the proper name or "initial baptist" may fix its reference by some description (possibly involving demonstratives and requiring supplementation by ostension, for example, "that bright object over there by the eastern horizon"). The second user may use the name with the intention of referring to whatever the first user was referring to, while perhaps ignorant of the original description. And so on in a historical chain. (Some commentators say causal chain, but it is important to note that there need not be any causal connection between initial baptist and thing named, which may be a mathematical object.) Kripke also offers an analogous picture of how a natural kind term comes to denote the kind of things it does.

Kripke also maintains the following doctrines about necessity, partly as corollaries to the above doctrines about naming. A true identity linking proper names (e.g., "Hesperus is Phosphorus") is necessary (as a consequence of rigidity, since even in a counterfactual situation each name will continue to denote the bearer it actually denotes, and therefore the two will continue to denote the *same* object, if they actually do so). But such an identity is not *a priori* (the identity of the heavenly body spotted at dawn and called "Phosphorus" with the one seen at dusk and dubbed "Hesperus" being an empirical astronomical discovery). Therefore, the metaphysical notion of necessity, "what could not have been otherwise," must be distinguished from epistemological notions like "what can be known *a priori* to be so."

There are other examples of metaphysical necessities, many involving natural kind terms: the facts of identity of heat with random molecular motion, of water with H_2O, of gold with the element of atomic number 79, and more; that a given object (e.g., a table) is composed of the material it is composed of (wood rather than ice); that a given person or organism has the ancestry he, she, or it does (e.g., that Elizabeth II is the daughter of George VI, and if he had had no daughter, she would never have been born). This is so even though in none of these examples does one have *a priori* knowledge. (There would be no internal logical contradiction in a tabloid press article claiming Elizabeth II to be the daughter of Harry Truman.) Historically, from Immanuel Kant to Gottlob Frege to Rudolf Carnap and beyond, necessity had tended to dwindle to aprioricity, which in turn had tended to dwindle to analyticity; Kripke's sharp reversal of this trend is perhaps his most important single contribution to philosophy.

See also Liar Paradox, The; Modal Logic; Philosophy of Language; Philosophy of Mind; Putnam, Hilary.

Bibliography

Wittgenstein, Ludwig. *Philosophical Investigations* (1953). Translated by G. E. M. Anscombe. Malden, MA: Blackwell, 1993.

WORKS BY KRIPKE

"Semantic Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94.

"Outline of a Theory of Truth." *Journal of Philosophy* 72 (1975): 690–716.

"Speaker's Reference and Semantic Reference." *Midwest Studies in Philosophy* 2 (1977): 255–276.

"A Puzzle about Belief." In *Meaning and Use: Papers Presented at the Second Jerusalem Philosophical Encounter, April 1976*, edited by Avishai Margalit. Dordrecht, Netherlands: Reidel, 1979.

Naming and Necessity. Cambridge, MA: Harvard University Press, 1980.

Wittgenstein on Rules and Private Language: An Elementary Exposition. Cambridge, MA: Harvard University Press, 1982.

John P. Burgess (1996, 2005)

KRISTEVA, JULIA
(1941–)

Julia Kristeva was born on June 24, 1941, in Sliven, Bulgaria. She was educated by French nuns, studied literature, and worked as a journalist before going to Paris in 1966 to do graduate work with Lucien Goldmann and Roland Barthes. While in Paris she finished her doctorate, was appointed to the faculty of the Department of Texts and Documents at the University of Paris VI (Denis Diderot) and began psychoanalytic training. Currently, Kristeva is Director of the Department of Science of Texts and Documents at the University of Paris VII, where she teaches in the Department of Literature and Humanities.

In her early writing, Kristeva is concerned with bringing the speaking body back into phenomenology and linguistics. In order to counteract what she sees as the necrophilia of phenomenology and structural linguistics,

which study a dead or silent body, Kristeva develops a new science that she calls *semanalysis*. She describes semanalysis as a combination of semiology (or Semiotics) from Ferdinand de Saussure, and psychoanalysis from Sigmund Freud. Unlike traditional linguistics, semanalysis addresses an element that is heterogeneous to language, the unconscious. The introduction of the unconscious into the science of signs, however, challenges the possibility of science, meaning, and reason. This is why Kristeva maintains that certain nineteenth-century poets whose work discharged unconscious drive force and emphasized the semiotic element of signification began a revolution in poetic language.

With semanalysis, Kristeva attempts to bring the speaking body, complete with drives, back into language. She does this both by putting language back into the body and by putting the body into language. She argues that the logic of signification is already present in the material body. In *Revolution in Poetic Language* she suggests that negation and identification—the two primary logical operations of language—are already operating within the body prior to the onset of signification: Expelling waste from the body prefigures negation and incorporating food into the body prefigures identification. The second way in which Kristeva brings the speaking body back to language is by maintaining that bodily drives make their way into language. One of Kristeva's major contributions to philosophy of language is her distinction between two heterogeneous elements in signification: the semiotic and the symbolic. Within Kristeva's writings, *semiotic* (*le sémiotique*) becomes a technical term that she distinguishes from *semiotics* (*la sémiotique*). The semiotic elements within the signifying process are the drives as they discharge within language. This drive discharge is associated with rhythm and tone. The semiotic has meaning but not does refer to anything. The symbolic, on the other hand, is the element of language that allows for referential meaning. The symbolic is associated with syntax or grammar and with the ability to take a position or make a judgment that syntax engenders.

Kristeva describes the relation between the semiotic and the symbolic as a dialectic oscillation. Without the symbolic there is only delirium, whereas without the semiotic, language would be completely empty, if not impossible. There would be no reason for people to speak if it were not for the semiotic drive force. The oscillation between the semiotic and the symbolic is both productive and necessary. The oscillation between rejection and stasis already existing within the material body produces the oscillation between semiotic and symbolic in the speaking subject.

In *The Powers and Limits of Psychoanalysis*, Kristeva revists the theme of revolution so prominent in her earlier work. In *Revolution in Poetic Language* Kristeva identifies the possibility of revolution in language—a revolution she deems *analogous* to social revolution—with (maternal) semiotic forces in avante-garde literature. In *Powers of Horror* this semiotic force of drives is not only associated with the maternal but more particularly with the abject or revolting aspects of the maternal. Here, the revolting becomes revolutionary through the return of the repressed (maternal) within (paternal) symbolic systems. Two decades later, in *The Sense and Non-Sense of Revolt*, Kristeva asks if revolt is possible today. In this book, volume one of *The Powers and Limits of Psychoanalysis*, she claims that within postindustrial and post-Communist democracies we are confronted with a new political and social economy governed by the spectacle within which it becomes increasingly difficult to think of the possibility of revolt. The two main reasons are that within media culture, the status of power and the status of the individual have changed. Kristeva argues that in contemporary culture there is a power vacuum that results in the inability to locate the agent or agency of power and authority or to assign responsibility. In a no-fault society, who or what can people revolt against?

In addition to the power vacuum, Kristeva identifies the impossibility of revolt with the changing status of the individual. The human being as a person with rights is becoming nothing more than an ensemble of organs that can be bought and sold or otherwise exchanged, what she calls the *patrimonial individual*. And, how can an ensemble of organs revolt? Not only is there no one or nothing to revolt against, but also there is no one to revolt. And without the possibility of revolt, especially the psychic revolt necessary for creativity, people are left with new maladies of the soul that make life seem meaningless.

In her *Female Genius* trilogy, Kristeva suggests that women with their attention to the sensory realm may provide an antidote for the meaninglessness that results from contemporary forms of nihilism. She argues that the genius of extraordinary women such as Hannah Arendt, Melanie Klein, and Colette help all women to see what is extraordinary in their own ordinary lives. Conversely, Kristeva maintains that the genius of everyday life is women's genius, particularly the genius of mothers; in creating new human beings, mothers are singular innovators, reinventing the child anew all the time. Kristeva

maintains that mothers may represent a safeguard against the automation of human beings.

See also Aesthetics, History of; Arendt, Hannah; Barthes, Roland; Feminism and Continental Philosophy; Feminism and the History of Philosophy; Feminist Aesthetics and Criticism; Feminist Philosophy; Freud, Sigmund; Language and Thought; Modernism and Postmodernism; Philosophy of Language; Psychoanalysis; Structuralism and Post-structuralism; Unconscious; Women in the History of Philosophy.

Bibliography

WORKS BY KRISTEVA

La révolution du langage poétique: L'avant-garde à la fin du XIXe siècle, Lautréamont et Mallarmé. Paris: Éditions du Seiul, 1974. Translated by Margaret Waller as *Revolution in Poetic Language* (New York: Columbia University Press, 1984).

Pouvoirs de l'horreur. Paris: Éditions du Seiul, 1980. Translated by Leon Roudiez as *Powers of Horror* (New York: Columbia University Press, 1982).

Histoires d'amour. Paris: Éditions Denoêl, 1983. Translated by Leon Roudiez as *Tales of Love* (New York: Columbia University Press, 1987).

Soleil noir: Dépression et mélancolie. Paris: Gallimard, 1987. Translated by Leon Roudiez as *Black Sun: Depression and Melancholy* (New York: Columbia University Press, 1989).

Étrangers à nous-memes. Paris: Fayard, 1989. Translated by Leon Roudiez as *Strangers to Ourselves* (New York: Columbia University Press, 1991).

Les nouvelles maladies de l'ame. Paris: Fayard, 1993. Translated by Ross Guberman as *New Maladies of the Soul* (New York: Columbia University Press, 1995).

Sens et non-sens de la révolte: Pouvoirs et limites de la psychanalyse I. Paris: Fayard, 1996. Translated by Jeanine Hermane as *The Sense and Non-Sense of Revolt* (New York: Columbia University Press, 2000).

La révolte intime: Pouvoirs et limites de la psychanalyse II. Paris: Fayard, 1997. Translated by Jeanine Herman as *Intimate Revolt: The Powers and Limits of Psychoanalysis, Volume II* (New York: Columbia University Press, 2002).

Le génie féminin, Tome premier, Hannah Arendt. Paris: Librairie Arthème Fayard, 1999. Translated by Ross Guberman as *Feminine Genius, Volume One, Hannah Arendt* (New York: Columbia University Press, 2001).

Le génie féminin, Tome II, Melanie Klein. Paris: Librairie Arthème Fayard, 2000. Translated by Ross Guberman as *Feminine Genius, Volume Two, Melanie Klein* (New York: Columbia University Press, 2001).

The Portable Kristeva. 2nd ed. Edited by Kelly Oliver. New York: Columbia University Press, 2002.

WORKS ABOUT KRISTEVA

Beardsworth, Sarah. *Julia Kristeva: Psychoanalysis and Modernity*. Albany, NY: SUNY Press, 2004.

Lechte, John, and Maria Margaroni. *Julia Kristeva: Live Theory*. London: Continuum Press, 2004.

McAfee, Noelle. *Julia Kristeva*. New York: Routledge, 2004.

Oliver, Kelly, ed. *Ethics, Politics and Difference in Kristeva's Writing*. New York: Routledge Press, 1993.

Oliver, Kelly. *Reading Kristeva: Unraveling the Doublebind*. Bloomington: Indiana University Press, 1993.

Kelly Oliver (2005)

KROPOTKIN, PËTR ALEKSEEVICH
(1842–1921)

Pëtr Alekseevich Kropotkin, the geographer and libertarian philosopher, was the principal exponent of the theories of anarchist-communism. He was born of a line of Russian princes who claimed descent from Riurik, the reputed founder of the Russian Empire. His father was a general, and he himself seemed destined for a military career. He was educated in the Corps of Pages and served as personal attendant to Tsar Alexander II. When the time came for him to choose a career, Kropotkin applied for a commission in the Mounted Cossacks of the Amur and went to Siberia because he felt his chance of serving humanity was greater there than in Russia. He had already come under the influence of liberal ideas through reading the clandestinely distributed writings of Aleksandr Herzen.

In Siberia Kropotkin carried out an investigation of the Russian penal system, which aroused in him a revulsion against the effects of autocratic government. During the early 1860s he led a series of expeditions into the untraveled regions of Siberia and, on the basis of his observations, developed an original and influential theory concerning the structure of the mountains of Asia. He also made important discoveries regarding the glacial ages and the great desiccation of east Asia, which resulted in the onset of barbarian wanderings.

In the solitude of the Siberian wastes, Kropotkin's thoughts turned more and more toward social protest. In 1865 the exiled poet M. L. Mikhailov introduced him to the writings of the French anarchist Pierre-Joseph Proudhon, and in 1866 Kropotkin resigned his commission in protest against the execution of a group of Polish prisoners who had tried to escape.

For some years he devoted himself to science, and in 1871 he was exploring the eskers of Finland when he was offered the secretaryship of the Russian Geographical Society. It was the moment of decision. Kropotkin was already feeling the urge to "go to the people" that affected many of the conscience-stricken Russian noblemen of the

1870s, and he decided to abandon science. In 1872 he visited Switzerland to make contact with exiled Russian liberals and revolutionaries. After listening to many radical views, he went to the Jura, where the watchmakers were fervent disciples of Mikhail Bakunin. "When I came away from the mountains, after a week's stay with the watchmakers, my views upon socialism were settled; I was an anarchist" (*Memoirs of a Revolutionist*).

In Russia Kropotkin joined the underground circle led by Nikolai Chaikovskii. In 1874 he was arrested and imprisoned in the Peter and Paul Fortress. Two years later he made a sensational escape and returned to western Europe, where he became an active worker in the rising anarchist movement. In 1879 he founded *Le révolté*, the most important anarchist paper to appear since the end of Proudhon's journalistic career in 1850, and in 1881 he took part in the London International Anarchist Congress, which founded the celebrated but short-lived "Black International." In 1882 he was arrested by the French authorities and was tried at Lyons along with a number of French anarchists. He was sentenced to five years imprisonment for alleged membership in the International Workingmen's Association. The sentence aroused wide international protest, and Kropotkin was released early in 1886. He went to England, where he lived until he returned to Russia after the 1917 revolution.

Kropotkin's career in western Europe was sharply altered by his arrival in England. On the Continent, from 1876 to 1886, he had been a revolutionary agitator, conspiring, lecturing, pamphleteering, and taking part in radical demonstrations. His writings were mainly periodical pieces for *Le révolté*. At first they were topical, but by 1880 Kropotkin was already developing the theory of anarchist-communism in a series of articles later incorporated in two books—*Paroles d'un révolté* (Paris, 1885) and *La conquête du pain* (Paris, 1892).

ANARCHIST-COMMUNISM

The doctrine of anarchist-communism differed from the collectivism preached by Bakunin and his followers in the 1860s in that it considered the need of the consumer rather than the achievement of the producer as the measure for distribution. In the vision of the anarchist-communist, the free-distribution warehouse would replace the earlier systems evolved by Proudhon and retained by the collectivists, which determined the worker's due either by hours of labor or quantity of production. Also, the anarchist-communists laid particular stress on the commune (in the sense of locality), rather than the industrial association, as the unit of social organization. In other respects—their rejection of the state, their stress on federalism, their emphasis on direct rather than parliamentary action, their denunciation of political forms—they did not differ profoundly from other schools of anarchism.

SOURCES. Although he became its leading exponent, Kropotkin did not originate anarchist-communism. The form of distribution embodied in the theory dates back at least as early as Thomas More's *Utopia* (1515–1516), and it appeared in a modified form in François Marie Charles Fourier's Phalansterian communities. The geographer Élisée Reclus, a former Phalansterian, appears to have brought the idea with him when he came to anarchism; it was first developed in writing by François Dumartheray, a Geneva artisan who helped Kropotkin in the founding of *Le révolté*. But Kropotkin developed the theory and, in *La conquête du pain*, he tried to show how it would work. This benign vision of an anarchist future reflects not only the optimism of Kropotkin's views, but also the benevolence of his character. For, although he always paid homage to the ideas of violent revolution, he did so against his nature; as Lev Tolstoy shrewdly remarked, "His arguments in favour of violence do not seem to me the expression of his opinions, but only of his faith to the banner under which he has served all his life."

ANARCHISM AND SCIENCE

When he reached England, Kropotkin moved into a world where he was respected by people in all walks of life. His achievements as a geographer were remembered; he was honored by learned societies; his articles were published in scientific journals; and his books were welcomed by respectable publishers. He did not abandon his ideals, but his role changed from that of agitator to that of writer and libertarian philosopher.

The most important books Kropotkin wrote during this period were his autobiography, *Memoirs of a Revolutionist* (New York, 1899), and *Mutual Aid: A Factor of Evolution* (London, 1902). *Mutual Aid*, together with *Modern Science and Anarchism* (London, 1912), shows Kropotkin attempting to base anarchist theory on a scientific foundation. These books reveal him as a devoted evolutionist, to the extent that he explains revolutions as part of the natural process by which man, as a social animal, evolves. He sees revolutions arising obscurely in the consciousness of the people and punctuating the slow tenor of progress by sudden mutations in social organization, while he views anarchism as a backward trend toward a natural order that has been perverted by the emergence of

authoritarian institutions. Man is naturally social, he suggests; therefore he does not need government, which itself perpetuates the unequal conditions that breed crime and violence. In their sociality, human beings resemble the more successful species of animals that depend for their survival on cooperation among their members. This idea is the core of *Mutual Aid*, which is an attempt, based largely on the arguments of K. F. Kessler, to reform evolutionary theory by demonstrating that the neo-Darwinians wrongly stressed competition as a factor in evolution, to the exclusion of cooperation. In biological terms, his point was well taken; the appearance of *Mutual Aid* led to modifications in evolutionary theory. But Kropotkin never convincingly welded his ideal of mutual aid to his anarchistic love of freedom, since he ignored the extent to which customs restrict liberty in most societies in which nongovernmental cooperation dominates the pattern of life.

Kropotkin's departure to Russia in 1917 led to tragic disappointment. He found himself out of touch with Russian realities and isolated during the events that led to the October Revolution. He retired to the village of Dmitrov outside Moscow, where he spent his last years writing. He denounced the Bolshevik dictatorship and the terror it imposed. When he died in 1921, his funeral was the last great demonstration against communist rule.

ETHICS

Kropotkin's last years were spent on the uncompleted *Etika* (*Ethics*), which was published posthumously in Moscow in 1922. In part a history of ethical theories, this book seeks to present ethics as a science. In developing his naturalistic viewpoint, Kropotkin shows the emergence of morality among animals as an outgrowth of mutual aid and demonstrates its extension into human society, where it acquires a disinterestedness that goes beyond mere equality. He sees morality as the extension of human good will beyond equity and justice. The historical parts of *Ethics* are admirable, but the work is incomplete; Kropotkin's own ethical system is barely worked out.

See also Anarchism; Bakunin, Mikhail Aleksandrovich; Communism; Evolutionary Ethics; Evolutionary Theory; Fourier, François Marie Charles; Herzen, Aleksandr Ivanovich; Libertarianism; More, Thomas; Proudhon, Pierre-Joseph; Russian Philosophy; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

WORKS BY KROPOTKIN

Anarchism: A Collection of Revolutionary Writings. Edited by Roger N. Baldwin. Mineola, NY: Dover Publications, 2002.

The Conquest of Bread. New York: New York University Press, 1972.

The Great French Revolution. Montreal: Black Rose Books, 1989.

Kropotkin's Revolutionary Pamphlets: A Collection of Writings. New York: B. Blom, 1968.

Russian Literature: Ideals and Realities. Montreal and New York: Black Rose Books, 1991.

Works. Vols. 7–11 edited by George Woodcock. Montreal and New York: Black Rose Books, 1992–1995.

WORKS ON KROPOTKIN

Cahm, Caroline. *Kropotkin and the Rise of Revolutionary Anarchism, 1872–1886*. Cambridge, U.K., and New York: Cambridge University Press, 1989.

Marshall, Peter. *Demanding the Impossible: A History of Anarchism*. London: HarperCollins, 1992.

Morland, David. *Demanding the Impossible? Human Nature and Politics in Nineteenth-Century Social Anarchism*. London and Washington, DC: Cassell, 1997.

Morris, Brian. *Kropotkin: The Politics of Community*. Amherst and New York: Humanity Books, 2004.

Woodcock, George, and Ivan Avakumovic. *Peter Kropotkin: From Prince to Rebel*. Montreal: Black Rose Books, 1990.

George Woodcock (1967)
Bibliography updated by Vladimir Marchenkov (2005)

KRUEGER, FELIX
(1874–1948)

Felix Krueger, a German philosopher and psychologist, was born in Poznán and received his doctorate in 1897 from the University of Munich, where he studied under Hans Cornelius and Theodor Lipps. After working as an assistant at the Physiological Institute in Kiel he became a *Privatdozent* at Leipzig under Wilhelm Wundt. From 1906 to 1908 Krueger held a professorship at Buenos Aires, where he organized the development of scientific psychology in Argentina and left lasting traces of his views and activities. After returning to Leipzig he was called to Halle to succeed Hermann Ebbinghaus. In 1912–1913 Krueger was an exchange professor at Columbia University. In 1917, after three years of military service, he returned to Leipzig as Wundt's successor. At Leipzig Krueger founded the second Leipzig school of psychology, whose basic principles were designated as a genetic psychology of wholeness and structure (*genetische Ganzheits- und Strukturpsychologie*). In 1928 he received an honorary doctorate from Wittenberg College, Springfield, Ohio. In

1935 Krueger was appointed rector of Leipzig University. He immediately became involved in political conflicts and was removed from the rectorship and for some time forbidden to lecture; in 1935 he retired prematurely from academic life. Krueger edited two series of psychological works, "Neue psychologische Studien" and "Arbeiten zur Entwicklungspsychologie," from 1914 and 1926, respectively. Early in 1945 he moved to Switzerland.

Krueger's first work, a philosophical one, was *Der Begriff des absolut Wertvollen als Grundbegriff der Moralphilosophie* (The concept of the absolutely valuable as the basic concept of moral philosophy; Leipzig, 1898). In this work he presented a critique of Immanuel Kant running counter to that of Neo-Kantianism. He tried to show that there was a material vein in the formal ethics of Kant himself, and he stressed that ethical responsibility is moored in the person, in his "energy of evaluation" (*Energie des Wertens*) and in his attitude toward values (*Werthaltung*), which Krueger understood as the "core structure" of personality or character.

After this work Krueger turned to empirical and experimental psychology, in which he became known particularly for his new theory of consonance and dissonance based on the influence of the different tones and for experiments in phonetics and the psychology of speech. In connection with this work he began to develop, as early as 1900, a theory of psychological wholeness, arising from the exhibition of emotional and physiognomic experiencing, which he characterized as a quality of complexes (*Komplexqualität*) parallel to Christian von Ehrenfels's Gestalt qualities (*Gestaltqualität*). Together with his English friend and student (who was, nevertheless, older than he), Charles Spearman, Krueger introduced into psychology the calculus of correlation including the first reflections on factor analysis.

In 1915, in *Über Entwicklungspsychologie, ihre historische und sachliche Notwendigkeit* (On developmental psychology, its historical and factual necessity) Krueger developed a theory of cultural origins departing from Wundt's psychology of peoples and carried it further in *Zur Entwicklungspsychologie des Rechts* (The developmental psychology of law; "Arbeiten zur Entwicklungspsychologie," No. 7, Munich, 1926). In 1918 and (in English) in 1927, Krueger presented sketches for a theory of the emotions, which he defined as the *Komplexqualitäten* of one's total experience, that is, as supersummative qualities not to be confused or identified with gestalt.

These various strands, including his old moral philosophy, were united by Krueger in 1923 in a theory of structure, which was both critically related to and opposed to the thought of Wilhelm Dilthey. Krueger defined structure as the new scientific conception of the mind, as "the organismic construct of psychophysical wholeness," that is, as the basis of events in experience in the form of disposition, attitude and readiness, inclination, habit, and capability. The existence and individuality of personal structure can be demonstrated particularly in experiences of personal significance and "depth," but also in the subjective predispositions or preconstellations of perception, thought, memory, etc. Structure is the bearer of development and of personal identity. Besides personal structure there are social and "objective" intellectual structures. Formally, the structure of the experienced gestalt, which exists in becoming, can be compared to the "actual genesis" (or microgenesis) of the gestalt. The development of man, like that of animals, arises from qualitatively complex, pregestalt experience and is only gradually differentiated into an articulated gestalt and into rational clarification. Krueger's last work, *Die Lehre von dem Ganzen* (The doctrine of the whole; Bern, 1948), began with psychology but culminated in cosmology.

There are four main points in Krueger's philosophical psychology: holism (opposition to associationism, emotionism (or emphasis on feeling and emotion), social evolutionism, and antiphenomenalism (structural personalism). Krueger's genetic *Ganzheitspsychologie* was carried on by many of his outstanding students. Shortly after his death it was characterized as a "re-establishment of the science of the mind" in the full sense of the word, as opposing both mere introspectionism and mere behaviorism. It is the radical rejection of atomism, mechanism, sensationalism, and phenomenalism (psychologism) of traditional psychology, whose loss of credit among academic psychologists is largely due to Krueger. The slogans and basic ideas of *Ganzheitspsychologie* have also stimulated and fertilized related fields, particularly aesthetics and education.

See also Dilthey, Wilhelm; Ehrenfels, Christian Freiherr von; Emotion; Gestalt Theory; Holism and Individualism in History and Social Science; Kant, Immanuel; Latin American Philosophy; Lipps, Theodor; Neo-Kantianism; Personalism; Psychology; Wundt, Wilhelm.

Bibliography

WORKS BY KRUEGER

In English

"Consonance and Dissonance." *Journal of Philosophy, Psychology and Scientific Method* 10 (1913).

"Magical Factors in the First Development of Human Labor." *American Journal of Psychology* 24 (1913).

"The Essence of Feeling." In *Feelings and Emotions.* Worcester, MA: Clark University Press, 1928.

In German

Zur Philosophie und Psychologie der Ganzheit. Edited by E. Heuss. Berlin and Heidelberg: Springer, 1953. Writings by Krueger dating from 1918 to 1940, with a complete bibliography.

WORKS ON KRUEGER

Buss, Onko. *Die Ganzheitspsychologie Felix Kruegers.* Munich, 1934.

Odebrecht, Rudolf. *Gefühl und Gestalt: Die Ideengehalt der Psychologie Felix Kruegers.* Berlin: Junker und Dünnhaupt, 1929.

Wellek, Albert. *Das Problem des seelischen Seins: Die Strukturtheorie Felix Kruegers,* 2nd ed. Meisenheim and Vienna, 1953.

Wellek, Albert. *Die Widerherstellung der Seelwissenschaft im Lebenswerk Felix Kruegers.* Hamburg, 1950.

Albert Wellek (1967)
Translated by Tessa Byck

KUHN, THOMAS

(1922–1996)

Educated at Harvard University (SB, 1943; PhD in physics, 1949), Thomas Kuhn taught at Harvard (1951–1956), University of California, Berkeley (1956–1964), Princeton University (1964–1979), and Massachusetts Institute of Technology (1979–1991). His book *The Structure of Scientific Revolutions*, first published in 1962 (2nd. ed., 1970), continues to stimulate discussion among historians and philosophers of science even as its concepts of "paradigm" and "paradigm shift" have been adopted by a great diversity of writers, often at some remove from their source in Kuhn's book.

CONCEPTUAL SCHEMES, PARADIGMS, AND NORMAL SCIENCE

At Harvard Kuhn became the protégé of its president, James B. Conant, to whom he dedicated the first edition of *Structure of Scientific Revolutions*. Conant's concept of "conceptual scheme," applied especially to the chemical revolution's phlogiston and oxygen theories, reappeared in Kuhn's first book, *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought* (1957), and was one of the principal sources of Kuhn's all-important paradigm concept. His evolving understanding of that concept also reflected a pivotal experience in 1947, in which he suddenly appreciated that Aristotle could not properly be understood from the perspective of post-Galilean physics, but only from within Aristotle's own context of problems, concepts, and assumptions. Kuhn's early conviction that such systems of scientific thought can only be understood holistically and that a scientist's appreciation for a radically new system comes in a flash of insight underlay his notions of the incommensurability of paradigms and of the gestalt switch that marks the transition from one paradigm to another.

Kuhn announced his central problem as "the nature of science and the reasons for its special success" (1970, p. v). He forged his concept of "paradigms"—glossed here as "universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners" (p. viii)—in part as a way to understand why there is less disagreement among natural scientists over fundamentals than there is among social scientists and psychologists. Kuhn rejected the view that scientific knowledge grows incrementally through the accumulation of individual facts, laws, and theories. He linked his rejection of demarcationist issues—what distinguishes good science from error or superstition—to his insistence that superseded conceptual systems like Aristotelian dynamics and phlogistic chemistry were, in their context, no less scientific than currently accepted science.

Kuhn applied the term "normal science" to "research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice" (1970, p. 10). Paradigm-defining works like Aristotle's *Physics*, Isaac Newton's *Principia*, and Antoine Lavoisier's *Chemistry* were "sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity" and "sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve" (p. 10). Subsequent scientists (and students of science) study such works as "concrete models," whereby they become "committed to the same rules and standards for scientific practice" (p. 11). Strong commitment and broad consensus characterize the practitioners of Kuhnian normal science. The paradigm that defines that practice limits the questions worth asking and the experiments worth performing as it specifies the entities the world is composed of and the relevance of putative facts.

For Kuhn, most scientists are engaged in "mopping-up operations" resembling "an attempt to force nature into the preformed and relatively inflexible box that the

paradigm supplies" (p. 24). Kuhn likened normal science to "puzzle-solving": a solution must be assumed to exist for any problem worth addressing, and one knows ahead of time the general form the solution will take. Kuhn insisted that paradigms guide research not via rules and definitions but as models (later called "exemplars") of proper scientific practice. He associated his understanding with Michael Polanyi's concept of tacit knowledge and Ludwig Wittgenstein's notion that one can employ words without having reduced their meaning to some putative essence.

In the context of his discussion of anomalies and the emergence of scientific discoveries, Kuhn began to employ the terms "paradigm" and "paradigm change" in a broader sense closer to his and Conant's earlier "conceptual scheme," whereby his central example was the chemical revolution associated with Lavoisier's oxygen theory. Kuhn here insisted that unanticipated discoveries of new sorts of phenomena typically occur in response to the perception of anomaly with regard to the expectations of normal science. Kuhn likened scientists' response to anomalies to subjects in an experiment with playing cards who are asked to identify—among normal cards—black hearts and red spades, and who typically try unconsciously to assimilate those anomalies to the expected categories: "In science, as in the playing card experiment, novelty emerges only with difficulty, manifested by resistance, against a background provided by expectation" (1970, p. 64).

ANOMALIES, CRISES, AND PARADIGM SHIFTS

The point is of crucial importance. Anomalies enable scientists to isolate weaknesses within the dominant paradigm and to devise a solution that ultimately induces the scientific community to embrace a new and more effective paradigm. These are the "paradigm shifts" associated with the Copernican, Newtonian, and chemical revolutions. In Kuhn's view awareness of serious anomaly—always with regard to internal, technical issues, not to any of various external factors—leads to a period of crisis characterized by "the proliferation of competing articulations, the willingness to try anything, the expression of explicit discontent, the recourse to philosophy and to debate over fundamentals" (1970, p. 91)—that is, by what he termed "extraordinary science."

Although Kuhn recognized that "every problem that normal science sees as a puzzle can be seen, from another viewpoint, as a counterinstance and thus as a source of crisis" (p. 79), he offered no satisfactory explanation for why only some unsolved problems are perceived as anomalies, and why only some anomalies lead to crises. In his view no fundamental changes to a paradigm can come from the resources of normal science itself. The transition from one paradigm to another constitutes "a reconstruction of the field from new fundamentals, a reconstruction that changes some of the field's most elementary theoretical generalizations as well as many of its paradigm methods and applications" (p. 85). Kuhn likened such a paradigm shift to "a change in visual gestalt" (p. 84) and defined the associated "scientific revolutions" as "those non-cumulative developmental episodes in which an older paradigm is replaced in whole or in part by an incompatible new one" (p. 92).

INCOMMENSURABILITY AND RELATIVISM

In elaborating parallels between scientific and political revolutions, Kuhn introduced a number of ideas that would prove controversial. He argued that because they recognize no common higher authority, "the parties to a revolutionary conflict must finally resort to the techniques of mass persuasion, often including force" (1970, p. 93).

> Like the choice between competing political institutions, that between competing paradigms proves to be a choice between incompatible modes of community life. Because it has that character, the choice is not and cannot be determined merely by the evaluative procedures characteristic of normal science, for these depend in part upon a particular paradigm.… As in political revolutions, so in paradigm choice—there is no standard higher than the assent of the relevant community. (p. 94)

Such assertions led many to accuse Kuhn of making science a matter of might makes right, of mob psychology, where the techniques of political persuasion replace those of evidence and rational argument.

Because different paradigms make different ontological claims, define different problems as significant, and employ different standards of what properly belongs to science, "the normal-scientific tradition that emerges from a scientific revolution is not only incompatible but often actually incommensurable with that which has gone before" (1970, p. 103). Hence defenders of opposing paradigms, absent a shared set of values, "will inevitably talk through each other when debating the relative merits of their respective paradigms" (p. 109). Although Kuhn

resisted the charge of relativism, his position clearly relativizes scientific knowledge to the paradigm-dependent standards enforced by particular scientific communities, not to ostensibly objective experimental tests.

That implicit relativism was reinforced by Kuhn's insistence that scientists working within different paradigms see the world in profoundly different ways, that they effectively live in different worlds. Again, analogies—gestalt switches and experiments with inverting lenses and anomalous playing cards—were invoked to enhance the claim's plausibility. The transformation of vision that students undergo as they learn to read bubble-chamber photographs parallels "the shifts in scientific perception that accompany paradigm change" (1970, p. 117). The sudden and unstructured gestalt switch that accompanies a paradigm shift thrusts scientists into a world "incommensurable" with the one they had inhabited before. Kuhn's insistence that such transformations of vision are not reducible to a reinterpretation of individual stable data derived from his rejection of the possibility of a neutral observation language for science. In speaking of the "flashes of intuition through which a new paradigm is born" (p. 123), Kuhn transformed the gestalt switch from a metaphor to an operative element in the dynamics of scientific change. And in shifting the locus of conceptual change from the ostensibly objective externalities of experiment and argument to the psychological internality of a holistically unanalyzable gestalt switch, he seemed to many to undercut the epistemological legitimacy of science. Kuhn likened a revolutionary paradigm shift to a conversion experience that cannot be forced by logic and neutral experience. "Persuasion" and "conversion" are the terms that dominate Kuhn's discussion of paradigm shift.

Although he appealed to the greater problem-solving ability of postrevolution theories, Kuhn had no paradigm-independent way to define scientific progress and no way at all to address the question of the truth value of particular scientific claims. He sought to make this stance acceptable by appealing to an analogy between the historical development of science and Charles Darwin's rejection of the goal-directedness of evolution. The process by which one paradigm wins out over its competitors "is the selection by conflict within the scientific community of the fittest way to practice future science. … Successive stages in that developmental process are marked by an increase in articulation and specialization. And the entire process may have occurred, as we now suppose biological evolution did, without benefit of a set goal, a permanent fixed scientific truth, of which each stage in the development of scientific knowledge is a better exemplar" (1970, pp. 172–173). But goal-directedness is not the same thing as correspondence to the physical world, and although this may be a viable way to account for the history of science, it does not address the underlying epistemological question concerning the truth-likeness of scientific theories. In asking why scientific communities are able to reach consensus at all, Kuhn failed to assign a principal role to inputs from the physical world as he increasingly appealed to the sociology of scientific communities.

SCIENTIFIC AND LINGUISTIC COMMUNITIES

In the postscript appended to the second edition of *Structure of Scientific Revolutions*, Kuhn defended his original claims while effectively abandoning the term "paradigm." One important amplification was his appeal to the analogy between members of scientific and linguistic communities, whereby he urged "that men who hold incommensurable viewpoints be thought of as members of different language communities and that their communication problems be analyzed as problems of translation" (1970, p. 175). Like the acceptance of a new paradigm, Kuhn saw the transition accompanying translation into a new language as a qualitatively discontinuous conversion experience: "The conversion experience that I have likened to a gestalt switch remains, therefore, at the heart of the revolutionary process" (p. 204).

See also Aristotle; Galileo Galilei; Lavoisier, Antoine; Newton, Isaac; Paradigm-Case Argument; Philosophy of Science, Problems of; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY KUHN

The Copernican Revolution: Planetary Astronomy in the Development of Western Thought. Cambridge, MA: Harvard University Press, 1957. Rev. ed., 1979.

The Structure of Scientific Revolutions. 1962. 2nd ed. with postscript. Chicago: University of Chicago Press, 1970. 3rd ed. with index, 1996.

The Essential Tension: Selected Studies in Scientific Tradition and Change. Chicago: University of Chicago Press, 1977.

The Road since Structure: Philosophical Essays, 1970–1993, with an Autobiographical Interview, edited by James Conant and John Haugeland. Chicago: University of Chicago Press, 2000. Contains a complete bibliography of Kuhn's publications.

WORKS ABOUT KUHN

Barnes, Barry. *T. S. Kuhn and Social Science*. New York: Columbia University Press, 1982.

Buchwald, Jed Z., and George E. Smith. "Thomas S. Kuhn, 1922–1996." *Philosophy of Science* 64 (1997): 361–376.

Caneva, Kenneth L. "Possible Kuhns in the History of Science: Anomalies of Incommensurable Paradigms." *Studies in History and Philosophy of Science* 31 (1) (March 2000): 87–124.

Gutting, Gary, ed. *Paradigms and Revolutions: Appraisals and Applications of Thomas Kuhn's Philosophy of Science*. Notre Dame, IN: University of Notre Dame Press, 1980.

Hoyningen-Huene, Paul. *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Translated by Alexander T. Levine. Chicago: University of Chicago Press, 1993.

Lakatos, Imre, and Alan Musgrave, eds. *Criticism and the Growth of Knowledge*. Cambridge, U.K.: Cambridge University Press, 1970.

Nickles, Thomas, ed. *Thomas Kuhn*. Cambridge, U.K.: Cambridge University Press, 2003. Contains a short bibliography of important studies on Kuhn and his influence.

Kenneth L. Caneva (2005)

KÜLPE, OSWALD

(1862–1915)

A German psychologist, philosopher, and historian of philosophy, Oswald Külpe was born in Kandava, Latvia. After teaching history, Külpe entered the University of Leipzig in 1881, intending to continue in history. However, the lectures of Wilhelm Wundt stimulated his interest in philosophy and psychology, and after further studies in Berlin, Göttingen, and Dorpat (Russia), he returned to Wundt's seminar in 1886, receiving his doctorate the following year. In 1894 he was appointed extraordinary professor at Leipzig but left to accept a full professorship at Würzburg, where he founded a psychological laboratory. Külpe returned to Leipzig in 1896, and he subsequently held academic positions at Bonn and Munich. Primarily because of his work in organizing experimental laboratories, Külpe is regarded as a pioneer of experimental psychology in Germany. He died in Munich during World War I of influenza contracted while visiting wounded German soldiers.

PSYCHOLOGY AND EPISTEMOLOGY

Külpe's philosophical position, a form of critical realism, was closely related to his work in psychology. He came to regard the positivistic attempts of Ernst Mach and Richard Avenarius to reduce mental processes to sensations as incapable of accounting for the findings of introspective experiments. In one series of experiments, Külpe presented cards with nonsense syllables of varying colors and arrangements to subjects who were asked to report either the color, pattern, or number of items seen. Each person abstracted the features he had been instructed to report, remaining unconscious of the other features of the cards. Külpe concluded that the process of abstraction depends not only on the material presented to sensation but also on the subject's apprehension. This was taken to prove that sensations—as well as physical phenomena—must be distinguished from their apprehension. Thus he questioned the equation of "being" with "being perceived," even at the level of sensation.

Külpe abandoned the sensationalist psychology of contents in favor of a psychology recognizing both contents and acts of mind. Abstraction, he maintained, is a mental act or function that cannot be directly observed, but its occurrence is undeniable, even though it is discoverable only retrospectively. There exist both thought contents (*Gedanken*) and thought processes (*Denken*). The latter include the impalpable acts of thinking, meaning, and judging, which are not merely relations among contents but activities of the ego that transform the actualities (*Wirklichkeiten*) of consciousness into realities (*Realitäten*).

Külpe's position was thus hostile to both naive realism and idealism. Against the former, he argued that thought, although it does not produce the object of knowledge, is nevertheless genuinely spontaneous and creative in contributing to the realization of the object. His argument against idealism held that the facts of conscious experience require the existence of independent objects. When a scientist studies the maturation of an egg, for example, he assumes that this process takes place while no consciousness is directed upon it. Such continuity of development implies the object's independence of its being thought, a presupposition of every science.

Külpe used the word *awareness* (*Bewusstheit*) to indicate that the meanings of abstract words can be discovered in consciousness even when only the words themselves are perceivable entities. This thesis is an application of the theory that there exist impalpable (*unanschaulich*) or imageless contents of consciousness, a theory for which Külpe's "Würzburg school" of psychology was noted. Meanings can be experienced and objectified even without words or other signs. Although we cannot analyze precisely how these contents are given, retrospective acts make the world of meanings accessible to us. Külpe's indebtedness to Edmund Husserl and Franz

Brentano is evident. Mental acts provide knowledge of meanings, and the act of meaning (*das Meinen*) may be directed even to such objects as God, the soul, electrons, or atoms, which could not possibly be actualized in consciousness. The capacity for imageless thought is essential if thought is to relate itself to something independent of it. When one wants to imagine a certain structure, the particular image one has in mind is only representative of the structure; the image points beyond itself or is the occasion for such an intentional act.

AESTHETICS

In aesthetics, Külpe attempted to support Gustav Fechner's results concerning the golden section. Like Wundt, he maintained that the aesthetic pleasure produced by ideally proportioned objects results from mental economy. When the ratio of a whole to its larger part is the same as that of the larger to the smaller part, the perception involves the least effort combined with the greatest possible diversity.

Külpe attempted to further the development of experimental aesthetics by such methods as asking people to record their reactions to glimpses of slides showing works of art. His findings indicated no sympathetic empathy on the part of his subjects, thus opposing the contention of Theodor Lipps that such empathy (*Einfühlung*) is the basic condition of all aesthetic enjoyment. In the reports of his subjects Külpe found that form, orderliness, symmetry, and harmony were related to attractiveness. However, he recognized the limited validity of his findings, admitting that aesthetically inexperienced people might respond differently than his subjects. This reluctance to claim more for a theory than was warranted by experimental findings was characteristic of Külpe's work in psychology.

See also Aesthetics, History of; Avenarius, Richard; Brentano, Franz; Critical Realism; Fechner, Gustav Theodor; Husserl, Edmund; Idealism; Lipps, Theodor; Mach, Ernst; Psychology; Wundt, Wilhelm.

Bibliography

WORKS BY KÜLPE

Grundriss der Psychologie auf experimenteller Grundlage dargestellt. Leipzig: Engelmann, 1893. Translated by E. B. Titchener as *Outlines of Psychology, Based upon the Results of Experimental Investigation.* New York: Macmillan, 1895.

Einleitung in die Philosophie. Leipzig, 1895. Translated by W. B. Pillsbury and E. B. Titchener as *Introduction to Philosophy.* London, 1910.

Die Philosophie der Gegenwart in Deutschland. Leipzig: Teubner, 1902. Translated by M. L. Patrick and G. T. W. Patrick as *The Philosophy of the Present in Germany.* London: G. Allen, 1913.

Immanuel Kant. Leipzig, 1907.

Die Realisierung. Vol. I, Leipzig, 1912; Vols. II and III, edited by August Messer, Leipzig, 1920 and 1923.

Ethik und der Krieg. Leipzig: Hirzel, 1915. Külpe's defense of Germany's position in World War I.

Grundlagen der Aesthetik. Edited by S. Behn. Leipzig, 1921.

WORKS ON KÜLPE

Messer, August. *Der Kritische Realismus.* Karlsruhe, 1923.

Ogden, R. M. "Oswald Külpe and the Würzburg School." *American Journal of Psychology* 61 (1951): 4–19.

Arnulf Zweig (1967)

KUMAZAWA BANZAN

(1619–1691)

Kumazawa Banzan, a Japanese Confucianist of the Wang Yangming school, was born in Kyoto and died at Koga, Shimoda prefecture. Both he and his father were masterless samurai. Deciding to become a scholar, Kumazawa went to Nakae Tōju (1608–1648); in 1642 Nakae taught him the doctrine of Wang Yangming (in Japanese, Ōyōmei)—"innate knowledge" and cultivation of the mind. Kumazawa entered the service of Lord Ikeda Mitsumasa of Okayama, but his ideas, contrasting with the officially established doctrine, Zhu Xi neo-Confucianism, aroused suspicion. However, his character and practical ability were recognized, and Ikeda put him in charge of the fief. For seven years (1649–1656) he successfully brought forth administrative reforms that transformed Okayama into a model fief. Paramount among his accomplishments was his role in organizing the Okayama college. Yet the extreme nature of these reforms, even in monasteries, angered many. Moreover, there were rebellious samurai among his pupils. He decided to retire to the studious life of a teacher in Kyoto, but slander of his teaching forced him to move in 1667; he did pass eight quiet years (1679–1687) at Yadasan near Kōriyama. On the official request of the Tokugawa government, he presented a plan of reform (possibly in his *Daigaku wakumon*). Thereupon his enemies, especially Hayashi, the defender of Zhu Xi Confucianism, succeeded in having him confined at Koga.

Kumazawa is typical of the early Tokugawa nonconformists, who were beset by adversities that multiplied with success. His politico-economic ideas, which were indeed very bold for his times, were the real reason for his

difficulties. They are expressed in *Daigaku wakumon* (Some questions concerning the great learning), which is not a commentary on the Confucian classic "The Great Learning" but rather a tract on many subjects concerning how to rule the realm according to the Confucian precept of *jinsei,* or "benevolent rule." Both his unconventional proposals and his pragmatic attitude toward doctrine are striking.

See also Chinese Philosophy; Hayashi Razan; Japanese Philosophy; Nakae Tōju; Wang Yang-ming; Zhu Xi (Chu Hsi).

Bibliography

For a guide to primary sources, see the bibliography to the Japanese Philosophy entry. See also M. Fisher, "Kumazawa B., His Life and Ideas," *Transactions of the Asiatic Society of Japan,* 2nd series, 16 (1938): 221–259; 259–356. W. T. de Bary, ed., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 384–392.

Gino K. Piovesana, S.J. (1967)

KUNG-SUN LUNG

See *Gongsun Long*

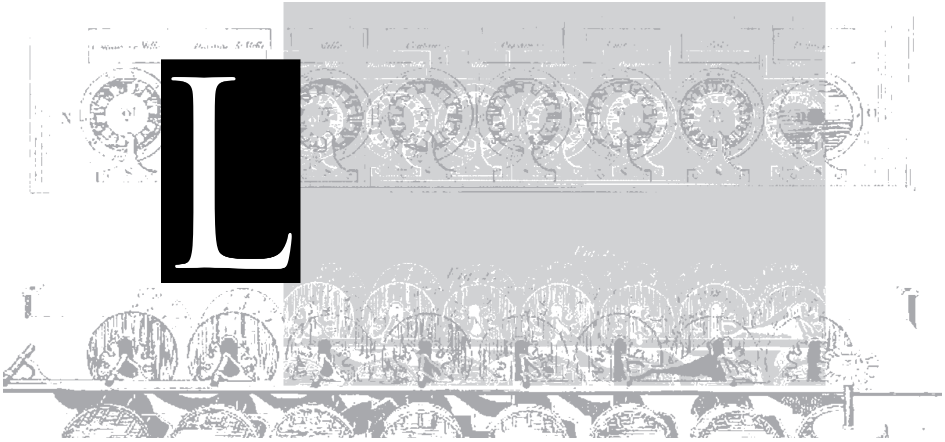

L

LAAS, ERNST

(1837–1885)

Ernst Laas, the German philosopher, was born in Förstenwalde. From 1872 on, he was professor in Strasbourg. His first important book, *Kants Analogien der Erfahrung* (Berlin, 1876), was a critical study both of Immanuel Kant and of "the foundations of theoretical philosophy"; but in his main work, *Idealismus und Positivismus* (3 vols., Berlin, 1879–1884), he launched a general attack on idealism, including Aristotle, René Descartes, Gottfried Wilhelm Leibniz, and especially Plato as its founder, as well as Kant. His purpose was to provide a remedy for the "discontinuity of philosophy"; that is, its failure to make progress over the centuries and its want of any clear standards. The remedy lay first of all in a new critical approach to the history of philosophy, which in the past had usually been at best merely scholarly and accurate. This new analysis revealed a basic dualism throughout the history of philosophy between the outlooks of Plato and Protagoras; and this revelation, in turn, permitted a revision of the judgment rendered in favor of Plato that had ever since benefited his followers at the expense of their opponents, such as the British empiricists. Laas referred specifically to J. S. Mill and cited approvingly a review of his own book on Kant that had compared it to Mill's *Examination of Sir William Hamilton's Philosophy.*

By "positivism" Laas meant, as was usual in Germany at the time, the tradition of Protagoras and the British empiricists, not the doctrine of Auguste Comte, whom Laas mentioned rarely and with little sympathy. Laas's position might more accurately, especially in English usage, be called neo-empiricism. It proposed to limit knowledge to the data of sense experience, thereby denying both a consciousness independent of the content of perception (insisting on the correlation of subject and object) and objects independent of the process of perception (asserting the instant changeability of objects of perception). At the same time Laas avoided the conclusions drawn by some empiricists, such as George Berkeley, by rejecting any version of subjective idealism (which would assert the superiority or exclusive reality of the perceiver vis-à-vis the objects of perception or sensation) even more vehemently than he rejected the objective idealism originated by Plato. He identified this idealistic tradition in logic with conceptual realism, in epistemology with a priori deductive rationalism, and in metaphysics with both spontaneous human creativity and superhuman teleology. He associated idealism with a mathematically

inspired desire to attain to the knowledge of absolutes and with the doctrines of innate ideas and final causes.

However, in his anxiety to escape from the "monstrous" notions of subjective idealism, as well as from "skepticism," "frivolity," and the "banal philosophy of common sense," Laas came close to a neo-Kantian position in postulating an ideal or total consciousness. Recognizing, with Mill, that the sum total of actual objects of sensation is insufficient to construct an intelligible world, he asserted that the world consists of the sum total of possible contents of perception, which would be vouchsafed to an ideal consciousness and which it is the task of philosophy to construct. Since facts (objects) exist independently of consciousness (although not of perception), including this ideal consciousness, Laas claimed in this way to have saved the possibility of scientific investigation of the physical world from "skepticism," even though that world is relative and variable.

Just as he quite openly sided even with idealism (particularly with Kant, whom he often cited sympathetically) rather than with epistemological skepticism, Laas also seeks to defend his ethical doctrine (mainly in Vol. II of *Idealismus und Positivismus*) against any imputation of relying on egoism. Here again, however, his main concern was to overcome what he saw as the Platonic tradition of asceticism founded on a set of absolute and transcendental ideals. For this he proposed to substitute a "positive" ethics for this world, based on its values as revealed by "enlightened self-interest." Laas acknowledged the founders of this ethical doctrine to be Epicurus, Claude-Adrien Helvétius, and Jeremy Bentham, but he diverged from them on the crucial point of egoism. He denied the identification of self-interest with egoism and held, rather, that self-interest dictates the performance of duties and the fulfillment of demands and expectations imposed on the individual by his environment. In this way, ethical values are the consequences of a particular social order. They acquire validity when they are judged, in the long run and by a considerable number of people, to be worthwhile. Laas characteristically listed as ethically desirable values security of employment, social harmony, the laws and institutions of the state, and cultural progress. These ethical teachings were the most influential part of his philosophy, affecting, in particular, the ideas of Theobald Ziegler and Friedrich Jodl.

See also Aristotle; Bentham, Jeremy; Berkeley, George; Comte, Auguste; Descartes, René; Empiricism; Epicurus; Ethical Egoism; Ethics, History of; Helvétius, Claude-Adrien; Idealism; Innate Ideas; Jodl, Friedrich; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Mill, John Stuart; Plato; Positivism; Protagoras of Abdera; Realism; Teleology.

Bibliography

In addition to the works cited in text, see Laas's *Literarischer Nachlass,* edited by B. Kerry (Vienna: Deutschen Worte, 1887).

For material on Laas, see Nikolaus Koch, *Das Verhältniss der Erkenntnistheorie von Ernst Laas zu Kant* (Würzburg, 1940) and Ludwig Salamonowicz, *Die Ethik des Positivismus nach Ernst Laas* (Berlin: H. Lücker, 1935).

W. M. Simon (1967)

LABERTHONNIÈRE, LUCIEN

(1860–1932)

Lucien Laberthonnière, the French philosopher of religion and a leading figure in the modernist movement in the Roman Catholic Church, was born at Chazelet (Indre). He studied for the priesthood and was ordained as an Oratorian in 1886. He then taught in various institutions, mainly in the college at Juilly, where he became rector in 1900. Laberthonnière was influenced by philosophies of life and action; he mentions Maine de Biran and Étienne Boutroux as the two philosophers who had most impressed him. Maurice Blondel's philosophy of action was another important formative factor, although Laberthonnière later found it moving too far toward intellectualism. He himself not only advocated a pragmatic point of view but also had an intense distaste for intellectualism and speculative philosophy. In particular, he had no sympathy for the attempted Thomist synthesis of faith and reason, believing that the task is not to conciliate these two but to choose between them. His teachings brought him into conflict with ecclesiastical authorities, and his principal writings were put on the Index in 1906. In 1913 he was prohibited from further publication.

Laberthonnière was not concerned with merely speculative philosophy that is constructed apart from life. He believed that the purpose of all philosophy is to give sense to life, and this motivation underlies even metaphysics, whether or not the metaphysician is aware of it. In the long run, the test of a philosophy must be its viability or its aptness for life, and the criterion of philosophical truth is a pragmatic one. We mistake the character of philosophy if we think of it as a theoretical enterprise resulting in a system of propositions linked together by abstract

logical principles. A philosophical doctrine has a moral as well as an intellectual character, so that a worthwhile philosophy has to be worked out by living. The test of its truth is whether it can be illuminating when brought to bear on the problems of life.

Although Laberthonnière apparently held that all philosophy has a pragmatic or existential motivation, even if this remains unconscious, he also believed that some philosophies have been much more successful than others in relating to life. The theme of one of his principal writings, *Le réalisme chrétien et l'idéalisme grec* (Paris, 1904), is the contrast between two supposedly extreme cases, Greek philosophy and Christian thought. Greek philosophy was concerned with abstract essences, conceived God as static and immutable, and proposed the life of pure contemplation as its ideal for man. In contrast to such idealism or intellectualism, Christianity is presented as a realism. Its concern is with the concrete life of action, and God himself is conceived as active, the living God of the Bible. Hence, the truth of Christianity cannot be reached by intellectual contemplation, as if it were something external to us. Such truth as Christianity teaches is concrete and intrinsic to life, so that we grasp it only in living and in re-creating this truth in ourselves. These ideas about religious truth had already found expression in Laberthonnière's *Essais de philosophie religieuse* (Paris, 1903), where it is maintained that the doctrines of religion are to be understood not as general truths of the same kind as scientific truths but as concrete truths that must be brought into experience and realized if we are to understand them and know their value.

Although these views lean strongly toward pragmatism, Laberthonnière did not think that religion could be reduced to a purely practical affair or that it could be adequately explicated in naturalistic terms. It is significant that in spite of the harsh treatment that he received from the Roman Catholic Church, he remained devoted to it and believed his philosophical views to be compatible with its teaching. If he went far toward abolishing the traditional distinction between the natural and the supernatural, this is not to be understood as the reduction of the latter to the former. Rather, it was Laberthonnière's conviction that the natural is itself already permeated by divine grace. Thus, we should look for God not in some upper or outer realm but in the immediate world, where he is active, and especially in the depth of human life itself.

See also Blondel, Maurice; Idealism; Maine de Biran; Modernism; Pragmatism; Realism; Thomism.

Bibliography

In addition to the writings mentioned above, Laberthonnière published *Positivisme et catholicisme* (Paris, 1911). His other work, including an important study of Descartes, appeared posthumously in *Oeuvres,* 2 vols. (Paris, 1948–1955).

See also M.-M. d'Hendecourt, "Laberthonnière," in *Revue de métaphysique et de morale* 63 (1961).

John Macquarrie (1967)

LABRIOLA, ANTONIO
(1843–1904)

Antonio Labriola, professor of philosophy in Rome from 1874 to 1904, was the first Italian Marxist philosopher. He wrote little, but that little was widely publicized by two disciples, Georges Sorel and Benedetto Croce; he exercised his extensive influence through lectures and discussions. Trained as a Hegelian in Naples, he became a Herbartian, more interested in Johan Friedrich Herbart's ethics and pedagogy than in his metaphysics. He discovered Marxism around 1890 and began a correspondence with Friedrich Engels that lasted until the latter's death and was published in *Lettere a Engels* (Rome, 1949). This discovery of Marxism was a decisive event in Italian intellectual life, for from it dates the introduction of Marxist theory into Italy's academic culture, where it still occupies a prominent place.

Labriola's articles on Marxism, published in Italy by Croce and in France by Sorel, were first collected in French, as *Essais sur la conception matérialiste de l'histoire* (Paris, 1897). Their publication established Labriola's international reputation as an expositor of Marxism. He wrote Sorel ten letters on the subject, published as *Discorrendo di socialismo e di filosofia* (Rome, 1897). These books were the first exposition of Marxism as an independent philosophy to be made by an academic philosopher. They have been widely used in later efforts to combat all varieties of philosophical revisionism, whether from neo-Kantian or positivist sources. The "return to Labriola," as recommended by Antonio Gramsci and as undertaken in Italy since 1950, has meant going back to the original innocence of a supposedly pure and independent Marxist philosophy, for Labriola claimed not to be an original thinker, and even less to be interested in developing or criticizing Marxism. He wanted to be simply an expositor and systematizer of a philosophy implicit in Karl Marx's work.

The philosophy he found in Marx's work closely resembled the Hegelian views that Labriola had defended in controversies with neo-Kantians before he had heard

of Marx. For example, he held that scientific socialism is not subjective criticism applied to things, but the statement of the self-criticism that is in things themselves. The only criticism of society is society itself, for there is an objective dialectic immanent in history, which progresses by contradictions. Socialism was no longer an aspiration or project (a view soon to be revived by neo-Kantian revisionists); it was the inevitable result of current contradictions in capitalist society. Labriola stressed the "scientific, objective" status of these assertions, in contrast to mere philosophies of history, which he dismissed as ideology. Historical materialism was no philosophy, but simply a method of research, a guiding thread like the Darwinian hypothesis.

Labriola, Croce, and Sorel were nicknamed the Holy Trinity of Latin Marxism, but the Roman professor came to feel that his spiritual sons were "going too far" in their development and criticism of the doctrine. They lacked that inflexible orthodoxy of which Labriola is the first eminent example in the Marxist tradition, and they touched off the revisionist controversy. That dispute broke out simultaneously in several countries, although Croce gave priority to his own and Sorel's writings. At all events, Eduard Bernstein in Germany, Sorel in France, Croce and Saverio Merlino in Italy, T. G. Masaryk in Prague, and the Fabians in England drew freely on each other's work, and Labriola found himself being quoted by and confounded with the "heretics." In a celebrated dispute, he broke publicly with Croce and Sorel, saying that revisionism was an international conspiracy organized by "scientific police-spies"—perhaps the first appearance of a philosophical terminology that was to become familiar later. Labriola never wrote on Marxism again. His earlier minor works, which include a *Socrate,* have been published by Croce (Bari, 1909) but are of small importance.

See also Continental Philosophy; Croce, Benedetto; Engels, Friedrich; Gramsci, Antonio; Herbart, Johann Friedrich; Historical Materialism; Marx, Karl; Marxist Philosophy; Masaryk, Tomáš Garrigue; Neo-Kantianism; Sorel, Georges.

Bibliography

WORKS BY LABRIOLA

Essays on the Materialistic Conception of History. Translated by C. H. Kerr. Chicago: Kerr, 1904.

Socialism and Philosophy. Translated by E. Untermann. Chicago: Kerr, 1907.

Opere, 3 vols. Edited by L. Dal Pane. Milan: Feltrinelli, 1959.

WORKS ON LABRIOLA

Bellamy, Richard. "Antonio Labriola." In *Modern Italian Social Theory: Ideology and Politics from Pareto to the Present.* Stanford, CA: Stanford University Press, 1987.

Bruzzo, S. *Il pensiero di Antonio Labriola.* Bari, 1942.

Dal Pane, L. *Antonio Labriola: la vita e il pensiero.* Rome, 1935.

Diambrini Palazzi, S. *Il pensiero filosofico di Antonio Labriola.* Bologna: Zanichelli, 1923.

Kolakowski, Leszek. "Antonio Labriola: An Attempt at an Open Orthodoxy." In *Main Currents of Marxism,* Vol. 2. Translated by P. S. Falla. Oxford: Oxford University Press, 1982.

Plekhanov, G. "The Materialist Conception of History." *Novoye Slovo* (September 1897). Also published separately. New York: International, 1940. Originally written as a review of Labriola's *Essays.*

Neil McInnes (1967)
Bibliography updated by Philip Reed (2005)

LA BRUYÈRE, JEAN DE

(1645–1696)

Jean de La Bruyère, the French author and moralist, was born in Paris, the son of a city official. After some legal training he apparently fell on hard times, but through the influence of Jacques Bénigne Bossuet he was appointed tutor to the grandson of the great Condé in 1684. After his tutorial functions were ended, he stayed on as librarian. The family seems to have been unpleasant; his colleagues, uncongenial; and the humiliations inflicted on him in this aristocratic society left a lasting mark. Elected to the Academy in 1693 after several unsuccessful attempts, he led a lonely and somewhat frustrated life, never marrying, making few friends, but showing passionate loyalty to those who, like Bossuet, won his respect.

La Bruyère's one famous work, the *Caractères*, reflects his personal experiences. Ostensibly modeled on the Greek *Characters* of Theophrastus, which La Bruyère translated and published in the same volume, the *Caractères* owes more to the quite different genre of La Rochefoucauld's *Maximes* and to the work of such contemporary moralists as Blaise Pascal and the Chevalier Antoine Gombault de Méré. Fifteen chapters somewhat arbitrarily group together epigrams (although La Bruyère explicitly disclaimed any intention of producing anything so authoritative as maxims), extended pen portraits (readily, and often wrongly, identified with living people) and brief moral essays, all arranged to cover, with considerable overlapping, the main characteristics and activities of contemporary society, from literary criticism to money lending, from sex to sermons. The last chapter, which, La Bruyère implausibly claimed, constituted the purpose

and culmination of the previous fifteen, is devoted to a defense of religion against the freethinkers. It combines in an agreeable rather than a compelling manner the stock arguments for God's existence from his visible effects in nature with others reminiscent of Pascal and drawn from human psychology. The length of the book more than doubled in the course of nine editions from 1688 to 1696, and it came to include more and more of the concrete and detailed description, based on acute observation and couched in brilliant style, which makes La Bruyère at once a distinctive and a distinguished author.

In La Bruyère's time the splendors of Louis XIV's reign had come to demand too high a price, both economically and morally, of those obliged to maintain it. La Bruyère, a bourgeois himself, soured by personal experience of aristocratic arrogance and temperamentally allergic to worldly frivolity, was unsparing in his criticism of the court, where methodical hypocrisy marked the lives of those enslaved by self-interest and the desire for royal favor.

Like Bossuet, his hero and patron, La Bruyère felt able to combine vehement attacks on social abuses, due certainly in fact (if not in theory) to royal absolutism as currently practiced, with fulsome eulogy of Louis himself, going so far as to assimilate respect for the prince to fear of God. A convinced Christian, he had a genuine social conscience, as is illustrated by his famous remarks about the pitiful condition of the peasants. He contrasted the elegant heartlessness of the nobles with the rough kindliness of the people, with whom, in the last analysis, he would wish to be classed. He was, however, neither egalitarian nor republican, but believed that inequality founded on order is divinely instituted; and it was on moral and religious grounds, not in the name of equality, that he dissociated himself from a society he regarded as irremediably corrupt.

In common with other moralists of the age, La Bruyère was fascinated by the discrepancy between appearance and reality in human behavior. He recorded how skill in playing the social game usurps the name and place of virtue, how fashion makes mock of convictions (a happily married couple finds it socially expedient to simulate infidelity), and how self-interest is the one constant motive of those who disguise it so ingeniously. He was, however, gloomy rather than hopeless about human nature, and did not despair of the potential goodness of men as yet uncontaminated by society. He also believed in the possibility of satisfactory human relationships, speaking with attractive warmth of love and friendship.

Moderate as well as modest, La Bruyère was saved by common sense from the clever cynicism that is purely destructive, and his work is characterized by a positive and humane quality underlying the bitterest criticism. Although the *Caractères* falls short of absolute greatness, it reflects with exceptional accuracy the wane of the *grand siècle*.

See also Appearance and Reality; Bossuet, Jacques Bénigne; Continental Philosophy; La Rochefoucauld, Duc François de; Moral Epistemology; Pascal, Blaise; Theophrastus.

Bibliography

La Bruyère's *Les caractères*, edited by G. Servois, appeared in 3 vols. (Paris, 1865–1882) and was also issued by Bibliothèque de la Pléiade (Paris, 1934). It was translated by H. van Laun as *The Characters* (Oxford, 1963).

Works on La Bruyère are A. Adam, *Histoire de la littérature française au XVIIe siècle* (Paris, 1956), Vol. V, Ch. 6; M. Lange, *La Bruyère critique des conditions et des institutions sociales* (Paris, 1909); P. Richard, *La Bruyère et ses Caractères* (Amiens, 1946); F. Tavera, *L'idéal moral et l'idée religieuse dans les Caractères* (Paris: Mellottée, 1941); Edward C. Knox, *Jean de la Bruyère* (New York: Twayne Publishers, 1973).

A. J. Krailsheimer (1967)
Bibliography updated by Tamra Frei (2005)

LACAN, JACQUES
(1901–1981)

Jacques Lacan is undoubtedly the most philosophical of psychoanalytic authors. He developed his psychoanalytic theory of subjectivity—as a ferocious critique of the modern metaphysical tradition—in direct dialogue with a number of major philosophical figures: Descartes, Kant, Heidegger, and many others.

Lacan never had any formal philosophical training. After studying medicine and psychiatry, he got involved in the surrealist movement in the early 1930s. Along with Sartre and Bataille, he participated in Alexandre Kojève's famous seminars on Hegel's *Phenomenology of Spirit*at the Ecole des Hautes Etudes en Sciences Sociales. Lacan joined the Société Psychanalytique de Paris in 1936. Both his theories—specifically his critique of ego psychology, which he carried out under the label of a "return to Freud"—and his practice of short psychoanalytic sessions caused discord within the French and the international psychoanalytic movement in the fifties. As a result of this rift, Lacan and his followers founded the Société française de psychanalyse in 1956 and later the Ecole freudienne in

1963. In the beginning of the fifties, Lacan also started to give seminars in Paris that not only attracted psychoanalysts but also a great number of philosophers such as Jean Hyppolite and Paul Ricoeur. In this way, psychoanalysis became a central force within French philosophical thinking of the second half of the twentieth century.

Lacan's theory of the mirror stage is his first original contribution to psychoanalytic thinking. This theory was first formulated at a conference in Marienbad in 1936. Although it is a reformulation of Freud's theory of narcissism, it has important consequences for the philosophical reflection on the status of the subject. Indeed, according to Lacan, the ego is an effect of an identification with an image (paradigmatically the mirror image) that represents an ideal of unity and completeness and that is not the ego itself: "*Je est un autre*." The ego is thus characterised by an alienation that cannot be undone. It gains access to itself only through the image of the other. In the mirror stage—and in all "imaginary" relations that function according to the same logic—the ego misrecognizes its difference from the image/ideal with which it identifies itself and of which it believes that it expresses its very essence.

Lacan's work of the 1930s and 1940s mainly consists of a detailed exploration of the characteristics and the dynamics of the mirror stage and the realm of the imaginary that is characterized by it. In this context, he specifically focuses on typical forms of human aggression. Human aggression is not primarily an effect of the frustration of vital needs. Indeed, since the ego structurally misrecognizes its difference from the image/ideal of the other with which it identifies itself, the latter also inevitably appears as an usurper that provokes aggressiveness. S/he indeed appears in the process at a place that seems to be rightfully mine. I desire what s/he desires because, on the basis of the identification, I am what s/he is. As a consequence, this desire is intrinsically conflictual. Lacan often refers in this context to Saint Augustine, who describes a scene in which a well-fed infant expresses uncontrollable anger at the sight of his baby brother being breastfed. This is a clear illustration of one of the meanings of Lacan's famous dictum that "desire is the desire of the other."

This intrinsic link between the mirror stage and human aggression explains why Lacan thinks of the former as an impasse that has to be overcome. The emergence of structuralism in the early fifties, and more particularly the publication of Levi-Strauss's *The Elementary Structures of Kinship* in 1949, allowed Lacan to explain once and for all how overcoming this impasse is possible. He now claimed that the symbolic order—the order of language and of the law—precedes and dominates the imaginary that is structured by it. Hence, the identification with the mirror image is only possible on the basis of a symbolic point of reference: "Look, that image in the mirror, that is little Jimmy."

Whereas in the thirties and forties Lacan mainly studied the dynamics of imaginary relations, during the fifties he focused on the relation between human beings and the symbolic order that he calls "the Other." Lacan turns to Hegel's idea that "the word is the murder of the thing." Entry into the symbolic order implies a loss of immediacy that desire tries to undo. This desire is essentially dependent on the symbolic order through which it takes shape. Humans desire in accordance with the symbolic systems in which they are born. Lacan shows, for instance, how the inability to write of one of his patients was linked to his youth in a Muslim country. When he was small, his father was accused of theft and, according to Islamic law, the hands of a thief should be cut off. This illustrates the second meaning of Lacan's dictum, "Desire is the desire of the Other." Here "the Other" indeed refers to the symbolic system—in the case of Lacan's patient: Islamic law—in which the subject participates without realizing its impact.

In the early 1960s, Lacan shifted his attention from the imaginary and the symbolic to the Real and the object a. Language consists, according to Lacan, of differentially determined signifiers whose meaning is completely dependent on the context in which they are used. Because there is no ultimate context that would end the production of meaning once and for all, the loss of immediacy can never be overcome or "sublated" in an ultimate synthesis. Something is irremediably lost and cannot be recuperated into the order of meaning (the imaginary and the symbolic). This is what Lacan calls the Real. This notion is intrinsically linked to Lacan's theory of the object a that is the cause (and not the telos) of desire. Examples of objects a include Freudian part-objects such as breast and feces as well as the voice and the gaze, which are paradigmatic examples of the object a, according to Lacan. The object a is a (dis)incarnation of the lack that causes desire: it gives the lack a bodily determination, on the one hand; at the same time, however, these objects cannot be grasped in the phenomenal world (when we reach for the gaze, we touch … the eye). In this way, they refer to the infinite character of human desire.

From the early 1960s onward, Lacan became more and more interested in topological figures like Borommean knots or rings. He believed that they could be used

to articulate the fundamental structures of human subjectivity. Lacan died in 1981 in Paris.

See also Psychoanalysis.

Bibliography

WORKS BY JACQUES LACAN

Ecrits: A Selection. Translated by Bruce Fink. New York and London: Norton, 2004.

The Ethics of Psychoanalysis. Translated by Dennis Porter. New York and London: Norton, 1997.

The Four Fundamental Concepts of Psychoanalysis (The Seminar of Jacques Lacan, Book 11). Translated by A. Sheridan. New York and London: Norton, 1998.

On Feminine Sexuality, the Limits of Love and Knowledge: The Seminar of Jacques Lacan, Book XX, Encore. Translated by Bruce Fink. New York and London: Norton, 1999.

WORKS ON JACQUES LACAN

Borch-Jacobsen, M. *Lacan: The Absolute Master*. Palo Alto: Stanford University Press, 1991.

Fink, Bruce. *Lacan to the Letter: Reading* Ecrits *Closely*. Minneapolis: University of Minnesota Press, 2004.

Rabaté, J. M. *The Cambridge Companion to Lacan*. Cambridge University Press, 2003.

Van Haute, Philippe. *Against Adaptation: Jacques Lacan's "Subversion" of the Subject. A Comment*. New York: Other Press, 2002.

Philippe van Haute (2005)

LACHELIER, JULES

(1832–1918)

Jules Lachelier, the French idealist, was born at Fontainebleau and studied at the École Normale Supérieure in Paris. He received his *docteur ès lettres* in 1871 and held various professorial and administrative positions in the French educational system until his retirement from the post of *inspecteur général* in 1900. Lachelier joined with his teacher Jean Gaspard Félix Ravaisson-Mollien in founding the neospiritualist movement in French philosophy, a movement opposed to what seemed to be the naive acceptance of science and the scientific attitude in all phases of life. Among those who have acknowledged Lachelier's influence are Émile Boutroux, Victor Brochard, Jules Lagneau, and Henri Bergson.

Lachelier advanced a number of skeptical arguments that tend to reduce objects to phenomena, phenomena to sensations, and, more generally, to resolve the external world into thought. Nevertheless, he retained the conviction that we live in a common, objective world. Accordingly, his philosophy is directed toward the conclusion that the objectivity of our knowledge and experience is derived from mind. He summarized his idealistic philosophy as the discovery of "a thought which does not think, suspended from a thought which thinks itself."

To avoid the pitfalls of both the empiricism and the spiritualism of his day, Lachelier attempted to provide a basis for induction in a philosophy of nature. His procedure consisted of a Kantian reflection upon the necessary conditions for the existence of the world as we know it. He began by observing that, if knowledge is to be possible, sensations must exhibit the same unities that are found in phenomena. By eliminating competing hypotheses, he found that the unifying element within any phenomenon, as well as the unifying element among phenomena, is established by the necessary relations operative in them and is expressed by the law of efficient causes. The necessity of this law cannot be discovered in sensations alone, in phenomena as such, or in their mere juxtaposition; nor can it be isolated in any locus from which mind is separated. It must be regarded, rather, as a kind of unconscious but logical thought diffused throughout nature. The mechanical linkages among events in nature reflect the logical relations in thought. Lachelier concluded that the unity of thought and the formal unity of nature are inverses of each other.

Given a series of phenomena, the law of efficient cause is sufficient to account for their organization in a mechanically interrelated series. But the questions remain: Why do whole phenomena occur? How are several series of mechanically ordered individual phenomenal objects coordinated into groups in order to form complex and recurrent phenomena? The question of recurrence involves the problem of induction and indicates that some principle—in addition to the law of efficient causes—must be found to explain the recurrence of phenomena. If we are neither to stretch the principle of efficient causes beyond reasonable bounds nor to supplement it with some occult principle ex machina, then we must suppose that the whole phenomenon—complex yet persistent—contains the reasons for its unity and recurrence. Lachelier, like Immanuel Kant, recognized a whole to be an end when the whole contained the reason for the organization of its parts. (A whole of this kind is illustrated in a stable chemical compound or in a living organism.)

Thus, in view of the fact that we indisputably are aware of phenomena which are harmonious and recurring complexes or wholes of this sort, Lachelier arrived at a second principle: The law of final causes. By its opera-

tion, sensa are grouped into perceptions of which we are actually aware, and thus they provide content and reality for the necessary but empty form of the universal mechanism. This law is to the matter of phenomena what the law of efficient causes is to their form. In these terms the distinction is drawn between the abstract existence of mechanical nature and the concrete existence of teleologically unified but contingent individuals. Since all actual objects are complex, they all presuppose the operation of the law of final causes. This law is, then, prior to the law of efficient causes in respect to actual existence.

These two laws are not on the same logical footing. Lachelier regards the law of efficient causes as proved. The proof is of the Kantian type. Given coherent experiences, this law, which is logic projected into phenomena, expresses the condition under which they cohere and are intelligible. The law of final causes, however, is not reached in the same way. Presumably, simple phenomena might remain logically ordered while being grouped in different ways. Their actual grouping into the harmonious and persistent unities that we experience is the consequence of a law which operates more like an act of will than like a formal or logical requirement. Thus, the law of final causes is said to be regulative only.

The twin laws of efficient and final causes provide the foundation for induction. Induction is thereby "founded" in the sense that it is partly proved or derived from the conditions for experience and partly justified as expressing a teleology of nature. The practice of induction, therefore, may be expected to be partly the logical deduction of events from previous events, and partly a "divining" that natural phenomena will cooperate with each other in a given way under given circumstances.

This foundation, however, is not ultimate. It does not explain why these two laws alone are the ordering principles of our existent world. Lachelier, in considering this point, observed that some organisms realize to a higher degree than others that harmony toward which nature moves. In fact, the law of final causes entails a whole hierarchy of beings that increase in order and harmony. The more complexly unified organisms in nature are not the chance products of accidentally unified simpler organisms. Rather, the simpler organisms, implicit in the more complex ones, are separated from them by a kind of "division and refraction."

The human being can free himself in thought from the particular mechanical conditions of phenomena. He has the capacity to separate some perceptions from others and, using them as symbols, to represent general properties of things. In his ability to abstract and generalize, the human being, although distinguished from all other things by this capacity, can be said to be in contact with the whole universe. The universe can be discovered again in thought but under a new condition, freedom. In addition, man is free because he can select the means and ends of his activity by reference to ideas. Hence, through man, the realm of final causes and the freedom that is its condition penetrate the organic and mechanical realms. Furthermore, without freedom it would be impossible to conceive of either mechanism or finality. Thus, the laws of efficient and final causality, upon which induction is founded, are themselves founded upon freedom—and freedom is the essential property of thought.

The process of founding induction within a philosophy of nature, therefore, consists partly in a demonstration and partly in a discovery of regulative rules. Finally, the process terminates in a metaphysics that affirms the basic reality of thought. This metaphysics is intended to found the philosophy of nature in the sense of providing a reason for belief in the unity of its laws and in its idealistic source. Lachelier's metaphysics of freedom is further developed in his article "Psychologie et métaphysique" (1885) and is given a religious dimension in "Le pari de Pascal" (1901).

See also Bergson, Henri; Continental Philosophy; Idealism; Induction; Kant, Immanuel; Ravaisson-Mollien, Jean Gaspard Félix.

Bibliography

WORKS BY LACHELIER

Oeuvres de Jules Lachelier. 2 vols. Paris: Alcan, 1933.

Lachelier, la nature, l'esprit, Dieu. Edited by Louis Millet. Paris, 1955.

The Philosophy of Jules Lachelier. Edited by Edward G. Ballard. The Hague: Nijhoff, 1960. Contains an introduction and translations by Ballard of "Du fondement de l'induction," "Psychologie et métaphysique," "Le pari de Pascal," and several short writings.

WORKS ON LACHELIER

Devivaise, C. "La philosophic religieuse de Jules Lachelier." *Revue des sciences philosophiques et théologiques* (1939): 435–464.

Mauchaussat, Gaston. *L'idéalisme de Lachelier.* Paris: Presses Universitaires de France, 1961.

Millet, Louis. *Le symbolisme dans la philosophie de Jules Lachelier.* Paris, 1959. Contains a comprehensive bibliography of Lachelier's writings and of the commentaries.

Noël, Georges. "La philosophie de M. Lachelier." *Revue de métaphysique et de morale* (1898): 230–259.

Séailles, Gabriel. *La philosophie de Jules Lachelier.* Paris, 1921.

Edward G. Ballard (1967)

LAKATOS, IMRE

(1922–1974)

LIFE

Imre Lakatos did important work in the 1960s and 1970s in the philosophy both of mathematics and science. He was born Imré Lipsitz in Debrecen Hungary, and by the time he left for England after the Hungarian Uprising in 1956, he had already lived a complex, charged, and controversial life. A convinced and influential Marxist, he had been unofficial leader of a group of young Jews in hiding from the Nazis after the invasion in 1944. As a high ranking official in the Ministry of Education after the war, he was involved in significant and controversial education reform before being arrested by the secret police in 1953 and held for three years under appalling conditions, sometimes in solitary confinement, in Recsk—the worst of the Gulag-style camps in Hungary.

He studied mathematics, physics, and philosophy at the University of Debrecen, graduating in 1944. He obtained a first PhD (with highest honors) from the Eötvös Collegium in 1947—this for a thesis on the sociology of science that he later insisted was worthless. After leaving Hungary in 1956, he obtained a Rockefeller Foundation grant to study for a second PhD at the University of Cambridge. From 1959 onward he regularly attended Karl Popper's seminar at the London School of Economics (LSE). Popper became the most important influence on him; amongst other things, Popper's Open Society views reinforced the decline of his faith in Marxism that had begun in 1956. Lakatos accepted a lectureship in logic at LSE in 1960 and was promoted to a personal chair (in Logic, with special reference to the philosophy of mathematics) in 1970. He was only fifty-one years old and still teaching at LSE at the sadly early time of his death from a heart attack in 1974.

PHILOSOPHY OF MATHEMATICS

Lakatos's Cambridge PhD thesis became the basis for his *Proofs and Refutations*. This work, published initially in the form of journal articles in 1963–1964 and in book form only posthumously in 1976, constitutes his major contribution to the philosophy of mathematics. A dialogue between a group of frighteningly bright students and their teacher, it reconstructs the process by which Euler's famous conjecture about polyhedra (that they all satisfy the formula: number of vertices plus number of faces minus the number of edges equals two) was proved and, in the process, heavily modified and transformed. Lakatos's claim was that although the eventual proof of the theorem in mathematics may be cast as a straightforward deduction, the process by which the proof is found is a more exciting process, involving counterexamples, reformulations, counterexamples to the reformulations, and careful analysis of failed proofs leading to further modifications of the theorem. Any number of interesting claims about both the history and philosophy of mathematics are thrown in to the mix—sometimes in the main text, sometimes in one of the voluminous footnotes. The work is a literary tour de force.

The extent to which *Proofs and Refutation*s represents a distinctive epistemological view that might challenge more traditional accounts in the philosophy of mathematics, such as logicism or formalism, is a controversial one. Lakatos sometimes described himself as extending Popper's fallibilist-falsificationist view of science into the field of mathematics, and there are even hints of Lakatos's Hegelian past in some of the claims about the autonomous development of mathematics. An alternative view, however, is that the main significance of his work is to cast light simply, though importantly, on the development of mathematics—on how mathematical truth is arrived at—and that it has nothing distinctive to say about the epistemological status of mathematical truths once they have been arrived at. But even if this alternative view is correct, there is a good of undoubtedly epistemological significance in some of the particular issues raised (for example, what he calls the problem of translation highlighting issues about how the formal systems, within which effectively infallible proof can be achieved, relate to the informal mathematics said to be captured by those formal systems).

PHILOSOPHY OF SCIENCE

As indicated, Lakatos thought of himself for some years as extending Popperianism, developed as an account of natural science, into the seemingly unlikely field of mathematics. However, he eventually began to discern faults in Popper's philosophy of natural science. Most significantly, in comparing Popper's views with those of Thomas Kuhn, Lakatos came to realize that Popper's view on the way that evidence impacts on scientific theories is seriously awry.

Lakatos claimed that science is best viewed as consisting not of single, isolated theories but rather of broader research programs. A hard core of principles characterizes such a program, but this needs to be supplemented by an evolving protective belt of more specific and auxiliary assumptions in order to come into contact

with experiment. When the latest theory produced by a program proves to be inconsistent with some empirical result, then the standard response of the program's proponents will be to retain the hard core and look to modify some element of the protective belt. This is a process much closer to Kuhn's idea of adverse experimental results being treated as anomalies than to the standard Popperian idea of falsification. However, while Popper used his framework to defend the idea that theory-change in science is a rational process, Lakatos believed that to accept Kuhn's account of paradigms and paradigm shifts was in effect to abandon the view that the development of science is rational. Kuhn's view, he (in)famously claimed, makes theory-change a matter of mob psychology. He was therefore led to make the important distinction between progressive and degenerating programs. The latest Newtonian theory was inconsistent with observations of Uranus's orbit; Newtonians reacted not by giving up the basic theory but by postulating a new planet.

Philip Gosse (1810–1888) realized that claim that God created the world essentially as it now is in 4004 BC is inconsistent with what Darwinians believed to be the fossil record; Gosse reacted not by surrendering the basic creationist theory (hard core), but by postulating that the alleged fossils were parts of God's initial creation. The first was a great scientific success; the second bears the clear hallmark of pseudoscience. Why? Lakatos's answer is that the Newtonian shift was progressive: It not only solved the anomaly of Uranus but made extra predictions (of the existence of a new and hitherto unsuspected planet) that could be tested empirically and were indeed confirmed (by the discovery of Neptune). Gosse's shift is degenerating: All it does is reconcile the basic creationist theory with observation but permits no independent test. The development of science consists of the replacement of degenerating programs by progressive ones. There are many other interesting aspects of the methodology, particularly concerning the role of heuristic principles, and of whether it does satisfactorily save the rationality of science.

See also Epistemology; Hegel, Georg Wilhelm Friedrich; Kuhn, Thomas; Logic, History of: Precursors of Modern Logic: Euler; Marx, Karl; Newton, Isaac; Philosophy of Science; Popper, Karl Raimund.

Bibliography

WORKS BY LAKATOS

With Alan Musgrave, eds. *Criticism and the Growth of Knowledge*. Cambridge, U.K.: Cambridge University Press, 1970.

Proofs and Refutations. The Logic of Mathematical Discovery, edited by John Worrall and E.G. Zahar. Cambridge, U.K.: Cambridge University Press, 1976.

The Methodology of Scientific Research Programmes and Mathematics, Science and Epistemology. Philosophical Papers. Vols. 1 and 2, edited by John Worrall and Gregory Currie. Cambridge, U.K.: Cambridge University Press, 1978.

With Paul Feyerabend. *For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence*, edited and with an introduction by Matteo Motterlini. Chicago, IL: University of Chicago Press. 1999.

WORKS BY OTHERS

Kampis, George, Ladislav Kvasz, and Michael Stölzner, eds. *Appraising Lakatos: Mathematics, Methodology and the Man*. Dordrecht, Netherlands: Kluwer, 2002.

Larvor, Brendon. *Lakatos: An Introduction*. London: Routledge, 1998.

John Worrall (2005)

LALANDE, ANDRÉ

(1867–1964)

André Lalande, the French philosopher, was born in Dijon and entered the École Normale Supérieure in 1885. He took his doctorate in 1899 and taught in *lycées* until he was appointed first to a lectureship and then, in 1904, to a chair of philosophy at the University of Paris.

Lalande was a rationalist whose whole life was devoted to the cause of international communication and the dissemination of knowledge. His constant preoccupation after 1902 was the launching, and subsequent reediting, of the *Vocabulaire technique et critique de la philosophie,* which aimed at the concise definition and standardization of philosophical terminology. His own philosophical work corresponds to this recognition and promotion of an interdependent humanity.

In his thesis of 1899, *L'idée directrice de la dissolution opposée à celle de l'évolution,* Lalande challenged Herbert Spencer's thesis that progress is evolutionary and differentiating, and held that, on the contrary, dissolution—or, as he later called it, involution—is more widespread and significant. Involution, or movement from the heterogeneous to the homogeneous, is observable in nature as entropy, or increase of randomness. In human life, however, this movement toward uniformity is fruitful and is served by reason, which, in scientific investigation, leads to the progressive subsumption of more and more classes of phenomena under fewer general laws.

Lalande disapproved of an imposed uniformity, which represents merely the transference from the indi-

vidual to the group of evolutionary, divisive drives. True reason ensures that although people feel differently, they shall think in the same way and thus understand each other even when they do not resemble each other. Lalande's concern was for the individual, whose uniqueness is sacrificed to function in a rigidly specialized and differentiated society. The application of reason to life in the technological field liberates the individual from his functional role, and the application of reason in the cultural field enables men to afford, and to benefit from, the diversity that is their birthright.

In *La raison et les normes* Lalande restated his involutionist case in the light of recent philosophies of "being-in-the-world." He took cognizance, for example, of the argument that geometrical, objective space is derived from the neuromotor "spaces" of man facing his tasks, but for Lalande the superiority of a common space amenable to conceptualization remained unimpaired. Similarly, he preferred chronological time to the "real" time of naive emotional experience. Lalande reaffirmed his universalist conception of rationality against more recent phenomenological thinking.

See also Continental Philosophy; Rationalism.

Bibliography

WORKS BY LALANDE

Lectures sur la philosophic des sciences. Paris, 1893; and Paris: Hachette, 1907.

L'idée directrice de la dissolution opposée à celle de l'évolution. Paris, 1899; revised and reissued as *Les illusions évolutionnistes.* Paris, 1930.

Quid de Mathematica vel Rationali vel Naturali Senserit Baconus Verulamius. Paris, 1899. Latin thesis.

Précis raisonné de morale pratique. Paris: Hachette, 1907.

Les théories de l'induction et de l'expérimentation. Paris: Boivin, 1929.

La raison et les normes. Paris, 1948.

Vocabulaire technique et critique de la philosophic, 8th ed. Paris: Presses Universitaires de France, 1960.

WORKS ON LALANDE

Lavelle, L. *La philosophic française entre les deux guerres.* Paris, 1942.

Smith, Colin. *Contemporary French Philosophy.* New York: Barnes and Noble, 1964.

Colin Smith (1967)

LAMARCK, CHEVALIER DE

(1744–1829)

Jean Baptiste Pierre Antoine de Monet, Chevalier de Lamarck, the French biologist and formulator of the first comprehensive theory of evolution, was born at Bazentin-le-Petit, a village in northeastern France. As a youth he studied briefly for the priesthood, but later withdrew to follow the family tradition of army service. While in Paris recovering from an injury and intermittently studying medicine, he met Jean-Jacques Rousseau, through whom he became interested in botany. This interest led to investigations that culminated in the publication of a large work on the flora of France, which brought Lamarck immediate fame and election to the Academy of Sciences. From 1783 to 1793 he held a small post at the Jardin du Roi, which was reorganized and expanded along lines proposed by Lamarck to include a museum of natural history and twelve professorial chairs. The last of these, for the study of invertebrates, went almost by default to Lamarck himself. Hence, at the age of fifty he began his indefatigable labors as a zoologist. These labors led to his conclusion, at some time between 1794 and 1802, that a transmutation of animal species had taken place. He expounded his views in a succession of important works: *Système des animaux sans vertèbres* (Paris, 1801), *Rechèrches sur l'organisation des corps vivans* (Paris, 1802), *Philosophie zoologique* (2 vols., Paris, 1809–1830, translated by H. Elliot as *Zoological Philosophy,* London, 1914), and *Histoire naturelle des animaux sans vertèbres,* (7 vols., Paris, 1815–1822). The significance of Lamarck's contribution was scarcely appreciated by his contemporaries. When he died at the age of eighty-five, blind and poor, he had become a forgotten man. His body was buried in a pauper's grave whose exact location is unknown.

SYSTEM OF NATURE

Lamarck aspired to produce a large-scale "system of nature" set in a deistic framework. He held that nature, "the immense totality of different beings," is neither eternal nor self-explanatory. It is the creation of a "Supreme Author" who brought matter into being and instituted the world order by means of laws that govern whatever happens. Within nature, change is universal. But nature in toto is unchangeable and "should be regarded as a whole constituted by its parts, for a purpose which its Author alone knows." This whole, however, is as distinct from the Creator as a watch is from the watchmaker. Hence, nature has productive powers of its own that the sciences can properly interpret in mechanical and materialistic terms. The system that Lamarck originally planned was to have included sections on physics, chemistry, meteorology, geology, and biology. Some of his writings did, in fact, discuss all these topics, but what

appeared can hardly be said to form a unified scheme. His attention was increasingly occupied by his reflections on living things, the science of which he named biology in 1802.

EVOLUTION

Lamarck effected a breakthrough to an evolutionary conception of nature by bringing together several lines of thought. His geological studies convinced him that Earth had endured for an immense span of time, during which it had undergone many changes of a gradual sort, especially in its surface features. His observation of fossils supported the conclusion that animal life had existed for a large part of geological time and had also undergone gradual changes. Hence, species must be mutable, and their apparent stability is due to man's limited time perspective. Furthermore, organisms are simply physical bodies whose parts are highly organized. Thus, Lamarck was opposed to vitalism. "Every fact or phenomenon observed in a living body," he held, "is … a physical fact or phenomenon, and a product of organization" (*Histoire naturelle des animaux sans vertèbres,* Vol. I, p. 53). Accordingly, he accepted the conclusion that a "spontaneous generation" of organisms had occurred. Animals and plants represent two independent lines stemming from two distinct types of spontaneous generation that utilized chemical materials differently. These materials are wholly inanimate and display none of the characteristic properties observed in the organisms they constitute.

PERFECTING POWER IN NATURE

The history of living things on Earth reveals a steady increase in the complexity of their organization, a process by which they have also been perfected. "Nature has produced all the species of animals in succession, beginning with the most imperfect or simplest, and ending her work with the most perfect." Man is the being who exemplifies the highest excellence of bodily organization, and he thereby provides "the standard for judging the perfection or degradation of other animal organizations." Lamarck's thought at this point was influenced by the idea of the "great chain of being," the infinitely graded series of forms from highest to lowest, which was a doctrine congenial to eighteenth-century deism. Since, in his evolutionary approach, the series came into existence from the bottom, Lamarck attributed it to a perfecting power inherent in nature. The postulating of this perfecting power is the feature of Lamarck's evolutionism that separates it most sharply from that of Charles Darwin.

CAUSES OF THE POWER OF EVOLUTION

If the environment were unchanging, the perfecting power of nature would produce a simple, linear sequence of organisms. But the environment is ceaselessly changing, and, as a result, evolution is "deflected" from a linear path into the "branching" pattern actually found among plants and animals. The mechanism by which the branching pattern is formed consists of a group of causal factors often mistakenly supposed to be the whole of Lamarck's theory, instead of just a part of it.

The causal factors are specified in several "laws"—two in *Philosophie zoologique* and four in *Histoire naturelle des animaux*—whose purport can be summarized as follows. The organs and habits by which animals maintain their adaptation to the environment are controlled by bodily fluids that are constantly in motion. Animals whose structure is so elementary that they have no faculty of feeling are acted on mechanically by environmental changes. New motions of the internal fluids are set up, and these give rise to adaptive alterations in the organs and habits. The case is different with animals whose structure is complicated enough to enable them to feel wants or needs (*besoins*). When the environment of these animals changes, new needs are felt, and each need, "exciting their inner feeling (*sentiment intérieur*), forthwith sets the fluids in motion and forces them toward the point of the body where an action may satisfy the want experienced" (ibid., p. 185). If a suitable organ already exists at that point, it is immediately incited to action. If not, the felt need gradually causes the organ to be generated, "provided the need be pressing and continuous." Everything thus acquired by an individual animal during its lifetime is preserved by heredity (*génération*) and transmitted to that individual's progeny. The operation of these causal factors, superimposed on the general perfecting tendency of nature, accounts for all that has happened in evolution.

MAN

Man's place in this theory was a topic that Lamarck understandably treated with caution. He stressed man's "extreme superiority" over other living things because of his possession of reason, although anatomically he differs only in degree from monkeys and apes. Is it not plausible to suppose that the differences have been "gradually acquired" over a long period of time? "What a subject for reflection," Lamarck commented, "for those who have the courage to enter into it!" He himself dared in a short section of *Philosophie zoologique* to outline a hypothetical

explanation of how apelike beings might "at length be transformed" into manlike beings, able to walk upright, to use tools, and to develop "the marvelous faculty of speaking." Throughout the process, changed habits would produce new wants and new capacities, until true human beings appeared. "Such are the reflections which might be aroused, if man were distinguished from animals only by his organization, and if his origin were not different from theirs." At this point Lamarck's courage apparently gave out.

ASSESSMENT

Despite the comprehensiveness of his outlook, Lamarck failed to formulate a unified theory of evolution. Therefore, he had to conclude that the diversification of plants and simple animals was due to mechanical factors alone, whereas in the case of complex animals an important psychological and teleological factor was operative. He held that no species had ever been totally extinguished, in spite of what the fossil evidence indicated, because he believed that the plan of the Supreme Author of the universe would not allow such wastage. His acceptance of the perfecting tendency obliged him to affirm that there are really two animal series: the grand one from simple to complex, and the particular, branching series that have deviated from it. Above all, his theory demanded not only that modifications acquired by parents during their lifetime should affect their offspring, but also that they should affect the same parts in the offspring as in the parents and should become a permanent hereditary feature in that line of descent, regardless of later modifying factors. Modern genetic research has shown strong, although perhaps not conclusive, reasons for believing that such an "inheritance of acquired characteristics" cannot occur. None of these difficulties, however, can detract from the greatness of Lamarck's contribution. "He first did the eminent service," Darwin remarked, "of arousing attention to the probability of all change in the organic world being the result of law, and not of miraculous interposition."

See also Darwin, Charles Robert; Darwinism; Evolutionary Theory; Laws of Nature; Rousseau, Jean-Jacques; Vitalism.

Bibliography

Cannon, H. G. *Lamarck and Modern Genetics.* Springfield, IL: Thomas, 1960.

Gillispie, C. C. "The Formation of Lamarck's Evolutionary Theory." *Archives internationale d'histoire des sciences* 35 (1956): 323–338.

Gillispie, C. C. "Lamarck and Darwin in the History of Science." *American Scientist* 46 (1958): 388–409.

Landrieu, Marcel. *Lamarck.* Paris: Société Zoologique de France, 1909.

Packard, A. S. *Lamarck, the Founder of Evolution.* New York: n.p., 1901.

Russell, E. S. *Form and Function.* Cambridge, U.K.: Murray, 1916. Ch. 13.

Wilkie, J. S. "Buffon, Lamarck and Darwin." In *Darwin's Biological Work,* edited by P. R. Bell, 262–307. Cambridge, U.K.: Cambridge University Press, 1959.

T. A. Goudge (1967)

LAMBERT, JOHANN HEINRICH

(1728–1777)

Johann Heinrich Lambert, the German mathematician, physicist, astronomer, and philosopher, was born in Mulhouse, Alsace. He taught himself mathematics, philosophy, and Asian languages; after 1748 he served as tutor in a Swiss family, traveling about Europe with his pupils for several years. He became a member of the Munich Academy in 1759 and of the Berlin Academy in 1764. In 1765 he was appointed by Frederick II as Prussian surveyor of public works. He did research in heat, light, and color and was the founder of the science of photometry. In mathematics Lambert demonstrated that π is an irrational number, and he introduced the conception of hyperbolic functions into trigonometry. In his *Kosmologische Briefe über die Einrichtung des Weltbaues* (Cosmological letters on the structure of the universe; Augsburg, 1761), Lambert proposed a cosmogonic hypothesis based on Isaac Newton's theory of gravitation; it was similar to the nebular hypothesis proposed earlier by Immanuel Kant in his *Allgemeine Naturgeschichte und Theorie des Himmels* (Königsberg and Leipzig, 1755) but unknown to Lambert.

Lambert's *Neues Organon, oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung von Irrtum und Schein* (New organon, or thoughts on the investigation and indication of truth and of the distinction between error and appearance; 2 vols., Leipzig, 1764) was an attempt to reform Wolffian logic. It was strongly influenced by the logical treatises of the Pietist philosophers A. F. Hoffmann and C. F. Crusius, and like their work it widened the field of logic to cover psychological and methodological questions. Although Lambert believed that metaphysics should follow a mathematical method, he assumed, like the Pietists and John

Locke, a multiplicity of elementary notions. The a priori sciences (pure theoretical and practical philosophy) should be constructed by combining these elementary notions mathematically. The final section of the *Neues Organon* discusses appearance and gives a theory of experimental and probable knowledge. It contains rules for distinguishing false (or subjective) appearance from true (or objective) appearance, the latter arising from true perception of the phenomenal world. As a blend of Leibnizian, Wolffian, Lockean, and Pietist elements the *Neues Organon* was neither more original nor more influential in its time than several Pietist treatises on logic or J. B. Basedow's *Philalethie.*

The lesser-known *Anlage zur Architektonik, oder Theorie des Einfachen und Ersten in der philosophischen und mathematischen Erkentniss* (Foundation of architectonic, or theory of the simple and primary elements in philosophical and mathematical knowledge; 2 vols., Riga, 1771) was a much more important work. In this work Lambert, dissatisfied with classical German and particularly Wolffian metaphysics, proposed a far-reaching reform through an analysis of the sources, genesis, and development of the basic concepts and axioms of metaphysics and their interrelations. Reacting also against sensationalism, skepticism, and the new schools of commonsense and popular philosophy, Lambert wished to save metaphysics by presenting it in a phenomenalistic manner (as J. N. Tetens and Kant were to do later).

Following Locke, Lambert assumed a certain set of concepts as given and then examined them. Once the analysis was completed, Lambert held, it would be possible to change from an empirical to a rationalistic procedure—the a priori deductive construction, modeled on the procedures of mathematics, of a body of general sciences that are true both logically and metaphysically. The deduced propositions of these sciences would then be applied to experience in the manner of applied mathematics. The joining of such propositions with rules abstracted from observation and experiments would give a foundation for truth in each of the particular sciences.

There were thus two main aspects to Lambert's philosophy, the analytic and the constructive. The former was the predominating interest in the *Anlage zur Architektonik.* This work consists largely of detailed discussions of, and subtle distinctions between, many of the most common simple notions and axioms and elementary interrelations discussed in traditional metaphysics. This refined analysis, too detailed even to be sampled here, exerted a great influence on Teten's mature work and on the making of Kant's *Kritik der reinen Vernunft.* Kant had earlier been much impressed by the *Neues Organon,* and acknowledged to Lambert in correspondence his interest in Lambert's analyses.

The second, constructive, aspect of Lambert's philosophy was an attempt to develop a mathematical logic (or "intensional calculus") for deducing propositions by an easy and exact method from the simple notions and axioms, once they have been established analytically.

See also Crusius, Christian August; Kant, Immanuel; Locke, John; Logic, History of; Metaphysics; Newton, Isaac; Tetens, Johann Nicolaus.

Bibliography

ADDITIONAL WORKS BY LAMBERT

Deutscher gelehrter Briefwechsel. 5 vols. Edited by Johann Bernoulli. Berlin, 1781–1787. Correspondence.

Logische und philosophische Abhandlungen. 2 vols. Edited by Johann Bernoulli. Berlin and Dessau, 1782.

Abhandlungen vom Criterium Veritatis. Edited by K. Bopp. Berlin: Reuther and Reichard, 1915.

Über die Methode, die Metaphysik, Theologie und Moral richtiger zu beweisen. Edited by K. Bopp. Berlin: Reuther and Reichard, 1918.

WORKS ON LAMBERT

Arndt, H. W. *Der Moglichkeitsbegriff bei Chr. Wolff und J. H. Lambert.* Göttingen, 1959. A mimeographed thesis.

Baensch, O. *J. H. Lambert und seine Stellung zu Kant.* Tübingen, 1902.

Eisenring, M. E. *Johann Heinrich Lambert und die wissenschaftliche Philosophie der Gegenwart.* Zürich, 1942.

Huber, D. *J. H. Lambert nach seinem Leben und Wirken.* Basel, 1829.

Krienelke, K. *J. H. Lamberts Philosophie der Mathematik.* Berlin, 1909.

Lepsius, Johann. *J. H. Lambert.* Munich, 1881.

Zimmermann, R. *Lambert der Vorgänger Kants.* Vienna, 1879.

Giorgio Tonelli (1967)

LAMENNAIS, HUGUES FÉLICITÉ ROBERT DE

(1782–1854)

Hugues Félicité Robert de Lamennais, the French ecclesiastic and philosopher, was born in Saint-Malo, Brittany, and died in Paris. Lamennais received the tonsure in 1809 but was not ordained a priest until 1816. His early works in defense of ultramontanism won him the approval of Rome, but it was not long before his inability to compromise in the interest of expediency led to his condemnation. Although never excommunicated, he voluntarily

relinquished all sacerdotal functions and died after refusing the last rites.

ULTRAMONTANISM

Lamennais's first influential work, *De la tradition de l'église sur l'institution des évêques* (Paris, 1814), written in collaboration with his brother Jean, was an attack on Gallicanism. Directly inspired by Vicomte de Bonald, it propounded three theses—the supremacy of the Church of Rome, papal infallibility in matters of doctrine, and the basic authority of tradition. It did not, however, grant the pope any sovereign rights in temporal matters. Lamennais's second work, the *Essai sur l'indifférence en matière de religion* (1817–1823) was welcomed enthusiastically in Catholic circles and received the approval of Leo XII. It took as its premises that no beliefs are without influence on the welfare of society and that religious beliefs are of primary importance in this respect. Hence, no man has the right to be neutral in religious disputes. Neutrality may arise from false notions of religion's place in life, from a failure to distinguish between orthodoxy and heresy, or from ignorance, lack of serious purpose, or simple sloth. Since no one can rightly maintain two antithetical ideas, there can be only one religious truth, one mouthpiece for it, and one tradition.

TRADITIONALISM

The traditionalism involved in this led to Lamennais's denial of the individual's rational powers, a denial that he clung to consistently. Our senses, feelings, and reason may lead to the truth, but only accidentally. Certitude can be acquired only by the common reason, that of the human race. One must therefore fuse his opinions with those of his fellow men and find the solution to his problems in faith, authority, and common sense. Trust in one's own insight is madness, as is eccentricity of behavior. But if one asks whence comes the authority of the general reason, the answer is, from God. God has entrusted it to the church, which speaks through the pope. No individual philosopher, even though he be a Descartes, can substitute his method for that based on revelation.

THE CONDEMNATION

So extreme a form of ultramontanism may have been logical, granted its premises, but it was politically inexpedient. Its anti-Gallicanism alone would have aroused resentment, but it was coupled with violent attacks on the French university system, the Charter, and certain personalities, such as Comte Denis de Frayssinous. Lamennais paid little attention to his critics, turned from them to the Vatican, and was shocked to receive in 1832 the encyclical *Mirari Vos,* which, without mentioning him by name, nevertheless condemned his ultramontanism on the ground that it disrupted the existing harmony between church and state. At the same time, it condemned freedom of conscience and opinion, which could lead only to freedom to err. Lamennais submitted but restricted his submission to questions of religion. During this period he also published his *Paroles d'un croyant* (1834), a series of prose poems that preached fraternity, freedom of association, and confidence in God and in prayer. This work was condemned outright in the encyclical *Singulari Nos* (1834).

PHILOSOPHY

In substituting "the Christianity of the human race" for that of the Vatican, Lamennais retained his traditionalism but abandoned his ultramontanism. His point of view was expressed in a three-volume work, the *Esquisse d'une philosophie* (1840), of which he published a fourth volume in 1846. It began with a theology, continued through a philosophical anthropology, aesthetics, and philosophy of science, and was to have been completed with a social philosophy. Lamennais's theology was Trinitarian and made the three persons of the Deity power, intelligence, and love, all interfused. Each realm of being reflected this triune nature, which was undemonstrable but demanded by the very nature of human thought. The work as a whole developed this thesis.

Lamennais's philosophy was Christian traditionalism minus ecclesiasticism, but with a philosophy of nature added. No man, he held, can assent to his own deductions if they are not in harmony with those of the whole human race, and the opinions of the human race will be found in tradition. The inconsistencies of tradition were never dwelt upon. His *Esquisse,* because of its Christian overtones, had no popularity in republican circles and, as for his Catholic associates, they felt little if any need for it.

See also Bonald, Louis Gabriel Ambroise, Vicomte de; Continental Philosophy; Traditionalism.

Bibliography

WORKS BY LAMENNAIS

De la tradition de l'église sur l'institution des évêques. 3 vols. Liège and Paris, 1814.

Essai sur l'indifférence en matière de religion. 4 vols. Paris, 1817–1823.

Défense de l'essai sur l'indifférence. Paris, 1821.

Paroles d'un croyant. Paris, 1834.
Oeuvres complètes. 12 vols. Paris: n.p., 1836–1837. Not complete.
Oeuvres choisies et philosophiques. 10 vols. Paris, 1837–1841.
Esquisse d'une philosophie. 4 vols. Paris, 1840–1846.
Amschaspands et Darvands. Paris, 1843.

WORKS ON LAMENNAIS

Duine, F. *La mennais.* Évreux, 1922. The most authoritative work, based in part on previously unpublished material.
Gibson, William. *The Abbé de Lamennais and the Liberal Catholic Movement in France.* London and New York: Longmans, Green, 1896.
Janet, Paul. *La philosophie de Lamennais.* Paris: Alcan, 1890.
Mourre, M. *Lamennais, ou l'hérésie des temps modernes.* Paris: Amiot-Dumont, 1955.
Spuller, Eugène. *Lamennais: étude d'histoire politique et religieuse.* Paris, 1892.

George Boas (1967)

LA METTRIE, JULIEN OFFRAY DE

(1709–1751)

Julien Offray de La Mettrie, the French physician and philosopher, was born in Saint-Malo, Brittany. After attending the Collège d'Harcourt, he studied medicine at the University of Paris, finally obtaining his doctor's degree from the Faculty of Rheims in 1733. He next went to Leiden to complete his training under the celebrated Dr. Hermann Boerhaave, whose iatromechanist doctrines were to have a decisive influence on his orientation in the philosophical, no less than in the medical, domain. Back in Saint-Malo as a practicing physician, La Mettrie undertook to popularize Boerhaave's teachings by translating into French a number of the latter's principal works. His marriage in 1739 to Marie-Louise Droneau proved unhappy and led before long to a separation. From 1743 to 1745 La Mettrie, as surgeon to the *Gardes Françaises* regiment, participated in several campaigns of the War of the Austrian Succession. The publication in 1745 of his first philosophical work, the *Histoire naturelle de l'âme*, brought him under severe official censure for his materialist views. This circumstance, along with an imprudent satire he wrote on the foibles of his medical colleagues, caused La Mettrie to exile himself to Holland. It was there that he published in 1747 *L'homme machine*, his best known and most influential book, whose atheistic and materialistic contents aroused even the liberal-minded Dutch to angry protest.

La Mettrie was fortunate enough, at this crucial moment, to find a protector in Frederick the Great, who invited him to Berlin. In Prussia he was appointed a member of the Royal Academy of Sciences, as well as "physician ordinary" and "reader" to the king. Profiting from the security of his position, he brought out, among other writings, *L'homme plante* (1748), *Le système d'Epicure* (1750), and *Discours sur le bonheur* (1750), each of which attested, in its own way, to the sort of scandalizing unorthodoxy of thought for which their author had already acquired a unique reputation. His numerous enemies, powerless to suppress either him or his ideas, contented themselves with a plethora of refutations that were too often irrelevant in substance or abusive in tone; in particular, they drew a portrait of La Mettrie himself as a monster of depravity. But apart from his theoretical advocacy and personal pursuit of a frankly hedonistic ideal and his delight in provoking or shocking those of a stiffly bourgeois or pious outlook, La Mettrie's character was actually far from deserving the ignominy heaped upon it. He died in 1751 of what was regarded by his contemporaries, somewhat unkindly, as the effects of overeating—a diagnosis exploited by his foes to prove both the practical dangers of materialism and the providential punishment reserved for atheists. Frederick II composed the eulogy that was read before the Berlin Academy. Besides his philosophical works, La Mettrie wrote several medical treatises of only minor value, a series of polemical and ironical pamphlets aimed at his critics, and three mordant, informative satires on what he considered to be the incompetence and "malpractice" of the doctors of the period, the best being his *Machiavel en médecine* (1748–1750).

"THE HISTORY OF THE SOUL"

In the *Histoire naturelle de l'âme*, directed against the metaphysical dualism of René Descartes, Nicolas Malebranche, Gottfried Wilhelm Leibniz, and their followers, La Mettrie contended that the soul owes its being to those specific organic forms, produced by a *force motrice* inhering in matter, on which the mental faculties and operations remain dependent. The "history of the soul" thus becomes an aspect of the body's history and falls under the authority, not of the metaphysician or theologian, but of the natural scientist. In this claim we have the fundamental attitude of La Mettrie, from which his originality as a philosopher would spring. His method of inquiry consisted in moving regularly from the empirical sphere of scientific facts and theories to that of philosophy proper—the latter being regarded, at least with respect to

epistemological and psychological problems, as the logical extension of such branches of knowledge as anatomy, physiology, chemistry, medicine, and the like. La Mettrie was perhaps the first "medical" philosopher in the complete and true sense—a designation suggesting at once the strengths and weaknesses peculiar to his thought.

In the *Histoire de l'âme*, La Mettrie sought to substantiate his naturalistic conception of the soul by means of two types of evidence, profusely cited, which tend to complement each other. Drawing, on the one hand, from the common fund of Lockean sensationalism (to which he gave, incidentally, a materialist meaning), La Mettrie argued that the contents of the mind—hence the mind itself—have no reality independently of the natural world in which sense impressions originate or of the sense organs by which these are transmitted. Utilizing, on the other hand, the technical data offered by the medical sciences of his time, he affirmed that the sensitive and intellectual activities of what is conventionally called the soul depend essentially on the structure and functions of the central nervous system, in general, and of the brain, in particular. Establishing a natural continuity from the external world through the sensory apparatus to the brain itself, La Mettrie identified the soul with a physically conditioned process in a way that allowed him to explain the various faculties of the soul, such as memory, reflection, imagination, the emotions, judgment, volition, solely in terms of their related organic causes.

However, a special feature of the *Histoire de l'âme* was its exposition of materialism within the conceptual framework of Aristotelian metaphysics. La Mettrie speculated that it is by virtue of the appropriate "material forms" and "substantial forms" that matter, actively organized by an intrinsic *force motrice*, realizes its potential attributes of a "vegetative soul" and a "sensitive soul"; each of these, in turn, he makes the "directing principle" of the biological or psychological functions coming under its sway. In presenting his empirico-physiological theory of mind under Scholastic auspices, La Mettrie intended, no doubt, to lend it some measure of metaphysical support, but probably more important was his wish to disarm the censorship by insisting—as he did throughout—on his theory's conformity with the prevailing orthodox tradition in Western philosophy. His strategy did not succeed very well, however, for the Aristotelianism on which he grafted his opinions served only to render them obscure and confused, yet apparently not quite obscure enough to prevent the authorities from recognizing and suppressing his "heretical" defense of materialism.

MAN A MACHINE

The thesis of *L'homme machine*, in asserting and illustrating the material dependence of the states of the soul uniformly on the corresponding states of the body, remains similar to that of the *Histoire de l'âme*, but its mode of expression and exact meaning are appreciably different. Composed in a lively, unmethodical, popular fashion, its exposition of materialism is effected not only without any metaphysical substructure but in a definitely antimetaphysical spirit. Its naturalistic view of man, consequently, is offered mainly as a general heuristic hypothesis necessary in the positive study of behavior, without the need being felt, beyond such a standpoint, to make mental processes reductively identical with their physiological causes. Concurrently La Mettrie proposed an experimental-inductive method, as opposed to the then prevalent apriorist ones, in the search for the principles of psychology. Discussing the organic basis of both vital and psychic events, he insisted on the mechanistic character of the causation involved. This important point was not brought out clearly in *Histoire de l'âme* because of the attempted materialization of the pseudo-Scholastic "souls" and "faculties."

In *L'homme machine* no essential distinction remained between the conscious and voluntary, as against the merely vital, involuntary, or instinctual activities of the "human machine"; the two types of activity are presumed explainable by the relative complexity of the mechanical structures responsible for their production. Thus La Mettrie could claim that his man-machine theory was the extension to its logical and empirical limits of the Cartesian animal-automaton doctrine. However, he must be credited with conceiving of the "living machine" in a manner that goes beyond the inadequacies of Descartes's passive and inert notion of mechanism. The organic machine that sustains the sensitive and mental life of the individual is defined by La Mettrie as a purposively self-moving and self-sufficient system, consisting of dynamically interrelated parts. It was typical of his empirical procedure that he found proof of the autonomous energy and internal finality of the organism in the physiological data of irritability. Following the pioneering researches of Albrecht von Haller, La Mettrie was among the first to understand the radical value of the capacity for irritability, and he succeeded in interpreting it with particular relevance for his thesis of psychophysical automatism.

Among the subsidiary themes of *L'homme machine*, the declaration of atheism was a new and significant development. On the one hand, it served a polemical and

propagandist aim against the religious enemies of La Mettrie's philosophical position. On the other hand, it was a logical outcome of the universal naturalism in which the man-machine theory was appropriately framed; the traditional belief in an Intelligent Creator was replaced by the concept of an active, self-creating nature.

In epistemology, La Mettrie's characteristic approach was to offer picturable analogies between mind and brain, suggesting (however crudely) the model of a "thinking machine" into which sense perceptions feed ideas in the form of coded symbols that are, in turn, stored, classed, compared, and combined by the cerebral apparatus in order to engender all the known varieties of thought. This mechanical ordering and manipulation by the brain of its symbolically represented contents prompted La Mettrie to consider that the fundamental faculty of the mind is "imagination."

Another feature of *L'homme machine* is its persistent tendency to assimilate human to animal nature with the aid of evidence drawn from the spheres of comparative anatomy and experimental psychology. The doctrine of free will, of course, becomes meaningless in the light of physiological necessity. The moral aspect of behavior is regarded as no less determined than its other aspects, although it should be noted that the man-machine theory, despite its context of universal determinism, leads to the affirmation of a hierarchy of individual values and capabilities, inasmuch as no two "machines" could ever be identical or equal. The problem of the moral or intellectual perfectibility of man, within the compass of La Mettrie's materialism, becomes primarily a medical problem, for its solution depends on the possibility of perfecting the state of the organism.

DISCOURSE ON HAPPINESS

In the *Discours sur le bonheur*, intended as a refutation of Senecan Stoicism, La Mettrie viewed the *summum bonum* of happiness in a manner no less individualistic than hedonistic. In consistence with his materialist premises, he described happiness as the optimum state of pleasurable well-being of the "man-machine." Underlying his entire treatment of the subject is the assumption that happiness was destined by nature as a benefit to be enjoyed by each and every person, regardless of moral, intellectual, or social preconditions of any sort; that is, the goal of happiness is divorced basically from such traditional considerations as vice and virtue, ignorance and knowledge, social status and responsibility. La Mettrie obviously conceived of the problem of happiness, seen from the perspective of medical ethics, as similar to—indeed, as a special instance of—the more comprehensive problem of health. Accordingly, he diagnosed the greatest threat to felicity to be "remorse," a morbid and "unnatural" symptom, which he proposed, ever faithful to the Hippocratic oath, to alleviate in all and sundry, including even conscience-ridden criminals; he remarked that the practical control of social behavior was a political matter and no business of his.

The *Discours sur le bonheur* was misinterpreted as a cynical inducement to vice and crime and, more than any of his works, gave to the author an enduring reputation for immoralism among philosophes and antiphilosophes alike.

MINOR WORKS

Among La Mettrie's minor works, perhaps the most curious is the *Système d'Epicure*. Its concern with ontogenesis and the origin of species represented a broadening of La Mettrie's materialism into an area of biological speculation which, at the time, was just beginning to excite interest. But his description of the "evolutionary" process, in which monstrous and unviable productions are supposed to have been eliminated in favor of the well-constituted types now extant, did little more than revive Lucretian memories.

In *L'homme plante*, La Mettrie's purpose was to stress the various parallelisms of structure and function between two such seemingly disparate things as the human organism and vegetable life. Reflecting his strong taste for analogical reasoning, it is an extreme confirmation of the "chain-of-being" idea, which it interprets in the sense of a uniform destiny for man and for all other living forms, excluding the possibility of a spiritual transcendence of nature.

Les animaux plus que machines is mainly a polemical piece directed against the school of animistic biology. By elaborating a mock defense of the opinion that a "soul" governs the animal economy, La Mettrie managed to expose, with the support of much physiological evidence, the absurdity and uselessness of such a hypothesis. The inference is that it would be equally ridiculous to claim that the operations of the human machine presuppose the agency of a "soul."

La Mettrie's philosophy, and in particular the man-machine doctrine central to it, has, owing to its very character, grown somewhat obsolete, together with the scientific documentation to which it was so intimately linked. The specific features of his mechanistic theory of mind might, in relation to what is now known or still

unknown about neural processes, seem naive, crude, superficial, and pretentious. Nevertheless, his was the first naturalistic rationale for, and technical application of, a consistently physiological method in psychology. And while his philosophic contribution remains circumscribed by the biomedical standpoint that shaped his thinking, the man-machine hypothesis may be said, within its proper limits, to have retained a basic validity and vitality. Despite La Mettrie's bad name in his own age, and the many attempts to suppress, disfigure, or discredit his ideas, he exerted (surreptitiously, on the whole) a considerable influence in the eighteenth-century milieu. Among those indebted to the man-machine conception and to the naturalistic overtones and consequences that accompanied its formulation, the most important were Denis Diderot, Baron d'Holbach, and Pierre-Jean Georges Cabanis. Long neglected after his death, La Mettrie has been recognized since the latter part of the nineteenth century as one of the major forerunners of modern materialism. His nonreductive form of materialism may be regarded as an early version of a theory that is widely advocated at the present time by, among others, Ernest Nagel and various American naturalists; and his view that human beings can be fruitfully considered as a certain type of machine has obvious similarities to the principles underlying the science of cybernetics.

See also Animal Mind; Aristotelianism; Atheism; Cabanis, Pierre-Jean Georges; Continental Philosophy; Descartes, René; Diderot, Denis; Happiness; Holbach, Paul-Henri Thiry, Baron d'; Leibniz, Gottfried Wilhelm; Malebranche, Nicolas; Materialism; Mind-Body Problem; Nagel, Ernest; Naturalism; Stoicism.

Bibliography

WORKS BY LA METTRIE

Oeuvres philosophiques. 3 vols. Berlin and Paris, 1796.

L'homme machine. Translated as *Man a Machine*. London: G. Smith, 1750. A later translation is *Man a Machine; including Frederick the Great's "Eulogy" … and extracts from "The Natural History of the Soul."* Translated with philosophical and historical notes by Gertrude C. Bussey. Chicago: Open Court, 1912.

WORKS ON LA METTRIE

Bergmann, Ernst. *Die Satiren des Herrn Maschine*. Leipzig, 1913. On La Mettrie's activities as a satirist.

Boissier, Raymond. *La Mettrie, médecin, pamphlétaire, et philosophe*. Paris, 1931. Mainly on the medical background of La Mettrie's thought.

Campbell, Blair. "La Mettrie: The Robot and the Automaton." *Journal of the History of Ideas* 31 (1970): 555–572.

Lange, F. A. *History of Materialism*. London: K. Paul, Trench, Trubner, 1925. Bk. I, Sec. IV, Ch. 2 has a good discussion of La Mettrie.

Lemée, Pierre. *Julien Offray de La Mettrie, médecin, philosophe, polémiste: sa vie, son oeuvre*. Mortain, France, 1954. Best biographical account.

Perkins, Jean. "Diderot and La Mettrie." *Studies on Voltaire and the 18th Century* 10 (1959): 49–100.

Pflug, Günther. "Lamettrie und die biologische Theorien des 18. Jahrhunderts." *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 27 (1953): 509–527.

Poritzky, J. E. *Julien Offray de La Mettrie, sein Leben und seine Werke*. Berlin, 1900. Best general treatment of La Mettrie's philosophy.

Thomson, Ann. "La Mettrie, Machines, and the Denial of Liberty." *Graduate Faculty Philosophy Journal* 22(1) (2000): 71–86.

Thomson, Ann. *Materialism and Society in the Mid-Eighteenth Century: La Mettrie's Discours Préliminaire*. Genève: Droz, 1981.

Vartanian, Aram. *La Mettrie's L'Homme Machine: A Study in the Origins of an Idea*. Princeton, NJ: Princeton University Press, 1960. Critical edition with introductory monograph and notes.

Vartanian, Aram. "Trembley's Polyp, La Mettrie, and Eighteenth-Century French Materialism." *Journal of the History of Ideas* 11 (1950): 259–286.

Wright, John P., and Paul Potter, eds. *Psyche and Soma: Physicians and Metaphysicians on the Mind -Body Problem from Antiquity to Enlightenment*. Oxford: Clarendon Press, 2000.

Aram Vartanian (1967)
Bibliography updated by Tamra Frei (2005)

LA MOTHE LE VAYER, FRANÇOIS DE

(1588–1672)

François de La Mothe Le Vayer, a French skeptical philosopher, was born in Paris, the son of a government official. He acquired his father's post when the latter died in 1625. His wife was the daughter of a Scottish intellectual, Adam Blackwood. During his early years La Mothe Le Vayer traveled widely in Europe. In 1639 he was elected to the Académie française and in 1647 was appointed preceptor to the Duke of Orléans. He was a prominent figure in avant-garde circles in Paris—in the group around Michel Eyquem de Montaigne's adopted daughter, Mademoiselle De Gournay; in the group of *libertins érudits* with Gabriel Naudé (1600–1653), Guy Patin (1601–1672), and Pierre Gassendi; in the scientific group around Marin Mersenne; and in the literary world of Molière (1622–1673; who jested at La Mothe Le Vayer in

Le Mariage forcé and other plays) and Savinien de Cyrano de Bergerac. His many writings on skepticism began with *Dialogues d'Oratius Tubero* (1630), followed by the *Discours chrétien de l'immortalité de l'âme* (1637, the year of René Descartes's *Discours de la méthode*), *De la Vertu des payens* (1642, published with Cardinal de Richelieu [1585–1642]as the sponsor), and a long series of skeptical essays on history and culture throughout the rest of his life.

Although his views are based primarily on those of Sextus Empiricus (whom he calls "le divin Sexte" and the author of "notre décalogue") and Montaigne, La Mothe Le Vayer's skepticism represents perhaps the most extreme type of antirationalism in the seventeenth century. He continually offers a wealth of evidence to show the variations in human moral behavior, the diversity of people's religious beliefs and practices, the vanity of scientific study, and the virtues of skepticism. He rarely develops his case theoretically by means of systematic arguments. Instead, he usually offers only illustrative materials, followed by a fideistic message that man can find truth only through faith, not through the use of his reason and senses.

In *Petit Traité sceptique sur cette façon de parler, n'avoir pas le sens commun* (1647) La Mothe Le Vayer contends that man does not understand the nature of even the most obvious things. All of one's information is relative to one's faculties. Even if there are any instruments for finding the truth about things, one, unfortunately, is unable to discover them. One's senses are unreliable, and one lacks any guaranteed criterion for distinguishing veridical experiences from others. Indubitable truths can be known only in heaven, not here and not through any human science.

These views are further developed in his *Discours pour montrer que les doutes de la philosophie sceptique sont de grande usage dans les sciences* (1669), where it is claimed, as the title shows, that the great service of Pyrrhonian skepticism for the sciences is that it can eliminate any serious concern with scientific research and that such research is a form of blasphemy. He asserts, without offering any real arguments, that logic is unreliable and physics only a problematical subject about which there are conflicting opinions. Nature is the free manifestation of God's will. Therefore, any attempt to restrict God's achievement to what man can measure and understand is an attempt to limit God's freedom and is hence blasphemous. When the scientists realize how uncertain their disciplines are, they should give them up and adopt skepticism, "the inestimable antidote against the presumptuous knowledge of the learned ones."

This complete skepticism should undermine the dogmatist's confidence and pride and lead him or her to the true faith: Christianity. In *La prose chagrine* (1661) La Mothe Le Vayer proclaims that of all the ancient philosophies, "there is no other that agrees so easily with Christianity as skepticism, respectful towards Heaven and submissive to the Faith." Had not St. Paul preached that skepticism was the way to salvation? The true Christian skeptic leaves his or her doubts at the foot of the altar and lives by faith.

La Mothe Le Vayer's anti-intellectual and destructive attack on human rational knowledge (presented almost obliviously to the scientific revolution going on around him, and especially to the achievements of Descartes) and his appeal to faith, although not introducing much that was new to skeptical argumentation, carried the Montaignian position to an absurd extreme. He denied any and all value to intellectual activities and left only blind faith. As a result, many commentators from Antoine Arnauld on have assumed that he was a pure libertine, undermining all bases for religion, and have classified him, partly on the basis of his risqué work *Hexaméron rustique* (1670), as an *incrédule volupteux* (sensual nonbeliever). His views, however, are compatible with his having been either a sincere Christian skeptic or a secret atheist undermining confidence in all views and beliefs—a genuine fideist or an irreligious doubter. His philosophical influence seems to have been more through personal contact than through any serious presentation of philosophical skepticism. As a representative of the skeptical view, he was still important in Pierre Bayle's time, but was forgotten for the most part thereafter.

See also Continental Philosophy; Descartes, René; Fideism; Gassendi, Pierre; Mersenne, Marin; Montaigne, Michel Eyquem de; Sextus Empiricus; Skepticism.

Bibliography

WORKS BY LA MOTHE LE VAYER

Oeuvres. 15 vols. Paris: L. Billaine, 1669.

WORKS ABOUT LA MOTHE LE VAYER

Bayle, Pierre. "Vayer." In *Dictionnaire historique et critique*, 2 vols. Rotterdam, Netherlands: Reinier Leers, 1697.

Charles-Daubert, Françoise. *Les Libertins érudits en France au XVIIe siècle*. Paris: Presses Universitaires de France, 1998.

Giocanti, Sylvia. "La Mothe Le Vayer: Modes de diversion sceptique." In *Libertinage et philosophie au XVIIe siècle*. Saint-Étienne, France: Publications de l'Université de Saint-Étienne, 1996.

Paganini, Gianni. "'Pyrrhonism tout pur' ou 'circoncis'? La dynamique du scepticisme chez La Mothe Le Vayer." In *Libertinage et philosophie au XVIIe siècle*, edited by Antony McKenna and Pierre-François Moreau. Saint-Étienne, France: Publications de l'Université de Saint-Étienne, 1996.

Pintard, René. *Le libertinage érudit dans la première moitié du XVIIe siècle*. 2 vols. Paris: Boivin, 1943.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Richard Popkin (1967, 2005)

LANDGREBE, LUDWIG
(1902–1992)

One of the most faithful followers of Edmund Husserl's phenomenology and phenomenological philosophy, Ludwig Landgrebe is equally known for his own contributions to these fields. What characterizes Landgrebe's work is his combination of philosophical issues and arguments with the precise delimitation of principles (or essential structures) whose rejection involves a contradiction. For example, he maintains that if one negates awareness, one nevertheless presupposes an awareness of this negation. (Landgrebe prefers the term "awareness" to "consciousness" due to the many traditional meanings associated with the latter.) As an assistant to Husserl, Landgrebe edited a number of Husserl's texts, including the classic *Experience and Judgment*. As a professor of philosophy at the university of Cologne, he formed a following of phenomenologists among whom are such notables as Klaus Held, Ulrich Klaesges, and Donn Welton. Landgrebe attracted students and audiences by his vast scholarship and personal modesty, both of which were seamlessly coupled with conceptual and logical clarity. While at home in all the modern speculative metaphysics, from Descartes through Kant, German Idealism, Nietzsche, and twentieth-century French thought, Landgrebe did not engage in speculative philosophy. When asked at DePaul University (Chicago) during a discussion of Husserl's understanding of the "subject" what its proper phenomenological status is, Landgrebe replied: "If I were to speculate, I would say that it is a monad with a window."He lectured widely in Latin America, the United States, and Eastern and Western Europe.

For Landgrebe, phenomenological philosophy is an effort to combine as clearly as possible an exposition of a given philosophical position, an analysis of the prejudgments or principles without which such a position could not be maintained, and an examination of the adequacy of the principles necessary to account for it within the context of the phenomena encountered in human awareness. For example, Kant proposes to account for all knowledge on the basis of a priori structures in the manner of a transcendental deduction that explicates the empirical domain; but he does not, on Landgrebe's view, account for the mode of awareness required to secure access to the a priori domain at issue in this deduction. Such awareness is required, according to Landgrebe, if the a priori, or any other epistemological, ontological, or even metaphysical conditions are to be evidentially legitimated. This does not mean that Landgrebe avoids treating such conditions in terms of their conceptual meanings; however, he maintains that anyone positing such conditions will also have to show the manner in which they are accessible to awareness, because failing this, the one positing them is placed in the untenable position of positing conditions that she is unaware of. Only the interrogation of such awareness will be able to decipher what is essential in each condition. Thus Landgrebe's analyses and investigations are designed to articulate the awareness implicit in the most fundamental experiences that open up what is essential in the experienced, lived world.

THEMES OF LANDGREBE'S WORK

Landgrebe's work has three major themes: philosophical anthropology, the basic structures of awareness, and history.

THE FIRST THEME. The first theme, according to Landgrebe, is called for by modern philosophical, cultural, and historical relativity. Within the latter, two claims are preeminent: (1) that different cultures, historical periods, and societies offer various, even clashing, interpretations of human beings; and (2) that modern scientific and technological thinking offers the means to "make" the human into something "new" or even radically different from what it has been previously. Landgrebe points out that these various views and proposed transformations of the human assume a tacit "essence" as far as awareness is concerned, which allows the difference inseparable from the different views and transformations to be directed toward something that appears to awareness as an invariant. Without the latter, no sense could be made of the claim that what "humans" are depends on cultural, historical, social, and even technical definitions and constructs. All these are different from one another. Yet simple differences would allow only the claim that at different times and in different places there were descriptively different creatures, which could only result in a catalogue of the various differing depictions. But even

those who claim that there are radical differences in cultures, societies, and histories, still insist in using the phrase "different humans"—and it is this phrase that implicates the appearance of an invariant across all differences.

THE SECOND THEME. Landgrebe's second theme maintains that neither empiricism, with its emphasis on the contingency of facts, nor rationalism, with its stress on conceptuality and universal necessity, is adequate to account for concrete human awareness. The former, with its succession of impressions, cannot account for the continuity and unity of experience. The latter, as is obvious from Kant, can account for neither the unity of experience without positing the "I think" accompanying all representations, nor for the individuality wherein such representations could be attributed as "mine." In terms of philosophical anthropology, empiricism reduces the human to a "*factum brutum*," whereas rationalism treats the factual human as an instance of a universal concept.

Landgrebe holds that this fails to account for the distinguishability of individual subjects from one another. To confront this issue, Landgrebe's investigation begins with the life-world and our direct experiential "opening" to it. Movement, in correlation to the things of the environment, is an epistemic requirement needed to form primal perspectivity, time awareness, and special formations. From the movement of the eyes that trace out the contours of things, to traveling around the planet, the focus and maintenance of the identity of anything is formed by body movements (kinaesthetic processes), movements that, for Landgrebe, manifest the body side of the transcendental subject. Moreover, various higher-level linguistic structures are formed at the level of movement, such as "if-then" implications: "If I want to see the other side, then I shall have to walk around the thing." Access to time and space is equally provided by bodily activities: "If I want to see the other side, I will have to be there and then." Activities also reveal the fundamental human "intentionality" and purposive understanding, including the instrumental selectivity of proper and unfitting factors leading to the fulfillment of purposes.

For Landgrebe, activities form habits and the primacy of the practical "I can" or "I cannot" perform something. They comprise singular "habits" (though not in the Humean sense of the association of ideas) that distinguish one individual from another. Such distinctions arise as activities oriented to common tasks wherein we begin to recognize our "otherness" on the grounds of what we can and cannot do, and not on the basis of the initial encounter with others as subjects or minds inside of bodies. Intersubjectivity is primarily formed at the level of bodily abilities such that we recognize ourselves and others on the basis of activities. The latter, in turn, are not arbitrary, but emerge in correlation to things that make "objective" demands on such activities. This means that the world is neither in doubt nor our construct. As "Euclidean beings," we must move around and not through things.

This claim must not be confused with any kind of realism or naturalism. According to Landgrebe, the natural presence of the world still requires an explication of the processes of awareness that are structurally distinct from the composition of things. Hence, metaphysical speculations might suggest that a special-temporal object is actually a flow of energies, or a slowly changing substance, but for awareness the thing is an X that is maintained as constant and given through the formation of movements and perspectives, of expectations of the next side and the unification of the previous side as sides of the same X. The X suggests the possibility of an indefinite ability to explore the given thing, an ability that is proper to it. One can see it from more perspectives, take it apart, and thus open up the "inner horizon" of the explored X. This complex process comprises the phenomena through which the real thing is experienced.

This level of primal awareness also opens up the "outer horizon" such that the thing is in a room, the room is in a house, the house is in the field, the field is in a region, and so on. The opening up of the external horizon is equally founded on the "I can," which is able to go on exploring and hence comprising an open space-time horizon which, while implicit in the initial awareness of the thing, opens up possibilities for exploration of the world. One may be aware that in one's own region there are hills, and more hills, but the horizon does not close; it is possible that beyond the hills there are deserts, lakes, flatlands, forests, cities, and strangers who "do things differently."

This horizon extends into indefinite possibilities that the ego can concretize by going from its region to that region "then" and discovering whether its intentional orientation toward the "that and then" region, say, as a possible desert, is concretized or disappointed. The ego expected a desert and there appeared a lake. Without such a horizon of possibilities and expectations there would be no mistakes. Empiricism and rationalism fail at this juncture. It needs to be said that at the level of movement the formation of horizons belonging to awareness involves a shift from direct perceptual fulfillment to an

open world-horizon whose possibilities can only be partially concretized in direct awareness. Hence, on the one hand, in this awareness there is a "consciousness" that suggests perceptual fulfillability, whereas on the other hand, this same awareness is experienced as a transcendental condition for the experience of the world as a totality, albeit one completely accessible to a singular subject in her engagements with the world. This state of affairs leads Landgrebe to his next step: historical awareness.

THE THIRD THEME. Historical awareness, Landgrebe's third theme, manifests itself as a horizon of the past achievements of others and their appropriation by the current inhabitants of the life world, including the manner that they may vary such achievements. At this level, Landgrebe raises the question concerning our experience of the historical past and rejects Hegelian dialectics, Marxian materialism, and empirical research. None can "travel to the past," except symbolically, and none can account for such would be symbolic understanding. Apart from that, such metaphysical "accounts" of history assume a continuous theoretical time without, however, offering any justification for its continuity. In this sense, we cannot think of history as a succession of events "in time" ruled by causes, or a deduction from the "eternity" of the "laws of dialectics" (either Hegelian or Marxian). Rather, history is an active engagement of making and building, of concrete projects based on what we can do and what others have done. What they have done is present to us such that we too could have acted and performed similar tasks, but we no longer do them in this way. We have acquired different abilities and hence have no necessary continuity with our predecessors.

The discontinuity does not imply that we are not open to the understanding of how they made things, what purposes are present in their buildings, implements, and comportment. We may learn some abilities from what they did, but also vary them in order to perform our own tasks. As was the case with the horizon of awareness, history comprises a horizon of what others have accomplished, thus extending our own horizon of possibility for transforming and varying our own abilities. This means that the historical others extend my perception and abilities, thus forming a "poli-centric" field of understanding. Our own perceptions would be limited without the others from whom we "borrow" perceptions and abilities and thus recognize our limitations and possibilities—all of which, indeed, are open to the future. This view prevents speaking of a singular historical aim. Some tasks are completed and discontinued, the accomplishments abandoned; others are taken up in part after the builders and makers have long since disappeared, and still others are postponed for the future. The historical horizon of possibilities cannot be concretized in a totality; hence this openness precludes any claim that history has a singular purpose.

For Landgrebe there is another level of historical awareness: the transcendental. This type of awareness comprises a way to access the modes of perception that others assumed in their understanding of the world. Thus, whereas we may not have any knowledge of Aristotle's psychological, social, political, and personal life, we can say, from his writings, that Aristotle regarded the world as composed of substances. Each substance could be regarded under specific categories accessible to him as well as to us. In this sense, historical awareness of others is not regarded psychologically or internally, but as a mode of awareness that comprises a transcendental orientation toward the world accessible to anyone. Even when we disagree with Aristotle or Plato, we also must be aware of the way Plato or Aristotle regarded the world. This type of awareness is already intersubjective and is a condition for the claim that our own awareness is limited and in turn extended through others. We can "borrow" Aristotle's mode of awareness and enhance our own. This illustrates the sense in which for Landgrebe we comprise a field of poli-centric awareness that has historical depth prior to specific temporal locations. From this vantage point Landgrebe avoids various theoretical dilemmas. If a social philosophy claims that all social life, including theoretical thinking, is a result of material conditions, then previous historical views would not be accessible to us, because we do not live under those conditions. In turn, the view that all theories are based on given material conditions is itself a specific theory that reflects current material conditions; as such, it would follow that such a theory cannot make a universal claim. The same holds for theories of history that are premised on the notion that history is a contingent fact and all necessary truths, even in logic, are a result of "historical development." A contingent fact cannot be posited as a ground of necessity, because its "sense" as such a fact excludes "necessity."

CONCLUSION

Finally, beyond advancing the above theoretical issues, Landgrebe also engaged in the controversies surrounding the issue of whether there is one life world that is presumed as a ground of various societies and cultures,

or whether such societies and cultures comprise distinct and, at times, incompatible life worlds. For Landgrebe this controversy reveals a most fundamental issue of awareness. If there were one life world, and we were completely immersed in it, then we would not be able to recognize our immersion in it. If there were more than one, we would then either belong to one or another and thus would interpret the other in terms of our own; hence, we would fail to recognize the distinction between them. If we can achieve access to both, then we cannot belong to either and must have an awareness of both and their differences. This awareness is taken for granted in all such comparative studies, and makes its appearance with Landgrebe's question: For whom are such life worlds given? This opens the discussion of transcendental awareness in its own right, apart from this or that (however radically different) content of such awareness. And it belongs to Landgrebe's enduring merit as a phenomenologist that this discussion can only be enriched by the consideration of his seminal researches into these three major themes of his work.

See also Husserl, Edmund; Phenomenology.

Bibliography

WORKS BY LANDGREBE

Philosophie der Gegenwart. Frankfurt: Ullstein Buecher, 1961.

Der Weg der Phäenomenologie. Guetersloh, Germany: Guetersloher Verlagshaus, 1963.

Major Problems in Contemporary European Philosophy, from Dilthey to Heidegger. Translated by Kurt F. Reinhardt. New York: F. Ungar, 1966.

Phäenomenologie und Geschichte. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 1968.

Perspektiven transcendentalphäenomenologisches Forschung; für Ludwig Landgrebe zum 70. Geburtstag von seinen Kölner Schülern. Hrsg. von Ulrich Klaesges und Klaus Held. The Hauge: M. Nijhoff, 1972.

Phaenomenologie Heute; Festschrift fuer Ludwig Landgrebe. Herausgeber Walter Biemel. The Hauge: M. Nijhoff, 1972.

Experience and Judgment: Investigations in a Genealogy of Logic. Translated by James S. Churchill and Karl Ameriks. Evanston, IL: Northwestern University Press, 1973.

The Phenomenology of Edmund Husserl: Six Essays, edited and translated by Donn Welton. Ithaca, NY: Cornell University Press, 1981.

Faktizität und Individuation: Studien zu den Grundlagen der Phänomenologie. Hamburg: Meiner, 1982.

Algis Mickunas (2005)
Burt Hopkins (2005)

LANGE, FRIEDRICH ALBERT

(1828–1875)

The German philosopher, historian, and sociologist Friedrich Albert Lange was born at Wald near Solingen. He studied at Duisberg, in Zürich, where he attended the lectures of a disciple of Johann Gottfried Herbart, and at Bonn. After receiving a degree at Bonn, he taught high school in Cologne, and in 1851 he became a university instructor at Bonn. His dissertation concerned the relation between theories of education and various worldviews. From 1858 to 1861 he taught school in Duisberg but resigned because of a government order forbidding teachers to participate in political agitation. Lange remained in Duisberg as a newspaper editor and secretary of the chamber of commerce. His socialist sympathies were not incompatible with a genius for finance. In 1866 he returned to Switzerland and in 1870 became professor of inductive logic at Zürich. He was appointed to a professorship at Marburg in 1873 and remained there until his death. The philosophical poems of Friedrich Schiller, on which he sometimes lectured, were said to be his final comfort.

Lange's importance in philosophy rests mainly on his brilliantly written *History of Materialism and Critique of Its Present Significance* (1866). This work gave support to the opponents of materialism and helped to stimulate the revival of interest in Immanuel Kant that led to the neo-Kantian schools of the last decades of the century. Less important philosophically, but a prominent part of Lange's versatile career, was his concern with social questions, as in *Die Arbeiterfrage* (1865), and his work for constitutional reform in the direction of democratic socialism.

Lange argued that materialistic theories of reality are just as guilty of transcending the proper limits of human knowledge as are the speculative systems of idealism. He appealed to Kant's arguments, rejecting the possibility of any metaphysical knowledge that pretends to take us beyond the sphere of experience. In his view, the attempt to comprehend the world as a whole is doomed to failure. But this criticism applied as much to the materialistic rejection of unobservable spiritual or mental agencies as to their defense. According to Lange, metaphysical theories belong to the realm of art and religion, a field governed by poetizing (*dichten*). This activity is not an illegitimate one, however. It is an essential human need, expressive of men's yearnings for an ideal realm. But reli-

gion and the speculative systems of metaphysics do not yield scientific knowledge or any substitute for it.

Lange saw materialism both as a demand for mechanistic explanations of natural phenomena and as a naive realism and dogmatic metaphysics. The first demand he considered valid, but the second, he held, had been refuted by Kant and by the development of physiological psychology. The demand that natural occurrences be explained in terms of material causes is a useful, even indispensable, postulate of scientific method. In attempting to explain human behavior, for instance, it is unreasonable to think of consciousness as intervening somewhere in the series of physical events from stimulus to brain, nerve, and muscular response. Mental processes are not members of this series.

While the only valid categories for science are those that, like space, time, and causality, render nature mechanistically intelligible, these categories have no proper role beyond that of organizing our sense experience. Along with the basic concepts of physics—matter, atom, force, physical object—they are the products of human invention. The Kantian theory of the a priori had shown this, while discoveries in the physiology of sensation proved that our knowledge is sifted through human sense organs. The scientist is not a passive recipient of data; the laws that he discovers are constructions whose objectivity is only an objectivity for us. Though the world which science presents is the cognitive realm valid for all men, there is also the individual's world of ideals. To confuse the two worlds is wrong, because each has its significance.

Lange's physiological interpretation of the categories was rejected by his neo-Kantian successors at Marburg, Hermann Cohen and Paul Natorp. His influence was very strong, however, on Hans Vaihinger, whose pragmatism owes much to Lange's concept of categories as no more than maxims of scientific method. Lange's rejection of all metaphysics placed him also in the positivistic tradition, and it is no surprise that he referred to Auguste Comte as "the noble Comte." Though Lange was critical of Ludwig Feuerbach, whom he regarded as only half emancipated from G. W. F. Hegel, his own sympathetic but noncognitivist view of religion and ideals is akin to the humanism of Feuerbach.

See also A Priori and A Posteriori; Cohen, Hermann; Comte, Auguste; Continental Philosophy; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Herbart, Johann Gottfried; Humanism; Kant, Immanuel; Materialism; Natorp, Paul; Neo-Kantianism; Schiller, Friedrich; Vaihinger, Hans.

Bibliography

WORKS BY LANGE

Die Arbeiterfrage. Winterthur, Switzerland, 1865.

Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart. 2 vols. Iserlohn and Leipzig, 1866. The edition published at Leipzig in 1902 has an introduction and critical appendix by Hermann Cohen. English translation by E. C. Thomas, 3 vols. London, 1877–1879, reprinted with an important introduction by Bertrand Russell, 1 vol. Edinburgh, 1925.

J. S. Mill's Ansichten über die Social Frage. Duisburg, Germany, 1866.

Logische Studien. Edited by H. Cohen. Iserlohn and Leipzig, 1877.

WORKS ON LANGE

Ellissen, O. A. *Friedrich Albert Lange.* Leipzig: Baedeker, 1891. Contains letters of interest.

Vaihinger, Hans. *Hartmann, Dühring, und Lange.* Iserlohn: Baedecker, 1876.

Vaihinger, Hans. *Die Philosophie des Als-Ob.* Berlin: Reuther & Reichard, 1911. Translated by C. K. Ogden as *The Philosophy of "As if."* New York: Harcourt Brace, 1924.

Arnulf Zweig (1967)

LANGER, SUSANNE K.

(1895–1985)

Susanne Langer was an American philosopher whose work remains significant because of her distinctive views on the philosophy of art, as expressed in her books *Philosophy in a New Key* (1942) and *Feeling and Form* (1953). Though now relatively neglected, various aspects of her views remain of interest, as shown by the following considerations concerning her most characteristic doctrines.

Langer rejects positivist views of meaning and thinking according to which only literal, scientific language has any objective significance—a view the consequence of which is that any other apparent kinds of meaning are mere subjective expressions of feeling (1957, ch. 4). Instead she argues that there is another kind of objective thinking that has a different kind of symbolic form. In place of the discursive, sequential structure of linguistic statements it uses a *presentational* symbolic mode, which communicates by showing rather than saying, as do images or pictures. Such presentational modes have their origin in low-level kinds of sensory experience, which provide the basis for the often metaphorical and imagistic experiences that underlie conscious thought (1957, chs. 4, 6).

As applied to the arts, Langer claims that all of the arts are to be explained in terms of such presentational symbolic forms. For example, pictures are able to communicate their content by showing or presenting—rather than by linguistically stating—their message (1957, ch. 4); while music, dance, and other art forms similarly present rather than state their meaningful content (1957, ch. 8; 1953).

But if such presentational forms do not communicate or express objective factual information, as do discursive linguistic forms, then what do they express? Langer's answer is that they express *feeling*—not the mere subjective feelings that the positivists rejected, but instead objective forms or structures of feeling that cannot be identified either with the betrayal of the personal feelings of an artist who creates an artwork, nor with the arousal of feelings in the audience who experience that work. For example, she says of music that it "is 'significant form,' and its significance is that of a symbol … which by virtue of its dynamic structure can express the forms of vital experience … Feeling, life, motion and emotion constitute its import" (1953, p. 32). Thus artistic symbolic forms communicate, in virtue of their structure, the same forms of feeling that occur in sentient life generally.

The above views, that art involves nondiscursive symbolic forms that primarily communicate feeling, have been much criticized (e.g., see Davies 1994, ch. 3 for incisive music-related criticisms). However, there remain other, more neglected aspects of Langer's theory that are harder to dismiss, such as her view that art involves what she calls "semblance" (1953, ch. 4), a seeming or illusory quality that is both experienced as such—"The 'otherness' that gives even a bona fide product like a building or a vase some aura of illusion" (1953, p. 46)—and which also implies the objective unreality or virtuality of those forms themselves. This quality of semblance enables Langer to distinguish between, for instance, the *actual* spatial qualities of a sculpture or building when considered purely as a physical object, and its *seeming* spatial qualities, which in part constitute, on her view, the perceptually experienced symbolic artwork itself.

To be sure, such an account seems to imply that artworks are relative to perception in some way (Khatchadourian 1978), hence raising questions about their objective status that Langer does not answer, but many would view her general insistence on the objectivity and cultural independence of the symbolic forms of artworks as being too strong in any case. Independently of such issues of objectivity and semblance versus reality, Langer's resulting analyses are sometimes of interest in their own right, such as her account of the ways in which sculptures are able to organize the spaces in which they occur—unlike paintings, whose spatial worlds are self-contained; this is an account that connects with other significant differences between sculptural and pictorial forms (Hopkins 2004).

In terms of the general classification of theories of art, Langer's theory is an unusual combination of a formalist and an expression theory in that her view is that all artworks express feeling in virtue of their specific symbolic form. Probably one reason for her current neglect is that she in turn neglects issues of artistic intention and individual expression that generally are thought to be at least relevant, if not central, issues in the philosophy of art. Nevertheless, whatever her theoretical flaws may be, Langer remains an engaging and insightful writer whose previous wide popularity is not hard to understand.

See also Aesthetics, History of.

Bibliography

WORKS BY LANGER

Philosophy in a New Key: A Study in the Symbolism of Reason, Rite and Art. Cambridge, MA: Harvard University Press, 1942. 3rd ed., 1957.

Feeling and Form: A Theory of Art. New York: Scribner, 1953.

WORKS ON LANGER

Davies, Stephen. *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press, 1994.

Hopkins, Robert. "Painting, Sculpture, Sight and Touch." *British Journal of Aesthetics* 44 (2004): 149–166.

Khatchadourian, Haig. "Movement and Action in the Performing Arts." *Journal of Aesthetics and Art Criticism* 37 (1978): 25–36.

John Dilworth (2005)

LANGUAGE

What is a "language"? Is it an internal component of a speaker's mind, or is it wholly dependent on our external behavior? Is it a matter of social practice, or are languages to be viewed as independently existing abstract objects? Arguments have been offered in favor of each of these conceptions.

Adherents to these different positions can agree that linguistic theories provide the most precise way of characterizing particular languages. A theory, or grammar, supplies a set of rules describing the semantic properties of the basic expressions and their permissible syntactic combinations into meaningful wholes. The disagree-

ments that arise concern the interpretation of linguistic theories and the nature of the linguistic objects and properties they describe.

Platonists, for instance, argue that languages are purely formal, or abstract entities, whose natures are fully specified by formal theories. For the Platonist, linguistics is a branch of mathematics. In contrast, mentalists see linguistics as a branch of cognitive psychology and take linguistic theories to be about the psychological states or processes of linguistically competent speakers. For others, linguistic theories can be seen as systematizing a vast range of facts about the behavior of an individual or community of speakers, with the rules describing regularities in individual or social practice.

For Platonists, such as J. Katz and S. Soames, languages with their properties of meaning and structure exist independently of speakers. A firm distinction is drawn between languages and linguistic competence: Theories of the former are not to be confused with theories of the latter. The formal properties of a language, on which its identity depends, owes nothing to its users. Speakers of those languages may be blind to some of its properties of meaning or structure, although these may be deduced from the theory. Moreover, languages with just these formal properties exist whether anyone speaks them or not. They may be defined, according to D. Lewis, as sets of expression meaning-pairs, with the set of human (or natural) languages making up a very small portion of the set of all possible languages. The task for Platonists is to explain what makes one rather than another of these abstract entities the language of a given individual or population. To explain this the Platonist must define an *actual-language relation* between speakers, or populations, and particular abstract objects (see Schiffer for discussion). This may depend, as Lewis thinks, on facts about the conventions that exist among a population of speakers. Or it may be based upon psychological facts about speakers' competence such as the claim that speakers have internalized a grammar that somehow generates either the set, or a subset, of the sentences described by the formal theory.

Mentalists, such as Noam Chomsky and Jerry A. Fodor, insist to the contrary that the best account of speakers' actual languages should fit the facts about the meanings and structures individuals actually give to expressions: Theories of language should be tailored to the contours of linguistic competence. Thus for Chomsky, a theory of language is a theory of a speaker's knowledge of language. The formal entities described by Platonists are just projections of the linguistic properties that speakers give to the expressions they produce and respond to. For mentalists, language is not in the world. The world contains only marks and sounds. Language is in the mind of speakers and consists in the assignments of meaning and structure given to particular marks and sounds.

For Chomsky, a grammar is a theory of the speaker's linguistic competence: An internalized system of rules or principles a person uses to map sounds to meanings. This is a body of tacit knowledge that the speaker puts to use in the production and comprehension of speech. It contains a largely innate, and species-specific, component common to all human language users. The workings of this component are described by universal grammar. Linguistic competence is just one of the factors affecting linguistic performance. Memory, attention, and other cognitive factors contribute to the actual production of speech. For Fodor, by contrast, the rules of grammar describe the actual psycholinguistic mechanisms at work in our production and comprehension of language. Language is just one of the perceptual modules, or sensory input systems, that serve our central cognitive processing.

In contrast to the Platonist and mentalist construals of language, behaviorists insist that grammars are merely theoretical representations of a speaker's practical abilities: The ability to use expressions in particular ways. For Willard Van Orman Quine, a language is a set of dispositions to verbal behavior. Quine argues that the only evidence for linguistic theory is linguistic behavior, and that many grammars will serve equally well to generate the set of sentences a speaker is disposed to produce and respond to. Thus grammars and the sentence structures they describe are construed as artifacts of theory. Chomsky denies that behavior provides the only evidence for testing theories of grammar. Psycholinguistic evidence and language acquisition are also relevant. He also argues that we could not have learned to produce and respond to so many novel sentences just on the basis of observed behavior. The data are too impoverished to support such inductive inferences: Sentences alike in surface structure differ in underlying levels of structure, and speakers respond to them differently. Chomsky concludes that speakers must bring their own internally generated representations of structure to bear on the evidence. Predictions of the sentences they find acceptable and unacceptable, and the interpretations they can and cannot allow, will be based on the fewest linguistic generalizations that fit the pattern of elicited data, and explain any gaps in the data. Claims about a speaker's grammar are thus based on inference to the best explanation about

the principles by which she generates structural descriptions (SDs) for the utterances she hears. A speaker's language is an internally generated set of structures.

Donald Davidson, like Quine, accepts that all facts about meaning must be exhibited in behavior. But unlike Quine, he holds that the assignments of meaning depend on facts about what the speaker believes and intends. Thus linguistic meaning cannot be reduced to behavior. The notions of belief and meaning are settled together by a total theory for interpreting what a speaker says and does.

Finally, is language an essentially social phenomenon? Michael Anthony Eardley Dummett argues that a language is a shared social practice upon which the possibility for communication among speakers depends. Lewis, although a Platonist, also argues that facts about the conventional regularities maintained by populations relate them to particular languages. Chomsky and Davidson, on the other hand, conclude that the fundamental notion of language is that of an individual's language, or idiolect. Differences in grammar and vocabulary between speakers ensure that no two speakers have exactly the same language: They can still communicate because there is often overlap in idiolects, and they can work out what others are saying. Chomsky distinguishes between E-languages, which are ill-assorted, externally described, and extensionally characterized social practices, and I-languages, which are the intensionally characterized, internalized grammars of individuals that assign SDs to expressions. For Chomsky, the former notion is ill-defined, so only the latter is of use in the scientific study of language. He argues that a language L cannot be identified apart from its structure, and the structure of L is the structure assigned to it by its speaker(s). He thus casts doubt on behaviorism. Many languages will share the same sounds: Whether a string of sounds is a sentence depends on how different speakers perceive those sounds. Relative to one structural assignment, the sound string may be grammatical, relative to another it may not. Quine's and Lewis's idea of a set of well-formed strings, which can be generated by different grammars, becomes problematic; instead we have a set of structures that speakers assign to sounds and signs. We might reconstruct the notion as follows: An E-language is a set of grammatical strings, where a "string" is grammatical if it has at least one structural description (SD), which is permitted by the I-language of some set of speakers in the sense that it conflicts with no principles of universal grammar (UG).

In the case of meaning, Tyler Burge has argued that word meaning depends on the social norms operating in the speaker's community; while Hilary Putnam stresses that the meaning and reference of natural terms are settled by a group of experts to whom ordinary speakers defer in their use of these terms. These social factors are compatible with the claim that the primary notion of language is that of an idiolect, as they concern vocabulary items only. Each of these different conceptions of language may coexist, all of them serving a different philosophical or scientific interest.

See also Behaviorism; Chomsky, Noam; Davidson, Donald; Dummett, Michael Anthony Eardley; Fodor, Jerry A.; Inference to the Best Explanation; Lewis, David; Meaning; Philosophy of Language; Putnam, Hilary; Quine, Willard Van Orman; Reference.

Bibliography

Burge, T. "Wherein Is Language Social." In *Reflections on Chomsky,* edited by A. George. Cambridge, MA: Blackwell, 1990.

Chomsky, N. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger, 1986.

Davidson, D. "Radical Interpretation." In his *Inquiries into Truth and Interpretation.* New York: Oxford University Press, 1984.

Dummett, M. "Language and Communication." In *Reflections on Chomsky,* edited by A. George. Cambridge, MA: Blackwell, 1990.

Fodor, J. "Some Notes on What Linguistics Is About." In *The Philosophy of Linguistics,* edited by J. Katz. New York: Oxford University Press, 1985.

Katz, J., ed. *The Philosophy of Linguistics.* New York: Oxford University Press, 1985.

Lewis, D. "Languages and Language." In his *Philosophical Papers,* Vol. 1. New York: Oxford University Press, 1983.

Putnam, H. "The Meaning of Meaning." In his *Philosophical Papers,* Vol. 2. New York: Cambridge University Press, 1975.

Quine, W. V. O. "Methodological Reflections on Current Linguistic Theory." In *Semantics of Natural Language,* edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.

Schiffer, S. "Actual-Language Relations." In *Philosophical Perspectives,* edited by J. Tomberlin. Atascadero, CA, 1973.

Soames, S. "Semantics and Psychology." In *The Philosophy of Linguistics,* edited by J. Katz. New York: Oxford University Press, 1985.

Barry C. Smith (1996)

LANGUAGE, ARTIFICIAL AND NATURAL

See *Artificial and Natural Languages*

LANGUAGE, PRIVATE

See *Private Language Problem*

LANGUAGE, RELIGIOUS

See *Religious Language*

LANGUAGE AND THOUGHT

Should questions about "thought"—about intentionality, beliefs, and concept possession, for example—be approached directly or, instead, indirectly via the philosophy of "language"? There are two slightly different ways in which questions about language and meaning might seem to offer illumination of issues concerning thought. One way relates to language that is explicitly about thoughts, as when someone says, "Bruce believes that boomerangs seldom come back." The idea that a philosophical investigation of thought should proceed via a study of the logical properties of language that is about thoughts is a particular case of a more general view that philosophy of language enjoys a certain priority over metaphysics.

The other way relates to the use of language to express thoughts, and this provides the topic for the present entry. Suppose that Bruce believes that boomerangs seldom come back, and expresses this thought in the English sentence: "Boomerangs seldom come back." Which takes priority, the meaning of the English sentence or the content of Bruce's thought?

A claim of priority is the converse of a claim of one-way dependence: X enjoys priority over Y if Y depends on X but X does not depend on Y. Thus, a question of the relative priority of X and Y has four possible answers: X has priority; Y has priority; X and Y are mutually dependent; X and Y are independent. But the question of the relative priority of thought and language is still unclear, until the relevant kind of priority has been specified. It is useful to distinguish three kinds of priority question: ontological, epistemological, and analytical (see Avramides 1989 for a similar distinction).

To say that thought enjoys *ontological priority* over language is to say that language is ontologically dependent on thought, while thought is not so dependent on language. That is, there can be thought without language, but there cannot be language without thought. To say that thought enjoys *epistemological priority* over language is to say that the route to knowledge about language (specifically, about linguistic meaning) goes via knowledge about thought (specifically, about the contents of thought), while knowledge about thought can be had without going via knowledge about language.

Donald Davidson denies both these priority claims. As for ontological priority, he argues (1975) that there cannot be thought without language: In order to have thoughts (specifically, beliefs), a creature must be a member of a language community, and an interpreter of the speech of others. As for epistemological priority, Davidson argues (1974) that it is not possible to find out in detail what a person believes without interpreting the person's speech.

Analytical priority is priority in the order of philosophical analysis or elucidation. To say that X is analytically prior to Y is to say that key notions in the study of Y can be analyzed or elucidated in terms of key notions in the study of X, while the analysis or elucidation of the X notions does not have to advert to the Y notions. On the question of the relative analytical priority of thought and language, there are, then, four positions to consider: two priority views, and two no-priority views.

PRIORITY FOR THOUGHT

A philosophical account of the content of thoughts—of intentionality—can be given without essential appeal to language, and the notion of linguistic meaning can then be analyzed or elucidated in terms of the thoughts that language is used to express. The analytical program of Paul Grice was aimed at an analysis of linguistic meaning in terms of the beliefs and intentions of language users, though Grice did not offer any account of the intentionality of mental states themselves (Grice 1989; see also Schiffer 1972). There are many proposals for explaining the intentionality of mental states without appeal to linguistic meaning, and these might be coupled with an elucidation of linguistic meaning in terms of mental notions. It is widely reckoned, however, that the Gricean analytical program cannot be carried through (Schiffer 1987).

PRIORITY FOR LANGUAGE

An account of linguistic meaning can be given without bringing in the intentionality of thoughts, and what a person's thoughts are about can then be analyzed in terms of the use of language. This view can be found in Michael Dummett's work (1973, 1991, 1993). If a theorist attempts to give a substantive account of linguistic mean-

ing in accordance with this view, then the resources that can be invoked are seriously limited, since the account cannot presume upon everyday psychological notions such as belief and intention. Because of this, it would not be surprising to find hints of behaviorism in work that is influenced by this view.

NO PRIORITY—INTERDEPENDENCE

There is no way of giving an account of either intentionality or linguistic meaning without bringing in the other member of the pair. The two notions have to be explained together. This is Davidson's view (Davidson 1984). He thus maintains an ontological, epistemological, and analytical no-priority position. While the three no-priority claims go together quite naturally, it is important to note that they are separable claims and that the analytical no-priority claim is not entailed by the ontological and epistemological no-priority claims.

NO PRIORITY—INDEPENDENCE

The notions of intentionality for mental states and of linguistic meaning are unrelated. This view might be defended if a language is considered as an abstract entity, composed of a set of expressions together with a function that assigns a value to each expression (a proposition to each sentence, for example). On such a conception, meaning is a purely formal notion. But for the notion of linguistic meaning as it applies to a public language in use, this fourth view is implausible.

See also Behaviorism; Davidson, Donald; Dummett, Michael Anthony Eardley; Grice, Herbert Paul; Intentionality; Language; Meaning; Philosophy of Language.

Bibliography

Avramides, A. *Meaning and Mind: An Examination of a Gricean Account of Language.* Cambridge, MA: MIT Press, 1989.

Davidson, D. "Belief and the Basis of Meaning." *Synthese* 27 (1974): 309–323.

Davidson, D. *Inquiries into Truth and Interpretation.* Oxford: Clarendon Press, 1984.

Davidson, D. "Thought and Talk." In *Mind and Language,* edited by S. Guttenplan. Oxford: Clarendon Press, 1975.

Dummett, M. *Frege: Philosophy of Language.* London: Duckworth, 1973.

Dummett, M. *The Logical Basis of Metaphysics.* Cambridge, MA: Harvard University Press, 1991.

Dummett, M. *The Seas of Language.* Oxford: Clarendon Press, 1993.

Dummett, M. *Truth and Other Enigmas.* London: Duckworth, 1978.

Grice, H. P. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press, 1989.

Schiffer, S. *Meaning* (1972). Oxford, 1988.

Schiffer, S. *The Remnants of Meaning.* Cambridge, MA: MIT Press, 1987.

Sellars, W. S., and R. M. Chisholm. "Intentionality and the Mental." In *Minnesota Studies in the Philosophy of Science,* edited by H. Feigl, M. Scriven, and G. Maxwell, Vol. 2: *Concepts, Theories, and the Mind-Body Problem.* Minneapolis: University of Minnesota Press, 1958.

Martin Davies (1996)

LANGUAGE OF THOUGHT

Simply stated, the language-of-thought thesis (LOT) holds that thinking (i.e., cognition) is carried out in a languagelike medium, where the thoughts that constitute thinking are themselves sentencelike states of the thinker. Since the demise of philosophical behaviorism in the early 1960s the LOT thesis has enjoyed considerable support as a central tenet of a more encompassing representationalist theory of mind (RTM). Proponents of RTM, led by Jerry Fodor, have mounted a sustained defense of LOT.

RTM offers an account of propositional attitudes—beliefs, desires, doubts, and so on—according to which propositional attitudes relate the possessor of the attitude to a mental representation (cf. Fodor 1981). Mental representations have both semantic and physically realized formal properties: They are semantically evaluable (e.g., as being true or false, as being about or referring to certain entities or properties); they stand in inferential relations to other mental representations; and, like words, pictures, and other representations, they also have certain formal properties (e.g., shape, size, etc.) in virtue of being physical, presumably neural, entities. Mental representations, and hence propositional attitudes, have their causal roles in thinking and behavior in virtue of their formal properties. Propositional attitudes inherit semantic properties from the mental representations that are one of their relata. RTM is silent as to what kind or sort of representation these mental representations are (cf. Fodor 1987, pp. 136–138).

LOT supplements RTM with a specific proposal or hypothesis about the character of mental representations: Like sentences of a language, they are structured entities, and their structures provide the basis for the particular semantic and causal properties that propositional attitudes exhibit. More specifically, they are syntactically structured entities, composed of atomic constituents

(concepts) that refer to or denote things and properties in the world. The semantic properties of a mental representation, including both truth conditions and inferential relations, are determined by the representation's syntactic structure together with the semantic properties of its atomic constituents. Mental representations, in other words, have a combinatorial semantics. The causal properties of a representation are similarly determined by the representation's syntactic structure together with the formal properties of its atomic constituents.

Three sorts of arguments have been advanced in support of LOT. The first makes much of the apparent semantic parallels between thoughts and sentences. Both beliefs and declarative sentences, for example, are typically meaningful, truth valued, and intentional (in the sense of being about something). Both stand in various inferential relations to other beliefs and assertions. One obvious explanation of these parallels is that thought has a languagelike character, individual thoughts a sentencelike structure. A second sort of argument focuses on the productivity and systematicity of thought. Thought, like language, is productive in the sense that there are indefinitely many, indefinitely complex thoughts. Whatever can be said can also be thought. Thought, like language, is also systematic in the sense that you can think one thought (e.g., that the child bit the monkey) if and only if one can also think certain other systematically related thoughts (that the monkey bit the child). Again, one obvious explanation is that thought has a languagelike character, individual thoughts a sentencelike structure. A third sort of argument claims that much cognitive scientific theorizing seems committed to LOT. Specifically, our best theories of rational choice, perception, and learning seem committed to the claim, not simply that cognition is a matter of the creation and manipulation of mental representations, but also that these representations are sentential in character. It is claimed, for example, that our best theories of learning are a species of hypothesis testing. But such a procedure, it is argued, presupposes the existence of a language, that is, a language of thought in which the hypothesis being tested is formulated.

Proponents of LOT readily concede that these arguments are not decisive. Each is an instance of inference to the best explanation, and as such each is vulnerable to refutation by some alternative explanation that does not appeal to a language of thought.

Critics of LOT find the foregoing sorts of arguments unpersuasive for any of a number of reasons. Either they believe that there are equally good explanations that do not appeal to a language of thought, or they deny the phenomena that LOT is said to explain, or they hold that the proposed explanations either rest on false presuppositions or are so sketchy and incomplete as not to merit the name, or they believe that these explanations have entailments so implausible as to impugn the explanatory premise that there exists a language of thought. Thus, for example, the argument from learning discussed above apparently entails that to learn a language one must already know a language. Many critics find in this entailment a reductio of LOT. Proponents such as Fodor, by contrast, have courageously embraced this entailment, arguing that all concepts, including, for example, our concept of a Boeing 747, are innate. Whatever the specific merits and defects of the arguments and counterarguments, it seems fair to say that the existence of a language of thought remains an open empirical question.

See also Behaviorism; Fodor, Jerry A.; Inference to the Best Explanation; Mental Representation; Philosophy of Mind; Propositional Attitudes: Issues in the Philosophy of Mind and Psychology.

Bibliography

DEFENDING LOT

Field, H. "Mental Representations." *Erkenntnis* 13 (1978): 9–61.

Fodor, J. *The Language of Thought.* New York: Crowell, 1975.

Fodor, J. *Representations.* Cambridge, MA: MIT Press, 1981.

Fodor, J. "Why There Still Has to Be a Language of Thought." In *Psychosemantics.* Cambridge, MA: MIT Press, 1987.

CRITICIZING LOT

Churchland, P. M., and P. S. Churchland. "Stalking the Wild Epistemic Engine." *Noûs* 17 (1983): 5–18.

Dennett, D. C. "Brain Writing and Mind Reading." In *Brainstorms.* Montgomery, VT: Bradford Books, 1978.

Dennett, D. C. "A Cure for the Common Code." In *Brainstorms.* Montgomery, VT: Bradford Books, 1978.

Loar, B. "Must Beliefs Be Sentences?" In *Proceedings of the PSA, 1982,* edited by P. Asquith and T. Nickles. East Lansing, MI: Philosophy of Science Association, 1983.

Schiffer, S. *Remnants of Meaning.* Cambridge, MA: MIT Press, 1987.

ADDITIONAL SOURCES

Aydede, M. "Language of Thought: The Connectionist Contribution." *Minds and Machines* 7 (1997): 57–101.

Barsalou, L. "Perceptual Symbol Systems." *Behavioral and Brain Sciences* 22 (1999): 507–569.

Cowie, F. *What's Within? Nativism Reconsidered.* Oxford: Oxford University Press, 1998.

Devitt, Michael. *Coming to Our Senses: A Naturalistic Program for Semantic Localism.* Cambridge, U.K.: Cambridge University Press, 1996.

Fodor, J. *Concepts: Where Cognitive Science Went Wrong.* Oxford: Oxford University Press, 1998.

Fodor, J. "Doing without What's Within: Fiona Cowie's Critique of Nativism." *Mind* 110 (2001): 99–148.
Knowles, J. "The Language of Thought and Natural Language Understanding." *Analysis* 58 (1998): 264–272.
Kosslyn, S. *Image and Brain.* Cambridge, MA: MIT Press, 1994.
Laurence, S., and E. Margolis. "Radical Concept Nativism." *Cognition* 86 (2002): 22–55.
Laurence, S., and E. Margolis. "Regress Arguments against the Language of Thought." *Analysis* 57 (1997): 60–66.
Leeds, S. "Perception, Transparency, and the Language of Thought." *Noûs* 36 (2002): 104–129.
Prinz, J. *Furnishing the Mind: Concepts and Their Perceptual Basis.* Cambridge, MA: MIT Press, 2002.
Rey, G. *Contemporary Philosophy of Mind: A Contentiously Classical Approach.* Oxford: Blackwell, 1997.
Vinueza, A. "Sensations and the Language of Thought." *Philosophical Psychology* 13 (2000): 373–392.

Robert J. Matthews (1996)
Bibliography updated by Alyssa Ney (2005)

LAO TZU

See *Laozi*

LAOZI

(sixth century BCE)

Laozi, according to the *Records of History* by Sima Qian, is believed to have been an elder contemporary of Confucius (551–479 BCE) and the author of the *Laozi* (*Daode jing* or *Tao-te-ching*), a work roughly five thousand characters long. This and other traditional accounts of Laozi and the date of his work have been seriously challenged, and various hypotheses about the authorship of the work and its date have been proposed. Nevertheless, three incomplete Guodian bamboo versions of the *Laozi* excavated in 1993 prove that the text was in circulation in the fourth century BCE and may have been composed still earlier.

Laozi is believed to be the first person in Chinese intellectual history to develop a brief theory on the source and grounds of the universe, represented by the concept of Dao (also commonly called Tao in Western writings).

DAO: SOURCE AND GROUNDS

Dao literally means "the way" and was often extended to cover the political or moral principles by which different schools expounded their ideas. Laozi attributed to this term a totally new meaning: "Dao produced the One, One produced the Two, Two produced the Three, and Three produced the ten thousand things" (chapter 42). Here, the One, Two, and Three do not indicate anything specific, just a general cosmological formula: from Nothing to Being, one to multitude, and simple to complex. This formula has been compared to the Big Bang theory of modern astrophysics. Dao is the primordial root of all beings and creatures, and all beings and creatures in turn depend on it. As the ultimate source and grounds of the universe, Dao would be termed a metaphysical, as opposed to an empirical, concept in European philosophy. But in Chinese philosophy there is no dichotomy between the metaphysical and the physical, the ontological and the axiological, the descriptive and the prescriptive, and so on. Dao runs through the whole universe and human life; it is both transcendent and immanent. As the model for human behavior and the object of ultimate concern for human beings, Dao is similar to God, but has nothing to do with will, feelings, or purpose. Dao runs through and embodies "ten thousands things," and *de* (power or virtue) is in each being. It can be said that Dao is a quasi-metaphysical concept, and de is its manifestation in all beings.

ZIRAN: THE CORE VALUE

The second key concept in Laozi's philosophy is *ziran*, or naturalness. Laozi advocates that "Man models himself on the earth, the earth models itself on heaven, heaven models itself on the Dao, and the Dao models itself on *ziran*" (chapter 25). The true meaning and message of this statement is that humans should practice the principle of naturalness, which involves allowing things to unfold without external coercion or, in the case of individual humans, without striving for things such as wealth and power. This permits actualizing natural harmony in human life and with one's surroundings. The word *ziran* comprises two parts: self (*zi*) and so (*ran*). Its basic meaning is "self-so." It may be rendered as naturalness to show its adjectival meaning and grammatical function as a noun.

One should not confuse ziran with Nature or the natural world. Ziran is used to indicate Nature in modern Chinese, but in classical contexts words such as *tian* (heaven), *di* (earth), and *wanwu* (ten thousand things) denoted the natural world. Some scholars relate ziran with Thomas Hobbes's (1588–1679) "state of nature," which is a hypothetical term for scientific argument and suggests that everyone is at war with everyone. Instead, Laozi's ziran is the ideal condition of human societies, namely natural harmony, and represents the highest principle and core value in his philosophy; it is embodied and

promoted by Dao. Natural harmony and order are valuable and desirable compared with humanly contrived order, which depresses human nature and arouses resistance and even inevitably leads to chaos. Human nature can only flourish within societies that have natural order, hence ziran is also the optimal condition of individuals. Laozi contends that "the sage should foster the *ziran* of the ten thousand things and dare not take action" (chapter 64). This leads us to the next fundamental concept: *wuwei*.

WUWEI: PRINCIPLED METHOD

Wuwei also comprises two parts: no (*wu*) and action (*wei*). Superficially, it means "no action at all," but in fact wuwei only negates some kinds of actions, not all. Obviously, "fostering the *ziran* of ten thousand things" is not the kind of action wuwei would exclude. The agent of wuwei in Laozi's theory is mainly the sage, the ideal model of rulers, who fosters potential in others instead of directly ordering, forcing, interfering, and interrupting. So there is a social and political message in Laozi that is absent in the Inner Chapters of the *Zhuangzi*. Laozi's wuwei implies two aspects: Its negative expression suggests preventing certain societal actions, such as oppression, confrontation, and strife, while its positive meaning advocates an alternative, sophisticated manner of behavior for better results of natural development and harmony in societies, such as fostering, assisting, and being patient. In his famous proclamation about "doing nothing yet leaving nothing undone," Laozi clearly reveals the positive objective of wuwei. "Doing nothing" is a means of realizing the end of "leaving nothing undone." Wuwei actually purports to be both a superior approach to and the consummate realization of human activity. It derives from comprehensive humanistic perspectives and considerations, not from fashions or trends of governance aimed at achieving immediate benefits.

Humans make two kinds of mistakes: One is not making enough effort, the other is overdoing. The former mistake is easy to remedy because it does not waste too many resources or shake morale. Correcting the second is more difficult, as in the case of environmental degradation. Here is an additional sense in which wuwei is reasonable and significant.

REVERSION: PARADOXICAL THINKING

Another distinctive feature of Laozi's philosophy is his dialectical or paradoxical thinking, which emerges through doctrines dealing with the unity and transformation of pairs of contradictions. One doctrine concerns the interdependence of opposite things and concepts. For example, "Calamity is that upon which happiness depends; happiness is that in which calamity is latent." Another is the reversibility of opposite sides, such as the "correct can become the perverse, and good may become evil." According to Laozi, all things are in motion and they are changing and proceeding toward their reverse. Thus, humility produces greatness, and ambitions bring about failure. Obverse and reverse sides often exchange positions. Things in both human societies and the natural world can work out to be the very opposite of our expectation and intention.

To sum up, wuwei is the methodological principle for fostering ziran, the core value in Laozi's system. Dao, as the ultimate source and grounds of the universe, is the quasi-metaphysical and axiological foundation for both wuwei and ziran, while the theory of dialectics supports ziran and wuwei from the perspective of human experience.

See also Chinese Philosophy: Daoism.

Bibliography

Chan, Wing-tsit. "The Natural Way of Lao Tze." In *A Source Book in Chinese Philosophy*, 136–176. Princeton, NJ: Princeton University Press, 1963. This includes a translation of the received version of the *Laozi*.

Graham, A. C. *Disputers of the Tao: Philosophical Argument in Ancient China*. La Salle, IL: Open Court, 1989.

Hansen, Chad. *A Daoist Theory of Chinese Thought: A Philosophical Interpretation*. New York: Oxford University Press, 1992.

Henricks, Robert G. *Lao Tzu's Tao Te Ching: A Translation of the Startling New Documents Found at Guodian*. New York: Columbia University Press, 2000. This is the translation of the bamboo versions excavated in 1993.

LaFargue, Michael. *Tao and Method: A Reasoned Approach to the Tao Te Ching*. Albany: State University of New York Press, 1994.

Laozi. *Tao Te Ching*. Translated by D. C. Lau. Hong Kong: Chinese University Press, 1989. This is the translation of silk versions excavated in 1973.

Liu, Xiaogan. "Naturalness (tzu-jan), the Core Value in Taoism: Its Ancient Meaning and Significance Today." In *Lao-tzu and the Tao-te-ching*, edited by Livia Kohn and Michael LaFargue, 211–228. Albany: State University of New York Press, 1998.

Liu, Xiaogan. "Non-action and the Environment Today: A Conceptual and Applied Study of Laozi's Philosophy." In *Daoism and Ecology: Ways within a Cosmic Landscape*, edited by N. J. Girardot, James Miller, and Liu Xiaogan, 315–340. Cambridge, MA: Center for the Study of World Religions, Harvard Divinity School, 2001.

Slingerland, Edward. *Effortless Action: Wu-wei as Conceptual Metaphor and Spiritual Ideal in Early China*. New York: Oxford University Press, 2003.

Xiaogan Liu (2005)

LA PEYRÈRE, ISAAC

(1596–1676)

Isaac La Peyrère, or Pereira, was born in Bordeaux, France, a Calvinist of Portuguese New Christian, or converted Jewish, background. He became the Prince of Condé's secretary. Apparently he was friendly with leading Parisian avant-garde intellectuals such as Pierre Gassendi, François de La Mothe Le Vayer, Hugo Grotius, Guy Patin (1601–1672), and Ménage. La Peyrère's first book, *Du rappel des juifs* (1643), deals with the conversion of the Jews, their potential return to Palestine, and the beginning of the Messianic Age. In 1644 he went to Denmark and gathered there the material for his *Relation de l'Islande* (1663) and *Relation du Groenland* (1647), both written as letters to La Mothe Le Vayer. His most famous works, the *Prae-Adamitae* and the *Systema Theologicum ex Prae-Adamitarum Hypothesi*, apparently written by 1643, were published in Amsterdam in 1655. Queen Christina of Sweden (1626–1689), whom he had recently met in Belgium, offered to pay the publishing expenses, so he took his manuscript to Amsterdam to get a printer. Five editions of these works were published almost immediately, and the book appeared not only in Latin but in English and Dutch. Among the early readers of his book was the young Spinoza. La Peyrère argued that the only consistent interpretation of certain biblical passages, and of the anthropological and historical evidence about the Chinese, Mexicans, Eskimos, and other peoples, is that there were men before Adam and that the Bible deals only with Jewish history and not world history. The effect of this work was like that of a bombshell to the seventeenth-century intellectual world. It appeared at almost the same time as Archbishop James Ussher's (1581–1656) proof, on the basis of biblical data, that the world was created in 4004 BCE. La Peyrère was immediately attacked and refuted on all sides. His book was burned in Paris, and he himself was arrested and kept in prison in Belgium for six months until he retracted his views and became a Catholic. He then went to Rome and begged the pope's forgiveness, publishing a formal retraction of his views. In 1659 he entered a religious order near Paris, where he remained until his death. Despite his official retractions, it is believed that he continued to hold to his pre-Adamite views. For example, Pierre Bayle cites a letter in which La Peyrère's religious superior is supposed to have said that "he was always writing books that … would be burned as soon as the good man died. La Peyrère was the best man in the world, the sweetest, who tranquilly believed very little."

La Peyrère's revolutionary work on the pre-Adamite theory had tremendous influence on seventeenth- and eighteenth-century thought. In raising the possibility that biblical data might only apply to Jewish history, he introduced a radical new conception of human development and led people to speculate on the relative merits of various cultures and religions. Further anthropological and geological studies, as well as investigations into comparative religion, soon led to the abandonment of biblical chronology and history as the framework for understanding all human history and led also to the beginning of higher criticism of the Bible by writers like Spinoza and Richard Simon and to the Enlightenment critiques of traditional religion. Pre-Adamism was a radical hypothesis in the Enlightenment that accounted for the variety of human beings.

Most writers for at least a century after La Peyrère seem to have been directly or indirectly aware of his pre-Adamite hypothesis and its extraordinary implications. In the nineteenth century La Peyrère's pre-Adamite hypothesis was developed into a racist view and finally into the ideology of British Israelites and of the Aryan Nation in the United States.

See also Bayle, Pierre; Continental Philosophy; Deism; Enlightenment; Gassendi, Pierre; Grotius, Hugo; La Mothe Le Vayer, François de; Philosophy of History; Simon, Richard; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY LA PEYRÈRE

Du rappel des juifs. 1643.

Relation de Groenland. Paris: A. Covrbe, 1647.

Prae-Adamitae. Amsterdam, 1655. Translated as *Men Before Adam*. London, 1956.

Relation de l'islande. Paris: Iolly, 1663.

WORKS ABOUT LA PEYRÈRE

Allen, D. C. *The Legend of Noah: Renaissance Rationalism in Art, Science, and Letters*. Urbana: University of Illinois Press, 1963.

Bayle, Pierre. "Peyrère, Isaac La." In *Dictionnaire historique et critique*. Rotterdam, Netherlands: Reinier Leers, 1695–1697.

Katz, David S., and Richard H. Popkin. *The Messianic Revolution: Radical Religious Politics to the End of the Second Millennium*. New York: Hill and Wang, 1998. See especially chapter 7.

McKee, David R. "Isaac La Peyrère, a Precursor of Eighteenth-Century Critical Deists." *PMLA* 59 (1944): 456–485.
Popkin, Richard H. *Isaac La Peyrère (1596–1676): His Life, Work and Influence*. Leiden, Netherlands: Brill, 1987.

Richard Popkin (1967, 2005)

LAPLACE, PIERRE SIMON DE

(1749–1827)

Pierre Simon de Laplace, the French astronomer and mathematician famous for his celestial mechanics and theory of probability, was born in Normandy. Upon coming to Paris, he attracted the attention of Jean Le Rond d'Alembert, who found him employment in the École Militaire. Here he taught mathematics to trainee artillery officers, among whom was Napoleon Bonaparte. When the revolutionary government established the École Polytechnique, Laplace was one of its founding professors. He served with distinction on many of the great committees of the French Academy of Sciences and of the government. He helped devise the meter, standardized weights and measures, and worked out an ingenious system of sampling to provide an economical and efficient census. The elegance of his mathematical work has yet to be rivaled, and his power of analysis is matched only by that of Isaac Newton and Joseph-Louis Lagrange. His philosophical opinions, especially those in his *Exposition du système du monde* (*The System of the World*) and *Essai philosophique sur les probabilités* (*A Philosophical Essay on Probabilities*), have a bluntness and clarity of expression that ensured their popularity.

Laplace's adult life was passed in conditions of civil strife and sometimes of chaos, but despite his revolutionary affiliations, the restoration of the Bourbons brought him neither poverty nor disgrace; he died honored by all, a newly created marquis. Against this background of political confusion, he came to believe that the theory of probability, properly and widely applied, would reduce most of the problems of society (like the attainment of justice) to something manageable; with the help of probability theory, he believed, a man of delicate intuition and wide experience could find practical solutions to most social difficulties.

Laplace's scientific work had a strong element of tidiness about it. It consisted largely of the final polishing of the Newtonian enterprise, knitting up its loose ends. Using the improved calculus devised by his colleagues, particularly Lagrange, he removed all known errors from, and explained all known anomalies in, the Newtonian cosmology and physics. It seemed to Laplace that there was no phenomenon that the improved and polished Newtonian physics was incapable of handling. He came to regard the enormous explanatory power of the system as practically a demonstration of its truth. New observations would only confirm it further, he thought, and their consequences were as certain as if they had already been observed.

What had produced this remarkable confidence was a series of complete successes. Newton had never been convinced of the stability of the solar system, which he suggested might need divine correction from time to time. Laplace showed, in effect, that every known secular variation, such as the changing speeds of Saturn and Jupiter, was cyclic and that the system was indeed entirely stable and required no divine maintenance. (It was this triumph that occasioned his celebrated reply to Napoleon's query about the absence of God from the theory; Laplace said that he had no need of that hypothesis.) He also completed the theory of the tides and solved another of Newton's famous problems, the deduction from first principles of the velocity of sound in air. Laplace added a very accurately estimated correction for the heating effect produced by rapidity of the oscillation, which was too short to allow the heat of compression to be dissipated.

DETERMINISM AND PROBABILITY

Not only was Laplace confident of the Newtonian theory, but he was also greatly struck by its determinist nature. Where one could gather accurate information about initial conditions, later states of a mechanical system could be deduced with both precision and certainty. The only obstacle to complete knowledge of the world was ignorance of initial conditions. Laplace's confidence in Newtonian theory is exemplified in the introduction to his *Philosophical Essay on Probabilities,* in which he envisaged a superhuman intelligence capable of grasping both the position at any time of every particle in the universe and all the forces acting upon it. For such an intelligence "nothing would be uncertain and the future, as the past, would be present to its eyes. The human mind offers, in the perfection which it has been able to give to astronomy, a feeble idea of this intelligence" (*Philosophical Essay,* p. 4).

But this ideal is difficult to attain, since we are frequently ignorant of initial conditions. The way to cope with the actual world, Laplace thought, is to use the theory of probability. The superhuman intelligence would

have no need of a theory of probability. Laplace would have regarded as ridiculous the idea that there could be systems that would react to stimuli in only more or less probable ways. He said, "The curve described by a simple molecule of air or vapor is regulated in a manner just as certain as the planetary orbits; the only difference between them is that which comes from our ignorance" (*Philosophical Essay,* p. 6). He then defined a measure of probability as follows:

> The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible … and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible. (*Philosophical Essay,* p. 6)

This is the definition of probability known today as the proportion of alternatives. Then as now, it involves the very tricky notion of equipossible cases. Laplace deals with this notion by glossing equipossible cases as those that "we may be equally undecided about in regard to their existence" (*Philosophical Essay,* p. 6).

This account does have its difficulties. Equal indecision is not at all easy to determine and may, in the end, hinge upon states of mind quite irrelevant to a sound estimate of probabilities. Throughout his study of probability Laplace refers to such subjective factors as honesty, good judgment, and absence of prejudice, which are required in using probability theory. However, he does give a much sounder criterion for its practice; it encourages one to reckon as equally possible those kinds of events instances of which we have no special reason to believe will occur. Equality of ignorance then becomes his criterion for equality of possibility. Laplace is quite happy about this, since he believed—perhaps rightly—that the proper occasion for the recourse to probability is ignorance of the initial conditions, the relevant theory, or both. Actual estimates of probability are made statistically. In his practical examples he appears to depend on a further distinction, which also seems correct. It is the distinction between the meaning of the statement of probability for a certain kind of event (that is, ratio of number of favorable to equipossible kinds of events) and the usual estimate of this probability, which is the relative frequency of actual events of the kind under consideration among all appropriate cases.

APPLICATIONS OF PROBABILITY

Laplace made several practical applications of probability theory. In science he applied it to the problem of sampling for the census and to the theory of errors; to both of these studies he made valuable contributions. He also believed that probability theory would have great utility in the moral sciences. He studied the optimum size for a jury to give the least doubtful verdict and the voting procedures of assemblies both on candidates for office and on propositions. He discussed the advantages and disadvantages of voting by ranking in order of merit and of voting by the knockout majority system. In this study and in his reflections on what it is reasonable to risk and in what kind of game, one gets the occasional glimpse of Laplace's basic moral principle, "Only bet on a reasonably sure thing."

PHILOSOPHY OF SCIENCE

In his philosophy of science and in his views on the nature of scientific method, Laplace expressed himself somewhat along the same lines as Newton, but more liberally. He saw quite clearly that science is not the accumulation of isolated and particular items of information. "It is by comparing phenomena together, and by endeavouring to trace their connection with each other, that he [man] has succeeded in discovering these laws, the existence of which may be perceived even in the most complicated of their effects" (*System,* Vol. I, p. 205). In searching for connections we do not need to shun hypotheses. Laplace said of hypotheses what Newton should have said, considering the use he made of them: that if we refuse to attribute them to reality and regard them merely as the means of connecting phenomena in order to discover the laws (which we correct according to further observations), they can lead us to the real causes or at least enable us to infer from observed phenomena those which given conditions ought to produce.

In fact, it is by excluding on the basis of decisive experiments all those hypotheses that are false that "we should arrive … at the true one." Ideally, Laplace sees scientific method as the formulation of generalizations of connection between phenomena, proceeding inductively from phenomena to laws (which are the ratios connecting particular phenomena), and from these to forces. When these forces reveal some general principle, that principle is verified by direct experience, if possible, or by examination of its agreement or disagreement with known phenomena.

Testing consists both of trying to formulate a deductive system based upon the highest hypotheses and

designed to explain the phenomena, "even in their smallest details," and of seeing whether the theory agrees with as varied and as numerous phenomena as are relevant to it. If a theory passes these tests, it "acquires the highest degree of certainty and of perfection that it is able to obtain."

Laplace saw that our confidence in predictions had to be based upon confidence in some principle of the uniformity of nature. The sources of his confidence in some principle of uniformity were twofold. First, there is the condition of the absence of interference. If there is no reason why a change should occur, a change will not occur—a principle deeply embedded in Newtonian science. As Laplace put it, "Being assured that nothing will interfere between these causes and their effects, we venture to extend our views into futurity, and contemplate the series of events which time alone can develop" (*System,* Vol. I, p. 206). Second, simplicity was to be regarded as a mark of future reliability. The principle of induction, said Laplace, is that "the simplest ratios are the most common." He said, too, "We judge by induction that if various events, movements for example, appear constantly and have been long connected by a simple ratio, they will continue to be subjected to it" (*Philosophical Essay,* p. 178). The theory of probability supplies a connection between the two sources of confidence, for, said Laplace, we conclude from the fact that a simple ratio is found among quantities in nature "that the ratio is due, not to hazard, but to a regular cause." Thus, if no other causes intervene, we may expect a likeness of effects, in fact, a uniformity of nature.

Summing up scientific method, Laplace said, "Induction, analogy, hypotheses founded upon facts and rectified continually by new observations, a happy tact given by nature and strengthened by numerous comparisons of its indications with experience, such are the principal means for arriving at truth" (*Philosophical Essay,* p. 176).

See also Alembert, Jean Le Rond d'; Determinism, A Historical Survey; Induction; Newton, Isaac; Probability; Scientific Method.

Bibliography

WORKS BY LAPLACE

Oeuvres complètes. 14 vols. Paris: Gauthier-Villars, 1878–1914.

Exposition du système du monde. Paris, 1798. Translated by J. Pond as *The System of the World.* London: R. Phillips, 1809. Contains two important conjectures: that the planets might have been formed by the condensation of a large, diffuse solar atmosphere as it contracted, and the hypothesis, since confirmed, that the nebulae are clouds of stars and that the Milky Way is our view of that nebula of which our sun is a star.

Traité de la mécanique céleste, 5 vols. Paris: Chez J.B.M. Duprat, 1799–1825. Translated by N. Bowditch as *Mécanique Céleste,* 4 vols. Boston: Hillard, Gray, Little, and Wilkins, 1829–1839.

Théorie analytique des probabilités. Paris: Courcier, 1812.

Essai philosophique sur les probabilités. Paris: Courcier, 1814. Translated by F. W. Truscott and F. L. Emory as *A Philosophical Essay on Probabilities.* New York: Wiley, 1902; New York: Dover, 1951. A semipopular introduction to the *Théorie analytique.*

WORKS ON LAPLACE

Andoyer, H. *L'oeuvre scientifique de Laplace.* Paris: Payot, 1922.

Whittaker, Sir E. "Laplace." *Mathematical Gazette* 33 (303) (1949): 1–12.

R. Harré (1967)

LAPSHIN, IVAN IVANOVICH

(1870–1952)

Ivan Ivanovich Lapshin, the Russian neo-Kantian philosopher, was born in Moscow and studied at the University of St. Petersburg under the leading Russian neo-Kantian, Aleksandr Vvedenskii. Lapshin pursued his studies abroad for some years after 1893, concentrating particularly on Kantianism in English philosophy. With the publication in 1906 of his dissertation and chief philosophical work, *Zakony myshleniia i formy poznaniia* (The laws of thought and the forms of cognition), he received his doctorate from the University of St. Petersburg and in 1913 was made professor of philosophy at that institution. Along with many other noted Russian scholars Lapshin was exiled from the Soviet Union in 1922; he settled in Prague, where he lived until his death. His many writings cover a broad range of topics in philosophy, psychology, literature, music, and art, and include Russian translations of works by William James.

In his chief work Lapshin developed an antimetaphysical position on Kantian grounds, arguing specifically that the "laws of thought" derive their necessity solely from their connection with the forms through which sensory objects are cognized and that, therefore, it cannot be known whether these laws apply beyond the bounds of possible experience. According to Lapshin the law of contradiction, for example, can be understood only in reference to space and time (which, contrary to Immanuel Kant, he held to be categories of the understanding rather than forms of sensibility); and since the categories of space and time do not necessarily apply to transempirical objects, neither does the law of contradic-

tion. Consequently nothing can legitimately be affirmed of "things in themselves," not even their existence.

Lapshin devoted little attention to problems of ethics and did not accept Kant's transition to a noumenal realm and to religious faith via the dictates of moral consciousness. In general he regarded metaphysics and religion as entirely without epistemological foundation and as obstacles to the progress and vitality of human thought.

Much of Lapshin's later philosophical work was concerned with questions of the psychology of creativity and with the epistemological basis of our knowledge of other minds. His two-volume study of creativity in philosophy (1922) was complemented by a number of other writings on creativity in literature and the arts.

As early as 1910 Lapshin had published a historical account of the problem of other selves, and in 1923 he presented his resolution of the problem in the article *Oproverzhenie solipsizma* (A refutation of solipsism). He argued that our sense of the immediate giveness of other selves is an illusion based on the projection of subjective impressions; other selves are hypothetical constructs, which can be called "immanently real" but cannot be shown to have transcendent reality.

See also James, William; Kantian Ethics; Kant, Immanuel; Neo-Kantianism; Russian Philosophy; Solipsism.

Bibliography

WORKS BY LAPSHIN

Zakony myshleniia i formy poznaniia. St. Petersburg, 1906.

Problema chuzhovo "ia" v noveishei filosofii (The problem of other selves in modern philosophy). St. Petersburg: Senatskaia, 1910.

Filosofiia izobreteniia i izobretenie filosofii (Philosophy of invention and the invention of philosophy). 2 vols. Petrograd, 1922; 2nd ed., Prague, 1924.

"Oproverzhenie solipsizma" (A refutation of solipsism). In *Trudy russkikh uchonykh za granitsei*. No. 1. Prague, 1923. A German translation is available: *Von der Ueberwindung des Solipsismus* Melbourne: D. Jakovenko, 1969.

Spor o svobode voli v sovremennoi filosofii (The debate on the freedom of the will in contemporary philosophy). Prague: Russkii svobodnyi universitet, 1941.

WORKS ON LAPSHIN

Losskii, N. O. *History of Russian Philosophy*. New York: International Universities Press, 1951.

Zen'kovskii, V. V. *Istoriia Russkoi Filosofii*. 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*. 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

LA ROCHEFOUCAULD, DUC FRANÇOIS DE
(1613–1680)

Duc François de La Rochefoucauld, the French epigrammatist and moral critic, was born in Paris; he was known as the prince de Marcillac until he succeeded his father in 1650. An incurable love of adventure and imprudent women brought him into early conflict with Cardinal Richelieu, who imprisoned him briefly in the Bastille in 1637. Contempt for Jules Mazarin, whose treatment he bitterly resented, led La Rochefoucauld to join the faction of the Cardinal de Retz when the Fronde broke out in 1648, but before the end of hostilities he had gone over to Condé's side and was seriously wounded in 1652.

In 1656 he was permitted to return from exile to Paris, where he lived until his death, which occurred after many crippling years of gout. During this period, he became a leading figure in salon society, where his closest friends were Mme. de Lafayette and Mme. de Sévigné, as well as in the Port Royal circle, which included Antoine Arnauld and Mme. de Sablé.

Shortly after his return to Paris he began his *Mémoires*, first Books III–VI (covering the Fronde), then Book II (on the years from 1642 to 1649), and finally Book I (on the years from 1624 to 1642). A grossly inaccurate pirated Dutch edition, which appeared in 1662, caused a great scandal, but the authentic text was not published until the nineteenth century. These *Mémoires*, although less ample and distinguished than those by Retz on the same events, are indispensable to an understanding of the *Maximes*, since they show the inconsistency, dishonesty, and superficiality characteristic of the aristocratic *Frondeurs*.

The *Maximes* were begun as a joint enterprise with Mme. de Sablé and Jacques Esprit (of the Port Royal Circle) and reflect a popular salon pastime, but after the appearance of a pirated Dutch edition in 1664, successive authorized editions followed from 1665 to 1678, considerably altering the scope and nature of the work. The contributions of La Rochefoucauld's friends, as well as maxims too closely resembling such models as Seneca and Montaigne, were deleted, and the original brief moral reflections that occupied a page or so were cut up into the present highly condensed epigrammatic form of a few lines.

The *Maximes* deal with human nature from a strictly human standpoint, all references to God and religion having been systematically removed. They give a lucid

and penetrating analysis of the manifold forms taken by self-interest, which, according to La Rochefoucauld, is the fundamental motive behind human behavior. He also claims that "reason is most often the dupe of the heart," so that human nature is a mass of capricious and unpredictable passions of physiological origin, and what commonly passes for virtue, when it is not pure accident, is really disguised, or unrecognized, vice. He shows little confidence in the Cartesian program of passions controlled by reason and will, and no confidence whatsoever in any concept of natural virtue such as that held by admirers of the virtuous pagans of antiquity. The *Maximes* stress the importance of self-analysis and being honest with oneself; without these qualities love and friendship are a hollow sham, and even with them they may be no more than exercises in egoism.

The predominantly pessimistic outlook reflected in the *Maximes* is partly relieved by the brilliance of the style and the subtlety of the analysis, and also partly by various qualified admissions that true friendship and genuine integrity (*honnêteté*), although rare, may occasionally be encountered. The growing pressure of conformism in a highly artificial society, the author's own experience of pointless heroism and shabby motives in the *Fronde*, and above all his proud and melancholy temperament serve to explain the harsh verdict of the *Maximes*. For all their abiding interest these epigrams remain the direct product and reflection of the age in which they were written. Some brief essays, portraits, and numerous letters constitute the rest of La Rochefoucauld's work.

See also Arnauld, Antoine; Continental Philosophy; Montaigne, Michel Eyquem de; Self-Interest; Seneca, Lucius Annaeus; Virtue and Vice.

Bibliography

La Rochefoucauld's writings may be found in his *Oeuvres complètes*, Bibliothèque de la Pléiade (Paris, 1957). More recent editions of the *Maximes* include an edition edited by J. Vallier (Lausanne, 1962) and the English translations of C. Fitzgibbon (London: Allan Wingate, 1957) and L. W. Tancock (Harmondsworth: Penguin, 1959).

For literature on La Rochefoucauld, see E. Magne, *Le vrai visage de La Rochefoucauld* (Paris, 1923). There is also an excellent chapter on him in A. Adam, *Histoire de la littérature française au XVIIe siècle*, Vol. 4, Ch. 2 (Paris, 1954). See also Susan R. Baker, "The Works of La Rochefoucauld in Relation to Machiavellian Ideas of Morals and Politics," *Journal of the History of Ideas* (44[1983]: 207–218); Henry C. Clark, "La Rochefoucauld and the Social Bases of Aristocratic Ethics," *History of European Ideas* (8[1987]: 61–76); Peter M. Fine, *Vauvenargues and La Rochefoucauld* (Manchester, U.K.: Manchester University Press, 1974); J. E. Parsons, Jr., "On La Rochefoucauld: Preliminary Reflections," *Interpretation* (2 [1971]: 126–142); Tilo Schabert, "The Para-Moral Principles of Early Modern Society: Contextual Reflections upon the Maxims of La Rochefoucauld," *History of European Ideas* (7[1986]: 67–84).

A. J. Krailsheimer (1967)
Bibliography updated by Tamra Frei (2005)

LAROMIGUIÈRE, PIERRE

(1756–1837)

Pierre Laromiguière, the French professor of philosophy, was born at Livignac in the district of Rouergue. As a young man, he was ordained a priest and exercised his ecclesiastical duties for a short period before becoming professor of philosophy successively at Carcassonne, Tarbes, Toulouse, and the Sorbonne. He was a close student of Étienne Bonnot de Condillac and an associate of the Idéologues but departed from the teachings of both in certain particulars. Excessively shy, he refused to propose his candidacy to the French Academy, though twice urged to do so, and confined his public appearances to the classroom. He died in Paris, one of the most esteemed and beloved of teachers. Among his more famous pupils were Victor Cousin and Théodore Jouffroy.

Laromiguière's disagreement with the school of Condillac arose over the question of the mind's passivity. He argued that if all our ideas were modifications of sensory material impressed upon us by external causes, it would be impossible to account for attention, comparison, and reason. These, he held, were essentially active. There is a fundamental distinction to be made, he said, between seeing and looking, listening and hearing, and the difference cannot be explained if the soul is a passive recipient of sensory stimuli. Activity was indefinable for Laromiguière, since it had no anterior ideas from which it could be derived. He seemed to believe that anyone hearing the term would grasp its meaning.

The three activities of the understanding were attention, comparison, and reasoning; and the three activities of the will corresponding to them were desire, preference, and freedom—the latter being the power to act or not to act. Laromiguière's insistence on the soul's activity was most welcome to his contemporaries, for it restored to men the autonomy that, they felt, Condillac had destroyed. While disagreeing with Condillac on this point, Laromiguière agreed with his predecessor that the primary business of philosophy was the analysis of ideas. In his best known and extremely popular work, *Leçons de philosophie,* which ran through six editions between 1815

and 1844, he assigned to metaphysics the single task of discovering the origin of all our ideas.

Laromiguière was particularly admired for the perfection of his literary style, the fame of which was acknowledged even by Hippolyte-Adolphe Taine.

See also Condillac, Étienne Bonnot de; Cousin, Victor; French Philosophy; Ideas; Jouffroy, Théodore Simon; Taine, Hippolyte-Adolphe.

Bibliography

WORKS BY LAROMIGUIÈRE

Projet d'éléments de métaphysique. Toulouse, 1793.

Sur les paradoxes de Condillac. Paris, 1805.

Leçons de philosophie sur les principes de l'intelligence. Paris, 1815–1818.

WORKS ON LAROMIGUIÈRE

Boas, George. *French Philosophies of the Romantic Period,* 33–42. Baltimore: Johns Hopkins Press, 1925.

Janet, Paul. "Laromiguière, la liberté de penser." *Revue philosophique et littéraire* 1 (1848): 253–263, 358–368.

Picavet, François. *Les idéologues,* 520–548, 552–567. Paris: Alcan, 1891. Second section referred to deals with Laromiguière's influence.

Taine, Hippolyte. *Les philosophes classiques du XIXe siècle en France.* Paris: Hachette, 1868. Ch. 1.

George Boas (1967)

LAROUI, ABDULLAH

(1935–)

A Moroccan intellectual born in 1935, Abdullah Laroui taught at Mohammed V University in Rabat and was one of a distinguished group of Moroccan thinkers such as M. A. Lahbabi and M. A. al-Jabri. His work involves a variety of theoretical and disciplinary perspectives, including history (his main professional discipline), sociology, philosophy, and literature. He has produced histories of the Maghreb (North Africa) and of the Arab world in general, and his writings on modernity and strike at the heart of many key issues that are important for Arab culture in the postcolonial world. He raised in particular the questions how history should be written and how to understand the cultural life of a group of people through understanding their histroy. This comes out as a much more complicated issue than might initially appear to be the case, and Laroui uses the conceptual machinery of both Ibn Khaldun and Machiavelli to try to position Arabi history within an appropriate theoretical context.

One of Laroyi's major achievement is in laying out the ambiguous nature of some of the key concepts of Arab culture in the contemporary world, including modernity, the state, authenticity, continuity, rationality, and tradition. He argues that the Arab world cannot adopt wholeheartedly the Western concept of the state since this is essentially a secular notion and pays little regard to the past, while for Arabs the link with Islam and their history is a crucial aspect of political legitimacy. In any case, the state is only a part of the whole Islamic *umma*, or community, and there is a notion of an Arab *umma* in which the state exists, and that produces a nexus of relationships for the concept of the state, that makes little sense of the Western notion. As with his predecessor Ibn Khaldun, Laroui has an approach to understanding society that makes it important to develop and clearly specify a theoretical perspective and put a particular social structure with in its historical and cultural context. In this way it is possible to say something both true and interesting about Arab ideology, while using the conceptual tools imported directly from outside the region is unlikely to be helpful.

Laroui's thought has moved from his earlier Marxism to produce a more nuanced approach to the philosophy of history and the understanding of culture. He is part of a significant movement in the modern Arab world that tries to define Arab culture and its unique features, while at the same time making use of theories from outside the region where they can shed light on the issue.

See also Ibn Khaldūn; Islamic Philosophy; Machiavelli, Niccolò; Marxist Philosophy; Philosophy of History.

Bibliography

Laroui, Abdullah. *Crisis of the Arab Intellectual: Traditionalism vs. Historicism*. Translated by D. Cammell. Berkeley: University of California Press, 1976.

Hasan Hanafi (2005)
Oliver Leaman (2005)

LASSALLE, FERDINAND

(1825–1864)

Ferdinand Lassalle, the German socialist, was born Ferdinand Lasal in Breslau, Silesia, of a middle-class Jewish family. The young Lassalle—he gallicized his name—was a poor and rebellious student. Quite early he indicated his persistent, but never conflicting, longings both to relieve the oppressed and to achieve aristocratic status. These

two desires illuminate the paradoxical nature of a man who championed the causes of oppressed workers and oppressed noblewomen with equal vigor. He corresponded regularly with Karl Marx, defended the honor of the Countess von Hatzfeldt in a lengthy and celebrated lawsuit, sought the acclaim of Berlin society, founded the Allgemeiner Deutscher Arbeiterverein (the first political party for German workers), dressed fastidiously, and died at the age of thirty-nine, from wounds suffered in a duel with Count von Racowitza.

Lassalle attended the universities of Berlin and Breslau, falling under the influence of Hegelian philosophy at the latter. But, although he had philosophic pretensions and sought the acclaim of philosophers, he preferred a life of action to one of theory; and his fame rests chiefly on his founding the Allgemeiner Deutscher Arbeiterverein in 1863, to which German Social Democrats still trace their origin.

Lassalle referred to his exposure to G. W. F. Hegel as his "second birth." He avidly consumed Hegel's works, as well as those of young Hegelians like David Friedrich Strauss and Ludwig Feuerbach. Hegel's reference to the ancient Ionian philosopher Heraclitus of Ephesus as his forerunner led Lassalle to study Heraclitus; and he sought to demonstrate that Heraclitus had forecast Hegelian ideas. Lassalle also aspired unashamedly to the fame that a major philological and philosophical work would provide for him in German society. He began his research while not yet twenty, but did not complete it until fifteen years later. Berlin academicians hailed the publication in 1858 of *Die Philosophie Herakleitos des Dunkeln von Ephesos,* but later critics have found grave defects in the work, most notably Lassalle's preoccupation with Hegel rather than Heraclitus.

Hegelian ideas dominated Lassalle's historical and economic thought as well. His historical and economic theories, although not carefully formulated, emerge most clearly from the works of his last years, when he was organizing the Arbeiterverein, especially *Das System der erworbenen Rechte* (Leipzig, 1861); *Arbeiter-Programm* (Berlin, 1862); *Über Verfassungswesen; Die indirekte Steuer und die Lage der arbeitende Klassen* (1863; reprinted Berlin, 1874); and *Herr Bastiat-Schulze von Delitzsch, der ökonomische Julian, oder Kapital und Arbeit* (Berlin, 1864). He shared with Marx the belief that revolutions are not "created" by revolutionaries, but occur as the result of a historical process. Men called revolutionaries are in fact merely the midwives to a new age produced in the womb of time. Lassalle described this process in Hegelian terms. A new social order, when it appeared, would rise on the wings of Hegelian ideas. The bourgeois idea of freedom had destroyed feudal solidarity in 1789. The bourgeoisie had liberated itself by reducing the state to the role of "nightwatchman." The proletariat would in turn liberate itself through association, at first within a political party that would demand and obtain universal suffrage from the state. Having achieved universal suffrage, the proletariat would use the power of the state to form great workers' associations or cooperatives. These would in turn liberate the worker from the cruel "iron law of wages" and achieve freedom for him.

Lassalle worked arduously at organizing the workers into a national political party. He did not intend to overthrow the state, but to use it. The idea of freedom would find eventual embodiment through the state. All previous conflicts would be synthesized in this final stage of history. Thus Lassalle accepted the Prussian state and perhaps even the Prussian monarchy. His position on the latter, as well as on private property, is ambiguous. Lassalle wrote and agitated under Prussian censorship and was constantly being tried for treasonable activity. His published works and public statements are therefore not always consistent with his private correspondence and conversations.

Lassalle's relationship with Marx waxed warm and cool. Lassalle undoubtedly admired Marx and sought the latter's approval, whereas Marx disapproved of much that Lassalle wrote and did. Marx regarded Lassalle as a friend, an informant, a creditor, a publishing agent, and an immature, pompous plagiarist. They broke off their correspondence before Lassalle's death.

See also Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heraclitus of Ephesus; Marx, Karl; Socialism; Strauss, David Friedrich.

Bibliography

COLLECTED WORKS OF LASSALLE

Gesammelten Reden und Schriften. 12 vols. Edited by Eduard Bernstein. Berlin, 1919–1920.

Nachgelassene Briefe und Schriften. 6 vols. Edited by Gustav Mayer. Stuttgart: Deutsche Verlagsanstalt, 1921–1925.

WORKS ON LASSALLE

Baron, S. *Die politische Theorie Ferdinand Lassalles.* Leipzig: Hirschfeld, 1923.

Bernstein, Eduard. "Ferdinand Lassalle und seine Bedeutung in der Geschichte der Sozialdemokratie." In Ferdinand Lassalle, *Reden und Schriften,* Vol. I. Berlin, 1892.

Brandes, Georg. *Ferdinand Lassalle.* Berlin, 1877. Translated by Eden and Cedar Paul. London: W. Heinemann, 1911.

Footman, David. *The Primrose Path, a Life of Ferdinand Lassalle.* London: Cresset Press, 1946. A good biography.

Knapp, Vincent J. "Ferdinand Lassalle on the State and Society: A Legacy to Welfare Statism." *Australian Journal of Politics and History*17 (3) (1971): 377–385.

Lewis, Vivian. "Ferdinand Lassalle, 1825–1864." *History Today* 15 (1) (1965): 58–64.

Oncken, Hermann. *Lassalle.* Stuttgart: Frommanns, 1904. An excellent biography.

Schirokauer, Arno. *Lassalle, die Macht der Illusion, die Illusion der Macht.* Leipzig: P. List, 1928. Translated as *Lassalle, the Power of Illusion and the Illusion of Power.* London: Allen and Unwin, 1931. A good biography.

Thier, Erich. *Rodbertus, Lassalle, Adolph Wagner.* Jena, Germany, 1930.

Sterling Fishman (1967)
Bibliography updated by Philip Reed (2005)

LATIN AMERICAN PHILOSOPHY

Latin American philosophy covers primarily the philosophy produced in the parts of the Americas that belonged to the Spanish and Portuguese empires after 1492. The Maya, Toltec, Aztec, and Inca civilizations engaged in some philosophical speculation in the form of religious myths and cosmological accounts prior to the arrival of the Europeans, but most of the records of these efforts were destroyed during the conquest. As happened with almost everything else in the wake of colonization, Iberians took control over the development of philosophy and scholastic philosophy became the most influential philosophical trend in the New World.

The encounter posed new challenges to European thought and initiated new developments in both Europe and Latin America. In Iberia, new issues, primarily concerned with the rights of conquered peoples and just war, took center stage, and the greatest Iberian philosophers of the times, Francisco de Vitoria (c. 1492–1546) and Francisco Suárez (1548–1617) addressed them. In the colonies, some attention was given to pre-Columbian worldviews, but these slowly receded into the background, making way for the concerns, first, of the Iberians living in the colonies and, then, of *criollos*—that is, native-born authors.

The colonial roots of Latin American philosophy helped set the stage for the emphasis on sociopolitical issues, such as human rights and social justice, which have been so central to the philosophical development in the region. In addition to addressing standard philosophical questions concerning the nature of being, knowledge, and value, Latin American philosophers have demonstrated a strong commitment to more concrete issues involving educational policy, political organization, and social reform. In contrast to their Anglo-American colleagues, many Latin American philosophers have developed their ideas not in technical articles and systematic treatises intended for specialized audiences, but in newspaper articles, essays, and even fiction, meant to be read by a broad public. This is consonant with the view that philosophy should be a tool for social change and has led some historians to speak of two trends in Latin American philosophy: academic and Latin Americanist. The first is inspired by European philosophy and is practiced in universities; the second is more autochthonous and extends beyond the boundaries of academia. Most authors cited in this entry belong to the first trend. Some of the best known thinkers who belong to the second are Eugenio María Hostos (Puerto Rico, 1839–1903), Justo Sierra (Mexico, 1848–1912), José Martí (Cuba, 1853–1895), José Enrique Rodó (Uruguay, 1871–1917), Jorge Luis Borges (Argentina, 1899–1986), Octavio Paz (Mexico, 1914–1998), Carlos Fuentes (Mexico, b. 1928), Mario Vargas Llosa (Peru, b. 1936), and Luis Castro Leiva (Venezuela, 1943–1999).

MAJOR PERIODS OF LATIN AMERICAN PHILOSOPHY

Four major periods in the history of Latin American philosophy stand out: colonial, independentist, positivist, and contemporary.

COLONIAL PERIOD (C. 1550–1750)

Latin American philosophy begins within the scholastic fold provided by the Iberian clergy sent by the Spanish and Portuguese Crowns to convert the indigenous inhabitants of the territories they had conquered. The main philosophical centers during the early colonial period were Mexico and Peru, the two places in which major empires had flourished and the Europeans found the gold and silver they coveted. The texts studied were those of medieval scholastics and of their Iberian commentators, and the issues addressed concerned logic, natural philosophy (physics), and metaphysics. The first author to publish systematic treatises on these topics was Alonso de la Vera Cruz (1504–1584), but it was Antonio Rubio's (1548–1615) *Logica mexicana* (*Mexican Logic*, 1603) that first gained prominence in the New World and Europe.

Although scholasticism was central to, and many thinkers continued to write within, this tradition, others were guided by humanist ideas. In particular, they were

concerned with the political and legal questions raised by the process of colonization. Arguably, the most important of these thinkers was Bartolomé de las Casas (1474–1566), a Dominican friar from Spain who became the leading champion of the rights of the Indians. His long life was devoted to arguing before the Spanish Crown that the indigenous groups of New Spain were not barbarians "in the strict sense"; they were no less human than Spaniards and so just as deserving of the same basic human rights.

Las Casas first brought up what became known in Spain as "the Indian Question." As early as 1515, he began to petition the Crown to enact laws that would eliminate the notorious *encomiendas*. This system gave Spanish settlers custody of groups of indigenous peoples who were then exploited in mining and agriculture. In 1550 an important debate took place in Valladolid, Spain, between the humanist Juan Ginés de Sepúlveda (1490–1573), a leading ideologue of the *Conquista*, and Las Casas. Las Casas argued that it was unjust to wage war against the indigenous peoples and to enslave them.

Las Casas's defense of the Indians reflects the influence of several sources. The thought of Aristotle (384–322 BCE), known as "the Philosopher" among scholastics, is behind several distinctions upon which Las Casas based his defense. Other sources included canon and Roman law, and such Christian thinkers as Augustine (354–430) and Thomas Aquinas (c. 1225–1274). Las Casas was scholastically trained and the scholastic method informs the structure and content of his rebuttal of Sepúlveda.

The debate between Las Casas and Sepúlveda raised questions concerning the natures of humanity and justice, issues that continue to shape Latin American philosophy to this day. Social injustice did not have only one face. Women also suffered oppression, although for most of the thinkers of the colonial period, this went unnoticed. Aristotle had claimed that women were inferior to men, and most of his scholastic followers did not question this view. But there were isolated voices that cried out against the claim that women were not fit for intellectual activity. One of the most eloquent and powerful of these voices was that of the nun Sor Juana Inés de la Cruz (Mexico, 1651–1695). A recurring theme in her writings is the image of a human being as a microcosm, and reflects the influence of neoplatonic philosophy upon her thought. Much of her writings, whether in prose or poetic form, displays a concern with the unjust position of women in colonial society.

INDEPENDENTIST PERIOD (C. 1750–1850)

A more complete break with scholasticism was attempted during the independentist period. This phase of Latin American thought receives its name from the political rationale articulated in the eighteenth century to gain independence from Spain and Portugal. The intellectuals engaged in this enterprise were men of action who used ideas for practical ends. The strong influence of Utilitarianism is reflected in their emphasis on progress and the attempt to employ ideas as tools for social change. Another source that shaped the period came from the liberal views of the French *philosophes*, who made reason a measure of legitimacy in social and political matters.

Most leading figures from this period were not philosophers in the strict sense. Simón Bolívar (Venezuela, 1783–1830), José Joaquin Fernández de Lizardi (Mexico, 1776–1827), Mariano Moreno (Argentina, 1778–1811), and José Cecilio del Valle (Honduras, 1780–1834) can be most accurately characterized as independence leaders rather than as philosophers. Instead of devoting their lives to developing systems of thought, they were more interested in concrete political and military action that would lead to the independence of the Iberian colonies.

Bolívar, known as the Liberator, successfully led northern South America to independence from Spain and was the founding father of five republics (Colombia, Venezuela, Ecuador, Peru, and Bolivia). Monumental deeds such as these are a central part of his legacy. Yet, his writings also helped to change the political structures in America and drew attention to the dangers confronting the newly liberated regions. In his *Carta de Jamaica* (*Jamaica Letter*, 1815), a call to independence from Spain, he complains of both a state of permanent infancy experienced by the nations of Spanish America and their dependence upon Europe. The problem of dependence is an enduring one in the Latin American philosophical tradition, shaping one of the most widespread strands of Latin American thought, the philosophy of liberation. In the *Carta*, Bolívar also touches upon the question of identity, another central theme of the tradition, prompted by the merging of indigenous and European populations and cultures.

Bolívar influenced thinkers of the contemporary period, such as the father of the Cuban revolution, José Martí (1853–1895), as well as the Mexican philosophers Samuel Ramos (1897–1959) and Leopoldo Zea (1912–2004). All of these thinkers devote considerable attention to the issues of liberation and cultural identity.

POSITIVISM (C. 1850–1910)

Once political independence had been achieved, a somewhat more stable period of philosophical activity, known as positivism, began. With the exception of scholasticism, positivism has been the most widespread and deeply rooted current in Latin American thought to date. The depth of its impact can be explained in historical terms: It took shape just in time to address the need for nation building in the region. Positivism was in part a response to the social, financial, and political needs of the newly liberated countries of Latin America.

The European father of positivism was the French philosopher Auguste Comte (1793–1857), who attempted to develop a rigorous and systematic understanding of human beings in both their individual and social dimensions. He emphasized experience over theoretical speculation and empirical science over metaphysics. The value of knowledge rests on its practical applications: Knowledge is a servant of action and should lead to the solution of concrete problems. This practical dimension was one of the most captivating aspects of Comte's thought for Latin Americans, for they wished to overcome anarchy, eradicate poverty and disease, and place their own countries on the path of progress.

Practical considerations, however, were not the only cause of positivism's success. Cultural and theoretical reasons also played a role. Since the colonial period, Latin American philosophy had been nurtured by scholasticism and, consequently, important practical issues had been neglected. Conceptual and terminological vagueness, excessive speculation, as well as unfounded and archaic dogmatism were predominant characteristics of much of the philosophy of the region. Positivists, by contrast, emphasized principles based on experience and logical rigor, and offered the assurance of progress, insisting that their claims rested on solid empirical evidence. With positivism, the leaders of the newly liberated republics thought they would finally leave not only the political legacy of the colonization behind, but the philosophical one as well.

The movement benefited greatly from the increasing prestige of science, because it proposed to limit its methods to those used by natural scientists. It was widely believed by those who favored this perspective that a new era had begun in which scientific study would make it possible to identify the causes of social evils and to eliminate them, just as medicine had begun to do with endemic diseases.

Comte's law of the three stages captured the attention of many Latin American intellectuals. According to this law, humanity passes through a theological, a metaphysical, and a scientific or positive stage. In the theological stage, the interpretation of reality is founded on prejudice and superstition. In the metaphysical, it is characterized by speculation, and facts are either ignored or not given adequate attention. Finally, in the positive stage, superstition and speculation are replaced by the establishment of facts, and knowledge is founded on experience.

Latin American thinkers applied Comte's law to the history of their own countries. An example of this application is found in the *Oración cívica* (*Civic Oration*, 1867) delivered by the Mexican Gabino Barreda in Guanajuato. With this text in mind, President Benito Juárez named Barreda member of a committee to draft a law, approved on December 2, 1867, that gave birth to public education in Mexico. The fact that another renowned teacher, Justo Sierra (1848–1912), succeeded to Barreda's position and continued to apply positivist principles to educational policy explains the strength that this perspective had in Mexico until the fall of the dictator Porfirio Díaz in 1911. Positivism was the official philosophy in Mexico during the twenty-seven-year dictatorship, and the government was guided by Comte's slogan "Order and Progress."

The chaos and backwardness that prevailed in some Latin American countries as the power vacuums left in the wake of colonial rule were filled by *caudillos* and other nondemocratic political leaders and structures helps to explain in part why positivist teachings captivated the minds of so many intellectuals and politicians. For example, the influence of positivism can be observed in the work of the Argentine thinker and statesman Domingo Faustino Sarmiento (1811–1888). His account of civilization in *Facundo, o civilización y barbarie* (*Facundo, or Civilization and Barbarism*, 1845) is shaped by positivist principles.

Each country of Latin America had its own particular way of appropriating positivism. Latin American positivism was shaped not only by Comte, but also by the social Darwinism expounded by the English philosopher Herbert Spencer (1820–1903). Comte had a stronger impact in Brazil, Mexico, and Chile, whereas Spencer was more influential in Argentina, Uruguay, and Cuba. In some cases, preference was given to one over the other for purely political reasons. In Cuba, for example, Enrique José Varona (1849–1933) rejected Comte's ideas because they did not favor the emancipation of Cuba from Spanish rule, and he adopted instead Spencer's notion of lib-

erty. In spite of these and many other national differences, one can speak of Latin American positivism as a unified, yet evolving movement in which the influence of Comte was greater toward the beginning of the period and that of Spencer predominated toward the end.

Juan Bautista Alberdi (Argentina, 1812–1884) and Andrés Bello (Venezuela, 1781–1865) were also influenced by positivism. Positivism's legacy in Argentina remained strong because it was never involved in any political movement, so it never came to be associated with dictatorships as, for example, was the case in Mexico. Furthermore, in Argentina positivism had an effective role in the development of educational institutions and, through the work of José Ingenieros (1877–1925), acquired renown in scientific and philosophical circles.

As in Argentina, positivism played a role in the development of Brazilian education. Nisia Floresta (1809–1885), one of the founders of the positivist movement in Brazil, was director of a school in Rio de Janeiro and, upon moving to Paris, established a close friendship with Comte. Furthermore, in Brazil, positivism was associated with the founding of the Republic in 1889, and the positivist motto, "Ordem e Progresso" (Order and progress), is inscribed on the Brazilian flag. Brazil was one of the last countries of Latin America to abolish slavery (1888) in part due to the positivist movement in Brazil. An understanding of Brazil's history in the eighteenth century is impossible without an appreciation of the role that positivist thinkers such as Raimundo Teixeira Mendes (1855–1927), Miguel Lemos (1854–1917), and Luis Pereira Barreto (1840–1923) played in the founding of the Republic.

CONTEMPORARY PERIOD (C. 1910–PRESENT)

The period following positivism is known as the contemporary period, and it can be broken down into three phases: the foundational stage, the period of normalcy, and the period of maturity.

FOUNDATIONAL STAGE (C. 1910–1940). This phase begins with the decline of positivism. The generation of thinkers who first rejected the central tenets of positivism became known as "the founders," a label coined by Francisco Romero (Argentina, 1891–1962). It included: Deústua, Alejandro Korn (Argentina, 1860–1936), Enrique Molina (Chile, 1871–1964), Carlos Vaz Ferreira (Uruguay, 1872–1958), Raimundo de Farias Brito (Brazil, 1862–1917), José Vasconcelos (Mexico, 1882–1959), and Antonio Caso (Mexico, 1883–1946), among others.

The general decline of positivism stems from several factors, and the national context must be taken into consideration insofar as the predominance of any particular cause varies from country to country. Still there were causes common to all Latin America. The first of these was the disappointment that Latin American intellectuals experienced when reality did not measure up to positivism's promises and aspirations. Immediate and assured results were envisioned and anxiously awaited, but progress was slow and uncertain. To uphold general principles and criteria for the study of social problems is one thing, but it is quite a different matter to develop effective, scientifically based procedures that can be applied in order to solve concrete problems. Stark reality shattered many illusions. The ideal of a scientific knowledge of society began to crumble in the face of difficulties, and the initial naive optimism gave way to corroding pessimism.

In addition, many thinkers began to discover fundamental theoretical shortcomings in positivism. The indiscriminate application of the principle of causality led positivists to deny freedom to human beings. Theoretical objections to determinism acquired momentum in the moral realm. No one can be held accountable for an act if it is determined, the critics of positivism claimed: positivism seemed to lead to an ethical dead end.

In particular, positivism seemed to spell disaster for aesthetic creation. If humans are not free, how can they be aesthetic agents? A mechanical explanation of the creative process factored out the very meaning of artistic creation, something that many Latin Americans found unacceptable. Deústua in particular responds by developing an aesthetic theory in his influential *Estética* (*Aesthetics*, 1923), in which aesthetic value is conceived as the source of all value. This "value of all values" as he calls it, is the product of free activity whose essential function consists in the creation and contemplation of an ideal aside from any practical intent. In contrast to the essentially instrumental character of other values, aesthetic value constitutes its own end, generating a completely disinterested activity, the creation of beauty.

Political considerations also factored into the general disenchantment with positivism. As already mentioned, in some countries such as Mexico, positivism was associated with a dictatorship that had been overthrown; in others such as Cuba, Comtian positivism was believed to support the colonial status quo against the possibility of independence to which many Cubans aspired. For countries that had suffered first under Spanish oppression and then under a succession of dictators, setting freedom

aside seemed too high a price to pay for the promise of progress. Indeed, freedom had become the battle flag, so if positivism could not make room for freedom, then positivism must be abandoned.

In 1909 a group of young intellectuals in Mexico, who later acquired well-deserved renown in the field of philosophy and literature, founded the "Ateneo de la Juventud" (Atheneum of youth). They studied the philosophical classics, especially Plato and Kant, and contemporary philosophers who had rejected positivism in Europe, such as Henri Bergson (1859–1941) and Benedetto Croce (1866–1952). The influence of Friedrich Nietzsche (1844–1900) and Arthur Schopenhauer (1788–1860), whose thought was a counterweight to the narrow, scientific emphasis of positivism, was also felt. Following these studies, lectures were given in which positivist doctrine was roundly criticized and new ideas were proposed.

Vasconcelos is one of the most influential figures of this generation. He was not only an accomplished philosopher, but, like so many other Latin American intellectuals, also a devoted educator and political activist. Much of his work focuses upon the meaning of Mexican culture in particular and the destiny of Latin America in general. In two of his most popular works, *La raza cósmica* (*The Cosmic Race*, 1925) and *Indología* (*Indology*, 1926), he claims that the future will be constituted by the cosmic race, a synthesis of the four basic races of the world that will emerge in the region of the Amazon basin and fulfill "the divine mission of America." He contrasts this to the ethnic egoism and nationalism that dominates in Anglo-Saxon culture and claims that the new race of which he speaks will be characterized by a universalist spirit based on love.

Two other key figures are Caso in Mexico and Korn in Argentina. They are particularly important because they functioned as influential teachers and mentors of the generation that followed them. The first developed a moral theory based on the principle that there are two basic attitudes toward existence: One is based on the notion of existence as economy, where action is dictated by maximum advantage with minimum effort; the other is guided by disinterest, where action is dictated by maximum effort with least concern for advantage. The first is a positivist morality, the second is a morality based on love. In a somewhat similar vein, Korn developed a philosophy of creative freedom inspired by Kant. Although the physical realm operates according to necessary laws, subjects can formulate ideals and act according to them, thus resisting the tyranny of nature.

Crucial influences in the overcoming of positivism and its legacy were vitalism and intuitionism, especially the versions imported from French philosophers such as Émile Boutroux (1845–1921) and Bergson. Vitalism was a metaphysical position that conceived reality in terms of life. Intuitionism was an epistemic view in which knowledge, particularly about values, is based on intuition. Representative of this move away from positivism's narrow approach was the work of Vaz Ferreira. In books such as *Conocimiento y acción* (*Knowledge and Action*, 1907), *Lógica viva* (*Vital Logic*, 1910), and *Fermentario* (*Fermentary*, 1938), he attacks the narrow, purely rational concept of knowledge that excludes the dynamic vitality of reality. He also pioneered the discussion of feminist issues in *Sobre feminismo* (*On Feminism*, written between 1914 and 1922, but first published in 1933).

An important force in the transition from positivism to vitalism was the Spanish philosopher José Ortega y Gasset (1883–1955). But there were also others, who belonged to what has come to be called the Generation of 98. The year 1898 marked the end of Spain's colonial empire, yet it also signaled the opening of a promising, new intellectual movement. The famous generation of 1898 gave Spain some of its most brilliant intellectuals, including two of its greatest philosophers, Ortega and Miguel de Unamuno (1864–1936).

Many thinkers in Spain, a country located at the geographical margins of Europe, struggled to be recognized as European, but Unamuno was more interested in developing the notion of *hispanidad* (Hispanicity) and to the *hispanización* (Hispanization) of Europe. The notion of *hispanidad* came to serve as an important bridge between the philosophy of Spain and Latin America. Interest in analyzing the meaning of *hispanidad* has continued into the twenty-first century, with philosophers in the United States developing arguments concerning rights for Hispanics and debating the very meaning of the term "Hispanic."

Like Unamuno, Ortega also made the intellectual, political, and social situation of Spain central to his philosophy. He developed what has become known as a "philosophy of circumstance," well captured in the famous lines: "*Yo soy yo y mis circunstancias y si no las salvo a ellas no me salvo yo*" (I am myself and my circumstances, and if I don't save them, I cannot save myself). The idea is that the self is not an entity apart from its context. Integral to this view is the notion that all knowledge is perspectival—that is, it is the expression of a view from a particular perspective. Ortega's perspectivism came to play a

critical role in the work of several Latin American philosophers.

THE FORGERS. The thought of the generation that followed the founders was shaped by ideas imported from Spain, France, and Germany, and Ortega is generally credited with having introduced them, particularly German philosophy, into Latin America. The extraordinary impact of Max Scheler (1874–1928) and Nicolai Hartmann (1882–1950) can be explained only through Ortega's influence. This group of philosophers has been characterized by Francisco Miró Quesada (Peru, b. 1918), as "the generation of forgers" because of the major role they played in setting the parameters for the subsequent development of Latin American philosophy. A major figure of this generation was Samuel Ramos (Mexico, 1897–1959). He focused upon Mexican culture, thereby inspiring interest in what is culturally unique to Latin American nations. Ramos' book, *El perfil del hombre y la cultura en México* (Profile of man and culture in Mexico, 1934), was the first attempt at interpreting Mexican culture. Francisco Romero was also an important thinker who developed an elaborate philosophical anthropology in his *Teoría del hombre* (Theory of man, 1952). He sought to frame a view of human beings in terms of universal notions such as intentionality and spirituality, rather than the culturally specific parameters used by Ramos.

Throughout the history of Latin American thought there has been a tension between philosophers who focus on the universal human condition and those who emphasize particular cultural circumstances. In Mexico, for example, many philosophers have discussed the impact of the colonization on the development of culture in Mexico. And in *Siete ensayos de interpretación de la realidad peruana* (Seven interpretative essays on Peruvian reality, 1928), the Peruvian Carlos Mariátegui (1894–1930) proposed an interpretation of Marxism that emphasized the particular conditions that characterized the Peruvian situation. This particularist tendency grew in part as result of a historical event that brought the Spanish and Latin American philosophical traditions into even closer contact with one another and heralded yet another stage in the latter.

The historical circumstances of Spain in the twentieth century were complicated, and part of the influence that Spanish thinkers came to have upon the development of philosophy in several Latin American countries can be attributed to the political upheaval caused by the Spanish Civil War (1936–1939) and the ensuing dictatorship of Francisco Franco (1939–1975). Many of the most important Spanish philosophers of this period were driven into exile during the years of Franco's oppressive dictatorship and several of them settled in Latin America.

THE "TRANSTERRADOS" AND THE PERIOD OF NORMALCY (C. 1940–1960). During the late 1930s and 1940s, due to the upheavals created by the Spanish Civil War, a significant group of thinkers from Spain arrived in Latin America. These philosophers became known as the *transterrados* (trans-landed). Seeking refuge from Franco's dictatorship, they settled in various countries of Latin America. Among them were Joaquín Xirau (1895–1946), Eduardo Nicol (1907–1986), José Ferrater Mora (1912–1991), José Gaos (1900–1969), Luis Recaséns Siches (1903–1977), and Juan D. García Bacca (1901–1992). Their presence helped to break some of the national barriers that had existed in Latin America before their arrival. The conception of *hispanidad* that they inherited from Unamuno and the need to establish themselves in their adopted land helped the process; they went from country to country, spreading ideas and contributing to an ever broadening philosophical dialogue. Their influence showed itself most strongly when the generation born around 1910 reached maturity.

Gaos was one of the most influential *transterrados*. He was a student of Ortega and became the teacher of one of Mexico's most important philosophers, Leopoldo Zea. Gaos encouraged Zea to study the history of Mexican thought, and this resulted in one of Zea's most important books, *El positivismo en México* (Positivism in Mexico, 1943). Through Gaos, Ortega had a strong influence on Zea's views. Following Ortega's insights that in order to understand ourselves, we must understand our circumstance, and that all knowledge is perspectival, Zea turned to the meaning of the Latin American circumstance for the development of the philosophy of the region.

Zea's philosophy was also influenced by Ramos's work. The latter's existential, psychoanalytic approach to the problem of cultural identity was transformed by Zea into a critique of philosophy and the articulation of a *mestizo* (mixed) consciousness. The term *mestizo* points to issues associated with race and culture, opening a philosophical discussion concerning the identity of persons who share Spanish and indigenous heritage. The source of this line of questioning can be traced back to the events following the colonization, when the Spaniards began to mix with the indigenous people to create what has come to be known as a *mestizo* race and culture. Zea's

notion of *mestizaje* had a strong influence on the Argentine philosopher Arturo Andrés Roig (b.1922) and the Peruvian Miró Quesada. The relation between these thinkers constitutes an example of a growing philosophical Pan-Americanism. During this period, philosophers from different countries in Latin America began to respond to each other and to interact critically with one another.

PERIOD OF MATURITY (C. 1960–PRESENT). This Pan-American trend continues to the present and has been further supported by the activities of various organizations founded to facilitate meetings and publications. From 1960 to the present, the level of philosophical activity in several Latin American countries has improved significantly. This is due, in part, to the institutionalization of philosophy. The number of national philosophical societies and of centers, institutes, faculties, and departments that have as their exclusive end the teaching and investigation of philosophy has increased substantially as have the number of philosophy journals. All of this activity has begun to awaken interest outside of Latin America, and indeed to give rise to a diversification of philosophical trends within Latin America itself. Three trends in particular illustrate the current situation of Latin America: philosophical analysis, liberation philosophy, and discussions of identity.

Philosophical analysis. Analytic philosophy is characterized by a preoccupation with language, a strong interest in logic, a positive attitude toward science, and a general mistrust of metaphysics. Its founders are G. E. Moore (1873–1958), Bertrand Russell (1872–1970), Ludwig Wittgenstein (1889–1951), and the members of the Vienna Circle. Analytical philosophy is often contrasted to Continental philosophy, which has its roots in France and Germany and is based on the thought of such figures as Hegel, Nietzsche, Jean-Paul Sartre (1905–1980), and Martin Heidegger (1889–1976). Continental philosophy was disseminated earlier and more widely than analytic philosophy in Latin America. Even in the early twenty-first century, authentic Latin American philosophy is often taken to be concerned exclusively with issues of Latin American cultural identity and liberation, and so to have little in common with the analytic tradition. Yet, this is a misconception.

The groundwork for the favorable reception of analysis in Latin America can be traced back to positivism. In the 1920s, key texts from the analytic tradition, such as G. E. Moore's *Ethics* and Bertrand Russell's *The Problems of Philosophy*, were translated into Spanish. While marginalized from its inception, philosophical analysis has provided a robust methodological alternative to Ortega's perspectivism and Continental philosophy.

The fruits of the interest in analytic philosophy became evident in the 1940s, when Vicente Ferreira da Silva (1916–1963) published a manual of mathematical logic in Brazil, and Miró Quesada published one in Peru. Miró Quesada has maintained a balanced view of philosophy throughout his career. His works, *Despertar y proyecto del filosofar latinoamericano* (The awakening and project of Latin American philosophy, 1974) and *El problema de la filosofía latinoamericana* (The problem of Latin American philosophy, 1976), testify to his view that philosophy must combine both solid philosophical analysis and a historical approach that takes into account the particular circumstances of Latin America.

In Buenos Aires, Hans A. Lindemann, who had connections to the Vienna Circle, brought attention to philosophical analysis in Argentina, as the work of Gregorio Klimovsky (b. 1922) and Julio Rey Pastor (1888–1962) illustrates. In *El punto de partida del filosofar* (Philosophizing's point of departure, 1945), Risieri Frondizi (Argentina, 1910–1983) offered a serious critique of logical positivism while displaying the influence of philosophical analysis.

In the 1960s, philosophical analysis was integrated into many philosophy departments throughout Latin America. Argentina continued to be a center of this kind of philosophy. Mario Bunge's (Argentina, b. 1919) *Causalidad* (Causality, 1961) and Tomás Moro Simpson's *Formas lógicas, realidad y significado* (Logical forms, reality and meaning, 1964) are examples of philosophical analysis in Argentina. In 1972 Eduardo Rabossi (b. 1930), who has published extensively on human rights, founded the *Sociedad Argentina de Analísis Filosófico* (SADAF), which, as its name indicates, is committed to the advancement of the analytic tradition and publishes the journal *Análisis Filosófico*.

The influence of philosophical analysis is also evident in Brazil, particularly in *Manuscrito*, a journal published by the Center of Logic, Epistemology and Philosophy of Science at Campinas. In Mexico, Alejandro Rossi (b. 1932), Fernando Salmerón (1925–1997), and Luis Villoro (b. 1922) founded *Crítica* in 1967, a journal devoted to discussions from an analytic perspective. And the Instituto de Investigaciones Filosóficas of Mexico has been actively engaged in supporting the work of analytic philosophy.

While we can speak of a period of stability in the development of philosophical analysis in Latin America, there has been widespread political instability in many countries of the region. As a result, many outstanding analytic philosophers had to leave Latin America. Bunge was a professor of philosophy at the University of Buenos Aires from 1957 to 1963, but has worked at McGill University in Montreal since 1966. Hector-Neri Castañeda (Guatemala, 1924–1991) worked in the United States for most of his career, and Ernest Sosa (b. 1940) and Jorge Gracia (b. 1942) came to the United States from Cuba. Apart from contributions in the history of philosophy, metaphysics, and hermeneutics, Gracia has published in many areas of Latin American philosophy, including the impact of philosophical analysis in Latin America. Sosa works primarily in epistemology and metaphysics and was recently elected president of the Eastern Division of the American Philosophical Association; he is active in the promotion of analytic philosophy in Latin America.

Philosophical analysis is generally recognized as an important philosophical current in Latin America and analytic philosophers have the support of several institutes and journals, but there is some animosity in some quarters against this philosophical approach. Indeed, some Latin American philosophers have explicitly accused analytic philosophers of turning a blind eye to social injustice and the pressing political and economic issues that plague the region.

The philosophy of liberation. One current within the Latin American philosophical tradition that puts social concerns at the center is the philosophy of liberation. For this movement, the fundamental task of philosophy consists in the social and national liberation from the unjust relations, such as that of dominator-dominated, which have traditionally characterized Latin American philosophy. The philosophy of liberation is rooted in the political discourse of marginalized and exploited segments of society.

This current grew out of liberation theology, which in turn began in Peru and Brazil. Its origins can be traced to the 1970s in Argentina, to a group of thinkers that included Arturo Andrés Roig (b. 1922), Horacio Cerutti Guldberg (b. 1950), and Enrique Dussel (b. 1934). Because of the political turmoil during this period, many of these philosophers were forced into exile, thus disrupting the continuity of the movement and leading to the creation of various distinct strands of the philosophy of liberation. In spite of differences, however, they share a common concern with what it means to do philosophy from the periphery—that is, from the condition of dependence that these thinkers claim characterizes Latin American culture. The philosophy of liberation has been influenced by Marxist and Catholic ideas and is one of the most active philosophical currents in Latin America.

Identity. The problem of identity in Latin American philosophy has two dimensions: the identity of Latin American thought and cultural identity. In dealing with these aspects of the problem, philosophers tend to favor either what may be called a national approach or a continental—in a purely geographical sense—approach. Mariátegui, for example, addressed Peru's reality, applying Marxist principles in order to solve the problems facing Peruvians, *not* Latin Americans in general. In contrast, the Cuban thinker José Martí addressed issues of *nuestra América* (our America), emphasizing what is common to all the nations that comprise the region. Both thinkers prepared the way for the exploration of what it means to speak of Latin America and of Latin American philosophy.

The question of the existence and character of Latin American philosophy was first explicitly raised by Zea and Frondizi in the 1940s, although related questions had been alluded to even earlier by Alberdi. According to Alberdi, a Latin American philosophy must have a social and political character intimately related to the most vital needs of the continent. Because he conceives philosophy as an instrument for social, political, and economic change, Alberdi rejects metaphysics and other "pure and abstract" philosophical fields.

Zea's work extended the discussion of the meaning of Latin American philosophy. His culturalist perspective, according to which philosophy is intimately related to the culture and history from which it emerges, has won many adherents. Supporters find in this approach to defining philosophy a way of opening space for contributions that do not fall under the umbrella of the European and Anglo-American philosophical traditions under whose shadows they tend to remain marginalized. Abelardo Villegas (Mexico, 1934–2001), Ricaurte Soler (Panama, 1932–1994), and Guillermo Francovich (Peru, 1911–1990) are just three of the many philosophers who have adopted Zea's view. In Venezuela, Ernesto Mayz Vallenilla (b. 1925) has addressed some of these issues both in his work and in his capacity as public educator.

The problem facing philosophers as they grapple with the issue of the identity of Latin American nations, peoples, and intellectual traditions has become even more complicated as these discussions have entered the United States. Philosophers concerned with the place of Hispanics or Latinos in the United States explore questions related to what happens to the identity of Mexicans, Cubans, Colombians, and other Latin Americans who

immigrate to the United States. Is there a term that can capture the identity of this diverse group? If so, which term? Or should we give up on the enterprise altogether? These questions take on particular relevance in light of the discussion of group rights. Latin American thinkers working in the United States on such issues include Ofelia Schutte (Cuba, b. 1945) and Gracia.

Latin American philosophy has a rich and variegated history. Latin American philosophers have a tradition of concern for the specific social and political problems that plague the population of the Americas. But they remain engaged with the universal concerns that have characterized philosophy since its inception—problems of truth, goodness, and justice, among others—that are not the product of any particular political structure, social context, or geographical location.

See also Aristotle; Augustine, St.; Bergson, Henri; Caso, Antonio; Comte, Auguste; Continental Philosophy; Croce, Benedetto; Farias Brito, Raimundo de; Hartmann, Nicolai; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Identity; Ingenieros, José; Kant, Immanuel; Korn, Alejandro; Liberation Theology; Logical Positivism; Molina Garmendia, Enrique; Moore, George Edward; Neoplatonism; Nietzsche, Friedrich; Ortega y Gasset, José; Plato; Positivism; Romero, Francisco; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Scheler, Max; Schopenhauer, Arthur; Sosa, Ernest; Spencer, Herbert; Suárez, Francisco; Thomas Aquinas, St.; Unamuno y Jugo, Miguel de; Utilitarianism; Varona y Pera, Enrique José; Vasconcelos, José; Vaz Ferreira, Carlos; Vitalism; Vitoria, Francisco de; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ardao, Arturo. "Assimilation and Transformation of Positivism in Latin America." *Journal of the History of Ideas* 24 (1963): 515–522.

Beuchot, Mauricio. *The History of Philosophy in Colonial Mexico*. Translated by Elizabeth Millán. Foreword by Jorge J. E. Gracia. Washington, DC: The Catholic University of America Press, 1998.

Crawford, William Rex. *A Century of Latin American Thought*. 3rd ed. New York: Praeger, 1966.

Cruz Costa, João. *A History of Ideas in Brazil: The Development of Philosophy in Brazil and the Evolution of National History*. Berkeley: University of California Press, 1965.

Dascal, Marcelo, ed. *Cultural Relativism and Philosophy: North and Latin American Perspectives*. Leiden, Netherlands: E. J. Brill, 1991.

Davis, Harold Eugene. *Latin American Social Thought: The History of Its Development since Independence, with Selected Readings*. Washington, DC: University Press of Washington, 1963.

Davis, Harold Eugene. *Latin American Thought: A Historical Introduction*. New York: Free Press, 1972.

Duran, Manuel, and William Kluback, eds. *Reason in Exile: Essays on Catalan Philosophers*. New York: Peter Lang, 1994.

Dussel, Enrique D. *Philosophy of Liberation*. Translated by Aquilina Martínez and Christine Morkovsky. New York: Orbis Books, 1985.

Fernández Retamar, Roberto. *Calibán and Other Essays*. Translated by Edward Baker. Minneapolis: University of Minnesota Press, 1980.

Ferrater Mora, José. *La filosofía en América*. Vol. 1, Actas del IX Congreso Interamericano de Filosofía. Caracas: Sociedad Venezolana de Filosofía, 1979.

Frondizi, Risieri. "Is There an lbero-American Philosophy?" *Philosophy and Phenomenological Research* 9 (1948–1949): 345–355.

Frondizi, Risieri. "On the Unity of the Philosophies of the Two Americas." *Review of Metaphysics* 4 (1951): 617–622.

Gaos, José, ed. *Antología del pensamiento de lengua española en la edad contemporánea*. Mexico City: Séneca, 1945.

Gaos, José. *En torno a la filosofía mexicana*. With an introduction by Leopoldo Zea. Mexico City: Alianza, 1980.

Gómez Martínez, José, ed. *Proyecto Ensayo Hispánico*. Available from www.ensayistas.org/.

Gómez Robledo, Antonio. *La filosofía en el Brasil*. Mexico: Imprenta Universitaria, 1946.

Gracia, Jorge J. E. *Filosofía hispánica: Concepto, origen y foco historiográfico*. Pamplona: Universidad de Navarra, 1998.

Gracia, Jorge J. E. *Hispanic/Latino Identity: A Philosophical Perspective*. Oxford: Blackwell, 2000.

Gracia, Jorge J. E., ed. *Latin American Philosophy Today*. Double issue of *Philosophical Forum* 20 (1-2) (1988–1989).

Gracia, Jorge J. E., and Mireya Camurati, eds. *Philosophy and Literature in Latin America: A Critical Assessment of the Current Situation*. Albany: State University of New York Press, 1989.

Gracia, Jorge J. E., and Iván Jaksić, eds. *Filosofía e identidad cultural en América Latina*. Caracas, Venezuela: Monte Avila, 1988.

Gracia, Jorge J. E., and Elizabeth Millán-Zaibert, eds. *Latin American Philosophy for the 21st Century: The Human Condition, Values, and the Search for Identity*. New York: Prometheus, 2004.

Gracia, Jorge J. E., and Eduardo Rabossi, Enrique Villanueva, and Marcelo Dascal, eds. *Philosophical Analysis in Latin America*. Dordrecht, Netherlands: Kluwer, 1984. More extensive Spanish edition: *El análisis filosófico en América Latina*. Mexico City: Fondo de Cultura Económica, 1985.

Hanke, Lewis. *Aristotle and the American Indians: A Study in Race Prejudice in the Mondern World*. Chicago: Henry Regnery, 1959.

Jaksi?, Iván. *Academic Rebels in Chile: The Role of Philosophy in Higher Education and Politics*. Albany, NY: State University of New York Press, 1989.

León-Portilla, Miguel. *Time and Reality in the Thought of the Maya*. 2nd ed. Norman: University of Oklahoma Press, 1988.

Liss, S. B. *Marxist Thought in Latin America*. Berkeley: University of California Press, 1984.

Löwy, M., ed. *Marxism in Latin America from 1909 to the Present: An Anthology*. Translated by M. Pearlman. Atlantic Highlands, NJ: Humanities Press, 1990.

Martí, Oscar R. "Is There a Latin American Philosophy?" *Metaphilosophy* 1 (1983): 46–52.

Mayz Vallenilla, Ernesto. *El problema de América*. Caracas, Venezuela: Ediciones de la Universidad Simón Bolívar, 1992.

Mendieta, Eduardo, ed. *Latin American Philosophy: Currents, Issues, Debates*. Bloomington: Indiana University Press, 2003.

Millán-Zaibert, Elizabeth, and Arleen Salles, eds. *The Role of History in Latin American Philosophy*. Albany, NY: State University of New York Press, 2005.

Miró Quesada, Francisco. "The Impact of Metaphysics on Latin American Ideology." *Journal of the History of Ideas* 29 (1963): 539–552.

Nuccetelli, Susana, and Gary Seay, eds. *Latin American Philosophy: An Introduction with Readings*. Upper Saddle River, NJ: Prentice Hall, 2004.

Nuccetelli, Susana. *Latin American Thought: Philosophical Problems and Arguments*. Boulder, CO: Westview Press, 2002.

Redmond, W. *Bibliography of the Philosophy in the Iberian Colonies of America*. The Hague: Nijhoff, 1972.

Rodó, José Enrique. *Ariel*. Translated by Margaret Sayers Peden. Foreword by James Symington. Prologue by Carlos Fuentes. Austin: University of Texas Press, 1988.

Roig, Arturo Andrés. "Nuestra América frente al panamericanismo y el hispanismo: la lectura de Leopoldo Zea." In *América Latina. Historia y destino. Homenaje a Leopodo Zea*, edited by Horacio Cerutti-Guldberg. Mexico City: UNAM, 1992: 279–284.

Roig, Arturo Andrés. *Teoría y crítica del pensamiento latinoamericano*. Mexico City: Fondo de Cultura Económica, 1981.

Romero, Francisco. *Sobre la filosofía en America*. Buenos Aires: Raigal, 1952.

Sáenz, Mario, ed. *Latin American Perspectives on Globalization: Ethics, Politics, and Alternative Visions*. Lanham, MD: Rowman & Littlefield, 2002.

Sánchez Reulet, A., ed. *Contemporary Latin American Philosophy: A Selection with an Introduction and Notes* Translated by W. R. Trask. Albuquerque: University of New Mexico Press, 1954.

Sasso, Javier. *La filosofía latinoamericana y las construcciones de su historia*. Caracas, Venezuela: Monte Ávila, 1998.

Schutte, Ofelia. *Cultural Identity and Social Liberation in Latin American Thought*. New York: State University of New York Press, 1993.

Stabb, Martin S. *The Dissenting Voice: The New Essay of Spanish America, 1960–1985*. Austin: University of Texas Press, 1995.

Stabb, Martin S. *In Quest of Identity: Patterns in the Spanish American Essay of Ideas, 1890–1960*. Chapel Hill: University of North Carolina Press, 1967.

Woodward, R. L., ed. *Positivism in Latin America: 1850–1900: Are Order and Progress Reconcilable?* Lexington, MA: Heath, 1971.

Zea, Leopoldo. *The Latin-American Mind*. Translated by James H. Abbott and Lowell Dunham. Norman: University of Oklahoma Press, 1963.

Zea, Leopoldo. *El pensamiento latinoamericano*. Barcelona: Ariel, 1976.

Zea, Leopoldo. *Positivism in Mexico*. Austin: University of Texas Press, 1974.

Zea, Leopoldo. *The Role of the Americas in History*. Savage, MD: Rowman & Littlefield Publishers, 1992.

Jorge J. E. Gracia (2005)
Elizabeth Millán-Zaibert (2005)

LAVATER, JOHANN KASPAR

(1741–1801)

Johann Kaspar Lavater, the German-Swiss poet, physiognomist, and theologian, was born in Zürich. He studied at the gymnasium there under the literary critics Johann Jakob Bodmer and Johann Jakob Breitinger. Later, in northern Germany, he attended the lectures of the Protestant pastor Johann Jakob Spalding, who, influenced by the Earl of Shaftesbury and the English moralists, sought to reconcile reason and sentiment and stressed the moral and religious conscience. While in northern Germany Lavater also met Johann Georg Sulzer, Moses Mendelssohn (whom he later tried to convert to Christianity), the dramatist Friedrich Gottlieb Klopstock, and other persons of note. Returning to Zürich in 1764, he held various posts in churches there from 1769 on. He traveled widely in Germany, was acquainted with many culturally important people, and was one of the most sought-after and famous persons of that time. As a poet, he published a volume of religious verse, *Christlicher Lieder* (1771), and two epic poems in the manner of Klopstock, *Jesus Messias* (1780) and *Joseph von Arimathia* (1794). Because of his opposition to the Zürich government, Lavater was forced to move to Basel in 1796. He returned, only to be wounded during the French capture of Zürich in 1799. He died of this wound in 1801.

Lavater is chiefly known as a physiognomist. His theories were expounded in two main works, *Von der Physiognomik* (Leipzig, 1772) and *Physiognomische Fragmente zur Beforderung der Menschenkenntnis und Menschenliebe* (Physiognomic fragments for furthering the knowledge and love of man; 4 vols., Winterthur, Switzerland, 1775–1778). Johann Gottfried Herder and Lavater's close and longtime friend Johann Wolfgang von Goethe both collaborated on the latter work. Lavater claimed independence from traditional physiognomy dating from the time of Aristotle, but his independence was chiefly a mat-

ter of superficial knowledge of the tradition. He supported the classical view that the human body is influenced in shape by the character of the person, and vice versa; but his criteria were inconsistent and confused. There were two main reasons for his unprecedented success: First, his lively and simple manner of exposition that followed the pattern of the "popular philosopher"; and second, the psychology of character at the base of his theory.

Lavater stressed "feeling" and such spiritual qualities as inspiration and creative genius, which were being widely discussed in the eighteenth century. The native language of genius, and of virtue and wisdom, could become known only by studying the human form. Man is the measure of truth. That which harmonizes in form with a man, and is a part of him, is what exists for him. There is no absolute truth, but only a subjective experiencing. Therefore feeling should be cultivated, as it is in the genius. Lavater's psychology of genius, which gave emotions a place beside reason, was an important link between Pietism and sentimentalism on the one hand, and *Sturm und Drang* on the other. Lavater was severely criticized—notably by Georg Christoff Lichtenberg—but his handsomely printed volumes, with their illustrations, and his complimentary analyses of various influential contemporary figures were widely read.

As a writer of religious and devotional literature, Lavater was equally influential. His religious views were based on a belief in inner light, making his subjectivism a mystical and sentimental anthropomorphic theology. God is what satisfies the needs of man. The Bible is historically true but it is to be interpreted subjectively. Lavater was strongly convinced of the magical force of grace and prayer, and was strongly interested in miracles and prophecies. He was therefore drawn to spiritualism and mesmerism.

See also Aristotle; Goethe, Johann Wolfgang von; Herder, Johann Gottfried; Lichtenberg, Georg Christoph; Mendelssohn, Moses; Miracles; Pietism; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Sulzer, Johann Georg.

Bibliography

ADDITIONAL WORKS BY LAVATER

Aussichten in die Ewigkeit in Briefen an Zimmermann. 2 parts. Zürich, 1768–1769; 2nd ed. and 3rd part, 1773; 4th part, 1778.

Geheimes Tagebuch von einem Beobachter seiner selbst. 2 vols. Leipzig, 1771–1773.

Pontius Pilatus oder die Menschen in allen Gestalten, oder die Bibel im Kleinen und der Mensch im Grossen. 4 parts. Zürich, 1782–1785.

Nathanael, oder die Göttlichkeit des Christenthums. Winterthur, Switzerland, 1786.

Ausgewählte Werke. Edited by E. Stähelin. 4 vols. Zürich: Zwingli, 1943.

WORKS ON LAVATER

Bracken, Ernst von. *Die Selbstbeobachtung bei Lavater, Beitrag zur Geschichte der Idee der Subjektivität im 18. Jahrhunderts.* Münster, 1932. Dissertation.

Forssmann, J. *Lavater und die religiösen Strömungen des 18. Jahrhunderts.* Riga, 1935.

Janentzky, Christian. *J. K. Lavater.* Frauenfeld, Switzerland, 1928.

Maier, Heinrich. *An der Grenze der Philosophie Melanchton, Lavater, D. F. Strauss.* Tübingen, 1909.

Muncker, Franz. *J. F. Lavater, Eine Skizze seines Lebens und Wirkens.* Stuttgart, 1883.

Vömel, Alexander. *J. K. Lavater 1741–1801. Ein Lebensbild.* Elberfeld, 1923.

Giorgio Tonelli (1967)

LAVELLE, LOUIS

(1883–1951)

Louis Lavelle, the French philosopher, was born in Saint-Martin-de-Villéréal, in southwestern France. He was professor of philosophy at the Sorbonne from 1932 to 1934 and at the Collège de France from 1941 until his death. In a time of reaction against speculative system-building, Lavelle boldly elaborated an extensive system combining elements of the French *philosophie de l'esprit* and existentialism. Convinced that the modern world needs basic security, Lavelle, like other existentialist thinkers, sought philosophical and moral certitude in the experience of the self, "pure inwardness," and "absolute existence." Unlike such philosophers as Jean-Paul Sartre, who "disintegrated" the human universe inherited from tradition, Lavelle, like Karl Jaspers and Karl Barth, attempted to "reintegrate" the basic experiences of humanity in a novel form. In his spiritualistic interpretation of the self Lavelle continued the French tradition of Nicolas Malebranche, Maine de Biran, Octave Hamelin, Henri Bergson, and Maurice Blondel.

METAPHYSICS OF PARTICIPATION

Metaphysics was for Lavelle "the science of spiritual inwardness." According to him, Immanuel Kant had shown that we cannot find true reality on the side of the object, or thing, because objects and the world they compose cannot have independent existence. The essence of things resides in their relation to a being for whom they are "objects." Consequently, in the search for true or

absolute reality we must turn toward the act of consciousness, the "inwardness" of the human being. Thus Lavelle's central preoccupation was to discover and describe the fundamental relation between our innermost being and the Absolute.

Lavelle pointed out that there is a "primitive act" upon which our very being depends, as well as the being of the entire world. It is our primordial experience of being part of the world, in which act we find ourselves also "participating" in something that infinitely transcends us—the Act (Absolute Being, God). From a subtle dialectic description of this spiritual act of "participation" flow the broad lines of Lavelle's doctrine.

ONTOLOGY OF SPIRITUALISTIC EXISTENTIALISM

The originality of Lavelle's conception of the nature of beings in their relation to Being consists in his introducing a dynamic and "actualistic" content into the traditional themes of Aristotelian ontology. His approach yields a finalistic and optimistic view of the universe and human destiny.

All experiences of humankind emerge against the background of the limited individual being, participating in the Absolute Being. By their relation to participation, which is constant and eternal, individual beings establish their relation to the world, and through the notions of essence and existence they establish their spiritual identity. The Absolute Being is pure actuality, the infinite source of existential dynamism, and an endless reservoir of all possible forms or essences, from which individual beings receive their own limited existence. In spite of this direct and continuous dependence of the individual on his source, actualism is reconciled in Lavelle's thought with temporal progression, dynamism with formal immobility, and human freedom is safeguarded by the self-creativity of the individual. Indeed, from the human point of view, participation is a pursuit of an ideal that constantly moves ahead of our efforts. In this pursuit we create our spiritual self, and our experiences, moving onward, progressively acquire a unique form. Our effort in life is meant to discover this form, which has its prototype in the reservoir of Being and is our spiritual essence. The accomplishment of our essence at our death means the radical passage from limited existence into transfinite Being. Thus participation appears as the means of humanity's ultimate redemption, toward which everything occurring in the universe converges.

The world is the interval that separates pure Act (Being) from the limited act of participation (human existence). Matter, in limiting the spirit, offers the resistance necessary for the self to transcend itself. The world comprises three modes of reality: the world of things, that of ideas, and that of individual beings (consciousnesses). The material world plays the necessary role of separating beings; ideas give spiritual meaning to things. The world of individual consciousnesses is necessarily conscious because the essence of the Absolute Being from which they proceed is itself perfect inwardness; as such it is eternally fecund and intended to communicate the creative act to beings which, in turn, propagate it in self-creation.

ETHICS OF CONSENT

In Lavelle's moral philosophy an unusual meaning is given to existential themes, such as freedom, human destiny, and solitude. Lavelle had a constructive conception of man's vocation and of the ideal of life.

Freedom is the essence of man. But whereas the Absolute Act is synonymous with absolute freedom, man, the participating act, is limited by the "natural spontaneity" of the instinct. Consequently, the life of the spirit, which he proposed as the ideal of human life, is a fighting toward gradual liberation from the passivity peculiar to instinct. We become fully human by subordinating natural spontaneity to reflection and rational discipline. Human freedom originates in this process; and this conversion of spontaneity into freedom is the real vehicle of participation. The spiritual being, like the Leibnizian monad, is endowed with potentialities for the accomplishment of its preestablished essence. Our vocation is to seek to make our actual selves coincide with the "better part of ourselves," which represents these potentialities. This self-searching and self-controlling effort presupposes an "act of consent" to our vocation of the spirit. In opposition to other existentialist thinkers who glorify the "exceptional instant," Lavelle rehabilitated everyday existence, seeing even in the least significant instant an opportunity for consent to the self-creative effort and, thereby, an opportunity for participation in the Absolute.

Finally, the theme of solitude was reconciled with that of human communion insofar as the ideal of wisdom was seen to lie in the union between a certain asceticism and everyday life and love.

See also Absolute, The; Aristotelianism; Barth, Karl; Being; Bergson, Henri; Blondel, Maurice; Continental Philosophy; Essence and Existence; Existentialism; Freedom; Hamelin, Octave; Jaspers, Karl; Kant, Immanuel; Maine de Biran; Malebranche, Nicolas; Sartre, Jean-Paul; Self.

Bibliography

WORKS BY LAVELLE

La dialectique du monde sensible. Strasbourg: Commission des publications de la Faculté des lettres, 1921.

La perception visuelle de la profondeur. Strasbourg: Commission des publications de la Faculté des lettres, 1921.

La dialectique de l'éternel présent. 3 vols. Vol. 1: *De l'être*. Paris: Alcan, 1928; Vol. 2: *De l'acte*. Paris, 1937; Vol. 3: *Du temps et de l'éternité*. Paris: Aubier, Éditions Montaigne, 1945.

La conscience de soi. Paris: Grasset, 1933.

La présence totale. Paris, 1934.

De l'acte. Paris: Aubier 1937; Paris: Flammarion, 1992.

Les puissances du moi. Paris, 1939.

L'erreur de Narcisse. Paris: Grasset, 1939.

Le mal et la souffrance. Paris, 1940.

La parole et l'écriture. Paris: L'artisan du livre, 1942.

La philosophie française entre les deux guerres. Paris: Aubier, 1942.

Du temps et de l'éternité. Paris, 1945.

Introduction à l'ontologie. Paris: Presses Universitaires de France, 1947.

Science, ésthetique, métaphysique: chroniques philosophiques. Paris: Albin Michel, 1967.

Psychologie et spiritualité. Paris: Albin Michel, 1967.

Règles de la vie quotidienn. Paris: Arfuyen, 2004.

TRANSLATIONS

The Meaning of Holiness. New York: Pantheon Books, 1956.

The Dilemma of Narcissus. Larson: Burdett, NY, 1993.

The Act of Presence, Works of Louis Lavelle. Introduction and tranlation into English of selected texts of Louis Lavelle by Robert Jones, posted at http://association-lavelle.chez.tiscali.fr/Robert%20Jones.pdf.

WORKS ON LAVELLE

Ainval, Christiane d'. *Une doctrine de la présence spirituelle: La philosophie de Louis Lavelle*. Paris: Nauwelaerts, 1967.

Bourbon-Busset, Jacques de. *La raison ardente*. Paris: Gallimard, 2000.

Chretien, Jean-Louis, *L'arche de la parole*. Paris: PUF, 1999.

Delfgaauw, B. M. I. *Het Spiritualistiche Existentialisme van Louis Lavelle*. Amsterdam, 1947.

Ecole, Jean, *Métaphysique de l'être, doctrine de la connaissance et philosophie de la religion chez Louis Lavelle*. Gênes: L'Arcipelago, 1994.

Foulquié, P. *L'existentialisme*, 109–122. Paris, 1946.

Grasso, Pier G. *L. Lavelle*. Brescia: La Scuola, 1949.

Nobile, O. M. *La filosofia di Louis Lavelle*. Florence: Sansoni, 1943.

Reymond, Christiane, *Autrui dans la Dialectique de l'éternel présent de Louis Lavelle*. Paris: Presses universitaires de France, 1972.

True, G. *De Jean-Paul Sartre à Louis Lavelle, ou désagrégation et réintegration*. Paris, 1946.

Anna-Teresa Tymieniecka (1967)

Bibliography updated by Tamra Frei (2005)

LAVOISIER, ANTOINE

(1743–1794)

Antoine Lavoisier played the central role in what has come to be known as the chemical revolution. He is credited with establishing that oxygen is an element and water its compound with hydrogen, refining experimental methods in chemistry, reforming chemical nomenclature along systematic lines, defining *element* operationally, and denying phlogiston a place in chemical explanation.

EARLY LIFE AND WORK

Lavoisier was born into a wealthy family of lawyers in 1743, and in preparation for a legal career attended the Collège des Quatre Nations (or Collège Mazarin), earning a baccalaureate in law in 1763. He pursued scientific interests under the guidance of the geologist Jean-Étienne Guettard (1715–1786), a family friend, and attended Guillaume-François Rouelle's (1703–1770) popular and influential lectures on chemistry and mineralogy at the Jardin du Roi. From 1763 Lavoisier assisted Guettard on field trips for the first geological survey of France. His first chemical work was a study of gypsum and plaster of Paris, which was read to the Academy of Sciences in 1765, to which he was elected in 1768. That year Lavoisier also joined the Ferme Générale, a private company collecting indirect taxes in return for a fixed payment to the Crown. This investment would secure his fortune, but also prove his downfall. In 1771 he married Marie Anne Paulze, the fourteen-year-old daughter of a senior member of the Ferme. Marie became a significant collaborator: She learned English to translate important scientific papers, assisted in the laboratory, and trained in the visual arts, providing the engravings for Lavoisier's *Traité Elémentaire de Chimie* (1789).

Lavoisier was active outside of chemistry, especially in economic and farming reform. As an academician, he pursued many technological projects in the service of the state, helping to investigate water supply and storage, food purity, ballooning, bleaching, and ceramics and to develop the metric system. From 1776 he was in charge of the production and administration of gunpowder, working from a laboratory in the Royal Arsenal.

Lavoisier's contributions to chemistry began at a time when advancing experimental techniques made clearer the atmosphere's active role in chemical reactions, but phlogiston, the principle of inflammability, still provided the prevailing framework for understanding combustion and calcination (the formation of metal oxides). In 1772 Louis-Bernard Guyton de Morveau (1737–1816)

reported to the Academy of Sciences that metals increase their weight on calcination. This was in tension with the phlogistonists' view that combustion and calcination involved loss of phlogiston to the air. Guyton de Morveau argued that the light phlogiston must "buoy up" the metal, but Lavoisier saw calcination instead as fixation of air in the calx. In the long and carefully constructed series of experiments that followed, Lavoisier studied the combustion and calcination of metals and nonmetals, measured the volumes of air absorbed or evolved, and weighed and investigated the solid products and the residual air. By 1778, drawing also on the experimental work of others, he was convinced that a particular component of air was involved in combustion, the "purest part of air" or "eminently respirable air," which combines with carbon to form fixed air (carbon dioxide). Lavoisier also noted during the 1770s that air was absorbed in the formation of phosphoric, sulfuric, and nitric acids and of fixed air, which was weakly acidic in solution. In papers read to the Academy of Sciences between 1776 and 1779 he concluded first that the acids were a chemical genus, containing air combined with different principles, and later that "eminently respirable air" contains the principle of acidity, which he called *principe oxigine* (later to become *oxygène*). Water he identified as oxygen combined with "inflammable air" (which he renamed hydrogen). Oxygen the gas was not *itself* the principle of acidity, though: Lavoisier saw gases also as a chemical genus, their common constituent being caloric, the matter of heat. Thus in combustion, substances combine with the oxygen principle, releasing caloric from oxygen gas, which explained why heat was evolved in the process. Experiments on animal respiration convinced him that respiration is a slow version of combustion, and in 1785 he extended his theory of acidity, accounting for the solution of metals in acids as wet calcination.

CHEMICAL REVOLUTION

These three theories—of combustion, acidity, and the gaseous state—gave Lavoisier a framework comprehensive enough to deny phlogiston its explanatory role. In 1785 he read "Réflexions sur le Phlogistique," a direct attack on the theory, to the Academy of Sciences. In 1787 he published, with Guyton de Morveau and others, a new nomenclature for chemistry, replacing a jumble of uninformative traditional names with a system for naming compounds based on their composition, reflecting the latest discoveries. This is largely still in use in modern chemistry.

Lavoisier published his most influential work, *Traité Elémentaire de Chimie*, in 1789. This combined a clear presentation of his own theories of gases, of combustion and acidity in part I, with (in parts II and III) a summary of less controversial material on acids, bases, and salts and on experimental methods. In the preface, he introduced his empirical definition of *elementhood*: rejecting the traditional speculations about the "simple substances," he proposed to treat as simple any substance that had not yet been decomposed in the laboratory.

After 1789 political revolution in France intervened increasingly in Lavoisier's activities, curtailing his scientific researches, though at first he was sympathetic to its aims. Scientific and administrative institutions of the ancien régime, in which he had played a prominent (though liberal and reforming) role, were successively abolished: the Ferme Générale in 1791 and the Academy of Sciences in 1793. Members of the Ferme were arrested in November 1793, and on May 8, 1794, were convicted of adulterating tobacco and withholding taxes from the government. Lavoisier was executed that same day, just after his father-in-law.

Lavoisier's achievement raises important historiographical and philosophical questions about progress in science. Lavoisier himself, writing in 1773, foresaw a revolution in chemistry, and his name appears throughout Thomas S. Kuhn's *Structure of Scientific Revolutions* (1970). In this technical sense the defeat of the phlogiston theory has been called a scientific revolution because: (1) it involved wholesale revision to theoretical interpretations of empirical evidence and accepted views of the relative simplicity of whole classes of substances (e.g., metals and their calxes); and (2) it was accompanied by a major reform of chemical nomenclature that embedded the oxygen theory in the very language of chemistry. The importance of his empirical definition of *elementhood* is less clear. It was not original to him, and it applied only selectively to his own list of elements.

See also Chemistry, Philosophy of.

Bibliography

WORKS BY LAVOISIER

Traité Elémentaire de Chimie. 2 vols. Paris: Cuchet, 1789. Translated by Robert Kerr as *The Elements of Chemistry* (Edinburgh, Scotland: William Creech, 1790), reprinted as David Knight (ed.) *The Development of Chemistry*, Vol. 2 (New York: Routledge/Thoemmes Press, 1998).

WORKS ABOUT LAVOISIER

Donovan, Arthur. *Antoine Lavoisier: Science, Administration, and Revolution*. Cambridge, MA: Blackwell, 1993.

Hendry, Robin Findlay. “Lavoisier and Mendeleev on the Elements.” In *Foundations of Chemistry* 7 (1) (2005): 31–48.

Holmes, Frederic Lawrence. *Antoine Lavoisier: The Next Crucial Year, or The Sources of His Quantitative Method in Chemistry*. Princeton, NJ: Princeton University Press, 1998.

Holmes, Frederic Lawrence. *Eighteenth-Century Chemistry as an Investigative Enterprise*. Berkeley: Office for History of Science and Technology, University of California at Berkeley, 1989.

Kitcher, Philip. *The Advancement of Science: Science without Legend, Objectivity without Illusions*. New York: Oxford University Press, 1993.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press, 1970.

Thagard, Paul. “The Conceptual Structure of the Chemical Revolution.” In *Philosophy of Science* 57 (2) (1990): 183–209.

Robin Findlay Hendry (2005)

LAVROV, PËTR LAVROVICH

(1823–1900)

Pëtr Lavrovich Lavrov was a Russian philosopher and social thinker, a major theoretician of Russian Populism and the leading exponent of a distinctive form of positivism in nineteenth-century Russian philosophy (also elaborated by Nikolai Mikhailovskii). Lavrov was born in Melekhov, the son of a landed gentleman and retired artillery officer. He was sent to the Artillery School in St. Petersburg in 1837 and received his commission upon graduating in 1842. In 1844 he joined the faculty of the Artillery School, and for more than twenty years (during which he rose to the rank of colonel), he taught mathematics and the history of science at military institutions in St. Petersburg. At the same time Lavrov read widely in philosophy and gained a reputation as a writer—first for his poetry and after 1858 for his scholarly essays in philosophy. In the 1860s, the increasing liberalism of his social views aroused the suspicion of the tsarist authorities. Arrested in 1866, he was exiled to the provinces in the following year. In 1870 he fled to Paris, where he played an active role in the Commune of 1871. After sojourns in London and Zürich, he settled in Paris in 1877. A friend of Karl Marx and Friedrich Engels, Lavrov became the voice of Russian socialism abroad and a revered figure in the international socialist movement. He died in Paris.

Lavrov developed an early interest in socialism through reading François Marie Charles Fourier and other leading socialists; he was particularly attracted to the ideas of Pierre-Joseph Proudhon and Aleksandr Herzen. Philosophically, Lavrov’s initial scientific orientation evolved in the direction of positivism rather than in the direction of the “materialism” that was prevalent in Russian radical circles of the day, among such thinkers as Nikolai Chernyshevskii and Dmitrii Pisarev. However, his positivistic philosophy was based more on German models than on Auguste Comte. Lavrov did not become acquainted with Comte’s writings until the middle of the 1860s; by then his thinking had been given strong direction by a close study of Immanuel Kant, G. W. F. Hegel, the neo-Kantian Albert Lange, and “young Hegelians” such as Ludwig Feuerbach and Arnold Ruge.

In his first important philosophical writings, which consisted of several long essays written between 1858 and 1861, Lavrov criticized materialism as a metaphysical system that unnecessarily restricts science to matter in motion. Distinguishing between material phenomena, conscious phenomena, and historical phenomena, he maintained that phenomena of the last two classes cannot be dealt with by the methods of the natural sciences. The phenomena of consciousness, in particular, require a “subjective,” introspective method, and furthermore, these phenomena must be regarded as scientifically primary, since every investigator must begin from the facts of his own consciousness. Calling this approach “anthropologism,” Lavrov developed it into a neo-Kantian positivism that, while it rejected supernaturalistic metaphysics and religion, did not reject moral imperatives. It stressed the thought and action of the free individual who finds in his own consciousness an absolute sanction to strive toward the realization of moral ideals such as individual dignity and social justice. While material phenomena are governed by universal natural laws, man’s conscious conviction that he is free is inescapable and thus may be taken as a foundation for practical philosophy. Moral ideals are ultimately grounded in man’s striving for pleasure, but in the consciousness of the cultivated individual they present themselves as nonegoistic, universal imperatives.

In his best-known philosophical work, *Istoricheskie pis’ma* (Historical letters), first published serially in the magazine *Nedelia* (Week) in 1868 and 1869, Lavrov continued his attack upon materialistic reductionism by applying “anthropologism” to history. Arguing that man can view history only “subjectively” and teleologically, he defined the goal of history as the physical, moral, and intellectual development of the individual. On this basis he maintained that the “critically thinking individuals” who have already achieved such development have a moral obligation to extend the opportunity for development to the masses, whose toil has given the privileged

few the leisure and the resources needed for self-cultivation. Lavrov asserted that in coming to understand the defects of existing social institutions and in actively striving to reform them, the "critically thinking individuals" both discharge their "debt to the people" and serve as the moving forces of history. He envisaged a future in which all social institutions will conform to man's natural needs and the coercive institutions of the state will be all but eliminated. *Istoricheskiye Pis'ma* had a great impact on the Russian revolutionary youth of the 1870s.

Lavrov was able to develop his socialist program more explicitly abroad, where he was free from tsarist censorship. From 1873 to 1876 he edited the journal *Vperyed!* (Forward!), the chief organ of Russian Populist socialism—a form of agrarian socialism, inspired by Herzen and Chernyshevskii, which stressed the Russian village commune and the possibility it afforded Russia of moving directly to a socialist order, thus bypassing the evils of capitalism. Lavrov's political theory was further elaborated in *Gosudarstvennyi element v budushchem obshchestve* (The state element in future society), published in London in 1876. Acknowledging the need for revolution, Lavrov at first stressed the value of preparatory education and propaganda. Later he came to condone revolutionary terrorism and was associated with the Russian extremist party, *Narodnaia volia* (The people's will).

In his later socialist views, which were closer to those of Marx, Lavrov gave more attention to class conflict and to the process of production, but he never adopted a fully Marxist view of history or social dynamics. His emphasis remained moralistic and individualistic, with its focus on the development and activity of the "critically thinking individual." The philosophical outlook reflected in Lavrov's *Istoricheskiye Pis'ma* remained fundamentally unchanged in his last major work, which consisted of two lengthy introductory volumes of an unfinished intellectual history titled *Opyt istorii mysly novogo vremeni* (Essay in the history of modern thought; Geneva, 1894).

See also Chernyshevskii, Nikolai Gavrilovich; Comte, Auguste; Engels, Friedrich; Feuerbach, Ludwig Andreas; Fourier, François Marie Charles; Hegel, Georg Wilhelm Friedrich; Herzen, Aleksandr Ivanovich; Kant, Immanuel; Lange, Friedrich Albert; Marx, Karl; Materialism; Mikhailovskii, Nikolai Konstantinovich; Pisarev, Dmitri Ivanovich; Positivism; Proudhon, Pierre-Joseph; Russian Philosophy; Socialism.

Bibliography

WORKS BY LAVROV

Sobranie sochinenii (Collected works). 11 numbers. Petrograd, 1917–1920.

Izbrannye sochineniia (Selected works). 4 vols. Moscow, 1934–1935.

Filosofiya i sotsiologiya. 2 vols. Moscow, 1965.

Historical Letters. Translated by James P. Scanlan. Berkeley, CA: University of California Press, 1967.

O religii (On religion). Moscow: Mysl, 1989.

WORKS ON LAVROV

Itenberg, B. S. *P. L. Lavrov v russkom revoliutsionnom dvizhenii* (P. L. Lavrov in Russian revolutionary movement). Moscow: Nauka, 1988.

Kazakov, A. P. *Teoriia progressa v russkoi sotsiologii kontsa XIX veka (P. L. Lavrov, N. K. Mikhailovskii, M. M. Kovalevskii)* (The theory of progress in late nineteenth-century Russian sociology [P. L. Lavrov, N. K. Mikhailovskii, M. M. Kovalevskii]). Leningrad: Izdatel'stvo leningradskogo universiteta, 1969.

Pomper, Philip. *Peter Lavrov and the Russian Revolutionary Movement*. Chicago: University of Chicago Press, 1970.

Semenkova, T. G. *Ekonomicheskie vzgliady P. L. Lavrova* (The economic views of P. L. Lavrov). Moscow: Vysshaia shkola, 1980.

Venturi, Franco. *Il populismo russo*. Rome, 1952. Translated into English by Frances Haskell as *Roots of Revolution*. New York, 1960.

Volodin, A. I. *Lavrov*. Moscow: Molodaia gvardiia, 1981.

Zenkovsky, V. V. *Istoriya Russkoy Filosofii*. 2 vols. Paris, 1948–1950. Translated into English by George L. Kline as *A History of Russian Philosophy*. 2 vols. New York and London, 1953.

James P. Scanlan (1967)

Bibliography updated by Vladimir Marchenkov (2005)

LAW, PHILOSOPHY OF

In addition to the detailed survey entries *Philosophy of Law, History of*, and *Philosophy of Law, Problems of*, the Encyclopedia includes the following entries in which legal theories and concepts are discussed: *Analytic Jurisprudence*; *Historical School of Jurisprudence*; *Justice*; *Legal Positivism*; *Legal Realism*; *Natural Law*; *Property*; *Punishment*; *Responsibility, Moral and Legal*; *Rights*; and *Sovereignty*. See "Philosophy of Law, History of," and "Philosophy of Law, Problems of," in the index for entries on philosophers and legal theorists who have concerned themselves especially with questions in the philosophy of law.

LAW, WILLIAM
(1686–1761)

William Law, the English devotional writer, controversialist, theologian, and mystic, was a fellow of Emmanuel College, Cambridge. As a nonjuror, he refused to take the oath to King George I and thus terminated his career at the university and in the church. For a time he was a tutor in the household of Edward Gibbon, grandfather of the historian. His later life was virtually without incident, and after years of retirement, he died in his native village of King's Cliffe, Northamptonshire.

Law is best known as a devotional writer and especially for his *A Serious Call to a Devout and Holy Life* (1728); but his importance in the history of thought lies elsewhere, in his resistance to latitudinarianism, his defense of morality, his attack on deism, and his mystical writings.

Law was a formidable controversialist, and in his *Three Letters to the Bishop of Bangor* (1717) he brought remorseless logic to bear on Benjamin Hoadly's lax view of the nature of the church. Bernard de Mandeville had contended in the *Fable of the Bees* that private vices are actually public benefits; Law subjected the work to rigorous examination and showed that the canons of morality cannot be understood in terms of such specious sophistries. His most serious and celebrated work was his attack on deism. In *Christianity as old as the Creation,* Matthew Tindal argued that reason is the only test of truth; insofar as Christianity is valid, it rests on rationalist principles that owe nothing to revelation. Law's *Case of Reason* was a closely argued refutation of the prevailing rationalism of the period. Human reason is not able, by itself, to encompass all knowledge, nor is it sufficient to test all truth. Those who exalt natural religion are exposed to the same criticism as those who accept revelation without question. The universe is less simple and the ways of God are more mysterious than the arrogance of rationalism admits. Law shared with George Berkeley and Joseph Butler the credit for terminating the active phase of the deistic controversy.

Law's later writings reflect the profound influence that mysticism (especially as expounded by Jakob Boehme) came to exercise over his thought. He reached the conclusion that real knowledge is "the communion of the knowing and the known." To convey his new insights, Law organized his teaching in the form of "myth." He believed that mysticism gives birth to symbols within which its truth can live. Law felt that he had penetrated to a deeper understanding of human nature and that it could best be interpreted through a grasp of the meaning of the myth of the Fall on the one hand and through an understanding of divine self-communication in love on the other ("Love is the first *Fiat* of God"). Law's mystical teaching about life was related to a restatement of orthodox Christianity. He expounded the atonement with great beauty and insight and believed that the Trinity was the most illuminating way to describe the self-unfolding of the Eternal.

Law's mystical writings were perplexing to thinkers of the eighteenth century (see John Wesley's letter to Law about mysticism), but his *Serious Call* exercised a profound influence at the time (especially on Samuel Johnson and John Wesley) and is still considered a classic work on the Christian life.

See also Berkeley, George; Boehme, Jakob; Butler, Joseph; Deism; Johnson, Samuel; Mandeville, Bernard; Mysticism, History of; Rationalism; Religion and Morality; Tindal, Matthew.

Bibliography

WORKS BY LAW

The Works of William Law. 9 vols. Edited by G. Moreton. London, 1892.

Selected Mystical Writings of William Law, edited by S. Hobhouse. London: C.W. Daniel, 1938.

WORKS ON LAW

Cragg, G. R. *Reason and Authority in the Eighteenth Century.* Cambridge, U.K.: Cambridge University Press, 1964.

Overton, J. H. *The Life and Opinions of the Rev. William Law.* London, 1881.

Talon, H. *William Law.* London: Rockliff, 1948.

Gerald R. Cragg (1967)

LAWS, SCIENTIFIC

On the standard picture there are three kinds of facts. Some facts cannot have been otherwise. These facts include the conceptual truths (e.g., the fact that Rebecca is taller than Abe if Abe is shorter than Rebecca) and the mathematical truths (e.g., that 2 + 1 = 3). The remaining facts (i.e., the "contingent" ones) are divided between the other two classes: (1) the laws of nature (and their contingent logical consequences), such as the fact that all copper objects are electrically conductive, and (2) the "accidents," such as the fact that Jones has ten fingers and the fact (one can suppose it is a fact, though humanity

may never discover it) that there never exists a solid gold cube larger than a cubic mile.

It is widely believed that one of science's chief goals is to discover the laws of nature. Philosophers have studied the role that the concept of natural law plays in scientific reasoning.

LAWS VERSUS ACCIDENTS: NECESSITY AND COUNTERFACTUALS

An accident just happens to obtain. A gold cube larger than a cubic mile could have formed, but the proper conditions for it to have done so happened never to arise. In contrast, it is no accident that an electrically insulating copper object never formed, since the natural laws prohibit such a thing. In short, events must conform to the laws of nature—the laws have a kind of necessity—whereas accidents are mere coincidences. The kind of necessity characteristic of laws (and their logical consequences) is usually called *nomic* or *physical* necessity to distinguish it from various stronger varieties of necessity (such as logical, conceptual, and metaphysical necessity) possessed by various facts that cannot have been otherwise.

Had Bill Gates wanted to build a large gold cube, then there would have been a gold cube exceeding one cubic mile. But even if Gates had wanted to build an electrically insulating copper object, all copper objects would still have been electrically conductive, since events are obliged to conform to the natural laws. In other words the laws govern not only what actually happens but also what would have happened under various circumstances that did not actually happen. The laws support *counterfactuals* (i.e., facts expressed by statements of the form "Had p been the case, then q would have been the case"). Consequently, scientists use the laws to ascertain, for example, the conditions that would have prevailed on Earth had Earth been ten times nearer to the Sun. The laws are preserved under this counterfactual supposition. In contrast, an accident would not still have held, had p been the case, for some p that is *nomically possible* (i.e., consistent with all the laws' logical consequences).

Counterfactuals are notoriously context sensitive. For example, when one is emphasizing how baseball pitching talent has declined over the years, one might correctly remark that were Babe Ruth playing in the major leagues today, he would hit an astounding 120 home runs in a single season. But in a different context, one might correctly remark that were the Babe playing today, he would hit only ten homers per season, since by now he would be an old man. Which facts are to be held fixed under some counterfactual supposition, and which are allowed to vary, depends somewhat on one's interests in entertaining that supposition. But it appears that in any context, the laws would still have held under every nomic possibility. This idea is sometimes called *nomic preservation*.

LAWS VERSUS ACCIDENTS: EXPLANATION AND INDUCTION

Laws have an explanatory power that accidents lack. For example, a certain powder burns with yellow flames, not another color, because the powder is a sodium salt and it is a law that all sodium salts, when ignited, burn with yellow flames. The powder had to burn with yellow flames considering that it was a sodium salt. This "had-to-ness" reflects the law's necessity. In contrast, that a couple has two children is not explained by the fact that all the families on the couple's block have two children, since this fact is accidental. Were a childless couple to move onto the block, this couple would not encounter an irresistible opposing force.

One believes that it would be mere coincidence if all U.S. presidents elected in years ending in 0 died in office. Hence, one's discovery that Warren Harding (elected in 1920) died in office fails to justify raising one's confidence that whoever was elected in 1840 died in office. A candidate law is confirmed differently: That one sample of a given chemical substance melts at 383 degrees (in standard conditions) is evidence, for every unexamined sample of that substance, that its melting point is 383 degrees (under standard conditions). This difference in inductive role between laws and accidents seems related to the fact that laws, unlike accidents, express similarities among things that reflect their belonging to the same *natural kind*. The electron, the emerald, and the electromagnetic force are all natural kinds, whereas the families on a block and the gold cubes do not form natural kinds (though gold objects, cubical and otherwise, constitute a natural kind).

DIFFICULTIES DISTINGUISHING LAWS FROM ACCIDENTS

The previous discussion is the standard view of the scientifically relevant differences between laws and accidents. Insofar as the same claims play all these special roles, scientific reasoning apparently recognizes an important distinction here, which philosophers label as the difference between accidents and laws. (Obviously, this distinction involves what laws do rather than which facts happen to be called "laws"; Archimedes' principle of buoyancy, the

axioms of quantum mechanics, and Maxwell's equations are all laws of physics.) However, it is notoriously difficult to capture the laws' special roles precisely.

For example, suppose one tries to distinguish laws from accidents on the grounds that laws support counterfactuals differently from accidents. That a car's maximum speed on a dry, flat road is a certain function of its gas pedal's distance from the floor is not a law (since it reflects accidental features of the car's engine). Nevertheless, this function supports counterfactuals regarding the car's maximum speed had the pedal been depressed to one-half inch from the floor, though not had certain changes been made to the engine. Indeed, all gold cubes would still have been smaller than a cubic mile even if Jones had been wearing a different shirt today. Of course, there are some nomic possibilities under which the gold-cubes generalization would not still have held. But circularity threatens if one uses the concept of nomic possibility to delimit the range of counterfactual suppositions under which a fact must be preserved for that fact to qualify as a logical consequence of the laws.

Likewise, a car's pedal-speed function, despite being accidental, can apparently be confirmed inductively. Moreover, when coupled with the road's condition and the pedal's position, it can explain the car's maximum speed. So although a fact's lawhood apparently makes a difference to science, it is difficult to identify exactly what difference it makes. This problem's stubbornness has led some philosophers to suggest that it is a mistake to distinguish laws sharply from accidents. There are merely various facts, each having a range of counterfactual suppositions under which it is preserved.

ARE THERE LAWS OUTSIDE OF FUNDAMENTAL PHYSICS?

Some so-called laws are plainly accidents—if they are true at all. Kepler's first law of planetary motion (that planets trace elliptical orbits) presupposes that the planets' masses happen to be negligible compared to the Sun's (since otherwise, the planets would be disturbed by their mutual gravitational influences) and that no body collides with a planet, knocking it out of its orbit. Some philosophers believe that the fundamental laws of microphysics (whatever they turn out to be) are the only genuine natural laws. This opinion is sometimes prompted by the fact that all events are ultimately nothing but the outcome of microphysical processes governed by the fundamental laws.

However, along with the laws of fundamental physics there might seem to be additional laws holding independent of the universe's microphysical details. The second law of thermodynamics, according to which the entropy of a closed system is likely to increase, seems not to reflect any peculiarities of the fundamental forces governing the universe's ultimate constituents; even if gravity had been twice as strong as it actually is, for example, the perfume molecules from a recently opened bottle would be more likely to spread quickly throughout the room than to remain in the bottle. Likewise, the principle of natural selection, according to which fitter traits are more likely to increase their prevalence in a closed population, seems like it would still hold whatever the laws of fundamental physics might have been.

Additionally, the second law of thermodynamics appears to require that certain initial microconditions be rare—for example, that the perfume molecules within recently opened bottles not usually be arranged so that whenever one molecule threatens to escape from the bottle, another happens to come along and knock it back inside. That the perfume molecules in recently opened bottles are indeed not so coordinated would seem to be an accident rather than a nomic necessity. Accordingly, perhaps the second law is not a law at all.

The principle of natural selection is perhaps also not a law, but a conceptual truth. That a trait is "fitter" in a given environment may simply mean that it is more likely to become increasingly common in subsequent generations. Nevertheless, both the second law of thermodynamics and the principle of natural selection appear to undergo inductive confirmation, to support counterfactuals, and to explain events in the manner of natural laws.

LAWS OF INEXACT SCIENCES: THE PROBLEM OF *CETERIS PARIBUS*

A "special" or "inexact" science (such as anatomy, ballistics, ecology, economics, marketing, or psychology) might appear to seek (or perhaps even to have already found) facts that in these sciences play the various roles characteristic of laws. However, there are three main obstacles to regarding Boyle's law (that the product of a gas's pressure P and its volume V is constant) as a law of gases, to regarding Gresham's law (that agents hoard sound money and spend currency of more dubious value) as a law of economics, or to regarding the area law (that larger islands have greater biodiversity) as a law of island biogeography. Each of these obstacles has persuaded some philosophers to deny that inexact sciences have laws.

First, any such "law" comes with a *ceteris paribus* qualification. Though *ceteris paribus* means roughly "all

other things being equal," a given qualification may be better captured as the idea that the specified correlation holds "normally," "in the ideal case," "in the absence of disturbing factors," or as long as certain other factors have certain values. The law is "hedged" in some way. For example, the gas in a container departs significantly from Boyle's law when its temperature is changed, some of the gas escapes, or the pressure is high. These circumstances are ruled out by the *ceteris paribus* proviso to Boyle's law. But what exactly does "PV is constant, *ceteris paribus*" mean?

If it means that PV is constant unless it is not, then the "law" is a trivial, noncontingent truth rather than an interesting discovery. If instead *ceteris paribus* is shorthand for a list of every factor allowed by fundamental microphysics and able to cause a gas's PV to vary, then Robert Boyle could not have discovered his law, since he did not know some of these factors (e.g., gas molecules adhering to the container's walls or attracting one another). Alternatively, some philosophers contend that Boyle's law describes only fictitious "ideal gases" that lack any interfering factors. But then it is unclear how observations of actual gases could confirm Boyle's law or how knowledge of Boyle's law could justify scientists in using it to predict the behavior of actual gases. Boyle had neither the concept of an ideal gas nor an account of what makes a gas ideal (e.g., that it consists of molecules without mutual attraction and occupying no finite volume). Such an account is not part of Boyle's law. Rather, the extent to which an actual gas has constant PV is explained by the extent to which it resembles an ideal gas.

Apparently then, *ceteris paribus* in Boyle's law refers only to the disturbing factors of which Boyle was aware (high pressure, changes to the gas's temperature, and so forth). There may be no complete list of these factors. Obviously (to shift examples), Gresham's law does not apply if the society is wiped out, if its members believe that hoarding the sounder currency causes illness, and so forth. Part of understanding Gresham's law is knowing how to recognize whether some factor qualifies as disturbing. One can catch on to which factors these are without having to read a complete list of them. (Nonexperts may even [in an attenuated sense] understand the *ceteris paribus* proviso without being able to tell themselves whether some factor qualifies as disturbing, just as they understand other technical terms: by virtue of knowing who the relevant experts are to whom they should defer.)

LAWS OF INEXACT SCIENCES: THE PROBLEM OF TRUTH

However, societal events are ultimately nothing but the outcomes of microphysical processes. Certain sequences of microevents permitted by the fundamental laws of physics involve a society's members hoarding the weaker currency and spending the sounder. In one such sequence each member of the society happens whenever he or she spends money to forget momentarily which currency is sounder, because as chance would have it, some neuron in each agent's brain behaves at that moment in a manner that the fundamental microlaws deem extremely unlikely, but nevertheless possible. The *ceteris paribus* proviso to Gresham's law does not rule out this freakish sequence of events, since economists surely do not need to grasp the subtleties of fundamental microphysics to understand the proviso to Gresham's law.

In other words, not all exceptions to a macrolevel "law" can be specified in the vocabulary of the macroscience. For example, it might require physics (or at least neurology) to specify certain circumstance in which an agent would depart from a psychological "law." The *ceteris paribus* proviso fails to cover those exceptions.

This is the second obstacle to regarding inexact sciences as having genuine laws: The alleged laws are false or, if true, merely accidentally so. Perhaps, however, one should relax the requirement that genuine laws be exceptionless in favor of holding that a law be sufficiently accurate for the relevant purposes. The proviso to Boyle's law neglects to mention a host of petty influences that make the PV of actual gases vary somewhat. Still, Boyle's law with its proviso (which rules out the major interfering factors—the ones that scientists cannot afford to neglect) is often enough close enough to the truth for various purposes in chemistry, theoretical and practical. Fully understanding Boyle's law requires knowing the range of purposes for which it can safely be applied. Likewise, the freakish sequence of neural events mentioned earlier is too rare to make Gresham's law unreliable for the purposes of economics.

The limited range of a special science's interests influence which facts qualify as laws of that science. Consider another example: The human aorta carries all the body's oxygenated blood from the heart to the systemic circulation. This reference to "the human aorta" (a generic), rather than to all or to most human aortas, apparently indicates that one is dealing here with a policy of drawing influences that, although fallible, is sufficiently reliable for certain purposes—in this case for forming expectations about medical patients in the

absence of more specific information regarding them. In human medicine this fact about the human aorta apparently functions as a law in connection with counterfactuals, explanations, and inductions. However, this aorta fact is merely an accident of natural history; it might not have held had evolutionary history taken a different path. Still, medicine does not treat evolutionary history as a variable. A physician might say that the shooting victim would not have survived even if he or she had been brought to the hospital sooner, since the bullet punctured his or her aorta and the human aorta carries all the body's oxygenated blood from the heart to the systemic circulation. (This aorta fact would still have held had the victim been brought to the hospital sooner.) But it would not be medically relevant to point out that the victim might have survived had evolutionary history taken a different course. Accordingly, that the human aorta carries all the body's oxygenated blood to the systemic circulation may be a law of human physiology even if it is an accident of physics.

LAWS OF INEXACT SCIENCES: THE PROBLEM OF NECESSITY

But (to shift examples) even if the law that larger islands have greater biodiversity (all other things—such as their distance from the mainland—being equal) is sufficiently reliable for the purposes of island biogeography, what makes this fact an island-biogeographical law? What makes it necessary? This is the third obstacle to inexact sciences having genuine laws of their own—that is, to their being autonomous.

Recall *nomic preservation*: that the laws would still have held under any counterfactual supposition that is logically consistent with every law. There appears to be no set of truths that is closed under logical consequence and that contains accidents (except the set of all truths) where every member of the set would still have been true under every counterfactual supposition that is logically consistent with every member of the set. Accordingly, it has been suggested that a truth n is a nomic necessity exactly when n belongs to a stable set, where a set is stable exactly when it includes every logical consequence of its members, it does not contain every truth, and its members are not only true but also all preserved under as broad a range of counterfactual suppositions as they could all logically possibly be—namely, under every supposition that is logically consistent with every member. On this view necessity involves possessing maximal invariance under counterfactual perturbations. No necessity is possessed by an accident, even one (such as a car's pedal-speed function) that would still have held under many counterfactual suppositions. (The set consisting of a car's pedal-speed function, with its logical consequences, is unstable since its members would not all still have held under engine alterations with which the pedal-speed function is logically consistent.) Stability allows one to draw a sharp distinction between laws and accidents. It also gives one a way to escape the circle involved in specifying the nomic necessities as the truths that would still have held under every nomic possibility.

This conception of nomic necessity can easily be relativized to particular sciences. Perhaps the area law belongs to a set of claims that are all sufficiently reliable for the purposes of island biogeography, where the set does not contain all such claims and where its members would all still have been sufficiently reliable under any counterfactual supposition that is not only consistent with all of them being reliable but also relevant to island biogeography. In that case the set's members are collectively as resilient under counterfactual suppositions relevant to island biogeography as they collectively could be. Therefore, they possess nomic necessity for island biogeography.

On this view a special science's laws need not include every detail of the fundamental microphysical laws. For example, biological species would still have been distributed according to the laws of island biogeography (if there are any such laws) even if creatures were made of a continuous rigid substance rather than molecules, contrary to microphysical laws. Whether a given special science is autonomous remains for scientific research to discover. Whether the fundamental microphysical laws are privileged among the natural laws (e.g., in having greater generality or being strictly true) remains philosophically controversial.

See also Causation: Philosophy of Science; Laws of Nature; Metaphysics; Theories and Theoretical Terms.

Bibliography

Armstrong, David A. *What Is a Law of Nature?* New York: Cambridge University Press, 1983.

Beatty, John. "The Evolutionary Contingency Thesis." In *Concepts, Theories, and Rationality in the Biological Sciences*, edited by Gereon Wolters and James G. Lennox. Pittsburgh, PA: University of Pittsburgh Press, 1995.

Carlson, Gregory N., and Francis Jeffry Pelletier, eds. *The Generic Book*. Chicago: University of Chicago Press, 1995.

Carroll, John W. *Laws of Nature*. New York: Cambridge University Press, 1994.

Carroll, John W., ed. *Readings on Laws of Nature*. Pittsburgh: University of Pittsburgh Press, 2004.
Cartwright, Nancy. *The Dappled World: A Study of the Boundaries of Science*. New York: Cambridge University Press, 1999.
Cartwright, Nancy. *How the Laws of Physics Lie*. New York: Oxford University Press, 1983.
Dretske, Fred. "Laws of Nature." *Philosophy of Science* 44 (1977): 248–268.
Earman, John, and John Roberts. "Ceteris Paribus, There Is No Problem of Provisos." *Synthese* 118 (1999): 439–478.
Earman, John, Clark Glymour, and Sandra Mitchell. *Ceteris Paribus Laws*. Dordrecht, Netherlands: Kluwer Academic, 2002.
Fodor, Jerry A. "Special Sciences, or the Disunity of Science as a Working Hypothesis." *Synthese* 28 (2) (1974): 97–115.
Gardner, Patrick L. *The Nature of Historical Explanation*. London: Oxford University Press, 1952.
Goodman, Nelson. *Fact, Fiction, and Forecast*, 4th ed. Cambridge, MA: Harvard University Press, 1983.
Hempel, Carl G. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, 1965.
Kincaid, Harold. *Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research*. New York: Cambridge University Press, 1996.
Lange, Marc. *Natural Laws in Scientific Practice*. New York: Oxford University Press, 2000.
Lewis, David K. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.
Lewis, David K. *Philosophical Papers.* Vol. 2. New York: Oxford University Press, 1986.
Loewer, Barry. "Humean Supervenience." *Philosophical Topics* 24 (1996) 101–127.
Mill, John Stuart. *A System of Logic.* 8th ed. New York: Harper, 1893.
Scriven, Michael. "Truisms as the Ground for Historical Explanations." In *Theories of History*, edited by Patrick L. Gardner. London: George Allen and Unwin, 1959.
Sklar, Lawrence. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. New York: Cambridge University Press, 1993.
Sober, Elliott. *Philosophy of Biology.* 2nd ed. Boulder, CO: Westview Press, 2000.
Van Fraassen, Bas C. *Laws and Symmetry*. New York: Oxford University Press, 1989.
Woodward, James. "Explanation and Invariance in the Special Sciences." *British Journal for the Philosophy of Science* 51 (2) (2000): 197–254.

Marc Lange (2005)

LAWS OF NATURE

The "laws of nature" are the general ways of working of the physical and mental world. Many natural scientists have as one of their great aims the uncovering of these laws. The topic of laws of nature has been the subject of vigorous discussion in contemporary philosophy. Three broad tendencies have emerged, with a number of important variations within these tendencies.

THE REGULARITY OR HUMEAN VIEW

Since the work of David Hume, at least, there have been many philosophers, particularly those in the empiricist tradition, who have tried to analyze both causes and laws (which they tend not to distinguish very clearly) in terms of mere regular successions or other regularities in the behavior of things. Laws tell us that, given a phenomenon of a certain sort, then a further phenomenon of a certain sort must occur in a certain relation to the first phenomenon. Particularly since the rise of quantum physics, this may be modified by saying that there must be a certain probability that the further phenomenon will occur. Regularity theorists see this "must" as mere universality: This is what always happens.

A great many difficulties have been raised against this position (for a fairly full listing see Armstrong 1983, pt. 1). The most important of these are as follows.

1. The intuitive difference between merely accidental uniformities and nomic (lawlike) uniformities. The traditional example is the contrast between the accidental uniformity that every sphere of gold has a diameter of less than one mile and the nomic uniformity that every sphere of uranium 235 has a diameter of less than one mile, because that diameter would ensure "critical mass" and the explosion of the sphere.
2. Laws of nature "sustain counterfactuals." If it is a law that arsenic is poisonous, then if, contrary to the facts, you had drunk arsenic, you would have been very sick. But from the fact that no human being of Neanderthal race ever spoke English, it by no means follows that if, contrary to fact, some of them had lived in an English-speaking society, they would not have spoken English. The uniformity that Neanderthals spoke no English does not sustain counterfactuals.
3. A regularity theorist cannot give a satisfactory solution to the problem of induction. If laws are mere regularities, what rational grounds have we for believing that observed uniformities will continue to hold in the future and for the unobserved generally?
4. A regularity theorist is likely to identify merely probabilistic laws with actually occurring frequencies. This identification is difficult, because

such laws do not actually rule out distributions with the "wrong" frequencies. All that probabilistic laws do is to make such frequencies improbable; they do not make them nomically impossible.

5. Science admits certain laws that may well have no positive instances falling under them. The most famous example is Isaac Newton's first law of motion. An uninstantiated law would have to be a vacuous uniformity, but there are far too many such uniformities all to be laws.

Those who continue to work in the regularity tradition try to meet these and other difficulties largely by distinguishing "good" uniformities that deserve to be called laws and "bad" ones that do not. There are two main approaches, the epistemic and the systematic.

Epistemic theorists emphasize the nature of the evidence that we have for claiming that certain uniformities obtain. References and criticism may be found in W. A. Suchting (1974), G. Molnar (1974), Fred Dretske (1977), and David M. Armstrong (1983). Brian Skyrms's (1980) resiliency account is a sophisticated epistemic approach. His basic idea is that we give assent to a generalization, and count it lawlike, only if we find it to hold under a wide variety of circumstances and conditions. For criticism of Skyrms, see Michael Tooley (1987) and J. Carroll (1990).

The systematic approach has been championed by David Lewis, explicitly basing himself on a suggestion made by F. P. Ramsey. Lewis says that "contingent generalization is a *law of nature* if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves the best combination of simplicity and strength" (1973, p. 73). Further discussion may be found in Lewis (1986). He himself finds that his greatest difficulties are associated with probabilistic laws. For criticism of Lewis see Armstrong (1983), Tooley (1987), and Carroll (1990).

STRONG LAWS

One who judges that no regularity theory of laws can succeed may wish to argue that laws are something stronger than mere uniformities or statistical distributions. Laws may be called strong if their existence entails the existence of the corresponding uniformities and so on but the reverse entailment fails to hold.

Traditional theories of strong laws tended to see these laws as holding necessarily. Given all the antecedent conditions, the consequent is entailed. In the days when Euclidean geometry was unchallenged, geometrical models were attractive. As with geometrical theorems, this necessity was thought to be discoverable, at least potentially, a priori. Granted that laws might in practice be discovered by experience, just as the Pythagorean theorem might be discovered by measuring and adding areas, it was still thought that a sufficiently powerful intellect might spell out the necessity involved without the aid of experience. This approach seems to have been abandoned by contemporary philosophers (though there are hints in Martin 1993). It now seems agreed, in general, and in agreement with regularity theorists, that the laws of nature can be discovered only a posteriori.

Upholders of strong laws do, however, differ among themselves whether these laws are contingent or necessary. The contingency view (also held by regularity theorists) is represented by Dretske (1977), Tooley (1977, 1987) and Armstrong (1983). These three evolved their rather similar views independently and almost simultaneously. Laws are argued to be dyadic relations of necessitation holding contingently between universals, schematically N(F,G). Such a relation entails the regularity that all Fs are Gs, but the regularity does not entail N(F,G). Dretske presents the central idea with particular clarity; Tooley and Armstrong develop the theory more fully. Tooley argues that the possibility of certain sorts of uninstantiated laws demands uninstantiated universals, leading him to what he calls a factual Platonism about universals. Armstrong, however, tries to get along with instantiated universals only.

The theory appears to be able to handle probabilistic laws (see Armstrong 1983, chap. 9; Tooley 1987, chap. 4). The connection between universals envisaged by the theory may be thought of as involving connections of differing strength holding between antecedent and consequent universals. The greatest strength, one (exactly one, not one minus an infinitesimal), represents the probability involved in an old-style deterministic law. The consequent universal must be instantiated if the antecedent universal is. Numbers between nought and one give the lesser probability of the consequent being instantiated under these conditions. This probability is an objective one. The antecedent universal, if instantiated, bestows an objective propensity, as some say, for the instantiation of the consequent.

An obvious cost of this sort of theory is that it must postulate universals. This is a stumbling block to many. But by far the most important criticism of this account has been developed by Bas van Fraassen (1989, chap. 5; see also the discussion-review of this book by him, Earman, Cartwright, and Armstrong 1993). He poses two

difficulties: the identification problem and the inference problem. The first is the problem of identifying in a non-circular way the nature of the necessitation relation supposed to hold between the universals involved in the law. The second is the problem, given a concrete account of this relation, of understanding why it is legitimate to infer from the fact that the universals are so related to the existence of corresponding uniformities or frequencies in the world. Van Fraassen argues that solving the one problem makes it impossible to solve the other. A clear account of the relation makes the inference problematic; a clearly valid inference makes the relation no more than something that validates the inference.

The view that laws of nature are necessities discovered a posteriori is developed by Sydney Shoemaker (1984) and Chris Swoyer (1982). They build on Saul Kripke's (1980) arguments for a posteriori knowledge of necessity, and Kripke hints that laws of nature may have this status. Their view depends upon taking a different view of properties from that found in Dretske, Tooley, and Armstrong. For the latter, properties are conceived of as "categorical" or self-contained entities. But for Shoemaker and Swoyer, properties, either singly or in combination, are nothing apart from the laws they enter into. They might be described as pure powers or dispositions to produce law-governed consequences.

On this view, therefore, if it is a law that property F ensures possession of property G, then it is the very essence of F so to ensure G, and so the law is necessary. That there are things having property F is contingent, but that F ensures G is necessary. It seems, then, that the dispute between contingent strong laws and necessary ones depends on the true theory of properties. See Richard Swinburne's (1983) critical comments on Shoemaker.

The view of properties just discussed might be called dispositionalism as opposed to categoricalism. There are theorists who favor a view of properties that gives them both a categorical and a dispositional (or power) side; see Evan Fales (1990) and C. B. Martin (1993). It is to be noted that both in pure dispositionalism and this mixed theory there is a strong tendency to regard laws as not fundamental but rather analyzable in terms of causal relations holding between individual events and particulars (singular causation). These causal relations, and so their laws, are determined by the nature of the dispositions or powers that particulars have.

ELIMINATIVISM ABOUT LAWS

The regularity theory of laws is a deflationary theory. It holds that there is less to being a law than one might naturally think. It also faces a number of serious difficulties. One response, rather typical of our age, is to meet the difficulties, not by proposing a strengthened theory of laws, but by taking the deflation further and arguing that there are no such things as laws. This is the position taken by van Fraassen in *Laws and Symmetry* (1989). A natural comparison is with eliminative materialism, which denies the existence of the mind in favor of the brain.

Van Fraassen begins with a systematic criticism, first of Lewis's version of the regularity theory, and then of various strong views. The rejection of laws he links to his "constructive empiricism," according to which the aim of science is not truth in general but only empirical adequacy, defined as truth with respect to what is observed. Beyond the observable, all that can usefully be done is the constructing of models that are in a deep way adequate to the phenomena and that may be true but about which we can have no special reason to think them true. In these constructions considerations of symmetry play an energizing role.

A certain skepticism about laws is also to be found in Nancy Cartwright's *How the Laws of Physics Lie* (1983). Her skepticism concerns the fundamental as opposed to more messy phenomenological laws. The former may explain better, but the latter are truer to the facts! The distinction she is concerned with is one made by physicists, and the "phenomenological" laws go far beyond van Fraassen's observables. Cartwright accepts these laws because the entities they deal with, though perhaps unobserved, appear to exist and to act as causes. In *Nature's Capacities and Their Measurement* (1989), she argues that the world is a world of singular causes, individual entities interacting with each other. The nature of these interactions is determined by the capacities of these entities. Her capacities seem close to the dispositions and powers that contemporary necessitarians identify with properties.

See also Armstrong, David M.; Cartwright, Nancy; Descartes, René; Dretske, Fred; Earman, John; Eliminative Materialism, Eliminativism; Empiricism; Geometry; Hobbes, Thomas; Hume, David; Induction; Kripke, Saul; Lewis, David; Locke, John; Newton, Isaac; Probability and Chance; Ramsey, Frank Plumpton; Shoemaker, Sydney; Van Fraassen, Bas.

Bibliography

Armstrong, D. M. *What Is a Law of Nature?* Cambridge, U.K.: Cambridge University Press, 1983.

Carroll, J. "The Humean Tradition." *Philosophical Review* 99 (1990): 261–276.

Carroll, J. *Laws of Nature.* Cambridge, U.K.: Cambridge University Press, 1994. Contains a useful set of references.
Cartwright, N. *How the Laws of Physics Lie.* Oxford: Clarendon Press, 1983.
Cartwright, N. *Nature's Capacities and Their Measurement.* Oxford Clarendon Press, 1989.
Dretske, F. "Laws of Nature." *Philosophy of Science* 44 (1977): 248–268.
Fales, E. *Causation and Universals.* London: Routledge, 1990.
Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.
Lewis, D. *Counterfactuals.* Cambridge, MA: Harvard University Press, 1973.
Lewis, D. *Philosophical Papers,* Vol. 2. New York: Oxford University Press, 1986.
Martin, C. B. "Power for Realists." In *Ontology, Causality, and Mind,* edited by J. Bacon, K. Campbell, and L. Reinhardt. Cambridge, U.K.: Cambridge University Press, 1993.
Molnar, G. "Kneale's Argument Revisited." In *Philosophical Problems of Causation,* edited by T. Beauchamp. Encino, CA: Dickenson, 1974.
Shoemaker, S. "Causality and Properties." In *Identity, Cause, and Mind.* Cambridge, U.K.: Cambridge University Press, 1984.
Skyrms, B. *Causal Necessity.* New Haven, CT: Yale University Press, 1980.
Suchting, W. A. "Regularity and Law." In *Boston Studies in the Philosophy of Science,* edited by R. S. Cohen and M. W. Wartofsky. Dordrecht: Reidel, 1974.
Swinburne, R. "Reply to Shoemaker." In *Aspects of Inductive Logic,* edited by L. J. Cohen and M. Hesse. Oxford, 1983. See also Shoemaker's paper in the same volume.
Swoyer, C. "The Nature of Natural Laws." *Australasian Journal of Philosophy* 60 (1982): 203–223.
Tooley, M. *Causation: A Realist Approach.* Oxford: Clarendon Press, 1987.
Tooley, M. "The Nature of Laws." *Canadian Journal of Philosophy* 7 (1977): 667–698.
van Fraassen, B. *Laws and Symmetry.* Oxford: Oxford University Press, 1989.
van Fraassen, B., J. Earman, D. M. Armstrong, and N. Cartwright. "A Book Symposium on Laws and Symmetry." *Philosophy and Phenomenological Research* 53 (1993): 411–444.

David M. Armstrong (1996)

LAWS OF NATURE [ADDENDUM]

Since David M. Armstrong's entry was first published in 1996, the philosophical debates he identifies have evolved in minor ways. There is still the central debate between the Humeans (the regularity theorists) and the anti-Humeans (the proponents of strong laws), and there are still those who choose to deny that there are any laws, to be eliminativists, rather than engage in the central debate. This addendum indicates how the literature has shifted focus to questions surrounding supervenience and whether laws govern. It also engages in a fuller discussion of the relationship between laws of nature and epistemological issues, including the role of laws in inductive inference and some skeptical challenges for both Humeans and anti-Humeans.

SUPERVENIENCE

The persevering Humean theory of laws is the systematic approach made popular by David Lewis and described briefly by Armstrong. (Versions of this account are also defended in Earman [1984] and in Loewer [1996].) A feature of this view prized by its supporters is its consistency with Humean supervenience, a thesis formulated various ways (see Earman and Roberts, 2005, Part I) but that basically maintains that the most fundamental nonmodal features of a universe fix everything else about it, including what its laws of nature are. Some anti-Humeans, the ones who think that laws are metaphysical necessities (e.g., Bigelow, Ellis, and Lierse 1992) accept that the laws do supervene for the mere reason that metaphysical necessities hold no matter what. But most anti-Humeans, the ones who believe that it is a contingent matter of fact whether something is a law, think that what the laws are does not supervene.

Michael Tooley (1977) asks one to suppose that there are exactly ten different kinds of fundamental particles. So, there are fifty-five possible kinds of two-particle interactions. Suppose also that fifty-four of these kinds of interactions have been studied and fifty-four laws have been proposed and thoroughly tested. It just so happens that there are no interactions between the last two kinds of particles ever. These final two kinds of particles are arbitrarily labeled as X and Y particles. What is interesting about this example is that it seems that many different X-Y interaction laws are consistent with all the events that take place: There might be a law that, when X and Y particles interact, the particles are destroyed; but instead there might be a law that, when X and Y particles interact, the particles bond.

Tooley's example presents a problem for Humeans. Consider what a certain, simple, regularity account would say about the seeming possibility that it is a law in the ten-particle world that when X and Y particles interact they bond. The account holds that P is a law of nature if and only if P is a true, contingent, universal generalization. It implies the absurdity that both of the X-Y regularities mentioned earlier are laws. This is absurd because such annihilation and bonding events are incompatible. The

account fails to pick out the bonding regularity as the law because the concepts invoked do not differentiate between bonding and annihilation. Because the two X-Y regularities are both true, contingent, universal generalizations, they both get counted as laws. Prima facie, Lewis's more sophisticated account is faced with the same problem. Not only do the two key X-Y regularities not differ regarding their truth, their contingency, or their logical form, they also do not differ regarding their simplicity or their strength. So, arguably, either these two regularities would both belong to all the true deductive systems with a best combination of simplicity and strength or else neither of them would.

Eliminativists may acknowledge the intuitiveness of Tooley's example but may see that as reason to think there are no laws. There is also a certain nonstandard way of being Humean that sidesteps the issue in something of a similar fashion. Barry Ward, in "Humeanism without Humean Supervenience: A Projectivist Account of Laws and Possibilities" (2002), maintains that lawhood sentences are not fact-stating, that they serve a different role, one of projecting some noncognitive attitude.

DO LAWS GOVERN?

The standard approach for Humeans is to somehow deny that Tooley's example and other cases to the same effect (see Carroll 1994, pp. 57–85) are genuinely possible. They hold that the intuition that the lawfulness of each of the two X-Y interaction principles is consistent with the events of the ten-particle world is somehow misleading or ill founded. Helen Beebee, in "The Non-governing Conception of Laws of Nature" (2000), suggests that the source of the intuition is a certain veneration of a conception of laws that holds that laws govern the course of history (also see Loewer 1996, pp. 115–117). The idea is that, if one comes to the debate with the governing conception in mind, one is likely to find nonsupervenience examples convincing, but using this conception to reject Humean analyses of lawhood is to beg the question because it is a conception Humeans reject. Having their own conception of laws not as governing but as summarizing, Humeans insist that at most one of the two X-Y interaction laws is consistent with the events of the ten-particle world. Anti-Humeans are sometimes accused of relying on a nonscientific or even theistic or legalistic conception of what it is to be a law, though the anti-Humeans themselves will insist, on the contrary, that it is their conception that is the scientific conception and that it needs no theistic or legalistic underpinning. Nevertheless, one idiosyncratic anti-Humean, John Foster, in *The Divine Lawmaker: Lectures on Induction, Laws of Nature, and the Existence of God* (2004), provides fuel for the Humean fire by arguing that lawful regularities ultimately must be explained in terms of the agency of God.

LAWS AND INDUCTION

In "The New Riddle of Induction" (1983) Nelson Goodman argues that the difference between laws of nature and accidental uniformities is linked with the problem of induction via the concept of lawlikeness. Lawlikeness is whatever additional characteristic a universal generalization needs, aside from truth, to be a law. Goodman claims that, if a true generalization is accidentally true (and so not lawlike), then an instance of the generalization does not confirm the generalization.

There are examples that threaten Goodman's contention. Suppose a brand new die will be thrown twice and then destroyed. Also suppose that one is interested in whether it will come up six both times it is tossed. It is thrown the first time and it does land six. Notice that this single instance has increased dramatically the probability that all tosses of this die will land six. Before the first toss, that probability was one out of thirty-six. Now that the first toss has landed six, that probability has gone up to one-sixth. So, apparently, Goodman's claim is mistaken; observation of one instance of the generalization that every toss of this die will land six has provided confirmation even though the generalization is not lawlike.

One natural response to this kind of example maintains that probability raising is not the notion of confirmation that Goodman had in mind. The temptation is to hold instead that a generalization is lawlike if observed instances confirm that the generalization also holds for unexamined cases. (Notice that, in the die example, the first roll landing six does not raise the probability that the second roll will land six.) But this alternative is not right either. Maybe you know that Sam sorts his coins by putting nickels in one pocket, dimes in another, and so on; he is fanatical that way. You do not know, however, which pocket he keeps his nickels in. Sam shows you one of the coins from his left-front pocket and you see that it is a nickel. Evidently this instance of the generalization that all the coins in Sam's left-front pocket are nickels has confirmed that all the coins in this pocket are nickels. Given your background knowledge, you seem perfectly justified in believing that the generalization is true even though it is not lawlike. You also have reason to believe that the other coins in his left-front pocket, the ones you have not examined, are nickels.

Followers of Goodman have their work cut out for them. The biggest obstacle involves background beliefs, as is evidenced by the two examples just given where prior knowledge that the die would only be tossed twice and prior knowledge of Sam's fanatical coin sorting played a major role. Here is the basic problem: The confirmation of a generalization or its unexamined instances is sensitive to the background assumptions in place. So much so, that with assumptions of the right sort, just about anything can be confirmed irrespective of its lawlikeness. Elliot Sober's "Confirmation and Law-likeness" (1988) presents a series of difficult cases and concludes by expressing this concern about background assumptions. Marc Lange's *Natural Laws in Scientific Practice* (2000, pp. 111–142) takes up the challenge, in part, by further refining the relevant notion of confirmation, characterizing what he takes to be an intuitive notion of inductive confirmation, and arguing that only those generalizations that are not believed not to be lawlike can be inductively confirmed.

A SKEPTICAL CHALLENGE FOR HUMEANS

Sometimes the idea that laws have a special role to play in induction serves as the starting point for an anti-Humean criticism of Humean analyses. Fred I. Dretske (1977, pp. 261–262) and Armstrong (1983, pp. 52–59) adopt a view according to which induction involves an inference to the best explanation (also see Foster 2004). On its simplest construal, the view describes a pattern that begins with an observation of instances of a generalization, includes an inference to the corresponding law (this is the inference to the best explanation), and concludes with an inference to the regularity itself or to some conclusion about its unexamined instances. The complaint lodged against Humeans is that, on their view of what laws are, laws are not suited to explain their instances and so cannot sustain the required inference to the best explanation. After all, if laws are summaries of their instances, then they cannot explain their instances. Does the fact that all Fs are Gs explain why this F is a G? It is hard to see how it could; that this F is G is part of what makes it true that all Fs are Gs.

A SKEPTICAL CHALLENGE FOR ANTI-HUMEANS

Sometimes very different skeptical considerations are used by Humeans against the anti-Humeans. Prompted by examples like Tooley's, it can seem that the events of the actual world fail to determine what the laws are. Scientists and philosophers believe it is a law that all signals have speeds less than the speed of light, but if the course of our actual history is consistent with this generalization not being a law, with it being a remarkable coincidence, then can anyone really know that it is a law of nature?

John Earman and John Roberts formalize this kind of reasoning in "Contact with the Nomic: A Challenge for Deniers of Humean Supervenience about Laws of Nature (Part II): The Epistemological Argument for Humean Supervenience" (2005). Let T be a theory that posits at least one law. Label one of the laws L and reformulate the theory as the conjunction that L is a law of nature and X. (So, X is the rest of the theory aside from the part that posits L as a law.) Let T* be the theory that L is true, not a law of nature, and X. So, T and T* cannot both be true; they differ on whether L is a law, though they agree on L's truth. Then, the argument is straightforward:

(1) If Humean supervenience is false, then no empirical evidence can favor T or T* over the other

(2) If no empirical evidence can favor T or T* over the other, then one cannot be epistemically justified in believing that T is true

(3) If Humean supervenience is false, one cannot be epistemically justified in believing that T is true

Earman and Roberts make no assumptions about T other than that it takes something to be a law. So, if Earman and Roberts's argument is sound and Humean supervenience is false, then no one is justified in believing that any proposition is a law. It is only a short step from there to the conclusion that no one knows what any of the laws are. As Earman and Roberts deny skepticism about laws, they ultimately see this as an argument for Humean supervenience.

In response, the anti-Humean will appeal to the history of philosophy, citing instances of epistemological questions that have befuddled metaphysical issues. George Berkeley's idealism is a dramatic example. Faced with René Descartes's skeptical investigations, Berkeley advanced the untenable metaphysical position according to which material objects are nothing more than collections of ideas. So, Humeans have to make it clear that they are not making a mistake parallel to the one made by Berkeley. Also, it is important to remember that, in some sense, it could be that nobody knows what the laws are; it is obvious that there are a great many ways that our world could be such that scientists will not discover every fact, and so our world could be such that no one will discover any of the laws.

Nevertheless, Humeans and anti-Humeans alike are right to balk at a conclusion to the effect that we can never know what the laws are. If a philosophical position on lawhood, be it the identification of laws with regularities or the denial of Humean supervenience or something else, has the consequence that no one ever knows anything to be a law, then the defender of that position should be concerned. Part of the motivation for engaging in a philosophical investigation of laws is the seeming truism that science has as one of its aims to uncover the laws. If it should turn out that science is bound not to result in knowledge of the laws, then philosophy would have led us to at least a disturbing take on the nature of the world.

See also Armstrong, David M.; Berkeley, George; Descartes, René; Dretske, Fred; Earman, John; Goodman, Nelson; Hume, David; Induction; Lewis, David; Natural Law; Supervenience.

Bibliography

Armstrong, David M. *What Is a Law of Nature?*. New York: Cambridge University Press, 1983.

Beebee, Helen. "The Non-governing Conception of Laws of Nature." *Philosophy and Phenomenological Research* 61 (2000): 571–594.

Bigelow, John, Brian Ellis, and Caroline Lierse. "The World as One of a Kind: Natural Necessity and Laws of Nature." *British Journal for the Philosophy of Science* 43 (1992): 371–388.

Carroll, John W. *Laws of Nature*. New York: Cambridge University Press, 1994.

Carroll, John W., ed. *Readings on Laws of Nature*. Pittsburgh, PA: Pittsburgh University Press, 2004.

Dretske, Fred I. "Laws of Nature." *Philosophy of Science* 44 (1977): 248–268.

Earman, John. "Laws of Nature: The Empiricist Challenge." In *D. M. Armstrong*, edited by Radu J. Bogdan. Dordrecht, Netherlands: D. Reidel, 1984.

Earman, John, and John Roberts. "Contact with the Nomic: A Challenge for Deniers of Humean Supervenience about Laws of Nature (Part I) Humean Supervenience." *Philosophy and Phenomenological Research* 71 (2005).

Earman, John, and John Roberts. "Contact with the Nomic: A Challenge for Deniers of Humean Supervenience about Laws of Nature (Part II): The Epistemological Argument for Humean Supervenience." *Philosophy and Phenomenological Research* 71 (2005).

Foster, John. *The Divine Lawmaker: Lectures on Induction, Laws of Nature, and the Existence of God*. Oxford, U.K.: Clarendon Press, 2004.

Goodman, Nelson. "The New Riddle of Induction." In *Fact, Fiction, and Forecast*. 4th ed., 59–83. Cambridge, MA: Harvard University Press, 1983.

Lange, Marc. *Natural Laws in Scientific Practice*. New York: Oxford University Press, 2000.

Loewer, Barry. "Humean Supervenience." *Philosophical Topics* 24 (1996): 101–127.

Sober, Elliott. "Confirmation and Law-likeness." *Philosophical Review* 97 (1988): 93–98.

Tooley, Michael. "The Nature of Laws." *Canadian Journal of Philosophy* 7 (4) (1977): 667–698.

Ward, Barry. "Humeanism without Humean Supervenience: A Projectivist Account of Laws and Possibilities." *Philosophical Studies* 107 (2002): 191–218.

John W. Carroll (2005)

LAWS OF THOUGHT

The term "laws of thought" traditionally covered the principles of identity, of contradiction, of excluded middle, and occasionally the principle of sufficient reason. Whereas these principles were frequently discussed from the time of the Greeks until the beginning of the twentieth century, the term has become obsolete, for at least two good reasons. One is the great and confusing variety of meanings with which it has been used, the other is the now generally acknowledged fact that no viable system of logic can be constructed in which the principles of identity, contradiction, and excluded middle would be the only axioms. Typical discussions of these principles are to be found, for example, in Friedrich Ueberweg's *System der Logik* and in H. W. B. Joseph's *Introduction to Logic*. In the following discussion the principle of sufficient reason, which, unlike the others, cannot be interpreted as a principle of formal logic, will not be dealt with.

The three laws of thought have in the main been conceived of as descriptive, prescriptive, or formal. As *descriptive* laws, they have been regarded as descriptive (*a*) of the nature of "being as such," (*b*) of the subject matter common to all sciences, or (*c*) of the activity of thinking or reasoning. As *prescriptive* laws, they have been conceived of as expressing absolute or conventional standards of correct thinking or reasoning. As *formal* laws, they have been held to be propositions which are true in virtue of their form and independently of their content, true *in* all possible worlds, or true *of* any objects whatsoever, whether these objects exist or not. Distinctions between these conceptions are often blurred, since they depend on implicit and often unclear assumptions about the relations between factual, normative, and metaphysical propositions: It is, for example, rarely investigated either to what extent various kinds of rules depend for their satisfiability on what is the case or to what extent logic is or can be free from metaphysical presuppositions or implications.

All these very different conceptions of the laws of thought are compatible with their traditional formulations, which lack the precision now achievable by means of the axiomatization and formalization of theories. Examples of typical, traditional formulations are: For the law of identity, *A* is *A*; everything is what it is; every subject is its own predicate. For the law of contradiction, *A* is not not-*A*; judgments contradictorily opposed to each other cannot both be true. For the law of excluded middle, everything is either *A* or not-*A*; judgments opposed as contradictories cannot both be false, nor can they admit the truth of a third or middle judgment, but one or the other must be true, and the truth of the one follows from the falsehood of the other. An obvious ambiguity concerning the law of identity is connected with the question whether *is* is to be taken as expressing equality or as the copula between subject and predicate, and, in the latter case, whether or not it implies the existence of the subject. Again, the term *not* admits of different interpretations according to different metaphysical and logical assumptions about negation.

DESCRIPTIVE INTERPRETATIONS

METAPHYSICAL INTERPRETATION. For Aristotle, who discussed the laws of thought in his logical and metaphysical works, they are primarily descriptive of being as such and only secondarily standards of correct thinking. It is thus a metaphysical or ontological impossibility that "the same can and cannot belong to the same in the same reference" (*Metaphysics* III, 2, 2), from which it follows as a rule of correct thought and speech that it is incorrect to assert that "the same is and is not" (*Metaphysics* IV, 6, 12). Aristotle produced seven "proofs" to demonstrate the indispensability of the law of contradiction. With a similar intention, formal logicians are nowadays wont to show that its negation implies any proposition whatever (and thus also the law of contradiction itself) by some such reasoning as the following: (1) To assume that the law of contradiction is false is to assume for some proposition *p* that *p* and not-*p* are both true. (2) From the truth of *p* it follows that "*p* or *x*" is also true, where *x* is an arbitrary proposition and "or" is used in the nonexclusive sense of "and/or." (3) From the truth of "*p* or *x*" and the truth of not-*p* the truth of *x* follows. But *x* is an arbitrary proposition for which, for example, the law of contradiction may be chosen.

Aristotle's defense of the law of contradiction as descriptive of "being as such" includes implicitly a defense of the metaphysical principle of identity against Heraclitus, who held it possible for the same thing to be and not to be and who explained the concept of becoming as implying the falsehood of the principle that everything is what it is. Before Aristotle this metaphysical principle had been defended by Parmenides.

Aristotle's arguments for the truth of the principle of excluded middle are again metaphysical. They are connected with his rejection of the Platonic doctrine that attempts to mediate between Heraclitus and Parmenides. The changing sensible and material objects, which in Plato's phrase "tumble about between being and nonbeing," are placed by Plato between the eternal Forms, which fully and truly exist, and that which does not exist at all, that is, they are "a third" between being and nonbeing. The metaphysical principle of excluded middle, as understood by Aristotle, excludes any such third. This principle has sometimes been taken to imply fatalism: Since of any two contradictory statements one must be true, of any two contradictory statements about the future one must be true, so that, it is argued, the future is wholly determined. In a famous passage about "the sea fight tomorrow" Aristotle refutes this argument: It is, he points out, necessary that the sea fight will or will not take place tomorrow. But it is not true that it will *necessarily* take place tomorrow or *necessarily* not take place tomorrow. Indeed the logical necessity of a disjunction "*p* or not-*p*" does not imply that either *p* or not-*p* is a necessary proposition.

METAPHYSICAL REFUTATION. Heraclitus, Parmenides, Plato, and Aristotle conceived of the laws of thought as controversial metaphysical principles, and just as Aristotle attempted their justification on metaphysical grounds, so did G. W. F. Hegel, Karl Marx, and Friedrich Engels attempt their refutation on metaphysical grounds. Hegel's attack was based on his distinction between abstract understanding, which petrifies and thus misdescribes the ever-changing "dialectical" process that is reality, and reason, which apprehends its true nature. Hegel objected to the principle that *A* is *A* or, what for him amounts to the same thing, that *A* cannot at the same time be *A* and not-*A* because "no mind thinks or forms conceptions or speaks in accordance with this law, and … no existence of any kind whatever conforms to it" (*Die Encyclopädie der philosophischen Wissenschaften*). For Hegel contradiction is not a relation that holds merely between propositions but one which is also exemplified in the real world, for example, in such phenomena as the polarity of magnetism, the antithesis between organic and inorganic matter, and even the complementarity of complementary colors. With such an interpretation it becomes possible for him to assert that "contradiction is the very moving principle

of the world" and that "it is ridiculous to say that contradiction is unthinkable." Aristotle's metaphysics corresponds to a logic in which the metaphysical principles of identity, contradiction, and excluded middle have their logical counterpart in corresponding laws of reasoning. The counterpart of Hegel's rejection of these metaphysical principles is not any traditional logical theory but a "dialectical" logic, or dialectics.

The Hegelian point of view was adopted by Marx and Engels with the difference that they conceived reality not as ideal but as material. Engels, unlike Hegel, did not even acknowledge the law of identity as valid for the abstractions of mathematics. His arguments, based on the alleged structure of the differential and integral calculus, seem—at least today—confused. He held, for example, that under certain circumstances straight lines and curves are literally identical.

EMPIRICAL INTERPRETATION. From the conception of the laws of thought as descriptive of "being as such," whatever this may mean precisely, we must distinguish the conception of them as empirical generalizations of very high order. This view was most clearly expressed by John Stuart Mill in his *System of Logic* (London, 1843). Thus, he regarded the principle of contradiction as one "of our first empirical generalizations from experience" and as "originally founded on our distinction between belief and disbelief as two different mutually exclusive states" (*System of Logic*, Book II, Ch. 7). He similarly argued that the empirical character of the law of excluded middle follows from, among other things, the fact that it requires for its truth a large qualification, namely "that the predicate in any affirmative categorical proposition must be capable of being meaningfully attributed to the subject, since between the true and the false there is the third possibility of the meaningless" (Book II, ch. 7). Mill's view must not be taken to imply that the laws of thought are psychological laws, describing the processes of thought—a view which rests on a confusion between thinking and correct thinking.

PRESCRIPTIVE INTERPRETATIONS

REGULATIVE INTERPRETATION. Another interpretation of the laws of thought regards them as in some sense prescriptive—based on some absolute authority, by analogy with moral laws, or based on conventions admitting of possible alternatives, by analogy with municipal laws. Traces of the former view are, for example, still found in J. N. Keynes's *Formal Logic*, one of the last valuable treatises on traditional formal logic. According to the preface of this work, logic deals with the laws regulating the processes of formal reasoning purely as "regulative and authoritative" and as affording criteria for the discrimination between valid and invalid reasoning.

CONVENTIONALIST INTERPRETATION. Versions of the conception that all logical principles are based on conventions have rarely been worked out with sufficient care. According to A. J. Ayer's *Language, Truth and Logic* (1936; 2nd ed., 1946) every logical principle is based on conventions. Thus "not (*p* and not-*p*)" is logically necessary because the use of "and" and "not" is governed by certain linguistic conventions, which are neither true nor false. Yet given these conventions the proposition "not (*p* and not-*p*)," that is, the law of contradiction, is necessarily true. Ayer and those who have held similar views never consider the question whether, and to what extent, linguistic conventions depend on some nonconventional framework which restricts one's freedom to formulate, accept, or reject them. Can one, for example, by adopting suitable conventions for the use of "or" and "not" really think or speak in contravention of the principle that under the usual conventions is expressed by "not (*p* and not-*p*)"?

Conventionalism is most plausible when it explains the necessity of alternative systems of definitions and of alternative systems of logic as being based on conventions, in the sense of rules whose acceptance is not obligatory. In the case of the law of contradiction no alternative is conceivable, so that the "convention" on which it is based would have to be obligatory in a sense in which the other conventions are not. However, an admission of "conventions obligatory for all thinkers" would bring conventionalism much nearer to views of logic which, at least prima facie, it seems to reject.

FORMAL INTERPRETATIONS

LEIBNIZ AND KANT. According to Gottfried Wilhelm Leibniz there are two kinds of truths, truths of fact and truths of reason, truths of reason being true in all possible worlds and therefore descriptive of facts in such a way that not even God can change them. Leibniz regarded as a necessary and sufficient condition for a truth's being a truth of reason, and thus logically necessary, that its analysis should reveal it to depend wholly on propositions whose negation involves a contradiction, that is, on identical propositions (see, for example, *Monadology*, Secs. 31–35). He even held, in the second letter to Samuel Clarke, that the law of contradiction is "by itself sufficient" for the demonstration of "the whole of arithmetic

and geometry." Although the thesis that all logical, as well as all mathematical, truths are demonstrable by means of the law of contradiction alone is, from the point of view of contemporary knowledge, mistaken or at least obscurely expressed, the characterization of logical truths as true in all possible worlds is still the root of the Bolzano-Tarski definition of logical validity.

Although Immanuel Kant opposed the Leibnizian doctrine that the truths of mathematics are logical truths, he adhered to the principle of contradiction as the supreme principle of all logical truths or, more precisely, as the "general and wholly sufficient principle of all analytical knowledge." Since the truth of such knowledge in no way depends on whether or not the objects which are referred to exist, the principle of contradiction is a necessary but not a sufficient condition of factual knowledge. What is true of possible objects must be true of all actual ones—what is true in all possible worlds must be true of the actual one. But since the converse statement is false, the principles of formal logic cannot be an "organon" of any particular science, that is, a means for attaining knowledge of its subject matter. (See *Kritik der reinen Vernunft*, 2nd ed., introduction to Part II of "Transcendentale Elementarlehre.")

CONTEMPORARY LOGIC. In contemporary logical theory the conception of "true in every possible world" or "true of any objects whatever" has been sharpened into the conception of valid statement forms and valid statements which are well formed in accordance with the precisely formulated syntactical rules of elementary logic—propositional calculus, quantification theory, and theory of identity. A distinction is made between the logical particles, or constants, on the one hand and nonlogical constants and variables on the other. The logical constants are (1) "$\neg$," "$\vee$," "$\wedge$," and other connectives, whose intended interpretations are, respectively, "not," "or," "and," and so on, conceived as connecting true or false propositions so as to form other true or false propositions in such a way that the truth or falsehood of any compound statement depends only on the truth or falsehood of the component statements; (2) the quantifiers "$\forall$" and "$\exists$," the intended interpretation of which is such that "$\forall x\ Px$" and "$\exists x\ Px$" mean, respectively, that for a well-demarcated domain of individuals, which may be finite or infinite, every element x has the predicate P, and that there exists an individual x which possesses P (in addition to such monadic predicates as Px, such dyadic predicates as Pxy and polyadic predicates are also admitted, so that, for example, $\forall x\ \exists y\ Pxy$, $\forall x\ \forall y\ Pxy$, and so on, are also admitted); and (3) the sign "=" with the intended interpretation as identity of individuals.

The nonlogical constants are (*a*) names of specific individuals, such as "Socrates" or, indeterminately, "x_0," (*b*) names of specific predicates, such as "green" or, indeterminately, "P_0" where two predicate names which are truly asserted of the same individuals of the given domain are regarded as naming the same predicate, (*c*) names of specific statements, such as "Socrates is mortal" or, indeterminately, "p_0." The variables are individual variables such as "x," predicate variables such as "P," and statement variables such as "p." Variables are either free (or, more precisely, free for substitution by names of corresponding constants) or bound by a quantifier so that, for example, "Px" contains a free individual variable and a free predicate variables, whereas $\forall x\ Px$ contains only a free predicate variable.

A well-formed formula of elementary logic that contains free variables is a statement form. A statement form is valid if—with the intended interpretation of the logical constants—every substitution instance of it is valid in every nonempty domain, provided that every individual, predicate, and statement variable is replaced by the same individual, predicate, and statement constant wherever it occurs in the statement form. Clearly the laws of thought are valid statement forms in, for example, the following formulations: Principle of identity: $x = x$. Principle of contradiction: $\neg(p \wedge \neg p)$, $\forall x \neg (Px \wedge \neg Px)$. Principle of excluded middle: $p \vee \neg p$, $(\forall x\ Px) \vee (\exists x \neg Px)$. It is equally clear that many other well-formed formulas such as $\neg\neg p \vee \neg p$ are valid. Valid statement forms that contain only statement variables have been called tautologies by Wittgenstein.

The great precision and clarity given to the conception of the laws of thought as principles of formal logic has, however, not lifted them out of the range of philosophical controversy. Thus, intuitionist philosophers of mathematics argue that the principle of excluded middle is valid only for finite domains and that the extension of its validity to the nonfinite domain of arithmetic is based on the mistaken notion of an actually infinite domain of natural numbers, a notion that unjustifiably assimilates the number sequence to a finite class of objects. Similarly, they deny the validity of other classically valid statement forms, such as $\neg \neg p \rightarrow \neg p$.

The results of modern mathematical logic have deprived the laws of thought of their privileged status as the supreme principles of all logical truths. But since these results do not imply that there is only one true logic, the choice between classical elementary logic, intuitionist

logic, and perhaps some other logical theories still depends, at least at the present time, on extralogical, philosophical arguments.

See also Bolzano, Bernard; Determinism, a Historical Survey; Mathematics, Foundations of; Semantics.

Bibliography

Alston, William P. "The Role of Reason in the Regulation of Belief." In *Rationality in the Calvinian Tradition*, edited by N. Wolterstorff et al. Lanham, MD: University Press of America, 1983.

Aristotle. *Metaphysics*.

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1936; 2nd ed., 1946.

Engels, Friedrich. *Herrn Eugen Dührings Umwälzung der Wissenschaft*. Leipzig, 1878. Translated by E. Burns as *Herr Eugen Dühring's Revolution in Science*. London, 1934.

Georg, Rolf. "Bolzano and the Problem of Psychologism." In *Husserl's Logical Investigations Reconsidered*, edited by Denis Fisette. Dordrecht: Kluwer, 2003.

Goddard, L. "Laws of Thought." *Australasian Journal of Philosophy* 37 (1959): 28–40.

Hacker, P. M. S. "Frege and the Later Wittgenstein." In *German Philosophy since Kant*, edited by Anthony O'Hear. New York: Cambridge University Press, 1999.

Hahn, Hans. *Logik, Mathematik und Naturerkennen*. Vienna, 1933. Partially translated by Arthur Pap as "Logic, Mathematics and Knowledge of Nature" in *Logical Positivism*, edited by A. J. Ayer, 147–161. Glencoe, IL: Free Press, 1959.

Halperin, Thoedore. "Boole's Abandoned Propositional Logic." *History and Philosophy of Logic* 5 (1984): 39–48.

Hegel, G. W. F. *Die Encyclopädie der philosophischen Wissenschaften*. Heidelberg, 1817. Partially translated by W. Wallace as *The Logic of Hegel*, Oxford: Clarendon Press, 1874, and *Hegel's Philosophy of Mind*, Oxford: Clarendon Press, 1894.

Husserl, Edmund. *Logische Untersuchungen.* 2 vols. Halle, 1900–1901. Chs. 4 and 5.

Jeffrey, Richard. *The Logic of Decision*. New York: McGraw-Hill, 1965.

Joseph, H. W. B. *Introduction to Logic*. Oxford: Clarendon Press, 1906; 2nd ed., 1925.

Kant, Immanuel. *Kritik der reinen Vernunft.* 2nd ed. Riga, 1787.

Katz, Jerrold. "Analyticity, Necessity, and the Epistemology of Semantics." *Philosophy and Phenomenological Research* 56 (1997): 1–28.

Keynes, J. N. *Studies and Exercises in Formal Logic*. London and New York, 1884.

Kornblith, Hilary. "The Laws of Thought." *Philosophy and Phenomenological Research* 52 (1992): 895–911.

Leibniz, G. W. *Monadology*, in *The Monadology and Other Writings*. Translated by Robert Latta. Oxford, 1898.

Musgrave, A. "George Boole and Psychologism." *Sciencia* 107 (1972): 593–608.

Nagel, Ernest. "Logic without Ontology." In *Naturalism and the Human Spirit*, edited by Y. H. Krikorian. New York: Columbia University Press, 1944. Reprinted in Nagel's *Logic without Metaphysics*. Glencoe, IL: Free Press, 1956.

Russell, Bertrand. *The Problems of Philosophy*. London: Williams and Norgate, 1912.

Slater, B. H. "Wittgenstein's Later Logic." *Philosophy* 54 (1979): 199–209.

Sober, Elliot. "Psychologism." *Journal for the Theory of Social Behavior* 8 (1978): 165–191.

Ueberweg, Friedrich. *System der Logik*. Bonn, 1857. Translated by T. M. Lindsay as *System of Logic*. London: Longmans Green, 1871.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. London: Routledge, 1922.

Woods, John, and Douglas Walton. "Laws of Thought and Epistemic Proofs." *Idealistic Studies* 9 (1979): 55–65.

S. Körner (1967)
Bibliography updated by Benjamin Fiedor (2005)

LEBENPHILOSOPHIE

See *Philosophical Anthropology*

LE CLERC, JEAN
(1657–1736)

Although Jean Le Clerc, the philosopher and Arminian theologian, was not a major figure, he had a considerable influence on eighteenth-century French philosophy. He championed rational religion, which was later widely accepted, and was also the first disciple of John Locke, whose work he introduced to Continental audiences. Through his learned reviews, the *Bibliothèque universelle et historique* (1686–1693), the *Bibliothèque choisie* (1703–1713), and the *Bibliothèque ancienne et moderne* (1714–1727), he stated and defended Locke's views.

Raised in Geneva during a period of strife over the Calvinist dogma of predestination, Le Clerc was a confirmed rationalist when he left the Geneva Academy. He believed that the fundamentals of Christianity (God's existence and the divinity of Scripture) are capable of demonstration. Scripture must be rationally interpreted; one cannot believe what conflicts with rational truths, and doctrines over which rational men disagree are not essentials of faith. For this doctrine, Le Clerc was expelled from Geneva in 1683.

He went first to England and then settled permanently in Holland, a haven for political and religious exiles. He found a spiritual home in the rationalistic Remonstrant Church and soon became professor of Hebrew, philosophy, and belles-lettres at the Remon-

strant College at Amsterdam. At this time Le Clerc met John Locke, then in exile, and acquired a systematic philosophy. In 1688, two years before the English publication of the *Essay concerning Human Understanding,* he printed a long French summary in his *Bibliothèque universelle.* He also helped popularize many other English writers and published a long review of George Berkeley's *New Theory of Vision* in the *Bibliothèque choisie* (1711).

Le Clerc's philosophy was purely Lockean. He rejected innate ideas, used the notion of abstract ideas, and continued the critique of the idea of substance. He opposed René Descartes, Nicolas Malebranche, Benedict de Spinoza, and Gottfried Wilhelm Leibniz because their theories claim knowledge beyond human ideas. However, whereas Locke was indifferent to the rise of radical skepticism, Le Clerc was quite critical of it. He vigorously asserted the reality of human knowledge, although restricting its scope, and tried to refute each of the leading skeptics (Sextus Empiricus, Michel Eyquem de Montaigne, Pierre-Daniel Huet, and Pierre Bayle). He became involved in an acrimonious dispute with Bayle, who argued that the conflict between fundamental Christian doctrines and the principles of reason must be considered as a basis for skepticism regarding reason. To Le Clerc, Bayle's view led to irreligion or fanaticism. He insisted that reason is the criterion of truth and that faith and reason are compatible.

See also Arminius and Arminianism; Bayle, Pierre; Berkeley, George; Descartes, René; Determinism, A Historical Survey; Huet, Pierre-Daniel; Innate Ideas; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Montaigne, Michel Eyquem de; Rationalism; Sextus Empiricus; Skepticism, History of; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY LE CLERC

Logica, sive Ars Ratiocinandi. Amsterdam, 1694.

Ontologia, sive de Ente in Genere & Pneumatologia. Amsterdam, 1694.

Physica, sive de Rebus Corporeus libri quinque. Amsterdam, 1696.

Vita et Opera ad Annum 1711. Amsterdam, 1711. This work includes the philosophical texts listed separately above.

WORKS ON LE CLERC

Barnes, Annie. *Jean Le Clerc et la république des lettres.* Paris: E. Droz, 1938. Contains details on Le Clerc's often stormy career.

Colie, Rosalie. *Light and Enlightenment, a Study of the Cambridge Platonists and Dutch Arminians.* Cambridge, U.K.: Cambridge University Press, 1957. Contains a discussion of Le Clerc's thought.

Haag, Eugene, and Émile Haag. *La France protestante.* Paris: J. Cherbuliez, 1846. This work and the Barnes mentioned above contain bibliographies of Le Clerc's diverse works on philosophy, theology, history, philology, and biblical criticism.

Phillip D. Cummins (1967)

LE CLERC, JEAN [ADDENDUM]

Although embroiled in many intellectual controversies throughout his career, Jean Le Clerc's repeated disputes with Pierre Bayle were of greatest significance. Bayle had argued that Christianity strengthens Pyrrhonian skepticism in that a number of axioms of logic and metaphysics are contradicted by Christian dogmas. Therefore, even self-evidence (*évidence*) is not an infallible criterion of truth, since these axioms are self-evident, yet false. In the same vein, Bayle argued that one cannot conceive how an omnipotent and omnibenevolent God could allow human suffering either here or in the afterlife. Le Clerc replied that God's justice demands that those who freely choose to sin be punished, but conceded that the torments of hell might not be eternal.

Similarly, when Le Clerc championed the notion of plastic natures—insentient, immaterial substances causally responsible for the organization of animal bodies—Bayle argued that the hypothesis undermined the most compelling argument for God's existence by severing the conceptual connection between complex effects and conscious design.

Underlying their debates were two fundamentally different conceptions of the relation between faith and reason. For Le Clerc, Bayle's insistence on the irrationality of Christianity constituted a thinly veiled attack on religion itself, whereas Bayle saw Le Clerc's demand that scripture be interpreted according to rational principles as ultimately leading to deism or atheism.

See also Atheism; Bayle, Pierre; Deism; Faith; Pyrrho; Reason.

Bibliography

Golden, Samuel A. *Jean LeClerc*. New York: Twayne, 1972. Study of Le Clerc's career as editor of a number of highly successful learned journals.

Rosa, Susan E. "Ralph Cudworth in the *République des Lettres*: The Controversy about Plastick Nature and the Reputation

of Pierre Bayle." *Studies in Eighteenth-Century Culture* 23 (1994): 147–160.

Todd Ryan (2005)

LEGAL POSITIVISM

In many discussions of the nature of law the terms "legal positivism" and "natural law" are assumed to be the names of rival theories. In fact, each of these designations stands for a number of different and logically distinct doctrines, with the unfortunate result that in many disputes between "positivism" and "natural law" the precise point of conflict is unclear and the classification of a legal theorist as a "positivist" may afford very little indication of the nature of his theory. Thus, what is called the imperative theory of law, that is, the view that laws are commands, is usually treated as a central tenet of legal positivism; but although Jeremy Bentham and John Austin held this view, Hans Kelsen (usually regarded as the most uncompromising of modern legal positivists) held neither this view nor its corollary, that international law is not really law but a mere species of morality. Similarly, "legal positivism" is sometimes used as a designation for a thesis concerning the nature of moral judgments, including those made about the justice or injustice or the goodness or badness of human laws. This is the thesis (sometimes termed "noncognitivism") that such judgments cannot be established by reasoning but are merely expressions of human feelings or choices or "prescriptions." Kelsen held this view of moral judgments but Bentham and Austin did not. Bentham and Austin were both utilitarians who considered that moral judgments could be rationally established by the application of the test of utility, which according to Austin was also an "index" of God's commands.

A variety of other doctrines about law, besides those mentioned above, have been described as "positivist." These include the doctrine that although law and morals may often overlap or be causally related, there is no necessary or conceptual connection between them; the doctrine that judicial decisions are or should be deducible by logical means from legal rules and involve no choice or creative activity on the part of the judge; and the doctrine that there is an absolute moral obligation to obey the law, however morally iniquitous it may be.

The etymology of the word *positivism* and cognate expressions offers little guidance to its use in the philosophy of law. Since at least the fourteenth century, the expression "positive law" has been used to refer to laws laid down or made by human beings in contrast to natural or divine law, which is regarded as something discovered and not made by man. But the expression "positive law" has also long been used to refer to any law brought into being by a command or act of will and so includes the law of God as well as human legislation. More recently, the use of the expression "legal positivism" has been colored by the philosophical sense of "positivism" introduced by Auguste Comte. In this sense a "positivist" doctrine is one according to which nothing can be truthfully (or in later versions, meaningfully) said to exist unless it is in principle observable by human beings.

More important for legal theory than the etymology of the word is the identification and classification of the principal issues in relation to which philosophers of law or legal theorists have advanced views commonly styled positivist. Five such issues may be distinguished, and the discussion of these constitutes the remainder of this entry.

POSITIVISM AS A THEORY OF A FORM OF LEGAL STUDY

Bentham, Austin, and Kelsen, while differing as noted above on certain points, agreed that there is an important branch of legal study distinguished by two features: that it is not concerned with any ideal law or legal system but only with actual or existent law and legal systems; and that its concern with law is morally, politically, and evaluatively neutral. The object of this form of legal study is the clarification of the meaning of law, the identification of the characteristic structure of a legal system, and the analysis of pervasive and fundamental legal notions, such as right, duty, ownership, or legal personality. Bentham, Austin, and Kelsen were all concerned to distinguish such an "analytical" jurisprudence, as this form of legal study is now called, from critical or evaluative studies of the law, and they have stressed the importance of this distinction. However, none of these theorists—though the contrary is sometimes suggested—considered that analytical jurisprudence excluded critical or evaluative studies of the law or rendered them unimportant.

It should be observed that belief in the importance of analytical studies of the law does not strictly entail belief in other forms of legal positivism, though in fact it has usually been associated with one or more of these other forms. It is also true that not all morally or evaluatively neutral studies of the law need take an analytical form. Many sociological descriptions of the operation of law and society, and many sociological theories of the causal connection between law and other social phenomena are

also evaluatively neutral, at least in intention. Hence, some of these, too, have at times been regarded as forms of positivism.

POSITIVISM IN THE DEFINITION OF LAW

The definition of law as the command of the "sovereign" is no doubt the most prominent example of a form of positivism. But the expression "positivist" is also used in a wider sense to include any doctrine according to which law is defined as the expression of human will or as man-made, even if it does not take the form of a command. Thus, both the doctrines known in American jurisprudence by the loose title of "legal realism," according to which only decisions of courts and the predictions of such decisions are law, and those theories of international law which insist that it is composed exclusively of rules originating in custom or in agreements between states are usually described as positivist. It is to be noted, however, that both Bentham and Austin, who defined law as the command of a sovereign, extended the notion of a command to include both customary law and judge-made law. For this purpose they invoked the idea of a "tacit," or "indirect," command resting on the principle that whatever the sovereign permits he commands.

POSITIVISM AS A THEORY OF THE JUDICIAL PROCESS

Sometimes the term "legal positivism" is used to refer to the view that correct legal decisions are uniquely determined by preexisting legal rules and that the courts either do or should reach their decisions solely by logical deduction from a conjunction of a statement of the relevant legal rules and a statement of the facts of the case. This is sometimes referred to as the "automatic" or "slot-machine" conception of the judicial process; but it is doubtful whether any Anglo American writer who is usually classified as a positivist would subscribe to any such view. It is true, however, that Bentham and Austin thought that the area of choice allowed to judges by a system of case law was excessive and led to great uncertainty, and they claimed that this could and should be drastically reduced by classification and codification of the law in clear and detailed terms. But they were both well aware of the fact of judicial legislation and creative activity, and as noted above, they sought to reconcile this fact with their definition of law as the command of the sovereign by using the idea of a tacit command. The doctrine that a judge should not exercise choice in his decision of cases but should merely be the mouthpiece of previously existing law is to be found in the works of eighteenth-century writers not usually classed as positivists, such as Baron de Montesquieu's *L'esprit des lois.* They looked upon this doctrine as a corollary of the doctrine of the separation of powers and as a protection of the individual against arbitrary decisions, uncertainty, and privilege.

POSITIVISM AS A THEORY OF LAWS AND MORALS

It seems that all writers classed as positivists have subscribed to the view that unless the law itself provides to the contrary, the fact that a legal rule is morally iniquitous or unjust does not entail that it is invalid or not law. This view may also be expressed as the claim that no reference to justice or other moral values enters into the definition of law. "The existence of law is one thing: its merit or demerit another" (Austin). "Legal norms may have any kind of content" (Kelsen). Such a denial of a necessary or definitional connection between law or legal validity and morality is perhaps the principal point of conflict between legal positivism and theories of natural law. For nearly all variants of the latter refuse to recognize as law or legally valid rules that violate certain fundamental moral principles. It is, however, important to remember that this denial of a necessary connection between law and morals is compatible with the recognition of many other important connections between them. Thus few, if any, positivists have denied that the development of the law has in fact been influenced by morality or that moral considerations should be taken into account by legislators and also by judges in choosing between competing interpretations or conflicting claims as to what the law is.

POSITIVISM AND THE OBLIGATION TO OBEY LAW

If positivism has become a pejorative term, it is very largely because it has been identified by some critics with the claim that where a legal system is in operation, there is an unconditional moral obligation to obey the law, however unjust or iniquitous it may be. This claim may be based either on the view that there is a moral obligation to obey law as such or on the belief that the actual existence of a legal system, however oppressive or unjust, provides large numbers of human beings with a minimum of peace, order, and security and that these are values that no individual is morally justified in jeopardizing by resistance to the law. The German legal theorist K. M. Bergbohm, perhaps the best-known legal positivist in continental Europe in the nineteenth century, held this view; but though he in fact also subscribed to other forms

of legal positivism described above, this view is logically quite independent of them. Utilitarian positivists, such as Bentham and Austin, held that resistance to law might be justified in extreme cases, but before this step was taken, careful calculations in terms of utility were necessary to ascertain that a balance of good over evil was likely to result. They criticized the doctrine of natural law and natural rights not because they believed that there was an unconditional obligation to obey the law, but because in their view these doctrines presented standing temptations for men to revolt without making such calculations of the consequences.

See also Austin, John; Bentham, Jeremy; Comte, Auguste; Kelsen, Hans; Natural Law; Noncognitivism; Philosophy of Law, History of; Philosophy of Law, Problems of; Positivism.

Bibliography

WORKS BY LEGAL POSITIVISTS

Austin, John. *Essay on the Uses of the Study of Jurisprudence.* London, 1863.

Austin, John. *Lectures on Jurisprudence, or the Philosophy of Positive Law.* London, 1863.

Austin, John. *The Province of Jurisprudence Defined.* London, 1832.

Bentham, Jeremy. *A Fragment on Government.* London: T. Payne, P. Elmsly, and E. Brooke, 1776.

Bentham, Jeremy. *Introduction to the Principles of Morals and Legislation.* London, 1789.

Bentham, Jeremy. *The Limits of Jurisprudence Defined.* New York: Columbia University Press, 1945.

Bentham, Jeremy. *The Theory of Legislation,* 2 vols. London: K. Paul, Trench, Trubner, 1931.

Bergbohm, K. M. *Jurisprudenz und Rechtsphilosophie.* Leipzig: Duncker and Humblot, 1892.

Kelsen, Hans. *The General Theory of Law and State.* Cambridge, MA: Harvard University Press, 1945.

WORKS ON LEGAL POSITIVISM

Fuller, Lon. "Positivism and Fidelity to Law." *Harvard Law Review* 71 (1958): 630–672.

Gray, J. C. *The Nature and Sources of Law.* New York, 1909.

Hart, H. L. A. *The Concept of Law.* Oxford: Clarendon Press, 1961.

Holmes, O. W. "The Path of the Law." In his *Collected Legal Papers.* Boston, 1920.

Morison, W. L. "Some Myths about Positivism." *Yale Law Journal* 68 (1958): 217–222.

Pound, Roscoe. "Mechanical Jurisprudence." *Columbia Law Review* 8 (1908): 605ff.

Roguin, E. *La science juridique pure.* Paris: Sousanne, 1923.

Ross, Alf. *On Law and Justice.* London: Stevens, 1958.

Stone, J. *The Province and Function of Law.* London: Stevens, 1947.

Numerous articles critical of various forms of legal positivism may be found in the *Natural Law Forum,* published since 1955 by the Notre Dame Law School.

H. L. A. Hart (1967)

LEGAL POSITIVISM: ANGLO-AMERICAN LEGAL POSITIVISM SINCE H. L. A. HART

HART'S POSITIVISM AND DWORKIN'S INITIAL OBJECTIONS

Herbert Lionel Adolphus Hart's version of legal positivism, developed in *The Concept of Law* and refined in the Postscript to the second edition, has been the centerpiece in the development of contemporary legal positivism as well as the focal point of the strongest and most interesting objections to it in the philosophy of law. Hart's own argument builds on the work of positivists who preceded him. In particular, Hart seeks to address and correct the main shortcomings he identifies in the theories of law offered by John Austin and Hans Kelsen. Both Austin and Kelsen thought that laws are a distinguishable subset of norms, identifiable by their possession of an intrinsic and necessary property. In other words, they saw the fundamental project of jurisprudence as determining what it is to be *a law*. For Austin, laws are orders or commands backed by a threat of sanction issued by a sovereign. The threat of sanction distinguishes a command from a plea or a request while the fact of issuance by a sovereign distinguishes legal commands from all other commands. A sovereign, properly so called, is a person who has secured the habit of obedience from the vast majority of the populace and who is not in the habit of obeying anyone. Like Austin, Kelsen holds that sanctions are both intrinsic to law and necessary for their existence. Unlike Austin, however, he argues that although sanctions are imposed on citizens, legal norms are directives to officials to impose sanctions against citizens in the event that they behave in ways identified as sanctionable by the law.

One of Hart's most important claims is that laws are irreducibly of at least two sorts: (1) power-conferring, or secondary, rules; and (2) primary rules that impose obligations to act or to forbear from acting. Power-conferring rules are not themselves reducible to rules that impose obligations because the failure to comply with a power-conferring rule renders one's action a nullity in law and is

neither the basis of liability nor a liability itself. An act nullified is one that does not have the legal effect a successful exercise of the legal power would have produced. A failure to exercise a power according to the required formula is not a failure to comply with an obligation to act or to forbear from doing so. Nullification, in turn, is no sanction. In one fell swoop, then, Hart undermines both the claim that legal rules are of one type and that what is distinctive of the type is the presence of sanctions.

Power-conferring, or secondary, rules are themselves of two kinds: private and public. Private secondary rules empower those governed by law to alter the normative relations among themselves (e.g., as in transferring private property through contracts, marital agreements, and wills) and a legal power to call upon the state's resources to enforce those relations. Public power-conferring rules create and regulate the offices to which legal authority attaches, and their existence is implicated in the very idea of private power-conferring rules as well. For Hart, the most important secondary rule is what he terms the rule of recognition. This rule sets forth the conditions that must be satisfied in order for a norm to constitute part of the community's law and in so doing constitutes an identity condition of a legal system. Beyond that, the rule of recognition both confers a power and imposes a duty on certain officials to evaluate conduct in the light of the norms that satisfy the criteria of law it sets forth.

It is important to Hart's account that rules are expressed in general terms. Such terms possess a core of settled meaning and a penumbra of disputed meaning. This maps on to the distinction between those cases in which no competent speaker of a language can legitimately deny that a rule applies (the core of a rule) and those cases in which reasonable, competent speakers of a language can legitimately disagree (the penumbra). In cases falling within the core of a rule, the law settles the matter and a judge is under a duty to apply the law to the facts at hand. In cases falling within the penumbra of a rule, there is no settled law on the matter and a judge must exercise discretion. The scope of judicial duty is fixed by settled meaning or practice. Where either meaning or practice runs out, judicial discretion—a rationally constrained power, not a license—enters. Some contend that Hart believed that the function of law is to guide conduct by reasons, but this is not his view. Whereas the heart's natural function is to pump blood, the law has no such function. Instead, it can serve any number of human ends and purposes—some laudatory and others evil. Hart's view is that whatever business law does, it goes about it by regulating conduct through rules that are reasons. That is its mode of operation, not its function.

For Hart, law is to be understood in terms of its structure—in the union of primary and secondary rules—not in terms of its having a natural function. Thus, Hart's positivism can be identified with the following tenets: laws are rules; legal rules are of two irreducible types—power conferring and duty imposing. Wherever there is law, there is a rule of recognition that sets out the criteria for the laws of a community. That rule is part of the identity conditions of a legal system. All rules are expressed in general terms, and the set of norms that satisfy the criteria of legality are finite. Thus, there will be gaps in meaning and in settled or controlling law. Discretion is inevitable.

Beyond his particular disagreements with them, Hart's positivism shares with its predecessors a view about the relationship between law and morality. This is the view that there is no necessary connection between law and morality—the so-called separability thesis. Most commentators take the separability thesis to be the sine qua non of legal positivism, but some recent work on legal positivism has raised doubts about the centrality of this claim to the field. Still, there is no denying that Hart was committed to it.

For our purposes, the important features of Hart's positivism are its commitment to judicial discretion and the rule of recognition. The set of binding legal standards in any community is determined by the criteria of legality in that community. Those criteria are set forth in a rule of recognition whose existence and content is fixed by the critical, reflective attitude (the internal point of view) of officials—in particular, judges. The set of standards or rules that satisfy these criteria is finite and thus, in principle, there will be disputes that are not resolved by available legal resources. Judges will be required, therefore, to exercise discretion: a legal power to rely on nonlegal standards, some of which will be moral standards, to resolve legal disputes. This is the so-called discretion thesis.

DWORKIN'S FIRST OBJECTIONS

Discretion is inevitable where settled law runs out. It is this feature of Hart's account that Ronald Dworkin exploits and which forms the basis of his first and most famous objections to Hart's account. Cases in which the *law runs out* are *hard cases*. Because discretion is a constrained power, judges must decide hard cases by appealing to rationally defensible standards. *Ex hypothesi*, such standards are not part of the law, but they must be prin-

ciples or other standards that constitute good reasons or justifications for the decisions a judge reaches at law. Dworkin notes that judges do not act as if the standards to which they appeal in deciding hard cases are optional for them. If we allow, instead, that the moral principles to which judges appeal in deciding hard cases are sometimes legal, binding because they are law, and law because they express an appropriate dimension of morality, one must abandon not only the discretion thesis but the other basic tenets of positivism as well—such is the nature of the relations among them.

First, if the moral *principles* to which judges appeal in hard cases are even sometimes law, then it cannot be true that all laws are *rules*. Second, if these principles are law, then they are not law in virtue of their satisfying the rule of recognition. They have no institutional source in that sense. This suggests that their legality depends on their content, in particular, on their moral value or worth. But if they are binding in virtue of the fact that they express an appropriate aspect of fairness or justice, then the separability thesis—which appears to reject the claim that legality can depend on morality—must be abandoned. So, too, must the rule of recognition, for it is not true that wherever there is law, there is a Master Rule that determines fully a community's binding legal standards.

EXCLUSIVE AND INCLUSIVE LEGAL POSITIVISM

Positivists have adopted one of two approaches to Dworkin's objections. Common to both is a willingness to grant one of Dworkin's main premises, namely, that at least in some cases the moral standards judges apply are binding on them, not optional. These approaches differ with respect to the second premise: In virtue of their being binding on officials, are those standards part of the community's law?

Those who reject the second premise are exclusive legal positivists (exclusivists); those willing to accept both premises are inclusive legal positivists (inclusivists, incorporationists, or soft positivists). Though he does not employ these labels, Joseph Raz (1939–) is most often cited as the leading positivist of the first sort whereas Jules Coleman (1947–) is usually cited as the most prominent advocate of the latter approach. Both approaches reject Dworkin's claim that the binding nature of moral principles undermines positivism, but for importantly different reasons.

Raz emphasizes a significant distinction between a norm being binding on an official (e.g., a judge) and its being binding in virtue of its being the law of his community. Laws of jurisdiction *A* may, under certain conditions, make laws of jurisdiction *B* binding on officials in jurisdiction *A*. That is not enough to make the laws of jurisdiction *B* laws of jurisdiction *A*. Understood in this way, Raz's argument is that Dworkin has not made the case that moral principles are part of the law. In fact, however, Raz advances the much stronger claim that if moral principles are binding on officials, they can only bind in the way that norms of other jurisdictions do. This stronger argument relies only on general considerations regarding the relationship between the concepts of law and authority and from no distinctive commitments of legal positivism. Instead, Raz begins with a putative conceptual truth about law: that it necessarily claims to be a legitimate authority. As long as governance by law is not an incoherent idea, then law must be the sort of thing of which the claim to authority could be true—even if, as a factual matter, it always turns out to be false. This feature of the claim to authority constrains the kind of thing law can be, but the exact constraints it imposes depends on what the claim to being a legitimate authority entails.

On Raz's account, an authority mediates between persons and reasons in such a way that in accepting an authority, an agent is (with rare exceptions) precluded from appealing directly to the reasons that would justify the authority's directives. If one appeals to what Raz calls the *dependent reasons* in order to identify what the law is or to determine its content, then one vitiates the law's claim to authority. Since the dependent reasons on which law relies and which justify laws are moral reasons, it follows that morality itself cannot be a condition or ground of law. Instead, all law must have what Raz calls a social source. Thus, the claim to authority in conjunction with Raz's theory of authority entails what has come to be called the sources thesis. Some positivists, most notably John Gardner (1965–), advance the view that the core of legal positivism is the sources thesis but note that the thesis itself derives from no claim of legal positivism at all. It derives, instead, from the conjunction of a conceptual claim about law and a theoretical and quite general, if controversial, theory of the meaning of authority.

The inclusivist grants both of Dworkin's premises: that moral standards can sometimes be legally binding and that they can sometimes be part of the community's law. His strategy is to show that none of Dworkin's objections to legal positivism are entailed by accepting these two premises.

Coleman argues that it cannot be Dworkin's view that all moral principles are law merely in virtue of their content or moral merits. That would make all of morality

part of the law everywhere and always. Put another way, there must be some institutional or social fact that makes some moral principles part of the law in the jurisdictions in which they are law. But then it would have to be that fact that makes them law, not their individual merits. Coleman then argues that there is no reason in principle why the relevant social or institutional fact that renders moral principles law could not be agreement among officials to count those moral principles as law. Nothing in legal positivism, in any case, precludes that.

THE NEXT PHASE

It is important to draw a distinction that is rarely explicitly made but which is central to understanding the various current disputes among positivists and between positivists and their critics. This is the distinction between the grounds, or sources of law, on the one hand, and the grounds of the grounds of law, on the other. The grounds, sources, or criteria of law refer to the *test* for legality within a community. Both the sources thesis and inclusivism are claims about possible constraints on the grounds of law. The sources thesis claims that all such grounds must be social sources; the inclusivist denies that. Dworkin's argument in "The Model of Rules" is that morality can be a ground of law, which is incompatible with the sources thesis and thus with exclusive legal positivism but not with inclusivism.

The fact that inclusivism shares with Dworkin the view that morality can be a source of law has led Dworkin to chide inclusivists, especially Coleman, for having an underdeveloped version of his view. The criticism cuts no ice, however. The core of legal positivism is not a claim about the grounds of law; it is a claim about the grounds of the grounds of law. And on this score, both exclusivists and inclusivists agree that only social facts—facts about individual behavior and attitude—can be the grounds of the grounds of law. What distinguishes positivists from one another is what they take this claim about the grounds of the grounds of law to allow. Inclusivists believe that commitment to social facts at this level does not impose any constraints on potential grounds of law. They hold, moreover, that nothing about the nature of law or authority does either. In contrast, exclusivists hold that facts about the nature of law, in particular its role in our practical lives, imposes constraints on potential grounds of law. Thus, exclusivists accept the sources thesis whereas inclusivists do not. That does not render the inclusivist a proto-Dworkinian.

The sources thesis is said to follow from the conjunction of the claim that law necessarily claims authority in conjunction with a particular account of what that claim requires. In order to escape this implication, the inclusivist might reject the conceptual and/or the theoretical claim, reject the alleged relationship between them, or accept both but deny that they entail the sources thesis. Coleman, for one, rejects both of the premises of the argument, but he is willing to accept them because, he argues, they do not in fact entail the sources thesis.

Roughly, his argument is this: Whereas appealing to the dependent reasons that purport to justify an authoritative directive in order to determine the directive or its content undermines the directive's claim to authority, it does not follow that every moral principle offered as a condition of legality for norms must be among the dependent reasons that justify the particular directive in question. For instance, a clause making *equal protection of the law* a condition of legality for every putative legal rule need not be a reason that justifies any particular law—one outlawing murder, for example. There is nothing in the logic of law that precludes reasons *R1* through *R5* being the grounds that justify a law *L*, but reasons *R6* through *R9* being the conditions of the legality of *L*. The mere fact of appealing to moral premises as the conditions of law does not mean that one is appealing to the reasons that would be offered to justify the law.

As Raz characterizes it, the theory of authority claims that one cannot appeal to moral principles to identify the law or to determine its content. This is an epistemic constraint on identifying law and determining its content. But inclusivism is a theory about the grounds of law: the conditions that make law *determinate*—what makes it the law. Inclusivism is, in a broad sense, an ontological or metaphysical theory that may well be untouched by the epistemic constraints that are said to fall out of the appropriate theory of authority.

DWORKIN AND THE POSITIVISTS REDUX

Dworkin has not been persuaded by either the inclusivist or the exclusivist strategies, but he has focused primarily on the inclusivist strategy. Recall that Coleman responds to Dworkin's objections to Hart by noting that moral facts cannot be law merely in virtue of their being moral facts, for that would render all moral facts legal facts. Instead, the positivist need not deny that moral facts can be legal facts; he need only argue that moral facts are legal facts in virtue of certain legally significant social facts about them—typically acceptance among officials of their status as part of the law. Any such account of the

way in which social facts make moral facts part of the law is going to be a form of legal positivism.

Dworkin has responded in two ways: one positive, the other critical. On the one hand, he has developed an argument, expounded primarily in *Law's Empire*, designed to show how certain kinds of moral facts, not just social facts, turn other moral facts into law. This is his *theory of constructive interpretation*. On the other hand, he has offered a variety of interesting and important objections to the inclusivist claim that moral facts can be grounds of law and, ultimately, to the claim that they can be grounds of law in virtue of certain social facts or practices.

The point of departure for some of these objections is the idea that morality is inherently controversial, and this fact about it is incompatible with the claim that it can be a condition of legality. At one point, Dworkin argued that positivism holds that the function of law is to guide conduct, in part by resolving disputes and disagreements about what one ought to do. This is one reason for his oft-repeated view that legal positivists identify law with plain or hard fact—the sort of thing one could determine with near certitude by looking it up in a book. Morality is too controversial for its inclusion in law to serve this function. This objection has no force. Presumably, one role of morality in our lives is to guide conduct. If morality is capable of guiding conduct, therefore, it is capable of guiding conduct whether it is part of the law or not. Beyond that, nothing in legal positivism would suggest that legal disputes must be resolved in a way that is essentially uncontroversial.

A more serious version of the objection maintains that the essentially controversial nature of morality means that judges applying the criteria of legality that includes morality will often disagree, and this level of disagreement is incompatible with the positivist claim that at the foundation of law resides a social rule, namely, the rule of recognition. A social rule has two dimensions: shared, convergent behavior and a shared critical, reflective attitude toward that behavior. One consequence of allowing that morality might be a condition of legality is widespread disagreement among officials that is incompatible with the requisite agreement necessary for the criteria of law to be determined by a social rule among officials.

Note that the key here is the connection between the grounds of law and the grounds of the grounds of law. If morality can be a condition of legality (a ground of law), as inclusivists claim, then the foundation of law (the grounds of the grounds of law)—the rule of recognition—cannot be a social rule. The claim that the grounds of the grounds of law are social facts precludes moral facts from being among the grounds of law. Positivism cannot have it both ways. The problem with this objection is that it treats disagreement about whether the grounds of law are satisfied as if it entailed disagreement about what the grounds of law are. You and I can disagree about whether someone is intelligent while agreeing that we should only hire intelligent people for our company. Disagreement about whether the conditions of legality are satisfied is perfectly compatible with agreement about what those conditions are. If any sort of agreement is required in order for there to be a rule of recognition, it is agreement of the second sort.

At one point, Dworkin responded that all disagreements about whether a rule applies could be formulated as disagreements about what the rule is, thereby raising doubts about whether the distinction Coleman points to between agreement about the criteria and agreement about its applications is helpful. Dworkin's response has proven unpersuasive because it identifies a rule with the set of its instances, which cannot be a plausible understanding of what it is to be a rule. Dworkin eventually adopted a more interesting line of attack, which begins by reflecting on a more general philosophical concern regarding the sort of agreement that is necessary in order for disagreements to make sense; that is, for individuals actually to be disagreeing with one another rather than merely talking past one another. He associates legal positivism with the view that judges and other officials must agree on the grounds of law in order for their disagreements about what satisfies those grounds to be meaningful.

If he has diagnosed the commitments of legal positivism correctly, Dworkin may have identified a powerful objection to it since it appears as if legal positivism cannot explain the possibility of disagreement about what the grounds of law are. Yet such disagreement is a significant feature of legal practice. At bottom, Dworkin has a picture of legal positivism that is very likely warranted by Hart's formulations and much of the positivist literature that has followed. In this picture, legal positivism represents a certain architectural rendering of law and of legal practice. Law is a closed normative system whose boundaries are determined by the scope of agreement. It has a set of initial premises (we can think of these as rules of recognition) whose existence depends on agreement about what they assert or prescribe. Once these rules are in place, we can have a practice called law. These rules make the practice possible and are both inside and out-

side the practice. They are part of the law, but not, as it were, in the mix. If anything, they are like Carnapian *meaning postulates*. They are not subject to revision from within the practice; they are immune in that sense. If the practice fails to achieve the aims we have for it, we can change the ground rules and have a new practice. But we do not have a practice that changes from within. We can disagree about what falls within the practice, but we cannot disagree about what the rules are that constitute the practice itself.

Dworkin's most interesting objections to legal positivism can be recast as trying to shed doubt on this *architecture* of the law. For Dworkin, the fluidity of the boundaries between law and other normative systems, between what is inside and outside law, his deep anti-Archimedianism, are all different ways of getting at the same problem. Dworkin is Willard Van Orman Quine to legal positivism's Rudolph Carnap.

As Coleman and others have argued, the heart of legal positivism is the claim that the grounds of the grounds of law are social facts—facts about behavior and attitudes. It is not obvious that this claim entails either of Dworkin's objections to positivism, namely, the claim that positivism cannot account for disagreement about the grounds of law or the related claim that positivism imposes a Carnapian architecture on legal practice. In fact, the Quinean picture of legal practice is completely compatible with positivism.

ANOTHER PICTURE

If, for the sake of simplicity, we use the term *rule of recognition* to refer to the grounds of law, then positivism can hold that the rule of recognition is at the *center* of law, not at the foundation of law. Thus, the rule of recognition is in the mix, subject to revision and even abandonment on the same grounds as are other rules and standards within law. There is a distinction worth emphasizing between the *existence* conditions of a rule and its *revisability* conditions. The rule of recognition's special status, moreover, is not a function of it being at the foundation of law but of its inferential importance. Much of the rest of the law of particular communities makes sense inferentially in virtue of the rule of recognition. As the importance of various grounds of law to inferences that warrant other settled areas of the law diminishes, the likelihood of revising that *ground of law* increases.

Nor is there a distinctive problem in understanding disagreement among participants about the grounds of law, for the rule of recognition is not rigidly fixed by agreement. The complexity of any particular legal practice is likely to mean that there will be different and quite varied views about which putative grounds of law are more or less central to the practice in place.

More importantly, the claim that the content of the rule of recognition is fixed by social facts does not entail that the content of the rule, or its proper formulation, is transparent or otherwise available to officials—that is, those whose conduct is regulated by it. No more so must the content of the rule be transparent than must the rule governing the use of personal pronouns be accessible to ordinary speakers of a language whose speech it governs. If transparency of the rule to officials is required to coordinate their behavior—and it is an empirical question whether in fact it is—sharing the rule in the sense Dworkin attributes to positivism is no more than an efficiency condition of law and not a theoretical commitment of legal positivism. There is nothing in the idea that law rests ultimately on social facts that is itself incompatible with disagreement about the grounds of law. Still less does the claim that law is created by social facts alone entail the architectural view of law that renders the grounds of law immune to revision from within.

Positivism need not necessarily be understood in the way sketched here. Rather, the above sketch is designed to suggest only that the social facts thesis does not render positivism vulnerable to the charge that it cannot explain disagreement about the grounds of law. Nor does the social facts thesis leave positivism vulnerable to the charge that it pictures law as *shut off* from or *bounded* by other normative systems (other legal systems and other schemes of regulating conduct) in the way the architectural picture of a definite inside and outside of law suggests. Whether this line of argument on behalf of positivism proves ultimately persuasive remains very much an open question.

See also Austin, John; Carnap, Rudolf; Dworkin, Ronald; Epistemology; Hart, Herbert Lionel Adolphus; Kelsen, Hans; Legal Positivism; Quine, Willard Van Orman.

Bibliography

Bratman, Michael. *Faces of Intention*. Cambridge, U.K.: Cambridge University Press, 1999.

Campbell, Tom D., ed. *Legal Positivism*. Aldershot, U.K.: Ashgate Publishing, 1999.

Coleman, Jules. "Constraints on the Criteria of Legality." *Legal Theory* 6 (2) (2000): 171–183.

Coleman, Jules. "Incorporationism, Conventionality, and the Practical Difference Thesis." In *Hart's Postscript*. Edited by Jules Coleman, 99–148. Oxford: Oxford University Press, 2001.

Coleman, Jules. "Negative and Positive Positivism." In *Markets, Morals, and the Law*. Cambridge, U.K.: Cambridge University Press, 1988.

Coleman, Jules. *The Practice of Principle*. Oxford: Oxford University Press, 2001.

Coleman, Jules. "Second Thoughts and Other First Impressions." In *Analyzing Law: New Essays in Legal Theory*. Edited by Brian Bix, 257–322. Oxford: Oxford University Press, 1998.

Coleman, Jules, and Brian Leiter. "Legal Positivism." In *A Companion to Philosophy of Law and Legal Theory*. Edited by Dennis Patterson, 241–260. Oxford: Blackwell, 1996.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

Dworkin, Ronald. "Objectivity and Truth: You'd Better Believe It." *Philosophy and Public Affairs* 25 (1996): 87–139.

Dworkin, Ronald. "The Model of Rules I," in *Taking Rights Seriously*, 14–45. Cambridge, MA: Harvard University Press, 1997.

Dworkin, Ronald. "The Model of Rules II," in *Taking Rights Seriously*, 46–80. Cambridge, MA: Harvard University Press, 1997.

Finnis, John. "On the Incoherence of Legal Positivism." *Notre Dame Law Review*. 75 (2000): 1597–1611.

Finnis, John. "The Truth in Legal Positivism." In *The Autonomy of Law: Essays on Legal Positivism*. Edited by Robert P. George, 195–214. Oxford: Clarendon Press, 1996.

Gardner, John. "Legal Positivism: 5 ½ Myths." *American Journal of Jurisprudence* 46 (2000): 199–227.

George, Robert P., ed. *The Autonomy of Law: Essays on Legal Positivism*. Oxford: Clarendon Press, 1996.

Green, Leslie. "Legal Positivism." *Stanford Encyclopedia of Philosophy*, 2003. Available from http://plato.stanford.edu/entries/legal-positivism.

Green, Leslie. "Positivism and Conventionalism." *Canadian Journal of Jurisprudence* 12 (1999): 35–52.

Hart, H. L. A. *The Concept of Law*. Edited by Penelope A. Bulloch and Joseph Raz. Oxford: Clarendon Press, 1994.

Himma, Kenneth Einar. "H. L. A. Hart and the Practical Difference Thesis." *Legal Theory* 6 (1) (2000): 1–43.

Himma, Kenneth Einar. "Inclusive Legal Positivism." In *The Oxford Handbook of Jurisprudence and Legal Philosophy*. Edited by Jules Coleman and Scott Shapiro, 125–65. Oxford: Oxford University Press, 2001.

Kramer, Matthew. *In Defense of Legal Positivism*. Oxford: Clarendon Press, 1999.

Leiter, Brian. "Legal Realism, Hard Positivism, and the Limits of Conceptual Analysis." In *Hart's Postscript*. Edited by Jules Coleman, 355–370. Oxford: Oxford University Press, 2001.

Leiter, Brian, ed. *Objectivity in Law and Morals*. Cambridge, U.K.: Cambridge University Press, 2001.

Lyons, David. "Principles, Positivism, and Legal Theory." *Yale Law Journal* 87 (1997): 415–435.

Marmor, Andrei. "Exclusive Legal Positivism." In *The Oxford Handbook of Jurisprudence and Legal Philosophy*. Edited by Jules Coleman and Scott Shapiro, 104–124. Oxford: Oxford University Press, 2001.

Marmor, Andrei, ed. *Law and Interpretation*. Oxford: Oxford University Press, 1995.

Perry, Stephen. "Varieties of Legal Positivism." *Canadian Journal of Law and Jurisprudence* 9 (1996): 361–381.

Postema, Gerald. "Coordination and Convention at the Foundation of Law." *Journal of Legal Studies* 11 (1982): 165–203.

Raz, Joseph. *The Authority of Law: Essays on Law and Morality*. Oxford: Clarendon Press, 1979.

Raz, Joseph. *Ethics in the Public Domain*. Oxford: Clarendon Press, 1995.

Shapiro, Scott J. "Law, Morality, and the Guidance of Conduct." *Legal Theory* 6 (2) (2000): 127–170.

Shapiro, Scott J. "Laws, Plans, and Practical Reason." *Legal Theory* 8 (2002): 387–441.

Shapiro, Scott J. "On Hart's Way Out." In *Hart's Postscript*. Edited by Jules Coleman, 149–192. Oxford: Oxford University Press, 2001.

Soper, Philip. "Legal Theory and the Obligation of a Judge: The Hart/Dworkin Dispute." *Michigan Law Review* 75 (1977): 473–519.

Soper, Philip. *A Theory of Law*. Cambridge, MA: Harvard University Press, 1984.

Waldron, Jeremy. "Normative or Ethical Positivism." In *Hart's Postscript*. Edited by Jules Coleman, 410–434. Oxford: Oxford University Press, 2001.

Waluchow, W. J. "Authority and the Practical Difference Thesis: A Defense of Inclusive Legal Positivism." *Legal Theory* 6 (1) (2000): 45–81.

Waluchow, W. J. *Inclusive Legal Positivism*. Oxford: Clarendon Press, 1994.

Zipursky, Benjamin C. "The Model of Social Facts." In *Hart's Postscript*. Edited by Jules Coleman, 219–270. Oxford: Oxford University Press, 2001.

Jules L. Coleman (2005)

LEGAL REALISM

Beginning about 1920, an iconoclastic group of American legal writers, led by K. N. Llewellyn, Walter Wheeler Cook, Jerome Frank, Herman Oliphant, and Underhill Moore, denounced the established legal tradition as formalistic and conservative. That tradition, they charged, wrongly saw the law as a complete and autonomous system of logically consistent principles, concepts, and rules. To apply the law was to unfold the ineluctable implications of those rules. The judge's techniques were socially neutral, his or her private views irrelevant; judging was more like finding than making, a matter of necessity rather than choice. The realists, by contrast, saw legal certainty as rarely attainable and perhaps even undesirable in a changing society. In their view the paramount concern of the law was not logical consistency but socially desirable consequences. Law was an instrument of government, and jurisprudence should focus less on legal concepts than on social facts.

BASIS OF LEGAL REALISM

According to the realists, legal decisions were not compelled; choice was necessary at every step. Just as lawmakers built their ideological preferences into a statute, judges built theirs into their formulation of "the facts" of a case. Legal concepts represented nothing more than tentative decisions to consider diverse cases identical with respect to a given concern. Unless readjusted continually, such concepts could be rendered irrelevant by changing circumstances and purposes.

Realism meant opposition to illusion or pretense, sometimes to abstractions or appearances. Judges had always made law, but now, the realists insisted, they must know and say that they did. They must acknowledge their responsibility instead of attributing their choices, through tortured technicalities, to the compulsions of legal doctrine. (Oliver Wendell Holmes, the favorite judge of the realists, had said that law becomes more civilized as it becomes more self-conscious.) If the judges' latent motives and official reasons were reconciled, their judgments would be not only more honest but more informed. Moreover, assumptions about the nature of law could then be considered in the light of scientific knowledge of the actual workings of legal institutions, and assumptions about social policy could be seen in the light of scientific knowledge about society. The realists took the possibility of such scientific knowledge for granted. They further assumed that society had its own mechanisms for effecting changes and that, in general, the law should reflect social change, not shape it. Yet the realists avoided the conservative conclusions that usually accompany this view, for being, above all, reformers, they believed that the constant flux of modern society required a legal system flexible enough to match its pace.

Like other iconoclasts, the realists saw rationalization and self-deception beneath traditional claims to objectivity. They "saw through" appearances, theories, and justifications to underlying motivations or functions. They sought complete candor. The natural sciences provided their model; John Dewey, their philosophical vocabulary. They also drew on the tradition of sociological jurisprudence, which in both Europe and America had already prescribed the study of society as the proper way to discover social preferences beneath the neutral forms of the law.

These ideas were more influential in America because of the unique power of American judges to declare statutes unconstitutional. Moreover, the differences among the states in their approach to identical legal problems frustrated belief in inexorable solutions. Thus, Holmes had asserted since the 1870s that "the true grounds of decision are considerations of policy and of social advantage." John Chipman Gray saw the sources of law brought to life only in the crucial act of judicial interpretation; he believed that since courts have the last word, "all law is judge-made law." At the turn of the twentieth century Roscoe Pound attacked "mechanical jurisprudence," distinguished "law in action" from "law in books," and conceived of a sociological jurisprudence that would increase legal sensitivity to social needs and to the social effects of legal rules.

The realists were distinctive, however, in their preoccupation with the processes of judicial decision, with how law is made. They put forward a theory of precedent starting from Llewellyn's assertion that "a case stands not for one thing, but for a wide variety of things." Following Dewey, for whom a judgment was always somebody judging something, they stressed the crucial position of the judge who decided whether a case was "the same" as a previous case—that is, which similarities between them should be considered important. Skeptical of principles abstracted from a particular factual context, the realists found support in the common-law tradition that principles should evolve from rather than precede the disposition of particular cases. They trusted the judge's trained reaction to the entire set of facts before him—his "intuition of experience" (Oliphant), which depended on "knowing how" rather than "knowing that"—much more than they trusted the justification he supplied in his opinion. They therefore wanted precedents to be based on what a court actually did in response to a particular set of facts, not on its language. But emphasizing particularity means getting less direction from previous cases, for facts vary enormously. The more that precedent presupposes factual similarity between cases, the fewer its applications; future judges are freer. No two cases are identical, and if any distinction distinguishes, no precedents are possible. Logically, it is always open to a judge to decide either way, to see a previous case as a precedent or not. Some realists therefore concluded that every decision was a "free" moral decision. This conclusion, shorn of the analysis of the logic of precedent behind it and interpreted simply as giving judges greater discretionary power than the traditional view allowed them, was seen by most of the legal community as the essential message of legal realism.

RISE OF LEGAL REALISM

Grant Gilmore has related the realist's view of precedent to the remarkable increase, starting around 1890, in the

amount of litigation and in the proportion of cases reported, an increase that threatened to inundate a system depending on “a comfortable number of precedents, but not too many.” According to Gilmore, the realists responded to this crisis by allowing fewer cases to count as precedents. In this way, Gilmore has noted, legal realism was part of the major social developments of 1880–1930, notably the rise of urbanism and modern industrialism and technology; during this period realism was not confined to the law. The search for fact, for concreteness, for the truth behind appearances, can be found everywhere—in literature, in painting, in social criticism. Consider, for example, the salient characteristics of a movement quite unrelated to legal theory—progressive education. For both progressive education and legal realism, pursuing “reality” meant going from theoretical formulas to what worked in practice, from books to life, from text to context, from passively and mechanically transmitting a received tradition to actively and flexibly responding to each pupil or case. Both progressive education and legal realism flourished in the 1930s during the New Deal. Both can be seen as to some extent a response to sheer numbers, to universal education and the increase in litigation, respectively.

INFLUENCE OF LEGAL REALISM

Throughout the law the realists contributed to greater candor about the social bases of decision. They also suggested specific improvements in practical areas of the law—for example, Charles E. Clark on covenants’ running with the land, Cook on conflict of laws, Arthur Corbin on contracts, Leon Green on torts, and Llewellyn on sales. On the other hand, they underestimated the role of generalization and of justification in the law. Dewey had distinguished clearly between the “logic of inquiry” and the “logic of exposition,” between an argument’s source and its persuasiveness. Yet the realists often pointed to a judge’s psychological processes or social background as if they were demonstrating the irrelevance of her justifications or the speciousness of her claim to be applying rules.

Realism is especially inadequate if taken to be the comprehensive explanation or theory of the nature of law suggested by the definitional form of certain central realist slogans. Thus, realists constantly endorsed Holmes’s statement that “the prophecies of what the courts will do in fact … are what I mean by the law.” This remark can be accepted as a paradoxical emphasis on the individual discretion inherent in applying “open-textured” concepts to particular circumstances; accordingly, the exercise of individual discretion becomes part of any adequate concept of law. But Holmes’s remark cannot be accepted if it is read as an assertion that the best understanding of legal reality derives from equating law with prediction. A predictive viewpoint obscures the role of legal rules as guides to conduct. As H. L. A. Hart said, “legal rules function as such in social life: they are *used* as rules not as descriptions of habits or predictions” (*The Concept of Law,* Oxford, 1961, pp. 134–135). If the normative character of legal rules were not generally accepted, our concept of law would be entirely different.

However, it may be that attempts, like that of the realists, to jolt accepted habits of thought must rely on paradox and exaggeration. Pierre-Joseph Proudhon said, “Property is theft,” knowing full well the immediate sense in which “property” is not “theft” at all. In jurisprudence the very distortion frequently produces the insight; we often learn more from a caricature than from a photograph.

See also Dewey, John; Hart, Herbert Lionel Adolphus; Philosophy of Law, History of; Philosophy of Law, Problems of; Proudhon, Pierre-Joseph; Realism.

Bibliography

Fisher, William W., III, Morton J. Horwitz, and Thomas A. Reed. *American Legal Realism.* New York: Oxford University Press, 1993.

Frank, Jerome. *Law and the Modern Mind,* 2nd ed. New York: Coward-McCann, 1949.

Fuller, Lon. “American Legal Realism.” *University of Pennsylvania Law Review* 82 (1934): 429–462.

Gilmore, Grant. “Legal Realism: Its Cause and Cure.” *Yale Law Journal* 70 (1961): 1037–1048.

Horwitz, Morton J. *The Transformation of American Law, 1870–1960: The Crisis of Legal Orthodoxy.* New York: Oxford University Press, 1992.

Kalman, Laura. *Legal Realism at Yale: 1927–1960.* Chapel Hill: University of North Carolina Press, 1986.

Llewellyn, K. N. *Jurisprudence.* Chicago: University of Chicago Press, 1962.

Martin, Michael. *Legal Realism: American and Scandinavian.* New York: Lang, 1997.

Purcell, Edward A., Jr. “American Jurisprudence between the Wars: Legal Realism and the Crisis of Democratic Theory.” In *American Law and the Constitutional Order,* edited by Lawrence Friedman and Harry Scheiber. Cambridge, MA: Harvard University Press, 1978.

Twining, William. *Karl Llewellyn and the Realist Movement.* London: Weidenfeld and Nicholson, 1973.

Yosal Rogat (1967)
Bibliography updated by Philip Reed (2005)

LEGAL RESPONSIBILITY

See *Punishment; Responsibility, Moral and Legal*

LEHRER, KEITH
(1936–)

Keith Lehrer was born January 10, 1936, in Minneapolis, Minnesota. He attended the University of Minnesota from 1953 to 1957, earning his BA in Philosophy magna cum laude. His teachers at Minnesota included Alan Donagan, John Hospers, Michael Scriven, Mary Shaw, May Brodbeck, Herbert Feigel, and Wilfred Sellers. Lehrer went on earn his AM in 1959 and his PhD in 1960, both in philosophy, at Brown University, where his teachers included Roderick Chisholm, John Ladd, John Lenz, Stephan K(rner, Vincent Thomas, Wesley Salmon, and Richard Taylor. Chisholm supervised Lehrer's master's thesis on epistemology, and Taylor supervised Lehrer's doctoral dissertation on free will. Chisholm's and Taylor's continuing support were not due to Lehrer's agreement with their positions: Lehrer was then and continues to be a coherence theorist in epistemology, whereas Chisholm was a foundationalist, and Lehrer has always endorsed compatibilism, whereas Taylor was a libertarian.

Lehrer is best known for his work on free will, theory of knowledge, rational consensus, and the philosophy of Thomas Reid. His earliest philosophical works clearly reflect the ordinary language and common sense approaches to philosophy that he learned first from Hospers at Minnesota, and then through the influence of Reid, partly gained indirectly from Reid's influence on both Chisholm and Taylor. Lehrer's first published article (1960) was a common sense defense of the claim that humans can know they have free will simply through introspection.

Despite his lifelong commitment to compatibilism, many of Lehrer's earliest works were critical of various analyses of freedom intended to defend that view—particularly hypothetical analyses of freedom (e.g., that S is free to do X just in case S would do X if S tries to do X). Lehrer's argument against such analyses is that the conditional might apply to S, but S might lack some advantage necessary for exemplifying the antecedent of the conditional. So, for example, it could be true that S would do X if S tried to do X, but because of some phobia or other disadvantage, S could never actually try to do X. One may thus have control over external circumstances, but not have control over oneself, and such a disadvantage leaves one unfree.

His own first defenses of the compatibility of freedom and determinism were based upon a possible worlds analysis of freedom. His work in the late twentieth and early twenty-first centuries analyzes freedom in terms of a power preference that is a preference for having the preference structure one has concerning an action. To insure freedom, one must have that power preference because one prefers to have it. The preference for the preference structure must be the primary explanation of one's having it.

Lehrer is one of the best known proponents of a coherence theory of knowledge. On Lehrer's view, coherence consists in a cognitive system that is able to meet critical objections to the acceptance of a target proposition. Although his first analyses included a standard belief condition, Lehrer later argues that acceptance rather than belief should constitute the relevant condition, partly because the former involves a decision one makes. One's epistemic mission is to accept what is true and not to accept what is false. One cannot decide what to believe at a given moment; but one can decide what to accept in the pursuit of one's epistemic mission.

The ability of a background system to meet critical objections to the accepting of something one accepts provides personal justification. Lehrer first construed this background system as consisting only in states of acceptance designed to pursue the subject's epistemic mission; he expanded this view of the background system, which he later calls the "evaluation system," to include preferences and reasonings.

Much of Lehrer's earliest work in epistemology critiqued various attempts to solve the Gettier problem. The Gettier problem shows that one can have convincing justification of a true belief and yet not have knowledge because some part of the justification is false, where if that part were removed or replaced by the truth, one would no longer qualify as justified. Where such problems in justification exist, the justification is "defeated," and defeasibility theorists seek to solve the Gettier problem by formulating and explicating as a necessary condition the stipulation that one who knows has undefeated justification for what one knows. Defeasibility remains a central concern in Lehrer's most recent work in epistemology, *Theory of Knowledge* (2000), according to which knowledge is the product of true belief that is personally justified on the basis of coherence with the evaluation system, where such justification is undefeated. Undefeated justification, according to Lehrer, is a kind of justi-

fication that cannot be refuted by pointing out errors in the evaluation system (2000).

Lehrer has also offered a number of criticisms of recent "naturalistic" or externalist approaches to knowledge and justification, on the general ground that reliable cognitive mechanisms or ways of believing that track truth without cognitive self-evaluation are insufficient for knowledge. Lehrer's development of this element of his epistemology derives from his interest in the philosophy of Thomas Reid. Lehrer noted within Reid's system a metaprinciple according to which our faculties and the principles thereof are trustworthy. Lehrer applies this same principle to allow the knower to meet critical and skeptical objections, while also immunizing his own analysis of knowledge—despite its requirement for cognitive self-evaluation—against the KK-regress (namely, that one's knowing requires knowing that one knows, that one knows that one knows one knows, and so on ad infinitum) (1990, 2000).

The theory of rational consensus, which Lehrer developed with Carl Wagner (1981), was an attempt to incorporate a social component into the theory of rationality—another echo of Reid's common sense approach to philosophy. Social rationality, in Lehrer's and Wagner's theory, results from the evaluations people make of others, expressed mathematically as weights. They argue that under plausible conditions of evaluation social convergence would yield rational consensus. Lehrer went on to unify his work on justification and preference in *Self Trust* (1997), in which he sought to explain the trustworthiness of the self in terms of rationality, theoretical and practical as well as personal and social. In this and in his epistemology, Lehrer claims that complete explanation will contain a loop of the sort Lehrer first found in Reid's philosophy. There is a fundamental choice, according to Lehrer, between starting with unexplained first principles or, instead, maximizing explanation by including a principle of trustworthiness, which explains both why people are justified in accepting everything else that they accept and also why people are justified in accepting the principle itself. The effectiveness of the explanation depends on the wider system of explanation as well, and not simply on the "keystone principle" of self-trust.

See also Chisholm, Roderick; Coherentism; Determinism and Freedom; Epistemology; Epistemology, History of; Freedom; Reid, Thomas; Salmon, Wesley; Self; Sellars, Wilfrid.

Bibliography

Bogdan, Radu, ed. *Keith Lehrer*. Dordrecht, Netherlands: D. Reidel, 1980. Several essays by various authors on various aspects of Lehrer's philosophy, with a single essay in reply as well as a 101-page self-profile by Lehrer himself.

Brandl, Johannes, et al., eds. *Metamind, Knowledge, and Coherence: Essays on the Philosophy of Keith Lehrer*. Amsterdam: Rodopi, 1991. Several essays by various authors on various aspects of Lehrer's philosophy, each with a reply by Lehrer.

Lehrer, Keith. "Can We Know that We Have Free Will by Introspection?" *The Journal of Philosophy* 57 (1960): 145–157. A defense of freedom on the ground that human possession of freedom is entirely obvious from introspection.

Lehrer, Keith. *Metamind*. Oxford: Clarendon Press, 1990. A study of ways in which people's epistemic and preference assessments can also apply at metalevels, and to themselves.

Lehrer, Keith. *Self Trust: A Study of Reason, Knowledge and Autonomy*. Oxford: Clarendon Press, 1997. Brings together into one theoretical presentation most of Lehrer's views about knowledge, justification, rationality, and preference.

Lehrer, Keith. *Theory of Knowledge*. 2nd ed. Boulder, CO: Westview Press, 2000. Explicates and critiques several different conceptions of knowledge and justification, and argues for Lehrer's own version of the coherence theory.

Lehrer, Keith, and Carl Wagner. *Rational Consensus in Science and Society*. Dordrecht, Netherlands: D. Reidel, 1981. Argues that rational consensus can be reached by assigning weights to others' evaluations.

Olsson, Erik J., ed. *The Epistemology of Keith Lehrer*. Dordrecht, Netherlands: Kluwer, 2003. Essays by various authors with a reply by Lehrer.

Nicholas D. Smith (2005)

LEIBNIZ, GOTTFRIED WILHELM
(1646–1716)

The German polymath Gottfried Wilhelm Leibniz made significant contributions to philosophy, logic, mathematics, physics, jurisprudence, politics, the mechanical arts, and history. He worked as a diplomat, an engineer, an attorney, and a political advisor. He corresponded with queens and emperors and with the most eminent intellectuals of the age. Yet his reputation as a philosopher depends largely on texts that were unpublished at the time of his death, including some never intended for publication. Besides well-known works such as the *Discourse on Metaphysics*, *First Truths*, *New Essays*, and *Monadology*, there are thousands of pages of other texts, many of which are still unpublished. Interpreting these vast writings is a daunting task, best approached by attending closely to the historical and cultural context in which he

was working and by taking into consideration as many texts as possible. Against the background of Leibniz's long, complicated life, it is possible to trace the development of his philosophical views, from his earliest essays in Leipzig in the 1660s to his last letters written in Hanover fifty years later.

The sheer volume of Leibniz's writings, combined with the fact that some are published and some are not, can sometimes make citing Leibniz seem complicated. The standard edition of his works is *Gottfried Wilhelm Leibniz: Sämtliche Schriften und Briefe*, which was published by the Akademie Verlag in 1923. To simplify citations in this text, it is abbreviated throughout simply as "A," followed by series, volume, and page number (see "A" in the Abbreviations section of the Bibliography at the end of this essay for full publication information). When an English-language translation exists, it follows a colon at the end of the German-language information. In addition to the abbreviation for that primary work, other prominent texts on Leibniz's life and works have also been abbreviated in the in-text citations that appear throughout this essay—a list of those abbreviations and full publication information for every one of them is provided in the Abbreviations section at the very beginning of the Bibliography.

It should be noted that, in regard to Leibniz's philosophical texts, as of mid-2005, only those written up to June 1690 had been published; for texts written after that date that are referenced in this essay, the best available edition has been cited. Finally, works by Leibniz that are divided into short sections have been cited by section number instead of by page number.

LIFE

Leibniz was born in the Lutheran city of Leipzig on July 1, 1646 to Friedrich Leibniz (1597–1652), professor of moral philosophy at the University of Leipzig and the son of a noblewoman and his third wife, Catharina Schmuck (1621–1664), the daughter of a celebrated jurist. An orphan, Schmuck was raised by Johann Hopner, professor of theology, as well as by Quirinus Schacher, professor of law. Upon Friedrich's death in 1652, Schmuck committed herself to the education of her son and his sister, Anna Catharina (1648–1672). As a very young boy, Leibniz was given access to his father's library where by his own account he taught himself Latin and read poetry, history, theology, and some Aristotelian philosophy.

On graduating from the Nicolai School, Leibniz entered the University of Leipzig in April 1661, aged fourteen. He studied ancient languages and literature and heard lectures in mathematics (mainly Euclid) and philosophy. Although the new mechanical philosophy of René Descartes, Thomas Hobbes, and Pierre Gassendi had not been embraced by the professors in Leipzig, there was a diverse intellectual culture available there. Johann Adam Scherzer (1628–1683), professor of philosophy, Hebrew, and theology, published on a wide range of topics, including Kabbalistic theology while Jakob Thomasius (1622–1684) promoted an eclectic mixture of Platonism, Aristotelianism, and other prominent historical schools. Thomasius was an unusually careful historian of philosophy, keen to distinguish between the true and false proposals of the various philosophical *sects*. As the father of Christian Thomasius, who (with Christian Wolff) is often credited as founding the German enlightenment, Jakob Thomasius occupies an important place in the development of German philosophy. Thomasius supervised Leibniz's bachelor's thesis titled *Disputatio metaphysica de principio individui* (Metaphysical Disputation on the Principle of Individuation), which Leibniz defended and published in 1663. The thesis argues for a monadic account of substantial individuation, a position that prefigures his mature views.

Leibniz spent the summer of 1663 at the University of Jena studying under Erhard Weigel (1625–1699), professor of mathematics. Weigel was more progressive than the professors at Leipzig and included mechanical physics within his eclectic mixture, combining Euclid, Aristotle, and the *new philosophers* in an attempt to construct the true philosophy. He returned to Leipzig in October 1663 and received his bachelor of law degree in 1665 under professors Schacher and Bartholomäus Schwendendörffer. In 1666, he published *Dissertatio de arte combinatoria* (Dissertation on the Combinatorial Arts). It contains his first thoughts on the universal characteristic and related logical issues. He planned to pursue legal studies at Leipzig but was refused admission (probably because of his age) and went instead to the University of Altdorf, near Nuremberg, where he quickly earned a doctorate. His dissertation *Disputatio de casibus perplexis in jure* (Disputation on Difficult Cases in Law, 1668) was so well written and defended that the Altdorf faculty immediately offered him a professorship.

Leibniz declined the Altdorf professorship and chose, instead, a life of public service. In Mainz, he impressed Baron Johann Christian von Boineburg (1622–1672), a pious Catholic, distinguished diplomat, and minister to the archbishop of Mainz, Elector Johann Philipp von Schönborn (1605–1673). Boineburg became Leibniz's patron and employed him as an assistant, attor-

ney, librarian, and foreign advisor. In this last capacity, Leibniz produced a lengthy work supporting Schönborn's candidate for the Polish throne. The Catholic Boineburg encouraged the Lutheran Leibniz to pursue ecumenical and conciliatory projects, and he began a project, *Demonstrationes Catholicae* (*Catholic Demonstrations*), aimed at devising a metaphysics consistent with Catholic and Lutheran doctrines. He worked on the *Catholic Demonstrations* between 1668–1671 and returned to it in 1679. Although never completed, it contains his earliest essays on central metaphysical topics.

Besides pursuing peace in politics and religion, the young Leibniz was committed to philosophical peace. In an effort to offer a conciliatory method in philosophy, he prepared a new edition of Marius Nizolius' (1498–1576) 1553 work, *De veris principiis, et vera ratione philosophandi contra pseudophilosophos* (On true principles, and the true method of philosophizing against the false philosophers). Also, between 1669 and 1671, he composed a series of notes titled *Elementa juris naturalis* (*Elements of Natural Law*), in which he discusses theology, metaphysics, and ethics. These notes cover a wide range of topics, including divine and human justice, knowledge, and universal harmony. At this time he began a correspondence with the Duke of Brunswick Johann Friedrich (1652–1679), presenting his views about the souls or *vital* principles in nature, to which he attached important theological essays on the immortality of the soul and the resurrection of the body.

In 1671 Leibniz published two related works that constitute his first extended account of the laws of motion and their metaphysical foundations. The first, the *Hypothesis physica nova* (*New Physical Hypothesis*), subtitled *Theoria motus concreti* (*Theory of Concrete Motion*), he dedicated to the Royal Society of London; the second, the *Theoria motus abstracti* (*Theory of Abstract Motion*), he dedicated to the French Academy of Sciences. Together these works, which employ the Hobbesian notion of *conatus* along with the indivisibles of authors such as Bonaventura Cavalieri (c. 1598–1647), propose a physical system, including a creation story and laws of collision, which relies on the notion of momentary minds. Thus, by 1671 he had already begun to think of minds as the only source of motion and activity in the world; minds in nonhuman substances are momentary while human minds persist and have memory. This attempt to combine an original account of mind with a Hobbesian notion of *conatus* reveals his conciliatory tendencies as well as his capacity to engage in contemporary discussions in physics.

In 1671 Leibniz and Boineburg devised an elaborate plan to divert a pending European war. With secret papers in hand, Leibniz traveled to Paris in March 1672 to meet with a representative of King Louis XIV but arrived too late. Despite this failed diplomatic undertaking, he remained in Paris, at first to promote other political plans of his mentor and then, upon Boineburg's sudden death at the end of 1672, to pursue philosophical peace. He stayed in Paris until 1676 and struggled to stay longer, arguing that the pursuit of science in the service of humanity could be better achieved there than in Hanover, where the Duke of Brunswick had recently employed him.

Leibniz's four years in Paris were enormously productive. In the fall of 1672, he met the Dutch mathematician Christiaan Huygens (1629–1695) who immediately recognized the young man's talent and guided his mathematical studies. Although his education had not acquainted him with recent developments in mathematics, he devoted himself to study and by the fall of 1675 had laid the foundations of his calculus. During his lifetime he suffered from accusations that he had stolen the insights that led to his discovery of the differential and integral calculus from Isaac Newton. But twentieth-century historians of science exonerated him from these charges, showing that he arrived at the calculus independently of Newton.

In early 1673 Leibniz traveled briefly to England on a political mission and met mathematicians and natural philosophers, including Robert Hooke (1635–1703), Robert Boyle, and Henry Oldenburg (1619–1677), secretary of the Royal Society. Back in Paris, he finished a lengthy dialogue, *Confessio Philosophi* (*Philosopher's Confession*), in which he discusses the problem of evil, a topic that would engage him until his death. He also wrote an essay "De vera methodo philosophiae et theologiae ac de natura corporis" ("On the True Method in Philosophy and Theology and on the Nature of Body," in which he restates his fundamental methodological concerns and insists that neither mechanical physics nor mathematics speaks directly to what is most important, namely, the good of the soul and the truths of theology. In 1675 he designed and demonstrated a calculating machine and befriended the young mathematician Ehrenfried Walther von Tschirnhaus, who introduced him to the philosophy of Benedict (Baruch) de Spinoza. At the same time he began work on a group of notes, given the title *De Summa Rerum* (*On the Greatest of Things*), in which he discusses a diverse group of theological and metaphysical topics.

Partly due to prejudices against his religion and nationality, Leibniz failed to attain appropriate employment in Paris, and in 1676 he reluctantly accepted an offer from Johann Friedrich to serve as librarian and adviser at the court of Hanover. In October he traveled from Paris to London and Holland before proceeding to Hanover where he took up residency in December. During his journey he composed a dialogue, *Pacidius Philalethi Prima de Motu Philosophia* (Pacidius to Philalethes: A First Philosophy of Motion), which concerns the problem of the continuum and offers an account of motion. In London he met with Oldenburg again and also John Collins (1624–1683), mathematician and librarian of the Royal Society, who showed him some of Newton's papers. In Holland he met with prominent Dutch mathematicians and scientists, including the microscopists Jan Swammerdam (1637–1680) and Antoni van Leeuwenhoek (1632–1723). He talked at length with Spinoza and possibly saw a draft of Spinoza's *Ethics*

Settled in Hanover Leibniz continued to work in logic, metaphysics, theology, and mathematics. He met visiting scholars and theologians (including Nicolaus Steno [1631–1686]) and wrote a dialogue on free will, *Dialogue entre Poliandre and Théophile*. He took notes on Spinoza's *Ethics*, then newly published, corresponded with Nicolas Malebranche on metaphysics, and returned to the *Catholic Demonstrations* and his work on the universal characteristic. He studied chemistry and made detailed proposals to Johann Friedrich about administrative matters, including the expansion of mining in the Harz mountains. Besides technical tasks involved with the mines, he was much occupied in 1678–1679 with logical topics. He composed a series of highly original notes, given the title *Calculus Universalis* (*Universal Calculus*, in which he tries to formulate a logical calculus. Underlying this work is again his interest in methodology as a means of leading people to the truth and thereby effecting peace. Inspired by the multivolume *Encyclopedia* by Johann Heinrich Alsted (1588–1638), he planned his own encyclopedia project. Also during this time he made a breakthrough in his work on dynamics, defending the notion of force as against the Cartesian principle of conserved motion.

The sudden death of Johann Friedrich and the succession of his brother, Ernst August (1629–1698), in 1680, marked the end of this period of intense productivity. Leibniz remained on good terms with the duke and developed a close friendship with the duke's wife, Sophie, Duchess of Brunswick (1613–1714), with whom he corresponded on political, theological, and philosophical topics. The new duke, who would later become elector, encouraged Leibniz's technical and political schemes but was less receptive to academic matters and left the philosopher much less time to develop his own projects. Leibniz was assigned the burdensome task of compiling a history of the House of Brunswick, with the aim of establishing descent from the wealthy Italian house of Este. This project occupied him until his death (by which time, for all his efforts, he had only reached the year 1005).

Between 1680 and 1686 Leibniz worked primarily on logic and on the Harz mining project designing windmills and other equipment. When Leipzig professor Otto Mencke (1644–1707) began publishing a scholarly journal the *Acta Eruditorum*, with the aim of introducing new ideas to German scholars, Leibniz applauded the project and became a frequent contributor on scientific topics. During this time he began another attempt to formulate a logical calculus and renewed his work on the reconciliation of Protestantism and Catholicism. In that context he began a correspondence with Landgrave Ernst von Hessen-Rheinfels (1623–1699), a Catholic eager to promote religious peace.

Caught in a snowstorm for a few days in the Harz mountains in early 1686, Leibniz took advantage of the free time to compose one of his most famous works, the *Discours de métaphysique* (*Discourse on Metaphysics*). It represents his first attempt to summarize the main ideas of his philosophy. He asked Landgrave Ernst to send a synopsis to Antoine Arnauld, and thus began one of the most interesting philosophical correspondences of the seventeenth, or any other, century. Arnauld's criticisms forced Leibniz to explain and expand upon some of his most fundamental ideas.

Leibniz was disappointed when the duke abandoned the Harz mining project but immediately began planning a trip to research the history of the House of Brunswick. In October 1687 he set out on an extended tour of the southern German states, Austria, and Italy. His official duty was to research family history; his personal desire, encouraged by Landgrave Ernst, was to promote religious and political peace. He visited public archives and personal libraries and talked with politicians, monks, and cardinals. During his residence in Vienna, he met the Austrian emperor, to whom he recommended, among other things, the reorganization of the economy, the formation of a general research library, and the establishment of an insurance fund; he worked on proposals for an Imperial College of History; for reforming the coinage of Austria, Brunswick, and Saxony; and for lighting the streets of the

city. And he wrote an important paper on motion later published in the *Acta Eruditorum*.

Leibniz spent a year in Italy traveling as far south as Naples and meeting with prominent intellectuals along the way. In Rome (April—November 1689), he made contact with leading Italian scientists, Jesuits (including Claudio Grimaldi [1638–1712], who had lived in China and with whom Leibniz corresponded), and Jansenists. Visits to the Physical-Mathematical Academy led to a treatise on dynamics, *Dynamica de potentia et legibus naturae corporeae* (Dynamics: Concerning the force and laws of natural bodies), which has two parts, one on abstract and the other on concrete dynamics. In Modena he arranged a marriage between the House of Modena and one of Duke Friedrich's daughters. In Venice he met the scientist and Jesuit Michel Angelo Fardella (1650–1708), with whom he later corresponded on philosophical topics.

Before leaving Italy Leibniz wrote a long (last) letter to Arnauld in which he develops further details of his metaphysics. At about the same time, he composed one of his most well-known texts, *Primae Veritates* (*First Truths*). Written on Italian paper, the paper (given the title *Principia Logico-Metaphysica* by the academy editors) dates from the time during—or soon after—his trip to Italy. The four-page essay is a neat summary of his most fundamental philosophical principles, which are outlined in a form interestingly different from previous presentations. Leaving Venice in March 1690, he traveled through Vienna, Prague, Leipzig, and other cities before returning to Hanover. In Vienna he wrote an important paper on motion and gravity titled *De causa gravitatis, et defensio sententiae auctoris de veris naturae legibus contra cartesianos* (On the cause of gravity), which was published in the *Acta Eruditorum* in May. When he arrived back in June 1690, he had been away for more than two and a half years.

Upon his return Leibniz felt the need to justify his lengthy and relatively expensive trip and so committed a good deal of time to his history of the House of Brunswick. In 1690 he became director of the ducal library in Wolfenbüttel, a position that he held for the rest of his life. During the early 1690s he maintained his close relationship with Sophie, by this time Electress of Hanover, published often (especially on mathematical and dynamical topics) in the German *Acta Eruditorum* and the French *Journal des Sçavans*, continued old correspondences, and began new ones (for example, with Johann Bernoulli [1667–1748]). His relations with members of the Royal Society, which had never been unproblematic, took an unfortunate turn when he was accused of using Newton's work as the basis for his own calculus. In March 1693 he wrote directly to Newton about the topic.

In the 1690s Leibniz exchanged several letters with Paul Pellisson-Fontanier (1624–1693), which were then shared with interested parties, including Sophie and her Catholic sister, Marie de Brinon. These letters addressed differences between Catholic and Protestant theology and the possibility of unification among the churches. The well-known physician, Kabbalist, and Quaker sympathizer, Francis Mercury van Helmont (1614–1698) visited Hanover and spent several days lecturing Leibniz and Sophie about his views. Becoming more and more fascinated with reports from Jesuits in China about the science and mathematics of that culture, Leibniz published *Novissima Sinica* (Latest news from China) in 1697, which is an edition of letters and reports from the Jesuit's mission there. For Leibniz the reports from China supported his assumption that there is an underlying truth that all people seek and that could be glimpsed, regardless of religion.

At each stage of his life, Leibniz worked on many diverse projects and wrote thousands of notes on philosophy, mathematics, science, and theology. As an intellect he was in constant motion. It is therefore striking that he published so little. After the texts of 1670–1671, he did not publish a general account of his views until 1695 when his *Système nouveau de la nature et des la communication de substances, aussi bien que de l'union qu'il y a entre l'âme et le corps* (New system of nature), a relatively brief account of a part of his metaphysics, appeared in the French *Journal des Savants*. This led to discussions with prominent Cartesians and others, including Simon Foucher and Basnage de Beauval (1692–1708).

In 1695 Leibniz was promoted to privy counselor of justice, a high-ranking position at court. However, he was not entirely content, complaining that he had little time for new ideas and projects and that, apart from Electress Sophie, there was no one with whom he could discuss intellectual matters. Ernst August died in early 1698 and was succeeded by his eldest son, Georg Ludwig (1660–1727) (later George I of England). Georg Ludwig had little patience either for Leibniz's slow progress on the history of the House of Brunswick or for his other *invisible* projects, and Leibniz received less financial support and freedom of movement. But his friendship with Sophie continued, and his relations with her daughter, Sophie Charlotte, Electress of Brandenburg (and soon to be Queen of Prussia) also became close. Sophie Charlotte

often asked Leibniz to act on her behalf, and she supported him in his successful attempt to set up the Berlin Society of Sciences in 1700. As founding president, he wrote its charter.

At this point Leibniz was ready to publish further details of his system of *preestablished harmony*. One of the most important accounts, *De Ipsa Natura* (On Nature Itself), appeared in *Acta Eruditorum* in 1698 and contains his first use of the term *monad*. These publications led to important intellectual exchanges with Pierre Bayle, Burchard de Volder (1643–1709), Lady Damaris Masham (1658–1708), Bernoulli, Bernard le Bovier de Fontenelle, Bartholomew des Bosses (1668-1728), Wolff (who became a kind of disciple), and others.

In the final years of the seventeenth century, Leibniz engaged again in controversy over the invention of the calculus. He was also drawn into secret diplomacy with the English court over the royal succession. Sophie Charlotte and he frequently conversed and exchanged letters about political and philosophical matters. After her sudden death in 1705, he wrote a memorial on topics they had discussed, which subsequently became his *Essais de Théodicée* (*Theodicy*), dedicated to her. Published in 1710, the *Theodicy* is the longest and most prominent publication of his life. In it he attempts to reconcile the goodness of God, the freedom of human kind, and the origin of evil. Its central claim, that this is the best of all possible worlds, was subsequently ridiculed by François-Marie Arouet de Voltaire in *Candide*.

By 1705 Georg Ludwig had lost all patience with Leibniz and forbade him to leave Hanover without permission until the history of the House of Brunswick was complete. Besides visits to nearby Wolfenbüttel, he spent much of his time over the next few years on the history and political relations among the courts in England, Hanover, and Brandenburg. But despite these duties, he began a study of John Locke's *Essay concerning Human Understanding* and wrote essays, some of which he published, on philosophy. As a result of his critical respect for Locke, he composed a lengthy dialogue between a Lockean and a Leibnizian but chose not to publish this text, *Nouveaux essays sur l'entendement humain* (New Essays on Human Understanding) because Locke died in 1704, around the time the work was finished.

In his last years Leibniz continued as librarian of Wolfenbüttel, political adviser, and historian. In 1711 he met Russian Czar Peter the Great (1672–1725) who wanted to engage him on legal and scientific matters. In 1713 Leibniz traveled to Vienna where the Austrian emperor appointed him imperial privy counselor and agreed to create a Society of Sciences. From Vienna he counseled friends in Hanover and Wolfenbüttel though dislike for him at court had increased. When Georg Ludwig became King George I of England in 1714, Leibniz returned to Hanover in hopes of seeing his employer. They missed one another, but the king left instructions insisting that the history of the House of Brunswick be finished. Despite these pressures and encroaching ill health, he began new correspondences—with Nicolas Remond in Paris and Samuel Clarke, an English Newtonian. He also wrote *Principes de la nature et de la grace, fondés en raison* (The Principles of nature and grace, based on reason); the *Discours sur la théologie naturelle des Chinois* (Discourse on the natural theology of the Chinese), in which he shows the connections between Chinese thought and his own *true* philosophy; and the *Monadology*, perhaps his most famous work, providing a summary of the basic tenets of his later philosophy.

Leibniz suffered from gout and by 1714 was severely affected. In the last months of his life, he developed sores on his right leg. Distrusting physicians he refused to see a doctor when he suffered an attack of kidney stones. Working constantly he died in bed on 14 November 1716. By this time he was so out of favor with the court that only a handful of people attended his funeral. Because few of his works were published during his lifetime, it was only in the later part of the eighteenth-century that the extent of his genius began to be understood and acknowledged. It would be left to twentieth-century scholars to uncover the extraordinary breadth of his contributions in physics, mathematics, logic, theology, and philosophy.

PHILOSOPHICAL CORPUS

Among the writings of great early modern thinkers, Leibniz's are unusually problematic. Descartes, Galileo, Spinoza, Hobbes, Malebranche all produced brilliant explications of their philosophies. But there is no single exposition of Leibniz's metaphysics replete with extended arguments and details. He published little during his lifetime and no published text (e.g., *A New System of Nature*, the *Theodicy*) provides a thorough-going account of his philosophy. Although there are a number of identifiable *main* texts, it remains unclear how to treat them since they differ noticeably from each other and were written over many years.

Leibniz wrote more pages—in Latin, French, and German—than most scholars can read in a lifetime. Stored in Hanover after his death, his papers were unorganized, unedited, and undated. The main part of his philosophical corpus has not been available in a standard

edition. The early editions of his philosophical work—a late eighteenth-century edition by L. Dutens and a late nineteenth-century collection by C. I. Gerhardt—are incomplete and sometimes inaccurate. The Prussian Academy of Science (now the German Academy of Science) began to publish the standard edition of his papers in 1923, but it has taken decades to cover even the main works in philosophy. The publication of the remainder is expected to take until 2050. It is surely difficult to acquire a broad understanding of his writings when only a small selection is available.

Leibniz's philosophical writings pose additional problems. First, many of them are hastily written personal notes, often both incomplete and undated. As he himself wrote about his papers: "Instead of treasure … you will only find ashes; instead of elaborate works, a few sheets of paper and some poorly expressed vestiges of hasty reflections, which were only saved for the sake of my memory" (A VI i 533). Second, even in the publications and letters sent to the great philosophers of Europe, he had specific methodological reasons for not being forthright about his views: His goal was to avoid preaching in an attempt to engage his reader. By such means he hoped to nudge the wayward soul toward the truth. In a frank moment in 1676 he writes: "A metaphysics should be written with accurate definitions and demonstrations, but nothing should be demonstrated in it apart from that which does not clash too much with received opinions. For in that way this metaphysics can be accepted; and once it has been approved then, if people examine it more deeply later, they themselves will draw the necessary consequences" (A VI iii 573–574: Pk 93). Finally, given his astonishing erudition, it is difficult to reconstruct the conceptual framework of his writings. Not only did he use major parts of the history of philosophy without citation or explanation, he thought that it was a *good thing* to combine ideas taken from the great philosophical systems. One of the main reasons that it is so difficult for us to recognize the borrowed doctrines and transformed assumptions in his writings is that he made such abundant use of the entire history of philosophy as it was conceived in the seventeenth century.

Due to the difficulties posed by Leibniz's writings, texts such as the *Discourse on Metaphysics*, *First Truths*, *New System*, *New Essays*, and *Monadology*—all of which suited twentieth-century philosophical tastes—became his canonical writings. As important as these writings are, they do not represent the extraordinary range and quirky diversity of his ideas. He is rarely explicit about the precise relations among his ideas, but he is clear about the fact that they are tightly connected. At the end of his life, he insists: "My principles are such that they can hardly be torn apart … whoever knows one well knows them all" (G II 412: L 599).

In an attempt to reveal the breadth of Leibniz's philosophical system and the connections among core doctrines, this article cites a diverse group of texts selected from all the main periods of his life. He borrowed ideas from the whole history of philosophy, and so before considering some of his philosophical ideas, we will situate them in their proper historical context.

METHODOLOGY AND INTELLECTUAL HARMONY

The early Renaissance philosopher Giovanni Pico della Mirandola, formulates in his *On the Dignity of Man* (1486) one of the defining statements of the conciliatory methodology of many humanist thinkers. Pico recommends that the seeker of truth study all the masters of philosophy. Once the truths in each philosophical tradition are discovered, they will be combined into a comprehensive philosophy. One of the main points of Pico's project is to show that *a concord* can be forged between the philosophies of Plato and Aristotle. For intellectual conciliators such as Pico, the doctrines of the prominent philosophical traditions, despite their apparent differences, can be made to form a coherent philosophical system.

In the aftermath of the Thirty Years War, whose battles were fought mostly on German soil, this methodology of peace was extremely attractive, especially to German thinkers, many of whom had witnessed the devastation and horrors of the war firsthand. As a young man Leibniz committed himself to his own form of conciliatory eclecticism. Like Pico he thought that the fundamental truths were (mostly) those offered by the illustrious ancient thinkers. Some of his basic metaphysical beliefs were taken directly from the Aristotelian, Platonist, and mechanical philosophies: that a substance is something wholly self-sufficient, that each creature is an emanation of God's essence, and that all corporeal features are to be explained mechanically. But he also went beyond Pico in his commitment to a philosophy that is consistent with specific Christian doctrines, such as those of the Eucharist and the resurrection of the body. His grand philosophical system is the result of the clever interweaving of ancient and modern assumptions.

In 1671 Leibniz published an edition of a text by the sixteenth-century humanist Mario Nizolio (1488–1567). He wrote a lengthy introduction to Nizolio's 1553 book

On the True Principles and the True Method of Philosophy, Against the Pseudo-philosophers. Both Nizolio's text and Leibniz's introduction discuss the proper way of philosophizing. It is significant that Leibniz attached to his introduction a slightly revised version of his April 1669 letter to Jakob Thomasius. The letter thereby became the young man's first published text on a contemporary metaphysical topic. Instead of being yet another philosopher "lusting for novelty," Leibniz seeks to find the "interconnections among doctrines" (A VI ii 426). He presents what he calls a "reformed philosophy," a philosophy that combines the "rule" of the new mechanical physics and the metaphysics of Aristotle (A VI II 434: L 94). He focuses on corporeal substances and reforms Aristotle's notions of substantial form and matter so that they accommodate the mechanical physics. By demoting the mechanical notion of matter as extended stuff to Aristotelian prime matter, he cleverly constructs a theory of substance that has the structure of the Aristotelian notion and yet is consistent with mechanical explanations in physics. He happily concludes that by such means, the mechanical philosophy "can be reconciled with Aristotle's" (A VI ii 435: L 95). The details of his views about substance would change over the years, but the basic structure of this theory of substance, developed as a synthesis of Aristotelianism and mechanism, would remain the same.

In his *New Essays*, written in response to Locke in the early years of the eighteenth century, Leibniz reflects on the methodology that produced his philosophy: "This system appears to unite Plato with Democritus, Aristotle with Descartes, the scholastics with the moderns, theology and morality with reason. It seems to take the best from all quarters and then goes further than anyone has done before" (A VI vi 71–73). His concern with intellectual harmony emerges also in his concern to engage his readers and interlocutors so as to enlist them in his march toward truth. In a letter of March 1678, he explains:

> I am concerned, as are all who wish to hold a middle ground, not to seem too much inclined toward either of the two opposed adversaries. Whenever I discuss matters with the Cartesians … I extol Aristotle where he deserves it and undertake a defense of the ancient philosophy, because I see that many Cartesians read their one master only … and thus unwisely impose limits on their own ability. … I think that the two philosophies should be combined and that where the old leaves off, the new should begin."
> (A II i 402: L 190)

For Leibniz the true metaphysics will be consistent with Christian doctrine and constructed from the underlying truths in the great philosophical systems. An underappreciated aspect of his brilliance is his ability to gather ideas from different philosophical sources and make them his own.

GOD AND CREATION

Like other prominent thinkers of the seventeenth century, Leibniz believed in a perfectly good Supreme Being who created and maintained the world and whose existence could be proven. He sometimes employed versions of the cosmological argument for God's existence. For example, in the *Monadology* (1714), he argues for God a posteriori based on the harmonized diversity of the world and the fact that there are contingent beings whose "final or sufficient reason" must be in a "necessary being" (§39, §45). But his favorite argument is an original version of the ontological argument, which is critical of Descartes's version and based on the mere possibility of God: "Since nothing can prevent the possibility of what is without limits, without negation, and consequently without contradiction, this by itself is sufficient for us to know the existence of God a priori" (§45).

Like many of his contemporaries, Leibniz owed a number of his assumptions about God as creator of the world to an ancient (mostly Platonist) tradition. From prominent professors at the University of Leipzig, he acquired a solid education in Platonism. The version of this ancient philosophical "sect" taught in Leipzig was one inspired by the third-century Platonist, Plotinus (c. 204–270) and by Jewish Kabbalism. Many of his most fundamental assumptions about knowledge, mind, plenitude, the nature of creation, and the relations among substances are rooted in this tradition. Two assumptions that he embraced as a young man are as follows:

GOD AND EMANATION. There is an ultimately good, perfectly self-sufficient, and thoroughly unified Supreme Being on which everything else depends and which itself depends on nothing. God's mind contains a number of Ideas or attributes (say, the Idea of Justice), which are the perfect essences of things (these are roughly based on Plato's theory of Ideas) and which are used as models for created things. The Idea or attribute of God is emanated to a creature in such a way that neither God nor God's attribute is depleted in any way while the creature acquires the attribute, though in an inferior manner. The emanative process is continual so that a creature instantiates a divine attribute if and only if God emanates the

attribute to the creature. For many Platonists, a corollary of this causal theory of emanation is that every product of the Supreme Being contains all the attributes (and hence the essence) of God though the product instantiates each of those attributes in a manner inferior to the way in which they exist in the Supreme Being. Justice as conceived by God is perfect; justice as instantiated by Socrates is not. Leibniz summarizes the position in §14 of the *Discourse on Metaphysics*: "It is evident that created substances depend upon God, who preserves them and who even produces them continually by a kind of emanation."

PLENITUDE AND SYMPATHY. The divine essence is emanated not just to each creature but to the whole of creation. The principle of plenitude develops from the idea that the more of the divine essence in the world—and hence of being and goodness—the better. Although the principle of plenitude suggests that there will be as much diverse being as possible (the more being, the better the world), this diversity of being must also be properly unified (the more unity, the better the world). One of the results of this unity among the parts of the world is a cosmic sympathy. Here the idea is that each part of the world is *in sympathy* with all the others. In other words, the principle of plenitude was supposed to imply that God fills creation with as much being as possible and unifies those diverse beings as much as possible. Such a diverse and unified world was supposed to engender wonder, delight, and awe in human observers. In the *Monadology*, he agrees with the ancient philosopher Hippocrates who claimed that all things are in sympathy with one another: everything "is affected by anything that happens in the universe, to such an extent that he who sees all can read in each thing what happens everywhere, and even what has happened or what will happen, by observing in the present what is remote in time as well as in space" (§61).

These ancient Platonist assumptions about emanation, plenitude, and sympathy inform much of Leibniz's thinking about the world. They inspire his theory of universal harmony, many of his views about mind, his account of knowledge, his solution to the problem of evil, and his views about the mirroring and expressing of substance. In this section, we consider the core doctrines closely related to these Platonist assumptions. As we will see, Leibniz remains committed to these doctrines throughout his philosophical career.

UNIVERSAL HARMONY. Leibniz first articulates the doctrine of universal harmony in a series of notes titled *Elements of Natural Law*, written between 1668 and 1671. As he summarizes the idea for Arnauld in 1671: "I define … harmony as diversity compensated by identity" (A II i 173–174: L 150). By the time he wrote the *Discourse on Metaphysics* in 1686, he had come to formulate the doctrine in terms of hypotheses though the underlying idea is still the same. In §6 of the *Discourse*, he explains: "God has chosen the most perfect world, that is, the one which is at the same time the simplest in hypotheses and the richest in phenomena." According to Leibniz, the single, unified, and perfect Supreme Being freely chooses to emanate the divine attributes to creatures; God remains transcendent while all creatures become an imperfect instantiation of God's attributes. Because God emanates the divine essence to all its products, he describes God as the reason (*ratio*) of the world and the one (*unum*) in it.

Universal harmony entails that God relates to the world and to each creature in it in two ways. God is the multiplicity in the world insofar as the divine essence is variously manifested in the vast diversity of creatures and in the diversity of the perceptions of each creature, but God is also the unity insofar as each created thing is a unified instantiation of the divine essence (although a manifestation of the essence far inferior to that of God) and therefore related to and reflective of all the others. The world is full of various perceptions of the world or *phenomena* because the world contains an infinity of different expressions of the divine essence. Leibniz's notion of universal harmony forms the basis for his mature theory of pre-established harmony.

PLENITUDE, DIFFERENCE, AND PRINCIPLE OF THE IDENTITY OF INDISCERNIBLES. From 1676 on Leibniz is increasingly explicit about the significance of the principle of plenitude. In a series of notes written in Paris titled *On the Greatest of Things*, he writes: "I take as a principle … that the greatest amount of essence that can exist does exist" (A VI iii 472: Pk 21). He never wavers from this commitment to plenitude. In *On the Ultimate Origination of Things* of 1697, he explains that God is the *reason*, or source, of things and argues that "there is a certain urge for existence or (so to speak) a straining toward existence in possible things or in the possibility of essence itself; in a word essence in and of itself strives for existence" (G VIII 303: AG 150). For Leibniz, the world is not just very full, it is as full of being as it can possibly be, consistent with harmony. As for his contemporaries Spinoza and Anne Conway, infinity is for Leibniz a mark of the fullness of being. Whereas Spinoza assigns God or nature an infinity of attributes, both Conway and Leibniz make each portion of the world infinitely full. In 1676 he claims

that every part of the world, regardless of how small, "contains an infinity of creatures," which is itself a kind of "world" (A VI iii 474: Pk 25). He emphasizes the same point later in *First Truths* (1689): "Every particle of the universe contains a world of an infinity of creatures" (VI iv [B] 1647–1648: AG 34). For Leibniz there is an aesthetic aspect to this elaborate harmony among the infinity of creatures. As he puts the point in the *Monadology*:

> the author of nature has been able to practice this divine and infinitely marvelous art, because each portion of matter is not only divisible to infinity, as the ancients have recognized, but is also actually subdivided without end, each part divided into parts …; otherwise, it would be impossible for each portion of matter to express the whole universe" (§65).

Nor is Leibniz content merely to fill the world with being. He argues that in order to contribute to the world's diversity, each created thing must be essentially distinct from every other. One of his most famous principles, the principle of the identity of indiscernibles, demands that no two substances are exactly alike. He writes in *Discourse*: "It is not true that two substances can resemble each other completely and differ only in number" (§9). Although he is not explicit about the importance of the principle until the late 1680s and then formulates it in a variety of ways, the basic idea is straightforward enough: There is always more than a mere numerical difference between substances. Two eggs might seem perfectly similar but they will not differ merely numerically; there will always be something true of one egg that is not true of the other. In *First Truths* he argues: "In nature, there cannot be two individual things that differ in number alone. For it certainly must be possible to explain why they are different, and that explanation must derive from some difference they contain" (A VI iv [B] 1645: AG 32). As he puts it in the *Monadology*: "It is also necessary that each monad be different from each other. For there are never two beings in nature that are perfectly alike, two beings in which it is not possible to discover an internal difference" (§9).

What the principle of the identity of indiscernibles claims is fairly clear; why he wanted to make such a claim is less so. His commitment to the principle of plenitude and theory of emanation offers insight into his underlying motivation. For Leibniz, as for many theists, the goodness of the world is a function of the diversity of beings as well as the order among them. Given that each creature contains the divine essence, the world will be better if it is as full of diverse emanations of the divine nature as is consistent with unity and harmony. His principle of the identity of indiscernibles pushes this intuition to its logical extreme: By demanding that no two substances (that is, no two emanations of the divine essence) be the same, he thereby increases the amount of diversity in the world. The principle of the identity of indiscernibles is a neat way of insisting on difference of the required sort.

MIRRORS AND EXPRESSIONS. The image of the mind as a mirror is a permanent fixture of Leibniz's mature thought. He first develops this idea in the *Elements of Natural Law* (1668–1671). Consider the following passage: "Since every mind is like a mirror, there will be one mirror in our mind, another in other minds. Thus, if there are many mirrors, that is, many minds recognizing our goods, there will be a greater light, the mirrors blending the light not only in the [individual] eye but also among each other. The gathered splendor produces glory" (A VI i 464: L 137). By such means, he goes beyond the plenitude and sympathy of his Platonist predecessors. He does not just maximize creatures and the assumed sympathetic relations among them, he heightens their connections by making each substance a mirror of all the others because each mind is (unconsciously) aware of all the others.

In the notes written in Paris in 1676, he develops his growing commitment to plenitude in a number of directions. For Leibniz, in *On the Greatest of Things*, each mind eternally mirrors the entirety of the world, and each does so from its own perspective. That is, consistent with the principle of the identity of indiscernibles, no two substances mirror the world from the same perspective. To elucidate his point he offers an analogy that he will use for the rest of his philosophical career: In the same way that travelers approaching a town from different directions see the town from different perspectives, so each mind approaches the world from a different perspective. For Leibniz it is important that each mind has a unique view of the world for "in this way a wonderful variety arises" (A VI iii 524: Pk 85). As he summarizes the point in *On the Greatest of Things* in 1676: "A most perfect being is one that contains the most. Such a being is capable of ideas and thoughts, for this multiplies the varieties of things, like a mirror"(A VI iii 475: Pk 29).

Forty years later Leibniz sets out the same claims, employing the same analogies, in the *Monadology*: "This interconnection or accommodation of all created things to each other, and each to all the others, brings it about that each simple substance has relations that express all

the others, and consequently, that each simple substance is a perpetual, living mirror of the universe" (§56).

> Just as the same city viewed from different directions appears entirely different and, as it were, multiplied perspectively, in just the same way it happens that, because of the infinite multitude of simple substances, there are, as it were, just as many different universes, which are, nevertheless, only perspectives on a single one, corresponding to the different points of view of each monad. … And this is the way of obtaining as much variety as possible, but with the greatest order possible, that is, it is the way of obtaining as much perfection as possible. (§57–58)

As these quotations suggest, there are close connections between the mirroring activity of minds and Leibniz's mature doctrine of expression. In various texts and in various ways, he claims that each substance expresses God, each substance expresses the world, and each substance expresses every other substance. After years of analysis of the texts, scholars have remained unclear about the implications and interconnections of these claims and about how exactly the doctrine of expression relates to the idea of minds as mirrors. The 1676 Paris notes, *On the Greatest of Things*, help solve some of the most recalcitrant problems by revealing the underlying motivation behind the doctrine. Each substance is an emanation of God's essence, and in this sense each shares the same essence. Each emanation will differ from every other by *expressing* the divine essence differently: "The essence of all things is the same," and they differ "only in the manner of their expression" (A VI iii 573: Pk 95). To explain his point he compares the essence of God to a number that can be expressed in an infinity of ways, each of which is a more or less clear expression of the essence. For the number 6, whether the expression is 3+3, 3×2, or 4+2, each is an expression of the same thing although "no one can doubt that the one expression differs from the other" (A VI iii 518: Pk 77). In the same way that the number 6 may be thought to contain its full essence, so God contains perfectly the divine essence. Whether the expression of 6 is 2+4, 3×2, 36–32+2, or any of the other infinite means of expressing it, each is a more or less clear expression of the same thing. Similarly, each substance—whether a human, roach, or chimpanzee—is a more or less clear expression of the divine essence. Leibniz concludes: "So do things differ from each other and from God" (A VI iii 519: Pk 77).

The arithmetical analogy makes it easier to see how expression works. Each substance expresses God insofar as it expresses the divine essence; each expresses the world insofar as the world just is the totality of expressions of God; and finally, each substance expresses every other insofar as each is a more or less clear expression of the same thing. The *Discourse on Metaphysics* employs expression to great effect: "Every substance is like a complete world and like a mirror of God or of the whole universe, which each one expresses in its own way, somewhat as the same city is variously represented depending upon the different positions from which it is viewed" (§9). He goes on to add that substances are "different expressions of the same universal cause, namely, God," where "the expressions vary in perfection" (§15).

Nor should we worry that creatures have become "little Gods." Although in the *Monadology* Leibniz is happy to describe human minds as "images of the divinity itself" (§83), he always distinguishes between the perfection of God and the limitations of creatures. In the *Monadology*, he insists that "what is limited in us is limitless" in God (§30), and argues: "God alone is the primitive unity or the first simple substance; all created or derivative monads are products, and are generated, so to speak, by continual fulgurations of the divinity from moment to moment, limited by the receptivity of the creature, to which it is essential to be limited" (§47).

GOD, MIND, AND KNOWLEDGE

The Platonism of Leibniz's professors bequeathed to him central concerns relating to mind. In the *Phaedo* Plato argues that it is "the divine-like" nature of the soul that guarantees its self-sufficiency, vitality, and unity. Because the soul remains "always the same as itself," it is immortal. The body, because it is never the "same as itself," is mortal (80a–e). Subsequent Platonists had to explain how the soul and the body could be causally related. Among the explanatory alternatives, the fifteenth-century Platonist Marsilio Ficino offered a version of one that influenced Leibniz strongly. In his *Platonic Theology*, Ficino uses the causal theory of emanation to bind the body to the soul. According to Ficino, the soul, which is "always alive," emanates its "vivifying" and "indivisible power" to its body so that it "causes life to be diffused" and thereby creates a harmony of components. As the unifying power of God is to the world, so is the soul to the body (Book II, chapter 3).

Besides a Platonist account of the soul and its relation to the body, the young Leibniz also took up a Platonist epistemology according to which the only true objects of knowledge (as opposed to opinion) are the eternal and immutable Ideas. Many Platonists placed the

Ideas within the soul, where they remain, waiting to become objects of conscious thought. Although Platonists differed about the precise role played by the senses in the acquisition of knowledge, most agreed that the process of coming to know the Ideas was one of removing oneself from the mutable world of the senses and letting one's understanding (*intellectus*) grasp the immutable Ideas within. For some Platonists cosmic sympathy aids in this pursuit of knowledge; the same Ideas that are implanted in souls are also evident in the harmony among creatures in the world. Theists often reinterpreted Plato's realm of Ideas as the mind of God and the Ideas as paradigms employed by God in creation. Acquisition of knowledge of these Ideas is a necessary step toward knowledge of God, to be achieved both by turning away from the world to the immutable ideas within and by attending to the connections among all things.

In some notes written during his stay in Venice in 1690, Leibniz summarizes this Platonist stance: "Each thing is so connected to the whole universe, and one mode of each thing contains such order and consideration with respect to the individual modes of other things, that in any given thing, indeed in each and every mode of any given thing, God clearly and distinctly sees the universe as implied and inscribed." Due to this connection among things:

> "when I perceive one thing or one mode of a thing, I always perceive the whole universe confusedly; and the more perfectly I perceive one thing, the better I come to know many properties of other things from it. And from this perfect consonance of things there also arises the greatest harmony and beauty of the universe, which exhibits to us the power and wisdom of the Highest Maker." (AG 103)

MIND AND ACTIVITY. From the beginning of his philosophical career, Leibniz associates activity with mind. Whether he calls these principles of activity souls, minds, substantial forms, or monads, the idea is always that the only sources of activity in the world are divine-like principles that have the power to generate unity, self-sufficiency, and vitality. In a note of 1671, he argues: "Just as God thinks things … because they follow from his nature, so does Mind. … Mind and God do not differ except that one is finite and the other infinite" (A VI ii 287–288). In the *Monadology*, he notes: "that souls, in general, are living mirrors or images of the universe of creatures, but that minds are also images of the divinity itself, or of the author of nature, capable of knowing the system of the universe … each mind being like a little divinity in its own realm" (§83).

For a short period in 1670–1671, Leibniz distinguished between the momentary minds in nature and conscious minds. His published treatises the *New Physical Hypothesis* and *Theory of Abstract Motion* of 1671 employ momentary minds as the cause of the motion in bodies to great effect. By 1676 his commitment to the plenitude has led him to make all minds eternal: "Every mind is of endless duration" and "is indissolubly implanted in matter. …There are innumerable minds everywhere" which "do not perish" (A VI iii 476–477: Pk 31). In *On the Greatest of Things* minds act constantly and constitute self-sufficient beings that are eternal and indestructible by anything but God. Human minds are created by God and then exist eternally. Nonhuman minds exist from the beginning of the world to its end. Despite appearances to the contrary, Fido the dog does not die but shrinks down to *an invisible core* of substance from which it activates another substance, and so on for all of eternity. This remained Leibniz's view: "There is never total generation nor, strictly speaking, perfect death, death consisting in the separation of the soul. And what we call *generations* are developments and growths, as what we call deaths are enfoldings and diminutions" (*Monadology* §73).

MARKS AND TRACES. The eternity of all mind-like active things is not an obviously plausible theory. Leibniz endorsed it because the eternity of minds adds significantly to the plenitude and harmony of the world. While developing his opinions about plenitude in *On the Greatest of Things*, he hit upon the idea that each mind-like creature eternally perceives the entirety of the world. Each mind "senses all the endeavors" of all the other minds in the whole history of the world; "no endeavor in the universe is lost; they are stored up in the mind, not destroyed" (A VI iii 393: Pk 47). He came to believe that plenitude requires that each moment in the eternity of the world contain its whole history: past, present, and future. Minds not only sense all the present activities of all the minds in the world, they also retain a memory or *trace* of them: "It is not credible that the effect of all perceptions should vanish" (A VI iii 510: Pk 61). Each mind "retains the effect of what precedes it" and also "has a quality of such a kind as to bring this [state or effect] about" (A VI iii 491: Pk 51).

Thus, in 1676 Leibniz develops a version of his doctrine of marks and traces according to which each mind at every moment includes an effect or trace of all it has done as well as a quality or mark of all it will do. In §8 of

the *Discourse*, he offers the soul of Alexander as an example: "There are vestiges of everything that has happened to him and marks of everything that will happen to him and even traces of everything that happens in the universe, even though God alone would recognize them all" (A VI iv [B] 1534: AG 41). By making minds eternal, allowing them to sense all endeavors, and assigning them traces of all that has gone before and marks of all that will occur, he makes each mind a mirror of the entire course of the world at every moment in time. Each mind reflects or mirrors the entire world at every moment of the mind's eternal existence. In *Discourse* §15, he summarizes the point in terms of expression: Each substance is of "infinite extension insofar as it expresses everything" (A VI iv [B] 1646). By such means he agrees with Plato "who taught that our soul expresses God, the universe, and all essences" (*Discourse*, §27).

GOD AND KNOWLEDGE. Throughout his life Leibniz was keen to acquire information about the world and to contribute to the sciences of his time. He studied history, designed machines, proposed lighting systems, created insurance programs, and contributed to the development of modern physics. Underlying all these enterprises, however, was his commitment to a Platonist epistemology according to which the divine Ideas are instantiated in the creatures in the world and exist in human minds innately. He summarizes this view in §28 of the *Discourse*: "The essence of our soul is a certain expression, imitation or image of the divine essence … and of all the ideas comprised in it."

From the very beginning of Leibniz's philosophical reflections on universal harmony, he recognizes its epistemological significance. In *Elements of Natural Law* (1668–1671), he presents for the first time the main steps that must be taken to acquire knowledge of fundamental truths. Since the goal of human life is to recognize the beauty and harmony in things, and harmony consists in *consonance* beneath apparent *dissonance*, we must learn to see beyond the dissonance. Once we abstract from the confusion of things and begin to recognize the underlying order of the world, the journey to this ultimate knowledge has begun. The first objects of knowledge are our innate Ideas, each of which is also an Idea in God's mind and so also instantiated in the world. By grasping one of these Ideas in the right way, we begin the process of knowing God and the ultimate nature of things. The goal of life is to recognize that everything is an emanation of God and hence a proper object of love. In a 1671 letter to Arnauld, he concludes this part of the project: "I show that it is the same thing to love others and to love God, the seat of universal harmony" (A II i 173–174: L 150).

In the *Philosopher's Confession* (1672–1673), Leibniz clarifies and expands upon the relation between universal harmony and knowledge: "The nature of mind is to think; therefore, the harmony of the mind will consist in thinking about harmony; and the greatest harmony of the mind or happiness will consist in the concentration of universal harmony, i.e., of God, in the mind" (A VI iii 116–117). The goal of life is to intuit the essence of God, which is evident in the "universal harmony" of the world. The means to this goal is to grasp "the eternal and immutable … Ideas" (A VI iii 120). The journey to knowledge begins when one "withdraws from the senses and draws back into his own mind." After a sincere "struggle toward the truth," "a stroke of light" may appear "as a split in the darkness" (A VI iii 120–121). Through the proper approach to the world, it is possible to be "admitted to God, i.e., universal harmony," to grasp it "in a single stroke of vision," and thereby to have "delight without end" (A VI iii 139). However, because minds are mostly "deformed" and exist "in shadow," many fail to recognize the "wondrous" interconnections among things (A VI i 464–465).

Leibniz remained committed to this form of innatism throughout his life. Thirty years after the *Elements of Natural Law*, he criticized the empiricism of Locke's *Essay concerning Human Understanding* in his own *New Essays*, noting that innate ideas distinguish us from beasts. According to Leibniz: "This is how ideas and truths are innate in us, as natural inclinations, dispositions, habits, or potentialities." Agreeing with Plato, he maintains: "The soul contains from the beginning the source of several notions and doctrines, which external objects awaken on certain occasions." Endorsing Paul's approach to knowledge, he quotes Paul's Letter to the Romans (2:15): "The law of God is written in our hearts" (A VI vi 49–52: AG 292–294).

Universal harmony increases the possibility for knowledge; the mirroring of minds increases it still more. For Leibniz *the wisdom of God* requires that creatures mirror one another and thereby add to the beauty and harmony of the world. He was motivated to convert the world into a harmony of mirroring substances at least partly in order to maximize the likelihood of such reflective awareness. The mirroring of minds increases variety and harmony because each mind encompasses the whole of existence. In *On the Greatest of Things*, each mind perceives the entire world at every moment of its eternal existence: "It seems to me that every mind is omniscient

in a confused way, that any mind perceives simultaneously whatever happens in the entire world" (A VI iii 524: Pk 85). In developing these views about plenitude and harmony, he reasons that it is good to maximize the number of diverse creatures in the world; it is even better to maximize the perception of that infinity of good things by making each creature mirror every other; but it is best to maximize the harmony among creatures by making all minds connected to all others at all moments in the eternity of the world.

Leibniz is rarely as explicit about the close relation between emanation and knowledge as he is in *On the True Mystical Theology*, a German text written (probably) in the final years of the seventeenth century. He begins with the metaphysics of universal harmony and its related epistemology: "Every perfection flows immediately from God. Only the inner light that God himself kindles in us has the power to give us a right knowledge of God." But it is not easy to acquire this knowledge: "The divine perfections are concealed in all things, but very few know how to discover them there. Hence there are many who are learned without being illumined, because they believe not God or the light but only their earthly teachers or their external senses and so remain in the contemplation of imperfections." Each created thing or "self-being" is from God and is therefore "a single self-sufficient" and "indestructible thing."

This separateness from God makes it difficult to recognize the divinity within us, but in our connectedness to God, it becomes easy: "God is the easiest and the hardest being to know." We can find "the essential truth" by seeking out the attributes of God: "The knowledge of God is the beginning of wisdom, the divine attributes are the primary truths for the right order of knowledge." Once we acquire knowledge of an attribute of God, which is present within us as an innate idea, we begin to approach "the essential light," which is "the eternal Word of God, in which is all wisdom, all light, indeed the original of all beings and the origin of truths. Without the radiation of this light no ones achieves true faith, and without true faith no one attains blessedness." He summarizes: In each mind "there lies an infinity, a footprint or reflection of the omniscience and omnipresence of God." Were we to acquire this "right knowledge of God," we would thereby attain "all wisdom, all light, indeed the original of all beings and the origin of all truth" (Guhrauer, 411–412: L 367–369).

LOGIC, TRUTH, AND PEACE

Biographers have claimed that as a boy Leibniz became dissatisfied with the categories of Aristotelian logic. Whatever truth there is in this, the youthful Leibniz joined the growing debate about the possibility of a universal language and a formal system for determining truth. For many seventeenth-century philosophers, the hope was to construct "an alphabet of human thought" that would form the basis for a universal language and a means of identifying truths. Leibniz intended to find a way to assign letters or numbers to the elements of thought so as to produce, "through the analysis of words" a means of judging the truth of all statements in the language. In *Dissertation On the Combinatorial Art* (1666), a young Leibniz begins work on this project, which he calls "the universal characteristic."

Although scholars have often treated Leibniz's account of logic and truth independently of his views about God and emanation, the two parts of his philosophy are closely related. The divine Ideas are the source of all truths, and human minds contain these Ideas innately, so the analysis of truth will involve these Ideas. Opening one of the main sections of *Dissertation On the Combinatorial Art*, he explains: "To begin at the top, Metaphysics treats being and the affections of being" (VI i 170: L 76). In 1671 he observes that although we are "conquerors of the world," we cannot have real knowledge until the mind has clarity about itself (A VI i 459). Leibniz's account of emanation and divine Ideas constitute a major part of the foundation for his program in logic because the ideas innate in us are also those emanated by God in the creations of the world. This connection persists in his thought until the very end; in the *Monadology* he observes that our mind contains "knowledge of eternal and necessary truths … thus in thinking of ourselves we think of being" and "of the immaterial and of God himself" (§29–30).

The relation between being and truth motivates other projects related to language. As with many of his contemporaries, Leibniz was fascinated with the evolution of languages since the "original language" of Eden. Many assumed that the language spoken by Adam and Eve made the truth more perspicuous and so attempted to recreate it. He went beyond most of his contemporaries in his fascination with the Chinese—both their language and culture. Like many of the Jesuit missionaries in China, he believed that the (apparently) extraordinary insights of the Chinese proved that the *elements of truth* were available to any who knew how to seek them and

that the identification of such truths would promote universal communication and eventually universal peace.

GOD, EVIL, AND THE BEST

PHILOSOPHER'S CONFESSION. Written within a year of his arrival in Paris, the *Philosopher's Confession* is a dialogue in which Leibniz discusses at length and for the first time the problem of evil, a problem that, together with a group of related problems, would engage his attention for the next forty years. The problem is ancient: How can the evil in the world (say, the suffering of innocents) be reconciled with the existence of an infinitely powerful, just, and good Supreme Being? Already in 1672–1673, he has a solution, one that would remain an important part of his thinking: The goodness of God is sufficient reason to create a world that is the best possible, and (apparent) evil is a necessary part of such a world. His solution is embedded in his notion of universal harmony: The world is the best and most harmonious possible despite the fact that its enormous diversity includes events that often suggest otherwise.

In order to explain how this world is best, it was necessary to develop a more thorough-going account of creation. Leibniz did this in the *Philosopher's Confession*. The divine intellect contains an unspecified number of eternal and immutable Ideas that constitute the divine essence and that God wills to instantiate in the world. That is, the essence of God "contains" the "nature of the things themselves" (A VI iii 124). But the essence of God does not necessitate *this* nature of things. Rather, God selects among possible versions of the divine essence and then emanates the selected version so as to create and sustain the world. He refers to these versions as possible *series of things*; he will later call these possible worlds. Each individual created thing is an instantiation of the (selected) divine essence. Further, God has a sufficient reason for choosing each thing, and each thing has a sufficient reason for acting as it does. He summarizes his position: "The present state of things depends on the series of things. The series of things depends on the universal harmony. The universal harmony depends on those well -known eternal and immutable ideas themselves … contained in the divine intellect" (A VI iii 131). God is "the sufficient and complete reason" for the world (A VI iii 123). God understands this world to be most harmonious and thereby has sufficient reason to choose it.

Leibniz's *best possible world* solution to the problem of evil gives rise to further problems: One concerns (what scholars sometimes call) the *author of sin*; another concerns the status of human freedom. On Leibniz's account, God causes evil, for God creates the best series of things, including many things that are, when considered in themselves, bad or sinful. In the *Philosopher's Confession* he responds to this problem by pointing out that God takes no delight in the existence of evil and hence is not properly thought to will it. In later works, he came to regard this response as inadequate. According to Leibniz, there is a sufficient reason for every thing that happens in the world. As we will see below, this principle plays an important role in his thinking about the world. When applied to the problem of human freedom, the principle commits him to determinism. For Leibniz, the will is never free of antecedent causes and in that sense it is always determined. But he is also a compatibilist in the sense that, just as God's perfect freedom does not involve lack of determination by the divine essence, so human freedom does not require undetermined choices. Freedom requires only spontaneity, or more exactly, the sort of spontaneity possessed by rational substances.

In both the *Elements of Natural Law* and *Philosopher's Confession*, Leibniz's approach to the problem of evil also has an epistemological aspect. The nature of universal harmony makes the acquisition of knowledge both more difficult and more *glorious*. Because there is a struggle, there will be some who fail. Yet the world is a better place because of the struggle to recognize the harmony among all things. When one sees an "unexpected" unity "where no one would suspect a connection" (A VI i 484–485), there is more delight and happiness. "The most confused discord fits into the order of the most exquisite harmony unexpectedly, as a painting is set off by shadow, as the harmony due to dissonances transforms the dissonances into consonance" (A VI iii 126). "Given that the whole is pleasing, it does not follow that each part is pleasing. … Only the whole is pleasing, only the whole is harmonious" (A VI iii 130). For Leibniz the beauty and goodness of the whole justifies the *apparent* ugliness and evil of some parts. In the end, the world is better because apparent disorder will "unexpectedly" reveal "the wonderful reason" behind this "greatest" of symmetries (A VI iii 122).

THEODICY. Leibniz's last extended treatment of the problem of evil restates many of the themes from the *Philosopher's Confession*, written almost forty years earlier. The *Theodicy* is a long, digressive work, devoted mainly to the topics listed in its subtitle: the goodness of God, human freedom, and the origin of evil. But the book also functions as a defense of the consistency of faith and reason. It is divided, rather arbitrarily, into three *essays*, preceded by an author's preface and a "Preliminary Dis-

sertation on the Conformity of Faith with Reason," and succeeded by various appendices.

Much of the *Theodicy* consists of Leibniz's responses to other authors, Bayle in particular. His own metaphysical system is in the background. His idealism, for example, is barely mentioned at all. But the characteristic themes of his philosophical theology nevertheless dominate the text, and it is in the *Theodicy* that his most complete response to the problem of evil is found. That response is, at its core, the same as the response that he gave in the *Philosopher's Confession*: that this is the best, that is, the most harmonious of all possible worlds; that the evils within it are not to be judged apart from the entire series of things; that God's perfection requires that only the best possible world be created; that humans therefore cannot reasonably wish that things had been different; that happiness is to be sought through understanding the perfection of God, the creator of all things, and the perfection of all the things that God has created.

The problem of the author of sin, to which Leibniz had given only a weak response in the *Philosopher's Confession*, is in the *Theodicy* handled with much more verve and power. He distinguishes between God's antecedent and consequent will. God wills each possible thing antecedently in proportion to its perfection. But some possible things are not compossible with others, so not all God's antecedent willings can be realized. God's consequent, that is, final and decisive, will is the existence of that series of things that realizes as much perfection as possible. To this account is added an Augustinian idea of metaphysical evil as mere privation or limitation. Thus, God does not will evil at all, for God's willing is directed only toward the perfection in things, and imperfections are nothing at all, and so not even possible objects of will.

The *Theodicy* contains extensive discussion of freedom, including many objections to so-called *freedom of indifference*—the capacity to choose between alternatives that are equally advantageous (or disadvantageous). Leibniz's commitment to the principle of sufficient reason rules out any such capacity, even in the case of God—a conclusion that plays a significant role in some of the argument in his later correspondence with Clarke. He allies himself with Augustine and the Thomists in holding that everything is determined and with Aristotle in requiring as conditions of freedom only spontaneity and intelligence. The rejection of a contracausal account of freedom also reflects Lutheran doctrine, and one of the declared goals of the *Theodicy* is to provide an account of human freedom on which Catholics and Protestants can agree.

As in other writings Leibniz struggles in the *Theodicy* to give an account of contingency that avoids necessitarianism. Absolute or metaphysically necessary truths exclude any alternative; they rely on the principle of noncontradiction. This kind of necessity is incompatible with freedom, and not even God is free with respect to these truths. Thus, according to Leibniz, God was not free to create spaces with fewer or more dimensions than three, for such spaces are logically impossible. Physical and moral necessity, by contrast, resting on the principle of sufficient reason, is not incompatible with freedom. God is free in choosing to create the best possible world because there are other worlds that are possible in themselves (even though God, being perfect, would not in fact create them); rational creatures are free in the choices they make if there are other options (even though, given preceding causes, they will not in fact choose them). His compatibilist account of freedom appears here in its starkest form: Both divine and human freedom require only the bare logical possibility of some alternative course of action. God is perfectly free because perfectly rational; humans are imperfectly free because less than perfectly rational. Acting against or without reason is, for Leibniz, the paradigm case of unfreedom.

This compatibilism, even if acceptable, leaves little room for contingency, and scholars have long argued the question whether Leibniz manages to avoid the claim that everything that happens, happens necessarily. His standard answer, given many times in the *Theodicy*, is that it depends what sort of necessity is intended. Nothing happens by logical necessity except when the opposite involves a contradiction; everything happens by moral necessity, for unless this entire *series of things* were the uniquely best, God would lack a sufficient reason to create it. It is nevertheless hard to see how any other series of things is ever possible given the necessary existence and perfection of God. Here the tension between his Platonism and the voluntarism of the Christian tradition is at its greatest.

Leibniz himself seems never to have wavered from the underlying optimism of his account of *the best of all possible worlds*. He often notes that he knows no one as happy as he. He summarizes the source of his contentment in a letter to Queen Sophie Charlotte:

> But the consideration of the perfection of things, or, what is the same, of the supreme power, wisdom, and goodness of God, who does everything for the best, that is, with the greatest order, is sufficient to make all reasonable people content, and to convince them that contentment

> should be greater to the extent that we are disposed to follow order or reason." (AG 192)

Leibniz's optimism, and his claim that this is the best of all possible worlds, was viciously satirized by Voltaire in *Candide*. But Voltaire's Dr. Pangloss, the representative of optimism, is a very unreliable guide to Leibniz, or even to the Leibnizianism of his disciple Wolff. Leibniz, from the *Philosopher's Confession* on, insists that the best possible world is not best in all of its parts. By the time of the *Theodicy*, he has a battery of arguments against the kinds of objections that Voltaire advances. But Voltaire's short and witty tale is a far better read than the long and, at times, tedious *Theodicy*, so it is not surprising that its *argument* is better liked.

SUBSTANCE, MATTER, AND NATURE

At the very end of his life, Leibniz explains that in order to understand the intellectual *discoveries* of others, it is often necessary "to detect the source of their invention" (G III 568). In presenting his views about God, creation, mind, activity, knowledge, and harmony, it is helpful to detect their Platonist sources. In order to understand his *discoveries* about the natural world, it will be necessary to detect the sources of his *invention*.

ARISTOTELIANISM AND MECHANISM. For most of his life Leibniz takes there to be two kinds of basic, natural entities, or substances. The first sort is a corporeal substance constituted of two *principles of nature*: one active, one passive. Corporeal substances are analogous to organisms: They are active, unified things with a material component or body and an organizing principle. The second kind of substance is variously called "mind, soul, spiritual substance," and "substantial form[s]." Although these are the active things in nature, which are tied to a material component of some sort, they are themselves also substances. Toward the end of his life, Leibniz began to call the ultimate components of nature monads. In the world of his monadology, there are only mind-like simple substances in various collections.

The Aristotelian philosophy offered the raw materials for Leibniz's account of substance; the new mechanical philosophy constituted the basis for his physics. Although he transformed those philosophies to suit his own philosophical and theological needs, he remained wedded to (what he considered to be) Aristotle's basic insights about the self-sufficiency of substances and to the mechanists' commitment to explain corporeal phenomena in terms of matter and motion.

For most Aristotelian philosophers, natural objects are constituted of two principles, matter and form, and natural events are explained in terms of the actualization of the potency of these two principles. When Leibniz began constructing his own philosophy in the mid-1660s, there was a new explanatory model available, one that had greatly diminished the power of the scholastic model. According to the mechanical philosophy (as it came to be called), nature is composed of matter—whether the extended stuff (*res extensa*) of Descartes, the atoms of Gassendi, or one of the many less popular accounts of corporeity—whose actions and movements cause and explain all the phenomena of nature. For the mechanist all physical phenomena are to be explained in terms of some kind of matter and motion. Although these thinkers disagreed about how to define the material component in nature, they all took it to be void of substantial forms.

Despite the genuine innovation of the new mechanical philosophy, it failed to solve adequately a number of important theological and metaphysical problems. By the middle of the seventeenth century, especially in the Protestant areas of northern Europe, a number of conciliators took it upon themselves to *reform* the Aristotelian philosophy rather than abandon it. Different reformers had different recipes for mixing the old with the new, but they all combined some part of the mechanical physics with Aristotelian metaphysics. Each claimed that, when properly understood, the Aristotelian philosophy could comfortably accommodate mechanical philosophy. Like these reformers, Leibniz also recognized very early on that the Aristotelian theory of substance could easily accommodate the new mechanical physics and thereby explain the phenomena.

The Aristotelian philosophy appealed to the young Leibniz for several reasons. At the heart of the Platonized Aristotelianism that his mentor, Jakob Thomasius, bequeathed to him stands the idea that nature is constituted of individual corporeal substances whose substantial forms act to compose a divinely arranged harmony. From the beginning of his philosophical career, Leibniz embraced the assumption that everything in the world acts to instantiate the good. Unlike those of his contemporaries who rejected final causation, he embraced the Aristotelian idea that nature moves toward the good. For Leibniz, an Aristotelian account of substance formed a secure foundation for such a rational, harmonious, and good world although it needed to be *reformed* to fit mechanical explanations in physics. He committed himself to the Aristotelian and mechanical philosophies as a youth and maintained this commitment until his death.

In the *Monadology* he writes: "Souls act according to the laws of final causes. … Bodies act according to the laws of efficient causes or of motions. And these two kingdoms, that of efficient causes and that of final causes, are in harmony with each other" (§79).

Leibniz had excellent metaphysical reasons to accept a major part of the Aristotelian philosophy. But he had other incentives as well. From the perspective of war-ravaged Germany, Aristotelianism must have seemed to Leibniz the safest bet as a philosophy of religious reconciliation. The doctrinal declarations of contemporary Catholics were framed in Aristotelian terms while Aristotelianism survived in Lutheran cities such as Leipzig. Aristotelian notions of substance thus presented themselves as ideal both for understanding the divinely arranged harmony in the world and for working toward religious and political harmony within it.

SUBSTANCE, SELF-SUFFICIENCY, AND THE REFORMATION OF THE MECHANICAL PHILOSOPHY. The young Leibniz intended to transform the Aristotelian notion of substance so that it would accommodate mechanical physics. For Leibniz, the mechanical physics of philosophers such as Descartes, Hobbes, Gassendi, and Galileo reduces to the following claims: There is some sort of matter or extended stuff (*res extensa*), which is (somehow) moved and whose arrangements both cause and explain the corporeal features of individual bodies; therefore, a body is organized *res extensa*, and all corporeal features are reducible to the arrangements of such extended stuff. Leibniz was *never* satisfied with the metaphysical foundations offered by leading proponents of the mechanical physics; the physical explanations of particular phenomena seemed adequate, but the metaphysical underpinnings of those explanations did not.

Leibniz's most fundamental assumption about the natural world is that it is composed of substances, each of which has its own source of activity by means of which it is constituted as a self-sufficient, unified thing. The material stuff of the mechanical philosophers did not have its own internal source of activity and so was neither self-sufficient nor properly unified; it therefore could not by itself constitute genuine substances. In his earliest comments about substances, Leibniz explains that because the corporeal substance of the mechanists "is not self-sufficient … an incorporeal principle must be added" (A VI i 490: L 110). This incorporeal principle is a substantial form or mind that organizes the matter and thereby makes it into a unified, self-sufficient thing. He corrects the mistakes of the mechanists by making substance active, allowing it to be both causally and explanatorily complete. He demotes the matter of mechanical physics to the status of the passive principle in substance and insists that the active mind or substantial form organizes the passive principle so as to make a unity with it.

The result is an individual corporeal substance that can act as the cause and explanation of its own (at least) basic features. Although the details of his views about substance will continue to evolve over the course of his long philosophical career (e.g., he comes to conceive the passive principle as itself constituted of mind-like substances and eventually prefers to construct the world entirely out of monads), he never wavers from his commitment to the causal and explanatory autonomy of the fundamental entities of nature. It is this robust self-sufficiency that is his most profound debt to the metaphysics of Aristotle. And it is this robust self-sufficiency that inspired many of the core doctrines of his mature thought.

THE METAPHYSICS OF SUBSTANCE BEFORE 1680

For much of the twentieth century, scholars maintained that Leibniz developed his theory of substance in the 1680s. Except for a few scattered works—mostly those in logic and physics—his earlier texts were either neglected or dismissed as juvenilia. However, close attention to writings from the 1660s and 1670s reveals that Leibniz developed his theory of substance much earlier. In this section we consider the most important of the early texts.

ORIGINAL ASSUMPTIONS ABOUT SUBSTANCE, ACTIVITY, AND SELF-SUFFICIENCY. During the mid-1660s, Leibniz worked on a number of related projects in law, logic, and theology. Encouraged by the distinguished German statesmen Boineburg, he began composition of the *Catholic Demonstrations* in 1668. The work, as Leibniz conceived it then, was to consist of a series of philosophical prolegomena and four parts. The prolegomena were to contain the *elements of philosophy*, that is, the *first principles* of metaphysics, logic, mathematics, physics, and practical philosophy, while the four parts were to be demonstrations of the existence of God, the immortality of the soul, the Christian mysteries (e.g., the Eucharist), and the authority of the church and scripture. The work was designed to offer a metaphysics that would cohere with Catholic and Lutheran doctrine and thereby effect a reconciliation between the two churches. But another sort of reconciliation is promoted within the work, for when Leibniz began the *Catholic Demonstrations*, he was

committed to a version of Aristotelian philosophy as he interpreted it and also to a mechanical account of the phenomena of nature.

The theological writings indicate exactly how his reconciliation of these two philosophies evolved in his attempt to explain the theological doctrines of the Eucharist, the immortality of the soul, and so on. He takes the Aristotelian notion of substantial form as the active principle of nature and combines it with the mechanical notion of passive extended stuff as the passive principle to create a coherent *reformed* Aristotelianism. At work in these theological essays are a number of philosophical assumptions. The most important of these are as follows (except for the Principle of Sufficient Reason, the names are not his):

- The *principle of substantial activity* assumes that a being is a substance if and only if it subsists per se, and a being subsists per se if and only if it has a principle of activity within its own nature.
- The *principle of sufficient reason* assumes that there is a complete or sufficient reason for everything.
- A *complete reason* for a state or feature *f*: (1) constitutes the necessary and sufficient condition for *f*; (2) is perspicuous in that, in those cases where one can understand it, one sees exactly why *f* as opposed to some other state of affairs came about; (3) is such that in those cases when a full account of it can be given, that account constitutes a complete explanation of *f*; and (4) does not require a reason of the same type.
- The *logical assumption* claims that, for any state or feature *f*, the logically necessary and sufficient conditions of *f* exist and in theory can be articulated.
- The *intelligibility assumption* claims that those conditions are in theory intelligible.
- The *substantial nature assumption* claims that every substance has a nature that contains the set of necessary and sufficient conditions or the complete reason for those features that strictly belong to it, and moreover, those conditions are in theory intelligible.

The precise status of these assumptions in the *Catholic Demonstrations* and related early texts is unclear. They constitute the underlying principles of Leibniz's discussions during this period. Although in the texts of 1668–1671 they may have the status of working hypotheses, they continue to inform and direct his thinking about metaphysical matters for years to come. Some of his most characteristic doctrines directly develop from these assumptions.

SUBSTANTIAL FORMS AND ACTIVITY. While developing his account of substances as the fundamental entities of nature in 1668–1671, Leibniz was also working on the *Elements of Natural Law*. As his views about universal harmony evolved, he integrated his Platonist assumptions about activity, emanation, and unity into the Aristotelian and mechanical assumptions about self-sufficiency, substantial forms, and matter. He assumes that substantial forms are divine-like and possess the kind of metaphysical powers described by Ficino. The idea here is that God continually emanates the divine essence to each individual mind and furnishes each mind with its own source of activity thereby generating unity and self-sufficiency. He suggests that each active thing acts constantly according to a *reason* given it by God: "Just as God thinks things … because they follow from his nature, so does Mind" (A VI ii 287–288). By being Godlike the active principles or substantial forms possess divine-like features, such as unity and self-sufficiency. They also act according to a divinely arranged *reason* (A VI i 534).

The principle of substantial activity reveals the close relation between substancehood and activity: Anything that possesses its own source of activity will be self-sufficient and hence substantial. In Paris, Leibniz develops this idea so that mind-like, active things are indestructible and the source of the individuality, unity, and identity of the corporeal substances of the world. No active creature is ever without a body or passive principle; only divine mind is "devoid of body" (A VI iii 100). God "arranged all things from the beginning" (A VI iii 477: Pk 31) so as to give each created substance a *rule* or set of instructions by means of which it acts (VI iii 483: Pk 39). As he summarizes his position:

> There are certainly many and important things to be said … about the principle of activity or what the scholastics called substantial form, from which a great light is thrown on Natural Theology and … the mysteries of faith. The result is that not only souls but all substances can be said to exist in a place only through the operation of their active principle, that souls can be destroyed by no power of body; and that every power of acting exists from the highest mind whose will is the final reason for all things, the cause being universal harmony; that God as creator can unite the body to the soul, and that in fact, every finite soul is embodied, even the angels are not excepted. (A VI iii 158)

In the pre-Paris period, minds are considered constantly active and therefore self-sufficient, unified things. In *On the Greatest of Things*, written during his final year in Paris, Leibniz develops and expands on the relations between activity, self-sufficiency, unity, and divisibility: "whatever acts cannot be destroyed" naturally, and yet "whatever is divided is destroyed" (A VI iii 392–393: Pk 45–47). Mind or substantial form acts as the "cement" in a corporeal substance and thereby guarantees that its passive principle will not be divided (A VI iii 474: Pk 27). Consistent with the theory of corporeal substance developed earlier, the mind-like substantial form acts constantly through its passive principle to create a single "unsplittable" thing, which Leibniz sometimes calls an "atom" (A VI iii 393: Pk 47). This atom or unified thing is a corporeal substance constituted of an active and a passive principle. Consistent with the substantial nature assumption, the nature of the substance acts as the necessary and sufficient condition of its features. In 1676, then, the activity of mind individuates the substance, unifies it, and makes it eternal. Throughout a substance's eternal existence, it is its active principle that will organize its passive principle so as to constitute its eternally self-sufficient nature.

In these early years the persistence of the substantial nature through various changes is especially important to Leibniz because of his concern for developing a metaphysics consistent with Christian doctrine. The doctrine of resurrection, for example, gives rise to the question: How can it be the same human substance that persists through the radical changes in a human life, then dies, and then is resurrected? He explains that the mind "is firmly planted in a flower of substance [that] subsists perpetually in all changes" and that can be "diffused" through a greater or less expanse of the original body (A VI iii 478–479: Pk 33). The mind-like principle of activity acts as the *cement* of the substance and forms the unity that persists through all substantial changes, including even bodily death and resurrection. In a letter to Johann Friedrich of 1671, he explains that in the same way that "God is diffused through everything," so mind is diffused through its body; just as the activities of God do not diminish the divine essence, so too the mind acts on its body "without being diminished" (II i 113).

It is clear from these texts of 1670–1676 that Leibniz believes he has hit upon an account of substance that comfortably accommodates the severe metaphysical demands of Christian doctrine, the physical explanations of the mechanists, and the Aristotelian commitment to the causal completeness and self-sufficiency of substance. Although the details of his position are in flux and will shift over time, the basic structure of this account of substance will not vary until the development of the world of the monadology. For Leibniz, a corporeal substance is a self-sufficient and unified thing that results from a substantial form activating and organizing its passive principle. The substantial form acts constantly on its passive principle by *a set of instructions* given it by God. The passive principle is the substantial form's *instrument of acting*. The unity is what results from the constant activity of the active principle on the passive one, thereby forming an organized unified thing.

MATTER, EXTENSION, AND PASSIVITY. Within weeks of entering the University of Leipzig, at the age of fourteen, Leibniz had a major philosophical insight. He recalls walking in some woods near his home and "deliberating whether I should keep the substantial forms" or convert to mechanism. In the end he decided to accept the physical explanations of the mechanical philosophers as opposed to those of the scholastics and thereby "to apply" himself to mathematics (G III 606: L 655). The young Leibniz thus assumes that the passive principle in corporeal substances is material, like the *res extensa* of Descartes. For the next few years he maintains that the active principle or substantial form takes this passive extended stuff, organizes it into an individual body, and thereby creates a unified thing or corporeal substance.

In the theological essays of 1668–1671, he conceives the union between the active and passive principles as involving constant activity, where the mind-like substantial form cannot "act outside itself" except through its passive principle (A VI i 533–534). The unity here is analogous to that in organisms in the sense that if the activity involved in maintaining an organic unity stops, so does the unity. When the maintenance of the organization ceases (e.g., the heart stops, the liver no longer functions), the unity of the substantial form and matter does so as well (e.g., the entity dies, the formerly organized body becomes a heap of decaying flesh). The nature of organic unities also helps us to understand what he means when he says that the active principle cannot act outside itself except through the passive: In order to act externally, the source or cause of the organization has to act through the passive principle that it organizes.

In the 1670s Leibniz became dissatisfied with this account of passivity. There were several problems. First, the mechanical account of body could not easily accommodate important theological doctrines, such as the Eucharist and resurrection of the body. According to the

Lutheran account of the *mystery* of the Eucharist, the body of Christ and the body of the bread exist side by side. However, if the body of Christ is a collection of extended stuff, it is unclear how it can be distinct from and coextensive with the extended stuff that constitutes the matter of the bread. Leibiniz argues: "For if body and space are one and the same, how can we avoid the consequence that in different spaces or places there must be different bodies" (A VI iii 157–158). He concludes that the views of the mechanists, who believe that *the essence of body consists in extension*, are therefore incompatible with the miracle of the Eucharist. He also argues that since, according to Descartes and other mechanists, each body is constituted of extended stuff and since all extended stuff is essentially the same, it becomes enormously difficult to give any particular body (say, Christ's body) a stable identity. Leibniz concludes: "One cannot say…why it is called the body of Christ rather than bread, to which it is very similar" (A II i 170). Nor, to take the case of another Christian doctrine, can one say how to identify and individuate bodies at the time of the resurrection.

Another problem facing Leibniz's early account of the passive principle in corporeal substance is less overtly theological. According to the principle of plenitude as he interpreted the ancient doctrine, the world is as full of diverse being as possible. But according to the version of Platonism that Leibniz learned as a university student, matter is uniform, divisible, unreal stuff. In the *Phaedo*, Plato describes it as "as unintelligible, soluble and never consistently the same" (80e). Matter lacks all unity and activity; it contributes nothing positive to the world. It follows from these Platonist assumptions that the world would be made better by filling it with mind-like unified things and stripping it entirely of extended passive matter.

There has been much disagreement among scholars about when Leibniz does finally strip the world of extended stuff. Once we take seriously Leibniz's interest in Platonism and his concern to solve the theological problems posed by doctrines such as the Eucharist and resurrection of the body, it seems relatively clear that he abandons extended stuff while still in Paris although he remains undecided about what exactly to put in its place. In the Paris texts he asks as many questions as he answers: "Since mind is something that has a certain relation to some portion of matter, it must be stated why it extends itself to this portion and not to all adjacent portions; or why it is that some body, and not every body, belongs to it in the same way" (A VI iii 392: Pk 45). In 1676 he did not have consistent answers to these questions; the texts are unclear about the precise nature of the passive principle in substances. However, one of the hypotheses that he entertained is that bodies are themselves unextended collections of mind-like substances whose only actions are perceptual states.

BODY AND FORCE. The young Leibniz embraced mechanical physics, according to which the features of bodies are to be explained in terms of the broadly geometrical properties of their parts—whether these are tiny indivisible atoms or infinitely divisible stuff—whose configurations shift and change through motion and whose motion changes through collision. When he published his *New Physical Hypothesis* and *Theory of Abstract Motion* in 1671, he agreed with the standard mechanical account of collision as the only means by which bodies naturally change motion. His *abstract* account of motion is offered in terms of the Hobbesian notion of conatus, defined here as "an indivisible, nonextended part of motion" and as "the beginning and end of motion" (A VI ii 264–265: L 139–140). In 1671 he agreed with Descartes that "all power in bodies depends on speed." If two bodies with unequal speeds collide, they will move together after the collision in the direction of the faster body with a speed that is the difference between the two (A VI ii 228). By the time he met Spinoza in the autumn of 1676, he had begun to question features of this mechanical account, and in particular, the law of the conservation of motion proposed by Descartes.

In the winter of 1677–1678, Leibniz takes some observations made by Huygens about impact and transforms them into a notion central to his thought. He decides that force or power of action must be conserved in collision between bodies rather than mere speed. By January 1678 he has hit upon the proper account of this force: mv^2 (mass times velocity squared). Given the importance of this insight, it is odd that he does not publish any part of his findings until 1686, and even then, in his *Brief Demonstration*, he merely criticizes Descartes's conservation principle and ONLY hints at his own account. Over the next few years, he will work out the details of his dynamics, especially in response to Newton's *Principia Mathematica* (1687).

Leibniz's discovery of mv^2 was enormously important and radically changed his account of the physical world. As he explains in the *Specimen of Dynamics* (1695), he was forced to recognize that in physics, purely geometrical notions were inadequate: "We must add to material mass a certain superior and so to speak formal principle.

Whether we call this principle form or entelechy or force does not matter so long as we remember that it can only be explained through the notion of force" (GM VI 241: AG 124–135). He notes the easy fit between an Aristotelian approach to substance (whose principle of activity is often described as *form or entelechy*) and the new notion of force. Leibniz had hit upon the basic features of his Aristotelian account of substance in the late 1660s. With the development of his dynamics, all he had to do was to redescribe the active principle in nature. The mind-like substantial forms in nature were now responsible for more than just the activity of creatures; they were also responsible for their force.

THE PRINCIPLE OF SUFFICIENT REASON. Leibniz is well known for his commitment to the principle of sufficient reason, which he often calls his *great principle*. As early as 1668 he assumes that God always has a reason for *choosing* one state of affairs rather than another and that this reason must be sufficient. In 1671 he calls the principle a *first* truth; and by way of demonstration, he adds: "Everything that is has all its requisites" since a state of affairs will not exist unless all its requisites "are given. … Consequently, everything that is has a sufficient reason" (A VI ii 483). Later in his career he articulates the principle in various ways, often in terms consistent with his account of truth. In the *Monadology*, for example, he presents it as the principle "by virtue of which we consider that we can find no true or existent fact, no true assertion, without there being a sufficient reason why it is thus and not otherwise, although most of the time these reasons cannot be known to us" (§32).

Leibniz's early commitment to the principle is matched by his early application of the principle to God as the sufficient reason of the world and to the natures of substances as the sufficient reason for their features. According to the substantial nature assumption, every substance has a nature that contains the set of necessary and sufficient conditions or the complete reason for its features. But a question arises about which features are covered here. If the nature of a substance is so complete as to contain the sufficient reason for all the features of the substance, then the principle of sufficient reason and the substantial nature assumption together bring us to the brink of two of his more startling metaphysical claims. The first is phenomenalism; the second preestablished harmony.

PREESTABLISHED HARMONY AND PHENOMENALISM. Although Leibniz does not use the term *preestablished harmony* until the 1690s (in the 1680s he calls it the theory of *concomitance*), there is significant evidence that he adopted its constitutive tenets in the 1670s and perhaps as early as 1671. The doctrine of preestablished harmony holds that each substance acts out of its own nature (spontaneity), that no substance causally interacts with any other substance (world apartness), and yet that each substance in the world parallels the activities of all the other substances perfectly (parallelism). The theory is closely related to another component of his mature philosophy: phenomenalism. The phenomenalism of the mature Leibniz, what is sometimes called *well-founded phenomenalism*, includes at least the following two claims: Bodies are phenomenal objects and so our perceptions of them arise from our own internal nature; and our perceptions nonetheless correspond to (parallel) the activities of real (unextended and mind-like) substances and in that sense are *well founded*.

The *New System* of 1695 summarizes the doctrines: "We must say that God originally created the soul (and any other real unity) in such a way that everything must arise for it from its own depths, through a perfect *spontaneity* relative to itself, and yet with a perfect *conformity* relative to external things." Since our perceptions are "internal perceptions in the soul itself" they "must arise because of its own original constitution," which is "given to the soul from its creation," and "constitutes its individual character. … This is what makes every substance represent the whole universe" from its own point of view, and "makes the perceptions or expressions of external things occur in the soul at a given time, in virtue of its own laws, as if in a world apart, and as if there existed only God and itself." In the perfectly harmonious world chosen by God, "there will be a perfect agreement among all these substances, producing the same effect that would be noticed if they communicated" (G IV 484-85: AG 143-44).

There is much, though scattered, evidence in the texts of the 1670s that Leibniz adopted most of the claims constitutive of phenomenalism and preestablished harmony early on. Neither preestablished harmony nor phenomenalism came to him suddenly. Rather, their core claims emerged gradually out of his attempts to solve the theological and philosophical problems that most concerned him. As he reflected on problems in ethics, law, theology, physics, and metaphysics, he developed his account of universal harmony and substance in an attempt to solve those problems. Preestablished harmony and phenomenalism resulted from the convergence of these solutions. These elaborate metaphysical doctrines were the most elegant way to solve a diverse group of dif-

ficult problems, to capture the rationality and goodness of God, and to reconcile ancient and modern ideas.

Preestablished harmony may be seen to result from the combination of universal harmony, the self-sufficiency of substances, and the mirroring of substances. According to universal harmony, God emanates the divine essence to every creature. The unity of the world is due to the fact that all creatures express the same thing: its multiplicity to the fact that each creature expresses the divine essence in a different way. The substantial nature assumption may be taken to entail that the complete reason for all the features of a substance is contained in its nature, in which case the complete reason for its perceptual states is contained there as well. The conjunction of the substantial nature assumption and universal harmony suggests spontaneity: For each substance, the manner of its expression of the divine essence will be contained *in its nature*. Further, if we assume that the substantial nature of a substance contains the necessary and sufficient conditions for each and every feature of it, then it seems to follow that the cause of every feature of the substance is *contained* in its nature, which is consistent with world apartness and the idea that there is no causal interaction among substances.

Finally, the theory that each substance mirrors all the others resembles the tenet of parallelism. Indeed, the parallelism of well-founded phenomenalism and preestablished harmony seems to be an extension of the Platonist notion of sympathy: Each substance, in its manifestation of the divine essence, is in perfect sympathy—for Leibniz, in perfect coordination—with every other. The doctrine of marks and traces is itself an elaboration of this notion of sympathy; it is also closely related to the idea that each substance is a world apart. Preestablished harmony is fundamentally emanation and sympathy perfectly organized in the self-sufficient substantial natures of the created world.

In the *Discourse* Leibniz implies that preestablished harmony is the blending of just these assumptions, and he acknowledges its close relation to his phenomenalism: "It is very evident that created substances depend upon God" who "produces them continually by emanation." In order to manifest divine "glory," God creates various substances to "express the universe." It follows from this account of God's relation to the world that "each substance is like a world apart, independent of all other things, except for God" from "whom all individuals emanate continually." By acting on us, God arranges things so that "all our phenomena, that is, all the things that can ever happen to us, are only consequences of our being" such that these phenomena are "in conformity with the world which is in us." It follows that "the perceptions or expressions of all substances mutually correspond" although each expression differs from every other. Finally, "if I were capable of considering distinctly everything that happens or appears to me at this time, I could see in it everything that will ever happen or appear to me" (A VI iv [B] 1549-51: §14).

Whether or not Leibniz commits himself to phenomenalism in the 1670s, he surely toys with the position. During his Paris period he often reduces the existence of bodies to the consistency of perceptions and concludes: "It does not follow that there exists anything but perception, and the cause of this perception and its consistency." The cause of perception is such that: "a reason can be given for everything and everything can be predicted" (A VI iii 511: Pk 63-65). From the perspective of conscious beings, in order to explain existence, it is unnecessary to resort to outside bodies; rather, we can reduce all existence to the consistency of perceptions, where the latter includes both the consistency of the perceptions within a mind and the coordination among minds: "We sense or perceive that we exist; when we say that bodies exist, we mean that there exist certain consistent perceptions, having a particular constant cause" (A VI iii 512: Pk 67).

In these and related texts of 1676, Leibniz seems to extend the substantial nature assumption to encompass all the features of substances, including their perceptual states. The suggestion is that God gives each substance a set of instructions or *rule* that makes each substantial nature the sufficient cause of all its features, including its perceptions. Thus, consistent with spontaneity and world-apartness, all the features of a substance are caused by its nature and there is no causal interaction among substances. Consistent with parallelism, "existence consists in" the coordinated perceiving of objects so that "several people perceive the same." It is "not necessary either that we act on them or that they act on us, but only that we perceive with such conformity" (VI iii 511: Pk 63). As a "perfect mind" God "arranged all things from the beginning" so as to make them "most harmonious" (A VI iii 474–476: Pk 25–29). For Leibniz in these texts of 1676, a major theme in this harmony is God's coordination of the perceptions among minds. Indeed: "Without sentient beings, nothing would exist. Without one primary sentient being, which is the same as the cause of all things, nothing would be perceived" (A VI iii 588: Pk 113). As he writes to Malebranche in 1679, "I have always been con-

vinced … that strictly speaking bodies do not act on us" (A II i 472-73: L 210).

THE METAPHYSICS OF SUBSTANCE, SECOND STAGE

Written during a snow storm in the Harz mountains in 1686, the *Discourse on Metaphysics* is the first general account of Leibniz's mature metaphysics. He sent a synopsis to Arnauld and thereby began the well-known correspondence between these two great seventeenth-century thinkers. Although not published during his lifetime, the *Discourse* and the correspondence with Arnauld, together with the terse summary of metaphysics contained in *First Truths*, have been favorites of twentieth-century Leibniz scholars. These texts have received a large amount scholarly attention, some of which is excellent. But we now know that many of their most important doctrines developed years earlier. For the most part, the *Discourse* and *First Truths* are summaries of doctrines extant in the 1670s, and what is new in them develops neatly from earlier views.

SUBSTANCES, SUBJECTS, AND TRUTH. In 1900 Bertrand Russell published a book in which he argued that Leibniz's metaphysics developed from his logic and theory of truth. For much of the twentieth century, scholars agreed with Russell that the theory of truth offers the key to Leibniz's philosophy and that the theory of substance developed out of that theory. With access to more of his writings and through attention to the sources of his ideas, it is clear that the core of his metaphysics—the account of substance and the theory of universal harmony—developed several years before the theory of truth. So, though the mature Leibniz sometimes puts the theory of truth front and center, it developed out of his views about the self-sufficiency, intelligibility, and explanatory completeness of substances; it was a consequence of those other views, not their source.

In 1676 Leibniz begins to emphasize subjects as the bearers of features. This is an important clarification of claims contained in the core metaphysics and constitutes a step toward the development of his conception of truth. One of his basic, Aristotelian assumptions is that substances are causally and explanatorily self-sufficient (at least with regard to their primary features). Another is that the relation between a feature and the substance to which it belongs is both logical and intelligible. These logical and intelligibility assumptions imply, for any feature of a substance, that the substance contains the logically necessary and sufficient conditions for that feature, that these conditions are in theory intelligible, and therefore that the truth of the attribution of the feature to the substance is in theory discoverable in the nature of that substance. When he extends the substantial nature assumption to cover all features, he commits himself to a truth-conferring relation between a substance and its features; a feature is truly predicated of a substance if and only if the nature of the substance contains the complete reason of that feature.

As Leibniz began to refine his views about the relation between the attributes of God and their instantiation in the world in the spring of 1676, he took his first steps toward the development of the idea that truth is a matter of relations among concepts. In *On the Greatest of Things*, he notes the metaphysical significance of substances as subjects or bearers of predicates and of truth as grounded in the relation between substances and their states: "It is a wonderful fact that a subject is different from forms or attributes. This is necessary because nothing can be said about forms on account of their simplicity; therefore, there would be no true propositions unless forms were united to a subject" (A VI iii 514: Pk 69). Once he has hit upon the idea that a substance is a subject in which a modification of the divine attributes has been placed, and once he sees truth in terms of the relation between a subject and such attributes, the materials are in place for the concept containment theory of truth. That there is a close connection between his metaphysical views about self-sufficiency and his theory of truth is clear. in a text of 1676 we find one of his first attributions of completeness to substance: "A substance or complete Being is for me that which alone involves all things, or for the perfect understanding of which, no other thing needs to be understood" (A VI iii 400: Pk 109).

By the spring of 1676, the metaphysical underpinnings of the theory of truth are in place, including the claim that there is a hierarchy of subjects. First there is God, who is the subject of all simple attributes; then there are creatures, each of which is the subject of a partial expression of those attributes. According to Leibniz: "The essence of God consists in the fact that he is the subject of all compatible attributes" or forms while it is the nature of created "subjects" to be "conceived through forms" (A VI iii 514: Pk 69–71). Before creation the Supreme Being conceives the fully articulated essence for each individual substance. It follows that all true statements about the active things in the world will be statements about a substance as a subject and its relation to one of the predicates contained in its complete concept. In such a world all basic truths about the created world involve the inclusion

of a predicate in the concept of a subject. For Leibniz, all the truths about an individual substance are contained in its nature.

Against this metaphysical background, it is unsurprising that, when Leibniz began working on logical matters in his early years in Hanover, he concluded that all truths were a matter of concept containment. For Leibniz, all there is in the world are divine attributes and their combinations. In a striking passage of 1676, he acknowledges this point: "There is the same variety in any kind of world, and this is nothing other than the same essence related in various ways, as if you were to look at the same town from various places, or, if you relate the essence of the number 6 to the number 3, it will be 3x2 or 3+3, but if you relate it to the number 4 it will be 6/4=3/2, or 6=4x3/2" (A VI iii 523: Pk 83). In a world in which everything is constituted of combinations of divine attributes, it is not difficult to think of truth in terms of concept containment.

In April 1679 Leibniz produced a series of papers titled *On the Universal Calculus* that treat a number of questions related to formal validity and in which he first proposes a concept containment account of truth. Underlying these discussions is the idea that an affirmative categorical proposition is true just in case the concept of its predicate is contained in the concept of its subject. He takes true propositions to signify "nothing other than some connection between predicate and subject" in the sense that "the predicate is said to be in the subject, or contained in the subject" (A VI iv [A] 197: L 236). In the complexities of the logical papers of the late 1670s, we can discern the development of the fascinating view that a theory of truth for categorical affirmative propositions will settle the truth conditions for all propositions.

SUBJECTS AND TRUTH IN THE *DISCOURSE ON METAPHYSICS*. The *Discourse* of 1686 is also governed by the series of assumptions found in the early works about activity, self-sufficiency, identity, difference, and the nature of substance although some of the terminology has changed. The most original argument in the text concerns what scholars often call the *logical notion* of substance. This account is introduced in one of the most famous paragraphs in Leibniz's writings. He begins §8 of the *Discourse* with a summary: "To distinguish the actions of God from those of creatures we explain the notion of an individual substance." He then makes two new observations. First, he notes that "it is evident that all true predication has some basis in the nature of things and that, when a proposition is not an identity, that is, when the predicate is not explicitly contained in the subject, it must be contained in it virtually." Second, he suggests that from this account of truth it follows that "it is the nature of an individual substance or a complete being … to have a notion so complete that it is sufficient to contain and to allow us to deduce from it all the predicates of the subject to which this notion is attributed" (A VI iv [B] 1539–1540). That is, an individual substance has a complete concept that contains all the predicates that can truly be predicated of it.

From these observations about substance Leibniz drew support for his doctrine of marks and traces: There must be something within each substance in virtue of which every predicate is presently true of it and which also provides the basis for the deduction of all the predicates that will ever be true of it, that is, traces of all the features that it has possessed in the past and marks of all those that it will possess in the future. He then begins § 9 of the *Discourse* by noting that "from this" account of substance follow "several notable paradoxes." Among others he lists the indestructibility of substances and the identity of indiscernibles (A VI iv [B] 1541-42).

SUBJECTS AND TRUTH IN *FIRST TRUTHS*. Roughly four years after the *Discourse*, Leibniz wrote a brief essay, usually titled *First Truths*, in which he presents many of his core ideas in terse logical fashion. Although we now know that *First Truths* was written either during or soon after his year-long stay in Italy (A VI iv [B] 1643), scholars in the early part of the twentieth century assigned the text an earlier date (around 1686), and this encouraged the belief that his metaphysics developed out of his theory of truth rather than the other way round. But even if the metaphysics of substance came first, it is nonetheless significant that he came to see the theory of truth as so fundamental.

In *First Truths* Leibniz begins with the account of truth, explaining that in true propositions, the predicate is "always in the subject." This *inclusion* means that all true propositions are identities, some of which are implicit and others explicit. That is, for some identities (for example, *A* is *AB*), the inclusion in the subject is explicit; for others (for example, Alexander defeated Darius) it is implicit, and a more thorough analysis of the concept *Alexander* is required. He goes on to claim that "a wonderful secret" about the difference between necessity and contingency lies hidden here. He believes that contingency is a matter of implicit inclusion; necessity a matter of explicit inclusion. All truths are a priori in the sense that the concept of the predicate is contained in the con-

cept of the subject. But some of these truths are more explicit than others. Those that are not explicit are contingent. After presenting his theory of truth, he claims first that the principle of sufficient reason *directly follows* from it (A VI iv [B] 1645: AG 31). Having given an account of that principle, he runs through all the major tenets of his metaphysics as though they follow from these considerations. Consistent with the substantial nature assumption, he insists: "No created substance exerts a metaphysical action or influx on any other" because "what we call causes are only concurrent requisites" (VI iv [B] 1647: AG 33).

Leibniz's claim that all true predication involves the containment of the predicate in the subject threatens to collapse the distinction between necessary and contingent truths. His stock response to this threat was to distinguish, as in the *Discourse*, between explicit and virtual containment or, as in *First Truths*, between explicit and implicit inclusion. But many critics (including Arnauld) have not been convinced. What does it mean to say that a predicate is contained in a subject virtually or implicitly rather than explicitly? His principal answer to this question, probably developed in the late 1680s in part as a reaction to Arnauld's objections, relies upon a distinction between finite and infinite analysis. Necessary truths are those where the containment of the predicate in the subject is revealed after only finitely many steps of conceptual analysis; a corresponding analysis in the case of a contingent truth would require infinitely many steps and cannot be completed by any finite mind. Only God can see to the end of an infinite analysis. Though some scholars have suggested that this infinite-analysis account of contingency was later abandoned by him, it is to be found in the *Theodicy* (1710) and also in a letter to Louis Bourguet (1678–1742) written in the last year of his life.

Infinite analysis, though it provided Leibniz with a way of distinguishing necessary and contingent truths, raised difficulties for his project of developing the universal characteristic: If contingent truths required an infinite analysis to show that a predicate is contained in the concept of its subject, then even if conceptual connections could be represented numerically, the calculations required to demonstrate them could not be carried out, at least not by any finite mind. He seems largely to have given up on the project after 1690. In the *Monadology* he makes the distinction this way:

> There are also two kinds of *truths*, those of *reasoning* and those of *fact*. The truths of reasoning are necessary and their opposite is impossible; the truths of fact are contingent, and their opposite is possible. When a truth is necessary, its reason can be found by analysis, resolving it into simpler ideas and simpler truths until we reach the primitives." (§33)

First Truths derives another typical Leibnizian doctrine, that there are no purely relational properties, from the concept-containment account of truth: "*There are not purely extrinsic denominations.* … For it is necessary that the notion of the subject denominated contain the notion of the predicate. And consequently, whenever the denomination of a thing is changed, there must be a variation in the thing itself." Here the metaphysical presuppositions that lie behind the notion of substance as self-sufficient extend, through the theory that truth consists in conceptual containment, to cover all predications whatsoever. Another Leibnizian doctrine follows immediately: "Every individual substance contains in its perfect notion the entire universe and everything that exists in it, past, present, and future. For there is no thing on which one cannot impose some true denomination from another thing, at the very least a denomination of comparison and relation." It is not surprising that presented with this text, Russell was inclined to see the theory of truth as the heart of Leibniz's mature philosophy. But even in that text, he remarks of the claim that there are no purely relational properties that: "I have shown the same thing in many other ways, all in harmony with one another" (VI iv [B] 1646: AG 32–33).

UNITY AND AGGREGATES. For Leibniz, one of the main goals of the *Discourse* and related texts is to tempt philosophers such as Arnauld away from Cartesianism and toward the metaphysics of (what he will soon call) preestablished harmony. It is not surprising, therefore, that he is keen to note the various weaknesses of the Cartesian account of corporeal substance. As a means to this goal, he is concerned to show that something whose essence consists merely of *res extensa* is inadequate as a substance. He develops an argument for his account of corporeal substances that has roots in his early views and that highlights a weakness in the Cartesian account of corporeal substance.

Leibniz's early assumption, captured in the principle of substantial activity, is that anything substantial will have its own principle of activity. He also believes that activity alone can generate self-sufficiency and unity. In 1676 he begins to connect self-sufficiency and completeness. He distinguishes substances or "complete things" from bodies or things "with figures." In order to have a "perfect understanding" of a substance, one must only

understand the substance or "complete being" itself. But a "figure is not of this kind, for in order to understand from what a figure of such and such a kind has arisen, there must be a recourse to motion. Each complete being can be produced in only one way: that figures can be produced in various ways is enough to indicate that they are not complete beings" (A VI iii 400: Pk 115). In the 1680s he stresses that there will be something real in extension only if there are self-sufficient, unified things. He also begins to describe bodies as aggregates or collections of substances and to distinguish them from a real, single substance. He summarizes the point in 1690: "A BODY [*sic*] is not a substance but an aggregate of substances, since it is always further divisible, and any given part always has another part, to infinity." Therefore: "It is contradictory to hold that a body is a single substance, since it necessarily contains in itself an infinite multitude, or an infinity of bodies, each of which, in turn, contains an infinite number of substances." From this it follows that:

> Over and above a body or bodies, there must be substances, to which true unity belongs. For indeed, if there are many substances, then it is necessary that there be one true substance. Or, to put the same thing another way, if there are many created things it is necessary that there be some created thing that is truly one. For a plurality of things can neither be understood nor can exist unless one first understands the thing that is one, that to which the multitude necessarily reduces." (Foucher de Careil 319: AG 103)

Arnauld wonders what constitutes the difference between a corporeal substance or unity and an aggregate. in response leibniz insists in his letter of April 1687 that some individuals are fundamental but others are not. The latter are aggregates, which are divisible, destructible, and temporary. They admit of degrees in the sense that they can be more or less unified and more or less divisible (e.g., a pile of rocks is more divisible than a piece of marble). The former are substances, which have a substantial forms, each of which creates a *living unity*. There is no reality to an aggregate above and beyond the reality of the entities that make it up. He insists that the unity that bodies or aggregates have is *imaginary*; a perceiving mind may see them as though they were a single thing. He writes to Arnauld that aggregates "have their unity in our mind only, a unity founded on the relation or modes of true substances" (G II 97: AG 86). Aggregates are logical constructions from modes and states of the entities aggregated.

As scholars have long noted, neither the *Discourse* nor the correspondence with Arnauld contains a clear account of exactly how a substantial form confers unity and identity on its substance. But the underlying assumption here, consistent with Leibniz's original views about self-sufficiency and the unifying powers of mind-like things, is that a substantial form confers unity and identity on its substance by acting constantly in relation to its passive principle. In the 1680s he believed that the human soul acts on its body by *concomitance* where the idea is that the two act in perfect preestablished parallelism. He writes in 1690:

> Hence, since I am truly a single indivisible substance, unresolvable into many others, the permanent and constant subject of my actions and passions, it is necessary that there be a persisting individual substance over and above the organic body. This persisting individual substance is completely different from the nature of body, which, assuming that it is in a state of continual flux of parts, never remains permanent, but is perpetually changed." (Foucher de Careil 320: AG 104)

MIND-BODY UNION AND PREESTABLISHED HARMONY. There are reasons to believe that Leibniz understood the relation between mind and body in terms of preestablished harmony as early as the 1670s. But it is not until the texts of the 1690s that he put this account of union front and center. In *A New System of the Nature and Communication of Substances, and of the Union of the Soul and the Body*, published anonymously in the *Journal des Savants* in 1695, he offers his account as an improvement over that of Descartes. He explains that it was the problem of "the union of soul and body" that led him to reject Descartes's philosophy and to recognize the need to "rehabilitate the substantial forms" (G IV 482–483: AG 142–143).

Here we have yet another approach to the core metaphysics, cleverly constructed to engage his audience—many of whom would have been quite interested in Cartesianism of one sort or another—on one of the weakest elements in the Cartesian system. The rhetorical hook here is that Cartesian dualism cannot adequately account for the mind-body union whereas preestablished harmony can. In the *New System* Leibniz declares that the great benefit of his metaphysics is that it offers a neat account of the world while at the same time explaining mind-body interaction. Because "it is not possible for the soul or any other true substance to receive something from without," the mind acts out of its own "depths," but

with perfect "spontaneity" and in perfect "conformity" to everything external to it, including the substances that make up its body. While each substance expresses the whole universe in its own way, the soul is related to the "organized mass that is its body" more "closely" than to other external things. Both the soul and the substances that constitute its body will express one another more closely than they do other "external" things. He concludes that this "hypothesis" displays "the marvelous idea of the harmony of the universe and the perfection of the works of God" (G IV 485–486: AG 143–144).

According to Leibniz the solution to the problem of the interaction between mind and body resides in the harmony constructed by God between the mind and its body. The mind wills to move its finger and the finger moves in perfect preestablished coordination. As he famously puts it, they are coordinated like two clocks constructed "from the start with so much skill and accuracy that one can be certain of their subsequent agreement." Their "sympathy" is guaranteed by the "divine artifice" that has given each substance its "very own law … from the beginning" (G IV 498–499: AG 148). In the *Monadology*, he writes: "According to this system, bodies act as if there were no souls (though this is impossible); and souls act as if there were no bodies; and both act as if each influenced the other" (G VI 621: §81).

METAPHYSICS OF SUBSTANCE, MONADOLOGY

Scholars generally agree that by the time of the *Monadology*, Leibniz holds that the created world is constituted entirely of mind-like monads and that extended things are phenomenal. But there has been a good deal of discussion about when Leibniz gave up the extended substances of his youth. Some scholars have claimed that when he began to construct his own philosophical ideas they were based on a version of mental monism while others have dated the commitment to phenomenalism to the *Discourse* and the correspondence with Arnauld. Until all the writings of the period 1690–1716 have been thoroughly edited and published, there is little chance of solving this mystery. But whenever the phenomenalism begins, there can be no doubt that the notion of corporeal substance plays a key role in the *Discourse* and correspondence. Whether the passivity in such substances is constituted of extended force or collections of mind-like substances, there are corporeal substances constituted of active and passive principles. At some point after 1700, he seems to have become less convinced that the basic entities of the world should be modeled on organisms conceived as combinations of substantial forms and passive principles. In the late 1690s, perhaps in response to criticisms leveled by Arnauld, he begins to emphasize the simplicity of substances, which he now sometimes calls monads, and to reduce everything in the world to these simple, mind-like monads and their perceptions. He writes to De Volder: "Considering the matter carefully, it must be said that there is nothing in the world except simple substances and in them, perception and appetite" (G II 270: L537).

MONADOLOGY. While he was in Vienna, Leibniz wrote this, the most famous of all his works, three years before his death. Written for a friend, he intended it as a summary of his philosophy. Although he did not publish it during his lifetime, generations of scholars have taken it to be the most complete and accurate account of his philosophy. He begins the work with a series of definitions: The monad is "a simple substance that enters into composites—simple, that is, without parts." Monads are the "true atoms" or "elements" of nature and can form aggregates. The activities of monads are of two sorts; they have perceptions and appetitions. "The passing state which involves and represents a multitude in the unity or in the simple substance is nothing other than what one calls perception"; "The action of the internal principle which brings about the change or passage from one perception to another can be called *appetition*" (§14, §15). Although there is a good deal of discussion among scholars about the notion of appetition, it seems closely related to the *reason* or rule of action of the early period. It is the internal feature of the substance that drives it forward, determining its next state on the basis of its present state.

The monad itself may be taken to be another version of his original notion of substance as what is fundamentally unified and self-sufficient. In a related text of 1714, he explains that the Greek term "*monas* signifies unity, or what is one" (G VI 598: AG 207). While there is no doubt that many of the terms and some of the details are new, much of the text merely explicates standard Leibnizian doctrines. We find the various assumptions whose inspiration was originally Platonist. Each monad is an emanation of God, offers a unique perspective on the world, mirrors the universe, and is an indestructible and eternally active thing. He writes: "[Human] minds are images of the divinity itself, or of the author of nature, capable of knowing the system of the universe … each mind being like a little divinity in its own realm" (§83). We find the commitment to the assumptions whose source was Aristotelian: The self-sufficiency of substance now makes

them *windowless*, but they constitute the fundamental entities whose natures anchor the theory of truth, the notion of a complete substance, the expression theory, the perfect coordination and harmony among things. Because each simple substance has its own entelechy, they can act as "the sources of their internal actions" (§18). Because "every present state of a simple substance is a natural consequence of its preceding state, the present is pregnant with the future" (§22).

Thus, the *Monadology* fits neatly into the sometimes subtle but always interesting evolution of Leibniz's views about substance. From the late 1660s to the last years of his life, these fundamental entities constitute the basis for his account of nature. And regardless of the evolution of his ideas about substance, he persists in seeing them as a perfectly rational and divine ordained harmony.

SUMMARY

Few thinkers in the history of philosophy have written so much, thought so deeply, and contributed so profoundly to so many areas. The vastness of Leibniz's texts, the difficulty of his thought, and the quirkiness of some of his ideas make him both a difficult and delightful philosopher to study. As more and more of his works are published, there will be more gems to discover and more interconnections to discern. Not only does Leibniz offer profound philosophical insights, he is admirable as someone who thought deeply about the history of philosophy and the need for intellectual and political peace. As he wrote at the end of his life: "I have tried to uncover and unite the truth buried and scattered under the opinions of all the different Philosophical Sects, and I believe that I have added something of my own which takes a few steps forward. … I flatter myself to have penetrated into the Harmony of these different realms" (G III 606: L 655).

See also Aristotle; Arnauld, Antoine; Augustine, St.; Bayle, Pierre; Boyle, Robert; Cartesianism; Clarke, Samuel; Conway, Anne; Descartes, René; Epistemology; Ficino, Marsilio; Fontenelle, Bernard Le Bovier de; Foucher, Simon; Galileo Galilei; Gassendi, Pierre; Hippocrates and the Hippocratic Corpus; Hobbes, Thomas; Kabbalah; Locke, John; Luther, Martin; Malebranche, Nicolas; Metaphysics; Newton, Isaac; Philosophy; Pico Della Mirandola, Count Giovanni; Plato; Russell, Bertrand Arthur William; Spinoza, Benedict (Baruch) de; Thomasius, Christian; Thomism; Tschirnhaus, Ehrenfried Walter von; Voltaire, François-Marie Arouet de; Wolff, Christian.

Bibliography

ABBREVIATIONS

A: Akademie der Wissenschaften, eds. *Gottfried Wilhelm Leibniz: Sämtliche Schriften und Briefe*. Berlin: Akademie Verlag, 1923–. (Capital roman numerals represent series number; lower case roman numerals represent volume number; arabic numerals represent page number).

AG: Ariew, Roger, and Daniel Garber, eds. *G. W. Leibniz: Philosophical Essays*. Indianapolis, IN: Hackett, 1989.

G: Gerhardt, C. I., ed. *Die Philosophischen Schriften von Leibniz*. 7 vols. Berlin: Wiedmann, 1875–1890. Reprinted, Hildesheim: Olms, 1965.

GM: Gerhardt, C. I., ed. *Mathematische Schriften*. 7 vols. Berlin: A. Asher/Halle: H. W. Schmidt, 1848–63. Reprinted, Hildesheim: Olms, 1962.

Guhrauer: Guhrauer, G. E., ed. *Leibniz' Deutsche Schriften*. 2 vols. Vol. I, 410. Berlin:1838–1840.

L: Loemker, Leroy E., ed. *G. W. Leibniz: Philosophical Papers and Letters*. 2nd ed. Dordrecht, Netherlands: Reidel, 1969.

Pk: Parkinson, G. H. R., ed. *G.W. Leibniz: De summa rerum: Metaphysical Papers 1675–76*. New Haven, CT: Yale University Press, 1992.

EDITIONS AND TRANSLATIONS

The standard of Leibniz's original texts is the Academy edition (A above). The projected completion date is 2050. For a discussion of the editorial project, see Christia Mercer on "Gottfried Wilhelm Leibniz: *Sämtliche Schriften und Briefe*, edited by Akademie der Wissenschaften, Berlin: Akademie Verlag, 1923–, Series VI, volume 4," in the *Times Literary Supplement* Oct. 18, 2002, 7–9.

Other than the Academy edition, the best editions of original texts are G and GM (above). Also helpful are:

Works

Opera Omnia. 6 vols., edited by Ludovici Dutens, Geneva: De Tournes, 1768. Reprinted, Hildesheim: Georg Olms, 1989.

Leibniz' Deutsche Schriften. 2 vols., edited by G. E. Guhrauer, Berlin: 1838–1840.

Nouvelles lettres et opuscules inédits de Leibniz, edited by Alexandre Foucher de Careil. Paris: Ladrange, 1857.

Textes inédits d'après des manuscrits de la Bibliothèque provinciale d'Hanovre. 2 vols., edited by Gaston Grua. Paris: Presses Universitaires de France, 1948.

Leibniz Korrespondiert mit China: Der Briefwechsel mit dem Jesuiten Missionaren (1689–1714), edited by Rita Widmaier, Frankfurt: Klostermann, 1990.

Standard English-language translations of texts are L and AG (above) and the following:

Translated Works

Selections, edited by P. P. Wiener, New York: Scribner's, 1951.

New Essays on Human Understanding, edited and translated by Peter Remnant and Johathan Bennett. Cambridge, U.K.: Cambridge University Press, 1981.

De Summa Rerum: Metaphysical Papers 1675–1676, edited by G. H. R. Parkinson. New Haven, CT: Yale University Press, 1992.

Writings on China, edited and translated by Daniel J. Cook and Henry Rosemont Jr. Chicago: Open Court, 1994.

G. W. Leibniz: Philosophical Texts. Edited and translated by R. S. Woolhouse and Richard Francks, Oxford: Oxford University Press, 1998.

The Labyrinth of the Continuum: Writings on the Continuum Problem, 1672–1676, edited and translated by Richard T. W. Arthur. New Haven, CT: Yale University Press, 2001.

Confessio Philosophi: Papers concerning the Problem of Evil: 1671–1678. In *The Yale Leibniz*, edited by R. C. Sleigh Jr. New Haven, CT: Yale University Press (forthcoming)

PROMINENT SECONDARY LITERATURE IN ENGLISH

Adams, Robert M. *Leibniz: Determinist, Theist, Idealist*. Oxford: Oxford University Press, 1994.

Aiton, Eric J. *Leibniz: A Biography*. Bristol, U.K.: Adam Hilger, 1985.

Broad, C. D. *Leibniz: An Introduction*. Cambridge, U.K.: Cambridge University Press, 1975.

Brown, Stuart, *Leibniz*. Sussex, U.K.: Harvester, 1984.

Brown, Stuart, ed. *The Young Leibniz and his Philosophy: 1646–1676*. Dordrecht, Netherlands: Kluwer Academic, 1999.

Coudert, Allison. *Leibniz and the Kabbalah*. Boston: Kluwer Academic, 1995.

Deleuze, Gilles. *The Fold: Leibniz and the Baroque*. Translated by Tom Conley. Minneapolis: University of Minnesota Press, 1993.

Garber, Daniel. "Leibniz and the Foundations of Physics: The Middle Years." In *The Natural Philosophy of Leibniz*, edited by D. Okruhlik and J. R. Brown, 27–130. Dordrecht, Netherlands: D. Reidel, 1985.

Hartz, Glenn. "Leibniz's Phenomenalisms." *The Philosophical Review* 101 (1992): 511–549.

Hartz, Glenn. "Why Corporeal Substances Keep Popping Up in Leibniz's Later Philosophy." *British Journal for the History of Philosophy* 6 (2) (1998): 192–207.

Ishiguro, H. *Leibniz's Philosophy of Logic and Language*. 2nd ed. Ithaca, NY: Cambridge University Press, 1990.

Jolley, Nicholas, ed. *The Cambridge Companion to Leibniz*. New York: Cambridge University Press, 1995.

Jolley, Nicholas. "Leibniz and Phenomenalism" *Studia Leibnitiana* 18 (1986): 38–51.

Jolley, Nicholas. *The Light of the Soul: Theories of Idea in Leibniz, Malebranche, and Descartes*. Oxford: Clarendon Press, 1990.

Kulstad, Mark, "Causation and Pre-established Harmony in the Early Development of Leibniz's Philosophy." In *Causation in Early Modern Philosophy: Cartesianism, Occasionalism, and Pre-established Harmony*, edited by Steven Nadler, 93–118. University Park, PA: Pennsylvania State University Press, 1993.

Kulstad, Mark. *Leibniz on Apperception, Consciousness, and Reflection*. Munich: Philosophia, 1991.

Leclerc, Ivor, ed. *The Philosophy of Leibniz and the Modern World*. Nashville, TN: Vanderbilt University Press, 1973.

Lodge, Paul, ed. *Leibniz and His Correspondents*. Cambridge, U.K.: Cambridge University Press, 2004.

Lodge, Paul. "Leibniz's Commitment to Pre-established Harmony in the Late 1670s and Early 1680s." *Archiv für Geschichte der Philosophie* 80 (3) (1998): 292–320.

Loemker, Leroy E. "Leibniz and the Herborn Encyclopedists." *Journal of the History of Ideas* 22 (1961): 323–338.

Loemker, Leroy E. "Leibniz's Conception of Philosophical Method." In *The Philosophy of Leibniz and the Modern World*, edited by Ivor Leclerc, 135–157. Nashville, TN: Vanderbilt University Press, 1973.

Loemker, Leroy E. *Struggle for Synthesis*. Cambridge, MA: Harvard University Press, 1972.

Mates, Benson, *The Philosophy of Leibniz: Metaphysics and Language*. Oxford: Oxford University Press, 1986.

McCullough, Lawrence B. *Leibniz on Individuals and Individuation*. Dordrecht, Netherlands: Kluwer Academic, 1996.

Mercer, Christia. *Leibniz's Metaphysics: Its Origins and Development*. New York: Cambridge University Press, 2001.

Mercer, Christia, with Robert C. Sleigh Jr. "Metaphysics: The Early Period to the *Discourse on Metaphysics*." In *The Cambridge Companion to Leibniz*, edited by Nicholas Jolley, 67–123. New York: Cambridge University Press, 1995.

Mondadori, Fabrizio. "A Harmony of One's Own and Universal Harmony in Leibniz's Paris Writings." *Studia Leibnitiana Supplementa* 18 (1978): *Leibniz à Paris (1672–1676)* 151–168.

Mondadori, Fabrizio. "Mirrors of the Universe." In *Leibniz: Die Frage nach Subjektivitat*, edited by Renato Cristin, 83–106. Stuttgard: Franz Steiner Verlag, 1994.

Mondadori, Fabrizio, "On Some Disputed Questions in Leibniz's Metaphysics." *Studia Leibnitiana* 25 (2) (1993): 153–173.

Mondadori, Fabrizio, "Reference, Essentialism, and Modality in Leibniz's Metaphysics." *Studia Leibnitiana* 5 (1) (1973): 73–101.

Nadler, Steven, ed. *Causation in Early Modern Philosophy*. University Park: Pennsylvania State University Press, 1993.

Parkinson, G. H. R. *Logic and Reality in Leibniz's Metaphysics*. Oxford: Clarendon Press, 1965.

Rescher, Nicholas. *Leibniz: An Introduction to his Philosophy*. Totowa, NJ: Rowman and Littlefield, 1979.

Rescher, Nicholas. *Leibniz's Metaphysics of Nature*. Dordrecht, Netherlands: D. Reidel, 1981.

Riley, Patrick. *Leibniz' Universal Jurisprudence: Justice as the Charity of the Wise*. Cambridge, MA: Harvard University Press, 1996.

Russell, Bertrand. *A Critical Exposition of the Philosophy of Leibniz*. Northampton: John Dickens. 1967.

Rutherford, Donald. *Leibniz and the Rational Order of Nature*. New York: Cambridge University Press, 1995.

Sleigh, R. C. Jr. *Leibniz and Arnauld: A Commentary on Their Correspondence*. New Haven, CT: Yale University Press, 1990.

Wilson, Catherine. *Leibniz's Metaphysics: A Historical and Comparative Study*. Princeton, NJ: Princeton University Press, 1989.

Wilson, Catherine. *The Invisible World: Early Modern Philosophy and the Invention of the Microscope*. Princeton, NJ: Princeton University Press, 1995.

Woolhouse, Roger. *Descartes, Spinoza, Leibniz: The Concept of Substance in the Seventeenth Century Metaphysics*. London: Routledge, 1993.

Woolhouse, Roger, ed. *Leibniz's "New System" (1695)*. Florence: Leo S. Olschki Editore, 1996.

For recent literature in other languages, see Christia Mercer's *Leibniz's Metaphysics* and issues of the *Leibniz Review*.

ADDITIONAL RESOURCES

Alsted, Johann Heinrich. *Encyclopaedia, septem tomis distincta*. Herborn: 1630.

Ficino, Marsilio. *Theologia Platonica, de immortalitate animorum*. Paris: 1559. Reprinted in 1995 in Hildesheim, Germany, by Georg Olms Verlag; p. 43. There is a translation of some of this material by Luc Deitz in *Cambridge Translations of Renaissance Philosophical Texts*, edited by Jill Kraye, pp. 30–36.

Ficino, Marsilio. *Platonic Theology*. 2 vols. Volume 1 edited by J. Hankins, translated by M. J. B. Allen. Cambridge, MA: Harvard University Press, 2002.

Locke, John. *Essay concerning Human Understanding*. London: 1690.

Newton, Isaac. *Principia mathematica philosophiae naturalis*. London: 1687.

Newton, Isaac. *The* Principia*: Mathematical Principles of Natural Philosophy*. Translated by I. Bernard Cohen and Anne Whitman. Berkeley: University of California Press, 1999.

Nizolio, Mario. *De Veris Principiis et Vera Ratione Philosophandi contra Pseudophilosophos*. Parma, Italy: 1553. There was an edition edited by G. W. Leibniz published in Frankfort in 1670.

Pico della Mirandola, Giovanni. *De Hominis Dignitate, Heptaplus, De Ente et Uno, e Scritti Vari*, edited by Eugenio Garin. Florence: Vallecchi,1942.

Pico della Mirandola, Giovanni. *On the Dignity of Man, On Being and One, Heptaplus*, Translated by Charles G. Wallis, Paul J. W. Miller, and Douglas Carmichael. Indianapolis, IN: Bobbs-Merrill, 1965.

Plato *Phaedo*. In *Plato, Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.

Plotinus. *Enneads*. Translated by A. H. Armstrong. Cambridge, MA: Harvard University Press, 1990.

Spinoza, Benedictus de. *Ethica*. In *Opera*. Vol. 2, edited by Carl Gephardt, Heidelberg: C. Winter, 1925.

Voltaire, François-Marie Arouet de. *Candide*. 1759.

Christia Mercer (2005)

LENIN, VLADIMIR IL'ICH

(1870–1924)

Lenin was a Marxist revolutionary, Russian Communist political leader, and major contributor to the philosophy of dialectical materialism. Although his mentor Georgii Valentinovich Plekhanov is considered the father of Russian Marxism, Lenin's distinctive version of the doctrine (later dubbed *Marxism-Leninism*) was considered authoritative by the Soviet Communist leadership and had an immense impact on Russia and the world through most of the twentieth century.

LIFE

Lenin was born Vladimir Ilich Ulyanov into the family of a well-to-do school official in Simbirsk, Russia. He enrolled in the University of Kazan in 1887, the same year his elder brother Alexander was executed for involvement in a plot to kill Tsar Alexander III. Lenin was soon expelled from the university for taking part in student disturbances, but he gained admission to the University of St. Petersburg as an external student and in 1892 graduated with a degree in law. His activity in Marxist and other radical circles, beginning in 1888 in Kazan and Samara and continuing in St. Petersburg from 1893, led to his imprisonment in 1895, followed by banishment to eastern Siberia in 1897. Allowed to leave Siberia in 1900, he promptly fled to western Europe. For most of the next seventeen years he worked in various locations outside Russia, writing and conspiring with fellow Russian Marxists to promote the overthrow of the tsarist regime in their homeland.

From the time of the formation of the Russian Social Democratic Labor Party in 1898, Lenin worked tirelessly to gain control of the group and mold it into a militant Marxist revolutionary force. His ideas and aggressive political tactics brought him into bitter conflict with other leading Russian Marxists, including Plekhanov, but eventually his Bolshevik faction of the party became dominant (the Russian term *bol'shevik* means a member of the majority) and in continuing intraparty struggles he consolidated his position as both a theoretician and a leader. After the February Revolution of 1917 he was able, with the help of German military authorities, to travel to Petrograd, where in the October Revolution he led the Bolsheviks in seizing control of the Russian government. In power for six stormy years, marked by attempted assassination, civil war, famine, and the formation of the Union of Soviet Socialist Republics, Lenin suffered a series of strokes beginning in 1922 and died in January 1924.

PHILOSOPHICAL WRITINGS

Lenin's philosophical activity extended from his student days to the Bolshevik revolution, and throughout this period its character was determined by his dogmatic materialism and his devotion to the theory and practice of Marxist social reconstruction as he understood it. His writings are strongly polemical in style, exemplifying the Leninist concept of *partiinost* (partisanship, party spirit) in philosophy.

Lenin first studied the writings of Karl Marx and his colleague Friedrich Engels systematically in 1888 and

1889. One of his earliest works, *What the "Friends of the People" Are and How They Fight the Social-Democrats* (1894)—directed against the Russian Populists, such as Nikolai Mikhailovsky—shows Lenin's general acceptance of dialectical materialism, the materialist conception of history, and the characteristic concepts of Marxist socialism. The distinctively Leninist element already evident is the strong emphasis on action, on the need to combine theory with revolutionary practice. Lenin asserted that the objective, necessary character of the laws of social change in no way destroys the role of active individuals in history. Thus, unlike those Marxists who feared that Russia was not sufficiently developed for a socialist revolution, Lenin stressed the need for expeditiously organizing the revolution, focusing on the proletariat (not the peasantry) as the leading revolutionary class. This activist approach was carried further in subsequent writings, chiefly *What Is to Be Done?* (1902)—the first work in which he used the pseudonym *Lenin*)—and *One Step Forward, Two Steps Back* (1904). In both of these works Lenin elaborated the need for a clandestine, militant, centralized, highly disciplined party to unify and direct the proletariat.

Lenin's book *Materialism and Empirio-Criticism* (1909)—the principal philosophical work published during his lifetime—is directed against a group of Russian writers, including Alexander Bogdanov and Anatoly Lunacharsky, who attempted to supplement Marxism with the phenomenalistic positivism of Richard Avenarius and Ernst Mach. Characterizing their position as a form of subjective idealism (and thus as inimical to Marxism), Lenin defended dialectical materialism on the chief points at issue, particularly the status and character of matter and the nature of knowledge. Opposing the view that matter is a construct of sensations, Lenin argued that it is ontologically primary, existing independently of consciousness. Likewise, space and time are not subjective modes of ordering experience but objective forms of the existence of matter. Opposing the view that discoveries of modern science cast doubt on the objectivity of matter, Lenin distinguished between scientific conceptions of matter, which are provisional and relative because no constituent of a material thing can be regarded as indivisible or irreducible, and the philosophical conception, according to which matter is simply the objective reality known to our senses. The only property of matter to which philosophical materialism is committed, according to Lenin, is the property of existing objectively. In epistemology, Lenin opposed the so-called hieroglyph theory of Plekhanov, according to which sensations are signs of an external reality that they do not necessarily resemble, and developed a strictly realist position, the copy theory, according to which sensations depict or mirror the real world. On this basis Lenin defended the possibility of objective truth, emphasizing practical experience as its test.

Dialectics, which Lenin had long considered the heart of Marxism, is treated most fully in *Philosophical Notebooks* (1933), a posthumous compilation of notebook entries and fragments dating chiefly from 1914 to 1916, including his extracts from, and comments on, a number of works by other thinkers, above all Georg Wilhelm Friedrich Hegel's *Science of Logic*. Lenin showed a high regard for the Hegelian dialectic, which he found thoroughly compatible with materialism, and he asserted that dialectics, logic, and the theory of knowledge are identical. In his conception of dialectics Lenin departed from Engels in laying the greatest stress not on the transition from quantity to quality but on the struggle of opposing (contradictory) forces or tendencies within every natural object and process; Lenin saw this struggle as the internal basis of all change, and thus as the core of dialectics.

Lenin's last major works are concerned mainly with economic and political aspects of the revolutionary transition from capitalism to communism. In *Imperialism, the Highest Stage of Capitalism* (1917) he argued that capitalism had reached its final, monopolistic phase and was ripe for overthrow, but that, because of the uneven development of capitalism in different countries, socialism would not triumph in all or most countries simultaneously, as Marx had expected. In *State and Revolution* (1918), directed against the supposed opportunism of Marxist rivals such as Plekhanov and Karl Kautsky, Lenin elaborated on the Marxist thesis that the state is an instrument of class domination. He laid special stress on a number of points not fully developed by Marx or Engels: One was the need for shattering the bourgeois state machinery and establishing a proletarian state or dictatorship of the proletariat, and another was the distinction between the lower phase of communism, in which people are rewarded in proportion to the work they perform and the state is still needed to repress remnants of the former exploiting classes, and the higher phase, in which rewards are proportional to peoples' needs and the state will wither away because all class antagonisms have been eliminated.

INFLUENCE

Throughout the twentieth-century Communist world (including China), Lenin was regarded as a philosophical

luminary of the first magnitude. In the Soviet philosophical pantheon, he was considered the greatest thinker in history after Marx. Formal education in philosophy in Russia under Communist rule was structured on the premise that Lenin's pronouncements were beyond criticism; his writings were published in vast editions in all major and many minor languages, making him the most widely published philosophical thinker of the twentieth century. In one sense he was also the most influential: Although his conceptual contributions to the development of philosophy as an intellectual discipline were negligible, his ideas provided the impetus and the rationale for policies that materially and often tragically affected the lives of millions of people.

The attack on Stalinism begun by Nikita Khrushchev in 1956 did not immediately disturb the cult of Lenin, whose principles Joseph Stalin was said to have betrayed, not implemented. With the introduction of *perestroika* in the mid-1980s under Mikhail Gorbachev, however, significant responsibility for the flaws and evils of the Soviet system was traced back to Lenin himself; in particular, his theory of imperialism and his fixation on class antagonisms to the neglect of common human interests and moral values were criticized at the highest levels. When the Soviet Union collapsed at the end of 1991, the Communist Party lost political power and Lenin's philosophical authority all but evaporated in Russia and the former Soviet bloc. Thereafter, educational curricula in Russia and Eastern Europe were reworked to eliminate the vestiges of Marxism-Leninism, and the publication of Lenin's writings ceased.

See also Communism; Dialectical Materialism; Engels, Friedrich; Marxist Philosophy; Russian Philosophy.

Bibliography

Harding, Neil. *Lenin's Political Thought*. 2 vols. New York: St. Martin's, 1977–1981.

Kolakowski, Leszek. *Main Currents of Marxism: Its Rise, Growth, and Dissolution.* Vol. 2: *The Golden Age*. Translated by P. S. Falla. Oxford: Clarendon Press, 1978.

Lenin, Vladimir Ilich. *Collected Works*. 45 vols. Moscow: Progress Publishers, 1960–1970.

Reference Index to V. I. Lenin, Collected Works. 2 vols. Moscow: Progress Publishers, 1978–1980.

Scanlan, James P. *Marxism in the USSR: A Critical Survey of Current Soviet Thought*. Ithaca, NY: Cornell University Press, 1985.

Service, Robert. *Lenin: A Biography*. Cambridge, MA: Harvard University Press, 2000.

Service, Robert. *Lenin: A Political Life*. 3 vols. Bloomington: Indiana University Press, 1985–1995.

Wetter, Gustav A. *Dialectical Materialism: A Historical and Systematic Survey of Philosophy in the Soviet Union*. Translated by Peter Heath. London: Routledge and Kegan Paul, 1958.

Wolfe, Bertram D. *Three Who Made a Revolution: A Biographical History*. 4th rev. ed. New York: Dell, 1964.

James P. Scanlan (1967, 2005)

LEONARDO DA VINCI

(1452–1519)

Leonardo da Vinci, the Florentine artist, scientist, and inventor, was born at Vinci in Tuscany, the natural son of a notary, and died near Amboise, France. At his death he left a sizable collection of notebooks that were subsequently scattered in the various libraries of Europe. From 1881 on, many of these notebooks have been published. They consist of notes and jottings on various topics: mechanics, physics, anatomy, physiology, literature, and philosophy. They contain, moreover, plans and designs for machines that frequently have suggested Leonardo's "precursive genius." There are machines of war and of peace, flying machines based on the flight of birds, a parachute, a helicopter, tools and gadgets of all kinds. Leonardo's notebooks are also full of methodological notations on the procedures of scientific inquiry and philosophical considerations about the processes of nature. Undoubtedly many of the arguments that he discussed were taken from the philosophical literature of the time, especially from the writings of the Ockhamists; however, a coherent and complete philosophical scheme cannot be found in the notes, whose chronological order is extremely uncertain. Pierre Duhem held that Leonardo was mainly inspired by the doctrines of Nicholas of Cusa, but recent studies tend to emphasize his dependence on Marsilio Ficino. Leonardo lived in Florence for the first thirty years of his life and subsequently returned there many times.

Leonardo's *Treatise on Painting* (published 1651) reveals the artist and the scientist united in one personality. Painting, which he placed above all other arts, aims at representing the work of nature to the senses. Thus it extends to the surfaces, the colors, and the forms of natural objects, which science studies in their intrinsic forms. The beauty that painting seeks in things is the proportion of the things themselves, and proportion is also the object of the scientific consideration of nature. According to Leonardo, understanding nature means understanding the proportion that is found not only in numbers but also in sounds, weights, times, spaces, and any natural power

whatever. Both art and science have the same object, the harmonious order of nature, which art represents to the senses and science expresses in its laws.

Leonardo held that the two pillars on which science stands are experience and mathematical calculation. As an "unlettered man" (as he called himself) he had contempt for those who, instead of learning from experience, claimed to learn from books (the commentators and followers of Aristotle). He contrasted his work as an inventor with their work of "trumpeting and reciting the work of others." "Wisdom is the daughter of experience," he said. Experience never deceives, and those who lament its deceitfulness should lament their own ignorance because they demand from experience what is beyond its limits. The judgment of experience can be mistaken; and the only way to avoid error is to subject every judgment to mathematical calculation and to use mathematics unrestrictedly to understand and demonstrate the reasons for the things that experience manifests. Mathematics is therefore, according to Leonardo, the basis of all certitude, since without recourse to mathematics it is impossible to put an end to the verbal disagreements of what he called the sophistic sciences—that is, the philosophical disputes about nature.

The privilege accorded to mathematics was most certainly a legacy from Platonism. Leonardo took from Plato's *Timaeus* and Ficino's commentary on it the doctrine that the elements of natural bodies are geometric forms; thus the efficacy of mathematics as an instrument of investigation was justified for him by the fact that nature itself is written in mathematical characters and that only those who know the language of mathematics can decipher it. This is the major contribution that ancient Platonism made to the formation of modern science. Nicolas Copernicus and Galileo Galilei shared this obviously metaphysical doctrine that, however, strongly contributed to launching science from its origins to its mathematical organization. It helped bring scientific consideration from the domain of quality (of natures or essences) to that of quantity by permitting consideration of the natural object as measurable; that is, in the extremes, by reducing the objectivity of nature to its measurability.

However, if the order of nature is a mathematical order, then it is a necessary order; and this necessity is, according to Leonardo, the only true "miracle" of nature: "O wondrous and awesome necessity! With your law you constrain all effects to result from their causes by the shortest path, and according to the highest and irrevocable law every natural action obeys you with the briefest operation." The phrases "by the shortest path" and "with the briefest operation" refer to another feature of the necessary order of nature: its simplicity. Nature follows the shortest or simplest path in its operations. It does not like useless loitering, and this also reveals the mathematical character of its structures. Necessity and simplicity of nature exclude the presence of arbitrary or miraculous forces, as well as the efficacy of magic and of those forces to which it appeals.

Guided by these criteria, Leonardo could arrive at and formulate important theorems and principles of statics and dynamics. The theorem of the composition of forces, the principle of inertia, and the principle of action and reaction are the most notable of these formulations, which, of course, he did not state in the precise form that they received later from René Descartes and Isaac Newton. Nevertheless, they demonstrate his genius for moving from the limited work of the inventor to the generalizations of the scientist.

See also Aesthetics, History of; Copernicus, Nicolas; Descartes, René; Duhem, Pierre Maurice Marie; Ficino, Marsilio; Galileo Galilei; Mathematics, Foundations of; Newton, Isaac; Nicholas of Cusa; Ockhamism; Plato; Platonism and the Platonic Tradition.

Bibliography

Leonardo's manuscripts have been published with photographic reproductions by Charles Ravaisson-Mollien, 6 vols. (Paris, 1881–1891). They have also been published in *Codex Atlanticus*, edited by G. Piumati (Milan, 1894–1904) and *I manoscritti e i disegni di Leonardo da Vinci*, published by the Reale Commissione Vinciana (Rome, 1923–1930). The best collection of selections is J. P. Richter, *The Literary Work of Leonardo da Vinci*, 2 vols. (London, 1883; 2nd ed., 1939). *Leonardo da Vinci on Painting: A Lost Book*, edited and translated from the *Codex Vaticanus Urbinas*, No. 1270, and from the *Codex Leicester* by Carlo Pedretti (Berkeley: University of California Press, 1964), includes a preface by Kenneth Clark and some material never published before.

Works on Leonardo include the following: Pierre Duhem, *Études sur Leonardo da Vinci*, 3 vols. (Paris, 1906–1913); E. Solmi, *Leonardo* (Florence, 1900); C. Luporini, *La mente di Leonardo* (Florence: Sansoni, 1953); Eugenio Garin, *Medioevo e Rinascimento* (Bari: Laterza, 1954), pp. 311ff., and *Cultura filosofica del Rinascimento italiano* (Florence: Sansoni, 1961), pp. 388ff.; I. B. Hart, *The World of Leonardo da Vinci* (London: Macdonald, 1961); James Beck, *Leonardo's Rules of Painting: An Unconventional Approach to Modern Art* (New York: Viking Press, 1979); and Boris Kouznetsov, "The Rationalism of Leonardo Da Vinci and the Dawn of Classical Science," *Diogenes* (69 [1970]: 1–11).

Nicola Abbagnano (1967)
Translated by Nino Langiulli
Bibliography updated by Tamra Frei (2005)

LEONT'EV, KONSTANTIN NIKOLAEVICH
(1831–1891)

Konstantin Nikolaevich Leont'ev was a Russian writer, philosopher, critic, and publicist. Like almost all important nineteenth-century Russian authors, Leont'ev came from a family of landowners. He was trained in medicine at the University of Moscow and served for three years as an army doctor in the Crimean war. After the war he took the post of family doctor on a country estate in the province of Nizhnii-Novgorod, married, and published his first novel, *Podlipki* (1861). In 1863 he entered the Russian diplomatic service and worked for eight years as a consular official on the island of Crete and the Balkans. After a cure from dysentery, he underwent a spiritual crisis and spent a year (1871–1872) in a Greek monastery on Mount Athos. Soon after he left the consular service, and he returned to Russia where he worked as a journalist in various cities and a censor of literature in Moscow. In 1887 he decided to renounce the secular world, was officially divorced from his wife, and retired to the Optyna Pustyn' cloister in the province of Tula. Shortly before his death he took monastic vows and died a monk in the Trinity-St. Sergius Monastery near Moscow.

Although Leont'ev can be considered one of the brilliant representatives of nineteenth-century Russian culture, on a par with Alexander Herzen, his work is not very well known. His novels and stories have hardly been translated and his philosophical and political views only scantily studied. The main reason for this seems to be his odd, maverick-like personality, which expressed itself in views so paradoxical and extreme that it is almost impossible to weld them together and to integrate them with the main ideas of his age.

Leont'ev was torn between an amoral aestheticism and the intense desire for saving his soul by the ascetic renunciation of the world. The protagonist of almost all his novels (among which, apart from *Podlipki, V svoem kraiu* [In my own land, 1864], and *Egipetskii golub'* [The Egyptian dove, 1881–1882]) is a narcissistic superhero (more or less identical with Leont'ev himself) who takes delight in all things beautiful and considers it his duty to lead a poetic life. "Ethics does not coincide with aesthetics: otherwise it is impossible to approve the beauty of Alcibiades, of a diamond, of a tiger." Which is better: "the bloody and spiritually exuberant age of the Renaissance, or contemporary Denmark, Holland, Switzerland—humble, prosperous, moderate?" (*Sobranie sochinenii*, Vol. I, p. 282; 414). However, the hero is dissatisfied with his actual self as he realizes his own limitations and the vanity of his sensuous experience and of the world he has enjoyed so much.

It is this latter attitude that made Leont'ev severely criticize contemporary writers such as Fëdor Dostoevsky, Lev Tolstoy, and Vladimir Solov'ëv. In the essay "*Nashi novye khristiane: F. M. Dostoevskii i graf Lev Tolstoi*" ("Our new Christians: F. M. Dostoevsky and Count Lev Tolstoy, 1882) he ridiculed the *rose-colored Christianity* of these authors. By promising paradise on earth (just like the utopian socialists), Leont'ev stated, they introduced heretical, humanistic elements into their religious views, making God a diluted God of love instead of a God of fear. However, in another essay he made a brilliant analysis of Tolstoy's novels, in particular praising *War and Peace*.

Leont'ev is best known for his aesthetic approach to history and his uncompromising criticism of his own age, which according to him, was dominated by equality and its unavoidable counterpart mediocrity. Just as such thinkers as de Maistre, Comte Joseph de Maistre, Thomas Carlyle, Friedrich Nietzsche, and John Stuart Mill, Leont'ev rejected the industrial revolution of the nineteenth century, which had led to democracies in which there was no place for great men and intense, *creative* contradictions. In his collection of essays *Vostok, Rossiia i slavianstvo* (The East, Russia, and Slavdom, 1885–1886), which included his main work "*Vizantizm i slavianstvo*" (Byzantinism and Slavdom, 1875) he developed a *biological* theory of the evolution of history. Each historical cycle comprises three periods: a period of childhood, or *primitive simplicity*; a second period of adulthood, characterized by differentiation and *flourishing complexity*; and a final period of old age, which through decline and disintegration leads to a *secondary simplicity*.

According to Leont'ev Europe was already in its third phase, the first being the period of the barbarian invasions, the second the High Middle Ages. As clear signs of the contemporary decay, he considered the disappearance of class distinctions and the dominance of bourgeois culture, the culture of the *average man*. Since the time of Peter the Great (1672–1725), this European *leveling interfusion* had infected Russia. Russia's salvation, he maintains, lies in reversing this process, which can only be done by defending its prime institutions, autocracy and orthodoxy, and promoting a situation in which "despotism, danger, strong passions, prejudices, superstitions, fanaticism …, in a word everything to which the nineteenth century is opposed" (*Sobranie sochinenii*, Vol. VIII, p. 98) could flourish. More extreme and reactionary than

the older Slavophiles such as Aleksei Khomiakov and the brothers Ivan and Petr Kireevskii, Leont'ev had no scruples about supporting strict censorship and political repression in order to reverse the *pernicious* process of democratization. However, he with great insight foretold the excrescences of the "fixed equality" of communism, which "through a series of combinations with other principles must gradually lead, on the one hand, to a decreased mobility of capital and property, and, on the other, to a new juridical inequality, to new privileges, to restrictions on individual freedom, and to *compulsory* corporate groups, clearly defined by laws—probably even to new forms of personal slavery or serfdom." (Edie, Scanlan, Zeldin 1965, p. 278).

Leont'ev is often called the Russian Nietzsche. With his pessimistic view on the development of European culture and society, he can be seen as a forerunner of Oswald Spengler. In Russia interest in his work has grown considerably since the 1990s. Biographical data, his complete works, and criticism about him (in Russian) can be found on the web at http://knleontiev.narod.ru.

See also Carlyle, Thomas; Dostoevsky, Fëdor Mikhailovich; Khomiakov, Aleksei Stepanovich; Kireevskii, Ivan Vasil'evich; Maistre, Comte Joseph de; Mill, John Stuart; Nietzsche, Friedrich; Solov'ëv (Solovyov), Vladimir Sergeevich; Spengler, Oswald; Tolstoy, Lev Nikolaevich.

Bibliography

WORKS

Edie, James M., James P. Scanlan, and Mary-Barbara Zeldin, eds. *Russian Philosophy*. Vol. II. Chicago: Quadrangle Press, 1965.

Sobranie sochinenii (Collected works). 9 vols. Moscow: Sablina, 1912–1914. (Reprint of vols. 1–4, Würzburg, Germany: JAL, 1975).

Egyptian Dove: The Story of a Russian. Translated by George Reavey. New York: Weybright and Talley, 1969.

Zapiski otshel'nika (The notes of a hermit). Moscow: Russkaia kniga, 1992.

STUDIES

Berdyaev, N. *Leontiev*. Translated by George Reavey. London: Centenary Press, 1940.

Lukashevich, S. *Konstantin Leontev (1831–1891). A Study in Russian "Heroic Vitalism."* New York: Pageant Press, 1967.

Ivask, Yurii *Konstantin Leont'ev: Zhizn' i tvorchestvo* (Konstantin Leont'ev: life and works). Bern and Frankfurt, Germany: Herbert Lang, 1974.

Rzhevsky, Nicholas "Leontiev's Prickly Rose." *Slavic Review* 35 (1976): 258–268.

Willem G. Weststeijn (2005)

LEOPARDI, COUNT GIACOMO
(1798–1837)

Count Giacomo Leopardi, the Italian poet and prose writer, was one of five children born to Count Monaldo Leopardi and Marquise Adelaide Antici, in Recanati, near Ancona. His brief and anguished existence was plagued both by continuous illnesses (among them rachitis, which made him a hunchback) and the bigotry of his parents, who refused him financial support. A liberal and an agnostic, he yearned to leave the "bodiless, soulless, lifeless" ancestral abode where he had spent all his time devouring books; learning Latin, Greek, Hebrew, and a number of modern languages; and translating and writing critical essays on the classics, history, and astronomy. A fellow philologist, Pietro Giordani, opened to him the world beyond his "savage native town." Afterward, he traveled to Rome, Milan, Bologna, Pisa, Florence, and Naples, never venturing beyond the Alps because of his frail constitution, and even refusing the Dante Alighieri chair offered to him by the University of Bonn. Often he returned to Recanati, only to leave after a short stay. Nature and beauty offered him moments of precious calm, but these few instants could not dispel the physical and metaphysical oppression that, for Leopardi, seemed to weigh upon the world. Everywhere reality proved a bitter disillusionment. Several devoted publishers and friends offered him various jobs and forms of subsistence, but generally to little avail. The poet both expected and invoked death, which came to him in Naples in 1837, shortly after he had dictated his last poem.

THE *CANTI*

As Elme Marie Caro said, Leopardi wanted to be, deserved to be, and was a philosopher. He did not come to philosophy through poetry, or to poetry through philosophy; his poetry is his philosophy. While Leopardi's prose works (the magnificently cogent *Operette morali,* 1827; the diary called the *Zibaldone,* 1898–1900; and the copious correspondence, or *Epistolario,* published posthumously) reflect the melancholy meditations of a thinker concerned with universal sorrow, the most fulfilling expression of his thoughts is to be found in his poetry, the *Canti* (1831, 1835, 1845). The *Canti* complement and complete the *Operette,* because in expression and content they constitute an organic outgrowth of the nature and orientation of Leopardi's philosophy.

PESSIMISM

Leopardi's philosophy, which should not be viewed as a methodically pondered and presented system, has been labeled skeptical and pessimistic, a philosophy of despair. Indeed, it dwells upon the triumph of evil over good and of nature over man, the mystery and insignificance of our mortal existence, the anguish of our miseries, the extinction of youth, and the lure of death. As Arthur Schopenhauer recognized, "No one has treated these subjects more fundamentally and exhaustively in our day than Leopardi." Given the limited dissemination of Schopenhauer's *Die Welt als Wille und Vorstellung* (1819) at that time, it is unlikely that Leopardi read the work or that he met the author. It is certain, however, that Schopenhauer read Leopardi's poems; yet while he mentions them, he in no way indicates whether they influenced the development of his own thought.

Yet the similarities run deep. Leopardi characterized life—this life we love, not for itself but, erroneously, for its promise of happiness—under the rubrics of sorrow (*dolore*), or unhappiness (*infelicità*), and tedium (*noia*). By means of this perspective, he was able to discard many cherished notions. Assuming the hapless state of humanity, the notions of patriotism and heroism vanish as follies, as does the glory of genius, which the poet had once assiduously pursued and which later, like Eduard von Hartmann, he relegated to the category of illusions. As for love and beauty, they entice soul and senses cruelly, since their ephemerality brands them as colossal deceptions. Nature, which according to Leopardi is the mysterious principle of being, closely related to Hartmann's concept of the Unconscious as a neutral absolute, answers none of man's queries about the secret of things; it is undecipherable, mechanical, unreasoning and unreasonable, and at times brutally hostile toward men. Man, then, is nothing; if he is something, he is so by virtue of being his own greatest enemy. In the *Operette morali,* Schopenhauer's gloomy picture of life as a gory chase in which men scramble for spoils differs only moderately from Leopardi's description of Prometheus's and Momus's journey.

Death as nonbeing is therefore, like love during its moment of existence, a thing of beauty. Death as suicide, however, solves nothing because it constitutes not a negation of existence but rather, as Schopenhauer asserted, an act directed against the accidental portion of unhappiness that creeps into human existence. Moreover, the future holds no promise, and *progress* and *perfectibility* are empty words.

EVIL

Leopardian pessimism differs from Schopenhauer's on two questions: the principle of evil and the remedy of evil. Leopardi refused to consider the problem of the necessity of evil and, in any case, would not have ascribed evil to a principle, such as Will or the Unconscious, simply because he believed that evil is an empirical datum and does not require metaphysical or transcendental explanation. He felt the existence of evil and saw only gross arbitrariness in those who attempt to show why it must exist, or who make a transcendent dialectics of the universal law of suffering. Historical pessimism, which stems from the "restless creative mind" of men who boldly oppose unconquerable nature, and cosmic pessimism, through which evil, inherent in nature, subjugates man, are fundamentally interrelated in Leopardi's philosophy and preclude all thought of remedy. The individual's only recourse is stoic dignity—resignation, silence, and scorn. "Of what value is our life, except to despise it?" In this respect, Leopardi was a precursor of German pessimism.

Schopenhauer also upheld Stoic dignity, but for Leopardi dignity was less a remedy for suffering than an instinctive and protective reaction that neither alters suffering nor consoles the sufferer. Schopenhauer even found some consolation in the Buddhist ideal of nirvāṇa, which Leopardi could not. And while Schopenhauer could derive a sense of pride from his belief that the more developed the organism, the greater its misery, Leopardi, even when speaking of man's nobility, could not find in it any basic gratification. The degree to which both men felt a sense of compassion differed: Leopardi's pity, although less central to his ethics than *Mitleid* was to Schopenhauer's, was still less condescending and more sympathetic than Schopenhauer's.

Leopardi held to the inexorability of destiny and nature's blind subservience to it—subservience which fails to take into account man's struggle and misery. Everything, therefore, is deceit; the only truth lies in nothingness. For Leopardi, what counts is the philosophical negation of life, both in its effective pains and in its false felicities. Only in this way can one claim to demonstrate moral consistency—through the affirmation of a negative totality.

ILLUSIONS AND REALITY

Reason, then, in Leopardi is tantamount to negation. Illusions are merely dreams, substances insofar as they may be considered "essential ingredients" of living, "half-real things." Since all that is real comes to nothing, Leopardi

inverted the concept of reality and asserted that only the illusory is real. In claiming this, he did not suggest that reality is a mere phenomenon concealing a noumenon. On the contrary: The reality of the world in which man lives and which has meaning for him is neither rational nor spiritual, but natural and imaginary; it is a reality that is necessarily maintained by what we call illusions. Beyond it lies complete negation. Hence Leopardi professed the opposite of the instinctive noumenalism of man's mind. The world is real in relation to the absence of those other substances that we seek under the heading of truth. Just as the world is arbitrary, so men's beliefs, desires, hopes, and "certainties" (justice, science, virtue, freedom, idealism) are merely groundless illusions. Leopardi despised theological, dogmatic, spiritualistic philosophies, along with any form of presumptuous optimism.

RELIGION

The philosophy outlined above precluded religious faith. Leopardi might assent to the Scriptures' theory of man's decadence, but he could not admit Christian Providence or the Resurrection. Yet although he is unhappy (*infelice*), the poet is not irreligious. His "atheism" bespeaks the combined awareness of the necessity and of the absence of God—in short, of the impossibility of hope. Escape into pleasure is self-deceiving ("pleasure is a subjective speculation and is unreal"), for we seek the idea of pleasure more than we seek pleasure itself; indeed, the latter does not exist. The resulting tedium closely approaches Martin Heidegger's *Angst,* which reflects the experience of nothingness.

VALUE OF LIFE

Because Leopardi is an artist and poet, the immensity of his despair loses its bitterness in a melancholy and fraternal contemplation of existence. Despair allows him to understand the value of human life, although in the long run life is a "useless misery." As a measure of exiguous man's infinite desires against the infinity of being, tedium itself (that is, enthusiasm, heroism, and desperation successively experienced and resulting in a sense of nothingness) seemed to him "the greatest sign of grandeur and nobility in human nature." He recognized illusion as a positive value, offsetting negation and "the infinite vanity of all things." This kind of deception is of value to man, since it constitutes his only justifiable pleasure. Despite it, or actually because of it, Leopardi called for brotherly solidarity and compassion, not out of love of God, but out of a desire to combat the cruelty of destiny and of nature.

What Leopardi finally did was to negate negation, thus creating what he called an ultraphilosophy. He developed a philosophy about philosophy (namely, that we should not philosophize) that rejects reason. For, wrote Leopardi, "As [Pierre] Bayle said, in metaphysics and morals reason cannot edify, only destroy." But by denying itself, reason in a sense vindicates its own power and worth. While exposing the pains and infirmities of existence, Leopardi makes us love the very objects of his despair. By glorifying illusion, art, in the pureness of its beauty (which supersedes the misery of all material things), becomes the most important postulate of ultraphilosophy. Art transfigures sorrow and, by not limiting its own strength and freedom, converts that sorrow into human greatness—a greatness that constitutes the triumph of free creative power and of infinite strength.

See also Beauty; Evil, The Problem of; Hartmann, Eduard von; Heidegger, Martin; Illusions; Life, Meaning and Value of; Pessimism and Optimism; Schopenhauer, Arthur.

Bibliography

WORKS BY LEOPARDI

The critical edition of Leopardi's collected works is *Tutte le opere,* edited by Francesco Flora, 5 vols. (Milan, 1937–1949). For English translations, see *Essays, Dialogues and Thoughts,* translated by James Thomson (New York, 1905?); *The Poems of Leopardi,* a translation of all the *Canti* by Geoffrey L. Bickersteth (Cambridge, U.K.: Cambridge University Press, 1923); *Translations from Leopardi* by R. C. Trevelyan (Cambridge, U.K.: Cambridge University Press, 1941); *Giacomo Leopardi: Poems,* translated and with an introduction by Jean-Pierre Barricelli (New York: Las Americas, 1963).

WORKS ON LEOPARDI

For literature on Leopardi, see Giovanni Amelotti, *Filosofia del Leopardi* (Genoa, 1937); Aristide Baragiola, *Giacomo Leopardi: filosofo, poeta e prosatore* (Strasbourg, 1876); Elme Marie Caro, *Le pessimisme au XIXe siècle* (Paris, 1880); Karl Vossler, *Leopardi* (Munich: Musarion, 1923); Giovanni Gentile, *Poesia e filosofia di Giacomo Leopardi* (Florence: Sansoni, 1939); Iris Origo, *Leopardi: A Biography* (London: Oxford University Press, H. Milford, 1935); J. H. Whitfield, *Giacomo Leopardi* (Oxford: Blackwell, 1954); and G. Singh, *Leopardi and the Theory of Poetry* (Lexington: University of Kentucky Press, 1964).

Jean-Pierre Barricelli (1967)

LEQUIER, (JOSEPH LOUIS) JULES
(1814–1862)

(Joseph Louis) Jules Lequier, or Léquyer, the French philosopher, was born at Quintin in Brittany. He was educated there and in Paris at the *collège* of St. Stanislas and the École Polytechnique. An intensely religious though extremely heterodox Roman Catholic, Lequier devoured the literatures of philosophy and theology, and although none of his own work was published during his lifetime, he wrote voluminously and also translated Sir Humphry Davy's autobiography. Jean Wahl has made interesting comparisons between certain aspects of the thought of Lequier and Søren Kierkegaard, although neither could actually have influenced the other. However, Lequier directly influenced Charles Renouvier, who always considered him his "master in philosophy," and through Renouvier he attracted the attention of William James. Renouvier later published Lequier's book, *La recherche d'une première vérité* (Paris, 1865).

Lequier's philosophy aimed at but never achieved systematic wholeness; its essential theses, however, may be restated in four interrelated doctrines. First, Cartesian methodological doubt must be genuine, not feigned, and unless it is employed in good faith, one is likely to err in doubting real evidence, just as, without methodological doubt one is likely to err in allowing unwarranted belief. Accordingly, doubt has no privileged status over belief. Ability to attain truth as well as falsehood must underlie the quest for truth, and freedom is thus a condition of the possibility of knowing truth as well as of being mistaken.

Second, freedom is a "double dilemma." Either causal necessity or freedom is a fundamental truth, and each doctrine must be asserted either necessarily or freely. If necessity is the true doctrine, my affirmation thereof is *eo ipso* necessary, but since neither doubt nor belief relative to evidence would function in that determination, doubt results. If necessity is true but I affirm freedom, then in addition to my inconsistency (for my affirmation is made necessarily), there is only a subjective foundation for knowledge and morality. Given the truth of determinism, erroneous as well as true judgments are necessary, and any supposed distinction between them is illusory. According to the hypothesis of freedom, if I freely affirm global necessity I am fundamentally inconsistent. Finally, if I affirm freedom under the same hypothesis, not only is my affirmation consistent with the hypothesis but I have a foundation for knowledge and morality. Under the double dilemma, the only satisfactory alternative is freely to affirm freedom—Lequier's "first truth." Freedom is essentially the power to add some novel reality to the existing world. Causality must be explained through freedom and not vice versa.

Third, the data that are present to a given event of consciousness arise out of the past relative to that event; they are past actualities but present potentialities for the internal character of that event of consciousness out of which a determining decision is made. Human consciousness is a succession of self-creative events, each of which is given its ancestor selves as well as other data, and each of which is partially *causa sui,* a "dependent independence." Thus, the totality of causal conditions of any human experience does not make this experience necessary, but only possible, while internal decision makes it contingently actual. All choice-making contains some arbitrary element.

Fourth, in extending these doctrines to theology, and taking as axiomatic the concept that freedom, responsibility, and moral and religious values depend upon choice-decisions, Lequier holds that an omniscient God need not know future contingents, since, in relation to any divine experience, they are not yet existent. To be knowable is to be determinate, and if all were known "from eternity," then all would be eternally determinate, and time and choice-making would be illusions. Also, since contingents are unequivocally in part *causa sui,* they are not wholly dependent on divine power. Far from viewing divine power as absolute total control, Lequier insists that the only power worthy of God is the far greater one of creating self-creators. Real choice in the world is incompatible with all-embracing necessity, and it is neither metaphysically requisite nor religiously desirable that God be wholly immutable and eternal. God must have a temporal aspect in order to come to know contingents as they are realized; thus he remains always omniscient in knowing all there is to know. Lequier's theology is thus that of an eternal-temporal being, his omniscience and omnipotence being relative to the irreducible contingency and self-creativity in the world.

Lequier's philosophy bears various striking resemblances to themes in Samuel Alexander, Henri Bergson, Nikolai Berdyaev, Émile Boutroux, William James, Kierkegaard, C. S. Peirce, and A. N. Whitehead.

See also Alexander, Samuel; Berdyaev, Nikolai Aleksandrovich; Bergson, Henri; Boutroux, Émile; Cartesianism; Consciousness; Freedom; James, William;

Kierkegaard, Søren; Peirce, Charles Sanders; Renouvier, Charles Bernard; Whitehead, Alfred North.

Bibliography

For works by Lequier, see *Oeuvres complètes,* edited by Jean Grenier (Neuchâtel: Baconnière, 1952).

Literature on Lequier includes Émile Callot, *Propos sur Jules Lequier* (Paris: Rivière, 1962); Jean Grenier, *La philosophie de Jules Lequier* (Paris: Société d'édition "Les Belles lettres," 1936); Charles Hartshorne and William L. Reese, eds. *Philosophers Speak of God* (Chicago: University of Chicago Press, 1953), the only material by Lequier now available in English; Adolphe Lazareff, *Vie et connaissance* (Paris: Vrin, 1948); Xavier Tilliette, *Jules Lequier ou le tourment de la liberté* (Paris, 1964); and Jean Wahl, *Jules Lequier* (Paris, 1948), which contains an introduction and selections.

Harvey H. Brimmer II (1967)

LE ROY, ÉDOUARD

(1870–1954)

Édouard Le Roy, the French philosopher of science, ethics, and religion, was born in Paris and studied science at the École Normale Supérieure. He passed the *agrégation* examination in mathematics in 1895 and took a doctorate in science in 1898. Le Roy became a *lycée* teacher of mathematics in Paris but was soon drawn to philosophical problems through an interest in the philosophy of Henri Bergson. He succeeded Bergson, to whose thought his own was deeply indebted, as professor of philosophy at the Collège de France in 1921 and was elected to the French Academy in 1945.

In a series of articles titled "Science et philosophie" (*Revue de métaphysique et de morale* 7 [1899]: 375–425, 503–562, 706–731, and 8 [1900]: 37–72), Le Roy took a pragmatic view of the nature of scientific truth, a view more or less shared by his contemporaries Bergson, Jules Henri Poincaré, and E. Wilbois. Scientific laws and even scientific "facts," Le Roy maintained, are arbitrary constructs designed to meet our needs and to facilitate effective action in pursuit of those needs. Scientific reason, in other words, distorts reality in the interests of practical action. The scientific facts on which induction is based are artificially extracted from the continuous flow of happenings and experiences and built up into convenient (rather than "true") thought structures, which constitute "the grammar of discourse" and enable us to talk about, and deal with, what would otherwise be "the amorphous material of the given." Thus, in reacting against scientific mechanism, Le Roy presented an extreme view of mind as the creator of its own reality.

Le Roy took the same pragmatic view of discursive religious truth in *Dogme et critique* (Paris, 1906). His views were supported by the Catholic modernists and condemned as dangerous in a papal encyclical. Le Roy held that the validity of dogmas cannot be proved, nor do they profess to be provable; they depend upon a rigid and externally imposed authority; their expression and frame of reference is that of medieval philosophy; and they are alien to, and incompatible with, the body of modern knowledge. For these four reasons they are unacceptable to the modern mind as truths. Nevertheless, they possess a pragmatic value; they fulfill a purpose, in this case a moral one. "Although mysterious for the intelligence in search of explanatory theories," Le Roy held, "these dogmas lend themselves nonetheless to perfectly specific formulation as directives for action." Christianity is thus not a system of speculative philosophy, but a set of stated or implied injunctions, a way of life. For example, the belief in a personal God demands that our relation to him resemble our relation to a human person. The doctrine of the resurrection of Christ teaches that we should behave in relation to him as if he were alive today.

Le Roy's misgivings concerning religious dogmas arose because the dogmas seemed to him irreconcilable with a homogeneous system of rational knowledge. In a pragmatic and relativist conception of truth such incompatibility should not be significant. However, the criterion of truth, for Le Roy, was neither use nor coherence, but "life" itself, dynamic and self-developing. Scientific theory is useful distortion, religious teaching a source of moral action, and both are arbitrary in their choice of concepts and symbols. Genuine knowledge is a kind of self-identification with the object in its primitive reality, uncontaminated by the demands of practical need. Intuition, not discursive thought, is the instrument of such knowledge, and the criterion of truth is that one should have lived it; otherwise, according to Le Roy, one ought not to understand it. This, as L. Susan Stebbing rightly pointed out, altogether removes the criterion from rational criticism, since life is both truth and the criterion of truth.

Le Roy's philosophy culminated in moral and religious concerns, as is seen in Volume 2 of his posthumously published *Essai d'une philosophie première* (2 vols., Paris, 1956–1958). His position is similar to Bergson's in *Les deux sources de la morale et de la religion.* The élan vital that animates us takes the form of an "open," that is, indeterminate, moral demand. This generalized

obligation is the essence of the self as a free and self-creating agent. Le Roy stated that "to believe is to perceive a spiritual exigency and to act under its inspiration." The open nature of the exigency "beyond any ideal capable of being formulated" places Le Roy's view in the same category as much recent morality of authenticity. The agent is constantly transcending the determinate in the direction of some necessarily unspecified self-fulfillment. Because morality implies precepts and precepts imply universalizability, the notion of a morality that cannot be formulated would seem to be self-defeating. In his conception of a moral quest Le Roy, in fact, seemed to presuppose the Christian values to which he subscribed.

See also Bergson, Henri; Laws, Scientific; Modernism; Philosophy of Science; Poincaré, Jules Henri; Religion; Stebbing, Lizzie Susan.

Bibliography

ADDITIONAL WORKS BY LE ROY

Une philosophic nouvelle: Henri Bergson. Paris: Alcan, 1912.

L'exigence idéaliste et le fait de l'évolution. Paris: Boivin, 1927.

Les origines humaines et l'évolution de l'intelligence. Paris: Boivin, 1928.

Le problème de Dieu. Paris, 1929.

La pensée intuitive. 2 vols. Paris: Boivin, 1929–1930.

Introduction à l'étude du problème religieux. Paris, 1944.

La pensée mathématique pure. Paris, 1960.

WORKS ON LE ROY

Gagnebin, S. *La philosophie de l'intuition. Essai sur les idées d'Édouard Le Roy.* Paris, 1912.

Olgiati, F. *Édouard Le Roy e il problema di Dio.* Milan, 1929.

Stebbing, L. Susan. *Pragmatism and French Voluntarism.* Cambridge, U.K.: Cambridge University Press, 1914.

Colin Smith (1967)

LE SENNE, RENÉ

(1882–1954)

René Le Senne, the French spiritualistic philosopher, was born in Elbeuf in Normandy. From 1903 to 1906 he was a pupil of Frédéric Rauh and Octave Hamelin at the École Normale Supérieure, where he passed the *agrégation* examination in philosophy in 1906. He obtained his doctorate in 1930 with a thesis titled *Le devoir* (Duty). After holding provincial teaching posts he was appointed to the Lycée Louis-le-Grand in Paris and, in 1942, to a chair of moral philosophy at the University of Paris. He distinguished himself as joint editor, with Louis Lavelle, of the series of works published in the collection "Philosophie de l'esprit." In 1948 he was elected to the Académie des Sciences Morales et Politiques.

The conception of philosophy underlying the "Philosophie de l'esprit" was traced by Le Senne to the Cartesian tradition, which, he held, identified existence with the act of thought and regarded existence as dependent upon a transcendent and infinite being. This tradition, according to Le Senne, was threatened both by positivism, which discounts the self-creating principle that raises man above causally determined physical nature, and by an excessive modern subjectivism, which makes man the measure of all things. Against these threats to the French "psycho-metaphysical" tradition Le Senne and Lavelle launched their series, in what they conceived as a kind of philosophicomoral mission, a reassertion of metaphysical philosophy against antiphilosophy.

Like much of recent French thought, Le Senne's work evokes not so much René Descartes as Maine de Biran. The essence of the self is consciousness of action against the resistance and limitation of reality. This could be rendered: I will, or I strive, therefore I am. Thus, personality for Le Senne was "existence as it is formed by the double cogito: hindered by obstacles, elevating itself by and towards value." Man participates in absolute and transcendent value. Although value outruns him and is not wholly his creation, it is made determinate by him in a given, concrete situation.

Reality, then, is at once the organ of self-creation and an obstacle to it. In a sense it degrades value, yet it actualizes value by making it determinate. We are, moreover, called back to awareness of the value-creating source in which we participate. This is a spiritual flow, or upsurge (*essor*). "Some obstacle has to break the continuity of the upsurge before the self, concentrating upon it the body's energy, begins to will." The willing self owes its being and consciousness to the obstacles it encounters. We participate in a world of absolute value and a world of brute reality and create ourselves unceasingly through them.

See also Cartesianism; Descartes, René; Essence and Existence; French Philosophy; Hamelin, Octave; Lavelle, Louis; Maine de Biran; Positivism.

Bibliography

WORKS BY LE SENNE

Introduction à la philosophie. Paris, 1925.

Le devoir. Paris: Alcan, 1930.

Le mensonge et le caractère. Paris: Alcan, 1930.

Obstacle et valeur. Paris: Aubier, 1934.

Traité de morale générale. Paris: Presses Universitaires de France, 1942.

Traité de caractérologie. Paris: Presses Universitaires de France, 1945.

La destinée personnelle. Paris: Flammarion, 1951.

La découverte de Dieu. Paris, 1955.

WORKS ON LE SENNE

Paumen, J. *Le spiritualisme existentiel de René Le Senne.* Paris: Presses Universitaires de France, 1949.

Pirlot, J. *Destinée et valeur. La philosophie de René Le Senne.* Namur, Belgium: Secrétariat des Publications, Facultés Universitaires, 1953.

Vax, L. "Pensée souffrante et pensée triomphante chez René Le Senne." *Critique* 12 (1956): 142–152.

Colin Smith (1967)

LEŚNIEWSKI, STANISŁAW

(1886–1939)

Leśniewski, Stanisław (1886–1939) was one of the founders of the Warsaw School of logic, which flourished from 1919 to 1939. He was the author of a highly original system for the foundations of mathematics, and one of the most innovative and unorthodox logicians of the twentieth century.

LIFE AND INFLUENCE

Leśniewski was born in Serpukhov, Russia, and received his schooling in Irkutsk. After studying at German universities, including Leipzig and Munich, he moved in 1910 to Lwów where he studied philosophy with Kazimierz Twardowski and obtained his doctorate in 1912. Leśniewski published several papers before the First World War, which he spent in Moscow. His preoccupation with the logical antinomies, which began in 1911 when he read Jan Łukasiewicz's book *On the Principle of Contradiction in Aristotle*, shifted his interests permanently from philosophy of language to the logical foundations of mathematics. In 1919 he was appointed professor of the Philosophy of Mathematics at the University of Warsaw. From then until his early death from cancer he was at the center of developments in mathematical logic in Poland, first developing his systems, then from 1927 publishing his results. Leśniewski's notes, correspondence, and a monograph on the antinomies were destroyed in the 1944 Warsaw Uprising: After the war several of his surviving students worked to reconstruct the lost results.

Leśniewski's sole doctoral student Alfred Tarski—Leśniewski boasted proudly of having one hundred percent geniuses as doctoral students—inherited many of his teacher's attitudes, but Tarski's increasing willingness to embrace platonistic set theory for the sake of metamathematical results caused tensions between them. Other pupils such as Jerzy Słupecki, Bolesław Sobociński, Czesław Lejewski, and Henry Hiż remained closer to Leśniewski's views, but their influence was limited. Quine's concern with ontological commitment and the meaning of the quantifiers probably went back to discussions he had with Leśniewski in 1933 on the interpretation of higher-order quantification. Because of the inconvenience of his systems, his forbidding perfectionism, and the idiosyncrasy of his positions, Leśniewski's work remained outside the mainstream, but some aspects became widely influential outside Poland. These include: the object language/metalanguage distinction, exact canons of definition, the theory of semantic categories, and mereology.

FORMAL SYSTEMS

After learning about Russell's Paradox, Leśniewski set himself to produce an antinomy-free foundation for mathematics. Disconcerted by the inexactitudes of Russell's and Whitehead's *Principia mathematica*, he initially forswore logical symbolism and formulated his views in highly regimented Polish, but in 1920 Leon Chwistek persuaded him to formalize his work, which he did with unprecedented precision. The logical order of Leśniewski's three systems is the reverse of the chronological order of their discovery. Leśniewski diagnosed an ambiguity in the notion of class which he made responsible for Russell's Antinomy, and in 1916 developed the theory of concrete classes, later renamed *mereology*. Then he set about formalizing the underlying logic of names, predicates and higher-order functors, which he called *ontology*, axiomatizing it in 1920. Finally he formalized the theory of sentences, connectives, and quantification which underlay the other theories, calling the resulting system *protothetic*. The axiomatization of protothetic was assisted by Tarski's 1923 discovery that conjunction could be defined in terms of material equivalence and universal quantification. Leśniewski and others improved the results through the 1920s, and he published accounts of protothetic in a series of German papers, and mereology in a Polish series.

MEREOLOGY

Mereology (from Greek *meros*, part), a formal theory of the part-whole relation and cognate concepts, is Leśniewski's nominalistically acceptable partial substitute

for set theory. It understands classes as concrete wholes literally composed of their members. Classes (now usually called mereological sums or fusions) are identical when they have the same parts, so the same sum may be determined by different pluralities of members—for example a chess board is the sum of its squares, but also of its ranks or its files. Sums exist if their members do, and a sum need not be spatiotemporally connected. There is no null or empty sum. Mereology can be axiomatized in many ways using many different undefined constants, its axiom(s) being added to ontology. The following perspicuous four-axiom system uses the primitive notion pt( ), meaning *part of*, and understood to include the case of identity or improper part; the variables in this system are all nominal (intended singular variables being capitalized), and the ontological notion presupposed is the copula "ϵ" of singular inclusion:

M1 $\forall AB\ [\ A \epsilon \text{pt}(B) \rightarrow B \epsilon B\ ]$

M2 $\forall ABC\ [\ ((A \epsilon \text{pt}(B)\ \&\ B \epsilon \text{pt}(C)) \rightarrow A \epsilon \text{pt}(C))\ ]$

M3 $\forall Ab\ [\ A \epsilon \text{Sm}(b) \leftrightarrow (A \epsilon A\ \&\ \forall C\ [\ C \epsilon b \rightarrow C \epsilon \text{pt}(A)\ ]\ \&\ \forall D\ [\ D \epsilon \text{pt}(A) \rightarrow \exists EF\ [\ E \epsilon b\ \&\ F \epsilon \text{pt}(E)\ \&\ F \epsilon \text{pt}(D)]\ ]\ )\ ]$

M4 $\forall Ab\ [\ A \epsilon b \rightarrow \text{Sm}(b) \epsilon \text{Sm}(b)\ ]$

These axioms say: (M1) that whatever has a part is an individual; (M2) that parthood is transitive; (M3) define the sum of all the *b*s as that unique individual *A* which overlaps all and only *b*s; and state (M4) that if there is at least one *b* then the sum of all *b*s exists and is unique. Mereology is consistent relative to protothetic. It is independent of this system whether or not there are atoms—that is, objects without proper parts.

Mereology was the first system rigorously formulated by Leśniewski and remains the most thoroughly investigated. Its principles are much weaker than those of set theory, although some of its assumptions, especially the general sum principle M4, have been questioned on philosophical grounds. Especially when based on standard predicate logic rather than Leśniewski's ontology, mereology has come into standard use in ontology and cognitive science.

ONTOLOGY

Mereology presupposes ontology, so called because Leśniewski took it to formulate several meanings of *be*. He intended it as a modernized term logic of the sort formulated by Ernst Schröder, and in its admittance of empty and plural terms it is closer to traditional logic than to Frege-Russell predicate logic, whose terms are all singular. Like mereology, ontology can be based on many different primitives, but the most frequently used is the one chosen by Leśniewski, namely the singular inclusion functor "ϵ." The basic sentence-form "*A* ϵ *b*," readable as "*A* is a *b*" but best read perhaps as "*A* is one of the *b*s" captures the distributive rather than collective sense of "class": "*A* is a member of the class of the *b*s" just means "A is one of the *b*s" and no individual called "the class of the *b*s" is assumed.

Leśniewski's original (1920) axiom, though not the shortest, remains the most perspicuous:

$$\text{O1}\quad \forall Ab\ [\ A \epsilon b \leftrightarrow (\ \exists C\ [\ C \epsilon A\ ]\ \&\ \forall DE\ [\ (D \epsilon A\ \&\ E \epsilon A) \rightarrow D \epsilon E\ ]\ \&\ \forall F\ [\ F \epsilon A \rightarrow F \epsilon b]\ )\ ]$$

This says that *A* is a *b* if and only if (1) there is at least one *A*, and (2) there is at most one *A*, and (3) every *A* is a *b*. This axiomatic equivalence, which constitutes a sort of implicit self-definition of "ϵ" mirrors the analysis of singular definite descriptions by Russell, as can be seen by reading "*A*ϵ *b*" as "the *A* is a *b*." Existential import in ontology is located in the functor "ϵ." rather than the quantifiers: "*A*ϵ *b*" is only true if an *A* exists.

The axiom is not ontology's only source of logical power. Leśniewski allows new constants to be defined, and as these are introduced, new semantic categories of expression and thereby new categories of bindable variable become available. Each category of expression is subject to a principle of extensionality, and so the system grows in logical strength, ascending as required to higher types of variable. Thus although the axiom binds only nominal variables, later theses allow variables for predicates and other higher-order functors to be bound. Because Leśniewski allows plural names, his first-order calculus is equivalent in logical strength to standard monadic second-order predicate calculus. There is no axiom of choice in ontology, but a directive can be formulated allowing choice principles to be stated for each higher logical type (semantic category). Like mereology, ontology is consistent relative to protothetic. Despite its expressive power, ontology is ontologically neutral in that no thesis stating the existence of an individual can be derived. It is thus true of the empty universe as well as others.

Ontology is in many ways Leśniewski's most innovative system, combining features of traditional, Schröderian, and Fregean logic with a potential expressive power

equivalent to that of the simple theory of types. Nevertheless, apart from some exploitation for purposes of historical comparison, and some development by Lejewski and others, it has found few supporters.

PROTOTHETIC

The basis of Leśniewski's logic is protothetic, a bivalent propositional calculus to which may be added propositional functors of any order, and incorporating the theory of quantification. It is equivalent in potential to a system of propositional types. Leśniewski, following Peirce, took quantification as embodying cardinally unconstrained conjunction and disjunction, and as part of the basis of logic rather than attaching primarily to nominal variables. The quantifiers ∀ and ∃ bind variables of any category. Leśniewski experimented from 1921 onwards with different axiomatic and combinatorial bases for protothetic. He chose material equivalence as sole undefined connective because he formulated definitions as object-language equivalences, and he developed the calculus of equivalent statements. But an intuitive axiomatization of prothetic using implication is:

P1 $\forall pq\ [\ p \rightarrow (q \rightarrow p)\ ]$

P2 $\forall pqrf\ [\ f(r\ p) \rightarrow (f(r\ \forall s[s]) \rightarrow f(r\ q))\ ]$

Quantifier apart, the first thesis is familiar from propositional calculus. The second exploits a variable f for functors taking two propositional arguments, with "$\forall s[s]$" a standard false sentence. Like ontology, protothetic derives much of its strength from the rules permitting the formulation of new definitions and extensionality principles for higher types. Each propositional type is finite in its extensions, starting from the basic types of sentences, which has just two extensions, the True and the False, so in principle the quantifiers can be replaced by computational principles running through the extensions for each type considered in a sentence. Leśniewski took great pains over protothetic but it remains the least discussed of his logical systems.

Though his published works covered mainly his own systems, with incidental but incisive criticism of such contemporaries as Russell and Whitehead, von Neumann and Zermelo, in his Warsaw lectures Leśniewski ranged more widely, finding single axioms for general and abelian group theory, developing Peano's axioms, investigating inductive definitions, comparing mereology with Whitehead's theory of events, and criticizing Łukasiewicz's many-valued logic.

PHILOSOPHICAL METALOGIC

Though an unprecedentedly exact formalizer, Leśniewski deplored all formalism. Having come to logic through regimented ordinary language, he understood his logical systems throughout as interpreted with a determinate intended meaning, and intended his theses as general truths. From his first paper Leśniewski scrupulously distinguished use from mention of expressions, and literally failed to understand writings where this distinction was not observed, notably *Principia*. By contrast he admired and extolled the great rigor of Frege's formal systems, notwithstanding their inconsistency. Leśniewski's strictures on quotation were inherited and made influential by Tarski.

Leśniewski criticized Twardowski's platonism and strove to make his logical systems compatible with nominalism. This meant treating systems not as abstract entities but as concrete collections of physical inscriptions, growing in time by the addition of new inscriptions called *theses*. Because the systems as they develop allow new expressions to be introduced via definitions, and new types of variable to become available for quantification, the regulation of their growth had to be precise but schematic. Leśniewski achieved this by formulating for each system regulatory directives allowing new theses to be introduced. These directives are self-adjusting in that what they allow expands as the system grows. Leśniewski considered faulty definition to be responsible for the logical antinomies, and by bringing definitions within the system as object-language equivalences—rather than metalinguistic abbreviations—kept them under tight control. The highly complex directives for adding definitions in protothetic and ontology are Leśniewski's proudest achievement. To formulate them and the other directives governing substitution, quantifier distribution, modus ponens, and extensionality required a sequence of more than fifty metalogical definitions called *terminological explanations*. In his everyday logical working however, Leśniewski used an unofficial system of natural deduction from assumptions, understood as delivering an outline which could, if necessary be transformed into a proof according to the directives. This he never formalized. The complexities of formulating general terminological explanations and directives for variable-binding operators were beyond even Leśniewski, and he had to rest content in his official system with a sole syncategorematic operator, the universal quantifier.

The formulation of the directives employed Leśniewski's notion of semantic categories, a systematic logical grammar inspired by Frege's practice and

Husserl's theory of *Bedeutungskategorien*, and intended as an ontologically parsimonious alternative to type theory. Though not codified by Leśniewski, the subsequent systematization by Ajdukiewicz and others has made this part of mainstream logic and linguistics. The basic category of protothetic was the sentence, to which ontology added the basic category of name. Mereology required no new categories or directives.

Leśniewski had definite ideas about the intellectual economy of logic. A system ought to have as few primitive notions, axioms, and directives as possible; the axioms ought to be as short as possible, logically independent, and organic—that is, not contain provable theses as subformulas. The search for ever shorter axioms was a general feature of the Warsaw School, which Leśniewski and his followers sometimes pursued at the expense of defending controversial aspects of the systems, such as their interpretation of quantification, their radical nominalism, and their thoroughgoing extensionalism.

Leśniewski's avowed metaphysical neutrality combined with his liberal use of quantifiers to bind non-nominal variables drew criticism from Quine. Leśniewski rejected Quine's accusation of platonism, and on reflection Quine came to regard Leśniewski's quantifiers as substitutional, committing not to corresponding entities, but to expressions to be substituted for variables bound by a quantifier. That Leśniewski cannot understand the quantifiers objectually is clear because a standard empty name "$\wedge$" can be substituted normally for bound variables: From the true "no $\wedge$ exists" one may validly infer "for some *a*, no *a* exists," so "for some" ($\exists$) cannot mean "there exists." In the light of its subsequent development, the substitutional interpretation fits Leśniewski no better than the objectual, because it would commit him to an infinity of platonic expression types. Comparison with standard accounts is complicated by Leśniewski's incriptional understanding of expressions and the import of his directives, which are conditional prescriptions rather than categorical descriptions. The directive "If *A* is the last thesis belonging to the system then a thesis *B* may be added if for some thesis *C* preceding *A*, *B* is a result of substitution from *C* into *A*" quantifies only over extant tokens. The question remains how expressions employed in a logical system (including the quantifiers) have their meanings. On this Leśniewski remains silent. How to theorize metalogically about meaning and truth within Leśniewski's strictures remains perhaps the biggest open question concerning his systems.

See also Logical Paradoxes; Łukasiewicz, Jan; Mereology; Syntactical and Semantic Categories; Tarski, Alfred; Twardowski, Kazimierz.

Bibliography

WORKS BY LEŚNIEWSKI

The only collected edition of Leśniewski's works is their English translation:

Leśniewski, Stanisław. *Collected Works*. 2 vols., edited by Stanisław J. Surma, Jan T. J. Srzednicki, and Dene I. Barnett, with an annotated bibliography by V. Frederick Rickey. Dordrecht, Netherlands: Kluwer, 1992. Rickey's bibliography lists the original Polish locations of these writings.

Srzednicki, Jan T. J., and Stachniak, Zbigniew, eds. *S. Leśniewski's Lecture Notes in Logic*. Dordrecht, Netherlands: Kluwer, 1988. Contains edited transcriptions of student notes from lecture courses Leśniewski gave in Warsaw.

WORKS ABOUT LEŚNIEWSKI

Luschei, Eugene C. *The Logical Systems of Leśniewski*. Amsterdam, Netherlands: North-Holland, 1962. The most comprehensive monograph on Leśniewski.

Miéville, Dénis. *Un développement des systèmes logiques de Stanislaw Lesniewski: Protothétique – Ontologie – Méréologie*. Berne, Switzerland: Lang, 1984. The nearest thing to an elementary introduction to Leśniewki.

Miéville, Dénis and Dénis Vernant, eds. *Stanislaw Lesniewski Aujourd'hui*. Grenoble, France: Université Pierre Mendès France, 1995. A collection of essays including an informative survey by Czesław Lejewski, "Remembering Stanislaw Lesniewski," pp. 25–66.

Srzednicki, Jan T. J., V. Frederick Rickey, and J. Czelakowski, eds. *Leśniewski's Systems: Ontology and Mereology*. Dordrecht, Netherlands: Kluwer 1984. Seminal commentaries and improvements by Lejewski, Słupecki, Robert E. Clay and others.

Srzednicki, Jan T. J., and Zbigniew Stachniak, eds. *Leśniewski's Systems: Protothetic*. Dordrecht, Netherlands: Kluwer, 1998. A similar compilation of important commentaries and improvements in protothetic.

Woleński, Jan. *Logic and Philosophy in the Lvov-Warsaw School*. Dordrecht, Netherlands: Kluwer, 1989. Definitive history, with a copious bibliography, situating Leśniewski in the context of the logical movement in Poland.

Peter Simons (2005)

LESSING, GOTTHOLD EPHRAIM

(1729–1781)

Gotthold Ephraim Lessing, the German dramatist and critic, was born at Kamenz in Saxony. The son of a scholarly Lutheran pastor, he was sent to study theology at Leipzig University. There, however, he absorbed the popular rationalism of the Enlightenment, whose leading

contemporary exponent was the Leibnizian Christian Wolff, of Halle. Lessing was influenced in the same direction by his friends from Berlin, Christoph Friedrich Nicolai and Moses Mendelssohn, and by the writings of the English deists, many of which had been translated into German. Although literature, and especially the drama, became Lessing's supreme interest, he was to return to theology in the last decade of his life. He has no special claim to being ranked as a philosopher of originality and distinction, but with regard to the diffusion of certain ideas and attitudes among educated minds, his historical influence is preeminent. He was above all a critic, and his attitude may be described as one of "passionate detachment." His nonconformity made him appear to be perennially restless; he was never permanently satisfied to adopt the conventional opinions of society, always preferring to be in a "minority of one." The movement of his mind carried him beyond his parents' theological beliefs and the commonplace deism of his twenties until, through his invocation of Benedict de Spinoza, he eventually prepared the way for the romantic reaction against the Enlightenment.

LITERATURE AND ART

Lessing's approach to the drama was based on his conviction that it was urgently necessary to break the tyrannical dominance over German literature exerted by the established French classicism—a trend that was encouraged by Frederick II of Prussia. In Lessing's eyes, the effect of this French influence was the suppression of the native German genius. In a series of "literary letters" (*Briefe, die neueste Literatur betreffend*, Leipzig, 1759–1765), written in cooperation with Nicolai and Mendelssohn, Lessing exhorted German writers to turn their backs on the artificial perfections of Pierre Corneille and Jean Racine; he claimed that they should take as their stylistic model the bold naturalism of William Shakespeare, whom Voltaire had characteristically dismissed as a "drunken savage."

Lessing's best-known work of criticism is his *Laokoon, oder, über die Grenzen der Malerei und Poesie* (*Laocoön, or the Bounds of Painting and Poesie*, Berlin, 1766). Judged as constructive thinking about the nature of art, it is a disappointing work, although it is noteworthy in that it contains the first explicit statement of the concept of "art for art's sake." Moreover, its overt thesis—that painting works by forms and colors in space, while poetry belongs to a quite different category in that it sets out to describe successive moments in time—is not only inadequate, since it fails to take account of lyric poetry and indeed of all poetry that describes states of mind, but also much less original than Lessing implied. But it is significant that the *Laokoon* takes the form of a critique of Lessing's German, English, and French predecessors; he could not write well without a target to attack. In the *Laokoon*, Lessing's main critique was directed against Johann Joachim Winckelmann and the latter's idealization of "noble simplicity and quiet grandeur." Lessing was prepared to acknowledge that this ideal may hold good for painting, which, he claimed, is exclusively concerned with the beauty of physical form. But he wholly denied its validity or relevance for judging poetry, which is concerned with action and passion. *Laokoon*, like much that Lessing wrote, has a subtle undercurrent of irony and polemic, the thrust of which, on the surface, is not apparent to the rapid reader. Although Lessing took as his text a famous piece of ancient sculpture, his essay is more an oblique sermon about literature than an aesthetic analysis of the visual arts by a critic with a real understanding of, or even sympathy for, his subject. Its essential thesis is a warning that Winckelmann's neoclassical ideals must not constrict the freedom of the poet, who, unlike the painter, is primarily concerned with passionate action.

Lessing's writings on art and literature do not constitute a serious analysis and critique of aesthetic experience. But his work was directed toward liberating the artist from all the limiting rules and conventions of artificial formality. Lessing was not in any sense a romantic writer, but because of his demand for the free expression of natural feelings and his retrospective interest in antiquity, he occupies an important place among the forces that made German romanticism possible. The significance of Lessing's role as a precursor of the romantic movement emerges even more prominently in his treatment of religious problems. He initiated the endeavor to discover within the immanent order of the world those values that had been derived by traditional Christianity from a transcendental view of the universe.

HISTORY AND THEOLOGY

Lessing inherited from his father strong scholarly and historical interests. By temperament antipathetic to all partisan historiography, he published a series of *Rettungen* (Vindications) in 1754, in which he defended historical figures to whom ecclesiastical historians, for dogmatic reasons, had not been quite fair. These essays are quite characteristic of Lessing's nature and cast of mind. Written with suppressed passion and permeated with a profound sense of engagement, they nevertheless remain uncommitted to any personal judgment either for or against the doctrinal beliefs of those whom he was vindi-

cating. His neutrality toward Christianity never took the form of quasi-Gibbonian irony. He always wrote as one wholly sympathetic to Christian ethical ideals, but coolly reserved toward dogmatic formulas that breed unreasoning prejudice and the negation of humane values.

The turning point of Lessing's life occurred in 1769, when he became librarian for the duke of Brunswick at Wolfenbüttel. In 1773 he began to publish essays on historical theology based on the Wolfenbüttel manuscripts. Earlier, during a three-year residence in Hamburg from 1766 to 1769 as a theater critic, Lessing had met the deist Hermann Samuel Reimarus (1694–1768), whose daughter had lent him the manuscript of an unpublished book by her father titled *Apologie oder Schutzschrift für die vernunftigen Verehrer Gottes* (Apology for rational worshipers of God). In 1774, and from 1777 to 1778, Lessing printed extracts from this work as fragments from the writings of an anonymous and unidentifiable deist whose manuscripts had presumably been found in the Wolfenbüttel library ("Wolfenbüttler Fragmente eines Ungenannten," in *Beitrage zur Geschichte und Literatur*).

The last and most important fragment precipitated a violent controversy with a Hamburg pastor, Johann Melchior Goeze, and effectively initiated the long nineteenth-century quest for the Jesus of history behind the Christ of faith. Reimarus was a believer in natural religion, but he was skeptical about revelation. His objections to traditional Christianity presuppose that biblical inerrancy is essential to faith. Lessing sometimes wrote as if he shared this assumption and sometimes as if he did not, so that it is not possible to arrive at a strictly coherent view on this point.

In his more cynical moments, Lessing treated liberal theology, such as that represented by J. S. Semler of Halle, with hostile contempt, on the ground that it was deceptively credible; he preferred to "defend" orthodoxy as being so patently absurd that by defense it would be sooner ended. Strictly as a scholar, Lessing was Semler's inferior; nevertheless, Lessing's genuinely scholarly instinct, combined with his inner detachment from the entrenched positions of the contemporary theological schools, as well as from those of the Enlightenment, enabled him to begin the critical study of the sources of the Synoptic Gospels (a fundamental question on which Reimarus had naively said nothing) with his pioneer essay, *Neue Hypothese über die Evangelisten als bloss menschliche Geschichtsschreiber betrachtet* (New Hypothesis concerning the Evangelists Regarded as Merely Human Historian). This was written from 1777 to 1778 and first printed in 1784 in Lessing's *Theologische Nachlass*.

Prevented by the duke of Brunswick from indulging in theological controversies, Lessing put his theology into a play, *Nathan der Weise (Nathan the Wise*, 1779) which was a plea for religious indifferentism on the ground that what is required of man is not an assent to the propositions of a creed, but sincerity, brotherly love, and tolerance. It is not easy to discover precisely what Lessing's positive beliefs were, so little did he commit himself, either in published writings or even in private correspondence, to any positive avowal of convictions. But he certainly accepted the commonplace thesis of the Enlightenment that the quintessence of Christianity, hidden beneath the accretions of theology, consists in universal brotherhood and a basic moral code. Like many rationalists of his age, he passed for a time into Freemasonry, though he emerged disillusioned with what was for him evidently a pale substitute for Christianity. In one sense, it could be said that Lessing spent his life hoping that Christianity was true and arguing that it was not. But his basic attitude toward religious belief was neither one of affirmation nor of denial; it took the form of an impassioned question.

Lessing was the first modern writer explicitly to emphasize that even if conclusions about historical events were more certain than they are, any religious affirmation based upon them involves a transition to another plane of discourse, that of faith. He was torn between the idea of revelation as the communication of timeless propositional truths, and the untidiness and irrationality of history. "Accidental truths of history can never become the proof of necessary truths of reason" (*Über den Beweis des Geistes und der Kraft*, 1777). Events and truths belong to altogether different categories, and there is no logical connection between one and another. Lessing's statement of this antithesis presupposes on the one hand the epistemology of Gottfried Wilhelm Leibniz, with its sharp distinction between necessary truths of reason (mathematically certain and known a priori) and contingent truths (known by sense perception), and on the other hand the thesis of Spinoza's *Tractatus Theologico-Politicus*, that the truth of a historical narrative, however certain, cannot give us the knowledge of God, which should be derived from general ideas that are in themselves certain and known. Lessing's own way out of the dilemma was to conceive the role of religious belief in the historical process as a relative state in the advance of humanity toward maturity, a thesis that he argued at length in the tract *Die Erziehung des Menschengeschlechts* (The Education of the Human Race; Berlin, 1780). Lessing thus became the father both of the "post-Christian" consciousness expressed in nineteenth-century posi-

tivism, and of the liberal religion of thinkers such as Samuel Taylor Coleridge and Frederick Denison Maurice.

There is more relativism than skepticism in Lessing's view. He did not think that absolute truth is revealed; but even if it were, and even if he were capable of apprehending it, he would not have wished to apprehend it. Adapting an aphorism of Clement of Alexandria, Lessing declared:

> The worth of a man does not consist in the truth he possesses, or thinks he possesses, but in the pains he has taken to attain that truth. For his powers are extended not through possession but through the search for truth. In this alone his ever-growing perfection consists. Possession makes him lazy, indolent, and proud. If God held all truth in his right hand and in his left the everlasting striving after truth, so that I should always and everlastingly be mistaken and said to me, Choose, with humility I would pick on the left hand and say, Father grant me that; absolute truth is for thee alone. (*Eine Duplik*, K. Lachmann and F. Muncker, eds., Vol. XIII, p. 23)

THE MOVE TO IMMANENTISM

Several fragmentary notes found among Lessing's papers, and published in 1784 by his brother Karl in *Theologischen Nachlass*, disclose the extent of Leibniz's influence. Lessing's interest was always most deeply aroused by Leibniz's references to theology and ethics. One of these pieces, written by Lessing about 1753, "Das Christentum der Vernunft" (The Christianity of Reason), foreshadowed a section of *Die Erziehung* in its attempt at making a speculative restatement of the doctrine of the Trinity, with the help of Leibnizian ideas on the hierarchy of being and the harmony of the monads. But there is a strong admixture of Spinoza in Lessing's conception of this harmony; he did not think of it as something preestablished by a Creator who is a superobject behind and beyond phenomena, but rather as being itself God, so that the perfect continuum of existents, in which there can be no gap, is indistinguishable from the perfection of the divine being. Similarly, in the brief notes titled "Ueber die Wirklichkeit der Dinge ausser Gott" (On the Reality of Things outside God; written in 1763, published in 1795 in Karl Lessing's *Lessings Leben*), Lessing denied the thesis of traditional theism that the created world exists independently of its Creator, in the sense of being distinct from him. Lessing urged that nothing can be outside the divine mind, and that there need be no hesitation before the conclusion that, since ideas of contingent things are themselves contingent, there is contingency even in God. These aphoristic fragments hardly amount to a coherent system. They show Lessing looking toward Spinoza, whom he had studied in his years at Breslau from 1760 to 1765, for a solution to some of the problems left unanswered by Leibniz.

Leibniz had formally asserted the freedom of the will, though it was doubted by Pierre Bayle and others whether Leibniz's libertarian assertions were in fact fully compatible with his philosophical principles. Lessing agreed with Spinoza that free will is a superfluity and an illusion. In 1776 Lessing published the *Philosophische Aufsätze* ("Philosophical Papers") of Karl Wilhelm Jerusalem, with the intention of making a protest against Johann Wolfgang von Goethe's *Werther*, with its description of Jerusalem's suicide. In a note to Jerusalem's third essay Lessing commented on his wisdom in recognizing that freedom is nothing but a cause of anxiety and fear, and that the recognition of necessity and destiny as beneficent is the only way to true happiness. "I thank my God," Lessing added, "that I am under necessity, that the best must be." The notion that the moralist has anything to fear from deterministic philosophies is just a mistake.

In 1785 at Breslau, Friedrich Heinrich Jacobi published his *Ueber die Lehre des Spinoza in Briefen an der Herrn Moses Mendelssohn* (Letters to Moses Mendelssohn on Spinoza's Doctrine), in which he disclosed that at Wolfenbüttel in July 1780, he had been told by Lessing, seven months before Lessing's death, that he could not believe the old transcendental metaphysic, and that he unreservedly accepted the pantheism of Spinoza—"There is no other philosophy." Jacobi was astonished to hear Lessing add that the determinism of Spinoza was no obstacle to him, and indeed that he had no desire for free will. Jacobi's revelations precipitated a furious controversy known as the *Pantheismusstreit*. The Enlightenment had derived from Bayle's *Dictionnaire historique et critique* such an unflattering picture of Spinoza that Jacobi's attribution of Spinozistic views to Lessing seemed like a shocking libel of a dead man. Moses Mendelssohn was moved to write an irate reply, in which he denied that Lessing was a pantheist and a determinist. Although not all of Jacobi's deductions were correct, the substantial accuracy of his account of what Lessing said is sufficiently vindicated by the fragments found among Lessing's papers. Lessing's final creed was a belief in an immanent destiny, with no room either for the concept of transcendence or for special revelation in any form; he believed in a determined pattern of cause and effect extending not

only throughout the physical order of nature, but also to morality and "the realm of ends."

Lessing's legacy to posterity was therefore to give an impetus to the notion of historical inevitability, especially in *Die Erziehung des Menschengeschlechts* at the end of which he even toyed with speculations about the transmigration of souls—obviously because this concept seemed to him more compatible with his historical determinism than the traditional eschatology connected with the Christian ideas of freedom and of personality.

The strong influence of Lessing is manifested in the history of religious thought in the nineteenth century. It can be traced particularly in the work of Søren Kierkegaard, whose *Concluding Unscientific Postscript* took its starting point from Lessing's statement about the intellectually impossible leap from the contingent truths of history to the necessary truths of divine revelation. The other, more liberal, side of Lessing was reflected in Coleridge, whose work was even suspected of being a plagiarism of Lessing's. In the field of literature and art, Lessing's attack on French classicism opened the way for the romantic ideal of free self-expression and naturalism, while his final theological position of Spinozistic immanentism clearly foreshadowed Friedrich Daniel Ernst Schleiermacher's *Speeches on Religion* (*Reden über die Religion*, 1799). His consciousness of living in an age of humanist maturity anticipated the Hegelian and Comtian estimates of religion as a useful, though now surpassed, stage in the education of humanity toward something higher and truer. Probably Lessing did as much as anyone to encourage among the educated European minds of his time an attitude of critical doubt that would lead to passionate engagement, rather than impersonal remoteness.

See also Aesthetic Experience; Clement of Alexandria; Deism; Enlightenment; Jacobi, Friedrich Heinrich; Kierkegaard, Søren Aabye; Leibniz, Gottfried Wilhelm; Mendelssohn, Moses; Nicolai, Christian Friedrich; Pantheismusstreit; Positivism; Rationalism; Reimarus, Hermann Samuel; Schleiermacher, Friedrich Daniel Ernst; Spinoza, Benedict (Baruch) de; Coleridge, Samuel Taylor; Voltaire, François-Marie Arouet de; Winckelmann, Johann Joachim; Wolff, Christian.

Bibliography

WORKS BY LESSING

Selected Prose Works. Translated by E. C. Beasley and Helen Zimmern. London, 1879.

Gesammelte Werke. 23 vols., edited by Karl Lachmann and Franz Muncker. 1886–1924.

Werke, edited by Julius Petersen and W. von Olshausen. Berlin, 1925–1935.

Hamburgischen Dramaturgie, edited by G. Waterhouse. Cambridge, U.K., 1926.

Lessings Gesammelte Werke. 10 vols., edited by Paul Rilla. Berlin: Aufbau-Verlag, 1954–1958.

Laocoön. Translated by E. A. McCormick. New York, 1962.

WORKS ON LESSING

Aner, K. *Die Theologie der Lessingzeit*. Halle: Niemeyer, 1929.

Bernstein, J. M., ed. *Classic and Romantic German Aesthetics*. New York: Cambridge University Press, 2003.

Cassirer, Ernst. *Die Philosophie der Aufklärung*. Tübingen: Mohr, 1932. Translated by F. C. A. Koelln and J. P. Pettegrove as *The Philosophy of the Enlightenment*. Princeton, NJ: Princeton University Press, 1951.

Chadwick, Henry. *Lessing's Theological Writings*. Palo Alto, CA, 1956.

Goetschel, Willi. *Spinoza's Modernity: Mendelssohn, Lessing, and Heine*. Madison: University of Wisconsin Press, 2004.

Gombrich, E. H. "Lessing." *Proceedings of the British Academy* 43 (1957): 133–156.

Hazard, Paul. *La pensée européenne au 18 ème siècle. De Montesquieu à Lessing*. Paris, 1946. Translated by J. Lewis May as *European Thought in the 18th Century*. London, 1954.

O'Flaherty, James C. *The Quarrel of Reason with Itself: Essays on Hamann, Michaelis, Lessing, Nietzsche*. Columbia: Camden House, 1988.

Ritchie, Gisela F. "Contributors to the Genesis of Europe: Gotthold Ephraim Lessing and His Followers." *History of European Ideas* 19(1–3) (1994): 425–430.

Wessel, Leonard P. *G. Lessing's Theology, Reinterpretation: A Study in the Problematic Nature of the Enlightenment*. The Hague: Mouton, 1977.

Henry Chadwick (1967)
Bibliography updated by Tamra Frei (2005)

LEUCIPPUS AND DEMOCRITUS

Leucippus and Democritus were the earliest Greek atomists. The originator of the atomic theory, Leucippus (fifth century BCE), must be considered a speculative thinker of the first order, but to Democritus (c. 460–c. 370 BCE) must go the credit for working out the detailed application of the theory and supporting it with a subtle epistemology. Moreover, the range of Democritus's researches surpassed that of any earlier philosopher, and he appears to have been an original and, for his day, advanced ethical thinker.

We have very little biographical data for Leucippus. Epicurus is even reported to have said that there was no

philosopher Leucippus, but the evidence of Aristotle decisively refutes this opinion (if, indeed, Epicurus did not merely intend to deny Leucippus's philosophical importance). Leucippus was probably born at Miletus; reports associating him with Elea or Abdera should be taken as reflecting views concerning his philosophical affiliations rather than as reliable evidence for his birthplace. He was presumably older than Democritus. His book *On Mind* may have been directed partly against Anaxagoras, and according to Theophrastus, Diogenes of Apollonia derived some of his theories from Leucippus. All this suggests that Leucippus was a slightly younger contemporary of Anaxagoras and that his main philosophical activity fell some time within the broad limits of 450–420 BCE.

Democritus was born at Abdera. He described himself in the *Little World-System* as a young man in the old age of Anaxagoras; Diogenes Laërtius says that he was forty years younger than Anaxagoras. On this evidence the date given for his birth by Apollodorus (in the 80th Olympiad, 460–456 BCE) is generally preferred to that suggested by Thrasylus (the third year of the 77th Olympiad, 470–469 BCE). He is variously reported to have lived between 90 and 109 years. To judge from the number of his writings, his literary activity extended over a considerable period, but we have no means of assigning different works to different times in his life. His statement that he wrote the *Little World-System* 730 years after the fall of Troy (Diogenes Laërtius, *Lives* IX, 41) is of little value since we cannot tell which of several possible chronologies for the Trojan War Democritus accepted.

Many stories, most of them apocryphal, relating to Democritus's life and character circulated in antiquity. There are the accounts of his saving the Abderites from a plague, of his dying by voluntarily abstaining from food, and of his reputation as the "Laughing Philosopher." The tradition that he traveled extensively is, however, more plausible and better grounded. The authenticity of the fragment (299) in which he claimed to be the most widely traveled of his contemporaries is disputed, and the genuineness of the five books dealing with foreign travel mentioned by Diogenes Laërtius (for example, *A Voyage round the Ocean*) has also been doubted. But evidence concerning his travels goes back to Theophrastus (see Aelian, *Varia Historia* IV, 20), and the reports that he visited such places as Egypt, Chaldea, and the Red Sea (see Diogenes Laërtius, *Lives* IX, 35) may well have a sound basis in fact.

All that has been preserved of the original writings of Leucippus and Democritus is a poor selection of isolated quotations, most of which derive from the ethical works of Democritus. For the atomic theory itself we rely on reports in Aristotle, Theophrastus, and later doxographers, who were often unsympathetic to the views of the atomists. In most of the principal texts referring to Leucippus, his doctrines are not clearly distinguished from those of Democritus, and the precise contribution of each philosopher is in question. Aristotle, however, undoubtedly treated Leucippus as the founder of atomism (*De Generatione et Corruptione* 325a23ff.), and we may reasonably attribute both the principles of the physical theory and a fairly complex cosmogony to him. Democritus evidently elaborated the atomic theory and was responsible for the detailed account of sensible qualities, besides going far beyond Leucippus both in the range of his scientific inquiries and in his interest in moral philosophy.

WRITINGS

Only two works are ascribed to Leucippus, *On Mind,* from which our sole surviving quotation comes, and the *Great World-System,* which may be attributed to Leucippus on the authority of Theophrastus (Diogenes Laërtius, *Lives* IX, 46), although Thrasylus later assigned it to Democritus.

Democritus, on the other hand, wrote some sixty-odd works, the titles of which provide valuable evidence of the scope of his interests. The main works were cataloged by Thrasylus into thirteen tetralogies. Two tetralogies are devoted to ethics and four to physics (including *Little World-System, On the Planets, On Nature, On the Nature of Man, On the Senses,* and *On Colors*). These were followed by nine works not arranged in tetralogies—for example, *Causes of Celestial Phenomena, Causes concerning Seeds, Plants and Fruits,* and three books of *Causes concerning Animals.* Three tetralogies are classified as mathematics, two deal with music and literature, and two consist of technical works, including treatises on medicine, agriculture, painting, and warfare. Nine other miscellaneous works, mostly concerning travel, are also mentioned by Diogenes Laërtius but are less certainly authentic as they were not included in Thrasylus's catalog.

Democritus's style is described by Cicero as elegant (*De Oratore* I, 11, 49) and lucid (*De Divinatione* II, 64, 133), and an anecdote recorded by Diogenes Laërtius (*Lives* IX, 40) implies that his works already had wide circulation by the time of Plato.

THE ATOMIC THEORY

The basic postulate of Greek atomism in its original form was that atoms and the void alone are real. The differences between physical objects, including both qualitative differences and what we think of as differences in substance, were all explained in terms of modifications in the shape, arrangement, and position of the atoms. Aristotle illustrates these three modes of difference with the examples A and N, AN and NA, and ⌶ and H.

This theory was already interpreted by Aristotle as an answer to the Eleatic denial of change and movement. Other post-Parmenidean philosophers had countered this denial in different ways, but both Empedocles and Anaxagoras had assumed a variety of elemental substances, on the one hand, the four "roots," on the other, an original mixture containing every kind of natural substance. In postulating a single elemental substance, Leucippus remained closer to Parmenides' own conception. In common with Parmenides' One Being the individual atoms are ungenerated, indestructible, unalterable, homogeneous, solid, and indivisible. Leucippus may be said to have postulated an infinite plurality of Eleatic ones, and he may even have been directly influenced by Melissus's argument (Fr. 8) that "if there were a plurality, they would have to be as the One is." Leucippus also agreed with the Eleatics that without void movement is impossible. Yet whereas the Eleatics denied the existence of the void, or "what is not," Leucippus maintained that not only "what is" (the atoms), but also "what is not" (the void), must be considered real. Leucippus thereby reinstated both plurality and change; the void is that which separates the atoms and that through which they move.

The atoms are infinite in number, dispersed through an infinite void. Their shapes are infinitely various, there being no reason that any atom should be of one shape rather than another. Democritus, at least, also allowed differences in the sizes of the atoms, but whether he thought any atom large enough to be visible seems doubtful. Late sources that report that atoms are unlimited in size as well as number (Diogenes Laërtius, *Lives* IX, 44) or which suggest the possibility of an atom the size of the world (Aëtius, *Placita* I, 12, 6) are difficult to credit in view of the testimony of Aristotle, who apparently believed that for both Leucippus and Democritus the atoms are all so small that they are invisible.

The atoms are in continuous motion. Aristotle, among others, objected that the atomists did not explain the origin of movement or say what kind of movement is natural to the atoms. However, they evidently assumed that the motion of the atoms is eternal, just as the atoms themselves are, and they perhaps drew no clear distinction between original and derived motion. Although Epicurus was later to suggest that atoms naturally fall vertically, the earlier atomists probably did not consider movement in any particular direction prior to movement in any other. Weight for them, it seems, was not a primary property of the atoms nor a cause of their interactions, although in a developed cosmos the atoms have "weight" corresponding to their size (and the weight of compound bodies varies according to the proportion of atoms and void they contain).

The movements of the atoms give rise to constant collisions whose effects are twofold. Sometimes, the atoms rebound from one another; alternatively, when the colliding atoms are hooked or barbed or their shapes otherwise correspond, they cohere and thus form compound bodies. Change of all sorts is accordingly interpreted in terms of the combining and separating of atoms, which themselves remain unaltered in substance. The compound bodies thus formed possess various sensible qualities—color, taste, temperature, and so on—and Democritus undertook a detailed exposition relating these qualities to specific atomic configurations.

COSMOGONY. Evidence concerning Leucippus's cosmogony comes mainly from Diogenes Laërtius (*Lives* IX, 31ff.). The process begins when a large group of atoms becomes isolated in a great void. There they conglomerate and form a whirl or vortex in which atoms of similar shape and size come together. In this vortex the finer atoms are squeezed out into the outer void, but the remainder tend toward the center, where they form a spherical mass. More atoms are drawn into this mass on contact with the whirl, and some of these are ignited by the speed of the revolution, thus forming the heavenly bodies. Earth is formed by atoms that cohere in the center of the mass. The cosmogonical process is not unique. The atomists argued that since atoms and the void are infinite, there are innumerable worlds. These worlds are not all alike, however; Democritus held that some worlds have no sun or moon and that some lack moisture and all forms of life (Hippolytus, *Refutatio* I, 13, 2f.).

Several features of this account are obscure, and two apparently conflicting criticisms were leveled against it in antiquity—first, that although the atomists asserted that the cosmogonical process came about by necessity, they did not explain what this necessity was (Diogenes Laërtius, *Lives* IX, 33); second, that they maintained that it occurred spontaneously (Aristotle clearly has the atom-

ists in mind when he considered this view, *Physics* 196a24ff.).

But Aristotle's judgment should be taken as referring primarily to the atomists' exclusion of final causes; in Aristotelian terms the atomists held that the world arose spontaneously because they denied that it was intelligently planned. Leucippus explicitly stated that "nothing happens at random, but everything for a reason and by necessity" (Fr. 2), and throughout their cosmology the atomists not only excluded purpose or design but also assumed that every event is the product of a definite, theoretically determinable cause. Thus, they doubtless conceived the vortex to arise from certain mechanical interactions between the colliding atoms, although it is unlikely that they attempted to say precisely how this came about. Democritus illustrated his doctrine that like things tend to come together with examples drawn from both the inanimate and the animate sphere (Fr. 164). And like many of the pre-Socratics, the atomists constructed their cosmogony in part on an embryological model. The outer envelope of the world was likened to a membrane, and in both Leucippus's cosmogony and Democritus's embryology the process of differentiation apparently takes place from the outside (see Aristotle, *De Generatione Animalium* 740a13ff.).

ASTRONOMY AND BIOLOGY. Leucippus's astronomical theories are surprisingly retrograde. He accepted the old Ionian picture of a flat earth, tilted toward the south, and he believed that the sun is the most distant of the heavenly bodies. Democritus's theories were generally less crude, and he attempted rational explanations of a wide variety of obscure phenomena. He accepted Leucippus's account of Earth with only minor modifications (Aëtius, *Placita* III, 10, 5) but corrected his notion of the relative positions of the heavenly bodies, observing, for example, that the planets are not equidistant from Earth and placing Venus between the sun and moon. Among other topics on which some of Democritus's theories are recorded are the behavior of the magnet, the nourishment of the embryo, and the relative longevity of different types of plants. Of his biological doctrines the notion that the seed is drawn from the whole of the body (the pangenesis theory) was particularly influential (Aëtius, *Placita* V, 3, 6).

SOUL, KNOWLEDGE, AND SENSATION. Our evidence concerning the atomists' psychological and epistemological doctrines derives very largely from Democritus, although his theory of the soul was probably developed from ideas outlined by Leucippus. This theory was a materialist one in line with the principles of atomism. Democritus conceived of the soul as consisting of spherical atoms, this being the shape best adapted to penetrate and move things. Fire, too, is composed of spherical atoms, and he evidently subscribed to the common Greek belief in the connection between life and heat, now interpreted in terms of the similarity in the shapes of soul atoms and fire atoms. The soul atoms tend to be extruded from the body by the pressure of the surrounding air, but this process is counteracted by other soul atoms that enter the body with the air we breathe; life depends on this continuous replenishment.

Our main source for Democritus's theory of knowledge is Sextus Empiricus. Several of the fragments that he quotes appear to express an extreme skepticism—for instance, "We know nothing truly about anything" (Fr. 7). However, Fragment 11 shows that Democritus was no outright skeptic. There he distinguished between two modes of cognition; the senses provide what is called a "bastard" knowledge but contrasted with this is a "legitimate" knowledge, which operates on objects too fine for the senses to perceive. Clearly, "legitimate" knowledge relates to atoms and the void, which alone are real; the objects of sensation, on the other hand, exist "by convention" (Fr. 9). The doctrine enunciated in the fragments is that sense perception is not trustworthy, and Aristotle's repeated statement that the atomists found truth in appearance (*De Generatione et Corruptione* 315b9ff.) should be understood as an interpretative comment based on Aristotle's own conception of the distinction between sensibles and intelligibles. Yet although we must rely on reasoning to attain knowledge, Democritus acknowledged that the mind derives its data from the senses (Fr. 125). Not a pure intellectualist like Parmenides, a crude sensationalist like Protagoras, nor a complete skeptic as Gorgias made himself out to be, Democritus advocated critical reflection on the evidence of the senses as our best means of approaching the truth; yet since thought itself, like sensation, involves physical interactions between atoms, it, too, is subject to distortion, and even "legitimate" knowledge is at best, it seems, only opinion (Fr. 7).

Democritus's detailed accounts of the five senses were reported and criticized at length by Theophrastus (*De Sensibus* 49–82). According to Alexander (*In Librum de Sensu* 24, 14ff.), Leucippus already held that physical objects constantly emit images that effect vision on entering the eye. Democritus modified and complicated this doctrine by suggesting that images from both the object and the eye itself meet and imprint the air in front of the

eye. Each of the other senses, too, is produced by contact between the organ and images deriving from the object, and thought was analogously explained as the contact between soul atoms and images coming from outside the body. But not content merely to assert in general terms that secondary qualities are due to differences in the shapes and sizes of the atoms, Democritus also proposed a detailed account relating specific tastes, colors, smells, and so on to specific shapes. Thus, an acid taste is composed of angular, small, thin atoms and a sweet taste of round, moderate-sized ones. Democritus's primary colors—black, white, red, and greenish yellow—were similarly associated with certain shapes and arrangements of atoms, and other colors were derived from combinations of these four. For all its crudities Democritus's theory may claim to be the first fully elaborated account of the physical basis of sensation.

MATHEMATICS. Democritus's interest in mathematics is apparent from the titles of fives works dealing with mathematical subjects, and we are told, for example, that he discovered the relation between the volumes of a pyramid and a prism with the same base and equal height. We have, however, little evidence on the part of his mathematical work that related directly to the atomic theory. The atoms are definitely conceived of as physically indivisible (on the grounds that they are solid and contain no void), but it is not clear whether they are absolute minima in the sense of being mathematically indivisible. Epicurus later distinguished between atomic bodies (which are physically indivisible but logically divisible) and the "minima in the atom." But Aristotle appears to have assumed that Leucippus and Democritus themselves drew no distinction between the limits of physical and mathematical divisibility (*De Generatione et Corruptione* 315b28ff.), and he considered that their atomic theory necessarily conflicted with the mathematical sciences (*De Caelo* 303a20ff.). Unless Aristotle has completely misrepresented the atomists, it would appear that Democritus was unaware of any inconsistency in holding both (1) that the atoms have different shapes and sizes and (2) that they are mathematically as well as physically indivisible. But it must be repeated that the evidence on which to convict or absolve Democritus of this gross confusion is scanty.

ETHICS. Although serious doubts have been raised concerning the transmission of the ethical fragments of Democritus, most scholars now consider that the majority of those accepted by Hermann Diels and Walther Kranz may be used as a basis from which to reconstruct his ethics. There remain, however, wide disagreements on the nature and value of his moral teaching. Alongside the fragments that convey traditional sentiments (for example, on the dangers of fame and wealth if not accompanied by intelligence) we find others that expound notions far in advance of the popular morality of the day, as, for instance, the doctrine that it is one's own consciousness of right and wrong, not fear of the law or public opinion, that should prevent one from doing anything shameful (Frs. 181, 264). And sayings such as Fragment 45 ("The wrongdoer is more unfortunate than he who is wronged") express views more commonly associated with Socrates than with Democritus.

The ethical ideal is termed "well-being" or "cheerfulness," which is to be gained through uprightness and a harmonious life. Although Democritus clearly implied that life without pleasure is not worth living and even said that pleasure is the mark of what is expedient (Fr. 188), it is the higher pleasures of the soul that we should cultivate, not those of the body. Sensual pleasures are condemned as short-lived. He repeatedly stressed that we should moderate our desires and ambitions, become self-sufficient, and be content, in the main, with simple pleasures. Yet Democritus was no quietist. Rather, he recognized that worthwhile objects are to be achieved only through effort (Frs. 157, 182).

One of the salient features of Democritus's ethics is his rejection of supernatural sanctions of behavior. In part, he seems to have rationalized belief in the gods as a mistaken inference from terrifying natural phenomena (Sextus, *Adversus Mathematicos* IX, 24), and yet he did not dismiss notions of the gods entirely, for he appears to have related certain such ideas to images, some beneficent, some harmful, that visit humans (Fr. 166). Religious sanctions are, however, rigorously excluded from his ethics. He refuted those who concocted fictions concerning the afterlife (Fr. 297), and he spoke with apparent irony of those who prayed to Zeus as "king of all" (Fr. 30). Equally, he castigated those who invented chance as an excuse for their own thoughtlessness or who failed to recognize that their misfortunes stemmed from their own incontinence (Frs. 119, 234). Throughout his ethics he may be said to have set high standards of personal integrity and social responsibility.

The question of the relation between Democritus's ethics and his physics has been much debated. In some respects, such as in the idea that excesses "cause great movements in the soul"—that is, presumably, in the soul atoms (Fr. 191)—his ethics reflect a psychology that is based on his physical theories. Whether we should expect

other aspects of the atomic doctrine to be in evidence in the ethical fragments seems very doubtful. Democritus clearly did not feel (nor need he have felt) that the notion of necessity in his physics (the belief that every event has a definite cause to be sought in the interactions of the atoms) conflicted with his doctrine of moral responsibility in the sphere of human behavior. His denial of supernatural sanctions in his ethics parallels his rejection of teleology in his cosmology. And his ethics have in common with his epistemological theory that he argued against an unreflecting acceptance of the evidence of the senses concerning what is pleasant just as much as concerning the nature of reality as a whole.

SOCIOLOGY AND POLITICS. The only indication we have of Democritus's political leanings is the idealistic but otherwise rather inconclusive Fragment 251: "Poverty under democracy is as much to be preferred to so-called happiness under tyrants as freedom to slavery." It has, however, been conjectured that the account of the origin of civilization preserved in Diodorus (*Bibliotheca Historica* I, 8) owes much to Democritus. According to this, primitive peoples originally gathered in groups for the sake of mutual protection from wild animals, and subsequently language and the arts were also invented under the spur of human needs. It is very uncertain how far this reproduces Democritus's ideas, but there is some evidence in the fragments that he maintained a naturalistic theory of civilization and progress and excluded teleological explanations here, as he did elsewhere in his philosophy. Fragment 144 may be taken to suggest that he believed that the earliest arts (although not some of the later ones) were products of necessity, and in Fragment 154 he argued that humans learned many of their skills by copying the behavior of animals.

The theory founded by Leucippus and developed by Democritus was the most coherent and economical physical system of its day, and the history of its influence can be traced from the fourth century BCE to modern times. Although Plato mentioned neither Leucippus nor Democritus, the *Timaeus* is markedly indebted to their thought. Even Aristotle, who rejected atomism outright, conceded that of all his predecessors Democritus was the most notable physicist. Later, the Epicureans championed atomism against the continuum theory of the Stoics. Leucippus's theory, in origin primarily an answer to the Eleatic arguments against change, was the first clear formulation of the doctrine that matter exists in the form of discrete particles, and as such it may legitimately be considered the prototype of modern theories of the discontinuous structure of matter, even though the nature of such theories, the problems they are intended to resolve, and the methods used to establish them all differ fundamentally from those of ancient atomism.

See also Alexander of Aphrodisias; Anaxagoras of Clazomenae; Aristotle; Atomism; Cosmology; Diodorus Cronus; Diogenes Laertius; Diogenes of Apollonia; Empedocles; Epicureanism and the Epicurean School; Epicurus; Gorgias of Leontini; Parmenides of Elea; Plato; Pre-Socratic Philosophy; Protagoras of Abdera; Quantum Mechanics; Sextus Empiricus; Stoicism; Theophrastus.

Bibliography

The extant fragments of Leucippus and Democritus and the principal reports and commentaries in ancient authors are collected in Hermann Diels's *Die Fragmente der Vorsokratiker,* Vol. II, 6th ed., with additions by Walther Kranz, ed. (Berlin, 1952). There is an English translation of the fragments in Kathleen Freeman, *Ancilla to the Pre-Socratic Philosophers* (Cambridge, MA: Harvard University Press, 1948).

More recent collections of texts are S. Luria, *Democritea* (Leningrad, 1970; original texts of fragments and testimonia with Russian translation and commentary) and C. C. W. Taylor, *The Atomists: Leucippus and Democritus. Fragments: A Text and Translation with a Commentary* (Toronto: University of Toronto Press, 1999; fragments in Greek with facing English translation, testimonia in translation, commentary).

A classic monograph on Leucippus and Democritus is Cyril Bailey, *The Greek Atomists and Epicurus* (Oxford, 1928). Lucid, brief expositions of their thought are found in G. S. Kirk, J. E. Raven and M. Schofield, *The Presocratic Philosophers,* 2nd ed. (Cambridge, U.K., 1983) and in R. D.McKirahan Jr., *Philosophy before Socrates* (Indianapolis and Cambridge, 1994). W. K. C. Guthrie, *A History of Greek Philosophy,* vol. 2 (Cambridge, U.K.: Cambridge University Press, 1965), contains a full discussion and extensive bibliography. Other significant general studies are E. Zeller, *Die Philosophie der Griechen,* 6th ed., revised and enlarged by W. Nestle (Leipzig, 1920), Part I, Sec. 2, pp. 1038–1194; Wilhelm Schmidt in Wilhelm Schmidt and Otto Stählin, eds., *Geschichte der griechischen Literatur* (Munich, 1948), Part I, Sec. 5, pp. 224–349; V. E. Alfieri, *Gli atomisti* (Bari, 1936) and *Atomos Idea,* 2nd ed. (Florence, 1979); J. Barnes, *The Presocratic Philosophers,* 2nd ed. (London and Boston: Routledge and K. Paul, 1982), Chapters 17, 19 (*b*), 20, 21 (*c*), 23 (*d*), 24 (*e*); P.-M. Morel, *Démocrite et la recherche des causes* (Paris, 1996); and J. Salem, *Démocrite: Grains de poussière dans un rayon de soleil* (Paris: CNRS-Éditions, 1996).

A full discussion of the chronological and biographical data on Democritus is contained in D. O'Brien, "Démocrite d'Abdère" in R. Goulet, ed., *Dictionnaire des Philosophes Antiques,* vol. 2, pp. 649–715 (Paris: CNRS-Éditions, 1994).

PHYSICS AND COSMOLOGY

Barnes, J. "Reason and Necessity in Leucippus." In *Proceedings of the First International Conference on Democritus,* edited by L. Benakis. Vol. 1, pp. 141–158. Xanthi, 1984.

Bury, R. G. "The Origin of Atomism." *Classical Review* 30 (1916): 1–4.

Dyroff, Adolf. *Demokritstudien.* Leipzig: Dieterich, 1899.

Fritz, Kurt von. *Philosophie und sprachlicher Ausdruck bei Demokrit, Plato und Aristoteles.* New York, n.d. [1939].

Furley, D. J. *Cosmic Problems.* Cambridge, U.K., and New York: Cambridge University Press, 1989. Chapters 7–9.

Furley, D. J. *The Greek Cosmologists.* Cambridge, U.K., and New York: Cambridge University Press, 1987. Chapters 9–13.

Furley, D. J. "Weight and Motion in Democritus' Theory." *Oxford Studies in Ancient Philosophy* 1 (1983): 193–209.

Hammer-Jensen, Ingeborg. "Demokrit und Platon." *Archiv für Geschichte der Philosophie,* n.s. 16 (1910): 92–105, 211–229.

Kersehensteiner, Jula. "Zu Leukippos A 1." *Hermes* 87 (1959): 441–448.

Löbl, R. *Demokrits Atomphysik.* Darmstadt: Wissenschaftliche Buchgesellschaft, 1987.

Löwenheim, Louis. *Die Wissenschaft Demokrits.* Berlin: Simion, 1914.

Makin, S. *Indifference Argument.* Oxford and Cambridge, MA: Blackwell, 1993.

McDiarmid, J. B. "Theophrastus *De Sensibus* 61–62: Democritus' Theory of Weight." *Classical Philology* 55 (1960): 28–30.

Mugler, Charles. "Sur Quelques Particularités de l'atomisme ancien." *Revue de philologie,* 3rd series 27 (1953): 141–174.

O'Brien, D. *Theories of Weight in the Ancient World.* Vol. 1: *Democritus: Weight and Size.* Paris: les Belles Letters; and Leiden: Brill, 1981.

Sambursky, Samuel. "Atomism versus Continuum Theory in Ancient Greece." *Scientia* 96 (1961): 376–381.

EPISTEMOLOGY

McDiarmid, J. B. "Theophrastus *De Sensibus* 66: Democritus' Explanation of Salinity." *American Journal of Philology* 80 (1959): 56–66.

McKim, R. "Democritus against Scepticism: All Sense-Impressions Are True." In Benakis op. cit., vol. 1, pp. 281–290.

Weiss, Helene. "Democritus' Theory of Cognition." *Classical Quarterly* 32 (1938): 47–56.

MATHEMATICS

Bodnár, I. M. "Atomic Independence and Indivisibility." *Oxford Studies in Ancient Philosophy*16 (1998): 35–61.

Furley, D. J. *Two Studies in the Greek Atomists.* Princeton, NJ, 1977. Study 1, "Indivisible Magnitudes."

Makin, S. "The Indivisibility of the Atom." *Archiv für Geschichte der Philosophie* 71 (1989): 125–149.

Luria, Salomo. "Die Infinitesimaltheorie der antiken Atomisten." *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik,* Part 2 (Studies) 2, Sec. 2 (Berlin, 1932–1933): 106–185.

Mau, Jürgen. *Zum Problem des Infinitesimalen bei den antiken Atomisten.* Berlin: Akademie-Verlag, 1954.

Philippson, Robert. "Democritea." *Hermes* 64 (1929): 175–183.

Vlastos, Gregory. "Minimal Parts in Epicurean Atomism." *Isis* 56 (1965): 121–147.

Zubov, V. P. "K Voprosu o Matematicheskom Atomisme Democrita." *Vestnik Drevnei Istorii* 4 (1951): 204–208.

ETHICS

Kahn, C. H. "Democritus and the Origins of Moral Psychology." *American Journal of Philology* 106 (1985): 1–31.

Langerbeck, Hermann. ΔΟΞΙΣ ΕΠΙΡΥΣΜΙΗ. *Neue Philologische Untersuchungen,* Vol. X. Berlin, 1935.

Laue, Heinrich. *De Democriti Fragmentis Ethicis.* Unpublished dissertation, University of Göttingen, 1921.

Laue, Heinrich. "Die Ethik des Demokritos." *Sokrates,* n.s. 11 (1923–1924): 23–28, 49–62.

Luria, Salomo. *Zur Frage der materialistischen Begründung der Ethik bei Demokrit.* Berlin: Akademie-Verlag, 1964.

McGibbon, Donal. "Pleasure as the 'Criterion' in Democritus." *Phronesis* 5 (1960): 75–77.

McGibbon, Donal. "The Religious Thought of Democritus." *Hermes* 93 (1965): 385–397.

Natorp, Paul. *Die Ethika des Demokritos.* Marburg: Elwert, 1893.

Philippson, Robert. "Demokrits Sittensprüche." *Hermes* 59 (1924): 369–419.

Stewart, Zeph. "Democritus and the Cynics." *Harvard Studies in Classical Philology* 63 (1958): 179–191.

Taylor, C. C. W. "Pleasure, Knowledge and Sensation in Democritus." *Phronesis* 12 (1967): 6–27.

Vlastos, Gregory. "Ethics and Physics in Democritus." *Philosophical Review* 54 (1945): 578–592; 55 (1946): 53–64.

Warren, J. *Epicurus and Democritean Ethics.* Cambridge, U.K., and New York: Cambridge University Press, 2002.

SOCIOLOGY AND POLITICS

Aalders, G. J. D. "The Political Faith of Democritus." *Mnemosyne,* 4th series, 3 (1950): 302–313.

Procopé, J. F. "Democritus on Politics and the Care of the Soul." *Classical Quarterly,* n.s. 39 (1989): 307–331 and 40 (1990): 21–45.

Vlastos, Gregory. "On the Pre-history in Diodorus." *American Journal of Philology* 67 (1946): 51–59.

G. E. R. Lloyd (1967)
Bibliography updated by C. C. W. Taylor (2005)

LEVINAS, EMMANUEL
(1906–1995)

Emmanuel Levinas was born in Kaunas, Lithuania, of Jewish parents. His education familiarized him with the Hebrew Bible and the Russian novelists. After having studied at the gymnasiums in Kaunas and Charkow, Ukraine, he traveled to Strasbourg, where he studied philosophy from 1924 to 1929. He spent the academic year of 1928–1929 in Freiburg, where he attended the last seminars given by Edmund Husserl and the lectures and seminars of Martin Heidegger. His dissertation, *La théorie de*

l'intuition dans la phénoménologie de Husserl, was published in 1930. In 1930 Levinas settled in Paris, where he worked for the Alliance Israélite Universelle and its schools located throughout the Mediterranean. In 1947 he became the director of the École Normale Israélite Orientale, the training facility for teachers of those schools. In 1961 he was appointed professor of philosophy at the University of Poitiers and in 1967 at the University of Nanterre. In 1973 he moved to the Sorbonne, where he became an honorary professor in 1976. Levinas died on December 25, 1995, a few days before his 90th birthday.

WORKS

Until World War II most of Levinas's writing focused on introducing the phenomenology of Husserl and Heidegger into France. His early commentaries on their work were collected in *En découvrant l'existence avec Husserl et Heidegger* (1949). His first personal essay was the article "De l'évasion" (1935), whose central question was whether it is possible to evade the totalizing tendency of being. The search for an answer coincided with the beginning of his criticism of Heidegger's ontology. Levinas's first personal book, with the anti-Heideggerian title *De l'existence à l'existant* (*From Existence to Existents* or *From Being to Beings*), was published in 1947. In the same year he gave a lecture series under the title *Le temps et l'autre* (*Time and the Other*), in which some central thoughts of his later work are anticipated. A part of *De l'existence à l'existant* to which Levinas later refers with approval is its phenomenology of *il y a* ("there is"), that is, being in its most general and indeterminate or empty sense, preceding all determination, order, and structure. Levinas describes it as a formless and obscure night and a silent murmur, an anonymous and chaotic atmosphere or field of forces from which no being can escape. It threatens the existing entities by engulfing and suffocating them. As such, being is horrible, not because it would kill—death is not an evasion from it—but because of its depersonalizing character. All beings are caught in the anonymity of this primordial materiality—much different from the giving essence of *es gibt* as described by Heidegger.

The work that made Levinas famous is *Totalité et infini. Essai sur l'extériorité* (1961). As an attack on the entirety of Western philosophy, including Heidegger's ontology, this work tries to show why philosophy has not been faithful to the most important facts of human existence and how its basic perspective should be replaced by another one. The "totality" of the title stands for the absolutization of a panoramic perspective from which reality is understood as an all-encompassing universe. All kinds of relation, separation, exteriority, and alterity are then reduced to internal moments of one totality. Borrowing from Plato's *Sophist,* Levinas affirms that Western philosophy reduces the other (*to heteron*) to "the Same" (*tauton*). The resulting tautology is an egology because the totalization is operated by the consciousness of an ego that does not recognize any irreducible heteronomy.

The relative truth of the ego's autonomy is shown in a phenomenology of the way in which human beings inhabit the world. Levinas characterizes this "economy" (from *oikos* = house, and *nomos* = law) as vitality and enjoyment of the elements. Implicitly polemicizing against Heidegger's description of *Dasein*'s being-in-the-world, he focuses on the dimension of human eating, drinking, walking, swimming, dwelling, and laboring, a dimension more primordial than the handling of tools and much closer to the natural elements than scientific or technological objectification.

The infinite (*l'infini*), which Levinas contrasts with the totality, is another name for "the Other" insofar as this does not fit into the totality. In order to determine the relation between consciousness, the totality, and the infinite, Levinas refers to René Descartes's *Meditations on the First Philosophy,* in which Descartes insists on the fact that the idea of the infinite is original and cannot be deduced from any other idea. It surpasses the capacity of consciousness, which in it "thinks more than it can think" (see Levinas's *Collected Philosophical Papers,* p. 56). The relation between the ego and the infinite is one of transcendence: The infinite remains exterior to consciousness, although this is essentially related to its "height."

The concrete sense of the formal structure thus indicated is shown through a phenomenology of the human other, whose "epiphany" reveals an absolute command: As soon as I am confronted, I discover myself to be under an absolute obligation. The fact of the other's existence immediately reveals to me the basic ought of all ethics. On this level is and ought are inseparable. Instead of the other (*l'autre* or *autrui*), Levinas often uses the expressions "the face" (*le visage*) or "the speech" (*la parole,* also *le langage*) because the other's looking at me and speaking to me are the two most striking expressions of the other's infinity or "height." As the relation between an economically established ego and the infinite other, the intersubjective relation is asymmetrical: The other appears primarily not as equal to me but rather as "higher" and commanding me. I am responsible for the other's life, a responsibility that puts infinite demands on

me, but I cannot order another to give his or her life for me.

In his second major work, *Autrement qu'être ou audelà de l'essence* (1974), Levinas continues his analyses of the relationship between the ego and the other but now emphasizes the basic structure of the ego, or rather of the "me" in the accusative, as put into question, accused, and unseated by the other. The relationship is described as nonchosen responsibility, substitution, obsession, being hostage, persecution. Subjectivity (the "me" of *me voici*) is determined as a nonchosen being-for-the-other and, thus, as basically nonidentical with itself, a passivity more or otherwise passive than the passivity that is opposed to activity. Subjectivity is primarily sensibility, being touched and affected by the other, vulnerability.

In the course of his analyses Levinas discovered that the other, me, and the transcendence that relates and separates them do not fit into the framework of phenomenology: Neither the other nor I (me) is phenomenon; transcendence does not have the structure of intentionality. Through phenomenology Levinas thus arrived at another level of thinking. He did not join Heidegger's call for a new ontology, however.

In *Autrement qu'être* Levinas gives a new description of the way being "is": *Esse* is *interesse*; being is an active and transitive "interestingness" (*intéressement*), which permeates all beings and weaves them together in a network of mutual interest. If ontology is the study of (this) being, it is not able to express the other, transcendence, and subjectivity. Transcendence surpasses being. Appealing to Plato, who characterized the good as *epekeina tès ousias,* Levinas points at transcendence, infinity, and otherness as "otherwise" and "beyond" the realm of being (or essence).

The other, subjectivity, and transcendence—but then also morality, affectivity, death, suffering, freedom, love, history, and many other (quasi-)phenomena—resist, not only phenomenology and ontology, but all kinds of objectification and thematization. As soon as they are treated in a reflective discourse, they are converted into a said (*dit*). The saying (*dire*), in which the "otherwise than being" (that which is not a phenomenon, a being, or a theme) addresses itself to an addressee, is lost in the text of the said. However, thematization and objectification are inevitable, especially in philosophy and science, but also in the practical dimensions of law, economy, and politics. The organization of justice cannot do without generalization and grouping of individuals into totalities. The transition from the asymmetrical relation between the other and me to the generalities of justice is founded in the fact that the other human who, here and now, obligates me infinitely somehow represents all other humans.

How does the intersubjective and asymmetric transcendence differ from the relationship to God? "Otherness," "infinity," and "beyond" do not apply to God in the same way as to the human other. God is neither an object nor a you; no human being can meet with God directly, but God has left a trace. The infinite responsibility of the one for the other refers to an election that precedes freedom. In coming from an immemorial, anachronical "past," responsibility indicates the "preoriginary" "illeity" of God. The *il* or *ille* of "the most high" is sharply distinguished from the chaotic anonymity of *il y a*; the dimensions of economy, morality, and justice separate the indeterminacy of being from the beyond-all-determinacy of God. However, as the practical and theoretical recognition of the relationship between God and humans, religion cannot be separated from ethics: The only way to venerate God is through devotion to human others.

Besides the two books summarized here, Levinas wrote many articles. Most of these were collected in *Humanisme de l'autre homme* (1972), *De Dieu qui vient à l'idée* (1982), *Hors sujet* (1987), and *Entre nous* (1991).

Like all other philosophers, Levinas has convictions that cannot be reduced to universally shared experiences, common sense, or purely rational principles. In addition to his philosophical work he wrote extensively on Jewish questions from an orthodox Jewish, and especially Talmudic, point of view. In his philosophical writings he quotes the Bible perhaps as often as William Shakespeare or Fëdor Dostoevsky, but these quotations are not meant to replace philosophical justification of his assertions. Phenomenological rigor and emphasis are typical of his method, even where he points beyond the dimensions of phenomena and conceptuality.

See also Consciousness in Phenomenology; Descartes, René; Dostoevsky, Fyodor Mikhailovich; Heidegger, Martin; Husserl, Edmund; Infinity in Theology and Metaphysics; Ontology; Phenomenology; Plato.

Bibliography

A complete bibliography of primary and secondary texts published between 1929 and 1989 is given in Roger Burggraeve, *Emmanuel Levinas; une bibliographie primaire et secondaire (1929–1985) avec complément 1985–1989.* Leuven: Peeters, 1990.

The most important philosophical books of Levinas are:

La théorie de l'intuition dans la phénoménologie de Husserl. Paris: Alcan, 1930; 2nd ed., 1963. Translated by A. Orianne

as *The Theory of Intuition in Husserl's Phenomenology.* Evanston, IL: Northwestern University Press, 1973.

De l'existence à l'existant. Paris, 1947; 2nd ed., 1978. Translated by A. Lingis as *Existence and Existents.* The Hague: Nijhoff, 1978.

Le temps et l'autre. Montpellier, 1979 (2nd ed. of Levinas's contribution to *Le choix, le monde, l'existence* [Paris, 1948]). Translated by R. Cohen as *Time and the Other.* Pittsburgh: Duquesne University Press, 1987.

En découvrant l'existence avec Husserl et Heidegger. Paris, 1949; 2nd ed., Paris: Vrin, 1967. Partially translated by A. Lingis in *Collected Philosophical Papers (v. infra).*

Totalité et Infini. Essai sur l'extériorité. The Hague: Nijhoff, 1961. Translated by A. Lingis as *Totality and Infinity: An Essay on Exteriority.* Pittsburgh: Duquesne University Press, 1969.

Humanisme de l'autre homme. Montpellier: Fata Morgana, 1972.

Autrement qu'être ou au-delà de l'essence. The Hague: Nijhoff, 1974. Translated by A. Lingis as *Otherwise than Being or Beyond Essence.* The Hague: Nijhoff, 1981.

De Dieu qui vient à l'idée. Paris: Vrin, 1982.

Collected Philosophical Papers. Translated by A. Lingis. Boston: Nijhoff, 1987. Contains the English translation of twelve thematic essays from several volumes and journals.

Hors sujet. Montpellier, 1987.

Entre nous: Essais sur le penser-à-l'autre. Paris: Bernard Grasset, 1991.

SECONDARY LITERATURE

Bernasconi, R., and S. Critchley, eds. *Re-reading Levinas.* Bloomington: Indiana University Press, 1991.

Bernasconi, R., and D. Wood, eds. *The Provocation of Levinas: Rethinking the Other.* London: Routledge, 1988.

Chalier, C., and M. Abensour, eds. *Emmanuel Levinas.* Paris: Editions de l'Herne, 1991.

Cohen, R., ed. *Face to Face with Levinas.* Albany: State University of New York Press, 1986.

Greisch, J., and J. Rolland, eds. *Emmanuel Levinas. L'éthique comme philosophie première.* Paris, 1993.

Peperzak, A., ed. *Ethics as First Philosophy: The Significance of Emmanuel Levinas for Religion, Literature and Philosophy.* New York: Routledge, 1995.

Peperzak, A. *To the Other: An Introduction to the Philosophy of Emmanuel Levinas.* West Lafayette, IN: Purdue University Press, 1993.

Wyschogrod, E. *Emmanuel Levinas: The Problem of Ethical Metaphysics.* The Hague: Nijhoff, 1974.

Adriaan Peperzak (1996)

LÉVY-BRUHL, LUCIEN

(1857–1939)

Lucien Lévy-Bruhl, the French philosopher and social anthropologist, was educated at the University of Paris and the École Normale Supérieure. He occupied the chair of philosophy at the Lycée Louis-le-Grand from 1885 to 1895, when he became *maître de conférences* at the Sorbonne; in 1908 he was appointed titular professor. In 1916 he became editor of the *Revue philosophique.*

Lévy-Bruhl's early work was devoted to the history of philosophy, particularly that of Auguste Comte. While still under the influence of Comte and also of Émile Durkheim, he published *La morale et la science des moeurs* (Paris, 1903; translated by E. Lee as *Ethics and Moral Science,* London, 1905). It stressed the need for detailed empirical studies of the diverse moral attitudes and ideas of different societies as well as the adaptation of these ideas to the social structure of the group. He considered such a description and explanation as a preliminary to a possible applied science of morals, which would give men the same power to modify social life as physical technology gives them over natural phenomena.

Lévy-Bruhl did not develop this idea of a moral technology but devoted most of his life to investigating an extremely wide range of anthropological data derived from the reports of other observers. The interest of his work lies in the theoretical ideas that he applied to this material.

Lévy-Bruhl argued that the behavior of men in primitive societies must be understood in terms of Durkheim's concept of "collective representations," which are emotional and mystical rather than intellectual. The primitive man's world is dominated by occult powers, and his thought is "prelogical," following a law of participation and quite indifferent to what civilized man would regard as self-contradictions. For example, the members of a totemic group may regard themselves as actually identical with their totem, as belonging to a continuum of spiritual powers, rather than as existing as distinct individuals. Prelogical concepts imply no systematic unity but "welter, as it were, in an atmosphere of mystical possibilities" (*How Natives Think,* Ch. 3). Space, for instance, is conceived, not as a homogeneous whole, but in terms of the mystical ties binding each tribe to a particular region, the structure of the ties being understood in terms of the various occult forces to which the life of the tribe is subject.

Primitive man is similarly indifferent to conceptions of causality as understood in civilized cultures. For him there is no natural order within which perceptible phenomena are causally interconnected, but, equally, nothing happens by chance. Events are brought about directly, not through any mechanism of secondary causes; they are effected by the imperceptible denizens of an occult realm who have no definite spatiotemporal location and who may be felt as present in several places simultaneously.

Durkheim's followers have criticized Lévy-Bruhl for failing to bring out the connections between primitive collective representations and social structure. He has also been accused of overstressing the extent of prelogical elements in primitive thought. In attempting to reconcile the existence of fairly highly developed arts and crafts in primitive tribes with his denial that such tribes thought at all in terms of logical and causal connections, he held that such manual skills are not based on reasoning but "are guided by a kind of special sense or tact," refined by experience without benefit of reflection. Lévy-Bruhl's most serious philosophical shortcoming, perhaps, is his failure to see anything problematic about the nature of logic itself and the role it plays in civilized life. His identification of logical thought with the thought of Western civilization prevented him from perceiving many important continuities and analogies between primitive and civilized attitudes and practices.

See also Comte, Auguste; Durkheim, Émile; History and Historiography of Philosophy; Logic, History of; Philosophical Anthropology.

Bibliography

ADDITIONAL WORKS BY LÉVY-BRUHL

History of Modern Philosophy in France (1899). New York: B. Franklin, 1971.

La philosophie d'Auguste Comte. Paris: Alcan, 1900. Translated by K. de Braumont-Klein as *The Philosophy of Auguste Comte.* London: Sonnenschein, 1903.

Les fonctions mentales dans les sociétés inférieures. Paris: Alcan, 1910. Translated by L. A. Clare as *How Natives Think.* London: Allen and Unwin, 1926.

La mentalité primitive. Paris: Alcan, 1922. Translated by L. A. Clare as *Primitive Mentality.* New York: Macmillan, 1923.

L'ame primitive. Paris: Alcan, 1927. Translated by L. A. Clare as *The "Soul" of the Primitive.* London: Allen and Unwin, 1928.

La morale et la science des moeurs (1927). 16th ed. Paris: Presses Universitaires de France, 1971.

Le surnaturel et la nature dans la mentalité primitive. Paris: Alcan, 1931. Translated by L. A. Clare as *Primitives and the Supernatural.* London: Allen and Unwin, 1936.

La mythologie primitive. Le monde mythique des Australiens et des Papous. Paris: Alcan, 1935.

Les carnets de Lucien Lévy-Bruhl. Paris: Presses Universitaires de France, 1949.

The Notebooks on Primitive Mentality. Oxford: Blackwell, 1975.

WORKS ON LÉVY-BRUHL

Cailliet, Émile. *Mysticisme et "mentalité mystique." Étude d'un problème posé par les travaux de M. Lévy-Bruhl sur la mentalité primitive.* Paris, 1938.

Leroy, Olivier. *La raison primitive. Essai de réfutation de la théorie du prélogisme.* Paris: Geuthner, 1927.

Peter Winch (1967)

LEWIS, CLARENCE IRVING
(1883–1964)

Clarence Irving Lewis, the American epistemologist, logician, and moral philosopher, was born in Stoneham, Massachusetts, and educated at Harvard University (AB, 1906; PhD, 1910). He taught at the University of California from 1911 to 1920 and at Harvard from 1920 until his retirement in 1953; after 1930 he was the Edward Pierce professor of philosophy. He delivered the Carus Lectures in 1945 and the Woodbridge Lectures in 1954.

Lewis was a student and critic of modern extensional systems of logic and developed a modal logic based on the notion of strict implication. In epistemology and ethics, he was a pragmatic Kantian.

Lewis internalized within himself the great dialogue on knowledge and reality which began with René Descartes and continued with the British empiricists, Immanuel Kant and the German idealists, and the American pragmatists. It may be said that this tortuous development, both in its long history and in the intellectual life of Lewis, is the attempt of the modern mind to achieve consistency and adequacy in its conceptual foundations.

The basic commitments of any philosopher, whether formulated or not, concern the nature and modes of knowledge; they not only determine what is philosophically problematic for him but also determine how intelligibility can be achieved. Lewis modifies the classical certainty theory of knowledge, which maintains that knowing is an infallible state of mind. He contends that it does not make sense to talk about knowledge where there is no possibility of error. Knowing, according to him, is an assertive state of mind that is subject to appraisal as correct or incorrect by virtue of its relationship to what it is about, and also subject to appraisal as justified or unjustified in terms of its grounds or reasons. Thus the apprehension of a sensory given, or, in other words, the occurrence of an appearance, the classical paradigm of empirical knowledge, is not regarded by Lewis as knowledge, for there is no possibility of error. The apprehension of the appearance and its existence are indistinguishable.

Yet Lewis's departure from the tradition is not great. He, too, insists that at the foundation of our knowledge structure there must be certainty and that this is found in knowledge of sensory appearances. This certainty, however, does not reside in the apprehension of the given. Sensory appearances may be linguistically reported in "expressive" language, which denotes and signifies only appearances. Although there can be no error in the apprehension of a given appearance, it is possible to tell lies

about it. Therefore, such reports are statements with truth-values. But still there is no knowledge, for no judgment is made in which the person could be in error. Knowledge is born at the level of what Lewis calls "terminating judgments," which are of the form "'*S* being given, if *A* then *E*,' where [in expressive language] '*A*' represents some mode of action taken to be possible, '*E*' some expected consequence in experience, and '*S*' the sensory cue." For example, there being a red patch in my visual field, if I seem to turn my head to the left, the red patch moves to the right. Such a judgment is not merely the apprehension of a given, or the linguistic expression of such. It embodies a prediction that the red patch will be displaced to the right if the specified condition is fulfilled, which Lewis contends is conclusively verified by the occurrence of the mentioned appearances.

Thus Lewis locates certainty in verified terminating judgments, which are about sensory appearances. Furthermore, he claims that all knowledge about the world is grounded in and derived from such certainties. Although this is more sophisticated than the traditional empiricist's account, it comes to much the same subjectivistic conclusion, namely, that the direct objects of knowledge are subjective and private, and therefore falls heir to all the problems of modern subjectivism. Lewis's major works are devoted to the central and toughest of these problems: how to make intelligible, from within these epistemological commitments, empirical knowledge of the objective world; a priori knowledge, including mathematics, logic, and philosophy itself; and value claims and normative judgments.

EMPIRICAL KNOWLEDGE OF THE OBJECTIVE WORLD

The paradigm of empirical knowledge for Lewis is the verified terminating judgment. It alone can be conclusively verified. All other empirical judgments are nonterminating. They may be shown to be probable but cannot be established with certainty. The probability value they have is conferred upon them by the verification of terminating judgments that they entail. Therefore, a necessary condition for a nonterminating judgment to be confirmable in any degree, and thus meaningful, is for it to entail terminating judgments.

Any statement that purports to be about objects other than appearances, such as physical objects, is nonterminating, and insofar as it is confirmable and therefore meaningful, it entails terminating judgments, which are about appearances only. It would seem that the full meaning of such a statement would be expressible in the terminating statements entailed by it and that, since these statements are about appearances only, the physical-object statement itself would really refer only to appearances. This would be phenomenalism.

Lewis resists this conclusion. He gives two arguments for realism. The first is that although a physical-object statement is intensionally equivalent to an inexhaustible set of terminating statements and the terms in the latter refer only to appearances, the terms in the physical-object statement genuinely denote physical objects. Thus we have two sets of statements, phenomenalistic and physical-object statements. For each physical-object statement there is a set (although inexhaustible) of phenomenalistic statements intensionally equivalent to it. By confirming the phenomenalistic set we confirm its equivalent physical-object statement with the same degree of probability. Yet the two are about radically different kinds of objects, and from knowledge of appearances we derive knowledge of physical objects.

This argument turns upon his theory of meaning. Lewis distinguishes four modes of the meaning of terms: (1) *denotation*, "the class of all actual things to which the term applies" (for example, the denotation of "man" is the class of all actual men, past, present, and future); (2) *comprehension*, "the classification of all possible or consistently thinkable things to which the term would be correctly applicable" (for example, the comprehension of "man" includes not only actual men but those who might have been but were not, like the present writer's sisters, since he has none); (3) *signification*, "that property in things the presence of which indicates that the term correctly applies, and the absence of which indicates that it does not apply" (for example, the property "rationality" is often regarded as included in the signification of "man"), and (4) *intension*, which consists of (*a*) linguistic intension or connotation, all other terms which must be applicable to anything to which the given term is applicable (for example, "animal" must be applicable to anything to which "man" is applicable); and (*b*) sense meaning, the criterion in mind, an imagined operation "by reference to which one is able to apply or refuse to apply the expression in question in the case of the presented, or imagined, things or situations" (for example, the sense meaning of "kilogon" is the imagined operation of counting the sides of a plane figure and the completion of the operation with the count of 1,000). Since he regards "propositions," statements with the assertive factor extracted (for example, "Mary's baking pies"), as terms, these modes of meaning apply to them as well. He further distinguishes between the "holophrastic" meaning of a statement, its

meaning as a whole, and its "analytic" meaning, the meaning of its terms.

His argument is that although the holophrastic intensional meaning of a physical-object statement is the same as that of a set of phenomenalistic statements, the physical-object statement and its corresponding set of phenomenalistic statements are different in their analytic denotive meaning, the former denoting physical objects and the latter appearances.

Lewis rightly maintains that any two expressions that have the same intension have the same signification. Yet if a term denotes a physical object, it must signify a physical-object property. Therefore such a term could not have the same signification as a phenomenalistic term. Hence it seems that a physical-object statement could not have the same intension as a set of phenomenalistic statements.

Lewis senses this difficulty and seeks to avoid it by speaking of intension, in the form of sense meaning, as "that in mind which refers to signification." Appearances are said to signalize objective properties or states of affairs. Yet he gives no account of how this is possible for beings who can apprehend only appearances. How can appearances, as simple occurrences, be signs of anything other than other appearances? It would seem that the only way out of this subjectivistic trap is to regard appearances not as simple occurrences or objects of apprehension but as intentional in nature, as experiences of physical objects that embody truth claims about them which can be assessed as true or false on the basis of their consistency or lack of it with the claims of other experiences.

Lewis's second argument for realism turns upon the interpretation of "if … then …" in terminating judgments. He regards it as a contrary-to-fact conditional, that is, he claims that the truth of the conditional as a whole is independent of the truth-value of the antecedent and therefore may be significantly asserted when the antecedent is known to be false. Therefore, since it does not express a logical relation of entailment or a truth-functional relationship, it must express a real connection, perhaps causality, that holds between the facts or states of affairs located or referred to by the antecedent and the consequent of the conditional sentence. Belief in a real world, he maintains, is belief in such contrary-to-fact conditionals.

It is not clear how this is an argument for realism. Why must independent physical objects be assumed to account for the contrary-to-fact character of terminating judgments? Why couldn't the "real" connection hold between kinds of appearances?

Furthermore, if terminating judgments are to be interpreted in the manner of the contrary-to-fact conditional, does this not compromise their conclusive verifiability? It would seem to introduce an element of generality that would transcend any specific sequence of subjective experiences. In fact, Lewis himself, for other reasons, held that no terminating judgment of the form "*S* being given, if *A* then *E*" is strictly entailed by a physical-object statement. The most we can say, he concluded, is "If *P* [physical-object statement], then when presentation *S* is given and act *A* is performed, it is more or less highly probable that *E* will be observed to follow." Since the statement is inconclusive, it seems that he has given up the terminating character of "terminating" judgments.

Lewis has not, it seems, made a convincing case for realism from within his phenomenalistic foundations. Some have concluded that it is impossible to do so and that the only way out of phenomenalism is to abandon the subjectivistic starting point itself.

A PRIORI KNOWLEDGE

The a priori disciplines, namely, mathematics, logic, and philosophy, were the stronghold of classical rationalism. They were regarded as yielding knowledge, grounded in rational intuition, about the essential and necessary structure of the world. Empiricists, for the most part, claim that such knowledge is only intralinguistic, that it consists of analytic truths, which are said to be uninformative about the world.

Lewis subscribes to the view that all "a priori truth is definitive in nature and rises exclusively from the analysis of concepts." Unlike many empiricists, however, he is not content with merely characterizing a priori knowledge as analytic. For him, concepts, their logical relations, and their relation to the data of sense and the structure of the world are highly problematic. He regards concepts, logical relations, and a priori truths arising from them as the peculiar characteristics of mind. He sets them in contrast with the given data of sense experience, which he regards as brute fact, unlimited and unaffected by the conceptual structure. But these givens would be unintelligible without the a priori criteria of classification provided by mind, criteria which are involved not only in talk about things but even in the experience of objects. Thus, the necessary connections of concepts are embedded in perception, and analytic truths, far from being trivial and only intralinguistic, formulate the a priori structure of the world as experienced and known.

Our basic conceptual structure, and thus our a priori truths, are not fixed and eternal. They consist of deep-seated attitudes grounded in decisions that are somewhat like fiats in certain respects and like deliberate choice in others. There is nothing in our conceptual structure that is not subject to change in the face of continuing experience. This includes such basic decisions as the decision that whatever is to count as real, in contrast with the hallucinatory, must stand in causal relations with other real things. Even the laws of logic, "the parliamentary rules of intelligent thought and action," are subject to change. The only test applicable is pragmatic, the achievement of intelligible order with simplicity, economy, and comprehensiveness in a way that will be conducive to the long-run satisfaction of human needs. Thus, Lewis holds to a pragmatic theory of a priori truth but not of empirical truth.

Philosophy, according to Lewis, is a reflective, critical study of mind and its a priori principles as found in "the thick experience of everyday life," and thus in "the structure of the real world which we know." Although it studies what is implicit in experience, it is analytic and critical in method rather than descriptive. Its function is not only to formulate the conceptual structure built into experience and thought but to sharpen and to correct it. Thus philosophical claims may be analytic in character, like "There is an intelligible order in the objective world." Lewis takes this statement to be analytic on the ground that an intelligible order is an essential mark of the objective world. Whatever lacks a certain minimum order is only subjective, private experience, like dreams and hallucinations. Philosophical claims also may be critical and revisionary, recommending some change in our categorial attitudes, such as "Only the physical is real."

Lewis's theory of the a priori places the conceptual framework between two sets of givens, the presentations of sense, to which concepts apply to yield empirical knowledge, and the values in terms of which the a priori structure is pragmatically tested. It seems that both sensory experiences and values would have to be free of a priori assumptions in order to serve the function ascribed to them. This is a difficult doctrine to maintain.

VALUE CLAIMS AND NORMATIVE JUDGMENTS

The ultimate test of the a priori conceptual framework, according to Lewis, is "the long-run satisfaction of our needs in general." It would seem that value judgments would have to be independent of the conceptual framework that is being pragmatically tested if the test is to be clear-cut and not beg the question. But obviously this would be impossible in the case of basic issues. Although Lewis does not face the problem in these terms, he may be said to blunt the criticism by locating values among sense presentations and by invoking unavoidable imperatives that would be operative in any conceptual framework.

Value, in its most primitive sense, has to do with sense presentations. It is not so much a specific phenomenal quality as a mode or aspect of the given, namely, the given as gratifying or grievous. The only thing that is intrinsically good is liked or wanted subjective experience. In addition to the immediately found intrinsic value of an experience, it may be said to have contributory value by virtue of the contribution it makes to the total value quality of the conscious life of which it is a part. Such a life, he contends, is not simply a sum of its parts. So the contributory value of an experience is quite different from its intrinsic value. Objects of experience are said to be extrinsically good or bad according to their capacity to produce experiences which are satisfying or unpleasant.

Thus, for Lewis, value knowledge is a form of empirical knowledge. There are both terminating and nonterminating value judgments. The former are subjective statements of intrinsic and contributory value; the latter are objective statements about extrinsic values. Judgments of right and wrong, however, are not empirical in character. They are determinable only by reference to rules or principles that refer to values in their prescriptions. He regards the basic rational imperative to be so to think and so to act that later you will not be sorry. The only way this can be achieved is for decisions to be guided by objective knowledge rather than merely by the affective quality of immediate experience. In the area of morals, this requires that we respect others as the realities we know them to be, "as creatures whose gratifications and griefs have the same poignant factuality as our own; and as creatures who, like ourselves, find it imperative to govern themselves in light of the cognitive apprehensions vouchsafed to them by decisions which they themselves reach, and by reference to values discoverable to them."

Any attempt to prove the validity of such principles can only appeal to an antecedent recognition of them. They must be recognized by all who make decisions, all who think and act. Genuine skepticism with regard to judgments of right and wrong, good and bad, would be impossible, for on such a basis even doubt itself would be meaningless.

The question remains: Is the conceptual framework in which normative and value knowledge is formulated pragmatically testable, and, if so, just what could such a pragmatic test amount to? If it is not so testable, then it would seem that in the end Lewis is not a pragmatist after all.

See also A Priori and A Posteriori; Descartes, René; Kant, Immanuel; Knowledge, A Priori; Meaning; Modal Logic; Phenomenalism; Pragmatism; Propositions; Rationalism; Realism; Value and Valuation.

Bibliography

WORKS BY LEWIS

A Survey of Symbolic Logic. Berkeley: University of California Press, 1918.

Mind and the World-Order. New York: Scribners, 1929.

Symbolic Logic. New York: Appleton-Century, 1932. Written with C. H. Langford.

An Analysis of Knowledge and Valuation. La Salle, IL: Open Court, 1946.

The Ground and Nature of the Right. New York: Columbia University Press, 1955.

Our Social Inheritance. Bloomington: Indiana University Press, 1957.

WORKS ON LEWIS

Baylis, C. A. "C. I. Lewis, *Mind and the World-Order.*" *Journal of Philosophy* 27 (1930): 320–327.

Boas, George. "Mr. Lewis's Theory of Meaning." *Journal of Philosophy* 28 (1931): 314–325.

Chisholm, Roderick M. "The Problem of Empiricism." *Journal of Philosophy* 45 (1948): 512–517.

Chisholm, Roderick M., Herbert Feigl et al. *Philosophy.* Englewood Cliffs, NJ: Prentice-Hall, 1964.

Ducasse, C. J. "C. I. Lewis' *Analysis of Knowledge and Valuation.*" *Philosophical Review* 57 (1948): 260–280.

Firth, R., R. B. Brandt et al. "Commemorative Symposium on C. I. Lewis." *Journal of Philosophy* 61 (1964): 545–570.

Frankena, William. "C. I. Lewis on the Ground and Nature of the Right." *Journal of Philosophy* 61 (1964): 489–496.

Frankena, William. "Lewis's Imperatives of Right." *Philosophical Studies* 14 (1963): 25–28.

Henle, Paul. "Lewis's *An Analysis of Knowledge and Valuation.*" *Journal of Philosophy* 45 (1948): 524–532.

Kusoy, B. K. *Kant and Current Philosophical Issues.* New York, 1961.

Pratt, J. B. "Logical Positivism and Professor Lewis." *Journal of Philosophy* 31 (1934): 701–710.

Schilpp, P. A. *The Philosophy of C. I. Lewis.* Vol. XIII, Library of Living Philosophers. La Salle, IL: Open Court, 1966.

Stace, W. T. "C. I. Lewis: *An Analysis of Knowledge and Valuation.*" *Mind,* n.s., 57 (1948): 71–85.

White, M. G. "Value and Obligation in Dewey and Lewis." *Philosophical Review* 58 (1949): 321–329.

E. M. Adams (1967)

LEWIS, C. S. (CLIVE STAPLES)

(1898–1963)

C. S. Lewis was a British teacher, writer, and critic. He was born and raised in Belfast but spent most of his academic career at Oxford. After having volunteered for the army and subsequently getting wounded, in 1917, he returned to Oxford and took first class honors in "Greats" (philosophy and classics). Shortly thereafter he taught philosophy at Oxford as a substitute for Edgar Carrit, his former tutor in philosophy while Carrit was on leave as a visiting professor at the University of Michigan. Finding no opportunity for teaching in classics or philosophy, and having also gotten first class honors in English, Lewis was elected to a fellowship in English at Magdalene College, where he taught for thirty years. Toward the end of his academic career he was appointed to a newly created Chair of Medieval and Renaissance English at Cambridge. His strictly academic work was concentrated on the ideas rather than the literary forms of medieval and renaissance English writers.

Early in his career at Oxford Lewis became a convert, first to theism and then to Christianity. During World War II he was asked to give lectures about Christianity on the BBC: Printed in book form, these were the basis of his most famous popular work, *Mere Christianity* (2001 [1942]). Other popular works were *The Problem of Pain* (2001 [1940]), *Miracles* (2001 [1947]), and *The Screwtape Letters* (2001 [1942]). In 1945, Lewis argued with G. E. N. Anscombe about a claim in Chapter 5 of *Miracles* that naturalism is self-refuting, for it says that all our thoughts are ultimately traceable to the blind working of chance and that no thought is valid if it can be fully explained as the result of irrational causes. Anscombe distinguished between "irrational" causes and "nonrational" causes and argued that being the result of "nonrational" causes does not make our reasoning invalid. Lewis, in reply, says the "valid" in the logician's sense is not the correct word for what he meant and distinguished between "reasons" and "causes" (Hooper 1979). Some have thought that he lost that argument. He revised the chapter of *Miracles* which Anscombe had criticized, and Anscombe, at least, felt that the revision answered much of her original objections (Purtill 2004). Late in his life (in 1957), Lewis married Helen Joy Davidman, who was dying of cancer. She surprisingly (and perhaps miraculously) recovered and they spent three happy years together.

After her death, Lewis wrote (anonymously) *A Grief Observed* (2001 [1961]), which some scholars have held

demonstrates that he had lost his faith, or at least his belief in the rational justification of Christianity. However, a more careful reading shows that his own description of Christianity to a friend is true: "It ends in faith, but begins with the blackest of doubts en route" (unpublished letter quoted in Purtill 2004, p. 25). It is useful to compare this book with two of his later works: *Till We Have Faces*, a fictional account of a woman who began writing a book as a complaint against the gods (the account is set in classical times) and "ends in faith"; and *Letters to Malcolm* (2002 [1964]), which touches on some of the same themes.

Lewis was, in this author's judgment as well as the judgment of other critics, a great master of English prose and a powerful writer of fiction with underlying religious themes: the seven books of the *Chronicles of Narnia*, the "space" trilogy, and *Till We Have Faces*.

Philosophically speaking, Lewis's work, both nonfiction and fiction, has a number of characteristics:

(1) He argues for his points on the basis of reason and experience. As he says in an essay, "There is, of course, no question … of belief without evidence … or in the teeth of evidence … if anyone expects that, I certainly do not" (Lewis 2001 [1955], p. 17);

(2) He thinks of faith as a rational acceptance and of "temptations against faith" as emotional reactions when we find it would be much more convenient *not* to believe;

(3) He accepts miracles and uses them as evidence for Christianity, first refuting the arguments of Hume and others against the *possibility* of miracles or the possibility of knowing them, and then arguing historically that miracles have occurred;

(4) Miracles, as Lewis defines them, depend on the existence of God. Lewis argues for God's existence using variations of the moral argument and the argument from design, especially a version of what Victor Reppert has called "the argument from reason" which argues that to really trust our reason we need the existence of God. For the moral argument, Lewis agrees with other philosophers that "if God does not exist anything is permitted" and by contraposition that "if not everything is permitted [as he argues from our moral experience] then God must exist."

(5) Lewis contrasts Christianity with other forms of belief—such as naturalism, Hinduism, and so on—and argues that Christianity explains the facts of experience better than other forms of belief.

(6) Lewis grants that the problems of moral and natural evil are the most powerful against a belief in a loving, omnipotent God, and addresses both in *The Problem of Pain* and elsewhere.

Professional philosophers may find many of Lewis's arguments oversimplifications; Lewis would probably grant this for his more popular works, which were intended for intelligent nonprofessionals. However, this leads to a situation where philosophical argument can begin. What are the alleged oversimplifications and how can they be repaired? Lewis's experience with Anscombe showed he was capable of doing this, as does his work in less popular works addressed to academic or clerical audiences.

The talent that made him a good writer of fiction carries over to his nonfictional works; he is a poet, as well as a logician, and employs a gift for metaphor and analogy in his statements of arguments. Lewis has been called "perhaps the twentieth century's most popular proponent of Faith based on reason" (Nicholi 2002, p. 3). Many opponents of Christianity have taken Lewis's arguments seriously, especially those scholars who, such as Antony Flew, wish to be fair to Christianity and try to refute its best arguments. Many supporters of Christianity, both nonprofessional and academic alike, would give Lewis major credit for beginning the process that led them to Christianity.

See also Anscombe, Gertrude Elizabeth Margaret; Experience; Evil, The Problem of; Hume, David; Immortality; Miracles; Reason.

Bibliography

WORKS BY LEWIS

The Problem of Pain (1940). San Francisco: HarperCollins, 2001.

The Screwtape Letters (1942). San Francisco: HarperCollins, 2001.

Mere Christianity (1943). San Francisco: HarperCollins, 2003.

The Great Divorce (1945). San Francisco: HarperCollins, 2001.

Miracles (1947). San Francisco: HarperCollins, 2001.

"An Obstinacy in Belief." In *The World's Last Night* (1955). New York: Harcourt Harvest Books, 2002.

Till We Have Faces (1956). San Diego, CA: Harcourt Harvest Books, 2002.

A Grief Observed (1961). San Francisco: HarperCollins, 2001.

Letters to Malcolm (1964). New York: Harvest Books, 2002.

WORKS ON LEWIS

Green, Roger Lancelon, and Walter Hooper. *C. S. Lewis: A Biography*. San Diego: Harcourt Brace, 1995.

Hooper, Walter. "Oxford Bonnie Fighter." In *C. S. Lewis at the Breakfast Table*, edited by James T. Como. New York: Macmillan, 1979.

Kreeft, Peter. *C. S. Lewis for the Third Millennium: Six Essays on the Abolition of Man*. San Francisco: Ignatius Press, 1994.

Nicholi, Armand M. *The Question of God: C. S. Lewis and Sigmund Freud Debate God, Love, Sex, and the Meaning of Life*. New York: Free Press, 2002.

Purtill, Richard. *C. S. Lewis's Case for the Christian Faith*. New York: Harper Collins, 1981. New. ed., San Francisco: Ignatius Press, 2004.

Purtill, Richard. "Did C. S. Lewis Lose his Faith?" In *A Christian for All Christians: Essays in Honour of C. S. Lewis*. London: Hodder and Staughton, 1990.

Reppert, Victor. *C. S. Lewis's Dangerous Idea: A Philosophical Defense of Lewis's Argument from Reason*. Downers Grove, IL: Intervarsity Press, 1998.

Richard Purtill (2005)

LEWIS, DAVID

(1941–2001)

David Lewis was born in Oberlin, Ohio. He studied as an undergraduate at Swarthmore College before gaining a PhD in philosophy from Harvard University in 1967 where he studied with Willard van Orman Quine. His first job was at the University of California at Los Angeles, where he worked from 1966–1970, before moving to Princeton University where he worked for the rest of his career. Lewis published four monographs and more than one hundred papers, most of which have been gathered into five volumes of his collected papers. Lewis made contributions to virtually every area of contemporary Anglo-American philosophy but is probably best known for his contributions to metaphysics, in particular, his work on modality (necessity and possibility) and possible worlds and also his theories of laws of nature, causation, and chance. His work in the philosophy of mind has also been influential, as has his work on conventions and language.

MODALITY

Some of Lewis's best known work is in the metaphysics of modality: that is, his account of the nature of necessity and possibility. Lewis thought it was important to make sense of what we are doing when we talk about different possibilities that seem open, or when we say that certain facts (such that 2+2=4) are necessary, or that it is impossible for them to be otherwise. Lewis held, along with others whose claims about possibility and necessity were to be understood as implicitly generalizing over possible worlds, complete ways things could be: To say something was possible was to say that it occurred in at least one possible world, and for something to be necessary was for it to obtain in all possible worlds. Where Lewis was nearly unique was his account of what these other possible worlds were.

According to Lewis, possible worlds were large spatiotemporal regions filled with objects and events of the same kind as those in our world, except, of course, that every possible sort of thing is found in one world or other. So Lewis's worlds contain people and trees and galaxies and tables; but also dragons, extra-spatiotemporal dimensions, ghosts, and so on. This construal of possible worlds became known as *modal realism* Despite the counterintuitive nature of this theory, Lewis showed that it brought with it many advantages, and he argued that attempts to construe possibilities as some sort of abstract object (*ersatzism* about possible worlds, in Lewis's vocabulary), failed to provide an analysis of modality, and many varieties suffered crippling internal problems.

Lewis also suggested a novel way of dealing with *de re* necessities and possibilities (possibilities or necessities for an object rather than as concerning the status of a proposition). Lewis argued that these were best analyzed using counterpart theory: where what is possible for me is what happens for one of my counterparts in another world. Since Lewis held that, strictly speaking, each possible individual was part of only one possible world, he could avoid some of the puzzles about trans-world criteria for identity. In addition, counterpart theory allowed more flexibility than literal identity would. Lewis argued, for instance, that the counterpart relation need not be transitive (so something that could happen to one of my counterparts need not be something that is possible for me) though a failure of transitivity is harder to understand if it is literal identity across possible worlds that is required for *de re* possibility (i.e., if something has to be literally happening to me in some other possible world in order for it to be possible for me). Lewis also allowed that there were multiple counterpart relations, which might give rise to multiple kinds of *de re* possibility for an object. So, for example, what the counterparts of an object are when that object is considered as a statue might be different from what the counterparts of that object are when the object is considered to be a lump of bronze. Lewis could thus allow that what we appropriately say is possible for the statue is different from what we appropriately say is possible for the piece of bronze even though the two objects might nevertheless be identical.

COUNTERFACTUALS, LAWS, CAUSATION, AND CHANCE

Issues about contrary-to-fact-conditionals, laws of nature, causation, and chance are often thought to be connected, and Lewis's contributions to these topics formed a unified neo-Humean system. Lewis's book *Counterfactuals* (1973a) offered an analysis of conditionals of the form *if* p, *it would have been the case that* q in terms of possible worlds: a conditional such as *if dolphins had had legs, they would have walked on land* is true if and only if the *nearest* possible world where dolphins have legs is one where dolphins walk on land. *Nearness*, in turn, is analyzed as overall similarity in salient respects: it is thus context-dependent, and Lewis had more to say about what sort of similarity is significant for particular sorts of these so-called counterfactual conditionals, for example the ones employed in causal reasoning.

This analysis of counterfactual claims has several advantages. It is formally tractable, yielding a logic of counterfactual judgments with some initially surprising features that do seem to correspond to features of our ordinary counterfactual judgments. For example, Lewis's system delivers the result that *strengthening the antecedent* is invalid: that is, the inference (if *p* then *q*), therefore (if *p* and *r* then *q*) is invalid. But consider this argument: If I leave now, I will catch the train; therefore, if I leave now and am assassinated on the way I will catch the train. The premise might well be true and the conclusion false if I run no real risk of being assassinated.

In addition, since the analysis of these conditionals does not itself appeal to, for example, dispositions or causation, it leaves the way free for counterfactual analyses of other puzzling parts of metaphysics. And, indeed, Lewis championed a counterfactual analysis of causation: At a first pass, an event *C* causes an event *E* if and only if both *C* and *E* occur and had *C* not occurred, *E* would not have occurred either. A lot more than this first pass is required for an adequate counterfactual account of causation: Sometimes *E* would have happened in any case, even without *C*, for example, if *E* is overdetermined. Lewis experimented with a number of counterfactual theories of causation: Their development can be seen in Lewis 1973b (and see especially the postscripts in Lewis 1986b), Lewis 1979a, and most recently Lewis 2000 and Lewis 2004.

The connection between counterfactuals and causation, on the one hand, and laws of nature, on the other hand, is slightly circuitous in Lewis, but it is another key connection in his overall system. Lewis defended a regularity theory of laws of nature: Following Ramsey, Lewis held that the laws of nature were given by the set of generalizations that provided the best tradeoff of simplicity and strength in capturing the goings-on of the world. Since the laws supervene on the patterns of particular matters of fact, at this point, at least, Lewis's metaphysical posits are minimal.

Even though the laws are only descriptions of certain privileged regularities, they make a difference to which counterfactuals are true, in Lewis's system, because similarity with respect to whether our laws hold is one of the most important components in the kind of similarity relevant for the nearness relation between possible worlds central to Lewis's analysis of counterfactuals. So when some event *A* would follow as a matter of law from another event *B*, the nearest world where *A* occurs will be one where *B* also does. Thus mere patterns of particular occurrences give rise to laws of nature, counterfactual dependencies, and so to causation—at least, if Lewis is right. Lewis extended his regularity framework to handle objective chances as well: Another member of the *nomic family* was explained, ultimately, in terms of regularities in particular events.

MIND

Lewis made contributions to several areas of the philosophy of mind. First in importance is his defense of an identity theory of the mind. Lewis characterized mental states according to the role attributed to them in our ordinary *folk* understanding of the mind: A belief, for example, is a state that tends to go together with desires with certain contents to produce certain sorts of actions. Folk psychology, when articulated, describes causal *roles* for each different sort of mental state (beliefs, desires, pains, emotions), and these roles are interdefined so that the typical causal profile of a belief is specified partly in terms of other beliefs it tends to cause, partly in terms of perceptions that tend to cause it, how it interacts with desires, and so on.

Armed with this role statement of the typical causes and effects of mental states, Lewis then argued that mental states are identical to those physical states in us that play these causal roles: So Tom's belief that it is raining is identical to the brain state of his that is typically caused by the sight and sound of rain, and typically goes together with other brain states to yield umbrella-grabbing behavior, and so on. This may well mean that which type of physical state is identical to which type of mental state may vary from subject to subject: Lewis says that which physical state is identical to a given mental state depends on what causal roles that state typically plays in the kind

of thing that has it. So in humans pain will be a certain sort of brain state while in advanced robots it may be some electronic state, and in extraterrestrials it might be a matter of how gases are distributed within internal bladders.

This typical-for-the-species criterion allows both for *multiple realizability*: pain-in-aliens or pain-in-invertebrates may not be the same physical state as pain-in-humans; and we loosen the behaviorist insistence that pain must be the state, whatever it is, that produces pain behavior since, for example, an atypical human may have the state that typically causes pain reactions but makes no outward show of it, or even engages in some nonstandard behavior (imagine a *madman* who whistles, but shows no discomfort, whenever he is in the state that produces pain behavior in normal humans). The view is still a type–type identity theory, according to Lewis, because, for example, the type pain-in-humans can be identified with a particular physical property (e.g. C-fiber-firing-in-humans) even though there is no unified physical type corresponding to *pain* or *belief* simpliciter.

LANGUAGE AND CONVENTION

In Lewis's first book, *Convention* (1969), he developed a theory of conventions as patterns of mutual expectation and conditional intentions. Roughly, according to Lewis, there is a convention in a population to act in a certain way in certain circumstances if everyone does tend to behave in that way, everyone has the conditional intention to continue behaving that way, conditional on everyone else so acting, and this is common knowledge. Finally, there must also be some other alternative action that people are deciding against: We all breathe oxygen, intend to continue and know that we intend to continue, but this does not make our practice of oxygen-breathing conventional since we all have no choice. He claimed these patterns could arise fairly spontaneously (certainly without the existence of an explicit agreement) and that they tended to arise to solve coordination problems: common cases of collective action where everyone does better by coordinating their activities than if everyone does their own thing. (A decision about which side of the road to drive on is an example: The most important thing is not whether people drive on the left or right hand side but, rather, that either everyone does the one or everyone does the other.)

Lewis argued that we could understand what it was for a population to use a language as a matter of convention, in his sense. In "Language and Languages" (1975), Lewis explained how we could integrate the formal, abstract theories of languages as functions from expressions to truth-conditions, on the one hand, with theories of language that concentrate on practices of language usage. Lewis also made significant contributions to the formal theory of language and philosophical semantics—his "General Semantics" (1970) is a prime example. Lewis was also responsible for a lot of work exploring the role of context in language: His "Scorekeeping in a Language Game" (1979b) is a classic in this area.

See also Hume, David: Metaphysics; Philosophy of Language; Philosophy of Mind; Quine, Willard Van Orman; Ramsey, Frank Plumpton.

Bibliography

MONOGRAPHS BY DAVID LEWIS

Convention: A Philosophical Study. Cambridge MA: Harvard University Press, 1969.

Counterfactuals. Oxford: Basil Blackwell, 1973a.

On the Plurality of Worlds. Oxford: Basil Blackwell, 1986a.

Parts of Classes. Oxford: Basil Blackwell, 1990.

VOLUMES OF COLLECTED PAPERS

Philosophical Papers. Vol I. Oxford: Oxford University Press, 1983.

Philosophical Papers. Vol. II. Oxford: Oxford University Press 1986b.

Papers in Philosophical Logic. Cambridge. U.K.: Cambridge University Press, 1998.

Papers in Metaphysics and Epistemology. Cambridge, U.K.: Cambridge University Press, 1999.

Papers in Ethics and Social Philosophy. Cambridge, U.K.: Cambridge University Press, 2000.

OTHER WORKS CITED

"General Semantics." *Synthese* 22 (1970): 18–67.

"Causation." *Journal of Philosophy* 70 (1973b): 556–467.

"Languages and Language." *Minnesota Studies in the Philosophy of Science* 7 (1975): 3–35.

"Counterfactual Dependence and Time's Arrow." *Nôus* 13 (1979a): 455–476.

"Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979b): 339–359.

"Mad Pain and Martian Pain." In *Readings in the Philosophy of Psychology*, Vol. 1, 216–232. Cambridge MA: Harvard University Press, 1980.

"Humean Supervenience Debugged." *Mind* 103 (1994): 473–90.

"Causation as Influence." *Journal of Philosophy* 97 (2000): 182–97.

"Causation as Influence" (extended version). In *Causation and Counterfactuals*, edited by J. Collins, N. Hall, and L. Paul, 75–106. Cambridge, MA: MIT Press, 2004.

Daniel Nolan (2005)

LI AO
(774?–836)

Li Ao is perhaps the thinker in Tang China (618–907) who contributed most to a new version of Confucian philosophy that addressed issues of human nature and spiritual cultivation. By Li's time, questions in this area had been left to Buddhism and Daoism for centuries, whereas the intellectual elite in general considered Confucianism solely the authority in family and political lives. Li's importance as a thinker comes entirely from a single treatise: the *Fuxing shu* (Writings on returning to one's true nature). It is arguably the first post-Han (206 BCE–220 CE) text that gave an original treatment on the topics of human nature and spirituality from a Confucian stance.

The theme of the *Fuxing shu* is how to become a sage, the Confucian ideal of personality. Li holds that a sage is a person who has realized his "nature" (*xing*), the character of which can be described as "sincerity" (*cheng*). The nature of human beings is bestowed on them by heaven, and all people share the same nature. The reason why hardly anyone becomes a sage is that people's "emotions" (*qing*) obscure their true nature.

As to the method of becoming a sage, Li contends that if one quiets down and thus clarifies one's emotions, one's nature will be revealed and will direct one's life. One can then naturally act in a proper manner—that is, in accord with Confucian behavioral norms. The central point here is that the true nature of humans only exists in the state of tranquility. Yet tranquility of one's nature is not equivalent to suspension of emotions, because the latter will inevitably shift to a state of movement. People should learn to respond to the world directly with their true nature. The nature that is at the same time tranquil and able to have a full control of one's life exists beyond the level of emotions.

At least two issues regarding the *Fuxing shu* deserve attention here. The first is the subject of this treatise. The search for sagehood through self-cultivation was a significant notion in classical Confucianism, but went almost absent after the Han. It was owing to the Buddhist concern with Buddhahood that the perfection of human existence through spiritual cultivation became a major issue in medieval Chinese thought. Li's revival of a dormant Confucian subject is in itself an indication that the *Fuxing shu* represents a Confucian response to the centuries-old dominance of Buddhism and Daoism in the realm of metaphysical and spiritual philosophy. Li's project anticipates the endeavor of neo-Confucianism in Song times (960–1279).

Then there is the much studied and debated issue: the sources of the originality of the *Fuxing shu*. It is clear that medieval Buddhism and Daoism not only gave birth to the theme of Li's treatise, but also affected its ideas in a substantial way. The sharp contrast between "nature" and "emotions" is a case in point. This distinction is not a salient feature of classical and Han Confucianism. Even for those Confucian thinkers believing that moral values were rooted in the essence of human beings, goodness did not just exist in one's nature. It was more important to realize people's moral potential in their actual lives filled with all kinds of emotions. Simply put, in early Confucianism there was no such notion that a return to one's nature, defined as the original state of human existence, represented the perfection of human life. This idea, which is at the core of Li's theory, owed its origins principally to classical and religious Daoism. The most crucial formative force behind this idea seems to be the fundamental Daoist belief that the ideal state of life lies in its reunion with its roots—indeed with the "primordial breath" (*yuanqi*) of the universe.

Although Li borrows heavily from religious ideas current in his time, it is unmistaken that his aim is providing a theoretical framework for a Confucian way of self-cultivation. Li emphasizes that once revealed, one's nature will lead to correct knowledge and actions, that is, those in line with Confucian values. One may say that Li uses a great deal of Buddhist and Daoist material to build a Confucian house. He was one of the rare individuals in the history of ideas to really make a breakthrough.

See also Confucianism; Han Yu.

Bibliography

Barrett, Timothy Hugh. "Buddhism, Daoism and Confucianism in the Thought of Li Ao." PhD diss. Yale University, 1978.

Barrett, Timothy Hugh. *Li Ao: Buddhist, Taoist, or Neo-Confucian*. New York: Oxford University Press, 1992.

Chen, Jo-shui. "'Fuxing shu' sixiang yuanyuan zaitan: Han Tang xinxing guannian shi zhi yizhang." *Zhongyang yanjiuyuan lishi yuyan yanjiusuo jikan* 69 (3) (1998): 423–482.

Jo-shui Chen (2005)

LIAR PARADOX, THE

Attributions of truth and falsehood under certain conditions generate the "liar paradox." The most famous illustration of this comes from the Epistle to Titus, in which

St. Paul quotes approvingly a remark attributed to Epimenides: "One of themselves, even a prophet of their own, said, The Cretans are always liars, evil beasts, slow bellies. This witness is true" (King James version). Let us suppose that Epimenides, the Cretan prophet, did say that the Cretans are always liars, and let us consider the status of his utterance—call it *E*—under the following two conditions. (1) A Cretan utterance counts as a lie if and only if it is untrue. (2) All Cretan utterances, except perhaps *E*, are untrue. Now, if *E* is true, then, since *E* is a Cretan utterance, not all Cretan utterances are untrue. Hence, Cretans are not always liars (by (1)), and so *E* must be untrue. On the other hand, if *E* is untrue, then indeed all Cretan utterances are untrue (by (2)). Hence, Cretans are always liars (by (1)), and so *E* is true after all. Both the hypotheses, that *E* is true and that *E* is not true, yield, therefore, a contradiction. Yet the steps in the argument are all apparently valid, and the initial setup is not impossible. This is the liar paradox.

The paradox was discovered by Eubulides of Miletus (fourth century BCE) and has exercised logicians down the ages to the present time. (See Bocheński 1961, Spade 1988.) For principally two reasons, interest in the paradox was especially great in the twentieth century. First, arguments similar to that found in the liar wreaked havoc in several prominent logical systems (e.g., those of Gottlob Frege and Alonzo Church). This prompted a search for systems that were immune from paradox. Second, the rise of semantical studies created a need for a better understanding of the notions of truth, reference, and the like. The notions are fundamental to semantical investigations, but the paradoxes reveal a profound gap in our understanding of them. (The notion of reference, like other semantical notions, exhibits, under certain conditions, paradoxical behavior.)

The liar and related paradoxes raise a number of difficult conceptual problems. One is the normative problem of designing paradox-free notions of truth, reference, and the like. Another is the descriptive problem of understanding the workings of our ordinary, paradox-laden notions. The work on the paradoxes in the first half of the twentieth century is, perhaps, best viewed as addressing the normative problem. The work in the second half is best viewed as addressing the descriptive problem. Some of this work is outlined below.

Let us sharpen the descriptive problem a little. For simplicity, let us restrict our attention to a fragment, *L,* of our language that contains no problematic terms other than "true." All other terms in *L* have, let us suppose, a classical interpretation. How should "true" be interpreted? A natural demand is that the interpretation must validate the T-biconditionals, that is, all sentences of the form,

(T) "*B*" is true if and only if (iff) *B*,

where *B* is a sentence of *L*. The argument of the liar paradox shows, however, that every possible classical interpretation of "true" is bound to make some T-biconditionals false. (This is a version of Alfred Tarski's indefinability theorem.) How, then, should we interpret "true"? Should we abandon the natural demand? Or the classical framework? Or the naive reading of the T-biconditionals? Essentially, the first course is followed in the contextual approach, the second in the fixed-point approach, and the third in the revision approach.

THE CONTEXTUAL APPROACH

This approach takes "true" to be a context-sensitive term. Just as the interpretation of "fish this long" varies with contextually supplied information about length, similarly, on the contextual approach, with "true": Its interpretation also depends upon contextual information. There is no consensus, however, on the specific information needed for interpretation. In the levels theory due to Tyler Burge and Charles Parsons, the context supplies the level at which "true" is interpreted in a Tarskian hierarchy of truth predicates. In the Austinian theory of truth developed by Jon Barwise and John Etchemendy, the relevant contextual parameter is the "portion" of the world that a proposition is about. In the singularity theory of Keith Simmons, the relevant information includes certain of the speaker's intentions.

Contextual theories assign to each occurrence of "true" a classical interpretation, though not the same one to all occurrences. This has several characteristic consequences: (1) Occurrences of "true" do not express global truth for the entire language (by Tarski's indefinability theorem). They express instead restricted or "quasi" notions of truth; the former possibility is realized in the levels theory, the latter in the singularity theory. (2) Truth attributions, even paradoxical ones, have a classical truth-value. Paradox is explained as arising from a subtle, unnoticed, shift in some contextual parameter. (3) Classical forms of reasoning are preserved. But caution is in order here: Whether an argument exemplifies a classically valid form turns out to be nontrivial. For example, the argument "*a* is true, $a = b$; therefore, *b* is true" exemplifies a classically valid form only if "true" is interpreted uniformly, but this is nontrivial on the contextual approach.

THE FIXED-POINT APPROACH

This approach interprets "true" nonclassically. It rests on an important observation of Saul Kripke, Robert Martin, and Peter Woodruff. Consider again the language *L,* and assign to "true" an arbitrary partial interpretation $\langle U, V\rangle$, where *U* is the extension and *V* the antiextension (i.e., the objects of which the predicate is false). We can use one of the partial-valued schemes (say, Strong Kleene) to determine the sentences of *L* that are true (U'), false (V'), and neither-true-nor-false. This semantical reflection defines a function, κ, on partial interpretations; $\kappa(\langle U, V\rangle) = \langle U', V'\rangle$. The important observation is that κ has a fixed point: There exist $\langle U, V\rangle$ such that $\kappa(\langle U, V\rangle) = \langle U, V\rangle$.

Certain partial-valued schemes have a least fixed point, which is a particularly attractive interpretation for "true." It is also the product of an appealing iterative construction: We begin by supposing that we are entirely ignorant of the extension and the antiextension of "true"; we set them both to be ∅ (the null set). Despite the ignorance, we can assert some sentences and deny others. The rule "Assert '*B* is true' for all assertible *B*; assert '*B* is not true' for all deniable *B*" entitles us to a new, richer interpretation, $\kappa(\langle\emptyset, \emptyset\rangle)$, for "true." But now we can assert (deny) more sentences. The rule entitles us to a yet richer interpretation $\kappa(\kappa(\langle\emptyset, \emptyset\rangle))$. The process, if repeated sufficiently many times, saturates at the least fixed point.

Under fixed-point interpretations, the extension of "true" consists precisely of the truths and the antiextension of falsehoods. The T-biconditionals are, therefore, validated. They are not, however, expressible in *L* itself: fixed points exist only when certain three-valued functions, including the relevant "iff," are inexpressible in *L.*

THE REVISION APPROACH

This approach holds truth to be a circular concept. It is motivated by the observation that truth behaves in a strikingly parallel way to concepts with circular definitions. Suppose we define *G* thus:

$$x \text{ is } G =_{\text{Df}} x \text{ is a philosopher distinct from Plato } \textit{or} \; x \text{ is Plato but not } G.$$

The definition is circular, but it does impart some meaning to *G. G* has, like truth, unproblematic application on a large range of objects. It applies to all philosophers distinct from Plato and fails to apply to nonphilosophers. On one object, Plato, *G* behaves paradoxically. If we declare Plato is *G,* then the definition rules that he is not *G*; if we declare he is not *G,* the definition rules that he is *G.* This parallels exactly the behavior of truth in the liar paradox.

The revision account of truth rests on general theories of definitions, theories that make semantic sense of circular (and mutually interdependent) definitions. Central to these theories are the following ideas. (1) A circular definition does not, in general, determine a classical extension for the definiendum (the term defined). (2) It determines instead a rule of revision. Given a hypothesis about the extension of the definiendum *G,* the definition yields a revised extension for *G,* one consisting of objects that satisfy the definiens (the right side of the definition). (3) Repeated applications of the revision rule to arbitrary hypotheses reveal both the unproblematic and the pathological behavior of the definiendum. On the unproblematic the revision rule yields a definite and stable verdict, irrespective of the initial hypothesis. On the pathological this ideal state does not obtain.

The ingredient needed to construct a theory of truth once we have a general theory of definitions is minimal: It is just the T-biconditionals, with "iff" read as "$=_{\text{Df.}}$" This reading was suggested by Tarski, but, as it results in a circular definition, it can be implemented only within a general theory of definitions. Under the reading, the T-biconditionals yield a rule of revision. Repeated applications of this rule generate patterns that explain the ordinary and the pathological behavior of truth. The revision approach thus sees the liar paradox as arising from a circularity in truth. The approach has been developed by, among others, Anil Gupta, Hans Herzberger, and Nuel Belnap.

The three approaches, it should be stressed, do not exhaust the rich array of responses to the paradoxes in the twentieth century.

See also Church, Alonzo; Correspondence Theory of Truth; Frege, Gottlob; Kripke, Saul; Logical Paradoxes; Plato; Russell, Bertrand Arthur William; Tarski, Alfred; Types, Theory of.

Bibliography

Antonelli, A. "Non-Well-Founded Sets via Revision Rules." *Journal of Philosophical Logic* 23 (1994): 633–679.

Barwise, J., and J. Etchemendy. *The Liar: An Essay on Truth and Circularity.* New York: Oxford University Press, 1987.

Bocheński, I. M. *A History of Formal Logic.* Notre Dame, IN, 1961.

Chapuis, A. "Alternative Revision Theories of Truth." *Journal of Philosophical Logic* 24 (1996).

Epstein, R. L. "A Theory of Truth Based on a Medieval Solution to the Liar Paradox." *History and Philosophy of Logic* 13 (1992): 149–177.

Gaifman, H. "Pointers to Truth." *Journal of Philosophy* 89 (1992): 223–261.

Gupta, A., and N. Belnap. *The Revision Theory of Truth.* Cambridge, MA: MIT Press, 1993.

Koons, R. C. *Paradoxes of Belief and Strategic Rationality.* Cambridge, U.K.: Cambridge University Press, 1992.

Martin, R. L., ed. *The Paradox of the Liar,* 2nd ed. Reseda, CA, 1978. Contains a useful bibliography of material up to about 1975; for later material consult the bibliography in Gupta and Belnap, 1993.

Martin, R. L., ed. *Recent Essays on Truth and the Liar Paradox.* New York: Oxford University Press, 1984. Contains the classic papers of Parsons, Kripke, Herzberger, and others; a good place to begin the study of the three approaches.

McGee, V. *Truth, Vagueness, and Paradox.* Indianapolis: Hackett, 1991.

Priest, G. *In Contradiction.* Dordrecht: Nijhoff, 1987.

Russell, B. "Mathematical Logic as Based on the Theory of Types." In *Logic and Knowledge.* London: Allen and Unwin, 1956.

Sainsbury, R. M. *Paradoxes.* Cambridge, U.K.: Cambridge University Press, 1988.

Simmons, K. *Universality and the Liar.* New York: Cambridge University Press, 1993.

Spade, P. V. *Lies, Language, and Logic in the Late Middle Ages.* London: Variorum, 1988.

Tarski, A. "The Semantic Conception of Truth." *Philosophy and Phenomenological Research* 4 (1944): 341–376.

Visser, A. "Semantics and the Liar Paradox." In *Handbook of Philosophical Logic,* edited by D. Gabbay and F. Guenthner, Vol. 4. Dordrecht: Reidel, 1989.

Yablo, S. "Hop, Skip, and Jump: The Agonistic Conception of Truth." In *Philosophical Perspectives,* edited by J. Tomberlin, Vol. 7. Atascadero, CA: Ridgeview, 1993.

Yaqūb, A. M. *The Liar Speaks the Truth.* New York: Oxford University Press, 1993.

Anil Gupta (1996)

LIBERALISM

By definition, a liberal is one who believes in liberty, but because different people at different times have meant different things by liberty, "liberalism" is correspondingly ambiguous. The word was first heard in a political sense in England in the early nineteenth century, when "liberals" were thus named by their Tory opponents. Indeed, they were first called *liberales,* and the Spanish form was used "with the intention of suggesting that the principles of those politicians were un-English" (see *Shorter Oxford English Dictionary*). This was ironical, since the word *liberal* had been adopted by the Spaniards for policies they regarded as essentially English—that is, the Lockean principles of constitutional monarchy, parliamentary government, and the rights of man. In any event, the Englishmen who were called liberals (though as late as 1816 Robert Southey was still calling them *liberales*) rejoiced in the name, and what was intended to be a pejorative quickly proved to have a distinctly pleasing flavor, perhaps partly because its other significance, the Shakespearean sense of liberal as "gross" or "licentious," had given way to the modern sense of liberal as "bountiful," "generous," or "open-hearted."

ENGLISH LIBERALISM

Traditional English liberalism has rested on a fairly simple concept of liberty—namely, that of freedom from the constraints of the state. In Thomas Hobbes's memorable phrase, "The liberties of subjects depend on the silence of the law." In general, however, English liberals have always been careful not to press this notion to anarchist extremes. They have regarded the state as a necessary institution, ensuring order and law at home, defense against foreign powers, and security of possessions—the three principles John Locke summarized as "life, liberty and property." English liberals have also maintained that the law can be used to extend the liberties of subjects insofar as the law is made to curb and limit the activities of the executive government. Thus, for example, the English laws of habeas corpus, of bail, and of police entry and arrest all constrain or restrain the executive and, in so doing, increase the freedom of the people. Some instruments of constitutional law have a similar effect.

The traditional form of English political liberalism naturally went hand in hand with the classical economic doctrine of laissez-faire. Toward the end of the nineteenth century, however, certain radical movements and certain English liberal theorists, such as Matthew Arnold and T. H. Green, developed, partly under foreign, left-wing influences, a different—as they claimed, a broader—concept of freedom, which was, to a large extent, to prove more popular in the twentieth century than traditional English liberalism with its economic gospel of laissez-faire. The central aim of this new school was utilitarian—namely, freeing men from misery and ignorance. Its exponents believed that the state must be the instrument by which this end was to be achieved. Hence, English liberal opinion entered the twentieth century in a highly paradoxical condition, urging, on the one hand, a freedom that was understood as freedom from the constraints of the state and, on the other, an enlargement of the state's power and control in order to liberate the poor from the oppressive burdens of poverty. In the political

sphere this contradiction in the liberal ideology ended in the disintegration of the British Liberal Party. With the defeat of Prime Minister Herbert Henry Asquith, a disciple of the philosopher T. H. Green and an adept at reconciling contradictions, the British Liberal Party broke into two, the right-wing, or laissez-faire, element joining forces with conservatism and the radical, *étatiste* element merging with socialism. Only a "rump" remained.

FRENCH LIBERALISM

The ambiguity of the word *liberalism* is more marked in French than in any other European language. Some writers hold that as a result of events in France since the time of Louis XIV, the French people have been divided into two political camps: One that supports the Roman Catholic Church, traditional social patterns, and the Syllabus of Pius IX (1864) and one that opposes the church and favors parliament, progress, and the rights of man. Historians who see France in these terms call one side *conservateur,* the other *libéral.* Opposed to this view are those historians who see not two, but at least three, continuing traditions in French political thought: on the right, royalism and conservativism; on the left, socialism, anarchism, syndicalism, and communism; in the center, liberalism. In the first of these two analyses, *libéralisme* is understood to embrace all the creeds of the left; according to the second analysis, *libéralisme* is a political doctrine at variance with the creeds of the left.

Again, one can distinguish two distinct—indeed, opposing—schools among French theorists who claim to be liberal. One is the Lockean liberalism of Voltaire, Baron de Montesquieu, and Benjamin Constant (in effect, also that of François Guizot and the July monarchy of Louis Philippe)—the liberalism of the minimal state, individualism, and laissez-faire. But there is a second liberalism, represented by the masters of the French Revolution and by the youthful Napoleon Bonaparte, which is democratic, Rousseauesque, and *étatiste.* Whereas Lockean liberalism understands freedom as being left alone by the state, the other liberalism sees freedom as ruling oneself through the medium of a state that one has made one's own.

Both these schools of *libéralisme* contributed something to the ideology of the French Revolution, and the often unperceived contradiction between them may also be said to have contributed to the intellectual confusion of those times. The fall of Napoleon was the signal for a return to the more purely Lockean style of liberalism. Benjamin Constant not only insisted that Jean-Jacques Rousseau's concept of liberty was an illusory one but also maintained that "*Du Contrat Social* [1762] so often invoked in favour of liberty, is the most formidable ally of all despotisms." Constant and his friends desired only to reproduce in France the Lockean Glorious Revolution of 1688. In 1830 they believed they had succeeded; Louis Philippe was enthroned on the basis of an understanding very like that on which William and Mary had been crowned in England. Politicians such as Guizot, who called themselves Libéraux, were put in charge of the kingdom. The result was not inspiring. A new bourgeoisie basked in the liberty the Lockean state introduced; the great were diminished, but the poor were not elevated. A rebellion came from the left in 1848, and the right replied with Napoleon III. Henceforth, there were few self-styled Libéraux of any importance in French politics and no liberal party. When new parties were formed later in the century, the name chosen by the center was Republicain rather than Libéral. This is not to say that liberalism died in France in 1848; rather, the word *libéralisme* thereafter ceased to call to the minds of French-speaking people any clear or distinct idea.

In 1912 Émile Faguet published a celebrated work, *Le libéralisme,* in which he took a rigidly Lockean position. "The state," he wrote, "is an evil; a lesser evil than anarchy, but nevertheless to be limited to the tasks of securing public order and safety through the justiciary, police and army." Several critics at the time attacked Faguet's definition as being outmoded; nevertheless, the definition of *libéralisme* in the 1935 edition of the *Dictionnaire de l'Académie Française* is, like Faguet's, thoroughly Lockean; it defines *libéralisme* in terms of the citizen's right to freedom of thought and to protection from government interference in private and business affairs.

One of several French theorists who attacked Faguet's exposition of liberalism (and, by implication, the academy's definition) was Jean de Grandvilliers. "How the word 'liberalism' is perverted by those who treat it as synonymous with individualism!" he wrote in *Essai sur le libéralisme allemand* (1925). "We can only reply by giving the word its true meaning." According to Grandvilliers, the true meaning of liberalism is to be found in a policy of extending the liberty of the people; he maintained that the intervention of the state is not only a useful, but also a necessary, means to achieve that end. Grandvilliers is thus a champion of the *étatiste* school of liberalism, which derived its concept of liberty from Rousseau and which argued that as long as the state belongs to the people, the enlargement of the power of the state is equally an enlargement of the power, and therefore the freedom, of its citizens.

GERMAN LIBERALISM

The word *liberal* was first heard in Germany in 1812, going there, as it went to England, from Spain. But the last years of Napoleon's power marked the decline of one tradition of German liberalism and the beginning of a new one. For in Germany, as elsewhere, we may discern not a single doctrine of liberalism but at least two main, conflicting schools, which again may be classified as the Lockean and the *étatiste.* The older German tradition was not merely derivatively Lockean; it also had contributed much to the formulation of Locke's own thought. In the sixteenth century it was a German philosopher, Johannes Althusius, who proclaimed that sovereignty derived from the people, and it was the German *Naturrechts* school of jurists that provided the bridge between the Stoic concept of *jus naturale* and the Lockean doctrine of the rights of man. But Locke, in turn, influenced the eighteenth-century German liberals, among whom Wilhelm von Humboldt was perhaps the most conspicuous. The very title of his book *Ideen zu einem Versuch, die Grenzen der Wirksamkeit des Staates zu bestimmen* (Ideas toward an investigation to determine the proper limits of the activity of the state; 1792), reveals his preoccupation with limited sovereignty and the minimal state. In this work Humboldt argued that the function of the state is not to do good but to ward off evil, notably the evil that springs from man's disregard for his neighbors' rights. The state, he said, "must not proceed a step further than is necessary for the mutual security of citizens and protection against foreign enemies; for no other object should it impose restrictions on freedom." Eighteenth-century Germany also had several liberal economists, including Christian Kraus, who considered that Adam Smith's *Wealth of Nations* (1776) was the most important book after the Bible.

In the nineteenth century a new school of liberalism, which was first and foremost nationalistic, arose in Germany. The freedom it stood for was the freedom of Germany, and the condition of the realization of this national freedom was the unification of Germany. Thus, whereas the old Lockean liberals were against the state, the new nationalist liberals wanted to create a greater state. The French declaration of 1789 proclaimed the rights of man; the German liberals inspired in 1848 a declaration of the rights of the German people. The new German liberals thought in terms of collective, rather than individual, rights. Thus, the *étatiste* German liberals saw nothing incongruous in sending a mission in 1849 from the Frankfurt parliament to Berlin to offer the crown of all Germany to a Prussian monarch, Friedrich Wilhelm, who detested democracy and who, in any event, grandly announced that he did not take crowns from commoners.

The difficulty of understanding in what sense this new German liberalism rested on a principle of freedom is that of understanding what it was that its votaries were demanding freedom from. Indeed, for many German liberals it was not a question of freedom from anything. German metaphysics of the same period was working out a concept of freedom that had nothing to do with resisting constraint. Guido de Ruggiero, a sympathetic Italian historian of German liberalism wrote:

> The eternal glory of Kant is to have demonstrated that obedience to the moral law is freedom.… It was the great merit of [G. W. F.] Hegel to have extracted from the Kantian identification of freedom with mind, the idea of an organic development of freedom, coinciding with the organisation of society in its progressively higher and more spiritual forms.… The State, the organ of coercion *par excellence,* has become the highest expression of liberty. (*History of European Liberalism*)

The idea that true freedom is to be found in obedience to the morally perfected state gave a theoretical justification (of a highly abstract kind) to the nineteenth-century German liberals' pursuit of liberty in submission to a strong and unified nation-state. But these high-thinking theorists never recovered from Friedrich Wilhelm's snub in 1849. Germany got its unity, but it was the imperialists, not the new liberals, who achieved it, and it was Otto von Bismarck, rather than Immanuel Kant, who gave the unified nation its political ethos. After the defeat of the Nazi regime in 1945, however, there was some revival of the Lockean type of liberalism in Germany.

AMERICAN LIBERALISM

In the United States the word *liberal* has never enjoyed the prestige it has in the United Kingdom, for in America there has never been, as there has in England, a national liberal party. The short-lived Liberal Republican Party of the 1870s was without a coherent program. Horace Greeley, its presidential candidate, was at once a socialist, spiritualist, vegetarian, and total abstainer; his personality led many Americans of his time to associate the word *liberal* with a visionary crank, and some still do. F. O. Matthiessen wrote in 1948: "In our nineteenth-century political life we had no such formulated division as that between the Conservatives and Liberals in England.… The key word seized upon by our native radical movement of the eighties and nineties, that of the Populists,

was not 'liberal' but 'progressive'" (*From the Heart of Europe,* New York, 1948, p. 90). Again, whereas in Vernon Louis Parrington's *Main Currents in American Thought* the word *liberal* occurs on almost every page, Parrington's pupil Henry Steele Commager never once uses the words *liberal* and *liberalism* in his continuation volume, *The American Mind* (New Haven, CT, 1950).

Just as in France the word *liberal* had been used by some writers for almost any kind of left-wing opinion, so in America the word *liberal* was widely adopted after the Great Depression as a soubriquet for "socialist." In *The Liberal Imagination,* Lionel Trilling defined liberalism as meaning, among other things, "a belief in planning and international co-operation, especially where Russia is in question." This definition may not have been wholly authorized by common usage, but there can be no doubt that the word *liberal* has come to be associated in the American public's mind with *étatiste* and left-wing ideologies rather than with the Lockean notions of laissez-faire and mistrust of organized power.

Indeed, it was one of Parrington's arguments in *Main Currents in American Thought* that American liberalism, as he called it, had always been concerned with democracy in a way that Locke and his English followers had not. Yet even before the emergence of twentieth-century left-wing liberalism, two rival creeds, both of which could reasonably be called liberal, contended for political supremacy. The first, as Parrington pointed out, was close to the "English philosophy of *laissez-faire,* based on the assured universality of the acquisitive instinct and postulating a social order answering the needs of the abstract 'economic man' in which the state should function in the interests of trade." The second liberalism was Rousseauesque rather than Lockean. It was "based on the conception of human perfectibility" and looked toward an egalitarian democracy "in which the political state should function as the servant to the common well-being."

The dominant political sentiment of the American tradition derives something from both these kinds of liberalism, for it has combined a Lockean attachment to liberty from the state with a Rousseauesque belief in democracy and equality. Nevertheless, perhaps it is still not quite respectable to be an avowed liberal in America. This may be partly because there has been no traditional support for a liberal party. It is also partly because not only socialists, but also communists and communist sympathizers, have not ceased to assume the title "liberal" rather than a more explicit expression of their political commitment.

A remarkable variety of political structures has been thought by different philosophers to embody liberty, and a correspondingly mixed company has shared the name "liberal." In singling out certain main streams or schools of liberal thought, one has to be mindful of the divergences that exist even among those which can be usefully grouped together. One might broadly divide philosophers of freedom into those who think that to be free is to be able to do what one wants to do and those who think that to be free is to do what one ought to do. By a similar method, one might divide liberals into those who see freedom as something that belongs to the individual, to be defended against the encroachments of the state, and those who see freedom as something which belongs to society and which the state, as the central instrument of social betterment, can be made to enlarge and improve. It remains to be said that some of the greatest names in the history of liberal thought, including John Stuart Mill himself, are strangely poised between these two positions.

See also Althusius, Johannes; Arnold, Matthew; Censorship; Green, Thomas Hill; Hobbes, Thomas; Humboldt, Wilhelm von; Kant, Immanuel; Libertarianism; Liberty; Locke, John; Mill, John Stuart; Montesquieu, Baron de; Rights; Rousseau, Jean-Jacques; Smith, Adam; Sovereignty; Voltaire, François-Marie Arouet de.

Bibliography

Adler, Mortimer J. *The Idea of Freedom.* New York: Doubleday, 1958.

"Alain." *Le citoyen contre les pouvoirs.* Paris: Éditions du Sagittaire, 1926.

Berlin, Isaiah. *Two Concepts of Liberty.* Oxford: Clarendon Press, 1958.

Cranston, Maurice. *Freedom.* London: Longmans, Green, 1953.

Duclos, Pierre. *L'évolution des rapports politiques.* Paris: Presses Universitaires de France, 1950.

Faguet, Émile. *Le libéralisme.* Paris, 1912.

Grandvilliers, Jean de. *Essai sur le libéralisme allemand.* Paris, 1925.

Halévy, Élie. *La formation du radicalisme philosophique.* Paris, 1935.

Hallowell, John H. *The Decline of Liberalism.* London: K. Paul, Trench, Trubner, 1946.

Hartz, Louis. *Liberal Tradition in America.* New York: Harcourt Brace, 1955.

Hayek, Friedrich A. von. *The Constitution of Liberty.* Chicago: University of Chicago Press, 1960.

Hobhouse, L. T. *Liberalism.* New York: Holt, 1911.

Konvitz, M. R., and C. L. Rossiter, eds. *Aspects of Liberty.* Ithaca, NY: Cornell University Press, 1958.

Laski, Harold J. *The Rise of European Liberalism.* London: Allen and Unwin, 1936.

Martin, B. Kingsley. *French Liberal Thought in the Eighteenth Century.* London: Benn, 1929.

Neill, T. P. *Rise and Decline of Liberalism.* Milwaukee: Bruce, 1953.

Parrington, Vernon Louis. *Main Currents in American Thought.* 3 vols. New York: Harcourt Brace, 1927–1930.

Ponteil, F. *L'éveil des nationalités.* Paris, 1960.

Popper, Karl R. *The Open Society and Its Enemies.* 2 vols. London: Routledge, 1945.

Ruggiero, Guido de. *Storia del liberalismo.* Bari, Italy, 1925. Translated by R. G. Collingwood as *History of European Liberalism.* London: Oxford University Press, 1927.

Sartori, Giovanni. *Democratic Theory.* Detroit, MI: Wayne State University Press, 1962.

Schapiro, J. Salwyn. *Liberalism.* Princeton, NJ, 1953.

Thomas, R. H. *Liberalism, Nationalism and the German Intellectuals.* Chester Springs, PA, 1953.

Trilling, Lionel. *The Liberal Imagination.* Garden City, NY: Doubleday, 1953.

Waldeck-Rousseau, P. M. R. *L'état et la liberté.* Paris, 1906.

Watson, G., ed. *The Unservile State.* London: Allen and Unwin, 1957.

Maurice Cranston (1967)

LIBERALISM [ADDENDUM]

The theory of liberalism and political philosophy in general were dramatically revitalized by the publication of John Rawls's *A Theory of Justice* (1971). In that work Rawls adopts the social contract model of political theory but with several key innovations. Rawls states that his theory seeks to capture the essence of the social contract theories of Locke, Rousseau, and Kant and develop the core idea of the contract to deal with traditional criticisms of the contract model of political legitimacy. The social contract model of justifying political authority has as its core idea that the basic principles of justice are the object of an original agreement that free and rational persons concerned to further their own interests would accept in an initial position of equality.

THE ORIGINAL POSITION

For Rawls, the first question for political theory is to specify what initial conditions are right for deciding the question of justice. His answer is what he calls the "original position," a hypothetical state of nature or situation without a government designed to be the conceptual context within which the basic principles of justice will be considered and the main outlines of the distribution of rights and duties will be defined and agreed upon. Rawls sees the original position as a heuristic device or a thought experiment used to rethink and clarify our intuitions about what justice is and what the basic structure of a just society would consist in. Rawls's concept of the original position has the following important components: the "veil of ignorance"; definition of the "people" in the original position; and, general knowledge that includes knowledge of the circumstances of justice and knowledge of the main competing theories of justice.

The overall design of the original position is based on what Rawls calls "considered judgments." These are judgments where moral capacities are likely to be manifested without distortion or prejudice. These judgments, for example, would include the beliefs that slavery, religious intolerance, and racism are wrong and ideas about fairness and human equality. Rawls terms his understanding of justification in ethics "reflective equilibrium." This model of justification rejects traditional foundationalist ideas of justification that hold that there are self-evident ethical principles from which one can derive specific moral rules and principles of justice and accepts a more coherentist model of justification. This coherence paradigm of justification holds that a theory is justified if one's considered judgments and moral and political principles cohere in a consistent belief system. By "equilibrium" Rawls means that one's judgments and principles are compatible and by "reflective" he means that one is fully and rationally aware of what our judgments and principles are and their derivation.

The veil of ignorance is a central feature of the original position. This imaginary veil is necessary, Rawls argues, because it excludes information that is not morally relevant or is a product of factors that are unjust and could be a source of prejudice. This means that information about one's social class, wealth, sex, race, abilities, personality, intelligence, particular conception of the good, health and the specific circumstances of one's society are excluded.

Though members of the original position are not allowed specific information about themselves, they are allowed certain general information. Members of the original position consider themselves free, equal, rational, and self-interested. As free, Rawls means no one is under the authority of another and as equal he means each has the same rights to make choices and decisions. As rational, Rawls means that people understand the ideas of justifying beliefs with evidence and that one should choose the most appropriate means to achieve one's goals. As self-interested, Rawls does not mean that people in the original position are selfish but rather that they are interested in their own welfare.

Members of the original position know that they need what Rawls calls "primary goods"; namely, certain basic rights, liberties, opportunities, income, wealth, and self-respect. These Rawls considers necessary means for whatever goals one may have and as such provide the motivation element in the deliberation in the original position.

Participants in the original position are also allowed knowledge of the circumstances of justice and the main competing theories of justice. The circumstances of justice include the notion that individuals coexist with roughly equal physical and mental powers, but are morally and intellectually limited with similar needs but different life goals in a world of moderate scarcity of resources.

The main competing theories of justice the members of the original position focus on are that of Rawls's theory, which he will call "justice as fairness," and utilitarianism. Rawls claims that people of the original position would reject utility as the principle of justice because, according to Rawls, the theory may allow injustice to a few to maximize utility overall.

Given his characterization of the hypothetical choice situation, Rawls believes members of the original position would agree to two principles of justice. Rawls's first principle states: "Each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others" (1971, p. 60). The second principle holds: "Social and economic inequalities are to be arranged so that they are both, a) reasonably expected to be to everyone's advantage, and b) attached to positions and offices open to all" (1971, p. 60). According to Rawls, members of the original position would also decide that the first principle has priority over the second and cannot be sacrificed to realize the second principle more fully.

The first part of the second principle Rawls calls the "difference principle," and it requires that all inequality in economic matters benefit all members of society, especially the least advantaged. The second part of the second principle requires what Rawls calls "fair equality of opportunity." Fair equality of opportunity requires not only that there are no legal obstacles for any position in society, but it would also provide for equal starting social conditions for all. Rawls believes that all people should have an equal chance to achieve any position in society regardless of what their family background, their social class, sex, religion, and ethnic background happen to be. Government would have to make sure that people have such equal opportunity by providing an equally good education and other services intended to prevent great social inequality in income, opportunity, and wealth.

Rawls admits that to implement his two principles may mean a large role for government, but he does not demand that either socialism or capitalism would necessarily be agreed upon in the original position. He considers this an empirical decision that social conditions and economic efficiencies would dictate.

Rawls's theory has been praised and critiqued. Many applauded its robust defense of welfare liberalism, concern for the poor, and the central importance of fair equality of opportunity. Others were pleased by the interdisciplinary nature of Rawls's work and a style accessible to the ordinary educated person.

CRITICISMS OF RAWLS'S THEORY

Critics of Rawls's theory come from the political right and left. Those on the right feel he overemphasizes equality and puts too much power in the hands of government. Libertarians such as Robert Nozick (1974) claimed he has reduced liberty too greatly at the expense of equality and allowed for the violation of the right to property by allowing increased taxation of the rich to help the poor.

Critics of Rawls from the left believe he has allowed too much inequality. Socialists believe that Rawls should have realized that capitalism allows too much power in the hands of the capitalists who would control government to promote their interests. Marxists also claimed Rawls's theory of human nature is biased in favor of human nature as it exists in an alienated form under capitalism based on competitive individualism and overlooking the great power of social class in limiting freedom. Other critics such as James Sterba (2004) have questioned certain specific elements of the theory. Sterba claims the difference principle would in fact not be chosen in the original position. Sterba argues that members of the original position would choose a guaranteed minimum rather than the difference principle, but Sterba then extends that minimum to distant peoples and future generations which, he believes, will have the effect of greater equality.

Still other critics such as Michael Sandel (1982) believe that liberal philosophers such as Rawls place too much emphasis on individual rights and not enough on the role of the community and individual responsibility.

RAWLS'S LATER WORK

Rawls's later work, *Political Liberalism* (1993), still defends his principles of justice but also attempts to

address some of the criticisms of his earlier work. The goals of *A Theory of Justice* (1971) were, according to Rawls, to develop justice as fairness as a superior moral and political theory to utilitarianism and use the social contract model to do so. The problem with these goals, Rawls explains in *Political Liberalism*, is that he was endorsing a comprehensive doctrine similar to Kantianism that is problematic in a world of incompatible doctrines none of which can be rationally determined to be correct. A "comprehensive doctrine" is defined by Rawls as a doctrine that encompasses all central values and beliefs about life. In this sense the main world religions and philosophical systems such as utilitarianism are comprehensive doctrines. By contrast, a "political conception" is a set of ideas that applies only to the political realm and does not assume any comprehensive doctrine but rather uses ideas found in the political culture of a society.

As Rawls puts it, the problem for his theory and political philosophy in general is: "How is it possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical and moral doctrines?" (1993, p. xviii). Rawls hopes to solve this problem of reasonable pluralism of comprehensive doctrines by establishing the following: (1) to distinguish more clearly the difference between a comprehensive doctrine from a political one; (2) to emphasize the importance of stability in a well-ordered society in a world of reasonable pluralism of comprehensive doctrines; (3) to clarify that justice as fairness is not a comprehensive but a political doctrine; and (4) to show that political liberalism assumes and is compatible with a pluralism of reasonable comprehensive doctrines.

The idea of a well-ordered society is central to Rawls's answer to the problem of pluralism. For Rawls, well-ordered society is a stable society that, when realized, generates its own support from the citizenry being accepted as a fair system of cooperation based on publicly recognized rules agreed by all as just. A well-ordered society must also be one where the conception agreed to is limited to the political because of three facts. First, there is what Rawls calls the diversity of reasonable comprehensive religious, philosophical, and moral doctrines found in modern societies. The second fact is that to maintain one comprehensive doctrine as the correct one would entail the use of coercive physical state power. Third, a secure democratic government must be freely supported by at least a majority of its citizens.

Rawls believes a well-ordered society is possible because there is a limited agreement about political justice in the political culture of democratic societies. This agreement he calls an "overlapping consensus." This consensus is not a mere "modus vivendi" according to Rawls; that is, it is not merely the result of negotiation of self-interested parties, but rather it is agreed to on moral grounds found in the differing comprehensive doctrines. As such, Rawls calls his political theory of liberalism "freestanding" in that it is not based on any comprehensive doctrine.

On the one hand, many philosophers praised *Political Liberalism* as a major work dealing with the postmodern world of pluralism and ideological diversity. Critics, on the other hand, claimed that Rawls has merely assumed an overlapping consensus among comprehensive theories when in fact there is no such consensus. Others prefer his earlier work because they feel there is a need for some foundations to justify the theory that seems to be lacking in the new presentation.

Discussion of Rawls's work continues, but there is a growing consensus that his contributions to the field of political theory of liberalism will stand as a major addition to the canon of political philosophy for a long time to come.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism.

Bibliography

Barry, B. *The Liberal Theory of Justice: A Critical Examination of the Principal Doctrines in a Theory of Justice*. Oxford: Clarendon, 1973.

Dworkin, R. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press, 1977.

Freeman, S. *The Cambridge Companion to Rawls*. Cambridge, U.K: Cambridge University Press, 2003.

Freeman, S., ed. *Rawls: Collected Papers*. Cambridge, MA., 1999.

Gould, C. *Rethinking Democracy*. Cambridge, 1988.

Grcic, J. *Ethics and Political Theory*. Lanham, MD: University Press of America, 2000.

Martin, R. *Rawls and Rights*. Lawrence, KS: University Press of Kansas, 1985.

Nozick, R. *Anarchy, State and Utopia*. New York, 1974.

Rawls, John. *Justice as Fairness: A Restatement*, edited by E. Kelly. Cambridge, MA: Belknap Press of Harvard University Press, 2001.

Rawls, John. *The Law of Peoples: With, "The Idea of Public Reason Revisited."* Cambridge, MA: Harvard University Press, 1999.

Rawls, John. *Lectures on the History of Moral Philosophy*, edited by B. Herman. Cambridge, MA: Harvard University Press, 2000.

Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993. Rev. paperback ed., 1996.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press, 1971. Rev. ed., 1999.

Sandel, M. *Liberalism and the Limits of Justice*. Cambridge, U.K.: Cambridge University Press, 1982.

Sterba, J. *Triumph of Practice over Theory in Ethics*. New York: Oxford University Press, 2004.

Joseph Grcic (2005)

LIBERATION IN INDIAN PHILOSOPHY

The concept of liberation presupposes someone's state of bondage and anticipates the possibility of his or her release into a state of freedom. From the philosophical perspective bondage marks the human predicament of leading a precarious existence in an unstable world. In Indian philosophy the state of bondage is termed *saṁsāra* (global flow) and understood as a beginningless process of life of beings who are born, die, and are constantly reborn. This process is governed by the eternal law called in mainstream Hinduism *sanātana dharma*. This expression is *multivalent*, having several layers of meaning; Indian thinkers regard it as a matrix encompassing reality in its totality. In Buddhism *dharma* occurs without the attribute "everlasting," but is understood as being beyond time.

The multivalency of *sanātana dharma* gives it at least three meanings. First, as the eternal law it represents an impersonal force inherent in everything so that reality is orderly rather than chaotic; processes of reality follow the law of cause and effect. Second, the aspect of timelessness of *dharma* implies the view that even the phenomenal reality has no conceivable beginning and end, but keeps renewing itself in cycles. In other words, the global world process—including the present universe—has no fixed origin, such as a creative act of God, and will never come to an end to be replaced by the eternal "new earth and new heaven" after a day of judgment. Rather, it undergoes periodic renewals: At the beginning of each period the world process starts with the emergence (*sṛṣṭi*) of the universe from its hidden dimension into the state of manifestation; in the course of its duration (*sthiti*) it evolves to a peak, followed by decline and end in universal dissolution into the unmanifest state (*pralaya*) called cosmic night. After a period of latency, the whole process starts again.

The lives of individual beings proceed within this global framework from birth to adulthood, old age, death and rebirth in a never-ending round of *saṁsāric* existences. During the cosmic night they subsist in a kind of limbo or oblivion. Third, the concept of *dharma* also refers to the timeless and absolute reality beyond the manifested one; it represents the final goal of religious and philosophical quest equated with the ultimate truth. This truth is eternal, outside time, and independent of the changeable phases of the phenomenal reality manifested within time. The manifestation of the eternal truth or law within the universe dominated by time does not make the world everlasting in the sense of a lineal duration, but provides for its cyclic nature, its recurring rise and fall.

The concept of *dharma* understood as the absolute truth and ultimate reality has still another connotation—that of consciousness, awareness, or intelligence. Truth makes sense only if it is known. Indian philosophy, unlike Western science, has never conceived of reality without consciousness. Thus, a verse in one of the earliest Indian philosophical texts (1500–1000 BCE), the creation hymn of the *Ṛg Veda* (10, 129, 4), describes the primordial oneness (*tad ekam*) as experiencing desire (*kāma*), the earliest seed of its mind (*manaso retaḥ*), which led to manifestation. The dimension of consciousness as an inherent quality of reality in its ultimate state evokes two fundamental insights. First, the idea of the ultimate personality (*puruṣottama*), albeit an infinite one, conceived as the personality of God, the free agent behind the world process, although not an omnipotent one. Second, it suggests that the individual human consciousness, being an instance of the universal dimension of consciousness, has—despite its present limitations—the potential of grasping reality on the ultimate level: Man has the capability to develop an understanding and vision of the absolute truth. Extricating himself thus from the conditionality of his phenomenal existence and attaining final liberation (*mokṣa*, *mukti*), he enters the timeless dimension of the absolute without having to participate in the world process and undergo repeated incarnations.

While in bondage, he is governed by *sanātana dharma* in all its aspects. Its aspect of causality operates in human life on a higher level as the law of karma, which is much more complicated than the law of cause and effect in the material universe, yet it can be expressed in the simple saying "as you have sown, so you will reap." Every volitional act in thought, speech, or deed generates a force that produces sooner or later—in one's present life or some future existence—results that shape one's external circumstances and appearance, forming one's character and determining one's fortune. The aspect of timelessness of *dharma* makes the lives of individual beings in the sequence of reincarnations appear to be without a con-

ceivable beginning and end. However, the aspect of *dharma* as the timeless and absolute reality beyond the manifested world lends individual beings an affinity with the ultimate truth and the potentiality of realizing it by direct conscious experience, which brings about the termination of their bondage and the attainment of liberation outside time.

This necessitates entering a spiritual path, a training to deepen one's perception of reality up to the point of the final vision. Volitional input is essential for this purpose—as it is also within the karmic process to sow only wholesome deeds to earn future good results. The spiritual path was eventually systematized and became known as yoga.

The previous outline is valid in principle for all schools of Indian philosophy, including the earlier phases of Indian thought before the formation of philosophical systems. Despite the difference in terminology and sophistication of language, the ideas occur even in the oldest strata of Vedic scriptures in mythological guise, although nineteenth-century pioneers of Vedic scholarship failed to recognize them.

THE VEDAS AND UPANIṢADS

The *Ṛg Veda* uses the verb *muc* (hence *mokṣa* and *mukti*) in the creation myth when the god Indra periodically liberates the cosmic waters (= creative forces) from the clutches of the demon Vṛtra (10, 104, 9; 1, 32, 11; 4, 22, 7), thereby enabling the manifestation of the universe. As to humans, they are subjected to successive lives (*anūcīnā jīvitā*, 4, 54, 2), so liberation for them means being granted immortality (*amṛta*, *amṛtatva*). It is therefore ardently prayed for: "Lead us to immortality!" (5, 55, 4) "May I be released from death, not reft of immortality!" (7, 59, 12) "Place me in that deathless, undecaying world … make me immortal" (9, 113, 7–11). Certain "long-haired ascetics" (*keśins*) even claimed to have won immortality during their lifetime: "Due to our sagehood we have mounted upon the winds, only our bodies do you mortals see" (10, 136, 3). The pleas for immortality show that everlasting life was not automatically granted even if one reached heaven as a result of good deeds (10, 14, 8) and religious fervor (*tapas*, 10, 54, 2). Repeated death (*punarmṛtyu*) lurked even there as is later asserted by Śatapatha Brāhmaṇa (10, 4, 3, 10), so the search for immortality continues.

The ideas of rebirth under cosmic law and liberation from it are subsequently clearly spelled out in the oldest Upaniṣads (700–600 BCE): "One becomes pure by pure actions, bad by bad ones" (Bṛhadāraṇyaka Upaniṣad 3, 2, 13), and when one dies, knowledge (*vidyā*), deeds (*karmāṇi*), and previous experience (*pūrva prajñā*) follow one (4, 4, 2). One may live in higher worlds while the merits of one's actions last, but eventually returns to this world (4, 4, 4–6). But one has affinity with the Ultimate; one's inner self (*ātman*) is, at bottom, identical with the core of reality (*brahman*, 4, 4, 5). When one realizes it and can proclaim "I am *brahman*," one becomes the self of everything, including gods (1, 4, 10), and is freed from reincarnation. Thus, liberation is the result of the direct knowledge of one's inmost self and thereby of the inner essence of everything else brought about by meditational effort (*dhyāna*) and by renouncing external desires. Later Upaniṣads started developing methods of acquiring the liberating knowledge, thus foreshadowing the classical system of Yoga.

Two schools of thought and practice outside the Vedic tradition, Jainism and Buddhism, also systematized the path. Both emerged from the circles of wanderers (*śramaṇas*) striving for liberation from the round of rebirths by asceticism. In contrast to the Brahmanic tradition, they regarded the state of liberation as beyond description and used the negative term *nirvāṇa* (blowing-out) for it.

JAINISM

The term used in the teachings of Jina Mahāvīra (599–467 BCE) for individual beings is *jīva* (animate substance, soul, spirit-monad) or *ātman*. In its pure form a *jīva* is perfect, omniscient, eternal, and formless and enjoys unlimited energy and infinite bliss. When he succumbs to the influx (*āśrava*) of passions (*kaṣāya*) from the phenomenal world of modalities (*saṁsāra*), the *jīva* takes shape, assuming a body born from his actions (*karmaṇa-śarīra*), and he loses his perfection and becomes a mundane pilgrim (*saṁsāri*) through innumerable forms of life whose quality is determined by the ethical quality of his actions. Good actions secure his temporary well-being in *saṁsāra*, but do not lead to liberation. Of bad actions injury to life is the most detrimental one. Liberation (*mokṣa*, *mukti*, *nirvṛti*) is achieved by purging off (*nirjarā*) of karmic burdens accumulated by past actions and stopping (*saṁvara*) further influxes by renunciation so that the soul rises above involvement in any actions. In the last stages of ascetic practice (*tapas*), the abstention from action may involve stopping even intake of food and drink; liberation is reached on the point of death by starvation. If the *saṁsāri* achieves liberation before death, he becomes a perfect one (*siddha*) or a *tīrthaṅkara* (fordmaker, the teacher of others). Discarnate *siddhas* in *nir-*

vāṇa enjoy four infinite accomplishments: knowledge, vision, strength, and bliss. The Jain elaborate path to liberation shows overlaps with the Buddhist one and with Patañjali's Yoga.

BUDDHISM

Early Buddhist sources largely abstain from conceptual descriptions of the nature of beings, liberation, and ultimate reality. The Buddha (563–483 BCE) of the Pāli Canon maintained noble silence about such issues and focused pragmatically on analysis of the existential situation of man as it is accessible to everybody's experience and on practical procedures for gaining liberation and direct knowledge of true reality; called awakening or enlightenment (*bodhi*), this achievement does not include omniscience as in Jainism. Man's experience of himself is described in terms of five constituent groups of clinging (*upādānakkhandhas*):

(1) Bodily awareness or the experience of having a form (*rūpa*)

(2) Feelings (*vedanā*) that are pleasant, unpleasant, or neutral

(3) Perception (*saññā*) experienced through six channels—the five senses and the mind, the latter having the function of coordinating the fivefold sensory data into conceptually grasped objects

(4) Inner volitional dynamism described as the group of mental coefficients (*saṅkhāras*), such as instincts, urges, desires, wishes, decisions, and aspirations

(5) Consciousness (*viññāṇa*) or the direct awareness of being conscious of visual and other sensory objects and of mental images and concepts

None of these constituents represents the inner core, substance, or soul (*atta/ātman*) of the personality—they are *anatta*—and no such core is either postulated or denied. The structural unity of the personality is expressed by the term *nāmarūpa* (name and form), occasionally also *puggala* (Sanskrit: *pudgala*) or *purisa* (Sanskrit: *puruṣa*); its constituents constantly change, yet its individuality is preserved by its continuity as a process: The Buddha frequently referred to his and others' past lives.

Bondage to the round of births and deaths governed by the laws of karma results from ignorance (*avidyā*, *moha*) of the true nature of reality (*dhamma*). Beings are then subject to craving (*taṇhā*, *lobha*) directed to fleeting and basically substanceless pleasurable experiences and develop hate (*dosa*) if somebody obstructs their aims. The beginning of the individuals' *saṁsāric* sojourn cannot be found, but liberation is possible when beings realize its unsatisfactoriness, recognize desire as its cause, understand that renouncing desire will free them from rebirth, and embark on the path toward that final goal.

This is the gist of the Buddha's "four noble truths," the fourth one being the eightfold path of systematic training, the first comprehensive formulation of a liberating technique. On reaching liberation a Buddha's disciple becomes an *arahat* (worthy one) and is equal to the Buddha in the acquired state of freedom, while the Buddha surpasses him in wisdom, thus enabling him to be the "teacher of gods and men." Individuals who attain liberation on their own without the guidance of a buddha become solitary enlightened ones (*paccekabuddhas*), who do not assume a teaching mission. Early Buddhism does not admit descriptions of or speculation about the state of a liberated one (*tathāgata*) after death. Here, too, the Buddha maintains "noble silence," expressly denying only the validity of the four alternatives put to him by questioners, namely that he "is," "is not," "both is and is not," and "neither is nor is not." "The final truth (*dhamma*) is deep, unfathomable, understood only by the wise" (Majjhima Nikāya 72)—an Enlightened One.

Despite this injunction, speculation did not cease and some Hīnayāna schools of thought, including Theravāda, interpreted the Buddha's description of personality factors (*khandhas*) as unsubstantial (*anatta*) to mean denial of an inner core or any other feature that would lend individuals identity in successive lives and continuity into *nirvāṇa*. This was challenged by the Pudgalavāda school, which maintained that personality (*pudgala*) as such is as eaqually undefinable as *tathāgata* and that it is independent of the individual's status, whether bound or liberated, which means that it persists throughout successive lives and into *nirvāṇa*. This doctrine was adopted by many sects and remained influential for centuries.

Mahāyāna schools of thought do not appear to have had problems with personal continuity. Innumerable *tathāgatas* are active from within their spheres of influence (*buddhakṣetras*), helping beings to liberation, assisted by *bodhisattvas*, individuals developing ten perfections (*pāramitās*) on the path to buddhahood that proceeds through ten stages (*bhūmis*). Some *bodhisattvas* vow not to enter final *nirvāṇa* until all beings are liberated "down to the last blade of grass," an innovation that envisages universal liberation. This is viewed as possible on the basis of the philosophy of emptiness (*śūnyavāda*), which developed as a result of meditational experience: The mind, emptied of all contents derived from sensory

perception and conceptual activity, can make the final breakthrough into *nirvāṇa*, which is equally empty because it is inaccessible to sensory perception and undefinable. Thus, emptiness (*śūnyatā*) came to be regarded as underlying both *saṁsāra* and *nirvāṇa*, making them, at bottom, identical. Liberation occurs by shifting one's perspective.

Such tendencies to hypostatize *śūnyatā* were checked by Nāgārjuna (flourished c. 150–250), the protagonist of the Mādhyamaka school, who used the dialectical method to refute conflicting theses; truth lay in the middle, but beyond dialectics. It is accessible only to direct vision—as the Buddha taught. Tendencies to hypostatization appeared also in the Vijñānavāda school, which regards pure consciousness as the basis for not only *saṁsāra* but also *nirvāṇa*, since its achievement cannot but be a conscious experience. *Saṁsāric* phenomena are mental constructs projected from the universal storehouse consciousness (*ālaya-vijñāna*), yet the emptiness and purity of the root consciousness (*mūla-vijñāna*) and of a liberated one's consciousness remain unaffected.

HINDU SYSTEMS OF PHILOSOPHY

During the golden age of Indian civilization under the Gupta dynasty (320–510), philosophical discussions flourished between various schools of thought. Six of them came to be recognized as valid Hindu angles of viewing (*dṛṣṭi*, hence *darśaṇa*) of reality and were systematized. All accept the basic teaching about *saṁsāric* bondage and the desirability of liberation, but differ in ontological conceptions and methodical approaches.

(1) Pūrva-Mīmāṁsā (original elucidation) regards the Vedas as eternal and pursues the path of ritual action (*karma-mārga*), which parallels cosmic processes and terrestrial events governed by the inherent law of *ṛta* (the Vedic equivalent of *dharma*), which is independent of any divine agency. Right rituals achieve anything, including rebirth in the highest existential spheres and liberation, although in advanced stages of the path ritual is interiorized and becomes a process of meditation.

(2) Vaiśeṣika (discrimination) is a kind of natural philosophy focusing on classifying reality into categories (*padārthas*). Reality is subjected to the invisible law (*adṛṣṭa dharma*) operating also in the ethical sphere independently of God (*īśvara*), an eternally free, omniscient spirit (not a creator) who can assist beings on the path of knowledge (*jñána-mārga*) based on a meditational analysis of *saṁsāric* categories that leads to liberation from them.

(3) Nyāya (guidance) analyses logical and epistemological processes that supply beings with their picture of the world. In testing its validity, Nyāya thinkers discovered syllogism that, however, required verification by experience. Logical analysis is the start of the path of knowledge (*jñāna-mārga*). It sharpens the mind, preparing it for meditational viewing, which culminates in direct knowledge of the final truth equaling liberation.

(4) Sāṅkhya (enumeration) is a dualistic metaphysical system with no God. It recognizes an infinite number of originally pure and free eternal spirits (*puruṣas*) and the creative force of nature (*prakṛti*), which conjures up the world process for *puruṣas*. As they show interest in this spectacle, *prakṛti* creates for them bodies with senses and mental functions. The *puruṣas*, fascinated by the antics of *prakṛti*, identify with their *prakṛtic* personalities and forget their true status. When a *puruṣa* recognizes this bondage, he can liberate himself by mentally discriminating between *prakṛtic* evolutes and his original pure consciousness; this is a variety of *jñāna-mārga*. His worldly personality dissolves and he regains total freedom in isolation (*kaivalya*) from *prakṛti*.

(5) Yoga (union) as one of the six *darśaṇas* is chiefly a systematic eightfold path to liberation called classical Yoga, expounded by Patañjali (second century BCE). However, chapter 4 of his *Yoga Sūtras* shows that it had been a philosophical system in its own right before its ontology was overshadowed by Sāṅkhya. Still, it retained the notion of God (*īśvara*), an eternally free *puruṣa* who may assist other *puruṣas* (entangled in *saṁsāra*) struggling for liberation but is neither the Creator nor the focus of a religious cult. The discipline of the Yoga path aims at experiencing liberation as autonomy (*kaivalya*) from limiting forms of existence, accompanied by the final vision of or cognitive unification with the totality of truth (*dharmamegha-samādhi*).

(6) Uttara Mīmāṁsā (higher elucidation) or Vedānta (end of Veda, meaning Upaniṣads, its base) split into three subschools. In the Advaita (nondualistic) Vedānta of Śankara (700?–750?) *brahman*, the Upaniṣadic source and core of the manifested universe, is regarded as the sole reality; the individual bondage in *saṁsāra* is an illusion (*māyā*). Liberation is achieved when this illusion is dispersed by treading

the path of knowledge (*jñāna-yoga*) that culminates in *samādhi* experienced as the unity of being, consciousness, and bliss (*sat-cit-ānanda*). The liberated one realizes that he is and has always been *brahman* and that nothing else really exists. The Viṣiṣṭa Advaita (qualified nondualistic) Vedānta of Rāmānuja (c. 1077–1137) interprets the Upaniṣadic *brahman* as the eternal God who created the world out of his own subtle body by transforming it into a gross one. Beings are attributes of God, but possess their own self-conscious existence. They retain it even when liberated in mystic union with God accomplished with his grace (*prasāda*) after surrendering to him on the path of devotion (*bhakti-mārga*). Upaniṣadic passages with traces of a dualistic worldview (foreshadowing Sāṅkhya) enabled even the Dvaita (dualistic) Vedānta of Madhva (c. 1199–c. 1278) to claim Vedic authority for its interpretation. It accepts the eternal existence of *prakṛti* and the plurality of *jīvas*, who retain their individuality even in the state of liberation granted as God's grace to those who live pure lives and embrace *bhakti-mārga*. Others may transmigrate in *saṁsāra* forever. Some evildoers may even reach a point past redemption and face eternal damnation in infinite remoteness from God.

A modern approach to liberation appears in the writings of Aurobindo (1872–1950). He envisioned a new phase in the world's evolution: if enough individuals prepare themselves through yoga for receiving the cosmic consciousness, then they could bring about the spiritualization of the earth or even the whole universe. This idea of universal liberation has its origin in the vow of Mahāyāna *bodhisattvas* to liberate all beings "down to the last blade of grass."

See also Brahman; Causation in Indian Philosophy; God/Isvara in Indian Philosophy; Karma; Knowledge in Indian Philosophy; Meditation in Indian Philosophy; Mind and Mental States in Buddhist Philosophy; Negation in Indian Philosophy; Self in Indian Philosophy.

Bibliography

Blackstone, Kathryn R. *Women in the Footsteps of the Buddha: Struggle for Liberation in the Therigatha*. London: Curzon Press, 1998.

Buswell, Robert E., and Robert M. Gimello. *Paths to Liberation: The Mārga and Its Transformations in Buddhist Thought*. Delhi, India: Banarsidass, 1994.

Chakraborti, Haripada. *Asceticism in Ancient India*. Calcutta, India: Punthi Pustak, 1973.

Châu, Thích Thiên. *Les sectes personnalistes (Pudgalavādin) du bouddhisme ancien*. Paris: Université de la Sorbonne Nouvelle, 1977.

Dange, Sadashiv A. "Metempsychosis and the Ṛgveda." *Journal of the University of Bombay* 43 (79) (1974): 1–12.

Eckel, Malcolm David. *To See the Buddha: A Philosopher's Quest for the Meaning of Emptiness*. Princeton, NJ: Princeton University Press, 1992.

Fort, Andrew O. "Going on Knowledge? The Development of the Idea of Living Liberation in the Upanishads." *Indian Philosophy* 22 (4) (1944).

Fort, Andrew O. *Jīvanmukti in Transformation: Embodied Liberation in Advaita and Neo-Vedanta*. Albany: SUNY Press, 1998.

Gethin, Rupert. *The Buddhist Path of Awakening: A Study of the Bodhi-Pakkiyā Dhammā*. Leiden, Netherlands: Brill, 1992.

Gonda, Jan. *The Vision of the Vedic Poets*. The Hague, Netherlands: Mouton, 1963.

Harvey, Peter. "The Nature of the Tathāgata." In *Buddhist Studies, Ancient and Modern*, edited by P. Denwood and A. Piatigorsky. London: Curzon Press, 1983.

Jacobsen, Knut A. *Prakṛti in Sāṃkhya-Yoga: Material Principle, Religious Experience, Ethical Implications*. Delhi, India: Banarsidass, 2002.

Jaini, Padmanabha Shrivarma. *The Jaina Path of Purification*. Delhi, India: Banarsidass, 1979.

Johansson, R. *The Psychology of Nirvana*. London: Allen and Unwin, 1969.

Klostermaier, Klaus K. *Mythologies and Philosophies of the Salvation in the Theistic Traditions of India*. Waterloo, Canada: Wilfrid Laurier University Press, 1984.

Lad, Ashok Kumar. *A Comparative Study of the Concept of Liberation in Indian Philosophy*. Burhanpur, India: Girdharlal Keshavdas, 1967.

Larson, Gerald James, and Ram Shankar Bhattacharya. "Sānkhya: A Dualist Tradition in Indian Philosophy." In *Encyclopedia of Indian Philosophies*. Vol. 4. Princeton, NJ: Princeton University Press, 1987.

Miller, Jeanine. *The Vedas*. London: Rider, 1974.

Miller, Jeanine. *The Vision of Cosmic Order in the Vedas*. London: Routledge and Kegan Paul, 1985.

Oberhammer, Gerhard. *La Délivrance, dès Cette Vie (jīvanmukti)*. Paris: Édition-Diffusion de Boccard, 1994.

Oldmeadow, Harry. "Delivering the Last Blade of Grass: Aspects of the *Bodhisattva* Ideal in the Mahāyāna." *Asian Philosophy* 7 (3) (1997): 181–194.

Orofino, Giacomella, trans. *Sacred Tibetan Teachings on Death and Liberation*. Bridport, U.K.: Prism Press, 1999.

Verpoorten, Jean-Marie. *Mīmāṁsā Literature*. Wiesbaden, Germany: O. Harrassowitz, 1987.

Welbon, G. R. *The Buddhist Nirvāṇa and Its Western Interpreters*. Chicago: University of Chicago Press, 1968.

Werner, Karel. "*Bodhi* and *Arahattaphala*: From Early Buddhism to Early Mahāyāna." *Journal of the International Association of Buddhist Studies* 4 (1981): 70–84.

Werner, Karel. "Indian Conceptions of Human Personality." *Asian Philosophy* 6 (2) (1996): 93–107.

Werner, Karel. "Indian Concepts of Human Personality in Relation to the Doctrine of the Soul." *Journal of the Royal Asiatic Society* 1988 (1): 73–97.

Werner, Karel. "The Longhaired Sage of Ṛg Veda 10, 136: A Shaman, a Mystic, or a Yogi?" In *The Yogi and the Mystic: Studies in Indian and Comparative Mysticism*, edited by Karel Werner, 33–53. London: Curzon Press, 1989, repr. in 1994.

Werner, Karel. "A Note on *Karma* and Rebirth in the Vedas." *Hinduism* 83 (1978): 1–4.

Werner, Karel. "Symbolism in the Vedas and Its Conceptualisation." In *Symbols in Art and Religion: The Indian and the Comparative Perspectives*, edited by Karel Werner, 27–45. London: Curzon Press, 1990.

Werner, Karel. "The Vedic Concept of Human Personality and its Destiny." *Journal of Indian Philosophy* 5 (1978): 275–289.

Werner, Karel. *Yoga and Indian Philosophy*. Delhi, India: Banarsidass, 1977, reprint in 1980 and 1998.

Whicher, Ian. *The Integrity of the Yoga Darśana*. New York: SUNY Press, 1998.

Whicher, Ian, ed. *Yoga: Tradition and Transformation*. Richmond, Australia: Curzon, 2001.

Karel Werner (2005)

LIBERATION THEOLOGY

Liberation theology is the name of a movement that arose in the churches, both Catholic and Protestant, of Latin America during the last third of the twentieth century. It also describes a theological trend that is found, often under different names and with somewhat different emphases across the world, as black theology in the United States and South Africa, as Dalit theology in India, as Minjung theology in Korea, and elsewhere in other forms.

THEOLOGY

The earliest and still definitive statement of the movement is *A Theology of Liberation: History Politics, and Salvation* (1988) by Gustavo Gutiérrez. The basic principles it sets forth are:

> (1) Theology is critical reflection on Christian praxis. Faith, charity, and commitment to God and to others in the struggle for humanity and justice are primary. Theology relates this praxis to the sources of revelation and the history of the church.
>
> (2) Biblical revelation commits the church to God's "preferential option for the poor." The poor are, by their condition, involved in a struggle to realize their humanity and to become "subjects of their own history," against the political, economic, and social powers that marginalize and oppress them. This struggle is revolutionary, not reformist. The church belongs with the poor in the midst of it, doing theology in a revolutionary situation.
>
> (3) The struggle of the poor for social justice is a work of human self-creation that finds its source, meaning, and hope in God's work. Salvation history is at the heart of human history, in creation, covenant, Christ's incarnation, and the coming kingdom of God. Political liberation is a partial salvific event, a historical realization of the kingdom, that looks forward to its ultimate fulfillment by divine grace operating in the human struggle, informing its character and directing it toward ever larger goals of human community.

This is still its basic structure. In its development and spread, however, three major issues have arisen.

CRITIQUE: DEFINING THE POOR

First, how are the poor defined? The Latin American theologians clearly have a dependent economic class in mind, created by exploiting landlords, industrialists, and bankers, along with their political and military agents. This definition, in terms of the dehumanizing dynamics of the capitalist system and class struggle against it, clearly borrows from Karl Marx. José Miguez Bonino (1976) acknowledges this explicitly as do many others. The Vatican, though affirming a preferential option for the poor, has been severely critical of this tendency to identify the poor of scripture with the proletariat that Marx defined. Liberation theologians claim, however, that this analysis is the secular expression in modern industrial society of a theme in Christian history that finds its source in the Hebrew prophets and the incarnation of Christ: the saving work of God liberating the people from the economic and political power of organized human sin. *The Kairos Document, Challenge to the Church: A Theological Comment on the Political Crisis in South Africa*—(1986), without appealing to Marx, makes the same argument concerning the apartheid system, calling it prophetic theology, as opposed to (a) state theology, which justifies the status quo, and (b) church theology, which is cautiously critical but without social analysis or a strategy for revolutionary change. Minjung theology in Korea focuses on a politically oppressed people (minjung), given hope by biblical history and promise, to strive for their liberation in a messianic kingdom where Jesus the suffering servant is lord. For Dalit theology in India, like American black theology, it is a subjugated minority, the outcastes (the dalits), to which the promise of God comes, in their conflict with an oppressive majority. Black theology draws especially on the Exodus of the

Israelites from Egypt to legitimate black people's fight for freedom.

All these movements agree that liberation is the basic theme of the Christian message. All see political, economic, cultural, and even religious powers as the instruments of oppression against which they struggle in God's name. They differ in their perception of how the poor are defined and which powers are their primary antagonists. The power analysis that Marxism provides is determinative for some and secondary for others. All of them, however, incorporate it into a more subtle and insightful guide that scripture provides to Christian understanding of the poor and to action that will realize God's promise.

CRITIQUE: THE QUESTION OF TRUTH

Second, how is the truth claim of liberation theology validated? This question arises on two levels. First, the hermeneutic of suspicion, which probes the roots of all truth claims in social experience and defines theology as a reflection on social praxis, owes much to Marx. It contradicts the teaching of St. Thomas Aquinas about the universality of reason and natural law as perfected, not destroyed by revelation. It reflects, however, the reformation understanding of reason distorted by human sin and is rooted, liberation theologians would claim, in the way God is known in the biblical history of calling, covenant, and promise.

The question remains, then, how divine revelation corrects and redeems the self-understanding also of the poor. How is truth, beyond the interests of one social group, known? Juan Luis Segundo (1976) describes the process as an expanding hermeneutical circle. Experience of reality from the perspective of the poor leads to ideological suspicion toward received structures of authority, morals, and dogma. This leads to a new awareness of God, which in turn creates a new hermeneutic for interpreting the biblical story. One does not escape ideology through this circle. But biblical revelation at one pole and the human condition of the poor at the other direct and correct it toward political and spiritual liberation. Paulo Freire develops the same line of thought as a teaching method in *Pedagogy of the Oppressed* (1970), with its emphasis on learning to be human in Christian-base communities through defining and struggling against oppressive powers while being transformed by God's saving love in the struggle.

CRITIQUE: SIN AND HOPE

Third, is liberation theology a universal message that offers hope to all, or a theology of and for the oppressed only? Vatican critiques, primarily in Pope John Paul II's speech to the Latin American bishops at the 1979 Puebla Conference in Mexico and in two "Instructions" from the Congregation for the Doctrine of the Faith 1984 and 1986, were especially strong on this point. (cf. A.T. Hennelly, *Liberation Theology: A Documentary History*, 1990). Authoritative for theology is not contemporary social analysis but the truth of the saving gospel of Christ revealed in scripture and interpreted by church tradition. The human situation must be understood in the light of the experience of the church through the ages as it responds in faith to God and the world. In this context one understands that the basic bondage is not just political oppression, but slavery to sin in all forms, that preferential option for the poor is concern for all who are caught in this bonda, and that Jesus's transforming, peacemaking, pardoning and reconciling love is the true liberation. Therefore, the church cannot sanction the violence of class war. It cannot identify God with historical achievement. It cannot understand freedom only as political.

REPLIES TO CRITICS

To these and to other criticisms, also from Protestant sources, liberation theologians reply variously. In replying to critics in his introduction to the revised edition (1988) of *A Theology of Liberation* Gutiérrez clearly addresses the community of the whole church with a call to join the poor in their struggle for liberation, confident of the reign of God, which is for all. Liberation, he says, is salvation on three levels: freedom from economic and political oppression, personal transformation, and ultimately redemption from sin. It is a movement with both historical and eschatological dimensions. However, his view of the church is less hierarchical and institutional than the Vatican critique. His emphasis on praxis as response to faith is also more social and historical.

Others, in their contexts, deal with the question in various ways. The *Kairos Document* calls the church to struggle against tyranny with appropriate force, with the hope that the coming reign of the risen Christ offers, but also with love for the oppressor and justice for all. Both Dalit theology in India and Black theology in the United States are more exclusively focused on the minority group whose faith they seek to express. Dalits, they claim, have their own participation in the liberating presence of the suffering Christ. They can only bear witness to God's promise for all people if they are not integrated into the ethos of the majority, of Hindu India, or even of the Christian church dominated by other castes. Similarly, for James H. Cone (1969, 1975), Christ's affirmation of black

people is central to God's liberating purpose, and salvation for white people means identifying with this experience. Minjung theologians speak in and for the church, but they understand the experience of the people of God and the suffering messiah in the Bible as offering God's promise and hope to the suffering people of Korea today. It is the minjung who are the messianic people.

These theologies differ in their identification of oppressed peoples seeking liberation, though they communicate with and learn from one another. Their views on the relation between these peoples and the church are not the same, though all have grown out of the church and speak to it. They are not always of one mind about the use of violence in the struggle against oppressive powers, though they all would condemn hatred and seek nonviolent methods where possible. They do not all agree about the relation between the struggle of the poor for political, economic, and social liberation and the ultimate freedom promised in the coming of the kingdom of God. But for all of them Christian faith is fundamental. This means for them God's special concern for the poor in their fight for justice and freedom, God's identification with them in the servanthood and suffering of Christ, and God's promise of a world in which both oppressed and oppressors will be freed from power and domination. The movement has been called utopian, a term that Gutiérrez accepts as a provisional expression of Christian hope. Whether it is also realistic, history must judge.

See also Marxist Philosophy; Marx, Karl; Natural Law; Philosophy of Religion; Reason; Reformation; Revelation; Thomas Aquinas, St.

Bibliography

Cone, James H. *Black Theology and Black Power*. New York: Seabury Press, 1969.

Cone, James H. *God of the Oppressed*. New York: Seabury Press, 1975.

Freire, Paulo. *Pedagogy of the Oppressed*. Translated by Myra Bergman Ramos. New York: Herder and Herder, 1970.

Gutiérrez, Gustavo. *The Power of the Poor in History*. Translated by Robert R. Barr. Maryknoll, NY: Orbis Books, 1983.

Gutiérrez, Gustavo. *A Theology of Liberation: History, Politics, and Salvation*. Translated and edited by Sister Caridad Inda and John Eagleson. Maryknoll, NY: Orbis Books, 1973. Rev. ed., 1988.

Hennelly, Alfred T., ed. *Liberation Theology: A Documentary History*. Maryknoll, NY: Orbis Books, 1990.

The Kairos Document, Challenge to the Church: A Theological Comment on the Political Crisis in South Africa—. Grand Rapids, MI: Eerdmans, 1986.

Kim Yong Bock, ed. *Minjung Theology: People as the Subjects of History*. Singapore: Commission on Theological Concerns, Christian Conference of Asia, 1981.

Miguez Bonino, José. *Christians and Marxists: The Mutual Challenge to Revolution*. Grand Rapids, MI: Eerdmans, 1976.

Miguez Bonino, José. *Doing Theology in a Revolutionary Situation*. Philadelphia, PA: Fortress Press, 1975.

Nirmal, Arvind P., ed. *A Reader in Dalit Theology*. Madras, India: Gurukul Lutheran Theological College and Research Institute for the Department of Dalit Theology, 1994.

Rowland, Christopher, ed. *The Cambridge Companion to Liberation Theology*. New York: Cambridge University Press, 1999.

Segundo, Juan Luis. *The Liberation of Theology*. Translated by John Drury. Maryknoll, NY: Orbis Books, 1976.

Charles C. West (1996, 2005)

LIBER DE CAUSIS

The *Liber de Causis* (or *Liber Aristotelis de Expositione Bonitatis Purae;* Book of Causes) is a Latin translation of an Arabic work that is derived from the "Elements of Theology" of Proclus (fifth century CE). The author of the Arabic work is unknown; some scholars consider it the twelfth-century composition of David the Jew (Abraham ibn Daud or Avendeath) at Toledo, while others believe it an eighth- or ninth-century product of a school of Neoplatonism in the Near East, possibly stemming from a still earlier Syriac source.

At least one Latin translation appeared before 1187, probably the product of the Toledan translator, Gerard of Cremona. The work then came to be ascribed variously to David, al-Fārābī, or Aristotle. By 1255 the Parisian Faculty of Arts, considering it a work of Aristotle, included it in the curriculum.

Among the many doctrines contained in the 211 chapters, or Propositions, of Proclus's "Elements of Theology," the following should be noted. Proclus uses the term *theology* to mean Neoplatonic metaphysics. The latter describes the necessary procession of the world, or being, from its ultimate origins. The most important of these originative principles are: first, the gods; second, the pure spirits, or Intelligences; third, souls. The supreme god, or the One, is not describable as "being," yet it is the universal cause of every being. Before producing Intelligences, the One effects a pair of opposite principles, Limit and Infinity, and then a series of subordinate gods, or "henads," which have the causal function of Plato's Forms. The immediate effect of each principle, whether the latter be a god, a spirit, or a soul, is an attribute that is

both similar to, and yet more specific than, its source. The particularity of the effect is due to its recipient. Consequently, it is difficult for the reader to see how the One can produce all things without the cooperation of its subordinates.

The thirty-two propositions of the *Liber de Causis* summarize this material with the following changes: (1) the multitude of deities (Limit, Infinity, and henads) is eliminated and divinity is reserved to the One alone; (2) the first cause is described as "being" and its causality as "creation." These changes suggest that the Neoplatonic author was either Jewish, Islamic, or Christian. Nevertheless, because the causes of Proclus act solely from the necessity of their natures and are mutually interdependent, it is questionable whether the *Liber de Causis* actually presents a monotheistic theory of free creation.

After reading William of Moerbecke's Latin translation of the "Elements of Theology" (*Elementatio Theologica,* 1268), St. Thomas Aquinas noticed for the first time that the *Liber de Causis* was not a work of Aristotle, but a modification of Proclus. Unfortunately, this discovery had to be made again during the Renaissance.

The doctrines in the *Liber de Causis* influenced many thinkers, among them: William of Auvergne, Roger Bacon, Albert the Great, John Duns Scotus, and Meister Eckhart.

See also Albert the Great; al-Fārābī; Aristotle; Bacon, Roger; Duns Scotus, John; Eckhart, Meister; Neoplatonism; Proclus; Renaissance; Thomas Aquinas, St.; William of Auvergne.

Bibliography

EDITIONS

Bardenhewer, Otto. *Die pseudo-aristotelische Schrift über das reine Gute, bekannt unter dem Namen Liber de Causis.* Freiburg im Breisgau, 1882.

Steele, Robert. *Opera Hactenus Inedita Rogeri Baconi.* Fasc. 102, "Questiones supra *Librum de Causis.*" Oxford and London: Clarendon Press, 1935.

ON THE *LIBER DE CAUSIS*

Anawati, Georges C. "Prolégomènes à une nouvelle édition du *De Causis* arabe (Kitāb al-hayr al-maḥḍ)." *Mélanges Louis Massignon* 1 (1956): 73–110.

Doresse, J. "Les sources du *Liber de Causis.*" *Revue de l'histoire des religions* 13 (1946): 234–238.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages,* 236–237. New York: Random House, 1955.

Michael W. Strasser (1967)

LIBERTARIANISM

Libertarians like to think of themselves as defenders of liberty. For example, Friedrich A. von Hayek sees his work as restating an ideal of liberty for "We are concerned with that condition of men in which coercion of some by others is reduced as much as possible in society" (1960, p. 11). Similarly, John Hospers believes that libertarianism is "a philosophy of personal liberty—the liberty of each person to live according to his own choices, provided that he does not attempt to coerce others and thus prevent them from living according to their choices" (1971, p.5). And Robert Nozick (1974) claims that, if a conception of justice goes beyond libertarian "side-constraints," it cannot avoid the prospect of continually interfering with people's lives.

Libertarians have interpreted their ideal of liberty in two basically different ways. Some, following Herbert Spencer (1820–1903), have taken a right to liberty as basic and have derived all other rights from this right to liberty. Others, following John Locke, have taken a set of rights, including typically a right to life and a right to property, as basic and have defined liberty as the absence of constraints in the exercise of these rights. Both groups of libertarians regard liberty as the ultimate political ideal, but they do so for different reasons. For Spencerian libertarians liberty is the ultimate political ideal because all other rights are derived from a right to liberty. For Lockean libertarians liberty is the ultimate political ideal because liberty is just the absence of constraints in the exercise of people's fundamental rights.

SPENCERIAN AND LOCKEAN LIBERTARIANS

Consider the view of Spencerian libertarians, who take a right to liberty to be basic and define all other rights in terms of this right to liberty. According to this view liberty is usually interpreted as being unconstrained by other persons from doing what one wants or is able to do. Interpreting liberty this way, libertarians like to limit constraints to positive acts (i.e., acts of commission) that prevent people from doing what they otherwise want or are able to do. In contrast, welfare liberals and socialists interpret constraints to include, in addition, negative acts (acts of omission) that prevent people from doing what they otherwise want or are able to do. In fact, this is one way to understand the debate between defenders of *negative liberty* and defenders of *positive liberty*. This is because defenders of negative liberty interpret constraints to include only positive acts of others that prevent people

from doing what they otherwise want or are able to do, while defenders of positive liberty interpret constraints to include both positive and negative acts of others that prevent people from doing what they otherwise want or are able to do.

Suppose then we interpret constraints in the manner favored by libertarians to include only positive acts by others that prevent people from doing what they otherwise want or are able to do. Libertarians go on to characterize their political ideal as requiring that each person should have the greatest amount of liberty commensurate with the same liberty for all. From this ideal they claim that a number of more specific requirements, in particular a right to life, a right to freedom of speech, press, and assembly, and a right to property, can be derived.

Here, it is important to observe that the libertarian's right to life is not a right to receive from others the goods and resources necessary for preserving one's life. It is not a right to welfare: It is simply a right not to be killed unjustly. Correspondingly, the libertarian's right to property is not a right to receive from others the goods and resources necessary to meet one's basic needs, but a right to acquire goods and resources either by initial acquisitions or by voluntary agreements.

Of course, libertarians would allow that it would be nice of the rich to share their surplus goods and resources with the poor. Nevertheless, they deny that government has a duty to provide for such needs. Libertarians claim that some good things, such as providing welfare to the needy, are requirements of charity rather than justice. Accordingly, failure to make such provisions is neither blameworthy nor punishable. As a consequence, libertarians contend that such acts of charity should not be coercively required. For this reason they are opposed to any coercively supported welfare program.

For a similar reason libertarians are opposed to coercively supported opportunity programs. This is because the basic opportunities one has under a libertarian conception of justice are primarily a function of the property one controls, and since unequal property distributions are taken to be justified under a libertarian conception of justice, unequal basic opportunities are also regarded as justified.

The same opposition to coercively supported welfare and equal opportunity programs characterizes Lockean libertarians, who take a set of rights, typically including a right to life and a right to property, as basic and then interpret liberty as being unconstrained by other persons from doing what one has a right to do. According to this view a right to life is simply a right not to be killed unjustly; it is not a right to receive welfare. Correspondingly, a right to property is a right to acquire property either by initial acquisitions or by voluntary transactions; it is not a right to receive from others whatever goods and resources one needs to maintain oneself. Understanding a right to life and a right to property in this way, libertarians reject both coercively supported welfare programs and equal opportunity programs as violations of liberty.

A PARTIAL DEFENSE

In support of their view libertarians advance examples of the following sort. The first two are adapted from Milton Friedman (1962), and the last one is from Robert Nozick (1974).

In the first example you are to suppose that you and three friends are walking along the street and you happen to notice and retrieve a $100 bill lying on the pavement. Suppose a rich fellow had passed by earlier throwing away $100 bills, and you have been lucky enough to find one of them. Now, according to Friedman, it would be nice of you to share your good fortune with your friends. Nevertheless, they have no right to demand that you do so, and hence, they would not be justified in forcing you to share the $100 bill with them. Similarly, Friedman would have us believe that it would be nice of us to provide welfare to the less fortunate members of our society. Nevertheless, the less fortunate members have no right to welfare, and hence they would not be justified in forcing us to provide such.

The second example, which Friedman regards as analogous to the first, involves supposing that there are four Robinson Crusoes, each marooned on four uninhabited islands in the same neighborhood. One of these Crusoes happens to land on a large and fruitful island, which enables him to live easily and well. The others happen to land on tiny and rather barren islands from which they can barely scratch a living. Suppose one day they discover the existence of each other. Now, according to Friedman, it would be nice of the fortunate Crusoe to share the resources of his island with the other three Crusoes, but the other three Crusoes have no right to demand that he share those resources, and it would be wrong for them to force him to do so. Correspondingly, Friedman thinks it would be nice of us to provide the less fortunate in our society with welfare, but the less fortunate have no right to demand that we do so, and it would be wrong for them to force us to do so.

In the third example Nozick asks us to imagine that we are in a society that has just distributed income

according to some ideal pattern, possibly a pattern of equality. We are further to imagine that in such a society someone with the talents of Wilt Chamberlain or Michael Jordan offers to play basketball for us provided that he receives one dollar from every home game ticket that is sold. Suppose we agree to these terms, and two million people attend the home games to see this new Wilt Chamberlain or Michael Jordan play, thereby securing for him an income of $2 million. Since such an income would surely upset the initial pattern of income distribution whatever that happened to be, Nozick contends that this illustrates how an ideal of liberty upsets the patterns required by other conceptions of justice and calls for the rejection of these conceptions of justice.

THE MINIMAL OR NIGHT-WATCHMAN STATE

Libertarians think that only a minimal or night-watchman state can be justified in terms of their ideal of liberty. The libertarian argument for the minimal or night-watchman state begins with the acceptable premise that voluntary agreements represent an ultimate ideal for social interaction. This ideal, libertarians contend, finds its fullest expression in a market economy where buyers and sellers, employers and employees, voluntarily agree to exchange the goods they possess. Thus, it is assumed that the requirements for voluntary agreements between persons with unequal resources are easily satisfied in a market economy. As long as alternative contractual arrangements make it possible for buyers and sellers, employers and employees, to take their business elsewhere, libertarians believe that agreements reached in market transactions are completely voluntary. On these grounds libertarians claim that the only significant role left for the state is to prevent and rectify departures from a market economy resulting from fraud, theft, or the use of force. Any more extensive role for the state, they contend, would restrict people's liberty; that is to say, it would restrict liberty understood negatively as the absence of interference by other persons. Accordingly, libertarians conclude that only a night-watchman state can be justified in terms of an ideal of negative liberty.

The libertarian argument for the night-watchman state also seeks to show that other social ideas cannot justify a more extensive state. Libertarians either maintain that other social ideals purporting to justify a more extensive state are themselves without justification, or they claim that these social ideals have lower priority when compared with the ideal of negative liberty. But there are not always agreements as to which critical approach is appropriate. Thus with respect to an ideal of equality, Nozick (1974) and Hayek (1960) adopt different approaches: Nozick maintains that an ideal of equality has not been effectively justified, while Hayek maintains that the ideal has some validity but that negative liberty is the superior ideal. Allowing for such disagreements, both critical approaches could also be used by libertarians against various conceptions of positive liberty.

Nozick even goes so far as to claims that taxation of earnings from labor is on a par with forced labor. Still, libertarians are not similarly sensitive to the loss of liberty that occurs in the marketplace. For example, when an employer decides to lay someone off, Hospers (1971) claims that the employer is simply deciding against continuing a voluntary exchange and is not restricting the person's liberty. Likewise, Hayek (1960) claims that as long as workers who are laid off can find alternative employment their liberty is not being restricted. But how can requiring a person to pay $500 into a social security program under threat of greater financial loss infringe on the person's liberty when requiring a person to take a job paying $500 less under threat of greater financial loss does not infringe on the person's liberty? Surely it would seem that if one requirement restricts a person's liberty, the other will also.

To distinguish these cases, some libertarians claim that only intentional interference by others restricts a person's liberty. Requiring a person to pay $500 into a social security program under threat of greater financial loss, they contend, is intentional interference by others and hence restricts the person's liberty, while requiring a person to take a job paying $500 less under a similar threat is but the unintended result of individuals trying to better themselves in a market economy and hence does not restrict the person's liberty. But whether interference with a person's life is intentional or not is relevant only when determining the extent to which others are responsible for that interference. Although people are clearly more responsible for actions done intentionally they can still be responsible for actions done unintentionally, especially if they were morally negligent and should have foreseen the consequences of their actions. Since moral responsibility can extend to both intentional and unintentional interference with a person's life, there seems to be no reason for not considering both types of interference to be restrictions of a person's liberty. What is crucial to liberty as a social ideal is whether people are morally responsible for interfering with a person's life irrespective of whether that interference is intentional or not.

A BASIC DIFFICULTY

A basic difficulty with the libertarian's conception of justice is the claim that rights to life and property, as the libertarian understands these rights, derive from an ideal of liberty. Why should we think that an ideal of liberty requires a right to life and a right to property that excludes a right to welfare? Surely it would seem that a right to property (as the libertarian understands it) might well justify a rich person's depriving a poor person of the liberty to acquire the goods and resources necessary for meeting his or her basic nutritional needs. How then could we appeal to an ideal of liberty to justify such a deprivation of liberty? Surely we couldn't claim that such a deprivation is justified for the sake of preserving a rich person's freedom to use the goods and resources he or she possesses to meet luxury needs. By any neutral assessment it would seem that the liberty of the deserving poor not to be interfered with when taking from the surplus possessions of the rich what they require to meet their basic needs would have priority over the liberty of the rich not to be interfered with when using their surplus possessions to meet their luxury needs. But if this is the case, then a right to welfare, and possibly a right to equal opportunity as well, would be grounded in the libertarian's own ideal of liberty.

See also Justice; Liberty; Locke, John; Nozick, Robert; Philosophy of Economics; Responsibility, Moral and Legal; Rights; Social and Political Philosophy; Socialism.

Bibliography

Boas, David. *Libertarianism: A Primer*. New York: Free Press, 1997.

Friedman, Milton, with Rose D. Friedman. *Capitalism and Freedom*. Chicago: University of Chicago Press, 1962.

Hayek, Friedrich A. von. *The Constitution of Liberty*. Chicago: University of Chicago Press, 1960.

Hospers, John. *Libertarianism*. Los Angeles: Nash, 1971.

Machan, Tibor. *The Passion for Liberty*. Lanham, MD: Rowman & Littlefield, 2003.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Sterba, James P. *Justice for Here and Now*. New York: Cambridge University Press, 1998.

James Sterba (2005)

LIBERTY

One of the central concerns of social and political philosophy has been the issue of what limits, if any, there are to the right of the state to restrict the "liberty" of its citizens. Unless one is convinced of the truth of anarchism, there are some actions with which the state may legitimately interfere, and unless one accords no value to personal liberty, there are some actions the state must leave to the discretion of the individual. One of the tasks of political philosophy is to develop and elaborate a theory to determine where these boundaries lie.

In his classical defense of liberalism—*On Liberty*—John Stuart Mill gave one influential answer to this question. The only reason that could justify the use of coercion against a person is to prevent harm to other people. Such a reason might not be decisive—it might be that the use of coercion would be ineffective or too costly or would violate the rights of privacy—but it brings the action in question within the scope of legitimate state power.

Other reasons, according to Mill, do not justify legal coercion. One cannot restrict someone's actions because they are harmful to that person; paternalism is not legitimate. One cannot restrict someone's actions because they are wrong or immoral (but not harmful to others); legal moralism is not legitimate. One cannot restrict someone's actions because his or her character would be improved by doing so; moral paternalism is not legitimate.

Obviously, a theory that puts such heavy weight on the notion of harm gives rise to disputes about the nature and limits of that notion. If conduct is offensive to others, does that count as harming them? If not, do we need a separate principle to justify prohibiting offensive conduct such as public nudity or racist graffiti? If we are competing for a job and you get it, am I harmed by this? Does only physical damage count as harm or emotional damage as well? Am I harmed by simply knowing that behind the walls of your house you are engaged in activities that I would find repulsive or wicked? If someone defaces the flag, is anyone harmed by this? If I consent to some action that is otherwise damaging to me, am I still harmed? Can I be harmed after my death—for example, by attacks on my reputation?

One of the most fully developed views that seeks to provide answers to these and similar questions is that of Joel Feinberg. He argues that any notion of harm that is going to play a role in answering normative questions will itself be normative in character. He accordingly defines the notion of harm in terms of a wrongful setback to a person's interests. To some extent, naturally, this shifts philosophical attention to the concept of interests.

PATERNALISM

The normative issue raised by paternalism is when, if ever, the state or an individual is entitled to interfere with a person for that person's good. Examples of laws that have been justified in paternalistic terms include requiring motorcyclists to wear helmets, forcing patients to receive blood transfusions against their wishes, or requiring individuals to save for their retirement (Social Security).

The reasons that support paternalism are those that support any benevolent action—promoting the welfare of a person. The reasons against are those that militate against any interference with the autonomy of individuals—respect for their desire to lead their own lives. Normative debates about the legitimacy of paternalism involve disputes about many issues including the nature of welfare (can we produce good for a person against that individual's preferences and evaluations?), the correctness of various normative theories (consequentialism vs. autonomy or rights-based theories), and the relevance of hypothetical consent (in Mill's famous example of the man walking across a bridge that, unknown to him, is about to collapse, we may stop him, since he would not want to cross the bridge if he knew its condition).

LEGAL MORALISM

The issue of whether the state may enforce morality—the subject that was brought to philosophical prominence by the debate between Lord Devlin and H. L. A. Hart—is present in discussions of the legalization of homosexuality, pornography, surrogate motherhood, and active euthanasia. The focus of such discussion is not the harm of such activities but their immorality and whether if they are immoral that is sufficient reason for the state to proscribe them. Since it is clearly the case that one of the grounds for proscribing murder is its immorality, the question arises as to what it might mean to deny that the state should take morality into account in limiting liberty. The best answer is that we may distinguish within the immoral different realms—for example, matters having to do with rights as opposed to matters having to do with ideals of conduct. Those who are opposed to the enforcement of morality are really opposed to enforcing certain areas of morality. Much of the discussion goes on under the heading of the "neutrality" of the liberal state.

See also Anarchism; Consequentialism; Euthanasia; Feinberg, Joel; Hart, Herbert Lionel Adolphus; Liberalism; Mill, John Stuart; Paternalism; Rights; Social and Political Philosophy.

Bibliography

Acton, Lord. *Essays in the History of Liberty.* Edited by J. R. Frears. Indianapolis: Liberty Classics, 1985.

Berlin, Isaiah. *Four Essays on Liberty.* Oxford: Oxford University Press, 1969.

Devlin, P. *The Enforcement of Morals.* London: Oxford University Press, 1965.

Dworkin, G. "Paternalism." *Monist* 56 (1972): 64–84.

Feinberg, J. *The Moral Limits of the Criminal Law.* 4 vols. New York: Oxford University Press, 1984–1988.

Hart, H. L. A. *Law, Liberty, and Morality.* Stanford, CA: Stanford University Press, 1963.

Mill, J. S. *On Liberty.* London, 1959.

Miller, David, ed. *Liberty.* Oxford: Oxford University Press, 1991.

Gerald Dworkin (1996)
Bibliography updated by Philip Reed (2005)

LICHTENBERG, GEORG CHRISTOPH

(1742–1799)

Georg Christoph Lichtenberg, the German satirist, scientist, and philosopher, studied mathematics and science at the University of Göttingen and was a professor there from 1767 to the end of his life. On two occasions Lichtenberg visited England. His impressions from these visits are recorded in his diaries and letters.

Lichtenberg's original contributions to mathematics and to pure and experimental science are not of great importance. The Lichtenberg figure in the theory of electricity was named after him. He was very successful as a teacher; among his pupils were Alexander von Humboldt and Christian Gauss. It has been said that his fame as a lecturer and demonstrator surpassed that of any other German scientist of his time.

His literary reputation with his contemporaries rested mainly on his satirical criticism of the writers of the *Sturm und Drang* movement and of the Swiss clergyman Johann Lavater's quasi-scientific psychology of character. Lichtenberg's own favorites and models in art were Englishmen: William Shakespeare; David Garrick, the actor; and William Hogarth, the painter. His analyses and descriptions of Garrick on the stage and his detailed "explanations" of Hogarth's etchings have become famous. Most of Lichtenberg's literary output during his lifetime appeared in two periodicals, of which he was the editor, the *Göttinger Taschen-Calender* and the *Göttingisches Magazin der Wissenschaften und Litteratur.*

By far the most valuable part of Lichtenberg's literary work, however, consisted of his "aphorisms," or scattered thoughts on psychological, philosophical, scientific, and many other topics. They were written down in notebooks but were never systematically arranged by the author. Nor were they used as raw material to any great extent for the more systematic work that Lichtenberg was constantly planning but never carried out. *Vermischte Schriften,* a comprehensive selection of his remarks, was published soon after his death.

Philosophically, Lichtenberg was not attached to any school or movement. The thinkers who made the deepest impressions on him were Benedict de Spinoza and Immanuel Kant. It is noteworthy that Lichtenberg was an early reviver of the great Jewish philosopher and one of the first to understand and acknowledge the revolutionary significance of Kant's transcendental philosophy. Furthermore, the versatility of his philosophical intellect is shown by his acute understanding of the work of Jakob Boehme.

Lichtenberg has had but a modest influence on the development of thought, but it is evident from the observations of Kant and Alexander von Humboldt, among others, that his contemporaries greatly prized his philosophical intellect. Subsequent generations were first made aware of his status as an independent thinker through the observation of Ernst Mach (in his *The Analysis of Sensations*) that Lichtenberg had anticipated the empiriocritical solution of the ego problem with his critique of the Cartesian *cogito ergo sum.* (In another work, "Die Leitgedanken meiner naturwissenschaftlichen Erkenntnislehre" [The primary ideas of my scientific epistemology], 1919, p. 5, Mach even hinted that he had been influenced by Lichtenberg.) Moreover, the affinity of Lichtenberg's ideas with modern linguistic philosophy has been indicated by various writers, for example, Friedrich Waismann in the preface to Moritz Schlick's *Gesammelte Aufsätze* and Richard von Mises in *Positivism.*

PHILOSOPHY OF MATHEMATICS

Lichtenberg, in contradistinction to Kant, distinguished sharply between pure and applied mathematics and separated mathematics as a logicodeductive formalism from mathematics as a theory of reality.

The truths of pure mathematics are not only certain in a strict sense but are derived (in principle) independently of experience and empirical observation. A blind man, for instance, could discover the laws of light by means of the calculus, for as soon as the fundamental facts of refraction and reflection are discovered experimentally, "the whole of dioptrics and catoptrics becomes a purely geometrical problem," which can be treated without further knowledge of natural processes. For this reason the ideal form of a scientific theory is that of a logicodeductive system. Lichtenberg stated: "The aim of the physicists is to prepare the way for mathematics."

In his conception of pure mathematics, Lichtenberg approached the notion of the analytical, or tautological, character of mathematical truths. He did not take a positive stand on Kant's view of the synthetic a priori character of mathematics, but it is evident from his remarks that he viewed it with suspicion.

Mathematics shapes its own world. The business of the physicist is to decide which "of the innumerable suppositions possible" is the single true one. The results of mathematical deduction cannot be asserted in advance to agree with the results of physical inquiry. "Their agreement is a purely *empirical* coincidence, nothing else." (It is apparent from his manuscript that Lichtenberg ascribed great importance to this remark.) Thus Lichtenberg renounced all a priori claims concerning the application of mathematics to reality.

Instead of being astounded at the actual success of mathematics in the exploration of natural phenomena, Lichtenberg emphasized the approximate character of mathematical laws of nature and warned of the temptation to read more mathematics into things than is actually there. "All mathematical laws that we find in nature, despite their beauty, are doubtful to me." The forms in which nature covers herself are too manifold and changeable to be comprehended exhaustively by our own conceptual apparatus. These thoughts, which had come early to Lichtenberg, were closely connected with his highly developed talent for observation and his acute feeling for the concrete.

It is characteristic that the work with which Lichtenberg qualified for his professorship was devoted to the study of an alleged discrepancy between theory and experience. This work, "Considerations about Some Methods for Removing a Certain Difficulty in the Calculation of Probability in Gambling" (not mentioned in J. M. Keynes's bibliography in his *Treatise on Probability*), concerned a famous problem of the theory of probability, the so-called Petersburg paradox, which engaged many leading mathematicians of the eighteenth century, among others, Jean Le Rond d'Alembert and Daniel Bernoulli. It is erroneous, however, to see in this problem, as Lichtenberg and others have done, a contradiction between the mathematical calculus and the actual course of events.

Recognition is due Lichtenberg for his scientific genius in being one of the first to see the possibility of denying, without contradiction, the Euclidean axioms. That between two points only one straight line can be drawn is indeed an accepted axiom, but it is by no means necessary. One can also conceive of the possibility that several distinct lines might pass through the same two points. The manner in which Lichtenberg attempted to show this possibility was, indeed, less significant: He imagined one could take arcs with the radii ∞, ∞^2, ∞^3, and so on, so that they proceed through two fixed points, describing distinct straight lines.

Interestingly enough, Lichtenberg also expressed some thoughts about the deflection of light through gravitation. As an adherent of Isaac Newton's corpuscular theory, he assumed that light has mass, from which it follows that a beam of light must deviate from a straight path because of its weight. "Light alone appears to be an exception (viz., to the curved path of most bodies); however, since it is probably heavy, it will be deflected as a result."

EPISTEMOLOGY OF THE EXACT SCIENCES

Lichtenberg realized the great significance of the discovery of structural identities among qualitatively different domains of theoretical research into nature. His idea of *paradigmata* (patterns), according to which processes were to be "declined," seems to have approached James Clerk Maxwell's view of the significance of analogy and to have anticipated the concept of isomorphism. Lichtenberg called discovery through *paradigmata* the most fruitful of all the heuristic devices of science. As an example of an application he suggested that one might use Newton's *Optics* as a model in the theory of the calcination of metals.

Lichtenberg had a clear view of the logic of constructing hypotheses: "If we want to understand nature," he said, "we must begin with sensible appearances." Hypotheses that transcend the evidence of the senses may only be constructed insofar as they can be tested within the domain of appearances. Concepts whose presence or absence in the individual case can never be demonstrated but only assumed are not permissible in science. The concept of ether in physics belongs to this category. The ether, which "no one has seen or felt, … condensed, rarefied, etc.," is like the notion of the world soul: Since it has no experiential consequences, it must be eliminated once and for all from a rational physics.

In spite of his opposition in principle to hypothesis making in physics, Lichtenberg did not agree with the view that all assumptions should be discarded if, although they have testable consequences, they do not literally correspond to sensible reality. Assumptions of this kind may nonetheless be useful as pictures of complicated courses of events, and thus facilitate the application of mathematics to nature. (The notion of "picture," reminiscent of Heinrich Hertz's *Principles of Mechanics,* turns up often in Lichtenberg.) "If someone could make a clock that presented the movements of the heavenly bodies as exactly as actually obtains, would he not deserve much credit, even though the world does not operate by means of cog-wheels? Through this machine he could discover many things that he would not have believed to be present in it." In addition to such mechanical models, the two theories of light and atomic theory also belong to this category.

The truth content of scientific assumptions of the type mentioned above is proportional to their explanatory power and to their relative simplicity. Lichtenberg quite aptly noted that with theories as complex as that of light "it can no longer be merely a question of what is true, but of what manner of explanation is the simplest." And he added, "The door to truth is through simplicity." Moreover, his speculation that one could attempt to combine the corpuscular and wave theories sounds very modern.

The falsification of such hypotheses can not be established beyond question by empirical circumstances. A single negative instance does not in general make it necessary to renounce a comprehensive scientific theory that has otherwise been well confirmed. "One should take special note of contradictory experiences," wrote Lichtenberg, "until there are enough of them to make constructing a new system worthwhile."

SOUL AND MATTER, REALISM AND IDEALISM

Early in his career Lichtenberg rejected the idea of the soul as a substance. Before enough was understood to explain the phenomena of the world scientifically, spirits were accepted as explanations of phenomena. As our knowledge of the physical world increased, however, the boundaries of the spiritual realm shrank until finally "that which haunts our body and produces effects in it" was the only thing left that required a ghost for an explanation. The case of the "soul" is like that of phlogiston: In the end both substances dissipate into nothingness. What remains is a "bare word" comparable to the word *state* (*Zustand*), to which, however, one may at least attribute

heuristic value as a picture and as a type of idea innate in the human being.

According to Lichtenberg, the thesis of materialism is "the asymptote of psychology." In psychology, he linked himself closely with the materialistic-mechanistic association theories of the Englishmen David Hartley and Joseph Priestley. A one-to-one correspondence obtains between the mental occurrence and the state of the brain, so that the former can, in principle, be inferred from the latter.

Lichtenberg, however, did not accept metaphysical materialism. Parallel with his critique of the concept of the soul went a critique of the concept of matter. Soul and inert matter are mere abstractions, he wrote in a letter in 1786; we know of matter and of soul only on the basis of the *forces* (*Kräfte*) through which they manifest themselves and "with which they are identical." We postulate for these forces in one case "an inert receptacle and call it matter." Through such a hypostatization, which is just a "chimera" of the brain, arises "the infamous dualism in the world": the division of being into body and soul, spirit and matter. But in reality everything is one.

This acknowledgment of monism still bore a metaphysical character. It is probable that the influence of Spinoza had its effect on the position taken by Lichtenberg in 1786, since the letter of that year referred directly to Spinoza. But we may observe that, much earlier, Lichtenberg had expressed the same opinion almost word for word. However, it is not impossible that the influence of Spinoza was already at work then. Even in his earliest books of aphorisms there were remarks of a Spinozistic character, although the name of the great thinker was not mentioned.

Later, Lichtenberg's monism took a more epistemological turn in that he clearly indicated how the basis of his monistic system should be interpreted. "We are aware only of the existence of our sensations, ideas, and thought," he said and expressed the same thought with the words, "Everything is feeling (*Gefühle*)." We experience a part of our impressions as dependent upon us, another as independent of the perceiving subject: in this way we arrive at the difference between the inner and outer worlds.

To argue from sensations to an "ego" as their bearer, as René Descartes does, is not logically warranted. Lichtenberg remarked very perceptively: "One should say, 'There is thinking,' just as one says, 'There is lightning.'" To say *cogito* is to say too much; for as soon as one translates it as "I think," it seems necessary to postulate an ego. Lichtenberg's earlier critique of the idea of the soul culminates here in a critique of the self, somewhat reminiscent of the position of David Hume.

It took considerable effort on Lichtenberg's part to attain clarity on the question of how we proceed from our sensations to things outside us. He perceived the significance of the problem from his study of Kant, and in his treatment of it we can generally discern Kant's influence.

At first it was very difficult for Lichtenberg to rid himself of the idea that something in the actual world might correspond to our representations, although we can have "no conception at all of the true nature of the outside world." But later he recognized that the question "whether things outside ourselves really exist and exist as we see them" is in fact "completely meaningless." It is just as foolish as asking whether the color blue is really blue. We are compelled by our nature—this compulsion he termed, with Kant, *die Form der Sinnlichkeit* (form of sensibility)—to express ourselves in such way that we speak of certain objects of our perception as being outside ourselves and of others as being within us. "What is outside? What are objects *praeter nos?* What is the force of the preposition *praeter?* It is a purely human invention; a name to indicate a difference from other things which we call 'not-*praeter nos.*'" "There is probably no one in the world who does not perceive this *difference,* and probably no such person will ever exist; and for philosophy that is enough. Philosophy need not go beyond this."

Is not this standpoint "idealism"? Lichtenberg clearly perceived that, just as his critique of the idea of the "soul" did not result in metaphysical materialism, so his attitude toward the question of the reality of the outer world should not be confused with metaphysical idealism. Rather his doctrine stood beyond idealism and materialism in their traditional senses. "It is truly of little consequence to me whether one wants to label this idealism. Names have no significance. It is at least an idealism which, through idealism, acknowledges that there are things outside us." What more can one ask? For human beings, "at least for the philosophical ones," there is no other reality than the one so constituted. It is true that one is satisfied in ordinary life with some other, "lower station," but whenever one begins to philosophize, one cannot but accept this enlightened point of view. "There is no other alternative," he concluded.

LICHTENBERG'S CONCEPTION OF PHILOSOPHY

"Our entire philosophy," wrote Lichtenberg, "is a correction of linguistic usage." What he meant by that is espe-

cially evident in his treatment of the question of realism. As indicated above, Lichtenberg's conception should not be understood as an attempt to deny the existence of things outside ourselves. That would have been a senseless undertaking. His intention was only to discover the meaning of the distinction between outer and inner objects by clearly presenting the facts that underlie this distinction. It turns out that the root of the traditional difficulty about the question of realism is that in ordinary life we attach a contradictory meaning to the expression "outside ourselves." When we have become conscious of this contradiction and have undertaken the proper correction of our linguistic usage, the difficulty vanishes of itself.

Philosophy, then, is a critique of language. Its goal, however, is not definitions of concepts. Lichtenberg was not of the opinion that one could, for philosophical use, replace the common language with an ideal language, perhaps in the sense of Gottfried Wilhelm Leibniz's *characteristica universalis.* Attempts to reform the nomenclature of the sciences did not find much favor with him. "To clarify words does not help," he said. Why? Because the interpretation of the clarified concepts takes place, in the final analysis, in the vernacular. But the vernacular, by its nature, is imbued with our false philosophy. The rectification of colloquial usage, which leads to true philosophy, is thus undertaken in the language of false philosophy: "We are therefore constantly teaching true philosophy with the language of the false one." The common philosophy, then, always maintains a certain superiority over the enlightened one, for the former is in possession of the "declensions and conjugations" of our language, and these are not changed by the clarification of meanings of words. "The invention of language preceded philosophy, and it is just this that makes philosophy difficult, particularly when one wishes to make it understandable to those who do not think much for themselves. Philosophy, whenever it speaks, is forced to speak the language of nonphilosophy.… Pure philosophy still imperceptibly enjoys the pleasure of love with the impure (and cannot avoid doing so)."

The philosopher, then, speaks with the words of the common language about things that are beyond it. He is thus compelled to express himself, to a certain degree, in metaphors (*Gleichnissen*). He is supposed to direct our attention with his sentences to the false logic of our language, so that we learn to see the world correctly. He does not teach us a new language but helps us to express ourselves clearly with our own. "The peasant," said Lichtenberg, "uses all the sentences of the most abstract philosophy, only they are entangled, hidden, confined, latent, as the physicist and chemist say; the philosopher gives us the pure sentences."

It should be evident from the above that Lichtenberg anticipated the conception of philosophy that has been represented in the twentieth century by Ludwig Wittgenstein. Wittgenstein knew Lichtenberg's work well and esteemed it highly. It is hardly possible, however, to speak of Lichtenberg as an influence on the philosophy of Wittgenstein. Nevertheless, a rare congeniality between the two men can be noted—not only in view of their conceptions of philosophy but also in view of their entire intellectual talents and temperaments.

See also Alembert, Jean Le Rond d'; Boehme, Jakob; Descartes, René; Hartley, David; Idealism; Kant, Immanuel; Keynes, John Maynard; Lavater, Johann Kaspar; Leibniz, Gottfried Wilhelm; Mach, Ernst; Materialism; Mathematics, Foundations of; Maxwell, James Clerk; Newton, Isaac; Priestley, Joseph; Schlick, Moritz; Spinoza, Benedict (Baruch) de; Wittgenstein, Ludwig Josef Johann.

Bibliography

Lichtenberg's works in German include *Vermischte Schriften,* edited by L. C. Lichtenberg and F. Kries, 5 vols. (Göttingen: Dieterich, 1800–1803); *Physikalische und mathematische Schriften,* edited by L. C. Lichtenberg and F. Kries, 4 vols. (Göttingen, 1803–1806); *Neue Original-Ausgabe,* 14 vols. (Göttingen, 1844–1853); *Aus Lichtenbergs Nachlass,* edited by Albert Leitzmann (Weimar, 1899); *Aphorismen, nach den Handschriften,* edited by Albert Leitzmann, 5 vols. (Berlin, 1902–1908); *Lichtenbergs Briefe,* edited by Albert Leitzmann and Carl Schüddelkopf, 3 vols. (Leipzig, 1901–1904).

Lichtenberg's remarks on questions of mathematics and physics have been printed only in part. It is most unfortunate that all of his notes from the years 1779–1788 and the greater part of those from 1793–1796, which existed at the time of the first edition of the *Vermischte Schriften,* had been lost when Albert Leitzmann, in the beginning of the twentieth century, edited the *Aphorismen, nach den Handschriften.* This loss greatly complicates the task of reconstructing the course of development of Lichtenberg's thought. The selection of aphorisms in the *Vermischte Schriften* shows that some of his most important philosophical remarks were among those subsequently lost.

For literature on Lichtenberg, see J. Dostal-Winkler, *Lichtenberg und Kant* (Munich, 1924); P. Hahn, *Georg Christoph Lichtenberg und die exakten Wissenschaften* (Göttingen: Vandenhoeck and Ruprecht, 1927); F. H. Mautner, "Amintors Morgenandacht," in *Deutsche Vierteljahrschrift für Litteraturwissenschaft und Geistesgeschichte* 30 (1956); F. H. Mautner and F. Miller, "Remarks on G. C. Lichtenberg, Humanist-Scientist," in *Isis* 43 (1952); A. Neumann, "Lichtenberg als Philosoph und seine Beziehungen zu Kant," in *Kantstudien* 4 (1900); A.

Schneider, *Georg Christoph Lichtenberg, précurseur du romantisme,* Vol. I, *L'homme et l'oeuvre,* Vol. II, *Le penseur* (Nancy, 1954); J. P. Stern, *Lichtenberg: A Doctrine of Scattered Occasions* (Bloomington: Indiana University Press, 1959); and G. H. von Wright, "Georg Christoph Lichtenberg als Philosoph," *Theoria* 8 (1942): 201–217, of which the present entry is an adaptation.

Georg Henrik von Wright (1967)
Translated by David H. DeGrood and Barry J. Karp

LIEBERT, ARTHUR

(1878–1946)

Arthur Liebert, the German neo-Kantian philosopher, was born Arthur Levi in Berlin. The son of a merchant, he spent six years in business after completing his secondary education in 1895. He then entered the University of Berlin, where he received his doctorate in 1908. After teaching at the Berlin Handelshochschule, Liebert lectured at the University of Berlin, becoming extraordinary professor in 1925. From 1918 to 1933 he was coeditor with Paul Menzer of *Kantstudien,* which became under their guidance an instrument of growing international cooperation in philosophy. Forced to leave Germany in 1933, when the National Socialists came to power, he was appointed professor of philosophy at the University of Belgrade and there founded the journal *Philosophia: Philosophorum Nostri Temporis Vox Universa,* which appeared at irregular intervals from 1936 to 1939. When the German armies invaded the Balkans, he found refuge in England, where he published *Das Wesen der Freiheit* (1944) and, together with other refugees, organized the Freier deutscher Kulturbund in Grossbrittanien. At the end of World War II he returned to his restored professorship at Berlin, but he died shortly thereafter.

Liebert was influenced by the realistic interpretation given Immanuel Kant at Berlin by Friedrich Paulsen, Alois Riehl, and especially by Wilhelm Dilthey, who stressed the historical aspects of the *Geisteswissenschaften* (cultural sciences). Within this realistic neo-Kantian orientation, Liebert turned to the ethical problems of value and freedom and to the search for a dialectic movement of ethical and metaphysical categories in history. Many of his writings, particularly in his later years, were devoted to the promotion of worldwide philosophical cooperation as "the free guardian of freedom" and particularly to the development of a philosophical organization, "an Areopagus of mankind," within which the new humanism was to be promoted. This is the theme of "On the Duty of Philosophy in Our Age" (*Von der Pflicht der Philosophie in unserer Zeit*), published during his exile in 1938.

Liebert's philosophical efforts to work out his critical metaphysics as a dialectic were to have taken the form of a large work titled *Geist und Welt der Dialektik,* of which only the first volume, *Grundlegung der Dialektik,* appeared (Berlin, 1929). To be distinguished from science, philosophy must accept as its field not simply being (*Sein*) but value (*Geltung*), for being not merely *is,* but *is valid* (*gilt*), or validates, itself. In opposition to the Baden neo-Kantians, Liebert rejected obligation (*Sollen*) as the ground of value, finding a new basis for metaphysics in the Kantian concern for the validation of judgments. "The right of metaphysics and the right to a metaphysics," he wrote, "flow from the idea and right of philosophy itself." The task of metaphysics thus becomes that of a historical "critical phenomenology" that "tests its own possibility and justification and derives its presuppositions and conclusions through reason."

Such a metaphysics does not merely use dialectic as the basis of metaphysical criticism but is itself dialectic. Its categories must include both philosophical ideas and the social and cultural contexts out of which they arise. "The idea of dialectic is at once the a priori condition and the definitive force (*massgebende Kraft*) for the construction of metaphysics, and also the distinctive instrument for penetrating into the nature of metaphysics, and for studying and understanding it." This dialectic must include within the scope of its critical and dynamic movement four motives: the intellectual, moral, aesthetic, and religious. Metaphysics is no longer "ontological-dogmatic" but "actualistic-critical"; the movement of its categorical structures of value combines temporal and supratemporal viewpoints. Its task is apparently never completed, because historical change outgrows the adequacy of every a priori structure. In particular, the modern world with its conflicts prevents a return to the classical humanizing harmonies of thought; the historical-normative dialectic that modern life calls forth must take the form of tragedy.

Liebert's lectures and seminars were devoted to the development and illustration of this conception of metaphysics. The *Grundlegung der Dialektik* provided only an introduction, in which Liebert traced the beginnings of the metaphysical dialectic in the thought of his contemporaries—practitioners of the *Geisteswissenschaften*; metaphysicians and theologians; and neo-Kantians and neo-Hegelians.

The Kantian identification of freedom with reason remained for Liebert the fixed a priori point of view of his

"actualistic-critical" metaphysics. He persistently attacked the currently popular forms of *Lebensphilosophie* as relativistic, irrational, and sacrificing philosophical freedom. Philosophers were called upon to fulfill their vocation by turning to metaphysics and ethics as guides for individual and organizational action against the forces of irrationalism and cultural decay.

Liebert's thought has received little attention since his death. His most important writings are those in which he sought to formulate the principles of his own historical metaphysics of value.

See also Dilthey, Wilhelm; Geisteswissenschaften; Kant, Immanuel; Metaphysics; Neo-Kantianism; Paulsen, Friedrich; Riehl, Alois; Value and Valuation.

Bibliography

ADDITIONAL WORKS BY LIEBERT

Das Problem der Geltung. Berlin, 1906.

Der Geltungswert der Metaphysik. Berlin: Reuther and Reichard, 1915.

Wie ist kritische Philosophie überhaupt möglich? Leipzig, 1919.

Die geistige Krisis der Gegenwart. Berlin: Heise, 1923.

Von der Pflicht der Philosophie in unserer Zeit. Zürich, 1946.

L. E. Loemker (1967)

LIEBMANN, OTTO
(1840–1912)

Otto Liebmann, the German neo-Kantian philosopher, was born at Löwenberg (Lwowek Slaski), Silesia, and became successively *Privatdozent* at Tübingen (1865), extraordinary professor at Strassburg (1872), and professor at Jena (1882). He served as a volunteer during the siege of Paris in 1870 and 1871 and published a memoir of his experiences.

In a *Festschrift* dedicated to Liebmann on his seventieth birthday, various thinkers discussed the aspects of his work that were of particular interest to them. Each interpreted him differently; for example, Bruno Bauch stressed transcendental-methodological aspects, Erich Adickes empirical openness, Wilhelm Windelband critical-metaphysical insight. Such variegated criticism was not without foundation, for Liebmann's thought had many facets and did not evolve so much as oscillate between impulsive outbursts and great restraint, passing from problem to problem.

In his notable early book, *Kant und die Epigonen* (1865), Liebmann swept aside the academic philosophy of his day and preached a return to Immanuel Kant. He simplified Kantian thought and streamlined the post-Kantian systems. The essence of the Kantian revolution, he claimed, was the discovery of the transcendental, which, however, must be freed from the *caput mortuum* of the thing-in-itself. The systematic effort of the great successors of Kant failed because Johann Gottlieb Fichte's Ego, Friedrich Schiller's Absolute, G. W. F. Hegel's Spirit, Johann Friedrich Herbart's "reals," and Arthur Schopenhauer's Will all represent the thing-in-itself, whereas J. F. Fries mistook the transcendental for the psychological. For Liebmann the only reality, immanent in consciousness and sufficient, is experience, which is both empirical reality and transcendental ideality. But could such simplified views be unequivocally developed?

In a subsequent essay, *Über den individuellen Beweis für die Freiheit des Willens* (1866), Liebmann dealt with the freedom of the will, in opposition to Schopenhauer. Are we, it can be asked, on the level of the transcendental or of the individual ego in dealing with this problem? Reexamining the question in 1901 (*Gedanken und Tatsachen,* Vol. II, p. 88), he referred it to the individual.

In *Über den objektiven Anblick* (1869) Liebmann distinguished three factors in perception: the sensitive, the intellectual, and the transcendent. The transcendent factor in perception "is the relationship between an unknown *X* and a likewise unknown *Y,* which appears to us as our body, and from which in turn there spring into our consciousness those sensitive qualities which our intellect transforms, according to a priori laws, into perceptible nature, a phenomenon of the external material world" (p. 153). In this work the thing-in-itself is not eliminated; on the contrary, two things-in-themselves—*X* and *Y*—are admitted.

Liebmann's major works, *Analysis der Wirklichkeit* (1876) and *Gedanken und Tatsachen* (2 vols., 1882–1907), are collections of problems, not only in the critique of knowledge but also in *Naturphilosophie,* psychology, aesthetics, and ethics. In all of these, self-consciousness recognizes its limits; but the resulting agnosticism is superseded by a program of "critical metaphysics."

In this connection Liebmann denounced as a *doktrinäre Fiktion* the neo-Baconian ideal (or idol) of pure experience, itself a notion that Liebmann took from Richard Avenarius and from the evolutionary genetic psychology of Herbert Spencer and others. Every experience and every science, Liebmann claimed, is possible only by means of certain nonempirical premises, such as

the principles of real identity, of the continuity of existence, of constant causality or legality, and of the temporal continuity of becoming, or, in general, by means of fundamental a priori forms or principles, which constitute the organization of human cognitive powers but from whose transcendental validity by no means necessarily follows its transcendent reality.

Liebmann distinguished three types of theories, which seek explanatory principles in the immediate empirical data, in hypotheses by which the phenomena are deduced, or in absolute metaphysical realities. He rejected the first and third, and admitted the hypotheses, if and as long as the facts confirm them. This is true not only of scientific but also of philosophical theories, especially of critical metaphysics as a "strict discussion of human views, human hypotheses on the essence of things." Liebmann concentrated on the theories of science and their metaphysical pronouncements or assumptions. He claimed, for example, that the biological point of view is more than a mere postulation of an as-if; it is a positive affirmation of entelechies. Darwinism abounds with metascientific problems and teleological claims; but not even the transcendental philosopher can escape the problems posed by nature, with its own immanent logic (*Weltlogik*), its dynamic causality that achieves an increase in perfection, even though he knows that every hypothesis and system is a product of the specifically human thinking apparatus. A study of space and time that Liebmann undertook to come to grips with non-Euclidean viewpoints led him to problems that appeared to Windelband as idle fancies.

In dealing with the problem of the multiplicity of subjects, Liebmann developed but did not elaborate upon a distinction between three conceptions of the ego: the metaphysical substrate, an objective never attained by dogmatic metaphysics; the individual ego, a tacit assumption of psychology; and the transcendental ego, a "typical" subject of the intelligence of the human species and a fundamental condition of the empirical world. The problem of psychophysical parallelism led him to postulate a coincidence of natural and logical laws on the metaphysical plane of *natura naturans,* but he did not draw the necessary methodological distinctions to adequately treat this problem.

See also Avenarius, Richard; Darwinism; Determinism and Freedom; Fichte, Johann Gottlieb; German Philosophy; Hegel, Georg Wilhelm Friedrich; Herbart, Johann Friedrich; Kant, Immanuel; Natural Law; Neo-Kantianism; Schiller, Friedrich; Schopenhauer, Arthur; Windelband, Wilhelm.

Bibliography

WORKS BY LIEBMANN

Kant und die Epigonen. Stuttgart: Schober, 1865; 2nd ed., Berlin: Reuther & Reichard, 1912.

Über den individuellen Beweis für die Freiheit des Willens. Stuttgart: Schober, 1866.

Über den objektiven Anblick. Stuttgart, 1869.

Zur Analysis der Wirklichkeit. Strassburg, 1876; 4th enlarged ed., Strassburg: Trübner, 1911.

Über philosophischen Tradition. Strassburg, 1883.

Die Klimax der Theorien. Strassburg, 1884.

Gedanken und Tatsachen, 2 vols. Strassburg, 1882–1901; 2nd ed., Strassburg: Trübner, 1899–1904.

Immanuel Kant. Strassburg, 1904.

WORKS ON LIEBMANN

Meyer, Adolf. *Über Liebmanns Erkenntnislehre und ihr Verhältniss zu Kant.* Jena, Germany, 1916. Dissertation.

Zum 70. Geburtstag Otto Liebmanns. A Festschrift in *Kantstudien* 15 (1910). Contains works by Adickes, Bauch, Hans Driesch, Windelband, and others.

Mariano Campo (1967)
Translated by Robert M. Connolly

LIFE, MEANING AND VALUE OF

To the questions "Is human life ever worthwhile?" and "Does (or can) human life have any meaning?" many religious thinkers have offered affirmative answers with the proviso that these answers would not be justified unless two of the basic propositions of most Western religions were true—that human life is part of a divinely ordained cosmic scheme and that after death at least some human beings will be rewarded with eternal bliss. Thus, commenting on Bertrand Russell's statement that not only must each individual human life come to an end but that life in general will eventually die out, C. H. D. Clark contrasts this "doctrine of despair" with the beauty of the Christian scheme. "If we are asked to believe that all our striving is without final consequence," then "life is meaningless and it scarcely matters how we live if all will end in the dust of death." According to Christianity, on the other hand, "each action has vital significance." Clark assures us that "God's grand design is life eternal for those who walk in the steps of Christ. Here is the one grand incentive to good living.… As life is seen to have purpose and meaning, men find release from despair and the fear of death" (*Christianity and Bertrand Russell,* p. 30). In a similar vein, the Jewish existentialist Emil Fackenheim claims that "whatever meaning life acquires" is derived from the encounter between God and man. The meaning thus

conferred upon human life "cannot be understood in terms of some finite human purpose, supposedly more ultimate than the meeting itself. For what could be more ultimate than the Presence of God?" It is true that God is not always "near," but "times of Divine farness" are by no means devoid of meaning. "Times of Divine nearness do not light up themselves alone. Their meaning extends over all of life." There is a "dialectic between Divine nearness and Divine farness," and it points to "an eschatological future in which it is overcome" ("Judaism and the Meaning of Life").

Among unbelievers not a few maintain that life can be worthwhile and have meaning in some humanly important sense even if the religious world view is rejected. Others, however, agree with the religious theorists that our two questions must be given negative answers if there is no God and if death means personal annihilation. Having rejected the claims of religion, they therefore conclude that life is not worthwhile and that it is devoid of meaning. These writers, to whom we shall refer here as "pessimists," do not present their judgments as being merely expressions of certain moods or feelings but as conclusions that are in some sense objectively warranted. They offer reasons for their conclusions and imply that anybody reaching a contradictory conclusion is mistaken or irrational. Most pessimists do not make any clear separation between the statements that life is not worthwhile and that life is without meaning. They usually speak of the "futility" or the "vanity" of life, and presumably they mean by this both that life is not worth living and that it has no meaning. For the time being we, too, shall treat these statements as if they were equivalent. However, later we shall see that in certain contexts it becomes important to distinguish between them.

Our main concern in this entry will be to appraise pessimism as just defined. We shall not discuss either the question whether life is part of a divinely ordained plan or the question whether we survive our bodily death. Our question will be whether the pessimistic conclusions are justified if belief in God and immortality are rejected.

SCHOPENHAUER'S ARGUMENTS

Let us begin with a study of the arguments offered by the pessimists, remembering that many of these are indirectly endorsed by religious apologists. The most systematic and probably the most influential, though in fact not the gloomiest, of the pessimists was Arthur Schopenhauer. The world, he wrote, is something that ought not to exist: The truth is that "we have not to rejoice but rather to mourn at the existence of the world; that its non-existence would be preferable to its existence; that it is something which ought not to be." It is absurd to speak of life as a gift, as so many philosophers and thoughtless people have done. "It is evident that everyone would have declined such a gift if he could have seen it and tested it beforehand." To those who assure us that life is only a lesson, we are entitled to reply: "For this very reason I wish I had been left in the peace of the all-sufficient nothing, where I would have no need of lessons or of anything else" (*The World as Will and Idea,* Vol. III, p. 390).

Schopenhauer offers numerous arguments for his conclusion. Some of these are purely metaphysical and are based on his particular system. Others, however, are of a more empirical character and are logically independent of his brand of metaphysical voluntarism. Happiness, according to Schopenhauer, is unobtainable for the vast majority of humankind. "Everything in life shows that earthly happiness is destined to be frustrated or recognized as illusion." People either fail to achieve the ends they are striving for or else they do achieve them only to find them grossly disappointing. But as soon as a man discovers that a particular goal was not really worth pursuing, his eye is set on a new one and the same illusory quest begins all over again. Happiness, accordingly, always lies in the future or in the past, and "the present may be compared to a small dark cloud which the wind drives over the sunny plain: before and behind it all is bright, only it itself always casts a shadow. The present is therefore always insufficient; but the future is uncertain, and the past is irrevocable" (ibid., p. 383). Men in general, except for those sufficiently rational to become totally resigned, are constantly deluded—"now by hope, now by what was hoped for." They are taken in by "the enchantment of distance," which shows them "paradises." These paradises, however, vanish like "optical illusions when we have allowed ourselves to be mocked by them." The "fearful envy" excited in most men by the thought that somebody else is genuinely happy shows how unhappy they really are, whatever they pretend to others or to themselves. It is only "because they feel themselves unhappy" that "men cannot endure the sight of one whom they imagine happy."

On occasions Schopenhauer is ready to concede that some few human beings really do achieve "comparative" happiness, but this is not of any great consequence. For aside from being "rare exceptions," these happy people are really like "decoy birds"—they represent a possibility that must exist in order to lure the rest of humankind into a false sense of hope. Moreover, happiness, insofar as it exists at all, is a purely "negative" reality. We do not

become aware of the greatest blessings of life—health, youth, and freedom—until we have lost them. What is called pleasure or satisfaction is merely the absence of craving or pain. But craving and pain are positive. As for the few happy days of our life—if there are any—we notice them only "after they have given place to unhappy ones."

Schopenhauer not infrequently lapsed from his doctrine of the "negative" nature of happiness and pleasure into the more common view that their status is just as "positive" as that of unhappiness and pain. But he had additional arguments that do not in any way depend on the theory that happiness and pleasure are negative. Perhaps the most important of these is the argument from the "perishableness" of all good things and the ultimate extinction of all our hopes and achievements in death. All our pleasures and joys "disappear in our hands, and we afterwards ask astonished where they have gone." Moreover, a joy that no longer exists does not "count"—it counts as little as if it had never been experienced at all:

> That which *has been* exists no more; it exists as little as that which has *never* been. But of everything that exists you may say, in the next moment, that it has been. Hence something of great importance in our past is inferior to something of little importance in our present, in that the latter is a *reality,* and related to the former as something to nothing. ("The Vanity of Existence," in *The Will to Live,* p. 229)

Some people have inferred from this that the enjoyment of the present should be "the supreme object of life." This is fallacious; for "that which in the next moment exists no more, and vanishes utterly, like a dream, can never be worth a serious effort."

The final "judgment of nature" is destruction by death. This is "the last proof" that life is a "false path," that all man's wishing is "a perversity," and that "nothing at all is worth our striving, our efforts and struggles." The conclusion is inescapable: "All good things are vanity, the world in all its ends bankrupt, and life a business which does not cover its expenses" (*The World as Will and Idea,* Vol. III, p. 383).

THE POINTLESSNESS OF IT ALL

Some of Schopenhauer's arguments can probably be dismissed as the fantasies of a lonely and embittered man who was filled with contempt for humankind and who was singularly incapable of either love or friendship. His own misery, it may be plausibly said, made Schopenhauer overestimate the unhappiness of human beings. It is frequently, but not universally, true that what is hoped for is found disappointing when it is attained, and while "fearful envy" of other people's successes is common enough, real sympathy and generosity are not quite so rare as Schopenhauer made them out to be. Furthermore, his doctrine that pleasure is negative while pain is positive, insofar as one can attach any clear meaning to it, seems glaringly false. To this it should be added, however, that some of Schopenhauer's arguments are far from idiosyncratic and that substantially the same conclusions have been endorsed by men who were neither lonely nor embittered and who did not, as far as one can judge, lack the gift of love or friendship.

DARROW. Clarence Darrow, one of the most compassionate men who ever lived, also concluded that life was an "awful joke." Like Schopenhauer, Darrow offered as one of his reasons the apparent aimlessness of all that happens. "This weary old world goes on, begetting, with birth and with living and with death," he remarked in his moving plea for the boy-murderers Richard Loeb and Nathan Leopold, "and all of it is blind from the beginning to the end" (*Clarence Darrow—Attorney for the Damned,* edited by A. Weinberg, New York, 1957). Elsewhere he wrote: "Life is like a ship on the sea, tossed by every wave and by every wind; a ship headed for no port and no harbor, with no rudder, no compass, no pilot; simply floating for a time, then lost in the waves" ("Is Life Worth Living?," p. 43). In addition to the aimlessness of life and the universe, there is the fact of death. "I love my friends," wrote Darrow, "but they all must come to a tragic end." Death is more terrible the more one is attached to things in the world. Life, he concludes, is "not worth while," and he adds (somewhat inconsistently, in view of what he had said earlier) that "it is an unpleasant interruption of nothing, and the best thing you can say of it is that it does not last long" ("Is the Human Race Getting Anywhere?," p. 53).

TOLSTOY. Lev Tolstoy, unlike Darrow, eventually came to believe in Christianity, or at least in his own idiosyncratic version of Christianity, but for a number of years the only position for which he could see any rational justification was an extreme form of pessimism. During that period (and there is reason to believe that in spite of his later protestations to the contrary, his feelings on this subject never basically changed) Tolstoy was utterly overwhelmed by the thought of his death and the death of those he cared for and, generally, by the transitory nature of all human achievements. "Today or tomorrow," he

wrote in "A Confession," "sickness and death will come to those I love or to me; nothing will remain but stench and worms. Sooner or later my affairs, whatever they may be, will be forgotten, and I shall not exist. Then why go on making any effort?" Tolstoy likened the fate of man to that of the traveler in the Eastern tale who, pursued by an enraged beast, seeks refuge in a dry well. At the bottom of the well he sees a dragon that has opened its jaws to swallow him. To escape the enraged beast above and the dragon below, he holds onto a twig that is growing in a crack in the well. As he looks around he notices that two mice are gnawing at the stem of the twig. He realizes that very soon the twig will snap and he will fall to his doom, but at the same time he sees some drops of honey on the leaves of the branch and reaches out with his tongue to lick them. "So I too clung to the twig of life, knowing that the dragon of death was inevitably awaiting me, ready to tear me to pieces.… I tried to lick the honey which formerly consoled me, but the honey no longer gave me pleasure.… I only saw the unescapable dragon and the mice, and I could not tear my gaze from them. And this is not a fable but the real unanswerable truth."

These considerations, according to Tolstoy, inevitably lead to the conclusion that life is a "stupid fraud," that no "reasonable meaning" can be given to a single action or to a whole life. To the questions "What is it for?" "What then?," "Why should I live?" the answer is "Nothing can come of it," "Nothing is worth doing," "Life is not worthwhile."

What ways out are available to a human being who finds himself in this "terrible position"? Judging by the conduct of the people he observed, Tolstoy relates that he could see only four possible "solutions." The first is the way of ignorance. People who adopt this solution (chiefly women and very young and very dull people) have simply not or not yet faced the questions that were tormenting him. Once a person has fully realized what death means, this solution is not available to him. The second way is that of "Epicureanism," which consists in admitting the "hopelessness of life" but seizing as many of life's pleasures as possible while they are within reach. It consists in "disregarding the dragon and the mice and licking the honey in the best way, especially if much of it is around." This, Tolstoy adds, is the solution adopted by the majority of the people belonging to his "circle," by which he presumably means the well-to-do intellectuals of his day. Tolstoy rejects this solution because the vast majority of human beings are not well-to-do and hence have little or no honey at their disposal and also because it is a matter of accident whether one is among those who have honey or those who have not. Moreover, Tolstoy observes, it requires a special "moral dullness," which he himself lacked, to enjoy the honey while knowing the truth about death and the deprivations of the great majority of men. The third solution is suicide. Tolstoy calls this the way of "strength and energy." It is chosen by a few "exceptionally strong and consistent people." After they realize that "it is better to be dead than to be alive, and that it is best of all not to exist," they promptly end the whole "stupid joke." The means for ending it are readily at hand for everybody, but most people are too cowardly or too irrational to avail themselves of them. Finally, there is the way of "weakness." This consists in seeing the dreadful truth and clinging to life nevertheless. People of this kind lack the strength to act rationally and Tolstoy adds that he belonged to this last category.

STRENGTHS OF THE PESSIMIST POSITION. Is it possible for somebody who shares the pessimists' rejection of religion to reach different conclusions without being plainly irrational? Whatever reply may be possible, any intelligent and realistic person would surely have to concede that there is much truth in the pessimists' claims. That few people achieve real and lasting happiness, that the joys of life (where there are any) pass away much too soon, that totally unpredictable events frequently upset the best intentions and wreck the noblest plans—this and much more along the same lines is surely undeniable. Although one should not dogmatize that there will be no significant improvements in the future, the fate of past revolutions, undertaken to rid man of some of his apparently avoidable suffering, does not inspire great hope. The thought of death, too, even in those who are not so overwhelmed by it as Tolstoy, can be quite unendurable. Moreover, to many who have reflected on the implications of physical theory it seems plain that because of the constant increase of entropy in the universe all life anywhere will eventually die out. Forebodings of this kind moved Bertrand Russell to write his famous essay "A Free Man's Worship," in which he concluded that "all the labors of the ages, all the devotion, all the inspiration, all the noonday brightness of human genius, are destined to extinction in the vast death of the solar system, and the whole temple of man's achievement must inevitably be buried beneath the debris of a universe in ruins." Similarly, Wilhelm Ostwald observed that "in the longest run the sum of all human endeavor has no recognizable significance." Although it is disputed whether physical theory really has such gloomy implications, it would perhaps be wisest to assume that the position endorsed by Russell and Ostwald is well-founded.

COMPARATIVE VALUE JUDGMENTS ABOUT LIFE AND DEATH

Granting the strong points in the pessimists' claims, it is still possible to detect certain confusions and dubious inferences in their arguments. To begin with, there is a very obvious inconsistency in the way writers like Darrow and Tolstoy arrive at the conclusion that death is better than life. They begin by telling us that death is something terrible because it terminates the possibility of any of the experiences we value. From this they infer that nothing is really worth doing and that death is better than life. Ignoring for the moment the claim that in view of our inevitable death nothing is "worth doing," there very plainly seems to be an inconsistency in first judging death to be such a horrible evil and in asserting later on that death is better than life. Why was death originally judged to be an evil? Surely because it is the termination of life. And if something, *y,* is bad because it is the termination of something, *x,* this can be so only if *x* is good or has positive value. If *x* were not good, the termination of *x* would not be bad. One cannot consistently have it both ways.

To this it may be answered that life did have positive value prior to one's realization of death but that once a person has become aware of the inevitability of his destruction life becomes unbearable and that this is the real issue. This point of view is well expressed in the following exchange between Cassius and Brutus in William Shakespeare's *Julius Caesar* (III.i.102–105):

> CASSIUS. Why he that cuts off twenty years of life—
> Cuts off so many years of fearing death.
>
> BRUTUS. Grant that, and then is death a benefit:
> So are we Caesar's friends that have abridged
> His time of fearing death.

There is a very simple reply to this argument. Granting that some people after once realizing their doom cannot banish the thought of it from their minds, so much so that it interferes with all their other activities, this is neither inevitable nor at all common. It is, on the contrary, in the opinion of all except some existentialists, morbid and pathological. The realization that one will die does not in the case of most people prevent them from engaging in activities which they regard as valuable or from enjoying the things they used to enjoy. To be told that one is not living "authentically" if one does not brood about death day and night is simply to be insulted gratuitously. A person who knows that his talents are not as great as he would wish or that he is not as handsome as he would have liked to be is not usually judged to live "inauthentically," but on the contrary to be sensible if he does not constantly brood about his limitations and shortcomings and uses whatever talents he does possess to maximum advantage.

There is another and more basic objection to the claim that death is better than life. This objection applies equally to the claim that while death is better than life it would be better still not to have been born in the first place and to the judgment that life is better than death. It should be remembered that we are here concerned with such pronouncements when they are intended not merely as the expression of certain moods but as statements that are in some sense true or objectively warranted. It may be argued that a value comparison—any judgment to the effect that *A* is better or worse than *B* or as good as *B*—makes sense only if *both A* and *B* are, in the relevant respect, in principle open to inspection. If somebody says, for example, that Elizabeth Taylor is a better actress than Betty Grable, this seems quite intelligible. Or, again, if it is said that life for the Jews is better in the United States than it was in Germany under the Nazis, this also seems readily intelligible. In such cases the terms of the comparison are observable or at any rate describable. These conditions are fulfilled in some cases when value comparisons are made between life and death, but they are not fulfilled in the kind of case with which Tolstoy and the pessimists are concerned. If the conception of an afterlife is intelligible, then it would make sense for a believer or for somebody who has not made up his mind to say such things as "Death cannot be worse than this life" or "I wonder if it will be any better for me after I am dead." Achilles, in the *Iliad,* was not making a senseless comparison when he exclaimed that he would rather act

> … as a serf of another,
> A man of little possessions, with scanty means of subsistence,
> Than rule as a ghostly monarch the ghosts of all the departed.

Again, the survivors can meaningfully say about a deceased individual "It is better (for the world) that he is dead" or the opposite. For the person himself, however, if there is no afterlife, death is not a possible object of observation or experience, and statements by him that his own life is better than, as good as, or worse than his own death, unless they are intended to be no more than expressions of certain wishes or moods, must be dismissed as senseless. At first sight the contention that in the circumstances under discussion value comparisons between life and death are senseless may seem implausible because of the widespread tendency to think of death as a shadowy kind

of life—as sleep, rest, or some kind of homecoming. Such "descriptions" may be admirable as poetry or consolation, but taken literally they are simply false.

IRRELEVANCE OF THE DISTANT FUTURE

These considerations do not, however, carry us very far. They do not show either that life is worth living or that it "has meaning." Before tackling these problems directly, something should perhaps be said about the curious and totally arbitrary preference of the future to the present, to which writers such as Tolstoy and Darrow are committed without realizing it. Darrow implies that life would not be "futile" if it were not an endless cycle of the same kind of activities and if instead it were like a journey toward a destination. Tolstoy clearly implies that life would be worthwhile, that some of our actions at least would have a "reasonable meaning," if the present life were followed by eternal bliss. Presumably, what would make life no longer futile as far as Darrow is concerned is some feature of the destination, not merely the fact that it is a destination; and what would make life worthwhile in Tolstoy's opinion is not merely the eternity of the next life but the "bliss" that it would confer—eternal misery and torture would hardly do. About the bliss in the next life, if there is such a next life, Tolstoy shows no inclination to ask "What for?" or "So what?" But if bliss in the next life is not in need of any further justification, why should any bliss that there might be in the present life need justification?

THE LOGIC OF VALUE JUDGMENTS. Many of the pessimists appear to be confused about the logic of value judgments. It makes sense for a person to ask about something "Is it really worthwhile?" or "Is it really worth the trouble?" if he does not regard it as intrinsically valuable or if he is weighing it against another good with which it may be in conflict. It does not make sense to ask such a question about something he regards as valuable in its own right and where there is no conflict with the attainment of any other good. (This observation, it should be noted, is quite independent of what view one takes of the logical status of intrinsic value judgments.) A person driving to the beach on a crowded Sunday, may, upon finally getting there, reflect on whether the trip was really worthwhile. Or, after undertaking a series of medical treatments, somebody may ask whether it was worth the time and the money involved. Such questions make sense because the discomforts of a car ride and the time and money spent on medical treatments are not usually judged to be valuable for their own sake. Again, a woman who has given up a career as a physician in order to raise a family may ask herself whether it was worthwhile, and in this case the question would make sense not because she regards the raising of a family as no more than a means, but because she is weighing it against another good. However, if somebody is very happy, for any number of reasons—because he is in love, because he won the Nobel Prize, because his child recovered from a serious illness—and if this happiness does not prevent him from doing or experiencing anything else he regards as valuable, it would not occur to him to ask "Is it worthwhile?" Indeed, this question would be incomprehensible to him, just as Tolstoy himself would presumably not have known what to make of the question had it been raised about the bliss in the hereafter.

It is worth recalling here that we live not in the distant future but in the present and also, in a sense, in the relatively near future. To bring the subject down to earth, let us consider some everyday occurrences: A man with a toothache goes to a dentist, and the dentist helps him so that the toothache disappears. A man is falsely accused of a crime and is faced with the possibility of a severe sentence as well as with the loss of his reputation; with the help of a devoted attorney his innocence is established, and he is acquitted. It is true that a hundred years later all of the participants in these events will be dead and none of them will then be able to enjoy the fruits of any of the efforts involved. But this most emphatically does not imply that the dentist's efforts were not worthwhile or that the attorney's work was not worth doing. To bring in considerations of what will or will not happen in the remote future is, in such and many other though certainly not in all human situations, totally irrelevant. Not only is the finality of death irrelevant here; equally irrelevant are the facts, if they are facts, that life is an endless cycle of the same kind of activities and that the history of the universe is not a drama with a happy ending.

This is, incidentally, also the answer to religious apologists like C. H. D. Clark who maintain that all striving is pointless if it is "without final consequence" and that "it scarcely matters how we live if all will end in the dust of death." Striving is not pointless if it achieves what it is intended to achieve even if it is without final consequence, and it matters a great deal how we live if we have certain standards and goals, although we cannot avoid "the dust of death."

THE VANISHED PAST. In asserting the worthlessness of life Schopenhauer remarked that "what has been exists as little as what has never been" and that "something of great

importance now past is inferior to something of little importance now present." Several comments are in order here. To begin with, if Schopenhauer is right, it must work both ways: If only the present counts, then past sorrows no less than past pleasures do not "count." Furthermore, the question whether "something of great importance now past is inferior to something of little importance now present" is not, as Schopenhauer supposed, a straightforward question of fact but rather one of valuation, and different answers, none of which can be said to be mistaken, will be given by different people according to their circumstances and interests. Viktor Frankl, the founder of "logotherapy," has compared the pessimist to a man who observes, with fear and sadness, how his wall calendar grows thinner and thinner as he removes a sheet from it every day. The kind of person whom Frankl admires, on the other hand, "files each successive leaf neatly away with its predecessors" and reflects "with pride and joy" on all the richness represented by the leaves removed from the calendar. Such a person will not in old age envy the young. "'No, thank you,' he will think. 'Instead of possibilities, I have realities in my past'" (*Man's Search for Meaning,* pp. 192–193).

This passage is quoted not because it contains any great wisdom but because it illustrates that we are concerned here not with judgments of fact but with value judgments and that Schopenhauer's is not the only one that is possible. Nevertheless, his remarks are, perhaps, a healthy antidote to the cheap consolation and the attempts to cover up deep and inevitable misery that are the stock in trade of a great deal of popular psychology. Although Schopenhauer's judgments about the inferior value of the past cannot be treated as objectively true propositions, they express only too well what a great many human beings are bound to feel on certain occasions. To a man dying of cancer it is small consolation to reflect that there was a time when he was happy and flourishing; and while there are undoubtedly some old people who do not envy the young, it may be suspected that more often the kind of talk advocated by the prophets of positive thinking is a mask for envy and a defense against exceedingly painful feelings of regret and helplessness in the face of aging and death and the now-unalterable past.

THE MEANINGS OF THE "MEANING OF LIFE"

Let us now turn to the question whether, given the rejection of belief in God and immortality, life can nevertheless have any "meaning" or "significance." Kurt Baier has called attention to two very different senses in which people use these expressions and to the confusions that result when they are not kept apart. Sometimes when a person asks whether life has any meaning, what he wants to know is whether there is a superhuman intelligence that fashioned human beings along with other objects in the world to serve some end—whether their role is perhaps analogous to the part of an instrument (or its player) in a symphony. People who ask whether history has a meaning often use the word in the same sense. When Macbeth exclaimed that life "is a tale/Told by an idiot, full of sound and fury,/Signifying nothing," he was answering this cosmic question in the negative. His point evidently was not that human life is part of a scheme designed by a superhuman idiot but that it is not part of any design. Similarly, when Fred Hoyle, in his book *The Nature of the Universe* (rev. ed., New York, 1960), turns to what he calls "the deeper issues" and remarks that we find ourselves in a "dreadful situation" in which there is "scarcely a clue as to whether our existence has any real significance," he is using the word *significance* in this cosmic sense.

On the other hand, when we ask whether a particular person's life has or had any meaning, we are usually concerned not with cosmic issues but with the question whether certain purposes are to be found in his life. Thus, most of us would say without hesitation that a person's life had meaning if we knew that he devoted himself to a cause (such as the spread of Christianity or communism or the reform of mental institutions), or we would at least be ready to say that it "acquired meaning" once he became sufficiently attached to his cause. Whether we approve of what they did or not, most of us would be ready to admit—to take some random examples—that Dorothea Dix, Louis Pasteur, V. I. Lenin, Margaret Sanger, Anthony Comstock, and Winston Churchill led meaningful lives. We seem to mean two things in characterizing such lives as meaningful: We assert, first, that the life in question had some dominant, overall goal or goals that gave direction to a great many of the individual's actions and, second, that these actions and possibly others not immediately related to the overriding goal were performed with a special zest that was not present before the person became attached to his goal or that would not have been present if there had been no such goal in his life.

It is not necessary, however, that a person should be devoted to a cause, in the sense just indicated, before we call his life meaningful. It is sufficient that he should have some attachments that are not too shallow. This last expression is of course rather vague, but so is the use of

the word *meaning* when applied to human lives. Since the depth or shallowness of an attachment is a matter of degree, it makes perfectly good sense to speak of degrees of meaning in this context. Thus, C. G. Jung writes that in the lives of his patients there never was "sufficient meaning" (*Memories, Dreams, Reflections,* New York and Toronto, 1963, p. 140). There is nothing odd in such a locution, and there is equally nothing odd in saying about a man who has made a partial recovery from a deep depression that there is now again "some" meaning in his life.

Although frequently when people say about somebody that his life has or had meaning, they evidently regard this as a good thing, this is not invariably the case. One might express this point in the following way: Saying that attachment to a certain goal has made a man's life meaningful is not tantamount to saying that the acts to which the goal has given direction are of positive value. A man might himself observe—and there would be nothing logically odd about it—"As long as I was a convinced Nazi (or communist or Christian or whatever) my life had meaning, my acts had a zest with which I have not been able to invest them since, and yet most of my actions were extremely harmful." Even while fully devoted to his cause or goal the person need not, and frequently does not, regard it as intrinsically valuable. If challenged he will usually justify the attachment to his goal by reference to more fundamental value judgments. Thus, somebody devoted to communism or to medical research or to the dissemination of birth-control information will in all likelihood justify his devotion in terms of the production of happiness and the reduction of suffering, and somebody devoted to Christianity will probably justify his devotion by reference to the will of God.

Let us refer to the first of the two senses we have been discussing as the "cosmic" sense and to the second as the "terrestrial" sense. (These are by no means the only senses in which philosophers and others have used the word *meaning* when they have spoken of the meaning or meaninglessness of life, but for our purposes it is sufficient to take account of these two senses.) Now if the theory of cosmic design is rejected it immediately follows that human life has no meaning in the first or cosmic sense. It does not follow in the least, however, that a particular human life is meaningless in the second, or terrestrial, sense. This conclusion has been very clearly summarized by Baier: "Your life or mine may or may not have meaning (in one sense)," he writes, "even if life as such has none (in the other).… The Christian view guarantees a meaning (in one sense) to every life, the scientific view [what we have simply been calling the unbeliever's position] does not in any sense" (*The Meaning of Life,* p. 28). In the terrestrial sense it will be an open question whether an individual's life has meaning or not, to be decided by the particular circumstances of his existence. It may indeed be the case that once a person comes to believe that life has no meaning in the cosmic sense his attachment to terrestrial goals will be undermined to such an extent that his life will cease to be meaningful in the other sense as well. However, it seems very plain that this is by no means what invariably happens, and even if it did invariably happen the meaninglessness of a given person's life in the terrestrial sense would not logically follow from the fact, if it is a fact, that life is meaningless in the cosmic sense.

This is perhaps the place to add a few words of protest against the rhetorical exaggerations of certain theological writers. Fackenheim's statement, quoted earlier, that "whatever meaning life acquires, it derives from the encounter between God and man" is typical of many theological pronouncements. Statements of this kind are objectionable on several grounds. Let us assume that there is a God and that meetings between God and certain human beings do take place; let us also grant that activities commanded by God in these meetings "acquire meaning" by being or becoming means to the end of pleasing or obeying God. Granting all this, it does not follow that obedience of God is the only possible unifying goal. It would be preposterous to maintain that the lives of all unbelievers have been lacking in such goals and almost as preposterous to maintain that the lives of believers never contain unifying goals other than obedience of God. There have been devout men who were also attached to the advance of science, to the practice of medicine, or to social reform and who regarded these ends as worth pursuing independently of any divine commandments. Furthermore, there is really no good reason to grant that the life of a particular person becomes meaningful in the terrestrial sense just because human life in general has meaning in the cosmic sense. If a superhuman being has a plan in which I am included, this fact will make (or help to make) my life meaningful in the terrestrial sense only if I know the plan and approve of it and of my place in it, so that working toward the realization of the plan gives direction to my actions.

IS HUMAN LIFE EVER WORTHWHILE?

Let us now turn to the question of whether life is ever worth living. This also appears to be denied by the pessimists when they speak of the vanity or the futility of human life. We shall see that in a sense it cannot be estab-

lished that the pessimists are "mistaken," but it is also quite easy to show that in at least two senses that seem to be of importance to many people, human lives frequently are worth living. To this end, let us consider under what circumstances a person is likely to raise the question "Is my life (still) worthwhile?" and what is liable to provoke somebody into making a statement like "My life has ceased to be worth living." We saw in an earlier section that when we say of certain acts, such as the efforts of a dentist or a lawyer, that they were worthwhile we are claiming that they achieved certain goals. Something similar seems to be involved when we say that a person's life is (still) worthwhile or worth living. We seem to be making two assertions: First, that the person has some goals (other than merely to be dead or to have his pains eased) which do not seem to him to be trivial and, second, that there is some genuine possibility that he will attain these goals. These observations are confirmed by various systematic studies of people who contemplated suicide, of others who unsuccessfully attempted suicide, and of situations in which people did commit suicide. When the subjects of these studies declared that their lives were no longer worth living they generally meant either that there was nothing left in their lives about which they seriously cared or that there was no real likelihood of attaining any of the goals that mattered to them. It should be noted that in this sense an individual may well be mistaken in his assertion that his life is or is not worthwhile any longer: He may, for example, mistake a temporary indisposition for a more permanent loss of interest, or, more likely, he may falsely estimate his chances of achieving the ends he wishes to attain.

DIFFERENT SENSES OF "WORTHWHILE." According to the account given so far, one is saying much the same thing in declaring a life to be worthwhile and in asserting that it has meaning in the "terrestrial" sense of the word. There is, however, an interesting difference. When we say that a person's life has meaning (in the terrestrial sense) we are not committed to the claim that the goal or goals to which he is devoted have any positive value. (This is a slight oversimplification, assuming greater uniformity in the use of "meaning of life" than actually exists, but it will not seriously affect any of the controversial issues discussed here.) The question "As long as his life was dedicated to the spread of communism it had meaning *to him,* but was it really meaningful?" seems to be senseless. We are inclined to say, "If his life had meaning to him, then it had meaning—that's all there is to it." We are not inclined (or we are much less inclined) to say something of this kind when we speak of the worth of a person's life. We might say—for example, of someone like Adolf Eichmann—"While he was carrying out the extermination program, his life *seemed* worthwhile to him, but since his goal was so horrible, his life *was not* worthwhile." One might perhaps distinguish between a "subjective" and an "objective" sense of "worthwhile." In the subjective sense, saying that a person's life is worthwhile simply means that he is attached to some goals that he does not consider trivial and that these goals are attainable for him. In declaring that somebody's life is worthwhile in the objective sense, one is saying that he is attached to certain goals which are both attainable and of positive value.

It may be held that unless one accepts some kind of rationalist or intuitionist view of fundamental value judgments one would have to conclude that in the objective sense of "worthwhile" no human life (and indeed no human action) could ever be shown to be worthwhile. There is no need to enter here into a discussion of any controversial questions about the logical status of fundamental value judgments. But it may be pointed out that somebody who favors a subjectivist or emotivist account can quite consistently allow for the distinction between ends that only seem to have positive value and those that really do. To mention just one way in which this could be done: One may distinguish between ends that would be approved by rational and sympathetic human beings and those that do not carry such an endorsement. One may then argue that when we condemn such a life as Eichmann's as not being worthwhile we mean not that the ends to which he devoted himself possess some nonnatural characteristic of badness but that no rational or sympathetic person would approve of them.

THE PESSIMISTS' SPECIAL STANDARDS. The unexciting conclusion of this discussion is that some human lives are at certain times not worthwhile in either of the two senses we have distinguished, that some are worthwhile in the subjective but not in the objective sense, some in the objective but not in the subjective sense, and some are worthwhile in both senses. The unexcitingness of this conclusion is not a reason for rejecting it, but some readers may question whether it meets the challenge of the pessimists. The pessimist, it may be countered, surely does not deny the plain fact that human beings are on occasions attached to goals which do not seem to them trivial, and it is also not essential to his position to deny (and most pessimists do not in fact deny) that these goals are sometimes attainable. The pessimist may even allow that in a superficial ("immediate") sense the goals which people try to achieve are of positive value, but he would add that because our lives are not followed by eternal bliss

they are not "really" or "ultimately" worthwhile. If this is so, then the situation may be characterized by saying that the ordinary man and the pessimist do not mean the same by "worthwhile," or that they do mean the same in that both use it as a positive value expression but that their standards are different: The standards of the pessimist are very much more demanding than those of most ordinary people.

Anybody who agrees that death is final will have to concede that the pessimist is not mistaken in his contention that judged by his standards, life is never worthwhile. However, the pessimist is mistaken if he concludes, as frequently happens, that life is not worthwhile by ordinary standards because it is not worthwhile by his standards. Furthermore, setting aside the objection mentioned earlier (that there is something arbitrary about maintaining that eternal bliss makes life worthwhile but not allowing this role to bliss in the present life), one may justifiably ask why one should abandon ordinary standards in favor of those of the pessimist. Ordinarily, when somebody changes standards (for example, when a school raises or lowers its standards of admission) such a change can be supported by reasons.

But how can the pessimist justify his special standards? It should be pointed out here that our ordinary standards do something for us which the pessimist's standards do not: They guide our choices, and as long as we live we can hardly help making choices. It is true that in one type of situation the pessimist's standards also afford guidance—namely, in deciding whether to go on living. It is notorious, however, that whether or not they are, by their own standards, rational in this, most pessimists do not commit suicide. They are then faced with much the same choices as other people. In these situations their own demanding standards are of no use, and in fact they avail themselves of the ordinary standards. Schopenhauer, for example, believed that if he had hidden his antireligious views he would have had no difficulty in obtaining an academic appointment and other worldly honors. He may have been mistaken in this belief, but in any event his actions indicate that he regarded intellectual honesty as worthwhile in a sense in which worldly honors were not. Again, when Darrow had the choice between continuing as counsel for the Chicago and North Western Railway and taking on the defense of Eugene V. Debs and his harassed and persecuted American Railway Union, he did not hesitate to choose the latter, apparently regarding it as worthwhile to go to the assistance of the suppressed and not worthwhile to aid the suppressor. In other words, although no human action is worthwhile, some human actions and presumably some human lives are less unworthwhile than others.

IS THE UNIVERSE BETTER WITH HUMAN LIFE THAN WITHOUT IT?

We have not—at least not explicitly—discussed the claims of Schopenhauer, Eduard von Hartmann, and other pessimists that the nonexistence of the world would be better than its existence, by which they mean that a world without human life would be better than one with it.

ARGUMENTS OF A PHENOMENOLOGIST. Some writers do not think that life can be shown to have meaning in any philosophically significant sense unless an affirmative answer to this question can be justified. Thus, in his booklet *Der Sinn unseres Daseins* the German phenomenologist Hans Reiner distinguishes between the everyday question about what he calls the "need-conditioned" meaning of life, which arises only for a person who is already in existence and has certain needs and desires, and the question about the meaning of human life in general. The latter question arises in concrete form when a responsible person is faced with the *Zeugungsproblem*—the question whether he should bring a child into the world. Reiner allows that a person's life has meaning in the former or "merely subjective" sense as long as his ordinary goals (chiefly his desire for happiness) are attained. This, however, does not mean that his life has an "objective" or "existential" (*seinshaft*) meaning—a significance or meaning that "attaches to life as such" and which, unlike the need-conditioned meaning, cannot be destroyed by any accident of fate. The philosopher, according to Reiner, is primarily concerned with the question of whether life has meaning in this objective or existential sense. "Our search for the meaning of our life," Reiner writes, "is identical with the search for a logically compelling reason (*einen einsichtigen Grund*) why it is better for us to exist than not to exist" (*Der Sinn unseres Daseins,* p. 27). Again, the real question is "whether it is better that mankind should exist than that there should be a world without any human life" (p. 31). It may be questioned whether this is what anybody normally means when he asks whether life has any meaning, but Reiner certainly addresses himself to one of the questions raised by Schopenhauer and other pessimists that ought to be discussed here.

Reiner believes that he can provide a "logically compelling reason" why a world with human life is better than one without it. He begins by pointing out that men differ

from animals by being, among other things, "moral individuals." To be a moral individual is to be part of the human community and to be actively concerned in the life of other human beings. It is indeed undeniable that people frequently fail to bring about the ends of morally inspired acts or wishes, but phenomenological analysis discloses that "the real moral value and meaning" of an act does not depend on the attainment of the "external goal." As Immanuel Kant correctly pointed out, the decisive factor is "the good will," the moral intent or attitude. It is here that we find the existential meaning of life: "Since that which is morally good contains its meaning and value within itself, it follows that it is intrinsically worth while. The existence of what is morally good is therefore better than its non-existence." (*Der Sinn unseres Daseins,* pp. 54–55). But the existence of what is morally good is essentially connected with the existence of free moral individuals, and hence it follows that the existence of human beings as moral agents is better than their nonexistence.

Unlike happiness, which constitutes the meaning of life in the everyday or need-conditioned sense, the morally good does not depend on the accidents of life. It is not within a person's power to be happy, but it is "essentially" (*grundsätzlich*) in everybody's power to do what is good. Furthermore, while all happiness is subjective and transitory, leaving behind it no more than a "melancholy echo," the good has eternal value. Nobody would dream of honoring and respecting a person for his happiness or prosperity. On the other hand, we honor every good deed and the expression of every moral attitude, even if it took place in a distant land and among a foreign people. If we discover a good act or a good attitude in an enemy we nevertheless respect it and cannot help deriving a certain satisfaction from its existence. The same is true of good deeds carried out in ages long past. In all this the essentially timeless nature of morality becomes evident. Good deeds cease to exist as historical events only; their value, on the other hand, has eternal reality and is collected as an indestructible "fund." This may be a metaphysical statement, but it is not a piece of "metaphysical speculation." It simply makes explicit what the experience of the morally good discloses to phenomenological analysis (*Der Sinn unseres Daseins,* pp. 55–57).

REPLIES TO REINER. There is a great deal in this presentation with which one could take issue. If one is not misled by the image of the ever-growing, indestructible "fund," one may wonder, for example, what could be meant by claiming that the value of a good deed is "eternal," other than that most human beings tend to approve of such an action regardless of when or where it took place. However, we are here concerned primarily with the question whether Reiner has met the challenge of the pessimists, and it seems clear that he has not. A pessimist like Schopenhauer or Darrow might provisionally grant the correctness of Reiner's phenomenological analysis of morality but still offer the following rejoinder: The inevitable misery of all or nearly all human beings is so great that even if in the course of their lives they have a chance to preserve their inner moral natures or their good will, the continued torture to which their lives condemn them would not be justified. Given the pessimist's estimate of human life, this is surely not an unreasonable rejoinder. Even without relying on the pessimist's description of human life, somebody while accepting Reiner's phenomenological analysis might reach the opposite conclusion. He might, for example, share the quietist strain of Schopenhauer's teachings and object to the whole hustle and bustle of life, concluding that the "peace of the all-sufficient nothing"—or, more literally, a universe without human life—was better in spite of the fact that moral deeds could not then be performed. Since he admits the "facts" of morality on which Reiner bases his case but considers the peace of the all-sufficient nothing more valuable than morality, it is not easy to see how an appeal to the latter would show him to be mistaken. What phenomenological analysis has not disclosed, to Reiner or, as far as is known, to anybody else, is that doing good is the only or necessarily the greatest value.

WHY THE PESSIMIST CANNOT BE ANSWERED. The conclusion suggests itself that the pessimist cannot here be refuted, not because what he says is true or even because we do not know who is right and who is wrong but because the question whether a universe with human life is better than one without it does not have any clear meaning unless it is interpreted as a request for a statement of personal preference. The situation seems to be somewhat similar to what we found in the case of the question "Is my life better than my death?" when asked in certain circumstances. In some contexts indeed when we talk about human life in general, the word *better* has a reasonably clear meaning. Thus, if it is maintained that life for the human race will be better than it is now after cancer and mental illness have been conquered, or that human life will be better (or worse) after religion has disappeared, we understand fairly well what is meant, what facts would decide the issue either way. However, we do not really know what would count as evidence for or against the statement "The existence of human life as such is better than its nonexistence." Sometimes it is

claimed that the question has a fairly clear meaning, namely, whether happiness outweighs unhappiness. Thus, von Hartmann supports his answer that the nonexistence of human life is better than its existence, that in fact an inanimate world would be better than one with life, with the argument that as we descend the scale of civilization and "sensitivity," we reach ever lower levels of misery. "The individuals of the lower and poorer classes and of ruder nations," he writes, "are happier than those of the elevated and wealthier classes and of civilized nations, not indeed because they are poorer and have to endure more want and privations, but because they are coarser and duller" (*Philosophy of the Unconscious,* Vol. III, p. 76). The "brutes," similarly, are "happier (i.e., less miserable)" than man, because "the excess of pain which an animal has to bear is less than that which a man has to bear." The same principle holds within the world of animals and plants:

> How much more painful is the life of the more finely-feeling horse compared with that of the obtuse pig, or with that of the proverbially happy fish in the water, its nervous system being of a grade so far inferior! As the life of a fish is more enviable than that of a horse, so is the life of an oyster than that of a fish, and the life of a plant than that of an oyster. (Ibid.)

The conclusion is inevitable: The best or least undesirable form of existence is reached when, finally, we "descend beneath the threshold of consciousness"; for only there do we "see individual pain entirely disappear" (*Philosophy of the Unconscious,* Vol. III, pp. 76–77). Schopenhauer, also, addressing himself directly to the "*Zeugungsproblem,*" reaches a negative answer on the ground that unhappiness usually or necessarily outweighs happiness. "Could the human race continue to exist," he asks (in *Parerga und Paralipomena,* Vol. II, pp. 321–322), if "the generative act were … an affair of pure rational reflection? Would not rather everyone have so much compassion for the coming generation as to prefer to spare it the burden of existence, or at least be unwilling to take on himself the responsibility of imposing such a burden in cold blood?" In these passages Schopenhauer and von Hartmann assume that in the question "Is a world with human life better than one without human life?" the word *better* must be construed in a hedonistic or utilitarian sense—and the same is true of several other philosophers who do not adopt their pessimistic answer. However, while one may stipulate such a sense for "better" in this context, it is clear that this is not what is meant prior to the stipulation. Benedict de Spinoza, for example, taught that the most miserable form of existence is preferable to nonexistence. Perhaps few who have directly observed the worst agonies and tortures that may be the lot of human beings or of animals would subscribe to this judgment, but Spinoza can hardly be accused of a self-contradictory error. Again, Friedrich Nietzsche's philosophy is usually and quite accurately described as an affirmation of life, but Nietzsche was very careful not to play down the horrors of much of life. While he did not endorse Schopenhauer's value judgments, he thought that, by and large, Schopenhauer had not been far wrong in his description of the miseries of the human scene. In effect Nietzsche maintained that even though unhappiness is more prevalent than happiness, the existence of life is nevertheless better than its nonexistence, and this surely is not a self-contradiction.

It is important to point out what does not follow from the admission that in a nonarbitrary sense of "better," the existence of the human race cannot be shown to be better than its nonexistence: It does not follow that I or anybody else cannot or should not prefer the continued existence of the human race to its nonexistence or my own life to my death, and it does not follow that I or anybody else cannot or should not enjoy himself or that I or anybody else is "irrational" in any of these preferences. It is also impossible to prove that in some nonarbitrary sense of "better," coffee with cream is better than black coffee, but it does not follow that I cannot or should not prefer or enjoy it or that I am irrational in doing so. There is perhaps something a trifle absurd and obsessive in the need for a "proof" that the existence of life is better than its nonexistence. It resembles the demand to have it "established by argument" that love is better than hate.

Perhaps it would be helpful to summarize the main conclusions reached in this essay:

(1) In certain familiar senses of "meaning," which are not usually regarded as trivial, an action or a human life can have meaning quite independently of whether there is a God or whether we shall live forever.

(2) Writers such as Tolstoy, who, because of the horror that death inspires, conclude that death is better than life, are plainly inconsistent. Moreover, the whole question of whether my life is better than my death, unless it is a question about my preference, seems to be devoid of sense.

(3) Those who argue that no human action can be worthwhile because we all must eventually die

ignore what may be called the "short-term context" of much of our lives.

(4) Some human lives are worthwhile in one or both of the two senses in which "worthwhile" is commonly used, when people raise the question of whether a given person's life is worthwhile. The pessimists who judge human life by more demanding standards are not mistaken when they deny that by their standards no human life is ever worthwhile. However, they are guilty of a fallacious inference if they conclude that for this reason no human life can be worthwhile by the usual standards. Nor is it clear why anybody should embrace their standards in the place of those commonly adopted.

(5) It appears that the pessimists cannot be answered if in order to answer them one has to be able to prove that in some nonarbitrary sense of the word *better,* the existence of life is better than its nonexistence. But this admission does not have any of the gloomy consequences that it is sometimes believed to entail.

See also Baier, Kurt; Death; Happiness; Hartmann, Eduard von; Jung, Carl Gustav; Kant, Immanuel; Lenin, Vladimir Il'ich; Meaning; Nietzsche, Friedrich; Ostwald, Wilhelm; Pessimism and Optimism; Russell, Bertrand Arthur William; Schopenhauer, Arthur; Suicide; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

The position that human life cannot be meaningful without religious belief is defended in James Martineau, *Modern Materialism and Its Relation to Religion and Theology* (New York, 1877), and more recently in C. H. D. Clark, *Christianity and Bertrand Russell* (London: Lutterworth Press, 1958), and in E. L. Fackenheim, "Judaism and the Meaning of Life," in *Commentary* 39 (1965): 49–55. Substantially similar views are expounded in Paul Althaus, "The Meaning and Purpose of History in the Christian View," in *Universitas* 7 (1965): 197–204, and in various publications by Viktor Frankl. Frankl's *Man's Search for Meaning* (Boston: Beacon Press, 1963) contains a full list of his own writings as well as those of his followers, most of whom may be described as practicing "pastoral psychology." A proreligious position is also advocated by William James in "Is Life Worth Living?," in *The Will to Believe and Other Essays in Popular Philosophy* (New York: Longmans, Green, 1897), and *The Varieties of Religious Experience* (New York: Longmans, Green, 1902), Chs. 6 and 7. A milder version of the same position is presented in Chad Walsh, *Life Is Worth Living* (Cincinnati, no date).

Schopenhauer's views about the "vanity" of life are stated in Vol. I, Bk. IV, of *Die Welt als Wille und Vorstellung* (Leipzig, 1818), translated by R. B. Haldane and J. Kemp as *The World as Will and Idea,* 3 vols. (London: Trubner, 1883), and in several of his pieces included in *Parerga und Paralipomena,* 6th ed., edited by J. Frauenstädt, 2 vols. (Berlin, 1851). Three of his essays that bear most closely on the subject of the present entry—"On the Vanity and Suffering of Life," "On the Sufferings of the World," and "The Vanity of Existence"—are available in an English translation by T. Bailey Saunders in *The Will to Live—Selected Writings of Arthur Schopenhauer,* edited by Richard Taylor (Garden City, NY: Doubleday, 1962).

Eduard von Hartmann's position is stated in Vol. III of *Die Philosophie des Unbewussten,* 3 vols. (Berlin, 1869), translated by W. C. Coupland as *Philosophy of the Unconscious* (London, 1884), in *Zur Geschichte und Begründung des Pessimismus* (Berlin, 1892), and in *Philosophische Fragen der Gegenwart* (Leipzig: Haacke, 1885), Ch. 5. Clarence Darrow's pessimism is expounded in *The Story of My Life* (New York: Scribners, 1932) and in two pamphlets, "Is Life Worth Living?" and "Is the Human Race Getting Anywhere?" (Girard, KS, no date). Tolstoy's views are stated in "A Confession," in *A Confession, the Gospel in Brief and What I Believe,* translated by Aylmer Maude (London: Oxford University Press, H. Milford, 1940). Gloomy implications are derived from the second law of thermodynamics by Bertrand Russell in "A Free Man's Worship" (1903), which is available in several books, perhaps most conveniently in Russell's *Mysticism and Logic* (London: Allen and Unwin, 1918), and by Wilhelm Ostwald, *Die Philosophie der Werte* (Leipzig: Kröner, 1913). F. P. Ramsey, in "How I Feel," in *The Foundations of Mathematics and Other Logical Essays* (London: K. Paul, Trench, Trubner, 1931), agrees with Russell and Ostwald about the physical consequences of the second law but does not share their gloomy response. Stephen Toulmin, "Contemporary Scientific Mythology," in Toulmin et al., *Metaphysical Beliefs* (London: SCM Press, 1957), questions whether the second law has the physical consequences attributed to it by Russell, Ostwald, and many others. L. J. Russell, "The Meaning of Life," *Philosophy* 28 (1953): 30–40, contains some interesting criticisms of the view that eternal existence could render any human actions meaningful.

The fullest discussions of the questions of the meaning and value of life by contemporary analytic philosophers are Kurt Baier, *The Meaning of Life* (Canberra, 1957), parts of which are reprinted in *Twentieth Century Philosophy—The Analytic Tradition,* edited by Morris Weitz (New York: Free Press, 1966); Ronald W. Hepburn, *Christianity and Paradox* (London: Watts, 1958), Ch. 8; and Antony Flew, "Tolstoi and the Meaning of Life," in *Ethics* 73 (1963): 110–118. Baier, Hepburn, and Flew support the position that life can be meaningful even if there is no God and no afterlife. This position is also defended in Eugen Dühring, *Der Werth des Leben* (Leipzig, 1881), Chs. 6–7, and more recently in Bertrand Russell, *The Conquest of Happiness* (New York: Liveright, 1930), Ch. 2; Ernest Nagel, "The Mission of Philosophy," in *An Outline of Man's Knowledge of the Modern World,* edited by Lyman Bryson (New York: McGraw-Hill, 1960); Sidney Hook, "Pragmatism and the Tragic Sense of Life," in *Proceedings and Addresses of the American Philosophical Association* 33 (1960): 5–26; Karl R. Popper, *The Open Society and Its Enemies,* 2 vols. (5th rev. ed.,

London: Routledge and K. Paul, 1966), Vol. II, Ch. 25; and Kai Nielsen, "Examination of an Alleged Theological Basis of Morality," in *Iliff Review* 21 (1964): 39–49. Jean-Paul Sartre and Albert Camus are frequently (and rather inaccurately) described as "nihilists," but in effect they also take the position that although the universe is "absurd," human life can be meaningful. Sartre's views are found in *Being and Nothingness,* translated by Hazel E. Barnes (New York: Philosophical Library, 1956), Pt. 4. Camus's views are stated in *The Myth of Sisyphus and Other Essays,* translated by Justin O'Brien (New York: Knopf, 1955). Views very similar to those of Sartre and Camus are advocated by Flew and R. W. Hepburn in their BBC discussion "Problems of Perspective," which is printed in *Plain View* 10 (1955): 151–166. C. D. McGee, *The Recovery of Meaning—An Essay on the Good Life* (New York: Random House, 1966), Ch. 1, contains a lively and detailed discussion of some of the issues treated in the present entry. The author reaches similar conclusions but devotes far more attention to the "malaise" that inspires questions about the meaning of life. In a similar vein, Ilham Dilman, "Life and Meaning," in *Philosophy* 40 (1965): 320–333, concentrates on the psychological situations that prompt people to ask whether their own lives or the lives of others have meaning. Moritz Schlick, *Vom Sinn des Lebens* (Berlin, 1927), is concerned primarily with psychological questions, arguing that modern life tends to be spoiled by overemphasis on the achievement of distant goals. Sigmund Freud in several places alludes to the question of the meaning of life and usually dismisses it as senseless and pathological. "The moment a man questions the meaning and value of life," he wrote in a letter to Marie Bonaparte, "he is sick.… By asking this question one is merely admitting to a store of unsatisfied libido to which something else must have happened, a kind of fermentation leading to sadness and depression" (*Letters of Sigmund Freud,* translated by James Stern and Tania Stern, edited by E. L. Freud, New York: Dover, 1960, p. 436).

The Polish Marxist Adam Schaff deals with some of the issues discussed in the present entry in his *A Philosophy of Man* (London: Lawrence and Wishart, 1963). Schaff's views are criticized from a Christian point of view in Christopher Hollis, "What Is the Purpose of Life?," in *Listener* 70 (1961): 133–136. There is a discussion of the "meaning of life" from the point of view of fascism in Mario Palmieri, *The Philosophy of Fascism* (Chicago: Dante Alighieri Society, 1936). The "phenomenological" position of Hans Reiner, which was discussed in the final section of this entry, is stated in his *Der Sinn unseres Daseins* (Tübingen, 1960). Other more recent German works include *Sinn und Sein,* edited by Richard Wisser (Tübingen: Niemeyer, 1960); Reinhart Lauth, *Die Frage nach den Sinn des Daseins* (Munich, 1953); and Johannes Hessen, *Der Sinn des Lebens* (Cologne, 1933).

Psychological studies of people who attempted or who committed suicide are contained in Margarethe von Andics, *Suicide and the Meaning of Life* (London: Hodge, 1947); Louis I. Dublin and Bessie Bunzel, *To Be or Not to Be* (New York: H. Smith and R. Haas, 1933); and E. Stengel, *Suicide and Attempted Suicide* (London: MacGibbon and Kee, 1965).

Will Durant, *On the Meaning of Life* (New York: R. Long & R.R. Smith, 1932), consists of answers by various eminent men, including Mohandas Gandhi, H. L. Mencken, Russell, and George Bernard Shaw, to the question of what they take to be the meaning of life.

Paul Edwards (1967)

LIFE, MEANING AND VALUE OF [ADDENDUM]

Paul Edwards primarily addresses the "pessimist view" that if there is no God and death is final, life has no meaning. The focus here will be on subsequent philosophical work and on issues he leaves unaddressed. Some account of nonmonotheistic religion (Buddhism, Daoism, Confucianism, and Advaita Vedanta Hinduism) should be given, especially since religious perspectives are now taken more seriously by many in the analytic philosophical tradition.

Thomas Nagel (1986) argues both that (1) human life viewed objectively is insignificant though viewed subjectively is significant and that (2) it is our capacity to recognize both (1) and our constitutional self-absorption, which makes us irreducibly absurd and our lives ironic. Against (1) David Wiggins (2002) argues that for our strivings to matter, even subjectively, there must be something we can "invest with overwhelming *importance*," and that this entails both that values are objective, though "lit up by the focus one brings to the world", and that happiness is not supremely important. Robert Nozick (1989) shares this view and imagines a hermetically sealed "experience machine" that can undetectably provide apparently real and happy experiences involving others. Would a life be better lived inside the machine in a state of perpetual happiness or outside, with the tribulation and joys of genuine connection to others? Nozick argues for the latter, distinguishing between intrinsic value and meaning. The measure of a thing's intrinsic value is the degree of its diversity and the degree of the organic unity of that diversity. Meaning comes from a thing's connection to other things with intrinsic value—the greater their value and the stronger the connection, the greater the meaning. Thus, value is proportional to both internal integration and the strength of external connections to things of great value.

Turning to religious accounts, Philip Quinn (2000) distinguishes axiological and teleological questions. He argues that an integrated life might have intrinsic or "axiological value" though it lacked any overt connection to a

transcendent reality that would give it "teleological value." Keith Ward argues, "What distinguishes a religious view of nontrivial purpose is that all positive human purposes are subordinate to the one objective purpose of attaining the supreme goal of union with, or fulfilling relation to, the supreme value… " (Ward 2000, p. 20). Still, a life may be intrinsically valuable, worthwhile (to oneself) and reasonably happy, without being specially valuable or meaningful, and it may be meaningful and, thus, valuable without being somehow ultimately valuable.

Now if there is a transcendent entity such as God, Brahman, Nirvana, or the Dao, then one's connectedness to it, assuming a connection of great diversity to other centers of value such as people, would bring an ultimate or teleological meaning to life, obviating the absurdity that Nagel posits. Asian philosophical traditions view normal human nature as inadequate though improvable through self-reflection and right moral action; the Indic traditions, however, offer a bleaker view of human nature than the Chinese. For both, a life of value and meaning is only possible if one aligns the self with the underlying relational structure of reality: the Dao (Chinese traditions) or the Dharma (Hinduism and Buddhism). Daoism holds that meaning is achieved by returning to the natural self in accordance with the Dao, whereas Confucianism, which developed religious notions of a transcendent reality after Mahayana Buddhism entered China, emphasizes refinement, also holding that, as Tu Wei-Ming puts it, "we can realize the ultimate meaning of life in ordinary human existence" (Tu 1985, p. 60).

For Indic traditions, an illusory view of the self leads to an attachment to this life, preventing the attainment of meaning through a transcendent being (Brahman) or a state (Nirvana), "the fullness of being" and the "fullness of emptiness," respectively. Within Hinduism, Sankara's Advaita Vedanta holds Atman (the true self) identical to Brahman. The majority Hindu position of *bhakti yoga*, espoused by Ramajuna and closer to monotheism, focuses on the key teaching that Eliot Deutsch identifies in the *Bhagavad Gita*: that nonattachment is only possible via a new attachment to that of greatest value. "One overcomes the narrow clinging to results … only when that passion is replaced by one directed to the Divine" (Deutsch 1968, p. 163). Still, even if there is a transcendent reality, the best is not an enemy of the good, so axiological and teleological meaning can be compatible and even inextricably linked. "A man who knows his own nature will know Heaven" says Mencius, and Masao Abe articulates the meaning of life in Mahayana Zen this way: "For the sake of wisdom, do not abide in *samsara* (this life); for the sake of compassion, do not abide in *nirvana*" (Abe 2000, p. 161).

See also Brahman; Buddhism; Buddhism—Schools: Chan and Zen; Chinese Philosophy: Daoism; Indian Philosophy; Mencius; Nagel, Thomas; Nirvāṇa; Nozick, Robert; Wiggins, David.

Bibliography

Abe, Masao. "The Meaning of Life in Buddhism." In *The Meaning of Life in the World Religions*. Edited by Joseph Runzo and Nancy M. Martin. Oxford: Oneworld Publications, 2000.

Deutsch, Eliot. *The Bhagavad Gita*. Lanham: University Press of America, 1968.

Klemke, E. D., ed. *The Meaning of Life*. 2nd ed. New York: Oxford University Press, 2000.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Nozick, Robert. *The Examined Life*. New York: Simon and Schuster, 1989.

Quinn, P. L. "How Christianity Secures Life's Meanings;" In *The Meaning of Life in the World Religions*. Edited by Runzo and Martin. Oxford: Oneworld, 2000.

Tu, Wei-Ming. *Confucian Thought: Selfhood as Creative Transformation*. Albany: State University of New York Press, 1985.

Ward, Keith. "Religion and the Question of Meaning." In *The Meaning of Life in the World Religions*. Edited by Runzo and Martin. Oxford: Oneworld, 2000.

Wiggins, David. "Truth, Invention, and the Meaning of Life." *Proceedings of the British Academy* LXII (1976). Reprinted in *Needs, Values, Truth*. 3rd ed. Oxford: Clarendon Press, 2002.

Joseph Runzo (2005)

LIFE, ORIGIN OF

Two explanations dominated prescientific thinking about the origin of life: special creation and spontaneous generation. According to the former view, supernatural intervention was essential for the creation of life; according to the latter, living organisms could form spontaneously—for example, from the mud of the Nile. Not surprisingly, special creation was usually favored as an explanation of the origin of humans and the higher animals, whereas spontaneous generation seemed adequate to explain the origin of insects, frogs, and even mice.

The theory of spontaneous generation came under attack in the seventeenth century when the Italian scientist Francesco Redi showed that maggots do not arise spontaneously in rotting meat but develop from eggs laid

by flies. The spontaneous generation controversy persisted for another two hundred years or so until the classic experiments of Louis Pasteur convinced almost everyone that even microorganisms appear only as the descendents of similar microorganisms. This posed the problem of the origin of life in its modern form: How were the first organisms generated from abiotic matter?

The generally accepted answer to this question was provided by the theory of evolution through natural selection as proposed by Charles Darwin and Alfred Russell Wallace. Darwin in the final paragraph of the first edition of "On the Origin of Species" suggests that the whole complex world of life has evolved from one or a few simple kinds of organism that were formed on the Earth long ago. "There is a grandeur in this view of life with its several powers, having been originally breathed into a few forms or into one, and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved" (1859, chapter 14).

Darwin never published his thoughts on the origin of those earliest organisms, probably to avoid upsetting his wife, but in a much-quoted letter he speculates that life may have emerged "in some warm little pond with all sorts of ammonia and phosphoric salts, light, heat, electricity, etc., present" (1959 [1898], pp. 202–203). Thus Darwin thought that, long ago, a complex mixture of organic molecules was formed spontaneously on the Earth "in some little pond," and that they supported the appearance of the first simple living organisms. After that, the evolution of the whole biosphere was the consequence of natural selection acting on those earliest organisms and their descendants. Modern research on the origin of life is largely concerned with filling in the details of Darwin's scenario.

THE NATURE OF THE PROBLEM

The Earth is about 4.6 billion years old. The dating of the earliest fossil microorganisms remains somewhat controversial, but it seems almost certain that organisms not unlike modern bacteria or algae were already present on the Earth about 3.5 billion years ago. During the first half-billion years of Earth's history repeated impacts of comets, asteroids, and other interplanetary objects would have sterilized the Earth's surface, so up to half a billion years was available for the evolution of complex life from abiotic origins. There is no reason to doubt that this was long enough.

DNA sequencing and the comparison of the genomes of different organisms have revolutionized human understanding of the evolutionary relationships between the varied forms of life. While many details remain to be worked out, people already have a reasonable picture of the nature of the last common ancestor of all life, and a fairly detailed outline of the sequence in which the different fossil and extant forms of life evolved from it. The many gaps in the picture are likely to be filled in during the early twenty-first century. The outstanding problem, therefore, is that of the origin of the first living, replicating microorganisms. Most scientists believe that they originated on the Earth, although the possibility that they were brought here from elsewhere in the solar system cannot be dismissed out of hand.

EARLY EXPERIMENTAL STUDIES

The modern era of experimental origin-of-life studies began in 1953 with the classical experiments of Harold Urey and Stanley Miller. Alexandre Ivanovich Oparin in 1924 had suggested that the organic material needed to get life started was formed in the atmosphere of the Earth when the atmosphere was still reducing. Miller, then a graduate student working with Urey, tested this hypothesis by passing an electric discharge through a "reducing atmosphere" of methane, water, and ammonia. To the surprise of his contemporaries, Miller was able to detect among the products substantial amounts of several of the amino acids that are present in proteins. This was the first successful experiment designed to demonstrate that important components of contemporary living organisms are readily formed from simple starting materials under prebiotic conditions.

In the years following Miller's experiment, most of the organic molecules that are central to molecular biology were obtained by related methods. The discovery by Juan Oro that adenine, a component both of nucleic acids and of ATP, the energy currency of the cell, could be formed from a simple solution of ammonium cyanide was particularly impressive. However, this whole approach came under attack when it was realized that the atmosphere of the Earth could never have been as strongly reducing as Miller and Urey assumed. Whether it was ever sufficiently reducing to support similar chemistry, even if less efficiently, is uncertain.

A second possible source of the organic material needed to permit the origin of life was identified in the carbonaceous chondrites, a common class of meteorite. Careful chemical analysis showed that these stones contained abundant organic material, including amino acids

and the nucleic acid bases. Many scientists believe that meteorites, comets, and interplanetary dust provided much of the organic material for the origin of life.

In the late twentieth century another possible source of prebiotic organic material was identified, namely the deep-sea vents. In the vents, superheated water containing large amounts of metal sulfides comes into contact with cold seawater causing the sulfides to precipitate. Laboratory experiments suggest that metal sulfides can act as catalysts for the formation of a mixture of a variety of organic molecules from volcanic gases. Clearly, there are several possible sources of the prebiotic organic material needed for the origin of life, but it is not clear which of them was most important.

THE RNA WORLD

The most important recent advance in our understanding of the origin of life is the realization that there once was an RNA world. The modern biological world depends on a complex, interacting system of proteins and nucleic acids in which proteins are needed to replicate nucleic acids, but the formation of proteins depends on the prior presence of nucleic acids. It is now known that the DNA/RNA/protein world was preceded by a much simpler world in which RNA, without the help of proteins, fulfilled both a genetic and a functional role.

It is now clear from laboratory experiments that RNA molecules are capable of evolution by natural selection and are capable of catalyzing a variety of difficult chemical reactions. In particular it has been possible to evolve an RNA catalyst that carries out the most important step involved in RNA replication. It seems probable, therefore, that RNA catalysts (ribozymes) once supported a fairly complex form of life, without the help of proteins. Thus the problem of the origin of life is simplified: How were the first replicating molecules of RNA synthesized on the primitive Earth?

Attempts to demonstrate the synthesis of RNA under prebiotic conditions have met with some success, but formidable difficulties remain. The monomeric components of RNA, ribonucleotides, are complicated organic molecules made up from a sugar, a heterocyclic purine or pyrimidine base, and an inorganic phosphate group. The prebiotic syntheses of the two organic components that have been reported are relatively inefficient and nonspecific, and the combination of the three elementary components to form ribonucleotides is complicated by several troublesome side reactions. A great deal of novel chemistry needs to be discovered before a plausible prebiotic synthesis of the nucleotides can be claimed. A number of scientists are working on the problem.

The formation of long polymers from ribonucleotides is another difficult step in the synthesis of RNA. However, substantial successes have been achieved in model systems. The most extensive studies make use of an abundant clay mineral, montmorillonite, to catalyze the polymerization of an analog of the activated nucleotides that are used in the enzymatic synthesis of RNA. This work emphasizes the important role that minerals are likely to have played in the origin of life. It seems probable that many of the most difficult reactions needed to get life started occurred on mineral surfaces rather than in solution.

The replication of DNA or RNA is dependent on specific base-pairing between adenylic acid and uridylic or thymidylic acid and between guanylic acid and cytidylic acid. Base pairing is an intrinsic property of the nucleotide bases, so that a preformed strand of RNA (DNA) will align the complementary mononucleotides in the correct sequence even in the absence of a protein enzyme. If the nucleotides are presented in an activated form suitable for incorporation into polymers, a preformed RNA (DNA) strand, therefore, will bring about the nonenzymatic synthesis of a new complementary strand. This process is known as template-directed synthesis.

Template-directed synthesis is a central theme in many scenarios for the origin of the RNA world. It has been shown, for example, that a great variety of RNA sequences can be "copied," that is a great variety of sequences will catalyze the synthesis of their complements, converting single-stranded RNA to double-stranded RNA. Thus mineral catalysis of the formation of long single-stranded RNA molecules followed by template-directed copying could, in principle, have assembled a complex mixture (library) of double-stranded RNA on the primitive Earth, but only if a supply of ribonucleotides was available.

It is possible to propose a scenario for the origin of the RNA world by optimistic extrapolation of the available experimental evidence. First nucleotides were formed abiotically; they condensed together on mineral surfaces to give single-stranded RNA that was then copied by template-directed synthesis to give a "library" of double-stranded RNA molecules. Among these was one that included an RNA polymerase that was able to get efficient RNA replication started.

The serious obstacles to the prebiotic synthesis of RNA have led many researchers to propose a different kind of solution to the problem of the origin of the RNA world. They believe that one or more much simpler biochemical worlds preceded the RNA world and "invented" RNA. The search for such simple worlds is just beginning, but there are already a number of RNA-like polymers that, although they are somewhat simpler than RNA, look as though they could have functioned as genetic systems. The search for even simple systems is an active field.

SUMMARY

It is generally accepted that once a replicating genetic polymer appeared on the early Earth, evolution through natural selection could account for the appearance of ever more complex organisms, and finally of the familiar biosphere. It is known that one such evolving world, the RNA world, preceded the world of DNA, RNA, and proteins. Scientists do not know how the RNA world came into existence. There are several theories, but none is as yet supported by strong experimental evidence. Ongoing research should provide an answer sometime in the early twenty-first century.

See also Darwinism.

Bibliography

Botta, O., and J. L. Bada. "Extraterrestrial Organic Compounds in Meteorites." *Surveys in Geophysics* 23 (2002): 411–467.

Darwin, C. R. "Letter to J. D. Hooker, [1 February] 1871. In *The Life and Letters of Charles Darwin*. Vol. II, ed. by F. Darwin [1898]. New York: Basic Books, 1959, pp. 202–203.

Darwin, Charles. *On the Origin of Species*.. London: John Murray, 1859.

Farley, J. *The Spontaneous Generation Controversy from Descartes to Oparin*. Baltimore: Johns Hopkins University Press, 1977.

Gilbert, W. "The RNA World." *Nature* 319 (1986): 618.

Johnston, W. K., P. J. Unrau, M. S. Lawrence, M. E. Glasner, and D. P. Bartel. "RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension." *Science* 292 (2001): 1319–1325.

Joyce, G. F. "Nonenzymatic Template-Directed Synthesis of Informational Macromolecules." *Cold Spring Harb Symp Quant Biol* 52 (1987): 41–51.

Kasting, J. F., and L. L. Brown. "The Early Atmosphere as a Source of Biogenic Compounds." In *The Molecular Origins of Life*, edited by A. Brack, 35–56. New York: Cambridge University Press, 1998.

Miller, S. L. "A Production of Amino Acids Under Possible Primitive Earth Conditions." *Science* 117 (1953): 528–529.

Oparin, A. I. *The Origin of Life*. Translated by S. Morgulis. New York: Macmillan, 1938

Orgel, L. E. "The Origin of Life on the Earth." *Scientific American* 271 (1938): 52–61.

Oro, J., and A. Kimball. "Synthesis of Adenine from Ammonium Cyanide." *Biochem Biophys Res Commun* 2 (1960): 407–412.

Steitz, T. A., and P. B. Moore. "RNA, the First Macromolecular Catalyst: The Ribosome is a Ribozyme." *Trends Biochem Sci* 28 (2003): 411–418.

Leslie E. Orgel (2005)

LIFE, SANCTITY OF

See *Popper-Lynkeus, Josef; Suicide*

LIPPS, THEODOR

(1851–1914)

Theodor Lipps, a psychologist and philosopher, was born in Wallhalben, Rhineland-Palatinate. He studied theology and natural science at Erlangen, Tübingen, Utrecht, and Bonn. He obtained academic positions in Bonn (1884), Breslau (1890), and finally in Munich (1894), where he remained until his death. There he was a full professor and the teacher of Johannes Daubert and Alexander Pfänder, the founding members of the Munich circle of phenomenology. Lipps published voluminously on a large variety of topics, though his orientation in philosophy was consistently a psychological one.

In *Basic Facts of Mental Life* (1883) Lipps states his conception of philosophy as follows: "Inner experience is the basis for psychology, logic, aesthetics, ethics, and the adjunct disciplines, including metaphysics in the sense in which it is permissible to speak of it. We regard all these disciplines now as philosophical, and at least in the main they fill what is usually viewed as the range of tasks that we especially honor with the name of philosophical ones. Their objects are presentations, sensations, and volitional acts, and no intelligent person denies that such objects are different from the subject matters of other sciences and therefore require their own manner of scientific treatment" (p. 3). Thus he conceived of philosophy as equivalent to or based on psychology, with an emphasis on "inner perception."

This psychologal style of philosophy is also evident in Lipps's views on logic. These in particular became subject to attack in Husserl's critique of psychologism. By no means, however, was the close tie between philosophy and psychology unusual for the late nineteenth and early twentieth centuries. The empirically minded psychological turn that occurred in the German-speaking world at that time was an attempt to establish philosophy as a sci-

ence amid the skepticism that was rife in the aftermath of the collapse of the speculative systems of German Idealism. Although Lipps's philosophical endeavors arose in this context, his approach to psychology differs significantly from the approaches of most of his contemporaries. He was, for example, willing to allow not only for inner perception but also for introspection or self-observation (*Selbstbeobachtung*), a notion that was unacceptable to many other philosopher-psychologists of the time, most notably Franz Brentano and his orthodox followers, for whom inner perception can never become self-observation.

The subject matter of psychology, according to Lipps, consists of conscious experiences (*Bewusstseinserlebnisse*), which always belong to an ego (*Ich*). It is, moreover, to be an empirical science. The ego to which conscious experiences belong—which is not to be confused with the soul (*Seele*)—is empirically given just as these experiences themselves are. "And the ego," Lipps significantly adds, "can intentionally direct its gaze upon itself. It can itself be an 'object.'. It can grasp and cognize itself" (1909, p. 6). Although this acceptance of the notion of self-observation put him at odds with Brentano and other contemporaries, Lipps had much in common with Brentano, Dilthey, and others insofar as he distinguished between two aspects of psychology as an empirical science: one descriptive and analytical, the other explanatory. The latter can involve physiological considerations and laboratory experiments in order to provide causal explanations of how conscious experiences arise, whereas the former makes no use of physiology or experimentation. It was this descriptive or "pure" psychology that primarily interested Lipps.

Lipps's most outstanding and enduring contribution is his concept of empathy (*Einfühlung*). This idea is of special importance because it was adopted and critically revised in such phenomenological theories of intersubjectivity as those developed by Husserl and Edith Stein. By means of empathy we come to know not only other minds but also other important objects of experience, such as those belonging to organic nature and works of art. One empathizes when one puts oneself in the place of—and even to some extent imitates—someone or something else. Lipps asserted with particular emphasis, contrary to some of his critics, that our knowledge of other minds is first and foremost grounded in empathy and thus in feelings rather than in purely intellectual operations. The pervasive role that he gave to empathy in his wide-ranging philosophical investigations naturally led to panpsychism in metaphysics.

The philosophy that Lipps developed out of his psychological studies was by no means subjectivistic or relativistic. This was certainly not the case with his logic. Moreover, in both aesthetics and ethics he thought it was possible to formulate universally valid prescriptions on the basis of psychology. As the science of the beautiful—of that which evokes or is suited to evoke the feeling of beauty (*Schönheitsgefühl*)—aesthetics aims to establish the psychological conditions under which such a feeling arises. Ethics, according to Lipps, is concerned with universally valid morality (*Sittlichkeit*) as opposed to the morals (*Moral*) of this or that historical period, nation, class, or individual. A Kantian influence is evident in his ethical reflections, in which the moral person (*sittliche Persönlichkeit*) is given the status of the highest good. In spite of this influence from Kant, however, Lipps presented a philosophical viewpoint that should be considered on its own merits and not merely in the shadow of his predecessors.

See also Brentano, Franz; Husserl, Edmund; Pfänder, Alexander; Phenomenology.

Bibliography

Bokhove, Niels, and Karl Schuhmann. "Bibliographie der Schriften von Theodor Lipps." *Zeitschrift für philosophische Forschung* 45 (1991): 112–130.

Kesserling, Michael. "Theodor Lipps (1851–1914): Ein Beitrag zur Geschichte der Psychologie." *Psychologische Beiträge* 7 (1962): 73–100.

Lipps, Theodor. *Ästhetik. Psychologie des Schönen und der Kunst*. 2 vols. Hamburg/Leipzig: Leopold Voss, 1903/06.

Lipps, Theodor. *Grundtatsachen des Seelenlebens*. Bonn: Max Cohen & Sohn (Fr. Cohen), 1883.

Lipps, Theodor. *Leitfaden der Psychologie*. 3rd ed. Leipzig: Wilhelm Engelmann, 1909.

Lipps, Theodor. "Zur Einfühlung." In *Psychologische Untersuchungen*. Vol. 2. Edited by Theodor Lipps. Leipzig: Wilhelm Engelmann, 1913.

Robin D. Rollinger (2005)

LIPSIUS, JUSTUS

(1547–1606)

Justus Lipsius, the Flemish humanist, classical philologist, and literary critic, foremost interpreter of Stoicism in the later Renaissance, and the founder of modern neo-Stoicism, exercised a strong influence on later moral thought. Born near Louvain, he spent most of his life in exile. At the age of twenty-four, he renounced the Catholicism of his native land, accepting the chair of his-

tory and eloquence at the Protestant University of Jena (1572). After two years, he returned—ostensibly as a repentant Catholic and loyal Brabantian. Again forced to flee—this time to the Calvinist Dutch—and abjuring Catholicism a second time, he accepted the chair of history at Leiden (1579). Harassed constantly by political and religious pressures, he went to the University of Louvain, becoming one of its most prominent scholars.

The vicissitudes of his life began during the time of civil war in the Low Countries. His *Tacitus* appeared at Louvain the year after his return from Jena (1575), as did his *Antiquae Lectiones*. These commentaries on Plautus signaled his adoption of a literary style modeled after Plautus, Tacitus, and Seneca. Lipsius was profoundly influenced by the thought and prose style of Seneca and devoted the remainder of his life to the study of Stoicism. This work of Lipsius, in turn, influenced Michel Eyquem de Montaigne, Guillaume du Vair, and Pierre Charron, and in England, Francis Bacon and Joseph Hall.

The victories of Don John of Austria (Gembloux, 1578) caused Lipsius to flee to the home of his friend Christophe Plantin, and then from Antwerp to Leiden, where he became a Calvinist. Here appeared *De Constantia* (1584), an introduction to Stoicism and his most famous work. Another well-known work, *Politicorum Libri Sex* (1589), led to a bitter dispute over its advocacy of severe methods to curb unrest. His position again became intolerable; finally, he made his peace with the Jesuits (and his old friend Martin Delrio) at Mainz (1591) and returned to Catholic Europe. He accepted the chair of history and Latin literature at Louvain (1592) and was also appointed professor of Latin at the Collegium Trilingue. He published several pieces on miracles as testimonials of faith, which added little to his fame. A projected *Fax Historica*, on Greco-Roman history and the histories of the Jews, Egyptians, and others, was never completed, although several parts were published. His last works were *Manuductio ad Stoicam Philosophiam* (1604), a miscellany of Stoic moral doctrines and survey of the *Paradoxa;* and *Physiologia Stoicorum*, a careful study of the Stoic logic and physics (1604). These make clear that Lipsius was responsible for a restored Stoic philosophy and particularly for the reemphasis on natural philosophy. Although he counted himself more an eclectic than an orthodox follower of any school, Lipsius attempted to show in these works that there was no real difficulty in reconciling the Stoic *fatum* with the Christian emphasis on free will (whereas in *De Constantia*, this possibility had been rejected).

See also Bacon, Francis; Charron, Pierre; Humanism; Montaigne, Michel Eyquem de; Moral Epistemology; Renaissance; Seneca, Lucius Annaeus; Stoicism.

Bibliography

WORKS BY LIPSIUS

Opera Omnia. 4 vols. Wesel, 1675.

Tvvo Bookes of Constancie, edited by R. Kirk. Translated by Sir John Stradling. New Brunswick, NJ: Rutgers University Press, 1939.

WORKS ON LIPSIUS

Corbett, Theodore G. "The Cult of Lipsius: A Leading Source of Early Modern Spanish Statecraft." *Journal of the History of Ideas* 36 (1975): 139–152.

Croll, Morris W. "Juste Lipse et le mouvement anti-cicéronien à la fin du XVII[e] siècle." *Revue du seizième siècle* 2 (1914): 200–242.

Long, A. A. "Stoicism in the Philosophical Tradition: Spinoza, Lipsius, Butler." In *Hellenistic and Early Modern Philosophy*, edited by Jon Miller. Cambridge, U.K.: Cambridge University Press, 2003.

Saunders, J. L. *Justus Lipsius: The Philosophy of Renaissance Stoicism*. New York: Liberal Arts Press, 1955.

Zanta, L. *La renaissance du stoicisme au XVI[e] siècle*. Paris: Champion, 1914.

Jason L. Saunders (1967)
Bibliography updated by Tamra Frei (2005)

LITERATURE, PHILOSOPHY OF

The concepts of fiction and of literature are distinct. On the one hand, there are nonfictional literary works—essays, memoirs, biographies, histories, writings about nature, and even philosophy. Perhaps we should also include some letter collections, diaries, and journals. On the other, there are nonliterary fictions both within and apart from the world of art. Cinema is full of fictional stories. Paintings represent fictional scenes. Advertising, whatever the medium it employs, often presents us with fictions. However, the concepts of fiction and literature are intertwined.

The paradigmatic literary works have steadily drifted toward being exclusively works of fiction: novels, stories, poems, and plays. When David Hume wanted to make his mark as a man of letters, he chose history and philosophy as his media. By comparison, Jean Paul Sartre made his literary mark with novels and plays while establishing his reputation as a philosopher with the contemporary equivalent of treatises and inquiries. Does this shift in lit-

erature's center of gravity reflect something important about it? Is there something about the value of literature that makes fictional works most apt to contain such value or is there perhaps an overlap between the value of fiction and literary value? We will discuss both concepts here, beginning with philosophical issues concerning fiction.

WHAT IS FICTION?

There are at least two senses of the word *fiction* that are easy to run together, but need to be distinguished for our present purpose. In one sense, a fiction can simply be a type of falsehood as when one says, "Your PhD is a fiction." By contrast, if one says that *Middlemarch* is a fiction, they are not saying that there is no such novel. They are saying that it is a certain type of book, story, or representation. It is true that there is such a book, story, or representation.

Unlike ambiguous words such as *bank*, there is probably some connection between the two senses of fiction, which explains the ease with which they are run together. Works of fiction typically contain an element of unreality. In reality, there is no such town as Middlemarch and no such people as the characters Dorothea or Casaubon who in the fiction inhabit the town. On another level, it is important to realize that the logical or semantic relationship between the two senses of fiction is loose. Fictions in the first sense can be lies and always involve falsehood. Works of fictions—a class of representations—are never lies, although they might just conceivably contain an intentional falsehood. They can refer to real things such as historical personages (Julius Caesar, Napoleon Bonaparte) and actual places (Rome, Moscow), and can contain truths about them.

The sense of fiction that primarily interests us is the second one, which refers to a class of works: works of fiction. Our job is to figure out what characterizes it and makes it distinct from other representations. Some have attempted to define fiction as a type of linguistic discourse. (Gale 1971, Urmson 1976). We know in advance that this is inadequate because of the ample existence of nonlinguistic fiction (see K.Walton [1990], for a survey and critique of this view). A second popular approach is to think of fiction as a form of pretense (though with no intent to deceive). This is on the right track, but the trick is to identify the right kind of pretense.

One might think that the standard function of a mode of representation such as language is to inform us about the actual world, to assert or show us things about it. Fiction could then be thought of as something derived from this standard use. Instead of actually asserting something, a fictional story or its author pretends to assert it (Searle 1975). The problem with this version of the pretense view is that it is not always the right description of what artists are doing in their works. Consider a clear case of pretense: Someone is pretending to sing by lip-synching. They are doing one thing in order to pretend to do another. Is Eliot pretending to describe a real town by representing one that does not exist? That does not seem right. To adequately describe what Eliot is doing, it is enough to say that she is writing about an imaginary town. The problem is to say what *about* means in the previous sentence.

The make-believe view offers an answer (Walton 1990). In order to understand this view one has to recognize that make-believe is being used in a restricted, somewhat technical sense. Make-believe in the relevant sense involves two special features. First it involves props. Props are publicly accessible objects that guide imaginings. If, for example, children are playing school with dolls, the dolls are props. Second, make-believe, unlike some other imaginings, operates according to underlying rules about these props, which authorize or mandate certain imaginings. For example, the game of school might operate according to the rule that the number of students in the classroom is equal to the number of dolls arranged in a certain way.

According to the make-believe view, a work of fiction—whether it be a painting, novel, or poem—is a work that is intended or has the function of being a prop in a game of make-believe. What makes *Middlemarch* fictional is that it is a work—a novel in this case—intended to be or having the function of being a prop of the kind described above. It prescribes that we imagine certain things: that there is a town inhabited by such and such people. This is the sense in which it is *about* a town and its inhabitants.

The make-believe view has become one of the most widely held views about the nature of fiction (Currie 1990, Lamarque and Olsen 1994, Levinson 1996, Walton 1990), but it seems to count works as fictional items that are usually not so considered. Suppose that one writes an autobiography, but in such a way that the reader can vividly imagine the events of the writer's life. Then it appears that this work fulfills two functions. One is to inform the reader about the writer's life. A second is to enable the reader to engage in the kind of guided imagining that is constitutive of make-believe in the technical sense. Something similar happens in certain works of history and journalism, as well as nonfiction novels such as Truman Capote's *In Cold Blood*. All these works are props

that authorize certain imaginings. There are some who claim that because of this, these works are fictional even if the primary purpose lies elsewhere. (Walton 1990). But this is counterintuitive: Historical novels are fictional; history is not, even if it uses techniques that produce guided imaginings. How are we to express the difference?

It might be suggested that fiction always presents some things to the imagination that are placed there simply for the purpose of being imagined. Whether or not they express truths or refer to items in the actual world is irrelevant to their proper functioning in the work. This need not be true in the nonfiction works just mentioned. This does not mean that fictions cannot contain some elements that are meant to express truth, or pick out actual people, places, or other things. It just means that not everything in the fictional work so functions. Even in a historical novel, where every character picks out a real person from the past, we are to imagine certain doings or conversations without worrying whether they occurred. So on the present proposal, something is a fiction if it is a work that is intended for, or has the function of, being a prop in a game of make-believe, and at least some of the things it mandates to be imagined are placed in the work just for the sake of being imagined.

FICTIONAL CHARACTERS

Works of fiction prescribe us to imagine people and their doings. Because of this, we say that such works, or their authors, create characters, about which we talk when describing and interpreting fictions. How should we understand such talk? How literally should we take it?

Consider the sentence, "Dorothea walked about the house with delightful emotion." A sentence such as this one normally refers to someone and says something about her, but if it appears in a work of fiction not about any actual personage, as this one does in *Middlemarch*, does it still refer to someone or, at least, to something? There are three answers to this question that currently have serious advocates: 1) the sentence still refers to someone, but a someone who does not exist; 2) the sentence does not refer to anyone, but it does refer to something (viz., a fictional character); 3) the sentence does not refer to anyone or anything.

Some proponents (Walton 1990, Lewis 1978) of the make-believe view hold the last view. Such a sentence refers to no one, but we make-believe that it does. In contrast, if we encounter in a work of fiction the sentence, "There was once a woman who was very happy," we may just make believe that there is some happy woman, at least until we are told more about her. An alternative way of putting the third position is to say that, although the original sentence refers to nothing, it is fictionally true or true in the story that it does. (Adams, Stecker, Fuller 1997)

Proponents of the second view believe that such things as fictional characters actually exist. (Howell 1979, Lamarque and Olsen 1994, Thomasson 1999, van Inwagen 1977, Wolterstorff 1980). They posit such things more to explain the things we say about fiction than to explain fiction itself. They might even agree that the original sentence, as it occurs in the story, does not refer to anyone or anything. But in creating a work of fiction, we also create other things including characters. We can then go on to talk about them, compare them to other characters, quantify over characters, and so on. Consider the claim that Hamlet is one of the most enigmatic characters in literature. Here we appear to be saying something about Hamlet, not merely making believe something. Characters are not people, although fictional works speak of them as if they are.

The plausibility of this view hangs on whether we actually gain something by assuming fictional characters exist, that is denied to those who claim that we merely make-believe that people are being referred to in fictions, or who claim that it is merely true in the fiction, but not in reality, that such reference occurs. The latter would say that the enigmatic thing is what make believe the play *Hamlet* prescribes or what is true in the play. One thing that is gained by positing the existence of characters is a convenient way to express ourselves when we talk about fictions. The paraphrases of statements about characters in terms of what is true in a story or what make-believe is prescribed by a story will always be more cumbersome. In practice, we will always prefer character-talk. But that does not settle the question whether character-talk really refers to characters rather than works.

On the second view, characters are not what they appear to be. They are not princes, lovers, or detectives. They are not male or female. They are not people. Presumably they are abstract entities of some sort, the properties of which are all, in one way or another, parasitic on the properties of the works in which they appear. Dorothea has the property of being a character in *Middlemarch*. She, or rather it, also has the property of being ascribed the property of walking about with delightful emotion on a certain occasion. But Dorothea does not actually have the property of walking about with delightful emotion. Proponents of the first view find this counterintuitive. They claim that Dorothea is a person capable of ambulating, feeling emotion, and having a gender.

Middlemarch refers to her. In general, fictional works really do refer to people and other things, only often they are fictional people and other fictional things (i.e., people and things that do not exist) (Dilworth 2004, Parsons 1980, Zalta 1983, Zemach 1997). Dorothea and Hamlet do not exist, according to them. In this they agree with some of those who hold the make-believe view. But the people in this camp do not think that that is a reason to deny that we refer to fictional things. In fact, their chief claim is that we can refer to what does not exist including fictional people and places.

The straightforward way in which the first view treats characters is refreshing after the cumbersome paraphrases of the third view and the metaphysical abstractions of the second. Unsurprisingly, the straightforwardness comes with a cost: a highly unorthodox conception of reference. What is fair to call the majority view (which obviously does not mean it is the true one) is that one can only refer to what exists. When we refer to something, we pick it out, and what does not exist cannot be picked out because there is nothing to be picked out. If there were something, it would exist. The things we refer to are distinguished from others in virtue of their properties or characteristics, but nothing can have properties unless it exists in the first place. Existence is not just another property, but is the condition for having properties. What does (did, or will) not exist is nothing and so cannot have properties. If the first view is to get off the ground, it would have to show that the orthodox conception of reference is mistaken. Currently, there is no consensus about which of these views is the most plausible, but rather a lively, ongoing debate.

THE PARADOX OF FICTION

Whatever is the correct view regarding fictional characters, once we become imaginatively involved in stories, we develop feelings and attitudes that appear to be directed toward creatures of fiction. We commonly say that we fear Dracula, despise Casaubon, or admire Sherlock Holmes. Yet there is something paradoxical about this. Feeling fear normally involves believing both that there is something to be feared, and that it poses a danger. We do not believe that Dracula actually exists, or that he poses a danger. Yet we feel fear nevertheless.

None of the views about fictional characters discussed in the preceding section offers a solution to this paradox. Two of the three deny that characters exist. They lead us into, rather than resolve, the problem. Those who claim fictional characters exist, deny that they are people, monsters, or anything else that could stir us to feel as we do. Characters, on this view, are abstract entities, and fearing them would be akin to fearing the number five.

The paradox of fiction has provoked an enormous literature, and many proposed resolutions. Three will be discussed here. The first denies that the object of fear is really fictional. (Charlton 1984, Paskins 1977). When we say we despise Casaubon or admire Holmes, we mean that we despise or admire people like them. We despise self-absorbed people who care nothing even for those close to them. We admire people with intellects (but not necessarily opium addictions) such as Holmes. Factualism, as this view is sometimes called, has some truth to it, but it cannot solve the whole problem. We don't fear creatures such as Dracula because we have no more of a belief in vampires in general than we do in Dracula in particular. Equally important, many of the feelings we develop in the course of taking in a fictional work, are guided by the specific things we imagine as we do this, and for this reason do not generalize beyond the fiction. As Anna Karenina approaches the railroad station, we hope she will turn away rather than enter and throw herself under the train. This is not the hope that despairing lovers will turn away from train stations, or, more generally, will refrain from suicide.

A second view is a further development of the make-believe approach to fiction. (Walton 1990, Levinson 1996). The basic idea here is that fear of Dracula, for example, occurs within the game of the make-believe we play when watching a Dracula movie. Hence, it is not literally fear, any more than our thought that Dracula lives in Transylvania is literally a belief. Our make-believe may, nevertheless, be phenomenologically indistinguishable from fear. That is, it can involve the same physiological changes in the body, we can experience similar feelings, and we may have an attenuated desire to duck, hide, or flee. What we lack is the beliefs that we have with real fear, and the full range of desires and behavioral tendencies.

The last view, known as the thought theory, rebels at the idea that what we feel are not real emotions—for example, real fear. The chief claim here is that emotions such as fear and pity do not require a belief in the existence of the object of these emotions. The emotions can be caused by vivid imaginings as well (Carroll 1990, Dadlez 1997, Feagin 1996, Gron 1996, Lamarque 1996, Yanal 1999).

It is not clear that we need to take these last two views as offering genuinely distinct theories (Currie 1997). Proponents of the thought theory must admit that when imaginings cause fear, it is different in some important respects than belief-induced fear. In addition to the dif-

ference in propositional attitude (believing that versus imagining that) there are cognitive and behavioral differences as well. Proponents of the make-believe view are willing to admit we feel a real emotion, but deny it is literally fear or pity. So it is not clear that the argument between the make-believe view and the thought theory amounts to more than a dispute over the name we should give to the feelings that arise in our imaginative encounters with fiction. They appear to agree about the nature and cause of those feelings.

WHAT IS LITERATURE?

The nature of literature is just as much a matter of controversy as the nature of fiction. However, it is now widely accepted that certain definitions will not work. In the first half of the twentieth century there was the hope that literature could be defined as a special way of using language. Literature uses defamiliarized language, drawing attention to its own literary devices. (Beardsley 1958, Jakobson 1960, Wellek and Warren 1973). But on the one hand, literary works can adopt the form of any kind of writing, from the scientific report to the advertising jingle. And on the other, all sorts of nonliterary uses of language can be rife with literary devices such as figures of speech, rhetorical techniques, implicit meanings, and so on.

Three proposals will be considered for defining literature. The first defines literature in terms of a role it plays in society or a community within society. Something is a work of literature, on this view, if it is a piece of writing that fulfills this role. Different theorists in this camp define the relevant role differently. For some, the relevant community is the community of critics, and the relevant role is that of being deemed worthy, or simply being the object, of critical attention (Fish 1980). For others, the relevant community is society at large, and the relevant role is sustaining the structure of power in the society (Eagleton 1983). It is not clear, however, that this approach can succeed in defining literature, whatever insights underlie it. Consider the first version. Who are the critics in question and what does critical attention consist in? They are the literary critics of course rather than the interpreters of philosophical texts (unless they are literary interpreters of those texts from the right academic departments). There are two dangers here and it is virtually impossible to avoid both. One danger is circular definition. The critics are those whose job it is to attend to a certain body of works—works of literature. Alternatively, the critics are those who use certain techniques—but those techniques can and sometimes are used on all sorts of things so that we get the extension of literature quite wrong.

A second approach asserts literature is a practice. Writers, readers, critics all enter into this practice by attempting to create, enjoy, or facilitate the appreciation of literary aesthetic value (Lamarque and Olsen 1994). To avoid circularity, literary aesthetic value is cashed out as the value to be found in the experience of a subject or story that has a humanly interesting content in virtue of embodying one or more perennial themes and that is given a complex form suitable to developing such a theme.

What seems right about this approach is the claim that the creation of literature is imbedded in a social practice with distinctive aims, institutions, and traditions. What is controversial about the approach is its conception of the practice in terms of aiming at a single kind of value in a way that has remained unchanged, at least since ancient Greece. When one thinks of all the various items that are relatively uncontroversial examples of literature, from ancient classics to eighteenth century essays to contemporary poetry, one must wonder whether the formula proposed by this definition really encompasses all of literature.

An alternative is to think of literature as a practice defined by an evolving set of values or functions and central art forms. Currently, these forms are the novel, short story, drama, and poetry, and in addition to their aesthetic value, we also characteristically value them in other ways such as for fulfilling certain cognitive functions, and for providing opportunities for open-ended interpretation. Anything that belongs to such an art form and is seriously intended to provide one or another of these values is a work of literature, but so are other pieces of writing that fulfill these valuable functions to a significant degree whether or not they are in one of the central literary forms. Finally, it should be recognized that our current concept of literature has itself evolved from earlier predecessor concepts, such as those of fine writing (belle lettres) and the ancient Greek or Latin classic. Items that fall under these predecessor concepts also belong to literature by a principle of inclusion implicit in our current concept. (Stecker 1996).

CRITICISM AND INTERPRETATION

Criticism is the blanket term for writing about or commenting on individual literary (or art) works. Being a blanket term, it covers different kinds of projects. One of the oldest kinds exists to orient an audience to new literary (artistic) productions as they appear. In doing this,

this kind of criticism fulfills a variety of distinct functions. It will typically identify the sort of work under discussion (e.g., an experimental novel in the manner of so and so), and acquaint a potential reader with important features of the work such as its style, plot, themes, and characters. Often implicit in these descriptions is an appreciative response (positive or negative) by the critic leading to an explicit evaluation of the work. The contemporary review is an example of this sort of criticism.

A different activity—that of analyzing and interpreting literary works—became a central critical activity in the twentieth century. This had a variety of causes. One was the rise of English and, more generally, literary studies, as an academic discipline. This generated a series of debates about the nature, content, value, and proper reception of such works, which associated a work with a great variety of ways of taking or reading it—in essence, a great variety of interpretations. Another factor was the growing prominence of difficult avant-garde works that are simply hard to understand. For such works at least, it is natural to turn to analysis and interpretation in order to understand and appreciate them. However, once we see how such analysis generates unexpected meanings or significance in these works, one suspects it might do so in any work, making any literary work a candidate for interpretation.

There are a variety of parameters along which approaches to interpreting literary works diverge. One that arose early on and has remained prominent concerns the significance of authorial intention in interpretation. Is the meaning of a work identical to such intentions, do they resolve ambiguities and other uncertainties in the work, or are they absolutely irrelevant to correctly interpreting it? Those who originally disagreed on this matter (Beardsley 1958 and 1970, Hirsch 1967) nevertheless did agree that the purpose of interpreting a work is to understand it better and that there is one best understanding that can in principle be attained. Notice there are two claims here: one about aim, one about number. These provide two further parameters about which literary theorists disagree.

Regarding the proper aim of interpretation, there are a variety of views. We have already mentioned one: understanding (Carroll 1992, Iseminger 1992, Juhl 1980, Margolis 1980, Stecker 2003). In some works, it is just difficult to grasp what is going on, and this can happen at all sorts of levels. A work can be hard to understand because of its historical or cultural distance from its audience. Alternatively, features of its style may make it difficult. There are poems where it is hard to understand what the individual lines mean. There are novels and stories where it is hard simply to follow the plot. There are others where, while it is clear that a certain series of events have transpired, there are different ways in which one could understand their significance in the story. More commonly, one knows what happens in a story or what the lines of a poem say, but one does not grasp their point or the point of various bits. There are many other ways in which one may feel one's understanding of a work is inadequate, but in all such cases one turns to interpretations of a work for greater clarity.

An alternative to understanding as the aim of interpretation, is appreciation. (Davies 1982, Goldman 1990, Lamarque 2002). The point of interpretation on this view is to create ways of taking works that enhance their aesthetic value, or that guide the reader to an appreciative experience. Just whether, and precisely how, these two aims really differ is debatable: How can one lead a reader to an appreciative experience, without offering a way of understanding a work by organizing certain features of it around a theme, by describing a character as representing a type of person, identifying the point of a series of images, and so on? The difference may be in the way one evaluates interpretations. If one's aim is understanding, perhaps one hopes to get things right, to give a correct or true interpretation, whereas if one aims to enhance the value of the work or an experience of it, the test of an interpretation is in the aesthetic enjoyment it offers to readers.

Those who think the aim of interpretation is enhanced appreciation, also tend to be pluralists about the number of acceptable interpretations a work can bear. Interpretations that are considered acceptable within this camp range from those strictly constrained by conventions in place when the work was created (Davies 1996) to a virtual free play with a text (Barthes 1989). Among those who claim that the aim of interpretation is understanding, some, such as M. Beardsley and E.D. Hirsch, are monists arguing there is a single ultimately correct understanding of a work, whereas others are pluralists. A number of writers argue that meaning is relative to the constantly changing historical moment in which the work is received (Gadamer 1975, Margolis 1980), to the responses of readers in the face of textual indeterminacy (Iser 1980), or to the assumptions of critical communities (Fish 1980; Carrier 1991).

All such relativist views imply pluralism regarding correct understanding, although pluralism does not imply relativism. An alternative to relativism about a work's meaning is a pluralism about the acceptable aims

of interpretation (Stecker 1997 and 2003). Not all interpretation aims at recovering the meaning of a work. Some legitimately aims at enhancing appreciation, making a work significant to a contemporary audience, or to filling in indeterminacies in optional ways. These projects can clearly be pursued in a plurality of ways. By contrast if one's aim is to recover the intention with which the work was made, that may be a more monistic project. Perhaps, among these interpretive aims, there is one that attempts to identify a historically correct understanding of a work. There are currently a variety of proposals about what this might be (Carroll 2000, Levinson 1996, Stecker 2003).

THE VALUE OF FICTION AND LITERATURE

At the beginning of this entry, we noted that, though fiction and literature are not the same thing, the paradigmatic literary forms today are all types of fiction: poetry, the novel, the short story, and the drama. The question we raised then and turn to now is what it is about the value of literature that makes fictional work the most typical to possess that value. Is it that the value of fiction and literature tend to overlap?

The philosophical debate about the value of literature might be aptly described as between those who answer this last question affirmatively and those who answer it negatively. Fiction, clearly, can serve all sorts of purposes, and we might value it for its function in almost any of these. The chief vehicle by which it achieves these valuable purposes is imaginative engagement (i.e., the make-believe that is intimately involved in the reception of fiction). Whether or not imaginative engagement is valuable in itself, it can quickly lead to things we clearly value (e.g., the pleasure of following a story and imaginatively participating in its world).

In addition to such pleasures, imaginative engagement can also be valuable in other ways. It is plausible that it can enhance valuable abilities: to make fine discriminations, to put ourselves in the shoes of others (to empathize), and to refine the ability to identify emotional and other psychological states. A fiction also might at least contribute to acquiring propositional knowledge. What is true in a fictional world is commonly at least possibly true in the actual world. Thus we can acquire knowledge of possibilities or conceptions of how things may be. A fiction may strongly suggest that something is not only possible, but that it actually is that way, and this may help us to learn about the way things not only might be, but are.

Clearly, all of these valuable traits of fiction can be possessed by literary works, fictional or not, but we can go further and say that literary fictions are the most likely to possess, in the highest degree, the cognitive values just mentioned. While not everyone would accept this, the more controversial issue concerns whether such traits add to the literary or artistic value of these works. A view that denies this claims instead that literary value resides wholly in the aesthetic experience a work offers, where this experience is fairly narrowly conceived. For example, one view that has been vigorously defended is that the aesthetic value of a work lies in its ability to create a complex form that explores a theme of perennial human interest (Lamarque and Olsen 1994). The appreciative experience, which determines the extent to which a work possesses aesthetic value, consists in following the development of the theme in the complex formal structure of the work. What is no part of the literary value is any insight the work might offer regarding the truth about the issues it explores.

This view has the virtue of serving as a corrective to the rejection of the relevance of the aesthetic, even suspicion about its place among the central human values, that has infected large swaths of literary theory (Eagleton 1983, Scholes 1978). However, even as an account of the aesthetic value of literature, it is far too narrow. For one thing, the perennial themes—fate, free will, nature versus nurture—just are not the organizing features of all literature. Some works are more concerned with characters, some with telling a riveting story, some with exhibiting an emotion, some with precise description, and so on. Perhaps we can say that every literary work offers a conception of some aspect of human experience, and when it is good literature, it does so in such a way that one can experience what it would be like if that conception were true (Stecker 1997). However, having said this, it becomes fairly obvious that it is perverse to deny that a further way that literature can be valuable is in the cognitive value of the conceptions offered. They can be valuable for getting it right, but also for suggesting new ways of thinking or experiencing, fruitful conjectures, as it were, even if they turn out to be ultimately wrong. After all we value philosophical works for just this reason, and there are many literary works that have overtly philosophical aims.

Just as fiction can be valuable in many ways, pluralism about literary value also seems to be the most sensible view. When literary works are evaluated not only for the aesthetic experience they offer, but the cognitive, ethical, art-historical value that they possess—to mention only some additional parameters that are relevant—we

are still evaluating them as literary works. Those who argue that interpretations of literary works should maximize the opportunities to appreciate them should welcome this point of view because it opens up so many new avenues from which such appreciation can develop.

See also Art, Interpretation of; Derrida, Jacques; Gadamer, Hans-Georg; Hermeneutics; Structuralism and Post-Structuralism.

Bibliography

Adams, Fred, Robert Stecker, and Gary Fuller. *The Semantics of Fictional Names*. Pacific Phiosophical Quarterly. 78 (1997): 128–48.

Barthes, Roland. *The Pleasure of the Text*. New York: Farrar, Strauss, Giroux, 1989.

Beardsley, M. *Aesthetics: Problems in The Philosophy of Criticism*. Harcourt: Brace and World, 1958.

Beardsley, M. *The Possibility of Criticism*. Detroit, MI: Wayne State University Press, 1970.

Carrier, David. *Principles of Art History*. University Park, PA: Penn State Press, 1991.

Carroll, N. "Art, Intention, and Conversation. " In *Intention and Interpretation*, edited by G. Iseminger. Philadelphia: Temple University Press, 1992.

Carroll, N. "Interpretation and Intention." *Metaphilosophy*. 31 (2000): 75–95.

Carroll, N. *The Philosophy of Horror, or Paradoxes of The Heart*. New York: Routledge, Chapman, and Hall, 1990.

Charlton, W. "Feeling for the Fictitious." *British Journal of Aesthetics* 24 (1984): 206–16.

Currie, G. *The Nature of Fiction*. Cambridge, MA: Cambridge University Press, 1990.

Currie, G. "The Paradox of Caring." In *Emotion and The Arts*, edited by M. Hjort and S. Laver. Oxford: Oxford University Press, 1997.

Dadlez, E. *What's Hecuba to Him? Fictional Events and Actual Emotions*. University Park: Pennsylvania State University Press, 1997.

Davies, S. "Interpreting Contextualities." *Philosophy and Literature* 20 (1996): 20–38.

Davies, S. "The Relevance of Painters' and Writers' Intentions." *Journal of Aesthetics and Art Criticism* 41 (1982): 65–76.

Dilworth, J. *The Double Content of Art*. Amherst, NY: Prometheus Books, 2004.

Eagleton, T. *Literary Theory*. Oxford: Blackwell, 1983.

Emt, J. "On the Nature of Fictional Entities." In *Understanding the Arts*, edited by J. Emt and G, Hermeran. Lund University Press, 1992.

Feagin, S. *Reading with Feeling*. Ithaca, NY: Cornell University Press, 1996.

Fish, Stanley. *Is There a Text in This Class?* Cambridge, MA: Harvard University Press, 1980.

Gadamer, H. *Truth and Method*. New York: Crossroads, 1975.

Gale, R. "The Fictive Use of Language." *Philosophy* 46 (1971): 324–339.

Goldman, Alan. "Interpreting Art and Literature." *Journal of Aesthetics and Art Criticism*. 48 (1990): 205–214.

Gron, E. 1996. "Defending Thought Theory from a Make-Believe Threat." *British Journal of Aesthics* 36 (1996): 311–312.

Hirsch, E. D. *Validity in Interpretation*. New Haven, CT: Yale University Press, 1967.

Howell, R. "Fictional Objects: How They Are and How They Aren't." *Poetics* 8 (1979): 129–177.

Iseminger, Gary. 1992. "An Intentional Demonstration?" In *Intention and Interpretation*, edited by G. Iseminger. Philadelphia: Temple University Press, 1992.

Iser, W. *The Act of Reading: A Theory of Aesthetic Response*. Baltimore, MD: Johns Hopkins University Press, 1980.

Jakobson, R. "Closing Statement: Linguistics and Poetics." In *Style in Language*, edited by T. Soebok. Cambridge, MA: MIT Press, 1960.

Juhl, P. D. *Interpretation: An Essay in The Philosophy of Literary Criticism*. Princeton, NJ: Princeton University Press, 1980.

Lamarque, P. "Appreciation and Literary Interpretation." In *Is There a Single Right Interpretation?*, edited by Michael Krausz. University Park, PA: Penn State Press, 2002.

Lamarque, P. *Fictional Points of View*. Ithaca, NY: Cornell University Press, 1996.

Lamarque, P., and S. Olsen. *Truth, Fiction, and Literature*. Oxford: Oxford University Press, 1994.

Levinson, J. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996.

Lewis, D. "Truth in Fiction." *American Philosophical Quarterly* 15 (1978): 37–46.

Margolis, J. *Art and Philosophy*. Brighton: Harvester Press, 1980.

Nussbaum, M. *Love's Knowledge*. Oxford: Oxford University Press, 1990.

Parsons, T. *Nonexistent Objects*. New Haven, CT: Yale University Press, 1980.

Paskins, B. "On Being Moved by Anna Karenina and Anna Karenina." *Philosophy* 52 (1977): 344–377.

Scholes, R. "Toward a Semiotics of Literature." In *What Is Literature?*, edited by P. Hernandi. Bloomington: Indiana University Press, 1978.

Searle, John. "The Logical Status of Fictional Discourse." *New Literary History* 6 (1975): 319–332.

Stecker, R. *Artworks: Definition, Meaning, Value*. University Park, PA: Penn State Press, 1997.

Stecker, R. *Interpretation and Construction: Art, Speech, and The Law*. Oxford: Blackwell, 2003.

Stecker, R. "What Is Literature?" *Revue Internationale de Philosophie* 50 (1996): 661–694.

Thomasson, Amie. *Fiction and Metaphysics*. Cambridge, MA: Cambridge University Press, 1999.

Urmson, J. O. "Fiction." *American Philosophical Quarterly* 13 (1976): 153–157.

van Inwagen, Peter. "Creatures of Fiction." *American Philosophical Quarterly* 14 (1977): 299–308.

Walton, K. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press, 1990.

Wellek, R., and A. Warren. *The Theory of Literature*. 3rd ed. London: Penguin, 1973.

Wolterstorff, Nicholas. *Works and Worlds of Art*. Oxford: Oxford University Press, 1980.

Yanal, R. *Paradoxes of Emotion and Fiction*. University Park: Pennsylvania State University Press, 1999.

Zalta, E. *Abstract Objects*. Dordrecht: Reidel, 1983.

Zemach, Eddy. *Real Beauty*. University Park: Pennsylvania State University Press, 1997.

Robert Stecker (2005)

LITTRÉ, ÉMILE

(1801–1881)

Émile Littré, the French linguist and positivist philosopher, was born in Paris. From an early age Littré was interested in medicine and languages; and he received training in both. He is now best known for his *Dictionnaire de la langue française* (4 vols., Paris, 1863–1872) and his edition (with Charles Robin) of Pierre Hubert Nysten's *Dictionnaire de médecine, de chirurgie, de pharmacie, de l'art vétérinaire et des sciences qui s'y rapportent* (Paris, 1885). He was also prominent in radical political journalism (in *Le national* of Armand Carrel) and in freethinking circles. He became a member of the Académie des Inscriptions in 1838 and of the Académie Française in 1871, the latter over the violent objections of Bishop Dupanloup of Orléans. Littré was elected a deputy in 1871 and a senator for life in 1875.

These various activities and contacts enabled Littré to be unusually successful in his principal philosophical activity, the propagation of Auguste Comte's Positivism. He began to read Comte's *Cours de philosophie positive* in 1840, wrote a series of articles on it in *Le national* in 1844 and 1845 (published separately under the title *De la philosophie positive* in 1845 and later reprinted in his *Fragments de philosophie positive et de sociologie contemporaine* in Paris in 1876), and for a time became Comte's "principal disciple" and heir apparent as Director of Positivism and High Priest of the Religion of Humanity. Littré broke with Comte in 1852, however, over a combination of personal and political disagreements. Thereafter he took an increasingly independent line on Comte's doctrine as well, forming a loose group of disciples—distinct from the orthodox Comtian school—that found its principal expression in the journal *La philosophie positive,* started by Littré (with G. N. Vyrubov, the Russian positivist) in 1867. Littré himself contributed numerous important articles to the journal, but his position is stated most clearly in his *Auguste Comte et la philosophie positive* (Paris, 1863).

Littré's fundamental proposition was that during the 1840s, partly for personal reasons, Comte had abandoned the positive method for the sake of a "subjective" method that vitiated all his subsequent work. Littré proposed to cleanse Positivism of the "aberrations" of Comte's "second career" by propagating the doctrine in the pure, scientific form of the *Cours.* He insisted that "there is only one stable point and that is science." Positivism as a scientific philosophy is in one aspect a system, "which comprehends everything that is known about the world, man, and societies," and in another aspect a method, "including within itself all the avenues by which these things have become known." It has, however, a practical purpose as well: to provide a "demonstrable rallying point" and a "definite direction" for humankind. Littré differed from Comte in doubting whether Positivism was yet sufficiently advanced to serve as a basis for social and political action. He also, among other things, denied ethics its place at the apex of the hierarchy of the sciences, which Comte in his later years had given it; for Littré, ethics was not an autonomous science at all. On the other hand, Littré was inclined, against Comte, to admit psychology as an independent discipline. Littré remained committed to the evolution of the positivist Religion of Humanity into a "spiritual power" but rejected Comte's prescriptions for its actual institutionalization.

Littré and his group often found it difficult to elaborate a consistent doctrine, largely because Comte's system had in fact been conceived as a unity very early in his career, and it was therefore wrong and illogical to divide his life and work in half.

See also Comte, Auguste; Positivism; Psychology.

Bibliography

Littré's important works also include *Conservation, révolution et positivisme* (Paris, 1852; 2nd ed., 1879) and *La science au point de vue philosophique* (Paris, 1873).

For information on Littré, see É. Caro, *M. Littré et le positivisme* (Paris, 1883), which is hostile.

W. M. Simon (1967)

LLOYD, GENEVIEVE

(1941–)

Born in Cootamundra in New South Wales, Australia, Genevieve Lloyd studied philosophy at the University of Sydney and then at Oxford. Her DPhil, awarded in 1973, was on Time and Tense. From 1967 until 1987 she lectured at the Australian National University, and it was during this period that she developed her most influential

ideas and wrote *The Man of Reason: "Male" and "Female" in Western Philosophy*, which was published in 1984. In 1987 she was appointed to the Chair of Philosophy at the University of New South Wales and was the first female professor of philosophy appointed in Australia.

Lloyd's contribution to feminist thought owes a good deal to Simone de Beauvoir. This is despite the fact that in *The Man of Reason* she is critical of Beauvoir's adoption of the pursuit of transcendence as the ideal of human excellence. Lloyd argues in this book that the historical notion of transcendence involves overcoming the body, which is represented as feminine, and so is a suspect value for women. At the same time her analysis of the symbolic meaning of philosophical concepts echoes Beauvoir. She follows Beauvoir in representing symbols as fundamentally dualistic, citing the Pythagorean table of opposites alluded to by Beauvoir in *The Second Sex*. Both agree that for the Pythagoreans, the male is associated with order and the right, light, and rational realm while the female corresponds to chaos and the left, dark, and irrational side of being. In an article published in *Australian Feminist Studies* in 1989, Lloyd explains that when Beauvoir speaks of woman as *other* she "is talking about the way culture has constructed the feminine—about its symbolic content" (p. 17). Likewise, Lloyd has been concerned with the ramifications of male power in the construction and control of symbolic structures. Unlike Beauvoir, however, she finds problematic the adoption by women of values traditionally symbolized as masculine. Yet she also shies away from a full endorsement of those strands of feminism of difference, which celebrate the body, emotion, and unreason as sources of essentially female values.

Though emphasizing the metaphorical association of reason with the male and reason's opposites and inferiors with the female, Lloyd is careful to avoid claiming that reason is *literally* male. In her concluding remarks to *The Man of Reason*, she says: "The claim that Reason is male need not at all involve sexual relativism about truth, or any suggestion that principles of logical thought valid for men do not hold also for female reasoners" (Lloyd 1984, p. 109). Nevertheless, she wants to avoid treating the maleness of reason as a *mere metaphor* that can easily be stripped away from the ideal of rationality. Alluding to Michèle Le Doeuff's (1948–) claim that the metaphors and images used by philosophers constitute a philosophical imaginary of marginalized tropes integral to the commitments of a text, she undermines the distinction between the literal and metaphorical. Elsewhere she evokes Jacques Derrida's deconstruction of the philosophical distinction between literal truth and metaphorical embellishment. However, Lloyd has not developed a detailed analysis of the relationship between metaphor, literal truth, rational argument, and literary effect, and this lends a certain obscurity to her position.

Despite having inspired Lloyd's line of argument, Le Doeuff has been critical of Lloyd's analysis of Francis Bacon's metaphors, arguing that the association between reason and masculinity discussed by Lloyd and found in the twentieth-century translation of Bacon is not to be found in Bacon's Latin original. She suggests in *The Sex of Knowing* that in general, historical claims that women are irrational are (false) literal claims intended to undermine women's intellectual authority.

Lloyd argues that Cartesian dualism is particularly problematic for feminism, and in her edited collection *Feminism and History of Philosophy*, sums up this suspicion. "What made the Cartesian philosophy suspect for feminists was its association with the doctrine of dualism—the rigid separation of minds and bodies as utterly distinct kinds of being. The dichotomy came to be seen as reinforcing the denigration of women, in association with the body, in opposition to the ideal of reason associated with "male 'transcendence'" (Lloyd 2000, p. 9). In her later work, Lloyd urges the fruitfulness for feminism of Benedict de Spinoza's treatment of the mind as an idea of the body, which she interprets as an ontological doctrine that undermines the polarities of the Cartesian tradition. During the 1990s she turned to working on Spinoza and published a number of books on his thought.

It is nevertheless questionable whether Cartesian dualism is literally a suspect metaphysical doctrine for feminists or whether Spinoza's form of monism would serve women better. Feminist historians such as Margaret Atherton (1943–) and Hilda Smith (1941–) have argued that historically, dualism has favored feminism. Even prior to René Descartes, women such as Christine de Pizan (1365–1431) were able to point to the immateriality of the soul as evidence that women's souls were the same as men's and so women were men's spiritual equals. Moreover, Descartes's method, with its reliance on reason and clear and distinct ideas, was accessible to women who had not had a university education. Descartes was not himself a misogynist; he took seriously the arguments of his correspondent the Princess Elizabeth of Bohemia (1630–1714), and his philosophy ushered in a period during which significant numbers of women engaged with the new philosophy.

The impact the perennial but by no means universal association between the mind and a masculine master ought to have on one's views concerning the literal mate-

riality of the soul remains obscure. Lloyd's seminal critique of the rhetoric of the male philosophical tradition has been widely influential. The consequences that one should draw from that critique, and its significance for feminism and metaphysics, remain contested.

See also Bacon, Francis; Beauvoir, Simone de; Cartesianism; Derrida, Jacques; Descartes, René; Pythagoras and Pythagoreanism; Spinoza, Benedict (Baruch) de.

Bibliography

Atherton, Margaret. "Cartesian Reason and Gendered Reason." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte Witt, 19–34. Boulder, CO: Westview Press, 1993.

Beauvoir, Simone de. *The Second Sex*. Translated by H. M. Parshley. Harmondsworth: Penguin, 1983.

Broad, Jacqueline. *Women Philosophers of the Seventeenth Century*. Cambridge, U.K.: Cambridge University Press, 2002.

Le Doeuff, Michèle. *The Sex of Knowing*. London: Routledge, 2003.

Lloyd, Genevieve, ed. *Feminism and History of Philosophy, Oxford Readings in Feminism*. Oxford: Oxford University Press, 2002.

Lloyd, Genevieve. " Feminism in History of Ideas: Appropriating the Past." In *The Cambridge Companion to Feminism in Philosophy*, edited by Miranda Fricker and Jennifer Hornsby, 165–172. Cambridge, U.K.: Cambridge University Press, 2000.

Lloyd, Genevieve. "Maleness, Metaphor, and the Crisis of Reason." In *A Mind of One's Own*, edited by Louise M. Antony and Charlotte Witt, 69–83. Boulder, CO: Westview Press, 1993.

Lloyd, Genevieve. *The Man of Reason: "Male" and "Female" in Western Philosophy*. London: Methuen, 1984.

Lloyd, Genevieve. *Part of Nature: Self-knowledge in Spinoza's Ethics*. Ithaca, NY: Cornell University Press, 1994.

Lloyd, Genevieve. "Rationality." In *A Companion to Feminist Philosophy*, edited by Alison M. Jaggar and Iris Marion Young, 165–172. Oxford: Blackwell, 1998.

Lloyd, Genevieve. *Spinoza and the Ethics, Routledge Philosophy Guidebook*. London: Routledge, 1996.

Lloyd, Genevieve. "Texts, Metaphors, and the Pretensions of Philosophy." *Monist* 69 (1986): 87–102.

Lloyd, Genevieve. "Woman as Other: Sex, Gender, and Subjectivity." *Australian Feminist Studies* 10 (1989): 13–22.

Pizan, Christine de. *The Book of the City of Ladies*. Translated by Earl Jeffrey Richards. London: Picador, 1983.

Smith, Hilda L. *Reason's Disciples: Seventeenth Century English Feminists*. Urbana: University of Illinois Press, 1982.

Karen Green (2005)

LOCKE, JOHN

(1632–1704)

John Locke, English empiricist and moral and political philosopher, was born in Wrington, Somerset. Locke's father, an attorney and for a time a clerk to the justices of the peace in Somerset, fought on the parliamentary side in the first rebellion against Charles I. Locke was reared in a liberal Puritan family and early learned the virtues of temperance, simplicity, and aversion to display. Though his father was severe and remote from him in early youth, as Locke matured they became close friends.

In 1646 Locke entered Westminster School, where he studied the classics, Hebrew, and Arabic. Little time was given at Westminster to science and other studies, and its harsh discipline, rote learning, and excessive emphasis on grammar and languages were later condemned by Locke.

In 1652 Locke was elected to a studentship at Christ's Church, Oxford. He received his BA in 1656 and remained in residence for the master's degree. He was not happy with the study of Scholastic philosophy and managed to inform himself of many new areas of thought. As a master, Locke lectured in Latin and Greek and in 1664 was appointed censor of moral philosophy.

His father's death in 1661 left Locke with a small inheritance and some independence. During these years he became acquainted with many men who were to have a profound influence upon his life. From Robert Boyle, Locke learned about the new sciences and the corpuscular theory, as well as the experimental and empirical methods. Confronted with the choice of taking holy orders, continuing as a don, or entering another faculty, Locke chose medicine. Though well trained, he never practiced medicine, nor was he permitted to take the medical degree, which would have permitted him to teach the profession, until 1674, although in 1667 he began to collaborate with the great physician Thomas Sydenham.

In 1665 Locke was sent on a diplomatic mission accompanying Sir Walter Vane to the elector of Brandenburg at Cleves. He subsequently rejected a secretaryship under the earl of Sandwich, ambassador to Spain, and returned to Oxford. It was at this time that his interests began to turn seriously to philosophy. Descartes was the first philosopher whom Locke enjoyed reading and the first to show him the possibility of viable alternatives to the Schoolmen.

Locke had met Lord Ashley, earl of Shaftesbury, in 1662 at Oxford. They found much pleasure in each other's company, and the astute Shaftesbury quickly recognized Locke's talents. In 1667 he invited Locke to live

with him in London as his personal physician. Later Locke served him well in many other capacities. Under Shaftesbury Locke found himself in the center of the political and practical affairs of the day. He assisted Shaftesbury in the framing of a constitution for the colony of Carolina. For a time he was secretary for the presentation of benefices and then secretary to the Council of Trade and Plantations. Locke was always at home in the world of practical affairs, and many of his philosophical attitudes reflect this interest. At the same time he became a fellow in the Royal Society, where he continued to be in touch with learning.

Locke, never robust in health, in 1675 went on a prolonged visit to France, where he made many friends and came into contact with the foremost minds of his day. His studies and criticisms of Descartes were deepened under the influence of various Gassendists.

In 1679 Locke returned to an England torn by intense political conflicts. Shaftesbury, who had become the leader of the parliamentary opposition to the Stuarts, alternated between political power and impotence. The close association with Shaftesbury brought Locke under suspicion; he was kept under surveillance. Shaftesbury was tried for treason in 1681, but acquitted. He subsequently fled England for Holland, where he died in 1683. Locke, at Oxford, uncertain of his position and fearing persecution, also fled England, arriving in Holland in September 1683. The king had demanded that Locke be deprived of his studentship at Oxford, and news of this demand caused Locke to prolong his stay. After the death of Charles II and the ascension of James II to the throne, the duke of Monmouth attempted a rebellion, which failed. Locke was denounced as a traitor, and the crown demanded of the Dutch that he be returned to England. No great effort was made to comply with the demand, and Locke remained in Holland.

During his stay in Holland, Locke again acquired a wide circle of distinguished friends and wrote extensively. He contributed an article as well as reviews to the *Bibliothèque universelle* of Jean Leclerc; these were his first published works. He wrote in Latin the *Epistola de Tolerantia*, which was published anonymously in 1689 and translated as the *First Letter concerning Toleration*. He also worked assiduously on *An Essay concerning Human Understanding*, which he had been writing off and on since 1671. In 1688 the *Bibliothèque universelle* published an abstract of the *Essay*.

These activities did not prevent him from being deeply engaged in politics. The plot to set William of Orange on the throne of England was well advanced in 1687, and Locke was, at the very least, advising William in some capacity. The revolution was accomplished in the fall of 1688, and in February 1689 Locke returned to England, escorting the princess of Orange, who later became Queen Mary.

In 1689 and 1690 Locke's two most important works, *An Essay concerning Human Understanding* and *Two Treatises of Government*, were published. From 1689 to 1691 Locke shuttled between London and Oates, the home of Sir Francis and Lady Masham, the daughter of Ralph Cudworth. He had declined an ambassadorial post only to accept a position as commissioner on the Board of Trade and Plantations. Apparently his practical wisdom was invaluable, for when he wished in 1697 to resign because of ill health, he was not permitted to do so. He remained until 1700, serving when he could, although his health was extremely poor.

In 1691 Locke made Oates his permanent residence at the invitation of Lady Masham. It was, for the aging Locke, a place of refuge and joy; there he received visits from Newton, Samuel Clarke, and others. These were productive years for Locke. *Some Thoughts concerning Education* appeared in 1693. The second edition of *Essay* was published in 1694. In the following year *The Reasonableness of Christianity* was published anonymously. He answered criticism of it in *A Vindication of the Reasonableness of Christianity* (London, 1695) and in a second *Vindication* in 1697. From 1697 to 1699 Locke engaged in an epistolary controversy with Edward Stillingfleet, bishop of Worcester.

However, Locke's health steadily failed him. After 1700, when the fourth edition of *Essay* appeared, he remained almost constantly at Oates. He was engaged in editing *Two Treatises of Government*, for no edition which pleased him had yet appeared. In his last years he wrote extensive commentaries on the epistles of St. Paul, which were published posthumously. On October 28, 1704, while Lady Masham was reading the Psalms to him, Locke died. Lady Masham wrote of him, "His death was like his life, truly pious, yet natural, easy and unaffected."

CHARACTER. The Lovelace Collection of Locke's personal papers in the Bodleian Library, Oxford, shows that Locke's character and personality were more complex than had been suspected. The great affection and respect which so many men and women had for him are testimony to his charm and wisdom. That he was modest, prudent, pious, witty, and eminently practical was long known. But he was also extremely secretive and apparently given to excessive suspicion and fears. When his life-

long friend, James Tyrrell, voiced his suspicion that Locke had written *Two Treatises*, Locke was evasive and would not admit the fact. When he suspected that Tyrrell was spreading the report that Locke was the author, Locke angrily demanded an explanation. At the same time, Locke showed great affection for many friends and a real fondness for children. In maturity he could not abide religious intolerance or suffer tyranny. He was passionately devoted to truth and strove constantly to state the truth as he saw it, but always with a caution that distrusted all dialectic, even his own, when it appeared to go beyond common sense.

INFLUENCES ON LOCKE. Locke's philosophy is grounded in medieval thought, though he, like Descartes, turned away from it as far as possible. The Cambridge Platonists, notably Ralph Cudworth and Benjamin Whichcote, influenced him greatly with respect to religious tolerance, empirical inquiry, and the theory of knowledge. Locke was indebted to Richard Hooker in his political thought. Hobbes probably influenced him somewhat, though Locke was concerned not to be classed as a Hobbist. The two most important philosophical influences upon him were Descartes and Pierre Gassendi. From Descartes he learned much that is incorporated in *Essay*, and in Gassendi and the Gassendists he found support to challenge the doctrine of innate ideas and the radical rationalistic realism of Descartes. Gassendi helped to convince Locke both that knowledge begins in sensation and that intellect, or reason, is essential to the attainment of truth and knowledge.

AN ESSAY CONCERNING HUMAN UNDERSTANDING

Locke's position in the history of Western thought rests upon *An Essay concerning Human Understanding* and *Two Treatises of Government*. He spent long years working out the thought of each, and he carefully and lovingly revised and corrected them for subsequent editions. Locke wrote two drafts of his *Essay* in 1671, and in 1685 he wrote a third. The first edition, though dated 1690, appeared in late 1689. During the years between 1671 and 1689, Locke revised and reorganized many of his original concepts. In response to criticisms of the first edition of *Essay*, he introduced a number of changes in subsequent editions. This long period of gestation and Locke's subsequent modifications of his initial public statement disclose primarily the refinement and clarification of his philosophy by way of certain important additions, but never by a radical or fundamental departure from his basic position.

From the first appearance of *An Essay concerning Human Understanding* Locke was criticized for being inconsistent in his theory of knowledge, vague in the presentation and development of many of his ideas, and wanting in thoroughness in developing other ideas. But these criticisms have in no way diminished either the importance or the influence of *Essay* on subsequent thinkers. By no means the first of the British empiricists, Locke nonetheless gave empiricism its firmest roots in British soil, where it still proudly flourishes. It must be remembered that Locke was also a rationalist, though one of quite different orientation from such Continental thinkers as Descartes, Spinoza, and Malebranche. In Locke many strands of traditional thought are rewoven into a new fabric. Subsequent thinkers, notably Berkeley, Hume, and Kant, perhaps fashioned more coherent and consistent systems, but it is doubtful whether they were more adequate to what Locke might have called the plain facts.

Locke's tendency toward inconsistency can be seen in his definition of knowledge as "the perception of the connection and agreement, or disagreement and repugnancy, of any of our ideas" (*Essay*, IV.i.2). This is plainly incompatible with his later contention that we have intuitive knowledge of our own existence, demonstrative knowledge of God's existence, and sensitive knowledge of the existence of particular things. Nonetheless, Locke would not abandon his position for the sake of consistency alone. He was persuaded that common sense and the facts justified his conviction and that whatever faults there were in his position lay in the difficulty of stating a coherent theory of knowledge, not in the reality of things. If this made him an easy prey to a skillful dialectician, like Berkeley, it also left him closer to the common conviction of most of us when we think about anything other than epistemology. It is this viewpoint, almost unique in philosophy, that accounts for the abiding interest in Locke's thought and the great extent of his influence despite the shortcomings of his work.

PURPOSE OF AN ESSAY. In "Epistle to the Reader" Locke related that some friends meeting in his chamber became perplexed about certain difficulties that arose in their discourse about a subject (left unnamed). He proposed that before they could inquire further, "it was necessary to examine our own abilities and see what objects our understandings were, or were not, fitted to deal with." This discussion in 1670 or 1671 first started Locke on the inquiries that were to continue intermittently for twenty years. What Locke first set down for the next meeting is not known, unless it was Draft A (1671) of *An Essay con-*

cerning Human Understanding. That the initial suggestion became the abiding purpose of *Essay* is clear from Locke's assertion that his purpose was "to inquire into the original, certainty, and extent of human knowledge, together with the grounds and degrees of belief, opinion, and assent" (I.i.2). At the same time he disavowed any intention to examine "the physical consideration of the mind, … wherein its essence consists, or by what motions of our spirits or alterations of our bodies we come to have any sensation by our organs or any ideas in our understandings, and whether those ideas do in their formation any or all of them depend on matter or no" (I.i.2).

Locke did not, in fact, offer any detailed or explicit accounts of these matters. He would have considered that a subject for natural philosophy. Nonetheless, he did, as indeed he had to, deal with the physical considerations of the mind, as well as all the other matters mentioned.

From the outset Locke was persuaded that our understanding and knowledge fall far short of all that exists; yet he was equally certain that men have a capacity for knowledge sufficient for their purposes and matters enough to inquire into. These convictions, pragmatic and utilitarian, set Locke apart from most of the other major philosophers of the seventeenth century, who, impressed by the new developments in mathematics and the new physical sciences, boldly plunged ahead with a rationalistic realism in the belief that their new methods would enable them in large measure to grasp reality. Locke saw that the very advances made in the new sciences put reality farther from the reach of the human mind. This did not make Locke a nominalist or an idealist in any modern sense; rather, he persistently affirmed the real objective existence of things or substances. What he denied was that the human understanding could know with certainty the real essences of substances. If "ideas" stand between reality and the understanding, it is to link them, even if only under the form of appearances. It is not to obliterate any connection between them or to justify a negation of substance—God, mind, or matter.

IDEAS. The key term in Locke's *Essay* is "idea," which he defined as "… whatsoever is the object of the understanding when a man thinks, … whatever is meant by phantasm, notion, species, or whatever it is which the mind can be employed about in thinking" (I.i.8). Any object of awareness or of consciousness must be an idea. But then how can we have any knowledge of anything other than ideas and their relationships? It is true that Locke spoke of ideas as the "materials of knowledge." Yet knowledge itself, when possessed and made the object of the mind, must be an idea. For example, to perceive that *A* is equal to *B* is to perceive the agreement between *A* and *B*. This agreement as perceived must be an idea, or it cannot be an object of the mind when it thinks. Despite this difficulty Locke clung tenaciously to his term "idea" in his disputes with Stillingfleet. He actually intended something other than he stated, namely, that knowledge is an operation, an activity of the mind, not initially one of its objects. It would have served his purpose better had he spoken of "knowing" rather than of "knowledge," even though this would not have entirely removed the difficulty, since to set the mind at a distance where we may look at it, in order to know what knowledge is, is still to have an idea.

Locke, however, went beyond ideas to assume the real existence of things, substances, actions, processes, and operations. Ideas, except when they are the free constructs of the mind itself, signify and represent, however imperfectly, real existences and events. So deep was Locke's conviction on this point that no argument could shake him, although he constantly tried to remove the difficulties implicit in his definitions of "ideas" and "knowledge." This conviction is evident in the first two books of *Essay*, in which Locke inquired into the origin of our ideas.

NO INNATE IDEAS. It was Locke's central thesis, developed extensively in Book II of *Essay*, that we get all our ideas from experience. The whole of the first book is given to an overlong criticism, at times not germane to the subject, of the doctrine that we have innate ideas and innate knowledge.

Locke contended that there are no innate principles stamped upon the mind of man and brought into the world by the soul. In the first place, the argument that people have generally agreed that there are innate ideas, even if true, would not demonstrate the innateness of ideas. Moreover, there are no principles to which all give assent, since principles such as "Whatever is, is" and "It is impossible for the same thing to be and not to be" are not known to children, idiots, and a great part of mankind, who never heard or thought of them. Locke here assumed that innateness was equivalent to conscious perceiving and argued that to be in the mind is to be perceived or to be readily recalled to perception. Locke allowed that there is a capacity in us to know several truths but contended that this lent no support to the argument that they are innate.

To argue that all men know and assent to certain truths when they come to the use of reason proves noth-

ing, since they will also come to know many truths that are not innate. It would appear, then, that all truth is either innate or adventitious. Again, why should the use of reason be necessary to discover truths already innately in the mind? Locke allowed that the knowledge of some truths is in the mind very early, but observation shows such truths are about particular ideas furnished by the senses; for example, a child knows the difference between the ideas of sweet and bitter before it can speak and before it knows abstract ideas. Even assent at first hearing is no proof of innateness, for many truths not innate will be assented to as soon as understood.

On the contrary, the senses first furnish us with particular ideas, which the mind by degrees becomes familiar with, remembers, and names. The mind subsequently abstracts from these particular ideas and gives names to general ideas. Thus, general ideas, general words, and the use of reason grow together, and assent to the truth of propositions depends on having clear and distinct ideas of the meaning of terms. Locke held it to be evident that particular propositions are known before the more universal and with as much certainty.

We have natural faculties or capacities to think and to reason. This is not, however, the same thing as having innate ideas, for if anyone means by innate ideas nothing but this natural capacity, he uses terms, according to Locke, in a manner plainly contrary to common usage.

In a similar fashion, Locke argued that we have no innate moral or practical principles, for there is no universal agreement about such principles; great varieties of human vice have been at one time or place considered virtues. We all have a desire for happiness and an aversion to misery, but these inclinations give us no knowledge or truth. Locke was persuaded that there are eternal principles of morality, which men may come to know through the use of reason about experience. This, however, is far from proving them innate.

In the third chapter of Book I Locke argued that no principles can be innate unless the ideas contained in them are innate, that is, unless men can be conscious of them. Impossibility and identity are hardly innate, yet without them we cannot understand the supposedly innate principle of identity, that it is impossible for the same thing both to be and not be. Similarly, the proposition that God is to be worshipped cannot be innate, for the notion of God is so diverse that men have great difficulty agreeing on it, while some men have no conception of God whatsoever.

Locke's target. Who was Locke criticizing in his long and repetitious attack on the doctrine of innate ideas? Was the position he denounced held by anyone in the form in which he presented the theory? Why did he examine the question at such length?

Since *Essay* was first published tradition has held that Locke's target was Descartes and the Cartesians. Certainly Leibniz thought so, as did others after him. In the late nineteenth century, critics pointed to Locke's own rationalism and noted that his recognition of men's natural faculties and innate powers to think and reason is not far from the position of Descartes, who wrote, "Innate ideas proceed from the capacity of thought itself," and "I never wrote or concluded that the mind required innate ideas which were in some sort different from its faculty of thinking." Various other possible objects of Locke's attacks were suggested, the Cambridge Platonists, certain groups in the universities, and various clergymen. Recently R. I. Aaron has argued persuasively that the older tradition, that Descartes, the Cartesians, and certain English thinkers were the targets of Locke's attack, is the correct one and that Locke was not simply striking at a straw man of his own making.

Reasons for attacking innate ideas. Locke suggested that the doctrine of innate ideas lends itself to a certain authoritarianism and encourages laziness of thought, so that the foundations of knowledge are not likely to be examined. The expression "innate ideas" is an unfortunate one and admittedly extremely vague. It carries with it the suggestion that certain ideas and knowledge are, in Locke's sense, imprinted on the mind and are in no way dependent on experience. Certainly there are passages in Descartes which strongly suggest that certain ideas are innately in the mind, and more than a few thinkers took this to be Descartes's meaning. Furthermore, Locke wished to prepare the ground for his own thesis that all ideas and all knowledge are acquired. If he overemphasized the crude sense of the theory of innate ideas, he also showed that even the refined doctrine is unnecessary in accounting for knowledge.

There is another point that Locke discussed later in *Essay*. Descartes asserted that the essence of the mind is to think. To Locke this meant that the mind could not both be and not think. He argued that the mind does not think always and that its real essence cannot be thinking. If the mind thinks always, either some ideas must be innate or the mind comes into being only after it has been furnished with ideas by experience. Neither alternative was acceptable to Locke.

SOURCE OF IDEAS. Locke, in his positive thesis in Book II, valiantly and sometimes awkwardly endeavored to show that every idea we have is ultimately derived from experience, either from sensation or reflection. Locke began by asserting that a man is conscious of two things, the fact "that he thinks" and "the ideas" in the mind about which he thinks. Locke's initial concern was with the question of how a man comes by his ideas; and he made an assumption in terms of several similes. "Let us then suppose the mind to be, as we say, white paper, void of all characters, without any ideas. How comes it to be furnished? … Whence has it all the materials of reason and knowledge?" (II.i.2). Locke replied to his own questions that we get all our ideas from experience, the two fountainheads of which are sensation and reflection. Our senses are affected by external objects (bodies) and afford us ideas, such as yellow, white, heat, cold, soft, hard, bitter, and sweet. Perceiving the operations of our own minds when we reflect, we are furnished with ideas of perception, thinking, doubting, believing, reasoning, knowing, and willing.

The ideas that are furnished by experience are the materials of reason and knowledge. These materials are either the immediate objects of sense, such as color, or the unexamined but direct awareness of such acts as doubting or knowing. Locke's meaning becomes explicitly clear in his account of solidity. He held that we get the idea of solidity by touch. "That which … hinders the approach of two bodies, when they are moving one towards another, I call solidity" (II.iv.1). He sharply distinguished this sense from the purely mathematical use of the term. Impenetrability is an acceptable alternative name for solidity. It is clearly distinct from space and hardness. After an extensive discussion Locke stated, "If anyone asks me what this solidity is, I send him to his senses to inform him. Let him put a flint or a football between his hands and then endeavour to join them, and he will know" (II.iv.6). All philosophical and scientific discourse about solidity, however complex and sophisticated it may be, must ultimately refer back to that from which it began, namely the experience or sensation we have when we put something such as a flint or a football between our hands. Similarly, we cannot by discourse give a blind man the idea of color or make known what pain is to one who never felt it. All knowledge about the physics of light and color or sound refers back to what we perceive when we see and hear. It is in this sense, then, that we get all our ideas from sensation and reflection. Locke nowhere, however, suggested that we can or should stop there. Once the mind is furnished with ideas, it may perform various operations with them.

IDEAS AND THE REAL WORLD. Throughout the first book of *Essay* Locke assumed the real existence of an external physical world and the substantial unity of a man in body and mind. He undoubtedly accepted the thesis that the external physical world is corpuscular and acts by bodies in motion that possess only those qualities which Locke called primary. Locke spoke of secondary qualities as powers in bodies to produce in our minds ideas that are signs of these powers but that in no way resemble the powers that produce them. Often he suggested that if we had the means of observing the minute motions of the particles making up gross bodies, we might have a clearer notion of what we mean when we call secondary qualities powers. Locke's position here is physical realism. It is not simply a manner of speaking. The ideas we have do represent real things outside of us and do constitute the links by which we know something of the external physical world.

Identity. Among the bodies that exist are those of plants, animals, and men. Existence itself constitutes the principle of individuation. Identity is not applied in the same way to a mass of matter and a living body. The identity of an oak lies in the organization of its parts, which partake of one common life. So it is with animals. Again, "the identity of the same man consists: viz. in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body" (II.xxvii.6).

Origin of sensation. With these controlling hypotheses in *Essay* in view, we may return to Locke's invitation to consider the mind as a blank sheet of paper without any ideas. Is a mind without ideas anything but a bare capacity to receive ideas? If we ask what a man is without ideas, we can say he is an organized body existing in a world of other bodies and interacting with them. Experience is a matter of contact of the organized human body with other bodies before it is a matter of sensation or perception. Not every body impinging on our body gives rise to sensation; if it does not, we take no notice of it. However, if some external bodies strike our senses and produce the appropriate motions therein, then our senses convey into the mind several distinct perceptions. How this takes place Locke avoided considering, but that it takes place he was certain; a man, he asserted, first begins to think "when he first has any sensation" (II.i.23).

SIMPLE AND COMPLEX IDEAS. Locke proceeded to distinguish between simple and complex ideas. A simple idea is "nothing but one uniform appearance or conception in the mind, and is not distinguishable into different

ideas" (II.ii.1). A color seen, a sound heard, warmth felt, an odor smelled, are all simple ideas of sense. Once it is furnished with a number of simple ideas, the mind has the power to repeat, compare, and unite them into an almost infinite variety of combinations; but it is utterly incapable of inventing or framing a new simple idea. Thus, with respect to simple ideas the mind is mostly passive; they are simply given in experience. The ideas are given not in isolation from each other but in combinations, as when we simultaneously feel the warmth and softness of wax or the coldness and hardness of ice; nevertheless, simple ideas are distinct from each other in that the mind may mark off each from the other, however united the qualities may be in the things that cause the simple ideas in the mind. Moreover, only those qualities in things that produce ideas in us can ever be imagined at all. Thus, our knowledge of existence is limited by the ideas furnished by experience. Had we one sense less or more than we now do, our experience and knowledge would be respectively decreased or increased.

We have certain ideas, such as color or odor, from one sense only; others, like figure and number, from more than one sense. Reflection alone provides us with experience of thinking and willing. Other ideas, such as pleasure, pain, power, existence, and unity, we have from both sensation and reflection.

PRIMARY AND SECONDARY QUALITIES. Locke made a second basic distinction—between primary and secondary qualities. In doing so he clearly went beyond ideas. He wrote, "Whatsoever the mind perceives in itself, or is the immediate object of perception, thought, or understanding, that I call idea; and the power to produce any idea in our mind, I call quality of the subject wherein that power is" (II.viii.8). Primary qualities, he argued, are utterly inseparable from body. They are known to be primary because sense constantly finds them there if body can be perceived at all, and the mind by critical reflection finds them inseparable from every particle of matter. Solidity, extension, figure, and mobility are all primary qualities. Our ideas of these qualities resemble the qualities themselves, and these qualities really exist in body, whether or not they are perceived. Berkeley was to show that to speak of resemblance supposes that a comparison, an observation, can be made. Locke was aware of the difficulty, as is shown in his *Examination of Malebranche*. Apparently he believed it was the only explanation plausible in spite of its difficulties.

Secondary qualities, in Locke's terms, were nothing but powers to produce various sensations. Bodies do so by the action of their bulk, figure, and texture, and by the motion of their insensible parts on our senses. Somehow they produce in us such ideas as color, odor, sound, warmth, and smell. These ideas in no way resemble the qualities of bodies themselves. They are but signs of events in real bodies. Locke also frequently called these ideas secondary qualities. He would have been clearer had he called them sensory ideas of secondary qualities, preserving the distinction between qualities as attributes of a subject and ideas as objects in the mind. A third class of qualities (sometimes called tertiary) is the power of a body to produce a change in another body, for example, the power of the sun to melt wax.

Nowhere is Locke's physical realism more evident than in his distinction between primary and secondary qualities. Whatever epistemological difficulties the distinction might entail, Locke was persuaded that the new physics required it. Indeed, the distinction was made by Boyle, Descartes, Galileo, and others before him and was thoroughly familiar in his day. Admittedly there is a problem in the assertion that a certain motion in body produces in us the idea of a particular color. Nevertheless, Locke was persuaded that it was so. In such difficult cases Locke fell back upon the omnipotence and wisdom of God and the fact that our knowledge is suited to our purpose.

IDEAS OF REFLECTION. Locke observed that perception is the first faculty of the mind and without it we know nothing else. Hence, the idea of perception is the first and simplest idea we have from reflection. What perception is, is best discovered by observing what we do when we see, hear, or think. Locke added that judgment may alter the interpretation we make of the ideas we receive from sensation. Thus, if a man born blind gains his sight, he must learn to distinguish between a sphere and a cube visually, though he can do so readily by touch. By habit the ideas of sensation are gradually integrated into the unified experience of complex ideas, and by judgment we come to expect things that look a certain way to also feel or smell a certain way. It is worth noting that Locke was persuaded that animals have perception and are not, as Descartes held, mere automatons.

Memory and contemplation. The second faculty of the mind that Locke held indispensable to knowledge is the retention manifested in both contemplation and memory. Contemplation consists in holding an idea before the mind for some time. Memory, however, gave Locke some difficulties. He asserted that "our ideas being nothing but actual perceptions in the mind—this laying

up of our ideas in the repository of the memory signifies no more but this: that the mind has a power in many cases to revive perceptions which it has once had, with this additional perception annexed to them, that it has had them before" (II.x.2). The inadequacy of this statement is at once evident. It proposes no more than a kind of subjective conviction that may often be in error. Locke's analysis of memory was more psychological than philosophical. He passed over the consideration of how memory is possible at all and the criteria by which a true memory may be distinguished from a false memory. He did say, however, that attention, repetition, pleasure, and pain aid memory and are the conditions under which memory is strengthened or weakened. Again he asserted that animals have memory.

Other ideas of reflection. Other faculties of the mind are discerning and distinguishing one idea from another, comparing and compounding, naming, and abstracting. Locke considered each point also in respect to animals, holding, for example, that animals compare and compound ideas only to a slight extent and do not abstract ideas at all. At the conclusion of this chapter (II.xi.15) Locke asserted that he thought he had given a "true history of the first beginnings of human knowledge."

COMPLEX IDEAS. Locke next considered complex ideas. Just as the mind observes that several combinations of simple ideas are found together, so too, it can by its own action voluntarily join several simple ideas together into one complex idea. There are three categories of complex ideas—modes, substances, and relations. Modes are dependencies or affections of substances. Simple modes are variations or different combinations of one simple idea, whereas in mixed modes several distinct ideas are joined to make a complex idea. Ideas of substances represent distinct particular things subsisting in themselves. Complex ideas of relation consist in comparing one idea with another.

This classification is not entirely satisfactory because ideas of modes invariably entail relations in the broadest sense. Locke seems to have been closer to Aristotle than to modern usage in his employment of the term "relation." Under modes Locke included space, duration and time, number, infinity, motion, sense qualities, thinking, pleasure and pain, power, and certain mixed modes. Under substance he placed the idea of substance in general, the ideas of particular substances, and collective ideas of substances. In the category of relation, he considered a number of ideas, including cause and effect, relations of place and time, identity and diversity, and others that he classified as proportional, natural, instituted, and moral.

The greater number of these concepts have in other philosophies been credited with some a priori and extraempirical character. They are not direct objects of sensory experience; and they appear to have a certainty not found in the mere coexistence of sensory ideas. They are more abstract and universal than the simple ideas of sensation and reflection. Locke's broad use of the term "ideas" tends to confuse and obscure the distinction between sensory percept and concept. Nevertheless, Locke undertook to show how the mind actively constructs these complex ideas, abstract and conceptual though they may be, out of the materials of knowledge, the simple ideas of sensation and reflection. In this undertaking Locke's rationalism was most evident, for he held that while the mind constructs complex ideas, it cannot do so arbitrarily. In this sense, Locke could claim for them an objective reality.

The mode of space. Examination here will be limited to only those complex ideas that are most important and difficult. Among modes, only space, duration, number, thinking, and power will be considered. Locke contended that the modifications of a simple idea are as much distinct ideas as any two ideas can be. Space in its first manifestation is a simple idea, since in seeing and touching we immediately perceive a distance between bodies and the parts of bodies. Though the idea of space constantly accompanies other sensory ideas, it is distinguishable from them. All our modes of the idea of space derive from the initial sensory experience. Thus space considered as length is called distance, considered three-dimensionally is capacity, considered in any manner is termed extension. Each different distance, especially when measured by stated lengths, is a distinct idea, including the idea of immensity, which consists in adding distance to distance without ever reaching a terminus. So too, figure allows an endless variety of modifications of the simple idea of space. Place is distance considered relative to some particular bodies or frame of reference.

Locke disagreed with Descartes's assertion that extension is the essence of matter, although he agreed that we cannot conceive of a body that is not extended. But a body has solidity, and solidity is distinct from the notion of space; for the parts of space are inseparable in thought and in actuality and are immovable, whereas a solid body may move and its parts are separable. Descartes's argument that the physical universe is a plenum was dismissed by Locke as unsound, for there is no contradiction in the conception of a vacuum. If body is not infinite, we can

conceive of reaching out beyond the physical limits of the universe to a place unoccupied by matter. The idea of pure space is necessarily infinite, for we can conceive of no limit or terminus to it. Locke professed not to know whether space was a substance or an accident and offered to answer the question when the ideas of substance and accident were clarified. He was more confident of the idea of pure space than he was of the traditional philosophical categories. Locke placed a great load on the simple idea of space, and by the activity of his reason he went beyond the bounds of possible experience.

Duration and time. The idea of duration is broader than that of time. If we consider the train of ideas that passes through our minds, we observe that one idea constantly succeeds another, and so we come by the idea of succession. By reflection we acquire the idea of duration, which we may then apply to motion and sensory ideas. Where there is no perception of the succession of ideas in our minds, there is no sense of time. Locke insisted that motion does not furnish us with the idea of duration, and he directly opposed Aristotle's definition that "time is the measure of motion with respect to before and after."

Once we have the idea of duration, we need a measure of common duration. Time is the consideration of duration marked by certain measures such as minutes, hours, and days. The most convenient measures of time must be capable of division into equal portions of constantly repeated periods. We cannot be certain of the constancy of motions or of the time spans they measure. Locke was concerned with liberating time from motion. Consequently, he argued that we must consider duration itself as "going on in one constant, equal, uniform course; but none of the measures of it which we make use of can be known to do so" (II.xiv.21). Once time is liberated from motion, Locke held, we can conceive of infinite duration even beyond creation. Thus we can expand by endless addition the idea of duration to come to the notion of eternity.

Were it not for the implicit realism of Locke's arguments, it would be possible to agree with those scholars who have seen in his arguments about duration and expansion a vague groping for a position somewhat similar to Kant's a priori aesthetic. For both men, space becomes the framework of body, and duration or time the structure of the mind, or the inner sense.

Number. The idea of unity is everywhere suggested to the mind, and no idea is more simple. By repeating it we come to the complex modes of number. Once we have learned to perform this operation, we cannot stop short of the idea of infinity. Locke regarded both finite and infinite as modes of quantity. Because we are able to apply the idea of number to space and time, we are capable of conceiving of them as infinite. The idea of infinity is essentially negative, since we come to it by enlarging our ideas of number as much as we please and discover that there is no reason ever to stop. We may know that number, space, and duration are infinite, but we cannot positively know infinity itself. Locke insisted that however remote from the simple ideas of sensation and reflection these ideas may be, they have their origin in those simple ideas.

The modes of thinking. Locke gave only casual and formal attention to the modes of thinking, such as sensation, remembrance, recollection, contemplation, attention, dreaming, reasoning, judging, willing, and knowing. Equally superficial was his consideration of modes of pleasure and pain, which consisted of little more than definitions of various emotions.

Power. The chapter on power is the longest in *Essay*, and Locke felt obliged to rewrite portions of it time and again, for each new edition. It is evident that power is not perceived as such. Locke observed that the mind, taking note of the changes and sequences of our ideas and "concluding from what it has so constantly observed to have been, that the like changes will for the future be made in the same things, by like agents, and by the like ways … comes by that idea which we call power" (II.xxi.1). From this it hardly seems that the idea of power is a simple idea, unless Locke meant no more than that the idea of power is only the observation of the regular order and connection of our ideas. But Locke wrote that "since whatever change is observed, the mind must collect a power somewhere able to make that change, as well as a possibility in the thing itself to receive it" (II.xxi.4). Here the idea of power is a necessary idea of reason, grounded in certain other experiences. Locke never made clear this distinction. He admitted that the idea of power included some kind of relation but insisted that it was a simple idea.

Power is both passive and active. Whether or not matter has any active power, Locke pointed out, we have our idea of active power from the operations of the mind itself. We find by direct observation that we have the power to begin, continue, or stop certain actions of our minds and motions of our bodies. This power we call will, and the actual exercise of this power, volition, or willing. Action is voluntary or involuntary insofar as it is or is not consequent upon the order or command of the mind.

Locke proceeded to explore the ideas of will, desire, and freedom in terms of the idea of power. "The idea of

liberty is the idea of a power in any agent to do or forbear any particular action, according to the determination or thought of the mind, whereby either of them is preferred to the other" (II.xxi.8). Where this power is absent, a man is under necessity. Locke consequently dismissed as unintelligible the question of whether or not the will is free. The only intelligible question is whether or not a man is free. Freedom is one power of an agent and will is another; one power cannot be the power of another. "As far as this power reaches, of acting or not acting, by the determination of his own thought preferring either, so far is a man free" (II.xxi.21). Freedom then, for Locke, was the absence of constraint. If we distinguish will from desire, we cannot make the mistake of thinking the will is free.

What then determines the will with respect to action is some uneasiness in a man that may be called the uneasiness of desire. Good and evil work on the mind but do not determine the will to particular actions. The only thing that can overcome the uneasiness of one desire is the greater uneasiness of another. The removal of uneasiness is the first and necessary step to happiness. Since it is present desire that moves the will to action, good and evil contemplated and known in the mind can move us to action only when that knowledge is accompanied by a greater uneasiness than any other. Since we have many desires and can have knowledge of desired good in the future as well as feared evil, we can suspend the pursuit of any desire until we have judged it. Thus, government of our passions is possible whenever there is a greater uneasiness in not doing so. This power is the ground on which we hold men responsible for their actions. Good and bad are nothing but pleasure or pain, present or future. Error in choice is usually due to the greater strength of present pleasure or pain in comparison with future pleasure and pain. A true knowledge of what contributes to our happiness can influence a choice only when to deviate from that choice would give greater uneasiness than would any other action. Thus it is possible to change the pleasantness and unpleasantness of various actions by consideration, practice, application, and custom.

Locke's conception of power, like his ideas of cause and effect, was inadequate and vague. It was both a simple idea and a complex one; it was the notion of regular sequence and that of efficacious cause; and it was at once given and a priori. The rational and empirical elements in Locke were at war here. Locke was at his best in showing how the word "power" is commonly used. His analysis of the will and freedom was likewise involved in difficulties. The will is not free and thus man's actions are determined; but at the same time we can suspend the execution of any desire by our judgment. Locke was aware of these difficulties, but he saw no satisfactory alternative.

Mixed modes. Mixed modes are made by the mind and are exemplified by drunkenness, a lie, obligation, sacrilege, or murder. To a great degree we get these ideas by the explanation of the words that stand for them.

SUBSTANCE. Of all the ideas considered by Locke none gave him more difficulty than that of substance, and nowhere was his empiricism more in conflict with his rationalism. The diverse trends of Locke's thought concerning substance and the problems he raised prepared the ground for Berkeley, Hume, Kant, and many others who struggled with the same questions. At every opportunity throughout *Essay* he returned to consider particular substances and the general idea of substance. Locke held that we are conversant only with particular substances through experience; yet his rationalism and realism would not permit him to abandon the general idea of substance.

The mind is furnished with many simple ideas by the senses, and it observes by reflection that certain of them are constantly together. It then presumes that these belong to one thing and for convenience gives them one name. In this way the mind arrives at the complex idea of particular substances, such as gold, which we observe to be yellow and malleable, to dissolve in aqua regia, to melt, and not to be used up in fire. A substance so defined gives us only a nominal definition.

Locke added that "not imagining how these simple ideas can subsist by themselves, we accustom ourselves to suppose some substratum wherein they do subsist, and from which they do result; which therefore we call substance" (II.xxiii.1). This idea of a substratum is extremely vague, and Locke called it a "something we know not what." Our ideas do not reach, and we cannot have, a knowledge of the real essence of substances. Nonetheless, Locke continued to believe that real essences do exist, although our knowledge comes short of them.

Our knowledge of corporeal substances consists of ideas of the primary and secondary qualities perceived by the senses and of the powers we observe in them to affect or be affected by other things. We have as clear an idea of spirit as of body, but we are not capable of knowing the real essence of either. Locke observed that we know as little of how the parts of a body cohere as of how our spirits perceive ideas or move our bodies, since we know nothing of either except our simple ideas of them. Locke

even suggested that God could if he wished, as far as we know, add to matter the power to think, just as easily as he could add to matter a separate substance with the power to think.

Even our idea of God is based on simple ideas that are enlarged with the idea of infinity. God's infinite essence is unknown to us. We can only know that he exists.

RELATIONS. The mind can consider any idea as it stands in relation to any other; and thus we come by ideas of relation, such as father, whiter, older. Frequently, the lack of a correlative term leads us to mistake a relative term for an absolute one. Locke distinguished the relation from the things related and appears to have made all relations external. Indeed, he held that many ideas of relation are clearer than ideas of substances; for example, the idea of brothers is clearer than the perfect idea of man.

Though there are many ideas and words signifying relations, they all terminate in simple ideas. There is a difficulty here. If the idea of relation is not a simple idea or a combination of simple ideas, then it is distinct from them. Like the general idea of substance, it is a concept derived from reason. No doubt the mind is capable of comparing the relation of one idea with another, but our perception of this operation must have for its object either a simple idea or the operation itself. On this point Locke was obscure and evasive and avoided the difficulties by the vague assertion that all relations terminate in simple ideas.

Causation. The relation to which Locke first turned was cause and effect. His discussion was inadequate and marked by the duality found in his consideration of other ideas. We observe the order and connection of our ideas and the coming into existence of things and qualities. In pointing this out Locke was on strictly empirical grounds. When, however, he defined cause as "that which produces any simple or complex idea," and "that which is produced, effect" (II.xxvi.1), he went beyond experience and rested his argument on reason. Locke undoubtedly saw the difficulties of his position. He was concerned, on the one hand, to show how we have the ideas of cause and effect from experience. On the other hand, he was not satisfied with a mere sequence theory. The difficulty arose, as it did with power and substance, because he was persuaded that there is a reality beyond the ideas manifest to us. It is a reality, however, about which he could say little in terms of his representationalism.

Identity and diversity. Under relation Locke also examined identity and diversity, by which he meant the relation of a thing to itself, particularly with respect to different times and places. As was stated above, the identity of a plant, an animal, or a man consists in a participation in the same continued life. To this Locke added an examination of personal identity. He argued that personal identity is consciousness of being the same thinking self at different times and places. Locke also discussed other relations, such as proportional, natural, instituted, and moral, which are not essential to the main argument of *Essay* and which will, therefore, not be discussed here.

The remaining chapters of Book II of *Essay* are devoted to "Clear and Obscure, Distinct and Confused Ideas," "Real and Fantastical Ideas," "Adequate and Inadequate Ideas," "True and False Ideas," and "The Association of Ideas." All of them have merit in clarifying other parts of *Essay* but add little that is new and not discussed elsewhere. Consequently, they will be passed over.

LANGUAGE. At the end of Book II of *Essay* Locke related that he had originally intended to pass on to a consideration of knowledge. He found, however, such a close connection between words and ideas, particularly between abstract ideas and general words, that he had first to examine the "nature, use, and signification" of language, since all knowledge consists of propositions. Book III, therefore, was incorporated into *Essay*.

The merits of Book III are the subject of some controversy. Most scholars have dismissed it as unimportant and confused. Some, such as Aaron, see many merits in it despite its manifest inadequacies.

The primary functions of language are to communicate with our fellow men, to make signs for ourselves of internal conceptions, and to stand as marks for ideas. Language is most useful when general names stand for general ideas and operations of the mind. Since all except proper names are general, a consideration of what kinds of things words stand for is in order. "Words, in their primary' or immediate signification, stand for nothing but the ideas in the mind of him that uses them" (III.ii.2). We suppose they stand for the same ideas in the minds of others. Words stand for things only indirectly. General words stand for general ideas, which become general by separation from other ideas and from particular circumstances. This process Locke called abstraction.

Definition. Definition by genus and differentia is merely a convenience by which we avoid enumerating various simple ideas for which the genus stands. (In this, Locke prepared the way for descriptive definition, which makes no pretense of defining the real essence of things.) It follows that general or universal ideas are made by the

understanding for its own use. Thus the essences of so-called species are nothing but abstract ideas. Locke asserted that every distinct abstract idea is a distinct essence. This must not be taken in a Platonic sense, for it is the mind itself that makes these abstract ideas. If essences are distinguished into nominal and real, then with respect to simple ideas and modes there is no difference between nominal and real essence. In substances, they are decidedly different, in that the real essence of substance is unknowable to us.

Names. Locke asserted that the names of simple ideas are not definable. One wonders, Is blue a general idea? If so, what is this blue as against that blue? What is separated out? What retained? Locke never examined these questions, with the result that his conception of abstraction is vague and vacillating. Locke gave several distinct meanings to such terms as "general ideas" and "universal ideas," shifting from one meaning to another and never clarifying them.

Complex ideas consisting of several simple ideas are definable and intelligible provided one has experience of the simple ideas that compose them. Without experience how can a blind man understand the definition of a rainbow?

Simple ideas are "perfectly taken from the existence of things and are not arbitrary at all" (III.iv.17). Ideas of substances refer to a pattern with some latitude, whereas ideas of mixed modes are absolutely arbitrary and refer to no real existence. They are not, however, made at random or without reason. It is the name that ties these ideas together, and each such idea is its own prototype.

Since names for substances stand for complex ideas perceived regularly to go together and supposed to belong to one thing, we necessarily come short of the real essences, if there are any. One may use the word "gold" to signify the coexistence of several ideas. One man may use the term to signify the complex idea of A and B and C. Another man of more experience may add D, or add D and leave out A. Thus, these essences are of our own making without being entirely arbitrary. In any case, the boundaries of the species of substances are drawn by men.

Connective words In a brief chapter, "Of Particles," Locke pointed out that we need words signifying the connections that the mind makes between ideas or propositions. These show what connection, restriction, distinction, opposition, or emphasis is given to the parts of discourse. These words signify, not ideas, but an action of the mind. Again a difficulty arises. If "is" and "is not" stand for the mind's act of affirming or denying, then either the mind directly apprehends its own actions in some way or we do have ideas of affirmation or denial. If we do have ideas of the mind's acts, then these words ought to signify the ideas of these acts; if we do not have ideas that these words signify, then either we do not apprehend them or something besides ideas is the object of the mind when it thinks. The remainder of Book III concerns Locke's thoughts on the imperfection of words, the abuse of words, and his suggested remedies for these imperfections and abuses.

KNOWLEDGE

The first three books of *Essay* are largely a preparation for the fourth. Many scholars see a fundamental cleavage between Book II and Book IV. Yet Locke saw no conflict between the two books, and whatever split existed in Locke's thought runs throughout *Essay*, as J. W. Yolton and others have pointed out. An effort can be made to reconcile Locke's empiricism and his rationalism, his grounding of all ideas and knowledge in experience and his going beyond experience to the existence of things.

Many of Locke's difficulties stem from his definition of "idea." It is so broad that anything perceived or known must be an idea. But Locke showed, in Books I and II, that we get all our ideas from experience, not in order to claim that nothing exists except ideas, but to show that there is an alternative to the theory of innate ideas. For Locke, experience is initially a contact of bodies and subsequently a reflection of the mind. He never doubted the existence of an external physical world, the inner workings of which are unknown to us.

Sources of knowledge. There are two sources of knowledge—sensation and reflection. The ideas we have from reflection are in some important ways quite different from those we have from sensation. In Book II Locke asserted that the mind "turns its view inward upon itself and observes its own actions about those ideas it has (and) takes from thence other ideas" (II.vi.1). The important point here is that in reflection the mind observes its own action. It is true that Locke spoke of modes of the simple ideas of reflection, such as remembering, discerning, reasoning, and judging. Nonetheless, if the mind does observe its own action, then something more than ideas are the object of the mind in reflection, or else ideas of reflection are somehow importantly different from the ideas of sensation. This point will show up in a consideration of Locke's theory of knowledge.

Propositions. Locke defined knowledge as "the perception of the connection and agreement, or disagree-

ment and repugnancy, of any of our ideas" (IV.i.2). This agreement or disagreement is in respect to four types: identity and diversity, relation, coexistence or necessary connection, and real existence. Perceiving agreement or disagreement is quite different from just barely perceiving the ideas that are said to agree or disagree. Strictly speaking, this perception must be a distinct idea of either agreement or disagreement. Yet this was not Locke's meaning. Where there is knowledge, there is judgment, since there can be no knowledge without a proposition, mental or verbal. Locke defined truth as "the joining or separating of signs, as the things signified by them do agree or disagree one with another" (IV.v.2). There are two sorts of propositions: mental, "wherein the ideas in our understandings are, without the use of words, put together or separated by the mind perceiving or judging of their agreement or disagreement" (IV.v.5); and verbal, which stand for mental propositions.

Judgments. In this view, ideas are the materials of knowledge, the terms of mental propositions. They are, insofar as they are given in sensation and reflection, the subject matter of reflection. If perception of agreement or disagreement in identity and diversity is the first act of the mind, then that act is a judgment. If we infallibly know, as soon as we have it in our minds, that the idea of white is identical with itself and different from that of red, and that the idea of round is identical with itself and different from that of square, we must distinguish between the bare having of these ideas and the knowledge of their identity and diversity. The knowledge of their identity and diversity is a judgment. It is reflective, and in it the mind perceives its own action or operation. There can be no distinction between the judgment and the idea of it. This is perhaps Locke's meaning, which is unfortunately obscured by his broad use of the term "idea." This perception of its own action is quite distinct from the abstract idea of the power of judgment. We may be uncertain as to how the mind makes judgments, what determines it to judge, or in what kind of a substance this power inheres, but we may be sure that in the actual making of a true judgment the mind perceives its own act. This position may be beset with difficulties, but it makes some sense out of Locke's definition of knowledge.

Degrees of knowledge. Locke recognized two degrees of knowledge, in the strict sense of the term—intuition and demonstration. Of the two, intuition is more fundamental and certain. "The mind perceives the agreement or disagreement of two ideas immediately by themselves, without the intervention of any other" (IV.ii.1). Such knowledge is irresistible and leaves no room for hesitation, doubt, or examination. Upon it depends all the certainty and evidence of all our knowledge. Here, clearly, what the mind perceives is not any third idea, but its own act. In demonstration the mind perceives agreement or disagreement, not immediately, but through other mediating ideas. Each step in demonstration rests upon an intuition. This kind of knowledge is most evident in, but is not limited to, mathematics.

A third degree of knowledge is "employed about the particular existence of finite beings without us, which going beyond bare probability and yet not reaching perfectly to either of the foregoing degrees of certainty, passes under the name of knowledge" (IV.ii.14). Locke called this sensitive knowledge. Fully aware of the dialectical difficulty entailed in this position, he grounded his reply to critics on common sense. The differences between dreaming and waking, imagining and sensing, are strong enough to justify this conviction. Hunger and thirst should bring a skeptic to his senses. For Locke, it was enough that common sense supported him, for he always took sensory ideas to be signs or representations of something beyond themselves.

Limits of knowledge. Locke asserted that knowledge extends no farther than our ideas and, specifically, no further than the perception of the agreement or disagreement of our ideas. We cannot have knowledge of all the relations of our ideas or rational knowledge of the necessary relations between many of our ideas. Sensitive knowledge goes only as far as the existence of things, not to their real essence, or reality. Two examples were given. In the first, Locke argued that though we have the ideas of circle, square, and equality, we may never find a circle equal to a square and know them to be equal. In the second, he observed that we have ideas of matter and thinking but may never know whether mere material being thinks. This has been discussed earlier.

In his controversy with Stillingfleet, Locke never abandoned this latter thesis. And throughout this section (IV.iii) Locke showed that many relations of coexistence give us no certainty that they will or must continue to be so. He seemed persuaded that the continued discovery of new knowledge suggests that there are vast horizons of reality that we may advance upon but can never reach. With respect to the relations between abstract ideas we may hope to advance very far, as in mathematics. To this he added the belief that a demonstrable science of morality is possible. On the other hand, he held that we can have no certain knowledge of bodies or of unembodied spirits.

Knowledge of existents. Locke argued that though our knowledge terminates in our ideas, our knowledge is real. "Simple ideas are not fictions of our fancies, but the natural and regular productions of things without us, really operating upon us; and so carry with them all the conformity which is intended; or which our state requires" (IV.iv.4). On the other hand, he argued: "All our complex ideas, except those of substances, being archetypes of the mind's own making, not intended to be copies of anything, nor referred to the existence of anything, as to their originals, cannot want any conformity necessary to real knowledge" (IV.iv.5).

Universal propositions, the truth of which may be known with certainty, are not concerned directly with existence. Nonetheless, Locke argued that we have intuitive knowledge of our own existence. Here the argument is much the same as Descartes's, and it is valid only if we accept the view that the mind in reflection perceives its own acts. This knowledge of our own existence has the highest degree of certainty, according to Locke.

We have a demonstrable knowledge of God's existence, Locke held. He used a form of the Cosmological Argument: Starting with the certainty of his own existence, he argued to the necessary existence of a being adequate to produce all the effects manifest in experience. The argument assumed the reality of cause, the necessity of order, and the intelligibility of existence.

Of the existence of other things, as has been shown, we have sensitive knowledge. Locke felt the inconsistency of his position on this matter, yet accepted what he believed common sense required. We know of the coexistence of certain qualities and powers, and reason and sense require that they proceed from something outside themselves. Throughout these arguments about existence Locke went beyond his own first definition of knowledge.

PROBABILITY. The remaining portions of *Essay* are concerned with probability, degrees of assent, reason and faith, enthusiasm, error, and the division of the sciences. Though Locke's treatment of probability is inadequate, he recognized its importance. The grounds of probability lie in the apparent conformity of propositions with our experience and the testimony of others. Practical experience shows us that our knowledge is slight, and action requires that we proceed in our affairs with something less than certainty.

Faith was, for Locke, the acceptance of revelation. It must be sharply distinguished from reason, which is "the discovery of the certainty or probability of such propositions or truths, which the mind arrives at by deduction made from such ideas which it has got by the use of its natural faculties, viz. by sensation or reflection" (IV.xviii.2). Though reason is not able to discover the truth of revelation, nevertheless, something claimed to be revelation cannot be accepted against the clear evidence of the understanding. Thus, enthusiasm sets reason aside and substitutes for it bare fancies born of conceit and blind impulse.

Error. Error cannot lie in intuition. Locke found four sources of error: the want of proofs, inability to use them, unwillingness to use them, and wrong measures of probability. Locke concluded *Essay* with a brief division of science, or human knowledge, into three classes—natural philosophy, or φυσικὴ practical action and ethics, or προκτικὴ, and οημειωτκὴ, or the doctrine of signs.

INFLUENCE OF *ESSAY*. Many minds of the seventeenth century contributed to the overthrow of the School philosophies and the development of the new sciences and philosophies. Descartes and Locke between them, however, set the tone and direction for what was to follow. Certainly Locke was the most prominent figure in the early eighteenth century, the indispensable precursor of Berkeley and Hume as well as a fountainhead for the French Encyclopedists. If it is said that the two strains of Cartesian rationalism and Lockian empiricism met in Kant, it can be added that Hume built on Locke's foundation and Kant formalized much that was first a vague groping in Locke. Though Locke was not a wholly satisfactory thinker, his influence on thought in England and America has never completely abated, and even now there appears to be a revived interest in *Essay*.

POLITICAL THOUGHT

Locke's earliest known political writings were *Essays on the Law of Nature*, written in Latin between 1660 and 1664 but not known until the Lovelace Collection was examined in 1946. They were first published in 1954 with a translation by W. von Leyden. Though much in these essays appears in *An Essay concerning Human Understanding* and *Two Treatises of Government*, there remain many points at which the early essays are in conflict with parts of both later works. This fact and the bother of translating them may have deterred Locke from publishing them, despite the urging of Tyrrell. Since von Leyden can find no evidence of direct influence of these essays on anyone other than Tyrrell and Gabriel Towerson, the student of Locke is referred to von Leyden's publication for additional information.

TWO TREATISES. *Two Treatises of Government* appeared anonymously in 1690, written, it is said, to justify the revolution of 1688, or, according to the preface, "to establish the Throne of Our Great Restorer, our present King William; to make good his Title, in the Consent of the People." Locke acknowledged his authorship only in a codicil in his will listing his anonymous works and giving to the Bodleian Library a corrected copy of *Two Treatises*. He never felt that any of the editions printed during his lifetime had satisfactorily rendered his work. Only in 1960 did Peter Laslett publish a critical edition based on the Coste master copy of *Two Treatises*.

THE FIRST TREATISE. It has long been suspected that the first treatise was written in 1683 and that the second treatise was written in 1689. Laslett has presented much evidence to show that the second treatise was the earlier work, written between 1679 and 1681. If his thesis is correct, it was a revolutionary document, whose purpose was not primarily to philosophize but to furnish a theoretical foundation for the political aims and maneuvers of Shaftesbury and his followers in their struggle with Charles II. Only further scholarly probing will resolve this question.

In his preface, Locke stated that the greater part of the original work had been lost. He was satisfied that what remained was sufficient, since he had neither the time nor the inclination to rewrite the missing sections. The evidence is clear that it was portions of the first treatise that were lost.

The first treatise is a sarcastic and harsh criticism of Sir Robert Filmer's *Patriarcha*, which argued for the divine right of kings. Locke's treatise is more of historical than philosophical importance. It argued that Adam was not, as Filmer claimed, divinely appointed monarch of the world and all his descendants. Neither was the power of absolute monarchy inherited from Adam. Adam had no absolute rights over Eve or over his children. Parents have authority over children who are dependent upon them and who must learn obedience as well as many other things for life. The function of the parent is to protect the child and to help him mature. When the child comes to maturity, parental authority ends. In any case, the relation of parent and child is not the same as that of sovereign and subject. Were Filmer right, one would have to conclude that every man is born a slave, a notion that was utterly repugnant to Locke. Even if Filmer were correct, it would be impossible to show that existing rulers, especially the English kings, possess legitimate claims to their sovereignty by tracing it back to lawful descent from Adam.

THE SECOND TREATISE. Locke began the second treatise with the proposition that all men are originally in a state of nature, "a state of perfect freedom to order their actions, and dispose of their possessions, and persons as they think fit, within the bounds of the Law of Nature, without asking leave, or depending upon the Will of any other man" (II.ii.4). Although Locke sometimes wrote as if the state of nature were some period in history, it must be taken largely as a philosophical fiction, an assumption made to show the nature and foundation of political power, a fiction at least as old as Plato's treatment of the Prometheus myth in the *Protagoras*. It is a state of equality but not of unbounded license. Being rational and being a creature bound by God, man must be governed by the law of nature.

Natural law. Though the concept of the law of nature is as old as antiquity, it flourished in the seventeenth century in the minds of a considerable number of ethical and political thinkers. In general it supposed that man by the use of reason could know in the main the fundamental principles of morality, which he otherwise knew through Christian revelation. Locke was extremely vague about the law of nature, but in his *Essays on the Law of Nature* he held that that law rests ultimately on God's will. Reason discovers it. It is not innate. When, however, Locke spoke of it as "writ in the hearts of all mankind," he suggested some kind of innateness. There are obvious difficulties here, for sense and reason may fail men, even though the law of nature is binding on all. Moreover, the various exponents of the law of nature differ on what it consists of, except that it presupposes the brotherhood of man and human benevolence.

State of nature. In a state of nature, according to Locke, all men are bound to preserve peace, preserve mankind, and refrain from hurt to one another. The execution of the law of nature is the responsibility of each individual. If any man violates this law, he thereby puts himself in a state of war with the others, who may then punish the offender. The power that one man may hold over another is neither absolute nor arbitrary and must be restrained by proportion. The state of nature was for Locke a society of men, as distinct from a state of government, or a political society.

Social contract. There are certain inconveniences in a state of nature, such as men's partiality and the inclination on the part of some men to violate the rights of others. The remedy for this is civil government, wherein men

by common consent form a social contract and create a single body politic. This contract is not between ruler and ruled, but between equally free men. The aim of the contract is to preserve the lives, freedom, and property of all, as they belong to each under natural law. Whoever, therefore, attempts to gain absolute power over another puts himself at war with the other. This holds in the political state as well as the state of nature. When a ruler becomes a tyrant, he puts himself in a state of war with the people, who then, if no redress be found, may make an appeal to heaven, that is, may revolt. This power is but an extension of the right of each to punish an aggressor in the state of nature. Unlike Hobbes, Locke was persuaded that men are capable of judging whether they are cruelly subjected and unjustly treated. Since one reason for men entering into the social contract is to avoid a state of war, the contract is broken when the sovereign puts himself into a state of war with the people by becoming a tyrant.

Slavery. Curiously, Locke justified slavery on the grounds that those who became slaves were originally in a state of wrongful war with those who conquered them and, being captive, forfeited their freedom. Apart from being bad history, this argument ignores the rights of the children of slaves. Locke's inconsistency here may mercifully be passed over.

Property. Property was an idea that Locke used in both a broad and a narrow sense. Men have a right to self-preservation and therefore to such things as they need for their subsistence. Each man possesses himself absolutely, and therefore that with which he mixes his labor becomes his property. "God has given the earth to mankind in common." No man has original, exclusive rights to the fruits and beasts of the earth. Nevertheless, man must have some means with which to appropriate them. This consists of the labor of his body and the work of his hand. By labor, man removes things from a state of nature and makes them his property. Without labor, the earth and things in general have but little value. However, only so much as a man improves and can use belongs to him, nor may a man deprive another of the means of self-preservation by overextending his reach for property.

Though the right to property is grounded in nature, it is not secured therein. It is one of the primary ends of the state to preserve the rights of property, as well as to make laws governing the use, distribution, and transference of property. In communities or countries under government, there are fixed boundaries to the common territory, and there is land and property held in common which no one may appropriate to himself and to which those not members of the community have no right at all. Money, being something that does not spoil, came into use by mutual consent, serving as a useful means of exchange. At the same time it made possible the accumulation of wealth greater than warranted by need or use.

Political society. Having established several rights and duties belonging to men by nature and having shown certain inconveniences and disadvantages of the state of nature, Locke turned to political society. The first society consists of the family, whose aims are not initially or primarily those of political society, but which may be included under political society.

In political society "any number of men are so united into one Society, as to quit everyone his Executive power of the law of nature, and to resign it to the public" (II.vii.89). The legislative and executive powers are "a right of making laws with penalties of Death, and consequently all less Penalties, for the regulating and preserving of property, and of employing the force of the community, in the execution of such laws, and in the defense of the commonwealth from foreign injury, and all this only for the public good" (II.i.3). By the social contract men give up, not all their rights, but only the legislative and executive right they originally had under the law of nature. This transference of power is always subordinate to the proper and true ends of the commonwealth, which are "the mutual preservation of their lives, liberties and estates."

Each man must voluntarily consent to the compact either explicitly or implicitly. An individual who at age of discretion remains a member of the community tacitly consents to the compact.

Since the compact is made between the members of the community, sovereignty ultimately remains with the people. The sovereign, in the form of a legislative body, and executive, or both, is the agent and executor of the sovereignty of the people. The community can act only by the rule of the majority, and everyone is bound by it, because an agreement of unanimity is virtually impossible. It is the people who establish the legislative, executive, and judiciary powers. Thus, an absolute monarch is incompatible with civil society.

Locke's theories so far are compatible with either monarchy, oligarchy, or democracy so long as it is recognized that ultimate sovereignty lies with the people. He believed that a constitutional monarchy with executive power, including the judiciary, in the hands of the monarch, and legislative powers in a parliamentary assembly elected by the people was the most satisfactory form of government. The supreme power he held to be

the legislative, for it makes the laws that the executive must carry out and enforce. Whenever the executive violates the trust that he holds, no obligation is owed him and he may be deposed. The legislature may also violate its trust, though Locke believed it less likely to do so. Whenever this occurs, the people have a right to dissolve it and establish a new government. For this reason a regularly elected legislative body is desirable.

Rebellion. Locke explicitly recognized, as the events during his lifetime had shown, that men may become tyrants to those whom they were bound to serve. It may be a king, an assembly, or a usurper that claims absolute power. In such cases the people have a right to rebellion if no other redress is possible. Locke was not unmindful of the fact that the executive needs latitude and prerogative so that he may govern, and that the legislative body must deliberate and make laws that they believe to be in the public good. The right to rebellion is warranted only in the most extreme conditions, where all other means fail. Locke did not believe that men would lightly avail themselves of this power, for men will suffer and endure much before they resort to rebellion.

In transferring to the government the right to make and execute law and make war and peace, men do not give up the natural light of reason, by which they judge good and evil, right and wrong, justice and injustice. In specific laws or executive decisions judgment must be allowed to the legislature and the executive. If, however, a long train of acts shows a tyrannical course, then men, judging that the sovereign has put himself into a state of war with them, may justly dethrone the tyrant. On the other hand, the legislative and executive power can never revert to the people unless there is a breach of trust.

The dissolution of government is not the dissolution of society. The aim of revolution is the establishment of a new government, not a return to a state of nature. The dissolution of a government may occur under many circumstances, but foremost among them are when the arbitrary will of a single person or prince is set in place of the law; when the prince hinders the legislature from due and lawful assembly; when there is arbitrary change in elections; when the people are delivered into subjection by a foreign power; and when the executive neglects and abandons his charge. In all such cases sovereignty reverts to the society, and the people have a right to act as the supreme power and continue the legislature in themselves, or erect a new form, or under the old form place sovereignty in new hands, whichever they think best. On the other hand, "the power that every individual gave the society, … can never revert to the individuals again, as long as the society lasts" (II.xix.243). As theory, Locke's second treatise is full of inadequacies, but its magnificent sweep of ideas prepared the ground for popular and democratic government.

EDUCATION AND RELIGION

Locke's thought on education and religion was not presented in strictly philosophical terms. It was, however, deeply rooted in the fundamental concepts of *Essay* and *Two Treatises*. His works in these areas display clearly the liberal bent of his mind as well as his love of freedom, tolerance, and truth. His attitude was pragmatic and based on considerable psychological insight into the motives, needs, passions, and follies of men. *Some Thoughts concerning Education*, several letters on toleration, and *The Reasonableness of Christianity* profoundly affected educational and religious thought in the eighteenth century and after. Two of these works, *Some Thoughts concerning Education* and the first *Letter on Toleration*, continue to be fresh and relevant.

EDUCATION. When Locke was in Holland, he wrote a number of letters to Edward Clark advising him on the education of his son, a young man of no particular distinction. Locke had in mind the education of a gentleman who would one day be a squire. In 1693 Locke modified these letters somewhat and published the contents as *Some Thoughts concerning Education* in response to "so many, who profess themselves at a loss how to breed their children." His thought was marked by a ready understanding of, and warm sympathy with, children. Three main thoughts dominate the work. First, the individual aptitudes, capacities, and idiosyncrasies of the child should govern learning, not arbitrary curricular or rote learning taught by the rod. Second, Locke placed the health of the body and the development of a sound character ahead of intellectual learning. In the third place, he saw that play, high spirits, and the "gamesome humor" natural to children should govern the business of learning wherever possible. Compulsory learning is irksome; where there is play in learning, there will be joy in it. Throughout he placed emphasis on good example, practice, and use rather than on precepts, rules, and punishment. The work was an implicit criticism of his own education at Westminster and Oxford, which he found unpleasant and largely useless.

Writing almost as a physician, Locke advised "plenty of open air, exercise, and sleep; plain diet, no wine or strong drink, and very little or no physic; not too warm and strait clothing; especially the head and feet kept cold,

and the feet often used to cold water and exposed to wet." The aim in all was to keep the body in strength and vigor, able to endure hardships.

Locke urged that early training must establish the authority of the parents so that good habits may be established. The prime purpose is the development of virtue, the principle of which is the power of denying ourselves the satisfaction of our desires. The child should be taught to submit to reason when young. Parents teach by their own example. They should avoid severe punishments and beatings as well as artificial rewards. Rules should be few when a child is young, but those few should be obeyed. Mild, firm, and rational approval or disapproval are most effective in curbing bad behavior. Children should be frequently in the company of their parents, who should in turn study the disposition of the child and endeavor to use the child's natural desire for freedom and play to make learning as much like recreation as possible. High spirits should not be curbed, but turned to creative use. Curiosity too should be encouraged, and questions should be heard and fairly answered. Cruelty must always be discouraged and courageousness approved.

As the child grows, familiarity should be increased so that the parent has a friend in the mature child. Virtue, breeding, and a free liberal spirit as well as wisdom and truthfulness were the goals set by Locke in all his advice. Affection and friendship were for him both means and ends of good education.

Learning, though important, Locke put last. First, he would have the child learn to speak and read his own language well by example and practice, not by grammar. In the study of all languages, he would put off the study of grammar until they can be spoken well. He would begin the learning of a second modern language early. Reluctantly he would allow a gentleman's son to learn Latin, but he did not recommend much time on Greek, Hebrew, Arabic, rhetoric, or logic, which constituted the curricula of the universities of his day. Rather, time should be given to the study of geography, arithmetic, astronomy, geometry, history, ethics, and civil law. Dancing he encouraged, and music as well, in moderation. He was less sympathetic to poetry. Remarkably, he urged that everyone learn at least one manual trade and make some study of accounting. Finally, travel was valuable if not done before one could profit by it.

If much of this is familiar and even trite, it must be remembered that Locke was among the first to formulate these ideas. His influence on educational thought and practice was enormous and is still very much with us in its fundamental outlook and method.

RELIGION. Locke saw some merits in all the competing claims of various religious groups. He also saw the destructive force that was released when these claims sought exclusive public dominion at the expense of individual conscience. He looked in several directions at once. This tendency has earned for him the reputation of being timorous and compromising. Nonetheless, it is on this trait of mind that much of his great influence and reputation rests. For Locke, fidelity to the evidence at hand always outweighed cleverness, consistency, and dialectic. It is the chief testimony to his claim that truth was always his aim, even when he might have won an easy victory by dogmatic consistency.

Locke's writings on religion are voluminous. When he died he was working on extensive commentaries on the *Epistles of St. Paul*, as well as a draft of a fourth *Letter on Toleration*. Earlier he had written and published three letters on toleration, *The Reasonableness of Christianity* (1695), and two *Vindications* (1695 and 1697) of the latter work. Moreover, Locke's three letters to Stillingfleet, the bishop of Worcester, are concerned with religious questions as well as epistemological ones.

Religious tolerance. Locke's first *Letter concerning Toleration* stated his position clearly, and he never deviated from it substantially. It was originally written in Latin as a letter to his Dutch friend Philip van Limborch. In 1689 it was published on the Continent in Latin, and in the same year a translation of it by William Popple appeared in English.

Locke was not the first to write in advocacy of religious toleration. His was, however, a powerful, direct, and passionate plea. It was linked with *Essay* by its recognition of the limits of human knowledge and human fallibility, and with *Two Treatises* by his deep commitment to individual rights and freedom.

Locke took toleration to be the chief characteristic mark of the true church, for religious belief is primarily a relation between each man and God. True religion regulates men's lives according to virtue and piety, and without charity and love religion is false to itself. Those who persecute others in the name of Christ abjure his teachings, seeking only outward conformity, not peace and holiness. Who can believe that in torture and execution the fanatic truly seeks the salvation of the soul of his victim? Moreover, the mind cannot be forced or belief compelled. All efforts to force or compel belief breed only hypocrisy and contempt of God. Persuasion is the only lever that can truly move the mind.

A church is "a voluntary society of men, joining themselves together of their own accord in order to the public worshipping of God in such manner as they judge acceptable to Him, and effectual to the salvation of their souls." It is sharply distinct from a state, or commonwealth. The state is concerned with the public good, protecting life, liberty, and property. It has no authority in matters of the spirit. "Whatever is lawful in the commonwealth cannot be prohibited by the magistrate in the church."

It is to be doubted that any man or group of men possess the truth about the one true way to salvation. In the Scriptures we have all that may reasonably be claimed by Christians to be the word of God. The rest are the speculations and beliefs of men concerning articles of faith and forms of worship. Sincere and honest men differ in these matters, and only tolerance of these differences can bring about public peace and Christian charity. Jews, pagans, and Muslims are all equally confident in their religious faith. Mutual tolerance is essential where such diversity exists. This is most evident when we observe that it is the most powerful party that persecutes others in the name of religion. Yet in different countries and at different times power has lain in the hands of different religious groups. It is physical power, not true faith, which decides who is persecuted and who persecutes.

Throughout Locke's argument the liberty of person and the liberty of conscience are decisive. He limited this liberty only by denying to religion the right to harm directly another person or group or to practice clearly immoral rites. By a curious and probably prudential exception, he denied tolerance to atheists, because promises, covenants, and oaths would not bind them, and to any church so constituted "that all those who enter into it do thereby *ipso facto* deliver themselves up to the protection and service of another prince."

Despite these limitations, Locke's letter moved subsequent generations to a greater spirit of tolerance in religious matters. It is still part of the liberal democratic ideal and transcends the time of its composition.

Faith and reason. *The Reasonableness of Christianity* and *Vindications* are works more bound to Locke's own time. Locke was probably neither a Socinian nor a deist, even though certain deists and Unitarians found comfort and inspiration in his work. He was a sincere Christian, who tried to diminish the flourishing schisms and sects by proposing a return to the Scriptures and an abandonment of the interminable theological disputes of his day. He accepted the divine inspiration of the Bible. Nevertheless, he held that even revelation must be tested by reason. In the New Testament, Christianity is rational and simple. The core of Christian faith lies in the belief in the fatherhood of God, the divinity of Christ the Messiah, and the morality of charity, love, and divine mercy. Justification by faith means faith in Christ, whose essential revelation is that God is merciful and forgives the sinner who truly repents and strives to live a life of Christian morality. The Mosaic law, God's mercy, and Christian morality are all consonant with human reason. Revelation discloses to man what unaided reason could not discover—the mysteries, the Virgin Birth, the Resurrection, the divinity of Christ. But when disclosed, these do not violate the canons of reason. Here as elsewhere, Locke's emphasis on reason was circumscribed, reason must be followed where possible, but it does not carry us far enough by itself.

Locke's influence was wide and deep. In political, religious, educational, and philosophical thought he inspired the leading minds of England, France, America, and to some extent, Germany. He disposed of the exaggerated rationalism of Descartes and Spinoza; he laid the groundwork for a new empiricism and advanced the claims for experimentalism. Voltaire, Montesquieu, and the French Encyclopedists found in Locke the philosophical, political, educational, and moral basis that enabled them to prepare and advance the ideas that eventuated in the French Revolution. In America, his influence on Jonathan Edwards, Hamilton, and Jefferson was decisive. Locke's zeal for truth as he saw it was stronger than his passion for dialectical and logical niceness, and this may account for the fact that his works prepared the ground for action as well as thought.

See also Animal Mind; Authority; Berkeley, George; Boyle, Robert; Cambridge Platonists; Clarke, Samuel; Cudworth, Ralph; Descartes, René; Edwards, Jonathan; Empiricism; Encyclopédie; Ethics, History of; Filmer, Robert; Gassendi, Pierre; Hobbes, Thomas; Hooker, Richard; Hume, David; Jefferson, Thomas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Malebranche, Nicolas; Montesquieu, Baron de; Natural Law; Newton, Isaac; Personal Identity; Philosophy of Education, History of; Social Contract; Spinoza, Benedict (Baruch) de; Stillingfleet, Edward; Voltaire, François-Marie Arouet de; Whichcote, Benjamin.

Bibliography

WORKS BY LOCKE

Essays on the Law of Nature. Translated from the Latin and edited by W. von Leyden. Oxford, 1954. Also gives an account of the Lovelace Collection.

An Early Draft of Locke's Essay (Draft A). Edited by R. I. Aaron and J. Gibb. Oxford, 1936. Valuable as a study of the development of *Essay*.

An Essay concerning the Understanding, Knowledge, Opinion and Assent (Draft B). Edited by B. Rand. Cambridge, MA: 1931. Valuable, but superseded by manuscript in Bodleian Library.

Epistola de Tolerantia. Gouda, 1689. Translated by William Popple as *A Letter concerning Toleration*. London, 1689. Several defenses appeared in the 1690s and fragments of a fourth in 1706.

An Essay concerning Human Understanding. London, 1690; 2nd ed. with large additions, London, 1694; 3rd ed., London, 1695; 4th ed., with large additions, London, 1700; 5th ed., with many large additions, London, 1706. Best modern edition, from which all quotes in this article are taken, is a reprint of the fifth edition, J. W. Yolton, ed., 2 vols. New York and London, 1961.

Two Treatises of Government. London, 1690. The critical and collated edition of Locke's corrected copy by Peter Laslett (Cambridge, 1960) surpasses all previous editions.

Some Considerations of the Consequences of the Lowering of Interest and the Raising of the Value of Money. London, 1692. Two additional papers on money appeared in 1695.

Some Thoughts concerning Education. London, 1693.

The Reasonableness of Christianity. London, 1695. Defenses of this work were published in 1695 and 1697.

A Letter to the Right Rev. Edward Ld. Bishop of Worcester, concerning Some Passages relating to Mr. Locke's Essay of Humane Understanding. London, 1697. Two further letters appeared, in 1697 and 1699.

A Paraphrase and Notes on the Epistles of St. Paul to the Galatians [etc.]. London, 1705.

Posthumous Works of Mr. John Locke, 6 vols. London, 1706.

The Remains of John Locke. Edited by E. Curl. London, 1714.

Works of John Locke. 3 vols. London, 1714; 10th ed., 10 vols., London, 1801.

The Correspondence of John Locke and Edward Clarke. Edited by B. Rand. Oxford, 1927.

For the remainder of Locke's published and unpublished papers, consult the works listed below by Aaron, Christopherson, von Leyden, Long, Ollion, and Yolton. See von Leyden and Long particularly for the Lovelace Collection.

WORKS ON LOCKE

Biographies

Bourne, H. R. Fox. *Life of John Locke*. 2 vols. London: H. S. King, 1876. Excellent, but inadequate since Lovelace Collection became available.

Cranston, Maurice. *John Locke, a Biography*. London and New York: Longmans, Green, 1957. A thorough study of the life of Locke, using all materials available at present.

King, Lord Peter. *The Life and Letters of John Locke*. London: H. Coburn, 1829. Not good, but contains original material.

Critical Commentaries

Aaron, R. I. "Locke's Theory of Universals." *PAS*, Vol. 33 (1932/1933). Useful and enlightening.

Aaron, R. I. *John Locke*. Oxford: Oxford University Press, 1937; rev. ed., Oxford, 1955. Best general commentary.

Adamson, J. W. *The Educational Writings of John Locke*. Cambridge, U.K.: 1922.

Bastide, C. *John Locke, ses théories politiques et leur influence en Angleterre*. Paris: Leroux, 1906. Still valuable work on Locke's political philosophy.

Christopherson, H. O. *A Bibliographical Introduction to the Study of John Locke*. Oslo, 1930. Incomplete.

Clapp, J. G. *Locke's Conception of the Mind*. PhD diss., Columbia University. New York, 1937.

Cranston, Maurice. "Men and Ideas; John Locke." *Encounter* 7 (1956): 46–54.

Czajkowski, C. J. *The Theory of Private Property in Locke's Political Philosophy*. Notre Dame, IN, 1941. Useful on the labor theory of value.

DeMarchi, E. "Locke's Atlantis." *Political Studies* 3 (1955): 164–165.

Gibson, James. *Locke's Theory of Knowledge*. Cambridge, U.K.: Cambridge University Press, 1917. Emphasizes Locke's rationalism.

Gibson, James. *John Locke*. British Academy, Henriette Hertz Lecture, 1933.

Gierke, Otto von. *Naturrecht und deutsches Recht*. Translated and edited by Ernest Barker as *Natural Law and the Theory of Society*. 2 vols. Cambridge, U.K.: 1934. A major study.

Gough, J. W. *John Locke's Political Philosophy. Eight Studies*. Oxford, 1950. Important.

Jackson, Reginald. "Locke's Distinction Between Primary and Secondary Qualities." *Mind* 38 (1929): 56–76.

Jackson, Reginald. "Locke's Version of the Doctrine of Representative Perception." *Mind* 39 (1930): 1–25.

James, D. G. *The Life of Reason: Hobbes, Locke, Bolingbroke*. London and New York: Longman, Green, 1949.

Krakowski, E. *Les sources médiévales de la philosophie de Locke*. Paris: Jouve, 1915. One of the few studies of influences on Locke.

Lamprecht, S. P. *The Moral and Political Philosophy of John Locke*. New York: Columbia University Press, 1918.

Laslett, Peter. "The English Revolution and Locke's Two Treatises of Government." *Cambridge Historical Journal* 12 (1956). Interesting and controversial.

Leibniz, G. W. *Nouveaux essais sur l'entendement humain*. Leipzig and Amsterdam, 1765. Translated by A. G. Langley as *New Essays concerning Human Understanding*. New York and London: Macmillan, 1896. An important critique of Locke by a contemporary.

Leyden, W. von. "John Locke and Natural Law." *Philosophy* 31 (1956). A useful examination.

Long, P. *A Summary Catalogue of the Lovelace Collection of Papers of John Locke in the Bodleian Library*. Oxford, 1959.

O'Connor, D. J. *John Locke*. Harmondsworth, U.K.: Penguin, 1952.

Ollion, H. *Notes sur la correspondance de John Locke*. Paris: A. Picard, 1908.

Polin, R. *La politique de John Locke*. Paris, 1960. Interesting contrast to Laslett.

Pollock, Sir Frederick. "Locke's Theory of the State." In his *Essays in the Law*. London: H. M. Stationery Office, 1922. Ch. 3.

Ryle, Gilbert. *Locke on the Human Understanding*. Oxford: Oxford University Press, 1933.

Smith, N. K. *John Locke*. Manchester, U.K.: Manchester University Press, 1933.

Ware, C. S. "The Influence of Descartes on John Locke." *Revue internationale de philosophie* (1950): 210–230.

Webb, T. E. *The Intellectualism of Locke. An Essay*. Dublin: W. McGee, 1857. Presents Locke as a precursor to Kant.

Yolton, J. W. "Locke's Unpublished Marginal Replies to John Sergeant." *Journal of the History of Ideas* 12 (1951): 528–559.

Yolton, J. W. "Locke and the Seventeenth-Century Logic of Ideas." *Journal of the History of Ideas* 16 (1955): 431–452.

Yolton, J. W. *Locke and the Way of Ideas*. Oxford: Clarendon Press, 1956. A careful study.

Yolton, J. W. "Locke on the Law of Nature." *Philosophical Review* 67 (1958): 477–498.

RECENT EDITIONS OF LOCKE'S WORKS

Goldie, Mark, ed. *John Locke: Selected Correspondence*. Oxford: Oxford University Press, 2002.

Nuovo, Victor, ed. *John Locke: Writings on Religion*. Oxford: Clarendon Press, 2002.

RECENT WORKS ON LOCKE

Bennett, Jonathan. "Ideas and Qualities in Locke's *Essay*." *History of Philosophy Quarterly* 13 (1996): 73–88.

Bolton, Martha Brandt. "Locke, Leibniz, and the Logic of Mechanism." *Journal of the History of Philosophy* 36 (1998): 189–213.

Coventry, Angela. "Locke, Hume, and the Idea of Causal Power." *Locke Studies* 3 (2003): 93–111.

Downing, Lisa. "The Status of Mechanism in Locke's *Essay*." *Philosophical Review* 107 (1998): 381–414.

Garrett, Don. "Locke on Personal Identity, Consciousness, and 'Fatal Errors'." *Philosophical Topics* 31 (2003): 95–125.

Jacovides, Michael. "Locke's Resemblance Theses." *Philosophical Review* 108 (1999): 461–496.

Jacovides, Michael. "The Epistemology under Locke's Corpuscularianism." *Archiv für Geschichte der Philosophie* 84 (2002): 161–189.

Langton, Rae. "Locke's Relations and God's Good Pleasure." *Proceedings of the Aristotelian Society* 100 (2000): 75–91.

Lowe, E. "J. Locke: Compatibilist Event-Causalist or Libertarian Substance-Causalist?" *Philosophy and Phenomenological-Research* 68 (2004): 688–701.

Martin, Raymond. "Locke's Psychology of Personal Identity." *Journal of the History of Philosophy* 38 (2000): 41–61.

McCann, Edwin. "Locke's Theory of Substance Under Attack!" *Philosophical Studies* 106 (2001): 87–105.

Newman, Lex. "Locke on the Idea of Substratum." *Pacific Philosophical Quarterly* 81 (2000): 291–324.

Ott, Walter. *Locke's Philosophy of Language*. Cambridge, U.K.: Cambridge University Press, 2003.

Owen, David. "Locke and Hume on Belief, Judgment and Assent." *Topoi* 22 (2003): 15–28.

Rogers, G. A. J., ed. *Locke's Philosophy: Content and Context.* Oxford: Clarendon Press, 1996.

Rozemond, Marleen, and Gideon Yaffe. "Peach Trees, Gravity and God: Mechanism in Locke." *British Journal for the History of Philosophy* 12 (2004): 387–412.

Stuart, Matthew. "Locke on Natural Kinds." *History of Philosophy Quarterly* 16 (1999): 277–296.

Stuart, Matthew. "Locke's Colors." *Philosophical Review* 112 (2003): 57–96.

Thiel, Udo, ed. *Locke: Epistemology and Metaphysics.* Brookfield: Ashgate: 2002.

Tuckness, Alex. *Locke and the Legislative Point of View: Toleration, Contested Principles, and the Law*. Princeton, NJ: Princeton University Press, 2002.

Wilson, Robert A. "Locke's Primary Qualities." *Journal of the History of Philosophy* 40 (2002): 201–228.

Wolterstorff, Nicholas. *John Locke and the Ethics of Belief.* New York: Cambridge University Press, 1996.

Yaffe, Gideon. *Freedom Worth the Name: Locke on Free Agency*. Princeton: Princeton University Press, 2000.

Yolton, Jean S. *John Locke: A Descriptive Bibliography*. Bristol: Thoemmes, 1998.

Yolton, John. "Locke's Man." *Journal of the History of Ideas* 62 (2001): 665–683.

James Gordon Clapp (1967)
Bibliography updated by Don Garrett (2005)

LOCKE, JOHN [ADDENDUM]

John Locke has been, for the last three decades, the subject of a rapid expansion of interest, stimulated by Oxford University Press's Clarendon edition of his works. The eight-volume edition of Locke's correspondence has opened new areas of information and exploration. So far in that series, we have definitive editions of *Essay* (including editions of the drafts and other relevant writings), the work on education, his paraphrases of St. Paul's epistles, and the papers on money, and well as *The Reasonableness of Christianity*, and the journals (again, opening a vast and important insight into Locke's reading, book buying, travels, opinions), and other works will follow. These editions, and the research that went into their production, have provided new resources for work on almost all aspects of Locke's life and writings, as well as material relating to his intellectual environment.

Antedating the Clarendon series was another medium for interest in Locke: *The Locke Newsletter*, founded and edited by Roland Hall. Beginning in 1970, published once per year (more or less), the newsletter has published articles on all aspects of Locke's thought. Included in each number is a list of recent (as of 1996) books and articles on Locke in many languages. This is a valuable source for keeping up to date on the publications about Locke. Another source of information on publications about Locke is the *Reference Guide* by Yolton and Yolton. Two other bibliographic resources are Attig's listings of Locke editions and the much fuller descriptive

bibliography of all editions of Locke's publications by Jean S. Yolton. The latter, a work long overdue, describes many different copies of Locke editions, which were located and examined in many different libraries and countries.

Among the topics in Locke's *Essay*, three have received special attention: the representative theory of perception, personal identity, and matter theory. The first of these in recent discussions has involved a debate over the nature of ideas: Are they special entities (e.g., images) standing between perceivers and objects, or are they simply the means for our access to the physical world? On the second topic it is becoming increasingly recognized that memory is not the crux of Locke's concept of person; it is consciousness, a wider and richer process (one with clear moral overtones) that focuses our awareness of self. A person for Locke is a moral being composed of the thoughts, feelings, and actions performed throughout a life. Consciousness is not a property of some immaterial substance, at least not so far as we can discover. The third topic has been given detailed attention via Locke's use of the corpuscular theory (see Alexander 1985). Some recognition has been given to Locke's movement toward the Newtonian concept of matter as force and power. Locke anticipated this development in his talk of the qualities of body being primarily powers. The substantiality of matter begins to fade under Locke's analysis of primary and secondary qualities. The chapter on power in *Essay*, the power of persons and the power of matter, is the longest and most complex chapter in that work (see J. W. Yolton 1993).

Locke's social and political thought has received even more attention throughout the decades, especially during the 1980s and 1990s. Laslett's early dating of *Two Treatises* and his locating that work in its historical context have been developed by writers such as Dunn, Harris, and Marshall. The central role of property and the relation of that concept to the person is generally recognized (see Tully 1980). His *Two Treatises* elaborates a concept of property that starts with each person's having property in his person. Acquisition of other possessions is a function of that original self-property. The tension between the interests and rights of the individual and those of society (or the community of mankind) is much discussed (see especially Gobetti 1992). The focus on consciousness as defining the person in his *Essay* indicated the central place of the individual in Locke's civil society. At the same time majority decisions were allowed to restrain individual actions. The power of the people is sanctioned by a social contract that obliges the ruler or legislative body to act for the good of the citizens, in conformity with the laws of nature. The interconnections between Locke's moral views and his social and political thought have been discussed by Marshall (1984). The issue of religious toleration has focused some of the recent treatments of Locke's political and religious writings, but all of the toleration writings by Locke await their inclusion in the Clarendon editions.

Locke's religious interests in the Bible and in what is required of a Christian have been clarified by recent studies (e.g., Wainwright's edition of the *Paraphrases*, 1987), but this area will be further illuminated when the Clarendon edition of *Reasonableness* appears. Locke's relation to the Latitudinarians and the role of original sin in his thinking have been explored by Spellman (1988). Coleman's (1983) systematic study of Locke's moral theory set the stage for some of the recent attention to this aspect of Locke's thought.

Another newly developing area of Locke studies concerns the reception of his doctrines in Europe, especially in France. The difficulties the French had with the term "consciousness" when translating this English term have been interestingly analyzed by Davies (1990). Reactions to Locke's books in French-language journals and the impact of his doctrines (especially thinking matter) on Enlightenment thinkers have been presented by several writers (Hutchison 1991; Schøsler 1985, 1994; J. W. Yolton 1991). The full story of the reception of Locke's doctrines in Europe (especially in Germany, Portugal, and Holland) in the eighteenth century has yet to be written. Fruitful research programs are waiting for scholars. A number of collections of articles can be consulted to fill out this brief sketch of newer developments in Locke studies (Chappell 1994, Harpham 1992, Thompson 1991).

See also Consciousness; Perception; Personal Identity; Power; Primary and Secondary Qualities; Property; Social Contract; Toleration.

Bibliography

RECENT EDITIONS OF LOCKE'S WORKS

De Beer, E. S., ed. *The Correspondence of John Locke*. 8 vols. Oxford: Oxford University Press, 1976–189.

Kelly, P., ed. *Several Papers Relating to Money, Interest, and Trade*. Oxford: Oxford University Press, 1991.

Laslett, P., ed. *Two Treatises of Government*. 3rd ed. Cambridge U.K.: Cambridge University Press, 1989.

Nidditch, P. H., ed. *An Essay concerning Human Understanding*. Oxford: Oxford University Press, 1975.

Nidditch, P. H., and G. A. J. Rogers, eds. *Drafts for the Essay concerning Human Understanding and Other Philosophical*

Writings. Vol. 1. Oxford: Oxford University Press, 1990. Includes drafts A and B: other volumes planned.

Wainwright, A. P., ed. *Paraphrases and Notes on the Epistles of St. Paul.* 2 vols. Oxford: Oxford University Press, 1987.

Yolton. J. W., and J. S. Yolton, eds. *Some Thoughts concerning Education.* Oxford: Oxford University Press, 1989.

WORKS ON LOCKE

Alexander, P. *Ideas, Qualities, and Corpuscles: Locke and Boyle on the External World.* Cambridge, U.K., and New York: Cambridge University Press, 1985.

Ashcraft, R. *Revolutionary Politics and Locke's "Two Treatises of Government."* Princeton, NJ: Princeton University Press, 1986.

Attig, J. C., ed. *The Works of John Locke: A Comprehensive Bibliography from the Seventeenth Century to the Present.* Westport, CT: Greenwood, 1985.

Ayers, M. *Locke*. 2 vols. London, 1991.

Chappell, V., ed. *The Cambridge Companion to Locke.* Cambridge, U.K., and New York: Cambridge University Press, 1994.

Colman, J. *John Locke's Moral Philosophy.* Edinburgh: Edinburgh University Press, 1983.

Davies, C. *"Conscience" as Consciousness: The Self of Self-Awareness in French Philosophical Writing from Descartes to Diderot*. Studies on Voltaire and the Eighteenth Century, 272. Oxford, 1990.

Dunn, J. *The Political Thought of John Locke: An Historical Account of the Argument of the "Two Treatises of Government."* Cambridge, U.K.: Cambridge University Press, 1969.

Franklin, J. H. *John Locke and the Theory of Sovereignty: Mixed Monarchy and the Right of Resistance in the Political Thought of the English Revolution.* Cambridge, U.K., and New York: Cambridge University Press, 1978.

Gobetti, D. *Public and Private: Individuals, Households, and Body Politic in Locke and Hutcheson.* London, 1992.

Goyard-Fabre, S. *John Locke et la raison raisonnable.* Paris: Bordas, 1986.

Harpham, E. J., ed. *John Locke's "Two Treatises of Government": New Interpretations.* Lawrence: University of Kansas Press, 1992.

Harris, I. *The Mind of John Locke: A Study of Political Theory in Its Intellectual Setting.* Cambridge, U.K., and New York: Cambridge University Press, 1994.

Hutchison, R. *Locke in France, 1688–1734.* Studies on Voltaire and the Eighteenth Century, 290. Oxford: Voltaire Foundation of the Taylor Institution, 1991.

Mackie, J. L. *Problems from Locke.* Oxford: Clarendon Press, 1976.

Marshall, J. *John Locke: Resistance, Religion, and Responsibility.* Cambridge Studies in Early Modern British History. Cambridge, U.K., and New York: Cambridge University Press, 1994.

Passmore, J. "Locke and the Ethics of Belief." *Proceedings of the British Academy*, 64 (1978): 185–208.

Rogers, G. A. J. "Locke, Anthropology, and Models of the Mind." *History of the Human Sciences* 6 (1993): 73–87.

Rogers, G. A. J., ed. *Locke's Philosophy: Content and Context.* Oxford: Clarendon Press, 1994.

Schochet, G. J. "Toleration, Revolution, and Judgment in the Development of Locke's Political Thought." *Political Science* 40 (1988): 84–96.

Schøsler, J. *La bibliothèque raisonnée (1728–1753): Les Réactions d'un périodique français à la philosophie de Locke au XVIIIe siècle.* Odense: L'universite D'ordense, 1985.

Schøsler, J. "Le Christianisme raisonnable et le débat sur le 'Socianisme' de John Locke dans la presse française de la première moitié du XVIIIe siècle." *Lias* 21 (2) (1994): 295–319.

Schouls, P. A. *Reasoned Freedom: John Locke and Enlightenment.* Ithaca, NY: Cornell University Press, 1992.

Spellman, W. M. *John Locke and the Problem of Depravity.* Oxford: Clarendon Press, 1988.

Thiel, U. *Locke's Theorie der personalen Identität.* Bonn, 1983.

Thompson, M. P., ed. *John Locke und Immanuel Kant.* Berlin: Duncker and Humblot, 1991.

Tomida, Y. "Idea and Thing: The Deep Structure of Locke's Theory of Knowledge." *Analecta Husserliana* 66 (1995): 3–143.

Tuck, R. *Natural Rights Theories: Their Origin and Development.* Cambridge, U.K., and New York: Cambridge University Press, 1979.

Tully, J. *A Discourse on Property: John Locke and His Adversaries.* Cambridge, U.K., and New York: Cambridge University Press, 1980.

Vaughn, K. I. *John Locke, Economist and Social Scientist.* Chicago: University of Chicago Press, 1980.

Vienne, J.-M. *Expérience et raison: Les fondements de la morale selon Locke.* Paris: Vrin, 1991.

Walmsley, P. "Locke's Cassowary and the Ethos of the Essay." *Studies in Eighteenth-Century Culture* 22 (1992): 253–267.

Walmsley, P. "Dispute and Conversation: Probability and the Rhetoric of Natural Philosophy in Locke's *"Essay." Journal of the History of Ideas* 54 (1993): 381–394.

Winkler, K. P. "Locke on Personal Identity." *Journal of the History of Philosophy* 29 (1991): 201–226.

Wood, N. *The Politics of Locke's Philosophy: A Social Study of "An Essay concerning Human Understanding."* Berkeley: University of California Press,1983.

Yolton, J. S. *John Locke: A Descriptive Bibliography.* Bristol, U.K.: Toemmes Press, 1996.

Yolton, J. S., and J. W. Yolton. *John Locke: A Reference Guide.* Boston: G. K. Hall, 1985.

Yolton, J. W. *Locke and the Compass of Human Understanding: A Selective Commentary on the Essay.* Cambridge, U.K.: Cambridge University Press, 1970.

Yolton, J. W. *Locke and French Materialism*. Oxford: Clarendon Press, 1991.

Yolton, J. W. *A Locke Dictionary.* Oxford and Cambridge, MA: Blackwell, 1993.

John W. Yolton (1996)

LOGIC, COMBINATORY

See *Combinatory Logic*

LOGIC, HISTORY OF

The mainstream of the history of logic begins in ancient Greece and comes down through the Arabian and European logic of the Middle Ages and through a number of post-Renaissance thinkers to the more or less mathematical developments in logic in the nineteenth and twentieth centuries. In the period after the fall of Rome many of the ancient achievements were forgotten and had to be relearned; the same thing happened at the end of the Middle Ages. Otherwise this Western tradition has been fairly continuous. Indian and Chinese logic developed separately. Today logic, like other sciences, is studied internationally, and the same problems are treated in the Americas, western and eastern Europe, and Asia and Australasia. The story of the development of logic will be told here under the following headings:

Bibliography

HISTORY OF LOGIC

(In general, texts by and studies on individual logicians are listed only if they do not appear in the bibliographies to the separate entries on these people. Most works cited in the body of this entry are not repeated below.)

General Works

Bocheński, I. M. *Formale Logik*. Freiburg: Alber, 1956. Translated and edited by Ivo Thomas as *History of Formal Logic*. Notre Dame, IN, 1961. Contains extensive bibliographies.

Enriques, Federigo. *L'évolution de la logique*. Paris, 1926.

Kneale, William C., and Martha Kneale. *The Development of Logic*. Oxford: Clarendon Press, 1962.

Kotarbiński, Tadeusz. *Wyklady z Dziejów Logiki* (Lectures on the history of logic). Lodz: Zaklad Narodowy im Ossolinskich we Wroclawiu, 1957. Translated into French as *Leçons sur l'histoire de la logique*. Paris: Presses Universitaires de France, 1964.

Prantl, Carl. *Geschichte der Logik im Abendlande*. 4 vols. Leipzig: Hirzel, 1855–1870; reprinted (3 vols.), Graz, 1955. Covers the history up to 1600, with numerous quotations from the sources.

Scholz, Heinrich. *Abriss der Geschichte der Logik*. Berlin, 1931. Translated by K. F. Leidecker as *Concise History of Logic*. New York: Philosophical Library, 1961.

Works on Post-Medieval and Later Developments

Jørgensen, Jørgen. *A Treatise of Formal Logic*, Vol. I: *Historical Development*. Copenhagen: Levin and Munksgaard, 1931.

Jourdain, P. E. B. "The Development of the Theories of Mathematical Logic and the Principles of Mathematics." *Quarterly Journal of Pure and Applied Mathematics* 41 (1910): 324–352; 43 (1912): 219–314; 44 (1913): 113–128.

Lewis, C. I. *A Survey of Symbolic Logic*. Berkeley: University of California Press, 1918; reprinted, New York, 1960.

Nidditch, P. H. *The Development of Mathematical Logic*. New York: Free Press of Glencoe, 1962.

Prior, A. N. *Formal Logic*, 2nd ed. Oxford: Clarendon Press, 1962.

Shearman, A. T. *Development of Symbolic Logic*. London: Williams and Norgate, 1906.

(A. N. P.)

ANCIENT LOGIC

THE BEGINNINGS

Logic as a discipline starts with the transition from the customary use of certain logical methods and argument patterns to the reflection on and inquiry into these and their elements, including the syntax and semantics of sentences. In antiquity, logic as a systematic discipline begins with Aristotle. However, discussions of some elements of logic and a focus on methods of inference can be traced back to the late fifth century BCE.

SYNTAX AND SEMANTICS. Some of the Sophists classified types of sentences (*logoi*) according to their force. So Protagoras (485–415), who included wish, question, answer, and command (Diels-Kranz 80.A1, Diogenes Laertius 9.53–4), and Alcidamas (pupil of Gorgias, fl. fourth century BCE), who distinguished assertion (*phasis*), denial (*apophasis*), question, and appellation (Diogenes Laertius 9.54). Antisthenes (mid-5th–mid-4th cent.) defined a sentence as "that which indicates what a thing was or is" (Diogenes Laertius 6.3, Diels-Kranz 45) and stated that someone who says what is speaks truly (Diels-Kranz 49). Perhaps the earliest surviving passage on logic is found in the *Dissoi Logoi* or Double arguments (Diels-Kranz 90.4, c.400 BCE). It is evidence for a debate over truth and falsehood. Opposed were the views that: (1) truth is a—temporal—property of sentences, and that a sentence is true (when it is said), if and only if things are as the sentence says they are when it is said, and false if they are not; and (2) truth is an atemporal property of what is said, and that what is said is true if and only if the things are the case, and false if they are not the case. These are rudimentary formulations of two alternative corre-

spondence theories of truth. The same passage also displays awareness of the fact that self-referential use of the truth-predicate can be problematic—an insight also documented by the discovery of the Liar paradox by Eubulides of Miletus (mid fourth century BCE) shortly thereafter.

Some Platonic dialogues contain passages whose topic is indubitably logic. In the *Sophist*, Plato analyzes simple statements as containing a verb (*rhēma*, which indicates action) and a name (*onoma*, which indicates the agent) (Soph.261E–262A). Anticipating the modern distinction of logical types, he argues that neither a series of names nor a series of verbs can combine into a statement (Soph.262A–D). Plato also divorces syntax (*what is a statement?*) from semantics (*when is it true?*). Something (e.g., *Theaetetus is sitting*) is a statement if it both succeeds in specifying a subject and says something about this subject. Plato thus determines subject and predicate as relational elements in a statement and excludes statements containing empty subject expressions. Something is a true statement if with reference to its subject (Theaetetus) it says of what is (e.g., sitting) that it is. Something is a false statement if with reference to its subject it says of something other than what is (e.g., flying) that it is. Here Plato produces a sketch of a reductionist theory of truth (Soph.262E–263D; cf. also Crat. xxx). He also distinguishes negations from affirmations and takes the negation particle to have narrow scope: It negates the predicate, not the whole sentence (Soph.257B–C). There are many passages in Plato where he struggles with explaining certain logical relations. For example, his theory that things participate in Forms corresponds to a rudimentary theory of predication; in the *Sophist* and elsewhere, he grapples with the class relations of exclusion, union, and coextension; also with the difference between the *is* of predication (being) and the *is* of identity (sameness); and in *Republic* 4 he anticipates the law of noncontradiction. But his explications of these logical questions are cast in metaphysical terms and so can, at most, be regarded as protological.

ARGUMENT PATTERNS AND VALID INFERENCE. Pre-Aristotelian evidence for reflection on argument forms and valid inference are harder to come by. Both Zeno of Elea (c.490 BCE) and Socrates (470–399) were famous for the ways in which they refuted an opponent's view. Their methods display similarities with *reductio ad absurdum*, but neither of them seems to have theorized about their logical procedures. Zeno produced arguments (*logoi*) that manifest variations of the pattern *this* (that is, the opponent's view) *only if that. But that is impossible. So this is impossible*. Socratic refutation was an exchange of questions and answers in which the opponents would be led, on the basis of their answers, to a conclusion incompatible with their original claim. Plato institutionalized such disputations into structured, rule-governed, verbal contests that became known as dialectical argument. The development of a basic logical vocabulary for such contests indicates some reflection upon the patterns of argumentation.

The fifth and fourth centuries BCE also see great interest in fallacies and logical paradoxes. Besides the Liar, Eubulides is said to have been the originator of several other logical paradoxes, including the Sorites. Plato's *Euthydemus* contains a large collection of contemporary fallacies. In attempts to solve such logical puzzles, a logical terminology develops here, too, and the focus on the difference between valid and invalid arguments sets the scene for the searching for a criterion of valid inference. Finally, it is possible that the shaping of deduction and proof in Greek mathematics that begins in the later fifth century BCE served as an inspiration for Aristotle's syllogistic.

ARISTOTLE

Aristotle is the first great logician in the history of logic. His logic was taught by and large without rival from the fourth to the nineteenth centuries CE. Aristotle's logical works were collected and put in a systematic order by later Peripatetics who titled them the *Organon* or *tool* because they considered logic not as a part but rather an instrument of philosophy. The *Organon* contains, in traditional order, the *Categories*, *De Interpretatione*, *Prior Analytics*, *Posterior Analytics*, *Topics*, and *Sophistical Refutations*. In addition, *Metaphysics* Γ is a logical treatise that discusses the principle of noncontradiction, and some further logical insights are found scattered throughout Aristotle's other works. Some parts of the *Categories* and *Posterior Analytics* would today be regarded as metaphysics, epistemology, or philosophy of science rather than logic. The traditional arrangement of works in the *Organon* is neither chronological nor Aristotle's own. The original chronology cannot be fully recovered since Aristotle often inserted supplements into earlier writings at a later time. However, by using logical advances as criterion, we can conjecture that most of the *Topics*, *Sophistical Refutations*, *Categories*, and *Metaphysics* Γ predate the *De Interpretatione*, which in turn precedes the *Prior Analytics* and parts of the *Posterior Analytics*.

DIALECTICS. The *Topics* provide a manual for participants in the contests of dialectical argument as instituted in the Academy by Plato. Books 2–7 provide general procedures or rules (*topoi*) about how to find an argument to establish or refute a given thesis. The descriptions of these procedures—some of which are so general that they resemble logical laws—clearly presuppose a notion of logical form, and Aristotle's *Topics* may thus count as the earliest surviving logical treatise. The *Sophistical Refutations* are the first systematic classification of fallacies, sorted by what logical flaw each type manifests (e.g., equivocation, begging the question, affirming the consequent, *secundum quid*) and how to expose them.

SUB-SENTENTIAL CLASSIFICATIONS Aristotle distinguishes things that have sentential unity through a combination of expressions (a horse runs) from those that do not (horse, runs); the latter are dealt with in the *Categories* (the title really means Predications). They have no truth value and signify one of the following: substance (*ousia*), quantity (*poson*), quality (*poion*), relation (*pros ti*), location (*pou*), time (*pote*), position (*keisthai*), possession (*echein*), doing (*poiein*), and undergoing (*paschein*). It is unclear whether Aristotle considers this classification to be one of linguistic expressions that can be predicated of something else, or of kinds of predication, or of highest genera. In *Topics* 1 Aristotle distinguishes four relationships a predicate may have to the subject: It may give its definition, genus, unique property, or accidental property. These are known as predicables.

SYNTAX AND SEMANTICS OF SENTENCES. When writing the *De Interpretatione*, Aristotle had worked out the following theory of simple sentences: A (declarative) sentence (*apophantikos logos*) or declaration (*apophansis*) is delimited from other pieces of discourse such as prayer, command, and question by its having a truth value. The truth bearers that feature in Aristotle's logic are thus linguistic items. They are spoken sentences that directly signify thoughts (shared by all humans) and, through these, indirectly, things. Written sentences in turn signify spoken ones. Sentences are constructed from two signifying expressions that stand in subject-predicate relation to each other: a name and a verb (Callias walks) or two names connected by the copula *is*, which cosignifies the connection (Pleasure is good) (*Int.* 3). Names are either singular terms or common nouns. Both can be empty (Cat. 10, *Int.* 1). Singular terms can only take subject position. Verbs cosignify time. A name-verb sentence can be rephrased with the copula (Callias is [a] walking [thing]) (*Int.* 12). As to their quality, a sentence is either an affirmation or a negation, depending on whether it affirms or negates its predicate of its subject. The negation particle in a negation has wide scope (Cat. 10). Aristotle defines truth separately for affirmations and negations: An affirmation is true if it says of that which is that it is; a negation is true if it says of that which is not that it is not (Met.Γ. 7((1011b25ff). These formulations can be interpreted as expressing either a correspondence or a reductionist conception of truth. Either way, truth is a property that belongs to a sentence *at a time*. As to their quantity, sentences are singular, universal, particular, or indefinite. Thus Aristotle obtains eight types of sentences, which are later dubbed *categorical sentences*; the following are examples, paired by quality:

Singular:	Callias is just.	Callias is not just.
Universal	Every human is just.	No human is just.
Particular:	Some human is just.	Some human is not just.
Indefinite:	(A) human is just.	(A) human is not just.

Universal and particular sentences contain a quantifier, and both universal and particular affirmatives are taken to have existential import. The logical status of the indefinites is ambiguous and controversial (*Int.* 6–7).

Aristotle distinguishes between two types of sentential opposition: contraries and contradictories. A contradictory pair of sentences (*antiphasis*) consists of an affirmation and its negation (that is, the negation that negates of the subject what the affirmation affirms of it). Aristotle assumes that—normally—one of these must be true, the other false. Contrary sentences are such that they cannot both be true. The contradictory of a universal affirmative is the corresponding particular negative; that of the universal negative the corresponding particular affirmative. A universal affirmative and its corresponding universal negative are contraries. Aristotle thus has captured the basic logical relations between monadic quantifiers (*Int.* 7).

Since Aristotle regards tense as part of the truth bearer (as opposed to merely a grammatical feature), he detects a problem regarding future tense sentences about contingent matters and discusses it in the famous chapter nine of his *De Interpretatione*: Does the principle that, of an affirmation and its negation one must be false, the other true, apply to these? What, for example, is the truth value now of the sentence *There will be a sea battle tomorrow*? Aristotle may have suggested that the sentence has no truth value now and that bivalence thus does not hold—despite the fact that it is necessary for there either to be or not to be a sea battle tomorrow, so that the principle of excluded middle is preserved.

NONMODAL SYLLOGISTIC. Aristotle's nonmodal syllogistic, the core of which he develops in the first seven chapters of Book One of his *Prior Analytics*, is the pinnacle of his logic. Aristotle defines a syllogism as "an argument (*logos*) in which, certain things having been laid down, something different from what has been laid down follows of necessity because these things are so." This definition appears to require that (1) a syllogism consists of at least two premises and a conclusion, (2) the conclusion follows of necessity from the premises, and (3) the conclusion differs from the premises. Aristotle's syllogistic covers only a small part of all arguments that satisfy these conditions.

Aristotle restricts and regiments the types of categorical sentence that may feature in a syllogism. The admissible truth bearers are now defined as each containing two different terms (*horoiv*) conjoined by the copula, of which one (the predicate term) is said of the other (the subject term) either affirmatively or negatively. Aristotle never comes clear on the question whether terms are things (nonempty classes) or linguistic expressions for these things. Only universal and particular sentences are discussed. Singular sentences seem excluded, and indefinite sentences are mostly ignored.

Another innovation in the syllogistic is Aristotle's use of letters in place of terms. The letters may originally have served simply as abbreviations for terms, as we can see for example in his *Posterior Analytics*, but in the syllogistic they seem mostly to have the function either of schematic term letters or of term variables with universal quantifiers assumed but not stated. Where he uses letters, Aristotle tends to express the four types of categorical sentences in the following way (with common later abbreviations in brackets):

"A holds of every B"	(A*a*B)
"A holds of no B"	(A*e*B)
"A holds of some B"	(A*i*B)
"A does not hold of some B"	(A*o*B)

Instead of *holds* he also uses *is predicated*.

All basic syllogisms consist of three categorical sentences in which the two premises share exactly one term, called the middle term, and the conclusion contains the other two terms, sometimes called the extremes. Based on the position of the middle term, Aristotle classified all possible premise combinations into three figures (*schemata*): The first figure has the middle term (B) as subject in the first premise and predicated in the second; the second figure has it predicated in both premises; the third has it as subject in both premises:

A holds of B	B holds of A	A holds of B
B holds of C	B holds of C	C holds of B

A is also called the major term and C the minor term. Each figure can further be classified according to whether or not both premises are universal. Aristotle went systematically through the fifty-eight possible premise combinations and showed that fourteen have a conclusion following of necessity from them. His procedure was this: He assumed that the syllogisms of the first figure are complete and not in need of proof since they are evident. By contrast, the syllogisms of the second and third figures are incomplete and in need of proof. He proves them by reducing them to syllogisms of the first figure and thereby *completing* them. For this he makes use of three methods: (1) Conversion (*antistrophē*)—a categorical sentence is converted by interchanging its terms. Aristotle recognizes and establishes three conversion rules: "from A*e*B infer B*e*A"; "from A*i*B infer B*i*A"; and "from A*a*B infer B*i*A." All second- and third-figure syllogisms but two can be proved by premise conversion. (2) *Reductio ad impossibile* (*apagogē*)—the remaining two are proved by reduction to the impossible, where the contradictory of an assumed conclusion together with one of the premises is used to deduce by a first-figure syllogism a conclusion that is incompatible with the other premise. Using the semantic relations between opposites established earlier, the assumed conclusion is thus established. (3) Exposition (*ekthesis*)—this method, which Aristotle uses additionally to (1) and (2), is controvertible both as to what exactly it was and as to whether it is proof.

For each of the thirty-four premise combinations that allow no conclusion, Aristotle proves by counterexample that they allow no conclusion. As overall result, he acknowledges four first-figure syllogisms (later called Barbara, Celarent, Darii, Ferio), four second-figure syllogisms (Camestres, Cesare, Festino, Baroco), and six third-figure syllogisms (Darapti, Felapton, Disamis, Datisi, Bocardo, Ferison); these were later called the modes or moods of the figures. (The names are mnemonics: e.g., each vowel indicates in order whether the first and second premises and the conclusion were sentences of type *a*, *e*, *i*, or *o*.) Aristotle implicitly recognized that by using the conversion rules on the conclusions we obtain eight further syllogisms (AnPr. 53a3–14), and that of the premise combinations rejected as nonsyllogistic, some (five, in fact) will yield a conclusion in which the minor term is predicated of the major (AnPr. 29a19–27, Fapesmo, Fris-

esomorum, Firesmo, Fapemo, Frisemo). Moreover, in the *Topics*, Aristotle accepted the rules "from A*a*B infer A*i*B" and "from A*e*B infer A*o*B." By using these on the conclusions, five further syllogisms could be proved though Aristotle did not mention this.

Going beyond his basic syllogistic, Aristotle reduced the third and fourth first-figure syllogisms to second-figure syllogisms, thus de facto reducing all syllogisms to Barbara and Celarent; later, in the *Prior Analytics*, he invokes a type of cut-rule by which a multipremise syllogism can be reduced to two or more basic syllogisms. From a modern perspective, Aristotle's system can be represented as an argumental natural deduction system *en miniature*. It has been shown to be sound and complete if one interprets the relations expressed by the categorical sentences set theoretically as a system of nonempty classes as follows: A*a*B is true iff the class A contains the class B. A*e*B is true iff the classes A and B are disjoint. A*i*B is true iff the classes A and B are not disjoint. A*o*B is true iff the class A does not contain the class B. The vexing textual question of what exactly Aristotle meant by *syllogisms* has received several rival interpretations, including one that they are a certain type of conditional propositional form. Most plausibly, perhaps, Aristotle's complete and *incomplete* syllogisms taken together are understood as formally valid premise-conclusion arguments; and his complete and *completed* syllogisms taken together as (sound) deductions.

MODAL LOGIC. Aristotle is also the originator of modal logic. In addition to quality and quantity, he takes categorical sentences to have a mode; this consists of the fact that the predicate is said to hold of the subject either actually or necessarily or possibly or contingently or impossibly. The latter four are expressed by modal operators that modify the predicate, for example: "It is possible for A to hold of some B"; "A necessarily holds of every B."

In *De Interpretatione* (12–13), Aristotle:

(1) Concludes that modal operators modify the whole predicate (or the copula, as he puts it), not just the predicate term of a sentence;

(2) States the logical relations that hold between modal operators, such as that "it is not possible for A not to hold of B" implies "it is necessary for A to hold of B";

(3) Investigates what the contradictories of modalized sentences are and decides that they are obtained by placing the negator in front of the modal operator.

(4) Equates the expressions *possible* and *contingent*, but wavers between a one-sided interpretation (where necessity implies possibility) and a two-sided interpretation (where possibility implies nonnecessity).

Aristotle develops his modal syllogistic in chapters eight to twenty-two of the first book of his *Prior Analytics*. He settles on two-sided possibility (contingency) and tests for syllogismhood all possible combinations of premise pairs of sentences with necessity (N), contingency (C), or no (U) modal operator: NN, CC, NU/UN, CU/UC, and NC/CN. Syllogisms with the last three types of premise combinations are called mixed modal syllogisms. Apart from the NN category, which mirrors unmodalized syllogisms, all categories contain dubious cases. For instance, Aristotle accepts:

A necessarily holds of all B.

B holds of all C.

Therefore A necessarily holds of all C.

This and other problematic cases were already disputed in antiquity, and since the mid-1930s, they have sparked a host of complex, formalized reconstructions of Aristotle's modal syllogistic. As Aristotle's theory is conceivably internally inconsistent, the formal models that have been suggested may all be unsuccessful.

THE EARLY PERIPATETICS: THEOPHRASTUS AND EUDEMUS

Aristotle's pupil and successor Theophrastus of Eresus (c. 371–c. 287 BCE) wrote more logical treatises than his teacher, with a large overlap in topics. Eudemus of Rhodes (later fourth century BCE) wrote books titled *Categories*, *Analytics*, and *On Speech*. Of all these works only a number of fragments and later testimonies survive, mostly in Aristotle commentators. Theophrastus and Eudemus simplified some aspects of Aristotle's logic and developed others where Aristotle left us only hints.

IMPROVEMENTS AND MODIFICATIONS OF ARISTOTLE'S LOGIC. The two Peripatetics seem to have redefined Aristotle's first figure so that it includes every syllogism in which the middle term is subject of one premise and predicate of the other. In this way, five types of nonmodal syllogisms only intimated by Aristotle later in his *Prior Analytics* (Baralipton, Celantes, Datibis, Fapesmo, and Frisesomorum) are included, but Aristotle's criterion that first-figure syllogisms are evident is given up (fr. 91). Theophrastus and Eudemus also improved Aristotle's modal theory. Theophrastus

replaced Aristotle's two-sided contingency by one-sided possibility so that possibility no longer entails nonnecessity. Both recognized that the problematic universal negative ("A possibly holds of no B") is simply convertible (fr. 102A). Moreover, they introduced the principle that in mixed modal syllogisms the conclusion always has the same modal character as the weaker of the premises (frs.106 and 107), where possibility is weaker than actuality, and actuality than necessity. In this way Aristotle's modal syllogistic is notably simplified, and many unsatisfactory theses, like the one mentioned above, disappear.

PROSLEPTIC SYLLOGISMS. Theophrastus introduced the so-called prosleptic premises and syllogisms (fr. 110). A prosleptic premise is of the form:

$$\text{For all X, if } \Phi(X)\text{, then } \Psi(X)$$

where $\Phi(X)$ and $\Psi(X)$ stand for categorical sentences in which the variable X occurs in place of one of the terms. For example:

1) A <holds> of all of that of all of which B <holds>.

2) A <holds> of none of that which <holds> of all B.

Theophrastus considered such premises to contain three terms, two of which are definite (A, B) and one indefinite (*that*, or the bound variable X). We can represent (1) and (2) as:

$$\forall X \;\; B\mathit{a}X \rightarrow A\mathit{a}X$$

$$\forall X \;\; X\mathit{a}B \rightarrow A\mathit{e}X$$

Prosleptic syllogisms then come about as follows: They are comprised of a prosleptic premise and the categorical premise obtained by instantiating a term (C) in the antecedent *open categorical sentence* as premises, and the categorical sentences one obtains by putting in the same term (C) in the consequent *open categorical sentence* as conclusion. For example:

A <holds> of all of that of all of which B <holds>.

B holds of all C.

Therefore, A holds of all C.

Theophrastus distinguished three figures of these syllogisms, depending on the position of the indefinite term (also called *middle term*) in the prosleptic premise; for example (1) produces a third-figure syllogism, (2) a first-figure syllogism. The number of prosleptic syllogisms was presumably equal to that of types of prosleptic sentences: With Theophrastus's concept of the first figure, these would be sixty-four (that is, 32+16+16). Theophrastus held that certain prosleptic premises were equivalent to certain categorical sentences, for example, (1) to "A is predicated of all B." However, for many, including (2), no such equivalent can be found, and prosleptic syllogisms thus increased the inferential power of Peripatetic logic.

FORERUNNERS OF *MODUS PONENS* AND *TOLLENS*. Theophrastus and Eudemus considered complex premises that they called *hypothetical premise* and that had one of the following two forms (or similar):

If something is F, it is G.

Either something is F or it is G. (with exclusive *or*)

They developed arguments with them that they called "mixed from a hypothetical premise and a probative premise" (fr. 112A) These arguments were inspired by Aristotle's syllogisms from a hypothesis (*An.Pr.* 1.44); they were forerunners of *modus ponens* and *modus tollens* and had the following forms: (frs.111 and 112):

If something is F, it is G.	If something is F, it is G.
a is F.	*a* is not G.
Therefore, *a* is G.	Therefore, *a* is not F.
Either something is F or it is G.	Either something is F or it is G.
a is F.	*a* is not F.
Therefore, *a* is not G.	Therefore, *a* is G.

Theophrastus also recognized that the connective particle *or* can be inclusive (fr. 82A); and he considered relative quantified sentences such as those containing *more*, *fewer*, and *the same* (fr. 89), and seems to have discussed syllogisms built from such sentences, again following up upon what Aristotle said about syllogisms from a hypothesis (fr. 111E).

WHOLLY HYPOTHETICAL SYLLOGISMS. Theophrastus is further credited with the invention of a system of the later so-called *wholly hypothetical syllogisms* (fr. 113). These syllogisms were originally abbreviated term-logical arguments of the kind:

If [something is] A, [it is] B.

If [something is] B, [it is] C.

Therefore, if [something is] A, [it is] C.

and at least some of them were regarded as reducible to Aristotle's categorical syllogisms, presumably by way of the equivalences to "Every A is B," and so on. In parallel to Aristotle's syllogistic, Theophrastus distinguished three

figures, each of which had sixteen modes. The first eight modes of the first figure are obtained by going through all permutations with "not X" instead of "X" (with X for A, B, C); the second eight modes are obtained by using a rule of contraposition on the conclusion:

> (CR) From "if X, Y" infer "if the contradictory of Y then the contradictory of X"

The sixteen modes of the second figure were obtained by using (CR) on the schema of the first premise of the first figure arguments, for example:

> If [something is] not B, [it is] not A.
>
> If [something is] B, [it is] C.
>
> Therefore, if [something is] A, [it is] C.

The sixteen modes of the third figure were obtained by using (CR) on the schema of the second premise of the first figure arguments, for example:

> If [something is] A, [it is] B.
>
> If [something is] not C, [it is] not B.
>
> Therefore, if [something is] A, [it is] C.

Theophrastus claimed that all second- and third-figure syllogisms could be reduced to first-figure syllogisms. If Alexander of Aphrodisias reports faithfully, any use of (CR) that transforms a syllogism into a first-figure syllogism was such a reduction. The large number of modes and reductions can be explained by the fact that Theophrastus did not have the logical means for substituting negative for positive components in an argument. In later antiquity, after some intermediate stages, and possibly under Stoic influence, the wholly hypothetical syllogisms were interpreted as propositional-logical arguments of the kind:

> If p, then q.
>
> If q, then r.
>
> Therefore, if p, then r.

DIODORUS CRONUS AND PHILO THE LOGICIAN

In the later fourth to mid third centuries BCE, a loosely connected group of philosophers, sometimes referred to as dialecticians and possibly influenced by Eubulides, conceived of logic as a logic of propositions. Their best-known exponents were Diodorus Cronus and his pupil Philo (sometimes called Philo of Megara) although no writings of theirs are preserved. They each made groundbreaking contributions to the development of propositional logic, in particular to the theories of conditionals and modalities.

A conditional (*sunēmmenon*) was considered as a nonsimple proposition comprised of two propositions and the connecting particle *if*. Philo, who may be credited with introducing truth-functionality into logic, provided the following criterion for their truth: A conditional is false *when* and only *when* its antecedent is true and its consequent is false, and it is true in the three remaining truth-value combinations. The Philonian conditional resembles material implication except that—since propositions were conceived of as functions of time that can have different truth values at different times—it may change its truth value over time. For Diodorus, a conditional proposition is true if it neither was nor is possible that its antecedent is true and its consequent false. The temporal elements in this account suggest that the possibility of a truth-value change in Philo's conditionals was meant to be improved on. With his own modal notions (see below) applied, a conditional is Diodorean true now if and only if it is Philonian true at all times. Diodorus's conditional is thus reminiscent of strict implication. Philo's and Diodorus's conceptions of conditionals led to variants of the *paradoxes* of material and strict implication—a fact the ancients were aware of (S.E.M. 109–117).

Philo and Diodorus each considered the four modalities possibility, impossibility, necessity, and nonnecessity. These were conceived of as modal properties or modal values of propositions, not as modal operators. Philo defined them as follows:

> Possible is that which is capable of being true by the proposition's own nature … necessary is that which is true, and which, as far as it is in itself, is not capable of being false. Non-necessary is that which as far as it is in itself, is capable of being false, and impossible is that which by its own nature is not capable of being true (Boethius, *In librum Aristotelis De interpretatione: secunda editio*, p. 234).

Diodorus's definitions were these: "Possible is that which either is or will be <true>; impossible that which is false and will not be true; necessary that which is true and will not be false; non-necessary that which either is false already or will be false" (Boethius, *In librum Aristotelis De interpretatione: secunda editio*, p. 234). Both sets of definitions satisfy the following standard requirements of modal logic:

> (1) Necessity entails truth and truth entails possibility;

(2) Possibility and impossibility are contradictories, and so are necessity and nonnecessity;

(3) Necessity and possibility are interdefinable;

(4) Every proposition is either necessary or impossible or both possible and nonnecessary.

Philo's definitions appear to introduce mere conceptual modalities whereas with Diodorus's definitions, some propositions may change their modal value. Diodorus's definition of possibility rules out future contingents and implies the counterintuitive thesis that only the actual is possible. Diodorus tried to prove this claim with his famous Master Argument, which sets out to show the incompatibility of (1) "every past truth is necessary," (2) "the impossible does not follow from the possible," and (3) "something is possible which neither is nor will be true" (Epictetus *Discourses* II.19). The argument has not survived, but various reconstructions have been suggested. Some affinity with the arguments for logical determinism in Aristotle's *De Interpretatione* 9 is likely.

THE STOICS

The founder of the Stoa, Zeno of Citium (335–263 BCE), studied with Diodorus. His successor, Cleanthes, (331–232) tried to solve the Master Argument by denying that every past truth is necessary and wrote books—now lost—on paradoxes, dialectics, argument modes, and predicates. Both philosophers considered logic as a virtue and held it in high esteem, but they seem not to have been creative logicians. By contrast, Cleanthes's successor, Chrysippus of Soli (c.280–207), is without doubt the second great logician in the history of logic. It was said of him that if the gods used any logic, it would be that of Chrysippus, and his reputation as a brilliant logician is amply testified. Chrysippus wrote more than 300 books on logic, on virtually any topic contemporary logic concerns itself with, including speech act theory, sentence analysis, singular and plural expressions, types of predicates, demonstratives, existential propositions, sentential connectives, negations, disjunctions, conditionals, logical consequence, valid argument forms, theory of deduction, propositional logic, modal logic, tense logic, logic of suppositions, logic of imperatives, ambiguity, and logical paradoxes—in particular, the Liar and the Sorites (Diogenes Laertius 7.189–199). Of all these, only two badly damaged papyri have survived, luckily supplemented by a considerable number of fragments and testimonies in later texts, in particular in Diogenes Laertius, book 7, sections 55–83; and Sextus Empiricus, *Outlines of Pyrrhonism*, book 2, and *Against the Mathematicians*, book 8 (both of which appear in *Works*). Chrysippus's successors, including Diogenes of Babylon (c.240–152) and Antipater of Tarsus, appear to have systematized and simplified some of his ideas, but their original contributions to logic seem small. Many testimonies of Stoic logic do not name any particular Stoic. Hence the following paragraphs simply talk about the Stoics in general; but we can be confident that a large part of what has survived goes back to Chrysippus.

LOGICAL ACHIEVEMENTS BESIDES PROPOSITIONAL LOGIC. The subject matter of Stoic logic are the so-called sayables (*lekta*): They are the underlying meanings in everything we say and think, but—like Gottlob Frege's senses—subsist also independently of us. They are distinguished from linguistic expressions: What we *utter* are those expressions, but what we *say* are the sayables (Diogenes Laertius 7.57). There are complete and deficient sayables. Complete sayables, if said, do not make the hearer feel prompted to ask a question(Diogenes Laertius 7.63). They include assertibles (the Stoic equivalent for propositions), interrogatives, imperativals, inquiries, hypotheses, and more. The accounts of the different complete sayables all had the general form *a so-and-so sayable is one in saying which we perform an act of such-and-such*. For instance: An imperatival sayable is one in saying which we issue a command; an interrogative sayable is one in saying which we ask a question; a declaratory sayable (that is, an assertible) is one in saying which we make an assertion. Thus, according to the Stoics, each time we say a complete sayable, we perform three different acts: we utter a linguistic expression, we say the sayable, and we perform a speech-act.

Assertibles (*axiōmata*) differ from all other complete sayables by having a truth value: At any one time they are either true or false. Truth is temporal and assertibles may change their truth value. The Stoic principle of bivalence is hence temporalized, too. Truth is introduced by example: the assertible "it is day" is true *when* it is day, and at all other times false (Diogenes Laertius 7.65). This suggests a reductionist view of truth, as does the fact that the Stoics identify true assertibles with facts but define false assertibles simply as the contradictories of true ones (S.E.M. 8.85).

Assertibles are simple or nonsimple. A simple *predicative* assertible, such as *Dion is walking*, is generated from the predicate *is walking*, which is a deficient assertible since it elicits the question *who*, and a nominative case (Dion's individual quality or the correlated sayable), which falls under the predicate (Diogenes Laertius 7.63

and 70). There is thus no interchangeability of predicate and subject terms as in Aristotle; rather, predicates—but not the things that fall under them—are defined as deficient and thus resemble propositional functions. It seems that whereas some Stoics took the Fregean approach that singular terms had correlated sayables, others anticipated the notion of direct reference. Concerning demonstratives, the Stoics took a simple *definite* assertible such as *this one is walking* to be true when the person pointed at by the speaker is walking (S.E.M. 100). When the thing pointed at ceases to be, so does the assertible though the sentence used to express it remains (Alex.Aphr.An.Pr. 177–8). A simple *indefinite* assertible such as *someone is walking* is said to be true when a corresponding definite assertible is true (S.E.M. 98). Aristotelian universal affirmatives ("Every A is B") were to be rephrased as "If something is A, it is B" (S.E.M. 9.8–11). The past tense assertible *Dion walked* is true when there is at least one past time at which *Dion is walking* was true. The negation of *Dion is walking* is *(It is) not (the case that) Dion is walking*, and not *Dion is not walking*. The latter is analyzed in a Russellian manner as *Both Dion exists and not: Dion is walking*' (Alex.Aphr.An.Pr. 402).

SYNTAX AND SEMANTICS OF COMPLEX PROPOSITIONS. Thus the Stoics concerned themselves with several issues we would place under the heading of predicate logic; but their main achievement was the development of a propositional logic, that is, of a system of deduction in which the smallest substantial unanalyzed expressions are propositions, or rather, assertibles.

The Stoics defined negations as assertibles that consist of a negative particle and an assertible controlled by this particle (S.E.M. 103). Similarly, nonsimple assertibles were defined as assertibles that either consist of more than one assertible or of one assertible taken more than once (Diogenes Laertius 7.68–9) and that are controlled by a connective particle. Both definitions are recursive and allow for assertibles of indeterminate complexity. Three types of nonsimple assertions feature in Stoic syllogistic. Conjunctions are nonsimple assertibles put together by the conjunctive connective *both … and … and …* . They have two or more conjuncts, all on a par. Disjunctions are nonsimple assertibles put together by the disjunctive connective *either … or … or …*. They have two or more disjuncts, all on a par. Conditionals are nonsimple assertibles formed with the connective *if …, …*; they consist of antecedent and consequent (Diogenes Laertius 7.71–2). What type of assertible an assertible is is determined by the connective particle that controls it, that is,that is, that has the largest scope. *Both not p and q* is a conjunction; *Not both p and q* a negation. Stoic language regimentation asks that sentences expressing assertibles always start with the logical particle or expression characteristic for the assertible. Thus the Stoics introduced an implicit bracketing device similar to that used in Jan Łukasiewicz's (1878–1956) Polish notation.

Stoic negations and conjunctions are truth-functional. Stoic (or at least Chrysippean) conditionals are true when the contradictory of the consequent is incompatible with its antecedent (Diogenes Laertius 7.73). Two assertibles are contradictories of each other if one is the negation of the other (Diogenes Laertius 7.73) or when one exceeds the other by a pre-fixed negation particle (SE M 8.89). The truth-functional Philonian conditional was expressed as a negation of a conjunction: that is,that is, not as *if p, q* but as *not both p and not q*. Stoic disjunction is exclusive and non-truth-functional. It is true when necessarily precisely one of its disjuncts is true. Later Stoics introduced a non-truth-functional inclusive disjunction (Gellius.N.A. 16.8.13–14).

Like Philo and Diodorus, Chrysippus distinguished four modalities and considered them as modal values of propositions rather than modal operators; they satisfy the same standard requirements of modal logic. Chrysippus's definitions are: An assertible is possible when it is both capable of being true and not hindered by external things from being true. An assertible is impossible when it is either not capable of being true <or is capable of being true, but hindered by external things from being true>. An assertible is necessary when, being true, it either is not capable of being false or is capable of being false but hindered by external things from being false. An assertible is nonnecessary when it is both capable of being false and not hindered by external things <from being false> (Diogenes Laertius 7.75). Chrysippus's modal notions differ from Diodorus's in that they allow for future contingents and from Philo's in that they go beyond mere conceptual possibility.

ARGUMENTS. Arguments are—normally—compounds of assertibles. They are defined as a system of at least two premisses and a conclusion (Diogenes Laertius 7.45). Syntactically, every premise but the first is introduced by *now* or *but*, and the conclusion by *therefore*. An argument is valid if the (Chrysippean) conditional formed with the conjunction of its premises as antecedent and its conclusion as consequent is correct (S.E.P.H. 2.137, DL 7.77). An argument is *sound* (literally: *true*) when in addition to being valid, it has true premises. The Stoics defined so-called argument modes as a sort of schema of an argu-

ment (Diogenes Laertius 7.76). A mode of an argument differs from the argument itself by having ordinal numbers taking the place of propositions. A mode of the argument:

> If it is day, it is light.
> But it is not the case that it is light.
> Therefore it is not the case that it is day.

is

> If the 1st, the 2nd.
> But not: the 2nd.
> Therefore not: the 1st.

The modes functioned first as abbreviations of arguments that brought out their logically relevant form and second, it seems, as representatives of the form of a class of arguments.

STOIC SYLLOGISTIC. Stoic syllogistic is an argumental deductive system consisting of five types of indemonstrables or axiomatic arguments and four inference rules, called *themata*. An argument is a syllogism precisely if it either is an indemonstrable or can be reduced to one by means of the *themata* (Diogenes Laertius 7.78). Syllogisms are thus certain types of formally valid arguments. The Stoics explicitly acknowledged that there are valid arguments that are not syllogisms but assumed that they could be somehow transformed into syllogisms.

All basic indemonstrables consist of a nonsimple assertible as leading premiss and a simple assertible as coassumption and have another simple assertible as conclusion. They were defined by five standardized metalinguistic descriptions of the forms of the arguments (S.E. M. 8.224–5; D.L.7.80–81):

> (1) A first indemonstrable is an argument composed of a conditional and its antecedent as premises, having the consequent of the conditional as conclusion.

> (2) A second indemonstrable is an argument composed of a conditional and the contradictory of its consequent as premises, having the contradictory of its antecedent as conclusion.

> (3) A third indemonstrable is an argument composed of a negated conjunction and one of its conjuncts as premises, having the contradictory of the other conjunct as conclusion.

> (4) A fourth indemonstrable is an argument composed of a disjunctive assertible and one of its disjuncts as premises, having the contradictory of the remaining disjunct as conclusion.

> (5) A fifth indemonstrable, finally, is an argument composed of a disjunctive assertible and the contradictory of one of its disjuncts as premises, having the remaining disjunct as conclusion.

Whether an argument is an indemonstrable can be tested by comparing it with these metalinguistic descriptions. For instance:

> If it is day, it is not the case that it is night.
> But it is night.
> Therefore it is not the case that it is day.

comes out as a second indemonstrable, and

> If five is a number, then either five is odd or five is even.
> But five is a number.
> Therefore either five is odd or five is even.

as a first indemonstrable. For testing, a suitable mode of an argument can also be used as a stand-in. A mode is syllogistic, if a corresponding argument with the same form is a syllogism (because of that form). However, there are no five modes that can be used as inference schemata that represent the five types of indemonstrables. For example, the following are two of the many modes of fourth indemonstrables:

Either the 1st or the 2nd.	Either the 1st or not the 2nd.
But the 2nd.	But the 1st.
Therefore not the 1st.	Therefore the 2nd.

Although both are covered by the metalinguistic description, neither can be singled out as *the* mode of the fourth indemonstrables: If we disregard complex arguments, there are thirty-two modes corresponding to the five metalinguistic descriptions; the latter thus prove noticeably more economical.

Of the four *themata* only the first and third are extant. They, too, were metalinguistically formulated. The first *thema*, in its basic form, was:

> When from two <assertibles> a third follows, then from either of them together with the contradictory of the conclusion the contradictory of the other follows (Apul.*Int*. 209.9–14).

This is an inference rule of the kind today called antilogism. The third *thema*, in one formulation, was:

> When from two <assertibles> a third follows, and from the one that follows <that is, the third> together with another, external assumption, another follows, then this other follows

from the first two and the externally coassumed one (Simp.Cael. 237.2–4).

This is an inference rule of the kind today called cut rule. It is used to reduce chain syllogisms. (The second and fourth *themata* are also cut rules, and reconstructions of them can be provided since we know what arguments they. together with the third *thema*, were thought to be able to reduce.) A reduction shows the formal validity of an argument by applying to it the *themata* in one or more steps in such a way that all resultant arguments are indemonstrables. This can be done either with the arguments or their modes (S.E. M. 8.230–8). For instance, the argument mode:

> If the 1^{st} and the 2^{nd}, the 3^{rd}.
> But not the 3^{rd}.
> Moreover, the 1^{st}.
> Therefore not: the 2^{nd}.

can be reduced by the third *thema* to (the modes of) a second and a third indemonstrable as follows:

> When from two assertibles ("If the 1^{st} and the 2^{nd}, the 3^{rd}." and "But not the 3^{rd}.") a third follows ("Not: both the 1^{st} and the 2^{nd}."—by a second indemonstrable) and from the third and an external one ("The 1^{st}.") another follows ("Not: the 2^{nd}."—by a third indemonstrable), then this other ("Not: the 2^{nd}.") also follows from the two assertibles and the external one.

The second *thema* reduced, among others, arguments with the following modes (Alex.Aphr.An.Pr. 164.27–31):

Either the 1^{st} or not the 1^{st}.	If the 1st, if the 1^{st}, the 2^{nd}.
But the 1^{st}.	But the 1^{st}.
Therefore the 1^{st}.	Therefore the 2^{nd}.

The Peripatetics chided the Stoics for allowing such useless arguments, but the Stoics rightly insisted that if they can be reduced, they are valid. The four *themata* can be used repeatedly and in any combination in a reduction. Thus propositional arguments of indeterminate length and complexity can be reduced. Stoic syllogistic has been formalized, and it has been shown that the Stoic deductive system shows strong similarities with relevance logical systems such as those by Storrs McCall. Like Aristotle, the Stoics aimed at proving nonevident, formally valid *arguments* by reducing them by means of accepted inference rules to evidently valid *arguments*. Thus, although their logic is a propositional logic, they did not intend to provide a system that allows for the deduction of all propositional-logical truths but, rather, a system of valid propositional-logical arguments with at least two premises and a conclusion. Nonetheless, it is evidenced that the Stoics independently recognized many simple logical truths, including excluded middle, double negation, and contraposition.

LOGICAL PARADOXES. The Stoics recognized the importance of both the Liar and the Sorites paradoxes (Cic.*Acad*. 2.95–8, Plut.*Comm.Not*. 1059D–E, Chrys.*Log. Zet*.col.IX, S.E.*M*.1.68&7.244-246&7.416.). Chrysippus may have tried to solve the Liar as follows: There is an uneliminable ambiguity in the Liar sentence ("I am speaking falsely," uttered in isolation)between the assertibles (1) 'I falsely say I speak FALSELY' and (2) 'I *am* speaking falsely' (that is, I am doing what I'm saying), of which at any time the Liar sentence is said precisely one is true, but it is arbitrary which one: (2) entails (3) 'I *am* speaking truly' and is incompatible with (2) and (4) I truly say I speak falsely' (2) entails (4) and is incompatible with (1) and (3). Thus bivalence is preserved. Chrysippus's stand on the Sorites seems to have been that vague borderline sentences uttered in the context of a Sorites series have no assertibles corresponding them, and that it is obscure to us where the borderline cases start, so that it is rational for us to stop answering while still on safe ground. The latter remark suggests Chrysippus was aware of the problem of higher-order vagueness. Again, bivalence of assertibles is preserved.

LATER ANTIQUITY

Very little is known about the development of logic from c. 100 BCE to c. 250 CE. It is unclear when Peripatetics and the Stoics began taking notice of the logical achievements of each other. Sometime during that period, the terminological distinction between *categorical syllogisms*, used for Aristotelian syllogisms, and *hypothetical syllogisms*, used not only for those by Theophrastus and Eudemus but also for the Stoic propositional-logical syllogisms, gained a foothold. In the first century BCE, the Peripatetics Ariston of Alexandria and Boethus of Sidon wrote about syllogistic. Ariston is said to have introduced the so-called *subaltern* syllogisms (Barbari, Celaront, Cesaro, Camestrop and Camenop) into Aristotelian syllogistic (Apul.*Int*. 213.5–10), that is, the syllogisms one gains by applying the subalternation rules (that were acknowledged by Aristotle in his *Topics*):

> From "A holds of every B" infer "A holds of some B"

From "A holds of no B" infer "A does not hold of some B"

to the conclusions of the relevant syllogisms. Boethus suggested substantial modifications to Aristotle's theories: He claimed that all categorical syllogisms are complete and that hypothetical syllogistic is prior to categorical (Gal.Inst.Log. 7.2), although we are not told prior in which way. The Stoic Posidonius (c.135–c.51 BCE) defended the possibility of logical or mathematical deduction against the Epicureans and discussed some syllogisms he called *conclusive by the force of an axiom*, which apparently included arguments of the type "As the 1st is to the 2nd, so the 3rd is to the 4th; the ratio of the 1st to the 2nd is double; therefore the ratio of the 3rd to the 4th is double," which was considered conclusive by the force of the axiom "things which are in general of the same ratio, are also of the same particular ratio" (Gal. *Inst. Log*.18.8). At least two Stoics in this period wrote a work on Aristotle's *Categories*. From his writings we know that Cicero was knowledgeable about both Peripatetic and Stoic logic; and Epictetus's discourses prove that he was acquainted with some of the more taxing parts of Chrysippus's logic. In all likelihood there existed at least a few creative logicians in this period, but we do not know who they were and what they created.

The next logician of rank, if of lower rank, of whom we have sufficient evidence is Galen (129–199 or 216 CE), whose greater fame was as a physician. He studied logic with both Peripatetic and Stoic teachers and recommended to avail oneself of parts of either doctrine, as long as it could be used for scientific demonstration. He composed commentaries on logical works by Aristotle, Theophrastus, Eudemus, and Chrysippus, as well as treatises on various logical problems and a major work titled *On Demonstration*. All these are lost except for some information in later texts, but his *Introduction to Logic* has come down to us almost in full. In *On Demonstration*, Galen developed, among other things, a theory of compound categorical syllogisms with four terms, which fall into four figures, but we do not know the details. He also introduced the so-called relational syllogisms, examples of which are "A is equal to B, B is equal to C; therefore A is equal to C" and "Dio owns half as much as Theo; Theo owns half as much as Philo. Therefore Dio owns a quarter of what Philo owns." (Gal. *Inst. Log*. 17–18). All relational syllogisms Galen mentions have in common that they are not reducible in either Aristotle's or Stoic syllogistic, but it is difficult to find further formal characteristics that unite them all. In general, in his *Introduction to Logic*, he merges Aristotelian Syllogistic with a strongly Peripatetic reinterpretation of Stoic propositional logic.

The second ancient introduction to logic that has survived is Apuleius's (second century CE) *De Interpretatione*. This Latin text, too, displays knowledge of Stoic and Peripatetic logic; it contains the first full presentation of the square of opposition, which illustrates the logical relations between categorical sentences by diagram. Alcinous, in his *Handbook of Platonism* 5, is witness to the emergence of a specifically Platonist logic, constructed on the Platonic notions and procedures of division, definition, analysis, and hypothesis, but there is little that would make a logicians heart beat faster. Sometime between the third and sixth century CE, Stoic logic faded into oblivion to be resurrected only in the twentieth century in the wake of the (re)discovery of propositional logic.

The surviving, often voluminous, Greek commentaries on Aristotle's logical works by Alexander of Aphrodisias (fl. c.200 CE), Porphyry (234–c.305), Ammonius Hermeiou (fifth century), John Philoponus (c. 500), and Simplicius (sixth century), and the Latin ones by Anicius Manlius Severinus Boethius (c.480–524) have their main importance as sources for lost Peripatetic and Stoic works. Still, two of the commentators deserve special mention: Porphyry, for writing the *Isagoge* or *Introduction* (that is, to Aristotle's *Categories*), in which he discusses the five notions of genus, species, differentia, property, and accident as basic notions one needs to know to understand the *Categories*. For centuries, the *Isagogē* was the first logic text a student would tackle, and Porphyry's five predicables (which differ from Aristotle's four) formed the basis for the medieval doctrine of the *quinque voces*.

The second is Boethius. In addition to commentaries, he wrote a number of logical treatises, mostly simple explications of Aristotelian logic, but also two very interesting ones: (1) His *On Topical Differentiae* bears witness of the elaborated system of topical arguments that logicians of later antiquity had developed from Aristotle's *Topics* under the influence of the needs of Roman lawyers. (2) His *On Hypothetical Syllogisms* systematically presents wholly hypothetical and mixed hypothetical syllogisms as they are known from the early Peripatetics; it may be derived from Porphyry. Boethius's insistence that the negation of "If it is A, it is B" is "If it is A, it is not B" suggests a suppositional understanding of the conditional, a view for which there is also some evidence in Ammonius, but that is not attested for earlier logicians. Historically, Boethius is most important because he translated all of Aristotle's *Organon* into Latin, and thus

these texts (except the *Posterior Analytics*) became available to philosophers of the medieval period.

See also Alcinous; Alexander of Aphrodisias; Antisthenes; Aristotle; Boethius, Anicius Manlius Severinus; Chrysippus; Cicero, Marcus Tullius; Cleanthes; Diodorus Cronus; Diogenes Laertius; Epictetus; Epicureanism and the Epicurean School; Frege, Gottlob; Galen; Gorgias of Leontini; Peripatetics; Philo of Megara; Philoponus, John; Plato; Porphyry; Posidonius; Protagoras of Abdera; Sextus Empiricus; Simplicius; Socrates; Stoicism; Theophrastus; Zeno of Citium; Zeno of Elea.

Bibliography

TEXTS AND TRANSLATIONS

Ackrill, J. L., trans. *Aristotle's Categories and De Interpretatione*. Oxford: Oxford University Press, 1961.

Alcinous. *Enseignement des doctrines de Platon*. Edited and translated into French by John Whittaker. Paris: Les Belles Lettres, 1990.

Alcinous. *The Handbook of Platonism*. Translated by John Dillon. Oxford: Oxford University Press, 1993.

Apuleius. *Peri Hermeneias*. Edited by Claudio Moreschini. Stuttgart/Leipzig: Teubner, 1991.

Barnes, Jonathan, ed. *The Complete Works of Aristotle*, 2 vols. Princeton: Princeton University Press, 1984.

Barnes, Jonathan, trans. *Porphyry's Introduction*. Oxford: Oxford University Press, 2003.

Boethius. *De hypotheticis syllogismis*. Edited and translated into Italian by L. Obertello. Brescia, Italy: Paideia, 1969.

Boethius. *In librum Aristotelis De interpretatione: secunda editio*. Edited by C. Meiser. Leipzig: Teubner, 1880.

Burnet, J., ed. *Platonis Opera*, vols. I–V. Oxford: Oxford University Press, 1901–1991.

Cooper, John M., and D. S. Hutchinson, eds. *Plato Complete Works*. Indianapolis, IN: Hackett, 1997.

Diels, Hans, and Walther Kranz, eds. (DK) *Die Fragmente der Vorsokratiker*. Berlin: Weidmann, 1951.

Diogenes Laertius. *Lives of the Philosophers*, 2 vols. Edited by M. Marcovich. Stuttgart & Leipzig: Teubner, 1999.

Döring, Klaus, ed. *Die Megariker. Kommentierte Sammlung der Testimonien*. Amsterdam: Grüner. 1972.

Dorion, Louis-Andre, trans. *Aristote: Les refutations sophistique*. Paris: J. Vrin, 1995.

Galen. *Institutio Logica*. Edited by C. Kalbfleisch. Leipzig: Teubner, 1896.

Giannantoni, G. *Socratis et Socraticorum Reliquiae*, 4 vols. *Elenchos* 18, Naples, Italy: Bibliopolis, 1990.

Huby, P. M., and Dimitri Gutas, eds. and trans. *Theophrastus: Logic*. In *Theophrastus of Eresus: Sourses for his Life, Writings, Thought and Influence*, edited by William W. Fortenbaugh, 114–275. Leiden, U.K.: Brill, 1992.

Hülser, Karlheinz, ed. *Die Fragmente zur Dialektik der Stoiker*, 4 vols. Stuttgart-Bad Cannstadt: Frommann-Holzboog, 1987–1988.

Kieffer, John Spangler, trans. *Galen's Institutio Logica*. Baltimore, MD: Johns Hopkins University Press, 1964.

Robinson, Thomas M. ed. *Contrasting Arguments: An Edition of the* Dissoi Logoi. London: Arno Press, 1979.

Ross, W. D., et al., eds. *Aristotelis Opera*. Oxford: Oxford University Press, 1890–1991.

Sextus Empiricus. *Works*. Edited by H. Mutschmann and J. Mau. Leipzig: Teubner, 1914–1961.

Smith, Robin, trans. *Aristotle, Topics I, VIII, and Selections*. Oxford: Oxford University Press, 1997.

Smith, Robin, trans. *Aristotle's Prior Analytics*. Indianapolis, IN: Hackett, 1989.

Stump, Eleanor, trans. *Boethius's 'De topicis differentiis'*. Ithaca, NY: Cornell University Press, 1978.

Weidemann, Hermann, trans. *Aristoteles, De Interpretatione*. Berlin: Akademie Verlag, 1994.

Works of Alexander of Aphrodisias, Porphyry, Ammonius, Simplicius, and Philoponus are found in: *Commentaria in Aristotelem Graeca*. Edited by H. Diels. Berlin: Reimer 1882–1909.

SECONDARY LITERATURE: GENERAL

"Ancient Logic, Overview." *Stanford Encyclopedia of Philosophy*. Available from http://www.plato.stanford.edu/.

Barnes, Jonathan. *Truth, etc.*. Oxford: Oxford University Press, 2005.

Kneale, Martha, and William Kneale. *The Development of Logic*. Oxford: Oxford University Press, 1962.

SECONDARY LITERATURE: THE BEGINNINGS

Frede, Michael. "Plato's *Sophist* on False Statements." In *The Cambridge Companion to Plato*. Edited by Kraut Richard. Cambridge, U.K.: Cambridge University Press, 1992.

Kapp, Ernst. *Greek Foundations of Traditional Logic*. New York: Columbia University Press, 1942.

Mueller, Ian. "Greek Mathematics and Greek Logic." In *Ancient Logic and its Modern Interpretation. Proceedings of the Buffalo Symposium on Modernist Interpretations of Ancient Logic, 21 and 22 April, 1972*. Edited by John Corcoran. Dordrecht, Holland: Reidel, 1974.

Robinson, Richard. *Plato's Earlier Dialectic*, 2nd ed. Ithaca, NY: Cornell University Press, 1953.

Salmon, W. C. *Zeno's Paradoxes*. Indianapolis, IN: Hackett, 2001.

SECONDARY LITERATURE: ARISTOTLE

Corcoran, John. "Aristotle's Natural Deduction System." In *Ancient Logic and its Modern Interpretation. Proceedings of the Buffalo Symposium on Modernist Interpretations of Ancient Logic, 21 and 22 April, 1972*. Edited by John Corcoran. Dordrecht, Holland: Reidel. 1974.

Frede, Dorothea. "The Sea-Battle Reconsidered." *Oxford Studies in Ancient Philosophy* 3 (1985): 31–87.

Frede, Michael. "Categories in Aristotle." In *Essays in Ancient Philosophy*, 29–48. Minneapolis: University of Minnesota Press, 1987.

Lear, Jonathan. *Aristotle and Logical Theory*. Cambridge, U.K.: Cambridge University Press, 1980.

Patterson, Richard. *Aristotle's Modal Logic: Essence and Entailment in the Organon*. Cambridge, U.K., Cambridge University Press, 1995.

Patzig, Günther. *Aristotle's Theory of the Syllogism*. Translated by Jonathan Barnes. Dordrecht, Holland: Reidel, 1969.
Primavesi, Oliver. *Die aristotelische Topik*. Munich, Germany: Beck, 1996.
Smiley, Timothy. "What Is a Syllogism?" *Journal of Philosophical Logic* 1 (1974): 136–154.
Smith, Robin. "Logic." In *The Cambridge Companion to Aristotle*. Edited by Jonathan Barnes. Cambridge, U.K.: Cambridge University Press, 1994.
Striker, Gisela. "Aristoteles über Syllogismen 'Aufgrund einer Hypothese.'" *Hermes* 107 (1979): 33.
Striker, Gisela. "Modal vs. Assertoric Syllogistic." *Ancient Philosophy* Spec. issue (1994): 39–51.
Whitaker, C. W. A. *Aristotle's De Interpretatione: Contradiction and Dialectic*. Oxford: Oxford University Press, 1996.

SECONDARY LITERATURE: THEOPHRASTUS AND EUDEMUS

Barnes, Jonathan. "Theophrastus and Hypothetical Syllogistic." In *Aristoteles: Werk und Wirkung* I. Edited by J. Wiesner. Berlin, 1985.
Bobzien, Susanne. "Pre-Stoic Hypothetical Syllogistic in Galen." *Bulletin of the Institute of Classical Studies* Suppl (2002):57–72.
Bobzien, Susanne. "Wholly Hypothetical Syllogisms." *Phronesis* 45 (2000): 87–137.
Bochenski, I. M. *La Logique de Théophraste*. Fribourg, Switzerland: Publications de l'Universite de Fribourg en Suisse, 1947.
Lejewski, Czesław. "On Prosleptic Syllogisms." *Notre Dame Journal of Formal Logic* 2 (1961): 158–176.

SECONDARY LITERATURE: DIODORUS CRONUS AND PHILO THE LOGICIAN

Bobzien, Susanne. "Chrysippus' Modal Logic and its Relation to Philo and Diodorus." In *Dialektiker und Stoiker*. Edited by K. Döring and Th. Ebert. Stuttgart: Franz Steiner, 1993.
"Dialectical School." Stanford Encyclopedia of Philosophy. Available from http://www.plato.stanford.edu/.
Prior, A. N. *Past, Present, and Future*, chaps. II.1–2, III.1. Oxford: Clarendon Press, 1967.
Sedley, David. "Diodorus Cronus and Hellenistic Philosophy." *Proceedings of the Cambridge Philological Society* 203 (1977): 74–120.

SECONDARY LITERATURE: THE STOICS

Atherton, Catherine. *The Stoics on Ambiguity*. Cambridge Classical Studies. Cambridge, U.K.: Cambridge University Press, 1993.
Bobzien, Susanne. "Chrysippus and the Epistemic Theory of Vagueness." *Proceedings of the Aristotelian Society* 102 (2002): 217–238.
Bobzien, Susanne. "Stoic Logic." In *The Cambridge History of Hellenistic Philosophy*. Edited by J. Barnes, J. Mansfeld, and M. Schofield. Cambridge, U.K.: Cambridge University Press, 1999.
Bobzien, Susanne. "Stoic Syllogistic." *Oxford Studies in Ancient Philosophy* 14 (1996): 133–192.
Brunschwig, Jacques. "Remarks on the Classification of Simple Propositions in Hellenistic Logics." In *Papers in Hellenistic Philosophy*. Cambridge, U.K.: Cambridge University Press, 1994.
Cavini, Walter. "Chrysippus on Speaking Truly and the Liar." In *Dialektiker und Stoiker*. Edited by K. Döring and Th. Ebert. Stuttgart: Franz Steiner, 1993.
Crivelli, Paolo "Indefinite Propositions and Anaphora in Stoic Logic." *Phronesis* 39 (1994): 187–206.
Frede, Michael. *Die stoische Logik*. Goettingen: Vandenhoeck & Ruprecht, 1974.
Frede, Michael. "Stoic vs. Peripatetic Syllogistic." *Archiv für Geschichte der Philosophie* 56 (1975): 132–154.
Gaskin, Richard. "The Stoics on Cases, Predicates, and the Unity of the Proposition." In *Aristotle and After*. Edited by R. Sorabji. London: Institute of Classical Studies, 1997.
Lloyd, A. C. "Definite Propositions and the Concept of Reference." In *Les Stoïciens et leur logique*, edited by Jacques Brunschwig. Paris: 1978.
Long, A. A. "Language and Thought in Stoicism." In *Problems in Stoicism*. Edited by A. A. Long. London: Athlone Press, 1971.
Schenkeveld, D. M. "Stoic and Peripatetic Kinds of Speech Act and the Distinction of Grammatical Moods." *Mnemosyne* 37 (1984): 291–351.

SECONDARY LITERATURE: LATER ANTIQUITY

Barnes, Jonathan. "A Third Sort of Syllogism: Galen and the Logic of Relations." In *Modern Thinkers and Ancient Thinkers*. Edited by R. W. Sharples. Boulder, CO: Westview Press, 1993.
Barnes, Jonathan. *Logic and the Imperial Stoa*. Leiden: Brill, 1997.
Bobzien, Susanne. "Hypothetical Syllogistic in Galen—Propositional Logic Off the Rails?" *Rhizai Journal for Ancient Philosophy and Science* 2 (2004): 57–102.
Bobzien, Susanne. "The Development of *Modus Ponens* in Antiquity: From Aristotle to the 2nd Century AD." *Phronesis* 47 (2002): 359–394.
Ebbesen, Sten. "Boethius as an Aristotelian Commentator." In *Aristotle Transformed—The Ancient Commentators and their Influence*. Edited by R. Sorabji. London: Duckworth, 1990.
Ebbesen, Sten. "Porphyry's Legacy to Logic." In *Aristotle Transformed—The Ancient Commentators and Their Influence*. Edited by R. Sorabji. London: Duckworth, 1990.
Sullivan, W. M. *Apuleian Logic. The Nature, Sources and Influences of Apuleius' Peri Hermeneias*. Amsterdam: North Holland, 1967.

Susanne Bobzien (2005)

LOGIC AND INFERENCE IN INDIAN PHILOSOPHY

By the fifth century BCE great social change was taking place in India and a period of intense intellectual activity came into being. Rational inquiry into a wide range of topics was under way, including agriculture, architecture, astronomy, grammar, law, logic, mathematics, medicine, phonology, and statecraft. Aside from the world's earliest extant grammar, Pāṇini's (c. 400 BCE) *Aṣṭādhyāyī*, however, no works devoted to these topics actually date from

this period. Nonetheless, scholars agree that incipient versions of first extant texts on these topics were being formulated.

One text dating from this period and important to tracing the development of logic in classical India is a Buddhist work, Moggaliputta Tissa's *Kathā-vatthu* (Points of controversy; third century BCE), which exhibits awareness of the fact that the form of argument is crucial to its being good. The text gives the refutation of some 200 propositions over which the Sthavīravāda, one of the Buddhist schools, disagreed with various Buddhist schools. The treatment of each point comprises a debate between a proponent and opponent. Throughout book 1, chapter 1, one finds refutations of precisely the following form:

Proponent:	Is A B?
Opponent:	Yes.
Proponent:	Is C D?
Opponent:	No.
Proponent:	Acknowledge defeat, since if A is B, then C is D.

The author clearly presumes it to be self-evident, first, that it is wrong to hold inconsistent propositions and, second, that the propositions assented to—corresponding to the propositional schemata of α, $\neg B$, $\alpha \rightarrow B$—are indeed inconsistent.

The first 500 years of the Common Era saw the redaction of treatises devoted to the systematic exposition of the technical subjects mentioned earlier, as well as of philosophical treatises in which proponents of diverse religious traditions put forth systematic versions of their worldview. These latter works bear witness, in a number of different ways, to the intense interest of the period in argumentation. To begin with, the authors of many of these texts submit arguments and, in doing so, explicitly appeal to such well-known logical principles as those of noncontradiction, of excluded middle and of double negation, though they adduce them, not as principles of logic, but as self-evident ontic facts. Thus, the Buddhist philosopher Nāgārjuna (c. 150–250) often invokes an ontic principle of noncontradiction, saying such things as "when something is a single thing, it cannot be both existent and non-existent" (*Mūlamadhyamakakārikā* chapter 7, verse 30), which is clearly reminiscent of Aristotle's own ontic formulation of the principle of noncontradiction, namely, "that a thing cannot at the same time be and not be" (*Metaphysics* book 3, chapter 2996b29–30).

Next, many of the arguments formulated correspond to such well-recognized rules of inference as *modus ponens* (i.e., from α and $\alpha \rightarrow B$, one infers B), *modus tollens* (i.e., from $\neg B$ and $\alpha \rightarrow B$, one infers $\neg\alpha$), disjunctive syllogism (i.e., from $\neg\alpha$ and $\alpha \vee B$, one infers B), constructive dilemma (i.e., from $\alpha \vee B$, $\alpha \rightarrow \gamma$ and $B \rightarrow \gamma$, one infers γ), categorical syllogism (i.e., from $\alpha \rightarrow B$ and $B \rightarrow \gamma$, one infers $\alpha \rightarrow \gamma$), and *reductio ad absurdum* (i.e., if something false follows from an assumption, then the assumption is false). This last form of argument, termed *prasaṅga* in Sanskrit, is extremely common. Indeed, so common are such arguments in Nāgārjuna's works that his follower, Buddhapālita (470–540), took all of Nāgārjuna's arguments to be *prasaṅga* arguments. As a result, Buddhapālita and his followers were and are referred to as *prāsaṅgikas* (absurdists).

Finally, many of the texts are either devoted to, or have passages devoted to, the enumeration, definition, and classification of public discussion, or debate (*vāda*). The same texts or passages also identify the parts of argument, the flaws found in poor arguments, including such fallacies as circularity (*anyonya-āśraya*, reciprocal dependence) and infinite regress (*an-avasthā*, ungroundedness), as well as quibbles (*chala*) and sophistical refutations (*jāti*) (see Solomon 1976, vol. 1, chapter 5). They also set down ways in which a discussant's behavior warrant his or her being judged the loser of the debate (*nigraha-sthāna*) (see Solomon 1976, vol. 1, chapter 6).

One of the earliest examples of an argument in a form that clearly adumbrates the canonical form the classical Indian inference eventually takes is found in a passage in the *Caraka-saṃhitā* (CS book 2, chapter 8, section 31), a medical text, which defines an argument to have five parts: the proposition (*pratijñā*), the ground or reason (*hetu*), the corroboration (*dṛṣṭānta*), the application (*upanaya*), and the conclusion (*nigamana*). The following is an example:

Proposition:	The soul is noneternal
Ground:	because it is detectable by the senses.
Corroboration:	It is like a pot.
Application:	As a pot is detectable by the senses, and is noneternal, so is the soul detectable by the senses.
Conclusion:	Therefore, the soul is noneternal.

This form of the argument clearly reflects the debate situation. First, one propounds a proposition, that is, one sets forth a proposition to be proved. One then states the ground, or reason, for the proposition one is propound-

ing. Next, one corroborates with an example the connection implicit between the property mentioned in the proposition and the property adduced as its ground. The immediately ensuing step, the application, spells out the analogy between the example and the subject of the proposition. Notice that this part of the argument retains the vestiges of the analogical reasoning that is no doubt its predecessor. Finally, one asserts the proposition.

As was obvious to the thinkers of this period, not all arguments of this form are good arguments. However, no clear criteria are set forth whereby good arguments or inferences can be distinguished from bad ones. At best, some authors simply list good arguments, as does the Buddhist idealist Asanga (flourish fourth–fifth century CE) in a section at the end of a chapter of his *Yogācārabhūmi-śāstra* (Treatise on the stages of yogic practice). Other works provide lists of both good and bad arguments, the latter often referred to as nongrounds (*a-hetu*) or pseudogrounds (*hetu-ābhāsa*) (see Solomon 1976, vol. 1, chapter 7). It is difficult to be sure what the basis for the classification is in these early texts. In the *Nyāya-sūtra* (Aphorisms on logic), a work attributed to Gautama Akṣapāda (flourished second century CE), the author gives neither a definition nor an example. Even in cases where definitions and examples are given, as in the *Caraka-saṃhitā* mentioned earlier, the modern reader is rarely sure of what is intended.

Other passages from these earliest texts treat inference. In these passages inference is taken to be knowledge of one fact arising from knowledge of another. Often, as in the passages of the *Caraka-saṃhitā* (CS book 1, chapter 11, sections 21–22) and the *Nyāya-sūtra* (NS book 1, chapter 1, aphorism 5), no mention is made of any knowledge of what links the two facts. Moreover, the classification of inference in these two texts seems to be based on characteristics completely extrinsic to the logical features of the inferences adduced, for example, according to whether the property permitting the inference precedes, is simultaneous with, or succeeds the property to be inferred.

In contrast, passages from other texts of this period provide definitions of inference that require, besides knowledge of the two states of affairs, knowledge of the relation linking the two. However, instead of providing a formal relation, they provide a miscellany of material relations. The *Ṣaṣṭi-tantra* (Sixty doctrines), which is attributed by some to Pañcaśikha (flourished second century BCE) and by others to Varṣaṇya (fl. after the second century CE), enumerates seven such relations, while the *Vaiśeṣika-sūtra* (Aphorisms pertaining to individuation; VS book 9, aphora 20), a text attributed to Kaṇāda (flourished first century CE), enumerates five: the relation of cause to effect, of effect to cause, of contact, of exclusion, and of inherence. In each of these texts the miscellany of material relations serves to classify inferences. Thus, although in these two works the parts of an inference have been made explicit, the formal connection among these parts remained implicit.

The works of the Buddhist philosopher Vasubandhu (flourished fourth century CE) seem to be the earliest extant works that provide a formal characterization of the inference. He holds that inference has only three parts: a subject (*pakṣa*) and two properties, the property to be established (*sādhya*) in the subject and another that is the ground (*hetu*). Exploiting an idea ascribed by his coreligionist Asanga in his *Shùn Zhèng Lùn* to an unknown school (thought by at least one scholar to be the Sāṃkhya school), Vasubandhu maintained that a ground in an inference is a proper one if, and only if, it satisfies three conditions—the so-called *tri-rūpa-hetu* (the grounding property *hetu* in its three forms). The first form is that the grounding property (*hetu*; H) should occur in the subject of an inference (*pakṣa*; p). The second is that the grounding property (H) should occur in those things similar to the subject insofar as they have the property to be established (*sādhya*; S). And third, the grounding property (H) should not occur in any of those things dissimilar from the subject insofar as they lack the property to be established (S). These conditions can be viewed as a partial specification of the validity of inferences of this form:

Thesis:	p has S.
Ground:	p has H.
Indispensability:	Whatever has H has S.

The first condition corresponds to the premise labeled *ground* in the schema above, while the second two correspond to the premise labeled *indispensability*. In his *Vāda-vidhi* (Rules of debate) Vasubandhu makes clear that the relation, knowledge of which is necessary for inference, is not just any in a miscellany of material relations, but a formal relation, which he designates, in some places, as *a-vinā-bhāva*—literally, not being without (compare to the Latin expression sine qua non)—and in others, as *nāntarīyakatva*—literally, being unmediated.

The following are two examples of inferences satisfying the previous schema:

Thesis:	p has fire.
Ground:	p has smoke.
Pervasion:	Whatever has smoke has fire.
Thesis:	p is a tree (i.e., has tree-ness).
Ground:	p is an oak (i.e., has oak-ness).
Pervasion:	Whatever is an oak (i.e., has oak-ness) is a tree (i.e., has tree-ness).

The previous schema is the one that Buddhist thinkers insisted on for all sound inference or argument. Brahmanical thinkers came to insist on a the form found in the *Caraka-saṃhitā*, but with the form of the application modified to express a universal claim, thereby giving it the same logical core as the form accepted by the Buddhists.

It is important to note that, no matter how different the metaphysical assumptions of the various philosophical schools, they all used a naive realist's ontology to specify the states of affairs used to study inference. According to this view, the world consists of individual substances or things (*dravya*), universals (*sāmānya*), and relations between them. The fundamental relation is the one of occurrence (*vṛtti*). The relata of this relation are known as substratum (*dharmin*) and superstratum (*dharma*), respectively. The relation has two forms: contact (*saṃyoga*) and inherence (*samavāya*). So, for example, one individual substance, a pot, may occur on another, say the ground, by the relation of contact. In this case the pot is the superstratum and the ground is the substratum. Or, a universal, say brownness, may occur in an individual substance, say a pot, by the relation of inherence. Here, brownness, the superstratum, inheres in the pot, the substratum. The converse of the relation of occurrence is the relation of possession.

Another important relation is the relation that one superstratum bears to another. This relation, known as pervasion (*vyāpti*), can be defined in terms of the occurrence relation. One superstratum pervades another just in case where ever the second occurs the first occurs. The converse of the pervasion relation is the concomitance relation.

As a result of these relations, the world embodies a structure: If one superstratum, designated as H, is concomitant with another superstratum, designated as S, and if a particular substratum, say p, possesses the former superstratum, then it possesses the second. This structure is captured by both the inferential schema for Buddhist thinkers and the inferential schema for Brahmanical thinkers.

Dignāga (flourished fifth century CE), another Buddhist philosopher, consolidated and systematized the insights into the formal basis of inference found in Vasubandhu's works. First, distinguishing between inference for oneself and inference for another, he made explicit what had previously been only implicit, namely, that inference, the cognitive process whereby one increases one's knowledge, and argument, the device of persuasion, are but two sides of a single coin. Second, he undertook to make the three forms of the grounding property (*tri-rūpa-hetu*) more precise, pressing into service the Sanskrit particle *eva* (only). And third, and perhaps most strikingly, he created the *hetu-cakra* (wheel of reasons), a three-by-three matrix, set up to classify pseudogrounds in light of the last two forms of the three forms of a proper ground. On the one hand, there are the three cases of the grounding property (H) occurring in some, none, or all of substrata where the property to be established (S) occurs. On the other hand, there are the three cases of the grounding property (H) occurring in some, none, or all of substrata where the property to be established (S) does not occur. Letting S be the substrata in which S occurs and $\overline{S}$ be the substrata in which S does not occur, one arrives at the following table:

H occurs in:	all S all $\overline{S}$	all S no $\overline{S}$	all S some $\overline{S}$
H occurs in:	no S all $\overline{S}$	no S no $\overline{S}$	no S some $\overline{S}$
H occurs in:	some S all $\overline{S}$	some S no $\overline{S}$	some S some $\overline{S}$

Dignāga's works set the framework within which subsequent Buddhist thinkers addressed philosophical issues pertaining to inference and debate. Thus, Śankarasvāmin (flourished sixth century CE) wrote a brief manual of inference for Buddhists, called the *Nyāya-praveśa* (Beginning logic), based directly on Dignāga's work. Not long thereafter, Dharmakirti (flourished seventh century CE), the great Buddhist metaphysician, also elaborated his views on inference and debate within the framework found in Dignāga.

Dharmakirti made at least two contributions to the treatment of inference. Recall that one of the developments found in Vasubandhu's work was the identification of the formal contribution of what corresponds with the premise labeled indispensability in the inferential schema above making explicit that the corresponding relation is a formal one. One of Vasubandhu's terms for it, namely, *a-vinābhāva* (not being without), made it clear that infer-

ence involves some form of necessity. The question raised by Dharmakirti is: What is the basis for the necessity? Recognizing that the necessity does not arise from a simple enumeration of cases, Dharmakirti postulated two relations to vouchsafe the necessity of inference: causation (*tadutpatti*) and identity (*tādātmya*). A second contribution was his attempt to bring knowledge of absences, or roughly negative facts, within the purview of inference.

Another important Buddhist thinker who treated inference was Dharmottara (flourished eighth century CE), who wrote a useful commentary on Dharmakirti's widely read *Nyāya-bindu*.

Dignāga not only had a profound influence on his Buddhist followers but he also influenced his non-Buddhist contemporaries and their followers. It would be wrong, however, to conclude that every adoption of ideas similar to those used by Dignāga in his works should be attributed to him. After all, one cannot be certain that Dignāga's contemporaries did not arrive at similar ideas independently or that they might not have got their ideas from sources common to them and Dignāga. In any event, Praśastapāda (flourished sixth century CE), an adherent of the Vaiśeṣika school and a near contemporary of Dignāga, also defined inference in a way that not only made clear its formal nature but also used the quantificational adjective *sarva* (all) to make the formal connection precise.

At the same time, some authors of this period seem to have retained a view of inference akin to the one found in the *Ṣaṣṭi-tantra* and the *Vaiśeṣika-sūtra*, in which the formal role of what corresponds with the inferential schema's pervasion (*vyāpti*) had yet to have been identified. This is true both of Vātsyāyana (flourished fifth century CE), the author of the earliest extant commentary on the *Nyāya-sūtra* and of Śabara (flourished sixth century CE), the author of the earliest extant commentary on Jaimini's *Mīmāṃsā-sūtra*. However, it was not long before the advocates of both *Nyāya* and *Mīmāṃsā* adapted to the formal view of inference. On the one hand, one finds that the *Mīmāṃsā* thinker Kumārila Bhaṭṭa (flourished 730 CE), adopted, without special comment, the formal perspective. On the other hand, one also finds that, though the *Nyāya* thinker Uddyotakara (flourished late sixth century CE) argued vigorously against many of Dignāga's views, he nonetheless advocated a view that presupposed the formalization found in Dignāga's works. Thus, Uddyotakara classified grounds (*hetu*) as: concomitant (*anvaya*), where nothing distinct from particular substratum p (in the inferential schema) fails to have the property S; exclusive (*vyatireka*), where nothing distinct from p (in the inferential schema) has the property S; and both concomitant and exclusive, where some things distinct from p have the property S and some fail to have the property S. This classification becomes the standard classification for the adherents of *Nyāya* during the scholastic period.

See also Knowledge in Indian Philosophy; Mind and Mental States in Buddhist Philosophy; Negation in Indian Philosophy; Truth and Falsity in Indian Philosophy; Universal Properties in Indian Philosophical Traditions.

Bibliography

Annambhaṭṭa. *Tarkasaṁgraha-dīpikā on Tarkasaṁgraha by Annaṁbhaṭṭa*. Translated by Gopinath Bhattacharya. Calcutta, India: Progressive, 1976. Originally published under the title *Tarka-saṁgraha*.

Bochenski, I. M. *Formale Logik*. Freiburg, Germany: Karl Alber, 1956. Translated and edited by Ivo Thomas as *A History of Formal Logic*. 2nd ed. New York: Chelsea House, 1970.

Ganeri, Jonardon, ed. *Indian Logic: A Reader*. Richmond, U.K.: Curzon, 2001.

Gillon, Brendan S. "Dharmakiirti and His Theory of Inference." In *Buddhist Logic and Epistemology: Studies in the Buddhist Analysis of Inference and Language*, edited by Bimal K. Matilal and Robert D. Evans, 77–87. Dordrecht, Netherlands: D. Reidel, 1986.

Gillon, Brendan S., and Martha Lile Love. "Indian Logic Revisited: *Nyaayaprave'sa* Reviewed." *Journal of Indian Philosophy* 8 (1980): 349–384.

Potter, Karl H. "Introduction to the Philosophy of Nyaya-Vaisesika." *Encyclopedia of Indian Philosophies*, Vol. 2. Edited by Karl H. Potter, et al. Princeton, NJ: Princeton University Press, 1977.

Randle, H. N. *Indian Logic in the Early Schools: A Study of the Nyāyadarśana in Its Relation to the Early Logic of Other Schools*. London: H. Milford, 1930.

Solomon, Esther A. *Indian Dialectics: Methods of Philosophical Discussion*. 2 vols. Ahmedabad, India: B. J. Institute of Learning and Research, 1976–1978.

Stcherbatsky, Fedor Ippolitovich. *Buddhist Logic*. 2 vols. New York: Dover, 1962.

Vidyabhusana, Satis Chandra. *A History of Indian Logic: Ancient, Medieval, and Modern*. Calcutta, India: Calcutta University, 1921.

Brendan S. Gillon (2005)

CHINESE LOGIC

Systematic argument in Chinese philosophy began with the Moist school, founded in the fifth century BCE by the first anti-Confucian thinker, Mozi (c. 470–c. 391 BCE). He laid down three tests for the validity of a doctrine: ancient authority, common observation, and practical

effect. At first the controversies of the various schools over moral and political principles led to increasing rigor in argument; then to an interest in dialectic for its own sake, as evidenced in Hui Shih's paradoxes of infinity and in Kung-sun Lung's sophism "A (white) horse is not a horse"; and still later to the antirationalism of the Daoist Zhuangzi (born c. 369 BCE), who rejected all dialectic on the grounds that names have only an arbitrary connection with objects and that any point of view is right for those who accept the choice of names it assumes.

LOGIC OF MOISM

In the third century BCE the Moists responded to Zhuangzi's skepticism by systematizing dialectic in the "Moist Canons" and the slightly later *Ta-ch'ü* and *Hsiao-ch'ü*.

"MOIST CANONS." The "Canons" confined dialectic to questions of the form "Is it this or is it not?" or, since they assumed that the proposition is merely a complex name for a complex object, "Is it or is it not the case that … ?" (The form is distinguished in Chinese by a verbless sentence with a final particle, not by a verb "to be.") In true dialectic the alternatives are paired ("Is it an ox or not?") so that one and only one fits the object. Dialectic excludes such questions as "Is it an ox or a horse?" (it may be neither) and "Is it a puppy or a dog?" (it may be both). Its solutions are absolutely right or wrong; being or not being "this," unlike being long or short, is not a matter of degree, since nothing is more "this" than this is. The Moists further argued that it is self-contradictory to deny or to affirm all propositions: the statement "All statements are mistaken" implies that it is itself mistaken, and one cannot "reject rejection" without refusing to reject one's own rejection.

Names are of three types, distinguished by their relations to "objects," which are assumed to be particular. "Unrestricted" names (such as "thing") apply to every object. Names "of kinds" (such as "horse") apply to every object resembling the one in question. "Private" names (for example, the proper name "Tsang") apply to one object. Whether a name fits an object is decided by appeal to a "standard." There may be more than one standard for an object; for "circle" the standard may be a circle, one's mental picture of a circle, or a compass. Some standards fit without qualification: A circle has no straight lines. Some fit only partially: In deciding whether someone is a "black man" it is not enough to point out his black eyes and hair. The "Canons" began with seventy-five definitions, evidently offered as "standards," of moral, psychological, geometrical, and occasionally logical terms. An example of a definition of a logical term is "'All' is 'none not so'" (supplemented in the *Hsiao-ch'ü* by "'Some' is 'not all'"). The first of the series is "The 'cause' is what is required for something to happen." ("*Minor cause*: With this it will not necessarily be so; without this it necessarily will not be so. *Major cause*: With this it will necessarily be so.") The "Canons" also distinguish the senses of twelve ambiguous terms. Thus, "same" is (1) identical ("two names for one object"), (2) belonging to one body, (3) together, and (4) of a kind ("the same in some respects").

"TA-CH'Ü" AND "HSIAO-CH'Ü." The Moist *Ta-ch'ü* further refined the classification of names. Names indicating "number and measure" cease to apply when their objects are reduced in size; when a white stone is broken up it ceases to be "big," although it is still "white." Names indicating "residence and migration" do not apply when the population moves, as in the case of names of particular states ("Ch'i") or of kinds of administrative divisions ("country"). The claim that one knows *X* only if one knows that an object is *X* applies only to names indicating "shape and appearance" ("mountain," but not "Ch'i" or "county").

The *Ta-ch'ü*, and still more the *Hsiao-ch'ü*, also showed a shift of interest from the name to the sentence and to the deduction of one sentence from another. The Chinese never analyzed deductive forms, but the Moists noticed that the formal parallelism of sentences does not necessarily entitle us to infer from one in the same way as from another, and they developed a procedure for testing parallelism by the addition or substitution of words. For example, "Asking about a man's illness is asking about the man," but "Disliking the man's illness is not disliking the man"; "The ghost of a man is not a man," but "The ghost of my brother is my brother." In order to reconcile the execution of robbers with love for all men some Moists maintained that although a robber is a man, "killing robbers is not killing men." Enemies of Moism rejected this as sophistry, on the assumption that one can argue from "A robber is a man" to "Killing robbers is killing men," just as one can argue from "A white horse is a horse" to "Riding white horses is riding horses." The *Hsiao-ch'ü* replied that there are second and third sentence types of the same form, which do not allow such an inference—for example, "Her brother is a handsome man," but "Loving her brother is not loving a handsome man"; "Cockfights are not cocks," but "Having a taste for cockfights is having a taste for cocks." A four-stage procedure

was used to establish that "A robber is a man" belongs to the second type:

> (1) Illustrating the topic ("robber") with things ("brother," "boat") of which formally similar statements may be made.
>
> (2) Matching parallel sentences about the illustrations and the topic—for instance, "Her brother is a handsome man, but loving her brother is not loving a handsome man"; "A boat is wood, but entering a boat is not entering wood"; "A robber is a man, but abounding in robbers is not abounding in men, nor is being without robbers being without men."
>
> (3) Adducing supporting arguments for the last and most relevant parallels by expanding them and showing that the parallelism still holds: "Disliking the abundance of robbers is not disliking the abundance of men; wishing to be without robbers is not wishing to be without men."
>
> (4) Inferring, defined as "using its [the topic's] similarity to what he [the person being argued with] accepts in order to propose what he does not accept": "Although a robber is a man, loving robbers is not loving men, not loving robbers is not not loving men, and killing robbers is not killing men."

XUNZI

Outside the Moist school only the Confucian Xunzi (c. 313–c. 238 BCE) left a consecutive treatise on logical questions. According to his "Correct Use of Names" the purpose of names is to point out objects, thereby distinguishing the noble from the base and the similar from the different. Names are fixed by convention and are mutable, but to use them idiosyncratically when their usage is fixed is a crime akin to falsifying weights and measures. Objects are different if they differ in place although not in form; they remain the same if they change in form without dividing. Objects of the same kind are perceived by the senses as similar and are given the same name. Names may be of any degree of generality; we may assimilate objects under the name "thing" or distinguish them as "bird" and "beast." (Like the Moists, Xunzi took for granted a nominalist position.) The sentence is a series of names conveying one idea, and a name is understood when we grasp both the object to which it points and its interconnections in the sentence.

Xunzi distinguished three sorts of fallacies, which he illustrated with unexplained examples (two are explained by his refutations of them in his "Treatise of Corrections"). Fallacies that abuse names are exposed by an appeal to established usage, and fallacies that abuse objects are exposed by an appeal to the evidence of the senses. The first fallacy, "confusing names by misuse of names," Xunzi illustrated by "To be insulted is not disgraceful." This is a violation of the established use of "disgrace" in two senses, for social and for moral degradation. The second fallacy, "confusing names by misuse of objects," was exemplified by "Our genuine desires are few." Xunzi criticized this as a factual error about humankind. The third fallacy is "confusing objects by misuse of names." Kung-sun Lung (born 380 BCE) had defended the sophism "A (white) horse is not a horse" on grounds which assume that the question is one of identity, not one of class membership. Xunzi would presumably have replied simply that a white horse is commonly called a "horse."

LATER LOGICAL THOUGHT

The classical period of Chinese philosophy ended about 200 BCE. The next important movement, the neo-Daoism of the third and fourth centuries CE, revived the study of the sophists and the Moist "Canons." Indian treatises on logic were available in translation from the seventh century on; Buddhists wrote commentaries on them during the Tang dynasty (618–907), and in Japan they have continued to do so. But there is little evidence of progress by either Daoists or Buddhists. Neo-Confucianism, the main philosophical movement after the Song dynasty (960–1279), entirely neglected logical inquiry.

CHINESE NEGLECT OF LOGIC

It is well known that almost all Chinese philosophical "systems" are practical, moral, or mystical philosophies of life, indifferent to abstract speculation. It is therefore not surprising that Chinese thinkers have cared little for the forms of reasoning, except under the pressure of the acute controversies of the third century BCE. What is surprising is the almost exclusive interest of Chinese philosophers in the problem of names and the fact that even those who advanced from the name to the sentence studied the parallelism of sentences rather than their analysis.

A reason for this interest can be found in the Chinese language, which organizes uninflected words solely according to word order and the placing of particles. Without the inflections that expose the structure of Sanskrit, Greek, or Arabic sentences and encourage the simultaneous growth of grammar and logic the Chinese sentence, until recently, almost defied analysis; the Chinese have been lexicographers but not grammarians. On the other hand, strict parallelism of clauses—in which

noun is matched with noun, adjective with adjective, adverb with adverb, verb with verb—is part of the ordinary resources of the Chinese language and easily calls attention to the logical dangers of formal parallelism.

See also Chinese Philosophy; Gongsun Long; Hui Shi; Mozi; Proper Names and Descriptions; Xunzi; Zhuangzi.

Bibliography

CHINESE LOGIC

Chao, Y. R. "Notes on Chinese Grammar and Logic." *Philosophy East and West* 20 (2) (1970): 137–154.

Cheng Chung-ying. "On Implication (*tse*) and Inference (*ku*) in Chinese Grammar and Chinese Logic." *Journal of Chinese Philosophy* 2 (1975): 225–244.

Chmielewski, Janusz. "Notes on Early Chinese Logic." *Rocznik Orientalistyczny* 26 (1) (1962): 7–21; 26 (2) (1963): 91–105; 27 (1) (1963): 103–121; 28 (2) (1965): 87–111; 29 (2) (1965): 117–138; 30 (1) (1966): 31–52; 31 (1) (1968): 117–136; 32 (2) (1969): 83–103.

Graham, Angus C. *Disputers of the Tao*. La Salle, IL: Open Court, 1989.

Graham, Angus C. *Later Mohist Logic, Ethics and Science*. Hong Kong: Chinese University Press, 1978; reprinted, 2003.

Hansen, Chad. *Language and Logic in Ancient China*. Ann Arbor: University of Michigan Press, 1983.

Harbsmeier, Christoph. *Science and Civilisation in China*, Vol. 7: *The Social Background*, Part 1: *Language and Logic in Traditional China*. Cambridge, U.K.: Cambridge University Press, 1998.

Hu Shi. *The Development of the Logical Method in Ancient China*. Shanghai: Oriental, 1922.

Lau, D. C. "On Mencius' Use of the Method of Analogy in Argument." *Asia Major* 10 (1963). Reprinted in D. C. Lau, tr. *Mencius*. London: Penguin, 1970, pp. 235–263.

Lau, D. C. "Some Logical Problems in Ancient China." *PAS*, n.s. 53 (1952–1953): 189–204.

Lenk, Hans, and Gregor Paul, eds. *Epistemological Issues in Classical Chinese Philosophy*. Albany: State University of New York Press, 1993.

Leslie, Donald. *Argument by Contradiction in Pre-Buddhist Chinese Reasoning*. Canberra: Australian National University, 1964.

Reding, Jean-Paul. *Les fondements philosophiques de la rhétorique chez les sophistes grecs et chez les sophistes chinois*. Bern: Peter Lang, 1985.

A. C. Graham (1967)
Bibliography updated by Huichieh Loy (2005)

LOGIC IN THE ISLAMIC WORLD

Arabic logic, like the rest of medieval Arabic science and philosophy, is entirely Western and has nothing to do with Oriental philosophy. It developed wholly in the wake of the classical Greek tradition as preserved in and transmitted through late Greek Aristotelianism. The present account briefly traces the evolution of Arabic logic from its inception in the late eighth century to its stultification in the sixteenth century, mentioning only the most important trends, figures, and achievements. Information on individual writers can be found in Carl Brockelmann's monumental *Geschichte der arabischen Litteratur*, cited hereafter as *GAL* (2 vols.—I, II—Weimar, 1890; Berlin, 1902; 2nd ed., Leiden, 1943–1949; 3 supp. vols.—SI, SII, SIII—Leiden, 1937–1942).

TRANSMISSION OF GREEK LOGIC TO THE ARABS

After their conquest of Syria-Iraq the Arabs came into contact with Greek learning as it continued to be nursed by various Christian sects—primarily the Nestorians and the Monophysites, or Jacobites—that had transplanted there (via such centers as Antioch, Edessa, and Nisibis) the Hellenistic scholarship of Alexandria. Thus, the first writers on logic in Arabic were Syrian Christian scholars, and their tradition of logical studies—closely linked to medicine—was transferred to an Arabic-language setting and laid the foundation for the development of Arabic logic.

The Syriac expositors of Aristotelian logic arrived at the following standard arrangement of logical works: *Isagoge* (by Porphyry), *Categories, De Interpretatione, Prior Analytics, Posterior Analytics, Topics, De Sophisticis Elenchis, Rhetoric*, and *Poetics*. These nine works were thought of as dealing with nine distinct branches of logic, each based on its own canonical text. This construction of Aristotelian logic was taken over by the Arabs, resulting in the following organization of the subject matter of logic:

Branch	Arabic Name	Basic Text
(1) Introduction	*al-īsāghūjī*	*Isagoge*
(2) Categories	*al-maqūlāt*	*Categories*
(3) Hermeneutics	*al-'ibārah*	*De Interpretatione*
(4) Analytics	*al-qiyās*	*Prior Analytics*
(5) Apodictics	*al-burhān*	*Posterior Analytics*
(6) Topics	*al-jadal*	*Topics*
(7) Sophistics	*al-mughālitah* (*or al-safsatah*)	*De Sophisticis Elenchis*
(8) Rhetoric	*al-khitābah*	*Rhetoric*
(9) Poetics	*al-shi'r*	*Poetics*

The totality of this organon was referred to as the nine books of logic, or as the eight books with the *Poetics* (or sometimes *Isagoge*) excluded. The first four of these logical treatises were apparently the only ones translated into Syriac prior to 800 and into Arabic prior to 850. They were called the four books of logic, and they constituted

the object of logical studies in the basic curriculum of the Syrian academies.

Arabic translations of Aristotle's logical treatises and of several Greek studies and commentaries on them prepared the ground for the first indigenous Arabic writer on logic, the philosopher Abū-Yūsuf Yaʿqūb ibn Isḥāq al-Kindī (c. 805–873; *GAL*, I, pp. 209–210). His logical writings, however, probably amounted to little more than summaries of the writings of others about the Aristotelian texts.

SCHOOL OF BAGHDAD

In the late ninth and the tenth centuries Arabic logic was virtually the monopoly of a single school of logicians centered at Baghdad. The founders of this school belonged to a closely knit group of Syrian Christians, including the teachers of Abū Bishr Mattā ibn Yūnus and the teachers of these teachers. Its principal continuators were the pupils of Abū Bishr's pupil Yaḥyā ibn ʿAdī and the pupils of these pupils. Virtually all of these men—with the notable exception of al-Fārābī, a Muslim—were Nestorian Christians.

Abū Bishr Mattā ibn Yūnus (c. 870–c. 940; *GAL*, I, p. 207) was the first specialist in logical studies to write in Arabic. He produced the first Arabic translations of *Posterior Analytics* and *Poetics* and translated several Greek commentaries on Aristotelian works (such as Themistius on *Posterior Analytics*). In addition he wrote logical commentaries and treatises of his own, which unfortunately have not survived.

Abū Naṣr al-Fārābī (c. 873–950; *GAL*, I, pp. 210–213) was perhaps the most important logician of Islam. His commentaries, only a fraction of which survive, covered the entire Aristotelian *Organon* in great detail. All later Arabic logicians—even those who, like Avicenna, have opposed al-Fārābī's influence—have seen Aristotle through his eyes. Among the points of special interest in the commentaries of al-Fārābī are (1) a strong emphasis on *ecthesis* (the setting out of terms) as a principle of syllogistic reduction, (2) an increased resort to noncategorical (for instance, hypothetical and disjunctive) types of syllogism, (3) an elaborate treatment of inductive uses of syllogistic reasoning, especially the application of the categorical syllogism in argument by analogy, and (4) a detailed treatment of the problem of future contingency, providing for a reading of Chapter 9 of *De Interpretatione* that does not deny prior truth status to future contingents (anticipating the position of Peter Abelard).

Yaḥyā ibn ʿAdī (893–974; *GAL*, I, p. 207), who studied logic and philosophy with both Abū Bishr and al-Fārābī, not only translated Greek works from Syriac into Arabic but also taught virtually half of the Arabic logicians of the tenth century. He wrote various independent works (including a commentary on *Prior Analytics* that devoted special attention to modal syllogisms), almost none of which have survived.

The three principal achievements of this school of Baghdad are (1) completion of the series of Arabic translations of Greek logical works, (2) the masterly commentaries of al-Fārābī (and possibly others) on the logical treatises of Aristotle, and (3) the elaborate study of certain extra-Aristotelian topics by Abū Bishr Mattā and al-Fārābī (for instance, theory of "conditional," or hypothetical and disjunctive, syllogisms along lines already found in Boethius, and the syllogistic reduction of inductive modes of argument).

AVICENNA AND HIS INFLUENCE

Despite the demise of the school of Baghdad around 1050, the ultimate survival of logical studies in Islam was assured by the fact that logic had, through the mediation of medicine, become an integral constituent of the Arabic medicophilosophical tradition as taken over from the Syrian Christians. From a quantitative standpoint the eleventh century was a low ebb in the history of Arabic logic. Yet this period produced perhaps the most creative logician of Islam, the great Persian scholar Abū ibn Sīnā, known as Avicenna (980–1037; *GAL*, I, pp. 452–458).

Avicenna made a daring innovation. Although greatly indebted to the school of Baghdad, he had nothing but contempt for it because it regarded logic as the study of the Aristotelian texts. Avicenna disapproved of this orientation toward the text rather than the subject. For him, and for the tradition he dominated, a logic book was no longer a commentary on Aristotle but an independent, self-sufficient treatise or handbook that covered the ground after its own fashion. Avicenna's masterpiece is a series of treatises in his monumental *Kitāb al-shifāʾ* dealing with the nine parts of the Arabic logical organon.

An example of Avicenna's originality is the following: In Aristotle and in the Stoics one finds a temporal construction of the modality of necessity that construes "All *X*'s are *necessarily Y*'s" as "At any time *t* all *X*'s-at-*t* are *Y*'s-at-*t*." This construction works well for, say, "All men are necessarily animals" but clearly not for "All men necessarily die." Avicenna distinguished between such cases as:

(1) At every time during its existence every *X* is a *Y* ("All men are necessarily animals").

(2) At most times during its existence every *X* is a *Y* ("All men are necessarily breathing beings").

(3) At some time during its existence every *X* is a *Y* ("All men are necessarily dying beings").

He then constructed a detailed theory of syllogistic inference from temporally modalized propositions of this sort.

Avicenna styled his own work in logic (and philosophy) as Eastern, in deliberate contrast with the Western approach of the school of Baghdad. This Eastern logic espoused by Avicenna differs from that of, say, al-Fārābī not so much in matters of substance as in emphasis and in willingness to depart from Aristotelian precedent. Thus, Avicenna imported into his logic a certain amount of material derived probably from Galen (including an at least grudging recognition of the fourth figure of the categorical syllogism) and certainly from the Stoics (for example, quantification of the predicate of categorical propositions, elaboration of quality and quantity for "conditional" propositions, and a treatment of singular propositions in the manner of the Stoics).

Avicenna's call to study logic from independent treatises rather than via the Aristotelian texts met with complete success in Eastern Islam. Only in Muslim Spain did the tradition of Aristotelian studies of the school of Baghdad manage—for a time—to survive.

LOGICIANS OF ANDALUSIA

During the late eleventh and the twelfth centuries Andalusia (Muslim Spain) was the principal center of logical studies in Islam. Muḥammad ibn ʿAbdūn (c. 930–c. 995; Heinrich Suter, *Die Mathematiker und Astronomen der Araber und ihre Werke*, Leipzig, 1900–1902, no. 161; not in *GAL*), a Spanish Muslim who studied medicine and philosophy in Baghdad, was instrumental in transplanting to Córdoba the teachings of the school of Baghdad in Aristotelian logic. In the medicological tradition of Andalusia these teachings stayed alive for more than two and a half centuries, surviving well past their extinction in Eastern Islam.

Abū'l-Ṣalt (1068–1134; *GAL*, I, pp. 486–487) wrote an influential logic compendium that follows al-Fārābī closely; like most other Spanish Arab logicians, he seems to have had special interest in modal syllogisms. The detailed study of the writings of Aristotle was revitalized by Ibn Bājja (or Avempace; c. 1090–1138; *GAL*, I, p. 460), who wrote an important series (extant but unpublished) of discussions of Aristotle's works based on the commentaries of al-Fārābī.

Ibn Rushd (or Averroes; c. 1126–c. 1198; *GAL*, I, pp. 461–462) was unquestionably the most important of the Arabic logicians of Spain. His elaborate commentaries on the treatises of Aristotle's logical *Organon* rival (and conceivably surpass) those of al-Fārābī in their detailed understanding of Aristotle's logic. Averroes stands, as he considered himself to stand, heir to the masters of the school of Baghdad and successor to the heritage of al-Fārābī.

Among the points of special interest in the Aristotelian commentaries of Averroes are (1) certain historical data—for instance, regarding Galen's origination of the fourth syllogistic figure—taken from the last writings of al-Fārābī, (2) anti-Avicennist polemics that afford us a view of the points of dispute between Avicenna and his opponents, (3) the detailed account of the Aristotelian theory of modal syllogisms, and (4) in general, his effort to systematize as unified doctrine the teachings of the Aristotelian *Organon*.

After Averroes the logical tradition of Muslim Spain entered a period of decline. Arabic logic became extinct in Spain because there—in contrast to Eastern Islam, where logic achieved a modus vivendi with religious orthodoxy—popular and theological hostility toward logic and philosophy as an integral part of "alien learning" continued unabated.

QUARREL OF THE EASTERN AND WESTERN SCHOOLS

Avicenna's criticisms of the school of Baghdad and his shift away from Aristotelian orthodoxy were not received with universal acceptance. A Western school arose to oppose Avicenna's innovations. Its principal exponents were the prolific Persian scholar Fakhr al-Dīn al-Rāzī (1148–1209; *GAL*, I, pp. 506–508) and his followers al-Khūnajī (1194–1249; *GAL*, I, p. 463) and al-Urmawī (1198–1283; *GAL*, I, p. 467). These logicians not only offered detailed criticisms of Avicenna's departures from Aristotle but also wrote handbooks of logic that became standard textbooks both during the lifetime of their school and later.

Opposed to these Westerners, the school of the Easterners, which supported Avicenna, continued to be active throughout the thirteenth century. Its leading exponent was the eminent and versatile Persian scholar Kamāl al-Dīn ibn Yūnus (1156–1242; *GAL*, SI, p. 859). His position

was supported by his pupils al-Abharī (1200–1264; *GAL*, I, pp. 464–465) and Naṣīr al-Dīn al-Ṭūsī (1201–1274; *GAL*, I, pp. 508–512), as well as by the pupils of the last-named scholar, especially the logician al-Qazwīnī al-Kātibī (c. 1220–c. 1280; *GAL*, I, pp. 466–467). These logicians produced polemical treatises to attack the theses of the Westerners, as well as textbooks and handbooks to facilitate the teaching of logic according to their conceptions.

Amid this disputation and textbook writing the logical treatises of Aristotle were completely lost sight of. In effect, Avicenna carried the field before him; in Eastern Islam, Aristotle's logical writings were utterly abandoned. Ibn Khaldūn (1332–1406) could lament, "The books and methods of the ancients are avoided, as if they had never been, although they are full of the results and useful aspects of logic." The handbooks of the two thirteenth-century schools provided a basis for all future study in Islam, completely replacing the works of Aristotle. But very little produced at this stage has any significance for logic as a science rather than as a field of instruction.

FINAL PERIOD

The period 1300–1500 may be characterized as the final period of Arabic logic, when its ossification became complete. It was a time not of creative logicians but of teachers of logic writing expository commentaries and supercommentaries on the thirteenth-century handbooks, now basic to all Arabic instruction in logic.

Underlying this development was the effort of al-Tustarī (c. 1270–c. 1330; *GAL*, SI, p. 816) and his disciple al-Taḥtānī (c. 1290–1365; *GAL*, II, pp. 209–210) to effect an arbitration between the Eastern and Western schools. As a result, later Arabic logicians were free to draw on both sectors of the tradition and to use the handbooks of both schools for the teaching of logic. The flood of glosses and supercommentaries on commentaries on the thirteenth-century logic handbooks marks the final, disintegrative phase of the evolution of logic in Islam.

CONTRIBUTIONS OF ARABIC LOGIC

Some of the original contributions made by the Arabic logicians to logic as a science are (1) al-Fārābī's syllogistic theory of inductive argumentation, (2) al-Fārābī's doctrine of future contingency, (3) Avicenna's theory of "conditional" propositions, (4) Avicenna's temporal construction of modal propositions, and (5) Averroes's careful reconstruction of Aristotle's theory of modal syllogistic. Many of the prominent "innovations" of medieval Latin logic are in effect borrowings or elaborations of borrowings of Arabic ideas (for example, the distinction between the various modes of *suppositio* and the distinction between modality *de dicto* and *de re*).

However, in speaking of the "original contributions" of Arabic logic two qualifications are necessary. In the first place, our knowledge of late Greek logic is so incomplete that any "original" item of Arabic work could turn out to be a mere elaboration of a Greek innovation. Second, an emphasis on originality in discussing Arabic logic is somewhat misplaced in that all the Arabic logicians—even Avicenna, the most original of them all—viewed their logical work as the reconstruction of a Greek teaching rather than as an enterprise of innovation.

See also al-Fārābī; al-Kindī, Abū-Yūsuf Yaʿqūb ibn Isḥāq; Aristotelianism; Aristotle; Averroes; Avicenna; Boethius, Anicius Manlius Severinus; Ibn Bājja; Naṣīr al-Dīn al-Ṭūsī; Porphyry.

Bibliography

ARABIC LOGIC

A complete bibliography of Arabic logic can be found in Nicholas Rescher, *The Development of Arabic Logic* (Pittsburgh: University of Pittsburgh Press, 1964). On the transmission of Greek logic to the Arabs, see Max Meyerhof, "Von Alexandrien nach Baghdad," in *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, Philosophisch-historische Klasse 23 (1930): 389–429. The conflict between logic and Islamic religion is detailed in Ignaz Goldziher, "Stellung der alten islamischen Orthodoxie zu den antiken Wissenschaften," in *Abhandlungen der Königlichen Preussischen Akademie der Wissenschaften*, Philosophisch-historische Klasse, Jahrgang 1915 (Berlin, 1916). For the Arabs' familiarity with Aristotle's logical works, see R. Walzer, "Aristū" (Aristotle), in *Encyclopedia of Islam* (London, 1960), Vol. I; and Ibrahim Madkour, *L'organon d'Aristote dans le monde arabe* (Paris: Vrin, 1934).

Some representative Arabic logical texts accessible in European languages are D. M. Dunlop, translations of several logical *opuscula* of al-Fārābī in *Islamic Quarterly*, 2 (1955)–5 (1959); Nicholas Rescher, *Al-Fārābī's Short Commentary on Aristotle's "Prior Analytics"* (Pittsburgh, 1963); A. M. Goichon, *Avicenne: Livre de directives et remarques* (Paris, 1951); Mohammad Achena and Henri Massé, *Avicenne: Le livre de science*, Vol. I (Paris, 1955); *Aristotelis Opera cum Averrois Commentariis* (Venice, 1550 and later; 1562–1574 ed. reprinted photographically, Frankfurt: Minerva, 1962).

Substantive study of the contributions of Arabic logicians has only begun. In addition to Prantl, *Geschichte der Logik im Abendlande*, Vol. II, above, consult T. J. de Boer, "Mantiḳ," in *Encyclopedia of Islam*, 1st ed.; and Nicholas Rescher, *Studies in the History of Arabic Logic* (Pittsburgh: University of Pittsburgh Press, 1963).

Nicholas Rescher (1967)

LOGIC IN THE ISLAMIC WORLD [ADDENDUM]

For more on everything in the entry, see especially Hans Daiber's *Bibliography of Islamic Philosophy* (1999). Few scholars would now accept that Arabic logic is "entirely Western"; it grew out of Greek texts, but developed differently from both Hellenistic and Latin logic.

TRANSMISSION OF GREEK LOGIC TO THE ARABS

Research on the translation of the books of the *Organon* and their attendant commentaries is presented in summary essays in Goulet (1989–2003, pp. 502ff).

THE SCHOOL OF BAGHDAD

The leading representative of the textual Aristotelianism of Baghdad was al-Fārābī, and much of his extant work is now either edited or translated (see Lameer 1994).

AVICENNA AND HIS INFLUENCE

The many new editions, translations, and studies of Avicenna are listed by Jules L. Janssens (1999). An attempt to deal philosophically with his modal syllogistic is made by Paul Thom (2003, chapter 4 and idem). See also his essay "Logic and Metaphysics in Avicenna's Modal Syllogistic" (forthcoming).

LOGICIANS OF ANDALUSIA

Averroes, though without much influence in the Islamic world, is the most acute of the Andalusian logicians. See Thom (2003, chapter 5) for a philosophical treatment of his later modal syllogistic.

QUARREL OF THE EASTERN AND WESTERN SCHOOLS

There certainly were major differences among the post-Avicennan logicians, but Nicholas Rescher's use of "Eastern" and "Western" schools to gather them into opposing camps is misleading (see Street 2004, pp. 567ff).

FINAL PERIOD

One cannot assume the tradition ossified because its most common genre became the commentary. The task ahead is to read and appraise the profusion of texts written from the 900s until after the colonial invasions of the nineteenth century. For a study of the attitudes to logic in this period, see Khaled El-Rouayheb's "Sunni Muslim Scholars on the Status of Logic, 1500–1800" (2004).

See also al-Fārābī; Averroes; Avicenna; Islamic Philosophy; Rescher, Nicholas.

Bibliography

Daiber, Hans. *A Bibliography of Islamic Philosophy*. Leiden, Netherlands: Brill, 1999.

El-Rouayheb, Khaled. "Sunni Muslim Scholars on the Status of Logic, 1500–1800." *Islamic Law and Society* 11 (4) (2004): 213–232.

Goulet, Richard, ed. *Dictionnaire des philosophes antiques*, Vol. 1. Paris: Editions du Centre national de la recherche scientifique, 1989–2003.

Janssens, Jules L. *An Annotated Bibliography on Ibn Sīnā: First Supplement (1990–1994)*. Louvain-la-Neuve: Fédération Internationale des Instituts d'études médiévales, 1999.

Lameer, Joep. *Al-Fārābī and Aristotelian Syllogistics: Greek Theory and Islamic Practice*. Leiden, Netherlands: Brill, 1994.

Street, Tony. "Arabic Logic." In *Handbook of the History of Logic, Vol. 1: Greek, Indian, and Arabic Logic*, edited by Dov M. Gabbay and John Woods, 523–596. Amsterdam: Elsevier, 2004.

Thom, Paul. "Logic and Metaphysics in Avicenna's Modal Syllogistic." In *The Unity of Science in the Arabic Tradition: Metaphysics, Logic and Epistemology and Their Interactions*, edited by Shahid Rahman, Tony Street, and Hassan Tahiri. Dordrecht, Netherlands: Kluwer-Springer, forthcoming.

Thom, Paul. *Medieval Modal Systems: Problems and Concepts*. Aldershot, U.K.: Ashgate, 2003.

Tony Street (2005)

MEDIEVAL (EUROPEAN) LOGIC

Although some elementary work was done in the ninth and tenth centuries it was not until the end of the eleventh century that medieval logic really began to develop a character of its own. It started as glosses and commentaries on some of a small number of texts that had survived from antiquity. These included Boethius's translations of Porphyry's *Isagoge*, Aristotle's *Categories* and *De interpretatione*, and two works written by Boethius himself, a treatise, *De Topicis Differentiis*, on topical inference based on the work of Themistius and Cicero, and another, *De divisione*, devoted to the various forms of division employed in logic. In the thirteenth century these works were collectively known as the *logica vetus*.

In addition logicians at the beginning of the twelfth century possessed Boethius's very extensive commentaries on the *Isagoge*, *Categories*, and *De interpretatione*, his two-part epitome of Aristotle's *Prior Analytics*, 1–7, *Introductio ad syllogismos categoricos* and *De syllogismo categorico*, his treatise on hypothetical syllogisms, *De*

hypotheticis syllogismis, and his commentary on Cicero's *Topica, In Topica Ciceronis*.

Also important in the early development of logic were Marius Victorinus's *De diffinitionibus*, Saint Augustine's *De dialectica*, and, at least in the ninth and tenth centuries, *De decem categoriae*, a fourth-century Latin translation of a Greek paraphrase of Aristotle's *Categories* attributed to Augustine. In addition, Priscian's *Institutiones grammaticae*, with the eleventh- and twelfth-century glosses on it known as the *Glossulae*, were an important influence in the twelfth century on the development of philosophical semantics and in particular of theories of the substantive verb *to be*.

Boethius's translations of Aristotle's *Sophistical Refutations*, *Topics*, and *Prior Analytics* were recovered before the middle of the twelfth century. Along with the translation of the *Posterior Analytics* made then by James of Venice they provided logicians with what was distinguished from the *logica vetus* as the *logica nova*. Apart from the *Sophistical Refutations*, however, it was not until the beginning of the thirteenth century that the works of the *logia nova* had a significant impact on the development of logic. Although some parts of Avicenna's logical works were translated into Latin, unlike other areas of philosophy, Arabic writing had little impact on the development of logic.

From the middle of the twelfth century logicians developed their discipline in various ways and produced works characteristic of what would much later be referred to as the *logica modernorum*. These dealt, for example, with the properties of terms, and in particular the theory of supposition, syncategorematic words, modality, *obligationes, insolublia*, consequences, and sophisms of various kinds, each of which is discussed in this entry.

THE BOETHIAN BACKGROUND

Based as it was upon the texts of the *logica vetus* medieval logic included a great deal that has to do with ontology and philosophical semantics rather than with logic more narrowly construed as the theory of valid argumentation. Boethius gave medieval logicians much of their terminology but his commentaries on Aristotle and even more so his own works are essentially elementary, often confused, and sometimes inconsistent. It was these, however, which provided twelfth-century logicians with the material from which they constructed their new formal and philosophical logics. In particular, the remarkable developments they made in theory of inference had their beginnings in reflection on Boethius's *De Topicis Differentiis* and *De hypotheticis syllogismis*.

TOPICAL INFERENCE. Medieval logic at least in the first half of the twelfth century was characterized by an intense interest in conditional propositions and in the nature of topical inference as formulated by Boethius in *De Topicis Differentiis*. Logicians at this time were not generally concerned to regiment arguments into the modes and figures of the categorical syllogism but everywhere they classified inferences in accordance with lists of topics, based upon those given by Boethius.

In his treatise Boethius proposes to show how arguments may be discovered to settle any given question. What has to be found, he claims, is what Cicero, in his *Topica*, calls an "*argumentum*"—defined as a "reason which brings conviction where something is in doubt." An argument (*argumentatio*) is the expression in speech or writing of the proof of a conclusion constructed with the required *argumentum*. A *locus*, or topic, is the "site," or "source," of *argumenta* (*Diff. Top*. I, 1174D).

Argumenta are invoked by Boethius to warrant the enthymematic inference of a categorical conclusion from categorical premisses or the direct proof of a conditional proposition. In each case what is needed is a principle that is not itself provable, called by Boethius a *maximal proposition*, and a relevant fact about the items mentioned in the conclusion. For example, by appealing to the maximal proposition "a genus is predicated of whatever its species is predicated" and the truth that animal is the genus of human being we may either infer from the premiss that Socrates is a human being the conclusion that he is an animal or, directly, the corresponding conditional.

The various relationships which Boethius holds may exist between the predicate and subject of a true categorical proposition or between the antecedent and consequent of a true conditional provide him with his *loci* (*Diff. Top*. II, 1186C). With each *locus* there are associated all the maximal propositions warranting inferences which may be made on the basis of that relationship. The enthymeme above, for example, would be characterized as holding "from species," that is, in virtue of the relationship in which a species stands to its genus.

Boethius gives the lists and classifications of the *loci* provided by both Themistius and Cicero. They are divided into those which are intrinsic, that is, having to do only with the things themselves about which a question is asked, and those which are extrinsic, having no such connection with them. (*Diff. Top*. II, 1186D) Examples of intrinsic *loci* are that from species, given above, and that from what is defined, for which one maximal proposition is: "of that of which what is defined is not

predicated, the definition is not predicated." Examples of extrinsic *loci* are that from authority, which justifies inferences from the authority of the majority of people, or the relevant experts, and *loci* from various kinds of opposition.

Argumenta drawn from the locus from authority are not necessary according to Boethius but they are probable in the sense of being generally convincing. Where Aristotle had taken probability and necessity to be properties of the premises and conclusion of a dialectical syllogism, however, Boethius takes them to characterize the nature of the inference from the premiss, or premisses, to the conclusion of an argument and the corresponding connection between the antecedent and consequent of a conditional (*Diff. Top*. I, 1180C).

THE THEORY OF CONDITIONAL PROPOSITIONS. In *De Topicis Differentiis*, Boethius classifies conditional propositions according to the quality of the antecedent and consequent. He accepts what we would now call a principle of contraposition and so maintains that a topical relationship warrants a conditional of the form "if something's an *A*, then it's a *B*" if and only if it warrants one of the form "if something's not a *B*, then it's not an *A*," where *A* and *B* are general terms such as "human being" and "animal." Conditionals of the form "if something's an *A*, then it's not a *B*" are true, he maintains, only for items which are "opposites," that is, opposed exclusively but not exhaustively. For example, "if something's a human being, then it's not a donkey." Those of the form "if something's not an *A*, then it's a *B*" hold only for items which are "immediates," that is, opposed exclusively and exhaustively. For example, "if something's not well, then it's ill" (*Diff. Top. I*, 1179C).

With *De hypotheticis syllogismis* Boethius provided twelfth-century logicians with an account of the logic of certain conditional and disjunctive propositions but neither he nor any other ancient source provided them with what we would recognize as a propositional logic. Boethius had no clear understanding of the nature of either propositionality or propositional operation (Martin 1991).

In his general treatment of compound propositions in his long commentary on *De interpretatione*, Boethius thus denies that the copulative conjunction "and" does anything other than punctuate a list (*2 In Peri. Herm.*, 5, 109). In the same work he also explicitly rejects the Stoic practice of preposing a negative particle to a categorical proposition as ambiguous between the negation of the subject and predicate terms (*2 In Peri. Herm.*, 10, 261–2).

Without a notion of propositionality, Boethius has no notion a propositional form or of the substitution of propositional contents into propositional contexts to obtain new contents of arbitrary complexity. In *De hypotheticis syllogismis* he thus lists all the various kinds of hypothetical syllogism which he accepts for each different quality of the component categorical propositions. There, just as everywhere else where Boethius employs it, the negative particle preposed to a conditional never takes the whole of the following conditional proposition for its scope but always acts only on the consequent.

Boethius designates a conditional as affirmative if its consequent is affirmative and negative if it is negative no matter what the quality of the antecedent (*Hyp. Syll.* 1.9.6). The only compound propositions he considers are simple conditionals and disjunctions, that is those whose components are both categorical, and compound conditionals of which one or more component is a simple conditional. The most complex form of conditional he considers has simple conditionals for both its antecedent and consequent. These compound conditionals, again, have nothing to do with propositional logic as it is now understood. Conditionalized instances of contraposition, for example, are not true instances of the form since Boethius requires for the truth of "if (if something's an *A*, then it's a *B*), then (if something's a *C*, then it's a *D*)" that both "if something's an *A*, then it's a *C*" and "if something's a *B*, then it's a *D*" are true (*Hyp. Syll.* 3.9.1).

In *De hypotheticis syllogismis* Boethius gives the basic truth-condition for a conditional proposition, or consequence (*consequential*), which will be accepted throughout the middle ages. To "destroy" such a proposition, that is, to show that it is false, he says, one must show that it is possible for the consequent to be false when the antecedent is true. A conditional is thus true only if the truth of the antecedent is inseparable from that of the consequent. A simple disjunction, "something's an *A* or it's a *B*," is equivalent, according to Boethius, to a simple conditional with a negative antecedent and affirmative consequent and so holds only for terms connected to one another as immediates (*Hyp. Syll.* 1.3.3).

In addition to stating the inseparability condition for their truth Boethius makes a distinction between conditionals which has profound consequences for the development of medieval logic and metaphysics. He claims that a relation of consequence may be indicated with either "*si*" ("if") or equivalently with "*cum*." The latter, however, usually means *when*, or *whenever* in Latin and that is how it is translated here.

The truth of an antecedent, Boethius notes, will be inseparable from that of a consequent if both are necessarily true even if there is no explanatory connection between them as, for example, "whenever (*cum*) fire is hot, then the heavens are spherical." Boethius does not notice, however, nor does any other ancient source available in the twelfth century, that the inseparability requirement is apparently also satisfied by any conditional whose antecedent is impossible, or whose consequent is necessary.

Boethius designates as "accidental consequences" conditionals formed with "whenever" which meet the inseparability requirement merely on account of the truth-value of their components. He contrasts them with "natural consequences," formed with "if," in which the truth of the antecedent is inseparable from that of the consequent in virtue of an explanatory connection between them. For example "if something's a human being, then its an animal" (*Hyp. Syll.* 1.3.6).

Finally, although Boethius correctly observes that Aristotle wrote nothing about hypothetical syllogisms, he takes from *Prior Analytics*, II. 4, as basic for the logic of conditional propositions what has been called Aristotle's Principle: No two conditionals of the form "if something's *A*, then its *B*" and "if something's not *A*, then its *B*" can both be true (*Hyp. Syll.* 1.4.1).

ABELARD AND THE DISCOVERY OF PROPOSITIONALITY

Peter Abelard, the first significant, and arguably the greatest, of all medieval logicians taught in Paris at various times between 1101 and 1140. Although most logical writing which we have from the twelfth century has been transmitted anonymously and with no certainty about its date of production, very fortunately both Abelard's own survey of logic, the *Dialectica*, written probably around 1116, his *Logica*, consisting of commentaries on Porphyry, Aristotle, and Boethius, written around 1120, and his *Glossulae* on Porphyry, written in the 1120s, have survived more or less intact. The following account of logic in the first half of the twelfth century is thus mainly an account of Abelard's work. He was, however, certainly not the only logician active at the time and much of his writing consists of arguments against sophisticated but unnamed opponents.

Most important, Abelard understood the distinction between the propositional content of a sentence and the force with which it is uttered (Martin 2004). The propositional content "that Socrates is running," for example, may be asserted with an assertive utterance of "Socrates is running" or it may contribute to the meaning without itself being asserted in an assertive utterance of the conditional "if Socrates is running, then he is moving." Since Boethius treats "proposition" (*propositio*) and "assertion" (*enuntiatio*) as synonyms, however, it was rather difficult for Abelard to formulate clearly the distinction for an assertion between force and content.

Abelard uses the term "proposition" (*propositio*) to refer to a token propositional sentence. In his early writings he borrows from Priscian the expression "the being of the thing" (*essentia rei*) to speak about propositional content and identifies it with a state-of-affairs. In later writings he refers rather to the *dictum* of a proposition, that is, to "what is said" with it. For Abelard it is *dicta* which are in the first place the bearers of truth and falsity and so, for example, a conditional is true if and only if the truth of the *dictum* of the consequent follows from the truth of the *dictum* of the antecedent.

The distinction between force and content, which Peter Geach has called the *Frege Point* in deference to its supposed discoverer, is crucial for the development of genuinely propositional logics. Abelard saw this and consequently rejected Boethius's views on copulative conjunction. To the contrary, he insists that a copulative conjunction of propositions is itself a single proposition and may thus be subject to a further propositional operation. "It's not the case that (*p* and *q*)" where "*p*" and "*q*" are propositions is just as much a single proposition, he insists, as "it's not the case that (if *p*, then *q*)."

ABELARD'S TWO NEGATIONS. Negation is the simplest propositional operation. If it is defined truth-functionally, it takes any propositional content and produces another, its contradictory, false if the first is true and true if it is false.

The invention of this operation in Latin logic cannot quite be claimed with certainty for Abelard. It is possible that it was used by his predecessors since it appears in very limited way in a discussion of the appropriate way to negate a simple conditional proposition in the *Dialectica* of Garlandus Compotista, apparently written in the second decade of the twelfth century roughly contemporary with Abelard's *Dialectica*.

Abelard, however, is the first Latin writer known to us who discusses propositional negation in general and applies it both to simple and compound propositions (Martin 2004). He distinguishes, indeed, two kinds of negation. First, and principally, propositional negation, which he calls "destructive" negation, and which has the whole of the following propositional content for its

scope. Second, and derivatively, a negation, which he refers to as "separative" which in the case of affirmative categoricals is obtained by negating the predicate (*Dial.* II.2, 173 sq.). Abelard follows Boethius in classifying conditionals as affirmative according the quality of their consequents. The separative negation of a given affirmative conditional is obtained by negating its consequent either destructively or separatively.

A necessary condition for the truth of both an affirmative categorical and its separative negation is that the subject term is not empty. There is no such requirement for the truth of its destructive negation.

With this distinction between negations Abelard constructs an account of the relationships between quantified propositions which results in effect in a rectangle of opposition rather than the famous square of Aristotle as Boethius understood it. Aristotle gives "not every *A* is *B*" as the contradictory opposite of "every *A* is *B*" in *De interpretatione* but in the *Prior Analytics* "some *A* is not *B*" and according to Boethius the meaning is the same.

Abelard, however, argues that "some *A* is not *B*" is not the contradictory of "every *A* is *B*" but rather "it is not the case that every *A* is *B*." He thus avoids the problem typically raised against Aristotle's logic of quantified terms, that since it requires for the truth of a universal affirmation that the subject term is not empty, given there are no chimeras, an affirmation such as "every chimera is conversing" is false. It follows that its contradictory is true. Since "some chimera is not conversing" is true, however, only if the subject term is not empty, there must be some chimeras for it to be true of! For Abelard this is not a problem since on his account both propositions are false (*Log. "Ingred." sup. Perierm.* 7, 408–11).

THE MANIPULATION OF MODALITY. Once the notion of propositional content was available the difference between two different interpretations of modal propositions could be formulated precisely. In his *Dialectica* Abelard notes that a mode may appear in a categorical proposition either as an adverb or an adjective as, for example, in "Socrates is possibly a bishop" and "that Socrates is a bishop is possible" (Knuutila 1993). Abelard holds that though they differ syntactically these two propositions are semantically equivalent and it is the first which properly expresses the intended meaning since possibility is properly attributed to things (*de rebus*) (*Dial.* II.2, 191sq.). The adverb serves to indicate that the inherence of the predicate in the subject is modified in some way. Later medieval logicians will refer to this as the *de re* reading of the modal claim.

In the case of true *de re* claims about possibility there is of course no actual inherence to modify and Abelard holds that such propositions are true just in case the nature of the subject is compatible with the predicate. Human nature is compatible with being a bishop so "Socrates is possibly a bishop" is true even though he never has been nor never will be one (*Dial.* II.2, 193).

Abelard records that one of his masters proposed an alternative account of propositions with adjectival modes. They are to be understood, he held, as claims about the possibility, necessity, etc. of the sense (*de sensu*), that is the propositional content, of the simple propositions from which they "descend." Against this interpretation Abelard, in effect, argues that if we substitute for a given propositional content an equivalent one, the truth-value of the proposition will remain the same. Since universal negatives convert simple, "no blind man is a seeing man" is equivalent to "no seeing man is a blind man." While his opponents accept, however, that "no blind man is possibly a seeing man" is true, since they agree that the blind do not regain their sight, they claim that "no seeing man is possibly a blind man" is false. The *de sensu* reading, however, requires them to have the same truth value (*Dial* II.2, 196).

Although he maintains in the *Dialectica* the *de sensu* reading is in general not the proper way to interpret modal propositions, Abelard does allow that it is correct for the adjectival modes "true" and "false" since these, he argues, they are properly predicated of propositional contents (*Dial.* II.2, 204–6).

Abelard discusses the same questions at length in his *Logica* in commenting on Aristotle's account of the relations between modalities in *De interpretatione*, 12. He notes, in the first twelfth-century reference to the *Sophistical Refutations*, that the distinction he is interested corresponds to that made by Aristotle between reading a proposition such as "a standing man is possibly sitting" in a composite (*per compositionem*), or a divided way (*per divisionem*). Here, however, Abelard does not insist on the reading *de rebus* but rather works out in detail the relations between modal claims of both kinds (Abelard 1958, 13).

ABELARD ON ENTAILMENT. In his logical works Abelard sought to unify into a single theory of inference the disconnected remarks on topics and the consequence relation which he found in Boethius (Martin 2004). To do this he provides a new general definition of a *locus* as the force of, or as we would say, the warrant for an entailment (*vis inferentiae*) (*Dial.* III.1, 253). He then devotes hun-

dreds of pages of his logical works to investigating the role of *loci* thus conceived in proving conditionals and validating the corresponding enthymemes.

According to Abelard, a proposition *p* entails a proposition *q*, just in case the corresponding conditional, or consequence, *if p, then q* expresses a relationship of following, or "consecution" (*consecutio*). For this to be so, he holds, the sense of the antecedent, that is, its propositional content, must contain that of the consequent. Abelard characterizes this kind of connection as necessary but insists that it must be distinguished from the satisfaction of the inseparability condition which it guarantees, and which alone provides only the necessity of what he calls association (*comitatio*) (*Dial.* III.2, 459).

Entailments are divided by Abelard into the perfect and the imperfect. Perfect entailments satisfy the containment requirement in virtue of the form, or structure, of the propositions involved. Imperfect entailments are those in which the sense of the antecedent contains that of the consequent but does not do so in virtue of their form (*Dial.* III.1, 253).

Abelard makes the notion of perfection, and so form, more precise, and anticipates modern definitions of logical truth, by giving as a necessary condition for perfect entailment that consecution is preserved through all uniform substitutions of terms or propositional contents. He does not, however, regard the condition as sufficient and, in particular, although he classifies the conditionalizations of all valid categorical and hypothetical syllogisms as perfect, he holds that instances of the principle of reflexivity, *if p, then, p*, are imperfect, presumably because they fail to have a canonical syllogistic form. Like all other imperfect entailments, according to Abelard, they must thus be warranted as instances of an appropriate maximal proposition (*Dial.* III.1, 255).

By far the greatest part of Abelard's *Dialectica* is concerned with establishing just which conditional propositions express imperfect entailments. Boethius in *De Topicis Differentiis* says that he will explore which *loci* are suited to which syllogisms and according to Abelard this led some logicians to hold that even the canonical syllogistic figures needed topical warrants. He and his mid-twelfth century followers known, probably because of their views on universals, as the *Nominales*, rejected this. They held rather that putative principles cited to support categorical and hypothetical syllogism are simply their metalinguistic formulation as rules. They contain no term indicating a topical relationship, that is no *locus differentia*, upon which the inference in question rests (*Dial.* III.1, 256–263).

Imperfect entailments, according to Abelard, are conditionals and the corresponding enthymemes, which satisfy the two conditions necessary and sufficient for following for a restricted range of terms. The topical difference specifies the relevant substitution class and the maximal proposition warrants the inference for substitutions from that class. For example, the conditional "if Socrates is a human being, then Socrates is an animal" is true and so are all substitutions for "human being" and "animal" which stand in the relationship of species to genus. For example, "if Socrates is a pearl, then Socrates is a stone," warranted by the maximal proposition "of whatever a species is predicated, so is its genus" (*Dial.* III.1, 315).

NECESSITY. Abelard's main task in his discussion of topical inference is to establish just which topical relations and which maximal propositions warrant true conditionals. He argues in the *Dialectica* that since what is being proved are conditional propositions, even though their surface form may be categorical, maximal propositions must in fact be general conditionals "containing" each of the proved conditionals as their instances. His treatment of this question involves a sophisticated discussion of how relative pronouns function in quantified propositions and the rules for logically manipulating them.

Since Boethius had allowed that some *argumenta* are probable but not necessary certain of Abelard's contemporaries had, he tells us, accepted as true any conditionalization of an enthymeme supported by a probable maximal proposition. In particular they took to be true conditionals warranted by maximal propositions which guarantee the inseparability of association but not the following or consecution which Abelard requires for entailment (*Dial.* III.1, 271 sq.).

Against them Abelard invokes the principle from the *Prior Analytics* mentioned above. His opponents accept conditionals warranted by appeal to the *locus* from immediates and the maximal proposition "of that from which one of a pair of immediates is removed the other is predicated." They must thus accept the following argument: [I1] if something does not exist, then it is not well (by the *locus* from part to whole, since "not-well" is predicated of all non-existent things as well as all existing things which are not well); [I2] if something is not well, then it is sick (from immediates); [I3] if something is sick, then it exists (from part to whole); so, by transitivity, [I4] if something does not exist, then it is sick, and thus [I5] if something does not exist, then it exists. [I5], how-

ever, contradicts Aristotle's principle and, Abelard maintains, is obviously impossible (*Dial.* III.1, 276).

Abelard investigates in detail various proposals to modify [I2] to block the embarrassing inference while retaining its warrant from immediates. In particular he considers various ways of adding what he calls a "temporal" qualification, indicated with "when" ("*cum*"), to form propositions such as "if (when something's an animal, it's not well), then it's sick."

Boethius, as noted, claims that "if" and "when" are equivalent as indicators of a conditional connection and in *De hypotheticis syllogismis* he invariably gives the conditional components of compound conditionals with "when." For example, "if (when something's an *A*, it's a *B*), then it's a *C*." This practice allows Abelard to treat the embedded propositions as temporal rather than conditional in interpreting Boethius claims about the hypothetical syllogism (*Dial.* IV.1, 472 sq.).

The problem for Abelard is that having insisted that one destroys a conditional by showing that it is possible for the antecedent to hold without the consequent, Boethius apparently assumes that an affirmative simple conditional and the corresponding negative conditional are contradictory opposites. He thus claims to be valid, for example, syllogisms of the form "if (when something's an *A*, its a *B*), then it's a *C*, but it's not a *C*; therefore when something's an *A*, it's not a *B*."

Abelard in the end rejects Boethius's account of the hypothetical syllogism. In this case, for example, he maintains, contrary to Boethius, that the valid argument is rather an instance of *modus tollens* (*if p, then q, not:q; therefore not:p*) which concludes with the propositional negation of the antecedent: "if (when something's an *A*, it's *B*), then it's a *C*, but it's not a *C*; therefore it is not the case that (when something's an *A*, it's *B*)." Abelard thus, in effect, replaces Boethius's account of the hypothetical syllogism with a genuinely propositional theory which takes *modus ponens* (*if p, then q, p; therefore q*) and transitivity (*if p, then q, if q, then r; therefore if p, then r*) as basic principles and *modus tollens* as a derived principle and holds that all uniform substitution instances, no matter how complex, are valid (*Dial.* IV.1, 498 sq.).

Abelard was unable to save Boethius's account of the hypothetical syllogism and so he replaced it with the correct one. Apparently no one else could to do any better and *De hypotheticis syllogismis* disappeared from the logic curriculum some time in the twelfth century. It is not until Walter Burley (1274–1344) published *De puritate artis logicae* in about 1325 that hypothetical syllogisms were discussed in any detail again, and there the conditional premisses are always simple conditionals.

RELEVANCE. Abelard accepts that the *locus* from immediates and many others guarantee the inseparability of association, but he also requires a relevant connection between antecedent and consequent for the conditional to be true (Martin 2004). He does not, however, insist on relevance for the validity of an argument. So long as it is impossible for the premisses to be true and the conclusion at the same time false, true premisses will guarantee a true conclusion and that is all that an argument is asked to produce. Abelard thus denies as a general principle what we would now call the Deduction Theorem, that an argument *p; therefore q* is valid if and only if the corresponding conditional *if p, then q* is true (*Dial.* III.2, 455).

Abelard's distinction between association and following or consecution as two kinds of necessary connection is based on the account given in the *Isagoge* of the relationship between substances and their accidents. According to this a substance does not require a particular accident in order to exist and so accidents are separable from their subjects. The problem is that while a given substance may undergo a change with respect to certain of its accidental features there are others, according to Porphyry, which must always be present. Blackness, for example, in the case of crows, and the property of being able to laugh in the case of humans. Neither of these are included in the account of what it is to be a crow or to be human but there is no natural possibility of their subjects existing without them. Such "inseparable" accidents can, however, it is claimed, be removed in the sense that we can conceive of a crow without conceiving its blackness. They are thus contrasted with definitional features which are included as part of its essence, in the definitional account of what it is to be a particular kind of thing (*Log. "Ingred." sup. Porph.* 6, 93).

Abelard's two necessities are a generalization of this distinction between actual and conceptual inseparability. He points out in his own discussion of inseparable accidents that although the antecedent and consequent of "if Socrates is a stone, then Socrates is a pearl," are inseparable, a pearl being classified as a kind of stone, nevertheless the conditional is false. The antecedent and consequent are inseparable, and Abelard is the first medieval logician we know of to make this point, merely because the antecedent is impossible. He goes to point out that if the inseparability of association were sufficient as well as necessary for following, then any conditional with an impos-

sible antecedent would be true. For example, "if Socrates is a stone, then Socrates is a donkey" (*Dial.* III.1, 285).

Abelard does not, however, formulate the famous principles that anything follows from an impossibility and that a necessity follows from anything. He could not be expected to do so, however, since given his definition of following they are false.

Abelard believes that his own account of the semantics of the conditional generates what we would today call a connexive logic, a logic, that is, for which no proposition can entail or be entailed by its contradictory opposite. These principles entail, Abelard recognizes, both the propositional version of Aristotle Principle and what we may call Abelard's Principle: No two conditionals of the form *if p, then q* and *if p, then not:q*, can both be true.

Abelard accepts simplification (*if (p and q), then p* and *if (p and q), then q*), contraposition (*if (if p, then q), then (if not:q, then not:p)*), and transitivity (*if p, then q, if q, then r; therefore if p, then r* is valid). Suppose, then, that Abelard's Principle is false for some *p* and *q*, that is both (1) *if p, then q* and (2) *if p, then not: q* are true. But then if (3) *if (p and not:q), then p* is true and likewise (4) *if q, then not:(p and not:q)*, we may infer by transitivity that *if (p and not:q), then not:(p and not:q)*, an instance of *if p, then not:p*, which Abelard insists is a paradigm of impossibility. Abelard's Principle is thus necessarily true and he gives a similar argument to prove Aristotle's Principle (*Dial.* III.1, 290).

From these principles there follows the most characteristic feature of the logical theory advocated by Abelard and the *Nominales*: No conditional can be true of which the antecedent and the consequent differ in quality. For example if *if p, then not: q* were true, for some *p* and *q*, then *if* (*p and q*), then not:(*p and q*) would true by transitivity and contraposition.

Most famously Abelard argued against the *locus* from opposites in this way. If the *locus* warranted a true conditional then the conditional "if Socrates a human being, then Socrates is not a donkey" would we be true and we could infer the impossibility "if Socrates is a human being and a donkey, then it is not the case that Socrates is a human being and a donkey." He sees too, and explicitly acknowledges, that it follows from the principles of his logic that the conditional principle of double negation (*p if and only if not:not:p*) is false in both directions (*Dial.* II.2, 179).

Unfortunately Abelard's various intuitions about the propositional connectives are inconsistent (Martin 1987). In particular the principles which he holds to govern negation are incompatible with simplification. This point seems to have been first noticed the 1130s by Alberic of Paris who confronted Abelard with the following argument: The conditional [A0] "if Socrates is a human being, then he is an animal" is a paradigm of entailment according to Abelard. He must also accept each of the following: By simplification [A1] if Socrates is human and Socrates is not an animal, then Socrates is not an animal; by contraposition, [A2] if Socrates is not an animal, then Socrates is not a human being; again by contraposition, [A3] if Socrates is not a human being, then it is not the case that Socrates is human being and Socrates is not an animal; so by transitivity, [A4] if Socrates is human being and Socrates is not an animal, then it is not the case that Socrates is a human being and not an animal—contradicting a fundamental principle of Abelard's logic. Alberic's proof of inconsistency precipitated a crisis in the history of logic.

THE PARISIAN SCHOOLS AND THE CRISIS OVER THE CONDITIONAL

In middle decades of the twelfth century a number famous logicians were active at Paris and with each of was associated a school (Martin 1987). In some cases very substantial treatises have survived from these schools, illustrating that this was a period of intense activity in logic. Unfortunately most of these and certainly the most important are still unpublished. The schools may be distinguished by their response to Alberic's proof of the inconsistency of Abelard's system.

Abelard's own followers, the *Nominales*, continued to maintain the correctness of his account of the conditional and the connexive principles. Their strategy seems to have been to take negation to be a cancellation of content so that nothing follows from *p and not:p* rather than both *p* and *not:p*.

The followers of Alberic, the *Montani*, so-called because their school was located on Mont Ste. Geneviève, held that the argument failed because the conjunction of contraries in [A1] undermined the relationship on which [A0] was based. In a different context Abelard himself anticipates this objection to impossible antecedents and argues at length against it that since the antecedent is not asserted, and the argument is formally valid, the conclusion follows.

The school of Gilbert of Poitiers, the *Porretani*, held that the problem lay in the unrestricted principle of simplification. They required, as do twentieth century connexive logics, that both conjuncts play a role in such an inference. The most surprising response was that of the

followers of Robert of Melun, the *Melidunenses*, who took as their basic principle for the logic of the conditional the rule "nothing follows from the false."

The solution that eventually won the day, however, was that proposed by the followers of Adam of the Little Bridge, the *Parvipontani*, so called again because of the location of their school in Paris. They accepted that the argument was sound because they apparently held that inseparability alone is both necessary and sufficient for the truth of a conditional. Aristotle's Principle thus fails when the consequent is necessary and Abelard's when the antecedent is impossible.

John of Salisbury tells us in his *Metalogicon* (1159) that one of his students, William of Soissons, had gone on to join the *Parvipontani* and discovered the twelfth-century version of one of the twentieth century's most famous arguments, the proof that *ex impossibli quodlibet*, the so-called paradox of strict implication, according to which anything follows from an impossibility (*Metalogicon* II.10).

In his *De naturis rerum* written at the end of the twelfth century Alexander Neckham gives the argument as follows: [S1] if Socrates is a human being and Socrates is not a human being, then Socrates is a human being; [S2] if Socrates is a human being, then Socrates is a human being or Socrates is a stone; [S3] if Socrates is a human being and Socrates is not a human being, then Socrates is not a human being; therefore [S4] if Socrates is a human being and Socrates is not a human being, then Socrates is a stone (*De Naturis Rerum* cixxiii, 288–89).

The outcome of the crisis provoked by Alberic was a complete change in the understanding of the logical connectives. John of Salisbury tells us that he could not conceive why any one would think that anything follows from an impossibility but according to Alexander Neckham nothing was more obvious.

Abelard had insisted that a genuine connection was required for the truth of conditionals and disjunctions. Alexander's argument, on the other hand, assumes only inseparability for the conditional and much less for the disjunction. [S2] is the so-called Principle of Addition characteristic of the disjunction defined as true if one of the disjuncts is true. The disjuncts are no longer required to be related as immediates.

The conditional and disjunction were standardly defined in this way for the rest of the middle ages. Until the end of the thirteenth century, however, a contrast continued to be drawn between an accidental consequence which held wherever the inseparability condition was met and a natural consequence in which the sense of the antecedent contained the consequent. This stronger connection was needed because it was necessary to reason about impossibilities.

THE RECEPTION OF THE *LOGICA VETUS* AND THE DEVELOPMENT OF THE *LOGICA MODERNORUM*

Some time towards the end of the twelfth century the various different schools disappeared as the independent masters formed themselves into the corporation that became the University of Paris. Teaching and research in logic was the preserve there of the Faculty of Arts and its results appear in the introductory textbooks of the *logica modernorum*. To the traditional topics these add extensive discussions of fallacies and the properties of terms.

FALLACIES. Although Abelard had some limited access to the *Sophistical Refutations* it was not until around 1140 that the analysis of fallacies became a major concern for logicians. From the beginning, however, a short list was available in Boethius's discussion of Aristotle's remark in *De interpretatione* 6, that the putative negation of a given proposition may fail to have the required opposite truth value because the subject or predicate terms have different meanings in the two propositions (De Rijk 1962–1967).

Although Boethius's list of the ways in which this might occur ceased to be of much interest once the *Sophistical Refutations* were easily available, one of his fallacies was particularly important for the later development of logic. With no further explanation Boethius gives as an example of what he calls *univocation* the propositions "*homo ambulat*" ("human being walks") and "homo non ambulat" ("human being does not walk"). He claims that they are true together when the first is true of an individual, or particular man, and the second is true of "special man."

Abelard notes that univocation arises because the context in which a term is used may affect its meaning. For example, since medieval Latin has no articles or quotation marks it cannot distinguish between the occurrences of "*homo*" in "*homo est albus*," "*homo est vox*," and "*homo est species*," in the way in which we distinguish in their translations between "a human is white," "'human' is a word," and "human is a species" (De Rijk 1962–1967, I, pp. 51–56).

Logicians in the second half of the twelfth century commented at length on and refined Aristotle's account

of fallacy in the *Sophistical Refutations*. By the end of the century the results of their work are clear in theology where the theory of fallacy is frequently invoked to explain and resolve errors in argumentation. In addition to the standard fallacies logicians also developed as a special form of argument the idea of counter instances (*instantiae*) which they found in the *Sophistical Refutations*, *Topics*, and *Prior Analytics*. With these the principles advocated by one or another of the schools were shown to lead to a conclusion which was unacceptable to it.

Once the works of the *logica nova* were available logicians seem to have turned their attention from the theory of consequences and topical inference to issues in philosophical semantics. Here a distinction was made between categorematic words, or terms, that is words which on their own can be the subject or predicate of a categorical proposition, and all other words which can occur in any kind proposition. The latter were called syncategorematic words.

THE PROPERTIES OF TERMS. Termist logic, so called because of its interest in the semantical properties of terms, seems to have developed in rather different ways in Paris and Oxford. The most famous Parisian termist was certainly Peter of Spain (c. 1205–1277), whose *Tractatus*, or *Summulae logicales*, written around 1235, was much commented on and remained the standard introductory text in logic in continental Europe and Scotland for the rest of the middle ages. It seems, however, not to have been greatly used in England, where the University of Oxford had its own textbooks. The *Introductiones in logicam* (c. 1245) by William of Sherwood (c. 1210–c. 1270) perhaps also belongs in the Oxford tradition. Another text belonging to the Parisian tradition is the *Summa Lamberti* (c. 1255) of Lambert of Auxerre (fl. 1250s) on which the following remarks are based.

IMPOSITION AND SIGNIFICATION. Medieval logicians developed their philosophical semantics in the first place from Boethius's commentaries on the first chapter of *De interpretatione*: Spoken words are introduced to bring to mind mental items, understandings (*intellectus*), which are obtained from the things which exist in the extra-mental world and are likenesses of them. For substantial common terms such as "human being" the corresponding understandings are the mental correlates of the forms which in the world make individuals to be the kinds of things that they are. For accidental terms such as "whiteness" they are the forms which cause individuals to have the accidental features that they do.

Words were held to acquire their meaning through acts of baptism, known as *imposition* (*impositio*), or *institution* (*institutio*) (Kretzmann et al. 1982, ch. 9). In the case of individual humans literally so. For general terms the impositor introduces a name in the presence of a paradigmatic sample with the intention that all and only individuals of the kind in question bear the same name. Adam's naming of the beasts of the field and the fowls of the air (Genesis 2:19) provided a suitable example. Although medieval accounts of imposition do not seem to have been very developed there are obvious similarities to modern causal theories of reference.

The immediate and proper signification of a common term is the understanding constituted when it uttered in the mind of a listener who speaks the language. Just what a given philosopher thought about the things understood and their relationship to individuals in the world depended on where he stood on the question of universals. Lambert, for example, was a realist. The term "human being," he claims, signifies immediately the understanding of the form which makes humans to be human and mediately the form itself. It does not signify individual human beings (*Logica*, 206).

SUPPOSITION. "Supposition" is used in the thirteenth century to refer to what earlier writers had called "appellation," it is a property which an already significant term has in virtue of its use. Corresponding to the three different contextual meanings recognized in the fallacy of univocation there are three forms of supposition. With no change in the signification established by its original imposition, the term "*homo*" thus supposits, or stands for three different kinds of things in the propositions "*homo est albus*," "*homo est vox*," and "*homo est species*."

In the first, according to Lambert, "*homo*" has personal supposition because it stands for the individuals "contained under" the form which it indirectly signifies. In the other two, he says, its supposition is simple (*Logica*, 209). In the second it stands for the thing which the term signifies indirectly—a form according to Lambert, and a "universal thing" according to Peter of Spain. In the third proposition the terms stands for itself.

William of Sherwood gives a slightly different classification. According to him in the third proposition "*homo*" has material supposition and in the other two formal supposition. In the first this formal supposition is personal and in the second it is simple (*Introductiones*, 75).

Personal supposition is the semantical property which most interested logicians since their task was to say

in general what determines the truth or falsity of a given proposition and to do so they needed to decide what the terms in the proposition stand for.

THE DIVISIONS OF SUPPOSITION. Treatises on the properties of terms make many distinctions and precisions within personal supposition. Supposition properly speaking is a property of a substantive noun which it has when it stands for something. An adjective in use, on the other hand, couples something and so is said to have the property of copulation.

Supposition in general, according to Lambert, is either natural and accidental. The imposition of a term connects it mediately with a form and, at a second remove, prior to any contextual determination to all the individuals which have done, do, or will share in that form. These are what it naturally supposits for (*Logica*, 208).

Accidental supposition is supposition determined by context and may, as noted, according to Lambert, be simple, or personal. Personal supposition is further divided into discrete supposition, the supposition had by proper names, and common supposition, the supposition of common terms.

The common supposition of a term such as "human being" is further determined by its interaction with the syncategorematic words of quantity and quality, and may be either determinate or confused. Logicians offered various accounts of these forms of supposition but by the fourteenth century typically explained them in terms of their inferential relations (Kretzmann et al. 1982, ch. 9)

Supposition is determinate when the term is the subject of an indefinite or particular affirmative, such as "a human being is running" and "some human being is running." Here we may descend from the particular or indefinite proposition to the propositional disjunction of singulars whose subjects are the *supposita* of the common term and ascend from any one of those singulars to the general proposition. So from "some human being is running" we may infer "Socrates is running or Plato is running or …" and from the truth of any one of the disjuncts we may infer that some human being is running.

In confused supposition, a common term stands for all its *supposita* together. It may do this in one of two ways, either as with the subject of a universal affirmative where the supposition is distributive, and one may descend to, and ascend from, the propositional conjunction of each of the corresponding singulars. For example from "every human being is running" to "Socrates is running and Plato is running and …" and conversely.

The other form of confused supposition, merely confused supposition, is exemplified by a common term occurring as the predicate of a universal affirmative proposition. Here the term again stands for all supposita but taken together in such way that one can descend only to the predicate disjunction but ascend from any singular. For example from "every human being is an animal" to "every human being is (this animal or that animal or …) and from "every human being is this animal" to every man is an animal.

Negation distributes any simple term to which it is applied, so both the subject and predicate of no man is running, that is, every man is not running, have confused and determinate supposition (Lambert, *Logica*, 210).

Historians have puzzled about the relationship between supposition theory and modern quantification theory but this seems to miss the point. Supposition theory does not aim to state truth-conditions for propositions but to determine which of the supposita of a term occurring in a proposition someone uttering it should be understood as referring to and in what way.

AMPLIATION AND RESTRICTION. The propositions given above to illustrate the divisions of supposition all have simple subjects and predicates with the verb in the present tense and not modified in any way. A term is said to appellate those of its *supposita* which actually exist and in the case of all these propositions appellation and supposition coincide. The qualification of a substantive with an adjective restricts its supposition to suitably qualified things. In "a white human being is running," for example, "human being" has determinate supposition only for those of its *appellate* which are white (Lambert, *Logica*, 226).

Tense affects the supposition of terms by ampliating them to stand for *supposita* other than their *appellata*, though these may also be included in the supposition. For example in "an old man was a young man" the predicate term has merely confused supposition for those of its *supposita*which existed in the past but do not now exist. The subject term has determinate supposition for its *appellate* and its past *supposita*.

There is no suggestion in the twelfth century termists named that a term might supposit for possibilia which never exist. Lambert and Peter of Spain hold, for example, that in the modal proposition "some man might be the Antichrist" "man" supposits for past and future men (Lambert, *Logica*, 228). Ampliation to pure possibilia is allowed, however, in the *Summa logicae* (c. 1324) of William of Ockham (c. 1285–1349) and *Summulae de*

dialectica (1330s) of John Buridan (c. 1300– c. 1360). The change in theory of ampliation reflects a radically new conception of possibility introduced in the work of John Duns Scotus (c. 1265–1308) at the beginning of the fourteenth century. Against the assumption that all possibilities must be realized in time Scotus famously argued for the logical possibility that things could now be otherwise than they in fact are and so that there are possibilities that are never realized.

SYNCATEGOREMATIC WORDS. Both Peter of Spain and William of Sherwood as well as other termist logicians produced treatises entirely devoted to syncategorematic words (Kretzmann et al. 1982, ch. 11). These treatises do not deal with all words that are not categorematic but only with a relatively small and fairly standard set. In addition to the definition by exclusion, syncategorematic words are further characterized as semantically incomplete in that they acquire a signification only by being combined in some way with categorematic terms. For this reason they are said to be *consignificant*.

It is in the treatises on *syncategoremata* that termist logicians deal with the difficult words whose presence may affect the validity of a principle of inference and allow the construction of sophisms. As, for example, in the proof by Sherwood that no man lectures at Paris unless he is an donkey: "A man lectures at Paris unless he is an donkey" is a false conditional since the antecedent "a man is not a donkey" is necessarily true and consequent may be false. Therefore the contradictory of the conditional is true (*Syncategoremata* 82–3). In the fourteenth century such puzzles and their resolutions were collected together in separate works devoted to grammatical, logical, including modal and epistemic, and physical sophisms. Their resolution often required that the inner structure of a syncategorematic term be exposed by what was called *exposition*. "Socrates is beginning to be white," for example, might be expound as 'Socrates is not now white and after now Socrates will be white' leading on to a discussion of tense, change, and the structure of time.

Included among the *syncategoremata* in these treatises we find the propositional connectives and confirmation the twelfth century insight into their nature had not been lost. William of Sherwood, for example, discusses both negation and the copulative conjunction. He clearly distinguishes, extinctive, or propositional negation and argues that if the conjunction "Socrates is running and Plato is arguing" is negated with a preposed particle the result is true just in case one of the coupled propositions is false (*Syncategoremata* 86).

MODISM

In the last quarter of the thirteenth century the termist semantics of supposition was replaced by what is known as modism, or speculative, that is, theoretical, grammar (Marmo 1994, Kelly 2002). The proponents of this theory, the *modisti*, for example Martin of Dacia (d. 1304), Boethius of Dacia (fl. 1275), and Thomas of Erfurt (fl. 1300) were concerned to say something more general about the meaning of both categorematic and syncategorematic terms than their termist predecessors. They held that all meaningful words are characterized by certain modes of signifying and that these correspond to the traditional parts of speech. Corresponding to each modes of signifying, is a mode of understanding, and a mode of being.

According to the modists a proper name like "Socrates" as well as signifying Socrates, carries information about the essential character of what it signifies. It signifies it as a substance, for example, in the *modus substantiae*, though not as an existent, since we use nouns to speak about presently non-existent and fictional items. A verb, on the other hand, signifies what it signifies in the mode of change and becoming. Grammatical features which were regarded as less fundamental, for example, number and tense, were held to correspond to accidental modes of signifying, understanding, and being (Kretzman et al. 1982, ch. 13).

On the basis of their distinction between modes the *modisti* developed an account of grammatical congruity—the modes have to fit together in the right way. They sought to go beneath the surface structure of their language to locate the underlying relationship between the components of propositions. Their idea was that the order required by Latin grammar did not properly represent the real relationships between the things signified. Though twelfth century logicians had already explored some of these ideas especially with regard to pronouns, the *modisti* deserve credit for being the first to attempt to develop a systematic theory of syntax.

Although the modists distinguished between the full signification of a word including its mode of signifying and the things in the world to which it applies, they made no use of the idea of supposition. They seem not to have developed an account of the contextual dependence of reference to compete with that of termists and in the end it was the semantics of termism which won the day (Kretzman et al. 1982, ch. 13).

OBLIGATIONS

The earliest treatises on what were known as obligations (*obligations*) date from the second half of the twelfth century (Martin 1993). In obligational disputation one participant, the respondent, is required to agree to a hypothesis and to reply consistently with it in the face of questions put to him by the opponent. The aim of the opponent is make the respondent contradict himself.

The most important form of obligation was the one known as *positio*, in which the opponent posits to be true something which is in fact false. In the twelfth and thirteenth centuries it had two forms depending on whether the *positum* was false but possibly true, possible *positio*, or an impossibility, impossible *position*. The original motivation for the latter seems to have been Boethius's proposal in *De hypotheticis syllogismis* that an impossibility be posited in order to see what follows (*Hyp. Syll.* I.2.6).

The earliest surviving treatise on impossible *positio*, the *Tractatus Emmeranus*, recognizes that no coherent argumentation is available under such an hypothesis if one accepts that anything follows from an impossibility. It stipulates instead that reasoning in impossible *positio* should rely only on consequences in which the consequent is contained in the antecedent and so not employ those with an affirmative antecedent and negative consequent—the theory uniquely characteristic of Abelard and the *Nominales* (De Rijk 1974). Later treatments of *impossible position* require only that they be conducted using consequences satisfying the containment condition.

In accounts of possible *positio* written before 1330s the respondent's answers are required to be consistent with everything that has gone before. He must thus concede a *propositum* which follows from the conjunction of the *positum* with all *proposita* already conceded and the contradictories of those which have been denied and deny a *propositum* whose contradictory follows from this conjunction. A *propositum* is irrelevant if neither it nor its contradictory follows from the conjunction and the respondent is required, if it is true, to concede it and, if it is false, to deny it (Kretzmann et al. 1982, ch. 16A).

A well conducted *positio* thus yields a set of propositions cotenable with the original *positum* and so an account of how the world might be. In treatises on possible *positio* written before the beginning of the fourteenth century we find a rule to the effect that if *n* is the present time, the *propositum* "*n* is the present time" must be denied, since it is not possible for things now to be other than now they are. Duns Scotus rejects this principle in setting out his new account of possibility and it is no longer found in fourteenth century accounts of *positio*.

Possible *positio* provides a way of testing the respondent's reasoning skills but also of constructing alternative possible world-histories. This application is common in fourteenth century treatments of reconciliation of divine foreknowledge with the possibility that things might be otherwise than they will be.

In the mid-1330s a group of logicians at Oxford proposed modifications to the principles of *position*. Richard Kilvington (c. 1305–1361) in his *Sophismata* required that the respondent answer an irrelevant positum not in accordance with his beliefs about its actual truth-value but rather in accordance with the beliefs he would have if the *positum* were true. Kilvington noticed that these may well differ if the *positum* refers to the respondent's epistemic states (Kretzmann et al. 1982, ch. 16B).

Roger Swineshead (d. 1356) went much further in his *Obligationes* (1340s?) and proposed what became known as the "new response" (Kretzmann et al. 1982, ch. 16B). For reasons which remain obscure he required the respondent simply to concede a *propositum* if it follows from the *positum* alone and to deny it if is incompatible. Everything else is irrelevant. This change, however, undermines the constructive character of *position* since, for example, if some false proposition *p* is posited and *q* is an irrelevant truth, the respondent must concede both *p* and *q* when they are proposed but go on to deny their conjunction *p and q*. Swineshead's account of *position* seems to have enjoyed some limited success but it is not mentioned after the end of the fourteenth century.

INSOLUBLES

The most famous example of what medievals called insolubles, sentences difficult but not impossible to solve, is the Liar: "This sentence is false" (Spade 1988). The difficulty is to assign it a truth-value since it seems that if it is true, then it is false, and if it is false, then it is true. The problem is first noticed the middle ages in the *Ars disserendi* of Adam of the Little Bridge published in 1132 and its medieval origins may well lie in reflection on possible *positio*.

Both the *Tractatus Emmeranus* and another treatise from the second half of the twelfth century, the *Obligationes Parisiensis* (De Rijk 1975), note that if a respondent accepts as a *positum* "the *positum* is false" or an equivalent, then the opponent will be able to force him to contradict himself (Martin 1993). Both works go on to discuss propositions such as "a falsehood is conceded"

which may be certainly be posited but cannot then be conceded as the rules of *position* require since if it is, it becomes a Liar. The appropriate response, they claim, is to reply "You are not saying anything" (*nugaris*).

The earliest known treatise entirely devoted to the Liar, the *Insolubilia Monacensis*, from roughly the same date, adopts the same solution, voiding (*cassatio*): A self-referential utterance of "this sentence is false" fails to assert anything (De Rijk 1966). This solution continued to be invoked in the thirteenth century but is no longer employed in the heyday of insoluble literature, the first half of the fourteenth century.

Many different solutions were proposed to the problem and Thomas Bradwardine (c. 1295–1349) lists eight others besides *cassatio* in his *Insolubilia* (Spade 1988). These include, for example:

1) *Secundum quid et simpliciter* (qualified and unqualified): Distinguish between the qualified and unqualified possession of a property as Aristotle does in the *Sophistical Refutations* discussing the puzzle of a man who takes an oath to break his oath. The Liar is false without qualification, but relatively true.

2) *Transcasus* (change of situation): The claim made in uttering the Liar refers to an instant before the utterance. The Liar is simply false since the speaker said nothing then.

3) *Restrictio* (restriction): The supposition of the term "false" in the Liar is restricted to standing only for sentences other than the Liar or sentences equivalent to it. Since uttering the Liar utters only that sentence, it is simply false.

Bradwardine rejected all the theories in his list and offered a new one which set the agenda for later discussions. He maintained, first, that a proposition is true if it signifies things *only* as they are but is false if it signifies things as other than they are—it may well also signify them as they are. Second, he held, and seems to have been the first to do so that a proposition signifies just what follows from it. Bradwardine concluded that if a proposition signifies itself to be false, then it signifies itself to be true. The Liar thus signifies itself to be both true and false and so is false (Roure 1970).

CONSEQUENCES

Treatises devoted to consequences seem to be product of the fourteenth century and, although one was written by the great Parisian logician John Buridan, they are almost exclusively a British production. The second or third decade of the fourteenth century marks a turning point in the history of consequences as important as the resolution of the twelfth-century crisis (Martin 2005).

Duns Scotus was not a logician but he put logic to the service of metaphysics when he located a formal distinction between any two items which are actually inseparable but conceptually separable. If being *B* follows accidentally but not naturally from being *A*, then being *A* is formally but not existentially distinct from being *A*.

Ockham's rejection of the formal distinction seems to explain his introduction of an entirely new theory of consequences. In his *Summa logicae* rather than distinguishing between natural and accidental consequences by appealing to *loci* which guarantee containment in contrast to those which do not, he takes basic logical distinction to be between what he calls material and formal consequences (*Sum. Log.* III.3.1).

All consequences must satisfy the inseparability requirement. Material consequences satisfy it merely in virtue of truth-values of the antecedent and consequent and so include all the paradoxical consequences. Formal consequences hold in virtue of there being a connection between antecedent and consequent guaranteed by a middle, another name for a *locus*. The middle, however, is required only to guarantee non-trivial inseparability.

There is thus no logical distinction between consequences for Ockham corresponding to that between natural and accidental consequences. It is replaced by an appeal to the epistemological notion of evidence but this does not partition the class of true consequences in the way the natural—accidental distinction does. Nor, more importantly, can it be used to argue for the formal distinction.

In an alternative classification of consequences Ockham invokes a distinction already made the thirteenth century to consequences which satisfy the Inseparability condition in virtue of the necessity of the present. He holds that if the conjunction *p and not:q* is now false but at some time will be true, the truth of the antecedent is now inseparable from that of the consequent and so *if p, then q* is a consequence *ut nunc* (as-of-now). If *p and not:q* is false at all times, past, present, and future, according to Ockham, *if p, then q* is a simple consequence (*Sum. Log.* III.3.1).

Ockham's new theory of consequences seems to have very rapidly supplanted the old one and natural consequences are not mentioned in logic texts after the first quarter of the fourteenth century. Nor for that matter is impossible *positio*.

While Ockham's examples of the middles which provide the guarantee of formal consequence are all what we would classify as formal in that they hold for all uniform substitution instances of terms, his practice indicates that some middles hold only for limited classes of terms. This possibility is absent in later writers such as Buridan who explicitly defines formal consequence in terms of the uniform substitution of any terms satisfying the inseparability conditions.

By the middle of the fourteenth century the logic of consequences is thus fully formal in the modern sense and treatises on the subject contain many of the rules recognized in classical modal propositional logic.

THE LOGIC OF MODALITY

While the *Prior Analytics* offered logicians nothing on categorical syllogisms not already available in Boethius what Aristotle had to say about modal forms was extremely problematic (Lagerlund 2000). The difficulty is that he accepts modal conversion principles such as accidental conversion: *if every A is necessarily a B, then some B is necessarily an A* but also claims that while *every B is necessarily C and every A is B; therefore every A is necessarily C* is valid *every B is a C and every A is necessarily B; therefore every A is necessarily C* is not. The conversion seems only to hold only if the modality is understood in the composite sense while the claim about the syllogisms requires the divided sense.

The first known medieval solution is found in the commentary on the *Prior Analytics* written Robert Kilwardby (1215–1279) in the 1240s (Thom 2003). Aristotle had designated as *per se* predications in which the subject contains the predicate and Kilwardby claims that modality may be uniformly construed in the divided sense if the conversion principles are restricted to those in which the antecedents are per se predications. Thus "every man is necessarily an animal" converts accidentally with "some animal is necessarily human' but "every literate (man) is necessarily a man" does not convert in this way with "every man is necessarily literate." Kilwardby thus makes just the distinction between modal claims that was made between natural and accidental consequences.

Ockham in his *Summa logicae* explores the relationship between divided and composite readings on the basis of his claim that these do not differ in the case singular propositions (Normore 1999). He derives syllogisms for composite modals by applying to categorical syllogism the principles of modal inference, for example "if the premisses are all necessary, then so is the conclusion." Ockham goes on to examine syllogisms formed with divided modals and with mixtures of both divided and composite (*Sum. Log.* III.1.20–46). He holds that divided claims are equivocal. Thus in "every *A* is possibly *B*," according to Ockham, the predicate is always ampliated by the mode but the supposition of subject may be understood to be only for what are now actually *A* or as ampliated for what can be *A*.

The most important development in syllogistics in the middle ages is in the work of Buridan. Buridan goes beyond Ockham in taking the theory of the syllogism to be simply an instance of the general theory of formal consequence (King 1985). He shows how the validity of the moods of the categorical syllogism can be proved from basic principles governing the semantics of general terms. The theory of modal syllogism with composite modality is, as with Ockham, quite straightforward. Buridan's treatment of divided modals is complex and of great interest since it reveals his attitude to the iteration of modalities and seems to commit him to the same principles as that of the modern system of strict implication known as S5.

Treatises on each of the subjects mentioned above continued to be produced through the fourteenth and fifteenth centuries by vast numbers of logicians. None of them, however, were of the stature of Abelard, Ockham, or Buridan, and originality in logic gave way at the end of the period to mere pedantry.

See also Abelard, Peter; Aristotle; Augustine, St.; Avicenna; Boethius, Anicius Manlius Severinus; Boetius of Dacia; Bradwardine, Thomas; Buridan, John; Burley, Walter; Cicero, Marcus Tullius; Conditionals; Duns Scotus, John; Gilbert of Poitiers; John of Salisbury; Kilwardby, Robert; Kilvington, Richard; Liar Paradox, The; Modality and Language; Peter of Spain; Porphyry; Proper Names and Descriptions; Propositions; Swineshead, Richard; Themistius; William of Ockham; William of Sherwood.

Bibliography

PRIMARY WORKS

Abaelard. Peter. *Dialectica*, edited by Lambertus De Rijk. Assen: Van Gorcum, 1970.

Abaelard. Peter. *Peter Abelard's Philosophische Schriften*, edited by B. Geyer. *Beiträge zur Geschichte der Philosophie des Mittelalters* 21, 1–4. Münster: Aschendorff: 1919–1927.

Abaelard. Peter. *Scritti di Logica*, edited by Mario Dal Pra. Florence: La Nuova Italia, 1969.

Abaelard. Peter. *Twelfth Century Logic II: Abaelardiana Inedita*, edited by Lorenzo Minio-Paluello. Rome: Edizioni Di Storia E Litteratura, 1958.

Boethius, Anicius Manlius. *Commentarium in Librum Aristotelis Perihermenias*. 2nd ed., edited by Karl Meiser. Leipzig: Tuebner, 1880.

Boethius, Anicius Manlius. *De Hypotheticis Syllogismis*, edited by Luca Orbetello. Brescia: Paideia, 1969.

Boethius, Anicius Manlius. *De Topicis Differentiis*. Translated with notes and essays on the text by Eleonore Stump. Ithaca, NY: Cornell University Press, 1978.

Buridan, Jean. *Sophisms on Meaning and Truth*. Translated by Theodore Scott. New York: Appleton-Century-Crofts, 1966.

Buridan, Jean. *Summulae De dialectica*. An annotated translation with an introduction by Gyula Klima. New Haven, CT: Yale University Press, 2001.

Burley, Walter. *On the Purity of the Art of Logic*. Translated by Paul Vincent Spade. New Haven, CT: Yale University Press, 2001.

Compotista, Garlandus. *Dialectica*, edited by Lambertus De Rijk. Assem: Van Gorcum, 1959.

De Rijk, Lambertus. *Logica Modernorum: A Contribution to the History of Early Terminist Logic*. 2 vols. Assen: Van Gorcum, 1962–1967.

De Rijk, Lambertus. "Some Notes on the Mediaeval Tract *De Insolubilibus*, with an Edition of a Tract Dating from the End of the Twelfth Century." *Vivarium* 4 (1966): 83–115.

De Rijk, Lambertus. "Some Thirteenth Century Tracts on the Game of Obligation I." *Vivarium* 12 (1974): 94–123.

De Rijk, Lambertus. "Some Thirteenth Century Tracts on the Game of Obligation II." *Vivarium* 13 (1975): 22–54.

Hughes, George. *John Buridan on Self-Reference: Chapter Eight of Buridan's Sophismata Translated, with an Introduction and a Philosophical Commentary*. Cambridge, MA: Cambridge University Press, 1982.

John of Salisbury. *Metalogicon*, edited by D. McGary, J. Hall, and K. Keats-Rohan. Turnhout: Brepols, 1991.

King, Peter. *John Buridan's Logic: The Treatise on Supposition, the Treatise on Consequences*. Dordrecht: Reidel, 1985.

Kilvington. Richard. *The Sophismata of Richard Kilvington*. Introduction, translation, and commentary by Norman Kretzmann and Barbara Ensign Kretzmann. Cambridge, MA: Cambridge University Press, 1990.

Lambert of Auxerre. *Logica (Summa Lamberti) a cura di Franco Alessio*. Firenze: La Nuova Italia, 1971.

Neckham, Alexander. *De Naturis Rerum*, edited by Thomas Wright. London: Longman, 1863.

Peter of Spain. *Syncategoreumata*. Translated by Joke Spruyt. Edited by Lambertus De Rijk. Leiden: Brill, 1992.

Peter of Spain. *Tractatus, called afterwards Summule logicales*, edited by Lambertus De Rijk. Assen: Van Gorcum, 1972.

Roure, Marie Louise. "La problématique des propositions insolubles au XIIIe siècle et au début du XIVe, suivie de l'édition des traités de W. Shyreswood, W. Burleigh et Th. Bradwardine." *Archives d'histoire doctrinale et littéraire du moyen age* 37 (1970): 205–326.

William of Ockham. *Summa Logicae*, edited by Philotheus Boehner, Gedeon Gál, and Stephan Brown. In *Opera Philosophica*. Vol. 1., edited by Guillelmi de Ockham. St. Bonaventure: The Franciscan Institute, 1974.

William of Sherwood. "De Introductiones in Logicam des Wilhelm von Shyreswood." *Sitzungsberichte der Bayerischen Akademie der Wissenschften, Phil.-hist. Abteilung* Heft 10, Munich, 1937.

William of Sherwood. "The Syncategorematat of William of Sherwood," edited by J. Reginald O'Donnell. *Medieval Studies* 3 (1941): 47–93.

William of Sherwood. *William of Sherwood's Introduction to Logic*. Translated by Norman Kretzmann. Minneapolis: University of Minnesota Press, 1966.

William of Sherwood. *William of Sherwood's Treatise on Syncategorematic Words*. Translated by Norman Kretzmann. Minneapolis: University of Minnesota Press, 1968.

SECONDARY WORKS

Bochenski, Innocentius. *A History of Formal Logic*. Edited and translated by Ivo Thomas. Notre Dame, IN: University of Notre Dame Press, 1961.

Boh, Ivan. *Epistemic Logic in the Later Middle Ages*. London: Routledge, 1993.

Kelly, Loius G. *The Mirror of Grammar: Theology, Philosophy, and the Modistae*. Amsterdam: Benjamins, 2002.

Kneale, William, and Martha Kneale. *The Development of Logic*. Oxford: Clarendon Press, 1984.

Knuutila, Simo. *Modalities in Medieval Philosophy*. London: Routledge, 1993.

Kretzmann, Norman, Anthony Kenny, and Jan Pinborg. *The Cambridge History of Later Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1982.

Kretzmann, Norman, Anthony Kenny, and Jan Pinborg. *Meaning and Inference in Medieval Philosophy: Studies in Memory of Jan Pinborg*. Boston: Kluwer Academic Publishers, 1988.

Kretzmann, Norman, and Eleonore Stump, eds. *Logic and the Philosophy of Language*. Cambridge, U.K.: Cambridge University Press, 1988.

Lagerlund, Henrik. *Modal Syllogistics in the Middle Ages*. Leiden: Brill, 2000.

Martin, Christopher J. "Embarrassing Arguments and Surprising Conclusions in the Development of Theories of the Conditional in the Twelfth Century." In *Gilbert de Poitiers et ses contemporains*, edited by Jean Jolivet and Alain de Libera. Naples: Bibliopolis, 1987.

Martin, Christopher J. "Logic." In *The Cambridge Companion to Abelard*, edited by Jeff Brower and Kevin Guilfoy. Cambridge, U.K.: Cambridge University Press, 2004.

Martin, Christopher J. "Formal Consequence in Scotus and Ockham: Towards an Account of Scotus' Logic." In *Duns Scot à Paris 1302–2002*. Brepols: Turnhout, 2005.

Martin, Christopher J. "The Logic of Negation in Boethius." *Phronesis* 36 (1991): 277–304.

Martin, Christopher J. "Obligations and Liars." In *Sophisms in Medieval Logic and Grammar*, edited by Stephen Read. Dordrecht: Kluwer, 1993.

Marmo, Costantino. *Semiotica e linguaggio nella scolastica: Parigi, Bologna, Erfurt, 1270–1330*. La semiotica dei modisti. Roma Istituto storico italiano per il Medio evo, 1994.

Moody, Ernest. *The Logic of William of Ockham*. London: Sheed & Ward, 1935.

Moody, Ernest. *Truth and Consequence in Medieval Logic*. Amsterdam: North-Holland Publishing Co., 1953.

Normore, Calvin. "Some Aspects of Ockham's Logic." In *The Cambridge Companion to Ockham*, edited by Paul Spade. Cambridge, U.K.: Cambridge University Press, 1999.

Pinborg, Jan. *Logik und Semantik im Mittelalter: ein Überblick*. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1972.

Pinborg, Jan. *Medieval Semantics: Selected Studies on Medieval Logic and Grammar*. London: Variorum Reprints, 1984.

Spade, Paul. *Lies, Language, and Logic in the Later Middle Ages*. London: Variorum Reprints, 1988.

Thom, Paul. *Medieval Modal Systems: Problems and Concepts*. Aldershot: Ashgate, 2003.

Yrönsuuri, Mikko, ed. *Medieval Formal Logic: Obligations, Insolubles, and Consequences*. Dordrecht: Kluwer Academic Publishers, 2001.

Christopher J. Martin (2005)

INTERREGNUM (BETWEEN MEDIEVAL AND MODERN)

The interregnum between medieval scholastic logic and modern mathematical logic may be taken as having begun about the middle of the fifteenth century. There is no clear mark of division; the change was a shift away from the characteristic interests of the twelfth to the fifteenth century, with nothing of comparable importance arising to take their place. At the same time, certain less desirable trends in scholastic logic were perpetuated. The result is that formal logic was reduced almost entirely to a very imperfectly presented syllogistic. Medieval influences continued to operate in the early years of the sixteenth century, and medieval authors were still sometimes read in the seventeenth, but by the time that William of Ockham's *Summa Logicae* was printed at Oxford in 1675, no one had written creatively in the idiom of scholastic logic for many years.

The interregnum was characteristically sterile, a cause for despondency when one thinks of the large place logic continued to occupy in the educational curriculum and of the innumerable writers who put manuals of logic on the market. The tendency to publish at all costs was encouraged by the post-Reformation and post-Tridentine growth of universities, colleges, and seminaries.

VALLA

The first author to consider is the humanist Lorenzo Valla (1407–1457), best remembered for his writing on the forged donation of Constantine. In his *Dialecticarum Libri Tres* (1441), Valla gave no definitions of syllogistic figures and moods, evidently assuming that the reader would know about these. His aim was to confine the syllogistic to the first two figures, without the five moods of Theophrastus and Eudemus. To do this he would have had to reject subalternation, conversion, and reductio ad absurdum. About subalternation he was inconsistent; conversion he rejected as lacking brevity, ease, pleasantness, and utility; reductio ad absurdum he largely neglected. The five offending moods were called "Agrippine births," and of them all the most monstrous was "Frisemomorum, forsooth!"

Here we see the common humanist objection to the barbarity of scholastic terminology, but of course Valla was not objecting merely to comparatively recent Scholastics. His fullest invective was saved for the six moods of the third figure, which he thought insane and never found in use, unlike the first-figure and second-figure moods, which he accepted as dictated by nature to everyone, "even peasants, even women, even children." The standard means of reduction are but "remedies for sick syllogisms." The standing of the third figure would remain a point of dispute for a hundred years, until Ramus undercut Valla's argument by declaring that the figure was in obvious fact very commonly used (*Institutionum Dialecticarum Libri Tres*, Paris, 1554). Thus, Philipp Melanchthon (*Compendiaria Dialectices Ratio*, Basel, 1521) could not make up his mind on the subject.

MELANCHTHON

In Melanchthon (1497–1560), a most influential writer, the rhetorical approach to logic already appeared at a high state of development, although he retained some Aristotelian doctrine. The rhetorical tradition, derived from Cicero and Quintilian, had a place, albeit a very subordinate one, in scholastic logic. We can see it beginning to predominate in the *Dialectica ad Petrum de Medicis* (edited by D. M. Inguanez and D. G. Muller, Monte Cassino, 1943; composed about 1457), by Joannes Argyropulos, who held that the detail of the theory of *suppositio*, which was the distinctive and most original scholastic contribution to logic, offered almost nothing to oratorical practice.

Thus, scholastic logic, which in its origins had borrowed considerably from grammar, began to yield to the third member of the trivium, rhetoric. Accordingly Melanchthon declared the fruit of dialectic to be the ability to speak with propriety and exactness on any theme, and he expounded the Ciceronian syllogism, with its five parts—*propositio, approbatio, assumptio, assumptionis approbatio*, and *complexio*—before the Aristotelian. (A century later a similar five-part syllogism, with proposition, reason, example, application, and conclusion, came into favor in the New Nyāya school of Indian logic.) In

general, Melanchthon said, the natural reasoning common to the learned, children, and ordinary people is to be preferred to the "rancid commentaries of dialecticians." From this time on it was often felt desirable to include comparative lists of terminology, ancient and modern, as was done by a commentator on Rodolphus Agricola in 1538, by John Seton in 1572, and by John Sanderson in 1589.

RAMUS

The syllogistic as a deductive system underwent considerable attrition in the rhetorical treatment of logic, but this cannot be ascribed exclusively to the new interests. John Dolz's *Sillogismi* (Paris, 1511), a work of purely scholastic inspiration, methodically examines arguments in the different moods and figures as though they had nothing to do with one another. Dolz gave thirty-two sets of objections to Barbara before going on to Celarent "to avoid prolixity." Although logic applied to itself was by no means unknown in Scholasticism, the idea of a closed logical system was little developed, and hence the piecemeal treatment so characteristic of the scholastic *sophismata* was easily extended to encroach on the systematic character of syllogistic. The fact that Aristotle began by presenting syllogisms in lists probably also contributed to this encroachment.

The process of fragmentation was given new impetus by Pierre de la Ramée (Peter Ramus, 1515–1572). This great master of Latin rhetorical style and innovator of educational theory developed a massive attack on the Aristotelian tradition in logic and an alternative corpus of logical material that quickly gave rise to a widespread Ramist scholasticism.

ATTACK ON ARISTOTELIAN TRADITION. Ramus's *Animadversiones Aristotelicae* (Paris, 1556) tells in twenty books how Ramus turned from the clarity of Plato to the comparative chaos of Aristotle. Pretending to be analytical, Aristotle was almost completely deficient in that (Ramist) analysis that consists in systematic definition and division, and his doctrines are not supported by examples (are not, in fact, established by rhetorical syllogisms!). These are the standards Ramus applied as he worked through the *Prior Analytics* in his Book VII, firing off a broadside at every detail of Aristotelian or scholastic doctrine that occurred to him on the way. The typically rhetorical teaching that experience, observation, and usage are the proper guides in logic is prominent. Variables seldom make their appearance in this milieu, but Ramus's express attack on abecedarian examples—which, being examples of nothing, can be adapted to nothing—is remarkable.

RAMIST LOGIC. The *Dialecticae Libri Duo* (Paris, 1556) is divided between invention, or discovery, and judgment, a distinction derived immediately from Agrippa and mediately from Cicero and Boethius. This distinction had been recalled among Scholastics—for example, at the opening of Kilwardby's popular thirteenth-century commentary on the *Prior Analytics*, often printed under the name of Giles of Rome. Like Descartes, whose methodological ideas supplanted his own, Ramus could not escape his antecedents. The first book covers topics, or loci; the second expounds the Ramist syllogistic, divided into the contracted syllogism (an enthymematic version of the Aristotelian third figure) and the explicated syllogism (comprising the second and first figures, in that order). There are no signs of quantification, all unquantified propositions that are not singular being deemed universal. A mood is general if it contains no singular term, special if it contains one, and proper if it contains two. Examples are taken from classical rhetoric and poetry; the propriety of such sources was vigorously attacked by a little-known anti-Ramist, Thomas Oliver of Bury, in his *De Sophismatum Praestigiis Cavendis* (Cambridge, U.K., 1604), on the ground that logic has very little place in poetry or forensic oratory.

This whole early version of an ordinary-language approach to logic was admirably countered by Gisbertus Isendoorn (*Cursus Systematicus*, Oxford, 1658). Writing directly against the famous Cambridge Ramist George Downame, Isendoorn said (p. 613): *Observa … orationem et popularem discurrendi usum non esse mensuram et normam Logicae, sed rectam rationem et accuratam artem viamque concludendi* (Mark that popular speech and usage are not the standard and norm of logic, but right reason and an exact method of reaching conclusions).

MANUALS OF LOGIC

With all the effort of the mid-sixteenth century to simplify logic, it is not surprising that vernacular manuals began to appear, although sparsely, at that time. In England there were Thomas Wilson's *The Rule of Reason* (London, 1551), Ralphe Lever's *The Arte of Reason rightly termed Witcraft* (London, 1573), Abraham Fraunce's *The Lawiers Logike* (London, 1588), and Thomas Blundevile's *The Arte of Logicke* (London, 1599); in France there was Philippes Canaye's treatise *L'organe* (Paris, 1589). Little further seems to have been published in English until John Newton's *The English Academy* (London, 1677).

Wilson's pioneer effort is interesting chiefly for its novel terminology; for example, the major, minor, and middle terms are called the "terme at large," the "severall terme," and the "double repeate." Blundevile introduced an arithmetical syllogism and used a catechetical method. This method had been used by Matthias Flacius Illyricus in *Paralipomena Dialectices* (Basel, 1558; composed 1550), which gives a very detailed treatment of the venerable *pons asinorum*. Canaye's book was also devoted largely to the *pons asinorum*, being distinguished by the dissection of the traditional rectangular figure into two circular ones. The same subject had been dealt with in Christopher Corner's *Ratio Inveniendi Medium Terminum* (Basel, 1549), which set a new standard of scholarship by appending a Greek text of relevant chapters of Aristotle. Thus, Aristotelian subjects were being pursued, in somewhat new ways, at the same time that the widespread Ramist innovations were taking hold.

Something of the same development can be seen in commentaries on the *Prior Analytics*, from the sixteenth-century editions of Kilwardby, through the work of Lefèvre d'Étaples (Faber Stapulensis), with his emphasis on tabular presentation; that of Agostino Nifo (Niphus Suessanus), who professed to follow the Greek commentators but wrote a long treatise on conversion in the scholastic manner; Burana's urbane commentary, with lengthy appendixes by his teacher Bagolinus and an interesting prefatory glimpse of the logical curriculum in a north Italian university; Monlorius's commentary, relatively brief but careful; to that of Pacius, with its businesslike presentation, schemes, and figures, a work praised by Sir David Ross in his own commentary. Within this developing tradition of Aristotelian scholarship we may also put the *Apparatus Syllogistici Synopsis* of Joannes Albanus (Bologna, 1620), which elaborately examined the crescent-shaped and triangular diagrams that descended from Greek sources to the Aristotelians of the Renaissance.

In a field in which syllogistic occupied so large a place one must note widespread incompetence in the matter of classification by figure. This is, of course, a point settled by definition, as Lorenzo Maiolo (*Epiphyllides in Dialecticis*, Venice, 1497) and John Wallis (*Institutio Logicae*, Oxford, 1687) saw. These two were exceptional, however. Franciscus Titelmans (*De Consideratione Dialectica Libri Sex*, Paris, 1544) found the distinction between major and minor premises a hard thing for youths; Richard Crakanthorp (*Logicae Libri Quinque*, London, 1622) omitted the fourth figure without rejecting it and found it hard to determine the number of moods. The basic trouble was that the later medievals, following a lead given by Boethius, defined the major premise as the first stated, the major term as the extreme therein, and so on, whereas Philoponus had defined the major term as the predicate of the conclusion, the major premise as the premise containing the major term, and so on. Each of the schemes can be worked out consistently, but they give different classifications and arc mutually incompatible. This was seldom understood; it was a common fault to speak of indirect conclusions in connection with Philoponian definitions or to define with Philoponus and then take, for example, Balnama as fourth figure, instead of first figure with transposed premises.

In the Oxford logicians one does not find twenty-four moods in four figures correctly worked out on a Philoponian basis until Henry Aldrich (*Artis Logicae Compendium*, Oxford, 1691; this first edition was anonymous). The principles of the matter remained so little understood that even Augustus De Morgan (*Formal Logic*, 1847) could say, "Consider the fourth and first figures as coincident and the arbitrary notion of arrangement by major and minor vanishes," and W. S. Jevons (*Elementary Lessons in Formal Logic*, 1876) described fourth-figure syllogisms as ill arranged and imperfect and unnatural in form. "Unnatural" as a description of fourth-figure syllogisms was first used by Averroes, and his opinion was reinforced by Giacomo Zabarella (1533–1589); both meant to make a point of genuine formal logic, but they used some phrases that permitted a psychological interpretation. Sir William Hamilton's treatment of the matter (*Lectures on Logic*, 1860, Vol. IV), with lists of authors for and against the fourth figure and indirect moods of the second and third, is useless without knowledge of these authors' definitions and therefore of what they were favoring or opposing. A writer of a very different style was John Hospinianus (1515–1575), who proceeded on a combinatory basis and found that by admitting singular and indefinite propositions to the syllogistic and by identifying certain moods, he could obtain thirty-six valid moods out of a possible 512.

Extremely influential on manuals of the eighteenth and nineteenth centuries was *Logique, ou l'art de penser* (1662; *The Port-Royal Logic*), by Antoine Arnauld and Pierre Nicole. Even Aldrich, who disliked its novel terminology and Cartesian standpoint, may well have been prompted by it to his strict deductive treatment, for he shows no acquaintance with any other likely influence. The authors' epistemological interests certainly contributed much to the psychologism that was soon to infect logic, but such headings as conception, judgment,

and reasoning were not new in promoting this tendency. Canaye had already spoken of syllogism as the third operation of the mind, which *leaves* the premises and *arrives* at the conclusion. Such terminology is symptomatic of a change that occurred in the mid-seventeenth century. The Port-Royal section on method—a most popular subject in this period—more explicitly opened the way to the discursive excesses that would soon masquerade as logic, culminating, perhaps, in Henry Kett's *Logic Made Easy, or A Short View of the Aristotelic System of Reasoning, and Its Application to Literature, Science, and the General Improvement of the Mind* (Oxford, 1809).

A book praised by Leibniz and rather above the average, although not completely out of the common rut, is the *Logica Hamburgensis* (Hamburg, 1638), by Joachim Jung, or Jungius. One notable feature of this book is the marking of the lines of a syllogistic demonstration by letters, which are then used as references for showing by what principles which line follows from which others. Such a rather exact method of proof was very exceptional in logic before modern times, but contemporaneously with Jung, Pierre Hérigone introduced a similar method in mathematics (*Cursus Mathematicus*, Paris, 1634–1637). Jung was thoroughly acquainted with the possible use of contraposition as a means of syllogistic proof but was no more successful in his discussion of the fourth figure than so many others had been. Under the medieval heading of consequences he noted the argument *a recto ad obliquum*, which can be found in Aristotle's *Topics* II, 8, 114a18.

Some considerations, usually brief, of such standard medieval subjects as consequences and supposition theory continued to appear—for instance, those of Chrysostom Javellus (*Compendium Logicae*, Lyons, 1580), Robert Sanderson (*Logicae Artis Compendium*, Oxford, 1618), and Henry Aldrich—but these were exceptions. Arnold Geulincx hoped to repopularize such treatises by his *Logica Fundamentis Suis a Quibus Hactenus Collapsa Fuerat Restituta* (Leiden, 1662). He was able to relate alternation, conjunction, and negation by means of their truth conditions according to the laws that are often called after De Morgan or William of Ockham but that go back, at least in part, to the *Syncategoremata* of Peter of Spain. These laws were also known to the mathematician Gerolamo Saccheri, whose *Logica Demonstrativa* (Turin, 1697) is outstandingly original in its high degree of organization, its reflections on the assumptions necessary to logic, and its use of indirect proof, in the pattern of the so-called *mirabilis consequentia*, to the effect that what follows from its own negation is true. Unfortunately the few signs of revival and advance discernible at the close of the seventeenth century did not produce any general or permanent result, and even the work of Leibniz met with little response.

See also Agrippa; Aristotelianism; Aristotle; Arnauld, Antoine; Averroes; Boethius, Anicius Manlius Severinus; Cicero, Marcus Tullius; De Morgan, Augustus; Descartes, René; Geulincx, Arnold; Giles of Rome; Hamilton, William; Jevons, William Stanley; Jungius, Joachim; Kilwardby, Robert; Leibniz, Gottfried Wilhelm; Melanchthon, Philipp; Nicole, Pierre; Ramus, Peter; Theophrastus; Valla, Lorenzo; William of Ockham.

Bibliography

INTERREGNUM

Dürr, Karl. "Die Syllogistik des Johannes Hospinianus (1515–1575)." *Synthese* 9 (1952): 472–484.

Hamilton, William. "Logic: The Recent English Treatises on That Science" (1833). In *Discussions on Philosophy and Literature, Education and University Reform*, 116–174. London and Edinburgh: Longman, Brown, Green and Longmans, 1852.

Howell, W. S. *Logic and Rhetoric in England, 1500–1700*. Princeton, NJ: Princeton University Press, 1956.

Ong, W. J. *Ramus, Method, and the Decay of Dialogue*. Cambridge, MA: Harvard University Press, 1958.

Risse, Wilhelm. *Die Logik der Neuzeit*, Vol. I (1500–1640). Stuttgart and Bad Cannstatt: Frommann, 1964.

Thomas, Ivo. "The Later History of the Pons Asinorum." In *Contributions to Logic and Methodology in Honor of J. M. Bocheński*. Amsterdam: North-Holland, 1965.

Thomas, Ivo. "The Liar: A New Historical Detail." *Notre Dame Journal of Formal Logic* 6 (1965): 201–208.

Thomas, Ivo. "Medieval Aftermath: Oxford Logic and Logicians of the Seventeenth Century." In *Oxford Studies Presented to Daniel Callus*. Oxford: Clarendon Press, 1964.

Thomas, Ivo. "The Setting of Classical Logic." *Notre Dame Scholastic* 101 (1960): 16–17.

Ivo Thomas (1967)

PRECURSORS OF MODERN LOGIC

Modern logic, or the logic that is loosely called "mathematical," began in a serious and systematic way with Augustus De Morgan's *Formal Logic* and George Boole's *Mathematical Analysis of Logic*, both published in 1847.

But a number of earlier writers were already "modern" in spirit, and of these, four stand out especially sharply—Leibniz, Euler, Lambert, and Bolzano.

See also Bolzano, Bernard; Boole, George; De Morgan, Augustus; Lambert, Johann Heinrich; Leibniz, Gottfried Wilhelm.

(A. N. P.)

LEIBNIZ

Gottfried Wilhelm Leibniz (1646–1716) was distinguished in many fields, but in none more than in logic. There, however, his worth was not fully appreciated until the twentieth century. He early began to investigate Aristotelian syllogistic and never completely escaped from the syllogistic point of view. In 1666 he wrote a *Dissertatio de Arte Combinatoria*, a juvenile work that was not free of mistakes, as he later realized, but that showed a new, high sense of organization and a genuine feeling for formal logic, very rare at the time. In one part of this book Leibniz worked out for himself the calculations of Hospinianus (1560) relative to the possible and the valid moods of syllogism. He differed from Hospinianus in making singular propositions equivalent to universal ones, as did Wallis and Euler. He arrived at twenty-four strictly Aristotelian syllogisms, six in each of four figures, which he arranged in a neat tableau suggestive of certain deductive relationships. Leibniz's standard method of proof in this context was reductio ad absurdum, as suggested to him by his teacher Jakob Thomasius (1622–1684), author of *Erotemata Logica* (Leipzig, 1670), but he also recognized the need for conversion. He wrongly credited Ramus with a method actually known in the thirteenth century, the device of proving laws of conversion and subalternation by means of syllogism and the laws of identity "All *a* is *a*" and "Some *a* is *a*."

Leibniz often returned to syllogistic and was periodically vexed by semantic considerations, namely whether to think of the matter in extension or in intension—whether in "All *a* is *b*" it is the *a*'s which are said to be contained in the *b*'s or the property *a* which contains the property *b*. Leibniz had something of a fixation on the intensional approach, although he often suspected that extension was more effective and logically satisfactory. One thing that pushed him in the direction of extensionality was a fondness for experimenting with spatial interpretations. Thus, we find several attempts at diagrammatic representation, some using ruled and dotted lines and some using circles. He found it impossible to carry through such interpretations when thinking in intension.

THEORY OF COMBINATIONS. The theory of combinations is highly relevant to logic. Chrysippus is reported to have shown some interest in combinations, Kilwardby and others in the thirteenth century repeatedly made combinatory summaries of assertoric and modal syllogistic, and semantic interpretations of logical formulas in finite domains employ the theory. Besides the syllogistic computations described, Leibniz considered how many predicates can be truly asserted of a given subject or how many subjects set under a given predicate. Such problems need some preliminary arrangements, and Leibniz supposed that a composite concept is analyzable into a number of ultimate simples, just as an integer is uniquely decomposable into its prime factors. Correlating the simple concepts with prime numbers, we can say that a predicate is truly attributable to its subject if the product associated with the predicate divides that associated with the subject. The essentials of this idea have been used in modern times to obtain a decision procedure for syllogistic, and unique decomposition into primes plays an essential part in Gödel numbering.

UNIVERSAL LANGUAGE. The idea of decomposing concepts into "prime factors" suggested to Leibniz the possibility of following up the initial steps toward a universal language taken by John Wilkins (1668), Jean Joachim Becher (1661), George Dalgarno (1661), Athanasius Kircher (1663), and others. He wanted such a language not merely to be practically or commercially useful, as were many of the pioneer efforts, but to be logically constructed so as to have general scientific import. Leibniz later distinguished a universal language from a logical calculus and desired to base his language on a thorough analysis of the communicative function of the various parts of speech, tenses, suffixes, and so on (an anticipation of modern theories of syntactical categories), and at one point (*Analysis Linguarum*, 1678) he envisaged a basic Latin rather in the style of C. K. Ogden and I. A. Richards's basic English.

In saying that nouns express ideas and verbs express propositions Leibniz radically altered the Aristotelian basis of the distinction and gave, in germ, the concept of a propositional function. Such reflections led him to a reductionist program, with adverbs reduced to (derived from) adjectives and adjectives to nouns, and with the copula taken as the only fundamental verb. He recognized that particles, connectives, and prepositions are of especial importance to linguistic structure. In taking us out of the syllogistic area this theory recalls the medieval doctrine of syncategorematic terms and Thomas Aquinas's analysis of many prepositions, while it adumbrates the logic of truth-functional connectives and of relations. Leibniz knew that not all arguments are syllogistic, in this

matter acknowledging a debt to Jung, but the dominance of a syllogistic point of view in Leibniz's thought is shown by his curious distinction between syllogistic and "grammatical" consequences.

This part of Leibniz's thought constitutes a distinct chapter in the history of the relations between grammar and logic. Grammar had been influential in the constitution of scholastic logic, but in the interregnum it had yielded to the third member of the medieval trivium, rhetoric, as a dominant power. In the projects for a universal and rational language we see grammar reasserting itself. But Leibniz was not content to confine logic to the "trivial" arts.

LOGICAL CALCULUS. The idea that logic might be quadrivial, and notably mathematical, was not new with Leibniz. Leibniz considered Aristotle to have been, in his logic, the first to write mathematically outside mathematics (letter to Gabriel Wagner, 1596). Roger Bacon (thirteenth century)—who also wished to reduce the trivial art of grammar to the quadrivial one of music—stated in his *Opus Maius* that "all the predicaments depend on the knowledge of quantity, with which mathematics deals, and therefore the whole of logic depends on mathematics." It is in the light of this that one should read the statement in his *Communia Mathematica* that "the mere logician cannot accomplish anything worthwhile in logical matters" (*nihil dignum potest purus logicus in logicalibus pertractare*). William of Ockham had been of the opposite opinion, and in *De Sacramento Altaris* he described mathematicians as among those less skilled in logic. Ramón Lull had written a combinatorial work, *Ars Magna* (which captured Leibniz's imagination, though he soon came to understand its deficiencies), and Thomas Hobbes had elaborated suggestively, if ineffectively, on the theme "by ratiocination I mean computation" ("Computatio Sive Logica," in *De Corpore*).

There is little doubt, however, that Leibniz's ideas, which far outstripped in detail and understanding any earlier hints, were his own spontaneous creation. "While I was yet a boy with a knowledge only of common logic, and without instruction in mathematics, the thought came to me, I know not by what instinct, that an analysis of ideas could be devised, whence in some combinatory way truths could arise and be estimated as though by numbers" (*Elementa Rationis*). He was thereafter constantly occupied with such notions and attempted to contrive an alphabet of thought, or *characteristica universalis*, which would represent ideas in a logical way, not things in a pictorial way, and would be mechanical in operation, unambiguous, and nonquantitative; this alphabet of thought would be a means of discovery, a support to intuition, and an aid in ending disputes.

Leibniz regarded his great invention of the infinitesimal calculus (1675) as emerging from such researches, and the calculus led him to reflect still more intently on the properties desirable in such a characteristic. Exactly what he meant by "mechanical" and "calculation" is still in question, and he no doubt underestimated the task he set himself, but the imaginative fervor with which he always wrote of it reveals, as we can now appreciate, a true prophetic instinct. He often used an image from mythology to summarize his intentions, saying that his method was to be a *filum Ariadnes*, a thread of Ariadne. Many authors had long envisaged logic as a Cretan maze in need of such a clue—and that this should be so in an age when logic was scarcely existent does them little credit—but from the pen of Leibniz the allusion was more than a literary elegance and condensed a program of "palpable demonstrations, like the calculations of arithmeticians or the diagrams of geometers." (For Leibnizian references to the *filum*, see Louis Couturat, *La Logique de Leibniz*, pp. 90–92, 124; for other authors, see Ivo Thomas, "Medieval Aftermath.")

ENCYCLOPEDIA. One may ask what the theory of combinations was meant to combine, what the logical calculus was meant to calculate with, or where the analyses presupposed by the unified language of science were to be found. Leibniz was not content to leave such analysis in the state of a general project. The enormous range of his knowledge and interests, which included unity in religion, international relations, cooperation among scientists and scholars, and jurisprudence, as well as the not unrelated ordering of thought, prompted his lasting interest in the construction of an encyclopedia. T. Zwinger's *Theatrum Vitae Humanae* (1565) and Johann Heinrich Alsted's *Encyclopaedia* (1608) provided Leibniz with a basis for early schematisms, and sketches and fragments from about 1668 to the end of his life show an unceasing interest in the plan, which he believed had failed of completion through his own distractions and the lack of younger assistants. Appeals to monarchs and to learned societies met with little response. The project was, of course, a gigantic one, impossible of immediate fulfillment, but it should not be supposed that Leibniz thought it could be perfected quickly. Rather, its elaboration was to proceed gradually, along with that of the universal language and a calculus of logic. In later drafts this calculus took an ever more prominent place.

STRUCTURE OF THE CALCULUS. The main stages (1679, 1686, 1690) of Leibniz's many experiments in logical algebra have often been expounded and commented on. Here only some laws which were constant features will be mentioned.

(1) *a* is *a*;

(2) If *a* is *b* and *b* is *c*, then *a* is *c*.

Propositions of the form "*a* is *b*" are intended as universal affirmatives, "All *a* is *b*," which Leibniz normally thought of as meaning that the property *a* contains the property *b*. Sometimes he wrote "*a* contains *b*" instead of "*a* is *b*." Accordingly, rule (1) is one of the syllogistic laws of identity which, as was said above, he used from the start in syllogistic demonstrations, and rule (2) is the Barbara syllogism. Today we know that by means of the calculus of quantifiers and some definitions all asserted laws of the syllogistic can be obtained from rules (1) and (2) alone. Leibniz lacked those aids, but he admitted negative terms that obey the laws

(3) *a* is interchangeable with not-not-*a*;

(4) *a* is *b* if and only if not-*b* is not-*a*.

Rule (4) is the law of contraposition familiar to the Scholastics and, for Leibniz, most recently given prominence by Jung. From rules (1) to (4), with some definitions and Leibniz's favorite method of reductio ad absurdum, the whole syllogistic can be obtained. Leibniz did not use exactly that method but adopted at one time a rather similar one based on a restatement of rule (1), *a* = *aa*, and rule (5), below. Identity has the substitutive property described below; "*a* is *b*" is made equivalent to "*a* = *ab*"; and "Some *a* is *b*" is written "*Sa* = *b*." Compound terms such as *ab* were thought of as signifying the addition of properties *a* and *b*. They obey the laws

(5) *ab* is *a*;

(6) *ab* is *b*;

(7) If *a* is *b* and *a* is *c*, then *a* is *bc*.

It has been pointed out by Karl Popper that if rules (5) and (6) are made the premises of the mood Darapti, we have the conclusion "Some *a* is *b*." This does not render the system inconsistent, but it does show that the system is already more extensive and more trivial than Leibniz presumably intended. From rules (1), (2), (5), (6), and (7) it is easy to deduce, as Leibniz did,

(8) If *a* is *bc*, then *a* is *b*, and *a* is *c*, which is the converse of (7), and

(9) If *a* is *b*, then *ac* is *bc* (using rules 2, 5, 6, and 7);

(10) If *a* is *b* and *c* is *d*, then *ac* is *bd* (using rule 9 twice and then rule 2).

Rule (10), which was known to Abelard in the twelfth century, Leibniz called *praeclarum theorema*, a very notable theorem.

Identity of terms was introduced in various ways, but always so that it was equivalent to the conjunction of "*a* is *b*" and "*b* is *a*" and so that identical terms could be substituted for one another in all contexts of the calculus. The first definition in the *Non Inelegans Specimen Demonstrandi in Abstractis*, for instance, posits that *a* = *b* holds if and only if *a* and *b* can be substituted for each other without altering the truth of any statement. The "only if" part is commonly called the principle of the identity of indiscernibles; for its converse W. V. Quine has suggested "the indiscernibility of identicals." As a principle of general application it has given rise to much discussion, although it is normally accepted in logic. While it is commonly attributed to Leibniz, Aristotle presented it in essentials in the *Topics* (VII, 1, 152a31 ff.) and *De Sophisticis Elenchis* (Ch. 24, 179a37 ff.).

An algebraic calculus requires that substitution for variables be possible, and Leibniz explicitly recognized this, in what was certainly the clearest statement in logic of the principle up to his time. Some medievals—Albert the Great, for instance—had shown their understanding of the generality conferred by variables when they called them "transcendental terms." Three more laws important for the calculus were known to Leibniz, following from rules (1), (5), (6), and (7):

(11) *ab* is *ba* (using 5, 6, and 7);

(12) *a* is *aa* (using 5);

(13) *aa* is *a* (using 1 and 7).

In the course of his experiments Leibniz came to see that particular propositions have existential import, whereas universals may not, and it was a puzzle to him what the existential import might be—factual existence or logical possibility—and whether it was built into his system or had to be further provided for. This problem had been raised by medieval logicians from the time of Abelard. One of Leibniz's solutions—that subalternation is invalid if the universal states a relation of concepts and the particular states a matter of fact but holds if we stay in one of those domains—is essentially that of Paul of Venice, who required the subjects of both propositions to have the same *suppositio*.

At a late stage Leibniz used the addition sign in place of, and with the sense of, multiplication; that is, he used *a*

+ *b* instead of *ab*. But he knew that such expressions could be interpreted as logical disjunctions, and there is also an early hint that the calculus could be interpreted propositionally, the antecedent of a conditional being said to contain the consequent. This hint may serve as a summary indication of Leibniz's position in the history of logic. Aristotle had used "antecedent" and "consequent" for "subject" and "predicate"; among medievals (such as Abelard and Kilwardby) it is often hard to tell whether the words were used of propositions or of terms; Leibniz offered a glimpse of the two domains as distinct but analogous. If his work had not gone long unpublished (we still have no complete edition), we might not have had to wait so long for the full light of Boolean day.

Bibliography

Baylis, C. A. Review of various articles on the identity of indiscernibles. *Journal of Symbolic Logic* 21 (1960): 86.

Couturat, Louis. *La Logique de Leibniz d'après des documents inédits*. Paris: Alcan, 1901; reprinted, Hildesheim, 1961.

Dürr, Karl. *Leibniz' Forschungen im Gebiet der Syllogistik. Leibniz zu seinem 300. Geburtstag*. Berlin, 1949.

Kauppi, Raili. *Über die Leibnizsche Logik mit besonderer Berücksichtigung des Problems der Intension und der Extension*. Helsinki, 1960.

Rescher, Nicholas. "Leibniz's Interpretation of His Logical Calculi." *Journal of Symbolic Logic* 19 (1954): 1–13. Reviewed by M. A. E. Dummett in *Journal of Symbolic Logic* 21 (1960): 197–199.

Ivo Thomas (1967)

EULER

The noted mathematician Leonhard Euler (1707–1783) is remembered in logic chiefly for his geometrical illustrations of syllogistic, "Euler's diagrams" or "Euler's circles." Similar devices were used by J. C. Sturm (1661), Leibniz (see Bocheński, *History of Formal Logic*, plate facing p. 260), Joachim Lange (1712), and Gottfried Ploucquet (1759), and in a very general way the idea of spatial illustration goes back at least to Juan Luis Vives, who used triangles to illustrate the Barbara syllogism ("De Censura Veri," in *Opera*, Basel, 1555). But because of Euler's fame as a mathematician and the popularity of his charming *Lettres à une princesse d'Allemagne* (the relevant letters are CII ff., dated 1761) such diagrams are traditionally named for him.

Euler used proper inclusion for the universal affirmative proposition, exclusion for the universal negative, and intersection for both the particulars. If his interpretation is followed systematically, it correctly decides the validity or invalidity of all three-term syllogisms with all terms distinct but fails for the laws of identity and contradiction and for degenerate syllogisms depending on them. Apparently nobody developed full syllogistic along these lines until J. D. Gergonne (1816), whose five relations give a complete system and can indeed be defined by three of them (see Ivo Thomas, "Eulerian Syllogistic," and references supplied there), but not by Euler's three. The extensional approach evidenced by Euler's interpretation of the universal affirmative was a healthy influence.

Euler also lent his authority to the doctrine that singular propositions are equivalent to universal ones (*Lettres*, CVII), a thesis propounded by John Wallis (from 1638; see Appendix to his *Institutio Logica*, Oxford, 1687). Bertrand Russell severely criticized this doctrine as confusing class membership with inclusion, but of course we can get an inclusive proposition equivalent to a membership proposition by taking the unit class of the singular subject.

See also Leibniz, Gottfried Wilhelm; Ploucquet, Gottfried; Propositions; Russell, Bertrand Arthur William; Vives, Juan Luis.

Bibliography

Faris, J. A. "The Gergonne Relations." *Journal of Symbolic Logic* 20 (1955): 207–231.

Gardner, Martin. *Logic Machines and Diagrams*. New York: McGraw-Hill, 1958.

Hamilton, William. *Lectures on Logic*. London, 1860.

Hocking, W. E. "Two Extensions of the Use of Graphs in Elementary Logic." *University of California Publications in Philosophy* 2 (2) (1909): 31–44.

More, Trenchard. "On the Construction of Venn Diagrams." *Journal of Symbolic Logic* 24 (1959): 303–304.

Thomas, Ivo. "Eulerian Syllogistic." *Journal of Symbolic Logic* 22 (1957): 15–16.

Thomas, Ivo. "Independence of Faris-Rejection-Axioms." *Notre Dame Journal of Formal Logic* 1 (1959): 48–51.

Venn, John. *Symbolic Logic*. 2nd ed. London: Macmillan, 1894.

Ivo Thomas (1967)

LAMBERT AND PLOUCQUET

Johann Heinrich Lambert (1728–1777), German physicist, mathematician, and astronomer, devoted a number of essays to the enterprise of making a calculus of logic, which he evidently thought of in connection with the tree of Porphyry. His standpoint is, as is usual with the early investigators, intensional. Let *a* and *b* be any concepts, *a* + *b* their combination into a compound concept, and *ab* their common part. The letters γ and δ can be multiplied with conceptual variables, so that $a\gamma$ and $a\delta$ are read as "the genus of *a*" and "the difference of *a*." The intended meaning suggests that γ and δ are descriptive operators;

yet Lambert sometimes treated them as though they were placeholders for generic or differential concepts. At any rate Lambert, following an elementary intuition, posited $a = a\gamma + a\delta = a(\gamma + \delta)$. Wanting to descend the tree to subordinate species as well as to ascend to superordinate genera and differences, he used the notation $a\gamma^{-1}$ or a/γ, which should mean "the genus under *a*." Waiving the fact that a concept containing *a* may be an ultimate species, we reflect that although $a\gamma$ is unique, $a\gamma^{-1}$ may not be so. This accounts for the trouble that Lambert found in applying multiplication and division, for $(a/\gamma)\gamma$, "the genus of a species of *a*," is identical with *a* whereas $(a\gamma)/\gamma$, "a species of the genus of *a*," need not be *a* itself. Lambert used subtraction to obtain the removal of a concept. He did not account for the appearance of coefficients and, in general, did not question the logical appropriateness of the algebraic operations to which his basic intuitions gave rise. Boole met with similar difficulties but reflected on them.

In syllogistic Lambert started not from the Aristotelian relations but from the five that are now attributed to Gergonne. This is feasible, but Lambert failed to achieve a satisfactory notion for the mutual exclusion of two terms. His most promising innovation lay in his attention to the relative product, but he did not develop this in any practical way.

Lambert, like Leibniz, experimented with sets of ruled and dotted lines to illustrate the relationships of syllogistic terms, in part trying to correct the defect in Euler's circles of not allowing for $a = b$. Some stages of his investigations were criticized by his correspondents G. J. von Holland (whose extensional standpoint was remarkable for the time) and Gottfried Ploucquet (1716–1790), both of whom were making their own efforts to evolve a logical calculus. Ploucquet, who was a teacher of Hegel, claimed independence of Euler in his use of closed figures—he used squares (1759)—and seems to have been the first to base his syllogistic on thoroughgoing quantification of the predicate. One of his notations, "$A \rangle B$" for "No *A* is *B*," strangely, enjoyed some popularity.

See also Boole, George; Hegel, Georg Wilhelm Friedrich; Lambert, Johann Heinrich; Leibniz, Gottfried Wilhelm; Ploucquet, Gottfried; Porphyry.

Bibliography

Bök, F. A., ed. *Sammlung der Schriften welche den logischen Calcul des Herrn Professor Ploucquet betreffen.* Tübingen, 1773.

Dürr, Karl. "Die Logistik Johann Heinrich Lamberts." In *Festschrift … Dr. A. Speiser.* Zürich, 1945.

Venn, John. *Symbolic Logic.* London, 1881; 2nd ed., London: Macmillan, 1894.

Ivo Thomas (1967)

BOLZANO

The most important logician of the first half of the nineteenth century was Bernard Bolzano (1781–1848). His views are closest to those of Leibniz, who preceded him by more than a century (Bolzano was sometimes called the Bohemian Leibniz). Although he quoted often and extensively from philosophers and logicians of his own generation and the preceding one, among them Kant, Salomon Maimon, Hegel, J. F. Fries, J. G. E. Maass, and K. L. Reinhold, he did this almost always in order to criticize them, and rightly so from our modern point of view, because orders of magnitude separate Bolzano as a logician from his contemporaries.

One may doubt whether he deserves to be called a forerunner of mathematical logic and modern semantics. His approach is in many respects rather crude and old-fashioned in comparison with those of George Boole and Gottlob Frege, one and two generations later, respectively. But many points first made by Bolzano look strikingly modern. Unfortunately most of these were either not noticed or not understood during his lifetime or were forgotten by later generations.

For Bolzano logic was mainly the theory of science. To investigate science he used a partly formalized language consisting of ordinary German extended by various types of constants and variables, as well as by certain technical terms which for the most part he was at great pains to define as carefully as possible.

The fundamental entities with which logic has to deal, according to Bolzano, are terms and the propositions they constitute. These abstract entities are carefully distinguished from the corresponding linguistic and mental entities. Because a single proposition can be expressed in an indefinite number of ways, Bolzano's first aim was to normalize such linguistic expressions, to reduce all of them to canonical forms prior to their purely formal treatment.

Bolzano's solution was highly idiosyncratic. Deviating radically from tradition, he claimed that all sentences (complex and compound sentences as well as simple ones) are reducible to the single form "*A* has *b*," where "*A*" is the subject term, "*b*" the predicate term, and "has" the copula. Although this reduction works reasonably well with such sentences as "John is hungry," which can easily be rendered as "John has hunger," it sounds less

convincing in the case of reducing "This is gold" to "This has goldness" (although Bolzano presented reasons why such words as "goldness" had not been created in natural languages) and still less so when "John is not hungry" is reduced to "John has lack-of-hunger." The reduction of the compound sentence "Either *P* or *Q*" to "The-term-One-of-*P*-and-*Q*-is-true has the-property-of-being-a-singular-term" or "The-term-One-of-*P*-or-*Q*-is-true has nonemptiness" (depending on whether the original expression "Either … or …" is interpreted from its context as denoting exclusive or inclusive disjunction) looks rather strange in its verbal formulation, although it looks much less strange in some appropriate symbolism. And reducing "Some *A* is *B*" to "The-term-An-*A*-which-is-*B* has nonemptiness" may appear fantastic at first sight, although it looks much more familiar when symbolized as $A \cap B \neq 0$. Nevertheless, Bolzano did not attempt to present a full set of rules for such conversions and relied instead on the reader's willingness to believe in the existence of such reductions after being shown how to perform them on certain representative samples, including some rather recalcitrant cases.

This reduction played a small role in the further development of Bolzano's work in logic. His major innovation was his introduction of the technique of variation into what amounts essentially to the logical semantics of language, even though the semantic approach, in its modern sense, was foreign to him. Starting with a proposition, true or false, he investigated its behavior with regard to truth and falsehood under substitution for any of its terms of all other fitting (that is, propositionhood-preserving) terms. (In modern terminology, he investigated all models of sentential forms.) When the number of such variants was finite he defined the degree of validity of a proposition with respect to one or more of its constituent terms as the ratio of the number of its true variants to the number of all variants. When this ratio is 1, the proposition is universally valid; when 0, universally contravalid; when greater than 0, consistent.

After extending these notions to propositional classes Bolzano was able to define an amazing number of interesting, and sometimes highly original, metalogical notions, including compatibility, dependency, exclusion, contradictoriness, contrariety, exclusiveness, and disjointness. By far the most important notion introduced in this way is that of derivability with respect to a given class of terms, defined as holding between two propositions *P* and *Q* if and only if *Q* is consistent and every model of *Q* is a model of *P* with respect to this class of terms; with respect to propositional classes it is defined similarly. This definition differs only in the unfortunate consistency clause from Tarski's definition, given in 1937, of what he called the consequence relation.

Kant had defined an "analytic" affirmative judgment as one in which the predicate concept was already contained in the subject concept. Rejecting this definition as clearly inadequate for explicating logical truth, Bolzano defined a proposition to be analytically true when universally valid with respect to at least one of its constituent terms, analytically false when universally contravalid, etc., and as analytic when either analytically true or analytically false. Bolzano was aware that this definition of analytical truth was too broad as an explication of logical truth, and he therefore went on to define a proposition as being logically analytic when (again in modern terminology) all its descriptive (extralogical) constituent terms occur in it vacuously, an anticipation of a well-known definition by W. V. Quine (1940).

Bolzano's views of probability are also strikingly modern. To define the probability of the proposition *M* on the assumptions *A, B, C, D,* · · · (with respect to certain terms *i, j,* · · ·) he used the relative degree of validity of *M* with respect to *A, B, C, D,* · · ·, which he defined as the ratio of the number of true variants of the set *M, A, B, C, D,* · · · to the number of true variants of the set *A, B, C, D,* · · ·. This conception, tenable, of course, only when the numbers involved are finite, is an important refinement of Laplace's well-known conception of probability, standard in Bolzano's time, in that it elegantly sidesteps the problem of circularity involved in the notion of equipossibility.

See also Bolzano, Bernard; Boole, George; Frege, Gottlob; Fries, Jakob Friedrich; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Laplace, Pierre Simon de; Leibniz, Gottfried Wilhelm; Maimon, Salomon; Quine, Willard Van Orman; Reinhold, Karl Leonhard; Semantics; Tarski, Alfred.

Bibliography

Bar-Hillel, Yehoshua. "Bolzano's Propositional Logic." *Archiv für mathematische Logik und Grundlagenforschung* 1 (1952): 65–98.

Berg, J. *Bolzano's Logic.* Stockholm: Almqvist and Wiksell, 1962.

Bolzano, Bernard. *Grundlegung der Logik.* Hamburg, 1964. A useful selection by Friedrich Kambartel from the first two volumes of the *Wissenschaftslehre.*

Bolzano, Bernard. *Wissenschaftslehre,* 4 vols. Sulzbach, 1837; Leipzig: Meiner, 1929–1931 (edited by W. Schulz).

Scholz, Heinrich. "Die Wissenschaftslehre Bolzanos." *Abhandlungen des Frieśchen Schule,* n.s., 6 (1937): 399–472.

Yehoshua Bar-Hillel (1967)

MODERN LOGIC

THE BOOLEAN PERIOD

The eighteenth-century and early nineteenth-century logicians considered in the preceding section were all Continental Europeans, and those who were also philosophers, namely Leibniz and Bolzano, were representatives of Continental rationalism. The British empiricism of the same period produced no logicians. On the contrary, it was antilogical. The empiricists attacked formal logic—by which they meant the attenuated syllogistic to which much of the science had shrunk during the interregnum—as trivial and sometimes as circular. This antilogicism largely echoed John Locke, whose scornful treatment of logic in his *Essay concerning Human Understanding* had provoked one of Leibniz's minor defenses of it, in the *Nouveaux Essais*. In the early nineteenth century the common logic was rescued from oblivion by Richard Whately but was not enlarged by him. Its enlargement, however, came soon after and, despite the British antilogical tradition, was at first largely a British affair, spreading later to the United States (C. S. Peirce) and then to Germany (Ernst Schröder).

See also Bolzano, Bernard; Empiricism; Leibniz, Gottfried Wilhelm; Locke, John; Peirce, Charles Sanders; Rationalism; Whately, Richard.

Bibliography

Appraisals of Hamilton's logic are to be found in John Venn, *Symbolic Logic*; C. I. Lewis, *A Survey of Symbolic Logic*; Jørgen Jørgensen, *A Treatise of Formal Logic*, Vol. I; and A. N. Prior, *Formal Logic*. Hamilton's principles, if not his practice, have found a recent defender in P. T. Geach, *Reference and Generality* (Ithaca, NY: Cornell University Press, 1962); his quantification has been analyzed in some detail by Władysław Bednarowski, "Hamilton's Quantification of the Predicate," in *PAS* 56 (1956): 217–240.

Except for *Formal Logic*, all the works by De Morgan cited in the section on him are reprinted in *On the Syllogism, and Other Logical Writings*, edited by Peter Heath (London: Routledge and K. Paul, 1966). The best modern accounts of De Morgan are to be found in Lewis, op. cit.; Jørgensen, op. cit.; and Prior, op. cit.

On Boole, see Lewis, op. cit.; Jørgensen, op. cit.; and William C. Kneale and Martha Kneale, *The Development of Logic*. The best earlier expositions are by P. E. B. Jourdain, in "The Development of the Theories of Mathematical Logic and the Principles of Mathematics"; Venn, op. cit.; and Alexander MacFarlane, in *Principles of the Algebra of Logic* (Edinburgh: Douglas, 1879). A. T. Shearman, *Development of Symbolic Logic*, is also useful.

Jørgensen, Lewis, and Jourdain all give some account of Jevons. For earlier criticism, see Shearman, above, and F. H. Bradley, *Principles of Logic* (London: K. Paul, Trench, 1883). On Jevons's machine, see Wolfe Mays and D. P. Henry, "Jevons and Logic," in *Mind* 62 (1953): 484–505; and Martin Gardner, *Logic Machines and Diagrams*.

There are appraisals of Venn in the works by Jørgensen, Jourdain, and Shearman. On Venn diagrams, see Lewis, op. cit.; Prior, op. cit.; Gardner, op. cit.; J. N. Keynes, *Studies and Exercises in Formal Logic*, 4th ed. (London: Macmillan, 1906); and almost any elementary logic text.

In addition to his book, MacColl published a series of seven papers, "The Calculus of Equivalent Statements," in *Proceedings of the London Mathematical Society* (1877–1898); a series of eight papers, "Symbolic Reasoning," in *Mind* (1880–1906); and "The Existential Import of Propositions," in *Mind* 30 (1905): 401–402, 578–580. On MacColl, see Jourdain, op. cit.; Prior, op. cit.; and Bertrand Russell's review of MacColl's *Symbolic Logic and Its Applications*, in *Mind* 30 (1906): 255–260.

Most of Peirce's logical writings are to be found in Vols. II, III, and IV of his *Collected Papers*, edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks, 8 vols. (Cambridge, MA: Harvard University Press, 1931–1958), but there is a discussion of logical paradoxes in Vol. V, Book 2, Paper 3, and one of the history of logic in Vol. VII, Book 2, Ch. 3, Sec. 10. His most developed and comprehensive logical paper is "On the Algebra of Logic: A Contribution to the Philosophy of Notation," Vol. III, Paper 13. "The Critic of Arguments," Paper 14 in the same volume, is comparatively easy reading and has a purple patch on rhemes and demonstratives. Peirce's existential graphs, which he thought were his most important contribution to logic, are the subject of Vol. IV, Book 2. *The Collected Papers* do not include some of Peirce's contributions to the *Century Dictionary*, such as the very suggestive article "Syllogism."

(A. N. P.)

HAMILTON. The nineteenth-century revival of logic in Britain, inaugurated by Whately and continued by, among others, George Bentham, chrétien, and Solly, owed much of its later impetus to the cosmopolitan learning and reforming zeal of Sir William Hamilton (1788–1856). A severely critical article by Hamilton on Whately and his followers, in the *Edinburgh Review* (1833; reprinted in his *Discussions*, London and Edinburgh, 1852), established his authority in the field, which was chiefly exercised thereafter in oral teaching from his Edinburgh chair. His scattered and largely polemical writings, including even the posthumous *Lectures on Logic* (Edinburgh and London, 1861), give a very imperfect account of his system, which acquired such order as it possessed from the works of his pupils and disciples: William Thomson and H. L. Mansel at Oxford; T. S. Baynes, John Veitch, and William Spalding in Scotland; and Francis Bowen in America. Hamilton's main service was to insist, following Kant, on the formal nature of logic and to break with the prevailing European tradition by exhibiting its forms primarily as relations of extension

between classes. He also attempted to maintain a parallel logic of intension (or comprehension) for concepts, as the inverse of extension, but this approach, like others of its kind, was a predictable, if pardonable, failure.

Hamilton's most celebrated innovation, though it was far from being his invention, was the "thoroughgoing quantification of the predicate." By attaching the quantifiers "all" ("any") and "some" to the predicate, he obtained eight propositional forms, in place of the AEIO of tradition:

(1) All *A* is all *B*.

(2) All *A* is some *B*.

(3) Some *A* is all *B*.

(4) Some *A* is some *B*.

(5) Any *A* is not any *B*.

(6) Any *A* is not some *B*.

(7) Some *A* is not any *B*.

(8) Some *A* is not some *B*.

If "some" is read as "some only," these are all simply convertible and can thus be represented as the affirmations or denials of equations. The syllogisms made up of such propositions arrange themselves, tidily enough, into 108 valid moods, 12 positive and 24 negative, in each of 3 figures (Hamilton rejected the fourth). With this arrangement, a consolidated rule of inference, and a quasi-geometrical symbolism to depict it all, Hamilton claimed to have effected a major simplification—indeed, completion—of the Aristotelian scheme.

These hopes were not borne out in the sequel. His own vacillations in the use of "some" and neglect of the differences between "all" and "any" threw even professed Hamiltonians into confusion, and the status of his propositional forms (not to mention the validity of some of his syllogisms) was much disputed. The first, for example, has no contradictory in the set and appears (on the ordinary view of "some") to be a compound of (2) and (3). The two particular affirmatives, (3) and (4), found acceptance with some writers, such as Thomson and Spalding; but of the new negatives, (6) made few friends, and (8) none at all; since it is compatible with any of the others, it says so little as to be well-nigh vacuous. A more serious objection is that since forms (1) to (5) represent all the possible ways in which two classes can be related in extension (that is, the Gergonne relations), the last three must necessarily be ambiguous or redundant.

See also Hamilton, William; Kant, Immanuel; Mansel, Henry Longueville; Whately, Richard.

P. L. Heath (1967)

DE MORGAN. The above criticisms of Hamilton's system are primarily due to Augustus De Morgan (1806–1871), whom Hamilton, in 1846, had misguidedly accused of plagiarizing his quantification. In the famous and protracted controversy that ensued, De Morgan was led into a thorough dissection of the whole system, and subsequent critics, from Mill, Peirce, and Venn onward, have taken most of their ammunition from him.

Though greatly superior as to insight and technical ability, the logic of De Morgan has affinities with that of his rival in that it, too, lays stress on the autonomy of logic and on the extensional point of view. It equally shares Hamilton's interest in reforming and enlarging the traditional syllogistic, an enterprise now outdated, which has caused it to fall into unmerited neglect. Apart from his early *Formal Logic* (London, 1847; 2nd ed., Chicago, 1926), the bulk of De Morgan's logical writings are to be found in five memoirs (plus a sixth, still unpublished) contributed to the *Cambridge Philosophical Transactions* between 1846 and 1862. The *Syllabus of a Proposed System of Logic* (London, 1860) gives a cursory account of his scheme, as does his article "Logic" in the *English Cyclopaedia* (*Arts and Science Division*, V, London, 1860, pp. 340–354).

The basis of common logic, for De Morgan, consists in relations of partial or total inclusion, or exclusion, among classes. Where information about a majority of class members is available or where, as in the "numerically definite" syllogism, precise numbers are given, it is possible, as he shows, to draw valid conclusions of a non-Aristotelian type. But these conditions are seldom realized. A more radical departure is the admission into ordinary propositions of negative terms and class names (symbolized by lower-case letters), such that a term *X* and its "contrary" *x* between them exhaust the "universe of discourse" (a useful device that has since been generally adopted). Assuming these classes to have at least notional members, it follows that two classes and their contraries can be related in eight possible ways:

(1) All *X*'s are *Y*'s.

(2) All *x*'s are *y*'s.

(3) All *X*'s are *y*'s.

(4) All *x*'s are *Y*'s.

(5) Some *X*'s are *Y*'s.

(6) Some *x*'s are *y*'s.

(7) Some *X*'s are *y*'s.

(8) Some *x*'s are *Y*'s.

These can be rewritten without negative symbols as:

(1) All *X*'s are *Y*'s.

(2) All *Y*'s are *X*'s.

(3) No *X*'s are *Y*'s.

(4) Everything is either *X* or *Y*.

(5) Some *X*'s are *Y*'s.

(6) Some things are neither *X*'s nor *Y*'s.

(7) Some *X*'s are not *Y*'s.

(8) Some *Y*'s are not *X*'s.

Of these the contradictory pairs are (1) and (7), (2) and (8), (3) and (5), and (4) and (6). Since the distribution of terms is given or implied throughout, these forms are simply convertible by reading them in reverse. "Contraversion" (or obversion) is obtained by altering the distribution of a term, replacing it by its contrary, and denying the result. "All *X*'s are *Y*'s" becomes successively "No *X*'s are *y*'s," "All *y*'s are *x*'s," and "Everything is either *x* or *Y*." The procedure is the same for the other seven forms, making 32 possibilities in all.

De Morgan's rule of syllogism is either that both premises should be universal or, when only one is, that the middle term should have different quantities in each. Inference takes place by erasing the middle term and its quantities. Since, including the syllogisms of weakened conclusion, there are 4 basic patterns, and since 3 terms and their contraries can be paired off, in premises and conclusion, in 8 different ways, there are 32 valid syllogisms, of which half have two universal premises and 8 a universal conclusion.

To remedy the "terminal ambiguity" whereby the undistributed term in the universal "All *X*'s are *Y*'s" may refer indifferently to some or all of the *Y*'s, De Morgan investigated the complex propositions produced by combining pairs of elementary forms. It is in this connection that he gives the well-known rules for negation of conjunctions which have since received his name—though he did not, in fact, invent them.

In endeavoring to patch up Hamilton's quantified system De Morgan made further distinctions between "cumular" (collective) and "exemplar" (distributive) forms of predication; struggled, unavailingly, to bring the intensional interpretation of terms (as attributes) into line with the extensional and to subsume both under a pure logic of terms (the "onymatic" system); and explored in passing such nontraditional forms of inference as the syllogisms of "undecided assertion" and "transposed quantity." More important is his recognition that the copula performs its function in inference, not as a sign of identity, but only through its role as a transitive and convertible relation.

De Morgan's generalization of the copula leads on, in his fourth Cambridge memoir, to a pioneer investigation of relations in general, which is the foundation of all subsequent work in the field. He there distinguishes a relation (say, "lover of") from its denial, its contrary, and its converse ("loved by"); proceeds to compound relations, or relative products ("*L* of *M* of"), and to quantified versions of these ("*L* of every *M*," "of none but *M*'s," etc.); and discusses a variety of equivalences that hold between these different sorts of relations and the rules for their discovery and manipulation. The purpose of this, typically enough, was to exhibit the syllogism in its most general form, as a series of combinations of relations. Despite the ingenuity and resource with which he treated it, this devotion to the syllogism was something of a weakness in De Morgan's work. It tethered him too closely to tradition, so that it was not until others exploited them that his own most fruitful discoveries were seen for what they were.

See also De Morgan, Augustus; Hamilton, William; Mill, John Stuart; Peirce, Charles Sanders; Venn, John.

P. L. Heath (1967)

BOOLE. George Boole (1815–1864) was the founder of modern mathematical logic. Nevertheless, few of his ideas are currently accepted in mainstream logic in the forms originally proposed by him. His learned and fertile mind conceived of several important hypotheses, the testing and modification of which changed the face of logic irrevocably. One of his most important hypotheses was that every proposition can be expressed using an algebraic equation suitably reinterpreted: that logic and algebra share a common uninterpreted formal language and thus also that they have similar problem types and similar methods.

The universal affirmative, or A proposition, "Every square is a rectangle" was expressed by $x = xy$, where x is the class of squares, y the class of rectangles, and xy the "Boolean or logical product" of x with y, the class of common members of x and y. The universal negative, or E proposition, "No rectangle is a circle" was expressed by $yz = 0$, where z is the class of circles and 0 is the empty

class—an idea Boole introduced into logic. The conclusion "No square is a circle," $xz = 0$, which Aristotle and previous logicians deduced in one "intuitive" step, was derived by Boole using a chain of algebraic manipulations—illustrating another of his hypotheses, namely that on some level reasoning was mechanical or algorithmic.

He used 1 for the universe, or "universe of discourse," a ubiquitous expression in modern logic that Boole coined. He used the minus sign for "logical subtraction": $1 - x$ is the class of objects in the universe that are not in the class x. Using the above symbols, expression of the particular affirmative, or I proposition, "Some rectangle is a square" and the particular negative, or O proposition, "Some rectangle is not a square" as *inequalities* would have been easy: $yx \neq 0$ and $y\ (1 - x) \neq 0$. This is a point that Boole never mentioned and probably did not notice—Boole's hypothesis was that algebraic *equations* were sufficient. Instead, he conceived of a logical operator, now called Boole's vee, or the vee, which was to produce from a class x a resultant class vx supposed by him to be "indefinite in every respect except that it contains some individuals of the class [x in this case] to whose expression it is prefixed." Using the vee, Boole "expressed" the above $vy = vx$ and $vy = v\ (1 - x)$. The vee itself as well as the two "translations" have been criticized by later logicians—mainstream logic has not adopted Boole's vee, although its similarity to other more recent nonstandard operators has been noted—for example, the Hilbert epsilon.

Using the algebraic formal language, Boole was able to express several "laws of thought" analogous to laws of algebra; indeed some were expressed by the same equations used for laws of algebra—for example, the commutative law $xy = yx$. He employed his laws of thought in two unprecedented ways. First, regarding the equations as conditions on "unknowns," he created a wholly new theory of logical equation-solving using the laws of thought the way laws of algebra are used in numerical equation-solving. Second, regarding the most basic of his laws of thought as laws of logic, he created an axiomatization of logic. Boole realized that no "class logic" as such could treat the arguments now dealt with in truth-functional proposition logic. To meet this deficiency he proposed an ingenious reinterpretation of his system that, in his view, transformed it into something akin to propositional logic. In the process, he discovered key ideas now incorporated into modern truth-function logic, establishing himself as the first modern figure in any history of propositional logic. These are but three of Boole's many revolutionary innovations.

See also Aristotle; Boole, George; Propositions.

John Corcoran (2005)

JEVONS. It was the aim of William Stanley Jevons (1835–1882), himself a pupil of De Morgan, to render Boole's calculus more simple and "logical" by removing those of its features that he found "mysterious" and by reducing its operations to mechanical routine. He also professed, officially, to reject the extensional standpoint in favor of a "pure logic" of terms, or "qualities," though the result in practice was still effectively a class or propositional logic, conceived rather in the manner of De Morgan's "onymatic" system. These views are set forth in two pamphlets, *Pure Logic* (London and New York, 1864) and *The Substitution of Similars* (London, 1869; both reprinted in *Pure Logic and Other Minor Works,* London, 1890), and at greater length in *The Principles of Science* (2nd ed., London, 1887) and *Studies and Exercises in Deductive Logic* (London, 1884).

Jevons takes over the Boolean notations for conjunction and identity (AB, $A = B$) and admits negative classes, which he symbolizes, like De Morgan, by a small *a,* but makes no use of 1, the universal class, and dismisses as uninterpretable both the operations of subtraction and division and the various ill-favored symbols—$(1 - x)$, x/y, 0/0, 1/0, etc.—that result from their use. In the case of disjunction (written + or, more generally, ·|·)Jevons follows the minority view of De Morgan and a few others in proposing to read it inclusively, so that $A + B$ is permitted to have common members, and $A + A = A$ (law of unity). The importance of this reform, almost universally accepted since, is that it abolishes the need for numerical coefficients, establishes the symmetry between conjunction and disjunction exhibited, for example, in De Morgan's laws, $\overline{AB} = a + b$ and $\overline{A + B} = ab$, and makes possible such other useful rules of simplification as the "law of absorption," $A + AB = A$.

Jevons conceives of classes as groups of individuals, and of propositions about such classes, or about qualities, as equations asserting a complete or partial identity between them. Thus, "All A is B" identifies all A's with those that are B—that is, $A = AB$—and the corresponding E-proposition is $A = Ab$. He symbolizes particular propositions, on occasion, by an arbitrary prefix, but pays little attention to them—or, indeed, to the problems of quantification in general. Inference consists merely of what he calls the "substitution of similars"—that is, the replacement of any term by another, stated in a premise to be identical to it. Thus, $A = AB$ and $B = BC$ yield, by substitution, $A = ABC = AC$, the conclusion.

Of more interest is the Jevonian method of indirect inference, based on what he calls the "logical alphabet." This alphabet, which amounts to no more than a Boolean expansion of 1, is constructed by listing all the possible combinations of the terms *A, B, C,* etc., together with their negatives, thus:

ABC	*aBC*
ABc	*aBc*
AbC	*abC*
Abc	*abc*

Any given premise, say $A = BC$, on being combined with each line in turn will be found inconsistent with some—that is, will yield an expression equal to 0. These lines being struck out, the remainder give the conclusion, though it still remains to consider the "inverse problem" (which Jevons saw but did not solve) of expressing the results in a single concise formula. Particular propositions are somewhat troublesome to handle on this scheme, which actually works better for propositions than for classes. But with many terms the process soon becomes tedious in either case, and it was to remedy this that Jevons invented his "logical abacus" and "logical piano," contrivances which operate mechanically on the same principle, namely the employment of the premises to eliminate inconsistent combinations from a matrix already set up on the machine. The development of the modern computer has revived interest in Jevons's pioneer device and in his very able description of its workings. For the rest, Jevons's "equational logic," though famous in its day, is now remembered chiefly for the technical improvements on Boole's procedure that it helped to bring into use.

See also Logic Machines.

P. L. Heath (1967)

VENN. The logic of John Venn (1834–1923), sketched briefly in the *Princeton Review* (1880) and more fully elaborated in his *Symbolic Logic* (London, 1881), shows a greater understanding of George Boole's intentions and a better acquaintance with the historical background than had yet been displayed by anyone else. Though he did not suppose the new methods to have any great practical advantage over the old, he saw no reason, either, to suspect them of being anything more than a generalization of traditional practices, couched, for convenience, in a mathematical form. He therefore resisted the Jevonian simplifications and was at pains to bring out the logical significance of such operations as subtraction and division, though the latter is admitted to merit inclusion more on grounds of consistency than for any use made of it in the reasoning of everyday life.

Venn's own account of the matter proceeds from what he calls the "compartmental," or "existential," view of logic, whose purpose is to set out the possible ways in which the four classes designated by *x, y*, and their negatives, in combination, may have one or more of their components empty. Omitting the case where all four compartments are unoccupied, this yields fifteen forms of proposition, compared with the four that arise on the traditional, or predication, view, whereby an attribute is asserted or denied of a class, and the five that emerge from diagrammatic consideration of the ways in which two nonempty classes may include, exclude, or overlap one another. Each view has its merits, in Venn's opinion, the choice between them being ultimately a conventional one.

This leads Venn to the discussion of another vexed issue, the "existential import" of propositions. Traditional logic must in consistency assume that its classes have members and nonmembers alike, and its universal propositions are thereby rendered hypothetical. To Venn it was clearer what the universal denies than what it asserts, and he therefore proposed to write A, "All *x* is *y*," as $x\bar{y} = 0$ and E, "No *x* is *y*," as $xy = 0$. These propositions are definite, yet they do not require members in *x* or *y* to make them true, since they deny only the existence of members in the common class. Particular propositions do, however, imply the presence of members in each class, since they contradict the universals; they are therefore to be written I, $xy \neq 0$, and O, $x\bar{y} \neq 0$, respectively. This was an improvement on Boole's use of indefinite symbols and has since been generally adopted, though one consequence of it (also noted by Hugh MacColl) is that subalternation ceases to be valid and that the "syllogisms of weakened conclusion" which depend on it have therefore to be rejected.

Venn was not much enamored of the syllogism, but he deserves the gratitude of all beginners in the subject for what is probably his best-known contribution to logic, the diagrams that bear his name. These are, in effect, graphical representations of the algebraic processes introduced by Boole and mechanically illustrated in Jevons's alphabet: The partitioning of a universe in terms of the possible combinations of *x, y*, and so on, and the elimination of those subdivisions inconsistent with the premises given. For two terms a pair of intersecting circles (*x* and *y*) on a ground give the four compartments $xy + x\bar{y} + \bar{x}y + \bar{x}\bar{y} = 1$ (Figure 1). Three interlaced circles (Figure 2) depict the eight combinations of Jevons's table,

FIGURE 1

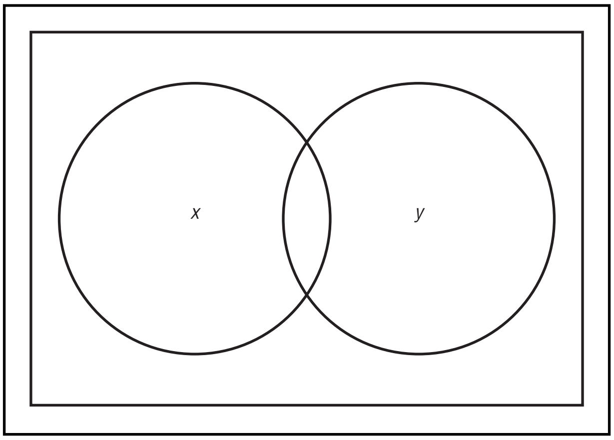

FIGURE 2

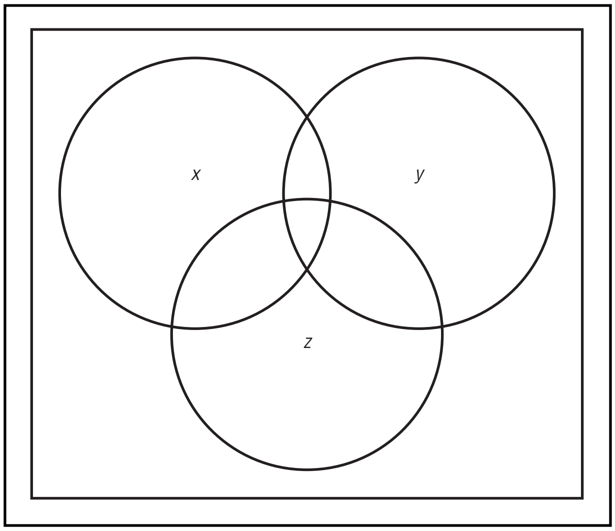

given earlier. The effect of a universal premise is to declare one or more compartments to be empty, shown by shading the area in question. A particular premise indicates that one or more compartments have occupants, shown by a cross (which may lie ambiguously on the boundary between two areas). The conclusion can then be read off, in various ways, by inspection. By the use of ellipses the same principle can be employed for up to five terms, but it then becomes unwieldy, especially in the "inverse problem" of formulating the outcome, so that one or another of the square diagrams devised by later authors is at that stage generally preferable. With suitable modifications the method can also be extended to the calculus of propositions. Though Venn did not carry this extension far, he was led by it to an early realization of the truth-functional character of the relation of material implication.

The merit of Venn's work lies not in its original departures, which are few, but rather in the light it throws on the obscurities of Boole's procedure and in its very careful and fair discussion of opposing views.

See also Boole, George; Jevons, William Stanley; Venn, John.

P. L. Heath (1967)

CARROLL. The contributions of Lewis Carroll (Charles L. Dodgson, 1832–1898) to logic consist of several pieces published between 1887 and 1899. *The Game of Logic* (London, 1887) is a book written for young people to teach them to reason logically by solving syllogisms using diagrams and colored counters. His diagrammatic method is a visual logic system that we know now to be sound and complete.

In *Symbolic Logic, Part I* (London, 1896) Carroll developed two formal methods to solve syllogisms and sorites. The first is the Method of Underscoring that is dependent on his idiosyncratic algebraic notation that he called the Method of Subscripts. The second is his Method of Diagrams, which he extended to handle more than three terms (classes), but without providing examples. However, his diagrammatic system is an improvement over that of his contemporary, John Venn, because first, unlike Venn's system, Carroll's can handle existential statements. Second, as A. Macula showed in 1995, diagrams for ten terms (sets) or more can be drawn more easily than Venn diagrams for a large number of sets. Finally, the diagrams are self-similar and can be generated by a linear iterative process. Carroll used his method to reduce the nineteen or more valid forms of inference codified by medieval Aristotelian logicians first to fifteen forms and then to just three formulas.

Carroll published two pieces in the journal *Mind*. The first, "A Logical Paradox" (N. S. vol. 3, 1894, 436–438) is an example of hypothetical propositions. W. W. Bartley III remarks in the second edition of his book, *Lewis Carroll's Symbolic Logic* (1986, p. 505) that for about eighty years eminent logicians and philosophers failed to see this problem as little more than a routine exercise in Boolean algebra. Of the eleven questions Dodgson sent to *The Educational Times* (ten on mathematical topics) the substance of one, Question 14122, (February 1, 1899, vol. lii, p. 93) on his logical paradox, had appeared as a "Note" to his 1894 *Mind* article. H. MacColl and H. W. Curjel provided (different) solutions. The second piece in *Mind*, "What the Tortoise Said to Achilles" (N. S. vol. 4, 1895, 278–280) is a humorous example of an important problem about logical inference that Carroll was perhaps the

first to recognize: the rule allowing a conclusion to be drawn from a set of premises cannot itself be treated as an additional premise without generating an infinite regress.

We see in Bartley's 1986 publication of Carroll's lost book, *Symbolic Logic, Part II*, that Carroll introduced two additional methods of formal logic. The first, the method of barred premises, a direct approach to the solution of problems involving multiliteral statements is an extension of his Method of Underscoring. The second and most important, the Method of Trees, a mechanical test of validity using a *reductio ad absurdum* argument, is the earliest modern use of a truth tree to reason in the logic of classes. It uses one inference rule (binary resolution) and a restriction strategy (set of support) to improve the efficiency of the construction. His tree method is a sound and complete formal logic system for sorites.

See also Carroll, Lewis; Logic Diagrams; Venn, John.

Bibliography

Abeles, Francine F. "Lewis Carroll's Formal Logic." *History and Philosophy of Logic* 26 (2005): 33–46.

Bartley, William W., III, ed. *Lewis Carroll's Symbolic Logic*. 2nd ed. New York: Clarkson N. Potter, 1986.

Beth, Evert W. *The Foundations of Mathematics*. Amsterdam: North Holland, 1965.

Macula, Anthony J. "Lewis Carroll and the Enumeration of Minimal Covers." *Mathematics Magazine* 69 (1995): 269–274.

Wos, Larry, et al. *Automated Reasoning*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

Francine F. Abeles (2005)

PEIRCE. The logical work of Charles Sanders Peirce (1839–1914) was an unusual blend of the traditional and the modern. His early paper "Memoranda concerning the Aristotelian Syllogism," read and distributed in 1866, adapted to the second and third syllogistic figures Kant's description of first-figure reasoning as the subsumption of a case under a rule, and in later papers he exhibited analogy and induction as probabilistic weakenings of the second and third figures thus conceived. In 1867, independently of Jevons, Peirce improved Boole's logical algebra by identifying logical addition with the inclusive rather than the exclusive sense of "either-or." In 1870, inspired by De Morgan's pioneer work on the logic of relations, he extended Boole's method of algebraic analogy to this discipline, noticed that there are three-termed as well as two-termed relations, and introduced the sign "—<" for class inclusion, considered an analogue of the arithmetical "≤."

In 1880, Peirce began to use the symbol "—<" indifferently for class inclusion, implication, and the "therefore" of inference. It became one of his persistent themes that the distinction between terms, propositions, and inferences is of little logical importance. For him all propositions are, in the end, implications (this thesis is bound up with his pragmatic theory of meaning) and as such are simply inferences deprived of an element of assertiveness; terms, at least general terms, are propositions deprived of a subject. General terms are "rhemes," or, as we would now say, "open sentences," sentences with gaps where names might go. Such sentences with gaps are in a broad sense relative terms, the number of gaps indicating what Peirce called the "adinity" of the relation. Thus, "— loves —" represents a "dyadic" relation,"— gives — to —" a "triadic" one, and so on. Extending this conception downward, Peirce described an ordinary predicative term, such as "— is a man," as representing a "monadic" relation and a complete sentence, with no gaps at all, as representing a "medadic" one.

As Frege did with his "concepts," Peirce compared his "rhemes" to unsaturated chemical radicals having various valencies. Unlike Frege, however, he did not subsume rhemes under functions, like "The square of —," as the special case in which the value of the function for a given argument is a truth-value. Frege's procedure underlined the resemblance between a completed proposition and a name; for Peirce a completed proposition was rather a special case of a predicate. Nevertheless, Peirce pioneered (in 1885) the use of truth-value calculations in establishing logical laws and also foreshadowed many-valued logic by suggesting that there might be an infinity of degrees of falsehood, with truth as the zero.

A gap in a rheme may be filled, in the simplest case, by what Peirce called an "index." He divided signs into indices, which operate through some physical connection with what they signify; icons, which operate through some resemblance to what they signify; and symbols, which acquire their meaning by convention. An ordinary proper name is an "icon of an index"; it is (when uttered) a noise that resembles the noise that was made when we were introduced to the person named. A simple index would be, for example, a demonstrative pronoun accompanied by a pointing gesture. Peirce regarded the phrase "demonstrative pronouns" as an inaccurate description—it would be more appropriate to call a noun a "pro-demonstrative." A common noun, for Peirce, is only an inseparable element in a rheme (for example, "man" in "is a man").

Instead of directly filling a gap in a rheme with an index, we may say either "I can find you an object such that it—" ("is a man," "loves Susan," etc.) or "Take anything you like and it —" ("is mortal if human," etc.). These are the particular and universal quantifiers, which Peirce introduced into his logic—independently of Frege, but with some debt to his own student O. H. Mitchell—in 1883. He represented them with the mathematical symbols "Σ" and "∏" for continued sums and products. If we write "$a = 0$" for "a is false" and "$a \rightarrow 0$" for "a is true," $\Sigma_i a_i$ or "For some individual i, a_i" will have for its value the sum of the values of the possible a_i's and therefore will be $\rightarrow 0$ (that is, true) if and only if at least one of the a_i's is $\rightarrow 0$, whereas $\prod_i a_i$ or "For any individual i, a_i" will have for its value the product of the values of the possible a_i's and therefore will be $\rightarrow 0$ if and only if all of the a_i's are $\rightarrow 0$. Peirce was aware of the possibility of putting any quantified expression into what is now called prenex normal form, with all the quantifiers at the beginning. He also, in what he called second-intentional logic, quantified over variables other than those standing for indices.

Every implication, Peirce came to believe, has an implicit or explicit initial quantifier—that is, is of the form $\prod_i(a_i \prec b_i)$, "For any i, if a_i, then b_i." The i's may be either ordinary individuals *of* which our a and b may be true, or instants *at* which they may be true, or possible states of affairs *in* which they may be true; for example, "If it rains it pours" may mean "For any instant *i*, if it rains at *i*, it pours at *i*" or "For any possible state of affairs *i*, if it rains in *i*, it pours in *i*." But in the latter case we may consider wider or narrower ranges of possibility, and if we limit ourselves to the actual state of affairs, the quantifier may be dropped.

Peirce made several attempts to define negation in terms of implication, and in 1885 he produced a set of axioms for the propositional calculus with implication accepted as an undefined operator and negation defined as the implication of a proposition from which anything at all would follow. This was the second set of axioms sufficient for the propositional calculus to be produced in the history of the subject (the first being Frege's of 1879) and the first set to use the curious law $((a \prec b) \prec a) \prec a$, now called Peirce's law. But Peirce experimented with other types of systems also, and in 1880 he anticipated H. M. Sheffer in showing that all truth-functions can be defined in terms of "Neither — nor —" and "Not both — and —." The "not" within a proposition (as opposed to "It is not the case that —," governing the whole), which forms the "negative propositions" of traditional logic, he regarded as expressing the relation of otherness, and he worked out what properties of this relation are reflected in traditional logical laws. For example, the law of contraposition, "'Every *A* is a *B*' entails that whatever is not a *B* is not an *A*," follows from the mere fact that otherness is a relation, for whatever relative term *R* may be, if every *A* is a *B*, then whatever is an *R* (for instance, an other) of every *B* is an *R* of every *A*.

Peirce thought it desirable that logical formulas should reflect the structure of the facts or thoughts which they express and so be, in his sense, "icons"—that is, signs operating by resemblance to what they signify—and he sought constantly to develop symbolisms that were genuinely "iconic." In his later years he came to regard this as best achieved by a system of diagrams which he called "existential graphs." Typically, he attempted to represent his graph for "If *A* then *B*" as basic, but in fact his diagrams are most easily understood as starting from the representation of "and" by juxtaposition and of "not" by enclosure in a bracket or circle or square. $(A(B))$, which is his graph for "If *A* then *B*," reads off naturally as "Not both *A* and not *B*." Rules of inference are represented as permissions to alter the graphs by insertions and erasures; for example:

(R1) We may insert or remove double enclosures at will, provided that there is no symbol caught between the two enclosures; for instance, we may pass from *A* to $((A))$, i.e., to "Not not *A*," and back, but not from $(A(B))$ to AB.

(R2) Any symbol may be removed from an evenly enclosed graph (including a completely unenclosed one) or added to an oddly enclosed one; for instance, we may pass from AB, i.e., "*A* and *B*," to *A*, or from $(A(BC))$ to $(A(B))$, i.e., from "If *A* then both *B* and *C*" to "If *A* then *B*," or from (A) to (AB), i.e., from "Not *A*" to "Not both *A* and *B*."

(R3) We may repeat a symbol across an enclosure immediately interior to the symbol's own, and if a symbol is already thus repeated, we may remove it from the inner enclosure; for instance, we may pass from $(A(B))$ to $(A(AB))$, i.e., from "If *A* then *B*" to "If *A* then both *A* and *B*," or from $A(AB)$ to $A(B)$, i.e., from "*A* and not both *A* and *B*" to "*A* and not *B*."

If a graph is such that these permissions will enable us to transform it into any graph at all, that graph is "absurd" and its negation a logical truth. For example, $A(A)$, "Both *A* and not *A*," leads by R2 to $A((B)A)$, where *B* is any graph you please, and this leads by R3 to $A((B))$, this by R2 to $((B))$, and this by R1 to *B*. Hence, $(A(A))$, "If *A* then *A*," is a logical truth. For clarity Peirce suggested drawing rectangular enclosures, with evenly enclosed

symbols written on the left and oddly enclosed ones on the right. For example, Figure 3 is a representation of (*A*(*B*(*C*)), "If *A* then (*B* but not *C*)." This arrangement makes it clear that Peirce was, in effect, setting up what are nowadays called "semantic tableaux," in the manner of E. W. Beth.

Peirce also thought of logical truth as represented by the blank sheet on which his graphs were drawn and absurdity by an enclosure with nothing but the blank graph sheet inside it. Since by R2 we may inscribe anything whatever in such an otherwise blank enclosure, this enclosure would in fact represent an absurdity in the previous sense of a graph that can be transformed into any graph whatsoever. "If *A* then absurdity," Peirce's favorite definition of "Not *A*," would then be strictly "(*A*((   )))" ("If *A* then *B*" is "(*A*(*B*))," and here we put "(   )" for *B*), but this assumes that in representing the absurd as "(   )" we already understand simple enclosure as negation, and in attempting to modify his symbolism in ways which would avoid this assumption Peirce was led into occasional unnecessary trouble.

Although Peirce was one of the inventors of bound variables, in his graphs for quantified formulas he explicitly dispensed with them in favor of what he called "lines of identity," a device recently put to the same purpose, though informally, by W. V. Quine and Peter Geach. A monadic rheme may be written as "— *A*" or "*A* —," the single valency line being close enough to be thought of as part of the symbol, and on its own this symbol is read as "Something is *A*." If "— *B*" is added to this, the whole, "*A* — — *B*," of course, means "Something is *A* and something is *B*." But if the valency lines are joined by a "line of identity," to give us "*A* —— *B*," this means "Something is *A* and *that same thing* is *B*," or "Something is at once *A* and *B*." In the common systems this identification of the subjects of which *A* and *B* are predicated is effected by attaching these predicates to the same bound variable, thus: "For some *x*, *x* is *A* and *x* is *B*." Again, "If anything is *A* then *that same thing* is *B*" is distinguished in the common systems from the more indefinite "If anything is *A* then something is *B*" by writing the former with a common bound variable, thus: "For any *x*, if *x* is *A* then *x* is *B*." In Peirce's graphs this is done by tightening "(— *A* (— *B*))" to "(*A*(*B*))" or "(*A* —(*B*))." To give some examples with dyadic rhemes, "Every *A* is an *R* of some *B*" comes out as "(*A* —(*R* — *B*))"; "Some *B* is *R*'d by every *A*" as "(*A* —(*R*))— *B*"; and "Every *A* is an *R* of itself" as "(*A* —(*R*))" or "(*A*——(*R*))."

This "Beta part" of Peirce's graphs, of course, contains special rules for the transformation of lines of identity. For example, the additions and erasures of terms permitted by R2 may be extended to terms attached to others by lines of identity; thus, we may pass from "*A* —— *B*," "Something is at once *A* and *B*," to the plain "— *B*," "Something is *B*." Peirce said that the blank sheet—which is left here when "*A* —" with its line of identity is removed and which represents accepted truth when considered as a medad—represents an accepted existent when considered as a monad.

FIGURE 3

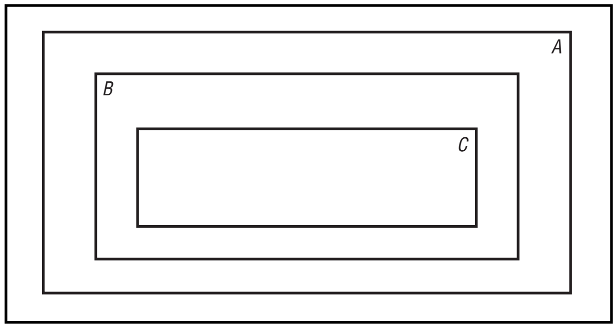

Since lines of identity may themselves be treated as dyadic rhemes and subjected to enclosure, the graphs cover identity theory and, therefore, the arithmetic of specific integers, as well as the theory of first-order quantification. For example, "There are at least two *A*'s" will be "*A* —( )— *A*"—that is, "Something is an *A*, and something that *is not* that thing is also an *A*." But the graphs do not readily lend themselves to the representation of higher-order quantifications, such as "Some qualities belong to everything and others to some things only," although Peirce made some rather clumsy efforts in this direction. More successful, but only adumbrated in outline, was his extension of his method to modal logic by using separate sheets for different possible worlds. This procedure is very like that now adopted by S. A. Kripke) and also echoes medieval theories of "ampliation."

There is probably no logical writer who has been more rich in original suggestions than Peirce, and his papers are a mine that has still to be fully worked. He was, at the same time, more aware than any of his contemporaries of the contributions made by their ancient and medieval predecessors. He held and persuasively supported a theory that Aristotle had anticipated (in a chapter of the *Prior Analytics* (now missing) later derivations of simple conversion from the laws of identity and syllogism, and he saw the significance of the Megarian controversy over the nature of implication and of the distinctions drawn by the Schoolmen in their theory of *consequentiae*.

Peirce's immediate circle in America included two logicians of some distinction: O. H. Mitchell, from whom Peirce derived the germ of his device of quantification, and Christine Ladd Franklin (1847–1930), who used eight "copulae" to construct De Morgan's eight categorical forms and exhibited syllogisms in different figures as derivable from "inconsistent triads," or "antilogisms." An antilogism states that a certain three propositions—for example, "Every *Y* is a *Z*," "Every *X* is a *Y*," and "Not every *X* is a *Z*"—cannot all be true: hence (syllogism 1), the first and second jointly imply the denial of the third; also (syllogism 2), the first and third jointly imply the denial of the second; also (syllogism 3), the third and second jointly imply the denial of the first.

See also Boole, George; De Morgan, Augustus; Existence; Frege, Gottlob; Jevons, William Stanley; Kant, Immanuel; Modal Logic; Peirce, Charles Sanders; Quine, Willard Van Orman.

A. N. Prior (1967)

THE HERITAGE OF KANT AND MILL. The development of logic, at least of formal logic, in the nineteenth century was largely independent of the general development of philosophy during the same period. Of the logicians considered in the preceding section only C. S. Peirce and perhaps William Hamilton were of importance in branches of philosophy other than logic, and the persons who were of most importance in other branches of philosophy contributed nothing whatsoever to technical developments of the sort here described. These persons did not ignore logic altogether, however, nor did competent logicians entirely ignore them. It will be helpful, therefore, to break the chronological order at this point and to glance back at these philosophical developments and influences.

In the nineteenth century, as in the eighteenth, there were divergent Continental and British philosophical influences, but the Continental stream, stemming from Immanuel Kant (1724–1804), was now not so much rationalistic as idealistic, and in logic it was increasingly antiformal, antimathematical, and antitechnical. Kant himself could not be described as antiformal; he had a quite exalted view of the place of formal logic in philosophy. Unfortunately, however, he thought of formal logic not as a field for new developments but as the first science to have reached perfection—it had reached perfection, he said, with the work of Aristotle. Even Kant's "Aristotelianism" was of the sadly truncated variety that had been characteristic of the interregnum. Slightly systematizing what he took to be Aristotelian logic, he divided "judgments" according to their "quantity" into universal, particular, and singular; according to their quality into affirmative (*X* is *Y*), negative (*X* is-not *Y*) and infinite (*X* is not-*Y*); according to what he called "relation" into categorical, hypothetical (that is, conditional), and disjunctive; and according to modality into apodictic (asserting necessity or impossibility), assertoric, and problematic (asserting possibility). The division according to quality is particularly absurd; where would one put, for example, the forms "*X* is-not not-*Y*" and "Not-*X* is *Y*"? More influential was his subdivision of affirmative categoricals into "analytic," in which the predicate concept is implicitly contained in the subject concept, and "synthetic," in which it is not. "Body is extended," for example, is analytic because what is meant by a body is precisely an extended substance.

The empiricism that had characterized British philosophy in the eighteenth century was still in evidence in the nineteenth in the work of John Stuart Mill (1806–1873), but Mill was not, as the eighteenth-century British empiricists had been, antilogical or antimathematical. He did not personally advance the young science of mathematical logic, but he was not hostile to it, and in the later nineteenth century it was possible for J. N. Keynes and W. E. Johnson to develop a logical style that was indebted almost equally to Mill and to the mathematicians.

Mill's own formal logic, like Kant's, was rather thin, and for details he referred his readers to Richard Whately; the greater part of his *System of Logic* (London, 1843) is devoted to what would now be called scientific method. Its first two books, however, contain well-developed theories about the meaning of various types of words and sentences and about the nature of syllogistic reasoning. It may be added here that the propositions corresponding to what Kant called analytic judgments were described by Mill as "merely verbal."

In the later nineteenth century there was considerable crossing of geographical and philosophical boundaries. Christoph Sigwart (1830–1904), in Germany, was indebted to Mill as well as to Kant; Franz Brentano (1838–1917), in Austria, owed much to Mill and nothing at all to Kant. The antimathematical logical tradition of Kant and G. W. F. Hegel was carried further in England by F. H. Bradley and Bernard Bosanquet, just when logic as an exact science was being given in Germany a new impetus by Gottlob Frege.

See also Aristotle; Bosanquet, Bernard; Bradley, Francis Herbert; Brentano, Franz; Empiricism; Frege, Gottlob; Hamilton, William; Hegel, Georg Wilhelm Friedrich;

Johnson, William Ernest; Kant, Immanuel; Logic, Traditional; Mill, John Stuart; Peirce, Charles Sanders; Sigwart, Christoph; Whately, Richard.

Bibliography

TEXTS

Hegel, G. W. F. *Wissenschaft der Logik*, Vol. I: *Die objektive Logik*, 2 vols. Nuremberg, 1812–1813. Vol. II: *Die subjektive Logik*. Nuremberg, 1816. Translated by W. H. Johnson and L. G. Struthers as *The Science of Logic*, 2 vols. London, 1929.

Kant, Immanuel. *Kritik der reinen Vernunft*. Riga: Hartknoch, 1781; 2nd ed., 1787. Translated by Norman Kemp Smith as *Critique of Pure Reason*. London: Macmillan, 1929.

Lotze, R. H. *Logik*. Leipzig, 1880. Translated by Helen Dendy as *Logic*. Oxford, 1884.

Mill, J. S. *An Examination of Sir William Hamilton's Philosophy*, 2nd ed. London: Longman, Green, Longman, Roberts and Green, 1865.

Mill, J. S. *A System of Logic*. London, 1843; 8th ed., 1872.

Sigwart, Christoph. *Logik*, 2 vols. Tübingen, 1873–1878. Translated by Helen Dendy as *Logic*, 2 vols. London, 1890.

A. N. Prior (1967)

KEYNES. John Neville Keynes (1852–1949) was for a large part of his long life registrar of the University of Cambridge. His first contribution to logic was an article in *Mind* in 1879, in which he defended formal logic as a substantial discipline distinguishable alike from the philosophical logic being pursued by the heirs of Kant and Hegel, from the "empirical" (largely inductive) logic developed by the heirs of J. S. Mill, and from the mathematical logic lately started on its career by Boole and De Morgan.

In 1884, Keynes's view of the subject was exhibited in greater detail in the first edition of his *Studies and Exercises in Formal Logic*. This work dealt, in the traditional manner, successively with terms, judgments, and syllogisms, but it had a fourth part in which essentially Boolean material was presented as a logic of categorical propositions with conjunctive, disjunctive, and negative terms and conjunctive and disjunctive compounds of these propositions. Each chapter in the book consisted of a number of well-constructed exercises, sometimes with introductory remarks and often with lengthy comments. Part I, on terms, was much influenced by the treatment of names in Book I of Mill's *System of Logic*. Part II was distinguished by a very judicious discussion, in Chapter 8, of the problems raised by Brentano and Venn about the existential import of categorical propositions.

In successive revisions and enlargements in 1887, 1894, and 1906 the chapters took on the more normal shape of extended discussions with exercises at the end. Part IV (on compound and complex propositions) was transformed into a long appendix, and much new material was incorporated. W. E. Johnson, in the preface to his own logic, was able to refer to the final product as "Dr. Keynes's classical work, in which the last word has been said on most of the fundamental problems of the subject." To this result Johnson himself generously contributed; he and Keynes had frequent and regular discussions of logical problems, and many of the footnotes in Keynes's third and fourth editions express his indebtedness to Johnson. For example, Keynes owed to Johnson the distinction between "conditional" and "true hypothetical" propositions that Russell later dealt with more precisely as one between formal and material implication.

Keynes's literary style was of singular clarity and distinction, and he dealt urbanely but decisively with the many sophistries and confusions that were current, especially among logical writers of a broadly idealist stamp, such as Bosanquet and Bradley. At the same time, he paid attention, particularly in his final edition, to the broadly "intensional" considerations to which these writers were perhaps more sensitive than many whose standards of logical rigor were higher. He handled modal distinctions, for example, with the same neatness and skill which he brought to other topics, and he anticipated C. I. Lewis in drawing attention to what are now called the paradoxes of strict implication.

The development of Keynes's thought from edition to edition, as he brought it to bear on one topic after another, is fascinating to examine. For instance in dealing with what Mill called the connotation and denotation of general names he distinguished even in the first edition between (1) the connotation proper—that is, the set of attributes that we select by convention as those that an object must have if we are to give the name to it—and (2) the totality of attributes possessed in common by all the attributes to which the name applies. In the second edition he suggested that for (2) we might use the Port-Royalists's term "comprehension." Thus, the connotation being selected by convention, objective facts determine the name's denotation, that is, which objects have the attributes entitling them to the name, and further objective facts determine the comprehension, that is, which attributes beyond the connotation these objects have in common. But in the third edition Keynes noted that we might alternatively fix the application of a name by an "exemplification," a selection not of attributes but of objects, with respect to which we decide that we will give a certain name to anything which possesses all the attrib-

utes that these objects have in common (making an exception, as Johnson reminded Keynes that we would have to do, of such attributes as that of having been selected for this purpose). When we proceed this way convention fixes the exemplification, and the facts determine the comprehension and then the denotation.

See also Boole, George; Bosanquet, Bernard; Bradley, Francis Herbert; Brentano, Franz; De Morgan, Augustus; Existence; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Lewis, Clarence Irving; Mill, John Stuart; Modal Logic; Russell, Bertrand Arthur William; Venn, John.

A. N. Prior (1967)

JOHNSON. Keynes's collaborator William Ernest Johnson (1858–1931) did not publish Part I of his own *Logic* until 1921 (Part II, 1922; Part III, 1924), although he had published a series of three articles titled "The Logical Calculus" in *Mind* in 1892 (17: 3–30, 235–250, 340–357) and two titled "The Analysis of Thinking" in *Mind* in 1918 (27: 1–21, 133–151). In the first series the variables in Boolean equations were explicitly given the propositional interpretation, the logical product ("x and y") being represented by juxtaposition and negation by a superimposed bar. The logical product and negation being taken as primitive, "If x then y" is defined as "Not (x and not y)"—that is, $\overline{x\bar{y}}$—the logical sum "x or y" as "Not (not x and not y)," and universal and particular quantification as continued logical multiplication and addition. "The Analysis of Thinking" is more philosophical and seems to reflect the influence of G. F. Stout's *Analytic Psychology*.

Johnson's *Logic* exhibits an attractive combination of the formal elegance of his 1892 articles with the philosophical penetration of those of 1918. In some ways—for example, in his extensive discussion of "problematic induction" (that is, scientific generalization)—he played Mill to Keynes's Whately. His book is now best known for its development of the distinction between "determinables" and "determinates," in Part I, Chapter 11. A "determinable" is one of the broad bases of distinction that may be found in objects, such as color, shape, size. Under each of these fall more or less determinate characteristics, such as red, blue, and so on, under color (and scarlet, crimson, etc. as more determinate forms of red). Johnson used this distinction as the basis of many further developments. In Part II, Chapter 10, for example, Johnson discussed what he called "demonstrative induction," in which a universal conclusion is deduced from a singular premise by the help of an "all-or-nothing" proposition. From "Either every S is P or every S is not P" and "This S is P" we can infer "Every S is P." A natural extension is the form of reasoning in which the major premise asserts that every S exhibits the same determinate form of the determinable P (for instance, every specimen of a given element has the same atomic number) and the minor that this S exhibits the determinate form p of this determinable; hence, every S is p. (Cf. Mill on "uniform uniformities" in his *System of Logic*, Book III, Ch. 4, Sec. 2.)

Johnson presented many critical asides concerning Russell's *Principles of Mathematics*, the most valuable being in Part II, Chapter 3, "Symbolism and Functions."

See also Mill, John Stuart; Russell, Bertrand Arthur William; Stout, George Frederick; Whately, Richard.

Bibliography

WORKS ON JOHNSON

Acute and careful, as well as captious, criticism of Johnson is contained in H. W. B. Joseph's "What Does Mr. W. E. Johnson Mean by a Proposition?," in *Mind* 36 (1927): 448–466, and 37 (1928): 21–39. Johnson's views are also discussed in A. N. Prior's "Determinables, Determinates and Determinants," in *Mind* 58 (1949): 1–20, 178–194. There is a fine informative obituary of Johnson by C. D. Broad in *Ethics and the History of Philosophy* (London: Routledge, 1952).

A. N. Prior (1967)

FROM FREGE TO GÖDEL

Twentieth-century logic, and even late nineteenth-century logic, cannot be properly understood without some acquaintance not only with earlier nineteenth-century logic but also with nineteenth-century mathematics. The final section of our survey therefore begins with a sketch of the influence of nineteenth-century mathematics on the major logical developments of both the Boolean and the more recent periods. This will be followed by discussions of particular logicians.

Bibliography

(Secondary material on the logicians discussed in the final section of the article is to be found mainly in the first two parts of the bibliography.)

Carnap, Rudolf. *Foundations of Logic and Mathematics*. Chicago: University of Chicago Press, 1939. This is Vol. I, no. 3, of *International Encyclopedia of Unified Science*.

Carnap, Rudolf. *Introduction to Semantics*. Cambridge, MA: Harvard University Press, 1942.

Carnap, Rudolf. *Meaning and Necessity*. 2nd, enl. ed. Chicago: University of Chicago Press, 1956.

Church, Alonzo. *Introduction to Mathematical Logic*. Vol. I. Princeton, NJ: Princeton University Press, 1956.

Church, Alonzo. "A Note on the Entscheidungsproblem." *Journal of Symbolic Logic* 1 (1936): 40–41. See also the correction, ibid., 101–102.

Church, Alonzo. "An Unsolvable Problem of Elementary Number Theory." *American Journal of Mathematics* 58 (1936): 345–363.

Frege, Gottlob. *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black. New York: Philosophical Library, 1952.

Gentzen, Gerhard. "Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlen-theorie." *Mathematische Annalen* 119 (1943): 140–161.

Gentzen, Gerhard. "Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie." *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, n.s., 4 (1938): 19–44.

Gentzen, Gerhard. "Untersuchungen über das logische Schliessen." *Mathematische Zeitschrift* 39 (1934): 176–210, 405–431. Translated into French, with commentary, by R. Feys and J. Ladrière as *Recherches sur la déduction logique*. Paris, 1955.

Gentzen, Gerhard. "Die Widerspruchsfreiheit der reinen Zahlentheorie." *Mathematische Annalen* 112 (1936): 493–565.

Gödel, Kurt. Articles on intuitionist logic and number theory. *Ergebnisse eines mathematischen Kolloquiums* 4 (1933): 34–38, 40.

Gödel, Kurt. *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory*. Princeton, NJ: Princeton University Press, 1940; rev. ed., 1951.

Gödel, Kurt. "What Is Cantor's Continuum Problem?" *American Mathematical Monthly* 54 (1947): 515–525.

Herbrand, Jacques. "Sur le Problème fondamental de la logique mathématique." *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie*, Classe III, 24 (1931): 12–56.

Hilbert, David. *Gesammelte Abhandlungen*. 3 vols. Berlin: Springer, 1932–1935.

Hilbert, David, and Paul Bernays. *Grundlagen der Mathematik*. 2 vols. Berlin: Springer, 1934–1939.

Post, Emil. "Formal Reductions of the General Combinatorial Decision Problem." *American Journal of Mathematics* 65 (1943): 197–215.

Post, Emil. "Recursively Enumerable Sets of Positive Integers and Their Decision Problems." *Bulletin of the American Mathematical Society* 50 (1944): 284–316.

Post, Emil. "Recursive Unsolvability of a Problem of Thue." *Journal of Symbolic Logic* 12 (1947): 1–11.

Quine, W. V. *Set Theory and Its Logic*. Cambridge, MA: Harvard University Press, 1963.

Ramsey, F. P. *The Foundations of Mathematics and Other Logical Essays*, edited by R. B. Braithwaite. New York: Harcourt Brace, 1931.

Skolem, Thoralf. "Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre." In *Wissenschaftlicher Vorträge, gehalten auf dem Fünften Kongress der Skandinavischen Mathematiker in Helsingfors 1922*,217–232. Helsinki, 1923.

Skolem, Thoralf. "Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen." In *Skrifter Utgit av Videnskapsselskapet i Kristiania, I. Matematisk-naturvidenskapelig Klasse 1919*, no. 4, 1–36. Oslo, 1920.

Skolem, Thoralf. "Über die Nicht-Charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen." *Fundamenta Mathematicae* 23 (1934): 150–161.

Skolem, Thoralf. "Über die Unmöglichkeit einer vollständigen Charakterisiergung der Zahlenreihe mittels eines endlichen Axiomsystems." *Norsk Matematisk Forenings Skrifter*, series 2 (10) (1933): 73–82.

Skolem, Thoralf. "Über einige Grundlagenfragen der Mathematik." In *Skrifter Utgit av det Norske Videnskaps-akademi i Oslo, I. Matematisk-naturvidenskapelig Klasse*, no. 4, 1–49. Oslo, 1929.

Wittgenstein, Ludwig. *Philosophische Bemerkungen*, edited by Rush Rhees. Oxford: Blackwell, 1964.

(A. N. P.)

NINETEENTH-CENTURY MATHEMATICS. Mathematics in the nineteenth century was characterized by reorganization in every field, effected both by generalization, which led to the viewing of areas once considered discrete as special instances of the same general case, and by the examination of foundations, either in terms of basic concepts or by an axiomatic approach. Apart, therefore, from any specific contributions that mathematicians made to modern logic, the atmosphere was highly favorable to an explicitly logical investigation both of mathematics in general and of its various branches, including, by the end of the century, mathematical logic itself. At the same time, the growth of abstract algebra encouraged the persistence of Leibniz's ideal of mathematizing deductive logic; his ideas, although most were unpublished, maintained a steady, if at first tenuous, foothold. Thus, the early mathematical logicians, having caught the idea of a new kind of algebra, tended to work on it as a specialized branch of mathematics. By the end of the nineteenth century it had become an instrument sufficiently perfected to be able to discard its traditional algebraic appearance, even to forget momentarily its self-concern, and to apply itself to the articulation of the increasingly well-organized mathematical material. Only in the twentieth century did it catch up to its own axiomatic origins and fruitfully rejoin its algebraic ones.

Peacock. As early as 1821, A. L. Cauchy (1789–1857), in his influential *Cours d'analyse* (Paris, 1821, introduction, p. ii), attacked the current use of algebraic reasonings in geometry because "they tend to make one

attribute an indefinite range to the algebraic formulas, while in reality most of these formulas hold uniquely under certain conditions, and for certain values of the quantities concerned." This thought was adopted, in a more positive version, by George Peacock (1791–1858) in *A Treatise on Algebra* (2 vols., London, 1842–1845), elaborating a work of 1830. Instead of merely rejecting such illegitimate, or at any rate unjustified, extensions of the ranges of algebraic formulas, he distinguished between two kinds of algebra, arithmetical and symbolic.

> Arithmetical algebra is the science which results from the use of symbols and signs to denote numbers and the operations to which they may be subjected; those numbers or their representatives, and the operations upon them, being used in the same sense and with the same limitations as in common arithmetic. [In symbolical algebra] the symbols which are used are perfectly general in their representation, and perfectly unlimited in their values; and the operations upon them, in whatever manner they are denoted, or by whatever name they are called, are universal in their application. (Vol. I, Ch. 1)

The relationship of the two is more fully explained in the introduction:

> The generalizations of arithmetical algebra are generalizations of reasoning not of form. … Symbolical algebra adopts the rules of arithmetical algebra, but removes altogether their restrictions. … It is this adoption of the rules of the operations of arithmetical algebra as the rules for performing the operations which *bear the same names* in symbolical algebra, which secures the absolute identity of the results in the two sciences so far as they exist in common. … This principle, in my former Treatise on Algebra, I denominated the "principle of the permanence of equivalent forms."

Peacock expressed his conviction that the convention by which such permanence had been commonly assumed had both delayed the emergence of his symbolical algebra as a science in its own right and resulted in consequent confusion and false reasoning such as Cauchy had complained of. By contrast to arithmetical algebra, "the results of symbolical algebra, which are not common to arithmetical algebra, are generalizations of form, and not necessary consequences of the definitions" which introduce special conditions according as the variables denote lines, forces, periods of time, and so on.

Boole. It is not hard to see the influence of Peacock's thoughts on George Boole. In the introduction to *The Mathematical Analysis of Logic* (1847), Boole wrote:

> Those who are acquainted with the present state of the theory of symbolical algebra, are aware, that the validity of the process of analysis does not depend upon the interpretation of the symbols which are employed, but solely upon the laws of their combination. Every system of interpretation which does not affect the truth of the relation supposed, is equally admissible. … That to the existing forms of analysis a quantitative interpretation is assigned, is the result of the circumstances by which those forms were determined, and is not to be construed into a universal condition of analysis. It is upon the foundation of this general principle, that I purpose to establish the calculus of logic, and that I claim for it a place among the acknowledged forms of mathematical analysis, regardless that in its object and in its instruments it must at present stand alone.

In this passage we see mathematical logic struggling to be born, aware of its parentage, but still uncertain, as it continued to be for some time, of its status. Boole himself interpreted his calculus in relation to both classes and propositions. Thus, "The symbol $1 - x$ selects those cases in which the proposition X is false" (*The Mathematical Analysis of Logic*, "Of Hypotheticals"), and "Let us for simplicity of conception give to the symbol x the particular interpretation of *men,* then $1 - x$ will represent the class of 'not-men'" (*An Investigation of the Laws of Thought,* London, 1854, Ch. 3 in Prop. iv).

Peacock's work drew increased attention to the formal properties of operations, and Boole regarded his subject from this point of view.

> The laws we have established … are sufficient for the base of a calculus. From the first of them it appears that the elective symbols are *distributive,* from the second that they are *commutative*; properties which they possess in common with symbols of *quantity,* and in virtue of which, all the processes of common algebra are applicable to the present system." (*The Mathematical Analysis of Logic,* "First Principles")

These terms actually antedate Peacock; they may have been introduced by F. J. Servois (see *Annales des mathématiques,* 5 [1814]: 93). "Associativity" has been ascribed

to Sir William Rowan Hamilton (see Hermann Hankel's *Theorie der complexen Zahlensysteme,* Leipzig, 1867).

Gergonne. The new trend in algebra was already evidenced by the "Essai de dialectique rationelle" (in *Annales des mathématiques* 7 [1816–1817]: 189–228) of J. D. Gergonne (1771–1859). In this he wrote:

> In the same way that an algebraic calculation can be carried out without one having the least idea about the meaning of the symbols on which one is operating, it is possible to follow a course of reasoning without any knowledge of the meaning of the terms in which it is expressed, or without adverting to it if one knows it.

Such a formalistic approach would have been more in order when fields of application were better charted, and Karl Weierstrass was still fighting for this point of view many years later. Gergonne later did important work on duality in geometry, which shows again his ability to distinguish structure from interpretation. He offered a new analysis of the fundamental ideas of syllogistic and used an inverted *C* for inclusion, now standardized as the hook, ⊂.

De Morgan. Augustus De Morgan, a contemporary of Peacock and Boole, took a special interest in the organization of mathematics for didactic purposes. After *Elements of Arithmetic* (1830) he wrote *On the Study and Difficulties of Mathematics* (1831), *First Notions of Logic* (1839), which was designed to help beginning students of geometry, and *Formal Logic* (1847). In *Trigonometry and Double Algebra* he investigated symbolic calculuses. A remarkable text ("On the Syllogism, III") shows De Morgan striking out element after element in the material proposition "Every man is animal" till he is left with *X*——*Y,* showing the "pure form of the judgment"; thus, he made a start on the extension of the mathematical notion of function, to which Boole, Peirce, and most notably Frege also contributed. De Morgan's right parenthesis, as used in "*X*)" to mean "every *X*," yielding "*X*)*Y*"—that is, every *X* is *Y*—is reminiscent of Gergonne's inverted *C,* although Gergonne's symbol means "is contained in" and operates on two terms rather than one.

Grassmann. One of the creators of a new form of algebra was H. G. Grassmann (1809–1877). Grassmann's *Ausdehnungslehre* (Leipzig, 1844; rev. ed., 1862), fundamental to vector analysis, anticipated W. R. Hamilton's work through its greater generality and influenced Alfred North Whitehead's *A Treatise on Universal Algebra with Applications* (Cambridge, U.K., 1898). Giuseppe Peano's *Calcolo geometrico* (Turin, 1888) was written "according to the *Ausdehnungslehre* of H. Grassmann, preceded by the operations of deductive logic."

Non-Euclidean geometry. In geometry the great breakthrough was the effective creation of non-Euclidean systems. The chief figures were János Bolyai (1802–1860), Nikolai Ivanovich Lobachevski (1793–1856), and Bernhard Riemann (1826–1866). Bolyai's work on non-Euclidean geometry was titled *Appendix Scientiam Spatii Absolute Veram Exhibens; A Veritate aut Falsitate Axiomatis XI Euclidei (A Priori Haud Unquam Decidenda) Independentem.* Written in 1823, it was published in 1833 at Maros-Vásárhely in the second volume of the *Tentamen* of his father, F. Bolyai. Lobachevski wrote *Geometrische Untersuchungen zur Theorie der Parallellinien* (Berlin, 1840), an elaboration of ideas first presented in a lecture delivered at Kazan in 1826. Riemann's inaugural lecture *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen* (1854) was published at Göttingen in 1867. Each seems to have done his work independently of the others, but behind all of them appears the great, although in this matter somewhat enigmatic, figure of Karl Friedrich Gauss (1777–1855), friend of Bolyai's father and of Lobachevski's teacher Bartels and teacher of Riemann. Gauss's correspondence shows him long to have had ideas on the subject, and to him we owe the word *non-Euclidean* (in a letter to Taurinus, 1824).

Bolyai, as the title of his work indicates, simply dropped Euclid's axiom of parallels; Lobachevski adopted its denial. Both required the infinity of the straight line. Riemann, approaching the matter from an analytic point of view, wished to determine the general conditions of spaces in which the measure of distance would remain everywhere constant and figures could move freely without deformation. He was thus led to consider spaces of constant curvature and more than three dimensions, with Euclidean space a special case. Riemann's work was immediately taken up by Hermann von Helmholtz (1821–1894), in *Über die thatsachlichen Grundlagen der Geometrie* (1868–1869) and *Über die Thatsachen, die der Geometrie zu Grunde liegen* (1868), and was further refined by Sophus Lie (1842–1899). Lie was one of the principal developers of the theory of groups, which Felix Klein (1849–1925) applied to geometry in his *Erlanger Programm, Vergleichende Betrachtungen über neuere geometrische Forschungen* (Erlangen, 1872; translated by M. W. Haskell as "A Comparative Review of Recent Researches in Geometry," in *Bulletin of the New York Mathematical Society* 2 [1892–1893]: 215–249).

Independence. Though Bertrand Russell (in 1897), Whitehead (in 1898), and David Hilbert (in 1899) all

wrote on geometry, and Hilbert's later foundational work (*Grundlagenforschung*) provided the basis for all subsequent investigations, these pioneers of mature mathematical logic failed to secure independence for their propositional axioms. This is remarkable after all the attention that had been devoted to the independence of Euclid's axiom of parallels. Frege, too, failed in this matter. Alessandro Padoa, in 1901, gave directives for establishing the independence of concepts within an axiom system—an idea that influenced Peano—but no general method for securing the independence of propositional axioms was attained until Jan Łukasiewicz (1925) and Paul Bernays (1926), independently, found the method of interpretation by matrices.

Many-valued logics and proof theory. Non-Euclidean geometries are often mentioned in discussions of the status of many-valued logics, but they appear to have had no direct influence. (Łukasiewicz was brought to the idea by Aristotle's *Peri Hermeneias.*) It is likely that the theory of groups (closed systems of operations)—which was already finding widespread application by the end of the nineteenth century—and the rise of different algebras did much to create the climate of thought in which proof theory, and in general the metalogical investigation of the properties of entire deductive systems, could be developed. Such investigation seems to be one of the most notable characteristics differentiating mathematical logic from the logic of any other period. Proof theory stems mainly from Hilbert.

Schröder. The early, algebraic period of mathematical logic ended with Ernst Schröder (1841–1902). After a paper on algorithms for solving equations (1870) and a textbook on arithmetic and algebra, Schröder devoted himself more and more to the algebra of logic, his two chief works being *Der Operationskreis des Logikkalküls* (Leipzig, 1877) and *Vorlesungen über die Algebra der Logik* (3 vols., Leipzig, 1890–1905). Much of his work was a tidying up of the past. He discarded Boole's subtraction and division, which were subject to too many restrictions to be satisfactory inverse operations; used (as had W. S. Jevons) the sign of addition in the sense of inclusive rather than exclusive alternation; and introduced at the beginning a sign for inclusion. In this last matter he independently duplicated Frege's abandonment of the algebraic form in *Begriffsschrift* (Halle, 1879), which later became standard with *Principia Mathematica* (3 vols. Cambridge, U.K., 1910–1913). But Schröder remained interested in the solution of equations; his results for the Boolean system were taken over by Whitehead in *A Treatise on Universal Algebra.* Like Peirce, Schröder noticed a duality between logical multiplication and addition and similarly between the null and the universal classes. Duality in geometry had been brought to the fore by J. V. Poncelet (1822), enunciated with greater generality by Gergonne (1827), and skillfully exploited by Jakob Steiner (1830).

Schröder explicitly rejected those syllogisms that are invalid when the terms are null, Boole having merely passed them over. Besides using the method of 1 – 0 evaluation, which goes back to Boole, he developed a process of reduction to normal form. Schröder introduced two novelties. Unlike those of his contemporaries mentioned above, he was interested in independence, wishing particularly to have the distributive law independent of his other axioms, and he was thus brought to perhaps the first idea of a nondistributive lattice. He also had a clear view of the need for a theory of logical types:

> By that process of arbitrary selection of classes of individuals of the manifold originally envisaged, there arises a new, much more extensive manifold, namely that of the domains or classes of the previous one. … [It] is necessary from the start that among the elements given as individuals there should be no classes comprising as elements individuals of the same manifold. (*Vorlesungen,* Vol. I, p. 247)

This foreshadows Russell's vicious-circle principle.

Schröder worked on Peirce's algebra of dyadic relatives as an extension of Boole's algebra, but the result was unsatisfactory, and, indeed, by the time Peirce reviewed it Schröder had already abandoned the algebraic form (though not the name) in favor of what is essentially first-order functional calculus. The Schröder-Bernstein theorem, to the effect that if each of two classes is similar to a part of the other, then they are similar to each other, was proved by Schröder in 1896 and independently by Felix Bernstein in 1898.

Peano. Schröder deplored the lack of use for the logical tool he had developed and experimented with the application of his theory of relation to Dedekind's chains. Giuseppe Peano, primarily interested in the rigor of mathematical proof, applied Schröder's instrument to comprehensive mathematical material in successive volumes of his *Formulaire de mathématiques* (5 vols., Turin, 1892–1908). He prefaced the work with a section on mathematical logic (a phrase that he originated), distinguished class membership from inclusion, which Schröder had not done, and expressed all theorems as implications rather than as equations. He still did not iso-

late propositional logic as a deductive preliminary, but he stated a generalized form of *modus ponens,* to the effect that a true proposition could be suppressed when it occurred as an antecedent or as part of a conjunction of antecedents in a theorem.

Peano had already obtained his five axioms of arithmetic, which contain the principle of mathematical induction, by 1889, when he published *Arithmetices Principia Nova Methodo Exposita.* The year before, J. W. R. Dedekind had reached substantially the same result in *Was sind und was sollen die Zahlen?* (Brunswick, Germany, 1888) with the induction principle provable, however, owing to his having started further back in logic, with sets and projections, rather than with sets, number, and successor. Frege, as Dedekind did not know at that time, had gone still further in the same direction. The fact that Peano, even in 1908, did not refer to either Frege or Dedekind but explicitly left the possibility of defining "number" an open question may indicate that he continued to be interested in logic more as a means of attaining brevity and rigor, and an occasional new insight, than as material from which the basic arithmetical notions might be constructed.

Cantor. Peano did draw on the theory of sets of Georg Cantor (1845–1918), including Cantor's proofs that the algebraic numbers can be put in one-to-one correspondence with the positive integers and that the real numbers cannot be so made to correspond (the "diagonal" proof). Cantor's work had grown out of a reorganization of analysis parallel to that of algebra and geometry. He was influenced, of course, by the work of Cauchy, Riemann, and Hankel on functions of complex variables, but his principal predecessor was Karl Weierstrass (1815–1897), who was greatly interested in foundational matters, especially in regard to irrational numbers and points of condensation of infinite sets. Cantor became convinced that without extending the concept of number to actually infinite sets it would hardly be possible to make the least step forward without constraint. The arithmetic that he thus created was welcomed by Frege; its influence is widely apparent and was acknowledged in Russell's *Principles of Mathematics* (Cambridge, U.K., 1903), which plotted the future progress of *Principia Mathematica.*

See also Aristotle; Boole, George; Cantor, Georg; De Morgan, Augustus; Frege, Gottlob; Geometry; Helmholtz, Hermann Ludwig von; Hilbert, David; Jevons, William Stanley; Łukasiewicz, Jan; Many-Valued Logics; Peano, Giuseppe; Peirce, Charles Sanders; Proof Theory; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bibliography

R. C. Archibald, "Outline of the History of Mathematics," in *American Mathematical Monthly* 56 (1949), cites standard histories of nineteenth-century mathematics. See also the general histories of logic listed above, as well as Roberto Bonola, *Non-Euclidean Geometry* (New York: Dover, 1955); Alonzo Church, "Schroder's Anticipation of the Simple Theory of Types," in *Erkenntnis* 9 (1939): 149–152; Georg Cantor, *Contributions to the Founding of the Theory of Transfinite Numbers,* translated, with an introduction, by P. E. B. Jourdain (Chicago: Open Court, 1915); Ettore Carruccio, *Mathematics and Logic in History and in Contemporary Thought,* translated by Isabel Quigly (Chicago: Aldine, 1965); H. B. Curry, *Foundations of Mathematical Logic* (New York: McGraw-Hill, 1963); and Giuseppe Peano, *Formulario Mathematico* (Turin, 1908; facsimile reprint, Rome: Edizioni Cremonese, 1960).

Ivo Thomas (1967)

FREGE. Modern logic began with the publication in 1879 of the *Begriffsschrift* of Gottlob Frege (1848–1925). In the *Begriffsschrift* we find for the first time a comprehensive treatment of the ideas of generality and existence, because sentence forms which were hitherto accommodated only by complicated ad hoc theories are here provided with an adequate symbolization by the device of quantification, rules for which are adjoined to the first complete formalization of the classical propositional calculus. The result closely approximates a modern formal axiomatic theory. It meets Frege's aim of a codification of the logical principles used in mathematical reasoning, although the rules of inference (substitution and *modus ponens*) and the definition of other logical constants in terms of the primitives (negation, implication, the universal quantifier, and identity) are not explicitly formalized but are mentioned as obviously justified by reference to the intended interpretation. A proof of completeness was not to be had in Frege's day, but he demonstrated the power of his system by deriving a large number of logical principles from his basic postulates and took an important step toward the formulation of arithmetical principles by showing, with the aid of second-order quantification, how the notion of serial order may be formalized.

After the *Begriffsschrift*, Frege's next major work was *Die Grundlagen der Arithmetik* (Breslau, 1884), an analysis of the concept of cardinal number presented largely in nontechnical terms. It opens the way for Frege's theories with a devastating criticism of the views of various writers on the nature of numbers and the laws of arithmetic. Difficulties encountered in the analyses of number find

explanation and resolution in the celebrated claim that a statement of number contains an assertion about a concept. To say, for instance, that there are three letters in the word *but* is not, on Frege's view, to attribute a property to the actual letters; it is to assign the number 3 to the concept "letter in the word 'but'." If we now say that two concepts *F* and *G* are numerically equivalent (*gleichzahlig*) if and only if there is a one-to-one correspondence between those things which fall under *F* and those which fall under *G*, we can define the number that belongs to a concept *F* as the extension of the concept "numerically equivalent to the concept *F*."

In terms of this definition any two numerically equivalent concepts, such as "letter in the word 'but'" and "letter in the word 'big'," can be seen to determine the same extension, and therefore the same number, and it remains only to specify concepts to which the individual numbers belong. In sketching this and subsequent developments Frege found that the notions used appear to allow of resolution into purely logical terms. He concluded that it is probable that arithmetic has an a priori, analytic status, a view that places him in opposition to Immanuel Kant, who held that propositions of arithmetic were synthetic a priori, and to J. S. Mill, who regarded them as inductive generalizations.

In papers published after the *Grundlagen*, Frege turned his attention to problems of a more general philosophical nature, and the development of his thought in this period led to a revised account of his logic, which is incorporated in his most ambitious work, *Die Grundgesetze der Arithmetik* (2 vols., Jena, Germany, 1893–1903), in which he extended and formalized the theory of number adumbrated in the *Grundlagen*. In the *Begriffsschrift* he had rejected the traditional subject-predicate distinction but had retained one predicate, "is a fact" (symbolized "$\vdash$"), which indicated that the judgment which it prefaced was being asserted. In his essay "Über Sinn und Bedeutung" this view was abandoned on the ground that the addition of such a sign, conceived as a predicate, merely results in a reformulation of the same thought, a reformulation which in turn may or may not be asserted.

The logic of the *Grundgesetze* is based on Frege's theory of sense and reference, the interpretation of the symbolism of the *Begriffsschrift* being modified accordingly. The formal system of the *Begriffsschrift* is further changed by replacing certain of the axioms with transformation rules, but a more important innovation is the extension of the earlier symbols to cover classes. Corresponding to any well-defined function $\Phi(\xi)$ is the range, or course of values (*Wertverlauf*), of that function, written $\dot{\epsilon}\Phi(\epsilon)$, which Frege introduced via an axiom stipulating that $\dot{\epsilon}\Phi(\epsilon)$ is identical with $\epsilon\psi(\epsilon)$ if and only if the two associated functions $\Phi(\xi)$ and $\psi(\xi)$ agree in the values which they take on for all possible arguments ξ. In particular, this axiom licenses the passage from a concept to its extension, the course-of-values notation providing a means of representing classes and foreshadowing Bertrand Russell's class-abstraction operator, $\hat{z}(\phi z)$. Another device that found a close analogue in Russell's logic is Frege's symbol $\backslash\xi$. If a course of values ξ has a unique member, then $\backslash\xi$ is this member; otherwise $\backslash\xi$ is the course of values ξ itself. In the first case $\backslash\xi$ provides a translation of expressions of the form "the *F*" and so corresponds to Russell's description operator, $(\imath x)(\phi x)$; the second case ensures that when ξ has no unique member, $\backslash\xi$ is nevertheless well defined.

The preliminary development of logic and the theory of classes is followed by the main subject of the *Grundgesetze*, the theory of cardinal number, developed with respect to both finite and infinite cardinals. The theory of real numbers is begun in the second volume but the treatment is incomplete, and Frege was probably loath to advance further in this direction after learning, while the second volume was in the press, that the very beginnings of his theory harbored a contradiction. This contradiction, discovered by Russell, resulted from the axiom allowing the transition from concept to class, an axiom in which Frege had not had the fullest confidence. Russell's communication is discussed in an appendix to the second volume, where an emended version of the axiom is put forward. This emendation was not, in fact, satisfactory, and although Frege apparently did not know that a contradiction could still be derived, he eventually abandoned his belief that the program of the *Grundgesetze* could be carried out successfully and claimed that geometry, not logic, must provide a basis for number theory.

See also Frege, Gottlob; Kant, Immanuel; Mill, John Stuart; Russell, Bertrand Arthur William.

Bede Rundle (1967)

PEANO. Giuseppe Peano (1858–1932), professor of infinitesimal analysis at Turin and a prolific writer on a wide range of mathematical topics, contributed to the early development of both logicism and the formalism to which it is partly opposed. His first book, published under the name of a former teacher, Angelo Genocchi, was devoted to the calculus and featured a careful, systematic treatment of the subject that contrasted favorably with customary texts in rejecting loosely phrased defini-

tions and theorems and in substituting rigorous proof for appeals to intuition. Peano was particularly insistent that the acceptability of a mathematical proposition should depend not on its intuitive plausibility but on its derivability from stated premises and definitions, and he devised a remarkable illustration of the way in which what appears evident to intuition may nonetheless be contradicted by formally incontrovertible considerations. This is his well-known space-filling curve, introduced in 1890 in the paper "Sur une Courbe, qui remplit toute une aire plaine" (*Mathematische Annalen* 36 [1890]: 157–160). About ten years earlier Camille Jordan had defined a curve as a continuous and single-valued image of the unit segment. This definition accords well enough with our intuitive conception of a curve, but Peano showed that a curve in conformity with this definition could in fact pass through every point in a square based on the unit segment and so would appear as a uniformly shaded surface if plotted on a graph.

Convinced that the development of mathematics must proceed independently of intuitive considerations, Peano embarked upon a program of refounding the various branches of mathematics. Not only geometry and analysis, where we are particularly inclined to make an appeal to what can be grasped pictorially, but even elementary number theory was to be purified of commonsense preconceptions. The entities of a mathematical theory (numbers, points, and so forth) would have to enter into the theory not as idealizations of objects given to intuition but as postulated or defined entities, having only those properties which are explicitly listed or which can be grounded on the initial definitions. To ensure the exclusion of misleading intuitive associations, Peano devised a new symbolic language in which to formalize definitions and other postulates. Principles of reasoning employed within mathematics, as well as conceptions forming the substance of mathematical theories, are transcribed into the new notation. It is at this point that mathematical logic enters into Peano's work, and although he did not carry the development of his system very far, the basic ideas and notation were taken over by Whitehead and Russell as a starting point for the system of logic presented in great detail in *Principia Mathematica*.

Also important for subsequent developments was Peano's presentation of arithmetic. It is based on a set of postulates known as the Peano axioms, although, as has been noted, Richard Dedekind had published them earlier. The axioms were intended to free the concept of number from dependence on intuition. The essentials of Peano's treatment are embodied in these five axioms:

(1) 0 is a number.

(2) The successor of any number is a number.

(3) No two numbers have the same successor.

(4) 0 is not the successor of any number.

(5) Any class which contains 0 and which contains the successor of n whenever it contains n includes the class of numbers.

The Peano axioms are commonly taken as a basis for the arithmetic of the natural numbers, supplemented by recursive definitions of such arithmetical operations as addition, multiplication, and exponentiation. Peano himself made considerable use of recursive definition, an analogue, for definitions, of the axiom of mathematical induction given by (5), which allows us to calculate the value of a function $f(n)$ step by step, given an explicit definition of $f(0)$ along with a definition of $f(n')$ in terms of $f(n)$—here "n" means "the successor of n." Thus, for addition Peano provided the two recursion equations $a + 0 = a$ and $a + n' = (a + n)'$. Rewriting the second of these as $a + (n + 1) = (a + n) + 1$, we can see that we have here a particular case of the associative law for addition, $x + (y + z) = (x + y) + z$, which can in fact be derived from the recursion equations by means of axiom (5). Multiplication is defined in similar fashion by means of the equations $a \cdot 0 = 0$ and $a \cdot b' = a \cdot b + a$, and once more familiar arithmetical laws can be extracted by means of induction.

With the assistance of a number of colleagues, including Cesare Burali-Forti, Peano succeeded in reformulating much of existing mathematical theory in accordance with his criteria of rigor and precision, the results of these investigations appearing in the journal *Rivista di Matematica* (later also *Revue de mathématiques* and *Revista de mathematica*) from 1891 to 1906 and in Peano's *Formulaire de mathématiques* (5 vols., Turin 1892–1908). The detailed coverage of algebra, arithmetic, set theory, geometry, and other branches of mathematics argues convincingly for Peano's approach, but it is questionable whether it vindicates a formalist philosophy of mathematics, since further metamathematical investigation, notably by Thoralf Skolem, has shown that if Peano's axioms are embedded in an axiomatization of set theory, they do not serve to characterize the natural numbers to the exclusion of other progressions. At the same time, it should be noted that Peano was not himself concerned with advancing either a formalist or a logicist philosophy; his approach was determined by a desire for

technical improvements in the presentation of mathematics.

See also Mathematics, Foundations of; Peano, Giuseppe; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bede Rundle (1967)

WHITEHEAD AND RUSSELL. In *The Principles of Mathematics*, published in 1903, Bertrand Russell (1872–1970) set out to establish the logicist view that "all pure mathematics deals exclusively with concepts definable in terms of a very small number of fundamental logical concepts, and that all its propositions are deducible from a very small number of logical principles" (2nd ed., p. xv) and also to explain "the fundamental concepts which mathematics accepts as indefinable" (ibid.). In the *Principles* this program is pursued with minimal recourse to symbolism, the systematic formal presentation being reserved for a proposed second volume. What in fact appeared as the sequel was the classic *Principia Mathematica* (3 vols., Cambridge, U.K., 1910–1913), written in collaboration with Alfred North Whitehead. The subject matter of *Principia Mathematica* considerably overlaps that covered by Frege in his *Grundgesetze der Arithmetik*, a work to which the authors acknowledge their chief debt on questions of logical analysis; in some respects, such as the demarcation between logical and metalogical theses, *Principia Mathematica* falls short of the standards of rigor observed in Frege's masterpiece. The symbolism adopted in *Principia Mathematica* derives largely from Peano, and the development of arithmetic and the theory of series is based on the work of Cantor.

We shall concentrate on the most important feature distinguishing *Principia Mathematica* from Frege's work, the attempt to avoid the contradictions which Russell found implicit in the fifth axiom of the *Grundgesetze*. This axiom licensed the transition from a concept to its extension and from an extension to the concept, a transition that appears to do no more than give formal expression to a platitude. For instance, the proposition "Stravinsky is a member of the class of composers" appears to be no more than a circumlocution for "Stravinsky is a composer." In general, it would seem reasonable to lay down as a law that x is a member of the class of ϕ's if and only if x is ϕ—in Russellian notation, $x \epsilon \hat{z}(\phi z) . \equiv . \phi x$. But despite its platitudinous appearance, this principle turns out to harbor a contradiction, since corresponding to the concept "is not a member of itself" we have the class of all such things—that is, the class of all classes which are not members of themselves—and if we now ask whether this class is or is not a member of itself, we find that either way a contradiction arises: If it is a member of itself, then it satisfies the defining condition of such members, so it is not a member of itself, and if it is not a member of itself, it belongs to the class of such classes and so is a member of itself.

This contradiction was noted by Russell in 1901, and in subsequent years finding ways to avoid it formed one of his major concerns. His final analysis, incorporated into *Principia Mathematica*, attributed the contradiction, along with a number of analogous paradoxes, to a mode of reasoning involving a vicious circle, a circle that arises when we postulate a collection of objects containing members definable only by means of the collection as a whole. Russell regarded such collections as illegitimate totalities, to be avoided by observing his "vicious-circle" principle, "Whatever involves *all* of a collection must not be one of the collection." Appealing to this principle, Russell claimed that the values of a function cannot contain terms definable only by means of the function, and in place of an indiscriminate application of functions to arbitrary arguments he defined an ascending hierarchy of types, beginning with individuals and progressing through functions of individuals, functions of functions of individuals, and so forth, the only arguments which a function can significantly take being those of the immediately preceding type. In particular, a class cannot significantly be taken as an argument to its defining function, and the derivation of Russell's paradox is accordingly obstructed by ruling out both "$x \epsilon x$" and its negation as ill-formed.

Apart from enabling us to block the derivation of paradoxes, Russell claimed, the theory of types based on the vicious-circle principle has a certain consonance with common sense. However, the principle itself (in the various nonequivalent forms given by Russell) can be challenged on the ground that it rules out circular procedures which are in no way vicious.

If the vicious-circle principle is rejected, it is natural to regard Russell's paradox as no more than a straightforward contradiction, the absurdities resulting from the abstraction schema $(\exists x)(y)(y \epsilon x \equiv \phi(y))$ being no different in kind and requiring no different an explanation from those yielded by $(\exists x)(y)(Fyx \equiv \phi(y))$, where the membership relation is replaced by an arbitrary dyadic predicate. On this view the problem of finding consistent instances of the abstraction schema reduces to the analogous problem for the uninterpreted version, but although such an approach has its merits, it loses sight of an important feature of the system which the vicious-circle

principle shapes via the theory of types. That is, the form of theory which the principle determines conforms to a natural conception of classes according to which they are, or at least could be, generated by a step-by-step procedure, the superstructure of classes of classes of classes, and so on, resting ultimately on the initial elements of lowest type. On the other hand, although it is natural to conceive of a domain of classes as initially secured by such a procedure, it would seem equally natural to relax this constructivist approach to the extent of allowing the specification of particular classes in the domain to proceed by characterizations in terms of the given totality, provided only that the consequent reflexivity does not embody a contradiction.

See also Cantor, Georg; Frege, Gottlob; Logical Paradoxes; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bede Rundle (1967)

POST. Besides provoking reactions in the form of rival philosophies of mathematics, the work of Whitehead and Russell stimulated new technical developments. For example, although Whitehead and Russell made free use, in *Principia Mathematica*, of the notions of truth-value and truth-function, they failed to incorporate these notions into a systematic technique for evaluating formulas of the propositional calculus. Such a technique, the method of truth tables, was presented by Emil Post (1897–1954) in his dissertation of 1920, published as "Introduction to a General Theory of Elementary Propositions" in the *American Journal of Mathematics* (43: 163–185) in 1921, the year in which Wittgenstein independently presented the same method in his *Tractatus Logico-Philosophicus*. The method dates back, in fact, to Peirce, but Post considered truth tables in their application not only to classical logic but also to systems in which any number of values are allowed, the primitive connectives of *Principia Mathematica*, "$\sim$" and "$\vee$," having in these systems the generalized analogues "$\sim_m$" and "$\vee_m$," where $\sim_m P$ takes the values $t_2, t_3, \cdots, t_m, t_1$ as P takes the values $t_1, t_2, \cdots, t_m$, and $P \vee_m Q$ takes that of the two values assigned to P and Q which bears the lesser subscript. Classical two-valued logic is accordingly a particular case of the many-valued logics so constructed. Post provided definitions of consistency and completeness, and for the first time a formulation of the propositional calculus was proved to have these properties, the method of truth tables providing a basis for the proofs.

In his 1920 dissertation Post showed how both truth tables for classical logic and associated postulate sets may be generalized. These postulate sets were treated as uninterpreted formal systems, an approach which Post maintained and extended in the direction of even greater generality in later works, where the derivation of theorems from axioms is represented as the production of strings—that is, finite sequences of symbols—from certain other strings of specified form. Most mathematical theories can be transcribed into the canonical forms admitted by Post, and he was able to show that the rules of any theory so expressed can be reduced to productions of a particularly simple type, a reduction that greatly simplifies investigations into the syntax of formal systems.

This approach leads directly to a formulation of recursive enumerability (a set is recursively enumerable if its members can be generated as the values of an effectively calculable function) and thence to one of recursiveness (a set is recursive if both it and its complement are recursively enumerable); Post provided illuminating proofs of results concerning decidability and related topics and introduced and developed a number of important concepts in this field. In 1947 he showed the recursive unsolvability of the word problem for semigroups. That is, he proved that it is impossible to determine whether or not two arbitrarily given strings are equivalent (where A and B are equivalent if B can be obtained from A by starting with A and applying a finite sequence of specified operations prescribing the production of one string from another). This result, published independently and in the same year by A. A. Markov, is an interesting example of the resolution, by techniques of mathematical logic, of an outstanding problem in the field of mathematics proper.

See also Russell, Bertrand Arthur William; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bede Rundle (1967)

RAMSEY. Frank Plumpton Ramsey (1903–1930), a brilliant Cambridge philosopher and logician, attempted to give a satisfactory account of the foundations of mathematics in accordance with the method of Frege, Russell, and Whitehead, defending their view that mathematics is logic while proposing revisions in the system of *Principia Mathematica* suggested by the work of Wittgenstein.

According to Russell, pure mathematics consists of "the class of all propositions of the form 'p implies q' where p and q are propositions containing one or more variables, the same in the two propositions, and neither p nor q contains any constants except logical constants" (*The Principles of Mathematics*, p. 3). Ramsey agreed with this definition insofar as it characterizes the generality

that is a feature of pure mathematics, but he claimed that it takes no account of an equally important mark of mathematics, its tautological character. The term *tautological* in the relevant sense derives from Wittgenstein, who applied it to formulas of the propositional calculus which come out true no matter what combinations of the values *true* and *false* are assigned to the component propositions. Ramsey extended the term to apply to valid formulas of the predicate calculus. Thus, the formula "$(x) . \phi x : \supset : \phi a$" is tautological, since "ϕa" expresses one of the possibilities which go to make up the possibly infinite conjunction abbreviated by "$(x) . \phi x$."

Admittedly we cannot write down the fully expanded versions of quantified formulas, but this inability does not affect the tautological character of truths formulated in the compressed notation. Similarly, Ramsey maintained, the inability of human beings to list the members of an infinite class is no bar to our conceiving of classes whose members could be indicated only in this way and not via the specification of a defining predicate. Indeed, the possibility of such indefinable classes is an essential part of the extensional attitude of modern mathematics, and Ramsey regarded the neglect of this possibility in *Principia Mathematica* as one of the work's three major defects. Thus, as interpreted in the system of *Principia Mathematica* the multiplicative axiom (axiom of choice) is logically doubtful, but on an extensional view of classes it is, according to Ramsey, an evident tautology.

The second major defect that Ramsey found in *Principia Mathematica* concerns Russell's attempt to overcome the paradoxes, in particular his postulation of the axiom of reducibility. Ramsey accepted the simple theory of types as an unquestionably correct measure for avoiding the logical contradictions, such as Russell's paradox and the Burali-Forti paradox, but he claimed that the contradictions that the hierarchy of orders had been introduced to avoid are of no concern either to logic or to mathematics. These contradictions—for instance, the Richard paradox and Weyl's contradiction concerning the word *heterological*—cannot be stated in logical terms alone but contain some further reference to thought, language, or symbolism. Rejecting Russell's conception of orders, Ramsey put forward a less restrictive theory based on his extensional view of propositional functions. Just as "$(x) . \phi x$" represents an infinite conjunction of atomic propositions "$\phi a . \phi b . \cdots$" so "$(\phi)\phi a$" expands to "$\phi_1 a . \phi_2 a . \cdots$" and similarly with disjunctions replacing conjunctions for existential quantifiers. Accordingly, if we start with truth-functions of atomic formulas, then no matter how often or in what respect we generalize upon them, we shall never pass to propositions significantly different from these elementary truth-functions; the only difference will lie in the notation introduced with the quantifiers. There is consequently no need for the axiom of reducibility—which, Ramsey claimed, could anyhow be false—and although the resultant theory countenances definitions of propositions in terms of totalities to which they belong, such definitions are in Ramsey's eyes no more vicious than an identification of a man as the tallest in a group of which he is a member.

The third great defect in *Principia Mathematica* which Ramsey proposed to rectify concerns Russell's definition of identity, according to which it is impossible for two objects to have all their properties in common. Ramsey held that this consequence shows that identity has been wrongly defined, and he advanced a definition of "$x = y$" designed to render the phrase tautological when x and y have the same value and contradictory otherwise.

See also Frege, Gottlob; Identity; Logical Paradoxes; Ramsey, Frank Plumpton; Russell, Bertrand Arthur William; Types, Theory of; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bede Rundle (1967)

BROUWER AND INTUITIONISM. The intuitionist conception of mathematics was developed by the Dutch mathematician Luitzen Egbertus Jan Brouwer (1881–1966). According to Brouwer mathematics is not a system of formulas and rules but a fundamental form of human activity, an activity that has its basis in our ability to abstract a conception of "twoness" from successive phases of human experience and to see how this operation may be indefinitely repeated to generate the infinitely proceeding sequence of the natural numbers. In the system of mathematics based on this primordial intuition, language serves merely as an aid to memory and communication and cannot of itself create a new mathematical system; our words and formulas have significance only insofar as they are backed by an essentially languageless activity of the mind. In particular, the wording of a theorem is meaningful only if it indicates the mental construction of some mathematical entity or shows the impossibility of the entity in question. Brouwer's conception of proof as essentially mental is useful as a corrective to a narrow formalist account that would construe proof as proof in a given formal system, although his psychologism is philosophically questionable—Wittgenstein's work has rendered more than doubtful the thesis that language is only an incidental accompaniment to thought, required solely for purposes of memory and

communication. What is important in intuitionism is not so much its psychologistic features as its emphasis on constructibility and the form of mathematics which its criterion of meaningfulness determines.

Implicit in classical mathematics is the notion that to know the meaning of a statement it is sufficient to know the conditions under which the statement is true or false, even though these conditions may be such that we could never be in a position to determine whether or not they held. The possibility of a gap between what can be meaningfully stated and what can be recognized either as true or as false is not admitted by the intuitionists. On their theory we can know the meaning of a statement only when we can recognize a proof of it; indeed, to understand a statement simply is to know what constitutes a proof or verification of that statement.

This emphasis on verification leads to an explanation of the logical constants and of a number of mathematical concepts that results in the rejection or reinterpretation of large parts of classical mathematics. Thus, whereas in classical mathematics the truth-table definition is adequate to giving the meaning of the constant "$\vee$" ("or"), for the intuitionist we can explain the meaning of a statement of the form "$A \vee B$" only by indicating under what conditions we should be warranted in asserting such a statement. These conditions are that we should be warranted in asserting A or that we should be warranted in asserting B, and it is clear that neither condition may hold, even when A is the negation of B.

Assume, for instance, that A is an existentially quantified statement, $\exists xP(x)$, with the quantifier ranging over the natural numbers. To suppose that this holds is to suppose that we can actually construct a number with the required property. On the other hand, what is it to suppose that $\exists xP(x)$ is false? It cannot mean that a case-by-case examination of the numbers will provide a refutation of the statement, since a case-by-case investigation of an infinite totality is not a real possibility—it is a picture to which the classical mathematician is wedded by a mistaken analogy with finite totalities. But if $\sim\exists xP(x)$ is to have a meaning which we can grasp, it can mean only that there is a contradiction in the idea of a number's having the property P. Given this explanation of the sense of the proposition and its negation, we are obliged to abandon Aristotelian logic as no longer trustworthy in this context, for asserting the disjunction $\exists xP(x) \vee \sim\exists xP(x)$ is tantamount to asserting that we either are in a position to construct a suitable number or can show the impossibility of such a construction. We are not entitled to assert a priori that at least one of these possibilities must obtain, but to do so would simply be to commit ourselves to the unfounded belief that all mathematical problems are solvable.

This insistence on the identification of existence with constructibility can be traced back to Leopold Kronecker (1823–1891), and a precise formulation of principles of intuitionist logic was carried out in 1930 by a pupil of Brouwer's, Arend Heyting (1898–1980). Several branches of mathematics have been redeveloped from the intuitionist standpoint, but the reconstructions are often complicated, and in some cases, particularly where set-theoretic notions are involved, there has been a question of outright rejection, rather than reconstruction, of classical mathematics. Thus, impredicative definitions, hierarchies of transfinite numbers, and nonconstructive postulates such as the axiom of choice (and hence the well-ordering theorem), while important classically, are rejected in toto by the intuitionists, a rejection which has led many mathematicians to discount the claims of intuitionism without giving sufficient attention to the arguments, admittedly often obscurely expressed, on which they are based.

See also Brouwer, Luitzen Egbertus Jan; Intuitionism and Intuitionistic Logic.

Bede Rundle (1967)

HILBERT AND FORMALISM. The leading exponent of the formalist philosophy of mathematics was David Hilbert (1862–1943), who pioneered in a development of logic known as proof theory or metamathematics. From the time of his first papers on the foundations of mathematics, Hilbert stressed the importance of the axiomatic method and its superiority over the genetic approach, by which concepts are extended piecemeal as the need arises. Once a theory is axiomatized, however, it invites a number of general questions concerning the logical relations holding between its propositions, and Hilbert was soon to consider as central among such questions the problem of establishing consistency, or freedom from contradiction. Hilbert did not himself think that there was any support for the allegations of inconsistency in analysis, as made by Hermann Weyl. Nevertheless, he wished to consolidate once and for all the foundations of mathematics and to give them such clarity that the axiom of choice would be as perspicuous as the simplest arithmetical truth. To this end he needed to devise consistency proofs. He had, in 1899, shown the consistency of Euclidean geometry relative to the theory of real numbers, but proofs of this form do no more than shift the problem of consistency to the

system to which the original theory has been reduced. Some new, more direct method seemed to be called for.

Despite his confidence in the consistency of classical mathematics, Hilbert contended that operating in an abstract way with mathematical concepts had proved insecure, and his remedy was to interpret number theory as relating to the observable domain of such signs as 1, 11, 111. Elementary number theory is thereby assured of a concrete interpretation—"3 > 2," for example, can be understood as asserting that the concatenation of three strokes extends beyond the concatenation of two strokes. However, the possibility of such an interpretation does not extend to all branches of classical mathematics, for such entities as transfinite cardinals do not allow of representation as sequences of strokes.

Hilbert's solution to this difficulty was to treat such numbers as "ideal" elements. Thus, appealing to Kant, he argued that one precondition for the application of logical laws is a domain of extralogical concrete objects, given in actual perception and capable of being exhaustively surveyed. Nowhere in nature is an actual infinity to be found; therefore, whereas for finite numbers a perceptually given basis could be given, transfinite numbers had a place in mathematics only as ideal elements, much like the ideal factors introduced to preserve the simple laws of divisibility for algebraic whole numbers. Such a reduction was, Hilbert claimed, a natural extension of the work of Weierstrass, who had shown that reference to infinity in the context of calculus involved merely a *façon de parler*, replaceable by a theory of limits requiring a potential infinite rather than an actual one. Similarly, the infinities introduced by Cantor, though apparently irreducible, had to be shown to be indispensable, and arguments proceeding via the infinite had to be replaced by finite methods that achieve the same goal. Again, since the transfinite enters with the use of unbounded quantifiers, statements containing these had to be regarded as ideal statements.

With this approach Hilbert hoped to partially vindicate classical mathematics against the attacks of the intuitionists. Complete vindication, however, required a proof of consistency, and the method that Hilbert proposed for obtaining such a proof is closely related to his method for providing elementary number theory with a sound basis. That is, just as he had considered numbers as sequences of strokes, so he now regarded formulas and proofs as sequences of uninterpreted signs. In this way he provided a concrete subject matter for a proof of consistency, a proof that was to invoke only logical principles whose security and perspicuity are equal to the security and perspicuity of the perceptually given domain on which they are to operate.

Thus, the consistency of some given formalization of a branch of mathematics could be unquestionably established if it could be shown by finite combinatorial methods that no manipulation of the symbols which represents a passage from axioms to theorems could result in the derivation of the expression "0 = 1" or of some other concatenation of symbols which, when interpreted, is seen to be an absurdity. The theory itself might contain symbols for transfinite cardinals and other ideal elements, but this would be no obstacle to a consistency proof, since in such a proof we are required only to treat these symbols as perceptually given objects and to show that they will never figure in a formula whose negation is also provable. On the other hand, Hilbert believed that although nonfinitary concepts are allowable within mathematics proper, they are not to be countenanced in the theory of proof that is to ensure consistency.

The formalist school, which included Wilhelm Ackermann, Paul Bernays, and John von Neumann, succeeded in establishing a number of metamathematical results of considerable significance, but without completing Hilbert's original program, for although successively stronger systems of arithmetic were proved consistent, no proof was forthcoming for the full system required by classical number theory. And, indeed, results obtained by Kurt Gödel in 1931 indicate that no finitary consistency proof is possible, since any proof of consistency must make an appeal to principles which are more general than those provided by the system and accordingly are as much open to question as those principles whose consistency we wish to establish. Attempts were subsequently made to prove consistency by means which were as close to being finitary as possible, notably by Gerhard Gentzen in 1936, but even if "finitary" were thought to apply to the methods used—in this case an application of transfinite induction—it would not follow that classical mathematics had been vindicated against the intuitionists, since to their way of thinking the mere consistency of mathematics would not suffice to confer a clear meaning on the crucial concepts of classical mathematics.

See also Gödel's Incompleteness Theorems; Hilbert, David; Kant, Immanuel; Mathematics, Foundations of; Neumann, John von; Proof Theory; Weyl, (Claus Hugo) Hermann.

Bede Rundle (1967)

LÖWENHEIM. A number of significant results concerning the first-order functional or predicate calculus (with

identity) date from a paper published in 1915 by Leopold Löwenheim (1878–1957), a mathematician of Schröder's school. In this paper, "Über Möglichkeiten im Relativkalkül" (*Mathematische Annalen* 76 [1915]: 447–470), Löwenheim showed how the problem of deciding the validity of formulas in this calculus reduces to the problem of determining the validity of formulas in which only two-place predicate letters occur. Since (from the point of view of decidability) such formulas are accordingly no less general than arbitrary formulas of the calculus, we know from a later result, by Alonzo Church, that the decision problem for this class is unsolvable. However, Löwenheim was able to provide a decision procedure for a more restricted class of formulas, those in which only one-place predicate letters occur. He also showed that no formula of this restricted class could be valid in every finite domain, yet not be valid in an infinite domain, and his most famous result, known as Löwenheim's theorem, states that any formula of the full calculus which is valid in a denumerable domain is valid in every nonempty domain.

Although it is not difficult to show that if a formula is valid in a given domain, it is valid in any smaller domain, we cannot in general claim that validity in a given domain establishes validity in a larger domain. But as Löwenheim recognized, a formula may be valid in every domain comprising only finitely many of the natural numbers, yet not be valid in the domain of all natural numbers. The significance of Löwenheim's result is thus that validity in a denumerable domain guarantees validity not simply in any smaller domain but in domains which, like that of the real numbers, are of even greater cardinality than the set of natural numbers.

Bede Rundle (1967)

SKOLEM. The Norwegian mathematician Thoralf Skolem (1887–1963) made extensive contributions to the development of logic, maintaining a steady output of important papers from 1920 until his death. Skolem's first major result was an extension of the above-mentioned theorem of Löwenheim that if a formula of the first-order functional calculus (with identity) is valid in a denumerably infinite domain, it is valid in every nonempty domain and that, equivalently, if such a formula is satisfiable at all, then it is satisfiable in a domain comprising at most a denumerable infinity of elements. In 1920, Skolem generalized this theorem to the case of classes (possibly infinite) of formulas, establishing that if a class of formulas is simultaneously satisfiable, then it is satisfiable in a denumerably infinite domain. Skolem's proof makes use of the axiom of choice and the Skolem normal form of a formula—a type of prenex normal form in which no universal quantifier precedes an existential quantifier—but both these devices were subsequently dropped, and a more constructive version of the proof was given in 1928, a version which led to the developments of Herbrand and to Gödel's completeness proof.

Skolem was led by his work on Löwenheim's theorem to consider set-theoretic concepts as in a certain sense relative. This view derives from the fact that suitable axiomatizations of set theory can be written in the notation of first-order logic, the only symbol foreign to this logic—the epsilon of membership—being replaced by a dyadic predicate letter. The result is a set of formulas which, if consistent, has by Löwenheim's theorem an interpretation within a denumerably infinite domain. At the same time, within the system of set theory we can establish, by Cantor's theorem, the existence of nondenumerably infinite sets. This apparent conflict between the magnitude of the sets in the axiomatic theory and the more limited domain in which it is modeled is known as the Löwenheim–Skolem paradox. Skolem's way out of this paradox was to suggest that the distinction between denumerable and nondenumerable be taken as relative to an axiom system, a set which is nondenumerable in a given axiomatization perhaps being denumerable in another.

The possibility of an enumeration not available within the original axiom system has led to the description of Löwenheim's theorem as the first of the modern incompleteness theorems, but Skolem's resolution of his paradox does not represent the only possibility. In the first place, it is not clear how the required enumeration could be devised even outside the system in question. To take an analogous case, Cantor's theorem shows that the members of a set containing three elements cannot be paired off with the members of the power set of this set. Since the power set in this case contains eight elements, Cantor's result is in no way surprising, but there is no inclination to say that further mappings might be devised which would yield a one-to-one correspondence between the three-member set and the eight-member set. In the second place, Löwenheim's theorem does not require us to suppose that the axiomatized theory guarantees an enumeration of the sets, since the reinterpretation of the original symbolism with respect to a denumerable domain results in a revision of the propositions implying or asserting the existence of a nondenumerable infinity of sets. By hypothesis, such propositions go over into propositions which hold in the denumerable model, but although their truth is preserved, their original meaning

is altered: they could not without contradiction assert the nondenumerability of the new model.

The set-theoretic relativism that Skolem inferred from the Löwenheim–Skolem theorem led him to doubt whether mathematical concepts could be completely characterized axiomatically, and in 1934 he published a result confirming these doubts by demonstrating that no categorical system of postulates for the natural numbers can be expressed in the notation of quantification theory. Any attempt to give a unique characterization of the natural numbers by means of propositions expressed in this notation is bound to fail, even if a denumerable infinity of such propositions is allowed, since there will always be other systems of entities conforming to the structure so defined. Although this result was uncongenial to those who had hoped to delineate the numbers from a formalist standpoint, the nonstandard models which are yielded by such proofs have become increasingly important, and their application to such topics as independence proofs and mathematical analysis promises to be fruitful.

Skolem also made important contributions to the theory of recursive functions. His work in this field dates from a pioneering paper of 1923, in which he sought to develop arithmetic in a logic-free calculus. Essentially this meant the elimination of quantifiers, an elimination that Skolem proposed to effect by the extensive use of recursive definitions. For instance, instead of defining "$a < b$" as "$(\exists x)(a + x = b)$," we can avoid the use of the existential quantifier by means of the joint stipulation of (*i*) $-(a < 1)$ and (*ii*) $a < (b + 1) \leftrightarrow (a < b) \vee (a = b)$. In this and subsequent papers Skolem advanced such reductions as part of a finitistic program for securing the basis of arithmetic.

Also important are Skolem's contributions to set theory. The Zermelo–Fraenkel system is commonly presented with his modifications, and in his last years he took up the study of set-theoretic contradictions from the standpoint of systems of many-valued logic.

See also Cantor, Georg; Many-Valued Logics; Set Theory.

Bede Rundle (1967)

HERBRAND. Despite a tragically short life—he was killed in a mountaineering accident in 1931 at the age of twenty-three—Jacques Herbrand made substantial contributions to the development of mathematical logic, especially to investigations in the metatheory of logic that were the particular concern of Hilbert and his school. The bulk of Herbrand's contributions is to be found in his University of Paris dissertation of 1930, *Recherches sur la théorie de la démonstration* (published in *Travaux de la Société des Sciences et des Lettres de Varsovie*, Classe III (33) [1930]: 33–160). This work has much in common with the later "Untersuchungen über das logische Schliessen" of Gerhard Gentzen, but the presentation of proofs and explanations is much less perspicuous than Gentzen's, and even now some aspects of Herbrand's work await further clarification and elaboration.

Herbrand's starting point is the system of classical propositional logic presented in Whitehead and Russell's *Principia Mathematica*, but the extension of this to functional calculi of first and higher orders is effected by the addition of further rules in place of axioms. The resultant calculi, in which mathematical theories may be embedded, are investigated from a Hilbertian proof-theoretic viewpoint, with emphasis on such syntactic notions as derivability and to the exclusion of semantic questions that cannot be given a finistic interpretation. New proofs are given of a number of results already known, such as those concerning solvable cases of the decision problem, and for the first time the idea and proof of the deduction theorem is presented for a particular system of logic. That is, Herbrand showed that a necessary and sufficient condition for the derivability of a proposition P in his theory with hypotheses H is that $H \supset P$ should be derivable in the logic without hypotheses.

Herbrand's most powerful result concerns the necessity and sufficiency of certain conditions for the provability of a quantificational schema. He showed, in fact, that such a schema is provable if and only if a quantifier-free tautology of a prescribed form is constructible from it. The proofs of the various theses that go to make up this result are somewhat complicated, but the form of tautology which is associated with a provable formula can be indicated in the following way: First, a given quantificational schema is so transformed that each quantifier has its minimum scope or, alternatively, each has its maximum scope. Taking just the first case, then, all the quantifiers are placed initially and have a scope that extends to the far right of the formula.

Suppose we are given a schema in this form—for example, $\exists x\, \exists y\, \forall z\, [Fx \rightarrow (Fy\ \&\ Fz)]$—which we shall call "$A$." The necessary and sufficient condition of A's holding is that it be false that for any x and y there is a value of z for which the matrix of A is false. Accordingly, if x and y both take the value a_1, say, there must be some value a_2 of z that results in the falsity of the matrix if A is to be false; further, if x takes the value a_1 and y takes the value a_2, then for some value a_3 of z the matrix must be false; again, if x takes the value a_2 and y takes the value a_1, the matrix must

be false for some value a_4 of z, and so on. But if we find that at least one of the substitution instances of the matrix so generated must be true, we have shown the failure of a necessary condition for the falsity of A. In fact, this is the outcome of the present example, since the disjunction of the cases so far considered, $[Fa_1 \rightarrow (Fa_1 \& Fa_2)] \vee [Fa_1 \rightarrow (Fa_2 \& Fa_3)] \vee [Fa_2 \rightarrow (Fa_1 \& Fa_4)]$, is a tautology—thus, if Fa_2 is true, the first disjunct is true, and if Fa_2 is false, the last disjunct is true.

Herbrand showed how such disjunctions can be constructed from a formula in prenex normal form with the quantifiers occurring in any number and order. He showed, too, that the original formula can be retrieved from such a disjunction by the application of a few simple rules, without use of *modus ponens*. And, indeed, it is clear from the example given that the only rules required for the derivation of the original formula from the tautology are rules allowing for the insertion of quantifiers before the disjuncts and a rule allowing us to erase repetitions of identical disjuncts. The final result allows us to assert that the constructibility of a tautologous disjunction is both a necessary and a sufficient condition for the provability of the associated quantified schema.

In addition to shedding considerable light on the structure of quantification theory, Herbrand's theorem is the source of a number of important metatheoretic results. Löwenheim's theorem is an immediate consequence—accepted by Herbrand only when reinterpreted finitistically—and certain cases of the decision problem allow of simple resolution. Important for Herbrand's aims was the application of his theorem to the question of the consistency of arithmetic, and he was able to show that if we have a model for a set of hypotheses, an interpretation with respect to some domain under which all these hypotheses are true, then no contradiction can arise in the theory deduced from the axioms. Suppose hypotheses $H_1, H_2, \cdots, H_n$ give rise to a contradiction while having a true interpretation within the model. Since $H_1, \& H_2 \& \cdots \& H_n$ comes out true in the model, the model brings the negation of this conjunction out false, and if a formula is false in some domain, it is not associated with a quantifier-free tautology. If, on the other hand, $H_1, H_2, \cdots, H_n$ yield a contradiction, then $\sim(H_1 \& H_2 \& \cdots \& H_n)$ is provable and thus is associated with a tautologous disjunction. This form of consistency proof was discussed further by Herbrand in his later article "Sur la Non-contradiction de l'arithmétique" (*Journal für die reine und angewandte Mathematik* 166 [1931]: 1–8), and the same idea appears in Gentzen's "Untersuchungen über das logische Schliessen."

See also Quantifiers in Formal Logic; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bede Rundle (1967)

GÖDEL. Kurt Gödel (1906–1978), a major figure in the history of logic, is best known for his celebrated incompleteness theorem presented in "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I" (*Monatshefte für Mathematik und Physik* 38 [1931]: 173–198) and his associated proof of the impossibility of establishing the consistency of customary formulations of arithmetic by methods formalizable within the systems themselves. In addition to these results (discussed in the entry Gödel's Theorem), Gödel made important contributions to several other branches of logic, and prior to his 1931 paper he had already presented the first completeness proof for the first-order functional calculus (in "Die Völlstandigkeit der Axiome des logischen Funktionkalküls," *Monatshefte für Mathematik und Physik* 37[1930]: 349–360). Making use of a normal form devised by Thoralf Skolem, Gödel elaborated a proof along lines that were followed by Jacques Herbrand to a similar end in a publication of the same year (*Recherches sur la théorie de la démonstration*, in *Travaux de la Société des Sciences et des Lettres de Varsovie*, Classe III [33] [1930]: 33–160), but he went further than Herbrand in his method for showing how any unprovable formula may be falsified.

Intuitionistic as well as classical logic has been one of Gödel's major concerns, and his results in this field are of importance to an understanding of the formalizations of this logic initiated by Arend Heyting in 1930. The intuitionist propositional calculus is naturally thought of as a subsystem of classical logic, obtained by omitting from the latter those theses that are intuitionistically unacceptable. Gödel indicated that this picture could in a sense be reversed, since it is possible to define all two-valued truth-functions by means of the connectives for negation and conjunction, and he was able to show that any formula involving only these connectives is provable within intuitionistic logic if it is provable classically. Gödel showed, further, that even classical number theory, if suitably interpreted, can be thought of as included within intuitionistic number theory. He also proved that the intuitionist propositional calculus has no finite characteristic matrix. That is, although the two-valued truth tables for classical logic serve to verify all and only those theses provable in this logic, it is impossible, according to Gödel's result, to devise truth tables having any finite number of values that will perform the same service for the intuitionist system.

Two propositions that have been at the center of much investigation and controversy are the axiom of choice and Cantor's generalized continuum hypothesis. It was Gödel who proved that both are consistent with the axioms of set theory provided only that these axioms are themselves consistent. The axiom of choice is a highly nonconstructive axiom licensing the selection of an unspecified element from each of a (possibly infinite) family of sets and the formation of a set comprising just the elements so selected. The generalized continuum hypothesis, which in fact implies the axiom of choice, states that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$; that is, starting with $\aleph_0$, which is the number of the natural numbers, the series of increasingly higher cardinals is successively generated by raising 2 to the power of the preceding aleph. The system Σ of set theory that Gödel used derives from John von Neumann and Paul Bernays. Gödel showed that if it were possible to derive a contradiction from the axiom of choice and the continuum hypothesis in Σ, then the axioms of Σ alone would suffice for the derivation of a contradiction. This result is obtained by constructing a model Δ within Σ itself, where Δ is such that the propositions asserting that the axioms of Σ hold for Δ are demonstrable in Σ and the similar relativizations to Δ of the axiom of choice and the generalized continuum hypothesis are likewise demonstrable in Σ. Paul J. Cohen showed, in 1963, that the negations of these propositions are also consistent with the axioms of set theory. In other words, the axiom of choice and the generalized continuum hypothesis are now known to be independent of the other axioms of set theory.

See also Cantor, Georg; Gödel, Kurt; Neumann, John von; Set Theory.

Bede Rundle (1967)

SINCE GÖDEL

The pace of development in logic picked up rapidly after Gödel's incompleteness theorems, and five branches emerged: set theory, model theory, proof theory, computability theory, and nonclassical logics.

Gödel's theorems were formulated for type theory, but this was soon displaced as the framework for mathematics by Zermelo-Frankel set theory with choice (ZFC). Gödel's theorems still apply, and imply the existence of set-theoretic statements that can be neither proved nor disproved. Gödel himself showed that Cantor's continuum hypothesis cannot be disproved, and conjectured that it cannot be proved, as was established in the 1960s by Paul Cohen. Since then the search for new axioms to settle questions left open by ZFC has flourished.

Gödel's results on the unprovability of the consistency of a formal theory within the theory itself were followed by Tarski's work on the undefinability of truth for a formal language within the language itself. Tarski's work also for the first time gave a rigorous definition, in a *meta*-language, of truth for a sentence of formal language, relative to an interpretation, which is needed for a fully rigorous statement even of Gödel's earlier *completeness* theorem. With his truth definition Tarski laid the foundations for a general theory of models, a model of a formal theory being an interpretation that makes it true.

Gödel showed the unachievability of the original aim of proof theory: to establish the consistency of infinitistic mathematics by finitist means; but this leaves open the possibility of establishing *relative* consistency through the interpretation of ostensibly stronger in ostensibly weaker theories. Gödel himself contributed to this program, and in the mid-1930s the powerful new methods were introduced by Gerhardt Gentzen (1909–1945).

Gödel used in his work the auxiliary notion of a primitive recursive function, which include many but not all functions that are effectively computable in an intuitive sense. Two equivalent proposed characterizations of the full class of effectively computable functions followed. Recursive function theory was developed in collaboration with his student S. C. Kleene (1909–1994) by Alonzo Church, who proved there is no effectively computable function that will tell whether a given formula is logically valid. Turing machines were developed by Alan Turing, who proved the possibility in principle of a universal programmable computer, a possibility that began to be realized during the Second World War.

Gödel contributed not only to the areas just enumerated, which together constitute *mathematical logic*, but also to the study of modal and other nonclassical logics, often called *philosophical logic*. Mathematical logic was characterized by explosive growth after 1945. Philosophical logic grew more slowly until the development of a usable model theory for nonclassical logics with the work of Saul Kripke and others circa 1960, after which development speeded up and important connections with theoretical computer science emerged.

Much of the growth in all five branches has occurred in areas far removed from philosophy, but if the volume of philosophically oriented work has decreased in relative terms, still it has increased in absolute terms owing to the overall growth of logic.

See also Cantor, Georg; Church, Alonzo; Gödel, Kurt; Kripke, Saul; Logic, Non-Classical; Tarski, Alfred; Turing, Alan M.

John P. Burgess (2005)

GENTZEN. The first systematic formulations of the propositional and predicate calculi were presented axiomatically, on the analogy of certain branches of mathematics. In 1934, Gerhard Gentzen (1909–1945), a logician of Hilbert's school, published a formalization of logical principles more in accordance with the way in which these principles are customarily applied. (A similar approach was developed independently by S. Jaśkowski; see below, section on Polish logicians.) In illustrating his technique Gentzen considered how we might establish as valid the schema $(X \vee (Y \,\&\, Z)) \supset (X \vee Y) \,\&\, (X \vee Z)$. Assuming that the antecedent holds, either X is true, or Y & Z is true. In the former case we can pass to each of $X \vee Y$ and $X \vee Z$ and hence to their joint assertion. Assuming now Y & Z, we may infer Y, whence $X \vee Y$, and likewise Z, whence $X \vee Z$. In this case the conjunction is once more derivable. Since it is derivable from each disjunct of the original assumption, we may assert the implication unconditionally.

In this simple form of argument the justification of the schema has been broken down into a series of uncomplicated steps, each involving either the introduction or the elimination of a logical connective. Extracting the rules that were applied and supplementing them with similar rules governing the use of the other connectives, we arrive at a system of "natural" deduction—either *NJ* (intuitionist logic) or *NK* (classical logic). Gentzen considered the former more natural than the latter, but whichever we opt for, it appears that the resultant codification of logical principles is more natural, on at least two counts, than a codification presented in axiomatic fashion.

In the first place, we avoid the devious moves that may be necessary to establish a logical principle from an axiomatic basis and follow more closely a pattern of reasoning that we should intuitively adopt. In the second place, the conception of logic as a system of axioms and theorems adjoined to some given subject matter appears inappropriate, since, in their application to, say, a branch of mathematics, principles of logic function not as true statements forming part of the theory in question but as rules of inference allowing us to establish relations of consequence between propositions of the theory.

In addition to the systems *NJ* and *NK*, Gentzen devised related formalizations of logic, the *L*-systems, in which derivable formulas are shown to possess a particularly direct form of proof. These systems contain the "cut" rule, a generalized form of *modus ponens* that, like *modus ponens*, has the disadvantage that we cannot work back from a schema to premises from which it could have been derived. However, although the cut rule is crucial in showing the equivalence of the *L*-systems with the earlier *N*-systems, Gentzen showed that the cut rule can be eliminated from any proof in the *L*-systems. This powerful metatheorem simplifies the reconstruction of proofs of valid formulas, yielding a decision procedure for the propositional fragments of *LJ* and *LK* and greatly facilitating the search for proofs in the full calculi. Gentzen further applied his *Hauptsatz* to proofs of consistency; in particular, he showed one formalization of arithmetic to be noncontradictory. The formalization in question does not contain a schema of unrestricted induction, but in later works Gentzen remedied this defect, overcoming the obstacle to such proofs presented by Gödel's results by making use of a principle of transfinite induction which cannot be reduced to ordinary induction within the system. It is a matter of controversy whether such a proof represents the attainment of one of Hilbert's goals, a finitary consistency proof for classical number theory.

See also Gödel, Kurt; Hilbert, David; Induction.

Bede Rundle (1967)

CHURCH. From the beginning of the twentieth century questions concerning the decidability of logical and mathematical theories have held a special interest for logicians, mathematicians, and philosophers. A number of important concepts and far-reaching results in this field have come from Alonzo Church (1903–1995), author of a definitive text on logic and noted writer on the history of logic.

The notion of decidability is not one which a beginner in mathematics could explicitly formulate, but both this and related notions, such as that of effective calculability, have a place in the description of the most elementary mathematical concepts. Often our understanding of a particular numerical predicate is inextricably tied to our ability to determine whether or not an arbitrary number satisfies that predicate, and in many cases terms expressing the result of a calculation or computation can be fully grasped only by one who has the ability to carry out the sorts of computation in question. Thus, with the division of numbers into odd and even there is intimately associated a technique for determining which of these predicates applies to an arbitrary whole number; similarly, a person's grasp of the concepts of sum and product is

measured by his ability to calculate sums and products. But although the grasp of concepts and the mastering of techniques may go hand in hand in many cases, the symbolism of arithmetic allows us to formulate propositions whose truth-value may resist determination by any obvious methods of computing or reasoning in general, a situation that frequently arises with the introduction of unrestricted quantification.

Consider, for instance, the proposition *P*, "There is at least one odd perfect number." A perfect number is a number that is equal to the sum of its divisors, itself excluded. Thus, 6 is perfect, being equal to the sum of 1, 2, and 3; so is 28, being equal to the sum of 1, 2, 4, 7, and 14. Like "*x* is odd," the predicate "*x* is perfect" is a decidable predicate, in the sense that given any number *n*, we can, after a finite number of steps, respond with an unambiguous *yes* or *no* to the question "Is *n* perfect?" But although both of the predicates entering into *P* are decidable, the infinitude of the positive integers is an obstacle to an immediate determination of the truth-value of *P* which would make use of the decidability of these predicates together with a case-by-case examination of the integers. Indeed, proposition *P*, along with Fermat's last theorem, Goldbach's conjecture, and many other propositions of elementary number theory, has as yet been neither proved nor disproved. Accordingly we may well wonder whether it is possible to devise a technique that, when applied to an arbitrary proposition of this class, would enable us to determine the truth or provability of the proposition. Now, for all we know, any one of these outstanding problems may eventually be resolved, but Church showed that no general technique could be devised which would allow us to ascertain in an effective manner the truth or provability of an arbitrary arithmetical proposition.

By a direct application of the method of diagonalization (a procedure whereby a hypothesized function is shown to differ from each member of a class of functions of which it must be a member if it is to exist), Church demonstrated not simply that such a technique has proved elusive but that the supposition of its existence involves an absurdity. In this respect arithmetic contrasts with the propositional calculus, but although the propositional calculus does have a decision procedure—the method of truth-tables—Church showed that the first-order functional calculus fares no better than arithmetic, it being impossible to find a method that allows us to recognize as provable or refutable an arbitrary formula of this calculus. It may prove—indeed, in many cases it has already been shown—that fragments of these systems are decidable, but Hilbert's aim of a general technique which would banish ignorance from mathematics appears to be unattainable.

In demonstrating his theorem Church was obliged to provide a formal counterpart of the intuitive notion of effective calculability, and he proposed that this notion be identified with that of recursiveness. The notion of a recursive function (of positive integers) was introduced by Gödel, acting on a suggestion of Herbrand, and was analyzed in detail by S. C. Kleene. A function is said to be (general) recursive (a generalization of the notion of primitive recursive) if, roughly speaking, its value for given arguments can be calculated from a set of equations by means of two rules, one allowing the replacement of variables by numerals, the other allowing the substitution of equals for equals. As Church remarks, the intuitive status of effective calculability rules out any complete justification of his proposal (since known as Church's thesis), but he adduces reasons for regarding the identification as plausible, and the plausibility of his thesis has subsequently been reinforced by the discovery that despite their apparent dissimilarity, various alternative attempts to characterize the intuitive concept have all proved equivalent to that of general recursiveness.

Thus, at the time Church put forward his thesis the Church-Kleene notion of λ-definability was already known to provide an equivalent, and Turing's "computability," Post's "1-definability" and "binormality," and Markov's "computability" provide alternatives defined with respect to machines and combinatorial operations. It should be mentioned, however, that Church's thesis has not met with universal support; a summary and criticism of a number of objections can be found in Elliot Mendelson's "On Some Recent Criticism of Church's Thesis" (in *Notre Dame Journal of Formal Logic* 4 [1963]: 201–205).

See also Church, Alonzo; Gödel, Kurt; Number; Turing, Alan M.

Bede Rundle (1967)

TURING AND COMPUTABILITY THEORY. In the late 1930s Alan M. Turing was one of the founders of computability theory. His main contributions to this field were published in three papers that appeared in the span of a few years, and especially in his ground-breaking 1936–1937 paper, published when he was twenty-four years old.

As indicated by its title, "On Computable Numbers, with an Application to the Entscheidungsproblem," Turing's paper deals ostensibly with real numbers that are

computable in the sense that their decimal expansion "can be written down by a machine." As he pointed out, however, the ideas carry over easily to computable functions on the integers or to computable predicates.

The paper was based on work that Turing had carried out as a Cambridge graduate student, under the direction of Maxwell Newman (1897–1984). When Turing first saw a 1936 paper by Alonzo Church, he realized at once that the two of them were tackling the same problem—making computability precise—albeit from different points of view. Turing wrote to Church and then traveled to Princeton University to meet with him. The final form of the paper was submitted from Princeton.

In a space of thirty-six pages, the paper manages to accomplish the following six goals:

(1) A formalization of what it means to be "calculable by finite means," in terms of an idealized computing device—now known of course as a Turing machine.

(2) The construction of a "universal computing machine," which when supplied with the "standard description" of a machine *M* on its tape, will simulate the operation of *M*.

(3) The proof of the unsolvability of the halting problem and proofs of the unsolvability of other problems, such as the problem of deciding, given a machine *M*, whether or not *M* will ever print the symbol 0.

(4) Three kinds of arguments for Turing's thesis, that is, the claim that his formulation in terms of machines is successful in capturing the idea of "processes which can be carried out in computing" (249). It should be noted that Kurt Gödel and others have found Turing's arguments here completely convincing.

(5) A proof of Church's theorem that David Hilbert's Entscheidungsproblem can have no solution, that is, the problem of deciding whether or not a given formula is derivable in the predicate calculus is unsolvable.

(6) An outline, in an appendix, of the equivalence of computability by Turing machines to computability as formulated by Church in terms of the λ-calculus. (This proof was given in further detail in Turing's 1937 paper, "Computability and λ-Definability.")

Turing's paper remains a readable introduction to his ideas. How might a diligent clerk carry out a calculation by following instructions? The clerk might organize the work in a notebook. At any given moment his or her attention is focused on a particular page. Following the instructions, he or she might alter that page, and then might turn to another page. And the notebook is large enough that he or she never comes to the last page.

The alphabet of symbols available to the clerk must finite; if there were infinitely many symbols, then there would be two that were arbitrarily similar and so might be confused. One can then, without loss of generality, regard what can be written on one page of notebook as a single symbol. And one can envision the notebook pages as being placed side by side, forming a paper tape, consisting of squares, each square being either blank or printed with a symbol. At each stage of the work, the clerk—or the mechanical machine—can alter the square under examination, can turn attention to the next square or the previous one, and can look to the instructions to see what part of them to follow next. Turing describes the latter as a change of state of mind.

Turing writes, "We may now construct a machine to do the work" (251). Such a machine is of course now called a Turing machine, a phrase first used by Church in his review of Turing's paper in *The Journal of Symbolic Logic*. The machine has a potentially infinite tape, marked into squares. Initially, the tape is blank. (Alternatively, if one wants to compute some function, the input word or number can be written on the tape.) The machine is capable of being in any one of finitely many states (the phrase *of mind* being inappropriate for a machine).

At each step of the calculation, depending on its state at the time, the machine can change the symbol in the square under examination at that time, can turn its attention to the square to the left or to the right, and can then change its state to another state.

The program for this Turing machine can be given by a table. Where the possible states of the machine are $q_1, \ldots, q_r$, each line of the table is a quintuple (q_i, S_j, S_k, D, q_m) that is to be interpreted as directing that whenever the machine is in state q_i and the square under examination contains the symbol S_j, then that symbol should be altered to S_k and the machine should shift its attention to the square on the left (if $D = L$) or on the right (if $D = R$), and should change its state to q_m. For the program to be unambiguous, it should have no two different quintuples with the same first two components. One of the states, say q_1, is designated as the initial state—the state in which the machine begins its calculation. If one starts the machine running in this state, it might (or might not), after some number of steps, reach a state and a symbol for which its

table lacks a quintuple having that state and symbol for its first two components. At that point the machine halts.

In particular, where 0 and 1 are among the symbols in the alphabet, the machine might run on and on, sometimes printing a 0 or 1 on the tape (besides printing whatever other markers are needed for the computation). In this way, the machine might generate an infinite binary string. One can interpret this binary string as giving the binary expansion of a real number in the unit interval. Say that a real number is computable if it differs by an integer from a number in the unit interval whose binary expansion can be generated by some Turing machine.

Alternatively, if one wants the machine to compute a function, one can, after starting the machine with the input word or number on the tape, wait for it to halt and then look at the tape (starting with the square then under examination) to see what the output word or number is.

Turing shows how to construct a "universal computing machine" that, when supplied with the "standard description" of a machine *M* on its tape, will simulate the operation of *M*. This allows him to apply a diagonal argument to show that there can be no computable way to determine whether or not a given machine will continue to print 0's and 1's forever. In effect, he shows the unsolvability of the halting problem.

Turing argues that his formulation of the computability concept includes all sequences that would informally be considered to be computable. That is, he argues for what is now called Turing's thesis, the Church-Turing thesis, or Church's thesis, depending on the context. He gives three kinds of arguments: First, he shows how his machines can capture the informal idea of a step-by-step process, as indicated briefly earlier. Second, he shows that certain changes to his definition of a machine would have no effect at all on what sequences would be computable. And third, he gives examples of large classes of numbers that are computable: the real algebraic numbers, *e*, π, the real zeros of the Bessel functions, and so forth. Of course, as he emphasizes, only countably many real numbers can be computable.

Turing's 1939 paper, "Systems of Logic Based on Ordinals," is based on his PhD dissertation, written under Church's supervision during Turing's two-year stay at Princeton.

Gödel's incompleteness theorems had shown that any sufficiently strong formal system was incomplete, and in particular could not prove its own consistency. One can then add to this formal system the sentence expressing its consistency, thereby obtaining a stronger system. And this process can be iterated. The iteration can be transfinite, making use of ordinal notations for the constructive ordinals. This topic, which was later taken up by Solomon Feferman (1928–) in the 1950s, does not directly pertain to computability theory.

In the process, however, Turing introduced the important concept of computability relative to an oracle. He gave the basic definitions and indicated how his work on computability could be adapted to incorporate the idea of calculations that, at any stage, could utilize a hypothetical fixed body of information. This idea later led to work on the classification (of problems or of sets or of functions) according to degree of unsolvability. Moreover, the degrees of unsolvability are partially ordered, under what is now called Turing reducibility.

After 1939 Turing's work on computability stopped while Turing, now back in England, threw himself into wartime cryptographic work. There was an urgent need to break the German battlefield Enigma code. The success of Turing and the British cryptographic team was of enormous military importance throughout World War II. But nothing was known publicly about this work until it was declassified several years after Turing's suicide in 1954.

After the war Turing turned to computation topics, both practical and theoretical, outside the field of computability theory. On the practical side, he was involved in hardware and software design for early digital computers. On the theoretical side, he published important work on artificial intelligence.

See also Church, Alonzo; Computability Theory; Computationalism; Gödel, Kurt; Hilbert, David; Machine Intelligence; Turing, Alan M.

Bibliography

Church, Alonzo. "An Unsolvable Problem of Elementary Number Theory." *American Journal of Mathematics* 58 (1936): 345–363.

Turing, A. M. "Computability and λ-Definability." *The Journal of Symbolic Logic* 2 (1937): 153–163.

Turing, A. M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42 (1936–1937): 230–265. A correction, 43 (1937): 544–546.

Turing, A. M. "Systems of Logic Based on Ordinals." *Proceedings of the London Mathematical Society* 45 (3) (1939): 161–228.

Herbert B. Enderton (2005)

DECIDABLE AND UNDECIDABLE THEORIES. Suppose *T* is a theory (i.e., a set of sentences) in a formal language

L of logic. A decision procedure for *T* is a mechanical procedure for calculating whether any given sentence of *L* is a logical consequence of *T*. We say that *T* is decidable if it has a decision procedure and undecidable if not. The decision problem for *T* is to determine whether or not *T* is decidable. (One can avoid the slightly vague notion of a mechanical procedure by noting that a theory *T* is decidable if and only if the set of its logical consequences is computable.

Quantifier elimination and related model-theoretic techniques have yielded proofs that many important first-order theories are decidable. Examples are the theory of addition of integers (Presburger 1930), the theories of real-closed fields and algebraically closed fields (Tarski 1951), the theory of abelian groups (Szmielew 1955), and—if a number-theoretic conjecture of Schanuel is true—the theory of the field of real numbers with exponentiation (Macintyre and Wilkie 1996). The first theory shown to be undecidable was first-order Peano arithmetic; Kurt Gödel proved its undecidability in 1931. Many other undecidable theories are known, but the proofs of undecidability are all based directly or indirectly on Gödel's ideas. In 1970 Yuri V. Matiyasevich (1993) improved Gödel's result by showing that the set of diophantine sentences true in the natural numbers is not computable (a diophantine sentence is one of the form "There are natural numbers *m*, *n*, and so on such that *E* is true," where *E* is an arithmetical equation using *m*, *n*, and so on). Part 3 of "Decidable and Undecidable Theories" of J. Donald Monk (1976) gives many examples.

We say that a formal language *L* of logic is decidable if the empty theory in *L* is decidable—in other words, if there is a mechanical test to determine which sentences of *L* are valid. Gödel's ideas led to a proof that if *L* is a nontrivial first-order language, for example, with at least one binary relation symbol besides equality, then *L* is undecidable (Church 1936). Later research extended this result to various important sublanguages of first-order languages. But there are also decidable languages, for example, languages of propositional logic and a number of languages with monadic predicate symbols (e.g., the language of syllogisms). See Egon Börger, Erich Grädel, and Yuri Gurevich (1997) for full information on decidable and undecidable languages. After their book appeared, a new family of decidable languages was discovered, the guarded languages, whose decidability implies the decidability of various modal logics (see Grädel, Hirsch, and Otto 2002).

The decision problem for logical languages is also known as the Entscheidungsproblem. See Paulo Mancosu (1999, §8) on the place of this problem in early twentieth-century thinking about the foundations of mathematics, particularly within the school of David Hilbert.

See also Computability Theory; First-Order Logic; Gödel's Theorem; Model Theory.

Bibliography

Börger, Egon, Erich Grädel, and Yuri Gurevich. *The Classical Decision Problem*. Berlin: Springer, 1997.

Church, Alonzo. "A Note on the Entscheidungsproblem." *Journal of Symbolic Logic* 1 (1936): 40–41. Correction *ibid*: 101–102.

Gödel, Kurt. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I." *Monatshefte für Mathematik und Physik* 38 (1931): 173–198. Edited and translated in *Kurt Gödel. Collected Works Volume I: Publications 1929–1936*, ed. Solomon Feferman, et al, pp. 144–195. New York: Oxford University Press, 1986.

Grädel, Erich, Colin Hirsch, and Martin Otto. "Back and Forth between Guarded and Modal Logics." *ACM Transactions on Computational Logics* 3 (2002): 418–463.

Macintyre, Angus, and Alex Wilkie. "On the Decidability of the Real Exponential Field." *Kreiseliana*, edited by P. Odifreddi, pp. 441–467. Wellesley, MA: A.K. Peters, 1996.

Mancosu, Paolo. "Between Russell and Hilbert: Behmann on the Foundations of Mathematics." *Bulletin of Symbolic Logic* 5 (1999): 303–330.

Matiyasevich, Yuri V. *Hilbert's Tenth Problem*. Cambridge, MA: MIT Press, 1993.

Monk, J. Donald. *Mathematical Logic*. New York: Springer, 1976.

Presburger, Mojżesz. "Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt." *Comptes Rendus du Premier Congrès des Mathématiciens des Pays Slaves, Warszawa 1929*. Warsaw (1930): 92–101; supplementarty note *ibid*: 395.

Szmielew, Wanda. "Elementary Properties of Abelian Groups." *Fundamenta Mathematicae* 41 (1955): 203–271.

Tarski, Alfred. *A Decision Method for Elementary Algebra and Geometry*. Berkeley: University of California Press, 1951.

Tarski, Alfred, Andrzej Mostowski, and Raphael M. Robinson. *Undecidable Theories*. Amsterdam: North-Holland, 1953.

Wilfrid Hodges (2005)

MODEL THEORY

Tarski. The Polish-American logician Alfred Tarski (1901–1983) was born Alfred Teitelbaum in Warsaw; he changed his surname to Tarski in 1924. That same year he obtained his doctorate at the University of Warsaw for a thesis in logic under the supervision of Stanisław Leśniewski; he had also studied under Tadeusz Kotarbiński, Kazimierz Kuratowski, Jan Łukasiewicz, Stefan Mazurkiewicz and Wacław Sierpiński. At the University of Warsaw he was Docent and then Adjunct Professor from

1924 to 1939; simultaneously he taught in a high school from 1925 to 1939. From 1942 to his retirement in 1969 he held posts at the University of California at Berkeley.

Through his own work and that of his students, Tarski stands along with Aristotle and Frege as one of the creators of the discipline of logic. Andrzej Mostowski, Julia Robinson, Robert Vaught, Chen-Chung Chang, Solomon Feferman, Richard Montague, Jerome Keisler, and Haim Gaifman, among others, wrote their theses under his supervision.

In print Tarski was reluctant to place himself in any philosophical tradition. He described himself as "perhaps a philosopher of a sort." In 1930 he said that he agreed in principle with Leśniewski's "intuitionistic formalism," but in 1954 he reported that this was no longer his attitude. His philosophical reticence was certainly deliberate and reflected a view that careful formalization can resolve or at least avoid problems thrown up by philosophical speculation.

Tarski had many research interests within logic. He maintained most of them throughout his career and integrated them to an extraordinary degree. The setting of most of his work in Warsaw from 1926 to 1938 was the notion of a deductive theory. Such a theory develops a certain subject matter, starting from primitive terms together with axioms and proceeding by definition and logical deduction, all within a formally defined language. Tarski saw these theories as a paradigm for research in mathematical subjects. Like David Hilbert with his metamathematics, Tarski proposed to take the theories themselves as subject matter. But unlike Hilbert, Tarski did so by developing metatheories (that is, deductive theories about deductive theories) without any restriction to finitary means. For example the notion of "true sentence of the deductive theory T" must be defined in a metatheory T'. Tarski chose as primitive notions of T' those of T together with notions from set theory and syntax, and he showed how to write a definition in the metatheory which exactly characterizes the class of true sentences of T. He proposed similar metatheoretic definitions of "satisfies," "definable," and (with less confidence) "logical consequence." His later characterization of "logical notion" was published posthumously. His influential English exposition of his definition of truth in 1944 is still the best nontechnical introduction.

At the same time Tarski developed methodologies for creating deductive theories of particular topics, and for settling the decision problem for particular deductive theories. His method of elimination of quantifiers, based on work of Thoralf Skolem and others, guided him to an axiomatisation of the first-order theory of the field of real numbers. As a byproduct he found an algorithm for deciding the truth of first-order statements about the field of real numbers (or, as he later realized, any real-closed field). Responding to the work of Alonzo Church and Alan Turing on undecidability, Tarski developed methods for proving the undecidability of a deductive theory T by interpreting a known undecidable theory within T.

In the 1940s Tarski turned his attention to the application of metatheorems of logic in mathematics. In parallel with Anatolii Mal'tsev and Abraham Robinson, he showed that the compactness theorem of first-order logic could be used to prove purely mathematical facts. During the early 1950s he recast his notion of deductive theory to fit the new program. A deductive theory was no longer about a particular subject matter. Rather it was in a formal language with primitive symbols that could be interpreted as one pleases. An interpretation that makes all the axioms of the theory true is called a *model* of the theory. We can study those classes of structures which consist of all the models of a particular theory; in 1954 Tarski proposed the name *theory of models* for this line of research. Tarski adapted his definition of truth to define the relation "Sentence ϕ is true in structure A." He published this new model-theoretic truth definition in a joint paper with Vaught, which also included fundamental theorems about elementary embeddings between structures.

Particular theories that Tarski had studied in connection with quantifier elimination or undecidability became central to model-theoretic research. Some of them, such as the theories of real-closed fields and algebraically closed fields, remained central fifty years on. Tarski also stated several problems that strongly influenced the direction of model-theoretic research. For example he asked for a quantifier elimination for the field of reals with an exponentiation function, and for algebraic necessary and sufficient conditions for two structures to be elementarily equivalent.

Tarski's further contributions during his American period were perhaps more scattered but no less important. He was closely involved in the theory of large cardinals. He also worked with students and colleagues on relation algebras and cylindrical algebras. During the 1960s he studied finite axiomatisations of equational classes, picking up a theme from his work with Łukasiewicz during the 1920s on propositional logics. He never lost his interest in formal theories of geometry. Students of his recall that he looked back with particular pride to the work that he did during the 1940s with Bjarni

Jónsson on decompositions of finite algebras. With the help of colleagues in Europe and the United States, he was instrumental in the setting up of the series of International Congresses in Logic, Methodology and Philosophy of Science, which first met at Stanford in 1960.

See also Aristotle; Frege, Gottlob; Hilbert, David; Kotarbiński, Tadeusz; Leśniewski, Stanisław; Łukasiewicz, Jan; Model Theory; Montague, Richard; Set Theory.

Bibliography

Feferman, Anita Burdman, and Solomon Feferman. *Alfred Tarski: Life and Logic*. New York: Cambridge University Press, 2004.

Givant, Steven R. "A Portrait of Alfred Tarski." *The Mathematical Intelligencer* 13 (1991): 16–32.

Suppes, Patrick. "Philosophical Implications of Tarski's Work." *Journal of Symbolic Logic* 53 (1988): 80–91.

Tarski, Alfred. *O logice matematycznej i metodzie dedukcyjnej*, Biblioteczka Mat. 3–5. Lwów and Warsaw: Ksiąznica-Atlas, 1936. Translated and revised as *Introduction to Logic and to the Methodology of the Deductive Sciences*. New York: Oxford University Press, 1994.

Tarski, Alfred. "The Semantic Conception of Truth." *Philosophy and Phenomenological Research* 4 (1944): 13–47.

Tarski, Alfred. *Logic, Semantics, Metamathematics*. Translated by J. H. Woodger. 2nd ed., edited by John Corcoran. Indianapolis: Hackett, 1983.

Tarski, Alfred. *Pisma Logiczno-Filozoficzne*. Vol. 1, *Prawda*; Vol. 2, *Metalogika*, edited by Jan Zygmunt. Warsaw: Wydawnictwo Naukowe, 1995, 2001.

Tarski, Alfred. "What Are Logical Notions?" *History and Philosophy of Logic* 7 (1986): 143–154.

Tarski, Alfred. *Collected Papers*. 4 vols., edited by S. R. Givant and R. N. McKenzie. Basel: Birkhäuser, 1986.

Tarski, Alfred, and Robert L. Vaught. "Arithmetical Extensions of Relational Systems." *Compositio Mathematica* 13 (1957): 81–102.

Woleński, Jan. *Logic and Philosophy in the Lvov-Warsaw School*. Dordrecht: Kluwer, 1989.

Wilfrid Hodges (2005)

Robinson. Abraham Robinson (1918–1974) was a logician and mathematician. Born in Waldenburg (Silesia), he moved to Palestine in 1933, where he studied mathematics at the Hebrew University in Jerusalem and also joined the Haganah. In 1940 he fled to Britain as a wartime refugee and enlisted with the Free French Air Force. He took his PhD in London in 1949 while teaching aerodynamics at the Cranfield College of Aeronautics. He held posts successively in Toronto, Jerusalem, Los Angeles, and finally Yale, where he died of cancer. His eventful life is described by Joseph W. Dauben (1995).

Robinson's PhD thesis on applications of logic in mathematics led to an invitation to speak at the International Congress of Mathematicians in 1950. The talks of Robinson and Alfred Tarski at this congress became founding documents of the new discipline that Tarski named model theory. Throughout his career Robinson was one of the most fertile contributors of programs, techniques, and results to model theory.

Robinson's thesis contains his independent discovery of the compactness theorem for first-order languages of any cardinality. In the proof he introduced constant symbols to stand for the elements of the model to be constructed. He noticed that if these constant symbols corresponded to the elements of a given structure *A*, and the theory contained sentences expressing all the relations of the structure *A*, then any model of the theory would contain an isomorphic copy of *A*. This observation became the method of diagrams, which Robinson used systematically as a way of creating models of a theory with prescribed embeddings between them. Diagrams immediately became one of the fundamental techniques of model theory (for many applications, see Robinson 1963).

Robinson switched from one branch of mathematics to another with extraordinary ease. There were certain topics that he kept returning to from different angles. Two in particular were elementary embeddings and algebraically closed fields. Combining the two, he noted that every embedding between algebraically closed fields is elementary. He coined the term *model-complete* for theories whose models have this property and devised tests to show when a theory is model-complete.

Observing the role of algebraically closed fields in field theory, he looked for analogous structures within other classes. Model completions, model companions, infinite forcing companions, and finite forcing companions were notions that he proposed at various times as generalizations of algebraic closure. He identified the classes of real-closed fields and differentially closed fields as the model completions of the ordered fields and the differential fields, respectively, and axiomatized the class of differentially closed fields (though the usual axioms are an improved version due to Lenore Blum). In 1965 the notion of model completion played a central role in the proofs by James Ax and Simon Kochen, and independently by Yuri Ershov, of a number-theoretic conjecture of Emil Artin.

Around 1960 he noticed that any proper elementary extension of the field of real numbers contains infinitesimals. He quickly developed this insight into a powerful

and intuitively natural approach to mathematical analysis that he named nonstandard analysis. Nonstandard analysis is one of the few innovations in logic that were entirely the work of a single individual.

Not long before his death, Robinson collaborated with the number theorist Peter Roquette to apply model-theoretic methods in number theory. This work gave a first hint of the deep interactions between model theory and diophantine geometry that came to light in the 1990s, sadly too late for Robinson to contribute. In fact, Robinson died before he could take on board the stability theory pioneered by Michael Morley and Saharon Shelah, though his students, Greg Cherlin and Carol Wood, did contribute to this field, bringing with them Robinson's lifelong eagerness to apply model theory to algebra, algebraic geometry, and mathematics in general.

Though unable himself to believe in any kind of existence for infinite totalities, he strongly defended the right of mathematicians to proceed as if such totalities exist. His discussion (Robinson 1965) of mathematical and epistemological considerations that favor one or another of the traditional views in philosophy of mathematics is thoughtful but seems not to reveal a thoroughly worked out position. His anti-Platonistic attitude may have helped him to create nonstandard analysis by allowing him to be relaxed about what the "real" real numbers are.

In Robinson's *Selected Papers* (1979), the bibliography lists ten books, more than a hundred papers, and a film. One in seven of his papers are in wing theory and aeronautics.

See also Infinitesimals; Model Theory; Tarski, Alfred.

Bibliography

Dauben, Joseph W. *Abraham Robinson: The Creation of Nonstandard Analysis—A Personal and Mathematical Odyssey*. Princeton, NJ: Princeton University Press, 1995.

Robinson, Abraham. "Formalism 64." In *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science, Jerusalem 1964*, edited by Y. Bar-Hillel, 228–246. Amsterdam, Netherlands: North-Holland, 1965.

Robinson, Abraham. *Introduction to Model Theory and to the Metamathematics of Algebra*. Amsterdam, Netherlands: North-Holland, 1963.

Robinson, Abraham. *Selected Papers*. 3 vols, edited by H. J. Keisler et al. Amsterdam, Netherlands: North-Holland, 1979.

Wilfrid Hodges (2005)

SET THEORY SINCE GÖDEL See *Set Theory*

THE PROLIFERATION OF NONCLASSICAL LOGICS. The twentieth century, and especially its second half, was marked by a fairly spectacular proliferation of what are sometimes called nonclassical logics. To understand this, one needs to see the matter in its historical context. There have been three great periods in the history of European logic: ancient Greece, medieval Europe, and, starting toward the end of the nineteenth century, the current period. Each period has been marked by the production of novel theories of the nature and extent of logical validity. Thus, in the ancient period, Aristotle, the Megarian, and the Stoic logicians offered different accounts of validity, the conditional, and modality. The medieval period tried to reconcile some of the differences in their heritage, and in the process produced numerous different accounts of the nature of the connectives, consequence, and supposition. Not surprisingly, in both periods there was active and lively debate concerning the theories that were produced.

The periods between the great periods were characterized not just by a lack of interest in logic, but by a forgetting of much of the significant prior developments. In particular, all that remained of logic in about the middle of the nineteenth century—so-called traditional logic—was a somewhat bowdlerized form of the theory of the syllogism and some of its medieval accompaniments. It was at this time that mathematical logic came into existence. It was mathematical in two senses. The first is that the logicians who produced it were interested in the analysis of the reasoning of the mathematics of their time (and its foundations). The second is that they applied mathematical techniques to the subject in a novel way, such as those of abstract algebraic, set theory, and combinatorics.

Out of this, principally at the hands of Gottlob Frege and Bertrand Russell, developed a novel theory of logic. This was streamlined, organized, and simplified by a number of logicians in the first part of the twentieth century—notably, David Hilbert, Alfred Tarski, and Gerhard Gentzen. The result was an account of inference that was so much more powerful than traditional logic that is soon superseded it as the standard canon. This is so-called classical logic.

It had hardly appeared, however, before some logicians realized that a number of assumptions that were packed into it were contentious—especially once one goes beyond the kind of mathematical reasoning out of which classical logic arose. One of these was the principle of bivalence: that every (declarative) sentence is either true or false. In the 1920s the first many-valued logics

were produced by Jan Łukasiewicz, Emil Post, Tarski, and others. In many-valued logics, sentences can be assumed to be neither true nor false, both true and false, have an infinity of degrees of truth, and so on.

Another assumption that is packed into classical logic is truth-functionality: that the truth value of a compound sentence is a function of the truth values of its parts. This is obviously not true of modal notions, and in the 1920s Clarence Irving Lewis presented in axiomatic form the first modern systems of modal logic. Modal logic was given an enormous boost with the discovery of world-semantics by, in particular, Saul Kripke in the 1960s. This allowed for the production of logics for other non-truth-functional notions (so called intentional logics), such as tense-operators (by Arthur Prior), epistemic and doxastic notions (by Jaako Hintikka), and deontic notions (by Henrik von Wright).

Another early critique of classical logic was provided by mathematical intuitionists, such as Luitzen Brouwer and Arend Heyting, who, driven by the view that existence should not be asserted unless people can construct the object in question, produced a system of formal logic in which a number of propositional and quantifier inferences that are valid in classical logic fail.

In the second half of the century, various critics of classical logic attacked the account of the (material) conditional it employs (as had Lewis). This produced the relevant logics of Alan Anderson and Nuel Belnap, and the conditional logics of Robert Stalnaker and David Lewis. These logics both have world-semantics. The world semantics for relevant logics were produced by, in particular, Richard Routley (later Sylvan) and Robert Meyer. The central feature of such semantics (it can be seen in retrospect) is the deployment of the notion of an impossible world.

The principle of inference of classical logic that everything entails a contradiction came under attack in its own right by logicians in the same period, including Stanisław, Jaśkowski, Newton da Costa, and Graham Priest. This produced a number of paraconsistent logics, which may be many-valued, modal, relevant, or of other kinds.

The development of nonclassical logics received further momentum from the advent of computer science and information technology after the 1960s. This produced new constructivist systems (such as linear logic), intentional logics (such as dynamic logic), and paraconsistent logics (such as various resolution systems). Research in Artificial Intelligence has also produced new epistemic logics, as well as the whole new area of formal non-monotonic (i.e., non-deductive) inference.

Thus, at the start of the twenty-first century there is a wide range of logics embodying different metaphysical presuppositions and potential applications. What to make of this is another matter. Perhaps most obvious is that the revolution in logic that occurred around the turn of the twentieth century was not so much the production of a novel logical theory—important though this was. It was instead the deployment of mathematical techniques to logic in a novel way. This allowed the development of classical logic, but the techniques were so powerful and versatile that they could be used to produce many other logics as well.

Which of all these logics is right, and, indeed, the meaning of that question, are matters to be determined only by detailed philosophical argument. Such arguments have been much part of the philosophical landscape since about the middle of the twentieth century. Indeed, the twenty-first century is seeing disputes in philosophical logic of a depth and acuity not seen since medieval logic. Whatever their outcome, the presence of the multitude of logical systems serves to remind that logic is not a set of received truths, but a discipline in which competing theories concerning validity vie with each other. The case for each theory—including a received theory—has to be investigated on its merits.

See also Aristotle; Brouwer, Luitzen Egbertus Jan; Conditionals; Frege, Gottlob; Hintikka, Jaako; Hilbert, David; Intuitionism and Intuitionistic Logic; Kripke, Saul; Lewis, Clarence Irving; Lewis, David; Logic, Non-Classical; Łukasiewicz, Jan; Many-Valued Logics; Megarians; Modal Logic; Non-Monotonic Logic; Paraconsistent Logics; Prior, Arthur Norman; Relevance (Relevant) Logics; Russell, Bertrand Arthur William; Stoicism; Tarski, Alfred; Wright, Georg Henrik von.

Bibliography

Gabbay, Dov M., and John Woods, eds. *Handbook of the History of Logic*. Vol. 6: *Logic and Modalities in the Twentieth Century*. Amsterdam, the Netherlands: Elsevier, 2005.

Gabbay, Dov M., and John Woods, eds. *Handbook of the History of Logic*. Vol. 7: *The Many-Valued and Non-Monotonic Turn in Logic*. Amsterdam, the Netherlands: Elsevier, 2005.

Haack, Susan. *Deviant Logic, Fuzzy Logic: Beyond the Formalism*. Chicago, IL: University of Chicago Press, 1996.

Priest, Graham. *Introduction to Non-Classical Logic*. Cambridge, U.K.: Cambridge University Press, 2001.

Read, Stephen. *Thinking about Logic: An Introduction to the Philosophy of Logic*. Oxford, U.K.: Oxford University Press, 1994.

Graham Priest (2005)

KRIPKE AND KRIPKE MODELS. See *Kripke, Saul*

FRIEDMAN AND REVERSE MATHEMATICS. During the second half of the twentieth century, many mathematicians lost interest in the foundations of mathematics. One of the reasons for this decline was an increasingly popular view that general set theory and Gödel-style incompleteness and independence results do not have much effect on mathematics as it is actually practiced. That is, as long as mathematicians study relatively concrete mathematical objects, they can avoid all foundational issues by appealing to a vague hybrid of philosophical positions including Platonism, formalism, and sometimes even social constructivism. Harvey Friedman (born 1948) has continually fought this trend, and in 1984 he received the National Science Foundation's Alan T. Waterman Award for his work on revitalizing the foundations of mathematics.

One of Friedman's methods of illustrating the importance of foundational issues is to isolate pieces of mathematics that either display the incompleteness phenomenon or require substantial set theoretic assumptions and which most mathematicians would agree fall within the scope of the central areas of mathematics. For example, he has created numerous algebraic and geometric systems that make no explicit reference to logic but which, under a suitable coding, contain a logical system to which Gödel's incompleteness theorems apply. Furthermore, these systems look similar to many systems used by mathematicians in their everyday work. Friedman uses these examples to argue that incompleteness cannot be dismissed as a phenomenon that occurs only in overly general foundational frameworks contrived by logicians and philosophers.

Friedman has also done a large amount of work concerning the necessary use of seemingly esoteric parts of Zermelo-Frankel set theory and its extensions. He has found theorems concerning concrete objects in mathematics that require the use of uncountably many iterations of the power set axiom and others that require the use of large cardinal axioms. These investigations have culminated in what Friedman calls Boolean relation theory.

In his 1974 address to the International Congress of Mathematicians, Friedman started the field of reverse mathematics by suggesting a three-step method for measuring the complexity of the set theoretic axioms required to prove any given theorem T. First, formalize the theorem T in some version of set theory. (Typically a formal system called second order arithmetic is used.) Second, find a collection of set theoretic axioms S which suffices to prove T. Third, prove the axioms in S from the theorem T (while working in a suitably weak base theory). If the third step is successful, then the equivalence between S and T shows that S is the weakest collection of axioms which suffices to prove T. If the third step fails, then the second step must be repeated until a proof of T is found using only axioms that can be proved from T. Because the third step involves proving axioms from theorems as opposed to the usual action of proving theorems from axioms, this type of analysis is now called reverse mathematics. It is frequently possible to draw a number of foundational conclusions concerning a theorem T once the equivalent collection S of set theoretic axioms has been isolated.

See also Gödel's Incompleteness Theorems; Mathematics, Foundations of; Platonism and the Platonic Tradition; Reverse Mathematics.

Bibliography

Friedman, H. "Some Systems of Second Order Arithmetic and Their Use." *Proceedings of the 1974 International Congress of Mathematicians*, vol. 1, 235–242. Montreal: Canadian Math Congress, 1975.

Harrington, L. A., M. D. Morley, A. Scedrov, and S. G. Simpson, eds. *Harvey Friedman's Research on the Foundations of Mathematics*. Studies in Logic and the Foundations of Mathematics, vol. 117. Amsterdam: North Holland Press, 1985.

Peter Cholak (2005)
Reed Solomon (2005)

LOGIC, MATHEMATICAL

See *Logic, History of*

LOGIC, MODAL

See *Modal Logic*

LOGIC, NON-CLASSICAL

The purpose of this entry is to survey those modern logics that are often called "non-classical," classical logic being the theory of validity concerning truth functions and first-order quantifiers likely to be found in introductory textbooks of formal logic at the end of the twentieth century.

For the sake of uniformity I will give a model-theoretic account of the logics. All of the logics also have proof-theoretic characterizations, and in some cases (such as linear logic) these characterizations are somewhat more natural. I will not discuss combinatory logic, which is not so much a non-classical logic as it is a way of expressing inferences that may be deployed for both classical and non-classical logics. I will use A, B, … for arbitrary sentences; $\wedge$, $\vee$, $\neg$, and $\rightarrow$, for the standard conjunction, disjunction, negation, and conditional operators for whichever logic is at issue. "Iff" means "if and only if." For references see the last section of this article.

EXTENSIONS VERSUS RIVALS

An important distinction is that between those non-classical logics that take classical logic to be alright as far as it goes, but to need extension by the addition of new connectives, and those which take classical logic to be incorrect, even for the connectives it employs. Call the former *extensions* of classical logic, and the latter *rivals*. Thus modal logics, as now usually conceived, are extensions of classical logic. They agree with classical logic on the extensional connectives (and quantifiers if these are present) but augment them with modal operators. By contrast, intuitionist and relevant logics are more plausibly thought of as rivals. Thus $A \vee \neg A$ is valid in classical logic but not intuitionist logic, and $A \rightarrow (B \rightarrow A)$ is valid in classical logic but not relevant logic.

The distinction must be handled with care however. Modern modal logics can be formulated, not with the modal operators, but with the strict conditional, ⥽ (from which modal operators can be defined), as primitive; and *A*⥽(*B*⥽*A*) is not valid. From this perspective modal logic is a rival to classical logic (which is the way it was originally intended). Similarly it is (arguably) possible to add a negation operator, $\$$, to relevant logics which behaves as does classical negation. Classical logic is, then, just a part of this logic, identifying the classical $\neg A$ and $A \rightarrow B$ with the relevant $\$A$ and $\$A \vee B$, respectively. From this perspective, in a relevant logic, $\rightarrow$ and $\neg$ are operators additional to the classical ones, and relevant logic is an extension of classical logic.

What these examples show is that whether or not something is an extension or a rival of classical logic is not a purely formal matter but a matter of how the logic is taken to be applied to informal reasoning. If, in a modal logic, one reads *A*⥽ *B* as "if *A* then *B*" then the logic is a rival of classical logic. If one reads $A \rightarrow B$ as "if *A* then *B*" and *A*⥽*B* as "necessarily, if *A* then *B*," it is an extension. If, in a relevant logic, one reads $A \rightarrow B$ as "if *A* then *B*," and $\neg A$ as "it is not the case that *A*," the logic is a rival to classical logic; if one reads $\$A \vee B$ as "if *A* then *B*" and $\$A$ as "it is not the case *A*," it is an extension. (The examples also raise substantial philosophical issues. Thus both a relevant logician and an intuitionist are liable to deny that $\$$ is a connective with any determinate meaning.)

MANY-VALUED LOGICS

A central feature of classical logic is its bivalence. Every sentence is exclusively either true (1) or false (0). In *many-valued logics*, normally thought of as rivals to classical logic, there are more than two semantic values. Truth-functionality is, however, maintained; thus the value of a compound formula is determined by the values of its components. Some of the semantic values are *designated*, and a valid inference is one in which, whenever the premises are *designated*, so is the conclusion.

A simple example of a many-valued logic is that in which there are three truth values, 1, *i*, 0; and the truth functions for the standard connectives may be depicted as follows:

	$\neg$
*1	0
i	*i*
0	1

$\rightarrow$	1	*i*	0
1	1	*i*	0
i	1	1	*i*
0	1	1	1

$\vee$	1	*i*	0
1	1	1	1
i	1	*i*	*i*
0	1	*i*	1

$\wedge$	1	*i*	0
1	1	*i*	0
i	*i*	*i*	0
0	0	0	0

The only designated value is 1 (which is what the asterisk indicates). This is the Łukaziewicz 3-valued logic, Ł$_3$. If the middle value of the table for $\rightarrow$ is changed from

1 to *i* we get the Kleene 3-valued logic K_3. The standard interpretation for *i* in this logic is *neither true nor false*. If in addition *i* is added as a designated value, we get the paraconsistent logic *LP*. The standard interpretation for *i* in this is *both true and false*.

$Ł_3$ can be generalized to a logic, $Ł_n$, with *n* values, for any finite *n*, and even to one with infinitely many values. Thus the continuum-valued Łukasiewicz logic, $Ł_{\aleph}$, has as semantic values all real numbers between 0 and 1 (inclusive). Normally only 1 is designated. If we write the value of *A* as $\nu(A)$, $\nu(A \lor B)$ and $\nu(A \land B)$ are the maximum and minimum of $\nu(A)$ and $\nu(B)$, respectively; $\nu(\neg A)=1-\nu(A)$; $\nu(A \rightarrow B)=1$ if $\nu(A) \leq \nu(B)$ and $\nu(A \rightarrow B)=1-(\nu(A)-\nu(B))$ otherwise. Standardly the semantic values are thought of as degrees of truth (so that 1 is *completely true*). Interpreted in this way $Ł_{\aleph}$ is one of a family of many-valued logics called *fuzzy logics*.

MODAL LOGICS

Another family of non-classical logics maintains bivalence, but rejects truth-functionality. *Modal logics* augment the connectives of classical logic with the operators $\Box$ (it is necessarily the case) and $\Diamond$ (it is possibly the case). The truth-values of $\Box A$ and $\Diamond A$ depend on more than just the truth value of *A*.

Standard semantics for modal logics invoke a set of (possible) worlds, augmented with a binary relation, *R*. *wRw′* means, intuitively, that from the state of affairs as it is at *w*, the state of affairs as it is at *w′* is possible. (In first-order modal logics each world comes also with a domain of quantification.) The extensional connectives are given their usual truth conditions with respect to a world, but if we write the value of *A* at world *w* as $\nu_w(A)$:

$\nu_w(\Box A)=1$, iff for all *w′* such that *wRw′*, $\nu_{w'}(A)=1$

$\nu_w(\Diamond A)=1$, iff for some *w′* such that *wRw′*, $\nu_{w'}(A)=1$

Validity is defined in terms of truth preservation at all worlds. (This is for *normal* modal logics. *Non-normal* modal logics have also a class of non-normal worlds, at which the truth conditions of the modal operators are different.)

Different modal logics are obtained by putting constraints on *R*. If *R* is arbitrary we have the system *K*. If it is reflexive (validating $\Box A \rightarrow A$), we have *T*; if transitivity is also required (validating $\Box A \rightarrow \Box\Box A$), we have *S*4; if symmetry is added (validating $A \rightarrow \Box\Diamond A$), we have *S*5. (Alternatively, in this case, *R* may be universal: For all *w and w′*, *wRw′*.) If we have just the condition that every world is related to some world or other (validating $\Box A \rightarrow \Diamond A$), we have *D*.

The notion of possibility is highly ambiguous (logical, physical, epistemic, etc.). Arguably, different constraints on *R* are appropriate for different notions.

INTENSIONAL LOGICS

World semantics have turned out to be one of the most versatile techniques in contemporary logic. Generally speaking, logics that have world-semantics are called *intensional logics* (and are normally thought of as extensions of classical logic). There are many of these in addition to standard modal logics.

$\Box$ may be interpreted as "it is known that", in which context it is usually written as *K* and the logic is called *epistemic logic*. (The most plausible epistemic logic is *T*.) It may be interpreted as "it is believed that," in which case it is usually written as *B*, and the logic is called *doxastic logic*. (Though even the logic *K* seems rather too strong here, except as an idealization to logically omniscient beings.) $\Box$ may be interpreted as "it is obligatory to bring it about that," in which case it is written as *O*, and the logic is called *deontic logic*. The standard deontic logic is *D*.

One can also interpret $\Box$ as "it is provable that." The best-known system in this regard is usually known as *GL* and called *provability logic*. This logic imposes just two constraints on the accessibility relation. One is transitivity; the other is that there are no infinite *R*-chains, that is, no sequences of the form w_0Rw_1, w_1Rw_2, w_2Rw_3, … This constraint verifies the principle $\Box(\Box A \rightarrow A) \rightarrow \Box A$, but not $\Box A \rightarrow A$. The interest of this system lies in its close connection with the way that a provability predicate, *Prov*, works in standard systems of formal arithmetic. By Gödel's second incompleteness theorem, in such logics one cannot prove $Prov(\langle A \rangle) \rightarrow A$ (where $\langle A \rangle$ is the numeral for the gödel number of *A*); but Löb's theorem assures us that if we can prove $Prov(\langle A \rangle) \rightarrow A$ we can prove *A*, and so $Prov(\langle A \rangle)$. It is this idea that is captured in the characteristic principle of *GL*.

Another possibility is to interpret $\Box$ and $\Diamond$ as, respectively, 'it will always be the case that,' and 'it will be the case at some time that.' In this context the operators are normally written as *G* and *F*, and the logic is called *tense logic*. In the world-semantics for tense logics, worlds are thought of as times, and the accessibility relation, *R*, is interpreted as a temporal ordering. In these logics there are also past-tense operators: *H* and *P* ("it has always been the case that" and "it was the case at some time that," respectively). These are given the reverse truth conditions. Thus for example:

$\nu_w(HA)=1$, iff for all w' such that $w'Rw$, $\nu_{w'}(A)=1$

The past and future tense operators interact in characteristic ways (e.g., $A \rightarrow HFA$ is logically valid). The basic tense logic, K_t, is that obtained when R is arbitrary. As with modal logics, stronger systems are obtained by adding constraints on R, which can now represent the ideas that time is dense, has no last moment, and so on.

Of course it is not necessary to have just one family of intensional operators in a formal language: One can have, for example, modal and tense operators together. Each family will have its own accessibility relation, and these may interact in appropriate ways. Systems of logic with more than one family of modal operators are called *multi-modal*. One of the most important multi-modal logics is *dynamic logic*. In this there are operators of the form $[\alpha]$ and $\langle\alpha\rangle$, each with its own accessibility relation, R_α. In the semantics of dynamic logic, the worlds are thought of as states of affairs or of a computational device. The αs are thought of as (non-deterministic) actions or programs, and $wR_\alpha w'$ is interpreted to mean that starting in state w and performing the action α (or running the program α) can take one to the state w'. Thus $[\alpha]\, A$ ($\langle\alpha\rangle A$) holds at state w, just if performing α at w will always (may sometimes) lead to a state in which A holds. The actions themselves are closed under certain operations. In particular, if α and β are actions, so are $\alpha;\beta$ (perform α and then perform β); $\alpha \cup \beta$ (perform α or perform β, non-deterministically); α^* (perform α some finite number of times, non-deterministically). There is also an operator, ? ("test whether"), which takes sentences into programs. The corresponding accessibility relations are: $xR_{\alpha;\beta}y$ iff for some z, $xR_\alpha z$ and $zR_\beta y$; $xR_{\alpha\cup\beta}y$ iff $xR_\alpha y$ or $xR_\beta y$; $xR_{\alpha*}y$ iff for some $x=x_1, x_2, \ldots, x_n=y$, $x_0R_\alpha x_1$, $x_1R_\alpha x_2$, $\ldots$, $x_{n-1}R_\alpha x_n$; $xR_{A?}y$ iff ($x=y$ and $\nu_x(A)=1$). Because of the * operator, dynamic logic can express the notion of finitude in a certain sense. This gives it some of the expressive strength of second-order logic.

CONDITIONAL LOGICS

Another family of logics of the intentional variety was triggered by some apparent counter-examples to the following inferences:

$A \rightarrow B \vdash (A \wedge C) \rightarrow B$

$A \rightarrow B, B \rightarrow C \vdash A \rightarrow C$

$A \rightarrow B \vdash \neg B \rightarrow \neg A$

which are valid for the material conditional. (For example: "If you strike this match it will light; hence if you strike this match and it is under water it will light.") Logics of the conditional that invalidate such principles are called *conditional logics*. Such logics add an intentional conditional operator, >, to the language. In the semantics there is an accessibility relation, R_A, for every sentence, A (or one, R_X, for every proposition, that is, set of worlds, X). Intuitively $wR_A w'$ iff w' is a world which A holds but is, *ceteris paribus*, the same as w. The truth conditions for > are:

$\nu_w(A>B)=1$ iff for all w' such that $wR_A w'$, $\nu_{w'}(B)=1$

The intuitive meaning of R motivates the following constraints:

$wR_A w'$ then $\nu_{w'}(A)=1$

if $\nu_w(A)=1$,then $wR_A w$

Stronger logics in the family are obtained by adding further constraints to the accessibility relations. A standard way of specifying these is in terms of "similarity spheres"—neighbourhoods of a world containing those worlds that have a certain degree of similarity to it.

The natural way of taking a conditional logic is as a rival to classical logic (giving a different account of the conditional). Some philosophers, however, distinguish between indicative conditionals and subjunctive/counterfactual conditionals. They take the indicative conditional to be the material conditional of classical logic, and > to be the subjunctive conditional. Looked at this way conditional logics can be thought of as extensions of classical logic.

INTUITIONIST LOGIC

There are a number of other important non-classical logics that, though not presented originally as intentional logics, can be given world semantics. One of these is *intuitionist logic*. This logic arose out of a critique of Platonism in the philosophy of mathematics. The idea is that one cannot define truth in mathematics in terms of correspondence with some objective realm, as in a traditional approach. Rather one has to define it in terms of what can be proved, where a proof is something that one can effectively recognize as such. Thus, semantically, one has to replace standard truth-conditions with proof-conditions, of the following kind:

$A \vee B$ is provable when A is provable or B is provable.

$\neg A$ is provable when it is provable that there is no proof of A

$\exists x A(x)$ is provable when we can effectively find an object, n, such that $A(n)$ is provable

Note that in the case of negation we cannot say that $\neg A$ is provable when A is not provable: We have no effective way of recognizing what is not provable; similarly, in the case of the existential quantifier, we cannot say that $\exists x A(x)$ is provable when there is some n such that $A(n)$ is provable: we may have no effective way of knowing whether this obtains.

Proceeding in this way produces a logic that invalidates a number of the principles of inference that are valid in classical logic. Notable examples are: $A \vee \neg A$, $\neg\neg A \rightarrow A$, $\neg \forall x A(x) \rightarrow \exists x \neg A(x)$. For the first of these, there is no reason to suppose that for any A we can find a proof of A or a proof that there is no proof of A. For the last, the fact that we can show that there is no proof of $\forall x A(x)$ does not mean that we can effectively find an n such that $A(n)$ can be proved.

In the world-semantics for intuitionist logic, interpretations have essentially the structure of an $S4$ interpretation. The worlds are interpreted as states of information (things proved), and the accessibility relation represents the acquisition of new proofs. We also require that if $\nu_w(A)=1$ and wRw', $\nu_{w'}(A)=1$ (no information is lost), and if x is in the domain of quantification of w and wRw' then x is in the domain of quantification of w' (no objects are undiscovered). Corresponding to the provability conditions we have:

$\nu_w(A \vee B)=1$ iff $\nu_w(A)=1$ or $\nu_w(A)=1$

$\nu_w(\neg A)=1$ iff for all w' such that wRw', $\nu_{w'}(A)=0$

$\nu_w(\exists x A(x))=1$ iff for all n in the domain of w, $\nu_w(A(n))=1$

Unsurprisingly, given the above semantics, there is a translation of the language of intuitionism into quantified $S4$ that preserves validity.

Another sort of semantics for intuitionism takes semantic values to be the open sets of some topology. If the value of A is x, the value of $\neg A$ is the interior of the complement of x.

RELEVANT LOGIC

Another logic standardly thought of as a rival to classical logic is *relevant* (or *relevance*) *logic*. This is motivated by the apparent incorrectness of classical validities such as: $A \rightarrow (B \rightarrow B)$, $(A \wedge \neg A) \rightarrow B$. A (propositional) relevant logic in one in which if $A \rightarrow B$ is a logical truth A and B share a propositional parameter. There are a number of different kinds of relevant logic, but the most common has a world-semantics. The semantics differs in two major ways from the world semantics we have so far met.

First it adds to the possible worlds a class of logically impossible worlds. (Though validity is still defined in terms of truth-preservation over possible worlds.) In possible worlds the truth conditions of $\rightarrow$ are as for $\dashv$ in $S5$:

$\nu_w(A \rightarrow B)=1$ iff for all w' (possible and impossible) such that $\nu_{w'}(A)=1$, $\nu_{w'}(B)=1$

In impossible worlds the truth conditions are given differently, in such a way that logical laws such as $B \rightarrow B$ may fail at the world. This may be done in various ways, but the most versatile technique employs a three-place relation, S, on worlds. If w is impossible, we then have:

$\nu_w(A \rightarrow B)=1$ iff for all x,y such that $Swxy$, if $\nu_x(A)=1$, $\nu_y(B)=1$

This clause can be taken to state the truth conditions of $\rightarrow$ at all worlds, provided that we add the constraint that, for possible w, $Swxy$ iff x=y. With no other constraints on S, this gives the basic (positive) relevant logic, B. Additional constraints on S give stronger logics in the family. Typical constraints are:

$\exists x(Sabx$ and $Sxcd) \Rightarrow \exists y(Sacy$ and $Sbyd)$

$Sabc \Rightarrow Sbac$

$Sabc \Rightarrow \exists x(Sabx$ and $Sxbc)$

Adding all three gives the (positive) relevant logic, R. Adding the first two gives RW, R minus Contraction ($A \rightarrow (A \rightarrow B) \vdash A \rightarrow B$). The intuitive meaning of S is, at the time of this writing, philosophically moot.

The second novelty of the semantics is in its treatment of negation. It is necessary to arrange for worlds where $A \wedge \neg A$ may hold. This may be done in a couple of ways. The first is to employ the Routley * operator. Each world, w, comes with a "mate," w^* (subject to the constraint that w^{**}=w, to give Double Negation). We then have:

$\nu_w(\neg A)=1$ iff $\nu_{w^*}(A)=0$

(If $w=w^*$, this just delivers the classical truth conditions.) Alternatively, we may move to a four-valued logic in which the values at a world are *true only*, *false only*, *both*, *neither* ({1}, {0}, {1,0}, ∅). We then have:

$1 \in \nu_w(\neg A)$ iff $0 \in \nu_w(A)$

$0 \in \nu_w(\neg A)$ iff $1 \in \nu_w(A)$

The semantics of relevant logic can be extended to produce a (relevant) *ceteris paribus* conditional, >, of the kind found in conditional logics, by adding the appropriate binary accessibility relations.

DISTRIBUTION-FREE LOGICS

There are some logics in the family of relevant logics for which the principle of Distribution, $A\wedge(B\vee C)\vdash(A\wedge B)\vee(A\wedge C)$, fails. To achieve this the truth conditions for disjunction have to be changed. In an interpretation, let $[A]$ be the set of worlds at which A holds. Then the usual truth conditions for disjunction can be written:

$$\nu_w(A\vee B)=1 \text{ iff } w\in[A]\cup[B]$$

To invalidate Distribution, the semantics are augmented by a closure operator, $\mathfrak{C}$, on sets of worlds, x, satisfying the following conditions:

$$X\subseteq\mathfrak{C}(X)$$

$$\mathfrak{C}\mathfrak{C}(X)=\mathfrak{C}X$$

$$\text{if } X\subseteq Y \text{ then } \mathfrak{C}(X)\subseteq\mathfrak{C}(Y)$$

The truth conditions of disjunction can now be given as:

$$\nu_w(A\vee B)=1 \text{ iff } w\in\mathfrak{C}([A]\cup[B])$$

Changing the truth conditions for disjunction in *RW* in this way (and using the Routley * for negation) gives linear logic (*LL*). *LL* is usually formulated with some extra intentional connectives, especially an intentional conjunction and disjunction. These connectives can be present in standard relevant logics too. Intuitionist, relevant, and linear logics all belong to the family of *substructural logics*. Proof-theoretically, these logics can be obtained from a sequent-calculus for classical logic by weakening the structural rules (especially Weakening and Contraction).

Another logic in which distribution fails is *quantum logic*. The thought here is that it may be true (verifiable) of a particle that it has a position and one of a range of momenta, but each disjunct attributing to it that position and a particular momentum is false (unverifiable). The states of a quantum system are canonically thought of as members of a Hilbert space. In the world-semantics for quantum logic, the space of worlds is taken to be such a space, and sentences are assigned closed subsets of this. $[A\wedge B]=[A]\cap[B]$, $[A\vee B]=\mathfrak{C}([A]\cup[B])$, where $\mathfrak{C}(X)$ is the smallest closed space containing X; and $[\neg A]=[A]^{\perp}$. $X^{\perp}$ is the space comprising all those states that are orthogonal to members of X. (It satisfies the conditions: $X=X^{\perp\perp}$, if $X\subseteq Y$ then $Y^{\perp}\subseteq X^{\perp}$, and $X\cap X^{\perp}=\emptyset$.) In quantum logic $A\rightarrow B$ can be defined in various ways. Perhaps the most plausible is as $\neg A\vee(A\wedge B)$. (The subspaces of a Hilbert space also have the structure of a partial Boolean algebra. Such an algebra is determined by a family of Boolean algebras collapsed under a certain equivalence relation, which is a congruence relation on the Boolean operators. Partial Boolean algebras can be used to provide a slightly different quantum logic.)

PARACONSISTENT LOGICS

Before we turn to quantifiers there is one further kind of logic to be mentioned: *paraconsistent logic*. Paraconsistent logic is motivated by the thought we would often seem to have to reason sensibly from information, or about a situation, which is inconsistent. In such a case, the principle $A,\neg A\vdash B$ (*ex falso quodlibet sequitur*, Explosion), which is valid in classical logic, clearly makes a mess of things. A paraconsistent logic is precisely one where this principle fails.

There are many different families of paraconsistent logics—as many as there are ways of breaking Explosion. Indeed many of the techniques we have already met in this article can be used to construct a paraconsistent logic. The 3-valued logic *LP* is paraconsistent, as is the Łukasiewicz continuum-valued logic, provided we take the designated values to contain 0.5. The ways that negation is handled in relevant logic also produce paraconsistent logics, as long as validity is defined over a class of worlds in which A and $\neg A$ may both hold. Another approach (*discussive logic*) is to employ standard modal logic and to take A to hold in an interpretation iff A holds at some world of the interpretation. In this approach the principle of Adjunction ($A,B\vdash A\wedge B$) will generally fail, since A and B may each hold at a world, whilst $A\wedge B$ may not. Another approach ("positive plus") is to take any standard positive (negation free) logic, and add a non-truth-functional negation—so that the values of A and $\neg A$ are assigned independently. In these logics, the principle of Contraposition ($A\leftrightarrow B\vdash\neg B\leftrightarrow\neg A$) will generally fail. Yet another is to dualise intuitionist logic. In particular one can take semantic values to be the closed sets in some topology. If the value of A is X, the value of $\neg A$ is the closure of the complement of X.

SECOND-ORDER QUANTIFICATION

We now turn to the issue of quantification. In classical logic there are quantifiers $\forall$ and $\exists$. These range over a domain of objects, and $\forall xA(x)$ $[\exists xA(x)]$ holds if every [some] object in the domain of quantification satisfies $A(x)$. All the propositional logics we have looked at may be extended to first-order logics with such quantifiers. Other non-classical logics may be obtained by adding to these (or replacing these with) different kinds of quantifiers.

Perhaps the most notable of these is second-order logic. In this there are bindable variables (X, Y, …) that can stand in the place where a monadic first-order predicate can stand and which range over sets of objects in the first-order domain—canonically all of them. (There can also be variables that range over the n-ary relations on that domain, for each n, as well as variables that range over n-place functions. The second-order extension of classical logic is much stronger than the first-order version. It can provide for a categorical axiomatization of arithmetic and consequently is not itself axiomatizable.

Monadic second-order quantifiers can also be given a rather different interpretation, as plural quantifiers. The idea here is to interpret $\exists X\, Xa$ not as "There is a set such that a is a member of it," but as "There are some things such that a is one of them." The proponents of plural quantification argue that such quantification is not committed to the existence of sets.

OTHER SORTS OF QUANTIFIERS

There are many other non-classical quantifiers. For example one can have a binary quantifier of the form $Mx(A(x),B(x))$, "most As are Bs." This is true in a finite domain if more than half the things satisfying $A(x)$ satisfy $B(x)$. It is not reducible to a monadic quantifier plus a propositional connective.

Another sort of quantifier is a cardinality quantifier. The quantifier "there exist exactly n things such that" can be defined in first-order logic with quantification and identity in a standard way. The quantifier "there is a countable number of things such that" (or its negation, "there is an uncountable number of things such that") cannot be so defined—let alone the quantifier "there are κ things such that," for an arbitrary cardinal, κ. Such quantifiers can be added, with the obvious semantics. These quantifiers extend the expressive power of the language towards that of second-order logic—and beyond.

Another kind of quantifier is the branching quantifier. When, in first-order logic, we write:

$$\forall x_1 \exists y_1 \forall x_2 \exists y_2 A(x_1,x_2,y_1,y_2)$$

y_2 is in the scope of x_1, and so its value depends on that of x_1. To express non-dependence one would normally need second-order quantification, thus:

$$\exists f_1 \forall x_1 \exists f_2 \forall x_2 A(x_1,x_2,f_1(x_1),f_2(x_2))$$

But we may express it equally by having the quantifiers non-linearly ordered, thus:

$$\begin{matrix}\forall x_1 \exists y_1 \\ \forall x_2 \exists y_2\end{matrix}\ A(x_1, x_2, y_1, y_2)$$

As this would suggest, branching quantifiers have something of the power of second-order logic.

A quite different kind of quantifier is the substitutional quantifier. For this there is a certain class of names of the language, C. $\Pi x A(x)$ [$\Sigma x A(x)$] holds iff for every [some] $c \in C$, $A(c)$ holds. This is not the same as standard (objectual) quantification, since some objects in the domain may have no name in C; but first-order substitutional quantifiers validate the same quantificational inferences as first-order objectual quantifiers. Note that the notion of substitutional quantification makes perfectly good sense for any syntactically well-defined class, including predicates (so we can have second-order substitutional quantification) or binary connectives (so that $\Sigma x(AxB)$ can make perfectly good sense).

Finally in this category comes free quantifiers. It is standard to interpret the domain of objects of quantification (at a world) as comprising the objects that exist (at that world). It is quite possible, however, to think of the domain as containing a bunch of objects, some of which exist, and some of which do not. Obviously this does not change the formal properties of the quantifiers. But if one thinks of the domain in this way one must obviously not read $\exists x$ as 'there exists an x such that'; one has to read it simply as 'for some x'. Given this set-up, however, it makes sense to have existentially loaded quantifiers, $\forall^E$ and $\exists^E$, such that $\forall^E A(x)$ [$\exists^E A(x)$] holds (at a world) iff all [some] of the existent objects (at the world) satisfy $A(x)$. If there is a monadic existence predicate, E, these quantifiers can be defined in the obvious way, as (respectively): $\forall x(Ex \rightarrow A(x))$ and $\exists x(Ex \wedge A(x))$. Clearly, existentially loaded quantifiers will not satisfy some of the standard principles of quantification, such as $\forall^E x A(x) \rightarrow A(c)$, $A(c) \rightarrow \exists x^E A(x)$ (since the object denoted by 'c' may not exist). Some logics do not have the existentially unloaded quantifiers, just the loaded ones. These are usually called *free logics*.

NON-MONOTONIC LOGICS

It remains to say a word about one other kind of logic that is often categorized as non-classical. In all the logics we have been considering so far:

if $\Sigma \vdash A$ then $\Sigma \cup \Delta \vdash A$

(where Σ and Δ are sets of formulas): Adding extra premises makes no difference. This is called *monotonicity*. Logics in which this principle fails are called *non-monotonic logics*. Non-monotonic inferences can be thought of as inferences that are made with certain default assumptions. Thus I am told that something is a bird, and I infer that it can fly. Since most birds fly this is a reasonable con-

clusion. If, however, I also learn that the bird weighs 20 kg. (and so is an emu or an ostrich), the conclusion is no longer a reasonable one.

There are many kinds of non-monotonic logics, depending on what kind of default assumption is implemented, but there is a common structure that covers many of them. Interpretations, I, of the language come with a strict partial ordering, $\succ$ (often called a *preference ordering*). Intuitively, $I_1 \succ I_2$ means that the situation represented by I_1 is more normal (in whatever sense of normality is at issue) than that represented by I_2. (In particular cases it may be reasonable to suppose that $\succ$ has additional properties.) I is a *most normal model of* Σ iff every $B \in \Sigma$ holds in I, and there is no $J \succ I$ for which this is true. A follows from Σ iff A holds in every most normal model of Σ. As is clear a most normal model of Σ is not guaranteed to be a most normal model of $\Sigma \cup \Delta$. Hence monotonicity will fail. As might be expected there is a close connection between non-monotonic logics and conditional logics, in which the inference $A \rightarrow B \vdash (A \wedge C) \rightarrow B$ fails. Though non-monotonic logic has come to prominence in modern computational logic, it is just a novel and rigorous way of looking at the very traditional notion of non-deductive (inductive, ampliative) inference.

HISTORY, PERSONS, REFERENCES

We conclude this review of non-classical logics by putting the investigations discussed above in their historical context. References that may be consulted for further details are also given at the end of each paragraph. For a general introduction to propositional non-classical logics, see Priest (2001). Haack (1996) is a discussion of some of the philosophical issues raised by non-classical logics.

The first modern many-valued logics, the $Ł_n$ family, were produced by Jan Łukasiewicz in the early 1920s. (Emil Post also produced some many-valued logics about the same time.) Łukasiewicz's major philosophical concern was Aristotle's argument for fatalism. In this context he suggested a many-valued analysis of modality. Logics of the both/neither kind were developed somewhat later. Canonical statements of K_3 and LP were given (respectively) by Stephen Kleene in the 1950s and Graham Priest in the 1970s. $Ł_{\aleph}$ was first published by Łukasiewicz and Alfred Tarski in 1930. The intensive investigation of fuzzy logics and their applications started in the 1970s. A notable player in this area was Lotfi Zadeh. (Rescher 1969, Urquhart 2001– , Hájek 1998, Yager and Zadeh 1992.)

Modern modal logics were created in an axiomatic form by Clarence Irving Lewis in the 1920s. Lewis's concern was the paradoxes of the material conditional, and he suggested the strict conditional as an improvement. Possible-world semantics for modal logics were produced by a number of people in the 1960s, but principally Saul Kripke. The semantics made possible the systematic investigation of the rich family of modal logics. (Bull and Segerberg 2001– , Garson 2001– , Hughes and Cresswell 1996.)

The idea that the techniques of modal logics could be applied to notions other than necessity and possibility occurred to a number of people around the middle of the twentieth century. Tense logics were created by Arthur Prior, epistemic and doxastic logic were produced by Jaakko Hintikka, and deontic logics by Henrik von Wright. Investigations of provability logic were started in the 1970s by George Boolos and others. Dynamic logic was created by Vaughn Pratt and other logicians particularly interested in computation, including David Harrel, in the 1970s. (van Bentham 1988, Burgess 2001– , Thomason 2001– , Meyer 2001– , Åqvist 2001– , Boolos 1993, Harrel, Kozen, and Tiuryn 2001– .)

Conditional logics (with "sphere semantics") were proposed by David Lewis and Robert Stalnaker in the 1970s. They were formulated as multi-modal logics by Brian Chellas and Krister Segerberg a few years later (Harper, Stalnaker, and Pearce 1981, Nute and Cross 2001–).

The intuitionist critique of classical mathematics was started by Luitzen Egbertus Jan Brouwer in the early years of the twentieth century. This generated a novel kind of mathematics: intuitionist mathematics. Intuitionist logic, as such, was formulated by Arend Heyting and Andrei Kolmogorov in the 1920s. The intuitionist critique of mathematical realism was extended to realism in general by Michael Dummett in the 1970s (Dummett 1977, van Dalen 2001–).

Systems of relevant logic, in axiomatic form, came to prominence in the 1960s because of the work of Alan Anderson, Nuel Belnap and their students. World-semantics were produced by a number of people in the 1970s, but principally Richard Routley (later Sylvan) and Robert Meyer. The semantics made possible the investigation of the rich family of relevant logics. The four-valued semantics for negation is due to J. Michael Dunn (Dunn and Restall 2001– , Mares 2004).

Linear logic was produced by Jean-Yves Girard in the 1980s. Although many members of the class of sub-structural logics had been studied before, the fact that they could be viewed in a uniform proof-theoretic way, was not appreciated until the late 1980s. The formulation of quantum logic in terms of Hilbert spaces is due, essentially, to George Birkhoff and John von Neumann in the

1930s. The use of an abstract closure operator to give the semantics for non-distributive logics is due to Greg Restall. (Troelstra 1992, Restall 2000, Paoli 2002, Chihara and Giuntini 2001, Hughes 1989).

The first paraconsistent logic (discussive logic) was published by StanisŁaw Jaśkowski in 1948. Other non-adjunctive logics were later developed in the 1970s by Peter Schotch and Raymond Jennings. Newton da Costa produced a number of different paraconsistent logics and applications, starting with positive-plus logics in the 1960s. The paraconsistent aspects of relevant logic were developed by Priest and Routley in the 1970s. (Priest, Routley and Norman 1989, Priest 2001, Carnielli et al. 2001, Mortensen 1995).

Second-order quantification goes back to the origins of classical logic in the work of Gottlob Frege and Bertrand Russell. Its unaxiomatizability put it somewhat out of fashion for a number of years, but it made a strong come-back in the last years of the twentieth century. The notion of plural quantification was made popular by George Boolos in the 1980s. (Shapiro 1991, 2001–; Boolos 1984).

Quantifier phrases other than "some *A*" and "all *A*" are pervasive in natural language; and since Frege provided an analysis of the quantifier many different kinds have been investigated by linguists and logicians. Branching quantifiers were proposed by Jaakko Hintikka in the 1970s. Substitutional quantification came to prominence in the 1960s, put there particularly in connection with quantification into the scope of modal operators by Ruth Barcan Marcus. It was treated with suspicion for a long time, but was eventually given a clean bill of health by Kripke. Free logics were first proposed in the 1960s, by Karel Lambert and others (van der Does and van Eijck 1996, Barwise 1979, Kripke 1976, Bencivenga 2001–).

Non-monotonic logics started to appear in the logic/computer-science literature in the 1970s. There are many kinds. The fact that many of them could be seen as logics with normality orderings started to become clear in the 1980s (Shoham, 1988; Crocco, Fariñas del Cerro, and Herzig 1995; Brewka, Dix, and Konolige, 1997).

See also Aristotle; Brouwer, Luitzen Egbertus Jan; Combinatory Logic; Dummett, Michael Anthony Eardley; First-Order Logic; Frege, Gottlob; Fuzzy Logic; Gödel's Incompleteness Theorems; Hintikka, Jaako; Intensional Logic; Intuitionism and Intuitionistic Logic; Kripke, Saul; Lewis, Clarence Irving; Lewis, David; Łukasiewicz, Jan; Many-Valued Logics; Modal Logic; Neumann, John von; Non-Monotonic Logic; Platonism and the Platonic Tradition; Prior, Arthur Norman; Provability Logic; Quantifiers in Natural Language; Quantum Logic and Probability; Russell, Bertrand Arthur William; Second-Order Logic; Semantics; Tarski, Alfred; Wright, Georg Henrik von.

Bibliography

Åqvist, Leonard. "Deontic Logic." In Gabbay and Guenthner (2001–), vol. 8.

Bencivenga, Ermanno. "Free Logic." In Gabbay and Guenthner (2001–), vol. 5.

van Bentham, Johan. *A Manual of Intensional Logic*. Stanford, CA: CSLI, 1988.

van Bentham, Johan. "On Branching Quantifiers in English." *Journal of Philosophical Logic* 8 (1979): 47–80.

Boolos, George. *The Logic of Provability*. Cambridge, U.K.: Cambridge University Press, 1993.

Boolos, George. "To Be Is to Be the Value of a Variable (Or Some Values of Some Variables)." *Journal of Philosophy* 81 (1984): 430–449. Reprinted in Boolos's *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press, 1998.

Brewka, Gerhard, Jürgen Dix, and Kurt Konolige. *Non-Monotonic Logic: An Overview*. Stanford, CA: CLSI, 1997.

Bull, Robert A., and Krister Segerberg. "Basic Modal Logic." In Gabbay and Guenthner (2001–), vol. 3.*

Burgess, John A. "Basic Tense Logic." In Gabbay and Guenthner (2001–), vol. 7.*

Chihara, Maria-Luisa D., and Roberto Giuntini. "Quantum Logics." In Gabbay and Guenthner (2001–), vol. 6.*

Crocco, Gabriella, Luis Fariñas del Cerro, and Andreas Herzig. *Conditionals: From Philosophy to Computer Science*. Oxford: Oxford University Press, 1995.

van Dalen, Dirk. "Intuitionistic Logic." In Gabbay and Guenthner (2001–), vol. 5.*

van der Does, Jaap, and Jan van Eijck. "Basic Quantifier Theory." In *Quantifiers, Logic, and Language*, edited by Jaap van der Does and Jan van Eijck. Stanford, CA: CLSI, 1996.

Dummett, Michael. *Elements of Intuitionism*. Oxford: Oxford University Press, 1977.

Gabbay, Dov, and Franz Guenthner, eds. *Handbook of Philosophical Logic*. 2nd ed. Dordrecht: Kluwer Academic, 2001–. Articles marked * are (possibly) revised versions of chapters in the much shorter first edition, Dordrecht: Reidel, 1983–89.

Garson, James. "Quantification in Modal Logic." In Gabbay and Guenthner (2001–), vol. 3.*

Haack, Susan. *Deviant Logic, Fuzzy Logic: Beyond the Formalism*. Chicago: University of Chicago Press, 1996.

Hájek, Petr. *Metamathematics of Fuzzy Logic*. Dordrecht: Kluwer Academic, 1998.

Harper, William, Robert Stalnaker, and Glenn Pearce, eds. *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Dordrecht: Kluwer Academic, 1981.

Harrel, David, Dexter Kozen, and Jerzy Tiuryn. "Dynamic Logic." In Gabbay and Guenthner (2001–), vol. 4.*

Hughes, George, and Max Cresswell. *A New Introduction to Modal Logic*. London: Routledge, 1996.

Hughes, R. I. G. *The Structure and Interpretation of Quantum Mechanics*. Cambridge, MA: Harvard University Press, 1989.

Kripke, Saul. "Is There a Problem about Substitutional Quantification?" In *Truth and Meaning*, edited by Gareth Evans and John McDowell. Oxford: Oxford University Press, 1976

Mares, Edwin. *Relevant Logic*. Cambridge, U.K.: Cambridge University Press, 2004.

Meyer, John Jules. "Modal Epistemic and Doxastic Logic." In Gabbay and Guenthner (2001–), vol. 10.

Mortensen, Chris. *Inconsistent Mathematics*. Dordrecht: Kluwer Academic, 1995.

Paoli, Francesco. *Substructural Logics: A Primer*. Dordrecht: Kluwer Academic, 2002.

Priest, Graham. *Introduction to Non-Classical Logic*. Cambridge, U.K.: Cambridge University Press, 2001.

Priest, Graham, Richard Routley, and Jean Norman. *Paraconsistent Logic: Essays on the Inconsistent*. Munich: Philosophia Verlag, 1989.

Rescher, Nicholas. *Many-Valued Logic*. New York: McGraw Hill, 1969.

Restall, Greg. *An Introduction to Substructural Logics*. London: Routledge, 2000.

Shapiro, Stewart. *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press, 1991.

Shoham, Yoav. *Reasoning about Change*. Cambridge, MA: MIT Press, 1988.

Thomason, Richmond H. "Combinations of Tense and Modality." In Gabbay and Guenthner (2001–), vol. 7.*

Troelstra, Anne. *Lectures on Linear Logic*. Stanford, CA: CSLI, 1992.

Urquhart, Alasdair. "Basic Many-Valued Logic." In Gabbay and Guenthner (2001–), vol. 2.*

Yager, Ronald R., and Lotfi A. Zadeh. *An Introduction to Fuzzy Logic Applications in Intelligent Systems*. Dordrecht: Kluwer Academic, 1992.

Graham Priest (2005)

LOGIC, SYMBOLIC

See *Logic, History of*

LOGIC, TRADITIONAL

In logic, as in other fields, whenever there have been spectacular changes and advances, the logic that was current in the preceding period has been described as "old" or "traditional," and that embodying the new material has been called "new" or "modern." The Stoics described themselves as "moderns" and the Aristotelians as devotees of the "old" logic, in the later Middle Ages the more adventurous writers were called *moderni,* and since the latter part of the nineteenth century the immensely expanded logic that has developed along more or less mathematical lines ("mathematical logic," "symbolic logic," "logistics") has been contrasted with the "traditional" logic inherited from the sixteenth and seventeenth centuries. In every case the logic termed "old" or "traditional" has been essentially Aristotelian, but with a certain concentration on the central portion of the Aristotelian *corpus,* the theory of categorical syllogism—the logic of Aristotle himself having been rather less circumscribed than that of the "tradition," especially of the sixteenth to the nineteenth century.

THE LOGIC OF TERMS

To begin with the categorical syllogism, an inference, argument, or syllogism (traditionally, all arguments are assumed to be syllogistic) is a sequence of propositions (premises followed by a conclusion), such as "All animals are mortal; all men are animals; therefore, all men are mortal." Propositions, in turn, are built up from terms—for example, "animals," "mortals," "men." The traditional order of treatment, therefore, begins with the study of terms (or, in writers with a psychological or epistemological bias, ideas) and goes on to the study of propositions (or judgments), concluding with that of syllogisms (or inferences).

The terms from which the propositions principally studied in the traditional logic are built up are common nouns (*termini communes*), such as "man" and "horse," although some attention is also paid to singular terms, such as "Socrates," "this man," and "the man next door." Much of the traditional theory is devoted to the arrangement of common nouns in an order of comprehensiveness, and here a distinction is made between two aspects of their functioning—their "extension" (as the logicians of Port-Royal called it) or "denotation" (John Stuart Mill) and their "intension" (Sir William Hamilton), "comprehension" (Port-Royalists), or "connotation" (Mill). The extension or denotation of a common noun is the set of individuals to which it applies, its intension or connotation the set of attributes that an individual must possess for the common noun to be applicable to it. Thus, the connotation of the term *man* consists of the attributes of being an animal, being rational, and perhaps possessing a certain bodily form; its denotation consists of all objects that possess these attributes.

Broadly, the connotation of a term is its meaning, the denotation its application. The analysis of the meaning of a term is described as definition, and the breaking up of the set of objects to which it applies into subsets is described as division. The subsets of the set of individuals to which a

given term applies are called the species of the genus denoted by the given term. The attribute that marks off a particular species from others of the same genus is called its differentia. The species is said by scholastic logicians to "fall under" its genus, and the standard way of defining a species is by giving its genus and its differentia.

The ordering of terms into species and genera is often thought of as having an upper and a lower limit. The upper limit, or *summum genus,* will be a broad category such as "thing" (*substantia*)—horses are animals, animals are organisms, organisms are bodies, bodies are things. More abstract terms will come to an end in more abstract categories, such as "quality" or "relation" (scarlet is a species or kind of red, red is a color, color is a quality). The *infima species,* or lower limit, is a more difficult concept. Man, for example, is commonly given as an *infima species,* but are not men divisible into, for instance, dark-haired and fair-haired men? This is answered, from the point of view of intension, by dividing the attributes of an individual into those that constitute its essence or nature and those that are merely accidental, and genuine species are said to be marked off by "essential" attributes only; further subdivisions differentiated by "accidental" attributes, such as the color of a man's hair, are not counted as genuine species. This distinction is not recognized by some writers. Gottfried Wilhelm Leibniz counted all attributes of an individual as essential, so that someone would not be *that* individual if he were in the least respect different from what he is. At the other extreme, Mill said that "individuals have no essences," although he had a use for the term *essence* in connection with general terms: It is of the essence of being a man, for example, to be an animal, if being an animal is one of the attributes commonly employed in fixing the application of the word *man.*

An allied doctrine of Mill's is that the proper names of individuals, by contrast with common nouns, have no connotation, only denotation. We may not be able to think of a named individual without thinking of him as having certain attributes, but the purpose of a proper name is not to convey the fact that he has those attributes but only to identify him as *that* individual. This view has been criticized by various writers, on the ground, among others, that we cannot identify an object at all without knowing at least its *infima species.* Mill has also been criticized for using the same term, "denotation," both for the application of a common noun and for what is named by a proper name.

Common terms can be simple or complex. Some kinds of complexity are of logical interest—for example, the conjunctive combination exemplified by "blind man" (i.e., what is both blind and human) and the disjunctive combination exemplified by "man-or-beast." This kind of complexity is of interest because, for one thing, it links up with the previous topic, a blind man being a species (in the broad though not the narrow sense) of man and a man being a species (again in the broad sense) of man-or-beast (i.e., of animal). Again, the term "son-of-Philip" is compounded of the relative expression "son of" and the proper name "Philip," and this, too, links with the preceding topic, a son of Philip being a species (in the broad sense) of son. But the logical behavior of complex terms of these types is a topic of modern rather than traditional logic. Even traditional logic, however, has something to say about negative terms, such as "non-man" (i.e., what is not human), as will be shown in what follows.

The distribution of terms is a subject that will be more intelligible after propositions and syllogisms have been considered.

THE LOGIC OF PROPOSITIONS

OPPOSITION. The division of traditional logic called the logic of propositions is not to be confused with what is now called the propositional calculus. The propositional calculus studies the logical behavior of propositions formed from simpler propositions by means of various connectives (for example, "Either all men are liars or no men are"), as opposed to propositions formed not from other propositions but from terms (for example, "No men are liars"). The traditional logic of propositions or judgments, on the other hand, is chiefly concerned with the classification and simpler interrelations of precisely the second class of propositions, although it normally also touches on "compound" or "hypothetical" propositions, without going beyond their simplest types and the simplest inferences involving them.

Propositions not compounded of other propositions are called categorical. This word has the force of "unconditional," the implied contrast being with forms like "*If* all that the Bible says is true, all men are mortal" or "*Either* not all that the Bible says is true, *or* all men are mortal." Categoricals have a *subject term* and a *predicate term* ("men" is the subject term and "mortal" the predicate term of "All men are mortal") and are subdivided in two main ways—according to quantity, into universals ("All men are mortal," "No men are mortal") and particulars ("Some men are mortal," "Some men are not mortal"), and according to quality, into affirmatives ("All men are mortal," "Some men are mortal") and negatives ("No men are mortal," "Some men are not mortal"). These are often displayed in a square, with universals at the top, particulars at the bottom, affirmatives on the left, negatives on the right:

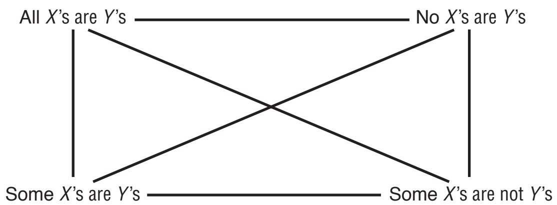

Universal affirmatives are called A-propositions, particular affirmatives I-propositions, universal negatives E-propositions, and particular negatives O-propositions (the vowels being taken from the words *affirmo* and *nego*). Two other "quantities" are commonly mentioned, namely *singular* and *indefinite.* Singular propositions, such as "Socrates is mortal," are a genuinely distinct type, which we shall touch upon at appropriate points; indefinites, such as "Men are mortal," seem merely to be universals or particulars in which the quantity is left unstated. The expressions other than terms which enter into these forms are called "syncategorematic"; they are divided into the signs of quantity "all" and "some" and the copulas "is" or "are" and "is not" or "are not." ("No" is both a sign of quantity and a sign of negation.)

These types of propositions—A, E, I, and O—are the traditional "four forms," and as a preliminary to logical manipulation it is customary to restate given sentences in some standard way that will make their quantity and quality immediately evident. The forms given above, with "all," etc., and with plural common nouns for terms, are the most widely used, but it is in some ways less misleading to use "every," etc., and the terms in the singular—"Every *X* is a *Y,*" "No *X* is a *Y,*" "Some *X* is a *Y,*" "Some *X* is not a *Y.*" What is important is to understand that "some" means simply "at least one"; "Some men are mortals" or "Some man is a mortal" must be understood as neither affirming nor denying that more than one man is a mortal and as neither affirming nor denying that all men are (i.e., "some" does *not* mean "only some").

A square of the type shown earlier is called a *square of opposition,* and propositions with the same terms in the same order may be "opposed" in four ways. Universals of opposite quality ("Every *X* is a *Y,*" "No *X* is a *Y*") are said to be *contraries*; these cannot be jointly true. Particulars of opposite quality ("Some *X* is a *Y,*" "Some *X* is not a *Y*") are said to be *subcontraries*; these cannot be jointly false. Propositions opposed only in quantity are said to be *subalterns,* the *subalternant* universal implying (without being implied by) the *subalternate* particular ("Every *X* is a *Y*" implies "Some *X* is a *Y,*" and "No *X* is a *Y*" implies "Some *X* is not a *Y*"). Propositions opposed in both quantity and quality ("Every *X* is a *Y*" and "Some *X* is not a *Y,*" and "No *X* is a *Y*" and "Some *X* is a *Y*") are *contradictories*; they cannot be jointly true or jointly false—the truth of a given proposition implies the falsehood of its contradictory; its falsehood implies the contradictory's truth.

EQUIPOLLENCE. Closely connected with the theory of opposition is that of the equipollence of propositions with the same terms in the same order but with negative particles variously placed within them. Since contradictories are true and false under reversed conditions, any proposition may be equated with the simple denial of its contradictory. Thus, "Some *X* is not a *Y*" has the same logical force as "Not every *X* is a *Y,*" and, conversely, "Every *X* is a *Y*" has the force of "Not (some *X* is not a *Y*)," or, to give it a more normal English expression, "Not any *X* is not a *Y.*" Similarly, "Some *X* is a *Y*" has the force of "Not (no *X* is a *Y*)" and "No *X* is a *Y*" that of "Not (some *X* is a *Y*)"—that is, "Not any *X* is a *Y.*" Also, since "no" conveys universality and negativeness at once, "No *X* is a *Y*" has the force of "Every *X* is not-a-*Y,*" and, conversely, "Every *X* is a *Y*" has the force of "No *X* is not-a-*Y.*" Writers with an interest in simplification have seen in these equivalences a means of dispensing with all but one of the signs "every," "some," and "no." Thus the four forms may all be expressed in terms of "every," as follows: "Every *X* is a *Y*" (A), "Every *X* is not-a-*Y*" (E), "Not every *X* is not-a-*Y*" (I), "Not every *X* is a *Y*" (O).

Of singular propositions all that need be said at this point is that they divide into affirmatives ("Socrates is mortal," "This is a man," "This man is mortal") and negatives ("Socrates is not mortal," etc.) and that when their subject is formed by prefixing "this" to a common noun (as in "This man is mortal"), the singular form is implied by the corresponding universal ("Every man is mortal") and implies the corresponding particular ("Some man is mortal"). Some of the traditional logicians attempted to assimilate singular propositions to particulars, some to assimilate them to universals, but these attempts are not very impressive, and it is one of the few merits of the Renaissance logician Peter Ramus that he and his followers treated them consistently as a type of their own.

CONVERSION OF PROPOSITIONS. With regard to pairs of propositions of the same form and with the same terms, but in reverse order—for example, "No *X* is a *Y*" and "No *Y* is an *X*"—these are sometimes equivalent and sometimes not. Where they are, as in the case just given, they are said to be *converses* of one another, and the forms are said to be convertible. E and I are convertible; A and O are not. That every man is an animal, for example, does not imply that every animal is a man, and that some animal is not a horse does not imply that some horse is not

an animal. Conversion, the inference from a given proposition to its converse ("Some men are liars; therefore, some liars are men"), is a type of immediate inference—that is, inference involving only one premise (as opposed, for instance, to syllogisms, which have two). Other immediate inferences are those from a given proposition to an "equipollent" form in the sense of the preceding section (for example, "Every man is mortal; therefore, not any man is not") and from a subalternant universal to its subalternate particular ("Every man is mortal; therefore, some man is mortal").

The conversion just described is "simple" conversion; with universals (even A, though it is not "simply" convertible) there is also a conversion *per accidens,* or *subaltern* conversion—that is, a legitimate inference to the corresponding particular form with its terms transposed. Thus, although "Every man is an animal" does not imply that every animal is a man, it does imply that some animal is.

Other forms of immediate inference arise when negative terms are introduced. The simultaneous interchange and negation of subject and predicate is called *conversion by contraposition,* or simply contraposition. It is a valid process with A's and O's, not with E's and I's. ("Every man is an animal" implies "Every non-animal is a non-man"—whatever is not an animal is not a man—and "Not every animal is a man" implies "Not every non-man is a non-animal," but "No horse is a man" does not imply "No non-man is a non-horse"; "Some *X* is a *Y*" is true and "Some non-*Y* is a non-*X*" false if the *X*'s and the *Y*'s overlap and between them exhaust the universe.) All of the four forms may be "obverted" (Alexander Bain's term)—that is, have their quality changed and the predicate negated ("Every *X* is a *Y*" implies "No *X* is a non-*Y*," "No *X* is a *Y*" implies "Every *X* is a non-*Y*," and similarly with the particulars). A variety of names are given to the results of repeated successive obversion and conversion.

THE LOGIC OF SYLLOGISM

A categorical syllogism is the inference of one categorical proposition, the conclusion, from two others, the premises, each premise having one term in common with the conclusion and one term in common with the other premise—for example:

Every animal is mortal;
Every man is an animal;
Therefore, every man is mortal.

The predicate of the conclusion (here "mortal") is called the major term, and the premise that contains it (here written first) the major premise. The subject of the conclusion ("man") is the minor term, and the premise that contains it (here written second) the minor premise. The term common to the two premises ("animal") is the middle term.

FIGURES AND MOODS. Syllogisms are divided into four figures, according to the placing of the middle term in the two premises. In the first figure the middle term is subject in the major premise and predicate in the minor; in the second figure predicate in both; in the third figure subject in both; in the fourth predicate in the major and subject in the minor. The following schemata, with *P* for the major term, *S* for the minor, and *M* for the middle, sum up these distinctions:

Figure 1	Figure 2	Figure 3	Figure 4
$M - P$	$P - M$	$M - P$	$P - M$
$S - M$	$S - M$	$M - S$	$M - S$
$S - P$	$S - P$	$S - P$	$S - P$

Within each figure, syllogisms are further divided into *moods,* according to the quantity and quality of the propositions they contain.

Not all of the theoretically possible combinations of propositions related as above constitute valid syllogisms, sequences in which the third proposition really follows from the other two. For example, "Every man is an animal; some horse is an animal; therefore, no man is a horse" (mood AIE in Figure 2) is completely inconsequent (even though all three propositions happen in this case to be true). During the Middle Ages those syllogistic moods that are valid acquired certain short names, with the mood indicated by the vowels, and all of them were put together in a piece of mnemonic doggerel, of which one of the later versions is the following:

Barbara, Celarent, Darii, Ferioque prioris;
Cesare, Camestres, Festino, Baroco secundae;
Tertia *Darapti, Disamis, Datisi, Felapton,*
Bocardo, Ferison habet. Quarta insuper addit
Bramantip, Camenes, Dimaris, Fesapo, Fresison.

Here Bocardo, for example, means the mood OAO in Figure 3, of which an illustration (C. S. Peirce's example) would be

Some patriarch (viz., Enoch) is not mortal;
Every patriarch is a man;
Therefore, some man is not mortal.

There is also a group of moods (Barbari and Celaront in Figure 1, Cesaro and Camestrop in Figure 2, Camenop in Figure 4) in which a merely particular conclusion is drawn although the premises would warrant our going further

and making the conclusion universal (the "subaltern" moods). The Ramists added special moods involving singulars (if we write S and N for affirmative and negative singulars, we have ASS and ESN in Figure 1, ANN and ESN in Figure 2 and SSI and NSO in Figure 3). It may be noted that every syllogism must have at least one universal premise, except for SSI and NSO in Figure 3—the so-called expository syllogisms, for example, "Enoch is not mortal; Enoch is a patriarch; therefore, not every patriarch is mortal." Moreover, every syllogism must have at least one affirmative premise, and if either premise is negative or particular, the conclusion must be negative or particular, as the case may be ("the conclusion follows the weaker premise," as Theophrastus put it, negatives and particulars being considered weaker than affirmatives and universals).

REDUCTION. The mnemonic verses serve to indicate how the valid moods of the later figures may be "reduced" to those of Figure 1—that is, how we may derive their conclusions from their premises without using any syllogistic reasoning of other than the first-figure type. (This amounts, in modern terms, to proving their validity from that of the first-figure moods taken as axiomatic.) In the second-figure mood Cesare, for example, the letter *s* after the first *e* indicates that if we *simply convert* the major premise we will have a pair of premises from which we can deduce the required conclusion in Figure 1, and the initial letter *C* indicates that the first-figure mood employed will be Celarent. An example of a syllogism in Cesare (EAE in Figure 2) would be

> No horse is a man;
> Every psychopath is a man;
> Therefore, no psychopath is a horse.

This conclusion may equally be obtained from these premises by proceeding as follows:

No horse is a man—*s*⟶No man is a horse;
Every psychopath is a man → Every psychopath is a man;
Therefore, no psychopath is a horse.

Here the right-hand syllogism, in which the first premise is obtained from the given major by simple conversion and the second is just the given minor unaltered, is in the mood Celarent in the first figure. Festino "reduces" similarly to Ferio, and Datisi and Ferison (in the third figure) reduce to Darii and Ferio, though in the third-figure cases it is the minor premise that must be simply converted. Darapti and Felapton reduce to Darii and Ferio by conversion of the minor premise, not simply, but *per accidens* (this is indicated by the *s* of the other moods being changed to *p*).

Camestres (Figure 2) and Disamis (Figure 3) are a little more complicated. Here we have not only an *s,* for the simple conversion of a premise, but also an *m,* indicating that the premises must be transposed, and a further *s* at the end because the transposed premises yield, in Figure 1, not the required conclusion but rather its converse, from which the required conclusion must be obtained by a further conversion at the end of the process. An example in Disamis would be the following:

> Some men are liars;
> All men are automata;
> Therefore, some automata are liars.

If we convert the major premise and transpose the two, we obtain the new pair

> All men are automata;
> Some liars are men,

and from these we may obtain in the first-figure mood Darii not immediately the conclusion "Some automata are liars" but rather "Some liars are automata," from which, however, "Some automata are liars" does follow by simple conversion.

Baroco and Bocardo are different again. In both of them neither premise is capable of simple conversion, and if we convert the A premises *per accidens* we obtain pairs IO and OI, and there are no valid first-figure moods with such premises—in fact, no valid moods at all with two particular premises. We therefore show that the conclusion follows from the premises by the device called reductio ad absurdum. That is, we assume for the sake of argument that the conclusion does not follow from the premises—that is, that the premises can be true and the conclusion false—and from this assumption, using first-figure reasoning alone, we deduce impossible consequences. The assumption, therefore, cannot stand, so the conclusion does after all follow from its premises.

Take, for example, the following syllogism in Baroco (AOO in Figure 2):

> Every man is mortal;
> Some patriarch (viz., Enoch) is not mortal;
> Therefore, some patriarch is not a man.

Suppose the premises are true and the conclusion is not. Then we have

> (1) Every man is mortal;
> (2) Some patriarch is not mortal;
> (3) Every patriarch is a man.

(This is the contradictory of the conclusion.) But from (1) and (3), in the first-figure mood Barbara, we may infer

(4) Every patriarch is mortal.

However, the combination of (2) and (4) is impossible. Hence, we can have both (1) and (2) only if we drop (3)—that is, if we accept the conclusion of the given second-figure syllogism.

It is possible to "reduce" all the second-figure and third-figure moods to Figure 1 by this last method, and although this procedure is a little complicated, it brings out better than the other reductions the essential character of second-figure and third-figure reasoning. Figure 1 is governed by what is called the *dictum de omni et nullo,* the principle that what applies to all or none of the objects in a given class will apply or not apply (as the case may be) to any given member or subclass of this class. As Immanuel Kant preferred to put it, first-figure reasoning expresses the subsumption of cases under a rule—the major premise states some affirmative or negative rule ("Every man is mortal," "No man will live forever"), the minor asserts that something is a case, or some things are cases, to which this rule applies ("Enoch and Elijah are men"), and the conclusion states the result of applying the rule to the given case or cases ("Enoch and Elijah are mortal," "Enoch and Elijah will not live forever"). Hence, in Figure 1 the major premise is always universal (that being how rules are expressed) and the minor affirmative ("Something *is* a case").

Second-figure reasoning also begins with the statement of a rule ("Every man is mortal") but in the minor premise denies that we have with a given example the result which the rule prescribes ("Enoch and Elijah are *not* mortal," "Enoch and Elijah *will* live forever") and concludes that we do not have a case to which the rule applies ("Enoch and Elijah cannot be men"). It combines, in effect, the first-figure major with the contradictory of the first-figure conclusion to obtain the contradictory of the first-figure minor (compare the "reduction" of Baroco). A second-figure syllogism, in consequence, must have a universal major, premises opposed in quality, and a negative conclusion. Its practical uses are in refuting hypotheses, as in medicine or detection ("Whoever has measles has spots, and this child has no spots, so he does not have measles"; "Whoever killed *X* was a person of great strength, and *Y* is not such a person, so *Y* did not kill *X*").

In the third figure we begin by asserting that something or other does not exhibit the result which a proposed rule would give ("Enoch and Elijah are *not* mortal," "Enoch and Elijah *will* live forever"), go on to say that we nevertheless do have here a case or cases to which the rule would apply if true ("Enoch and Elijah *are* men"), and conclude that the rule is not true ("Not all men are mortal," "Some men do live forever"). A third-figure syllogism, consequently, has an affirmative minor (the thing is a case) and a particular conclusion (the contradictory of a universal being a particular); its use is to confute rashly assumed rules, such as proposed scientific laws.

This rather neat system of interrelations (first clearly brought out by C. S. Peirce) concerns only the first three figures; it was not until the later Middle Ages, in fact, that a distinct fourth figure was recognized. The common division of figures assumes that we are considering completed syllogisms, with the conclusion (and its subject and predicate) already before us; however, the question Aristotle originally put to himself was not "Which completed syllogisms are valid?" but "Which pairs of premises will yield a syllogistic conclusion?" Starting at this end, we cannot distinguish major and minor premises as those containing, respectively, the predicate and subject of the conclusion. Aristotle distinguished them, in the first figure, by their comparative comprehensiveness and mentioned what we now call the fourth-figure moods as odd cases in which first-figure premises will yield a conclusion wherein the "minor" term is predicated of the "major." Earlier versions of the mnemonic lines accordingly list the fourth-figure moods with the first-figure ones and (since the premises are thought of as being in the first-figure order) give them slightly different names (Baralipton, Celantes, Dabitis, Fapesmo, Frisesomorum).

DISTRIBUTION OF TERMS. Terms may occur in A-, E-, I-, and O-propositions as distributed or as undistributed. The rule is that universals distribute their subjects and particulars distribute their predicates, but what this means is seldom very satisfactorily explained. It is often said, for example, that a distributed term refers to all, and an undistributed term to only a part, of its extension. But in what way does "Some men are mortal," for example, refer to only a part of the class of men? Any man whatever will do to verify it; if any man whatever turns out to be mortal, "Some men are mortal" is true. What the traditional writers were trying to express seems to be something of the following sort: A term t is distributed in a proposition $f(t)$ if and only if it is replaceable in $f(t)$, without loss of truth, by any term "falling under it" in the way that a species falls under a genus. Thus, "man" is distributed in

Every man is an animal;
No man is a horse;
No horse is a man;
Some animal is not a man,

since these respectively imply, say,

> Every blind man is an animal;
> No blind man is a horse;
> No horse is a blind man;
> Some animal is not a blind man.

On the other hand, it is undistributed in

> Some man is keen-sighted;
> Some man is not disabled;
> Every Frenchman is a man;
> Some keen-sighted animal is a man,

since these do not respectively imply

> Some blind man is keen-sighted;
> Some blind man is not disabled;
> Every Frenchman is a blind man;
> Some keen-sighted animal is a blind man.

In this sense A- and E- propositions do distribute their subjects and E- and O-propositions their predicates. John Anderson pointed out that the four positive results above may be established syllogistically, given that all the members of a species (using the term widely) are members of its genus—in the given case, that all blind men are men. From "Every man is an animal" and "Every blind man is a man," "Every blind man is an animal" follows in Barbara; with the second example the syllogism is in Celarent, with the third in Camestres, with the fourth in Baroco. Note, however, that the mere prefixing of "every" to a term is not in itself sufficient to secure its "distribution" in the above sense; for example, "man" is not distributed in "Not every man is disabled," since this does not imply "Not every blind man is disabled."

For a syllogism to be valid the middle term must be distributed at least once, and any term distributed in the conclusion must be distributed in its premise (although there is no harm in a term's being distributed in its premise but not in the conclusion). Many syllogisms can quickly be shown to be fallacious by the application of these rules. "Every man is an animal; every horse is an animal; therefore, every horse is a man," for example, fails to distribute the middle term "animal," and it is clear that any second-figure syllogism with two affirmative premises would have the same fault (since in the second figure the middle term is predicate twice, and affirmatives do not distribute their predicates). Other special rules for the different figures, such as that in Figures 1 and 3 the minor premise must be affirmative, can be similarly proved from the rules of distribution together with the rules of quality (that a valid syllogism does not have two negative premises, and that a conclusion is negative if and only if one premise is). Logicians have endeavored to prove some of these rules from others and to reduce the number of unproved rules to a minimum.

EULER'S DIAGRAMS. One device for checking the validity of syllogistic inferences is the use of certain diagrams attributed to the seventeenth-century mathematician Leonhard Euler, although their accurate employment seems to date rather from J. D. Gergonne, in the early nineteenth century.

From the traditional laws of opposition and conversion it can be shown that the extensions of any pair of terms *X*, *Y* will be related in one or another of five ways: (α) every *X* is a *Y* and every *Y* is an *X*, that is, their extensions coincide; or (β) every *X* is a *Y*, but not every *Y* is an *X*, that is, the *X*'s form a proper part of the *Y*'s; or (γ) every *Y* is an *X*, but not every *X* is a *Y*, that is, the *Y*'s form a proper part of the *X*'s; or (δ) some but not all *X*'s are *Y*'s and some but not all *Y*'s are *X*'s, that is, the *X*'s and *Y*'s overlap; or (ϵ) no *X*'s are *Y*'s and so no *Y*'s are *X*'s, that is, the *X*'s and *Y*'s are mutually exclusive. These five cases are represented by the following diagrams:

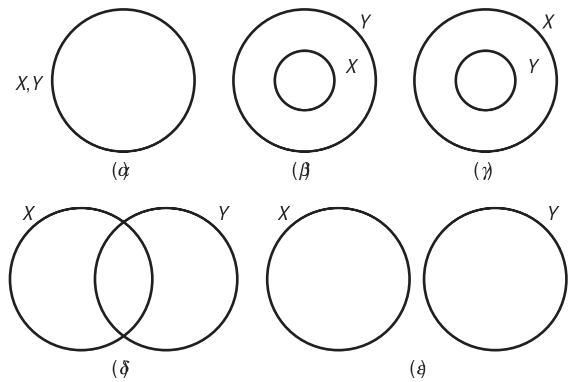

"Every *X* is a *Y*" (A) is true if and only if we have either (α) or (β); "Some *X* is not a *Y*" (O) if and only if we have either (γ) or (δ) or (ϵ); "No *X* is a *Y*" (E) if and only if we have (ϵ); and "Some *X* is a *Y*" (I) if and only if we have either (α) or (β) or (γ) or (δ). From these facts it follows that A and O are in no case true together and in no case false together, and similarly for E and I; that I is true in every case in which A is and also in two cases in which A is not, and similarly for O and E; that A and E are in no case true together but in two cases are both false; and that O and I are in no case both false but in two cases are both true. After working out analogous truth conditions for the forms with reversed terms, we will see that they are the same for the two I's and the two E's (showing that these are simply convertible) but not for the two A's and the two O's (showing that these are not). Given which of

the five relations holds between *X* and *Y* and which between *Y* and *Z*, we can work out by compounding diagrams what will be the possible relations between *X* and *Z*. For example, if we know that every *X* is a *Y* and every *Y* a *Z*, then we must have either $(\alpha)XY$ and $(\alpha)YZ$ or $(\alpha)XY$ and $(\beta)YZ$ or $(\beta)XY$ and $(\alpha)YZ$ or $(\beta)XY$ and $(\beta)YZ$; that is, we must have

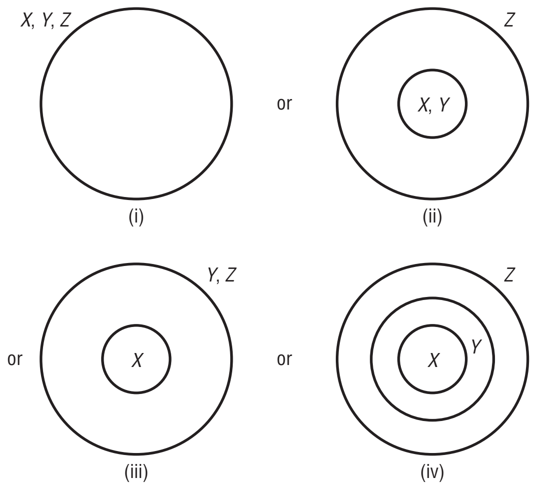

Inspection will show that for *X* and *Z* we have in every case either

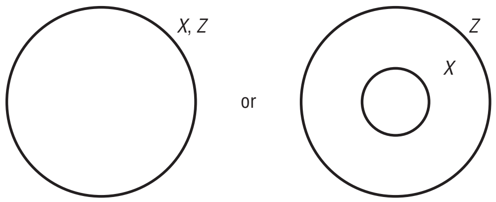

so in every case every *X* is a *Z*. Hence, Barbara is valid.

When employing this procedure it is essential to consider all the possible cases involved. Barbara is not validated, for example, by considering case (iv) alone, as popular expositions of this method sometimes suggest.

POLYSYLLOGISMS, ENTHYMEMES, AND INDUCTION. In an extended argument the conclusion of one inference may be used as a premise of another, and the conclusion of that as premise of a third, and so on. In presenting such an argument we may simply omit the intermediate steps and list all the premises together. For example, the sequence of categorical syllogisms "Every *X* is a *Y*, and every *Y* is a *Z*, so every *X* is a *Z*; and every *Z* is a *T*, so every *X* is a *T*" may be condensed to "Every *X* is a *Y*, every *Y* is a *Z*, and every *Z* is a *T*; therefore, every *X* is a *T*." Such a condensed chain of syllogisms is called a polysyllogism or sorites. The theory of chains of two syllogisms was thoroughly studied by Galen, as reported in an ancient passage unearthed by Jan Łukasiewicz. Galen showed that the only combinations of the Aristotelian three figures that could be thus used were 1 and 1, 1 and 2, 1 and 3, and 2 and 3. His discovery of these four types of compound syllogism was misunderstood by later writers as an anticipation of the view that single syllogisms may be of four figures.

Even when it is not a conclusion from other premises already stated, one of the premises of an inference may often be informally omitted (for example, "Enoch and Elijah are men; therefore, Enoch and Elijah are mortals"). Such a truncated inference is often called an enthymeme. This is not Aristotle's own use of the term, though he did mention that a premise is often omitted in the statement of an enthymeme in his sense. An Aristotelian enthymeme is a merely probable argument—that is, one in which the conclusion does not strictly follow from the premises but is merely made more likely by them. When the claim made for an argument is thus reduced, the normal rules may be relaxed in certain directions; in particular, the second and third figures may be used to yield more than merely negative results. Thus, Figure 2 may be used not only to prove that something is not a case falling under a given rule but also to suggest that it is one—to use a modern example:

> Any collection of particles whose movement is accelerated will occupy more space than it did;
>
> A heated gas will occupy more space than it did;
>
> Therefore, a heated gas may be a collection of particles whose movement is accelerated.

Figure 3 may be similarly used not only to prove that some rule does not hold universally but also to suggest that it does hold universally—for instance:

> *X, Y, Z* are all of them white;
> *X, Y, Z* are all of them swans;
> Therefore, perhaps all swans are white.

If the second premise here is strengthened to "*X, Y, Z* are all the swans there are," the conclusion will follow without any "perhaps" (of course, the new premise is in this case a false one, and the conclusion is also false). The form of inference

> *X, Y, Z,* etc., are all of them *P*'s;
> *X, Y, Z,* etc., are all the *S*'s there are;
> Therefore, all *S*'s are *P*'s

was called by Aristotle "induction"; more accurately, he used this term for a similar passage from all the sub-

species to their genus ("The *X*'s, the *Y*'s, and the *Z*'s are all of them *P*'s and are all the *S*'s; therefore, …"). He observed that the "conversion" of the second premise to "All the *S*'s are the *X*'s, the *Y*'s, and the *Z*'s" will turn such an induction into a syllogism in Barbara.

The term *induction* being extended in the more recent tradition to cover the merely probable inference given just previously, we distinguish Aristotelian induction by calling it "formal" or "perfect" induction or (as W. E. Johnson called it) "summary" induction. The Figure 2 type of merely probable inference is one of the things meant by the term "argument from"—or "by"—"analogy" (or just "analogy"); C. S. Peirce called it "hypothesis."

SKEPTICAL CRITICISMS OF SYLLOGISTIC REASONING. In the latter part of the nineteenth century, under the influence of J. S. Mill, textbooks of the traditional type came to have two main divisions, "formal" or "deductive" logic (dealt with more or less as above) and "inductive" logic or "scientific method." With the details of inductive logic we are not concerned here, but we may glance at the view of some writers that merely probable induction and analogy are the only genuine types of reasoning, "formal" or syllogistic reasoning being useless or spurious because it is inevitably circular, assuming in the premises what it sets out to prove as the conclusion.

The second-century skeptic Sextus Empiricus suggested that in the syllogism "Every man is an animal; Socrates is a man; therefore, Socrates is an animal," the only way to establish the major premise is by induction; however, if the induction is incomplete the examination of a new instance—for example, of Socrates—might prove it false, and if it is complete the conclusion ("Socrates is an animal") must already have been used in establishing it. This argument was repeated by such writers as George Campbell, in the eighteenth century, who supplemented it with another, to cover the case in which the major is established not by induction but simply by definition or linguistic convention: "Of course every man is an animal, for being an animal is part of what we mean by being a man." In this case it is the minor premise, "Socrates is a man," that cannot be established without first establishing the conclusion (that he is an animal). The same point was urged by another Scottish philosopher, Thomas Brown. It is allied to an argument used by Sextus to show not that syllogism is circular but that the major premise is superfluous. If, he said, every man is an animal because it follows from an object's being a man that it is an animal, then the allegedly enthymematic "Socrates is a man; therefore, Socrates is an animal" must be valid as it stands.

Richard Whately, answering Campbell's arguments in the early nineteenth century, complained that Campbell had confined himself to examples in which the syllogistic argument was indeed superfluous and countered them with some in which it was not—for example, the case of some laborers, ignorant of the fact that all horned animals are ruminant, digging up a skeleton which they, but not a distant naturalist, could see to be horned, the laborers and the naturalist thus separately providing premises which were both required to obtain the conclusion that the skeleton was of a ruminant animal. Whately admitted that the sense in which we may make a "discovery" by drawing a syllogistic conclusion is different from that in which we make a discovery by observation, but it can be a genuine discovery none the less; there are "logical" as well as "physical" discoveries.

After Whately, J. S. Mill took up the argument, but it is not entirely clear what side he was on. Sometimes he treated a universal major as already asserting, among other things, the conclusion:

> Whoever pronounces the words, All men are mortal, has affirmed that Socrates is mortal, though he may never have heard of Socrates; for since Socrates, whether known to be so or not, really is a man, he is included in the words, All men, and in every assertion of which they are the subject. (*System of Logic,* Book II, Ch. 3, p. 8, note)

"Included in the *meaning* of the words," he must have meant (for it is obvious that neither Socrates the man nor "Socrates," his name, forms any part of the words "All men"), but this contradicts Mill's own insistence that the meaning of general terms like "men" lies wholly in their "connotation" and that "All men are mortal" means that wherever the attributes of humanity are present, mortality is present, too. He rightly chided Brown, who thought that the meaning of "Socrates is mortal" (like that of "Socrates is an animal") is already contained in the minor premise "Socrates is a man," for failing to distinguish the actual connotation of "man" (i.e., the attributes by which its application is determined) from other attributes (such as mortality) which we may empirically discover these to be attended with, but his own view in the passage cited is similarly negligent.

Mill's main point, however, is different and more defensible. When careful and extensive observation warrants the conclusion that, say, all men are mortal, and we

then observe that the duke of Wellington is a man and conclude that he is therefore mortal, we have in effect an induction followed by a syllogism. Mill pointed out that if this procedure is justified at all, the introduction of the syllogistic major is superfluous. For if the original body of evidence really does warrant the inference that all men are mortal, it is certainly sufficient to warrant the inference that the duke of Wellington is mortal, given that he is a man. In other words, if we really are justified in the move from particular observations to the general proposition, and from there to new particulars, we would be equally justified in moving directly "from particulars to particulars."

What the syllogistic major does, Mill argued, is simply to sum up in a single formula the entire class of inferences to new particulars which the evidence warrants. That is, "All men are mortal" means, in effect, that if we ever find anyone to be a man we are justified in inferring, from the observations we have previously amassed, that he is mortal. "The conclusion is not an inference drawn *from* the formula"—that is, from "All men are mortal" thus understood—"but an inference drawn *according* to the formula" (ibid., p. 4). Mill here anticipated Gilbert Ryle's treatment of "lawlike statements" as "inference licenses" and echoed Sextus's point that it is inconsistent to require that such licenses be added to the premises of the inferences they permit, since what they license is precisely the drawing of the conclusion from those premises.

Mill in fact here shifted the discussion from Sextus's first skeptical "topic" to his second—from the charge of circularity to the question of what distinguishes a rule of inference from a premise. On this point more was said later in the nineteenth century by C. S. Peirce. Peirce, like Mill, distinguished sharply between the premise or premises from which, and the "leading principle" according to which, a conclusion is drawn. He also noted, as did Mill, that what is traditionally counted as a premise may function in practice as a "leading principle." But it need not, and, indeed, what is traditionally counted as a "leading principle" (say the *dictum de omni et nullo*) may sometimes be, conversely, treated in practice as a premise. Certainly, since *all men are mortal* (leading principle 1), we are justified in inferring the mortality of Socrates (or the duke of Wellington, or Elijah) from his humanity. But equally, since *all members of any class are also members of any class that contains the former as a subclass* (leading principle 2), we are justified in inferring the mortality of Socrates from his being a man *and* from men's being a subclass of mortals. For the very same reason (that all members of any class are also members of any class that contains the former as a subclass) we are justified in inferring the mortality of Socrates from his being a member of a subclass of the class of mortals *and* from the membership of any member of a class in all classes of which it is a subclass. In this last example we have one and the same proposition functioning as a premise and as a leading principle in the same inference (not merely, like "All men are mortal" in the preceding two examples, as a leading principle in one and a premise in another); to be capable of this, Peirce thought, is the mark of a "logical" leading principle.

It is not certain that Peirce's method of distinguishing "logical" from other sorts of "leading principles" will bear inspection. However, he seems to have established his basic point, that what it would be fatal to require in all cases—the treatment of a leading principle as a premise—we may safely permit in some. There may be useful and valid reasoning about subjects of all degrees of abstraction, including logic itself.

HYPOTHETICAL AND DISJUNCTIVE SYLLOGISMS. Traditional textbooks, aside from developing the theory of categorical propositions and syllogisms, have a brief appendix mentioning "hypothetical" (or "conditional") and "disjunctive" propositions and certain "syllogisms" to which they give rise.

"Hypothetical" syllogisms are divided into "pure," in which premises and conclusion are all of the form "If *p* then *q*" (notably the syllogism "If *p* then *q*, and if *q* then *r*; therefore, if *p* then *r*," analogous to Barbara), and "mixed," in which only one premise is hypothetical and the other premise and the conclusion are categorical. The mixed hypothetical syllogism has two valid "moods":

(1) *Modus ponendo ponens*: If *p* then *q*, and *p*; therefore, *q*.

(2) *Modus ponendo tollens*: If *p* then *q*, but not *q*; therefore, not *p*.

In both these moods the hypothetical premise is called the major, the categorical the minor. *Ponere,* in the mood names, means to affirm, *tollere* to deny. In (1), by affirming the antecedent of the hypothetical we are led to affirm its consequent; in (2), by denying its consequent we are led to deny its antecedent. The fallacies of "affirming the consequent" and "denying the antecedent" (i.e., of doing these things *to start with,* in the minor premise) consist in reversing these procedures—that is, in arguing "If *p* then *q*, and *q*; therefore, *p*" and "If *p* then *q*, but not *p*; therefore, not *q*."

"Disjunctive" syllogisms—that is, ones involving "Either-or" propositions—have the following two "mixed" moods:

(3) *Modus tollendo ponens*: Either *p* or *q*, but not *p*; therefore, *q* (or, but not *q*; therefore, *p*).

(4) *Modus ponendo tollens*: Either *p* or *q*, and *p*; therefore, not *q* (or, and *q*; therefore, not *p*).

Mood (4) is valid only if "Either *p* or *q*" is interpreted "exclusively"—that is, as meaning "Either *p* or *q* but not both"—whereas (3) is valid even if it is interpreted as "Either *p* or *q* or both." There is also a *modus tollendo ponens* with the simple "Not both *p* and *q*" as major and the rest as in (4).

DILEMMAS. Hypothetical and disjunctive premises may combine to yield a categorical conclusion in the *dilemma,* or "horned" syllogism (*syllogismus cornutus*), with its two forms:

(5) *Constructive*: If *p* then *r*, and if *q* then *r*, but either *p* or *q*; therefore, *r.*

(6) *Destructive*: If *p* then *q*, and if *p* then *r*, but either not *q* or not *r*; therefore, not *p.*

These basic forms have a number of variations; for instance, *q* in (5) may be simply "not *p*," making the disjunctive premise the logical truism "Either *p* or not *p*"; or *p* may imply *r* and *q* imply *s*, giving as conclusion "Either *r* or *s*" rather than the categorical *r*; or the disjunctive premise may be conditionalized to "If *s* then either *p* or *q*," making the conclusion "If *s* then *r*."

A typical dilemma is that put by Protagoras to Euathlus, whom he had trained as a lawyer on the understanding that he would be paid a fee as soon as his pupil won a case. When the pupil simply engaged in no litigation at all, Protagoras sued him for the fee. His argument was "If Euathlus wins this case, he must pay my fee by our agreement, and if he loses it he must pay it by the judge's decision (for that is what losing this case would mean), but he must either win or lose the case; therefore, in either case he must pay."

"Escaping between the horns" of a dilemma is denying the disjunctive premise; for example, Euathlus might have argued that he would neither win nor lose the case if the judge refused to make any decision. "Taking a dilemma by the horns" is admitting the disjunction but denying one of the implications, as Euathlus might have done by arguing that if he won he would still not be bound by the agreement to pay Protagoras, because this was not the sort of case intended in the agreement. "Rebutting" a dilemma is constructing another dilemma drawing upon the same body of facts but leading to an opposite conclusion. This is what Euathlus did, arguing that if he won the case he would be dispensed from paying by the judge's decision, and if he lost it the agreement would dispense him, so either way he was dispensed from paying. Rebuttal, however, is possible only if one of the other moves (though it may not be clear which) is also possible, for a single set of premises can lead by equally valid arguments to contradictory conclusions only if they contain some fault in themselves.

Dilemmatic reasoning obtains a categorical conclusion from hypothetical and disjunctive premises; the Port-Royalists pointed out that we may also obtain hypothetical conclusions from categorical premises. For in any categorical syllogism we may pass directly from one of the premises to the conclusion stated not categorically but conditionally on the truth of the other premise; for instance, from "Every man is mortal" we may infer that if Socrates is a man he is mortal, and from "Socrates is a man" that if every man is mortal Socrates is, and similarly with all other syllogisms. This "rule of conditionalization" is much used in certain modern logical systems.

TRADITIONAL AND MODERN LOGIC

Not only the "rule of conditionalization" but the whole subject of hypothetical and disjunctive reasoning fits more comfortably into modern than into traditional logic, being an inheritance from the Stoics, the first "modern" logicians, rather than from Aristotle. Traditionalists have often been worried at its finding any place at all in their general *corpus* and have sometimes attempted to justify it by "reducing" hypothetical and disjunctive propositions and syllogisms to "categorical" ones.

Disjunctives, to begin with, may be eliminated as a distinct form by equating "Either *p* or *q*" with the conditional "If not *p* then *q*," and the conditional form does sometimes look as if it might be a mere verbal variant of the categorical universal. This last is especially true where the conditional is introduced not by the plain "if" but by "if ever" or "if any"; "If ever a gas is heated it expands" and "If any gas is heated it expands" seem simply variants of "Every heated gas expands." But here the antecedent and consequent of the conditional are not, as J. N. Keynes put it, complete propositions with an "independent import"—"it expands" is not on its own a comprehensible sentence; the "it" refers back to the heated gas of the antecedent. Keynes suggested that the term *conditional* be used for precisely this type of "If-then" statement and the term *true hypothetical* confined to cases in which the

antecedent and consequent do have "independent import," such as "If Socrates is damned, then there is no justice in heaven." And the representation of "true hypotheticals" as categorical universals is not easy.

In modern logic, from the Stoics through some of the medieval *moderni* to the "logisticians" of our own century, "the stone which the builders rejected has been made the head of the corner." "Pure hypotheticals," together with other forms in which entire propositions are linked by various "connectives," have been made the subject of the most elementary part of logic, the propositional calculus. Aristotelian universals and particulars are built out of these forms (by means of prefixes called "quantifiers") rather than vice versa. (Details are given in the entries Logic, Modern and Russell, Bertrand, section on logic and mathematics.) The essential procedure is to read "Every *A* is a *B*" as "For every individual *x,* if *x* is an *A* then *x* is a *B*" and "Some *A* is a *B*" as "For some individual *x, x* is an *A* and *x* is a *B.*" Here, instead of a Keynesian "conditional" being explained as a categorical universal in disguise, the explanation is reversed, and the components which, as Keynes said, are "not propositions of independent import" are represented as "propositional functions" in which the place taken in a genuine proposition by an individual name is taken by a variable ("bound" by the initial quantifier "for all *x*"). But the "if" which links these components is the very same "if" which in the "pure hypotheticals" of the propositional calculus links genuine propositions. This "if" is not explained in terms of anything else (except perhaps other connectives) but is taken as fundamental.

In this way the traditional themes are not banished from modern logic but are incorporated into a much larger subject. When the Aristotelian forms are thus interpreted, however, their laws seem to require modification at some points. In particular, the A-form "For any *x,* if *x* is an *A* then *x* is a *B*" does not seem to imply the I-form "For some *x, x* is an *A* and *x* is a *B,*" for the former does not imply that any *x* in fact is an *A* (it says only that if any *x* is an *A* it is a *B*), whereas the latter does imply this (if some *x* both is an *A* and is a *B,* then that *x* is at least an *A*). This eliminates inference by subalternation and whatever else in the traditional theory depends on it, such as subaltern conversion and syllogisms, like Darapti, which require this for reduction to Figure 1.

Modern logic, however, is not at all monolithic in character, and the sketch just given is a little stylized, depicting modern logic not as a living discipline but rather as a new "tradition" that has displaced the old and against which there are already dissentient voices that give the older tradition a measure of justification (rather like that accorded to pre-Copernican astronomy by the more radical forms of relativity theory). We cannot go back to the prison that would confine all logic to the Aristotelian syllogism, but it is possible to defend (*a*) something like the view that the form "Every *X* is a *Y*" is more fundamental than either "For all *x, f*(*x*)" or "If *p* then *q*" and (*b*) the traditional ignoring (in inference by subalternation, etc.) of terms that have no application.

As to (*a*), we now know how to define both "for all *x*" and "if" in terms of a single undefined logical operator which amounts to "for all *x,* if"; for we can take as our fundamental logical complex the form "Anything such that α is such that β" and read "If *p* then *q*" as the special case of this in which α and β are "propositions with independent import," and "For all *x,* β" as the special case in which α is logically true anyway (for instance, in which it has the form "Anything such that β is such that β") and so can he ignored as a "condition" of β's truth. C. S. Peirce—at almost every point the most imaginative and flexible of the "moderns," although he died in 1914—always regarded some such reduction as possible in principle and saw the difference between the "terms" out of which categorical propositions are constructed and the "propositions" out of which we construct hypotheticals as a point of little logical importance.

Peirce, moreover, gave a highly modern justification for the traditional view that within syllogistic logic only the first figure is strictly necessary. Traditional methods of "reducing" other figures to the first do indeed involve another form of inference, namely conversion, and although this can be represented as a kind of enthymematic syllogism, it comes out as syllogism that is already in the second and third figures. For we do it by letting the term *B* be the same as *A* in the two syllogisms

No *C* is a *B* (i.e., an *A*);
Every *A* is a *B* (i.e., an *A*);
Therefore, no *A* is a *C*

(Cesare, Figure 2) and

Every *B* (i.e., *A*) is an *A*;
Some *B* (i.e., *A*) is a *C*;
Therefore, some *C* is an *A*

(Datisi, Figure 3). The replacement of *B* by *A* turns the universal affirmative premise into the logical truism "Every *A* is an *A,*" which can be dropped, and the conclusion into the converse of the remaining premise.

We can, however, derive second-figure syllogisms from first-figure ones by a variant of the reductio ad

absurdum method, employing nothing but Barbara in its terminal and propositional forms, the forms

(*a*) Every *A* is a *B,* and every *B* is a *C*; therefore, every *A* is a *C*; and

(*b*) If *p* then *q,* and if *q* then *r*; therefore, if *p* then *r,*

together with freedom to rearrange our premises and to "conditionalize" and "deconditionalize" conclusions, that is, to make such passages as that from (*a*) to, and to (*a*) from,

(*c*) Every *A* is a *B*; therefore, if every *B* is. a *C* then every *A* is a *C*

and from (*b*) to, and to (*b*) from,

(*d*) If *p* then *q*; therefore, if (if *q* then *r*) then if *p* then *r.*

As a special case of (*d*) we have

(*e*) If every *B* is a *C* then every *A* is a *C*; therefore, if (if every *A* is a *C* I am much mistaken) then if every *B* is a *C* I am much mistaken.

Forms (*c*) and (*e*) will take us from the premise to the conclusion of

(*f*) Every *A* is a *B*; therefore, if (if every *A* is a *C* I am much mistaken) then if every *B* is a *C* I am much mistaken.

But "If *X* then I am very much mistaken" just amounts to "Not *X,*" and (*f*) therefore amounts to

(*g*) Every *A* is a *B*; therefore, if not every *A* is a *C,* not every *B* is a *C,*

that is, a conditionalized form of Bocardo, Figure 3.

The equation of "Not *X*" with "If *X* then I am much mistaken" is Peirce's variant, at this point, of one account of denial. It makes it possible to present the other traditional forms as complexes of "if" and "every" (and "if" and "every," as was shown, are basically the same form of linkage), as follows:

Not every *X* is a *Y* (O) = If every *X* is a *Y* I am much mistaken.

No *X* is a *Y* (E) = Every *X* is not-a-*Y* = Every *X* is such that if it is a *Y* I am much mistaken.

Some *X* is a *Y* (I) = Not (no *X* is a *Y*) = If every *X* is such that if it is a *Y* I am much mistaken, then I am much mistaken.

Syllogisms, in all figures, involving these forms are derivable from Barbara by methods similar to that used to obtain Bocardo above, although the derivations will often be more complicated than the one given. For some of them we require Barbara in yet another form besides (*a*) and (*b*) above, namely the mixed terminal and propositional

Every *X* is a *Y*; therefore, anything such that if it is a *Y,* then *p,* is such that if it is an *X,* then *p,*

and a kind of terminal principle of *modus ponens,*

Whatever is an *X* is a thing such that if its being an *X* implies that *p,* then *p.*

Modern logic will not admit that Barbara gives us all the logic there is, but its techniques do bring out anew the extreme fecundity of this ancient form.

Turning now to the failure of certain traditional forms of inference when terms without application are employed, there have been two more recent lines of attack on the view that traditional logic is simply "wrong" in accepting such forms as "Every *X* is a *Y*; therefore, some *X* is a *Y.*" One, used by Łukasiewicz, is formalistic in character; it is a mistake, Łukasiewicz says, to interpret the traditional propositional forms in terms of modern quantification theory in the ways above indicated, or in any other ways. If we just take them as they stand, without interpretation, we can find a rigorous symbolism for them and show that the traditional laws form a self-consistent system; worries about their interpretation are extralogical. T. J. Smiley, on the other hand, thinks the interpretation of the traditional forms in quantification theory worth attempting but points out that quantification theory, as now developed, offers us wider choices of interpretation than was once thought. For quantification theory now handles cases of the form "For all *x,* *f*(*x*)" in which the range of the variable *x* is restricted to objects of some particular sort, each sort of object having its own type of variable. We need not, therefore, interpret "Every man is mortal," say, in the standard modern way as "For any individual object *x,* if that object is human it is mortal" but may read it, rather, as "For any *human* individual *m,* that human individual is mortal" (with no "ifs" about it). This interpretation, when embedded in a suitable theory of "many-sorted" quantification, will yield all the traditional results.

See also Negation.

Bibliography

Peter of Spain's *Summulae Logicales* (modern reprint, edited by I. M. Bocheński, Turin, 1947), Tractatus I–V and VII, is the best-known medieval compendium of the traditional

material. In the post-Renaissance epoch the most influential work has been the so-called Port-Royal logic, Antoine Arnauld and Pierre Nicole's *Logique, ou l'art de penser* (translated by T. S. Baynes as *The Port-Royal Logic,* Edinburgh: Sutherland and Knox, 1851). Richard Whately's crisp, homely, and pugnacious *Elements of Logic,* which appeared in successive editions in the first half of the nineteenth century, is another classic. But the most comprehensive treatment of logic along traditional lines is J. N. Keynes's *Studies and Exercises in Formal Logic* (London and New York, 1884; 4th ed., London: Macmillan, 1906).

J. S. Mill's views on the denotation and connotation of terms are developed in his *System of Logic* (London, 1843), Book I, Chs. 2, 5, and 6; his views on the uses of the syllogism are in Book II, Ch. 3. The views of C. S. Peirce are in his *Collected Papers,* edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks (8 vols., Cambridge, MA: Harvard University Press, 1931–1958), 2.455–516 and 3.154–197, and in his article "Syllogism" in the *Century Dictionary* (6 vols., New York: Century, 1889–1891).

For modern systematizations and interpretations, see Jan Łukasiewicz, *Aristotle's Syllogistic* (2nd ed., Oxford: Clarendon Press, 1957); J. C. Shepherdson, "On the Interpretation of Aristotle's Syllogistic," in *Journal of Symbolic Logic* 21 (1956): 137–147; and T. J. Smiley, "Syllogism and Quantification," in *Journal of Symbolic Logic* 27 (1962): 58–72.

A. N. Prior (1967)

LOGIC, TRANSCENDENTAL

See *Kant, Immanuel*

LOGICAL ATOMISM

See *Analysis, Philosophical; Russell, Bertrand Arthur William; Wittgenstein, Ludwig Josef Johann*

LOGICAL EMPIRICISM

See *Logical Positivism*

LOGICAL FORM

One can use sentences to present arguments, some of which are valid. Sentences are complex linguistic expressions that exhibit grammatical structure. And the grammatical properties of sentences need not be obvious. As discussed in this entry, certain arguments seem to be valid because the relevant premises and conclusions exhibit nonobvious logical structure. But this raises questions about what logical structure is and how it is related to grammatical structure.

PATTERNS OF REASONING

An ancient thought is that premises and conclusions have parts and that valid arguments exhibit valid forms, like the following: **Q** if **P**, and **P**; so **Q**. One can say that the variables (in bold) range over propositions, leaving it open for now what propositions are: sentences of some (perhaps unspoken) language, abstract states of affairs, or whatever. One can also assume that declarative sentences can be used, in contexts, to indicate or express propositions. But each sentence of English is presumably distinct from the potential premise/conclusion indicated with that sentence in a given context. Different speakers can use *I swam today* at different times to indicate various propositions, each of which could be expressed in other languages. Nonetheless, propositions seem to be sentence-like in some respects, especially with regard to being composite.

The conclusion of (1)

(1) Chris swam if Pat swam, and Pat swam; so Chris swam.

is evidently part of the first premise, which has the second premise as another part. But simple propositions, without propositional parts, also seem to have structure. Aristotelian schemata like the following are valid: Every *P* is *D*, and every *S* is a *P*; so every *S* is *D*. The italicized variables are intended to range over predicates—logical analogs of nouns, adjectives, and other classificatory terms (like *politician*, *deceitful*, and *senator*). Simple propositions appear to have subject-predicate structure; where a subject can consist of a predicate and a quantifier (indicated with a word like *every*, *some*, or *no*).

Medieval logicians explored the hypothesis that all propositions are composed of simple propositions and a few special elements, indicated with words like *or* and *only*. While they expected some differences between grammatical and propositional structure, the idea was that sentences reflect the important aspects of logical form. The medieval logicians also made great strides in reducing Aristotelian schemata to more basic inferential principles: one concerning replacement of a predicate with a less restrictive predicate, as in *Rex is a brown dog, so Rex is a dog*; and one concerning converse examples, like *Rex is not a dog, so Rex is not a brown dog*.

Nonetheless, traditional logic/grammar was inadequate. If Juliet kissed Romeo, then Juliet kissed someone. And predicates containing quantifiers were problematic.

If *respects some doctor* and *respects some senator* indicate nonrelational proposition-parts, like *is tall* and *is ugly*, then the argument indicated with (2)

> (2) Some patient respects some doctor, and every doctor is a senator; so some patient respects some senator.

has the following form, which is not valid: Some *P* is *T*, and every *D* is an *S*; so some *P* is *U*. One can introduce a variable *R* ranging over relations and offer schemata like the following: Some *P* **R** some *D*, and every *D* is an *S*; so some *P* **R** some *S*. But this is not a basic inference pattern; and such schemata do not capture the validity of inferences like the following: Every patient who met every doctor is tall, and some patient who met every doctor respects every senator; so some patient who respects every senator is tall. Relative clauses posed difficulties as well. If sentence (3) is true, so is sentence (4):

> (3) Every patient respects some doctor.

> (4) Every old patient respects some doctor.

But in (5) and (6) the direction of valid inference is reversed:

> (5) No lawyer who saw every patient respects some doctor.

> (6) No lawyer who saw every old patient respects some doctor.

FUNCTIONS AND ARGUMENTS

Gottlob Frege showed how to deal with these examples and more. But on his view, propositions have function-argument structure. Let *S* stand for the successor function. Frege interpreted the arithmetic expression *S*(3) as having a semantic value: the value of the relevant function given the relevant argument; that is, the number four. The division function can be represented as a mapping from ordered pairs of numbers to quotients: $Q(x, y) = x/y$. Functions can also be specified conditionally; consider the function that maps every even integer onto itself, and every odd integer onto its successor. On Frege's view, *Mary sang* indicates a proposition with the following structure: Sang(Mary). And he took the relevant function to be a conditional mapping from individuals in a given domain to truth values: $Sang(x) = \mathbf{t}$ if x sang, and **f** otherwise; where for each individual x, $Sang(x) = \mathbf{t}$ if and only if (iff) x sang, and $Sang(x) = \mathbf{f}$ iff x did not sing. The proposition that John admired Mary, like the proposition that Mary was admired by John, was said to have the following structure: Admired(John, Mary); where $Admired(x, y) = \mathbf{t}$ if x admired y, and **f** otherwise.

Frege's treatment of quantification departed more radically from tradition. Let F be the function indicated by *sang*, so that someone sang iff some individual x is such that $F(x) = \mathbf{t}$. Using modern notation, someone sang iff $\exists x[Sang(x)]$; where the quantifier binds the variable. Every individual in the domain sang iff F maps each individual onto **t**; in modern notation, $\exists x[Sang(x)]$. With regard to the proposition that some politician is deceitful, subject-predicate grammar suggests the division *Some politician / is deceitful*. But for Frege the logically important division is between the existential quantifier and the rest, with the quantifier binding two occurrences of its variable: $\exists x[P(x) \ \& \ D(x)]$; some individual is both a politician and deceitful. Likewise with regard to the proposition that every politician is deceitful: $\forall x[P(x)\rightarrow D(x)]$; everyone is such that if he or she is a politician then he or she is deceitful. In which case, *every politician* does not indicate a constituent of the proposition. Grammar also masks a logical difference between the existential and universal propositions: predicates are related conjunctively in the former, but conditionally in the latter.

The real power of Frege's logic is most evident in his discussion of how the proposition that every number has a successor is logically related to more basic arithmetic truths. But just consider the following analyses of (3a–6a):

> (3a) $\forall x\{P(x)\rightarrow \exists y[D(y) \ \& \ R(x,y)]\}$

> (4a) $\forall x\{[O(x) \ \& \ P(x)]\rightarrow \exists y[D(y) \ \& \ R(x,y)]\}$

> (5a) $\neg\exists x\{Lx \ \& \ \forall y[P(y)\rightarrow S(x,y)] \ \& \ \exists z[D(z) \ \& \ R(x,z)]\}$

> (6a) $\neg\exists x\{Lx \ \& \ \forall y\{[O(y) \ \& \ P(y)]\rightarrow S(x,y)\} \ \& \ \exists z[D(z) \ \& \ R(x,z)]\}$

Given Frege's rules of inference, (3a) implies (4a), while (5a) follows from (6a). Frege concluded that natural language is not suited to the task of representing propositions perspicuously. On his view, premises/conclusions have function-argument structure, which is often masked in natural language. But one can try to invent languages whose sentences depict true propositional structure.

Frege originally took propositional constituents to be the relevant functions and (ordered n-tuples of) entities that such functions map to truth-values. But he later refined this view, taking the sense of an expression to be a mode of presentation of the corresponding semantic value. Frege identified propositions—or what he called thoughts (*Gedanken*)—with senses of sentences in an ideal language, which allowed him to distinguish the

proposition that Hesperus is bright from the proposition that Phosphorus is bright. Thus, Frege could deny that the inference *Hesperus is Hesperus, so Hesperus is Phosphorus* is an instance of the valid form *P, so P*.

DESCRIPTIONS AND MISMATCH

One might think that the logical form of any proposition indicated with *The boy from Canada sang* is *Sang(b)*, where *b* stands for the individual in question. But this makes elements of the description logically irrelevant. And if the boy from Canada sang, then a boy sang. Moreover, *the* implies uniqueness (at least within a context). So Bertrand Russell (1919) held that a proposition expressed with *The boy sang* has the following structure: $\exists x\{Boy(x) \& \forall y[Boy(y) \rightarrow y = x] \& Sang(x)\}$; where the middle conjunct is one way, among many, of expressing uniqueness. According to Russell, even if a speaker refers to a certain boy when saying *The boy sang*, that boy is not a constituent of the indicated proposition—which has the form of an existential quantification, as opposed to a function saturated by the boy. In this respect, *the boy* is like *some boy*. Though on Russell's view, not even *the* indicates a propositional constituent. This extended Frege's idea that natural language is misleading, while letting Russell account for the meaningfulness of descriptions that describe nothing.

Let **Frank** be the proposition indicated (now) with *The (present) king of France is bald*. If **Frank** consists of some function saturated by an entity indicated with *The king of France*, there must *be* such an entity. But instead of appealing to nonexistent kings, or ways of presenting them, Russell held that **Frank** is of the form $\exists x\{K(x) \& \forall y[K(y) \rightarrow y = x] \& B(x)\}$. In which case, the true negation of **Frank** is not of the form $\exists x\{K(x) \& \forall y[K(y) \rightarrow y = x] \& \neg B(x)\}$. This invited the thought, developed by Ludwig Wittgenstein (1922, 1953) and others, that many philosophical puzzles might dissolve if one properly understood the logical forms of one's claims. Russell also held that one bears a special relation to constituents of propositions one can entertain and that one typically does not bear this relation to the individuals one refers to with names. This led Russell to say that names are disguised descriptions. On this view, *Hesperus* is associated with a complex predicate—say, for illustration, of the form $E(x) \& S(x)$. Then *Hesperus is bright* indicates a proposition of the form $\exists x\{[E(x) \& S(x)] \& \forall y\{[E(y) \& S(y)] \rightarrow y = x]\} \& B(x)\}$. It follows that Hesperus exists iff $\exists x[E(x) \& S(x)]$; and this was challenged by Saul Kripke (1980). But Russell could say that "Phosphorus is bright" indicates a proposition of the form $\exists x\{[M(x) \& S(x)] \& \forall y\{[M(y) \& S(y)] \rightarrow y = x]\} \& B(x)$; where $E(x)$ and $M(x)$ indicate different functions, specified in terms of evenings and mornings, leaving room to discover that $E(x) \& S(x)$ and $M(x) \& S(x)$ both indicate functions that map Venus alone to the truth-value **t**.

Positing unexpected logical forms thus had payoffs. But if mismatches between sentential and propositional structure are severe, one wonders how one manages to indicate propositions. This worry was exacerbated by increasing suspicion that talk of propositions is (at best) a way of talking about how one should regiment one's verbal behavior for purposes of scientific inquiry and that one should regiment natural language in first-order predicate calculus. From this perspective, associated with Willard Van Orman Quine (1950), mismatches between logical and grammatical form are to be expected. Another strand of thought, inspired by Wittgenstein's later work, also suggested that a single sentence could be used (on different occasions) to express different kinds of propositions. Peter Strawson (1950) argued, contra Russell, that a speaker could use an instance of *The F is G* to express a singular proposition about the F in the context at hand. Keith Donnellan (1966) contended that a speaker could even use an instance of *The F is G* to express a singular proposition about an individual that is not an F. Various considerations suggested that relations between spoken sentences and propositions are at best very complex and mediated by speakers' intentions.

With hindsight, though, one can see that the divergence between logical and grammatical form was exaggerated. Consider again the proposed regimentation of the proposition indicated with *Some boy sang*: $\exists x[Boy(x) \& Sang(x)]$. With restricted quantifiers, one can offer another logical paraphrase that parallels the grammatical division between *some boy* and *sang*. Let $\exists x{:}Boy(x)$ be an existential quantifier that binds a variable ranging over boys in the domain. Then $\exists x{:}Boy(x)[Sang(x)]$ means that for some individual x such that x is a boy, x sang. Likewise, $\forall x{:}[Tall(x) \& Boy(x)]\{Sang(x)\}$ is logically equivalent to $\forall x\{[Tall(x) \& Boy(x)] \rightarrow Sang(x)\}$. And $\exists x{:}[Boy(x) \& \forall y{:}Boy(y)[x = y]]\{Sang(x)\}$ means that for some boy x such that x is identical with every boy, x sang. Richard Montague (1974) offered a similar rewrite of Russell's hypothesis about the logical form of *The boy sang*. On this view, *The boy* corresponds to a propositional constituent, even though the boy referred to (if such there be) does not.

Still, the subject-predicate structure of *Mary trusts every doctor* diverges from the function-argument structure of $\forall y{:}Doctor(y)[Trusts(Mary, y)]$. Grammatically,

trusts and *every doctor* form a phrase; though logically, *trusts* combines with *Mary* and a variable to form a complex predicate that in turn combines with a restricted quantifier. Given Montague's (1974) techniques, one can provide algorithms that systematically associate quantificational sentences of natural language (described in subject-predicate terms) with Fregean propositional structures. But it seemed that mismatches between grammatical and logical form remained, at least in cases of complex predicates with quantificational constituents.

TRANSFORMATIONAL GRAMMAR AND LF

One must not, however, assume a naive conception of grammar when thinking about its relation to logic. For example, the grammatical form of a sentence need not be determined by the order of its words. Using brackets to indicate phrasal structure, one can distinguish sentence (7) from the homophonous sentence (8).

(7) {Mary [saw [the [boy [with binoculars]]]]}

(8) {Mary [[saw [the boy]] [with binoculars]]}

The direct object of (7) is *the boy with binoculars*, while in (8), *saw the boy* is modified by an adverbial phrase. And a leading idea of modern linguistics is that many grammatical structures are transformations of others.

Expressions often appear to be displaced from positions canonically associated with certain grammatical relations. In (9), *who* seems to be associated with the direct-object position of *saw*.

(9) Mary wondered who John saw

And (9) can be glossed as *Mary wondered which person is such that John saw him*. This invites the hypothesis that the structure of (9) is as shown in (9-SS), reflecting a transformation of the simpler expression shown in (9-DS):

(9-SS) {Mary [wondered [who$_i$ {John [saw (_)$_i$]}]]}

(9-DS) {Mary [wondered {John [saw who]}]}

where coindexing indicates a grammatical relation between the coindexed positions. The idea was that each sentence has a surface structure and a deep structure and that the former will differ from the latter when expressions like *who* are displaced as in (9). As an illustration of the kind of data relevant to such hypotheses about grammar, note that (10–12) are perfectly fine sentences, while (13) is not:

(10) The boy who sang was happy

(11) Was the boy who sang happy

(12) The boy who was happy sang

(13) Was the boy who happy sang

The ill-formedness of (13) is striking, since one can ask whether or not the boy who was happy sang. This suggests that (11-SS) is the result of a permissible transformation, but (13-SS) is not:

(11-SS) Was$_i$ {[the [boy [who sang]]] [(_)$_i$ happy]}

(13-DS) Was$_i$ {[the [boy [who [(_)$_i$ happy]]]] sang}

As transformational grammars were elaborated, many linguists posited another level of grammatical structure—LF, intimating logical form—obtained by displacing quantificational expressions. In particular, it was proposed that structures like (14-SS) were transformed, as in (14-LF):

(14-SS) {Pat [trusts [every doctor]]}

(14-LF) {[every doctor]$_i$ {Pat [trusts (_)$_i$]}}

Clearly, (15-LF) does not reflect the pronounced word order in English. But there is independent evidence for covert (inaudible) quantifier-raising in natural language. The suggestion was that each sentence has a PF (intimating phonological form) that determines pronunciation, and an LF that determines interpretation. On this view, the scope of a quantifier must be determined at LF, as in (14-LF). And one can say this, while also saying that the pronunciation of *Pat trusts every doctor* reflects the untransformed surface structure (14-SS). Many apparent examples of grammar-logic mismatches were thus rediagnosed as mismatches between different aspects of grammatical structure. This preserves the idea that surface appearances are often misleading with regard to propositional structure. But it also suggests that grammatical form and logical form converge, once one moves beyond traditional subject-predicate conceptions of structure with regard to both logic and grammar. And further simplification may be possible.

Given a conception of grammar according to which each sentence has a PF and an LF, perhaps involving different transformations, it is not obvious that one needs to posit other levels of grammatical analysis. Each expression of a natural language may just be a PF-LF pair that can be generated in accordance with certain constraints on how expressions can be combined and transformed. One can hypothesize that a sentence like (9) is formed in stages, including stages like those depicted in (9-DS) and (9-SS), without saying that any one stage is special in ways that deep structure and surface structure were said

to be. On this view, (10–12) correspond to natural ways of associating a PF with an LF, but the string of words in (13) does not. From this perspective, urged by Noam Chomsky and others, talk of PFs and LFs need not be understood in terms of interlevel transformations (Chomsky 1995, Hornstein 1995). Rather, PFs and LFs can be viewed simply as generable linguistic structures that reflect pronunciation and meaning. In which case questions about grammatical form and linguistic meaning are largely questions about LFs.

Nonetheless, there is still an important conceptual distinction between the linguist's notion of LF and the logician's notion of logical form. The LF of a sentence may, in various ways, underdetermine the structure of the proposition a speaker expresses with that sentence (in a given context). The LF may, however, provide a scaffolding that can be elaborated in particular contexts, with little or no mismatch between basic sentential and propositional structure. These issues remain unsettled. But discoveries of rich grammatical structure reinvigorated the idea that natural languages are semantically compositional.

Prima facie, *Every tall sailor respects some doctor* and *Some short boy likes every politician* exhibit common modes of linguistic combination. So a natural hypothesis is that the meaning of each sentence is somehow fixed by these modes of combination, given the word meanings. Inspired by Alfred Tarski's development of Frege in 1956, Donald Davidson (1967) conjectured that there are recursively specifiable theories of truth for natural languages. And while there are many apparent objections, the conjecture has been fruitful. This raises the possibility that talk of logical forms should be construed in terms of the structure(s) that speakers impose on words to understand natural language systematically. From this tendentious perspective, the phenomenon of valid inference would be largely a reflection of semantic compositionality.

At this point, many issues become germane. Given any sentence of natural language, one can ask interesting questions about its grammatical structure and what it can be used to say. (Modal claims and propositional attitude reports have been studied intensively.) It is not obvious how one should characterize meanings or logical relations. (Are theories of meaning theories of truth? Which valid inferences, if any, cannot be captured in first-order terms?) The role of context is large and ill understood. But it seems clear that the traditional questions—what kinds of structures do propositions and sentences exhibit, and how do thinkers who also speak relate these structures—must be addressed in terms of increasingly sophisticated conceptions of logic and grammar.

See also Events in Semantic Theory; Modality and Quantification; Semantics; Syntax.

Bibliography

Beaney, Michael, ed. *The Frege Reader*. Oxford, U.K.: Blackwell, 1997.

Boolos, George. *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press, 1998.

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.

Chomsky, Noam. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger, 1986.

Chomsky, Noam. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

Davidson, Donald. "Truth and Meaning." *Synthese* 17 (1967): 304–323.

Donnellan, Keith. "Reference and Definite Descriptions." *Philosophical Review* 75 (1966): 281–304.

Fodor, Jerry. "Propositional Attitudes." *The Monist* 61 (1975): 501–523.

Grice, H. Paul. "Logic and Conversation." In *Syntax and Semantics*. Vol. 3, edited by P. Cole and J. Morgan. New York: Academic Press, 1975.

Hornstein, Norbert. *Logical Form: From GB to Minimalism*. Oxford, U.K.: Blackwell, 1995.

Kneale, William, and Martha Kneale. *The Development of Logic* (1962). New York: Oxford University Press, 1984.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

May, Robert. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press, 1985.

Montague, Richard. *Formal Philosophy*. New Haven, CT: Yale University Press, 1974.

Preyer, Gerhard, and Georg Peters, eds. *Logical Form and Language*. New York: Oxford University Press, 2002.

Quine, W. V. *Methods of Logic*. New York: Henry Holt, 1950.

Russell, Bertrand. *Introduction to Mathematical Philosophy*. London: George Allen and Unwin, 1919.

Sainsbury, Mark. *Logical Forms*. Oxford, U.K.: Blackwell, 1991.

Strawson, Peter. "On Referring." *Mind* 59 (1950): 320–344.

Tarski, Alfred. *Logic, Semantics, Metamathematics*. 2nd ed. Translated by J. H. Woodger; edited by John Corcoran. Indianapolis, IN: Hackett, 1983.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. New York: Macmillan, 1953.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. Translated by D. Pears and B. McGuinness. London: Routledge and Kegan Paul, 1922.

Zalta, Edward. "Frege (Logic, Theorem, and Foundations for Arithmetic)." *The Stanford Encyclopedia of Philosophy*. Stanford, CA: Metaphysics Research Lab, Stanford University, 2003. Available at http://plato.stanford.edu/archives/fa112003/entries/frege-logic.

Paul M. Pietroski (2005)

LOGICAL KNOWLEDGE

"Logical knowledge" can be understood in two ways: as knowledge of the laws of logic and as knowledge derived by means of deductive reasoning. Most of the following is concerned with the first of these interpretations; the second will be treated briefly at the end. Furthermore, only deductive logic will be treated: As yet, there is no set of laws of inductive logic enjoying the kind of consensus acceptance accorded to deductive logic.

To begin with, we must specify what is a law of logic—not an entirely straightforward task. There are three, not all mutually exclusive, conceptions of logic laws. First, one could take them to be valid schemata (of statements), such as the familiar law of excluded middle, "*p* or not *p*". A second conception is that they are valid rules of inference, such as the familiar *modus ponens*—that is, from "$p \rightarrow q$" and *p* infer *q*. The third conception of logic law, due to Gottlob Frege and Bertrand Russell, takes them to be maximally general, true (not valid) second-order quantified statements (see Goldfarb, 1979). The following discussion is confined, by and large, to the second conception; but the philosophical problems canvassed arise with respect to the other conceptions as well.

In order to appreciate the problems involved in the analysis of knowledge of logical laws, note first that, however these laws are conceived, knowledge of them appears to be propositional. That is, to know a law of logic is to know that a rule of inference (or a schema) is valid (or a statement true). But, given the classical analysis of knowledge as justified true belief, it follows that knowledge of the validity of a rule of inference requires justification. There are two uncontroversially entrenched forms of justification: inductive and deductive justification. By the nature of inductive reasoning an inductive justification of validity shows, at best, that a rule of inference usually leads from true premises to a true conclusion (or that it is sufficiently highly likely to do so). This is too weak; a valid rule of inference, as noted above, necessarily leads from true premises to true conclusions. So it appears that the justification of validity must be deductive.

On the basis of this conclusion it can be shown that the justification of the validity of any rule of inference either is circular or involves an infinite regress. The argument has two parts. To begin with, there certainly are deductive justifications of rules of inference that raise no serious philosophical questions. Take the justification of the rule "existential specification" in Benson Mates's widely used *Elementary Logic*: "To justify this rule,… we observe that … we may … obtain the inference it permits [using certain basic rules] … Assuming … that the basic rules … are [valid], … the above description of how any [existential specification] inference can be made using only [those] rules … shows that [existential specification] is [valid], too" (Mates, 1972, p. 123). The rule is justified by explicitly assuming the validity of other rules, so the justification here is only relative. If all logical laws are justified in this way, then, plausibly, the justification of any given rule will be either circular, by explicitly assuming its own validity, or will involve an infinite regress.

One might conclude from this that there must be some set of rules that are not justified on the basis of the assumed validity of other rules. Let us call these rules fundamental. Unfortunately, there is a simple argument that the justification of fundamental rules will involve a similar circularity or infinite regress.

What counts as a deductive justification of a proposition depends on what forms of inference are taken to be valid. For, if any rule of inference used in an argument is invalid, then the argument could not constitute a deductive justification of anything. Let us formulate this point as: A deductive argument presupposes the validity of the rules of inference it employs. Given this formulation, we can state an intuitive principle: If an argument for the validity of a rule of inference presupposes the validity of that very rule, then the argument is circular. To distinguish this notion of circularity from the one used above, let us call this pragmatic circularity, and the former, direct circularity.

Suppose a fundamental rule of ρ is justified by an argument π. Now either π employs nonfundamental rules, or it does not. Suppose π employs a nonfundamental rule σ. By the first part of the argument, σ is justified by assuming the validity of fundamental rules. Again, either the justification of σ assumes the validity of ρ or it does not. Now assume further that if an argument employs a rule whose justification assumes the validity of another, then it presupposes the validity of the second. Thus, in the first case, the justification of ρ is pragmatically circular. In the second case, the justification of ρ presupposes the validity of a set of other fundamental rules.

Now suppose that π does not employ nonfundamental rules. Then, either it employs ρ or it does not. In the first case the justification is pragmatically circular. In the second, again, the justification of ρ presupposes the validity of a set of other fundamental rules. Hence, the justification of any fundamental rule either is pragmatically circular or involves an infinite regress. (See Goodman 1983, pp. 63–64; see also Bickenbach 1978, Dummett 1973, and Haack 1976.)

One might object to the notion of circularity of argument used in the second part of the argument. Unlike the more familiar variant of circularity, the conclusion in this case is not actually assumed as a premise but is presupposed by the inferential transitions. Thus, it is unclear that this sort of circular argument suffers from the principal difficulty afflicting the more familiar sort of circular argument, namely, that every conclusion is justifiable by its means.

This, however, is not a very strong objection. One might reply, to begin with, that pragmatically circular arguments are just as objectionable as directly circular ones in that both assume that the conclusion is not in question, by assuming its truth in the one case and by acting as if it were true in the other. Moreover, while it is unclear that every rule of inference is justifiable by a pragmatically circular argument, it is clear that such an argument can justify both rules that we take to be valid and rules that we take to be fallacies of reasoning. For example, the following is an argument demonstrating the validity of the fallacy of affirming the consequent (see Haack, 1976):

1. Suppose "$p \rightarrow q$" is true.
2. Suppose q is true.
3. By the truth table for "→," if p is true and "$p \rightarrow q$" is true, then q is true.
4. By (2) and (3), p is true and "$p \rightarrow q$" is true.
5. Hence, p is true.

Second, one might accept that deductive justification is not appropriate for fundamental logical laws but conclude that there is another kind of justification, neither deductive nor inductive, for these laws. There have been two proposals about a third kind of justification.

One proposal, due to Herbert Feigl (1963), claims that fundamental logical laws require pragmatic, instrumental justification. An immediate difficulty is, What counts as a pragmatic justification of a logical law? Surely, if there is anything that a rule of inference is supposed to do for us, it is to enable us to derive true conclusions from true premises. So, it looks as if to justify a logical law pragmatically is to show that it is suited for this purpose. And that seems to require showing that it is valid. Feigl is aware of this problem and argues that, in the context of a pragmatic justification, circularity is not a problem, since all that such a justification is required to do is provide a recommendation in favor of doing things in some particular way, not a proof that this way necessarily works. It is not clear, however, that this constitutes a compelling response to the philosophical problem of justifying deduction, since, far from needing a letter of reference before employing deductive reasoning, its use is inescapable.

Another proposal for a third kind of justification is due to J. E. Bickenbach (1978), who argues that rules of inference are justified because they "fit with" specific instances of arguments that we accept as valid; for this reason he calls this kind of justification "instantial." The problem with this approach is that, in the case of rules of inference having some claim to being fundamental, such as *modus ponens,* it is plausible that we take the validity of the rule to be conceptually prior to the validity of any instance of it. For example, in the case of *modus ponens,* where there appear to be counterinstances to the rule, such as the sorites paradox, we take the problem to lie not in *modus ponens* but in vague concepts. Hence, whatever force "instantial" justification has, it seems incapable of conferring on fundamental rules of inference the kind of conceptual status we take them to have.

One might simply accept the conclusion of the argument, that fundamental logical laws cannot be justified, as indicating the philosophical status of these laws: They are simply constitutive rules of our practice of deductive justification. That is, there is no such thing as deductive justification that fails to conform to these rules, just as there is no such thing as the game of chess in which the queen is allowed to move in the same way as the knight. This third response leads to at least two philosophical questions: (1) How do we identify the fundamental laws of logic? (2) Is there such a thing as criticism or justification, as opposed to mere acceptance of a deductive practice?

A natural way to answer the first question is to take the fundamental rules to be determined by the meanings of the logical constants. This answer has been developed in some detail by Dag Prawitz (1977) and Michael Dummett (1991). Following Gerhard Gentzen (1969), they take the natural deduction introduction and elimination rules for a logical constant to be determined by the meaning of that constant. (More detail on the answer is provided in the final paragraph of this article.) Part of an answer to the second question has been provided by A. N. Prior (1967) and Nuel Belnap (1961), who showed that there exist sets of rules of inference that we can recognize as internally incoherent.

This third response has the consequence that our relation to the fundamental laws of logic is not one of knowledge classically construed and, hence, is different from our relation to other laws, such as the laws of physics, or of a country.

We turn now to the notion of knowledge derived from deductive reasoning. The question this notion raises, first studied by J. S. Mill (1950, bk. 2, chap. 3), is to explain how deductive reasoning could be simultaneously necessary and informative. It is undeniable that we can understand the premises and the conclusion of an argument without knowing that the former implies that latter; this is what makes it possible for us to gain information by means of deductive reasoning. This fact does not by itself conflict with the necessity of deductive implication, since there is no conflict between the existence of something and our lack of knowledge thereof. But, a problem can arise if the explanation of the necessity of deductive implication entails constraints on the notion of understanding. The following are two ways in which the problem of deduction arises.

First, consider Robert Stalnaker's (1987) analysis of the notions of proposition and of understanding. The proposition expressed by a statement is a set of the possible worlds, the set of those worlds in which the proposition is true. To understand a statement is to know the proposition it expresses; hence, to understand a statement is to know which possible worlds are those in which the proposition it expresses is true. These claims have two consequences: First, that all necessary statements, and hence all deductive valid statements, express the same proposition, namely, the set of all possible worlds; second, to understand any necessary statement is to know that the proposition it expresses is the set of all possible worlds. From these consequences it would seem to follow that in virtue of understanding any valid statement, one would know that it is necessarily true. It seems plausible that if one understands the premises and the conclusion of a valid argument, then one must also understand the conditional whose antecedent is the conjunction of the premises and whose consequent is the conclusion. But if the argument is valid, so is this conditional. Hence, if an argument is valid, then anyone who understood its premises and conclusion would know that this conditional expressed a necessary truth. It is now plausible to conclude that one can know whether an argument is valid merely on the basis of understanding its premises and conclusion by knowing whether the corresponding conditional expressed a necessary truth.

Next, consider Dummett's (1973, 1991) analysis of deductive implication. According to this analysis, deductive implication is based on the meanings of the logical constants. Thus, for example, the fact that *p* and *q* imply "*p* and *q*" is explained by the fact that the meaning of "and" is such that the truth condition of "*p* and *q*" is satisfied just in case those of *p* and of *q* are. Similarly, the meaning of the existential quantifier is such that if the truth condition of "*a* is *F*" is satisfied, then so must the truth condition of "There is an *F*". Thus, corresponding to each logical constant, there is an account of the truth conditions of logically complex statements in which that constant occurs as the principal connective, in terms of the truth conditions of its substatements. This account explains the validity of rules of inference to those statements from their substatements and hence determines the set of fundamental rules, rules whose validity must be acknowledged by anyone who understands the meanings of the logical constants. But there are, as we have seen, cases in which we can understand the premises and the conclusion of an argument without knowing that the former implies the latter. So, how is deductive implication to be explained in those cases? This question is easy to answer if all the inferential transitions in these arguments are instances of fundamental rules determined by the senses of the constants. But the fact is otherwise; we acknowledge a number of rules of inference that are not reducible to fundamental rules. The problem is thus not an epistemological one; it arises because our conception of deductive implication includes rules whose necessity is not explainable on the basis of our understanding of the logical constants.

See also A Priori and A Posteriori; Dummett, Michael Anthony Eardley; Frege, Gottlob; Induction; Mill, John Stuart; Prior, Arthur Norman; Russell, Bertrand Arthur William.

Bibliography

Belnap, N. "Tonk, Plonk, and Plink." In *Philosophical Logic,* edited by P. F. Strawson. Oxford: Oxford University Press, 1967.

Bickenbach, J. E. "Justifying Deduction." *Dialogue* 17 (1979): 500–516.

Dummett, M. A. E. "The Justification of Deduction." In *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978.

Dummett, M. A. E. *The Logical Basis of Metaphysics.* Cambridge, MA: Harvard University Press, 1991.

Feigl, H. "De Principiis Non Disputandum…?" In *Philosophical Analysis,* edited by M. Black. Englewood Cliffs, NJ: Prentice-Hall, 1963.

Gentzen, G. "Investigations into Logical Deduction." In *Collected Papers,* edited by M. E. Szabo. Amsterdam: North-Holland, 1969.

Goldfarb, W. D. "Logic in the Twenties." *Journal of Symbolic Logic* 79 (1979): 1237–1252.

Goodman, N. *Fact, Fiction, and Forecast.* Cambridge, MA: Harvard University Press, 1983.

Haack, S. "The Justification of Deduction." *Mind* 85 (1976): 112–119.

Mates, B. *Elementary Logic.* 2nd ed. New York: Oxford University Press, 1972.

Mill, J. S. *A System of Logic.* In *Philosophy of Scientific Method,* edited by E. Nagel. New York: Hafner, 1950.

Prawitz, D. "Meaning and Proof: On the Conflict between Classical and Intuitionistic Logic." *Theoria* 48 (1977): 2–4.

Prior, A. N. "The Runabout Inference Ticket." In *Philosophical Logic,* edited by P. F. Strawson. Oxford: Oxford University Press, 1967.

Stalnaker, R. *Inquiry.* Cambridge, MA, 1987.

Strawson, P. F., ed. *Philosophical Logic.* Oxford: Oxford University Press, 1967.

Sanford Shieh (1996)

LOGICAL PARADOXES

A paradox is an argument that derives or appears to derive an absurd conclusion by rigorous deduction from obviously true premises. Perhaps the most famous is Zeno's paradox of the runner, who, before she can reach her destination, first has to reach the point halfway there, and who, before reaching the halfway point, has to reach the quarter point, before which she must reach the point one-eighth of the way to the destination, and so on. The conclusion is that no runner ever reaches her goal, or even gets started.

To contemporary ears the argument does not sound so irresistible, since we can attribute its appeal either to an ambiguity in the use of "never" ("at no point in time" *versus* "at no point in the sequence") or to a dubious hidden premise that it is impossible to perform infinitely many tasks in a finite time, perhaps because there is a positive minimum to the length of time each task requires. To the ancients, however, the paradox was deeply disturbing. The most influential response was that of Aristotle, who concluded that it was not possible to partition the runner's path into infinitely many parts. Any segment of the runner's course can be divided in two, so that there is no finite bound on how many pieces the path contains, but the process of partitioning the path never concludes in a path with infinitely many parts. The number of segments that make up the path is said to be *potentially infinite*. The moral Aristotle drew from Zeno is that there is, in nature or in mathematics, no *actual infinite*. "Potentially infinite" is not like "potentially hot." When we say that a poker is potentially hot, we mean that, at some time and circumstance, it could be actually hot, whereas when we say that a line is potentially infinite, we mean that it can always be made longer but not that there is any time at which it is actually infinite.

Aristotle's doctrine commanded wide adherence among philosophers and mathematicians, but toward the end of the nineteenth century it came widely to be seen as too restrictive. New mathematics embraced not only infinitely long lines, but also an analysis of a line as made up of infinitely many points, as well as infinite sets, infinite numbers, and infinite-dimensional geometry.

The new mathematics brought a spate of new paradoxes, which, in their formal structure, resemble the semantic paradoxes, the first of which appeared in the sixth century BCE when Epimenides, himself a Cretan, declared that Cretans always lie. Provided Epimenides' neighbors are sufficiently mendacious, we are driven to the conclusion that, if his statement is true, it is false, and if false, true. Deep problems, or perhaps a single deep problem in different manifestations, afflict the foundations of both mathematics and linguistics.

COUNTING BEYOND THE FINITE

Broadly speaking there were two reasons for repudiating of Aristotle's prohibition of the actual infinite. First as mathematics became vastly more general, finitistic techniques came to be seen as confining. The ancient Greeks had a marvelously sophisticated theory of polygons and conic sections, but a fully general theory of shapes requires such techniques as approximating an unruly curve by an infinite sequence of curves that are better behaved.

The second reason was the so-called arithmetization of geometry, brought about by the investigation of alternatives to Euclid's axiom that, given a line and a point not on the line, there is on their plane exactly one line through the point that never intersects the given line, no matter how far the two lines are extended. Once alternatives to Euclidean geometry emerged, one could no longer be fully confident that Euclid's axioms correctly described the world around us (and, indeed, these suspicions are confirmed by the general theory of relativity). The theory of real numbers remained at the center of modern mathematics, but since one could no longer identify the positive real numbers as the ratios of lengths of physical line segments, one was no longer sure what the theory referred to.

A strategy for answering this question can be found in William Hamilton's treatment of the complex numbers. Extending the real number system by introducing a fictitious solution to the equation "$x^2 + 1 = 0$" proves

enormously useful mathematically, but one cannot help fretting that addressing algebraic problems with make-believe solutions is more an exercise in wishful thinking than legitimate science. Hamilton proposed to soothe this consternation by taking "complex number" to refer not to new and ontologically dubious entities but to familiar mathematical objects thought of in a new way. Namely, we treat a complex number as an ordered pair of ordinary real numbers, with appropriate operations.

Hamilton's construction presupposed the real numbers, but we can apply the same technique to secure the real numbers on a firmer basis. The starting point is easy enough. We can identify the positive rational numbers as ordered pairs of relatively prime positive integers, but the rationals have gaps—$\sqrt{2}$, for instance—and the modern theory of continuity and limits requires a number system without gaps. More precisely we want assurance that the *least upper bound principle*, according to which every nonempty set of real numbers that is bounded above has a least upper bound, is satisfied. Richard Dedekind solved this problem by identifying the real number with pairs $\langle A,B\rangle$ that partition the rationals into nonempty, nonoverlapping sets with every member of A less than every member of B.

Dedekind's construction succeeded in securing the real numbers on a foundation that did not presuppose the truth of Euclidean geometry, but it required the unapologetic acceptance of infinite sets. It permitted real analysis to be seen as built upon a foundation in the theory of sets, and indeed set theory is widely perceived as providing a uniform foundation for all of mathematics.

The elevation of set theory to its central role received its greatest impetus from the work of Georg Cantor (1895, 1897) who extended elementary-school arithmetic so that infinite as well as finite sets could be counted. Doing so required him to confront Galileo's paradox. Two sets have the same number of elements if there is a one-one correspondence by which each member of one set is paired off with one and only one member of the other. It follows that there are just as many perfect squares as there are nonnegative integers, since we can pair off n with n^2. But it seems obvious that there are more nonnegative integers than there are squares. A lot more, in fact, since as N grows the proportion of perfect squares among the first N integers becomes vanishingly small. The moral Galileo drew from this is that the notions of more and fewer cannot be applied to the infinite.

Overcoming Galileo's paradox was largely a matter of raw intellectual courage. Cantor had to resolve to follow the computations where they led, no matter how strongly the results he obtained contravened the intuitions obtained from grade-school experience with finite numbers. Stipulating that the cardinal number of S is equal to the cardinal number of T (in symbols, $\#(S) = \#(T)$) if and only if S and T can be put in one-one correspondence and that $\#(S) \leq \#(T)$ if and only if S has the same cardinal number as a subset of T, we find that the familiar laws of order carry over directly, with one glaring exception. As Galileo's paradox illustrates, you can have $\#(S) \leq \#(T)$ even though T is a proper subset of S. Defining, for S and T disjoint, $\#(S) + \#(T) = \#(S \cup T)$, we find that the familiar laws of addition are largely upheld, but that particular computations yield wildly unexpected results. If we let $\aleph_0$ be the number of natural numbers, we find that $\aleph_0 + \aleph_0 = \aleph_0$ Similarly if we define $\#(S) \cdot \#(T)$ to be the cardinal number of the set of ordered pairs $\langle s,t\rangle$ with $s \in S$ and $t \in T$, we find $\aleph_0 \cdot \aleph_0 = \aleph_0$. In fact for any infinite numbers κ and λ, we have $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa,\lambda)$.

It is starting to look as if infinite arithmetic is remarkably easy: Whatever the question, the answer is "$\aleph_0$." This happy impression is dispelled when we turn to infinite exponentiation. Defining $\#(S)^{\#(T)}$ to be the number of functions from T to S, we find that $2^{\#(T)}$, which is the cardinal number of the *power set* of T (the set $\wp(T)$ of subsets of T), is strictly greater than $\#(T)$. That is, $2^{\#(T)} \geq (T)$ and $2^{\#(T)} \neq \#(T)$. That $\#(\wp(T)) \geq \#(T)$ is easy; use the function that takes an element x of T to $\{x\}$. To see that $\#(T) \neq \#(\wp(T))$, let F be a function from $\wp(T)$ to T. We want to see that F is not one-one, that is, that there exist distinct subsets U and V of T with $F(U) = F(V)$, which we do by assuming F were one-one and deriving a contradiction. Define a binary relation E on T by stipulating that xEy if and only if x is an element of some set W with $y = F(W)$. Then for any element x of T and any subset V of T, we have $xEF(V)$ if and only if Vx. (Why? If $xEF(V)$, then there is a W with $F(V) = F(W)$ and Wx; because F is one-one, we must have $V = W$, hence Vx. Conversely if Vx, then we can find our set W with $F(V) = F(W)$ and Wx by setting W equal to V.) In particular if we let R be the set of all elements of T that do not bear the relation E to themselves, we have, for any x, $xEF(R)$ if and only if Rx, which happens if and only if not xEx. Taking x equal to $F(R)$ reveals a contradiction.

In particular, the number of real numbers is $2^{\aleph_0}$ so that there are more real numbers than there are natural numbers. To see that the real numbers are equinumerous with the numbers in the interval from 0 to 1, use the function that takes x to $\frac{1}{2}(\frac{x}{|x|+1}+1)$. The proof that the real numbers between 0 and 1 are equinumerous with the sets of natural numbers uses the function that takes a real

number to the set of places in its binary decimal expansion where 1s appear, although one has to tinker a bit to make allowance for numbers like ⅝, which has two different binary decimal expansions, 0.10100000000 … and 0.10011111111 … .

There are two fundamental ways we apply the counting numbers: To measure the size of a set, and to mark positions in a queue. For the first purpose we employ the English nouns, "one," "two," "three," and so on, whereas for the second we use the adjectives "first," "second," "third." … To generalize the first concept into the infinite Cantor developed his theory of infinite cardinal numbers, and for the second he introduced a theory of infinite so-called *ordinal numbers*. First some definitions. A binary relation *L* on a set *S* is an *ordering* if it meets the following three conditions, for any *x*, *y*, and *z* in *S*: If *xLy* and *yLz*, then *xLz*; not *xLx*; and either *xLy*, *yLx*, or $x = y$. The usual way of ordering the real numbers is an ordering, but it doesn't distinguish a first, second, and third real number. In order for the members of an ordered set to be counted by ordinal numbers, a further condition is required: A binary relation *L* on a set *S* is *well-founded* just in case every nonempty subset *R* of *S* has an *L*-least element, an element *x* of *R* such that there in no element *y* of *R* with *yLx*; equivalently there is no infinite sequence $s_0, s_1, s_2, s_3, \ldots$ with $s_{n+1}Ls_n$. A well-founded ordering is a *well-ordering*. Cantor's second great innovation was to extend the notion of ordinal number to infinite well-orderings.

If L well-orders a set *S* and *M* well-orders *T*, an *order isomorphism* from *L* to *M* is a one-one correspondence that preserves the order relation, so that, for any *x* and *y* in *S*, we have: *xLy* if and only if $f(x)Mf(y)$. Two well-orderings have the same ordinal number if and only if they are order isomorphic. If α is the ordinal number of an ordering *L* on *S* and β is the ordinal number of an ordering *M* on *T*, we say that $\alpha \leq \beta$ if and only if *L* is order-isomorphic to an initial segment of *M*. This provides a well-ordering of the ordinals, which supplies for each ordinal α a well-ordering of the ordinals less than α; its ordinal number is α. If *L* is a well-ordering of a set *S* there is a unique ordinal number associated with each element *x* of *S* that marks its position, namely the ordinal number of the well-ordering we get by restricting *L* to $\{y \in S: yLx \text{ and } y \neq x\}$.

Ernst Zermelo discovered a deep connection between cardinal and ordinal numbers: The cardinal numbers are well-ordered, so that the infinite cardinals can be placed in a sequence, $\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ldots$, and every cardinal number has the form $\aleph_\alpha$, for some ordinal α.

MR. RUSSELL'S BARBER

The program of securing the theory of sets on a unified axiomatic basis was trenchantly pursued by Gottlob Frege. A prerequisite for such a program is a system of logic that is both highly powerful and fully explicit, and before Frege there was no such logic. Frege's program has a philosophical motive. He wanted to show that the laws of arithmetic are analytic, so that, by providing suitable definitions, the laws of arithmetic can be reduced to pure logic. The key idea is that to say Traveler is a horse and to say that Traveler is an element of {*x*: *x* is a horse} are two ways of saying the same thing, just like "Lee rode Traveler" and "Traveler was ridden by Lee."

The specific form taken by Frege's reduction of arithmetic to logic depends on his doctrine of concepts and objects. Proper names (such as "Traveler"), definite descriptions (such as "the horse Lee rode into battle"), and sentences (such as "Traveler is a horse") are *saturated* expressions, and they denote *objects*. Under Frege's rather eccentric usage, sentences are a species of name; they denote either the True or the False, which are objects. Open sentences, like "x is a horse" and "Lee rode x into battle," are *unsaturated*, and they denote *concepts*. When we complete an open sentence by replacing the variable by a name, we get a sentence that denotes either the True or the False. Open sentences are a special case of *function sign*, an unsaturated expression whose completion yields a name. A concept is a special kind of function, one that cannot take any values other than the True and the False. There are, in addition, functions that demand more than one argument, represented by such multiply unsaturated phrases as "*x* rode *y*," and so-called *second-level* functions that take ordinary functions as arguments.

The fundamental principle of Frege's set theory, his Basic Law V (Basic Laws I through IV are unexceptionable principles of logic), associates a set, the object {*x*: *Fx*}, with each concept *F* in such a way that, for any concepts *F* and *G*, {*x*: *Fx*} is equal to {*x*:*Gx*} if and only if, for every object *x*, we have *Fx* if and only if *Gx*. The left-to-right direction of this axiom is the axiom of extensionality, which has proven harmless. Extensionality is what distinguishes sets from properties. The property of being a human being is different from the property of being a featherless biped, even though {*x*: *x* is a featherless biped} = {*x*: *x* is a human being}. The right-to-left direction has proven deeply problematic. It asserts that the second-level function taking the concept *F* to {*x*: *Fx*} is one-one. On the basis of this axiom we can define "∈": $z \in y$ if and only if, for some *F*, $y = \{x: Fx\}$ and *Fz*, and we can derive the so-called *comprehension principle* that, for any *F* and

z, $z \in \{x: Fx\}$ if and only if Fz, just as in the proof of Cantor's theorem, and just as before, we can derive a contradiction by asking whether $\{x: x \notin x\}$ is an element of itself.

Bertrand Russell discovered the paradox and communicated it to Frege just as the second volume of Frege's monumental *Grundegesetze der Arithmetic* was going to press. Frege regarded it as devastating. "With the loss of my Basic Law V," he wrote in reply to Russell (1902, pp. 127–128), "not only the foundations of my arithmetic, but also the sole possible foundations of arithmetic, seem to vanish." Later scholarship, led by George Boolos, has seen the devastation as not quite so complete as Frege took it. Roughly speaking there are two principal components to the *Grundgesetze*: First the employment of Basic Law V (together with suitable definitions) to derive what Frege calls Hume's Principle, that, for any concepts F and G, the numbers associated with F and G are equal if and only if there is a one-one correspondence between the objects that fall under F and those that fall under G; and second the derivation from Hume's Principle of the fundamental laws of arithmetic. The latter component is a substantial mathematical accomplishment that is unharmed by Russell's paradox.

A couple of other set-theoretic paradoxes emerged at about the same time, one, due to Cantor, involving cardinal numbers, and the other, due to Cesare Burali-Forti, about ordinal numbers.

For Cantor's paradox, let V be the set of all sets. Since every set of sets is a set, $\wp(V)$ is a subset of V, and so $\#(\wp(V)) \leq \#(V)$. Yet Cantor's theorem tells us that $\#(\wp(V)) > \#(V)$. Cantor concluded from the contradiction that there is no set of all sets, invoking a distinction reminiscent of Aristotle's distinction of potentially and actually infinite. The sets measured by the $\aleph_\alpha$s are transfinite, whereas the set of all sets, if there were such a thing, would be absolutely infinite. There is no such thing as V because the sets do not form a completed whole. There is no absolute infinity in mathematics; absolute infinity is the province of God alone.

For Burali-Forti's paradox, consider that the ordinals are well-ordered, and so they have an ordinal number. Call it α. α also the ordinal of the collection of ordinals less than α and so there is an order-isomorphism f from the collection of all the ordinals to the collection of ordinals less than α. $f(\alpha) < \alpha$, and so, since the ordering on the ordinals is well-founded, there has to be a least ordinal β with $f(\beta) \neq \beta$. We have $f(\beta)$ = the least ordinal greater than all the members of $\{f(\gamma): \gamma < \beta\}$ = the least ordinal greater than all the members of $\{\gamma: \gamma < \beta\}$; = β.

Mirimanoff's paradox emerged a little later. Let us say that a set is *hereditarily well-founded* if it belongs to a collection C with the following properties: Every element of an element of C is an element of C; and the elements of C are well-founded (that is, the restriction to an element of C of the elementhood relation is well-founded). It is easy to verify that the collection of all hereditarily well-founded sets is hereditarily well-founded. But this gives us the absurd prospect of a well-founded set that is an element of itself.

Russell illustrated the logical structure of his paradox with an amusing example. Imagine a village whose barber (an adult male villager) shaves all and only the adult male villagers who do not shave themselves. A contradiction arises when we inquire whether the barber shaves himself, by reasoning exactly analogous to the thinking that gets Russell's paradox. Unlike the set-theoretic paradox, however, the puzzle about the barber has an easy solution. There can be no such barber, however plausible the story that said there was one sounded on first hearing. One would like to obtain a similar resolution to Russell's paradox, denying that there is such a set as $\{x: x \notin x\}$ (and, presumably, also that there are such sets as V, the set of all ordinals, and the set of all hereditarily well-founded sets) by restricting the range of open sentences that can be substituted for "F" in the comprehension principle. The trick is to do this in a principled, credible way that avoids contradictions while maintaining the set existence principles required to do mathematics. Before asking how this might be done, let us examine the semantic analogues of the set-theoretic paradoxes.

SEMANTIC PARADOXES

Semantics, as Alfred Tarski characterized it, is the branch of linguistics that studies the connections between expressions of a language and the things or states of affairs those expressions refer to. Its principal theoretical concepts are *truth*, *reference*, and *satisfaction*. A name, like "Traveler" or "Robert E. Lee's horse," refers to (or *names* or *denotes*) an object, in this case a stallion. An open sentence, like "x is a horse," represents a concept. The sentence got by substituting a name for the variable in an open sentence is true just in case the object referred to by the name falls under the concept represented by the open sentence. Because Traveler falls under the concept *horse*, the sentence "Traveler is a horse" is true. The reason for using the variable "x" to mark the place in the open sentence where a name needs to be supplied is to accommo-

date open sentences like "*x* rode *y* into battle," which represent concepts with more than one argument. The account of satisfaction needs to be complicated a bit to allow for such open sentences, but that need not concern us here.

Another, more substantial, complication is that one should not really speak of a sentence being true, but rather of a sentence being true in a language in a context, or perhaps of a sentence expressing, in a language in a context, a proposition that is true. "I am now riding Traveler" is true when Lee says it while riding his horse, but it is false in most other contexts. This complication too can be set aside here.

Here our concern is with the semantic paradoxes. Epimenides is credited with the earliest formulation, although it is doubtful that he recognized that his statement was paradoxical. Someone acutely aware of the paradox was Eubulides of Miletas (a contemporary of Aristotle), who asked, "A man says that he is lying; is what he says true or false?" Eubulides formulated other notorious paradoxes, among them the *Bald Man* (The observation that plucking a single hair from a man who is not bald will not make him bald leads, by multiple application, to the conclusion that not even a man with no hair at all is bald) and the *Hooded Man* (You know who your father is, and you do not know who the hooded man is, even though, unknown to you, the hooded man is your father; this violates the law of identity, which allows the exchange of names that denote the same thing).

To avoid fretting about indexicals and also to avoid consternation that only purposefully false statements count as "lies," let us consider what we may call the Liar Sentence, the sentence "The Liar Sentence is not true." We would naively expect the notion of truth to be governed by the (T)-schema, "'__________' is true if and only if __________;" "Traveler is a horse," for example, is true if and only if Traveler is a horse. However filling the blank with "The Liar Sentence is not true," and noting that "The Liar Sentence is not true" = the Liar Sentence, results in contradiction.

The Liar paradox does not have a direct set-theoretic analogue, but there are paradoxes involving satisfaction and reference that have such analogues. The analogue to Russell's paradox is due to Kurt Grelling. We would intuitively expect satisfaction to be governed by a principle exactly parallel to the comprehension principle in set theory, telling us, for example, that, for any *y*, *y* satisfies "*x* is a horse" if and only if *y* is a horse. The phrase "*x* is a horse" is not a horse, and so "*x* is a horse" does not satisfy itself, whereas "*x* is an open sentence" does satisfy itself. For any *y*, *y* satisfies "*x* is an open sentence that does not satisfy itself" if and only if *y* is an open sentence that does not satisfy itself. Taking *y* to be the open sentence "*x* is an open sentence that does not satisfy itself" yields a contradiction.

Cantor's paradox is obtained by generalizing Cantor's argument that there are more real numbers between 0 and 1 than there are positive integers, which proceeds by assuming for *reductio ad absurdum* that there were a list that enumerated all the real numbers between 0 and 1, then asking where on the list there appears the number *r* given by stipulating that the *n*th digit in the binary decimal expansion of *r* is equal to one if and only if zero is the *n*th digit in the binary decimal expansion of the *n*th number on the list. *Richard's paradox* invites us to consider, in particular, the list gotten by enumerating the English expressions that denote real numbers between 0 and 1 in alphabetical order. Cantor's argument gives us a real number between 0 and 1 that is not named by any expression of English. But is Cantor's number not named by the expression "the number *r*, between 0 and 1, such that the *n*th binary digit of *r* is equal to one if and only if zero is the *n*th binary digit of the number named by the alphabetically *n*th English phrase that names a real number between 0 and 1"?

The number of English expressions that name ordinal numbers is $\aleph_0$ since the expressions are finite strings of words from a finite vocabulary. There are more than $\aleph_0$ ordinals, so there are ordinals not named by an expression of English, and hence a least ordinal number not named by an expression of English. But "the least ordinal number not named by an expression of English" names it. This is *König's paradox*. *Berry's paradox* is the finitary version, got by noting that "the least natural number that cannot be named by an English expression of fewer than thirty syllables" is an English expression that names a natural number in twenty-eight syllables.

PRINCIPIA MATHEMATICA

Alfred North Whitehead and Bertrand Russell undertook to solve simultaneously both the set-theoretic paradoxes and the semantic paradoxes. The aim of their highly ambitious *Principia Mathematica* was to secure all of mathematics on a basis in pure logic. In their system the role hitherto played by sets was taken over by *propositional functions*, which are a kind of amalgam of Frege's concepts and Frege's propositions. *Propositions* are, for Frege, the objects of belief and judgment, the sort of thing referred to by "that" clauses in English. According to Russell if you prefix the word "that" to an open sentence, you

get an expression that names a propositional function, a function taking objects to propositions. If you supply an object as argument, the output is the proposition you would express if you substituted a name of that object into the open sentence. With Traveler as argument, the propositional function designated by "that Lee rode *x* into battle" yields the proposition that Lee rode Traveler into battle.

This account was considered and rejected by Frege on the basis that it would yield the result that the proposition that Traveler is white is identical to the proposition that the horse Lee rode into battle is white, in spite of the fact that someone who did not realize that Traveler is the horse Lee rode into battle might believe one but not the other. For Frege, the argument of the propositional function is not the horse Traveler but the sense (*Sinn*) of the name "Traveler." Russell thought he could thwart Frege's objection by a logical analysis according to which "the horse Lee rode into battle" is, despite appearances, not a denoting phrase. The relative merits of Frege's and Russell's conceptions of propositional functions have been much debated.

Whitehead and Russell proposed to avert the paradoxes by adopting the *vicious circle principle* (which they attribute to Henri Poincaré), according to which you cannot have a proposition that refers to itself since before you could formulate such a proposition you would have to already possess the proposition you were trying to formulate. You cannot have a propositional function that has itself or any of its values as argument, nor can you have a propositional function in whose formulation you are required to talk about the propositional function or any of its arguments. To formulate an analogue to Russell's paradox in terms of propositional functions, you would have to suppose that the phrase "that *x* is a propositional function not true of itself" denoted a propositional function, and such a propositional function would violate the vicious circle principle. Whitehead and Russell adopted a (maddeningly elaborate) formalism in which such phrases were grammatically ill-formed.

The vicious circle principle evades the paradoxes (as far an anyone knows), but it also rules out ordinary mathematics. If *r* is the least upper bound of a given collection of real numbers then it is the least element of a totality that includes *r* itself.

In *Principia Mathematica*, the propositional functions are arrayed in layers, where the level of a given function is determined by the levels of its possible arguments and also by the levels of the propositional functions utilized in defining the given function. A propositional function is said to be *predicative* if it is at the lowest level it could possibly be at, given what its arguments are. Either it is defined without referring to anything beyond its potential arguments (say, by giving a list), or the things referred to are sufficiently low-level that they do not affect the level of the function. Whitehead and Russell obtained the least upper bound principle by adopting the *axiom of reducibility*, according to which for every propositional function there is a predicative propositional function true of the same things. Given a collection *C* of real numbers, take a predicative function *Fx* that is coextensive with "*x* is an upper bound of *C*," we get the least upper bound of *C* as the least number that satisfies *Fx*.

The justification of the axiom of reducibility is purely pragmatic—it is needed for mathematics—and its adoption seriously undermines Whitehead and Russell's claim to have reduced mathematics to logic.

Once we have the axiom of reducibility on board, there in no longer any useful purpose in having the position in the hierarchy of a propositional function depend on the positions of the things we refer to in defining the function as well as on the positions of the potential arguments. We can obtain both the same mathematical results and the same degree of insulation from paradox simply by taking the type of a propositional function to be immediately above the types of its arguments, no matter how the propositional function is defined. Frank Ramsey first recognized this, effecting an enormous simplification in the system. W. V. Quine took the observation a step further, noticing that there was no longer any benefit in supposing coextensive propositional functions to be distinct, so that we could take the things the "Ramseyfied" theory was about to be sets and relations, rather than propositional functions of one or more variables.

The system can be further streamlined by replacing talk about binary relations between individuals with talk about sets of ordered pairs of individuals, replacing talk about ternary relations with talk about sets of ordered triples, and so on. There is no need to take ordered triples as primitive, since we can define $\langle a,b,c\rangle$ as the ordered pair $\langle a,\langle b,c\rangle\rangle$. In fact—this is an ingenious observation of Norbert Weiner—there is no need to take ordered pairs as primitive, since we can define $\langle a,b\rangle$ as $\{\{\{a\},\emptyset\},\{\{b\}\}\}$. This stipulation enables us to derive the principle that, for any *a*, *b*, *c*, and *d*, $\langle a ,b\rangle = \langle c,d\rangle$ if and only if $a=c$ and $b=d$, which is the only thing one ever needs to know about ordered pairs. The resulting system is arrayed in a simple hierarchy. There are individuals, sets of individuals, set of sets of individuals, and so on. It is this hierarchical struc-

ture, rather than the vicious circle principle, that precludes paradox.

ZERMELO-FRAENKEL SET THEORY

The most prominent of *Principia Mathematica*'s rivals originates, not in a philosophical analysis of methods of reasoning that lead to the paradoxes, but in a mathematical examination of ways of reasoning that never cause problems. Zermelo's 1904 proof that every set can be well-ordered, with its corollary that the cardinal numbers are well-ordered, met considerable resistance. It was felt that it skirted too close the edge of the newly discovered paradoxes. Poincaré, for example, complained about vicious circularity. Zermelo replied that the methods of forming sets that figure in the paradoxes are far removed from the methods that are gainfully employed by working mathematicians, and that the principles that figure in the deduction of the well-ordering theorem fall into the latter category. To make this reply more precise Zermelo wrote down axioms of set theory sufficient to derive the well-ordering theorem, in the hope that all could see that the proof required only well-established principles of workaday mathematics that had never been implicated in paradox.

Zermelo's axioms did not come equipped with a diagnosis of the paradoxes, but a couple of widely accepted further principles do have diagnostic import. Although Zermelo's axioms are immensely powerful, they omit some common and apparently harmless mathematical practices, like forming infinite sequences of the form $\langle s_0, s_1, s_2, s_3 \ldots \rangle$. Abraham Fraenkel proposed to rectify this situation by adopting the *replacement axiom schema*, which says that for any open sentence that defines a function, if the inputs to the function form a set, so do the outputs. In Frege's logic, in which there was one style of variables ranging over sets and other objects and another style ranging over functions, the replacement axiom would be expressed by saying the restriction of a function to a set domain invariably has a set as its range. Fraenkel is only able to produce a schema that applies to definable functions because his logical resources are restricted to the first-order predicate calculus, which only has variables ranging over individuals. The new principle is formulated as a formula that contains a schematic letter, so understood that the formulas of the language of set theory obtained by substituting a formula for the schematic letter are regarded as axiomatic. This retreat to first-order logic results from relinquishing the program of trying to produce a logic so powerful that set theory can be reduced to it. The existence of sets cannot be established by logic alone, and the first-order formulation makes set theory's existence assumptions fully explicit.

The replacement axioms assure us that an open sentence has a set as its extension unless the things that satisfy it are more numerous than the members of any set. The doctrine of *limitation of size* has it that things form a set unless there are too many of them, and that the paradoxes arise from attempts to form sets that are too large to hang together. The Burali-Forti paradox, for example, tells us that the ordinals are too numerous to form a set.

Mirimanoff's paradox distinguishes the hereditarily well-founded sets—what Mirimanoff calls the *ordinary sets*—from the others. It turns out that extraordinary sets (if there are any) are never needed for mathematics, and if we restrict our attention to ordinary sets, adopting an axiom that the elementhood relation is well-founded, an attractive picture of the universe of set appears. Sets are built up in stages. At the bottom level the so-called *urelements* are whatever non-sets you may want to count or measure. (For pure, as opposed to applied, mathematics, no urelements are needed.) At the second level are sets of urelements. At the third level are sets whose elements are urelements and sets of urelements. And so on. At each stage the available building blocks are the urelements and the sets built at earlier stages, and every set is constructed at some, possibly transfinite, stage. Because new sets and new ordinals are added at every stage, there is no stage at which one constructs a set that contains all the ordinals or a set that contains all the sets that do not contain themselves (which would be a set that contained all sets). The purported sets that threaten paradox never appear.

The two strategies for blocking the paradoxes—limitation of size and construction in stages—are by no means in conflict. Indeed one could argue (although it is certainly not obvious) that the stage construction ensures that no stage ever produces a set large enough to violate the limitation of size. On the other hand Peter Aczel (1988) has devised an alternative to standard set theory, provably consistent if ordinary set theory is, that upholds limitation of size but allows non-well-founded sets in great profusion.

The Whitehead-Russell system and the Zermelo-Fraenkel system are by no means the only extant responses to the set-theoretic paradoxes—Quine, for example, devised a method for restricting the comprehension principle so as to allow a universal set, without apparent contradiction—but they are the most prominent. Their rivalry is not so implacable as first appears. Indeed Kurt Gödel (1944/1983) has noted that we can think of the Zermelo-Fraenkel system as obtained from

Principia Mathematica, as simplified by Ramsey, Quine, and Weiner, by generalizing in two directions. First we allow the types to be cumulative, so that the possible elements at a type level are not just the sets of the immediately preceding type but sets of all the preceding types. Second we allow the type levels to extend into the transfinite.

One question the axiomatizations leave unanswered is how comprehensive the theory of sets is intended to be. Not all collections are sets. A stamp collection for example can survive the acquisition of a new stamp, whereas when you add a stamp to a set of stamps, you get a different set. But perhaps set theory accounts for all extensional collections. If not then the set theorist has two tasks: To say what extensional collections there are, and to say which of the extensional collections are sets. John von Neumann gave such an account. His theory has two kinds of *classes*, namely, sets and *proper classes*. Classes are made up of sets and urelements, and a class is a set if and only if it is not equinumerous with the class of all ordinals. The proper classes form the top element of the cumulative hierarchy.

Zermelo has a different perspective, reminiscent of Aristotle's doctrine that a line or the integers never form a completed whole. The universe of set theory does not form a completed whole, according to Zermelo. Candidates for the "universe" of set theory are only provisional, and one can always advance to a higher perspective from which a candidate universe is seen as a set within a larger universe.

SEMANTIC PARADOXES

In looking at possible solutions to the set-theoretic paradoxes, the semantic paradoxes have been set aside. One cannot happily say about the Liar Sentence the same thing one wants to say about the alleged set of all ordinals, namely, that there is no such sentence. One might want to say that what goes wrong with the Liar Sentence is that it does not express a proposition, but any satisfaction this gives us is short-lived. A propositionalist theory of truth has to account for two things, the truth conditions for propositions and the connection between a sentence and the proposition it expresses, and the latter relation remains troublesome. Consideration of the Propositional Liar Sentence ("The Propositional Liar Sentence does not express a true proposition") seems to force us the self-defeating conclusion that the Propositional Liar Sentence does not express a proposition, and hence that it does not express a true proposition.

The standard response to the semantic paradoxes was given by Alfred Tarski, who insists that, in developing a semantic theory, the language one employs (the *metalanguage*) must be richer in expressive power than the language one is talking about (the *object language*), so that one can never formulate the theory of truth, reference, and satisfaction for a language within the language itself.

A number of ingenious extensions of Tarski's basic idea have been developed. Notable among them is Saul Kripke's (1975) demonstration that, thinking of such troubled sentences as the Liar as neither true nor false, one can add the predicate "true" to a language and partition the sentences of the resulting language in such a way that a sentence *S* is counted as true, false, or undecided according as the statement that *S* is true is accounted true, false, or undecided. The enriched language cannot, however, express the equivalence of *S* with the statement that *S* is true, nor can it express the proof that the Kripke construction yields the equivalence. These things can only be said within a richer metalanguage. Indeed, in the best known version (there are a number of variants), the addition of the truth predicate results in a drastic restriction in the range of truth-preserving inferences, so that the object language has an enervated logic in which nothing resembling ordinary mathematical or philosophical reasoning can be carried on (as Solomon Feferman's investigations have made abundantly clear). Hartry Field (2003) has proposed enriching the Kripke construction by adding a new, nonclassical conditional, which behaves enough like the everyday "if…, then" to accommodate a substantial range of familiar inferences. With Field's novel interpretation of "if and only if," the (T)-sentences are all counted as true.

Field's construction is too complicated to describe here, but its key idea comes from revision theory, which employs full classical logic but regards the "if and only if" that appears in the (T)-sentences as a special connective that represents definitional equivalence. Developed by Anil Gupta and Nuel Belnap (1993) on a foundation laid down independently by Gupta and Hans Herzberger, revision theory treats the (T)-sentences as defining "true," and it ascribes to the "if and only if" of definition a special logic that allows for circular definitions. If "*F*" is defined by "*F*(*x*) if and only if __________," where the defined predicate "*F*" appears in the blank, and if *C* is proposed as a possible candidate for the extension of "*F*," then the map that takes *C* to {*x*: *x* satisfies __________ when *C* is taken as the extension of "*F*"} gives us, by iteration, better and better candidates for the extension of

"F." Throughout the revision process, the Liar sentences keeps flip-flopping between true and false, but sentences that have an intuitively correct truth value eventually settle down to their intuitively expected values, even in the presence of extensive self-reference and cross-reference. Tarski's restriction is still observed, inasmuch as the entire construction is developed within a richer metalanguage.

Tarski's doctrine works happily for formal languages, but it leaves us unable to understand how the notion of truth and reference apply to natural languages, since "true in English" is a phrase of English and not some unnatural metalanguage.

Some progress has been made by adapting the *Principia Mathematica* idea of subscripted "true"s. There is a predicate "true_0" that applies to sentences that contain no semantic notions; a predicate "true_1" that applies to sentences with no semantic notions other than "true_0;" "true_2" that applies to sentences with no semantic notion other than "true_0" and "true_1" and so on. Tyler Burge (1979) and Charles Parsons (1974) have proposed applying this notion to English by supposing that the English word "true" is ambiguous, and that disambiguating subscripts are tacitly ascribed by contexts in such a way that a truth attribution is supplied a subscript one greater than the maximum of those that appear in the sentences one is talking about. Eubulides's derivation of a contradiction is seen as committing a fallacy of equivocation, inasmuch as the tacit subscript changes during the course of his argument.

The *Principia*-inspired approach still has limitations. For one thing there does not appear to be any uniform, non-arbitrary way of coping with situations in which *A* talks about all the things *B* says at the same time *B* talks about all things *A* says. For another the description of the subscripting machinery lies outside the object language, so that we are provided with no good way of dealing with "This sentence is not true_α, for any α."

A different approach tries to consolidate the Liar paradox and the Bald Man by developing the idea that, if Harry is a borderline case of "bald," "Harry is bald" should be neither true nor false. "'Harry is bald' is either true or false" follows logically (defining falsity as truth of the negation) from the conditionals we get by substituting "Harry is bald" and "Harry is not bald" into the right-to-left direction of the (T)-schema. Allowing that "Harry is bald" is neither true nor false requires restricting the right-to-left direction of (T) so that it does not apply to such things as border applications of vague terms, which are semantically defective. This restriction yields an attractive response to the Bald Man, but the answer to the Liar is problematic. The left-to-right direction of (T) ("If 'The Liar Sentence is not true' is true then the Liar Sentence is not true") suffices to yield the conclusion that the Liar Sentence is untrue, and hence, because the right-to-left (T)-schema fails for it, that the statement that Liar Sentence is untrue is semantically defective. This result is unwelcome, because it tells us that a conclusion can be derived by rigorous, careful deduction from secure premises and still be semantically defective.

The argument of the previous paragraph is adapted from Richard Montague's (1960) argument that a Liar-type paradox can be obtained for necessity in place of truth. Also, as Montague and David Kaplan (1960) showed, for knowledge. It seems safe to say that the semantic paradoxes are currently less well-managed than the set-theoretic paradoxes.

See also Aristotle; Cantor, Georg; Frege, Gottlob; Galileo Galilei; Gödel, Kurt; Hamilton, William; Kaplan, David; Kripke, Saul; Liar Paradox, The; Montague, Richard; Neumann, John Von; Poincaré, Jules Henri; Quine, Willard van Orman; Ramsey, Frank Plumpton; Relativity Theory; Russell, Bertrand Arthur William; Set Theory; Tarski, Alfred; Whitehead, Alfred North; Zeno of Elea.

Bibliography

Aczel, Peter. *Non-Well-Founded Sets*. Stanford, CA: CSLI, 1988.

Aristotle. *Physics*. 2 vols. Translated by Philip H. Wicksteed and Francis M. Cornford. Cambridge, MA: Harvard University Press, 1929 and 1934.

Boolos, George S. *Logic, Logic, and Logic*. Cambridge, MA, and London: Harvard University Press, 1998.

Burali-Forti, Cesare. "Una Questione sui Numeri Transfiniti." *Rendiconti del Circolo Matematico di Palermo* 11 (1897): 154–164. English translation by Jean van Heijenoort in van Heijenoort, 104–111.

Burge, Tyler. "Semantical Paradox." *Journal of Philosophy* 76 (1979): 169–198. Reprinted in Martin, 83–117.

Cantor, Geog. "Beiträge zur Begründung der transfiniten Mengenlehre." *Mathematische Annalen* 46 (1895): 481–512 and 49 (1897): 312–356. English translation by Philip E. B. Jourdain. New York: Dover, 1955.

Dedekind, Richard. *Stetigkeit und irrationale Zahlen*. Braunschweig: F. Veiwig, 1872. English translation by Wooster W. Beeman in Dedekind *Essays on the Theory of Numbers*, 1–27. New York: Dover, 1963.

Feferman, Solomon. "Toward Useful Type-Free Theories, I." *Journal of Symbolic Logic* 49 (1984): 75–111. Reprinted in Martin, 237–287.

Field, Hartry. "A Revenge-Immune Solution to the Liar Paradox." *Journal of Philosophical Logic*32 (2003): 139–177.

Fraenkel, Abraham. "Über die Zermelosche Begründung der Mengenlehre." *Jahresbericht der Deutchen Mathematiker-Vereinigung* 3011 (1921): 97–98.

Fraenkel, Abraham. "Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre." *Mathematische Annalen* 86 (1922): 230–237.

Frege, Gottlob. *Begriffsschrift*. Halle: Nebert, 1879. Translated by Stefan Bauer-Mengelberg in van Heijenoort, 1–82.

Frege, Gottlob. "Funktion und Begriff." 1881. English translation by Peter Geach and Max Black in *Translations*, 21–41.

Frege, Gottlob. "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik* 50 (1892): 25–50. English Translation by Peter Geach and Max Black in *Translations*, 56–78.

Frege, Gottlob. *Grundgesetze der Arithmetik*, 2 vols. Jena: Pohle, 1893 and 1903. Partial translation by Montgomery Furth. Berkeley and Los Angeles: University of California Press, 1964.

Frege, Gottlob. "Letter to Russell." 1902. English translation by Beverly Woodward in van Heijenoort, 126–128.

Frege, Gottlob. *Translations from the Philosophical Writings*. 2nd. ed. Oxford: Basil Blackwell.

Galileo Galilei. *Discorsi e dimostrazioni matematiche*. Translated by Henry Crew and Alfonso de Salvio. New York: Dover, 1956.

Gödel, Kurt. "Russell's Mathematical Logic." In *The Philosophy of Bertrand Russell*, edited by Paul A. Schilpp, 125–153. Evanston and Chicago: Northwestern University Press, 1944. Reprinted in Paul Benacerraf and Hilary Putnam, eds., *Philosophy of Mathematics* 2nd ed., 447–469. Cambridge, U.K.: Cambridge University Press, 1983.

Grelling, Kurt, and Leonard Nelson. "Bemerkungen zu den Paradoxien von Russell und Burali-Forti." *Abhandlungen der Fries'schen Schule neue Folge* 2 (1908): 301–334.

Gupta, Anil. "Truth and Paradox." *Journal of Philosophical Logic* 11 (1982): 1-60. Reprinted in Martin, 175–235.

Gupta, Anil, and Nuel Belnap. *The Revision Theory of Truth*. Cambridge, MA: MIT Press, 1993.

Hallett, Michael. *Cantorian Set Theory and Limitation of Size*. Oxford: Clarendon, 1984.

Herzberger, Hans G. "Notes on Naïve Semantics." *Journal of Philosophical Logic* 11 (1992): 61-102. Reprinted in Martin, 133–174.

König, Julius. "Über die Grundlagen der Mengenlehre und das Kontinuumproblem." *Mathematische Annalen* 61 (1905): 156–160. Translated by Stefan Bauer-Mengelberg in van Heijenoort, 145-149.

Keefe, Rosanna, and Peter Smith, eds. *Vagueness: A Reader*. Cambridge, MA, and London: MIT Press, 1996.

Kripke, Saul A. "Outline of a Theory of Truth." *Journal of Philosophy* 72 (1975): 690–716. Reprinted in Martin, 53–81.

McGee, Vann. *Truth, Vagueness, and Paradox*. Indianapolis: Hackett, 1991.

Martin, Robert L., ed. *Recent Essays on Truth and the Liar Paradox*. Oxford and New York: Oxford University Press.

Mirimanoff, Dimitry. "Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles." *L'enseignement Mathématique* 19 (1917): 37–52.

Montague, Richard. *Formal Philosophy*. New Haven, CT: Yale, 1974.

Montague, Richard. "Syntactic Treatments of Modality, with Corollaries on Reflection Principles and Finite Axiomatizability." *Acta Philosophical Fennica* 16 (1963): 153–167. Reprinted in Montague, *Formal Philosophy*, 286–302.

Montague, Richard, and David Kaplan. "A Paradox Regained." *Notre Dame Journal of Formal Logic* 1 (1960): 79–90. Reprinted in Montague, *Formal Philosophy*, 270–285.

Parsons, Charles. "The Liar Paradox." *Journal of Philosophical Logic* 3 (1974): 381–412. Reprinted in Martin, 9–45.

Quine, Willard van Orman. "New Foundations for Mathematical logic." *American Mathematical Monthly* 44 (1937): 70–80. Reprinted in Quine, *From a Logical Point of View*. Cambridge, MA: 1953, 80–101.

Quine, Willard van Orman. *The Logic of Sequences: Harvard Dissertations in Philosophy*. New York: Garland, 1990.

Ramsey, Frank Plumpton. "The Foundations of Mathematics." *Proceedings of the London Mathematical Society* 25 (1925): 338–384. Reprinted in Ramsey, *The Foundations of Mathematics and Other Logical Essays*. London: Routledge and Kegan Paul, 1931: 1–61, and in Ramsey, *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1990, pp. 164–224.

Richard, Jules. "Les principes de mathématiques et le problème de ensembles." *Revue Générale des Sciences Pures et Appliquées* 16 (1905): 541. English translation by Jean van Heijenoort in van Heijenoort, 142–144.

Russell, Bertrand. "Letter to Frege." 1902. In van Heijenoort, 124–125.

Russell, Bertrand. "On Denoting." *Mind* n. s. 14 (1905): 479–493. Reprinted in Russell, *Logic and Knowledge*. London: Allen and Unwyn, 1956: 41–56.

Tarski, Alfred. "Grundlegung der wissenschaftenlichen Semantik." *Actes du Congrès International de Philosophie Scientifique* 3 (1930): 1–8. English translation by J. H. Woodger in Tarski, *Logic, Semantics, Metamathematics*, 401–408.

Tarski, Alfred. *Logic, Semantics, Metamathematics*. 2nd ed. Indianapolis: Hackett, 1983.

Tarski, Alfred. "Der Wahrheitsbegriff in den formalisierten Sprachen." *Studia Philosophica* 1 (1935): 261–405. English translation by J. H. Woodger in Tarski, *Logic, Semantics, Metamathematics*, 152–278.

Van Heijenoort, Jean, ed. *From Frege to Gödel*. Cambridge, MA, and London: Harvard University Press, 1967.

Von Neumann, John. "Eine Axiomatisierung der Mengenlehre." *Journal für die reine und angewandte Mathematik* 154 (1925): 219–240. English translation by Stefan Bauer-Mengelberg and Dagfinn Føllesdal in van Heijenoort, 393–413.

Whitehead, Alfred North, and Bertrand Russell. *Principia Mathematica*. 2nd ed. 3 vols. Cambridge, U.K.: Cambridge University Press, 1925 and 1927.

Wiener, Norbert. "A Simplification of the Logic of Relations." *Proceedings of the Cambridge Philosophical Society* 17 (1914): 387–390. Reprinted in van Heijenoort, 224–227.

Zermelo, Ernst. "Beweis, daß jede Menge wohlgeordnet werden kann." *Mathematische Annalen* 59 (1904): 514–516. English translation by Stefan Bauer-Mengelberg in van Heijenoort, 139–141.

Zermelo, Ernst. "Neuer Beweis für die Möglichkeit einer Wohlordnung." *Mathematische Annalen* 65 (1908): 107–28.

English translation by Stefan Bauer-Mengelberg in van Heijenoort, 183–198.

Zermelo, Ernst. "Über Grenzzahlen und Mengenbereiche." *Fundamenta Mathematicae* 16 (1930): 29–37.

Zermelo, Ernst. "Untersuchungen über die Grundlagen der Mengenlehre I." *Mathematische Annalen* 65 (1908): 261–281. English translation by Stefan Bauer-Mengelberg in van Heijenoort, 199–215.

Vann McGee (2005)

LOGICAL POSITIVISM

"Logical positivism" is the name given in 1931 by A. E. Blumberg and Herbert Feigl to a set of philosophical ideas put forward by the Vienna circle. Synonymous expressions include "consistent empiricism," "logical empiricism," "scientific empiricism," and "logical neopositivism." The name logical positivism is often, but misleadingly, used more broadly to include the "analytical" or "ordinary language" philosophies developed at Cambridge and Oxford.

HISTORICAL BACKGROUND

The logical positivists thought of themselves as continuing a nineteenth-century Viennese empirical tradition, closely linked with British empiricism and culminating in the antimetaphysical, scientifically oriented teachings of Ernst Mach. In 1907 the mathematician Hans Hahn, the economist Otto Neurath, and the physicist Philipp Frank, all of whom were later to be prominent members of the Vienna circle, came together as an informal group to discuss the philosophy of science. They hoped to give an account of science that would do justice—as, they thought, Mach did not—to the central importance of mathematics, logic, and theoretical physics, without abandoning Mach's general doctrine that science is, fundamentally, the description of experience. As a solution to their problems, they looked to the "new positivism" of Jules Henri Poincaré; in attempting to reconcile Mach and Poincaré they anticipated the main themes of logical positivism.

In 1922, at the instigation of members of the "Vienna group," Moritz Schlick was invited to Vienna as professor, like Mach before him (1895–1901), in the philosophy of the inductive sciences. Schlick had been trained as a scientist under Max Planck and had won a name for himself as an interpreter of Albert Einstein's theory of relativity. But he was deeply interested in the classical problems of philosophy, as Mach had not been.

Around Schlick, whose personal and intellectual gifts particularly fitted him to be the leader of a cooperative discussion group, the "Vienna circle" quickly established itself. Its membership included Neurath, Friedrich Waismann, Edgar Zilsel, Béla von Juhos, Felix Kaufmann, Feigl, Victor Kraft, Philipp Frank—although he was by now teaching in Prague—Karl Menger, Kurt Gödel, and Hahn. In 1926 Rudolf Carnap was invited to Vienna as instructor in philosophy, and he quickly became a central figure in the circle's discussions; he wrote more freely than the other members of the circle and came to be regarded as the leading exponent of their ideas. Carnap had been trained as a physicist and mathematician at Jena, where he had come under Gottlob Frege's influence. Like other members of the circle, however, he derived his principal philosophical ideas from Mach and Bertrand Russell.

Ludwig Wittgenstein and Karl Popper were not members of the circle but had regular discussions with its members. In particular, Wittgenstein was in close contact with Schlick and Waismann. Wittgenstein's *Tractatus Logico-Philosophicus* had a profound influence on the deliberations of the circle, where it was interpreted as a development of British empiricism.

The circle ascribed to Wittgenstein the "verifiability principle"—that the meaning of a proposition is identical with the method of verifying it—that is, that a proposition means the set of experiences that are together equivalent to the proposition's being true. Wittgenstein, they also thought, had shown how an empiricist could give a satisfactory account of mathematics and logic. He had recognized that the propositions of logic and mathematics are tautologies. (The logical positivists paid no attention to Wittgenstein's distinction between tautologies and identities.) They are "independent of experience" only because they are empty of content, not because, as classical rationalists had argued, they are truths of a higher order than truths based on experience.

In the German-speaking countries, the Vienna circle was a small minority group. For the most part, German-speaking philosophers were still committed to some variety of "German idealism." Neurath, with his strong sociopolitical interests, was particularly insistent that the circle should act in the manner of a political party, setting out to destroy traditional metaphysics, which he saw as an instrument of social and political reaction.

In 1928 the significantly named Verein Ernst Mach (Ernst Mach Society) was set up by members of the circle with the avowed object of "propagating and furthering a scientific outlook" and "creating the intellectual instru-

ments of modern empiricism." To welcome Schlick back to Vienna in 1929 from a visiting professorship at Stanford, California, Carnap, Hahn, and Neurath prepared a manifesto under the general title *Wissenschaftliche Weltauffassung, Der Wiener Kreis* (The Scientific World View: The Vienna Circle). This manifesto traced the teachings of the Vienna circle back to such positivists as David Hume and Mach, such scientific methodologists as Hermann Ludwig von Helmholtz, Poincaré, Pierre Maurice Marie Duhem, and Einstein, to logicians from Gottfried Wilhelm Leibniz to Russell, utilitarian moralists from Epicurus to John Stuart Mill, and to such sociologists as Ludwig Feuerbach, Karl Marx, Herbert Spencer, and Menger. Significantly absent were any representatives of the "German tradition"—even, although somewhat unfairly, Immanuel Kant.

In order to make its conclusions familiar to a wider world, the circle organized a series of congresses. The first of these was held in Prague in 1929 as a section of a mathematical and physical, not a philosophical, congress. It was jointly sponsored by the Ernst Mach Society and the Society for Empirical Philosophy, a Berlin group led by Hans Reichenbach and with such members as Walter Dubislav, Kurt Grelling and Carl Hempel, which stood close in its general approach to the Vienna circle.

Meanwhile, the international affiliations of the circle were increasing in importance. American philosophers like C. W. Morris emphasized the link between logical positivism and American pragmatism; Ernest Nagel and W. V. Quine visited Vienna and Prague. In Great Britain, logical positivism attracted the interest of such Cambridge-trained philosophers as L. Susan Stebbing and John Wisdom and the Oxford philosophers Gilbert Ryle and A. J. Ayer, the latter participating for a time in the deliberations of the circle. In France such philosophers of science as Louis Rougier were attracted by logical positivism, as were a group of neo-Thomists led by General Vouillemin, who welcomed the positivist critique of idealism. In Scandinavia, where the way had been prepared by the antimetaphysical philosophy of Axel Hägerström, a number of philosophers sympathized with the aims of the logical positivists; Eino Kaila, Arne Naess, Åke Petzäll, and Jørgen Jørgensen were prominent representatives of the international movement centering on logical positivism. The Polish logicians, especially Alfred Tarski, exerted a considerable influence on members of the circle, particularly on Carnap. German philosophers, except for Heinrich Scholz of Münster and the Berlin group, remained aloof. Undoubtedly, the organizational energies of the circle did much to bring into being in the 1930s an international community of empiricists; this was largely a consequence of the circle's isolation within the German countries themselves.

Meanwhile the circle was publishing. In 1930 it took over the journal *Annalen der Philosophie* and renamed it *Erkenntnis.* In the period from 1930 to 1940 it served as a "house organ" for members of the Vienna circle and their associates. In addition, the circle prepared a series of monographs under the general title *Veröffentlichungen des Vereines Ernst Mach* (from 1928 to 1934) and *Einheitswissenschaft* (edited by Neurath from 1934 until 1938).

During the 1930s, however, the Vienna circle disintegrated as a group. In 1931 Carnap left Vienna for Prague; in that year Feigl went to Iowa and later to Minnesota; Hahn died in 1934; in 1936 Carnap went to Chicago and Schlick was shot by a mentally deranged student. The meetings of the circle were discontinued. The Ernst Mach Society was formally dissolved in 1938; the publications of the circle could no longer be sold in German-speaking countries. Waismann and Neurath left for England; Zilsel and Kaufmann followed Feigl, Carnap, Menger, and Gödel to the United States. *Erkenntnis* moved in 1938 to The Hague, where it took the name *Journal of Unified Science*; it was discontinued in 1940. Logical positivism, too, disintegrated as a movement, absorbed into international logical empiricism.

CRITIQUE OF TRADITIONAL PHILOSOPHY

Mach denied that he was a philosopher. He was trying, he said, to unify science and, in the process, to rid it of all metaphysical elements; he was not constructing a philosophy. The general attitude of the Vienna circle was very similar. Schlick was the exception. With logical positivism, he argued, philosophy had taken a new turn, but logical positivism was nonetheless a philosophy. Carnap, in contrast, wrote that "we give no answer to philosophical questions and instead *reject all philosophical questions,* whether of Metaphysics, Ethics or Epistemology" (*The Unity of Science,* p. 21). Philosophy, on his view, had to be destroyed, not renovated.

Undoubtedly, this intransigent attitude to philosophy can in part be explained by the peculiar character of German idealism and its hostility to science. The logical positivists thought of themselves as extending the range of science over the whole area of systematic truth and as needing for that purpose to destroy the claim of idealist philosophers to have a special kind of suprascientific access to truth.

METAPHYSICS. Of the traditional branches of philosophy, the positivists rejected transcendental metaphysics on the ground that its assertions were meaningless, since there was no possible way of verifying them in experience. Nothing that we could possibly experience, they argued, would serve to verify such assertions as "The Absolute is beyond time." Therefore, the positivists held, it tells us nothing. The rejection of transcendental metaphysics was not a novelty; Hume had described transcendental metaphysics as "sophistry and illusion" and had alleged that it makes use of insignificant expressions; Kant and the neo-Kantians had rejected its claim to be a form of theoretical knowledge; Mach had sought to remove all metaphysical elements from science. But whereas earlier critics of metaphysics had generally been content to describe it as empty or useless or unscientific, the logical positivists took over from Wittgenstein's *Tractatus* the rejection of metaphysics as meaningless. The propositions of metaphysics, they argued, are neither true nor false; they are wholly devoid of significance. It is as nonsensical to deny as to assert that the Absolute is beyond time.

EPISTEMOLOGY. Neo-Kantians had sometimes suggested that philosophy could be reduced to epistemology or "theory of knowledge," which discussed such topics as "the reality of the external world." But assertions about the external world, the positivists argued, are quite as meaningless as assertions about the Absolute or about things-in-themselves. For there is no possible way of verifying the assertion that there is, or the assertion that there is not, an external world independent of our experience. Realism and idealism, considered as epistemological theses, are equally meaningless. So far as epistemology has any content, it reduces to psychology, to assertions about the workings of the human mind, and these have nothing to do with philosophy.

ETHICS. The logical positivists disagreed about ethics. Of course they all rejected any variety of transcendental ethics, any attempt to set up a "realm of values" over and above the world of experience. Assertions about values, thus conceived, fall within the general province of transcendental metaphysics and had therefore to be rejected as nonsensical. But whereas Schlick sought to free ethics from its metaphysical elements by converting it into a naturalistic theory along quasi-utilitarian lines, Carnap and Ayer argued that what are ordinarily taken to be ethical assertions are not assertions at all. To say that "stealing is wrong," for example, is neither, they suggested, to make an empirical statement about stealing nor to relate stealing to some transcendental realm. "Stealing is wrong" either expresses our feelings about stealing, our feelings of disapproval, or, alternatively (positivists' opinions differ about this), it is an attempt to dissuade others from stealing. In either case, "stealing is wrong" conveys no information.

PHILOSOPHICAL MEANINGLESSNESS. In general, the positivists explained, when they said of philosophical assertions that they were meaningless, they meant only that they lacked "cognitive meaning." Ethical and metaphysical assertions have emotional associations; this distinguishes them from mere jumbles of words. Such statements as "God exists" or "Stealing is wrong" are, on the face of it, very different from a collocation of nonsense syllables. But the fact remains, the positivists argued, that such "assertions" do not convey, as they purport to do, information about the existence or character of a particular kind of entity. Only science can give us that sort of information.

Not all philosophers, however, have devoted their attention to describing pseudo entities such as "the Absolute" or "values" or "the external world." Many of them have been mainly concerned with empirical-looking concepts such as "fact," "thing," "property," and "relation." Russell's lectures on logical atomism and Wittgenstein's *Tractatus* are cases in point.

Wittgenstein suggested, however, that the sections in the *Tractatus* in which he talked about facts, or attempted to show how propositions can picture facts, must all in the end be rejected as senseless—as attempts to say what can only be shown. For it is impossible in principle to pass beyond our language in order to discuss what our language talks about. Philosophy is the activity of clarifying; it is not a theory.

Schlick carried to its extreme Wittgenstein's *Tractatus* doctrine that philosophy is an activity. Philosophy, he suggested, consists in the deed of showing in what the meaning of a statement consists; that is, philosophy is a silent act of pointing. The ultimate meaning of a proposition cannot consist in other propositions. To clarify, therefore, we are forced in the end to pass beyond propositions to the experience in which their meaning consists.

This view won few adherents. It was generally agreed that philosophers could not avoid making the sort of ontological assertions Wittgenstein made in the *Tractatus* and that it is altogether too paradoxical to suggest that all propositions about, for example, the relation between facts and language are nonsensical, even if "important" nonsense. Neurath, in particular, insisted that nonsense

cannot be "important," cannot act as a ladder by which we arrive at understanding, as Wittgenstein had said.

STATEMENTS ABOUT LANGUAGE. Carnap suggested that Wittgenstein was mistaken in supposing that his ontological assertions were without any sense. They were, however, meaningful assertions about language, not about a world beyond language. No doubt, Carnap admits, ontological statements have the appearance of being about the world or, at least, about the relation between language and the world. But this is so only because they have been wrongly formulated in what Carnap calls "the material mode."

Carnap distinguishes three classes of sentences: object sentences, pseudo object sentences, and syntactical sentences. Any ordinary sentence of mathematics or science is an object sentence. Thus, for example, "Five is a prime number" and "Lions are fierce" are both object sentences. Syntactical sentences are sentences about words and the rules governing the use of words. For example, "Five is not a thing-word but a number-word" and "Lion is a thing-word" are syntactical sentences. Pseudo object sentences are peculiar to philosophy; they look like object sentences but if rightly understood turn out to be syntactical sentences. To understand them rightly we have to convert them from the "material mode" into the "formal mode," that is, from sentences that look as if they are about objects into sentences that are obviously about words. Examples are "Five is not a thing but a number" and "Lions are things." Once these sentences are converted out of the "material mode" into the corresponding "formal" (or syntactical) mode, they can be discussed; in the material mode they are quite undiscussable.

But how are syntactical disputes to be settled? Suppose one philosopher asserts and another denies that "numerical expressions are class-expressions of the second level"—Carnap's "translation" of "numbers are classes of classes"—how is it to be determined which is correct? All such statements, Carnap argues, are relative to a language; they are either statements about the characteristics of some existing language or proposals for the formation of a new language. Fully expressed, that is, they have the form "In language *L*, such-and-such an expression is of such-and-such a type." It can be immediately determined whether such a syntactical statement is true by examining the language in question.

PROBLEMS OF POSITIVISM

VERIFIABILITY. The course taken by the subsequent history of logical positivism was determined by its attempts to solve a set of problems set for it, for the most part, by its reliance on the verifiability principle. The status of that principle was by no means clear, for "The meaning of a proposition is the method of its verification" is not a scientific proposition. Should it therefore be rejected as meaningless? Faced with this difficulty, the logical positivists argued that it ought to be read not as a statement but as a proposal, a recommendation that propositions should not be accepted as meaningful unless they are verifiable. But this was an uneasy conclusion. For the positivists had set out to destroy metaphysics; now it appeared that the metaphysician could escape their criticisms simply by refusing to accept their recommendations.

Recognition of this difficulty led Carnap to suggest that the verifiability principle is an "explication," a contribution to the "rational reconstruction" of such concepts as metaphysics, science, and meaning, to be justified on the quasi-pragmatic grounds that if we ascribe meaning only to the verifiable we shall be able to distinguish forms of activity that are otherwise likely to be confused with one another. It is not, however, by any means clear in what way the verifiability principle can be invoked against a metaphysician who takes as his point of departure that his propositions clearly have a meaning. The most that can be said is that the onus is then on the metaphysician to distinguish his propositions from others that he would certainly have to admit to be meaningless.

A second set of problems hinged on the nature of the entities to which the verifiability principle applies. Since "proposition" had ordinarily been defined as "that which can be either true or false," it seemed odd to suggest that a proposition might be meaningless. Yet it was no less odd to suggest that a sentence—a set of words—could be verified, even if there was no doubt that it could be meaningless. Ayer suggested as an alternative the word *statement,* and he wrote as if the problem were a purely terminological one. But it is a serious question whether "true," "false," and "meaningless" are alternative descriptions of the same kind of occurrence or whether to describe a sentence as "meaningless" is not tantamount to denying that any statement has been made, any proposition put forward. This would have the consequence that we can consider whether a statement is verifiable only after we have settled the question of the meaning of the sentence used to make the statement.

The logical positivists themselves were much more concerned about the fact that the verifiability principle threatened to destroy not only metaphysics but also science. Whereas Mach had been happy to purge the sci-

ences, the logical positivists ordinarily took for granted the substantial truth of contemporary science. Thus, it was a matter of vital concern to them when it became apparent that the verifiability principle would rule out as meaningless all scientific laws.

For such laws are, by the nature of the case, not conclusively verifiable; there is no set of experiences such that having these experiences is equivalent to the truth of a scientific law. Following Frank Plumpton Ramsey, Schlick suggested that laws should be regarded not as statements but as rules permitting us to pass from one singular statement to another singular statement. In Ryle's phrase, they are "inference-licenses." Neurath and Carnap objected to this on the ground that scientific laws are used in science as statements, not as rules. For example, attempts are made to falsify them, and it is absurd to speak of "falsifying a rule." Furthermore, Carnap pointed out, ordinary singular statements are in exactly the same position as laws of nature; there is no set of experiences such that if I have these experiences there must be, for example, a table in the room.

For these and comparable reasons "verifiability" was gradually replaced by "confirmability" or by the rather stronger notion of "testability." Whereas at first the meaning of a proposition had been identified with the experiences that we would have to have in order to know that the proposition is true, now this was reduced to the much weaker thesis that a proposition has a meaning *only if it is possible to confirm it,* that is, to derive true propositions from it. Carnap, in accordance with his "principle of tolerance," was prepared to admit that a language might be constructed in which only verifiable propositions would count as meaningful. He was content to point out that such a language would be less useful for science than a language that admits general laws. But most positivists, interested as they were in the actual structure of science, simply replaced the verifiability principle by a confirmability principle.

If, however, the original principle proved to be too strong, the new principle threatened to be too weak. For, on the face of it, the new principle admitted as meaningful such metaphysical propositions as "Either it is raining or the Absolute is not perfect." Whether the confirmability principle can so be restated as to act as a method of distinguishing between metaphysical statements as meaningless and scientific statements as meaningful remains a question of controversy.

UNIFICATION OF SCIENCE. A further set of problems hinges on the question of what sort of things act as "verifiers" or "confirmers." One of Mach's main concerns, which the logical positivists shared, had been to unify science, especially by rejecting the view that psychology is about an "inner world" that is different from the "outer world" that physical science investigates. The doctrine that both physics and psychology describe "experiences" made such a unification possible. In his earlier writings Carnap tried to show in detail how "the world" could be constructed out of experience, linked together by relations of similarity. But then a new difficulty arose; one about how it is possible to show that one person's experiences are identical with another's. On the face of it, an experience-based science is fundamentally subjective; science is verified only at the cost of losing its objectivity.

To overcome this difficulty, Schlick drew a distinction between "content" and "structure." We can never be sure, he argued, that the content of our experience is identical with the content of any other person's experience, for example, that what he sees when he says that he sees something red is identical with what we see when we say we see something red. For scientific purposes, however, this does not matter in the slightest. Science is interested only in the structure of our experience, so that provided, for example, we all agree about the position of red on a color chart, it is of no importance whether our experience of red differs.

Yet Schlick still thought that such "experiences" are what gives content, meaning, to science, converting it from a conceptual frame into real knowledge. Thus, it appears that the ultimate content of science lies beyond all public observation. There is no way of verifying that another person is even experiencing a content, let alone a content that is like or unlike the content of my experience.

PHYSICALIST THEORIES. Profoundly dissatisfied with the conclusion that the ultimate content of scientific truths is private, Neurath was led to reject the view—which logical positivists had so far taken for granted—that it is "experiences" which verify propositions. Only a proposition, he argued, can verify a proposition. Carnap accepted this conclusion and developed the conception of a "protocol statement," the ultimate resting point of verifications, a statement of such a nature that to understand its meaning and to see that it is true are the same thing. Carnap still suggested, however, that a protocol statement records a private experience, even though every such statement—indeed every statement—can be translated into the public language of physics. Statements of the form "Here now an experience of red" can, he argued, be

translated into statements about the physical state of the body of the person who has the experience of red. (Subsequently this "physicalist" thesis was expressed in the weaker form, that every statement is linked by means of correspondence rules with the statements of physics.)

Neurath was still dissatisfied. Protocol statements, he argued, must form part of science as distinct from merely being translatable into its language. Otherwise, science still rests on essentially private experience. In fact, protocol statements must take some such form as "Otto Neurath reports that at 3:15 p.m. there was a table in the room perceived by Otto." The effect of this suggestion, as Schlick remarked with horror, is to leave open the possibility that the basic protocol statements may not be true. They, rather than some natural law with which they are incompatible, can be rejected as false. Schlick persisted in arguing that the ultimate confirmations of scientific propositions must be experiences of the form "here, now, blue"—which he described as "the only synthetic statements which are not hypotheses." Carnap came to agree with Neurath, however, that all synthetic statements are hypotheses.

At first, indeed, Carnap replied to Neurath by invoking his principle of tolerance. One has a free choice, he argued, between a language that incorporates protocol statements and a language into which they can be translated. Subsequently he has moved more and more in Neurath's direction. Statements of the form "the body Carnap is in a state of green-seeing," he now suggests, are sufficient to act as confirmations, and it is not necessary at any point to use the "phenomenal language" that Mach had thought to be the basic language of science. But Carnap still writes as if the issue between physicalist and nonphysicalist hinges on the choice of a language. Logical positivism, we might say, split into three groups, one asserting physicalism, the second rejecting it, and the third expressing a preference for the physicalist language.

In his *Logical Syntax of Language* Carnap had argued that all statements about the "meaning" or "significance" of statements are of the "pseudo object" type and should be translated into a syntactical form. Thus, for example, "This letter is about the son of Mr. Miller" has to be read as asserting that in this letter a sentence occurs which has the expression "the son of Mr. Miller" as its subject. This was a highly implausible doctrine, since, clearly, a letter can be about the son of Mr. Miller without using the phrase "the son of Mr. Miller." Under Tarski's influence Carnap decided that his original thesis had been unduly restrictive; philosophy had to refer to the semantical as well as the syntactical characteristics of language in order to give a satisfactory explication of, for example, the conception of "truth." Now Carnap found himself in opposition to Neurath. To try to pass beyond language to what language signifies, Neurath argued, is at once to reintroduce the transcendental entities of metaphysics. The subsequent development of semantics at Carnap's hands would have done nothing to relieve Neurath's qualms. Languages can be constructed, Carnap argues, in a variety of ways, and the question whether, for example, one accepts a language that includes names for abstract entities is a matter of practical convenience, not admitting of argument at any other level. The influence of Mach on Carnap's thinking has now been almost entirely dissipated; he writes, rather, in the spirit of a Poincaré or a Duhem.

THE INFLUENCE OF POSITIVISM

Logical positivism, considered as the doctrine of a sect, has disintegrated. In various ways it has been absorbed into the international movement of contemporary empiricism, within which the disputes that divided it are still being fought out. Originally, it set up a series of sharp contrasts: between metaphysics and science, logical and factual truths, the verifiable and the nonverifiable, the corrigible and the incorrigible, what can be shown and what can be said, facts and theories. In recent philosophy, all these contrasts have come under attack, not from metaphysicians but from philosophers who would in a general sense be happy enough to describe themselves as "logical empiricists." Even among those philosophers who would still wish to make the contrasts on which the logical positivists insisted, few would believe that they can be made with the sharpness or the ease that the logical positivists at first suggested.

Logical positivism, then, is dead, or as dead as a philosophical movement ever becomes. But it has left a legacy behind. In the German-speaking countries, indeed, it wholly failed; German philosophy, as exhibited in the works of Martin Heidegger and his disciples, represents everything to which the positivists were most bitterly opposed. In the United States, Great Britain, Australia, the Scandinavian countries, and in other countries where empiricism is widespread, it is often hard to distinguish the direct influence of the positivists from the influence of such allied philosophers as Russell, the Polish logicians, and the British "analysts." But insofar as it is widely agreed that transcendental metaphysics, if not meaningless, is at least otiose, that philosophers ought to set an example of precision and clarity, that philosophy should make use of technical devices, deriving from logic,

in order to solve problems relating to the philosophy of science, that philosophy is not about "the world" but about the language through which men speak about the world, we can detect in contemporary philosophy, at least, the persistence of the spirit that inspired the Vienna circle.

See also Absolute, The; Analysis, Philosophical; Ayer, Alfred Jules; Basic Statements; Carnap, Rudolf; Duhem, Pierre Maurice Marie; Einstein, Albert; Emotive Theory of Ethics; Empiricism; Epicurus; Epistemology, History of; Frege, Gottlob; Gödel, Kurt; Heidegger, Martin; Helmholtz, Hermann Ludwig von; Hempel, Carl Gustav; Hume, David; Kant, Immanuel; Language; Leibniz, Gottfried Wilhelm; Mach, Ernst; Metaphysics; Mill, John Stuart; Neo-Kantianism; Neurath, Otto; Planck, Max; Poincaré, Jules Henri; Popper, Karl Raimund; Positivism; Ramsey, Frank Plumpton; Reichenbach, Hans; Russell, Bertrand Arthur William; Ryle, Gilbert; Schlick, Moritz; Tarski, Alfred; Utilitarianism; Verifiability Principle; Wittgenstein, Ludwig Josef Johann.

Bibliography

The essays by leading logical positivists included along with an introduction by Ayer and an extensive bibliography in Alfred J. Ayer, ed., *Logical Positivism* (Glencoe, IL: Free Press, 1959), are the best introduction to the movement as a whole. Other representative writings by members of the Vienna circle and its associates include the following: Alfred J. Ayer, *Language, Truth and Logic* (London: Gollancz, 1936; 2nd ed., 1946). Rudolf Carnap, *Der Logische Aufbau der Welt* (Berlin: Weltkreis, 1928), translated by R. George as *The Logical Structure of the World* (London: Routledge and Kegan Paul, 1967); "Die physikalische Sprache als Universalsprache der Wissenschaft," in *Erkenntnis* 2 (5–6) (1932): 423–465, translated by Max Black as *The Unity of Science* (London: Kegan Paul, 1934); *Logische Syntax der Sprache* (Vienna: Springer, 1934), translated by Amethe Smeaton as *The Logical Syntax of Language* (London and New York: Kegan Paul, 1937); *Philosophy and Logical Syntax* (London: Kegan Paul, 1935); "Testability and Meaning," in *Philosophy and Science* 3 (4) (1936): 419–471, and 4 (1) (1937): 1–40; *Meaning and Necessity* (Chicago: University of Chicago Press, 1947); "Empiricism, Semantics and Ontology," in *Revue internationale de philosophie* 4 (11) (1950), reprinted in *Semantics and the Philosophy of Language,* edited by Leonard Linsky (Urbana: University of Illinois Press, 1952), pp. 208–228; and "Intellectual Autobiography" and "Reply to My Critics," in *The Philosophy of Rudolf Carnap,* edited by P. A. Schilpp (La Salle, IL: Open Court, 1963).

See Herbert Feigl and May Brodbeck, eds., *Readings in the Philosophy of Science* (New York: Appleton-Century-Crofts, 1953), and Herbert Feigl and Wilfrid Sellars, eds., *Readings in Philosophical Analysis* (New York: Appleton-Century-Crofts, 1949), for articles by Carnap, Feigl, Reichenbach, Hempel, Frank, and Zilsel. See also Philipp Frank, *Modern Science and Its Philosophy* (Cambridge, MA: Harvard University Press, 1949); Victor Kraft, *Der Wiener Kreis, Der Ursprung des Neopositivismus* (Vienna: Springer-Verlag, 1950), translated by Arthur Pap as *The Vienna Circle* (New York: Philosophical Library, 1953); Jørgen Jørgensen, *The Development of Logical Empiricism* (Chicago: University of Chicago Press, 1951); Charles W. Morris, *Logical Positivism, Pragmatism and Scientific Empiricism* (Paris: Hermann, 1937); Otto Neurath, *Einheitswissenschaft und Psychologie* (Vienna: Gerold, 1933) and *Le développement du Cercle de Vienne et l'avenir de l'empiricisme logique* (Paris: Hermann, 1935); Hans Reichenbach, *The Rise of Scientific Philosophy* (Berkeley: University of California Press, 1951); Karl Raimund Popper, *Logik der Forschung* (Vienna: Springer, 1935), translated with new appendices by the author, with the assistance of Julius Freed and Lan Freed, as *Logic of Scientific Discovery* (London: Hutchinson, 1959); Moritz Schlick, *Allgemeine Erkenntnislehre* (Berlin: Springer, 1918), *Fragen der Ethik* (Vienna: Springer, 1930), translated by D. Rynin as *Problems of Ethics* (New York: Prentice Hall, 1939), and *Gesammelte Aufsätze* (Vienna: Gerold, 1938), partially republished as *Gesetz Kausalität und Wahrscheinlichkeit* (Vienna: Gerold, 1948); and Ludwig Wittgenstein, *Tractatus Logico-Philosophicus,* German version as *Logisch-Philosophische Abhandlung* in *Annalen der Naturphilosophie* 14 (1921): 185–262, English translation by C. K. Ogden (London: Routledge, 1922; rev. ed., New York: Humanities Press, 1961); *Philosophical Investigations* (Oxford: Blackwell, 1953).

See also Richard von Mises, *Kleines Lehrbuch der Positivismus* (The Hague: W. P. van Stockum, 1939), translated as *Positivism: A Study in Human Understanding* (Cambridge, MA: Harvard University Press, 1951); G. Bergmann, *The Metaphysics of Logical Positivism* (London: Longmans Green, 1954); Frederick C. Copleston, *Contemporary Philosophy* (London: Burns and Oates, 1956); Ernest Nagel, *Logic without Metaphysics* (Glencoe, IL: Free Press, 1956); John Arthur Passmore, "Logical Positivism," in *Australasian Journal of Psychology and Philosophy* 21 (1943): 65–92, 33 (1944): 129–153, and 26 (1948): 1–19, and *A Hundred Years of Philosophy* (London: Duckworth, 1957); James O. Urmson, *Philosophical Analysis* (Oxford: Clarendon Press, 1956); and Julius R. Weinberg, *An Examination of Logical Positivism* (London: Kegan Paul, 1936).

John Passmore (1967)

LOGICAL TERMS

The two central problems concerning "logical terms" are demarcation and interpretation. The search for a demarcation of logical terms goes back to the founders of modern logic, and within the classical tradition a partial solution, restricted to logical connectives, was established early on. The characteristic feature of logical connectives, according to this solution, is truth-functionality, and the

totality of truth functions (Boolean functions from n-tuples of truth values to a truth value) determines the totality of logical connectives. In his seminal 1936 paper, "On the Concept of Logical Consequence," Alfred Tarski demonstrated the need for a more comprehensive criterion by showing that his semantic definition of logical consequence—the sentence σ is a logical consequence of the set of sentences Σ iff (if and only if) every model of Σ is a model of σ—is dependent on such a demarcation. (Thus suppose the existential quantifier is not a logical term, then its interpretation will vary from model to model, and the intuitively logically valid consequence, "Rembrandt is a painter; therefore there is at least one painter," will fail to satisfy Tarski's definition. Suppose "Rembrandt" and "is a painter" are both logical terms, then the intuitively logically invalid consequence, "Frege is a logician; therefore Rembrandt is a painter," will satisfy Tarski's definition.) Tarski, however, left the general demarcation of logical terms an open question, and it was not until the late 1950s that the first steps toward developing a systematic criterion for logical predicates and quantifiers were taken.

In his 1957 paper, "On a Generalization of Quantifiers," A. Mostowski proposed a semantic criterion for first-order logical quantifiers that generalizes Frege's analysis of the standard quantifiers as second-level cardinality predicates. Technically, Mostowski interpreted a quantifier, Q, as a function from universes (sets of objects), A, to A-quantifiers, Q_A, where Q_A is a function assigning a truth-value to each subset B of A. Thus, given a set A, the existential and universal quantifiers are defined by: for any $B \subseteq A, \exists_A(B) = T$ iff $B \neq \phi$ and $\forall_A(B) = T$ iff $A - B = \phi$. Intuitively, a quantifier is logical if it does "not allow us to distinguish between different elements" of the underlying universe. Formally, Q is logical iff it is invariant under isomorphic structures of the type $<A,B>$, where $B \subseteq A$; that is, Q is a logical quantifier iff for every structure $<A,B>$ and $<A',B'>$:if$<A,B> \cong <A',B'>$, then $Q_A(B) = Q_{A'}(B')$. Quantifiers satisfying Mostowski's criterion are commonly called *cardinality quantifiers,* and some examples of these are "$!\delta x$" ("There are exactly δ individuals in the universe such that …"), where δ is any cardinal, "Most x" ("There are more x's such that … than x's such that not …"), "There are finitely many x," "There are uncountably many x," and so forth.

In 1966, P. Lindström extended Mostowski's criterion to terms in general: A term (of type n) is logical iff it is invariant under isomorphic structures (of type n). Thus, the well-ordering predicate, W, is logical since for any A,A', $R \subseteq A^2$ and $R' \subseteq A'^2$: if $<A,R> \cong <A',R'>$, then $W_A(R) = W_{A'}(R')$. Intuitively, we can say that a term is logical iff it does not distinguish between isomorphic arguments. The terms satisfying Lindström's criterion include identity, n-place cardinality quantifiers (e.g., the 2-place "Most," as in "Most A's are B's"), relational or polyadic quantifiers like the well-ordering predicate above and "is an equivalence relation," and so forth. Among the terms not satisfying Lindström's criterion are individual constants, the first-level predicate "is red," the first-level membership relation, the second-level predicate "is a property of Napoleon," and so forth. Tarski (1966) proposed essentially the same division.

The Mostowski-Lindström-Tarski (MLT) approach to logical terms has had a considerable impact on the development of contemporary model theory. Among the central results are Lindström's characterizations of elementary logic, various completeness and incompleteness theorems for generalized (model-theoretic, abstract) logics, and so forth. (See Barwise and Feferman 1985). But whereas the mathematical yield of MLT has been prodigious, philosophers, by and large, have continued to hold on to the traditional view according to which the collection of (primitive) logical terms is restricted to truth-functional connectives, the existential and/or universal quantifier and, possibly, identity. One of the main strongholds of the traditional approach has been Willard Van Orman Quine, who (in his 1970 book) justified his approach on the grounds that (1) standard first-order logic (without identity) allows a remarkable concurrence of diverse definitions of logical consequence, and (2) standard first-order logic (with or without identity) is complete. Quine did not consider the logicality of non-standard quantifiers such as "there are uncountably many," which allow a "complete" axiomatization. L. H. Tharp (1975), who did take into account the existence of complete first-order logics with nonstandard generalized quantifiers, nevertheless arrived at the same conclusion as Quine's.

During the 1960s and 1970s many philosophers were concerned with the interpretation rather than the identity of logical terms. Thus, Ruth Barcan Marcus (1962, 1972) and others developed a substitutional interpretation of the standard quantifiers; Michael Anthony Eardley Dummett (1973) advocated an intuitionistic interpretation of the standard logical terms based on considerations pertaining to the theory of meaning; many philosophers (e.g., van Fraassen) pursued "free" and "many-valued" interpretations of the logical connectives; Jaako Hintikka (1973, 1976) constructed a game theoretic semantics for logical terms. In a later development, G.

Boolos (1984) proposed a primitive (non-set-theoretic) interpretation of "nonfirstorderizable" operators, which has the potential of overcoming ontological objections to higher-order logical operators (e.g., by Quine).

In the mid-1970s philosophers began to search for an explicit, general philosophical criterion for logical terms. The attempts vary considerably, but in all cases the criterion is motivated by an underlying notion of logical consequence. Inspired by Gerhard Gentzen's proof-theoretic work, Ian Hacking (1979) suggests that a logical constant is introduced by (operational) rules of inference that preserve the basic features of the traditional deducibility relation: the subformula property (compositionality), reflexivity, dilution (stability under additional premises and conclusions), transitivity (cut), cut elimination, and so forth. Hacking's criterion renders all and only the logical terms of the ramified theory of types genuinely logical. A. Koslow's (1992) also utilizes a Gentzen-like characterization of the deducibility relation. Abstracting from the syntactic nature of Gentzen's rules, he arrives at a "structural" characterization of the standard logical and modal constants. Both Koslow and Hacking incorporate lessons from an earlier exchange between A. N. Prior (1960, 1964) and N. Belnap (1962) concerning the possibility of importing an inconsistency into a hitherto consistent system by using arbitrary rules of inference to introduce new logical operators.

C. Peacocke (1976) approaches the task of delineating the logical terms from a semantic perspective. The basic property of logical consequence is, according to Peacocke, a priori. α is a logical operator iff α is a noncomplex n-place operator such that given knowledge of which objects (sequences of objects) satisfy an n-tuple or arguments of α, $<\beta_1,\ldots,\beta_n>$, one can know a priori which objects satisfy $\alpha(\beta_1,\ldots,\beta_n)$. Based on this criterion Peacocke counts the truth-functional connectives, the standard quantifiers, and certain temporal operators ("In the past …") as logical, while identity (taken as a primitive term), the first-order membership relation, and "necessarily" are nonlogical. Peacocke's criterion is designed for classical logic, but it is possible to produce analogous criteria for nonclassical logics (e.g., intuitionistic logic). T. McCarthy (1981) regards the basic property of logical constants as topic neutrality. He considers Peacocke's condition as necessary but not sufficient, and his own criterion conjoins Peacocke's condition with Lindström's invariance condition (MLT). The standard first-order logical vocabulary as well as various nonstandard generalized quantifiers satisfy McCarthy's criterion, but cardinality quantifiers do not (intuitively, cardinality quantifiers are not topic-neutral).

Sher (1991) considers necessity and formality as the two characteristic features of logical consequence. Treating formality as a semantic notion, Sher suggests that any formal operator incorporated into a Tarskian system according to certain rules yields consequences possessing the desired characteristics. Viewing Lindström's invariance criterion as capturing the intended notion of formal operator, Sher endorses the full-fledged MLT as delineating the scope of logical terms in classical logic.

The theory of logical terms satisfying Lindström's criterion has led, with various adjustments, to important developments in linguistic theory: a systematic account of determiners as generalized quantifiers (Barwise and Cooper, Higginbotham and May); numerous applications of "polyadic" quantifiers (van Benthem, Keenan); and an extension of Henkin's 1961 theory of standard branching quantifiers, applied to English by Hintikka (1973), to branching generalized quantifiers (Barwise and others).

See also Dummett, Michael Anthony Eardley; Frege, Gottlob; Hintikka, Jaako; Logic, History of; Marcus, Ruth Barcan; Model Theory; Prior, Arthur Norman; Quine, Willard Van Orman; Tarski, Alfred.

Bibliography

Barwise, J. "On Branching Quantifiers in English." *Journal of Philosophical Logic* 8 (1979): 47–80.

Barwise, J., and R. Cooper. "Generalized Quantifiers and Natural Language." *Linguistics and Philosophy* 4 (1981): 159–219.

Barwise, J., and S. Feferman, eds. *Model-Theoretic Logics.* New York: Springer-Verlag, 1985.

Belnap, N. "Tonk, Plonk, and Plink." *Analysis* 22 (1962): 130–134.

Boolos, G. "To Be Is to Be a Value of a Variable (or to Be Some Values of Some Variables)." *Journal of Philosophy* 81 (1984): 430–449.

Dummett, M. "The Philosophical Basis of Intuitionistic Logic." In *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978.

Hacking, I. "What Is Logic?" *Journal of Philosophy* 76 (1979): 285–319.

Henkin, L. "Some Remarks on Infinitely Long Formulas." In *Infinitistic Methods.* Warsaw, 1961.

Higginbotham, J., and R. May. "Questions, Quantifiers and Crossing." *Linguistic Review* 1 (1981): 41–79.

Hintikka, J. "Quantifiers in Logic and Quantifiers in Natural Languages." In *Philosophy of Logic,* edited by S. Körner, 208–232. Oxford, 1976.

Hintikka, J. "Quantifiers vs. Quantification Theory." *Dialectica* 27 (1973): 329–358.

Keenan, E. L. "Unreducible n-ary Quantifiers in Natural Language." In *Generalized Quantifiers,* edited by P. Gärdenfors. Dordrecht: Reidel, 1987.

Koslow, A. *A Structuralist Theory of Logic.* Cambridge, U.K.: Cambridge University Press, 1992.

Lindström, P. "First Order Predicate Logic with Generalized Quantifiers." *Theoria* 32 (1966): 186–195.

Marcus, R. Barcan. "Interpreting Quantification." *Inquiry* 5 (1962): 252–259.

Marcus, R. Barcan. "Quantification and Ontology." *Noûs* 6 (1972): 240–250.

McCarthy, T. "The Idea of a Logical Constant." *Journal of Philosophy* 78 (1981): 499–523.

Mostowski, A. "On a Generalization of Quantifiers." *Fundamenta Mathematicae* 42 (1957): 12–36.

Peacocke, C. "What Is a Logical Constant?" *Journal of Philosophy* 73 (1976): 221–240.

Prior, A. N. "Conjunction and Contonktion Revisited." *Analysis* 24 (1964): 191–195.

Prior, A. N. "The Runabout Inference-Ticket." *Analysis* 21 (1960): 38–39.

Quine, W. V. O. *Philosophy of Logic.* Englewood Cliffs, NJ: Prentice-Hall, 1970.

Sher, G. *The Bounds of Logic: A Generalized Viewpoint.* Cambridge, MA: MIT Press, 1991.

Tarski, A. *Logic, Semantics, Metamathematics,* 2nd ed. Indianapolis: Hackett, 1983.

Tarski, A. "On the Concept of Logical Consequence" (1936). In *Logic, Semantics, Metamathematics,* 2nd ed., 409–420. Indianapolis: Hackett, 1983.

Tarski, A. "What Are Logical Notions?" (1966). *History and Philosophy of Logic* 7 (1986): 143–154.

Tharp, L. H. "Which Logic Is the Right Logic?" *Synthese* 31 (1975): 1–21.

van Benthem, J. "Polyadic Quantifiers." *Linguistics and Philosophy* 12 (1989): 437–464.

van Fraassen, B. C. "Singular Terms, Truth-Value Gaps, and Free Logic." *Journal of Philosophy* 63 (1966): 481–495.

Gila Sher (1996)

LOGICAL TERMS, GLOSSARY OF

This glossary is confined, with few exceptions, to terms used in formal logic, set theory, and related areas. No attempt has been made to cover what is often called "inductive logic," although several terms in this field have been included for the convenience of the reader.

It should be noted that many topics dealt with very briefly here are treated in full in various other entries in this encyclopedia. Cross references to these will be enclosed in quotation marks; cross references to other glossary entries will be indicated by boldface italics (e.g., "see ***relation***").

abduction. (1) A syllogism whose major premise is known to be true but whose minor premise is merely probable. (2) C. S. Peirce's name for the type of reasoning that yields from a given set of facts an explanatory hypothesis for them.

abstraction. (1) In traditional logic, the process of deriving a universal from particulars. (2) In set theory, the process of defining a set as the set of all objects that have a particular property.

abstraction, axiom of (axiom of comprehension). An axiom in set theory stating that for any predicate *P,* there exists a set of all and only those objects that satisfy *P.* It was the unrestricted use of this axiom that led to the paradoxes of set theory.

abstract term. In traditional logic, a term that is a name of the common nature of many individuals, considered apart from them or from what distinguishes them from one another. A common example of an abstract term is "humanity."

accident. See ***predicables.***

actual infinite. The infinite regarded as a completed whole.

a fortiori. A nonsyllogistic mediate inference of the form "*B* is greater than *C; A* is greater than *B*; hence, *A* is greater than *C.*" It is clear that the validity of this argument follows from the transitivity of the relation "greater than," and therefore some authors extend the term to cover all relational syllogisms whose validity depends on the transitivity of the relation involved. See ***relation.***

aggregate. A collection of objects satisfying a given condition.

alephs. The symbols, introduced by Georg Cantor, that designate the cardinality of infinite sets (see entry "Set Theory"). *Aleph-null* ($\aleph_0$) designates the cardinality of the smallest infinite set, aleph-one ($\aleph_1$) the cardinality of the next largest infinite set, etc. See ***continuum hypothesis***; entry "Set Theory."

algebra of logic. A system in which algebraic formulas are used to express logical relations. In such a system many familiar algebraic laws that hold for numbers are not retained. The work of George Boole contains the first important example of an algebra of logic.

algorithm. A mechanical procedure for carrying out, in a finite number of steps, a computation that leads from certain types of data to certain types of results. See ***decision problem; effectiveness.***

alternation. See ***disjunction, exclusive.***

alternative denial. See ***Sheffer stroke function.***

ambiguity. Capability of being understood in two or more ways. The term is strictly applied only in cases where the possibility of different interpretation is due not to the expression itself but to some feature of the particular use of the expression; when this possibility is due to the expression itself the expression is called *equivocal.* Many authors, however, do not make this distinction.

amphiboly. An equivocation that arises not out of an equivocation in a word or phrase but because the grammatical structure of the sentence or clause leaves the place of the phrase in the whole not entirely determinate. An example is "The shooting of the hunters was finished quickly."

ampliation. In medieval logic, the extension of a common term from a narrow supposition to a wider one.

analogy. A comparison between two or more objects that indicates one or more respects in which they are similar. An *argument from analogy* is an inference from some points of resemblance between two or more objects to other such points. The method of *refutation by logical analogy* is a method for showing that an argument is fallacious by giving an example of another argument of the same form whose invalidity is immediately apparent.

analysis, mathematical. The theory of real and complex numbers and their functions.

analytic. Used of a proposition whose denial is self-contradictory. Such a proposition is true either by virtue of its logical form alone (in which case it is called a *logical truth,* or *logically necessary*) or by virtue of both its logical form and the meaning of its constituent terms. An instance of a logical truth is "It is raining or it is not raining"; an example of an analytic truth that is not a logical truth is "All bachelors are unmarried." Analytic propositions cannot be false and are therefore said to be *necessary truths.* Whether there are necessary truths that are not also analytic truths is a matter of much dispute. See entry "Analytic and Synthetic Statements."

ancestral relation. For a given relation R, the relation R^* that exists between two objects x and y if and only if y has every R-hereditary property that x has. A property is said to be *R-hereditary* when, if it is correctly predicated of b and if aRb, then it is also correctly predicated of a. For example, let R be the property "is the successor of." Then "is a natural number" (where this property also applies to 0) is R-hereditary, since if b is a natural number and a is the successor of b, then a is also a natural number. Given this fact, we can define the property "is a natural number" as the property of all objects that bear the ancestral relation to 0 for the relation "is the successor of"—that is, as the property of all objects that have every "is the successor of"-hereditary property that 0 has. One of these properties is "is a natural number," and therefore only the natural numbers can meet this definition.

It should be noted that the above definition is an example of an *impredicative definition,* since "is a natural number" is defined in terms of the class of "is the successor of"-hereditary properties, a class of which it is a member.

antecedent. The part of a hypothetical proposition that precedes the implication sign.

antilogism. A triad of propositions such that the joint truth of any two of the propositions implies the falsity of the third. Christine Ladd-Franklin's principle of the syllogism states that a valid syllogism is one whose premises taken with the contradictory of the conclusion constitute an antilogism. Thus, the syllogism whose premises are "All men are mortal" and "Socrates is a man" and whose conclusion is "Socrates is mortal" is a valid syllogism, for the joint assertion of any two of the three propositions that constitute the premises and the contradictory of the conclusion implies the falsity of the third proposition.

antinomy. See ***paradox.***

apodictic (apodeictic) proposition. See ***modality.***

appellation. In medieval logic a term is said to have appellation if it is applicable to some existing thing. Thus, "the present queen of England" has appellation, but "the present queen of the United States" does not.

A-proposition. In traditional logic, a universal affirmative categorical proposition. An example is "All men are mortal."

Archimedean property. The property of a system of numbers whereby for any two numbers a and b, if a is less than b, then there is a number c such that a multiplied by c is greater than b.

argument of a function. A member of the domain of a given function.

arithmetical predicate. A predicate that can be explicitly expressed in terms of the truth-functional connectives of propositional calculus, the universal and existential quantifiers, constant and variable natural numbers, and the addition and multiplication functions.

arithmetization of mathematics (arithmetization of analysis). The definition, which was developed by Karl Weierstrass, Richard Dedekind, and Georg Cantor, of the nonnatural numbers as certain objects construed out of

the natural numbers and set-theoretic objects and the corresponding reduction of the properties of the former to the properties of the latter.

arithmetization of syntax. The process of correlating the objects of a formal system with some or all of the natural numbers and then studying the relations and properties of the correlated numbers so as to gain information about the syntax of the formal system. This was done systematically by Kurt Gödel in the researches that led to his incompleteness theorems. See entry "Gödel's Theorem."

ars combinatoria. A technique of deriving complex concepts by the combination of relatively few simple ones, which are taken as primitive. This technique was proposed by Gottfried Wilhelm Leibniz as a valuable aid for the study of all subjects. He proposed the development of a universal language (*characteristica universalis*) containing a few primitive symbols in terms of which all other symbols would be defined. A universal mathematics (*mathesis universalis*)—that is, a universal system of reasoning—would then be added, and all subjects could be studied in this language. Leibniz program is often viewed as an early forerunner of the formalization of various disciplines.

assertion sign. The sign ⊦, introduced by Gottlob Frege to indicate in the object language that a proposition is being judged as true and is not merely being named. Some authors now use this sign in the metalanguage to express that the formula to which it is prefixed is a theorem in the object language.

assertoric proposition. See ***modality.***

associativity. The property of a relation *R* that consists in the identity of "$aR(bRc)$" and "$(aRb)Rc$," where *a, b,* and *c* are any elements of the field of *R*. Addition has this property, since "$a + (b + c)$" is the same as "$(a + b) + c$."

attribute. Although it is now often used synonymously with "property," this term was traditionally confined to the essential characteristics of a being.

Aussonderungsaxiom. An axiom in set theory, first introduced by Ernst Zermelo, which states that for any set *a* and any predicate *P,* there exists a set containing all and only those members of *a* that satisfy the predicate *P.*

axiom. A basic proposition in a formal system that is asserted without proof and from which, together with the other such propositions, all other theorems are derived according to the rules of inference of the system. See ***postulate.***

axiomatic method. The method of studying a subject by beginning with a list of undefined terms and a list of axioms and then deriving the truths of the subject from these postulates by the methods of formal logic.

axiom schema. A representation of an infinite number of axioms by means of an expression containing syntactical variables and having well-formed formulas as values. Every value of the expression is to be taken as an axiom.

axiom schema of separation. See ***Aussonderungsaxiom.***

Barbara. See ***mnemonic terms.***

Baroco. See ***mnemonic terms.***

biconditional. A binary propositional connective (↔, ≡), usually read "if and only if" (often abbreviated "iff"), whose truth table is such that "*A* if and only if *B*" is true when *A* and *B* are either both true or both false and is false when one is true and the other false. "*A* if and only if *B*" is equivalent to "if *A* then *B,* and if *B* then *A.*"

binary connective. See ***connective.***

Bocardo. See ***mnemonic terms.***

Boolean algebra. The first algebra of logic. It was invented by George Boole and given its definitive form by Ernst Schröder.

Boolean functions. Functions that occur in Boolean algebra. The more important ones are the class-union function, the class-intersection function, and the class-complement function.

bound occurrence of a variable. An occurrence of a variable *a* in a well-formed part of a formula *A* either of the form "for all *a, B*" or of the form "there is an *a* such that *B.*"

bound of a set. For a given relation *R,* a *lower* bound (or first element) of a set *a* is any member of *a* that bears the relation *R* to all members of *a*; an upper bound of *a* is any member of *a* to which all members of *a* bear the relation *R.* A *greatest lower bound* of a set *a* (or *infimum* of *a*) is a lower bound of *a* to which all lower bounds of *a* bear the relation *R*; a *least upper bound* of *a* (or *supremum* of *a*) is an upper bound of *a* that bears the relation *R* to all upper bounds of *a.*

bound variable. A bound variable of a formula *A* is a variable that has a bound occurrence in *A.*

Bramantip. See ***mnemonic terms.***

Burali-Forti's paradox. See ***paradox.***

calculus. Any logistic system. The two most important types of logical calculi are *propositional* (or senten-

tial) calculi and *functional* (or predicate) calculi. A propositional calculus is a system containing propositional variables and connectives (some also contain propositional constants) but not individual or functional variables or constants. In the *extended* propositional calculus, quantifiers whose operator variables are propositional variables are added. Among the *partial* propositional calculi, in which not all the theorems of the standard propositional calculus are obtainable, the most important are David Hilbert's *positive* propositional calculus (this contains all those parts of the standard propositional calculus that are independent of negation) and the *intuitionistic* propositional calculus (in this system axioms about negation acceptable from the intuitionistic point of view are added to the positive propositional calculus). A functional calculus is a system containing, in addition to the symbols of propositional calculus, individual and functional variables and/or constants, as well as quantifiers that take some of these variables and constants as their operator variables. In a *first-order* functional calculus (or *first-order logic*) the quantifiers have as their operator variables only individual variables, and the functions have as their arguments only individual variables and/or constants. In a *second-order* functional calculus (or second-order logic) the operator variables of the quantifiers can be functional variables. After that, each odd order adds functional variables and/or constants some of whose arguments are of the type introduced two orders below, and each even order allows the use of the variables introduced one order below as operator variables for the quantifiers. When there are no individual or functional constants present the functional calculus is called *pure*; when either is present it is called applied.

Camenes. See ***mnemonic terms.***

Camestres. See ***mnemonic terms.***

Cantor's paradox. See ***paradox.***

Cantor's theorem. The theorem stating that for any given set *a,* the power set of *a* has a greater cardinality than *a* has.

cardinality (power). For a given set, the cardinal number associated with it.

cardinal number. An object *a* that is associated with all and only the members of a set of equipollent sets. Various authors disagree on what this object is. The *Frege-Russell definition* of cardinal number is simply the identification of *a* with the set of equipollent sets.

Cartesian product. For a given set *a,* the set whose members are all and only the sets that contain one member from each member of *a.*

categorematic. In traditional logic, used of a word that can be a term in a categorical proposition. In contemporary logic, used of any symbol that has independent meaning. An example of a categorematic word is "men." Cf. ***syncategorematic.***

categorical proposition. See ***proposition.***

category. A general or fundamental class of objects or concepts about whose members assertions can significantly be made which differ from those that can significantly be made about nonmembers of this class. The two most famous lists of categories are those of Aristotle and Immanuel Kant. Aristotle's list comprises substance, quantity, quality, relation, activity, passivity, place, time, situation, and state. Kant's comprises unity, plurality, and universality (categories of quantity); reality, negation, and limitation (categories of quality); substantiality, causality, and reciprocity (categories of relation); and possibility, actuality, and necessity (categories of modality).

Celarent. See ***mnemonic terms.***

Cesare. See ***mnemonic terms.***

choice, axiom of (multiplicative axiom). An axiom in set theory stating that if *a* is a disjoint set which does not have the null set as one of its members, then the Cartesian product of *a* is different from the null set. It can be proved that this axiom is equivalent to the well-ordering theorem.

choice function. A function *R* whose domain includes (or, according to some authors, is identified with the set of) all the nonempty subsets of a given set *a* and whose value is a member of any such subset.

Church's theorem. The theorem, stated and proved by Alonzo Church, that there is no decision procedure for determining whether or not an arbitrary well-formed formula of the first-order functional calculus is a theorem of that system.

Church's thesis. The thesis that every effectively calculable function (effectively decidable predicate) is general recursive.

circular reasoning. See ***fallacy.***

class. (1) An aggregate. (2) In Gödel–von Neumann–Bernays set theory, where a distinction is made between sets and classes, a class is an object that can contain members but cannot be a member of any object. See ***set.***

classification. Two of the issues of concern to traditional logicians were the nature of the process of grouping individuals into classes of individuals (*species*), these classes into further classes, and so on (the process of classification), and the nature of the reverse process (the process of *division*)—breaking a class down into its subclasses, these into their subclasses, and so on, until the simplest classes are broken down into the individuals that are their members.

In the process of classification one begins with a group of individuals and arranges them into classes, called *infimae species,* none of which can be broken down into species but only into individuals. One then groups the *infimae species* into other classes, of which the *infimae species* are subclasses. (For any species the class of which it is a subclass is called the *proximum genus.*) The grouping continues until one reaches the class of which all the original individuals are members. This is the *summum genus,* and when one reaches it the process of classification is finished. (All the classes between the *infimae species* and the *summum genus* are called the *subaltern genera.*)

In the process of division one begins with the *summum genus* and breaks it down into its subclasses, continuing until one reaches the *infimae species.* Finally, these are broken down into the individuals that are their members.

Several rules were set up for classification and division: (1) at each step only one principle may be used for breaking down the classes or grouping them together; (2) no group may be omitted at any step; (3) no intermediate step may be omitted. When applied to division this last rule is known as the rule of *division non faciat saltum.*

A *dichotomy* is a form of division (or of classification) in which at each stage the genus is divided into species according to whether or not the objects possess a certain set of differentiae. The two species formed (*proxima genera*) are therefore mutually exclusive and jointly exhaustive.

closed sentence (closed schema). A sentence (or schema) that has no free variables.

closed with respect to (closed under) a relation. A set is closed under a relation *R* if and only if for all *a,* if *aRb* and if *a* is a member of the set, then *b* is a member of the set.

closure of a formula. A formula formed by placing before an original formula *A* quantifiers binding all variables that occur freely in *A*. A *universal* closure is the formula formed when only universal quantifiers are used, and an *existential* closure is the formula formed when only existential quantifiers are used.

collective term. In traditional logic, a term that denotes a collection of objects regarded as a unity. An example is "the Rockies."

combinatory logic. A branch of mathematical logic where variables are entirely eliminated, their place being taken by certain types of functions that are unique to this branch of logic.

commutativity. The property of a relation *R* that consists in the equivalence of *aRb* and *bRa,* where *a* and *b* are any elements of the field of *R.*

comparability, law of (law of trichotomy). The principle in set theory that the cardinality of two sets is always comparable; that is, for any two sets *a* and *b, a* is greater than *b* or equal to *b* or less than *b.*

complement of a set (negate of a set). The set of all and only those objects that are not members of a given set *a.*

completeness. The word *completeness* is used in varying senses. In the strongest sense (E. L. Post) a logistic system is said to be complete if and only if for any well-formed formula *A,* either *A* is a theorem of the system or the system would become inconsistent upon the addition of *A* as an axiom (without any other changes); in this sense propositional calculus, but not pure first-order functional calculus, is complete. In a second, weaker sense (Kurt Gödel) a logistic system is said to be complete if and only if all valid well-formed formulas are theorems of the system; in this sense both propositional calculus and pure first-order functional calculus are also complete. In a third, and still weaker, sense of completeness (Leon Henkin) a logistic system is said to be complete if and only if all secondarily valid well-formed formulas are theorems of the system; in this sense the pure second-order functional calculus and functional calculi of higher order are complete.

complete set. A set all of whose members are subsets of it.

composition, fallacy of. See ***fallacy.***

comprehension, axiom of. See ***abstraction, axiom of.***

computable function. See ***Turing-computable.***

conclusion. That which is inferred from the premises of a given argument.

concrete term. In traditional logic, a term that is the name of an individual or individuals. An example of such a term is "Socrates."

condition. A *necessary condition* is a circumstance in whose absence a given event could not occur or a given thing could not exist. A *sufficient condition* is a circumstance such that whenever it exists a given event occurs or a given thing exists. A *necessary and sufficient condition* for the occurrence of a given event or the existence of a given thing is therefore a circumstance in whose absence the event could not occur or the thing could not exist and which is also such that whenever it exists the event occurs or the thing exists.

This terminology is sometimes extended to the formal relations that exist between propositions. Thus, the truth of a proposition *A* is said to be a necessary condition for the truth of another proposition *B* if *B* implies *A,* and the truth of *A* is said to be a sufficient condition for the truth of *B* if *A* implies *B.*

conditional. See ***implication.***

conditional proof. A proof that begins by making certain assumptions, $A_1, A_2, \cdots, A_n$, deducing *B* from them, and then asserting on the basis of this the truth of the hypothetical proposition "if A_1, then if A_2, then if …, then if A_n, then *B.*" The *rule of conditionalization* is the rule that allows one to make this last step on the basis of the preceding ones.

conjunction. A binary propositional connective (&, **.**), usually read "and," whose truth table is such that "*A* and *B*" is false when *A* or *B* or both are false and is true when both are true.

connective. A symbol that is used with one or more constants or forms to produce a new constant or form. When the constants or forms are propositional ones the connective is known as a *propositional connective* (or *sentential connective*). The most common propositional connectives are negation, conjunction, disjunction, implication, and biconditional. They are classified as *singulary, binary,* etc., according to the number of propositional constants or forms with which they combine.

connotation. See ***meaning, Frege's theory of.***

consequence. Any proposition that can be deduced from a given set of propositions. Thus, given the set of propositions {*A,* if *A* then *B*}, the proposition *B* is a consequence of the set, since it can be deduced from the members of the set by one application of *modus ponens.*

consequent. The part of a hypothetical proposition that follows the implication sign or the "then."

consequentia. The name given by medieval logicians to a true hypothetical proposition. *Formal* consequentiae (those which hold for all substitutions of the categorematic terms) were distinguished from *material* consequentiae (those holding only for particular categorematic terms).

consistency. A set of propositions has consistency (or is consistent) when no contradiction can be derived from the joint assertion of the propositions in the set. A logistic system has consistency when no contradiction can be derived in it. Two syntactical definitions of the consistency of a logistic system are Alfred Tarski's, that a system is consistent if not every well-formed formula is a theorem, and E. L. Post's, that a system is consistent if no well-formed formula consisting of only a propositional variable is a theorem. There is, in addition, a semantical definition of consistency, according to which a set of propositions (or a logistic system) is consistent if there is a model for that set of propositions (or for the set of all the theorems of the system). It must not be assumed that any of these definitions are equivalent; in any case where it is claimed that they are, a proof is required.

constant. A symbol that, under the principal interpretation, is a name for something definite, be it an individual, a property, a relation, etc.

constructive existence proof. A proof of the existence of a mathematical object having a property *P* that gives an example of such an object or at least a method by which one could find such an example.

contingent. Logically possible. See ***logical possibility.***

continuity. An ordered dense class all of whose nonempty subsets which have an upper bound have a least upper bound has continuity (or is continuous). See entry "Continuity."

continuum hypothesis. The hypothesis, proposed by Georg Cantor, that the cardinality of the power set of a set whose cardinality is aleph-null ($\aleph_0$) is aleph-one ($\aleph_1$)—that is, that there is no set whose cardinality is greater than aleph-null but less than the cardinality of the power set of a set whose cardinality is aleph-null. The *generalized continuum hypothesis* is the hypothesis that for the cardinality of any infinite set, the next highest cardinality is the cardinality of its power set.

contradiction. The joint assertion of a proposition and its denial.

contradiction, law of. See ***laws of thought.***

contradictory. Two propositions are contradictory if and only if their joint assertion would be a contradiction. "All men are mortal" and "Some men are not mortal," for example, are contradictory propositions. Two terms are contradictory when they jointly exhaust a universe of dis-

course and are mutually exclusive. In the domain of natural numbers other than 0, for example, "odd" and "even" are contradictory terms. See ***contrary.***

contraposition. In traditional logic, a type of immediate inference in which from a given proposition another proposition is inferred that has as its subject the contradictory of the original predicate. (It should be noted that a change of quality is involved in some cases.) *Partial* contraposition results in a new proposition that is the same as the subject of the original proposition; *full* contraposition results in a predicate of the new proposition that is the contradictory of the subject of the original proposition. The process of contraposition (whether partial or full) yields an equivalent proposition only when the original proposition is an A- or O-proposition; when it is an E-proposition traditional logicians allowed for contraposition *per accidens* (or by limitation)—that is, contraposition plus a change in the quantity of the proposition from universal to particular—claiming that the proposition formed is equivalent to the original proposition. The process of contraposition yields no equivalent proposition when the original proposition is an I-proposition. See entry "Logic, Traditional."

contrary. Applied to two propositions that cannot both be true but can both be false. "All men are mortal" and "No men are mortal," for example, are contrary propositions. Also applied to two terms that are mutually exclusive, but need not be jointly exhaustive, in a universe of discourse. In the domain of natural numbers, for instance, "less than 7" and "more than 19" are contrary terms. See ***contradictory.***

contrary-to-fact (counterfactual) conditional. A conditional proposition whose antecedent is known to be false.

converse domain of a relation (range of a relation). For any relation *R,* the set of all objects *a* such that there exists an object *b* such that *bRa.*

converse of a relation (inverse of a relation). For any relation *R,* the relation R^* such that aR^*b if and only if *bRa.*

conversion. In traditional logic, a type of immediate inference in which from a given proposition another proposition is inferred that has as its subject the predicate of the original proposition and as its predicate the subject of the original proposition (the quality of the proposition being retained). The process of conversion yields an equivalent proposition only when the original proposition is an E- or I-proposition; when it is an A-proposition traditional logicians allowed for conversion *per accidens* (or by limitation)—that is, conversion plus a change in the quantity of the proposition from universal to particular. Thus, the E-proposition "No men are immortal" yields "No immortals are men," but the A-proposition "All men are mortal" can be converted only by limitation, yielding "Some mortals are men." The process of conversion yields no equivalent proposition if the original proposition is an O-proposition. See entry "Logic, Traditional."

copula. In traditional logic, the term that connects the subject and predicate in a categorical proposition. It is always a form of the verb "to be."

corollary. A proposition that follows so obviously from a theorem that it requires little or no demonstration.

counterfactual conditional. See ***contrary-to-fact conditional.***

course-of-values induction. An argument from mathematical induction such that in the induction step one proves that "if the property *P* holds for all numbers before *a,* it holds for *a* as well," where *a* is any number.

Darapti. See ***mnemonic terms.***

Darii. See ***mnemonic terms.***

Datisi. See ***mnemonic terms.***

decision problem. The problem of finding an algorithm (a *decision procedure*) that enables one to arrive, in a finite number of steps, at an answer to any question belonging to a given class of questions. For a logistic system in particular, this is the problem of finding a decision procedure for determining, for any arbitrary well-formed formula of the system, whether or not it is a theorem of the system.

A positive solution to a decision problem consists of a proof that a decision procedure exists. A negative solution to a decision problem consists of a proof that no such procedure is possible. An example of a positive solution is the proof that the truth tables provide a decision procedure for the propositional calculus; an example of a negative proof is Church's theorem.

decision procedure. See ***decision problem.***

Dedekind finite. See ***finite set.***

Dedekind infinite. See ***finite set.***

deducible. A set of propositions is said to be deducible from another set of propositions if and only if there is a valid deductive inference which has the latter set as its premises and the former set as its conclusion.

deduction. A form of inference such that in a valid deductive argument the joint assertion of the premises and the denial of the conclusion is a contradiction.

deduction theorem. For a given logistic system, the metatheorem that states that if there is a proof in the system of A_{n+1} from the assumptions $A_1, A_2, \cdots, A_n$, then there is also a proof in the system of the proposition "if A_n, then A_{n+1}" from the assumptions $A_1, \cdots, A_{n-1}$.

definiendum. That which is defined in a definition.

definiens. That which, in a definition, defines the definiendum.

definite descriptions, theory of. A definite description is a description which, by virtue of the meanings of the words in it, can apply to only one object. A standard example of a definite description is "the author of *Waverley.*" The theory of definite descriptions, introduced by Bertrand Russell, aims at eliminating definite descriptions. Unlike most other eliminative theories, Russell's does not attempt to offer a way of explicitly defining definite descriptions. Instead, it shows how in any given context the description together with the context can be eliminated in such a way that the resulting linguistic expression is equivalent to the original one. It is for this reason that Russell's theory is said to offer a way of contextually defining definite descriptions.

If we symbolize the definite description as "$(x)P$" ("the unique x such that P," where P is any well-formed expression), Russell's theory can be stated as follows (unless otherwise indicated, it will be supposed that the scope of the occurrence of a definite description is the smallest well-formed part of the formula that contains that occurrence of the definite description): Let us symbolize the scope of the definite description as M and the whole formula as A. M is replaced by the expression "$(\exists y)(z)[(Pz \equiv z = y).\ M']$," where y and z are the first two variables not occurring in A and M' is the result of substituting y for every occurrence of "$(\iota x)P$" in M. The resulting formula, A', is equivalent to A but lacks the definite description that we set out to eliminate.

The motivation for this theory is to be found in certain difficulties that arose for Russell's theory of meaning, the theory that the meaning of a term is its reference. It has been suggested, primarily by W. V. Quine, that since similar difficulties can arise for names in general, this theory should be extended to all names. Russell, however, thought that there was a class of names, *logically proper names,* for which these difficulties could not arise; he therefore favored retaining names of this class. See entry "Proper Names and Descriptions."

definition. The description or explanation of the meaning of a word or phrase. Various types of definitions have been distinguished by logicians. To begin with, there is the distinction between a *lexical* definition (a report of a meaning the word already has) and a *stipulative* definition (a proposal to assign a meaning to a word). One must also distinguish, with traditional logicians, the following techniques for defining: (1) *dictionary* definition, giving a word or phrase that is synonymous with the definiendum; (2) *ostensive* definition, giving examples of objects to which the word or phrase is properly applied; and (3) definition *per genus et differentiam,* giving the genus of the objects to which a word or phrase is properly applied and the differentiae that distinguish these objects from the other members of the genus. See ***predicables.***

Some new types of definition that have been discussed by contemporary logicians include (4) definition *by abstraction,* defining a class term by specifying the properties that an object must have in order to be a member of the class, and (5) *recursive* (*inductive*) definition, defining a number-theoretic function or predicate term by giving the value or values of the function or predicate when 0 is the argument and then giving the value or values when the successor of any number a is the argument in terms of a and the value when a is the argument (cf. ***recursive function***). Finally, one must distinguish (6) *contextual* definitions, which give meaning to the definiendum only in particular contexts, not in isolation.

definition, Aristotelian theory of. See ***predicables.***

demonstration (derivation). A deductive proof offered for a given set of propositions.

De Morgan's laws. The theorems of propositional calculus that assert the material equivalence of "not (A or B)" with "not-A and not-B" and "not (A and B)" with "not-A or not-B." De Morgan, in his book *Formal Logic,* did not actually state these laws; he gave, instead, the corresponding laws for the logic of classes. It should be noted that some of the medieval logicians stated these theorems for the logic of propositions.

denotation. See ***meaning, Frege's theory of.***

dense. Used of an ordered set such that between any two elements of the set there is another element of the set.

denumerable set. A set whose cardinality is aleph-null ($\aleph_0$). Some authors extend "denumerable" so as to make it synonymous with "enumerable."

derivable. See ***deducible.***

derivation. See ***demonstration.***

derived rule of inference. A metalinguistic theorem asserting that under certain conditions there is a proof in the object language for a certain type of well-formed formula. The point of such theorems is that they enable us to state that certain well-formed formulas are theorems of the object language without having to find a proof in the object language for these formulas.

descending induction. An argument that shows that a certain property holds for no number by demonstrating that if it held for any number, it must hold for a lesser number.

diagonal proof. The proof, given by Georg Cantor, that there are infinite sets that cannot be enumerated.

dichotomy. See ***classification.***

dictum de omni et nullo. The principle of syllogistic reasoning that asserts that whatever is distributively predicated (whether affirmatively or negatively) of any class must be predicated of anything belonging to that class.

difference of sets. For any two sets *a* and *b,* the set of all and only those objects that are members of *a* but not of *b.*

differentia. See ***predicables.***

dilemma. An argument whose major premise is the conjunctive assertion of two hypothetical propositions and whose minor premise is a disjunctive proposition. If the minor premise alternatively affirms the antecedents of the major premise, the dilemma is said to be *constructive*; if the minor premise alternatively denies the consequents of the major premise, the dilemma is said to be *destructive.* Constructive dilemmas are divided into *simple constructive* dilemmas (the antecedents of the major premise are different and the consequents are the same) and *complex constructive* dilemmas (both the antecedents and the consequents of the major premise are different). Destructive dilemmas are divided into *simple destructive* dilemmas (the consequents of the major premise are different and the antecedents are the same) and *complex destructive* dilemmas (both the consequents and the antecedents of the major premise are different).

Dimaris. See ***mnemonic terms.***

Disamis. See ***mnemonic terms.***

discreteness. The property possessed by all ordered sets that lack the property of continuity.

disjoint sets. Sets that have no members in common.

disjunction, exclusive (alternation). A binary propositional connective, one possible interpretation of "or," whose truth table is such that "*A* or *B*" is true if and only if one of the two propositions is true and the other false.

disjunction, inclusive. A binary propositional connective ($\vee$), one possible interpretation of "or," whose truth table is such that "*A* or *B*" is true in all cases except where both *A* and *B* are false.

distributed term. In a categorical proposition the occurrence of a term is distributed if and only if the term as used in that occurrence covers all the members of the class that it denotes. In a universal categorical proposition the subject is distributed; in a negative categorical proposition the predicate is distributed.

distributivity. The relation that exists between two relations R and R^* when "$aR(bR^*c)$" is identical with "$(aRb)R^*(aRc)$."

division. See ***classification.***

division non faciat saltum. See ***classification.***

domain of a relation. For any relation *R,* the set of all objects *a* such that there exists an object *b* such that *aRb.*

domain of individuals. For a given interpretation of a given logistic system, the set of objects that is the range of the individual variables.

duality. The relation that exists between two formulas that are the same except for the interchanging of the universal with the existential quantifier, the symbol for the null class with that for the universal class, sum of sets with product of sets, and conjunction with disjunction (where conjunction, disjunction, and negation are taken as primitive, all other propositional connectives being defined in terms of them). The two formulas are said to be the duals of each other. "A and B" and "A or B," for example, are duals.

dyadic relation. A two-place relation.

effectiveness. A notion is said to be effective if there exists an algorithm for determining, in a finite number of steps, whether or not the notion applies to any given object. For example, in a logistic system the notion of a proof is effective, since there is a mechanical procedure for determining, in a finite number of steps, whether or not in that system a given sequence of well-formed formulas constitutes a proof of another given well-formed formula.

element. A member of a given set.

elementary number theory. The theory of numbers insofar as it does not involve analysis.

empty set. See ***null set.***

entailment. The relation that exists between two propositions one of which is deducible from the other.

enthymeme. A syllogism in which one of the premises or the conclusion is not explicitly stated. An example of an enthymeme is the inference of "Socrates is mortal" from "All men are mortal," the missing premise being "Socrates is a man."

enumerable set. A set that either is finite or has a cardinality of aleph-null ($\aleph_0$). Cf. ***denumerable set.***

epagoge. In traditional logic, the process of establishing a general proposition by induction.

epicheirema. A syllogism in which one or more of the premises is stated as the conclusion of an enthymematic prosyllogism. See ***polysyllogism.***

episyllogism. See ***polysyllogism.***

E-proposition. In traditional logic, a universal negative categorical proposition. An example is "No men are mortal."

epsilon. In set theory, the name of the symbol (ϵ) for set-membership.

equality. A relation that exists between two or more sets, equated by some authors with *identity* and by others with *equivalence relation.*

equipollent. Used of sets between which there exists a one-to-one correspondence.

equivalence relation. A relation that is reflexive, symmetric, and transitive (see ***relation***). Identity is a standard example of an equivalence relation.

equivalent. Used of two propositions that are so related that one is true if and only if the other is true. Some authors also use this term, as applied to sets, synonymously with "equipollent."

equivocation. See ***fallacy.***

eristic. The art of fallacious but persuasive reasoning.

essence. See ***predicables.***

Euler's diagrams. The representations, generally attributed to Leonhard Euler, of relations among classes by relations among circles. See entry "Logic Diagrams."

excluded middle, law of. See ***laws of thought.***

existential generalization, rule of. The rule of inference that permits one to infer from a statement of the form "Property *P* holds for an object *a*" a statement of the form "There exists an object such that property *P* holds for it."

existential import. The commitment to the existence of certain objects that is entailed by a given proposition.

existential instantiation, rule of. The rule of inference that permits one to infer from a statement of the form "There exists an object such that property *P* holds for it" a statement of the form "Property *P* holds for an object *a*." Because this inference is not generally valid, restrictions have to be placed on its use.

existential quantifier. The symbol (*E*) or (∃), read "there exists." It is used in combination with a variable and placed before a well-formed formula, as in "(∃*a*) ______" ("There exists an object *a* such that ______").

extension. Although often used synonymously with "denotation," this term is sometimes used to refer to the set of species that are contained within the genus denoted by a given term. In the first sense the extension of "men" is the set of all men; in the second sense it is the set of sets into which humankind can be divided.

extensional. Used of an approach to a problem which in some respect confines attention to truth-values of sentences rather than to their meanings. Thus, a logic in which, for purposes of deductive relations, truth-values may be substituted for sentences is an extensional logic. Cf. ***intensional.***

extensionality, axiom of. An axiom in set theory stating that for any two sets *a* and *b,* if for all *c, c* is a member of *a* if and only if *c* is a member of *b,* then *a* is identical with *b.*

fallacy. An argument that seems to be valid but really is not. There are many possible types of fallacy; traditional logicians have discussed the following ones: (1) *accentus,* a fallacy of ambiguity, where the ambiguity arises from the emphasis (accent) placed on a word or phrase; (2) *affirmation of the consequent,* an argument from the truth of a hypothetical statement and the truth of the consequent to the truth of the antecedent; (3) *ambiguity,* an argument in the course of which at least one term is used in different senses; (4) *amphiboly,* a fallacy of ambiguity where the ambiguity involved is of an amphibolous nature; (5) *argumentum ad baculum,* an argument that resorts to the threat of force to cause the acceptance of the conclusion; (6) *argumentum ad hominem,* an argument that attempts to disprove the truth of what is asserted by attacking the asserter or attempts to prove the truth of what is asserted by appealing to the opponent's special circumstances; (7) *argumentum ad ignorantiam,* an argument that a proposition is true because it has not been shown to be false, or vice versa; (8) *argumentum ad misericordiam,* an argument that appeals to pity for the sake of getting a conclusion accepted; (9) *argumentum ad populum,* an argument that appeals to the beliefs of the multitude; (10) *argumentum ad verecundiam,* an argument in which an authority is

appealed to on matters outside his field of authority; (11) *begging the question* (*circular reasoning*), an argument that assumes as part of the premises the conclusion that is supposed to be proved; (12) *composition,* an argument in which one assumes that a whole has a property solely because its various parts have that property; (13) *denial of the antecedent,* an argument in which one infers the falsity of the consequent from the truth of a hypothetical proposition and the falsity of its antecedent; (14) *division,* an argument in which one assumes that various parts have a property solely because the whole has that property; (15) *equivocation,* an argument in which an equivocal expression is used in one sense in one premise and in a different sense in another premise or in the conclusion; (16) *ignoratio elenchi,* an argument that is supposed to prove one proposition but succeeds only in proving a different one; (17) *illicit process,* a syllogistic argument in which a term is distributed in the conclusion but not in the premises; (18) *many questions,* a demand for a simple answer to a complex question; (19) *non causa pro causa,* an argument to reject a proposition because of the falsity of some other proposition that seems to be a consequence of the first but really is not; (20) *non sequitur,* an argument in which the conclusion is not a necessary consequence of the premises; (21) *petitio principii,* see (11) *begging the question*; (22) *post hoc, ergo propter hoc,* argument from a premise of the form "*A* preceded *B*" to a conclusion of the form "*A* caused *B*"; (23) *quaternio terminorum,* an argument of the syllogistic form in which there occur four or more terms; (24) *secundum quid,* an argument in which a proposition is used as a premise without attention given to some obvious condition that would affect the proposition's application; (25) *undistributed middle,* a syllogistic argument in which the middle term is not distributed in at least one of the premises. See entry "Fallacies."

Felapton. See ***mnemonic terms.***

Ferio. See ***mnemonic terms.***

Ferison. See ***mnemonic terms.***

Fesapo. See ***mnemonic terms.***

Festino. See ***mnemonic terms.***

field of a relation. The union of the domain and the converse domain of a given relation.

figure. A way of classifying categorical propositions. According to most traditional logicians, since figure depends on the position of the middle term in the premises, there are four possible figures. In the first figure the middle term is the subject of the major premise and the predicate of the minor premise. In the second figure the middle term is the predicate of both premises and in the third figure the subject of both premises. In the fourth figure the middle term is the predicate of the major premise and the subject of the minor premise. Aristotle allowed only three figures and treated as being indirectly in the first figure those syllogisms that later logicians placed in the fourth. See entry "Logic, Traditional."

finitary method. The type of method to which David Hilbert and some of his followers restricted themselves in their metamathematical research. The clearest statement of the restrictions was made by Jacques Herbrand, who insisted that the following conditions be met: (1) One must deal only with a finite and determined number of objects and functions. (2) These are to be so defined that there is a univocal calculation of their values. (3) One should never affirm the existence of an object without indicating how to construct it. (4) One must never deal with the set of all the objects of an infinite totality. (5) That a theorem holds for all of a set of objects means that for every particular object it is possible to repeat the general argument in question, which should then be treated as only a prototype of the resulting particular arguments.

finite set (inductive set). A set that either is empty or is such that there exists a one-to-one correspondence between its members and the members of the set of all natural numbers less than a specified natural number. A set which is not finite is said to be *infinite.*

Richard Dedekind introduced a different characterization of finite and infinite sets. A *Dedekind finite* set is one that has no proper subset such that there exists a one-to-one correspondence between the elements of the set and the elements of that proper subset. A *Dedekind infinite* set (or *reflexive* set) is one that is not Dedekind finite. It can be shown that Dedekind's characterization is equivalent to the previous one; the proof, however, involves the axiom of choice.

first element of a set. See ***bound of a set.***

first-order logic. First-order functional calculus. See ***calculus.***

formalism. The doctrine, advanced as a program by David Hilbert and his followers, that the only foundations necessary for mathematics are its formalization and a proof by finitary methods that the system thus produced is consistent. See entry "Mathematics, Foundations of."

formalization. The construction of a logistic system whose intended interpretation is such that under it the truths of a given body of knowledge are the interpreted theorems of the system.

formalized language. A logistic system with an interpretation.

formally imply. A proposition *A* is said to formally imply a proposition *B* in a given logistic system if there is, in that system, a valid proof of *B* from *A* taken as a hypothesis.

formal system. See ***logistic system.***

formation rules. For a given logistic system, the rules that determine which combinations of symbols are well-formed formulas and which are not.

formula. For a given logistic system, any sequence of primitive symbols.

foundation, axiom of (Axiom der fundierung, axiom of regularity). An axiom in set theory stating that every nonempty set *a* contains a member *b* which has no member in common with *a.*

free occurrence of a variable. For a given variable *a* that occurs in a given well-formed formula *A,* an occurrence of *a* in no well-formed part of *A* which is of the form "For all *a, B*" or of the form "There exists an *a, B.*"

free variable. A free variable of a formula *A* is a variable in *A* that has no bound occurrence in *A*.

Fresison. See ***mnemonic terms.***

function. A many-one correspondence.

functional calculus. See ***calculus.***

future contingents, problem of. The problem, first discussed by Aristotle, of whether any contingent statement about the future has a truth-value prior to the time it refers to.

Galenian figure. The fourth syllogistic figure, supposedly introduced by Galen.

generalization, rule of. The rule of inference that allows one to infer from every proposition another proposition that is the same as the original one except that it is preceded by a universal quantifier binding any variable.

general term. A term that is predicable, in the same sense, of more than one individual.

Gentzen's consistency proof. The proof, first given by Gerhard Gentzen in 1936, of the consistency of classical pure number theory with the unrestricted-induction postulate. The proof employs transfinite induction up to the ordinal ϵ_0.

Gentzen system. A system of logic characterized by the introduction into the object language of a new connective (symbolized by →) that has properties analogous to the ordinary metalinguistic idea of "provable in the system." The rules of inference of such a system apply to *Sequenzen*—that is, to formulas of the form "$A_1, A_2, \cdots, A_n \rightarrow B_1, B_2, \cdots, B_m$," where *m* and *n* are equal to or greater than 0, and $A_1, A_2, \cdots, A_n, B_1, B_2, \cdots, B_m$ are formulas of ordinary logical systems.

genus. See ***predicables.***

Gödel-numbering. The assignment of a natural number to each entity of a formal system. See ***arithmetization of syntax.***

Gödel's completeness theorem. The theorem, first introduced by Kurt Gödel in 1930, that every valid well-formed formula of pure first-order functional calculus is a theorem of that system.

Gödel's incompleteness theorems. Two theorems that were first proved by Kurt Gödel in 1931. One states that any ω-consistent system adequate for elementary number theory is such that there is a valid well-formed formula of the system not provable in the system. J. B. Rosser, in 1936, extended this result to any consistent system. The second theorem states that any consistent system adequate for elementary number theory is such that there can be no proof of the consistency of the system within the system. See entry "Gödel's Theorem."

Gödel–von Neumann–Bernays set theory. The form of axiomatic set theory that avoids the paradoxes of set theory by distinguishing between sets (collections that can also be elements of other collections) and classes (collections that cannot be elements of other collections) and ensuring that all the objects leading to paradoxes (for example, the universal class) are classes and not sets.

Henkin's completeness theorem. The theorem, proved by Leon Henkin in 1947, that every secondarily valid well-formed formula of pure second-order functional calculus is a theorem of that system.

hereditary property. See ***ancestral relation.***

Hilbert program. See ***formalism.***

ideal mathematics. For David Hilbert, the nonfinitary part of mathematics, which, although necessary, was suspect and therefore required a consistency proof. See ***real mathematics.***

idempotency. A binary operation is idempotent if and only if that operation, when performed on any element with itself, results in just that element.

identically false. Used of a well-formed formula of propositional calculus whose truth-value is falsehood for all possible values of its constituent well-formed formulas.

identically true. Used of a well-formed formula of propositional calculus whose truth-value is truth for all possible values of its constituent well-formed formulas.

identity. A relation that holds only between an object and itself.

identity, law of. See ***laws of thought.***

identity of indiscernibles. Gottfried Wilhelm Leibniz's principle that two objects are identical if for every class, one object belongs to the class if and only if the other does. This is not to be confused with what W. V. Quine has called the *indiscernibility of identicals,* the principle that if two objects are identical, they belong to the same classes.

iff. A common abbreviation for "if and only if." See ***biconditional.***

ignoratio elenchi. See ***fallacy.***

image. The members of the converse domain of a relation that are values of the relation when its argument is a member of a set that is part of its domain.

immediate inference. An inference of a conclusion from a single premise. Traditional logicians discussed two types: (1) *opposition of propositions,* the inference, from the truth or falsity of one proposition, of the truth or falsity of another proposition having the same subject and predicate (such inferences involve contradictory, contrary, subalternate, and subcontrary propositions), and (2) *eductions,* the inference, from one proposition, of another differing from it in subject or predicate or in both (these involve obversion, conversion, contraposition, and inversion).

imperfect figures. The second and third syllogistic figures, the valid arguments of which, according to Aristotle, are such that their validity can be known only by their reduction to valid syllogisms in the perfect first figure.

implication (conditional). A binary propositional connective (→, ⊃), usually read "if-then," of which there are two major interpretations: (1) *Material implication.* Under this interpretation, "If *A* then *B*" is true in all cases except when *A* is true and *B* false. (2) *Strict implication.* Under this interpretation, "If *A* then *B*" is true only when *B* is deducible from *A. Philonian* implication is the Stoic version of material implication, and *Diodorean* implication is the Stoic interpretation of "if-then" according to which "If *A* then *B*" is true if whenever (in the past, present, or future) *A* is true, *B* is also true.

implicit definition. A set of axioms implicitly define the undefined terms in them by, in effect, confining the references of these terms to the intended ones. The axioms do this by stating conditions satisfiable by only one set of objects.

The idea that a set of axioms can implicitly define the undefined terms in them is usually credited to J. D. Gergonne (1819). It was once thought that the basic terms of arithmetic could be implicitly defined by the axioms (namely, Peano's postulates) containing them; however, it is now known that this cannot be done, since Peano's postulates admit of more than one interpretation.

impredicative definition. Definition of an object in terms of a totality of which it is a member. For an example of impredicative definition, see ***ancestral relation.***

inclusion. A relation that holds between two sets when all the members of one are members of the other. The relation of set-inclusion must be distinguished from that of set-membership.

inconsistent. Used of a set of propositions from which, or a logistic system in which, a contradiction can be derived.

indemonstrables. The Stoics' name for the axioms of their propositional logic.

independence. An axiom *A* of a given logistic system is independent (or has independence) if and only if in the system obtained by omitting *A* from the axioms of the given system, *A* is not a theorem. *A* rule of inference *R* of a given logistic system is independent if and only if in the system obtained by omitting *R* from the rules of inference of the given system, *R* is not a derived rule of inference.

indirect proof (reductio ad absurdum). An argument that proves a proposition *A* by showing that the denial of *A,* together with accepted propositions $B_1, B_2, \cdots, B_n$, leads to a contradiction. Strictly speaking, this fails to prove the truth of *A,* since one of the previously accepted premises may be false; the force of the argument therefore rests on using premises that are far better established than the denial of *A,* so that the denial of *A* will be rejected and *A* accepted.

individual (particular). (1) Anything considered as a unit. (2) In the theory of types, any member of the lowest type.

induction. Among acceptable inferences, logicians distinguish those in which the joint assertion of the premises and the denial of the conclusion is a contradiction from those in which that joint assertion is not a contradiction. The former are deductive inferences; inductive inferences are to be found among the latter.

Much has been written about the precise nature of inductive inferences, but few definite results have been obtained. It is likely that there is a wide variety of types of inductive inferences. Two quite different types are the inference from observational data to theoretical conclusions and the inference from the composition of a sample to the composition of a whole population.

induction, mathematical. An inference of the form "0 has the property *P*; if any natural number *a* has the property *P*, then its successor has the property *P*; therefore, every natural number has the property *P*." The first step is called the *basis,* or the *zero step,* of the induction, and the second is called the *induction step.*

inductive set. See ***finite set.***

inference. Derivation of a proposition (the conclusion) from a set of other propositions (the premises). When the inference is acceptable the premises afford good reasons to assert, or render certain, the conclusion.

infima species. See ***classification.***

infinite set. See ***finite set.***

infinity, axiom of. An axiom in set theory that guarantees the existence of an infinite number of individuals. This axiom takes various forms, all having in common the property of being valid in at least one infinite domain of individuals while not being valid in any finite domain of individuals.

initial ordinal. An ordinal that is not equipollent with any smaller ordinal.

insolubilia. The medieval name for antinomies. The antinomies that are usually referred to by this name are variants of the Liar paradox.

intension. A term sometimes used by traditional authors as synonymous with "connotation." In contemporary logical works "intension" has come to be synonymous with "sense." See ***meaning, Frege's theory of.***

intensional. (1) Used of an approach which in some respect considers the meaning as well as the truth-value of a formula. A characteristic of such systems is that some propositions in them are referentially opaque. Systems of modal logic are usually intensional systems.

(2) Used of a proposition that contains a referentially opaque part. Cf. ***extensional.***

intention, first (primary). In medieval logic, signs that signify things and not other signs are said to have first intention. See entry "Logic, Traditional."

intention, second (secondary). In medieval logic, signs that signify other signs and not things are said to have second intention. See entry "Logic, Traditional."

interpretation. An interpretation of a set *A* of well-formed formulas consists of a nonempty set (the *domain of the interpretation*) and a function which assigns to each individual constant appearing in any of the members of *A* some fixed element in the domain, to each *n*-place predicate letter appearing in any of the members of *A* some *n*-place relation in the domain, and to each *n*-place function letter appearing in any member of *A* some function whose arguments are *n*-tuples of elements of the domain and whose values are also elements of the domain. The individual variables are thought of as ranging over the elements of the domain, and the connectives are given some meaning. Such an interpretation provides meaning for the members of *A*.

The *principal* interpretation is the intended interpretation. The *secondary* interpretations of a set of well-formed formulas are all the interpretations, other than the principal one, such that under them all the members of the set are true.

intersection of sets (product of sets). The set of all the objects that are elements of all the sets $a_1, a_2, \cdots, a_n$ (symbolized "$a_1 \cap a_2 \cap \cdots \cap a_n$").

intuitionism. The doctrine, advanced by L. E. J. Brouwer and his followers, whose key thesis is that a mathematical entity with a particular property exists only if a constructive existence proof can be given for it. As a result the actual infinite is ruled out of mathematics, and only denumerably infinite sets, viewed as potentially infinite, are allowed. Furthermore, the law of excluded middle is rejected in the sense that when infinite classes are being dealt with, a disproof of a universal statement is not automatically a proof of its denial—that is, an existential statement. See entry "Mathematics, Foundations of."

intuitive set theory. The form of set theory that is based on an unrestricted use of the axiom of abstraction. The paradoxes of set theory were generated within a system of intuitive set theory.

inverse of a relation. See ***converse of a relation.***

inversion. In traditional logic, a type of immediate inference in which from a given proposition another proposition is inferred whose subject is the contradictory of the subject of the original proposition. See entry "Logic, Traditional."

iota operator. The definite description operator, ι. It is read: "The unique ______ such that ______."

I-proposition. In traditional logic, a particular affirmative categorical proposition. An example is "Some men are mortal."

joint denial. A binary propositional connective (↓) whose truth table is such that "*A* joint-denial *B*" is true if and only if both *A* and *B* are false. Joint denial and the Sheffer stroke function are the only binary propositional connectives that are adequate for the construction of all truth-functional connectives.

judgment. (1) The affirming or denying of a proposition. (2) The proposition affirmed or denied.

Lambert's diagrams. The representation, introduced by J. H. Lambert, of relations among classes by relations among straight lines.

law of logic. Any general truth of logic.

laws of thought. Three laws of logic that were traditionally treated as basic and fundamental to all thought. They were (1) *the law of contradiction,* that nothing can be both *P* and not-*P,* (2) *the law of excluded middle,* that anything must be either *P* or not-*P*; and (3) *the law of identity,* that if anything is *P,* then it is *P.*

lekton. The Stoic name for the sense of a formula.

lemma. A theorem proved in the course of, and for the sake of, the proof of a different theorem.

level (order). In the ramified theory of types, a class of objects that is composed of all and only those objects such that the definition of one of them requires no reference to a totality containing other members of the class. A hierarchy of levels is built up by beginning with the class of those objects that can be defined without reference to any totality and continuing with succeeding levels, members of each of which are defined in terms of totalities of objects of the previous level.

Liar paradox. See ***paradox, Epimenides' paradox.***

limit. For a given sequence of numbers, the number *a* such that for any arbitrarily small number *b* greater than 0 there exists a number *c* such that for any number *d* larger than *c* the absolute value of the difference between the *d*th member of the sequence and *a* is less than *b.*

limit number. An ordinal number that is not 0 and is such that if *a* is a member of it, then the successor of *a* is also a member of it.

limit ordinal. See ***limit number.***

logic. The study of the validity of different kinds of inference. This term is often used synonymously with *deductive* logic, the branch of logic concerned with inferences whose premises cannot be true without the conclusion's also being true. The other major branch of logic, *inductive* logic, is concerned with inferences whose premises can be true even if the conclusion is false.

logical fiction. The apparent denotation of a symbol that really has no denotation. Formulas containing such symbols are translatable into formulas containing no symbol or symbols that even appear to have this denotation.

logical form. It is commonly said that logic is concerned with the form, not the matter, of a proposition or argument. The distinction between form and matter is, however, seldom made precise; it can therefore best be seen by consideration of an example:

If it is raining, people will carry umbrellas.

It is raining.
People will carry umbrellas.

Analysis of this inference shows that it is valid because it is of the form "If *A,* then *B; A*; therefore, *B.*" The values of the variables make no difference in the validity of the argument. Formal logic is concerned with inferences, like this one, whose validity depends on their form.

As the example shows, the form of a proposition is nothing more than the result of substituting, in the proposition, free variables for the constants, whereas the *matter of a proposition* is that for which the variables are substituted. The form of an argument is the result of substituting, in all the premises and in the conclusion of the argument, free variables for constants.

In some contemporary works any formula that contains one or more free variables is called a form.

logical implication. The relation that holds between two propositions when one is deducible from the other.

logically necessary. See ***analytic.***

logical possibility (possible truth). A proposition that is not self-contradictory. Some authors restrict this term to propositions that are also not logically necessary.

logical truth. See ***analytic.***

logic diagram. A diagram used to represent logical relations. See entry "Logic Diagrams."

logicism. The doctrine, advanced by Gottlob Frege and Bertrand Russell, that all the concepts of mathematics can be derived from logical concepts through explicit definitions and all the theorems of mathematics can be derived from logical axioms through purely logical deduction. See entry "Mathematics, Foundations of."

logistic method. The method of studying a subject by formalizing it.

logistic system (formal system). A system whose primitive basis is explicitly stated in the metalanguage.

Löwenheim's theorem. See ***Skolem-Löwenheim theorem.***

major premise. In a categorical syllogism, the premise that contains the major term.

major term. In a categorical syllogism, the term that is the predicate of the conclusion.

many-one correspondence. A relation *R* such that for every element *a* of its domain there is only one member *b* of its converse domain such that *aRb.* "Son of" is a many-one correspondence since for every member of its domain (for every son) there is only one member of the converse domain (his father) of which it is true that the member of the domain is the son of the member of the converse domain.

many-valued logic. A system of logic in which each formula has more than two possible truth-values.

map of one set into another. A one-to-one correspondence between two sets whose domain is the first set and whose converse domain is a proper subset of the second set.

map of one set onto another. A one-to-one correspondence between two sets whose domain is the first set and whose converse domain is the second set.

material implication. See ***implication.***

mathematical induction. See ***induction, mathematical.***

matter of a proposition. See ***logical form.***

meaning, Frege's theory of. According to this theory, propounded by Gottlob Frege in 1892, the meaning of a proper name has two aspects, the *sense* and the *reference.* The reference of a proper name is that which it is a name of. Thus, the reference of "Sir Walter Scott" is Sir Walter Scott. Frege claimed that there must be, besides the reference, another aspect of the meaning of such a name. "Sir Walter Scott" and "the author of Waverley" have the same reference, but it would be most implausible to say that they have the same meaning. The aspect of meaning that distinguishes "Sir Walter Scott" from "the author of Waverley" is called the sense of the proper name.

It should be noted that this is a theory of the meaning of proper names, not common names. It is for common names that John Stuart Mill first introduced his distinction between *denotation* (the objects to which the common name is properly applied) and *connotation* (the characteristic or set of characteristics that determines to which objects the common name properly applies). Unlike Frege, Mill thought that the meaning of a proper name is simply that which it denotes.

mediate inference. An inference in which the conclusion follows from two or more premises.

membership. The relation that exists between a set and its elements. The relation of set-membership must be distinguished from the relation of set-inclusion.

mention of a term. An occurrence of a linguistic expression in quotation marks for the purpose of talking about that linguistic expression. For example, in "'Cicero' has six letters" it is not the orator himself but the word referring to him that is being discussed.

This is to be contrasted with *use of a term,* the occurrence of a linguistic expression for the purpose of talking about something other than the expression.

metalanguage. A language used to talk about an object language; a *meta-metalanguage* is a language used to talk about a metalanguage, and so forth. Derivatively, a proposition is said to be in the metalanguage if and only if it is about an expression in the object language.

metamathematics (proof theory). The study of logistic systems. Some authors restrict this term to investigations employing finitary methods.

metatheorem. A theorem in a metalanguage.

metatheory. The metamathematical investigations relating to a given logistic system.

method of construction. Bertrand Russell's name for the method of introducing new types of numbers by defining them in terms of previously introduced numbers and the usual logical and set-theoretic notation. Opposed to the method of construction is the *method of postulation,* whereby one introduces new types of numbers as primitive terms with appropriate axioms.

middle term. In a categorical syllogism, the term that occurs in both premises but not in the conclusion.

minor premise. In a categorical syllogism, the premise that contains the minor term.

minor term. In a categorical syllogism, the term that is the subject of the conclusion.

mnemonic terms. The names that the medieval logicians introduced for the valid syllogisms. One such term is "Barbara." The key for these mnemonics is as follows: The three vowels respectively indicate the three constituent propositions of the syllogism as A, E, I, or O. For

first-figure syllogisms the initial consonants are arbitrarily the first four consonants; for the other figures the initial consonants indicate to which of the first-figure syllogisms the syllogism in question may be reduced. Other consonants occurring in second-, third-, and fourth-figure mnemonics indicate the operation that must be performed on the proposition indicated by the preceding vowel in order to reduce the syllogism to a first-figure syllogism. The key for this is as follows: "*s*" indicates simple conversion, "*p*" indicates conversion *per accidens,* "*m*" indicates metathesis (interchanging of the premises), "*k*" indicates obversion, and "*c*" indicates *convertio syllogism* (that is, the syllogism is to be reduced indirectly). In mnemonic terms the only meaningless letters are "*r,*" "*t,*" "*l,*" "*n,*" and noninitial "*b*" and "*d.*" More elaborate mnemonics have been devised for syllogisms in which two or more of the premises exhibit modality. See entry "Logic, Traditional."

Mnemonic Terms

Name	Figure	Major premise	Major premise	Conclusion
Barbara	first	A	A	A
Baroco	second	A	O	O
Bocardo	third	O	A	O
Bramantip	fourth	A	A	I
Camenes	fourth	A	E	E
Camestres	second	A	E	E
Celarent	first	E	A	E
Cesare	second	E	A	E
Darapti	third	A	A	I
Darii	first	A	I	I
Datisi	third	A	I	I
Dimaris	fourth	I	A	I
Disamis	third	I	A	I
Felapton	third	E	A	O
Ferio	first	E	I	O
Ferison	third	E	I	O
Fesapo	fourth	E	A	O
Festino	second	E	I	O
Fresison	fourth	E	I	O

modality. (1) The characteristic of propositions according to which they can be described as "apodictic," "assertoric," or "problematic." An *assertoric* proposition asserts that something is the case; an *apodictic* proposition asserts that something must be the case; a *problematic* proposition asserts that something may be the case. This type of modality was called by the medieval logicians *modality sine dicto* (*de re*).

(2) The characteristic of propositions according to which they can be described as "necessary," "impossible," "possible," or "not-necessary." Medieval logicians called this type *modality cum dicto* (*de dicto*).

modal logic. The study of inferential relations among propositions which are due to their modality. Most logicians treat systems of modal logic as intensional, basing them upon strict implication. An alternative approach is to treat these systems as extensional, basing them upon a many-valued logic. See entry "Modal Logic."

model. An interpretation of a given set of well-formed formulas according to which all the members of the set are true. The *standard* model corresponds to the principal interpretation, and a *nonstandard* model corresponds to a secondary interpretation. See ***interpretation.***

modus ponendo tollens. An inference of the form "Either *A* or *B; A*; therefore, not-*B.*" This type of inference is valid only if "or" is interpreted as exclusive disjunction.

modus ponens. An argument of the form "If *A* then *B; A*; therefore, *B.*" Some authors use the term to designate the rule of inference that allows arguments of this form.

modus tollendo ponens. An argument of the form "Either *A* or *B*; not-*A*; therefore, *B.*"

modus tollens. An argument of the form "If *A* then *B*; not-*B*; therefore, not-*A.*" Some authors use the term to designate the rule of inference that allows arguments of this form.

mood. A way of classifying categorical syllogisms according to the quantity and quality of their constituent propositions.

multiplicative axiom. See ***choice, axiom of.***

name. In traditional logic, a word or group of words that can serve as a term in a proposition. A *general* name is one that can be significantly applied to each member of a set of objects, a *singular* name is one that can be significantly applied to only one object, and a *collective* name is one that can be significantly applied to a group of similar things regarded as constituting a single whole.

natural number. A member of a certain subset of the cardinal numbers. There are various ways of defining this subset so that it contains all and only the desired objects (namely 0, 1, 2, 3, · · ·); the most common way is to define it as the set of all objects that belong to all sets containing 0 and closed under the successor relation.

necessary condition. See ***condition.***

necessary truth. See ***analytic.***

negate of a set. See ***complement of a set.***

negation. A singulary propositional connective (¬, ¯, ~, –), usually read "not," whose truth table is such that "not-*A*" is true if and only if *A* is false.

negative name. In traditional logic, a name that implies the absence of one or more properties or that

denotes everything with the exception of some particular thing or set of things. An example of such a name is "non-Briton."

non sequitur. See ***fallacy.***

normal system of domains. A system of domains such that the axioms of second-order functional calculus are valid in them and the rules of inference of second-order functional calculus preserve validity in them.

null set (empty set). A set with no members.

number. See ***cardinal number; natural number; rational number; real number***; entry "Number."

object language. A language used to talk about things, rather than about other languages. Derivatively, a proposition is said to be in the object language if and only if it is not about any linguistic expression. "Socrates was a philosopher" is therefore in the object language, whereas "'Socrates' has eight letters" is not.

obversion. In traditional logic, a type of immediate inference in which from a given proposition another proposition is inferred whose subject is the same as the original subject, whose predicate is the contradictory of the original predicate, and whose quality is affirmative if the original proposition's quality was negative and vice versa. Obversion of a proposition yields an equivalent proposition when applied to all four types (A, E, I, and O) of propositions that traditional logicians considered. See entry "Logic, Traditional."

omega. The smallest infinite ordinal (denoted by ω), the order type associated with the set of all natural numbers as ordered in their natural order.

omega-complete. Used of a system which, if it contains the theorems that property *P* holds of 0, of 1, of 2, and so on, contains the theorem that *P* holds of all numbers.

omega-consistent. Used of a system which, if it contains the theorems that property *P* holds of 0, of 1, of 2, and so on, does not contain the theorem that *P* holds of all numbers.

one-many correspondence. A relation *R* such that for every member *a* of its converse domain, there is more than one object *b* that is a member of its domain such that *bRa.* "Father of" is an example of a one-many correspondence, since for every member of its converse domain (everyone who has a father) there is only one member of its domain (that person's father) such that the member of the domain is the father of the member of the converse domain.

one-to-one correspondence. A relation *R* such that for every member *a* of its converse domain, there is only one object *b* that is a member of its domain such that *bRa.* A one-to-one correspondence is said to be *order-preserving* if both its domain and its converse domain are simply ordered and if, for all *c* and *d* that are members of its domain and are such that *c* precedes *d* in the ordering of the domain, it is the case that their respective images *e* and *f* in the converse domain are such that *e* precedes *f* in the ordering of the converse domain.

open schema. A formula containing free individual and functional variables.

open sentence. A formula containing free individual variables.

operator. A symbol or combination of symbols that is syncategorematic under the principal interpretation of the logistic system it occurs in and that may be used with one or more variables and one or more constants or forms or both to produce a new constant or form. Universal and existential quantifiers are the most common examples of operators.

O-proposition. In traditional logic, a particular negative categorical proposition. An example is "Some men are not mortal."

order. See ***Level.***

ordered, partially. A set *a* is partially ordered if and only if there is a relation *R* such that for all *b, c,* and *d* that are members of *a,* (1) if *bRc* and *cRd,* then *bRd,* and (2) it is not the case that *bRb.*

ordered, simply. A set *a* is simply ordered if and only if there is a relation *R* such that *a* is partially ordered by *R* and for all *b* and *c* that are members of *a* and are not identical, either *bRc* or *cRb.*

ordered, well. A set *a* is well ordered if and only if there is a relation *R* such that *a* is simply ordered by *R* and for every nonempty subset of *a,* there is a first element of that nonempty subset.

ordered pair. For given objects *a* and *b,* the ordered pair (*a,b*) is the pair set of which one member is the unit set whose only member is *a* and the other member is the pair set whose members are *a* and *b.*

order-preserving. See ***one-to-one correspondence.***

order type. The set of all sets that are ordinally similar to a given set.

ordinally similar. Two or more sets are ordinally similar if and only if there exists between them a one-to-one order-preserving correspondence.

ordinal number. An order type of a well-ordered set.

pairing axiom. An axiom in set theory stating that for any two objects *a* and *b,* there is a set *c* whose members are *a* and *b* only.

pair set. A set that contains exactly two members.

paradox (antinomy). A statement whose truth leads to a contradiction and the truth of whose denial leads to a contradiction. Since F. P. Ramsey it has been customary to distinguish between *logical paradoxes* (often called *paradoxes of set theory*), which can arise in the object language because they involve only the usual logical and set-theoretic symbols, and *semantic paradoxes,* which can arise only in the metalanguage because they involve semantic concepts.

The most prominent logical paradoxes are the following: (1) *Russell's paradox.* Consider the set of all objects that are not members of themselves. Is that set a member of itself? If it is, then it is not. If it is not, then it is. (2) *Cantor's paradox.* Consider the set of all sets. Is it equal to or greater than its power set? If it is equal, then there is a contradiction, since there is a proof that the power set of any set is greater than the set itself. If it is not, then there is a contradiction, since the power set of any set is a set of sets and must therefore be a subset of the set of all sets, and there is a proof that the subset of a set cannot be greater than the set itself. (3) *Burali-Forti's paradox.* Consider the set of all ordinals. Does it have an ordinal number? If it does not, there is a contradiction, since by the "less than" relation it is well ordered, and there is a proof that all well-ordered sets have ordinal numbers. If it does, there is a contradiction, since it can be proved that the set's ordinal number must be both equal to and less than its image in the mapping of the set of all ordinals onto the set of all ordinals less than its own ordinal.

The most prominent of the semantic paradoxes are the following: (1) *Berry's paradox.* Consider the expression "the least natural number not namable in fewer than 22 syllables." Is the number it denotes namable in fewer than 22 syllables? If it is, there is a contradiction, since by definition it cannot be. If it is not, there is a contradiction, since we can produce a way of naming it in 21 syllables—the way we named it in stating this paradox. (2) *Epimenides' paradox.* Consider the sentence "This sentence is not true." Is it true? If it is, then it is not; if it is not; then it is. (3) *Grelling-Nelson paradox of heterologicality.* A predicate is heterological if the sentence ascribing the predicate to itself is false. Is the predicate "heterological" itself heterological? If it is, then it is not; if it is not, then it is. (4) *Paradox of the Liar.* See *Epimenides' paradox* (although the name is often used to refer to the nearly identical paradox beginning with the sentence "This statement expresses a lie"). (5) *Richard's paradox.* Consider the set of all real numbers between 0 and 1 that can be characterized in a finite number of English words. This set has only denumerably many members. It can be shown, in a manner very similar to Cantor's diagonal proof, that we can specify in a finite number of English words a number that cannot belong to the set. Does it belong to the set? If it does, there is a contradiction, since it cannot. If it does not, there is a contradiction, since it can be characterized in a finite number of English words, and all such numbers belong to the set. See entry "Logical Paradoxes."

paradoxes of material implication. These so-called paradoxes consist in the fact that if "if ______ then ______" is taken in the sense of material implication, then any proposition of that form is true if the antecedent is false no matter what the consequent is or if the consequent is true no matter what the antecedent is. Thus, "If Eisenhower were premier of France, then the moon would be made of cheese" and "If 2 + 2 = 17, then Johnson is the president of the United States" are both true propositions if "if-then" is interpreted in the sense of material implication.

paralogism. Any fallacious reasoning.

particular. See ***individual.***

Peano's postulates. A system of five postulates from which one can derive the rest of arithmetic. The five postulates are (1) 0 is a number; (2) the successor of any number is a number; (3) there are no two numbers with the same successor; (4) 0 is not the successor of any number; (5) every property of 0 also belonging to the successor of any number that has that property belongs to all numbers.

per accidens. Used of a predication to the subject of one of its accidents.

perfect figure. The first figure of the syllogism. According to Aristotle, this is the only figure to which the *dictum de omni et nullo* is directly applicable.

per se. Used of a predication to the subject of one of its essential attributes.

petitio principii. See ***fallacy***, (11) *begging the question.*

polysyllogism. A series of syllogisms so linked that the conclusion of one is a premise of another. In such a series a syllogism is said to be a *prosyllogism* if its conclu-

sion is a premise of the syllogism with which it is connected and an *episyllogism* if one of its premises is the conclusion of the syllogism with which it is connected. See ***sorites.***

possible truth. See ***logical possibility.***

post hoc, ergo propter hoc. See ***fallacy.***

postulate. Although often used synonymously with "axiom," this term is sometimes confined to the basic propositions of a particular discipline, with the axioms being the basic propositions common to all disciplines (for example, the laws of logic). The distinction arises only when one is concerned not merely with a formal system but also with its interpretation.

postulation, method of. See ***method of construction.***

potential infinite. The infinite regarded as a limiting concept, as something becoming rather than as something completed.

power. See ***cardinality.***

power set. The set of all subsets of a given set.

power-set axiom. An axiom in set theory stating that for any given set, its power set exists.

pragmatics. See ***semantics, formal.***

predicables. A classification of things and concepts as predicated of subjects, first made by Aristotle. His four predicables were definition, genus (in which he included differentia), proprium, and accident. Medieval logicians, following Porphyry, offered a list of five predicables—species, differentia, genus, proprium, and accident—which was adopted by most traditional logicians.

For Aristotle one defined a term by stating the *essence* of the object that it names (this statement is called the *definition*). The essence of a thing is that property which makes it the type of thing it is and not some other type of thing. The essence has two aspects: the *genus* is that which is predicable essentially of other kinds of things as well, and the *differentia* is that which is possessed essentially only by things of one type (members of one species) and not by things of any other type. Thus, in "Man is a rational animal" the genus is "animal," and the differentia is "rational."

Aristotle distinguished between the essence of a thing and other properties which belong only to that type of thing but are not part of its essence; such a property is called a *proprium.* The precise manner in which he hoped to make this distinction is not very clear. He also recognized that a thing might have a property that it need not have. He called such a property an *accident.*

predicate. Traditionally, the word or group of words in a categorical proposition that connote the property being attributed to the subject or denote the class which the subject is being included in or excluded from. The term is often extended, in contemporary works, to cover all words or groups of words that connote properties or relations in any type of proposition. Thus, in "All men are mortal" the predicate is "mortal."

predicate calculus. See ***calculus.***

predication. The attributing of a property to a subject.

premise. A member of the set of propositions, assumed for the course of an argument, from which a conclusion is inferred.

primitive basis. The list of primitive symbols, formation rules, axioms, and rules of inference of a given logistic system.

primitive symbols. Those symbols of a given logistic system that are undefined and are not divided into parts in the course of operating within the system. One can, following John von Neumann, divide these symbols into constants, variables, connectives, operators, and bracket-like symbols.

privative name. A name that implies the absence of a property where it has been or where one might expect it to be.

problematic proposition. See ***modality.***

product of sets. See ***intersection of sets.***

proof. For a given well-formed formula A in a given logistic system, a proof of A is a finite sequence of well-formed formulas the last of which is A and each of which is either an axiom of the system or can be inferred from previous members of the sequence according to the rules of inference of the system.

proof from hypothesis. A proof from a given set of hypotheses $A_1, A_2, \cdots, A_n$ in a given logistic system is a sequence of well-formed formulas the last of which is the conclusion of the proof and each of which is either an axiom of the system or one of $A_1, A_2, \cdots, A_n$ or a formula that can be inferred from previous formulas in the sequence by the rules of inference of the system.

proof theory. See ***metamathematics.***

proper class. An object which contains members but which cannot itself be a member of any object.

proper subset. A subset of a given set that is not identical with the given set.

proposition. There is no uniform use of the word *proposition* among logicians and philosophers. Many writers distinguish a proposition from a sentence; thus, "Socrates was a philosopher" and "Socrates war ein Philosoph" would be two different sentences that express the same proposition. Other writers use *sentence* and *proposition* interchangeably. To avoid some of the associations of the word *proposition* some contemporary philosophers abandon the term altogether in favor of *statement.* For a discussion of some of the philosophical controversies arising in this connection, see entry "Propositions." For present purposes it is assumed that the reader has a rough idea of what the term *proposition* means. This discussion will accordingly confine itself to an account of the different kinds of propositions distinguished by logicians.

Propositions may be classified in many ways. To begin with, one must distinguish *simple* (or *atomic* or *elementary*) propositions, propositions that do not have other propositions as constituent parts, from *compound* (or *molecular*) propositions, propositions that do have other propositions as constituent parts.

Among simple propositions the more important types are *categorical* (or *subject-predicate*) propositions, which affirm or deny that something has a property or is a member of a class, and *relational* propositions, which affirm or deny that a relation holds between two or more objects. A categorical proposition is *singular* when its subject is the name of an individual and *general* when its subject is the name of a property or class, affirmative when its predicate is affirmed of the subject and *negative* when its predicate is denied of the subject. A general categorical proposition is *universal* when it is talking about all the members of the subject class or all the objects that have the subject property and *particular* when it is talking about only some of the members of the subject class or some of the objects that have the subject property.

Among compound propositions the most important types are *alternative* (or *disjunctive*) propositions, which are of the form "*A* or *B*," *conditional* (or *hypothetical*) propositions, of the form "If *A* then *B*," *conjunctive* propositions, of the form "*A* and *B*," and *negative* propositions, of the form "Not-*A*." Many propositions that seem to be simple turn out under proper analysis to be compound. Such propositions are known as *exponible* propositions.

Kant, and many logicians following him, distinguished a class of *infinite* (or *limitative*) propositions, affirmative propositions with a negative term as predicate. This distinction has been challenged by many authors. A more widely accepted addition to our classification is the *indefinite* proposition, a proposition that is equivocal because no indication is given of whether it is universal or particular. Finally, modality provides still another means of classifying propositions.

propositional calculus. See ***calculus.***

propositional connective. See ***connective.***

propositional function. A function whose range of values consists exclusively of truth-values. Thus, "*a* is the father of George Washington" is a propositional function, since for any argument for *a*, the value of the whole unit is truth or falsehood, depending on whether or not the argument is the name of George Washington's father.

proprium. See ***predicables.***

prosyllogism. See ***polysyllogism.***

protothetic. A form of the extended propositional calculus, first introduced by Stanisław Leśniewski, to which have been added variables whose values are truth-functions and a notation for the application of a function to its argument or arguments, and in which the quantifiers are allowed to have variables of any kind as operator variables. In the *higher* protothetic, variables whose values are propositional functions of truth-functions are added.

proximum genus. See ***classification.***

quality of a proposition. The characteristic that makes a proposition affirmative or negative. Kant, and logicians following him, added a third type, infinite propositions. See ***proposition.***

quantification of the predicate. The prefixing of a sign of quantity, "some" or "all," to the predicate of a proposition in the same way as to the subject, a device introduced by Sir William Hamilton. The claim was that this would make explicit what was implicit in the proposition.

quantifier. An operator of which it is true that both the constant or form it is used with and the constant or form produced are propositions or propositional forms. Thus, an existential quantifier, when joined to a proposition or propositional form *A*, produces a new proposition or propositional form "$(\exists a)M$."

quantity of a proposition. The characteristic that makes a proposition universal or particular. Kant and others considered singular propositions as being a third, distinct type of quantity.

Quine's set theories. A group of set theories proposed by W. V. Quine, combining some of the features of type theory with some of the features of the Zermelo-Fraenkel

and Gödel–von Neumann–Bernays set theories. As in the set theories, the axiom of abstraction is not retained in its full power, and the formation rules of intuitive set theory are not modified; as in type theory, the notion of stratification is used, since in certain key axioms only stratified formulas generate sets.

range of a relation. See ***converse domain of a relation.***

range of values. The class of those things that are ambiguously named by a given variable.

rational number. A number that can be put into the form *a/b,* where *a* is any integer and *b* any natural number.

real mathematics. For David Hilbert, that part of mathematics that is finitary in character, has therefore a clear and intuitive meaning, and poses no problem about its foundation except for the fact that when ideal mathematics is adjoined to it the possibility of inconsistency arises. See ***ideal mathematics.***

real number. Any number which can be represented by an unending decimal.

recursive function. There are various types of recursive functions. In order to explain them we must first introduce some terminology: a *constant function* is a function that has the same value for all of its arguments; a *successor function* has as its value for any given argument the successor of that argument; an *identity function* is a function of *n* arguments whose value is always the *i*th argument. All such functions are known as *fundamental functions.*

A function of *n* arguments is defined by *composition* when, given any set of previously introduced functions of *n* arguments, the value of the new function is equal to the value of a previously introduced function whose arguments in any particular case are the values of each of the members of the set of functions when their arguments are the arguments of the newly introduced function in that particular case. In symbols, where *P* is the new function being defined by composition, $P(a_1, a_2, \cdots, a_n) = R(S_1(a_1, a_2, \cdots, a_n), S_2(a_1, a_2, \cdots, a_n), \cdots, S_m(a_1, a_2, \cdots, a_n))$, where *R* and $S_1, S_2, \cdots, S_m$ are previously introduced functions.

A function is defined by *recursion* in the following circumstances: (1) A value is assigned to the function for the case where one of its arguments is 0 in terms of a previously introduced function whose arguments, except for 0, are in any particular case all and only the arguments of the new function in that particular case. In symbols, where *P* is the new function and *R* the previously introduced function, $P(a_1, a_2, \cdots, a_n, 0) = R(a_1, a_2, \cdots, a_n)$. (2) A value is given to the new function when 0 is not one of its arguments and when one of its arguments is the successor of any number *b,* in terms of a previously introduced function *S,* whose arguments, except for the successor of *b,* are in any particular case all the arguments of the newly introduced function, *b* itself, and the value of the new function when its arguments are all and only the arguments already given for *S.* In symbols, $P(a_1, a_2, \cdots, a_n, b + 1) = S(a_1, a_2, \cdots, a_n, b, P(a_1, a_2, \cdots, a_n, b))$.

Any numerical function that is a fundamental function or can be obtained, by composition or recursion or both, from the fundamental functions by a finite sequence of definitions is a *primitive recursive numerical function.* A function *P* is introduced by the *least-number operator* if its value for a given set of arguments is the least number *b* such that the value of a previously introduced function *R,* whose arguments in any particular case are the arguments of *P* in that case and *b,* is equal to 0 provided that there is such a *b*; if there is no such *b,* the function is undefined for those arguments. In symbols, $P(a_1, a_2, \cdots, a_n)$ = the least *b* such that $R(a_1, a_2, \cdots, a_n, b) = 0$, provided that there is a *b* such that $R(a_1, a_2, \cdots, a_n, b) = 0$. Any numerical function that either is a fundamental function or can be obtained from the fundamental functions by a finite sequence of definitions by composition, recursion, and the least-number operator (when this operator is used in defining a general recursive function, it must be the case that for all $a_1, a_2, \cdots, a_n$ there is a *b* such that $R(a_1, a_2, \cdots, a_n, b) = 0$) is a *general recursive numerical function.*

recursively enumerable. Used of a set or class that is enumerated (allowing for repetitions) by a general recursive function. That is, there is a general recursive function whose converse domain has the same members as the set when its domain is the set of natural numbers.

recursive number theory. The development of number theory, instituted by Thoralf Skolem, in which no quantifiers are introduced as primitive symbols, in which universality is expressed by the use of free variables, and in which functions are introduced through definitions by recursion.

recursive set. A set that is enumerated (allowing for repetitions) by a general recursive function and whose complement is also enumerated (allowing for repetitions) by a general recursive function.

reducibility, axiom of. An axiom, introduced by Bertrand Russell and A. N. Whitehead in *Principia Mathematica,* which says that for any propositional function of

arbitrary level there exists a formally equivalent propositional function of the first level.

reductio ad absurdum. (1) See ***indirect proof.*** (2) The method of proving a proposition by showing that its denial leads to a contradiction. In this sense it is often known as a *reductio ad impossibile.*

reduction of syllogisms. The process whereby syllogisms in imperfect figures are expressed in the first figure. Reduction is *direct* when the original conclusion follows from premises in the first figure derived by conversion, obversion, etc., from premises in an imperfect figure. Reduction is *indirect* when a new syllogism is formed which establishes the validity of the original conclusion by showing the illegitimacy of its contradictory. See entry "Logic, Traditional."

reference. See ***meaning, Frege's theory of.***

referential opacity. An occurrence of a word or sequence of words such that one cannot in general supplant the word or sequence of words with another word or sequence of words that refers to the same thing while preserving the truth-value of the containing sentence. For example, although "9 is necessarily greater than 7" is true, the result of substituting for "9" a sequence of words that refers to the same thing, "the number of planets," is the false proposition "The number of planets is necessarily greater than 7." Therefore, in this occurrence "9" is referentially opaque.

reflexive relation. See ***relation.***

reflexive set. See ***finite set.***

regularity, axiom of. See ***foundation, axiom of.***

relation. This term is not adequately defined in traditional logic. The failure to offer an adequate definition is symptomatic of the lack of serious consideration, on the part of traditional logicians, of the significant differences between categorical and relational propositions. Augustus De Morgan and C. S. Peirce were the first logicians in the contemporary period to study the logic of relational propositions. Since their time this subject has become an important part of logic. In contemporary works, particularly in works on set theory, a relation is defined as a set of ordered pairs.

A relation *R* is *reflexive* if "*aRa*" holds for all *a* that are members of the field of *R, irreflexive* if "*aRa*" holds for no members of the field of *R,* and *nonreflexive* if "*aRa*" holds for some but not all members of the field of *R.* For example, "is a member of the same family as" is a reflexive relation, "is not a member of the same family as" is an irreflexive relation, and "loves" is a nonreflexive relation.

A relation *R* is *symmetric* if for all *a* and *b* that are members of the field of *R, aRb* if and only if *bRa, asymmetric* if for all *a* and *b* that are members of the field of *R, aRb* if and only if not-*bRa,* and *nonsymmetric* when "*aRb*" and "*bRa*" hold for some but not all *a* and *b* that are members of the field of *R.* For example, "is a member of the same family as" is a symmetric relation, "is a child of" is an asymmetric relation, and "is a brother of" is a nonsymmetric relation.

A relation *R* is *transitive* when for all *a, b,* and *c* that are members of the field of *R,* if *aRb* and *bRc,* then *aRc, intransitive* when for all *a, b,* and *c* that are members of the field of *R,* if *aRb* and *bRc,* then not-*aRc,* and *nontransitive* when if *aRb* and *bRc,* then "*aRc*" holds for some but not all of the *a, b,* and *c* that are members of the field of *R.* For example, "is a descendant of" is a transitive relation, "is a child of" is an intransitive relation, and "is not a brother of" is a nontransitive relation.

The foregoing classifications are said to apply to a relation in a set if the corresponding properties hold for all members of the field of a relation that are members of the set. A relation is *connective* in a set if for all distinct *a* and *b* that are members of the set, either *aRb* or *bRa.*

The study of relational propositions has raised many philosophical issues—and has greatly influenced discussions of older issues—about the nature of relations. On these matters, see entry "Relations, Internal and External."

replacement, axiom of (axiom of substitution). An axiom in set theory stating that for any set *a* and any single-valued function *R* with a free variable *b,* there exists *a* set that contains just the members *R*(*b*), with *b* being a member of *a.*

representative of a cardinal number. A set that has a given cardinal number as its cardinality.

Richard's paradox. See ***paradox.***

rule of inference (transformation rule). For a given logistic system, any rule in its metalanguage of the form "From well-formed formulas of the form $A_1, A_2, \cdots, A_n$, it is permissible to infer a well-formed formula of the form *B.*"

Russell's paradox. See ***paradox.***

Russell's theory of definite descriptions. See ***definite descriptions, theory of.***

Russell's vicious-circle principle. The principle according to which impredicative definitions are not allowed.

satisfiable. A well-formed formula that is satisfiable in some nonempty domain of individuals.

satisfiable in a domain. A well-formed formula is satisfiable in a given domain of individuals if and only if it has the value truth for at least one system of possible values of its free variables.

Schröder-Bernstein theorem. The theorem, first conjectured by Georg Cantor and proved by Felix Bernstein and Ernst Schröder, which states that if *a* and *b* are sets such that *a* is equipollent with a subset of *b* and *b* is equipollent with a subset of *a,* then *a* and *b* are equipollent.

scope of a quantifier. For a given occurrence of a quantifier as part of a well-formed part of a well-formed formula, the rest of that well-formed part.

secondarily satisfiable. Used of a well-formed formula that is satisfiable in some normal system of domains.

secondarily valid. Used of a well-formed formula that is valid in every normal system of domains.

second-order logic. Second-order functional calculus. See ***calculus.***

section of a set. See ***segment of a set.***

segment of a set (section of a set). The subset of a given set ordered by a given relation whose members are those members of the set that precede a given member in the given ordering.

selection set. A set that contains one member from each subset of a given set.

self-contradiction. A proposition that in effect both asserts and denies some other proposition.

semantical rule. Any rule in the metalanguage that concerns the meaning of expressions in the object language.

semantics, formal (semiotics). The study of linguistic symbols. Following C. W. Morris, it is customary to divide formal semantics into three areas: (1) *Syntax,* the study of the relations between symbols. The study of the ways in which the symbols of a given language can be combined to form well-formed formulas is one part of syntax. (2) *Semantics,* the study of the interpretation of symbols. Following W. V. Quine, it is customary to distinguish between the theory of reference, which studies the reference or denotation of symbols, and the theory of meaning, which studies the sense or connotation of symbols. (3) *Pragmatics,* the study of the relations between symbols, the users of symbols, and the environment of the users. Thus, the study of the conditions in which a speaker uses a given word is part of pragmatics. See entry "Semantics."

sense. See ***meaning, Frege's theory of.***

sentential calculus. See ***calculus.***

sentential connective. See ***connective.***

sequence. A function whose domain is a subset, not necessarily a proper one, of the set of natural numbers. Some authors extend the term to any function whose domain is ordered.

set. (1) An aggregate. (2) In Gödel–von Neumann–Bernays set theory, where a distinction is made between sets and classes, sets are those objects that can both contain members and be members of some other object.

Sheffer stroke function (alternative denial). A binary propositional connective (|), whose truth table is such that "*A* stroke-function *B*" is false if and only if *A* and *B* are both true. The Sheffer stroke function and joint denial are the only binary propositional connectives adequate for the construction of all truth-functional connectives.

simultaneously satisfiable. A class of well-formed formulas is said to be simultaneously satisfiable if there is some nonempty domain of individuals such that for all the free variables in all the formulas that are members of the class, there exists at least one system of values in that domain for which every formula in the class has the value truth.

singular term. A term that, in the sense in which it is being used, is predicable of only one individual. For example, any definite description is a singular term.

singulary connective. See ***connective.***

Skolem-Löwenheim theorem. In 1915, Leopold Löwenheim proved that if a well-formed formula is valid in an enumerably infinite domain, it is valid in every nonempty domain. A corollary is that if a well-formed formula is satisfiable in any nonempty domain, it is satisfiable in an enumerably infinite domain. In 1920, Thoralf Skolem generalized this corollary—and thus completed the theorem—by proving that if a class of well-formed formulas is simultaneously satisfiable in any nonempty domain, then it is simultaneously satisfiable in an enumerably infinite domain.

Skolem's paradox. The seemingly paradoxical fact that systems in which Cantor's theorem is provable, and which therefore have nondenumerable sets, must, by virtue of the Skolem-Löwenheim theorem, be satisfiable in an enumerably infinite domain.

sorites. A chain of syllogisms in which the conclusion of each of the prosyllogisms is omitted. If each of the conclusions forms the minor premise of the following episyl-

logism, the sorites is an *Aristotelian* sorites; if each of the conclusions forms the major premise of the following episyllogism, it is a *Goclenian* sorites.

sound. Used of an interpretation of a logistic system such that under the interpretation all the axioms either denote truth or always have the value truth, and all the rules of inference are truth-preserving.

species. See ***classification.***

square of opposition. A diagrammatic representation of that part of the traditional doctrine of immediate inferences between categorical propositions that went under the name of the opposition of propositions. See entry "Logic, Traditional."

stratification. The substitution of numerals for variables in a formula (the same numeral for each occurrence of a single variable) in such a way that the symbol for class-membership is flanked always by variables with consecutive ascending numerals.

subalternation. The relation between a universal and a particular proposition of the same quality. Traditionally this relation has been viewed in such a way that the universal proposition implies the particular proposition. The universal proposition is called the *subalternant*; the particular proposition is called the *subalternate.*

subaltern genera. See ***classification.***

subcontrary propositions. Two propositions that cannot both be false but may both be true. Any I- and O-propositions with the same subject and the same predicate form a pair of subcontrary propositions.

subject. The word or words in a categorical proposition that denote the object to which a property is being attributed or the class which is either included in or excluded from some other class.

subset. Any set *b* such that all the members of *b* are members of a given set *a.*

substitution, axiom of. See ***replacement, axiom of.***

substitution, rule of. A rule of inference that allows one to infer from a given formula *A* another formula *B* that is the same as *A* except for certain specified changes of symbols. The various rules of substitution differ in the types of changes they allow.

successor. For a given number, the number that follows it in the ordinary ordering of the numbers. In Peano's axiomatic treatment of arithmetic "successor" is treated as a primitive term. In the various set-theoretic treatments of arithmetic it is defined differently. For example, "the successor of *a*" is sometimes defined as the unit set whose only member is *a.*

sufficient condition. See ***condition.***

summum genus. See ***classification.***

sum of sets. See ***union of sets.***

sum set. For a given set *a,* the set whose members are all and only those objects which are members of members of *a.*

sum-set axiom. An axiom in set theory stating that for any set *a,* its sum set exists.

supposition. Roughly, the property of a term whereby it stands for something; the doctrine of supposition was extensively developed by the medieval logicians. *Material* supposition is possessed by those terms that stand for an expression, and *formal* supposition is possessed by those terms that stand for what they signify. Among terms having formal supposition, those that are common terms have *common* supposition, and those that are properly applicable to only one individual have *discrete* supposition. When in a given occurrence a common term stands for the universal, it has *simple* supposition; opposed to this is *personal* supposition, a property possessed by a common term in those occurrences where it stands for particular instances.

syllogism. A valid deductive argument having two premises and a conclusion. The term is often restricted to the case where both premises and the conclusion are categorical propositions that have between them three, and only three, terms. More careful authors distinguish this case by referring to it as a *categorical* syllogism. A *hypothetical* syllogism is one whose premises and conclusions are hypothetical propositions, and a *disjunctive* syllogism is one whose premises and conclusion are disjunctive propositions. All of these cases, where the three propositions are of the same type, are *pure* syllogisms. A *mixed* syllogism is one in which there occur at least two types of propositions.

A *strengthened* syllogism is one in which the same conclusion could be obtained even if we substitute for one of the premises that is a universal proposition its subalternate. Thus, the syllogism whose premises are "All men are mortal" and "All baseball players are men" and whose conclusion is "Some baseball players are mortal" is a strengthened syllogism, since it would have been sufficient to have as a premise "Some baseball players are men." A *weakened* syllogism is one whose premises imply a universal proposition but whose conclusion is the subalternate of that universal proposition. The above example is also an example of a weakened syllogism, since the premises, as they stand, imply "All baseball players are mortal."

symbol, improper. A symbol that is syncategorematic under the principal interpretation of the logistic system it occurs in. An example of such a symbol is "and."

symbol, proper. A symbol that is categorematic under the principal interpretation of the logistic system it occurs in. Any individual constant is a proper symbol.

symmetrical relation. See ***relation.***

syncategorematic. In traditional logic, used of a word which cannot be a term in a categorical proposition and which must be used along with a term in order to enter into a categorical proposition. An example of this is "all." In contemporary logic the term refers to any symbol that has no independent meaning and acquires its meaning only when joined to other symbols. Cf. ***categorematic.***

syntactical variable. A variable ranging over the names of symbols and formulas.

syntax. See ***semantics, formal.***

synthetic. Used of a proposition that is neither analytic nor self-contradictory.

systematic ambiguity (typical ambiguity). A convention, introduced by Bertrand Russell and A. N. Whitehead, whereby one does not specify the type or order to which the variables in a formula belong, thus allowing one formula to represent an infinite number of formulas, namely all those formulas that are exactly like it except for the fact that their variables are assigned orders and types in such a manner that the formula formed is well-formed according to the formation rules of the ramified theory of types.

tautology. A compound proposition that is true no matter what truth-values are assigned to its constituent propositions. Thus, "*A* or not-*A*" is a tautology, since if "*A*" is true, then the whole proposition is true, and if "*A*" is false, then "not-*A*" is true, and therefore the whole proposition is still true.

term. Traditionally, the subject or predicate in a categorical proposition. Some authors extend the word *term* to cover all occurrences of categorematic words or expressions which, although not propositions by themselves, are parts of a proposition.

tertium non datur. The law of excluded middle. See ***laws of thought.***

theorem. Any well-formed formula of a given logistic system for which there is a proof in the system.

theorem schema. A representation of an infinite number of theorems by means of an expression that contains syntactical variables and has well-formed formulas as values. Every value of the expression is to be taken as a theorem.

theory of types. The theory, introduced by Bertrand Russell and A. N. Whitehead in *Principia Mathematica,* which avoids the paradoxes of set theory by modifying the formation rules of intuitive set theory. In the *simple theory of types* the only modification is that every variable is assigned a number that signifies its type, and formulas of the form "*a* is a member of *b*" are well-formed if and only if *a*'s type-number is one less than *b*'s. In *ramified type theory* each variable is also assigned to a particular level, and certain rules are introduced about the levels of variables; these rules are such as to exclude classes defined by impredicative definitions. See entry "Types, Theory of."

tilde. The name of the symbol for negation (~).

token. A specified utterance of a given linguistic expression or a written occurrence of it. An expression-*type,* on the other hand, is an entity abstracted from all actual and potential occurrences of a linguistic expression. In "John loves John," for example, there are three word-tokens but only two word-types.

transfinite cardinals. All cardinal numbers equal to or greater than aleph-null ($\aleph_0$).

transfinite induction. A proof by course-of-values induction where the numbers involved are the ordinal numbers. This type of proof is important because it can be used to show that a property holds not only for the finite ordinals but for the transfinite ordinals as well.

transfinite ordinal. The order-type of an infinite well-ordered set.

transfinite recursion. A definition of a function by recursion in such a way that a value is assigned not only when the argument is a finite ordinal but also when it is a transfinite ordinal.

transformation rule. See ***rule of inference.***

transitive relation. See ***relation.***

transposition. A rule of inference that permits one to infer from the truth of "*A* implies *B*" the truth of "Not-*B* implies not-*A,*" and conversely.

trichotomy, law of. See ***comparability, law of.***

truth-function. A function whose arguments and values are truth-values. A compound proposition is said to be a truth-functional proposition if the connective that is adjoined to the constituent propositions to form the compound proposition has a truth-function associated with it. In such a case, since the only arguments of the function are truth-values, the truth-value of the compound proposition depends only on the truth-values of its constituent propositions.

truth table. A table that shows the truth-value of a compound proposition for every possible combination of the truth-values of its constituent propositions.

truth-value. One of two abstract entities, truth and falsehood, postulated in Fregean semantics to serve as the reference of true and false sentences. In many-valued logics other truth-values are introduced.

Turing-computable. Used of a function whose value for any given argument a Turing machine can compute. The notion of Turing computability, due to A. M. Turing, is often introduced as a way of making precise the notion of an effectively computable function.

Turing machine. A machine that is capable of being in any one of a finite number of internal *states* at any particular time. The machine is supplied with a linear tape divided into squares on which symbols (from a fixed finite alphabet) may or may not be printed. It scans one, and only one, square at any given time and can erase a symbol from the scanned square and print some other symbol on it. The machine's behavior (in terms of changing what is on the scanned square, changing its internal state, and moving the tape so as to scan a different square) is governed by a *table* of instructions that determines what the machine is to do, given any *configuration* (a combination of the state the machine is in and the symbol on the scanned square) of the machine.

type. (1) See ***token.*** (2) In the theory of types, a class of objects all of whose members are such that they can be members of the same object. The lowest type is composed of all individuals, the next type of all sets of individuals, and each succeeding type of sets whose members are objects of the immediately preceding type.

typical ambiguity. See ***systematic ambiguity.***

union of sets (sum of sets). The set whose members are all and only those objects that are members of at least one of two or more sets.

unit set. A set with only one member.

universal generalization, rule of. The rule of inference that permits one to infer from a formula of the form "Property *P* holds for an object *a*" a formula of the form "Property *P* holds for all objects." Because this inference is not generally valid, restrictions have to be placed on its use.

universal instantiation, rule of. The rule of inference that permits one to infer from a statement of the form "Property *P* holds for all objects" a statement of the form "Property *P* holds for an object *a*."

universal quantifier. The symbol ( ) or (∀), read "for all." It is used in combination with a variable and placed before a well-formed formula, as in "(*a*) ______" ("For all *a*, ______").

universal set. A set such that there is no object *a* that is not a member of the set.

universe of discourse. Those objects with which a discussion is concerned.

univocal. A linguistic expression is univocal if and only if it is neither ambiguous nor equivocal.

use of a term. See ***mention of a term.***

valid formula. A well-formed formula that is valid in every nonempty domain. A well-formed formula is said to be valid for a given domain of individuals if it is true for all possible values of its free variables.

valid inference. An inference the joint assertion of whose premises and the denial of whose conclusion is a contradiction.

value. A member of the range of values of a given variable.

value of a function. That member of the converse domain of a function with which a given argument is paired under the function.

variable. A symbol that under the principal interpretation is not the name of any particular thing but is rather the ambiguous name of any one of a class of things.

Venn diagram. A modification, first introduced by John Venn, of Euler's diagrams. The key differences between Euler's diagrams and Venn's diagrams stem from the fact that Venn, and many other logicians, wanted to deny the traditional assumption that propositions of the form "All *P* are *Q*" or "No *P* are *Q*" imply the existence of any *P*'s. For details, see entry "Logic Diagrams."

vicious-circle principle. See ***Russell's vicious-circle principle.***

well-formed formulas. Those formulas of a given logistic system of which it can sensibly be asked whether or not they are theorems of the system. In any particular system, rules are given that define the class of well-formed formulas and enable one to determine mechanically whether or not a given string of symbols is a well-formed formula of the system.

well-ordering theorem. The theorem stating that for any set there is a relation that well-orders it. See ***choice, axiom of.***

wff. A common abbreviation for "well-formed formula."

Zermelo-Fraenkel set theory. That form of axiomatic set theory that avoids the paradoxes of set theory by dropping the axiom of abstraction and substituting for it a set of axioms about set-existence.

Boruch A. Brody (1967)

LOGIC AND THE FOUNDATIONS OF MATHEMATICS

A very detailed account of main developments of logic will be found in *Logic, History of*. Brief explanations of many of the terms commonly used by logicians will be found in *Logical Terms, Glossary of*. The Encyclopedia also features the following articles dealing with questions in logic and the foundations of mathematics: *Artificial and Natural Languages*; *Combinatory Logic*; *Computability Theory*; *Computing Machines*; *Decision Theory*; *Definition*; *Existence*; *Fallacies*; *Geometry*; *Gödel's Theorem*; *Identity*; *Infinity in Mathematics and Logic*; *Laws of Thought*; *Logical Paradoxes*; *Logic Diagrams*; *Logic Machines*; *Many-Valued Logics*; *Mathematics, Foundations of*; *Modal Logic*; *Negation*; *Number*; *Questions*; *Semantics*; *Set Theory*; *Subject and Predicate*; *Synonymity*; *Syntactical and Semantical Categories*; *Types, Theory of;* and *Vagueness*. See "Logic" and "Mathematics, Foundations of," in the index for entries on thinkers who have made contributions in this area.

LOGIC DIAGRAMS

"Logic diagrams" are geometrical figures that are in some respect isomorphic with the structure of statements in a formal logic and therefore can be manipulated to solve problems in that logic. They are useful teaching devices for strengthening a student's intuitive grasp of logical structure, they can be used for checking results obtained by algebraic methods, and they provide elegant demonstrations of the close relation of logic to topology and set theory.

Leonhard Euler, the Swiss mathematician, was the first to make systematic use of a logic diagram. Circles had earlier been employed, by Gottfried Wilhelm Leibniz and others, to diagram syllogisms, but it was Euler who, in 1761, first explained in detail how circles could be manipulated for such purposes. Euler's contemporary Johann Heinrich Lambert, the German mathematician, in his *Neues Organon* (1764) used straight lines, in a manner similar to Euler's use of circles, for diagramming syllogisms.

FIGURE 1

Venn Diagram Applied to a Syllogism

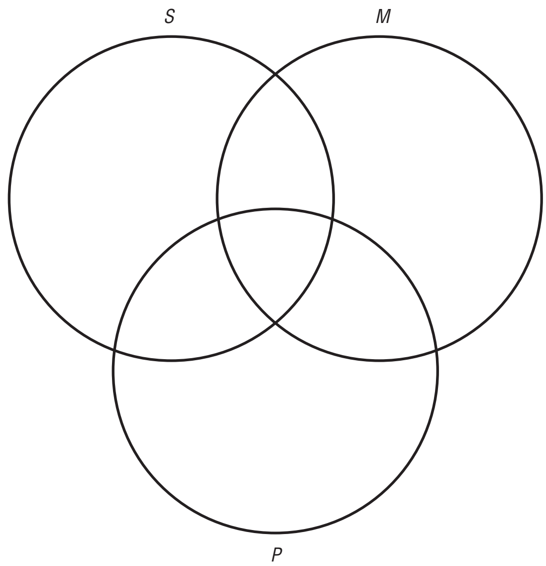

VENN DIAGRAMS

The Euler and Lambert methods, as well as later variants using squares and other types of closed curves, are no longer in use because of the great improvement on their basic conception which was introduced by the English logician John Venn. The Venn diagram is best explained by showing how it is used to validate a syllogism. The syllogism's three terms, *S, M,* and *P,* are represented by simple closed curves—most conveniently drawn as circles—that mutually intersect, as in Figure 1. The set of points inside circle *S* represents all members of class *S*, and points outside are members of class not-*S*—and similarly for the other two circles. Shading a compartment indicates that it has no members. An X inside a compartment shows that it contains at least one member. An X on the border of two compartments means that at least one of the two compartments has members.

Consider the following syllogism:

Some *S* is *M*.

All *M* is *P*.

Therefore, some *S* is *P*.

The first premise states that the intersection of sets *S* and *M* is not empty. This is indicated by an X on the bor-

FIGURE 2

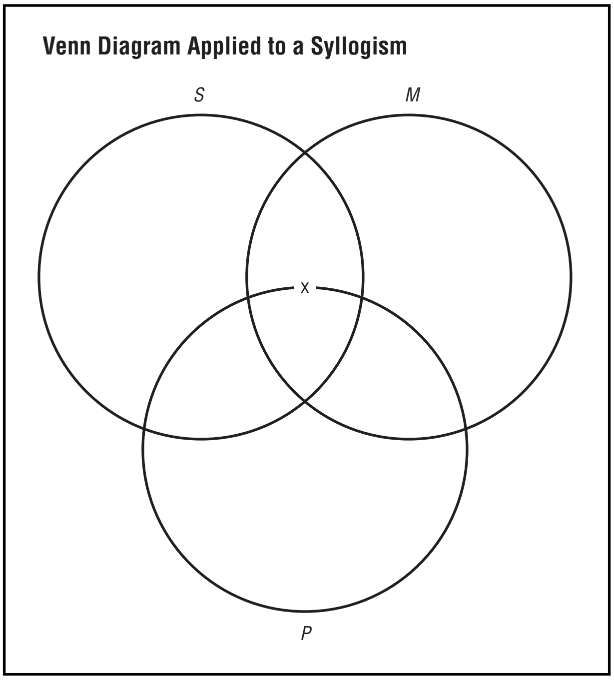

FIGURE 3

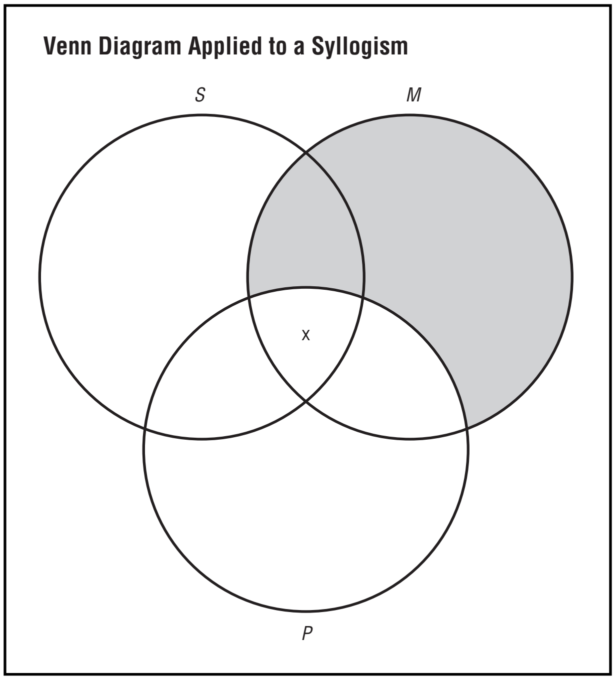

der dividing the two compartments within the overlap of circles *S* and *M* (Figure 2). The second premise states that the set indicated by that portion of circle *M* that lies outside of *P* is empty. When this area is shaded (Figure 3) the X must be shifted to the only remaining compartment into which it can go. Because the X is now inside both *S* and *P*, it is evident that some *S* is *P*; therefore, the syllogism is valid.

Venn did not restrict this method to syllogisms. He generalized it to take care of any problem in the calculus of classes, then the most popular interpretation of what is now called Boolean algebra. For statements with four terms he used four intersecting ellipses, as shown in Figure 4. Since it is not possible for five ellipses to intersect in the desired manner, statements with five or more terms must be diagramed on more complicated patterns. Various methods of forming nonconvex closed curves for Venn diagrams of statements with more than four terms have been devised.

RECTANGULAR CHARTS

Statements involving a large number of terms are best diagramed on a rectangle divided into smaller rectangles that are labeled in such a way that the chart can be manipulated efficiently as a Venn diagram. Many different methods of constructing such charts were worked out in the late nineteenth and early twentieth centuries, each with merits and defects. The first to be published was that of Allan Marquand in 1881. Figure 5 shows a Marquand chart for four terms. Alexander Macfarlane preferred a narrow strip, which he called a "logical spectrum," subdivided and labeled as in Figure 6. Later, in "Adaptation of the Method of the Logical Spectrum to Boole's Problem" (in *Proceedings of the American Association for the Advancement of Science* 39 [1890]: 57f.), Macfarlane used his chart for solving a complicated problem in George Boole's *Laws of Thought* (1854).

Other types of rectangular charts were devised by William J. Newlin, William E. Hocking, and Lewis Carroll. Carroll introduced his chart in a book for children, *The Game of Logic* (London and New York, 1886). Instead of shading compartments, he proposed marking them with counters of two colors, one for classes known to have members, the other for null classes.

An elaborate diagrammatic method designed to cover all types of logic, including modal logics, was devised in 1897 by the American philosopher Charles Sanders Peirce and later discussed in several brief, obscurely written papers. Although Peirce considered these "existential graphs," as he called them, his greatest contribution to logic, they aroused little interest among later logicians and have yet to be fully explicated and evaluated.

FIGURE 4

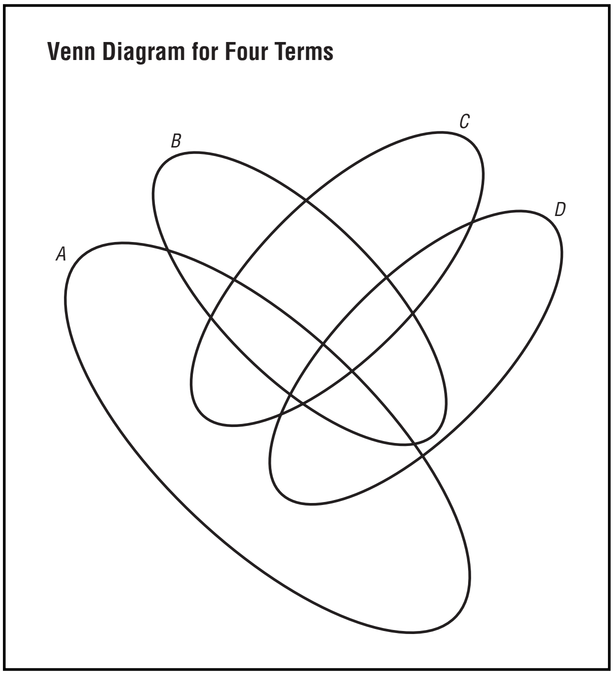

FIGURE 5

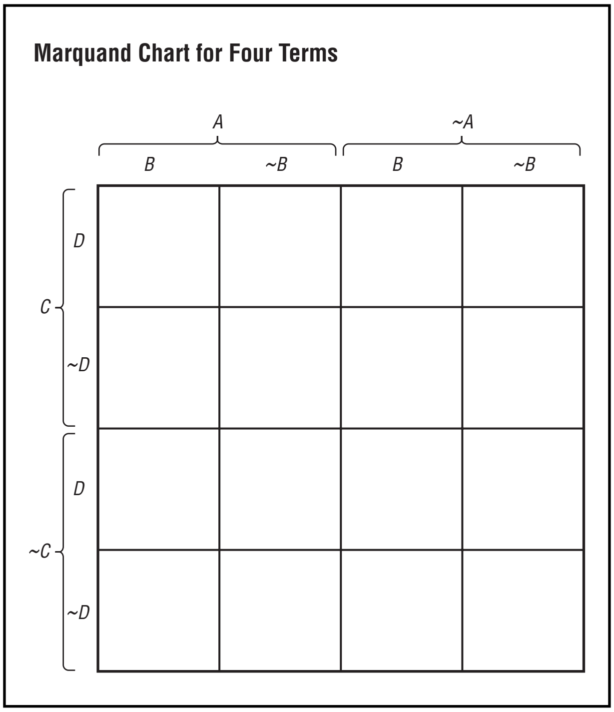

DIAGRAMS FOR THE PROPOSITIONAL CALCULUS

In the early twentieth century the class interpretation of Boolean algebra was supplemented by a more useful interpretation in which classes are replaced by propositions that are either true or false and related to one another by logical connectives. The Venn diagrams, as well as their chart extensions, work just as efficiently for the propositional calculus as for the class calculus, but cultural lag has prevented this fact from entering most logic textbooks. For example, the class statement "All apples are red" is equivalent to the propositional statement "If *x* is an apple, then *x* is red." The same Venn diagram is therefore used for both statements (Figure 7). Similarly, the class statement "No *A* is *B*" is equivalent to the propositional statement "Not both *A* and *B*," symbolized in modern logic by the Sheffer stroke. Both statements are diagramed as in Figure 8.

A major defect of the Venn system is that it is difficult to distinguish the shading of one statement from the shading of another, so that one loses track of individual premises. This is best remedied by diagramming each statement on a separate sheet of transparent paper and superposing all sheets on the same basic diagram. Such a method using cellophane sheets shaded with different colors and superposed on a rectangular diagram was devised by Karl Döhmann, of Berlin.

A network method for solving problems in the propositional calculus, designed to keep statements separate and to bring out visually the nature of the logical connectives, is given in Martin Gardner's *Logic Machines and Diagrams* (1958), Chapter 3. Each term is represented by two vertical lines, one for "true," the other for "false." A connective is symbolized by "shuttles" that connect truth-value lines in the manner indicated by the "true" lines of a truth table for that connective. Figure 9 shows the diagram for implication.

A "Boole table," devised by Walter E. Stuerman, also keeps individual statements separate and can be used for graphing any type of Boolean algebra. It combines features of Macfarlane's chart with Lambert's linear method. John F. Randolph has developed a simple method of handling a Marquand diagram by sketching nested cross marks and using dots to indicate nonempty compartments.

Although the Venn circles and their various chart extensions can obviously be given three-dimensional forms, no three-dimensional techniques for diagramming Boolean algebra have been found useful because of the extreme difficulty of manipulating solid diagrams. In this connection, however, mention should be made of a curious cubical chart, devised by C. Howard Hinton in 1904, that is constructed with 64 smaller cubes and used for identifying valid syllogisms.

FIGURE 6

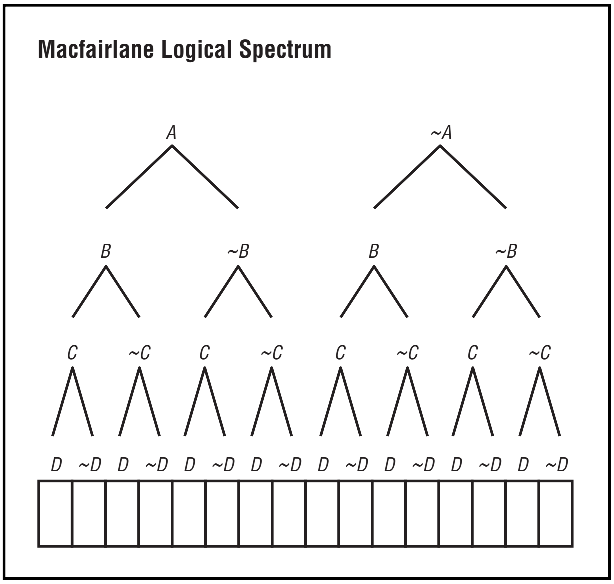

FIGURE 7

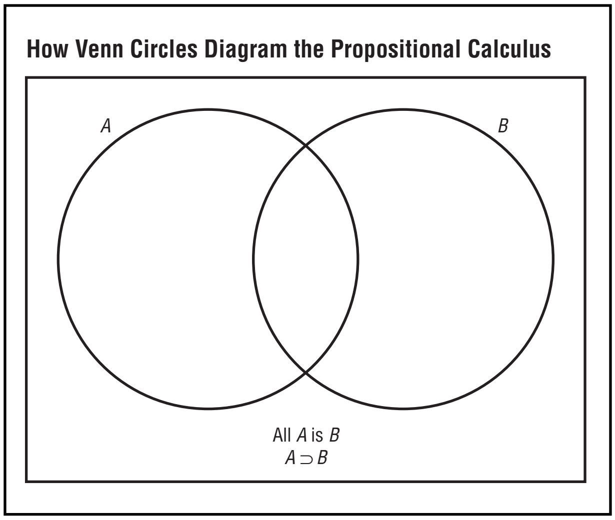

Boolean algebra is now known to be a special type of lattice, which in turn is a certain type of partially ordered set. A lattice diagram for a Boolean algebra of two terms is easily drawn, and although of little use in problem solving, it displays graphically many features of the propositional calculus.

In the logic of relations a large variety of useful diagrams have been widely used. The tree graph, for example, which goes back to ancient Greece, is an efficient way to indicate a familiar type of relation. Examples include the tree of Porphyry, found in medieval and Renaissance logics, the later tree diagrams of Peter Ramus, diagrams showing the evolution of organisms, family tree graphs, and graphs of stochastic processes in probability theory. The topological diagrams in Kurt Lewin's *Principles of Topological Psychology* (1936), as well as modern "sociograms," transport networks, and so on, may be called logic diagrams if "logic" is taken in a broad sense. However, such diagrams are now studied in the branch of mathematics called graph theory and are not generally considered logic diagrams. In a wide sense any geometrical figure is a logic diagram since it expresses logical relations between its parts.

AREAS FOR EXPLORATION

All diagrams for Boolean algebras work most efficiently when the statements to be diagramed are simple binary relations. Compound statements with parenthetical expressions are awkward to handle unless the statements are first translated into simpler expressions. Attempts have been made to extend the Venn diagrams and other types of Boolean graphs to take care of parenthetical statements directly, but in all cases the diagrams become too complex to be useful. Perhaps simpler methods will be found by which traditional diagrams can be made to accommodate parenthetical expressions.

Little progress has been made in developing good diagrammatic methods for minimizing a complex logical statement—that is, for reducing it to a simpler but equivalent form. Several chart methods for minimizing have been worked out. The closest to a diagrammatic technique is the Karnaugh map, first explained by Maurice Karnaugh in 1953. The map is based on an earlier diagram called the Veitch chart, in turn based on a Marquand chart.

Work on better methods of minimizing is still in progress. The work has important practical consequences because electrical networks can be translated into Boolean algebra and the expression minimized and then translated back into network design to effect a simplification of circuitry. It is possible that a by-product of new minimizing methods may be a diagrammatic method superior to any yet found.

Another field open to exploration is the devising of efficient ways to diagram logics not of the Boolean type, notably modal logics and the various many-valued logics.

See also Boole, George; Carroll, Lewis; Geometry; Hocking, William Ernest; Lambert, Johann Heinrich; Leibniz, Gottfried Wilhelm; Logic, History of; Logic Machines; Peirce, Charles Sanders; Porphyry; Ramus, Peter; Renaissance; Venn, John.

FIGURE 8

How Venn Circles Diagram the Propositional Calculus

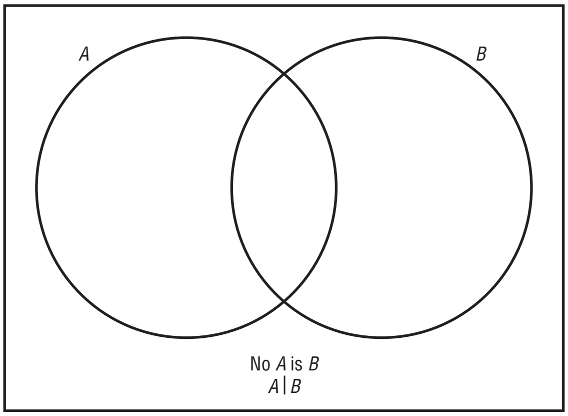

FIGURE 9

A Network Diagram of Implication

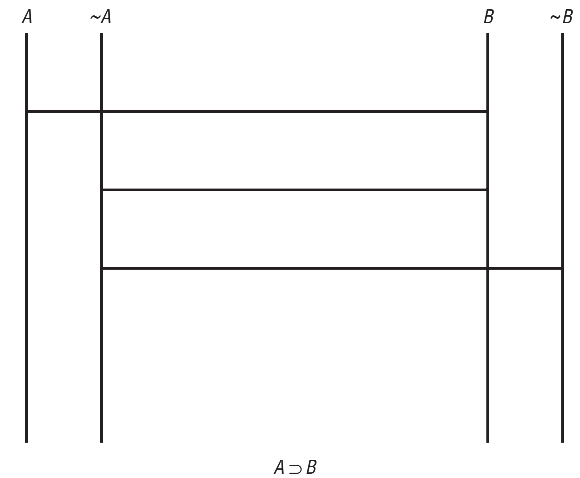

Bibliography

Venn's *Symbolic Logic,* rev. 2nd ed. (London: Macmillan, 1894), remains the best single source for the early history of logic diagrams. It also contains the fullest exposition of his own method, which he first presented in "On the Diagrammatic and Mechanical Representation of Propositions and Reasonings," *Philosophical Magazine* 10 (July 1880): 1–18. For more recent history and references, consult Martin Gardner, *Logic Machines and Diagrams* (New York: McGraw-Hill, 1958).

Euler's explanation of his circles can be found in his *Lettres à une princesse d'Allemagne,* Vol. II, letters 102–108 (St. Petersburg: Academie Impériale des Sciences, 1768). Lambert's method, in an improved form, is explained in J. N. Keynes, *Studies and Exercises in Formal Logic.* 4th ed. (London: Macmillan, 1906), p. 243.

For Venn diagrams of statements with more than four terms, see W. E. Hocking, "Two Extensions of the Use of Graphs in Elementary Logic," in *University of California Publications in Philosophy* 2 (2) (1909): 31–44; Edmund C. Berkeley, "Boolean Algebra and Applications to Insurance," in *Record,* American Institute of Actuaries 26 (1937): 373–414, and 27 (1938): 167–176 (reprinted, New York, 1952); Trenchard More Jr., "On the Construction of Venn Diagrams," in *Journal of Symbolic Logic* 24 (1959): 303–304; and Stephen Barr, *Experiments in Topology* (New York: Crowell, 1964), p. 206. A method of drawing five irregular but congruent convex pentagons intersecting in the required manner is explained by David W. Henderson in "Venn Diagrams for More Than Four Classes," *American Mathematical Monthly* 70 (1963): 424–426.

On rectangular Venn charts, see Allan Marquand, "A Logical Diagram for *n* Terms," in *Philosophical Magazine* 12 (1881): 266–270; Alexander Macfarlane, "The Logical Spectrum," in *Philosophical Magazine* 19 (1885): 286–289; William J. Newlin, "A New Logical Diagram," in *Journal of Philosophy, Psychology, and Scientific Methods* 3 (1906): 539–545; and the Hocking paper mentioned above. Carroll's method is further developed in his *Symbolic Logic* (London and New York, 1896; paperback ed., New York, 1958). Peirce's papers are reprinted in the *Collected Papers of Charles Sanders Peirce,* edited by Charles Hartshorne and Paul Weiss, Vol. IV (Cambridge, MA: Harvard University Press, 1933).

An explanation of how Venn circles can be used to handle problems in the propositional calculus appears in Gardner's *Logic Machines and Diagrams* (op. cit.), pp. 48–54. Döhmann explained his method in a privately printed booklet, *Eine logistische Farbenquadrat-Methode* (Berlin, 1962). Stuerman described his "Boole table" in "Plotting Boolean Functions," *American Mathematical Monthly* 67 (1960): 170–172, and in "The Boole Table Generalized," *American Mathematical Monthly* 68 (1961): 53–56. Randolph's method is given in "Cross-Examining Propositional Calculus and Set Operations," *American Mathematical Monthly* 72 (1965): 117–127. Hinton explained his cubical chart in his eccentric book *The Fourth Dimension* (London: Sonnenschein, 1904), pp. 100–106. John Evenden presented a two-term lattice diagram for Boolean algebra in "A Lattice-Diagram for the Propositional Calculus," *Mathematical Gazette* 46 (1962): 119–122.

Karnaugh first explained his map in "The Map Method for Synthesis of Combinational Logic Circuits," *Transactions of the American Institute of Electrical Engineers,* Part 1, 72 (1953): 593–599. E. W. Veitch explained his earlier diagram in "A Chart Method for Simplifying Truth Functions," *Proceedings of the Association for Computing Machinery* (May 2 and 3, 1952), 127–133.

Martin Gardner (1967)

LOGICISM

See *Mathematics, Foundations of*

LOGIC MACHINES

Because logic underlies all deductive reasoning, one might say that all computers are logic machines. In a wider sense, any mechanical device is a logic machine (for example, an eggbeater spins clockwise "if and only if" its crank turns clockwise). Generally, however, the term is restricted to machines designed primarily or exclusively for solving problems in formal logic. Although a digital computer, or even a punch-card data-processing machine, can be programmed to handle many types of logic, it is not considered a logic machine in the strict sense.

The rotating circles of Ramón Lull, thirteenth-century Spanish mystic, cannot be called logic machines even though they were used as reasoning aids. The first true logic machine was a small device called a "demonstrator," invented by Charles Stanhope, third Earl Stanhope, an eighteenth-century English statesman. By sliding two panels (one of gray wood, the other of transparent red glass) behind a rectangular opening, he could test the validity of traditional syllogisms, as well as syllogisms with such quantified terms as "Most of *a*" and "8 of 10 of *a*." Stanhope also used his device for solving elementary problems in what he called the logic of probability.

JEVONS'S MACHINE

The first logic machine capable of solving a complicated problem faster than a human could solve it without the aid of a machine was the "logical piano" invented by the nineteenth-century economist and logician William Stanley Jevons. The machine was built for him by a clockmaker at Salford in 1869 and first demonstrated by Jevons in 1870 at a meeting of the Royal Society of London. The device (now owned by the Oxford Museum of the History of Science) resembles a miniature upright piano, about three feet high, with a keyboard of 21 keys. On the face of the piano are openings through which one can see the 16 possible combinations of 4 terms and their negatives. A statement in logic is fed to the machine by pressing keys according to certain rules. Internal levers and pulleys eliminate from the machine's face all combinations of terms inconsistent with the statement. When all desired statements have thus been fed to the machine the face is inspected to determine what term combinations, if any, are consistent with the statements.

Jevons believed that this machine, designed to handle Boolean algebra, provided a convincing demonstration of the superiority of George Boole's logic over the traditional logic of Aristotle and the Schoolmen. John Venn's system of diagramming follows essentially the same procedure as Jevons's machine. In both cases the procedure gives what are today called the valid lines of a truth table for the combined statements under consideration. Neither the Venn diagrams nor Jevons's machine is capable of reducing these lines to a more compact form. This criticism of the machine was stressed by the English philosopher F. H. Bradley in his *Principles of Logic* (1883).

OTHER MECHANICAL DEVICES

Jevons's logical piano was greatly simplified by Allan Marquand, who built his first model in 1881, when he was teaching logic at Princeton University. Like Jevons's, Marquand's machine is limited to 4 terms, but the 16 possible combinations are exhibited on its face by 16 pointers, each with a valid and an invalid position, arranged in a pattern that corresponds to Marquand's chart for 4 terms (see the entry "Logic Diagrams," Figure 5). The number of keys is reduced to 10, and the device is about a third the height of Jevons's machine. Both Marquand and Jevons interpreted Boolean algebra primarily in class terms, but their machines operate just as efficiently with the propositional calculus.

A third machine of the Jevons type was invented in 1910 by Charles P. R. Macaulay, an Englishman living in Chicago. It is a compact, ingenious boxlike device with interior rods operated by tilting the box a certain way while pins on the side are pressed to put statements into the machine. Consistent combinations of four terms and their negatives appear in windows on top of the box.

A curious contrivance for evaluating the 256 combinations of syllogistic premises and conclusions was constructed in 1903 by Annibale Pastore, a philosopher at the University of Genoa. It consists of three wheels, representing a syllogism's three terms, joined to one another by an arrangement of endless belts appropriate to the syllogism being tested. If the syllogism is valid, all three wheels turn when one is cranked.

GRID CARDS

Logic grid cards are cards that can be superposed so that valid deductions from logical premises are seen through openings on the cards. A set of syllogism grid cards invented by the Englishman Henry Cunynghame, a contemporary of Jevons, was depicted by Jevons in Chapter 11 of *Studies in Deductive Logic* (London, 1884). A differently designed set is shown in Martin Gardner's "Logic Machines" (in *Scientific American* 186 [March 1952]: 68–73). A more elaborate set, indicating the nature of the

fallacy when a syllogism is invalid, can be found in Gardner's *Logic Machines and Diagrams* (New York, 1958) and Richard Lampkin's *Testing for Truth* (Buffalo, NY, 1962). Triangular-shaped grid cards, for binary relations in the propositional calculus, are described in Gardner's book and in H. M. Cundy and A. P. Rollett's *Mathematical Models* (2nd ed., Oxford, 1961; see pp. 256–258). Gardner described a simple way to make punch cards that can be sorted in such a manner as to solve logic problems in "Mathematical Games" (in *Scientific American* 203 [December 1960]: 160–168).

ELECTRICAL MACHINES

Marquand sketched an electrical circuit by which his machine could be operated, but the electrical version was probably never built. Benjamin Burack, a psychologist at Roosevelt College, Chicago, was the first actually to construct an electrical logic machine, in 1936. His device tested all syllogisms, including hypothetical and disjunctive forms. Since then many different kinds of electrical syllogism machines have been constructed.

In 1910, in a review in a Russian journal, Paul Ehrenfest pointed out that because a wire either carries a current or does not, it would be possible to translate certain types of switching circuits into Boolean algebra. Work along such lines was done by the Russian physicist V. I. Šestakov in 1934–1935, but his results were not published until 1941. Similar views were set forth independently in 1936, in a Japanese journal, by Akira Nakasima and Masao Hanzawa. It was the mathematician Claude E. Shannon, however, who impressed the engineering world with the importance of this isomorphism by his independent work, first published in 1938.

Shannon's paper inspired William Burkhart and Theodore A. Kalin, then undergraduates at Harvard University, to design the world's first electrical machine for evaluating statements in the propositional calculus. The Kalin-Burkhart machine was built in 1947. Statements with as many as twelve terms are fed into it by setting switches. The machine scans a truth table for the combined statements, and a set of twelve small bulbs indicates the combination of true and false terms for each truth-table row as it is scanned. If the combination is consistent with the statements, this is indicated by another bulb. The machine is thus an electrical version of Jevons's device but handles more complex statements and presents valid truth-table rows in serial time sequence rather than simultaneously.

A three-term electrical machine was built in England in 1949 without knowledge of the Kalin-Burkhart machine. Advances in switching components made possible more sophisticated logic machines in the United States and elsewhere during the early 1950s. Of special interest is a ten-term machine built at the Burroughs Research Center in Paoli, Pennsylvania, using the parenthesis-free notation of Jan Łukasiewicz.

DIGITAL COMPUTERS

While the special machines were being developed it became apparent that statements in Boolean algebra could easily be translated into a binary notation and analyzed on any general-purpose digital computer. As digital computers became more available, as well as faster and more flexible, interest in the design of special-purpose logic machines waned. Since 1955 almost all machine-aided investigations in logic have been conducted with digital computers. In 1960, Hao Wang described how he used an IBM 704 computer to test the first 220 theorems of the propositional calculus in *Principia Mathematica.* The machine's total running time was under three minutes.

The similarity between switching circuits and the nets of nerve cells in the brain suggests that the brain may think by a process that could be duplicated by computers. Much work is being done in programming computers to search for proofs of logic theorems in a manner similar to the heuristic reasoning of a logician—that is, by an uncertain strategy compounded of trial and error, logical reasoning, analogies with remembered experience, and sheer luck. The work is closely related to all types of learning machines. Such work may prove useful in exploring logics for which there is no decision procedure—or no known decision procedure—but no special machines have yet been built for such a purpose. Work is also under way on the more difficult problem of designing a machine, or programming a digital computer, to find new, nontrivial, and interesting theorems in a given logic.

Attempts have been made to design machines capable of reducing a statement in Boolean algebra to simpler form. A primitive minimizing machine was constructed by Daniel Bobrow, a New York City high school student, in 1952. At about the same time, Shannon and Edward F. Moore built a relay circuit analyzer that makes a systematic attempt to simplify circuits, a problem closely related to the logic minimizing problem.

No special machines are known to have been constructed for handling many-valued logics, but many papers have been published explaining how such machines could be built, as well as how digital computers

could be programmed to handle such logics. Kurt Gödel's undecidability proof has ruled out the possibility of an ultimate logic machine capable of following a systematic procedure for testing any theorem in any possible logic, but whether the human brain is capable of doing any kind of creative work that a machine cannot successfully imitate is still an open, much debated question.

See also Aristotle; Boole, George; Bradley, Francis Herbert; Computing Machines; Gödel, Kurt; Gödel's Incompleteness Theorems; Jevons, William Stanley; Logic, History of; Łukasiewicz, Jan; Lull, Ramón; Machine Intelligence; Venn, John.

Bibliography

For a history of logic machines, see Martin Gardner, *Logic Machines and Diagrams* (New York: McGraw-Hill, 1958), and Rudolf Tarján, "Logische Maschinen," in *Digital Information Processors,* edited by Walter Hoffmann (New York: Interscience, 1962), which includes a bibliography of 119 references. Stanhope's machine is discussed in Robert Harley, "The Stanhope Demonstrator," *Mind* 4 (1879): 192–210.

Jevons described his logical piano in "On the Mechanical Performance of Logical Inference" (1870), reprinted in his *Pure Logic and Other Minor Works* (London: Macmillan, 1890); a briefer discussion appears in his *Principles of Science* (London: Macmillan, 1874; reprinted, New York, 1958). Marquand discussed his machine in "A New Logical Machine," *Proceedings of the American Academy of Arts and Sciences* 21 (1885): 303–307. Marquand's machine was praised by C. S. Peirce in "Logical Machines," *American Journal of Psychology* 1 (November 1887): 165–170. Pastore wrote an entire book about his device, *Logica formale, dedotta della considerazione di modelli meccanici* (Turin, 1906). The book was reviewed at length by André Lalande in *Revue philosophique de la France et de l'étranger* 63 (1907): 268–275.

Wolfe Mays, "The First Circuit for an Electrical Logic-Machine," in *Science* 118 (1953): 281–282, and George W. Patterson, "The First Electric Computer, a Magnetological Analysis," in *Journal of the Franklin Institute* 270 (1960): 130–137, describe Marquand's sketches. A description of Macaulay's device is in his U.S. patent, No. 1,079,504, issued in 1913.

Burack first described his 1936 machine in "An Electrical Logic Machine," *Science* 109 (1949): 610–611. Shannon's historic paper on the isomorphism of Boolean algebra and certain types of switching circuits is "A Symbolic Analysis of Relay and Switching Circuits," in *Transactions of the American Institute of Electrical Engineers* 57 (1938): 713–723. The Kalin-Burkhart machine is described in Edmund C. Berkeley, *Giant Brains* (New York: Wiley, 1949). The English three-term machine is discussed in Wolfe Mays and D. G. Prinz, "A Relay Machine for the Demonstration of Symbolic Logic," *Nature* 165 (1950): 197. The Burroughs machine is discussed in A. W. Burks, D. W. Warren, and J. B. Wright, "An Analysis of a Logical Machine Using Parenthesis-Free Notation," *Mathematical Tables and Other Aids to Computation* 8 (1954): 53–57, and in William Miehle, "Burroughs Truth Function Evaluator," *Journal of the Association for Computing Machinery* 4 (1957): 189–192.

Hao Wang's description of his use of the IBM 704 is in "Toward Mechanical Mathematics," *IBM Journal of Research and Development* 4 (1960): 2–22, reprinted in *The Modeling of Mind,* edited by Kenneth Sayre and Frederick J. Crosson (Notre Dame, IN: University of Notre Dame Press, 1963). On computer simulation, see Allen Newell and H. A. Simon, "Computer Simulation of Human Thinking," in *Science* 134 (1961): 2011–2017, and references cited there. On the Shannon-Moore relay circuit analyzer, see Shannon and Moore's "Machine Aid for Switching Circuit Design," in *Proceedings of the Institute of Radio Engineers* 41 (1953): 1348–1351.

Martin Gardner (1967)

LOGOS

The Greek term *logos* is multiply ambiguous. The unabridged Greek dictionary gives five and a half long columns of definitions and examples. *Logos* is a noun corresponding to the verb *legein* (say), signifying, among other things, speech, statement, sentence, account, definition, formula, calculation, ratio, explanation, reasoning, and faculty of reason. Early studies of the term tended to talk about a concept of *logos*, as if there were some single concept or theory associated with it. In fact, the term was employed in different ways by different thinkers. Yet, there is a kind of interplay in concepts associated with the term that makes a single study worthwhile.

Scholars sometimes speak of a change from *mythos* to *logos*; roughly, a transition in expression from storytelling in myths, usually expressed in poetry, to scientific, philosophical, or historical accounts, usually expressed in prose. Philosophers of the sixth century BCE were among the first Western writers to compose treatises in prose. The new medium of expression permitted a more analytic and detached view of things, and it embodied a revolution in thinking about the world. Although *logos* (plural: *logoi*) could signify a story, increasingly *logoi* were taken to be scientific accounts in contrast to *mythoi* "stories" and *epea* "verses" (see Plato *Timaeus* 26e). But for the sophists, a *mythos* can be used to express a *logos* (Plato *Protagoras* 320c)—but only insofar as *logos* is seen as a more basic kind of explanation.

THE PRESOCRATICS

Logos soon came to signify something of the content of rational discourse as well as the medium, and it is this sense, or set of senses, that this entry will focus on. Hera-

clitus (c. 500 BCE) was the first philosopher to raise *logos* to the level of a principle. He opens his book by saying, "Of this *Logos*'s being forever do men prove to be uncomprehending, both before they hear and once they have heard it. For although all things happen according to this *Logos* they are like the unexperienced experiencing words and deeds such as I explain when I distinguish each thing according to its nature and show how it is" (fr. 1). Heraclitus's *logos* can be shared with people, and indeed he explicates it in his own treatise; but he anticipates that most people will fail to understand the message. "Although this *Logos* is common," Heraclitus writes, "the many live as if they had a private understanding" (fr. 2). Somehow the *logos* is publicly available but ignored by the many, who lack philosophical insight. The *logos* has a particular message, or implication: "Listening not to me but to the *Logos* it is wise to agree that all things are one" (fr. 50). Heraclitus regards the *logos* as transcending his own personal communication, and teaching the unity of things.

Heraclitus's *logos* is a kind of structural principle as well as a message, a reciprocal law of exchange. It has a kind of syntax like language that orders the changes of the world. Heraclitus plays with statements that are syntactically ambiguous, as if to show that the same words can make different statements, which at another level complement each other. So the world is based on a single structure that manifests itself in contraries. Language provides a model for the world.

In the early fifth century BCE, Parmenides presented an argument against change, in the form of a revelation from a goddess. Yet the goddess tells the narrator, "Judge by *logos* the contentious refutation spoken by me" (fr. 7). Here *logos* seems to mean something like *reasoning*, which clearly becomes the key to philosophical truth. For, despite the religious imagery and associations of his poem, Parmenides's message is above all an argument addressed to the reason.

In the latter half of the fifth century BCE the sophists traveled about Greece teaching practical skills to help young men succeed in politics and, above all, the art of public speaking. They saw a knowledge of *logos*—and especially, for them, the spoken word—as the key to controlling emotions and hence the reactions of audiences to a message. As Gorgias observed, "*Logos* is a great potentate, who by means of the tiniest and most invisible body is able to achieve the most godlike results" (fr. 11, section 8). Sophists composed contradictory arguments (*antilogikoi logoi*) on a single topic to teach skill in argumentation, and sometimes studied elements of language and argumentation.

PLATO AND ARISTOTLE

By the fourth century BCE *logos* is established not only as speech and the like, but as the faculty of reason. Speech becomes the manifestation of reason, and reason the source of speech. According to Plato an understanding of rhetoric presupposes a knowledge of souls—what would later be called psychology—and the use of dialectic to implant truth in souls (*Phaedrus*). In fact, thinking (*dianoia*) is just internal speech (*Sophist* 263e, *Theaetetus* 189e). Thus speech becomes a model for thought, and ultimately a representation for the world; for a sentence (*logos*), such as "Theaetetus is sitting," is true just in case it correctly describes an action or condition of Theaetetus (*Sophist* 263a–b). In another context, Plato suggests that one can more safely study the world in *logoi* than by means of sensations, and he consequently adopts a method of hypothesis (*Phaedo* 99d–100a).

The sign that one has knowledge is one's ability to give an account (*logos*) or explanation (*Phaedo* 76b), and one who can give an adequate account is a dialectician (*Republic* 534b). At one point Plato considers as a definition of *knowledge* "true judgment accompanied by an account [*logos*]," but rejects this in part because a satisfactory explanation of *logos* cannot be given independently of knowledge (*Theaetetus* 201c ff.). While the ability to give a rational account provides evidence of knowledge, the account is no mere component of knowledge.

Aristotle accepts Plato's view of the relation between language and the world along with some of Plato's terminology (*Categories* 2–4; *On Interpretation* 1–7). He recognizes, if somewhat obscurely, the two relationships that allow language to connect to reality: reference (*semainein*) and predication (*katêgoria*)—the latter primarily a link between a substance and its attributes, but mirrored in the link between grammatical subject and predicate. The basic unit of communication is the sentence (*logos*), which when it makes an assertion (*apophantikos*) is the bearer of truth or falsity. Whereas reference connects words with things, (grammatical) predication asserts that the things are connected in a certain way; if the assertion corresponds to the way things are, it is true; otherwise it is false. Building on this basic theory of language, Aristotle developed the first system of logic, showing how certain propositions follow logically from certain other propositions (*Prior Analytics*). Moreover, he conceived of a science as a set of propositions arranged in a logical order with axioms and definitions as starting points, and theorems as conclusions (*Posterior Analytics* I)—laying out this ideal structure that would be realized by the axiomatization of geometry a generation

or two after his death. Thus in a certain sense Aristotle saw the world as possessing a thoroughgoing logical structure that could be captured in language. Indeed, whereas contemporary logicians often think of logical systems as arbitrary human constructs, some of which are useful for capturing certain linguistic relationships, Aristotle thinks of his logic as having its basis in the nature of things (*Prior Analytics* I 27).

HELLENISTIC PHILOSOPHY

According to the Stoics, the world is ultimately composed of fire, which is identical with God. Fire pervades the world and functions as a world-soul. Reason (*logos*) is found in the world-soul, which orders and controls the world; it is the active principle and is identical with God (Diogenes Laertius 7.134). Soul is found in all animals, and in humans there is also a ruling principle that possesses reason. Thus *logos* in the human mind is like *logos* in the cosmos. Through the activity of fire, reason controls the creation and the history of the world. The world periodically perishes in a conflagration that turns all the elements back into fire, from which a new world arises, seeded by seminal *logos*, a structural principle that directs the cosmogony (Diogenes Laertius 136). The events of the world are ultimately under the control of reason, so that the world is governed by providence (Diogenes Laertius 138–9). The Stoics distinguish between uttered discourse (*prophorikos logos*) and internal discourse (*endiathetos logos*); the former humans have in common with parrots, but the latter is peculiar to humans (Sextus Empiricus *Against the Professors* 8.275).

Philo of Alexandria (early to mid-first century CE), combining Judaism and Platonism by using Plato's theory to explicate the Bible, recognizes *logos* as an image of the invisible God, and human beings as created in the image of the *logos* (*On Dreams* 239, *The Confusion of Tongues* 147). God also acts by his word, for "His word is his deed" (*The Sacrifices of Able and Cain* 65). The world is itself the product of a plan in the mind of God, consisting of the Platonic Forms (*On the Creation* 17–19), which are thus conceived of as present in the mind of God. From this model of the world the creator makes first an invisible world, then a visible one (29–36).

CHRISTIANITY AND NEOPLATONISM

The Gospel according to John begins by affirming the central role of the *Logos*, or Word: "In the beginning was the Word, and the Word was with God, and the Word was God. … All things were made by him; and without him was not any thing made that was made. … And the Word was made flesh, and dwelt among us" (John 1.1, 3, 14). It may be that the *Logos* of John derives from Jewish rather than Greek conceptions, yet the notion was close enough to Greek philosophical conceptions to allow early Christian thinkers to see in it a point of contact between their scriptures and pagan philosophy. They saw Philo as an inspired writer who shared their vision of the Word of God as an intermediary between God and humans. Jesus of Nazareth was the Word of God, who manifested the power of God on earth and prepared the way for his disciples to become sons of God (1.12).

In the mid-second century Justin Martyr identifies Jesus as the *Logos* that wise men, including philosophers, partake of. He finds references to the *Logos* in Plato's *Timaeus*, and more general instances of divine reason in Heraclitus and the Stoics (*First Apology* 5, 40; *Second Apology* 8, 10, 13). He explains that Christians "call Him [Jesus] the Word, because He carries tidings from the Father to men: but maintain that this power is indivisible and inseparable from the Father" (*Dialogue with Trypho* 128). In the most systematic statement of the early church fathers, Origen (third century), commenting on the opening lines of Hebrews, says that Jesus as Word is the invisible image of the invisible God—apparently apprehensible only by reason—who "interpret[s] the secrets of wisdom, and the mysteries of knowledge, making them known to the rational creation" (*On Principles* 1.2.6–7).

Plotinus borrowed from the Stoics at least the general conception of *logos* in a seed to account for the influence Soul has on the visible world. The world "was ordered according to a rational principle [*kata logon*] of soul potentially having throughout itself power to impose order according to rational principles [*kata logous*], just as the principles in seeds shape and form living creatures like little worlds" (*Enneads* 4.3.10). This also helps one understand how Mind orders things by comparing its operation to that of a seed with a rational principle; in such a way reason (*logos*) flows out from Mind to the world (3.2.2). And one can understand how timeless realities have foresight over the world of change by supposing that events unfold according to an archetype, which is effortlessly realized by the imposition on matter of rational principles (4.4.12). Indeed, Plotinus proclaims in a theodicy, "The origin [of events in the world] is *logos* and all things are *logos*," even if they seem to be irrational or evil to our limited view (3.2.15).

In the early fifth century Augustine argued that a word in the heart precedes the articulate word of speech. This inner word is a likeness of the Word of God, by whom God carried out the creation of the world, and

which came to be embodied in flesh in a way analogous to that in which the inner word becomes articulated in language. Thus the preverbal cognition that humans have in themselves an image of the Word of God (*On the Trinity* 15.11.20).

Although in Greek philosophy many different versions of how language, reason, and rational principles connect with the world can be found, what is remarkable is the widespread commitment to some view whereby reason is imbedded in the cosmos. Human reason does not simply impose some extraneous order on the world, but it discovers in nature a structure that mind has in common with the world.

See also Aristotle; Augustine, St.; Diogenes Laertius; Hellenistic Thought; Heraclitus of Ephesus; Neoplatonism; Parmenides of Elea; Patristic Philosophy; Philo Judaeus; Plato; Platonism and the Platonic Tradition; Plotinus; Semantics, History of; Sextus Empiricus; Sophists; Stoicism.

Bibliography

Aall, Anathon. *Geschichte der Logosidee in der griechischen Philosophie*. 2 vols. Leipzig, Germany: O. R. Reisland, 1896–1899. Reprint, Frankfurt, Germany: Minerva, 1968.

Boeder, Heribert. "Der fr ühgriechische Wortgebrauch von *Logos* und *Aletheia*." *Archiv f ür Begriffsgeschichte* 4 (1959): 82–112.

Frede, Michael, and Gisela Striker, eds. *Rationality in Greek Thought*. Oxford: Clarendon Press, 1996.

Kittel, Gerhard, et al. "Lego." In *Theologisches Wörterbuch zum Neuen Testament*, edited by Gerhard Kittel, vol. 4, 69–140. Stuttgart, Germany: W. Kohlhammer, 1942. Abridged in *Theological Dictionary of the New Testament*. Translated by Geoffrey W. Bromiley, 505–514. Grand Rapids, MI: W. B. Eerdmans, 1985.

Mourelatos, Alexander P. D. "Heraclitus, Parmenides, and the Naive Metaphysics of Things." In *Exegesis and Argument*, edited by E. N. Lee, A. P. D. Mourelatos, and R. Rorty. Assen, Netherlands: Van Gorcum, 1973.

Robb, Kevin, ed. *Language and Thought in Early Greek Philosophy*. La Salle, IL: The Hegeler Institute, 1983.

Daniel W. Graham (2005)

LOISY, ALFRED

(1857–1940)

Alfred Loisy, the French biblical exegetist, was the best-known and most controversial representative of the Modernist movement in France at the end of the nineteenth and beginning of the twentieth centuries. His scholarly investigation led him to the kind of destructive criticism of the Gospel narratives and Christian dogmas carried on earlier by such scholars as D. F. Strauss and Ernest Renan, whose lectures at the Institut Catholique Loisy attended from 1882 to 1885. Loisy's long career, from his entry into the priesthood in 1879 to shortly before his death, was one of much controversy and progressive estrangement from personal religion.

Loisy was born at Ambrière, Marne, and died at Ceffonds, Haute Marne. He became professor of Hebrew in 1881, and of Holy Scripture in 1889, at the Institut Catholique. Loisy's views on the date of the book of Proverbs soon aroused misgivings, and he was warned that continuation of such unorthodoxy would place him in danger of official censure.

Loisy's superior, Monsignor d'Hulst, was an enlightened man and not intolerant of the work of the modern critical school, but as head of the Institut Catholique he was in a responsible and difficult position. The head of the College of St. Sulpice had forbidden his students to attend the heterodox Loisy's lectures, and when in 1892 Loisy started his own periodical, *L'enseignement biblique,* for the instruction of young priests, d'Hulst felt obliged to urge caution. In 1892, soon after Renan's death, d'Hulst himself wrote an article on Renan in *Le correspondant.* Without condoning Renan's break with Catholicism, d'Hulst upheld his complaint, in *Souvenirs d'enfance et de jeunesse,* that the instruction given at such seminaries as St. Sulpice was out of touch with modern scholarship and the modern world. A further article by d'Hulst, aimed at promoting tolerance of the more searching kind of biblical criticism, gave offense in orthodox quarters, and d'Hulst felt obliged to clear his institute of any suspicion of unorthodoxy. Therefore, when Loisy continued to declare his critical independence of dogma and revelation, and to present a historical Jesus apart from the Christ of faith, he was forced to resign his chair in 1893.

As a reply to modernist exegesis, the pope issued the encyclical *Providentissimus Deus* (November 18, 1893), denying that error is compatible with divine authorship. Loisy wrote to Leo XIII, professing submission to the encyclical's demand that the truth of the Bible should not be questioned. His insincerity can be inferred, however, for his activities remained unchanged. In fact, on receiving a reply in a mollified tone that invited him to devote himself to less contentious studies, Loisy openly expressed his impatience.

Loisy criticized the Protestant scholar Carl Gustav Adolf von Harnack's *Wesen des Christentums* (Leipzig, 1900) in his *L'évangile et l'église* (Paris, 1902), which was condemned by the archbishop of Paris as undermining

faith in the authority of Scripture and the divinity of Jesus Christ. Loisy wrote an apology, *Autour d'un petit livre* (Paris, 1903), which, with four other works of his, was condemned by the Holy Office and placed on the Index in 1903. The papal secretary of state required the archbishop of Paris to demand that Loisy withdraw the five offending volumes, but Loisy refused.

He wrote in conciliatory terms to Pope Plus X, but the development of his religious ideas—or, in Catholic eyes, the disintegration of his faith—could ultimately lead only to his exclusion from the Roman communion. He regarded such mysteries as the incarnation of God as mere metaphors and symbols, and described his own religious belief as pantheistic, positivistic, or humanitarian rather than Christian. He conceived the basic problem facing the man torn between belief and doubt to be whether the world contains or embodies any spiritual principle apart from man's own consciousness.

In 1907 the papal secretary of state called upon Loisy to repudiate certain propositions, attributed to him and condemned in the decree *Lamentabili* (July 2, 1907), and to disown Modernism, condemned in Plus X's encyclical *Pascendi Dominici Gregis* (September 6, 1907). Loisy replied that where his views were not misrepresented in the decree, he felt obliged to stand by them, since he regarded them as true. The demands were repeated, and Loisy was required to submit within ten days. He still refused and was thereupon excommunicated.

Loisy's break with the church in 1908 put an end to what had become a false and increasingly impossible position. In 1909 he was appointed professor of the history of religion at the Collège de France, a chair that he held until 1927 and that allowed him to continue publishing in freedom. He published memoirs of his most controversial years in *Choses passées* (Paris, 1913).

His *Naissance du christianisme* (Paris, 1933) drew together and presented more intransigently views that he had held and expressed earlier, but his disbelief in the truth of the Gospel narratives and the Acts of the Apostles was now more pronounced. The supernatural elements were discredited, and the view of the historical Jesus was not very different from those of Strauss and Renan. A prophet appeared in Galilee and was crucified while Pontius Pilate governed Judaea. The rest—the alleged events of Jesus' life and his subsequent deification by his followers—belonged, for Loisy as for Renan, to the realm of myth and Messianic aspiration in search of its symbolic figure.

See also Harnack, Carl Gustav Adolf von; Modernism; Renan, Joseph Ernest; Strauss, David Friedrich.

Bibliography

ADDITIONAL WORKS BY LOISY

Histoire du canon de l'Ancien Testament. Paris: Letouzey and Ané, 1890.

Histoire du canon du Nouveau Testament. Paris: Maisonneuve, 1891.

Histoire critique du texte et des versions de l'Ancien Testament. Paris: Letouzey and Ané, 1892.

Le quatrième évangile. Paris: A. Picard, 1903.

Les évangiles synoptiques. Paris, 1908.

Jésus et la tradition évangélique. Paris, 1910.

À propos d'histoire des religions. Paris: Nourry, 1911.

Les mystères païens et le mystère chrétien. Paris: Nourry, 1914.

La paix des nations et la religion de l'avenir. Paris: Nourry, 1920.

La morale humaine. Paris: Nourry, 1923.

WORKS ON LOISY

Petre, M. D. *Alfred Loisy: His Religious Significance.* Cambridge, U.K.: Cambridge University Press, 1944.

Vidler, A. R. *The Modernist Movement in the Roman Church,* 67–139. Cambridge, U.K.: Cambridge University Press, 1934.

Colin Smith (1967)

LOMBARD, PETER

See *Peter Lombard*

LONGINUS (PSEUDO)

From its first publication in 1554 until the early nineteenth century, the fragmentary text *Peri hupsous* was all but unquestioningly attributed to Cassius Longinus, a Greek of the third century CE. Prevalent scholarly opinion now places the origin of the text in the first half of the first century; but, nothing beyond the text itself being known of its actual author, and nothing of comparable interest being known to have been written by the historical Longinus, the use of the latter name for the author of the text has stuck.

Problems of interpretation of the text begin with its title. Although the word *hupsēlos* is commonly translated as "sublime," Longinus, in contrast with modern writers, uses it neither as a quasi-technical term nor as the expression of an aesthetic concept coordinate with "the beautiful" but as an ordinary term of praise (even if special praise) for compositions of words. A case has even been made for taking the term, as he uses it, to signify nothing

more specific than greatness or excellence in discourse (Grube 1957).

However, there are also grounds for supposing that Longinus has in mind a specific literary virtue. The sublime, he says, is the echo of a great mind; its effect is to transport us; it fills us with joy and pride as if we ourselves had produced what we hear. Some of these passages may give the modern reader a faint tingle of Kant, but Longinus, however often he compares sublime writing to thunderbolts, volcanoes, and the like, never applies the word *hupsēlos* to natural phenomena—only to verbal productions. Even where he compares writing to painting and to music, he never attributes sublimity to products of those media, but only to works of words. In short, the Longinian *hupsos*, unlike the modern "sublime," is a quality of discourses only.

At the same time, Longinus holds that the power to produce sublime discourse is more a product of nature than of art (*technē*), the latter being understood as a teachable, specialized form of know-how.For Longinus, whatever contribution is made by the calculated use of figures, diction, and word arrangement—devices whose exposition in fact takes up the major part of *Peri hupsous*—the chief source of sublimity is the inborn power to form great conceptions. (Longinus's term for this power, *megalophuia*, usually translated "genius," literally means "great-naturedness.") Sublime discourse thus turns out to be discourse in which the natural greatness of the mind of the writer or speaker is seemingly imparted to the reader or hearer.

Peri hupsous enjoyed a great vogue following the publication of Nicholas Boileau's French translation in 1674, then passed out of fashion over the next century as the application of the term *sublime* shifted from discourse to nature, and the underlying conception from rhetoric to psychology. One can argue, however, that by explaining sublimity of discourse in terms of the nature of the author and the mind of the audience, Longinus himself provided the basis for that shift.

See also Aesthetics, History of; Beauty; Ugliness.

Bibliography

WORKS BY LONGINUS (PSEUDO)

On Great Writing (1957). Translated by G. M. A. Grube. Indianapolis: Hackett, 1991. Does a better job than most English translations of avoiding the imposition of alien meanings, though at the price of excluding the word "sublime."

"Longinus" on the Sublime, edited by D. A. Russell. Oxford: Clarendon Press, 1964. Greek text with commentary.

WORKS ON LONGINUS (PSEUDO)

Grube, G. M. A. "Notes on the *Peri Hupsous*." *American Journal of Philology* 78 (1957): 355–374. A collection of remarks on matters of translation and interpretation, particularly on the meaning of the word *hupsos*.

Macksey, Richard. "Longinus Reconsidered." *Modern Language Notes* 108 (1993): 913–934. A concentrated summary account of the text and its influence.

Olson, Elder. "The Argument of Longinus' *On the Sublime*." In *Critics and Criticism, Ancient and Modern*, edited by R. S. Crane. Chicago: University of Chicago Press, 1952. (Originally published in *Modern Philology* 39 (1942): 225–258.) A defense of Longinus's claim to be explaining the *art* (in the classical sense) of sublime writing.

Miles Rind (2005)

LOPATIN, LEO MIKHAILOVICH

See *Lopatin, Lev Mikhailovich*

LOPATIN, LEV MIKHAILOVICH

(1855–1920)

Lev Mikhailovich Lopatin, the Russian philosopher and psychologist, was one of a number of Russian thinkers—such as A. A. Kozlov—to advance a pluralistic idealism or personalism inspired by the monadology of Gottfried Wilhelm Leibniz. Lopatin was for many years professor of philosophy at Moscow University, president of the Moscow Psychological Society, and editor of the leading Russian journal, *Voprosy Filosofii i Psikhologii* (Problems of Philosophy and Psychology). He wrote extensively and is famous for the clarity and beauty of his style. His thought owed much not only to Leibniz (and to Rudolf Hermann Lotze) but also to his longtime friend, the Russian philosopher Vladimir Solov'ëv.

Lopatin held that every activity or process presupposes an agent. In his metaphysics there is a plurality of agents, which are spiritual entities (monads), supratemporal, and thus indestructible (since destruction involves cessation of existence in time). He held that God is related to this plurality as its unifying ground, but he did not develop fully the character of this relationship. Lopatin's chief contributions to the general doctrine of monads are his view of the substantiality of the individual spirit and his doctrine of "creative causality." According to the former, the individual spirit is neither a substance that is

separate from its phenomena nor a pure succession of absolute states; each of these conceptions is fundamentally self-contradictory. Rather, the spirit is a substance that is immanent in its phenomena; its phenomena are the direct realization of its nature. Each individual spirit, moreover, is a "creative" or productive cause; temporal, mechanical causality, and necessity, as well as all material properties—such as extension—are derivatives of the primary causality of supratemporal spirit.

Lopatin was the first of the Russian Leibnizians to give thorough attention to the moral sphere. The doctrine of creative causality gave him a basis for asserting the freedom of the will and for developing an ethical personalism in which moral phenomena represent the highest manifestation of the creative activity of individual spirit. Thus moral phenomena have metaphysical significance, and despite the evil and the inefficacy of good that we observe in the world, reality contains a moral order and is not "indifferent to the realization of the moral ideal."

Just as in ethics Lopatin maintained that unaided experience is not an adequate guide, so in epistemology generally, he discounted pure empiricism in favor of "speculative" principles, defining speculative philosophy as "the knowledge of real things in their principles and in their ultimate signification." Man's immediate inner experience is the source of his knowledge of real things, but philosophy works on this experience and goes beyond it through rational speculation.

See also Agent Causation; Ethics, History of; Kozlov, Aleksei Aleksandrovich; Leibniz, Gottfried Wilhelm; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY LOPATIN

Polozhitel'nyye zadachi filosofii (The positive tasks of philosophy). 2 vols. Moscow, 1886–1891.

Filosofskie kharakteristiki i rechi (Philosophical characterizations and speeches). Moscow: Academia, 1995.

'The Philosophy of Vladimir Soloviev." *Mind* 25 (October 1916): 425–460.

WORKS ON LOPATIN

Ognev, A. *Lev Mikhailovich Lopatin*. Petrograd, 1922.

Zen'kovskii, V. V. *Istoriia russkoi filosofii*. 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy*. 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

LOSEV, ALEKSEI FËDOROVICH
(1893–1988)

Aleksei Fëdorovich Losev was a Russian philosopher and classicist and the author of numerous works on ancient and early modern aesthetics, language, symbolism, myth, and music aesthetics. A native of Novocherkassk, he graduated from Moscow University in 1917 with degrees in philosophy and classical philology and later taught at the University of Nizhnii Novgorod and Moscow Conservatory. Before they ceased to exist in 1922 he attended the meetings of the Vladimir Sergeevich Solov'ëv (Solovyov) Religious-Philosophical Society and Nikolai Aleksandrovich Berdyaev's Free Academy of Spiritual Culture, where he met the leading figures of the so-called religious-philosophical renaissance.

During the 1920s Losev forged his own version of Christian neoplatonism for which he drew on ancient Platonists, Greek church fathers, German idealism (especially Friedrich Wilhelm Joseph von Schelling and Georg Wilhelm Friedrich Hegel), Russian religious-philosophical thought, and Edmund Husserl. The later, largely forced, assimilation of Marxism was neither purely cosmetic nor did it cause any fundamental shift in his outlook. Losev accepted the valuable aspects of Marxism but eschewed its limitations. In 1929 he secretly took monastic vows. Between 1927 and 1930 he published eight volumes on ancient philosophy, philosophy of language, mathematics, music aesthetics, and philosophy of myth. The last book in this series, *Dialektika mifa* (The dialectics of myth; 1930), became the cause of Losev's arrest and sentence of ten years in labor camps. He was freed, almost totally blind, in 1933 and for the next twenty years he was not allowed to publish his own work or teach philosophy. After teaching part time at provincial universities he became a professor at Moscow University in 1942 and was even awarded a doctorate in classical philology. The appointment was soon withdrawn, however, on charges of idealism and Losev was transferred to Moscow State Pedagogical Institute, where he remained until retirement. He resumed publishing in 1953 and eventually established himself, against considerable official resistance, as one of the most respected authors on ancient philosophy and culture in the Soviet Union. By the end of his life Losev's oeuvre included more than 30 monographs and 400 scholarly publications. Posthumous editions have increased this number almost twofold. The crowning achievement of his life's labor was an eight-volume study on ancient aesthetics—an original interpreta-

tion of antiquity without precedent, in scope and size, in world classical scholarship.

Losev's output over his lifetime is marked by a remarkable continuity. The themes of antiquity, language, symbol, myth, mathematics, and music remained constant from his earliest to his last publications. His vast oeuvre, however, still requires much study and any judgment of it must remain provisional at this stage.

METHOD

Losev elaborated his phenomenological-dialectical method in the 1920s and later only supplemented it by new influences, among which Marxism and structuralism were perhaps most notable. In Marxism Losev found support for his conviction about a meaningful link between socioeconomic and intellectual processes but, in contrast to Marxism, he did not reduce the latter to the former. He considered the eidos of classical phenomenology, which he described as "the integral semantic face (*smyslovoi lik*) of a thing" (1927a, p. 53), too static and supplemented it by establishing dynamic dialectical relations among its constituent parts. Like Hegel, Losev understood dialectics as the rhythm of both thinking and objective reality, but his own version of dialectics was derived largely from ancient and Christian neoplatonist sources.

Losev had a penchant for developing multilayered analytic structures of the phenomena that he studied. Often, the key element in these conceptual constructions is what he calls the dialectical tetraktis: the development of meaning via the four steps of unity, multiplicity, the ideal synthesis of the two, and, finally, the fact in which this synthesis is realized.

Throughout his life Losev argued strenuously against the dogmatic one-sidedness of both materialism and idealism and strove to position himself above these abstract divisions.

LANGUAGE

Central to Losev's entire outlook was the philosophy of language articulated in *Filosofiia imeni* (Philosophy of the name; 1927b). Losev's view was informed by onomatodoxy (*imiaslavie*), a trend in Orthodox theology centered on the veneration of God's name. He understood language in terms of ontological symbolism, that is, as access to the reality of being. "The name," he argued, "is life. … The mystery of the word consists precisely in that it is the tool of our intimate and conscious encounter with the inner life of things. … The world is created and is held together by the name and the word" (1993, pp. 617, 642, 746). The name, according to Losev, embraces being in its entirety, from the meonic formlessness of pure matter, to the rational, eidetic formation of all natural and social phenomena, to the suprarational regions of thinking where it passes into "noetic ecstasy" (pp. 676–677). Losev's meticulous gradations of this phenomenological-neoplatonist terrain are held together by a dialectical hierarchy of various "moments" in the structure of words. Later in life Losev attentively studied structuralism with which the eidetic aspect of his analysis of language had much in common. He consistently objected, however, to all nondialectical treatments of language, be they positivist, neo-Kantian, or structuralist.

SYMBOL

Losev's theory of the symbol was inspired by the thought of Pavel Aleksandrovich Florenskii and Viacheslav Ivanovich Ivanov, but it also absorbed other influences as it evolved from his early work, such as *Antichnyi kosmos i sovremennaia nauka*, to later writings, such as *Problema simvola i realisticheskoe iskusstvo* (The problem of symbol and realistic art, 1976). In the latter Losev analyses in detail the structure of symbol and argues that symbols are means of practical, creative "re-making of reality" (pp. 15–17). In Losev's view a symbol is the perfect fusion of inner meaning and its external expression. It is this balance that distinguishes it from allegory, where the image outweighs the abstract idea, or from a scheme, where the idea is rich but its representation arid.

MYTH

Losev regards myth as a necessary category of consciousness and defines it as "unfolded magical name" (2003, pp. 186–187)—a formula that highlights myth's verbal (narrative) form, personalistic nature, and the presence of the miraculous in it. As a story about reality it is distinct from poetry and art in general; as a prereflexive story about a miraculous reality, myth is distinct from science and metaphysics. Myths form the foundations of people's outlooks, Losev argues, and thus determine cultural and historical processes on the most fundamental level. He views the history of culture as a constant struggle among various mythologies, and one of his tasks is to uncover the inner logic of this process. "Whatever one's view of myth, any critique of mythology is always merely a profession of another, new mythology" (1927b, p. 771). According to Losev no historical epoch is free of mythology and, despite its hostility toward myth, modernity is emphatically mythological. Modern cosmology advances, he impugns, a vision of the world as an infinite dark void,

ruled by a "blind, deaf, and dead" monster, that is, matter. Losev's other targets among modern myths include titanic Prometheanism that he critiqued at length in *The Problem of Symbol and Realistic Art*. The key notion in Losev's critique of modernity was what he called in *Estetika Vozrozhdeniia* (Renaissance aesthetics; 1978) "the absolutization of the human subject."

The reconciliation of myth and philosophy is Losev's goal in his essay "Absoliutnaia mifologiia = absoliutnaia dialektika" (Absolute mythology = absolute dialectics; 1929–1930, published in 2000). Taken by themselves, both myth and dialectics are limited and an adequate outlook can be based, Losev insists, only on their synthesis. Dialectics inevitably comes up against the ultimate limit of rational cognition, and in the suprarational realm beyond this boundary it should be fused with mythology (p. 275). In his early period Losev found the optimal candidates for such a synthesis in the mythology of Eastern (Orthodox) Christianity and Russian religious-philosophical thought.

Losev applied his theoretical ideas to numerous analyses of specific myths ranging from ancient Greco-Roman to modern mythology (*Ocherki antichnogo simvolizma i mifologii* [Essays on ancient symbolism and mythology; 1930]; *Mifologiia grekov i rimlian* [The mythology of the Greeks and Romans; 1930s, published in 1996]; and *The Problem of Symbol and Realistic Art*).

ANTIQUITY

In *Istoriia antichnoi estetiki* (History of ancient aesthetics; 1963–1994) Losev's point of departure is that all ancient philosophizing, from pre-Socratics to Proclus, is based "on the intuitions of a thing, rather than of personhood" (*Istoria antichnoi filosofii v konspektivnom izlozhenii* [History of ancient philosophy: a conspectus], p. 155). He emphasizes the link between the "material-thingly" (*material'no-veshchestvennaia*) basis of thinking and ancient slave-owner economy but rather than a particular economic order the ultimate intuitive ground of ancient philosophy was "the sensible, material cosmos" (p. 15). From this impersonal absolute stems ancient fatalism that gradually evolves, via Stoics and other schools, toward providentialism. Ancient philosophy ends, Losev claims, when this original, astronomical intuition is replaced by the personalistic and historical vision of reality in Christianity. Losev argues for a dialectical view of this process, in which ancient philosophy grows out of specific mythological intuitions in the late archaic and early classical period, and in the end returns to embrace and justify this original mythology on rational grounds—only to yield to a new mythology and a philosophy that evolves on its basis.

MUSIC

Losev's philosophy of music combines Pythagoreanism and Romanticism, both refracted through his dialectical phenomenology. Eventually, he also explored Marxist themes, such as music and ideology—especially in his philosophical prose of the early 1930s (published posthumously). The culmination of Losev's early philosophy of music was *Muzyka kak predmet logiki* (Music as the subject of logic; 1927), where music is defined as the expression of "the life of numbers." In its depth this life is a total "coincidence of opposites" and "extreme formlessness" that defies all categories of the understanding (1990, p. 209). At the same time a musical work possesses an "eidetic completeness" (p. 269). Fused with the chaos of "pure musical being," this mathematically determined fullness of form makes music "the eidos of the alogical" (p. 279). Losev further evokes Plotinus to define time as "the alogical becoming of the number" (p. 328) and links this idea with the temporal nature of music. The closing passages of the book are devoted to deriving from these insights such elements of musical form as melody, rhythm, harmony, and even timbre. Losev both used and further elaborated his philosophy of music in a number of essays, written in the course of his lifetime, on specific composers, such as Nikolai Rimsky-Korsakov (1844–1908), Aleksandr Scriabin (1872–1915), and Richard Wagner (1813–1888).

Losev is one the key philosophers—perhaps *the* key philosopher—who preserved the continuity of Russian religious-philosophical tradition in Russia against a concerted effort by the Soviet regime to destroy it. In the post-Soviet period Losev emerged as one of the central figures of twentieth-century Russian thought—a position confirmed by numerous editions of his works and his broad influence on the current philosophical discourse. The significance of his work, however, reaches far beyond the Russian context. While the strikingly broad reach of his thought makes the recognition of his contribution difficult, it also comprises highly valuable insights into the nature of thinking, history, personhood, and expression.

See also Berdyaev, Nikolai Aleksandrovich; Florenskii, Pavel Aleksandrovich; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; Idealism; Ivanov, Viacheslav Ivanovich; Marxist Philosophy; Myth; Neoplatonism; Patristic Philosophy; Philosophy of Language; Platon-

ism and the Platonic Tradition; Pre-Socratic Philosophy; Proclus; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY LOSEV

Antichnyi kosmos i sovremennaia nauka (The ancient cosmos and modern science). Moscow: Author, 1927a.

Filosofiia imeni (Philosophy of the name). Moscow: Author, 1927b.

Istoriia antichnoi estetiki (History of ancient aesthetics), Vol. 1. Moscow: Vysschaia shkola, 1963; Vols. 2–8. Moscow: Iskusstvo, 1969–1994.

Iz rannkikh proizvedenii (Early writings). Moscow: Pravda, 1990.

Bytie, Imia, Kosmos (Being, name, cosmos). Moscow: Mysl', Rossiiskii otkrytyi universitet, 1993.

Forma, Stil', Vyrazhenie (Form, style, expression). Moscow: Mysl', 1995.

The Dialectics of Myth. Translated by Vladimir Marchenkov. London: Routledge, 2003.

Istoriia antichnoi filosofii v konspektivnom izlozhenii (A history of ancient philosophy: a conspectus), Moscow: Mysl', 1989.

Problema simvola i realisticheskoe iskusstvo (The problem of the symbol and realistic art), Moscow: Iskusstvo, 1976.

WORKS ABOUT LOSEV

"Aleksej Fedorovich Losev: Philosophy and the Human Sciences." *Studies in East European Thought* 56 (2–3) (June 2004) (Special Issue).

Haardt, Alexander. "Aleksei Losev and the Phenomenology of Music." In *Russian Thought after Communism: The Recovery of a Philosophical Heritage*, edited by James P. Scanlan, 197–205. Armonk, NY: M. E. Sharpe, 1994.

Haardt, Alexander. *Husserl in Russland: Phänomenologie der Sprache und Kunst bei Gustav Spet und Aleksej Losev*. Munich, Germany: Fink Verlag, 1993.

Jubara, Annett. *Die Philosophie des Mythos von Aleksej Losev im Kontext "Russischer Philosophie"* Wiesbaden, Germany: Harrassowitz Verlag, 2000.

Vladimir Marchenkov (2005)

LOSSKII, NIKOLAI ONUFRIEVICH

(1870–1965)

Nikolai Losskii (Lossky), a Russian religious philosopher, was born in the province of Vitebsk in western Russia. He studied history, philology, and natural sciences at St. Petersburg University (1891–1898), as well as philosophy under the neo-Kantian Aleksandr Vvedenskii (1856–1925). Losskii continued his philosophical education in Germany (1901–1903) with Wilhelm Windelband, Wilhelm Wundt, and Georg Müller. He received his master's degree in 1903, and his doctorate in philosophy four years later. From 1900 Losskii taught at St. Petersburg University, where he was appointed to a chair of philosophy in 1916. In 1921 Losskii was dismissed from the university for his religious beliefs, and in 1922 he was exiled by the Soviet government from the homeland. From 1922 to 1945 he settled in Czechoslovakia, where he taught in universities in Prague, Brno, and Bratislava. From 1946 Losskii lived in the United States and taught at St. Vladimir's Orthodox Theological Seminary in New York (1947–1950).

Losskii was a systematic philosopher and prolific writer whose works have been translated into many foreign languages. His writings cover most of the traditional philosophical disciplines, though he gave special emphasis to epistemology, metaphysics, and ethics. His philosophy is variously labeled as intuitivism, hierarchical personalism, or ideal-realism, depending on what part of his comprehensive system the commentator focuses on. The central idea of Losskii's philosophy is, in his own words, the insight that "everything is immanent in everything" (Zenkovsky 1953, p. 668). In his religious views Losskii adhered to Christian doctrine, though some of his views, such as his teachings about reincarnation and creation, seem incompatible with the Orthodox tradition.

In his epistemology, Losskii rejected the possibility of transcendent knowledge and affirmed that in the process of cognition, subject and object must be connected. In acts of knowing, the object of knowledge is not a representation of an entity but the actual entity itself. The subject or self becomes cognizant of the world of nonself by a special act that Losskii called "epistemological coordination." Although the object of knowledge is part of the process of knowing, the content of knowledge contains more than its own object; rather, it is the result of the subject's efforts at comparing and distinguishing. Hence, the truth that one can achieve in the cognitive process is never complete, because the process of differentiating, however strong it may be, always leaves unexplored some part of reality.

In Losskii's theory of knowledge, named "intuitivism," intuition is not merely one aspect of cognition, but permeates all cognitive processes. Though all knowledge is intuitive by nature, knowledge can be differentiated by the type of intuition. Losskii distinguished three types of intuition: sensuous, intellectual, and mystical, corresponding respectively to the real, ideal, and metalogical levels of existence.

In his ontology, Losskii defended an "organic," or holistic, worldview. In his view, any object constitutes a system by virtue of a principle that lies beyond that system. As a systemic unity, the world requires a principle that stands beyond it and represents its foundation. This principle is called "the Absolute" in philosophy and "God" in religion. No positive definition grasps the Absolute as such, but philosophers can study its manifestations in the created world.

In the created realm, Losskii distinguished three levels of reality: the real, the abstract, and the concretely ideal, the last of which consists of living agents, whom he sometimes referred to as concrete ideal entities, substances, or, more precisely, substantival agents. As compared with the abstract ideal, which includes, for instance, abstract relations, ideal entities are active agents who independently determine their own manifestations in time. The human self is one such substantival agent. As an entity that transcends space and time, it is responsible for creating psychic processes in time and realizing material events in a spatiotemporal framework.

In Losskii's view, God's creation stops with substantival agents, who are free to choose their own evolution. The original sin of self-centeredness, symbolically described in the Biblical story of the fall of Adam and Eve, does not signify that humanity once attained perfection and then freely lost it. The life of the spirit has to result from efforts exercised by the creature itself; otherwise the creature's freedom is falsified. Those substantival agents who choose selfishness and prefer their own interests to God's will must continue their evolution on the lower levels of reality and are subjected to a long and difficult process of redemption.

Since the universe is an integral holistic system, an organism, all substantival agents are interconnected with each other. Their consubstantiality is crowned with and headed by the cosmic substance, which Losskii, following the Solov'evian tradition, called "Sophia." Though not identified with the Absolute, this supreme substance, like all other creatures belonging to the created realm, is perfect and unites the multiplicity of creation into one cosmic whole. The kingdom of God, led by Sophia, represents the ontological basis of absolute values and the ultimate goal for every substantival agent. The existence of the spiritual kingdom makes it possible for fallen beings to restore their original divine identities and to partake of the heavenly life. In the kingdom of God, everyone is in harmony with all, and everyone is all. In the life of the kingdom of God, headed by Sophia, every member experiences constant growth in all possible dimensions that ideally complement and enrich one another.

Though Losskii wrote comparatively little on political philosophy, in his few articles on the subject he consistently stood for democratic values. According to him, in the course of an increasingly complex social life, the state is unified more securely by the dispersion of power and by constitutional limits on the absolute power of the monarch. The ultimate choice between monarchy and republic depends on which can best balance the united will of the nation with the rights and development of its members.

See also Intuition; Personalism; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich; Sophia; Windelband, Wilhelm; Wundt, Wilhelm.

Bibliography

WORKS BY LOSSKII

Mir kak osushchestvlenie krasoty: Osnovy estetiki (The world as the manifestation of beauty: principles of aesthetics). Moscow: Progress Publishers, 1996.

Chuvstvennaia, intellektual'naia i misticheskaia intuitsiia (Sensory, intellectual, and mystical intuition). Moscow: Respublika, 1995. Includes the following works: *Tipy mirovozzrenii: Vvedenie v metafiziku* (Types of worldviews: an introduction to metaphysics; 1931); *Chuvstvennaia, intellektual'naia i misticheskaia intuitsiia* (Sensory, intellectual, and mystical intuition; 1938); *Obshchedostupnoe vvedenie v filosofiiu* (A general introduction to philosophy; 1956), pt. 1 and pt. 3, chap. 21.

Bog i mirovoe zloi (God and worldly evil). Moscow: Respublika, 1994. Includes the following works: *Dostoyevski i ego khristianskoe miroponimanie* (Dostoyevsky and his Christian worldview; 1953); *Tsennost' i bytie: Bog i Tsarstvo Bozhie kak osnova tsennostei* (Value and existence: God and the kingdom of God as the basis of values; 1931); translated by S. Vinokooroff as pt. 1 of *Value and Existence*, by N. O. Lossky and J. S. Marshall (London: Allen and Unwin, 1935); *Bog i mirovoe zlo: Osnovy teoditsei* (God and worldly evil: principles of theodicy; 1941).

Istoriia russkoi filosofii. Moscow: Progress, 1994. Translated as *History of Russian Philosophy* (New York: International Universities Press, 1951). Includes Losskii's summary of his own philosophy.

Usloviia absoliutnogo dobra (Conditions of the absolute good). Moscow: Politicheskaia literatura, 1991. Includes the following works: *Usloviia absoliutnogo dobra: Osnovy etiki* (Conditions of the absolute good: principles of ethics; 1949); *Kharakter russkogo naroda* (The character of the Russian people; 1957).

Izbrannoe (Selected works). Moscow: Pravda, 1991. Includes the following works: *Obosnovania intuitivisma* (1906), translated by N. Duddington as *The Intuitive Basis of Knowledge: An Epistemological Inquiry* (London: Macmillan, 1919); *Mir kak organicheskoe tseloe* (1917), translated by N.

Duddington as *The World as an Organic Whole* (London: Oxford, 1928); *Svoboda voli* (1927), translated by N. Duddington as *Freedom of Will* (London: Williams and Norgate, 1932).

WORKS ON LOSSKII

Lossky, Boris, and Nadejda Lossky. *Bibliographie des œuvres de Nicolas Lossky*. Paris: Institut d'études slaves, 1978. A complete bibliography of works by Losskii and translations of his works. Includes a detailed chronology of his life.

Zenkovsky, V. V. *A History of Russian Philosophy*. Vol. 2, 630–676. Translated by George L. Kline. London: Routledge and Kegan Paul, 1953.

Mikhail Sergeev (2005)

LOTMAN, IURII MIKAILOVICH

(1922–1993)

Iurii Mikailovich Lotman was a specialist in the theory of literature and aesthetics, the history of Russian literature, semiotics, and study of culture. He was born in Petrograd (now St. Petersburg). In 1939 he commenced his studies in the philology department of Leningrad University. In the fall of 1940 he joined the army and fought in World War II from 1941 to 1945. In 1946 he continued his studies at the university, finishing them in 1950. Because of the anti-Semitic campaign in the Soviet Union, Lotman was not able to work in Leningrad and moved to Estonia. From 1950 to 1954 he taught at the Tartu Pedagogical Institute. In 1952 he defended his dissertation in philology on the ideas of A. N. Radishchev and N. M. Karamzin. In 1954 he was named docent of Tartu University, and from 1960 to 1977 he was the head of the Department of Russian Literature there. In 1961 he received a doctorate in philology by defending the dissertation titled *Puti razvitiia russkoi literatury preddekabristskogo perioda* (Paths of the development of Russian literature in the pre-Decembrist period).

FROM THE HISTORY OF LITERATURE TO SEMIOTICS

Lotman's chief historical works are devoted to the history of Russian literature from the eighteenth century to the mid-nineteenth century. He examines this literature in conjunction with other cultural phenomena, particularly philosophical thought, history, and sociopolitical life. From the beginning of the 1960s Lotman develops a structural-semiotic approach to the study of works of art, organized the publication of the series *Trudy po znakovym sistemam, Semiotika* (*Sign Systems Studies, Semiotics*), and directed regularly held "summer schools," conferences, and seminars on the semiotic study of various domains of culture. The combination of these activities, which included the participation not only of Tartu scholars but also of scholars from Moscow and other cities, became the internationally known Tartu-Moscow School of Semiotics (Grzybek 1989). The first issue of *Sign Systems Studies* included his *Lektsii po struktural'noi poetike* (Lectures on structural poetics) (Lotman 1964).

The works of Lotman and those of his colleagues and followers on the semiotic analysis of various cultural texts, including artistic texts in particular, are united by the idea of "secondary modeling systems," where the text is interpreted as a unity of models of objective and subjective reality, as well as in the capacity of a sign system secondary in relation to the signs of natural languages, which represent the "primary modeling system." Headed by Lotman, the "Tartu school" of semiotics continues the traditions of the Russian "formal school," especially Iurii Tynianov, and structural linguistics (Ferdinand de Saussure and Roman Jakobson), taking into account the efforts to develop semiotic structuralism in various countries. However, the Tartu school does not limit itself to the study of the formal structure of works of art; it focuses primarily on the semantics of sign structures (Lotman 1970, Shukman 1977). Together with his semiotic studies, Lotman also continues his historico-literary investigations, in which he employs a structural-semiotic methodology. The novelty of his work is that he attempts to combine structuralism with historicism, the premise being that a semiotician must also be a historian. Lotman's work in the history of literature is characteristically theory-laden.

FROM SEMIOTICS TO THE STUDY OF CULTURE

At the beginning of the 1970s Lotman arrived at the view that the semiotic object must be adequately understood not simply as a separate sign but as a text existing in culture—as a text constituting "a complex device storing multiple and diverse codes, capable of transforming received messages and of generating new ones, like an information generator possessing traits of intellectual personality" (Lotman 1981, p. 132). Taking this as his point of departure, Lotman considers culture itself in its semiotic aspect, in the multiplicity of its communicative connections (Lotman 1970–1973). By analogy with V. I. Vernadskii's concepts of "biosphere" and "noosphere," Lotman introduced the concept of "semiosphere," which is characterized by the limits of semiotic space, its struc-

tural heterogeneity and internal diversity, forming a structural hierarchy whose components are in a dialogic interrelationship (Lotman 1984). Lotman thus realized the transformation of the initial semiotics and overcomes its total schematism. But he did this not through post-structuralism and "deconstruction" (in the spirit of Jacques Derrida) but through a semiotic interpretation of cultural texts, taking into account their uniqueness, creative character, and intertextual dialogues. Not only is culture as a whole understood as a text, but any text is viewed as a product of culture.

Lotman's theoretical views take into account the development of contemporary scientific knowledge, especially information theory, cybernetics, the theory of systems and structures, the theory of the functional asymmetry of the brain, and the ideas of synergetics (Lotman 1990, 1992). At the same time these views also rely on the abundant material of world culture, primarily Russian culture, which is considered in its typological significance. Lotman's works on the history of Russian culture are of great value. Highly popular was his series of television broadcasts on Russian culture, aired posthumously in 1994 (Lotman 1994).

PHILOSOPHICAL POSITION

Lotman did not explicitly declare his philosophical views. In the presemiotic period of his activity, philosophy interested him only as an object of historical study. But semiotic and culturological studies presupposed a theoretico-philosophical self-definition. Lotman had a broad knowledge of philosophy and closely studied the ideas of Gabriel de Mably, Jean-Jacques Rousseau, and Alexander Radischev (Lotman 1958, 1960). He also identified in a masterful way the philosophical content of the work of literary artists (Lotman 1987, 1988). His own philosophical-methodological ideas underwent a specific evolution (Kim Soo Hwan 2003). In the 1960s the adherents of the "Tartu school" held positivist views, maintaining that semiotics was in fact their philosophy (Stolovich 1994). But later Lotman began to search for a philosophy that would correspond to his semiotic culturology. He turned to Leibniz's monadology, proposing that the semiosphere consists of a multiplicity of "semiotic monads" as intellectual units—that is, bearers of Reason. In his own words, "man not only thinks but also finds himself within a thinking space, just as a bearer of speech is always immersed in a certain language space." The existence of the external world is accepted, but it too is "an active participant in the semiotic exchange" (Lotman 1989, pp. 372, 375). God for Lotman is a universally significant phenomenon of culture. Although his attitude toward religion was respectful, he himself was a theological agnostic (Egorov 1999, pp. 236–237).

Lotman keenly absorbed the ideas of various thinkers, including Leibniz, Rousseau, Kant, Hegel, and Freud. In 1967 and 1971 he published in the journal *Semiotics* certain works of the Russian religious philosopher and scientist Pavel Florenskii, who had been repressed by the Soviet authorities. Lotman also reacted positively to Bakhtin's conception of dialogue (Egorov 1999, pp. 243–258). However, Lotman's own philosophical views cannot be reduced to any one system, be it Platonism (Vetik 1994), Kantianism (M. Lotman 1995), Hegelianism, or Marxism. His philosophical views can be defined as a type of "systemic pluralism," which presupposes the combination of heterogeneous ideological components in a specific system.

Living and being educated in the Soviet Union, Lotman could not fail to feel the influence of Marxism. He assimilated that aspect of Marxism that was related to Hegel's dialectic, the principle of historicism, and the social factor in the development of culture. But the ideological content of Marxism was alien to Lotman (Gasparov 1996, pp. 415–426). His structural-historical studies provoked the suspicion and displeasure of official circles (at the beginning of the 1970s, he was even interrogated by the KGB and his belongings were searched). At the same time his popularity grew immensely and he was considered a scholar of the first rank and a brilliant personality in intellectual circles both in the Soviet Union and abroad. He was elected as a corresponding member of the British Academy, as an academician of the Norwegian, Swedish, and Estonian academies of science, and as vice president of the International Semiotics Association. The institute for Russian and Soviet culture in Germany was named after him: Lotman-Institut für russische und sowjetische Kultur, Ruhr-Universität Bochum.

See also Aesthetics, History of; Bakhtin, Mikhail Mikhailovich; Cybernetics; Derrida, Jacques; Florenskii, Pavel Aleksandrovich; Freud, Sigmund; Hegel, Georg Wilhelm Friedrich; Hegelianism; Historicism; Information Theory; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Marxist Philosophy; Monad and Monadology; Neo-Kantianism; Platonism and the Platonic Tradition; Radishchev, Aleksandr Nikolaevich; Rousseau, Jean-Jacques; Russian Philosophy; Structuralism and Post-structuralism.

Bibliography

WORKS BY LOTMAN

Radishchev i Mabli [Radishchev and Mably]. 1958. In *Izbrannye stat'i* [Selected articles], by Yu. M. Lotman. Vol. 2, 100–123. Tallinn, Estonia: Aleksandra, 1992.

Russo i russkaja kul'tura v XVIII–nachala XIX veka [Rousseau and Russian culture from the 18th century to the the early 19th century]. In *Traktaty* [The treatises], by Jean-Jacques Rousseau. Moscow: Nauka, 1969.

Lektsii po struktural'noii poetike [Lectures on structural poetics]. Tartu, Estonia, 1964.

Struktura khudozhestvennogo teksta [The structure of the artistic text]. Moscow: Iskusstvo, 1970. Translated by Gail Lenhoff and Ronald Vroon as *The Structure of the Artistic Text*. Michigan Slavic Contributions 7. Ann Arbor: University of Michigan, Department of Slavic Languages and Literatures, 1977.

Stat'i po tipologii kul'tury [Articles on the typology of culture]. Vols. 1–2. Tartu, Estonia, 1970–1973.

Analiz poeticheskogo teksta [Analysis of the poetic text]. 1972. Translated by D. Barton Johnson as *Analysis of the Poetic Text*. Ann Arbor, MI: Ardis, 1976.

Semiotika kino i problemy kinoestetiki [Semiotics of cinema and problems of a film-aesthetics]. Tallinn, Estonia: Eesti raamat, 1973. Translated by Mark Suino as *Semiotics of Cinema*. Michigan Slavic Contributions. Ann Arbor: University of Michigan Press, 1976.

"*Semiotika kultury i poniatie teksta*" [Semiotics of culture and the concept of the text] [1981]. In *Izbrannye stat'i* [Selected articles], by Yu. M. Lotman. Vol. 1, p. 132. Tallinn, Estonia: Aleksandra, 1992.

"*O semiosfere*" [On the semiosphere]. 1984. In *Izbrannye stat'i* [Selected articles], by Yu. M. Lotman. Vol. 1, 11–24. Tallinn, Estonia: Aleksandra, 1992.

Sotvoreni? Karamzina [The creative works of Karamzin]. Moscow: Kniga, 1987.

Iz razmyshlenii nad tvorcheskoi evoliutsiei Pushkina [From reflections on the creative evolution of Pushkin]. 1988. In *Izbrannye stat'i* [Selected articles], by Yu. M. Lotman. Vol. 2, 473–478. Tallinn, Estonia: Aleksandra, 1992.

Kul'tura kak sub'ekt i sama-sebe ob'ekt. 1989. In *Izbrannye stat'i* [Selected articles], by Yu. M. Lotman. Vol. 3. Tallinn, Estonia: Aleksandra, 1993. Translated by Lotman Juri as "Culture as a Subject and an Object in Itself." *Trames* 1 (1) (1997): 7–16.

Universe of the Mind: A Semiotic Theory of Culture. Translated by Ann Shukman. Introduction by Umberto Eco. London: Tauris, 1990. 2nd ed., 2001.

Kultura i vzryv [Culture and explosion]. Moscow: Gnosis-Progress, 1992.

Besedy o russkoi kul'ture: Byt i traditsii russkogo dvorianstva; XVIII-nachalo XIX veka [Conversations on Russian culture: The life and traditions of the Russian nobility from the 18th century to the early 19th century). St. Petersburg: Iskusstvo, 1994.

Vnutri mysliashchikh system: Chelovek-Tekst-Semiosfera-Istoriia [Inside thinking systems: man-text-semiosphere-history]. Moscow: Iazyki russkoi kul'tury. 1996.

REFERENCES AND FURTHER READING

Egorov, B. F. *Zhizn' i tvorchestvo Iu M. Lotmana* [Life and creativity of Yu. M. Lotman]. Moscow: Novoe literaturnoe obozrenie, 1999.

Gasparov, M. L. *Lotman and Marxism* [1996]. In *Vnutri myslyashchikh system: Chelovek-Tekst-Semiosfera-Istoriya*, by Yu. M. Lotman. Moscow: Iazyki russkoi kul'tury, 1996.

Grzybek, P. *Studium zum Zeichenbegriff der sowjetischen Semiotik (Moskkauer und Tartuer Schule)*. Bochum, Germany, 1989.

Kim Soo Hwan. *Osnovnye aspekty tvorcheskoi evoliutsii Y. M. Lotmana* [The main aspects of Lotman's creative evolution]. Moscow: Novoe literaturnoe obozrenie, 2003.

Lotman, M. Yu. *Zametki o filosofskom fone tartuskoi semiotiki* [Notes on the philosophical background of semiotics in Tartu]. In: *Lotmanovskii sbornik*, vol. 1. 214–222. Moscow: ITS-Garant, 1995.

Shukman, Ann. *Literature and Semiotics: A Study of the Writings of Yu. M. Lotman*. Amsterdam: North-Holland, 1977.

Stolovich, L. N. "*A. F. Losev o semiotike v Tartu*" [A. F. Losev on semiotics in Tartu]. *Novoe literaturnoe obozrenie* 8 (1994): 99–104.

Vetik, R. "Platonism of Yu. Lotman." *Semiotika* 99 (1–2) (1994).

Leonid N. Stolovich (2005)
Translated by Boris Jakim

LOTZE, RUDOLF HERMANN
(1817–1881)

Rudolf Hermann Lotze, the German idealist metaphysician, was born in Bautzen He studied medicine and philosophy at the University of Leipzig, taking his doctorates in both fields. He studied mathematics and physics with E. H. Weber, W. Volckmann, and G. T. Fechner and philosophy with C. H. Weisse, who influenced him greatly. In 1841 he became instructor in medicine at Leipzig, where he subsequently taught philosophy. While at Leipzig he published two short works, the *Metaphysik* (Leipzig, 1841) and *Logik* (Leipzig, 1843), which adumbrated the essentials of his later philosophy. In 1844 Lotze succeeded Johann Friedrich Herbart as professor of philosophy at the University of Göttingen. He remained there until 1881, when he was called to the University of Berlin. Shortly after joining the faculty at Berlin, he contracted pneumonia and died.

Lotze pursued his interests in the medical sciences, psychology, philosophy, the arts, and literature throughout his life. As a result of his medical training, he developed a strong love for exact investigation and precise knowledge, but art and literature made him particularly

sensitive to the central role of feeling and value in the total life of a culture. He wanted nothing to interfere with the growth of the exact sciences in all areas of human experience, yet he insisted that both intellect and scientific knowledge were essentially the means and tools of feeling, emotion, and intuition.

NEW CONCEPTION OF METAPHYSICS

Like many thinkers born in the first two decades of the nineteenth century, Lotze faced three great schisms: the schism between science and Christianity, which was known at that time as the conflict between science and religion; the schism between reason and feeling; and the schism between knowledge and value. To Lotze, these schisms had to be rationally harmonized in some manner. It seemed impossible to him for any rational man to reject any one of the trinity that composes the total culture of man: science, art, and value. Each has its place in the life of man and the universe, and none can be eliminated without distorting and destroying that life. However, Lotze felt that their proper relationship cannot be established by the older metaphysical methods. There is no possibility of rationally deducing the basic categories and values of existence by any sort of logical dialectic, either Platonic or Hegelian. Knowledge of existence depends upon knowledge of fact acquired through observation and experimentation. Consequently, the empirical sciences are the proper investigators of existence. All that metaphysics can do is to analyze, clarify, and order those concepts and theories that the sciences create into as adequate a system as the facts permit. Metaphysics cannot go beyond this in any scientific sense. Nevertheless, Lotze admitted that metaphysics has another, broader purpose. The urge to be metaphysical is not to be found in metaphysics itself but in ethics, in the desire to know and attain some ultimate good. Thus, metaphysics involves speculating beyond what is scientifically warrantable in order to include that which drives men to write metaphysics: the experience of ultimate goodness.

Because metaphysics must be founded upon science, Lotze objected to any philosophical system that claimed completeness. All philosophical systems, like his own, must remain open and undogmatic. They must not even provide provisional answers to profound questions for which not even provisional answers exist. He thought it was far better for a philosopher to raise questions and to stimulate inquiry than to offer sterile answers lacking any reasonable foundation in fact.

IDEALISTIC MONADISM, MECHANISM, AND GOD

Lotze was essentially an idealist, but his idealism was tempered by his respect for science and his emphasis upon feeling as the dominant element guaranteeing meaning to the life of man.

From the beginning, Lotze considered thought as one aspect of the soul. Thought is aware only of ideas about reality; it does not know reality, for knowledge and reality can never be identical. Neither are experience and thought to be identified. Identity with reality or object can be achieved as experience, never as thought, for thought is purely representative. Truth is attained in ways different from thought. Thought must fuse with the total feeling experience, since it is in feeling that we have direct awareness of good and evil, beauty and ugliness, worth and unworth, contradiction and harmony; these rest in the soul's original capacity to experience pain and pleasure. Consequently, Lotze thought that feeling is the ultimate arbiter of consistency and intellectual harmony, the ultimate judge of the worth of anything, and the ultimate creator of imagination and its works. Moreover, feeling is the nisus that drives man to seek whatever total unity of comprehension and action is possible for him. His love of knowledge, goodness, and beauty arises from, and finds its fulfillment in, feeling. Thus, the essential nature of feeling is love, which constantly drives man toward a greater overall comprehension, of his life and the cosmos. "If … love did not lie at the foundation of the world … this world … would be left without truth and without law."

Feeling convinced Lotze that the world is psychical and thus consists of souls as well as a personal deity. A soul is not simply a stream of impressions united by memory; it is a substantival entity, causally related to the body and interacting with it. Nevertheless, the soul is the greater influence and governs the body in ways closed to it. Both soul and body act according to law, but the laws of bodies as such are purely physical. The laws of the soul are on a higher level; they are teleological and unite the physical and the mental. They do not contradict the laws of the physical world, but they do control and reorder them.

A personal deity, God, follows from the existence of souls and ends. How else can they be explained? The world and everything in it is the personal creation of God and the means by which he attains his ends and the ends of his creatures. However, Lotze tempered this conviction by insisting that God attains his ends through the mechanisms or causal nexuses that science discovers, and he

further insisted that these mechanisms are characteristic of both the living and the nonliving.

Lotze rejected the notion that telic and nontelic explanations are incompatible with each other. He opposed the older forms of vitalism popular in his day. He argued that mechanisms are simply the instruments, or tools, by which God accomplishes his ends in the world. Although it is true that God might have used other means, he preferred mechanism as a way of establishing universal law in both the physical and psychic realms.

Lotze argued further that the thesis of mechanism should not be identified with materialism. Mechanism does not imply the nonexistence of ends or of psychic beings; it implies only the existence of uniform modes by which things come into being. The world can just as easily be psychic as material, but the psychic interpretation is rationally to be preferred because it does not make a mystery or a paradox of the presence of feeling and values in the world.

In the *Mikrokosmus,* Lotze continued to elaborate this position. Mechanism is simply a method of research; it is not a fundamental explanation of life and mind. Only the most exhaustive survey of the life of man can provide such an explanation and relate his life to the cosmos and God. Furthermore, mechanism does not repudiate free will; it is simply the necessary condition for the will to express its autonomy.

For Lotze, there are three realms of observations: the realm of fact, the realm of universal law, and the realm of values, which serve as standards of meaning for the world. These realms are only logically separable; they cannot be separated in reality. Fact and law are the means, the mechanisms, by which values are attained in this world; they are also the means by which men discover that certain values are foolish, contradictory, unrealizable, or in other words, false. Since fact and universal law are not existentially separable from value, God must also be the creator of everything and the quintessence of whatever deserves to exist for its own sake. Moreover, since feeling is fundamental, a sort of pluralistic idealism in which the only realities are living spirits and God in interaction is justified. All other realities are so only secondarily, as manifestations of these spiritual activities.

Lotze ultimately accepted a variant of Leibnizian monadism as a correct interpretation of experience. There is no single unity or oneness to existence. Direct experience reveals an irreducible multiplicity of things. Reality is always in flux, always involving constant doing and suffering. Nevertheless, the flux, the doing, and the suffering occur within a fixed order, a preestablished harmony between God and the multitude of spirits.

Lotze recognized that this metaphysical theory is neither a logical deduction from experience nor completely intelligible, but he believed it to be a reasonable inference from the manner in which the valid experiential concepts of our thought, the flux of facts, and the order of values interconnect in our experience. To limit ourselves to what science understands is to exclude unjustifiably the realms of feeling and values, and to exclude the latter is to render our experience unintelligible.

PIETY

To Lotze, nature and the social life are the two fundamental sources for religious ideas. From nature we derive the concept of God; from our social life, the concepts of ethical living. Paganism has tended to emphasize the cosmological; Christianity, the ethical. Christianity has sought to fuse both into one complete theological scheme. In this it is mistaken. To Lotze, the ethical element in religion is far more significant than the cosmological (which can properly be left to science), even though the emphasis upon cosmology leads to recognition of God. The true mark of the religious man is not his cosmology but his feeling for, and search after, what ought to be, his passion for, and loyalty to, the highest possible ideals. This passion and loyalty, however, are not so much activistic as contemplative.

Piety, for Lotze, is found in the inner life, in a feeling for the holy that attains so high a state of intuitive comprehension that logic, reason, becomes futile and inessential. Piety lies beyond any sectarian interest, Christian or not, for it drives men to seek a totality of feeling in which truth is completely fused with goodness and beauty. In consequence, the holy is not merely what men think it is; it is rather the unattained, the beyond in our lives that is without contradiction, defect, or dissonance. It is manifested in the endless striving for that immortal sea that is the infinite parent of all things. In this contemplative striving, in this endless search for total harmony and for the ought-to-be lies the possibility of progressively uniting science, religion, and art. However, the overwhelming realization of this unity occurs only at particular moments when one is moved by the experience of total beauty. In such moments, one knows absolutely that the fusion has, as far as possible, been accomplished.

INFLUENCE

Lotze's influence in Germany, France, and England was considerable during his lifetime. Philosophers became

more empirical-minded, less dogmatic. More consideration was given to the feeling, experiential aspects of human life. Nevertheless, Lotze left few, if any, disciples, and no Lotzean school of philosophers arose.

In America, Lotze's influence during the 1870s and 1880s was felt both in church and philosophical circles. The Reverend Joseph Cook of Boston made him widely popular, hailing him as the seer who had made the microscope the instrument of immortality and science the humble servant of the Bible. Leading American philosophers such as B. P. Bowne, G. T. Ladd, and Josiah Royce were particularly influenced, for he offered them an experiential mode of reconciling their strong Christian commitments with the methods and conclusions of science.

See also Bowne, Borden Parker; Fechner, Gustav Theodor; Herbart, Johann Friedrich; Metaphysics; Psychology; Royce, Josiah; Value and Valuation.

Bibliography

The principal writings of Lotze include *Allgemeine Pathologie und Therapie als mechanische Naturwissenschaften* (Leipzig, 1842); *Medicinische Psychologie oder Physiologie der Seele* (Leipzig: Weidmann, 1852); *Mikrokosmus,* 3 vols. (Leipzig, 1856–1864), translated by E. Hamilton and E. E. C. Jones as *Microcosmus,* 2 vols. (Edinburgh. 1885–1886); and *Die Geschichte der Aesthetik in Deutschland* (Munich, 1868). Lotze planned a comprehensive account of his philosophy in three volumes titled *System der Philosophie.* The first part was *Logik* (Leipzig: Hirzel, 1874), and the second, *Metaphysik* (Leipzig: Hirzel, 1879). These were translated into English and edited by Bernard Bosanquet as *Lotze's System of Philosophy* (Oxford: Clarendon Press, 1884). The third volume, which was to have covered religion, art, and practical philosophy, was incomplete at the time of his death. See also *Kleine Schriften,* edited by D. Peipers, 3 vols. (Leipzig, 1885–1891).

For literature on Lotze, see Karl Robert Eduard von Hartmann, *Lotzes Philosophie* (Leipzig: Friedrich, 1888); Henry Jones, *A Critical Account of the Philosophy of Lotze* (Glasgow, 1895); J. W. Schmidt-Japing, *Lotzes Religionsphilosophie in ihrer Entwicklung …* (Göttingen, 1925); E. E. Thomas, *Lotze's Theory of Reality* (London: Longmans, Green, 1921); and Max Wentscher, *H. Lotze. Lotzes Leben und Werke,* Vol. I (Heidelberg: Winter, 1913).

Rubin Gotesky (1967)

LOVE

"Love" as a concept enters philosophy at one point through religion, particularly when the origin of the world is expressed as an act of procreation or the Creator is conceived of as loving his creation either as a whole or in part (i.e., the human race). But the concept of love is also a subject for philosophic meditation in regard to ethical problems. Love, as one of the most powerful of human impulses, was early seen to be much in need of control, especially if man as rational animal was to be able to use his rational capacities. Much of the ethical writing on love is designed to suggest some means whereby the pleasures and other values of loving may be preserved without entailing the supposed evils of intemperate sexuality. This type of speculation ran from Plato through the Neoplatonists—those of both the early Christian period and the Italian Renaissance. In the Platonic tradition love had a unique metaphysical status, for it existed in both the material and the ideal worlds. Love can take on many forms, from gross sexual passion to a devotion to learning, but, it was argued, the ultimate object of love is the beautiful. The goodness that God sees in his creation is its beauty and to feel the beauty of the world is to love it and its Creator.

CLASSICAL MYTHOLOGY

The word *eros* as it is found in Homer is not the name of a god but simply a common noun meaning "love" or "desire." In Hesiod's *Theogony* Eros becomes one of the three primordial gods, the other two being Chaos and Earth. Although Eros has no offspring and seems to play no role in the genealogy of the gods, he has the greatest power over his fellow immortals. He unnerves the limbs and overcomes the reason of both gods and men. When Aphrodite is born from the sperm of Uranus (Heaven), Eros and Himeros (desire, longing, lust) accompany her into the council of the gods. Whether Hesiod was talking in terms of personalized abstractions or was actually thinking of anthropomorphic beings is not clear, for the *Theogony* is a curious mixture of both kinds of expression. For the history of philosophy, the importance of Hesiod's brief mention of Eros lies in the attribution to him of a power that is the enemy of reason. Something similar is to be found in Sophocles' *Antigone* in the chorus that is sung just after Creon has announced that Antigone must die for having buried her brother's body. Eros is addressed as the god who has brought about Antigone's tragedy. He is described as unconquerable, destructive, roaming over the sea and among the dwellers of the wilderness. Neither the gods nor ephemeral humankind can escape him; he drives his victims to madness and turns the just to evil. An even stronger denunciation of the god may be found in Euripides' *Hippolytus,* along with the additional warning that whether one surrenders to love or refuses to capitulate to it, one is

doomed. And indeed, Phaedra, whose successors are obviously Vergil's Dido and Racine's Phèdre, became the prototype of a woman ruined by Eros.

Such poetic passages reflect certain observations about human nature and human behavior. They point to a struggle within man's psyche between a rational, controllable, prudent, and wise agent and an irrational, uncontrollable, mad, and foolish agent. When the former is in control, man will behave in praiseworthy fashion, but when the latter gains the upper hand, he will act like a beast. He will abandon reason that, according to most of the ancients, alone distinguishes him from the beasts. Although man also has an animal nature, to yield to its demands is to betray his essential nature. The notion that Eros might reinforce the human element in man does not appear in the pre-Platonic writers.

EARLY PHILOSOPHIC REFLECTIONS

The Greeks admitted several forms of love, including heterosexual and homosexual passion; parental, filial, and conjugal affection; fraternal feeling; friendship; love of country; and the love of wisdom. All were associated with either Eros or *Philia* (fondness or friendship). Love was believed to be a power capable of uniting people in a common bond. And since not only people but also animals and the elements were thus united, it was appropriate to conceive of this power as lodged in a single agent that governed the whole cosmos. According to Parmenides, Love was created by the goddess Necessity, and in the writing of Empedocles, love emerges as one of the two universal forces (the other being strife) that explain the course of cosmic history. These two agents—the one of union, the other of decomposition—are not simply names for the fact that composition and decomposition occur; on the contrary, love and strife are not resident in things but are external to them and act upon them. According to Empedocles, the cosmos, so to speak, is held in tension between the forces of harmony and disunion. Were the two forces to be synchronously present, the world would clearly be in a state of disorder. Hence, Empedocles introduced the idea of cycles into his philosophy, as well as the concept of world history as an alternation of the reigns of Love and Strife. When Love is in control, the elements form compounds out of which arise more complex units and, eventually, animate beings. In the primitive period of the cycle, men worship Aphrodite, are innocent of slaughter and, presumably, of war, and are, moreover, vegetarians. "The altar did not reek of the unmixed blood of bulls, but this was the greatest abomination among men, to snatch out the life and eat the goodly limbs" (Fragment 128). But when Strife is dominant, disorganization, the ultimate disaggregation of the elements, and war and all its attendant evils, take the place of the blessings of love. As far as we can tell from the surviving fragments, Empedocles believed that the cyclical process was everlasting.

The attribution of peace and harmony to the goddess Aphrodite (Empedocles' name for love) is clearly a renunciation of the early poets' idea of love. Empedocles' conception of her resembles the *alma* Venus of Lucretius. Yet she remains the goddess of sexual love, for sexual love has become one example of the universal power of union: It provides the philosopher with empirical evidence of a metaphysical principle.

PLATO. For a complete expression of a philosophic concept of love, one must turn to Plato's *Symposium.* Probably no other document in European literature has had as much influence on the philosophy of love. The various speeches that are reported in this dialogue represent points of view with which Plato does not always agree but which he apparently thought important enough to be presented as typical. These speeches range from an encomium of love's effect on morality to a description of its effect on knowledge. Phaedrus likened the passionate attachment between Achilles and Patroclus to the conjugal affection between Alcestis and Admetus. In both cases it is the lover, not the beloved, who has gained virtue through his or her love. In the following speech, by Pausanias, two kinds of love are distinguished, that of the heavenly Aphrodite and that of the earthly Aphrodite, or the love of the soul and the love of the body. The former is more likely to be the love of a young man (not a boy) at the time when his reason begins to develop and his beard begins to grow. In this speech honorable love is clearly the attraction that a man has for a virtuous soul and is fused in the mind of the speaker with philosophy itself, which is the love of wisdom. It is this honorable love that Eryximachus then describes as the source of harmony and the preserver of the good.

The conclusion drawn from these encomiums is that love is in essence the love of beauty and that beauty is nothing material; it is an ideal. But no man desires the ideal until he has been educated through philosophic training. In the final speech, which supposedly reports the philosophy of the seeress Diotima, we find that there is a scale of beauty, progressing from that of bodies through that of forms, thoughts, minds, institutions and laws, the sciences, to absolute or ideal beauty.

Beauty, for Plato, was the one bridge between the two realms of the material and the ideal, particulars and universals. (This appears clearly enough in the *Phaedrus*; what the *Symposium* adds is a discussion of the power that draws men to beauty in its many modes.) The two realms present not simply a duality of kind but also of value, for the ideal and the universal, which are perfect and eternal, are always to be preferred to the material and the particular. Sexual love itself, although lowest on the scale of love, is nevertheless the seed of ideal love, since what attracts a man to the beloved is beauty.

ARISTOTLE. Plato's account of love, insofar as it concerns friendship, was amplified by Aristotle in the eighth and ninth books of the *Nicomachean Ethics.* But Aristotle treated chiefly the ethical and psychological aspects of the matter. He also utilized the metaphor of the attractive power of love in explaining the motion of the planetary spheres, the Unmoved Mover being the beloved and the planetary system the lover. With important differences that will be mentioned below, the Unmoved Mover became a part of the Christian concept of God.

TRANSITION TO CHRISTIANITY

In the *Magna Moralia,* which was probably composed at least in part by Aristotle, it is written that "It would be strange if one were to say that he loved Zeus.… It is not love towards God of which we are in search … but love towards things with life, that is, where there can be a return of affection." God then is thought to be incapable of returning our love for him, assuming that we can have love for him. In fact, although there are myths in which gods and mortals have been in love with each other, the gods always first disguise themselves as mortals, as Aphrodite did when she fell in love with Anchises, or take on various other forms, which was the habit of Zeus. These myths all deal with sexual intercourse, not with friendship or paternal affection. Omitting the culture heroes, there was no god or goddess in ancient mythology who had any love for humankind. Prometheus is an exception, but he was punished for his help to mortals, and in all probability the historic Greeks thought of him as simply a personification of forethought.

There is no god in classical religion who could be called "our father in heaven." The attitude that Lucretius tried to foster in the minds of his fellow Romans was supposed to be an antidote to their fear of the gods. According to the legends, however, there was good reason to fear them. Ceres and Bacchus may have given men bread and wine, but most of the divinities did little more than take revenge on the human race for the injuries they had received from their fellow gods. In Judaism and Christianity, however, a new relationship to the divinity was established. As early as Deuteronomy 6:5 the commandment was laid down to love God "with all thine heart, and with all thy soul, and with all thy might," a commandment repeated by Jesus (Matthew 22:37) as the first and great commandment, followed by the second, "Thou shalt love thy neighbor as thyself." It will be observed that now love is not seen as a power that destroys man's reason, but rather, as an emotional attitude that can be voluntarily produced. It is praised in the Psalms (for example, 91:14) and also in the First Epistle to the Corinthians and the First Epistle of John (I John 4:16–20). Both epistles cite the power of love to heal discord and fear, and love is represented as a bond between God and man. According to the Gospel of John (3:16), it is because of God's love for the world that redemption is brought to man.

That man could love God, even if he could not love Zeus, had been seen by Philo Judaeus in his *Questions on Genesis* (XVIII, 16) in which he says that once a man has received a clear impression of God and God's powers, his soul is filled with longing for union with God. Thus, in the First Epistle of John, God is identified with love, "and he that dwelleth in love dwelleth in God, and God in him" (I John 4:16). This idea was also found in non-Christian theologians of Hellenistic times, for example, in the *Hermetica (Asclepius* II, Sec. 21), in which all things, including God, are said to be bisexual, a unity that is approximated by men and women in sexual love. This unity is admittedly incomprehensible and what "you might correctly call either Cupid or Venus." But in both Philo and the *Hermetica,* as in Plotinus and Cleanthes' "Hymn to Zeus," the original stimulus to the love of God is knowledge, not sexual love. In *Asclepius* (XIII, 9) the love of God is reduced to worship, sacrifice, prayer, and reverence, and these follow upon a knowledge of the divine nature. In Plotinus the union with God, although aided by ascetic practices, is nevertheless the climax of cognition. Since knowledge occurs only between similar beings, to know God is to be like him; since God is unique, one must become absorbed into his being in order to know him. This may seem to be suggested in the verses from the First Epistle of John cited above, but actually in John the love of God, although it unites man and God, is an act of will similar to the love for one's fellow man. It would presumably be made manifest by one's acts and one's faith; it is not the conclusion or fulfillment of a metaphysical system.

Although the Church Fathers came closest to an identification of God with Aristotle's Unmoved Mover and later Christian philosophers gave God the attributes of that ontological principle, there were differences that have too often been obscured. The Unmoved Mover was neither a person nor a creator; he was uniquely able to produce change without being altered himself, and he could thus suffer no emotions whatsoever. The biblical God was the very antithesis of this. But in order to give an analogy of the way in which the Unmoved Mover moves the world, Aristotle took recourse to the metaphor of the beloved who attracts the lover. This, of course, became in time Dante Alighieri's "love which moves that sun and the other stars." For Aristotle, however, the Unmoved Mover could not return the love of the beings who are below him. In Christianity, as in Judaism, it was essential that God love his creatures as they love him, and, as previously mentioned, love seems to have been thought of as subject to volition. According to Plato (to limit the discussion to him), love arose involuntarily at the sight of a beautiful body. A man's erotic education consisted in a denial, after an analysis of the nature of beauty, of the acts that usually follow such a sight. Once that denial became a part of a man's character, he could rise to allegedly nobler beauties until the final goal—the contemplation of absolute beauty completely detached from anything corporeal—was reached.

The early Christians had more confidence in man's will than had their pagan contemporaries. Both love of God and religious faith were thought to be subject to volition. The concept of believing in order to understand, as St. Augustine put it, was based on the assumption that belief was not the effect but the source of understanding. To what extent the early Christian writers were aware of the psychological effect of practicing certain rites, as Pascal later was, is difficult to say. But since great emphasis was put upon ceremonious expressions of devotion and upon the refusal to carry out pagan rites, we can assume that the practices were believed to induce the appropriate emotions. The most famous of such ceremonies was the Christian agape, in which the devout met to share a supper and to rejoice in their common beliefs. The word *agape* means both love and the object of love, although the pagan satires treated it as if it meant a sexual orgy. The participants in the agape probably thought of it as a ceremony of brotherly love commemorating the Last Supper, although according to the testimony of the Epistle of Jude (12), it was abused at a fairly early date. Whatever its origin and its primitive significance, it is clear that it was supposed to be a ceremony of affection, and it reinforced the friendliness that members of the same religion might be expected to have toward one another. Two emotional factors that seem to have been absent from paganism thus came into prominence in early Christianity—fraternal love as an essential of piety and filial love to a divine father, both of which were reciprocated. These forms of love were strengthened by the persecutions to which the early Christians were subjected—persecutions that bound them together in a special community and led to self-sacrifice in the various forms of martyrdom.

AUGUSTINE. Of the Church Fathers, it is St. Augustine who gives us the most detailed analysis of love, ranging from his youthful sexual escapades to his final love of God. The famous opening of Book II of the *Confessions* described his condition as one of utter subservience to the flesh. Just as he was capable of enjoying sin (in his case, petty theft), not for the loot it brought him but for the joy of sinning, so he enjoyed love not for the sake of his beloved, but for the sake of his own self-centered pleasure. He described in vivid terms the loathing that invaded him while satisfying his passion. The death of a dear friend aroused in him a realization of the egocentricity of his passion, and in planning to organize a small group of fellow Christians who would live in charity and share their belongings (a plan that came to nothing), he first approached unselfish love. Through self-knowledge he learned to look upon the eternal light and ultimately came to the complete love of God, which he described in the tenth book of the *Confessions.* The fruit of this love was knowledge of the divine. Whereas for Plato and Philo cognition led to love, for Augustine it was love that led to cognition. This theme was developed in the twelfth century by such writers as William of Saint Thierry and St. Bernard of Clairvaux.

MIDDLE AGES

The ecstatic loss of self that accompanies sexual love was also assumed to be one of the features of the beatific vision. It is apparent in mystical literature that erotic language is especially effective in communicating mystical experience, and the similarities between religious and sexual ecstasy are manifest in, for example, the Song of Solomon. One should not conclude, however, that the medieval mystics were actually aware of the similarity between the beatific vision and sexual union, for those who are supposed to have made "mystic marriages," like the two St. Catherines, had presumably never had a corporeal marriage. Nonetheless, in mysticism the climax of the love of God was self-annihilation, much as in the Indian *mithuna,* and although the church never encour-

aged mystic practices, it had to admit their importance when they led to the immediate knowledge of God.

Thus, love in itself became an object of study, and the casuistry of love was elaborated in textbooks and poems as early as the twelfth century. Most of these writings seem to have taken as their source the *De Amore* of André le Chapelain which, whether intended to be serious or not, was taken seriously by most of its readers. It would appear to be a manual on seduction and to have only the most remote relevance to love. The time of its publication, however, coincided with the appearance of many commentaries on the Song of Solomon, and its influence on the rituals of the courts of love has been admitted by most medievalists. As the etiquette of the courts of love developed, love became an end in itself and was not necessarily to be gratified by sexual experience. The lover was supposed to serve his lady with no recompense other than the consciousness of his having served her.

One can only guess at how faithfully the precepts of courtly love were carried out, but as a set of ideas they form an important part of European moral philosophy. By elevating women to a position of irrefragable sovereignty over men, the ideals of courtly love became interwoven with the religious ideal of unquestioned loyalty to church and to God. The sovereign woman became identified with the Blessed Virgin to whom were applied many of the epithets of the bride in the Song of Solomon—rose of Sharon, the closed garden, the tower of ivory—phrases whose symbolical meaning had already been elaborated by St. Bernard. In the thirteenth century the question of the relative primacy of God's reason and will was disputed. For those who believed in the primacy of God's will, it followed that obedience rather than understanding was to be given the higher value. This was also true of courtly love and of chivalry as a doctrine.

DANTE. The culmination of the medieval writing on love is, for modern readers, Dante's *Vita nuova.* However else this book may be interpreted, it is the story of how love that begins with the sight of a girl's beauty ends with a vision which Dante intimated was to be that of the *Divine Comedy.* For Dante the Johannine phrase "God is love" was of essential importance in religion. In ending the *Divine Comedy* with the love that moves the sun and the other stars, he identified his own love and all love with the love that the cosmos has for its Creator. His "new life" was not to be fulfilled in a union with the woman whom he loved but in her guiding him through paradise. Few words occur more frequently in the poems of Dante than "amore." Sometimes he seems to be writing in the vein of courtly love, sometimes in the mystical vein of St. Bernard, but in both cases love is represented as a force that attracts man to a nobler life. Dante does not overlook the sufferings of a man in love; indeed, he emphasizes them. But to suffer because of love appears to be analogous to the sufferings of the martyrs—an abnegation of the self for a value that transcends egoism.

RENAISSANCE NEOPLATONISM

In Plotinus a distinction was made between three forms of love—love as a god, as a daemon, and as a passion. The first of these was again divided into the celestial and terrestrial Aphrodite. The celestial Aphrodite inspires the love of ideas and is the soul of the intelligible world. The terrestrial Aphrodite presides over marriage and is the soul of the sensible world. Love as a demon is identified with the souls of individual human beings. As a passion it is the love of beauty in temperate men and the love of sexual pleasure in those who dwell exclusively in the material world of ugliness. All love, however, is the love of some degree of beauty. Plotinus adopted the scale of beauties that had been outlined in the *Symposium* and read into it a hierarchy of being. At the apex stood the One; the "way up" to the One led from the beauty of material objects to that of ideas. In this instance one sees again the fusion of the erotic passion with the ecstasy of the mystic vision. Paradoxically, an experience that is intimately associated with our bodily life was thought of as the one escape from it.

This complex of confused ideas permeated Renaissance Neoplatonism. Philosophers such as Marsilio Ficino and Count Giovanni Pico della Mirandola constantly emphasized the power of love to free the soul from its bodily prison. They took over the theme of the two Venuses, and they assigned separate human faculties to each. They gave different names to the kinds of love—namely, divine, human, and animal.

LEONE EBREO. The philosophy of love expounded by Ficino and Mirandola was most fully developed by Leone Ebreo (Judah Abrabanel) in his *Dialoghi d'amore* (1501–1502), a work that circulated extensively not only in Italy but (in translation) through all Europe. Leone tied together the religious, philosophic, and literary traditions into a single network of ideas.

In the *Dialogues* the two interlocutors are Philo and Sophia, obviously elements of the word *philosophia.* Philo is the lover, and Sophia is the beloved. The first dialogue distinguishes between love and desire and describes the various forms of love; the second discusses the presence

of love in all natural operations, from the synthesis of the four elements to the movements of the planetary spheres; and the third deals with the love of God as the force that holds the universe together. Thus, it is asserted that love is a single principle permeating all things, from the material through the spiritual, and that this principle is the dynamic factor in cosmic change. There is no difference in essence between the attraction the elements have for one another and the forms of love that exist in human beings. The appraisal of the kinds of love is based on the objects of love, and Leone, like most of his contemporaries, thought that wisdom was inherently more valuable than pleasure.

It should be noted that the concept of a single dynamic power, whether it was called love or force or attraction, became more and more widely used as time went on. Its most extreme form was the *Sehnsucht* ("longing") of some German romantic philosophers, the *Streben* ("striving") of Johann Gottlieb Fichte, and Novalis's endless and unfulfilled search for the blue flower. One of the characteristics of love, at least in the mind of Leone, is its inability ever to be satisfied. Though Philo in the *Dialogues* pleads with Sophia to tell him that she responds to his love, she will not do so.

MODERN PERIOD

During the seventeenth and eighteenth centuries the interest in love was largely psychological and was expressed mainly in novels, poems, and maxims. While love of neighbor and God was approved, sexual love was morally more problematic. The ideal of female chastity was still upheld; in English novels, such as those of Samuel Richardson, a man was allowed to love a woman as long as he did not infringe upon her virginity. Whereas André le Chapelain graded sexual relations according to the social ranks of the maiden and her seducer, Richardson put all men and women on the same level in this respect. Thus, love was democratized. Sexual love was not to be condoned unless sanctified by the sacrament of marriage.

In such French novels as *Le grand Cyrus* by Madeleine de Scudéry, *Les liaisons dangereuses* by Choderlos de Laclos, and *La nouvelle Héloïse* by Jean-Jacques Rousseau, one finds more subtle distinctions and analyses. These authors continue the Renaissance casuistry about the different kinds of love and their respective values, but it must be remembered that their psychology of love was developed against the background of Christian moral principles. There is a constant conflict between the fervent religious and moral desire not to satisfy one's longings (described in *La nouvelle Héloïse*) and an awareness of the almost unlimited force of the individual's erotic desires (treated in *Les liaisons dangereuses*).

SPINOZA. The *Ethics* of Benedict de Spinoza was published in Holland in 1677. In this posthumous work, as in earlier publications, Spinoza emphasized man's need of perfection—that is, the fulfillment of both his intellectual and his emotional powers, which indeed were not existentially separate. He maintained that the more adequate an idea, the more it is pleasing, liberating, and intrinsically human. The culmination of the ethical life—that is, the life devoted to freedom of the intellect—is found in the "intellectual love of God." This phrase may have come from Leone Ebreo, but the idea goes back to St. Augustine. Both the *Confessions* and the *Ethics* are built on premises that are discovered by the intuitive process. The God of Spinoza is far from being the God of St. Augustine, but the method of finding him in the inner life and becoming aware of his presence is curiously similar. Both philosophers present a similar paradox: One must lose oneself in order to find oneself, but in so doing, one finds that what one has really discovered is God.

OTHER WRITERS. The analysis of love now passes into the hands of psychologists. Comte Destutt de Tracy and the novelist Stendhal both wrote books on love in which they attempted to probe its motivation and its effects upon conduct, but neither attempted to do more than to discuss love as a sexual experience. Destutt de Tracy's *De l'amour* was not published until 1926, although it may have been known in manuscript form; Stendhal's *On Love,* however, was published in 1822, and although it had no popular success at the time, it was later widely read. In Germany, on the other hand, such books as Johann Wolfgang von Goethe's *Sorrows of Young Werther, Elective Affinities,* and, of course, *Faust* gave a quasi-religious tone to the sexual experience. The impossibility of attaining complete satisfaction led men of this tendency to idealize Don Juan as a perfectionist who seeks a goal that he can never reach, for the ideal is precisely that which ought to be and never is. K. W. F. Schlegel's *Lucinde* is a perfect example of this interpretation of love as the ever-sought and unrealizable ideal.

SCHOPENHAUER. Arthur Schopenhauer was unique in condemning all forms of love on the grounds that they tie one to the will-to-live. But he found this will even in the subanimate world of nature; thus, he was reverting to the ancient tradition of an omnipresent principle and was more interested in the metaphysical status of this princi-

ple than in the details of human psychology. Although Schopenhauer's condemnation of love follows from his general metaphysical position, he supplemented this condemnation with an essay, "Metaphysics of the Love of the Sexes," in which he tried to show that poets and novelists had recognized the evil of loving, although they had not formulated the abstract principles that would justify this point of view. Love drives men and women to suicide, madness, and extremes of sacrifice. Pointing out that he has no philosophic precedents to guide him, Schopenhauer flatly declares that all forms of love are rooted in sexuality and that, obviously, the existence of future generations depends upon its gratification. But the sexual instinct can disguise itself in various ways, especially as "objective admiration," although in reality the will-to-live is aiming at the production of a new individual. Because sexual union exists for the benefit of the species, not for the individuals involved in it, marriages should not be made for love but for convenience. Thus, he says, there is guilt in loving, for its culmination is simply the perpetuation of the will-to-live, with all its attendant miseries.

FREUD. Historically, Schopenhauer's influence on Sigmund Freud is more important than his theory of the will-to-live in itself. Freud renamed the will-to-live the libido and at one time even saw its goal as death. The concept of the death wish paralleled Schopenhauer's emphasis on art and pity as the two ways of escape from life, and it had no great success in psychological circles. The libido as a term for generalized desire, on the other hand, has become part and parcel of the terminology of psychodynamics. Like most philosophic concepts, it has been distorted by both its supporters and its adversaries, but by reintegrating humanity and its strivings into the natural world, it has revived in a new form the kernel of Diotima's speech in the *Symposium.* Freud, along with most Platonists, would deny this. However, since love in the *Symposium* is found not only in sexual attraction but also in scientific research and philosophic meditation, there is only a verbal difference between the two philosophies. Freud, to be sure, does not preach the denial of bodily love, but at the same time he never denied the need for self-restraint and self-discipline. Although he may have said that the scientist is dominated by an anal-erotic urge, he did not deprecate science in these terms; rather, he explained what he thought was its general etiology. He also opened the door to a franker discussion of human motivations, and his contribution to ethics can hardly be overestimated. He attempted to show men how to realize the ideal of self-knowledge that philosophers had advocated for centuries without indicating how one might attain it. By pointing out the universality of love in its various forms and suggesting how it becomes deformed and alienated from its natural goals, Freud laid the foundation for an ethics that would be freed from ecclesiastical dogmatism. Although his followers have modified some of his ideas, as was inevitable, they have not denied either the preeminence of the libido as a driving power in human affairs or its ability to mask itself. One cannot overlook Freud's contribution toward giving men the ability to understand both one another and themselves—a type of understanding that had been preached over the centuries but always on the assumption that human nature could be observed in conscious behavior.

As is always the case in intellectual history, ancient beliefs survive and take on new forms. This is as true of the history of the idea of love as it is of other ideas. It is obvious that although no one believes any longer in the myth of the two Aphrodites as anthropomorphic deities each of whom is accompanied by a special Eros, the distinction between the two still persists as the contrast between carnal and spiritual love. The First Epistle of John and the Gospel of John have been by no means discarded in the Occident, nor has the commandment to love God and one's neighbor been forgotten. *Caritas* as both brotherly love and charity is still preached, if not practiced, and the Neoplatonic notion that through love we shall have harmony and through harmony, peace, is as potent a force in social education as it has ever been. Philosophy sometimes takes as its goal the rationalization of common sense, or at least of widely held beliefs, and according to the available evidence, no one has ever maintained that the whole duty of man consists in hating, provoking disorder, and disobeying what are at various times called the laws of God or of nature. Philosophers writing on love have attempted in numerous ways, first, to describe the unique part it plays in human life; second, to seek its similarity to other impulses; third, to appraise the ends that it wishes to achieve; and finally, to work out a systematic account of all these distinctions and put them into a logical network of ideas.

See also Aristotle; Augustine, St.; Beauty; Bernard of Clairvaux, St.; Dante Alighieri; Destutt de Tracy, Antoine Louis Claude, Comte; Empedocles; Fichte, Johann Gottlieb; Ficino, Marsilio; Freud, Sigmund; Goethe, Johann Wolfgang von; Neoplatonism; Parmenides of Elea; Pascal, Blaise; Perfection; Pico della Mirandola, Count Giovanni; Plato; Plotinus; Rousseau, Jean-Jacques; Schlegel, Friedrich von; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de.

Bibliography

PRIMARY SOURCES

Aristotle. *Nicomachean Ethics.* Translated by W. D. Ross. Oxford, 1925. Books 7–8.

Augustine. *Confessions.* Translated by E. B. Pusey. London, 1907. Book 1, Ch. 12; Book II, Chs. 1 and 6; Book III, Ch. 1; Book IV, Ch. 6; Book VI, Ch. 14; Book VII, Ch. 10; Book VIII, Ch. 2; Book X, Ch. 7.

Bernard. *De Diligendo Deo*; Sermon, *In Cantica Canticorum.* In *Patrologia Latina,* edited by J. P. Migne. Paris, 1844–1864. Vols. 182–183, respectively.

Chapelain, André le. *De Amore.* Edited by E. Trojel. Havniae, 1892.

Cleanthes. *Hymn to Zeus.* Translated by George Boas in *Rationalism in Greek Philosophy.* Baltimore: Johns Hopkins Press, 1961. P. 251.

Dante. *Divine Comedy.* Translated by H. F. Carey, 2nd ed. London, 1819.

Dante. *Il convivio.* Translated by P. H. Wicksteed as *The Convivio of Dante Alighieri.* London: Dent, 1903.

Dante. *Vita nuova.* Translated by Dante Gabriel Rossetti as *The New Life of Dante Alighieri.* London, 1901.

Empedocles. Fragments 115, 128, 130. In *Early Greek Philosophy,* by J. Burnet, 3rd ed. London, 1920; New York, 1957.

Euripides. *Hippolytus.*

Freud, Sigmund. *A General Introduction to Psycho-Analysis.* Translated by Joan Riviere. Garden City, NY, 1943.

Freud, Sigmund. *Three Contributions to the Theory of Sex* and *Totem and Taboo.* In *The Basic Writings of Sigmund Freud.* New York: Random House, 1938.

Hermetica (Asclepius II, 21). Edited by A. D. Nock and A. J. Festugière. Paris, 1960. Vol. II.

Laclos, Choderlos de. *Les liaisons dangereuses.* Garden City, NY: Doubleday, 1961.

Leone Ebreo. *The Philosophy of Love.* Translated by F. Friedberg-Seelye and Jean H. Barnes. London: Soncino Press, 1937.

Lucretius. *De Rerum Natura,* Book I, 1–49. In *The Stoic and Epicurean Philosophers,* by W. J. Oates. New York: Modern Library, 1957.

Plato. *Symposium.* In *Plato: Dialogues,* edited by B. Jowett, 4th rev. ed. Oxford, 1953.

Pseudo-Aristotle. *Magna Moralia.* Translated by St. George Stock. Oxford, 1915. 1208 b, 31.

Schopenhauer, Arthur. *The World as Will and Idea.* Translated by R. B. Haldane and J. Kemp, 6th ed. London, 1907. See especially Vol. III, Ch. 44, "Supplements to the Fourth Book." London, 1909.

Spinoza, Benedict de. *Ethics,* Part 5. In *The Chief Works of Spinoza,* translated by R. H. M. Elwes, 2 vols. New York, 1955.

SECONDARY SOURCES

Bruyne, Edgar de. *Études d'esthétique médiévale.* Bruges: De Tempel, 1946. Vol. III, Book IV, Ch. 2.

Gilson, Étienne. *The Spirit of Mediaeval Philosophy.* Translated by A. H. C. Downes. New York, 1940. Ch. 14.

Mackail, J. W. *Select Epigrams from the Greek Anthology.* London, 1906. See especially the introduction, Sections vi–vii, pp. 32–41.

Panofsky, Erwin. *Studies in Iconology.* New York: Oxford University Press, 1939. Chapters 5–6.

George Boas (1967)

LOVE [ADDENDUM]

Since the middle of the twentieth century, analytic philosophers have taken diverse interests in love. Philosophers of mind have asked what kind of psychological state love is. A natural answer is that love is an emotion like any other. Some philosophers, however, find love to be an anomalous emotion, or even not to be an emotion at all. Most types of emotions seem to be triggered by, or partially to consist in, a belief that the emotion is warranted by some fact about its object. Fear of something, for example, typically involves the thought that the thing feared is dangerous or threatening. Love seems to be an exception, since it is unclear what fact about one's beloved might warrant one's love for this person. Some are willing to accept love as an emotion despite this anomaly, while others insist that love must be a psychological state of a different kind. The most commonly proposed alternative is that love is a desire, or set of desires, regarding one's beloved.

The view that love is an anomalous emotion stems from a perception that nothing warrants or justifies it. This raises a second issue that has occupied philosophers: whether there are reasons for love, and if so, what these reasons might be. The most natural candidates for reasons for love would seem to be properties or qualities of the beloved, such as wit, beauty, or kindness. Among many problems with this proposal, three have attracted especially close attention. First, some find the proposal fetishistic, or at least misdirected. It appears to represent love as focused on the beloved's accidental properties, rather than on that person's essence. Second, if one's reasons for loving the beloved are properties, then one's love ought to wane as the beloved loses those properties. This seems at odds with the thought, famously expressed by William Shakespeare, that "Love is not love/Which alters when it alteration finds." Finally, if one's reasons for loving the beloved are properties, then insofar as one's love is responsive to those reasons, it will soon migrate to another person with those properties in sufficient proportion. This too seems antithetical to love.

Impressed by some of these problems, Harry Frankfurt concludes that while love creates reasons, there are no reasons for love. Love is a structure of desires for which there is no antecedent justification. Love is focused on the particular person whom one loves; it is not a response to some generalizable, justifying property that the person has. Since Jane, say, is the particular person she is and she can neither lose this trait nor share it with anyone else, one's love for her does not alter as it alteration finds, nor does it transfer to her twin. David Velleman (1999), resisting Frankfurt's conclusion, suggests that love is a response to a justifying feature that is also identical with the beloved's essence: Jane's rational nature or capacity for valuation, for instance. However, this suggestion seems to leave one's beloved vulnerable to being replaced—indeed, replaced by any other person with a rational nature. A different strategy for avoiding Frankfurt's conclusion is to suggest that love is a response to the reasons provided by one's shared history with the person one loves. This would explain why one's love does not alter as the beloved's wit or beauty fades, and why one's love does not accept a substitute with whom no such history is shared. However, the appeal to shared history again threatens to make love focused on the beloved's accidental properties, rather than on that person's essence. It also seems to put the cart before the horse. Love seems to precede many relationships, rather than develop with them.

Moral philosophers have been particularly concerned that love, and similar attitudes such as friendship, are in tension with morality, at least as understood in certain theories. The tension is thought to arise because these moral theories—most notably, utilitarianism and Kantianism—require one to be impartial, that is, to give equal weight to everyone's interests. Love, in contrast, seems to impel one to be partial: to give greater weight to the interests of one's beloved. The tension has been thought to be more acute at the level of deliberation than at the level of action. While there may be utilitarian and Kantian justifications for permissions, or even requirements, to act as love directs, deliberating in terms of such justifications seems incompatible with love. This incompatibility has generally been seen as a problem for such moral theories, rather than as a problem for love. The incompatibility makes these moral theories seem self-defeating or overly demanding, or it reveals that they fail to take into account something of genuine value.

In defense of these moral theories, some philosophers have insisted that the incompatibility is only apparent. Indirect utilitarians have pointed out that while utilitarianism requires one to do what is best from an impartial standpoint, utilitarianism need not require one to deliberate in impartial terms. Indeed, there may be strong utilitarian reasons for not so deliberating. Kantians have similarly observed that the moral agent need not always be guided by specific reflection on what it is morally permissible to do. A less concessive Kantian response appears in Velleman's work. Love, he argues, is a "moral emotion," by which he seems to mean, at least in part, that love is animated by the same value that underlies morality itself.

Other philosophers, however, have insisted that the incompatibility is real. Some of these philosophers urge rejecting impartial moral theories, perhaps in favor of a virtue-based approach. Others see the incompatibility as casting doubt not on the impartiality of morality, but instead on its authority over our lives.

See also Friendship; Moral Psychology; Virtue and Vice; Virtue Ethics.

Bibliography

Frankfurt, Harry. *Necessity, Volition, and Love*. Cambridge, U.K.: Cambridge University Press, 1999.

Frankfurt, Harry. *The Reasons of Love*. Princeton, NJ: Princeton University Press, 2004.

Green, O. H. "Is Love an Emotion?" In *Love Analyzed*, edited by Roger E. Lamb. Boulder, CO: Westview Press, 1997.

Keller, Simon. "How Do I Love Thee? Let Me Count the Properties." *American Philosophical Quarterly* 37 (2) (2000): 163–173.

Kolodny, Niko. "Love as Valuing a Relationship." *Philosophical Review* 112 (2) (2003): 135–189.

Taylor, Gabriele. "Love." *Proceedings of the Aristotelian Society* 76 (1975–1976): 147–164.

Velleman, David. "Love as a Moral Emotion." *Ethics* 109 (2) (1999): 338–374.

Williams, Bernard. "Persons, Character, and Morality." In his *Moral Luck*. Cambridge, U.K.: Cambridge University Press, 1981.

Wolf, Susan. "Morality and Partiality." *Philosophical Perspectives* 6 (1992): 243–259.

Niko Kolodny (2005)

LOVEJOY, ARTHUR ONCKEN

(1873–1962)

Arthur Oncken Lovejoy, the American philosopher and historian of ideas, was born in Berlin, Germany, the son of the Reverend W. W. Lovejoy of Boston and Sara

Oncken of Hamburg. Educated at the University of California (Berkeley) and at Harvard, where he received his MA, Lovejoy began his teaching career at Stanford University (1899–1901) and then taught for seven years at Washington University in St. Louis. After short periods at Columbia University and the University of Missouri, he went to Johns Hopkins in 1910 as professor of philosophy, remaining there until his retirement in 1938. In 1927 he gave the Carus Lectures, published as *The Revolt against Dualism* in 1930, and the William James Lectures, published as *The Great Chain of Being* in 1933. Lovejoy was widely known as an epistemologist, a philosophic critic, a historian of ideas, and a man of action. He helped to organize the Association of American University Professors, in which he served for many years as chairman of the group that investigated all charges of violation of academic freedom. In this connection he wrote the article "Academic Freedom" for the *Encyclopaedia of the Social Sciences*.

Lovejoy's works fall into two main groups—those on epistemology and those on intellectual history—although he also wrote essays on ethics, religion, and social problems.

PHILOSOPHICAL WORKS

For many years Lovejoy confined his writings to articles, a great number of them critical. These were often directed against various forms of anti-intellectualism: "The Thirteen Pragmatisms" (1908), "Some Antecedents of the Philosophy of Bergson" (1913), and "The Paradox of the Thinking Behaviorist" (1922). These articles, however, were frequently examinations of certain contemporary movements in philosophy, such as the New Realism: "Reflections of a Temporalist on the New Realism" (1911) and "On Some Novelties of the New Realism" (1913). Some were even on the supposed philosophical implication of the theory of relativity: "The Travels of Peter, Paul and Zebedee" (1932) and "The Paradox of the Time-Retarding Journey" (1931).

It was not until 1930 that Lovejoy published his major work, *The Revolt against Dualism*, in which he attempted to defend epistemological dualism against the reigning modes of monism. He began by sketching what he called naive dualism, which assumes that (1) many possible objects of knowledge (*cognoscenda*) are at places external to the body of the percipient; (2) man must have real traffic with things that existed in the past and may exist in the future; (3) man can have knowledge of things as they would be if they were not directly known; (4) other minds and experiences exist; and (5) *cognoscenda* in other places and at other times are apprehensible by other knowers. The book analyzed this naive dualism and defended a corrected form of it. On the whole, although not in detail, Lovejoy was more interested in the duality of two existents (of two five-cent stamps, for instance) than qualitative duality such as of red and green. The duality of two things is demonstrated, he wrote, by the fact that one of the supposed pair has a spatial, a temporal, or a spatiotemporal position that is inconsistent with that empirically exhibited by the other. If, then, it can be shown that our ideas of objects have positions that can be shown not to be those of the objects, then the two cannot rightly be believed to be one. Qualitative duality would be demonstrated in analogous fashion, but the inconsistency would lie between two sets of qualities.

In his autobiographical essay, "A Temporalistic Realism," in Volume II of *Contemporary American Philosophy*, edited by G. P. Adams and W. P. Montague (London and New York, 1930), Lovejoy pointed out that one of his earliest philosophical theses was that experience itself is temporal. Any philosophical position that overlooks or denies this, or conflicts with it, would, in his opinion, be condemned as contradicting a manifest truth. (This does not, of course, assert that any philosophy—such as that of Henri Bergson—that admits the empirical reality of time is thereby proved.) The various forms of monism fail to evade, and cannot evade, the consequences of this fact. For instance, the date at which a visual datum occurs is not the date of the object that one is seeing. There is a time lag between the emission of light rays from a star and their arrival at the retina of a human eye, to say nothing of the arrival of the nerve current stimulated by them at the cerebral cortex, where it apparently causes a visual image to appear. Indeed, some stars that we perceive now may have become extinct many light-years ago. Analogous statements can be made about sound, odor, and taste.

Although Lovejoy also used other criteria, this criterion of duality suffices to establish existential duality between object and sensum. To deny the duality, Lovejoy asserted, would be equivalent to asserting that two particulars can each be in two places at the same time, that one particular has or consists of many shapes and other inconsistent qualities at the same time, that it has two dates in the same temporal order, that it can be at the same time both the beginning and the end of a causal series, and, finally, that error is impossible. Lovejoy discussed each of these theses in connection with epistemological positions widely held at the time the book was written: the New Realism, objective relativism, Alfred

North Whitehead's denial of simple location, and Bertrand Russell's epistemology as given in *The Analysis of Mind* (London and New York, 1921) and *The Analysis of Matter* (London and New York, 1927).

Lovejoy's dualism differed from that of the naive dualist in that the latter is likely to believe that his objects are qualitatively, if not existentially, identical with the objects of others. Our ideas, Lovejoy held, do not necessarily have properties identical with the properties of anything in the physical world, but we are not therefore condemned to know nothing whatsoever of that world. We cannot prove beyond doubt that some of the properties of our ideas are also properties of the physical world, but such is "a natural assumption which no one can prove to be false" (*Revolt against Dualism*, p. 273). Qualities that vary with percipients must be held to be subjective, but there are certain residual properties—extension, shape, relative position, temporal succession, and motion—that may reasonably be said to characterize both our ideas and their objects. The reasonableness of the hypothesis rests on its ability to give us grounds for framing a "coherent, simple, unifying, scientifically serviceable" set of hypotheses for explaining both the rise of our sensory data and their peculiar characteristics. It will, in short, account for a world that is causally efficacious, that exists between our perceptual moments, and that has a past and future independent of any percipients.

INTELLECTUAL HISTORY

To separate Lovejoy's philosophical views from his historical studies is artificial, for his philosophy is based on a wide knowledge of history, and his historiography is based on his belief in the existence and efficacy of ideas. However, such a distinction may be made for purposes of classification.

Lovejoy was the chief promoter in the United States of the historiography of ideas. His continuing interest in this area dated back at least to his monograph *The Dialectic of Bruno and Spinoza* (Berkeley, CA, 1904). He was the originator and first editor of the *Journal of the History of Ideas*. He studied such general ideas as romanticism, evolutionism, naturalism, and primitivism, showing the ambiguities resident in them and their ingression into fields that have no ostensible logical connection with them.

In the preface to *Essays in the History of Ideas*, Lovejoy defined his conception of the historiography of ideas: (1) It studies the presence and influence of the same ideas in very diverse provinces of thought and in different periods; thus, an idea that may have originated in logic may turn up in biology, or vice versa. (2) There are certain catchwords, such as *nature*, that have taken on new meanings over time, although the people using them are seldom aware of their ambiguities. The historian of ideas will analyze these various meanings as they occur. An example from fairly recent history (not one of Lovejoy's own) would be the eulogistic usage of the word *organic*. (3) It has also been noticed that a given author will prove susceptible to the emotional aura of certain terms and, probably because of this, will waver between a valid meaning of an idea and an incongruous meaning. It is usually assumed that the thought of a given writer must be consistent and unified; but by accepting this assumption, a historian may overlook precisely those thoughts expressed by a writer that were in fact influential. A fuller explanation of the program is given in Lovejoy's essay "The Historiography of Ideas," first published in 1938 and republished as the opening chapter in *Essays in the History of Ideas*.

THE GREAT CHAIN OF BEING. Lovejoy's most influential single contribution to the history of ideas is *The Great Chain of Being*. The idea whose fortunes he traced in this book was first expressed by Plato in the *Timaeus*. There Plato maintained that the Demiurge, being good, was not jealous and, not being jealous, wanted the world to lack nothing; therefore, if the world were to lack nothing, all possibilities must be realized. The realization of all possibilities is the great chain of being, and the principle it rests upon was called by Lovejoy the principle of plenitude.

This apparently simple idea, contained in a creation myth, was introduced into Christian theology through Neoplatonism and into cosmography by Hasdai Crescas with his supposition of many worlds, by Johannes Kepler, by Nicholas of Cusa with his theory of a boundless universe, and, above all, by Giordano Bruno with his open acceptance of the principle as it applies to stellar bodies. In Benedict de Spinoza it appeared as the doctrine that all ideas of God must be realized, and in Gottfried Wilhelm Leibniz as the principle of sufficient reason. Lovejoy showed how the principle entered into biological speculations in the eighteenth century and how it was "temporalized." In the idea of the great chain of being, which he presented with a richness of erudition, Lovejoy found one of the most fertile yet neglected ideas in Western philosophy and masterfully traced its ramifications and subsequent history.

PRIMITIVISM. A second dominant idea, the study of whose history Lovejoy initiated, is that cluster of notions known as primitivism. Primitivism has two forms—a

chronological form, exemplified in the myth of the Golden Age, and a cultural form, best exemplified in cynicism and in all attempts to rediscover the so-called natural life. Each of these forms has two subspecies, "hard" primitivism and "soft" primitivism. Hard primitivism maintains that the state of nature (man's primordial condition) was rugged and unencumbered with superfluities, a state very close to that of the legendary noble savage. Soft primitivism, on the contrary, maintains that the state of nature was agreeably gentle, that earth gave man her fruits spontaneously without any labor on his part, and that there was no private property and hence no covetousness, no war, no foreign trade, none of the complications that the arts and sciences introduce.

Lovejoy urged as early as 1917 that there would be more progress in philosophical studies if there were more cooperation among philosophers ("On Some Conditions of Progress in Philosophical Inquiry"). A documentary history of primitivism provided, it seemed, an ideal opportunity for such cooperation. Lovejoy and three other scholars formed a team and agreed to publish a four-volume work, to be titled *A Documentary History of Primitivism and Related Ideas*, covering the ground from early Greek times to the recent past. Of this projected work only one volume, *Primitivism and Related Ideas in Antiquity*, written by Lovejoy with George Boas, was completed, although a number of smaller works by various scholars came out as contributions to the subject. The published volume contained, along with documents and commentaries, two supplementary essays—"Primitivism in Ancient Western Asia," by W. F. Albright, and "Primitivism in Indian Literature," by P.-E. Dumont—and an appendix by Lovejoy—"Some Meanings of 'Nature.'" Although the original four-volume plan was never carried out, what did appear may have shown historians of philosophy that primitivism was a philosophic theme neglected by the historical tradition that had nevertheless permeated Occidental thought.

See also Bruno, Giordano; Crescas, Hasdai; History and Historiography of Philosophy; Kepler, Johannes; Leibniz, Gottfried Wilhelm; Neoplatonism; Nicholas of Cusa; New Realism; Realism; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bibliography

WORKS BY LOVEJOY

"The Thirteen Pragmatisms." *Journal of Philosophy* 5 (1908): 1–12, 29–39.

"Reflections of a Temporalist on the New Realism." *Journal of Philosophy* 8 (1911): 589–599.

"Some Antecedents of the Philosophy of Bergson." *Mind*, n.s. 22 (1913): 465–483.

"On Some Novelties of the New Realism." *Journal of Philosophy* 10 (1913): 29–43.

Bergson and Romantic Evolutionism. Berkeley, CA, 1914.

"On Some Conditions of Progress in Philosophical Inquiry." *Philosophical Review* 26 (1917): 123–163.

"The Paradox of the Thinking Behaviorist." *Philosophical Review* 31 (1922): 135–147.

The Revolt against Dualism: An Inquiry concerning the Existence of Ideas. La Salle, IL: Open Court, 1930.

"The Paradox of the Time-Retarding Journey." *Philosophical Review* 40 (1931): 48–68, 152–167.

"The Travels of Peter, Paul and Zebedee." *Philosophical Review* 41 (1932): 498–517.

Primitivism and Related Ideas in Antiquity. Written with G. Boas. Baltimore, MD: Johns Hopkins Press, 1935.

The Great Chain of Being: A Study of the History of an Idea. Cambridge, MA: Harvard University Press, 1936.

Essays in the History of Ideas. Baltimore, MD: Johns Hopkins Press, 1948. Contains a list of Lovejoy's articles up to and including 1947.

Reflections on Human Nature. Baltimore, MD: Johns Hopkins Press, 1961.

The Reason, the Understanding, and Time. Baltimore, MD: Johns Hopkins Press, 1961. Deals with the Romantic theory of knowledge.

Essays in the History of Ideas. Baltimore, MD: John Hopkins Press, 1970.

WORKS ON LOVEJOY

Boas, George. "A. O. Lovejoy as Historian of Philosophy." *Journal of the History of Ideas* 9 (4) (October 1948): 404–411.

Mandelbaum, Maurice. "Arthur O. Lovejoy and the Theory of Historiography." *Journal of the History of Ideas* 9 (4) (October 1948): 412–423.

Montague, W. P. "My Friend Lovejoy." *Journal of the History of Ideas* 9 (4) (October 1948): 424–427.

Montague, W. P. "Professor Lovejoy's Carus Lectures." *Journal of Philosophy* 25 (11) (May 24, 1928): 293–296.

Murphy, Arthur E. "Mr. Lovejoy's Counter-revolution." *Journal of Philosophy* 28 (1931): 29–42, 57–71.

Nicholson, Marjorie. "A. O. Lovejoy as Teacher." *Journal of the History of Ideas* 9 (4) (October 1948): 428–438.

Spencer, Theodore. "Lovejoy's *Essays in the History of Ideas.*" *Journal of the History of Ideas* 9 (4) (October 1948): 439–446.

Taylor, H. A. "Further Reflections on the History of Ideas: An Examination of A. O. Lovejoy's Program." *Journal of Philosophy* 40 (May 27, 1943): 281–299.

Wiener, Philip P. "Lovejoy's Rôle in American Philosophy." In *Studies in Intellectual History*, 161–173. Baltimore: Johns Hopkins Press, 1953. A collection of essays by a group of Lovejoy's colleagues.

George Boas (1967)
Bibliography updated by Michael J. Farmer (2005)

LOYALTY

"Loyalty," as a moral rather than a political concept, has received scant attention in philosophical literature. In fact, at the present time [1967] it seems banished from respectable ethical discussions, owing, no doubt, to its historical association with an obsolete metaphysics (idealism) and with such odious political movements as the extreme nationalism of Nazism. However, the supposed implications suggested by these disreputable associations are ill-founded. On the contrary, loyalty is an essential ingredient in any civilized and humane system of morals.

Philosophical issues regarding loyalty may be separated into the question of the object of loyalty, and the question of the moral value of loyalty.

THE OBJECT OF LOYALTY

Granted that loyalty is the wholehearted devotion to an object of some kind, what kind of thing is this object? Is it an abstract entity, such as an idea or a collective being? Or is it a person or group of persons?

The idealist contends that loyalty is "the willing and practical and thoroughgoing devotion of a person to a cause" (Josiah Royce, *The Philosophy of Loyalty,* p. 17). Its object is "a cause beyond your private self, greater than you are … impersonal and superpersonal" (ibid., pp. 19–20). As a cause it is something that transcends the individual, "an eternal reality." Apart from familiar metaphysical and logical objections to this concept of a superpersonal reality, this view has the ethical defect of postulating duties over and above our duties to individual men and groups of men. The individual is submerged and lost in this superperson not only ontologically but also morally, for it tends to dissolve our specific duties and obligations to others into a "superhuman" good.

Opposing the idealistic position is the view, characteristic of social atomism (empiricism or utilitarianism, for example), that denies any distinctive status to loyalty on the grounds that metaphysically there can be no such superpersonal entity to serve as its object. Insofar as the concept of loyalty has any validity at all, it reduces to other kinds of relations and dispositions, such as obedience or honesty. Most empiricists are inclined to agree with David Hume, however, that loyalty is a virtue that holds "less of reason, than of bigotry and superstition."

Thus, it is generally assumed that we must either accept the notion of a superperson or some other abstract entity as the object of loyalty or reject the notion of loyalty altogether as founded on an illusion. This assumption is open to question.

In answer to the idealists, it should be pointed out that in our common moral language, as well as historically, "loyalty" is taken to refer to a relationship between persons—for instance, between a lord and his vassal, between a parent and his children, or between friends. Thus, the object of loyalty is ordinarily taken to be a person or group of persons.

Loyalty is conceived as interpersonal, and it is also always specific; a man is loyal to *his* lord, *his* father, or *his* comrades. It is conceptually impossible to be loyal to people in general (to humanity) or to a general principle, such as justice or democracy.

The social atomist fails to recognize the special character and significance of the ties that bind individuals together and provide the basis for loyalties. Loyalty is not founded on just any casual relationship between persons, but on a specific kind of relationship or tie. The special ties involved arise from the twofold circumstance that the persons so bound are comembers of a specific group (community) distinguished by a specific common background and sharing specific interests, and are related in terms of some sort of role differentiation within that group. A friendship, a family, or such a highly organized group as a political, priestly, or military community illustrates the presence of these conditions. Special ties of this sort provide both the necessary and the sufficient conditions for a person to be a proper object of loyalty.

The impersonal or objective element mentioned by Royce and other idealists is explained by the fact that it is the ties, the mutually related roles, rather than any particular personal characteristics of the individuals involved that provide the grounds for loyalty. Why should I be loyal to *X*? Because he is *my R* (friend, father, leader, comrade). More purely personal characteristics of *X,* such as his kindness, courage, amiability, honesty, or spirituality cannot serve as *grounds* for loyalty. That the conditions of loyalty abstract from the personal characteristics of the individuals concerned does not, of course, entail that loyalty must relate to a superpersonal entity (cause, whole) any more than the fact that an algebraic formula contains a variable within it (such as *Fx*) entails that there must be some kind of supernumber to satisfy the function.

THE MORAL VALUE OF LOYALTY

Is loyalty something good in itself? Is it always good? Can there be bad loyalties?

On these questions the idealist takes an extreme position, for he holds that loyalty is the highest moral good. According to Royce, a man's wholehearted devotion

to a cause is *eo ipso* good and becomes evil only when it conflicts with other loyalties. The supreme good is loyalty to loyalty: "so choose and so serve your individual cause as to secure thereby the greatest increase of loyalty amongst men" (ibid., p. 121).

The view that loyalty has an inner value, "whatever be the cause to which this man is loyal," can be used to redeem the most evil acts of men. Such a belief outrages our moral feelings, for we want to say that a cause which demands injustice or cruelty as the price of devotion renders that devotion an evil in itself. It is impossible to separate logically the moral quality of devotion from the moral quality of its object, if that object is a cause. (Incidentally, a distinction must be made between devotion to a thoroughly evil purpose and devotion that is simply misdirected, in the sense that it is well-intentioned but wrong for some other reason.)

Even assuming that the problem of bad loyalties can be resolved by invoking "loyalty to loyalty," the idealist may still be accused of turning morality, which properly concerns man's relations to his fellows, into service of an abstract principle or a cause, thus treating man as a mere means rather than as an end-in-itself.

The social atomist, on the other hand, regards the moral value of loyalty, construed as devotion or obedience to persons or institutions, entirely as a function of its benign or mischievous consequences. This view, however, robs loyalty of any special moral significance. It fails to account, for example, for the admirable side of a mother's loyalty to her son even when, considering the total picture, it is not entirely justified morally.

We must ask what loyalty demands of a person. The etymology of the word *loyalty* gives a clue, for it comes from the French word *loi* and thus means something akin to *legality.* Loyalty, strictly speaking, demands what is morally due the object of loyalty. A loyal subject is one who wholeheartedly devotes himself to his duties to his lord. What is due or owed is defined by the roles of the persons concerned. The fact that loyalty gives what is due also explains why we can demand the loyalty of others.

It follows that mere blind obedience to every wish of the person who is the object of loyalty is not loyalty; it is a perversion of loyalty. There is no moral value to it at all, since it is not something that is morally due. A loyal Nazi is a contradiction in terms, although a loyal German is not.

There are, to be sure, conflicts of loyalties, but this fact does not entail that any of the loyalties involved are improper or invalid. It is simply a logical consequence of the fact that there are conflicts of duties; my duty to my parents may conflict with my duty to my wife or to my fellow countrymen. Sometimes there are clear ways of resolving these conflicts and sometimes there are not, but we cannot eliminate the problem of conflicting loyalties either by a metaphysical trick or by the mechanical application of a value calculus.

One final observation must be made concerning the distinction between loyalty and fidelity. Loyalty includes fidelity in carrying out one's duties to the person or group of persons who are the object of loyalty; but it embraces more than that, for it implies an attitude, perhaps an affection or sentiment, toward such persons. Furthermore, at the very least, loyalty requires the complete subordination of one's own private interest in favor of giving what is due, and perhaps also the exclusion of other legitimate interests. In this sense, loyalty may often be one-sided, although it need not be. If we could not count on the loyalty of others or give them our loyalty, social life would be not only bleak but also impossible.

See also Atomism; Empiricism; Hume, David; Idealism; Royce, Josiah; Utilitarianism.

Bibliography

Aristotle. "On Friendship." In *Nicomachean Ethics.* Books VIII and IX.

Bryant, Sophie. "Loyalty." In James Hasting's *Encyclopaedia of Religion and Ethics.* Vol. VIII, 183–188. New York: Scribners, 1916.

Rashdall, Hastings. *Theory of Good and Evil.* Vol. 1, 188 and 273. London: Oxford University Press, H. Milford, 1924. For the social atomist view of the moral value of loyalty.

Royce, Josiah. *The Philosophy of Loyalty.* New York, 1924.

Sidgwick, Henry. *Methods of Ethics,* 254. 7th ed. London, 1922. For the social atomist view of the moral value of loyalty.

John Ladd (1967)

LU CHIU-YÜAN

See *Lu Xiangshan*

LUCIAN OF SAMOSATA

(c. 115–c. 200)

Lucian of Samosata, the philosophical satirist and satirist of philosophy, was born at Samosata (Samsat) on the Euphrates and was educated there. He then studied rhet-

oric in Asia Minor, after which he was a lawyer for a while, toured Greece and Italy as a lecturer, and held a chair of literature in France. In middle age he settled in Athens, where he wrote and gave public readings of his most successful dialogues, many of which were on philosophical themes. Late in life he joined the staff of the Roman governor of Egypt. Nothing is known of his death except that it occurred after 180.

Lucian's philosophical position is not easy to define because he expresses contradictory attitudes, and his persistent irony and his obvious wish to entertain make it hard to know how seriously to take his statements. The contradictions have been used as a basis for several different theories of his intellectual development, but the chronological order of his works is too uncertain for any such interpretation to be wholly convincing.

In *The Fisher*, Lucian claimed to be a champion of philosophy, which he described elsewhere as a civilizing and morally improving study; however, he constantly criticized pseudo philosophers for their greed, bad temper, sexual immorality, and the general inconsistency between their preaching and their practice. The historical occasion for such attacks was the encouragement of philosophy by Marcus Aurelius, which had made philosophers almost as numerous as monks and friars were in the Middle Ages.

Lucian's favorite target was the Stoic, but he also savagely attacked such Cynics as Peregrinus, and in *The Sale of Lives* he made fun of every school. However, he sometimes wrote approvingly of individual philosophies. The *Nigrinus* appears to be a eulogy of Platonism, although this may be ironical or simply an excuse for satirizing Roman society. The *Cynicus* is a less ambiguous defense of Cynicism, and in several dialogues Lucian speaks through a character called Cyniscus or through that of the Cynic Menippus. Diogenes is once mentioned favorably, and in the *Alexander* there is enthusiastic praise for Epicurus, "a really great man who perceived, as no one else has done, the beauty of truth."

The *Hermotimus* rejects all philosophical systems on the grounds that they are mutually contradictory and thus cannot all be right, and life is too short to discover which of them is nearest to the truth. The wisest course is to get on with the business of living, guided by common sense. Tiresias in the *Menippus* gives the same advice.

In general, Lucian disliked philosophies that encourage superstition, such as Platonism and Stoicism, and preferred materialists like Democritus and Epicurus. Although he made fun of the Skeptics, he was temperamentally inclined to skepticism, or to an eclecticism of the kind described in the *Life of Demonax*.

His own positive ideas included a conception of society free from racial, social, and economic distinctions. He valued such human qualities as sincerity, courage, cheerfulness, and kindness; and he continually stressed the importance of facing facts, especially the fact of death.

Lucian's influence on later thought was exerted largely, but not entirely, through the medium of literary technique. He facilitated the spread of humanism in the sixteenth century by suggesting one of the basic themes (the absurdity of plutocracy) and some of the incidental jokes in Thomas More's *Utopia*, but his main contributions were the lighthearted manner, the form (a fantastic journey described in a familiar dialogue), and the trick of using proper names that etymologically imply nonexistence or nonseriousness. He also aided in the Reformation by providing literary precedents and humorous devices for the satire on ecclesiastics, theologians, monks, and superstitions in Desiderius Erasmus's *Encomium Moriae* and in the work of François Rabelais. Voltaire's *Candide* is Lucianic in both manner and theme (the refutation of philosophical theory by reality), and its final moral is identical with that of the *Menippus*. The *Conversation between Lucian, Erasmus and Rabelais in the Elysian Fields* shows that Voltaire regarded Lucian as one of his masters in the strategy of intellectual revolution.

Bacon called Lucian a contemplative atheist, and as such Lucian evidently interested David Hume, who described him as a very moral writer, quoted him with respect when discussing ethics and religion, and read him on his deathbed. Since then, professional philosophers have tended to ignore him, but perhaps his spirit is still alive in those who (as Bertrand Russell did), are prepared to flavor philosophy with wit.

See also Cynics; Diogenes of Sinope; Epicurus; Erasmus, Desiderius; Humanism; Hume, David; Leucippus and Democritus; More, Thomas; Platonism and the Platonic Tradition; Rabelais, François; Russell, Bertrand Arthur William; Skepticism; Stoicism; Voltaire, François-Marie Arouet de.

Bibliography

TEXTS

Luciani opera. 4 vols, edited by M. D. Macleod. Oxford Classical Texts, 1972–1987.

Harmon, A. M. With facing Greek text. 8 vols., 1968–1979, Loeb Classical Library. Vol. 6 translated by K. Kilburn; vols. 7–8 translated by M. D. Macleod.

Macleod, Matthew D. "Lucianic Studies since 1930." *Aufstieg und Niedergang der Römischen Welt* 2.34.2 (1994): 1362–1421.

STUDIES

Anderson, Graham. *Lucian: Theme and Variation in the Second Sophistic*. Leiden: Brill, 1976.

Baldwin, Barry. *Studies in Lucian*. Toronto: Hakkert, 1973.

Bernays, J. *Lukian und die Kyniker*. Berlin, 1879.

Billault, Alain. *Lucien de Samosate*. Lyon: Centre d'Études Romaines et Gallo-Romaines 13, 1994.

Bompaire, J. *Lucien écrivain: Imitation et création*. Paris: Bibliothèque des Écoles Françaises d'Athènes et de Rome, 1958.

Branham, R. B. *Unruly Eloquence: Lucian and the Comedy of Traditions*. Cambridge, MA: Harvard University Press, 1989.

Camerotto, A. *Le metamorfosi della parola: studi sulla parodia in Luciano di Samosata*. Pisa: Istituti editoriali e poligrafici internazionali, 1998.

Caster, M. *Lucien et la pensée religieuse de son temps*. Paris: Société d'édition "Les belles lettres," 1937.

Clay, Diskin. "Lucian of Samosata, Four Philosophical Lives (Nigrinus, Demonax, Peregrinus, Alexander Pseudomantis)." *Aufstieg und Niedergang der Römischen Welt* 2.36.5: 3406–3450.

Edwards, Mark J. "Lucian and the Rhetoric of Philosophy: The Hermotimus." *Acta Classica* 62 (1993): 195–202.

Georgiadou, A., and D. H. J. Lamour. *Lucian's Science Fiction Novel True Histories: Interpretation and Commentary*. Leiden: Brill, 1998.

Helm, R. *Lucian und Menipp*. Leipzig: Teubner, 1906; reprinted, Hildesheim, 1967.

Jones, C. P. *Culture and Society in Lucian*. Cambridge, MA: Harvard University Press, 1986.

Peretti, A. *Luciano: Un intellettuale greco contro Roma*. Florence: Nuova Italia, 1946.

Robinson, Christopher. *Lucian and His Influence in Europe*. Chapel Hill: University of North Carolina Press, 1979.

Tackaberry, W. H. *Lucian's Relation to the Post-Aristotelian Philosophers*. Toronto, 1930.

Paul Turner (1967)
Bibliography updated by David Konstan (2005)

LUCRETIUS

(?–c. 55 BCE)

Little is known of Lucretius (d. ca. 55 BCE [Donatus, *Life of Virgil*] or perhaps a few years later; cf. Hutchinson 2001) apart from his poem in six books, *On the Nature of Things* (*De rerum natura*), an exposition in Latin hexameters of the doctrines of the Greek philosopher Epicurus, who lived two centuries earlier. Saint Jerome, in his *Chronicle* (Olympiad 171.3), claims that he committed suicide as a result of taking a love potion, and that he wrote his poem "in intervals of insanity," presumably meaning between, rather than during, such episodes. Jerome also asserts that Cicero "emended" Lucretius' text, that is, corrected it for publication, after his death (as Jerome gives it) in 51/50. It is possible that this is an inference from a letter of Cicero's to his brother (2.9, February 54 BCE), in which he praises Lucretius' poem, though Cicero himself had translated the Greek poet Aratus into Latin hexameters, and might well have taken an interest in a fellow poet's work.

Internal evidence reveals some repetitions and inconsistencies (e.g., the doublet at 4.45–53 and 4.26–44), which Lucretius would doubtless have eliminated in a final version; Lucretius also states that he will treat in greater detail the nature and habitation of the gods (5.155), but no such passage survives. Some scholars have supposed that he planned to include it in a seventh or even later book, and that accordingly the poem as we have it is radically incomplete; in particular, Lucretius did not intend to conclude with the depressing spectacle of the Athenian plague (summary of views in Boyancé 1963: 79–83). But there are good justifications for this ending, and Lucretius could have changed his mind about the theological section, or treated it briefly within the compass of the poem as we have it. In the proem to Book 6 (91–94) he indicates plainly that he is approaching the end of the poem.

The Pre-Socratic philosophers Parmenides and Empedocles had written treatises in verse, and Empedocles' poem, which Lucretius regarded highly enough to deem its author "godlike" (1.716–741), may have borne the same title (*Peri phuseôs*, or perhaps the even closer *Peri phuseôs tôn ontôn*: Sedley 1998: 21–22; the title may not have been Empedocles' own: Schmalzriedt 1970), and may have extended to several thousand lines (Diogenes Laertius 8.77). Empedocles' proem was likely a model for Lucretius' own (Gale 1994: 59–74; Sedley 1998: 1–34). Later, the medium for philosophy was decidedly prose, and Epicurus himself was suspicious of poetry (fr. 229 Usener; cf. Gale 1994: 14–18). In the Hellenistic period (third–first centuries BCE), didactic poetry was composed on a variety of topics, from astronomy and farming to poisonous snakes, but these genre pieces were not usually intended to provide serious instruction; Lucretius' poem was. He succeeds remarkably in conveying rigorous arguments concerning such matters as the constitution of the universe, which for the Epicureans was composed solely of atomic matter and empty space (Books 1–2), the materialist basis of perception and cognition (Books 3–4), and the evolution of the earth and of human civilization (Book 5), along with such special topics as the nature of magnetism (Book 6), even as he strug-

gles with the relative poverty of the Latin philosophical vocabulary, as opposed to Greek (1.136–39, 832, 3.260; cf. Cicero *De finibus* 3.51).

Given the mainly fragmentary or hostile character of our sources concerning Epicurus' doctrines (three short essays by Epicurus in the form of letters are reproduced by Diogenes Laertius, Book 10), Lucretius provides the single extended exposition of Epicurean physics that survives by a follower of the school. Doubtless, the medium of verse imposed some limitations, and Lucretius' understanding of certain points was perhaps faulty, but the poem is immensely valuable for the history of philosophy. It is also a magnificent work of literature, shot through with a moral passion that brightens even the most painstaking arguments about atoms and void.

SOURCES AND ORIGINALITY

This said, it is obviously important to determine what sources Lucretius himself employed, and over this question there is considerable controversy. It is in principle possible that Lucretius relied on no particular text but composed an independent poetical treatise based on his immersion in Epicureanism (Clay 1983: 31). David Sedley (1998), in turn, has argued forcefully that Lucretius adhered principally to a single treatise by Epicurus—*On Nature*—and was almost completely indifferent to or unaware of more recent currents in Epicureanism, or of ongoing debates with other schools, above all the Stoics (he dubs Lucretius a "fundamentalist" in this respect; cf. Furley 1967). Other scholars have seen clear indications of later influences in Lucretius' poem, for example in his attack on skepticism (Vander Waerdt 1989, Lévy 1997), his account of socio-political evolution (Schrijvers 1996 detects the influence of Polybius' theory of constitutions), and his arguments against teleology (Schmidt 1990: 152–160). Some have found it implausible that Lucretius should have been wholly isolated from the contemporary revival of Epicureanism in Italy, and have sought to demonstrate parallels between Lucretius' poem and the treatises of Philodemus (Kleve 1997), burned and buried in the eruption of Vesuvius in 79 CE, but still partly legible. Evidence that Lucretius' poem was among the scrolls in Philodemus' library remains inconclusive.

No one denies that Lucretius composed more freely in the proems with which he prefaced each of the six books, where, for example, he speaks of Venus as the ancestress of the Romans (1.1), and often too in the conclusions, or that he sometimes resorted to other sources than Epicurus (e.g., the description of the plague, based closely on Thucydides 2.47–54; cf. the analysis of passionate love at the end of Book 4, esp. vv. 1121–1191, and the personification of Nature scolding the man who fears death at 3.931–977, both indebted to Greek styles of diatribe [Wallach 1976, Reinhardt 2002]). So too, his choice of imagery in technical passages is frequently his own, for instance his illustration of the flow of thin membranes or simulacra from the surface of objects by reference to the colors cast on the audience by the awnings stretched above a Roman amphitheater (4.75–83). Some passages are more difficult to decide. When Lucretius explains the drive to accumulate wealth as a function of the fear of death, he says that poverty is imagined the "antechamber to hell" (3.65–69). Is this a Lucretian metaphor, or a piece of Epicurean doctrine? So too, Lucretius affirms that the legendary torments in the underworld, like Tantalus' perpetual hunger and the Danaids' task of carrying water in leaky pails, are really images of the forever frustrated pursuit of wealth and power in this world (3.978–1023). This may be a poetical flourish, but conceivably it reflects a genuine Epicurean explanation of the fear of punishment in the afterlife (Konstan 1973: 13–27).

Apart from such passages, in its broad outline Lucretius' poem conforms to the subjects that we know Epicurus treated in his principal statement of his views, above all his *On Nature* (*Peri phuseôs*), of which some substantial, though lacunose, fragments have been recovered on papyrus (see Sedley 1998: 133 for a possible reconstruction). To all appearances, Lucretius set about to versify a treatise on the atomic theory, and its implications for human psychology and society. He did not incorporate into his poem substantial arguments from Epicurus' ethical writings (for example, *On Lives* or *On the End*; cf. Diogenes Laertius 10.30). What is more, he shows no interest in many of the issues with which Philodemus was concerned, such as rhetoric, literary theory, virtues and vices, governance, semiotics, or the right methods for training disciples, which became central concerns of the school after the founder's death. Nor does he engage systematically and polemically with later opponents of Epicureanism, or with dissident views within the school, as Philodemus does (cf. the debates that Cicero stages between Epicureans, Stoics, and Academics); if indeed there are traces of such controversies in his poem, it is nonetheless remarkable that the philosophers whom he refutes explicitly and at length are Empedocles, Heraclitus, Anaxagoras, and Democritus: no mention of later thinkers. His poem purports to present classical Epicureanism in a palatable but accurate form to a Roman public—sweetening the spoon of medicine, in Lucretius' image (1.936–950, 4.11–25). He describes himself as

planting his feet in Epicurus' footprints (3.3–4, 5.55–56), and this seems a fair statement of his intentions.

This fidelity to Epicurus' major exposition of his doctrine need not be taken as a sign of intellectual narrowness or a quasi-religious commitment to the word of the Master (Sedley 1998: 93). It was the custom of Hellenistic didactic poets to take as their source a scientific treatise, as Aratus, for example, did in his *Phaenomena* or *Constellations*, where he followed Eudoxus' work of the same name (fr. 3a Lasserre), even as he modelled his style on that of the archaic poet Hesiod. The Roman poet Ennius, whom Lucretius praises extravagantly despite his mistaken belief in the underworld (1.117–126), did something similar when he rendered into prose the pseudo-scientific narrative of Euhemerus. Lucretius was writing as much in the sophisticated Alexandrian tradition as in that of the pre-Socratic poet-philosophers.

It is a separate question whether Lucretius sometimes altered Epicurus' order of presentation, and with this his chain of reasoning, and whether he added to or modified the arguments of the Master here and there, either independently or by mining other works of Epicurus or early Epicureans. He seems to claim some responsibility for the sequence in which he presents a series of proofs (1.52, 3.419–420; cf. Clay 1983: 38). Sedley (1998: 148–152) speculates that Lucretius planned a more extensive rearrangement of topics, but did not live to finish revising the entire poem. Lucretius may have been influenced also by the order of subjects in standard collections of doctrines, whether doxographies or rhetorical disquisitions (Runia 1997; on rhetoric, Classen 1986: 371).

LUCRETIUS AND EPICUREAN DOCTRINE

No doubt, Lucretius' vivid analogies and images are not without philosophical interest, though some will have had antecedents in Epicurus' works or elsewhere; for example, the proof of atomic motion from the visible vibration of dust motes in a sunbeam (2.114–141), comparable to Brownian motion, was evidently already proposed by Democritus (cf. Aristotle *De anima* 404a3–4). The image of a flock of sheep on a distant hillside (2.317–322), by which Lucretius illustrates how a compound may be seen as proceeding slowly although its constituent particles are moving rapidly, was likely Lucretius' own. Epicureanism tended, more than other ancient schools, to admit proof by analogy—a principal means of inferring the properties of the invisible atomic world from perceptible events—and this favored the probative value of similes (cf. 2.112–113). Isolating philosophically significant innovations in Lucretius, however, is a delicate task, given the scrappy condition of his principal source or sources (even where he composed freely rather than drawing on specific texts), and a novel comparison does not necessarily constitute a new argument.

In the circumstances, there are several ways to proceed. First, one may identify arguments in Lucretius that have no known parallel in Epicurus' own writings or those of later Epicureans; these at least are possible candidates for Lucretian innovation. Second, one may demonstrate Lucretius' dependence on some other, non-Epicurean source, e.g., Polybius or Thucydides, bearing in mind that Lucretius' references to early writers may have been filtered through Epicurus. Third, one may note specifically Roman or personal nuances of the sort that alter or affect in some measure orthodox Epicurean doctrine. Finally, one may discover places where Lucretius seems to disagree with what we know to have been Epicurus' view. The last is certainly the most dramatic, and indeed there is one apparent case of such a discrepancy: Lucretius speaks of four components of the soul (3.231–245)—air, ether, fire, and an unnamed, superfine element—whereas Epicurus, in the *Letter to Herodotus* (63), mentions just three, and in somewhat different terms. It is hardly likely that Lucretius is silently introducing here a modification of Epicurean doctrine. Conceivably, he was simply mistaken; alternatively, and more probable, Epicurus' account in the *Letter* is compressed, and he elaborated the fuller view in the relevant, now lost, passage in *On Nature* (Sedley 1998: 71n47).

Given the state of Epicurean texts, Lucretius is often our best guide to Epicurean doctrine, especially since there is not sufficient reason to suppose that his treatment is original. For example, Lucretius appeals to the so-called swerve of atoms, by which they shift by a minimal amount in their downward course at no determinate time or place (2.216–293), to account for free will and also for the initial interaction of atoms, which could not have collided had they maintained their natural downward motion at uniform speed. The latter argument seems particularly weak, since there is no beginning to the Epicurean universe, but it may nevertheless have been broached by Epicurus himself (cf. Fowler 2002: 301–309). Lucretius' account of the development of human civilization departs from parallel treatments known from other writers (Cole 1967), among other ways by inserting passages on the origin of religion and of language; again, this sequence may very well go back to Epicurus himself (Konstan 1973: 44–55; Campbell 2003: 15–18, 283–293). But Lucretius inclines to multiplying arguments—for

example, he offers 28 or 29 different proofs for the mortality of the soul (3.417–614)—and it is plausible that he may have added several to the common Epicurean stock, especially since only one or two are attested in Epicurus (Boyancé 1986: 141–142).

Epicurus discouraged active participation in politics because it produced the kinds of psychological tensions that his teachings were designed to eliminate. Lucretius, however, expresses a desire for peace (1.21–49) so that Memmius, the Roman aristocrat to whom he addresses his poem, will not have to engage in public service (perhaps an allusion to his praetorship in 58 BCE; Hutchinson (2001) sees a reference to the civil war that began in 49 BCE, and dates the poem to this period); in this way, Memmius will be free to dedicate himself to philosophy and achieve the tranquillity that Epicureanism held to be the goal of life. Epicureanism had a certain vogue among Roman nobles who had no intention of giving up their political status and activities—Julius Caesar himself is said to have been an adherent —and Lucretius was no doubt adapting his advice here to the outlook and social realities of his time (as did Philodemus). Whether this represents a change of principle or simply a tactical shift of rhetoric is difficult to say (cf. Fowler 1989).

Epicurus affirmed that sex should be avoided, since it has never done any good and is often harmful (Diogenes Laertius 10.118, fr. 62 Usener; *VS* 51); he also discouraged marriage (Diogenes Laertius 10.119 [textually corrupt], Epictetus *Discourses* 3.7.19–20, etc.). Lucretius' attitude toward love and sex is not inconsistent with Epicurus' own, though Epicurus' surviving writings are not so fervent on the subject, but he appears, at the end of Book IV (1278–1287), to introduce a newly positive view of matrimony and parenthood (Nussbaum 1994: 185–187; Brown 1987: 87–91, 118–122 sees no discrepancy here between Epicurus and Lucretius). Again, the fear of death, and of punishment in the afterlife, was the central cause of mental perturbation, according to Epicurus, and here too Lucretius is wholly in agreement; but his approach seems "more personal and emotional" than Epicurus' (Segal: 1990: 6; cf. 27–33, 51–54, 113; for the arguments, see Warren 2004); indeed, the Roman poet Statius spoke of the "burning passion of learned Lucretius" (*Silvae* 2.7.76). Further, Lucretius' anguished distress at the needless suffering of his fellow men (2.14–19) lends his poetry a proselytizing fervor, and this, like the shuddering pleasure (3.25–30) he experiences at the vision of a world without a hell, may seem to admit into Epicureanism a passion at odds with its goal of quietude.

Why did Lucretius end his poem with the grisly description of the plague that struck Athens in 429 BCE? Some scholars have supposed that it serves as a "final exam" in Lucretius' course on Epicureanism, testing whether readers have learned the lesson that death holds no terrors. This seems an adequate explanation (cf. Commager 1957, Bright 1971), even without an explicit moral to point the message. The plague is an accelerated image of life itself, which invariably terminates in death. Since Epicurus taught that pleasure does not increase with length of time (*Principal Doctrines* 18–20; Lucretius 3.944–945, 1080–1081), a life cut short by illness is not cause for apprehension.

Lucretius' poem immediately became famous: Virgil (*Georgics* 2.490–492) wrote, "Blessed is he who is able to know the causes of things," with obvious reference to Lucretius (*Ovid Amores* 1.15.23–24). Its rediscovery in the Renaissance inspired philosophical didactic poetry down through the eighteenth century, when the genre came to an end.

See also Empedocles; Epicurus.

Bibliography

TEXTS AND TRANSLATIONS

Bailey, Cyril. *Titi Lucreti Cari De Rerum Natura Libri Sex*. 3 vols. Oxford: Oxford University Press, 1947.

Humphries, Rolfe. *The Way Things Are: The De Rerum Natura of Titus Lucretius Carus*. Bloomington: University of Indiana Press, 1969.

Melville, Sir Ronald. *Lucretius, On the Nature of the Universe*. With introduction and notes by D. and P. Fowler. Oxford: Clarendon Press, 1997.

Smith, Martin Ferguson. *Lucretius: De Rerum Natura*. With English translation by W. H. D. Rouse, revised with new text, introduction, and notes. Cambridge, MA: Harvard University Press (Loeb Classical Library), 1975.

STUDIES

Algra, Keimpe A., Pieter W. van der Horst, and David T. Runia, eds. *Polyhistor: Studies in the History and Historiography of Ancient Philosophy*. Presented to Jaap Mansfeld on his Sixtieth Birthday. Leiden: E. J. Brill, Philosophia Antiqua 72, 1996.

Algra, K.A., M.H. Koenen and P. H. Schrijvers, eds. *Lucretius and his Intellectual Background*. Amsterdam: Royal Netherlands Academy of Arts and Sciences, 1997.

Boyancé, Pierre. "La théorie de l'âme chez Lucrèce." In *Probleme der Lukrezforschung*, edited by Carl Joachim Classen, 131–150. Hildesheim: Olms, 1986 (orig. 1958).

Boyancé, Pierre. *Lucrèce et L'Épicurisme*. Paris: Presses Universitaires de France, 1963.

Bright, David F. "The Plague and the Structure of the *De rerum natura*." *Latomus* 30 (1971): 607–632.

Brown, Robert D., ed. *Lucretius on Love and Sex. A Commentary on De Rerum Natura IV, 1030–1287*. Leiden: E. J. Brill, 1987.
Campbell, Gordon. *Lucretius on Creation and Evolution: A Commentary on De Rerum Natura Book Five Lines 772–1104*. Oxford: Oxford University Press, 2003.
Classen, Carl Joachim. "Poetry and Rhetoric in Lucretius." In *Probleme der Lukrezforschung*, 331–373. Hildesheim: Olms, 1986.
Classen, Carl Joachim, ed. *Probleme der Lukrezforschung*. Hildesheim: Olms, 1986.
Clay, Diskin. *Lucretius and Epicurus*. Ithaca NY: Cornell University Press, 1983.
Cole, Thomas. *Democritus and the Sources of Greek Anthropology*. Cleveland: Press of Western Reserve University, 1967.
Commager, Henry Steele, Jr. "Lucretius' Interpretation of the Plague." *Harvard Studies in Classical Philology* 62 (1957): 105–121.
Fowler, Don P. "Lucretius and Politics." In *Philosophia Togata: Essays on Philosophy and Roman Society*, edited by Miriam Griffin and Jonathan Barnes, 120–150. Oxford: Clarendon Press, 1989.
Fowler, Don. *Lucretius on Atomic Motion. A Commentary on De Rerum Natura 2.1–332*. Oxford: Oxford University Press, 2002.
Furley, David J. *Two Studies in the Greek Atomists*. Princeton, NJ: Princeton University Press, 1967.
Gale, Monica. *Myth and Poetry in Lucretius*. Cambridge, U.K.: Cambridge University Press, 1994.
Hutchinson, G. O. "The Date of *De rerum natura*." *Classical Quarterly* 51 (2001): 150–162.
Kleve, Knut. "Lucretius and Philodemus." In *Lucretius and his Intellectual Background*, edited by Algra et al., 49–66. Amsterdam: Royal Netherlands Academy of Arts and Sciences, 1997.
Konstan, David. *Some Aspects of Epicurean Psychology*. Leiden: E. J. Brill, Philosophia Antiqua 25, 1973.
Lévy, Carlos. "Lucrèce avait–il lu Enésideme?" In *Lucretius and his Intellectual Background*, edited by Algra et al. 115–124. Amsterdam: Royal Netherlands Academy of Arts and Sciences, 1997.
Nussbaum, Martha C. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press, 1994.
Reinhardt, Tobias. 2002. "The Speech of Nature in Lucretius' De Rerum Natura 3.931–71." *Classical Quarterly* 52: 291–304.
Runia, David T. "Lucretius and Doxography." In *Lucretius and his Intellectual Background*, edited by Algra et al., 93–103. Amsterdam: Royal Netherlands Academy of Arts and Sciences, 1997.
Schmalzriedt, Egidius. *Peri physeos: Zur frühgeschichte der Buchtitel*. Munich: Fink Verlag, 1970.
Schmidt, Jürgen. *Lucrez, der Kepos, und die Stoiker*. Frankfurt am Main: Peter Lang, Studien zur klassischen Philologie 53, 1990.
Schrijvers, Piet H. "Lucretius on the Origin and Development of Political Life (De Rerum Natura 5.1105–1160)." In *Polyhistor: Studies in the History and Historiography of Ancient Philosophy*, 220–230. Presented to Jaap Mansfeld on his sixtieth birthday. Leiden: E. J. Brill, Philosophia Antiqua 72, 1996.
Schrijvers, Piet H. *Lucrèce et les sciences de la vie*. Leiden: Brill, Mnemosyne Supplement 186, 1999.
Sedley, David. *Lucretius and the Transformation of Greek Wisdom*. Cambridge, U.K.: Cambridge University Press, 1998.
Segal, Charles. *Lucretius on Death and Anxiety*. Princeton, NJ: Princeton University Press, 1990.
Vander Waerdt. "Colotes and the Epicurean Refutation of Skepticism." *Greek Roman and Byzantine Studies* 30 (1989): 225–267.
Wallach, Barbara Price. *Lucretius and the Diatribe against the Fear of Death*. Leiden: Brill, Mnemosyne Supplement 40, 1976.
Warren, James. *Facing Death: Epicurus and His Critics*. Oxford: Oxford University Press, 2004.

David Konstan (2005)

LU HSIANG-SHAN

See *Lu Xiangshan*

LUKÁCS, GEORG

(1885–1971)

Georg (György) Lukács, the Hungarian Marxist philosopher and literary critic, was professor of aesthetics and the philosophy of culture at the University of Budapest from 1945 to 1956. Lukács was born in Budapest into a rich and eminent family (before he became a communist he wrote under the family name "von Lukács"). He took a doctorate in philosophy in Budapest (1906) and then studied under Georg Simmel at Berlin and under Max Weber at Heidelberg. Since Lukács was recognized as one of Europe's leading literary critics when he joined the Communist Party of Hungary in December 1918, he was offered the post of people's commissar for culture and education in the communist regime of Béla Kun (March–August 1919). After the fall of Kun, Lukács took refuge in Vienna, where he edited the review *Kommunismus* and carried on a struggle with Kun (exiled in Moscow) for control of the Hungarian underground movement. Publication in Berlin in 1923 of Lukács's collection of essays, *Geschichte und Klassenbewusstsein*, decided the issue in favor of Kun—for the book was denounced as "deviationist." Lukács was ousted from the central committee of the Communist Party and from the editorship of *Kommunismus* after publishing his "self-criticism." He took refuge in Russia when Adolf Hitler came to power and, after a further and more thorough act

of self-criticism, worked in the Institute of Philosophy of the Soviet Academy of Science from 1933 to 1944. Returning to Hungary, he became a member of parliament and professor of aesthetics. In 1956 Lukács was a leader of the Petofi circle, which played a role in the anti-Russian insurrection, and then minister for culture in the short-lived Imre Nagy government. After the defeat of the revolution, Lukács was deported to Romania, but he was allowed to return to Budapest in April 1957 to live in retirement and to devote himself to a monumental work on aesthetics, of which one volume was published, in Hungarian.

AESTHETICS AND CRITICISM

Lukács's fame as one of the few philosophers produced by the Marxist movement rests on a book that he repudiated soon after its publication, *Geschichte und Klassenbewusstsein* (History and class consciousness). His later work—some thirty books and hundreds of articles—constitutes an attempt to found a Marxist aesthetic that could be used to criticize modernist, formalist, and experimental art in the name of socialist realism. This critical work entailed some confusion of literary criticism with political polemic, of which the following judgment on Kafka is typical: "no work of art based on *Angst* (anxiety) can avoid—objectively speaking—guilt by association with Hitlerism and the preparations for atomic war" (*The Meaning of Contemporary Realism*, p. 81). Lukács's influence as a critic has been intensely conservative, for he held that "realism is not one style among others; it is the basis of literature" (p. 48).

In his first aesthetic studies, *Die Seele und die Formen* (The soul and the forms) and *Die Theorie des Romans* (The theory of the novel), Lukács was still a neo-Kantian. He held that literature was the striving for expression of the irrational soul in and through an alien and hostile reality. He stressed the value of "inwardness" and the uselessness of society to the individual. These works have been claimed as among the sources of existentialism, but Lukács himself denounced them as "false and reactionary" upon his conversion to communism. Thereafter he contrasted Marxism, as a philosophy that integrated the individual in society, with all modern "philosophies of crisis and evasion," and in particular with existentialism, which isolated men outside social and economic relations.

Lukács's stress on social relationships became the basis of his aesthetics. Form, he argued, should be determined by content (therefore abstract art and formalism are degenerate), and "there is no content of which Man himself is not the focal point" (*The Meaning of Contemporary Realism*, p. 19). Since man exists only in a social and historical context, aesthetics inevitably is concerned with politics. If the subject of a work of art is man seen statically, then that work declines into subjectivism and allegory. Literature must be dynamic, setting characters in historical perspective in order that they might be shown as having direction, development, and motivation. For literature to be dynamic, the major historical movement of the day must be taken into account. In the twentieth century that movement was socialism. The only valid contemporary literary styles are socialist realism, which is practiced inside the socialist movement, and critical realism, which is practiced by authors sympathetic to socialism. Lukács's theories naturally entailed condemnation of most twentieth-century art, literature, and music, but they were fruitfully applied to the historical novel.

SOCIAL AND HISTORICAL ANALYSIS

Geschichte und Klassen-bewusstsein, the censored masterpiece of communist thought, became the classic text of Western Marxism as contrasted with Soviet orthodoxy. It led to a revaluation of Marxism by setting it in a Hegelian context. Lukács was the first to see that Karl Marx's theory of history and even his economics could be read as an application of the Hegelian dialectic. He did this a decade before the discovery and publication of Marx's *Economic and Philosophic Manuscripts of 1844*, which amply confirmed his theory, at least with regard to the young Marx. Having meanwhile disowned his book, Lukács could not claim credit for that brilliant piece of philosophical reconstruction, but he later could show the profound similarity between the philosophies of G. W. F. Hegel and Marx (*Der junge Hegel*). His idealist reading of Marx clashed with the accepted Leninist version, and, since Lukács worsened his case in 1923 by revealing the influence of Georges Sorel and Rosa Luxemburg on his thought, his book was condemned with a ferocity unusual even in communist polemics.

Lukács had rejected Friedrich Engels's and V. I. Lenin's conception of the Marxist dialectic as a set of laws applying to nature, and he rejected too the notion that historical materialism deduces all social and moral life from the economic base. Historical materialism and the dialectic, he said, both mean the same thing, namely that in society subject and object are one. When men know (or enter into any other relation with) social entities—whether these are institutions or economic goods or another age's culture—the relation established is not the sort of relation they have with the natural objects studied by physical science. Social

entities are reified personality or alienated spirit, while men themselves are the product of historical forces. The knower and the known, subject and object, are moments of one entity, society, and their relations are necessarily ambiguous, two-way, or dialectical.

Marx had said, "As personal interests become autonomous in the shape of class interests, the personal conduct of the individual becomes reified and alienated and thereby becomes a thing apart from him, an independent force." It is just such alienated forms of conduct that make up society. In the nineteenth century in particular, because of the development of industry, "material forces were saturated with spiritual life, while human existence was made animal, became a material force." Marx meant, said Lukács, that spirit had become thing and things were steeped in spirit, so that history was a fabric of meanings-become-forces. This dialectical relation of subject and object was most marked in the case of the proletariat because the proletariat had been reduced by capitalism to labor, a mere economic commodity, and yet it could still take cognizance of itself as a commodity by acquiring class consciousness. Thereupon, it saw through the supposed natural laws of economics and revolutionized capitalism. "For this class, self-knowledge means at the same time correct knowledge of the whole of society … so this class is at once subject and object of knowledge" (*Geschichte und Klassenbewusstsein*). Its self-knowledge is history knowing itself, and in that total clarity lies the promise of a return from alienation.

The difficulties raised by historical relativism—difficulties that had been seen by all who asked how Marxism alone among social opinions could escape being vitiated by its relation to a given class and age—can be resolved only by going right to the extreme of relativism. That is to say, historical materialism must be applied to itself until it is seen as relative and provisional. This means abandoning the notion of absolute truth and denying the complete opposition of true and false. History is a dialectical totality of knowers and things known, and every piece of culture, no matter how deformed by class position and historical situation, reflects that totality. Truth exists, but it exists only in the future tense; it is the presumptive totality to be attained by permanent self-criticism. "The criterion of truth is grasp of reality. But reality is not at all to be confounded with empirical being, what actually exists. Reality *is* not; it becomes—and not without the collaboration of thought" (*Geschichte und Klassenbewusstsein*). Rejecting the representative theory of knowledge made orthodox for Marxists by the examples of Engels and Lenin (the "concepts in our heads" are "true images of reality"), Lukács held that truth is not something to be reflected but something to be made by us by collaborating with what is new and progressive in historical forces. The vague notion of a moving totality of things, of the whole of history, is essential to this "relativization of relativism." Lukács did not clearly delineate this notion, but it evidently bears a resemblance to the Hegelian Absolute.

Lukács's three main doctrines—the dialectical unity of subject and object in society; the promise of a return from alienation when society, through the proletariat, attains self-knowledge; and the notion of truth as a totality yet to be achieved—were attractive to some Western existentialists. Lukács complained that their "treacherous" use of his work was a "falsification of a book forgotten for good reason." Another line of influence was through his former associate Karl Mannheim, who developed the relativization of all ideologies into the sociology of knowledge. Within the communist world, the only doctrine of Lukács's censored book to enjoy some surreptitious authority was his "proof" of the communist intellectual's duty to accept the Communist Party as the supreme expression of proletarian class consciousness and thus as endowed with the correct view of history. This doctrine Lukács himself practiced rigorously, even to the extent of repudiating his own major contribution to modern thought.

See also Aesthetics, History of; Communism; Critical Realism; Critical Theory; Engels, Friedrich; Existentialism; Hegel, Georg Wilhelm Friedrich; Historical Materialism; Kafka, Franz; Lenin, Vladimir Il'ich; Mannheim, Karl; Marxist Philosophy; Marx, Karl; Neo-Kantianism; Simmel, Georg; Socialism; Sorel, Georges; Weber, Max.

Bibliography

WORKS BY LUKÁCS

Philosophy

Die Seele und die Formen. Berlin: Fleischel, 1911.

Die Theorie des Romans. Berlin: Cassirer, 1920.

Geschichte und Klassenbewusstsein. Berlin: Malik, 1923.

Der junge Hegel. Zürich: Europa, 1948; rev. ed., Berlin, 1954.

Existentialismus oder Marxismus? Berlin: Aufbau, 1951.

Beiträge zur Geschichte der Aesthetik. Berlin: Aufbau, 1954.

Die Zerstörung der Vernunft. Berlin: Aufbau, 1955.

Literary Criticism

Essays über Realismus. Berlin: Aufbau, 1948. Translated by E. Bone as *Studies in European Realism*. London: Hillway, 1950.

Thomas Mann. Berlin, 1949. Translated by S. Mitchell as *Essays on Thomas Mann*. London: Merlin Press, 1964.

The Historical Novel. Translated by H. Mitchell and S. Mitchell. London: Merlin Press, 1962.

The Meaning of Contemporary Realism. Translated by J. Mander and N. Mander. London: Merlin Press, 1963.

WORKS ON LUKÁCS

Arato, A., and P. Breines. *The Young Lukács and the Origin of Western Marxism*. New York: Seabury Press, 1979.

Carbonara, C. *L'estetica del particolare di G. Lukacs*. Naples, 1960.

Goldmann, Lucien. *Le dieu caché*. Paris: Gallimard, 1955. Translated by P. Thody as *The Hidden God*. New York: Humanities Press, 1964.

Goldmann, Lucien. "Introduction aux premiers écrits de Georges Lukacs." In Lukács's *La théorie du roman*. Geneva, 1963.

Heller, A., ed. *Lukács Revisited*. Oxford: Blackwell, 1983.

Jung, W. *Georg Lukács*. Stuttgart: Metzler, 1989.

Kadarkay, A. *Georg Lukács Life, Thought, and Politics*. Oxford: Blackwell, 1991.

Lukács, Georg et al. *Georg Lukács: Zum Siebzigsten Geburtstag*. Berlin: Aufbau-Verlag, 1955.

Merleau-Ponty, Maurice. *Les aventures de la dialectique*. Paris: Gallimard, 1955. Translated as *Adventures of the Dialectic*. Evanston, IL: Northwestern University Press, 1972.

Mészáros, L. *Lukács' Concept of the Dialectic*. London: Merlin, 1972.

Watnick, Morris. "Relativism and Class Consciousness: Georg Lukács." In *Revisionism*, edited by Leopold Labedz. New York: Praeger, 1962.

Zitta, Victor. *Georg Lukács's Marxism*. The Hague: Nijhoff, 1964.

Neil McInnes (1967)
Bibliography updated by Thomas Nenon (2005)

ŁUKASIEWICZ, JAN
(1878–1956)

Jan Łukasiewicz, the Polish philosopher and logician, was born in Lvov. After studying mathematics and philosophy at the University of Lvov he was graduated in 1902 with a PhD in philosophy. Łukasiewicz taught philosophy and logic first at Lvov and from 1915 at the University of Warsaw. In 1918 he interrupted academic work to accept a senior appointment in the Polish ministry of education in Ignacy Paderewski's cabinet. At the end of that year, however, he returned to the university and continued as professor of philosophy until September 1939. During that period he served twice as rector of the university (1922/1923 and 1931/1932). Toward the end of World War II Łukasiewicz left Warsaw. After some time in Münster and then in Brussels, in 1946 he accepted an invitation from the Irish government to go to Dublin as professor of mathematical logic at the Royal Irish Academy, an appointment that he held until his death. Łukasiewicz held honorary degrees from the University of Münster and from Trinity College, Dublin. He was a member of the Polish Academy of Sciences in Kraków, the Society of Arts and Sciences in Lvov, and the Society of Arts and Sciences in Warsaw.

EARLY WRITINGS

Łukasiewicz studied under Kazimierz Twardowski, who was occupied with conceptual analysis. The rigorous, clear thinking Twardowski advocated is easily recognizable in the first major essays published by Łukasiewicz. Of these works, *O zasadzie sprzeczności u Arystotelesa* (On the principle of contradiction in Aristotle; Kraków, 1910) was one of the most influential books in the early period of the twentieth-century logical and philosophical revival in Poland. It must have stood high in the author's own estimation, for in 1955 he began translating it into English. The main point of the book is that in Aristotle's work one can distinguish three forms of the principle of contradiction: ontological, logical, and psychological. The ontological principle of contradiction is that the same property cannot both belong and not belong to the same object in the same respect. The logical principle says that two contradictory propositions cannot both be true, and the psychological principle of contradiction holds that no one can, at the same time, entertain two beliefs to which there correspond two contradictory propositions. Łukasiewicz supported his findings with quotations from the writings of Aristotle and then examined the validity of Aristotle's argumentation. One chapter brought to the notice of Polish readers Bertrand Russell's antinomy concerning the class of all classes that are not members of themselves. The appendix contains an elementary exposition of the algebra of logic, as well as an original and interesting methodological classification of the ways of reasoning, a problem with which at least two of Łukasiewicz's early papers were concerned.

Łukasiewicz's writings published before 1918 suggest that until that time he was in quest of topics to which he could devote all his intellectual resources. He found such topics in the logic of propositions and in the logic of the ancient Greeks. From 1918 onward, deviations from this double line of research are few and of little significance.

LOGIC OF PROPOSITIONS

MANY-VALUED LOGICS. The first and perhaps most important result obtained by Łukasiewicz in the logic of propositions was his discovery of three-valued logic in 1917. Our ordinary logic of propositions is two-valued, presupposing only two logical values, truth and falsity,

and it tacitly adheres to the principle of bivalence, that a propositional function holds of any propositional argument if it holds of the constant true proposition (usually symbolized by 1) and if it holds of the constant false proposition (represented by 2). If we use δ as a functorial variable that, when followed by a propositional argument, forms a propositional expression, then we can express the principle of bivalence by saying "if $\delta 1$ then if $\delta 2$ then δp," where p is a propositional variable. The meaning of the logical constants forming such expressions as, for instance, *Cpq* ("if *p* then *q*"), *Kpq* ("*p* and *q*"), *Apq* ("*p* or *q*"), and *Np* ("it is not the case that *p*") are, in two-valued logic, conveniently and adequately determined by means of the familiar two-valued truth tables:

$$C11 = C21 = C22 = 1$$
$$C12 = 2$$
$$K11 = 1$$
$$K12 = K21 = K22 = 2$$
$$A11 = A12 = A21 = 1$$
$$A22 = 2$$
$$N1 = 2$$
$$N2 = 1$$

In three-valued logic the principle of bivalence does not hold. It is replaced by the principle of trivalence, which presupposes three logical values: the constant true proposition represented by 1, the constant false proposition by 3, and the constant "possible" proposition by 2. The principle then says "if $\delta 1$ then if $\delta 2$ then if $\delta 3$ then δp." As a consequence the meanings of implication, conjunction, alternation, and negation have to be readjusted, and the following three-valued truth tables suggest themselves for the purpose:

$$C11 = C21 = C22 = C31 = C32 = C33 = 1$$
$$C12 = C23 = 2$$
$$C13 = 3$$
$$K11 = 1$$
$$K12 = K21 = K22 = 2$$
$$K13 = K23 = K31 = K32 = K33 = 3$$
$$A11 = A12 = A13 = A21 = A31 = 1$$
$$A22 = A23 = A32 = 2$$
$$A33 = 3$$
$$N1 = 3$$
$$N2 = 2$$
$$N3 = 1$$

In this logic alternation and conjunction can be defined as follows: *Apq* = *CCpqq*, and *Kpq* = *NANpNq*. All expressions involving only *C* and *N* and verified by the new truth tables can be constructed into a deductive system based on the axioms *CpCqp, CCpqCCqrCpr, CCCpNppp,* and *CCNpNqCqp*. This was shown by Mordchaj Wajsberg, who had studied logic under Łukasiewicz in Warsaw. Wajsberg's system, however, does not enable us to define all the functors available in three-valued logic. In particular the functor *T,* whose truth table says that $T1 = T2 = T3 = 2$, cannot be defined in terms of *C* and *N*. Jerzy Słupecki, who had also been a pupil of Łukasiewicz, subsequently proved that by adding *CTpNTp* and *CNTpTp* to Wajsberg's axioms we get a functionally complete system of three-valued logic, in which any functor can be defined.

The conception of three-valued logic was suggested to Łukasiewicz by certain passages in Aristotle. Purely formal considerations, such as those that led E. L. Post to comparable results, played a subordinate role in Łukasiewicz's thinking. By setting up a system of three-valued logic Łukasiewicz hoped to accommodate the traditional laws of modal logic. He also hoped to overcome philosophical determinism, which he believed was entailed by the acceptance of the bivalence principle and which he had always found repulsive. Interestingly enough, he modified his views in the course of time and saw no incompatibility between indeterminism and two-valued logic.

Once a system of three-valued logic had been constructed, the possibility of four-valued, five-valued, …, *n*-valued, and, finally, infinitely many-valued logics was obvious. At one time Łukasiewicz believed that the three-valued and the infinitely many-valued logics were of greater philosophical interest than any other many-valued logic, for they appeared to be the least arbitrary. In the end, however, he interpreted Aristotelian modal logic within the framework of a four-valued system.

The philosophical significance of the discovery of many-valued logic can be viewed in the following way: The laws of logic had long enjoyed a privileged status in comparison with the laws propounded by natural sciences. They had been variously described as a priori or analytic, the purpose of such descriptions being to point out that the laws of logic were not related to reality in the same way as were the laws of natural sciences, which had often been corrected or discarded in the light of new observations and experiments. The laws of logic appeared unchallengeable. By discovering many-valued logics Łukasiewicz showed that even at the highest level of generality—within the field of propositional logic—alternatives were possible. By adhering to the principle of bivalence or any other *n*-valence principle we run the same risk of misrepresenting reality that the scientist does when he offers any of his generalizations.

THE CLASSICAL PROPOSITIONAL LOGIC. Although Łukasiewicz contemplated the possibility that a nonclassical logic of propositions applied to reality, he made the classical propositional logic the principal subject of his research. He showed that the axiom systems of the calculus of propositions proposed by Frege, Russell, and Hilbert each contained a different redundant axiom. He proved that all the theses of the *CN*-calculus could be derived from the three mutually independent axioms *CCNppp, CpCNpq,* and *CCpqCCqrCpr.* He solved the problem of the shortest single axiom for the *E*-calculus and the *C*-calculus by showing that the *E*-calculus, whose only functor means "if and only if," with $E11 = E22 = 1$ and $E12 = E21 = 2$ as its truth table, could be based on any of *EEpqEErqEpr, EEpqEEprErq,* and *EEpqEErpEqr* and on no shorter thesis and by proving that *CCCpqrCCrpCsp* is the shortest thesis strong enough to yield the *C*-calculus. The first single axiom for *CN*-calculus, consisting of 53 letters, was discovered by Alfred Tarski in 1925. It was soon followed by a series of successive simplifications devised by Łukasiewicz and by Bolesław Sobociński. The latest in this series is a 21-letter axiom, *CCCCCpqCNrNsrtCCtpCsp,* discovered by C. A. Meredith, Łukasiewicz's Irish colleague. It is likely to prove to be the shortest possible axiom for the *CN*-calculus.

CONSISTENCY, COMPLETENESS, AND INDEPENDENCE. The metalogical study of deductive systems of the logic of propositions includes the study of consistency and completeness, and, in the case of systems based on several axioms, the mutual independence of the axioms has also to be considered. Independently of Post, Łukasiewicz developed both a method of proving consistency and one of proving the completeness of systems of the calculus of propositions. The completeness proof was based on the idea that if the system under consideration is not complete, there must be independent propositions, that is, propositions not derivable from the axioms of the system which on being adjoined to the axioms lead to no contradiction. If there are independent propositions, then there must be a shortest one among them. Following Łukasiewicz's method, one tries to show that any proposition that is meaningful within the system either is derivable from the axioms or is longer than another proposition inferentially equivalent to it. This method dispenses with the concept of "normal expressions" and is very useful for proving weak completeness of partial systems. Mutual independence of theses is usually established by an appropriate reinterpretation of the constant terms occurring in them. Many such reinterpretations have been provided by Łukasiewicz's many-valued logics. The wealth of metalogical concepts and theorems worked out in Łukasiewicz's logical seminar in Warsaw by Łukasiewicz himself, Tarski, Adolf Lindenbaum, Sobociński, and Wajsberg can best be seen in "Untersuchungen über den Aussagenkalkül," which summarizes the results obtained there between 1920 and 1930.

FUNCTORIAL CALCULUS. In Dublin, Łukasiewicz became interested in a two-valued calculus of propositions involving functorial variables. Since he used only functorial variables requiring one propositional argument to form a propositional expression, his new calculus was only a part of what Stanisław Leśniewski had called protothetic. A very strong rule of substitution invented by Łukasiewicz, together with the usual substitution rules for propositional variables, allows us, for instance, to use a thesis of the form $\delta\alpha$ to infer not only $N\alpha$ but also such theses as $Cp\alpha$, $C\alpha p$, $C\alpha CN\alpha p$, $C\alpha\alpha$, and α. By means of the new rule Łukasiewicz was able to base the calculus on the single axiom $C\delta C22C\delta 2\delta p$. This axiom is identical with the principle of bivalence, because $C22 = 1$. Meredith succeeded in showing that Łukasiewicz's axiom could be replaced by $C\delta\delta 2\delta p$ or by $C\delta pC\delta Np\delta q$. He was also able to prove completeness of the system.

ANCIENT LOGIC

Concurrently with his investigations of the logic of propositions Łukasiewicz was engaged in a thorough reappraisal of ancient logic. For centuries the logic of the Stoics had been regarded as a sort of appendage to the Aristotelian syllogistic. Łukasiewicz was the first to recognize in it a rudimentary logic of propositions. He found evidence that the main logical functions, such as implication, conjunction, exclusive disjunction, and negation, were known to the Stoics, who, following Philo of Megara, interpreted them as truth-functions, just as we do now. He pointed out that the Stoics, unlike Aristotle, had given their logic the form of schemata of valid inferences. Some of these schemata had been accepted axiomatically and others were rigorously derived from them. He subjected to severe but justified criticism the treatments of Stoic logic by such authorities as Carl Prantl, Eduard Zeller, and Victor Brochard. His preliminary investigations of medieval logic showed beyond doubt that in this field too there was room for fruitful research.

Equally successful was Łukasiewicz's inquiry into Aristotle's syllogistic. No sooner had he mastered the elements of symbolic logic for himself than he realized that the centuries-old traditional treatment of the Aristotelian

syllogistic called for revision. A new presentation of the logic of Aristotle was before long included in his regular lectures at the university and then published in *Elementy logiki matematycznej* (Elements of mathematical logic; Warsaw, 1929). Łukasiewicz completed a detailed monograph on the subject in Polish in the summer of 1939, but the manuscript and all printed copies were lost during the war. *Aristotle's Syllogistic* (1951) is a painstaking reconstruction undertaken by Łukasiewicz on his arrival in Dublin. The monograph can rightly be called revolutionary. In it Łukasiewicz argued that Aristotelian syllogisms are logical laws rather than schemata of valid inferences, as is taught in traditional textbooks. He put in historical perspective Aristotle's introduction of variables and, referring to a forgotten Greek scholium, gave a plausible explanation of the problem of the so-called Galenian figure. Among more formal results, we owe to Łukasiewicz the first modern axiomatization of syllogistic. The system he set up, based on the axioms *Aaa* ("every *a* is *a*"), *Iaa* ("*some a* is *a*"), *CKAbcAabAac,* and *CKAbcIbalac,* seems to be in perfect harmony with Aristotle's own treatment of the subject in the *Analytica Priora.* The axioms are jointly consistent and mutually independent. Moreover, Słupecki has ingeniously solved the decision problem for the system.

MODAL LOGIC

During the last few years of his life Łukasiewicz devoted much attention to modal logic. The results are presented in "A System of Modal Logic," and in the second edition of *Aristotle's Syllogistic* (1957) they serve as the basis for a critical examination of Aristotle's theory of modalities. Łukasiewicz's principal idea is that of "basic modal logic," obtained by adding to the classical calculus of propositions the axioms *CpMp* and *EMpMNNp* and by axiomatically rejecting *CMpp* and *Mp.* In these formulas *Mp* stands for "it is possible that *p.*" According to Łukasiewicz any modal system must contain basic modal logic as a part. This condition is fulfilled by the four-valued modal system based on *CδpCδNpδq* and *CpMp* as the only axioms, with *CMpp* and *Mp* axiomatically rejected.

The logical symbolism used in this entry was worked out by Łukasiewicz in the early 1920s. It requires no punctuation signs, such as brackets or dots, which from the point of view of metalogical investigations is its greatest merit. At the same time Łukasiewicz worked out a simple and perspicuous method of setting out proofs in the logic of propositions and in syllogistic. Both his symbolism and his proof technique have been adopted by many logicians outside Poland.

Łukasiewicz was not only a resourceful and imaginative scholar but also a gifted and inspiring teacher. He was one of the founders, and the life and soul, of the Warsaw school of logic. Tarski, Lindenbaum, Stanisław Jaśkowski, Wajsberg, Father Jan Salamucha, Sobociński, Słupecki, and Meredith have been his most outstanding pupils or collaborators.

See also Aristotle; Frege, Gottlob; Hilbert, David; Logic, History of; Modal Logic; Philo of Megara; Propositions; Russell, Bertrand Arthur William; Tarski, Alfred; Truth; Twardowski, Kazimierz.

Bibliography

PRINCIPAL WORKS BY ŁUKASIEWICZ

"Analiza i konstrukcja projecia przyczyny" (Analysis and construction of the concept of cause). *Przegląd Filozoficzny* 9 (1906): 105–179.

O zasadzie sprzeczności u Arystotelesa. Kraków, 1910.

Die logischen Grundlagen der Wahrscheinlichkeitsrechnung. Kraków: Akademie der Wissenschaften, 1913.

"Opojeciu wielkości" (On the concept of magnitude). *Przeglad Filozoficzny* 19 (1916): 1–70.

"O logice trójwartościowej" (On three-valued logic). *Ruch Filozoficzny* 5 (1920): 170–171.

"Logika dwuwartościowa" (Two-valued logic). *Przegląd Filozoficzny* 23 (1921): 189–205.

"Démonstration de la compatibilité des axiomes de la théorie de la déduction." *Annales de la Société Polonaise de Mathématique* 3 (1925): 149.

Elementy logiki matematycznej. Warsaw, 1929; 2nd ed., Warsaw, 1958. Translated by Olgierd Wojtasiewicz as *Elements of Mathematical Logic.* New York, 1963.

"Untersuchungen über den Aussagenkalkül." *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie,* Classe III 23 (1930): 30–50. In collaboration with Alfred Tarski. English translation by J. H. Woodger in *Logic, Semantics, Metamathematics,* by Alfred Tarski, 38–59. Oxford: Clarendon Press, 1956.

"Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagenkalküls." *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie,* Classe III 23 (1930): 51–77.

"Ein Vollständigkeitsbeweis des zweiwertigen Aussagenkalküls." *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie,* Classe III 24 (1931): 153–183.

"Zur Geschichte des Aussagenlogik." *Erkenntnis* 5 (1935–1936): 111–131.

"Der Äquivalenzenkalkül." *Collectanea Logica* 1 (1939): 145–169.

"Die Logik und das Grundlagenproblem." In *Les entretiens de Zurich sur les fondements et la méthode des sciences mathématiques 6–9 Décembre 1938,* 82–100. Zürich, 1941.

"The Shortest Axiom of the Implicational Calculus of Propositions." *Proceedings of the Royal Irish Academy* 52, Section A (1948): 25–33.

"On Variable Functors of Propositional Arguments." *Proceedings of the Royal Irish Academy* 54, Section A (1951): 25–35.

Aristotle's Syllogistic. Oxford, 1951; 2nd ed., Oxford: Clarendon Press, 1957.

"A System of Modal Logic." *Journal of Computing Systems* 1 (1953): 111–149.

Z *zagadnień logiki i filozofü* (Problems of logic and philosophy). Edited by Jerzy Słupecki. Warsaw: Panstwowe Wydawn, 1961. Essays. Contains bibliography.

In addition to the above, Łukasiewicz published about twenty papers and over fifty notes and reviews.

WORKS ON ŁUKASIEWICZ

Borkowski, Ludwik, and Jerzy Słupecki. "The Logical Works of J. Łukasiewicz." *Studia Logica* 8 (1958): 7–56.

Jordan, Z. A. *The Development of Mathematical Logic and of Logical Positivism in Poland between the Two Wars.* Oxford, 1945.

Kotarbiński, Tadeusz. "Jan Łukasiewicz's Works on the History of Logic." *Studia Logica* 8 (1958): 57–62.

Kotarbiński, Tadeusz. *La logique en Pologne.* Rome: Signorelli, 1959.

Prior, A. N. "Łukasiewicz's Contributions to Logic." In *Philosophy in the Mid-century,* edited by Raymond Klibansky, 53–55. Florence, 1958.

Sobociński, Bolesław. "In Memoriam Jan Łukasiewicz." *Philosophical Studies* (Ireland) 6 (1956): 3–49. Contains bibliography.

Sobociński, Bolesław. "La génesis de la Escuela Polaca de Lógica." *Oriente Europeo* 7 (1957): 83–95.

Czesław Lejewski (1967)

LULL, RAMÓN

(c. 1232–1316)

Ramón Lull (or Llull), the Franciscan philosopher, was born in Palma de Mallorca in the Balearic Islands. Lull received the education of a rich knight of the period, but was converted from dissipation to a devout life in about 1263. At that time Majorca was largely populated by Muslims, and Islam was still the great rival of Christianity. Lull resolved to dedicate himself to the conversion of Muslims and to seek martyrdom for their sake. After selling almost all his possessions and undertaking various pilgrimages, Lull spent nine years (c. 1265–1274) in Majorca, acquiring a profound knowledge of Arabic. In 1274 he had a vision that revealed to him the Principles on which his combinatory Art should be based. In 1275 James II of Majorca had Lull's early writings examined for orthodoxy, and in 1276 James founded at Miramar in Majorca a monastery where Franciscans could study Arabic and Lull's Art to prepare for missions to Islam.

Lull appears to have divided his time in the years 1276–1287 between Miramar and Montpellier. In 1287 he began a series of journeys to the courts of kings and popes with the hope of persuading them to support his missionary, his reforming, and (later) his crusading projects. Lull placed his hopes principally in the papacy and in the kings of France and Aragon. His only apparent success was when the Council of Vienne (1311–1312) ordained the creation of chairs for Hebrew, Arabic, and "Chaldean" in five centers. Lull also undertook missions to Tunis (1293), to Bougie, in Algeria (1307), and again to Tunis (1314–1315). The traditional account of his martyrdom at Bougie cannot be sustained. He seems to have died in Majorca before March 25, 1316. He has been beatified by the Roman Catholic Church.

In the years 1288–1289, 1297–1299, 1309–1311, and probably 1306, Lull taught at the University of Paris; he also lectured publicly at Naples and Montpellier. Starting about 1272, he began to write incessantly. Some 240 of his approximately 290 works have survived. About 190 are only preserved in Latin (over 100 of these Latin works remaining unpublished until recently), although most of them were originally written in Catalan. Some of his works were originally written in Arabic; all these Arabic versions, however, are lost.

The desire to bring about the conversion of Muslims and Jews, as well as pagan Tartars, which inspired Lull's ceaseless activity, also inspired his writings. The desire for the reunification of the church (divided into hostile East and West), and for the complete reunification of humankind, through Christianity, dominated Lull's life. Lull's Art and his whole philosophy are apologetic and Franciscan, aimed at conversion by peaceful persuasion. Lull's advocacy of an armed crusade came late in his life; it was intended as subsidiary to missions. Lull's life was a continual battle with Islam, not only in Spain and North Africa, but also, from 1298, in Paris, with the "Averroists." In opposition to the "double-truth" theory imputed to such rationalist philosophers as Boethius of Dacia and Siger of Brabant, whose master was Aristotle as interpreted by Averroes, Lull sought to reestablish the unity of truth in philosophy and theology.

THE *ARS COMBINATORIA*

According to Lull, God, insofar as he can be known to men, consists of a series of divine attributes, or "Dignities," which are also the absolute Principles of Lull's Art. These Dignities (in the later works goodness, greatness,

eternity, power, wisdom, will, virtue, truth, glory) are the instruments of God's creative activity, the causes and archetypes of all created perfection. The essence of the Art does not (as is often thought) consist in demonstration, but in the metaphysical reduction of all created things to the Dignities, which are Principles of knowing as well as of Being, and in the comparison of particular things between themselves in the light of the Dignities, by means of such relative predicates as difference, agreement, contrariety, beginning, middle, end, majority, equality, minority. The absolute and relative predicates together form the self-evident principles common to all the sciences. These principles are combined in circular figures, where letters are substituted for their names (*B* = goodness, and so on).

Lull's treatises on different sciences (cosmology, physics, law, medicine, astronomy, geometry, logic, psychology) are applications of his general Art. Lull made continual efforts to simplify and popularize his Art, from the primitive version in the *Ars Magna* of about 1274 to the final *Ars Generalis Ultima* of 1308. The latter work and also the *Arbre de ciència (Arbor Scientiae*) of 1296 are more philosophical and less polemical in purpose than the original Art. A vast encyclopedia that found favor in the Renaissance, the *Arbre* is an attempt to classify all knowledge under a unified plan. Lull's influence was acknowledged by Gottfried Wilhelm Leibniz in the later philosopher's search for the *caracteristica universalis* and *ars combinatoria,* which he hoped would make possible the deduction of all truths from basic concepts. Despite the clear analogies between the two systems, Leibniz only took over part of Lull's ideas, omitting Lull's original purpose of the Art as a means of converting infidels.

Lull was the first Christian philosopher of the Middle Ages to use a language other than Latin for his major works. Although he did not receive a university training, he enjoyed advantages denied to the great Scholastics. Of the three Mediterranean cultures of his time he knew Latin Christianity and Islam well and was aware of Greek Christianity. The basis of Lull's philosophy was Neoplatonic realism as transmitted through the Augustinian tradition: his exact use of John Scotus Erigena, Anselm, the Victorines, Bonaventure, and Roger Bacon is still debated. Lull was also familiar with the writings and beliefs of his Jewish and Muslim contemporaries.

All Lull's contemporaries shared a vision of the world based on Neoplatonism. The common belief in a hierarchy, or ladder, of creation, the theories of the four elements and of the spheres, the organization of reality by numerical-geometrical symbolism, the idea of man as a microcosm, were all incorporated by Lull into his system. That excellent scholars have seen the inspiration of Lull's theory of the Dignities in the Muslim *hadras* or in the Jewish kabbalist *sefirot* (both terms for the divine attributes) shows that Lull's doctrine (although of Christian derivation) provided a reasonable basis for a dialogue with the Muslim and Jewish elites. Much the same is true of the doctrine of correlative principles, developed in Lull's later works, by which each attribute unfolds into a triad of interconnected principles, agent, patient, and the action itself, expressing the relations between God, a creature, and God's action. Lull probably took this doctrine from the Arabic writer al-Ghazālī, whose *Logic* he translated. It is more probable that Lull derived the idea for the figures that illustrate his Arts from contemporary Spanish kabbalists or from the circular figures of Isidore of Seville's well-known cosmological treatise *De Natura Rerum* than from Ibn al-ʿArabī of Murcia, who has been suggested as his source.

Two of the most striking characteristics of Lull's philosophy and theology—his "rationalism" and his emphasis on the importance of action, shown in his constant appeals to Christian rules—owe their prominence in his system to its polemical inspiration. Lull's "necessary reasons," by which he proposed to "prove" the articles of faith, are reasons of congruence and analogy, not purely deductive principles. In opposition to Islamic scholastic theology (the *kalam*), which tried to demonstrate the Faith, Lull sought to show that the Muslim, who began with a belief in monotheism and the divine attributes, must proceed to Christianity. Despite the nondeductive character of his works, Lull's thought is deeply rational. Only seldom in his mystical writings does love eclipse the intellect or obscure its powers. For him, contemplation issues in action. *Blanquerna* and *Felix* are the first philosophical-social novels of Europe. In *Blanquerna* Lull sketched his plan for a *Pax Christiana,* a society of nations presided over by the papacy.

See also al-Ghazālī, Muhammad; Anselm, St.; Aristotle; Augustinianism; Bacon, Roger; Boetius of Dacia; Bonaventure, St.; Erigena, John Scotus; Ibn al-ʿArabī; Islamic Philosophy; Jewish Philosophy; Kabbalah; Leibniz, Gottfried Wilhelm; Logic, History of; Medieval Philosophy; Neoplatonism; Realism; Siger of Brabant.

Bibliography

Lull's Latin works may be found in a new critical edition, *Raimundi Lulli Opera latina* (vols. 1–5, Palma, 1959–1967; vols. 6–, Turnhout, Belgium, 1978–; at present 28 vols. have

appeared). Reprint editions have made two important earlier collections of Lull's Latin works available, one by the printer Zetzner, *Raimundus Lullus, Opera* (2 vols., Strasbourg, 1651; reprinted, Stuttgart, 1996), another by Ivo Salzinger, *Raymundi Lulli Opera omnia* (8 vols., Mainz, 1721–1742; reprinted, Frankfurt am Mainz, 1965). Dr. Viola Tenge-Wolf has undertaken the digitalization of the over 2,000 microfilms of Lullian manuscripts in the Raimundus-Lullus-Institut of the University of Freiburg im Breisgau.

Editions of Lull's Catalan works include *Obres de Ramón Lull* (21 vols., Palma and Barcelona, 1906–1950) and the supplementary *Nova edició de les obres de Ramon Llull* (Palma, 1990–, in progress). A handy edition is *Ramon Llull, Obres essencials* (2 vols., Barcelona, 1957–1960). *Selected Works of Ramon Lull* have been translated into English by Anthony Bonner (2 vols., Princeton, NJ, 1985; who has also published the Catalan texts in 2 vols., Palma, 1989; these volumes include an important chronological catalog of Lull's Works, pp. 1257–1304 in the English edition). A complete catalog of the printed editions of Lull's works has been published by Elíes Rogent and Estanislau Duràn, *Bibliografía de les impressions lul.lianes* (Barcelona, 1927).

For studies of Lull's life and work see Tomás y Joaquín Carreras y Artau, *Historia de la filosofía española,* vol. 1 (Madrid: Real Academia de Ciencias Exactas, Físicas y Naturales, 1939) pp. 231–640; vol. 2 (Madrid, 1943) contains a valuable history of Lullism. For an English biography of Lull see E. Allison Peers, *Ramon Lull: A Biography* (London, 1929), which should now be supplemented by Jocelyn N. Hillgarth, *Ramon Lull and Lullism in 14th-Century France* (Oxford, 1971). A general introduction to Lull's worldview is provided by Robert Pring-Mill, *El microcosmos lul.lià* (Palma: Editorial Moll, 1962; in German, Stuttgart, 2001). For the intellectual and religious context in which Lull's Art developed see Harvey J. Hames, *The Art of Conversion: Christianity and Kabbalah in the 13th Century* (Leiden, 2000). For Lull's influence in the sixteenth century see Frances A. Yates, *Lull & Bruno: Collected Essays* (2 vols., London, 1982).

Rudolf Brummer has published a *Bibliographia lulliana: Ramon-Llull-Schriftum, 1870–1973* (Hildesheim, 1976); which has been supplemented for the years 1974–1984 by Marcel Salleras i Carolà in *Randa* (Barcelona) 19 (1986): 153–198. Current bibliography may also be found in the review *Estudios lulianos* (now *Studia lulliana*), published in Palma since 1957.

Jocelyn Nigel Hillgarth (1967)
Bibliography updated by Charles Lohr (2005)

LUNACHARSKI, ANATOLI VASILYEVICH

See *Lunacharskii, Anatolii Vasil'evich*

LUNACHARSKII, ANATOLII VASIL'EVICH
(1875–1933)

Anatolii Vasil'evich Lunacharskii (also Lunacharsky), the Marxist philosopher and literary critic and Soviet administrator, joined the Russian Social Democratic Party in Kiev in 1892. Because of his political activities as a secondary school student, he was denied admission to Russian universities. He attended lectures at Kiev University and at the University of Zürich, where in 1894–1895 he studied under Richard Avenarius, who converted him to empiriocriticism. Lunacharskii returned to Moscow in 1897, was exiled to Vologda (1899–1902), and spent several years in western Europe between 1904 and 1917. He was the first Soviet people's commissar for education (1917–1929).

Lunacharskii's contributions to philosophy are concentrated in value theory (which he rather misleadingly called biological aesthetics), ethics, and philosophy of religion. Like the positivists, he denied the adjudicability of value disputes. "In order to show," he wrote, "that a given type of valuation is in its very root worse than another type, the scientist must oppose one criterion to another, but the choice between criteria is a matter of *taste*, not *knowledge*" ("K voprosu ob otsenke" [On the question of valuation], 1904, reprinted in *Etiudy*, Moscow, 1922, p. 55).

In ethics and social philosophy Lunacharskii was a "Nietzschean Marxist." He called himself an aesthetic amoralist and rejected the categories of duty and obligation, stressing instead free creative activity, the "artistic" shaping of ends and ideals. "Nietzsche," he declared, "and all the other critics of the morality of duty, have defended the autonomy of the individual person, the individual's right to be guided in his life solely by his own desires" ("'Problemy idealizma'…," [Problems of idealism…] in *Obrazovanie* 12 [2] [1903]: 133).

Lunacharskii called his individualism macropsychic, or "broad-souled," to distinguish it from "narrow-souled" (micropsychic) individualism. It approached collectivism in its stress on the historical community of the creators of culture.

Traditional religious attitudes and institutions, according to Lunacharskii, could and should be given a new, socialist content. The old religions—supernatural, authoritarian, "antiscientific"—must be replaced by a new religion that will be humanistic, libertarian, and "scientific." The building of socialism and the shap-

ing of the high human culture of the future will be a building of God (*bogostroitel'stvo*). "Scientific socialism," Lunacharskii declared, "is the most religious of all religions, and the true Social Democrat is the most deeply religious of men" ("Budushchee religii" [The future of religion], p. 23). The religion of God-building will soften the sting of mortality by intensifying man's awareness of the "universal connectedness of life, of the *all-life* which triumphs even in death" ("Eshche o teatre i sotsializme [Once more on the theater and socialism], in *Vershiny*, Vol. I, 1909, p. 213). The new religion, imparting a sense of "joyous union with the triumphant future of our species," will be full of drama and passion, having its own "saints and martyrs." It will be worthy to stand beside medieval Christianity in the "universal arsenal of art and inspiration" (*R. Avenarius: Kritika chistogo opyta v populiarnom izlozhenii A. Lunacharskovo* [R. Avenarius: *Critique of Pure Experience*, Expounded for the layman by A. Lunacharskii], Moscow, 1905, p. 154).

See also Avenarius, Richard; Marxist Philosophy; Marx, Karl; Positivism; Russian Philosophy; Socialism; Value and Valuation.

Bibliography

WORKS BY LUNACHARSKII

On Education: Selected Articles and Speeches. Moscow: Progress Publishers, 1981.

On Literature, and Art. Translated by Avril Pyman and Fainna Glagoleva. Moscow: Progress Publishers, 1973.

WORKS ON LUNACHARSKII

Fitzpatrick, S. *The Commissariat of Enlightenment: Soviet Organization of Education and the Arts under Lunacharsky, October 1917–1921*. Cambridge, U.K.: Cambridge University Press, 2002.

O'Connor, T. E. *The Politics of Soviet Culture: Anatolii Lunacharskii*. Ann Arbor, MI: UMI Research Press, 1983.

Tait, A. L. *Lunacharskii: Poet of the Revolution (1875–1907)*. Birmingham: Department of Russian Language and Literature, University of Birmingham, 1984.

George L. Kline (1967)

Bibliography updated by Vladimir Marchenkov (2005)

LUTHER, MARTIN

(1483–1546)

Martin Luther, the German theologian and leader of the Protestant Reformation, was born at Eisleben, Saxony. His father came of peasant stock, but established himself during Luther's boyhood as a successful copper miner in Mansfeld. From 1501 to 1505 Luther attended the University of Erfurt, and then, at his father's wish, he began the study of law; but a spiritual crisis, occasioned by a violent thunderstorm, induced him to enter the Erfurt monastery of the Augustinian Friars. Despite conscientious and even overscrupulous attention to his monastic duties, Luther was obsessed by dread of God's anger, and his superior tried to direct the young man's energies and undoubted ability into a scholar's calling. From 1512 he was biblical professor at the new University of Wittenberg, a position he held, despite interruptions, until his death.

THEOLOGICAL DEVELOPMENT

Three stages may be distinguished in Luther's theological development. Between 1512 and 1517, and probably (in the judgment of most scholars) not later than 1515, his biblical studies led to a theological reorientation, at the center of which was an interpretation of the justice of God in Romans 1:17, not as a divine attribute expressed in punishment and reward, but as the activity by which God makes men just ("justifies" them). This justice of God is identical with His grace: It is not conditional upon human merit, but is received by faith alone (faith itself being a work of God in man). The working out of this basic insight made Luther increasingly critical of late scholastic theology and of ecclesiastical abuses. The appearance of the Ninety-five Theses on indulgences (1517), although they were not intended as "un-Catholic," was interpreted by Luther's opponents as ecclesiastically disloyal and subversive. Luther had, indeed, touched on the heart of medieval piety, the sacramental system, since indulgences belonged to the sacrament of penance.

The second period of Luther's development, from 1517 to 1521, was marked by his struggle with the Roman authorities, during which he abandoned the theory of papal, and even ecclesiastical, infallibility. In his *Babylonian Captivity* (1520), he made a systematic attack on the sacramental system, reinterpreting a sacrament as, like preaching, a form of the divine Word, by which God offers man His justice and creates the response of faith. The "church" is defined, not in terms of hierarchical authority, but as the communion of those whom Christ rules with His Word, all of whom are priests. Luther's basic insight into the character of Christian justice (or righteousness) was sharpened during this same period by greater precision in the distinction (already made before 1517) between Law and Gospel. The Law of God can only demand and condemn; it cannot be used by man as a means of self-salvation through strict obedience. The

security of man before God lies solely in the Gospel, with its word of free forgiveness.

During the third period, after 1521, Luther's attention was turned to rival reformers who departed from him on particular points, or who demanded a more radical transformation of the church than he was prepared to countenance. Many of the radicals sought to establish communities in which the ethic of the Sermon on the Mount should be the sole rule of social conduct. Against them, Luther again argued for the distinction between Law and Gospel. Just as it is wrong to place Law between God and the conscience, so it is wrong to regulate society by the Gospel. The conscience needs the gospel of forgiveness, but society can only be founded upon the law of retributive justice (though Law should always be the agency of love). The two "realms," or "kingdoms," of Heaven and Earth—that is, the two ways in which God rules over the world of men—are not to be confused.

In his controversy with the humanist leader Desiderius Erasmus, which also belongs within the third stage of his development, Luther again believed himself to be fighting for the gospel of forgiveness. He acknowledged that Erasmus's selection of the theme to be debated—namely, the freedom of the will—came closer to the decisive issue than did the questions of the papacy, purgatory, and indulgences. Luther was not, of course, interested in the psychology of human action as such but in preserving his original insight into the agency of divine grace. He acknowledged a measure of human freedom in matters that do not concern salvation, but refused to make salvation depend at any point on the inherent possibilities of human nature. He therefore located the power of man's decision for God in the Gospel itself, and in the secret influence of the Holy Spirit. For Luther, this did not mean that God acts coercively, thereby doing violence to man's will, but that God is sovereign over the will and can direct it to His ends. Man acts voluntarily (that is, as he wills) even in those matters that concern his salvation. But the will itself is controlled by God. It cannot change itself from an evil to a good will: It must *be* changed under the influence of the Spirit.

Luther was not, of course, a philosopher. He was primarily a theologian, obliged by circumstances to become a rebel and a reformer. Indeed, it is often supposed that he was an implacable enemy of philosophy, and to this problem the remainder of this article will be devoted. It will appear how closely Luther's views on reason and philosophy are related to the central theological concerns (Christian justice and the two realms of Heaven and Earth) that have been sketched above.

ATTITUDE TOWARD PHILOSOPHY

It is not hard to document from Luther's own writings the common accusation that he was an anti-intellectualist. His description of reason as "the Devil's Whore" is well known, and he recommended that the faithful sacrifice reason, or slay it, as the enemy of God. Many have seen in this apparent antirationalism evidence of Luther's Ockhamist heritage, but this is an oversimplification of an intricate historical problem. Luther did not invariably decry reason. In his celebrated appearance before the Diet of Worms (1521) he seemed to appeal to a double norm—Scripture and reason. (He refused to recant unless convinced by "the testimonies of Scripture or by evident reason.") And sometimes he showered extravagant praise upon reason as the greatest of God's gifts, as the "inventress and mistress of all the arts, of medicine and law, of whatever wisdom, power, virtue and glory men possess in this life."

Luther accepted the traditional view that reason set man apart from the brute beasts and gave him dominion over the world. Clearly, the problem is to explain, not an extreme one-sidedness, but a strange ambivalence. And the appeal to Luther's alleged Ockhamist heritage cannot help to explain his attitude until the Ockhamist understanding of reason is itself clarified and the extent of Luther's overall dependence upon nominalism is carefully assessed. The persistent image of nominalist theology as antirational and un-Catholic requires reconsideration in the light of recent studies, and verbal echoes of nominalism in Luther's writing may prove of no great significance. In any case, the primary historical task is to examine Luther's actual utterances on reason and philosophy and to view them in relation to the inner structure of his thought.

THE CONCEPT OF REASON. The apparent ambiguities in Luther's utterances on reason can be explained, in part, by his fundamental distinction between the two realms of human existence. At one and the same time, man lives toward God in the Heavenly Kingdom and toward his natural and social environments in the Earthly Kingdom. Luther judges human reason to be an adequate instrument for dealing with earthly affairs, that is, the maintaining of physical subsistence (*oeconomia*) and the regulation of life in society (*politia*). In this realm, reason is legitimately exercised and affords the only light man needs. But in spiritual affairs the situation is quite different. Reason has no understanding of what it is that commends a man to God. Therefore God has given His Word (in the Scriptures), and reliance upon reason could, in

this realm, only be perverse and presumptuous. The way of salvation could never have been thought out by rational enquiry, for all God's works and words transcend reason. The Word of God is apprehended, not by reason, but by faith.

This does not mean that, for Luther, reason must be totally excluded from theology. He allowed for the possibility of taming reason's presumptuousness. It then becomes the handmaid of faith. Luther spoke of reason as illumined by faith, regenerated, or born anew. Sometimes the notion of regenerate reason tended to coalesce with the notion of faith itself. But generally, Luther seemed to think of regenerate reason as the human capacity for orderly thought being exercised upon material provided by the Word. Perhaps this is what he meant by the correlation of Scripture and reason in his answer before the Diet of Worms: He was willing to be persuaded either by direct biblical citations or by plain inferences from them. He certainly did not mean to set reason beside Scripture as an independent and supplementary source of theological knowledge.

The doctrine of the two realms provides, then, the framework for a threefold distinction by means of which Luther's various utterances on reason may, for the most part, be harmonized. We have to distinguish between natural reason, ruling within its own domain (the Earthly Kingdom); presumptuous reason, encroaching on the domain of faith (the Heavenly Kingdom); and regenerate reason, serving faith in subjection to the Word of God. Luther does not represent an anti-intellectualist dismissal of disciplined thought; he tries to formulate a theological critique of reason, in which the boundary lines of reason's competence are sharply drawn. Only in the second of these three contexts does reason appear as "the Devil's Whore." In the first it is the greatest of God's gifts; in the third, an excellent instrument of godliness.

It is necessary, however, to carry the analysis further and to show that Luther's invective against reason is focused upon a quite specific blunder that reason makes when it trespasses, unregenerate, upon the domain of faith. It then appears that the *sacrificium intellectus* for which he calls cannot be understood simply as an epistemological doctrine, but rests upon a more strictly theological (or soteriological) concern. For in many passages from his writings, what Luther meant to express by his colorful invective against reason, was his constant astonishment at the heart of his own gospel: the unconditioned character of God's grace. Reason must be "put to death" because it cannot comprehend the miracle of divine forgiveness, and therefore stands in the way of man's receiving the justice of God. Reason became identified in Luther's mind with the religious attitude of the natural (that is, unregenerate) man, who can conceive only of a strictly legalistic relationship to God. *Ratio* became virtually synonymous with a definite *opinio,* and it is by no means accidental that the two words can be found side by side in several passages. Nor, of course, was this usage wholly eccentric, since Lewis and Short's Latin-English dictionary gives as one of the meanings of *ratio* a "view or opinion resting upon reasonable grounds." And Luther fully acknowledged a certain reasonableness about the assumption that a just God must require "good works" as the precondition of communion with Him.

Consequently, the proclamation of an unconditioned grace—which demands nothing, save the acceptance of faith—can be greeted by reason only with incredulity. What needs to be "sacrificed," therefore, is not human rationality, without qualification, but rather the legalistic mentality of the natural man. As Luther put it, grace must "take us out of ourselves," and we must learn to "rise above reason." In short, Luther's concept of reason (at least, when his remarks about it are pejorative) is not formal, but material. *Ratio* is a concrete attitude rather than the faculty or structure of reasoning. When the natural man turns his thoughts to religion, he carries over into the Heavenly Kingdom presuppositions that, however appropriate in dealing with his social existence in the Earthly Kingdom, no longer apply. For the Kingdom of Christ is a realm, not of law, but of grace (*das Reich der Gnaden*).

THE CONCEPT OF PHILOSOPHY. Because Luther's views on reason are set in a theological context, they are not always directly relevant to the problem of faith and reason as the philosopher normally understands it. But Luther's standpoint certainly had consequences for the philosophy of religion, and more particularly for the problem of a natural theology. For Luther there could be no question of treating the truths of reason as a kind of foundation for the truths of revelation. The continuity between nature and grace, as presented in the classical scholastic scheme, is broken. There is no rational preamble to faith, because reason is not a neutral instrument for the discovery of objective truths; it is misled by its own bias and even corrupted by sin—that is, by the egocentricity of the unredeemed man. For man in sin actually prefers a God of law, upon whom he can establish a claim. Revelation does not confirm or supplement reason: It stands in contradiction to reason, until the natural man is "born anew." The religion of reason is not merely insufficient or imperfect, but perverted and erroneous. Luther

does not deny that a limited knowledge of God is available to reason; but the egocentricity of man in sin is a fatal defect, productive of idolatry and superstition. Reason makes God as it wills Him to be, and turns this natural knowledge into idolatry. The god of reason is a false God.

In general, Luther's direct statements about philosophy closely parallel his judgment on reason. As early as the Lectures on Romans (1515–1516) he had come to see his mission as a protest against philosophy, and his writings are interspersed with abusive descriptions of Aristotle ("the stinking philosopher," "the clown of the High Schools," "the blind pagan," etc.). Thomas Aquinas, who symbolized the attempt to synthesize Aristotle and the Christian faith, is treated with similar disrespect. Nevertheless, Luther could on occasion speak deferentially of philosophy and even of Aristotle. He approved of much that the Greek philosopher had written on social ethics and ranked Cicero's ethics even higher. He freely acknowledged that the Christian had much to learn from philosophy in this area.

The key to Luther's ambivalence lies, as with his concept of reason, in the distinction between the two realms. The boundaries are carefully drawn. Philosophy is an excellent thing in its own place, but if philosophical categories are transferred into theology, the result can only be confusion. Luther saw philosophy as tied to the empirical world (the Earthly Kingdom), whereas theology is concerned with things unseen (the Heavenly Kingdom). He was not, strictly speaking, hostile to Aristotle, but to the theological application of Aristotelianism by the Schoolmen. Of course, some of the Greek philosopher's doctrines already had a theological bearing (for example, on the immortality of the soul and on divine Providence). These Luther dismissed. But he approved Aristotle's treatises on the sermonic arts (logic and rhetoric) and, with qualifications, those on moral philosophy.

Perhaps the most important illustration of Luther's attitude toward Aristotle is afforded by his discussions of moral "habit" (Latin, *habitus*; Greek, *hexis*). In the *Nicomachean Ethics,* Aristotle taught that "we become just by performing just acts." Luther's opponents apparently gave this doctrine a theological application: That is, it was used to support the claim that good works must precede justification. In assailing the concept of habit, Luther is not offering a philosophical critique of Aristotle, but rejecting the theological application of Aristotelian doctrines. A philosophical theory belongs within the Earthly Kingdom. The Schoolmen mix the kingdoms.

COMPARISON WITH NOMINALISM. Luther's distinction between two spheres of knowledge (philosophy and theology) and between two organs of knowing (reason and faith) certainly invites comparison with late medieval Scholasticism. There is perhaps a prima facie probability that Luther's views on reason and philosophy were under the influence of the nominalists. His main instructors at Erfurt were nominalists, and it is noteworthy that Luther could speak of William of Ockham with apparent respect, even calling him "my dear master." He adopted the nominalist view of universals, and he explicitly owned a debt to the nominalist Pierre d'Ailly in the doctrine of the Real Presence. Other possible debts have been argued with more or less plausibility, although it can hardly be denied that Luther left nothing unchanged that he borrowed from others. At least the possibility is open that at the outset the sharp distinction between faith and reason may have been suggested to him by his familiarity with the Ockhamist school.

It may be that the separation of theology and philosophy in Luther is to be explained partly by his acceptance, along with the nominalists, of a strict Aristotelian concept of science. Against Thomas, Luther agreed with the nominalists that since theology rests upon assertions of faith, it cannot be classed as a science. Philosophy (which is the sum total of rational knowledge and embraces the various sciences) deals with the visible world, which is accessible to reason. Theology deals with an invisible world, accessible only to faith. Such points of agreement between Luther and the Ockhamists cannot, however, conceal the sharp differences between them. Quite apart from the fact that Luther developed a divergent concept of faith, his standpoint represents a different basic concern. The interest of the Ockhamists in the problem of faith and reason was primarily epistemological. Hence they devoted considerable thought to relating the cognition of reason to the cognition of faith, and sought in various ways to bridge the gap that they had apparently cut between the two. Nominalist theologians tried to comprehend both faith and reason within a single epistemological scheme. They regarded theological propositions (once established) as subject to rational scrutiny, believed that merely probable arguments could lead to faith when the will cooperates, and argued that revelation was given precisely to those who made maximum use of their rational capacities. Luther, on the other hand, was not interested in narrowing the epistemological gap. On the contrary, the problem for him was graver, because he allowed for the corruption of reason by human sinfulness. Hence his restrictions on reason, even if they were built on a nominalist view of science, go beyond it in

what is primarily a theological, rather than philosophical, concern.

THE THEORY OF "DOUBLE TRUTH." The nominalist distinction between the spheres of faith and of reason has commonly been interpreted as though there were a disharmony, or even a contradiction, between them. Indeed, the doctrine of a "double truth"—that is, that a proposition may be true in theology, but false in philosophy—has been attributed to the nominalist theologian Robert Holkot. Properly speaking, double truth seems never to have been a consciously adopted "doctrine" in the Middle Ages, but rather an accusation leveled against theological opponents. There does not seem to be adequate reason to attribute it to any of the nominalists. True, they admitted some apparent conflicts, for instance, that the Christian belief in the Trinity, when formulated according to the rules of Aristotelian logic, contained real contradictions. But this simply prompted the quest for a higher logic, which could embrace both the traditional Aristotelian rules and also the rules appropriate to the peculiarities of theological truth.

A doctrine of double truth could, however, be attributed to Luther with some plausibility, since he explicitly said that "the same thing is not true in different disciplines" (*Disputation on the Proposition, "The Word became flesh,"* 1539). But Luther himself did not use the expression "double truth," and a close inspection of his argument suggests that, despite appearances, he really had a rather different thesis in mind. What he was trying to defend might better be called a "theory of multiple meaning." Neither "twofold" nor "truth" quite pinpoints Luther's thesis, and perhaps even "manifold truth" (Bengt Hägglund's phrase) is still misleading. If we may paraphrase the drift of Luther's argument, he seems to be saying that *homo loquens* reflects and communicates, not by means of a single, universally valid language, but by means of several languages, which are relative to particular disciplines or areas of experience. Hence the meaning of a term or proposition is determined by the area of discourse: If transferred from one area of discourse to another, a term may acquire a different meaning, or have no meaning at all. To use Luther's own examples, it makes no sense to ask the weight of a line or the length of a pound.

Whether correct or not, this argument bears a close resemblance to ideas that played an important role in twentieth-century linguistic philosophy, and is therefore not likely to be dismissed as obscurantism or anti-intellectualism. Unfortunately, Luther's argument is not developed with adequate precision, either in this *Disputation* or elsewhere. But it is not an isolated argument. The basic thesis—that the same form of words may have different meanings in different disciplines—underlies many of his remarks about the relation of ethics and theology. For example, the proposition that fallen man can do no good is fundamental to Luther's teaching on justification. But Luther admits that this is true only in a theological, not in an ethical, context, for in each context the word *good* means something different. This is, perhaps, a statement of double truth, but only because it rests on a theory of multiple meaning. Thus interpreted, "double truth" does not imply contradiction, but excludes it, since real contradiction is possible only within a single realm of discourse. As Luther put it in the first thesis of the *Disputation*: "Although we must hold to the saying, 'One truth agrees with another,' nevertheless the same thing is not true in different disciplines."

See also Ailly, Pierre d'; Aristotelianism; Aristotle; Averroism; Cicero, Marcus Tullius; Erasmus, Desiderius; Faith; Holkot, Robert; Reason; Reformation; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS BY LUTHER

The definitive German edition of Luther's writings in Latin and German is *D. Martin Luthers Werke. Kritische Gesamtausgabe* (Weimar, 1883–). The most comprehensive English version is *Luther's Works: American Edition,* edited by Jaroslav Pelikan and Helmut T. Lehmann (St. Louis: Concordia, 1955–1986, 55 vols.).

WORKS ON LUTHER

Recent Studies

Three twentieth-century books dealing with Luther's views on reason and philosophy are Bengt Hägglund, *Theologie und Philosophie bei Luther und in der occamistischen Tradition. Luthers Stellung zur Theorie von der doppelten Wahrheit* (Lund: Gleerup, 1955); Bernhard Lohse, *Ratio und Fides. Eine Untersuchung über die Ratio in der Theologie Luthers* (Göttingen: Vandenhoeck and Ruprecht, 1958); and B. A. Gerrish, *Grace and Reason: A Study in the Theology of Luther* (Oxford: Clarendon Press, 1962). One of the most adequate treatments of Luther's intellectual background is still Otto Scheel, *Martin Luther, Vom Katholizismus zur Reformation,* Vol. I, 1st ed. (Tübingen: n.p., 1916); Vol. II, 3rd and 4th eds. (Tübingen, 1930).

Luther and Nominalism

The literature dealing with the general question of Luther's relation to nominalism is sketched in Leif Grane, *Contra Gabrielem. Luthers Auseinandersetzung mit Gabriel Biel in der Disputatio contra scholasticam theologiam 1517* (Copenhagen: Gyldendal, 1962). Grane's own discussion is focused on the theological rather than on the philosophical

points, as is the work of Reinhard Schwarz, *Fides, Spes und Caritas beim jungen Luther unter besonderer Berücksichtigung der mittelalterlichen Tradition* (Berlin: De Gruyter, 1962). The work of Heiko Augustus Oberman in *The Harvest of Medieval Theology: Gabriel Biel and Late Medieval Nominalism* (Cambridge, MA: Harvard University Press, 1963) is intended to lay the foundations for a study of nominalism in relation to the beginnings of Reformation theology.

Additional Background

For the wider aspects of Luther's thought, see the articles and bibliographies under "Luther" in *Die Religion in Geschichte und Gegenwart,* 3rd ed. (Tübingen, 1960), Vol. IV, pp. 480–523, which may be brought up to date by the annual listings of the *Luther-Jahrbuch.*

B. A. Gerrish (1967)

LUTHER, MARTIN [ADDENDUM]

The renaissance of Luther studies enjoyed by the twentieth century continues apace. The massive critical, or Weimar (WA), edition of his work has recently been finished in 127 volumes. Important interpretive works have been published and discussed, including a major three-volume theological biography by Martin Brecht (1985–1993). Despite, or perhaps because of, this wide variety of scholarship, even such a seemingly simple theme as faith and philosophy in Luther has no consensus among interpreters.

There has been a welcome re-reception of Luther by Catholic scholars, starting with Joseph Lortz in 1939. A significant ecumenical consensus was reached by evangelical and Catholic scholars on the occasion of Luther's 500th birthday, noting that Vatican II reflects many of the concerns Luther addressed in his own witness to the gospel. This larger ecumenical interpretation has led to studies that appreciate the more Catholic side of Luther as a reformer and teacher of the whole church, not excluding his doctrine of justification by faith.

The importance of understanding each of Luther's distinct writings within its own historical, institutional, and rhetorical context is a major virtue of modern Luther studies. Equally important is an understanding of Luther against his late-medieval background. This has led to a new appreciation for Luther's dependence upon nominalism, especially the school of Ockham (*via moderna*). Recent scholarship has documented Luther's use of philosophy and logic in his theological arguments, including elements of nominalist logic from Gabriel Biel and Pierre d'Ailly. Luther's strong language against reason, philosophy, and Aristotle were aimed at a particular target, namely, the scholastic theology of an earlier age (*via antiqua*). Unlike the Neo-Kantian and existentialist interpretations of Luther, recent scholars have argued that Luther nowhere has a complete condemnation of metaphysics or ontology in theological understanding. Indeed, some scholars now find a kind of ontology in Luther's conception of salvation.

At the heart of contemporary controversy surrounding the interpretation of Luther is the so-called Finnish school, including the work of Tuomo Mannermaa (2005). On this view, Luther taught that Christ is really and personally present in faith for the Christian. Justification is not simply alien, external, and forensic but also relational and ontological. The relationship between human being and the divine Trinity is understood not only as an external declaration of a righteousness that is not our own but also as the growth of Christ-like love through faith. The similarities of Luther's view thus understood, and the Eastern Orthodox notion of *theosis* (divinization), has been a key point in the Finnish school. Even given this new understanding, Luther consistently rejected philosophical ontology and scholastic metaphysics. When discussing the presence of Christ, he refused to go beyond what was promised in the Word. "But how He is present—this is beyond our thought; for there is darkness." (*Lectures on Galatians* [1535], WA 40/1:229). Thus, Luther appears to have used philosophical tools and concepts but refused to build theology on philosophical systems. For Luther, theology is grounded on the Word of God, not philosophical speculation.

See also Ailly, Pierre d'; Aristotle; Biel, Gabriel; Existentialism; Kant, Immanuel; Ockhamism; Ontology; Philosophy; Reason; William of Ockham.

Bibliography

Braaten, Carl, and R. W. Jenson, eds. *Union with Christ: The New Finnish Interpretation of Luther*. Grand Rapids, MI: Eerdmans, 1998.

Brecht, Martin. *Martin Luther*. 3 vols. Minneapolis: Fortress, 1985–1993.

Dieter, Theodor. *Der junge Luther und Aristoteles*. Berlin: de Gruyter, 2001.

Janz, Denis. *Luther and Late Medieval Thomism*. Waterloo, ON: Wilfrid Laurier University, 1983.

Joint Lutheran/Roman Catholic Study Commission. *Facing Unity*. Geneva: Lutheran World Federation, 1985.

Lortz, Joseph. *The Reformation in Germany*. New York: Herder & Herder, 1968 [1939].

Luther, Martin. *D. Martin Luthers Werke: Kritische Gesamtausgabe*. 4 parts in 127 vols. Weimar: H. Böhlau, 1883–2007.

Mannermaa, Tuomo. *Christ Present in Faith*. Minneapolis: Fortress, 2005.
White, Graham. *Luther as Nominalist*. Helsinki: Luther-Agricola-Society, 1994.
Wicks, Jared, ed. *Catholic Scholars Dialogue with Luther*. Chicago: Loyola University Press, 1970.

Alan G. Padgett (2005)

LU XIANGSHAN

(1139–1193)

Lu Xiangshan, also called Lu Jiuyuan, started the idealistic trend in Chinese philosophy. He emphasized the supremacy and self-sufficiency of the mind, contrary to his contemporary Zhu Xi, who stressed the need to discover reason and to acquire knowledge of the external world. He lived in the province of Jiangsi. His father was a respected member of the gentry, and from his early youth Lu was able to devote himself to the study of Confucius and Mencius. He disagreed with the views of the scholar Cheng Yi of the Northern Sung Dynasty.

Lu Xiangshan is known for the following:

> When a sage arises in the East,
> The mind is the same,
> And so is reason.

The same is true of sages born in the West, the North, and the South and of those born thousands of generations earlier and later. What he meant is that mind is the same the world over and at all times. From this fundamental thesis he drew the conclusions that mind has priority over all things and that reason has a universal validity.

Yang Jian, a disciple of Lu and a submagistrate, asked him, "What is the Original Mind?" Lu quoted the words of Mencius concerning the four kinds of virtues—*ren* (benevolence), *yi* (righteousness), *li* (decency), and *zhi* (knowledge)—and said, "This is the Original Mind." But Yang failed to understand what Lu meant. Some time after, a lawsuit was brought by a salesman of fans for Yang's verdict, and Yang again came to Lu with the same question. Lu answered, "In trying the case of the fan salesman, you were able to judge right that which is right and wrong that which is wrong. This is the Original Mind." Yang was then convinced that the mind is self-conscious and self-evident.

Lu was firmly convinced that there is a universal mind and a universal rationality: "What fills the universe is rationality; what the scholars should search for is to render the idea of rationality clear to all. The scope of rationality is boundless." He also quoted Cheng Hao's words, "The universe is great; yet it has its limitation," and then inferred from them that what is more perfect than the universe is rationality.

Again he said: "Rationality in the universe is so evident that it is never concealed. The greatness of the universe lies in the existence of rationality which is an order publicly followed and without partiality. Man with Heaven and Earth constitutes the triad. Why should one be egocentric and not in conformity with rationality?" Lu's main idea is that since each one has a mind and reason is inherent in mind, mind is reason. Furthermore, he says: "What is the happening of the universe is the ought-to-do-duty of man; what is the ought-to-do-duty is the happening of the universe."

See also Chinese Philosophy; Mencius; Rationality; Reason; Zhu Xi (Chu Hsi).

Bibliography

Works by Lu may be found in the typescript *The Philosophy of Lu Hsiang-shan, a Neo-Confucian Monistic Idealist,* translated by L. V. Cady in 2 volumes (New York: Union Theological Seminary, 1939), which also contains discussion.
Chan, Wing-tsit. *A Source Book in Chinese Philosophy.* Princeton, NJ: Princeton University Press, 1963.
De Bary, Theodore William. *The Message of the Mind in Neo-Confucianism.* New York: Columbia University Press, 1989.
Huang, Chin-hsing. "Chu Hsi versus Lu Hsiang-shan: A Philosophical Interpretation." *Journal of Chinese Philosophy* 14, no. 2 (1987): 179–208.
Huang, Siu-chi. *Lu Hsiang-shan: A Twelfth Century Chinese Idealist Philosopher.* New Haven, CT: American Oriental Society, 1944; Westport, CT: Hyperion Press, 1977.

Carsun Chang (1967)
Bibliography updated by Huichieh Loy (2005)

LYING

Lying may be defined as the making of a declarative statement to another person that one believes to be false, with the intention that the other person believe that statement to be true, and the intention that the person believe that one believes that statement to be true. Lying may be distinguished from other forms of intentional deception insofar as it involves the use of conventional signs arranged to make a statement. Intentional deception using natural signs, such as fake smiling, shamming a limp, or wearing a disguise, does not count as lying. Intentional deception using conventional signs that are neither spoken nor written, such as deceptively nodding

one's head, sending deceptive smoke signals, or deceptive signaling by semaphore, does count as lying, at least insofar as one is making a statement.

Lying requires that a statement be made; hence that form of deception that consists in withholding a statement from another person with the intention that the other person infer a believed falsehood—sometimes called a lie of omission or a concealment lie—does not count as lying. Exaggerating, being misleading, hedging, or being evasive, with the intention that the other person infer a believed falsehood, also does not count as lying. Lying does not require that the statement that is made is false, but it does require that the statement made is believed to be false rather than merely not believed to be true, or believed to be possibly false or probably false. Lying does not require that the other person is real, only that the other person is believed to be a person and is believed to be real. This does not resolve the questions of whether one can lie to no other person in particular (for example, by publishing a believed false account of an event), or whether there can be intrapersonal lying (for example, an earlier self lying to a later self).

The most important philosophical discussions of lying are to be found in St. Augustine, St. Thomas Aquinas, and Immanuel Kant. Aquinas differed from Augustine and Kant in holding that making a declarative statement to another person that one believes to be false is sufficient for lying; no further deceptive intention is needed. All three held that lying is wrong and that one should never lie; however they distinguished between not lying or being truthful, which is required, and being candid or volunteering believed truths, which is not. Augustine and Aquinas held that some lies, such as lies told to save the lives of innocents or lies told to avoid being defiled, that do not harm the particular person(s) lied to, are less egregious than other lies, such as malicious lies and lies told in the teaching of religion. All three argued that lying is a perversion of the faculty of speech, the natural end of which is the communication of thoughts. Augustine and Kant argued that in telling a lie one harms oneself, and undermines trust in society; hence there can never be a harmless lie. Kant also argued that a person cannot consent to being told a particular lie; hence in lying to another person one is necessarily treating that person as a mere means to one's end.

See also Deontological Ethics; Duty; Kantian Ethics; Moral Rules and Principles; Self-Deception; Virtue and Vice.

Bibliography

Adler, J. E. "Lying, Deceiving, or Falsely Implicating" *Journal of Philosophy* 94 (1997): 435–452.

Aquinas, St. Thomas. "Question 110: Lying." In *Summa Theologiae*, edited by Thomas Gilby. Garden City, NY: Image Books, 1969.

Augustine, St. "On Lying and Against Lying." In *Treatises on Various Subjects*, edited by R.J. Deferrari. New York: Fathers of the Church, 1952.

Bok, S. *Lying: Moral Choice in Public and Private Life*. 2nd ed. New York: Vintage Books, 1999.

Chisholm, R. M., and T. D. Feehan. "The Intent to Deceive." *Journal of Philosophy* 74 (1977): 143–159.

Frankfurt, H. G. "The Faintest Passion." In *Necessity, Volition, and Love*. Cambridge, U.K.; New York: Cambridge University Press, 1999.

Kant, I. *Immanuel Kant, Practical Philosophy*, edited by Mary Gregor and Allen W. Wood. Cambridge, U.K.: Cambridge University Press, 1996.

Mahon, J. E. "Kant on Lies, Candour, and Reticence." *Kantian Review* 7 (2003): 101–133.

Mannison, D. S. "Lying and Lies." *Australasian Journal of Philosophy* 47 (1969): 132–144.

Siegler, F. A. "Lying." *American Philosophical Quarterly* 3 (1966): 128–136.

James Edwin Mahon (2005)

LYOTARD, JEAN-FRANÇOIS
(1924–1998)

Born in Versailles, France, on August 10, 1924, Jean-François Lyotard was educated in Paris. As a child, Lyotard wanted to be a monk, painter, historian, or novelist, but settled a career in philosophy. He began teaching philosophy at the secondary school level in Constantine, Algeria, and later at La Flèche, France. From 1954 to 1966, Lyotard was a member of a leftist revolutionary group called Socialism ou Barbarie (either socialism or barbarism), eventually joining a splinter group called Pouvoir Ouvrier (Worker's Power) in 1964. He broke with the group in 1966 after becoming critical of Marxism's tendency toward universalism. He began work as a philosophy professor, and was employed at University of Paris X, Nanterre, during the student protests of May 1968. He gained a full position at the University of Paris VIII, Vincennes, where he spent many years and became an emeritus faculty member in 1987. He was also a founding member of the Collège International de Philosophie in Paris. With *The Postmodern Condition: A Report on Knowledge* (1979) he achieved international renown, and was guest lecturer at many universities throughout the world. On April 21, 1998, Lyotard died of leukemia in Paris. Lyotard's philosophical influences are diverse,

including research on topics in Marxism, psychoanalysis, aesthetics, continental and analytical philosophy. An overall theme throughout his works is the inability for a single theory to capture the whole of reality, typically stressing what has been left out or forgotten in a particular theory.

Lyotard's initial writings of the 1950s and early 1960s were political and focused on the Marxist concerns of *Socialism ou Barbarie*, with particular attention to the ending the French occupation of Algeria. Additionally, he published *La phénoménologie* (*Phenomenology*) that supports many aspects of phenomenology, but is critical of its tendency to prioritize the transcendental ego in isolation from the material concerns addressed in Marxism. After attending Jacques Lacan's lectures in the 1960s, Lyotard wrote his first major work, *Discours, figure* to complete his *doctorat d'etat*. Published in 1971, *Discours, figure* compares the approaches of structuralism and phenomenology by examining the relationship between textual words of reading, and the figural or visual image of seeing that resists signification and rational concepts. Lyotard argues that text and figure cannot be neatly separated from one another, and neither word nor image should be privileged. His next important work, *Libidinal Economy*, published in 1974, is strongly influenced by Freud, Marx, and Nietzsche, though Lyotard later recants his self-professed "evil book" (*Perigrinations*, 13). *Libidinal Economy* is a break from the rest of Lyotard's work because it retreats entirely from the intellectualism of rational concepts in favor of an examination of drives, affects, intensities, and energy flows that can be ordered in a variety of ways by society.

Lyotard attained fame with the publication of *The Postmodern Condition* in 1979, which was commissioned by the Quebec government to examine the status of knowledge in highly developed societies. The publication of this book catapulted Lyotard into the international spotlight. Often, Lyotard's use of the term "postmodernism" is misunderstood as a historical era following the modern period, though in *The Postmodern Condition* Lyotard insists that the postmodern occurs within the modern period as an "incredulity toward meta-narratives" (p. xxiv). For Lyotard, modernism relies upon meta-narratives that are overarching discourses that try to explain all phenomena according to their own terms.

Lyotard utilizes Ludwig Wittgenstein's terminology of "language games" during this period to suggest that different language games follow their own rules and cannot be adequately translated to one another. While scientific discourse is denotative, ethical discourse is prescriptive, and to translate the descriptive into the prescriptive would be analogous to translating the rules of chess into those of checkers. Universal grand narratives in modernity suppose that language games are indeed commensurable and result in a kind of "terror" that cannot accept other kinds of games. Lyotard questions the hierarchical priority of scientific and technological forms of knowledge in developed societies that exclude other types of knowledge. According to Lyotard, grand narratives cannot legitimate their authority, and the postmodern breaks through the modern when grand narratives lose their credibility. The epistemological questions raised in *The Postmodern Condition* turn toward political themes in *The Differend*.

Published in 1983, *The Differend: Phrases in Dispute* is thought to be Lyotard's most important work because of its elaboration of the central concept of the book, the "differend." Lyotard defines the *différend* as a "case of conflict, between (at least) two parties, that cannot be equitably resolved for lack of a rule of judgment applicable to both arguments" (p. xi). Lyotard uses the instance of proving the horror of the gas chambers at Auschwitz as his paradigmatic model of a *différend*. Revisionist historian Robert Faurisson denies that the Holocaust occurred because there are no victims who were eyewitnesses to the atrocity. In order for there to be an eyewitness, one would have to be a victim that survived the gas chambers, making it impossible to establish the crime according to Faurisson's criterion. This situation is used as a touchstone to examine various political scenarios in which the victim cannot establish the existence of an injustice, because his or her experience does not conform to present criterion for establishing a legitimate "injustice," and for that reason, the plaintiff becomes a victim of a further wrong. A *différend* follows the structure of a double bind, where it is impossible for the plaintiff to prove damage by the rules of current authority, and differs from litigation that can be established within the present rules. For Lyotard, the *différend* is signaled by a sublime feeling because it involves an overwhelming feeling of pleasure and a feeling of pain. The pain in the sublime comes from the inability to express the wrong of the *différend*, but the feeling of pleasure arises from the potential for the creation of new idioms of discourse that can express the wrong. Lyotard uses Kant's theory of aesthetical judgments of the sublime to describe a theory of political judgment where judgments are made without recourse to a universal rule. Because of the incommensurability of language genres, the *différend* cannot be eliminated for good, but one can bear witness to *différends* and even strain to hear their call.

Much of Lyotard's later work explores Kant's theory of the sublime in greater detail. Lyotard also published many important books of essays focusing on art, literature, history, technology, politics, and postmodernism, in addition to books on several other topics. According to Geoffrey Bennington (1988), Lyotard personally believed that his major works were *Discourse, figure*, *Libidinal Economy*, and *The Differend*.

See also Postmodernism.

Bibliography

WORKS BY LYOTARD

Discours, figure. Paris: Klincksieck, 1971.

Économie libidinale. Paris: Éditions de Minuit, 1974. Translated by Iain Hamilton Grant as *Libidinal Economy* (Bloomington: Indiana University Press, 1993).

La condition postmoderne: Rapport sur le savoir. Paris: Éditions de Minuit, 1979. Translated by Geoff Bennington and Brian Massumi as *The Postmodern Condition: A Report on Knowledge* (Minneapolis: University of Minnesota Press, 1984).

Au Juste. With Jean-Loup Thébaud. Paris: Christian Bourgois, 1979. Translated by Wlad Godzich as *Just Gaming* (Minneapolis: University of Minnesota Press, 1985).

Le différend. Paris: Éditions de Minuit, 1983. Translated by Georges Van Den Abbeele as *The Differend: Phrases in Dispute* (Minneapolis: University of Minnesota Press, 1988).

The Lyotard Reader, edited by Andrew Benjamin. Oxford, U.K.: Blackwell, 1989.

L'inhuman: Causeries sur le temps. Paris: Éditions Galilée, 1988. Translated by Geoffrey Bennington and Rachel Bowlby as *The Inhuman: Reflections on Time* (Stanford, CA: Stanford University Press, 1991).

Leçons sur l'analytique du sublime. Paris: Galilée, 1991. Translated by Elizabeth Rottenberg as *Lessons on the Analytic of the Sublime* (Stanford, CA: Stanford University Press, 1994).

WORKS ABOUT LYOTARD

Bennington, Geoffrey. *Lyotard: Writing the Event*. New York: Columbia University Press, 1988.

Readings, Bill. *Introducing Lyotard: Art and Politics*. London: Routledge, 1991.

Williams, James. *Lyotard and the Political*. London: Routledge, 2000.

Karin Fry (2005)

M

MACH, ERNST

(1838–1916)

Mach, Ernst, Austrian physicist and philosopher, was born at Turas near Brno, Moravia (now in the Czech Republic). As with many great figures, a profound psychological experience in youth had lasting effect. Mach describes it in *The Analysis of Sensations*:

> I have always felt it as a stroke of special good fortune that early in life, at about the age of fifteen, I lighted, in the library of my father, on a copy of Kant's *Prolegomena to Any Future Metaphysics*. The book made at the time a powerful and ineffaceable impression upon me, the like of which I never afterwards experienced in any of my philosophical reading. Some two or three years later the superfluity of the role played by "the thing in itself" abruptly dawned on me. On a bright summer day in the open air, the world with my ego suddenly appeared to me as *one* coherent mass of sensations, only more strongly coherent in the ego. Although the actual working out of this thought did not occur until a later period, yet this moment was decisive for my whole view.

Examination of Mach's life and work confirms this statement. Fired by the stimulus, he studied in Vienna and became professor of mathematics at Graz in 1864. In 1867 he took a chair of physics at Prague and in 1895 became professor of the history and theory of inductive science at Vienna. In 1901, he was appointed to the upper house of the Austrian parliament. His interests were extraordinarily wide: In physics he made contributions to acoustics, electricity, hydrodynamics, mechanics, optics, and thermodynamics, and in psychology to perception and aesthetics. William James, who met Mach in 1882, reported that he appeared to have read and thought about everything. At the start of the twentieth century, he and Henri Poincaré were the two outstanding popularizers of science in the world. Lenin's main philosophical work is an onslaught on Machian thought, which was highly regarded by Russian socialists who opposed Lenin. Albert Einstein's 1916 obituary of Mach includes this comment: "His direct joy in seeing and comprehending, Spinoza's amor dei intellectualis, was so overwhelming that in high old age he still stared at the world with the inquisitive eyes of a child in order to take simple delight in understanding the connection of things." On another occasion, Einstein (1949) praised Mach's "incorruptible skepticism."

MACH'S INFLUENCE

Mach gave his name to three things in science. A crude but revealing measure of his enduring significance is given by the number of Internet entries listed by Google, at the time of writing, for each of them: the Mach number (41,400), Mach's principle (2,820), and Mach bands (1,580). For comparison, the uncertainty principle of Heisenberg has 56,500 entries. Under Ernst Mach, one finds 92,100 entries. David Hume has 249,000, and Einstein 1,070,000.

Mach has been described as a superb experimentalist but unusual theorist. The Mach number is named after him because he was the discoverer of shock waves, which he observed directly in a brilliant early use of flash photography. He explained the sonic bang first heard in the Franco-Prussian war of 1870. For this outstanding work he was twice nominated for the Noble Prize near the end of his life. However, this was at a time when discoveries were flooding in, and he never received the prize he undoubtedly deserved. For his many other experimental researches—including the discovery of Mach bands in psychology—the reader is referred to Blackmore's biography cited at the end of this entry. This article is about his influence on philosophy of science and, more significantly, natural philosophy in the great tradition of the seventeenth century.

Mach's vivid holistic experience in youth became the unifying core of his *The Science of Mechanics: A Critical and Historical Account of Its Development*. Published in 1883 and widely read ever since, it argues fiercely for the primacy of empirical facts and the need to understand the contingent historical nature of progress in science. Mach was strongly antimetaphysical and questioned the foundations of all knowledge. Physical concepts are not immutable and should always be based on universally observed connections within phenomena. Newton had given a circular definition of mass; Mach replaced it with an operational definition based on the observed accelerations that interacting bodies impart to each other. Einstein recognized the key importance of Mach's approach in his own celebrated operational definition of simultaneity in the special theory of relativity in 1905.

Perhaps even more important than this influence was Mach's intense distrust of the invisible rigid structure of absolute space and time that Newton had introduced in his *Principia* in 1687 in order to formulate his first law motion. Now known as the law of inertia, it states that every body continues in a state of rest or uniform motion in a straight line unless acted upon by external forces. Absolute space was widely attacked as a dubious concept in Newton's time, above all by Gottfried Wilhelm Leibniz and George Berkeley. However, Mach was the first person to offer a plausible alternative to the framework that Newton had introduced on the basis of rather strong empirical evidence. Mach argued that the locally observable inertial motion of force-free bodies could in reality be "guided" by the integrated physical effect of the totality of matter in the universe rather than by absolute space. Einstein dubbed this idea *Mach's principle*. It was undoubtedly the greatest single stimulus that led to the creation of his general theory of relativity in 1915. Ironically, the actual status of Mach's principle within general relativity is still controversial, although the present writer believes that the theory is almost perfectly Machian when correctly understood.

Mach also had an influence, though far less decisive, on the discovery of quantum mechanics. By the early 1920s, many physicists had come to despair of ever finding a description of atomic phenomena within the traditional framework of space and time. Strongly influenced by Mach's contention that science should solely concern itself with connections between directly observable phenomena, and impressed by Einstein's "Machian" successes, the youthful Werner Heisenberg embarked on a radical approach. The single-sentence abstract of his 1925 paper in which he created quantum mechanics in a matrix representation reveals the depth of Mach's influence: "This paper," Heisenberg wrote, "attempts to create foundations for a quantum-theoretical mechanics that is based exclusively on connections between quantities that are in principle observable." (Heisenberg 1925, p. 879) Mach also had an influence on the formulation of the so-called Copenhagen interpretation of quantum mechanics by Heisenberg and Niels Bohr in 1926 and 1927. In a decidedly Machian manner, they argued that it was the job of science to establish correlations between phenomena and not to attempt a direct description of "reality."

MACH IN THE TWENTIETH CENTURY

Although Mach's ideas manifestly played a strongly positive role in the great discoveries of twentieth-century physics, his actual philosophy of science has had a mixed and generally negative reception. There is no doubt that he underestimated the value of pure theoretical speculation in scientific discoveries, especially in physics. There are many important discoveries that clearly could never have been made had theoreticians stuck rigidly to Mach's precept that the role of science is solely to establish directly the immediate connection of phenomena. They include general relativity, Erwin Schrödinger's wave-

mechanical formulation of quantum mechanics, and the modern theory of gauge interactions. Many working scientists now accept Karl Raimund Popper's contention that in physics at least significant progress is often made through a bold conjecture that can in no way be justified by direct experience. Instead, the theoretician relies on intuition and accumulated experience to create a conceptual framework from which conclusions are drawn deductively and then tested against observation. In this approach, which is alien to Mach's philosophy, theories are always tentative and liable to empirical refutation.

The weakness of Mach's approach can probably be attributed to two main factors. First, his youthful epiphany made him an idealist rather like Berkeley. The extent to which Mach claimed ontological primacy for direct sense perceptions comes out startlingly in the opening chapter of *The Analysis of Sensations*. The difficulty with such an approach, which does have intellectual coherence, is that it has hitherto proved impossible to go beyond purely qualitative statements. The interconnection of directly experienced phenomena is notoriously difficult to grasp, as is the nature of the phenomena themselves. The second factor is the age in which Mach lived and worked. Theories based on invisible mechanically operating microscopic constituents of matter and substances such as phlogiston and caloric had indeed had a dismal track record more or less up to Mach's time. However, Newton had already given striking examples of rigorous, mathematically based use of hypotheses and deduction, and in Mach's time theoreticians had considerably refined in their art. The twentieth century saw their skill increase still further with spectacular effect. In contrast, it is characteristic that Mach's desire to "see connections" led him to make the famous flash photographs of shock waves for which he so nearly won the Nobel Prize. This was the greatest direct triumph of his approach to science.

The article by Peter Alexander in the previous edition of this encyclopedia, with twice the length of this entry, goes into much more detail about the various aspects of Mach's philosophy of science. The present writer therefore felt it would be useful to concentrate on Mach's great influence in natural philosophy. Within the narrower confines of philosophy of science, Mach was described by Philipp Frank in his *Modern Science and Its Philosophy* as one of the "spiritual ancestors … and real master of the Vienna Circle." The Vienna Circle was influential. Mach was also an important inspiration for the operationalism of Percy W. Bridgman.

See also Berkeley, George; Bohr, Niels; Bridgman, Percy William; Einstein, Albert; Energy; Force; Heisenberg, Werner; Hume, David; James, William; Laws, Scientific; Leibniz, Gottfried Wilhelm; Lenin, Vladimir Il'ich; Logical Positivism; Mass; Motion; Newton, Isaac; Phenomenology; Poincaré, Jules Henri; Popper, Karl Raimund; Quantum Mechanics; Relativity Theory; Schrödinger, Erwin; Sensationalism; Space; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY MACH

Mach wrote numerous books, and it is a mark of his impact that several are still in print in English translations. Lack of space precludes a detailed bibliography, which can be found at the end of the article by Peter Alexander. Mach's best known work is *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt* (1883), translated as *The Science of Mechanics* by T. J. McCormack (LaSalle, IL: Open Court, 1960). Among his more physical writings, one can certainly recommend his *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Energie* (1872), translated by P. E. B. Jourdain as *History and Root of the Principle of the Conservation of Energy* (Chicago: Open Court, 1911) and the *Populärwissenschaftliche Vorlesungen*, translated by T. J. McCormack as *Popular Scientific Lectures* (Chicago: Open Court, 1894), which includes a beautiful account of his work on shock waves. Also interesting but of uneven standards are his *Space and Geometry* (Chicago: Open Court, 1894) and *Die Prinzipien der Wärmelehre*, the last of Mach's major books to be translated (*Principles of the Theory of Heat*. Dordrecht, Netherlands: D. Reidel, 1986). His most important book laying out his philosophy is undoubtedly *Die Analyse der Empfindungen* (1906) (*The Analysis of Sensations*, available from Dover Publications, 1959) and there is also *Erkenntnis und Irrtum* (1905) (*Knowledge and Error*. Dordrecht, Netherlands: D. Reidel, 1976). There is a valuable exhaustive list of Mach's scientific papers and books (and much secondary literature in German) in Joachim Thiele's "Ernst Mach-Bibliographie" published in *Centaurus* 8 (1963): 189–237.

WORKS ON MACH

Einstein's obituary of Mach appeared in the *Physikalische Zeitschrift*, Volume 17, No. 7, pp. 101–104, 1919. His comment about Mach's incorruptible skepticism appears in his "Autobiographical Notes" in *Albert Einstein: Philosopher-Scientist*, edited by P. Schilpp, New York: Harper and Row (1949), p. 1. Heisenberg's article that created the matrix formulation of quantum mechanics is: "Über quantentheoretische Umdeuting kinematischer und mechanischer Beziehungen," *Zeitschrift für Physik*, Vol. 33, No. 12, 879 (1925). Philipp Frank made his comment about Mach and the Vienna Circle in his book *Modern Science and Its Philosophy*, Harvard: Harvard University Press, 1950. The English-language secondary literature is extensive. Blackmore's biography *Ernst Mach: His Life, Work, and Influence* (Berkeley and Los Angeles: University of California Press, 1972) is a mine of information but uneven in the

discussion of his philosophy. Alexander's article is another useful guide to earlier literature, but in this modern age the scholar who really wishes to make an in-depth study of the literature is probably best advised to trawl the Internet. In 1988, the Charles University in Prague organized an excellent conference to mark the 150th anniversary of Mach's birth. The conference papers *Ernst Mach and the Development of Physics* (Prague: Karolinum, 1991) are a useful compendium but probably difficult to obtain. A special conference *Mach's Principle: From Newton's Bucket to Quantum Gravity* was held at Tübingen, Germany in 1993. The proceedings, edited by Julian Barbour and Herbert Pfister, were published in 1995 by Birkhäuser (Boston) and include the present writer's article arguing that general relativity is Machian and includes much other material by physicists, historians, and philosophers.

Julian Barbour (2005)

MACHIAVELLI, NICCOLÒ

(1469–1527)

Niccolò Machiavelli, the Italian politician and political thinker, is famous for his treatise on princeship titled *The Prince* (*Il principe*) and for a discussion of how to establish a good republican government, *The Discourses* (*Discorsi sopra la prima deca di Tito Livio*). Machiavelli also wrote poems and comedies (including the *Mandragola*), a *History of Florence,* and a book titled *Art of War.* They contain many original ideas and were widely read, but today these writings arouse interest mainly because their author was the man who, with *The Prince* and *The Discourses,* inaugurated a new stage in the development of political thought.

When Machiavelli wrote *The Prince* and *The Discourses,* he was aware that he was saying things about politics that had not been expressed before; in the introduction to *The Discourses* he stated that he was resolved "to open a new route which has not yet been followed by anyone." Nevertheless, Machiavelli would not have claimed to be a systematic political philosopher. *The Prince* was written in 1512–1513; the date of *The Discourses* is less certain, but it was certainly completed by 1517. Machiavelli was then in his forties and, in the preceding years of his life, he had been a practical politician who had never shown interest in becoming a political writer or in embarking on a literary career.

In 1498, after the expulsion of the Medici from Florence and the fall of Girolamo Savonarola, Machiavelli had entered the Florentine chancellery, where his special function was to serve as the secretary of The Ten, a group of magistrates charged with the conduct of diplomatic negotiations and the supervision of military operations in wartime. In this position Machiavelli carried out a number of diplomatic missions in Italy, France, and Germany. His ability attracted the attention of Gonfalonier Piero Soderini, the official head of the Florentine government, and Machiavelli became Soderini's confidant—his "lackey," according to Soderini's enemies. Machiavelli's close relationship with Soderini became a serious handicap when, in 1512, the republican regime was overthrown and the Medici returned to Florence. Other members of the chancellery were permitted to continue in office, but Machiavelli was dismissed and forced to withdraw to a small estate near Florence, where he lived in straitened economic circumstances.

It was at this time that Machiavelli turned to literary work in the hope that through his writings he would gain the favor of influential men who might help him to regain a position in the Florentine government. *The Prince* was dedicated to Lorenzo de' Medici, a nephew of Pope Leo X and the actual ruler of Florence. *The Discourses* was dedicated to members of the Florentine ruling group, and his *History of Florence* was written at the suggestion of Cardinal Giulio de' Medici, who in 1523 became Pope Clement VII. In the 1520s Machiavelli's efforts began to bear fruit. Clement VII entrusted him with a number of minor political commissions, and Machiavelli devoted himself to this kind of work, relegating the completion of his literary projects to the background. However, in 1527, before Machiavelli had been firmly reestablished in a political position—actually, at a moment when his future had again become uncertain because the Medici had once more been driven from Florence—he died.

Thus, Machiavelli's attitude in composing *The Prince* and *The Discourses* was not that of a disinterested scholar; his aims were practical and personal. He wanted to give advice that would prove his political usefulness, and he wanted to impress those who read his treatises. Therefore, Machiavelli was inclined to make numerous startling statements and extreme formulations. A characteristic example is his saying that the prince "must abstain from taking the property of others, for men forget more easily the death of their father than the loss of their patrimony" (*The Prince,* Ch. 17).

ARTS OF WAR

Machiavelli's statements were startling not only because of their form of presentation but also because of their content. One aspect of political affairs with which Machiavelli had been particularly concerned and in which he was especially interested was the conduct of military

affairs. He thought deeply about the reasons why the French had so easily triumphed over the Italians in 1494 and had marched from the north to the south of Italy without meeting serious resistance. Machiavelli's explanation was that the governments of the various Italian states, whether they were republican regimes or principalities, had used mercenary soldiers led by hired *condottieri.* He therefore recommended that in case of war the prince should lead his troops himself and that his army should be composed of his own men; that is, the Italian governments should introduce conscription. Moreover, Machiavelli polemicized against other favorite notions of his time on military affairs; for instance, he denied that artillery was decisive in battle or that fortresses could offer a strong defense against an invading army.

MORALS AND POLITICS

Machiavelli's rejection of traditional political ideas emerged most clearly in his discussions of the relation between morals and politics. The most revolutionary statements on these issues are found in chapters 15–19 of *The Prince,* which deal with the qualities a prince ought to possess. In the Mirror of Princes literature of the ancient world and of the Middle Ages, a prince was supposed to be the embodiment of human virtues; he was expected to be just, magnanimous, merciful, and faithful to his obligations, and to do everything that might make him loved by his subjects. Machiavelli objected to such demands. According to him, a prince "must not mind incurring the scandal of those vices without which it would be difficult to save the state, and if one considers well, it will be found that some things which seem virtues would, if followed, lead to one's ruin and that some others which appear vices result in one's greater security and well-being." This sentence and chapters 15–19 have frequently been understood as meaning that instead of being mild a prince ought to be cruel; instead of being loyal, treacherous; instead of aiming to be loved, he should aim to be feared. But this is a misunderstanding. A closer reading shows that Machiavelli admonishes a prince to disregard the question whether his actions would be called virtuous or vicious. A ruler ought to do whatever is appropriate to the situation in which he finds himself and may lead most quickly and efficiently to success. Sometimes cruelty, sometimes leniency, sometimes loyalty, sometimes villainy might be the right course. The choice depends on circumstances. To illustrate his point of view Machiavelli used as an example the career of Cesare Borgia, which he outlined in chapter 7 of *The Prince.*

Machiavelli's views have frequently been interpreted as meaning that wickedness is more effective than goodness. This distortion of his views has been regarded as the essence of Machiavelli's teaching, as identical with what later centuries called Machiavellism. It should be stated that Machiavelli was not concerned with good or evil; he was concerned only with political efficiency. His rejection of the *communis opinio*—whether in the special area of military affairs or in the general field of ethics—was a reflection of a new and comprehensive vision of politics. Before Machiavelli, the prevailing view had been that the task of government was distribution and maintenance of justice. Machiavelli believed that the law of life under which every political organization existed was growth and expansion. Thus, force was an integral, and a most essential, element in politics.

Machiavelli's interest in military affairs had its basis in his conviction that possession of a powerful and disciplined military force was a requisite for the preservation of political independence. Moreover, because political life was a struggle, the conduct of life according to Christian virtues could endanger political effectiveness; Christianity, by preaching meekness and selflessness, might soften men and weaken a political society. Machiavelli directed some very strong passages against the effeminacy to which Christianity had led. Political man needed not virtues but *virtù,* "vitality." The possession of *virtù* was the quality most necessary for a political leader, but according to Machiavelli both individuals and entire social bodies could and should possess *virtù.* That is why, in *The Prince,* Machiavelli could write a "handbook for tyrants," while in *The Discourses* he could advocate a free republican regime. Every well-organized, effective political organization must be permeated by one and the same spirit and must form an organic unit. There are few if any passages in Machiavelli in which he uses the word *state* (*stato*) in the modern sense of an organic unit embracing individuals and institutions. However, there can be no doubt that his concept of an organized society producing *virtù* among its members comes very close to the modern concept of state.

METHOD OF ARGUMENT

The new vision of the character of politics required a new method of political argumentation. Rules for the conduct of politics could not be formulated on the basis of theoretical or philosophical assumptions about the nature of a good society; successful political behavior could be learned only through experience. Machiavelli stated in his dedication of *The Prince* that he wanted to tell others

what he had "acquired through a long experience of modern events and a constant study of the past." Thus, experience was not limited to those events in which a person participated but embraced the entire field of history. To Machiavelli the most instructive period of the past was that of republican Rome. Machiavelli thought that, because the Romans succeeded in extending their power over the entire world, no better guide for the conduct of policy could be imagined than that of Roman history. It is indeed true that previous writers on politics, particularly the humanists, had used historical examples, but to rely exclusively on historical experience in establishing political laws was an innovation; Machiavelli's writings implied that every true political science ought to be based on history.

It has been said that, in rejecting the validity of the doctrines of theology and moral philosophy for the conduct of politics, Machiavelli established politics as an autonomous field. He could do so because he regarded political bodies not as creations of human reason but as natural phenomena. In Machiavelli's opinion all political organizations, like animals, plants, and human beings, are subject to the laws of nature. They are born, they grow to maturity, they become old, and they die. Well-organized political bodies might live longer than others, but even the best-constructed political society, even Rome, could not escape decline and death. This view of the instability and impermanence of all things gives Machiavelli's recommendations their particular tenor. Men or political bodies are entitled to use all possible means and weapons because the moments when they can flourish and triumph are brief and fleeting. Despite Machiavelli's claim that political success depended on acting according to the political laws he established in his writings, he was always conscious of the role of accident and fortune in human affairs.

INFLUENCE

It is of some importance to distinguish between the shocking novelty of Machiavelli's particular recommendations and his general concepts of politics, from which his practical counsels arose. Such a distinction helps to explain the contradictory reception his ideas found in the following centuries. Machiavelli's writings soon became known in Italy and then in other European countries, particularly France and England, although in 1559 his works were placed on the Index. Generally he was considered an adviser of cruel tyrants, an advocate of evil; Cardinal Reginald Pole said that Machiavelli wrote "with the finger of the Devil." Although nobody in the sixteenth century dared publicly to express anything but abhorrence, a school of political writers arose in Italy who explained that the criteria of a statesman's or ruler's actions were the interests of the state. These advocates of the doctrine of "reason of state"—even if they did not acknowledge their obligations to Machiavelli—followed the course Machiavelli had charted. The Enlightenment, with its belief in the harmony of morality and progress, could only condemn Machiavelli's view that political necessity permitted the neglect of ethical norms. An example is the *Anti-Machiavel* that Frederick II of Prussia composed as a young man. Some eighteenth-century thinkers, however, recognized truth in Machiavelli's approach to politics. For instance, Gabriel Bonnot de Mably and Jean-Jacques Rousseau admired Machiavelli because he had realized that the strength of a political organization depends on the existence of a collective spirit that is more than a summation of individual wills.

In the nineteenth century, students of Machiavelli, following the interpretation that the German historian Leopold von Ranke had given, did not believe that Machiavelli had wanted to separate ethics and politics. Because the last chapter of *The Prince* contains an appeal for the liberation of Italy from the barbarians, they assumed that Machiavelli had permitted the violation of moral rules only for the purpose of a higher ethical goal; that his purpose had been to point the way toward the foundation of a unified Italy. Thus, in the nineteenth century Machiavelli became respectable as the prophet of the idea of the national state. In the later part of the century Machiavelli was also referred to by those who wanted to free man from the oppressive shackles of traditional morality and believed that man's faculties could be fully developed only if he placed himself "beyond good and evil." Friedrich Nietzsche's superman was supposed to have "virtue in the style of the Renaissance, *virtù,* virtue free from morality."

See also Enlightenment; Nietzsche, Friedrich; Peace, War, and Philosophy; Political Philosophy, History of; Religion and Politics; Social and Political Philosophy; Rousseau, Jean-Jacques.

Bibliography

The literature on Machiavelli is very extensive. A more recent critical edition of his works is that edited by Sergio Bertelli and Franco Gaeta and published by Feltrinelli in its Biblioteca di classici italiani. So far four volumes containing Machiavelli's literary works and three volumes containing his *Legazioni e commissarie* have appeared (1960–1964). This edition provides a critical discussion of the Machiavelli literature. The best recent translation is Allan Gilbert, *Chief*

Works, and Others, 3 vols. (Durham, NC: Duke University Press, 1965).

Older biographies have become obsolete since the appearance of Roberto Ridolfi's *Vita di Niccolò Machiavelli* (Rome, 1954), translated by Cecil Grayson as *The Life of Niccolò Machiavelli* (Chicago: University of Chicago Press, 1963). Machiavelli's intellectual development is well analyzed by Gennaro Sasso in his *Niccolò Machiavelli: Storia del suo pensiero politico* (Naples: Nella sede dell'Istituto, 1958). For the relation of Machiavelli's thought to that of his contemporaries, see Felix Gilbert, *Machiavelli and Guicciardini* (Princeton, NJ: Princeton University Press, 1965). The main lines of the influence of Machiavelli's ideas on the political thought of later centuries are traced in Friedrich Meinecke, *Die Idee der Staatsräson in der neueren Geschichte* (Berlin: R. Oldenbourg, 1924), translated by Douglas Scott as *Machiavellism* (New Haven, CT: Yale University Press, 1957). For Machiavelli's impact on English political thought, see Felix Raab, *The English Face of Machiavelli: A Changing Interpretation 1500–1700* (London: Routledge and Paul, 1964).

Felix Gilbert (1967)

MACHIAVELLI, NICCOLÒ [ADDENDUM]

Many readers come to Machiavelli with their minds made up about who he was and what he espoused. A more balanced assessment must take into account many approaches to his work and possible influences from the classical world.

EVALUATING MACHIAVELLI

In order to evaluate Machiavelli one must first decide what he was doing and second decide how to balance the assessment of the texts. The traditional assessment of Machiavelli is "expedient egoist." Under this reading one would cite passages from *The Prince* in which rulers are advised to employ deceit and cruelty for the sake of political advantage. However, it is unclear whether these passages should be taken at face value or rather as an invective set in its European historical perspective: (a) a striving to connect to the past—particularly to the Roman Empire and its eloquent Republican spokesman, Cicero; and (b) a chafing with the Papal authority over the legacy of the Roman Empire—especially the bogus "Donation of Constantine." In this forged document the Roman Emperor Constantine supposedly granted the whole of the Roman Empire to the pope who, in turn, allowed the daily duties of running the secular to fall upon the emperor. This document sought to establish a legal claim for the pope's universal secular power. It could be that Machiavelli, in the first case, was interested in espousing the republican message of Cicero. It could also be, in the second case that Machiavelli was consciously breaking away from established forms of exposition in order to create another mode of political discourse.

In recent scholarship (over the second half of the twentieth century) Ernst Cassirer (1946) believed that Machiavelli espoused a clear and coherent argument based upon a vision that moved the modern world forward in a realistic fashion. Isaiah Berlin (1972 [1953]\) followed in asserting that Machiavelli put forth a cogent secular vision that was consistent. Leo Strauss (1958) agreed that the vision was consistent, but said that both from the points of view in *The Prince* and of *The Discourses on Livy* that Machiavelli was a teacher of evil (namely, an expedient egoist).

Certainly, the worldviews presented in *The Prince* and *The Discourses on Livy* appear both different and the same. They are *different* in that in the former case there seem to be many aphorisms that violate ethical laws whereas in the latter it seems that Machiavelli is concerned to uphold public morality—such as eliminating public corruption for the sake of the republic. One might reasonably ask whether the same person wrote both works.

However, they are *similar* in that they are both pragmatically oriented toward solving problems. Thus, we are faced with one interpretative option of which work should be seen as representing the author's "true vision"? Because of the caveats mentioned above, (a) and (b), some of the so-called "Cambridge School" (Pocock, Skinner, and Viroli) have accentuated the emphasis upon the rule of law, common good, and general republicanism as seen from the *Discourses* as evidence that Machiavelli was really a forward-thinking republican thinker.

If [this]\ reading of Machiavelli is correct, then he is a thinker who is not an advocate of expedient egoism, but rather is a thinker who saw various dead ends in the way political philosophy was being explored. To start anew he tries to jettison the views of the reigning paradigm and start afresh. This is an interesting interpretation, but it has one possible flaw: Machiavelli does not spend time on theoretical foundations. Any theory asserted to be present there must be read into the text. And so what theory might support his pragmatic observations?

Two candidates are Aristotle and Cicero. At this period of history, Aristotle's *Politics* and *Ethics* had recently been translated into Latin. Cicero had been the established authority (because of his association with

Rome and because of the rhetorical structure of the *Discourses*) and so most commentators seem to think that he is the dominant influence. Another argument along this line is that Aristotle thought that civic virtue derived from man fulfilling his nature in the context of society (so that politics flows from ethics). Machiavelli does not employ such an explanatory framework. Can his texts be read from this perspective? Perhaps, but it may be a stretch because the practicality of Machiavelli works against Aristotle's essentialism.

The candidate left standing is Cicero. If this is the case, then Cicero's presence best describes the Muse of Machiavelli. The structure of the *Discourses* seems to suggest this as it follows the classical rhetorical form: (for example, observe the titles of the first three chapters: (1) What Have been Universally the Beginnings of Any City Whatever, and What was That of Rome; (2) Of How Many Species Are Republics, and Which Was the Roman Republic; (3) What Accidents Made the Tribunes of the Plebs Be Created in Rome, Which Made the Republic More Perfect).

However, once this is accepted, then other results may follow. The postmodernists assert that literary constructions can substitute for a traditional exposition of the pursuit of a universal Truth (as per Aristotle). Instead, the use of Ciceronian rhetoric might resonate with the etymology of "rhetoric" à la speaking in public or engaging in discourse with others. If this understanding is correct, then the philosophy of discourse and *construction from discourse* as per Foucault, Ricoeur, Derrida, or Habermas might be more apropos than the use of rhetoric as a means of transmitting already settled truths.

Such an interpretation may have many advantages. First, it might resolve the contradictions between the various texts of Machiavelli. This is because contradictions are only a problem if one is creating a systematic work of philosophy, such as was aspired to by Aristotle or Thomas. Second, it might blunt the traditional "bad boy" image of Machiavelli by bringing him into the realm of merely revealing various approaches to the questions of what policies might be necessary for running a state. By bringing out various options and interacting with them, Machiavelli might be eschewing the conventional method of discourse (even though he employs traditional forms) in favor of creating a new realpolitik.

This reinvigorated conception would find its sources in the way politics are actually practiced. So, for example, *The Prince* might be seen not as a way things *ought to be*, but a description of the way things *are*. If we are to go anywhere, here is the starting point. Let us all accept this. And in the *Discourses* if Rome is a model of a civilization that worked well for a long time, then the focus should be upon what can be done to correct the flaws that brought it down. Under this sort of reading, the exploration of politics is not about creating treatises on political theory, but instead of initiating a dialogue among readers about the "deal points" in running a state. The completion of the text lies in the audience.

See also Aristotle; Berlin, Isaiah; Cassirer, Ernst; Cicero, Marcus Tullius; Derrida, Jacques; Foucault, Michel; Habermas, Jürgen; Political Philosophy, History of; Ricoeur, Paul; Social and Political Philosophy; Thomas Aquinas, St.

Bibliography

Bertman, Martin. "Justice with Particular Reference to Hobbes." *Kriterion* 42 (103) (2001): 58–70.

Blattberg, Charles. *From Pluralist to Patriotic Politics: Putting Practice First*. New York: Oxford University Press, 2000.

Cassirer, Ernst. *The Myth of the State*. New Haven, CT: Yale University Press, 1946.

Matthes, Melissa M. *The Rape of Lucratia and the Founding of Republics: Readings in Livy, Machiavelli, and Rousseau*. University Park: Pennsylvania State University Press, 2000.

WORKS BY MACHIAVELLI

Opere Complete, edited by S. Bertelli and F. Gaeta. Milan: Feltrinelli, 1960–1965.

The Prince. Translated by Harvey C. Mansfield, Jr. Chicago: University of Chicago Press, 1985.

Florentine Histories. Translated by Laura F. Banfield and Harvey C. Mansfield, Jr. Princeton, NJ: Princeton University Press, 1988.

Machiavelli: The Chief Works and Others, edited and translated by A. Gilbert. Durham, NC: Duke University Press, 1989.

Tutte. Le Opere. Edited by Mario Martelli. Florence: Sansoni, 1992.

Discourses on Livy. Translated by Harvey C. Mansfield, Jr. and Nathan Tarcov. Chicago: University of Chicago Press, 1996.

WORKS ON MACHIAVELLI

Ascoli, Albert Russell, and Victoria Kahn. "Introduction." In *Machiavelli and the Discourse of Literature*, edited by Albert Russell Ascoli and Victoria Kahn, 1–15. Ithaca, NY: Cornell University Press, 1993.

Berlin, Isaiah. "The Originality of Machiavelli." 1953. In *Studies on Machiavelli*, edited by Myron P. Gilmore. Florence: Sansoni, 1972.

Falco, Maria J., ed. *Feminist Interpretations of Niccolo Machiavelli*. University Park: Pennsylvania State University Press, 2004.

Femia, Joseph. "Machiavelli and Italian Fascism." *History of Political Thought* 25 (1) (2004): 1–15.

Femia, Joseph. *Machiavelli Revisited*. Cardiff: University of Wales Press, 2004.

Fiore, Silvia Ruffo. *Niccolo Machiavelli: An Annotated Bibliography of Modern Criticism and Scholarship*. New York: Greenwood Press, 1990.
Fischer, Markus. *Well-Ordered License: On the Unity of Machiavelli's Thought*. Lanham, MD: Lexington Books, 2000.
Flyvbjerg, Bent. *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*. Translated by Steven Sampson. Cambridge, U.K.: Cambridge University Press, 2001.
Gilbert, Felix. *Machiavelli and Guicciardini: Politics and History in Sixteenth-Century Florence*. New York: Norton, 1984.
Gilbert, Felix. "On Machiavelli's Idea of *virtù*." *Renaissance News* 4 (1951): 53–56.
Griffiths, Paul E. "Basic Emotions, Complex Emotions, Machiavellian Emotions." *Philosophy* 52 (2003): 39–67.
Hulliung, Mark. *Citizen Machiavelli*. Princeton, NJ: Princeton University Press, 1983.
Kahn, Victoria. "Habermas, Machiavelli, and the Humanist Critique of Ideology." *PMLA* 105 (1990): 464–476.
Kahn, Victoria. *Machiavellian Rhetoric: From the Counter-Reformation to Milton*. Princeton, NJ: Princeton University Press, 1994.
Knoll, Manuel. "Die Konservative Verantwortungsethik des Humanisten Niccolo Machiavelli." *Politisches Denken* (2003): 94–116.
Lang, Andre. "La Dialectique de la Fortune et de la Virtu chez Machiavel." *Diotima: Review of Philosophical Research*. 31 (2003): 179–188.
McCormick, John P. "Machiavelli against Republicanism: On the Cambridge School's 'Guicciardian Moments.'" *Political Theory* 31 (5) (2003): 615–643.
Montgomery-Blair, Brook. "Post-Metaphysical and Radical Humanist Thought in the Writings of Machiavelli and Nietzsche." *History of European Ideas* 27 (3) (2001): 199–238.
Nederman, Cary J. "Machiavelli and Moral Character: Principality, Republic and the Psychology of 'Virtu.'" *History of Political Thought* 21 (3) (2000): 349–364.
Pocock, J. G. A. *The Machiavellian Moment: Florentine Political Thought and the Antlantic Republican Tradition*. Princeton, NJ: Princeton University Press, 1975.
Skinner, Quentin. *Machiavelli*. New York: Hill and Wang, 1981.
Skinner, Quentin. "The Republican Ideal of Political Liberty." In *Machiavelli and Republicanism*, edited by Gisela Bock, Quentin Skinner, and Maurizio Viroli, 293–309. Cambridge, U.K.: Cambridge University Press, 1990.
Spackman, Barbara. "Politics on the Warpath: Machiavelli's *Art of War*." In *Machiavelli and the Discourse of Literature*, edited by Albert Russell Ascoli and Victoria Kahn, 179–194. Ithaca, NY: Cornell University Press, 1993.
Strauss, Leo. *Thoughts on Machiavelli*. Chicago: University of Chicago Press, 1958.
Tarlton, Charles. "'Azioni in modo l'una dall'altra': Action for Action's Sake in Machiavelli's The Prince." *History of European Ideas* 29 (2) (2003): 123–140.
Vatter, Miguel E. *Between Form and Event: Machiavelli's Theory of Political Freedom*. Dordrecht, Netherlands: Kluwer, 2000.
Viroli, Maurizio. *Machiavelli*. Oxford: Oxford University Press, 1998.

Michael Boylan (2005)

MACHINE INTELLIGENCE

Computers beat the best human chess players. Computers guide spacecraft over vast distances and direct robotic devices to explore faraway astronomical bodies. Computers outpace humans in many respects, but are they actually intelligent? Can they think? Even if one is skeptical about the mentality of today's computers, the interesting philosophical issue remains: Might computers possess significant intelligence someday? Indeed, might computers feel or even have consciousness? And, how would we know?

THE HISTORICAL DEBATE

These issues of machine intelligence are not new to philosophy. The debate about whether a machine might think has its philosophical roots in the seventeenth and eighteenth century with the development of modern science. If the universe is fundamentally materialistic and mechanistic, as the emerging scientific paradigm suggested, it would follow that humans are nothing more than machines. Possibly, other machines might be constructed that would be capable of thought as well. Thomas Hobbes (1588–1679), who advocated a materialistic, mechanistic view, argued that reasoning is reckoning and nothing more. Humans reason by calculation with signs involving addition, subtraction, and other mathematical operations. Hobbes took these signs to be material objects that have significance as linguistic symbols. Julien La Mettrie (1709–1751), another materialist and mechanist, speculated that it might be possible to teach a language to apes and to build a mechanical man that could talk.

Not every philosopher of that era agreed with such radical predictions. René Descartes (1596–1650) held that animals are, indeed, complex machines but as such, necessarily lack thought and feeling. People have bodies that are in themselves nothing but complex machines, but people also have minds, nonmaterial entities that are in time but not space, that interact with their bodies. On this dualistic conception, intelligence and consciousness of people exist only as part of their minds, not as part of their bodies. Constructing a nonhuman machine that by itself had intelligence or consciousness was an impossibility for Descartes. Descartes admitted that a machine could be built that might give an impression of possessing intelligence, but it would be only a simulation of real intelligence and could be unmasked as a thoughtless machine. In fact, Descartes offered two certain tests by which a machine can be distinguished from a rational

human being even if the machine resembled a human in appearance. First, although a machine may utter words, a machine will never reply appropriately to everything said in its presence in the way that a human can. Second, although a machine may perform certain actions as well as or even better than a human, a machine will not have the diversity of actions that a human has.

THE CONCEPTION OF COMPUTING MACHINES

The contemporary debate about the possibility of machine intelligence ignited with the advent of modern electronic computers that are accurate, reliable, fast, programmable, and complex. Nobody did more in the twentieth century to construct a coherent concept of computing and to generate the contemporary debate about the intellectual possibilities of computers than Alan Turing (1912–1954). Turing explained computability in terms of abstract mathematical machines, now called *Turing Machines*. A Turing machine consists of a potentially infinite tape, divided into individual cells, on which a read–write head travels either left or right one cell at a time. The read–write head follows instructions that are found in a table of transition rules. The table of transition rules is the program that directs the Turing machine. Each instruction in the table specifies for a given a state of the Turing machine and a particular symbol being read on the tape, what the read–write head should do (print a symbol, erase the symbol, move right, or move left), and which state the machine should go to next.

Turing showed how such simple, elegant machines could compute ordinary arithmetic functions, and he conjectured that anything that is effectively computable could be computed by such a machine. In addition, Turing developed the concept of a universal Turing machine that can compute what any Turing machine can compute. Turing also showed the limitations of his machines by demonstrating that some functions are not computable, even by a universal Turing machine. Turing's seminal work on computable numbers and computing machines provided much of the conceptual foundation for the development of the modern computer. During World War II Turing applied some of his theoretical insights in designing special computing equipment to decipher the German Enigma codes. After World War II Turing led efforts to design some of the earliest computers, including the Automatic Computing Engine (ACE) in 1945.

The concept of computing developed by Turing provided not only a theoretical foundation for computer science but also a theoretical framework for much of artificial intelligence and cognitive science. A central paradigm of these fields is that mental processes and, in particular, cognitive processes are fundamentally computational. Processes that constitute and demonstrate human intelligence and general mentality, such as perception, understanding, learning, reasoning, decision making, and action, are to be explained in terms of computations. On the computational view, a mind is an information processing device. In its strongest form the computational theory of the mind holds that an entity has a mind *if and only if* that entity has computational processes that generate mentality.

Three important aspects of the theory of computation support the possibility of machines possessing intelligence and various aspects of minds. First, computation is understood in terms of the manipulation of symbols. Symbolic manipulation can represent information inputted, information processed, information stored, and information outputted. If human intelligence depends on the ability to represent the world and to process information, then the symbolic nature of computation offers a promising environment in which to conceive and develop intelligent machines. Much, though not all, of machine intelligence work has been conducted within this framework.

Second, if intelligence and mentality are computational in nature, then it does not matter what material conducts the computations. The computational structures and processes are multiply realizable. They might be instantiated in human brains, in computers, or even in aliens comprised of a different assortment of chemicals. All may have mentality as long as they have the appropriate computational processes. Indeed, it is possible to have mixed systems comprised of different materials. Cochlear implants and bionic eyes send information to human brains from external stimuli. Humans with these implants hear and see although part of their processing channels are inorganic.

Third, the computational model suggests an account of the connection between mind and body that other theories of the mind leave mysterious. The computational model explains intelligence and overall mental activity on the basis of decreasingly complex components. A hierarchy of computational systems is hypothesized, each of which is made up of simpler computational systems, until at bottom—as in a computer—there is nothing but elementary logical components, the operations of which can be explained and easily understood in terms of physical processes.

THE TURING TEST

For many people the phrase *machine intelligence* is an oxymoron. Machines by their nature are typically regarded as unintelligent and unthinking. How could a mere machine demonstrate actual intelligence? Turing believed that computing machines could be intelligent but was concerned that our judgments of the intelligence of such machines would be influenced by our biases and previous experiences with the limitations of machines. In his seminal article, "Computing Machinery and Intelligence" (1950), Turing considered the question "Can machines think?" but did so by replacing that question with another. The replacement question is explained in terms of a game that he calls "the imitation game." The imitation game is played by a man (A), a woman (B), and a human interrogator (C). The interrogator is in a room apart from the other two and tries to determine through conversation which of the other two is the man and which is the woman. Turing suggested that a teleprinter be used to communicate to avoid giving the interrogator clues through tones of voice. In the game the man may engage in deception in order to encourage the interrogator to misidentify him as the woman. The man may lie about his appearance and preferences. Turing believed that the woman's best strategy in the game is to tell the truth.

After he explained how the imitation game is played in terms of a man, a woman, and a human interrogator, Turing introduced his replacement question(s). Turing said, "We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'" (Turing, 1950, p. 434). Although his proposed version of the imitation game, now called the *Turing test*, may seem straightforward, many questions have been raised about how to interpret it. For example, to what extent does gender play a role in the test? Some maintain that Turing intended, or should have intended, that the computer imitate a woman just as the man did in the original imitation game. The more standard interpretation of the test is that the computer takes the part of A but that the part of B is played by a human—a man or a woman. On the standard interpretation the point of the test is to determine how well a computer can match the verbal behavior of a human, not necessarily a woman. The examples of questions for the test that Turing suggested are not gender specific but rather more general inquiries about writing sonnets, doing arithmetic, and solving chess problems.

Turing neglected to elaborate on many details of his test. How many questions can be asked? How many judges or rounds of judging are there? Who is the average interrogator asking questions? What counts precisely as passing the test? And, importantly, what conclusion should be drawn from a Turing test if it were passed? Turing moved quickly to replace the initial question "Can machines think?" with questions about playing the imitation game. He suggested that the original question "Can machines think?" is "too meaningless to deserve discussion" (Turing, 1950, p. 442). He could not have been claiming that the question is literally meaningless, or his own replacement project would not make sense. What he was suggesting is that terms like *machine* and *think* are vague terms in ordinary speech, and what people typically associate with a machine is not something that has or perhaps could have intelligence. What he was proposing with his test is a way to make the overall question of machine thinking more precise so that at least in principle, an empirical test could be conducted. Still, the issue is left open as to exactly what passing the Turing test would establish. Could it ever show that a machine is intelligent or that a machine thinks or possibly even that a machine is conscious?

A widely held misconception is that Turing proposed the test as an operational definition of thinking or considered the test to give logically necessary and sufficient conditions for machine intelligence. Critics of the test frequently point out that exhibiting intelligent behavior in this test is neither a logically necessary nor logically sufficient condition for thinking. But this common objection against the test misses the mark, for Turing never said he was giving an operational definition and never argued that the test provided a logically necessary or sufficient condition for establishing machine intelligence. Indeed, Turing argued for the opposite position. He did not take his test to be a necessary condition for intelligence, for he readily admitted that a machine might have intelligence but not imitate well. He never maintained that passing the test is logically sufficient for intelligence or thinking by a machine. On the contrary, he argued that demanding certainty in knowledge of other minds would push one into solipsism, which he rejected.

A more plausible interpretation of the Turing test is to regard it as an *inductive* test. If a machine passed a rigorous Turing test with probing questioning on many topics, perhaps by different judges over a reasonably extended period of time, then good inductive evidence

for attributing intelligence or thinking to the machine might exist. Behavioral evidence is used routinely to make inductive judgments about the intelligence of other humans and animals. It would seem appropriate to use behavioral evidence to evaluate machines as well. In judging human-like intelligence linguistic behavior seems particularly salient. There would be no logical certainty in such a judgment any more than there is logical certainty in scientific testing in general, and revision of judgments in light of new evidence might be required. Regrettably, other evidence like relevant to a judgment of machine intelligence, such as evidence from non-linguistic behavior and evidence about the internal operation of the machine, cannot be directly gathered within the Turing test. Turing realized this, but thought it more important to eliminate bias so that a machine would not be excluded as intelligent simply because the person making the judgment knew it was a machine.

CRITICISMS OF THE TURING TEST

Turing himself considered and replied to a variety of criticisms of his test ranging from a theological objection to an extrasensory perception objection. At least two of the objections he discussed remain popular. One is the Lady Lovelace objection based on a remark by Ada Lovelace that Charles Babbage's Analytical Engine, a nineteenth-century mechanical computer, had no pretension to originate anything. A similar point is often made by claiming that computers only do what they are programmed to do. The objection is difficult to defend in detail because computers can surprise even their programmers, are affected by their input as well as their programming, and can learn. Of course, one might argue that, at bottom, computers are merely following rules and therefore are not creative. But to firmly establish this objection, one would need to show that, at bottom, humans are not merely following rules and that anything merely following rules cannot be creative.

Another objection that Turing considered is the mathematical objection that utilizes results in mathematical logic, such as Kurt Gödel's incompleteness theorem. This argument, later developed by J. R. Lucas (1961) and by Roger Penrose (1989), maintains that fundamental limits of logical systems are limits of computers but not of human minds. But, as Turing himself pointed out, it has not been established that these logical limits do not apply equally well to humans.

In addition to these classical criticisms, a number of contemporary objections to the Turing test have been advanced. Robert French (1990) has maintained that the test is virtually useless because there will always be subtle subcognitive behavior that will allow an interrogator to identify humans from machines. If true, the Turing test would be more difficult to pass and possibly not very useful, but this outcome would also enhance the potential inductive sufficiency of the Turing test if it were passed.

In another criticism of the Turing test, Ned Block (1981) has suggested that a computer program that worked as a conversation jukebox so that it gave a stored but appropriate response to every possible remark by an interrogator throughout a conversation would pass the test. Because the test occurs during a finite period and in that period only a finite, though very large, number of responses can be made, such a program seems logically possible. Whether such a program could exist in practice given the complexity of semantic relations in a conversation and the changing facts of the world is unclear, but even taken as a thought experiment, the success of the jukebox program would at most show that the Turing test does not provide a logically sufficient condition for the possession of intelligence, a position to which Turing agreed.

John Searle (1980) developed one of the most popular contemporary objections against machine intelligence: the Chinese Room Argument. Simply put, a computer program running on a digital machine is only manipulating symbols syntactically and necessarily lacks semantics. Thus, even if a machine passed a Turing test, it would not understand anything. A digital computer might simulate intelligence, but on Searle's view it would not have a mind. Some critics of this argument have suggested that humans acquire semantics through interaction with the environment, and possibly, machines equipped with sensory inputs and motor outputs could acquire semantics in this way as well. More telling, the Chinese Room Argument does not validly establish what it claims. Searle has maintained that a human brain has the causal powers to produce a mind; the Chinese Room Argument does not demonstrate that computer programs, once loaded and running on a physical machine, could not have similar causal powers.

The Turing test is a possible test for machine intelligence and one that has received enormous philosophical discussion, but it is not the only test. Normally, the intelligence of animals and other humans is tested and inferred by examining an entity's relevant behavior in various situations. Similarly, machine intelligence can be tested based on its ability to demonstrate such processes as understanding, reasoning, and learning regardless of how well it can imitate a human. Human intelligence is

not the only kind of intelligence. Along these lines Patrick Hayes and Kenneth Ford (1995) have argued that too much emphasis on passing the Turing test has actually been detrimental to progress in artificial intelligence.

THE FUTURE OF MACHINE INTELLIGENCE

Turing believed that human language and understanding of machines and mentality would shift by the year 2000, and indeed, the notion of a machine being intelligent is not as outlandish as it once was. In his 1950 article (p. 442) Turing also made a very famous specific prediction that has not fared as well. He said: "I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning." No computer has come close to meeting this standard in a rigorously conducted Turing test. But it should be noted that behind Turing's prophecy was a plan that has not come to pass. He imagined that one day a computer would learn just as a child does. And like a human the computer gradually would obtain a larger and larger understanding of the world. Machine learning in specific contexts has been a reality for decades, but general learning by a machine remains an elusive goal, and without it, the intelligence of machines will be limited.

The long-term future of machine intelligence is a matter of considerable philosophical debate. Here are four visions of the future that have been suggested. On the *android vision* some intelligent machines of the future will look like humans or at least resemble humans in their intellectual capacities. Because humans are the most intelligent creatures known, human intelligence is taken as the obvious standard. Turing's own proposals employed much of this vision. From this viewpoint it is sensible to ask whether robots someday will be the intellectual peers of humans and might deserve rights as rational beings. But some critics argue that computers will never be much like humans without similar emotional needs and desires. On the *slave vision* intelligent machines of the future will give humans increasingly sophisticated assistance but, like their not-so-intelligent predecessors, they will be slaves, possibly held in check by Isaac Asimov's well-known three laws of robotics (1991). On the *successor vision* machine intelligence will become increasingly sophisticated and machines will evolve beyond humans. Hans Moravec (1999) has argued that humans will be surpassed by machines in terms of intelligence within a relatively short time. Such machines might evolve and progress rapidly through a Lamarckian transmission of culture to the next generation. Finally, on the *cyborg vision*, advanced by Rodney Brooks (2002) and others, machine intelligence will increasingly be embedded in us. Machine intelligence will be used to augment our abilities and will blend into our nature. Machine intelligence will become part of our intelligence, and we will become, at least in part, intelligent machines.

See also Artificial Intelligence; Chinese Room Argument; Computationalism; Descartes, René; Gödel's Theorem; Hobbes, Thomas; Induction; La Mettrie, Julien Offray de; Solipsism; Turing, Alan M.

Bibliography

Asimov, Isaac. *Robot Visions*. New York: Penguin Books, 1991.
Block, Ned. "Psychologism and Behaviorism." *Philosophical Review* 90 (1981): 5–43.
Boden, Margaret. *The Creative Mind: Myths and Mechanisms*. London: Routledge, 2004.
Brooks, Rodney. *Flesh and Machines: How Robots Will Change Us*. New York: Pantheon, 2002.
Copeland, Jack, ed. *The Essential Turing*. New York: Oxford University Press, 2004.
Dennett, Daniel. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.
Descartes, René. *Discourse on Method*. Translated by Donald Cress. Indianapolis, IN: Hackett, 1980.
Dreyfus, Hubert, and Stuart Dreyfus. *Mind over Machine*. New York: Free Press, 1986.
French, Robert. "Subcognition and the Limits of the Turing Test." *Mind* 99 (1990): 53–65.
Gunderson, Keith. *Mentality and Machines*. 2nd ed. Minneapolis: University of Minnesota Press, 1985.
Haugeland, John. *Artificial Intelligence: The Very Idea*. Cambridge, MA: MIT Press, 1985.
Hayes, Partick, and Kenneth Ford. "Turing Test Considered Harmful." *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* (1995): 972–977.
Hobbes, Thomas. *Leviathan: With Selected Variants from the Latin Edition of 1668*, edited by Edwin Curley. Indianapolis, IN: Hackett, 1994.
Hodges, Andrew. *Alan Turing: The Enigma*. New York: Simon and Schuster, 1983.
La Mettrie, Julien. *L'homme Machine: A Study in the Origins of an Idea*. Princeton, NJ: Princeton University Press, 1960.
Lucas, J. R. "Minds, Machines, and Gödel." *Philosophy* 36 (1961): 120–4.
Moor, James. "The Status and Future of the Turing Test." *Minds and Machines* 11 (2001) 77–93.
Moravec, Hans. *Robot: Mere Machine to Transcendent Mind*. New York: Oxford University Press, 1999.
Penrose, Roger. *The Emperor's New Mind*. Oxford: Oxford University Press, 1989.
Searle, John. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (1980): 417–457.

Sterrett, Susan. "Turing's Two Tests for Intelligence." *Minds and Machines* 10 (2000): 541–559.
Turing, Alan. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433–460.

James H. Moor (2005)

MACINTYRE, ALASDAIR
(1929–)

Alasdair Chalmers MacIntyre was born in Glasgow, Scotland. He was philosophically trained at Manchester University and subsequently taught at Manchester, Leeds University, Oxford University, and the University of Essex before emigrating to the United States in 1970. Since then he has held teaching posts at Brandeis University, Boston University, Vanderbilt University, Duke University, and the University of Notre Dame.

By his late teens MacIntyre became sympathetic to Marxism as a theoretical articulation of the failures of contemporary social, economic, and political institutions, resulting in the publication of his first book, *Marxism: An Interpretation*, at the age of twenty-three. While never giving up his view that modernity merits wide-ranging criticism and that such criticism must come from a rationally defensible theoretical standpoint, he came to believe that Marxism lacked the necessary resources. What is needed, MacIntyre held, is a moral and political philosophy built on an adequate theory of human nature and the human good—though this theory would have to recognize that human nature and the human good are deeply historically conditioned. What is also needed is an adequate account of how such a theory can be shown to be rationally superior to its rivals—though, again, this account would have to recognize that standards of rationality in inquiry are themselves deeply historically conditioned. MacIntyre's mature philosophy, expressed in the series of books *After Virtue* (1981), *Whose Justice? Which Rationality?* (1988), *Three Rival Versions of Moral Enquiry* (1990), and *Dependent Rational Animals* (1999), respond to these perceived needs and exhibit the most noteworthy features of his work: his account of tradition-constituted rationality, his Aristotelian ethics of virtue, and his Aristotelian politics of local community.

TRADITION-CONSTITUTED RATIONALITY

MacIntyre's view is that the most salient feature of contemporary moral and political discourse is interminable disagreement. Defenders of rival views become ever more sophisticated in the development and advocacy of their theories, but there is no progress toward resolution of these disagreements. It seems to be the aspiration of participants in these debates to offer a defense of their respective theories that is acceptable to any rational agent. MacIntyre calls this aim of providing a defense of morality acceptable to rational agents as such The Enlightenment Project, and holds that, for all the substantive differences between figures such as David Hume and Immanuel Kant, it is their common objective to provide a basis for morality that commands rational acceptance by all. After all, one might think that the alternative is an unacceptable relativism whereby different theories are justified in terms of different standards, with no way to bring rival theories truly into competition.

MacIntyre's contribution is to argue for the existence of rival and incompatible standards of rational assessment while denying that this affirmation brings with it a commitment to relativistic conclusions. We are confronted with different traditions of rational inquiry, each with its own theories and standards for assessment of theories, and each with a history within which various positions have been forwarded, defended, and to whatever extent affirmed or rejected. There is no neutral rationality-as-such by which we can decide between these various competing traditions. But relativistic conclusions do not follow, MacIntyre argues, because it is always possible that one tradition can show itself superior to a rival tradition by showing that one's tradition fares better than the rival even in that rival's own terms.

MacIntyre's positive views in ethics and politics are versions of Aristotelianism. In keeping with his conception of rationality in inquiry, his basis for affirming these views is that Aristotelianism is more defensible than rival traditions, even on those rival traditions' own terms.

THE ETHICS OF VIRTUE

MacIntyre argues in *After Virtue* that of the classical moral theories presented by Hume, Kant, Jeremy Bentham, and John Stuart Mill, neither they nor their contemporary defenders offer anything like compelling reasons to affirm these theories, nor do we have any reason to think that such reasons are forthcoming. Should we then, MacIntyre asks, follow Friedrich Nietzsche in thinking that the institution of morality is a fraud, to be jettisoned as the institution of taboo was jettisoned?

MacIntyre holds that there is an alternative to the moral theories defended in the Enlightenment and in the wake of the Enlightenment: Aristotelianism. On Aristotle's view ethics deals with the transformation of human

beings from their immature condition into a condition that constitutes their true end, the realization of their specifically human potentialities, which realization occurs through the acquisition and exercise of various moral and intellectual virtues. Aristotelianism fell by the wayside during the Enlightenment—in part because of its close identification with Roman Catholic scholasticism and in part because of the discrediting of Aristotelian science in the Scientific Revolution—but MacIntyre argues that this rejection was unwarranted, for an ethics coming out of the Aristotelian tradition is the best hope for moral philosophy.

MacIntyre's original formulation of this virtue ethics in *After Virtue* defines the virtues in terms of those qualities of character and intellect that are necessary for one's achievement of goods specific to practices (for example, games, crafts, arts, sciences, and other complex activities), for the sustenance of one's quest for the good life, and for the maintenance of one's community and one's traditions. He does not there formulate his view as part of a teleological conception of human nature and, indeed, in that work, he treats it as a desideratum for a restated Aristotelian ethics that it not rest on such a *metaphysical biology*. But in later works, most clearly *Dependent Rational Animals*, MacIntyre argues that ultimately we have to understand the virtues in terms of just such a teleological conception—a version of Aristotelianism grounded in the work of Thomas Aquinas—and that this conception is not at odds with the well-founded claims of contemporary science.

THE POLITICS OF LOCAL COMMUNITY

MacIntyre's views in political philosophy are frequently labeled *communitarian*, but this is a mistake if by communitarian we mean the position that states should be in some way guided by the ideals of the value of community. MacIntyre's position is more radical, for he holds that every conception of politics that is built on the attempt to justify the state is doomed to failure. For the state—a hierarchically structured apparatus of political control—is not justifiable; all attempts to explain why the state is authoritative have failed. This does not mean that politics is an empty enterprise or that authority is inevitably illegitimate. It means, rather, that the goods of politics are realized not through the state but through much more local communities in which people can engage in genuine argument and have effective control over how their common life is structured. Only in local communities can the politics of the common good rather than that of individual advantage or class dominance be practiced. This emphasis on the necessarily local character of good politics also marks MacIntyre's views as Aristotelian.

See also Aristotelianism; Aristotle; Bentham, Jeremy; Communitarianism; Enlightenment; Hume, David; Kant, Immanuel; Marxist Philosophy; Mill, John Stuart; Nietzsche, Friedrich; Scientific Revolutions; Social and Political Philosophy; Thomas Aquinas, St.; Virtue Ethics.

Bibliography

WORKS BY MACINTYRE

Marxism: An Interpretation. London: SCM Press, 1953.

A Short History of Ethics (1966). 2nd ed. Notre Dame, IN: University of Notre Dame Press, 1998.

After Virtue: A Study in Moral Theory (1981). 2nd ed. Notre Dame, IN: University of Notre Dame Press, 1984.

Whose Justice? Which Rationality? Notre Dame, IN: University of Notre Dame Press, 1988.

Three Rival Versions of Moral Enquiry: Encyclopaedia, Genealogy, and Tradition. Notre Dame, IN: University of Notre Dame Press, 1990.

The MacIntyre Reader, edited by Kelvin Knight. Notre Dame, IN: University of Notre Dame Press, 1998.

Dependent Rational Animals: Why Human Beings Need the Virtues. Chicago: Open Court, 1999.

WORKS ON MACINTYRE

Horton, John, and Susan Mendus, eds. *After MacIntyre: Critical Perspectives on the Work of Alasdair MacIntyre*. Notre Dame, IN: University of Notre Dame Press, 1994.

Murphy, Mark C., ed. *Alasdair MacIntyre*. New York: Cambridge University Press, 2003.

Mark C. Murphy (2005)

MACKIE, JOHN LESLIE
(1917–1981)

John Leslie Mackie was born in Sydney, Australia, and educated under John Anderson at the University of Sydney, and at Oxford, where he graduated with a First in Literae Humaniores in 1940. After the war, he returned to an academic position in the University of Sydney, and in 1955 he took up the Chair in Philosophy at the University of Otago, in Dunedin, New Zealand. In 1959 he returned to the University of Sydney to replace Anderson in the Challis Chair. After five years he left for Great Britain, going first to fill the foundation Chair of Philosophy at the new University in York. In 1967 he became Fellow of University College, Oxford, and University Reader in 1978. He remained at Oxford until his death in 1981.

Mackie's work is characterized by an acute, unwearied, and always dispassionate analysis of alternative solutions to specific philosophical problems. Striving first for full clarity in the statement of the problem, he proceeds by careful exploration and appraisal of the arguments available in support of alternative proposed solutions. Mackie applied this analytic style of reasoning across a broad range of issues. He made contributions to, among other topics, logic—and particularly the understanding of logical paradoxes; to the nature of conditionals and the theory of causality; to the interpretation of counterfactual conditionals; to the theory of space and time; to the theological problem of evil; to the theory of ethics; to the relations between reason, morality, and law; to the philosophy of mind; to the philosophy of biology; and to the interpretation of Locke's epistemology and metaphysics, and of Hume's ethics.

For many years, Mackie published a succession of important articles, but no books. This pattern of publication was transformed in 1973 with the appearance of *Truth, Probability, and Paradox*, a collection of essays on logical themes. This was followed in rapid succession by *The Cement of the Universe* (1974), which presents his views on causation, and *Problems from Locke* (1976). In this work Mackie takes up a group of characteristically Lockean themes, including primary and secondary qualities, perception, substance, universals, identity, and innate ideas, and relates them to contemporary discussion of the same issues. In *Ethics, Inventing Right and Wrong* (1977) he presents a sustained argument for a distinctive error-projection account of human moral thinking, which was provided with some additional support in his extended discussion of Hume's moral theory, which appeared in a book of that name in 1980. Lastly, posthumously, *The Miracle of Theism* was published in 1982. Its subtitle—*For and Against the Existence of God*—sufficiently indicates its contents. Though scrupulously fair, Mackie himself was firmly convinced by the case for atheism. This burst of productivity propelled Mackie to the forefront among British philosophers of his generation, and his relatively early death, while still at the height of his powers, was keenly felt.

MACKIE'S THESES

Although contributing to many debates in the course of his career, Mackie is principally celebrated for four distinctive theses. The first, in philosophical theology, is his insistence, patiently argued over many years, that all the attempts to reconcile the existence of evil with the classical Christian conception of God as omnipotent, omniscient, and benevolent are failures, and that any plausible variations on them will fail also.

The second is in philosophical logic, in which Mackie argues that despite appearances, counterfactual conditionals are not actually propositions at all, but rather condensed and elliptically expressed arguments. The conditional's antecedent is the argument's premise, and its consequent is the conclusion. The counterfactual conditional is to be accepted if the argument is good as it stands, or can be made good by the supply of plausible understood additional premises.

The third thesis pertains to metaphysics, specifically causation. Recognizing that in almost every case the whole cause of an event involves multiple factors, Mackie proposed an account of causal factors. These, he held, are INUS conditions—that is: insufficient but necessary parts of unnecessary but sufficient conditions for the occurrence of the effect.

In ethics, the area of Mackie's fourth distinctive thesis, he argues that although the semantics of ordinary indicative moral discourse apparently require that there be moral facts in virtue of which human moral claims are true or false, there are no such moral facts. Moral discourse must therefore be explicated as arising from widespread error. The denial of objective moral facts is the aspect of his thought that most clearly shows the influence of his Andersonian education. Mackie argued that people's attitudes and feelings when considering their behavior and its effects lead them to assume, falsely, the existence of objective features of right or wrong, good or bad, in human situations, which correspond to, and validate, those attitudes and feelings. As there are no such validating properties, people must take on themselves the responsibility for the judgments they make.

In the years since his death, Mackie's philosophy has continued to be influential. In particular, his controversial views in ethics and philosophical theology continue to attract critical but respectful discussion.

See also Anderson, John; Causation: Metaphysical Issues; Conditionals; Counterfactuals; Evil, The Problem of; Hume, David; Locke, John; Logical Paradoxes; Noncognitivism; Philosophy of Biology; Philosophy of Mind; Space; Theism, Arguments For and Against; Time.

Bibliography

WORKS BY MACKIE

Truth, Probability and Paradox: Studies in Philosophical Logic. Oxford: Clarendon Press, 1973.

The Cement of the Universe: A Study of Causation. Oxford: Clarendon Press, 1974.

Problems from Locke. Oxford: Clarendon Press, 1976.

Ethics: Inventing Right and Wrong. Harmondsworth, U.K.: Penguin, 1977.

Hume's Moral Theory. London: Routledge and Kegan Paul, 1980.

The Miracle of Theism: Arguments For and Against the Existence of God. Oxford: Clarendon Press, 1982.

COLLECTIONS OF MACKIE'S ARTICLES

Mackie, J., and P. Mackie, eds. *Logic and Knowledge: Selected Papers. Vol. 1*. Oxford: Clarendon Press, 1985.

Mackie, J, and P. Mackie, eds. *Persons and Values; Selected Papers. Vol. 2*. Oxford: Clarendon Press, 1985.

WORKS ABOUT MACKIE

Honderich, T., ed. *Morality and Objectivity; a Tribute to J. L. Mackie*. London: Routledge and Kegan Paul, 1985. Contains a comprehensive Mackie bibliography.

Keith Campbell (1996, 2005)

MACROCOSM AND MICROCOSM

"Macrocosm" and "microcosm" are philosophical terms referring, respectively, to the world as a whole and to some part, usually man, as a model or epitome of it. According to one version of this ancient analogy, man and the universe are constructed according to the same harmonic proportions, each sympathetically attuned to the other, each a cosmos ordered according to reason. By an imaginative leap, the universe itself was thought to be, like man, living and conscious, a divine creature whose nature is reflected in human existence. Animism and panpsychism also regard the world as alive throughout, but the microcosm idea is distinct in emphasizing the unity or kinship of all life and thought in the world. If man is the microcosm of the universe, then not only is everything animated by *some* soul or other, but there is *one* world soul by which everything is animated. Thus, the followers of Pythagoras and Empedocles held, according to Sextus Empiricus, that "there is a certain community uniting us not only with each other and with the gods but even with the brute creation. There is in fact one breath pervading the whole cosmos like soul, and uniting us with them" (W. K. C. Guthrie, *A History of Greek Philosophy,* Vol. I, p. 278).

Because the word *kosmos* can mean order as well as world or world order, "microcosm" can signify not only man in relation to the universe (or in relation to the state, as in Plato's *Republic*) but also any part of a thing, especially a living thing, that reflects or represents the whole it belongs to, whenever there is a mirroring relation between the whole and each of its parts. Nicholas of Cusa's doctrine of individuals as "contractions" of the form of the universe is a microcosm theory, as is Gottfried Wilhelm Leibniz's theory of monads as "perpetual living mirrors of the universe"; similarly, to cite an example from nonphilosophical discourse, the composer Béla Bartók's collection of piano pieces *Mikrokosmos* is a little world of modern musical style and technique.

The idea of the microcosm appears in pre-Socratic philosophy in connection with the problem of relating the One and the Many. Taking all of nature to derive ultimately from a single common substance, they supposed it to have inherent in it a principle of motion and change (which they identified with life, soul). Since some of the resulting entities possess consciousness, so too must their source. And if the universal soul is eternal and divine, then the human soul, which is a "fragment" of the One, as the Pythagoreans held, must also be eternal and divine. The return of the individual soul to its divine origin could be realized by philosophical understanding of the cosmos; since like is known by like, as the cosmos becomes known the knower is assimilated to it. Thus, man is, and discovers himself to be, the part that most perfectly reveals the nature of the whole.

Man the microcosm is a commonplace of Greek thought from Anaximenes, the Pythagoreans, Heraclitus, and Empedocles to the Stoics and Neoplatonists. It is a staple theme for variation in the Orphic, Gnostic, and Hermetic texts and in the literature of mysticism, pantheism, and the occult. That man is the microcosm was, in the Renaissance, widely taken to mean that cosmic knowledge and influence might be achieved through contemplation of the powers and tendencies men find in their own imaginations. Such knowledge would be based not on mere inference from resemblance but rather on the kinship or identity of human life and consciousness with the forces governing nature as a whole.

The notion that man is the microcosm has always played both rational and mystical roles in Western thought. Well into the period of the scientific revolution, the microcosm was an image of the order and harmony pervading the world. Saying that the universe is controlled by a single principle (in the way that rational thought is the controlling principle in man) expressed the unified and self-regulating character of the world as understandable in its own terms, fit for scientific investigation. Similarly, human thought itself was conceived to be self-regulating and self-correcting—thus entered the

idea of the autonomy of reason that has played an important part in the history of rationalism and of Western philosophy generally. According to Plato's recollection doctrine, "All nature is akin, and the soul has learned everything, so that when a man has recalled a single piece of knowledge—*learned* it, in ordinary language—there is no reason why he should not find out all the rest" (*Meno* 81D, E). By recollection Plato meant the recovery of systematic knowledge of necessary truths from within oneself, but it is easy to see how it could also be thought of as an intuitive, nontheoretical process—a stream of consciousness leading to memory of past reincarnations or of the soul's celestial origin.

The thought that the universe is ordered not by chance but by one spiritual principle stimulated the wish for direct mystical union with this soul, and even for influence over things through it, as easily as it encouraged the pursuit of systematic understanding of the world. The first impulse produced such exalted sentiments as those lavished upon the universe in the Hermetic religious writings; the second pushed open the door to that underground world of magic, astrology, alchemy, and spiritualism that claimed to utilize the same unifying principles assumed in science and in the astral theology of the philosophers. Perhaps something may be said for a generous interpretation of this magical view of nature, which even in antiquity was distinguishable from its rationalistic and humanistic counterpart. For the practitioners of the occult and for their opponents, the view of the world as a "be-souled" creature was neither an isolated hypothesis nor an idle conceit; the microcosm was an almost omnipresent presupposition, the basis of the very language in which the phenomena whose explanation was sought were represented. Yet there were always philosophical skeptics, and often the same writers who affirmed the world soul or the microcosm—for example, Plotinus, Giovanni Pico della Mirandola, Johannes Kepler—also tried to restrict it in ways that precluded the possibility of undesirable magical application.

ANCIENT THOUGHT

In the *Timaeus* Plato presents a mythical account of the creation of the world according to which the world's soul and body are made by the Demiurge, who copies the Form of the ideal living creature (not itself any species of animate being but embracing the types of them all). The world soul is constructed according to a complex musical pattern, and, in order to be capable of thought, the elements of discourse—sameness, difference, and existence—are blended to form its mind. The body joined to the world soul is said to be unlike the human body or that of any animal in the world, being perfectly spherical, devoid of organs of sense, respiration, and ingestion; however, the processes of the universe are said to be reproduced even in the details of microcosmic processes, such as the moment of blood in humans. And because of the affinity between the divine part in humans and the thoughts and revolutions of the universe, the study of the rhythms of the macrocosm are recommended as a means of "correcting those circuits in the head that were deranged at birth."

A methodological discussion forms the context of a playful passage in the *Philebus* (27A–31B) in which the microcosm image also appears. All philosophers hold mind to be the king of heaven and earth, Socrates observes: "in reality they are magnifying themselves. And perhaps they are right." Socrates and Protarchus agree that the order of the world proves that the cosmos is governed by "Mind [*nous*] and a wondrous regulating Intelligence." Socrates argues further that the elements composing our bodies are but fragments produced and sustained by the elements in the universe. Because the unity of the elements in us makes up our bodies, the collective unity of elements in the universe must make up the world's body; because our bodies have souls, the body of the universe must have one, too; for where could our bodies have gotten their souls "if the body of the universe, which has elements the same as our own though still fairer in every respect, were not in fact possessed of a soul?" Strictly, this much of the argument concludes merely in the existence of a world soul that is the cause of the mixture of the body's elements—there is as yet barely a hint of the world soul's having a structure of its own apart from the body, of its being rationally ordered and the cause not just of all mixture but of all movement in the cosmos. Ultimately, the universal soul itself is said to be produced by Cause (later identified with Mind), yet this Mind cannot come into existence without soul (30C). To the extent that we can distinguish the Demiurge from the world soul (in the *Timaeus*), we can say that the Cause of the *Philebus* is probably more like the first of these.

Aristotle's physical system seems to have been designed to avoid the view of the cosmos as "besouled" or as alive in all its parts. Thus, in *De Caelo* the motion of the stars is explained not by any life in them but mainly in terms of the circular motion natural to the *aether* of which they are composed. In Book II (Ch. 2) Aristotle rejects the view that "it is by the constraint of a soul that it [the heaven] endures forever." The Demiurge as designer of the world is wholly excluded; no conscious-

ness is needed of the rational (but unpremeditated) pattern to which nature adheres. Although there is a reference (to the views of others) in the *Physics* (Book VIII, Ch. 2), which may be the first occurrence of the Greek expression for "microcosm," Aristotle seems not to have organized his conception of nature around the view of it as an organism in any significant way. (For a contrasting account, see W. K. C. Guthrie, "Man as Microcosm.")

What is missing in Aristotle reappears (partly under Heraclitus's influence) in the thought of the Stoics—the sense of the world as an animate and conscious continuum each part of which affects all others by its *sympathy,* its "sharing of experience" with the others. The doctrine of sympathies and antipathies among the parts of the world animal guided the physical research of the Stoics and predisposed them to accept and to attempt to rationalize the particulars of astrology and divination. And man as microcosm was the source of their efforts to locate the basis of human conduct in natural law; by playing one's assigned role in the cosmos, one's logos, his "inner self," would be linked to that of the whole (Hans Jonas, *The Gnostic Religion,* p. 248).

Plotinus, like the Stoics, treated the world as a single creature, "living differently in each of its parts." If the world soul of Plato's system is thought of as operating purposefully and consciously, and if the Nature of Aristotle's system is taken to work purposefully but unconsciously, we should say that for Plotinus the world as a whole is governed consciously yet produces individual things "as in a dream," spontaneously, without reasoning, choice, or calculation. According to Plotinus only a unity of soul among us could explain our sympathetic relations to one another, "suffering, overcome, at the sight of pain, naturally drawn to forming attachments" (*Ennead* IV, ix, 3). Plotinus denied that the unity he spoke of entailed the transference of a person's emotions to places outside his body; the souls of the sufferer and of the sympathizer do not feel as one. Rather, his model of unity is that of a science, where individual truths cannot be considered apart from the whole; "the whole is in every part: … The one detail, when it is matter of science, potentially includes all" (IV, ix, 5). In geometry, for example, "the single proposition includes all the items that go to constitute it and all the propositions which can be developed from it" (IV, ix, 5). Perhaps this very strict sense of unity, which asserts that each thing is internally connected with every other thing (or that there is one thing with which each is connected) has always been latent in the microcosm doctrine; if so, it is an aspect of the doctrine that seems to offer small encouragement to the search for the actual relations in nature. The question "Which things are causally connected, which are not?" has little point if all can affect all alike.

The general ancient view of the world as a perfect organism may have been responsible, as Samuel Sambursky suggests, for the insistence of ancient thinkers on the attempt to understand the world as a whole, in its entirety, and for their almost total avoidance of experimentation—the isolation of phenomena, or "dissection of nature," characteristic of modern science.

MEDIEVAL AND MODERN THOUGHT

Man as microcosm of the universe is not integral to Jewish and Christian doctrine in the way that it is to the Gnostic religious system, for example; thus, Philo Judaeus and Moses Maimonides employed the idea of the world soul only dialectically. In *The Guide of the Perplexed* (Pt. I, Ch. 72) Maimonides at first argues that the world is like a human being, but he then presents so many points of difference between the two that in the end it is clear that he considers the possession of a rational order to be their only common factor. As a cosmological view, the microcosm has little or no place in Augustine or in Thomas Aquinas, who treats it as a mere figure of speech. By contrast, Joseph ibn Zaddik states one of the microcosm's main attractions when he proposes to show how self-knowledge will lead to knowledge of the whole—a "short cut" through the study of man, bypassing the sciences. Bernard of Tours and other members of the school of Chartres assimilated the world soul of Plato's *Timaeus* to the Third Person of the Trinity. Drawing upon Bernard, Hildegard of Bingen, in her visionary writings, represented detailed correspondences between heavenly motions, winds, elements, humors, and bodily and spiritual states in the individual.

Plato had typically employed the microcosm image to portray the transformation of consciousness through theoretical knowledge of whatever cosmic order science reveals; Ibn Zaddik reverses the process, seeking to discover in man what the cosmic order must be. Where Plato stressed the dissimilarity between the living cosmos and the structure and functioning of any particular animal, including man, Hildegard dwells on their supposed similarity in picturesque detail. The idea that inner experience of human nature supplies a direct route to reality is prone to magical extension in a way that Plato's view is not, but it was this conception that took hold in medieval and Renaissance microcosm literature.

Renaissance speculation on the microcosm centered on the idea that human nature partakes of bodily, intel-

lectual, and divine existence, uniting in itself the whole of the sublunary, celestial, and supercelestial realms. Human consciousness, by which man can know all things, connects him with all things; consciousness is itself a link between thought and its objects. Through consciousness man can know and become all that he wills. A similar doctrine of connections drawn from the Kabbalah underlies the various magical theories of language which asserted that quasi-physical influences join names and things, beyond the conventions of the various natural languages. Partly controllable influences also form the structure of the elaborate identities and correspondences that Agrippa von Nettesheim and Paracelsus described between minerals, animals, heavenly bodies, psychic powers, and parts of the human body. Such influences are also involved in the interaction between thought and its objects that Giordano Bruno assumed in his search for direct awareness of the sympathies controlling nature through memory and the ideas of them in his imagination.

The occult "applications" of the microcosm idea did not survive the advance of the mechanistic worldview. By the eighteenth century, occult qualities, or anything that seemed like them—for example, action at a distance—were in such wide disrepute that even Isaac Newton, to avoid the appearance of being committed to an occult doctrine, refrained from expressing fully his theory of the mode of action of atomic "Central Forces." But in the second edition of the *Principia* (1713), he described the ether as "a certain most subtle spirit which pervades and lies hid in all gross bodies … by the force and action of which spirit the particles of bodies attract one another at near distances and cohere … and all sensation is excited, and the members of animal bodies move at the command of the will, namely by vibrations of this spirit"—a view not far from that of the Stoics, as Stephen Toulmin and June Goodfield remark (*The Architecture of Matter,* p. 195).

Even later, belief in psychic planetary action had not lost all ground; thus, Franz Anton Mesmer's explanation of "animal magnetism," or hypnosis, assumed a "responsive influence … between the heavenly bodies, the earth, and animated bodies," which the hypnotist drew upon. And the idea of a psychic force in the world beyond our immediate awareness, of which our conscious lives are parts or manifestations, endured, for example, in Johann Wolfgang von Goethe's Nature philosophy and in Arthur Schopenhauer's world will—ancestors of the concept of the unconscious. Perhaps some aspects of the microcosm idea can be found in Sigmund Freud's attempts to explain the instincts in man as repetitions of the reactions of living matter to drastic changes in the prehistoric environment. (Thus, we might say that man's instincts are a microcosm of his evolution.) Among the known "enforced alterations in the course of life … stored for repetition," Freud, along with Sándor Ferenczi, noted the drying up of the oceans which left life to adapt on land and the cultural development necessitated by the glacial epoch. These are reexperienced at birth, in the diphasic onset of man's sexual life, and in the latency period. Freud invokes the contending forces, Love and Strife, of Empedocles's "Cosmic phantasy," pointing out their similarity to Eros and Destructiveness, the two primal instincts of his biopsychical theory. These instincts, which "present the delusive appearance of forces striving after change and progress" actually impel the organism toward the reinstatement of earlier, more stable states, ultimately to inorganic existence. The originally biological principle that ontogeny recapitulates phylogeny has received very wide psychological extension in psychoanalysis; most recently, Carl Jung has (somewhat cryptically) identified his doctrine of the collective unconscious with that of "the microcosm containing the archetypes of all ideas."

Perhaps the microcosm image is not entirely the scientific dead end it has understandably been taken for; as early attempts to construct models of the embodied soul's structure, development, and dynamics, some versions of the image may stand to scientific psychological research as alchemy stands to chemistry.

See also Agrippa von Nettesheim, Henricus Cornelius; Anaximenes; Aristotle; Augustine, St.; Bernard of Tours; Bruno, Giordano; Chartres, School of; Empedocles; Freud, Sigmund; Goethe, Johann Wolfgang von; Heraclitus of Ephesus; Hildegard of Bingen; Ibn Zaddik, Joseph ben Jacob; Jung, Carl Gustav; Kabbalah; Kepler, Johannes; Leibniz, Gottfried Wilhelm; Maimonides; Neoplatonism; Nicholas of Cusa; Panpsychism; Paracelsus; Philo Judaeus; Pico della Mirandola, Count Giovanni; Plato; Plotinus; Pythagoras and Pythagoreanism; Schopenhauer, Arthur; Sextus Empiricus; Socrates; Thomas Aquinas, St.

Bibliography

Three useful histories of the microcosm theme are G. P. Conger, *Theories of Macrocosms and Microcosms in the History of Philosophy* (New York: Russell and Russell, 1922), which includes a survey of critical discussions up to 1922; Rudolph Allers, "Microcosmus, From Anaximandros to Paracelsus," in *Traditio* 2 (1944): 319–407; and W. K. C. Guthrie's "Man as Microcosm," in *Proceedings of the European Cultural Foundation* (Athens, 1966); all contain

many references. W. K. C. Guthrie's discussion of the microcosm, to which this article is indebted, in *A History of Greek Philosophy* (Cambridge, U.K.: Cambridge University Press, 1962–), Vol. I, is the most important one for the period covered; this volume, *The Earlier Presocratics and the Pythagoreans,* also contains valuable remarks on Plato and Aristotle. The microcosm in Plato is discussed by F. M. Cornford throughout his commentary on the *Timaeus* in *Plato's Cosmology* (London: K. Paul, Trench, Trubner, 1937); G. M. A. Grube discusses the microcosm as part of Plato's theory of the soul in *Plato's Thought* (London: Methuen, 1935), Ch. 4; see also F. M. Cornford, "Psychology and Social Structure in the Republic of Plato," in *Classical Quarterly* (1912): 247–265; R. Hackforth's translation of the *Philebus,* with commentary, in *Plato's Examination of Pleasure* (Cambridge, U.K., 1945); and Gregory Vlastos, "Anamnesis in the *Meno,*" in *Dialogue* 4 (2) (September 1965): 143–167, which interprets the recollection theory with comments on its connection with the doctrine of reincarnation. Possible oriental influences on Plato are discussed in A. Olerud, *L'idée de microcosmos et de macrocosmos dans la Timée de Platon* (Uppsala, 1951). Two valuable relevant studies of Aristotle are W. K. C. Guthrie's introduction to the text and translation of *Aristotle on the Heavens* (London, 1939) and Friedrich Solmsen's *Aristotle's System of the Physical World* (Ithaca, NY: Cornell University Press, 1960).

On the Stoics, see Samuel Sambursky, *The Physics of the Stoics* (Princeton, NJ: Princeton University Press, 1987.)

On Plotinus, see the introductions and translations in E. R. Dodds, *Select Passages Illustrating Neoplatonism* (London, 1923), and A. H. Armstrong, *Plotinus* (London: Allen and Unwin, 1953). Remarks bearing on the microcosm in ancient thought generally are contained throughout E. R. Dodds, *The Greeks and the Irrational* (Boston: Beacon, 1957); Hans Jonas, *The Gnostic Religion* (Boston: Beacon, 1963), especially Ch. 10, "The Cosmos in Greek and Gnostic Evaluation"; A.-J. Festugière, *Personal Religion among the Greeks* (Berkeley: University of California Press, 1954); E. A. Lippman, *Musical Thought in Ancient Greece* (New York: Columbia University Press, 1964); and Samuel Sambursky, *The Physical World of the Greeks* (London: Routledge and Paul, 1956). See also E. W. Beth, *The Foundations of Mathematics: A Study in the Philosophy of Science,* rev. ed. (New York, 1964), Chs. 1 and 2, "The Pre-history of Research into Foundations" and "Aristotle's Theory of Science."

Hildegard of Bingen's life and writings are examined in Charles Singer, *From Magic to Science* (New York: Dover, 1958), Ch. 6, "The Visions of Hildegard of Bingen," a rewritten chapter from *Studies on the History and Method of Science,* Vol. I (Oxford, 1917). Ernst Cassirer, *Individuum und Kosmos in der Philosophie der Renaissance* (Leipzig: Teubner, 1927), translated by Mario Domandi as *The Individual and the Cosmos in Renaissance Philosophy* (New York: Harper, 1963), is the standard discussion of the microcosm in Renaissance thought. On the difficult subject of Renaissance occult literature, see D. P. Walker, *Spiritual and Demonic Magic from Ficino to Campanella* (London: Warburg Institute, University of London, 1958). Three chapters in Frederick Copleston's *A History of Philosophy,* Vol. III, *Late Medieval and Renaissance Philosophy,* Part 2 (Westminster, MD: Newman Bookshop, 1953), are useful surveys; Ch. 15 discusses the microcosm in Nicholas of Cusa, Chs. 16 and 17 are on the philosophy of nature. An important interpretation of Bruno is Frances Yates, *Giordano Bruno and the Hermetic Tradition* (Chicago: University of Chicago Press, 1964). There are also interesting discussions in Alexandre Koyré, *Mystiques, spirituels, alchimistes du XVI*[e] *siècle allemand* (Paris, 1955), and in Werner Pauli, "The Influence of Archetypal Ideas on the Scientific Theories of Kepler," in *The Interpretation of Nature and the Psyche* (New York: Pantheon, 1955). Microcosm and macrocosm are discussed in the context of the idea of the chain of being in E. M. W. Tillyard, *The Elizabethan World Picture* (New York, 1941); see also W. C. Curry, *Shakespeare's Philosophical Patterns* (Baton Rouge: Louisiana State University Press, 1937). On the transition from animism to mechanism in science, see E. J. Dijksterhuis, *Mechanization of the World-Picture,* translated by C. Dikshoorn (Oxford: Clarendon Press, 1961); M. B. Hesse, *Forces and Fields* (London: T. Nelson, 1961); and Stephen Toulmin and June Goodfield, *The Architecture of Matter* (New York: Harper and Row, 1962).

A brief account of Mesmer's ideas can be found in Clark L. Hull, *Hypnosis and Suggestibility* (New York: Appleton-Century, 1933), pp. 6–11. Schopenhauer's doctrine of the microcosm and its influence on Ludwig Wittgenstein are discussed in Patrick Gardiner, *Schopenhauer* (Baltimore: Penguin, 1963). Wittgenstein's remark "I am my world. (The microcosm.)" appears in the *Tractatus,* but without the connection with the world-spirit doctrine it has in his *Notebooks* (pp. 84–85). Wittgenstein's idea of an internal connection between language, thought, and reality is discussed in Erik Stenius, *Wittgenstein's Tractatus* (Oxford, 1960), and Max Black, *A Companion to Wittgenstein's Tractatus* (Ithaca, NY: Cornell University Press, 1964).

A short discussion of the microcosm image as employed by Freud and other analysts is contained in Philip Rieff's introduction to *General Psychological Theory* (New York, 1963), which is a volume in the paperback edition of Freud's *Collected Papers*; see pp. 9–17. Freud discusses Empedocles in "Analysis Terminable and Interminable," in the volume *Therapy and Technique,* edited by Philip Rieff (New York, 1963), the paperback edition of Freud's *Collected Papers.* Jung's ideas are expressed in his *Naturklärung und Psyche* (Zürich: Rasche, 1952), translated by R. F. C. Hull as *The Interpretation of Nature and the Psyche* (New York: Harcourt Brace, 1955). Ch. 3 of his essay "Synchronicity: An Acausal Connecting Principle" contains numerous quotations from earlier microcosm literature.

Problems that arise in trying to characterize the universe as a unified whole (or as a "whole" at all) on the basis of information concerning only a part and in trying to treat scientifically the nature of a necessarily unique object are presented in D. W. Sciama, *The Unity of the Universe* (Garden City, NY: Doubleday, 1959), pp. 69–205. For further discussion and bibliography, see the Cosmology and Rationalism entries.

Donald Levy (1967)

MAILLET, BENOÎT DE

(c. 1656–1738)

Benoît de Maillet was a French diplomat, traveler, and natural scientist. Information concerning the place and date of his birth, details of his life, and the significance of his works is, at best, sketchy and contradictory. A member of the impoverished nobility, Maillet presumably received the customary classical education of the day. He seems to have led an apathetic existence until his appointment to the French consulate in Cairo at the age of thirty-six. As consul, he handled the king's business well and, for services rendered, was named ambassador to Ethiopia in 1702. He declined the honor, ostensibly for reasons of health but actually because his duties would be less concerned with Franco-Ethiopian relations than with the formidable task of converting the natives to Christianity. In 1707, at his own request, he left his post in Cairo to assume charge of the French consulate in Livorno, Italy. He was so successful as consul and later as inspector of French settlements in other parts of the Mediterranean that, upon his retirement in 1724, he received a handsome pension and spent the remaining fourteen years of his life in Marseille. There, besides attending to a large correspondence, most of which is now lost, he wrote several works, including *Description de l'Egypte* (1735) and the vastly more important *Telliamed, ou entretiens d'un philosophe indien avec un missionnaire françois* (1748), which appeared posthumously.

TELLIAMED

The years of Maillet's consulships, his travels in the Mediterranean basin, and his wide readings and careful observations formed much of the background for *Telliamed* (the author's name spelled backward). First published in Amsterdam, it was closely followed by other editions in both French and English, the most important being that of the Abbé Le Mascrier (1755). The work consists of a series of conversations in which Maillet, speaking through his Indian philosopher, Telliamed, puts forth various geological and biological speculations about Earth's cosmogony and its evolution—together with the organic beings it supported—into its present state. According to Maillet's system, Earth, product of a whirlpool of cosmic dust, was for countless ages entirely covered with swirling waters. As the waters gradually receded, the primordial mountains formed by the currents of these waters slowly emerged from the depths. The crashing of the waves against these mountains formed new mountains, and with the appearance of life in the seas, fossil strata were formed.

Primitive forms of aquatic life, produced in ever-increasing abundance through the aeons, underwent gradual modifications of structure and function in keeping with changing habits and new environments. Thus, creatures along the shallow coastal waters moved into the marshes and, after much trial and error, finally emerged with wings for flying or legs for walking. Beneath this speculation lay the work's basic theme that everything in the universe, through the processes of time, was undergoing constant change. Occasionally the author's boldly imaginative thought resulted in whimsy, which was interpreted by many of his critics as folly or childish fantasy.

Telliamed immediately became a center of controversy that extended well into the nineteenth century. Maillet's heretical views, which ran counter to the tenets of Genesis, aroused the theologians of the day, while many eighteenth-century rationalists and scientists, led by Voltaire, were violently opposed to his ideas on other grounds. Disparaging criticisms continued in the writings of such eminent men of science as Étienne Geoffroy Saint-Hilaire and Georges Cuvier, Nonetheless, Comte de Buffon, Denis Diderot, Chevalier de Lamarck, and Erasmus Darwin, among others, availed themselves of Maillet's theories as a starting point for even more daring concepts of their own.

Bibliography

Collier, Katherine. *Cosmogonies of Our Fathers.* New York, 1934.

Dufrenoy, Marie-Louise. *Benoît de Maillet, précurseur de l'évolution.* Paris, 1960.

Haber, Francis C. *The Age of the World. Moses to Darwin.* Baltimore: Johns Hopkins Press, 1959.

Kohlbrugge, J. H. F. "B. de Maillet, Lamarck und Darwin." *Biologisches Zentralblatt* 32 (1912): 508–518.

Wolf, A. *A History of Science, Technology and Philosophy in the XVIIIth Century.* 2nd ed, edited by Douglas McKie. London: Allen and Unwin, 1952.

Otis Fellows (1967)

MAIMON, SALOMON

(1753–1800)

Born in 1753 in a small village in Lithuania, Shlomo ben-Yehoshua later named himself "Maimon" after the great medieval Jewish philosopher Moses Maimonides. After being married at the age of eleven and fathering a child at fourteen, Maimon left his native country around 1778 in search of "Enlightenment." Following extraordinary adventures as a wandering beggar and scholar, Maimon

arrived in Berlin in March 1780. There he became acquainted with Moses Mendelssohn and his circle. Maimon formulated many of his views (on Judaism and religion in general, and on Spinoza) in overt or covert criticism of Mendelssohn. Until 1791 Maimon contributed to projects of Jewish Enlightenment (*Haskala*), which he wished to promote in the first place through scientific knowledge. Later he became estranged from Jewish affairs.

Maimon's rather coarse way of life, which offended both Jewish ceremonial law and bourgeois decorum, forced him to leave Berlin in 1783. From June 1783 until March 1785 he studied in a German high school in Altona (Hamburg), and improved his knowledge of German and mathematics as he also learned Latin, English, and French. Back in Berlin, Maimon was supported almost entirely by benefactors. By the end of 1789, following praise from Immanuel Kant, Maimon published his first German book, *Versuch über die Tranzscendentalphilosophie* (*An Essay on Transcendental Philosophy*), which is a critical commentary on Kant's *Critique of Pure Reason*. A prolific writer, Maimon produced a number of publications, including both books and journal articles. Almost all of Maimon's works are commentaries of a sort on different writers, a method of philosophizing that is certainly a legacy of his Jewish education. In 1795 Maimon met Graf Adolf Kalkreuth, later himself a philosophical writer, and moved into his house near Berlin, and later to his estate in Silesia. There Maimon died on November 11, 1800.

In his autobiography (1792–1793) Maimon interprets his life as a process of progressive formation (*Bildung*) leading from traditional orthodox Judaism in Lithuania to the center of Enlightenment in Berlin. This change is conceived as "spiritual rebirth" (*Gesammelte Werke* 1:301), a classical term of contemporary Pietism. This work inspired many later autobiographies of European Jews seeking Enlightenment ideals, and it sets the stage for Maimon's own brand of philosophy.

RATIONAL DOGMATISM AND EMPIRICAL SKEPTICISM

In the *Versuch*, Maimon describes his position as "rational dogmatism and empirical skepticism," and despite the oddity of the combination, this is an apt description of his views (*Gesammelte Werke* 2, 432). Maimon follows the rationalists (particularly B. Spinoza and G. W. Leibniz) in granting the principle of sufficient reason unlimited scope: There is nothing inexplicable in the world, and reason's demands are unconstrained. But at the same time, while we can be sure that the principle of sufficient reason in general holds universally, our finitude prevents us from knowing with any certainty whether any *particular* judgment we make about the world accords with this rational condition. As such, whereas rationalism is right about the nature of knowledge, skepticism infects particular knowledge claims.

The exception to this, Maimon claims, is mathematics, in which we can achieve certain knowledge, for here our situation is compared to the "divine": In our mathematical claims we create the contents of mathematical judgments, by constructing *a priori* the objects of geometry and arithmetic. Here we can be assured (although in geometry this is not always the case) that our concepts apply to objects, because the objects themselves are created according to the concepts. But whereas certainty is guaranteed in the field of mathematics, our empirical judgments do not rise to this level, because they can never be shown to possess the "determinable" relation between subject and predicate demanded of "real thought."

DETERMINABILITY AND REAL THOUGHT

According to Leibniz, analytic thought is governed by the law of identity or contradiction: The complete concept of the subject contains all predicates that can be truthfully predicated of it. All true propositions are hence either overtly or "virtually" analytic. Maimon maintains that if there are synthetic judgments *a priori* (as Kant holds), there must also be a principle of such judgments. Since Maimon rejects the thesis that synthesis is the result of the application of the understanding to intuition, he maintains that synthetic thought must have a principle in reason itself. This is his Law of Determinability. The principle distinguishes between the subject that can be thought by itself and the predicate that can be thought only in relation to a subject: It thus permits the synthesis "square table" and excludes "tablish square," because "table" can be thought by itself and the property "square" cannot.

A further, seemingly paradoxical component of the law of determinability is that in a "real synthesis" there is exactly *one* predicate for each subject term. It thus demands for "line" either "straight" or "curved" and excludes "sweet line." Finally, it positively determines that a "real synthesis" is only a synthesis that produces a new object. The hallmark of an object determined through real thought is that new consequences follow from it that flow neither from the subject nor from the predicate terms alone, but only from their synthesis. Thus a trian-

gle has certain "consequences" (e.g., that the sum of its internal angles equals two right angles), whereas the Pythagorean theorem is a further consequence of the synthesis of "triangle" and "right angle."

Real thought then depends on a determinable relation between subject and predicate, and this in turn can be guaranteed only in cases where an object is constructed according to a concept. While this occurs in geometry, the determinable relation cannot be shown to hold in cases in which we are passively given empirical objects through sensibility. Here it is also conspicuous that "real synthesis" is equivalent to the construction of the object itself, and that it proceeds from general to particular concepts. This and the unique relation between subject and predicate imply that if we could generate predications according to the law of determinability, we would be able to (re)construct the *entire* conceptual structure of the world. Because rationalism assumes that complete knowledge exhausts its object, the generation of this conceptual structure would be tantamount to the construction of the world. In mathematics we are hence similar to God (*Gesammelte Werke* 4:42). But this divinity is sharply limited: because proper knowledge consists in such determinable relations that in empirical cases we cannot produce but are merely given, most of what we think of as empirical human knowledge does not in fact deserve the name. Our beliefs about the merely encountered world of objects fail to meet the criteria of real knowledge. This is one source of Maimon's skepticism, which plays a key role in his critique of Kant.

QUID FACTI/QUID JURIS

The difficulty Maimon finds in Kant's views on synthetic judgments a priori centers on two crucial questions that drive the Transcendental Deduction of the Categories in the *Critique of Pure Reason*. There, Kant distinguishes between the *quid facti* (or, the question of fact) and the *quid juris* (or the question of right or warrant) of the use of the pure concepts of the understanding. The first question concerns whether we indeed have certain synthetic judgments due to the application of categories to intuitions, whereas the second asks about our right or justification in doing so. Kant is largely concerned with the second question, because he assumes, according to Maimon, that our experience reveals that we in fact have certain knowledge. But Maimon calls this assumption into question, by challenging the Kantian idea that our experience really involves supposedly objective and necessary claims such as "The sun warms the stone." Kant can assume this to be the case, but this will not convince the skeptic—yet the central argument of the Deduction needs just this supposition, Maimon argues, to establish that the categories are legitimately employed in experience. As a result, only by begging the question "*quid facti*" against a Humean skeptic can Kant's argument succeed.

Maimon also remains suspicious of Kant's answer to the question *quid juris*. Kant aims to show that the legitimate employment of the categories rests on the way in which they can be applied to the intuitive contents of experience delivered by the faculty of sensibility; this depends on his fundamental commitment to a model of experience that distinguishes between intuitions (which are singular and immediate) and concepts (which are general and mediate). Kant's system endorses a kind of cognitive dualism, in which the separate faculties of the understanding and sensibility each contribute distinct and ineliminable elements of cognition. Yet Maimon finds this dualism problematic, for it faces all the challenges and problems that traditionally confront other dualisms such as that between mind and body. For how can wholly separate faculties nonetheless interact in the way that cognition requires? Maimon claims that for this reason Kant's cognitive dualism cannot answer the *quid juris* in a satisfactory manner.

In Maimon's critique of Kant, his allegiances to both skepticism and rationalism come to the fore. The challenge to the *quid facti* draws upon a kind of Humean skepticism about the structure of experience, and calls into question the notion of experience with which Kant's project begins. The critique of the *quid juris* rests upon Maimon's rationalist commitments, for it demands that some sufficient reason or explanation be provided for what Maimon takes to be a wholly mysterious relation between concepts and intuitions. Maimon's challenge to Kant is so interesting and powerful precisely because of his odd brand of skepticism, for it allows him to mount simultaneous attacks on the critical system from both an empiricist and a rationalist position.

MAIMON AND THE TRADITION

Kant famously described Maimon as his most acute critic, and this admiration—even if tinged with occasional acrimony—was shared by a number of other figures in German philosophy. The renown provided by Kant's comments allowed Maimon to engage in conversations and disputes with a number of the leading lights of the day. Maimon corresponded with K. L. Reinhold (and later had a bitter falling out when Maimon published their letters without Reinhold's permission), and penned a series

of pseudonymous responses to the then-anonymous author of *Aenesidemus* (G. E. Schulze). In all of these works Maimon pressed his version of empirical skepticism and rational dogmatism.

Maimon's most lasting influence, however, was on J. G. Fichte, who shared Kant's respect for Maimon's intellect, and who saw more clearly than others the threat that Maimon's position posed for Kant's philosophy. Fichte's formulations of his *Wissenschaftslehre* in large part stand as attempts to meet the challenge Maimon posed to Kant, in particular to answer the charge that a dualistic model of cognition cannot explain how its disparate elements interact. Fichte's solution—which turns on rejecting Kant's model of cognition in favor of the positing activity of the Absolute-I—marked the beginning of Absolute Idealism, which reached its fruition in Schelling and Hegel. Maimon himself was certainly no Absolute Idealist—in fact, in his correspondence with Fichte he distances himself from Fichte's project—but his challenge to Kant provided an important goad in the development of Fichte's *Wissenschaftslehre*, and through him the systems of Schelling and Hegel. Thus Maimon's "dogmatic rationalism" found successors but neither his "empirical skepticism" nor his unique combination of both attracted adherents.

But Maimon's combination of skepticism and rationalism is of interest not simply as a historical step on the road from Kant to Hegel, but as a fascinating and often compelling position in its own right. Maimon's skepticism is unique in being based not upon a suspicion of the claims of rational inquiry, but, perhaps paradoxically, on an uncompromising commitment to the demands of reason. The challenge Maimon poses to all accounts of cognition is to explain how the understanding can apply to the contents of sensibility (whether *a priori* or *a posteriori*). Maimon ultimately resorts to a skeptical answer to this question, yet a nonskeptical response to the challenge he presents is something contemporary theories of cognition continue to struggle to meet.

See also Epistemology.

Bibliography

WORKS BY MAIMON

Gesammelte Werke. 7 vols., edited by Valerio Verra Hildesheim, Germany: Georg Olms, 1965–1976.

Versuch über die Transzendentalphilosophie (1790), edited by Florian Ehrensperger. Hamburg, Germany: Felix Meiner Verlag, 2004.

WORKS ABOUT MAIMON

Atlas, Samuel. *From Critical to Speculative Idealism: The Philosophy of Solomon Maimon*. The Hague: Martinus Nijhoff, 1964.

Beiser, Frederick. *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, MA: Harvard University Press, 1987.

Bergmann, Samuel. *The Philosophy of Salomon Maimon*. Translated by Noah Jacobs. Jerusalem: Magnes Press, 1967.

Bransen, Jan. *The Antinomy of Thought: Maimonian Skepticism and the Relation between Thoughts and Objects*. Dordrecht, Netherlands: Kluwer, 1991.

Engstler, Achim. *Untersuchungen zum Idealismus Salomon Maimons*. Stuttgart-Bad Cannstatt, Germany: Frommann-Holzboog, 1990.

Freudenthal, Gideon, ed. *Salomon Maimon: Rational Dogmatist, Empirical Skeptic*. Dordrecht, Netherlands: Kluwer, 2003.

Gueroult, Martial. *La philosophie transcendental de Salomon Maimon*. Paris: Vrin, 1929.

Kuntze, Friedrich. *Die Philosophie Salomon Maimons*. Heidelberg, Germany: Carl Winter, 1912.

Gideon Freudenthal (2005)
Peter Thielke (2005)

MAIMONIDES
(1135–1204)

Maimonides was the most celebrated Jewish philosopher of the Middle Ages. "Maimonides" is the Latinized cognomen of Moses son of Maimon. Also called RaMBaM, the acronym for Rabbi Moses ben Maimon, he was born in Córdoba, which belonged at that time to Muslim Spain. His father, Maimon son of Joseph, was a distinguished scholar versed in traditional Jewish lore. At the age of thirteen, Maimonides left his native town after it was conquered by the army of the Almohads, an intolerant Muslim sect. After various journeys he and his family settled in northern Africa, under the oppressive rule of the Almohads. In 1165 they went to Egypt, where Maimonides became a court physician and leader of the Jewish community. He died in Cairo.

Maimonides was and is regarded as an outstanding authority on Jewish religious law, the Halachah. His writings in this field include a commentary in Arabic on the Mishnah that contains a treatise on ethics known as "Eight Chapters" and a list of the thirteen fundamental dogmas of the Jewish faith as established by Maimonides; another of these works, known under the two titles *Mishnah Torah* and *Yad Hazakah,* is a voluminous codification of the Law written in Hebrew, whose first portion, the "Book of Knowledge," expounds a system of religious beliefs and is markedly influenced by philosophy.

The fact that a considerable portion of Maimonides' activity was devoted to legal doctrine is by no means irrelevant in a consideration of his philosophical attitude. In a sense this was a practical activity that can be assimilated to that of a statesman; it was accordingly consonant with Maimonides' Platonizing contention that certain superior individuals are able to combine a mode of existence given over to contemplation and intellection with a life of action.

Maimonides also wrote several medical treatises in Arabic. One of them, known as *Moses' Chapters* (*Fuṣūl Mūsā*), contains a critique of Galen, part of which deals with the Greek physician's animadversions on the Law of Moses. He also composed two popular tracts, "Treatise on Resurrection" and "Epistle to Yemen," the latter treatise rebutting the claims of a pseudo Messiah who had appeared in Yemen. Maimonides is also the author of one philosophical treatise on logic, composed in his early youth.

GUIDE OF THE PERPLEXED

Maimonides' reputation as a philosopher rests squarely upon his *Guide of the Perplexed* (*Dalālat al-Hāirīn* in Arabic), a work that its author did not regard as being of a philosophical nature. The "perplexed" to whom the *Guide* is supposed to have been addressed are men who are well grounded in the Jewish religious tradition and have some knowledge of certain philosophical sciences; the disciple to whom Maimonides addresses the "Introductory Epistle" at the beginning of the *Guide* is said to be conversant with logic and mathematics but not with physics or metaphysics. These semi-intellectuals are regarded by Maimonides as being in a state of mental confusion because they consider that the theses of the Greek sciences contradict religious faith. The word *hayra,* "perplexity," which is connected with the participle *hā'irīn* figuring in the title of the work under discussion, appears to have served as a technical term denoting the state of mind induced by a tug of war between two opposed beliefs. Both al-Farabi and, in the generation before Maimonides, the Jewish philosopher Abraham ibn Da'ud also used the term *perplexed* to describe people who hesitate between the conflicting claims of philosophy and religion. In one passage of the *Guide* Maimonides seems to indicate that his purpose in writing the work was to help such of the perplexed as were endowed with the requisite intellectual capacities to achieve a full knowledge of philosophical truths without giving up the observance of the religious commandments.

Maimonides, however, like his contemporary Averroes, was convinced that philosophy could constitute a terrible threat to the social fabric if a vulgarized version of its doctrines were to spread among ordinary people and destroy simple faith in authority. Systematic treatises, giving a step-by-step account of the Aristotelian doctrines, avoided this danger through recourse to technical terms and logical argumentation, which were incomprehensible to noninitiates. Maimonides employed another method, set forth in his introduction to the *Guide.* In the case of this work his very considerable gift for literary composition, which had enabled him to succeed in the extremely difficult task of producing a well-ordered code comprising the whole of Talmudic law, was called upon to disarrange and make a jumble of the systematic expositions of Aristotle and the Aristotelians. Maimonides makes it quite clear that in order to make understanding more difficult, he carefully tore apart conceptions that belong together. The reader is thus faced with the challenge of reconstructing the original whole out of pieces dispersed in various portions of the *Guide.* Maimonides even states that on certain points he deliberately makes two contradictory assertions. These and other precautions, which were intended to confuse readers of insufficient intellectual caliber or preparation, have turned the *Guide* into an enigma; any solution of the enigma can be impugned by an appeal to some statement of Maimonides' that may or may not have been meant to be taken at its face value.

INFLUENCES ON MAIMONIDES

There is a question whether the *Guide* was meant to be an apologetic attempt to render religion intellectually respectable by exposing the limitations of human reason, beyond which lies the domain of faith in things that may be true although they are unknown to philosophers; or, alternatively, whether it was meant to demonstrate that religion has a purely practical use. If the latter, then Maimonides meant to say that theoretical truth is essentially, although perhaps not completely, revealed by philosophy and to deny that religion has anything to offer except, in the most favorable cases, myths and parables to be interpreted with the help of scientific knowledge. A knowledge of the philosophical authors whose influence was avowed by Maimonides or may be discerned in his work may help to determine what actually was the main object of the *Guide.*

In a letter to Samuel ibn Tibbon, who translated the *Guide* into Hebrew, Maimonides wrote that he considered Plato's writings to be superseded by those of Aristotle, which are the root and foundation of all philosophy.

Nevertheless, he thought that Aristotle should be studied only with the help of the commentators Alexander of Aphrodisias, Themistus, and Averroes (a contemporary of Maimonides, who was not acquainted with the Muslim philosopher's commentaries at the time the *Guide* was written). Maimonides esteemed al-Farabi above all the other Islamic philosophers (a typical attitude of the philosophers of Spain), and also praised Ibn Bājja, the Muslim Spanish Aristotelian. His reaction to Avicenna, who was the dominant philosophical influence in the Islamic East, was ambivalent.

Maimonides does adopt certain conceptions of Avicenna's. Thus, his view that existence is an accident derives from Avicenna's fundamental doctrine that essences per se are neutral with respect to existence, which supervenes on them as an accident. However, in points that have an obvious bearing on religious beliefs, Maimonides sometimes does not hesitate to prefer Aristotelian notions, although they appear to be incompatible with the Jewish tradition prevalent in his time, to views that are more easily reconcilable with this tradition and that, through Avicenna's adhesion, were given the hallmark of philosophical respectability. To cite an outstanding example, Maimonides holds no brief for Avicenna's opinion that the individual human soul survives the death of the body and is immortal. Like Alexander of Aphrodisias and other Aristotelians, he considers that in man only the actual intellect—which lacks all individual particularity—is capable of survival. In adopting this view, Maimonides clearly shows that, at least on this point, he prefers the philosophical truth as he sees it, however opposed it may seem to be to the current religious conceptions, to the sort of halfway house between theology and philosophy which, in the severe judgment of certain Spanish Aristotelians—notably Averroes—Avicenna had sought to set up.

To cite another instance, Maimonides does not give the slightest indication of recognizing, as Avicenna did, the mystical ecstatic way to God as being on the same level as the way of the intellect (the Muslim philosopher may have claimed even more for it than simple equality). According to the *Guide,* the religious commandment enjoining the love of God entails the duty of knowing whatever may be known of him, for love is proportionate to the knowledge man has of the beloved.

THEORY OF DIVINE ATTRIBUTES

What kind of cognition of God is possible to man? The *Guide* sets forth at considerable length and with stronger emphasis than in Avicenna the doctrine of negative theology. According to this doctrine, nothing positive can be known about God, who has nothing in common with any other being. No predicate or descriptive term can legitimately be applied to him unless it is given a meaning that is wholly different from the one the term has in common usage and is purely negative. All statements concerning God considered in himself should, if they are to be regarded as true, be interpreted as providing an indication of what God is *not.* This applies even to the statement that God exists. Maimonides maintains that progress in this kind of negative knowledge is of considerable value, for it does away with false ideas concerning God.

On the other hand, the positive knowledge that man is capable of is concerned with quite a different domain; it deals not with God in himself but with his governance of nature, or, in other words, with the order obtaining in the cosmos and determining the events that occur in it. According to Maimonides' interpretation of Exodus 33, only this knowledge is granted to Moses, and such are the limitations of human science. As far as this conception is concerned, the acts of God may be identified with the operations of nature (or with historical happenings brought about by natural causes). Maimonides' view of the world being by and large Aristotelian, these operations are subject to the rule that they do not destroy but, rather, safeguard the perpetuity of the immutable order of nature, including the preservation of humankind and of the various other species of living beings.

Some of the operations of God (or of nature) seem, from the human point of view, to be beneficent, for instance, the operation that instills into progenitors the impulse to care for their young; others, such as earthquakes or large floods, seem destructive. Because of the anthropomorphic tendency, men witnessing happenings of the first kind speak of God as being merciful and may impute havoc and death to God's being vengeful. These are two of the so-called divine attributes of action. Quite evidently they are not concerned with the essence of God but reflect a purely human evaluation of God's, or nature's, actions. In contrast with other medieval Aristotelian philosophers, Maimonides does not recognize the divine attributes of relation.

DIVINE INTELLECTION

As the Aristotelian system of physics requires, and as Maimonides demonstrates by means of a number of proofs taken over from earlier philosophers, this world is dependent upon God (who is the Prime Mover); but, contrary to Aristotle's conception (already modified by

some of the late Greek Neoplatonists, whose views reached Maimonides through the Islamic philosophers), God is regarded as the efficient and formal as well as the final cause of the cosmos. This God is pure intellectual activity, to which (in Maimonides' view as well as in Aristotle's) man's intellection bears a certain resemblance. Indeed, Maimonides seems to go out of his way to point out this similarity. In this connection a comparison between a statement of his and one of al-Farabi's is instructive. In accordance with the doctrine of Book A of Aristotle's *Metaphysics,* the Muslim philosopher states quite unequivocally that it is because God intellects only himself that the subject, object, and act of divine intellection are identical. Maimonides, too, maintains this threefold identity with regard to God (*Guide,* Part I, Ch. 68); but he points out that it exists equally in the case of man's intellection of any object, for instance, a piece of wood, because according to an opinion of Aristotle, the actual intellect is identical with the object cognized by it. (This opinion was apparently quite unconnected with Aristotle's conception of God.) This comparison of man's cognition to God's, which argues similarity between the two, appears to be incompatible with Maimonides' negative theology. This point had already been made in the Middle Ages and must be taken into account in any interpretation of the *Guide.*

Furthermore, the fact that Maimonides uses as an example the intellection of a piece of wood seems to suggest that, unlike Aristotle and al-Farabi but in accordance with many of the medieval Aristotelians, he tends to believe that God cognizes not only himself but all the intelligibles. Since cognition involves identity, this conception would appear to entail the identification of God with the intelligible structure of the universe, regarded both as the subject and as the object of cognition. The argument does not entail the identification of matter with God or with an attribute of the Deity. To call Maimonides' position or its logical corollaries "pantheism" would therefore be to go beyond the evidence.

ORIGIN OF THE WORLD

A main theme of the *Guide* concerns the contradiction between the idea of God upon which Judaism is founded and the philosophical view of God. The philosophical view for Maimonides is the conception of God as an intellect rather than as described by the speculations of negative theology. Maimonides is fully aware of the crucial character of the issue and of the impossibility of achieving a true reconciliation between the philosophical and the religious points of view. He remarks in the *Guide* (Part II, Ch. 20): "For to me the combination between [the world] existing in virtue of necessity and being produced in time in virtue of a purpose in the world … comes near to being a combination of two contraries." Maimonides points out the "very disgraceful conclusions" that follow from the first opinion:

> Namely it would follow that the Deity, whom everyone who is intelligent recognises to be perfect in every kind of perfection, could as far as all beings are concerned, produce nothing new in any of them; if He wished to lengthen a fly's wing or shorten a worm's foot, He would not be able to do so. But Aristotle would say that He would not wish it and that it is impossible to will something different from what is; that it would not add to His perfection, but would perhaps from a certain point of view be a deficiency. (*Guide,* Part II, Ch. 22)

In Maimonides' interpretation of the Aristotelian position, God's will is assimilated to the divine Intellect, which is identical with God himself, and the world may be regarded as something like an intellection necessarily produced by this Intellect. A consequence of Aristotle's theory as understood by Maimonides is that every characteristic of things existing in the world must be supposed to have a cause grounded in the natural structure of the universe (as opposed to a supernatural cause not determined by this structure). It may be added that as far as bodies are concerned, Maimonides seems to believe that in cases in which a mechanistic explanation can be found, it might provide such a cause. If this were accepted, it would mean that no part of the natural order could be, or could ever have been, different from what it actually is, for its existence is guaranteed by the immutability of divine reason. In other words, the world could not have been created in time.

From this point of view Maimonides is quite consistent in describing temporal creation as the greatest of miracles and in stating that if this is admitted, the intellectual acceptance of other direct interventions of God in the natural course of events does not present any difficulties. Since it serves Maimonides' purpose to make out the best case possible for what he designates as the religious conception of God, he attempts to show that a structure of the universe that is necessary, because it is rationally determined in every respect, does not exist—or at least he seems to do so. In fact, he does not go beyond the demonstration, made at some length, that as far as the heavenly spheres are concerned, Aristotelian physics (although it gives satisfactory explanation of the phenomena of the

sublunar world) is incapable of propounding a comprehensive scientific theory that can be regarded as certain and that provides cogent proof for the assumption that the cosmic order could not be different from what it actually is. In this critique of Aristotle's celestial physics he is helped by the much-debated discrepancy that exists between Aristotle's natural science and the Ptolemaic system.

Maimonides also puts forward an argument of somewhat different character. He points out that man's knowledge of the order of nature is based on the empirical data of which he is cognizant. It is, however, conceivable that the existence of the data that are known to man had a beginning in time. No man who studies this problem should ignore this possibility, for if he does so, his case would be analogous to that of a person who disbelieves on empirical grounds—because he has met only adults—that human beings are brought into the world through birth after having been embryos.

Maimonides' critique of the inconsistencies and the insufficiency of the Aristotelian physics is pertinent within its scheme of reference. However, the doctrine of the eternity of the world does not rest exclusively upon physical theory. It is also corollary to the conception of God as Intellect, and Maimonides is aware of this. It is certainly significant, and it may be a deliberate omission, that when Maimonides is dealing with the problem of the eternity of the world in the *Guide,* he does not mention this conception although other portions of the work prove he had adopted it. Thus he does not allude to God as Intellect when he proclaims in the *Guide* (Part II, Ch. 25) that he does not accept the doctrine of the eternity of the world for two reasons: (1) because it has not been demonstrated; (2) because its adoption would be tantamount to destroying the foundations of the Law, for it would mean denying the claims of the prophets and rejecting the belief in miracles.

SOURCES OF KNOWLEDGE

That Maimonides rejected the doctrine of the eternity of the world partly because (as his second reason) it would have destroyed the foundations of religious law may appear to affirm the claim of religious belief to have a decisive voice in theoretical questions that are of paramount concern to it. That is, it may appear to affirm this claim, provided that the intellect is unable to reach a fully demonstrable conclusion with regard to the moot points. Clearly such a claim can have far-reaching implications. It could be argued that this position leads to the recognition of suprarational theoretical truths or, alternatively, to the assertion of validity of conclusions in the sphere of theory adopted only on the basis of practical reason. Maimonides himself, however, does not at all countenance such a demotion of theoretical reason. In the *Guide* (Part I, Ch. 2) he explains the superiority of theoretical reason, which is concerned with the difference between truth and falsehood, over practical reason, which deals with the distinction between good and evil. His allegorical interpretation of Adam's fall entails the conclusion that practical reason has the comparatively lowly function of curbing the appetite to which man is prone when he is not given over to theoretical contemplation.

As for prophecy and divine revelation, they cannot be regarded as sources of supraintellectual knowledge conceived as being independent of, and superior to, the system of sciences produced by theoretical reason. This comes out clearly in Maimonides' description of the characteristics peculiar to prophets. According to him, prophets must have both an outstanding intellectual capacity and an outstanding imaginative capacity. Given these two preconditions, and suitable conduct, prophecy is a natural phenomenon; the gift of prophecy can be withheld from a person having the required qualifications only by means of a miracle. The intellectual capacity of prophets is similar at least in kind to that of the philosophers; it enables them to receive what Maimonides terms a "divine overflow," an influx coming from the Active Intellect, which, according to the interpretation of the Aristotelian doctrine adopted by Maimonides, brings about the actualization of man's potential intellect. The Active Intellect is the last of the ten incorporeal Intellects; its special sphere of action is the sublunar world.

There is no suggestion that the conclusions reached by the prophets through the use of the intellect are in any way different from those of the philosophers, though the prophets may reach them more rapidly; all prophets are philosophers. This clearly applies also to Moses, in spite of a statement in the *Guide* that none of the author's assertions about the prophets pertain to Moses. In other writings Maimonides describes Moses as having attained union with the Active Intellect; according to the conception of certain Islamic Aristotelians, union with the Active Intellect represents the highest goal and is reached by the great philosophers.

Imagination is inferior to intellect for Maimonides, who was on this point an orthodox Aristotelian. Imagination enables the prophet to see veridical dreams and visions, for the divine overflow spills over from the intellectual to the imaginative sphere. But it certainly does not

give access to a supraintellectual truth. In fact the superiority of Moses over all other prophets is, according to Maimonides' interpretation, partly the result of the circumstance that in his prophecy he did not have recourse to imagination.

POLITICAL PHILOSOPHY

Religious revelation thus does not procure any knowledge of the highest truth that cannot be achieved by the human intellect; it does, however, have an educative role—as well as a political one. In Maimonides' words, "The law as a whole aims at two things: the welfare of the soul and the welfare of the body" (*Guide,* Part III, Ch. 27).

Because of the great diversity of human character, a common framework for the individuals belonging to one society can be provided only by a special category of men endowed with the capacity for government and for legislation. Those who have only a strong imagination, unaccompanied by proportionate intellectual powers, are not interested in the intellectual education of the members of the state which they found or govern. On the other hand, the foremost example of an ideal lawgiver is Moses.

The law instituted by Moses had to take into account the historical circumstances—the influence of ancient Oriental paganism—and had to avoid too great a break with universal religious usage. To cite one example, sacrifices could not be abolished, because this would have been an excessively violent shock for the people. In spite of these difficulties, however, Moses succeeded in establishing a polity to which Maimonides, in the "Epistle to Yemen," applies the term *al-madīna al-fādila* ("the virtuous city") used by the Muslim philosophers to designate the ideal state of Plato's *Republic*—a work that, perhaps mainly through the mediation of al-Farabi, had a considerable impact on Maimonides' political thought.

MORAL PHILOSOPHY

The polity is not alone in regulating men's actions in the best possible way. The Scriptures by which the polity is ruled also contain hints that may guide such human individuals as are capable of understanding its hints to philosophical truths. Some of these truths are to be discovered in the beliefs taught to all those who profess Judaism; these dogmas are for evident reasons formulated in a language adapted to the understanding of ordinary unphilosophical people. There are, however, other religious beliefs that, although they are not true, are necessary for the majority of the people, to safeguard a tolerable public order and to further morality. Such are the belief that God is angry with those who act in an unjust manner and the belief that he responds instantaneously to the prayer of someone wronged or deceived (*Guide,* Part III, Ch. 28). The morality suited to men of the common run aims at their exercising a proper restraint over the passions of the appetite; it is an Aristotelian middle-of-the-road morality, not an ascetic one. The ascetic overtones that are occasionally encountered in the *Guide* concern the philosopher rather than the ordinary man.

There is a separate morality for the elite, which is or should be called upon to rule, to which Maimonides alludes in the *Guide* (Part I, Ch. 54; Part III, Chs. 51 and 54). This ethical doctrine is connected with Maimonides' interpretation of what ought to be man's superior goal, which is to love God, and, as far as possible, to resemble him.

From the point of view of negative theology, love of God can be achieved only through knowledge of divine activity in the world, the only knowledge of God possible. This supreme goal can be reached through a study of natural science and of metaphysics, which appears to signify that the highest perfection can be attained only by a man who leads the theoretical life—the man whose superiority was proclaimed by Aristotle. However, Maimonides is at pains to show—and this seems to be a Platonic element in his doctrine—that the theoretical life can be combined with a life of action, as proved by the examples of the patriarchs and of Moses.

What is more, a life of action can constitute an imitation of God. For the prophetic legislators and statesmen endeavor to imitate the operations of nature, or God (the two are equivalent; the expression "divine or natural actions," which occurs in the *Guide,* may have been in Benedict de Spinoza's mind when he first spoke of *Deus sive natura*). Maimonides emphasizes two characteristics that belong both to the actions of God-nature and to the actions of superior statesmen. First, however beneficent or destructive—or, in ordinary human parlance, however merciful or vengeful—the actions in question appear to be, neither God nor the prophetic statesman is actuated by passions. Second, the activity of nature (or God) tends to preserve the cosmic order, which includes the perpetuity of the species of living beings, but it has no consideration for the individual. In the same way the prophetic lawgivers and statesmen, who in founding or governing a polity should imitate this activity, must have in mind first and foremost the commonweal, the welfare of the majority, and must not be deterred from following a politically correct course of action by the fact that it hurts individuals.

The imitation of the works of God (or of nature) by the prophets means (*Guide,* Part III, Ch. 32) that the prophets imitate in leadership the indirect and complicated way through which nature obtains its desired results, as seen, for instance, in the extremely intricate mechanism of living organisms. Maimonides calls this indirect method a "gracious ruse" of God and his wisdom; he may have taken the expression over from Alexander of Aphrodisias's work "Principle of the All" (extant only in Arabic translation). It is reminiscent, not only on the verbal plane, of G. W. F. Hegel's "Cunning of Reason." According to the *Guide,* Moses used the indirect method in making the sons of Israel wander for forty years in the desert instead of leading them straight to the land of Canaan, for he wanted the people to shed slavish habits and acquire in the hard school of the desert the warlike virtues necessary for conquest. He also used it in adapting the commandments to the historical and geographical circumstances.

INFLUENCE OF THE *GUIDE*

The *Guide* was first translated into Hebrew in Maimonides' lifetime, by Samuel ibn Tibbon and a little later by al-Harizi. Its first translation into Latin was also produced in the thirteenth century. Maimonides' injunction to follow his example in writing the Arabic text of the work only in Hebrew characters (and thus to prevent its being read by non-Jews) was not always observed. The work is mentioned by some later Muslim writers but does not appear to have had more than a very slight impact on Muslim thought.

In the period after Maimonides the *Guide* was the fundamental text of medieval Jewish thought and was much debated. In the thirteenth and fourteenth centuries it was violently denounced for being antireligious and as vehemently defended against this charge; commentaries upon it were written by Shem-Tov Falaquera, Joseph ibn Kaspi, Moses of Narbonne, Isaac Abravanel, and others, and its theses are discussed at length in such capital philosophical works as Gersonides' *Milhamot Adonai* (The wars of the Lord) and Hasdai Crescas's *Or Adonai* (Light of the Lord). At first blush it is therefore rather surprising that among Jewish philosophers, relatively few of Maimonides' disciples have been content to adopt his apparently agnostic attitude toward fundamental metaphysical problems and thus to leave what he believed to be a necessary loophole for religious belief. In fact, no doubt partly because of the unsystematic mode of exposition of the *Guide,* some philosophically minded commentators (notably Moses of Narbonne) expounded Averroes's conceptions rather than Maimonides' in their commentaries on the *Guide.* Other commentators—for example, Abravanel—often criticized him from a traditionalistic religious point of view.

The *Guide* had a strong influence on later Jewish philosophers, many of whom owe their introduction to philosophy to the *Guide.* This can be seen in Spinoza (a considerable portion of the *Tractatus Theologico-politicus* is devoted to a critique of Maimonides, although the explicit references to him are few) and in Salomon Maimon, who wrote a commentary on the *Guide.*

The influence of Maimonides on the medieval Christian Schoolmen seems to have been considerable; the matter has not yet been sufficiently investigated, though several studies dealing with the subject do exist. It may be noted that by elaborating the doctrine of suprarational truths the systems of Thomas Aquinas and of other Scholastics found a way of legitimating from a theoretical point of view Maimonides' decision to opt for the belief in temporal creation, because the existence of religion hinged on this belief's being generally accepted.

See also Alexander of Aphrodisias; al-Fārābī; Averroes; Avicenna; Crescas, Hasdai; Ethics, History of; Galen; Gersonides; Hegel, Georg Wilhelm Friedrich; Ibn Bājja; Jewish Philosophy; Maimon, Salomon; Medieval Philosophy; Plato; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.

Bibliography

WORKS BY MAIMONIDES

Le guide des égarés. Edited by Salomon Munk, 3 vols. Paris: A. Franck, 1856–1866. Arabic text and French translation, with many detailed notes. The French translation has been reedited. Paris: G.-P. Maisonneuve, 1960.

The Guide of the Perplexed. Translated with an introduction and notes by Shlomo Pines. Chicago: University of Chicago Press, 1963. Introductory essay by Leo Strauss.

WORKS ON MAIMONIDES

Altmann, Alexander. "Das Verhältnis Maimunis zur jüdischen Mystik." *Monatsschrift für Geschichte und Wissenschaft des Judentums* 80 (1936): 305–330.

Baron, Salo, ed. *Essays on Maimonides: An Octocennial Volume.* New York: Columbia University Press, 1941.

Diesendruck, Z. "Maimonides' Lehre von der Prophetie." In *Jewish Studies in Memory of Israel Abrahams,* edited by G. A. Kohut, 74–134. New York: Press of the Jewish Institute of Religion, 1927.

Diesendruck, Z. "Die Teleologie bei Maimonides." *Hebrew Union College Annual* 5 (1928): 415–534.

Epstein, I., ed. *Moses Maimonides: 1135–1204.* London: Soncino Press, 1935.

Guttmann, Jakob. *Der Einfluss der Maimonideschen Philosophic auf das christliche Abendland.* Leipzig, 1908.

Rohner, A. *Das Schöpfungsproblem bei Moses Maimonides, Albertus Magnus, und Thomas von Aquin.* Münster: Aschendorff, 1913.

Roth, Leon. *The Guide for the Perplexed, Moses Maimonides.* London, 1948.

Strauss, Leo. *Persecution and the Art of Writing.* Glencoe, IL: Free Press, 1952. Includes "The Literary Character of *The Guide for the Perplexed,*" also published in Baron's *Essays on Maimonides* (see above).

Strauss, Leo. *Philosophic und Gesetz.* Berlin: Schocken, 1935.

Strauss, Leo. "Quelques Remarques sur la science politique de Maimonide et de Farabi." *Revue des études juives* 100 (1936): 1–37.

Wolfson, H. A. "Hallevi and Maimonides on Design, Chance, and Necessity." *Proceedings of the American Academy for Jewish Research* 11 (1941): 105–163.

Wolfson, H. A. "Hallevi and Maimonides on Prophecy." *Jewish Quarterly Review,* n.s., 32 (1941–1942): 345–370, and n.s., 33 (1942–1943): 49–82.

Wolfson, H. A. "Maimonides and Halevi." *Jewish Quarterly Review,* n.s., 2 (1911–1912): 297–337.

Wolfson, H. A. "Maimonides on Negative Attributes." In *Louis Ginzberg Jubilee Volume,* edited by A. Marx et al., 419–446. New York: American Academy for Jewish Research, 1945.

Wolfson, H. A. "Maimonides on the Internal Senses." *Jewish Quarterly Review,* n.s., 25 (1934–1935): 441–467.

Wolfson, H. A. "The Platonic, Aristotelian, and Stoic Theories of Creation in Hallevi and Maimonides." In *Essays in Honour of the Very Rev. Dr. J. H. Hertz,* edited by Isidore Epstein, Ephraim Levine, and Cecil Roth, 427–442. London: E. Goldston, 1942.

Shlomo Pines (1967)

MAIMONIDES [ADDENDUM]

Since Shlomo Pines's entry, scholars have come to accept 1138, not 1135, as the year of Maimonides' birth. Some scholars also believe that the youthful treatise on logic (*Millot ha-Higayon*) is not by Maimonides. The major development in Maimonidean studies, however, is an interpretive one. Pines worked closely with Leo Strauss on the 1963 English translation of Maimonides' *Guide of the Perplexed*, which remains the best complete English version of his philosophical magnum opus. Strauss, who wrote the introductory essay to the translation, had an idiosyncratic way of reading many premodern thinkers, including Maimonides. In brief, Strauss understood Maimonides to be engaged in a vast project of deception, of concealing his real beliefs, in order that those incapable of understanding and accepting them not become perplexed and dislodged from their simple pieties.

Strauss's way of reading Maimonides finds its way into this article when Pines suggests that Maimonides was a closet Aristotelian who (really) believed in the eternity of the world. Never mind that Maimonides says the opposite to this; for the Straussian, this is just the point: to conceal one's real beliefs, and to suggest the opposite from what one explicitly argues for. There are still Straussian interpreters and interpretations, but they are in retreat. Philosophical scholars tend to rest content with mulling over the actual arguments that Maimonides presents. Further, in response to the Straussian position that there exists a deep divide between philosophy and the law (religion), between Athens and Jerusalem, recent scholars such as Isadore Twersky (1967) and David Hartman (1976) argue that, on the contrary, Maimonides grounds philosophy in the law and understands the law as subserving in large part suprapolitical ends.

Scholars seem less taken with the Maimonidean reaction to Avicenna (Ibn Sīnā) than Pines appears to be. The Islamic thinkers who have more recently emerged as significant for Maimonides are al-Fārābī and Ibn Bājja (Avempace). They tend to be important for their influence on Maimonides' moral and political theorizing. Pines is still good on Maimonides' practical philosophy. Especially to be noted is his insistence on a Platonic element in his view of the summum bonum. Often Maimonides is presented as endorsing Aristotle's view that human happiness is a function of contemplative activity alone. Pines rightly resists this, noting that Moses, the political prophet, is paradigmatic for Maimonides. Indeed, the end of the *Guide* makes clear that imitation of God mirrors God's providential care for the created world.

See also al-Fārābī; Aristotelianism; Aristotle; Avicenna; Ibn Bājja; Platonism and the Platonic Tradition.

Bibliography

WORKS BY MAIMONIDES

Crisis and Leadership: Epistles of Maimonides, edited by David Hartman. Translated by Abraham Halkin. Philadelphia: Jewish Publication Society, 1985.

Ethical Writings of Maimonides, edited and translated by Raymond L. Weiss and Charles E. Butterworth. New York: New York University Press, 1975.

The Guide of the Perplexed. Abridged edition, edited by Julius Guttmann. Translated by Chaim Rabin. London: East and West Library, 1952. Reprinted with a new introduction by D. H. Frank. Indianapolis: Hackett, 1995.

A Maimonides Reader, edited by Isadore Twersky. New York: Behrman House, 1972.

WORKS ON MAIMONIDES AND HIS PHILOSOPHY

Burrell, David. *Knowing the Unknowable God: Ibn Sina, Maimonides, Aquinas*. Notre Dame, IN: University of Notre Dame Press, 1986.

Cohen, Robert, and Hillel Levine, eds. *Maimonides and the Sciences*. Dordrecht, Netherlands: Kluwer, 2000.

Fox, Marvin. *Interpreting Maimonides: Studies in Methodology, Metaphysics, and Moral Philosophy*. Chicago: University of Chicago Press, 1990.

Hartman, David. *Maimonides: Torah and Philosophic Quest*. Philadelphia: Jewish Publication Society, 1976.

Kellner, Menachem. *Maimonides on Human Perfection*. Atlanta: Scholars Press, 1990.

Kellner, Menachem. *Maimonides on Judaism and the Jewish People*. Albany: State University of New York Press, 1991.

Kellner, Menachem. *Maimonides on the Decline of the Generations and the Nature of Rabbinic Authority*. Albany: State University of New York Press, 1996.

Kraemer, Joel, ed. *Perspectives on Maimonides: Philosophical and Historical Studies*. Oxford: Littman Library, 1991.

Lachterman, David. "Maimonidean Studies 1950–86: A Bibliography." *Maimonidean Studies* 1 (1990): 197–216.

Langermann, Yitzhak Z. "The Mathematical Writings of Maimonides." *Jewish Quarterly Review* 75 (1984): 57–65.

Leaman, Oliver. *Moses Maimonides*. London: Routledge, 1990.

Leibowitz, Yeshayahu. *The Faith of Maimonides*. New York: Adama Books, 1987.

Manekin, Charles. *On Maimonides*. Belmont, CA: Wadsworth, 2004.

Pines, Shlomo, and Yirmiyahu Yovel, eds. *Maimonides and Philosophy*. Dordrecht, Netherlands: Nijhoff, 1986.

Seeskin, Kenneth. *Maimonides: A Guide for Today's Perplexed*. West Orange, NJ: Behrman House, 1991.

Seeskin, Kenneth. *Searching for a Distant God: The Legacy of Maimonides*. New York: Oxford University Press, 2000.

Stern, Josef. *Problems and Parables of Law: Maimonides and Nahmanides on Reasons for the Commandments*. Albany: State University of New York Press, 1998.

Strauss, Leo. *Philosophy and Law: Contributions to the Understanding of Maimonides and His Predecessors*. Albany: State University of New York Press, 1994. Originally published as *Philosophie und Gesetz*. Berlin: Schocken, 1935.

Twersky, Isadore. *Introduction to the Code of Maimonides (Mishneh Torah)*. New Haven, CT: Yale University Press, 1980.

Twersky, Isadore. "Some Non-Halakic Aspects of the *Mishneh Torah*." In *Jewish Medieval and Renaissance Studies*, edited by A. Altmann, 95–118. Cambridge, MA: Harvard University Press, 1967.

Twersky, Isadore, ed. *Studies in Maimonides*. Cambridge, MA: Harvard University Press, 1991.

Weiss, Raymond. *Maimonides' Ethics: The Encounter of Philosophic and Religious Morality*. Chicago: University of Chicago Press, 1991.

Daniel H. Frank (2005)

MAINE DE BIRAN

(1766–1824)

Maine de Biran, the French statesman and philosopher, was born Marie François Pierre Gonthier de Biran, receiving the name "Maine" from the name of his family's property (le Maine). He attended the *collège* at Périgueux, dominated by the secular, moderate constitutional Royalists called *Doctrinaires*, and excelled there in mathematics. In 1784 he joined the king's guard and in 1789 was wounded defending Louis XVI in a mob uprising. To escape the Reign of Terror, he retired to his estate in 1793 and began intensive psychological and philosophical investigations. In 1797 he was elected to the Council of Five Hundred, and this election of a moderate royalist was a symptom of the beginning of the end of the Reign of Terror. This post and other public duties did not keep him from reaping the fruits of his earlier meditations. He became acquainted with the *Idéologues* Pierre-Jean Georges Cabanis and Comte Destutt de Tracy by winning first prize in an essay contest sponsored by the Institute of France with the essay *L'influence de l'habitude sur la faculté de penser* (*The Influence of Habit on the Faculty of Thinking*). He won membership in the institute in 1805 by gaining another first prize, for *Mémoire sur la décomposition de la penser* (The Analysis of Thought). While continuing to write outstanding philosophic and psychological essays, he intensified his political activities, became a member of the Chamber of Deputies, and was made commander in the Legion of Honor. Under the first restoration he returned to the National Assembly and was put in charge of liaison between the assembly and the king on financial matters. Despite these public activities, he was at the time of his death acknowledged by most of his distinguished contemporaries as their master (*maître à tous*) in philosophy.

His famous *Journal intime* reveals a melancholy, emotionally changeable person, of poor health, who was highly sensitive to climatic and personal surroundings. He spent much of his personal and philosophic life trying to understand and mitigate this sensitivity.

PHILOSOPHICAL DEVELOPMENT

Maine de Biran's philosophic development can be summarized briefly as a movement toward a more and more detailed conviction that man's inward experience is (1) different from his outwardly experienced "impressions," and (2) an important source and basis of knowledge. His most mature essays speak of an "inward sense" (*sens intime*) that reveals our experience of willed bodily move-

ment (*effort voulu*); in the course of his philosophic development he gave to this experience a more and more important role, progressively more subtly analyzed. The names of John Locke, Étienne Bonnot de Condillac, and Charles Bonnet, all of whom emphasized outward impressions as the ultimate source of knowledge, occurred as frequently in his early notes as did the name of Jean-Jacques Rousseau, whose "Profession of Faith of a Savoyard Vicar" in *Émile* had aroused Maine de Biran's interest in the "inner light" (*lumiére intérieure*).

But the outwardly oriented epistemologies of Condillac and Locke and their disciples, the *Idéologues*, soon grew less adequate for Maine de Biran, as did Bonnet's explanations of perception in terms of physiological mechanisms (explanations based upon outward "impressions"). After 1802 and his first great prize essay, *The Influence of Habit on the Faculty of Thinking*, which was similar in many ways to the writings of the *Idéologues*, Maine de Biran moved into his longest and most original period of philosophizing, during which he became quite critical of his former masters and developed and defended the key doctrine of his philosophy, that the *effort voulu* is a unique source of basic knowledge. In this stage he wrote *Mémoire sur la décomposition de la penser* (which won him membership in the Institute of France) and his most mature completed philosophic work, *Essai sur les fondements de la psychologie* (Essay on the Foundations of Psychology; 1812).

From 1814 to the end of his life he developed—but never with great precision—a doctrine derived from Immanuel Kant (by way of Maine de Biran's friend André Marie Ampère), a doctrine that identified "belief" (*croyance*) as one of the inner sources of knowledge. At first Maine de Biran spoke of belief as revealing the transphenomenal substance of things, and from 1815 on he applied this notion of a "faculty of belief" to problems of theology. According to Maine de Biran, *croyance*, like the *effort voulu*, originates inwardly, but—unlike voluntary bodily movement—is always passive; its function is to receive God's grace. Still, he continued to speak of the importance of the *effort voulu*; the doctrine of the significance of the faculty of belief in relation to religious matters was not a repudiation of the significance of the activistic, individualistic capacity of the *effort voulu* in matters of natural knowledge. In fact, during this last period, from 1814 to 1824, he wrote some of his finest essays developing his doctrine that the *sens intime* is a unique and important source of knowledge. Two of his outstanding works on this subject were *Examen des leçons de philosophie de M. Laromiquière* (An Examination of Laromiquière's Lessons in Philosophy; 1817) and his unfinished masterpiece, *Nouveaux Essais d'anthropologie* (New Essays in Anthropology; 1824), both of which cast much light on the doctrine of *effort voulu*. In fact this doctrine was far more thoroughly developed than the doctrine of *croyance*. Nevertheless, the emphasis given to belief in the last stage of his thought confirms the generalization that the whole tendency of his philosophic development was toward a more profound conviction that inward experience—whether of willed effort or of belief itself—is the richest basis of knowledge.

LEARNING AND EXPERIENCE

Condillac, the forerunner of the *Idéologues*, had insisted on clarifying terms and validating claims to knowledge by reference to simple, directly experienced outward "sensations" stripped of the increments of learning. The leader of the *Idéologues* in Maine de Biran's day, Destutt de Tracy, had continued Condillac's line of thought but had noticed that (1) some experiences get duller and vaguer by repetition, while others become more distinct; and that (2) there is a capacity to move our bodies voluntarily (Destutt de Tracy called it "*motilité*") that has a vital function in our learning to perceive objects. In addition, Destutt de Tracy's colleagues Cabanis and Bonnet had seen the importance of physiological conditions for an analysis of the human mind.

In his first prize-winning essay Maine de Biran developed all of these suggestions. He not only distinguished between outer impressions and felt effort, but he distinguished what he called "sensations" (such as tastes and smells), wherein the impression is vivacious and our voluntary bodily movement is minimal, from what he called "perceptions" (such as talking aloud and hearing ourselves), wherein the outward impresssion is less important than the inward experience of moving our organs.

But these distinctions might have no importance for an analysis of knowledge, he thought, if they do not help us to understand learning more fully. And so in his first essay he set about trying to discover whether habituation or repetition has a different effect on passive sensations than on active perceptions; if different effects were found to exist it could be assumed that the distinction between sensations and perceptions is important. He found that passively experienced sensations got vaguer with habituation, and perceptions that are involved with our willed bodily movement became more and more precise. Our sense of smell loses its refinement in a hothouse, but we walk, talk, play games better by practicing. Therefore, he concluded, in perceptions alone do we find the possibility

of learning, of moving from the passive sensational confusion of the infant to the subtle distinctions of the adult mind. If Condillac's passively received outward impressions were all that was available to consciousness, the repetition of these impressions would have resulted in a vague blur. The development of mind is linked with willed bodily movement, with perceptions.

One of our most important perceptions is our experience of speaking and hearing our own words; this is the most active perception, and the least dependent upon adventitious external impressions. Sounds uttered by us are among the first signs we know; they are outwardly experienced signs of our own inward actions, and it is the inward action that constitutes the meaning of the sign. There are other signs too: We learn to associate two or more external impressions as natural, or physical, signs of each other. But for Maine de Biran the sign-relationship most directly involved in human reasoning is the relationship between spoken words or conventional signs and our inwardly experienced effort to move our organs of speech. In the course of acquiring by habituation a more subtle and distinct way of talking we acquire a more subtle and distinct mentality. Maine de Biran never lost sight of natural sign-relationships between impressions or between images of impressions as part of our learning process, but he insisted that oral, conventional sign-relationships were basic to human mentality. To describe human thinking only in terms of associated images of outward impressions is to ignore speech, the faculty that makes human thought peculiarly human.

In 1812, in his "Essay on the Foundations of Psychology," Maine de Biran set out to find a primary experience, a *fait primitif* antecedent to all learning or habituation (Condillac had sought such a fact and had claimed to find it in outward sensations). Maine de Biran held that such a basic experience must satisfy three criteria: First, it must be within the limits of awareness (although he sometimes talked of unconscious perceptions); second, it must, of course, not be learned or deduced, but must be directly experienced; finally, it must be persistent, for knowledge must have a firmer basis than the passing moment. He rejected outward impressions and inward emotions and affections because they were fleeting, and he rejected the physiological findings he had once been attracted to because they were the results of inferences or deductions, not immediately experienced. In the end he adopted as his primary experience the *effort voulu* he had found to be so crucial to the learning process: We are aware of it, although sometimes not vivaciously; it is not itself learned, although we learn how to move various members skillfully; and this experience persists in various degrees of tension (ranging from sensations up to perceptions) throughout our waking life. The most lucidly developed part of Maine de Biran's philosophy is his explanation and defense of this triple claim involved in calling the *effort voulu* a primary experience.

SELFHOOD, CAUSALITY, AND LIBERTY

Philosophers such as Locke, Condillac, and the *Idéologues* had great difficulty accounting for our idea of a persistent, inwardly experienced self, because they assumed that experience was made up of nothing but fleeting, outward impressions. But the origin of this idea loses its mystery if we give our attention to our persistent, inward experience of our own willing against our varying bodily resistance to that willing. Throughout our lives we feel this relationship at the center of our experience in varying degrees of tension. The center is the self (*le moi*), the periphery, or the surrounding impressions, is the nonself. In fact, the unity of our own more or less resisting body as felt in the *sens intime* is the origin of our whole notion of unity or identity, whether it occurs in mathematics or elsewhere.

The felt relationship between the body and our more or less active willing to move that body is for Maine de Biran our basic experience of causation. In defending this claim he argued that the term *cause* cannot be explained by hazy references to "innate" ideas, or by question-begging, tautological assertions about effects presupposing causes; in this he agreed with David Hume. He also agreed with Hume that our disparate impressions do not reveal any instance of necessary connection. But he flatly disagreed with Hume's double assumption that outward impressions are basically similar to and are the origin of any inward experience we may have. Maine de Biran insisted that in our *sens intime* we find a unique, primary experience of necessary connection.

Hume's main objections to this claim occur in his *Enquiries concerning the Human Understanding and concerning the Principles of Morals*; he points out that in cases such as palsy or amputation we cannot be sure our own bodily movement will follow our willing. Moreover, the means by which the will and our body are united is, in Hume's word, "mysterious." How then can we be said to experience an instance of necessary connection when neither connection nor necessity is experienced here? Maine de Biran responded to these objections by using his basic distinction between impressions and the *effort voulu*, or between images, or copies of outward impressions, and our idea of inward felt effort. To the first objection he replied that bodily movement is simultaneous with the

willing that is its cause, and that if there is any failure or disappointment, it is the failure or disappointment of a plan involving memory and anticipatory images concerning a succession of experiences. Willed effort itself, involving the simultaneity of cause and effect, never fails; only plans involving successive outward impressions may fail. According to Maine de Biran, Hume mistakes our *pensées* for our *effort voulu*, confuses disparate outward impressions and their images with intimately related, inwardly simultaneous willing and movement.

Hume's second objection is that no connection or "means" connecting the will to the body is present in willed effort. By "means" Hume chiefly meant physiological means that can be demonstrated through outward impressions and derived hypotheses concerning the connection between the willed effort and bodily movement. Maine de Biran answered, however, that in the face of the plainly felt experience of inward causation, one need not ask for "connecting" entities deviously derived from a different sort of experience; Hume, in doing so, simply reasserted his old prejudice in favor of outward impressions and their images. No assertion concerning our physiological structures can diminish or put in question our inwardly experienced relationship between willing and our body. To say that it does is like claiming that remarks about a Caruso's anatomy diminish or put in question the greatness of his artistry. The greatness lies in the singing itself, just as our certainty in experiencing the *effort voulu* lies in this experience itself, not in any hypothetical structures based on quite different experiences. Finally, Maine de Biran pointed out that we apply the term *cause* or *necessary connection* to outward impressions by projecting our inward experience of simultaneity into the outward world of successive impressions; our original experience of causation or necessary connection is inward; all other uses of the term *causation* are derivative from it.

The certainty of the experienced relationship between will and bodily movement is the basis of man's liberty. Deterministic arguments that have been invoked to contest man's liberty depend on causal laws that are less certain than, and indeed irrelevant to, the experience of moving our bodies ourselves. Maine de Biran was willing to assert that in varying degrees strong motives or desires incline us to will certain movements. He was even willing to agree that our passions are sometimes overwhelming, for example, under the influence of hunger or fear, but he went on to say that there are times when the crucial causal factor in any action is our will, which is capable of rejecting any given desire or inclining motive. At those times we are free, and no dubious hypotheses concerning determining causes can hold up against the plain fact that we can and do withstand particular external or internal pressures. Our freedom does exist, although it is occasional and is tempered by the degree of inclination or pressure.

See also Ampère, André Marie; Bonnet, Charles; Cabanis, Pierre-Jean Georges; Causation; Condillac, Étienne Bonnot de; Destutt de Tracy, Antoine Louis Claude, Comte; Hume, David; Kant, Immanuel; Locke, John; Perception; Rousseau, Jean-Jacques.

Bibliography

WORKS BY MAINE DE BIRAN

The first published edition of Maine de Biran's works was *Oeuvres philosophiques de Maine de Biran*, edited by Victor Cousin, 4 vols. (Paris: Ladrange, 1841). This edition is incomplete and should be avoided, except by those who wish to account for the gross misunderstandings of Maine de Biran's thought that were current in the nineteenth century. The definitive edition of Maine de Biran's notes, essays, and letters is the one edited by Pierre Tisserand and Henri Gouhier: *Oeuvres de Maine de Biran*, 14 vols. (Paris: Alcan, 1920–1942). Gouhier has also edited the definitive edition of Maine de Biran's philosophically revealing *Journal intime* (Neuchâtel, 1954–1957). Only one of Maine de Biran's works has been translated into English—his first prize-winning essay, translated by Margaret Boehm as *The Influence of Habit on the Faculty of Thinking* (Baltimore: Williams and Wilkins, 1929).

WORKS ON MAINE DE BIRAN

No definitive biography has been written; the most detailed life now in print is that by Amable de La Vallett-Monbrun, *Maine de Biran: Essai de biographie historique et psychologique* (Paris, 1914). Sainte-Beuve's brief biography of him in *Causeries du Lundi*, Vol. VIII (Paris, undated), is famous for its eloquence.

On the development of Maine de Biran's philosophy three excellent books have been written. Henri Gouhier's *Les conversions de Maine de Biran* (Paris: Vrin, 1947) is the best account we have of the influences upon him. *Maine de Biran et son oeuvre philosophique*, by Victor Delbos (Paris: Vrin, 1931), is a lucid, impartial summary of the key works. *L'expérience de l'effort et de la grâce chez Maine de Biran*, by George Le Roy (Paris, 1934), uses a Bergsonian approach but even so is faithful and perceptive; it is the best consecutive account of his development. A perceptive, memorable account of his thought occurs in *French Philosophies of the Romantic Period*, by George Boas (Baltimore: Johns Hopkins Press, 1925).

A few useful works on specific topics include Henri Gouhier, "Maine de Biran et Bergson," in *Les études bergsoniennes*, Vol. I (Paris, 1948); Philip Paul Hallie, *Maine de Biran, Reformer of Empiricism* (Cambridge, MA: Harvard University Press, 1959); Jacques Paliard, *Le raisonnement*

selon Maine de Biran (Paris, 1925); Euthyme Robef, *Leibniz et Maine de Biran* (Paris, 1927); Ian W. Alexander, Ian W. "Maine De Biran and Phenomenology," *Journal of the British Society for Phenomenology* 1 [1970]: 24–37); Francis C. Moore, Francis C., *The Psychology of Maine De Biran* (Oxford: Clarendon Press, 1970); Serge J. Morin, "Maine De Biran: A New Dualism," *Philosophical Forum* (5[1974]: 441–459); Jean Pucelle, "The Meaning of Experience in Maine De Brian's Philosophy," *International Philosophical Quarterly* (13[1973]: 25–32); Christopher C. Rodie, "Delacroix, Maine De Biran, and the Aesthetics of Romanticism." *Dialogue* (17[1974]: 13–24).

Philip P. Hallie (1967)
Bibliography updated by Tamra Frei (2005)

MAISTRE, COMTE JOSEPH DE

(1754–1821)

Comte Joseph de Maistre, the Savoyard philosopher and diplomat, was born in Chambéry. After the conquest of Savoy by the French revolutionary forces, he retired to Lausanne, where he lived for three years, devoting himself mainly to writing his *Considérations sur la France* (1796), an attack on the political philosophy of republicanism. He was then summoned to Turin by the king of Sardinia and later moved to Cagliari, the capital of the very diminished kingdom of Sardinia. In 1802 he was appointed Sardinian minister plenipoteniary to St. Petersburg and remained there for fourteen years, composing his famous *Soirées de Saint-Pétersbourg*, which was not published until the year of his death.

ULTRAMONTANISM

De Maistre is best known for his ultramontanism and traditionalism, which are most forcibly stated in *Du pape*, written in 1817, although anticipated in certain details in his *Considérations sur la France*. His presuppositions were those of any medieval Roman Catholic—the church is a divine institution; its foundation was given to St. Peter; St. Peter was the first pope; his successors have inherited the powers conferred on him by Jesus Christ himself. The book opens with a demonstration of papal infallibility. Identifying the sovereignty of the pope with that of any secular ruler, de Maistre argued that sovereignty implies infallibility, since no ruler is sovereign whose decisions can be set aside or be subject to appeal. He thus made no distinction between executive competence and validity. As parliaments exist simply to inform the sovereign of matters of which he might not be aware or to make requests and express occasional desires, so the church councils have no power to do more than this. They are convoked and presided over by the pope, who is not bound by their decisions, for they have no real power of decision. The notion that matters of faith and doctrine can be decided by a council is as absurd as the notion that a parliament can actually rule. De Maistre maintained that when the pontiff speaks ex cathedra and without restraint to the church, he has never erred nor can he ever err in questions of faith. He might be constrained to make a false pronouncement, or he might be speaking merely as a man and not as a pope, but in his function as a sovereign monarch, it is impossible that he should ever be in error.

The reason we require any kind of government is that we are born corrupt, yet with a sense of morality. Our souls are thus in a state of conflict. Sovereigns exist in order to prevent the disasters that arise from this conflict and to keep order within the state. No man is capable of governing himself, for no man can spontaneously quell the evil that is in him; therefore, the power to do so must reside in the hands of one ruler who will be above criticism and have absolute power. This ruler, whether he is a king or a pope, does not rule by the consent of his people but because of their needs. Kings, although infallible in regard to their own provinces, are nevertheless subject to the laws of God, and the pope is the only possible judge of whether they have been faithful to them. The pope is the deputy of God, and when a secular ruler has erred, he can be deposed and his subjects can be freed from their oaths of allegiance to him by papal decree. This power, de Maistre maintained, has been used only rarely where hereditary sovereigns were involved; it was used more freely against elected sovereigns, such as the Holy Roman emperors, for they were chosen by man, not by God. The pope, it should be noted, does not interfere in purely secular problems of administration; his intervention is invoked only in morals and religion.

Nevertheless, the pope is not a universal sovereign, for his power is checked by the canons, the laws, the customs of nations, duty, fear, prudence, and opinion, "which governs the world." Is it not better, de Maistre asked, to settle disputes by the decision of a wise and prudent ruler, inspired by God himself, than by rebellions, civil wars, and all the evils that follow from them? Such an arbitrator will inevitably submit to the commands of duty and prudence, will be sensitive to custom and opinion, and will intuitively know which road to take when conflict arises.

TRADITIONALISM

A reader of *Du pape* will be impressed by de Maistre's use of tradition to justify his conclusions. The supremacy of the pope, he argued, has always been acknowledged, even by his critics. That is, they all admitted that he has done what de Maistre said he has the power to do, and, de Maistre added, no one except those who had suffered at his hands objected to his power. That something has always been done is to de Maistre proof that it has been done correctly. He even denied the right to liberty on the ground that slavery was the fate of most men until the rise of Christianity.

To de Maistre the human race is a single being, the soul of which is expressed in its language. Language develops, but so does tradition. The tradition of Catholicism is simply the fulfillment of the covenant God gave to Abraham; passed to Moses and then to Aaron, the high priest; and so on down to the promise made to Peter. But in every tradition, in spite of its development, there is a unity of idea, and the maintenance of that unity is entrusted to the pontiff.

ROYALISM

Concurrent with de Maistre's traditionalism was his royalism. He was so convinced of the need for absolute monarchs that he even maintained that since kings had a longer life expectancy than other men, royal families differ in nature from nonroyal families, as a tree differs from a shrub. A king is not a private individual and must not be judged as such. He is the nation in the same way that the pope is the church. Consequently, his power is also absolute, for when he speaks, it is the nation speaking through him. Kings alone preserve national unity. The word *unity* was a eulogistic term for de Maistre. To be unified is better than to be manifold; to remain the same is better than to change. And although de Maistre had to admit those changes that have obviously occurred and are not evil, he insisted on the unity that underlay them.

De Maistre usually carried his ideas to their logical conclusions. His famous apostrophe to the hangman in the *Soirées* is based on de Maistre's presupposition of the twofold nature of man. If the hangman is removed from society, order will give way to chaos, thrones will totter, and society will disappear. "God who is the author of sovereignty is also the author of punishment." He is the author of punishment so that corrupt man may still be redeemed. But if man is to be punished, there must be an absolute and unquestioned power to execute the punishment, and that power is the king's.

De Maistre was the first philosopher of the counterrevolution in France. With the vicomte de Bonald, he gave a set of arguments to legitimists and Catholics. But although de Maistre was admired by many for his consistency in both principle and inference, his variety of political philosophy was never popular, even during the restoration. The anti-intellectualism of François René de Chateaubriand and Mme. de Staël, as fully opposed to the extremes of revolution as was de Maistre's traditionalism, gained more adherents. Moreover, ultramontanism was disclaimed by the Vatican. This disclaimer, perhaps, was the main reason for the failure of de Maistre's thought to become popular in France.

See also Bonald, Louis Gabriel Ambroise, Vicomte de; Chateaubriand, François René de; Republicanism; Staël-Holstein, Anne Louise Germaine Necker, Baronne de; Traditionalism.

Bibliography

WORKS BY DE MAISTRE

Considérations sur la France. Neuchâtel, Switzerland, 1796.

Du pape. 2 vols. Lyons, 1819.

Soirées de Saint-Pétersbourg. 2 vols. Paris, 1821.

Oeuvres complètes. 14 vols. Lyons, 1884–1887.

The Works of Joseph de Maistre. Translated by Jack Lively. New York: Macmillan, 1965. Selections.

WORKS ON DE MAISTRE

Boas, G. *French Philosophies of the Romantic Period*. Baltimore: Johns Hopkins Press, 1925. See Ch. 3.

Ferraz, M. *Histoire de la philosophie en France au XIXe siècle*. Vol. II, *Traditionalisme et ultramontanisme*. Paris: Didier 1880.

Garrard, Graeme. "Joseph De Maistre's Civilization and its Discontents." *Journal of the History of Ideas* 57(3) (1996): 429–446.

Garrard, Graeme. "Rousseau, Maistre, and the Counter-Enlightenment." *History of Political Thought* 15(1) (1994): 97–120.

Gianturco, E. *Joseph de Maistre and Giambattista Vico*. Washington, DC, 1937.

Kochin, Michael S. "How Joseph De Maistre Read Plato's Laws." *Polis* 19(1–2) (2002): 29–43.

Kow, Simon. "Maistre and Hobbes on Providential History and the English Civil War." *Clio* 30(3) (2001): 267–288.

Laski, H. J. *Authority in the Modern State*. New Haven, CT: Yale University Press, 1919.

Lecigne, C. *Joseph de Maistre*. Paris, 1914.

Spektorowski, Alberto. "Maistre, Donoso Cortes and the Legacy of Catholic Authoritarianism." *Journal of the History of Ideas* 63(2) (2002): 283–302.

George Boas (1967)
Bibliography updated by Tamra Frei (2005)

MAJOR, JOHN
(1469–1550)

John Major, or Mair, was a Scottish theologian, active at the University of Paris for some years before and after he secured a license in theology in 1506. Major helped to revive, if only briefly, the spirit of fourteenth-century nominalism. He was entirely sympathetic with the approach of William of Ockham and Jean Buridan, even though he adopted some doctrines of John Duns Scotus and other realists.

Major came to Paris in 1493 after studying at Cambridge. He taught at the University of Paris for most of his lengthy career, with the exception of seven years at the Scottish universities of Glasgow and St. Andrews. When he arrived at Paris, scholasticism, pietism, and humanism were rivals within the university itself. Late medieval pietism was reflected in the ascetic discipline instituted at the Collège de Montaigu, the school that so repelled Desiderius Erasmus by its austerity and its logic-chopping. Major, with his frugal Scottish background, found the atmosphere of Montaigu less forbidding, and he responded with initial enthusiasm to its manner of disputing. He seems to have been little influenced by the sort of humanism being advocated at the time by Jacques Lefèvre d'Étaples, who stressed the value of knowing Aristotle and the Church Fathers in the original Greek. Major belonged to the scholastic tradition completely. His theological and philosophical works proceed entirely from a formal analysis of separate arguments. He made no use of Greek, although he clearly was conversant with Latin literature.

Major's earliest published work consisted of short treatises on terminist logic, published separately from 1500 to 1503, and then together at Lyons in 1505 as a commentary on Peter of Spain. Later he published commentaries on Aristotle's *Ethics* and *Physics*. In theology, he wrote commentaries on the *Sentences* of Peter Lombard and on the Gospels. All of these writings reflect his teaching duties, even in their style. Toward the close of his long life, Major complained mildly at having been forced to accommodate himself to the "manner of our ancestors" and admitted that students had not always found the disputatious style agreeable. In addition to the works already mentioned, Major wrote *A History of Greater Britain*, a landmark in the writing of Scottish history and a most unusual work for a nominalist theologian. Many passages in this work—such as those in defense of the "oaten bread" of Scotland or of ale as opposed to wine—suggest a personality by no means dry and pedantic. Nevertheless, Major's philosophical style has put off scholars, and his work still awaits total and mature evaluation. Almost all present-day accounts of Major continue to be colored by humanist criticisms of theology made in the spirit of Erasmus, with little sympathy for medieval logic.

See also Aristotle; Buridan, John; Duns Scotus, John; Erasmus, Desiderius; Logic, History of: Medieval (European) Logic; Medieval Philosophy; Patristic Philosophy; Peter Lombard; Peter of Spain; Pietism; William of Ockham.

Bibliography

A reliable, although sketchy, account of Major's philosophical opinions is given by Ricardo Garcia Villoslada in *La universidad de Paris durante los estudios de Francisco de Vitoria* (Rome: Universitatis Gregorianae, 1938), pp. 127–164. Carl Prantl, in *Geschichte der Logik im Abendlande* (Leipzig, 1927), Vol. IV, pp. 247–250, gives a few excerpts from Major's logical writings. Major's views on church matters (he was a conciliarist and champion of Gallicanism) are sometimes dealt with briefly in histories of political theory. The details of his life are presented in Aeneas J. G. Mackay's biography, prefixed to an English translation of *A History of Greater Britain* (Edinburgh: Edinburgh University Press, Scottish History Society, 1892), which also contains a bibliography of Major's writings. This bibliography needs to be supplemented, however, by the additions given by Hubert Élie, *Le traité "De l'infini" de Jean Mair* (Paris: Vrin, 1938); James F. Keenan, "The Casuistry of John Major: Nominalist Professor of Paris (1506–1531)," *Annual of the Society of Christian Ethics* (1993, pp. 205–221).

Neal W. Gilbert (1967)
Bibliography updated by Tamra Frei (2005)

MALCOLM, NORMAN
(1911–1990)

Norman Malcolm, one of America's best-known philosophers, was born in Selden, Kansas, in 1911. After studying philosophy with O. K. Bouwsma at the University of Nebraska, he enrolled as a graduate student at Harvard in 1933. The decisive period for Malcolm's career, however, was probably the time he spent at Cambridge University in 1938–1939, when he met G. E. Moore and Ludwig Wittgenstein. Although Moore exerted a strong influence on him, it is perhaps not unfair to say that most of Malcolm's published work was an attempt to understand Wittgenstein, to explain his thought to others, and to apply Wittgenstein's characteristic manner of approaching philosophical questions to areas the latter did not directly treat.

Malcolm's published work deals especially with the nature of necessary truth; empirical certainty; the connections between common sense, ordinary language, and philosophy; knowledge and perception; and such topics in the philosophy of mind as memory, dreaming, and the problem of other minds. He also wrote on topics in the philosophy of religion. What follows will be confined to the first three topics.

NECESSARY TRUTH

"Are Necessary Propositions Really Verbal?" and its companion piece, "The Nature of Entailment" (in *Knowledge and Certainty*), together form an interesting statement of the linguistic theory of the a priori. In the former, Malcolm points out that some philosophers (for example, C. D. Broad, Moore, and A. C. Ewing) hold that necessary propositions state very general truths about reality—for instance, that nothing is both red and green all over. Others (for example, A. J. Ayer and the early Wittgenstein) apparently believe that if necessary propositions state anything at all, they state truths about language; they are "merely verbal." Malcolm tries to show that, although it is false, literally speaking, that necessary propositions are merely verbal, there is nonetheless considerable merit in saying that they are. He argues this point by claiming that we learn necessary truths by observing how people use certain expressions. Finding out that a pair of propositions are equivalent, for example, is the same thing as finding out that some pairs of expressions are used interchangeably. What makes a given statement necessary is some empirical fact about linguistic usage. (Although Malcolm considers the objection that on this account any necessary statement turns out to be identical with or equivalent to some contingent statement about linguistic expressions, he does not, it seems, have a clear answer to it.) Accordingly, he says, it is false that necessary statements are merely verbal or are rules of grammar or are not really propositions; it is nonetheless worthwhile to say these things in that they prevent one from supposing, for example, that there are two kinds of facts or truths, necessary and contingent, a supposition that is, literally speaking, true but nonetheless misleading. Why? Perhaps Malcolm believed that in saying this one minimizes the vast and important difference between necessary and contingent truths, the difference being that the necessary truths depend upon or reflect facts of linguistic usage in a way that the contingent truths do not.

EMPIRICAL CERTAINTY

In "The Verification Argument" and "Certainty and Empirical Statements" (in *Knowledge and Certainty*), Malcolm objects to the view that no empirical statements are ever really certain. "The Verification Argument" is a careful, clear, and very impressive examination of the arguments philosophers (in particular, C. I. Lewis, who was a teacher of Malcolm's at Harvard) have offered for this skeptical view. Where *S* is any empirical statement, Malcolm points out that these arguments always invoke as a premise the claim that the consequences of *S* may not occur and deduce from this that it is not certain that the consequences of *S* will occur. What Malcolm shows is that there is no interpretation of the former statement according to which it both is true and entails the latter.

ORDINARY LANGUAGE

In several essays, Malcolm dealt with certain questions about the relationships between ordinary language, common sense, and philosophy. Essentially, what he says is that if a philosopher is investigating a concept of ordinary language (for example, *seeing*) and comes to conclusions at variance with ordinary language, then we may be sure that he has made a mistake. What is it to come to a conclusion that goes against ordinary language? One way of doing this is to hold that a sentence with an ordinary use expresses a logical impossibility: some philosophers, for example, appear to insist that it is logically impossible to see physical objects. We may recognize their error by noting that such sentences as "I see the table in the corner" have a perfectly good ordinary use and therefore cannot be self-contradictory. But it is impossible to convey the full power of Malcolm's arguments without a very detailed consideration of particular cases.

See also Ayer, Alfred Jules; Broad, Charlie Dunbar; Common Sense; Dreams; Lewis, Clarence Irving; Memory; Moore, George Edward; Ontological Argument for the Existence of God; Other Minds; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY MALCOLM

"Defending Common Sense." *Philosophical Review* (1949).

"Philosophy and Ordinary Language." *Philosophical Review* (1951).

"Dreaming and Skepticism." *Philosophical Review* (1956).

Ludwig Wittgenstein: A Memoir. London: Oxford University Press, 1958.

Dreaming. London: Routledge, 1959.

Knowledge and Certainty. Englewood Cliffs, NJ: Prentice Hall, 1963.

"Behaviorism as a Philosophy of Psychology." In *Behaviorism and Phenomenology,* edited by T. W. Wann. Chicago: William Marsh Rice University by the University of Chicago Press, 1964.

"Is It a Religious Belief That God Exists?" In *Faith and the Philosophers,* edited by John Hick. New York: St. Martin's, 1964.

"Scientific Materialism and the Identity Theory." *Dialogue* (1964).

Alvin Plantinga (1967)

MALEBRANCHE, NICOLAS
(1638–1715)

EARLY LIFE AND *RECHERCHE*

One of the major figures in post–René Descartes Cartesianism, Nicolas Malebranche was one of many children born to his mother, Catherine de Lauzon, the sister of a viceroy of Canada, and his father, also Nicolas Malebranche, a secretary to Louis XIII. As in the case of Descartes and Blaise Pascal, Malebranche was born in frail health. His particular afflictions were a severe malformation of the spine and weak lungs, and because of these conditions he needed to be tutored at home until the age of sixteen. Subsequently, he was a student at the Collège de la Marche, and after graduating he went to study theology at the Sorbonne. His education left him with a dislike of a scholasticism that focused on the work of Aristotle. Thus, in 1660 he decided to leave the universities and enter the Oratory, a religious congregation founded in Paris in 1611 by the Augustinian theologian Pierre Bérulle. At the Oratory Malebranche studied ecclesiastical history, linguistics, and the Bible, and with his fellow students he also immersed himself in the work of St. Augustine. Though judged to be merely a mediocre student, he was ordained a priest on September 14, 1664.

The same year he was ordained, Malebranche happened in a Paris bookstall upon a posthumous edition of Descartes's *Traité de l'homme* (*Treatise on Man*), which provides a sketch of a mechanistic account of the physiology of the human body. Malebranche's early biographer, Father Yves M. André, reports that he was so "ecstatic" on reading this account that he experienced "such violent palpitations of the heart that he was obliged to leave his book at frequent intervals, and to interrupt his reading of it in order to breathe more easily" (André 1970, pp. 11–12). Though André does not indicate why Malebranche was so moved, one can speculate that he had discovered in this text a way to investigate the natural world without relying on Aristotelian scholasticism. In any case, after his encounter with *L'homme* Malebranche devoted himself to a decade-long study of the Cartesian method and its results in mathematics and natural philosophy.

The principal fruit of this study was a two-volume work bearing the title *De la recherche de la vérité. Où l'on traitte de la nature de l'esprit de l'homme, et de l'usage qu'il en doit faire pour eviter l'erreur dans les sciences* (*The Search after Truth*, first published 1674–1675), in which is treated the nature of the human mind and the use that must be made of it to avoid error in the sciences. It is primarily this text that provides the basis for Malebranche's reputation in the early modern period. As its full title indicates, the *Recherche* focuses on the principal sources of human error and on the method for avoiding those errors and for finding the truth. The first five books enumerate the various errors deriving from the senses, — imagination, pure understanding, inclinations, and passions, respectively—and a sixth book is devoted to the Cartesian method of avoiding such errors through attention to clear and distinct ideas. The centerpiece of the third book, on pure understanding, is a defense of the claim that the ideas through which one perceives bodies exist in God. Tucked away in the final book, on method, is a critique of "the most dangerous error of the ancients," namely, the Aristotelian position that there are secondary causes in nature distinct from God.

The first volume of the *Recherche*, containing the first three books, was published in 1674 and drew an immediate response in 1675 from Simon Foucher, the canon of Sainte Chapelle of Dijon. Foucher was an "academic skeptic" who attacked the assumption that ideas in one can represent objects distinct from oneself (see Foucher 1969). The Cartesian Benedictine Robert Desgabets replied to Foucher by insisting that the Cartesian rule that clear and distinct ideas are true presupposes that one's thoughts correspond to real external objects. In brief prefaces added to various editions of the second volume of the *Recherche*, Malebranche chastised both thinkers for failing to read the work they were discussing, noting in particular that he had explicitly argued in the *Recherche* that the ideas one perceives exist in God rather than in oneself.

Malebranche solicited written responses to the *Recherche* modeled on the sets of objections published with Descartes's *Meditations*. Perhaps put off by Malebranche's harsh treatment of Foucher and Desgabets, his critics offered instead only informal objections channeled

through mutual friends. In 1678 Malebranche appended to the *Recherche* a set of sixteen *Eclaircissements*, or clarifications, that respond to these objections. Among the more important objections addressed are those that concern Malebranche's assertion that one has a freedom to "consent" to certain motives for action (Eclaircissement I), his claim that reason does not yield a demonstrative argument for the existence of the material world (Eclaircissement VI), his doctrine of the vision of ideas in God (Eclaircissement X), his conclusion that one knows one's own soul through a confused consciousness rather than through a clear idea of its nature (Eclaircissement XI), and his occasionalist thesis that God is the only true cause (Eclaircissement XV). In the 1678 edition there is a final Eclaircissement that defends the importance "not only for knowledge of nature but also for knowledge of religion and morals" of the view, only hinted at in the text of the *Recherche* itself, that God acts for the most part through "general volitions" (*volontez générales*), and that He acts though "particular volitions" (*volontez particulières*) only in the exceptional case of miracles.

NATURE ET GRÂCE AND THE DEBATE WITH ARNAULD

Malebranche developed his theory of divin action in his 1680 *Traité de la nature et de la grâce* (*Treatise on Nature and Grace*). He published this work over the objections of the Jansenist theologian and Cartesian philosopher Antoine Arnauld, who was disturbed by what he saw as Malebranche's denial of the claim in the Scriptures and Catholic tradition that God attends to particular details in matters of grace. Arnauld responded to the publication of *Nature et de la grâce* by publishing a response to Malebranche, and the ensuing battle between these two individuals became one of the major intellectual events of the day. Arnauld's opening salvo was the 1683 *Des vraies et des fausses idées* (*On True and False Ideas*), which attacks not *Nature et de la grâce* but the *Recherche* (see Arnauld 1990). His strategy here is to undermine Malebranche's influence in theological matters by revealing the inadequacy of his philosophical views. In particular, Arnauld attacks Malebranche's assumption that ideas are "representative beings" distinct from one's perceptions, offering instead the position, which he plausibly ascribes to Descartes, that ideas are simply aspects of the perceptual modifications of one's soul. This argument reflects a sympathy for Descartes's views that dates back to Arnauld's set of comments on the *Meditations*.

The same year that Arnauld presented his initial critique, Malebranche published the *Méditations chretiennes et métaphysiques* (*Christian and Metaphysical Meditations*), where "the Word" (i.e., the Second Person of the Trinity) offers a summary of Malebranche's system that highlights the central role that God plays in both metaphysics and morality. This work was in some ways a follow up to his 1677 *Conversations chrétiennes* (*Christian Conversations*). In this earlier text Malebranche presents a defense of the Christian religion that emphasizes the Augustinian theme of one's dependence on God for knowledge and happiness. In 1684 Malebranche further develop his views in moral philosophy in the *Traité de morale* (*Treatise on Ethics*), in which he argues that moral virtue requires a love of the "immutable order" that God reveals to those who seek to know it.

Also in 1684 Malebranche responded to Arnauld's *Idées*, and after a further exchange on the topic of the nature of ideas the debate turned to the religious issues of divine providence, grace, and miracles. The battle became increasingly bitter, and as a result of a campaign on the part of Arnauld and his supporters, Malebranche's *Nature et de la grâce* was put on the Catholic *Index librorum prohibitorum* (*Index of Prohibited Books*) in 1690 (the *Recherche* was added in 1709). The Malebranche-Arnauld polemic continued even after Arnauld's death in 1694, with the posthumous publication of two letters from Arnauld in 1699 and of Malebranche's responses to those letters in 1704.

ENTRETIENS AND DEBATES WITH LEIBNIZ AND RÉGIS

In 1688 Malebranche published his *Entretiens sur la métaphysique et la religion* (*Dialogues on Metaphysics and on Religion*), a concise summary of his main metaphysical doctrines of the vision in God and occasionalism that also addresses the problem of evil. In 1696 he appended to this text the *Entretiens sur la mort* (*Dialogues on Death*), which he composed after a life-threatening illness.

In 1692 Malebranche published a short study, the *Lois de la communication des mouvements* (*Laws of the Communication of Motions*), in which he endorses Descartes's law of the conservation of the quantity of motion but offers rules governing collision that, unlike Descartes's own rules, involve no appeal to a force in bodies to remain at rest. In correspondence with Malebranche, Gottfried Wilhelm Leibniz emphasized difficulties with Descartes's conservation law and that correspondence led Malebranche to insert into a 1700 edition of the *Lois* the claim that experience reveals the falsity of this law.

In 1693 Malebranche responded to the criticisms of the *Recherche* in the 1690 *Système de philosophie* (*System of Philosophy*) by the French Cartesian Pierre-Sylvain Régis. Régis defended an account of ideas similar to the one that Arnauld had defended against Malebranche during the 1680s, and Arnauld used the Régis-Malebranche exchange as an occasion to return to the issue of ideas during the last year of his life (on this exchange, see Schmaltz 2002, chapter 5). Despite their dispute, Malebranche and Régis were both appointed as honorary members of the French Académie des sciences when it was reorganized in 1699. Malebranche presented an inaugural lecture to the Académie that defends against Descartes an account of color in terms of the frequency of vibrations of light. In later published versions of the lecture Malebranche revised his discussion to take into account the theory of the nature of color in the work of the great English natural philosopher Sir Isaac Newton.

FINAL WORKS

In 1699 Malebranche published *Traité de l'amour de Dieu* (*Treatise on the Love of God*), along with *Trois lettres à Lamy* (*Three Letters to Lamy*), in which he rejects the claim of the Benedictine François Lamy (not to be confused with his Cartesian contemporary, the Oratorian Bernard Lamy) that passages from the *Traité de morale* and other texts support the quietist position, that moral action derives from a disinterested "pure love of God." This rejection of Lamy's quietism provided the basis for Malebranche's reconciliation with the French cleric and establishment figure Jacques-Bénigne Bossuet. Bossuet had earlier enlisted the aid of François de Salignac de la Mothe Fénelon in writing against Malebranche's occasionalism and his appeals to God's "general will," but later became a bitter enemy of Fénelon's quietism.

With the support of the apostolic vicar in China, Malebranche published in 1708 *Entretien d'un philosophe chrétien et d'un philosophe chinois, sur l'existence et la nature de Dieu* (*Dialogue between a Christian Philosopher and a Chinese Philosopher on the Existence and Nature of God*). In this text, Chinese philosophy is closely allied with the monism found in the early modern Dutch thinker, Benedict (Baruch) de Spinoza.

A sixth and last edition of the *Recherche* appeared in 1712, and in 1715 Malebranche published his final work, *Réflexions sur la prémotion physique* (*Reflections on Physical Premotion*), in which he responded to the claim of the abbé Laurent-François Boursier that occasionalism leads naturally to the Thomistic position that God determines one's actions by means of a "physical premotion." In his response, Malebranche defended the claim, present from the first edition of the *Recherche*, that one's free actions involve a "consent" that God does not determine.

NATURE OF IDEAS AND THE VISION IN GOD

In a section of the third book of the *Recherche* devoted to "the nature of ideas," Malebranche argues for his famous doctrine of the vision in God. More precisely, the thesis in this section is that one sees external objects by means of ideas in God. The argument for this thesis begins with the claim at the beginning of this section that "everyone agrees that we do not perceive objects external to us by themselves" since it can hardly be the case that "the soul should leave the body to stroll about the heavens to see the objects present there" (Malebranche 1997b, III-2.i.§1). Arnauld later took exception to this starting point, countering that "ideas, taken in the sense of representative beings, distinct from perceptions, are not needed by our soul in order to see bodies" (Arnauld 1990, p. 18). His main objection is that Malebranche stacks the deck in favor of his doctrine that one sees ideas of bodies in God by assuming from the start that these ideas are distinct from one's own perceptions.

In developing his own position, Arnauld appeals to Descartes's distinction in the Third Meditation between the formal reality of an idea as a perceptual modification of mind and its objective reality as a representation of an object. Arnauld insists that a representative idea is simply the objective reality of a perception, and thus not something distinct from that perception. However, it is important to note that Malebranche's definition of an idea does not rule out such a position from the start. As he himself insists to Arnauld, the claim that one must perceive external objects through ideas leaves open the question of whether an idea is "*a modality of the soul*, according to the opinion of M. Arnauld; an *express species*, according to certain philosophers, or an *entity created with the soul*, according to others; or finally *intelligible extension rendered sensible by color or light*, according to my opinion" (Malebranche 1958–1984, p. 6:95).

Malebranche's description of his own opinion goes beyond what can be found in the original edition of the *Recherche*. However, his description of the other alternatives is drawn directly from this text. In particular, Malebranche argues that there are only four alternatives to the conclusion that one sees bodies through ideas in God: (1) bodies transmit resembling species to the soul; (2) one's soul has the power to produce ideas when triggered by nonresembling bodily impression; (3) ideas are created

with the soul or produced in it successively by God; and (4) one's soul sees both the essence and the existence of bodies by considering its own perfections. Malebranche tells Arnauld that because this list constitutes "an exact division … of all the ways in which we can see objects" and because each of the alternative accounts yields "manifest contradictions," his argument from elimination serves to demonstrate the doctrine of the vision in God (Malebranche 1958–1984, p. 6:198f).

It is difficult to determine from the *Recherche* the precise source of the enumeration. However, Desmond Connell (1967) establishes that Malebranche's argument was drawn from the account of angelic knowledge in the work of the sixteenth-century Spanish scholastic Francisco Suárez. Particularly crucial for Malebranche's enumeration is Suárez's claim that angels must know material objects through species that God adds to their mind given that God alone can know them through His own substance. In light of this claim, one can take Malebranche's first three hypotheses to cover the various ways in which one can perceive bodies through immaterial species "superadded" to one's soul, and his fourth hypothesis to cover the possibility that one perceives bodies in the perfections of one's soul. In arguing against the last hypothesis Malebranche notes that because a finite being can see in itself neither the infinite nor an infinite number of beings (as Suárez argues in the case of angels), and because one in fact perceives both the infinite and infinity in external objects, it must be that one sees these objects by means of perfections contained in the only being that can possess an infinity of ideas, namely, God Himself.

Malebranche takes the conclusion here to confirm the view in "an infinity of passages" in Augustine that "we see God" in knowing eternal truths. This appeal to the Augustinian theory of divine illumination provides the basis for an argument for the vision in God that bypasses the unusual enumeration in the *Recherche*. This more direct argument is introduced in Eclaircissement X, where Malebranche urges that the ideas one perceives must exist in an "immutable and necessary Reason" because they are themselves immutable and necessary (Malebranche 1958–1984, p. 3:129f). Malebranche emphasizes that the Augustinian view that eternal truths derive from uncreated features of the divine intellect conflicts directly with the voluntarist conclusion in Descartes that these truths derive rather from God's free and indifferent will. Particularly in his exchanges with Arnauld, Malebranche attempts to present his doctrine of the vision in God as a natural consequence of Descartes's account of ideas. However, Malebranche's own Augustinian argument serves to show that Descartes could not have accepted this doctrine. Moreover, such an argument reveals the most fundamental reason for Malebranche's rejection of Arnauld's Cartesian identification of ideas with one's own perceptions. Because Malebranche identified these ideas with necessary and immutable essences, and because he held that these ideas derive their necessity and immutability from the divine intellect, he concludes that Arnauld's position can lead only to a radical subjectivism that renders impossible any sort of *a priori* knowledge of the material world.

INTELLIGIBLE EXTENSION AND EFFICACIOUS IDEAS

Eclaircissement X also introduces the notion of "intelligible extension" mentioned in Malebranche's claim to Arnauld quoted earlier concerning his own opinion. According to this text, God has a single ideal extension that serves to represent particular bodies to Him. Arnauld objects that this position involves a retraction of the claim in the *Recherche* that one perceives bodies by means of distinct ideas in God. In response, Malebranche insists that his view all along is that God represents particular bodies by means of His own simple "absolute being." For Arnauld, however, the view that God contains extension in this way is objectionable because it is connected to the heretical view in the work of Spinoza that God is extended substance. The charge of Spinozism reappears in Malebranche's 1713–1714 correspondence with one of his former students, J. J. Dortous de Mairan, who later became the secretary of the Paris Académie des sciences (for this correspondence, see Malebranche 1995). As in the case of Arnauld, so in this correspondence Malebranche vigorously denies this charge. In both cases he responds by emphasizing that the infinite and indivisible ideal extension that exists in God differs from the finite and divisible extension in the material world.

A final feature of Malebranche's doctrine of the vision in God is connected to the notion in his writings of the "efficacious idea" (*idée efficace*). This notion became entrenched in Malebranche's system around 1695, after his encounter with his Cartesian critic Régis (see Robinet 1965). In his *Système de philosophie* Régis challenges the claim in the preface to the *Recherche* that one's mind is united to God in a manner that "raises the mind above all things" and is the source of "its life, its light, and its entire felicity." While he grants the commonplace claim that God must create and conserve one's soul, Régis denies that one is enlightened by means of a union with ideas of

bodies in God. Rather, he insists that God conserves in one ideas that derive directly from the bodies they represent. In his 1693 *Réponse à Régis* (*Response to Régis*) Malebranche emphasizes his Augustinian position that one can be instructed as to the nature of bodies only through a union with God. However, he puts a new spin on this position when he notes that the union with God involves an "affecting" or "touching" of one's mind by God's idea of extension.

Already in the 1688 *Entretiens sur la métaphysique* Malebranche suggests that the union with God can be explicated in terms of a causal relation between God's ideas and one's mind. After 1695 he develops this suggestion by introducing the notion of "pure" or nonsensory intellectual perceptions that are produced by God's efficacious idea of extension. Still, he also stresses in this later period that such an idea is the causal source of one's sensations. One advantage of this extension of the doctrine of efficacious ideas to sensations is that it yields a fairly clear explanation of Malebranche's claim to Arnauld that an idea is "intelligible extension rendered sensible by color or light." Before 1695 Malebranche explained how intelligible extension is so rendered by appealing somewhat obscurely to the view that the soul "attaches" colors to a nonsensory idea. However, the theory of efficacious ideas allows him to say that this idea is rendered sensible by causing in one the appropriate sensations of light and color. The claim that one sees ideas in God is thus transformed into the claim that one's soul has intellectual and sensory perceptions that yield an understanding of the truth concerning bodies in virtue of their causal relation to God's idea of extension. One scholar concludes that while Malebranche starts with the vision *in* God, he ends with a vision *by* God (Alquié 1974, 209).

CARTESIAN DUALISM AND SENSATION

Malebranche tells Arnauld that it was Augustine's authority "which has given me the desire to put forth *the new philosophy of ideas*" (Malebranche 1958–1984, p. 6:80). By contrast, he emphasizes in the preface of the *Recherche* that Augustine failed to see that sensible qualities "are not clearly contained in the idea we have of matter," adding that "the difference between mind and body has been known with sufficient clarity for only a few years." The allusion here is to Descartes's discovery of an idea of matter that reveals that its nature consists in extension alone. This idea dictates that sensible qualities such as colors, tastes, and odors that are not reducible to modes of extension cannot exist external to mind. But since these qualities exist in the mind, and in particular in the mind's perception of the qualities, the mind itself must be distinguished from body. In this way the Cartesian idea of matter reveals "the difference between mind and body."

In the initial book of the *Recherche*, on the errors of the senses, Malebranche proposes that the erroneous belief of the Aristotelians as well as of Augustine that sensible qualities exist in bodies has its source in a misuse of "natural judgments" that help in the conservation of the human body. Here, he is following Descartes's account in the Sixth Meditation of the "teachings of nature," and in particular the claim there that the purpose of sensations is not to teach one about the nature of bodies but simply to inform one of what is beneficial or harmful to the human composite. Just as Descartes urged that erroneous beliefs about the nature of body can be avoided by attending to the clear and distinct perceptions of the intellect, so Malebranche counsels that one avoid error by attending to what the clear idea of matter reveals to one about the nature of body. As noted earlier, Malebranche has Augustinian reasons for saying that the idea that so instructs one exists in God. By his own admission, however, the conclusion that the idea that instructs one is an idea of extension derives from Descartes's discoveries.

Malebranche emphasizes that the clear idea of extension must be distinguished from one's confused sensations. One point he wants to make is that the idea exists in God while the sensations are only modifications of one's mind. However, his emphasis that this idea is "pure" or nonsensory indicates that one's experience of the material world has an intellectual component. His late doctrine of the efficacious idea involved the position that one has pure intellectual perceptions produced by God's intellectual idea of extension. But his mature position that this idea is also the cause of one's sensations allows for the claim that one's most basic sensory contact with the material world has an intellectual component.

Malebranche's doctrine of the vision in God also conflicts with Descartes's doctrine of the creation of the eternal truths. However, there are further departures from orthodox Cartesianism that are linked to two qualifications of this doctrine. The first qualification is that God's idea of extension can reveal only the nature of bodies and not their existence. This qualification is not explicit in the initial edition of the *Recherche*, which says only that the existence of properties of bodies external to one is "very difficult to prove" (Malebranche, 1997b, I.x.§1). Foucher objected that Malebranche has no good reason to affirm the external existence of these properties. In Eclaircissement VI, Malebranche urges that the idea of

extension does reveal the possible existence of the material world and that Descartes has shown that one has a probable argument for its actual existence deriving from one's natural propensity to believe that there are bodies. However, he concedes in this text—without crediting Foucher—that neither he nor Descartes can provide an argument from reason that demonstrates "with evidence" or "with geometric rigor" that this belief is true. His claim is that any conclusive argument must appeal to faith in the veracity of the report in the Scriptures that God has created the heavens and the earth.

According to the second qualification of the vision in God—which is found in the original edition of the *Recherche*—one perceives the nature of one's soul not through a clear idea in God, but only through a confused "consciousness or inner sensation" (*conscience ou sentiment intérieur*). Malebranche accepts the Cartesian commonplace that consciousness reveals immediately the existence of the soul. He allows that one knows the nature of one's soul to consist in thought; moreover, he embraces the Cartesian conclusion that the soul as a thinking substance is distinct from the body as an extended substance. Still, he insists that one knows that the soul is distinct from the body not by means of any direct insight into the nature of thought, but by seeing that thought is not contained in the idea of matter. More generally, Malebranche claims that one's lack of access to a clear idea of the soul is evident because one does not have knowledge of thought that matches one's knowledge of the mathematical features of bodies. This last point turns on its head Descartes's own conclusion in the Second Meditation that the nature of the human mind is "better known" than the nature of body; for Malebranche, it is the nature of body that is better known than the nature of mind.

In Eclaircissement XI Malebranche attempts to counter "the authority of Descartes" by arguing that the Cartesians themselves must admit that they have only a confused awareness of the nature of the sensory modifications of the soul. He notes that whereas the intellectual idea allows the various modes of extension to be related in a precise manner, there is no clear scale on which one can order one's sensations of different shades of the same color, not to mention one's sensations of sensible qualities of different kinds. Malebranche takes the confusion in the sensations to reveal a confusion in one's perception of the nature of the soul. He adds that Cartesians can discern that sensible qualities are modifications of an immaterial soul only by seeing that they are "not clearly contained in the idea we have of matter" (Malebranche 1958–1984, pp. 3:168, 170f).

OCCASIONALISM AND GENERAL VOLITIONS

Malebranche is known for his occasionalism, that is, his doctrine that God is the only causal agent and that creatures are merely "occasional causes" that prompt divine action. On the old textbook account, occasionalism was an ad hoc response to the purported problem in Descartes of how substances as distinct in nature as mind and body can causally interact. According to this account, Malebranche was driven by this problem with Cartesian dualism to propose that it is God who brings it about that one's sensations and volitions are correlated with motions in one's body.

However, occasionalism was already an old doctrine at the time that St. Thomas Aquinas wrote against it in the thirteenth century. Thomas indicated that the primary concern of the occasionalists was to strengthen the assertion of God's omnipotence. Though he allowed that God must "concur" with creatures in producing effects, he also claimed that there is reason to conclude that creatures are true secondary causes. For instance, he urged that it is more in accord with divine greatness to say that God communicates His power to creatures. Moreover, he claimed that it is simply evident to the senses that creatures have the power to bring about effects. Thomas also argued that if there were no natures in creatures that explain effects, then there could be no true scientific explanation of effects through their natural causes.

Malebranche was concerned to respond to all these arguments against occasionalism, particularly as they were developed in the work of scholastics such as Suárez. Against the first point that God's greatness requires the communication of His power, Malebranche counters that it is in fact idolatrous to attribute divine power to creatures. His argument that God alone can produce effects relies on the assumption that "a true cause … is one such that the mind perceives a necessary connection [*liaison nécessaire*] between it and its effects" (Malebranche 1997b, VI-2.iii). Malebranche claims that there is such a connection neither among bodily states, nor between bodily and mental states, nor among mental states. In all these cases one can deny the connections without contradiction. There can be a necessary causal connection in only one case, namely, the connection between the volitions of an omnipotent agent and its upshots. Thus, only such an agent, namely, God, can be a true cause.

In the *Entretiens sur la métaphysique* Malebranche offers a different argument based on Descartes's suggestion in the Third Meditation that God conserves the world by continuously creating it. The argument begins

with the claim that God must create bodies in some particular place and in determinate relations of distance to other bodies. If God conserves a body by creating it in the same place from moment to moment, that body remains at rest, and if He conserves it by creating it in different places from moment to moment, it is in motion. One cannot even create motion in one's own body. Rather, it is God who must produce it on the occasion of volitional states. Moreover, it is not motions in one's brain that cause one's sensory states, but God who produces them on the occasion of the presence of such motions.

Unlike the argument from necessary connection, this argument from continuous creation is for the most part restricted to the case of body. There is a good reason for this restriction since the argument depends on the premise—dictated by a Cartesian understanding of the nature of body in terms of extension alone—that particular bodies cannot exist without bearing determinate relations of distance among themselves. As noted, Malebranche denies that one has a clear knowledge of the nature of the soul. No consideration of the soul could therefore reveal that it can exist only with a determinate set of modes. Indeed, Malebranche allows for the view that God creates souls with an indeterminate inclination toward "the good in general." Even so, he insists that God must be the cause of "everything real" in one's soul on the grounds that such real effects can be produced only by the power of creation. In this way the argument from continuous creation converges on the conclusion, which Malebranche claims to find in Augustine, that all creatures depend entirely on God.

The second scholastic argument against occasionalism appealed to the purported fact that it is evident to the senses that creatures have causal power. For Malebranche, however, this argument is no more persuasive than the argument that bodies must have qualities such as colors and tastes since one's senses tell one that they do. As indicated earlier, Malebranche offers Cartesian grounds for thinking that the purpose of one's sensations is not to reveal the true nature of the material world, but to indicate what is helpful or harmful to one's body. Malebranche holds that one's attribution of causal powers to bodies manifests in particular an attachment to the body that is an effect of original sin. Because of this attachment, one takes objects in the material world to be a cause of one's happiness rather than God.

In Eclaircissement XV Malebranche responds to the scholastic point that occasionalism renders scientific explanation impossible by appealing to the fact that God is not an arbitrary agent, but acts in accord with His wisdom. This wisdom dictates that He act "almost always" by means of a "general and efficacious will." Such a will produces effects that are perfectly lawlike. For instance, God acts by a general will in producing changes in bodies in accord with the law of the communication of motion. Malebranche does allow that God can produce miracles by "particular volitions" that are not lawlike. However, he emphasizes that there are relatively few such volitions in God. Thus, one can offer scientific explanations that appeal to the laws of motion that reflect the nature of God's general will.

Malebranche was not the first Cartesian to endorse occasionalism. There were followers of Descartes, such as Louis de la Forge and Claude Clerselier, who stressed that God must be the cause of the communication of motion in bodily collisions given the passivity of Cartesian matter. These Cartesians attempted to preserve some room for the action of finite minds on the body, but the Cartesian Géraud de Cordemoy went further in claiming that only God can cause changes in the material world. However, none of these thinkers went as far as Malebranche in asserting that God must produce all real changes in nature. Moreover, Malebranche is distinctive in providing an explanation of God's action that distinguishes His general will from His particular volitions.

THEODICY AND FREEDOM

The presence of various evils in the world is problematic for any theist who claims that this world was created by a God who has infinite power, knowledge, and goodness. However, the problem is particularly acute for an occasionalist, such as Malebranche, who holds that God is the only true cause of effects in nature. Malebranche offers a theodicy that addresses the problem of evil by stressing that in the "order of nature" God acts for the most part through His general will. In *Nature et de la grâce* he starts by admitting that God could have acted by particular volitions to prevent natural evils such as malformed offspring (a fitting example given his own malformed spine), and thus could have produced a more perfect world than He actually did create. However, he urges that God could have done so only by departing from simple laws, thereby sacrificing the simplicity and uniformity of action that is a supreme mark of His wisdom. God produces the natural evils that follow from simple laws not because He wills those particular effects, but because He wills a world that best reflects His wisdom by possessing the most effects governed by the fewest laws.

In his *Réflexions* on Malebranche's *Nature et de la grâce* Arnauld objects to what he takes to be the sugges-

tion in his target text that God has concern only for general features of the world and does not will the details of His effects. For Arnauld, divine providence requires that God intend all the particularities of the world He creates. There is some controversy over whether Arnauld's critique is based on a proper interpretation of Malebranche. Certain commentators follow Arnauld in thinking that Malebranche's claim in *Nature et de la grâce* that God acts by relatively few general volitions involves a rejection of the position that He has volitions for each particular effect. Others insist that this claim says only that God has volitions in accord with general laws and that the doctrine of God's continual creation in the *Entretiens* in fact requires distinct volitions for distinct effects. Some evidence for the former view is provided by the fact that Malebranche emphasizes that the laws themselves are "efficacious" and that God employs relatively few volitions in producing effects in the order of nature.

Malebranche insists that God's general will is operative not only in the order of nature but also in the "order of grace." However, he notes that the production of effects in the latter order also involves human action that is free in the strong sense of not being determined by anything external to the agent. His appeal to this sort of freedom is in fact central to his solution to the problem of moral evil, that is, the compatibility of sin with God's goodness. According to Malebranche God is not responsible for sinful action since such action derives not from Him but from sinful agents. Arnauld objects that this solution is "more pelagian than anything in Pelagius" and that one must side with Augustine, who declares Pelagianism a heresy. Malebranche responds that he does not follow Pelagius in denying the importance of grace and that Augustine himself emphasizes one's freedom in action.

Malebranche also insists that it is obvious by "inner sensation" that one is genuinely free. However, there is some question whether this introspective report is compatible with Malebranche's occasionalist claim that God is the only real cause. As indicated earlier, Malebranche does hold that God alone is the cause of one's indeterminate inclination to love the good in general. However, he insists that one is free to "consent" to the stopping of that inclination at a particular object other than God. Such consent results in an "absolute and intrinsic" love of that object that is sinful given that this love is worthy only of God. The consent is free because one is always able to suspend consent and to search for objects more worthy of one's love. Malebranche claims that one's freedom to consent or suspend consent does not conflict with occasionalism since these acts produce no "real" or "physical" change in one's mind. Sometimes he suggests that consent is nothing real because it is involves merely resting with a particular good. One problem with this suggestion is that it makes it difficult to understand how taking the opposite course of suspending consent could also involve the production of nothing real. However, Malebranche sometimes indicates that both consent and suspense produce nothing real merely in the sense that they create neither new thoughts nor an increase in inclination. He also indicates that though God determines one's "natural love" for particular objects, he leaves undetermined our "free love" for such objects.

Although Malebranche himself is less than explicit on the point, he seems at times to have left at least some room for the position that one's consent involves the determination of one's free love, whereas one's suspense involves leaving that love in its indeterminate state. In neither case is there the production of a physical change because there is no creation of new thoughts or of an increase in inclination. Whether this reflects Malebranche's own considered view is, however, a matter of scholarly dispute.

MORAL THEORY AND SELF-LOVE

The theocentrism that is evident in Malebranche's doctrines of the vision in God and occasionalism would lead one to expect that God plays a central role in his moral theory. This expectation is borne out by his remarks in the *Traité de morale*. Indeed, Malebranche's two doctrines are prominent in this work. The vision in God is reflected in the insistence that moral duties are dictated by "relations of perfection" revealed in God's wisdom. As in the case of necessary truths concerning body, so in the case of moral truths Malebranche unequivocally rejects Cartesian voluntarism. The doctrine of occasionalism is reflected in Malebranche's insistence that God is one's greatest good because He alone can cause one's happiness. This point indicates that Malebranche takes moral action to require a consideration not only of abstract relations of perfection but also of the happiness of the self.

Malebranche starts from the Augustinian position that morality concerns the proper ordering of one's love. Given the importance of human freedom for his theodicy, it is not surprising that Malebranche insists that the love required for moral action involve the free exercise of the will. In his view, the "good will" is one that freely strives to be guided in action by objective relations of perfection that hold among the various objects of love. God is the most perfect being and hence the most worthy of

one's love, whereas human beings are more perfect than mere material beings and thus more worthy of one's love. When the intensity of one's love matches the order among perfections, one has a right love that provides the basis for virtue, that is, a habitual inclination to love objects according to their perfections.

Malebranche holds that because of original sin, one is inclined not to right love directed by one's perception of relations of perfection in God's wisdom, but to a disordered love directed by bodily pleasures deriving from the soul-body union. This is the counterpart to the disordered inclination of one's will to make judgments about the nature of the material world that are based on sensations deriving from the union. For Malebranche, a corrective to both of these disorders of the will is to attend to clear ideas that exist in God.

Malebranche sometimes suggested that disordered love of bodily pleasure derives from self-love. Encouraged by this suggestion, one of his followers, François Lamy, claimed that his position leads to the quietist view in Fénelon that moral conduct requires a "pure love of God" that involves no concern for the self or its pleasure. This position, which Lamy himself endorsed, was later condemned by the Catholic Church, due in large part to a campaign against Fénelon directed by his critic, Bossuet. But Malebranche insisted that such a position directly conflicts with his own view that pleasure itself is a good that is required as a motive for action. When critics such as Arnauld and Régis charged that this view results in hedonism, Malebranche responded that it is only ordered pleasures that bring the greatest good. This response is reflected in Malebranche's claim to Lamy that a disordered love of self is to be contrasted not with pure love of God, but with an ordered love that seeks happiness in the contemplation of the greatest good, God. In emphasizing the need for this sort of love of God, Malebranche was returning to his view in the preface to the *Recherche* that it is through a union with God that the mind "receives its life, its light, and its entire felicity."

HISTORICAL INFLUENCE

Malebranche's influence on seventeenth- and eighteenth-century philosophy was significant. This is clear in the case of Leibniz, who wrote to Malebranche in 1679 that "I enthusiastically approve of the two propositions that you put forward: namely, that we see all things in God and that bodies strictly speaking do not act on us." Moreover, Leibniz's discussion in his 1684 *Discours de la métaphysique* (*Discourse on Metaphysics*) bears an evident relation to Malebranche's *Nature et de la grâce*. Here, Leibniz follows Malebranche in insisting that God acts in accord with wisdom and that He selects from among an infinity of possible worlds that world that best reflects His perfection by balancing simple laws and variety of effects. Leibniz stresses, in line with Malebranche's views, that the simplicity constraint governs both laws of nature and laws of grace.

The *Discours* also includes a section in which Leibniz comments on the Arnauld-Malebranche debate on the nature of ideas and offers some complimentary remarks concerning the Malebranchean doctrine of the vision in God. In his 1710 *Théodicée*, Leibniz highlights his agreement with the claim in *Nature et de la grâce* that natural evil exists because God's wisdom dictates that He restrict himself to a "general will." However, he also charges in this text that Malebranche's occasionalism leads to a kind of Spinozism insofar as it denies the activity and thus the substantiality of creatures. Leibniz offers his "preestablished harmony," on which creatures have the power to cause alterations in their own states. This theory, which is anticipated in the *Discours*, distinguishes Leibniz's view from Malebranche's. However, Leibniz himself sometimes presents the preestablished harmony as an internal correction to the Malebranchean system that is in accord with Malebranche's own emphasis on the perfection of divine action in creation.

Malebranche's influence extended across the Channel, where he gained admirers such as John Norris, Thomas Taylor, and Arthur Collier. His views drew a more critical reception from John Locke, who wrote *Examination of Père Malebranche's Opinion of Seeing All Things in God*, which was published posthumously in 1706. Though Malebranche himself did not respond to this work, it later received a full reply from the Savoyard cardinal, Giacinto Sigismondo Gerdil, who would have been elected pope in 1800 were it not for the veto exercised by the Austrians on political grounds. In his *Défense du sentiment du P. Malebranche*, published in 1748, Gerdil urged that Malebranche's hypothesis that God causes one's perceptions is more intelligible than Locke's own hypothesis that passive matter is the cause of these states. Because of Gerdil's influence, Malebranche's views gained a following in Italy.

During the eighteenth century Malebranche also won the grudging respect of George Berkeley and David Hume. Berkeley indeed appeared to his critics to be a "Malbranchiste de bonne foi," a view that Berkeley himself counters when he writes in the third (1734) edition of his *Three Dialogues between Hylas and Philonous* that "there are no principles more fundamentally opposed

than [Malebranche's] and mine." Berkeley does differ from Malebranche in rejecting the existence of an external material world, in insisting that ideas exist in one's mind rather than in God's and in claiming that the senses reveal immediately the true nature of sensible objects. However, Berkeley follows Malebranche in rejecting the Aristotelian conception of nature and in attributing causal efficacy in natural interactions to God (though Berkeley does attempt, with questionable success, to leave room for the power of finite spirits to move their own bodies). Also, Berkeley holds with Malebranche that one's perceptions are related to certain "archetypes" in the divine mind that serve as the pattern for God's creation (Luce [1934] is the classic study of the relation between Berkeley and Malebranche).

In 1737 Hume wrote to his friend Michael Ramsey that he should prepare himself for "the metaphysical Parts" of the reasoning in the forthcoming *Treatise of Human Nature* (1739–1740) by reading "once over la Recherche de la Vérité of Pere Malebranche," along with selected works from Descartes, Berkeley, and Pierre Bayle. Malebranche is important primarily for the account of causation and causal belief in the *Treatise*. Hume relies there explicitly on Malebranche's argument for the negative conclusion that neither external nor internal experience affords one any idea of power. With Malebranche, Hume emphasizes the importance of necessary connection to the understanding of causation. Hume does reject Malebranche's own claim that God is the only real cause, noting in a famous passage from the *Enquiry concerning Human Understanding* (1748) that with such a claim "we are got into fairy land, long ere we have reached the last steps of our theory." Hume's preference is for a psychological account of causal belief that sticks closely to "common life and experience" and that emphasizes the central role of the imagination. Nonetheless, Hume's own discussion belies his remark in the *Enquiry* that "the glory of Malebranche is confined to his own nation, and to his own age."

See also Arnauld, Antoine; Aristotle; Augustine, St.; Bayle, Pierre; Berkeley, George; Bossuet, Jacques Bénigne; Cartesianism; Chinese Philosophy; Collier, Arthur; Descartes, René; Desgabets, Robert; Determinism and Freedom; Ethics, History of; Evil, The Problem of; Fénelon, François de Salignac de la Mothe; Foucher, Simon; General Will, The; Hume, David; Leibniz, Gottfried Wilhelm; Locke, John; Newton, Isaac; Norris, John; Pascal, Blaise; Pelagius and Pelagianism; Régis, Pierre-Sylvain; Spinoza, Benedict (Baruch) de; Spinozism; Suárez, Francisco; Thomism; Volition; Voluntarism.

Bibliography

WORKS BY MALEBRANCHE

Oeuvres complètes de Malebranche. 20 vols, edited by André Robinet. Paris: J. Vrin, 1958–1984.

Dialogue between a Christian Philosopher and a Chinese Philosopher on the Existence and Nature of God. Translated by Dominick A. Iorio. Washington, DC: University Press of America, 1980.

Treatise on Ethics. Trans. C. Walton. Dordrecht: Kluwer, 1993.

Malebranche's First and Last Critics: Simon Foucher and Dortous de Mairan. Translated by Richard A. Watson and Marjorie Grene. Carbondale: Southern Illinois University Press, 1995.

Dialogues on Metaphysics and on Religion. Translated by David Scott; edited by Nicholas Jolley. New York: Cambridge University Press, 1997a.

The Search after Truth. Translated and edited by Thomas M. Lennon and Paul J. Olscamp. New York: Cambridge University Press, 1997b.

WORKS ON MALEBRANCHE

Alquié, Ferdinand. *Le cartésianisme de Malebranche*. Paris: J. Vrin, 1974.

André, Yves M. *La vie du R. P. Malebranche*. Geneva: Slatkine Reprints, 1970.

Arnauld, Antoine. *On True and False Ideas*. Translated by Elmar J. Kremer. Lewiston, NY: Edwin Mellen Press, 1990.

Bardout, Jean-Christophe. *Malebranche et la métaphysique*. Paris: Presses Universitaires de France, 1999.

Brown, Stuart, ed. *Nicolas Malebranche: His Philosophical Critics and Successors*. Assen, Netherlands: Van Gorcum, 1991.

Chappell, Vere, ed. *Nicholas Malebranche*. New York: Garland, 1992.

Connell, Desmond. *The Vision in God: Malebranche's Scholastic Sources*. Louvain, Belgium: Nauwelaerts, 1967.

Easton, Patricia, Thomas M. Lennon, and Gregor Sebba, eds. *Bibliographia Malebranchiana: A Critical Guide to the Malebranche Literature into 1989*. Carbondale: Southern Illinois University Press, 1992.

Foucher, Simon *Critique de la recherche de la verité,* edited by Richard A. Watson. New York: Johnson Reprint, 1969.

Gueroult, Martial. *Malebranche*. 3 vols. Paris: Aubier, 1955–1959.

Jolley, Nicholas. *The Light of the Soul: Theories of Ideas in Leibniz, Malebranche, and Descartes*. Oxford, U.K.: Clarendon Press, 1990.

Luce, A. A. *Berkeley and Malebranche: A Study in the Origins of Berkeley's Thought*. London: H. Milford, 1934.

McCracken, Charles J. *Malebranche and British Philosophy*. Oxford, U.K.: Clarendon Press, 1983.

Moreau, Denis. *Deux cartésiens: La polemique entre Antoine Arnauld et Nicolas Malebranche*. Paris: J. Vrin, 1999.

Nadler, Steven, ed. *The Cambridge Companion to Malebranche*. Cambridge, U.K.: Cambridge University Press, 2000.

Nadler, Steven. *Malebranche and Ideas*. New York: Oxford University Press, 1992.

Pyle, Andrew. *Malebranche*. London: Routledge, 2003.

Radner, Daisie. *Malebranche: A Study of a Cartesian System*. Assen, Netherlands: Van Gorcum, 1978.

Robinet, André. *Système et existence dans l'oeuvre de Malebranche*. Paris: J. Vrin, 1965.

Rodis-Lewis, Geneviève. *Nicolas Malebranche*. Paris: Presses Universitaires de France, 1963.

Schmaltz, Tad M. *Malebranche's Theory of the Soul: A Cartesian Interpretation*. New York: Oxford University Press, 1996.

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Sebba, Gregor, ed. *Nicolas Malebranche, 1638–1715: A Preliminary Bibliography*. Athens: University of Georgia Press, 1959.

Walton, Craig. *De la Recherche du Bien: A Study of Malebranche's Science of Ethics*. The Hague: Nijhoff, 1972.

Tad M. Schmaltz (2005)

MALRAUX, GEORGES-ANDRÉ

(1901–1976)

Georges-André Malraux, the French author, critic, revolutionist, and statesman, was born in Paris to a well-to-do family. He studied at the Lycée Condorcet and the Institut des Langues Orientales and early in life developed an enduring interest in archaeology, art, and Oriental languages and thought. His life and writing were characterized by a restless, questioning, quasi-apocalyptic intensity that is fully understandable only in terms of the crisis with which Western thought was confronted in the first half of the twentieth century: At grips with a fast-accumulating mass of new knowledge, Western civilization was seeking to adjust to the violent changes that had disrupted its former social, intellectual, and spiritual framework of values.

In 1923 Malraux went on an archaeological expedition into the Cambodian jungle, and soon afterward he returned to the Orient to participate in the revolutionary struggle that was transforming the Asiatic world. He seems at the time to have been in sympathy with the Marxist ideology. *La tentation de l'occident* (Paris, 1926), his first serious work, is a fictional dialogue between a Chinese and a European intellectual and shows how decisive was his first encounter with the Orient. It intensified Malraux's self-styled obsession with the notions of civilization and culture. He was always vitally concerned with the problems of the life and death of civilizations; the specificity, irreducibility, and relativity of all cultures; their determining action in shaping the mental structures of individuals; and the bearing on his own cultural world of the observations and conclusions of historians and anthropologists such as Oswald Spengler and Leo Frobenius. This initial obsession was nourished and substantiated by Malraux's legendary familiarity with all realms of art (painting and sculpture in particular); his avid and exceptionally broad grasp of literature; and his addiction to passionate debate with leading personalities in Europe and the Orient. Although his thought was always concentrated on a present unremittingly interrogated, it developed within vast perspectives both in time and space.

In the late 1920s Malraux, as art editor for the Gallimard publishing firm in Paris, traveled widely in search of art treasures, while actively participating in the unavailing struggle of the European intellectuals against fascism, Nazism, and anti-Semitism. He later commanded a group of aviators for the Republican forces in the Spanish Civil War, was active in the French resistance after 1940, and became, first, minister of information, then minister of cultural affairs, in the cabinet of General Charles de Gaulle.

He was deliberately "committed" as a writer for intellectual reasons. Western science, he claimed, offers a set of relationships that define the cosmos but, by omitting the observer, it presents a cosmos in which man has no place. According to Malraux, psychoanalysis has revealed the blind, destructive forces at work within the self and has put into question the very notion of a fundamental human personality. To recover some concept of man, Malraux maintained that one must once again examine what man does, thereby redefining his powers. The image of the rational, detached observer—scientist or philosopher—placed outside the world he observes must therefore give way to the participant who is, as it were, a knot of relations with the world. Malraux often reiterated that man "is what he does." Participation therefore was the first and necessary stage in his search for definition.

The elucidation of an action is the theme of his novels. All revolve around the question, "What can a man best do with his life?"; all are animated by the same answer that is given in *Man's Hope*: "Transform into consciousness an experience as broad as possible." Writing is the medium through which this transformation takes place; hence the intensity of the process, the inner questioning, and the many-faceted debate that it embodies. His six widely read novels all are wrenched from stages of his own experience: *Les conquérants* (Paris, 1928); *La voie royale* (Paris, 1930); *La condition humaine* (Paris, 1933); *Le temps du mépris* (Paris, 1935); *L'espoir* (Paris, 1937);

and *Les noyers de l'Altenburg* (Lausanne, 1943), the first volume of a two-part novel whose second part was destroyed by the Nazis. These were followed by an impressive series of works on art: *Goya* (Geneva, 1947); *La psychologie de l'art* (3 vols., Geneva, 1947, 1949, 1950); *Le musée imaginaire de la sculpture mondiale* (3 vols., Paris, 1952, 1953, 1954); *Les voix du silence* (Paris, 1953); and *La métamorphose des dieux* (Paris, 1960). A number of reviews, prefaces, and speeches add to this abundant corpus of work.

Despite both the variety of his media and the obscurities inherent in his manner of writing, there is a remarkable degree of consistency and lucidity in Malraux's thought, questionable though many of his assumptions and examples may be. He posits as premise the definitive disappearance from Western civilization of the structure of values established by the Christian *Weltanschauung.* Western man is thus left face to face with a cosmos to which he cannot relate. However, he is still in possession of the inner drive that, since the Greeks, has structured his world—the need to create a coherent, intelligible image of man's fate that gives significance to each individual life. Hence the double burden of lucidity and anguish characteristic of our time, hence its "temptations." The most prevalent is the nihilism whereby Western man, living in a state of "metaphysical distraction," renounces his drive toward lucidity and submits to blind necessity and to natural and social conditioning. This, according to Malraux, is an intolerable reversion to the "demons," that is, to the blind animal instinct within us. Malraux also examined and partially rejected the Asian resorption of the individual into the cosmos (considered as divine). In preference to the Asian view, he sought to define man's power in his capacity to "leave a scar on the planet," to transform his environment. For a while he understood the process in terms of the Marxist theory of history.

Malraux's final view emerged from his meditations on art. It is a complex outlook related to the study of art styles and their migrations and metamorphoses, an approach that is characteristic of such art historians as Élie Faure and Henri Focillon. In brief, for Malraux a new planetary civilization that has destroyed all significant cultures is now in the making. The structures of values whereby each individual within a human society relates to the cosmos, to the community, and to his own actions now exist only as "relativized absolutes." This is the first agnostic civilization, the first that does not relate to some form of the divine. It also presents a new phenomenon, the "imaginary museum," in which all works of art—whatever their origin—are available, to be perceived as significant in themselves and not for what they once signified. For Malraux this universal presence and significance testifies to a fundamental power of humankind: the power to dominate and transcend fate and to create a universe in some way accessible to all men, who are thereby freed from time, death, and blind necessity. The privileged potential image of humankind, therefore, that Malraux detects as indicative of our present orientation is that of man as creator and as forger of his own freedom. Malraux thus formulated in new terms the age-old problem of freedom and destiny, to serve as the foundation for a new ethic. His work is fundamentally relevant in an age that is deeply preoccupied with the working of the mind, considered on one hand as a form of conditioned mechanism and on the other as a principle of free activity, order, and meaning.

See also Aesthetics, History of; Agnosticism; Art, Expression in; Marxist Philosophy; Nihilism; Spengler, Oswald; Value and Valuation.

Bibliography

WORKS BY MALRAUX

The works listed below are English editions of Malraux's works in the order in which they appear in the text.

The Temptation of the West. Translated with an introduction by Robert Hollander. New York: Vintage, 1961.

The Conquerors. Translated by W. S. Whale. New York: Harcourt Brace, 1929.

The Royal Way. Translated by Stuart Gilbert. New York: Smith and Haas, 1935.

Man's Fate. Translated by Haakon Chevalier. New York: Smith and Haas, 1934.

Days of Wrath. Translated by Haakon Chevalier. New York: Random House, 1936.

Man's Hope. Translated by Stuart Gilbert and Alistair MacDonald. New York: Random House, 1938.

The Walnut Trees of Altenburg. Translated by A. W. Fielding. London: Lehmann, 1952.

Saturn: An Essay on Goya. Translated by C. W. Chilton. London: Phaidon, 1957.

The Psychology of Art. Translated by Stuart Gilbert. Vol. I: *Museum without Walls*; Vol. II: *The Creative Act.* New York: Pantheon, 1949–1951.

The Voices of Silence. Translated by Stuart Gilbert. New York: Doubleday, 1953.

The Metamorphosis of the Gods. Translated by Stuart Gilbert. New York: Doubleday, 1960.

WORKS ON MALRAUX

Blend, Charles. *André Malraux: Tragic Humanist.* Columbus: Ohio State University Press, 1963. A biography and general critical study, with a bibliography of Malraux's work and a brief critical bibliography.

Frohock, Wilbur. *André Malraux and the Tragic Imagination.* Stanford, CA: Stanford University Press, 1952. A basic work.

Lewis, R. W. B., ed. *Malraux.* Englewood Cliffs, NJ: Prentice-Hall, 1964. A collection of critical essays.

Lyotard, Jean Francois. *Soundproof Room: Malraux's Anti-Aesthetics*. Translated by Robert Harvey. Stanford: Stanford University Press, 2001.

Raymond, Gino. *Andre Malraux: Politics and the Temptation of Myth*. Aldershot: Avebury; Brookfield: Ashgate, 1995.

Righter, William. *The Rhetorical Hero: An Essay on the Aesthetics of Andre Malraux*. New York: Chilmark Press, 1964.

Vandergars, André. *La jeunesse littéraire d'André Malraux.* Paris: Pauvert, 1964. Contains a wealth of information on Malraux's Indochinese activities.

Germaine Brée (1967)
Bibliography updated by Desiree Matherly Martin (2005)

MALTHUS, THOMAS ROBERT

(1776–1834)

Thomas Robert Malthus, the English economist and moral philosopher, is most famous for his contributions to population studies. In his *Principles of Political Economy* (1820) and in his controversies with David Ricardo, Malthus seems partly to have anticipated J. M. Keynes; and Keynes himself, in his *Essays in Biography*, generously remarked that "if only Malthus, instead of Ricardo, had been the parent stem from which nineteenth century economics proceeded, what a much wiser and richer place the world would be today!"

Malthus's work on population is contained in two books, misleadingly presented as if they were merely different editions of one. The first, best referred to as the *First Essay*, is actually titled *An Essay on the Principle of Population as It Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*. The second, best thought of as the *Second Essay*, was, with some reserve, offered by Malthus as a much extended second edition. But it was retitled *An Essay on the Principle of Population, or a View of Its Past and Present Effects on Human Happiness with an Inquiry into Our Prospects Respecting the Future Removal or Mitigation of the Evils Which It Occasions*. The *First Essay* is an occasional polemic against utopianism; the *Second*, a labored treatise full of detailed factual material. What they have in common is the same guiding and coordinating theoretical schema, although even this is in one respect importantly amended in the later book.

The fundamental principle is that unfreakish human populations possess a power of multiplying in a geometrical progression. The next step is to urge that this power always is and must be checked by countervailing forces; for, on the most optimistic supposition, means of subsistence could in the long run at best be increased only in an arithmetical progression. (The subsistence of checks could, of course, be inferred without recourse to this misleadingly arithmetized supposition, by referring directly to the fact that no human population ever does achieve its full multiplicative potential.) The questions then arise. What are these checks? what ought they to be?

Checks are classified in two different ways. First, they can be positive or preventive: the former by the time of the *Second Essay* being all causes of (premature) death; and the latter, correspondingly, all checks on the birth rate. The second classification is strongly normative: In the *First Essay* all checks must count as either misery or vice; but in the *Second Essay* a third option, moral restraint, is added. This is defined as "the restraint from marriage which is not followed by irregular gratifications." Malthus seems never to have entertained the possibility of restraint within marriage; and he categorically rejected any form of contraception, even within wedlock, as vice.

This scheme of ideas constituted an intellectual engine that was immensely powerful both for its primary purpose of confounding utopian optimism and for its secondary function of guiding social inquiry. We also have clear statements from both Charles Darwin and Alfred Russel Wallace that it was reading Malthus on population which independently led each to see the clue to the problem of the origin of species in natural selection through "a struggle for existence," a phrase used by Malthus himself. Against the utopians the argument was that our inordinate animal power of multiplication is bound—sooner or later, and usually sooner—to run up against the inexorably constricting walls of scarcity. All measures of intended amelioration which directly or indirectly encourage an increase of population that outstrips resources—and most do—will, in the not very distant end, merely multiply the number of bearers of misery and agents of vice. These harsh and gloomy conclusions were only modified, not upset, by the belated recognition of the option of moral restraint. For it was, and remains, hard to cherish high hopes from the preaching of such prudence; and in any society which did generally accept such preaching all but the richest would have to marry women nearing the evening of their reproductive powers.

It is, therefore, not surprising that generations of idealists hoping to reshape the present sorry scheme of things nearer to their heart's desire have released torrents of argument and abuse at "Parson Malthus" and his ideas. Yet, despite the apparent implication of his system—that God has placed humankind in a situation offering little promise of secure improvement—it would be wrong to assume that Malthus as a man or as a thinker was either insensitive or harsh. Compared with the optimistic utopians of his father's reading and acquaintance he could not but appear a jarring pessimist. But this was a matter of facing what he took to be the sober facts of the human condition, not of callous indifference to the relief of man's estate. To quote Keynes again, his work is really in "the tradition which is suggested by the names of Locke, Hume, Adam Smith, Paley, Bentham, Darwin and Mill, a tradition marked … by a prosaic sanity … and by an immense disinterestedness and public spirit." As against, say, Condorcet, who wrote of inevitable progress while under the shadow of the guillotine, Malthus was concerned first with finding what the facts are and then with discovering how, in the light of those perhaps recalcitrant facts, we are to do the best we can. It is no accident that in the first chapter of the *First Essay* he acknowledges a debt to David Hume and Adam Smith but not to the impossible and visionary Jean-Jacques Rousseau, whom his father had known and admired.

THEODICY

The same intellectual associations are seen in his theodicy. William Paley was one of the early converts to Malthus on population, and appropriately, Paley was one of Malthus's favorite theologians. So Malthus insists in the *First Essay* that "Evil exists in the world not to create despair but activity." (It was from this part of the work that Darwin and Wallace most directly derived the idea of a necessary struggle for existence.) What Malthus may have acquired from the dissenting Christians and Unitarians of his father's circle is a note of theological radicalism, a note not caught either by the hostile conventional left, represented then by Wiliam Cobbett and William Hazlitt, or by such sentimental conservative opponents as Samuel Taylor Coleridge and Robert Southey.

In the theodicy of the last chapter of the *First Essay* Malthus boldly steps away from Paley and from the whole tradition of Christian orthodoxy by insisting that "it is perfectly impossible to conceive that any … creatures of God's hand can be condemned to eternal suffering. Could we once admit such an idea, all our natural conceptions of goodness and justice would be completely overthrown, and we could no longer look on God as a merciful and righteous Being." (Malthus settles his own account with Christianity by accepting the Hobbist interpretation; that eternal death means eternal death and not eternal life in torment. The "doctrine of life and immortality which was brought to light by the gospel" is "the doctrine that the end of righteousness is everlasting life, but that the wages of sin are death." This plausible reading had been unanimously rejected by the orthodox Saints and Fathers, doubtless as being unacceptably merciful.)

CRITIQUE OF POPULATION THEORY

As a heuristic and explanatory scheme, the population theory resembles bits of classical physics, although it might also be usefully compared with that of Darwinism. The fundamental principle is like the first law of motion in that both describe not what does go on but what would go on if there were no counteracting forces; and in both cases the main theoretical function of the basic law is to generate questions about such forces and checks. Again, Malthus in classifying checks always aims at complete, exhaustive lists; and his arguments often depend on his appreciation that the values of the various checks considered as variables will be, for a given population, inversely connected: the bigger the sum of the preventive checks, the smaller the sum of the positive checks; and so on. These are similarities of which Malthus himself—thanks to his mathematical training at Cambridge—seems to have been aware. (It is doubtless to the same training that we owe his introduction of the supposition of the arithmetical progression to which, and to the consequent comparison of the two progressions, is due much of the appearance of "mathematical certainty" in his demonstrations.)

Malthus never tied up all the various minor logical loose ends in his original conceptual scheme, although he added important appendices to the third and fifth editions of his work in 1806 and 1817 and wrote the article "Population" for the 1824 supplement to the *Encyclopaedia Britannica* (revised and published separately as his last word in 1830). But the main objections to Malthus that emerged from the enormous controversy are two, one moral and one logical. The moral objection repudiates Malthus's total rejection of contraception. It is this repudiation, combined with acceptance of Malthus's warnings on the dangers of overpopulation, which makes a Neo-Malthusian. The suggestion sometimes heard that the spread of contraception has made Malthusian ideas obsolete should be seen as manifestly absurd. Contraception is one kind of preventive check; none at all would be

required if the multiplicative power was not still there to be checked.

The second objection insists on a distinction, which Malthus was forever inclined to overlook, between two senses of *tendency*. A tendency to produce something may be a cause which, operating unimpeded, would produce it. But to speak of a tendency to produce something may also be to say that the result is one that may reasonably be expected to occur in fact. This point seems to have been put against Malthus for the first time by Nassau Senior in his *Two Lectures on Population* (1831) and was grudgingly accepted. It was developed in the following year by Archbishop Whateley in *Lectures on Political Economy* (ninth lecture).

If both these objections are accepted, it becomes possible to recognize the Malthusian menace but to insist that the tendency to catastrophe does not have to be a tendency in the second sense—not if people can be persuaded to employ the means which science has and will put into our hands. Yet Malthus must have the last word. For it was he who most dramatically and powerfully drew attention to an absolutely vital fact, a fact that is still persistently and often disastrously ignored. It is, in the words of Senior, that "no plan for social improvement can be complete, unless it embraces the means both of increasing production, and of preventing population making a proportionate advance."

See also Keynes, John Maynard.

Bibliography

WORKS BY MALTHUS

An Essay on the Principle of Population as It Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers. London: J. Johnson, 1798. Facsimile edition (London, 1926); paperback edition with introduction by K. E. Boulding (Ann Arbor: University of Michigan Press, 1959).

An Essay on the Principle of Population, or a View of Its Past and Present Effects on Human Happiness, with an Inquiry into Our Prospects Respecting the Future Removal or Mitigation of the Evils Which It Occasions. London, 1803. There is an Everyman Library edition (London and New York); the version now in print has an introduction by M. P. Fogarty.

Glass, D. V., ed. *Introduction to Malthus*. London: Watts, 1953. Includes discussion by Glass and others and an appendix of two things by Malthus.

T. R. Malthus: The Unpublished Papers in the Collection of Kanto Gakuen University.

The Malthus Library Catalogue: The Personal Collection of Thomas Robert Malthus at Jesus College, Cambridge. New York: Pergamon Press, 1983.

The Pamphlets of Thomas Robert Malthus. New York: A.M. Kelley, 1970.

An Essay on the Principle of Population: Text, Sources and Background, Criticism. New York: Norton, 1976.

Principles of Political Economy. Cambridge; New York: Cambridge University Press, 1989.

The Works of Thomas Robert Malthus. Edited by E. A. Wrigley and David Souden. London: W. Pickering, 1986, 1798.

WORKS ON MALTHUS

Bonar, J. *Malthus and His Work*. 2nd ed. New York: Macmillan, 1924.

Flew, Antony. "The Structure of Malthus' Population Theory." In *Philosophy of Science: The Delaware Seminar*. Vol. I, edited by B. Baumrin. New York: Interscience Publishers, 1963.

Keynes, J. M. "Robert Malthus: The First of the Cambridge Economists." In *Essays and Sketches in Biography*. London, 1951.

Levin, S. M. "Malthus and the Idea of Progress." *Journal of the History of Ideas* 27 (1) (1966): 92–108.

Petersen, William. *Malthus*. Cambridge, MA: Harvard University Press, 1979.

Senior, Nassau William. *Two Lectures on Population*. London, 1831.

Turner, Michael Edward. *Malthus and His Time*. New York: St. Martin's Press, 1986.

Waterman, Anthony Michael C. *Revolution, Economics, and Religion: Christian Political Economy, 1798–1833*. Cambridge, U.K.; New York: Cambridge University Press, 1991.

Whateley, Richard. *Lectures on Political Economy*. London, 1832.

Winch, Donald. *Malthus*. Oxford; New York: Oxford University Press, 1987. New York: Cambridge University Press, 1997–2004.

Antony Flew (1967)

Bibliography updated by Michael J. Farmer (2005)

MAMARDASHVILI, MERAB KONSTANTINOVICH
(1930–1990)

Merab Mamardashvili was born September 15, 1930, in Gori, Gerorgia and died November 25, 1990, in Moscow. He was a philosopher most of whose creative life passed in Moscow and Tbilisi, Georgia, in the period from the 1950s through the 1980s. He was an original thinker who received world recognition. His main spheres of inquiry were the philosophy of consciousness, the theory of transformed forms of consciousness; classical and nonclassical forms of rationality; the phenomenology of life, love, and death; proof of the necessity of Cartesian, Kantian, and Husserlian themes as "elements" or dimensions of all philosophizing; problems of the existence, consciousness, and action of man under the conditions of

socialism and of the Soviet regime; contemporary civilization and the "anthropological catastrophe."

Mamardashvili graduated from the philosophy department of Moscow University in 1954 and completed his graduate studies there in 1957. He was on the editorial staff of the journals *Voprosy filosofii* [Questions of philosophy] (1957–1961) and *Problemy mira isotsializma* [Problems of the world and of socialism] (1961–1966). He then worked in a number of institutes of the Academy of Sciences (the Institute of the International Workers Movement and the Institute of the History of Natural Science and of Technology); from 1968 to 1974, he was associate editor-in-chief of *Voprosy filosofii*. From 1980 to 1990, he lived in Tbilisi, where he worked in the Institute of Philosophy of the Georgian Academy of Sciences. From 1972, he was a professor of philosophy.

Having been formed in the period of the "thaw" in the 1950s and having by the 1960s become an original thinker, an opponent of socialism and of the political regime existing then in the USSR (although without being an open dissident), Mamardashvili was compelled to expound his ideas not so much in published works that were subject to censorship, as in lecture courses, which attracted hundreds of listeners. In view of his ability to expound the most complex and recondite philosophical ideas in oral form, he was called "the Georgian Socrates."

Some of Mamardashvili's lecture courses were given in France, Italy, and other countries: He was fluent in a number of foreign languages. His popularity and his recognition as a talented philosopher grew. But the opposition of the authorities, who persecuted him, also grew. That is why during his life he was able to publish only three books: *Formy i soderzhanie myshleniia. K kritike gegelevskogo ucheniia o formakh poznaniia* (*Forms and Content of Thought. Toward a Critique of Hegel's Doctrine of the Forms of Knowledge*), Moscow, 1968; *Klassicheskii i neklassicheskii idealy ratsional'nosti* (*Classical and Nonclassical Ideals of Rationality*), Tbilisi, 1984; *Kaki ia ponimaiu filosofiiu* (*How I Understand Philosophy*), Moscow, 1990; as well as articles in journals and collected works. There is a principal difficulty in assimilating and evaluating Mamardashvili's philosophical ideas: The tape recordings of his lectures that served as the basis of the works published under his name after his death were edited and modified by the editors and publishers. Because of this, these books are secondary sources whose status is ambiguous: They are integral parts of Mamardashvili's philosophical heritage, but at the same time a number of specialists view them as inauthentic.

MAMARDASHVILI'S MAIN SPHERES OF INQUIRY AND HIS PRINCIPAL IDEAS

Mamardashvili dealt in four major spheres in his lifetime. His principal ideas and concepts are outlined below and a general explanation is given for his contribution to philosophy.

1. ANALYSIS OF CONSCIOUSNESS AND OF THE TRANSFORMED FORMS OF CONSCIOUSNESS IN THE WORKS OF KARL MARX. For Mamardashvili, as well as for a number of other influential philosophers of Russia of the Soviet period, reference to Marx became a means of struggle with the dogmas of dialectical and historical materialism, as well as a means of grounding his own ideas. In Mamardashvili's exposition, the chief of these are: "The Marxian schemata give rise to the elements of a series of theories: to the elements of (1) a theoretical model of the social conditionedness of consciousness; (2) a theory of fetishism and of the symbolics of the social in consciousness; (3) a theory of ideology (the socio-philosophical critique of ideology developed by Marx was subsequently transformed into that which is now called the sociology of knowledge as an academic discipline); (4) a theory of science and of free spiritual production as particular forms of active consciousness; (5) a theory of consciousness as an instrument of man's personal development and of his responsibility in the sphere of culture and historical activity" (*How I Understand Philosophy*, Moscow, 1990, pp. 299–300). Later Mamardashvili will say that he found his way to phenomenology not through Husserl but through Marx, who revealed "the phenomenological nature of consciousness, its quasi-objective character," but—in contradiction to the phenomenology of the twentieth century—always disclosed "behind phenomena" their causal origin and "the social system of communion, which the phenomena of consciousness serve" (p. 303).

To this is appended an interpretation of the concept of "the transformed forms of consciousness," which we already encounter in Marx, but to which Mamardashvili attributes a broader and more profound theoretical significance. According to Mamardashvili, the transformed forms are characterized by the fact that "the form of manifestation acquires an 'essential' significance, is particularized, and content is replaced in the phenomenon by another relation, which merges with the property of the material bearer (substrate) of the form itself (for example, in cases of symbolism) and takes the place of the real relations" ("*Forma prevrashchennaia*" [*Transformed Form*] in *Filosofskaia entsiklopediia* [*Philosophical Ency-*

clopedia], vol. 5, Moscow, 1970, p. 387). Examples of this are capitalized cost in the system of bourgeois economics (the case of an irrational transformed form); objective appearance: the movement of the sun and planets around the earth; the operation of sign forms of culture; memory and coding units in computers; and the symbolic processing of links of consciousness (according to Freud).

II. EXISTENTIALISM AND FRENCH MARXISM. It was early on that Mamardashvili began his polemic with existentialism and with French Marxism. He personally debated Sartre and Althusser. During the 1950s and 1960s, like these French authors whom he critically analyzed, Mamardashvili based his thought on Marx's conception, but he was also developing an original conception of society and man. At the center of Mamardashvili's positive analysis was a theory of personality and alienation which rejected Sartre's conception of nature, matter, and the material in socio-historical life: "Taking as his point of departure a phenomenological analysis, Sartre can see in the manifestations of social 'matter' (i.e., the fact of the existence in society of forces and relationships which are independent of individuals and their consciousness) only an extra-human and mysterious power, which bewitches people and their relationships and weaves together with them the thread of factual history" ("*Kategoriia sotsial'nogo bytiia i metod ego analiza v ekzistentsializme Sartra*" [*The Category of Social Being and Its Method of Analysis in Sartre's Existentialism*] in *Sovremennyi ekzistentializm* [*Contemporary Existentialism*], Moscow, 1966, p. 187).

III. COMPARATIVE ANALYSIS OF RATIONALITY. Mamardashvili devoted a number of his works to a comparative analysis of the classical and non-classical types or ideals of rationality. He discerned the specific character of the classical type of rationality in the following features: (1) the concept of the "objective" in the "classical" type was identified with the external (the spatial), while the spatial was identified with the material, which had important philosophical and methodological consequences; (2) "from within the physical theory, which investigates natural phenomena and comes to a certain objective and intelligible picture of the world, we cannot (from within this theory itself) understand those means which we use to construct this picture" (*The Classical and Non-classical Ideals of Rationality*, p. 5). The understanding of the physical world is bought at the cost of a "lack of scientific understanding" of conscious phenomena (although, as living beings, we freely live and orient ourselves in this sphere). Other features include the principles of classical rationality: "the principle of the continuity of reproducible experience," "the self-identity of the subject" (p. 9); and reliance on the concept of "phenomenon"; de-anthropomorphization. Non-classical rationality arises under the influence of the theory of relativity and quantum mechanics; and in the social and humanitarian disciplines, it arises under the influence of the theory of Marx's ideology, Husserl's phenomenology, and Freud's psychoanalysis. The main principles and procedures of non-classical rationality are: (1) phenomenon instead of appearance, for "I return to the phenomemological level, which prohibits us from discussing something without first bringing to a stop the premises of our objectifying thought …" (p. 50); (2) the refusal to accept the existence of some "preestablished world with ready-made laws and essences" (p. 64); (3) a complete and comprehensive understanding that consciousness is "one of the inalienable elements of the very object of investigation" (p. 79).

IV. INTERPRETATIONS OF DOCTRINES OF PROMINENT PHILOSOPHERS. The central place in Mamardashvili's philosophy is occupied by a particular interpretation of the doctrines of a number of prominent thinkers and cultural figures of the past (see the posthumous *Kartezianskie razmyshleniia* [Cartesian Meditations]; *Kantianskie variatsii* [Kantian Variations], Moscow, 1997; and *Lektsii o Pruste* [Lectures on Proust], Moscow, 1995). The originality of this interpretation consists in a free transition from an abstractly philosophical analysis of the doctrines of Descartes or Kant to an illumination of the socio-historical content as well as the trans-historical cultural, moral, aesthetic, and personal content contained in these doctrines. As a result, the philosophical consciousness is closely interwoven with the radical problems, contradictions, and crises of civilization, with orientations of the human personality that have meaning for life. This is realized, for example, in the historico-philosophical as well as socio-philosophical figure of the three "K's": "Kartesius" (Descartes), Kant, and Kafka.

In the interpretation of Descartes the central plane is occupied by the theme of *cogito*, which Mamardashvili calls "the phenomenon of all phenomena," as well as by the paths leading to *cogito*. The consciousness of *ego cogito* is interpreted, on the one hand, as a limit abstraction from all that is historically concrete, even from man, a limit abstraction which implies the "permissibility" and even the inevitability of transcendentalism (in the traditions of Descartes, Kant, and Husserl). On the other hand, this "improbable abstraction" is realized, after which it "becomes in a concealed manner the founda-

tion" of our physical knowledge and of the formulation of physical laws, although it is scarcely the case that we are always conscious of the "accomplishment" of the abstraction. Here, the abstraction of the transcendental *ego* acquires social, personal, and moral foundations and consequences. What this means is that thought is free and thus "paths of *coherent* space must be laid for thought, i.e., paths of open discussion (*glasnost*), mutual tolerance, formal legality…" (*Lektsii o Pruste*, p. 115).

The second "K" (Kant) in Mamardashvili's interpretation indicates the conditions under which man—a finite, mortal being, whose life could have become meaningless in the face of infinity—creates around himself a special world, a world which presupposes choice, evaluations, decisions; in other words, freedom. This is because everyone who is born not only enters the world of nature with its rigid causal connections, but also encounters and in part creates the world of "intelligible" objects. These latter, according to Mamardashvili, are "*images* of integralities," as if designs and projects of development.

The third "K" is a figurative reference to the "world of Kafka," i.e., to the penetration into the human world of certain "zombie-situations," attesting to the "degeneration" or "regressive variant" of the general K-principle: In opposition to *Homo sapiens*, n other words, to "man who knows good and evil," a "strange man," an indescribable man, enters the world of civilization. "Ridiculous, absurd, bizarre, dreamlike confusion and something otherworldly"—that is how Mamardashvili describes the actions of Joseph K. in Kafka's *Trial*, and this also goes for the situation of the absurd in human society. With the accumulation of the potential of the absurd in human history, including contemporary history, the result can be the most dangerous chaos of civilization, a kind of anthropological catastrophe. "Terrifying idols of passion, soil, and blood cover the world, concealing the hidden paths of order; and it is very difficult to tear oneself away from these idols, and to enter onto the radiant paths of thought, order, and harmony" (p. 210).

See also Cartesianism; Descartes, René; Existentialism; Freud, Sigmund; Husserl, Edmund; Kafka, Franz; Kant, Immanuel; Marx, Karl; Quantum Mechanics; Rationality; Relativity Theory; Russian Philosophy; Sartre, Jean-Paul.

Bibliography

Kruglikov, V., ed. *Kongenial'nost' mysli: O filosofe Merabe Mamardashvili* (*Congeniality of Thought: On the Philosopher Merab Mamardashvili*). Moscow, 1999.

Kruglikor, V., and J. Schokasov, eds. *Vstrecha s Dekartom* (*Encounter with Descartes*). Moscow, 1996.

Mamardashvili, Merab. *Estetika myshleniia* (*Aesthetics of Thought*). Moscow, 2000.

Mamardashvili, Merab. *Méditationes cartésiennes*. Solin, 1997.

Mamardashvili, Merab. *Moi opyt netipichen* (*My Experience Is Atypical*). Moscow, 2000.

Mamardashvili, Merab. *Soznanie i zivilisatija*. Moscow, 2004.

Mamardashvili, Merab. *Strela posnanija*. Moscow, 2004.

Mamardashvili, Merab. *Variationi Kantiane*. Torino, 2003.

Motroshilova, Nelly. *Raboty rasnych let. Izbrannye statji i esse*. Moscow, 2005.

Nelly Motroshilova (2005)
Translated by Boris Jakim

MANDEVILLE, BERNARD

(c. 1670–1733)

Bernard Mandeville, a physician and moralist, was probably born in Rotterdam, Holland, where he was baptized on November 20, 1670. His family was a distinguished one, his father, grandfather, and great-grandfather having been noted physicians. The family name was originally de Mandeville, but Mandeville dropped the "de" in later life. He was educated at the Erasmian School in Rotterdam and then attended the University of Leiden, where he studied philosophy and medicine. He was granted the degree of doctor of medicine in 1691. His medical specialty was the treatment of nerve and stomach disorders, or, as he called them, the "hypochondriack and hysterick passions." Dr. Johnson is said to have had a high regard for a treatise Mandeville wrote on these diseases.

A short time after taking his degree Mandeville visited London to learn English, and liking the country and the people, he chose to settle in England. Little is known about his English life beyond the bare facts that he married, that he had a son and a daughter, that he practiced medicine, and that he apparently had plenty of time for writing. His success as a writer is all the more remarkable when one remembers that English was his adopted language. His best-known work is *The Fable of the Bees,* with its slogan "private vices, public benefits." It called forth a number of replies from the outraged defenders of virtue, including George Berkeley in the *Alciphron* and Francis Hutcheson. The book was a regular source of public and private controversy in the eighteenth century. The notoriety that this work gained Mandeville doubtless explains why no very consistent account of his situation and character has come down to us from his contemporaries. But Benjamin Franklin, who once met Mandeville, reported

that he was "a most facetious and entertaining companion." Mandeville died at Hackney in England.

The Fable of the Bees was twenty-four years in the making. It began as a poem of 433 lines called "The Grumbling Hive: Or, Knaves Turn'd Honest" (London, 1705). The many bitter attacks on the poem caused Mandeville to produce several expositions, elaborations, and defenses of it, all of which grew, over the years, into the book *The Fable of the Bees; Or Private Vices, Public Benefits.* In its final form, the sixth edition (1729), the *Fable* consists of two parts. Part I is the original poem followed by several essays: (1)"An Enquiry into the Origin of Moral Virtue," consisting of twenty-two remarks on various lines or words in the poem, such as luxury, pride, and so on; (2) "An Essay on Charity and Charity Schools"; (3) "A Search into the Nature of Society"; and (4) "A Vindication of the Book" against a presentment of the grand jury of Middlesex and other abuse. Part II, which is as long as the first part, consists of six dialogues in which Cleomenes instructs Horatio in the true meaning of the *Fable.*

As might be expected in a book that was put together over a long period and whose later parts are a defense of the earlier, Mandeville's targets are several, and assessing the relative importance of his ideas is not easy. His economic doctrines are certainly more thoroughly worked out than his moral theories, and he wanted politicians to take his economic views seriously. Given that a politician desires the nation he governs to be great and wealthy and given that there is a large population to be kept in employment, then a certain kind of economic life must be permitted and even fostered. The production of necessities will neither employ very many people nor by itself make a nation great. Therefore, the production of luxuries must be permitted, and their consumption on the most lavish scale possible encouraged, thus simultaneously achieving splendor and full employment. Mandeville analyzes the making of hooped and quilted petticoats in order to show not only the opportunities for labor the manufacture of this luxury provides in itself, but also the subsidiary employments (shipwright, sailor, dye-finder, and so on) that fashion calls into being.

In "An Essay on Charity and Charity Schools" Mandeville gives some hint of the structure of the society that is required to produce a great and wealthy nation. In this essay, he opposes educating the poor on the grounds that knowledge enlarges and multiplies our desires and that the fewer things a person wishes for, the more easily may his necessities be supplied. As Mandeville understood the English economic system of the eighteenth century, it required a large number of laboring poor, and he feared that education would make them dissatisfied with their lot and would consequently disrupt the system.

But Mandeville goes on to show the mixed feelings that have always troubled the analytical observer of society who is also a decent human being. He tells us that he does not wish to be thought personally cruel, but he believes that proposing to educate the poor is "to be Compassionate to excess, where Reason forbids it, and the general Interest of the Society requires steadiness of Thought and Resolution." It is, he argues, no harder on the poor to withhold education from them, even though they may have "natural parts and genius" equaling the rich, than it is to withhold money from them as long as they have the same inclinations to spend as the rich have.

Mandeville strongly favored free trade, seeing clearly that in order for one nation to buy another's goods, it must be able to sell its own. Any restriction in international trade must cause the loss of markets, with a consequent fall in the level of employment at home. In the eighteenth century Mandeville's writings became the chief source of arguments in favor of the manufacture of luxuries and against restrictions on trade, either within a given nation or between nations. Adam Smith owed much to his knowledge of *The Fable of the Bees.*

Mandeville did not choose, however, to publish these economic doctrines in a straightforward way. Instead, he offered them in his moralizing poem, "The Grumbling Hive." The bees in the poem have many vices, but their society thrives. Mandeville's notion of vice is a threefold one. First, he has in mind such character traits as envy, vanity, love of luxury, and fickleness in diet, furniture, and dress. These traits make buyers eager to spend lavishly and consume prodigiously, so that they will soon be ready to spend again. Second, Mandeville calls vice that behavior necessary to profitable trade. The seller must conceal from the prospective buyer both the original cost of his goods and the lowest price at which he is willing to sell, while the buyer must conceal the highest price at which he will buy. Mandeville believes that success will certainly require deceit on the part of both buyer and seller, not to mention sharper practices that may descend to downright fraud. Third, Mandeville counts crime as a vice that provides public benefits. Thieves are valuable on two counts. The threat of them keeps locksmiths in business, and when they do succeed, they soon squander their gains, thus contributing to the circulation of wealth. Mandeville may therefore conclude, "The worst of all the Multitude/Did something for the Common Good." In this vein he regards even wars and natural disasters as

valuable to the economic system, for by destroying goods, they provide an opportunity for labor to replace them.

Against his claims for the social utility of vice Mandeville sets the following picture of virtue:

> It is certain that the fewer Desires a Man has and the less he covets, the more easy he is to himself … the more he loves Peace and Concord, the more Charity he has for his Neighbor, and the more he shines in real virtue, there is no doubt but that in proportion he is acceptable to God and Man. But let us be Just, what Benefit can these things be of, or what earthly good can they do, to promote the Wealth, the Glory and Worldly Greatness of Nations?

By a divine fiat the bees of the grumbling hive are all made honest, and their society declines into simplicity and insignificance.

Why did Mandeville present his economic doctrines in a poem praising vice, a poem that could only outrage his contemporaries? The most likely supposition is that in the first writing the motives of the moralist are uppermost. If English economic life is seen as it is and as it will be, then encouraging men to be honest and frugal is a disservice to both them and the continuation of the economic system. By praising those sorts of behavior that are ordinarily called vicious, Mandeville hoped to shock the moralist into seeing the world as it is. He gives the moralist the choice either of accepting the world as it is and changing his tune or of rejecting the world and admitting that the virtues the moralist praises require a context quite different from what is ordinarily supposed. What Mandeville takes to be economic truths thus become the basis for a program that is no less than the reform of moralizing.

As *The Fable of the Bees* grew, Mandeville came to offer bits of moral theory, largely because of his discovery of the writings of the Earl of Shaftesbury. He attacked Shaftesbury bitterly. He calls the claim that men may be virtuous without self-denial "a vast Inlet to Hypocrisy." He says that Shaftesbury's search for "a real worth and excellence" in things "is not much better than a Wild-Goose-Chace that is but little to be depended on." Mandeville's own view is that "our Liking or Disliking of things chiefly depends on Mode and Custom, and the Precept and Example of our Betters and such whom one way or other we think to be Superior to us. In Morals there is no greater certainty."

The organization of men into a society arises from the multiplicity of each man's desires and the need to overcome the great man's desires and the need to overcome the great natural obstacles that stand in the way of satisfying these desires. In society each man achieves his own ends by laboring for others. Under a government each member of society is rendered subservient to the whole, and all men, by cunning management, are made to act as one. The key to social organization is man's pride and his consequent delight in flattery. Thus, governors may flatter men into putting public interest before private interest, and men are led to be pleased with themselves for being virtuous. Indeed, this satisfaction is the reward for virtuous actions, and it is ultimately this feeling that makes virtue possible.

These doctrines place Mandeville in the moral-sense school, but his presentation of them is desultory and unsystematic. A successor, such as David Hume, would have been interested to find these views in the *Fable.* But there is something else in Mandeville's writings that is even more impressive—the large number of vignettes, anecdotes, and sketches that make the reader feel he is learning what people are really like and that must in the end make him a shrewder observer of human nature.

See also Berkeley, George; Franklin, Benjamin; Hutcheson, Francis; Johnson, Samuel; Moral Sense; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Smith, Adam; Virtue and Vice.

Bibliography

The premier modern edition of *The Fable of the Bees* is that prepared by F. B. Kaye, 2 vols. (London, 1924). Kaye's researches have provided us with a balanced account of Mandeville's life, and his introductory essay on Mandeville's thought and influence should be consulted.

Mandeville's other works include *The Virgin Unmask'd: Or, Female Dialogues betwixt an Elderly Maiden Lady, and Her Niece* (London: J. Morphew, 1709); *Free Thoughts on Religion, the Church, and National Happiness* (London: T. Jauncy and J. Roberts, 1720); *A Modest Defence of Publick Stews* (1724); *An Enquiry into the Causes of the Frequent Executions at Tyburn* (London: J. Roberts, 1725); *An Enquiry into the Origin of Honour, and the Usefulness of Christianity in War* (London: J. Brotherton, 1732); and *A Letter to Dion [Berkeley], Occasion'd by His Book Call'd Alciphron* (1732).

Elmer Sprague (1967)

MANI AND MANICHAEISM

Mani, "the apostle of God," founder of one of the most widely influential religions of the ancient world, was born in southern Babylonia about 216 CE. Little is definitely

known of his birthplace and parentage, since some statements should probably be discounted as malicious reports from his adversaries. He seems to have been of Persian descent and related, at least on his mother's side, to the royal house of Parthia, which was overthrown in 226 by the Sassanid Ardashir I. He is said to have received his first revelation at the age of twelve, but he did not receive his formal call to apostleship until he was twenty-four. His public activity began with a journey to India, where he founded his first community.

Upon the death of Ardashir in 241, Mani returned to Parthia, where he was welcomed by Ardashir's successor Shapur, for whom he wrote a book, the *Shapurakan.* When Shapur died thirty years later, Mani also enjoyed the favor of his successor, but when Bahram came to the throne in 272 the situation changed. Throughout Mani's career the Magian priests had been his most deadly enemies, and they now secured his impeachment and condemnation. He was executed about 276 CE, and his death apparently was followed by persecution of his adherents.

At least seven works have been ascribed to him, including the *Shapurakan,* another work titled "The Living Gospel," and the *Epistula Fundamenti,* which, on the evidence of Augustine, was used by north African Manichaeans as a handbook of doctrine. To these some Western authorities add the *Kephalaia,* which is extant in Coptic. Resources for the study of Manichaeism—once limited to the information supplied by such opponents as Augustine and Titus of Bostra and to excerpts in the works of Theodore bar Konai, in Hegemonius's *Acta Archelai,* and in such Arabic sources as the *Fihrist* of En-Nadim—had in the twentieth century been enriched by discoveries of original Manichaean documents in Turkestan and Egypt. The fragments discovered at Turfan include texts in several Iranian dialects, Turkish, and Chinese, while the Egyptian discovery includes Coptic versions of the *Kephalaia,* a psalmbook, and a collection of homilies.

THE SYSTEM OF MANI

The chief characteristic of Mani's system is a consistent dualism that rejects any possibility of tracing the origins of good and evil to one and the same source. Evil stands as a completely independent principle against Good, and redemption from the power of Evil is to be achieved by recognizing this dualism and following the appropriate rules of life. The opposition of God and Matter is seen in the realm of nature as the conflict of Light and Darkness, Truth and Error. The present world, and man in particular, presents a mixture of Good and Evil, the result of a breach of the original limits by the powers of evil. The whole purpose of the founding of the universe was to separate the two principles and restore the original state of affairs, rendering Evil forever harmless and preventing any future repetition of the intermingling.

It is the special task of the Manichaean, the man who has been brought to the light, to collaborate in this separation. Through the God-sent mind that is in him and that sets him apart from the other creatures, he must become aware of the mixture present in all things. He must thus discover the true meaning and significance of the world and conduct himself accordingly, in such a way as to avoid any further contamination of the light and promote its release from its mixture with the darkness. The death of the body is thus redemption; and true life is the release of the soul, which is light, from its imprisonment in the body and its return to its true abode.

The Manichaean myth begins with the two primal principles of Light and Darkness, each dwelling in its own realm, coeternal but independent. Perception of the Light excites envy, greed, and hate in Darkness, and provokes it to attack the Light. In response the Father of Greatness calls forth the Primal Man, who arms himself with five powers and descends to battle with the Darkness. He is defeated, however, and the five powers of Darkness devour a part of his light and thus bring the mixture into being. In some versions this is explained as part of a deliberate plan to satisfy the powers of Darkness temporarily by the cession of a portion of the light and thus to prevent further attack. The captive portion of light, the armor of the Primal Man, is identified with the soul, which thus becomes subject to the affections of Matter.

The Primal Man appeals to the Father of Greatness, who sends the Living Spirit to deliver him. The archons, or powers of Darkness, are now overcome (although they do not lose their power of action), and heaven and earth are made from their carcasses. From the purest part of the Light in the archons the sun and moon are formed, but even so only a small part of the Light has been delivered. A fresh appeal from the powers of Light leads the Father of Greatness to send a Third Messenger, whose appearance inspires the Darkness to produce Adam and Eve in the image of his glorious form and to enclose in them the Light still at its disposal. The creation of Eve has a special purpose, in that she is more subservient to the demons and serves as their instrument for the seduction of Adam. Procreation serves the ends of Darkness, since each birth means a further dispersal of the Light, another subject for the realm of Darkness, and a prolonging of the captivity of the Light. The powers of Light accordingly send Jesus

on a mission of revelation to Adam, who is still innocent but subsequently disobeys, is seduced by Eve, and so sets the chain of reproduction in motion. This protracts the drama of salvation, and with it the mission of Jesus, into the history of humankind. In one age the revelation comes to India through the Buddha, in another to Persia through Zoroaster, in a third to the West through the historical Jesus, and in the last age it comes through Mani himself, the apostle of the true God.

MANICHAEAN ETHICS

The cosmogonic myth provides the basis and substructure for the Manichaean ethics and hope of redemption. The ethics are rigorously ascetic: Since procreation only prolongs the reign of the powers of darkness, marriage must be rejected. The Manichaean must abstain from all "ensouled" things and eat only vegetables, so as to avoid, as far as possible, any injury to the Light. The full rigor of Manichaean ethics is reserved for the Elect, and the mass of adherents, the Hearers or Soldiers, are allowed to live under less rigorous rules. Correspondingly there is a difference in their destiny after death: The Elect pass at once to the Paradise of Light, but the Soldiers must return to the world and its terrors until their light is freed and they attain to the assembly of the Elect. The third class of men, the sinners who are outside the Manichaean religion, are doomed to remain in the power of Evil.

MANICHAEAN GNOSTICISM

It is clear that Manichaeism may be regarded as a form of Gnosticism. Indeed, it has been called "the most monumental single embodiment of the gnostic religious principle, for whose doctrinal and mythological representation the elements of older religions were consciously employed" (Hans Jonas, *The Gnostic Religion,* pp. 207f.). It differs, however, from such older forms of Gnosticism as Valentinianism in that here the dualism is from the beginning an integral part of the myth, and not the result of a development in the myth. In Jonas's words, "the tragedy of the deity is forced upon it from outside, with Darkness having the first initiative," whereas in the other type of Gnosticism, Darkness is the product of the divine passion, not its cause. Any attempt to identify the sources upon which Mani drew for the construction of his system is, however, fraught with difficulty, and it would be dangerous to try to establish any genetic relationship. For example, attempts have been made, on the basis of the statement that his father belonged to a Baptist sect, the *Mugtasila,* to forge a link with Mandaeism; but although Mandaean elements have been found in the Manichaean psalmbook, the identity of the *Mugtasila* with the Mandaeans, or of either with some still older Jewish or Jewish-Christian Baptist movement, is still a matter of debate.

Another possible link is with the Zervanite heresy in Zoroastrianism, but here again caution is necessary. (On this whole subject, see Carsten Colpe, *Die Religion in Geschichte und Gegenwart,* Sec. 5.) In a general way, it may be said that Mani incorporated Christian, Buddhist, and Zoroastrian elements into his religion, but Manichaeism seems to have adapted itself to the dominant religion of a particular area. Moreover, it has been held that he had little more than a hearsay knowledge of Christianity, although he had some acquaintance with the heresies of Bardesanes and of Marcion. It appears that he intended to found not merely a sect but a new religion that could embody the best of the older faiths, fusing elements from Buddhism, Christianity, and Zoroastrianism with his own teaching.

His success is evident from the fact that Manichaeism survived so long and for a time was a serious rival to Christianity. After Mani's death it spread through Syria into the West and spread eastward deep into central Asia. Centuries later Manichaean ideas were current among the Bogomiles in the Balkans (see Dmitri Obolensky, *The Bogomils*) and among the Albigenses and Cathari in Provence (see Steven Runciman, *The Mediaeval Manichee*). There may be debate as to the historical connection of these later movements with the original Manichaeism, but some influence appears beyond dispute. Nor should it be forgotten that Augustine himself was for a time an adherent of Manichaeism. A religion that could arouse the interest of such later thinkers as Pierre Bayle, David Hume, and Voltaire must be regarded as one of profound significance for the history of thought.

See also Augustine, St.; Bayle, Pierre; Buddhism; Christianity; Evil; Evil, The Problem of; Gnosticism; Hume, David; Valentinus and Valentinianism; Voltaire, François-Marie Arouet de; Zoroastrianism.

Bibliography

TEXTS AND TRANSLATIONS

Allberry, C. R. C. *A Coptic Manichaean Psalm-Book.* Stuttgart: Kohlhammer, 1938.

Asmussen, Jes. *X^{u}āstvānīft.* Copenhagen: Prostant apud Munksgaard, 1965.

Boyce, Mary. *The Manichaean Hymn Cycles in Parthian.* London: Oxford University Press, 1954.

Clark, Larry. "The Manichaean Turkic Pothi-Book." *Altorientalische Forschungen* 9 (1982): 145–218.

Gardner, Iain. *The Kephalaia of the Teacher.* Leiden: Brill, 1995.
Gardner, Iain, and Samuel N. C. Lieu. *Manichaean Texts from the Roman Empire.* Cambridge, U.K.: Cambridge University Press, 2004.
Klimkeit, Hans-Joachim. *Gnosis on the Silk Road.* San Francisco: Harper, 1993.
Koenen, Ludwig, and Cornelia Römer. *Der Kölner Mani-Kodex.* Opladen: Westdeutscher, 1988.
Polotsky, Hans Joachim. *Manichäische Homilien.* Stuttgart: Kohlhammer, 1934.
Schmidt-Glintzer, Helwig. *Chinesische Manichaica.* Wiesbaden: Otto Harrassowitz. 1987.
Sundermann, Werner. *Der Sermon vom Licht-Nous.* Berlin: Akademie-Verlag, 1992.
Sundermann, Werner. *Der Sermon von der Seele.* Turnhout, Belgium: Brepols, 1997.

STUDIES

BeDuhn, Jason David. *The Manichaean Body in Discipline and Ritual.* Baltimore: Johns Hopkins Press, 2000.
Bryder, Peter. *The Chinese Transformation of Manichaeism.* Löberöd: Plus Ultra, 1985.
Gulácsi, Zsuzsanna. *Manichaean Art in Berlin Collections.* Turnhout, Belgium: Brepols, 2001.
Jonas, Hans. *The Gnostic Religion.* 2nd ed. Boston: Beacon Press, 1963.
Lieu, Samuel N. C. *Manichaeism in the Later Roman Empire and Medieval China,* 2nd ed. Tübingen: Mohr-Siebeck, 1992.
Mikkelsen, Gunner. *Bibliographia Manichaica.* Turnhout, Belgium: Brepols, 1997.
Obolensky, Dmitri. *The Bogomils.* Cambridge, U.K.: Cambridge University Press, 1948.
Pedersen, Nils Arne. *Studies in the Sermon on the Great War.* Aarhus: Aarhus University Press, 1996.
Reeves, John. *Heralds of that Good Realm.* Leiden: Brill, 1996.
Runciman, Steven. *The Mediaeval Manichee.* Cambridge: Cambridge University Press, 1947.
Widengren, Geo. *Der Manichäismus.* Darmstadt: Wissenschaftliche Buchgesellschaft, 1977.
Wurst, Gregor. *Das Bemafest der ägyptischen Manichäer.* Altenberge: Oros, 1995.

R. McL. Wilson (1967)
Bibliography updated by Jason BeDuhn (2005)

MANNHEIM, KARL
(1893–1947)

Karl Mannheim, the German sociologist, was born in Budapest and died in London. He studied at Berlin and Paris, and at Heidelberg under Max Weber, and later taught at Heidelberg, Frankfurt am Main, and, after 1933, in London.

Mannheim's thought resembles that of such philosophers as Auguste Comte and G. W. F. Hegel, who believed that in the past man had been dominated by the historical process whereas in the future he would gain ascendancy over it. Mannheim was deeply influenced by Karl Marx, but he deviated from Marxism in asserting that a better society might be achieved by nonrevolutionary means and also in de-emphasizing the interpretation of the development of society as being semiautomatic and stressing the importance of conscious political effort. He was, in addition, decisively influenced by German historicism and Anglo-Saxon pragmatism. From the former he took the belief that history is the *ens realissimum*, while from the latter he derived his criterion of truth. Both positions pointed toward a radical relativism, which, however, he strove to overcome.

In his first and most important book, *Ideologie und Utopie*, Mannheim asserted that the act of cognition must not be regarded as the effort of a purely theoretical consciousness, because the human consciousness is permeated by nontheoretical elements arising both from man's participation in social life and in the streams and tendencies of willing which work themselves out contemporaneously in that life. The influence of these active factors is all-important; even the categorial structure of the intellect does not escape it. Mannheim therefore maintained that epistemology (as practiced, for instance, by Immanuel Kant) was outdated, and must be superseded by a new discipline, the sociology of knowledge.

According to Mannheim, this new discipline revealed that all knowledge (at any rate, knowledge of things human) was situation-bound (*situationsgebunden*)—that is, tied to a given constellation of sociohistorical circumstances. Each age develops its own style of thought, and comparisons between these styles are impossible, since each posits a different basic (or, so to speak, relatively absolute) sphere. Even within each age there are conflicting tendencies toward conservation, on the one hand, and toward change on the other. Commitment to conservation tends to produce "ideologies"—to falsify thought by excessive idealization of the past and overemphasis on the factors making for stability. Intentness on change is apt to produce "utopias," which overvalue both the future and factors leading to change.

Between ideology and utopia there is at least the possibility of completely realistic (*situationsgerecht*) thought that functions without friction within the given framework of life, and is set neither on pushing forward nor on holding back the development of society. But Mannheim places little emphasis on this possibility. He sees a very strong tendency toward the polarization of society into hostile camps. Only the comparatively uncommitted intelligentsia is likely to approach nearer the truth. From its special and particularly favorable vantage point, it

could, and should, elaborate a "total perspective" that would synthesize the conflicting contemporary world views and thereby neutralize, and to some extent overcome, their one-sidedness. Such a "dynamic synthesis" is the nearest possible approximation to a truly realistic attitude, within the limitations imposed upon a given epoch.

This estimate of human thought might seem to justify accusing Mannheim of skepticism, but Mannheim held himself innocent of the charge. To rebut it, he developed his doctrine of "relationism," which he opposed to skeptical relativism. Relationism, he argued, does not impugn the validity of an insight: It merely draws attention to the fact that the insight is dependent upon, and confined within, a specific sociohistorical situations. But this argument merely shifts the relativity, and does not remove it. Mannheim held that every sociohistorical situation is located at a specific point along a unilinear, ever-progressing and never-returning temporal continuum—history. Each situation is therefore unique, and the knowledge to which it gives birth, and which is true within it, is equally unique, bound to its time and place, and relative.

But Mannheim was not primarily concerned with the truth of propositions. Rather, he operated with a radically different conception of "truth." To him, truth is an attribute, not so much of discourse, as of reality. The individual who is in contact with the living forces of his age has the truth, or better, is in the truth—a conception that shows at once Mannheim's Marxism, his historicism, and his pragmatism. He was moving close to the belief that the traditional *adaequatio rei et intellectus* (correspondence of thought and reality) should be replaced by a new test, the *adaequatio intellectus et situs* (correspondence of thought and situation). He was interested in the genuineness, rather than in the truth (properly so called), of a given world view.

Mannheim was a confirmed progressivist, and he tended to prefer whatever was, at any time, emergent. After his immigration to England in 1933, he adopted a more practical and political orientation. He argued dialectically, especially in *Mensch und Gesellschaft im Zeitalter des Umbaus* (1935), that a completely unregulated society, such as he thought liberalism had created, was apt to produce its own opposite, totalitarian dictatorship. To secure the values of democracy, it was necessary to avoid the weaknesses of both liberalism and totalitarianism. As a viable synthesis, Mannheim advocated "planning for freedom," a social system that would ensure economic stability by regulating the more objective aspects of life, such as production, but at the same time grant freedom to men's subjective strivings (for example, in matters of taste), thereby releasing cultural creativity. In this context, Mannheim became interested in education as the prime means of radical democratization. Toward the end of his career, he began to feel that a modernized Christianity held out some hope for a new integration of society's value system, which had become splintered and self-contradictory.

See also Christianity; Democracy; Hegel, Georg Wilhelm Friedrich; Historicism; Ideology; Kant, Immanuel; Marxist Philosophy; Marx, Karl; Pragmatism; Sociology of Knowledge.

Bibliography

WORKS BY MANNHEIM

Ideologie und Utopie. Bonn: Cohen, 1929. Translated by Louis Wirth and Edward Shils as *Ideology and Utopia*. London: Routledge, 1936. The English edition also includes the article "Wissenssoziologie" from *Handwörterbuch der Soziologie*, edited by Alfred Vierkandt. Stuttgart: Enke, 1931.

Mensch und Gesellschaft im Zeitalter des Umbaus. Leiden: Sijthoff, 1935. Translated by Edward Shils as *Man and Society in an Age of Reconstruction*. London: Routledge, 1940.

Diagnosis of Our Time. London: Routledge, 1943.

Freedom, Power and Democratic Planning, edited by Hans Gerth and Ernest K. Bramstedt. London: Routledge, 1951.

Essays on the Sociology of Knowledge, edited by Paul Kecskemeti. London: Routledge, 1952.

Essays on Sociology and Social Psychology, edited by Paul Kecskemeti. London: Routledge, 1953.

Essays on the Sociology of Culture, edited by Ernst Mannheim and Paul Kecskemeti. London: Routledge, 1956.

Systematic Sociology, edited by J. S. Erös and W. A. C. Stewart. London: Routledge and Kegan Paul, 1957.

An Introduction to the Sociology of Education. Written with W. A. C. Stewart. London: Routledge, 1962.

Freedom, Power, and Democratic Planning. London: Routledge & K. Paul, 1968.

From Karl Mannheim, edited by Kurt H. Wolff. New York: Oxford University Press, 1971.

Die strukturanalyse der erkenntnistheorie. Vaduz/Liechtenstein: Topos Verlag, 1978.

Strukturen des Denkens, edited by David Kettler, Volker Meja, and Nico Stehr. Frankfurt am Main: Suhrkamp, 1980.

Structures of Tinking: Text and Translation, edited by David Kettler, Volker Meja, and Nico Stehr. Translated by Jeremy J. Shapiro and Shierry Weber Nicholsen. Boston, MA: Routledge & Kegan Paul, 1982.

Conservatism: A Contribution to the Sociology of Knowledge, edited by David Kettler, Volker Meja, and Nico Stehr. London: Routledge & Kegan Paul, 1986.

Collected Works of Karl Mannheim. London; New York: Routledge, 1997.

Selected Correspondence (1911–1946) of Karl Mannheim, Scientist, Philosopher, and Sociologist, edited by Éva Gábor. Lewiston, NY: Edwin Mellen, 2003.

WORKS ON MANNHEIM

Kettler, David, and Volker Meja. *Karl Mannheim and the Crisis of Liberalism: The Secret of these New Times*. New Brunswick, NJ: Transaction, 1995.

Lieber, Hans-Joachim. *Wissen und Gesellschaft*. Tübingen: Niemeyer, 1952.

Loader, Colin. *The Intellectual Development of Karl Mannheim: Culture, Politics, and Planning*. Cambridge, U.K.; New York: Cambridge University Press, 1985.

Longhurst, Brian. *Karl Mannheim and the Contemporary Sociology of Knowledge*. New York: St. Martin's Press, 1989.

Maquet, Jacques Jerome. *Sociologie de la connaissance*. Louvain: Institut de Recherches Économiques et Sociales, 1949. Translated into English by John F. Locke as *The Sociology of Knowledge*. Boston: Beacon Press, 1951.

Remmling, Gunter W. *The Sociology of Karl Mannheim: With a Bibliographical Guide to the Sociology of Knowledge, Ideological Analysis, and Social Planning*. New York: Humanities Press, 1975.

Schoeck, Helmut. "Die Zeitlichkeit bei Karl Mannheim." *Archiv für Rechts- und Sozialphilosophie* 38 (1949–1950): 371–382.

Simonds, A. P. *Karl Mannheim's Sociology of Knowledge*. Oxford: Clarendon, 1978.

Woldring, H. E. S. *Karl Mannheim: The Development of His Thought: Philosophy, Sociology and Social Ethics, with a Detailed Biography*. Assen: Van Gorcum, 1986.

Werner Stark (1967)
Bibliography updated by Michael J. Farmer (2005)

MANSEL, HENRY LONGUEVILLE
(1820–1871)

Henry Longueville Mansel, an English philosopher and divine, was educated at Merchant Taylors' School, London, and St. John's College, Oxford. He became tutor in his college, the first Wayneflete professor of moral and metaphysical philosophy at Oxford University in 1859, Regius professor of ecclesiastical history there in 1866, and dean of St. Paul's in 1868.

Mansel was at Oxford during the period when, after more than a century of slumbers, it was again beginning to take philosophy seriously. But whereas his Oxford contemporaries, such as Benjamin Jowett and T. H. Green, looked to Germany for their philosophy, Mansel looked to France and Scotland.

Indebted to various thinkers, especially to William Hamilton and Victor Cousin, Mansel was remarkably successful in assimilating their influences. When—as on the question of the perception of an external world—he occupied common ground with Hamilton, Mansel's version was marked by a superior clarity and relevance. Likewise, he more than did justice to what was genuinely original and valuable in Cousin's critique of John Locke's doctrine of judgment, making it the foundation of a subtle and thorough discussion of the relation of thinking to experience begun in the *Prolegomena Logica* and completed in the article "Metaphysics, or the Philosophy of Consciousness."

The point at issue was the relation of meaning to verification. Can we know a proposition to be true or false without first understanding the meaning of the terms involved, in the sense of being able to define each of them separately? Mansel dealt with this difficulty by making a sharp distinction between a *logical judgment,* in which the understanding of the terms precedes the judgment as to the truth or falsity of the proposition, and a *psychological judgment,* in regard to which this sharp distinction cannot be drawn, and in regard to which the understanding of the terms coincides with the judgment as to the truth of the proposition.

Mansel's main point was that the former sort of judgment must always, in the last analysis, rest upon the latter, of which the Cartesian *cogito* is the prime example. In this way the kind of clear-cut empirical knowledge with which science deals rests on the foundation of an essentially vague metaphysical knowledge embodied in the *cogito.* This doctrine, which descended through Cousin from Thomas Reid, was worked out by Mansel in the course of an excellent discussion of the problem of universals and particulars, contained in the article "Metaphysics." What nominalistic atomists had forgotten was that the individual thing is initially given in an essentially vague experience (for example, three objects seen in the far distance and just recognizably human) that withholds the details and reveals only general characteristics.

While this topic of the relation of thinking to experience was central in Mansel's work, he was equally stimulating on other questions. Somewhat in the French style, he held that the will, in the form of attention, forms an integral part of cognition. Following a suggestion of Dugald Stewart's, he tried to illuminate the difference between the presence and the absence of efforts of will by an interesting phenomenology of daydreaming and semiconsciousness. Again influenced by Reid, Mansel was aware—as few were in his time—of the complexities and difficulties of the problem of our knowledge of the existence of other minds, discussing it, appropriately enough, in connection with the moral judgment. Finally, Mansel

dealt interestingly with the distinction between philosophy and science. Philosophy deals with what he called facts of consciousness, whose distinctive feature is that their *esse* is *percipi,* in the sense in which René Descartes had said that, so far as philosophy is concerned, there is no difference between seeing something and thinking one sees it.

The result of this careful phenomenological analysis (the word *phenomenology* had been introduced by Mansel's masters, Hamilton and Cousin) was that Mansel saw human experience as inherently complex and mysterious. In the background of Mansel's philosophy there was always an explicit contrast with a rival kind of reductive analysis that regarded man as being as unmysterious in his inner workings as a pocket watch. This contrast was the key to the controversies aroused by Mansel's Bampton lectures, "The Limits of Religious Thought," delivered in 1858. Mansel held that reason tells us that if evil exists, then God cannot be both perfectly good and all-powerful. However, God's omnipotence and perfect goodness must be accepted as a matter of faith. Although God is perfectly good, we cannot know the nature of his goodness. Man's finite goodness cannot explain God's infinite goodness; they are the same by analogy, not identity.

Mansel's lectures were attacked by F. D. Maurice and Goldwin Smith, and by John Stuart Mill, who devoted Chapter 7 of his *Examination of Sir William Hamilton's Philosophy* to Mansel's views. Mill wrote, "I will call no being good, who is not what I mean when I apply that epithet to my fellow creatures, and if such a being can sentence me to hell for not so calling him, to hell I will go." Mansel replied in *The Philosophy of the Conditioned,* and Mill in turn replied in numerous footnotes in later editions of the *Examination,* listing Mansel first among his critics. For Mansel man's goodness was not clear and God's goodness was inscrutable; both were equally a mystery.

Mansel's *Letters, Lectures, and Reviews,* published posthumously, contains, among other things, interesting articles on the philosophy of language and on mathematical logic.

See also Cousin, Victor; Descartes, René; Green, Thomas Hill; Hamilton, William; Locke, John; Logic, History of; Mill, John Stuart; Phenomenology; Language, Philosophy of; Reid, Thomas; Stewart, Dugald.

Bibliography

WORKS BY MANSEL

Prolegomena Logica. Oxford: W. Graham, 1851.

"Metaphysics, or the Philosophy of Consciousness." In *Encyclopaedia Britannica,* 8th ed. 1857. Published separately, Edinburgh, 1860.

The Limits of Religious Thought. Oxford and London: J. Wright for J. Murray, 1858.

The Philosophy of the Conditioned. London and New York: A. Strahan, 1866.

Letters, Lectures, and Reviews. Edited by H. W. Chandler. London: J. Murray, 1873.

WORKS ON MANSEL

Burgon, J. W. *Lives of Twelve Good Men.* London, 1888. Vol. II, pp. 149–237.

Mill, J. S. *An Examination of Sir William Hamilton's Philosophy.* London: Longmans, Green, 1865. Ch. 7.

Stephen, Leslie. "H. L. Mansel." In *Dictionary of National Biography.* London, 1893. Vol. 36, pp. 81–83.

George E. Davie (1967)

MANY-VALUED LOGICS

An orthodox assumption in logic is that (declarative) sentences have exactly one of two values, true (1) and false (0). Many-valued logics are logics where sentences may have more than two values. Aristotle (*De Interpretatione*, chapter 9) was perhaps the first logician to countenance the thought that some sentences (future contingents) may be neither true nor false; Aristotle's ideas were discussed by many logicians in the Middle Ages. However, contemporary work on many-valued logics commenced with the work of the Polish logician Jan Łukasiewicz early in the twentieth century. One hundred years later there are many well-known many-valued logics, and the properties of such logics are well established. The logics have important philosophical applications (e.g., in articulating the views that some sentences are neither true nor false, or both true and false, or that truth comes by degrees). They also have important technical applications (e.g., in establishing various independence results).

In what follows, p, q, … will be used for propositional parameters (variables); A, B, … for arbitrary sentences; and Σ, Δ, … for sets of sentences. For references, see the last section of this entry.

ŁUKASIEWICZ LOGICS

To illustrate the notion of a formal many-valued logic, consider classical propositional logic with the following

connectives: ∧ (conjunction), ∨ (disjunction), ¬ (negation), and → (conditional). This may be formulated as follows. The set of semantic values, *Val*, is {0, 1}. The set of designated values, *Des*, is {1}. An evaluation, *v*, assigns every propositional parameter (pp), a member of *Val*. All formulas are then assigned such values recursively by the clauses:

$$v(\neg A) = 1 - v(A)$$
$$v(A \wedge B) = Min(v(A), v(B))$$
$$v(A \vee B) = Max(v(A), v(B))$$
$$v(A \rightarrow B) = 1 \quad \text{if } v(A) \leq v(B)$$
$$= 1 - (v(A) - v(B)) \text{ otherwise}$$

(*Max*(*x, y*) is the maximum of *x* and *y*; *Min*(*x, y*) is the minimum of *x* and *y*. $v(A \rightarrow B)$ takes the maximum value minus any amount one has to drop to get from *A* to *B*.) The inference from Σ to *A* is valid ($\Sigma \vDash A$) just if there is no evaluation that makes all the premises designated but not the conclusion (i.e., there is no *v* such that for all $B \in \Sigma$, $v(B) \in Des$, but $v(A) \notin Des$).

If everything is exactly the same, except that *Val* = {0, ½, 1}, one has the three-valued Łukasiewicz logic $Ł_3$. The semantic conditions for the connectives can be depicted in the form of tables, thus:

→	1	1/2	0
1	1	1/2	0
1/2	1	1	1/2
0	1	1	1

	¬
1	0
1/2	1/2
0	1

∨	1	1/2	0
1	1	1	1
1/2	1	1/2	1/2
0	1	1/2	0

∧	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	0
0	0	0	0

More generally, if $n>1$ and everything is the same, except that $Val = \{i/(n-1) : 0 \leq i \leq n-1\}$, one has the Łukasiewicz *n*-valued logic $Ł_n$. Finally, if everything is the same, except that *Val* = [0, 1] (the set of all real numbers between 0 and 1, inclusive), one has the Łukasiewicz continuum-valued logic $Ł_{\aleph}$. (The relationship between these logics is that $Ł_n$ is a [proper] sublogic of $Ł_m$ if and only if [iff] *m* divides *n*; and $Ł_{\aleph}$ is a [proper] sublogic of all the $Ł_n$. The logic in which *Val* is the set of rationals between 0 and 1 turns out to be equivalent to $Ł_{\aleph}$.)

BOTH/NEITHER LOGICS

The values of a many-valued logic need not be numbers (and the designated values do not need to be a singleton). In another well-known family of logics, *Val* = {1, *b*, *n*, 0}. (1 can be thought of as *true and only true*; 0 as *false and only false*; *b* as *both true and false*; and *n* as *neither true nor false*.) *Des* = {1, *b*}. One can order these values as follows:

1
↗ ↖
b *n*
↖ ↗
0

If *v* is an evaluation of the pps into *Val*, it is extended to all formulas by the following conditions:

$$v(A \vee B) = Lub\ \{v(A), v(B)\}$$
$$v(A \wedge B) = Glb\ \{v(A), v(B)\}$$

(*Lub X* is the least element of the lattice greater than or equal to every member of *X*. *Glb X* is the greatest element of the lattice less than or equal to every member of *X*.) The conditions for negation can be represented as follows:

	¬
1	0
b	*b*
n	*n*
0	1

$A \rightarrow B$ can be defined as $\neg A \vee B$. Note that all these conditions agree with classical logic when the values are just 0 and 1.

These semantics give the logic often called First Degree Entailment (FDE). If one ignores the value *n*, one gets the three-valued logic LP. If one ignores the value *b*, one gets the strong Kleene three-valued logic, K_3. FDE and K_3 have no logical truths; LP (and $Ł_3$) does. LP and FDE are paraconsistent (i.e., the inference $A, \neg A \vdash B$ is not valid); K_3 is not. FDE is a sublogic of both logics, but neither is a sublogic of the other (and all three are sublogics

of classical logic). The weak Kleene three-valued logic, B_3, is the same as K_3, except that any truth function with an *n* as an input gives *n* as an output.

For the first-order versions of all the logics in this section and the last, the quantifiers ∀ and ∃ can be thought of as the infinitary generalizations of ∧ and ∨, in the usual way. Thus, if *Dom*, is the domain of quantification, and every $d \in Dom$, has a name, c_d, (and if not just add them):

$$v(\forall xA\,(x)) = Glb\{v(A(c_d)) : d \in Dom\}$$

$$v(\exists xA\,(x)) = Lub\{v(A\,(c_d)) : d \in Dom\}$$

where the bounds are with respect to the appropriate orderings.

GENERAL DEFINITION

In general terms, in a semantics for a formal many-valued propositional logic, there is an arbitrary set of semantic values, *Val*. (If the cardinality of *Val* is *n*, the logic is called *n*-valued; if it is finite, the logic is called finitely many-valued; if it is infinite, the logic is called infinitely many-valued.) *Des*, the set of designated values, is an arbitrary subset of *Val*. Each *n*-ary connective in the language, #, is assigned an *n*-place (total) function, $f_\#$, with inputs and outputs in *Val*. An evaluation of the language, *v*, assigns each pp a member of *Val*. Semantic values are assigned to all sentences recursively by the equations $v(\#(A_1, \dots, A_n)) = f_\#\,(v(A_1), \dots, v(A_n))$. An inference is valid if there is no evaluation that makes all the premises designated and the conclusion undesignated. (Slightly more general definitions are also possible here.)

For quantifiers, a domain of quantification, *Dom*, and denotation function, δ, are added. For every constant *c*, $\delta(c) \in Dom$; if *P* is an *n*-place predicate, $\delta(P)$ is a (total) *n*-place function with inputs and outputs in *Dom*. $v(Pc_1, \dots c_n) = \delta(P)(\delta(c_1),\dots,\delta(c_n))$. Each quantifier, *Q*, is assigned a (total) function, f_Q, with inputs that are subsets of *Val* and outputs in *Val*. Assuming that each object in the domain has a name: $v(QxA(x)) = f_Q(\{v(A(c_d)) : d \in Dom\})$.

It is not difficult to check that any many-valued logic is a Tarski consequence relation. That is, it satisfies the following properties. (Here, Σ,Δ means $\Sigma \cup \Delta$; and set braces for singletons are omitted.)

If $A \in \Sigma$, $\Sigma \vDash A$

If $\Sigma \vDash A$ and $\Sigma \subseteq \Delta$, then $\Delta \vDash A$

If $\Sigma \vDash A$ and $\Delta, A \vDash B$, then $\Sigma, \Delta \vDash B$.

If $\Sigma \vDash A$, then any uniform substitution is valid.

(A uniform substitution is obtained by replacing each occurrence of any pp with the same formula.)

In many cases, the set of values (*Val*), together with the operations on it (the $f_\#$s), is a special case of an algebra of a certain kind. In classical logic, these are Boolean algebras; in the case of FDE, these are De Morgan algebras; and in the case of $Ł_\aleph$, these are MV algebras. Another notion of validity can be obtained by appealing to all the algebras of a kind. At this point, many-valued logic slides into algebraic logic.

PROOF PROCEDURES

All finitely many-valued logics are decidable (and *a fortiori* axiomatizable, though not necessarily finitely axiomatizable). A uniform algorithm is a generalization of truth tables (often there are more efficient ones). Consider all the possible assignments of values to the relevant pps. In each case, compute the values of the premises and the conclusion, and see if there is any assignment in which all the premises are designated and the conclusion is not.

A simple axiom system for $Ł_3$ is as follows:

$A \to (B \to A)$

$(A \to B) \to ((B \to C) \to (A \to C))$

$(\neg A \to \neg B) \to (B \to A)$

$((A \to \neg A) \to A) \to A$

The only rule of inference is *modus ponens* ($A, A \to B \vdash B$); $A \vee B$ is defined as $(A \to B) \to B$; and $A \wedge B$ is defined as $\neg(\neg A \vee \neg B)$. In each $Ł_n$ a family of *J*-functions can be defined, where $v(J_iA) = 1$ if $v(A) = i$, and $v(J_iA) = 0$ otherwise (*i*, here, being any value of the logic). These can be exploited to give a uniform procedure for producing an axiom system for each $Ł_n$. Similar techniques work for other finitely many-valued logics in which analogues of the J-functions can be defined. (Much technical effort has gone into investigating which functions can be defined in various many-valued systems.) An axiom system for $Ł_\aleph$ is obtained by replacing the last axiom cited earlier with:

$$((A \to B) \to B) \to ((B \to A) \to A)$$

If the designated values are changed to $[r, 1]$ (closed at the left end) or $(r, 1]$ (open at the left end), for some rational number, *r*, the systems are also axiomatizable. If *r* is an irrational number, they may not be.

Appropriate tableau and natural deduction systems for many-valued logics can often be found. For example, here is a tableau system for FDE. Lines of the tableau are of the form A: + or A: – . (Intuitively, + means "is designated" and – means "is not designated".) To test the inference $A_1, \ldots, A_n \vdash B$, start with lines of the form A_1: + , …, A_n: + , B: – . The rules are as follows (± can be disambiguated uniformly either way):

$$\begin{array}{ccc} \alpha\wedge\beta:+ & \alpha\wedge\beta:- & \neg(\alpha\wedge\beta):\pm \\ \downarrow & \swarrow\ \searrow & \downarrow \\ \alpha:+ & \alpha:-\quad \beta:- & \neg\alpha\vee\neg\beta:\pm \\ \beta:+ & & \end{array}$$

$$\begin{array}{ccc} \alpha\vee\beta:- & \alpha\vee\beta:+ & \neg(\alpha\vee\beta):\pm \\ \downarrow & \swarrow\ \searrow & \downarrow \\ \alpha:- & \alpha:+\quad \beta:+ & \neg\alpha\wedge\neg\beta:\pm \\ \beta:- & & \end{array}$$

$$\begin{array}{c} \neg\neg\alpha:\pm \\ \downarrow \\ \alpha:\pm \end{array}$$

A branch closes if it contains lines of the form A: + and A: – . Adding closure whenever there are lines of the form A: + and $\neg A$: +, gives K_3. Adding closure whenever there are lines of the form A: – and $\neg A$: –, gives LP. (Adding both gives classical logic.) The first-order versions of all the finitely many-valued logics already mentioned also have sound and complete proof procedures. However, first-order $Ł_\aleph$ is not axiomatizable. By contrast, the logics that are the same as $Ł_\aleph$, except that for some rational number, $r<1$, $Des = (r, 1]$ (open at the left end) or $[r, 1]$ (closed at the left end) are axiomatizable.

MANY-VALUED AND OTHER LOGICS

A number of important logics, notably intuitionist logic, standard modal, and relevant logics, are demonstrably not finitely many-valued. Specifically, suppose that a logic validates the inferences $\vdash A \rightarrow A$ and $A \vdash A \vee B$. Then for any $a, b \in Val$, $f_\rightarrow(a, a) \in Des$, and if $a \in Des$, $f_\vee(a, b) \in Des$. Now suppose that the logic is n-valued, and that $p_0, \ldots, p_n$ are distinct pps. Let A be the disjunction of all formulas of the form $p_i \rightarrow p_j$ (for $0 \le i \ne j \le n$). Consider any evaluation. For some i and j, p_i and p_j must have the same value; hence, $p_i \rightarrow p_j$, and so A, are designated. Hence, A is a logical truth. The logics just cited can be shown to have no logical truths of this form (where → is the intuitionist, strict, and relevant conditional, respectively).

However, nearly all logics have an infinitely many-valued semantics of a rather unilluminating kind. Consider the set of logical truths of any logic closed under uniform substitution. Let Val be the set of formulas of the language; $Des = \{A : \vdash A\}$; $f_\#(A_1, \ldots, A_n) = \#(A_1, \ldots, A_n)$. Then $\vdash A$ iff $\vDash A$.

[Proof: Suppose that A is a logical truth. Consider any interpretation, v. It is easy to check that $v(A)$ is A with every pp, p, replaced by $v(p)$. Since the logic is closed under uniform substitution $v(A)$ is a logical truth; that is, it is designated. Conversely, suppose that A is not a logical truth. Consider the interpretation, v, which maps every pp to itself. It is easy to check that $v(A) = A$, which is not designated.]

The construction can be extended to show that any Tarski consequence relation with finite sets of premises has a many-valued semantics iff it satisfies one condition. This is called uniformity, and is, loosely speaking, to the effect that pps not involved in an inference are irrelevant to it. Specifically, if $\Gamma, \Delta \vDash A$, then $\Gamma \vDash A$, provided that:

1.) Δ is nontrivial (that is, for some B, $\Delta \nvDash B$)

2.) No formula in Δ contains a pp that occurs in a formula in $\Gamma \cup \{A\}$

It should be noted that not all logics are uniform. In Ingebrigt Johansson's minimal logic, $\emptyset \cup \{p, \neg p\} \vDash \neg q$, but $\{p, \neg p\}$ is nontrivial, and $\emptyset \nvDash \neg q$.

The finiteness constraint can be dropped if the notion of uniformity is strengthened in an appropriate fashion. (Some interesting differences between single-conclusion inference and multiple-conclusion inference emerge in this case.)

PHILOSOPHICAL APPLICATIONS

Many-valued logics have been claimed to have numerous philosophical applications. Like all interesting philosophical matters, these applications are debatable.

Łukasiewicz interpreted Aristotle's argument in *De Intepretatione* (chapter 9) as showing that, though true statements about the past and present are now necessarily true, contingent statements about the future (such as "There will be a sea battle tomorrow") currently have an indeterminate truth status. He suggested deploying $Ł_3$ in an analysis of this situation, reading the truth values {1, ½, 0} as necessarily true, indeterminate, and necessarily false, respectively. As one would expect $A \vee \neg A$ is not logically valid in $Ł_3$.

Łukasiewicz suggested adding an operator to the language, □, representing necessity, whose truth conditions may be represented as follows:

	□
1	1
1/2	0
0	0

Its dual, possibility, ◇, that is, ¬□¬, is as follows:

	◇
1	1
1/2	1
0	0

This makes the inference $A \vdash \Box A$ valid—which is reasonable enough on the Aristotelian picture. However, it also makes the inference $\Diamond A, \Diamond B \vdash \Diamond(A \wedge B)$ valid—which it is not, even for Aristotle. (Just let B be $\neg A$.) As has already been seen, normal modal logics are not finitely many-valued.

Future contingents are just one example of sentences that have been suggested as being neither true nor false (truth value gaps). Others include: sentences with reference failure ("The king of France is bald," "3 = 1/0"), category mistakes and other "nonsense" ("This stone is thinking of Vienna"), paradoxical sentences of self-reference ("This sentence is false"), sentences attributing a vague property in a borderline case ("This is a child"—said of someone around puberty), and sentences unverifiable by the appropriate mathematical or scientific procedure ("There are ten consecutive '7's in the decimal expansion of π", "This electron has a velocity of exactly 100 m/sec").

It is often claimed that K_3 (or, sometimes, B_3) is the appropriate logic for such cases: Gappy sentences take the value n. (In the last case, quantum logic and intuitionist logic have also been suggested to handle the matter.) In these logics $A \vee \neg A$ is not a logical truth, but neither is anything else. In particular, then, $A \wedge \neg A$ is not a logical falsity. Even if "The king of France is bald" is neither true nor false, "The king of France is bald and not bald" would seem to be logically false.

One way around this problem is to deploy the method of supervaluations. If v is any K_3 evaluation, let μ be a supervaluation of v ($v \preceq \mu$) iff:

$$\mu(p) \text{ is never } n \text{, and if } v(p) \neq n,\ v(p) = \mu(p)$$

An important feature of this logic, not shared by $Ł_3$, is that if $v(A)$ is 1 or 0, and $v \preceq \mu$, then $\mu(A)$ has the same value.

Now define the supertruth-value, v_s of a sentence under v as follows:

$$\begin{aligned} v_s(A) &= 1 \text{ if for all } \mu \text{ such that } v \preceq \mu,\ \mu(A) = 1 \\ &= 0 \text{ if for all } \mu \text{ such that } v \preceq \mu,\ \mu(A) = 0 \\ &= n \text{ otherwise} \end{aligned}$$

Define an inference as supervaluation valid if it preserves supertruth-value 1. The inferences that are supervaluation-valid now turn out to be exactly those that are classically valid.

[Proof: If an inference is not classically valid, let v be an evaluation that makes the premises true and the conclusion false. But v is a K_3 evaluation and $v \preceq v$. Hence the inference is not supervaluation-valid. Conversely, suppose that an inference is not supervaluation valid. Then there is a K_3 valuation, v, such that every supervaluation of v gives all the premises the value 1, but not the conclusion. Hence, there is some supervaluation that gives all the premises value 1, but the conclusion value 0. This is a classical evaluation. Hence, the argument is classically invalid.]

On the other side of the street, it has been suggested that some sentences are both true and false (truth-value gluts). These include: paradoxical sentences of self-reference ("This sentence is false"), statements describing instantaneous transition states ("He is in the room"—said at the instant he is symmetrically poised between being in and out), statements of rights and obligations ("She is legally required to do such and such"—when the requirements are based on inconsistent legislation), and sentences attributing a vague property in a borderline case ("This is a child"—said of someone around puberty).

It is sometimes suggested that LP—or FDE if one wants to also take in the possibility of truth value gaps—is the appropriate logic for such cases. The glutty sentences take the value b. (Other paraconsistent logics have also been suggested for the job.) In these logics $A \wedge \neg A$ may take a designated value. In LP the negation of this is also a logical truth.

A way to regain classical logic with LP is by the use of subvaluations. Subvaluations and subvaluation validity are defined in the way dual to supervaluation (*b* replacing *n*, and *some* replacing *all*). In the case of subvaluations, one has the equivalence between classical validity and subvaluation validity only in the one-premise case. (But the duality between the two cases is exact. In a classical multiple-conclusion logic $A \vee B \vdash A, B$ is valid. It is not supervaluation-valid. The equivalence between classical and supervaluational validity holds only because in a single-conclusion inference, one is, in effect, disjoining all the conclusions. In the subvaluation case, this corresponds to conjoining all the premises, which reduces matters to the single premise case.) The technique of super/subvaluations can be generalized to FDE, where there are both gaps and gluts.

A weakness of both LP and FDE is that they do not have a detachable conditional, since $A, A \rightarrow B \nvDash B$. They can be augmented with such a conditional, though. Thus, the many-valued logic RM_3 augments LP with a detachable conditional, $\Rightarrow$, whose truth conditions can be represented as follows:

$\Rightarrow$	1	*b*	0
1	1	0	0
b	1	*b*	0
0	1	1	1

In the context of information processing, truth value gaps are often interpreted as incomplete information, and truth-value gluts as inconsistent information. While in the context of gaps and gluts, a word should be said about set theory. It is well known that the naive comprehension schema

$$x \in \{y : A(y)\} \leftrightarrow A(x)$$

leads to contradiction (and so triviality)—in the shape of paradoxes such as Russell's—when the underlying logic is classical. It has often been suggested that the principle might be consistent (or at least inconsistent but nontrivial) when the underlying logic is many-valued. Problems for such suggestions arise because the principle generates triviality if the logic contains contraction $((A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B))$ and *modus ponens*. Let $A(y)$ be $y \in y \rightarrow B$. Call the set that this defines *c*. Comprehension quickly gives: $c \in c \leftrightarrow (c \in c \rightarrow B)$. Contraction and *modus ponens* then give *B*. (This is Curry's paradox.) RM_3, K_3, B_3, and $Ł_n$ (for finite *n*) all contain *modus ponens* and, if not contraction, something closely related to it that will do the same job. However, the schema based on $Ł_{\aleph}$ is consistent. If the extensionality principle $(\forall x(x \in y \leftrightarrow x \in z) \rightarrow y = z)$ is added, though, then even $Ł_{\aleph}$ gives triviality. (Virtually the same comments can be made about the naive *T*-schema ("*A*" is true $\leftrightarrow$ *A*) when self-reference is present. Though here extensionality is, of course, not an issue.)

For a final example of the philosophical application of many-valued logics: It is often claimed that the appropriate semantics for a language with vague predicates is one with degrees of truth. Such logics now usually go under the rubric of fuzzy logics. $Ł_{\aleph}$ is a paradigm one such. (It is not the only one: $Ł_{\aleph}$ is one of a family of logics in which $Val = [0, 1]$. Each is based on a so-called t-norm—essentially a function stating the truth conditions for an appropriate conjunction connective.) The only logical inference that the simplest form of the Sorites paradox uses is *modus ponens*. This is valid in $Ł_{\aleph}$; but if one changes *Des* to, say, [0.8, 1], it is not. (Let $v(p) = 0.9 \in Des$, $v(q) = 0.7 \notin Des$. Then $v(p \rightarrow q) = 0.8 \in Des$.) Note that probability theory is not a many-valued logic. The probability of a compound sentence is not determined by the probabilities of it components. (Let *a* and *b* be independent fair coins. Let A_H be "Coin *a* will come down heads"; A_T be "Coin *a* will come down tails"; and B_H be "Coin *b* will come down heads." $\text{Prob}(A_H) = \text{Prob}(A_T) = \text{Prob}(B_H) = 0.5$. But $\text{Prob}(A_H \wedge A_T) = 0$ and $\text{Prob}(A_H \wedge B_H) = 0.25$.)

TECHNICAL APPLICATIONS

Many-valued logics have various technical applications. Perhaps the most important of these, in a philosophical context, is their use in proving independence results. Thus, suppose that one has some axiom system, *T*, and wishes to know whether some formula, *A*, is deducible in it. One way to show that it is not is to construct a many-valued logic such that all the axioms of *T* always take a designated value, and all the rules of *T* preserved designated values. It follows that all theorems always take designated values. If one can find an interpretation of the logic in which *A* does not take a designated value, it follows that it cannot be proved.

For example, the following is a set of axioms for the $\rightarrow/\neg$ fragment of the relevant logic often called RW (R minus contraction). The only rule of inference is *modus ponens*:

$A \rightarrow A$

$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$

$A \rightarrow ((A \rightarrow B) \rightarrow B)$

$\neg\neg A \rightarrow A$

$(A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$

Now consider the three-valued Łukasiewicz logic, $Ł_3$. One can check (e.g., by truth tables) that all the axioms always take the designated value and that *modus ponens* preserves that property. Now let C be the formula: $(p \rightarrow (p \rightarrow q)) \rightarrow (p \rightarrow q)$. Take an evaluation, v, in which $v(p) = ½$ and $v(q) = 0$. Computation verifies that $v(C) = ½$. Hence C is not provable. Since $v(\neg C) = ½$ as well, $\neg C$ cannot be proved either. Hence, C is independent of RW.

A much more technically demanding example of the use of many-valued logics to prove independence is in set theory. If one takes the values of the logic to be those of any Boolean algebra, taking the top value as the only designated value, and interprets the connectives and quantifiers in appropriate ways, the logic delivered is classical logic. Choosing the Boolean algebra in an appropriately set-theoretic way, one can also show that the axioms (and so theorems) of Zermelo Fraenkel set theory, ZF, take the designated value. Choosing the algebra in more cunning fashions, one can show that various important set-theoretic principles, such as the continuum hypothesis, do not receive designated values. Hence, ZF does not entail the continuum hypothesis.

HISTORY, PERSONS, AND REFERENCES

This entry concludes by putting the investigations discussed earlier in their historical context. Relevant references that may be consulted for further details are also given at the end of each paragraph. For a gentle introduction to many-valued logics, see Graham Priest (2001, chapters 7, 8, 11); for a more detailed introduction, see Alasdair Urquhart (2001); and for further detailed technical discussions, see Richard Hähnle (2001). J. Michael Dunn and George Epstein (1977) provide a bibliography of work on many-valued logics up to 1974.

The first modern many-valued logic was $Ł_3$. This, and its generalization to n-valued logics, $Ł_n$, were published by Łukasiewicz around 1920. At about the same time, the U.S. mathematician Emil Post was also constructing finitely many-valued logics. (The most significant feature of Post's systems is its treatment of negation. If the values of the n-valued logic are 0, 1, … , n–1, then $v(\neg A) =| 1 + v(A)|$ (Mod n). Philosophical applications of this many-valued logic are difficult to find.) The logic $Ł_{\aleph}$ was published by Łukasiewicz and Alfred Tarski in 1930. Much of the early investigation of many-valued logics and their axiomatizations were carried out by Polish logicians including Mordechaj Wajsberg and Jerzy Słupecki. Finding a demonstrably complete axiom system for $Ł_{\aleph}$ turned out to be a hard problem. Reputedly, it was solved by Wajsberg, but the first proofs to be published were by Alan Rose and Berkeley Rosser and by Chen Chung Chang in the late 1950s. The unaxiomatizability of first-order $Ł_{\aleph}$ was proved by Bruno Scarpellini in 1962. (Łukasiewicz 1970, Rosser and Turquette 1952, Wójcicki 1988, Malinowski 1993.)

Canonical statements of the other many-valued logics mentioned in this entry were given by the following: B_3, Dmitryi Anatol'evich Bochvar, 1939; K_3, Stephen Kleene, 1952; FDE and RM_3, Alan Ross Anderson and Nuel Belnap, 1975; LP, Graham Priest, 1979. (Rescher 1969, Priest 2001.)

The proof that intuitionist logic is not many-valued was first given by Kurt Gödel in 1933. The idea was applied to modal logic by James Dugunji in 1940. The earliest versions of the idea that every logic has a many-valued semantics are usually attributed to Adolf Lindenbaum in the 1920s. Generalizations are due to Jerzy Łos and Roman Suszko in 1958. (Hughes and Cresswell 1968, Shoesmith and Smiley 1978, Wójcicki 1988.)

The applicability of many-valued logics to the view that some sentences are neither true nor false was pursued by many people in the second half of the twentieth century. These include Richard Routley, Leonard Goddard, Saul Kripke, Kit Fine, and Scott Soames. Supervaluations were invented by van Fraassen in 1969. Toward the end of the twentieth century, their application to vagueness became a very standard idea. The application of many-valued logics to the view that some sentences are both true and false, though less popular, has been pursued by various paraconsistent logicians. These include Newton da Costa, Priest, Routley, and Dominic Hyde. The generalization of supervaluation to logics with gluts as well as gaps was developed by Achille Varzi in the 1990s. (Rescher 1969, Scott 1974, Haack 1978, Dunn and Epstein 1977, Humberstone 1998, Varzi 2000, Priest 2001.)

The possibility of basing the naive comprehension schema for sets on $Ł_{\aleph}$ was investigated by Thoralf Skolem and Chang in the 1950s. The consistency of the schema (and the inconsistency of extensionality) was proved by Richard White in 1979. (White 1979.)

Fuzzy logics and their applicability to vagueness have been investigated fairly intensely since about the 1970s, by many people, including Kenton Machina and Patrick Grim, and, on the technical side, Lotfi Zadeh, Petr Hájek,

and Daniele Mundici. (Keefe 2000, Hájek 1998, Cignoli, D'Ottaviano, and Mundici 2000.)

The use of many-valued logics in independence investigations goes back to the early years of the subject, though this has flourished with the proliferation of non-classical logics in the second half of the twentieth century. One of the earliest techniques for proving independence results in set theory is that of forcing, developed by Paul Cohen in the early 1960s. That similar things could be done with Boolean-valued models was realized by Robert Solovay, Dana Scott, and others a few years later. (Anderson and Belnap 1975, Bell 1985.)

See also Fuzzy Logic; Intuitionism and Intuitionistic Logic; Logic, History of; Logic, Non-Classical; Modal Logic; Paraconsistent Logics; Relevance (Relevant) Logics; Set Theory.

Bibliography

Anderson, Alan Ross, and Nuel D. Belnap Jr. *Entailment: The Logic of Relevance and Necessity*. Vol 1. Princeton, NJ: Princeton University Press, 1975.

Bell, John. *Boolean-Valued Models and Independence Proofs in Set Theory*. New York: Oxford University Press, 1985.

Cignoli, Roberto L. O., Itala M. L. D'Ottaviano, and Daniele Mundici. *Algebraic Foundations of Many-Valued Reasoning*. Dordrecht, Netherlands: Kluwer Academic, 2000.

Dunn, J. Michael, and George Epstein, eds. *Modern Uses of Multiple-Valued Logic*. Dordrecht, Netherlands: D. Reidel, 1977.

Haack, Susan. *Philosophy of Logics*. New York: Cambridge University Press, 1978.

Hähnle, Richard. "Advanced Many-Valued Logic." In *Handbook of Philosophical Logic*, Vol. 3. 2nd ed., edited by D. Gabbay and F. Guenthner. Dordrecht, Netherlands: Kluwer Academic, 2001.

Hájek, Petr. *Metamathematics of Fuzzy Logic*. Dordrecht, Netherlands: Kluwer Academic, 1998.

Hughes, George, and Max Cresswell. *Introduction to Modal Logic*. London: Methuen, 1968. Revised as *A New Introduction to Modal Logic*. London: Routledge, 1996.

Humberstone, Lloyd. "Many-Valued Logics, Philosophical Issues in." In *Routledge Encyclopedia of Philosophy*, Vol. 6., edited by E. Craig. London: Routledge, 1998.

Keefe, Rosanna. *Theories of Vagueness*. New York: Cambridge University Press, 2000.

Łukasiewicz, Jan. *Selected Works*, edited by L. Borkowski. Amsterdam, Netherlands; North-Holland, 1970.

Malinowski, Grzegorz. *Many-Valued Logics*. Oxford, U.K.: Clarendon Press, 1993.

Priest, Graham. *Introduction to Non-classical Logic*. New York: Cambridge University Press, 2001.

Rescher, Nicholas. *Many-Valued Logic*. New York: McGraw Hill, 1969.

Rosser, J. Barkely, and Atwell R. Turquette. *Many-Valued Logics*. Amsterdam, Netherlands: North-Holland, 1952.

Scott, Dana. "Does Many-Valued Logic Have Any Use?" In *Philosophy of Logic*, edited by S. Körner. Oxford, U.K.: Blackwell, 1974.

Shoesmith, David J., and Timothy J. Smiley. *Multiple Conclusion Logic*. New York: Cambridge University Press, 1978.

Urquhart, Alasdair. "Basic Many-Valued Logic." In *Handbook of Philosophical Logic*. Vol. 2. 2nd ed., edited by D. Gabbay and F. Guenthner. Dordrecht, Netherlands: Kluwer Academic, 2001. This is a revised version of "Many-Valued Logic" in Vol. 3 of the first ed (Dordrecht, Netherlands: D. Reidel, 1986).

Varzi, Achille. "Supervaluationism and Paraconsistency." In *Frontiers in Paraconsistent Logic*, edited by D. Batens et al. Baldock, U.K.: Research Studies, 2000.

White, Richard B. "The Consistency of the Axiom of Comprehension in the Infinite-Valued Predicate Logic of Łukasiewicz." *Journal of Philosophical Logic* 8 (1979): 509–534.

Wójcicki, Ryszard. *Theory of Logical Calculi: Basic Theory of Consequence*. Dordrecht, Netherlands: Kluwer Academic, 1988.

Graham Priest (2005)

MANY WORLDS/MANY MINDS INTERPRETATION OF QUANTUM MECHANICS

The many worlds/many minds formulations of quantum mechanics are reconstructions of Hugh Everett III's (1957a, 1957b, 1973) relative-state formulation of quantum mechanics. Each is presented as a proposal for solving the quantum measurement problem. Much of the philosophical interest in these theories derives from the metaphysical commitments they suggest. They illustrate the roles played by traditional metaphysical distinctions both in formulating and in evaluating physical theories. They also illustrate the range of metaphysical options one must consider if one wants a metaphysics that is consistent with the structure of the physical world suggested by the best physical theories.

The quantum measurement problem is a consequence of the orthodox quantum-mechanical representation of physical properties. In order to account for interference effects, the orthodox view requires that one allows for a physical system to be in a *superposition* of having mutually incompatible classical physical properties. An electron *e* might, for example, be in a superposition of being in New York City and being in Los Angeles. If the unit-length vector $(NYC)_e$ represents the electron

being in New York City and if the orthogonal unit-length vector $(LA)_e$ represents the electron being in Los Angeles, then the state of the electron in a superposition of being in each city is represented by

$$(S)_e = a(NYC)_e + b(LA)_e,$$

where *a* and *b* are complex numbers, such that *a*-squared plus *b*-squared equals one. On the orthodox view, the state represented by the unit-length vector $(S)_e$ is not a state where the electron is determinately in NYC, it is not a state where the electron is determinately in LA, it is not a state where the electron is determinately in both cities, and it is not a state where the electron is determinately in neither city. Rather, on the standard interpretation of states, an electron in state $(S)_e$ simply fails to have a determinate position.

While allowing for superpositions of classical properties explains the counterintuitive empirical results of interference experiments, it leaves a puzzle: If electrons are sometimes in such superpositions of position, then why do electrons have determinate positions whenever one looks for them? In its most general form the quantum measurement problem is to explain why physical systems exhibit quantum interference effects, which typically involves talk of superpositions, and to explain why people—when they look for them—always observe physical systems to have determinate physical properties.

LINEAR DYNAMICS AND COLLAPSE DYNAMICS

The standard von Neumann-Dirac collapse formulation of quantum mechanics (1955 [1932]) explains interference effects and definite measurement results by stipulating two dynamical laws. The *linear dynamics* describes the deterministic continuous evolution of the state of a physical system when no measurement of the system is made. It is this law that describes the evolution of physical systems in superpositions of classical properties and thus explains quantum interference effects. The *collapse dynamics* describes the random discontinuous evolution of the state when a measurement is made of the physical system. It is this law that explains how one gets determinate measurement records at the end of an observation and makes the standard statistical predictions. More specifically, in the case of the electron in state $(S)_e$, if an observer *M* looks for the electron in NYC, the collapse dynamics predicts that the state will instantaneously and randomly evolve from

$$(Ready)_m\ (a(NYC)_e + b(LA)_e),$$

a state where M is ready to look for *e* and *e* is in state $(S)_e$, to either ("In NYC")$_m$ $(NYC)_e$ (with probability (a^2), in which case *e* is now determinately in NYC and *M* determinately records this fact, or to ("Not in NYC")$_m$ $(LA)_e$ (with probability (b^2), in which case *e* is now determinately in LA and *M* determinately records that it is not found in NYC.

In order to understand the work done by the collapse dynamics in the standard theory, consider what would happen without the collapse of the quantum-mechanical state. In the measurement above, the linear dynamics predicts that the postmeasurement state of the observer who correlate their records perfectly with the position of the electron, written in the determinate record basis, is

$$(E) = a(\text{“In NYC”})_m\ (NYC)_e + b(\text{“Not in NYC”})_m\ (LA)_e.$$

On the standard interpretation of states, *M* here has no determinate measurement record. Rather, without the collapse dynamics, *M* ends up in an entangled superposition of finding and not finding the electron. This is presumably not what happens.

So, in the standard theory, the collapse dynamics is both responsible for the theory making the standard quantum statistical predictions and for the explanation of determinate measurement results. But because the physical state that results from applying the collapse dynamics to a system is typically different from the state that results from applying only the linear dynamics, the standard formulation of quantum mechanics is at best incomplete and arguably logically inconsistent on a strict reading—unless one can stipulate strictly disjoint conditions for when each dynamical law obtains. In the context of the standard collapse theory, solving the measurement problem would require one to stipulate exactly what interactions count as measurements and hence cause collapses.

Rather than stipulating when collapses occur, Everett's proposal for solving the quantum measurement problem involved denying that there are collapses. More specifically, Everett proposed simply dropping the collapse dynamics from the standard von Neumann-Dirac theory of quantum mechanics and taking the resulting pure wave mechanics as a complete and accurate description of all physical systems. Everett then intended to deduce the standard statistical predictions of quantum mechanics—the predictions that are explained by the collapse dynamics in the standard formulation of quantum mechanics—as subjective experiences of observers who are themselves treated as ordinary physical systems within the new theory. Dropping the collapse dynamics clearly eliminates potential conflict between the two

dynamical laws; but if one drops the collapse dynamics, one must then explain how we obtain determinate measurement results that exhibit the standard quantum statistics when the linear dynamics alone typically predict entangled postmeasurement superpositions such as $(E)_e$.

RECONSTRUCTING EVERETT'S THEORY

While it is clear that Everett intended for his relative-state formulation of quantum mechanics to explain why one gets determinate measurement results, it is unclear how this was supposed to work. There are several alternative reconstructions of Everett's theory in the literature, all designed to provide quantum mechanics without the collapse dynamics with determinate measurement records while somehow recovering the standard quantum statistics. The many worlds and the many minds formulations of quantum mechanics represent two general approaches to reconstructing Everett's relative-state formulation of quantum mechanics.

The splitting worlds formulation is perhaps the most popular version of the many worlds formulation. The splitting world formulation of quantum mechanics

> asserts that it makes sense to talk about a state vector for the whole universe. This state vector never collapses and hence reality as a whole is rigorously deterministic. This reality, which is described *jointly* by the dynamical variables and the state vector, is not the reality we customarily think of, but is a reality composed of many worlds. By virtue of the temporal development of the dynamical variables the state vector decomposes naturally into orthogonal vectors, reflecting a continual splitting of the universe into a multitude of mutually unobservable but equally real worlds, in each of which every good measurement has yielded a definite result and in most of which the familiar statistical quantum laws hold. (DeWitt and Graham 1973, p. v)

Proponents of this view admit that the metaphysical commitments it suggests are counterintuitive: "I still recall vividly the shock I experienced on first encountering this multiworld concept. The idea of 10^{100} slightly imperfect copies of oneself all constantly spitting into further copies, which ultimately become unrecognizable, is not easy to reconcile with common sense. Here is schizophrenia with a vengeance" (DeWitt and Graham 1973, p. 161).

But it is precisely these counterintuitive commitments that explain why observers end up recording determinate measurement results. On the splitting worlds formulation the universe splits whenever one makes a measurement in such a way that every physical possible result in fact determinately occurs in some future world. More specially, there is one world corresponding to each term in the expression of the quantum mechanical state when written in the theory's preferred basis. In choosing the preferred basis, one chooses a single preferred way from among the many different, mathematically equivalent, ways of representing quantum-mechanical states as the sum of mutually orthogonal unit-length vectors. On the splitting worlds formulation, the preferred basis is chosen so that each term in the expansion of the state describes a world where there is a determinate measurement record. The state (E) above describes two worlds: One where the observer *M* determinately records the measurement result "In NYC" and *e* is in fact in NYC and another where *M* determinately records "Not in NYC" and *e* is in fact in LA.

PROBLEMS WITH THE SPLITTING WORLDS FORMULATION. While the splitting worlds formulation of quantum mechanics does explain why there are determinate measurement records, it encounters other problems. A standard complaint is that the theory is ontologically extravagant. We presumably only ever need one physical world, our world, to explain our experiences. The reason for postulating the actual existence of a different physical world corresponding to each term in the quantum-mechanical state is that it allows one to explain our determinate experiences while taking the deterministically evolving quantum-mechanical state to be in some sense a complete and accurate description of the physical facts. But again one might wonder whether the sort of completeness one gets warrants the many-world ontology.

Another problem with the splitting worlds formulation concerns the statistical predictions of future events. The standard collapse formulation of quantum mechanics predicts that *M* will get the result "In NYC" with probability *a*-squared and the result "Not in NYC" with probability *b*-squared in the above experiment, and this is what is observed as relative frequencies for such experiments. Insofar as there will be two copies of *M* in the future, *M* is guaranteed to get each of the two possible measurement results. So, in this sense at least, the probability of *M* getting the result "In NYC" is one, which is simply not what is observed if both a and b are nonzero. A principle of indifference might lead one to assign a probability of one-half to each of the two possible meas-

urement outcomes. But not only would such a principle be difficult to justify here, probability one-half for each possible outcome is typically not what would be observed for such experiments as relative frequencies. So while the splitting worlds formulation explains why observers get determinate measurement records, as it stands, it makes no empirical predictions for the likelihood of future events.

In order to understand what one would have to add to the theory to get the standard quantum statistical predictions for future events, one might note that the question "What is the probability that *M* will record the result 'In NYC'?" is, strictly speaking, nonsense—unless one has an account of the transtemporal identity of the observer *M*. Because there is no rule that states which worlds are which at different times, the splitting worlds theory is prevented from making statistical predictions concerning an observer's future experiences. And not being able to account for the standard quantum probabilities is a serious problem because it was the successful statistical predictions of quantum mechanics that made quantum mechanics worth taking seriously in the first place.

Another problem for the splitting worlds formulation of quantum mechanics concerns the way worlds are supposed to split. In order to explain the determinate measurement records, one must choose a preferred basis so that observers have determinate measurement records in each term of the quantum-mechanical state when written in the preferred basis. The problem is that not just any basis will make records determinate in every world (consider, for example, a basis that includes the vector (E) above). Selecting the preferred basis to use determines when worlds split, and determining when worlds split is as difficult as trying to determine when the collapse occurs in the standard formulation of quantum mechanics. This is the preferred basis problem. This problem is closely analogous to the original measurement problem in the context of the standard collapse formulation of quantum mechanics.

A popular strategy for resolving the preferred basis problem is to try to find a criterion involving the interaction between a quantum-mechanical system and its environment that would dynamically select a preferred basis for a system. As a simple example of an environmental decoherence criterion, one might take the preferred basis of a system to be the one that represents the classical property of the system to which its environment becomes most strongly correlated, whatever this may be. Insofar as a measurement record is easily read, one might argue, the environment becomes strongly with the value of the record, so such a criterion would be expected select the determinate-record basis as preferred. One problem with having the environment of a system select the preferred basis, however, is that, in the case of the splitting worlds formulation at least, one presumably needs a preferred basis for the entire universe, which does have an environment.

THE MANY MINDS FORMULATION. David Albert and Barry Loewer's many minds formulation of quantum mechanics (1988) provides another approach for interpreting Everett's relative-state formulation of quantum mechanics. Everett said that his theory "is objectively continuous and causal, while subjectively discontinuous and probabilistic" (1973, p. 9). The many minds formulation of quantum mechanics captures this feature by distinguishing between an observer's *physical state* and its evolution, which is continuous and causal, and an observer's *mental state* and its evolution, which is discontinuous and probabilistic. This is a sort of hidden-variable theory, where the variable being added to the standard quantum-mechanical state is the mental states of observers. Stipulating determinate mental states solves the quantum measurement problem by directly providing observers with determinate, accessible measurement records.

In order to get the observer's complete mental state to supervene on her or his physical state, Albert and Loewer associate with each observer a continuous infinity of minds. The standard quantum-mechanical state always evolves in the usual deterministic linear way, but each mind evolves randomly, with probabilities determined by the particular mind's current mental state and the evolution of the quantum-mechanical state. In the experiment above, Albert and Loewer's mental dynamics predicts that the probability of each of the observer's minds becoming randomly associated with the result "In NYC" (the first term of (E)) is *a*-squared and that the probability of each becoming randomly associated with the result "Not in NYC" (the second term of (E)) is *b*-squared.

An advantage of the many minds formulation over the splitting worlds formulation is that here there is no physically preferred basis. One must choose a preferred basis in order to specify the mental dynamics completely, but this choice has nothing to do with any physical facts. Rather, it can be thought of as part of the description of the relationship between physical and mental states. Another advantage of the many minds formulation is that, unlike the splitting worlds formulation, it makes the standard probabilistic predictions for the future measurement results of each mind. Because the states of particu-

lar minds do not supervene on the physical state here, in order to talk about their states and how they evolve, one must suppose that individual minds have transtemporal identities, which in turn requires a commitment to a strong form of mind-body dualism. But it is also this strong dualism that makes the many minds theory one of the few formulations of quantum mechanics that resolves the quantum measurement problem and is manifestly compatible with special relativity.

One might wonder whether the sort of mental supervenience one gets in the many minds formulation (it is not the states of an observer's individual minds, but only the complete distribution of the states of all of these minds that can be taken to supervene on her or his physical state) is worth the trouble of postulating a continuous infinity of minds associated with each observer. Another option is to suppose that each observer has a single mind that evolves in the Albert and Loewer random way. But here one sacrifices all but the weakest sort of supervenience of mental states on physical states. Here the physical state would only tell one the probabilities of various mental states obtaining.

If one wants to avoid the mind-body dualism involved in the many minds formulation, one can use the evolution of minds to construct an alternative many worlds formulation. On one such theory, the many threads formulation of quantum mechanics, worlds do not split. Rather, one stipulates that there is one world corresponding to each possible trajectory of a single Albert-Loewer mind and that the history of that world is described by the history of the world that would be observed by the mind. Each observer then inhabits exactly one of these worlds, and the global quantum-mechanical state is used to assign prior epistemic probabilities to each physically possible world in fact being a given observer's world. These prior probabilities, concerning which possible world is an observer's actual world, might then be updated as she learns more about the history of her world. In the simplest case, she eliminates from contention all possible worlds that are incompatible with a particular observed event. Unlike the splitting worlds formulation, there is no special problem in understanding probabilities of future events on this account. A particular event is either going to happen or not in our world. The standard quantum probabilities here simply represent our posterior uncertainly concerning which world we in fact inhabit (Barrett 1999).

See also Quantum Mechanics.

Bibliography

Albert, D. Z. *Quantum Mechanics and Experience*. Cambridge, MA: Harvard University Press, 1992.

Albert, D., and B. Loewer. "Interpreting the Many Worlds Interpretation." *Synthese* 77 (1988): 195–213.

Barrett, J. *The Quantum Mechanics of Minds and Worlds*. Oxford: Oxford University Press, 1999.

DeWitt, B. S. "The Many-Universes Interpretation of Quantum Mechanics." In *Foundations of Quantum Mechanics*. New York: Academic Press, 1971. Reprinted in *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham, 167–218. Princeton, NJ: Princeton University Press, 1973.

DeWitt, B. S., and N. Graham, eds. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton, NJ: Princeton University Press, 1973.

Dowker, F. A., and A. Kent. "On the Consistent Histories Approach to Quantum Mechanics." *Journal of Statistical Physics* 83, nos. 5–6 (1983): 1575–1646.

Everett, H. *On the Foundations of Quantum Mechanics*. Thesis. Princeton University, New Jersey. 1957a.

Everett, H. "'Relative State' Formulation of Quantum Mechanics." *Reviews of Modern Physics* 29 (1957b): 454–462.

Everett, H. "The Theory of the Universal Wave Function." In *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham. Princeton, NJ: Princeton University Press, 1973.

von Neumann, J. *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton University Press, 1955. Translated by R. Beyer from *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer, 1932).

Jeffrey A. Barrett (2005)

MARBURG, SCHOOL OF

See *Neo-Kantianism*

MARCEL, GABRIEL

(1889–1973)

Gabriel Marcel, the French philosopher, dramatist, and critic, was born in Paris. His father, a highly cultured man, held important administrative posts in the Bibliothèque Nationale and the Musées Nationaux. Marcel's mother died when he was four. Raised in a home dominated by the cultured agnosticism of his father and the liberal, moralistic Protestantism of his aunt, and nurtured in a scholastic system concerned only with intellectual achievement, he later sought refuge in a modified type of idealism. The shaking experiences of World War I, during which he was an official of the Red Cross concerned with locating missing soldiers, brought home to him the failure of abstract philosophy to cope with the

tragic character of human existence. His conversion to Catholicism in 1929 did not substantially alter the direction of his thought, although it intensified his conviction that the philosopher must take into consideration the logic interior to faith and hope.

RELATIONSHIP TO EXISTENTIALISM

Marcel's name has most often been linked with "theistic existentialism." Because of the ambiguities of this term and the association of existentialism in the popular mind with Jean-Paul Sartre's philosophy, to which his is almost diametrically opposed, Marcel has preferred the designation "Neo-Socratic" for his thought. This should not obscure Marcel's contributions to existential philosophy or his similarity to other thinkers who are ordinarily associated with it.

Before publication of the major philosophical works of Karl Jaspers and Martin Heidegger, Marcel introduced into French philosophy, in his essay "Existence and Objectivity" (1925) and in his *Metaphysical Journal*, many of the themes that later became central to existentialism. Often making use of an independently developed phenomenological method, he dealt with such themes as participation, incarnation, man as being in the world, and the priority of existence over abstraction (the *cogito*) as a starting point for philosophy.

Marcel's critique of idealism and his defense of faith resemble Søren Kierkegaard's critique of G. W. F. Hegel; Marcel, however, refuses to allow that faith is an irrational leap or that the individual stands alone in his faith. Heidegger and Marcel explore much of the same terrain in seeking to restore the "ontological weight to human experience" (*Being and Having*, p. 103). They share a common view of the nature of truth and language. Marcel, however, unlike Heidegger, includes within his ontology the assurance of fulfillment that is part of faith's apprehension of God as Absolute Presence. In many ways Martin Buber has been Marcel's closest contemporary philosophical relative. Each has independently developed a philosophy of dialogue and communion in which the distinction between the relation of an I to a thou and an I to an it or a him plays a central part.

PHILOSOPHICAL METHOD

A great injustice is necessarily done in any summary account of Marcel's thought, for the charm and the convincing power of his conclusions are inseparable from his itinerant, tentative, and exploratory philosophical method. One of the most characteristic features of his thinking is the vigor with which he combated the spirit of abstraction and the conceptual sclerosis that he believes is an occupational hazard of systematic and academic philosophers. But despite his rejection of systematic philosophy, Marcel's work is based on an underlying principle of unity, or more accurately an underlying vision, which, seen dimly from the beginning, has been progressively more clearly apprehended. This vision, which is essentially both Platonic and Christian, expresses itself in the conviction that within the temporal and transient order *homo viator* is given a foretaste of eternal realities.

Marcel's philosophical explorations cannot be divorced from his dramatic writings or from his experimentation in music. His plays are not philosophical in the sense of being popular forums for the presentation of worked-out ideas. Rather, they present complicated situations in which persons find themselves trapped, challenged, and confused; and thus indirectly they explore the nature of the exile into which the soul enters as it becomes alienated from itself, from those it loves, and from God. Marcel believes that in music one finds a foretaste or presentiment of the perfect harmony and communion toward which all authentic human existence strives. Philosophy shares both in the tension that is the essence of drama and in the harmony which is the essence of music. Its starting point is a metaphysical "dis-ease" like that of a person in a fever who shifts around searching for a comfortable position. This search for a home in the wilderness, a harmony in disharmony, a transcendent source of assurance in a transient life takes place through a reflective process that Marcel calls secondary reflection.

THE NATURE OF THINKING

Marcel distinguishes two degrees or types of thinking, primary and secondary reflection. Primary reflection is characterized as abstract, analytical, objective, universal, and verifiable. The thinking subject in primary reflection is not the individual human person but the thinker qua mind (the *Bewusstsein überhaupt*). Primary reflection deals with the realm of the problematic. As the etymology of "problem" (*pro-ballo*) suggests, the distinguishing feature of the problematic approach to reality is the separation of the questioner from the data about which he questions. The data of primary reflection lie in the public domain and are equally available to any qualified observer. Once a problem is posed, primary reflection proceeds to abstract from the concrete data any elements that are not relevant to the solution of the particular problem under consideration. When a solution or an explanation has been found, the original curiosity and tension that motivated the thinker are alleviated.

Primary reflection, as exemplified in scientific and technical thought, has allowed us to possess and manipulate our world more completely and is therefore indispensable to human culture. However, intellectual and moral confusion results when primary reflection becomes imperialistic and claims the right to judge all knowledge and truth by criteria appropriate only to the realm of the objective and the problematic. When this happens, abstraction gives way to "the spirit of abstraction," the use of techniques gives way to technocracy, and the inexhaustible riches of a kaleidoscopic world are forced to conform to a black-and-white logic.

Secondary reflection is concrete, individual, heuristic, and open. Strictly speaking, it is concerned not with objects but with presences. Its contemplation begins not with curiosity or doubt but with wonder and astonishment. Hence, it is humble in its willingness to be conformed to categories created by that on which it is focused. It remains open to its object as a lover does to his beloved—not as a specimen of a class but as a unique being. This openness is not a methodological principle as in scientific thought but arises from the possibility of something new being created in the relationship. Secondary reflection is dialogical, not dialectical. Rather than searching for information about the other and dealing with it abstractly, secondary reflection seeks the revelation of total presence, whether the presence be that of my body, the world, the other person, or God. Thus, secondary reflection is brought to bear on data or questions from which the thinker as existing person cannot legitimately abstract himself: "Am I free?" "Is there meaning and value in life?" "Can I commit myself to this person?" In other words, secondary reflection is concerned not with problems but with mystery.

MYSTERY

According to Marcel, a mystery initially appears to be merely a problem that is difficult to solve. Reflection shows, however, that in dealing with a genuine mystery the distinction between subject and object, between what is in me and what is before me, breaks down. Faced with questions about freedom, the meaning of life, the existence of God, and so forth, no objective standpoint can be found from which a universally valid answer may be discovered. This does not mean that mystery is unknown or unknowable and lies in a realm of vague feelings over which thought has no grasp. Rather, knowledge of mystery presupposes an immediate participation, or what Marcel also calls a "blinded intuition," but this participation is understood only with the aid of a conceptual process. Unaided intuition is not an adequate philosophical instrument. However, secondary reflection penetrates into the mystery of existence and being only when it works in conjunction with love, fidelity, faith, and the other "concrete approaches." It yields a kind of knowledge and truth that, if unverifiable, nevertheless is confirmed as it illuminates our lives. Two foci of mystery may be distinguished, although never separated, in Marcel's thinking. The mystery of existence is dealt with in "concrete" philosophy and the mystery of being in "concrete" ontology.

CONCRETE PHILOSOPHY

Marcel denies that the detached, disincarnate, Cartesian *cogito* provides a possible starting point for a concrete philosophy. It is with the existing subject, the incarnate being who is already in the world, that philosophy must begin. The experience of the inexhaustible concreteness of the existing world can be neither deduced, doubted, nor demonstrated. Existence is not a thing, a quality, or a discrete content of thought that can be isolated and pointed out; rather it is that in which the subject participates and from which thought begins its quest for meaning. The assurance of existence that we have is not of the intellectual order but is an outcome of our direct participation in the world via sensation and feeling. Because sensation and feeling are inseparable from the body, our knowledge of existence is tied up with our being incarnate.

Incarnation is the "central given of metaphysic," the absolute starting point for an existential philosophy, because it is on the analogy of my experience of my body that the world is understood. I project into the world the sense of density and presence that I experience when I become aware of my own body. The world exists for me only in the measure that I am related to it in a way similar to the way in which I am related to my own body.

As I am not even ideally separable from my body, I am likewise inseparable from my situation. Those habitual surroundings and historical conditions that shape my life enter into the very fiber of what I am. Insofar as I recognize that my situation enters into the constitution of my being, and hence that I am not able to abstract myself from it completely and view it with the objective detachment of a spectator, I may speak of the family that nurtured me or of an illness that shaped me as having a mysterious character.

A concrete philosophy must also affirm the immediacy of our being with others. The principle of the intentionality of consciousness, Marcel holds, applies in our

relations both to persons and to the world. Philosophy begins not with *I am* but with *we are*.

The significance of this intersubjectivity will be determined by the type of relations that characterize one's life. The self who treats other persons as objects to be manipulated and used is condemned because of its egocentricity to live in a world lacking in ontological depth, and hence it will be prey to despair when the thrill of possession wears thin. To endeavor to allow the other person to become present as a thou is to enter into a relationship within which the assurance of fulfillment is received.

ONTOLOGY

No word used by Marcel is more difficult to define or richer in meaning than *being*. It refers neither to the sum total of all objects that exist nor to some universal substratum underlying all particulars. Being is eternal and inexhaustible. It is "that which does not allow itself to be dissolved by the dialectics of experience" (*Metaphysical Journal*, p. 181). Only by participation in being can isolation, despair, and tragedy be overcome. The quest for being is thus identical with the quest for salvation. To deny being is to say that "all is vanity," that nothing has intrinsic worth. To affirm being is to declare that corresponding to the deepest exigency of the human spirit is a fulfillment of which an earnest is given in experiences of creativity, joy, and love.

As defined by Marcel, the question of being cannot be approached objectively and problematically. Being can be affirmed only if I can discover within experience some presence that testifies to being. Two elements in human experience seem to offer such a testimony. First, at the heart of the human condition is an "ontological exigence," an impulse to transcendence that is present in all authentic human life, the exigence to penetrate to a level of experience saturated with meaning and value. The mere existence of such an exigence is no guarantee in itself that a corresponding satisfaction exists. It could be the case, as Sartre says, that man is a "useless passion." But Marcel attempted to show, by way of a phenomenological analysis, that certain experiences of love, joy, hope, and faith, as understood from *within*, present a positive testimony to the existence of an inexhaustible presence. This assuring presence, which might be called the immanence of being in human experience, is never a possession but is constantly created anew as an I enters into relations with an empirical thou or the Absolute Thou (God). Although the assurance of being never becomes conceptually clear, it provides the illumination making creative, open existence possible.

In what might be called Marcel's ontological personalism, the concrete approaches to being are identical with the approaches to other persons and to God. To enter into a loving relationship requires that a person exorcise the spirit of egocentricity and possession and become spiritually available (*disponible*) to others. A vow of creative fidelity is likewise necessary if the unconditional demands of love are to be satisfied. In approaching God, fidelity becomes faith and *disponibilité* becomes hope. In love, fidelity, hope, and faith man approaches the mystery of being and is overtaken with the assurance that he is accompanied by the eternal fulfilling Presence that he seeks to know.

See also Being; Existentialism; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Idealism; Jaspers, Karl; Kierkegaard, Søren Aabye; Personalism; Philosophy of Religion, History of; Sartre, Jean-Paul; Thinking.

Bibliography

WORKS BY MARCEL

A thorough and updated bibliography of works by Marcel, translations of his works, and works about Marcel is maintained by the Gabriel Marcel Society at http://www.lemoyne.edu/gms.

Philosophical Works

Journal métaphysique. Paris: Gallimard, 1927. Translated by Bernard Wall as *Metaphysical Journal*. Chicago: Regnery, 1952.

Être et avoir. Paris: Aubier, 1935. Translated by Katharine Farrer as *Being and Having*. London: Dacre Press, 1949.

Du refus à l'invocation. Paris: Gallimard, 1940. Translated by Robert Rosthal as *Creative Fidelity*. New York: Noonday Press, 1964.

Homo Viator. Paris: Aubier, 1945. Translated by Emma Craufurd. New York, 1962.

The Philosophy of Existence. Translated by Manya Harari. New York: Philosophical Library, 1949. Republished as *Philosophy of Existentialism*. New York: Citadel Press, 1961. Previously uncollected essays written between 1933 and 1946. Perhaps the best short introduction to Marcel's thought.

Le mystère de l'être. 2 vols. Paris: Aubier, 1951. Translated by G. S. Fraser and René Hauge as *The Mystery of Being*. 2 vols. Chicago, 1950.

Les hommes contre l'humain. Paris: La Colombe, 1951. Translated by G. S. Fraser as *Men against Humanity*. London: Harvill Press, 1952. Republished as *Man against Mass Society*. Chicago: Regnery, 1962.

L'homme problématique. Paris: Aubier, 1955.

Présence et immortalité. Paris: Flammarion, 1959.

The Existential Background of Human Dignity. Cambridge, MA: Harvard University Press, 1963.

Plays

"The Lantern." *Cross Currents* 8 (2) (1958): 129–143.

Three Plays. Translated by Rosalind Heywood and Marjorie Gabain. New York: Hill and Wang, 1958; 2nd ed., 1965. Contains "A Man of God," "Ariadne" (*Le chemin de Crête*), and "The Funeral Pyre" (*La chapelle ardente*; in the 2nd ed., "The Votive Candle").

Translations

Creative Fidelity. Translated, with an introduction, by Robert Rosthal. New York: Farrar, Strauss, 1964.

Presence and Immortality. Translated by Michael A. Machado. Pittsburgh: Duquesne University Press, 1967.

Tragic Wisdom and Beyond. Translated by Stephen Jolin and Peter McCormick. Evanston, IL: Northwestern University Press, 1973.

WORKS ON MARCEL

Applebaum, David. *Contact and Alienation, The Anatomy of Gabriel Marcel's Metaphysical Method*. Washington, DC: Center for Advanced Research in Phenomenology & University Press of America, 1986.

Busch, Thomas. *The Participant Perspective: A Gabriel Marcel Reader*. Lanham, MD: University Press of America, 1987.

Cain, Seymour. *Gabriel Marcel*. London: Bowes and Bowes, 1963.

Gallagher, Kenneth T. *The Philosophy of Gabriel Marcel* New York: Fordham University Press, 1963.

Lapointe, Francis H., and Claire C. Lapointe. *Gabriel Marcel and His Critics: An International Bibliography (1928–1976)*. NewYork: Garland, 1977.

Pax, Clyde. *An Existentialist Approach to God: A Study of Gabriel Marcel* The Hague: Martinus Nijhoff, 1972.

Prini, Pietro. *Gabriel Marcel et le méthodologie de l'invérifiable*. Paris, 1953.

Ricoeur, Paul. *Gabriel Marcel et Karl Jaspers*. Paris: Éditions du Temps Present, 1947.

Schilpp, Paul A., and Lewis E. Hahn. *The Philosophy of Gabriel Marcel*. Library of Living Philosophers Vol. XVII. La Salle, IL: Open Court, 1984.

Troisfontaines, Roger. *De l'existence à l'être*. 2 vols. Paris, 1953. Essentially a concordance of what Marcel has said on any subject. Contains a complete bibliography of Marcel's work until 1953.

Samuel McMurray Keen (1967)
Bibliography updated by Thomas Nenon (2005)

MARCION

(c. 85–c. 159)

Marcion was one of the most significant and, in a way, perplexing figures of the second century CE—significant both for founding the Marcionite Church and for providing the stimulus for the formation of the New Testament canon, and perplexing because of the difficulty of classifying him among contemporary thinkers. He is often called a Gnostic, and there are certainly distinct affinities with Gnosticism in his cosmology and soteriology; but his lack of a mythical anthropology and of any syncretistic tendency sets him apart.

A native of Sinope in Pontus, he was born c. 85 and must have died c. 159, since there is no suggestion in our sources that he survived until the reign of the emperor Marcus Aurelius (161–180). According to the ecclesiastical writer Hippolytus, Marcion was the son of a bishop, and indeed there are indications that he grew up within the Christian faith. Excommunicated by his own father because of his unorthodox views, he traveled first to Asia Minor, then to Rome (c. 138–140), where he was at first closely associated with the church. In 144 he was again excommunicated, and he founded a church of his own that was for a time a serious menace to "orthodox" Christianity.

Marcion was a Bible critic and theologian rather than a philosopher; indeed, Adolf von Harnack describes him as "fundamentally a Biblicist and an opponent of all philosophy." The root of his teaching lies in the Pauline antithesis of Law and Gospel, but he exaggerated this contrast to the extent of distinguishing the Creator (the God of the Old Testament) from the true God, in himself unknown and alien to this world but manifested in the person of Jesus. This conception of the "alienness" of the true God Marcion shared with the Gnostics, but for him this concept developed from the study of the Scriptures rather than from philosophical speculation. Rejecting allegorical interpretation, he was unable to reconcile the Old Testament description of God with the New Testament portrayal of God as the father of Christ. Unlike the Gnostics as well as some of his followers, Marcion himself held that the Creator is not evil but merely just. Only the true God is good, a God of love. From this initial contrast the whole of Marcion's system follows naturally. This world, which is the work of the Creator, is imperfect. The Jewish law, and indeed all positive morality, is a means by which the Creator exercises control over humankind and is therefore to be rejected. Marcion's conclusions, however, led not to licentious antinomianism but to asceticism: Marriage and sexual intercourse, for example, were prohibited as devices for the continued procreation of subjects of the Creator. Salvation is deliverance from the world and its God and is effected at the price of Christ's blood, solely by God's grace and not because the redeemed were considered "akin" to the supreme good God, as the Gnostics believed.

The gospel brought by Jesus was misunderstood and falsified by the apostles: Only Paul had the truth of the matter. Marcion therefore rejected not only the Old Tes-

tament but also those parts of the New Testament that, according to him, were contaminated by Judaism. His canon consisted of ten letters of Paul, beginning with Galatians, and an expurgated Gospel of Luke. He also set out his teaching in his *Antitheses,* which was largely composed of contrasts between the two Gods. Marcion's works have not survived, and we are dependent on information provided by his opponents (especially Tertullian) His followers (especially Apelles) later modified his teachings so that they were in closer conformity with ordinary Gnosticism. Some of the "Gnostic" elements in his own theology have been attributed to the influence of the second-century Gnostic Cerdo.

Marcionism was at its height in the latter half of the second century. Thereafter it tended to decline in the West, and the remnants of Marcionite churches were often absorbed into Manichaeanism. In the East it had a longer history, surviving down to the fifth century or later.

See also Cosmology; Gnosticism; Harnack, Carl Gustav Adolf von; Mani and Manichaeism; Tertullian, Quintus Septimius Florens.

Bibliography

Blackman, Edwin C. *Marcion and His Influence.* London: S.P.C.K., 1948.

Harnack, Adolf von. *Marcion: Das Evangelium vom fremden Gott.* Leipzig: Hinrichs, 1921. A classic.

Jonas, H. *The Gnostic Religion.* New York, 1958. Pp. 137ff.

Knox, J. *Marcion and the New Testament.* Chicago: University of Chicago Press, 1942.

R. McL. Wilson (1967)

MARCUS, RUTH BARCAN

Ruth Barcan Marcus, though she has published in a number of areas, is best known for her groundbreaking papers in modal and philosophical logic. In 1946 she initiated the first systematic treatment of quantified modal logic (see Barcan, 1946), therein provoking W. V. Quine's decades-long attack upon the meaningfulness of quantification into alethic modal contexts. The ensuing dispute focused attention on the phenomenon of referential opacity and led to important developments in logic, metaphysics, and philosophy of language. In subsequent papers Marcus extended the first-order formalization to second order with identity (Barcan, 1947) and to modalized set theory (Marcus, 1963, 1974). Particularly significant theses presented in these works were the axiom $\Diamond(\exists x)Fx \rightarrow (\exists x)\Diamond Fx$, known as the Barcan formula (Barcan, 1946), and the proof of the necessity of identity (Barcan, 1947; Marcus, 1961). It is of some historical interest that Marcus introduced the now standard "box" operator for necessity.

Marcus's response to criticisms of quantified modal logic took many forms and was a theme to which she returned repeatedly throughout her career. In her 1961 paper (and elsewhere) she sought to dispel certain puzzles about substitutivity of identity in modal contexts; she was an early advocate of a substitutional interpretation of the quantifiers for certain purposes (Marcus, 1961, 1962, 1972), as for example in modal and fictional discourse; she maintained that quantification into modal contexts involves no commitment to an objectionable essentialism (Marcus, 1961), and she later developed and defended a version of Aristotelian essentialism within a modal framework (Marcus, 1967, 1976). Finally, in the mid-1980s she offered an explicit defense of the metaphysical actualism that had informed her early papers in modal logic (Marcus, 1985–1986). Here once again Marcus employed an objectual interpretation of the quantifiers, construing our core modal discourse as counterfactual discourse about actual objects.

Allied doctrines of enduring significance either originated or evolved in other writings by Marcus. For example, she introduced a flexible notion of extensionality whereby languages and theories are extensional to the extent that they identify relatively stronger equivalence relations with relatively weaker ones (Marcus, 1960, 1961). She also proposed that ordinary proper names are contentless directly referential tags (Marcus, 1961). In so doing, Marcus rejected earlier "descriptivist" accounts, often associated with Gottlob Frege and Bertrand Russell, and laid the cornerstone of the so-called new theory of direct reference later elaborated by Saul Kripke, Keith Donnellan, David Kaplan, and others.

Writing in moral theory, Marcus exposed defects in the structure of standard deontic logic (Marcus, 1966). She also argued that moral dilemmas are real and, moreover, that their reality is compatible with the consistency of the moral principles from which they derive (Marcus, 1980). Reasoning from a straightforward analogue of semantic consistency, she called into question familiar arguments from the existence of moral dilemmas to ethical antirealism. The resulting account also yielded some second-order principles of conflict avoidance.

Finally, in a series of papers on the nature of belief (Marcus, 1981, 1983, 1990), Marcus rejected language-centered theories according to which beliefs are attitudes

to linguistic or quasilinguistic entities (sentences of English or "Mentalese," for instance). Her proposal was that an agent *X* believes that *S* if and only if *X* is disposed to respond as if *S* obtains, where *S* is a possible state of affairs and what is to count as such a response is a function of environmental factors and internal states such as *X*'s needs and desires. This object-centered theory, as opposed to the language-centered views of Donald Davidson and Jerry Fodor, for example, more naturally accommodates unconscious beliefs and beliefs of infralinguals and nonlinguals. It also accommodates a more robust notion of rationality and explains, as its rivals cannot, why a fully rational agent would not believe a contradiction. In the wide sense of the term, a rational agent is one who, among other things, strives to maintain the global coherence of the behavioral—that is, verbal as well as nonverbal—indicators of his beliefs. Thus, although a rational agent might assent to a contradiction, his assent would not "go over" into a belief. Indeed, upon discovering the contradiction, he would retract his earlier (contradictory) belief claim. On Marcus's view, just as one cannot know what is false, one cannot believe what is impossible.

Marcus was professor of philosophy and chair of the department at the University of Illinois at Chicago from 1964 to 1970, professor of philosophy at Northwestern University from 1970 to 1973, and the Reuben Post Halleck Professor of Philosophy at Yale, where she succeeded her mentor Frederick B. Fitch, from 1973 to the time of her retirement in 1992. In addition to her scholarly achievements Marcus changed the face of the philosophical profession by her efforts on behalf of women. Perhaps most noteworthy in this connection was the reform of hiring practices instituted by the American Philosophical Association during her tenure as an officer and subsequently as chairman of its National Board of Officers.

See also Davidson, Donald; Fodor, Jerry A.; Frege, Gottlob; Kaplan, David; Kripke, Saul; Logic, History of; Metaethics; Metaphysics; Modal Logic; Philosophy of Language; Quine, Willard Van Orman; Rationality; Russell, Bertrand Arthur William; Set Theory.

Bibliography

WORKS BY MARCUS

Barcan, R. C. "A Functional Calculus of First Order Based on Strict Implication." *Journal of Symbolic Logic* 11 (1946): 1–16.

Barcan, R. C. "The Identity of Individuals in a Strict Functional Calculus of First Order." *Journal of Symbolic Logic* 12 (1947): 12–15.

Marcus, R. B. "Extensionality." *Mind* 69 (273) (1960): 55–62.

Marcus, R. B. "Modalities and Intensional Languages." *Synthese* 13 (1961): 303–322.

Marcus, R. B. "Interpreting Quantification." *Inquiry* 5 (1962): 252–259.

Marcus, R. B. "Classes and Attributes in Extended Modal Systems." *Proceedings of the Colloquium in Modal and Many Valued Logic, Acta philosophica fennica* 16 (1963): 123–136.

Marcus, R. B. "Iterated Deontic Modalities." *Mind* 75 (300) (1966): 580–582.

Marcus, R. B. "Essentialism in Modal Logic." *Noûs* 1 (1967): 91–96.

Marcus, R. B. "Essential Attribution." *Journal of Philosophy* 67 (1971): 187–202.

Marcus, R. B. "Quantification and Ontology." *Noûs* 6 (1972): 240–250.

Marcus, R. B. "Classes, Collections, and Individuals." *American Philosophical Quarterly* 11 (1974): 227–232.

Marcus, R. B. "Dispensing with Possibilia." Presidential address, *Proceedings of the American Philosophical Association* 49 (1976): 39–51.

Marcus, R. B. "Moral Dilemmas and Consistency." *Journal of Philosophy* 77 (3) (1980): 121–135.

Marcus, R. B. "A Proposed Solution to a Puzzle about Belief." In *The Foundations of Analytic Philosophy,* edited by P. French, T. Uehling, H. K. Wettstein. Midwest Studies in Philosophy 6. Minneapolis: University of Minnesota Press, 1981.

Marcus, R. B. "Rationality and Believing the Impossible." *Journal of Philosophy* 75 (1983): 321–337.

Marcus, R. B. "Possibilia and Possible Worlds." Edited by R. Haller. *Grazer philosophische Studien* 25–26 (1985–1986): 107–132.

Marcus, R. B. "Some Revisionary Proposals about Belief and Believing." *Philosophy and Phenomenological Research* (Supplement 1990): 133–154.

Marcus, R. B. *Modalities.* New York: Oxford University Press, 1993.

Diana Raffman (1996)
G. Schumm (1996)

MARCUS AURELIUS ANTONINUS

(121–180 CE)

Marcus Aurelius Antoninus may have wielded more political power than any other person to have an entry in this encyclopedia. Born into a prominent Roman family in 121 CE, Marcus was adopted in 138 by Emperor Hadrian's heir, Antoninus Pius (at Hadrian's behest), and he succeeded Antoninus as emperor in 161. Marcus's reign is usually judged favorably; indeed, his death in 180 is often thought to end the golden age of the Roman Empire. But it was not all wine and roses: Marcus faced troubles on the frontiers of the empire, a devastating

plague, and worst of all, persistent wars that included the first Germanic invasion of Italy in centuries, a harbinger of invasions to come.

By late antiquity, Marcus Aurelius was most famous as a philosopher. This reputation has come to rest on his Greek writings *to himself*, best known in English as the *Meditations*. In Book One, Marcus offers an idealized account of the influences on his character, acknowledging gods, family, and teachers, including several philosophers, a grammarian, and the rhetorician Fronto, part of whose correspondence with Marcus survives. The remaining eleven books manifest no obvious organization and have puzzled many scholars. Their 473 chapters vary considerably in length and style, from maxims to minitreatises, with consolations, dialogues, and harangues thrown in. These chapters commonly feature first- or second-person pronouns and the imperative mood, with the aim of recommendation or rebuke. Sometimes Marcus articulates more theoretical doctrines of philosophy or even, though much less often, arguments, but even at his most explicitly theoretical, he does not stray far from commending or censuring. Scholars have conjectured that the *Meditations* are the scraps of an intended treatise, but this does not fit well with the text. Scholars have also tried to rearrange the chapters to impose a clearer organization, but no such reorganization has commanded broad acceptance. Instead, most scholars now take the *Meditations* as they are. On the consensus view, although the whole collection is informed by philosophical reflection, Marcus writes not to theorize but to bring his thoughts, feelings, and activities in line with the philosophical commitments he accepts.

The *Meditations* are therefore not like usual philosophical writing, and this is what makes them historically significant. Philosophers in ancient Greece and Rome often encourage others to engage in meditative exercises to cultivate a philosophical way of life (especially relevant is Epictetus, Diss. I 1.25), and Marcus's work is the best example of such exercises. It suggests that one does not cultivate a philosophical way of life by the detailed application of philosophical theory to particular dilemmas. That is why Marcus's exercises do not shed much light on the particulars of his life. When he does make practical precepts explicit, he states them in general terms that could apply to a shopkeeper in Kansas as well as a Roman emperor and in terms that target attitudes more than actions. So it seems that one cultivates philosophy by bringing about a general outlook that one will then put into action as the circumstances demand.

The philosophical outlook that Marcus cultivates is generally thought to be Stoic though he does not call himself a Stoic. His praise for Epictetus and his use of Stoic vocabulary encourage this thought, but by no means decisively, since he also cites Plato and Epicurus favorably, and by his time philosophers of many schools used Stoic vocabulary. Still, some of the most prominent themes of the *Meditations* are genuinely Stoic: strong contrasts between the value of one's mind, a part of the divine intelligence, and what is external to one's mind and indifferent to one's happiness (II 13, III 12); concerted efforts to reduce anger at others and to control impulses (II 1, VII 22); and regular insistence that one should help other members of the human community (V 33, VIII 59, IX 1.1, IX 23). Less distinctively Stoic is the persistent theme of death (II 12, III 3, IV 5, IV 6, IV 48, VI 28; XII 36), though this is a natural obsession if the *Meditations* were written (as the evidence suggests) in the last decade of Marcus' life and some of them at military camps. So on balance the impression of Stoic commitments is hard to deny.

To call Marcus a Stoic, though, one must use an undemanding litmus. First, Marcus shows very weak adherence to two-thirds of the traditional Stoic system. He ignores the epistemology, language, and formal logic of the Stoic study of reason (or logic; logikē), and he belittles the need to study nature (that is, to engage physics; physikē). He occasionally helps himself to the Stoic thought that the cosmos is providentially ordered (II 3, X 6, XI 18.1), but he is detached enough from this thought that he also tries repeatedly to claim that the same practical precept applies whether the world is providential or, as Epicurean atomism holds, not (VII 32, VIII 17, IX 39, X 6, XII 24). In general, Marcus's philosophical commitments do not much outrun his ethic.

Even Marcus's ethical reflections are so untheoretical as to suggest a departure from traditional Stoicism. For example, Stoic ethics traditionally relies on the thought that virtuous activity alone constitutes a happy life, and Stoics support this thought either by describing the natural development of concern for virtuous activity alone (and the concomitant stripping away of obstacles to a smooth flow of life) or by engaging directly in the question of what happiness is. Marcus, though, does not motivate his Stoic aims theoretically. Presumably, he does not need to. If he already has these aims, he needs only to reshape his attitudes to improve his pursuit of them. In this way, Marcus's special purpose leads him to pass over many of the issues, distinctions, and arguments of traditional Stoicism.

The Stoicism of the *Meditations* clearly owes much to Epictetus, but in its ruthless pursuit of getting the cast of mind right without dallying in logic, physics, or the distinctions among the things that are neither good nor bad, it also resembles still more Cynicizing versions of Stoicism as that of Aristo of Chios, the renegade Stoic of the third century BCE. Marcus might have been especially influenced by Aristo's work or by the Cynic revival in imperial Rome. Or perhaps the appearance of such affinity is due to his special purpose in the *Meditations* of trying to recast his general practical attitudes.

This purpose might also explain another characteristic of Marcus's Stoicism. It is often said that Marcus shows strong Platonist leanings, especially in the starkly dualistic way in which he contrasts the intellect in the soul with the body (IV 41, VIII 37, IX 24, XII 33) and the matter of the external torrent (V 10.2, VI 15, VIII 24, IX 36). Sometimes these leanings are attributed specifically to the influence of the Platonizing Stoic Posidonius (c. 135–c. 50 BCE), and sometimes they are said to anticipate Neoplatonism. But Marcus's occasionally dualistic talk and his hostility toward the body might be understood instrumentally as part of a regimen to correct his excessive attachment to his own body and not as a commitment to any dualism.

The *Meditations* were apparently not in wide circulation for several centuries after Marcus's death, and so they exhibit no obvious influence on the immediately subsequent history of philosophy. In modern times, however, the work has been widely admired, sometimes for its fresh glimpse into ancient Stoicism but more often for its intimate picture of an aging emperor's struggle with noble yet human goals, to be a better person, and to face death without fear or regret.

See also Aristo of Chios; Cynics; Epictetus; Epistemology; Ethics; Neoplatonism; Platonism and the Platonic Tradition; Posidonius; Stoicism.

Bibliography

TEXT, TRANSLATION, AND COMMENTARY

Farquharson, A. S. L. *Meditations of the Emperor Marcus Antoninus*. 2 vols. Oxford: Clarendon Press, 1944.

CORRESPONDENCE

Hout, M. P. J. van den, ed. *M. Cornelii Frontonis Epistulae*. Vol. 1. Leiden: Brill, 1954

BIOGRAPHY

Birley, Anthony. *Marcus Aurelius: A Biography*. Revised ed. New Haven: Yale University Press, 1987.

STUDIES

Asmis, Elizabeth. "The Stoicism of Marcus Aurelius." *Aufstieg und Niedergang der römischen Welt*. Part II, Vol. 36.3. Berlin: de Gruyter, 1989, pp. 2228–2252.

Brunt, P.A. "Marcus Aurelius in his *Meditations*." *Journal of Roman Studies* 64 (1974): 1–20.

Hadot, Pierre. *The Inner Citadel: The Meditations of Marcus Aurelius*. Translated by Michael Chase. Cambridge, MA: Harvard University Press, 1998.

Rutherford, R. B. *The Meditations of Marcus Aurelius: A Study*. Oxford: Clarendon Press, 1989.

Eric Brown (2005)

MARÉCHAL, JOSEPH
(1878–1944)

Joseph Maréchal, one of the most original and influential of Neo-Scholastic thinkers, was born at Charleroi, Belgium. He entered the Society of Jesus at the age of seventeen, and between 1895 and 1910, in spite of poor health, he not only successfully completed the long and exacting Jesuit course of studies in the humanities, philosophy, theology, and asceticism but also obtained his doctorate in the natural sciences from the University of Louvain (1905). After the completion of his Jesuit training, during the latter part of which he also taught biology to his younger confreres, he spent some time in Germany studying experimental psychology and psychotherapy. From the outset his main interest centered on the psychology of religious experience and its implications for metaphysics and the critical problem.

After the outbreak of war in 1914 he went to England with his Jesuit students. He did not begin teaching formally at the Jesuit scholasticate in Louvain until 1919. From then until 1935 he conducted courses in psychology, theodicy, and the history of modern philosophy. It was during these years that he published his most important works, the two-volume *Études sur la psychologie des mystiques* and the First, Second, Third, and Fifth Cahiers of the *Point de départ de la métaphysique* (the first three are somewhat abridged in his *Précis d'histoire de la philosophie moderne*). The Fourth Cahier, *Le systéme idéaliste chez Kant et les postkantiens*, was published posthumously in 1947 from manuscripts left by the author.

After 1935 and until his death Maréchal ceased teaching and writing, mostly because of poor health but partly because he felt that his work was misunderstood and ineffectual. Concerning "my epistemology," he remarked, "I have never had the means of exposing, orally

or by writing, my general conception of the problem of knowledge. The Fifth Cahier states once more this problem *in terms of Kant*, which retains something artificial demanded by immediate historical antecedents. My definitive position ought to appear only at the end of the Sixth Cahier, in which there remains a new stage to overcome" (*Mélanges Maréchal*, Vol. I, p. 13; all translations are the author's). Unfortunately, the Sixth Cahier was never published.

In an article, "À propos du Sentiment de presence chez les profanes et chez les mystiques," published in 1908, the year he was ordained a priest, and later reproduced in the first volume of his *Études sur la psychologie des mystiques* (2nd ed., pp. 67–122), Maréchal for the first time indicated the distinctive trend of his philosophical thought. He pointed out that "the judgment of presence properly speaking affirms a spatial relation between a subject and an object," implying their reality, which is conditioned by "(1) a certain unity of mind, realized by (2) the coordination of representations, (3) with the concurrence of feeling" (*Études*, p. 110). Because the existential judgment cannot be founded solely on sensible experience, in view of sensible illusions, or on subjective feeling, the "psychologists" arbitrarily assume the anteriority of the subjective over objective knowledge, thus creating the pseudocritical problem of the "bridge" from thought to reality, the solution of which is thus prejudiced in favor of idealism. According to Maréchal the terms of the problem should be reversed. A more simple and more logical procedure would be "to posit as a primitive fact the *real, affirmation*, and the *objective* and to seek how this fact, in being broken up, gives birth to the secondary notions of the *unreal*, of *doubt*, and of the *subjective*. We shall thus rediscover, with a certain number of modern psychologists and under the impulse of experience, the point of view—very clear but insufficiently analyzed—of ancient Thomistic psychology" (ibid.).

Maréchal's principal work is his Fifth Cahier. The first four cahiers present a historical exposition and critical analysis of the problem of knowledge prior to Immanuel Kant, in Kant, and in post-Kantian transcendental idealism and a "historical demonstration" of the Thomistic solution. A twofold antinomy emerges, of the sensibility and understanding and of the understanding and metaphysical reason. Kant resolved the first antinomy by refuting the exaggerated claims of both the empiricists and the rationalists and by effecting a synthesis of the sensibility and understanding. However, according to Maréchal, Kant failed to resolve the second antinomy because he did not take into consideration the role of finality and intellectual dynamism in objective knowledge, a failure revealed in his *Opus Postumum* and in Johann Gottlieb Fichte's finalism. Maréchal held that Thomas Aquinas's epistemology virtually contains the solution of the antinomy of the understanding and reason by their effective synthesis in terms of intellectual dynamism (though Thomas himself did not explicitly consider the modern critical problem). Hence, the Fifth Cahier, "Thomisme devant la philosophic critique," presents the Thomistic solution of the critical problem without pretending to present an anachronistic confrontation of Kant and Thomas.

Maréchal agreed with Kant that we have no intellectual intuition of the noumenal, but he denied Kant's conclusion that the noumenal is therefore unknowable to human reason. Even though the human mind is not intuitive, but only abstractive and constructive, in its knowledge, yet in virtue of its innate active dynamism to Absolute Being it attains the noumenal or metaphysical in its synthetic elaboration of the object of knowledge by the "active intellect."

The Fifth Cahier has two main divisions. The first part is an examination, according to the demands of modern criticism, of "the theory of knowledge in the framework of Thomistic metaphysics," which Maréchal aptly termed "a *metaphysical* critique of the object"; it is preceded by a "critical preamble," in which the author explains Thomas's "universal doubt" and refutation of skepticism. The second part is "a Thomistic critique of knowledge transposed to the transcendental plane" and therefore "a *transcendental* critique of the object," an attempt to go beyond Kant on the basis of Kant's point of departure and transcendental method, which seeks the a priori conditions of the possibility of the objective contents of human consciousness, viewed precisely as objective.

How does Maréchal's metaphysical critique of the object differ from his transcendental one? Both have as their initial point of departure the object immanent in the mind, the mental content directly revealed in consciousness, what René Descartes called the "objective reality" of the idea. However, according to the metaphysical critique, the presence of the object in the mind is intentional and therefore ontological or noumenal in its signification, whereas according to the transcendental critique there is present to the mind only a phenomenon. From either viewpoint, however, there can be no question but that this immanent object presents (1) a sensible aspect, (2) a conceptual aspect (involving the notes of universality and necessity), and (3) a transcendent aspect inex-

orably pointing toward Absolute Being. Unlike Kant, scholastic Thomism accepts the objective validity of the third aspect. As we shall presently see, the two critical approaches differ not as regards their philosophical methods but only as regards their formal object. The formal object of the metaphysical critique is being, viewed as being in all its fullness, universality, and necessity—namely, Absolute Being or God; the formal object of the transcendental critique is the phenomenon.

This is not to say that the transcendental method, as understood in too narrow a sense by Kant himself, does not differ from the metaphysical method of Thomism. The transcendental method seeks to determine the a priori conditions of the possibility of the "objective" contents of consciousness. But as Maréchal contended, the most important and salient of these a priori conditions (which Kant failed to recognize) is the intellectual dynamism of the subject, its activity in constructing the immanent object. This is revealed by "transcendental reflection," whereas "transcendental deduction" proves that the object immanent in consciousness cannot be truly "objective" except in terms of this a priori or objectivizing function of the dynamic intellect, whose formal object is Absolute Being. Needless to say, Kant himself never conceived the transcendental method in such a dynamic fashion. Thus, the most basic inconsistency of his methodology, according to Maréchal, is his stated purpose of disclosing by transcendental reflection the purely logical and static a priori conditions of knowledge, whereas, inadvertently or not, his procedure is often psychological and dynamic; he viewed the mind as constructive and synthetic, and therefore as active, but illogically concluded that the only a priori discoverable by transcendental reflection is purely logical, formal, and static. Hence, Maréchal refuted Kant in the first part of Cahier V by applying the transcendental method to the ontological object, thus legitimizing the Thomistic point of departure of metaphysics (namely, that the human mind directly attains the noumenal or intelligible in its necessary judgments), while in the second part he attempted to go beyond Kant's agnostic conclusions by proving the necessity of metaphysics, using this same transcendental method and basing the proof on Kant's own presupposition that the object immanent in consciousness is the phenomenal.

To constitute a noumenal "object in itself," that which is known must be something more than an abstract essence or form in the mind; it must go beyond the domain of *form* and be related to the sphere of *act*. An abstract essence can become a possible essence and therefore represent a real essence only when the immanent form becomes an act of the dynamism of the intellect, necessarily relating the abstract form to Absolute Being, as a partial fulfillment of this dynamism.

Maréchal was not maintaining "the ontological parologism" that the proposition "Truth is" is intuitive or analytical; rather, he held that what the discursive and abstractive intellect apprehends is that the connection between truth and being must be affirmed under pain of contradiction, when our intellectual dynamism to Absolute Truth is also apprehended. (The objective validity of our abstractive knowledge is thus assured.) Only the divine intellect is intuitive, but an abstractive intellect is capable of apprehending and reducing an abstracted form, inherent in the potentially intelligible data of sense, to act by virtue of its active dynamical tendency to Pure Act, thus approximating the perfection of the exemplary divine knowledge. Since our intellectual knowledge is not a purely passive reception of abstract forms, the self-consciousness of the synthesizing knowing subject as an intellectual dynamism is the key to Maréchal's doctrine on the objectivization of human knowledge.

Maréchal's distinction between the human intellect viewed as formally cognoscitive and the same intellect viewed as a natural being or entelechy (*ut res quaedam naturae*) is very important for an understanding of his epistemology of objectivization. The strictly intentional function of the abstractive intellect, whose formal object is being as such, must be basically identified with the entitative function of the same intellect viewed as a dynamic real tendency to Absolute Being or Truth. It is only in virtue of the intellect viewed as dynamic act that the formally cognoscitive and abstractive intellect can assimilate a representative form as objective being, that is, as a partial fulfillment of the intellect's natural dynamism to the acquisition of *all being*, the intuition of Being Itself.

Granted the sensible data, it is in the formation of the concept that the synthesizing function of the knowing subject reveals itself. Thus, metaphysical concepts present themselves in our consciousness as universal and necessary and therefore as connoting a relation to Absolute Being; though they may conceptually represent a multiplicity, they necessarily signify a universal, though analogical, unity of being that is intelligible only in terms of Absolute Being. How are we to explain these elements of universality and necessity?

In a Thomistic metaphysical critique of the object, the a priori is not simply a logical function, as in Kant. Rather, it designates, in terms of Maréchal's intellectual dynamism, an a priori that is at once both metaphysical

and psychological; for Maréchal the formal object of the intellect as a natural entelechy, or *res quaedam naturae*, is Absolute Being. On the conscious, elicitive, and formally cognoscitive level, being is necessarily presented as an abstract being as such, but such a representation, Maréchal contended, is possible only because the intellect naturally tends to Absolute Being as its natural entelechy or end on the preconscious and preelicitive level. The substantial unity of the knowing subject makes possible the "conversion to the phantasm," without which it could not make a judgment concerning the concrete individual.

Maréchal's transcendental critique of knowledge can be more readily understood when it is viewed in the light of his posthumously published Fourth Cahier, especially his remarks on Kant's *Opus Postumum* (pp. 225–326) and on Fichte's "Intellectual Intuition of Act or Dynamic Intuition" (pp. 348ff.) and his article "L'aspect dynamique de la méthode transcendentale chez Kant" (*Revue Néoscholastique* 42 [1939]: 341–384). In his analysis of Kant's *Opus Postumum* ("The Passage from the First Foundations of the Metaphysic of Nature to Physics")—which Kant once called his "masterpiece" but which was first published in 1920 by Erich Adickes under the title *Kants Opus Postumum, dargestellt und beurteilt*—Maréchal pointed out that Kant acknowledged that the "form" involved in human knowledge is not merely static or logical but dynamic and real in its implication. This same idea of intellectual dynamism is emphasized by Maréchal's analysis of Fichte's development of Kantianism, so much so that Maréchal has been accused of being too Fichtean and voluntaristic in his application of the Kantian transcendental method to the problem of knowledge. For Fichte, as for Maréchal, the self-reflecting self, the immediate intuition of the self as "a primary fact of consciousness … is the sole solid foundation of all philosophy" (Fourth Cahier, p. 349).

See also Descartes, René; Fichte, Johann Gottlieb; Idealism; Kant, Immanuel; Neo-Kantianism; Scotism; Thomas Aquinas, St.; Thomism.

Bibliography

WORKS BY MARÉCHAL

Le point de départ de la métaphysique. Leçons sur le développement historique et théorique du problème de la connaissance. 5 vols. Vols. I, II, III, Brugge and Paris, 1922–1923; Vol. IV, Brussels, 1947; Vol. V, Louvain and Paris, 1926.

Études sur la psychologic des mystiques. 2 vols. Vol. I, Brugge and Paris, 1924; Vol. II, Brussels, 1937. Translated in great part by Algar Thorold as *Studies in the Psychology of the Mystics*. London: Burns, Oates and Washburne, 1927.

Précis d'histoire de la philosophic moderne, Vol. I: *De la renaissance à Kant*. Louvain, 1933.

Mélanges Maréchal, Vol. I: *Oeuvres*. Brussels: Édition Universelle, 1950. Collected articles, with bibliography.

A Maréchal Reader. New York: Herder and Herder, 1970.

George Carver: Savant noir, ancien esclave. 1973.

Proust, prince des humoristes. Paris: Panthéon, 1996.

L'école de papa. Lyon: Editions Bellier, 1998.

Victor Hugo: Dictionnaire de ses idées, ses jugements, seshumeurs. Lyon: Bellier, 1999.

L'enseignement selon les anarchistes. Lyon: Bellier, 2000.

WORKS ON MARÉCHAL

Casula, M. *Maréchal e Kant*. Rome: Fratelli Bocca, 1955.

Gilson, Étienne. *Réalisme thomiste et critique de la connaissance*, 130–155. Paris: J. Vrin, 1939.

Hayen, A. "Un interpréte thomiste du kantisme: Joseph Maréchal." *Revue internationale de philosophie* (1954): 449–469.

Mélanges Maréchal. Vol. II (Paris, 1950), contains additional studies.

James I. Conway, S.J. (1967)
Bibliography updated by Michael J. Farmer (2005)

MARIANA, JUAN DE
(1535–1624)

Juan de Mariana, the neo-Scholastic political philosopher, was born at Talavera de la Reina, Spain, and died at Toledo. Entering the Society of Jesus at eighteen, he completed the Jesuit course of studies in philosophy and theology and taught theology in Rome from 1561 to 1569 and at Paris from 1569 to 1574. He then retired to Toledo to work on his "History" and other writings in practical philosophy. Mariana's *Historiae de Rebus Hispaniae* (Toledo, 1952; also published in elegant Spanish by the author, Toledo, 1601) was one of the first general histories of Spain. Also influential were his treatises *De Rege et Regis Institutione* (Toledo, 1599, translated by G. A. Moore as *The King and the Education of the King,* Washington, DC, 1948) and *De Mutatione Monetae* (On Changing the Value of Money), one of the *Tractatus Septem* (Cologne, 1609).

Accused of attacking the sovereign power of Spain in his criticism of its fiscal policies, Mariana was tried in 1609 by the Spanish Inquisition and acquitted. His philosophy is important for its handling of political, social, and economic problems. A strong advocate of the power of the people, Mariana argued that the citizens as a whole (*communitas civium*) are superior in power to the monarch. Men lived originally in an unorganized "state of

nature," not needing political institutions to maintain justice; all possessions were held in common, and men naturally cooperated for their common welfare (*De Rege,* Chs. 8 and 13). With advances in arts and sciences, a division of goods developed into private possession; thus arose jealousy, pride, and strife among men. Tired of the struggle for domination, men then made a pact, delegating the ruling power to certain leaders. (Note that Mariana antedates both Thomas Hobbes and Jean-Jacques Rousseau.) The basic enactments of law can be changed only by the manifest will of the people. If the king fails to rule in accord with the law, he may be deposed by the people using prudent judgment; physical force may be employed for this purpose. Mariana was accused of trying to justify tyrannicide; his views did not endear him to the Spanish monarchists.

See also Hobbes, Thomas; Political Philosophy, History of; Rousseau, Jean-Jacques; Scotism; Thomism.

Bibliography

Laures, J. *The Political Economy of Juan de Mariana.* New York: Fordham University Press, 1928. Contains Latin text of *De Mutatione Monetae.*

Lewy, Guenter. *Constitutionalism and Statecraft during the Golden Age of Spain: A Study of the Political Philosophy of Juan de Mariana, S.J.* Geneva: Droz, 1960.

Tallmadge, G. K. "Juan de Mariana." In *Jesuit Thinkers of the Renaissance,* edited by Gerard Smith, 157–192. Milwaukee: Marquette University Press, 1939.

Ullastres Calvo, A. "La teoria de la mutación monetaria del Mariana." *Anales de economia* (15) (1944): 273–304; (20) (1945): 437–471.

Vernon J. Bourke (1967)
Bibliography updated by Philip Reed (2005)

MARÍAS, JULIÁN
(1914–)

Julián Marías is the best-known and most productive of the post–Civil War philosophers in Spain who have sought to reconcile the doctrines of their teacher, José Ortega y Gasset, with traditional theism. Born in Valladolid in 1914, Marías studied under Ortega in Madrid just before the Civil War. When Ortega returned from exile in 1948, they jointly founded the Institute of Humanities in Madrid. Marías has taught at the institute and, as visiting professor, at various American universities. The bulk of his published work concerns the history of philosophy, mainly Spanish and scholastic philosophy. His general *Historia de la filosofía* (1941), which he wrote at the age of twenty-six, emphasizes the Aristotelian and scholastic traditions and gives a prominent position to Spanish thought. In *La escuela de Madrid* (The Madrid school; Buenos Aires, 1959), Marías presented the most comprehensive study available of such contemporary Spanish thinkers as Ortega, Miguel de Unamuno, Xavier Zubiri, and Manuel García Morente.

As a Catholic disciple of Ortega, who was explicitly irreligious and anti-Catholic, Marías gave a theistic interpretation of Ortega's "ratiovitalism" (a reconciliation of rationalism and the vitalist doctrines of the 1920s). In his major work, *Introducción a la filosofía* (1947), Marías argued that certain intellectual and spiritual "ultimates" are true biological needs of humankind. To be lived at all humanly, life requires, in addition to food and other animal necessities, "the possession of a radical and decisive certitude." That certitude serves as the foundation for numerous "partial truths." It harmonizes all our beliefs into a single clear perspective, and it also provides society with a ruling view that is needed for social stability. Men turn to philosophy for this certitude, so there is nothing more "practical," vital, or socially relevant than metaphysics, which is called upon to give men a standard to live by.

Marías accepts all the pragmatist, relativist, and historicist implications of vitalism, which usually have been regarded as destructive of religious convictions, and he argues from them back to the traditional religious outlook. Truth is what answers a vital need by removing the feeling of insecurity and perplexity. It is always relative to particular life situations and historical periods. Truth fragments into a multitude of relative truths, which contain concrete concepts as distinct from general concepts, which are obtained only by an arbitrary and schematizing process of abstraction. Yet, if the quest for completely satisfying, radical certainty is pressed tenaciously enough, it will lead beyond this complete nominalism to God, who appears as the ground or basis of being. Although the ego that carries on that quest was, for Ortega, the incarnation of "vital reason," for Marías it is the person who owns both vitality and reason. At death, that person, or soul, loses vitality and psychic activity but does not necessarily cease to exist. The mortality of the soul is a theory that remains in need of proof.

See also Ortega y Gasset, José; Rationalism; Unamuno y Jugo, Miguel de; Vitalism.

Bibliography

Works by Marías are *Obras* (Works), 10 vols. (Madrid: Revista de Occidente, 1959–1982), *Historia de la filosofía* (Madrid, 1941); and *Introdución a la filosofía* (Madrid, 1947), translated by Kenneth Reid and Edward Sarmiento as *Reason and Life* (New Haven, CT, 1956).

On Marías, see Alain Guy, "Julián Marías," in *Philosophes espagnols d'hier et d'aujourd'hui* (Toulouse, 1956), pp. 330–339.

OTHER RECOMMENDED WORKS BY MARÍAS

Meditaciones sobre la sociedad española. Madrid: Alianza Editorial, 1966.

Miguel de Unamuno. Cambridge, MA: Harvard University Press, 1966.

History of Philosophy. New York: Dover, 1967.

Nuevos ensayos de filosofía. Madrid: Ediciones de la Revista de Occidente, 1968.

Obras. Madrid: Revista de Occidente, 1969–1982.

Generations: A Historical Method. University: University of Alabama Press, 1970.

José Ortega y Gasset, Circumstance and Vocation. Norman: University of Oklahoma Press, 1970.

Metaphysical Anthropology; the Empirical Structure of Human Life. University Park: Pennsylvania State University Press, 1971.

Philosophy as Dramatic Theory. University Park: Pennsylvania State University Press, 1971.

La justicia social y otras justicias. Madrid: Seminarios y Ediciones, 1974.

A Biography of Philosophy. University: University of Alabama Press, 1984.

The Structure of Society. Tuscaloosa: University of Alabama Press, 1987.

Understanding Spain. Ann Arbor: University of Michigan Press; San Juan: Editorial de la Universidad de Puerto Rico, 1990.

La felicidad humana. Madrid: Alianza Editorial, 1995.

The Christian Perspective. Houston, TX: Halcyon Press, 2000.

Ortega y Gasset, José. *Meditations on Quixote*. Introduction and notes by Julián Marías, Urbana: University of Illinois Press, 2000, 1961.

Neil McInnes (1967)
Bibliography updated by Michael J. Farmer (2005)

MARITAIN, JACQUES

(1882–1973)

The French philosopher Jacques Maritain was a powerful force in twentieth-century philosophy and cultural life. The author of more than fifty philosophical works and of countless articles that appeared in the leading philosophical journals of the world, he was widely regarded as a preeminent interpreter of the thought of Thomas Aquinas and as a highly creative thinker in his own right.

Maritain, born in Paris, was reared in an atmosphere of liberal Protestantism. He attended the Sorbonne, where he fell briefly under the spell of teachers passionately convinced that science alone could provide all the answers to the questions that torment the human mind. It was at the Sorbonne that he met his wife-to-be, Raïssa Oumansoff, a young Russian-Jewish student who was to share his quest for truth and to become an intellectual and poet of real stature in her own right. She was also to collaborate with Maritain on a number of books. Soon disillusioned with the scientism of their Sorbonne masters, the two attended the lectures of Henri Bergson at the Collège de France. Bergson liberated in them "the sense of the absolute," and, following their marriage in 1904, they were converted (1906) to the Roman Catholic faith through the influence of Léon Bloy.

The years 1907 and 1908 were spent in Heidelberg, where Maritain studied biology under Hans Driesch. He was particularly interested at the time in Driesch's embryogenetic theory of neovitalism, a theory then little known in France. Upon returning to Paris, Maritain undertook the task of directing the compilation of a *Dictionary of Practical Life.* During the three years that he worked on this project, he also undertook a serious study of the writings of Thomas Aquinas. In 1914, he was appointed to the chair of modern philosophy at the Institut catholique de Paris.

From 1945 to 1948 Maritain was French ambassador to the Vatican. Afterward he taught at Princeton University until his retirement in 1956. He has also taught at the Pontifical Institute of Mediaeval Studies in Toronto, Columbia University, the Committee on Social Thought at the University of Chicago, and the University of Notre Dame. The Jacques Maritain Center was established at Notre Dame in 1958 for the purpose of encouraging research along the lines of his philosophy.

Maritain's thought is based on the principles of Aristotle and Thomas Aquinas but incorporates many insights found in other philosophers, both classical and modern, and also profits greatly from data supplied by such sciences of man as anthropology, sociology, and psychology.

THEORY OF KNOWLEDGE

The cardinal point in Maritain's theory of knowledge is his defense and critical elucidation of different ways of knowing reality. On the one hand, Maritain sees the richness and inexhaustibility of material reality as requiring that the mind let fall on it different noetic glances, each of which reveals to the mind a different universe of intelligi-

bility to be explored. There is, first of all, the universe of *mobile being*—being imbued with mutability—which constitutes the sphere of the knowledge of nature and which itself calls for both an empiriological analysis, that is, a spatiotemporal analysis oriented toward the observable and measurable as such (science of nature), and an ontological analysis, that is, an analysis oriented toward intelligible being, toward the very being and intelligible structure of things (philosophy of nature). There is, second, the universe of *quantity* as such, which constitutes much of the sphere of mathematics. And there is, finally, the universe of *being as being,* which constitutes the sphere of metaphysics.

Much of Maritain's energy was devoted to giving the philosophy of nature its epistemological charter, in contrast with many Thomists in a hurry who would have it almost totally eclipsed by metaphysics, and in contrast with the many scientists who think that the only object capable of giving rise to an exact and demonstrable science is that which is sense-perceivable and can be subjected to methods of experimental and mathematical analysis. Maritain's serious study of the work of modern physicists and biologists revealed to him that scientists are led by their science itself to discover within the mysterious universe of nature problems that go beyond the experimental and mathematical analysis of sensory phenomena. It also revealed to him that the conceptual lexicon of the scientist is radically different from the conceptual lexicon of the philosopher. For these reasons, Maritain emphasized the need for, and prerogatives of, both an ontological analysis and an empiriological analysis of the sensible real. He also worked out a theory of physicomathematical knowledge that relates this knowledge to what the Scholastics called intermediary sciences (*scientiae mediae*), sciences which straddle the physical order and the mathematical order and which have more affinity with mathematics than with physics as to their rule of explanation and yet at the same time are more physical than mathematical as to the terminus in which their judgments are verified.

On the other hand, Maritain saw the human mind as having another life than that of its conscious logical tools and manifestations: "there is not only logical reason but also, and prior to it, intuitive reason." There is indeed not only the Freudian unconscious of instincts, tendencies, complexes, repressed images and desires, and traumatic memories; there is also a spiritual unconscious or preconscious, the preconscious of the spirit in its living springs. The acts and fruits of human consciousness and the clear perceptions of the mind—in other words, the universe of concepts, logical connections, rational discursus, and rational deliberation—emerge in the last analysis from the hidden workings of this preconscious life of the spirit; but there also emerge from them many genuine knowings, and many affective movements, which remain more or less *sur le rebord de l'inconscient,* as Bergson would have said—on the edge of the unconscious. Among such knowings we have the various kinds of knowledge by inclination (knowledge through connaturality)—notably, poetic knowledge, the "natural" or prephilosophical knowledge of moral values, and mystical experience. Maritain felt it to be most incumbent upon us to recognize not only the different kinds or degrees of conceptual and discursive knowledge but also these different nonconceptual and "immediate" forms of knowledge.

METAPHYSICS

Maritain held the classical view that the object of metaphysics is *being as being,* and he stressed that it is in things themselves that metaphysics finds this object. It is the being of sensible and material things, the being of the world of experience, which is the immediately accessible field of investigation for metaphysics; it is this which, before seeking its cause, metaphysics discerns and scrutinizes—not as sensible and material but as being. Before rising to what may be a realm of spiritual existents, metaphysics must grasp empirical existence, the existence of material things—not as empirical and material but as existence.

For Maritain, at the starting point of metaphysics there lies an intuition, the "metaphysical intuition of being," which may be said to consist in the intellect's seeing—through an abstractive or eidetic (idea-producing) visualization—the intelligible value *being,* being in itself and in its essential properties. The word *intuition* here has caused much difficulty for some philosophers, but it seems to be demanded by the thought that Maritain was trying to express. What must somehow be preserved is, on the one hand, that it is as true to say that this "seeing" produces itself through the medium of the vital action of our intellect—of our intellect as vitally receptive and contemplative—as to say that we produce it; and, on the other hand, that it is being more than anything else that produces this "seeing."

In his scrutiny of the being of sensible and material things, Maritain presented a highly original treatment of what Thomists and others have long considered to be the first principles of speculative reason—the principles of identity, sufficient reason, finality, and causality. He

explained that the reality that is the object of the idea of being is richer than this idea, and it presses for multiplication in a manifold of notions, among them the notions of unity, of goodness, of truth: being is one, is good, is true. Each of these notions expresses to the mind nothing but being itself, to which it adds nothing but a conceptual difference. But precisely in virtue of this ideal element that differs from one to the other, these notions as such are different among themselves and are different from the notion of being; they are convertible notions but they are not identical with one another. There is thus a superabundance of being with regard to the notions in which it is objectified, and it is in terms of this superabundance that Maritain elucidated the intuitivity of the first principles.

When he turned his philosophical gaze to the problem of the "cause of being," Maritain was attentive both to specifically philosophical ways of establishing the existence of God and to nonphilosophical or prephilosophical ways of approaching God. Under the first heading he restates the five classical ways of Thomas Aquinas, divesting them of the examples borrowed from ancient physics and formulating them in a language more appropriate to modern times; then he proposes a "sixth way." In this "sixth way" we have first the complex primordial intuition, and later the rational and philosophical reflection, that the *I* who thinks, the *I* who is caught up in pure acts of intellect, cannot ever not have been, for both the intellect and the intelligible as such are above time: this *I* must always have existed, and in some personal existence, too, although not within the limits of its own personal being but rather in some transcendent and suprapersonal Being. Philosophical reflection can go on to establish how the *I* always existed in God, can establish that "the creature which is now I and which thinks, existed before itself eternally in God—not as exercising in Him the act of thinking, but as thought by Him."

But Maritain was quick to recognize prephilosophical approaches to God—the "natural," or instinctive and intuitive, approach proper to the first apperceptions of the human intellect, the approach through art and poetry, and the approach through moral experience. The inner dynamism of a man's first awakening to the intelligible value of existence causes him to see that the Being-with-nothingness that is both his own being and the being of the universal whole must be preceded by transcendent Being-without-nothingness. As concerns art and poetry, the poet or artist, in following the very line of his art, tends without knowing it to pass beyond his art; just as a plant, although lacking knowledge, directs its stem toward the sun, the artist, however sordid his life, is oriented toward the primary source of beauty. And finally, as concerns moral experience, when a man experiences, in a primary act of freedom, the impact of the moral good, and is thus awakened to moral existence and directs his life toward the good for the sake of the good, then he directs his life, without knowing it, toward the absolute Good. In this way he knows God vitally, by virtue of the inner dynamism of his choice of the good, even if he does not know God in any conscious fashion or through any conceptual knowledge.

MORAL PHILOSOPHY

One of the most provocative sides to Maritain's thought was his theory of "moral philosophy adequately taken." His contention was that moral philosophy—however vast, necessary, and fundamental be the part that natural ethics plays in it—must, if it is to be adequate to its object (the direction or regulation of human acts), take into account the data of revelation and theology concerning the existential state of man. Human conduct is the conduct of an existent, not simply the conduct of a nature. Consequently, the moral philosopher must take into account all data that contribute to make the existential condition of man genuinely known to us. He must take into account the data of ethnology, sociology, and psychology. And he must also take into account theological data. For, in fact, as a result of the present state of human nature, man has more propensity to evil than the man of pure nature by reason of the original sin and of the concupiscence that remains even in the just; and, on the other hand, he has incomparably stronger weapons for good, by reason of divine grace. Maritain recognized that the moral philosopher who does take this situation into account will not be a *pure* philosopher but maintained that he will still be able to use the method proper to philosophy and advance with steps, so to speak, of philosophy, not of theology.

Maritain's theory of natural law was elaborated against the background of anthropological data. He held that two basic elements must be recognized in natural law: the *ontological* and the *gnoseological*; and it is perhaps in considering the second of these two that Maritain made his most fecund insights. The chief point he wished to emphasize is that the genuine concept of natural law is the concept of a law that is natural not only in the sense that it is the normality of functioning of human nature or essence but also in the sense that it is naturally known, that is, known through inclination or through connaturality, not through conceptual knowledge and by way of

reasoning. The inclinations in question, even if they deal with animal instincts, are essentially human and, therefore, reason-permeated inclinations; they are inclinations refracted through the crystal of reason in its unconscious or preconscious life. And since man is a historical animal, these essential inclinations of human nature either developed or were released in the course of time; as a result, man's knowledge of natural law developed progressively and continues to develop. Thus, the fact that there is considerable relativity and variability in the particular rules, customs, and standards of different peoples is in no way an argument against natural law.

It belongs, of course, to moral philosophy to provide a scientific justification of moral values by a demonstrative determination of what is consonant with reason and of the proper finalities of the human essence and of human society.

SOCIAL AND POLITICAL PHILOSOPHY

Much of Maritain's effort was directed to working out the character of authentically Christian politics. He lays primary emphasis on man as being both an *individual* and a *person*—an individual by reason of that in him which derives from matter, and a person by reason of that in him which derives from his subsisting spirit. Man must live in society both because of his indigence as an individual and because of his abundance or root generosity as a person. As an individual, man is only a part, and as such he bears the same relation to society as the part bears to the whole. His private good as an individual is in everything inferior to the common good of the whole, so that an individual may even be required to risk his life for the sake of the good of the community. But as a person, man is a whole; and the whole that the person is surpasses the whole that society is, because the person, by reason of the subsistence of his spiritual soul, is destined for eternal union with the transcendent Whole, whereas the particular society in which the person lives, by reason of its not having a spiritual soul, is not destined for union with the transcendent Whole, but will die in time. Man is above and superior to political society, and the political community must recognize the person's orientation to an end above time and facilitate his attainment of it.

Maritain's social and political philosophy also manifested a keen sense of history. For Maritain as for Pindar, man must become what he is—man must "win his being"; man must become, in the psychological and moral order, in the social and political order, the person he is in the ontological order. Among the many truths related to this fundamental exigency of man's being is one that Maritain sees as of absolutely essential importance—the fact that human history is made up of periods, each of which is possessed of a particular intelligible structure, and therefore of particular basic requirements.

It is Maritain's contention that the historical climate of the modern world is quite different from that of the medieval world. For him, medieval civilization was a sacral civilization, by which he means that the historical ideal of the Middle Ages was principally controlled by two dominants: On the one hand, the idea or myth of fortitude in the service of God—the lofty aim was to build up a fortress for God on earth—and on the other hand, the concrete fact that temporal civilization had a largely ministerial role as regards the spiritual—the body politic was to a large extent a function of the sacred and imperiously demanded unity of religion. In contrast, modern civilization was for Maritain a secular civilization, by which he meant that the historical ideal of modern times is largely controlled by two other dominants: On the one hand, the idea or myth of the body politic as being by nature something of the natural order and something directly concerned, therefore, only with the temporal life of men and their temporal common good; and on the other hand, the concrete fact that in pursuing this temporal common good, modern man is most intent on the attainment of freedom and the realization of human dignity in social and political life itself.

Against the background of this view of medieval and modern civilizations, Maritain reflected at length on the nature of the democratic ideal. He saw democracy as the only way of bringing about a moral rationalization of politics, and he insisted that in order to accomplish this task democracy needs the quickening ferment of Gospel inspiration. But he also insisted, no less forcefully, that the "creed of freedom" that lies at the very basis of democracy is not a religious, but rather a civic or secular, one. Furthermore, this secular creed deals with practical tenets that depend basically on simple, "natural" apperceptions of which the human heart becomes capable with the progress of moral conscience and which can be similarly adhered to by minds that may differ greatly as to the speculative and theoretical justifications. In keeping with such a conception, Maritain repeatedly asserted that men belonging to very different philosophical or religious lineages can and should cooperate in the pursuit of the common good of political life. He also maintained that the supreme principles governing the relationship between church and state should be applied less in terms of the social power than in terms of the vivifying inspiration of the church: "the superior dignity of the Church is to find

its ways of realization in the full exercise of her *superior strength of all pervading inspiration.*" This reflects a most basic premise in all of Maritain's thought: that immutable principles admit of, and even call for, analogical applications in different existential situations.

PHILOSOPHY OF ART

From his earliest years Maritain was the friend and confidant of numerous artists, writers, poets, and musicians, and he was considered by many as having the finest aesthetic sensibility among the major figures of modern philosophy. His long reflection on almost every facet of the artistic process culminated in his monumental *Creative Intuition in Art and Poetry,* which grew out of six lectures given in 1952 at the National Gallery of Art, Washington, where he had been invited to deliver the initial series of the A. W. Mellon Lectures in the Fine Arts.

Maritain held, like Dante Alighieri, that human art continues in its own way the labor of divine creation. But he kept reminding the modern artist that human art cannot create out of nothing; it must first nourish itself on things, which it transforms in order to make a form divined in them shine on a bit of matter. Maritain would admit that the widespread effort toward "pure art" in the latter part of the nineteenth century may have been a beneficent phase after the exasperation of sensibility provoked by impressionism, but he affirmed that in the last analysis human art is doomed to sterility and failure if it cuts itself off from the existential world of nature and the universe of man.

The deepest concern of Maritain was with the nature of poetic knowledge and poetic intuition, that is, with the nature of the knowledge immanent in and consubstantial with poetry, poetry as distinct from art and quickening all the arts. He held that poetic knowledge is a typical instance of knowledge through connaturality. Poetic knowledge, as he saw it, is nonconceptual and nonrational knowledge; it is born in the preconscious life of the intellect, and it is essentially "an obscure revelation both of the subjectivity of the poet and of some flash of reality coming together out of sleep in one single awakening." This unconceptualizable knowledge comes about, Maritain maintained, through the instrumentality of emotion, which, received in the preconscious life of the intellect, becomes intentional and intuitive, and causes the intellect obscurely to grasp some existential reality as *one* with the self (of the knower) reality has moved; and at the same time the knower grasps all that which this reality calls forth in the manner of a sign. In this way the self is known in the experience of the world and the world is known in the experience of the self, through an intuition that essentially tends toward utterance and creation. Thus, in such a knowledge it is the object created—the poem, the painting, the symphony—in its own existence as a world of its own that plays the part played in ordinary knowledge by the concepts and judgments produced within the mind.

Poetic knowledge, then, is not directed toward essences, for essences are disengaged from concrete reality in a concept, a universal idea, and are an object for speculative knowledge. Poetic intuition is directed toward concrete existence as connatural to the soul pierced by a given emotion. In a passage of great beauty Maritain wrote:

> This transient motion of a beloved hand—it exists an instant, and will disappear forever, and only in the memory of angels will it be preserved, above time. Poetic intuition catches it in passing, in a faint attempt to immortalize it in time. But poetic intuition does not stop at this given existent; it goes beyond, and infinitely beyond. Precisely because it has no conceptualized object, it tends and extends to the infinite, it tends toward all the reality, the infinite reality which is engaged in any singular existing thing. (*Creative Intuition in Art and Poetry,* p. 126)

Maritain was admired even by those who may be of very different philosophical convictions. He was admired not only for his lifelong zeal for truth and impassioned commitment to freedom but also for his exceptional qualities as a person—his humility, his charity, his fraternal attitude toward all that is. He came to be recognized as one of the great *spirituels* of his time.

See also Aesthetic Judgment; Aristotle; Being; Bergson, Henri; Dante Alighieri; Driesch, Hans Adolf Eduard; Epistemology; Epistemology, History of; Ethics, History of; Metaphysics; Poetizing and Thinking; Social and Political Philosophy; Thomas Aquinas, St.

Bibliography

WORKS BY MARITAIN

Theory of Knowledge

La philosophie bergsonienne. Paris: Rivière, 1914; 3rd ed., Paris, 1948. Translated by Mabelle L. Andison and J. Gordon Andison as *Bergsonian Philosophy and Thomism.* New York: Philosophical Library, 1955.

Réflexions sur l'intelligence et sur sa vie propre. Paris: Nouvelle Librairie Nationale, 1924.

Le docteur angélique. Paris: Desclée De Brouwer, 1930. Translated by Joseph W. Evans and Peter O'Reilly as *St. Thomas Aquinas.* New York, 1958.

Distinguer pour unir, ou Les Degrés du savoir. Paris: Desclée De Brouwer, 1932; 4th ed., Paris, 1946. Translated under the supervision of Gerald B. Phelan as *The Degrees of Knowledge.* New York, 1959.

Le songe de Descartes. Paris: Corrêa, 1932. Translated by Mabelle L. Andison as *The Dream of Descartes.* New York: Philosophical Library, 1944.

La philosophie de la nature; Essai critique sur ses frontières et son objet. Paris: Téqui, 1935. Translated by Imelda C. Byrne as *Philosophy of Nature.* New York: Philosophical Library, 1951.

The Range of Reason. New York: Scribners, 1952.

Metaphysics

Sept Leçons sur l'être et les premiers principes de la raison spéculative. Paris: Téqui, 1934. Translated as *A Preface to Metaphysics: Seven Lectures on Being.* London and New York: Sheed and Ward, 1939.

Court Traité de l'existence et de l'existant. Paris: Hartmann, 1947. Translated by Lewis Galantière and Gerald B. Phelan as *Existence and the Existent.* New York: Pantheon, 1948.

Approches de Dieu. Paris: Alsatia, 1953. Translated by Peter O'Reilly as *Approaches to God.* New York: Harper, 1954.

Moral Philosophy

Science et sagesse. Paris: Labergerie, 1935. Translated by Bernard Wall as *Science and Wisdom.* New York: Scribners, 1940.

Education at the Crossroads. New Haven, CT: Yale University Press, 1943.

Neuf leçons sur les notions premières de la philosophie morale. Paris: Téqui, 1951.

La philosophie morale, Vol. I: *Examen historique et critique des grands systèmes.* Paris: Gallimard, 1960.

Social and Political Philosophy

Humanisme intégral. Paris: Aubier, 1936. Translated by M. R. Adamson as *True Humanism.* New York: Scribners, 1938.

La personne et le bien commun. Paris: Desclée De Brouwer, 1947. Translated by John J. FitzGerald as *The Person and the Common Good.* New York: Scribners, 1947.

Man and the State. Chicago: University of Chicago Press, 1951.

The Social and Political Philosophy of Jacques Maritain. Edited by Joseph W. Evans and Leo R. Ward. New York, 1955. Selected readings.

On the Philosophy of History. New York: Scribners, 1957.

Reflections on America. New York: Scribners, 1958.

Philosophy of Art

Art et scolastique. Paris: Librairie de l'Art Catholique, 1920; 3rd ed., Paris, 1935.

"Frontières de la poésie." In his *Frontières de la poésie et autres essais.* Paris: Rouart, 1935. This and the above work were translated by Joseph W. Evans as *Art and Scholasticism and the Frontiers of Poetry.* New York, 1962.

Situation de la poésie. Paris: Desclée De Brouwer, 1938. Written in collaboration with Raïssa Maritain. Translated by Marshall Suther as *The Situation of Poetry.* New York: Philosophical Library, 1955.

Creative Intuition in Art and Poetry. New York: Pantheon, 1953.

The Responsibility of the Artist. New York, 1960.

WORKS ON MARITAIN

Bars, Henry. *Maritain en notre temps.* Paris: B. Grasset, 1959.

Bars, Henry. *La politique selon Jacques Maritain.* Paris: Les Éditions ouvrières, 1961.

Croteau, Jacques. *Les fondements thomistes du personnalisme de Maritain.* Ottawa: Éditions de l'Université d'Ottawa, 1955.

Evans, Joseph W., ed. *Jacques Maritain: The Man and His Achievement.* New York: Sheed and Ward, 1963.

Gallagher, Donald, and Idella Gallagher. *The Achievement of Jacques and Raïssa Maritain: A Bibliography, 1906–1961.* New York: Doubleday, 1962.

Maritain, Raïssa. *Les grandes Amitiés,* 2 vols. New York: Éditions de la Maison Française, 1941. Vol. I translated by Julie Kernan as *We Have Been Friends Together.* New York: Longman, 1942. Vol. II translated by Julie Kernan as *Adventures in Grace.* New York: Longman, 1945.

The Maritain Volume of the Thomist. New York, 1943. Originally published as Vol. 5 (1943) of *The Thomist,* dedicated to Maritain on the occasion of his sixtieth birthday.

Phelan, Gerald B. *Jacques Maritain.* New York: Sheed and Ward, 1937.

Tamosaitis, Anicetus. *Church and State in Maritain's Thought.* Chicago, 1959.

Joseph W. Evans (1967)

MARITAIN, JACQUES [ADDENDUM]

Jacques Maritain died in Toulouse on April 28, 1973, as a professed religious of the Petits Frères de Jesus. His wife Raissa had died in 1959 when the couple was visiting France and from that point on Maritain's center of gravity was once again Europe. In Toulouse, he taught the brothers of his community and the published works that resulted are almost exclusively theological. Thus, Maritain continued to surprise: the quintessential layman became a professed religious, the philosopher became a theologian.

His reputation with many suffered when he published *The Peasant of the Garonne* in 1966. In the immediate wake of the ecumenical council dubbed Vatican II, Maritain was severely critical of developing trends in the Catholic Church. Teilhard de Chardin and phenomenology were major targets of his criticism. Some saw in this a retrogression, remembering *Antimoderne*. It helps to distinguish Maritain's political views from his Catholic faith. He held the latter with unswerving orthodoxy from the time of his conversion. It was otherwise with his political views. His long association with *Action Francaise*, so difficult to reconcile with his earlier socialism, was followed by a resurgence of his natural liberalism in

political matters. The conservatism of the *Peasant* is theological, not political.

Negative reactions to the *Peasant* are eclipsed by the upsurge of interest in Maritain during the latter part of the twentieth century. The Jacques Maritain Center at Notre Dame was founded in 1958 and seemed destined to become the repository of Maritain's papers. The bulk of his papers are to be found in Kolbsheim, the home of the *Cercle d'études Jacques et Raissa Maritain*. Under the general direction of René Mougel a magnificent sixteen volume *Oeuvres complètes* has appeared. There is another International Maritain Association centered in Rome under the aegis of Roberto Papini which has sponsored a score of publications and conferences, as well as a periodical, *Notes et Documents*. There are flourishing Maritain associations in Canada, the United States, and Latin America. Biographies have been written, collections of letters published, various monographs have appeared. A projected twenty volume set of Maritain's work in English is under way from the Jacques Maritain Center, whose web site at nd.edu can be consulted for other relevant materials.

Perhaps interest is strongest in his political, social, and aesthetic views. Given the contingency of the practical order this is surprising, perhaps, but would seem to attest to Maritain's knack of finding permanent values in the changing cultural landscape. His metaphysical views have their adherents still and there is a quickened interest on the part of physicists in Maritain's views of natural philosophy and natural science, as is evident in the institute founded by the physicist-philosopher Anthony Rizzi. Far from waning, interest in Maritain's thought seems to be increasing. For all that, it is perhaps not too much to say that it is his personality that continues to attract. Leon Bloy's line, "There is only one tragedy, not to be a saint," may seem a counter-cultural motto for a philosopher, but perhaps that is due to the all too exiguous character of recent philosophizing. In any case, as person as well as thinker, Jacques Maritain's influence is still strongly felt in the twenty-first century.

See also Aesthetics, History of; Liberalism; Phenomenology; Socialism; Teilhard de Chardin, Pierre.

Bibliography

WORKS BY MARITAIN

Jacques et Raissa Maritain Oeuvres Complétes, edited by René Mougel et. al. Strasbourg: Editions Universitaires Fribourg Suisse, 16 volumes, 1986–1999.

The Collected Works of Jacques Maritain, edited by Ralph McInerny, Bernard Doering, Fred Crosson. Notre Dame: Notre Dame University Press, 1997.

WORKS ABOUT MARITAIN

Barré, Jean-Luc. *Jacques et Raissa Maritain: Les Mendiants du Ciel*. Paris: Stock, 1995.

Doering, Bernard. *Jacques Maritain and the French Catholic Intellectuals*. Notre Dame, IN: University of Notre Dame Press, 1983.

McInerny, Ralph. *The Very Rich Hours of Jacques Maritain: A Spiritual Life*. Notre Dame: University of Notre Dame Press, 2003.

Schall, James V. *Jacques Maritain, The Philosopher in Society*. New York: Rowman and Littlefield, 1998.

Ralph McInerny (2005)

MARKOVIĆ, SVETOZAR
(1846–1875)

Svetozar Marković was a Serbian socialist, philosopher, and publicist. After prolonged uprisings between 1804 and 1815 had liberated Serbia from Turkey, a cultural revolution took place, led by the reformer of the Serbian language and orthography Vuk Karadžich (1787–1864), and socialist ideas began to spread. The first Serbian socialist writers were the economist and philosopher Živojin Žujović (1838–1870) and Svetozar Marković.

After technical studies in Belgrade, Marković continued his education in St. Petersburg, where he attended the lectures of Dmitri Pisarev and became acquainted with the ideas of the Russian revolutionary democrats. Marković went to France in 1869 and then to Zürich, where he became acquainted with the Western revolutionary workers' movement and with the works of Karl Marx. Marković became the correspondent for Serbia and the Balkans of the Marxist First International. In 1870 he returned to Serbia, where he gathered about himself a circle of young intellectuals and workers. He published *Radenik* (The Worker; 1871–1872), the first socialist newspaper in the Balkans, and later the newspapers *Javnost* (The Public) and *Glas Javnosti* (The Public Voice). After nine months' imprisonment for violating the press law, Marković, who had become seriously ill, was set free in 1875. He began publishing a new newspaper, *Oslobodjenje* (Liberation), but shortly afterward he died in Trieste.

The basic determinant of Marković's thought and activity was the Serbian social situation. The disoriented rural paupers and the small and unorganized urban proletariat had repudiated the patriarchal social order, but

they disagreed on the means of improving their lot. In search of ways to solve the social problems of his countrymen, Marković developed a socialist ideology. This theory was greatly influenced by the Russian revolutionary democrats Nikolai Chernyshevskii, Nikolai Dobrolyubov, and Pisarev, and later by Marx, but its main sources were materialist philosophy and the natural sciences—French eighteenth-century materialism (particularly Baron d'Holbach, Denis Diderot, and Jean Le Rond d'Alembert); the vulgar materialism of Friedrich Büchner, Karl Vogt, and Jacob Moleschott; the positivism of Auguste Comte and John Stuart Mill; and the scientists Charles Darwin, Ernst Haeckel, Wilhelm Wundt, and Ivan Mikhailovich Sechenov, the Russian physiologist. There are also traces in Marković's thought of the utopian socialists the Comte de Saint-Simon, François Marie Charles Fourier, and Étienne Cabet, as well as of other socialists such as Pierre-Joseph Proudhon and Louis Blanc.

ATHEISM AND MATERIALISM

Lacking a deep and systematic philosophical and sociopolitical education, Marković did not intend to become a philosopher or a literary figure but strove to be the ideologist and spiritual leader of a new trend in science and life—a publicist and propagator of new ideas. Nevertheless, his theoretical outlook was relatively original and presented an integral whole.

Marković's ideology embraced first of all the general principles of scientific atheism and natural-philosophical materialism expressed in the study "Realni Pravac u Nauci Iživotu" (The Realistic Trend in Science and Life; in the journal *Letopis Matice Srpske,* 1871–1872) and other works. From Chernyshevskii and Marx he borrowed the notion of the need for building up a philosophical theory as the basis of sociopolitical knowledge and practice. He called his view "scientific materialism and realism." All phenomena, as well as the processes of nature, society, and spiritual life, were interpreted in terms of matter and its laws. Nature and society were integrally connected. Only by means of science was the people's economic and political revival possible. Marković, like Marx, contrasted his view with Bakunin's. In spite of certain elements of mechanism and agnosticism in his outlook, Marković advocated the idea of dialectical development and an evolutionistic-materialistic theory of knowledge as the basis of the social struggle of the socialist movement.

In his interpretation of man and society, Marković drew upon Darwin, Comte, the French materialists, Ludwig Feuerbach, and Chernyshevskii. Morals is founded on knowledge and science, and the development of morals is affected by the development of man's needs through the socialization of instincts. Moral feelings are not innate; man becomes individually moral and socially more morally minded as society develops. Only by constant labor can man raise himself to a height unreachable by any other organism. Marković condemned the morals of bourgeois society as being founded upon the exploitation of the lower classes. Because morality is the indispensable consequence of the social machine, only a socialist revolution can bring about a new socialist morality. Seeing the primary goal of the future socialist society as the morality of its members, Marković termed his ethical socialism "idealistic realism." He did not conceive of the idea as being determined by matter, but spoke of the idea as the primary motive force in the development of society.

AESTHETICS

Believing that a spiritual revolution must precede the political and economic revolutions, Marković held that the social revival had to be supported by literature and art. In "Pevanje i Mišljenje" (Songs and Thought; *Matica,* 1868), "Realnost u Poeziji" (Reality in Poetry; *Matica,* 1870), and many other works, Marković expounded a materialist aesthetic modeled upon that of Chernyshevskii. Literature should be realistic and rational, expressing the genuine life, needs, and interests of the people, and should have an effect upon the general social revival. Marković's views decisively affected the development of Serbian literature, turning it toward Russian and western European realism.

SOCIOPOLITICAL VIEWS

In his voluminous book *Načelo Narodne Ekonomije* (The principles of the national economy [Belgrade, 1874]), written in the vein of J. S. Mill and Chernyshevskii, Marković praised Marx for his discovery of the law of social development, but he held that these laws could not be applied to Russia, Serbia, and other economically undeveloped countries, which, in Marković's opinion, could bypass capitalism and move from patriarchal cooperatives directly to socialism. Marković's teachings on society, state, and revolution, in spite of some elements of utopianism and historical idealism, showed a high degree of accuracy. Although he gave too much weight to the roles of social consciousness, science, and philosophy, and consequently to the revolutionary intelligentsia, in the development of socialist society, his program was revolutionary and democratic. In a series of works, especially in his most original work, *Srbija na Istoku* (Serbia in the East

[Novi Sad, 1872]), Marković defended the Paris Commune and criticized the capitalistic social system of western Europe and the narrowness of the bourgeois democracies. Marković was convinced that the transition to socialism was possible only by means of a revolution of the whole people against foreign invaders and native capitalist exploiters. He developed a fragmentary theory of the smashing of the bourgeois state in the socialist revolution and the withering away of the socialist state in the process of building communism. Like Marx, he held that only in conjunction with revolutionary practice could revolutionary theory solve the social problem. He perceived the significance of the class struggle in the West, but in backward Serbia he thought that the revolutionary intelligentsia could play a more decisive role than the proletariat. He advocated federation and self-government for the southern Slav nations. He also advocated a system of cooperatives.

Although Marković was more a revolutionary democrat than a Marxist, his teachings nevertheless united general Marxian principles concerning revolution with theories concerning the specific national character of Serbia. Moreover, they stressed the need for joint action on the part of the revolutionary intelligentsia, the peasantry, and the workers. Thus, Marković was the founder and leader of the Serbian socialist movement, as well as its theoretician, philosopher, aesthetician, and literary critic.

See also Aesthetics, History of; Aesthetics, Problems of; Alembert, Jean Le Rond d'; Atheism; Chernyshevskii, Nikolai Gavrilovich; Comte, Auguste; Darwin, Charles Robert; Diderot, Denis; Feuerbach, Ludwig Andreas; Fourier, François Marie Charles; Haeckel, Ernst Heinrich; Holbach, Paul-Henri Thiry, Baron d'; Marxist Philosophy; Marx, Karl; Materialism; Mill, John Stuart; Moleschott, Jacob; Pisarev, Dmitri Ivanovich; Proudhon, Pierre-Joseph; Realism; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Socialism; Wundt, Wilhelm.

Bibliography

WORKS BY MARKOVIĆ

Many of Marković's scientific and periodical papers are collected in *Sabrani Spisi,* 4 vols. (Belgrade, 1960–1965).

WORKS ON MARKOVIĆ

For literature on Marković, see Slobodan Jovanović, *Svetozar Marković* (Belgrade, 1904). Jovan Skerlić, *Svetozar Marković* (Belgrade, 1910); Veselin Masleša, *Svetozar Marković* (Belgrade, 1947); Dušan Nedeljković. "Lik Svetozara Markovića," in the journal *Glas SAN* (3) (1951): 200–207; Dimitrije Prodanović, *Shvantanje Svetozara Markovića o državi* (Belgrade, 1961).

ADDITIONAL SOURCES

McClellan, Woodford. *Svetozar Marković and the Origins of Balkan Socialism.* Princeton, NJ: Princeton University Press, 1964.

Stoianovich, Traian. "The Pattern of Serbian Intellectual Evolution." *Comparative Studies in Society and History* 1 (3) (1959): 242–272.

Stokes, Gale. "Svetozar Marković in Russia." *Slavic Review* 31(3) (1972): 611–625.

Andrija Stojković (1967)
Bibliography updated by Philip Reed (2005)

MARSILIUS OF INGHEN
(1340–1396)

Marsilius of Inghen was a scholastic theologian, writer on logical textbooks, and prolific commentator on Aristotle. He played an important role in the foundation of the University of Heidelberg. His significance rests not only on his commentaries on Aristotle—his advocacy and popularization of the new, nominalist logic and semantics—but also on an independent-minded theology that sometimes rejected post-Scotistic positions in favor of thirteenth-century positions (such as those of Thomas Aquinas or Bonaventure).

Marsilius of Inghen was a student at Paris, matriculating there in Arts in 1362, and then in Theology in 1366. At Paris, he was influenced by the thought of John Buridan, and he undertook significant administrative work, including rectorships (1367–68, 1371) as well as representation to the Papal court (1369, 1377–78). Marsilius's whereabouts are largely unknown between 1379 and the founding of the University of Heidelberg in 1386—except, that is, for a Nijmegen banquet he attended in 1382. From 1386 to 1392, he was a Master at Heidelberg—and was also an occasional Rector—up until his death in 1396 (Hoenen 1993, pp. 7–11; Santos Noya, 2000, Vol. 87, pp. 1–26).

He read the Sentences (the standard requirements to become a Master of Theology) from 1392 to 1394. Part of the preparation for this commentary was most likely done in Paris from 1367 to 1377. (Hoenen and Braakhuis 1993, pp. 39–57; Santos Noya 2000, Vol. 87, pp. 31–32).

Marsilius was a nominalist on universals. Like Ockham and Buridan he did not believe that universals exist outside the soul, and that the direct object of each science is merely the proposition in the mind. Real objects, he believed, are the objects of sciences via the signification of

the proposition. Marsilius's logic and semantics can be described reliably as Buridanist, albeit with some points of dissent and less detail. As well, he differed from Buridan on the division of supposition, the signification of chimera, his definitions of *ampliatio* and *appellatio*, and his non-adoption of *suppositio naturalis* (Bos 1983, p. 254).

Marsilius's natural philosophy is empiricist; he holds that the starting point of natural philosophy is sense data and per se known principles. From this point he then leaps from singular observations to a universal proposition if there is no expectation of a counterexample—due to the mind's inclination to truth. Thus, a causal connection can be held to be universal, though one has not experienced all its instances.

In his theology, he criticized both the Scotistic position that the Divine Ideas are formally distinct from the Divine Essence, and the Ockhamist thesis that the Ideas are identical with the objects that are known. He held the Thomistic theses that God's Ideas of created things are not distinct from his essence and that the difference between the divine attributes exists only in the human mind due to its finitude. He also held that natural reason can prove that God is the cause of all and knows created things. Marsilius brought together the critical semantico-logical tradition of the fourteenth century and the themes of thirteenth century theologians such as Aquinas and Bonaventure (Hoenen 1993, pp. 235–253).

In the fifteenth and sixteenth centuries, he was regarded as a great advocate of nominalism, and grouped with Buridan, Ockham and Gregory of Rimini. His logical treatises exist in many manuscripts, and were widely used as textbooks in the fifteenth century. His theology of grace and divine foreknowledge was well known and quoted by late-scholastic writers such as Vitoria, De Soto, Molina, and Suarez. His Aristotelian commentaries were also well known and cited up to the early-modern period. For example, both Leonardo da Vinci and Gallileo Galilei refer to his commentary on Aristotle's *De Generatione et Corruptione*.

See also Buridan, John; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS BY MARSILIUS OF INGHEN

Quaestiones super quattuor libros Sententiarum. Super primum quaestiones 1–7 and 8–21, edited by Manuel Santos Noya, et al. (Studies in the History of Christian Thought, vols. 87–88.) Leiden, Netherlands: Brill, 1999–.

Treatises on the Properties of Terms. A first critical edition of the *Suppositiones*, *Ampliationes*, *Apellationes*, *Restrictiones and Alienationes*, with introduction, translation, notes, and appendices by Egbert P. Bos. (Synthese Historical Library, vol. 22.) Dordrecht, Netherlands: D. Reidel, 1983.

WORKS ABOUT MARSILIUS OF INGHEN

Braakhuis, H. A. G., and Maarten J. F. M. Hoenen, eds. *Marsilius von Inghen*. Acts of the International Marsilius of Inghen Symposium. Nijmegen, Netherlands: Ingenium, 1992.

Hoenen, Maarten J. F. M. *Marsilius of Inghen: Divine Knowledge in Late Medieval Thought*. Leiden, Netherlands: Brill, 1993.

Hoenen, Maarten J. F. M., and Paul J. J. M. Baakker, eds. *Philosophie und Theologie des ausgehenden Mittelalters: Marsilius von Inghen und das Denken seiner Zeit*. Leiden, Netherlands: Brill, 2000.

C. F. Ledsham (2005)

MARSILIUS OF PADUA

(c. 1275/1280–1342)

Marsilius of Padua (Marsilio dei Mainardini), an Italian political theorist, was born between 1275 and 1280 and died in 1342. He probably studied medicine at the University of Padua. In 1313 he was rector of the University of Paris, where he met such leading Averroists as Peter of Abano and John of Jandun. He is chiefly famous for his antipapalist treatise *Defensor Pacis* (Defender of peace; 1324), a landmark in the history of political philosophy. When his authorship of this work became known in 1326, he was forced to flee to the court of Louis of Bavaria in Nuremberg; Pope John XXII thereupon branded him a heretic. Marsilius subsequently assisted Louis in various imperial ventures in Italy.

DEFENSOR PACIS

The primary purpose of the *Defensor Pacis* was to refute the papalist claims to "plenitude of power" as these claims had been advanced by Pope Innocent IV, Egidius of Rome, and others in the thirteenth and fourteenth centuries. So crushing was the refutation produced by Marsilius that it completely reversed the papalist position. The papal position had held that secular rulers must be subject to the papacy even in "temporal" affairs, so that they must be established, judged, and, if necessary, deposed by the pope. Marsilius, in contrast, undertook to demonstrate that the papacy and the priesthood in general must be subject not only in temporal, but even in "spiritual," affairs to the whole people and to the secular ruler acting by the people's authority. The powers of the

priesthood were to be reduced to the administration of the sacraments and the teaching of divine law, but even in these functions the priests were to be regulated and controlled by the people and its elected government. The upshot of Marsilius's doctrine was that the attempt to base human society on religious values under priestly control was decisively overthrown; instead, the way was opened for a purely secular society under the control of a popularly elected government. Hence, it is understandable that Marsilius has been hailed as a prophet of the modern world. His treatise exerted a marked influence during the period of the Reformation.

THEORY OF THE STATE. Equally as important as these revolutionary conclusions are the premises from which Marsilius derived them. These premises are found in his general theory of the state, which is noteworthy for its fusing of three distinct themes. The first is the Aristotelian teleological view of the state as subserving the good life. The various parts of the state, including government, are defined by the contribution they make to the rational "fulfillment" of men's natural desire for a "sufficient life." This fulfillment proceeds through the "proper proportioning" of men's actions and passions, ranging from nutritive and sensitive acts to appetitive and cognitive ones. The function of government is to regulate men's transitive acts in accordance with the law as a standard of justice. The first theme, then, stresses an affirmative and maximal utilitarianism—what is required for the attainment of the highest ends of the "sufficient life," the common benefit, and justice.

The second theme of Marsilius's political theory, in contrast, is a negative and minimal utilitarianism. It emphasizes the inevitability of conflicts among men and the consequent need for the formal instrumentalities of coercive law and government in order to regulate these conflicts. Without such regulation, Marsilius repeatedly insists, human society itself must be destroyed. In developing this theme, Marsilius presents a positivistic concept of law, which stands in contrast with his nonpositivistic conception of justice (a distinction often overlooked in discussions of his ideas). He holds that there are objective criteria of justice, which he characterizes in terms of Aristotle's analysis of rectificatory justice—moderating the excesses of men's transitive acts and "reducing them to equality or due proportion," thereby promoting the common benefit. But whereas Marsilius views law as a system of general rules concerned with the regulation of the same "excesses" and the resultant conflicts, as well as with other matters bearing on the common benefit, he emphasizes that these legal rules need not be based on "true cognitions of justice." On the contrary, laws may be based on "false cognitions of the just and the beneficial," so that Marsilius, unlike most medieval political philosophers, holds that justice is not a necessary condition of law. What is necessary is that the legal rules have coercive force, such that with regard to their observance "there is given a command coercive through punishment or reward to be distributed in the present world." These rules and the government that enforces them must be unitary in the sense that, if a society is to survive, it cannot have two or more rival coercive bodies of law and government.

The third theme of Marsilius's political theory is that the people is the only legitimate source of all political authority. It is the people, the whole body of citizens or its "weightiest part," that must make the laws either by itself or through elected representatives, and it is also the people that must elect, "correct," and, if necessary, depose the government. Marsilius presents many arguments for this republican position: (1) The whole people is intellectually and emotionally superior to any of its parts, so that only from its choice will emerge the best law and government, the ones most conducive to the common benefit, as against the ones that subserve the interests of some special group; (2) self-legislation is necessary for individual freedom; (3) only if the laws and government are chosen by the people will they be obeyed; and (4) that which affects all ought to be subject to approval by all.

Although all three themes of Marsilius's general political theory were found in earlier medieval political philosophers, no other philosopher had given the second and third themes as central a position as did Marsilius. As a result of this, although Marsilius's first theme—about the ends of the "sufficient life," the common benefit, and justice—persists throughout his treatise, it is overshadowed by his emphases on coerciveness as the essence of political authority and on the republican bases of all such authority. The full consequence of these emphases emerges in the applications he makes of his general political theory to the problems of ecclesiastical politics.

APPLICATIONS OF THE THEORY. In keeping with his first theme, Marsilius views the Christian priesthood as one of the parts of the state dedicated to achieving the "sufficient life" for all believers. Unlike the other parts of the state, however, the priesthood subserves the "sufficient life" to be attained primarily "in the future world" rather than the present one. Like the other Averroists, Marsilius manifests skepticism about the rational demonstrability of such a future life; nevertheless, he offi-

cially accepts the Christian doctrine that the future life is superior to the present life. He also holds, however, that secular and religious values are in basic opposition; here he seems to be applying in the realm of the practical the Averroist doctrine of the contrariety of reason and faith in theoretic philosophy.

Taken in conjunction with the maximal, affirmative utilitarianism of his first theme, accepting that the priesthood subserves the highest end of man would have required Marsilius to accept also the papalist doctrine that the "secular" government, subserving the lesser end of this-worldly happiness, must be politically subordinate to the priesthood. At this point, however, Marsilius's second and third themes have their effect. Since the essence of political authority is the coerciveness required for the minimal end of preserving society, it follows that the higher end subserved by the priesthood does not entitle it to superior political authority. The question of the order of political superiority and inferiority is thus separated from the question of the order of moral and religious values. What determines the order of political authority is not the greater excellence of one end over another but, rather, the specifically political need for unified coercive authority in order to prevent unresolved conflicts from destroying society. Hence, the secular government, as bearer of this coercive authority, must be politically superior to the priesthood. If the priests refuse to obey the government and its laws, then they must be compelled to do so, because such disobedience threatens that unity of coercive authority without which society cannot survive. Indeed, it is because of this disobedience and because of its claim to a rival, superior "plenitude of power," that Marsilius convicts the papacy of being the gravest enemy of civil peace. In this context Marsilius presents his whole critique of the papacy as an application to fourteenth-century conditions of Aristotle's book on revolutions (*Politics* V), dealing with the ways in which threats to civil peace may be avoided.

In addition to this political argument against diverse centers of coercive power in any society, Marsilius also stresses, from within the religious tradition itself, that religious belief, in order to be meritorious, must be purely voluntary. Hence, in order to fulfill its mission, divine law and the priesthood that teaches and administers it cannot be coercive in this world.

Marsilius's third theme, republicanism, also plays an important role in the political subordination of the priesthood and papacy. The only rules and persons that are entitled to the status of being coercive laws and government officials are those ultimately chosen by the people; hence, there can be no crediting the claims of divine law and the priesthood to a separate derivation of coercive political authority from God. It is true that Marsilius subsequently holds that secular rulers govern by divine right, but he views this only as a divine confirmation of the people's ultimate electoral authority. This republicanism operates not only in the relation of the priesthood to the secular state but also in its relation to religious affairs. Because the whole people is superior in virtue to any of its parts and because freedom requires popular consent or election, the priesthood itself must be elected by the people of each community rather than being appointed by an oligarchically chosen pope, and the pope himself must be elected by the whole of Christendom. Similarly, the whole people must elect general councils to provide authoritative interpretations of the meaning of divine law. In these ways Marsilius's general political theory leads to a republican structure for the church as against its traditional monarchic structure. In effect, this also means that the secular government, acting by the people's authority, secures hegemony over the priesthood and papacy in all spheres.

See also Aristotelianism; Aristotle; Averroism; John of Jandun; Medieval Philosophy; Political Philosophy, History of; Republicanism; Sovereignty; Utilitarianism.

Bibliography

The *Defensor Pacis* is available in two critical editions, one edited by C. W. Previté-Orton (Cambridge, U.K.: Cambridge University Press, 1928) and one edited by Richard Scholz, 2 vols. (Hanover: Hahn, 1932–1933), in the series *Fontes Juris Germanici Antiqui* of the series *Monumenta Germaniae Historica.* There is an English translation by Alan Gewirth, *Defensor Pacis* (New York: Columbia University Press, 1956).

A comprehensive bibliography of studies published to about 1950 of Marsilius's doctrines is contained in Alan Gewirth, *Marsilius of Padua and Medieval Political Philosophy* (New York: Columbia University Press, 1951), pp. 323–326.

Bibliographical information on subsequent studies will be found in Georges de Lagarde's important study, *Marsile de Padoue,* 2nd ed. (Paris, 1948), which is in the series *La naissance de l'esprit laique au déclin du moyen âge,* 3rd ed. (Paris, 1959–).

Alan Gewirth (1967)

MARSILIUS OF PADUA [ADDENDUM]

In order to understand Marsilius more fully, it is useful to examine both the classical influences upon his work and

the ways he applies his own principles in the minor works such as *Defensor Minor* and *De Translatione Imperii*.

MARSILIUS AND CICERO

Most discussions of *Defensor Pacis* concentrate upon Aristotle's *Ethics* and *Politics* (which had become available in translation around 1250 and 1260, respectively). Indeed, Marsilius employs the Aristotelian distinctions of the healthy types of civil constitution: monarchy, aristocracy, and polity and their complements the diseased constitutions: tyranny, oligarchy, and extreme democracy. However, though Aristotle is certainly the primary source of many of the distinctions in Part I of *Defensor Pacis*, there are other key influences as well. Among these is Cicero's doctrine of natural duty to others from his *De officiis*. Cary Nederham has argued that this sense of natural duty is the secular analogue to theological or Christian duty. The use of parallel justifications for why a person should be committed to the community follows the general structure of the book in which Part I creates a secular justification for politics whereas Part II elaborates the foundations of ecclesiastical duty.

The secular duty to the community is a natural duty so that every person in the state must fulfill the duties of friendship and of civic society—without regard to personal welfare. This duty extends to a concern for others and a duty to rescue and assist those in need. Because the source of the duty is natural to all people, there is no national restriction on this duty. Thus, it commits each person to exhibit concern beyond his own society to others internationally.

DEFENSOR MINOR AND *DE TRANSLATIONE IMPERII*

These works are more conventional and do not contain the split presentation of secular argument and theological exposition that characterized *Defensor Pacis*. These minor works are more conventional dealing with parsing the jurisdictions of theology and secular government. Though these works are not as well known as the *Defensor Pacis*, they are useful to help put Marsilius's major work into perspective. For example, one of the possible motivations for Marsilius's antipapal rhetoric (though Marsilius, himself, was a priest) might be Marsilius's alliance with the Bavarian King Ludwig IV. Ludwig wanted to expand his empire and move into Italy. (It should be remembered that at this time the Pope resided in Avignon, France.) Marsilius's writing was associated with Ludwig, who appointed Marsilius as spiritual vicar of Rome and himself as the Roman Emperor. However, this situation was short lived and soon both fled back to Germany.

The *Defensor Minor* and *De Translatione Imperii* fit into this context. They apply principles of *Defensor Pacis* to contemporary problems. For example, papal authority is questioned in regard to Ludwig's plan to marry his daughter to a close relative in order to stabilize his political prospects. Both Marsilius and William of Ockham were to weigh in on this question as an issue of authority. In *Defensor Minor* no new positions are forged, but are fine tuned so that they might be applied to cases such as the marriage of Ludwig's daughter.

Another example concerns the bogus "Donation of Constantine." In this forged document the Roman Emperor Constantine supposedly granted the whole of the Roman Empire to the Pope who, in turn, allowed the daily duties of running the secular to fall upon the emperor. This document sought to establish a legal claim for the pope's universal secular power. Marsilius argues against the Donation in both *Defensor Pacis* and *De Translatione Imperii*.

See also Aristotelianism; Aristotle; Cicero, Marcus Tullius; Duty; Political Philosophy, History of; William of Ockham.

Bibliography

Ashcraft, Richard. "Ideology and Class in Hobbes' Political Theory." *Political Theory* 6 (1978): 17–62.

Black, Antony. *Political Thought in Europe, 1250–1450*. Cambridge, U.K.: Cambridge University Press, 1992.

Black, Antony. "Society and the Individual from the Middle Ages to Rousseau: Philosophy, Jurisprudence, and Constitutional Theory." *History of Political Thought*. 1, (2) (Summer 1980): 145–166.

Luscombe, Davie E. "The State of Nature and the Origins of Society." In *The Cambridge History of Later Medieval Philosophy*, edited by N. Kretzmann, A. Kenny, and J. Pinborg. Cambridge, U.K.: Cambridge University Press, 1982.

McGrade, A. S. "Aristotle's Place in the History of Natural Rights." *Review of Metaphysics* 49 (4) (1996): 803–829.

WORKS BY MARSILIUS OF PADUA

Defensor Pacis, edited by C. W. Previté-Orton. Cambridge, U.K.: Cambridge University Press, 1928.

Defensor Pacis. Translated by Alan Gewirth. New York: Columbia University Press, 1956.

Oeuvres Minueres, edited by C. Jeudy and J. Quillet. Paris: Editions CNRS, 1979.

Writings on the Empire: Defensor minor and De translatione imperii, edited by C. J. Nederman. Cambridge, U.K.: Cambridge University Press, 1993.

WORKS ON MARSILIUS OF PADUA

Canning, Joseph. "The Role of Power in the Political Thought of Marsilius of Padua." *History of Political Thought* 20 (1) (1999): 21–34.

Coleman, Janet. "Medieval Discussions of Property: *Ratio* and *Dominium* According to John of Paris and Marsiglio of Padua." *History of Political Thought* 4 (Summer 1983): 209–228.

Condren, Conal. "Democracy and the *Defensor Pacis*: On the English Language Tradition of Marsilian Interpretation." *Il Pensiero Politico* 8 (1980): 301–316.

Condren, Conal. "Marsilius of Padua's Argument from Authority: A Study of Its Significance in the *Defensor Pacis*." *Political Theory* 5 (1977): 205–218.

D'Entrèves, A. P. *The Medieval Contribution to Political Thought: Thomas Aquinas, Marsilius of Padua, Richard Hooker*. Oxford: Oxford University Press, 1939.

Gewirth, Alan. "John of Jandun and the *Defensor Pacis*." *Speculum* 23 (1948): 267–272.

Gewirth, Alan. *Marsilius of Padua and Medieval Political Philosophy*. 2 vols. New York: Columbia University Press, 1951. Contains an extensive bibliography.

Gewirth, Alan. "Republicanism and Absolutism in the Thought of Marsilius of Padua." *Medioevo* 5 (1979): 23–48.

Grignaschi, Marc. "La Rôle de l'aristotélisme dans le 'Defensor Pacis' de Marsile de Padoue." *Revue d'Histoire ed de Philosophie Religieuse* 35 (1955): 310–340.

Kaye, Sharon. "Against a Straussian Interpretation of Marsilius of Padua's Poverty Thesis" *History of Philosophy Quarterly* 11 (3) (1994): 269–279.

Loffelberger, Michael. *Marsilius von Padua: Das Verhältnis Zwischen Kirche und Staat in "Defensor Pacis."* Berlin: Duncker and Humbolt, 1992.

Nederman, Cary J. "Character and Community in the *Defensor Pacis*: Marsiglio of Padua's Adaptation of Aristotelian Moral Psychology." *History of Political Thought* 13 (3) (1992): 377–390.

Nederman, Cary J. *Community and Consent: The Secular Political Theory of Marsiglio of Padua's 'Defensor Pacis'*. Lanham, MD: Rowman and Littlefield, 1995. Contains a good bibliography.

Nederman, Cary J. "From *Defensor Pacis* to *Defensor Minor*: The Problem of Empire I Marsiglio of Padua." *History of Political Thought* 16 (3) (1995): 312–329.

Nederman, Cary J. "Nature, Justice and Duty in the *Defensor Pacis*: Marsiglio of Padua's Ciceronian Impulse." *Political Theory* 18 (1990): 615–637.

Quillet, Jeannine. "L'aristotélisme de Marsile de Padoue." *Miscellanea Medievalia* 2 (1963): 696–706.

Quillet, Jeannine. "L'aristotélisme de Marsile de Padoue et ses rapports avec l'averoïsm." *Medioevo* 5 (1979): 81–142.

Spiers, Kerry E. "The Ecclesiastical Theory of Marsilius of Padua: Sources and Significance." *Il Pensiero Politico* 10 (1977): 3–21.

Tierney, Brian. "Marsilius on Rights." *Journal of the History of Ideas* 52 (1991): 3–17.

Wilks, Michael. "Corporation and Representation in the *Defensor Pacis*." *Studia Gratiana* 15 (1972): 251–292.

Michael Boylan (2005)

MARSTON, ROGER

(c. 1250–1303)

Roger Marston, the Augustinian Scholastic, was born in one of England's Marstons. He was educated at the Faculty of Arts and Theology at the University of Paris about 1270 and taught at Oxford and Cambridge between 1276 and 1285. He was the provincial of the English Franciscans between 1292 and 1298.

Marston's career may be characterized as a conscious effort to restore St. Augustine to his position as the great leader of Christian philosophers and theologians. In carrying out the proposals of his teacher, John Peckham (also an Augustinian), Marston exhibited a phenomenal knowledge of the writings of Augustine, as well as a fine sense of historical and textual criticism. He must have been attacked as an archconservative, because he defended himself by remarking that he did not cling to tradition out of mere habit, but that after a reasonable scrutiny of the evidence, he had formed opinions that harmonized the writings of the "saints" with the wisdom of the philosophers. Marston knew the Greek and Muslim philosophers, and interpreted them with a great deal of subtle skill, sometimes calling attention to fundamental ambiguities in their thought.

Marston needed all the resources at his command to counter the attacks directed against the Augustinian theory of divine illumination, which he deemed necessary to explain certitude. Since the attacks were made under the guise of Aristotle's authority, Marston attempted to reconcile Augustine's theory of knowledge with that of Aristotle, as seen through the latter's Islamic commentators. Thus, Roger claimed that the Eternal Light of Augustine is the same as the separate agent intellect of Avicenna and Averroes. However, the English friar would not allow man to be "dispossessed" of his own individual agent intellect, and hence he posits a double agent intellect: divine and human. This was one of the medieval solutions to the idealist-empiricist dilemma.

In the realm of the philosophy of nature, there was one doctrine of Thomas Aquinas to which Marston took serious exception—namely, the Thomistic contention that each individual being had but one form. Prior to Thomas, the far more common opinion had been that in material beings there was a plurality of forms. In man there were the forms of "vegetivity," "sensitivity," and "rationality," corresponding to the human functions of nutrition, sensation, and thought. Marston's solution to the question introduced a refinement that amounted to a synthesis of the Thomistic and traditional solutions,

although it favored the latter. There is one substantial form for each being, but that single form admits of various subordinate and persisting degrees, or grades. Marston's theory of the grades of the form is the first organized version of this theory that has come down to us from the Middle Ages.

With respect to the majority of his philosophico-theological tenets, Marston followed the lead of Bonaventure. With Bonaventure (and against Thomas), he considered an eternal creature an impossibility. Prime matter can exist apart from all forms by divine intervention, because God is the "Form of all things" who conserves his handiwork just as water conforms to the intricate convolutions of a mold, as long as it is contained by the mold. On the subject of God's foreknowledge of future human acts—a perennial problem in Christian philosophy—Marston remarks that since an individual's memory of a past event does not constrain his free will with regard to the past, neither does God's foreknowledge constrain his free will with regard to the future.

For a medieval, Marston has an unusually personal style, and his remarks are often a source of valuable information for the historian.

See also Agent Intellect; Augustine, St.; Augustinianism; Averroes; Avicenna; Bonaventure, St.; Empiricism; Idealism; Medieval Philosophy; Peckham, John; Thomas Aquinas, St.; Thomism.

Bibliography

WORKS BY MARSTON

Quaestiones Disputatae, edited by Franciscani non-nominati. Quaracchi, 1932: *Bibliotheca Franciscana Scholastica VII,* lxxx + 497pp.)

Quodlibeta Quatuor, edited by G. Etzkorn and I. Brady. Grottaferrata, 1994: *Bibliotheca Franciscana Scholastica XXVI,* 2nd ed., 87 + 550 pp.). Introduction by Ignatius Brady includes a thorough treatment of Marston's life, career, and works.

WORKS ON MARSTON

Etzkorn, G. "Roger Marston, O.F.M.: An Example of Thirteenth-Century Anti-Semitic Apologetics." *Cithara* 21 (1982): 3–16.

Etzkorn, G. "Roger Marston's Grades' Theory in the Light of His Philosophy of Nature." *Miscellanea Mediaevalia* 2 (1963): 535–542.

Hissette, R. "Roger Marston a-t-il professé l'hylémorphisme universel?" *Revue de Théologie Ancienne et Médiévale* 39 (1972): 205–223.

Magnani, P. *Il problema delle Rationes Seminales in Ruggero di Marston.* Milan: Marzorati, 1992.

Pérez-Estévez, A. "La materia en Roger Marston." *Veridad y Vida* 55 (1997): 217–220.

Girard J. Etzkorn (1967, 2005)

MARTINEAU, JAMES
(1805–1900)

James Martineau, an English philosopher and religious leader, was born in Norwich. He was a brother of Harriet Martineau, the novelist and economist. James Martineau attended school in Norwich and Bristol and went on to study for the ministry under the Unitarian auspices of Manchester New College at York. He accepted a call to a congregation in Dublin in 1828 and was married later the same year. In 1832 he became minister to a dissenting congregation in Liverpool. He occupied this post for twenty-five years, but for most of that period he was also teaching philosophy and other subjects at Manchester New College, and when the college was moved to London in 1857, he moved with it. From 1869 to 1885 he served as principal of the college. Despite the criticism aroused by his views on religious and theological matters, he was regarded as the foremost spokesman of Unitarianism in England and was revered by many in other religious groups as well for his impressive contributions to the literature of hymn, private prayer, and sermon.

In accordance with the then prevailing tendency of Unitarian thought, Martineau was brought up to accept the doctrines of associationism, egoism, and necessitarianism as taught by David Hartley and Joseph Priestley. In his early teaching he used works by James Mill and Thomas Brown as texts, but the difficulties he had in defending their views, together with his own growing sense of the inadequacy of their philosophy as a basis for a Christian outlook, led him rapidly toward a new general position. By 1839 he concluded that necessitarianism was incompatible with that sense of "the personal origin and personal identity of sin" which is central to Christianity. During the next half-dozen years he worked out the implications of this point. The results were first published in 1845 and 1846 in two long reviews (reprinted in *Essays, Reviews, and Addresses*) that outlined the positions he was to develop and defend for the rest of his life. Although he learned much from a year of study in Berlin in 1848 and 1849, German philosophy did not really change his thought. He remained far more a follower of Bishop Butler and Thomas Reid than of Immanuel Kant or G. W. F. Hegel.

At the basis of all of Martineau's constructive thought is the view that we must accept as true certain deliverances of consciousness that appear to give us directly information about the external world, the self, and morality. Neither Kant nor William Hamilton nor J. S. Mill seemed to him to have given us reason to distrust the intuitions of the mind, and since these intuitions present themselves as reliable, we are entitled to have faith in them until reasons against them are produced. Martineau's intuitionism is the philosophical counterpart of the very great emphasis he placed, in interpreting religion, on personal religious experience. It is in such experience, he held, that one must look for revelation, not in messages delivered by others nor in traditions preserved by organized groups. Philosophically, both epistemology and ethics lead directly to justifications of religious belief.

From the very start of knowledge, Martineau argued, we are aware of a self and a not-self, and we are aware of these not as simply passively there but as being actively related. We thus intuit ourselves as willing and the world, in turn, as an expression of will. The former intuition is at the basis of our understanding of causality, which cannot be explained in terms of succession of phenomena, and the idea of causality finds its mature expression in the belief that God is the noumenal cause of the phenomenal order. Science, which deals only with phenomena, cannot upset our belief in God, but the increasing unity of the laws and theories that science discovers acts as a confirmation of our intuitive belief in the unity of the cause of nature.

If the "natural" attributes of God, such as omnipotence and intelligence, are revealed through our experience of the external world, the moral attributes are revealed to us primarily in our moral experience. Martineau argued very carefully that the central subject of moral judgment is motives or "springs of action," not acts or consequences. He held that whenever there is more than one motive competing to direct our action, we are intuitively aware that one of the motives is higher than the others.

"The moral faculty," he said, "is not any apprehension of invisible qualities in external actions, not any partition of them into the absolutely good and absolutely evil, not any intellectual testing of them by rules of congruity or balances of utility, but a recognition, at their very source, of a scale of *relative* values lying within ourselves," relative because a given motive may be higher in relation to one alternative, lower in relation to another. To be good is to choose to act on the relatively higher motive. Once this choice is made, consideration of consequences comes in to aid in selecting the particular act that will best express the motive in the actual circumstances. It is the first choice only that is morally relevant, though the second is, of course, important. Since the moral value of both agent and act is wholly determined by his choice of motive, Martineau went to considerable pains to defend absolute freedom of the will. The arguments rely heavily on the concept of cause developed in his epistemology. In our own willing we learn something of the nature of God's activity; the realization that there is an authoritative demand on us to act on the relatively higher motive is the chief revelation of God within our moral experience. The authoritativeness of the demand can be explained only in theistic terms, and the content of the demand reveals to us God's moral nature.

Martineau's style is extremely florid and his exposition quite diffuse. In his epistemological and metaphysical writings he seems often to have missed the point of an opposing theory or to have been content with very weak arguments for his own. But his ethics, as an account of the ethics of motive, if not highly original, is in conception and in execution one of the finest that has ever been presented.

See also Brown, Thomas; Butler, Joseph; Egoism and Altruism; Epistemology; German Philosophy; Hamilton, William; Hartley, David; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Mill, James; Mill, John Stuart; Priestley, Joseph; Reid, Thomas.

Bibliography

Works by Martineau include *A Study of Spinoza* (London: Macmillan, 1882); *Types of Ethical Theory,* 2 vols. (Oxford: Clarendon Press, 1885); *A Study of Religion,* 2 vols. (Oxford: Clarendon Press, 1888); *The Seat of Authority in Religion* (London: Longmans, Green, 1890); and *Essays, Reviews, and Addresses,* selected and revised by Martineau himself (London: Longmans, Green, 1890–1891).

For an account of his life and philosophy, see James Drummond and C. B. Upton, *Life and Letters of James Martineau* (London: J. Nisbet, 1902); criticism of the ethics will be found in Henry Sidgwick, *Lectures on the Ethics of T. H. Green, Mr. Herbert Spencer, and J. Martineau* (London: Macmillan, 1902).

J. B. Schneewind (1967)

MARTINETTI, PIERO
(1872–1943)

Piero Martinetti, an Italian metaphysician, was professor of theoretical philosophy at the University of Milan from

1906 until 1931, when he resigned in protest against the oath imposed on university professors by the Fascist regime.

Martinetti sought to reestablish metaphysics as a valid science by a method whose validity would have to be recognized even by positivists. This project involved a refutation of positivism on its own grounds. The positivist attack on metaphysics, Martinetti claimed, is valid only against vulgar or dogmatic metaphysics. Scientific metaphysics meets all the requirements of scientific methodology. It adheres to data that all science must recognize; but it is no mere synthesis of the sciences, for it interprets scientific findings and determines their meaning rather than their mere truth. Consequently, a scientific metaphysics would achieve, on a posteriori grounds, successive unifications of empirical data until the Absolute was achieved.

The first of the successive levels in this projected unification is that of the "I" or self as a unity of sensuous consciousness. This is the constant flux of sense perception, the central point around which all perception is synthesized. At this stage no distinction is made between subject and object. The self at this level possesses a rudimentary transcendental character in the invincible conviction that its sense perceptions are identical with those of all possible subjects, but this persuasion is itself a mere datum.

This intimation of the transcendental and a priori provides a means of passage to the next level of synthesis, the logical level. But the a priori forms of synthesis are not a priori in the Kantian sense; they are "con-natural" with their empirical content. Among these forms are substance and cause, which unify respectively the coexistent and the successive. The movement from the sensible forms of unity to the logical forms is not itself a logical process; rather, it is entirely natural. Logic is the "science of the natural conformations of human thought," and logical relations are therefore empirical relations.

The third stage of synthesis, that of absolute unity, cannot be achieved in thought; it is implied in the dynamic of thought. We can have no speculative concept, but only a symbolic intuition, of it. However, it cannot be concluded, therefore, that our knowledge is limited to phenomena. The absolute unity is always present, although in an imperfect way, because it enters structurally into all levels of synthesis. This omnipresence of the Absolute Martinetti called mystical: "Our knowledge is a mystic unity with the eternal Logos."

This process of synthesis applies also to the practical order, whose transcendental principle is liberty. Morality exhibits a primary synthesis in the form of necessity freely achieved—a synthesis that is continued and extended by art and religion.

See also Absolute, The; A Priori and A Posteriori; Metaphysics; Positivism.

Bibliography

WORKS BY MARTINETTI

Introduzione alla metafisica, 2 vols. Milan, 1902–1904.

Emanuele Kant-Prolegomini. Turin, 1913. With a commentary by Martinetti.

La libertà. Milan, 1928.

Gesù Cristo ed il cristianesimo. Milan, 1934.

Ragione e fede. Turin: G. Einaudi, 1942.

WORKS ON MARTINETTI

Alessio, F. P. *L'idealismo religioso di Piero Martinetti.* Brescia: Morcelliana, 1950.

Gentile, Giovanni. "La teoria della conoscenza del Martinetti." In *Saggi critici,* 1st series. Naples, 1921.

Sciacca, M. F. *Piero Martinetti.* Brescia, 1943. Contains an excellent bibliography.

A. Robert Caponigri (1967)

MARTY, ANTON
(1847–1914)

Anton Marty was a professor of philosophy at the German University of Prague and for forty years a close associate of Franz Brentano. Marty's most important work is the *Untersuchungen zur Grundlegung der allgemeinen Sprachtheorie* (Halle, 1908), a treatise on the philosophy of language. His theory of meaning, or "semasiology," is based upon Brentano's descriptive psychology. From a contemporary point of view, the most interesting aspects of this theory are the distinction between categorematic and syncategorematic uses of words and the theory of emotive utterances.

Like Brentano, Marty appeals to the correctness of affirmation and rejection, and of love and hate (in a broad sense) to explicate the syncategorematic character of certain basic philosophical concepts. In the assertion "There is a horse," the words "a horse" *refer* to an object, but the words "there is" serve only to *express* the fact that the speaker is accepting or acknowledging the object. An object is said to have being if it may be correctly accepted; it has nonbeing if it may be correctly rejected; it is good if

it may be correctly loved; it is bad if it may be correctly hated; the necessary is that which may be correctly accepted a priori; the impossible is that which may be correctly rejected a priori.

Marty rejected the view of Bernard Bolzano and Alexius Meinong, according to which there are objects that may be said to "subsist" and not to "exist." But he did contend that the objects that may be said to "exist" may be classified as being either "real" or "nonreal." Examples of nonreal objects that exist are gaps, deficiencies, holes, space, time, and what Marty called the content of a judgment. (If the judgment "There are horses" is correct, then there exists that nonreal object that is the being of horses; if it is incorrect, then there exists that nonreal object that is the nonbeing of horses.) According to Marty, nonreal objects have no causal efficacy, and their existence is always a function of the existence of certain concomitant real objects. Brentano objected to this view on the ground that sentences ostensibly referring to such nonreal objects may be translated into sentences referring only to the real objects that Marty conceded to be their concomitants ("There is an absence of food in the larder" serves only to express the rejection of food in the larder) and that hence all such "irrealia" are superfluous. But where Marty restricted "real" to a subclass of things that exist, Brentano said that judgments about unicorns are also judgments about "real objects"; these judgments are about things that, if they were to exist, would be real (in Marty's sense of "real").

The word *good,* according to Marty, serves to express one's love of an object; "bad" serves to express one's hate of an object. Marty discussed the emotive function of ethical sentences in detail and noted the ways in which such sentences are related to commands, recommendations, questions, and optatives. However, unlike contemporary emotivists, Marty held with Brentano that the emotions expressed and incited by ethical sentences are emotions that are either correct or incorrect; his theory of ethical sentences could thus be said to be emotive and also objective. He discussed in detail the relations among emotive and nonemotive sentences and the respects in which sentences of the one type may presuppose sentences of the other (for example, a man who calls "Stop thief!" asserts implicitly that there is a thief and that he is trying to get away).

See also Bolzano, Bernard; Brentano, Franz; Emotive Theory of Ethics; Meinong, Alexius; Philosophy of Language.

Bibliography

Marty's posthumously published *Raum und Zeit* (Halle: Niemyer, 1916) sets forth a comprehensive theory of space, time, and causality. His writings also include *Über den Ursprung der Sprache* (Würzburg: Stuber, 1875). *Die geschichtliche Entwicklung des Farbensinnes* (Vienna, 1879); *Die logische, lokalistische und andere Kasustheorien* (Halle, 1910); *Gesammelte Schriften,* edited by Josef Eisenmeir, Alfred Kastil, and Oskar Kraus, 2 vols. (Halle: Niemyer, 1916–1920); and *Nachgelassene Schriften,* edited by Otto Funke (Bern, 1940–1950).

See also Oskar Kraus, *Anton Marty: sein Leben und seine Werke* (Halle, 1916); and *Die Werttheorien* (Brünn: Rudolf M. Rohrer, 1937).

Roderick M. Chisholm (1967)

MARULIĆ, MARKO

(1450–1524)

Marko Marulić, the Croatian poet, historian, and philosopher, was born in Split, Dalmatia. Marulić's epic, *Istorija Svete Udovice Judit* (The History of the Holy Widow Judith; Vinegia, 1521), is the oldest Croatian epic and the first printed Croatian literary work. Like all of Marulić's poetry, it is both epic and didactic. Marulić's philosophical works were written in Latin and translated into German, French, Italian, Portuguese, and other languages. His *De Institutione Bene Beateque Vivendi per Exempla Sanctorum,* first published in Venice in 1506, was reprinted, in the original or in translation, fifteen times in the sixteenth and seventeenth centuries. His *Evangelistarium* (Venice, 1516) was printed nine times.

Marulić was influenced by the Renaissance humanists and was also a student of the classical Greek philosophers, but he was at the same time an outstanding representative of then-modern Christian philosophical thought. He enriched Christian moral teaching with the abundant wealth of Stoic-Platonic moral thought and revived traditional philosophy in the spirit of humanism. Marulić regarded Epicurean and Stoic ethics as antithetically opposed and Stoic ethics as superior to Epicurean. In general, he rejected all forms of hedonism and utilitarianism, and with them ethical subjectivism and relativism.

Marulić's exposition of a Christian ethics combined with elements of Stoicism and Platonism was enlivened by examples from life. This original synthesis of ancient elements, rejuvenated by humanism, was greatly appreciated in its day, especially for its service in the Catholic fight against the Reformation. Although ethical problems

were Marulić's main concern, he also considered the fundamental problems of philosophy.

See also Epicureanism and the Epicurean School; Ethical Subjectivism; Ethics, History of; Hedonism; Humanism; Platonism and the Platonic Tradition; Reformation; Renaissance; Stoicism; Utilitarianism.

Bibliography

For additional philosophical works by Marulić, see *Quinquaginta Parabolae* (Venice, 1510); *De Humilitate et Gloria Christi* (Venice, 1519); and *Dialogus de Laudibus Herculis a Christianis Superacto* (Venice, 1524).

For works on Marulić, see *Zbornik Marka Marulića 1450–1950* (Zagreb, 1950), a commemorative volume honoring Marulić, published by the Yugoslav Academy of Arts and Sciences, with a complete bibliography of Marulić's works and of works about his life and writings.

Vladimir Filipović (1967)

MARX, KARL

(1818–1883)

Karl Marx was born in 1818 in the small German city of Trier in the Rhineland, then part of Prussia. He died in 1883 in London. His life thus spanned the better part of the nineteenth century, a time of rapid and profound economic, social, and political change in Western Europe and America. Philosophically, Marx can be seen as both the culmination of the tradition of German Idealism and its end. In this latter sense, and because most of his work consists of political, economic, and historical analysis, Marx has been taken as having moved beyond purely philosophical interests and investigations into the empirical realms of the social and historical sciences.

The primary goal of Marx's life and work, of course, was to facilitate the revolutionary overthrow of the capitalist system and to help give birth to the socialist society that he believed would inevitably follow the demise of capitalism. In the broadest sense, the project was to achieve the promise of human emancipation, a theme Marx inherited from Kant and Rousseau through German Idealism. Essential to this project was the understanding of the nature and limits of human reason, particularly as embodied in social institutions, a theme of critique also derived from Kant. Marx did not truly leave philosophy behind; he remained a philosopher whose project of liberation led him to increasingly empirical analyses of capitalist society and history. His central concerns are freedom, alienation, and critique, themes at the center of the tradition of German Idealism.

LIFE

Marx came from a Jewish family with rabbinical roots on both his paternal and maternal sides. His father, however, broke with his family and converted to Lutheranism. Karl, his eldest son, was baptized in 1824. After a year studying law at the university in Bonn, Marx transferred to the university in Berlin to study philosophy. He received his doctorate from the University of Jena in 1841, but because of his close association with the radical Young Hegelians, he was unable to secure an academic appointment. Instead of pursuing a career in philosophy, he began to work as a journalist, the only career in which he ever earned any income. Increasingly engaged in the radical politics of the day, in 1843 he moved to Paris, the political heart of Europe, where he did his first serious work in the relatively new field of political economy as well as continuing his critical work on Hegel. This early work, *Economic and Philosophic Manuscripts*, unfinished and not published until 1932, is important to understanding the transformation of Marx the Young Hegelian philosopher into Marx the historical materialist. These manuscripts contain his most extended discussion of alienation, a discussion that helps shed light on how this concept was developed in his later writings, including *Capital*. It was also at this time that Marx established his lifelong friendship and collaboration with Frederick Engels.

While living in Brussels from 1845 to 1848, Marx made his final break with Hegel and the Young Hegelians, including Feuerbach. The two most important pieces of work from this period were the "Theses on Feuerbach" and *The German Ideology* (in collaboration with Engels). Neither was published in Marx's lifetime. These works are often regarded as the first statements of historical materialism and related ideas that would be further developed in Marx's mature thought. What is perhaps Marx's most famous writing, *The Communist Manifesto*, was written with Engels in 1848 at the request the Communist League, an association of revolutionary German workers headquartered in London. Soon after its publication, revolutionary activity burst out across Europe. Eager to participate, Marx went first to Paris and then Cologne, but within a year, as it became clear that the revolution would not succeed, he settled in London. He lived there for the rest of his life.

While not absenting himself entirely from politics, Marx spent the better part of the next fifteen years

immersed in economic theory and history. In an effort to come to terms with recent events in France, he wrote *The Eighteenth Brumaire of Louis Bonaparte* in 1851–1852. Little else of note was published until the end of the decade, when he published *A Contribution to the Critique of Political Economy*. The preface to this work, often referred to as the "1859 Preface," contains his most famous succinct statement of historical materialism. Also during this period, he worked on manuscripts never published during his lifetime that have come to be known as the *Grundrisse*. These notebooks, which did not come to light until the middle of the twentieth century, are important for a number of reasons. They include what is the broadest outline of Marx's theoretical project, an early statement of the themes that became the focus of *Capital*, important points about Marx's method of working on texts, and insights into how Hegelian concepts such as alienation continued to be part of Marx's thinking.

In 1862 and 1863 Marx worked closely on the theories of political economists, Adam Smith and David Ricardo in particular, writing manuscripts later published as *Theories of Surplus Value*. This work culminated in 1867 with the publication of the first volume of his magnum opus, *Capital*. Marx continued to work on the remaining parts of this manuscript, never finishing them to his satisfaction. Engels published them only after Marx's death: Volume 2 in 1885 and Volume 3 in 1894.

Marx returned to more active political involvements in the 1860s, becoming one of the leaders in the The International Working Men's Association, formed in 1864. He remained politically active for the rest of his life, becoming recognized as the leading theoretician of the European working-class movement. Among his later notable writings are *The Civil War in France*, written as an address to mark the demise of the Paris Commune in 1871, and "Marginal Notes on the Program of the German Workers' Party," popularly known as "Critique of the Gotha Program," written in 1875 in an attempt to help unify the two major factions of the German working-class movement. These two later works are important for comments on the nature of society and the state in postrevolutionary, socialist society, a topic about which Marx wrote very little.

FREEDOM, ALIENATION AND CRITIQUE

Marx's philosophical views can be understood in terms of a series of central concepts: freedom, alienation and critique; historical materialism as a dialectical theory; the production of value and the problem of exploitation; and communism and the nature of a free society.

The chief good for Marx, as it was for Hegel, was freedom. For both, a fully free individual was autonomous, and this required rational understanding of and control over one's actions. Both Hegel and Marx appreciated that human emancipation, understood as autonomy, was a collective project. Individuals could be autonomous in the full sense only in a rational and free society. They differed concerning the conditions of a rational society and, in particular, whether the emerging commercial bourgeois society was rational and therefore yielded the conditions for human emancipation.

Marx followed Hegel in arguing that one major impediment to autonomy was a lack of understanding of one's self in relation to one's social world. Such a lack of understanding results in conditions of alienation wherein the individual is dominated to her detriment by states of affairs or objects that she has helped to produce but, in her misunderstanding, treats as independent of her. Conditions of alienation, the young Marx realized, undermined not only the possibility of freedom but created human misery and a sense of meaninglessness. Whereas Hegel analyzed alienation largely as a phenomenon of consciousness, Marx stressed the objective and social roots of alienation, locating its origins in the conditions of production and the nature of labor.

Part of the project of overcoming alienation, Marx realized, involved critique—philosophical analysis that reveals the nature and sources of the alienation and that allows the individual to break through the veils of misunderstanding. Importantly, since the young Marx also realized that alienation was produced by the conditions of social existence, he grasped that until these conditions were understood and overcome, critique alone could not free the world of the destructive consequences of alienation. If the project of emancipation were to be carried to success, Marx recognized that he would have to understand the conditions that give rise to alienation and how those conditions could be changed.

HISTORICAL MATERIALISM

Historical materialism is the theory Marx produced to explain the nature and sources of human alienation, oppression, and suffering and the possibility of attaining emancipation. In its fullest scope, historical materialism supplies an explanation of the central developments of human history, the series of stages of social development through which human societies have passed, and an account of the key dynamics determining the develop-

ment of any given social formation. Marx's theory of capitalism is the most crucial and developed part of the overall theory of historical materialism.

The starting point of historical materialism is the claim that the central project of human history is the production and reproduction of material life. Humans exist within nature as creatures of needs that can only be satisfied through interaction with nature, that is, through labor. The necessity of labor is a manifestation of the fact that the human condition has been one of scarcity. While all animals are in a similar circumstance and must interact with nature to satisfy their needs, humans are distinguished because they have the capacity to develop tools, technology in the broad sense, that allow them to better satisfy their needs. With the development of technology, new needs were created as were the possibility of satisfying them. The production of ever more powerful technology and ways of putting it to use to satisfy an ever greater array of needs and wants, that is, the growth of human productive power, is for Marx the main theme of human history.

Human productive activity involves three elements: raw materials from nature, technology, and human labor. Marx referred to the first two factors, the natural resources and technology, as the means of production. Combined together, the three elements provide the productive power or, as it is more often called, the forces of production. The forces of production, to be put to use, must be organized in terms of some set or other of social relations that determines who has access to and control over the technology, the activity of labor and the product of the labor. Marx refers to these social relations of power as the relations of production. Typically, those who dominate the relations of production appropriate a disproportionate share of the product and dominate society. Two groups of people can be designated: those who dominate the relations of production and have power over the conditions of labor and the product and those who lack control. This division is the basis of Marx's theory of class and the inherently antagonistic class relations of the dominant ruling class and the subordinate workers class.

Historical materialism claims that the forces of production tend to develop in power over time. For any given level of the forces of production, there will be a set of relations of production in terms of which the existing forces can be best utilized and developed. That is, the relations of production that exist at any give point in history will tend to be those that are best suited to the further use and development of the existing forces of production. The existing forces of production together with the set of relations of production in terms of which they are organized form an economic structure that Marx calls a mode of production. Historical development proceeds through determinant stages characterized by the prevailing mode of production. According to Marx, there have been three modes of production prior to capitalism: the ancient slave mode of production characteristic of Greece and Rome, the feudal mode of production, and the Asiatic mode of production that is found in ancient India and China and that, unlike the modes of production found in Europe, does not develop beyond itself.

Within a mode of production, the forces of production continue to develop within the constraints of the existing relations of production. At some point in the development of the forces, the existing relations are no longer optimal to the continued use and development of the forces and the relations break down, allowing a new set of relations to emerge. These points of transition between old and new relations of production are considered revolutionary periods; such periods need not be violent and swift. Marx well understood that the transition from feudalism to capitalism took several centuries. With the emergence of a new mode of production and new relations of production, the nature of power relations within the economic order changes, and a new ruling class comes into being.

All modes of production, then, are made up of a dominant and a subservient class, with the members in the latter class far out numbering those in the former. Given the obvious disparities in power and freedom class society involves, one may ask why have they been as stable for as long as they have? According to historical materialism, the economic or class relations of a society form the basic institution of that society. The other principal institutions, including the political, legal, religious, and cultural, constitute what Marx calls the superstructure of society and justify and reinforce the economic relations. The superstructural institutions that tend to exist at any given point are those that help to stabilize the base. For Marx, just as the level of development of the forces of production determines and explains the nature of the existing relations of production, so the existing relations of production determine and explain the nature of the superstructural institutions.

Part of the superstructure of a society consists of what Marx terms the realm of consciousness; that is, the prevalent ideas and values in a society. As with other aspects of the superstructure, these ideas and values are explained in terms of their role in stabilizing class relations and the base. When such beliefs and values are pro-

duced and propagated by professionals (academics, religious authorities, cultural critics, and the like), Marx refers to them as the ideology of the society. As Marx famously states it: the ruling ideas of an age are the ideas of the ruling class, and they serve the interests of that class. Morality and religion are part of the ideological superstructure, according to Marx. Hence his well-known disdain for them.

Ideological beliefs are not necessarily false, although typically they are. But even when not false, they serve to limit or mislead understanding—for instance, by suggesting that a certain condition is natural and not socially constructed. Thus ideology creates false consciousness. Insofar as the members of the subordinate class accept the ideology of their society, they are misled about the nature of their actions, their society, and the role they play in creating it. In this way, ideological mystification is a major factor in the creation of the experience of alienation and the subsequent loss of freedom. Alienated conditions of existence, conditions that involve the domination of people by the reality they have produced but do not understand, are built into the nature of class society.

THEORY OF EXPLOITATION

In all class societies, the ruling class dominates and exploits the labor of the subordinate class. Such exploitation is fairly evident in slave societies and in feudalism. Capitalism, however, presents a far more complex case. The wage laborer (to use Marx's terms, the proletarian who is a member of the proletariat, the class of wage laborers in capitalism) appears to voluntarily accept work and to be paid for each unit of labor (typically, the hourly wage). The focus of Marx's most sustained work was to unmask this ideological appearance and expose how and why the proletarian was exploited in a way at least as bad, and perhaps worse, than was the slave or serf. By explaining the nature of capitalist exploitation, Marx believed, he could explain the nature and limits of the capitalist mode of production and why it was doomed to be replaced by a socialist society.

Marx's theory of capitalist exploitation is complex; it is grounded on the crucial distinction between labor and labor power. The proletarian, hired by the capitalist, is paid for every hour of labor he performs. What he sells to the capitalist and what the capitalist buys is the capacity of the worker to labor for an agreed upon period of time, say, a ten hour day. During that period, the capitalist owns the worker's capacity to produce goods and can use that capacity in any way he wants. He can use it efficiently, making the worker work harder and produce more, or he can use it less efficiently. Since the capitalist already owns the other factors of production, the raw materials and machines and other technology, and now owns the labor that goes into producing the product, he owns all the factors of production and thus the entire product produced, which he then takes to the market to sell, hoping to return with profit.

Where does this profit come from? Marx asked. The answer resides in determining how commodities, goods produced to be sold in the market, get their prices. Marx used the labor theory of value, taken from Smith and Ricardo, to explain the nature of prices in terms of the labor necessary to produce the commodity. He extended the theory by treating labor power as a commodity that received a price, in this case called the wage, in the same way as other commodities. It is important to note here that human beings can produce under most circumstances more than they need to survive; they can produce a surplus. According to Marx's analysis, the wage (the price of labor power as a commodity) is determined by the value of what is necessary to keep the worker alive and able to work from day to day. The wage does not reflect the value of what the worker is able to produce, which includes both what is necessary and the surplus. Since in virtue of purchasing the worker's labor power and putting it to use as he wishes, the capitalist owns the entire product produced. The capitalist, that is, gets both the value of what is necessary for the worker to have in order to live and the surplus. The capitalist returns to the worker in the form of a wage, however, only the necessary value. He keeps the surplus, and it is this surplus that forms the basis of profit.

Marx noted that, according to the dominant values of capitalism, this exchange between capitalist and proletarian was neither unfair nor coercive. It is as fair and free as any other market exchange. Understanding morality as he did largely in its ideological function, Marx disdained moral critique and did not consider it important to morally condemn the exchange. What was important was to realize that through the process of exploitation, the worker produced, on the one hand, the wealth, privilege and power of the ruling class and, on the other, his own subordination, alienation and misery.

This analysis of the wage and profit is, one might say, Marx's microeconomics with a philosophical intent. His macroeconomic theory attempted to show how capitalism would, with increasing frequency, fall into various crises as the capitalists competing within the essentially anarchistic market struggled to maintain their profit. As

this process continued, the misery of the workers would only increase as well. As the proletariat struggled against worsening misery, their political consciousness would be awakened by the ideologues of their class perspective, political activists and theorists like Marx and Engels. The dual movements, of the capitalists struggling to keep the system going and the workers struggling with increased understanding to overcome it, would eventually culminate in a revolution, ending capitalism and instituting a socialist society. In accordance with the general theory of historical materialism, a successful revolution would happen at or after the point when capitalism was no longer the mode of production best suited to allow optimal use and further development of the forces of production. At that point, socialism would be the best mode.

The dialectical nature of historical materialism is illustrated in the internal dynamics of capitalism and how they are claimed to lead to the overcoming of capitalism. As Marx used the concept, appropriated from Hegel, a theory is dialectical insofar as it reflects and captures a dialectical process in the world. Dialectical processes, typically organic processes, unfold according to a logic of internal development until the present stage of the object or being is fully realized, at which point, again according to the internal logic of development of the object or being, a new stage emerges from the crises and failures of the previous stage. The conditions for the appearance of a successor stage develop in and as a result of the internal developments of the previous stage. Thus the developing nature and struggles of capitalism give rise to a unified and self-conscious proletariat able to forge a new mode of production in its class interests, which happen to be, according to Marx, universal interests.

COMMUNISM

Marx wrote very little about the nature of the mode of production he predicted would displace capitalism. It is clear, though, that he thought human emancipation would be fully realized only in communism, the second stage of postcapitalist society. The first stage following the socialist revolution, referred to at times as socialism, would be dominated by the proletariat—hence the well known phrase, "the dictatorship of the proletariat." Socialism would eliminate the private ownership of the means of production and the exploitation that accompanies private ownership. Technological progress would promote the accelerated development of the forces of production. At some future point, a level of productive forces would be attained that allowed humans to transcend scarcity and enter "the realm of abundance." Abundance refers to a condition in which all can satisfy their needs without depriving others of the satisfaction of needs and without having to spend the greater part of their time in undesirable, unfulfilling labor.

At this stage of human development, communism, all would be free to pursue truly human and creative activities that allowed each individual to fully realize himself or herself. Because all people would have equal access to the means of production, communism would be a classless society. Alienation and exploitation would be abolished. With conflicts over the distribution and fruits of labor eliminated or at least minimized, the primary source of social conflicts would likewise be eliminated, and there would be no need for state authority or for the distorting effects of ideology. There would be no further struggles of the sort that propelled the dialectic of history. Having provided the conditions for full human emancipation, communism would continue to allow optimal development of the forces of production. Hence, no mode of production beyond communism would be necessary or conceivable. Human life as a collective enterprise would gain a self-transparency that would allow humans for the first time to create with knowledge and intention their social fate. In this sense, Marx held that communism would be the end of history, or better perhaps, the beginning of truly human history.

See also Cosmopolitanism; Marxist Philosophy; Postcolonialism; Republicanism.

Bibliography

Avineri, Shlomo. *The Social and Political Thought of Karl Marx*. Cambridge, U.K.: Cambridge University Press, 1968.

Ball, T. and J. Farr, eds. *After Marx*. Cambridge, U.K.: Cambridge University Press, 1984.

Berlin, Sir Isaiah. *Karl Marx: His Life and Environment*. 4th ed. Oxford: Oxford University Press, 1978.

Brundey, Daniel. *Marx's Attempt to Leave Philosophy*. Cambridge, MA: Harvard University Press, 1998.

Buchanan, Allen. *Marx and Justice*. Totowa: Rowman and Littlefield, 1982.

Callinicos, Alex, ed. *Marxist Theory*. Oxford: Oxford University Press, 1989.

Callinicos, Alex. *The Revolutionary Ideas of Karl Marx*. London: Bookmarks, 1983.

Carver, Terrell, ed. *The Cambridge Companion to Marx*. Cambridge, U.K.: Cambridge University Press, 1991.

Carver, Terrell. *Marx's Social Theory*. Oxford: Oxford University Press, 1982.

Cohen, G. A. *History, Labour and Freedom: Themes from Marx*. Oxford: Clarendon Press, 1988.

Cohen, G. A. *Karl Marx's Theory of History: A Defence*. Princeton, NJ: Princeton University Press, 1978.

Cohen, M., T. Nagel, and T. Scanlon, eds. *Marx, Justice and History*. Princeton, NJ: Princeton University Press, 1981.

Draper, Hal. *Karl Marx's Theory of Revolution*. 2 vols. New York: Monthly Review Press, 1977–1978.

Elster, Jon. *Making Sense of Marx*. Cambridge, U.K.: Cambridge University Press, 1985.

Geras, Norman. *Marx and Human Nature*. London: New Left Books, 1983.

Giddens, Anthony. *A Contemporary Critique of Historical Materialism*. Berkeley: University of California Press, 1981.

Kolakowski, Leszek. *Main Currents of Marxism*. 3 vols. Oxford: Oxford University Press, 1978.

Little, Daniel. *The Scientific Marx*. Minneapolis: University of Minneapolis Press, 1986.

Lukes, Steven. *Marxism and Morality*. Oxford: Clarendon Press, 1984.

Marx, Karl. *Capital*. Vol. 1. Translated by Ben Fowkes. Pelican Marx Library. New York: Random House, 1977.

Marx, Karl. *Capital*. Vol. 2. Translated by David Fernbach. Pelican Marx Library. New York: Random House, 1981.

Marx, Karl. *Capital*. Vol. 3. Translated by David Fernbach. Pelican Marx Library. New York: Random House, 1981.

Marx, Karl. *Grundrisse: Foundations of the Critique of Political Economy*. Translated by Martin Nicolaus. Harmondsworth: Penguin Books, 1974.

Marx, Karl. *Marx Engels Collected Works*. New York: International Publishers, 1975–.

Marx, Karl. *Writings of the Young Marx on Philosophy and Society*. Edited and translated by Loyd D. Easton and Kurt H. Guddat. Garden City: Doubleday Anchor Books, 1967.

McLellan, David. *Karl Marx: His Life and Thought*. New York: Harper & Row, 1973.

Miller, Richard. *Analyzing Marx: Morality, Power and History*. Princeton, NJ: Princeton University Press, 1984.

Nielsen, Kai. *Marxism and the Moral Point of View*. Boulder, CO: Westview Press, 1989.

Peffer, R. G. *Marxism, Morality, and Social Justice*. Princeton, NJ: Princeton University Press, 1990.

Roemer, John, ed. *Analytical Marxism*. Cambridge, U.K.: Cambridge University Press, 1986.

Schmidt, Alfred. *The Concept of Nature in Marx*. London: New Left Books, 1971.

Schmitt, Richard. *Alienation and Freedom*. Boulder, CO: Westview Press, 2002.

Wheen, Francis. *Karl Marx*. London: Fourth Estate, 1999.

Wolff, Jonathan. *Why Read Marx Today?* Oxford: Oxford University Press, 2002.

Wood, Allen W. *Karl Marx*. 2nd ed. New York and London: Routledge, 2004.

Wright, Eric Olin, Andrew Levine and Elliot Sober, *Reconstructing Marxism: Essays on Explanation and the Theory of History*. London: Verso, 1992.

Lawrence H. Simon (2005)

MARXISM

Marxist theories, insofar as they are of philosophical interest, are discussed in detail in the entries *Dialectical Materialism*; *Historical Materialism;* and *Marxist Philosophy*. Various Marxist ideas are also discussed in the articles *Alienation*; *Communism*; *Dialectic*; *Ideology;* and *Socialism*. See "Marxism" in the index for thinkers who are usually regarded as Marxists.

MARXIST PHILOSOPHY

Marxist philosophy is the aggregation of philosophical ideas developed from various aspects of Karl Marx's social theory by later thinkers. Marx did not intend to write a philosophy and would have regarded "Marxist philosophy" as a contradiction in terms. He considered his work to be scientific, historical, and sociological, as opposed to "philosophical" divagations on social affairs, which he rejected as class-biased ideology. Moreover, he held that his social theory showed that philosophy was about to end. Philosophy, he said, was a symptom of social malaise and would disappear when revolution put society on a healthier foundation. The young Marx thought that this would happen because revolution would "realize" philosophy, would give solid reality to the ideal phantoms of reason, justice, and liberty that philosophers in sick societies consoled themselves with. The older Marx thought that revolution would destroy philosophy, would simply make it unnecessary, by bringing men back to the study of "the real world." Study of that world is to philosophy "what sexual love is to onanism." In either case Marx never varied in the opinion that the reign of philosophy over men's minds was drawing to a close. Thus, he naturally would not have contributed to its survival by writing a "Marxist philosophy."

MARXISM AND TRADITIONAL PHILOSOPHIES

Within a few years of Marx's death, however, there were attempts to turn Marxism into philosophy. These have continued ever since and, indeed, have gathered force since the discovery of Marx's earliest writings. There are two explanations for this posthumous transformation. First, there is the familiar paradox that efforts to get rid of philosophy by argument are themselves philosophical. Thus, Marx's antiphilosophy and the theory of historical materialism on which it is based blossomed into a veritable philosophical doctrine, to which Georg Lukács gave

consummate form. Second, after the empirical social sciences had taken from Marx's work all that was useful to them (and it was a great deal), there remained much dross—disproven prophecy, hasty generalization, and plain error. Instead of being discarded, as the errors and absurdities of Isaac Newton and Louis Pasteur were discarded in the physical and biological sciences, this nonempirical material was kept alive by a social movement committed to preserving intact the whole of Marx's legacy. It has been called Marxist philosophy.

Because Marxism is not explicitly a philosophy, those who have treated it philosophically have largely sought to find the philosophy to which it "corresponds," from which it "derives," or which it "implies." Solutions have been extremely varied and incompatible. Enrico Ferri put Marxism into the Spencerian system, and Karl Kautsky connected it with Darwinism. Eduard Bernstein and Max Adler found its philosophical complement in Immanuel Kant, and "Back to Kant!" became the slogan of the revisionists. Georgii Valentinovich Plekhanov noted Marx's Hegelian origins but preferred to ally Marxism with materialism, notably that of Ludwig Feuerbach. This opinion was widely accepted by Marxian political activists but was ardently combated by intellectuals. Otto Bauer said that Marxism could not be annexed by materialism because it was compatible with any philosophical doctrine, "including Thomism." Henri de Man essayed a combination of Marx and Freud, whereas the Marburg school of neo-Kantians made a synthesis of Kant's ethics and Marx's socialism. The Russians whom V. I. Lenin attacked in *Materialism and Empirio-Criticism* had married Marxism to the positivism of Ernst Mach and Richard Avenarius. Lenin himself followed Plekhanov in putting Marxism in the tradition of mechanist materialism, later adding a dialectical theory of development to distinguish it from classic materialism. Georges Sorel, René Berthelot, and various Italian writers found the extension of Marxism in pragmatism, and this view became influential in the United States through the writings of Sidney Hook. Antonio Gramsci and Giovanni Gentile, in their different ways, reacted against the "materialist debasement" of Marxism by coupling it with Italian neoidealism. The search for new philosophic settings for Marxism, such as existentialism, continues and is necessarily inconclusive.

The variety of opinions confirms that there is no Marxist philosophy. Nevertheless, some efforts to incorporate Marxism into philosophy are less successful than others, for Marxism is not philosophically neutral even if it does fail to define its position in respect to the major philosophical traditions. Least successful are alliances of Marxism with materialism, from Baron d'Holbach to L. Büchner, or with positivism, whether Mach's or Herbert Spencer's. The tendency of decades of criticism has been to show that the idealist content of Marx's thought is too dominant to allow those confusions. Conversely, the alliance that has proven most fruitful and that has grown in authority over the years is that between Marxism and the Hegelian dialectic. Though Antonio Labriola had noted this, it was ignored for more than a generation until Lukács insisted that Marx belonged in the Hegelian tradition. In this Lukács has been followed by Karl Mannheim, Herbert Marcuse, Lucien Goldmann, Jean-Paul Sartre, and Maurice Merleau-Ponty. Everywhere, Marxism's principal philosophical consequence has been to stimulate the study of G. W. F. Hegel. Otherwise, it has had singularly little effect on philosophy, even on pragmatism, with which it has evident affinities.

ORTHODOX MARXISM

The distinction between a materialist and an idealist reading of Marx does not exactly coincide with the division between the orthodoxy of the Communist parties and the independent criticism of the so-called Western Marxists, but the history of the subject must be told in terms of the latter division. The orthodox tradition begins with Friedrich Engels, not with Marx. It uses two principal texts, Engels's *Anti-Dühring* and Lenin's *Materialism and Empirio-Criticism.* The name of Marx is very seldom mentioned in these discussions, for Marx never explicitly stated the doctrines set out by Engels, taken over and interpreted by Lenin, and then dogmatically systematized by Joseph Stalin. He sometimes appeared to hold opinions resembling those they expressed—for example, the representationist theory of knowledge—yet his early manuscripts seem far removed in spirit from the materialism of these works. That is why the early works, which are the basis of most Marxist philosophy in the West, were dismissed by Soviet writers as juvenile hangovers from Hegelianism that the mature Marx disowned.

EPISTEMOLOGY. Orthodox Marxist philosophy has developed very little over the years, being accepted as much by Rosa Luxemburg as by Lenin, as much by Leon Trotsky as by Stalin, as much by Mao Zedong as by Nikita Khrushchev. Its epistemology is naive representationism: The "concepts in our heads" are images, reflections, or copies of "real things." Objections to that view have been familiar since Bishop Berkeley, but they are held by orthodox Marxists to be answered by a reference to practice. We can compare mental images and the things they

copy by noting our success or failure in manipulating those things. This manipulation is primarily economic activity or is affected by it, so it must differ for each technological age and each class. There is therefore no nonpartisan science. There is a contradiction here, for it is contended that the mind has exact copies of reality and yet its knowledge is historically relative. This is admitted but is circumvented by asserting that absolute knowledge is the historical goal but relative knowledge is the present plight.

METAPHYSICS. In metaphysics the orthodox doctrine distinguishes itself from classic materialism by insisting on dialectic process, as opposed to mechanism, in the development of things. Matter is subject to laws that are causal and determinist but not mechanist. It evolves toward the better and more complex, and it does so in a series of revolutionary jumps, in which accumulations of quantitative difference produce sudden qualitative changes after a period of tension and conflict. Matter is the unique reality. Chance does not exist, and there is no breach in this absolute monism. Mind is an epiphenomenon producing, in consciousness, reflections of matter. Matter does not determine mind directly, as the medical materialists said, but indirectly, by way of society. Society, too, develops dialectically, in revolutionary jumps that resolve its recurrent self-contradictions or internal conflicts. Human liberty consists in awareness of the necessity of social process.

RELIGION, ETHICS, AND AESTHETICS. Religion is doomed to disappear, being a symptom of unjust and self-negating social conditions. Ethics and aesthetics evolve as society changes, for there are no eternal, nonhistorical laws in either. Beauty is objective but appreciation is relative to class, so art is implicated in the class struggle.

In ethics the situation is more complex. At first the exclusion of eternal, suprahistorical laws was held to warrant amoralism, ethical indifference, or at least some experimentation in new ways of living. Soviet authorities found that attitude socially inconvenient, and eventually Stalin formally condemned all applications of historical relativism that suggested that the new polity could have a new ethics (or a special new logic). Since then the position has been that Marxist philosophy substantially accepts the ethical ideals preached in other contemporary societies but adds that only a communist nation can escape hypocrisy by living up to those ideals, by practicing what it preaches. Thus, not only is ethical innovation discouraged in communist countries, but ethical criticism in noncommunist countries—for instance, by existentialists—is strongly deplored as a diversion from the work of creating the social conditions for the application of the uncriticized ethical code common to all modern societies.

WESTERN MARXISM

The Western Marxists, whose first generation, in the 1920s, comprised Lukács, Karl Korsch, Bela Fogarasi, and Josef Revai, rejected the representationist theory of knowledge, but their quarrel with orthodoxy centered on the dialectic. On this issue the orthodox followed Engels, the Westerners the young Hegelian Marx.

Engels had posited the triadic dialectic of thesis, antithesis, and synthesis as an eternal law of cosmic development, applying as much to nature as to mind and society. Everywhere, one would find constant progress from lower to higher by way of objective tensions. The tensions are caused when something engenders its own opposite or negation and are resolved when the opposites merge in a synthesis (the negation of the negation). Engels's immediate successors, whether social democrats, revisionists, Austro-Marxists, or independent students such as Benedetto Croce and Sorel, could make nothing of these ideas and simply ignored the dialectic. At first Lenin did the same, in 1894 dismissing it as a "vestige of Hegelianism." However, he later adopted Engels's dialectic as the badge that distinguished Marxist materialism from classic or vulgar materialism. This dialectic embellishment of materialism has remained a point of honor with subsequent Marxist philosophers even when the dialectic is seldom applied or evoked. The law of the negation of the negation has found little use, and the examples of it offered by Engels, August Thalheimer, and Paul Sandor have been generally rejected by philosophers and scientists. Stalin formally declared that the other law of dialectic, the law of the transformation of quantity into quality, did not have universal scope but applied only to class-divided societies. With the two laws in effect discarded, orthodox Marxist materialism no longer has a characteristic theory of development. There remains only the law of the union of opposites, which serves to reconcile contradictions (and to justify inconsistencies).

The role of the dialectic in Western Marxism is very different. It does not operate in physical nature and is not a law at all. It concerns the relation between mind and social history. That relation comes to the fore because of an evident difficulty encountered by the historical relativism of Marx. If all knowledge is partial, provisional, relative, class-biased, and historically limited, then is this

not true of Marxism itself? The answer of Engels and Lenin was that everything was relative except a small number of absolutely true propositions that included logic and Marxist theory. Seeing the impossibility of maintaining this dualism of relative and absolute knowledge, Lukács abandoned absolute (or unconditionally true) knowledge and accepted the relative and partial character of all knowledge. The relation between our knowledge and all other worldviews that constitute cultural history is a dialectical one, meaning that none is completely true or completely false. More generally, all relations between subject and history are dialectical in the sense of being ambiguous, reciprocal relations that leave room for "contrary and inseparable truths." This is true because, on the one hand, the subject is a social and historical product and, on the other hand, because historical forces are alienated spirit, reified personality. There is conflict and tension between the two terms of that relation, and they will be removed by revolution, which will effect the synthesis of the two and will represent the triumph of the human spirit over the alienation or reification of its products. In this view the crux of historical materialism is the relation between mind and history, the dialectic relation between the personal subject and the apparently impersonal, material forces of society. In showing that those forces are really alienated personality, the theory denounces the objectification of spirit in inhuman institutions. It foresees the victory of spirit over that dehumanization.

Marxist historical materialism, said Lukács, thus criticizes itself according to its own principles. It comes to hold itself as provisional, as, at most, a progress toward a truth that is yet to be attained. Because this relativization seemed to lower Marxism from the status of a dogma to that of one ideology among others, it was no doubt the main reason for the condemnation of Western Marxist philosophy by the orthodox. Yet even the relativism of Lukács (and also of Karl Mannheim) still claims to have dogmatic knowledge of the whole of history, which is the total process into which all partial ideologies fit dialectically and which they all reflect more or less faithfully. With this notion of totality the relativists have brought back the Absolute that they first threw out in favor of the historically relative.

COMMON FEATURES

Because of a dualism in Marx's own thinking, which he never cared to resolve, Marxist philosophy has thus divided into two broad streams. On the one side, there is emphasis on the determinist, evolutionist, materialist, and sociological themes. On the other side, there is the idealist strain that looks forward to the deliverance of humanity from economic determinism. This idealist strain, stressing the primacy of present human activity over the solidified, alienated products of past human activity, has aptly been called titanism by Nikolai Berdyaev. It is a powerful factor in all modern Marxist thought—not only in Western Marxism, where it is explicit, but also in orthodox Soviet Marxism. After a profession of materialist faith, orthodox Marxism introduces the idealist element by attributing to matter a readiness to cooperate with progressive causes. (In other contexts such an attribution of spiritual purposes to matter is called magic.)

The two varieties of Marxist philosophy retain other common features. Both abandon the distinction between truth and falsity in favor of a relativist notion that sees truth as a historical goal and knowledge as never more than progress toward absolute truth. This relativist concept appears in all philosophical developments of Marxism, from Engels to Gramsci and Lukács. Moreover, both sorts of Marxist philosophy cling to the idea of an ultimate reality. Though this is called matter in one case and history in the other, the difference is not great wherever matter has tacitly been endowed with a purposefulness and spirituality (by evolving dialectically) that make it resemble history. Marxism started with the recognition of all things as events or processes that interact, and it emphasized, in the theory of historical materialism, some sorts of interaction that had been overlooked. In its philosophical extensions it has gone on from there to the concept of a moving totality of things to which single things are relative and within which single things have ambiguous, dialectical relations with one another. This view is as familiar to philosophers as the representationist theory of knowledge that Lenin revived and has been as thoroughly criticized. For this reason, among others, Marxist philosophy has seldom secured consideration or academic influence outside of countries where it is politically privileged.

See also Aesthetics, History of; Avenarius, Richard; Berkeley, George; Büchner, Ludwig; Darwinism; Engels, Friedrich; Existentialism; Feuerbach, Ludwig Andreas; Freud, Sigmund; Gentile, Giovanni; Gramsci, Antonio; Hegel, Georg Wilhelm Friedrich; Hegelianism; Holbach, Paul-Henri Thiry, Baron d'; Kant, Immanuel; Kautsky, Karl; Labriola, Antonio; Lenin, Vladimir Il'ich; Lukács, Georg; Mach, Ernst; Mannheim, Karl; Marx, Karl; Materialism; Merleau-Ponty, Maurice; Metaphysics; Neo-Kantianism; Newton, Isaac; Plekhanov,

Georgii Valentinovich; Positivism; Religion; Sartre, Jean-Paul; Sorel, Georges; Thomism.

Bibliography

WORKS BY MARX AND ENGELS

Engels, Friedrich. *Dialektik der Natur.* Berlin, 1927. Translated by Clemens Dutt as *Dialectics of Nature.* New York: International Publishers, 1940. This work was first written in 1872–1873.

Engels, Friedrich. *Herr Eugen Dührings Umwälzung der Wissenschaft.* Leipzig, 1878. Translated by E. Burns as *Herr Eugen Dühring's Revolution in Science.* London, 1935. The *AntiDühring.*

Engels, Friedrich. *Herr Eugen Dühring's der klassischen deutschen Philosophic.* Stuttgart, 1888. Translated as *Ludwig Feuerbach and the Outcome of Classical German Philosophy.* New York: International Publishers, 1934.

Marx, Karl. *Oekonomische-philosophische Ausgabe.* Frankfurt, 1932. Translated by Martin Milligan as *Economic and Philosophical Manuscripts of 1844.* Moscow and London, 1959.

Marx, Karl, and Friedrich Engels. *Die deutsche Ideologic.* Berlin, 1932. Translated anonymously as *The German Ideology.* Edited by S. Ryazanskaya. Moscow, 1964.

MARXIST WRITINGS

Gramsci, Antonio. *Il materialismo storico e la filosofia di Benedetto Croce.* Turin: Einaudi, 1940.

Lenin, N. *Materializm i Empirio-Krititsizm.* Moscow, 1908. Translated by David Kvitko and Sidney Hook as *Materialism and Empirio-Criticism.* New York, 1927.

Lukács, Georg. *Geschichte und Klassenbewusstsein.* Berlin: Malik, 1923. Translated as *Histoire et conscience de classe.* Paris, 1960.

Mannheim, Karl. *Ideologie und Utopie.* Bonn: Cohen, 1929. Translated by Louis Wirth and Edward A. Shils as *Ideology and Utopia.* London: Routledge, 1936.

Mao Zedong. *On Contradiction.* New York: International Publishers, 1953.

Mao Zedong. *On Practice.* New York, 1953.

Mao Zedong. *Selected Works,* 4 vols. New York: International Publishers, 1954–1956.

Plekhanov, Georgi. *Izbranniye Filosofskie Proizvedenia v Piati Tomakh.* Moscow, n.d. Translated as *Selected Philosophical Works.* London, 1961.

Plekhanov, Georgi. *Osnovnie Voprosi Marksizma.* 1910. Translated by Eden and Cedar Paul as *Fundamental Problems of Marxism.* London: M. Lawrence, 1929.

Stalin. *Concerning Marxism and Linguistics.* Moscow, 1950.

Stalin. *Leninism, Selected Writings.* New York, 1938. Contains the essay "Dialectical and Historical Materialism."

WORKS ON MARXISM

Berdyaev, Nikolai. *The Origin of Russian Communism.* Translated by R. French. London: Centenary Press, 1937. This work has never been published in Russian.

Bocheński, I. M. *Der sowjetrussische dialektische Materialismus (Diamat).* Bern and Munich, 1950. Translated by Nicholas Solluhub as *Dialectical Materialism.* Dordrecht, Netherlands, 1963.

Carew Hunt, R. N. *Marxism Past and Present.* New York: Macmillan, 1954.

Fetscher, I. *Der Marxismus: Seine Geschichte in Dokumenten, Vol. I, Philosophic, Ideologic.* Munich: R. Piper, 1963.

Hook, Sidney. *From Hegel to Marx.* London, 1936.

Hook, Sidney. *Towards the Understanding of Karl Marx.* New York: John Day, 1933.

Korsch, Karl. *Marxismus und Philosophie.* Leipzig, 1923; 2nd ed., 1930.

Leningrad Institute of Philosophy. *Textbook of Marxist Philosophy.* Translated by A. Moseley. London, 1937.

Lichtheim, George. *Marxism: An Historical and Critical Study.* London, 1961.

Marcuse, Herbert. *Soviet Marxism: A Critical Analysis.* New York: Columbia University Press, 1958.

Merleau-Ponty, Maurice. *Les aventures de la dialectique.* Paris: Gallimard, 1955.

Vranicki, P. *Historija marksizma.* Zagreb, 1961.

Wetter, Gustav. *Der dialektische Materialismus.* Vienna, 1952. Translated by Peter Heath as *Dialectical Materialism: A Historical and Systematic Survey of Philosophy in the Soviet Union.* New York: Praeger, 1958.

Neil McInnes (1967)

MARXIST PHILOSOPHY [ADDENDUM]

Post–World War II Marxist theory has been decisively shaped by historical changes: the growing irrelevance of orthodox Marxist political movements and the moral and economic decline (and eventual collapse) of the Soviet empire; the emergence of politically radical social movements based in nationalism, gender, and race rather than economic class; changes in the world capitalist economy including the emergence of globalization; and increasing environmental degradation. These developments are reflected in divergent formulations of historical materialism; the adaptation and transformation of Marxism by the new social movements and by seemingly culturally radical postmodern theories; neo-Marxist theories of contemporary capitalism; the cross-fertilization of religion and Marxism in Liberation Theology and its variants; and "eco-Marxism."

Western Marxists such as Herbert Marcuse, the early Jürgen Habermas, and Jean-Paul Sartre resisted the dogmatic and positivist versions of historical materialism found in Marx and in the Second and Third Internationals. These writers denied that a theoretical analysis of capitalist society could provide laws of historical development. Rather, they believed that, at best, economic theory could describe certain continuing contradictions in the social order, the resolutions of which

necessarily depended on the self-awareness and political organization of contending social groups. Given the rise of fascism out of the depression and the triumph of capitalist hegemony over the industrial working classes after World War II, political revolution could no longer be thought of as a direct consequence of predictable economic collapse. It was necessary to investigate social forces that seemed to make the working class not only politically passive but also psychically attached to bourgeois authority.

These forces included not just conscious beliefs, but unconscious personality structures; not just the experience of work, but the experiences of sexuality and family life as well. Consequently, Marxist theory had to encompass psychology and cultural theory as well as economics and politics. It was further claimed that any assimilation of Marxism into a natural-science model was itself an element in political totalitarianism. Habermas (1970) developed this position into a critique of "science and technology as ideology." When we identify social theory with natural science, he argued, we fail to distinguish between science's goal of controlling nature and social theory's goal of understanding and liberating human beings. As a result we end up treating people like things.

French Marxist Louis Althusser posed an influential counterview in 1969 arguing that, while different aspects of society did possess a "relative autonomy" from the economy, it was class structure that always determined historical outcomes "in the last instance." Claiming to present the scientific view of Marxism, and in a move that anticipated later developments of postmodern thought, Althusser asserted that subjectivity was an effect of social structures and not a primary constituent of them.

There have been many subsequent attempts to connect postmodernism and Marxism, including debates about the validity of the former's criticism of totalizing theories and grand historical narratives, about the compatibility of the two perspectives, and about the claim that postmodernism is itself simply "the cultural logic of late capitalism" (Jameson 1991).

Anglo-American philosophy has seen a sophisticated reformulation of some of Marx's original claims about the social primacy of technological development. Analytical Marxist G. A. Cohen developed a "functional" analysis in which a universal human drive to develop forces of production conditioned social relations to change to support such development. Other analytic Marxist philosophers attempted to articulate a distinct moral perspective in Marx to ground claims about the immorality of capitalist exploitation and to critique the individualism of the dominant liberal paradigms of writers such as John Rawls. This discussion has paralleled a rethinking of Marx's relationship to Georg Wilhelm Friedrich Hegel, and of the place of the concept of dialectic in Marxist philosophical and economic theory.

Because of the rise of radical social movements of racial minorities and women, Marxist theory was challenged to integrate accounts of patriarchy and racism with its traditional focus on class exploitation and technological development. Theorists argued that racism and sexism were not reducible to or simple consequences of class power. They were embedded in European culture and conferred certain limited privileges on the white and/or male working class itself. Rather than depending solely on the concept of economic exploitation, or on the traditional Marxist notion that the liberation of the working class would liberate all other subject groups, socialist or Marxist-feminist theorists and black liberationists analyzed the mutually supportive, conflicting, and at times disparate elements of class, racial, and gender domination.

From the 1960s to the 1990s the structural evolution of capitalism led to new versions of Marxist economic and sociological theory. Baran and Sweezy's analysis (1966) revealed how dominant sectors of the economy had become controlled by a small number of firms and that the classic price competition and overproduction oscillations of the nineteenth and early twentieth centuries had consequently given way to stagnation as a result of an unutilizable surplus. Other theorists (e.g., Wallerstein 1974–1980, furthered by Arrighi 1994) redefined capitalism as a capitalist "world-system" constituted by exploitative trade relations between a developed Euro-American core and an underdeveloped periphery.

Many writers claimed that the increased role of the state in the national economy mitigated the business cycle and redirected class struggles to competition over state resources. James O'Connor (1973) foresaw that contradictions between state support of capitalist accumulation and democratic legitimation would eventually cause a "fiscal crisis of the state." Habermas (1975), writing under the shadow of the political uprisings of the 1960s and 1970s, described conflict between ideals of democracy and equality and state support of capitalist accumulation as causing a "legitimation crisis."

Responding to the continued dominance of capitalism and the failure of almost all state-controlled communism, theorists of socialism have also raised the possibility of alternative forms of a socialist economy, especially a socialism in which consumer demand is allo-

cated by markets but is not at the same time controlled by private ownership of the forces of production. For a number of writers, the key issue is no longer the traditional idea of "socialization of the means of production," which to some extent left open the question of the structure of that socialization, but rather that of "economic democracy," which identifies the structure of socialization with certain fundamental political, as well as economic, changes (Schweickart 2002).

One defining aspect of contemporary society is the worsening environmental crisis, which Marxist theoreticians have responded to not only by using familiar Marxist concepts to explain it, but by positing, as did O'Connor in 1988, an "eco-Marxist" analysis in which capitalist destruction of the environment becomes the "Second Contradiction of Capital." In this view capitalism's tendency to destroy its own physical basis of production (through ecological devastation) now coexists with the resistance it generates from the labor force as a major source of its own undoing. While some authors, such as Andrew McLaughlin (1993) accuse Marxism of an industrialism that is a major source our environmental problems, Jonathan Hughes (2000) and others argue that Marxism is crucial to understanding how to solve them.

On a much different front, several Latin American theorists—inspired by the renewed emphasis on social justice of Vatican II (1964–1965), the spread of communist movements, and the appalling degree of poverty and repression surrounding them—created what they called "liberation theology" (Gutiérrez 1988). This was an attempt to join Christian social ethics (most importantly the "preferential option for the poor") with Marxist social theory. Essential to liberation theology was the belief that "the poor" could only enter history as fully human beings if a fundamental social transformation—virtually a social revolution (preferably without violence)—were to occur. While the development of liberation theology would certainly have surprised Marx—and eventually prompted stiff resistance from the church hierarchy—it provided a model for the cross-fertilization of politically left ideas with religious moral concerns. The use of a quasi-Marxist vocabulary was found in Martin Luther King, who both criticized imperialism and called for a "beloved community," and in the appearance of politically radical forms of theology focusing on women, race, gay and lesbian issues, poverty and war (see Gottlieb 2003). In the 1990s theologians' and institutionalized religion's dramatically sharpening response to the environmental crisis led to positions increasingly resonant with Marxism. Frequently their criticisms of the market, capitalism, and the global economy would not be out of place in a professedly Marxist journal or socialist party. Conversely, some (e.g., Gottlieb 2002) have argued that a sustained engagement with religious values of nonviolence, universal respect, self-examination, and humility would help compensate for widely shared limitations found in politically radical, including Marxist, perspectives.

Since the last decade of the twentieth century, all of these discussions have been shaped by globalization—an economic, social, and cultural phenomenon the very definition of which is the subject of intense debate. Early on, both the monopoly capital and the world-system models were challenged by the "global capitalism" perspective (Ross and Trachte 1990), which diagnosed an international economy dominated by multinational firms, intranational competition rather than a dominant Euro-American core, and increased power for capitalists as international mobility allows them to evade local labor movements, governments, and environmental regulations. In the twenty-first century, different accounts within a broadly Marxist paradigm include Hardt and Negri's analysis (2001) of globalization as a new form of imperialism or "empire" and Manuel Castells's discussion of globalization as a "network society." For Castells traditional forms of national sovereignty have become increasingly less relevant, and the separation of decisive social power from even remotely local control causes a dramatic resurgence in the importance of ascribed social identities, especially fundamentalist religion. The importance of globalizing economic structures has also led to increased interest in multidimensional global political resistance movements, which typically involve the working class, but also include peasants, community organizations, environmentalists, students, and progressive middle class. The possibility of an alternative to globalization, and the possible forms of that alternative, are perhaps the most critical questions for contemporary Marxism.

Despite the enormous variety of Marxist and Marxist-related writings since the middle of the twentieth century, there are significant areas that remain relatively unexplored. These include the relation of disability to other forms of social marginalization; sustained and honest examination of the subjective or psychological sources of leftist political failures (for example, how the character structure or ethical orientation of activists has caused destructive intragroup conflict and unnecessary antagonism of social groups outside the left); and a willingness

to question the Marxist premise that human beings are essentially rational.

In sum, Marxism continues to evolve and mutate, with many of its basic concepts (the critique of ideology, the analysis of capitalism) still essential to socially critical perspectives such as postmodernism and feminism. If it is now virtually impossible to delineate any simple Marxist orthodoxy, or to say where Marxism ends and other left perspectives begin, one can (as in other intellectual traditions) trace the historical roots of philosophical perspective and revolutionary social intent from Marx, through enormous historical change, to the Marxisms of the present.

See also Civil Disobedience; Communism; Cosmopolitanism; Fascism; Feminist Social and Political Philosophy; Habermas, Jürgen; Historical Materialism; Liberation Theology; Marx, Karl; Modernism and Postmodernism; Multiculturalsim; Postcolonialism; Racism; Rawls, John; Republicanism; Sartre, Jean-Paul; Socialism.

Bibliography

Althusser, Louis. *For Marx*. Translated by Ben Brewster. New York: Random House, 1969.

Arrighi Giovanni. *The Long Twentieth Century: Money, Power, and the Origins of Our Times*. London: Verso Books, 1994.

Baran, Paul A., and Paul M. Sweezy. *Monopoly Capital*. New York: Monthly Review Press, 1966.

Buchanan, Allen E. *Marx and Justice: The Radical Critique of Liberalism*. Totowa, NJ: Rowman & Littlefield, 1984.

Castells, Manuel. *The Information Age*. 3 vols. Cambridge, MA: Blackwell Publishers, 1996–2001.

Cohen, G. A. *Karl Marx's Theory of History: A Defence*. Princeton, NJ: Princeton University Press, 1978.

Eisenstein, Z., ed. *Capitalist Patriarchy and the Case for Socialist Feminism*. New York: Monthly Review Press 1979.

Gottlieb, Roger. S. *Joining Hands: Politics and Religion Together for Social Change*. Cambridge, MA: Westview, 2002.

Gottlieb, Roger S., ed. *Liberating Faith: Religious Voices for Justice, Peace and Ecological Wisdom*. Lanham, MD: Rowman & Littlefield, 2003.

Gutiérrez, Gustavo. *A Theology of Liberation: History, Politics, and Salvation*. Rev. ed. Maryknoll, NY: Orbis Books, 1988.

Habermas, Jürgen. *Legitimation Crisis*. Translated by Thomas McCarthy. Boston: Beacon Press, 1975.

Habermas, Jürgen. *Towards a Rational Society*. Translated by Jeremy J. Shapiro. Boston: Beacon Press, 1970.

Hardt, Michael, and Antonio Negri. *Empire*. Cambridge, MA: Harvard University Press, 2000.

Harvey, David. *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*. London: Blackwell Publishers, 1990.

Hughes, Jonathan. *Ecology and Historical Materialism*. New York: Cambridge University Press, 2000.

Jameson, Fredric. *Postmodernism, or the Cultural Logic of Late Capitalism*. Duke University Press, 1991.

Kuhn, Annette, and Ann Marie Wolpe, eds. *Feminism and Materialism: Women and Modes of Production*. London: Routledge and Kegan Paul, 1978.

Le Grand, Julian, and Saul Estrin, eds. *Market Socialism*. Oxford: Oxford University Press 1989.

Marcuse, Herbert. *Eros and Civilization: A Philosophical Inquiry into Freud*. New York: Vintage 1962.

Marcuse, Herbert. *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Boston: Beacon 1964.

McLaughlin, Andrew. *Regarding Nature: Industrialism and Deep Ecology*. Albany: State University of New York Press, 1993.

O'Connor, James. "Capitalism, Nature, Socialism: A Theoretical Introduction." *Capitalism, Nature, Socialism* 1 (1988).

O'Connor, James. *The Fiscal Crisis of the State*. New York: St. Martins, 1973.

Ollman, Bertell. *Dialectical Investigations*. New York: Routledge, 1993.

Reich, Michael. *Racial Inequality: A Political-Economic Analysis*. Princeton, NJ: Princeton University Press 1981.

Ross, Robert J. S., and Kent C. Trachte. *Global Capitalism: The New Leviathan*. Albany: State University of New York Press, 1990.

Ryan Michael. *Marxism and Deconstruction: A Critical Articulation*. Baltimore, MD: Johns Hopkins University Press, 1982.

Sergent, Lydia, ed. *Women and Revolution*. Boston: South End Press 1981.

Sartre, Jean-Paul. *Critique of Dialectical Reason*. Translated by Alan Sheridan-Smith; edited by Jonathan Rée. London: New Left Books, 1976.

Schweickart, David. *After Capitalism*. Lanham, MD: Rowman & Littlefield, 2002.

Smith, Tony. *The Logic of Marx's Capital: Replies to Hegelian Criticisms*. Albany: State University of New York Press, 1990.

Wallerstein, Immanuel. *The Modern World-System*. 2 vols. New York: Academic Press 1974–1980.

Roger S. Gottlieb (1996, 2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

MASARYK – NUSSBAUM

6 *volume*

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT

Editor in Chief

MACMILLAN REFERENCE USA

An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON
GALE™

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

ENCYCLOPEDIA OF
PHILOSOPHY
2nd edition

M

MASARYK, TOMÁŠ GARRIGUE
(1850–1937)

Tomáš Garrigue Masaryk, a Czech statesman and philosopher, and president of Czechoslovakia from 1918 to 1935, was born in Hodonín, Moravia. His political career belongs to history; of interest to students of philosophy is the fact that he studied philosophy at the University of Vienna from 1872 to 1876 under Franz Brentano. He spent the year 1876–1877 at Leipzig, where Wilhelm Wundt was his teacher and Edmund Husserl and Richard Avenarius were fellow students. In 1879 Masaryk became *Privatdozent* at Vienna, submitting *Der Selbstmord als sociale Massenerscheinung* (Vienna, 1881) as his habilitation thesis. In 1882 Masaryk became professor of philosophy at the Czech University in Prague, where he soon made his mark as a politician and writer in Czech. *Základové konkretné logiky* (The foundations of concrete logic; Prague, 1885; German translation, *Versuch einer concreten Logik,* Vienna, 1887) and *Otázka sociální* (The social question; Prague, 1898; German translation, *Die philosophischen und sociologischen Grundlagen des Marxismus,* Vienna, 1899) were followed by books on Czech history and politics and by an extensive Russian intellectual history, first published in German as *Russland und Europa* (2 vols., Jena, Germany, 1913; translated by Eden and Cedar Paul as *The Spirit of Russia,* 2 vols., London, 1919). World War I and the presidency of Czechoslovakia put an end to Masaryk's academic pursuits, but a book of memoirs, *Světová revoluce* (The world revolution; Prague, 1925; English translation, edited by H. W. Steed, *The Making of a State,* London, 1927) and *Hovory s T. G. Masarykem* (Conversations with T. G. Masaryk; 3 vols., Prague, 1931–1935) by Karel Čapek (English translations by M. and R. Weatherall, *President Masaryk Tells His Story,* London, 1934, and *Masaryk on Thought and Life,* London, 1938) reformulate his convictions impressively.

Masaryk was a practical philosopher who believed that philosophy should not only contemplate the world but also try to change it. He thus had little interest in problems of epistemology or cosmology. In his early life he reacted against German idealism and accepted British empiricism (David Hume) and French positivism (Auguste Comte). Later he argued for a type of realism that he called concretism. In every act of knowing, he believed, the whole man takes part. Concretism acknowledges not only reason but also the senses, the emotions, and the will—the whole experience of our consciousness. It is something like William James's radical empiricism

without the exceptional experiences admitted by James. But Masaryk's main interest was in sociology and philosophy of history.

Masaryk's realism was combined with a deep religious belief—Masaryk was a theist who found the Unitarianism of his American wife congenial—and a strong conviction of the immutable difference between right and wrong. Masaryk's thinking centered on the crisis of civilization caused by the decay of religion. He diagnosed the diseases of modern man (indifference, suicidal mania, violence, war, etc.) and prescribed remedies for them. He believed that sociology is the foundation of any further cultural advance but that its method must not be purely genetic and descriptive. Teleology, or explanation by purpose, is legitimate. The aim of history is the realization of the ideal of humanity. Masaryk's humanism was not, however, merely humanitarianism, although he often spoke of democracy as another term for his ideal. In spite of his sympathies for the concrete demands of socialism, Masaryk remained an individualist who disapproved of all forms of collectivism. He criticized Karl Marx as a blind worshiper of determinist science. Nevertheless, Masaryk exalted the role of the right kind of science. In *Základové konkretné logiky,* his philosophically most ambitious book, he classified the sciences and showed how they are internally related and coordinated. The task of philosophy is to create a worldview based on the results of the sciences. Masaryk desired a new "Advancement of Learning" that would save man from intellectual and moral anarchy.

Masaryk assigned an important role in the realization of his ideal to his own nation, the Czech, and interpreted its history, remembering the Hussites and the Bohemian Brethren as a preparation for this task. He thoroughly criticized Russia for being a breeding ground for all the European diseases, particularly romanticism and materialism. Fëdor Dostoevsky, whom he both admired and rejected as a thinker, was a lifelong concern. Masaryk always expressed the deepest sympathies for the English and American tradition of empiricism and moralism and, in politics, turned his nation resolutely toward the Anglo-Saxon West. In 1918 he liberated the Czechs not only politically but also intellectually.

See also Avenarius, Richard; Brentano, Franz; Comte, Auguste; Dostoevsky, Fyodor Mikhailovich; Empiricism; Humanism; Hume, David; Husserl, Edmund; James, William; Marx, Karl; Philosophy of History; Positivism; Teleology; Wundt, Wilhelm.

Bibliography

For information on Masaryk as a thinker, see the bibliography and articles in *Festschrift Thomas G. Masaryk zum 80. Geburtstag,* 2 vols., edited by B. Jakowenko (Bonn, 1930); W. P. Warren, *Masaryk's Democracy* (Chapel Hill: University of North Carolina, 1941); and René Wellek, "Masaryk's Philosophy," in *Essays on Czech Literature* (The Hague: Mouton, 1963).

René Wellek (1967)

MASS

The mass of a body is its inertia or resistance to change of motion. More precisely, it is a property of the body that determines the body's acceleration under the influence of a given force. Mass can therefore be measured either by the amount of force necessary to impart to the body a given motion in a given time or by the acceleration produced by a given force.

The absolute metric unit of mass is the gram, which is the mass of a body whose velocity increases by one centimeter per second each second if acted upon by a force of one dyne. Other common units are the kilogram (1,000 grams) and the pound (453.592 grams). For velocities that are small as compared with the speed of light, the mass of a body is a constant, characteristic of the body and independent of its location—in contrast to weight, which varies with the body's place on Earth or in the universe.

Although fundamental to science and, together with length and time, the basis of all measurements in physics, the concept of mass was unambiguously defined only at the end of the nineteenth century. However, its rudimentary sources, systematically employed long before by Isaac Newton and to some extent already by Johannes Kepler, can be traced back to early Neoplatonic ideas concerning the inactivity of matter as opposed to the spontaneity of mind. The ancient metaphysical antithesis of matter and spirit served as a prototype of the physical contrast of mass and force.

CONCEPT OF INERTIAL MASS

Antiquity, and Greek science in particular, had no conception of inertial mass. Even the idea of quantity of matter (*quantitas materiae*), the antecedent of inertial or dynamic mass, was foreign to the conceptual scheme of Aristotelian natural philosophy. Paradoxically, it was Neoplatonism and its admixtures of Judeo-Christian doctrines, with their emphasis on the spiritual and immaterial nature of reality, that laid the foundations for the

inertial conception of mass, which later became the basic notion of materialistic or substantial philosophy. To accentuate the immaterial, sublime source of all force and life in the intellect or God, Neoplatonism degraded matter to impotence and endowed it with inertia in the sense of an absolute absence of spontaneous activity. For Plotinus, Proclus, Philo, Ibn Gabirol, and the Platonic patristic authors, matter was something base, inert, shapeless and "plump," attributes that reappear in Kepler's characterization of matter as that which is too "plump and clumsy to move itself from one place to another."

The idea of a quantitative determination of matter different from, and ontologically prior to, spatial extension originated in scholastic philosophy in connection with the problem of the transubstantiation. The question of how accidents of condensation or rarefaction (volume changes) can persist in the consecrated *hostia* of the holy bread and wine of the Eucharist whereas the substances of the bread and the wine change into the Body and the Blood of Christ led Aegidius Romanus, a disciple of Thomas Aquinas, to the formulation of his theory of *duplex quantitas.* According to this theory matter is determined by two quantities; it is "so and so much" (*tanta et tanta*) and "occupies such and such a volume" (*et occupat tantum et tantum locum*), the former determination, the *quantitas materiae,* having ontological priority over bulk. Aegidius's early conception of mass as quantity of matter, expounded in his *Theoremata de Corpore Christi* (1276), was soon renounced and had little influence on the subsequent development of the concept of mass. It was primarily Kepler who ascribed to matter an inherent propensity for inertia in his search for a dynamical explanation of the newly discovered elliptical orbits of planetary motion; in need of a concept expressing the opposition intrinsic in matter to motory forces, Kepler formulated the inertial concept of mass. In his *Epitome Astronomiae Copernicanae* (1618) he declared that "inertia or opposition to motion is a characteristic of matter; it is stronger the greater the quantity of matter in a given volume."

A different approach to the same idea arose from the study of terrestrial gravitation. As soon as gravity was regarded no longer as a factor residing in the heavy body itself, as Aristotle taught, but as an interaction between an active principle, extraneous to the gravitating body, and a passive principle, inherent in matter, as Alfonso Borelli and Giovanni Baliani (author of *De Motu Gravium,* 1638) contended, the notion of inertial mass became a necessity for a dynamical explanation of free fall and other gravitational phenomena. Furthermore, Christian Huygens's investigations of centrifugal forces (*De Vi Centrifuga,* 1659; published in Leiden, 1703) made it clear that a quantitative determination of such forces is possible only if with each body is associated a certain characteristic property proportional to, but conceptually different from, the body's weight. Finally, the systematic study of impact phenomena, carried out by John Wallis, Sir Christopher Wren, and Huygens, enforced the introduction of inertial mass. With Newton's foundations of dynamics (*Principia,* 1687) these four categories of apparently disparate phenomena (planetary motion, free fall, centrifugal force, and impact phenomena) found their logical unification, through his consistent employment of the notion of inertial mass. Newton's explicit definition of this concept, however, as "the measure of quantity of matter, arising from its density and bulk conjointly" was still unsatisfactory from both the logical and the methodological points of view. It was probably the influence of Kepler or of Robert Boyle and his famous experiments on the compressibility of air that made Newton choose the notion of density as a primary concept in his peculiar formulation of the definition of mass, a formulation that was severely criticized in modern times, especially by Ernst Mach and Paul Volkmann.

LEIBNIZ AND KANT

Gottfried Wilhelm Leibniz's original conception of mass (1669), in contrast to Newton's, defined it as that property which endows primary matter with spatial extension and antitypy, or impenetrability. In his later writings, especially in his doctrine of monads, Leibniz associated mass with secondary matter and saw in it a property of a collection of substances (monads) resulting from their being a collection. Finally, recognizing the insufficiency of purely geometric conceptions to account for the physical behavior of interacting bodies, Leibniz departed from the Cartesian approach and accepted the dynamic, or inertial, conception of mass. The trend of Leibniz's ideas was brought to its final consequences by Immanuel Kant, with his rejection of the Newtonian *vis inertiae,* the dynamic opposition against impressed force. Refuting its legitimacy on the ground that "only motion, but not rest, can oppose motion," Kant postulated the law of inertia as corresponding to the category of causality ("every change of the state of motion has an external cause") and consequently defined mass as the amount of the mobile (*die Menge des Beweglichen*) in a given volume, measured by the quantity of motion (*Die metaphysischen Anfangsgründe der Naturwissenschaft,* 1786).

DEFINITION OF MASS

Under the influence of the Kantian formulation, often incompletely understood, and primarily owing to the fact that in spite of the universal use of the concept in science as well as in philosophy no clear-cut definition of mass was available, most authors defined mass as quantity of matter without specifying how to measure it. Toward the middle of the nineteenth century, with the rise of modern foundational research and the critical study of the principles of mechanics, the logical deficiency of such definitions became obvious. It was primarily Ernst Mach, preceded by Barré de Saint-Venant and Jules Andrade, who insisted on the necessity of a clear operational definition of mass. In an essay, "Über die Definition der Masse" (1867; published in 1868 in *Carl's Repertorium der Experimentalphysik,* Vol. 4, pp. 355–359), and in the *Science of Mechanics* (*Die Mechanik in ihrer Entwicklung, historisch-kritisch dargestellt,* Leipzig, 1883; translated by T. J. McCormack, La Salle, IL, 1942), Mach defined the ratio of the masses of two bodies that interact with each other but are otherwise unaffected by all other bodies in the universe as the inverse ratio of their respective accelerations ($m_1/m_2 = a_2/a_1$), thereby converting Newton's third law of action and reaction to a definition of mass. If a particular body is chosen as the standard unit of mass, the mass of any other body can be unambiguously determined by simple physical operations. The practical method of comparing masses by weighing is, of course, operationally still simpler but logically more complicated, since the notion of weight presupposes that of mass. Although Mach's definition is not quite unobjectionable, it has gained great popularity and is generally adopted in modern texts in science.

INERTIAL AND GRAVITATIONAL MASS

In addition to its inertial mass, every physical body possesses gravitational mass, which determines, in its active aspect, the strength of the gravitational field produced by the body and, in its passive aspect, the amount by which the body is affected by the gravitational field produced by other bodies. According to Newton's law of universal gravitation, the force of attraction is proportional to the inertial masses of both the attracting and the attracted bodies. The resulting proportionality of inertial and gravitational masses of one and the same body, experimentally confirmed by Newton, Friedrich Bessel, Roland von Eötvös, and others, remained in classical physics a purely empirical and accidental feature, whereas the strict proportionality between the active and the passive gravitational masses is a straightforward consequence of Newton's third law of action and reaction or, alternatively, of the very definition of inertial mass if the postulated interaction is of gravitational nature. In general relativity, however, the so-called principle of equivalence, which maintains the unrestricted equivalence between uniformly accelerated reference systems and homogeneous gravitational fields, implies the fundamental identity between inertial and passive gravitational masses. In addition, it can be shown that on the basis of general relativity the active gravitational mass of a body or dynamical system equals its inertial mass, so that in relativistic physics, in contrast to Newtonian physics, the identity of all three kinds of masses is a necessary consequence of its fundamental assumptions.

MASS AND ENERGY

Whereas general relativity led to an important unification of the concept of mass, special relativity already, with Albert Einstein's paper *Does the Inertia of a Body Depend upon Its Energy Content?* (1905; reprinted in *The Principle of Relativity,* New York, 1923), led to a vast generalization of the concept by showing the equivalence of mass and energy insofar as a body emitting radiative energy of an amount E loses mass to an amount of E/c^2, where c is the velocity of light. Subsequent research, especially in connection with energy transformations in nuclear physics, supported the general validity of the formula $E = mc^2$, according to which mass and energy are interconvertible and one gram of mass yields 9×10^{20} ergs of energy. It also became obvious that Antoine Lavoisier's law of the conservation of mass (1789) and Robert Mayer's (or Hermann Helmholtz's) law of the conservation of energy were only approximately correct and that it was the sum total of mass and energy that was conserved in any physicochemical process.

INFLUENCE OF THE ELECTROMAGNETIC CONCEPT

The way to these far-reaching conclusions of relativity had been prepared to some extent already by the introduction of the electromagnetic concept of mass at the end of the nineteenth century (by J. J. Thomson, Oliver Heaviside, and Max Abraham). It seemed possible on the basis of James Clerk Maxwell's electromagnetic theory to account for the inertial behavior of moving charged particles in terms of induction effects of purely electromagnetic nature. Walter Kaufmann's experiments (1902) on the deflection of electrons by simultaneous electric and magnetic fields and his determination of the slightly variable inertial mass of the electron seemed at the time to

support the hypothesis that the mass of the electron, and ultimately the mass of every elementary particle, is of purely electromagnetic nature. Although such eminent theoreticians as H. A. Lorentz, Wilhelm Wien, and Henri Poincaré accepted these ideas, according to which the whole universe of physics is but an interplay of convection currents and their radiation, with physical reality stripped of all material substantiality, the electromagnetic conception of mass had to make way for the relativistic concept as outlined above. Certain aspects of the electromagnetic conception of mass did survive, however, and reappeared in modern field theories—in particular the fundamental tenet that matter does not do what it does because it is what it is, but it is what it is because it does what it does.

See also Aristotle; Boyle, Robert; Energy; Ibn Gabirol, Solomon ben Judah; Kant, Immanuel; Kepler, Johannes; Leibniz, Gottfried Wilhelm; Mach, Ernst; Maxwell, James Clerk; Neoplatonism; Newton, Isaac; Patristic Philosophy; Philo Judaeus; Plotinus; Poincaré, Jules Henri; Proclus; Thomas Aquinas, St.

Bibliography

Bainbridge, K. T. "The Equivalence of Mass and Energy." *Physical Review* 44 (1933): 123.

Comstock, D. F. "The Relation of Mass to Energy." *Philosophical Magazine* 15 (1908): 1–21.

Jammer, Max. *Concepts of Mass in Classical and Modern Physics.* Cambridge, MA: Harvard University Press, 1961; Mineola, NY: Dover, 1997.

Lampa, A. "Eine Ableitung des Massenbegriffs." *Lotos* 59 (1911): 303–312.

Mach, Ernst. *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit.* Prague, 1872.

Pendse, C. G. "On Mass and Force in Newtonian Mechanics." *Philosophical Magazine* 29 (1940): 477–484.

Whittaker, E. T. "On Gauss' Theorem and the Concept of Mass in General Relativity." *Proceedings of the Royal Society,* A, 149 (1935): 384–395.

M. Jammer (1967)

MATERIALISM

Materialism is the name given to a family of doctrines concerning the nature of the world that give to matter a primary position and accord to mind (or spirit) a secondary, dependent reality or even none at all. Extreme materialism asserts that the real world is spatiotemporal and consists of material things and nothing else, with two important qualifications: first, space and time, or spacetime, must also be included if these are realities rather than mere systems of relations, for they are not material things in any straightforward sense. Second, materialism is fundamentally a doctrine concerning the character of the concrete natural world we inhabit, and it is probably best to set to one side controversies over abstract entities such as numbers, or geometric figures, or the relations of entailment and contradiction studied in logic. A strictly extreme materialism would undertake to show that, to the extent that any of these were genuine realities, they are all material in nature, but the issues raised by abstract entities will not be pursued here. It is with extreme materialist views in the concrete realm that this entry is concerned, and in what follows, "materialist" is to be understood in that sense.

Philosophers and scientists have had various views regarding the constitution and behavior of material objects and over whether every material thing is a body, or whether forces, or waves, or fields of force are also realities in their own right. Thus, the cardinal tenet of materialism, "Everything that is, is material," covers a range of different claims.

To accommodate these differences, a material thing can be defined as a being possessing many physical properties and no other properties, or as being made up of parts all of whose properties are physical. The physical properties are position in space and time, size, shape, duration, mass, velocity, solidity, inertia, electric charge, spin, rigidity, temperature, hardness, magnetic field intensity, and the like. The phrase "and the like" is important, for it indicates that any list of physical properties is open-ended. A material thing is one composed of properties that are the object of the science of physics. And physics is a developing science, in which new properties are still being discovered. The question "What counts as a physical property?" thus has no determinate answer. In consequence, there are also no fully determinate answers for the questions "What is a material thing?" and "What does materialism claim?"

This is less serious for materialism than may at first appear, for there is a broad consensus on which properties—among those already known—are the physical ones. And new properties that emerge from research in the physical sciences are, generally speaking, readily identified as belonging among the physical ones rather than representing an anomalous, nonphysical development. It is known well enough what is involved in claiming that something is a material reality, and therefore it is understood well enough what is involved in the various ver-

sions of extreme materialism, all of which assert that everything there is, is material.

The psychological characteristics people ascribe to themselves and to one another—consciousness, purposiveness, aspiration, desire, and the ability to perceive, for example—are not considered to be physical properties. So materialism differs from panpsychism, the doctrine that everything material is also at least partly mental or spiritual. Materialism denies the world's basic entities possess these psychological properties. Materialists add that there is no second class of nonmaterial beings in possession of such psychological properties and no others; there are no incorporeal souls or spirits, no spiritual principalities or powers, no angels or devils, no demiurges and no gods (if these are conceived as immaterial entities). Hence, nothing that happens can be attributed to the action of such beings.

The second major tenet of materialism is, accordingly, "Everything that can be explained can be explained on the basis of laws involving only the relevant physical conditions." The differences among materialists over the types of effect material things can have on one another make this second tenet another slogan covering a variety of particular doctrines. Further, although materialists have traditionally been determinists, holding that there is a physical cause for everything that happens, this is not strictly required by materialism itself. Recently, the appeal of determinism has been weakened by the development and success of quantum theory, and many contemporary materialists are not committed to determinism. It should also be mentioned that metaphysical materialism in no way involves an overzealous disposition to pursue money and tangible goods, despite the popular use of "materialistic" to describe this interest.

NATURE AND APPEAL OF MATERIALISM

The enduring appeal of materialism arises from its alliance with those sciences that have contributed most to an understanding of the world humans inhabit. Investigations in the physical sciences have a materialist methodology; that is, they attempt to explain a class of phenomena by appeal to physical conditions alone. The claim of materialists is that there is no subject matter that cannot be adequately treated with a materialist methodology. This claim cannot be established by any scientific investigation; it can be established, if at all, only by critical reflection on the whole range of human thought and experience.

Early philosophers proceeded dogmatically, aiming to prove the material nature of the world by mere reflection on what must be. Contemporary materialists are much more modest, offering the claim as a speculative but reasonable generalization from the progress of the physical sciences.

Materialism has been, traditionally, a minority view, indeed a rather daring and scandalous one, but it has made considerable progress over the past century, particularly among educated European peoples. There seem to be three main reasons for this. First, the rise of what might be called "cosmic naturalism"; there has been a decline in those aspects of religious conviction that involve appeal to providential or satanic interventions in the course of events, so that pestilence or climate change, for example, are not attributed to nonmaterial, supernatural forces. Second, the rise of "medical materialism"; the discovery of the biochemical mechanisms involved in neural functioning, and their links to psychological processes, so that it is now taken for granted that thinking, feeling, and the will are subserved by the nervous system, and can be altered by making physical changes by the use of drugs or electrodes. A malfunction of the mind is taken to be a malfunction of the brain. This is a kind of pragmatic materialism—the physical aspects are accorded primacy. Third, the rise of "electronic materialism"; recent years have witnessed an astonishing expansion in the range and sophistication of the mental tasks that digital machines can perform. Not only remembering, recalling, and calculating, but pattern recognition, estimation processes, problem solving, and learning new skills, which hitherto have been the exclusive preserve of living, conscious beings, are now routinely performed by electronic devices that, unless panpsychism is true, are purely physical structures. This has formed the background for an increasingly common assumption that mental activity is a special kind of physical process, which is one critical aspect of materialism.

Materialism remains, nonetheless, a striking and apparently paradoxical doctrine, for it insists that the only differences between human beings and grains of sand prove to be matters of energy flow and structural complexity. People have continued to embrace materialism in the face of the difficulties with which it is beset because it offers a comprehensive, unified account of the nature of reality that is economical, intelligible, and consistent with the most successful of the sciences.

HISTORY OF MATERIALISM

CLASSICAL PERIOD. Materialism has been a theme in European speculative thought from the earliest periods for which there is any record.

Ionian philosophers in the tradition of Thales (sixth century BCE) attempted to account for the origin and present state of the world by appeal to changes in the state of a fundamental underlying substance (the *arche*), which in most cases was held to be of a physical nature. Parmenides of Elea (fifth century BCE) vigorously defended a thoroughgoing monism, maintaining that the world is One, unchanging, eternal, homogeneous, indivisible, indestructible, and without any interior void.

These two threads of thought are combined in the true materialism of Leucippus and his pupil Democritus, who flourished at Abdera in the fifth century BCE. Between them they worked out the first clear conception of matter, the first clear restrictions on the kinds of natural interactions in which material particles could participate, and the first clear program of explanation by appeal to these material interactions alone. The "Great Diakosmos," a lost work written by one or the other (or both), expounded their position. Their basic idea was that the fundamental stuff was of just one kind (matter) and that the fundamental entities were material atoms that were of course by no means unique, but otherwise had all the characteristics of Parmenides' One. These atoms are in constant motion in a void that surrounds them.

Insofar as it can be reconstructed, their doctrine embraced the following theses:

(1) Nothing exists but atoms and empty space.

(2) Nothing happens by chance (for no reason at all); everything occurs for a reason and of necessity. This necessity is natural and mechanical; it excludes teleological necessitation.

(3) Nothing can arise out of nothing; nothing that is can be destroyed. All novelties are merely new combinations or separations of atoms.

(4) The atoms are infinite in number and endlessly varied in form, but uniform in composition, being made of the same stuff. They act on one another by pressure or collision only.

(5) The great variety of things that we encounter in the world is a consequence of the variety in number, size, shape, and arrangement of the atoms that compose them.

(6) The atoms have been in confused random motion from all eternity. This is their natural state and requires no explanation. (Some scholars dispute the attribution of random motion to the atoms and credit the "Great Diakosmos" with advancing the doctrine of an eternal fall through infinite space, which was later presented by Epicurus.)

(7) The basic mechanism whereby complex bodies are formed is the collision of two atoms, setting up a vortex. In the vortex motion is communicated from the periphery toward the center. In consequence, heavy atoms move to the center, and there form a body, which is dense relative to the collection of light atoms around the periphery. The vortex continually embraces any new atoms that come near it in their random motion, and it thus begins a world.

This materialist philosophy requires a mechanical account of human sensation. The Leucippus-Democritus account seems to have been ingenious, speculative, and false. Sensation occurs in the human soul, which, like everything else, is composed of atoms. Objects perceptible by the distal senses sight, hearing, or smell, give off effluences, or images, composed of fine, smooth atoms. There are channels in the eyes, ears, and nose along which these effluent atoms pass to collide with the atoms of the soul and produce sensation. Differences of color, in the case of vision, or of pitch, in the case of sound, are due to the varying smoothness or roughness of the incoming image atoms. With the contact senses touch and taste, it is the size and shape of the atoms on the surface of the perceived object that act on soul atoms in the skin or tongue.

Sensory qualities (for example, sweetness, bitterness, temperature, and color) are thus not qualities of the object perceived, which is a collection of atoms, possessed only of physical properties such as size, shape, mass, and hardness. The sensory qualities are, rather, the effects of that collection of atoms on us, that is, on our soul atoms. Here is an early appearance of the distinction between primary and secondary qualities, a distinction every subsequent materialist has also found it necessary to make.

Empedocles (fifth century BCE) founded a medical school in Acragas (Agrigento) in Sicily. His aim was to account in a naturalistic manner for the special features of this world, particularly for the specially organized matter to be found in living creatures. The first appearance of the famous four elements—earth, air, fire, and water—is in his theory. Empedocles seems to have believed that each of these four elements comprised a different type of atom. The creation and dissolution of the macroscopic objects of this world is brought about by the combination

and separation of these atoms by two fundamental forces, love and hate, or harmony and discord.

Under the influence of love and hate the world goes through an endless cycle from complete random separation of elements (the triumph of hate), through gradually increasing order, to a complete, calm, spherical, harmonious union (the triumph of love). Hate then begins to exert itself once more. Disintegration sets in, and ultimately the world returns to the state of complete separation of elements. The present state of the world lies between these two extremes. The existence of planetary systems and the origin of animals are thus explained as the influence of love.

Empedocles can be considered a true materialist only if love and hate are either inherent forces in the elemental atoms or themselves material elements with a cementing or corrosive effect on combinations of the other elements; however, he probably thought of them as blind, powerful gods. The rest of his system is similarly ambiguous. On the one hand, he believed in the transmigration of souls and adhered to some kind of Orphic mystery religion; on the other, he gave a mechanical account of sensation, held that the soul was composed of fiery atoms, and said that the blood around the heart is the thought of men. Empedocles' philosophy thus perpetuated the materialist tradition but not in a rigorous or consistent form.

The hostile misinterpretation of his ethics as unworthily hedonistic has made Epicurus (342–270 BCE) the most famous of classical materialists. In his middle age Epicurus came to Athens and founded a school where materialism was taught as the sole foundation of a good life, at once disciplined, calm, serene, and free from superstition.

He adopted the materialist metaphysics of the "Great Diakosmos" but gave a modified account of the origin of worlds. There are an infinite number of atoms falling vertically through an infinite space. In one construction of the Epicurean system the heavier, faster atoms occasionally strike the lighter, slower ones obliquely, giving them a slight lateral velocity. In another construction all atoms fall at uniform velocity, and the original deviations from parallel downward motion are left unexplained.

However caused, the original lateral deviations result in more collisions and deviations and the establishment of vortexes. From these vortexes ordered arrangements of atoms arise. The number of atoms and the time available are unlimited, so every possible arrangement of atoms must occur at some time or another. This world, with its marvelously organized living bodies, is thus just one of the infinite, inevitable arrangements into which the indestructible atoms must fall.

The only Roman author of note in the tradition of materialism is Lucretius (born c. 99 BCE), whose long didactic poem *De Rerum Natura* gives imaginative sparkle to the metaphysics of Epicurus. Lucretius adopted the second account of the fall of atoms through the void and appealed to some form of voluntary action to explain the original deviations from vertical descent. He thus introduced a nonmechanical source of motion, inconsistent with the remainder of his system.

Like Epicurus, Lucretius was motivated by a wish to free people from the burden of religious fear. He argued passionately and at length against the existence of any spiritual soul and for the mortality of humankind. These beliefs have been explicit features of materialism ever since.

SEVENTEENTH CENTURY. From the close of the classical period until the Renaissance the church and Aristotle so dominated European speculation that materialist theories virtually lapsed. The revival of materialism is attributable to the work of two seventeenth-century philosophers, Gassendi and Hobbes, who crystallized the naturalistic and skeptical movements of thought that accompanied the rediscovery of antiquity and the rise of natural science. Their most important forerunners were probably Telesio, Campanella, and Cyrano de Bergerac, all of whom attempted to combine materialistic views in physics with a psychology based on sensations.

Pierre Gassendi (1592–1655), who in the last part of his life taught astronomy at the Royal College in Paris, rejected the official Aristotelian philosophy of his time and set about the rehabilitation of Epicureanism. To bring the Epicurean system into closer conformity with Christian doctrine, he claimed that the atoms are not eternal but created. They are finite, not infinite, in number and are organized in our particular world by a providential determination of initial conditions.

Gassendi's materialism extended over physics and psychology, undertaking to account for all inanimate changes and for sensation on a materialist basis. He treated the coming into being of particular things as the accumulation of matter about a seed atom.

But his metaphysics was not, strictly speaking, materialistic, for outside the experienced world Gassendi admitted a creative and providential God and an immaterial and immortal intellect in man distinct from his corporeal soul. There are even some lapses in the physics,

too, for Gassendi spoke of gravitation as some kind of movement for self-preservation and allowed that growth from seed atoms may be controlled by formative principles other than the natural motions of atoms.

Thomas Hobbes (1588–1679) was much more consistent and uncompromising. In 1629 he discovered Euclidean geometry and was captivated by its method. During the years that followed he strove to work out a rational philosophy of nature on the Euclidean model.

Hobbes's aim was to discover by cunning analysis of experience the fundamental principles expressing the true nature of everything. The truth of these principles would be manifest to right reason and could thus serve as axioms from which a comprehensive theory of the nature of the world could be deductively derived.

The resulting system is almost pure materialism. Hobbes hoped to use the new non-Aristotelian physics of the seventeenth century as the basis for a final, complete account of reality. From definitions of space and motion he derived the laws of uniform motion. From these, together with a notion of the interaction of bodies, he hoped to proceed to an account of change, thence to an account of sensible change, thence to a theory of the senses and appetites of people, and finally to his notorious civil philosophy.

No part of the universe is not a body, said Hobbes, and no part of the universe contains no body. Hobbes was a plenist, holding all space to be filled by an intangible material ether if nothing else. This doctrine followed directly from his definition of a body as anything existing independently of human thought and having volume. Thus, Hobbes considered God to be a corporeal spirit difficult to distinguish from his eternal, immutable, omnipresent, embodied space, the pervasive ether.

All change in the universe consists in the motion of bodies, so all change reduces to change of position and velocity. Further, nothing can cause a motion but contact with another moving body. The substance of anything is body, and "incorporeal substance" is therefore a contradiction in terms. Hobbes thereby disposed of angels, the soul, and the God of orthodox theology. He departed from strict materialism, however, in his introduction of "conatus" and "impetus" (which are not physical properties) into his account of the initiation of motion and measurement of acceleration. Conatus also enters into Hobbes's account of human sensation and action. Sensations are motions in a person's body, and changes of sensation are changes of that motion. Sensory qualities are really within the perceiver, but by conatus a "phantasm" is projected from the observer onto the observed.

Hobbes was the first to take seriously the problems that language, thought, and logic pose for materialism. He developed a nominalist theory of language and took the subject matter of thought and inference to be phantasms of sense or abstractions from these phantasms. He held, for example, that to remember is to perceive one has perceived. But Hobbes did not make clear just what contact mechanism is at work in mental operations nor whether the phantasms are genuinely corporeal. Thus, in spite of his best efforts, it is doubtful that he developed a fully consistent materialism.

The influence of Gassendi and Hobbes was diminished by the prestige of their brilliant contemporary, Rene Descartes (1596–1650), who accepted a materialist and mechanical account of the inanimate world and the brute creation but insisted that men had immaterial, immortal spirits whose essential nature lay in conscious thought undetermined by causal processes. According to Descartes, there are in the world two quite different sorts of things, extended (material) substances and thinking (spiritual) substances, which are mysteriously united in the case of humankind. He thus crystallized the tradition of dualism (the doctrine that there are just two fundamentally different kinds of substance), which was until recently materialism's chief rival.

EIGHTEENTH CENTURY. In Epicurus and Lucretius one motive for working out a materialist philosophy was to provide an antidote for the all too prevalent religious terror of their times. With Hobbes, and again in eighteenth-century France, the corresponding motive was opposition to religious oppression. But in addition, rapid growth in physiological knowledge had given rise to the hope that a complete doctrine of man in purely physiological terms was possible and so generated a medical materialism that made the path of the metaphysicians smoother.

Ever since the time of Democritus, materialists had held that the soul consists of fine particles within the body. In the course of the eighteenth century this suggestion was taken up and amplified, and some attempt was made to give it an experiential basis.

An anonymous manuscript, the *Ame materielle*, written between 1692 and 1704, contains many ingenious explanations of mental function along Democritean lines. Pleasure and pain consist, respectively, of the flow of finer or coarser particles through the channels of the brain. The passions are a matter of the temperature of the

heart. Reason consists in the ordering of the soul's fine particles, and the effect of wine in its course through the body is to dislodge some of these fine particles from their proper places. The manuscript is panpsychic in its expression, crediting the atoms with a rudimentary consciousness and will, but it is materialist in substance, for these qualities are not credited with any causal power. The doctrines advanced were purely hypothetical and, as we now know, false. The *Ame materielle* had successors in Dr. Maubec's *Principes physiques de la raison et les passions de l'homme* (1709), which again gave a materialist vision of man a panpsychic dress and opposed Descartes's view of the mind as a thinking substance. During the middle years of the century, Denis Diderot's many unsystematic writings took progressively a more materialistic turn. Diderot's *Le reve de d'Alembert* is a striking hypothetical account of heredity, growth, and the simpler forms of animal behavior in terms of internal motions of living bodies.

The most famous medical materialist is Jean de la Mettrie (1709–1751), a doctor with a philosophical bent whose radical views obliged him to leave a fashionable practice in Paris to live in Holland and Prussia. In *L'homme machine* (1943 [1748]) he presented a view of the human being as a self-moving machine.

After criticizing all views of the soul as a spiritual entity, La Mettrie proceeded to review all the common-sense evidence for the physical nature of mental activity. He cited the effects of bodily needs, aging, and sleep; he pointed to the analogy of the human body to much "lower" forms, which were not supposed to harbor spiritual minds. Anticipating Pavlov, he spoke of the mechanical basis of speech and of the possibilities of educating deaf-mutes and anthropoid apes. He explained learning to perceive and learning to make moral judgments by appeal to modifications of the brain. Human action is accounted for by the then new doctrine of the stimulus irritability of muscles. La Mettrie embarrassed those who held that the soul is a spiritual unity governing all vital functions by observing the continuing function of organs removed from bodies, the muscular activity of dead or decapitated animals, and the ability of a bisected polyp to grow into two complete ones. He explained conscious sensation and the mental capacities of which we are introspectively aware by means of a magic-lantern analogy, but this was unsatisfactory, for the status of the images involved was not made clear.

The details of La Mettrie's physiology, depending as they do on supposed movements of nervous filaments, are false. However, his program of seeking in neural changes the explanation of mental activity has endured, and his claim that appeals to the actions of a spiritual soul can furnish only pseudo-explanations has gained wide support.

Jean Cabanis (1757–1808), a French doctor, continued this line of thought and in 1802 published *Rapports du physique et du moral de l'homme*, the most notable innovation of which was to treat the brain as analogous with the digestive system, making sensory impressions its aliments and thoughts its product. The great metaphysical materialist of the period is Paul Heinrich Dietrich d'Holbach (1723–1789), a German nobleman living in Paris. His work the *Systeme de la nature* was published under a false name "Mirabaud," with a false imprint "London" (Amsterdam) in 1770. This "Bible of all materialism" is speculative philosophy in the grand style; in it the antireligious motive is again uppermost. Holbach maintained that nothing is outside nature. Nature is an uninterrupted and causally determined succession of arrangements of matter in motion. Matter has always existed and always been in motion, and different worlds are formed from different distributions of matter and motion. Matter is of four basic types (earth, air, fire, and water), and changes in their proportions are responsible for all changes other than the spatiotemporal ones that motion without redistribution can accomplish.

Mechanical causes of the impact type, such as collision or compression, are the only intelligible ones, hence the only real ones. Because human beings are in nature and part of nature, all human actions spring from natural causes. The intellectual faculties, thoughts, passions, and will can all be identified with motions hidden within the body. In action outward motions of the limbs are acquired from these internal movements in ways we do not yet understand.

Holbach based the intellectual faculties on feeling and treated feelings as a consequence of certain arrangements of matter. Introspected changes are all changes in our internal material state. Thus, in remembering, we renew in ourselves a previous modification. He treated personal characteristics and temperament in terms of a person's internal structure and interpreted so-called free action not as motiveless action (an absurdity) but as action that, although seeming to flow from a free choice, actually springs from an ultimately unchosen modification of the brain. Holbach's theory of mind is also interesting because in dealing with wit and genius, it suggests the first behavioral analyses of mental concepts. As consistency required, he held the soul to be mortal. The purity of Holbach's materialism is marred only by his

admission of relations of sympathy, antipathy, and affinity among material particles, in addition to their unequivocally physical properties, the primary qualities, gravity, and inert force.

The revolution in chemistry that was effected by Joseph Priestley (1777) in England and Antoine-Laurent Lavoisier in France in the 1770s and 1780s was of importance for the later development of materialism, for it established chemistry as a strictly physical science. Since the beginning of the nineteenth century, all properly chemical explanations appeal only to material substances and their natural interactions. Such a chemistry has since been extended in biochemistry to cover all the processes of life, and the case for materialism has thereby been profoundly strengthened. Priestley himself nevertheless vigorously upheld an unorthodox version of Christianity, insisting that the existence of God and the resurrection of the body are not incompatible with a materialist and determinist view of the natural world.

NINETEENTH CENTURY. The philosophers of greatest influence in the nineteenth century—Kant, Fichte, Hegel, Schopenhauer, Lotze, and Mill, for example—were all of an idealist or phenomenalist bent. The dialectical materialism of Friedrich Engels and Karl Marx is not an extreme materialism of the kind discussed here.

Ludwig Buchner, a minor figure, deserves mention as the first to claim explicitly that materialism is a generalization from *a posteriori* discoveries. In *Kraft und Stoff* (1855) he claims that we have discovered (not proven *a priori*) that there is no force without matter and no matter without force.

There was during this period a continuation of inquiry and speculation on the physiological bases of mental function. Jacob Molescott (1852), Karl Vogt (1846, 1854), and Emil Du Bois-Raymond proceeded with the investigation of physiological processes along biophysical and biochemical lines. The most important developments were scientific findings that undermined the barrier between physical systems and living organisms and thus softened the natural resistances to materialistic theses.

In 1828 the synthesis of urea was achieved, and this refuted the idea that biochemistry was in some way special and distinct from chemistry. In 1847, Hermann Helmholtz established the conservation of energy in organic systems, making still less plausible any claims that living and nonliving systems could not possibly be comprehended in a single theory.

In 1859 Charles Darwin published his *Origin of Species*, in 1871 his *Descent of Man*. T. H. Huxley had produced *Man's Place in Nature* in 1863. These three works at last provided a plausible, empirically grounded case for two of the main planks of materialism, the claim that the organization of living things into forms admirably adapted for survival and reproduction can be explained without appeal to immanent or transcendent purposes, and the claim that humans are a part and product of the natural world. Since then biologists, physiologists, and pathologists have increasingly taken the truth of medical materialism for granted, couching their explanations in physicochemical terms without questioning the propriety or completeness of successful explanations in this form.

TWENTIETH CENTURY. The triumphant progress in the twentieth century of a materialistic biology and biochemistry has almost completely eliminated vitalist notions of living forms as governed by forces additional to, and distinct from, the purely physical forces operating on inanimate matter. The situation of earlier ages has been reversed; it now seems implausible to maintain that the vital functions of living organisms are different in kind from chemical (ultimately, physical) processes. In the attempt to demonstrate that something other than matter exists, it is on mind, rather than life, that the opponents of materialism now rely.

Early in the twentieth century, the behaviorist movement arose, in a development linked to the emergence of psychology as a distinct science in its own right, rather than a branch of the philosophy of mind. Many psychologists became disheartened by the difficulties involved in any introspective investigation of inner mental states, and turned to the study of behavior. In its analyses and explanations of human activities, behaviorist psychology relies as far as possible on publicly observable, physical phenomena of stimulus and response. Its aim was to expel the traditionally conceived inner, immaterial mind from psychology, and in this way was a profoundly materialistic development.

In the realm of the mind, a new challenge for immaterialists has also developed. The rise of cybernetics (the abstract theory of machines) and its applications in computing machinery threatens the idea of a special status for mental activity. The gathering and interpretation of information, the employment of stored information, successful and spectacular problem solving, even analogues of fatigue, overload, and confusion, hitherto all found only among complex living organisms, are now displayed by computing hardware, that is, by material structures all

of whose operations can be explained in terms of physical properties alone.

Approaching the issue from the opposite direction, experimental study of the nervous systems of animals and of ourselves is showing, in ever-increasing detail, how artificially induced physical changes in the electrochemical state of the nervous system issue in changes in the subject's mental activity. Displays of emotion, performance in perception and recall, and anxiety and tension are being tied down to brain function in this way.

During the twentieth century, there were in fact three distinct movements of a materialistic stamp in the philosophy of mind. In the 1920s and 1930s some logical positivists, led by Rudolph Carnap (1932–1933) and Otto Neurath, espoused an epistemic materialism. They held that the meaning of any statement consists in the directly testable statements deducible from it (the protocol sentences). In order for language and meaning to be public and shared, these protocol sentences must be intersubjectively testable. However, because no statement about one individual's experience or thought or other inner psychological state can be tested by anyone else, only sentences referring to the physical properties of physical entities are intersubjectively testable in the required way. Now, because most statements about minds are incontestably meaningful, they must, despite appearances to the contrary, in fact refer to physical properties and entities, even though translations of them into physical terms cannot be provided. In this way the philosophy of language led to a behaviorist materialism.

The beginnings of translation into behavioral terms was offered for some psychological expressions—for example, "is happy"—by directing attention to the way in which the use of such expressions is taught. A key element in teaching such an expression is to point to people behaving happily. In this emphasis on the conditions under which an expression can be learned, the positivists anticipated the favorite strategy of Ludwig Wittgenstein (1953) and moved away from complete dependence on their general doctrines of meaning and verification.

During the middle years of the twentieth century, the analytic behaviorists, in particular Gilbert Ryle (1949) and his followers, offered to show that descriptions of states of mind are essentially dispositional, so that attributions of intention and intelligence, choice and desire, excitement and fear, and other mental states are all to be understood as attributions of a disposition to behave in a characteristic manner in appropriate circumstances. Dispositions are held by most thinkers to issue from some standing or recurrent underlying state, and with these analytic behaviorists the relevant states underlying human mental life were assumed to be states of the body. Their manifest intention to exorcise any spiritual soul—as Ryle would put it, any "ghost in the machine"—places them in the materialist tradition.

Wittgenstein, although he disdained the title behaviorist, belongs to the same group. He insisted that in any acceptable analysis of a mental concept the description of a person's state of mind must make reference only to publicly detectable features of the organism and its behavior. His many subtle discussions of mental concepts are all attempts to identify the patterns of behavior whose display would constitute being in a given state of mind. To attribute that state of mind to someone is to attribute a disposition to display the relevant pattern of behavior. The alternative analysis that interprets the various states of mind as states and processes in a spiritual soul is, according to Wittgenstein, not merely false, it is unintelligible.

On two key points the analytic behaviorists were not convincing. First, if mental states are dispositions to display particular patterns of behavior, they cannot be causes of the behavior in question. It cannot be that a man's anger made him shout, for the shouting is itself just an aspect of the anger. Nor can a woman's pride have made her stubborn. Yet this causal link between a mental state and the characteristic behavior pattern that springs from it, is at the heart of how we understand one another.

Second, some inner mental episodes, such as afterimages, pains, sudden unsought recollections, dreams, or flashes of insight, resist any plausible dispositional analysis. The mind does seem to be a collection of categorical states, items, or events in addition to a cluster of dispositions. The effort to correct both these weaknesses, first the denial of any categorical component, and later the denial of any causal power to the mind, was a significant factor in materialism's subsequent development.

The third group of twentieth-century materialists embraced a theory of mind known as central-state physicalism, from which contemporary materialism derives. The central-state physicalists held that although it may be that some mental states can be understood dispositionally, there are many mental states, items, or events that must be accorded a straightforwardly categorical status. These categorical mental states turn out to be, as a matter of contingent fact, states of the central nervous system. To introspective awareness they do not seem to be neural states, but the explanation for this is that the nervous system is presented to itself in an opaque or covert fashion.

The mind has many aspects, and mental life underpins almost every distinctively human capacity. Most of our distinctive capacities have been pointed to as showing that a living human being must be something more than a mere assemblage of atoms. To understand ourselves, we cannot do without the concepts of perception, belief, and intelligence; action, decision, and choice; motive, drive, and need; feeling, emotion, and mood; temperament and character. We will also need to treat of consciousness and self-consciousness. The task for materialists is to explain how merely material structures could exhibit all these mental attributes. In attempting this, two basic approaches were at first adopted, the behavioral and the topic neutral.

Behavioral strategy. The central-state physicalists were able to appropriate the earlier work of the behaviorists and accept that the attribution to an organism of some of the mental predicates (for example intelligence, equanimity, or ambition) is in reality the attribution of a disposition to behave in a characteristic way under suitable conditions. The organism displaying the behavior, the form the behavior takes, and the conditions under which it is manifested, are all specifiable in purely physical terms. Moreover, the remarkable subtlety and complexity of human behavior, which until the twentieth century appeared to surpass anything of which a mere machine could be capable, no longer has such immaterialist implications, for now the development of electronic machines suggests that the ability to duplicate human performance is possible. In particular, the self-monitoring features of conscious behavior can be displayed by material systems.

Topic-neutral strategy. Many mental states resist the behavioral strategy: being in pain, seeing a color, or feeling depressed, for example. For these, a different claim was made: To attribute such a state is to assert that there is present within the organism some state or process that typically arises from a particular kind of stimulus and/or typically issues in a characteristic kind of behavior. A burning pain, for example, is a state of a person typically arising from excessive heat on the skin, and characteristically issuing in applications of soothing cream to the affected part. Mental predicates of this kind have been called *topic-neutral* because they do not specify the nature of the inner state in question. The inner state is not described either as material or as immaterial. To say that someone is in pain, the argument runs, does not of itself imply that the experience belongs to a immaterial mind. It implies only that the person is in some central state or other, arising from the states and processes in the sensory system (input), and issuing in certain behavior patterns (output). When we attempt to identify this central state, we find that the sensory system provides inputs to the organism's central nervous system, which in turn sets in train the muscular movements required for any type of behavior. If inner states admit of the topic-neutral treatment, they, too have no immaterialist implications.

Among early central-state physicalists, some, such as Paul K. Feyerabend (1963) and Hilary Putnam, claimed only that this is the most promising line for investigation to now take. Others, such as U. T. Place (1956), J. J. C. Smart (1959, 1963), and Herbert Feigl (1958), went further and held that any alternative dualist view is already frankly incredible.

CONTEMPORARY MATERIALISM

During the later years of the twentieth century, under the influence chiefly of David Armstrong (1968) and David Lewis (1972), the topic-neutral strategy was taken up and developed. The behavioral strategy became less prominent, as more and more mental attributions were interpreted as asserting that the organism was in an appropriate categorical state. And the role of the mental as the causal bridge between stimulus and response was taken up and emphasized. Mental states came to be regarded as theoretical constructs and assimilated to other theoretical entities more familiar from other sciences, as philosophers adopted a third strategy for accounting for mental descriptions in a material world.

CAUSAL/THEORETICAL STRATEGY. In a complete departure from the behaviorist viewpoint, which saw mentality as a matter of the outer effects of stimuli, the new position is that the really essential thing about any mental state is its causal role, as the crucial inner intermediary between input and response. The idea is that the activity of conscious living beings calls for explanation, and the most appropriate explanations will attribute to such organisms inner states, produced by environmental and remembered elements, and producing behavior that, in the light of the organism's beliefs, is best suited to fulfilling its purposes.

So the mind becomes an inner, theoretical entity, the that-which-best-accounts-for the phenomena of conscious behavior. The analogy was drawn with the gene in biology, that-which-best-accounts-for the phenomena of heredity, and with lightning, that-which-best-accounts-for flashes, thunder, and some kinds of storm damage.

Then, still following the analogy, the research question becomes that of finding which element in the world

turns out to fill the theoretical role in question. Structures crucially involving the DNA molecule, as it turns out, best account for heredity. Electrical discharges, as it turns out, best account for the flashes, rumbles, and damage of electrical storms. This is a matter of the contingent identification of underlying structures and processes as the causal bases for patterns of observed phenomena. So with the mind: It is the central nervous system (brain, optic nerve, spinal chord, and some other components) that, as it turns out, fulfills the mind's causal role as the intermediary and clearinghouse between the inputs, many of which we know as experience, and the outputs that consist in purposive activity.

In this way functionalism, the dominant form of contemporary materialism, developed. It has two components. The first component is a theory of the mind, which asserts that the essential feature of the mind is its causal role, and identifies the different states of mind—beliefs, fears, plans, twinges, and so forth—in terms of their particular places in the whole mental causal scheme. This theory of the nature of mind lends itself to materialism, but is not itself materialist. It is topic-neutral, allowing for any of a number of views of what it is that provides the causal bridge between inputs and responses. The second component in functionalist materialism is the theoretical identification of the mind with the central nervous system. This is a contingent assertion about what minds turn out to be in this world. As such, it is vulnerable to various empirical developments, as all substantial empirical claims should be.

OBJECTIONS TO MATERIALISM

THE POSSIBILITY OF SCIENTIFIC REFUTATION. Materialism is a strong version of naturalism. It asserts that everything whatsoever that occurs in this world is the result of the operation of physical forces in accord with physical laws. So a spectacular and unequivocal divine intervention in the course of nature, such as the Apocalypse and the Day of Judgment as described in the book of Revelation, would spell the end of materialism as a credible philosophy.

Less spectacular developments could have the same impact. The firm establishing of parapsychological powers (telepathy, clairvoyance, or psychokinesis) would do so, for by definition any paranormal phenomenon involves knowledge or action by a mind in defiance of physical law. So also would developments in neural science that uncovered variations in effectual states of mind without any appropriate change in states of the central nervous system. Or changes in the central nervous system linked to changes in mental state, such as forming a new resolution, that systematically violate the probabilities for neural change that physical laws set forth and that defy any modification to accepted physical laws.

Materialism, being vulnerable in these ways, remains to that extent speculative. But whereas a watching brief needs to be kept over the progress of scientific investigations, it is fair to say that there is at present no serious threat from these quarters. The credibility of positive paranormal results has, if anything, diminished in the course of the past half century. And we are very far indeed from being able to assert that the activity of the brain is physically anomalous. Quite the contrary; so far, no apparent violations of physical law have been found.

THEOLOGY. Materialism not only holds that there are no supernatural interventions in the course of nature, but that there are no divine beings of any kind. To defend materialism on these points, one must first show that there is no valid deductive argument for the existence of a necessary being, then sustain the view that this world does not call for a divine creator as the best explanation for its existence and character.

Next, one must deny that religious experience reveals a supernatural realm, as vision provides access to a physical one. Adopting the skeptical empiricists' critique, one can argue that religious experience is not sufficiently uniform, widespread, and unanimous to warrant abandoning the natural modes of explanation that have served so well in all other enquiries, especially as supernatural hypotheses face peculiar difficulties when it comes to putting them to the test. The materialist position is strengthened by the promise of continued success in finding concrete natural explanations of religious experience through developments in sociology, psychology and physiology.

If these positions can be established, claims to the existence of God and the occurrence of miracles are established neither by argument nor in experience and so must be considered as interpretative hypotheses laid upon the experienced world. The materialist must again urge that in framing hypotheses, as in seeking explanations, there is no sufficient reason for deserting the natural for the supernatural. In such circumstances as these considerations of parsimony exclude all supernatural entities from any reasonable ontology.

Materialists must show that, contrary to the claims of Spiritualists and Buddhists, there is no sufficient reason to believe in survival of bodily death or in reincarnation. And indeed there are plausible arguments that both

doctrines rest on untenable views of the self. These arguments do not impugn the possibility of bodily resurrection, but that is compatible with materialism.

METAPHYSICS. Materialism has in the past been assailed as incomplete. Even if, in a great advance on its predecessors, modern cosmology does provide explanations for the origin, persistence, and motion of the fundamental particles, it provides none for the initial conditions from which these derive. Nor does materialism make intelligible why each fundamental interaction has had one result and not another. The reply, now widely accepted, is that all chains of explanation must eventually come to a terminus and that to seek to go beyond contingent truths concerning the items and processes in this world is to go hunting a mare's nest.

THE MIND AND HUMAN EXPERIENCE. There is no doubt that our own conscious experience provides the greatest intuitive challenge to materialism. C. D. Broad in *The Mind and Its Place in Nature* (1925) formulates many people's reaction to the suggestion that mental events are physical events, such as molecular movements, taking place in our body:

> About a molecular movement it is perfectly reasonable to raise the question "Is it swift or slow, straight or circular and so on?" About the awareness of a red patch it is nonsensical to ask whether it is a swift or slow awareness, a straight or circular awareness, and so on. Conversely, it is reasonable to ask whether it is a clear or a confused awareness, but it is nonsense to ask of a molecular movement whether it is a clear or a confused movement. Thus the attempt to argue that "being a sensation of so and so" and "being a bit of bodily behavior of such and such a kind" are just two names for the same characteristic is evidently hopeless. (p. 623)

Indeed, this attempt is hopeless, but it is not one a materialist must make. We need to distinguish the process of being aware from the item of which we are aware. The two "names" that materialists claim to name the same thing are "subject *S* having sensation *P*" and "subject *S* undergoing bodily changes *Q*," and it has become clear since Broad wrote that what is or is not nonsensical is not an immediate deliverance of introspection, but an issue in the fashioning of concepts to improve theories of the world. As for *P*, which is the item of which S is aware—what Broad calls the sensation *S* has—there would be no absurdity if this could be dealt with by a topic-neutral strategy. We are aware that something is going on in us, which deserves the description "red patch," but according to the topic-neutral strategy, the nature of what is going on is not part of what we are conscious of. The fact of the matter, according to the materialists, is that we have a covert presentation of bodily changes *Q* to the person *S*, who is having the sensation. Nevertheless, the two main stumbling blocks for functionalist materialism both concern the character of our inner life.

The qualia problem. The topic-neutral or causal/theoretical strategies may well be satisfactory for those inner states that have no special "feel" about them, such as deciding. We can decide to do something, and be aware that we have decided, but that awareness carries no special feel or twinge or glow with it. We are aware that something is going on in us, something that will have an impact on how we behave by bringing a new causal factor into our life. But that state, and our awareness of that state, reveal nothing about its nature as material or immaterial. Decisions and intentions are thus favorable candidates for a topic-neutral analysis—so, too, is doing mental arithmetic, where the process leads to changes in what one will say or do, but carries no other inner characteristics that one is aware of.

The case is otherwise, however, with sensations and feelings. To see a red patch is to be aware of an inner state that has a redness about it, that sets it apart from the green and blue patches we see. This difference is not obviously a difference in how we discriminate the two items, and react to them, as is brought out by the spectrum-shift arguments, which point out that although your outward color-vision behavior may match mine, you may see reds as I see pale pinks, or blues as I see greens.

To be in love is certainly to be in a state apt to issue in a characteristic pattern of behavior. But it is more than that; there is a complex of feelings involved that do not of themselves involve behavioral differences, but differences in consciousness, by comparison with those not in love.

To be angry, or in pain, or delighted, carry special sensations or feelings with them too. All such sensations or feelings are known as *qualia*, and the qualia problem is the problem of fitting them in to a materialist world view. It is notorious that when you are seeing something green, and therefore experiencing a sensation of green, there is no green physical surface anywhere inside you. The sharp pangs of pain are similarly elusive—the neural activities have been found that occur when pain is felt, but the painfulness of pain does not seem to be present among them.

Qualia seem to be an important part of being conscious. They seem to make a difference to how we speak and act, yet they stand outside the network of physical causation, neither taking energy in their production, nor having any force to apply to change the world. They challenge the deep materialist commitment to the physical closure of the natural world. If only physical items can have physical effects, then qualia cannot even produce our awareness of them, nor our capacity to describe them, which makes them paradoxical items indeed.

There have been attempts to account for them behavioristically, as dispositions to act and react in particular ways. Perhaps the most promising materialist suggestion is that the intrinsic qualities of sensations are in reality purely schematic and enable us only to distinguish one sensation from another. Inner states notoriously elude direct characterization. Our attempts to describe them often proceed by comparison with other sensations directly or ultimately picked out by reference to their stimulus and/or response. For example, we describe smells as of cinnamon or of rotten eggs (stimulus) and as appetizing or nauseating (response); we speak of pains as jabbing, burning, or like "pins and needles." Feelings of anger, shame, pride, and fear are all described in terms of bodily temperature.

If the sameness or difference of inner states but not their nature is given introspectively, sensations could well be states of the nervous system typically connected with stimulus and/or response, even though we are not aware of this. This strategy for dealing with qualia faces the problems of spectrum-shift arguments, because two sets of sensations, tastes for example, could be shifted relative to one another along a spectrum, yet perform equally well in informing us of the sameness or difference, and typical causes and effects, of our inner states.

The qualia problem was long emphasized by F. C. Jackson (1998) in a series of influential articles. His most recent stance is the "there must be a solution" solution: Somehow, qualia must be reconcilable with materialism, even if we cannot see how.

The insight problem. The second currently most acute problem for materialism concerns the nature of human insight and understanding. When we learn to speak a language, we acquire the ability to conduct a conversation satisfactorily; that is, to make appropriate responses to the speech of others, to initiate conversations using sounds the other recognizes and responds to. But to properly understand, more than linguistic competence is required. This was dramatized by John Searle (1992) in his "Chinese Room" argument: If someone who had no understanding of Chinese but who could recognize Chinese characters were shut away in a room, and provided with pieces of Chinese—questions and so forth—through a mailbox, that person could, using a computerized dictionary for example, choose appropriate Chinese-character responses. This is a linguistic competence that does not include understanding and is clearly deficient by comparison with the capacity of a genuine Chinese speaker. The missing component, understanding or insight, proves just as elusive as do the qualia to materialist studies of the nervous system.

PHILOSOPHY. Materialism faces several other more general objections, for the most part of a logical kind, that must be faced.

The argument from self-destruction. A popular argument for disposing of materialism is this:

> All doctrines concerning the nature of the world are arrived at by inference.
>
> Thus, *a fortiori*, materialism is so reached.
>
> But if materialism is true, inference is a causally determined process in people's brains, and not a rational process.
>
> Materialism is therefore a doctrine arrived at by non-rational causal processes.
>
> Thus, if it is true, there can be no reason to think it so.

This argument has a long history, being found in Epicurus and developed and defended by J. B. S. Haldane (1932) and Karl Popper (1977). Nevertheless, it is invalid. That the course of a given process of inferring was determined by the structure of a brain does not entail that it was an unreasonable inference. Nor does it entail that there could be no ground for thinking it reasonable. We can see that this is so, by comparing reasoning in people with calculating in adding machines. The result reached is a causal consequence of the structure of the machine; it is nonetheless a correct one, and one we are entitled to rely on. Haldane later retracted his argument (1954).

Asymmetrical knowledge of physical and mental states. Another common argument against materialism points out that,although ordinary people can recognize thoughts and feelings and intentions, they are completely ignorant of processes in the central nervous system, and so the mental occurrences cannot be identified with any such physical events. Friedrich Paulsen, for example, argued to this effect in chapter one of his *Introduction to Philosophy* (1895 [1892]).

This argument is also, as it stands, invalid. It is like arguing that because the police know some of the characteristics of a man who committed a crime but do not know anything about John Smith, John Smith could not possibly be the man who committed the crime. A similar reply is provided by Place and Smart in articles cited in the bibliography.

The argument would be valid if another premise were added: In introspection the full nature of mental events is disclosed. But there is no good reason for thinking this premise is true.

A variation of this argument claimed that introspective knowledge of our own mental states is incorrigible, whereas no knowledge of anything physical is incorrigible, so mental states cannot be physical. This argument faded from view after Armstrong exposed its weakness: We can and do make mistakes about our own inner mental states.

The general nature of human reason. Keith Gunderson (1964) revived an argument of Descartes's to the effect that men are not machines, even cybernetic machines, and therefore not merely material. In all known machines the matching or surpassing of a human intellectual ability is a specific outcome of a specific structure. Each skill is a skill at some specific task and no other. But in human beings, intellectual skills are generalized and come in clusters; human reason is a tool for all circumstances. Thus, it is not proven that the human skill and that of the machine arise from a like inner structure. On the contrary, the reasonable conclusion is that the machine's skill and the human skill are to be explained in different ways—that is, a person is not any kind of machine.

The reply available to materialists is that this argument is premature. The simulation of human performance by material assemblages is in its infancy. There seems no reason to suppose a machine with generalized skills impossible.

Intentionality. Unlike the situation with anything physical, in the realm of the mind there are relations that can exist even in the absence of one of their terms. These are the intentional relations, which include intending, believing, hoping, fearing, and desiring. The argument from intentionality rests on this peculiarity and may be put this way:

> A peculiarity of many mental states is their essential connection with an object. In intending, I must intend something, and in hoping, I must hope for something.
>
> However, whereas when I kick something, the thing I kick must exist, the thing intended or the thing hoped for may or may not have any real existence.
>
> In this way some mental states differ essentially from all physical states.
>
> Thus, materialism cannot be true.

The materialist reply to this argument is that intentional "relations" are strictly speaking not relations but monadic states that are identified by reference to what would fulfill them or constitute their exercise. These are possible states or circumstances that, were they actual, would be material. It is a further question, however, whether the existence of mere, unactualized possibilities is compatible with a strict materialism.

Logical connections between distinct existences. The essential link between a mental state and the behavior to which it gives rise has also been seen to rule out materialism:

> Where an intention is carried out, both the intention and the thing intended exist.
>
> They are two different things.
>
> Nevertheless, they are logically connected, because what was carried out makes the intention what it was.
>
> But any two different physical items are only contingently connected.
>
> Hence, mental states cannot be physical items.
>
> Materialists urge in rebuttal that this is a consequence of the peculiarly causal character of mental states and has its counterpart in the uncontroversially physical realm.

Thus if we describe arsenic causally as a lethal poison, there is a logical connection between drinking the lethal poison arsenic and dying, even though the arsenic, the drinking, and the dying, are all distinct existences.

NONREDUCTIVE MATERIALISM

Despite the progress made in rebutting the classical objections to materialism, and despite the current popularity, in English-speaking philosophy, of functionalist physicalism as a philosophy of mind, uneasiness remains that materialism accords insufficient recognition to consciousness and its highest expressions—music, literature, love, and fine feeling generally, as well as culture, morality, and religious aspiration. In response to this, there have been some attempts at a softer materialism that tries

to accord to the physical a primary but not exclusive place. While everything depends on the physical, it does not reduce to the physico-chemical, but rather supervenes upon it. The most thorough attempt in this direction is J. F. Post's *The Faces of Existence* (1987). A further step away from extreme materialism is taken in Nicholas Maxwell's *The Human World in the Physical Universe* (2001), which advocates a dual-aspect position while clinging to the central materialist claim that the universe is a closed system, in which the only causally effective forces are the physical ones.

See also Philosophy of Mind.

Bibliography

GENERAL HISTORIES OF MATERIALISM

Lange, Frederick Albert. *Geschichte des Materialismus* (1865). Translated by E. C. Thomas as *The History of Materialism*. London: Truebner, 1877–1892. This classic is by far the most important secondary source in the history of materialist theories. All English editions since the London edition of 1925 include an introduction by Bertrand Russell titled "Materialism."

Lewes, George Henry. *The Biographical History of Philosophy*. London: C. Knight, 1845. A lively and idiosyncratic history, with good comments on materialist philosophers.

Vitzthum, R. *Materialism: An Affirmative History and Definition*. Amherst, NY: Prometheus, 1995. Accessible work, emphasizing the scientific aspects of the development of materialism.

CLASSICAL PERIOD

Aristotle. *Metaphysica*. Translated by W. D. Ross as *Metaphysics* Oxford: Clarendon Press, 1924. Most easily accessible source for ancient world opinion on materialistic Greeks.

Burnet, John. *Early Greek Philosophy*. 4th ed. London: A & C Black, 1930.

Burnet, John. *Greek Philosophy; Thales to Plato*. London: Macmillan, 1914. Standard account of ancient philosophy.

Lucretius. *De Rerum Natura*. Classic presentation, in Latin verse, expanding and occasionally modifying the doctrines of Epicurus. Translated by Rolfe Humphries as *The Way Things Are*. Bloomington: Indiana University Press, 1969. Another edition, translated by R. E. Latham. Harmondsworth, England: Penguin, 1961.

SEVENTEENTH CENTURY

Brandt, Frithiof. *Den mekaniske Naturopfattelse hos Thomas Hobbes*. Translated by Vaughan Maxwell and Annie Fairstoll as *Thomas Hobbes' Mechanical Conception of Nature*. London: Librairie Hachette, 1928. Detailed and definitive.

Cyrano de Bergerac, Savinien de. *Etats et empires de la lune* and *Etats et empires du soleil*. 1657, 1662. Translated by Richard Aldington as *Voyages to the Moon and the Sun*. New York: Orion Press, 1962. These fictional works epitomize the seventeenth-century movement toward materialism.

Gassendi, Pierre. *Animadversiones in Decimum Librum Diogenis Laerti*. Lyons, France, 1649.

Gassendi, Pierre. *De Vita et Moribus Epicuri Libri Octo*. Lyons, France, 1647.

Gassendi, Pierre. *Philosophiae Epicuri Syntagma*. Lyons, France, 1649.

Gassendi, Pierre. *Syntagma Philosophicum*. Lyons, France, 1658. These four works appeared together in *Opera Omnia*. Lyons, France, 1658. Facsimile reprint Stuttgard-Bad-Cannstatt, frommann, 1964. In these four works Gassendi expounds and defends Epicurus (except where his views conflict with Catholic doctrine.)

Hobbes, Thomas. *English Works*, edited by William Molesworth. London: Bohn, 1839.

Hobbes, Thomas. *Latin Works*, edited by William Molesworth. London: Bohn, 1839.

Mintz, Samuel I. *The Hunting of Leviathan*. Cambridge, U.K.: Cambridge University Press, 1962. Narrates contemporary reactions to Hobbes.

Peters, Richard. *Hobbes*. Harmondsworth: Penguin, 1956. Interesting and comprehensible introduction.

Rochot, Bernard. *Les travaux de Gassendi sur Epicure et sur l'atomisme 1619–1658*. Paris: Libraire Philosophique, 1944. Most accessible modern treatment of Gassendi.

Spink, John S. *French Free-Thought from Gassendi to Voltaire*. London: University of London, Athlone Press, 1960.

EIGHTEENTH CENTURY

Cabanis, Pierre-Jean-Georges. *Rapports du physique et du moral de l'homme*. Paris, 1802. English translation by M. D. Saidi. Baltimore: Johns Hopkins Press, 1981.

Diderot, Denis. *Oeuvres completes*. Paris: Garnier Freres, 1875.

Holbach, Paul Heinrich Dietrich d'. *Systeme de la nature … par Mirabaud*. 1770. Translated by H. D. Robinson as *The System of Nature*. Edited by Denis Diderot. Boston: J. P. Mendum, 1853. Classic treatise.

La Mettrie, Julien Offray de. *L'homme machine*. 1748. Translated and edited by Ann Thom as *Machine Man*. Cambridge, U.K.: Cambridge University Press, 1996. There is also a critical edition with notes by Aram Vartanian. Princeton, 1960.

Maubec. *Principes de la raison et les passions de l'homme*. Manuscript held in the Bibliotheque nationale in Paris (B.N. n.a.f.r.14709).

Priestley, Joseph. *Disquisitions Relating to Matter and Spirit*. London: J. Johnson, 1777.

NINETEENTH CENTURY

Buchner, Ludwig. *Kraft und Stoff* (1855). Translated by J. Frederick Collingwood as *Force and Matter*. London, Truebner, 1884.

Darwin, Charles. *The Descent of Man*. New York: Appleton, 1871.

Darwin, Charles. *The Origin of Species*. London: John Murray, 1859. Classic treatise.

Moleschott, Jakob. *Der Kreislauf des Lebens*. Mainz, Germany: Von Zabern, 1857.

Passmore, John. *A Hundred Years of Philosophy*. London: Duckworth, 1957. Readable and authoritative guide.

Vogt, Karl. *Kohlerglaube und Wissenschaft*. Giessen, Germany: Ricker, 1855.

Vogt, Karl. *Physiologische Briefe*. Giessen, Germany, 1854.

TWENTIETH CENTURY

Armstrong, David M. *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul, 1968. A modern classic, first systematic exposition of central-state physicalism.

Carnap, Rudolf. "Psychologie in physikalischer Sprache" (1932–1933). Translated by Frederick Schick as "Psychology in Physical Language," in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959. Epistemic materialism.

Eliot, Hugh. *Modern Science and Materialism*. London: Longmans, Green, 1919.

Feigl, Herbert. "The 'Mental' and the 'Physical.'" In *Minnesota Studies in the Philosophy of Science*. Vol. 2, edited by Herbert Feigl et al. Minneapolis: University of Minnesota Press, 1958. Identifies mental states with brain processes.

Feyerabend, Paul K. "Materialism and the Mind-Body Problem." *Review of Metaphysics* 17 (1) (1963): 49–66. Claims materialism is the only plausible view.

Hook, Sidney, ed. *Dimensions of Mind: A Symposium*. New York: New York University Press, 1960. Deals with the relations of mind, body, and machinery from many points of view.

Lewis, David K. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (1972): 249–258.

Nagel, Ernest. "Are Naturalists Materialists?" In his *Logic without Metaphysics*. Glencoe, IL: Free Press, 1957.

Popper, Karl, and John C. Eccles. *The Self and Its Brain*. New York: Springer International, 1977. Full-scale attempt at a Cartesian dualism. Includes discussion of the argument from self-refutation.

Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47 (1956): 44–50. Pioneering paper in central state physicalism.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949. Readable and influential advocacy of analytic behaviorism.

Skinner, B. F. *Science and Human Behavior*. New York: Macmillan, 1953. Major statement of methodological behaviorism.

Smart, J. J. C. *Philosophy and Scientific Realism*. London: Routledge and Kegan Pual, 1963. Amplification of Smart's materialism.

Smart, J. J. C. "Sensations and Brain Processes." *Philosophical Review* 68 (1959): 141–156. Classic and influential pioneering physicalist paper.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Blackwell, 1953. Influential but difficult reappraisal of the logic of mental concepts.

CONTEMPORARY MATERIALISM

Braddon-Mitchell, David, and Frank Jackson. *The Philosophy of Mind and Cognition*. Oxford: Blackwell, 1996. Best modern introduction, with discussion of the qualia problem.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996. Presents a moderate dualist position.

Churchland, Paul. *A Neurocomputational Perspective*. Cambridge, MA: MIT Press, 1989. Shows how information systems can exhibit mental powers.

Dennett, Daniel. *Consciousness Explained*. Boston: Little, Brown, 1991. Functionalist account of consciousness.

Gillett, Carl, and Barry Loewer, eds. *Physicalism and Its Discontents*. New York: Cambridge University Press, 2001. Sympathetic treatment of difficulties for materialism, with useful bibliography.

Jackson, Frank, ed. *Consciousness*. Aldershot, U.K.: Ashgate, 1998. Collection of articles, including some on qualia.

Kim, Jaegwon. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

Lycan, William G. *Consciousness and Experience*. Cambridge, MA: MIT Press, 1996. Functionalist view.

Lycan, Willaim G., ed. *Mind and Cognition, An Anthology*. 2nd ed. Malden, MA: Blackwell, 1999. Collection of influential articles.

Maxwell, Nicholas. *The Human World in the Physical Universe: Consciousness, Free Will, and Evolution*. Lanham, MD: Rowman & Littlefield, 2001. Presents a dual-aspect reconciliation of physical science and conscious experience.

McGinn, Colin. *The Problem of Consciousness*. Oxford: Blackwell, 1990. Argues the problem is beyond our competence to resolve.

Papineau, David. *Philosophical Naturalism*. Oxford: Blackwell, 1993. Concerns supervenience and physicalism.

Poland, Jeffrey. *Physicalism; The Philosophical Foundations*. Oxford: Clarendon, 1994. General and systematic presentation of the philosophy of materialism.

Post, John F. *The Faces of Existence*. Ithaca, NY: Cornell University Press, 1987. Propounds nonreductive materialism.

Rosenthal, David, ed. *The Nature of Mind*. New York: Oxford University Press, 1991. Collection of influential articles.

Searle, John. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press, 1992. Antireductionist materialism.

Thau, Michael. *Consciousness and Cognition*. Oxford: Oxford University Press, 2002.

PHILOSOPHICAL OBJECTIONS

Armstrong, David M. "Is Introspective Knowledge Incorrigible?" *Philosophical Review*. 72 (1963): 417–432.

Broad, Charlie Dunbar. *The Mind and Its Place in Nature*. London: Harcourt, Brace, 1925. Readable review of various doctrines of the mind.

Ducasse, Curt John. *Nature, Mind, and Death*. La Salle, IL: Open Court, 1951. Dualist attack on materialism.

Foster, John L. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*. London: Routledge, 1991.

Gunderson, Keith. "Descartes, La Mettrie, Language and Machines." *Philosophy* 39 (1964): 193ff. Revives argument from the generalized nature of reason.

Haldane, J. B. S. *The Inequality of Man*. London: Chatto and Windus, 1932. Presents the argument from self-refutation.

Haldane, J. B. S. "I Repent an Error." *The Literary Guide*. (April 1954): 7 and 29. Retracts argument from self-refutation.

Lewis, C. S. *Miracles*. New York: Macmillan, 1947. Expounds argument against materialism as self-refuting.

Paulsen, Friedrich. *Einleitung in die Philosophie*. 1892. Translated by Frank Thilly as *An Introduction to Philosophy*. New York: Holt, 1895. Idealist attack on materialism.

Plantinga, Alvin. "Probability and Defeaters." *Pacific Philosophical Quarterly* 84 (3) (2003): 291–298. Argues that materialism is self-subverting.

Keith Campbell (1967, 2005)

MATERIALISM, DIALECTICAL

See *Dialectical Materialism*

MATERIALISM, HISTORICAL

See *Historical Materialism*

MATHEMATICS, FOUNDATIONS OF

The study of the foundations of mathematics comprises investigations, though probably not all possible investigations, that consist of general reflection on mathematics. The subject naturally proceeds by singling out certain concepts and principles as "fundamental" and concentrating attention on them, but of course the identification of fundamental concepts and principles is itself based on foundational research or may be revised in the light of it.

In this entry considerable emphasis will be placed on philosophical questions about mathematics, which undoubtedly belong to foundations. However, many, perhaps most, foundational investigations are mainly mathematical. In the last hundred years an important role has been played by mathematical logic. We shall not give a detailed exposition of mathematical logic, but we hope that our discussion will give an idea of the relation between the logical problems and results and the philosophical problems and an idea of some of the results of recent work in logic.

Two of the main qualities for which mathematics has always attracted the attention of philosophers are the great degree of systematization and the rigorous development of mathematical theories. The problem of systematization seems to be the initial problem in the foundations of mathematics, both because it has been a powerful force in the history of mathematics itself and because it sets the form of further investigations by picking out the fundamental concepts and principles. Also, the systematic integration of mathematics is an important basis of another philosophically prominent feature, its high degree of clarity and certainty. In mathematics systematization has taken a characteristic and highly developed form—the axiomatic method—which has from time to time been taken as a model for systematization in general. We shall therefore begin our main exposition with a discussion of the axiomatic method.

Foundational research has always been concerned with the problem of justifying mathematical statements and principles, with understanding why certain evident propositions are evident, with providing the justification of accepted principles that seem not quite evident, and with finding and casting off principles which are unjustified. A natural next step in our exposition, then, will be to consider mathematics from an epistemological point of view, which leads us to examine mathematics as a primary instance of what philosophers have called a priori knowledge. In this connection we shall give some logical analysis of two very basic mathematical ideas, class and natural number, and discuss the attempts of Gottlob Frege and Bertrand Russell to exploit the intimate relation between these two ideas in order to prove that mathematics is in some way a part of logic. We shall also discuss Immanuel Kant's views on the evidence of mathematics and other conceptions of a priori knowledge. (The word *evidence* will often be used in this entry in a way that is unusual outside philosophical writings influenced by the German tradition, to mean "the property of being evident"—German, *Evidenz*.)

The growth of modern mathematics, with its abstract character and its dependence on set theory, has caused the problem of evidence to be focused on the more particular problem of platonism. It is in this development and the accompanying growth of mathematical logic that modern foundational research has centered.

Throughout the nineteenth century, mathematicians worked to make arithmetic and analysis more rigorous, which required axiomatization and an attempt to use the concepts of the theory of natural numbers as a basis for defining the further concepts of arithmetic and analysis. The manner in which this axiomatization and definition was undertaken was platonist, in the sense that both numbers and sets or sequences of numbers were treated as existing in themselves. The development of set theory by Georg Cantor provided a general framework for this work and also involved even greater abstraction and even stronger platonist assumptions.

The growth of mathematical logic introduced as further elements the axiomatization of logic (the basic step in which was completed by Frege in 1879), the effort to incorporate the axiomatization of logic into that of mathematics, and the accompanying tendency, on the part of Frege and Giuseppe Peano, to interpret rigorous axiomatization as formalization. Frege carried the development much further by undertaking to develop the whole of

arithmetic and analysis in a formal system that is essentially a system of set theory.

At the turn of the twentieth century the entire development reached a crisis with the discovery of the paradoxes of set theory, which showed that the concept of class or set as it was then being used had not been sufficiently clarified. Much of the foundational research of the early twentieth century—and not only in the axiomatization of set theory—was directed at problems posed or believed to have been posed by the paradoxes.

In that period emerged three general viewpoints, each of which had its own program based on a distinctive attitude toward the question of platonism. The most radical was intuitionism, based on L. E. J. Brouwer's critique of the whole idea of platonism. In contrast to Brouwer, David Hilbert had a firm commitment to the patronizing tendency in mathematics, but he held epistemological views that were fundamentally in accord with Brouwer's critique of platonism. Making use of the fact that no matter how platonist the mathematics formalized, questions of provability in a formal system are meaningful from a narrow constructivist point of view, Hilbert's school sought to secure the foundations of platonist mathematics by metamathematical investigation of formalized mathematics—in particular, by a proof of consistency. This viewpoint was called formalism, although the designation is misleading, since Hilbert never maintained that even platonist mathematics could be simply defined as a "meaningless" formal system.

Proponents of the third viewpoint, logicism, whose leading figure was Russell, continued to believe in Frege's program of reducing mathematics to logic. Accepting this program involved taking some platonist assumptions as intuitively evident.

A great deal of work in mathematical logic was directed toward clarifying and justifying one or another of these points of view. We might mention Brouwer's (informal) results on the impossibility of constructively proving certain theorems in analysis, Arend Heyting's formalization of intuitionist logic, the development of finitist proof theory by Hilbert and his coworkers, and Russell and A. N. Whitehead's *Principia Mathematica* as a much further development of mathematics within a system of set theory.

Nonetheless, the trichotomy of logicism, formalism, and intuitionism has probably never been the best classification of points of view in foundations. It does not take account of one of the philosophically most important problems, that of predicativity, or of some mathematical developments—such as the development of the semantics of logic by Leopold Löwenheim, Thoralf Skolem, Kurt Gödel, and Alfred Tarski—which were crucially important for later work. At any rate the schools no longer really exist. All of them had programs that encountered serious difficulties; further experience with set theory and the axiomatizations of Ernst Zermelo and Russell deprived the paradoxes of their apparently apocalyptic character; and specialized work in mathematical logic led more and more to the consideration of problems whose significance cut across the division of the schools and to looking at the results of the schools in ways which would be independent of the basic controversies. A decisive step in this development came in the early 1930s, with the discovery of Gödel's incompleteness theorem and the coming of age of formal semantics.

Some areas of the foundations of mathematics will be passed over here—in particular, we shall not go far into the significance of the fact that mathematics has applications to the concrete world, although historically the relation between mathematics and its applications has been very close, and the present sharp distinction between pure and applied mathematics is a rather recent development. For instance, we shall omit a special consideration of geometry. If the pre-twentieth-century view that geometry is a purely mathematical theory that nonetheless deals with actual space is correct, then the omission is unjustified. However, even the question whether this view still has something to be said for it is more intimately related to the philosophy of physics than to the problems on which we shall concentrate. Geometry as understood today by the pure mathematician, as the general study of structures analogous to Euclidean space, raises no philosophical problems different from those raised by analysis and set theory.

§1. THE AXIOMATIC METHOD

As we said, we shall begin our discussion with the axiomatic method. Consideration of the notion of an informal axiomatic system leads to the notions of formalization and formal system. Through this process, especially through the last step, mathematical theories become themselves objects of mathematical study. The exploitation of this possibility is perhaps the specifically modern move in the study of the foundations of mathematics and has led to an enormous enrichment of the subject in the last hundred years.

1.1. AXIOMATIZATION. Ever since Euclid, axiomatizing a theory has meant presenting it by singling out certain

propositions and deducing further ones from them; if the presentation is complete, it should be the case that all statements which could be asserted in the theory are thus deducible. Axiomatization has also come to mean a similar reduction of vocabulary, in that certain notions should be taken as primitive and all further notions which are introduced in the development of the theory should be defined in terms of the primitive ones. In essence this is the conception of an axiomatized theory that prevails today, although it has been developed in different directions.

There are important ambiguities concerning the means of deduction and definition to be admitted in the development of the theory. Here informal axiomatics always makes use of some general background that can be used in developing the theory but is not itself included in the axiomatization. In modern mathematics this background typically includes logic and arithmetic and usually also analysis and some set theory. For example, in an axiomatic theory concerning objects of a certain kind, one permits oneself very quickly to make statements about sequences and sets of those objects, to introduce concepts defined in terms of the primitives of the theory by means of these general mathematical devices, and to make inferences that turn on laws of arithmetic, analysis, or set theory. Such notions often enter into the statement of the axioms themselves. We shall presently say more about the significance of this procedure.

It might seem natural to require provisionally that the means of deduction and definition be restricted to those of pure logic, for logic is supposed to contain those rules of correct inference which have the highest degree of generality and which must be applied in all sciences. We would then regard an axiomatization as only partial if deductions from it required the use of methods of the special sciences—in particular, branches of mathematics (likewise if, in addition to the primitives, notions other than purely logical ones entered into the definitions). An axiomatic theory would then consist of just those statements that are deducible by purely logical means from a certain limited set of statements and of the statements that can be obtained from these by definitions expressible purely logically in terms of the primitives.

It seems possible that such an axiomatic system was the objective toward which Euclid was striving. He evidently did not intend to allow himself general mathematical notions, such as arithmetical ones, for he included propositions involving such notions among his axioms and undertook to develop some of number theory from the axioms in Books VII–IX. Even some of Euclid's well-known failures to achieve this degree of rigor—for example, his assuming in his very first proof that two circles with the center of each lying on the circumference of the other will have two points of intersection—might have arisen because he saw them as immediate deductions from the meaning of the concepts involved. Of course, a rigorous theory of definition would require definitions to be given or axioms to be explicitly stated in such a way that such deductions do proceed by mere logic.

A perfectly satisfactory axiomatization in this form certainly was not possible in Euclid's time; it probably had to wait for two developments that did not take place until the late nineteenth century, Frege's discovery and axiomatization of quantification theory and the Dedekind-Peano axiomatization of arithmetic. (Nonetheless, considerable progress was made prior to these developments.)

This remark points to a limitation of the conception we are considering, for it does not give a meaning to the idea of an axiomatization of logic itself, although such axiomatization has played a vital role in modern foundational studies. Appreciation of this point leads to the concept of a formal system, but before we consider this concept let us observe a consequence of the axiomatization of a theory.

1.2. THE ABSTRACT VIEWPOINT. Suppose a theory is so completely axiomatized that all concepts of the special theory which are used in statements and deductions are explicitly given as primitives and all special assumptions underlying the proofs are disengaged and either stated among or deduced from the axioms. This means that the validity of the deductions does not at all depend on the actual meaning of the primitive terms of the special theory. It follows that the formal structure determined by the primitive concepts and the axioms can have a more general application than they have in the given special theory, in the sense that we could by any choice of interpretation of the primitive terms obtain a deductive system of hypotheses concerning some subject matter, even though the hypotheses will in many cases be false.

This fact is of crucial importance in the study of axiom systems. We can then think of a *model* of an axiomatic theory as a system of objects and relations that provides references for the primitive terms so that the axioms come out true. We can think of axiomatization as having proceeded with a particular model in mind, but this need not have been the case; at any rate, interest attaches to the study of other possible models. (Although we may, in this discussion, allow means of deduction that

go beyond pure logic, it ought to be the case that if a proposition is deducible from the axioms of the theory, then it must be true in all models of the theory. It might be reasonable to take this as a sufficient condition of deducibility, but if so it seems that the notion of model will have to have a relativity comparable to that of the notion of deducibility.)

For example, suppose we consider absolute geometry—that is, Euclidean geometry without the parallel postulate. Then any model either of Euclidean geometry or of the standard non-Euclidean geometries will be a model of absolute geometry. If the parallel postulate is deducible from the other axioms of Euclidean geometry—that is, from the axioms of absolute geometry—then it must be true in every model of absolute geometry. The construction of models for non-Euclidean geometries showed that this is not the case. We call an axiom of a system independent if it is not deducible from the others. Thus, if the theory obtained by dropping an axiom $\mathcal{A}$ has a model in which $\mathcal{A}$ is false, then $\mathcal{A}$ is independent.

Another possibility, which has been much exploited in modern mathematics, is to replace a system of primitive terms and axioms by what amounts to an explicit definition of a model of the axioms. Thus, suppose Euclidean geometry is formulated with two primitive predicates (following Alfred Tarski in "What Is Elementary Geometry?," 1959):

"$\beta(x,y,z)$",

meaning "x, y, and z are collinear, and y lies between x and z or $y = x$ or $y = z$," and

"$\delta(x,y,z,w)$",

meaning "x is the same distance from y as z is from w." (The variables here range over points, which in the informal theory must be thought of as a primitive notion.) Then we can define a *Euclidean space* as a triple $\langle S,B,D \rangle$, where S is a set of entities called "points," B a ternary relation on S, and D a quaternary relation on S, such that the axioms of Euclidean geometry hold. Then to any theorem proved from these axioms corresponds a statement of the form "Every Euclidean space is such that … ." A number of attempts to characterize mathematical structures axiomatically have led in a similar way to explicit definitions of abstract types of structure. This is regarded, for more than historical reasons, as a fruit of the axiomatic method. The search for an axiomatic basis for a mathematical theory is also the search for a formulation of the arguments in a fashion which will make them more generally applicable, giving them a generality which can be expressed in the definition of a general type of structure.

1.3. FORMALIZATION. Whereas one development of the axiomatic method tends to the replacement of axioms by definitions, another leads to the conception of a formal system. One result of the axiomatization of a theory was that the meaning of the primitive terms became irrelevant to the deductions. If we carry this abstraction from meaning to its limit, we can cover the case of axiomatizations of logic and resolve once and for all the question of what means of deduction are to be allowed. That is, we put into the construction of an axiom system a complete specification of all the means of inference to be allowed (for example, logic and basic mathematics) in the form both of further axioms and of rules of inference that allow us to infer from statements of certain given forms a statement of another given form. If this is done with utmost rigor, so that use can be made of only as much of the meaning of the terms as is specified in axioms and explicit definitions, then the system is specified simply in terms of the designs of the "linguistic" forms in which it is expressed. "Linguistic" is put in quotation marks because, invariably, much of the language has been replaced by an artificial syntax. We are left with a specification of certain strings of symbols as "axioms" and certain rules, each of which allows us to "infer" a new string from certain prior ones. The strings which we can obtain from axioms by successive application of the rules can be called *theorems*.

A proper explanation of the concept of a formal system requires somewhat more apparatus. The exactness of this procedure requires that the strings of symbols used be constructed out of preassigned material, which we can assume to be a finite list of symbols. Among the strings of these symbols we single out a subclass that we call formulae (or well-formed formulae, wffs), which are those strings to which, in an interpretation, we would give a meaning. (The non-wffs correspond to ungrammatical sentences.) Then a certain class of formulae is singled out as the axioms. The class of theorems can be defined as the closure of the axioms under certain operations; that is, rules of the following form are specified:

(R_i). If $\mathcal{A}_1, \cdots, \mathcal{A}_{r_i}$ are theorems and $\mathcal{R}_i(\mathcal{A}_i, \cdots \mathcal{A}_{r_i})$, then $\mathcal{B}$ is a theorem, where $\mathcal{R}_i$ is some relation on strings of the symbols of the system.

So the definition of *theorem* is an inductive definition with the clauses (R_i) and

every axiom is a theorem.

In this setting we can resolve another ambiguity of our original rough conception of axiomatization. The question arises concerning what conditions a class of statements must satisfy to be appropriate as the axioms of an axiomatic theory. Various epistemological desiderata, such as self-evident truth for the intended model, are put aside once we take the abstract point of view. Another requirement that has been found natural in the past is that both individual axioms and the class of axioms as a whole should have a certain simplicity. What there is in the way of general theory about the simplicity of individual axioms has not played much of a role in investigations of the foundations of mathematics, although much effort has been expended in replacing individual axioms with simpler ones or in finding systems of axioms which have particular advantages of "naturalness" for intended applications.

In order to characterize the important axiom systems which have been used in the past we shall have to place some limitation on the class of axioms. In the traditional cases the class has been finite. However, the formalization of such an axiomatic system can give rise to an infinite system—for example, if we take as axioms all instances of a certain schema.

The limitation which is used instead of a finite class of axioms is based on the fact that the notions of formula, axiom, and theorem are to be syntactically specified. Then the requirement is that there be a mechanical, or effective, procedure for deciding whether a given formula is an axiom and whether a given inference (of a formula from finitely many premises) is correct according to the rules of inference. This requirement is natural in the light of the idea that a proof of a statement in an axiomatic theory should contain all the mathematically significant information needed to show that the statement is indeed assertible in the theory. That would not be the case, it is argued, if something beyond mechanical checking were needed to determine the correctness of the proof. (It should be pointed out, however, that generalizations of the concept of formal system in which this condition is not satisfied are frequently used in mathematical logic.)

The notion of a formal system gives the highest degree of generality, in that there is no element of the symbolism whose interpretation is restricted. Indeed, it permits much of what we might want to say about an axiomatic theory to be formulated without reference to interpretation, since the formulae, axioms, and rules of inference are specified without reference to interpretation, and what is a theorem is then defined, again without such reference. An entire division of the theory of formal systems—what is usually called syntax—can thus be built up with no more than a heuristic use of interpretation. In particular, the intensional notions—concept, proposition, etc.—relied on so far in the informal exposition can be eliminated.

The concept of a formal system also brings to the formulation of the theory the highest degree of precision, at the cost of a still further idealization in relation to the concrete activities of mathematicians. Furthermore, the concept not only gives a refined formulation to axiomatizations and allows a mathematical study of axiom systems of a more general scope than was possible without it but also makes possible a precise formulation of differences about mathematical methods. Carrying the axiomatic method to this limit makes possible a new approach to a wide variety of questions about the foundations of mathematics.

Inasmuch as axiomatization is a rendering of a theory in a more precise formulation (if not a singling out of some particular aspect of the theory), the axiomatized theory cannot be identified in every respect with what has gone before. It can replace, however, what has gone before and actually has done so in many cases. The passage from axiomatization to formalization is in an important respect more radical than the various stages of informal axiomatization, and we can therefore regard a formalization of a theory as not so much a more precise formulation of the theory as an idealized representation of it. The process of replacing expressions of natural language by artificial symbols, which goes on in all mathematical development, is here carried to an extreme. For example, we lay down by a definition what are "formulae" and "proofs" in the system, whereas informally we rely for the notion of sentences on our more or less unanalyzed linguistic sense, and for proofs we rely on this sense, on mathematical tradition, and on intuitive logic. In particular, formulae and formal proofs are of unbounded length and complexity, without regard to the limits of what we can perceive and understand.

With this goes the fact that the basic general notions with which we operate in formulating and reflecting on theories—sentence, proposition, deduction, axiom, inference, proof, definition—are replaced in the formalized version by specifically defined, more or less simplified and idealized substitutes. In particular, although we "interpret" formalized theories, the relation between a sign or a formal system and its reference in some model is a "dead" correspondence, an aspect of a purely mathematical relation between two systems of objects. This enables one to avoid the intractable problems of how lin-

guistic expressions come to have "meaning" and, with it, reference and is therefore an extremely valuable piece of abstraction. But it is an abstraction; moreover, it does not mean that the informal linguistic and intellectual apparatus disappears altogether, since it will still be used in the setting up and investigation of the formalized theory. In fact, one of the results of formalization is a sharper separation between what is within the theory and what belongs to discourse about it—that is, to the metatheory. If the metatheory is in turn axiomatized and then formalized, the same situation arises at the next-higher level.

The importance of this observation is difficult to assess, but it is relevant to a number of problems we shall discuss later—in particular, attempts to argue from results of mathematical logic to philosophical conclusions.

§2. EPISTEMOLOGICAL DISCUSSION

2.1. A PRIORI KNOWLEDGE. We shall now put the matter of axiomatization and formalization aside and consider mathematics from the point of view of general epistemology. The guiding thread of our discussion will be the fact that a powerful tradition in philosophy has regarded mathematics, or at least a part of it, as a central case of a priori knowledge. This means that reflection on mathematics has been at the center of philosophical discussion of the concept of a priori knowledge.

The characteristics of mathematics which have led to the conclusion that mathematics is a priori are its abstract character and accompanying enormous generality and its great exactitude and certainty, which, indeed, have traditionally been considered absolute. Thus, even before setting forth a developed logical analysis of the concept of number, we find that the effort to interpret "2 + 2 = 4" as a hypothesis that can be checked by observation runs into obvious obstacles. It is perhaps not so vital that the statement refers to abstract entities, numbers, which are not the sort of thing we observe. The concept of number certainly does apply to empirically given objects, in the sense that they can be counted and that the numbers thus attributed to them will obey such laws as "2 + 2 = 4." Therefore, the proposition could so far be taken as a law concerning such entities. Even then its range of application is so enormous, extending over the entire physical universe, that it seems evident that if it were taken as a hypothesis, it would be stated and used in a more qualified way, at least by critically minded scientists. In other words, the certainty that we attribute to elementary arithmetical propositions would be quite unwarranted if they were laws based on observation. Even in the case of mathematical principles to which we do not attribute this degree of certainty, such as the axiom of choice and the continuum hypothesis, the possible "contrary evidence" would arise from the deductive development of the theory involved (in the examples, set theory), not from observation.

Moreover, it seems that we ought to be able to conceive of a possible observation which would be a counterinstance. Although it is perhaps not evident that this is impossible, the ideas that come to mind lead either to descriptions of doubtful intelligibility or to the description of situations where it seems obviously more reasonable to assume some other anomaly (such as miscounting or the perhaps mysterious appearance or disappearance of an object) than to admit an exception to "2 + 2 = 4."

Another difficulty is that the concept of number must apply beyond the range of the concrete entities which are accessible to observation; such abstract entities as mathematical objects must be subject to counting, and this seems also to be the case for transcendent entities.

The foregoing considerations could be developed into decisive arguments only with the help of both a more developed formal analysis of number and a more detailed discussion of the relation between arithmetical laws and actual counting and perhaps also of the role of mathematics in empirical science. In any case, they do not tell against another form of the denial that arithmetic is a priori, the view that arithmetical laws are theoretical principles of a very fundamental sort, which we are therefore far more "reluctant to give up" in a particular situation than more everyday beliefs or impressions or even than fundamental theoretical principles in science. Such a view would nonetheless take it to be conceivable that in response to some difficulty in, say, particle physics a new theory might be formulated which modified some part of elementary arithmetic.

2.2. MATHEMATICS AND LOGIC. The above considerations show why it is necessary to add technical analysis to the epistemological discussion. We shall take as our guiding thread the attempt to show that mathematics—in particular, arithmetic—is a part of logic. This attempt has led to some of the most important results in the logical analysis of mathematical notions. The view that mathematics can be reduced to logic is one of the principal general views on the foundations of mathematics which we mentioned earlier; it goes generally by the name of logicism, and its classic expression is in the writings of Frege and Russell.

Even if successful, the reduction of mathematics to logic could not by itself give an account of how there can be a priori knowledge in mathematics, for it would only reduce the problem of giving such an account to the corresponding problem with regard to logic. Nonetheless, the a priori character of mathematics has traditionally been found perhaps slightly less certain than that of logic. The obvious fact that one of the primary tasks of mathematics is the deductive development of theories has been found to be one of the most powerful supports of the claim that mathematics is a priori. We can expect that a successful reduction of mathematics to logic will simplify the problem of a priori knowledge, and not only by replacing two problems by one. Logic is more unavoidable: We cannot get anywhere in thinking without using logical words and inferring according to logical rules. This would suggest that logic is in fact more basic than mathematics and more certainly a priori. (It would also suggest that philosophical treatments of logic are more liable to circularity.) Moreover, in the course of history philosophers have invoked sources of evidence for mathematics which are at least apparently special, such as Kant's pure intuition. Thus, a reduction of mathematics to logic might make superfluous certain difficult epistemological theories.

The claims of logicism are based in large part on mathematical work in axiomatics. A number of nineteenth-century investigations showed that the basic notions of analysis—for example, rational, real, and complex number—could be defined, and the basic theorems proved, in terms of the theory of natural numbers and such more general notions as class and function. At the same time, axiomatic work was done in the arithmetic of natural numbers, culminating in the axiomatization of Richard Dedekind (1888) and Peano (1889). The movement toward formalization began somewhat later, with the work of Frege and of the school of Peano.

Thus, the effort to reduce mathematics to logic arose in the context of an increasing systematization and rigor of all pure mathematics, from which emerged the goal of setting up a comprehensive formal system which would represent all of known mathematics with the exception of geometry, insofar as it is a theory of physical space. (But of the writers of that generation only Frege had a strict conception of a formal system.) The goal of logicism would then be a comprehensive formal system with a natural interpretation such that the primitives would be logical concepts and the axioms logical truths.

We shall be guided by Frege's presentation, although he did not go very far in developing mathematics within his system and of course the system turned out to be inconsistent. Nonetheless, it is already clear from Frege's work how to define the primitives and prove the axioms of a standard axiomatization of arithmetic. We shall begin with some discussions of the notions of number and class, which are crucial for the reduction and for the foundations of mathematics generally.

2.3. COUNTING AND NUMBER. In order to be clearer about the concept of number, we might start with the operation of counting. In a simple case of carefully counting a collection of objects, we perhaps look at and point to each one successively, and with each of these directions of the attention we think of or pronounce one of a standard series of symbols (numerals) in its place in a standard ordering of these symbols. We are careful to reach each of these objects once and only once in the process. We thus set up a one-to-one correspondence between the objects and a certain segment of the series of numerals. We say that the number of objects in the collection is __________, where the blank is filled by the last numeral of the series.

Before pursuing this matter further, let us examine the series of numerals itself. We have certain initial symbols and rules for constructing further symbols whose application can be iterated indefinitely. We could simplify the situation in actual language and suppose that there is one initial symbol, say "|," and a generating operation, concatenation of another "|," so that the numerals will be |, ||, |||, ||||, · · ·, It is not clear, however, that it is merely a matter of "practical convenience" that ordinary numerals are, in the long run, considerably more condensed: If a string of several million "|'s" were offered as a result of counting, one would have to count them to learn what the number was.

However, it is worth asking whether the pure notion of natural number requires more than the possibility of generating such a string of symbols. By "symbols" do we mean here blobs of ink? Only with certain reservations. The particular blobs which we have produced are not at all essential; if we write others—|, ||, |||, ||||, · · · —they will do just as well. In fact, we could have chosen symbols of quite different forms and still have produced something equivalent for our purposes, such as +, ++, +++, · · ·, or something not consisting of marks on paper at all, such as sounds, which are, of course, actually used. As long as it is capable of representing to us the process of successive generation by which these sequences of symbols are produced, anything will do—any collection of perceptible

objects that can be placed in one-to-one order-preserving correspondence with our first sequence of symbols.

Thus, the blobs of ink serve as the representatives of a quite abstract structure. This abstraction allows us (even on a subordinate level) to disregard some limitations of the blobs besides their particularity and accompanying boundedness to a particular place and time. They are constructed according to a procedure for generating successive ones, and what matters is the structure embodied in the procedure, not any particular limitations that might be encountered in carrying it out. On a sufficiently abstract level we say that we can continue to generate symbols indefinitely, although life is too short, paper and ink run out, the earth perhaps disintegrates, etc.

Here we have already taken the step of introducing abstract entities. In a weak form this could be represented as taking certain abstract equivalence relations between entities (e.g., marks on paper) as criteria of identity for new kinds of entities (e.g., symbols as types or, further, numbers). But we have already reached a point where more is involved, since the abstract entities which are represented by all the marks of a given equivalence class belong to a series which can be continued far beyond any practical possibility of constructing representatives. We can create a "pseudo-concrete" model by appealing to space, time, and theoretical physics, but then we are already depending on abstract mathematical objects. Given that we do think of numerals as referring to numbers, it is natural to introduce the apparatus not only of identity but also of quantification. Certain uses of such quantification, however, will involve still stronger presuppositions than we have uncovered up to now, and we shall discuss these when we consider platonism and constructivism.

2.4. **AXIOMS OF ARITHMETIC**. We have so far taken for granted that the natural numbers are obtained by starting with some initial element 0 and iterating an operation of "successor" or "adding 1." This is the basis for an especially simple axiomatization of the theory of natural numbers, that of Dedekind and Peano, in which the primitives are "0," "number" ("*NNx*"), and "successor" (which we shall give as a relation: "*Sxy*" means "*y* is successor of *x*"). Then the axioms are

(1) $$NN0.$$

(2) $$NNx \supset (\exists!;y)(Nny \;\&\; Sxy).$$

(3) $$\neg S0x.$$

(4) $$Sxz \;\&\; Syz. \supset x = y.$$

(5) $$(F)[F0 \;\&\; (x)(y)(Fx \;\&\; Sxy\;.\supset Fy)\;.\supset (x)(NNx \supset Fx)].$$

In (5), "(*F*)" may be read "for all properties *F*," but for the present we shall not discuss just what this means. We do not need to suppose that precisely what properties there are is determined in advance, but we have to acknowledge that if it is not determined what properties there are, then it may not be determined precisely what natural numbers there are.

We could think of the natural numbers as given by a kind of inductive definition:

(*a*) *NN*0.

(*b*) If *NNx*, then *NN*(*Sx*).

(*c*) Nothing is a natural number except by virtue of (*a*) and (*b*).

However, in this case we have to suppose that the successor relation is given in such a way that axioms (2), (3), and (4) are evident. We might think of "0" as represented by "|" and the successor function as represented by the addition of another "|" to a string. Then there is apparently an appeal to spatial intuition in regarding these axioms as evident. In that event the induction principle (5) will be in some way a consequence of (*c*). It could be regarded simply as an interpretation of (*c*), or one might argue, as Ludwig Wittgenstein apparently did at one time (see Friedrich Waismann, *Introduction to Mathematical Thinking*, Ch. 8), that the meaning of all natural numbers is not given to us by such specifications and our independent concept of "all" and that the induction principle functions as a criterion for a proposition's being true of all natural numbers.

2.5. **THE CONCEPT OF CLASS (SET)**. Before we discuss further the notion of number it is necessary to give some explanation of the notion of class or set. We shall consider two explanations, one suggested by Cantor and one suggested by Frege.

2.5.1. *Frege's explanation.* Instead of the term *class* or *set*, Frege used the phrase "extension of a concept." Frege's usage is based on the tendency to regard the predicates of a language as standing in quantifiable places—

John is a Harvard man.

Henry is a Harvard man.

∴ John and Henry have something in common—

and the tendency to derive from general terms abstract singular terms, which are usually explained as referring to properties or attributes.

These two tendencies can be separated. Frege regarded predicates in context as in fact referring, but to concepts, not to objects. Concepts, like the predicates themselves, have argument places; Frege called both predicates and concepts "unsaturated" because only with the argument place filled by an object (in the case of a predicate, a proper name) could they "stand by themselves." A notation which expresses his conception is that of the second-order predicate calculus, in which the above conclusion might be symbolized (misleadingly) as $(\exists F)[F(\text{John})\ \&\ F(\text{Henry})]$. An expression which is syntactically appropriate for denoting an object cannot denote a concept, and vice versa.

The extension of a concept, then, is simply an object associated with the concept in such a way that if two concepts apply to the same objects, they have the same extension—that is,

(6) $$\hat{x}Fx = \hat{x}Gx. \equiv (x)(Fx \equiv Gx),$$

where $\hat{x}Fx$ is the extension of the concept F. This is essentially Frege's famous axiom V (*Grundgesetze der Arithmetik*, Vol. I, p. 36; Frege's notion of concept can interpret the quantifiers in our axiom 5).

2.5.2. *Cantor's explanation.* Cantor characterized a set as "jedes Viele, welches sich als Eines denken lässt, d.h. jeden Inbegriff bestimmter Elemente, welcher durch ein Gesetz zu einem Ganzen verbunden werden kann" ("every many, which can be thought of as one, that is, every totality of definite elements which can be combined into a whole by a law"; *Gesammelte Abhandlungen*, p. 204). "Unter einer 'Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die 'Elemente' von M genannt werden) zu einem Ganzen" ("By a 'set' we understand any collection M of definite well-distinguished objects of our intuition or thought, which are called the 'elements' of M, into a whole"; p. 282).

It is virtually impossible to explain Cantor's idea of set without using words of the same general type, only vaguer ("collection," "multitude," *Inbegriff*). We can perhaps approach it by mentioning a few ways in which multitudes are thought of as unities: by being thought of by means of a predicate—that is, by being brought under a concept in Frege's sense—so that Frege's extensions could perhaps be regarded as sets, or by being in some way brought to the attention at once, even without the intervention of language; in particular, a finite number of objects of perception can constitute a set. That the objects must be "determinate and well-distinguished" means that it must be determinate what the elements are, that identity and difference be well-defined for the elements, and that a set must be determined by its elements.

One is inclined in this connection to think of a set as "composed" of its elements, but this is not essential and might lead to confusion of a set with a spatiotemporal sum, but a portion of space or time (for example, a geometric figure) can be partitioned in a number of ways, so the sets of the parts will be different but the sum will always be the same.

The picture of finite sets can be extended in such a way that one might imagine an "arbitrary" infinite set independent of any predicate. Suppose it is to be a set S of natural numbers. We go through the natural numbers one by one deciding for each n whether n is a member of S ($n \in S$) or not. Although the determination takes infinitely long, it is determined for each n whether $n \in S$. (Or we might imagine its being done all at once by God.)

2.5.3. *Difficulties in these conceptions.* Both Cantor's and Frege's conceptions of sets have difficulties which did not come clearly to the consciousness of logicians and set-theorists until the discovery of the set-theoretical paradoxes, discussed below. We shall merely mention here a source of difficulty. In both theories a set or extension is supposed to be an object, capable of being itself a member of sets. Cannot this give rise to circularities—that is, that a set is formed from or constituted by certain objects, among them itself?. (Or, in Frege's terms, among the objects in the range of the quantifiers on the right side of formula 6 are $\hat{x}Fx$ and $\hat{x}Gx$ themselves, so that the identity condition for these objects, which from Frege's point of view was part of their essence, seems to depend on particular facts about them.)

We shall not say anything at the moment about the particular form the difficulties take or about how to resolve them. We shall continue to use second-order quantification somewhat vaguely; one can interpret the variables as ranging over Frege's concepts, in most cases over classes or even over intensional entities, as might have been suggested by our original word "property."

2.6. FREGE'S ANALYSIS OF NUMBER. We can now proceed to the main steps of Frege's argument for the thesis that arithmetic is a part of logic. Frege observed that a necessary and sufficient condition for, say, the number of F's (which we shall write as "N_xFx") to be the same as the number of G's is that there should be a one-to-one corre-

spondence of the *F*'s and the *G*'s. (In that case we say they are numerically equivalent.) This criterion, which is quite general—that is, not restricted to the case where there are only finitely many *F*'s or *G*'s—had already been exploited by Cantor to generalize the notion of cardinal number to infinite classes. It can be justified by our discussion of counting and number, above.

On the basis of a one-to-one correspondence between the *F*'s and $\{1, \cdots, n\}$ we are prepared to say that the number of *F*'s is *n*. But no such correspondence can then exist with $\{1, \cdots, m\}$ for any $m \neq n$, and if by the same criterion there are *n* *G*'s, then by composition we can set up a one-to-one correspondence between the *F*'s and the *G*'s. If there are *m* *G*'s for $m \neq n$, we cannot. So we say that there are *n* *F*'s if and only if a one-to-one correspondence exists between the *F*'s and $\{1, \cdots, n\}$, and in that case there are *n* *G*'s if and only if there is a one-to-one correspondence between the *F*'s and the *G*'s. Writing "there are *n* *F*'s" as "$(\exists x)_n Fx$," we have that if $(\exists n)[(\exists x)_n Fx]$,

(7) $N_x Fx = N_x Gx. \equiv$ the *F*'s and the *G*'s are numerically equivalent.

Since we have no independent criterion for the case where there are infinitely many *F*'s, we take (7) to be true by definition in that case. We then have Frege's criterion.

Frege then defined a relation *H* as a one-to-one correspondence of the *F*'s and the *G*'s if and only if for every *F* there is exactly one *G* to which it bears the relation *H* and vice versa—in symbols,

(8) $(x)[Fx \supset (\exists! y)(Gy \ \&\ Hxy)] \ \&\ (y)[Gy \supset (\exists! x)(Fx \ \&\ Hxy)]$,

where "$(\exists! x)(\cdots x \cdots)$" can be defined in first-order logic:

(9) "$(\exists! x)(\cdots x \cdots)$" for "$(\exists x)[\cdots x \cdots \ \&\ (y)(\cdots y \cdots \supset y = x)]$".

Thus, numerical equivalence can be defined by a formula "$(\exists H)\mathcal{S}(H,F,G)$," where "$\mathcal{S}(H,F,G)$" is an abbreviation for a first-order formula, namely, the expansion of (8) in terms of (9).

The relation of numerical equivalence is an equivalence relation; Frege's idea was, in effect, to define cardinal numbers as the equivalence classes of this relation. This definition, however, requires a powerful use of the notion of extension which is allowed by his axiom (6). In other words, $N_x Fx$ is to be the extension of the concept concept numerically equivalent to the concept *F*—that is, we define

(10) "$N_x Fx$" for "$\hat{G}(\exists H)\mathcal{S}(H,G,F)$".

(In fact, in the *Grundgesetze*, Frege avoided applying the extension operator to a second-order variable by appeal to formula 6: *G* can be replaced by its extension. We define "$\hat{G}\mathcal{F}(G)$" as $\hat{y}(\exists G)[y = \hat{x}Gx \,.\, \mathcal{F}(G)]$".)

Formula (10) gives a definition of Cantor's general concept of cardinal number, so we can prove (7); no further use of axiom V is needed for the definition of the natural numbers and the proof of the axioms (1)–(5). We now define Peano's primitives—"0," "*Sxy*" ("*y* is the successor of *x*"), and "*NNx*" ("*x* is a natural number"):

(11) "0" for "$N_x(x \neq x)$,"

for then (7) yields $N_x Fx = 0 \equiv \neg(\exists x)Fx$.

Intuitively, $n + 1 = N_x(x = 0 \vee \cdots \vee x = n)$; this result will be reached if we define "*Sxy*" as follows:

(12) "*Sxy*" for "$(\exists F)\{y = N_w Fw \ \&\ (\exists z)[Fz \ \&\ N_w(Fw \ \&\ w \neq z) = x]\}$".

Intuitively, the number of *F*'s is one more than the number of *G*'s if there is an *F* such that the number of the *rest* of the *F*'s is precisely $N_x Gx$. Definition (12) implies that in this case $S(N_x Gx, N_x Fx)$.

The remaining primitive is defined by an ingenious device (already present in Frege's *Begriffsschrift*), which yields mathematical induction: we want to define "*NNx*" so that something true of 0 and of the successor of anything of which it is true is true of every natural number—that is,

(13) $F0 \ \&\ (x)(y)(Fx \ \&\ Sxy. \supset Fy) \,.\supset (x)NNx \supset Fx)$.

But this will be immediate if we define "*x* is a natural number" as "*x* falls under every concept *F* which 0 falls under and which is such that any successor of whatever falls under it also falls under it"—that is,

(14) "*NNx*" for "$(F)\{F0 \ \&\ (x)(y)(Fx \ \&\ Sxy. \supset Fy) \,.\supset Fx\}$".

To prove the other axioms: (1) is immediate from (14); that *S* is one-to-one and that 0 is not the successor of anything follow from (12) together with (7).

2.7. DIFFICULTIES IN LOGICISM. The first difficulty with Frege's construction is certainly the use Frege made of the notion of extension. We have alluded to difficulties with the ideas of set theory; they affected Frege's system through Russell's deduction in 1901 of a contradiction from (6). (For Russell's initial exchange of letters with Frege, see van Heijenoort, 1967). We shall discuss Rus-

sell's paradox and other paradoxes and the difficulties of the concept of class below.

Nonetheless, it turns out that a reasonably secure system of set theory can be developed in any one of a number of ways that are more than sufficient for the definition of Peano's primitives and proof of his axioms. In fact, no part of the axiomatic apparatus of a system of set theory which gives rise to any doubts as to consistency is really necessary for this reduction; we can say that if the development in set theory of a branch of mathematics necessarily involves the stronger and more problematic parts of set theory, this is due to the nature of the branch of mathematics itself, not the reduction to set theory.

This success is not without loss for the development of arithmetic: it seems that in the more natural set-theoretical systems (the theory of types, Zermelo's set theory) no definition of "N_xFx" can be given with the same appearance of naturalness as in (10). The consequences of Russell's theory of types are more serious: The numbers must be duplicated at each type. What one usually ends up doing is identifying the numbers in a somewhat arbitrary way with a sequence of sets of the required order type.

Given that all this has been done, in what sense is the enterprise a reduction of arithmetic to set theory, and in what sense is it a reduction to logic? To take up the last question first, obviously the construction does not reduce arithmetic to logic unless the principles of the set theory involved can count as logical principles. The notion of class is not very far removed from concepts which played a role in traditional logic; from that point of view it is not at all evident why the first-order predicate calculus, which is already a considerable extension of the traditional formal apparatus, should count as logic and the theory of classes should not.

One difference is that whereas a valid formula of first-order logic will yield a truth if the quantifiers are interpreted to range over any domain of objects whatsoever, and without regard to its cardinal number in particular, set theory involves existence assumptions, so the domain over which the quantifiers range must be large enough to contain representatives for the sets whose existence is implied by the formula in question. In Frege's procedure these assumptions were embodied in the admission as a term of an abstract "$\hat{x}Fx$" for any predicate "F," and simple nonparadoxical instances of (6) already require that Frege's universe contain infinitely many objects.

Frege, of course, regarded (6) as a logical principle, a view which was fairly well refuted by its inconsistency. It would be much more reasonable to regard set theory as logic if its existence assumptions all followed from a single general principle, such as (6). But the analysis of the foundations of set theory stimulated by the paradoxes points to the opposite conclusion: Any very definite system of existential postulates will prove incomplete in the sense that it is always possible to construct further existential postulates that are stronger (in the sense of first-order, or even second-order, logic). Moreover, these postulates assume a character not unlike principles of construction, so it is at least as natural to consider them hypothetical and analogical extensions of "constructions in pure intuition" as it is to consider them principles of logic. At any rate, if logic consists of the necessary principles of all coherent reasoning, then it seems evident that the stronger principles of set theory do not have this character; it is far from certain even that the weaker ones have it (perhaps even that all of first-order logic does). This being so, a reduction of arithmetic to set theory does little to increase the security and clarity of the foundations of arithmetic.

2.8. KANT'S VIEW. One of the purposes that Frege, Russell, and many later proponents had in mind in seeking to reduce arithmetic to logic was to show that no appeal to sensible intuition was necessary in arithmetic, as had been claimed by such empiricists as John Stuart Mill and by Kant in his theory of a priori intuition. Let us consider whether this purpose has been accomplished. Since Kant's view constitutes an independent effort to explain the a priori character of arithmetic, and since it is part of an extremely influential general philosophy, it deserves special mention.

Kant began by insisting that mathematical judgments (at least the most characteristic ones) were synthetic, rather than analytic. We shall not enter into the question of just what he meant by that. Provided that one remembers that the scope of logic was much narrower for Kant than it is for us, it is plausible to suppose that his claim that mathematical judgments are synthetic implies that the propositions of a mathematical theory cannot be deduced from logical laws and definitions. The case of Kant's principal example, the geometry of space, seems clear, given, for instance, the fact that there are consistent geometrical theories which differ with respect to certain fundamental principles, such as the parallel postulate. (Even here, however, one might claim that the difference in principles corresponds to a difference in the meanings of the primitive terms. In application to real space this

comes down to the question of "conventionalism" in geometry. W. V. Quine is probably right in holding that one cannot, in general, decide the question whether such a difference is merely a difference of meaning.)

The case of arithmetic presents a certain similarity if we deny that set theory is logic. The proofs in the set-theoretic development even of such elementary arithmetical laws as "2 + 2 = 4" depend on existential axioms of these theories. However, this does not mean that we can come as close to clearly conceiving the falsity of these principles as we can for the principles of geometry. Although we can easily enough set up a domain in which the existence postulates will fail, it is not clear that this counts as conceiving that the numbers 0, 1, 2, · · · should not exist.

Kant went on to maintain that the evidence of both the principles of geometry and those of arithmetic rested on the "form of our sensible intuition." In particular, he said that mathematical demonstrations proceeded by "construction of concepts in pure intuition," and thus they appealed to the form of sensible intuition. Mathematical proof, according to Kant, required the presentation of instances of certain concepts. These instances would not function exactly as particulars, for one would not be entitled to assert anything concerning them which did not follow from the general concept. Nonetheless, conclusions could be drawn which were synthetic, because the construction of the instance would involve not merely the pure concept as of an abstract structure but also its "schematism" in terms of the general structure of our manner of representing objects to ourselves.

Thus, geometric figures would obey the axioms of geometry even though these axioms were not provable by analysis of the concepts. At the same time, the constructions would serve to verify any existence assumptions involved. (Indeed, instead of existential axioms Kant spoke of postulates asserting the possibility of certain constructions.)

In the case of arithmetic Kant argued that in order to verify "7 + 5 = 12" one must again consider an instance, this time in the form of a set of five objects, and add each one in succession to a given set of seven. It seems that although the five objects may be quite arbitrary, even abstract, they will, if not themselves present to perception, be represented by symbols which are present and which exhibit the same structure. In fact, we find this structure even in the symbolic operations involved in the formal proofs of "7 + 5 = 12" either within a set theory or directly from axioms for elementary number theory—or even in the proof of the formula of *first*-order logic

$$\textbf{(15)}(\exists x)_7 Fx \,\&\, (\exists x)_5 Gx \,\&\, (x)\neg(Fx \,.\, Gx) \,.\supset (\exists x)_{12}(Fx \vee Gx),$$

which is the key to the proof of "7 + 5 = 12" in Frege's construction. We think of "$(\exists x)_n(Fx)$" expanded as follows:

$$\text{“}(\exists x)_0 Fx\text{” for “}\neg(\exists x)Fx\text{”.}$$

$$\text{“}(\exists x)_{n+1} Fx\text{” for “}(\exists x)[Fx \,\&\, (\exists y)_n(Fy \,\&\, y \neq x)]\text{”.}$$

The arguments for the claim that intuition plays an essential role in mathematics are inevitably subjectivist to a degree, in that they pass from a direct semantical consideration of the statements and of what is required for their truth to a more pragmatic consideration of the operations involved in understanding and verifying them (and perhaps even "using" them, in a broad sense) and to a metalinguistic reflection on formulae and proofs as configurations of symbols. Gottfried Wilhelm Leibniz had already emphasized the essential role of calculation with symbols in mathematics, and to Kant this role became an argument for the dependence of mathematics on sensible intuition.

We can see why the arguments must have this subjectivist character if we notice the complete abstractness of both set theory and arithmetic, which talk of objects in general in terms of logical operations (propositional combination, quantification) which are equally general. Even the specifically mathematical objects (sets and numbers) are subjected by the theory only to certain structural, relational conditions, so that they are not, as it were, individually identified by the theory. The content thus does not suggest any direct sensory verification; indeed, it seems that any proposition which is susceptible of such verification must contain some particular reference to space or time or to objects or properties which by nature occur only in space and time. Although it is Frege's construction and the development of set-theoretic mathematics which make this fact clear, Kant apparently was aware of it in the case of arithmetic, which he related closely to the pure categories and therefore to logic.

Nevertheless, it does not seem, at least in the light of philosophical and mathematical experience, that we can directly verify these propositions, or even understand them, independently of the senses. Determining the precise nature of the dependence of the operations of the mind in general on the senses is one of the central difficulties of all philosophies. But it is hard to maintain that we understand mathematical structures, or even the general notion of object which underlies them, without at least starting with a sensible representation, so that con-

crete explanations make use both of embodiments of the structures by perceptible objects and of reflection on symbolism. For instance, explanations of the notion of class can either make use of an appeal to language, as Frege's explanation does, or begin with the notion of a group of perceptible objects. (Indeed, it seems that even in the second case an appeal to language is sooner or later indispensable.)

Perhaps more decisive than these rather vague considerations is the fact that we cannot carry on any even fairly elaborate reasoning in mathematics without, as it were, placing ourselves at the mercy of a symbolic representation. Prior to the construction of a proof or calculation we do not know the answer to any substantial mathematical question. That the proof can be constructed, that the calculation turns out as it does, is, as it were, brute fact without which one cannot see any reason for the mathematical state of affairs being what it is. In *Über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral*, Kant gave this as his principal reason for asserting that mathematics proceeds by representing concepts in intuition, and in the *Critique of Pure Reason* the idea is again suggested in the discussion of "7 + 5 = 12" and the remarks about "symbolic construction" in algebra.

One might argue that the existence of a natural number n is verified by actually constructing a sequence of numerals up to that point. Such a construction provides a representation for the numbers up to n. It is noteworthy that either it or a mental equivalent is necessary for a full and explicit understanding of the concept of the number n. This gives some plausibility to the view that the possibility of such a representation rests on the "form of our sensible intuition," since everything belonging to the content of the particular realization is nonessential. It is perhaps permissible to speak, as Kant did, of "pure intuition," because we are able to take the symbols as representing or embodying an abstract order. This conception could be extended to the intuitive verification of elementary propositions of the arithmetic of small numbers. If these propositions really are evident in their full generality, and hence are necessary, then this conception gives some insight into the nature of this evidence.

However, the above description already ceases to apply when we pass to the construction, by a general rule, of the sequence of natural numbers and therefore when we consider large numbers, which we must describe in terms of general rules. Besides the "factor of abstraction" signalized in our being able to use sensory representations in thinking about the abstract structures they embody, there is also a factor of higher generality and the accompanying possibility of iteration, so that the sequence of natural numbers extends far beyond those represented by numerals it is possible actually to construct. Here the sense of the notion of "form of intuition" is less clear. Kant's idea, however, must surely be that the larger numbers are conceived only as an extension of the structures of our actual experience. The fact that the forms in question are, according to Kant, those of space and time means that the abstract extension of the mathematical forms embodied in our experience parallels an extension of the objective world beyond what we actually perceive.

Kant connected arithmetic with time as the form of our inner intuition, although he did not intend by this to deny that there is no direct reference to time in arithmetic. The claim apparently was that to a fully explicit awareness of number goes the successive apprehension of the stages in its construction, so that the structure involved is also represented by a sequence of moments of time. Time thus provides a realization for any number that can be realized in experience at all. Although this view is plausible enough, it does not seem strictly necessary to preserve the connection with time in the necessary extrapolation beyond actual experience. However, thinking of mathematical construction as a process in time is a useful picture for interpreting problems of constructivity (discussed below).

Kant's view enables us to obtain a more accurate picture of the role of intuition in mathematics, but, at least as developed above, it is not really satisfying, because it takes more or less as a fact our ability to place our perceptions in a mathematically defined structure and to see truths about this structure by using perceptible objects to symbolize it. The great attraction of Kantianism comes from the fact that other views seem unable to do any better: Frege, for example, carried the epistemological analysis less far than Kant in spite of his enormously more refined logical technique.

2.9. **CONVENTIONALISM**. Attempts to avoid dogmatism completely while still affirming the existence of a priori knowledge in mathematics have been made on the basis of conventionalism, the characteristic logical positivist view of a priori knowledge. This view in effect rejects the question of evidence in mathematics: Mathematical statements do not need evidence because they are true by fiat, by virtue of the conventions according to which we specify the meanings of the words occurring in

mathematics. Mathematics is therefore "without factual content" or even "empty."

Before we proceed to discuss this view we should distinguish it from two others which are associated with logical positivism, the view that mathematical statements are true by virtue of the meanings of the words in them and the view that they are analytic. The doctrine that mathematical statements are true by virtue of the meaning of the words they contain is somewhat vague and is likely to reduce to the doctrine that they are analytic, to conventionalism, or to something compatible with Kantianism or even with some form of direct realism. If there are objective relations of meaning which hold not merely by fiat, then there is as much need in this view for an account of the evidence of our knowledge of them as there is for the evidence of mathematics itself.

The view that mathematics is analytic has generally been associated on one side with logicism and on the other with conventionalism. The definitions of "analytic" that have been given have been such that logical truths were automatically analytic. If the thesis that mathematics is analytic was to say more than the thesis of logicism, the definitions had to be taken as explicating a concept which had a more direct epistemological significance, usually truth by virtue of meanings or truth by convention. (Once this has been done, the connection with logicism seems less important, in spite of the importance that the logical positivists attributed to it. Thus, one may explain the claim that the axioms of set theory are analytic by saying that they are "meaning postulates" in Carnap's sense, but one could argue equally well that the axioms of number theory are meaning postulates. Logicism was important to the logical positivists for other reasons: the reduction served as a methodological paradigm; it served the "unity of science.")

That the propositions of mathematics should be true by convention in a strong sense, that one should actually have set up conventions which determine that they should be true, seems possible only for "rational reconstructions" of mathematics by explicit construction of an axiom system and identification of the system with mathematics. If such a procedure could be carried out, there would still be room for discussion of the sense in which it showed that the mathematics practiced by those who are not interested in foundations is true by convention.

The usual conventionalist position appeals to rules specifying that certain propositions are to be true by convention or, more often, to rules of another sort (such as semantical rules of an interpreted formal system), from which it can be deduced that certain statements are true, the nature of the premises being such that they can be called conventions governing the use of expressions. (For example, the truth of any statement that is a substitution instance of a theorem of the classical propositional calculus can be deduced from the information contained in the truth tables for the propositional connectives. Then if the truth tables are regarded as semantical rules specifying the meanings of the connectives, then the theorems of classical propositional logic thus become true by virtue of these rules.)

In the simplest case—that of simply laying down, by rules or in individual instances, that certain sentences are to be taken as expressing true statements—something more seems to be required to justify this procedure as attributing "truth" to "statements." No serious philosopher, however, has been content to leave the matter at that.

Nonetheless, the procedure of specifying by rules runs into a difficulty essentially independent of the form of the rules and the manner in which they are interpreted. This difficulty, which was pointed out forcefully by Quine early in his career (in "Truth by Convention") and is perhaps implicit in remarks by Frege, is that the passage from the general statements which are the actual explicit conventions to the truth by convention of specific statements involves inference. So something essentially logical is not, on the face of it, reduced to convention by the analysis. The inferences will assume properties of generality (for example, the properties of the universal quantifiers) and of the conditional, since the rules will in all probability be of the form of conditionals—for instance, they may say that if a statement satisfies certain conditions, then it is true by convention. In the example that we gave, one needs in addition the laws of contradiction and of excluded middle: Application of the truth tables already supposes that each statement has one, and only one, of the two truth-values.

Quine showed that the attempt to regard the rules by which this inference proceeds as themselves valid by convention leads to an infinite regress. For example, suppose a rule is *modus ponens:* from "p" and "$p \supset q$" infer "q". This could be stated as the convention:

(16) If A and C are true and C is the result of substituting A for "p" and B for "q" in "$p \supset q$", then B is to be true.

Now, suppose that for some A' and B' we have proved that A' and C' are true by convention, where

(17) C' is the result of substituting A' for "p" and B' for "q" in "$p \supset q$".

Then we have also

(18) A' is true;

(19) $A' \supset B'$ is true.

Therefore, by (16) and *modus ponens, B*' is true. However, in order to represent this inference as proceeding according to the convention, it is necessary to make another application of *modus ponens*, and so on.

The above argument would not prevent this form of conventionalism from being applied to further parts of mathematics, particularly to existential axioms. In view of the equivalences between derivability statements in logic and elementary propositions in number theory, as well as the above-mentioned element of brute fact in the existence of a derivation, it is not likely that such an approach will work for elementary number theory. But with the stronger axiom systems for set theory the view is on somewhat firmer ground, in that such axioms are often not justified by appeal to direct evidence and "pragmatic" criteria have played a role in the selection of axioms.

Nonetheless, the procedure also has much in common with the setting up of a hypothetical theory in science, and, indeed, as Alfred North Whitehead and Russell already emphasized, the axioms are subject to a sort of checking by their consequences, since some propositions deducible from them are decidable by more elementary and evident mathematical means. It is not evident that if a system of axioms is replaced by another because its consequences come into conflict with intuitive mathematics, the meaning of "set" has changed and the original axioms can be interpreted according to a previous meaning so as to remain true. Moreover, set theory proceeds on the assumption that the truth-value of statements is determinate in many cases where it is not determined by the axioms—that is, by the conventions.

Quine, in fact, now argues, apparently even in the case of elementary logic, that there is no firm ground for distinguishing between making such principles true by convention and adopting them as hypotheses ("Carnap and Logical Truth"). This is as much an extension of conventionalism to the whole of science as a rejection of it in application to mathematics.

2.9.1. *Wittgenstein's view.* At this point we must consider the possibility that a priori truths, even the elementary ones, are thought of as true by convention, not in the sense that they may be made so by an explicit convention actually set up but in the sense that the conventions are, as it were, implicit in our practice with the logical and mathematical vocabulary. It might still be argued that the principles of mathematics are not in that way sufficiently distinguished from the principles of natural science or from other rather deep or fundamental principles that we firmly accept. But this objection could be met by a more detailed descriptive analysis of how logical and mathematical words are used.

However, this type of conventionalism must be careful not to slip into the situation of the more explicit conventionalism of requiring a necessary connection between general intentions and their application in particular statements which is not itself accounted for by the conventions. It appears that the only philosopher who has really faced these challenges has been Ludwig Wittgenstein, in his later period. In connection with Wittgenstein it would probably be better to speak of "agreement" than convention, since the reference to explicit conventions or to "decisions" seems metaphorical, as a picture which is contrasted with that against which he is arguing rather than as a fundamental theoretical concept. It is agreement in our actions—e.g., what we say follows from what—that is essential. We should also be cautious in attributing to Wittgenstein any explanatory theory of logical and mathematical knowledge, in view of his disclaimers of presenting a theory.

Even with these qualifications Wittgenstein's view seems highly paradoxical, for in order to avoid the above-mentioned pitfall the analysis in terms of agreement must extend even to the connection between general rules and their instances. This seems to be the point of the famous discussion of following a rule in Wittgenstein's *Philosophical Investigations*. What ultimately determines what is intended in the statement of a rule are facts of the type of what is actually accepted in the course of time as falling under it.

Wittgenstein (I, 185) gave the example of instructing someone in writing down the terms of the sequence of natural numbers 0, 2, 4, · · ·, $2n$, · · ·. At the start the instructor does not actively think that when the time comes the pupil is to write 1,000, 1,002, 1,004, · · ·, rather than 1,000, 1,004, 1,008, · · ·. Wittgenstein regarded it as conceivable that the pupil might do the second on the basis of a misunderstanding which we just could not clear up. Moreover, it is, as it were, just a fact of natural history that normally, in such a case, we accept the first and reject the second—indeed, continue in that way ourselves. It appears, further, that the same issue can arise for steps in the sequence which have been written before, since the recognition of symbols as tokens of an already understood type is itself an application of a rule (see I, 214).

Wittgenstein's criticism seems directed particularly against certain psychological ideas associated with platonism and Kantianism. The manner in which the steps of writing numerals are determined by the rule cannot be explained by appealing to one's understanding of the relations of abstract entities expressed in the rule or even to the intentions of the instructor. According to Wittgenstein the criterion of how the pupil does understand the rule lies in the steps which he in fact takes. And what makes them right or wrong is their agreement or disagreement with what we do.

The steps are indeed determined by the rule, in the sense that at each stage there is only one number we accept as correct, and the force of social custom directs us to expand the series in the way we do. But this does not mean that Wittgenstein considered his appeals to custom and training as constituting a fully satisfactory explanation of either the agreement that exists or the fact that we feel "compelled" by the rule, for it is because we are made as we are that we react to custom and training as we do.

The paradoxical nature of Wittgenstein's position can perhaps be brought out by considering the case of a complex mathematical proof which contains steps which no one has thought of before. The proof may lead to a quite unexpected conclusion. Yet each step is recognized by every trained person as necessary, and their combination to form the proof is entirely convincing. (This is, of course, not inevitably the case: proofs as published can be obscure or doubtful and can rest on principles about which there are difficulties.) In spite of the fact that it is in principle possible for an irresolvable disagreement to arise at each point, this does not happen: Irresolvable disputes among mathematicians are only about fundamental principles and about taste. Nonetheless, Wittgenstein, in *Remarks on the Foundations of Mathematics*, used the metaphor of decision in speaking of our acceptance of the proof and spoke of the proof as providing a new criterion for certain concepts; his terminology suggests change of meaning.

The vast extent of the agreement on which mathematics rests seems to have astonished Wittgenstein; indeed, it is hard to understand, on his view, how such agreement is possible and why contradictions arise so seldom. We may be faced here with natural facts, but they are facts which show an extremely regular pattern.

Wittgenstein devoted a good deal of attention in the *Remarks* to discussions of calculation and proof, their relation to mathematical truth, and the ways in which they resemble and differ from experiment. In a number of examples he revealed an outlook which resembles Kant's in seeing a construction either of figures or of arrangements of formulae or propositions as essential to a proof. To the problem concerning how such a singular construction can serve to establish a universal and necessary proposition Wittgenstein suggested a quite different answer: In accepting the proof we accept the construction as a paradigm for the application of a new concept, so that, in particular, we have new criteria for certain types of judgments. (For example, if we have determined by calculation that $25 \times 25 = 625$, then a verification that there are 25×25 objects of a certain kind is also accepted as verifying that there are 625.) The same question arises in connection with the possibility of conflict in these criteria as arose in connection with agreement.

We shall close at this point our discussion of the a priori character of mathematics and the attempts to justify and explain it. In the sense that the concepts of mathematics are too general and abstract to refer to anything particular in experience, their a priori character is evident, at any rate after a certain amount of logical analysis of mathematical concepts. The a priori evidence of mathematics, on the other hand, is perhaps not raised, by our discussion, above the level of a somewhat vague conviction. In the case of the more powerful forms of set theory one is probably forced to admit that the evidence is less than certainty and therefore to admit that there is an analogy between the principles involved and the hypotheses of a scientific theory. In the case of arithmetic and elementary logic, however, this conviction can withstand the objections that might be posed, but in view of the difficulties we have discussed in relation to various accounts, it seems still not to have been analyzed adequately.

§3. PLATONISM AND CONSTRUCTIVISM

The discussion in the preceding section suggests that the problem of evidence in mathematics will appear to differ according to the part of mathematics being emphasized. The form which discussion of these differences has tended to take is a distinction between two broad methodological attitudes in mathematics, which we shall call platonism and constructivism. This section will be devoted to a discussion of these attitudes.

3.1. PLATONISM. We begin with platonism because it is the dominant attitude in the practice of modern mathematicians, although upon reflection they often disguise this attitude by taking a formalist position. Platonism is the methodological position that goes with philosophical realism regarding the objects mathematics deals with.

Mathematical objects are treated not only as if their existence is independent of cognitive operations, which is perhaps evident, but also as if the facts concerning them did not involve a relation to the mind or depend in any way on the possibilities of verification, concrete or "in principle."

This is taken to mean that certain totalities of mathematical objects are well defined, in the sense that propositions defined by quantification over them have definite truth-values. Thus, there is a direct connection between platonism and the law of excluded middle, which gives rise to some of platonism's differences with constructivism.

It is clear that there is a connection between platonism and set theory. Various degrees of platonism can be described according to what totalities they admit and whether they treat these totalities as themselves mathematical objects. These degrees can be expressed by the acceptance of set-theoretic existence axioms of differing degrees of strength.

The most elementary kind of platonism is that which accepts the totality of natural numbers—i.e., that which applies the law of excluded middle to propositions involving quantification over all natural numbers. Quite elementary propositions in analysis already depend on this law, such as that every sequence of rational numbers either tends to the limit 0 or does not, which is the basis for the assertion that any real number is either equal to 0 or not. We shall see that not even this assertion is immune to constructivist criticism.

What is nowadays called classical analysis advances a step further and accepts the totality of the points of the continuum or, equivalently, the totality of subsets of the natural numbers. The equivalence between these totalities and their importance in mathematics were brought out by the rigorous development and "arithmetization" of analysis in the nineteenth century. We recall that the theories of (positive and negative) integers and rational numbers can be developed from the theory of natural numbers by means of the notion of ordered pair alone and that this notion can in turn be represented in number theory. A general theory of real numbers requires general conceptions of a set or sequence of natural numbers to which those of a set or sequence of rational numbers can be reduced.

Following Paul Bernays ("Sur le platonisme dans les mathématiques") we can regard the totality of sets of natural numbers on the analogy of the totality of subsets of a finite set. Given, say, the numbers $1, \cdots, n$, each set is fixed by n independent determinations of whether a given number belongs to it or not, and there are 2^n possible ways of determining this. An "arbitrary" subset of the natural numbers is fixed by an infinity of independent determinations fixing for each natural number whether it belongs to the subset or not. Needless to say, this procedure cannot be carried out by a finite intelligence. It envisages the possibility of sets which are not the extensions of any predicates expressed in a language.

3.1.1. ***Impredicative definitions.*** The strength of the assumption of the totality of arbitrary subsets of the natural numbers becomes clear if we observe that it justifies impredicative definitions, definitions of sets or functions in terms of totalities to which they themselves belong. A predicate of natural numbers involving quantification over all sets of natural numbers will have a well-defined extension, which will be one of the sets in the range of the quantifier.

Such definitions have been criticized as circular (for example, by Henri Poincaré), but they do not seem so if we understand the sets as existing independently of any procedure or linguistic configuration which defines them, for then the definition picks out an object from a preexisting totality. The resistance that impredicative definitions met with arose partly because their acceptance clashes with the expectation that every set should be the extension of a predicate, or at least of a concept of the human mind.

Given any definite (formalized) notation, we can by Cantor's diagonal method define a set of natural numbers which is not the extension of a predicate in the notation. Thus, no procedure of generating such predicates by continually expanding one's notation can possibly exhaust the totality. And the idea that every set is the extension of a predicate has little sense if it is assumed that in advance of the specification of notations there is a totality of possible predicates which can be arrived at by some generating procedure.

If the statements of classical analysis are interpreted naively, then quite elementary theorems, such as that every bounded set of real numbers has a least upper bound, require impredicative definitions. Nonetheless, in *Das Kontinuum*, Hermann Weyl proposed to construct analysis on the basis of mere platonism with respect to the natural numbers. He proposed an interpretation under which the least upper bound theorem is true. Later interpretations have preserved more of the statements of classical analysis than Weyl's, and it is an involved technical question how much of it can be given a natural predicative interpretation (see below).

3.1.2. ***Set theory and the paradoxes.*** Set theory as developed by Cantor and as embodied in the present standard systems involves a higher degree, or variety of degrees, of platonism. The axiom system of Zermelo and its enlargement by Fraenkel (which is called the Zermelo-Fraenkel system), for example, allows the iteration of the process of forming the set of all subsets of a given set and the collection into a set of what has been obtained by iterated application of this or some other generating procedure. This latter allows the iteration into the transfinite. If we assume we have transfinite ordinal numbers, then we can generate a transfinite succession of "universes" U as follows: Let $\mathcal{P}(A)$ be the set of all subsets of the set A.

U_0 = a certain class, perhaps empty, of "individuals."

$$U_\alpha + 1 = \mathcal{P}(U_\alpha) \cup U_\alpha.$$

U_α = the union of all U_β, for $\beta < \alpha$, if α is a limit ordinal.

Then for certain ordinals α the U_α will form models for the different systems of set theory ($U_\omega + \omega$ for Zermelo's set theory, without Fraenkel's axiom of replacement).

The paradoxes of set theory imply that we must accept some limitations on forming totalities and on regarding them in turn as mathematical objects—that is, as sets. If, for example, the totality of sets is a well-defined set, then it seems that it will be reasonable to ask of each set x whether it is a member of itself ($x \in x$) or not and to form $\hat{x}(x \notin x)$, the set of all sets which are not members of themselves. This will satisfy

$$(y)[y \in \hat{x}(x \notin x) \,.\!\equiv y \notin y],$$

which implies

$$\hat{x}(x \notin x) \in \hat{x}(x \notin x) \,.\equiv.\, \hat{x}(x \notin x) \notin \hat{x}(x \notin x).$$

a contradiction. This is Russell's paradox, the most shocking, because the most elementary, of the paradoxes of set theory.

On the same basis one can ask for the cardinal number of the set of all sets, which we shall call S. Then $\mathcal{P}(S)$, the set of all subsets of S, will have a cardinal number no greater than that of S, because $\mathcal{P}(S) \subseteq S$. But by Cantor's theorem the cardinal number of $\mathcal{P}(S)$ is properly greater than that of S (Cantor's paradox, 1895).

If the totality O of ordinals is a set, then, since it is well-ordered, there will be an ordinal number γ that represents its order type. But then O will be isomorphic to the set of ordinals less than γ—that is, to a proper initial segment of itself. This is impossible: γ must be the greatest ordinal, but there is no obstacle to forming $\gamma + 1$ (Burali-Forti's paradox, 1897).

These paradoxes do not imply that we have to stop or otherwise limit the process, described above, of generating larger and larger universes. On the contrary, we must never regard the process as having given us "all" sets. The totality of sets, and hence the totality of ordinal numbers, cannot be the terminus of a well-defined generating process, for if it were we could take all of what we had generated so far as a set and continue to generate still larger universes.

Thus, suppose we consider the arguments for the paradoxes applied to a particular U_α, as if it were the universe of all sets. The construction precludes $x \in x$, so $\hat{x}(x \notin x)$ is just U_α itself. But $U_\alpha \notin U_\alpha$ and hence is disqualified as a set. The same consideration applies to Cantor's paradox. Burali-Forti's paradox is avoided because the passage from U_α to $U_{\alpha+1}$ always introduces well-orderings of higher order types. Thus, for no α can U_α contain "all" ordinals, no matter how the ordinals are construed as sets. (A very natural way of construing them would be such that α occurs in $U_{\alpha+1}$ but not in U_β for any $\beta \leq \alpha$. But then only for certain ordinals will U_α contain an ordinal for each well-ordered set in U_α.)

For some time after they were first discovered, the paradoxes were viewed with great alarm by many who were concerned with the foundations of mathematics. In retrospect this seems to have been because set theory was still quite unfamiliar; in particular, the distinction between the customary reasonings of set theory and those that led to the paradoxes was not very clear. The opposition that set theory had aroused had not yet died down. However, the marginal character of the paradoxes has seemed more and more evident with time; the systems which were soon devised to cope with the paradoxes (Russell's theory of types and Zermelo's set theory, both published in 1908) have proved satisfactory in that they are based on a reasonably clear intuitive idea, and no one today regards it as a serious possibility that they (or the stronger Zermelo-Fraenkel system) will turn out to be inconsistent. This does not mean that the security and clarity of set theory are absolute; in the sequel some of the difficulties will become apparent.

The above-described sequence of universes uses general conceptions of set and ordinal but applies the characteristic move of platonism only one step at a time. It renounces what Bernays calls "absolute platonism," the assumption of a totality of all mathematical objects which can be treated as itself a customary mathematical object—for example, a set. Such a conception seems def-

initely destroyed by the paradoxes. The totality of sets can be compared with Kant's "Ideas of Reason": it is an "unconditioned" or absolute totality which just for that reason cannot be adequately conceived by the human mind, since the object of a normal conception can always be incorporated in a more inclusive totality. From this point of view there is an analogy between the set-theoretic paradoxes and Kant's mathematical antinomies.

If we assume that every set will appear in one of the U_α, we have a conception which is adequate for all of modern mathematics except, perhaps, the recent theory of categories. The conception is by nature imprecise: there are limitations on our ability to circumscribe both what goes into the power set of a given set and what ordinals there are. It is perhaps unreasonable to apply classical logic to propositions involving quantification over all sets, since such an application seems to presuppose that it is objectively determined what sets (and a fortiori, on this conception, what ordinals) there are. Nonetheless, this additional idealization does not seem to have caused any actual difficulties.

This way of conceiving sets combines two of Russell's early ideas for resolving the paradoxes—the theory of types and the theory of "limitation of size." What are rejected as sets are the most inclusive totalities, such as the entire universe. (Our talking of "totalities" while rejecting them as sets is not incompatible with our conception; as John von Neumann observed, all that is necessary is to prohibit them from belonging to further classes. Von Neumann's observation was the basis for some new set theories, the principal one being that of Bernays and Gödel.) Moreover, the sets are arranged in a transfinite hierarchy: One can assign to each set an ordinal, its type or, as it is now called, rank, which will be the least ordinal greater than the ranks of its members. We have thus a transfinite extension of the cumulative theory of types. But we have dropped the more radical idea from which Russell proceeded: that each variable of a system of set theory should range over objects of a specified type, and that "$x \in y$" is meaningless unless the range of "y" is of a type one higher than that of "x," so that, in particular, "$x \in x$" is meaningless.

3.1.3. ***Predicativism.*** In the first twenty-five years or so after the discovery of the paradoxes a number of more radical proposals for their elimination were presented. These generally amounted to some further attenuation of platonism. We shall first consider the program of eliminating impredicative definitions, which amounts to a restriction of platonism to the natural numbers. This was the outcome of the general views of Poincaré and Russell. Russell's original theory, the ramified theory of types, which formed the basis of *Principia Mathematica*, was directed to the elimination of impredicative definitions, which he held to involve a "vicious circle" and to be responsible for the paradoxes. The effect was, however, nullified by his axiom of reducibility.

A greatly simplified version of the ramified theory is as follows: One has variables, each of which is assigned a natural number as its level, and the predicates of identity and membership. The logic is the usual quantification theory, except that in the rules for quantifiers allowance must be made for levels. Since the levels can be cumulative, we could have for the universal quantifiers the following:

(20) $(x^i)Fx^i \supset Fy^j$ if $j \leq i$;

(21) From "$p \supset Fy^i$" infer "$p \supset (x^i)Fx^i$," where for "p" only something not containing free "y^i" can be substituted.

The axioms are those of identity, extensionality, and the following schema of class existence:

(22) If "F" represents a predicate which does not contain free x^{i+1}, any free variables of level $> i + 1$, or any bound variables of level $> i$,

$$(\exists x^{i+1})(y^i)(y^i \in x^{i+1} \equiv Fy^i).$$

One effect of this axiom is that a predicate involving quantification over objects of level n need not have an extension of level n. Therefore, the axiom does not assert the existence of any impredicative classes; in fact, it is compatible with the idea that classes are constructed by the construction of predicates of which they are the extensions.

Russell's actual theory combined that of a hierarchy of levels, applied in this case to "propositional functions," the objects over which the variables of a higher-order logic were to range, with the "no class" theory, the introduction of locutions involving classes by contextual definition in terms of propositional functions. In order to derive classical mathematics, however, he wanted to avoid dividing the classes into levels. This he did by postulating the axiom of reducibility, which asserts that for every propositional function there is a function of the lowest possible level (compatible with the nature of its arguments) extensionally equivalent to it. Russell admitted that this axiom was equivalent to the existence of classes, and he has never been satisfied with it. In effect, it yields even impredicatively defined classes and destroys the effect of the hierarchy of levels.

A formalization of mathematics on the basis of the ramified theory is the most natural formalization if a platonist theory of classes is repudiated but classical logic admitted. The construction of the natural numbers leads to the difficulty that the class quantifier needed to reduce induction to an explicit definition is no longer available. One must either assume the natural numbers or have a hierarchy of different concepts of natural number.

A ramified theory with the natural numbers as individuals and the Peano axioms would be a natural formalization of the mathematics allowed by platonism with respect to the natural numbers. But there is in principle no reason not to extend the hierarchy of levels into the transfinite. The question of the limits of predicative mathematics has become identical with the question of the transfinite ordinals that can be predicatively introduced.

We have said that quite elementary proofs in analysis already require impredicative definitions when naively interpreted. Nonetheless, from recent work it appears that a good deal of classical analysis is susceptible of a natural predicative interpretation, which, however, fails for some theorems. One can, on this basis, give a good approximation to classical analysis, but not to the whole of it. That part of mathematics which depends essentially on still more powerful set theory is completely lost. It seems that it would not be reasonable to insist on this limitation unless there were some quite powerful reason for rejecting platonism. We shall discuss some possible reasons later.

3.2. **CONSTRUCTIVISM**. We shall now consider the complete rejection of platonism, which we shall call constructivism. It is not a product of the situation created by the paradoxes but rather a spirit which has been present in practically the whole history of mathematics. The philosophical ideas on which it is based go back at least to Aristotle's analysis of the notion of infinity (*Physics*, Bk. III). Kant's philosophy of mathematics can be interpreted in a constructivist manner, and constructivist ideas were presented in the nineteenth century—notably by Leopold Kronecker, who was an important forerunner of intuitionism—in opposition to the tendency in mathematics toward set-theoretic ideas, long before the paradoxes of set theory were discovered.

Our presentation of constructivism relies heavily on the "intuitionism" of Brouwer, presented in many publications from 1907 on, but the ideas can also be found to some extent in other critics of platonism, including the French school of Émile Borel, Poincaré, and Henri Lebesgue, although in their work predicativity played a greater role than constructivity. These writers did not arrive at a very consistent position, but they contributed mathematically important ideas. L. E. J. Brouwer reached and developed a conclusion from which they shrank: that a thoroughgoing constructivism would require the modification of classical analysis and even of classical logic.

3.2.1. *Intuitionism.* Constructivist mathematics would proceed as if the last arbiter of mathematical existence and mathematical truth were the possibilities of construction. "Possibilities of construction" must refer to the idealized possibility of construction mentioned in the last section. Brouwer insisted that mathematical constructions are mental. The possibilities in question derive from our perception of external objects, which is both mental and physical. However, the passage from actuality to possibility and the view of possibility as of much wider scope perhaps have their basis in intentions of the mind—first, in the abstraction from concrete qualities and existence; second, in the abstraction from the limitations on generating sequences. In any case, in constructive mathematics the rules by which infinite sequences are generated are not merely a tool in our knowledge but part of the reality that mathematics is about.

Why this is so can be seen from the problem of assertions about the infinite. We have suggested that the generation of a sequence of symbols is something of which the construction of the natural numbers is an idealization. But "construction" loses its sense if we abstract further from the fact that this is a process in time which is never completed. The infinite in constructivism must be "potential" rather than "actual." Each individual natural number can be constructed, but there is no construction which contains within itself the whole series of natural numbers. To view the series *sub specie aeternitatis* as nonetheless determined as a whole is just what we are not permitted to do.

Perhaps the idea that arithmetic rests on time as a form of intuition lies behind Brouwer's insistence on constructivity interpreted in this way. One aspect of sensibility from which we do not abstract in passing from concrete perception to its form is its finite character. Thus, whatever one may think of the notion of form of intuition, Brouwer's position is based on a limitation, in principle, on our knowledge: Constructivism is implied by the postulate that no mathematical proposition is true unless we can in a nonmiraculous way know it to be true.

Because of its derivation from his own philosophical account of mathematical intuition Brouwer called his position, and the mathematics which he constructed on

the basis of it, intuitionism. We shall use this name for a species of constructivism which answers closely to Brouwer's ideas.

In spite of the "potential" character of the infinite in mathematics, we shall not renounce assertions about all natural numbers or even, with some reservations, talk of infinite classes. A proposition about all natural numbers can be true only if it is determined to be true by the law according to which the sequence of natural numbers is generated. This Brouwer took to be equivalent to its possessing a proof. Thus, the intensional notions of "law" and "proof" become part of the subject matter of mathematics.

A consideration of existential propositions connects the broad philosophical notion of constructivity with the general mathematical notion. Roughly, a proof in mathematics is said to be constructive if wherever it involves the mention of the existence of something, it provides a method of "finding" or "constructing" that object. It is evident that the constructivist standpoint implies that a mathematical object exists only if it can be constructed; to say that there exists a natural number x such that Fx is to say that sooner or later in the generation of the sequence an x will turn up such that Fx. If x depends on a parameter y, this x must be determinable from y on the basis of the laws of the construction of the numbers and of the constructions involved in F. Proving $(\exists x)Fx$ means showing how to construct x, so one can say that the proof is not complete until x has been exhibited. (But then "proof" is used in an idealized sense.) To prove $(y)(\exists x)Fxy$ must involve giving a general method for finding x on the basis of y.

This point of view leads immediately to a criticism of the basic notions of logic, particularly negation and the law of excluded middle. That "$(x)Fx$" is true if and only if it can be proved does not mean that "$(x)Fx$" is a statement about certain entities called proofs in the way in which, on the usual interpretation, it is a statement about the totality of natural numbers. According to Brouwer we can assert "p" only if we have a proof; the hypothesis that $(x)Fx$ is the hypothesis that we have a proof, and it is a reasonable extrapolation to deny that we can say more about what "$(x)Fx$" asserts than is said in specifying what is a proof of it. The explanation of "$\neg(x)Fx$" as "$(x)Fx$ cannot be proved" does not satisfy this condition. Brouwer said instead that a proof of "$\neg p$" is a construction which obtains an absurdity from the supposition of a proof of "p."

An immediate consequence of this interpretation is that the law of excluded middle becomes doubtful. Given a proposition "p," there is no particular reason to suppose that we shall ever be in possession either of a proof of "p" or of a deduction of an absurdity from "p." Indeed, if the general statement of the law of excluded middle is taken as a mathematical assertion, a proof of it will have to yield a general method for the solution of all mathematical questions. Brouwer rejected this possibility out of hand.

It is evident that such a point of view will lead to changes in quite basic parts of mathematics. Many instances of the law of excluded middle, where the propositions involved can be shown constructively to be systematically decidable, will be retained. But Brouwer rejected even very elementary instances in classical analysis. Let the sequence r_n of rational numbers be defined as follows: if there is no $m \leq n$ such that the mth, $(m + 1)$st, $(m + 2)$d terms of the decimal expansion of π are each 7, then $r_n = 1/2^n$; if there is such an m, then $r_n = 1/2^k$, where k is the least such m. Then r_n constructively defines a real number r. But a proof of either $r = 0$ or $r \neq 0$ would tell us whether or not there are three 7's in the decimal expansion of π. Thus, we cannot assert either $r = 0$ or $r \neq 0$.

For a satisfactory constructivist theory of analysis, an analysis is needed of the notion of an arbitrary set or sequence of natural numbers. Brouwer's analysis gives additional distinctiveness to intuitionism. Such a sequence is thought of as generated by a succession of independent determinations or "free choices," which may be restricted by some law. Obviously the succession of choices must be thought of as never being complete. In the absence of a law a statement about a sequence can be true only if it is determined to be true by some finite initial segment of the sequence. The consequence of this is that a function defined for all sequences of natural numbers whose values are integers must be continuous. It also leads to sharper counterexamples to the law of excluded middle: It is absurd that for all sequences α, either $(x)(\alpha(x) = 0)$ or $\neg(x)(\alpha(x) = 0)$. We can also sharpen the result of the preceding paragraph and state generally that not every real number is equal to or different from 0.

The intuitionist point of view thus leads to a distinctive logic and to a distinctive theory of the foundations of analysis. The latter contains another distinctive principle, the bar theorem, obtained by analyzing the requirement that if a function is defined for all sequences, there must be a constructive proof of this fact. It is roughly equivalent to the proposition that if an ordering is well-founded, transfinite induction holds with respect to it. Nonetheless, intuitionism is far from having shown itself capable of the same rich development as classical mathe-

matics, and it is often very cumbersome. Important as it is in itself, it does not provide a sufficient motive for renouncing platonism.

3.2.2. *Finitism.* So far our account of constructivism has been based entirely on Brouwer's intuitionism. However, intuitionism is not the only possible constructivist development of mathematics. Indeed, it makes some quite powerful assumptions of its own. As we have said, the intuitionists make the notions of construction and proof a part of the subject matter of mathematics, and the iteration of logical connectives, especially, renders it possible to make quite elaborate and abstract statements involving construction and proof. Thus, intuitionist mathematics seems to rest not merely upon intuition but upon rather elaborate reflection on the notion of intuitive construction. (It also does not obviously exclude impredicativity, since what counts as a proof of a given proposition can be explained in terms of the general notion of proof.) A constructivist might feel that intuitionism leads from the Scylla of platonist realism to the Charybdis of speculative idealism.

A weaker and more evident constructive mathematics can be constructed on the basis of a distinction between effective operation with forms of spatiotemporal objects and operation with general intensional notions, such as that of proof. Methods based on operation with forms of spatiotemporal objects would approximate to what the mathematician might call elementary combinatorial methods or to the "finitary method" which Hilbert envisaged for proofs of consistency. Formal systems of recursive number theory, in which generality is expressed by free variables and existence by the actual presentation of an instance or (if the object depends on parameters) a function, will accord with this conception if the functions admitted are sufficiently elementary—for example, primitive recursive functions. In such formalisms any formula will express a general statement each instance of which can be checked by computation. For this reason classical logic can be used. Moreover, the concept of free choice sequence can be admitted so that some analysis can be constructed.

The precise limits of this conception are perhaps not clear, although it is evident that some constructive arguments are excluded. The conception does not allow full use of quantifiers but probably does allow a limited use of them.

3.2.3. *The Hilbert program.* If one accepts the idea that from a philosophical point of view constructivist conceptions are more satisfactory than platonist conceptions—more evident or more intelligible—one is not necessarily constrained to abandon classical mathematics. The way is still open to investigating classical mathematics from a constructive point of view, and it may then prove to have an indirect constructive sense and justification.

Such an investigation was the objective of the famous program of Hilbert, which was the third main animating force—with logicism and intuitionism—in foundational research in the period before World War II. The possibility arises first from the fact that classical mathematics can be formalized (though not completely; we shall consider this fact and its implications later). Once it has been formalized, one can in principle drop consideration of the intended meaning of the classical statements and simply consider the combinations of the symbols and formulae themselves. Thus, if the proof of a certain theorem has been formalized in a system *S* (say Zermelo-Fraenkel set theory), it is represented as a configuration of symbols constructed according to certain rules. Whether a configuration is a proof can be checked in a very elementary way.

The concepts by which a formal system is described belong, in effect, to finitist mathematics. For example, the consistency of the system is the proposition that no configuration which is a proof will have a last line of a certain form—for example, $\mathcal{A}$ & $\neg\mathcal{A}$. Nonetheless, although in the mathematical study we abstract from the intended interpretation, this interpretation certainly guides the choice of the questions in which we are interested.

Hilbert sought to establish classical platonist mathematics on a firm foundation by formalizing it and proving the consistency of the resulting formalism by finitist means. The interest of the question of consistency depends on the fact that the formulae of the system represent a system of statements; that is, even if the meanings of the platonist conceptions are highly indeterminate, statements in terms of them are introduced according to an analogy with "real" (i.e., finitist) statements which is intended to preserve at least the notions of truth and falsity and the laws of logic.

In fact, Hilbert had a further motive for his interest in consistency: the fact that platonist mathematics is an extension of an extrapolation from finitist mathematics. Certain elementary combinatorial notions are also embodied in the formalism; formulae involving them express "real statements." Hilbert thought of the other formulae as expressing "ideal statements"—analogous to the ideal elements of projective geometry—introduced to give greater simplicity and integration to the theory. Within the system they have deductive relations to the

real statements. It would be highly undesirable that a formula of the system should be seen by elementary computation to be false and yet be provable. One might hope to prove by metamathematical means that this would not happen. In the central cases a proof of consistency is sufficient to show that it would not. Thus, suppose we extend a quantifier-free recursive number theory by adding quantifiers and perhaps also second-order quantifiers. A proof of the consistency of the resulting system will show that no false numerical formula (stating a recursive relation of particular integers) will be provable. In fact, it will yield a constructive proof of any formula of the original system provable in the extension, in this sense showing the use of "ideal" elements to be eliminable. Since Hilbert it has been pointed out (chiefly by Georg Kreisel) that many further results relevant to the understanding of nonconstructive mathematics from a constructivist point of view can be obtained from consistency proofs.

Hilbert hoped to settle the question of foundations once and for all, which for him meant establishing the platonist methods of set theory on a firm basis. His hope was founded on two expectations: that all of mathematics (at least all of analysis) could be codified in a single formal system and that the consistency of this system could be proved by methods so elementary that no one could question them. He was disappointed of both these expectations as a result of Gödel's incompleteness theorems (1931). Work on the program has nonetheless continued, with the limitations that one has to work with formalisms which embody only part of the mathematics in question and that the proofs must rely on more abstract, but still constructive, notions; and the work in finitist proof theory has achieved valuable results, some of which will be discussed later.

§4. MATHEMATICAL LOGIC

Our remaining considerations on the subjects of the two preceding sections fit best into an independent discussion of mathematical logic as a factor in the study of the foundations of mathematics. Before World War II an important part of the work in logic was directed toward establishing, in the service of some general position such as logicism or intuitionism, a more or less final solution to the problems of foundations. Certain particular results, and probably also a more diffuse evolution of the climate of ideas, have discouraged this aim. Today nearly all work in mathematical logic, even when motivated by philosophical ideas, is nonideological, and everyone acknowledges that the results of this work are independent of the most general philosophical positions.

Starting from the axiomatic method in a more general sense, mathematical logic has become the general study of the logical structure of axiomatic theories. The topics selected from the great variety of technical developments for discussion here are Gödel's incompleteness theorems, recursive function theory, developments related to Hilbert's program, foundations of pure logic, and axiomatic set theory.

4.1. GÖDEL'S INCOMPLETENESS THEOREMS. Research in mathematical logic took quite new directions as a result of the discovery by Kurt Gödel, in 1930, of his incompleteness theorems. According to the first theorem (as strengthened by J. B. Rosser in 1936) any formalism S that is sufficiently powerful to express certain basic parts of elementary number theory is incomplete in the following sense: A formula $\mathcal{A}$ of S can be found such that if S is consistent, then neither $\mathcal{A}$ nor $\neg\mathcal{A}$ is provable in S. The conditions are satisfied by very weak systems, such as the first-order theory Q whose axioms are the Peano axioms for the successor function and the recursion equations for addition and multiplication. (This system is formalized in first-order logic with equality, having successor, addition, and multiplication as primitive function symbols. The axioms are versions of our axioms (1)–(4), recursion equations for addition and multiplication, and an axiom which says that every number not equal to 0 is the successor of something.) They are satisfied by extensions of systems that satisfy them and therefore by the full elementary number theory Z (the first-order version of the Dedekind-Peano axiomatization, obtained from Q by adding induction: in place of the second-order axiom (5) one adds all results of substituting a predicate of the formalism for "F" in (7), by analysis, and by axiomatic set theories in which number theory can be constructed. They are also satisfied by formalizations of intuitionist theories. Evidently adding further axioms offers no escape from this incompleteness, since the new theories will also satisfy the conditions of the theorem.

One of the conditions necessary for some general statements of the theorem is that which we mentioned earlier, that proofs can be checked mechanically. This must be interpreted more precisely in terms of one of the concepts of recursive function, discussed below.

The technique of Gödel's proof is of great interest and has since found wide application. It consists of a mapping of the syntax of the theory into the theory itself, through assigning numbers to the symbols and formulae

of the system. Any syntactical relation will then be equivalent to some relation of natural numbers. For the crucial relation "$\mathscr{X}$ is a proof in S of the formula $\mathscr{A}$" the corresponding relation $P(x,a)$ can be expressed in the theory, and certain things about it can be proved in S. Then the undecidable formula $\mathscr{A}$ is a formula which has a number k such that what $\mathscr{A}$ says (about numbers) is equivalent to the unprovability of the formula number k, i.e., $\mathscr{A}$. (1) Then if only true formulae are provable, $\mathscr{A}$ is unprovable. But then $\mathscr{A}$ is true. Therefore, (2) by the same assumption $\neg\mathscr{A}$ is also unprovable. This appeal to the notion of truth was replaced in Gödel's detailed argument by the condition that S be consistent for (1) and ω-consistent for (2). By changing the formula Rosser showed that the assumption of ω-consistency could also be replaced by that of consistency.

The proof that if S is consistent, then $\mathscr{A}$ is unprovable is finitist. If S and the mapping of its syntax into S satisfy some further conditions, the argument can be formalized in S. This yields the second theorem of Gödel. If S is consistent, then the formula which, under the above mapping, corresponds to the consistency of S is unprovable in S.

The first theorem implies not only that mathematics as a whole cannot be codified in a single formal system but also that the part of mathematics that can be expressed in a specific formal notation cannot be so codified. This fact undermines most attempts at a final solution to the problem of foundations by means of mathematical logic. The second theorem was a blow to the Hilbert program in particular. The methods that the Hilbert school envisaged as finitary could apparently be codified in first-order number theory Z; indeed, that they can be so codified seems fairly certain, even though the notion of finitary methods is not completely precise. Therefore, not even the consistency of Z is provable by finitary means. Moreover, the consistency of stronger and stronger systems requires stronger and stronger methods of proof.

There has been much discussion of the broader philosophical implications of Gödel's theorem. We shall not enter into the discussion of such questions as whether the theorem shows the falsity of any mechanistic theory of mind. It should be remarked that there are a number of connections between the surpassing of any given formal system by possible means of proof and the inexhaustibility phenomena in the realm of mathematical existence. Gödel's argument can be viewed as a diagonal argument parallel to that by which Cantor proved that no countable set of sets of natural numbers can exhaust all such sets. Peano's axioms are categorical if the range of the quantifiers in the induction axiom (5) includes all classes of natural numbers, but in the context of a formal system one can use only the fact that induction holds for classes definable in the system, of which there are only countably many. In set theory the addition of axioms asserting the existence of very large classes can make decidable previously undecidable arithmetical formulae.

4.2. RECURSIVE FUNCTION THEORY. A number of problems in mathematical logic require a mathematically exact formulation of the notion of mechanical or effective procedure. For most purposes this need is met by a concept of which there are various equivalent formulations, arrived at by several writers. The concept of (general) recursive definition, introduced in 1931 by Jacques Herbrand and Gödel, was the first. A function of natural numbers which is computable according to this conception (the "computation" consists of the deduction of an evaluation from defining equations by simple rules) is called a general recursive, or simply a recursive, function. Other formulations are that of λ-definability (Alonzo Church), computability by Turing machine (A. M. Turing), algorithms (A. A. Markov), and different notions of combinatorial system (Emil Post and others).

The concept of recursive definition has proved essential in decision problems. Given a class of mathematical problems defined by some parameter, is there an effective algorithm for solving each problem in the class? As an example consider the tenth problem of Hilbert: Given a polynomial with integral coefficients, is there a general method that tells us whether it has a zero among the integers? If such a question can be resolved in the affirmative, the resolution can generally be reached on the basis of the intuitive conception of an algorithm: If one can invent the procedure, then it is generally clear that the procedure is effective. But to give a negative answer to such a question one needs some idea of the possible effective procedures. The development of recursive function theory has made possible a large number of results asserting the nonexistence of decision procedures for certain classes of problems. This way of interpreting the results depends on a principle known as Church's thesis, which says that the mathematical conception of an effectively computable function in fact corresponds to the intuitive idea—i.e., that a number-theoretic function is (intuitively) effectively computable if and only if it is recursive.

An important type of decision problem is that concerning provability in formal systems. Given a formal system S, is there an algorithm for deciding whether a given

formula $\mathscr{A}$ is a theorem of *S*? If there is, then *S* is said to be decidable. Although quite interesting examples of decidable systems exist, the systems to which Gödel's first incompleteness theorem applies are undecidable. In fact, Gödel's type of argument can also be used to prove that first-order logic is undecidable (as by Church in 1936).

Another important aspect of recursive function theory is the classification of sets and functions according to different principles related to recursiveness. One such principle, stated in terms of the complexity of possible definitions by recursive predicates and quantifiers (the Kleene-Mostowski hierarchy), not only is of wide application in logic but is closely related to older topological classifications. One can single out the arithmetical sets (those sets definable from recursive predicates by quantification over natural numbers alone), the hyperarithmetical sets (a certain transfinite extension of the arithmetical hierarchy—in effect, those sets definable in ramified analysis with levels running through the recursive ordinals), and the analytic sets (those sets definable from recursive predicates by quantification over numbers and functions, or sets, of natural numbers). The recursive ordinals, singled out by Church and Kleene, can most readily be characterized as the order types of recursive well-orderings of the natural numbers.

The theory of recursive functions is evidently valuable for explicating different notions of constructivity and for comparing classical and constructive mathematics. A constructive proof of a statement of the form "$(x)(\exists y)Fxy$" should yield an effective method of obtaining y from x. For example, Kleene and his collaborators have shown that any statement provable in formalized intuitionist number theory and analysis has a property called "realizability," which amounts roughly to interpreting "$(x)(\exists y)Fxy$" as asserting the existence of a recursive function giving y in terms of x. Although it is also intuitionistically meaningful, the construction gives a classical interpretation of the intuitionist formalisms. It also allows a sharpening and extension of Brouwer's counterexample technique. Certain classically provable formulas can be shown not to be realizable and therefore not to be provable in the intuitionist formalisms Kleene considers.

A problem arises with regard to the relation between the concept of recursive function and the fundamental concepts concerning constructivity—for instance, the concept of intuitionism. One cannot interpret Church's thesis as explicitly defining "effectively computable function" and therefore as giving the meaning of the intuitionist quantifiers. For by definition a function is general recursive if there is a set of equations from which for each possible argument one can compute the value of the function for that argument, a statement of the form "$(x)(\exists y)Fxy$." If this is interpreted constructively, the proposed definition is circular. The relation between "function constructively proved to be everywhere defined" and "general recursive function" is still not clear. One can ask whether every intuitionistically everywhere-defined number-theoretic function is general recursive or whether every (classically) general recursive function can be proved constructively to be such. Neither question has yet been resolved.

4.3. DEVELOPMENT OF THE HILBERT PROGRAM. For the study of constructivity it is also important to study more restricted types of recursive definition that can be seen by definite forms of argument to define functions. This is particularly important for the extended Hilbert program.

Gödel's second incompleteness theorem meant that the consistency even of elementary number theory Z could not be proved by the methods envisaged by Hilbert. A number of consistency results of the sort envisaged by Hilbert have since been obtained by stronger constructive methods. Gödel and Gentzen proved independently (and finitistically) that if intuitionistic first-order arithmetic is consistent, then so is classical first-order arithmetic. The proofs were based on a quite simple method of translating classical theories into intuitionist theories which is of wide application—for example, to pure logic. One renders an atomic formula P by $\neg\neg P$ (in elementary number theory, equivalent to P itself). If $\mathscr{A}$, $\mathscr{B}$ are translated into $\mathscr{A}°$, $\mathscr{B}°$, respectively, then $\mathscr{A} \vee \mathscr{B}$ is translated by $\neg\neg(\mathscr{A}° \vee \mathscr{B}°)$, $(\exists x)\mathscr{A}$ by $\neg\neg(\exists x)\mathscr{A}°$, $\mathscr{A} \supset \mathscr{B}$ by $\neg(\mathscr{A}° \,\&\, \neg\mathscr{B}°)$, $\mathscr{A} \,\&\, \mathscr{B}$ by $\mathscr{A}° \,\&\, \mathscr{B}°$, $\neg\mathscr{A}$ by $\neg\, \mathscr{A}°$, and $(x)\mathscr{A}$ by $(x)\mathscr{A}°$. Evidently the translation not only proves relative consistency but also gives each provable formula an intuitionist meaning according to which it is intuitionistically true. If $\mathscr{A}$ is a quantifier-free formula of number theory, or if it is composed with conjunction, negation, and universal quantification only, then if it is provable in Z, it is intuitionistically provable. This translation can easily be extended to ramified analysis. Since intuitionistically the consistency of the intuitionist systems follows from their soundness under the intended interpretation, the consistency of the classical systems has been intuitionistically proved.

A sharper result was obtained in 1936 by Gerhard Gentzen. New proofs, with various advantages and refinements, have since been found by several workers.

Gentzen proved the consistency of *Z* by adding to finitist arithmetic the assumption that a certain recursive ordering of natural numbers, of order type ϵ_0 (the least ordinal greater than ω, ω^ω, ω^{ω^ω}, $\cdots$), is a well-ordering. This assumption could be proved in intuitionist ramified analysis using set variables only of level 1 but could not in elementary number theory.

Gentzen's result has made it possible to extract further information about the power of elementary number theory. Kreisel obtained information about the relation between elementary number theory and certain quantifier-free arithmetics and also obtained a characterization of the functions which can be proved in *Z* to be general recursive.

A corresponding result for ramified analysis for finite levels was obtained by Lorenzen in 1951 and sharpened by Kurt Schütte. It was extended by Schütte to transfinite levels.

On the basis of these results we can say that constructive consistency proofs are available for all of predicative mathematics. In well-defined senses they are the best possible results (for instance, the above-mentioned ordinal ϵ_0 cannot be replaced by a smaller one). Nonetheless, efforts to give such a proof for impredicative classical analysis, not to speak of axiomatic set theory, have proved fruitless.

Results of quite recent research have shed considerable light on this situation. Clifford Spector (1962) proved the consistency of classical analysis relative to a quantifier-free theory (Gödel 1958) of primitive recursive functionals of arbitrary finite types, enriched by a new schema for defining functionals by "bar recursion." This amounted to generalizing Brouwer's bar theorem to arbitrary finite types. Such generalized bar recursion has not found a constructive justification, but the method has led to consistency proofs by the original bar theorem for subsystems of analysis which are, according to a reasonable criterion, impredicative.

Kreisel (1963) has shown that intuitionist analysis, with the bar theorem and a strong schema of "generalized inductive definitions" included, does not suffice to prove the consistency of classical analysis. Such a proof requires an essential extension of constructive methods beyond the established intuitionist ones.

Solomon Feferman and Schütte have given an analysis of the notion of predicativity according to which established intuitionist methods go beyond predicative ones. According to their conception, inductive definitions such as that of the class *O* of numbers representing the recursive ordinals are impredicative.

What has been the fate of the Hilbert program? Put most broadly, its objective was to secure the foundations of platonist mathematics by a constructive analysis of classical formal systems. The incompleteness phenomena have made it impossible, in dealing with stronger and stronger systems, to avoid the introduction of more and more abstract conceptions into the metamathematics. However interesting the information obtained about the relation between these conceptions and the platonist ones, it is not evident that these conceptions are in all respects more secure. Moreover, in the present state of research it is not certain that strong enough constructive methods can be found even to prove the consistency of classical analysis.

This state of affairs is unfavorable to those methodological views seeking to restrict mathematics to the methods which have the greatest intuitive clarity. It is evident that such methods will not suffice to resolve certain mathematical questions whose content is extremely simple, namely those concerning the truth of certain statements of the form "$(x)Fx$," where "*F*" stands for a primitive recursive predicate of natural numbers. Proponents of the views in question seem forced to admit that even such questions can be objectively undetermined.

4.4. FOUNDATIONS OF LOGIC. An important result concerning pure logic obtained in finitist metamathematics is a theorem, or cluster of related theorems—including Herbrand's theorem (1931) and Gentzen's theorem (1934)—to the effect that the proof of a formula of first-order logic can be put into a normal form. In such a normal-form proof the logical complexity of the formulae occurring in the proof is in certain ways limited in relation to the complexity of the conclusion; for instance, no formula can contain more nested quantifiers than the conclusion. The proof is, as it were, without detours, and *modus ponens* is eliminated. As a consequence, a quantifier-free formula deduced from quantifier-free axioms can be proved by propositional logic and substitution, which implies all the consistency results proved by the Hilbert school before the discovery of Gödel's theorem. Gentzen's theorem also applies to intuitionist logic and to other logics, such as modal logics.

These theorems, which are the fundamental theorems of the proof theory of quantification theory, are closely related to the fundamental theorem of its semantics, Gödel's completeness theorem. Every formula not formally refutable has a model—in fact, a model in which

the quantifiers range over natural numbers; i.e., there are denumerably many individuals. This can be strengthened to the following: If *S* is any set (finite or infinite) of formulae of first-order logic, it has a denumerable model unless some finite subset of *S* is inconsistent—that is, unless the conjunction of the subset's members is formally refutable (Skolem-Löwenheim theorem).

This theorem has some quite startling consequences: in particular, it applies if *S* is the set of theorems of some system of set theory. Then if the system is consistent, *S* has a denumerable model even though *S* may contain a theorem which asserts.the existence of nondenumerable sets. That is not a contradiction: If n represents a nondenumerable set in the model, there will indeed be only countably many *m*'s such that $m \in n$ is true in the model, but the assertion "*n* is nondenumerable" will be true in the model because the model will not contain an object representing the function that enumerates the objects *m* for which $m \in n$ is true in the model. The model is denumerable only from "outside."

This is an example of a model which is nonstandard in that it differs in some essential way from the intended one. The Skolem-Löwenheim theorem also implies the existence of nonstandard models for systems of number theory. In fact, there is a nonstandard model even for the set *S* of all true formulae of elementary arithmetic. The number sequence cannot be characterized up to isomorphism by any countable set of first-order formulae.

The existence of denumerable models of set theory illustrates how essential the platonist conception of set, particularly of the set of subsets of a given set, is to set theory. If there is no more to the platonist conception than is specified in any particular formal system, then apparently the cardinal number of a set cannot be objectively determined. Indeed, the cardinal number of a set depends on what mappings there are and therefore on what sets there are.

The acceptance of this relativity has been urged by many, including Skolem. A fully formalist conception would give rise even to the relativity of the natural numbers themselves.

The completeness theorem and the construction of nonstandard models are fundamental tools in a now rapidly developing branch of logic called model theory. This subject can be viewed as a development of logical semantics, but what is perhaps distinctive about the point of view underlying recent work is that it regards a model of a formal theory as a type of algebraic structure and, in general, that it integrates the semantic study of formal systems with abstract algebra. Model theory takes mathematical logic a long way from the philosophical issues with which we have been mainly concerned, in particular by taking for granted a strong form of platonism. The leaders of this development have, in fact, emphasized the application of metamathematical methods to problems in ordinary mathematics.

There are other investigations concerning the foundations of pure logic. For example, we have mentioned that there can be no decision procedure for quantification theory. Nonetheless, there is interest in the question of what subclasses of formulae are decidable. As a striking result in this direction we might mention the proof of A. S. Kahr, E. F. Moore, and Hao Wang (1962) that the existence of models of formulae of the form "$(x)(\exists y)(z)M(x,y,z)$" (or, equivalently, the provability of formulae of the form "$(\exists x)(y)(\exists z)M(x,y,z)$" where "$M(x,y,z)$" is an arbitrary quantifier-free formula, is undecidable. The development of appropriate concepts of model and completeness proofs for modal logics and intuitionist logic has come to fruition in recent years. In the case of the completeness of intuitionist logic, the situation is unclear. E. W. Beth (1956) has given a construction of models in terms of which he proves classically the completeness of intuitionist quantification theory. On the other hand, Kreisel has shown that the completeness of intuitionist logic cannot be proved by methods available in present intuitionist formal systems and, indeed, that it is incompatible with the supposition that all constructive functions of natural numbers are recursive.

4.5. AXIOMATIC SET THEORY. We shall not undertake here to survey the different axiomatic systems of set theory. We shall, however, mention some developments in the metamathematics of set theory, developments concerning the axiom of choice and Cantor's continuum problem.

The axiom of choice asserts (in one formulation) that for every set *A* of nonempty sets no two of which have a common element, there exists a set *B* which contains exactly one element from each of the sets in *A*. This axiom became prominent when Zermelo used it in 1904 to prove that every set can be well-ordered. Although it was much disputed, it came to be applied more and more, so that entire theories of modern abstract mathematics depend essentially on it. Naturally the question arose whether it was provable or refutable from the other axioms of various systems of set theory. A. A. Fraenkel (1922) showed that it could not be proved from Zermelo's axioms, provided that the axioms allowed individ-

uals—that is, objects which are not sets—in the range of the quantifiers.

The continuum problem appears to be an elementary problem in the arithmetic of cardinal numbers: Is there a cardinal between $\aleph_0$, the cardinal of the integers, and $2^{\aleph_0}$, that of the continuum; stated otherwise, does the continuum contain subsets of cardinal number different from that of the continuum and that of the integers? If the answer is negative, then $2^{\aleph_0} = \aleph_1$, the first cardinal larger than $\aleph_0$, and the cardinal of the first noncountable well-ordering. Cantor's conjecture that $2^{\aleph_0} = \aleph_1$ is called the continuum hypothesis.

Gödel, in 1938, proved that the axiom of choice and a generalization of the continuum hypothesis are consistent with the other axioms. The argument applies to a number of different systems, including the Zermelo-Fraenkel system (ZF). What is proved (finitistically) is that if, say, ZF is consistent, it is likewise consistent with a new axiom, the axiom of constructibility, which implies the axiom of choice and the generalized continuum hypothesis. For the constructible sets, which are the sets obtained by extending the ramified hierarchy of types through all the ordinals, can be proved in the system to satisfy all the axioms plus the axiom of constructibility, which says that every set is constructible. In terms of models, any model of ZF contains a subclass that is a model in which all sets are constructible. The constructible sets are of interest on their own account; Gödel has remarked that the idea behind them is to reduce all impredicativities to one special kind, the existence of large ordinals. However, he does not consider the axiom of constructibility plausible.

Thus, it has been known for some time that the axiom of choice and the continuum hypothesis are not refutable from the other axioms. More recently, Paul J. Cohen proved that they are not provable either. That is, if, say, ZF is consistent, it remains so by adding the negation of the axiom of choice or by adding the axiom of choice and the negation of the continuum hypothesis. Starting from Gödel's ideas, Cohen developed a quite new method for constructing models, which has led very quickly to a large number of further independence results.

The situation with respect to the axiom of choice and the continuum problem raises anew the question of how definite our idea of a set is, whether or not such a question as the continuum problem has an objectively determinate answer. Most mathematicians today find the axiom of choice sufficiently evident. But the continuum hypothesis—perhaps because of its more special character and because of the fact that the analogy of the infinite to the finite on which the conception of the set of all subsets of a given set is based does not suggest a justification of it—is left much more uncertain by considerations of intuitive evidence or plausibility. The role of the Skolem-Löwenheim theorem in Gödel's and Cohen's constructions might encourage the idea that the continuum hypothesis is in fact undetermined. Gödel himself believes that it is false and hopes that an axiom will be found which is as evident as the axiom of choice and which suffices to refute the continuum hypothesis. At present no one seems to have a good idea of what such an axiom would be like. It would have to be of a different character from the usual strong axioms of infinity, to which the method of Gödel's consistency proof applies.

The question of the continuum hypothesis is thus very close to the general epistemological question concerning platonism. If the general conceptions of set and function are given in some direct way to the mind, if, to echo René Descartes, the idea of the infinite is in one's mind before that of the finite, there is no reason to expect a comparatively simple question like the continuum problem to be unanswerable. If, on the other hand, the platonist conceptions are developed by analogies from the area where we have intuitive evidence, if they are "ideas of reason" which, without having an intuition corresponding to them, are developed to give a "higher unity" which our knowledge cannot obtain otherwise, then it would not be particularly surprising if the nature of sets were left indeterminate in some important respect and, indeed, could be further determined in different, incompatible ways.

SUPPLEMENT (2005)

The period since 1967 has seen considerable work in all areas of the foundations of mathematics. This is most notable on the mathematical side. These developments will be discussed before turning to philosophical work.

§5. MATHEMATICAL LOGIC

Of the extensive work since the 1960s, that dealing with formalized axiomatic theories is most central to the foundations of mathematics, although there might now be more debate than earlier about the centrality of the axiomatic method. For some time mathematical logic has been divided into Proof theory, Model theory, Computability (recursion) theory, and Set theory (see the entries on those subjects), although of course there are important interconnections. Model theory and com-

putability theory are more purely mathematical, although their methods are important for the other two areas, and some applications (such as nonstandard analysis) are of foundational interest.

One upshot of work in Proof theory is that strong subsystems of classical analysis (second-order arithmetic) have been analyzed by means that are in some sense constructive but much more powerful and abstract than was envisaged in the early history of the subject. A possibly clearer foundational gain was achieved by another proof-theoretic program, which can trace its roots to Hermann Weyl's (1918) attempt to reconstruct classical analysis predicatively. The work of Harvey Friedman, Stephen Simpson, and others, surveyed in Simpson (1998), showed that many standard theorems of analysis (and of other branches of mathematics) can, if suitably formulated, be proved in weak systems. The method of Reverse mathematics (q.v.) made it possible to calibrate exactly what axiomatic power was needed to prove a particular theorem.

The most striking developments have been in set theory, where Paul Cohen's proof in 1963 of the independence of the axiom of choice and the continuum hypothesis touched off an explosion of research. Cohen's method of forcing proved of wide applicability. In the following years, many more independence results were found in all areas of set theory and its applications. In particular, many classical conjectures were shown both consistent with and independent of the standard axiom system ZFC (or ZF in cases where the axiom of choice sufficed to prove a statement).

This body of work might suggest to a philosopher a vast indeterminacy in the concept of set or of the universe of sets, a random-seeming collection of logical relations among statements independent of ZF or ZFC. However, there is more order than this picture would suggest. The existence of important independent statements would suggest seeking new axioms, and in fact progress has been made by developing the consequences of two kinds of new axioms: strong axioms of infinity (axioms asserting the existence of certain large cardinals) and special cases of the axiom of determinacy.

The large cardinal axioms that have been studied have turned out to be linearly ordered by consistency strength (see §6 of the entry on Set theory), and this has made it possible to determine the consistency strength of other independent statements. In particular this is true of the game-theoretic axiom of determinacy. The assumption PD that the latter holds for projective sets of real numbers (roughly those definable by quantification over reals) implied solutions to the classical problems of descriptive set theory, the study of these sets. PD (and more) was shown to follow from strong large cardinal axioms.

Although this result left the continuum problem untouched, it did show that a program of investigating new axioms along lines proposed by Kurt Gödel in the 1940s could settle an important class of open problems. The large cardinal axioms implying PD have the desirable feature that their consequences in second-order arithmetic cannot be altered by forcing. W. Hugh Woodin's (2001) approach to the continuum problem (see §6 of the entry on Set theory) aims to extend this result to a higher level. But it is not regarded even by Woodin himself as a definitive solution, and even the question whether the continuum hypothesis has a determinate truth-value remains open.

§6. APPROACHES TO PHILOSOPHY OF MATHEMATICS

In 1967 philosophy of mathematics was largely ancillary to logic, and discussion centered either on logical results or on the earlier foundational programs that had contributed to the development of mathematical logic. Since then it has become more a subject in its own right. It has been influenced by the general tendencies moving the philosophy of science away from logic. In particular, historical studies have assumed a larger role, and many such studies have been of developments not close to logic.

In the earlier entry, the philosophical problems discussed concern the analysis of basic mathematical concepts (such as natural number) and the identification and justification of mathematical principles. The term *foundations* naturally suggests that focus. But the philosophy of mathematics can and does contain inquiries of other kinds. It has been charged with concerning itself only with elementary mathematics. This charge is not correct; for example, identifying the axioms required for conclusions in set theory is a matter of high-level mathematical research, and in general the justification of axioms is not independent of knowledge of the theories developed from them.

But it is true that an inquiry into basic concepts and principles will be selective in its attention to the elaboration of mathematics in current and earlier research. And one may well seek philosophical understanding of aspects of mathematical practice of a different kind. One influential strand of work of this kind is that inaugurated by Imre Lakatos, particularly in his book *Proofs and Refutations* (1976). Lakatos studied a classic theorem of Leon-

hard Euler (1707–1783) relating the number of vertices, faces, and edges of a polyhedron and brought to light difficulties that had been found with proofs of it over a period of time and the refinements of the statement of the theorem that had resulted. An underlying idea was that mathematical knowledge is more fallible than a certain traditional picture has it, for a different reason from those that might be suggested by difficulties with basic principles. For reasons of space, this sort of inquiry will not be pursued here, but it should be recognized that this strand of philosophy of mathematics has grown relative to the whole since 1967.

§7. LOGICISM AND THE NEO-FREGAN PROGRAM

In §2, much attention is paid to the project of reducing arithmetic to logic and the analysis of number. Logicism in its earlier forms has not been revived, but a kind of neologicism has become an active program. It was observed that the axioms of arithmetic could be derived in second-order logic from the criterion (7) in §2.6, with numerical equivalence defined as in (8). (This is briefly sketched after (12), but the most difficult case, the proof that every natural number has a successor, is omitted.) (7) thus formulated has come (misleadingly) to be called Hume's principle (HP). The second-order theory with the number operator N_xFx and HP as a nonlogical axiom is called Frege arithmetic (FA). In 1983 Crispin Wright gave the proof that the Dedekind-Peano axioms of second-order arithmetic are provable in FA using Frege's definitions, but this was in essentials proved by Gottlob Frege and has come to be called Frege's theorem. Intuitively, Frege uses the definition of N_xFx in terms of extensions only to derive HP, and then the work is done by that principle. Richard G. Heck Jr. showed in 1993 that this was essentially true of Frege's proofs in *Grundgesetze*. Several logicians showed that FA is consistent if second-order arithmetic is.

Wright's neo-Fregean proposal is to take FA as basic arithmetic. It is a logical construction of arithmetic only if the notion of cardinal number is a logical notion and HP is a principle of logic. As a proof that arithmetic is a part of logic the construction seems to be question-begging. Still, it generated a lot of discussion by Wright and others of the status of abstraction principles like HP, which take an equivalence relation of entities of one kind as a criterion of identity for entities of another kind. Wright's initial idea seems to have been that HP is something close to a definition, although it is not an explicit definition and does not meet the usual standard for a contextual definition, that it should enable the term introduced to be eliminated by paraphrase of contexts in which it occurs. A fatal difficulty for this idea is that HP can be true relative to a domain of individuals only if the domain is infinite. Wright and his collaborators continued to argue that HP is analytic. Others have doubted that a principle that implies the existence of an infinite sequence of objects could be analytic. Another difficulty is that Frege's inconsistent axiom V is an abstraction principle, and other abstraction principles that seem plausible are either inconsistent or can be satisfied only in a finite domain.

The program of axiomatizing parts of mathematics by abstraction principles is of independent logical interest, and work has been done on analysis, and preliminary work on set theory. Kit Fine (2002) carried out an extensive analysis of abstraction principles, to distinguish those that introduce inconsistency from those that do not.

§8. PLATONISM

Since World War II, the view that classical mathematics is seriously threatened by the known paradoxes or by other unknown ones has virtually disappeared. Platonism as described in §3 has been widely accepted as a mathematical method. Taking the language of classical mathematics at face value, as implying the existence of abstract mathematical objects, even forming uncountable and still larger totalities, and allowing reasoning using both the law of excluded middle and impredicative definitions, is probably a default position among philosophers and logicians. This can be called default platonism. It is in relation to such a view, whether accepting it or rejecting it, that much of the work in the philosophy of mathematics since 1967 has concentrated on ontological problems. How might this position be rejected?

§9. CONSTRUCTIVISM

In §3.2, platonism is contrasted principally with constructivism. Intuitionism and other forms of constructivism did not accept the reasoning characteristic of classical mathematics, in the case of intuitionism the law of excluded middle.

A significant development in this area is the argument in favor of intuitionist logic based on considerations of the philosophy of language presented by Michael Dummett (1973). This has, however, had more influence on discussions of realism as a general philosophy than on the foundations of mathematics specifically. Important metamathematical work on intuitionistic theories was done especially in the 1960s and 1970s. An important

development is the development of intuitionistic-type theories that are of much greater expressive power than traditional intuitionistic theories. That of Per Martin-Löf (1984) is the most developed. But although intuitionistic logic has proved to have wide application, intuitionism has declined significantly as a general approach to mathematics, competing with classical mathematics. Another constructive approach to mathematics, pioneered by Errett Bishop (1967), has been developed by several mathematicians. Although it has been more active in the last generation than intuitionism, philosophers have been more interested in the latter, perhaps justifiably because what is philosophically interesting about the Bishop approach is shared with intuitionism, and L. E. J. Brouwer and other intuitionists did more to develop philosophical arguments for their position.

§10. NOMINALISM

The term *platonism* is also used so that the view contrasts with nominalism. Since 1980 or so that opposition has been more prominent among philosophers, especially in North America. This is perhaps fundamentally due to the great influence of scientific naturalism on all theoretical parts of philosophy.

The traditional way in which nominalism rejects default platonism is by not taking the language of mathematics at face value and seeking to paraphrase it in such a way that commitment to abstract mathematical objects is avoided. Programs of this kind have been pursued especially since the 1980s, but it has proved essential to enlarge traditional nominalist resources in at least one of two ways: allowing points and possibly regions of space-time as physical or allowing modality. It is then possible to reconstruct a considerable amount of classical mathematics, at least if one accepts a controversial thesis of George Boolos (1998) that his reading of the language of monadic second-order logic by means of the English plural does not involve commitment to such entities as sets, classes, concepts, or pluralities. What has been achieved in this sort of reconstruction is surveyed in John P. Burgess and Gideon Rosen, *A Subject with No Object* (1997).

A bolder proposal was made by Hartry H. Field (1980, 1989): Where he parted from default platonism was in rejecting the view that statements of classical mathematics, taken at face value with regard to meaning, are true and even that mathematics aims at truth. He sought to account for the apparent objectivity of mathematics by viewing it instrumentally, as a device for making inferences within scientific theories. The role of truth is taken over by conservativeness: Given a nominalistic scientific theory T, a mathematical theory M is conservative if adding its resources to those of T does not enable the derivation of conclusions in the language of T that were not already derivable. This committed him to giving nominalistic versions of scientific theories, and (with the previously mentioned assumption about points and regions of space-time) he was able to give such a version of the Newtonian theory of gravitation. Difficulties stand in the way of carrying out this program for modern physical theories.

§11. STRUCTURALISM

Two related intuitions about modern mathematics are widely expressed: that it is the study of (abstract) structures and that mathematical objects have no more of a nature than is expressed by the basic relations of a structure to which they belong. The structuralist view of mathematical objects is a development of the second intuition. Its relation to default platonism is ambiguous. Some versions, which can be called eliminative structuralism, reject one part of that view, taking the language of mathematics at face value, by proposing paraphrases that eliminate reference to mathematical objects or at least to the most typical mathematical objects. Others take the structuralist idea as an explication of what the reference to objects in standard mathematical language amounts to. This noneliminative type of structuralism offers an ontological gloss on default platonism rather than a modification or rejection of it.

A simple case of an eliminative structuralist analysis is a translation of the language of second-order arithmetic into that of pure second-order logic. Suppose A is a sentence of second-order arithmetic. Since arithmetical operations such as addition and multiplication are second-order definable, it can be assumed that A contains as only primitives N (natural number), S (successor), and 0. The structure of the natural numbers is characterized by a second-order sentence with these primitives, the conjunction P of these axioms. If A is provable, the sentence $P \rightarrow A$ is provable by pure logic. If A is true, it is valid in the standard semantical sense. One can regard $P \rightarrow A$ (or the result of replacing N, S, 0 by variables) as a translation of A that eliminates reference to numbers. The translation has the difficulty that if there is no structure satisfying the axioms, then $P \rightarrow A$ and $P \rightarrow \neg A$ are both vacuously true. The translation seems to presuppose that P is satisfiable.

One version of structuralism would allow sets as basic objects. This would be a natural way of developing

the first intuition, understanding structures as set-theoretic constructs. But a general structuralist view of mathematical objects would naturally aim not to exempt sets from structuralist treatment. At this point modality has been introduced. In the previous example, the assumption that it is possible that there are *N*, *S*, and 0 satisfying *P* is sufficient, since $P \rightarrow A$ can be strengthened to $\Box(P \rightarrow A)$. The modal structuralism of Geoffrey Hellman (1989) is a version of eliminative structuralism relying on this idea. It includes a detailed treatment of set theory. (An approach had been sketched earlier by Hilary Putnam [1967].)

What these constructions accomplish depends on the status of second-order logic, a question that arises also for the neo-Fregean program and for nominalism. Concerning this there has been much debate. Regarding set theory, there is the additional problem that the presupposition of the possibility of the structure is of a structure of such large cardinality that it could not be witnessed by objects that are in any sense concrete or physical, so that the claim of the construction to eliminate reference to mathematical objects can be questioned.

Other versions of structuralism are suggested by remarks of Willard Van Orman Quine (1969) and of some earlier writers. Noneliminative structuralisms have been worked out in some detail by Michael D. Resnik (1997), Stewart Shapiro (1997), and Charles Parsons (1990). Concerning these views, there is debate about the status of structures, as well as about questions about identity.

§12. ROBUST PLATONISM?

A more robust type of platonism is expressed in Gödel's remark that "the set-theoretical concepts and theorems describe some well-determined reality, in which Cantor's conjecture must be either true or false" (1964, p. 260). Such a view would be supported by whatever general considerations support philosophical realism. But something more is demanded, a certain clarity and unambiguity of set-theoretical concepts and quantification over sets. Gödel wished to argue that the continuum hypothesis (CH) must be either true or false, even though he was unable to determine which. What might reinforce his claims would be a development (such as the work of Woodin [2001]) that determines the truth-value of CH. However, the assumptions of such a result might then be incorporated into a less robust platonist view. Perhaps the greater value of Gödelian realism is as a regulative principle: one is more likely to find answers to mathematical questions if one assumes at the outset that there are answers to be found.

That decisive philosophical arguments can be given for such a realistic stance is unlikely. An alternative is to say that default platonism applied to mathematics as it develops represents the limit of what one should claim about the determinateness of the reality described by mathematical theories. This would be the application to mathematics of the naturalistic stance recommended by Quine in many writings, but without his privileging of empirical science. Such a view was advanced by Hao Wang (1974) and more recently by Penelope Maddy (1997).

Gödel's confidence in set-theoretic concepts has not been universally shared; in particular Solomon Feferman (1998, 1999) has defended a skeptical view, influenced by the earlier predicativist tradition.

§13. EPISTEMOLOGICAL PROBLEMS

In the 1967 entry, the epistemological discussion centered on the question whether mathematics can be shown to be *a priori*. It seems that there has been no decisive advance on this question, so others will be concentrated on here.

Paul Benacerraf (1973) raised in rather abstract terms a problem about mathematical knowledge: If default platonism is true, how can one have mathematical knowledge? One response would be to start from the fact that one evidently does have mathematical knowledge and then question the assumptions that generate the problem. One assumption made in Benacerraf's original formulation, the causal theory of knowledge, is relatively easy to reject. To demand a causal relation between objects referred to in a proposition for knowledge of that proposition seems to stack the deck in advance against abstract objects, and the causal theories that were current when he wrote have not stood up well in general epistemology. But one can see the problem in more general terms: Can one give an epistemology for mathematics that is naturalistic? The most fruitful approach might then be to examine actual mathematical knowledge and to consider what sort of explanation of it makes sense and whether it then meets some standard of naturalism.

No explicit program of this kind has been carried far. One place where one might naturally look for naturalistic explanation is psychology, and there has been a considerable amount of research on the development of concepts of number in young children. Although the questions are often framed in terms of the concept of set, it is not clear that that is essential or that ontology is at all

central to the formulation of the problems. It can be argued that mathematical ontology only arises at a more advanced state of the development of mathematical competence than the children investigated have reached.

When one does consider even the mathematics taught in elementary college courses, then what one has to go on is history and the reflection of mathematicians (and sometimes philosophers) on the justification of their claims. That some basic statements and inferences are rationally evident seems an inescapable assumption. Examples would be simple logical inferences and the most elementary axioms of set theory, such as the pairing axiom. It does not mean that this evidence does not get crucial reinforcement from the development of theories based on these evident starting points or that the latter can never be revised in the light of the further development of knowledge. Other assumptions might become evident when an edifice of knowledge has been built up; that might be true of higher-level set-theoretic axioms such as power set and choice. What possible explanations of rational evidence would count as naturalistic is a question that has not been much explored. But now any grounds for holding that no acceptable explanation is possible would have to rely on *a priori* presuppositions.

A less abstract and perhaps more interesting epistemological question arises particularly for higher set theory. It is suggested by the indispensability argument mentioned earlier. Whatever one thinks of rational evidence in general, it is already diminished when one reaches the usual axioms for the mathematics applied in science, as is indicated by the issues about the law of excluded middle raised by Brouwer, and those about impredicativity raised by Poincaré (1908) and Weyl (1918, 1919). However, a long history of successful application convinces one, for example, that the classical mathematics of the continuum is necessary for science and at least as well established as basic physics itself. This is the claim made by the indispensability argument, and it had been suggested earlier by Bertrand Russell and then Gödel that axioms could derive their evident character from the theory they give rise to. Among the applications of mathematics, however, are those within mathematics. Gödel's view apparently was that much of mathematics (including some higher set theory) could be seen to be evident in an *a priori* way, not contaminated by evidence derived from application in empirical science. However, particularly in higher set theory axioms could obtain additional justification through the theories constructed on their basis, and such justification would be possible for stronger axioms, such as the stronger large cardinal axioms that have been proposed, where a convincing intrinsic justification is not available.

Gödel's view and the indispensability argument have in common that the justification of mathematical axioms can rest at least to a certain degree on their consequences. However, for Gödel this is compatible with the status of mathematics as rational knowledge independent of experience, whereas for the main proponents of the indispensability argument, Quine and Putnam (1971), it is not. The indispensability argument clearly runs out before higher set theory. Empirical science makes no use of it, and indeed it has been argued that from the proof theorist's point of view the mathematical theories that are applied in science are weak.

Since few are satisfied with intrinsic justifications for the strongest axioms of infinity, and little such justification is claimed for determinacy axioms, the accepted solution to the classical problems of descriptive set theory rests on assumptions whose justification depends on the theory they give rise to (see Martin 1998). The same would have to be admitted for any solution to the continuum problem that can be expected in the forseeable future.

§14. HISTORICAL STUDIES

Practically every aspect of the history of the foundations of mathematics has seen some intensive scholarly study in the period since 1967. With respect to Immanuel Kant, a decisive development was Michael Friedman's *Kant and the Exact Sciences* (1992), which integrated Kant's philosophy of mathematics with his philosophy of physics and gave the strongest version of the logical view of the role of intuition in mathematics pioneered by Evert Willem Beth (1959) and Jaakko Hintikka (1974). Younger scholars have followed up Friedman's work, often criticizing aspects of it. In particular they have explored the relation of Kant's thought about mathematics to the mathematics of his own time and earlier and to the philosophy of his immediate predecessors.

One strand of work on Frege, of which Boolos and Heck (see Demopoulos 1995) have been the leaders, has worked out perspicuously the mathematical content of Frege's work, particularly in *Grundgesetze*. Another strand has emphasized his conception of logic and how it differs from our own conception of logic. A third has drawn connections of Frege to nineteenth-century developments in mathematics, particularly geometry.

The foundations of mathematics as an object of special study arose from the revolution in mathematics in

the nineteenth century, particularly developments in its second half: the rigorization of the methods of analysis, the beginning of set theory and of abstract methods, the rise of modern logic, and the role assumed early in the twentieth century by the paradoxes. Every aspect of this development has been the subject of scholarly study. The same holds of later developments such as Russell's logic, Brouwer's intuitionism, the Hilbert program, and the work of the Vienna Circle. Space does not permit describing this work, but in the bibliography selective references have been given.

See also Aristotle; Brouwer, Luitzen Egbertus Jan; Cantor, Georg; Carnap, Rudolf; Church, Alonzo; Constructivism and Conventionalism; Descartes, René; First-Order Logic; Frege, Gottlob; Geometry; Gödel, Kurt; Gödel's Theorem; Hilbert, David; Infinity in Mathematics and Logic; Intuitionism and Intuitionistic Logic; Kant, Immanuel; Knowledge, A Priori; Logic, History of; Logical Paradoxes; Mill, John Stuart; Modal Logic; Neo-Kantianism; Neumann, John von; Nominalism, Modern; Peano, Giuseppe; Poincaré, Jules Henri; Proof Theory; Quantifiers in Formal Logic; Quine, Willard Van Orman; Realism and Naturalism, Mathematical; Russell, Bertrand Arthur William; Second-Order Logic; Set Theory; Structuralism, Mathematical; Tarski, Alfred; Turing, Alan M.; Types, Theory of; Weyl, (Claus Hugo) Hermann; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

TEXTBOOKS

Boolos, George, John P. Burgess, and Richard C. Jeffrey. *Computability and Logic*. 4th ed. New York: Cambridge University Press, 2002. The first edition of this book (by Boolos and Jeffrey) was published in 1974.

Enderton, Herbert B. *A Mathematical Introduction to Logic*. 2nd ed. San Diego: Harcourt/Academic Press, 2001. The first edition of this book was published in 1972.

George, Alexander, and Daniel J. Velleman. *Philosophies of Mathematics*. Oxford, U.K.: Blackwell, 2002.

Shapiro, Stewart. *Thinking about Mathematics*. New York: Oxford University Press, 2000.

Shoenfield, Joseph R. *Mathematical Logic*. Reading, MA: Addison-Wesley, 1967. Reprinted Urbana, IL: Association for Symbolic Logic, 2001.

COLLECTIONS OF PAPERS

Benacerraf, Paul, and Hilary Putnam, eds. *Philosophy of Mathematics: Selected Readings*. 2nd ed. (with revised selection). Cambridge, U.K.: Cambridge University Press, 1983.

Ewald, William, ed. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. 2 vols. Oxford, U.K.: Clarendon Press, 1996.

Gödel, Kurt. *Collected Works*, 5 vols. Edited by Solomon Feferman et al. New York: Oxford University Press, 1986–2003. Vol. I, *Publications 1929–1936*, 1986. Vol. II, *Publications 1938–1974*, 1990. Vol. III, *Unpublished Essays and Lectures*, 1995. Volumes. IV and V, *Correspondence*, Oxford: Clarendon Press, 2003.

Hart, W. D., ed. *The Philosophy of Mathematics*. New York: Oxford University Press, 1996.

Shapiro, Stewart, ed. *The Oxford Handbook of Philosophy and Mathematics and Logic*. New York: Oxford University Press, 2005.

Van Heijenoort, Jean, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press, 1967.

§1. The Axiomatic Method

1.1. Axiomatization

Beth, Evert Willem. *The Foundations of Mathematics: A Study in the Philosophy of Science*. Amsterdam, Netherlands: North-Holland, 1959. Parts II and III.

Euclid. *The Thirteen Books of Euclid's Elements*, 3 vols. 2nd ed. Translated by Thomas L. Heath. New York: Dover, 1956. The first edition of this book was published in 1908.

1.2. The abstract viewpoint

Hilbert, David. *Grundlagen der Geometrie*. 10th ed. Stuttgart, Germany: Teubner, 1968. Translated as *Foundations of Geometry* (Chicago: Open Court, 1971). The first German-language edition of this book was published in 1899.

Tarski, Alfred. "What Is Elementary Geometry?" In *The Axiomatic Method*, edited by Leon Henkin, Patrick Suppes, and Alfred Tarski. Amsterdam, Netherlands: North-Holland, 1959.

1.3. Formalization

Church, Alonzo. *Introduction to Mathematical Logic*. Princeton, NJ: Princeton University Press, 1956.

Hilbert, David, and Paul Bernays. *Grundlagen der Mathematik*, 2 vols. 2nd ed. Berlin: Springer, 1968–1970. The first edition of this book was published in 1934–1939.

Kleene, Stephen Cole. *Introduction to Metamathematics*. New York: Van Nostrand, 1952.

Tarski, Alfred. *Introduction to Logic*. New York: Oxford University Press, 1941.

§2. Epistemological Discussion

2.2 Mathematics and logic

Frege, Gottlob. *Grundgesetze der Arithmetik*, 2 vols. Hildesheim, Germany: Olms, 1962. Partly translated by Montgomery Furth as *The Basic Laws of Arithmetic: Exposition of the System* (Berkeley: University of California Press, 1964). The first German-language edition of this book was published in 1893. Vol. 2, 1903.

Frege, Gottlob. *Die Grundlagen der Arithmetik*. Edited by Christian Thiel. Hamburg, Germany: Meiner, 1986. Translated by J. L. Austin as *The Foundations of Arithmetic: A Logico-mathematical Enquiry into the Concept of Number*. 2nd ed. (Oxford, U.K.: Blackwell, 1953). The first German-language edition of this book was published in 1884.

Russell, Bertrand. *Introduction to Mathematical Philosophy*. London: Allen and Unwin, 1919.

Russell, Bertrand. *The Principles of Mathematics*. 2nd ed. London: Allen and Unwin, 1937. The first edition of this book was published in 1903.

Whitehead, A. N., and Bertrand Russell. *Principia Mathematica*, 3 vols. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1925–1927.

2.4. Axioms of arithmetic

Dedekind, Richard. *Was sind und was sollen die Zahlen?* 3rd ed. Braunschweig, Germany: Vieweg, 1911. The first edition of this book was published in 1888. Translation in Ewald.

George and Velleman, above.

Hilbert and Bernays, above.

Kleene, *Introduction*, above.

Peano, Giuseppe. *Arithmetices principia nova methodo exposita*. Turin, Italy: Bocca, 1889. Translation in van Heijenoort.

Waismann, Friedrich. *Introduction to Mathematical Thinking*. New York: Ungar, 1951.

2.5 The concept of class (set)

Benacerraf and Putnam, above. Papers in part IV of 2nd ed. on the concept of set.

Cantor, Georg. *Contributions to the Founding of the Theory of Transfinite Numbers*. Chicago: Open Court, 1915. Translation of papers written from 1895 to 1897.

Cantor, Georg. "Foundations of a General Theory of Manifolds." In Ewald 1996. (With some other short texts.) Translation of a paper of 1883.

Cantor, Georg. *Gesammelte Abhandlungen*. Edited by Ernst Zermelo. Hildesheim, Germany: Olms, 1962. The first publication of this book was in 1932.

Fraenkel, A. A., Yehoshua Bar-Hillel, and Azriel Lévy. *Foundations of Set Theory*. 2nd ed. Amsterdam: North-Holland, 1973.

Gödel, Kurt. "The Present Situation in the Foundations of Mathematics." In Gödel 1986–2003, Vol. III, 45–53.

2.6 Frege's analysis of number

Frege, Gottlob. *Begriffschrift*. Hildesheim, Germany: Olms, 1964. The first publication of this book was in 1879. Translation in van Heijenoort.

Frege, Gottlob. The Foundations of Arithmetic, above.

Frege, Gottlob. Grundgesetze der Arithmetik, above.

George and Velleman, chapter 2.

2.7.Difficulties in logicism

Carnap, Rudolf. "Die logizistische Grundlegung der Mathematik." *Erkenntnis* 2 (1931): 91–105. Translation in Benacerraf and Putnam.

Poincaré, Henri. *Science et méthode*. Paris: Flammarion, 1908. Translated by Francis Maitland as *Science and Method* (New York: Nelson, 1914).

Wittgenstein, Ludwig. *Remarks on the Foundations of Mathematics*. Rev. ed. Translated by G. E. M. Anscombe. Edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe. Cambridge, MA: MIT Press, 1978.

2.8 Kant's view

Kant, Immanuel. *Kritik der reinen Vernunft*. Riga, Russia: N.p., 1781. Translated by Paul Guyer and Allen W. Wood as *Critique of Pure Reason* (New York: Cambridge University Press, 1998). See especially Introduction (in 2d ed.), Transcendental Aesthetic, Axioms of Intuition, and Discipline of Pure Reason in Dogmatic Use.

Kant, Immanuel. "Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral" (1764). In *Gesammelte Schriften*, Vol. 2, 272–301. Berlin: G. Reimer, 1902. Translated by David Walford in *Theoretical Philosophy, 1755–1770* (New York: Cambridge University Press, 1992).

2.9 Conventionalism

Carnap, Rudolf. *Logische Syntax der Sprache*. Vienna, Austria: Springer, 1934. Translated as *Logical Syntax of Language* (London: Routledge and Kegan Paul, 1937).

Carnap, Rudolf. *Meaning and Necessity*. 2nd ed. Chicago: University of Chicago Press, 1956.

Gödel, Kurt. "Is Mathematics Syntax of Language?" In Gödel 1986–2003, Vol. III, 334–362. The introduction by Warren Goldfarb (324–334) questions the interpretation of Carnap as a conventionalist.

Quine, W. V. "Carnap and Logical Truth" (1960). In *The Ways of Paradox and Other Essays*. 2nd ed. Cambridge, MA: Harvard University Press, 1976.

Quine, W. V. "Truth by Convention" (1936). In *The Ways of Paradox and Other Essays*. 2nd ed. Cambridge, MA: Harvard University Press, 1976.

Wittgenstein, Ludwig. *Lectures on the Foundations of Mathematics, Cambridge, 1939*. Edited by Cora Diamond. Ithaca, NY: Cornel University Press, 1976.

Wittgenstein, Ludwig. *Philosophical Investigations*. 2nd ed. Translated by G. E. M. Anscombe. Oxford, U.K.: Blackwell, 1958.

Wittgenstein, Ludwig. Remarks on the Foundations of Mathematics, above.

§3. Platonism and Constructivism

3.1. Platonism

Bernays, Paul. "Sur le platonisme dans les mathématiques." *L'enseignement mathématique* 34 (1935): 52–69. Translation in Benacerraf and Putnam.

Church, Alonzo. "Comparison of Russell's Solution of the Semantical Antinomies with That of Tarski." *Journal of Symbolic Logic* 41 (1976), 747–760. Gives a precise formulation of the ramified theory of types, taking account, as §3.1.3 does not, of its intensional aspect.

Feferman, Solomon. "Systems of Predicative Analysis." *Journal of Symbolic Logic* 29 (1964): 1–30. Part 1 surveys work on predicativity up to 1963.

Gödel, Kurt. "Russell's Mathematical Logic" (1944). In Gödel 1986–2003, Vol. II, 119–141.

Gödel, Kurt. "What Is Cantor's Continuum Problem?" (1947, 1964), In Gödel 1986–2003, Vol. II, 254–270.

Russell, Bertrand. "Mathematical Logic as Based on the Theory of Types" (1908). In *Logic and Knowledge: Essays 1901–1950*, edited by Robert Charles Marsh. New York: Macmillan, 1956.

Weyl, Hermann. "Der *circulus vitiosus* in der heutigen Begründung der Analysis." *Jahresbericht der Deutschen Mathematiker-Vereinigung* 28 (1919): 85–92. Translated as appendix to *The Continuum*.

Weyl, Hermann. *Das Kontinuum*. Leipzig, Germany: Veit, 1918. Translated as *The Continuum* (Kirksville, MO: Thomas Jefferson University Press, 1987).

§3.2 Constructivism

Gödel, Kurt. "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes." *Dialectica* 12 (1958): 280–287. Translated in Gödel 1986–2003, Vol. II, 240–251. Note also the expanded English version (1972) in the same volume.

Heyting, Arend. *Intuitionism: An Introduction*. Amsterdam, Netherlands: North-Holland, 1956.

Hilbert, David. "Die Grundlagen der Mathematik." *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität* 6 (1928): 65–85. Translation in van Heijenoort.

Hilbert, David. "Über das Unendliche." *Mathematische Annalen* 95 (1926): 161–190. Translation in van Heijenoort.

Hilbert and Bernays, above.

Tait, W. W. "Finitism." In *The Provenance of Pure Reason*. New York: Oxford University Press, 2005. This article originally appeared in the *Journal of Philosophy* (1981).

Troelstra, A. S. *Principles of Intuitionism*. Berlin: Springer, 1969.

Weyl, Hermann. *Philosophy of Mathematics and Natural Science*. Princeton, NJ: Princeton University Press, 1949.

4.1. Gödel's incompleteness theorems

Gödel, Kurt. "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I." *Monatshefte für Mathematik und Physik* 38 (1931): 173–198. Translated in Gödel 1986–2003, Vol. I, 144–195.

Smullyan, Raymond M. *Gödel's Incompleteness Theorems*. New York: Oxford University Press, 1992.

Gödel's incompleteness theorem is treated in each of the general works on mathematical logic cited above and in many other works.

4.2. Recursive function theory

Kleene, *Introduction*, above.

Rogers, Hartley, Jr. *Theory of Recursive Functions and Effective Computability*. New York: McGraw-Hill, 1967.

4.3. Development of the Hilbert program

Gentzen, Gerhard. *Collected Papers*. Edited by M. E. Szabo. Amsterdam, Netherlands: North-Holland, 1969.

Gentzen, Gerhard. "Die Widerspruchsfreiheit der reinen Zahlentheorie." *Mathematische Annalen* 112 (1936): 493–565. Translated in Gentzen 1969.

Gödel, Kurt. "Zur intuitionistischen Arithmetik und Zahlentheorie." *Ergebnisse eines mathematischen Kolloquiums* 4 (1933): 34–38. Translated in Gödel 1986–2003, Vol. I, 286–295.

Kreisel, Georg. "Hilbert's Programme." *Dialectica* 12 (1958): 346–372.

Kreisel, Georg. "On the Interpretation of Non-finitist Proofs." *Journal of Symbolic Logic* 16 (1951): 241–267; 17 (1952): 43–58.

Schütte, Kurt. *Proof Theory*. New York: Sprinter-Verlag, 1977.

Spector, Clifford. "Provably Recursive Functionals of Analysis." In *Recursive Function Theory, Proceedings of Symposia in Pure Mathematics*, Vol. 5, 1–27. Providence, RI: American Mathematical Society, 1962.

Takeuti, Gaisi. *Proof Theory*. 2nd ed. Amsterdam, Netherlands: North-Holland, 1987. This edition contains important appendices by other authors.

4.4. Foundations of logic

Bell, J. L., and A. B. Slomson. *Models and Ultraproducts*. Amsterdam, Netherlands: North-Holland, 1969.

Gentzen, Gerhard. "Untersuchungen über das logische Schliessen." *Mathematische Zeitschrift* 39 (1934): 176–210, 405–431. Translated in Gentzen 1969.

Gödel, Kurt. "Die Vollständigkeit der Axiome des logischen Funktionenkalküls." *Monatshefte für Mathematik und Physik* 37 (1930): 349–360. Translated in Gödel 1986–2003, Vol. I, 102–123.

Herbrand, Jacques. *Écrits logiques*. Edited by Jean van Heijenoort. Paris: Presses Universitaires de France, 1968. Translated as *Logical Writings* (Dordrecht, Netherlands: D. Reidel, 1971).

Prawitz, Dag. *Natural Deduction: A Proof-Theoretical Study*. Stockholm, Sweden: Almqvist and Wiksell, 1965.

Skolem, Thoralf. *Selected Works in Logic*. Edited by J. E. Fenstad. Oslo, Norway: Universitetsforlaget, 1970. Some of the important papers are translated in van Heijenoort with significant introductions.

Tarski, Alfred. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. 2nd ed. Translated by J. H. Woodger. Edited by John Corcoran. Indianapolis, IN: Hackett, 1983. The first edition was published in 1956.

4.5. Axiomatic set theory

Cohen, Paul J. "The Independence of the Continuum Hypothesis." *Proceedings of the National Academy of Sciences, USA* 50 (1963): 1143–1148; 51 (1964): 105–110.

Cohen, Paul J. *Set Theory and the Continuum Hypothesis*. New York: Benjamin, 1966.

Kunen, Kenneth. *Set Theory: An Introduction to Independence Proofs*. Amsterdam, Netherlands: North-Holland, 1980.

Scott, Dana. "A Proof of the Independence of the Continuum Hypothesis." *Mathematical Systems Theory* 1 (1967), 89–111. A lucid presentation using the alternate method of Boolean-valued models.

§5. Mathematical Logic

Barwise, Jon, ed. *Handbook of Mathematical Logic*. Amsterdam, Netherlands: North-Holland, 1977.

Kanamori, Akihiro. *The Higher Infinite*. 2nd ed. Berlin: Springer, 2003.

Simpson, Stephen G. *Subsystems of Second-Order Arithmetic*. Berlin: Springer, 1998.

Woodin, W. Hugh. "The Continuum Hypothesis." *Notices of the American Mathematical Society* 48 (2001): 567–576, 681–690.

See the bibliographies of Computability theory, Model theory, Proof theory, and Set theory.

§6. Approaches to the Philosophy of Mathematics

George, Alexander, and Daniel J. Velleman. *Philosophies of Mathematics*. Oxford, U.K.: Blackwell, 2002.

Grosholz, Emily, and Herbert Breger, eds. *The Growth of Mathematical Knowledge*. Dordrecht, Netherlands: Kluwer Academic, 2000.

Lakatos, Imre. *Proofs and Refutations*. New York: Cambridge University Press, 1976.

Shapiro, Stewart. *Thinking about Mathematics*. New York: Oxford University Press, 2000.

§7. Logicism and the Neo-Fregean Program

Boolos, George. *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press, 1998.

Demopoulos, William, ed. *Frege's Philosophy of Mathematics*. Cambridge, MA: Harvard University Press, 1995.

Fine, Kit. *The Limits of Abstraction*. New York: Oxford University Press, 2002.

Hale, Bob, and Crispin Wright. *The Reason's Proper Study*. Oxford, U.K.: Clarendon Press, 2001.

Wright, Crispin. *Frege's Conception of Numbers as Objects*. Aberdeen, Scotland: Aberdeen University Press, 1983.

§9. Constructivism

Bishop, Errett. *Foundations of Constructive Analysis*. New York: McGraw-Hill, 1967.

Bishop, Errett, and Douglas Bridges. *Constructive Analysis*. Berlin: Springer, 1985.

Dummett, Michael. "The Philosophical Basis of Intuitionistic Logic" (1973). In *Truth and Other Enigmas* (London: Duckworth, 1978).

Martin-Löf, Per. *Intuitionistic Type Theory*. Naples, Italy: Bibliopolis, 1984.

Troelstra, A. S., ed. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*. Berlin: Springer, 1973.

Troelstra, A. S., and D. van Dalen. *Constructivism in Mathematics*, 2 vols. Amsterdam, Netherlands: North-Holland, 1988.

§10. Nominalism

Burgess, John P., and Gideon Rosen. *A Subject with No Object: Strategies for Nominalist Reconstruction of Mathematics*. Oxford, U.K.: Clarendon Press, 1997.

Field, Hartry H. *Realism, Mathematics, and Modality*. Oxford, U.K.: Blackwell, 1989.

Field, Hartry H. *Science without Numbers*. Princeton, NJ: Princeton University Press, 1980.

Malament, David. Review of *Science without Numbers*, by Hartry H. Field. *Journal of Philosophy* 79 (1982): 523–534.

§11. Structuralism

Hellman, Geoffrey. *Mathematics without Numbers*. Oxford, U.K.: Clarendon Press, 1989.

Parsons, Charles. "The Structuralist View of Mathematical Objects." *Synthese* 84 (1990): 303–346.

Putnam, Hilary. "Mathematics without Foundations." *Journal of Philosophy* 64 (1967): 5–22.

Quine, W. V. *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Resnik, Michael D. *Mathematics as a Science of Patterns*. Oxford, U.K.: Clarendon Press, 1997.

Shapiro, Stewart. *Philosophy of Mathematics: Structure and Ontology*. New York: Oxford University Press, 1997.

§12. Robust Platonism?

Feferman, Solomon. "Does Mathematics Need New Axioms?" *American Mathematical Monthly* 106 (1999): 99–111.

Feferman, Solomon, et al. "Does Mathematics Need New Axioms?" *Bulletin of Symbolic Logic* 6 (2000): 401–446.

Feferman, Solomon. *In the Light of Logic*. New York: Oxford University Press, 1998.

Maddy, Penelope. *Naturalism in Mathematics*. Oxford, U.K.: Clarendon Press, 1997.

Maddy, Penelope. *Realism in Mathematics*. Oxford, U.K.: Clarendon Press, 1990.

Tait, W. W. "Truth and Proof: The Platonism of Mathematics." *Synthese* 69 (1986): 341–370.

Wang, Hao. *From Mathematics to Philosophy*. London: Routledge and Kegan Paul, 1974.

§13. Epistemological Problems

Benacerraf, Paul. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–679.

Hauser, Kai. "Is Cantor's Continuum Problem Inherently Vague?" *Philosophia Mathematica* (III) 10 (2002): 257–285.

Maddy, Penelope. "Indispensability and Practice." *Journal of Philosophy* 89 (1992): 275–289.

Martin, Donald A. "Mathematical Evidence." In *Truth in Mathematics*, edited by H. G. Dales and G. Oliveri, 214–231. Oxford, U.K.: Clarendon Press, 1998.

Parsons, Charles. "Reason and Intuition." *Synthese* 125 (2000): 299–315.

Putnam, Hilary. *Philosophy of Logic*. New York: Harper and Row, 1971.

Putnam, Hilary. "What Is Mathematical Truth?" In *Mathematics, Matter, and Method*. 2nd ed. New York: Cambridge University Press, 1979.

§14. Historical Studies

Boolos, George. *Logic, Logic, and Logic*, above. Section on Frege Studies.

Demopoulos, William, ed. *Frege's Philosophy of Mathematics*, above.

Dummett, Michael. *Frege: Philosophy of Mathematics*. Cambridge, MA: Harvard University Press, 1991.

Ewald, William B., ed. *From Kant to Hilbert*, above.

Ferreiros, José. *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*. Basel, Switzerland: Birkhäuser, 1999.

Friedman, Michael. *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press, 1992.

Goldfarb, Warren. "Frege's Conception of Logic." In *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy*, edited by Juliet Floyd and Sanford Shieh, 25–41. New York: Oxford University Press, 2001.

Goldfarb, Warren, and Thomas Ricketts. "Carnap's Philosophy of Mathematics." In *Science and Subjectivity*, edited by David Bell and Wilhelm Vossenkuhl. Berlin: Akademie-Verlag, 1992.

Gödel, Kurt, *Collected Works*, above. Introductory notes by various writers.

Hallett, Michael. *Cantorian Set Theory and Limitation of Size*. Oxford, U.K.: Clarendon Press, 1984.

Heck, Richard G., Jr. "The Development of Arithmetic in Frege's *Grundgesetze der Arithmetik*" (1993) and other essays reprinted in Demopoulos.

Hintikka, Jaakko. *Knowledge and the Known. Historical Perspectives in Epistemology*. Dordrecht, Netherlands: Reidel, 1974.

Kneale, William, and Martha Kneale. *The Development of Logic*. Oxford, U.K.: Clarendon Press, 1962.

Mancosu, Paolo, ed. *From Brouwer to Hilbert: The Debate on Foundations of Mathematics in the 1920s*. New York: Oxford University Press, 1998.

Parsons, Charles. "Platonism and Mathematical Intuition in Kurt Gödel's Thought." *Bulletin of Symbolic Logic* 1 (1995): 44–74.

Posy, Carl J., ed. *Kant's Philosophy of Mathematics: Modern Essays*. Dordrecht, Netherlands: Kluwer Academic, 1992.

Ricketts, Thomas. "Logic and Truth in Frege." *Aristotelian Society Supplementary Volume* 70 (1996): 121–140.

Shabel, Lisa. "Kant on the 'Symbolic Construction' of Mathematical Concepts." *Studies in the History and Philosophy of Science* 29 (1998): 589–621.

Shabel, Lisa. "Kant's Philosophy of Mathematics." In *The Cambridge Companion to Kant and Modern Philosophy*. Edited by Paul Guyer. New York: Cambridge University Press, forthcoming.

Sieg, Wilfried. "Hilbert's Programs, 1917–1922." *Bulletin of Symbolic Logic* 5 (1999): 1–44.

Sieg, Wilfried, and Dirk Schlimm. "Dedekind's Analysis of Number: Systems and Axioms." *Synthese* forthcoming.

Sutherland, Daniel. "Kant on Arithmetic, Algebra, and the Theory of Proportion." *Journal of the History of Philosophy* forthcoming.

Tait, William. *The Provenance of Pure Reason*. New York: Oxford University Press, 2005. See especially essays 5 and 10–12.

Van Heijenoort, Jean, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press, 1967.

Charles Parsons (1967, 2005)

MATHER, COTTON

(1663–1728)

Cotton Mather, scholar, clergyman, and author, was the oldest son of Increase Mather, one of the leading figures in the Puritan theocracy in Massachusetts. The younger Mather was so precocious that he entered Harvard College at the age of twelve and was graduated at fifteen. Because he stammered, he felt unqualified to preach and therefore began to study medicine. After a few years, however, he overcame his speech handicap and became the assistant to his father at the Second Church, Boston. Ordained in 1685, he remained in the service of the Second Church for the rest of his life.

Mather was disappointed in many of the major quests of his life. Partly because he associated himself politically with the unpopular royal governor, Sir William Phips, partly because of the diminished prestige of the Puritan clergy, and partly because of his own often unpleasant personal qualities he lost the power to wield significant influence in public affairs. When he greatly desired to succeed his father, who retired in 1701 as president of Harvard College, he was not selected. Convinced that Harvard no longer represented the true Calvinist faith, he threw himself energetically into the foundation of Yale College, but its presidency was not offered to him until 1721, when he declined the position because of his age.

Mather's intellectual attitudes during his earlier years were extremely narrow, for he moved within the confines of a strict Puritan worldview; later, however, he became more tolerant of the differing beliefs of others. Finally, especially in his *Christian Philosopher* (1721), he moved close to the natural religion characteristic of the Age of Reason. He interpreted the theological doctrine of divine Providence in philosophical terms by asserting that the order of the universe was planned for man's good by an all-wise, all-good God. Man's appreciation of natural Beauty and his application of reason to observations drawn from nature are sufficient to prove the existence and beneficence of God. His scientific communications to the Royal Society of London led to his election as a fellow in 1713, one of the first Americans to be so honored. He was one of the earliest in the colonies to advocate inoculation against smallpox, and he ably defended his position in several pamphlets. The change in his mental attitude thus epitomizes the alteration in the intellectual life that pervaded his milieu.

Nowhere is this duality more apparent than in Mather's involvement in the witchcraft epidemic in Salem. He attempted to make a "scientific" study of the cases, but he came to the conclusion that they could be treated by prayer and fasting. He warned the judges in witchcraft trials to proceed very cautiously against the suspects and to be particularly careful in admitting "spectral evidence," yet in his *Wonders of the Invisible World* (1693) he argued that the verdicts in the Salem trials were justified. By 1700, however, he changed his mind about the fairness of the trials. In regard to the suspicion of witchcraft, as in other respects, Mather stood uneasily between traditional faith and the new scientific outlook.

See also Philosophy of Religion, History of; Scientific Method.

Bibliography

Mather's most important works (of more than 450 published) are *Magnalia Christi Americana, or the Ecclesiastical History of New England* (London, 1702); *Essays to Do Good* (Boston, 1710; originally titled *Bonifacius*, Boston, 1710); and *Christian Philosopher* (London: E. Matthews, 1721). Kenneth B. Murdock has edited, with introduction and notes, *Selections from Cotton Mather* (New York: Harcourt Brace, 1926; new ed., 1960).

Discussion of Mather may be found in Ralph P. and Louise Boas, *Cotton Mather, Keeper of the Puritan Conscience* (New

York and London: Harper, 1928); Barrett Wendell, *Cotton Mather, the Puritan Priest* (New York: Dodd Mead, 1891; new ed., with introduction by Alan Heimert, New York: Harcourt Brace, 1963); and Otho T. Beall and Richard Shryock, *Cotton Mather, First Significant Figure in American Medicine* (Baltimore: Johns Hopkins Press, 1954).

OTHER RECOMMENDED WORKS

An Edition of Paterna, Cotton Mather's Previously Unpublished Autobiography, Complete, with Introduction and Notes. Edited by Ronald A. Bosco. Diss., University of Maryland, 1975.

Diary of Cotton Mather. Salem, MA: Higginson, 1997.

Mather Microfilm: Part I, The Papers of Cotton Mather. Boston: Massachusetts Historical Society, 1973.

The Papers of Cotton Mather. Boston, MA: Massachusetts Historical Society; Worcester, MA: American Antiquarian Society, 1981.

Paterna: The Autobiography of Cotton Mather. Delmar, NY: Scholars' Facsimiles & Reprints, 1976.

Scheick, William J., ed. *Two Mather Biographies: Life and Death and Parentator*. Cranbury, NJ: Associated University Presses, 1989.

Selected Letters of Cotton Mather. Compiled with commentary by Kenneth Silverman. Baton Rouge: Louisiana State University Press, 1971.

The Witchcraft Delusion in New England; Its Rise, Progress, and Termination, as Exhibited by Dr. Cotton Mather in The Wonders of the Invisible World, and by Mr. Robert Calef in His More Wonders of the Invisible World. Compiled by Samuel Gardner Drake. New York: B. Franklin, 1970.

J. L. Blau (1967)
Bibliography updated by Michael J. Farmer (2005)

MATTER

The term "matter" and its cognates ("material," "materialist," "materialistic," and the like) have played active parts in philosophical debate throughout intellectual history. Natural philosophers have studied material objects and contrasted them with such immaterial agencies as energy and fields of force; metaphysicians and mathematical philosophers have distinguished the material or tangible aspects of things from their formal or intangible aspects, their physical properties from their geometrical ones. Again, the terms "matter" and "material" have played a humble part not only in science but also in moral philosophy and even theology. Matter has thus been placed in opposition to life and mind, soul and spirit, and a preoccupation with worldly pleasures and bodily comforts, as opposed to the "higher" pleasures of the mind, has been condemned as "materialistic" and unworthy of spiritual beings. In thinking about matter, accordingly, the question of how far—if at all—these various distinctions can actually be justified and reconciled must always be borne in mind.

This question immediately poses a historical problem, for ideas about matter have not been static. On the contrary, they have been subject to continual development, and it is highly doubtful whether one can isolate a single concept of matter shared by, say, Anaximander and Thomas Aquinas, Democritus and René Descartes, Epicurus and Albert Einstein. Thus, for instance, a seventeenth-century philosophical thesis about the relations between mind and matter must be interpreted in relation to seventeenth-century ideas about physics and chemistry. Such a thesis can be transplanted into the intellectual environment of the twentieth century only by taking into account changes in the fundamental concepts of science during the intervening years. We must therefore consider how the concept of matter has been progressively refined and modified in the course of intellectual history.

GREEK PHILOSOPHY

As far as we can judge from the surviving texts and the testimony of Aristotle, the idea of a constituent or material ingredient (*hyle*) common to things of all kinds was a central concept of the Ionian school of philosophy. The Ionian philosophers, beginning with Thales of Miletus, disagreed about the nature of this common ingredient. Some likened it to water, others to air or breath, others to fire; some insisted that it could have no properties analogous to those of any familiar substance but must be entirely undifferentiated or unlimited. Yet they agreed, at any rate, in their statement of the basic philosophical problem: "What universal, permanent substance underlies the variety and change of the physical world?"

It would be a mistake, however, to think of the Ionians as materialists in the modern sense. As they conceived it, the universal material of things was far from being brute, inorganic, passive, mindless stuff intrinsically devoid of all higher properties or capabilities. Water, for instance, was, for them, not a sterile, inorganic chemical but a fertilizing fluid, and in their system it was quite open to consideration whether the basic stuff of the world might not be provided by either spirit (pneuma) or mind (nous). At this initial stage in philosophical speculation, indeed, the questions preoccupying philosophers cut across many of the distinctions that later generations were to treat as fundamental.

We first find these distinctions being drawn explicitly and insisted on by the Athenian philosophers, following the examples of Plato and Aristotle. For instance, Plato

and his fellow mathematicians at the Academy explained the properties of homogeneous material substances in one way, those of organized, functional systems in another. Like the Sicilian philosopher Empedocles, they classified material substances into four contrasted states or kinds—solid (earth), aeriform (air), liquid (water), and fiery (fire)—but they added a novel mathematical theory to account for the contrasted properties of these four kinds of substance. Each kind, they supposed, had atoms of a distinct geometrical shape, and they hypothetically identified these shapes with four of the five regular convex solids—tetrahedron, cube, octahedron, and icosahedron—whose mathematical properties had been studied by Plato's associate Theaetetus. (The fifth solid, the dodecahedron, they associated with the twelve constellations of the outer heavens.) The characteristic properties of organisms, on the other hand, they explained in functional rather than material terms. The form of any bodily organ must be accounted for as reflecting its role in the life of the organism; this form should be thought of as created specifically to perform a particular function as effectively as the available materials permitted.

Aristotle went further. He distinguished sharply between the material substance of which an object was composed and the form imposed on it, and he questioned whether the characteristic properties of any substance or system could be usefully explained in either atomistic or geometrical terms. In order to understand the properties and behavior of any individual object, it was first necessary to recognize it as an object of a particular kind. Each kind of object existing in nature had properties determined by its own special form or essence, so that any universal primary stuff (*hyle*) must be devoid of any particular distinguishing characteristic. For Aristotle and his followers the problem of distinguishing substances became primarily a matter of taxonomy, of qualitative classification, rather than a quantitative, physicochemical problem. Weight, from this point of view, was just one possible quality among others. Aristotle's views went beyond those of Plato in one other respect that was to have profound implications for cosmology. He drew a clear distinction between the sublunary world, whose objects were composed of the four terrestrial elements—earth, air, fire, and water—and could be created and destroyed, and the superlunary or celestial world of the outer heavens, whose inhabitants were composed of the quintessence (fifth essence) and exempted from change and decay. Of all terrestrial things only the souls of rational beings in any way shared this immutability.

LATER CLASSICAL AND MEDIEVAL PERIODS

Subsequent philosophers—whether in Hellenistic Alexandria (200 BCE–550 CE), the Islamic centers of learning (650–1150), or the newly founded universities of western Europe (950–1500)—introduced a number of variations into the debate about matter without adding any fundamentally new themes. For both the Stoics and the Epicureans, ideas about matter were closely associated with religious beliefs. Epicurus and his followers—notably, the Roman poet Lucretius—developed the more fragmentary speculations of Democritus and Leucippus about the atomic structure of matter into a complete philosophical system. But the atoms of the Greek philosophers differed from those of nineteenth-century European science in three crucial respects. First, they had an indefinitely large range of sizes and shapes instead of a limited number of fixed forms, one for each chemical "element." Next, they interacted only by direct contact or impact rather than by exerting forces of attraction or repulsion on one another. And, finally, they existed in special varieties—atoms of magnetism, of life, of mind, and of soul—to explain all sorts of activities—physical, biological, psychological, and even spiritual. The collisions and conjunctions of these atoms were regarded by Epicurus as an autonomous physical process, for his fundamental aim was to attack any belief in external interference by divine agencies in the affairs of the natural world.

The Stoics, such as Zeno of Citium and Chrysippus, rejected atoms in favor of three kinds of continuous physical medium or spirit (pneuma) for both scientific and religious purposes. The pneuma was an integrative agency, analogous to a field of force, capable of maintaining a stable pattern of properties and behavior in a physical system; in addition, it was capable of existing in separation from the solid and liquid frame of the "body" and could probably be identified with the soul. Instead of rejecting the traditional deities, like the Epicureans, the Stoics reinterpreted them as incorporeal agencies comparable to the pneuma. Yet though the Stoics and the Epicureans differed about many things, they agreed that every agency capable of producing physical effects—even the mind—must be regarded as a material body (soma). As a result for Lucretius pure mind was composed of very smooth and mobile atoms; for Chrysippus it consisted of undiluted fire.

The alchemical philosophers, for their part, introduced an experimental element into the study of matter. Beginning with the Democritean Bolos of Mendes (c. 200

BCE), going on through Maria the Jewess and Zozimos of Alexandria (second and third centuries CE), the alchemists exploited the traditional craft techniques of the Middle Eastern metallurgists, dyers, and jewelers and attempted to find ways of separating and isolating the essences or spirits in things. In this way they were led to contrast volatile and chemically active substances, such as alcohol and ether (spirits), with solid and passive ones, such as earths and *calces* (bodies). The association of the soul and the body in living creatures was thus treated as analogous to the association of volatile and gaseous with solid and earthy substances in a chemical compound. When freed from this association, incorporeal spirits naturally tended to rise toward the heavens and corporeal bodies to sink to the earth, a fact that apparently harmonized with the traditional Aristotelian contrast between the celestial and terrestrial worlds.

Nevertheless, philosophers and theologians in the strictly orthodox Aristotelian tradition rejected Stoic, Epicurean, and alchemical ideas as being excessively materialistic. In their view the soul was not in any way a subject for chemical or quasi-chemical speculation. The forms or essences of things were not themselves composed of any material stuff, even of the highly tenuous kinds conceived by the Stoics and alchemists. Accordingly, for Thomas Aquinas and the other philosophers of the high Middle Ages, the relation between matter and form was a problem in metaphysics or theology rather than one in natural philosophy.

NEW THEORIES: 1550–1750

Thus, the revival of the physical sciences during the Renaissance started from a position in which no single doctrine about the nature of matter was clearly established and generally accepted. All supporters of the new mechanical philosophy were attracted to an atomistic or corpuscular view of matter, but most of them took care to dissociate themselves from the original atomistic doctrines of Democritus and Epicurus, which were still suspected of having atheistical implications. Thus, Johannes Kepler explained the crystalline structure of snowflakes by reference to a geometrical theory of atoms modeled on that of Plato, Galileo Galilei embraced atomism as a physical embodiment for the points of geometry, and Descartes treated all matter as corpuscular in structure, at the same time denying the theoretical possibility of a void or vacuum. All of them regarded such mechanical interactions as collisions as the basic model for physical processes and sought to build up a theory of forces (dynamics) capable of explaining the established generalizations about the motions of physical objects.

However, attempts to work out an effective and comprehensive system of physical theory without going beyond the categories of atomism inherited from the Greeks encountered a number of difficulties. These sprang ultimately from the dual axiom that any agency capable of producing physical effects must be composed of a corresponding type of material object and that these objects could influence one another only by direct mechanical action, which required that the bodies be in contact. To deny the first half of this axiom implied accepting the notion of nonmaterial physical agencies; to deny the second implied accepting action at a distance. Both these notions were widely rejected as being incompatible with sound natural philosophy.

The immediate outcome of this dual axiom was to commit the advocates of the new mechanical corpuscular philosophy to a proliferation of new kinds of atom—for instance, magnetic, calorific, and frigorific corpuscles—introduced to account for the corresponding physical phenomena of magnetism, heat, cold, and so on. Although some philosophers, including Descartes, saw the possibility of cutting down the types of atoms—for example, by explaining heat as a consequence of the internal agitation of the material atoms composing hot bodies—even Descartes felt bound to accept that light, magnetism, and the like were carried by subtle fluids made up of corpuscles of insensible weight. Matter, he declared, came in three kinds, of which only "third matter" was subject to gravity and thus had any weight.

An indirect but even more profound outcome of the corpuscularian axiom was to support Descartes's fundamental division between mind and matter as absolutely distinct substances. The least plausible element in traditional atomism had been its psychology. Christian theology had added its own objections to any explanation of mental activity that regarded the mind as composed of atoms, no matter how light or mobile, for this, it was generally agreed, came perilously close to denying the immortality of the soul. The new physical science of the seventeenth and eighteenth centuries accordingly limited its aim. The realm of nature consisted of material bodies interacting mechanically by contact and impact and could be studied by science. The realm of spirit—including, at least, the intellectual activities of human beings—was a distinct and separate object of speculation to which the categories of physical science were not directly relevant. Much of the debate in subsequent epistemology can be traced to this point.

Accordingly, for two hundred years beginning around 1700, the concept of matter kept a central place in physical theory but was set aside as irrelevant to the study of mind. In physics the first major break with traditional ideas came through the work of Sir Isaac Newton. By his theories of dynamics and gravitation, Newton established a sharp distinction between material objects in a strict sense, whose mass conferred on them both inertia and weight, and forces, which were a measure of the way in which material objects interacted rather than a special kind of material thing. In the case of gravity, as he showed in his *Philosophiae Naturalis Principia Mathematica* (1687), these forces had to be supposed capable of acting over distances of many million miles, though Newton himself was inclined to believe that some invisible mechanical link existed by which the sun, for instance, exerted its gravitational action on the planets. In the later editions of his *Opticks* (especially those published after Gottfried Wilhelm Leibniz's death in 1716) he extended this idea to explain other physical phenomena. Electrical, magnetic, and chemical action also, he argued, might prove to be manifestations of forces of attraction and repulsion acting across the spaces between the massive corpuscles of bodies. Thus, the traditional system of atoms and the void was amended to become a theory of material corpuscles interacting by centrally directed forces.

CLASSICAL PHYSICS

Newton's program for natural philosophy made its way only slowly to begin with, but it met with no grave check until the late nineteenth century. At first, his insistence on mass as the essential property of matter was not found universally convincing. Others continued to regard extension, impenetrability, weight, or the capacity to produce physical effects as the indispensable criterion. As a result, throughout the eighteenth century there was an element of cross-purposes in debates about the corporeal nature of, for example, light and fire. Two developments particularly helped to clarify the intellectual situation and established the Newtonian categories as the basis of physical science. First, Antoine Lavoisier and his followers—notably, John Dalton—demonstrated that the phenomena of chemistry as well as those of physics could be unraveled on the assumption that all genuine material substances possessed mass and were composed of corpuscles or atoms. Second, the mathematical work of Leonhard Euler and his successors transformed Newton's account of forces of attraction and repulsion into the modern theory of fields of force.

After 1800, then, physical scientists went ahead rapidly with the experimental and mathematical work that culminated in the so-called classical physics and chemistry of the late nineteenth century. In this system the agents responsible for physical action were divided into two sharply contrasted categories. On the one hand, there was matter; this consisted of massive atoms that combined to form molecules in accordance with the principles of chemical combination. The mechanical energy associated with the motion of the molecules within any body accounted for its temperature; the fields of force between them explained gravitational, electric, and magnetic attraction and repulsion. On the other hand, there were those agencies—such as light and radiant heat—that apparently lacked both mass and weight and that were transmitted in the form of waves across the empty space between the material atoms. Gravitation apart, these various agencies turned out, as was shown by James Clerk Maxwell's electromagnetic theory of light, to be all of one general kind. By combining the established theories of the electrical and magnetic fields of force into a single mathematical system having the same degree of generality as Newton's dynamics, Maxwell demonstrated that electromagnetic waves would share the known properties of light and radiant heat and would move across space with the same velocity that had actually been measured in the case of light. This interpretation gained greatly in strength when Heinrich Hertz used an intermittent electrical spark to produce artificial electromagnetic waves, the so-called radio waves.

Though devoid of mass, these various forms of radiation nevertheless carried energy. Numerically, the sum total of all forms of energy in any isolated system (like the sum total of the masses of all the material bodies involved) was apparently conserved unchanged throughout all physical and chemical changes. As a result it seemed for several decades that the whole of natural philosophy could successfully be built on the central distinction between matter and energy and on the two independent axioms of the conservation of mass and the conservation of energy. Thus, Newton's program for physical science came close to being finally fulfilled in classical physics and chemistry.

TWENTIETH-CENTURY RECONSIDERATIONS

This intellectual equilibrium was short-lived. As Sir John Squire put it:

> Nature and all her Laws lay hid in Night.
> God said "Let Newton be, and all was Light."

> It could not last. The Devil, shouting "Ho!
> Let Einstein be," restored the *status quo.*

To do Einstein justice, the difficulties in the classical system that he resolved had been considered residual embarrassments for some time, and many of the conceptual changes for which he argued have since established themselves as indispensable features of physical theory. Still, they did undoubtedly have the effect of blurring the sharp distinctions and tidy certitudes of nineteenth-century science.

The effect of these conceptual changes on our concept of matter has been profound. Physicists have been compelled to reconsider and modify all the fundamental planks in the program enunciated for natural science by the mechanical philosophers of the seventeenth century. To begin with, Einstein displaced the seventeenth-century model of mechanical action as the universal pattern for intelligible physical processes by a new model based on electromagnetic theory. The embarrassments facing physicists in the 1890s arose, he showed, from a mathematical conflict between Maxwell's theory of electromagnetism and the mechanics of Galileo and Newton. Einstein circumvented these difficulties in his theory of relativity by giving priority to the theory of electromagnetic fields and by amending the principles of Newtonian mechanics to conform to the Maxwellian pattern. As a result the attitudes of a representative late nineteenth-century physicist, such as William Thomson, Lord Kelvin (who declined to accept Maxwell's theories, declaring that he could embrace a physical explanation of a phenomenon wholeheartedly only if he could make a mechanical model to demonstrate it), have since come to seem excessively narrow.

As a result of this initial change, however, certain other fundamental elements in classical physics have had to be called in question. The absolute distinction between matter and energy, for instance, has gone by the board. It now appears that any quantity of energy (E) is in certain respects equivalent to a proportional quantity of mass ($m = E/c^2$, where c is Maxwell's constant, equal to the measured velocity of electromagnetic radiation); that for theoretical purposes the twin conservation principles of nineteenth-century physics and chemistry should be joined in a single axiom, according to which the sum total of energy and mass (combined according to the formula $E + mc^2$) was conserved in all physical processes; and that in appropriate circumstances a quantity of electromagnetic energy can be transformed into the corresponding quantity of matter or vice versa. This implication was confirmed in the 1930s from a detailed study of individual actions between atomic nuclei and other particles, and it was dramatically reinforced by the explosion of the first atomic bombs, whose energy was derived from the marginal loss of mass involved in the nuclear fission of such heavy elements as uranium.

Meanwhile, the earlier contrast between matter, which was assumed to exist in discrete atomic units, and radiation, which traveled in the form of continuous waves, was under criticism for quite different reasons. First, Max Planck showed that bodies exchanged light-energy in the form of bundles or wave-packets. Einstein, going further, argued that electromagnetic energy always existed in the form of these photons. Then, in the early 1920s, Louis de Broglie put forward the idea that the subatomic particles into which Niels Bohr and Ernest Rutherford had analyzed the fundamental material units of earlier chemistry might themselves manifest some of the properties of wave-packets. This was confirmed in 1927, when it was shown that a beam of electrons passed through a crystal lattice produced a diffraction pattern just as a beam of light of the corresponding wavelength and velocity would have done. By the 1960s it began to appear that matter-particles might differ from the energy-packets of light or other kinds of radiation only in having part of their energy frozen in the form of inertial mass.

Finally, the theory of quantum mechanics, first formulated between 1926 and 1932 by Werner Heisenberg, Erwin Schrödinger, and P. A. M. Dirac, has radically undercut one last presupposition, which had underlain physical science since the time of Galileo. From 1600 on, the fundamental units of matter—whether called corpuscles, particles, or atoms—had been regarded as intrinsically brute, inert, and passive. They might be constituted in such a way that they are capable of exerting forces on one another by virtue of their relative motions and positions, but one had to seek the ultimate source of this capacity—as of their motion—in God who created them. (This was one point on which Newton, Descartes, and Maxwell all agreed.) Since 1926 the final unit of analysis in physics has ceased to bear any serious resemblance to these inert corpuscles. Instead, the quantum physicists begin with certain wave functions or eigenfunctions, which characterize the activity of, say, an electron or an atom as much as they do its structure and position. Just as mass has ceased to be entirely distinct from energy, so the particles of Newton's physics have ceased to be absolutely distinct from the forces of attraction and repulsion acting between them. On the contrary, according to the principles of contemporary physical theory,

every kind of fundamental particle—whether of matter or energy—should be associated with a corresponding mode of interaction and force field. Photons, electrons, mesons, nucleons—all these have a dual aspect, being characterized partly by their inertial mass or intrinsic energy and partly by their pattern of interaction with the environment. One outstanding and at present unsettled question is whether the transmission of gravitational forces, from which the whole notion of a field began, also involves the propagation of particles ("gravitons") at a finite speed. If it proves that "gravitons" do in fact exist and travel at the same speed as photons, this will tie up one of the more notorious loose ends of mid-twentieth-century physics.

IMPLICATIONS OF NEW THEORIES

Today almost all the axioms of earlier natural philosophy have been qualified, if not abandoned. Mass has ceased to be the essential, unalterable characteristic of all physical objects and now appears to be one variant of the wider category of energy. No longer can any determinate amount of this energy be localized with absolute precision (Heisenberg's principle), and we are left with a picture of a natural world whose fundamental elements are not so much passive bricks as units of activity. This transformation—as Samuel Sambursky has argued—involves a reaction against the axioms of seventeenth-century physics as radical as the Stoics' rejection of the atomism of Epicurus. Indeed, Sambursky points out, there is a strong parallel between the two reactions. As in the Stoic theory, physicists today also consider matter essentially active rather than passive and explain its behavior as the outcome of patterns of energy and excitation associated with any given state or condition.

The full implications of this change for our other ideas are beginning to become apparent only now. In biology, at any rate, a considerable change has come about since 1950 by the extension of physical theories about molecular structure into the fields of genetics, embryology, and bacteriology. Here the intimate association of structure and function characteristic of modern subatomic theory is reproduced in the association of specific biological activities with particular configurations (and, thus, eigenfunctions) of the complex molecules involved. The extensions of the new ideas about matter into the theory of organic development and human behavior are still at a speculative stage.

This much can, however, be said. During the centuries that have elapsed since the revival of natural philosophy at the Renaissance, the concept of matter has changed its character quite fundamentally. In the present state of scientific thought, accordingly, all earlier questions about, for instance, the relation of matter, life, and mind need to be entirely reconsidered. When, for instance, Descartes classified matter and mind as distinct substances, he was putting the concept of mind and mental activities in opposition to a concept of matter as inert extension, a concept that is now discredited. To that extent the extreme dualism of Descartes's philosophy has been not so much refuted by later science as made irrelevant; its categories no longer fit our situation.

Similarly, other long-standing debates concerning, for example, the reality of the material world or the relation between material objects and our sensations will need to be reappraised in the light of changes in our concept of matter. But this is a task for the future.

See also Anaximander; Aristotle; Atomism; Bohr, Niels; Chrysippus; Descartes, René; Dynamism; Empedocles; Energy; Epicurus; Einstein, Albert; Ether; Galileo Galilei; Heisenberg, Werner; Hertz, Heinrich Rudolf; Kepler, Johannes; Lavoisier, Antoine; Leucippus and Democritus; Mass; Maxwell, James Clerk; Newton, Isaac; Plato; Renaissance; Schrödinger, Erwin; Thales of Miletus; Thomas Aquinas, St.; Zeno of Citium.

Bibliography

In general, this article follows the argument of Stephen Toulmin and June Goodfield, *The Architecture of Matter* (London: Harper and Row, 1962), in which the development of the concept of matter is fully analyzed but discussed without serious technicalities. For the various periods covered here the reader is referred to the following works.

GREEK PHILOSOPHY

S. Sambursky, *The Physical World of the Greeks* (London: Routledge and Paul, 1956), is an outstanding survey for the general reader. W. K. C. Guthrie, *A History of Greek Philosophy,* Vol. 1 (Cambridge, U.K.: Cambridge University Press, 1962), and G. S. Kirk and J. E. Raven, *The Presocratic Philosophers* (Cambridge, U.K.: Cambridge University Press, 1957), are up-to-date scholarly discussions of the Ionian natural philosophers. F. M. Cornford, *Plato's Cosmology* (London: K. Paul, Trench, Trubner, 1937), is the most convenient existing version of the *Timaeus,* in which Plato's views about matter are expounded. J. H. Randall Jr., *Aristotle* (New York: Columbia University Press, 1960), provides an illuminating account of that philosopher's scientific ideas; it is useful for the nonspecialist.

LATER CLASSICAL AND MEDIEVAL PERIOD

S. Sambursky's *The Physics of the Stoics* (London: Routledge and Paul, 1959) and *The Physical World of Late Antiquity* (London, 1962) complete the story begun in his *Physical World of the Greeks* (see above). Cyril Bailey, *The Greek*

Atomists and Epicurus (Oxford, 1928), and A. J. Hopkins, *Alchemy, Child of Greek Philosophy* (New York: Columbia University Press, 1934), are scholarly but readable: Both books remain stimulating and full of interest. E. J. Holmyard, *Alchemy* (London, 1957), and A. C. Crombie, *Medieval and Early Modern Science* (Garden City, NY: Doubleday, 1959), are readable popular surveys.

NEW THEORIES: 1550–1750

H. T. Pledge, *Science since 1500* (London: H. M. Stationery Office, 1939; reprinted, New York: Harper, 1959), and A. R. Hall, *From Galileo to Newton* (London, 1963), are general histories, both of which include useful material on the new theories. Mary B. Hesse, *Forces and Fields* (Edinburgh, 1961); Marie Boas, *Robert Boyle and Seventeenth Century Chemistry* (Cambridge, U.K.: Cambridge University Press, 1958); Hélène Metzger, *Les doctrines chimiques* (Paris, 1923) and *Newton, Stahl, Boerhaave* (Paris: F. Alcan, 1930); I. Bernard Cohen. *Franklin and Newton* (Philadelphia: American Philosophical Society, 1956); and E. J. Dijksterhuis, *The Mechanization of the World Picture,* translated by C. Dikshoorn (Oxford: Clarendon Press, 1961), are scholarly books dealing in a penetrating way with more detailed aspects of the subject.

CLASSICAL PHYSICS

Edmund Whittaker, *History of the Theories of Aether and Electricity,* 2 vols. (Edinburgh, 1951–1953), and Mary B. Hesse, *Forces and Fields* (see above), are the best specialist surveys. For the general reader Charles C. Gillispie, *The Edge of Objectivity* (Princeton, NJ: Princeton University Press, 1960), N. R. Campbell, *What Is Science?* (London: Methuen, 1921; reprinted, New York: Dover, 1952), Albert Einstein and Leopold Infeld, *The Evolution of Physics* (New York: Simon and Schuster, 1938), and George Gamow, *Biography of Physics* (New York, 1963), may be selected from many others as being particularly useful.

TWENTIETH-CENTURY RECONSIDERATIONS

A great many books of general interest have been published about the twentieth-century transformation in physical theory. Apart from Einstein and Infeld, op. cit., and Gamow, op. cit., one of especial merit is Banesh Hoffmann, *The Strange Story of the Quantum* (New York: Harper, 1947). Many of the physicists directly involved have written interestingly about the changes—notably, Werner Heisenberg, *Philosophical Problems of Nuclear Science* (London: Faber, 1952). The analogy between Stoic matter theory and wave mechanics is pursued in Sambursky, *The Physics of the Stoics* (see above).

Stephen E. Toulmin (1967)

MATTER AND PROBLEMS OF PERCEPTION

See *Appearance and Reality; Illusions; Perception; Phenomenalism; Primary and Secondary Qualities; Realism; Sensa*

MATTHEW OF ACQUASPARTA

(c. 1237–1302)

Matthew of Acquasparta, the Italian Franciscan scholastic philosopher and theologian, was born in Acquasparta, near Todi in Umbria, possibly of the illustrious Bentivenghi family. In 1254 he entered the Franciscan order, and about 1268 he began studies at the University of Paris, where he was profoundly influenced by Bonaventure's system. Matthew was lector in the Studium Generale at Bologna (at least for the year 1273–1274), and in 1276 he became master in theology at Paris. From 1279 to 1287, he was lector Sacri Palatii in Rome, succeeding John Peckham. He was general of the order from 1287 to 1289. In 1288 he was made cardinal, and in 1291 he was named bishop of Porto and Santa Rufina. Matthew died at Rome, where he is buried in the church of Ara Coeli.

DOCTRINE

Matthew taught and wrote during the time of conflict between the Augustinian–Franciscan doctrinal tradition and the rising Thomistic Aristotelianism. In this far-reaching controversy he proved himself to be exceptionally well-versed in Augustine's doctrines and in general a faithful follower of Bonaventure. Although he incorporated a few Aristotelian elements, Matthew's system in its entirety shows that he was among the purest adherents of Augustinianism in the last quarter of the thirteenth century. He had a calm, balanced mind, a sober style, and an exact manner of formulating his ideas. In discussion he was generally modest and perceptive. With these qualities he often achieved, at least in his *Quaestiones Disputatae de Fide et de Cognitione,* a level comparable to that of the greatest thinkers of his age.

In his theory of knowledge Matthew taught that our intellect knows the individual object not only by reflection, as St. Thomas Aquinas held, but also by a direct perception, which precedes the formation of an abstract idea. By virtue of this perception, the intellect forms a *species singularis* of the concrete object with all the richness of detail it possesses in reality. In this way the mind prepares for knowledge of the essence of the object. Similarly, the soul knows its own existence and habits not only by reasoning and by reflection but also by a direct and intimate intuition. In *Quaestiones Disputatae de Cognitione,* Matthew presented a personal solution to the controversial question of the activity of the knowing subject. Rejecting the impressionism of Bonaventure and Thomas Aquinas, the innatism of Thomas of York and

Roger Bacon, and the pure activism of William of Auvergne and John Peckham, Matthew defended a semi-activism, not an occasionalism. Whereas according to pure activism the *species intentionalis* is completely (matter and form) caused by the knowing subject, according to Matthew the matter comes from the object, the form from the subject. This opinion, however, was soon contested by Roger Marston as contradicting both Aristotle and Augustine.

Matthew defended the theory of divine illumination almost in the same manner as did Bonaventure. The purely human faculties for knowing the extramental world do not give us either clear understanding or certainty. We need the aid of the divine *rationes aeternae* (divine ideas) to illuminate our mind during the process of knowledge. God is not simply the creator of human intelligence; he also conserves it and concurs in each of its actions. This collaboration of God by means of the divine illumination is possible because man in his mind bears a special likeness to his creator. Our intellect is illumed by the divine light that contains the eternal ideas and is the ground of all created beings. The divine light is not the object itself of our knowledge but the moving principle that leads us to the true knowledge of the created world. Following the Augustinian doctrine, Matthew believed that the object of knowledge never determines the election of the will.

Among Matthew's other philosophical theses, the following are worthy of mention. Matthew, like Bonaventure, rejected the possibility of a creation from eternity; the spiritual beings (souls and angels) are necessarily composed of matter and form, because if they were composed simply of essence and existence (as Thomas Aquinas taught), this would not account for their contingency. Also, the process of coming to existence must be explained by the Augustinian theory of the *rationes seminales.* The "being body" (*esse corporale*) constitutes a plurality of forms. The two elements of the beings, matter and form, are together the cause of individuality. Matthew upheld the Ontological Proof of the existence of God; he also argued that the knowledge of God that we attain through faith is compatible with scientific knowledge. Matthew was particularly interested in problems concerning the relations between the natural order and the supernatural order.

IMPORTANCE

Matthew is undoubtedly to be ranked among the great scholastic thinkers. His importance, however, lies not so much in the originality of his thought as in the fact that he is, after Bonaventure, the ideal representative of Augustinianism. The only philosophers that are known to have been directly influenced by him are Roger Marston and Vitalis of Furno.

See also Aristotelianism; Aristotle; Augustine, St.; Augustinianism; Bacon, Roger; Bonaventure, St.; Marston, Roger; Medieval Philosophy; Ontological Argument for the Existence of God; Peckham, John; Thomas Aquinas, St.; Thomas of York; William of Auvergne.

Bibliography

PRIMARY SOURCES

Matthew's most important philosophical works are *Introitus ad S. Scripturam* (1268–1269), in *Bibliotheca Franciscana Scholastica Medii Aevi,* Vol. I (Quaracchi, 1903; 2nd ed., 1957), pp. 3–21; *Introitus ad S. Theologiam* (probably 1271–1272), ibid., pp. 22–33; Commentarius in I, II, et III Sententiarum (1271–1272), unedited manuscript at Quaracchi; *Quaestiones Disputatae* (1267–1287), which is almost complete in *Bibliotheca Franciscana Scholastica Medii Aevi,* Vols. I–II, XI, XVII–XVIII (Quaracchi, 1903–1961); *Quaestiones de Anima VI,* in *Archives d'histoire doctrinale et littéraire du moyen âge,* edited by A.-J. Gondras, Vol. XXIV (Paris, 1958), pp. 203–352; *Quaestiones Disputatae de Anima XIII,* edited by A.-J. Gondras, *Études de philosophie médiévale* 50 (1961); VI Quodlibeta, about 90 questions (1276–1279), unedited manuscript at Quaracchi; Concordantiae Super IV Libros Sententiarum, unedited manuscript at Quaracchi.

SECONDARY SOURCES

Beha, H. M. "Matthew of Acquasparta's Theory of Cognition." *Franciscan Studies* 20 (1960): 161–204; 21 (1961): 1–79, 383–465.

Bettoni, E. "Rapporti dottrinali fra M. d'Acquasparta e G. Duns Scoto." *Studi francescani* 15 (1943): 113–130.

Bonafede, G. "Il problema del 'lumen' nel pensiero di Frate M. d'Acquasparta." *Rivista rosminiana* 31 (1937): 186–200.

Doucet, V. *Matthaei ab Aquasparta, Quaestiones Disputatae de Gratia.* Quaracchi, 1935. See pp. 11*–163* for a general introduction.

Pacchierini, L. *La dottrina gnoselogica de M. d'Acquasparta.* Naples, 1949.

Pegis, A.-C. "Matthew of Aquasparta and the Cognition of Non-being." In *Scholastica Ratione Historico-Critica Instauranda,* 463–480. Rome: Pontificum Athenaeum Antonianum, 1951.

Simoncioli, F. "Il concetto di legge in M. d'Acquasparta." *Studi francescani* 56 (1959): 37–50.

A. Emmen, O.F.M. (1967)

MAUPERTUIS, PIERRE-LOUIS MOREAU DE

(1698–1759)

Pierre-Louis Moreau de Maupertuis, the French scientist and philosopher, was born in Saint-Malo, Brittany. Elected in 1723 to the Académie des Sciences (and to the Royal Society in 1728), he first became known for his work in geometry. The expedition that he led to Lapland in 1736 to measure a degree of meridian near the pole helped finally to prove that Earth was an oblate spheroid. With his early introduction of Newtonian theories into France, Maupertuis became a leading exponent among the *philosophes* of the ideal of experimentalism as opposed to the overly deductive method in science associated with the Cartesian tradition. In 1744 Frederick II of Prussia asked him to reorganize the Berlin Academy of Sciences and later appointed him as its president (1746–1759). The remainder of his career was intimately linked to the activities of this group, and the growth of the academy into an important center of research owed much to his efforts.

PRINCIPLE OF LEAST ACTION

Maupertuis's famous principle of least action, which contributed signally to the systematization of mechanics, was formulated in "Recherche des loix du mouvement" (1746) as follows: "Whenever any change occurs in nature, the quantity of action employed for this is always the smallest possible"—the "quantity of action" being proportional to the product of the mass of a body and its velocity and the distance traversed. Among the heated controversies provoked by this notion, Samuel Koenig's unfair (although understandable) attribution of it to Gottfried Wilhelm Leibniz brought about a scandalous quarrel and lifelong enmity between Maupertuis and Voltaire. But all this proved irrelevant to the historic value of the principle of least action, which, clarified progressively by the applications it found in the works of Leonhard Euler, Joseph Lagrange, William Hamilton, Hermann Ludwig von Helmholtz, and others, emerged ultimately as a basic concept in the mathematical analysis of dynamic systems.

COSMOLOGICAL ARGUMENT. In the *Essai de cosmologie* (1750), Maupertuis's extension of the principle of least action to the much debated problems of theodicy offered a compromise solution between the radical antifinalism of contemporary materialists and the naive finalism of those who saw God's wisdom in every manifestation of design in nature, however trivial or self-contradictory. By claiming that an actual mathematical equation showed God's regulation of nature through the parsimony of kinetic means employed in the production of all physical events, Maupertuis succeeded in giving an original and seemingly scientific version of the Cosmological Argument. But his assumption that there is logical necessity as such in the existence of mechanical laws, which was consistent with the example of René Descartes and Leibniz, typified a rationalist attitude that, though prevalent at the time, was already undermined by those who, like David Hume, alleged a merely empirical necessity for physical causation. Although Maupertuis's distrust of metaphysical reasoning led him to present his cosmological argument not as demonstrably certain, but only as the best that the imperfect human intellect was capable of, it remained perhaps less plausible than ingenious, particularly since it was affirmed without sufficient regard either to the epistemological difficulties it incurred or to the possible nontheological interpretations of its underlying minimal concept. Coming late in a current of thought that was to yield before long to new orientations in philosophy, the *Essai de cosmologie* had a limited historical impact. It was, in fact, in a form essentially free of teleological meanings that the principle of least action exercised its considerable influence on the development of physicomathematical science.

BIOLOGY: THE STRUCTURE OF MATTER

A different science, biology, inspired Maupertuis's next major work (1751), the *Dissertatio Inauguralis Metaphysica de Universali Naturae Systemate* (known also as the *Système de la nature*). Study of the problem of heredity had led Maupertuis to reject, in the *Vénus physique* (1745), the then reigning doctrine of preformation and to favor instead a theory of epigenesis using the law of attraction. But he had subsequently found this theory inadequate and had despaired altogether of accounting mechanistically for the origins and nature of life. In the *Dissertatio Inauguralis*, therefore, he sought to explain the formation of living things by supposing that all the elementary particles of matter are individually endowed in a proportionately elementary degree with "desire, aversion, and memory," by virtue of which they combine to form organic entities.

Such a notion, no less than that of least action, betrays a marked Leibnizian background in Maupertuis's thinking, despite his outspoken criticism of the metaphysics of Leibniz. It is true, nevertheless, that Mauper-

tuis did not assign the metaphysical status of the monads to his "percipient particles" but, rather, presented them as part of a general biological hypothesis; he accounted for the elemental coexistence of physical and psychic properties in nature by reference to a common unknowable substance. Thus, the philosophical basis of his biological theorizing may be described as either an "atomistic dualism" or a "corpuscular psychism," sustained by a phenomenological accord between matter and its presumed psychic qualities. These ideas were misinterpreted in materialistic terms by Denis Diderot and contributed indirectly to the eventual success of naturalism in biology. Since Maupertuis's metabiological conception was also intended to explain the structural transformations of the various species by a process of genetic mutation, it merged, in that respect too, with an important current of evolutionist speculation that grew in France after about 1750.

EPISTEMOLOGY

The views of Maupertuis in epistemology can be judged from a number of his writings. While, like Étienne Bonnot de Condillac and most of the *philosophes*, he agreed with John Locke that sensation is the source of all our knowledge, his position was appreciably more sophisticated, probably because of his encounter with the Berkeleian critique. If this critique did not quite win him over to subjectivism, he at least became convinced that experience offers no more than the disjointed fragments of a merely phenomenal reality and that the substance presumed to excite in the mind the perceptions that in turn are projected cognitively toward the natural world remains itself beyond objective determination. Maupertuis ascribed even the evidence of mathematics not to any intrinsic veracity of such knowledge but to the fact that it is based on the repetition (*réplicabilité*) of certain simple ideas that consist of identical units and are abstracted from the heterogeneous totality of sensory impressions. In the same spirit, his *Réflexions philosophiques sur l'origine des langues et la signification des mots* (1748) raises the equally crucial question of the linguistic prefigurations of sense experience, from which scientific reasoning is unable completely to escape.

ETHICS

Maupertuis's principal excursion into ethics, *Essai de philosophie morale* (1749), tried somewhat overambitiously to reconcile the Stoic, Epicurean, and Christian schools but succeeded only in reaching an eclectic view characterized by the author's own pessimism concerning the chances of human felicity. It offered, however, an early instance of the application of arithmetic to the problem of happiness by its attempt to express, in the analogy of statics, the equations of a "hedonistic calculus."

IMPORTANCE

Generally, the thought of Maupertuis pursued the aim, shared by many of his contemporaries, of linking philosophy more concretely than in the past with the content of the particular sciences. Instead of presenting an overall logical coherence, his work contributes various philosophical essays reflecting the different points of departure dictated by his primarily scientific interests. The cosmological thesis, speculative biology, and moral opinions of Maupertuis remained largely separate from each other; moreover, Maupertuis himself was often in the curious but historically symptomatic predicament of searching earnestly for metaphysical solutions while disbelieving in their possibility. Having elaborated the principle of least action and the notion of percipient particles of matter in a rather ambiguous zone between metaphysics proper and scientific theory, it is not surprising that he should have suffered much unmerited neglect from historians both of philosophy and of science. But it is now recognized that Maupertuis had a significant, even if secondary, role in the maturing of modern physics and biology alike, as well as in the transition of philosophical thinking from classical metaphysics to the critical position adopted by Immanuel Kant.

See also Condillac, Étienne Bonnot de; Cosmological Argument for the Existence of God; Descartes, René; Geometry; Hamilton, William; Helmholtz, Hermann Ludwig von; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Meier, Georg Friedrich; Pessimism and Optimism; Scientific Method; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY MAUPERTUIS

Lettres. 1752. Place of publication unknown.

Examen philosophique de la preuve de l'existence de Dieu. In *Memoirs of the Berlin Academy of Sciences*. 1756.

Oeuvres. 4 vols. Lyons, 1756.

WORKS ON MAUPERTUIS

Abelé, Jean. "Introduction à la notion d'action et au principe de l'action stationnaire." *Revue des questions scientifiques* 119 (1948): 25–42.

Bachelard, Suzanne. *Les polémiques concernant le principe de moindre action au XVIII^{e} siècle*. Paris: Université de Paris, Palais de la Découverte, 1961.
Beeson, David. *Maupertuis: An Intellectual Biography*. Oxford: Voltaire Foundation, 1992.
Brunet, Pierre. *Étude historique sur le principe de la moindre action*. Paris, 1938.
Brunet, Pierre. *Maupertuis, étude biographique* and *Maupertuis: l'oeuvre et sa place dans la pensée scientifique et philosophique de XVIII^{e} siècle*. 2 vols. Paris: Blanchard, 1929. An authoritative study of Maupertuis's life and thought.
Crombie, A. C. "Maupertuis, précurseur du transformisme." *Revue de synthèse* 78 (1957): 35–56.
Feher, Marta. "The Role of Metaphor and Analogy in the Birth of the Principle of Least Action of Maupertuis (1698–1759)." *International Studies in the Philosophy of Science* 2 (1988): 175–188.
Gossman, L. "Berkeley, Hume and Maupertuis." *French Studies* 14 (1960): 304–324.
Guéroult, Martial. "Note sur le principe de la moindre action chez Maupertuis." In *Dynamique et métaphysique leibniziennes*, 215–235. Strasbourg, 1934.
Maglo, Koffi. "The Reception of Newton's Gravitational Theory by Huygens, Varignon, and Maupertuis: How Normal Science May Be Revolutionary." *Perspectives on Science* 11(2) (2003): 135–169.
Ostoya, Paul. "Maupertuis et la biologie." *Revue d'histoire des sciences* 7 (1954): 60–78.
Terrall, Mary. *The Man Who Flattened the Earth: Maupertuis and the Sciences in the Enlightenment*. Chicago: The University of Chicago Press, 2002.
Vartanian, Aram. "Diderot and Maupertuis." *Revue Internationale De Philosophie* 38 (1984): 46–66.

Aram Vartanian (1967)
Bibliography updated by Tamra Frei (2005)

MAXWELL, JAMES CLERK
(1831–1879)

James Clerk Maxwell, the British physicist, came from a well-known Scottish family, the Clerks; his father adopted the name Maxwell on inheriting an estate originally belonging to that family. Maxwell was educated at Edinburgh University and the University of Cambridge, becoming a fellow of Trinity College in 1855. In 1856 he won the Adams Prize at Cambridge for an essay in which he demonstrated that the rings of Saturn would be unstable if they were continuously solid or fluid and that they must be composed of discrete and separated parts. Maxwell was professor of natural philosophy at Marischal College in Aberdeen from 1856 to 1860 and professor of natural philosophy and astronomy at King's College in London from 1860 to 1865. His first paper on electromagnetism appeared in 1856; his electromagnetic field theory with the derivation of the velocity of light was first published in 1861–1862 and in more rigorous form in 1865; and he began work on the kinetic theory of gases in 1860. From 1865 to 1871 Maxwell remained at his country estate in Scotland where he worked on his *Treatise on Electricity and Magnetism,* which summarized the subject and his contributions thereto. In 1871 he became the first occupant of the Cavendish chair of experimental physics at Cambridge, supervised the construction of the Cavendish Laboratory, and later guided the first research done there. During this period he edited the works of Henry Cavendish. During his lifetime Maxwell also did research on color vision, mechanics, and other topics, and although his fame rests on his theoretical achievements, his experimental work was noteworthy.

THE ELECTROMAGNETIC FIELD

Maxwell's greatest contribution to fundamental physics was his concept of the electromagnetic field, a concept that underwent much modification both in the course of his own researches and at the hands of his successors. In modern terms, a field—such as the electric field—is a condition in the space surrounding charged bodies that determines the force that a unit electric charge would experience if it were placed at any point. In field theory all actions are regarded as transmitted from point to point by the contiguous modification of the field between the points, and the field is regarded as the seat of energy. Contemporary physics is dominated by the field-theoretic viewpoint, whether or not it is reinterpreted in terms of quantum theory.

Maxwell aimed at embodying in mathematical notation the ideas of Michael Faraday and, in particular, Faraday's fruitful concept of lines of force. In this Maxwell was inspired by the work of William Thomson (later Lord Kelvin), who had demonstrated the mathematical analogy between the problems of heat flow and of the distribution of static electricity. Maxwell developed similar analogies in his first paper on the subject, "On Faraday's Lines of Force" (1855–1856), drawing separate analogies for different aspects of electromagnetism: between electrical and fluid currents, and between electric or magnetic lines of force and fluid currents. While suggestive, such an endeavor was of course not a unified theory. "I do not think," he wrote, "that we have any right at present to understand the action of electricity, and I hold that the chief merit of a temporary theory is, that it shall guide experiment, without impeding the progress of the true theory when it appears." The beginning of the paper is of interest as a statement of method; Maxwell points out the pitfalls of commitment to a mathematical formula, in

which case "we entirely lose sight of the phenomena to be explained," or to a physical hypothesis, the irrelevant parts of which are liable to carry one beyond the truth. He advocates instead the use of physical analogy, "that partial similarity between the laws of one science and those of another which makes each of them illustrate the other."

In his "On Physical Lines of Force" (1861–1862), Maxwell's electromagnetic field theory appears for the first time, presented as a deduction from a detailed model of the ether. Magnetic lines of force are represented as molecular (microscopic) vortices in this ether, the matter of the ether whirling around in planes normal to the direction of the lines of force, so that the latter is the direction of the axes of the vortices. Maxwell found that in this fashion he could represent the properties of lines of force needed for magnetostatics, that is, that the lines should tend to contract along their length and repel each other laterally. But how can neighboring vortices spin in the same sense, since their neighboring boundaries move in opposite directions, and how are these motions initiated and communicated through the ether? Maxwell assumed a layer of tiny idle wheels between each pair of vortex cells in the ethereal substance. These wheels can rotate freely, so that a uniform magnetic field is represented by the vortex cells all spinning at the same rate and in the same sense, and the interspersed wheels rotating in place in the opposite sense. The idle wheels can also move from place to place in a conductor, but they are constrained to rolling contact without slipping with the neighboring vortices. The translatory motion of the wheels is identified with the electric current and used to explain the manner in which a magnetic field is created by an electric current (Hans Christian Ørsted's discovery); it also is used to account for electromagnetic induction. Furthermore, in a dielectric, including the vacuum, the wheels are not free to move in translation, but can only be displaced slightly against the elastic forces of the material of the cells. This action of displacement is the displacement current that forms the new term Maxwell added to previous results, while transforming all of them into his theoretical language. Maxwell then proceeded to calculate the velocity of propagation of transverse waves in his elastic ether. The speed of these waves was proportional to the ratio between the electromagnetic and electrostatic units of charge.

The factor of proportionality between the speed of the waves and the ratio of the units depended in this calculation on the specific model chosen for the ether; the argument showing the two terms to be equal cannot be regarded as very satisfactory. In "A Dynamical Model of the Electromagnetic Field" (1865), the electromagnetic field equations are presented directly without recourse to the ether model, and the relation between velocity of waves and ratio of electrical units is derived directly from the equations. Since, according to Wilhelm Weber and Friedrich Kohlrausch (1857), the ratio between the units was 3.11×10^8 meters/sec., whereas, according to Armand Fizeau, the speed of light was 3.15×10^8 meters/sec., Maxwell drew the important conclusion that light consisted of waves in the electromagnetic ether. This finally gained general acceptance when Heinrich Hertz generated electromagnetic waves by electrical means and showed that they had all the properties of light except that they were of much lower frequency, a result of the conditions of generation.

In his later papers Maxwell no longer relied on specific models of the ether. In the *Treatise* he wrote:

> The attempt which I then [in "On Physical Lines of Force"] made to imagine a working model of this mechanism must be taken for no more than it really is, a demonstration that mechanism may be imagined capable of producing a connexion mechanically equivalent to the actual connexion of the parts of the electromagnetic field. The problem of determining the mechanism required to establish a given species of connexion between the motions of the parts of a system always admits of an infinite number of solutions.

Nevertheless, he still regarded the underlying phenomena as motions and stresses in the mechanical ether, maintaining that the energy of magnetism "exists in the form of some kind of motion of the matter in every portion of space," apparently of a vortical character. Maxwell's views differ from those of the twentieth century in the following ways: The electromagnetic field was not regarded as a separate dynamic entity from matter, that is, a material ether; ordinary matter was treated macroscopically, phenomenologically, rather than from the atomic point of view; and the role of charge in the theory was ambiguous. Late in the nineteenth century H. A. Lorentz combined Maxwell's field theory with Continental conceptions of atomicity of charge to establish the classical theory of the dualism of matter and field.

KINETIC THEORY OF GASES

Also of fundamental importance was Maxwell's work on the kinetic theory of gases. In deriving the experimental gas laws, previous investigators had made the simplified

assumption that all the gas molecules moved with the same speed. In "Illustrations of the Dynamical Theory of Gases" (1860), Maxwell first derived the equilibrium distribution of the velocities of the molecules: the components of the velocity along a given direction are distributed according to Carl Friedrich Gauss's error law. This paper also contained the startling result, later demonstrated experimentally, that the viscosity (internal friction) of a gas should be independent of its density. Maxwell wrote two other pathfinding papers on the kinetic theory; their main subject was the derivation of the transport coefficients of a gas (coefficients of diffusion, viscosity, and thermal conductivity) and, in the last of them, the discussion of radiometric phenomena.

Maxwell's work on the kinetic theory may be regarded as constituting the first important introduction of statistical reasoning into physics and the first steps in the development of statistical mechanics, later continued by Ludwig Boltzmann and Josiah Gibbs. In statistical mechanics the use of statistics is not a manifestation of any indeterminism in the purported fundamental laws of nature, as it is in quantum physics; rather it is the reflection of our ignorance of the exact motions of the enormous number of molecules in any macroscopic system. The very immensity of this number (there are about 6×10^{23} hydrogen atoms in one gram of hydrogen) and the minuteness of the individual molecules give assurance that in ordinary experiments the measurable properties will be statistical in character and thus will be exactly the properties singled out by a statistical theory.

Maxwell's demon, a hypothetical being that apparently could reverse the tendency of isolated systems toward increase of disorder or entropy and so would violate the second law of thermodynamics, appears in his *Theory of Heat* (London, 1872, pp. 308–309). The thermal equilibration of neighboring vessels containing gas, representing a state of maximum disorder, could be destroyed by a being capable of seeing the individual molecules of the gas who acts so as to let only the faster molecules in one container pass through a small hole into the other, and the slower ones in the latter to pass in the reverse sense. Since the temperature is determined by the mean energy of motion of the molecules, this process would result in the gas in one vessel becoming warmer than that in the other, without any interference from outside the system. The demon has been exorcised by L. Brillouin and others (see Brillouin's *Science and Information Theory,* New York, 1956, Ch. 13). To obtain the information about an approaching molecule that the demon needs in order to decide whether or not to open the hole, the demon must absorb at least one quantum of light, the energy of which is reasonably greater than the mean energy of the quanta of thermal radiation that are always present. The absorption of this quantum demonstrably leads to a greater increase in entropy in the total system (including the demon) than the decrease obtained by properly manipulating the hole.

See also Boltzmann, Ludwig; Energy; Ether; Faraday, Michael; Gibbs, Josiah; Matter; Motion; Philosophy of Physics; Quantum Mechanics.

Bibliography

The *Scientific Papers of James Clerk Maxwell,* including his semipopular lectures but not the *Treatise* and other books, appear in two volumes edited by W. D. Niven (Cambridge, U.K.: Cambridge University Press, 1890). See in particular his Bradford address, "Molecules," Vol. II, pp. 361–377, in which he expresses most lucidly his religious and metaphysical position. The *Treatise on Electricity and Magnetism,* 3rd ed., edited by J. J. Thomas, was published in 1892 at Oxford. The standard biography is Lewis Campbell and William Garnett, *The Life of James Clerk Maxwell* (London: Macmillan, 1882).

Arthur E. Woodruff (1967)

MAYA

See *Indian Philosophy*

MCCOSH, JAMES

(1811–1894)

James McCosh, an influential representative of "commonsense realism," was born in southern Ayrshire, Scotland. He was educated at Glasgow and Edinburgh universities. McCosh was licensed for the ministry in 1834 and served as a pastor of the Established Church of Scotland until 1850, when he was appointed professor of logic and metaphysics at Queen's College of Belfast. In 1868 he came to America to serve as president of the College of New Jersey (now Princeton University), a position he held until 1888.

McCosh's philosophical outlook was in its largest features inherited from the "Scottish school" of Thomas Reid, Dugald Stewart, and others. On one side this meant the denial that our beliefs about the external world rest on any dubious inferences, causal or otherwise, from immediately presented ideas. Those beliefs are rather the natural, noninferential accompaniments of sensation,

and their general reliability cannot sensibly be questioned. On another (and for McCosh, more important) side, commonsense philosophy meant apriorism. In *The Intuitions of the Mind, Inductively Investigated* (London and New York, 1860), McCosh undertook to enumerate certain fundamental principles (such as principles of causation and moral good) that belong to the constitution of the mind. Although persons are not necessarily or normally aware of these very general truths, their particular cognitions and judgments are regulated by them. In saying that these principles are to be discovered "inductively" McCosh did not mean that they are inductive generalizations. Certainly one is led to these principles by reflection on experience. But once before the mind, the principles are recognized as self-evidently and necessarily true. McCosh's realism, unlike that of H. L. Mansel and William Hamilton, was relatively free of the influence of Immanuel Kant. Thus, in *An Examination of Mr. J. S. Mill's Philosophy* (London and New York, 1866), McCosh defended Hamilton's intuitional philosophy against Mill's criticism but took care to disassociate himself from the former's "agnostic" view that man's knowledge is limited to the finite.

The most original aspect of McCosh's philosophy was his effort to accommodate evolution and Christian theism. In one of his earliest works, *The Method of the Divine Government, Physical and Moral* (Edinburgh, 1850), he opposed the view that God's design exhibits itself entirely in the lawful development of nature. Such a view, he thought, amounted to a denial of divine providence. Divine government proceeds instead by a combination of law and particular, spontaneous interventions. When *The Origin of Species* appeared (1859), McCosh found it natural to identify his "special providences" with Charles Darwin's "chance variations." In *Christianity and Positivism* (New York and London, 1871) he argued that evolution, properly understood, is not only compatible with a divine design but in fact magnifies the Designer. Unlike Darwin, McCosh found nothing abhorrent in the notion that God employs the struggle for survival as a technique of creation. He was confident that success in that struggle was a matter of moral rather than physical strength.

McCosh's writings enjoyed considerable popularity, particularly among the evangelical clergy who found in them a way of dealing with the difficulties raised by science and science-inspired philosophies.

See also Common Sense; Darwin, Charles Robert; Darwinism; Hamilton, William; Kant, Immanuel; Mansel, Henry Longueville; Mill, John Stuart; Realism; Reid, Thomas; Stewart, Dugald.

Bibliography

Apart from those already mentioned, McCosh's chief works are *The Supernatural in Relation to the Natural* (New York: R. Carter, 1862); *The Scottish Philosophy, Biographical, Expository, Critical, from Hutcheson to Hamilton* (London, 1874); and *First and Fundamental Truths, Being a Treatise on Metaphysics* (New York: Scribners, 1889). An extensive bibliography by Joseph H. Dulles is appended to the autobiographical *The Life of James McCosh,* edited by William Milligan Sloane (New York: Scribners, 1896).

Douglas Arner (1967)

MCDOUGALL, WILLIAM
(1871–1938)

William McDougall, a British-American proponent of hormic psychology, was born in Chadderton, England, the second son of a chemical manufacturer. He was educated at schools in England and Germany, and at Manchester and Cambridge universities, where he received first-class honors in biology. In 1897 he qualified in medicine at St. Thomas's Hospital, London. While working there with Charles Scott Sherrington, he read William James's *Principles of Psychology,* and returned to Cambridge to study psychology on a fellowship from St. John's College. He joined the Cambridge Anthropological Expedition (1899) to Torres Straits, collaborating with W. H. R. Rivers in sensory researches and with Charles Hose in anthropological studies, which resulted in *The Pagan Tribes of Borneo* (London, 1912). He worked at Göttingen with G. E. Müller and subsequently joined the psychology department of University College, London, under James Sully, where he published researches supporting Thomas Young's theory of color vision against those of H. L. F. von Helmholtz and Ewald Hering (*Mind* 10 [1901]: 52–97, 210–245, 347–382). In London, and in Oxford from 1904 as Wilde reader in mental philosophy, McDougall worked on reflexes, inhibition, and psychophysical relationships. In *Physiological Psychology* (London, 1905) he combined James's view of instinctive action and emotion as objective and subjective aspects of the excitement of inherited perceptual dispositions with Sherrington's theory of the nervous system as integrator of reflex and instinctive-impulsive actions. McDougall explained subjectivity and purposiveness through R. H. Lotze's "psychoneural parallelism," postulating psychic currents induced in etherlike soul-stuff by neural activity.

McDougall first outlined his hormic psychology in *An Introduction to Social Psychology* (London, 1908). He derived human behavior from instincts, which are innate psychophysical dispositions with specific cognitive, affective, and conative aspects (for example, perception of danger, fear, flight). In adult humans, instincts operate indirectly through socially acquired patterns, the sentiments, in which object(s) and instinct(s) have become enduringly associated. Sentiments increasingly remote from innate instincts are exemplified, for instance, by parental love, family feeling, patriotism. In the growth of character the developing sentiments become hierarchically ranged round a master sentiment (or ruling passion) whose nucleus in a stable character is the self-regarding sentiment.

In *Body and Mind* (London, 1911), subtitled *A History and Defense of Animism,* McDougall reviewed psychophysical theories. To explain heredity and evolution, memory and learning, the "body-memory" of growth and repair, and parapsychological evidences of personal survival, he now discarded Lotzean parallelism, and declared himself, unfashionably, a dualist, interactionist, vitalist, animist, and Lamarckian.

In World War I McDougall enlisted as a French army ambulance driver but was drafted into the Royal Army Medical Corps. His command of a British shellshock unit provided the limited clinical material for his *Abnormal Psychology* (see below). In 1920 he became professor of psychology at Harvard, and in 1927 professor of psychology at Duke University. His American period was one of immense literary productivity. *The Group Mind* (New York, 1920) essayed to complete McDougall's social psychology by applying the hormic theory to "national mind and character." It was a work of subjective sociopolitical criticism rather than of objective scientific psychology, and resembled his many books of polemic and propaganda on national and international policy, from *Is America Safe for Democracy?* (New York, 1921) to *World Chaos* (London and New York, 1931). In these he advocated racial eugenics, a subsidized intellectual aristocracy, and a world air police, to defend the finest (explicitly North European–American) type of civilization.

In *An Outline of Psychology* (New York and London, 1923), *An Outline of Abnormal Psychology* (New York and London, 1926), and *Character and the Conduct of Life* (New York and London, 1927), McDougall elaborated his theory of personality built from sentiments that are powered by instincts, themselves channels of biological purposive energy (horme). The self-regarding sentiment governs conduct according to guidelines formed through identifications with admired persons or abstract ideals. Within the self-regarding sentiment, moral sentiments (conscience) control crude instinctive impulses, and thus, in McDougall's view, individual free will is truly exercised. The ordered hierarchy of sentiments completes the integration of personality. In *Abnormal Psychology,* McDougall reproached both Sigmund Freud and Carl Jung for neglecting the integration of personality—at that time Freud's "superego" and Jung's "self" were not yet formulated.

McDougall's theory still had to explain the occurrence of autonomous complexes apparently outside the hierarchy, and of dissociated activities and "multiple" personalities. Rejecting Freud's determinism, McDougall considered these unconscious mental functions purposive and goal-seeking. He then combined his personality theory with a revised view of body-mind relationships in an elaborate monadic theory based upon that of Gottfried Wilhelm Leibniz. Every personality is integrated as a converging hierarchy of monads, each "potentially a thinking striving self, endowed with true memory." A supreme monad "which each of us calls 'myself'" exercises control by telepathic communication through the hierarchy. Failure of integration allows pathological conflicts, automatisms in sleep or hypnosis, or even revolt of a subordinate monad as a dissociated personality.

McDougall left open the question whether monads might be perceptible through the senses, and he considered the monadic theory to be consistent with either a monistic or a dualistic psychophysical theory. To reconcile a presumably purposive mind with an apparently causally determined body, he suggested that there might be two types of monad, one goal-seeking and the other cause-following, that were somehow interconnected, or one single series of monads with two aspects, causalistic and finalistic. Thus McDougall reconciled his theory both with causal-mechanistic schemes of neurophysiological levels (Sherrington) and with more purposive views, neurological (Henry Head, *Studies in Neurology,* London, 1920) and psychological (hormism). However, he too hastily equated biological purpose (horme) with individual goal-seeking will, and acquired self-control with the capacity for choice and responsibility in conduct.

Once a noted experimental physiologist, McDougall later based hormic psychology increasingly upon his purposivist metaphysical beliefs, little upon verifiable observation or experiment. His great experimental work at Duke was designed to test Chevalier de Lamarck's hypothesis of evolution by inheritance of acquired char-

acteristics. Eventually, after ten years and twenty-three animal generations, McDougall reported an apparently inherited facilitation of learning in laboratory rats. Subsequent workers have not confirmed his results.

A lucid and persuasive writer, McDougall wielded great if temporary influence, and guided many English-reading students toward dynamic, biological, and social psychology. His weaknesses were his fondness for intellectual and verbal solutions to empirical problems, and his temptation to premature systematization. Admiration tinges the epigram that, had the Creator but paused to consult William McDougall, there had been no need of redemption.

See also Darwinism; Freud, Sigmund; Helmholtz, Hermann Ludwig von; James, William; Jung, Carl Gustav; Lamarck, Chevalier de; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Macrocosm and Microcosm; Panpsychism; Psychology; Racism; Vitalism.

Bibliography

ADDITIONAL WORKS BY MCDOUGALL

Psychology, the Study of Behavior. London: Williams and Norgate, 1912.

Ethics and Some Modern World Problems. New York and London, 1924.

The Battle of Behaviourism. London: Kegan Paul, 1928. Written with J. B. Watson.

Modern Materialism and Emergent Evolution. New York: Van Nostrand, 1929.

Psycho-analysis and Social Psychology. London: Methuen, 1936.

WORKS ON MCDOUGALL

Greenwood, Major, and May Smith. "William McDougall." In *Obituary Notices of Fellows of the Royal Society,* Vol. III, 39–62. London, 1939–1941. Contains a complete bibliography.

Nicole, J. Ernest. *Psychopathology.* London, 1930; 3rd ed., 1942. Chs. 15, 16, 21.

Woodworth, R. S. *Contemporary Schools of Psychology.* New York: Ronald Press, 1931. Ch. 6.

J. D. Uytman (1967)

MCDOWELL, JOHN

(1942–)

John McDowell, a professor of philosophy at the University of Pittsburgh, was born in Boksburg, South Africa. After receiving his bachelor's from the University College of Rhodesia and Nyasaland he was awarded a Rhodes scholarship to New College, Oxford, where he earned a second bachelor's in 1965 and a master's in 1969. In 1966 he became a fellow of University College, Oxford, where he remained until he joined the faculty at the University of Pittsburgh in 1986. McDowell is a fellow of both the British Academy and the American Academy of Arts and Sciences.

With the rise of modern science there emerged a view of the world that is radically different from that of everyday life, a view sometimes described as "the view from nowhere." This new view was made possible, McDowell argues, by a new clarity regarding natural scientific understanding. Modern natural science explains things not by giving reasons to show that they are somehow better that way but by subsuming them under discoverable physical laws; it understands things by locating them within the realm of law as it contrasts with what Wilfrid Sellars calls the space of reasons. Because modern scientific understanding focuses on explanation by appeal to (physical) laws rather than to reasons, the world as revealed in the view from nowhere is "disenchanted," empty of meaning and value, indeed, of all distinctively human significance. One of the most pervasive themes in McDowell's work (whether in the philosophy of language, the philosophy of mind, metaphysics, epistemology, or ethics) is that philosophers since René Descartes have mistakenly assumed that respectable philosophy must begin with the view from nowhere, and thereby with a conception of nature as the realm of law, rather than with the everyday view from here and its much richer conception of nature.

Consider an ordinary sign, say a stop sign. In day-to-day life one knows how to follow such a sign. But how, the philosopher asks, can one follow the rule expressed by the sign given that what is presented is itself a mere thing, merely a piece of painted metal? It can seem natural to answer that the sign expresses a rule, tells one how to go on, only under an interpretation, that independent of an interpretation of that bit of matter as a stop sign, the sign just stands there. But this cannot be right, McDowell argues following Ludwig Wittgenstein, because any interpretation—say an utterance of the sound *stop*—will be similarly inert unless provided with an interpretation. The right response is to reject the assumption that what is presented is a mere thing. One can learn to conceive the sign as a mere thing independent of all human concerns, just as one can learn to conceive nature in a way that is independent of sensory experience. (One can learn to take the view from nowhere.)

But that capacity is essentially late; it cannot be understood except against the backdrop of one's everyday ability to follow rules such as that expressed in a stop sign.

Indeed, thinking of a sign as a mere thing is itself a matter of rule-following: from the perspective afforded by the view from nowhere, the sign tells one how it is to be thought, namely, as a particular bit of stuff shaped in a certain way. Although the view from nowhere involves pure cognition rather than bodily action, one needs in that case as well the notion of going on in light of a conception of correctness, of thinking one way rather than another on the basis of an understanding of the thing about which one thinks.

Knowing how to follow a rule is at least in some cases a perceptual skill, the ability to see an expression of the rule (e.g., a stop sign) as telling one how to go on. In his masterwork *Mind and World* (1994) McDowell argues more generally that experience, conceived as the capacity to take in manifest facts (e.g., to see that things are thus and so), is an essential component in any adequate conception of cognition. According to his diagnosis the modern unquestioned assumption that natural scientific understanding is the only acceptable mode of access to nature leads philosophers to begin with the mistaken idea that the space of reasons within which thought operates is dualistically opposed to nature. As a result, modern philosophy falls into an oscillation between two equally unsatisfactory conceptions of cognition: on the one hand, an empty coherentism that eschews the notion of experience altogether, and on the other hand, what Sellars calls the "Myth of the Given," the idea that brute impacts of the sort described in physics might provide a perceiver with reasons for belief.

Rejecting the assumption that generates the oscillation, McDowell urges that what is needed instead is the Kantian conception of experience as inextricably involving both sensibility and understanding. Because experience so conceived is at once passive, that is, receptivity in operation, and conceptually articulated, it can serve rationally to constrain one's thought about what is the case, and thereby to explain the empirical contentfulness of thought. As McDowell also argues, the capacity for experience so conceived is essentially second nature; it is acquired only in the course of one's acculturation into natural language, where natural language is itself to be understood as a repository of tradition, the embodiment of the possibility of an orientation to the world.

In his writings on ethics McDowell argues that modern philosophers have a fundamentally distorted conception of practical reason grounded in their scientistic understanding of nature and that this conception has blinded them to the insights of the ancient Greeks. The capacity to act virtuously, he argues following Aristotle, essentially involves the capacity to take in objective moral facts, where this latter capacity—like the capacity to take in nonmoral facts—is acquired in the course of one's acculturation. It follows that the rationality, and so the desirability, of a life of virtue cannot be established from the outside, independent of how a virtuous person sees things. Critical reflection in ethics, as in any other domain, is Neurathian, possible only from within the tradition one inherits.

Although mostly written in the form of essays, McDowell's work systematically addresses many of the deepest philosophical perplexities that can arise on reflection about human being in the world and the nature and place of language in human life. His writings provide a diagnosis and a cure for the ills of modernity, and a rich, subtle, and profoundly moral vision of what it is to be human.

See also Aristotle; Descartes, René; Ethics, History of; Metaethics; Philosophy of Language; Philosophy of Mind; Rule Following; Sellars, Wilfrid.

Bibliography

WORKS BY MCDOWELL

Plato. *Theaetetus*. Oxford, U.K.: Clarendon Press, 1973. A translation with extensive notes.

Mind and World. Cambridge, MA: Harvard University Press, 1994. The text of McDowell's 1991 Locke Lectures, with a long afterword.

"Having the World in View: Sellars, Kant, and Intentionality." *Journal of Philosophy* 95 (9) (1998): 431–491. A revised version of McDowell's 1997 Woodbridge Lectures.

Meaning, Knowledge, and Reality. Cambridge, MA: Harvard University Press, 1998. A collection of essays in the philosophy of mind, the philosophy of language, metaphysics, and epistemology.

Mind, Value, and Reality. Cambridge, MA: Harvard University Press, 1998. A collection of essays in ethics, the philosophy of mind, and metaphysics.

WORKS ON MCDOWELL

De Gaynesford, Maximilian. *John McDowell*. Cambridge, U.K., and Malden, MA: Polity Press, 2004.

Smith, Nicholas H., ed. *Reading McDowell: On Mind and World*. London and New York: Routledge, 2002. A collection of essays, with responses by McDowell.

Stjernberg, Fredrik, ed. *Theoria* 70 (2/3) (2004). A special issue devoted to essays on McDowell's work, with responses by McDowell.

Thornton, Tim. *John McDowell*. Chesham, U.K.: Acumen, 2004.

Danielle Macbeth (2005)

MCGILVARY, EVANDER BRADLEY

(1864–1953)

Evander Bradley McGilvary, an American realist philosopher, was born in Bangkok, Siam. He received his B.A. from Davidson College in 1884, his M.A. from Princeton in 1888, and his Ph.D. from the University of California in 1897. He was appointed assistant professor of philosophy in California and then Sage professor of ethics at Cornell (1899–1905). From 1905 to 1924 he was professor of philosophy and head of the department at the University of Wisconsin, and in the year 1912–1913 he was the president of the American Philosophical Association. He was the Howison lecturer in 1927, the Mills lecturer in 1928, and the Carus lecturer in 1939.

PHILOSOPHICAL ORIENTATION

McGilvary's "first impulse" toward philosophy was a reaction against the theology in which he was schooled. He came under the Hegelian influence of George Howison at California, and his writings from 1897 to 1903 reflect this influence. But McGilvary, like other Hegelians of his time, eventually found Hegelianism unacceptable. From the start McGilvary held the view that every part of the world is what it is by virtue of its organic relation to every other part. And when he broke with Hegelianism, he took with him this theory of relations and the characteristically Hegelian view that two antagonistic ideas always suggest a third that synthesizes the truth of each.

Realist philosophers in America during the first two decades of the twentieth century were struggling to formulate an epistemology that would do justice both to those elements in experience that are clearly in the objective world and to those dependent upon the experiencing organism. Taking William James's thesis that "the world is as it is experienced," the non-Hegelian new realists developed a monistic realism, but it always threatened to become panobjectivism. In reaction the critical realists set forth a dualistic realism that always threatened to become pansubjectivism. In his "perspective realism" McGilvary sought to combine the truth of new realism with the truth of critical realism. He, too, took James's thesis as his starting point and sought to combine epistemological monism with epistemological dualism and the theory of external relations with the theory of internal relations. McGilvary's synthesis of the objective and the relative—like John Dewey's and A. N. Whitehead's—was dubbed "objective relativism" by A. E. Murphy.

To effect the synthesis of monism and dualism, McGilvary developed his theory of perspectives. It is summarized in the first three postulates of perspective realism: (1) "In our sense-experience there is presented to us in part the real world in which we all in common live"; (2) "Every particular in the world … is what it is only because of its context"; (3) "In the world of nature any 'thing' at any time is, and is nothing but, the totality of the relational characters, experienced or not experienced, that the 'thing' has at that time in whatever relations it has at that time to other 'things.'" McGilvary first hinted at such a theory in 1907, but he did not systematically state it until twenty years later, and in 1939 it became the core of his Carus lectures, *Toward a Perspective Realism.* This work is the key to understanding McGilvary's philosophy, and it grew out of his early thinking about the nature of consciousness.

THE NATURE OF CONSCIOUSNESS

McGilvary believed that the question of the precise nature of consciousness was the fundamental question of philosophy. Like other realists, he agreed with James that consciousness is a relation. Since it was his view that things are what they are only in their relations to other things, he could not agree with realists who claimed that this relation was external. Consciousness, he held, is that relation by which anything becomes an experience. It is a unique kind of "togetherness" of, or between, things. It is neither a spatial nor temporal togetherness, nor is it any other distinguishable relation. The peculiar relation of feeling binds external objects together into an experiential unity we call "consciousness," "awareness," or "experiencing."

McGilvary thought this togetherness may have been what Immanuel Kant meant by the synthetic unity of apperception. It has a unique center of reference in the body of the experiencing organism. This centering gives to the relation of togetherness a character and coloring all its own. Hence, consciousness exists in individualized instances, like other relations, yet each instance produces an individuality generically different from that of any other individualized relation. Each instance is its own kind of betweenness.

As he developed this theory, McGilvary increasingly described consciousness in terms of perspectives. In addition to the familiar perceptual perspectives of space and time, he said, consciousness is characterized by intellectual, moral, and aesthetic perspectives. All these perspectives have both a physical and an "epiphysical," a dynamic and an "epidynamic," causal and noncausal quality. The

most distinctive characteristic of these perspectives is the absence of energy transaction between their station point (the organism) and objects in the perspective. The peculiar "epidynamic" relatedness of a perspective does not "go over" to the object or do anything to it. Yet it does "go over" in the way any other relation "goes over" from one term to another. It is a conditioning relatedness that is not itself a cause of the physical existence of its objects, nor is it itself an object in the relation complex. Thus, a perspective (seeing, for example) is not an act of the organism on its object. If it were, it would be difficult to understand how an organism can see now what antedates the seeing, such as a star that may have exploded aeons ago. Like the verb "to relate," the verb "to see" does not name an act performed on the objects seen, any more than "having" a grandfather is an act performed on him. Physical objects become a field of vision when light from them stimulates an organism through its eyes, just as grandparents become grandparents only when a grandchild is born.

The organism, then, is a condition of vision, and as such it is not one of the members or terms in the relationship, just as common parents are a condition for the relationship of brotherhood but are not members in that relationship. Seeing the star that no longer exists is no more difficult for McGilvary to explain than how being an ancestor of a president of the United States is a quality that comes to belong to persons who die before the event that permits ascribing that characteristic to them. In the same way the perspective realist can hold that the physical object that initiated the series of physical conditions that ended in a perception of attributes occupying the position of that object still does not have those attributes. These attributes, however, can be considered part of the real world resulting from a real and natural relation between the organism and external objects. Not all physical qualities, then, are causally conditioned. Sense qualities, for example, can be considered part of their object but are not causally related to the organism that senses them.

It is the same for McGilvary with memory or knowledge of the past. The pastness of an event is not independent of all external standpoints. The pastness of consciousness is retro*spective,* a particular kind of perspectivity, but not retro*active.* Consciousness also is prospective, another kind of perspectivity, but not active on the future. This is the "epiphysical" or "epidynamic" quality of the consciousness relation that distinguishes it from other physical, dynamic, causal relations that act on their objects. Perspectives do not exist if that means being in space and time. Nor do they subsist. The being of a perspective is its being between—"inter-sistence," McGilvary called it—and each perspective is its own kind of "inter-sistence."

But it is not clear whether McGilvary thought that each perspective is an instance of consciousness and whether perspectives go to make up what we call consciousness. Nor does he show us how to distinguish between what the organism contributes to the perspective, as its station point, and what is there independent of the organism. At times he said nothing is there independent of the organism, for the organism is the necessary condition of any perspective. But when Dewey said that the logical forms of our knowledge cannot be read back into nature (because they come into being only when inquiry is instituted and are only modes of operating upon subject matter), McGilvary disagreed. He argued that any logical form that serves to solve a problematic situation serves that purpose because it is actually the form of the subject matter under investigation, not of the subject matter as it was immediately experienced when inquiry started but as successful inquiry shows the subject matter to have been in the natural world.

It is doubtful, then, that McGilvary, like the other objective relativists, was any more successful than other realists in doing justice to the objective and the relative found in experience.

McGilvary's few articles on ethics present familiar positions, but none of them is developed systematically, nor did McGilvary apply his perspective realism beyond epistemological and ontological problems.

See also Consciousness; Dewey, John; Hegelianism; Howison, George Holmes; James, William; Murphy, Arthur Edward; Realism; Whitehead, Alfred North.

Bibliography

WORKS BY MCGILVARY

"Pure Experience and Reality: A Re-assertion." *Philosophical Review* 16 (May 1907): 266–284.

"The Physiological Argument against Realism." *Journal of Philosophy* 4 (October 1907): 589–601.

"Realism and the Physical World." *Journal of Philosophy* 4 (December 1907): 683–692.

"Experience and Its Inner Duplicity." *Journal of Philosophy* 6 (April 1909): 225–232.

"Experience as Pure and Consciousness as Meaning." *Journal of Philosophy* 8 (September 1911): 511–525.

"The Relation of Consciousness and Object in Sense Perception." *Philosophical Review* 21 (March 1912): 152–173.

"A Tentative Realistic Metaphysics." In *Contemporary American Philosophy,* edited by G. P. Adams and W. P. Montague, 2 vols. New York: Macmillan, 1930. Vol. II, pp. 109–132.

"The Revolt against Dualism." *Philosophical Review* 40 (May 1931): 246–265.

"Perceptual and Memory Perspectives." *Journal of Philosophy* 30 (June 1933): 109–132.

Toward a Perspective Realism. La Salle, IL: Open Court, 1956. His 1939 Carus lectures and his only book.

Between 1918 and 1926 McGilvary published only two book reviews. His publications after 1926 display a new interest in and command of mathematical physics. Two of these papers are reprinted in his book. In all he published 48 articles and 23 reviews in addition to 81 articles in *The New International Encyclopedia* (New York, 1902).

WORKS ON MCGILVARY

Murphy, Arthur E. "McGilvary's Perspective Realism." *Journal of Philosophy* 56 (February 1959): 149–165. Reprinted in *Reason and the Common Good,* edited by W. H. Hay et al. Englewood Cliffs, NJ: Prentice-Hall, 1963.

Oliver, Donald W. "The Logic of Perspective Realism." *Journal of Philosophy* 35 (April 1938): 197–208.

Thomas Robischon (1967)

MCTAGGART, JOHN MCTAGGART ELLIS

(1866–1925)

John McTaggart Ellis McTaggart, a British metaphysician, was born in London, the son of Francis and Caroline Ellis. (His father later took the name McTaggart to fulfill a condition for inheriting a bequest.) He attended school at Clifton and went on to Trinity College, Cambridge, where he took first-class honors in the moral science tripos in 1888. He was made a fellow of Trinity in 1891. The next year he paid a visit to New Zealand, where his widowed mother lived, and there he met Margaret Elizabeth Bird, whom he married in 1899, during a second visit to New Zealand. Thereafter he resided at Cambridge. Active in the affairs of his college and the university, he was a busy and successful teacher from 1897 until he retired in 1923. He died suddenly in January 1925.

McTaggart's philosophy is a peculiar and quite personal variety of Hegelian idealism. Ultimate reality, he held, is spiritual: It consists entirely of individual minds and their contents. He understood this in a way that excludes space, time, and material objects from reality. What appear to us as being these things are really minds and parts of the contents of minds, but we "misperceive" these entities in a systematic way, and this misperception is the source of the whole apparent universe. Despite the unreality of time, McTaggart argued, there is an important sense in which it is true to say that individual persons are immortal, and that they are reincarnated in a succession of (apparent) bodies. He also held that in reality persons stand in relations either of direct perception, and consequently love, or of indirect perception, and consequently affection, to one another. Love is, indeed, the basically real emotional state. There is, however, no God in this heavenly city, for McTaggart did not think there is any reason to believe that there is or even can be an overarching mind that includes individual minds like ours but is still in some sense an individual mind itself. McTaggart was, in addition, a determinist, though he held that determinism is not incompatible with the existence of valid judgments of moral obligation.

On these basic points McTaggart never changed his mind. He argued in support of them both in his early writings on G. W. F. Hegel and in his great systematic work, *The Nature of Existence*. The main difference between his earlier and his later work is that in the former the arguments are dialectical in a Hegelian manner, whereas in the latter they are more straightforwardly deductive.

WRITINGS ON HEGEL

McTaggart's commentaries on Hegel are all more or less critical of Hegel, and none is entirely reliable as pure exegesis. Two deal primarily with Hegelian methodology. The essays on the dialectic defend Hegel's method against what McTaggart took to be common misunderstandings and criticisms and offer an account of the way in which the Absolute Idea works to move thought from stage to stage. The *Commentary on Hegel's Logic* is a detailed and very careful examination of the validity of each step in the logical development of the categories. McTaggart frequently found Hegel to be mistaken or confused about his transitions and in some cases offered alternative modes of development.

The essays on cosmology are among McTaggart's most interesting work. He here discussed, more fully than anywhere else, a number of concrete topics—such as the moral criterion, sin, the organic nature of society, and the relations between Christianity and Hegelianism—in the light of his metaphysical position. He brought out his differences, not only with Hegel, but with many of the British Hegelians as well. And in the concluding chapter he presented with great clarity and power what is essentially his mature view of the relations between selves in ultimate reality.

SOME DOGMAS OF RELIGION

In *Some Dogmas of Religion* McTaggart examined, in a careful but nontechnical manner, a number of dogmas that are especially relevant to Christianity. (By *dogma* he meant "proposition having metaphysical significance.") He argued that dogmas of some sort are essential to any religion and that we must have reasoned proof of a dogma before we can be justified in believing in it. Then, without claiming to give conclusive arguments (for these would involve a whole metaphysical system) he argued in favor of immortality, preexistence, and determinism, criticized the belief in a personal and omnipotent God, and attacked some of the arguments that have been alleged to support this belief. Finally, he tried to show that there is much less connection than is frequently held to be between the truth of theism and improved chances for personal happiness.

NATURE OF EXISTENCE

McTaggart's metaphysical system is presented in two parts. In the first, contained in Volume I of *The Nature of Existence*, he gave an extended argument to show that whatever exists must be of a certain nature and must, therefore, satisfy a certain requirement, to be explained below. In the second part, occupying Volume II, he examined various types of entities that our present experience shows us as existing to determine whether these entities can satisfy the requirement; he attempted to account for the apparent existence of those entities that do not really exist; and he evaluated the practical importance of the results he had thus reached.

The argument of Volume I is almost entirely a priori. McTaggart appealed to experience for only two propositions: that something exists, and that what exists has parts. His argument proceeds through the following stages: First, McTaggart offered a proof of the principle of the Identity of Indiscernibles. Second, he argued that every substance must have a "sufficient description," that is, a description that uniquely identifies the substance and contains no reference to substances that are only identified (as by pointing or by the use of purely referring expressions), not described.

He next moved to the assertion that every substance, without exception, must be divisible into parts that are themselves substances, and hence into parts within parts to infinity. The crucial argument is then presented. The principle that every substance must have a sufficient description together with the principle that every substance is infinitely divisible into further substances would entail a contradiction unless the substances in question were such that from the nature of any existing substance there follow sufficient descriptions of all of its parts within parts to infinity. This can occur, McTaggart showed, if the substance stands in a certain extremely complex relation to its parts, which he called the relation of "Determining Correspondence"; it can occur, he held, in no other way. Hence, whatever exists—and we know that something does exist—must satisfy the conditions necessary for it to stand in Determining Correspondence relations to its parts.

In Volume II McTaggart denied the existence of material objects, space, judgments, inferences, sense data, and certain other mental contents, on the ground that entities of these types cannot satisfy the conditions required for them to stand in Determining Correspondence relations. His denial of the existence of time, however, rests on a quite different argument. This argument is McTaggart's most widely discussed contribution to philosophy. Briefly, it is as follows: Temporal positions and events may be ordered either as earlier-later or as past-present-future. Ordered the first way, they form what McTaggart called a *B*-series; ordered the second way they form an *A*-series. In the first stage of the argument McTaggart tried to show that the *A*-series characteristics "past," "present," and "future" are essential to the existence of time. He assumed it to be admitted that change is essential to time, and he argued that unless the *A*-series characteristics can change, nothing can change. The *B*-series characteristics cannot change, for if an event is ever earlier than another, it is always earlier; and neither can the other characteristics of events change, for if it is ever true that an event is, for instance, the death of a queen, then it is always true that this event is the death of a queen. Hence, without the *A*-series there cannot be time, and in the second stage of the argument McTaggart tried to show that a vicious infinite regress is involved in affirming the existence of a series ordered by *A*-series characteristics. Each member of such a series must have all the *A*-series characteristics, he said, but those characteristics are incompatible. If we try to remove the contradiction by saying that each member possesses all the characteristics *at different times*, we are presupposing the existence of different moments of time at which the *A*-series characteristics are possessed. But each of these moments, to be temporal, must itself possess all of the *A*-series characteristics, which, again, is impossible; the attempt to relieve this contradiction by appeal to yet another set of moments only gives rise to another set of contradictions, and so on.

McTaggart's complicated and difficult account of the relations between appearance and reality centers on the concept of a *C*-series, analogous to the *B*-series in having its members related by an asymmetrical and transitive relation, but timeless. The model for the *C*-series relationship is the concept of "inclusion," and the terms that are included in and inclusive of each other are perceptions, that is, parts of spirits. McTaggart argued that reality must be structured so as to form a set of related inclusion series that, however, are misperceived as temporal series. He drew the further conclusion that time had a first moment and will have a last moment.

McTaggart went on to discuss the question of the value of the universe, both in its prefinal stages and at the stage when the appearance of time has ceased. Taking both "good" and "evil" to stand for simple, unanalyzable characteristics, and arguing that only what is spiritual can have value, he found that in the prefinal stages the relative proportions of good and evil will fluctuate considerably, though we can be confident that on the whole the proportion of good will steadily increase. In the final stage we will exist in a "timeless and endless state of love" far more profound and powerful than anything we now have any inkling of. We shall, McTaggart said, "know nothing but our beloved, those they love, and ourselves as loving them," and this will be our ultimate and unshakable satisfaction. If McTaggart's metaphysics thus concludes with a vision that he himself was not unwilling to call mystical, it is at least a vision that springs from one of the most brilliantly conceived and carefully executed attempts any philosopher has ever produced to grasp the nature of reality in purely rational terms.

See also Hegel, Georg Wilhelm Friedrich; Time.

Bibliography

WORKS BY MCTAGGART

Studies in the Hegelian Dialectic. London: Cambridge University Press, 1896; 2nd ed., Cambridge, 1922.

Studies in Hegelian Cosmology. London: Cambridge University Press, 1901; 2nd ed., Cambridge, 1918.

Some Dogmas of Religion. London: Arnold, 1906; 2nd ed., London, 1930.

A Commentary on Hegel's Logic. London: Cambridge University Press, 1910.

The Nature of Existence. 2 vols. Vol. I, London: Cambridge University Press, 1921; Vol. II, edited by C. D. Broad. London: Cambridge University Press, 1927.

Essays in *Philosophical Essays*, edited by S. V. Keeling. London: Arnold, 1934.

Human Immortality and Pre-existence. New York: Kraus Reprint, 1970.

Philosophical Studies, edited by S. V. Keeling. South Bend, IN: St Augustine's Press, 2000.

WORKS ON MCTAGGART

C. D. Broad delivered an obituary address to the British Academy that contains an admirable summary of McTaggart's work. It was published in the society's *Proceedings* for 1927 and reprinted in the second edition of *Some Dogmas of Religion*, as well as in Broad's *Ethics and the History of Philosophy* (London: Routledge, 1952). *John McTaggart Ellis McTaggart*, by G. Lowes Dickinson (London: Cambridge University Press, 1931), contains more information about McTaggart's life, in addition to interesting reminiscences and a chapter by S. V. Keeling on McTaggart's metaphysics. The standard commentary is C. D. Broad's exhaustive *Examination of McTaggart's Philosophy* (Vol. I, London: Cambridge University Press, 1933; Vol. II, in two parts, London: Cambridge University Press, 1938). This work is discussed at length by R. L. Patterson in *The Philosophy of C. D. Broad* (New York: Tudor, 1959). (For discussion of McTaggart on time, see the bibliography for the *Time* entry.)

OTHER RECOMMENDED WORKS

Airaksinen, Timo. *The Ontological Criteria of Reality: A Study of Bradley and McTaggart*. Turku, Finland: Turun Yliopisto, 1975.

Cesarz, Gary L. *Substance and Relations in McTaggart's Philosophy: A Re-examination of His Basic Principles*. PhD diss., University of New Mexico, 1988.

Farmer, David John. *Being in Time: The Nature of Time in Light of McTaggart's Paradox*. Lanham, MD: University Press of America, 1990.

Gale, Richard M. *The Language of Time*. London, Routledge & Kegan Paul; New York, Humanities Press, 1968.

Geach, P. T. *Truth, Love, and Immortality: An Introduction to McTaggart's Philosophy*. Berkeley: University of California Press, 1979.

Mellor, D. H. *Real time*. Cambridge, U.K., and New York: Cambridge University Press, 1981.

Rochelle, Gerald. *Behind Time: The Incoherence of Time and McTaggart's Atemporal Replacement*. Aldershot, Hants, U.K.; Brookfield, VT: Ashgate, 1998.

Rochelle, Gerald. *The Life and Philosophy of J. McT. E. McTaggart, 1866–1925*. Lewiston, NY: E. Mellen, 1991.

Schulz, James Allen. *McTaggart's Theory of Substance*. PhD diss., Northwestern University, 1974.

J. B. Schneewind (1967)
Bibliography updated by Michael J. Farmer (2005)

MEAD, GEORGE HERBERT

(1863–1931)

George Herbert Mead, the American pragmatist philosopher, was born in South Hadley, Massachusetts. He received his BA from Oberlin College in 1883 and did graduate work at Harvard in 1887–1888, where he studied under Josiah Royce and William James. From 1888 to

1891 he studied psychology and philosophy in Europe. He was married in 1891 and in the same year was appointed instructor at the University of Michigan. In 1892 he joined the staff of the University of Chicago and later became chairman of its philosophy department.

A major figure in American pragmatism, Mead has also had a large influence on psychologists and social scientists. Many thinkers, including Alfred North Whitehead and John Dewey, regarded Mead as a creative mind of the first magnitude. He published relatively few papers, however, and died before he was able to develop his many original ideas into an integrated philosophy. Large segments of his books were collated from his unfinished manuscripts and from his students' notes and hence are repetitious, unsystematic, and difficult.

Mead's main philosophic themes may be classified as follows: (1) the emergence of mind and self from the communication process between organisms (often termed his "social behaviorism"), discussed in *Mind, Self and Society*; (2) the psychological genesis of scientific categories in purposeful acts, discussed in *The Philosophy of the Act*; and (3) the social conception of nature and the location of reality in the present, discussed in *The Philosophy of the Present.*

SOCIAL BEHAVIORISM

Mead's thought stemmed from the impact of Darwinism on nineteenth-century ideas. Man was regarded as an organism functioning in accordance with natural laws. This approach opposed traditional philosophy and theology and sought to understand human nature by the methods of experimental science. The theory of evolution also gave impetus to the conception of the universe as a process rather than as a set of fixed, unalterable essences that remain invariant over time. In psychology the process concept was expressed in functionalism, which sought to comprehend all mental phenomena not as structures, traits, or attributes of the mind but as relations between the organism and its environment. These ideas were taken up by behavioristic psychology, which dismissed introspection as unscientific and confined itself to experimental data, particularly the responses of organisms to stimuli under varying conditions.

Mead challenged many of the crudities of behaviorism. In rejecting introspection, this school tended to regard it as a nonexistent phenomenon, since it could not be studied experimentally. Mead's social behaviorism sought to widen behaviorism to include the introspectively observed phenomena of consciousness. For Mead stimulus and response are meaningful only when viewed as aspects of communication; they cannot be studied in abstraction from the social process in which actions occur. Furthermore, organisms do not merely respond mechanically and passively to stimuli. Rather, the individual purposefully selects its stimuli. Mead here opposed associationism; the organism is a dynamic, forceful agent, not a mute receptacle for ideas that are later associated. For Mead organism and environment mutually determine each other. Mind emerges from this reciprocal determination.

Mead's naturalistic conception of introspection was based on the viewpoint that an idea is the early, inner stage in an ongoing act directed toward an environmental goal. The mistake of the behaviorists was to study merely one part of the complete act, the last, overt stage, thereby ignoring the initial phase of the act, which occurs privately, within the organism.

According to Mead actions occur within a communicative process. The initial phase of the overt stage of an act constitutes a gesture. A gesture is a preparatory movement that enables other individuals to become aware of the intentions of the given organism. The rudimentary situation is a conversation of gestures, in which a gesture on the part of the first individual evokes a preparatory movement on the part of the second, and the gesture of the second organism in turn calls out a response in the first person. On this level no communication occurs. Neither organism is aware of the effect of its own gestures upon the other; the gestures are nonsignificant. For communication to take place, each organism must have knowledge of how the other individual will respond to his own ongoing act. Here the gestures are significant symbols.

Communication is also based on the fact that actions are organized temporally. The consequences of behavior (final phases of the act) are present in imagery during the early phases of the action and control the nature of the developing movement. There are usually several alternative ways of completing a movement that has been started. Since the final phases of the act control the ongoing movement, the organism can select one of these alternative ways of conjoining means with the end. In this manner rational conduct is possible. Where organisms use significant symbols, the role of the other individual controls the ongoing act. In advance of our completion of a social action, we anticipate the response of the other individual. Since our behavior is temporally organized, the imported role of the other may cause us to select a course of action that is different from what we originally intended.

Mind is the ability of an organism to take the role of the other toward its own developing behavior. Reflexivity, the ability of a person to reflect upon himself, is the necessary condition for the emergence of mind within the social process. With reflexivity the social act is imported within the individual and serves to alter the person's ongoing acts. A complete social act can be carried out internally without external movements necessarily occurring. Mead denotes the internalized role of the other as the "me." Each organism has an "I," which is a capacity for spontaneity. The "I" is expressed when the individual alters his ongoing response or creates a new response to the "me." Individuality and originality arise from the inner conversation between the "I" and the imported role of the other. An inner forum comes to exist, consisting of a dialogue between the "I" and the "me." This inner rehearsal of projected actions constitutes introspection, or thinking.

In the organized group situation, such as is exemplified in games, the individual learns to take into himself the entire social organization which now exerts internal control over his ongoing acts. The "generalized other" is the group's attitudes imported into the individual. It is here that social institutions enter into an individual's thinking as a determinative factor and cause him to develop a complete self. Now the inner forum becomes an inner dialogue between the person and the group.

The religious experience occurs in situations where each person becomes closely identified with the other members of the group. In common efforts, such as in teamwork, where a sense of closeness develops among everyone involved, a feeling of exaltation arises. Here Mead refers to a "fusion" of the "I" and the "me."

Mead's social psychology is similar to the psychoanalytic theories of Sigmund Freud and Harry Stack Sullivan in that it conceives personality as arising from the internalization of the roles of other persons and relates inner conflict to the tension between the spontaneous forces of the person and the introjected demands of society. The temporal organization of the act, stressed by Mead, is also a key concept in automatic control machinery and digital computers, where the later stages of a process feed back upon the earlier phases, modifying the ongoing process.

PHILOSOPHY OF SCIENCE

Mead sought to find the psychological origin of science in the efforts of individuals to attain power over their environment. The notion of a physical object arises out of manipulatory experience. Perception is coordinated with the ongoing act: When we approach a thing we wish to manipulate, the imagery of handling that thing is present in the distance perception. Here again there is a temporal organization of the act, in that the later phase of the action, the contact experience, is present in the earlier stage when we are merely perceiving the distant object. Perception involves the readiness of the organism to manipulate the thing when the intervening distance has been traversed. The reality of a thing is in the consummatory phase of the act, the contact experience, and this reality is present in the experience of perceiving that thing at a distance.

There is a social relation to inanimate objects, for the organism takes the role of things that it manipulates directly or that it manipulates indirectly in perception. For example, in taking (introjecting or imitating) the resistant role of a solid object, an individual obtains cognition of what is "inside" nonliving things. Historically, the concept of the physical object arose from an animistic conception of the universe.

Contact experience includes experiences of position, balance, and support, and these are used by the organism when it creates its conceptions of the physical world. Our scientific concepts of space, time, and mass are abstracted from manipulatory experience. Such concepts as that of the electron are also derived from manipulation. In developing a science we construct hypothetical objects in order to assist ourselves in controlling nature. The conception of the present as a distinct unit of experience, rather than as a process of becoming and disappearing, is a scientific fiction devised to facilitate exact measurement. In the scientific worldview immediate experience is replaced by theoretical constructs. The ultimate in experience, however, is the manipulation and contact at the completion of an act.

COSMOLOGY

The Philosophy of the Present develops the conception that reality always exists in a present. However, as it is experienced, the present involves both the past and the future. A process in nature is not a succession of instantaneous presents or a sequence of spatial points. Instead there is both spatial and temporal duration, or continuity.

The developing action is the basis of existence. It is true that as we look back the present is determined by the past. But each new present, as it passes into the next present, is a unique emergent. A new future also arises as the result of the emerging present. Hence, we are always reconstructing our pasts and restructuring our future.

Novelty stretches out in both directions from the present perspective.

Every object in the universe is seen from the perspective of a particular individual. What is seen from one person's perspective may be different from that which is seen by another individual. Mead was not solipsistic, however, for although a person sees nature only from his own perspective, he is able to import within himself the perspectives of others. Reality is the integration of different perspectives. Mead made use of the theory of relativity to project his theory of sociality and mind into nature. Sociality is the ability to be in more than one system at a time, to take more than one perspective simultaneously. This phenomenon occurs in emergence, for here an object in the process of becoming something new passes from one system to another, and in the passage is in two systems at the same time. During this transition, or transmutation, the emergent entity exists on two levels of nature concomitantly.

Mead's philosophy has been compared with that of Martin Buber. Although their approaches stem from different traditions, both thinkers have a social conception of nature and conceive of the self as arising from a social matrix. Certain affinities between Mead and Edmund Husserl have been suggested, in that the mind's reflexive examination of itself is an effort to describe the constitution and foundation of experience.

See also Behaviorism; Buber, Martin; Darwinism; Dewey, John; Evolutionary Theory; Experience; Freud, Sigmund; Husserl, Edmund; James, William; Natural Law; Pragmatism; Royce, Josiah; Whitehead, Alfred North.

Bibliography

WORKS BY MEAD

The Philosophy of the Present, edited by Arthur E. Murphy. Chicago: Open Court, 1932. Mead's Carus Lectures. Prefatory remarks by John Dewey.

Mind, Self and Society from the Standpoint of a Social Behaviorist, edited by Charles W. Morris. Chicago: University of Chicago Press, 1934. Based on Mead's lectures in social psychology. Introduction by Morris. Contains a listing of Mead's writings.

Movements of Thought in the Nineteenth Century, edited by Merritt H. Moore. Chicago: University of Chicago Press, 1936. Notes from course lectures.

The Philosophy of the Act, edited by Charles W. Morris, in collaboration with John M. Brewster, Albert M. Dunham, and David L. Miller. Chicago: University of Chicago Press, 1938. Unpublished papers and lecture notes. Introduction by Morris.

The Social Psychology of George Herbert Mead, edited by Anselm Strauss. Chicago: University of Chicago Press, 1956. Introduction by Strauss.

The Social Psychology of George Herbert Mead, Part Six: Self, edited by Anselm Strauss. Chicago, University of Chicago Press, 1965.

Movements of Thought in the Nineteenth Century, edited by Merrit H. Moore. Chicago: University of Chicago Press, 1967.

George Herbert Mead: Essays on His Social Philosophy, edited by John w. Petras. New York: Teachers College Press, 1968.

On Social Psychology: Selected Papers. Rev. ed, edited by Anselm Strauss. Chicago: University of Chicago Press, 1972.

Miller, David L. *George Herbert Mead: Self, Language, and the World*. Austin: University of Texas Press, 1973.

Selected Writings, edited by Andrew J. Reck. Chicago: University of Chicago Press, 1981.

The Individual and the Social Self: Unpublished Work of George Herbert Mead, edited by David L. Miller. Chicago: University of Chicago Press, 1982.

George's Page: A Document Repository for the Work of George Herbert Mead with Resources to Support Research on His Contribution to Social Psychology, edited by Lloyd Ward and Robert Throop. St. Catherines, ON: The Mead Project, Dept. of Sociology, Brock University, 1998–.

Essays in Social Psychology, edited by Mary Jo Deegan. New Brunswick: Transaction Publishers, 2001.

WORKS ON MEAD

Cook, Gary A. *George Herbert Mead: The Making of a Social Pragmatist*. Urbana: University of Illinois Press, 1993.

Joas, Hans. *G. H. Mead: A Contemporary Re-examination of His Thought*. Cambridge, MA: MIT Press, 1985.

Lee, Grace Chin. *George Herbert Mead: Philosopher of the Social Individual.* New York: King's Crown Press, 1945. Includes bibliography of secondary literature.

Natanson, Maurice. *The Social Dynamics of George Herbert Mead.* Washington, DC: Public Affairs Press, 1956. Discusses affinities with Husserl's phenomenology. Includes bibliography. Introduction by Horace M. Kallen.

Pfuetze, Paul E. *The Social Self.* New York: Bookman Associates, 1954. Comparisons between Mead, Buber, and psychoanalysis.

William H. Desmonde (1967)

Bibliography updated by Michael J. Farmer (2005)

MEANING

What is it for a sentence—or a substantial expression, such as a word or phrase—to have a particular "meaning" in a given language? While it is widely agreed that the meaning of a sentence, phrase, or word must have something to do with the way that the expression is used by speakers of the language, it is not at all obvious how to move from that vague idea to a precise answer to our question. One problem is that utterances of a given sentence might be used to convey all manner of messages,

many of which would be far removed from what we intuitively regard as the literal linguistic meaning of the sentence. Any account of meaning in terms of use must find a way to avoid having every innovative or idiosyncratic feature of use registered as an aspect of meaning. There are two ideas about linguistic meaning that might help with this problem. One is the idea that linguistic meaning is a matter of convention. The other is the idea that linguistic meaning is compositional; that is, the linguistic meaning of a sentence depends in a systematic way on the meanings of the words and phrases from which the sentence is constructed.

LINGUISTIC MEANING IS CONVENTIONAL

To define the meaning of a sentence as the message or messages that the sentence is, or can be, used to convey is inadequate, because too inclusive. In order to exclude the innovative or idiosyncratic features of language use, we might reach for the notion of a rule of language: What it is for a sentence to mean that *p* is for there to be a rule saying that the sentence is to be used (or may be used) to convey the message that *p*. However, if a rule is something that is formulated explicitly (in language), then the proposal may just reintroduce the notion of linguistic meaning; and that would be unsatisfactory if the project is to define or analyze the notion of linguistic meaning in other terms. So, instead of the notion of an explicitly formulated rule we can make use of the notion of a convention, defined as a rationally self-perpetuating regularity (Lewis, 1969). The resulting proposal is that what it is for a sentence S to mean that *p* in the language of a given population is for there to be a convention in that population to use utterances of S to convey the message that *p*.

LINGUISTIC MEANING IS COMPOSITIONAL

The term *theory of meaning* can be applied to two very different kinds of theory. On the one hand, there are semantic theories that specify the meanings of the expressions of some particular language; on the other hand, there are metasemantic theories that analyze or explain the notion of meaning. We should expect the idea that meaning is compositional to be reflected in semantic theories. The way in which the meanings of sentences depend on the meanings of words and phrases should be revealed in a semantic theory by having the meaning specifications for whole sentences derived logically from more basic principles that specify the meanings of words and phrases.

Many features of the messages conveyed by the use of a sentence will not be seen simply as the results of contributions to meaning made by the words in the sentence—contributions that would be repeated in other sentences—but rather as the products of interaction between the meaning of the sentence and other background assumptions. (The study of this interaction is called pragmatics. See Davis, 1991.) It is true, for example, that a letter of reference that says only, "Mr. X's command of English is excellent, and his attendance at tutorials has been regular" is likely to convey the message that Mr. X is not a talented philosopher (Grice, 1975). But this message is not the logical product of the meanings of the words and phrases used. Rather, the letter writer is able to convey that message by relying on shared assumptions about what information would be relevant in the circumstances. (See Grice's early [1961] proposals about pragmatics.)

TWO APPROACHES TO THE STUDY OF MEANING

These ideas, that meaning is conventional and compositional, can be seen at work in two important approaches to the study of linguistic meaning, on which this article focuses. One is Herbert Paul Grice's program for analyzing the concept of literal linguistic meaning in terms of psychological notions such as belief and intention (Grice, 1989). The other is Donald Davidson's project of illuminating the notion of meaning by considering how to construct compositional semantic theories for natural languages (Davidson, 1984).

GRICE'S ANALYTICAL PROGRAM

The Gricean analytical program can be regarded as having two stages (for overviews, see Avramides, 1989; Neale, 1992). The first stage aims to characterize a concept of speaker's meaning that corresponds, roughly, to the idea of conveying, or attempting to convey, a particular message (Grice, 1957, and other papers, 1989). The second stage then aims to use the concept of speaker's meaning, along with the notion of a convention, to build an analysis of literal linguistic meaning. (In fact, Grice himself did not introduce the notion of convention, but used a slightly different idea. See Grice, 1989; Lewis, 1969, 1975; Schiffer, 1972.)

The basic idea of the first stage of the program is that an agent who is attempting to convey a message—perhaps the message that it is time for tea—makes an utterance (which might or might not be linguistic in nature) with the intention that the hearer should come to believe

that it is time for tea and should believe it, at least in part, in virtue of recognizing that this is what the utterer intends him or her to believe. The analysis of speaker's meaning was refined and complicated in the face of counterexamples (Grice, 1989; Strawson, 1964; Schiffer, 1972), but it retained the crucial feature of not itself importing the notion of literal meaning. This feature is shared by the analysis of convention as a rationally self-perpetuating regularity, and so the prospects are good that the analysis of meaning resulting from Grice's program can meet the requirement of noncircularity.

PROBLEMS WITH GRICE'S PROGRAM. Grice's program does, however, face a number of serious objections. One problem concerns the application of the program to sentences that are never used at all—perhaps because they are too long or too implausible. Clearly, the Gricean analysis of literal meaning cannot be applied directly to these sentences. If we want to say that there is, nevertheless, a fact of the matter as to what unused sentences mean, then we seem bound to appeal to the meanings of the words and phrases from which unused sentences are built. But now we come to the most serious problem for the program, namely, how to analyze the notion of meaning as it applies to subsentential expressions.

Parties to a convention know what the relevant regularity is, and their belief that they and others have conformed to the regularity in the past gives them a reason to continue conforming to it. Thus, the Gricean program involves crediting speakers of a language with knowledge about regularities of use. While this is plausible in the case of the use of complete sentences, it is problematic when we move to subsentential expressions. Words and phrases are used in complete sentences, and they make a systematic contribution to the meanings of the sentences in which they occur. Regularities of use for words and phrases are regularities of contribution to the messages that sentences are used to convey. But spelling out in detail how words and phrases (and ways of putting them together) contribute to the meanings of complete sentences is a highly nontrivial project. So, it is not plausible that every speaker of a language knows what these regularities of contribution are.

The problem for the Gricean program is that it seems bound to attribute to ordinary language users knowledge that they do not really have. It may be that we can deal with this problem by invoking some notion of tacit (Chomsky, 1986) or implicit (Dummett, 1991, 1993) knowledge (Loar, 1981). But the dominant consensus—and the view of one of the most authoritative exponents of Grice's program (Schiffer, 1987)—is that the project of analyzing literal meaning in terms of intentions and beliefs cannot be completed.

DAVIDSON AND TRUTH-CONDITIONAL SEMANTICS

Any metasemantic theory can be used to provide conditions of adequacy on semantic theories. Thus, consider the Gricean metasemantic proposal:

Sentence S means that *p* in the language of population G if and only if (iff) there is a convention in G to use utterances of S to convey the message that *p*.

And suppose that a semantic theory for a particular language L delivers as one of its meaning specifications:

Sentence S1 means (in L) that wombats seldom sneeze.

Then, according to the metasemantic proposal, one necessary condition for the correctness of the semantic theory is that there should be a convention in the population of L-speakers to use utterances of S1 to convey the message that wombats seldom sneeze.

This kind of transposition can be carried out in the opposite direction too. Any condition of adequacy on semantic theories can be reconfigured as a partial elucidation of the concept of meaning—or of whatever other concept plays a key role in the semantic theory—and a great deal of philosophical work on the concept of meaning proceeds by considering constraints on semantic theories. Davidson's work (1984) provides an important example of this approach.

THE TRUTH-CONDITIONAL FORMAT. As we introduced the notion, a semantic theory is a theory that tells us what expressions mean. It is natural to suppose, then, that the key concept used in a semantic theory will be the concept of meaning, and that the format of the meaning specifications for sentences will be either:

The meaning of sentence S = *m*

or else:

Sentence S means that *p*

according as meanings are or are not regarded as entities. But Davidson (1967) rejects both these formats, and argues instead for the truth-conditional format:

Sentence S is true if and only if *p*.

His argument comes in two steps.

The first step is intended to rule out the idea that, to each word, each phrase, and each sentence, there should be assigned some entity as its meaning. This step proceeds by showing that, under certain assumptions about the assignment of entities, all true sentences would be assigned the same entity. (The argument that is used here is sometimes called the Frege argument.) Clearly, no such assignment of entities could be an assignment of meanings, since not all true sentences have the same meaning. However, it is possible to resist this first step by arguing that an assignment of meanings would not conform to the assumptions that are needed to make the Frege argument work.

Even though the first step is controversial, the second step in Davidson's argument remains important for anyone who begins by favoring the format:

Sentence S <u>means that</u> *p*.

We said that, given the compositionality of meaning, we should expect that, in a semantic theory, the meaning specifications for whole sentences will be derived from more basic principles that specify the meanings of words and phrases. But Davidson points out that the logical properties of the "means that *p*" construction raise problems for the formal derivation of meaning specifications for sentences. In contrast, the truth-conditional format is logically well understood. And from the work of Alfred Tarski on certain formal languages (1944, 1956) we can carry over methods for deriving truth-condition specifications for sentences from axioms that assign semantic properties to words and phrases.

CONDITIONS OF ADEQUACY. If what a semantic theory tells us about each sentence of a language is to be cast in the truth conditional format:

Sentence S is true if and only if *p*

then what are the conditions of adequacy on semantic theories? We have already seen an adequacy condition on the internal structure of a semantic theory; namely, that it should reveal how the truth conditions of complete sentences depend on the semantic properties of words and phrases. But what conditions must the truth condition specifications themselves meet, in order to be correct?

Tarski imposed, in effect, the condition that the sentence that fills the "*p*" place should translate (or else be the very same sentence as) the sentence S. (This is Tarski's Convention T [1956].) This condition of adequacy can be transposed into a partial elucidation of the concept of truth in terms of the concept of translation. The concept of translation is sufficiently closely related to the concept of meaning that we can move from here to a partial elucidation of truth in terms of meaning:

If a sentence S means that *p* then S is true iff *p*.

But we cannot shed any light on the concept of meaning itself without bringing in extra resources.

The key notion that Davidson introduces is that of "interpretation." We imagine using the deliverances of a semantic theory to help interpret the linguistic behavior of speakers. For these purposes, we can abstract away from the details of the format, and use deliverances in the schematic form:

Sentence S __________ *p*

to license the redescription of utterances of a sentence S as linguistic acts of saying or asserting that *p*. Now, by providing a way of understanding speakers' specifically linguistic behavior, a semantic theory can play a part in the project of interpreting, or making sense of, them. So, any constraints on the project of overall interpretation of people can be reconfigured as partial elucidations of the key concepts used in semantic theories.

Two suggestions for overarching constraints on interpretation emerge from Davidson's work. One possible constraint is that speakers should be so interpreted that what they say and believe about the world turns out to be by and large correct. This is the "principle of charity" (Davidson, 1967, 1973). The other possible constraint—widely reckoned to be more plausible—is that speakers should be so interpreted that what they say and believe about the world turns out to be by and large reasonable or intelligible. This is sometimes called the "principle of humanity" (see Wiggins, 1980).

In the imagined project of interpretation, the deliverances of a semantic theory are used in schematic form. For these purposes, at least, it does not matter whether the semantic theory uses the "means that *p*" format or the "is true if and only if *p*" format. So we can, if we wish, say that the constraints on interpretation shed light on the concept of meaning and thence—by way of the connection between meaning and truth—on the concept of truth.

MEANING AND USE

We began from the vague idea that meaning has something to do with use, and have focused on two approaches to the study of meaning, both of which lay stress upon

such notions as conveying the message that *p*, saying that *p*, and asserting that *p*. Both approaches take the basic way of specifying the meaning of a sentence to involve a "that *p*" clause, and both permit the straightforward connection between meaning and truth. However, there are other ways to develop the idea of a link between meaning and use. For example, we might regard knowing the meaning of a sentence as knowing how to use it appropriately. Or we might say that knowing the meaning of a sentence is knowing under what circumstances a speaker would be warranted in using the sentence to make an assertion. Many of these ways of linking meaning with use do not lead to specifications of meaning by way of a "that *p*" clause, and so do not support the direct transfer of elucidation from the concept of meaning to the concept of truth. It is to metasemantic theories of this kind that the term "use theory of meaning" is usually applied. Use theories of meaning are often coupled with the claim that there is nothing substantive to be said about the concept of truth (see Field, 1994; Horwich, 1990, 1995).

See also Chomsky, Noam; Davidson, Donald; Dummett, Michael Anthony Eardley; Frege, Gottlob; Grice, Herbert Paul; Intention; Philosophy of Language; Pragmatics; Reference; Semantics; Strawson, Peter Frederick; Tarski, Alfred; Truth.

Bibliography

Avramides, A. *Meaning and Mind: An Examination of a Gricean Account of Language.* Cambridge, MA: MIT Press, 1989.

Chomsky, N. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger, 1986.

Davidson, D. *Inquiries into Truth and Interpretation.* New York: Oxford University Press, 1984.

Davis, S. *Pragmatics: A Reader.* New York: Oxford University Press, 1991.

Dummett, M. *The Logical Basis of Metaphysics.* Cambridge, MA: Harvard University Press, 1991.

Dummett, M. *The Seas of Language.* Oxford: Clarendon Press, 1993.

Field, H. "Deflationist Views of Meaning and Content." *Mind* 103 (1994): 249–285.

Grice, H. P. "The Causal Theory of Perception." *Proceedings of the Aristotelian Society,* supp. vol. 35 (1961): 121–152. Reprinted in Grice, 1989.

Grice, H. P. "Logic and Conversation." In *Syntax and Semantics,*Vol. 3: *Speech Acts,* edited by P. Cole and J. Morgan, 41–58. London, 1975. Reprinted in Grice, 1989.

Grice, H. P. "Meaning." *Philosophical Review* 66 (1957): 377–388. Reprinted in Grice, 1989.

Grice, H. P. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press, 1989.

Horwich, P. "Meaning, Use and Truth." *Mind* 104 (1995): 355–368.

Horwich, P. *Truth.* Oxford: Blackwell, 1990.

Lewis, D. *Convention.* Cambridge, MA, 1969.

Lewis, D. "Languages and Language." In *Language, Mind and Knowledge,* edited by K. Gunderson. Minneapolis: University of Minnesota Press, 1975.

Loar, B. *Mind and Meaning.* Cambridge, U.K.: Cambridge University Press, 1981.

Neale, S. "Paul Grice and the Philosophy of Language." *Linguistics and Philosophy* 15 (1992): 509–559.

Schiffer, S. *Meaning* (1972). Oxford: Clarendon Press, 1988.

Schiffer, S. *The Remnants of Meaning.* Cambridge, MA: MIT Press, 1987.

Strawson, P. F. "Intention and Convention in Speech Acts." *Philosophical Review* 73 (1964): 439–460.

Tarski, A. "The Concept of Truth in Formalized Languages." In *Logic, Semantics, Metamathematics,* edited by A. Tarski, 152–278. Oxford: Clarendon Press, 1956.

Wiggins, D. "What Would Be a Substantial Theory of Truth?" In *Philosophical Subjects,* edited by Z. van Straaten, 189–221. Oxford: Clarendon Press, 1980.

Martin Davies (1996)

MEASUREMENT AND MEASUREMENT THEORY

Metrology in general and measurement theory in particular, have grown from various roots into fields of great diversity in the natural and social sciences, engineering, commerce, and medicine. Informally, and in its widest empirical sense, a measurement of a property, exhibited by stereotype objects in variable degrees or amounts, is an objective process of assigning numbers to the objects in such a way that the order-structure of the numbers faithfully reflects that of degrees or amounts of the measured property. Measuring instruments with pointers and calibrated scales for reading are the basic empirical means by which numerical assignments are realized. Abstractly, a particular way of assigning numbers as measures of extents of a property in objects is called a *quantity scale*. In the natural sciences, the results of measurement on a quantity scale are expressed in the form of denominate numbers, each comprised of a numerical value (magnitude) and a physical unit. Nominalists support the view that the results of measurement are not denominate numbers but numerals and perhaps other symbols.

CLASSICAL TEMPERATURE MEASUREMENT

To illustrate this morass of preliminary definitions, consider classical temperature measurement. Temperature is a local thermodynamic property of physical substances,

linked to the transfer of thermal energy (heat) between them. From the standpoint of statistical mechanics, heat in a physical substance is a macroscopic manifestation of the random motion of the constitutive atoms or molecules. An increase of temperature in the substance matches the increase of rate of molecular motion, so that temperature can be rigorously conceived as a measure of the kinetic energy of molecules.

It is important to emphasize that classical temperature measurement does not depend on any of these deep underlying physical theories. In 1592 Galileo Galilei was able to measure temperature in a theory-independent way, using the contraction of air that drew water up a calibrated tube. Approximately a century later, Daniel G. Fahrenheit invented the mercury-in-glass thermometer, again without understanding energy conservation laws that were discovered and firmly established only after 1850. These remarks, however, are not all that obvious and must be taken with a grain of salt. Precise construction of thermometers and their calibration certainly relies on theories of heat and the correct representation of (freezing and boiling) reference points. Immediately a foundational question arises: Is measurement theory-laden? The answer to this question is subtle and depends on how measurement is modeled. Because modeling of numerical quantification of measurable properties makes no commitments to and assumptions about quantitative laws and substantive scientific theories, a straight answer must be in the negative. However, measurement theory addresses many issues that go well beyond the construction of quantity scales, including prominent relationships among quantity scales of measurable properties, studied by well-established scientific theories.

From the inception of quantifying temperature and other variable properties, the concept of measurement has proved to be a steady source of methodological difficulties. For example, it would be false to conclude that today it was twice as warm as yesterday because today the local temperature at noon was balmy ninety degrees and it was only forty-five degrees yesterday. The inference may appear correct because on the Fahrenheit scale indeed there is 90°F = 2 × 45°F. But to the opposite effect, a meteorologist equipped with a Celsius thermometer observed at the same site that the temperature today was 32.2°C and it was 7.2°C yesterday, inferring that today's temperature was approximately 4.6 times higher than yesterday. Based on the familiar conversion formula *b*°C = 5/9(*a*°F – 32) from Fahrenheit to the Celsius scale, the meteorologist quickly obtains the equalities 32.2°C = 90°F and 7.2°C = 25°F, further corroborating that today's temperature on the Celsius scale is not twice as high as it was yesterday. Simple physical experiments show that it is not meaningful to make scale-independent comparative statements of the form above —"yesterday was *n* times as warm as today," if the temperature is measured traditionally on an *interval* scale (including Celsius, Fahrenheit, Reaumur, and Rankine) in the sense of Stanley Smith Stevens (1960) and the definition recalled below. Science has little use for observational statements whose truth depends on the choice of quantity scales. In all cases of quantitative observation, the main interest is in those measurement data that are invariant under scale transformations. Louis Narens discusses many other examples of a similar nature in his *Theory of Meaningfulness* (2002).

A performance of any empirical observation is usually a complex activity that is impossible and (fortunately) unnecessary to report completely. The structure of a measurement-based observation that an experimenter is able to extract and analyze formally with some success is best captured by a measurement model. For example, in the simplest and best-known physical situation of temperature measurement, the experimenter assumes that the temperature-bearing entities (e.g., substances in vessels) can, at least conceptually, be identified and distinguished one from another, and then appropriately labeled or described. As common in other branches of mathematics, the experimenter next conceives of collecting such labels or mathematical descriptions of substances into a set, to be called a *measurement domain* and denoted M. Because this domain furnishes a mathematical basis for modeling the scale structure of measurable properties, care must be exercised in its selection. To simplify the preceding pedantic language in what follows the discussion will often refer to M as a domain of substances, objects, or events, when in actuality we mean a set of their mathematical labels or descriptions.

Galileo and Fahrenheit were able to order effectively many substances at given time instances in accordance with their exhibited degrees of the temperature property, here denoted t, without recourse to any antecedently established thermodynamical theories. This suggests that the scaling model of temperature measurement should be based on a designated comparative relation $\leq_t$, where the associated atomic formula "$x \leq_t y$" is meant to express that substance y is at least as warm as substance x, for all substances x and y belonging to the underlying domain M.

THE MEASUREMENT MODEL

The resulting deceptively simple *measurement model*, commonly symbolized by the ordered pair $(M, \leq_t)$, captures the ordering of substances with respect to degrees of their temperature property t at a specified time instant. It should be clear that a similar model can be used to characterize the comparison of substances with respect to their mass property. In many measurement-theoretic applications, the foregoing comparative relation $\leq_t$, henceforth abbreviated to $\leq$, enjoys the following pair of measurability properties for all elements x and y in the given domain M:

(i) *Transitivity*: If $x \leq y$ and $y \leq z$, then $x \leq z$.

(ii) *Connectedness*: $x \leq y$ or $y \leq x$.

We associate with every comparative relation $\leq$ a canonical *indiscernibility* equivalence relation $\approx$, defined by

$$x \approx y \text{ iff } x \leq y \text{ and } y \leq x$$

for all x and y in M. Here the notation "iff" is a standard abbreviation for "if, and only if." Under the foregoing intended interpretation, the atomic formula "$x \approx y$" encodes the fact that substances x and y have the same degree of temperature. It should be obvious that the relation $\approx$ partitions the domain M into equivalence classes of substances, where each class contains precisely those substances whose degrees of temperature coincide.

At this point we may ask: What are measurement models good for and how do we know that they are adequate? In measurement theory, measurement models have four basic functions: upholding *numerical representation*, specifying the *uniqueness* of representation, and capturing quantitative and qualitative *meaningfulness*.

REPRESENTATIONAL ROLE OF MEASUREMENT MODELS. In their *representational* role, measurement models provide a mathematical basis for numerical quantification of extents, degrees, or amounts of measurable properties of objects. For example, in the case of temperature measurement, the possibility of numerical quantification of the variable temperature property t comes down to the existence of a *quantity scale*, rendered precise by a real-valued function, denoted $\Phi: M \to R$, that assigns to each substance x in M a unique real number $\Phi(x)$ in R (interpreted as the degree of temperature of substance x) in such a way that the numerical order in the host field (R, $\leq$) of real numbers agrees with the comparative relation $\leq$ specified in the measurement model. Formally, we have the order-embedding representational condition

$$x \leq y \text{ iff } \Phi(x) \leq \Phi(y)$$

for all x, y in M. In general, there is no guarantee that an order-embedding function Φ exists. A major task of representational measurement theory is to find a body of empirically meaningful constraints—constraining the structure of $(M, \leq)$, usually called the *representation axioms*, such that they are necessary and sufficient for the existence of a quantity scale (order-embedding function) Φ. The preceding transitivity and connectedness properties are usually included in the collection of representation axioms, but generally they are not sufficient for the existence of a quantity scale. In essence, this is the way the experimenter expects to achieve a theoretically justified passage from qualitative observations (x is t-er than y) to quantitative data that may be processed further by various computational and statistical means. It should be clear that the foregoing low-complexity measurement model is totally ineffective in characterizing the measurement of television violence, unemployment, and many other highly complicated attributes studied in the social sciences.

Not surprisingly, quantity scales (if they exist) are seldom unique. We have already seen that two arbitrary temperature measurement scales $\Phi': M \to R$ (e.g., for Celsius degrees) and $\Phi: M \to R$ (e.g., for Fahrenheit degrees) are always linked via functional composition of the form $\Phi'(x) = f(\Phi(x))$ for all substances x, where $f: R \to R$ is an affine (positive linear) *permissible transformation*, specified by $f(r) = ar + b$ with $a > 0$ for all real numbers r. From the standpoint of algebra, the totality of permissible transformations between temperature quantity scales forms a numerical affine group. In general, a property is said to be measured on an *interval* scale provided that its family of permissible transformations is the affine group. Along similar lines, a property is measured on a *ratio* scale just in case its family of permissible transformations is the similarity group of all functions $f: R \to R$, specified by $f(r) = ar$ with $a > 0$ for all real numbers r. So the apparent relativism and arbitrariness in the choice of measurement methods and accompanying quantity scales are factored out by invoking pertinent scale-transformations. In addition to guaranteeing the existence of a quantity scale, representation axioms specify the correct group of permissible transformations between scales. Thus if the experimenter intends to draw conclusions about objective temperature values, he or she must consider the associated affine group of scale-transformations and ensure that they preserve all numerical relationships of interest.

DETECTION OF MEANINGLESS OBSERVATIONAL STATEMENTS. Measurement models are instrumental in detecting meaningless observational statements; meaningfulness has long been a favorite of measurement theorists. We begin with the simplest characterization. Given a binary numerical relation ρ on the real line R, we say that ρ is *quantitatively meaningful* for the measurement model (M, ≤) just in case for all quantity scales Φ', Φ: $M \to$ R the equivalence

$$\Phi'(x)\ \rho\ \Phi'(y) \text{ iff } \Phi(x)\ \rho\ \Phi(y)$$

holds for all elements x and y in M. It is easily seen that this definition automatically generalizes to n-place relations. For example, for any pair of temperature scales Φ' (e.g., Celsius) and Φ (e.g., Fahrenheit) the equivalence

$$\Phi'(x) - \Phi'(y) < \Phi'(z) - \Phi'(w) \text{ iff } \Phi(x) - \Phi(y) < \Phi(z) - \Phi(w)$$

holds for all substances x, y, z, and w. The concept of quantitative meaningfulness is extremely useful in determining the applicability of statistical concepts (including sample averages and standard deviation) in the world of measurement data.

There is a closely related concept of qualitative meaningfulness that is based on the notion of automorphism. Recall that an order-embedding map α: $M \to M$ of the domain of a measurement model (M, ≤) to itself is called a measurement automorphism precisely when it is one-to-one and onto. Briefly, a binary relation ρ on the measurement domain M is said to be qualitatively meaningful for the model (M, ≤) provided that for each measurement automorphism α: $M \to M$ and for all x and y in M the equivalence

$$x\ \rho\ y \text{ iff } \alpha(x)\ \rho\ \alpha(y)$$

holds. Less formally, a binary relation ρ on M is measurement-theoretically meaningful for (M, ≤) if the exact identity of ρ-related objects is irrelevant. The only thing that matters is that the objects in M possess the measured property in equal amounts. In general, quantitative and qualitative meaningfulness are not coextensive. The notion of qualitative meaningfulness is important in delineating the class of model-definable relations. It is easy to check that the omnipresent indiscernibility relation ≈ is qualitatively meaningful for (M, ≤).

REPRESENTATION AXIOMS. Finally, in addition to securing a quantity scale and its uniqueness (up to permissible transformations), representation axioms of a measurement model can also be viewed as capturing the overall *empirical content* under consideration, encountered in testing the measurement model's adequacy. In this context, measurement axioms are usually classified into *rationality* (design) axioms (including transitivity)—assumed to be automatically true under the intended interpretation; *structural* (technical) axioms (e.g., the Archimedean axiom), crucial in establishing powerful representation theorems; and various testable *empirical* axioms, characterizing (often in a highly idealized way) specific measurement methods.

To appreciate the striking simplicity of measurement models, it is important to realize that these models represent the observational structure of a measurable property in such a way that most of the empirical detail of the actual observation is ignored. Here the experimenter is interested only in a basic abstraction that is based on comparisons of extents of given measurable properties, sufficient for a suitable order-preserving numerical quantification.

REPRESENTATIONAL THEORY OF MEASUREMENT

Measurement theory in general (as a branch of applied mathematics) and representational measurement theory in particular, are mainly based on work summarized in *Foundations of Measurement* (vol. 1, 1971) by David Krantz and others; *Foundations of Measurement* (vol. 2, 1989) by Patrick Suppes and others; and in *Foundations of Measurement* (vol. 3, 1990) by Duncan Luce and others. These authors use a model-theoretic (semantic) conception of empirical theories. In brief, instead of conceiving measurement theory as a deductively organized body of empirical claims, the semantic conception views a theory as a way of specifying a class of set-theoretic relational structures that represents various aspects of reality. The principal objectives of measurement theory are the study of set-based models of measurable properties of empirical objects, maps between them, and the representation of measurement models in terms of convenient numerical structures, with special regards to the relationships between the latter and affiliated quantitative theories of empirical objects.

Representational measurement theory studies many species of measurement models. In his *Physics: The Elements*, Norman Campbell (1920) noted that in modeling *extensive* properties (including, e.g., length, area, volume, mass, and electric charge), the above specified order-theoretic measurement model (M, ≤) has a powerful algebraic enrichment, typically symbolized by (M, ≤, ∘), where ∘ is a binary composition operation on M, satisfy-

ing the following partially testable empirical conditions for all x, y, z, and w in M:

(i) *Commutativity*: $x \circ y \approx y \circ x$.

(ii) *Associativity*: $(x \circ y) \circ z \approx x \circ (y \circ z)$.

(iii) *Monotonicity*: $x \leq y$ iff $x \circ z \leq y \circ z$.

(iv) *Positivity*: $x \leq x \circ y$ and not $x \circ y \approx x$.

(v) *Strongly Archimedeanness*: If $x \leq y$ and not $x \approx y$, then for any z and w there exists a positive integer n such that $n\bullet\ x \circ z \leq n\bullet\ y \circ w$, where $n\bullet\ x$ is defined inductively by setting $1\bullet\ x = x$ and $(n + 1)\bullet\ x \approx n\bullet\ x \circ x$.

In the case of length measurement, the measurement domain M consists of suitable and to some extent idealized length-bearing entities (e.g., straight, rigid rods) that can be properly identified and distinguished one from another. Because length measurement is modeled within a classical framework, relativistic reminders that length is not an intrinsic property of rods but something relational—relative to inertial reference frames—will not be of concern.

To measure length in a basic way, independently of any application of laws, the experimenter *operationalizes* the comparative "at least as long as" relation $\leq$ by placing two rods side by side in a straight line, with one end of the rods coinciding, and observing which one extends at the other end. In this manner the experimenter has an effective way of determining whether the relational formula "$x \leq y$" holds for virtually any pair of rods x and y in M. Of course if rod x is a physical part of rod y or is equal to y, then the validity of "$x \leq y$" is accepted by default. The composition $x \circ y$ of rods x and y is understood to be the rod obtained by the operation of placing rods x and y end to end in a straight line. Thus we take the abutted combination of rods x and y to be the whole composite rod $x \circ y$.

We know from David H. Krantz and others (1971, p. 73) that the representation axioms above are necessary and sufficient for the existence of a real-valued, order-embedding, additive scale function Φ: $M \to$ R, satisfying the representational condition

$$\Phi(x \circ y) = \Phi(x) + \Phi(y)$$

for all x, y in M. We see that the representation axioms not only justify a numerical quantification of amounts or extents of measurable properties, they capture the structure of the associated extensive measurement process itself.

In his basic concepts of measurement Brian Ellis (1966) addresses the question whether the preceding interpretation of composition operation $\circ$ is intrinsic to physical measurement of length or is perhaps just a convenient convention. Ellis points out that the representation axioms listed above remain valid even if the experimenter uses an orthogonal concatenation of rods. Specifically, this time the composite rod $x \circ' y$ is obtained somewhat artificially as a rod formed by the hypotenuse of the right triangle, whose sides are the rods x and y. Thus here the experimenter is abutting x and y perpendicularly rather than along a straight line. Not surprisingly, because the operational peculiarities of respective compositions in a straight line versus orthogonally are not visible in the representation axioms, the corresponding enriched measurement models $(M, \leq, \circ)$ and $(M, \leq, \circ')$ are measurement-theoretically indiscernible. Ellis holds a conventionalist view of measurement, in the sense that measurable properties do not exist independently of their methods of measurement.

The technical problem of "$x \circ x$" is circumvented by using an unlimited supply of copies of x (so that $x \circ x \approx x \circ y$, where $x \approx y$) or by passing to a partial composition operation. Ontological objections against using models with infinitely many objects are obvious. Another problem is whether the comparative relation $\leq$ and composition $\circ$ of a measurement model $(M, \leq, \circ)$ are directly observable. Scientific realists in particular argue that in general the representation axioms treat the empirical structures of measurement models as something decisively theoretical.

There are several ways to develop a general theory of *derived* measurement. In some ways the most natural place to start is with the notion of fundamental measurement, covered earlier. A measurable property is said to be *fundamental* or *basic* provided that its measurement does not depend on the measurement of anything else. Simply, a measurement theorist starts with a measurement model $(M, \leq, \circ)$ of a basic property together with the characterizing representation axioms and then proves the existence and uniqueness of the quantity scale. No other measurement models are needed.

In contrast, a *derived* measurable property is measured in terms of other previously established quantity scales and measurement models. A classical example in physics is density, measured as a ratio of separate measurements of mass and volume. To avoid conceptual confusion, it is not suggested that a fundamental measurement of density is impossible. When mass and volume are known, there are offsetting advantages to working

with a derived notion of density. Another question is whether any measurement is truly basic.

A BRIEF HISTORY OF MEASURING DEVICES

It is invariably difficult to trace the origins of measurement devices. Weights and measures were among the earliest tools, invented and used in primitive societies. Ancient measurements of length were based on the use of parts of the human body (e.g., the length of a foot, the span of a hand, and the breath of a thumb). Time was measured by water clocks, hourglasses, and sundials.

The earliest weights were based on objects frequently weighed (e.g., seeds, beans, and grains). Comparisons of capacities of containers were performed indirectly by filling gourds and vessels with plant seeds—which were later counted—and water. These *qualitative* measurement methods, used in conjunction with crude balance scales, formed a basis of early commerce. There was an enormous proliferation of local and national measurement systems and units (e.g., Egyptian around 3000 BCE; Babylonian around 1700 BCE; Greek in 500 BCE; and Roman around 100 BCE). Romans adapted the Greek system that was later adopted with local variations throughout Europe as the Roman Empire spread. As these methods of associating numbers with physical objects were growing, it became possible to compare the objects *abstractly* by comparing the associated numbers and to combine them by manipulating numbers. In the presence of standardized units accepted by the whole community it became possible to replace accidental comparatives of the form “five times the width of my finger” with more universal but still unit-dependent “3.75 inches.”

In England in the early thirteenth century, measures and weights (strongly influenced by the Roman system) quickly evolved along the lines of strict standardization. In France, standardization of measures and weights came several centuries later. In 1670 Gabriel Mouton, a French priest, proposed the establishment of a decimalized metrology of weights and measures. The unit of length that was finally decided on was one ten-millionth part of a meridional quadrant of the earth. Weight of a cubic decimeter of distilled water at maximum density temperature of 4°C was adopted as the kilogram. (During the second half of the twentieth century there was a shift away from standards based on particular artifacts toward standards based on stable quantum properties of systems.) The adoption of the metric system in France and generally in Europe was slow and difficult, until the International Bureau of Weights and Measures, formed in 1875, recommended the universal adoption of the MKS metric system in European countries that was subsequently signed in seventeen states. In the modern SI (Système International d’Unites) version of the metric system, there are seven base units (length, mass, time, temperature, electric charge, luminous intensity, and phase angle) from which all other units of measurement are derived.

One impressive feature of modern science is the rapidity with which new measuring instruments are being developed. For example, in the case of time measurement, and starting from imprecise ancient water clocks and hourglasses, people in the Middle Ages built town clocks (maintained by hand) to display local time. In 1656 Christian Huyghens built the first accurate pendulum clock; less than a century later John Harrison presented the first nautical chronometer. In 1928 Joseph Horton and Warren Morrison built the first quartz crystal oscillator clock. And finally, in 1950, Harold Lyons developed an atomic clock based on the quantum mechanical vibrations of the ammonia molecule. Cesium atomic clocks measure time with an accuracy of 10^{-15} seconds.

Experimental science has progressed thanks in great part to the speedy development of highly accurate measuring devices in nearly all branches of science, engineering, and medicine. The symbiotic relationship between theoretical research and measurement methodology continues to be a fundamental factor in the development of science. Philosophically, measurement is important because it provides empirical foundations for the construction of quantitative scientific theories, necessary for reliable prediction and explanation of vast categories of empirical phenomena.

See also Decision Theory; Experimentation and Instrumentation; Quantum Mechanics; Suppes, Patrick.

Bibliography

Campbell, Norman. *Physics: The Elements*. Cambridge, U.K.: Cambridge University Press, 1920.

Ellis, Brian David. *Basic Concepts of Measurement*. London: Cambridge University Press, 1966.

Heisenberg, Werner. *The Physical Principles of the Quantum Theory*. Translated into English by Carl Eckart and Frank C. Hoyt. Chicago: University of Chicago Press, 1930.

Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement*. Vol. 1: *Additive and Polynomial Representations*. New York: Academic Press, 1971.

Luce, R. Duncan, David H. Krantz, Patrick Suppes, and Amos Tversky. *Foundations of Measurement*. Vol. 3: *Representation, Axiomatization and Invariance*. New York: Academic Press, 1990.

Narens, Louis. *Theories of Meaningfulness*. Mahwah, NJ: Erlbaum, 2002.

Stevens, Stanley Smith. "On the Theory of Scales of Measurement." In *Philosophy of Science*, edited by Arthur Danto and Sidney Morgenbesser, New York: Meridian, 1960.

Suppes, Patrick, David H. Krantz, R. Duncan Luce, and Amos Tversky. *Foundations of Measurement.* Vol. 2: *Geometrical, Threshold and Probabilistic Representations*. New York: Academic Press, 1989.

Zoltan Domotor (2005)

MEDIAVILLA, RICHARD OF

See *Richard of Mediavilla*

MEDICAL ETHICS

A basis for medical ethics can be found in the Hippocratic oath. These ethics, in sum, emphasize that doctors should keep confidences, soothe their patients' suffering, and not overstep their medical abilities. The limitations of physicians set the limits of the code. With fewer limits, there are more issues to discuss: surrogate motherhood; allocation of expensive but lifesaving modalities; an emphasis on privacy and autonomy and an evaluation of the medical system itself. A caveat is necessary. The discussion of medical ethics that follows is based on the present day American system of medical practice. While much of the ethics and ethos of medicine crosses cultures, other issues may not. For example, particular questions concerning paternalism especially related to truth telling are often culture specific. Also, the American legal system, at least according to some, encourages malpractice suits against physicians leading to interesting questions about how best to practice medicine.

A standard set of topics in medical ethics are: abortion, euthanasia, confidentiality, truth telling, medico-legal jurisprudence, genetics and medicine, allocation, experimentation and informed consent, suffering, and guilt. Each area can be associated with a basic question.

Issues in medical ethics tend to arise not from questions about moral theory but from practical and clinical concerns. Failure to take this fact into account can lead to analyses that bear little resemblance to principles or rules that can be applied in clinical practice. One important difference between typical questions that arise from moral theory and those that arise in medical contexts is the lack of disinterest that one usually finds in medical contexts where a disinterested perspective is probably unrealistic. One cannot be disinterested in a beautiful but possibly battered infant. One cannot be disinterested in the pain and suffering of a terminally ill patient in virtually unmitigatable pain who asks to be allowed to die. But even if disinterested, a physician need not, therefore, be uninterested or uncaring. Indeed, physicians almost always have emotional investment in cases such as these. Whether they should or should not is another issue (a question that concerns medical education and human nature), but they do.

Even so, the moral principles appealed to are traditional ones. Do not cause pain unnecessarily. Keep promises and tell the truth, except when obvious harm will result from doing so. Do not interfere with the lives of people unless they ask for this sort of help. Do not be so selfish that the good of others is never considered. Thus, despite the glittering high technology of the modern day hospital, the dramatic emergency room, the life and death feeling of the neonatal intensive care unit, the vulnerability often felt in the examining room, medical ethics is still ethics. What follows is a description of some central issues in medical ethics.

PATERNALISM AND THE GEORGETOWN MANTRA

To say that *A* acts paternalistically toward *B* involves five beliefs on the part of *A* about the action aimed at *B*: (1) It is done for the good of *B*; (2) *A* is qualified to perform the act; (3) the action violates a specifiable moral rule; (4) the most important factor is the good of *B*; (5) *B* believes that no outside help is needed. Justifying a paternalistic action requires that it be clear that *B* would be irrational not to want the action forced and that *A* be willing to accept as a general rule something such as, "In all cases like this, a paternalistic action is allowable" (see Bernard Gert's and Charles Culver's *The Justification of Paternalism*; for an overview of the issue see the entry under *paternalism* in the Stanford Encyclopedia of Philosophy, available from http://plato.stanford.edu/entries/paternalism/).

The four part approach to medical ethics, often referred to (after the home of its proponents) as the Georgetown mantra suggests that all medical ethics decisions can be seen from the standpoint of playing off autonomy, beneficence, non-maleficence, and justice one against the other with the goal being, in each case, to get just the right balance. The four parts represent principles: respect persons rights to decide for themselves; help those

in need; avoid harming others; fair treatment, given what is owed. The mantra, popularized in *The Principles of Bioethics* by Tom Beauchamp and James Childress, was presented as a midlevel set of principles between theory, from which they were derivable, and practice. Whether using an approach based on the four principles overemphasized the application of principles to the detriment of, and need for, an overarching theoretical approach, is an ongoing debate, one especially relevant to the pedagogy of medical ethics (see Koppelman [1999]).

Because physicians make ethical decisions to some extent based on their medical school courses in medical ethics, pedagogy has always played an important role in medical ethics. Initially, most medical ethics courses were based on extrapolations from an analytic approach to ethics. There are at least two other approaches. One stresses phenomenology, the other stresses the view that patients are best understood in terms of their unfolding stories or narratives thus diminishing the role of analytic type approaches to medical ethics. The use of literature in teaching medical ethics is a natural consequence of seeing medical ethics in this manner. This essay shall discuss neither the narrative approach to medical ethics, the phenomenological approach, nor the pedagogy of medical training (on the narrative approach, see Howard Brody's *Stories of Sickness* and A.H. Hawkins's *Literature, Medical Ethics, and Epiphanic Knowledge*; for the phenomenological approach, see Zaner [1981]); for pedagogy, see the journal, *Academic Medicine*; for a critique of some uses of literature as well as a defense of an analytic approach to medical ethics, see Zucker [2006]).

THE DOCTOR-PATIENT RELATIONSHIP

The issue of paternalism is closely related to questions about the norms governing the doctor-patient relationship. Different models have been proposed to characterize this relationship. Most are based on some version of a contract and so rights are important. The business model: Here the patient gives up rights (privacy, for example) and money. For this, the patient receives service (health care). The engineering model: the doctor as a mechanic. Just as one leaves an automobile with the mechanic after trying to describe the problem, the patient tells the mechanic-physician what seems to be wrong and, in effect, leaves. Here, once the physician knows the problem, the patient is treated more like an automobile and less like a person. The patient trades the right to be treated like a person for a tune-up from the doctor—in the hope that this is the best route to running smoothly.

The priest to supplicant model: The doctor has access to important information to which the patient has no access. On this model, getting better is like having one's soul saved by a priest. Staying within the church requires that you follow the rituals required of you by the priest. Getting better requires that you follow the doctor's instructions. On this view, self-help programs would be discouraged. On the collegial model, the stress is on the partnership between the physician and patient. They are partners with a common goal: the health of the patient. On this model, each side trusts the other; each has confidence in the other. The physician suggests treatment, the patient agrees or says why not, so that a compromise can be reached.

The covenant model is not based on a contract. It stresses the dedication of the physician to the goals of medicine. Among the highest of these goals are eliminating disease and alleviating pain. The covenant model focuses on trust, concern, and sympathy. It emphasizes the caring relationship. To many, the appeal to such ideals characterizes the medical profession.

These medical models are ambiguous in the following sense. Are they descriptive or normative? These models are not meant to be an exact replica of reality. Rather, each is, to some extent, heuristic; meant to highlight an aspect of doctor-patient relations making them easier to analyze (on doctor patient models, see E. J. and L. L. Emanuel's *Four Models of the Physician-Patient Relationship*).

ABORTION

The ethical questions concerning abortion have to do with the justification of killing in a medical context. The first line of defense permitting abortion is the claim that what is killed is not the sort of thing that is (or should be) protected by traditional rules against killing. A second line of defense is seeing abortion as a help to the pregnant woman who wants the abortion. It is even possible to see abortion as a help to a fetus whose life, if not aborted, would be one of pain, degeneration, and death (e.g., infants with Tay-Sachs disease). A third line of defense views abortion as a public health issue. That is, history shows that some pregnant women will seek abortions. If abortions were illegal or very difficult to get, only the rich would be able to get safe abortions. This would be unfair as well as pose health risks to the poor. In a situation, where abortion is contemplated as an option, the question from medicine's standpoint, whether explicit or implicit, is: What is the moral status of a fetus (see entry on "Abortion")?

EUTHANASIA

There are situations where a physician might be asked (desire) to let a patient die, might be asked (desire) to help a patient to die, or might be asked (desire) to outright kill a patient. The usual reasons are unmitigatable pain (except through rendering the patient unconscious); irrevocable loss of meaningful consciousness (permanent vegetative state); irrevocable loss of some ability held so dear to the patient that death is preferable (see entry on "Euthanasia").

While it may be rational to prefer death to constant, unremitting pain, it still may be unethical for a physician to allow such a patient to die (by withdrawing or never starting life-sustaining therapy) when that patient can be kept alive. It should be noted that sometimes, the pain referred to is not so much the pain of physiology gone awry as it is the emotional distress caused by the loss of quality of life. That is, a return to baseline may not be possible and, to some people, a new and restricted life is not worth living.

The blunter version of the euthanasia question is: Should a physician kill a patient under any of these circumstances even with the permission of the patient, even where the patient begs to be killed or allowed to die? Writing a prescription for a lethal drug dose and giving it to a patient knowing that it will be used to commit suicide is considered physician assisted suicide. Some consider it a violation of medical ethics. Even if care is taken in establishing the legal and moral rules for physician assisted suicide, this can still be seen as irrelevant to the ethical evaluation. Appeal to the medical tradition does not support assisted suicide as a legitimate form of practice but there is no reason to think that tradition must be obeyed, that no new traditions can be initiated. The clearest example of traditions changing is the shift toward autonomy and consent in medicine—paternalism certainly had been the rule.

It has been argued that medicine has no room at all for intentional killing or letting patients die (see Thomasma and Pellegrino [1993]). The argument can be supported by religion but it need not be. The argument can be based on the nature of the medical profession and what most patients come to expect from physicians. The argument—by no means an uncontested one—is that letting physicians kill patients (or allowing physicians to let patients die) would erode patients' trust that nothing will be done to them that is not in their best interests. The argument goes on to claim that allowing physicians to kill some patients will create nagging suspicion: Will I be next?

The profession of medicine is dedicated to preserving life and make it better. Therefore, medicine should not aim at ending life. Here there is a clash between individual patient rights and physician rights to discharge what may be seen as the obligations of the profession.

The Council on Ethical and Judicial Affairs of the American Medical Association (AMA) updated its Do Not Resuscitate (DNR) guidelines to include two reasons for withholding Cardio-Pulmonary Resuscitation (CPR): (1) The action would be medically futile; (2) the patient has requested no CPR. The guidelines also suggest that physicians talk to their patients about the possibilities of cardiac arrest and the need for CPR. The idea is to have an informed patient taking an active part in the decision-making. Physicians—the AMA guidelines say—are obligated to honor the wishes of the patient (or named surrogate) except where it is clear to the physician that the CPR would be futile. The definition of *futile* is: (a) unlikely to restore cardiac or respiratory function; or (b) unlikely to achieve stated patient goals.

The guidelines allow the physician to enter DNR in the record because of futility but only if the patient or surrogate is fully informed. *Fully* includes explaining why and what the alternatives are if the patient still wants CPR. Of course, sometimes it is not the patient who wants everything done. Sometimes, it is a family member.

Part (b) of the AMA suggested definition of futility (viz., not likely to achieve stated patient goals) would allow for a patient to demand CPR for just a few hours more of life when that, but only that, was likely to occur. This can be seen as counterproductive in that it is a waste of resources and offers false hope to patients (Lo [1991] offers a standard defense of this view). Judging a hope false on allocation grounds may well beg a question against the role of autonomy in medical practice.

CONFIDENTIALITY AND TRUTH TELLING

Confidentiality goes hand-in-hand with privacy and truth telling. During a visit, a physician may ask personal questions such as "Are you sexually active?" Physicians expect truthful answers. Truthfulness is insured by the tacit understanding that answers will be kept private and used only to help the patient. Where the clear well-being of a third party (or parties) is jeopardized by keeping a confidence, there is at least the presumption that the confidence can be violated (on this, see *Tarasoff v. The Regents of the State of California*). Contagious diseases are just one kind of example. People with seizure disorders and drivers of public vehicles who have high blood pressure

would not have their driver licenses suspended if physicians never reported this information. These sorts of cases bring up a related question.

Should a physician be put in a situation where privacy and confidentiality are likely to be compromised? Physicians working for industry or for government can be in a situation where they are expected to reveal what would otherwise be kept confidential. In cases like these, what counts as a confidence is determined by the sort of physician one is. Physicians doing health exams for insurance companies or school boards cannot keep certain conditions private. Physicians working for factories are expected to identify malingerers. Should these be seen as violations of confidentiality? Do they undercut the very professionalism of the physician? A true malingerer does hurt everybody by collecting undeserved benefits. But should it be the role of any physician to protect the economic interest of a company and its workers?

Physicians who work for the armed services or as team physicians in organized sports can find themselves in the odd situation of patching someone up in order to have that person go back into battle or back onto the playing field only to risk more injury. Some physicians in the armed forces may find themselves as consultants to interrogators. The justification here is that in this capacity the physicians are behavioral scientists and therefore freed of their usual ethical obligations because those obligations are based on clinical medicine (see Bloche and Marks [2005] for an analysis of this type situation). Are such physicians in conflict with the higher goals implicit in the covenant view of the doctor-patient relationship? Put another way, is the covenant view of the doctor-patient relation, even if meant merely as normative, a realistic normative picture? What are realistic values for the medical profession? This question is the crux of medical ethics.

"Should physicians ever not tell the truth?" is a question related to the justification of paternalism. The usual context for questioning the necessity of truth telling is along the lines of withholding some information that the physician knows the patient (or a third party) would like to know (e.g., your son has a sexually transmitted disease [STD]); or deflecting a question such as "What do you think it is, doctor?" because the doctor thinks the answer is not one that the patient really wants to hear. Where truth telling and confidentiality conflict, confidentiality almost always will take precedence. Whether it should, is another question. The nondirective counseling favored by most genetic counselors may sometimes be open to being interpreted as withholding truthful replies.

MEDICO-LEGAL JURISPRUDENCE

There is no issue in medical ethics that does not have a legal version of it—a case brought to court. The theory behind most decisions is personal injury law. In medical malpractice, one must show damage that was caused by care that was less than standard.

There have been many cases that can be considered to be landmarks. *Tarasoff v. The Regents of the State of California*, decided in 1976, found that a psychiatrist was negligent in not warning a third party that she might be at risk from a patient. This decision changed the form of consent in psychiatry and clinical psychology limiting the confidentiality that can be offered a patient in therapy. Less dramatic but almost as far reaching is *Helling v. Carey*, which helped determine standards of care against which to judge physicians; on surrogate motherhood; *in the Matter of Baby M*; on abortion, *Roe v. Wade*; on brain death and persisitent vegetative state, *In the Matter of Karen Quinlan, An Alleged Incompetent* and *Cruzan v. Director, Mo. Health Dpeartment*; on privacy, *Griswold v. Connecticut*; on informed consent, *Canterbury v. Spence*.

Medical malpractice has an allocation aspect to it. Some specialties are sued much more than others. The usual reasons cited are the high-risk patients seen and the high expectations of many of these patients (here is an overlap of consent and malpractice; appropriate consent should include a realistic statement of expected outcomes). Rather than continue paying for high malpractice coverage to insurers and rather than risk what they take to be unfair assaults on their integrity, specialists will retire early or relocate to areas with low malpractice rates. Legislation proposed to limit awards in malpractice cases can be seen as trying to limit suits filed. But such legislation can also be viewed as aiding insurance companies who cover physicians (as well as aiding less than fully competent physicians).

GENETICS AND MEDICINE (GENOMIC MEDICINE)

Until the recent successes of the human genome project, issues in medical genetics revolved around genetic counseling and a what now might be termed proto-genetic engineering. Patients, sometimes referred to as clients in the genetic counseling context, almost always ask: Why did this happen to me? Should I have another child? What do you think this is? Directive counseling would answer these questions explicitly, sometimes before they were asked. Nondirective counseling deflected them as best as possible. The justification for the nondirective approach is that any directive counseling smacked of paternalism,

at the least, and eugenics, the attempt to change the gene pool through selective breeding of humans, at the worst. The nondirective approach grew out of the fact that advances in genetics that made genetic counseling a viable specialty coincided with the connection made between the eugenics movement in the United States and the use of eugenics in Nazi Germany. Eugenics is implicit in any directed program of genetic counseling and prenatal diagnosis along with selective abortion, thus the preference for nondirective counseling (negative eugenics weeds out unwanted genetically controlled traits; positive eugenics encourages the proliferation of desired genetically controlled traits).

The major question connected to genetics via the Human Genome Project is in what ways would we like to be better—and just how much better? And not just for us, for our progeny. Talk about what the good life is and even how best to reach it has a long history. But now there is promise that it is attainable via genetic engineering, that we will be able to choose or redesign our genes so that we will have more control over our ability to live the good life. Nurture plays a role but having the possibility of controlling the raw material of nature gives us a head start on nurture. We can be taller, shorter, thinner, more muscular, more musical, more mathematical, and so on. Again, even a head start is better than the level playing field—if these are our goals. John Rawls proposed that because what he termed *natural assets* are not distributed according to moral worth, a principle of redress was needed as a way to compensate people slighted by the natural lottery. Such a principle of redress would have to be implicit in the control over the natural lottery (on this whole topic, see Buchanan, Brock, and Daniels [2000] as well as Rawls [1971]).

Cloning humans, cloning stem cells, methods for prenatal genetic selection (including genetically engineering our progeny) raise issues that reflect those from abortion, euthanasia, privacy, and allocation. Answering: "What sort of person do we want our child to be?" or "What sort of people do we want in general and how much should we spend to get them?" are variations of age-old ethical issues. If some genetic changes are actually crucial to what we are as humans, then there are issues of defining personhood involved.

ALLOCATION

Allocation issues are divided into microallocation (who gets what) and macroallocation (how should health care itself be distributed). These two questions straddle the line between economics, social and political philosophy, and ethics. The question is one of a proper distribution of goods, where some baseline version of health is a minimal good and maximum health is the maximum good. Any decision of how to distribute these goods will also determine in part what we take the profession of medicine to be. Given that resources—time and money, as well as organs, fetal tissue, hospitals, operating rooms, and so on—are limited, it is difficult to decide how to distribute health care in a just manner. Why should some people get more and better health care than others? It certainly does happen. Is it because of planning or is it just the luck of the draw? Should something as important as health care be left to luck? The question is how to deal with the reality and the necessity.

Daniel Callahan (2000) has argued that many of our worst allocation problems are traceable to what he terms the research agenda of medicine, an agenda to cure everything to extend life as a goal in and of itself. Daniel Callahan thinks medicine should have another major goal. He offers three alternative principles. First, research should focus on premature death, ones before sixty-five, according to the U.S. government. Callahan gives a looser formulation. He says: "[any death is premature if it occurs] before a person has lived long enough to experience a typical range of human possibilities and aspirations: to work, to learn, to love, to procreate, and to see one's children grow up and become independent adults" (Callahan 2000, p. 654).

Second, research should aim at reducing poor quality of life at the last stages of life. Third, clinicians should be persuaded that helping a patient to a peaceful death is just as important as fighting for life to the end, against all odds. Callahan says that as ideals, helping a patient to a peaceful death and fighting for life against all odds are of equal value because, in the end, we all die. It is here that this perspective on allocation overlaps euthanasia issues.

A program of allocation based on autonomy and tolerance in a laissez-faire driven economy, where economics plays an important role in health care means some people will get less health care and suffer for it. In such situations, one would be forced to say (after H.T. Engelhardt in his Shattuck Lecture of 1984) that this is unfortunate but not unfair. If, however, justice demands more of an equitable distribution of needed goods, and health is one such good, then the unfortunate begins to blend into the unfair.

The lifeboat offers an interesting model for both macro and micro allocation. How many lifeboats should any ship carry? In a crowded lifeboat, should anyone have to go overboard to save the majority? What is the best

strategy for saving the majority, for getting the most moral result? How is such a decision to be made? Should there be prearranged rules, should there be deviations allowed (many toddlers aboard, no sailors), should the rules be made during times of stress (a storm, rising seas)? For a lifeboat case, see *United States v. Holmes*.

FREE AND INFORMED CONSENT IN CLINICAL AND EXPERIMENTAL MEDICINE

The gold standard for medical experimentation is the randomized, double-blind, and placebo controlled experiment with a statistically predetermined cutoff point. There is no such thing in clinical studies or in science in general as definitive results, per se. All results are definitive enough, against a background of assumptions and goals. Design and ethics go together. A poorly designed experiment will waste resources and, where there is risk, will put subjects at risk for no good reason. Consent is an ethically necessary part of any experiment. The consent must be free and informed. Subjects cannot be under so much emotional or physical pressure that they feel that they must consent. They must believe it when told that their deciding not to enter a study will not affect their treatment. This freedom from felt coercion overlaps the informed in free and informed consent because it is unlikely that someone under the previously mentioned stresses would (or could) fully understand the information given. The benefit from an experiment must at least promise some gain to the subject or future patients proportional to whatever is the risk of harm. The gain may be limited only to the knowledge that one has helped some future people.

To highlight some issues, consider work done by Dr. Saul Krugman at Willowbrook. Many children at the Willowbrook State School in New York developed hepatitis because of poor sanitary conditions. Newly admitted children were separated from other children, kept in clean quarters but fed the virus collected from infected children. Careful follow-up on these children revealed that there were two strains of hepatitis, one more communicable than the other. In defense of the experiment, it was pointed out that children were likely to get hepatitis anyway and that as subjects they received better care than they would otherwise. Parents had given consent but the reward for consent was immediate admission instead of a long wait (Munson 2003). Willowbrook exemplifies clashes between a physician's obligations to society—clean up Willowbrook; obligation to patients—find a cure or preventive for hepatitis; obligation to science—find out more about hepatitis, even if a cure is not immanent. It also highlights consent issues. How can one get truly free and informed consent for these subjects or their parents? Recent experiments utilizing genetic therapy have been halted because of excess morbidity and deaths. In these instances, there was great risk, but taking the risk was the only route to possible freedom from disease.

Free and informed consent is part of any clinical encounter as well. The principles insuring free and informed consent for subjects also apply to patients in everyday clinical situations. Patients must be treated with up-to-date therapies that are aimed specifically at their condition. Treating a contagious disease affects others but does not affect the principle that it is the patient in front of the physician who ought to be the target for therapy. Patients must be told what they are asked to accept as therapy and why. They must believe that they can ask questions as well as ask for a second opinion without jeopardizing their treatment. And, of course, risk in therapy must be proportional to gain. An often overlooked point is that some patients do not want much, if any, information. In such cases, doctors have to gauge just how little information they can safely (medically and legally) refrain from giving verbally (where consent forms are needed, information is written, and the question would be how carefully and explicitly the material should be explained to the patient).

PAIN AND SUFFERING

Sometimes, medicine can do no more than to alleviate pain. Sometimes, physicians cannot even diagnose the underlying problem. But if they can relieve pain, they have discharged what might be called a minimum obligation. This is the sort of obligation that is captured in the old saying "Above all, do no harm." Sometimes the only way to pursue this end is by listening to a patient ask, and ruminate on, the Jobian question, "Why is this happening to me?" Perhaps this aspect of medical ethics is the one that takes it furthest from traditional philosophy.

See also Bioethics; Euthanasia; Genetics and Reproductive Technologies.

Bibliography

American Medical Association Council on Ethical and Judicial Affairs. *Code of Medical Ethics*, sec. 2.035. Chicago: AMA Press, 1997.

Annas, G., and M. Grodin, eds. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*. New York: Oxford University Press, 1992.

Annas, G. *American Bioethics: Crossing Human Rights and Health Law Boundaries*. New York: Oxford University Press, 2004.

Annas, G. *Standard of Care, the Law of American Bioethics*. New York: Oxford University Press, 1993.

Annas, George. "Culture of Life Politics at the Bedside—the Case of Terri Schiavo." *New England Journal of Medicine* (2005).

Beauchamp, T., and J. Childress. *Principles of Biomedical Ethics*. 5th ed. New York: Oxford University Press, 2001.

Beaufort de, I. "Patients in a Persistent Vegetative State—A Dutch Perspective." *New England Journal of Medicine* (2005).

Bloche, G., and J. Marks. "Doctors and Interrogators at Guantanamo Bay." *New England Journal of Medicine* (2005).

Boyle, P., and K. O'Rourke, eds. *Medical Ethics: Sources of Catholic Teachings*. Washington, DC: Georgetown University Press, 1999.

Brody, Howard. *Stories of Sickness*. 2nd ed. New York: Oxford University Press, 2002.

Buchanan, A., D. Brock, and N. Daniels. *From Chance to Choice: Genetics and Justice*. Cambridge, U.K.: Cambridge University Press, 2000.

Callahan, Daniel. "Death and the Research Imperative." *New England Journal of Medicine* 342 (2000): 654–656.

Charon, R. ed. *Stories Matter—the Role of Narrative in Medical Ethics*. Philadelphia: Taylor and Francis, 2002.

Code of Medical Ethics 2004–2005. Chicago: American Medical Association Press, 2004.

Cohen, C., ed. *New Ways of Making Babies: The Case of Egg Donation (Medical Ethics Series), National Advisory Board on Ethics in Reproduction*. Bloomington: Indiana University Press, 1996.

Emanuel, E. J., and L. L. Emanuel. "Four Models of the Physician-Patient Relationship." *Journal of the American Medical Association* 267 (16) (1992): 2221–2226.

Engelhardt, H. T. *The Foundations of Bioethics*. 2nd ed. New York: Oxford University Press, 1996.

Engelhardt, H. T. "Shattuck Lecture—Allocating Scarce Medical Resources and the Availability of Organ Transplantation; Some Moral Presuppositions." *New England Journal of Medicine* 311 (1984): 66–71.

Frame, J. *Medical Ethics: Principles, Persons, and Problems (Christian Perspectives)*. Phillipsburg, NJ: Presbyterian and Reformed, 1988.

Hawkins, A. H. "Literature, Medical Ethics, and Epiphanic Knowledge." *The Journal of Clinical Ethics* (5) (1994): 283–290.

Hilyard, B. *U.S. Supreme Court and Medical Ethics: From Contraception to Managed Health Care (Paragon Issues in Philosophy)*. St. Paul, MN: Paragon House, 2004.

Hoffmaster, B., B. Freedman, and G. Fraser. *Clinical Ethics, Theory, and Practice*. Clifton, NJ: Humana Press, 1989.

Holmess, H., and L. Purdy. *Feminist Perspectives in Medical Ethics*. Bloomington: Indiana University Press, 1992.

Jonsen, A., M. Siegler, and W. Winslade. *Clinical Ethics: A Practical Approach to Ethical Decisions in Clinical Medicine*. 4th ed. New York: McGraw-Hill, 2002.

Kass, L., and J. Wilson. *The Ethics of Cloning*. Washington DC: AEI Press, 1998.

Katz, Jay, ed. *Experimentation with Human Beings; The Authority of the Investigator, Subject, Professions, and State in the Human Experimentation Process*. New York: Russell Sage Foundation, 1972.

Koppelman, L., ed. *Building Bioethics—Conversations with Clouser and Friends on Medical Ethics*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1999.

Kuhse, H., P. Singer, eds. *Bioethics*. Malden, MA: Blackwell, 1999.

La France, A. *Bioethics: Health Care, Human Rights, and the Law*. New York: Matthew Bender, 1999.

Lo, B. "Unanswered Questions about DNR Orders." *Journal of the American Medical Association* 265 (1991): 1874–1875.

Magnus, D., G. McGee, and A. Caplan, eds. *Who Owns Life?*. Amherst, New York: Prometheus, 2002.

Marquis, D. "Why Abortion is Immoral?" *The Journal of Philosophy* 86 (4) (1987).

Munson, R. *Intervention and Reflection: Basic Issues in Medical Ethics*. 7th ed. Belmont, CA: Wadsworth, 2003.

Pence, G. *Classic Cases in Medical Ethics: Accounts of Cases That Have Shaped Medical Ethics, with Philosophical, Legal, and Historical Backgrounds*. New York: McGraw-Hill, 2003.

Pellegrino, E. and D. Thomasma. *The Virtues in Medical Practice*. New York: Oxford University Press, 1993.

Plomer, A. *The Law of Ethics and Medical Research: International Bioethics and Human Rights*. New York: McGraw-Hill Humanities/Social Sciences/Language, 2003.

Rachels, James. "Active and Passive Euthanasia." *New England Journal of Medicine* 292 (1975): 78–80.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Rispler-Chaim, V. *Islamic Medical Ethics in the Twentieth Century (Social, Economic, and Political Studies of the Middle East and Asia)*. Leiden, Netherlands: Martinus Nijhoff, 2003.

Steinberg, A., and F. Rosner. *Encyclopedia of Jewish Medical Ethics: A Compilation of Jewish Medical Law on All Topics of Medical Interest*. Nanuet, NY: Feldheim, 2003.

Steinbrook, Robert. "Physician-Assisted suicide in Oregon—An Uncertain Future." *New England Journal of Medicine* 346 (2002): 460–464.

Thomasma, D., and P. Marshall. *Clinical Medical Ethics: Cases and Readings*. Boston: University Press of America, 1995.

Thompson, L. "Human Gene Therapy: Harsh Lessons, High Hopes." *FDA Consumer Magazine* (2000).

Tomlinson, T., and H. Brody. "Futility and the Ethics of Resuscitation." *Journal of the American Medical Association* 264 (1990): 1276–1280.

Veatch, R., ed. *Cross-Cultural Perspectives in Medical Ethics*. 2nd ed. Boston: Jones and Bartlett, 2000.

Veatch, R. *Medical Ethics (Jones and Bartlett Series in Philosophy)*. Boston: Jones and Bartlett, 1997.

Widdershoven, G. "Beyond Autonomy and Beneficence: The Moral Basis of Euthanasia in the Netherlands." Available from http://www.ethical-perspectives.be/viewpic.php?LAN=E&TABLE=EP&ID=52.

Zaner, R. *The Context of Self: A Phenomenological Inquiry Using Medicine As a Clue*. Athens: Ohio University Press, 1981.

Zucker, A. "Medical Ethics as Therapy." *Medical Humanities* (2006).

Arthur Zucker (2005)

MEDIEVAL AND EARLY CHRISTIAN PHILOSOPHY

In addition to the general article *Medieval Philosophy*, the Encyclopedia features the following articles having discussions of early Christian and medieval schools and movements: *Apologists*; *Augustinianism*; *Averroism*; *Byzantine Philosophy*; *Carolingian Renaissance*; *Chartres, School of*; *Gnosticism*; *Ockhamism*; *Patristic Philosophy*; *Saint Victor, School of*; *Scotism;* and *Thomism*. Particular aspects of early Christian and medieval thought are discussed in the Encyclopedia's general entries, including *Ethics, History of*; *Islamic Philosophy*; *Jewish Philosophy*; *Logic, History of*; *Metaphysics, History of*; *Mysticism, History of*; *Semantics, History of;* and *Universals, A Historical Survey*. See also *Christianity*; *Illumination;* and *Liber de Causis*. See "Medieval Philosophy" and "Christianity" in the index for entries on important figures in this area.

MEDIEVAL PHILOSOPHY

"Medieval philosophy" began with the African Christian Augustine of Hippo (354–430), whose life and writings reflected the unsettled state of the declining Roman Empire long before the commencement of the Middle Ages proper. His rich and many-sided works display the Platonic otherworldliness of his theories of knowledge and world history. According to Augustine's vision, the true cosmic plan unfolds in the history of the City of God, and the local accidents of the Earthly City are of little account in comparison. Correspondingly, true wisdom and virtue are obtainable only in the light of the Christian faith and by the prevenience of divine grace; human nature, grossly corrupted since the Fall, is in need of a correspondingly complete divine remaking. Whereas for Plato and Aristotle the fulfillment of human capacities required the possession of a high degree of sophisticated intelligence, for Augustine such fulfillment depended on rightness of the will and the affections. These two features, a radical view of the transforming power of grace and a voluntaristic accent, may be regarded as the kernel of Augustinianism, at least insofar as it affected subsequent thought. The tremendous influence of Augustine on medieval thought is matched by that of Ancius Manlius Severinus Boethius, whose grandiose plan was to transmit to the Latin West the works of Plato and Aristotle—a plan rudely cut short by his execution in 524. However, he accomplished the translation of Aristotle's logical works into Latin; his commentaries on some of them, and on the Neoplatonist Porphyry's introduction (*Isagoge*) to the *Categories* of Aristotle, were immensely influential in shaping the technical Latin vocabulary and turns of expression that prevailed in the Middle Ages, so much so that any appreciation of medieval thought must inevitably be inadequate without a thorough acquaintance with Boethius's logical output.

The intervention of the Dark Ages presented Western scholars with a gigantic task of rethinking and reconstruction. During these centuries of insecurity and uprootedness there was little intellectual endeavor, apart from the exceptional work of the Neoplatonist John Scotus Erigena in the ninth century. The logical, theological, and classical inheritance slumbered insecurely within the libraries of threatened Western monasteries. When Anselm of Canterbury (1033–1109) began to exploit Boethian logic in order to render his Christian faith intelligible, he had no immediate predecessor who in any way approached his stature as a thinker. Author of the Ontological Argument and fully alive to the power of linguistic analysis as a tool for clarifying conceptual problems, Anselm was the father of Scholasticism. Working within an Augustinian framework, Anselm and other logical theologians of the eleventh and twelfth centuries attempted to bring into order and coherence the body of doctrine to which they were committed by Holy Writ, dogmatic pronouncements, and the works of earlier authoritative church writers. The formidable dimensions of the enterprise were well known to them, as is shown in the lists of clashing antitheses made explicit in the *Sic et Non* (For and Against) of the ill-fated logician Peter Abelard (1079–1142). A systematic collection of authoritative opinions, the *Sentences*, upon which all subsequent medieval thinkers exercised their logical and philosophical ingenuity in the form of commentary, was compiled by Peter Lombard (c. 1095–1160).

While the Latin West, employing a predominantly logical Aristotelianism, was engaged in the tasks described above, as well as in controversy on the topic of universals, the more advanced Islamic civilization spreading from the Middle East possessed the whole body of Aristotle's works. These received development, commentary, and a Neoplatonic flavor at the hands of a series of subtle thinkers, among whom were al-Fārābī (c. 873–950), Avicenna (980–1037), and Averroes (c. 1126–c. 1198). From about the middle of the twelfth century on, Latin translations of their works became available; and through these, as well as through translation directly

from the Greek, Western thinkers eventually knew all of Aristotle's writings.

The Jewish philosophers Solomon ben Judah ibn Gabirol (c. 1021–1058 or 1070) and Moses Maimonides (1135–1204) also contributed to the intellectual ferment of the thirteenth century, which was accompanied by the establishment of universities within which members of the recently founded orders of Dominican and Franciscan friars were soon competing with secular masters for professorships. Generally speaking, the Dominicans, following the lead of Thomas Aquinas (c. 1224–1274), attempted to assimilate Aristotle by adopting a framework within which divine grace was seen as completing and fulfilling human nature, rather than dramatically abrogating it in the Augustinian manner. Consequently, the Thomistic tradition represented a separation, at least in principle, of philosophy from theology and a more optimistic view of human nature, society, and the civil state, coupled with opposition to those Latin Averroists who were prepared to compartmentalize their thought to the extent of claiming that on certain points philosophy (Aristotle, as interpreted by Averroes) demonstrated conclusions incompatible with their personal Christianity. Those who preferred to remain within the Augustinian stream, especially St. Bonaventure (c. 1217–1274), John Duns Scotus (c. 1266–1308), and William of Ockham (c. 1285–1349), nevertheless increasingly absorbed elements of the new Aristotelianism. Concerned as they were with the sense in which theology could be a science (a form of knowing), Duns Scotus and William of Ockham evinced a tendency to bring epistemological considerations more to the forefront of their work.

NATURE OF SCHOLASTICISM

ARISTOTELIAN EMPIRICISM: MATTER, FORM, AND SUBSTANCE. Medieval philosophy and logic are aspects of an effort to resolve conceptual puzzles (often, but not always, theologically inspired) and to underpin such resolutions with a satisfactory theory of how things are and why they are as they are. The dominant theory, although subjected to multiple variations and modifications during the medieval period, was basically Aristotelian and therefore involved an ultraempiricist effort (not always successful) to resist the abrogation of the pretheoretical commonsense aspect of the world by the theoretical. Before the consideration of any theory, whether scientific or metaphysical, human beings are inevitably confronted with a world populated by a multiplicity of diverse kinds and sorts of beings that are subject to generation, change, and death. These diverse beings are understood to the extent that "why?" questions about them or their kinds can be answered; they are the objects of evaluation insofar as they or their qualities, quantities, states, or relations are characterized as good, bad, and so on.

In accordance with the nonabrogatory policy, a technical vocabulary is required such that the pretheoretical picture does not forfeit its basic sense by relativization to a more fundamental theory that demands radical revision of that picture. For example, an ultraempiricist account of how things are must always leave place for the attribution of a literal (and not merely metaphorical) sense to questions regarding the "makings" of sense objects, states of affairs, or processes. The term *matter* represents an attempt to guarantee such a literal sense—it is the general reply to the always sensible question (in the context mentioned) "What is it made out of?" The detailed replies to such questions—"wood," "stone," "bones and flesh," "clay," "cloth," and so forth—all mention makings or materials out of which something is made, physical antecedents that are among the necessary conditions of a thing's being.

In the same context, however, explanations of why things are as they are can be given by reference to the kinds or sorts to which those things belong; for example, "Horses are self-moving because they are animals, and all animals are self-moving." Here a feature of a particular sort of being (horse) is explained by reference to its general kind (animal), and it is the notion of "form" (with its alternative medieval vocabulary, "nature," "essence," "quiddity") that represents a reminder of the fact that things fall into distinguishable sorts (species) that can in turn be subsumed under broader kinds (genera). Since truistic explanations can be given in terms of sorts and kinds, the form or essence is said to be the principle of the intelligibility, or explanation-worthiness, of things; and such general definitions as "Man is rational animal" are said to hold true in regard to the formal aspect of things. Whether or not the definitions are true of things in a scientific sense is of little import to the philosophical notion of form: Its point is to ensure the nonabrogation, by a general theory of how things are, of the pretheoretical picture of the diversity of things; realization of this point may lie behind Aquinas's agnosticism concerning the scientific value of such formal definitions.

It is plain that the replies to questions about the makings (matter) of things still involve a formal aspect, since not only are explanations in terms of the definitions of wood, stone, and the other sorts of material mentioned still possible, but it is also possible sensibly to ask what the wood or stone is made out of, or what "stuff" endures

when wine becomes vinegar. In order to do justice to such possibilities—and to the pretheoretical conviction that in processes of change the successive sorts that occur are not totally new creations but rather a sequence of diverse activizations of a common substratum—the notion of "prime" matter is employed; this is matter as mere substratum, totally devoid of any formal aspect. Prime matter was viewed schematically, by a kind of extrapolation, as pure susceptibility upon which the various formal actualities supervene, and was said to be by some medievals the principle of individuation, whereby form, the principle of intelligibility and generality, is concretized to the particularity of the various individual "this-es" that belong to a given sort. Thus, one might say that a horse is an *equinizing* of prime matter, a stone is a *petrifying* of prime matter, and so on; this use of verblike nouns helps to bring out the fact that form is act, or actuality, as opposed to the mere susceptibility of prime matter. These verblike nouns are constant, since it never makes sense to say of a horse, for example, that is it more horse or less horse (using "more" and "less" in a nonquantitative sense). Some actualizations, however, are variable, such as whiteness; one can say of a white object that it is (or becomes) more white or less white.

The real correlates of certain of the constant actualizations are called substances, objects that are pretheoretically recognized as being constantly what they are over the whole span of their existence. A horse does not become a horse, and on ceasing to be a horse, it simply ceases to be, whereas a white object can be something that becomes white in varying degrees and may cease to be white, but it is not on that account said to cease to exist. When adjectival terms such as *white* are used to denote subjects in sentences, such as "A white thing is coming down the road," it always makes sense (although in many instances it may be superfluous) to ask a question like "What is the thing that is white and is coming down the road?" This is true because such terms leave open the possibility of asking a question regarding the nature of the "something else" (*aliquid aliud*, as Aquinas has it) that is qualified (in this instance by the whiteness). When the "something else" is a substance, such as "horse," the possibility of a further question having a similar sense, but with the substance name in place of the adjective, vanishes. For example, one would not ask, "What is the thing that is a horse and is coming down the road?" Thus, this notion of substance is unlike that with which John Locke was concerned; for him it did make sense, even when a substantial sentence subject had been used, to carry on with requests for information about what he called a "something besides."

TECHNICAL LANGUAGE, MEANING, AND UNIVERSALS. Much of medieval philosophical and logical discourse involved the endowment of old words with new senses, as part of the artificialization of natural language that is characteristic of the Schoolmen, who, according to Locke, "covered their ignorance with a curious and inexplicable web of perplexed words." The Scholastics were in fact to some extent aware of the exigencies of discourse of this sort, which constitutes a kind of halfway house between the sort of philosophy that is careful to use only a completely jargon-free natural language, and the sort that is prepared to use the resources of some totally artificial language (such as those of modern symbolic logic) as a set of coordinates whereby sense and senselessness may be distinguished. When discussing the technical sense of "in" in sentences such as "Qualities inhere *in* substances," Boethius had distinguished no fewer than nine ways in which the word *in* could be used. It was clear to him that the *man* of the technical sentence "Man is a species" does not play the same role as does the name *man* in "Socrates is a man"; if it did, then one should be able to use these two sentences as premises whence "Socrates is a species" (which is false or nonsensical) could be inferred.

How, then, are such terms as *man, animal, genus*, and *species*, as they occur in sentences like "Man is a species" and "Animal is a genus," to be understood? These are sentences of a sort that must occur in the discussion of the principles of those definitions described as efforts to do justice to the formal aspect of things. Interpretation of such sentences as consisting of two names joined by *is* naturally leads to the question, transmitted by Boethius when commenting on Porphyry, of what the things are that these names name. Are the things named by such specific or generic names extramental entities additional to individual human beings and animals? An affirmative answer represents one medieval form of the option for a "realist" position in the problem of universals, and throughout the period thinkers were divided on this topic. Certain early medieval antirealists, such as Roscelin and Garland the Computist, developed a solution that had been suggested by Boethius: Words such as *species* and *genus*, said Boethius, may be interpreted as "names of names" (*nominum nomina*), so that "Man is a species" should be analyzed as "'Man' is a species," with *species* naming the word *man* and indicating that it is predicable specifically of many individuals. Herein lies one of the roots of the logical doctrine developed during the thirteenth and fourteenth centuries, the doctrine of *suppositio*.

Roscelin and Garland went further than Boethius and regarded *man* in "Man is a species" not as a mentioned name (a mentioned significant utterance) but as a mere utterance (*vox*) undergoing mention; thus St. Anselm accused Roscelin of having reduced universals to the "breath of an utterance" (*flatus vocis*). Other antirealists, observing that this extreme nominalism (as it is usually called) failed to account for the success of language as a representation of the formal aspect of things, adopted an intermediate position, according to which the universal is a natural (as opposed to a merely conventional) mental sign, or concept; such a position was designed to secure the objective reference of the universal while avoiding commitment to the plethora of extra entities demanded by realism. Abelard, Aquinas, and Ockham may be credited with having held, each in his own way, a doctrine of this type.

EXTENT OF THE ARTIFICIALIZATION OF LANGUAGE. There are several facets of the general medieval concern with the study of meaning. In the writings of Anselm of Canterbury, for example, there is an immensely powerful and pervasive realization that the overt, apparent, or grammatical form of an utterance need not show its implicit, true, or logical form—a realization whose revival has been most prominently reinitiated in our own age by Bertrand Russell. Again and again Anselm's writings contain the contrast between forms of speech that are allowed by the loose texture of ordinary language (*usus loquendi*) and the forms to which a strict attention to the exact sense (*significatio per se*) commits one; the loose texture is methodically explored, and the results of this exploration are applied to the elucidation of difficulties raised by forms of speech found in Holy Writ and ordinary language. In their technical explanations Anselm and his successors felt compelled to make innovations that violated the grammar of the natural language (Latin) in which they wrote; for instance, in expressing the objective counterparts of assertions concerning the meaning of adjectival (as opposed to substantival) words, Anselm used the novel formula "Literate is literacy," which in its Latin version (*Grammaticus est grammatica*) is about as full of scandals, from the point of view of ordinary Latin grammar, as any three-word sentence could be.

Naturally the classicists of the time, like their counterparts of the sixteenth century, took alarm at these monstrous impurities of language; a classicist rearguard action is shown in the *Metalogicon* of John of Salisbury (c. 1115–1180), who at one point explicitly argues against mixtures of abstract and concrete of the kind put forth by Anselm. A better-known example of this technical development, resulting in nonsense in respect to ordinary language, is found in Aquinas's assertion that a man *is* neither his humanity nor his existence, whereas God *is* both his essence (divinity) and his existence; these claims involve a like mixture of concrete and abstract nouns that in nontechnical speech just cannot be connected by the same "is" (or "is not").

BREAKDOWN OF COMMUNICATION. The semiartificial language of the Scholastics was excessively clumsy, and, in the absence of the precise definitional control that goes with a totally artificial language, required for its tolerably safe employment an intuitive power extending beyond the ordinary; even when this has been achieved, the history of the period demonstrates that there is no guarantee that communication will be maintained. For example, skill in the use of such language probably reached its peak in the writings of Duns Scotus, the Subtle Doctor. He rejected the theory that matter is the principle of individuation on the grounds that this attribution leaves the individual lacking in total intelligibility and even makes problematic the possibility of an omniscient being's (God's) radical understanding of the individual object. He therefore posited that individuation is performed not by a material, but by a formal, principle; for example, by "Socrateity" in respect of the individual Socrates, and in general by the "thisness" (*haecceitas*) appropriate to each individual "this." We have already observed the connection between form and intelligibility presupposed in this operation, an operation that raises a further phase of the universals controversy and at the same time exemplifies the breakdown in communication.

Ockham criticized the Scotist thing-centered formal distinction (*distinctio formalis a parte rei*) alleged to hold between the universal nature in question (humanity in the case of a human being) and the individuating formal principle (Socrateity) that makes the individual into *this* individual. Ockham was at a loss to see how this distinction could be thing-centered (*a parte rei*) and yet not commit its proponent to the admission of extra entities (humanity, Socrateity) over and above, and distinct from, individuals, in spite of the fact that the existence of universals as extra entities of this sort was denied by Scotus.

It has already been suggested that form may be best expressed by means of verblike nouns (*equinizing, petrifying*); hence, the abstract nouns often used to express formal principles could be viewed as being more verblike than namelike—a position taken by Aquinas from Boethius and apparently recognized by other Scholastics.

If this view is accepted, then the statement that the Socrateity of Socrates is distinct from his humanity may be interpreted, using appropriate verblike forms, as asserting that *Socratizing* is not identical with *humanizing*, an analysis that yields a true thing-centered distinction and yet does not send one on a vain search for extra named entities over and above the man Socrates; this offers at least one way in which the Scotist contention may be consistently understood.

But Ockham assumed, in effect, that any distinction that holds in respect of things (a "real" distinction) can only be like that which holds between, for example, Socrates and Plato and that is expressed by a sentence such as "Socrates is not Plato," wherein "Socrates" and "Plato" are names (as opposed to the verblike *Socratizing* and *humanizing*). When, therefore, Ockham encountered the further Scotist tenet that although a thing-centered formal distinction holds between Socrateity and humanity (for example), it is nevertheless not the case that a real distinction holds between the two, he assumed that "Socrateity" and "humanity" could be treated in the same way as such names as Socrates, Plato, Cicero, and Tully, and that even as the negation of a real distinction between Tully and Cicero amounts to a statement of their real identity as the same individual object, so also the denial of a real distinction between Socrateity and humanity amounts to a statement of real identity of this sort. In point of fact, however, once the verblike nature of the form-expressing words *Socrateity* and *humanity* has been grasped, it becomes clear that a denial of a real distinction between Socrateity and humanity should be understood as the rejection of any attempt to treat those form expressions as though they were pure names. The whole weight of Ockham's subsequent attack, aimed as it was at the consequence that the Scotists were in such contexts stating the denial of a real identity (one framed in terms of names, as opposed to verbs) is therefore totally misplaced.

The same blindness, combined with the theological premise that God is omnipotent, and hence can effect anything that does not involve a contradiction, also played havoc with other distinctions patiently established by earlier thinkers. For example, the distinction between essence and existence, some of whose associated theses were described above as embodying novel uses of words, was attacked on the grounds that the essence of a thing (a man's humanity) and its existence are (if a *real* distinction holds between them) two things distinct in the way that Socrates and Plato are two distinct things. In consequence, the Ockhamists considered themselves licensed to assert that the admission of a *real* distinction between essence and existence has as a consequence the possibility of God's omnipotence producing something's essence without at the same time producing its existence, or vice versa, however, this is patently absurd, and therefore (they concluded) there is no real distinction between essence and existence.

In the presence of such misplaced criticism it is obvious that scholastic thought could have been better expressed in a fully artificial language, armed with precise definitions and a greater capacity for generating and identifying new parts of speech than that of the semiartificial language that was used.

REACTION AGAINST TECHNICAL ARTIFICIALIZATION. Although the artificialization of natural language for the expression of technical truths beyond the capacity of natural language proceeded apace from the time of Anselm, the final major philosophical reaction, brought about by communication difficulties, was in the opposite direction. Ockham's attitude to the contrast between ordinary and technical discourse was the polar opposite of Anselm's attitude at the opening of the period. For Anselm, accounts of meaning could and did call for the use of, or have as consequences, technical assertions that were either nonsense from the point of view of ordinary usage, or at least involved radical departures therefrom—and his successors were similarly venturesome.

Ockham, although likewise constantly conscious of the contrast between ordinary speech and the technical forms of speech used by his predecessors, nevertheless placed propriety of expression on the side of ordinary speech, and not on the technical side, except in those instances where the novel locutions of his forerunners could be explained away or disarmed as mere stylistic ornament. His lists of sentences that are false if taken literally (*de virtute sermonis*) because words are not therein used properly (*secundum proprietatem sermonis*) are catalogs of the sort of technical assertions that for Anselm and following thinkers had been a necessary consequence of the special requirements of logical and philosophical discourse, and that for them enshrined propriety to a degree to which the looseness of ordinary speech could not aspire. This reversal of attitude, symptomatic of the breakdown of communication in terms of semiartificial language, did not, of course, immediately prevail, it was combated at great length, for instance, by John Wyclyf (c. 1320–1384). Nevertheless, Ockham's attitude, reinforced by Renaissance philology, ultimately triumphed and was represented in the strictures of Locke on "the frivolous

use of uncouth, affected, and unintelligible terms" that made philosophy "unfit or uncapable to be brought into well-bred company and polite conversation."

ETHICS AND POLITICS. Augustine's severe view of the effects of the Fall of man resulted in a largely negative view of the civil state. He held that save in the ideal case of a Christian commonwealth, earthly states are merely coercive institutions that would not exist had man not fallen, and serve simply to issue punishments and remedies for the corruption of human nature. Correspondingly, divine grace is seen by Augustine as playing a dramatically elevating part in the reformation and reordination of the will. However, the thirteenth-century revival of full Aristotelianism, coupled with the Thomist view of grace as a completion rather than an abrogation of nature, allowed that civil subordination was natural to man, would exist even if the Fall had not taken place, and hence could not be written off as an extraneous penal imposition; the state possesses a positive value in its own right.

Aquinas's enormously detailed philosophical anthropology constituted the foundation of his version of Aristotelian humanist ethics and politics, to which he attempted to give a Christian completion; it cited the perfection and fulfillment of human nature in the intellect rather than in the will: Accordingly, he viewed law as essentially a rule of right reason, rather than as a species of will-based command. This doctrine was in conflict with the teachings of the Augustinian voluntarists such as Ockham, whose view has endured through Thomas Hobbes and John Austin down to modern times. Aquinas's system of rationally based natural law as a measure of the value of human actions in general, and of human law in particular, was in opposition to the absolutist tendencies evident in the coalescence of revived Roman law with Augustinianism, which were to come to final fruition in the sovereign nation-state of our own era. The distinction between the righteous prince (who remains within the bounds of the law) and the tyrant (who puts himself above the law) had been trenchantly enunciated by John of Salisbury, was supported by the non-Roman medieval legal tradition, and clearly presupposes limits to the powers of the chief legal authority.

It is clear that Aquinas's natural-law theory supports this limiting attitude and justifies resistance to tyranny; he was therefore faced with the task of coming to terms with those features of Roman law (to be emphasized in the Renaissance) according to which the prince is above the laws. This he did by distinguishing between the coercive power (*vis coactiva*) and the directive, or rationally qualifying, power (*vis directiva*) of law: In respect of the first the prince is above the law, but in respect of the second he is voluntarily subject to it. In his theory of law Aquinas directly influenced Richard Hooker, to whom Locke admitted his indebtedness.

It is in connection with Aquinas's defense of the right of resistance, as well as in his prima facie puzzling assertions on the relation of the papacy to civil power, that we may best see how he attempted to resolve the perennial problem of the relation between political principle and political fact through the use of exceptive (*nisi forte* …) clauses. Instead of rigidly carrying through principle to the bitter end and at all costs, without any regard for concrete or historical facts (in the manner, one might say, of Plato in the *Republic*), Aquinas suggested that the most rational course would be to make appropriate accommodations with local conditions, if necessary by recourse to empirically based anticipation of the results of political action. For example, it follows from natural law that tyranny may rightly be resisted by force; this justification of rebellion may be acted upon, said Aquinas, except perhaps (*nisi forte*) when the facts of the case make it plain that the revolution will generate worse evils than the tyranny that it is designed to displace. Again, in religious matters he declared that the ecclesiastical power is to be obeyed rather than the civil, and in civil matters the lay power is to be obeyed rather than the ecclesiastical, except perhaps (*nisi forte*) in the special case of the two powers' being amalgamated in one person, such as the Roman pontiff.

Commentators discussing this last example, and not armed with a realization of the significance of its exceptive (*nisi forte*) structure, have inferred from it that Aquinas here committed himself to an extreme papalist position that would endow the pope with the fullness of spiritual and temporal power. However, once the significance of that structure has been gathered from the many other available textual examples, the conclusion may be drawn that Aquinas taught the separation of these powers as a matter of principle, yet he also observed the local fact that insofar as the pope is a temporal ruler of papal territory, he, exceptionally, holds both spiritual and temporal power. A like adaptability may be seen in Aquinas's concession that the secondary precepts of natural law are mutable in accordance with changing historical conditions and in his recommendation that laws should be tailored to fit the type of population for which they are intended; to attempt to legislate a people into full virtue is futile.

Augustinianism in general, and the Augustinian theory of law as essentially will-based command, received impetus and encouragement from the archbishop of Paris's condemnation in 1277 of certain Aristotelian theses of Arabic philosophical complexion, a condemnation that also bore upon some Thomist positions. The tendency of Averroism had been toward a pantheism that diminished the freedom of God in the act of creation. Aquinas's claim that moral evaluation consists of rational assessments based upon the intrinsic nature of the cases in question was also susceptible of being interpreted as constituting a restriction on divine omnipotence. Accordingly, Duns Scotus and Ockham, in varying degrees, claimed that the rules governing the attribution of Tightness or wrongness to human actions were contingent in relation to the absolute power of God; the consequent contingency of connection between deed and merit has caused some historians to assume that in Augustinian thought one may find the basis of Martin Luther's doctrine of justification by faith alone, as well as a source for the legal aspects of the Hobbesian theory of sovereignty.

SCIENCE AND PHILOSOPHY. Although the nonabrogatory policy of medieval philosophy outlined above served well enough to ensure that philosophers took seriously the fully human realm of reasons, purposes, hopes, and so forth, thus avoiding the split between the thinker as a human being and the thinker as a philosopher, the extrapolation of that policy's attendant ultraempiricism to sciences such as physics and cosmology tended to a greater or lesser extent to inhibit their development as practical tools. A prime and early example of such ultraempiricist inhibition is to be found in the refusal of the second-century astronomer Ptolemy to consider a suncentered planetary system because it so obviously is at variance with things as we find them to be, a refusal that was espoused by most but not all medieval philosophers. On this point Ptolemy was in agreement with the physics-based cosmology of Aristotle, but in general he represented a rival tradition, that of the mathematicians, who were usually regarded by the medievals as devisers of ingenious fictions that served merely to "save the observed appearances." Mathematical theories were accordingly believed to lack the necessity attributable to the vast and coherent background of Aristotelian physics and metaphysics, and this attitude prevailed until the time of Galileo Galilei.

However, there was some support for the development of mathematical physics, insofar as it relies on thought experiments as opposed to exact experiment, in the very competent medieval enlargements on a point whose root lay ultimately in Aristotle's *Categories*; there, when attempting to differentiate between substances (such as man, tree, stone) and qualities (such as whiteness, roundness, hardness), Aristotle pointed out that the latter are susceptible of degree, while the former are not. To this remote starting point much of modern mechanics owes its origin, for through speculation on the various kinds, rates, and degrees of "intension" and "remission" of qualities, the ideas of constant motion and acceleration and deceleration (uniform or nonuniform), and their relations to time and distance were thoroughly explored by fourteenth-century philosophers, such as those of Merton College, Oxford. Nicholas Oresme (c. 1325–1382) related these aspects of motion to their graphical expressions and anticipated infinitesimal calculus and coordinate geometry. Herein lies the starting point of certain segments of Galileo's mechanics.

See also Abelard, Peter; al-Fārābī; Anselm, St.; Aristotelianism; Aristotle; Artificial and Natural Languages; Augustine, St.; Augustinianism; Austin, John; Averroes; Averroism; Avicenna; Boethius, Anicius Manlius Severinus; Bonaventure, St.; Duns Scotus, John; Erigena, John Scotus; Galileo Galilei; Hobbes, Thomas; Ibn Gabirol, Solomon ben Judah; Islamic Philosophy; Jewish Philosophy; John of Salisbury; Logic, History of; Luther, Martin; Maimonides; Mathematics, Foundations of; Neoplatonism; Ontological Argument for the Existence of God; Oresme, Nicholas; Pantheism; Peter Lombard; Plato; Porphyry; Realism; Roscelin; Russell, Bertrand Arthur William; Scotism; Socrates; Sovereignty; Thomas Aquinas, St.; Thomism; Universals, A Historical Survey; William of Ockham; Wyclyf, John.

Bibliography

SOURCES, TEXTS, COMMENTARIES, TRANSLATIONS

Al-Farabi. *Al-Farabi's Commentary and Short Treatise on Aristotle's De interpretatione*. Translated by F. Zimmermann. London, 1981.

Al-Farabi. *Al-Farabi's Philosophy of Plato and Aristotle*. Translated by M. Mahdi. Ithaca, 1969; rev. ed. 2001.

Albert the Great. *Opera Omnia*. Edited by P. Jammy. 21 vols. Lyon, 1651.

Albert the Great. *Summa Theologiae*. Edited by D. Siedler. Munster, 1978.

Anselm of Canterbury. *Opera Omnia*. Edited by F. Schmitt. 6 vols. Stuttgart-Bad Cannstatt, 1968.

Aquinas, Thomas. *S. Thomae Aquinatis Doctoris Angelici Opera Omnia*. Rome, 1882.

Aquinas, Thomas. *Summa contra Gentiles*. Translated by A. Pegis et al. 5 vols. Notre Dame, IN, 1975.

Aquinas, Thomas. *Summa Theologiae*. Edited by T. Gilby et al. 61 vols. London, 1964–1980.

Augustine.*Against the Academicians and The Teacher*. Translated by P. King. Indianapolis, 1995.

Augustine. *Confessions*. Translated by H. Chadwick. Oxford, 1991.

Augustine. *Confessions*. Text, translation, and commentary by J. O'Donnell. Oxford, 1992.

Augustine. *De genesi ad litteram*. Translated by J. Taylor. New York, 1982.

Augustine. *Of True Religion*. Translated by J. Burleigh. Chicago, 1964.

Augustine. *On Christian Doctrine*. Translated by D. Robertson. Indianapolis, 1958.

Augustine. *On the Free Choice of the Will*. Translated by T. Williams. Indianapolis, 1993.

Augustine. *On the Trinity*. books 5–8, Translated by G. Matthews. Cambridge, 2002.

Augustine. *The City of God against the Pagans*. Translated by H. Bettenson. London, 1972.

Averroes. *On the Harmony of Religion and Philosophy*. Translated by G. Hourani. London, 1976.

Averroes. *Middle Commentary on Aristotle's De anima*. Edited by and translated by A. Ivri. Provo, UT, 2002.

Boethius. *De Consolatione philosophiae*. Edited by C. Moreschini. Munich, 2000.

Boethius. *The Consolation of Philosophy*. Translated by J. Relihan. Indianapolis, 2001.

Bonaventure. *Opera Omnia*. 10 vols. Ad Claras Aquas, 1882–1902.

Bonaventure. *Saint Bonaventure's Disputed Questions on the Mystery of the Trinity*. Translated by Z. Hayes. St. Bonaventure, NY, 1979.

Maimonides, Moses. *Ethical Writings of Maimonides*. Translated by R. Weiss and C. Butterworth. New York, 1975.

Maimonides, Moses. *The Guide to the Perplexed*. Translated by S. Pines. 2 vols. Chicago, 1963.

Scotus, John Duns. *Opera omnia*. Edited by C. Balic et al. Vatican City, 1950–.

Scotus, John Duns. *God and Creatures: The Quodlibetal Questions*. Translated by F. Alluntis and A. Wolter. Princeton, NJ, 1975.

Scotus, John Duns. *Questions on the Metaphysics of Aristotle*. Translated by G. Etzkorn and A. Wolter. St. Bonaventure, NY, 1997.

Scotus, John Duns. *Philosophical Writings*. Translated by A. Wolter. Edinburgh, 1962.

William of Ockham. *Opera philosophica et theological*. 17 vols. Edited by G. Gál et al. St. Bonaventure, NY, 1967–1988.

William of Ockham. *Philosophical Writings*. Translated by P. Boehner, rev. F. Brown. Indianapolis, IN, 1990.

HISTORIES

Armstrong, A. H. ed. *The Cambridge History of Later Greek and Early Medieval Philosophy*. Cambridge, 1967.

Gilson, E. *History of Christian Philosophy in the Middle Ages*. London, 1955.

Nasr, S. H. and O. Leaman, eds. *History of Islamic Philosophy*. 2 vols. London, 1996.

Schmidtt, C. B. and Q. Skinner, eds. *The Cambridge History of Renaissance Philosophy*. Cambridge, 1988.

Leaman, O. and D. Frank, eds. *History of Jewish Philosophy*. London, 1997.

STUDIES

Augustine

Brown, P. *Augustine of Hippo*. Berkeley, CA, 1967; rev. ed., 2000.

Gilson, E. *The Christian Philosophy of St. Augustine*. New York, 1960.

Kirwan, C. *Augustine*. London, 1989.

Rist, J. *Augustine: Ancient Thought Baptized*. Cambridge, 1994.

Boethius

Marenbon, J. *Boethius*. New York, 2003.

Al-Farābī

Druart, T.-A. "Le sommaire du livre des 'Lois' de Platon (Gawami' Kitab al-Nawamis li-Aflatun) par Abu Nasr al-Farabi." *Bulletin d'Etudes Orientales* 50 (1998) 109–55.

Galston, M. *Politics and Excellence: The Political Philosophy of Alfarabi*. Princeton, 1990.

Mahdi, M. *Alfarabi and the Foundations of Islamic Political Philosophy*. Chicago, 2001.

Ibn Sinā

Gutas, D. *Avicenna and the Aristotelian Tradition. Introduction to Reading Avicenna's Philosophical Works*. Leiden, 1988.

Wisnovsky, R. *Aspects of Avicenna*. Princeton, 2002.

Anselm of Canterbury

Southern, R. *Saint Anselm: A Portrait in a Landscape*. London, 1990.

Peter Abelard

Marenbon, J. *The Philosophy of Peter Abelard*. Cambridge, 1997.

Peter Lombard

Colish, M. *Peter Lombard*. 2 vols. Leiden, 1994.

John of Salisbury

Wilks, M. *The World of John of Salisbury*. Oxford, 1984.

Ibn Rushd

Endress, G. and J. Aertsen, eds. *Averroes and the Aristotelian Tradition. Sources, Constitution, and Reception of the Philosophy of Ibn Rushd (1126–1198)*. Leiden, 1999.

Renan, E. *Averroès et l'averroïsme*. Paris, 1852; 3rd rev. ed. 1866; repr. 1997.

Moses Maimonides

Buijs J. *Maimonides: A Collection of Critical Essays*. Notre Dame, IN, 1988.

Hyman, A. ed. *Maimonidean Studies*. 4 vols. New York, 1991–1996.

Kraemer, J. ed. *Perspectives on Maimonides: Philosophical and Historical Studies*. London, 1996.

Albert the Great

Weisheipl, J. *Albertus Magnus and the Sciences*. Toronto, 1980.

Bonaventure

Gilson, E. *The Philosophy of St. Bonaventure*. New York, 1938.

Thomas Aquinas

Benez, D. *The Primacy of Existence in Thomas Aquinas*. Translated by B. Llamzon. Chicago, 1966.

Finnis, J. *Aquinas*. Oxford, 2000.

MacDonald, S. and E. Stump, eds. *Aquinas' Moral Theory*. Ithaca, NY, 1999.

McInerny, R. *Ethica Thomistica*. Rev. ed. Washington, DC, 1997.

O'Connor, D. *Aquinas and Natural Law*. London, 1967.

Wippel, J. *The Metaphysical Thought of Thomas Aquinas: From Finite Being to Uncreated Being*. Washington, DC, 2000.

John Duns Scotus

Adams, M. ed. *The Philosophical Theology of John Duns Scotus*. Ithaca, NY, 1990.

Cross, R. *The Physics of Duns Scotus*. Oxford, 1998.

William of Ockham

Adams, M. *William Ockham*. 2 vols. Notre Dame, IN, 1987.

Other studies

Alverny, M.-T. "Le Cosmos symbolique du XIIe siècle." *Archives d'Histoire Doctrinale et Littéraire du Moyen Age*. 20 (1953) 31–81.

Black, A. *Political Thought in Europe 1250–1450*. Cambridge, 1992.

Craig, W. *The Problem of Divine Foreknowledge and Future Contingents from Ockham to Suarez*. Leiden, 1988.

D'Ancona Costa, C. *La Casa della sapienza. La transmissione della metafisica greca e la formazione della filosofia araba*. Milan, 1996.

Davidson, H. *Alfarabi, Avicenna, and Averroes, on Intellect: Their Cosmologies, Theories of the Active Intellect and Theories of Human Intellect*. New York, 1992.

Gracia, J. *Introduction to the Problem of Individuation in the Early Middle Ages*. Munich, 1984; 2nd ed. 1988.

Grant, E. ed. *A Source Book in Medieval Science*. Cambridge, MA, 1974.

Grant, E. *God and Reason in the Middle Ages*. Cambridge, 2001.

Grant, E. *The Foundations of Modern Science in the Middle Ages: Their Religious, Institutional, and Intellectual Contexts*. Cambridge, 1966.

Gutas, D. *Greek Thought, Arabic Culture: The Greco-Arabic Translation Movement in Baghdad and Early 'Abbasid Society (2nd–4th/8th–10th Centuries)*. London, 1998.

Jacobi, K. ed. *Argumentationstheorie: Scholastische Forschungen zu den logischen und semantischen Regeln korrekten Folgerns*. Leiden, 1993.

Knuuttila, S. *Modalities in Medieval Philosophy*. London, 1993.

Lovejoy, A. *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, MA, 1936.

Maier, A. *Studien zur Naturphilosophie der Spätmittelalter*. 5 vols. Rome, 1952–1968.

McGrath, A. *Iustitia Dei*. 2 vols. Cambridge, 1986.

Miethke, J. *De potestate papae. Die päpstliche Amtskomptetenz im Widerstreit der politischen Theorie von Thomas von Aquin bis Wilhem von Ockham*. Tübingen, 2000.

Pasnau, R. *Theories of Cognition in the Later Middle Ages*. Cambridge, 1997.

Skinner, Q. *The Foundations of Modern Political Thought*. 2 vols. Cambridge, 1978.

Desmond Paul Henry (1967)
Bibliography updated by Scott Carson

MEDITATION IN INDIAN PHILOSOPHY

Meditation as a distinct practice in Indian philosophy appears in a variety of texts from the third century before the common era as well as in sculptural depictions that date from 3500 BCE. The quintessential manual on meditation, the *Yoga Sūtra*, was composed by approximately 200 CE and includes philosophical positions and meditation techniques from the Sāṃkhya, Jaina, and Buddhist traditions.

Early depictions of meditating figures were found in the excavations of Mohenjodaro and Harappa, Indus Valley cities that date from 3500 BCE Sculptures and steatite seals show people with half-closed eyes sitting in the lotus posture. In some seals, animals surround a meditating figure, indicating a shamanic, totemic origin of this tradition.

The earliest text of Indian literature, the *Ṛg Veda*, which dates from at least 1500 BCE, mentions longhaired ascetics and, amidst hundreds of hymns extolling various gods and goddesses, lays out the philosophical foundations for later traditions of meditation. *Ṛg Veda* (1:164.20) describes two birds in the same tree, one eating sweet berries while the other merely witnesses. This theme repeats itself in the *Muṇḍaka Upanishad* (3:l:l) and the *Śvetāśvatara Upanishad* (4:6) and is expressed in the *Bhagavad Gītā* themes of the lower nature subject to constant change and activity (*pṛakrti*) and the higher nature or inner true self (*puruṣa* or *ātman*). The worldview presented in this early metaphor delineates two major modalities of engagement with the world. One aspect freely and unreflectively participates in and contributes to the world. The other aspect remains aloof and transcendent, as a spectator or onlooker.

Sāṃkhya philosophy, articulated by the philosopher Ishvarakrishna in the early centuries of the common era, delineates a cosmology based on this dynamic tension between the processes of activity and witnessing. The realm of activity includes psychological states (*bhāva*), operations of the mind (*manas*), sense and motor capacities (*indriya*), as well as the subtle and gross elements (*bhūta*) that manifest as discrete, concrete objects. By understanding and harnessing the karmically influenced outflows that arise when the witnessing consciousness becomes intrigued and defined by the particularity found in the manifest realm of activity, one gains mastery over and release from compulsive behavior, resulting in liberation (*kaivalyam*). This philosophy undergirds the system of Yoga, which presents a variety of meditation tech-

niques to accomplish the goal of liberation. Yoga also appears within non-Vedic traditions such as Jainism, Buddhism, Sufism, and Sikhism.

THE *YOGA SŪTRA*

The *Yoga Sūtra* of Patañjali (c. 200 CE) defines Yoga as the restraint of the fluctuations of the mind (*yogaś-citta-vṛtti-nirodha?*). The application of Yoga allows for the gradual diminishment of karmic influences, referred to as seeds (*bīja*) or residues (*saṃskāra*). Yoga specifies five aspects of defilement that must be controlled: ignorance, egoism, attraction, repulsion, and a desire for life to continue. By following the practices of Yoga, including meditation, karma dissipates. The practitioner reshapes his or her identity, abandoning attachment to fixed behaviors. By drawing inward, one reaches deeper self-understanding and approaches a state of lucidity and purification.

Numerous meditation practices can be found in the texts of Yoga, Buddhism, Jainism, and Sikhism. Different objects of meditation are listed, including fixing one's attention on Īśvara through the use of mantra. Patañjali defines Īśvara, sometimes referred to as a deity, as a special soul or *purusa* who has never been tainted by the actions of karma. By fashioning such an ideal through the imagination, one can then strive to emulate this rarefied being. For a Jaina, this state of Īśvara is symbolized through the twenty-four great teachers (Tīrthaṅkāra). For a Buddhist, Lord Buddha serves the same function. In the Hindu bhakti or devotional tradition, fixing one's attention on any one of a variety of deities can result in karmic purification, with Krishna and Rama being the most frequently worshipped Vaisnava deities and Siva and Ganesh and the Goddess Kāli the object of devotion for Saivites. For the Sikhs, the highest soul cannot be named and exists outside time (*akal*). However, the ten Sikh gurus, beginning with Guru Nanak, serve as objects with worship because of their teachings. Patañjali, through his concept of chosen deity (*iạ ṭa devatā*), suggests that the meditative procedures engaged in order to purify oneself carry more significance than the actual object of one's meditation.

Several other practices are listed in the *Yoga Sūtra* that do not require the presence of an inspirational, theistic object of devotion. They include becoming one-pointed in one's activities, regulating one's breath, experiencing inner radiance, reflecting on an auspicious dream, or "meditation as desired" (1:39). Patañjali puts forward a progressive technique, where one begins with a gross, outward object (*vitarka*) and then takes it inward, seeing its relationship with and grounding in one's mental constructs. One then moves on to more subtle aspects of one's psychological conditioning (*vicāra*), focusing on the patterns of past karma that tend to govern one's personality. By applying meditation techniques of focusing and calming the mind, and by probing into the root causes of one's motivations, one gradually gains the ability to move into a seedless state of pure being, referred to as *nirbīja samādhi*.

ETHICS

Ethics plays a crucial role in the meditation systems of India. Buddhists refer to these practices as perfections. Yogis and Jainas share a list of common vows. By holding to nonviolence (*ahiṃsā*) one engenders an atmosphere of well-being that brings calm and solace to others. By holding to truth, one's word corresponds to reality. Through not stealing, one gains appreciation of all that exists without seeking to appropriate or horde it for oneself. By abandoning sexual obsession, one makes the world safe from one's designs and manipulations. By giving up the acquisition of things, one can learn to understand one's motivations and past predilections. These five vows, common to nearly all India's meditative paths, allow for the deconstruction of destructive habits and the active construction of a safe, ethically-grounded world. For the Buddhists and the Yogis, a purified person naturally exhibits enlightened behavior and is friendly (not jealous) toward successful people, compassionate (not scornful) toward those who suffer, happy (not envious) for those who are meritorious, and retain their equanimity (do not become hateful) in regard to those who lack virtue.

PRACTICE

Meditation enables the practitioner to avoid the repetition of behavior that can be harmful to oneself and others. Indian philosophy, particularly as found in Buddhism, Sāṃkhya, and Yoga, claims that due to desire or thirst (*kāma/tṛṣṇa*) one engages in actions (karma) prompted by the residues of past actions (*saṃskāra*) that lead to repeated difficulty, darkness, and even despair (*duḥkha*). By the application of meditation and meditative ethical practices, one can cultivate an alternate way of being (*prati-pakṣa-bhāvanā*)rooted in purity. By withdrawing the outward flows of the mind and the senses and reversing the tendency to be defined by external objects and realities, one can become free of psychological entanglements and social expectations, achieving the status of a solitary hero, in charge of one's own reality. The word Jina, an epithet for Vardhamana Mahavira, the

twenty-fourth and most recent Tīrthaṅkāra of the Jaina tradition, indicates that he was a great vanquisher, one who conquered his past karma to establish himself as a model for others to emulate. Similarly, the enlightenment of the Buddha is cloaked with martial symbolism, with Siddhartha defeating the evil Mara in a great test of wills.

Meditation results in the accumulation of powers, ranging from enhanced language-learning abilities and physical beauty to memory of one's past lives. Through focusing on the interior energy of the body, one gains intimacy with the various subtle energy centers (*cakras*) that correlate with locations along the spine. These include vortexes of the earth-connected eliminative function, sexuality, and power found in the respective areas of the anus, the sexual organs, and the solar plexus. Above these three lower functions, one finds the seat of compassion in the heart, an array of emotions in the area of the throat, the third eye representing insight between the eyebrows, and in the area above the skull, a magnificent lotus. Through meditation techniques associated with Tantra and popularized from the eighth century forward, one systematically advances from the lower *cakras* toward the higher ones, bring about the ascent of a force known as the *kundalini*. However, whether the philosophy originates from Yoga, Buddhism, or Jainism, all traditions state that the powers (*siddhi*) must not distract one from the ultimate goal of self-purification.

Indian systems of meditation mandate the presence of a qualified teacher guru in order to engage in this variety of techniques. A well-qualified guru, in addition to knowing the mechanics, guides the student through the pitfalls of self-aggrandizement and periodic disappointment. Discovering one's past history can be fraught with frightful memories; the guru assists the disciple in this process of self-discovery. The Jaina tradition of past-life stories and the Buddha's narration of his past births in the Jataka tales, demonstrate that human action derives from ignorant, self-serving motivations, unless one has made a commitment to strive for purification. As shown in the paradigmatic case of the life of the Buddha, a realization of the fleeting nature of reality will often prompt a potential meditator to seek out instruction on how to achieve and maintain peace of mind. In the case of the Buddha, he studied various techniques for six years under two different renowned teachers before he entered into *nirvāṇa* and subsequently decided to teach others how to overcome their own personal difficulties through meditation. Guru Nanak (1469–1539), living in a time of great strife between Hindus and Muslims, underwent a miraculous transformation that prompted him to develop a new way of meditation that transcended both traditions. Modern day Yoga and meditation practices offer pathways of self-cultivation through the purification of the body, the emotions, and one's way of being situated in the world. These traditions all trace their origins back to an original teacher, whether Swami Vivekananda or Krishnamacarya for many schools of Yoga or to the Buddha himself for Buddhist meditators.

The philosophical texts on meditation in each of the traditions outline different paths and offer different catalogues of the karma that must be overcome. The *Yoga Sūtra* and its commentaries outlines five states of mind, five afflictive karma categories, seven levels of *samadhi*, a threefold path and an eightfold path of practice, and a tenfold ethical system. The core texts of Buddhism set forth an eightfold path and a fivefold assessment of the nature of reality that further subdivides into either seventy-five or one hundred constituent features. The Theravada texts outline nine meditations on objects with form and four formless meditation states. The *Tattvārtha Sūtra*, the foundational meditation text of Jainsim, describes 148 forms of karma known as *p̣rakrtis* and a fourteen-step analysis of states of increasing purification.

Meditation constitutes an important aspect of Indian philosophy. It requires an active engagement of the world through ethics. It requires the cultivation of a body that can sit for long periods of time. It also requires protracted states of introspection in order to gain mastery over the mind. Meditation comprises a comprehensive system of purification that, regardless of the particular theological context or philosophical point of view, serves to diminish negative karma and bring about states of equanimity.

See also Brahman; God in Indian Philosophy; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Mind and Mental States in Indian Philosophy; Negation in Indian Philosophy; Self in Indian Philosophy; Truth and Falsity in Indian Philosophy.

Bibliography

Chapple, Christopher Key, and Eugene P. Kelly, Jr., trs. *The Yoga Sūtra of Patanjali: An Analysis of the Sanskrit with Accompanying English Translation*. New Delhi: Sri Satguru Publications, 1990.

King, Winston. *Theravāda Meditation: The Buddhist Transformation of Yoga*. University Park: Pennsylvania State University Press, 1980.

Ñāṇamoli, Bhikku, tr. *The Path of Purification (Vissudhimagga) of Buddhaghosa*. Comombo: Semage, 1964.

Roebuck, Valerie J., tr. *The Upaniṣads*. London: Penguin Books, 2003.

Sargeant, Winthrop, tr. *The Bhagavad Gītā*. Albany: State University of New York Press, 1990.

Tatia, Nathmal, tr. *That Which Is: Tattvārtha Sūtra*. San Francisco: HarperCollins, 1994.

Christopher Key Chapple (2005)

MEGARIANS

The Megarians flourished during the fourth and the early third centuries BCE. They derived their name from their connection with Megara on the Isthmus (a city one day's walk west of Athens). They constituted a 'philosophical school' only in a weak sense: no shared lifestyle, no rigid body of doctrine. Since no work of any Megarian has survived, knowledge of them must rely on fragments and reports of other authors.

The earliest Megarian was Euclides of Megara. Diogenes Laertius (2.106) reports that Euclides' followers "were called 'Megarians,' then 'Eristics,' and later 'Dialecticians.'" Modern scholars traditionally understood this report as indicating that a single school had three successive labels. However, in 1977 David Sedley argued that the three labels designated three distinct groups of philosophers that were influenced to some extent by Euclides but, far from constituting a single school, were in competition with one another. Sedley's reconstruction has won widespread, although not universal, scholarly approval. The present entry will cover all those thinkers who have traditionally been regarded as Megarians, including Eristics and Dialecticians, except the Dialecticians Diodorus and Philo, who have separate entries.

Euclides of Megara was probably born after 450 BCE and died before 365 BCE. A pupil of Socrates, he also studied Parmenides' writings. He is mentioned by Plato in the *Phaedo* (59b–59c), where he is portrayed as present at Socrates' death, and in the *Theaetetus* (142a–143c), where he is described conversing with Terpsion, another early Megarian. After Socrates' death, Plato and some of his companions fled Athens to stay for awhile with Euclides at Megara. Euclides authored six dialogues: *Lamprias*, *Aeschines*, *Phoenix*, *Crito*, *Alcibiades*, and a *Discourse on Love*. We know little of Euclides' philosophical views. He claimed that the good is one although it is called by many names (such as 'wisdom,' 'God,' and 'mind'), and that the contrary of the good is mere nonbeing: he thus seems to have borrowed Socratic views in ethics and combined them with Eleatic monism. He attacked proofs by opposing their conclusions, not their premises (he probably did this by reducing to absurdity the conclusions, wherein an influence of the methods of Zeno of Elea can be detected), and he rejected arguments from parallel cases.

Euclides had numerous pupils: Dionysius of Chalcedon, Dioclides of Megara, Thrasymachus of Corinth, Ichthyas, and Clinomachus of Thurii, who founded the Dialectical school. According to Diogenes Laertius (2.112), Clinomachus was "the first who wrote about assertibles, predicates, and the like." Later, in Stoic logic, assertibles and predicates are two of the main types of sayables, incorporeal items that are signified by utterances of linguistic expressions and are themselves neither thoughts nor linguistic expressions (specifically, assertibles and predicates are what is signified, respectively, by utterances of declarative sentences and predicative expressions). It is unclear how much of the Stoic views about assertibles and predicates was already held by Clinomachus, but it cannot be ruled out that the basics were already in place.

According to some sources, one of Euclides' pupils was named 'Bryson.' Modern scholars disagree on whether there was exactly one thinker answering to this name, and whether he is the same as the one who introduced a method for squaring the circle which was criticized by Aristotle.

Later Dialecticians were Polyxenus (to whom the authorship of a 'third man' argument against forms is ascribed) and Eubulides of Miletus. Since he taught Demosthenes and wrote a defamatory book against Aristotle, Eubulides was probably born in the second half of the fourth century BCE. According to Diogenes Laertius (2.108), he fathered seven arguments: the Liar, the Disguised, the Electra, the Veiled, the Heaper, the Horned, and the Baldhead. These arguments, in question-and-answer form, were extensively discussed by later Hellenistic philosophers.

It is not clear whether Eubulides' version of the Liar had already the devastating self-referential character of modern versions. For instance, we cannot rule out that Eubulides' version was presented roughly as follows: The questioner makes an obviously false statement, adds the remark 'I am speaking falsely,' and then asks whether he is speaking truly or falsely—both answers can be regarded as correct with regard to different statements made by the questioner. Note that all ancient versions of the Liar turn on the sentence 'I am speaking falsely' (modern versions instead turn on 'This sentence is false' or variants thereof). The Heaper heaps questions concerning heaps: 'Does one grain constitute a heap?' 'Do two grains constitute a heap?' 'Do ten thousand grains constitute a heap?'

One is likely to answer the first question negatively, and then, on the assumption that the addition of a single grain cannot transform what is not yet a heap into one, is induced to answer negatively each of the following.

The Baldhead was probably an alternative formulation of the same puzzle. On the basis of Lucian (*Vitarum Auctio*, 22–23), we can plausibly reconstruct the Veiled as follows: 'Do you know your father?—Yes.—If I set a veiled man before you and I ask you whether you know him, what do you answer?—That I do not know him.—But the veiled man is your father. So, you both know and do not know your father.' The Disguised and Electra were probably variants of the Veiled. On the basis of Diogenes Laertius (7.187), we can plausibly reconstruct the Horned as follows: 'If you have not lost something, do you still have it?—Yes.—Have you lost horns?—No.—Then you still have horns.'

Pupils of Eubulides were Euphantus of Olinthus, Apollonius, surnamed 'Cronus' (his pupil Diodorus inherited this surname from him), and Alexinus of Elis, whose fondness of controversy earned him the nickname 'Elenxinus' ('Refuter'). Some sources describe Alexinus as a Dialectician, others as an Eristic. Active around 300 BCE, he wrote a book *On Education* and works against other thinkers, Aristotle and Zeno of Citium among them. Alexinus attacked Zeno by taking arguments of his and constructing unpalatable 'parallels,' namely arguments that were isomorphic to Zeno's and had plausible premisses but absurd conclusions. For instance, Zeno had offered the following argument: 'What is rational is better than what is not rational; but nothing is better than the universe; therefore, the universe is rational' (Sextus Empiricus, *Adversus Mathematicos*, 9.104).

Alexinus constructed the following parallel: 'What is poetic is better than what is not poetic; but nothing is better than the universe; therefore the universe is poetic' (Sextus Empiricus, *Adversus Mathematicos*, 9.108). Zeno was thereby left with two options: either claim that his argument is valid whereas Alexinus's parallel is not, or claim that all the premisses of his argument are true whereas at least one of Alexinus's parallel is not. The first option was hard to follow because the two arguments are extremely similar (in fact, neither of them is valid in first-order logic as it stands, but becomes such if an uncontroversial premise is added: 'Something is rational' in the case of Zeno's argument, 'Something is poetic' in the case of Alexinus's parallel). Sextus Empiricus (*Adversus Mathematicos*, 9.109–110) reports that Zeno's followers chose the second option: they insisted that all the premisses of Zeno's argument are true but one of Alexinus's parallel is not.

Little is known of Panthoides, a Dialectician who flourished around 300–280 BCE. The last Megarian about whom we are relatively well informed is Stilpo of Megara, who probably lived between 360 and 280 BCE. According to Diogenes Laertius (2.113), "so far did he excel everyone else in inventiveness and sophistry that nearly the whole of Greece was looking at him and Megarizing." He had many pupils, Zeno of Citium and Menedemus of Eretria among them, and wrote many dialogues. According to Plutarch (*Adv. Colotem*, 23, 1120a), Stilpo claimed that what is predicated must be identical with what it is predicated of. For example, goodness cannot be predicated of a man because it is not identical with him, nor can running be predicated of a horse because it is not identical with it. Stilpo's attack on predication recalls a position criticized by Plato in the *Sophist* (251a–c), and therefore lends plausibility to identifying Plato's target with some Megarian earlier than Stilpo. Stilpo attacked forms. One of his arguments can perhaps be reconstructed on the basis of Diogenes Laertius (2.119) and Alexander of Aphrodisias (*Commentary on Aristotle's Metaphysics*, 84, 7–14). Suppose that individual perceptible men and the form Man were the only men. It is surely true that man speaks. But who is then the man who speaks? Nobody: for it is none of the particular perceptible men (for why should it be this one rather than this one?), and it is not the form Man (for forms do not speak). If we want to avoid denying that man speaks, we must give up the assumption that individual perceptible men and the form Man are the only men, and therefore introduce a 'third man.' This seems to undermine our motivation for assuming there is the form Man.

According to Diogenes Laertius (2. 115), when Demetrius Poliorcetes had taken Megara and wanted Stilpo to list the items he had lost, "he said that he had lost nothing of his own: for nobody had subtracted his learning, and he still had reason and knowledge." This anecdote suggests that for Stilpo the only human goods are moral and intellectual attainments, which are inalienable (a view close to that of the Cynics).

In the *Metaphysics* (9. 3, 1046b29–32), Aristotle attributes to unnamed Megarians the view that a thing has the capacity to do something when and only when it is actually doing it. For example, whenever the builder is building, he also has the capacity to build, but when he is not building, he lacks the capacity to build. We are unable to link this view to any specific Megarian, and the ideas

about modality we can ascribe to Diodorus Cronus and Philo do not chime with it.

See also Alexander of Aphrodisias; Cynics; Diodorus Cronus; Diogenes Laertius; Hellenistic Thought; Parmenides of Elea; Philo of Megara; Plato; Plutarch of Chaeronea; Socrates; Sextus Empiricus; Stoicism; Zeno of Citium; Zeno of Elea.

Bibliography

COLLECTIONS OF FRAGMENTS

Döring, K. *Die Megariker. Kommentierte Sammlung der Testimonien*. Amsterdam: Grüner 1972.

Giannantoni, Gabriele. *Socratis et Socraticorum reliquiae*. 4 vols. Naples: Bibliopolis, 1990.

SECONDARY LITERATURE

Döring, K. "Gab es eine Dialektische Schule?" (Was there a Dialectical School?). *Phronesis* 34 (1989): 293–310.

Makin, Stephen. "Megarian Possibilities." *Philosophical Studies* 83 (1996): 253–276.

Montoneri, Luciano. *I Megarici: Studio storico-critico e traduzione delle testimonianze antiche*. Catania, Italy: University of Catania, 1984.

Müller, Robert. *Introduction à la pensée des Mégariques*. Paris: Vrin, 1988.

Müller, Robert. *Les Mégariques: fragments et témoignages*. Paris: Vrin, 1985.

Schofield, Malcolm. "The Syllogisms of Zeno of Citium." *Phronesis* 28 (1983): 31–58.

Sedley, David. "Diodorus Cronus and Hellenistic Philosophy." *Proceedings of the Cambridge Philological Society* 203 (1977): 74–120.

Zeller, Eduard. *Die Philosophie der Griechen in ihrer geschichtlichen Entwicklung dargestellt*. 5th ed. Leipzig, Germany: Reisland, 1920—1923, 2.1, p. 244–275.

Paolo Crivelli (2005)

MEIER, GEORG FRIEDRICH
(1718–1777)

Georg Friedrich Meier was a German philosopher and aesthetician. A pupil of Alexander Gottlieb Baumgarten, Meier succeeded Baumgarten as extraordinary professor at the University of Halle in 1740 and became a full professor in 1748, holding that position until his death.

Meier, a prolific writer, developed and commented on Baumgarten's doctrines as an extension and revision of Wolffianism and went far beyond Baumgarten in the reform of Wolffianism. His treatises, used as textbooks in many universities, were perspicuous, sophisticated, and modern renderings of Wolffian doctrine; by their thorough discussion of basic concepts and attention to details they give one of the best insights into the Wolffian system and its problems. Christian Wolff's and Baumgarten's ideas were rendered more fluid by Meier's work, establishing connections between disparate problems and establishing new distinctions. Meier's style was closer to the style of the "popular philosophers" than to that of orthodox Wolffians, and he made little use of the Wolffian mathematical method in philosophy.

Meier's *Vernunftlehre* introduced into the traditional frame of Wolffian logic lengthy psychological and methodological discussions like those of the Pietist philosophers A. F. Hoffmann and C. F. Crusius. He also presented a detailed typology of concepts. In a marked departure from Wolff, he stressed the limits of the human understanding, devoting an entire work to the subject (*Betrachtungen über die Schranken der menschlichen Erkenntniss*).

Meier's *Metaphysik*, although in general rather close to Baumgarten, shows the same individual features. For instance, in empirical psychology Meier advocated a subjectivism like that of Crusius. He held that the nature of our understanding determines what we can or cannot think. This determination, like the principle of *cogitabilis* in Crusius, is the foundation of the principle of identity.

Meier devoted several pamphlets to the immortality of the soul, which he held could not be theoretically demonstrated. Any a priori proof of God's existence must be completed by an a posteriori one. And in general Meier would not extend the power of reason much beyond basic truths and human experience.

Meier's most typical work was his *Anfangsgründe aller schönen Künste und Wissenschaften* (Principles of All Beautiful Arts and Sciences). He was opposed to the classical thesis that art imitates nature. He stressed the importance of sensitivity (the "lower faculty") and the indispensability of a knowledge of the beautiful within one's whole outlook on the world. Besides Baumgarten, whose views it is difficult to extricate from Meier's because of their close collaboration, Meier was influenced by the Swiss critics Johann Jakob Bodmer and Johann Jakob Breitinger and by English aestheticians. Like Baumgarten, he gave the term *aesthetics* a broad interpretation and, like Baumgarten's, his work contains an extensive discussion of scientific methodology.

See also Aesthetics, History of; Baumgarten, Alexander Gottlieb; Crusius, Christian August; Identity; Scientific Method; Wolff, Christian.

Bibliography

PRINCIPAL WORKS BY MEIER

Anfangsgründe aller schönen Künste und Wissenschaften. 3 vols. Halle, 1748–1750.

Gedanken über die Religion. Halle, 1749.

Vernunftlehre. Halle, 1752.

Philosophische Sittenlehre. 5 vols. Halle, 1753–1761. A much extended version of Baumgarten's *Ethica Philosophica*.

Metaphysik. 4 vols. Halle, 1755–1759.

Betrachtungen über die Schranken der menschlichen Erkenntniss. Halle, 1775.

WORKS ON MEIER

Bergmann, E. *Die Begründung der deutschen Aesthetik druch A. G. Baumgarten und G. F. Meier*. Leipzig: Röder and Schunke, 1911.

Böhm, Hans. "Das Schönheitsproblem bei G. F. Meier." *Archiv für die gesamte Psychologie* 56 (1926).

Langen, S. G. *G. F. Meier*. Halle, 1778.

Makkreel, Rudolf. "The Confluence of Aesthetics and Hermeneutics in Baumgarten, Meier, and Kant." *Journal of Aesthetics and Art Criticism* 54(1) (1996): 65–75.

Giorgio Tonelli (1967)
Bibliography updated by Tamra Frei (2005)

MEINECKE, FRIEDRICH

(1862–1954)

Friedrich Meinecke, the German historian and political philosopher, was small in stature and somewhat frail but remained mentally very vigorous and intellectually prolific until his death at the age of ninety-two. His great charm and influence were due partly to his erudition, partly to his modesty, and partly to two conflicting tendencies in his thinking that he continually sought to reconcile.

One of these tendencies was his patriotism and loyalty to Germany's best traditions of the past. As a boy he had been thrilled by the sight of the victorious German troops marching home through the Brandenburg Gate after the Franco-Prussian War. Later he admired the skill with which Otto von Bismarck established the long-desired unification of his country and saw with pride Germany's industrial and commercial expansion into a great power. After studying under the Prussian nationalist historian J. G. Droysen, Meinecke became an archivist and published in rapid succession several valuable historical works, including accounts of the German uprising against Napoleon Bonaparte and a two-volume biography of Hermann von Boyen, one of the leading figures in the reorganization and liberalization of Prussia in the early nineteenth century. In 1893 he was appointed an editor of the leading German historical journal, *Historische Zeitschrift,* a post that he filled with distinction for forty years until ousted by the Nazis.

The second tendency in Meinecke's thinking asserted itself in 1901 when he became deeply occupied with the problems of European political philosophy. In that year he was promoted to a teaching position at the University of Strassburg, later moving to Freiburg. Here in these two cities in the beautiful Rhine valley Meinecke's eyes were opened to the charm of the countryside. His talks with the Roman Catholic population and scholars and his contact with French culture widened his outlook and quickened his philosophical interests. These were his happiest years. In 1914 he was appointed to a permanent professorship at Berlin.

Meinecke's dual preoccupation with liberal culture and with Prussia found expression in a perceptive account of German development. *Weltbürgertum und Nationalstaat* (1908) examines the views of many cosmopolitan liberals and political leaders and, at the same time, analyzes the characteristics and pretensions of the Prussian state, which had been exaggerated by G. W. F. Hegel. It was supplemented by some two dozen articles written by Meinecke in the following years and reprinted in *Preussen und Deutschland* (1918).

Can reason of state justify the employment of might against right? May a state properly do things that are ethically forbidden to the ordinary citizen? Does it enjoy a code of morals above and beyond that of the private individual? Meinecke's classic treatment of these old but perennial questions, *Die Idee der Staatsräson in der neueren Geschichte* (1924), examines meticulously the actions of various European rulers and statesmen and the writings of numerous political theorists from Niccolò Machiavelli to Heinrich von Treitschke. Meinecke comes to the conclusion that, since power is the essence of its existence, the state is justified in using such means as are necessary to maintain and even extend its power, but that this power is limited by the state's obligation to protect the rights of its citizens and to promote their cultural and material welfare. It is, however, practically impossible to draw a precise line between state egoism and ideal morality.

Meinecke always preferred to till a small area where he could closely observe concrete facts and deal with them in a rigorously critical scientific manner. For Leopold von Ranke and Jakob Burckhardt he had the highest regard. He rejected the grandiose theoretical constructions of Karl Lamprecht, Oswald Spengler, and Arnold Toynbee. If he could be said to have had any one primary underlying

thought, it would be that of individuality—the unique individual character of every event, person, social group, nation-state, or idea. In addition he believed in evolution—the capacity of every individuality for development either by growth or decay. Hence his preoccupation with Machiavelli, Cardinal Richelieu, Freiherr vom Stein, Friedrich Schleiermacher, Wilhelm von Humboldt, Johann Wolfgang von Goethe, Joseph Maria von Radowitz, Bismarck, and Adolf Hitler. Meinecke's conceptions of individuality and evolution contributed to the new way of historical thinking, now known as "historicism," which developed in the age of Johann Gottfried Herder and Goethe and which Meinecke minutely unfolded in *Die Entstehung des Historismus* (1936). Historicism dealt a sharp blow to unquestioning belief in absolute values, optimistic positivism, religious creeds, and natural law. It opened wide the floodgates of relativism. Meinecke, however, was not unaware of the aberrations resulting from historicism and tried to counteract them by repeatedly insisting that the only sure and safe guide to morality and conduct is the individual's own conscience.

With the advent to power of the Nazis, Meinecke was forced to retire from active teaching, and under their tyranny he suffered spiritual agony and physical hardship. He might have escaped abroad as did so many others; but he remained in the country hoping to hasten Hitler's downfall and by his own advice and influence to help to lead Germany back to its older and better traditions. He was a close personal friend of General Beck and had some inkling of the plots to get rid of Hitler, but did not participate actively in them. His last contribution to an understanding of German history and his own interpretation of it was his little volume *Die deutsche Katastrophe* in 1946. Later, when the University of Berlin fell under communist control he took the lead in founding the new Free University in West Berlin, of which he was appropriately chosen rector.

See also Burckhardt, Jakob; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Historicism; Humboldt, Wilhelm von; Machiavelli, Niccolò; Political Philosophy, History of; Schleiermacher, Friedrich Daniel Ernst; Spengler, Oswald; Toynbee, Arnold Joseph.

Bibliography

WORKS BY MEINECKE

The principal works of Friedrich Meinecke are *Weltbürgertum und Nationalstaat* (Munich and Berlin, 1908; 7th ed., 1929); *Preussen und Deutschland* (Munich and Berlin, 1918); *Die Idee der Staatsräson in der neueren Geschichte* (Munich and Berlin: Oldenbourg, 1924), translated by D. Scott as *Machiavellism: The Doctrine of Raison d'État and Its Place in Modern History* (London: Routledge and Kegan Paul, 1957); and *Die Entstehung des Historismus,* 2 vols. (Munich and Berlin: Oldenbourg, 1936). *Die deutsche Katastrophe* (Wiesbaden: Brockhaus, 1946), translated by Sidney B. Fay as *The German Catastrophe* (Cambridge, MA: Harvard University Press, 1950, and Boston, 1962), written at the moment of Germany's utter defeat and deepest despair, contains Meinecke's penetrating reflections on the preceding hundred years, the causes of the Nazi disaster, and his faith in the future. Two short autobiographical volumes are *Erlebtes, 1862–1901* (Leipzig: Koehler and Amelang, 1941) and *Strassburg-Freiburg-Berlin, 1901–1914* (Stuttgart: Koehler, 1949). A six-volume edition of part of his works was published for the Friedrich Meinecke Institute of the University of Berlin between 1957 and 1962; this edition contains a volume of his correspondence and a reprint, with valuable editorial introductions and notes, of his more important writings. See also Meinecke's *Cosmopolitanism and the National State,* translated by Robert B. Kimber (Princeton, NJ: Princeton University Press, 1970); *Historism: The Rise of a New Historical Outlook,* translated by J. E. Anderson (London: Routledge and Kegan Paul, 1972); and *The Age of German Liberation, 1795–1954,* edited by Peter Paret, translated by Paret and Helmuth Fischer (Berkeley: University of California Press, 1977).

WORKS ABOUT MEINECKE

One of the best books on Meinecke and historicism is Walther Hofer, *Geschichtschreibung und Weltanschauung: Betrachtungen zum Werk Friedrich Meineckes* (Munich: Oldenbourg, 1950). A bibliography of writings by and about Meinecke may be found in the *Historische Zeitschrift,* Vol. 174, 503–523.

ADDITIONAL SOURCES

Pois, Robert A. *Friedrich Meinecke and German Politics in the Twentieth Century.* Berkeley: University of California Press, 1972.

Sterling, Richard W. *Ethics in a World of Power: The Political Ideas of Friedrich Meinecke.* Princeton, NJ: Princeton University Press, 1958.

Wolfson, Philip J. "Friedrich Meinecke (1862–1954)." *Journal of the History of Ideas* 17 (1956): 511–525.

Sidney B. Fay (1967)
Bibliography updated by Philip Reed (2005)

MEINONG, ALEXIUS
(1853–1920)

Alexius Meinong studied under Franz Brentano at the University of Vienna from 1875 through 1878 and taught at the University of Graz from 1882 until his death. In 1894 he established at Graz the first laboratory for experimental psychology in Austria. Some of his psychological

writings fall within this area, but most pertain to what Brentano called descriptive psychology. The philosophical works, referred to below, also pertain to descriptive psychology.

Meinong's most important contributions to philosophy concern the theory of objects, the theory of assumptions, the theory of evidence, and the theory of value. He also discussed, at considerable length, the nature of the emotions and their relation to intellectual phenomena, imagination, abstraction, wholes and other "complex objects," relations, causality, possibility, and probability.

THEORY OF OBJECTS

The two basic theses of Meinong's theory of objects (*Gegenstandstheorie*) are (1) there are objects that do not exist and (2) every object that does not exist is yet constituted in some way or other and thus may be made the subject of true predication. Traditional metaphysics treats of objects that exist as well as of those that merely subsist (*bestehen*) but, having "a prejudice in favor of the real," tends to neglect those objects that have no kind of being at all; hence, according to Meinong, there is need for a more general theory of objects.

Everything is an object, whether or not it is thinkable (if an object happens to be unthinkable then it is something having at least the property of being unthinkable) and whether or not it exists or has any other kind of being. Every object has the characteristics it has whether or not it has any kind of being; in short, the *Sosein* (character) of every object is independent of its *Sein* (being). A round square, for example, has a *Sosein*, since it is both round and square; but it is an impossible object, since it has a contradictory *Sosein* that precludes its *Sein*.

Of possible objects—objects not having a contradictory *Sosein*—some exist and others (for example, golden mountains) do not exist. If existence is thought of as implying a spatiotemporal locus, then there are certain subsistent objects that do not exist; among these are the *being* of various objects and the *nonbeing* of various other objects. Since there are horses, there is also the being of horses, the being of the being of horses, the nonbeing of the nonbeing of horses, and the being of the nonbeing of the nonbeing of horses. And since there is no Pegasus, there is the nonbeing of Pegasus, as well as the being of the nonbeing of Pegasus and the nonbeing of the being of Pegasus.

Meinong's theory must be distinguished from both Platonic realism, as this term is ordinarily interpreted, and the reism, or concretism, of Brentano and Tadeusz Kotarbiński. (Meinong noted that since his view is broader than realism, it might properly be called objectivism.) Thus, the Platonic realist could be said to argue: "(*P*) Certain objects that do not exist have certain properties; but (*Q*) an object has properties if and only if it is real; hence (*R*) there are real objects that do not exist." The reist, or concretist, on the other hand, reasons from not-*R* and *Q* to not-*P*; that is, he derives the contradictory of Plato's first premise by taking Plato's second premise along with the contradictory of Plato's conclusion. But Meinong, like Plato and unlike the reist, accepted both *P* and *R*; unlike both Plato and the reist, he rejected *Q* by asserting the independence of *Sosein* from *Sein*; and therefore, again unlike both Plato and the reist, he said that the totality of objects extends far beyond the confines of what is merely real (*das Universum in der Gesamtheit des Wirklichen noch lange nicht erschöpft ist*).

This doctrine of *Aussersein*—of the independence of *Sosein* from *Sein*—is sometimes misinterpreted by saying that it involves recourse to a third type of being in addition to existence and subsistence. Meinong's point, however, is that such objects as the round square have no type of being at all; they are "homeless objects," to be found not even in Plato's heaven. Bertrand Russell objected that if we say round squares are objects, we violate the law of contradiction. Meinong replied that the law of contradiction holds only for what is real and can hardly be expected to hold for any object, such as a round square, that has a contradictory *Sosein*.

Russell's theory of descriptions is often thought to constitute a refutation of the doctrine of *Aussersein*; actually, however, his theory merely presupposes that Meinong's doctrine is false. According to Meinong, the two statements "The round square is round" and "The mountain I am thinking of is golden" are true statements about nonexistent objects; they are *Sosein* and not *Sein* statements. The distinction between the two types of statements is most clearly put by saying that a *Sein* statement (for example, "John is angry") is an affirmative statement that can be existentially generalized upon (we may infer "There exists an x such that x is angry") and a *Sosein* statement is an affirmative statement that cannot be existentially generalized upon; despite the truth of "The mountain I am thinking of is golden," we may not infer "There exists an x such that I am thinking about x and x is golden." Russell's theory of descriptions, however, presupposes that every statement is either a *Sein* statement or the negation of a *Sein* statement and hence that there are no *Sosein* statements. According to Russell, a statement of the form "The thing that is F is G" may be

paraphrased as "There exists an *x* such that *x* is *F* and *x* is *G*, and it is false that there exists a *y* such that *y* is *F* and *y* is not identical with *x*." If Meinong's true *Sosein* statements, above, are rewritten in this form, the result will be two *false* statements; hence Meinong could say that Russell's theory does not provide an adequate paraphrase.

An impossible object, as indicated above, is an object having a *Sosein* that violates the law of contradiction. An *incomplete object*, analogously, is one having a *Sosein* that violates the law of the excluded middle. Of the golden mountains, which most readers will think of on reading the paragraph above, it will be neither true nor false to say that they are higher than Mount Monadnock. And some objects are even more poorly endowed. For example, if I wish that your wish will come true, then the object of my wish is whatever it is that you happen to wish; but if, unknown to me, what you wish is that my wish will come true, then this object would seem to have very little *Sosein* beyond that of being our mutual object. Meinong said that such an object is a *defective object* and suggested that the concept may throw light upon some of the logical paradoxes.

The theory of complexes—that is, the theory of wholes and other such "objects of higher order"—upon which Meinong wrote at length, also falls within the theory of objects.

None of the objects discussed above is created by us, nor does any of them depend in any way upon our thinking. Had no one ever thought of the round square, it would still be true *of* the round square that it does not exist; the round square need not be thought of in order not to exist. We draw these objects, so to speak, from the infinite depths of the *Ausserseienden*, beyond being and not-being.

THEORY OF ASSUMPTIONS

Meinong's theory of assumptions, or suppositions, is set forth in *Über Annahmen* ("On Assumptions"; first ed., Leipzig, 1902; 2nd ed., Leipzig, 1910). The theory is best understood by contrasting it with two theses held by Brentano, to which Meinong's theory may be said to be a reaction. The first of Brentano's theses is that of reism, or concretism, referred to above: Every object is a concrete thing; there are no objects such as the being of horses or the nonbeing of unicorns; the object of a judgment, therefore, is not a proposition, fact, or state of affairs; it is, rather, a certain concrete thing that the judgment may be said either to accept or to reject. And according to the second of Brentano's theses, there are basically only two types of intellectual attitudes we can take with respect to any object: We can simply think about the object, in which case it is the object of a thought or idea, or we can take an intellectual stand with respect to the object, either accepting it or rejecting it, in which case it becomes the object of a judgment. Meinong rejected both these theses of Brentano.

The object of a judgment, according to Meinong, is not a concrete thing; it is an "objective" (*Objektiv*). "That there are horses," for example, designates an objective—an object of higher order, containing horses as a kind of constituent. (Thus, the nonexisting, nonsubsisting round square is a constituent of that subsisting objective that is the nonbeing of the round square.) Assumptions, like judgments, take objectives as their objects.

What Meinong intended by his term *assumption* (*Annahme*) is most clearly exemplified in deliberation: "Suppose I were to do *A*. What would happen then? And now suppose I were not to do *A*. What would happen then?" Assumptions belong to a category falling between ideas and judgments. Like mere ideas, they do not themselves involve commitment, belief, or conviction; therefore, as such, they do not involve any possibility of error. Like judgments, they are concerned with objectives (in the above example, with what is designated by "I shall do *A*"), which are either true or false (it is either true or false that I shall do *A*); and, like judgments, assumptions involve either affirmation ("Suppose I do *A*") or denial ("Suppose I do not do *A*"), but affirmation or denial without commitment.

Meinong argues that only by reference to assumptions can we understand such phenomena as the nature of inference, our apprehension of negative facts, communication in general, desire, art, and the nature of play and of games. *Über Annahmen*, which is probably Meinong's best book, contains important material on these and many other topics.

THEORY OF EVIDENCE

The concept of evidence involves three dichotomies: (1) direct and indirect; (2) a priori and a posteriori; and (3) "evidence for certainty" and "evidence for presumption." Meinong's conception of the first two dichotomies is similar to that of Brentano. Thus there are axioms of mathematics and logic and the theory of objects, which are directly evident and a priori; and there are facts of "inner perception"—for example, the fact that I am making such-and-such an assumption, or the fact that I take something to be a tree—which are directly evident and a posteriori. (Any psychological process that "presents" an object to us, as memory may be said to present certain

objects of the past, is also a process that "presents itself"; "self-presentation" is thus the source of that evidence which is direct, certain, and a posteriori.) These directly evident judgments may confer evidence upon certain other judgments, which are then said to be indirectly evident.

For Meinong, paradigm cases of what is a priori evident would be expressed by "Round squares are both round and square" and "red is different from blue." Every a priori judgment has four characteristics: It is grounded in the nature of its object (*gegenständlich begründet*); it is certain; it is necessary; and it does not take into consideration the question whether its object exists. (Brentano had said that every a priori judgment is a judgment to the effect that a certain type of object does not exist.)

An evident presumption (*Vermutung*) may be directly evident but not certain. The concept is needed, according to Meinong, in order for us to understand memory, perception, and induction. In each of these three cases we have a source of knowledge that cannot be impugned as such but may on occasion mislead us. A particular memory judgment, for example, may not be certain, but it may be evident, especially if it is supported by other memory judgments, by perceptual judgments, or by inductive inferences from such judgments; analogously, this holds for any particular perceptual judgment or any particular inductive conclusion. Such items of a posteriori knowledge may be compared with the cards in a pack, "no one of which is capable of standing up by itself, but several of which placed together can serve to hold each other up. Or, for something more solid, consider a stack of weapons in the field." A consequence of this theory of evident presumptions is that a false judgment may yet be evident, a consequence that Brentano took to be absurd. Evidence does not guarantee truth; but, according to Meinong, evidence resembles truth in that if a judgment is evident, then its being evident—its *Evidentsein*—as well as the *Evidentsein* of this *Evidentsein*, and so on ad infinitum, is also evident.

An essential part of Meinong's epistemology is his theory of "emotional presentation" There is an analogy between the way in which we come to know, say, that the temperature is high and the way in which we come to know that the temperature is agreeable. Meinong proposed, as a "heuristic principle," that we try to carry the analogy as far as possible. If it is by means of a subjective feeling that we perceive the temperature to be agreeable, it is also by means of a subjective sensation that we perceive the temperature to be high. In neither case is the subjective experience the object of the presentation; in neither case is our apprehension a matter of inference or of reasoning from effect to cause. "The sense in which the sky is said to be 'beautiful,' for example, is precisely that in which it is said to be 'blue.' But the experience by means of which the first property is presented plays an important role in our psychical life in addition to that of enabling us to grasp something else. This fact is reflected in our language; we refer to the one experience directly, but in the other case we must go round about, by way of the object that is presented, and use some such expression as 'experience of blue.'" Meinong noted that the traditional arguments against a "subjectivistic" or "psychologistic" interpretation of ordinary sense perception apply equally to any such interpretation of emotional presentation.

THEORY OF VALUE

In the final version of his theory of value, Meinong made use of the theory of emotional presentation considered above, as well as of Brentano's doctrine of correct and incorrect emotion—that is, the doctrine according to which emotions, like judgments, may be said to be correct or incorrect, justified or unjustified, and according to which certain things may thus be said to merit or be worthy of certain emotions.

The basic concept of value theory is not that of desire, interest, or utility, but that of value feeling (*Wertgefühle*). Value feelings take objectives as their objects, more particularly, objectives consisting of the being or nonbeing of certain objects. One type of value feeling is *Seinsfreude*, pleasure or joy in the existence or being of a certain object; another type is *Seinsleid*, displeasure or sorrow with respect to the existence or being of a certain object. But the feelings of joy and sorrow may also be directed toward nonexistence and nonbeing; hence there are four fundamental types of value feeling, which may be illustrated by reference to the nature of good and evil. The good is that which merits *Seinsfreude* if it exists and *Nichtseinsleid* (sorrow with respect to its nonexistence) if it does not exist; evil, on the other hand, merits *Seinsleid* if it exists and *Nichtseinsfreude* (joy with respect to its nonexistence) if it does not exist. Meinong noted that human beings are not consistent in their emotional reactions. For example, as far as our health and ordinary comforts are concerned, we experience considerable *Nichtseinsleid* when they are absent, but not the appropriate amount of *Seinsfreude* when they are present.

Our actions have moral qualities other than those of being good, bad, or indifferent. Meinong introduced four moral categories, which he explicated by reference to

good and bad. Actions that are good may be either meritorious or simply required; those that are bad may be either excusable or inexcusable. (Meinong's terms are, respectively, *verdienstlich, correct, zulässig*, and *verwerflich*.) One may say of any act that performance is meritorious if and only if nonperformance is bad but excusable; nonperformance is meritorious if and only if performance is bad but excusable; performance is required if and only if nonperformance is inexcusable; and nonperformance is required if and only if performance is inexcusable. Given this "law of omission" (*Unterlassungsgesetz*), Meinong's concepts of meritorious, required, excusable, and inexcusable, respectively, approximate what are sometimes called the supererogatory, the obligatory, misdeeds that are venial, and misdeeds that are not venial. According to one of Meinong's followers (Ernst Schwarz), these four moral concepts are related to the concept of justified or correct emotion in the following way: The meritorious is that which it is incorrect to blame and incorrect not to praise; the required is that which it is incorrect to blame, correct to praise, but not incorrect not to praise; the merely excusable is that which it is incorrect to praise, correct to blame, and not incorrect not to blame; and the inexcusable is that which it is incorrect to praise and incorrect not to blame.

See also Brentano, Franz; Epistemology, History of; Ethical Objectivism; Kotarbiński, Tadeusz; Logical Paradoxes; Nonexistent Object, Nonbeing; Plato; Platonism and the Platonic Tradition; Propositions; Psychology; Realism; Russell, Bertrand Arthur William; Value and Valuation.

Bibliography

WORKS BY MEINONG

Meinong summarized his principal philosophical conclusions in *Die deutsche Philosophie der Gegenwart in Selbstdarstellungen*, edited by Raymund Schmidt (Leipzig, 1921), Vol. I, pp. 91–150. His purely psychological writings can be found in the first volume of *Gesammelte Abhandlungen* (2 vols.; Leipzig: Barth, 1913–1914).

Theory of Objects

Meinong's theory of objects is discussed in "Über Gegenstandstheorie" (1904); this article was reprinted in Vol. II of his *Gesammelte Abhandlungen* and translated as "The Theory of Objects" in *Realism and the Background of Phenomenology*, edited by Roderick M. Chisholm (Glencoe, IL: Free Press, 1960). The theory is discussed also in *Über die Stellung der Gegenstandstheorie im System der Wissenschaften* (Leipzig, 1907) and, indeed, in almost all Meinong's writings after 1904.

Epistemology

His most important epistemological writings are *Zur erkenntnistheoretischen Würdigung des Gedächtnisses* (1886), reprinted in Vol. II of *Gesammelte Abhandlungen; Über die Erfahrungs-grundlagen unseres Wissens* (Berlin, 1906); *Über Möglichkeit und Wahrscheinlichkeit* (Leipzig: Barth, 1915); and *Über emotionale Präsentation* (Vienna, 1917).

Value Theory

Meinong's principal writings in value theory are *Psychologischethische Untersuchungen zur Werththeorie* (Graz: Leuschner and Lubensky, 1894) and the posthumously published *Zur Grundlegung der allgemeinen Werththeorie* (Graz: Leuschner and Lubensky, 1923).

WORKS ON MEINONG

Among the most useful writings on Meinong are Bertrand Russell, "Meinong's Theory of Complexes and Assumptions," three articles in *Mind* 13 (1904): 204–219, 336–354, and 509–524; J. N. Findlay, *Meinong's Theory of Objects and Values* (2nd ed., Oxford: Clarendon Press, 1963); G. Dawes Hicks, "The Philosophical Researches of Meinong," in *Critical Realism* (London: Macmillan, 1938); and Konstantin Radakovic et al., *Meinong-Gedenkschrift* (Graz: "Styria" Steirische Verlagsanstalt, 1952).

Important material on all aspects of Meinong's philosophy can be found in *Philosophenbriefe: Aus der wissenschaftlichen Korrespondenz von Alexius Meinong mit Fachgenossen seiner Zeit*, edited by Rudolf Kindinger (Graz: Akademischen Druck- u. Verlagsanstalt, 1965).

OTHER RECOMMENDED WORKS

Barber, Kenneth F. *Meinong's Hume Studies: Translation and Commentary*. PhD diss., University of Michigan, 1967.

Grossmann, Reinhardt. *Meinong*. London; Boston: Routledge & K. Paul, 1974.

Lambert, Karel. *Meinong and the Principle of Independence: Its Place in Meinong's Theory of Objects and Its Significance in Contemporary Philosophical Logic*. Cambridge, U.K., and New York: Cambridge University Press, 1983.

Schubert Kalsi, Marie-Luise. *Alexius Meinong's Elements of Ethics: With Translation of the Fragment Ethische Bausteine*. Dordrecht; Boston: Kluwer Academic Publishers, 1996

Sylvan, Richard. *Exploring Meinong's Jungle and Beyond: An Investigation of Noneism and the Theory of Items*. Canberra: Research School of Social Sciences, Australian National University, 1980.

Works by Meinong

Gesamtausgabe. Herausgeber: Rudolf Haller und Rudolf Kindinger. Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1968–1978.

Abhandlungen zur Werttheorie. Rudolf Kindinger. Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1968.

Abhandlungen zur psychologie. Rudolf Kindinger. Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1969.

On Emotional Presentation. Evanston, IL: Northwestern University Press, 1972.

Über philosophische Wissenschaft und ihre Propädeutik; Über die Stellung der Gegenstandstheorie im System der Wissenschaften; Über die Erfahrungsgrundlagen unseres Wissens; Zum Erweise des allgemeinen Kausalgesetzes. Graz: Akademische Druck- u. Verlagsanstalt, 1973.

On Assumptions. Berkeley: University of California Press, 1982.
Meinong Reconstructed versus Early Russell Reconstructed: A Study in the Formal Ontology of Fiction. PhD diss., Indiana University, 1986.
Uber Gegenstandstheorie; Selbstdarstellung. Hamburg: F. Meiner, 1988.
Alexius Meinong und Guido Adler: Eine Freundschaft in Briefen. Amsterdam; Atlanta, GA: Rodopi, 1995.

Roderick M. Chisholm (1967)
Bibliography updated by Michael J. Farmer (2005)

MELANCHTHON, PHILIPP
(1497–1560)

Philipp Melanchthon, the German reformer, was born at Bretten, Baden, and died at Wittenberg. He was a grandnephew of the great humanist Johannes Reuchlin, who encouraged him in his studies and deeply influenced his outlook. After studying at Heidelberg and Tübingen, Melanchthon, on Reuchlin's recommendation, became professor of Greek at Wittenberg. Because of his persuasiveness in interpreting the humanist spirit, this appointment marked the beginning of a new era in German education. At Wittenberg, Melanchthon collaborated closely with Martin Luther. He helped him both in translating the Bible and in giving systematic shape to the new theology that until that time had existed in a highly subjective form. Melanchthon's task was to reduce this theology to exact form and to set it forth as an integrated and persuasive system. In 1521 Melanchthon published his *Loci Communes Rerum Theologicarum,* a work that in its various editions was one of the most influential manuals of Protestant theology.

During the rest of his career, Melanchthon was much occupied with controversy and debate. In many of the famous conferences of the Reformation era, his influence was thrown on the side of moderation and peace. He was closely identified with some of the most important formularies of the period, such as the Augsburg Confession.

Such activities involved even a man of conciliatory spirit in vigorous debate, and Melanchthon's position in the history of thought is largely determined by the controversies in which he took part. Two of these demand consideration.

The Adiaphoristic controversy was concerned with "indifferent matters"—that is, religious practices or theological beliefs on which flexibility or compromise might be permissible. Melanchthon was unfairly charged with including among the "adiaphora" such major questions as justification by faith. Melanchthon did not minimize the importance of essentials, but he was inclined to veil them beneath a conscious indefiniteness of expression. This deliberate obscurity extended to many matters that were intensively canvassed in the sixteenth century. He was willing to concede that good works are necessary to salvation, but not in the way in which the connection had traditionally been taught. He was prepared to recognize seven sacraments, but only if most of them were regarded as rites that have no inherent efficacy in securing salvation. Later he retreated from the permissive position he had adopted on the "adiaphora" and maintained a strict interpretation of the doctrines set forth in the *Loci Communes.*

More acute and more important was the controversy about synergism. Here the central issue was the relation between God's grace and man's will in regeneration. In his early period, Melanchthon, strongly influenced by Luther and deeply impressed by the experience of dependence upon God, severely restricted the role of man's will. To defend free will was to rob God's grace of its unique supremacy. But Melanchthon naturally tended to adopt a mediating outlook, and ethical issues were of great importance to him. Desiderius Erasmus, in his controversy with Luther concerning free will, had advanced views that served to modify Melanchthon's position. Melanchthon was now prepared to recognize the part played in conversion by man's will. The position that he reached (called synergism) precipitated a violent debate. Melanchthon's own statements were ambiguous and lacking in precision. His supporters (Johan Pfeffinger and Viktorin Strigel, for instance) and his opponents (Nikolaus von Amsdorf and Matthias Flacius Illyricus) were very explicit indeed. Synergism, however, can best be understood as an ethical protest against attitudes that paralyze the conscience and leave the church powerless in its struggle against moral chaos. Melanchthon's concern with God's moral purity led him to the belief that the problems of evil and of human responsibility have been aggravated by an extreme doctrine of predestination. He therefore abandoned the decree of eternal reprobation. The cause of sin lies in man himself; the hardening of his heart is due to his own perversity. Man has a real measure of responsibility for his spiritual condition. Man's will, therefore, can cooperate with God's grace, and does so. The human will, of course, is never the primary cause of man's regeneration—the Spirit of God and the preaching of the Word always maintain the initiative—but man's will is specifically granted a place, and unless there is consent on man's part there can be no effective regeneration. Melanchthon guarded himself against the charge of Pelagianism, but nevertheless he was accused of yielding to

this heresy. The violence of the controversy was due to the seriousness of the issues involved. A wide range of theological views had to be reexamined, and every aspect of the Christian doctrine of man and of salvation was involved. The controversy was finally silenced by the Formula of Concord, which ruled against the Melanchthonist position.

See also Erasmus, Desiderius; Evil, The Problem of; Logic, History of; Luther, Martin; Pelagius and Pelagianism; Reformation.

Bibliography

WORKS BY MELANCHTHON

"Works." In *Corpus Reformatorum,* edited by K. G. Bretschneider and E. Bindseil, Vol. I–XXVIII. Brunswick, Germany, 1834–1860.

Supplementa Melanchthoniana. Leipzig: Haupt, 1910.

The Loci Communes of Philipp Melanchthon. Translated by C. L. Hill. Boston, 1944.

WORKS ON MELANCHTHON

Hammer, W. *Die Melanchthonforschung im Wandel der Jahrhunderte,* 3 vols. Gütersloh, Germany: Gerd Mohn, 1967–1981.

Hartfelder, K. *Philipp Melanchthon als Preceptor Germaniae.* Berlin: Hofmann, 1889.

Hildebrandt, Franz. *Melanchthon, Alien or Ally?* Cambridge, U.K.: Cambridge University Press, 1946.

Kusukawa, Sachiko. *The Transformation of Natural Philosophy: The Case of Philip Melanchthon.* Cambridge, U.K.: Cambridge University Press, 1995.

Richards, J. W. *Philipp Melanchthon: The Protestant Preceptor of Germany.* New York, 1898.

Wengert, Timothy. *Human Freedom, Christian Righteousness: Philip Melanchthon's Exegetical Dispute with Erasmus of Rotterdam.* Oxford: Oxford University Press, 1998.

Gerald R. Cragg (1967)
Bibliography updated by Christian B. Miller (2005)

MELISSUS OF SAMOS
(fifth century BCE)

Melissus of Samos, the Greek Eleatic philosopher, led the Samian fleet against the Athenians and defeated them (Plutarch, *Pericles* 26, quoting a lost work of Aristotle). The date of the battle was 441–440 BCE, and this is the only reliable date in the biography of Melissus. He was said to have been a pupil of Parmenides, but this may be an inference from his work, which gives ample evidence of dependence on Parmenides.

Portions of Melissus's book titled *On Nature or What Exists,* written in prose, were quoted and preserved by the Aristotelian commentator Simplicius. The total length of these fragments is a little under one thousand words—enough to provide evidence of the content and quality of Melissus's argument. No other fragments survive. The pseudo-Aristotelian treatise *On Melissus, Xenophanes and Gorgias* (c. first century CE) adds nothing useful.

Melissus's argument, as revealed by the fragments, was similar to Parmenides' in method and results, although it differed in some details. The starting point is the contradictoriness of descriptions of change. Any change ultimately implies the generation of something from nothing or its destruction into nothing, and Melissus, with Parmenides, held both of these to be impossible on the ground that "nothing" is absolutely nonexistent and unthinkable. Hence, what exists must have existed always and must continue to exist (Melissus seems to view eternity as a continual existence through time, whereas Parmenides thought of a timeless present).

From the eternity of what exists, Melissus deduced its spatial infinity. He argued that if what exists did not come into existence, it had no beginning or end, and being without beginning or end, it must be limitless or infinite. He seemed not to have noticed the ambiguity of "beginning" and "end" (or else his defense of the move from time to space has been lost); this is presumably the basis of Aristotle's criticism of the argument (*De Sophisticis Elenchis* 167b13 and 168b35), although he does not make it quite explicit.

From the spatial infinity of what exists, Melissus deduced its unity. If there are two things in existence, each must limit the extent of the other; there cannot be more than one limitless thing in existence. Thus, Melissus chose a different route to the monism of Parmenides—indeed, according to most interpreters of Parmenides, this route was closed to him since, unlike Melissus, he held that what exists is spatially limited. But this is a dubious interpretation of Parmenides.

Next, Melissus argued that if what exists is one, it cannot have parts and must therefore be incorporeal because any solid body has actual or imaginable parts. Moreover, what exists cannot vary in density since this, according to Melissus, could come about only if one area contained less of being—and hence more of nonbeing—than another, and nonbeing is absolutely nonexistent. For similar reasons there is no motion, since there is no "give" anywhere in the plenum (this is an argument against motion that may not have been used by Parmenides). Every form of change—whether of size, order, or qual-

ity—means the coming into existence of something that previously was nothing, or the annihilation of something that exists, and these are ruled out by the first stage of Melissus's argument.

In the eighth fragment Melissus applies his own criteria of existence to the plural beings of the sensible world. If these things, such as air and fire, exist, then they must be just what our senses tell us they are and nothing else. But our senses tell us that they do change into something else. Our senses must therefore be wrong about this; hence, we can conclude that they were wrong initially in telling us that things are many and not one. The sensible world is therefore illusion.

Melissus was the least important of the Eleatics. Zeno's arguments proved more influential than his, and Parmenides was the original genius who pioneered the way. If Melissus has any claim to special historical importance that is not shared by the other Eleatics, it is perhaps that by applying Eleatic criteria to the plural beings posited by his opponents, he produced a formula (in Fr. 8) that led Leucippus directly to the concept of atoms. In the absence of complete texts it is wiser to refrain from pronouncing on Melissus's originality. Aristotle criticized both Parmenides and Melissus for bad arguments (*Physics* 186a6) and was more severe on Melissus, but perhaps that was because Melissus's clear style made him an easier target.

See also Aristotle; Change; Eternity; Infinity in Mathematics and Logic; Leucippus and Democritus; Parmenides of Elea; Plutarch of Chaeronea; Space; Zeno of Elea.

Bibliography

Fragments of Melissus's writings in Greek with German translations have been published in Diels-Kranz, *Fragmente der Vorsokratiker,* 10th ed., Vol. I (Berlin, 1960); English translations, in J. Burnet, *Early Greek Philosophy,* 4th ed. (London: A. and C. Black, 1930).

Selected texts with English translation and commentary are in G. S. Kirk and J. E. Raven, *The Presocratic Philosophers* (Cambridge, U.K.: Cambridge University Press, 1957); the commentary should be treated with caution, especially on the subject of infinity and Melissus's relation with the Pythagoreans. The same is true of J. E. Raven, *Pythagoreans and Eleatics* (Cambridge, U.K.: Cambridge University Press, 1948).

See also Harold Cherniss, *Aristotle's Criticism of Presocratic Philosophy* (Baltimore: Johns Hopkins Press, 1935), and G. E. L. Owen, "Eleatic Questions," in *Classical Quarterly* 10 (1960).

ADDITIONAL SOURCES

Barnes, J. *The Presocratic Philosophers,* 2nd ed., 180–185, 194–230, 298–302. London: Routledge and Kegan Paul, 1982.

Bicknell, P. "Melissus' Way of Seeming." *Phronesis* 27 (1982): 194–201.

Booth, N. B. "Did Melissus Believe in Incorporeal Being?" *American Journal of Philology* 79 (1958): 61–65.

Burnet, J. *Greek Philosophy. Thales to Plato,* 4th ed., 320–329. London, 1930.

Cherniss, H. *Aristotle's Criticism of Presocratic Philosophy,* 2nd ed., 61–72, 402–403. New York: Octagon, 1964.

Curd, P. *The Legacy of Parmenides: Eleatic Monism and Later Presocratic Thought,* 202–204, 206–216, 224–227. Princeton, NJ: Princeton University Press, 1998.

Furley, D. *The Greek Cosmologists,* Vol. 1: *The Formation of the Atomic Theory and Its Earliest Critics,* 110–114. Cambridge, U.K.: Cambridge University Press, 1987.

Furley, D. "Melissus of Samos." In *Ionian Philosophy,* edited by K. Boudouris, 114–122. Athens, 1989.

Guthrie, W. K. C. *A History of Greek Philosophy,* Vol. 2: *The Presocratic Tradition from Parmenides to Democritus,* 101–121. Cambridge, U.K.: Cambridge University Press, 1965.

Jouanna, J. "Rapports entre Mélissos de Samos et Diogène d'Apollonie, à la lumière du traité hippocratique De natura hominis." *REA* 67 (1965): 306–323.

Kirk, G. S., J. E. Raven, and M. Schofield. *The Presocratic Philosophers,* 2nd ed., 390–401. Cambridge, U.K.: Cambridge University Press, 1983.

Loenen, J. H. M. M. *Parmenides, Melissus, Gorgias: A Reinterpretation of Eleatic Philosophy.* Assen: Royal VanGorcum, 1961.

Owen, G. E. L. "Plato and Parmenides on the Timeless Present." *Monist* 50 (1966): 317–340. Reprinted in *The Pre-Socratics,* edited by A.P. D. Mourelatos, 271–292, Garden City, NY: Anchor Press, 1974, and in G. E. L. Owen, *Logic, Science and Dialectic. Collected Papers in Greek Philosophy,* edited by M. Nussbaum, 27–44, Ithaca, NY: Cornell University Press, 1986.

Reale, G. *Melisso. Testimonianze e frammenti. Introduzione, traduzione e commento.* Biblioteca di Studi Superiori. Filosofia antica, 50. Florence: Nuova Italia, 1970.

Sedley, D. "Parmenides and Melissus." In *The Cambridge Companion to Early Greek Philosophy,* edited by A. A. Long, 125–131. Cambridge, U.K., and New York: Cambridge University Press, 1999.

Solmsen, F. "The 'Eleatic One' in Melissus." *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen* (Amsterdam) n.s. 32 (8) (1969): 219–233.

Vitali, R. *Melisso di Samo, "sul mondo o sull' essere": Una interpretazione dell'Eleatismo.* Urbino, Italy: Argalìa, 1973.

Vlastos, G. Review of J. E. Raven, *Pythagoreans and Eleatics. Gnomon* 23 (1953): 29–35.

David J. Furley (1967)
Bibliography updated by Richard D. McKirahan (2005)

MEMORY

Remembering is one of the most characteristic and most puzzling of human activities. In particular, personal memory—the ability mentally to travel back into the past, as leading psychologist Endel Tulving puts it—often has intense emotional or moral significance: It is perhaps the most striking manifestation of the peculiar way human beings are embedded in time, and of humans' limited but genuine freedom from their present environment and immediate needs. Memory has been significant in the history of philosophy as much in relation to ethics and to epistemology as in theories of psyche, mind, and self.

The philosophy of memory is a fascinating, diverse, and underdeveloped area of study, which offers difficult but rewarding connections not only with psychology and the cognitive sciences, but also with the social sciences and political theory, and with literature and the arts. Outside philosophy, interest in memory increased massively and disproportionately in the late twentieth century in both the neurocognitive sciences and the humanities, driven both by internal developments within disparate disciplines and by wider social and cultural concerns about trauma and recovered memories, about the politics of forgetting and collective responsibility, about memory loss in an aging population, and about the manipulation, control, ownership, and protection of individual memory. The widespread and troubled fascination in Western culture with this last set of concerns in particular, and with challenging associated questions about moral psychology and personal identity, is suggested by the success of films like *Bladerunner* (1982), *Memento* (2000), and *Eternal Sunshine of the Spotless Mind* (2004).

As a result, just as in other areas of the philosophy of mind, it has become increasingly difficult to cordon off a set of questions about memory, or methods for its study, which are uniquely or primarily philosophical. Some philosophers treat memory as a case study in philosophy of science, asking for example whether the psychology of memory might be reducible to the neuroscience of memory. Others begin with the phenomenology of memory, the ordinary experiences and practices of remembering; others still inquire into cross-cultural or historical differences in these practices. It seems likely, further, that psychopharmacological influences on memory, and their potential misuse, will make memory a central topic in the emerging fields of neuroethics and philosophy of psychiatry. This entry covers more traditional philosophical issues about the nature of memory, but includes some consideration of the need for a broader framework that can encompass the neural, embodied, psychological, and social aspects of remembering.

FORMS OF REMEMBERING

When a person is remembering, there are many different activities he or she may be engaged in, and the expression of the individual's memory can take many different forms. One reminisces with old allies about shared experiences; one finally calls to mind that obscure fact; one mindlessly cycles off down the lane, despite not having been on a bike for years; one sits alone and ruminates on one joyful or agonizing moment long ago; one gathers with others to commemorate a significant occasion; one writes or fashions something in memory of a person or an event; a photo, an odd memento, or a long-forgotten melody suddenly immerses a person in the emotions of another time.

It is not easy to pinpoint just what is common across this range of activities, and some philosophers have argued that not all of them involve true memory. But the present-day consensus in both philosophy and psychology is that there are at least three distinct forms of remembering that can helpfully be detected in the variety of ordinary experience.

First, in remembering specific events or episodes from an individual's personal past he or she draws on *personal memory* (also known as *experiential* or *event memory*): For example, one remembers walking down by the river with a friend that spring afternoon. Psychologists often call this *episodic memory*, or sometimes *autobiographical memory*.

A different form of memory is naturally expressed with a "that" complement: One remembers that Aristotle was Alexander's tutor. This *factual* or *semantic memory* is akin to simple belief, and the remembered facts can be about events in the remote past, or indeed the future, as well as personally experienced events. One can factually remember details one has been told about one's early life, for example, for which one has no personal memory, no sense at all of what the past experiences were like.

In English, and many other languages, people sometimes contrast things that they "just know" from what they genuinely (personally) remember, thus treating personal memory as the basic or essential kind of memory. But in other contexts people are happy to talk also of remembering facts, and to attribute their general beliefs about the world to "memory" in a broader sense.

Personal/episodic remembering and factual/semantic remembering are both forms of *declarative memory*, in which individuals seek to hook up to reality, to represent the world or the past. Although remembering activities often have quite different functions as well, under normal circumstances such memories aim at truth. This is so even though, as both scientific and common-sense psychology increasingly suggest, people do not always get there. The point is not that memory necessarily or even reliably achieves this aim, but that one's ordinary practices include a general commitment to its reliability in doing so. For example, an individual may or may not in fact have walked by the river with a friend that spring afternoon, and Aristotle may or may not actually have been Alexander's tutor. But if one is sincerely expressing that personal memory, or that factual memory, one is (among other things) making a claim about what happened.

In these declarative forms of memory, the content of one's memory can in principle —at least in central cases—be articulated. But when a person wonders if a friend remembers how to play the flute, or how to drive a car, the person is asking not about the friend's personal or factual memories, but about his or her skills or embodied memories. Philosophers have often talked here of *habit memory*, while psychologists identify these cases as types of *procedural memory*, where this category is also taken to include more basic/primitive forms of conditioning and associative learning.

Procedural memory has been sharply divided from declarative memory for a number of reasons: Perhaps most important is the case of H.M., an epileptic patient who suffered terrible amnesia after brain surgery in the 1950s. H.M., who had lost his hippocampus and other brain structures now known to be central to declarative memory, was no longer able to lay down event memories, so that he would forget everything minutes after its occurrence, and lose any clear sense of time passing. Yet H.M. was still able to learn new games, and to improve his performance at new perceptual-motor skills, despite having no idea each time that he had ever tried them before.

Procedural memory is philosophically important for a number of reasons, although habits and skilled activities have been little studied. For example, neither philosophers nor psychologists have a clear grip on the various ways that personal memory and other high-level cognitive processes interact with remembered embodied skills. Competition and coordination between the different memory systems can both occur. On the one hand, skilled performers in dance or sport know that their motor habits often run best in a groove, when not consciously or verbally controlled: yet the skills involved are robust and flexible, unlike more primitive forms of procedural memory, and can sometimes be directly shaped by mood, context, verbal instruction, and conscious decision.

These conceptual, grammatical, and experiential distinctions between personal, factual, and habit memory have in contemporary cognitive psychology been developed into theories of distinct memory systems. There is considerable disagreement about the psychological status of these systems, and about whether the distinction between episodic and semantic memory, in particular, should be characterized by reference merely to the kind of information in question, or by an essential phenomenological difference. Since there is little agreement more generally about what a psychological system or module is, or about the nature of any putative natural kinds in psychology, these debates about memory systems are likely to be resolved only in conjunction with progress on broader questions in philosophy of psychology.

PERSONAL MEMORY

An individual's capacity to conjure up experiences, emotions, and events from long in the past involves the same kind of memory as the mundane ability to keep track of just what he or she has been doing, feeling, and thinking in the last day or week. Personal remembering does not seem to be distinguished from other related activities—imagining, dreaming, factual remembering, for example—by the level of sensory detail or vividness which it involves: some memories, after all, are both faint and fragmentary, while some scenes of fantasy can be richly imagined. Memory capacities, even in their normal and reliable functioning, are both fallible and selective: human beings don't need either total or precise recall to maintain sufficient coherence and continuity of self over time, for personal memory works in part through an ongoing condensing, editing, and summarizing of life experiences, on which people draw in specific autobiographical narratives. One's narratives or other memory expressions can be public or private, and they can be more or less under one's control, either smoothly tailored to specific audiences or emerging in involuntary fragments.

Personal remembering is a context-sensitive activity from the start. As young children build on their earlier abilities to understand typical sequences of events, their capacity to remember particular past experiences is supported and shaped by adults. Joint attention to the shared past emerges in an interactive social environment, as chil-

dren come to see that there can be different perspectives on the same past time. Spontaneous self-conscious thought about the personal past is a gradual development out of these memory-sharing practices, which can vary considerably in nature, frequency, and significance across contexts and cultures. One condition for the full emergence of such self-conscious thought about his or her own past experiences, which may be surprisingly late, is that the child picks up the causal connections between events in time, and within the child's own history. Some grasp of the temporal asymmetry of experience is needed to understand that, in principle at least, remembered events can be integrated on a connected temporal dimension. Children's personal memory, then, is a highly sophisticated achievement closely linked not only to their emerging self-awareness and understanding of other minds, but also to their recognition that they cannot change the past, and that their current and future actions are unique and irrevocable.

Because early personal remembering is socially situated in this way, it is also tightly meshed with emotional and social/moral development. Key social practices, such as promising and forgiving, and some central complex emotions, such as grief, love, and regret, depend essentially on personal memory and on one's grasp of temporal relations. The point here is not just that the fallible but more-or-less reliable operation of memory in two or more people is needed to give those people current informational access to the past times at which their paths have crossed. Memory's affective tone and influence means that, in addition to its role in retaining the past, it also has a forward-looking function, as Richard Wollheim argued in his *Thread of Life* (1984): Remembering can keep what happened in the past alive, giving it significance for one's ongoing relationships and projects. According to this view, memory is not just a means for checking on the continuity of the self over time, but also itself partly produces or creates personal identity: As Wollheim puts it, the past affects people in such a way that they become creatures with a past.

The particular ways in which, through memory, individuals deal with events and experiences that are no longer present varies according to context and aim. Most dramatically, for example, legal contexts impose demands and standards on the memory narratives witnesses must produce that differ greatly from the norms operating in other remembering activities. But questions about the reliability of memory and about its mechanisms arise in many different circumstances just because memory, with its orientation to truth, is in these ways intimately involved in both personal identity and significant social practices. Two connected lines of thought have raised the most serious concerns about people's access to the past in remembering: philosophical views about representations and memory traces, and psychological accounts of the constructive nature of remembering.

THEORIES OF MEMORY

People can, sometimes, remember past events and experiences in the absence of immediate external cues or prompts to memory. It is natural, then, to think that somehow individuals carry around with them what they will need in order to remember when circumstances are right. Even one's ordinary conception of memory, C. B. Martin and Max Deutscher argued in their influential causal analysis *Remembering* (1966), requires the existence of an appropriate causal link between one's past experience and one's present remembering. Although the notion of the "memory trace" has appeared in many strange metaphors and theories in the history of philosophy and the history of science, it need be no richer than this idea of a state that causally connects experience and remembering in a certain way. This causal analysis embeds the theory of memory in the broader *representational theory of mind* which has come to characterize mainstream philosophy of cognitive science; however the bare invocation of memory traces is compatible with many quite different views about their nature and operation.

However, even this basic view about memory traces, in the eyes of its critics, engenders serious problems about the nature of a person's access to the past. If the past is thus truly lost, so that a person can only make contact with it by examining certain representations in the present, critics complain, there is a real danger of scepticism, to be countered by affirming that the person is in fact aware of the past directly in memory. The ensuing, long-running debate between representative realists or indirect realists, who accept memory representations, and direct realists or phenomenologists who reject them, is exactly parallel to that found in theories of perception. Although the dichotomous nature of this debate no longer fits the range of positions available, and many quite different views are often condensed by critics into a monolithic target, there is some common ground.

Contemporary trace theorists tend to work in a broadly materialist framework, and do not in general think of traces as *direct* objects of awareness from which the nature of the past is consciously inferred at the personal level of psychological analysis. If complex noncon-

scious processes, operating subpersonally on representations which may themselves be partial and context-sensitive, are involved in the shaping and constructing of the contents of memory, this does not mean that the experience of remembering is indirect. On this point, the positive direct realist contribution is convincing: Remembering, under normal circumstances, is a kind of immersion in which one has a pre-reflective confidence.

But this idea that an individual typically inhabits the memory, rather than judging and assessing it for plausibility and coherence, is in fact entirely compatible with the existence and involvement of subpersonal mechanisms operating on enduring but modifiable traces. Such mechanisms can be typically reliable even if they are fallible in particular instances. To raise a general skeptical worry again at this point against the invocation of memory representations would be unrealistically to demand incorrigible access to the past, to seek a blanket guarantee of accuracy in memory. Such blunt certainty about memory was expressed, for example, by the eighteenth-century Scots philosopher Thomas Reid, the most ardent critic of philosophies of "ideas" or "traces," who wrote that "those things really did happen which I distinctly remember" (Reid, *Essays*, 1849, p. 444). But this renders the indisputable evidence—both everyday and scientific—of errors in memory quite mysterious, and thereby threatens to erode commonsense realism about the past.

Theorists who posit memory traces are also criticized for adherence to what is seen as an arbitrary metaphor of "storage," unfortunately entrenched in the philosophy of memory since Plato's *Theaetetus*. The bare retention of capacities or dispositions to act or respond in certain ways, the critics complain, implies nothing about the means by which such capacities are retained: Storage is a mistakenly concrete way of thinking, as if each memory had to be stashed away separately, like sacks of grain in a storehouse or fixed entries in an archive. Ludwig Wittgenstein, for example, mocked the static but inaccessible inner records he identified in the psychological theories of his time: In notes of 1935–1936, he wondered "whether the things stored up may not constantly change their nature" (Stern 1991, p. 204).

In some invocations of memory representations, each trace has indeed been treated as distinct, with each single remembered item mapped on to one storage element. Such atomist or localist representational schemes make control over the contents of memory easier to imagine or achieve: The remembered items are passive, and must be manipulated or altered by an external executive. In this separation of data from process, ordinary digital computers exemplify the localist memory scheme: But what is "stored" in human memory displays more intrinsic dynamics than this, tending in some contexts naturally to interfere, blend, and generalize without deliberate or voluntary control. But just as such computers do not exhaust possible computational devices, so localist representational schemes are not essential to the general framework of memory traces. Both historical theories of memories as patterned flows of "animal spirits" through the pores of the brain, and contemporary connectionist models in cognitive science employ distributed (rather than localist) representation: What can be distinctly remembered need not be held distinctly or independently, since each item is spread or "superposed" across many elements in a system or network. This entry examines the implications of these distributed models of memory after setting them in the context of recent developments in cognitive psychology and the cognitive sciences.

REMEMBERING AND THE COGNITIVE SCIENCES

The recent history of the sciences of memory offers a sharp contrast and corrective to the stereotyped image of cognitive science as a scientistic quest to reduce the human mind to the dull mechanism of digital computers. Memory research was one of the first areas to be taken out of the lab in the 1980s and 1990s, as psychologists sought to address the kinds of memory that matter in everyday life (such as autobiographical memory), and to find ecologically valid methods of studying such memories outside artificial isolated situations. The difficulty facing philosophers or scientists with an urge toward synthesis is not that psychological results are irrelevant to wider concerns about memory, but that the daunting diversity of methods and traditions even within cognitive psychology makes it hard to see how different levels of explanation might relate to one another. There are issues of considerable interest for the philosophy of science in understanding the connections between neuroscientific and cognitive-psychological descriptions and methods; and, equally, robust and philosophically intriguing research traditions on autobiographical memory in developmental, personality, and social psychology. This entry briefly examines ideas about the constructive nature of remembering that seem to have direct relevance to concerns about truth in memory.

Remembering is a multifaceted activity that takes place in the present, and so the best explanatory frameworks for understanding it will attend closely to the con-

text of recall, rather than simply investigating the nature of encoded traces. Memories are often compiled or constructed for particular purposes when needed, not held fully formed. There is room for considerable internal plasticity in memory traces, which are (on the connectionist model) always composites shaped by the entire history of their network. The most dramatic work on construction in memory has come not from connectionist modeling, however, but from the research on suggestibility and false memory by Elizabeth Loftus and her colleagues (2003). Misleading information from external sources can be incorporated into personal recollection. Confident, entrenched childhood "memories" of spilling a bowl of punch at a wedding, for example, or of gazing long at an exceptionally colorful mobile in the days after birth, can be elicited artificially in certain circumstances. This work is partly motivated by a wish to confirm the possibility of false confessions, in which individuals may come sincerely and passionately to believe that they have committed horrible crimes in the past; but the mechanisms in play are just the ordinary and normally robust processes of shaping and generalizing memories to make them fit. Although Loftus has adopted the high moral tone of a crusade, ongoing careful investigation of individual differences and integration of these results with social and personality psychology promises a much richer picture of the conditions which make different kinds of distortion more likely.

Again, the point of this research is not to show, implausibly, that reliability in memory is impossible or unlikely. Psychologists assume that understanding the mechanisms of distortion will also throw light on the processes involved in veridical remembering. Reliability and accuracy are not transparent notions here. Prereflective confidence in personal memories can, and in certain contexts should, coexist with attention to the other evidence about the past which is often available, and care for the defeasible but subtle and robust capacities to winnow evidence that individuals have developed in the rich and complex social context of early memory-sharing and memory-using practices.

SOCIAL MEMORY AND SHARED MEMORY

The general constructive picture of remembering can be accepted while acknowledging that external influences—particularly social influences—on memory need not inevitably lead to error. As Sue Campbell argued in her powerful philosophical responses to the "memory wars," there are vital features of relational interaction with others that contribute positively to practices of good remembering, both in development and in adult social life: To treat the true unit of memory as the isolated individual, free from the distorting influence of other people, is to miss the value we often appropriately place on negotiating the past—both the personal past and the shared past—in company.

Indeed a need for attention to shared remembering and social remembering in both psychology and philosophy can be motivated from within the broadly constructivist framework itself. It is because one's internal memory is partial and context-sensitive, and does not naturally retain information in distinct and unchanging form between experience and recollection, that one relies so pervasively and—in the main— successfully on external social and technological scaffolding. A challenge for psychologists is to find ways to study shared memories that do not focus solely on the conformity induced or sought by powerful external authorities; and a challenge for philosophers is to construct a social ontology of memory by which to understand the diverse ways in which people manage to hook their incomplete inner systems of traces with the vast social and cultural resources in which cognition is situated.

Mark Rowlands (1999) and Rob Wilson (2004), for example, have suggested specific ways in which external symbol systems—in their many distinct historical and cultural forms—allow individuals to leave information and skills out in the world, saving on the resources and capacity required for biological memory. Drawing on the more precise invocations of terms like *social memory* and *collective memory* in the contemporary social sciences, this distributed cognition framework suggests that researchers can study the transmission of particular representations across different individuals and media, and the specific forms of interplay between group dynamics and individual recall. It also promises to throw better light on the influential work on memory by the French sociologist Maurice Halbwachs (1980).

Halbwachs's notion of the collective memory is often cited by contemporary social scientists and historians as deeply anti-psychological, or as sociologically determinist: but in fact his work focuses on the incomplete or shrouded nature of the individual's memory, which (outside of dreams) must be sculpted and completed within a social framework, which provides the context and the means for the construction of a specific recollection. Philosophical analysis can potentially be of immense service to empirical disciplines like cognitive anthropology and historical theory in the study of memory by

showing how case studies of remembering activities in particular times and places might be embedded in robust broader theories of memory. So in addition to the long-standing philosophical concerns about truth and the self previously outlined, it is likely that philosophical attention will increasingly engage, through topics like memory, with the urgent challenge of connecting the cognitive sciences and the social sciences.

See also Cognitive Science; Computing Machines; Moral Psychology; Personal Identity; Philosophy of Mind; Plato; Reid, Thomas; Time, Consciousness of; Wittgenstein, Ludwig Josef Johann.

Bibliography

Bartlett, Frederic. *Remembering: A Study in Experimental and Social Psychology*. Cambridge, U.K.: Cambridge University Press, 1932.

Bursen, Howard. *Dismantling the Memory Machine*. Dordrecht: Reidel, 1978.

Campbell, John. "The Structure of Time in Autobiographical Memory." *European Journal of Philosophy* 5 (1997): 105–118.

Campbell, Sue. *Relational Remembering: Rethinking the Memory Wars*. Lanham, MD: Rowman and Littlefield, 2003.

Campbell, Sue. "Models of Mind and Memory Activities." In *Moral Psychology: Feminist Ethics and Social Theory*, edited by Peggy DesAutels and Margaret Urban Walker. Lanham, MD: Rowman & Littlefield, 2004.

Carruthers, Mary. *The Book of Memory*. Cambridge, U.K.: Cambridge University Press, 1990.

Casey, Edward S. *Remembering: A Phenomenological Study*. Bloomington: Indiana University Press, 1987.

Connerton, Paul. *How Societies Remember*. Cambridge, U.K.: Cambridge University Press, 1989.

Deutscher, Max. "Remembering 'Remembering.'" In *Cause, Mind, and Reality*, edited by J. Heil, 53–72. Dordrecht: Kluwer, 1989.

Draaisma, Douwe. *Metaphors of Memory: A History of Ideas about the Mind*. Translated by Paul Vincent. Cambridge, U.K.: Cambridge University Press, 2000.

Draaisma, Douwe. *Why Life Speeds Up as You Get Older: How Memory Shapes Our Past*. Translated by Arnold and Erica Pomerans. Cambridge, U.K.: Cambridge University Press, 2004.

Engel, Susan. *Context Is Everything: The Nature of Memory*. New York: W. H. Freeman, 1999.

Foster, Jonathan K., and Marko Jelicic, eds. *Memory: Systems, Process, or Function?* New York: Oxford University Press, 1999.

Hacking, Ian. *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton, NJ: Princeton University Press, 1995.

Halbwachs, Maurice. *On Collective Memory*. Translated by F. J. and V. Y. Ditter. New York: Harper & Row, 1980.

Hamilton, Andy. "False Memory Syndrome and the Authority of Personal Memory-Claims: A Philosophical Perspective." *Philosophy, Psychiatry, and Psychology* 5 (1998): 283–297.

Hoerl, Christoph. "Memory, Amnesia, and the Past." *Mind and Language* 14 (1999): 227–251.

Hoerl, Christoph, and Teresa McCormack, eds. *Time and Memory: Philosophical and Psychological Perspectives*. New York: Oxford University Press, 2001.

Hyman, Ira E., and Elizabeth F. Loftus. "Errors in Autobiographical Memory." *Clinical Psychology Review* 18 (1998): 933–947.

Krell, David Farrell. *Of Memory, Reminiscence, and Writing: On the Verge*. Bloomington: Indiana University Press, 1990.

Loftus, Elizabeth F. "Our Changeable Memories: Legal and Practical Implications." *Nature Reviews Neuroscience* 4 (2003): 231–234.

Malcolm, Norman. *Memory and Mind*. Ithaca, NY: Cornell University Press, 1977.

Margalit, Avishai. *The Ethics of Memory*. Cambridge, MA: Harvard University Press, 2002.

Martin, C. B., and Max Deutscher. "Remembering." *Philosophical Review* 75 (1966): 161–196.

Martin, M. G. F. "Perception, Concepts and Memory." *Philosophical Review* 101 (1992): 745–763.

McClelland, James L., and David E. Rumelhart. "A Distributed Model of Human Learning and Memory." In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 2, edited by J. L. McClelland and D. E. Rumelhart, 170–215. Cambridge, MA: MIT Press, 1986,

Misztal, Barbara. *Theories of Social Remembering*. Maidenhead: Open University Press, 2003.

Neisser, Ulric. "The Ecological Study of Memory." *Philosophical Transactions of the Royal Society B: Biological Sciences* 352 (1997): 1697–1701.

Nelson, Katherine. "Self and Social Functions: Individual Autobiographical Memory and Collective Narrative." *Memory* 11 (2003): 125–136.

Owens, David. "The Authority of Memory." *European Journal of Philosophy* 7 (1999): 312–329.

Owens, David. "A Lockean Theory of Memory Experience." *Philosophy and Phenomenological Research* 56 (1996): 319–332.

Reid, Thomas. "Essays on the Intellectual Powers of Man." In *The Works of Thomas Reid*, edited by W. Hamilton. Edinburgh: McLachlan, Stewart, & Co., 1849.

Ricoeur, Paul. *Memory, History, Forgetting*. Translated by Kathleen Blamey and David Pellauer. Chicago: Chicago University Press, 2004.

Rowlands, Mark. *The Body in Mind: Understanding Cognitive Processes*. Cambridge, U.K.: Cambridge University Press, 1999.

Rubin, David C. *Memory in Oral Traditions: The Cognitive Psychology of Epic, Ballads, and Counting-Out Rhymes*. New York: Oxford University Press, 1995.

Schacter, Daniel. *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books, 1996.

Schacter, Daniel. *The Stranger Behind the Engram: Theories of Memory and the Psychology of Science*. Hillsdale, NJ: Lawrence Erlbaum, 1982.

Schechtman, Marya. "The Truth about Memory." *Philosophical Psychology* 7 (1994): 3–18.

Sorabji, Richard. *Aristotle on Memory*. London: Duckworth, 1972.

Stern, David G. "Models of Memory: Wittgenstein and Cognitive Science." *Philosophical Psychology* 4 (1991): 203–218.

Sutton, John. *Philosophy and Memory Traces: Descartes to Connectionism*. Cambridge, U.K.: Cambridge University Press, 1998.

Tulving, Endel. "Episodic Memory: from Mind to Brain." *Annual Review of Psychology* 53 (2002): 1–25.

Tulving, Endel, and F. I. M. Craik, eds. *The Oxford Handbook of Memory*. New York: Oxford University Press, 2000.

Warnock, Mary. *Memory*. London: Faber, 1987.

Wilcox, Stephen, and Stuart Katz. "A Direct Realist Alternative to the Traditional Conception of Memory." *Behaviorism* 9 (1981): 227–239.

Wilson, Robert A. *Boundaries of the Mind*. Cambridge, U.K.: Cambridge University Press, 2004.

Wollheim, Richard. *The Thread of Life*. Cambridge, U.K.: Cambridge University Press, 1984.

John Sutton (2005)

MENASSEH (MANASSEH) BEN ISRAEL

(1604–1657)

Menasseh (Manasseh) ben Israel, the Jewish scholar, philosopher, and theologian, was probably born in Madeira. His father, a victim of the Spanish Inquisition, escaped with his family to La Rochelle and then to Amsterdam, where Menasseh studied in the growing Jewish community. At eighteen he became a teacher and preacher. Although very successful in his rabbinical career, Menasseh could not support his family with his salary and so became a printer, establishing Holland's first Hebrew press. He printed his own first published work, an index to the *Midrash Rabbah* (1628). Most of his subsequent works are in Spanish, Portuguese, or Latin.

Menasseh's vast erudition in Jewish and Christian theology and philosophy and classical and contemporary literature attracted notice in 1632, when the first part of his *El Conciliador* appeared in Frankfurt (the second, third, and fourth parts appeared in Amsterdam, 1641–1651; the book was translated into English by E. H. Lindo, London, 1842). This work attempted to reconcile the apparent conflicts and contradictions in the Bible and brought Menasseh into the company of Gerhard Johannes and Isaac Vossius, Hugo Grotius, and many other scholars, who came to regard him as the leading expositor of Jewish thought to the Christian world. He corresponded with Christian and Jewish scholars everywhere, and many came to Amsterdam to confer with him.

Menasseh ben Israel was greatly interested in the Jewish and Protestant kabbalistic, mystical, and Messianic views of his time and was involved with some of the strangest seventeenth-century visionaries. This led to his most famous work and the best-known episode of his career. A Portuguese Jew from South America told him of finding some of the lost tribes of Israel in the jungles there. Using this material and other "data," Menasseh ben Israel published his *Hope of Israel* in Latin, Spanish, and English (1650), in which he argued that because the Israelites were spread almost everywhere on Earth, the Messianic age was at hand. If the Jews were readmitted to England, then all might be ready for the Messiah. Several influential Puritans, including Oliver Cromwell, held similar views, and they invited Menasseh ben Israel to London to discuss the readmission of the Jews. Menasseh ben Israel stayed in England from 1655 to 1657, but after much controversy no official solution emerged, although the unofficial readmission of Jews to England did begin. Disappointed, Menasseh ben Israel died shortly after leaving England.

Although his works are not of the first rank, Menasseh ben Israel was extremely influential in developing and disseminating a modernized form of Jewish learning and in making Christian scholars aware of then-current streams of Jewish thought.

See also Grotius, Hugo; Jewish Philosophy; Kabbalah.

Bibliography

The Hope of Israel appeared in Spanish in 1650 (Amsterdam) and was translated by Menasseh ben Israel into Latin the same year. An English translation by M. Wall appeared in London in 1650. The latest English edition was published in London in 1901.

Cecil Roth, *A Life of Menasseh ben Israel* (Philadelphia: Jewish Publication Society of America, 1934), contains an excellent bibliography of works by and about Menasseh ben Israel. Also consult the articles "Manasseh ben Israel" in *Jewish Encyclopedia* (London and New York, 1904), Vol. VIII, pp. 282–284, and *Dictionary of National Biography* (Oxford, 1959–1960), Vol. XII, pp. 898–899.

OTHER RECOMMENDED TITLES

Coppenhagen, J. H., ed. *Menasseh ben Israel (1604–1657): A Bibliography*. Jerusalem: Misgav Yerushalayim, 1990.

Méchoulan, H., and R. H. Popkin. *Menasseh ben Israel and His World*, edited by Y. Kaplan. Leiden: Brill, 1989.

Roth, C. *A Life of Menasseh ben Israel*. Philadelphia: Jewish Publications Society, 1934.

Richard H. Popkin (1967)

Bibliography updated by Oliver Leaman (2005)

MENCIUS

(fourth century BCE)

Mencius, a Chinese philosopher, is often regarded as the most important Confucian thinker after Confucius. He lived in the Warring States period, during which China was divided into different states with their own rulers, often waging war against each other. He traveled from state to state to convert rulers to the teachings of Confucius. At the same time, he also combated other influential movements of thought, especially those associated with Mozi and Yang Zhu (fifth to fourth century BCE). One's main access to his thinking is through the *Mengzi* (Mencius), probably compiled by his disciples or disciples of his disciples. The text was subsequently edited and shortened by Zhao Qi in the second century CE, and this is the version of the text available today.

Elaborating on Confucius's teachings, Mencius highlighted four ethical attributes: *ren* (benevolence, humaneness), *yi* (propriety), *li* (observance of rites), and *zhi* (wisdom). *Ren* has to do with love or concern for others and involves a reluctance to cause harm and the capacity to be moved by the suffering of others. The scope of such concern includes not just human beings but also certain kinds of animals, and there is a gradation in *ren* in that one has special concern for and obligations to those closer to oneself. *Ren* results from cultivating the special love for parents that everyone shares as an infant and the affective concern for others shown in the well-known Mencian example of one's commiseration for the infant on the verge of falling into a well.

The earlier use of *yi* refers to a proper regard for oneself and distancing oneself from disgrace, involving such things as not brooking an insult. Mencius retained this use of *yi*, but disgrace for him is measured not by ordinary social standards but by ethical standards, and *yi* has to do with a firm commitment to such standards. One regards what falls below such standards as potentially tainting oneself and insists on distancing oneself from such occurrences even at the expense of death. One example is that of a beggar starving to death, who would reject food given with abuse despite the resulting loss of life. According to Mencius everyone shares responses of this kind, which provide the starting point for cultivating *yi*.

Li originally referred to rites of sacrifice and later to rules of conduct governing ceremonial behavior as well as behavior in other social contexts. Mencius continued to use *li* in this way, and in addition used it to refer to an ethical attribute having to do with the observance of *li*. This attribute involves a general disposition to follow *li*, as well as a mastery of the details of *li* that enables one to follow *li* with ease. It also involves one's observing *li* with the proper attitude and mental attention, such as reverence in interacting with others or sorrow in mourning.

In early Chinese thought, *xin*, which refers to the physical heart, is regarded as the site of both cognitive and affective activities. It is translated as "heart" or "mind," and sometimes as "heart/mind." *Xin* can form certain directions, which can take the form of long-term goals in life or more specific intentions. The fourth ethical attribute, *zhi*, involves having proper directions of the heart/mind, which in turn requires an ability to assess situations without adhering to fixed rules of conduct. This discretionary judgment may lead one to deviate from established rules of *li*, and may also guide one's behavior in situations in which no general rule is applicable.

For Mencius, these four ethical attributes result from people cultivating four kinds of predispositions of the heart/mind. These include commiseration, the sense of shame, a reverential attitude toward others, and the sense of right and wrong. He referred to these as the four "sprouts" or "beginnings" and regarded the four ethical attributes as growing from these predispositions in the way that a plant grows from a sprout. Besides commiseration and the sense of shame, he also regarded love for parents and obedience to elder brothers as the starting point for cultivating *ren* and *yi*, respectively. His view that the heart/mind has these ethical predispositions provides the basis for his response to the Moist and Yangist challenges.

Mozi advocated the doctrine of indiscriminate concern for everyone. He did not believe that human beings have the appropriate predispositions to begin with and thought that one could restructure one's motivations accordingly after endorsing this doctrine. In the absence of such predispositions, the practice of indiscriminate concern seems humanly impossible, a point seized on by Mozi's opponents. By contrast, Mencius thought that human beings have ethical predispositions that relate to the ethical ideal in the way that a sprout relates to a full-grown plant. Such predispositions contain within them a direction of growth and provide the appropriate emotional resources that one can draw on to achieve the ideal.

The Yangists advocated nourishing *xing* (nature), a term referring to the direction of growth or development of a thing. They understand the *xing* of human beings in biological terms, such as living to an old age, and regarded it as the proper direction of development for humans. Mencius rejected the biological conception of

xing, instead, *xing* is constituted by the ethical direction implicit in the predispositions of the heart/mind. The view that *xing* has an ethical direction is expressed in his well-known slogan that *xing* (human nature) is good.

Although the heart/mind has the relevant ethical predispositions, they need to be nourished for them to flourish, and one should also guard against the various factors that can potentially harm their growth. Mencius often highlighted the senses as something that can lead one astray. The senses operate automatically—when they come into contact with their ideal objects, they are just pulled along unreflectively by these objects. By contrast, the heart/mind can reflect on what is proper and can halt any course of action it regards as improper. The heart/mind should constantly exercise these capacities to ensure that one progresses in an ethical direction.

One may also be led astray by erroneous doctrines, such as Mohist and Yangist teachings, which Mencius explicitly opposed. One may also be led astray by problematic desires. For example, in a series of dialogues between Mencius and King Xuan of the state of Qi, the king referred to his great desire to expand territories and his feverish desires for wealth, women, and display of valor. These desires not only led the king to harsh policies but also led him to rationalizations about his inability to be caring toward his people. Mencius's response was to try to steer the king toward seeing that a more caring policy toward the people is not only compatible with the king's desires but actually enables their attainment in a higher form. For example, a king who seeks to be invincible can do so by practicing *ren* government, thereby drawing the allegiance of the people. He will become invincible not in the sense of superior military strength, but in the sense of being without opposition.

While Mencius's teachings competed for influence with other kinds of Confucian teachings for several hundred years after his time, he eventually came to be regarded as the true transmitter of Confucius's teachings. Zhu Xi included the *Mengzi* as one of the Four Books, which became canonical texts of the Confucian tradition. Mencius also came to be regarded as the greatest Confucian thinker after Confucius himself, and his teachings have been influential on the development of Confucian thought in the Song (960–1279), Ming (1368–1644), Qing (1644–1912) Dynasties, and up to modern times.

See also Chinese Philosophy; Confucius; Mozi; Yang Zhu; Zhu Xi (Chu Hsi).

Bibliography

Chan, Alan K. L., ed. *Mencius: Contexts and Interpretations*. Honolulu: University of Hawaii Press, 2002.

Graham, A. C. "The Background of the Mencian Theory of Human Nature." In *Studies in Chinese Philosophy and Philosophical Literature*, 7–66. Albany: SUNY Press, 1990. Originally published in the journal *Tsing Hua Journal of Chinese Studies* 6 (1967).

Graham, A. C. *Disputers of the Tao: Philosophical Argument in Ancient China*. La Salle, IL: Open Court, 1989.

Lau, D. C., trans. *Mencius*. London: Penguin, 1970.

Legge, James, trans. *The Works of Mencius*. 2nd ed. Oxford, U.K.: Clarendon Press, 1895.

Nivison, David S. *The Ways of Confucianism: Investigations in Chinese Philosophy*. La Salle, IL: Open Court, 1996.

Shun, Kwong-loi. *Mencius and Early Chinese Thought*. Stanford, CA: Stanford University Press, 1997.

Kwong-loi Shun (2005)

MENDELSSOHN, MOSES

(1729–1786)

Moses Mendelssohn, the greatest Jewish philosopher in the eighteenth century, was born in Dessau, the son of a poor Jewish copyist of sacred scrolls. His first studies were devoted to the Bible, the Talmud, and Maimonides' *Guide for the Perplexed.* He followed his teacher Rabbi David Fränkel to Berlin in 1745, where he learned to read German and Latin while living in great poverty. In 1750 he became a tutor in the household of the Jewish silk manufacturer Isaak Bernhard; he was later a bookkeeper and ultimately a partner in Bernhard's firm. In Berlin Mendelssohn became a close friend of G. E. Lessing, C. F. Nicolai, and Thomas Abbt. After 1755 his reputation as a philosopher and critic grew rapidly throughout Germany. By his contemporaries he was regarded as eminently kind and virtuous, and because of his wisdom and ugliness he was called "The Jewish Socrates." Lessing is said to have modeled the character of Nathan in his drama *Nathan der Weise* upon Mendelssohn. In 1763 Mendelssohn's *Abhandlung über die Euidenz in den metaphysischen Wissenschaften* (Essay on Evidence in Metaphysical Science; Berlin, 1764) won a prize from the Berlin Academy, and he was later elected to the academy, although his appointment was never confirmed.

In spite of his Jewish extraction, Mendelssohn's development as a philosopher was notably German in character; he was influenced mainly by Gottfried Wilhelm Leibniz, Christian Wolff, Alexander Baumgarten, G. F. Meier, his Berlin friends, and among foreign philosophers, by John Locke, the earl of Shaftesbury, Edmund

Burke, Jean Baptiste Dubos, and Pierre-Louis Moreau de Maupertuis.

Mendelssohn was a typical "popular philosopher." He was empirically minded, refrained from final systematizations of his theories, wrote in an easy and attractive style, and was mainly interested in aesthetics, psychology, and religion (although he also discussed methodological and metaphysical questions). His contribution to the emancipation of the Jews was significant. Because of the continuous evolution of his ideas, a summary of his views can only cover the general trends of his thought. He exerted a great influence not only upon his closest friends but upon his whole generation in Germany, and upon Immanuel Kant in particular.

Aesthetics and psychology were, in Mendelssohn's mind, closely interrelated. He continued the work of Baumgarten and Meier, but amalgamated their doctrines with the tenets of English and French aesthetics translated into the terminology of German psychology. Generally attributed to Mendelssohn is the first clear distinction between Beauty and metaphysical perfection: He held that Beauty was an inferior, subjective kind of perfection. Metaphysical perfection consists in unity in a multiplicity. Aesthetic perfection arises out of the limits of human understanding. Man is unable to conceive, as God can, the real, supreme unity in the enormous variety of things. He must therefore content himself with introducing an artificial unity (uniformity) into some objects in order to be able to perceive them as wholes; and this is beauty.

In this way, Mendelssohn began a trend away from Baumgarten's and Meier's aesthetic objectivism toward a subjective aesthetics that soon dominated German aesthetics: A beautiful object is not necessarily perfect in itself, but must be perfect in its capacity to be perceived. The perception of Beauty strengthens the representative activity of the soul and makes it more perfect, thus causing a feeling of pleasure. The perception of Beauty causes intuitive knowledge; in its highest stage it becomes the "aesthetic illusion" in which, for example, fable appears as reality. Mendelssohn's conception of Beauty permitted him to explain the pleasurable effect of tragedy and of the sublime, whose distinction from Beauty he was the first in Germany to explain clearly. In tragedy, murder is the representation of a morally and metaphysically imperfect event, but its representation may be subjectively perfect. Mendelssohn, clearly under the influence of Burke, held that in the sublime, the pleasure in awareness of immensity of distance, size, or number is mixed with some pain because of our inability to comprehend it completely. In both cases, aesthetic pleasure is the result of the "mixed feeling" (*vermischte Empfindung*) arising in our soul: Even if some element of the perception is unpleasant, the perception as a subjective whole is pleasurable.

Mendelssohn's study of the perception of Beauty led him to introduce a doctrine of mental faculties that was later adopted in modified form by Kant and others. Mendelssohn held that aesthetic feelings must be attributed to a faculty different from intellect and desire, a faculty that he called the faculty of approval (*Billigungsvermögen*). The beauty of an object escapes us if we subject it to a process of analysis and definition; therefore, experience of the beautiful cannot be an object of knowledge. A beautiful object gives us aesthetic pleasure even if we do not possess the object; thus, the approval of Beauty must be distinct from desire. Metaphysical perfection, unlike Beauty, is both known by intellect and an object of desire.

Beauty is produced by genius. Genius does not imitate nature, but "idealizes" it; that is, it exhibits natural objects as God would have created them if his aim had been aesthetic and not metaphysical perfection. Genius is independent of rules because it establishes its own rules. A genius's procedure is instinctive.

Mendelssohn believed that both the existence of God and the immortality of the soul could be demonstrated. Although his *Morgenstunden oder Vorlesungen über das Daseyn Gottes* (Morning Hours, or Lectures on the Existence of God; Berlin, 1785) was written in awareness of Kant's previously published *Kritik des reinen Vernunft,* in it Mendelssohn accepted both the Ontological Argument and the Argument from Design.

Mendelssohn's *Phädon oder über die Unsterblichkeit der Seele* (*Phaedo,* or on the Immortality of the Soul; Berlin, 1767) was a dialogue on immortality in imitation of Plato's *Phaedo.* The soul is a simple substance and therefore indestructible. The soul might nevertheless lose its consciousness, but the divine wisdom and goodness of God would not allow this to happen.

Mendelssohn's plans to publish a work commemorating Lessing, who had died in 1781, prompted Friedrich Heinrich Jacobi to write to Mendelssohn asking whether he knew that Lessing was a Spinozist. The resulting quarrel, which soon involved Johann Georg Hamann, Johann Gottfried Herder, and Johann Wolfgang von Goethe as well as Mendelssohn and Jacobi, is discussed in the entry "Pantheismusstreit."

Mendelssohn had been challenged in 1769 by the Swiss physiognomist and religious writer Johann Kaspar Lavater either to demonstrate the falsity of Christian rev-

elation or to become a convert to Christianity. Mendelssohn's answer was that the deism of the Enlightenment, which he had developed into a universal religion of reason, was in fact identical with Judaism. In his *Jerusalem oder über religiöse Macht und Judentum* (Jerusalem, or on Religious Power and Judaism; 2 vols., Berlin, 1783), Mendelssohn supported religious and political toleration, and advocated separation of church and state and civil equality for the Jews. He always fought against both advocates of anti-Semitism and conservative Jews for a cultural and political union of Christians and Jews.

See also Pantheismusstreit.

Bibliography

WORKS BY MENDELSSOHN

Philosophische Gespräche. Berlin, 1755.

Briefe über die Empfindungen. Berlin, 1755.

Betrachtungen über die Quellen und die Verbindungen der schönen Künste und Wissenschaften. Berlin, 1757.

Moses Mendelssohn an die Freunde Lessings. Berlin, 1786.

Werke. 7 vols, edited by E. G. B. Mendelssohn. Leipzig, 1843–1844.

Gesammelte Schriften, edited by D. Elbogen, J. Guttmann, and E. Mittwoch. Berlin: Akademie-Verlag, 1929–.

Jerusalem. Translated by Allan Arkush. Hanover, NH: University Press of New England, 1983.

Philosophical Writings. Translated and edited by D. Dahlstrom. Cambridge, U.K.: Cambridge University Press, 1997.

WORKS ON MENDELSSOHN

Altmann, Alexander. *Moses Mendelssohn: A Biographical Study.* Tuscaloosa: University of Alabama, 1973.

Arkush, Allan. *Moses Mendelssohn and the Enlightenment.* Albany: State University of New York Press, 1994.

Bamberger, F. *Die geistige Gestalt M. Mendelssohns.* Frankfurt, 1929.

Cahn, N. *M. Mendelssohns Moralphilosophie.* Giessen, 1921.

Cassirer, Ernst. "Die Idee der Religion bei Lessing und Mendelssohn." In *Festgabe zum zehn jährigen Bestehen der Akademie für die Wissenschaft des Judentums,* 22–41. Berlin, 1929.

Cohen, B. *Über die Erkenntnislehre M. Mendelssohns.* Giessen, 1921.

Goldstein, L. *M. Mendelssohn und die deutsche Aesthetik.* Königsberg, 1904.

Hoelters, Hans. *Der Spinozistische Gottesbegriff bei M. Mendelssohn und F. H. Jacobi, und der Gottesbegriff Spinozas.* Bonn, 1938.

Kayserling, Moses. *Moses Mendelssohn, sein Leben und sein Wirken.* Leipzig, 1862; 2nd ed., 1888.

Morgan, Michael. "Mendelssohn." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman. London: Routledge, 1997.

Pinkus, F. *M. Mendelssohns Verhältniss zur englischen Philosophie.* Würzburg, 1929.

Richter, L. *Philosophie der Dichtkunst, M. Mendelssohns Aesthetik.* Berlin, 1948.

Ritter, J. H. *Mendelssohn und Lessing.* 2nd ed. Berlin: Steinthal, 1886.

Sander, D. *Die Religionsphilosophie Moses Mendelssohns.* Erlangen, 1894.

Zarek, O. *M. Mendelssohn.* Amsterdam: Querido, 1936.

Giorgio Tonelli (1967)
Bibliography updated by Oliver Leaman (2005)

MENTAL CAUSATION

There is mental causation whenever a mental state, event, process, or activity has a causal effect. The pursuit of our lives seems replete with mental causation. It may thus seem as obvious that it occurs as we pursue our lives. But how mental causation is possible is not obvious. And therein lies a philosophical tale. Any attempt to explain how it occurs must engage the mind-body problem.

René Descartes (1596–1650) maintained that there is body-to-mind causation when we perceive our surroundings, and mind-to-body causation when we act. But one of the most serious charges leveled again his substance dualism, according to which the mind is an immaterial substance that is not extended in space, is that it leaves unexplained *how* mental states and events (etc.) have causal effects on our bodies. Descartes held that the locus of mind-body causal interaction is in the brain (specifically, in the pineal gland). His contemporary, Princess Elisabeth of Bohemia, asked how states of, or changes in, a substance not extended in space (the mind) could causally affect states of, or changes in, a substance extended in space (the brain or pineal gland), and declared such causal interaction too incredible to believe. The absence of a satisfactory answer to her "how-question" contributed to the demise of Cartesian substance-dualism (Watson 1987).

Many contemporary philosophers hold that to have a mind is not to possess an immaterial substance, but rather to possess certain capacities, such as the capacity to think and/or to feel. Brains serve somehow as the material basis of such capacities. (Whether an artificial brain could so serve is the question of whether artificial intelligence is possible.) But because of the many apparent differences between mental and physical properties, some philosophers, while rejecting Cartesian substance dualism, nevertheless embrace Cartesian property dualism. They hold that while there are no immaterial substances, mental properties are distinct from physical properties, and are related to certain of them by irreducible laws of

nature. This view faces the question of how an individual's having a mental property could exert any causal influence on the course of events. Given the absence of a reality underlying both mental and physical reality, an individual's having a mental property would have to exert a direct causal influence on its initial effects in the brain, one unmediated by any mechanism.

The year 1870 marked more than a century of increasingly detailed investigation of human physiology. In that year, Ewald Herring declared at his lecture to the Imperial Academy of Sciences in Vienna that brain physiologists should make "the unbroken causative continuity of all material processes an axiom of [their] system of investigation" (translated and quoted in Butler 1910, pp. 64–65). It remains an axiom of neurophysiology. The fact that there are no "gaps" in physiological brain processes for mental events to fill led Thomas Huxley (1874) to maintain we are "conscious auotmata": conscious events accompany certain physiological brain events as dual effects of other physiological events, but are causally inert. Trained as a medical doctor, William James (1890) appropriated the term *epiphenomena*, a medical term for symptoms of diseases, for mental phenomena that while caused, lack causal efficacy. James Ward (1903) coined the term *epiphenomenalism* for the view that mental phenomena have no causal effects.

The view that mental phenomena are epiphenomena has a dense air of paradox. Epiphenomenalists maintain that we are merely under the illusion that there is mental causation. But, on their view, the illusion could not give rise to our belief in mental causation, for that would require mental causation. Moreover, on pain of inconsistency, they cannot take themselves to have been led to the doctrine by theoretical reasoning, for their being so led would involve mental causation. Indeed, reasoning itself seems to be a causal process. It should thus come as no surprise that virtually no contemporary philosophers who acknowledge the reality of the mental espouse the view that no mental states or events have causal effects. But the question of how they have effects remains.

Some philosophers combine the rejection of Cartesian substance dualism with the rejection of mental and physical event dualism, while nevertheless embracing Cartesian property dualism. C. D. Broad (1925) examined a dual-aspect theory of events, according to which physiological events in "the mind-brain" (1925, p. 439) have two independent aspects, one mental, the other physiological, the two linked by contingent fundamental laws. In discussion of the view, he formulated epiphenomenalism as a disjunctive doctrine: "mental events either (a) do not function at all as cause factors; or that (b) if they do, they do so in virtue of their physiological characteristics, and not in virtue of their mental characteristics" (p. 473). If, rather than being accompanied by mental events, certain physiological events have mental characteristics, and so are mental events, then it seems, on the evidence, that they function as cause factors in virtue of their physiological characteristics, but not their mental ones. The mental qua mental seems causally inert.

Donald Davidson (1970) proposed the doctrine of anomalous monism: every particular mental event is a physical event, but there are no strict psychological or psychophysical laws, and mental characteristics are irreducible to physical characteristics. He did not, however, embrace Cartesian property dualism, which is committed to fundamental psychophysical laws. Moreover, he regarded talk of properties as pleonastic; strictly speaking, there are only predicates, not properties. He held that since mental events (i.e., events mental predicates are true of) are causes or effects, they fall under strict physical laws, and so are physical events because physical predicates that figure in the relevant strict laws are true of them. Still the causal relation, he emphasized, is extensional: if two events are causally related, they are so related however they are described. There is no qua-causation.

Many philosophers hold that properties are distinct from predicates, and indeed that predicates apply to things only in virtue of the properties that things have. And they hold that although the causal relation is indeed extensional, it is nevertheless the case that events enter into causal relations in virtue of certain of their properties. The weighs-less-than relation is extensional: If *a* weighs less than *b*, then it does so however *a* and *b* are described. Still *a* weighs less than *b* in virtue of something about each of them, namely their respective weights—their respective masses in the gravitational context in question. Anomalous monism entails the denial of token epiphenomenalism. But its proponents must answer the charge of commitment to type epiphenomenalism, the thesis that no events are causally related in virtue of falling under mental types (McLaughlin 1989, 1994; Kim 1993; Sosa 1993; see also Davidson 1993).

In the early twentieth century, the atomic view of matter was vindicated, and in the 1930s a quantum mechanical explanation of chemical bonding was provided, dispelling the idea that there are fundamental chemical forces; and later monumental advances in organic chemistry and molecular biology led to the demise of any form of vitalism (McLaughlin 1992). It is

now generally held, on empirical grounds, that: for any (caused) microphysical event *P* there is a distinct microphysical event *P** that causally determines the objective probability of *P* (if determinism is true, that probability will be 1).

This thesis has been called by various names in the literature, including "the closure of the physical." Given this thesis, if Cartesian property dualism is correct, then it seems that an individual's having a mental property could have microphysical effects only if it causally overdetermined those effects. Such overdetermining psychophysical causal transactions would be fundamental in that they would be unmediated by any mechanism. While that may fall within the realm of logical possibility, it is hard to see how the view that it actually occurs could be justified (Kim 1998).

Many contemporary philosophers hold that there is a stronger dependence of mental properties on microphysical properties than Cartesian property dualism allows. There is no received formulation of the dependency. But one leading view is that it is captured by the following *supervenience* thesis: any minimal physical duplicate of the actual world is a duplicate *simpliciter* of it (Jackson 1998). A physical duplicate of the actual world is any world that is exactly like the actual world in every microphysical respect, in respect to its worldwide pattern of distribution of microphysical properties and relations, its worldwide pattern of distribution of microphysical objects, its microphysical laws of nature, and so on. A minimal physical duplicate of the actual world is any physical duplicate of it that contains nothing other than what is metaphysically required to be a physical duplicate of it.

While the supervenience thesis is incompatible with Cartesian property dualism, it does not entail that every property is a microphysical property. The thesis entails that any minimal physical duplicate of the actual world will have exactly the same worldwide pattern of distribution of properties as the actual world. But, as should be made clear below, that does not require that every property be a microphysical property. Indeed, one can embrace the supervenience thesis while holding a kind of property pluralism, according to which not only mental properties, but properties that figure in the laws of the special sciences—economics, psychology, biology, and even most of chemistry—are not microphysical properties. Some proponents of the supervenience thesis are property pluralists and hold, in addition, (token) event and state pluralism, on the grounds that events and states are property exemplifications. They thus hold that mental events, and events within the domains of the special sciences, are not microphysical events. Let us label this kind of "nonreductive physicalism," which combines the supervenience thesis with property and event pluralism, "NRP."

NRP theorists acknowledge that every event is such that its objective probability is causally determined by some microphysical event occurring across some cross section of its backward light cone. But they deny that this excludes higher-level events from being causes. Some defend this denial by distinguishing causation from causal determination (Yablo 1992). They hold that to be causally related, events must be appropriately proportional, and that microphysical events are typically disproportional to the higher-level events they causally determine, and are thus disqualified as causes of those events. On this view, when the turning of a key causes a lock to open, some microphysical event will causally determine that the lock opens. But it will not be a cause of the lock's opening. The reason is that it contains too much superfluous detail to be suitably proportionate to the opening of the lock. Had the key turning occurred without that microphysical event, the lock would still have opened. The key turning thus "screens off" the microphysical event vis-à-vis the lock's opening. Of course, in the counterfactual situation that is stipulated, some other microphysical event will underlie the key turning and cause the microphysical event underlying the lock's opening. But it is claimed that is so because higher-level causal transactions are implemented by lower-level ones, and ultimately by microphysical ones.

One charge against this view is that it mistakes causal explanation for causation. Any microphysical event that causally determines the opening of the lock causes it. Nevertheless, an explanation of why the lock opened in terms of a microphysical cause would be an extremely poor one indeed in a typical context since it would contain far too many details that are superfluous to understanding why the lock opened. But whether that charge can be justified remains a matter of dispute. The dispute turns on controversial issues about the nature of causation and the individuation of events.

Many NRP theorists hold that every event is caused by some microphysical event that determines its objective probability. They maintain, nevertheless, that higher-level events are causes. One concern about this view is that if higher-level events were causes, then their effects would include microphysical events. If my decision to walk into the next room causes me to walk into the next room, a result will be that many of the physical particles making

up my body at the time of the decision will end up in the next room. The decision would be a cause (though not of course a sufficient cause) of the movements of the particles. Such "downward causation" is regarded by some philosophers as untenable (Kim 1998). NRP theorists respond that while the movements of the particles are in a sense causally overdetermined, such overdetermination is not the objectionable sort to which the interactionist Cartesian property dualist is committed. For the psychophysical causal interactions are not fundamental: They are implemented by causal transactions between microphysical events. Mechanics can ignore them. Still some critics charge that the fact that the microphysical event was brought about by another microphysical event leaves no work for the decision to do in bringing it about (Kim 1998). Some NRP theorists reply that this sort of worry is based on a productive conception of causation, and that we should eschew such a conception as unrealistic (Loewer 2002). They maintain that this sort of overdetermination can be accommodated by a kind of regularity account of causation (Melynk 2003), or a kind of counterfactual account of causation (Loewer 2002). This strand of the debate also leads to issues concerning causation and event individuation.

Given the supervenience thesis, any minimal physical duplicate of the actual world would have the same worldwide pattern of distribution of mental events and special science events as the actual world. Why is that the case if mental and special science events are not microphysical events? The leading NRP answer is that all mental and special science events are *realized* by microphysical events and such realization guarantees this result. While there is no received view of realization, the leading notion is the functionalist notion, according to which the realization relation is the relation of role-occupancy: a realization is a role-player. This idea, however, has been implemented in two different ways (see Block 1980). Role-functionalism implements it one way; filler-functionalism implements it in another (see McLaughlin forthcoming).

According to role-functionalism, every event token of a mental type *M* is a higher-order event token, an event of participating in some event or other that occupies a certain role *R*, which includes a causal role. Events that occupy *R* realize *M* events, that is, realize events that are exemplifications of *M*. On this view, higher-order events are never identical with lower-order events. Thus, even if mental events are always realized by microphysical events, no mental event is a microphysical event; similarly, for special science events. This event pluralism is compatible with the supervenience thesis because the basic roles could be filled by microphysical events that fill them in virtue of microphysical laws and conditions.

But NRP theorists would nevertheless face a problem in embracing role-functionalism, for there is a serious question of whether higher-order events have causal effects. While every second-order event is realized by a first order event that has causal effects, a serious question remains whether second-order events themselves have effects. The role-functionalist idea seems most plausible for abilities, but abilities themselves seem not to have causal effects, rather their bases or realizations do. The role-functionalist idea has, however, also been interestingly applied to constituted dispositional states, such as water-solubility, water-absorbency, fragility, ductability, and the like (Jackson, Pargetter, and Prior 1982; Prior 1985). For something to be water-soluble is (arguably) for it to be in some state that, under appropriate conditions, would cause it to begin to dissolve when immersed in a liquid. The state that has the causal role of producing the maninfestation of the disposition (dissolving) is the basis (realization) of the disposition. (Being composed of sodium chloride is one such basis; but the dispositional property is multiply realizable.) It is, however, the basis of water-solubility that causes the substance to dissolve when immersed in water, not the disposition—*if* the disposition is indeed a second-order state (other accounts of such states are possible). On this role-functionalist conception, the substance's being water-soluble seems to just be the fact that there is some state of it that would (in appropriate circumstances) result in its dissolving were it immersed in water.

The concern, then, is that if (token) mental states and events were functional states and events (i.e., higher-order states and events), they would have no causal effects (Jackson 1996, McLaughlin forthcoming). That would not exclude them from being causally explanatory. The claim that a substance dissolved in water because it is water-soluble provides some information about the causal chain leading to its dissolving (see Prior 1985). But the NRP theorist is after higher-level causation, not just causal explanation. Thus, the NRP theorist must respond to this concern with a compelling account of causation according to which functional states indeed have causal effects. Suffice it to note that the claim that functional states are inefficacious does not presuppose a productive conception of causation (see Lewis 1986).

According to filler-functionalism, an event is of mental type *M* if and only if it occupies or plays a certain role *R*, where *R* includes a causal role. On this view, an event

token *realizes* role *R* by occupying the role—by filling it. For an event to be of type *M* is just for it to fill the role. Thus, if *E* occupies *R*, then *E* is thereby of type *M*. Since the role includes a causal role, filler-functionalists reject token-epiphenomenalism. Note that if, on a particular occasion, event token *E* is the occupant of *R*, then "*E* is the *M* event" will be a contingent statement of identity, like "Benjamin Franklin is the inventor of bifocals." (The description "the *M* event" will, like the description "the inventor of bifocals," be nonrigid: it will pick out different things in some possible worlds from those that it picks out in others.)

It may well be that tokens of various types of events can occupy role *R*, and thus be realizations of *M*; if so, then *M* is multiply realizable. Moreover, events of some type *N* can realize *M*, even when *N* itself is multiply realizable. That will be the case when an event is of type *N* if and only if it fills a role *R**, which includes *R* as a proper sub-role (Shoemaker 1994). If, on a particular occasion, an event realizes *M* in virtue of being an *N* event, and realizes *N* in virtue of being a *C* event, then, on that occasion, the *C* event is the *N* event, the *N* event is the *M* event, and so the *C* event is the *M* event.

Notice, then, that, when conjoined with the thesis that every mental event is realized by some microphysical event, filler-functionalism entails that every mental event is a microphysical event. And indeed the filler-functionalist explanation of why any minimal physical duplicate of the actual world will have the same worldwide pattern of distribution of mental (and special sciences) events as the actual world is that the only basic fillers of the roles are microphysical events, which fill them solely in virtue of microphysical laws and conditions. Events are of different orders only relative to types. (Moreover, the ordering here, it has been pointed out, is not one of scale [Kim 1998].) The filler-functionalist account of realization will not serve the NRP theorist's purposes. On the filler view, every event is a microphysical event, and it is ultimately in virtue of microphysical event types that events enter into causal relations. Mental event types are not microphysical event types, both because of actual multiple microphysical realization, and because of the logical possibility of realization without microphysical realization. Nevertheless, they are relevant to whether events of one sort cause events of another since they implicitly type events in terms of patterns of causal relations. And that may very well make them indispensable to certain causal explanations. But whether such a view is correct turns, of course, not only on the nature of causation and the individuation of events, but also on the nature of mental (and special science) properties.

Problems remain, moreover, that are specific to the mental. Some philosophers maintain that neither a role nor a filler-functionalist view is tenable for mental states with qualitative or phenomenal characters: states such that it is like something for the subject of the state to be in the state (e.g., the state of feeling pain). And some embrace Cartesian property dualism for phenomenal mental properties ("qualia"; Chalmers 1996, Kim 2005). They thus reject the psychophysical supervenience thesis. They hold that there could be an exact physical duplicate of the actual world that, unlike the actual world, is entirely devoid of phenomenal consciousness (a "zombie world"; Chalmers 1996). But they do not deny the closure of the physical. And they acknowledge that they may thus very well have to hold that an individual's having a phenomenal property has no causal effects. Suffice it to note that even this restricted epiphenomenalism has an air of paradox. It entails, for instance, that our feeling of pains never cause our pain-behavior, or even our beliefs that we are in pain.

Moreover, even if Cartesian property dualism is rejected for all mental properties, problems remain. Intentional mental states are explanatory, in part, by virtue of their propositional contents. For example, the content that *there is a snake in the room* figures essentially in both the rationalizing explanation, "He decided not to enter because he believed there was a snake in the room," and the nonrationalizing explanation, "He began to quiver because he feared that there was a snake in the room." The leading theories of content, however, are externalist theories, according to which the content of a mental state fails to supervene on intrinsic states of the subject (Putnam 1975, Burge 1979). On these views, two intrinsic duplicates (e.g., an inhabitant of Earth and her *doppelgänger* on Twin Earth) could be in intentional states with different contents. Indeed, according to some externalist theories, content depends on historical context (Dretske 1988), and according to others, on social context (Burge 1979). Some philosophers maintain that such highly relational properties are causally irrelevant to behavior, and so must play a noncausal explanatory role. But some philosophers defend the view that intentional states cause behavior, despite being essentially extrinsic (Yablo 1999). Others claim that wide content is causally explanatory because it provides information about the causal history of the agent's behavioral dispositions (Dretske 1988). And others contend that intentional states have an externalist or wide content in virtue of hav-

ing a "narrow content" in a causal environmental context, and that it is narrow content that is causally relevant to behavior (Jackson 1996). There are other views as well that are as yet less explored. Suffice it to note that these content issues too are matters of ongoing philosophical investigation.

See also Anomalous Monism; Artificial Intelligence; Broad, Charlie Dunbar; Cartesianism; Consciousness; Content, Mental; Davidson, Donald; Descartes, René; Dualism in the Philosophy of Mind; Elisabeth, Princess of Bohemia; Functionalism; Huxley, Thomas Henry; James, William; Kim, Jaegwon; Mind-Body Problem; Nonreductive Physicalism; Philosophy of Mind; Putnam, Hilary; Qualia; Supervenience.

Bibliography

Block, N. *Readings in Philosophy of Psychology*. Vol. I. Cambridge, MA: Harvard University Press, 1980.

Broad, C. D. *The Mind and Its Place in Nature*. New York: Harcourt Brace, 1925.

Burge, T. "Individualism and the Mental." *Midwest Studies in Philosophy* 4 (1979): 73–121.

Butler, S. *Unconscious Memory*. London: A. C. Fifield, 1910.

Chalmers, D. J. *The Conscious Mind*. New York: Oxford University Press, 1996.

Descartes, René. *Meditions on First Philosophy* (1641). Edited by Stanley Tweyman. Ann Arbor: Caravan Books, 2002.

Davidson, Donald. "Mental Events." In *Experience and Theory*, edited by L. Foster and J. W. Swanson. University of Massachusetts Press and Duckworth, 1970. Reprinted in Donald Davidson, ed. *Actions and Events*. Oxford: Clarendon Press, 1980.

Davidson, Donald. "Thinking Causes." In *Mental Causation*, edited by John Heil and Alfred Mele. Oxford: Clarendon Press, 1993.

Dretske, F. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press, 1998.

Huxely, T. H. "On the Hypothesis That Animals are Automata, and Its History." *Fortnightly Review* 16 (1874): 555–580. Reprinted in Huxley, *Collected Essays*, Vol. 1. New York: J. A. Hill, 1904.

Jackson, F. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Jackson, F. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.

Jackson, F. "Mental Causation: The State of the Art." *Mind* 105 (1996): 377–413.

Jackson, F., R. Pargetter, and E. Prior. "Functionalism and Type-Type Identities." *Philosophical Studies* 42 (1982): 209–225.

James, W. *The Principles of Psychology* (1890). Cambridge, MA: Harvard University Press, 1983.

Kim, J. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.

Kim, J. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press, 2005.

Kim, J. *Supervenience and Mind: Selected Philosophical Essays*. New York: Cambridge University Press, 1993.

Lepore, E., and B. Loewer. "Mind Matters." *Journal of Philosophy* 84 (1987): 630–642.

Lewis, D. "An Argument for the Identity Theory." *Journal of Philosophy* 63 (1966): 17–25.

Lewis, D. "Events." In *Philosophical Papers* II, edited by D. Lewis, 241–269. Oxford: Oxford University Press, 1986.

Loewer, B. "Comments on Jaegwon Kim's Mind and the Physical World." *Philosophy and Phenomenological Research* LXV (2002): 655–662.

Macdonald, C., and G. Macdonald. "How to be Psychologically Relevant." *Debates on Psychological Explanation*. Vol. I, 60–77. Oxford: Blackwell, 1995.

McLaughlin, B. P. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989): 209–235.

McLaughlin, B. P. "The Rise and Fall of British Emergentism." In *Emergence or Reduction?*, edited by A. Beckermann, H. Flohr, and J. Kim, 49–93. Berlin and New York: Walter de Gruyter, 1992.

McLaughlin, B. P. "On Davidson's Response to the Charge of Epiphenomenalism." In *Mental Causation*, edited by John Heil and Alfred Mele, 27–40. Oxford: Clarendon Press, 1993.

McLaughlin, B. P. "Epiphenomenalism." In *A Companion to the Philosophy of Mind*, edited by S. Guttenplan, 277–289. Oxford: Blackwell, 1994.

McLaughlin, B. P. "Is Role-Functionalism Committed to Epiphenomenalism?" *Journal of Consciousness Studies*. Forthcoming.

Melnyk, A. *A Physicalist Manifesto*. Cambridge, U.K.: Cambridge University Press, 2003.

Prior, E. *Dispositions*. Scots Philosophical Mongraphs. Aberdeen: Aberdeen University, 1985.

Putnam, H. "The Meaning of 'Meaning,'" In *Mind, Language and Reality*, edited by H. Putnam, 215–271. Cambridge, U.K.: Cambridge University Press, 1975.

Shoemaker, S. *Identity, Cause, and Mind: Philosophical Essays*. Cambridge, U.K.: Cambridge University Press, 1994.

Sosa, E. "Davidson's Thinking Causes." In *Mental Causation*, edited by edited by J. Heil and A. Mele. Oxford: Clarendon Press, 1993.

Ward, S. L. "The Conscious Automism Theory." Lecture XIII of *Naturalism or Agnosticism*, vol. 2, 34–64. Adam and Charles Black, 1895/1903.

Watson, R. A. *The Breakdown of Cartesian Metaphysics*. Atlantic Highlands, NJ: Humanities Press, 1987.

Yablo, S. "Mental Causation." *Philosophical Review* 101 (1992): 245–280.

Yablo, S. "Wide Causation." *Philosophical Perspectives* 11 (1999): 251–281.

Brian P. McLaughlin (1996, 2005)

MENTAL CONTENT

See *Content, Mental*

MENTAL IMAGERY

See *Imagery, Mental*

MENTAL-PHYSICAL DISTINCTION

The distinction between the mental and the physical is central both to commonsense thinking about the world and to many philosophical, scientific, and religious theories. Perhaps it is as important to human thought as the distinction between fact and value, and between the empirical and the a priori. This entry will focus both on the role of the distinction in analytic philosophy and on various proposals about how it is to be understood.

The mental/physical distinction plays a role in two main areas of philosophy. First, in philosophy of mind, many arguments and issues are formulated in terms of it. Philosophers who advance physicalist theories about the mind argue that phenomenal consciousness (for example) is a physical phenomenon similar in kind to electricity or sexual reproduction; dualists deny this, saying that what we have here are two fundamentally different sorts of thing or two different characteristics of things. Second, in the philosophy of science and related parts of metaphysics, there is the issue of how to formulate the picture of the world that is presented to us by modern science. Many contemporary philosophers assume that this picture is in essence a physicalist one, and mean by this that the world-\view implicit in modern science bears important affinities with the materialism (also known as physicalism) of the seventeenth and eighteenth centuries, in particular that of La Mettrie and Hobbes. A natural assumption is that to properly evaluate whether the worldview of modern science really *is* a kind of physicalism, and to fully understand the related dispute in philosophy of mind between physicalism and dualism, one would need to clarify the mental/physical distinction. So what exactly *is* it?

There seems to a tacit general understanding of the mental/physical distinction but no rigorous idea of how it is to be drawn *exactly*—the implicit understanding has not been made explicit. That we understand the distinction in some sense is indicated by the fact that we spontaneously sort various features or characteristics of people or animals into two lists, the mental and the physical. So, to focus on a particular person Jones, we have on the mental side the fact that he knows where his car keys are, has itchy feet, wants tickets to the opera, and so on. On the physical side, we have the fact that he weighs 170 pounds, is currently located in Detroit, Michigan, is moving in such and such a direction with such and such a speed, and so on. The problem comes when we try to say in any detail what the occurrences of "and so on" mean. What precisely places a feature in the mental list, and what distinguishes those on the mental list from those on the physical? What groups *weighing 170 pounds* together with *being currently located in Detroit, Michigan*, and sets it apart from *having itchy feet*? Or take some other property of Jones not mentioned so far: for example, that his brain is releasing certain hormones into his bloodstream—is *it* mental or physical? If, as it seems natural to say, it is physical, what makes it so?

There is no shortage of proposals in the literature about what makes it so, and more generally about how to understand the mental/physical distinction, but all of them face problems, and none commands widespread assent. What immediately follows is a brief catalogue. The first, and historically the most important, proposal is that of Descartes (1641). Descartes said that being physical (or material) is just being extended in space; likewise, he said, the essence of the mind is to think, to engage in the activity of thinking. Descartes went on to argue that, if this is the way to draw the mental/physical distinction, dualism in philosophy of mind is true. This clarification of the distinction is straightforward, but it also has a number of drawbacks. First, we think of matter as something that *occupies* space, rather than being identical to space—but Descartes notoriously makes no room for such a distinction. Second, there are intuitively physical forces—such as the force of gravity—that would not be classified as physical from Descartes' point of view. Third, the idea that the essence of the mind is to think apparently excludes mental states that are sensory rather than cognitive and those that do not involve some sort of mental activity.

The second proposal—one might view it as an updated version of Descartes—draws the mental/physical distinction by appealing to two ways in which we find out about the world: introspection and perception. On this view, something is mental just in case we can find out about it, at least in principle, by introspection, whereas something is physical just in case we can find out about it, at least in principle, by perception. But this proposal faces difficulties also. One problem is that many things that seem intuitively physical are not directly available to perception even in principle—for example, subatomic particles. One might weaken the criterion and say that something is physical just in case we can find out about it *either* by perception *or* by inference from perception. But

the problem now is that the mental states of other people are such that we can find out about *them* by inference from perception; hence the weakened account entails the physicality of those mental states. Another problem with this second proposal is that it is not clear what the category of introspection is. Introspection seems to be the faculty by which we find out about our own *mental* goings-on—but this drains the idea of content.

The third proposal, prominent in the work of Thomas Nagel (1974, 1986), explains the mental/physical distinction as a special case of the contrast between the subjective and the objective. One obvious problem here is that the distinction between the subjective and objective is itself unclear; it is no advance to take *subjective* to mean "mental." But Nagel himself interprets the distinction as concerning different conditions of understanding: An objective truth or fact is one that can be understood from more than one point of view, whereas a subjective truth or fact is one that can be understood from at most one point of view. One objection to this is that there are psychological phenomena that are objective in Nagel's sense; presumably, the psychological properties attributed to humans by theoretical as opposed to folk psychology are as objective as any anything else. (These properties are not available to introspection either—and this causes a problem for the previous proposal, too.) A second objection is that the distinction between mental and physical is now a distinction *within* the realm of things that can be understood. But it is quite unclear that something is physical only if it is understandable.

The two proposals we have just considered inherit from Descartes the idea that we need criteria *both* for the mental and the physical. But contemporary philosophers have also explored the more cautious idea that one might define directly what it is for something to be physical, leaving aside the question of what it is for something to be mental. Hence, the fourth proposal is that something is physical just in case it is the sort of thing that physical theory tells us about or perhaps is entailed by the sort of thing physical theory tells us about. The basic objection to this view is Hempel's dilemma (Hempel 1969; see also Crane and Mellor 1990). Hempel's dilemma is that if the physical theory in question is contemporary physics, this proposal entails that physicalism is obviously false—after all, nobody believes that contemporary physics is complete; on the other hand, if the physical theory in question is some idealized or future physics, then the proposal entails that physicalism is empty—after all, who knows what some idealized or future physics will include? Some (for example, Smart 1974) respond by asserting that it is rational to believe that contemporary physics is complete. Although there is something right about this—surely it *is* rational to believe contemporary physics—the implicit suggestion that we should *define* the physical in terms of contemporary physics is implausible. Medieval impetus physics (for example) is a false and outmoded theory, but the property that objects have according to it—namely, impetus—is a physical property nonetheless.

According to a fifth proposal—sometimes called the paradigm physical object view—something is physical just in case it is the sort of thing required by or entailed by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents (Block 1980; see also Feigl 1967). The basic idea of this view is that we have some paradigms of physical objects—trees, stones, planets, toasters—and that the physical is whatever you need to explain them. One problem with this view is that it is circular—it explains the physical in terms of physical objects. (The same problem afflicts the previous proposal, which defines the physical in terms of physical theories.). Another problem for this view is that if physical objects turned out very different from how they appear—if, for example, they had a spiritual essence—physicalism and idealism would on this view be indistinguishable.

Perhaps it is unsurprising on reflection that the proposals just reviewed run into difficulties; they are all attempts at saying something positive about what the physical consists in. The sixth proposal is the negative one of saying that *physical* just means "nonmental" (for example, Levine 2001). One problem with this idea is that it assumes some criterion or mark of what it is to be mental; for example, that something is mental just in case it has phenomenal character or intentionality or both. And someone might question or reject both proposals either singly or in combination. But the more serious problem for the *via negativa* is that, construed as a definition of the physical, it gets things quite wrong. A vitalist, for example, thinks that living things instantiate properties—*élan vital*—which are both nonmental and nonphysical. However, while vitalism might be as false and outmoded as medieval impetus physics, it is not self-contradictory.

In view of the fact that every extant proposal about how to clarify the mental/physical distinction faces problems, it is natural to wonder whether there *is* any clear distinction here at all. Perhaps this is a distinction that we draw in ordinary thought but is something that should be done away with in serious scientific or philosophical descriptions of the world. That is the proposal that a number of people have found themselves drawn to, including Chomsky (2000).

One response to this sort of scepticism is that it is driven by overly high standards of clarity. True, it is hard to clarify the mental/physical distinction, but this difficulty does not mean that there is no such distinction—for the same thing might be said for many interesting distinctions and concepts. A different (but consistent) response asks us to look again at why we wanted a clarification of the mental/physical distinction in the first place. If the answer is intellectual curiosity, the Chomksian view is as reasonable as any other. But Chomskian skepticism gains much of its power from the further idea that various intellectual projects in philosophy of mind and science *make no sense unless* the mental/physical distinction can be clarified. But in fact it is not clear that this is so. Earlier we noted that various projects in philosophy of mind and science are formulated in terms of the mental/physical distinction. But it does not follow that the distinction is *essential* to these projects. If the mental/physical distinction can be shown to play only an illustrative or inessential role in these projects, then skepticism about the distinction itself—whether or not it is warranted—will not be as consequential as it would otherwise appear to be.

See also Chomsky, Noam; Descartes, René; Dualism in the Philosophy of Mind; Hempel, Carl Gustav; Hobbes, Thomas; Idealism; La Mettrie, Julien Offray de; Nagel, Thomas; Philosophy of Science, History of; Philosophy of Science, Problems of; Physicalism.

Bibliography

Block, N. "Troubles with Functionalism." In *Readings in the Philosophy of Psychology*, vol. 1, edited by N. Block. Cambridge, MA: Harvard University Press, 1980.

Chomsky, Noam. *New Horizons in the Study of Language and Mind*. Cambridge, U.K.: Cambridge University Press, 2000.

Crane, T., and D. H. Mellor. "There Is No Question of Physicalism." *Mind* 99 (1990): 185.

Descartes, Rene. "Meditations on First Philosophy." In *The Philosophical Writings of Rene Descartes*, edited by John Cottingham. Cambridge, U.K.: Cambridge University Press, 1985. First published in 1641.

Feigl, H. *The "Mental" and the "Physical."* Minneapolis: University of Minnesota Press, 1967. First published 1958.

Hempel, C. "Reduction: Ontological and Linguistic Facets." In *Essays in Honor of Ernest Nagel*, edited by S. Morgenbesser, P. Suppes, and M. White. New York: St. Martin's Press, 1969.

Levine, J. *Purple Haze*. New York: Oxford University Press, 2001.

Nagel, T. *The View from Nowhere*. New York: Oxford, 1986.

Nagel, T. "What Is It Like to Be a Bat." *Philosophical Review* 4 (1974): 435–450.

Smart, J. J. C., "The Content of Physicalism." *Philosophical Quarterly* 28 (1974): 239—241.

Daniel Stoljar (2005)

MENTAL REPRESENTATION

"Mental representations" are the coin of contemporary cognitive psychology, which proposes to explain the etiology of subjects' behavior in terms of the possession and use of such representations. "How does a subject manage to move through her darkened bedroom without stumbling over the furniture? She has an accurate mental representation of the room's layout, knows her initial position in the room, and is able to use this representation, in roughly the way a mariner uses a chart, to navigate through the room." "How does a sighted subject manage to recover information, available in the retinal image, about 'what's where' in her environment? She computes a series of representations, using information present in the retinal image, that eventuates in a three-dimensional representation of the distal objects present in the subject's visual field." "Why do native speakers of English have difficulty recognizing the grammaticality of so-called garden-path sentences such as 'The horse raced past the barn fell'? In recovering the meaning of a sentence, a speaker first constructs a representation of the syntactic structure of the sentence. In the case of garden-path sentences, the parsing processes that construct this representation mistakenly take the sentence's subject noun phrase to be a complete sentence, thus concluding that the entire sentence is ungrammatical." Cognitive ethologists offer similar explanations of many animal behaviors: Foraging red ants are said to practice a form of dead reckoning to maintain a representation of their current location relative to their nest, which they use to find their way back; migratory birds are said to navigate using representations of various sorts (celestial, magnetometric, topographic, etc.) that are either innate or learned as juveniles.

If, as these explanations apparently assume, mental representations are real entities that play a causal role in the production of a subject's behavior, then presumably it makes sense to ask about the form in which the information contained in these representations is encoded. This question has been the focus of considerable debate, especially with respect to mental imagery. Descriptionalists argue that, subjective impressions to the contrary notwithstanding, all mental representation, including mental imagery, is descriptional in form; mental repre-

sentations are said to represent in a way similar to the ways linguistic descriptions represent. Descriptionalists subscribe to a language of thought hypothesis, according to which all human cognition is conducted in a quasi-linguistic medium. Pictorialists, by contrast, argue at least some mental representations, notably those involved in mental imagery, represent in ways similar to the ways pictures represent. The issues in dispute here are not straightforwardly empirical. Neither party believes that we literally have descriptions or pictures in our heads; rather, their claims are about similarities to the respective ways that pictures and descriptions represent. But it is precisely these similarity claims that render this debate obscure. What are the respective ways that pictures and descriptions represent, and what are the salient similarities such that if they hold they would justify characterizing mental representations as being of one form rather than the other? It is not obvious that there is a definitive answer to either of these questions.

To describe the representations to which psychological and ethological explanations appeal as mental is not to imply that their possessors are conscious of them; typically the representations are nonconscious or subconscious. Nor is it to imply that these representations are nonphysical; there is no commitment here to dualism. Psychologists and ethologists presume that the representations to which their explanations appeal are neurologically realized, physical structures. The point of describing the representations as mental is simply to emphasize the particular explanatory role that these representations play in these explanations. The explanations undertake to explain a kind of purposive behavior on the part of a subject, in which the particular behavior exhibited by the subject is typically modulated in a characteristic fashion, not only by the goal or purpose of the behavior, but also by the environment in which the behavior is exhibited. Thus, for example, our subject's movement through her darkened bedroom is modulated by her knowledge of the current layout of the room. The mental representations that figure in these explanations serve two distinct explanatory roles: (1) They explain why a subject behaves in one way rather than another—she behaves as she does because she currently has this particular representation rather than another, and this representation is causally efficacious in the etiology of her behavior—and (2) they explain how the subject's behavior manages to be modulated (in characteristic ways) by her environment. Mental representations are able to play this dual explanatory role by virtue of possessing both physico-formal and semantic (intentional) properties that are linked in such a way as to ensure that a subject's environment can modulate her behavior. Basically, the cognitive processes that make use of mental representations are causally sensitive to the physico-formal properties of these representations that encode their semantic properties in much the way that sound-reproduction processes are sensitive to the physico-formal properties of records, tapes, and CDs.

Commonsense psychological explanations of behavior standardly appeal to beliefs, desires, intentions, and other so-called propositional attitudes (e.g., "Jones went to the refrigerator because he wanted a beer and believed there to be one there"). Behaviorists and eliminativists have challenged the legitimacy of these explanations, arguing that propositional attitudes either do not exist or do not figure in the etiology of behavior. Impressed with the prominent explanatory role of mental representations in cognitive psychological and ethological explanations, many philosophers of mind, notably Jerry Fodor, have proposed establishing the materialistic respectability of these explanations by appeal to the notion of mental representation. Their strategy is to explicate propositional attitudes in terms of mental representations. They defend a doctrine called the representational theory of mind (RTM), which holds that possessing a propositional attitude (e.g., believing that it is sunny today) is a matter of having a mental representation that (1) expresses the propositional content of that attitude (viz., that it is sunny today) and (2) plays a causal-functional role in the subject's mental life and behavior characteristic of the attitude in question (viz., the characteristic role of beliefs in modulating goal-satisfying behavior). More formally, for any organism O, any attitude A toward the proposition P, there is a mental representation MR such that MR means that (expresses the proposition that) P and a relation R (which specifies the characteristic causal-functional role of the MRs that are associated with a given A); and O bears attitude A to P if and only if O stands in relation R to MR. So formulated, RTM is silent as to the form of the mental representations that express the propositional contents of attitudes; proponents of RTM, however, invariably assume that these representations are syntactically structured entities, composed of atomic constituents (concepts) that refer to or denote things and properties in the world. More colorfully, these representations are sentences in the language of thought. The structure and meaning of these sentential representations purportedly explain the particular semantic and causal properties that propositional attitudes exhibit.

RTM is clearly realist in its construal of propositional attitudes: It purports to explain, not only what they are, but also how they could have both the causal and seman-

tic properties that common sense attributes to them (viz., of being causally efficacious in the production of other thoughts and of behavior, and of being semantically evaluable, as, e.g., true or false). RTM is equally realist in its construal of mental processes, which, it holds, are causal sequences of the tokenings of mental representation. These sequences are said to be proof-theoretic in character, with the sequential states in a thought process functioning like premises in an argument. Thought processes are, like arguments, generally truth preserving.

Proponents of RTM claim to find strong empirical support for the doctrine in the apparent explanatory (and predictive) successes of cognitive science, whose theories are heavily committed to the existence of mental representations. Critics tend to dismiss this claimed support, arguing that what is at issue is not whether there are mental representations but whether there are mental representations with the particular properties demanded by RTM. Critics argue that propositional-attitude contents cannot always be paired with mental representations in the way that RTM requires: A subject may bear a certain attitude to a proposition but lack, among the many mental representations that cognitive scientific theories attribute to her, any mental representation of that particular proposition. Thus, for example, more than one critic has pointed out that, while David Marr's computational theory of early vision (see his *Vision* [1982]) attributes to the visual system the assumption that objects in the visual field are rigid in translation, the theory does not attribute to the visual system an explicit representation of that assumption; rather, the assumption is implicit in the operation of visual processes. Proponents, for their part, have tended to dismiss such counterexamples as "derivative" cases, arguing that RTM nonetheless holds for what they term the "core" cases of propositional attitudes. Such a response presumes that there is a non–question-begging characterization of the class of core cases. It also presumes that the class so characterized includes those propositional attitudes that figure in the commonsense psychological explanations that RTM is intended to vindicate. It remains an open question whether either of these presumptions can be met.

Other critics of RTM have challenged the doctrine's apparent commitment to "classical" cognitive architectures that presume a principled distinction between mental representations, on the one hand, and the computational processes that are defined over these representations, on the other. These critics point out that connectionist computational models of cognition do not preserve such a distinction, so that, if, as these critics presume, cognitive architecture is connectionist rather than classical, then RTM is untenable. Not surprisingly, proponents of RTM have been in the forefront of efforts to demonstrate that cognitive architecture is not connectionist.

Still other critics of RTM have focused on the semantics of the postulated mental representations, arguing that, if RTM is to provide a materialistic vindication of explanations that appeal to propositional attitudes, it must be possible to provide a "naturalistic" semantics, a theory of content, for these representations. By such a semantics these critics understand a materialistic account, invoking no intentional or semantic notions, of how it is possible for mental representations to have the semantic properties that they do (of being about things in the world, of being truth valued, etc.). There is general agreement among critics and proponents alike that none of the proposed naturalistic semantics is adequate, but, where critics see in these failures the symptoms of RTM's untenability, proponents see the beginnings of a difficult but eventually successful research project. There is disagreement among critics as to the import for cognitive science itself of there possibly being no naturalistic semantics for mental representations. Some argue that it would impugn the claimed explanatory role of mental representations; others argue that it would not. Whatever the upshot of these arguments, the untenability of RTM would not in and of itself impugn the explanatory role of mental representations in cognitive science, since that commitment to mental representations does not entail RTM. One can perfectly well be a representationalist in the way that most cognitive scientists are without also being a proponent of RTM.

See also Cognitive Science; Connectionism; Eliminative Materialism, Eliminativism; Imagery, Mental; Language of Thought; Mental Causation; Philosophy of Mind.

Bibliography

Block, N., ed. *Imagery.* Cambridge, MA: MIT Press, 1981. A collection focusing on debate between descriptionalists and pictorialists regarding mental imagery.

Crane, T. *Mechanical Mind: A Philosophical Introduction to Minds, Machines, and Mental Representation.* London: Routledge, 2003.

Devitt, M. *Coming to Our Senses: A Naturalistic Program for Semantic Localism.* Cambridge, U.K.: Cambridge University Press, 1996.

Dretske, F. *Naturalizing the Mind.* Cambridge, MA: MIT Press, 1995.

Dretske, F. *Perception, Knowledge and Belief.* Cambridge, U.K.: Cambridge University Press, 2000.
Field, H. "Mental Representation." *Erkenntnis* 13 (1978): 9–61.
Fodor, J. A. *Concepts: Where Cognitive Science Went Wrong.* Oxford: Oxford University Press, 1998.
Fodor, J. A. *The Elm and the Expert.* Cambridge, MA: MIT Press, 1994.
Fodor, J. A. "Fodor's Guide to Mental Representation." *Mind* 94 (1985): 55–97.
Fodor, J. A. "The Persistence of the Attitudes." In *Psychosemantics.* Cambridge, MA: MIT Press, 1987.
Gunther, Y., ed. *Essays on Nonconceptual Content.* Cambridge, MA: MIT Press, 2003.
Jackendoff, R. *Languages of the Mind: Essays on Mental Representation.* London: Bradford, 1995.
Marr, D. *Vision.* San Francisco: Freeman, 1982.
Matthews, R. "Is There Vindication through Representationalism?" In *Meaning in Mind: Fodor and His Critics,* edited by B. Loewer and G. Rey. Oxford: Blackwell, 1991.
Matthews, R. "Troubles with Representationalism." *Social Research* 51 (1984): 1065–1097.
McDowell, J. *Mind and World.* Cambridge, MA: Harvard University Press, 1994.
Pylyshyn, Z. "The Explanatory Role of Representations." In *Computation and Cognition.* Cambridge, MA: MIT Press, 1984.
Pylyshyn, Z. *Seeing and Visualizing: It's Not What You Think.* Cambridge, MA: MIT Press, 2003.
Sterelny, K. *The Representational Theory of Mind: An Introduction.* Oxford: Blackwell, 1991.
Stich, S. *Deconstructing the Mind.* Oxford: Oxford University Press, 1996.
Stich, S., and T. Warfield, eds. *Mental Representation: A Reader.* Oxford: Blackwell, 1994. An excellent collection of papers on theories of content for mental representation; includes Field (1978) and Fodor (1985).
Strawson, G. *Mental Reality.* Cambridge, MA: MIT Press, 1994.
Thau, M. *Consciousness and Cognition.* Oxford: Oxford University Press, 2002.
Tye, M. *Consciousness, Color, and Content.* Cambridge, MA: MIT Press, 2000.

Robert J. Matthews (1996)
Bibliography updated by Alyssa Ney (2005)

MERCIER, DÉSIRÉ JOSEPH

(1851–1926)

Désiré Joseph Mercier, a Thomist philosopher and Roman Catholic cardinal, was born in the Walloon section of Brabant, Belgium. At the end of his secondary education, Mercier decided to study for the priesthood; he studied philosophy and theology at the Malines Seminary for five years and subsequently at the University of Louvain. Ordained in 1874, he received the licentiate (equivalent to the current doctorate) in theology in 1877. The same year he was named professor of philosophy at the Malines Seminary, where he taught logic and psychology for the next five years.

The famous encyclical, *Aeterni Patris,* of Pope Leo XIII, urging the restoration of scholastic, particularly Thomistic, philosophy, was published in 1879. In 1882 a chair of Thomistic philosophy was established at Louvain, and Mercier was named to this post.

For the next several years, Mercier taught courses in the various branches of philosophy, always attempting to relate Thomism to contemporary issues; in the course of this effort, Mercier became convinced that the task of making Thomism a living philosophy would require the combined efforts of many specialists. Hence, he conceived the notion of establishing a special institute of philosophy, with the aim not only of offering courses in Thomistic thought but also of providing the staff and facilities for a genuine research center. After considerable difficulty the Institute of Philosophy was established in 1889 as an integral part of the University of Louvain, with Mercier as its first president. The Philosophic Society of Louvain (still active) was founded by Mercier in 1888; in 1894 this organization founded the philosophical quarterly *Revue néo-scolastique* (still published under the title of *Revue philosophique de Louvain*), with Mercier as its editor.

From 1893 to 1906, Mercier's life was intimately bound up with that of the institute. His teaching activity continued; he published widely; and in the face of many difficulties, he worked incessantly to build and maintain the quality of the institute. His success in this area is measured by the fact that Louvain quickly became an internationally recognized center for philosophical work, attracting students from all over the world.

In 1906 Mercier's career in philosophy was interrupted by his being named archbishop of Malines; he was made cardinal the following year. From this time until his death, Cardinal Mercier's immense energies were directed toward the organizational and pastoral duties of his office. The seven volumes of his *Oeuvres pastorales* (Louvain, 1911–1928) give some indication of the extent of his writings on pastoral, religious, and theological matters. Chief among his interests were social, political, and scientific questions affecting religious life, the liturgy, and church unity. In 1921, at Malines, he initiated the "conversations" with members of the Anglican Church, which continued at intervals until his death.

World War I broke out during Cardinal Mercier's episcopate, and he became a national and international

leader in resisting German imperialism and in articulating the moral rights of peoples and nations during times of war. His death was the occasion of worldwide tributes to Mercier's immense moral stature and influence as an outstanding philosopher, ecclesiastic, and citizen of the world.

MERCIER'S PHILOSOPHY

An examination of the life of Cardinal Mercier makes it evident that one dimension of his importance for the history of philosophy must be related to his key role in organizing and developing the Institute of Philosophy at Louvain. It becomes equally evident, however, that this dimension cannot be divorced from his originality and depth as a philosopher. Moreover, the significance of Mercier as a philosopher can be fully seen only in the context of the state of philosophy among Roman Catholic thinkers and teachers in Catholic institutions in the latter half of the nineteenth century, on the one hand, and in the light of Mercier's response to and understanding of the papal encyclical *Aeterni Patris,* on the other. Although there were scattered efforts at a renewal of Thomistic thought during this period, philosophy in Catholic circles was by and large eclectic and superficial. Little serious effort had been made to meet either the challenge of Immanuel Kant or the positivism of Auguste Comte and the skepticism of David Hume and the British empiricists. Consequently, Catholic philosophy was generally in serious disrepute.

It is in this setting that the publication of *Aeterni Patris* must be viewed. This encyclical has been misinterpreted by Catholic and non-Catholic thinkers alike as calling for a return to the letter of thirteenth-century thought and as representing ecclesiastical approval, even sanction, of a particular philosophical doctrine. Recent scholarship has amply demonstrated the falsity of both these views and shows Leo XIII's intent to have been a renewal and articulation of a philosophy organically linked to a great philosophical tradition and compatible with Christian faith but rethought in relation to contemporary problems and issues (see J. Collins in *Leo XIII and the Modern World,* edited by Edward T. Gargan, New York, 1961, pp. 181–209).

No one seems to have caught the spirit of this intent or to have grasped the urgency and challenge of the intellectual crisis of the time more accurately than Cardinal Mercier. Perhaps this can best be seen by a brief exposition of Mercier's thought in three crucial areas: the nature of the philosophical endeavor in itself and in its relation to revealed truth and theology, the relation of Thomistic thought to modern philosophy, and the relation of philosophy to the discoveries of modern science.

For Mercier, philosophy is essentially an effort of reason reflecting on the data of experience. Included in this view is a strong affirmation that philosophy must take its point of departure and find its ultimate grounding in the evidence of the real, objective world, in contradistinction to all forms of idealism and theories of innate ideas. The role of reason is likewise strongly emphasized by Mercier, especially in his opposition to positivism. For him, philosophy must be scientific in the classical Aristotelian sense; the mind is capable of going beyond the contingent order of the factually given and of finding real, general necessity and order underlying the sensibly grasped world. Hence, Mercier makes a strenuous effort to reestablish the viability of a realistic metaphysics in the face of the Kantian critique and the severe limitations placed on reason by Comtian positivism. The doctrine of abstraction and the legitimate use of the analytic and synthetic activity of the mind constitute the operative principles in this effort. Nevertheless, philosophy for Mercier is a highly personal endeavor that must always remain open and be capable of organic growth in the light of new evidence. Thus, Thomistic philosophy is held by him as "neither an ideal which one is forbidden to surpass nor a barrier fixing the limits of the activity of the mind"; rather, it is a source of philosophical inspiration that provides a framework for entering into genuine dialogue with the contemporary situation.

Mercier is in fundamental agreement with St. Thomas Aquinas in expressing confidence in the impossibility of real contradiction between revealed doctrine and philosophically established truth. Revealed truth functions for him as an extrinsic negative norm, but it provides neither the motivation for adherence to a philosophical truth nor a source of evidence or knowledge for the philosopher in his proper task. Thus, Mercier emphasizes the essential autonomy, the rigorously rational character, the intrinsic openness, and the need for internal growth of philosophy.

In his writings Mercier is manifestly impatient with the general tendency of his immediate predecessors among Roman Catholic philosophers to opt for one of two general positions—a superficial eclecticism or a dogmatic and naive realism based on common sense. In sharp contrast to these positions, Mercier felt it absolutely essential to examine the whole of modern philosophy with great sympathy and to integrate its sound insights into an integral and rethought Thomism. This principle did not, however, prevent Mercier from being highly crit-

ical of the various contemporary philosophical positions. His polemical writings are directed against fideism, traditionalism (the view that human reason without the aid of revelation necessarily falls into error), voluntarism, sentimentalism, pragmatism, Cartesianism, positivism, and Kantian critical philosophy. He argued strenuously against the Cartesian principle of universal methodic doubt and against Cartesian dualism, undertaking to show that the Thomistic doctrine of the substantial unity of man could overcome the difficulties to which this dualism gives rise.

Positivism and Kantian philosophy, however, occupied most of Mercier's attention, and it was in relation to these views that Mercier developed his own epistemology (in *Critériologie générale,* 1899), which represents one of his most original contributions to the renewal of Thomistic thought. Against the positivist theories of H. A. Taine, John Stuart Mill, Herbert Spencer and Comte, which he undertook to refute in detail, Mercier insistently affirmed the primacy of the criterion of reason and the absolute value of "ideal judgments." Although the positivists of his day were his principal adversaries, Kant was probably the modern philosopher whom he most admired. His understanding of Kant was limited, however, to the interpretation of his times, and his criticism centers on what he considered to be the psychological subjectivism, hence relativism, of Kant. In the final analysis, then, he feels that both Kantian critical philosophy and positivism lead to skepticism and agnosticism. His response was an attempt to establish a realistic metaphysics on the basis of a sophisticated epistemological critique and a development of a theory of certitude. In his own systematic thought, it is not clear that Mercier fully succeeded in formulating what he intended—that is, a middle term between empiricism and rationalism—for his effort begins with a vigorous defense of the absolute certitude of ideal judgments, and from this position he attempts to establish the degree of certitude proper to judgments of experience. In choosing this starting point, Mercier is forced to infer the reality of the external world on the basis of an ideal principle of causality. Nevertheless, it remains a fact that Mercier's epistemology in its attempt to establish a viable, realistic metaphysics represented a major advance in Thomistic thought.

Apart from his epistemology the most original and commanding dimension of Mercier's thought concerned the relation between philosophy and science. In this area he strongly advocates the necessity for philosophy to be intimately acquainted with the findings of modern science. His own efforts in this area were devoted to a synthesis of the new science of psychology and traditional philosophy; the detail with which he undertook to understand the work of such contemporary psychologists as Wilhelm Wundt and the developments in medical psychology were radically new for his time. Although he clearly held that science and philosophy represent two different modes of thought and although he attributed some real autonomy to science, Mercier probably did not fully appreciate the theoretical component of science (this is hardly surprising given the state of the psychological sciences and the philosophy of science in his day). Hence, his synthesis represents an attempt to understand the facts and laws established by science in the light of metaphysical principles. Once again, however partial Mercier's particular solution to this problem may be, it represents a major advance over the earlier tendency of scholastic philosophy to develop in complete isolation from contemporary thought.

Mercier's own philosophical work represents, then, a vigorous and sustained effort to rethink traditional Thomistic thought in the light of contemporary thought on all fronts; moreover, the spirit of this effort was embraced by colleagues whom Mercier chose to staff the Institute of Philosophy. The true philosophical importance of Mercier must be judged by the caliber of philosophical research and writing that has emanated from the Louvain Institute from his day to the present.

See also Cartesianism; Comte, Auguste; Empiricism; History and Historiography of Philosophy; Hume, David; Kant, Immanuel; Mill, John Stuart; Neo-Kantianism; Positivism; Pragmatism; Rationalism; Taine, Hippolyte-Adolphe; Thomas Aquinas, St.; Thomism; Voluntarism; Wundt, Wilhelm.

Bibliography

WORKS BY MERCIER

For a complete bibliography of Mercier's writings, see *Revue néo-scolastique* 28 (1926): 250–258. Mercier wrote extensively for this and other philosophical journals, and much of his polemical writing appears in articles. His major books were written primarily as textbooks and frequently appeared in several mimeographed forms before publication; the published books were revised and frequently reprinted.

The following are his principal works: "La psychologie expérimentale et la philosophie spiritualiste," in *Bulletin de la Classe des Lettres et des Sciences Morales et Politiques et de la Classe des Beaux-Arts* (Brussels, 1900), which was translated by E. J. Wirth as *The Relation of Experimental Psychology to Philosophy* (New York: Benziger, 1902); *Psychologie,* 2 vols. (Louvain and Paris, 1892; 11th ed., 1923); *Logique* (Louvain and Paris, 1894; 7th ed., 1922);

Métaphysique générale ou ontologie (Louvain and Paris, 1894; 7th ed., 1923); *Les origines de la psychologie contemporaine* (Louvain and Paris, 1897; 5th ed., 1922), which was translated by W. H. Mitchell as *Origins of Contemporary Psychology* (New York, 1918); *Critériologie générale* (Louvain and Paris, 1899; 7th ed., 1918).

Mercier collaborated with M. de Wulf and D. Nys in writing *Traité Elémentaire de philosophie,* 2 vols. (Louvain and Paris, 1905; 5th ed., 1920), translated by T. L. Parker and S. A. Parker as *A Manual of Modern Scholastic Philosophy,* 3rd ed., 2 vols. (London: K. Paul, Trench, Trubner, 1928).

STUDIES ON MERCIER

The definitive personal and intellectual biography of Mercier is by L. de Raeymaeker, *Le Cardinal Mercier et l'Institut Supérieur de Philosophie de Louvain* (Louvain: Publications Universitaires de Louvain, 1952), which also contains a detailed account of the founding and history of the institute. The best critical study of Mercier's thought is in G. Van Riet, *L'epistémologie thomiste* (Louvain: Editions de l'Institut SupÈrieur de Philosophie, 1946), pp. 135–178. Also to be noted is L. Noel, "Le psychologue et le logicien," *Revue néoscolastique* 28 (1926): 125–152. Probably the best biography in English is by J. Gade, *The Life of Cardinal Mercier* (New York: Scribners, 1934).

Alden L. Fisher (1967)

MEREOLOGY

"Mereology" (from Greek *meros,* "part") is the theory (often formalized) of part, whole, and cognate concepts. The notion of part is almost ubiquitous in domain of application, and for this reason Edmund Husserl assigned its investigation to formal ontology. Aristotle observed that the term *part* was used in various ways, as for a subquantity, a physical part (leg of an animal), a part in definition (animal is part of man), a part in extension (man is part of animal). Part concepts had obvious applications in geometry and were among Euclid's undefined terms. Several senses of "part" are expressible using the preposition "in," but not all uses of "in" express parthood.

Until the twentieth century it was generally assumed that the concept of part was sufficiently clear not to require elucidation, but gradually the need for a formal treatment became apparent. Euclid's maxim that the whole is greater than the part appeared to be contradicted by infinite classes, for example. In 1901 Husserl proposed a general theory of part and whole and distinguished several kinds of parts, notably dependent and independent parts. Explicit formal theories of part and whole were developed around 1914 to 1916 by Alfred North Whitehead and Stanisław Leśniewski, who worked independently of each other. They had different motivations: Whitehead wanted an empirical basis for geometry, whereas Leśniewski wished to offer a paradox-free class theory. Mereology was later formulated within first-order predicate logic by H. S. Leonard and Nelson Goodman, who called it "the calculus of individuals." Mereology has often been employed by nominalists as a partial substitute for set theory, but it is not intrinsically a nominalistic theory: Part relations are definable via endomorphisms in many mathematical domains.

The most natural basic concept of mereology is that of a (proper) part to its (larger) whole. A coincident of an object is the object itself or something that shares all parts with it. An ingredient of an object is a part or coincident of it. Two objects overlap if and only if they share an ingredient, and they are disjoint if and only if they do not. The relation of part to whole has some minimal formal properties: It is (1) existence entailing; (2) asymmetrical; (3) transitive; and (4) supplementative. That means (1) that if one thing is part of another, if either the part or the whole exists, so does the other; (2) that if one thing is part of another, the second is not part of the first; (3) that a part of a part of a whole is itself a part of the whole; and (4) that if an object has a part, it has another part disjoint from the first. Principles (3) and (4) have occasionally been doubted, (4) unconvincingly. Some meanings of "part" are not transitive; for example, a hand is said to be part of the body, but an arbitrary chunk of flesh is not, and for such concepts counterexamples to (3) may sound plausible, but only because they restrict the general (and transitive) concept, to mean, for example, organ, functional part, immediate part, assembly component.

Beyond such minimal properties mereologists often make further assumptions. Very often it is assumed that objects with the same ingredients are identical: Such a mereology is extensional. Extensionality makes good sense for homogeneous domains such as regions of space or masses of matter, but some objects of distinct sorts seem to be able to coincide, at least temporarily, without identity. Another assumption often made is that any two objects make up a third, indeed that any nonempty collection of objects constitutes a single object, their mereological sum. The minimal properties together with extensionality and this general-sum principle constitute the classical mereology of Leśniewski and Leonard/Goodman: It is as rich in parts as an extensional theory can be, differing algebraically from Boolean algebra only in lacking a null element. It does, however, have an ontologically maximal object or universe, the sum of all there is, which by extensionality is unique. Whitehead denied that there was a universe: For him every object is part of something greater, so he rejected the sum principle. Whitehead also

denied there are atoms, that is, objects without parts: For him, every object has a part. This antiatomism, together with supplementarity, ensures that every object has nondenumerably many parts. Whitehead thus denies geometrical points, and his method of extensive abstraction is directed to logically constructing substitutes for points out of classes of extended objects, an idea also carried through by Alfred Tarski. As the examples indicate, the issue whether atomism or antiatomism holds is independent of general mereology. Formally, the best worked-out forms of mereology are those of Leśniewski and his followers; they have shown that any of a wide range of mereological concepts may be taken as sole primitive of the classical theory.

Beyond extensional mereology attention has focused on the combination of mereological notions with those of space, time, and modality. Thus, Whitehead and a number of more recent authors combine mereological with topological concepts to define such notions as two regions' being connected, or their abutting (externally or internally), using mereology as its modern authors intended, as an alternative framework to set theory. When time is considered, matters become more complex. Some objects have temporal parts, including phases, and perhaps momentary temporal sections. States, processes, and events (occurrents) are uncontroversial cases of objects that are temporally extended, but many modern metaphysicians apply the same analysis to ordinary things such as bodies and organisms, giving them a fourth, temporal dimension, though this view is not uncontested. Whether or not continuants (spatially extended objects with a history but not themselves temporally extended) are thus reduced to occurrents, a number of chronomereological concepts may be defined and applied, such as temporary part, initial part, final part, permanent part, temporary overlapping, growth, diminution, and others, though their formulation will vary as applying to occurrents or continuants.

Embedding mereological notions within a modal framework likewise opens up a wider range of concepts such as essential part, accidental part, dependent part, accidental overlapping. Combining these in their turn with temporal notions allows the definition of concepts such as accidental permanent part, essential initial part, and so on. In general, where mereological notions are enriched with others, their interactions become multifarious and lose the algebraic elegance of the classical theory while gaining in applicability and usefulness.

In modal mereology much attention has been paid to R. M. Chisholm's thesis of mereological essentialism, which states that every part of a continuant is both essential and permanent to that continuant (though, conversely, a part may outlast the whole and need not have it as whole). Chisholm's position is presaged in Gottfried Wilhelm Leibniz and Franz Brentano. Since it appears to be contradicted by everyday experience of such things as rivers, mountains, organisms, and artifacts, it is natural for Chisholm to regard such mereologically fluctuating things as not "real" continuants but as *entia successiva,* supervenient upon successions of continuants for which mereological essentialism holds.

The ubiquity and importance of mereological concepts ensure them a growing place within cognitive science and formal representations of commonsense knowledge, and there is no doubt that mereology is firmly established as a part of formal ontology.

See also Aristotle; Brentano, Franz; Chisholm, Roderick; Cognitive Science; Goodman, Nelson; Husserl, Edmund; Leibniz, Gottfried Wilhelm; Leśniewski, Stanisław; Metaphysics; Tarski, Alfred; Whitehead, Alfred North.

Bibliography

Chisholm, R. M. *Person and Object: A Metaphysical Study.* London: Allen and Unwin, 1976.

Husserl, E. "On the Theory of Wholes and Parts." In *Logical Investigations,* 2 vols. London: Routledge, 1970.

Leonard, H. S., and N. Goodman. "The Calculus of Individuals and Its Uses." *Journal of Symbolic Logic* 5 (1940): 45–55.

Leśniewski, S. "On the Foundations of Mathematics." In *Collected Works.* Dordrecht: Kluwer Academic, 1992.

Simons, P. M. *Parts: A Study in Ontology.* Oxford: Clarendon Press, 1987.

Whitehead, A. N. "Principles of the Method of Extensive Abstraction." In *An Enquiry concerning the Principles of Natural Knowledge.* Cambridge, U.K.: Cambridge University Press, 1919.

Peter Simons (1996)

MERLEAU-PONTY, MAURICE
(1908–1961)

Maurice Merleau-Ponty, a French philosopher associated with existential phenomenology, was the youngest philosopher ever to be appointed to the chair once occupied by Henri Bergson at the Collège de France. Merleau-Ponty was born in Rochefort-sur-Mer on March 14, 1908. His father died early in his childhood; he and his brother

and sister were raised by his mother. He attended the Lycée Louis-le-Grand and then the École Normale Supérieure earning his aggregation in 1930. He taught in lycées and then was mobilized in the Fifth Infantry Regiment, and served as a second lieutenant from 1939 until demobilization in 1940. During the occupation he participated in the Résistance. After the liberation in 1945 he taught at the Université de Lyon; during this time he, together with Jean-Paul Sartre, founded the avant-garde journal, *Les temps modernes*. In was also in 1945 that his major work, the *Phenomenology of Perception* was published.

Merleau-Ponty is known primarily for developing an ontology that recognizes the philosophical significance of the human body and for his success in overcoming the dualism that has plagued European philosophy from its inception, but these endeavors also include significant contributions to post-structuralist linguistics, political theory, developmental psychology, and aesthetics. His early interest in the resonance between the emergent school of gestalt psychology and the phenomenology of Edmund Husserl and Martin Heidegger led to a radical reassessment of transcendental philosophy. He died abruptly on May 3, 1961, at the age of fifty-three, leaving his last major manuscript, *Le visible et l'invisible*, unfinished. Claude Lefort has edited the extant text, four chapters and an appendix, and published it together with extensive working notes dated from January 1959 to March 1961.

THE LIVED BODY

Merleau-Ponty revolutionized European thinking about the body—which since ancient Greece had taken it to be either insignificant or a detriment to knowledge—by demonstrating its constitutive role in the process of human understanding. He showed, for example, that it is through bodily motility that the various adumbrations or perspectival views of an object can be synthesized into a unitary whole. Human understanding of objective space, the three-dimensional Cartesian grid of depth, breadth, and height, is an abstraction from lived space—space articulated by the body's capacity to move purposively, to grasp things, to maintain the equilibrium that allows for stable visual coordinates, and to interrogate its environment. Furthermore, the body's ability to perceive the world is grounded in the body's double role as sensor and sensed, capable of being both subject and object of experience: One could not touch an object were one not oneself, as body—an object capable of being touched; nor could one see were her or his eyes not themselves objects located within the surroundings to which they are sensitive. The classical dualism, which views the body and other worldly objects as disjunct from the mind as the subject or agency of disembodied thought, is replaced with Merleau-Ponty's model of corporeal intentionality in which the body is revealed as having an intelligence of its own, manifest in reflex as in habitual activities, which allows it to interact with the world at a level prior to the reflexivity of deliberate conceptualization.

REVERSIBILITY THESIS

The transcendental role of the body, its ability to project its organizational schemas into the world, is inseparable from the body's own status as physical object subject to the worldly forces impinging upon it. These roles are inseparable, but not coincident. There is a divergence of the body as sensing from the body as sensed: The finger that touches the thumb or is touched by it does not form an identity with the thumb; rather the two bodily parts coexist in an ambiguous relationship of reversibility within the encompassing matrix of bodily being-in-the-world. Finger and thumb can reverse roles, the erstwhile sensor becoming the sensed, just as the hand that feels the table can sense itself being touched by the table. Yet neither of these roles would be possible were it not for the other.

THE FLESH OF THE WORLD

Merleau-Ponty takes the reversibility of subject and object roles in the case of human flesh as emblematic of a global manner of being which he designates as chiasm or intertwining. The term *flesh* is generalized to encompass worldly being as such. The world is taken as an arena of interaction in which every entity is what it is in relation to every other. This is not a pan-animism, but rather an attempt to rectify the post-Socratic reduction of nature to inert materiality in a movement of thought which is as consonant with the ancient concept of *physis* as it is with the contemporary notion of world as ecosystem. The figure of the chiasm, the intersection marking the point at which things touch each other as they cross, refers to the dynamics of worldly unfolding or global temporality in which the interaction of things brings about change. The brute or savage being of the world, the factuality of its transcendence, is counterbalanced with the relatedness of its denizens apparent in the relatively abiding structures human intelligence organizes under the heading of science. Humans are that aspect of the flesh of the world that is capable of the reflective relationship of conceptualization or understanding, but other aspects of the world betray other forms of corporeal reflexivity in the complex

of interaction that encompasses organic cycles, weather systems, geological formations, and so forth as each of these contributes and responds to all the others.

VISIBLE AND INVISIBLE

Merleau-Ponty's thesis of the primacy of perception evolves from the middle phase of his thinking when he published the *Phenomenology of Perception* and set forth the view that "the perceived world is the always presupposed foundation of all rationality, all value and all existence" to later phases in which this thesis had to be expanded to accommodate the findings of extensive analyses of language based on his unique interpretation of the philosophical significance of Saussurean semiotics. There is controversy regarding his later thinking on the relative primacy of language and perception, but general agreement that the relationship between the two is that of intertwining: language, conceived as sign system, may be conceived as an invisible nexus of relations that is apparent in the visible world and is itself perceptible in speech and writing. The controversy centers on two questions regarding origins or foundations. Does the invisible structure of language reflect organization perceived in the world or does it constitute that nexus of relations? The second question challenges the legitimacy of asking the first: Is it possible to separate perception from language in such a way that one could even ask about the primacy of one with respect to the other?

Merleau-Ponty regards language as flesh, akin to the flesh of the body in its reflexivity—its relatedness to itself and world—but "less heavy, more transparent." In general, the structure of the visible-invisible relation can be defined as *asymmetrical reversibility*: Just as the object one touches can be seen although its tactile aspect remains invisible as such, so can the hidden or horizonal aspects of a given theme be brought into focal vision but only through the loss of its horizonality.

POLITICS

Merleau-Ponty's thinking in general is dynamic and emergent; it is unified by an elusive paradigm he would never have captured even if he lived longer than he did. Nowhere is this questing more apparent than in his political thought. He was always a critical reader of Marx—although he refrained from revisionism as long as he could—and was highly suspicious of the Communist revolution, although he initially endorsed its humanist goals. When Merleau-Ponty died at the height of his powers, he was working toward what may be called an ethics of expression and reversibility, and the direction of this thought can be seen articulating itself as early as his chapter on "Freedom" in the *Phenomenology of Perception*.

The issue that dominated left-wing politics in France—indeed, Europe at large and the USSR—had to do with the tension between party leadership and domination, on the one hand, and the emergence of an increasingly self-conscious proletariat anxious to take up the reins of history, on the other. Was the role of the Central Committee to take charge? Or to take its bidding from the workers of the world? Was the dialectical movement of history objectively determined by materiality? Or subjectively articulated in contests at the level of ideality?

Merleau-Ponty refused to take sides, but sought to undercut the polarity and find a means to embrace the truths to be found on both ends of the spectrum. "The world," he writes, "is already constituted, but also never completely constituted; in the first case we are acted upon, in the second we are open to an infinite number of possibilities. But this analysis is still abstract, for we exist in both ways *at once*. There is, therefore, never determinism and never absolute choice, I am never a thing and never bare consciousness. … It is impossible to determine precisely the 'share contributed by the situation' and the 'share contributed by freedom'" (p. 453). In short, it is through the expression of his situation on the part of the individual worker and his recognition of others in the same plight that solidarity is formed and action can be undertaken. The worker can benefit from guidance from above, but the task of gaining freedom and overcoming the forces that resist it cannot be displaced on to others, else the worker is reduced to slavery again, this time at the hands of his or her liberators.

This idea of circumscribed freedom was in direct opposition to the thesis of radical freedom then espoused by Merleau-Ponty's colleague and cofounder of *Les temps modernes*, Jean-Paul Sartre. This conflict at the level of ideas came to a head in the early 1950s with the disclosure of the atrocities being committed by Stalin in Russia. How to respond? Sartre maintained solidarity with the Communist Party; Merleau-Ponty distanced himself from both, and resigned from the editorial staff of the journal in 1953. The political writings in *Sense and Non-Sense* (1964 [1948]) were written before this break, and the critical reflections on Marxism (including a chapter on "Sartre and Ultrabolshevism") titled *Adventures of the Dialectic* was published in 1955. In the later *Humanism and Terror* (1969 [1947]), Merleau-Ponty sought to put the dialectical thinking of Hegel and Marx in historical perspective, transcend it, and point in a new direction. His conclusion constitutes another step in the direction

of the ethics of expression and reversibility mentioned above. "To seek harmony with ourselves and others, in a word, truth, not only in … solitary thought but through the experience of concrete situations and in a living dialogue with others apart from which internal evidence cannot validate its universal right, is the exact contrary of irrationalism, since it accepts our incoherence and conflict with others as constants but assumes we are able to minimize them. It rules out the inevitability of reason and well as that of chaos" (1969 [1947], p. 187).

In his last and unfinished work, *The Visible and the Invisible* (1968 [1964]), Merleau-Ponty returns to the subject of dialectical thought, espouses the thought that ideality and materiality intertwine in a movement of history that can move in the direction of minimizing conflict, but explicitly repudiates the formalism that informs the work of Hegel, Marx, and Sartre in a misguided attempt to impose an abstract structure on the unpredictable and messy historical process in which situated human freedoms collide and intertwine. It is also in this work that he begins to articulate the notion of reversibility, his own response to the Husserlian doctrine of foundation (*Fundierung*).

PSYCHOLOGY

From the earliest of his writing until the last, Merleau-Ponty maintained the thesis of the irreducibility of the figure-ground or theme-horizon structure articulated by gestalt theory. This thesis holds that perception and cognition are fundamentally relational, hence stand in opposition to such standpoints as that of sense-data theory based on the notions of perceptual atoms, elemental simples, or discrete qualia.

In the *Phenomenology of Perception*, Merleau-Ponty offers an extended case study of Schneider, a World War I soldier debilitated by a shrapnel wound in the occipital region of his brain. The point of the study is to demonstrate the inadequacy of the standpoints of empiricism or physicalism, on the one hand, and intellectualism or transcendentalism, on the other, to provide an accurate description of Schneider's afflictions, which are neither purely physiological nor purely intentional but involve a degeneration of the lived body resulting in aberrant forms of substitution behavior in such domains as sexual responsiveness, existential spatiality, motility, expression, and memory.

Merleau-Ponty is unique among phenomenologists in reinterpreting Freudian notions regarding the unconscious in a positive way and integrating them within his own body of theory. This appropriation involved some modification, to be sure, specifically that of asserting a continuity between conscious and unconscious aspects of human experience at the level of prereflective horizonality. Merleau-Ponty steers a middle course between Freud's relatively mechanistic account of such phenomena as repression, which attributes it to an autonomous function of censorship and dissemblance, and Jean-Paul Sartre's relatively voluntaristic account, which attributes repression to an act of self-deception on the part of a consciousness recoiling from the implications of its own freedom. Merleau-Ponty interprets behavior traditionally subsumed under the heading of repression in terms of a process of habituation operating at prepersonal or unreflective levels in which the body's response to worldly events becomes sedimented as a style of contending with a domain of existence permeated with negative significance. Thus, the aphonia and anorexia of a girl whose family has forbidden her to see her lover is understood, neither as a reversion to an infantile phase of oral sexuality, as Freud would have it, nor as a recoil from responsibility in the mode of magical transformation, as Sartre would have it, but as a refusal of coexistence, a withdrawal from the communal world of eating and talking, which acquires the autonomy of a habit exacerbated by former habitualities favoring oral modes of responding to the world.

In addition to his interests in gestalt psychology and Freudian psychoanalysis, Merleau-Ponty was also well-acquainted with the work done by his sometime colleague Jean Piaget in developmental psychology and the work of Jacques Lacan, a contemporary known for his reinterpretation of Freudian themes along semiological lines. There are frequent references to Piaget in *The Structure of Behavior* (1963 [1942]) and the *Phenomenology of Perception*, and an extended response to Lacan's seminal thinking on the mirror stage in a late essay titled "The Child's Relations with Others." Perhaps Merleau-Ponty's greatest contribution to psychological theory lies in his articulation of an ontological framework capable of consolidating the findings of thinkers across the full spectrum of ideologies from eidetic analysis to experimental and behavioral research: He unremittingly refused to endorse the radical distinctions between the a priori and the a posteriori, between transcendental and empirical approaches, which have functioned to isolate the various schools through polarized opposition.

AESTHETICS

Merleau-Ponty revivifies the ancient Greek sense of the term *aesthetics* by focusing on the perceptual foundations

of art rather than concerning himself with judgments of taste. In accordance with his thesis of the primacy of perception, he regards the artist as one who seeks to respond to the world as it manifests itself perceptually rather than to superimpose preconceived conceptual structures upon the world. For example, classical Renaissance painting attempts to render depth on a two-dimensional surface by applying the laws of perspective. Such laws reduce depth to a mere rotation of breadth, seeing it from the side, and overlook the existential or lived aspect of depth as the dimension of exploration and mystery. In classical painting the eye of the artist is fixed and static, whereas in perception the artist's body is spatially mobile and not delimited to an instant of time.

Cézanne, Merleau-Ponty's favorite exemplar, renders depth in his paintings of Mont St. Victoire by using broad, blurred strokes in the foreground, clearer ones in the mid-ground, and an ethereal mistiness in the distance. In his still life paintings, table tops, vases, carafes of wine, and the like are portrayed as a moving eye would see them, not as a photograph would array them from a single point. The painting of galloping horses titled *Derby at Epsom* by Theodore Géricault shows the quadrupeds with their legs extended forward and backward, a distortion of the actual positions of legs in equine movement that succeeds in imparting motion to the animals rather than suspending them awkwardly in the air as a fixed frame, instantaneous representation would. The distortion is actually truer to what people perceive in the extended duration of the lived moment.

Artists have the ability to see what theoretical presuppositions lead people to overlook, and this allows them to bring the invisible to visibility, hence to bring the painting to life. Artists paint what they see rather than what they know of an object. Renoir visually interrogates the water he sees in the Mediterranean sea at Cassis to enable him to paint *The Bathers* in a pool in a sylvan setting. He sees the play of light through the fluid surfaces of the dynamic element that is invisible to the eye of the observer who can only see what he or she thinks is actually there. Artists train themselves to see the speck of light on the glistening surface of eyes that are, themselves, seeing. It is the invisibility of that speck of light to Fra Lippo Lippi, for example, who paints the eye as he thinks it truly is anatomically, that makes the persons in his portraits appear moribund.

The reversibility of seer and seen crosses as in a chiasm with the reversibility of the invisible and the visible. Artists attuned to their own visibility can paint their subjects seeing them and thereby depict the subjectivity of the subject that remains invisible to those who think that in perceiving others people see only their material bodies. Perception, however, is—or can be—truer to living bodies than Cartesian philosophy that reduces human flesh to *res extensa* and conceives *res cogitans* as invisible.

See also Aesthetic Experience; Aesthetics, History of; A Priori and A Posteriori; Art, Representation in; Bergson, Henri; Cartesianism; Dialectical Materialism; Empiricism; Freedom; Freud, Sigmund; Gestalt Theory; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Humanism; Husserl, Edmund; Lacan, Jacques; Marxist Philosophy; Marx, Karl; Nomos and Phusis; Perception; Perception, Contemporary Views; Phenomenology; Physicalism; Piaget, Jean; Qualia; Rationality; Sartre, Jean-Paul; Unconscious.

Bibliography

PRINCIPAL WORKS BY MERLEAU-PONTY

Humanisme et terreur, Essai sur le problème communiste. Paris: Gallimard, 1947. Translated by John O'Neill as *Humanism and Terror: An Essay on the Communist Problem* (Boston: Beacon Press, 1969).

"La conscience et l'acquisition du langage." *Bulletin de psychologie* 236 (18) (1964): 3–6. Translated by Hugh J. Silverman as *Consciousness and the Acquisition of Language* (Evanston, IL: Northwestern University Press, 1973).

"La Nature": Notes, cours du collège de France. Compiled and with notes by Dominique Séglard. Paris: Editions du Seuil, 1995. Translated by Robert Vallier as *"Nature": Course Notes from the Collège de France* (Evanston, IL: Northwestern University Press, 2003).

La prose du monde. Paris: Gallimard, 1969. Translated by John O'Neill as *The Prose of the World* (Evanston, IL: Northwestern University Press, 1973).

La structure du comportement. Paris: Presses universitaires de France, 1942. Translated by Alden L. Fisher as *The Structure of Behavior* (Boston: Beacon Press, 1963).

Le visible et l'invisible. Edited by Claude Lefort. Paris: Gallimard, 1964. Translated by Alphonso Lingis as *The Visible and the Invisible* (Evanston, IL: Northwestern University Press, 1968).

Les aventures de la dialectique. Paris: Gallimard, 1955. Translated by Joseph Bien as *Adventures of the Dialectic* (Evanston, IL: Northwestern University Press, 1973).

Phénoménologie de la perception. Paris: Gallimard, 1945. Translated by Colin Smith as *Phenomenology of Perception* (London; New York: Routledge, 1962).

Résumés de cours, Collège de France, 1952–1960. Paris: Gallimard, 1968. Translated by John O'Neill as *Themes from the Lectures at the Collège de France, 1952–1960* (Evanston, IL: Northwestern University Press, 1970).

Sens et non-sens. Paris: Gallimard, 1948. Translated by Hubert L. Dreyfus and Patricia Allen Dreyfus as *Sense and Non-Sense* (Evanston, IL: Northwestern University Press, 1964).

Signes. Paris: Gallimard, 1960. Translated by Richard C. McCleary as *Signs* (Evanston, IL: Northwestern University Press, 1964).

WORKS ON MERLEAU-PONTY IN ENGLISH

Dillon, M. C. *Merleau-Ponty's Ontology*. 2nd ed. Evanston, IL: Northwestern University Press, 1997.

Johnson, Galen. *The Merleau-Ponty Aesthetics Reader*. Evanston, IL: Northwestern University Press, 1993.

Madison, G. B. *The Phenomenology of Merleau-Ponty*. Athens: Ohio University Press, 1981.

Silverman, Hugh J., and James Barry, eds. *Texts and Dialogues*. Atlantic Highlands, NJ: Humanities Press, 1992.

M. C. Dillon (2005)

MERSENNE, MARIN

(1588–1648)

Marin Mersenne, a French mathematician, philosopher, and scientist, was one of the most influential figures of the scientific and philosophical revolutions of the seventeenth century. Although he is remembered primarily for his relationship with René Descartes, he was a significant figure in his own right and also, through his immense correspondence, publications, and personal acquaintances, a key figure in coordinating and advancing the work of the new philosophers and scientists.

He was born at Oizé, France, and studied at Le Mans and later at the Jesuit college of La Flèche, from 1604 to 1609. (Descartes, eight years his junior, was there from 1604 to 1612, but their friendship began later, around 1623.) He next studied in Paris and then entered the pious and austere order of the Minims. After further theological studies Mersenne taught philosophy at a convent in Nevers until 1619, when he was sent back to Paris by his order. He remained there until his death in 1648, except for some trips to The Netherlands, Italy, and the French provinces. His Parisian monastic cell was the center of the European scientific world, as scholars, scientists, philosophers, and theologians often made their way to Mersenne's quarters.

MERSENNE'S PUBLICATIONS

From 1623 to 1625 Mersenne published several enormous polemical works attacking all sorts of Renaissance outlooks and figures, ranging from atheists, deists, kabbalists, astrologers, and numerologists to Pyrrhonists. These writings include the *Questiones Celeberrimae in Genesim* (1623), *L'impiété des deists, athées et libertins de ce temps, combatuë, et renversée* (1624), and *La vérité des sciences contre les septiques ou Pyrrhoniens* (1625). The last work, more than a thousand pages long, was the culmination of this phase of Mersenne's career and the beginning of the scientific phase that was to continue until his death. Thereafter, his writings were on all sorts of scientific and mathematical subjects (including the famous *Harmonie universelle* [1636–1637] on the theory of music, harmonics, and acoustics) and were compendiums of the knowledge in these areas.

Mersenne became involved in the publication of fundamental works of his friends or correspondents, such as Galileo Galilei's *Mechanics* (translated by Mersenne), the objections to Descartes's *Meditations* (gathered by Mersenne), Herbert of Cherbury's *De Veritate* (in a translation by Mersenne), Thomas Hobbes's *De Cive* (the publication of which was arranged by Mersenne), and François de La Mothe Le Vayer's *Discours sceptique sur la musique* (published in Mersenne's *Questions harmoniques*). He also carried on a monumental correspondence that provides a magnificent running record of the intellectual revolution of the time. Mersenne was actively interested in an enormous range of scientific and pseudoscientific questions, from the most complex ones in physics, mathematics, and Hebrew philology to such ones as "How high was Jacob's ladder?" and "Why do wise men earn less money than fools?"

His major philosophical contributions were his massive refutation of skepticism, *La vérité des sciences*, and his later discussions of the nature of scientific knowledge. *La vérité des sciences* is a dialogue between a skeptic, an alchemist, and a Christian philosopher (Mersenne). The skeptic uses his arguments to show that alchemy is not a true science. When he broadens his attack to encompass all claims to knowledge of the real nature of things, Mersenne's Christian philosopher offers his own resolution to the skeptical crisis, starting with a detailed examination of Sextus Empiricus's *Outlines of Pyrrhonism*. He repeatedly contends that although the Pyrrhonian arguments may show that one cannot know the real nature of things, one can gain knowledge of the apparent, phenomenal world in terms of how it seems to one and how the various appearances are related. Although one's sense experiences vary and although one cannot tell what objects are really like, one can find laws that enable one to connect and, thus, to predict experiences. Although one cannot find any absolutely certain first principles, one can discover enough indubitable ones to enable one to construct systematic information about one's experienced world. "This limited knowledge suffices to serve us as the guide for our actions." One is able to know something—namely, the sciences of phenomena—and this has ade-

quate pragmatic value for one in this life. Francis Bacon was trying to find out too much and was raising too many insoluble skeptical problems with his Idols. Instead, the ultimate answer to skepticism was to show how much one could and did, in fact, know. The last 800 pages of the work is a listing of what is known in mathematics and mathematical physics—until the Pyrrhonist gives in. He has been conquered not by being refuted but by being shown what sort of knowledge one can have once one grants that knowledge about reality is unattainable.

"CONSTRUCTIVE OR MITIGATED SKEPTICISM"

Mersenne was willing to accept the skeptic's claims but was unwilling to see them establish that nothing can be known. Instead, he saw an epistemological skepticism as the prelude to a "constructive or mitigated skepticism" which would allow a scientific and systematic development of the truths of the sciences of the empirical world. The rest of Mersenne's life was devoted to his religious duty, exploring in phenomenalistic terms what could be known about the world God had made. Mersenne's immense contribution to the scientific revolution was the result of his positive views. Although he had originally portrayed skepticism as one of the greatest menaces to humankind, he continued to insist in his scientific tracts that one can gain no certain knowledge about reality but can study only the surfaces of things as they appear to one and employ mathematics as a hypothetical system about things. Like his close friend Pierre Gassendi (in whose arms he died), Mersenne saw scientific endeavors as a via media between complete skepticism and dogmatism. Mersenne tended to emphasize the antiskeptical aspect of this view, whereas Gassendi tended to emphasize the antidogmatic one.

In his formulations of the new science Mersenne was probably the first to use a mechanical model to account for the world that one experiences and to develop a thoroughgoing phenomenalism (although hardly as well worked out as Gassendi's) adequate to state the findings and assumptions of modern science. Mersenne's lifelong devotion to science and scientists can apparently be attributed to their common quest for more information and understanding of the phenomenal world. Hence, Mersenne could see in Descartes a major contributor to the scientific revolution but could see nothing important in his metaphysical revolution. Descartes, Hobbes, Herbert of Cherbury, Gassendi, Blaise Pascal, Galileo, and others were, for Mersenne, together in seeking the truth of the sciences, although some of them still had illusions that more truth than that could be discovered. For Mersenne, science had no metaphysical foundations and needed none. "Until it pleases God to deliver us from this misery," one can find no ultimate knowledge, but one can, if one is not destructively skeptical, proceed to gain and use scientific knowledge.

See also Bacon, Francis; Descartes, René; Galileo Galilei; Gassendi, Pierre; Herbert of Cherbury; Hobbes, Thomas; La Mothe Le Vayer, François de; Pascal, Blaise; Scientific Revolutions; Sextus Empiricus; Skepticism, History of.

Bibliography

WORKS BY MERSENNE

Correspondance du P. Marin Mersenne. 17 vols., edited by Paul Tannery, Cornélis de Waard, and René Pintard. Paris: G. Beauchesne, 1933–1988.

L'impiété des déistes, athées et libertins de ce temps. Stuttgart, Germany: Frommann, 1975.

Questions inouyes, Questions harmoniques, Questions théologiques, Les méchaniques de Galilée, Les préludes de l'harmonie universelle. Paris: Fayard, 1985.

La vérité des sciences contre les sceptiques ou pyrrhoniens, edited and annotated by Dominique Descotes. Paris: Champion, 2003.

Traité de l'harmonie universelle, edited by Claudio Buccolini. Paris: Fayard, 2003.

WORKS ABOUT MERSENNE

Chappell, Vere. *Grotius to Gassendi*. New York: Garland, 1992.

Dear, Peter. *Mersenne and the Learning of the Schools*. Ithaca, NY: Cornell University Press, 1988.

Lenoble, Robert. *Mersenne, ou, La naissance du mécanisme*. Paris: J. Vrin, 1943.

Popkin, Richard H. *History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Sorell, Tom. *The Rise of Modern Philosophy: The Tension between the New and Traditional Philosophies from Machiavelli to Leibniz*. Oxford, U.K.: Clarendon Press, 1993.

Vickers, Brian. *Occult and Scientific Mentalities in the Renaissance*. New York: Cambridge University Press, 1984.

Richard Popkin (1967, 2005)

MESLIER, JEAN

(1664–1729)

Jean Meslier, perhaps the least restrained freethinker of the French Enlightenment, is also one of the most notorious examples of apostasy. As curé of the village of Etrépigny in Champagne from 1689 to his death, Meslier lived in complete obscurity, attending to his pastoral duties. But under the innocuous exterior of the humble

Catholic priest, there seethed a violent hatred and passionate disavowal of the religion that it was his ironic profession to serve. Having resolved sometime in the 1720s to compose his only work, the *Testament,* with the aim of keeping it secret until his death, he felt free to vent fully the anti-Christian, atheistic, revolutionary—indeed, anarchistic—sentiments that he had been obliged to suppress beneath a lifelong mask of prudent duplicity. The available biographical facts are unfortunately too meager to clarify this extraordinary personality. It is known, however, that on one occasion Meslier's abhorrence of injustice and persecution brought him into bitter conflict with the local nobility and, indirectly, almost into rebellion against the archbishop of Rheims, who, siding (as might be expected) with feudal privilege in the dispute, had castigated the morally outraged but powerless curate.

EDITIONS OF THE *TESTAMENT*

The three autograph originals of the *Testament* addressed by its author to posterity were succeeded, in eighteenth-century France, by a profusion of manuscript copies that circulated briskly in the philosophical underworld of forbidden literature. The prolixity and other stylistic shortcomings of the work resulted, however, in its being edited in the form of various abridgments that proved more suitable for dissemination. The most important of these summaries was, without question, the *Extrait des sentiments de Jean Meslier,* prepared by Voltaire and published in 1762. This first printed version of the apostate priest's opinions was often reprinted, especially under the rubric of Baron d'Holbach's *Le bon sens du curé Meslier*—a combination of one of his own atheistic tracts and of the *Extrait*—which saw many editions well into the nineteenth century. The integral text of the *Testament* was not published until 1864.

THOUGHT

Meslier's entire critique follows from the assumption that religion is basically a political means whereby those in power consolidate their control over the vastly greater number of weak and poor members of society. All religious dogmas, beliefs, and rituals, supposedly devised by the ruling class as instruments of government, are considered to be nothing but errors and superstitions serving to dupe and paralyze the victims of tyranny, holding them in ignorant fear and keeping them from any effective action to alleviate their misery by overthrowing their oppressors.

Meslier thought primarily in terms of economic exploitation, asserting that the opulence and power of the few are, thanks to the protection of civil and religious laws, acquired and maintained at the expense of the near destitution of the people. There is little doubt that, in adopting this general view, he was motivated by deep feelings of sympathy for the sufferings of the poor, with whom he came into daily contact. His condemnation of Christianity therefore had at its root the eminently Christian virtue of pity for the downtrodden and helpless, joined, however, to a fiercely un-Christian zeal to right secular wrongs.

Although Meslier condemned all religions, he attacked Christianity in particular. The bulk of the *Testament* is devoted to fastidious refutations of the many different types of argument by which the "truth" of Christian revelation was presumed demonstrable. Meslier examines and rejects, in turn, the validity of faith, the historicity of miracles, the authenticity of Scripture, the authority of tradition, the accuracy of biblical prophecies, the testimony of martyrdom, the morality of eternal rewards and punishments, and the meaningfulness of such dogmas as the Trinity, the Incarnation, and transubstantiation. The *Testament* is, indeed, a compendium of the historical, exegetical, textual, and logical objections concerning the essentials of the Christian creed discussed in the critical and apologetic literature from the time of Pierre Bayle through the early decades of the eighteenth century. Meslier was conversant with this literature, and although there is relatively little in his criticism that is entirely new with him, the forcefulness, breadth, and intransigence of his "case against Christianity," together with its politicoeconomic basis, give his work a unique character.

Moreover, Meslier did not stop at exposing the fallacies of Christian belief and the social abuses of institutional religion but boldly pursued his train of thought to the affirmation of a materialistic system in which all phenomena can be traced to a physical basis and are subject to the laws of mechanics. He advocated atheism as the only outlook consistent with the interests of the majority of humankind in its struggle against the lust for domination of the unscrupulous few. Among the sources of the *Testament,* special importance should be given to Michel Eyquem de Montaigne's skeptical treatment of time-honored social practices, to the philosophy of Benedict de Spinoza, and to the Epicurean-Cartesian vision of a mechanistic, naturalistic universe in which the supernatural—particularly the doctrines of divine creation and spiritual immortality—no longer found any place.

INFLUENCE

The impact of Meslier's ideas still has to be studied carefully. During the eighteenth century it was merely his negation of Christianity that proved appealing, and his socioeconomic protest, with its overtones of popular revolution, went largely unheeded. Contrary to the philosophes' estimate of Meslier as compatible with middle-class *bon sens,* some Marxists have been able to see in him an audacious spokesman for the economically repressed class of peasants and urban workers and the advocate of socialistic and egalitarian reform of society. But even if this was the true spirit of Meslier's thought, it did not play its intended role, for his influence was largely assimilated into the mainstream of Enlightenment ideology, with its predominantly bourgeois, liberal, and deistic polemic directed at Christianity. Seen in retrospect, the principal weakness of Meslier's anti-Christian *summa* is his oversimplification of the extreme psychological and cultural complexity of the religious phenomenon and its social applications. Moreover, his ardent wish forever to abolish injustice and wretchedness from the world by the expedient (in his own words) of "hanging and strangling with the bowels of the priests all the nobles and rulers of the earth" was no less utopian than fanatical. Nevertheless, Meslier's indignant and savage denunciation of religion was meaningful at the historical moment that inspired and shaped it, when the Roman Catholic Church of France, owing to its official status and immense riches, actually had a vested interest in the perpetuation of political and economic institutions related to the feudal oppression and exploitation of the people.

See also Bayle, Pierre; Cartesianism; Clandestine Philosophical Literature in France; Enlightenment; Epicureanism and the Epicurean School; Holbach, Paul-Henri Thiry, Baron d'; Montaigne, Michel Eyquem de; Religion and Politics; Spinoza, Benedict (Baruch) de.

Bibliography

Marchal, Jean. *L'étrange figure du curé Meslier.* Charleville, France, 1957.

Morehouse, Andrew. *Voltaire and Jean Meslier.* New Haven, CT: Yale University Press, 1936.

Petitfils, E. *Un socialiste-révolutionnaire au commencement du XVIIIe siècle, Jean Meslier.* Paris, 1908.

Porchnev, B. F. *Jean Meslier, et les sources populaires de ses idées.* Moscow, 1955.

Spink, J. S. *French Free-Thought from Gassendi to Voltaire.* London: University of London, Athlone Press, 1960.

Le testament de Jean Meslier. 3 vols, edited by Rudolf Charles. Amsterdam, 1864.

Wade, I. O. *The Clandestine Organization and Diffusion of Philosophical Ideas in France from 1700 to 1750.* Princeton, NJ: Princeton University Press, 1938.

Aram Vartanian (1967)

METAETHICS

Judgments to the effect that certain things (or certain classes of things) are good or bad, right or wrong, or just or unjust, are first-order ethical judgments. Metaethics addresses second-order questions about the meaning and status of moral judgments, for example, "What does it mean to say that something is good or bad, or right or wrong?", "Are moral judgments statements that purport to be true or false?", and "In what sense, if any, can moral judgments be true or false (or correct or incorrect)?" Metaethical questions have been discussed throughout the history of philosophy, but systematic work on metaethics began early in the twentieth century with the publication of G. E. Moore's *Principia Ethica* (1993).

The first half of this entry discusses theories about the meaning of moral judgments, specifically, Moore's theory, the Franz Brentano–A. C. Ewing (1899–1973) theory, emotivism, Richard Hare's prescriptivism, Philippa Foot's theory normative relativism, and Allan Gibbard's (1942–) expressivism. The second half addresses the question of whether moral judgments are objectively true or false; it explains and assesses (some of) the main arguments for and against the view that moral judgments are objectively true or false. Questions about the truth of moral judgments are distinct from questions about moral knowledge. If moral judgments are not true or correct, then there is no such thing as moral knowledge. Moral knowledge is possible only on the assumption that there is something *to know* (i.e., moral truths). However, the view that there are objective moral truths does not imply that we have knowledge of them; it is compatible with moral skepticism, the view that there are objective moral truths but we cannot know what they are.

I. THEORIES OF MEANING

Twentieth-century work on metaethics begins with questions about the meaning of moral judgments.

1. MOORE'S OPEN QUESTION ARGUMENT. In *Principia Ethica* Moore claims that the concepts of intrinsic goodness and badness are the most fundamental moral concepts. He says that the concepts of right and wrong can be defined in terms of "good" and "bad." Moore

writes: "To assert that a certain line of conduct is, at a given time, absolutely right or obligatory, is obviously to assert that more good or less evil will exist in the world, if it be adopted than if anything else be done instead" (Moore 1993, p. 77). Moore argues that goodness (the property denoted by the word "good") is indefinable. By this he means that the property of goodness cannot be analyzed into constituent properties or elements. Goodness is a simple, ultimate property like the property of being yellow as opposed to a complex property such as the property of being a horse. Being a horse is analyzable in terms of other constitutive properties such as having a head, a heart, four legs, four hooves, and so on. Moore defends the claim that goodness is a simple, unanalyzable property with his "open question argument." This argument can be summarized as follows: Consider any definition of good or goodness according to which goodness is identical with a complex property (*P*). It will always make sense to ask if *P* is good (it is an open question whether *P* is good since it is not self-contradictory to deny that *P* is good). However, it makes no sense to ask whether *P* is *P* (this is not an "open question"—it is self-contradictory to deny that *P* is *P*). Therefore, goodness cannot be identical with *P*.

Moore considers several specific definitions of good, which he subjects to the open question argument. He considers the view that good means pleasure, the view that good means what we desire, and the view that good means what we desire to desire. Clearly the question "Is pleasure good?" is not equivalent to the question "Is pleasure pleasant?" The statement "That which we desire to desire is good" is not equivalent to the statement "That which is good is good." Moore's open question test seems to work very nicely for these and many other definitions of good, but he gives no reason to think that *every possible definition of goodness* fails his open question test. Another serious problem with Moore's argument is that he assumes that goodness cannot be identical with a property *P* unless the statement that *P* is good is analytic or true by definition. It is analytic (true by definition) that pleasure is pleasant, but it is not analytic that pleasure is good. Many contemporary philosophers contend that this assumption has been refuted. According to Hilary Putnam (1975) and Saul Kripke (1972), certain natural properties are identical with each other even though statements to the effect that the properties are identical are not analytic (not true by definition). For example, water is H_2O (the property of being water is identical with the property of being H_2O) even though the statement that water is H_2O is not true by definition.

2. **MOORE'S POSITIVE VIEWS**. Moore is a *cognitivist* (i.e., he holds that moral judgments are statements that ascribe properties to things). By contrast, *noncognitivism* is the view that moral judgments are not statements that ascribe properties to things. Moore claims that goodness is a simple, unanalyzable property. Goodness is not a natural property like redness that can be perceived or apprehended through the five senses. Nonetheless, Moore claims that we can have direct intuitive knowledge of this property. This view is problematic. It is open to debate whether any such property exists. The quality of goodness that Moore posits is elusive; it is difficult to know what he is referring to. Many people, on careful introspection, report that they do not intuit any such property. Another serious problem for Moore's view is that it seems to be unable to account for the fact that moral judgments give or purport to give us reasons to act in certain ways as opposed to others. To say that something is intrinsically good implies that we have reasons to choose or prefer it. But it is unclear why this should be so if Moore's theory is true. It is not clear why we should care whether or not our actions produce or fail to produce instances of the nonnatural properties that Moore postulates and claims are identical to the properties of goodness and badness. (Even if they exist, it is not clear that these nonnatural properties are "reason providing" in the way that something must be in order to be the property of goodness or badness.)

Moore's open question argument is one of the most influential arguments in the history of philosophy. Noncognitivist ethical theories, such as emotivism and prescriptivism, arose in a context of philosophical debate in which it was widely assumed that: (a) Moore has shown that goodness is indefinable, and (b) Moore's own positive view is untenable—"good" does not refer to a simple nonnatural property that we directly intuit. Some philosophers concluded that moral terms do not refer to any properties at all and that moral judgments are not statements that ascribe properties to things.

3. **BRENTANO AND EWING**. Brentano (1969) and Ewing (1947) agree with Moore that moral terms refer to "nonnatural" properties, but they give a very different description of the nature of those properties. They hold that the most fundamental moral properties are nonnatural *relational* properties of "fittingness" or "appropriateness" that hold between objects/ properties and attitudes toward them. Ewing holds that the relation of fittingness is unanalyzable and that our apprehension of it is self-evident. In *The Definition of Good* he says: "Certain characteristics are such that the fitting response to whatever

possesses them is a proattitude, and that is all there is to it" (Ewing 1947, p. 172). According to Brentano, to be good is to be an appropriate (fitting) object of love, and to say that one thing is better than another is to say that it is correct to prefer it to the other. Ewing holds that to say that something is good means that it ought to be the object of a favorable attitude. This theory arguably avoids the second objection to Moore's theory noted above. According to Brentano and Ewing, it is part of the meaning of moral judgments (e.g., judgments to the effect that something is good or bad, or right or wrong) that they claim to give us reasons to have favorable or unfavorable attitudes about certain things and reasons to choose certain things over others.

4. **EMOTIVISM**. Emotivism is the view that moral judgments are *expressions* of attitudes rather than statements that ascribe properties to things. Favorable (unfavorable) moral judgments about something express favorable (unfavorable) attitudes about it. "Lincoln was a good man" means roughly "Yea Lincoln." So understood, emotivism denies the obvious phenomenon of moral disagreement. Suppose that you claim that Stalin was a good man and I claim that he was a bad man. If moral judgments were mere expressions of attitudes, then this could not constitute a disagreement. We might both agree that you like Stalin and I dislike him. Similarly, it is not a disagreement if you express your fondness for a particular flavor of ice cream and I express my distaste for that same flavor.

Alfred Ayer (1952) and Charles Stevenson (1944) defend more sophisticated versions of emotivism. According to Ayer, to make a moral judgment is to express an attitude with the intention of influencing the attitudes or actions of other people: "Lincoln is a good man" means roughly "Yea Lincoln, catch the wave." In cases of moral disagreement, each party is attempting to alter the attitudes of the other. Stevenson holds that moral disagreement involves a disagreement in attitudes (the parties to the disagreement have incompatible attitudes about something), and each party is attempting to change the attitudes of the other party about the thing in question. Stevenson says that "*X* is good" means roughly "I approve of *X*; do so as well."

These revised versions of emotivism still do not afford a satisfactory account of moral disagreement. In cases of moral disagreement, people not only disagree in their attitudes and try to cause others to *share their attitudes*, they assert that their own attitudes are correct or justified and that the attitudes of those who disagree with them are mistaken. If two people disagree about whether or not Stalin was a good man, each claims that the other's attitudes about Stalin are mistaken or inappropriate.

Ayer on moral reasoning. Moral disputes often involve disagreements about factual questions. Ayer says that, to the extent that moral disagreements involve disagreements about "factual" questions, they can be rationally debated. For example, we can rationally debate whether the institution of capital punishment deters murder and whether it frequently results in the execution of innocent people. Sometimes, however, moral disagreements are based on differences in basic moral principles. Utilitarians believe that we should always do whatever will have the best consequences. Some people are unconditional pacifists. They believe that killing people is always wrong no matter what, even if killing saves many lives and produces much better consequences than not killing. Utilitarians and unconditional pacifists accept incompatible basic moral principles. According to Ayer, when people disagree about matters of basic principle, their disagreements cannot be rationally debated or rationally resolved. (Gibbard's *Wise Choices and Apt Feelings*, discussed below, is a recent development of emotivism.)

5. **HARE'S PRESCRIPTIVISM**. One of the notable features of this theory is that Hare offers a systematic reply to Ayer's claims about the limits of moral reasoning. Hare claims that moral judgments are prescriptions that are universalizable and overriding. Prescriptions are commands, or imperatives, for example: "Don't lie!" and "Shut the door!" Since commands are not statements that are true or false, prescriptivism is a noncognitivist theory. To say that moral judgments are universalizable means that if one makes a moral judgment about a particular case, then one must make the same judgment about any cases that are similar to it in all morally relevant respects. (If I say that it is morally permissible for me to lie to my customers in a certain situation, then I am committed to the view that it would be permissible for others to lie to me in relevantly similar situations.) To say that moral judgments are overriding means that a person who makes moral judgments takes the prescriptions expressed by them to override any conflicting nonmoral considerations, such as considerations of prudence, etiquette, and the law. According to Hare statements of the form: "It is morally wrong all things considered for you to do *X*, but, nevertheless, you would be justified in doing *X*," are self-contradictory. On Hare's view it is also inconsistent to say that it is morally wrong (all things considered) for you to do *X* but still command or advise you to do *X*.

In response to Ayer's claims about the limits of moral reasoning, Hare would say that sometimes we can argue against another person's moral judgments by showing that they are *inconsistent*. Hare claims that requirements of consistency severely constrain the kinds of moral judgments we can make. Suppose that a dishonest plumber claims that it is morally permissible (or even obligatory) for him to defraud his customers and bill them for unneeded repairs that cost them thousands of dollars. To be consistent, he must say that it would be morally permissible (obligatory) for others to defraud him and those he cares about in relevantly similar (hypothetical and/or actual) cases. Since moral judgments are prescriptive, he is committed to prescribing that others defraud him and those he loves in relevantly similar circumstances. Consistency also requires that he refrain from objecting if others defraud him and those he loves in such cases.

In *Freedom and Reason*, Hare considers the case of a Nazi who claims that it is his moral duty to kill Jews. To be consistent, the Nazi must hold that others should kill him if he is Jewish. Suppose that we show the Nazi that, unbeknownst to him, he and his wife are Jewish. To be consistent, the Nazi must say (and *mean*) "All right, send me and my family to an extermination camp." Since moral judgments are overriding, the Nazi cannot consistently make any commands or pleas to the contrary. Hare thinks that few people can be consistent Nazis. Hare allows that a Nazi could be consistent if he or she so hates Jews that he or she sincerely holds that, (in Hare's words): "'Jews are such an abomination that I and my whole family, if we were Jews, should be sent to the gas chamber'" (Hare 1963, p. 172). One can be a consistent Nazi if one is willing to have one's moral principles applied against one's own interests and the interests of those one loves.

Hare gives no reason to think that the Nazi's distinctive moral views are false or mistaken, just that it is difficult to be a consistent Nazi. This concession bodes ill for Hare's theory of moral reasoning since if it cannot establish the correctness of the view that the Nazi's actions are wrong, it is doubtful that it can establish the correctness of *any* ethical judgments. However, Hare is too quick to concede the limits of his arguments in this case. It may be possible to be a consistent Nazi provided that one has a very great hatred of Jews and desires their extermination more than the continuation of one's own life and the lives of one's loved ones. The obvious question to ask here is whether such hatred is rational, and it seems that it is not. Such hatred depends on numerous false beliefs about the characteristics of Jews and their responsibility for the ills of the world. A Nazi could be consistent provided that she is willing to have her principles applied against herself and her loved ones, but a Nazi could not be *both* consistent and adequately informed about matters relevant to her moral convictions. (See Hare's *Moral Thinking* [1981] for a later development of his views.)

As Hare himself notes, his consistency arguments apply only to people who make moral judgments. They do not apply to amoralists who refrain from making moral judgments. Hare's consistency arguments cannot show why we should not be amoral. Even more worrisome for Hare's purposes is that people who employ alternative normative concepts are able to endorse horrendous acts such as the extermination of the Jewish people. Suppose that a Nazi rejects the concepts of morally right and wrong actions in favor of a "Code of Honor" according to which it is "honorable" to kill Jews. Hare's theory does not give us any basis for criticizing such views.

This last possibility raises an important and somewhat neglected set of issues. There are many different alternative normative concepts (e.g., concepts of moral obligation and right and wrong, concepts of virtue, and concepts of honor). People are free to employ any of these concepts and order their lives in accordance with them; people are also free to reject any of these concepts. Philosophers who write about metaethics need to say much more about the *choices* we make in accepting and rejecting various normative concepts. They also need to say much more about the question of how, if it all, we can justify the choices we make in accepting/wielding certain concepts rather than others. (See Friedrich Nietzsche [1967, 1988], Hare [1981], Simon Blackburn [1993], John Mackie [1977], and Bernard Williams [1985] for discussions that shed light on these issues; also see the discussions of Foot, Gibbard, and "Incommensurability" below.)

6. FOOT AND THICK MORAL CONCEPTS. Both emotivism and prescriptivism imply that the concepts of good and bad and right and wrong have no fixed descriptive meaning. One can consistently apply these terms to any things (or any actions). For example, it is perfectly consistent to say that it is morally obligatory to clasp and unclasp one's hands every half hour or to say that bringing it about that the number of hairs on one's head is an even number is a great intrinsic good. These are sincere coherent ethical judgments provided that the person who makes them has the attitudes they express or is willing to consistently universalize the prescriptions that they express. Foot (1978) argues that this is a serious mistake

because there are limits to the things to which "good" and other moral terms can be consistently applied. She thinks that emotivism and prescriptivism make a mistake that is comparable to the sort of mistake one would make if one said that being proud of something consists simply in having a certain sort of attitude about it and that, in principle, *anything* could be the object of one's pride.

Pride is not just a feeling, a welling up in one's chest that one can have about anything. The object of one's pride must be: (a) one's own somehow, and (b) an achievement or something that one takes to be good. Even if I puff up my chest and feel it welling up as I look at the sky, it is not correct to say that I am (feeling) proud of the sky unless there is some background belief that explains why I think that it is somehow *mine*. Foot claims that something can be called a "good action" only if it satisfies one of the following conditions: (a) It is the fulfillment of a special duty derived from a role or promise or (b) it exemplifies a virtue. It follows that we cannot say that twiddling one's thumbs four times each day, for instance, is a morally good action in the absence of special reasons for thinking that it fulfills a duty or exemplifies a virtue.

Foot's own analysis of the meaning of moral judgments is a combination of noncognitivism and (naturalistic) cognitivism. She claims that moral judgments have both evaluative meaning (they express attitudes and guide actions) and descriptive meaning. Moral concepts that have both kinds of meaning are called "thick" concepts. Emotivism and prescriptivism claim that the concepts of right and wrong and good and bad are "thin" concepts. Thin normative concepts have no fixed descriptive meaning, only evaluative meaning.

Foot's theory aptly describes the meaning of the terms we use to refer to moral virtues and vices. Terms such as "generous," "cowardly," and "honest" are thick concepts—they have both evaluative and descriptive meaning. The words "generous" and "honest" commend or express favorable attitudes about the things to which they are applied. There are clear descriptive criteria for using such terms. It is a misuse of language to apply them to things that do not satisfy those criteria. It would be a misuse of the word "generous" to apply it to someone who never gives any tangible goods or time or effort to other people even though that person has a great deal of money and leisure time and many opportunities to help others in need. Foot's theory helps us to frame some important questions: "In ordinary language, are the terms 'good' and 'bad' and 'right'and 'wrong' thick or thin concepts?" "If these concepts are thin concepts, should we dispense with them in favor of thick concepts?" Foot seems to think that the concepts of good and bad and right and wrong are thick concepts. By contrast, some proponents of "virtue ethics," including Williams (1985), think that the concepts of good and bad and right and wrong are thin concepts. However, they think that these thin concepts should be dropped or greatly downplayed in favor of the thick concepts that refer to virtues and vices.

Thick concepts mandate particular evaluations of certain kinds of things. Many thick concepts encapsulate objectionable evaluations (e.g., ethnic slurs). The word "n___" only applies to people of African origin; it cannot be correctly applied to Chinese or Europeans. The word "n___" also expresses contempt for Africans. Those who do not think that Africans, qua Africans, are worthy of contempt do not use the word "n___" (or do not use the word nonironically). (Similar comments apply to all other "ethnic slur terms.") Honor is another example of a thick concept that many people have reasonably chosen to abandon because they reject the evaluations implicit in its use. Given certain concepts of honor, it is dishonorable for me to not to challenge you to a duel to the death if you insult me or show me disrespect. Given other concepts of honor, it is dishonorable for me not to kill my sister if she is raped. These examples make it clear that, for any thick concept that we employ, we should be open to criticisms of the evaluations implicit in that concept and consider the possibility that they are mistaken and cease employing the concept. Thus, it is at least arguable that we need higher-level thin concepts in terms of which to assess the evaluations implicit in the thick concepts we use and encounter.

7. **NORMATIVE RELATIVISM**. This theory is defended by many anthropologists, including Ruth Benedict (1887–1948) and William Sumner (in Moser and Carson 2001). They claim that "*X* is morally right" means roughly "*X* is approved of by my society." This view is open to very serious objections. It implies that statements such as "Slavery is morally wrong, even though my society approves of slavery" are self-contradictory. But such statements are not self-contradictory. A person can criticize or dissent from the moral standards of her own society without contradicting herself. Normative relativism also implies that many ostensible moral disagreements between members of different societies are not genuine disagreements. Suppose that I am a member of a society that approves of the institution of slavery and you are a member of a society that disapproves of slavery. I claim that slavery is just and morally permissible. You object

and claim that slavery is unjust and impermissible. Surely this is a moral disagreement; what I say contradicts what you say. But, according to normative relativism, there is no disagreement in this case. My statement is perfectly consistent with your statement, and both statements are *true*—it is true that my society approves of slavery and true that your society disapproves of slavery.

8. GIBBARD'S EXPRESSIVISM. This theory is a recent development of emotivism that incorporates elements of the Brentano–Ewing theory. Gibbard (1990) analyzes moral judgments as claims about the rationality (or aptness) of feelings of guilt and anger. What a person does is *morally wrong* if and only if it is rational for him to feel guilty for having done it and for others to be angry at him for having done it. He defends an expressivist/emotivist analysis of rationality. According to Gibbard, to say that something (an act, belief, or feeling) is rational is to express one's acceptance of norms that permit it. Unlike Moore, Hare, and Foot, Gibbard does not claim to be offering an analysis of the ("ordinary language") meaning of moral terms. Rather, he describes his theory as a *proposal* about how to use normative concepts. Evidently, not every society has a normative system that includes norms for guilt and anger. Thus, on Gibbard's narrow construal of "morality," not every society has a moral code. Gibbard raises important questions about our choices between alternative normative concepts. Among other things, he asks about the value of morality (narrowly construed): Would we be better off with a normative code in which norms for guilt and anger didn't play a central role? Gibbard offers an answer to Nietzschean criticisms about the value of morality. He says that moral norms help coordinate guilt and anger. Guilt assuages anger and thereby helps promote peace between human beings. Normative codes that do not include norms for guilt will not be able to assuage anger and promote reconciliation between human beings as well as *moral* codes.

II. MORAL TRUTH, MORAL REALITY

We now turn to questions about moral truth and moral reality. In what sense, if any, can moral judgments be true or false or correct or incorrect? Are moral judgments objectively true or false in the way that we take ordinary "factual" statements to be? Is there a moral reality or something else in virtue of which moral judgments are true or false (or correct or incorrect)? We cannot answer these questions simply by appealing to theories of meaning. When we ask whether moral judgments are true or false, we are not simply asking about what we mean or *claim* when we make moral judgments. We are asking whether there is anything that backs up our moral judgments and makes them true or correct.

COGNITIVISM, NONCOGNITIVISM, AND THE TRUTH OF MORAL JUDGMENTS. Cognitivists hold that moral judgments are statements that *purport* to be true. This is compatible with the view that there are no moral facts (no moral reality) that back up our moral judgments and make (some of) them true or correct. Mackie (1977) holds such a view, which he calls an "error theory" of morality. Mackie is a cognitivist who claims that moral judgments are statements that assert or presuppose the existence of objective values. However, he claims that since objective values do not exist, *all moral judgments are false*.

Noncognitivists hold that moral judgments are not statements that purport to be true or false. Strictly speaking, noncognitivists cannot say that moral judgments are true or false. However, they can still say that moral judgments possess something that closely resembles truth or falsity. Emotivists can say that moral judgments are reasonable or unreasonable depending on whether the emotions or attitudes they express are reasonable or unreasonable. (At the very least attitudes and emotions can be unreasonable if they are based on false beliefs.) Prescriptivists can also make sense of something resembling the idea of moral truth. In *Moral Thinking*, Hare claims that there are certain moral judgments that an informed, consistent person must endorse, provided that he or she makes any moral judgments at all. These are judgments that we can reject only by opting out of moral discourse altogether.

MORAL OBJECTIVISM. Our ordinary notion of truth is a notion of *objective truth*. If something is true, then it is *true for everyone* (and true for everyone, everywhere, at all times). (Thus, it is misleading to use the word "truth" as many relativists do when they claim that the truth of moral judgments is "relative to" different people so that a moral judgment that is "true for" one person may not be "true for" another.) Let us use the term "objectivism" to refer to the view that moral judgments are objectively true or false (or objectively correct or incorrect in some sense that closely resembles truth or falsity). We should distinguish between the view that there is an objectively correct answer to *every* moral question and the view that there are objectively correct answers to *some*, but not all, moral questions. Call the former view "unqualified objectivism" and the latter view "qualified objectivism." Call the view that there are no objectively correct answers to any moral questions "unqualified nonobjectivism."

Unqualified objectivism implies that for any moral question, for example: "Was it right for Ms. Jones to have an abortion in April 1999?" there is an objectively correct answer and that anyone who gives a conflicting answer is mistaken. Qualified objectivism holds that there are some moral questions about which there are objectively correct answers and other moral questions about which there are no objectively correct answers. Unqualified nonobjectivism implies that, for any moral question, there is no objectively correct answer to that question.

MORAL REALISM. Statements such as "the earth is less than 100,000,000 miles from the sun" are true in virtue of facts that hold independently of what we believe or desire. *Moral realism* is the view that there are moral facts in virtue of which moral judgments are objectively true or false and that these facts are logically independent of the beliefs, attitudes, emotions, or preferences of rational beings and independent of the beliefs, attitudes, emotions, or preferences that rational beings would have in hypothetical situations (e.g., the moral beliefs that someone would have if she or he were fully informed about relevant facts). Moral nonrealism is the view that there are no independent moral facts. The truth of moral realism would guarantee the truth of moral objectivism, but one can be a moral objectivist without being a moral realist. Immanuel Kant, Roderick Firth (1917–1987), and Michael Smith (1954–) (see below) are moral objectivists but not moral realists. Hare's *Moral Thinking* also defends nonrealist moral objectivism.

IIA. ARGUMENTS AGAINST MORAL OBJECTIVISM

We now turn to arguments against moral objectivism.

1. DISAGREEMENT. Moral disagreement is widespread among ostensibly sane and rational people. Consider the following argument:

> (1) There is disagreement among rational people about the answers to all (some) moral questions.
>
> (2) If there is disagreement among rational people about the answer to a question, then there is no objectively correct or objectively true answer to that question.

Therefore, unqualified (qualified) nonobjectivism is true. (There are no moral questions [there are only some moral questions] for which there is an objectively correct answer.)

The cogency of this argument depends on the account of moral truth or correctness that the objectivist gives. If moral realism is true, then there are moral facts in virtue of which moral judgments are objectively true that are independent of what we believe. So, if moral realism is true, then moral objectivism is true, and the phenomenon of moral disagreement among rational people is not a serious objection to moral objectivism. Similarly, disagreement between reasonable people does not constitute any kind of objection to the view that ordinary historical judgments are objectively true or false. Consider the question: "Did Lee Oswald fire any of the shots that killed President Kennedy?" Rational people disagree about the answer to this question, and, at the present time, it may be impossible to know for certain what the answer is. In spite of the disagreement, there is an objective fact of the matter—either Oswald fired some of the fatal shots or else he did not. However, the phenomenon of ethical disagreement among ostensibly rational and well-informed people constitutes a serious argument against attempts to defend moral objectivism by appeal to theories of rationality because such theories claim that objective moral truths are constituted by the *agreement* of rational people. One standard rejoinder to the objection about disagreement is the claim that "*ideally* rational" or "*fully* rational" people would not have moral disagreements. (See the discussion of the ideal observer theory below.)

Digression: Disagreement vs. incommensurabilit. Moral disagreement should not be confused with moral incommensurability. Two people can disagree about whether an act is right or wrong only if they share the concepts of right and wrong action. Often, the differences between the moral views of different societies constitute cases of incommensurability rather than disagreement. Many philosophers and cultural anthropologists (including Nietzsche and Gertrude Anscombe) claim that the concept of moral obligation is unique to Judaism, Christianity, Islam, and the civilizations that developed from those religions. If this is true, then it is doubtful that Genghis Khan (c. 1167–1227) and his warriors possessed our concept of a moral obligation and morally right and wrong actions. I condemn the Mongol destruction of Iraq in the thirteenth century as a morally wrong action. Many or most of the Mongols who took part in this did so in good conscience, but it would probably be incorrect to say that they thought that what they did was morally right (they could not have this belief unless they employed the concepts of right and wrong actions). Even though we do not disagree about the *moral rightness* of this action, there is surely some kind of disagreement here. My moral judg-

ments and the attitudes they endorse are contrary to many of Genghis Khan's attitudes; Genghis Khan clearly approves of (and thinks it correct to approve of) actions that I disapprove of and condemn as morally wrong.

2. **THE APPEAL TO TOLERANCE**. This argument goes roughly as follows: "We should be respectful and tolerant of other people's moral views; however, moral objectivism implies that many people's views are mistaken and thus not worthy of respect or toleration." This argument is widely accepted and motivates many people to reject moral objectivism, but, on examination, it is a very weak argument that few philosophers take seriously. First, endorsing moral objectivism does not commit one to being intolerant of the moral views of others. If I am an unqualified objectivist, then I think that there are objectively correct answers to every moral question. However, my being an objectivist does not entail that I claim to know what those answers are. Nor does my being an objectivist entail that I think that the views of others who disagree with me are worthy of disrespect or suppression. An objectivist can claim that objectively true moral principles require tolerance and respect for the views of others. Second, nonobjectivism does not imply that we should be tolerant. Nonobjectivists can endorse first-order moral principles that permit or require them to be intolerant of the views of others. All that follows from nonobjectivism is that one's moral judgments, *whatever they happen to be*, are not objectively true or false or objectively correct or incorrect.

3. **MORAL EXPLANATIONS AND MORAL FACTS**. Gilbert Harman argues that it is unnecessary to posit the existence of moral facts in order to explain phenomena. Thus, moral facts are superfluous entities—there is no reason to suppose that they exist.

> [O]bservation plays a role in science that it does not play in ethics. The difference is that you need to make assumptions about certain physical facts to explain the occurrence of the observations that support a scientific theory, but you do not seem to need to make assumptions about any moral facts to explain the occurrence of the so-called moral observations I have been talking about. In the moral case, it would seem that you need only make assumptions about the psychology or moral sensibility of the person making the moral observation. In the scientific case, theory is tested against the world.
>
> (HARMAN 1977, P. 6)

Among the phenomena that moral facts might explain are the moral judgments we make and the moral sentiments we feel (e.g., feelings of guilt and indignation). According to Harman, these phenomena are fully explained by our psychology and the fact that we accept certain moral principles; we do not need to assume the existence of moral facts or assume that those principles are true. Harman gives the following example: Someone tortures an animal. You believe that this action is wrong and you feel moral indignation. Harman says that we do not need to postulate moral facts in order to explain *your* belief and *your* indignation. They are explained by the fact that you were taught certain moral principles and have a certain psychological make up. Your *accepting* the moral principles in question is necessary to explain your beliefs and your indignation, but *we* do not need to assume that your moral principles are *true*. By contrast, in science, we can justify the postulation of entities by their ability to explain our observations of the world. The postulation of atoms helps to explain such things as Geiger counters and nuclear bombs. (Nicholas Sturgeon offers an influential reply to this argument. See below.)

IIB. ARGUMENTS IN FAVOR OF MORAL OBJECTIVISM

Most contemporary philosophers who defend moral objectivism do so on one of the following three grounds:

1. **THE APPEAL TO MORAL REALISM**. As explained earlier, the truth of moral realism would guarantee the truth of moral objectivism. Some versions of moral realism claim that moral properties are "nonnatural" properties. Such theories are widely criticized on the grounds that the entities that they postulate do not exist or that those entities cannot plausibly be identified with moral properties (see above). In light of these criticisms of nonnaturalist versions of moral realism, recent attempts to defend naturalistic versions of realism are particularly noteworthy.

Sturgeon's naturalistic realism. Sturgeon claims that moral facts are constituted by natural facts. Sturgeon holds that moral properties are identical with natural properties, but he does not take statements asserting the identity of moral properties and natural properties to be analytic. He claims that his view is invulnerable to Moore's open question argument. Sturgeon also attempts to answer Harman's argument about explanation. He defends his theory on the grounds that moral facts help explain certain phenomena. Sturgeon offers the following

example: Hitler's moral depravity helps explain why he started World War II and ordered the Holocaust. If Hitler had not been morally depraved, he would not have started World War II, and he would not have ordered the Holocaust.

Harman's thesis implies that the supposed moral fact of Hitler's being morally depraved is irrelevant to the explanation of Hitler's doing what he did. To assess this claim, we need to conceive a situation in which Hitler was *not* morally depraved and consider the question whether in that situation he would still have done what he did. My answer is that he would not, and this answer relies on a (not very controversial) moral view: that in any world at all like the actual one, only a morally depraved person could have initiated a world war, ordered "the final solution," and done any number of other things Hitler did. That is why I believe that if Hitler had not been morally depraved, he would not have done those things and hence that the fact of his moral depravity is relevant to an explanation of what he did (Sturgeon "Moral Explanations," in Sayre-McCord 1988, p. 249). Sturgeon's arguments have generated a very lively debate. (Criticisms by Terrance Horgan and Mark Timmons (1951–) are particularly noteworthy.)

2. **THEORIES OF RATIONALITY**. Many nonrealists claim that there are objective moral facts that are constituted by facts about what it is rational for people to believe or desire or "will." Kant, Hare, and Christine Korsgaard defend such views. Kant holds that moral truths (truths about what is right and wrong) are truths about what we can rationally and consistently will. For Kant moral truths are truths of reason.

The ideal observer theory (IOT) uses the idea of an ideally rational moral judge or ideal observer as a standard for the truth of moral judgments. According to Firth's version of the IOT, a favorable moral judgment about *X* ("*X* is good/right") is (objectively) true provided that all ideal observers would feel approval for *X*. An unfavorable moral judgment about *X* ("*X* is bad/wrong") is (objectively) true provided that all ideal observers would feel disapproval for *X* (Hospers and Sellars, pp. 200–221). In *Ethical Theory*, Richard Brandt (1959) defends a different version of the IOT. Brandt says that a moral judgment *X* is objectively true provided that all ideal observers would accept or believe *X*. (David Hume and Adam Smith also defend versions of the IOT.) Firth ascribes the following characteristics to an ideal observer:

(1) Omniscience or knowledge of all nonmoral facts;

(2) omnipercipience, or the ability to imagine vividly any events or states of affairs, including the experiences of others;

(3) disinterestedness, that is, not having any interests or desires that involve essential reference to particular persons or things;

(4) dispassionateness, that is, not having any emotions that are directed upon objects because they are believed to have essentially particular features;

(5) consistency;

(6) normality "in other respects."

Firth thinks that all ideal observers would feel approval and disapproval for the same things. Given this, and given his version of the IOT, unqualified moral objectivism is true. Brandt thinks that ideal observers would agree about the answers to some, but not all, moral questions. He thinks that the IOT commits us to qualified objectivism. If both Firth and Brandt are mistaken and ideally rational moral judges could disagree in their attitudes or judgments about *every moral question*, then the IOT commits us to unqualified nonobjectivism.

Ideal observers are characterized as "informed" or "fully informed." Brandt says that ideal observers must possess all information "relevant to" the issues they judge. Firth notes difficulties in determining which facts are and are not relevant to answering a given moral question. He contends that there is no way to determine which facts are and are not relevant to a given moral question without presupposing answers to controversial moral questions that the IOT is supposed to provide rather than presuppose. Because of this he feels compelled to say that an ideal observer is *omniscient* with respect to all nonmoral facts. There is an unintended irony in Firth's characterization of an ideal observer as a *human being* who is omniscient "but otherwise normal." Humans are very far from being omniscient. It is not clear that it makes sense to talk about how you or I would react if we were omniscient. An omniscient being would have to be God or some kind of deity. If we press this point, then the IOT starts to look a lot like a divine will theory of morality.

Michael Smith's *The Moral Problem* (1994) defends a theory that closely resembles the IOT. With qualifications, Smith holds that to say that an action is morally right means that we have normative reason to do it. What a person has normative reason to do is what he or she *would desire to do* if fully rational (being fully rational includes having no false beliefs and having all relevant true beliefs). Smith stresses that what I have normative

reason to do (in my actual circumstances) depends on what my fully rational self would want (or advise) my actual self to do. (Smith's theory is sometimes called an "ideal advisor theory.") Smith thinks that his theory implies that there are (objective) moral facts. On his account, the judgment "it is right for *S* to do *X*" is objectively true provided that we would *all desire* that *S* do *X* if we were fully rational. Smith thinks that an *ideally rational* process of reflection and debate between people who initially disagree in their desires about what people should do is likely to yield (complete) agreement.

According to Smith, thick moral concepts such as "honesty" and "treachery" reveal considerable agreement about what is right and wrong. The common use of such concepts reveals that nearly everyone agrees that acts of treachery are wrong (other things equal) and that acts of honesty are right (other things equal). Smith also argues that much seemingly intractable moral disagreement has its origins in (unreasonable) appeals to religious authority. (Others take a very different view about the relevance of religion to these issues; see below.) Smith says that the case for moral objectivism ultimately depends on the outcome of (first-order) debates in normative ethics. In order to determine "whether or not there *really are* moral facts," we must "engage in normative ethical debate and … see where the arguments that we give ultimately lead us" (Smith 1994, p. 202). "The real question is whether we will, by engaging in such debate, come up with answers to moral questions that secure the free agreement of those who participate" (Smith 1994, p. 201).

3. **THE APPEAL TO GOD'S WILL**. Some hold that God's will constitutes the objective standard for the truth of moral judgments. The view that God created human beings for certain purposes is one way of making sense of the widely held view that ethical theories should be based on theories of human nature or the "telos" of human life. The most well-known theory that attempts to base (objective) morality on God's will is the divine command theory. The traditional divine command theory (TDCT) holds that God's commands constitute the ultimate standard of right and wrong. What makes an act morally obligatory is that God commands it; what makes an act morally permissible is that God permits it; what makes an act is morally wrong is that God forbids it.

There are a several standard objections to the TDCT. These objections are widely regarded as fatal or decisive. (1) The TDCT implies that nothing anyone does can be morally obligatory or morally wrong unless God exists and commands and forbids us to do certain things. However, certain actions would be right or wrong even if God did not exist. (2) The TDCT implies that *any act* would be right if God commanded us to do it. But certain acts (e.g., acts of cruelty or murder) would be wrong even if God commanded us to perform them. (3) The TDCT implies that what is wrong/obligatory is wrong/obligatory *because* God forbids/commands it. The TDCT does not allow us to say that *God forbids murder because murder is wrong*. Rather, the TDCT implies that *murder is wrong because God forbids it*. Thus, the TDCT implies that God has no reason to command one thing rather than another, and God's arbitrary commands cannot be the basis for genuine moral obligations.

These may be fatal objections to the TDCT. However, it does not follow that all theories that attempt to make God's will the basis for an objective morality are subject to fatal objections. Robert Adams (1937–) has formulated a modified version of the TDCT that avoids all of the objections to the TDCT. Adams's modified TDCT can be stated roughly as follows:

> If there is a loving God then: (1) an action is obligatory if, and only if, a loving God commands it; (2) an action is morally permissible if, and only if, a loving God permits it; and (3) an action is morally wrong if, and only if, a loving God forbids it. If there does not exist a loving God, then the rightness or wrongness of actions is determined in some other way.

Adams (1987) holds that if there is a loving God, then right and wrong are determined by God's commands; if there does not exist a loving God, then right and wrong are determined in some other way. Thus, Adams's theory avoids the first objection. Adams's modified TDCT also avoids the second objection. It does not imply that we would be morally obligated to obey God's commands if God commanded cruelty. If God commanded cruelty for its own sake, he would thereby show himself to be unloving. Adams's theory does not imply that we would be obligated to follow the commands of a cruel or unloving God. Adams cannot say that God commands what he commands *because* it is morally right (independently of being commanded by God). But Adams is *not* committed to the view that God's commands are arbitrary or that God has no reason to command one thing rather than another. Adams can say that God commands what he commands because of his loving nature and because he is omniscient. (See Linda Zagzebski's *Divine Motivation Theory* [2004] for a very different sort of religiously based moral theory.)

Suppose that there exists an omniscient loving God who created human beings for certain purposes. Suppose *also* that moral realism is false and there are no independent moral facts to which God's will must conform on pain of error. Given all of this, it is plausible to regard God's will and purposes as objective standards of morality; God's standpoint for assessing things is arguably more authoritative than that of a maximally rational human being. If moral realism is false and God does not exist, then the most promising basis for defending moral objectivism is by appeal to a theory of rationality.

See also Anscombe, Gertrude Elizabeth Margaret; Ayer, Alfred Jules; Brandt, R. B.; Brentano, Franz; Constructivism, Moral; Emotive Theory of Ethics; Error Theory of Ethics; Ethical Naturalism; Ethical Relativism; Ethics; Ethics, History of; Foot, Philippa; Hare, Richard M.; Harman, Gilbert; Hume, David; Ideal Observer Theories of Ethics; Internalism and Externalism in Ethics; Intuitionism, Ethical; Kant, Immanuel; Kripke, Saul; Mackie, John Leslie; Moore, George Edward; Moral Realism; Moral Skepticism; Nietzsche, Friedrich; Noncognitivism; Objectivity in Ethics; Projectivism; Putnam, Hilary; Rationalism in Ethics (Practical Reason Approaches); Smith, Adam; Stevenson, Charles L.; Sumner, William Graham; Williams, Bernard.

Bibliography

Adams, Robert. *The Virtue of Faith*. Oxford: Oxford University Press, 1987.

Ayer, A. J. *Language, Truth, and Logic*. 2nd ed. New York: Dover, 1952.

Blackburn, Simon. *Essays in Quasi-Realism*. Oxford: Oxford University Press, 1993.

Brandt, Richard. *Ethical Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1959.

Brandt, Richard. *A Theory of the Good and the Right*. Oxford: Oxford University Press, 1979.

Brentano, Franz. *The Origin of Our Knowledge of Right and Wrong*, edited by Roderick Chisholm. Translated by Roderick Chisholm and Elizabeth Schneewind. New York: Humanities Press, 1969.

Ewing, A. C. *The Definition of Good*. New York: Macmillan, 1947.

Foot, Philippa. *Virtues and Vices*. Berkeley: University of California Press, 1978.

Gibbard, Allan. *Wise Choices, Apt Feelings*. Cambridge, MA.: Harvard University Press, 1990.

Hare, R. M. *Freedom and Reason*. Oxford: Oxford University Press, 1963.

Hare, R. M. *The Language of Morals*. Oxford: Oxford University Press, 1952.

Hare, R. M. *Moral Thinking*. Oxford: Oxford University Press, 1981.

Harman, Gilbert. *The Nature of Morality*. Oxford: Oxford University Press, 1977.

Horgan, Terrance, and Mark Timmons. "Troubles on Twin Moral Earth: Moral Queerness Revived." *Synthese* 92 (1992): 221–260.

Hurka, Thomas. *Perfectionism*. Oxford: Oxford University Press, 1993.

Korsgaard, Christine, with G. A. Cohen, et al. *The Sources of Normativity*, edited by Onora O'Neill. Cambridge, U.K.: Cambridge University Press, 1996.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1972.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Middlesex, U.K.: Penguin, 1977.

Moore, G. E. *Principia Ethica*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1993.

Nietzsche, Friedrich. *On the Genealogy of Morals*. In *On the Genealogy of Morals and Ecce Homo*, translated by Walter Kaufmann. New York: Vintage Books, 1967.

Nietzsche, Friedrich. *The Twilight of the Idols and the Anti-Christ*. Translated by R. J. Hollingdale. Middlesex, U.K.: Penguin, 1968.

Putnam, Hilary. "The Meaning of 'Meaning.'" In *Mind, Language, and Reality: Philosophical Papers*. Vol II. Cambridge, U.K.: Cambridge University, Press 1975.

Ross, W. D. *The Foundations of Ethics*. Oxford: Clarendon Press, 1939.

Ross, W. D. *The Right and the Good*. Oxford: Clarendon Press, 1930.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. New York: Dover, 1966.

Smith, Michael, *The Moral Problem*. Oxford: Blackwell's, 1994.

Stevenson, Charles L. *Ethics and Language*. New Haven, CT: Yale University Press, 1944.

Warnock, G. J. *The Object of Morality*. London: Methuen, 1971.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: 1985.

Zagzebski, Linda. *Divine Motivation Theory*. Cambridge, U.K.: Cambridge University Press, 2004.

ANTHOLOGIES

All of the following anthologies contain extensive bibliographies:

Darwall, Stephen, Allan Gibbard, and Peter Railton, eds. *Moral Discourse and Practice*. Oxford: Oxford University Press, 1997.

Hospers, John, and Wilfrid Sellars, eds. *Readings in Ethical Theory*. 2nd ed. Englewood Cliffs: Prentice-Hall, 1970.

Moser, Paul K., and Thomas Carson, eds. *Moral Relativism: A Reader*. Oxford: Oxford University Press, 2001.

Sayre-McCord, Geoffrey, ed. *Essays on Moral Realism*. Ithaca, NY: Cornell University Press, 1988.

Thomas L. Carson (2005)

METAMATHEMATICS

See *Mathematics, Foundations of*

METAPHOR

"Metaphors" have an emotive force and aesthetic dimension that have long been recognized. What has made metaphor so compelling to contemporary philosophers, however, has been its importance to cognition. Aesthetics and philosophy of religion are no longer the sole province of the study of metaphor. Instead, most of the research is located in philosophy of language, philosophy of science, and cognitive science. The ubiquity of metaphor and its contribution to all forms of discourse, the apparent anomaly of metaphor in light of standard accounts of language, and the increased interest by philosophers in providing theories for natural (rather than formal or artificial) languages have made an account of metaphor an important criterion of adequacy for theories of language. The limits of literality have similarly been felt in accounts of science and cognition. Max Black's (1962) seminal work connecting the use of scientific models to metaphors opened an area of inquiry now pursued by psychologists and cognitive scientists as well as philosophers of science. Some philosophers join questions of the role of metaphor in science to debates concerning scientific realism (Boyd, 1979; Hesse, 1970). The work emanating from theories of language and theories of science and cognition converge in concerns about meaning change, computer modeling of discovery processes, linguistic competencies, creativity, and religious discourse (Soskice, 1985).

While many questions remain, a few issues have been settled. The view of metaphor as an isolated word or phrase that is an occasional, unsystematic, and deviant phenomenon in language valued for its rhetorical force but disdained for its ability to mislead or be used in place of proper argument has been challenged. Metaphors have come to be understood as syntactically complex (Black, 1962; Tirrell, 1991) attributions that may or may not be grammatically deviant (Stern, 1985). In the tradition of I. A. Richards (1936) and Black, metaphors are generally taken to implicate entire conceptual domains or semantic fields (Kittay, 1987) through which a metaphor is interpreted, extended, and even systematically integrated into the language (Lakoff and Johnson, 1980). They either exploit some similarity between the metaphorically used term (the vehicle or source) and the concept spoken of (the topic or target) or create or intimate a similarity. While the similarity appealed to in earlier discussions pertained to intrinsic properties or properties associated with vehicle and topic, similarity has increasingly come to mean a relational or structural similarity—akin to models and analogies—between the contexts or domains (Black, 1962; Goodman, 1968) implicated in the metaphor.

While earlier debates concerned metaphor's cognitive value, current debates accept its cognitive function and ask if this function is properly assigned to metaphoric meaning and whether it is a distinctive form of cognition not reducible to other forms such as the capacity to recognize similarity and make comparisons. The outcome of the debate is important to the nature of language, of thought, and of epistemic enterprises such as science. If metaphors have meaning, then a theory of language must explain how such meaning is determined, and any account of mind in which linguistic capacity plays a central role for cognition must similarly explain how cognitive faculties make use of, and make possible, metaphorical thought. Similarly, if the use of metaphorical language in knowledge domains such as science is not reducible to literal language, then we need metaphor in order to understand and explain what is knowable. Furthermore, if we need metaphor to access scientific knowledge, as well as for aesthetic or evocative purposes, then the domains such as art and religion may be more akin to science—or related in more interesting ways—than we have presumed (Fleischacker, 1994). But if metaphors perform their cognitive function without generating a distinctive meaning, then theories of language that are based on literal language suffice; metaphoric contributions to cognition are assimilable to other, already understood or accepted cognitive abilities; the cognitive role of metaphor would be valuable only as heuristic (although, in the case of combinatorially complex problems, the heuristic contribution of metaphor itself may be irreplaceable), and we maintain a clear delineation between the scientific and the poetic.

The position propounding metaphoric meaning and the cognitive irreducibility of metaphor was staked out by Black and has been buttressed by arguments and evidence gathered by philosophers of science, cognitive psychologists, philosophers of language, and linguists. However, the parsimony of the opposing position, and its elegant articulation by Donald Davidson (1978), continues to make it attractive, despite the counterintuitive claim that metaphors have no meaning and the weighty evidence of metaphor's importance in all cognitive endeavors.

Philosophers claiming that metaphors have meaning generally begin by accepting some version of the interaction theory of metaphor but have utilized the resources of many different semantic theories (e.g., possible-world semantics [Bergman, 1982; Hintikka and Sandu, 1994], semantic-field theory [Kittay, 1987], cognitive semantics

[Gibbs, 1994; Lakoff and Johnson, 1980; Sweetser, 1990], a componential semantics [Levin, 1977], a Wittgensteinian semantic, and David Kaplan's semantics for demonstratives [Stern, 1985]). Some use speech-act theory, claiming that metaphors are a feature of speaker meaning rather than sentence meaning (Searle, 1981) or that metaphors are, in the end, elliptical similes after all (Fogelin, 1988).

Newer comparison theories, versions of the theory that metaphors are elliptic similes or implicit comparisons and so do not have a distinctive meaning, explore the notion of figurative rather than literal similarity (Glucks and Keysar, 1990; Ortony, 1979). Some of these approaches offer a causal theory, opposing it to a semantic theory, claiming that metaphors cause us to make comparison by "intimating similarities" and have a causal effect of creating intimacy among speaker and listener (Cohen, 1978; Cooper, 1986). Questions remain concerning the relation between metaphor and literal language (e.g., Can the distinction be drawn in a clear fashion? Is the interpretative process the same or different? Is language originally metaphorical or literal?) and other nonliteral languages (see Hintikka and Sandu, 1994; Jakobson, 1960).

The importance of metaphor in science was stressed by Mary Hesse (1970), who developed the understandings of metaphors as systematic analogies in which the "neutral"—that is, unexplored analogical relations—provide a distinctive source for predictive claims. Dedre Gentner (1982), a cognitive psychologist, along with her associates has identified features, such as systematicity and higher-order relations, that make some metaphors more productive for cognitive purposes than others.

Noting the affinity between metaphor and analogy has permitted a number of researchers in philosophy and psychology to make headway with computational approaches to metaphor—a promising tool for testing theories of metaphor and for understanding the extent to which accounts of metaphor are amenable to formal and precise accounts (Holyoak and Thagard, 1989; Steinhart and Kittay, 1994). Making use of advances in our understanding of metaphor, theorists have explored the role of metaphor in creativity, in language acquisition and concept formation, and in both the consolidation and the breakdown of habituated patterns of thought such as cultural prejudice. These latter developments (which have especially been taken up by feminist philosophers and other social critics) bring the question of the cognitive role of metaphor full circle, reconnecting it to its rhetorical force.

See also Aesthetics, History of; Aesthetics, Problems of; Black, Max; Cognitive Science; Davidson, Donald; Goodman, Nelson; Hintikka, Jaako; Kaplan, David; Philosophy of Language; Philosophy of Religion; Philosophy of Science, History of; Philosophy of Science, Problems of.

Bibliography

Bergman, M. "Metaphorical Assertions." *Philosophical Review* 91 (1982): 229–245.

Black, M., ed. *Models and Metaphors.* Ithaca, NY: Cornell University Press, 1962.

Boyd, R. "Metaphor and Theory Change: What Is 'Metaphor' a Metaphor For?" In *Metaphor and Thought,* edited by A. Ortony. Cambridge, U.K.: Cambridge University Press, 1979.

Cohen, T. "Metaphor and the Cultivation of Intimacy." *Critical Inquiry* 5 (1978): 3–12.

Cooper, D. *Metaphor.* Oxford: Blackwell, 1986.

Davidson, D. "What Metaphors Mean." In *On Metaphor,* edited by S. Sacks. Chicago: University of Chicago Press, 1979.

Fleischacker, S. "Frustrated Contracts, Poetry, and Truth." *Raritan* 13 (4) (Spring 1994): 47–70.

Fogelin, R. *Figuratively Speaking.* New Haven, CT: Yale University Press, 1988.

Gentner, D. "Are Scientific Analogies Metaphors?" In *Metaphor: Problems and Perspectives,* edited by D. S. Maill. Atlantic Highlands, NJ: Humanities Press, 1982.

Gibbs, R. *The Poetics of Mind: Figurative Thought, Language, and Understanding.* Cambridge, U.K.: Cambridge University Press, 1994.

Glucksberg, S., and B. Keysar. "Understanding Metaphorical Comparisons: Beyond Similarity." *Psychological Review* 97 (1990): 3–18.

Goodman, N. *Languages of Art.* Indianapolis: Bobbs Merrill, 1968.

Hesse, M. *Models and Analogies in Science.* Notre Dame, IN: University of Notre Dame Press, 1966.

Hintikka, J., and G. Sandu. "Metaphor and Other Kinds of Nonliteral Language." In *Aspects of Metaphor,* edited by J. Hintikka. Dordrecht: Kluwer Academic, 1994.

Holyoak, K. J., and P. Thagard. "Analogical Mapping by Constraint Satisfaction." *Cognitive Science* 13 (1989): 295–355.

Jakobson, R. "Closing Statement: Linguistics and Poetry." In *Style in Language,* edited by T. A. Sebeok. Cambridge, MA: Technology Press of MIT, 1960.

Kittay, E. F. *Metaphor: Its Cognitive Force and Linguistic Structure.* Oxford: Clarendon Press, 1987.

Lakoff, G., and M. Johnson. *Metaphors We Live By.* Chicago: University of Chicago Press, 1980.

Levin, S. R. *The Semantics of Metaphor.* Baltimore: Johns Hopkins University Press, 1977.

Ortony, A. "The Role of Similarity in Similes and Metaphors." In *Metaphor and Thought,* edited by A. Ortony. Cambridge, U.K.: Cambridge University Press, 1979.

Richards, I. A. *The Philosophy of Rhetoric.* Oxford: Oxford University Press, 1936.

Searle, J. "Metaphor." In *Philosophical Perspectives on Metaphor,* edited by M. Johnson. Minneapolis: University of Minnesota Press, 1981.

Soskice, J. M. *Metaphor and Religious Language.* Oxford: Clarendon Press, 1985.

Steinhart, E., and E. F. Kittay. "Generating Metaphors from Networks: A Formal Interpretation of the Semantic Field Theory of Metaphor." In *Aspects of Metaphor,* edited by J. Hintikka. Dordrecht: Kluwer Academic, 1994.

Stern, J. "Metaphor as Demonstrative." *Journal of Philosophy* 82 (1985): 677–710.

Sweetser, E. *From Etymology to Pragmatics: The Body-Mind Metaphor in Semantic Structure and Semantic.* Cambridge, U.K.: Cambridge University Press, 1990.

Tirrell, L. "Reductive and Nonreductive Simile Theories of Metaphor." *Journal of Philosophy* 88 (1991): 337–358.

Eva F. Kittay (1996)

METAPHOR [ADDENDUM]

This addendum confines itself to general accounts of the nature of verbal metaphor, setting aside work on such more specialized questions as whether metaphors are paraphrasable and such more general and speculative questions as whether the nonverbal arts provide convincing examples of nonverbal metaphor.

SEMANTIC TWIST THEORIES

Semantic twist theories follow Beardsley in holding that a metaphor is a sentence in which a relation of tension or incongruity obtains among the standing meanings of its constituent words and phrases, a tension which is relieved when some of these meanings (those of what Max Black called the *focus*) change or "twist" so as to come into harmony with the others (those of the *frame*). Semantic twist theories have been devised to fit many different conceptions of meaning and of verbal incongruity (Kittay 1989, Ricoeur 1979, Skulsky 1992). Such theories have trouble accounting for sentences one takes to be metaphors despite the availability of a completely apt and pertinent literal reading, sentences one might call *twice-apt*. An example is the joke epitaph a friend composed for Thomas Hobbes: *This is the true philosopher's stone*.

PRAGMATIC TWIST THEORIES

Pragmatic twist theories (Grice 1989, Searle 1979, Sperber and Wilson 1985/6) hold that when we indulge in metaphor, we use words and phrases with their standard literal meanings to *say* one thing, yet we are taken to *mean—taken as intending to convey*—something else. To put it another way, *our sentence as used by us* means one thing, *we in using it* mean something else—where both "things" are straightforwardly propositional in character. Only by attributing some special meaning to *us* can listeners portray *our utterance* as an intelligible, cooperative contribution to a shared conversational enterprise. Metaphor becomes a mode of overt insinuation, akin to conversational implicature, loose talk, and indirect speech acts. (Theories of this second kind likewise have difficulty accounting for twice-aptness.)

COMPARATIVISM

A new and more robust form of comparison theory (Fogelin 1988) holds that a metaphor "A is (a) B" is an elliptically presented comparison of its primary subject (A) to its secondary subject (B, or Bs in general), where this comparison is to be taken in a distinctively *figurative* manner, as a *simile*. Whether one takes it literally or figuratively, a comparison "A is like (a) B" is true just in case A shares sufficiently many of (a) B's most salient properties. Understanding metaphor becomes a matter of identifying a distinctively figurative way of deciding which properties of (a) B count as salient for present conversational purposes and how many of them count as sufficiently many.

BRUTE FORCE THEORIES

Brute force theories (Davidson 1984, White 1996) hold that in metaphor no words go missing and neither words nor speakers mean anything out of the ordinary. Instead, an utterance that would otherwise be pointless or unaccountable produces what Richard Moran (1989) calls a "framing effect": listeners are induced to view or consider or experience a primary subject A in a special light afforded by the sheer mention, in the midst of a discourse devoted to A, of the secondary subject B.

CONCEPTUAL THEORIES

Conceptual theories (Lakoff 1993, Fauconnier and Turner 2002) hold that verbal metaphor is a manifestation of pervasive modes of thinking wherein people "map" one conceptual domain (e.g., love affairs with their successive stages) onto another (e.g., journeys with their successive stops) or "blend" the systems of terms in which they conceive two different domains.

SEMANTIC ACCOUNTS

An assortment of recent semantic theories (Stern 2000, Walton 1993, Hills 1997) rehabilitate metaphorical truth and metaphorical sentence content outside the confines

of verbal opposition theories by drawing on more general accounts of pretense, presupposition, and demonstrative thought.

See also Beardsley, Monroe C.; Black, Max; Events in Semantic Theory; Hobbes, Thomas; Presupposition; Semantics.

Bibliography

Cavell, Stanley. "Aesthetic Problems of Modern Philosophy." In *Must We Mean What We Say? A Book of Essays*. New York: Scribners, 1969.

Cohen, Ted. "Notes on Metaphor." *Journal of Aesthetics and Art Criticism* 34 (3) (1976): 249–259.

Davidson, Donald. "What Metaphors Mean." In *Inquiries into Truth and Interpretation*, 245–264. Oxford: Clarendon Press, 1984.

Derrida, Jacques. "White Mythology." In *Margins of Philosophy*, translated by Alan Bass. Chicago: University of Chicago Press, 1982.

Fauconnier, Gilles, and Mark Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books, 2002.

Fogelin, Robert J. *Figuratively Speaking*. New Haven, CT: Yale University Press, 1988.

Gibbs, Raymond J. *The Poetics of Mind: Figurative Thought, Language, and Understanding*. New York: Cambridge University Press, 1994.

Goatly, Andrew. *The Language of Metaphors*. London: Routledge, 1997.

Grice, H. Paul. "Logic and Conversation." In *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press, 1989.

Hills, David. "Aptness and Truth in Verbal Metaphor." *Philosophical Topics* 25 (1) (1997): 117–153.

Kittay, Eva Feder. *Metaphor: Its Cognitive Force and Linguistic Structure*. Oxford: Clarendon Press, 1989.

Lakoff, George. "The Contemporary Theory of Metaphor." In *Metaphor and Thought*. 2nd ed., edited by Andrew Ortony. Cambridge, U.K.: Cambridge University Press, 1993.

Moran, Richard. "Seeing and Believing: Metaphor, Image, and Force." *Critical Inquiry* 16 (1) (1989): 87–112.

Ortony, Andrew. *Metaphor and Thought*, 2nd ed. Cambridge, U.K.: Cambridge University Press, 1993.

Ricoeur, Paul. "The Metaphoric Process as Cognition, Imagination, and Feeling." In *On Metaphor*, edited by Sheldon Sacks. Chicago: University of Chicago Press, 1979.

Searle, John R. "Metaphor." In *Expression and Meaning*. Cambridge, U.K.: Cambridge University Press, 1979.

Skulsky, Harold. *Language Recreated: Seventeenth-Century Metaphorists and the Act of Metaphor*. Athens: University of Georgia Press, 1992.

Sperber, Dan and Deirdre Wilson, "Loose Talk." *Proceedings of the Aristotelian Society* 86 (1985/6), 153-171.

Stern, Josef. *Metaphor in Context*. Cambridge, MA: MIT Press, 2000.

Walton, Kendall L. "Metaphor and Prop Oriented Make-Believe." *European Journal of Philosophy* 1 (1) (1993): 39–57.

White, Roger M. *The Structure of Metaphor: The Way the Language of Metaphor Works*. Oxford: Blackwell, 1996.

David Hills (2005)

METAPHYSICS

Physics is the scientific investigation of the fundamental nature of physical being. Metaphysics—at least within that tradition that traces itself back to Aristotle's eponymous treatise—is the philosophical investigation of the even more fundamental nature of being as such. Metaphysics is concerned with the contours of the categories of entity postulated or presupposed by any possible, acceptable, account of the world, whether of the physical world or of any other aspect of the world. The task of metaphysics is to lay out a complete, coherent ontology, embracing all that is necessary to capture the correct account of the world in any of the special inquiries—whether they be empirical, mathematical, modal, or moral.

THE CHANGING METHODS OF METAPHYSICS

Traditionally, metaphysics was practiced as a top-down, a priori discipline, with Euclidean geometry as its model. The metaphysician begins with self-evident principles of a highly general nature, together with appropriate definitions, and proceeds to draw out the necessary consequences.

This approach is clearly exemplified in the work of two prominent eighteenth-century metaphysicians, Gottfried Wilhelm Leibniz and Benedict (Baruch) de Spinoza. Leibniz spun metaphysical gold out of the dross of the *principles of noncontradiction* and sufficient reason: His entire *Monadology* (1965), replete with an infinite collection of possible worlds, with the actual world (the best of all possible worlds) consisting of a myriad of mutually reflecting, simple, mind-like substances. Spinoza was even more self-consciously imitating Euclid, but his conclusions are almost diametrically opposed to those of Leibniz. Spinoza's ontology comprises exactly one substance (God-or-Nature), of which the mental and the physical realms are two aspects, and everything about the one Substance is absolutely necessary—only the actual is really possible.

In the light of its lofty aim, the conflicting conclusions of its practitioners, and their exaggerated claims to have achieved the aim with completeness and certainty, it

is perhaps unsurprising that the discipline of metaphysics, so practiced, has been regularly contested. Empiricists, led by David Hume, have often attacked a priori metaphysics, contrasting its lackluster or conflicting results with the astonishing successes of empirical sciences, on the one hand, and of mathematics on the other.

At the end of the eighteenth century, Immanuel Kant, in response to Hume's critique, attempted a partial vindication of a priori metaphysics. According to Kant, metaphysics can play a legitimate role as handmaid to science and a less straightforward role in upholding ethics. Through an analysis of the cognitive needs of thinking, sensing beings, it can establish the presuppositions of Newtonianism—Euclidean space, absolute time, deterministic causation, and enduring interacting substances obeying conservation laws. In addition, if a metaphysical hypothesis—the existence of God or the freedom of the will—is required for the smooth and effective operation of morality, then that may be legitimately adopted as though it were true, as a postulate of practical rationality. Kant's compromise evidently failed to rein in the metaphysical spirit, his work unleashing a century's worth of metaphysical system-building in an increasingly problematic idealist tradition.

In the late nineteenth century, the appetite for idealist metaphysics began to fade. A realist assault on this tradition was launched by Alexius Meinong, Bertrand Russell, Gottlob Frege, and George Moore, and their style of argumentation, as much as the content of their conclusions, was influential in shaping the twentieth century's more circumspect approach to metaphysics. Rather more radically a group of scientifically minded thinkers—inspired by Ludwig Wittgenstein's *Tractatus* (1922), rallying under the banner of Logical Positivism and brandishing a verificationist criterion of meaning—declared all metaphysical discourse completely meaningless. They argued that sentences that cannot be either verified by observation or proven by pure logic and are not merely beyond our knowing but are strictly speaking, *meaningless*. Echoing Hume, they denied any legitimate space for metaphysics between a posteriori science and a priori logic. The shortcomings of Logical Positivism were rapidly exposed (mostly by its adherents and fellow empiricists such as Karl Popper), but its offspring—ordinary language philosophy—cast over the metaphysical enterprise a pall that did not lift until the 1960s. Metaphysics, cautiously revived by heirs of both movements (albeit with notable differences in methodology detailed below), is once again a flourishing discipline in the early twenty-first century.

Contemporary metaphysics is characterized by a bottom-up approach rather than the traditional top-down approach. The contemporary metaphysician begins with a problem or puzzle, often generated by some basic data or the consequences of such data. The different sources of this basic data characterize two broad traditions. One tradition—championed by Moore, mediated midcentury by philosophers such as P. F. Strawson, Arthur Prior, and Roderick Chisholm and embraced by contemporary philosophers such as Frank Jackson—takes as prime data the deliverances of everyday discourse and commonsense, so called "Moorean facts": for example, that I have two hands; that there is a piece of cheese in my left hand and a stick of chalk in my right; that the chalk and the cheese are distinct things; that cheese and chalk have the same color, and so on.

A different tradition, traceable back through the empiricists (such as Russell and Rudolph Carnap), mediated by Willard Quine, and embraced by philosophers such as John Smart, John Mackie, and David Armstrong, is less impressed with commonsense data. It takes the serious data to be constituted by the presuppositions and deliverances of extraordinarily successful scientific theories: that there is no role for the flow of time in a fundamental account of the world; that the fundamental laws are probabilistic rather than deterministic; that simultaneity is relative to motion; and that space–time may be non-Euclidean. The presuppositions and deliverances of the mathematical disciplines essential to science are also treated as serious data: for example, that there numbers, and an infinite class of such; that there are functions from numbers to numbers and that the infinite class of such functions is vastly bigger than the infinite class of numbers; that there can be no complete axiomatization of mathematical truth, and so on.

The two traditions overlap, of course, as exemplified in the work of prominent metaphysicians like David Lewis. Lewis (1981, 1986) draws extensively on both kinds of data, seeking an ontology compatible with and explanatory of both. However, if that's not possible, the data from the sciences usually trump those of commonsense.

To say that contemporary metaphysics is bottom-up is not to saddle it with a crude inductivism—the fallacious inference of general theories from finite data. The task of the contemporary metaphysician is not so much to *prove* an ontology, either from high-level first principles or from lower-level data, as to *propose* an ontology to accommodate and explain the data, to resolve apparent conflicts by explaining away the appearance of such, or

explain why the data are misleading. The methodology is less like that of pure mathematics and more like that of science—conjecture and refutation—with the difference being the kind of data that require accommodation or furnish a counterexample.

Given a finite amount of data, the number of potentially adequate metaphysical theories seems limited only by the imagination of practicing metaphysicians. To decide between theories we need more than data accommodation. Metaphysicians typically subscribe to *Occam's Razor*—the injunction to refrain from multiplying entities beyond necessity. The *Razor*, read as an endorsement of ontological abstemiousness, is sometimes considered a license to slash entities without regard for a complementary principle—the injunction to refrain from eliminating entities that are necessary. Necessary for what? For accommodating and explaining the data. The upshot of these two principles is, then, that a theory must explain the data; and, of two theories that both explain the data, the theory with fewer ontic commitments is to be preferred.

So, we begin with a domain of discourse—such as mind, or mathematics, or morality—and note that, on the surface at least, it supplies data that posit or presuppose an ontology. Our ordinary mind-talk, for example, presupposes mental states (experiences, thoughts, desires, emotions) along with physical states, and a rich network of causal interactions between them. Mathematics posits numbers, classes, functions, spaces, and a rich array of other abstract objects. Morality presupposes goods and evils, rights and obligations, virtues and vices. But there is often a problem with the entities posited or presupposed. For example, if the mind is something over and above the physical, how can it causally interact with the physical without violating physical laws? And if it is difficult to understand how the mind could affect physical states, it is even more difficult to see how numbers, existing outside space and time, could affect the mind. Whence, then, our knowledge of numbers? Finally, a good would have to be something the mere recognition of which would engage the will, and nothing (some will aver) could do that. The question arises, then, whether such things as minds, numbers, and goods should be counted among the indispensable building blocks of the universe. Is it coherent to postulate them? Are they consistent with the rest of what we know? And even if it is coherent, do we really need them to accommodate the data? Can they be explained, or explained away, in a complete, consistent account of the world? Already with the posing of such seemingly unavoidable questions, the enterprise of metaphysics is up and running.

A SPECTRUM OF METAPHYSICAL APPROACHES

Whenever entities are posited to explain the data provided by some domain of discourse, three broad responses, differing in ontological commitment, are possible. At one end of the spectrum we have *realism*, at the other, *antirealism*; between, we have *determinationism*. Each can be divided into two subcategories.

The realist with respect to a domain accepts both the discourse and the data at face value; affirms the necessity of the entities postulated to explain the data, and adds that the entities really are basic—they are additional to (or "over and above") whatever else there may be. Realism comes in two broad varieties. Transcendent realism locates the posited entities outside the spatiotemporal, causal order. By contrast, immanent realism locates the entities within the spatiotemporal, causal order, typically ascribing them an indispensable causal role.

Most realists about numbers and other abstract entities have been of the transcendent variety, but recently, some number-realists have embraced immanence, espousing a role for abstract entities in the causal network. A transcendent realist theory of value is also usually ascribed to Plato—with the Form of the Good, like all the Forms, eternal and unchanging, existing "over and above" the transient realm of particular contingent beings. It is not hard to find naturalist theories of value that ascribe them a causal role, but as we will see below, there is an important sense in which naturalism about value is not fully realist—it does not posit value "over and above" the natural realm. A version of immanent value realism holds, like Platonic realism, that value is real, that it is something "over and above" the purely natural realm, but adds that value plays a causal role with respect to the motivational states of sensitive beings.

At the other end of the spectrum we have antirealism. The antirealist repudiates the entities in question, maintaining that the discourse that delivers the data is fundamentally misleading. But the data can be misleading in one of two very different ways. The data are recorded and delivered in what appear to be genuine, truth-bearing (or assertoric) claims. What masquerades as truth-bearing claims, however, might really be something else, and the nonassertoric antirealist says just that. Rather than being truth-bearing assertions, they might be expressions of desire, or moves in a language game, or instruments in the derivation of genuinely truth-bearing

assertions. In ethics, nonassertoric antirealism is *noncognitivism*; in the philosophy of science, *instrumentalism*; in the philosophy of math, *formalism*; in the philosophy of mind, the *intentional stance theory*; and so on. Nonassertoric antirealism allows apparent reference to the purported entities while denying there are any such.

The assertoric antirealist, by contrast, accepts that the discourse consists of genuine truth-bearing assertions but rejects those assertions as untrue. In the metaphysics of the mind, this constitutes *eliminativism* concerning the entities. In the metaphysics of morality, it is known as the *error theory*. An increasingly popular variant of assertoric antirealism, especially with respect to abstract entities such as numbers, is *fictionalism*. The fictionalist thinks that the relevant claims are untrue but also thinks that there is a point in continuing the discourse just as though they were.

Between realism and antirealism lie a collection of approaches that share a doctrine of *determination*. Like the realist, the determinationist acknowledges the discourse and the data that suggest the disputed entities. Like the antirealist, however, the determinationist denies that the disputed entities are basic, holding that the truth about the higher level is fully determined by the truth about ontologically more fundamental entities. Determinationism also comes in two varieties, *reductive* and *nonreductive*. The reductionist holds either that the disputed entities are reducible to more basic entities (entity reduction) or that all the facts about the disputed entities are reducible to facts about undisputed entities (fact reduction).

A necessary and sufficient condition for entity education is that the (apparently) higher-level entities are identical to lower-level entities, that properties of the reducible entities are identical to properties of the lower-level entities, and consequently that truths about the reducible entities turn out to be truths about the entities to which they reduce. The reduced entities are nothing but the lower-level entities to which they are reduced. Thus, for example, *logicism* claims that numbers are reducible to classes: The number zero, for example, is simply identical to the empty class; the number one is identical to the class of all singleton classes, and so on. The identity theory of the mind claims that mental states are identical to physical states of the brain. The ethical naturalist claims that moral properties (such as the rightness of actions) are identical to natural properties (such as maximizing expected happiness).

A classic example of reduction without entity reduction can be found in Russell's justly famous theory of descriptions. In his *Principles of Mathematics*, Russell embraced Meinong's theory of nonexistent objects: that there are genuine objects—*possibilia* like the golden mountain and the King of France, and *impossibilia* like the round square—which have a range of features (the golden mountain is made of gold) but which lack the crucial feature of existence. In "On Denoting"—which set the tone for twentieth century analytic philosophy—Russell repudiated this ontology by showing that phrases that apparently denote such possibilia are not really denoting phrases at all. They do not denote particulars, and *they do not denote anything else*. Russell shows us a way of dispensing with nonexistent objects, but unlike the eliminativist, he does not repudiate the data or the discourse that suggest them. Rather, he shows how to translate the data into facts about properties. Nonexistents disappear from Russell's ontology theory, but the data that suggested them are fully accommodated. This is a kind of reduction without being a reduction of the problematic entities. It is a reduction of the *facts* about the purported entities while the entities are repudiated or "analyzed away." Let's reserve the term *fact reduction* for those cases in which every fact about some purported entities is equivalent to a fact about some other entities; there is no entity reduction, and the purported entities are repudiated.

Finally, we have nonreductive determinationism, which has gained considerable currency through the notion of supervenience in philosophy of mind. All determination theories affirm that there can be no difference in one kind of entity without a difference in another, more basic kind. For example, a widely held view is that there can be no difference in the moral without some difference in the natural. Another is that there can be no difference in the mental without some difference in the physical. The higher-level entities are thus determined by the lower-level entities. What is characteristic of a supervenience theory as such is that it posits this determination, does not repudiate the higher-level entities, but also denies the reducibility of the higher-level entities to the lower-level entities.

Supervenience is naturally located between reductionism and realism. The supervenience theorist agrees with the realist and the fact reductionist that the higher-level entities cannot be reduced to the lower-level entities but agrees with the entity reductionist that the higher-level entities are not ontologically basic. There is thus a sense (weaker than the reductionist sense) in which supervening entities are "nothing over and above" the basic entities, but there is also a sense in which the supervening entities, while falling short of the independently

real, enjoy some kind of autonomy denied reducible entities.

Those sympathetic to physicalism (viz., there is nothing over and above the physical) but skeptical that mental state kinds are identical to physical state kinds are attracted to the thesis of the nonreductive supervenience of the mental. Those sympathetic to naturalism (viz., there is nothing over and above the purely natural) but skeptical that moral properties are identical to natural properties are attracted to the nonreductive supervenience of the moral on the natural.

Supervenience theories share considerable common ground with *emergence* theories—lately enjoying something of a revival—and it is interesting that supervenience is popular in domains (such as philosophy of mind and philosophy of biology) where emergence theories also seem a promising compromise between realism and reductionism.

Whether or not there is logical space for nonreductive determinationism has not yet been satisfactorily settled. Like other attempts to forge middle paths between two clear alternatives, a supervenience theory embodies a certain instability, suggesting to some that, in the end, the supervenience advocate will either be forced to embrace the reductionism eschewed or lapse into a form of realism.

One particularly important determination theory is worth singling out for special attention since it has played a pivotal role in the history of metaphysics—namely, determination by the mental (or *mind-dependence*). Broadly speaking, this is *idealism*, and it is a perennially attractive option—indeed, so attractive that idealism has often been taken to be *the* rival to realism. Bishop George Berkeley famously claimed that physical objects are nothing but (are identical to) congeries of experiences. The notorious problem of maintaining the intermittently observed tree in the quad in uninterrupted existence led Berkeley to posit an *omniobserver*, someone to keep a perpetual eye on things. A different response to this problem moves beyond actual experiences to various potential experiences. Physical differences that go undetected may be *detectable* by observers under suitable hypothetical conditions. (*If* you were in the quad, or having in-the-quad experiences, *then* you would have tree experiences.) An idealist could add those conditional states to the determination base. This move, from Berkeleian idealism to phenomenalism, might be a move from entity reduction to fact reduction, or it might be a move from entity reduction to supervenience. If physical objects "disappear" in the final analysis leaving behind the truths that appear to be about them, then we have fact reduction. If physical objects are not identified with anything else, reference to physical objects as genuine entities remains, and the totality of facts about such objects is determined by the actual and conditional facts about experiences, then we have a version of supervenience.

Faced with the fact that actual minds have various cognitive shortcomings, the idealist may also want to tidy up both actual and potential mental states in various ways. The physical facts are held to be determined not by the actual mental states of existing observers but by the mental states that *ideal observers* would have if they were ideally placed. Hence, variations on the basic idealist theme of mind dependence include *positivism*, *ideal limit* theories of truth, and related accounts such as *internal realism*. Many who regard Berkeleian idealism about the physical world as deeply implausible have embraced some version of idealism in other domains—with respect to mathematical entities, theoretical entities, God, possibilities, colors, values, and universals. This three-fold classification helps explain why there is a certain amount of confusion in debates about realism since *antirealism* and *nonrealism* (the disjunction of antirealism and determinationism) and are not usually defined or carefully distinguished.

A PROBLEM IN METAPHYSICS: UNIVERSALS AND PARTICULARS

These patterns of opposition and compromise—realism versus antirealism, with determination seeking a middle way—have played out across the metaphysical spectrum. They find a particularly clear expression, however, in a problem of central concern to metaphysics since its very inception—the problem of universals and particulars, a problem the intrinsic interest of which proves prefatory to its myriad applications. For one's attitude toward universals and particulars has profound implications for one's attitude to a host of other problems—those of abstract entities, change, time, causation, identity, possibility, value, and morality.

Consider a stick of chalk (*A*), a wedge of cheese (*B*), and a chunk of chocolate (*C*). *A* is chalk and *B* is not chalk—it follows that *A* and *B* are not identical. This is an application the principle of the *indiscernibility of identicals*, often associated with Leibniz: If entity *X* is identical to entity *Y*, then anything true of *X* is also true of *Y*. This in turn is a consequence of the *principle of non contradiction*: that no proposition can be both true and false. So much seems straightforward, yet this sort of easy observation intersects with a second, equally obvious fact, to

create a puzzle: *A* is yellow and *B* is yellow. How can *A* and *B*, clearly distinct, be *the same*—that is, yellow? How can they be both the same and not the same? That's puzzling.

Whenever we strike an apparent contradiction (*A* and *B* are the same and are not the same), it is natural to make a distinction. *A* and *B* are not *numerically* the same (they are two) but they are *qualitatively* the same (they instantiate the property *yellowness*). Properties are instantiated by particulars one at a time—they are *monadic* universals. But there are also universals that characterize couples and triples. *Resemblance*, for example, is not a property that particulars have or lack. Resemblance involves pairs—it is a dyadic relation. And *betweenness*, which characterizes triples, is a *triadic* relation, and so on. The problematic data can be accommodated, and the apparent inconsistency explained away, by an ontology that posits two radically different categories of being—particulars and universals—and a relation of instantiation (itself a universal) holding between particulars and universals.

Responses to this two-tiered ontology have traditionally been categorized as either *nominalist* or *realist*, with a third category—*conceptualist*—sometimes thrown rather awkwardly into the mix. The six-fold schema set out above suggests that the space of possibilities is much richer. With respect to universals, one might be antirealist (assertoric or nonassertoric), determinationist (reductive or nonreductive), or realist (transcendent or immanent). Further, any of those positions might be combined with one of the six distinct approaches to particulars. So in all there are thirty-six possible combinations, not just two or three. For example, one might be a transcendent realist about both universals and particulars, or an immanent realist about one and a reductionist about the other, or a reductionist about both universals and particulars (invoking some third category of entity to which both reduce), and so on. Not all of these combinations have been embraced, but many have. For the purposes of illustrating the approaches and the arguments that characterize them, a few of the more commonly held positions are sketched.

REALISM ABOUT PARTICULARS, NONREALISM ABOUT UNIVERSALS

A particular, unlike a universal, lacks the mysterious capacity to be "fully present" in distinct particulars. We feel we understand particulars, perhaps because we are experientially acquainted with them. We do seem to be acquainted with concrete particulars—such as bits of chalk, cheese, and chocolate. But there are other purported entities that strike us as particular rather than universal that are not like these—for example, numbers, classes, propositions, and possibilities. And there are others that are difficult to categorize—such as space and time.

Particularism embraces realism about particulars—there are particulars, and particulars are not reducible to anything more basic—and adds that the *only* basic entities are particulars. It is thus a one-tier ontology with both considerable simplicity and commonsense in its favor. Particularism affirms that anything at all (universals, numbers, classes, possibilities, causation, space, time) is either eliminable or reducible to particulars. *Concrete* particularism is more austere, restricting fundamental being to concrete particulars, the paradigms of which are physical objects. The concrete particularist is a realist about concrete particulars and is typically an immanent realist, assigning particulars both spatiotemporal location and a role in the causal order. There are versions of transcendent realism about particulars although typically transcendent particulars will be abstract rather than concrete (for example, numbers, classes, and possibilities).

In what follows particularism is combined with five different versions of nonrealism about universals. These five accounts have all been called "nominalisms," but the nomenclature is not particularly perspicuous. Sometimes nominalism connotes *concrete particularism*; at others, the broader doctrine of *particularism*. Sometimes it connotes *antirealism* about universals; at others, *determinationism*. *Predicate nominalism*, which holds that there are just particulars and the words (names) we call them by, is perhaps the clearest candidate for the title. *Mereological nominalism* combines concrete particularism with reduction of universals. *Conceptualism*, sometimes called "concept nominalism," is a version of idealism—reduction of universals to mental particulars. *Extensionalism*, sometimes called "class nominalism," is a reduction of universals not to concrete particulars but to classes of concrete particulars. Finally, *resemblance theories* are determinationist and may be of either the reductive or nonreductive kind. The determination base includes concrete particulars and resemblances between them.

PREDICATE NOMINALISM AS AN ELIMINATIVIST PARTICULARISM. The most austere version of nominalism—often called *word* or *predicate nominalism*—is naturally construed as concrete particularism combined with eliminativism about universals. It holds that there are concrete particulars and there are the predicates

(words) we apply to them, and that is all. Simply put: Things are yellow because we call them "yellow." We call things by the same name—*A* and *B* are both called "yellow"—but there is no need to postulate a universal of yellowness instantiated by all things so called. (Distinct people are called "Brian," but we do not postulate a universal *Brianness*.)

The predicate nominalist repudiates universals rather than reducing them to particulars but accommodates the data by means of the following equivalences:

A is yellow ⇔ *A* is called "yellow."

B is yellow ⇔ *B* is called "yellow."

C is brown ⇔ *C* is called "brown."

On the left-hand side we have some Moorean facts, which apparently presuppose concrete particulars (*A*, *B*, *C*) and universals (yellowness; brownness). However, the right-hand side supplants reference to a property with reference to a predicate. Since it is implausible to identify yellowness with the predicate "yellow," the position is plausibly construed as denying the universal yellowness: There is a word "yellow," we call a bunch of things "yellow," and if we call them "yellow," they are yellow (end of story).

Some criticisms of word nominalism are worth sketching because they involve argument kinds that crop up repeatedly in this area. The word nominalist's explanation of the data seems backwards. Distinct things are not made yellow by virtue of being called "yellow;" they are called "yellow" because they are one color, conventionally dubbed "yellow." If we call a person "Brian," then for any other similar person—for example, his identical twin—it is a matter of separate convention whether we also call him "Brian." But if we call a color sample on a chart "yellow," and another sample is qualitatively indistinguishable from that, the original convention covers the second color sample, and we call it "yellow," too.

Secondly, since there are only concrete particulars, words must also be concrete particulars. How many words are there on the following line?

yellow, yellow

There are two answers: two and one. There are two concrete words but only one dictionary entry at issue. We must distinguish between *word-tokens* and *word-types*. On the one hand, word-tokens are spatially located, unrepeatable, concrete particulars. A word-type, on the other hand, has distinct word-tokens as instances. Word-types look very like universals. Elimination of one universal—yellowness—has only been achieved at the cost of accepting another—the word-type "yellow."

To eliminate this word-type, the nominalist might deploy the complex predicate "called 'tokens of the word-type "yellow."'" But this launches a regress, for the same problem resurfaces for applications of this new predicate. Universals are eliminated at one level only to have them pop up in the shape of word-types at the next. Only if word-types at all levels are eliminable in favor of word-tokens will universals be exorcised, but that would require an infinite supply of word-tokens, and there are just not enough to go around. The nominalist might invoke *possible* word-tokens here, but that would launch the nominalist beyond the actual world into the problematic outer space of possibility—not a happy place for nominalists to venture.

Finally, chalk is *called* "yellow," but the cheese is also *called* "yellow." *Calling* involves many distinct particular-word pairs—it seems to be a dyadic relation linking particulars and word-tokens, and relations are universals. If one tries to eliminate this relation in favor of another word, we are again launched on a tiresome regress.

There is an alternative construal of predicate nominalism according to which yellowness is eliminated in favor of a different property—*being called "yellow."* But this would be neither elimination of universals (since both properties have an equal claim to being universals) nor a reduction of universals to particulars. It is simply a proposal to economize within the class of universals itself. As Lewis noted in a different context (possibility), it is not just the number of entities that fall within an ontological category that matters to ontic simplicity, but more importantly, it is the number of basic ontological categories countenanced. If there is something unsatisfactory about the *category* of universal, then whether you admit one, or a million, or an infinite array is immaterial—your attempt to eliminate universals fails.

MEREOLOGICAL NOMINALISM AS ENTITY REDUCTION. Predicate nominalism is an eliminativist version of concrete particularism, but there are reductionist versions too—*mereological nominalism*, for example. Mereology is the theory of the part–whole relation. Some particulars are parts of other particulars. The top of the pen and the body of the pen are both proper parts of the pen, and the pen is the mereological sum of the body and the top. The pen continues to exist even when the top is removed and its parts are separated. Perhaps the chalk and the cheese are also parts of a particular, a spatially scattered whole made partly of chalk and partly of cheese.

Starting with A, B, and C, there are at least four such distinct, albeit overlapping, mereological sums—$(A+B)$, $(B+C)$, $(A+C)$ and $(A+B+C)$—yielding at least seven distinct concrete particulars in total. Note that sameness of parts entails sameness of whole: $((A+B)+A)$, for example, has just two parts, A and B, and so is identical to $(A+B)$. If there are n basic (nonoverlapping) concrete particulars, then, assuming the principle of *unrestricted mereological composition* (that every sum of concrete particulars is itself a concrete particular), there are 2^n—1 concrete particulars. The principle is controversial, but one way of characterizing concrete particularism is this: The principle of unrestricted composition places an upper bound on the collection of entities.

Let $S(F)$ be the mereological sum of all particulars to which predicate F applies—S(yellow) is the sum of all yellow things. Plausibly, something is yellow if and only if it is a part of S(yellow). This suggests the following analysis:

X is $F \Leftrightarrow X$ is a part of $S(F)$

Note that the analysis assumes that for any predicate F, the sum of all things with F is also a concrete particular. However, it does not assume unrestricted composition—for there may be collections of particulars to which no single predicate applies. This mereological analysis, if adequate, might allow us to identify the property F-ness with $S(F)$, a concrete particular.

Mereological nominalism accommodates quite a lot of data about properties such as *water* and *yellow*. But take two related properties such as *water* and *single* H_2O *molecule* that involve the part–whole relation—a quantity of water has parts that are single H_2O molecules. Mereological particularism entails that for something to be water, it has to be a part of S(water), and for something to be a single H_2O molecule, it has to be part of the S(single H_2O molecule). But these two sums are *identical*—sum all quantities of water, and sum all single molecules of water, and you arrive at the same whole. So mereological nominalism entails that to be water just is to be a single molecule of water—and that, unfortunately for the theory, is just false.

CONCEPTUALISM AS IDEALISM ABOUT UNIVERSALS. The attraction of conceptualism is that it reduces universals to something concrete, particular, and also mind-friendly: *concepts*. Concepts are things that the mind can get a handle on whereas universals may be problematic in that respect. The beginning idea is that the cheese is yellow if and only if the cheese falls under the concept of yellow. Quite generally, where $C(F)$ is the concept of F, we have:

X is $F \Leftrightarrow X$ falls under $C(F)$.

We can thus explain the data by appealing to a concept—an apparently familiar mental particular—rather than a mind-independent universal.

Our mental vocabulary (such as belief, thought, desire) suffers a pervasive state-content ambiguity. *My belief that the cheese is yellow* might be my state of believing, as in: "My belief that the cheese is yellow, given my aversion to yellow cheese, made me refuse it." But it might also be the content of my belief, as in: "My belief is just the same as yours: that the cheese is yellow." Our believings are distinct entities, but what we believe here is the same. Believings are mind-dependent (which believings there are depends on who believes what), but the common content of distinct believings does not depend on what you or I believe.

Concepts also suffer the state-content ambiguity—there are concept-graspings and there are the concepts grasped. Concepts in the state-sense are mind-dependent (which concept-graspings exist depends on who is grasping what), but the contents of such graspings do not appear to be mind-dependent.

The conceptualist may eschew concepts as contents of graspings in favor of a myriad different and individual graspings. But your grasping of yellow has something in common with mine. What is that? If we apply conceptualism to this datum then, as with predicate nominalism, we are launched on a regress and there will not be enough particular concept graspings to accommodate all the data.

EXTENSIONALISM: REDUCTION OF UNIVERSALS TO CLASSES. For every predicate F that may or may not apply to a particular, there is the class of all and only the particulars to which F applies in fact, the extension of F—$E(F)$. In our example, class $\{A,B\}$ is the extension of *yellow*, $\{C\}$ the extension of brown. Something is yellow if and only if it is in the class of yellow things. So the following is an apparently necessary equivalence:

X is $F \Leftrightarrow X$ is a member of $E(F)$.

This suggests identifying F-ness with its extension, $E(F)$. (After abandoning concrete particularism, Quine adopted extensionalism.)

Extensionalism is sometimes called "class nominalism," but the postulation of classes marks a real departure from concrete particularism. Classes may be particular, but they are not concrete. Classes are "over and above" the concrete particulars that are their members, and Nelson

Goodman's criterion explains why. Starting with *A*, *B*, *C* there are at most seven concrete particulars. However, there are many more classes. There are seven nonempty classes of particulars. Each one pairs off with a mereological sum—{*A*,*B*} with *A*+*B*, {*A*} with *A*, and so on—but we cannot identify these classes with the corresponding concrete particulars. Classes are individuated by membership: They are identical if and only if they have exactly the same members (the principle of extensionality). So the singleton class {*A*} is a distinct entity from its sole member *A*. A piece of cheese is not a class consisting of a piece of cheese. Quite generally a class is not the mereological sum of its members.

Once we acknowledge classes of classes, the hierarchy of classes "over and above" the concrete particulars, *A*, *B*, *C* explodes into a vast and infinitely intricate structure, one massively exceeding the modest seven-member ontology of mereological sums countenanced by the concrete particularist. The two-membered class {{*A*,*B*},*A*}, for example, is distinct from the two-membered class {*A*,*B*}. The former has a member {*A*,*B*} that the latter lacks. Contrast this with ((*A*+*B*)+*A*) and (*A*+*B*)—the same concrete particular.

Extensionalism, a radical departure from nominalism, thus has plenty of resources—but does it have enough to do justice to the data? Being a chordate is the property of being a heart-bearing animal; being a renate, the property of being a kidney-bearing animal. These are distinct properties. As it happens, these two properties have the same extension. Extensionalism thus entails that they are one and the same property. As generous as the ontology of classes is, it is not generous enough. This *coextension problem* is a classic example of an argument against a reduction thesis. The reduction base is shown to be insufficiently rich to capture all relevant entities.

A different criticism suggests that there are too many classes as well. A universal involves sameness. There are, however, "arbitrary" collections of concrete particulars exhibiting no genuine sameness. If the sameness of *A* and *B* (yellowness) reduces to the fact that both are members of {*A*,*B*}, why does not the fact that *A* and *C* are both members of {*A*,*C*} yield a genuine sameness there? The class theorist could bite this bullet and accept that all classes are universals. (Bullet-biting is a rather common response to recalcitrant data.) A different response would be to block the counterexample by declaring that only certain "natural" classes are genuine or have what it takes to be universals. (This response exhibits an ad hocness that is arguably worse than bullet-biting.)

An explanatory asymmetry argument against extensionalism is often deployed. It is claimed that *A*'s being yellow *explains* *A*'s membership of the class of yellow particulars, not the other way around. This claim contradicts the extensionalist's claim that these facts are really one and the same. The extensionality principle, however, entails that if *X* is a member of a class *C*, then *that very class* could not have lacked *X*—any class *C** that lacks *X* is *necessarily* distinct from *C*. That *A* is a member of {*A*,*B*} is necessary. That *A* is yellow is, by contrast, contingent. No contingent fact explains a necessary fact, and so the argument fails. Even though it fails, it suggests a different argument. It follows, by the indiscernibility of identicals, that a contingent fact (such as *A*'s being yellow) cannot be a necessary fact (such as *A*'s being a member of {*A*.*B*}). But extensionalism entails they are the same fact. Call this the *necessary extension* problem.

The coextension and necessary extension problems are closely related. The reason chordate and renate cannot be identified with the class that is their common extension is that their extensions *might* well have differed from each other. And that presupposes that they have their extensions contingently, not necessarily. Since chordate and renate differ in their possible extensions, one way of modifying extensionalism would be to expand the reduction base to include possibilia. Two accounts have predominated. One embraces possible but nonactual particulars and takes the extension of property *P* to be the class of actual and nonactual particulars that have *P*. (This presupposes that particulars are world-bound—no particular appears in more than one possible world.) Another is to include possible but nonactual *worlds* with common or overlapping domains of particulars. In each possible world *W*, the property *P* has an extension in *W*. The property *P* thus induces a function *F*(*P*) from worlds to extensions—and a reductionist might identify *P* with the function *F*(*P*). These accounts are both reductionist, presupposing different accounts of possibility. They go well beyond the domain of concrete particulars entertained by traditional extensionalists, and they are both known as *intensional* accounts of universals.

Despite the richness of the framework of worlds and functions, however, it may still not be rich enough to capture all the data. Being *triangular* is the property of being a plane figure with three angles. Being *trilateral* is the property of being a plane figure with three sides. Because each logically necessitates the other, these properties induce the very same function from worlds to classes of concrete particulars (or classes of possibilia), and so even these intensionalist accounts render them identical. If

they are not identical, then we need something more discriminating than functions from worlds to classes of particulars with which to identify them.

RESEMBLANCE THEORIES: REDUCTION AND SUPERVENIENCE. An important group of theories claim that property facts are determined by facts about resemblance. A crude version of the resemblance theory invokes paradigms, and resemblance to such paradigms—namely, where $P(F)$ is a specified paradigm of F:

X is $F \Leftrightarrow X$ resembles $P(F)$.

The shortcomings of the paradigm account are numerous. It entails, for example, that the designated paradigm of yellowness is *necessarily* yellow (since everything necessarily resembles itself). It also entails that anything resembling $P(F)$ in any respect at all is F. A far more promising account draws on the notion of *similarity circle*—a class such that all the members of the class resemble every member of the class, and nothing outside the class also resembles every member of the class. Provided that there is sufficient variety in particulars, similarity circles carve out what are, intuitively, the genuine universals without the necessity for privileging any particular. This might be regarded as a reduction of universals to classes of particulars plus resemblances, but it can also be regarded as a supervenience thesis: Properties supervene on a basis consisting of resemblance and the domain of particulars. There can be no difference in properties without some difference in the structure of resemblances—same resemblance structure, same properties.

As Russell famously noted, any account that grounds properties in resemblance faces a problem. Resemblance is a relation between particulars and as such seems to be a universal. It might be considered an ontological saving to reduce myriad universals to one, but as noted in the context of word nominalism, what is important is the number of nonempty ontological *categories*.

This criticism of resemblance theories can be generalized to any attempt to reduce universals to something "else." Suppose we reduce property P to some entity Reduct(P): the class of Ps; the mereological sum of Ps; the concept of P; the similarity circle that corresponds to P, and so on. The reductionist says that for X to be P is for X to bear some suitable relation R to Reduct(P). But the reductionist is then forced either to admit one universal—the relation R—or to apply the theory to R itself, launching an unhappy regress.

Resemblance theorists might employ the tu quoque, charging that the realist faces a similar regress. Assume realism: For X to be yellow is for X to instantiate (I) the universal yellowness (Y). But then, for X to instantiate Y is for a certain triple—X,Y,I—to stand in a relation, I^*. I^* cannot be I. For one thing I, is a dyadic relation, and I^* is a triadic relation. For another, this would involve a relation taking itself as one of its own *relata*. So I^* is distinct from I. We can repeat the argument to obtain a third relation I^{**}, and so on. So the realist is thus as much involved in a regress as any of the reductionist rivals.

The realist might appeal to the category response: The regress is damaging to the particularist because it shows that the category of relation cannot be done away with. That's an internal inconsistency. But the realist about universals does not object to that category being nonempty, and even an infinite class of distinct instantiation relations constitutes no embarrassment for realism as such. Of course, a realist who wants to keep the number of universals down to a small or finite collection might be embarrassed.

Finally, the resemblance theorist may run out of the kind of variety in actual objects required to set properties apart. (Renate and chordate are still coextensive.) To increase the variety and block a coextension objection, the resemblance theorist might take a now familiar tack—embracing relations of resemblance between *possible* as well as actual particulars. As with the related attempts to deflect counterexamples by invoking possibilia, this constitutes a significant and not entirely unproblematic expansion of the reduction base. Certainly it violates the original nominalistic spirit that inspired it.

REALISM ABOUT BOTH PARTICULARS AND UNIVERSALS

One explanation for this apparent failure to eliminate or reduce universals is realism—that universals are neither eliminable, nor reducible, nor supervenient. That this is no proof of realism is obvious—we may not have exhausted all possible alternatives. However, realism about universals conjoined with realism about particulars does explain these failures, as well as providing an explanation of the ubiquitous Moorean facts of predication.

TRANSCENDENT REALISM ABOUT UNIVERSALS. What is often called *ante rem* realism, or Platonic realism, is a transcendent realism: that irreducible universals exist of necessity, beyond contingency in general, and beyond the contingent causal network in particular. One powerful explanatory principle, typically embraced by transcendent realists (Plato perhaps) states that any

meaningful predicate, whether simple or complex, applies to things in virtue of designating a genuine universal. So not only contingent predicates such as "black" and "raven" designate universals, so, too, do predicates that apply of necessity (such as "self-identical"); predicates that apply to nothing (such as "unicorn"); predicates that apply to nothing of necessity (such as "self-distinct"); and finally, not only simple predicates (such as "black" and "raven") designate universals, but so do complex predicates (such as "black raven," "not black," "black or raven," "black if and only if a raven," and so on). Since predicates apply to universals themselves (e.g., *yellowness is a pretty color*), universals are instances of other universals. This *unrestricted transcendent realism* makes the domain of universals a largely a priori affair.

Perhaps the greatest threat to unrestricted transcendent realism is Russell's paradox. Particulars have properties (such as being honest or cowardly), but properties also have properties (honesty is virtue, chalkiness is a universal, a piece of chalk is not a universal), and those properties have properties in turn (virtue is good, being a universal is something all universals have in common, being a particular is a universal not a particular).

By unrestricted realism the two predicates—*being a universal* and *being a concrete particular*—designate two universals, *U* and *P*. All universals have *U* in common. *U* is a universal, and so *U* itself has *U*. *U* is *self-predicating*. However, *P* is not a concrete particular (it is a universal), and so *P* is *non-self-predicating*. Given unrestricted realism, the meaningful predicate *non-self-predicating* designates a universal, *N*. Each universal either has *N* or lacks *N*. If *N* has *N*, then *N* is non-self-predicating—but then *N* does not self-predicate and so *N* lacks *N* (contradiction). If *N* lacks *N*, then *N* is not non-self-predicating; that is, *N* is self-predicating, and so *N* has *N* (contradiction again).

Russellian paradoxes can be constructed for just about any account of universals, including the most austere version of predicate nominalism. (The predicate "short" is called "short," but the predicate "long" is not called "long." Call the former "self-predicating" and the latter "non-self-predicating." Is the predicate "non-self-predicating" called "non-self-predicating"? Paradox ensues.) Russellian paradoxes are thus too pervasive, the realist might claim, for them to be peculiarly damaging to realism. Still, short of embracing paradoxes, the realist has an obligation to deflect them.

Any adequate realist answer to Russellian paradoxes must involve some restriction on the predicates: namely, not every apparently meaningful predicate necessarily designates a universal. Russell's theory of types is a classic restriction. Type theory stratifies entities. Simplifying somewhat: particulars are type-0 entities, properties of particulars are type-1 entities, properties of (type-1) properties are type-2 entities, and so on. A type-0 entity may either have or lack a type-1 property, and a type-1 property may have or lack a type-2 property, but no property either has or lacks a property of the same type. A property is always one type higher than the highest type of entity to which it can be sensibly applied or denied. Thus, the question of whether a universal *P* has or lacks *itself* does not arise. It is neither true nor untrue that *P* lacks *P*. The very attempt to apply *P* to itself is a *category* error (like the attempt to apply the color green to the number 7), and so the predicates "self-predicating" and "non-self-predicating" are literally *meaningless*. (The notion of a category error took on a life of its own long after Russell's theory of types lost the attention of most philosophers.) The paradox is blocked because there are no universals of self-predication or of non-self-predication.

The chief worry about a type theory is that it rules itself out as unsayable—to state it one must violate its strictures. Take the claim *no universal can be applied to, or denied of, a universal of the same type*. This makes a perfectly general claim about all universals. What type does the universal of being a universal (*U*) belong to? It cannot consistently be assigned a level. It is this problem that undergirds the famous theme of Wittgenstein's *Tractatus*, that philosophy consists of things that can be shown but cannot be said.

IMMANENT REALISM ABOUT UNIVERSALS. An important group of restricted realist theories trace their ancestry to Aristotle. Not every predicate picks out a universal, and it is a contingent matter, to be settled a posteriori, what universals there are. *In rebus* realism is a version of immanent realism. It begins with the simple idea that universals exist only in their instances. If a universal is not instantiated, it does not exist. Consequently, a predicate must apply to a particular for it to designate a genuine universal. This *instantiation condition* prohibits universals such as unicornhood. It also rules out various truth-functional combinations of universals. Even if *black* and *raven* are both universals, the predicate "not black and a raven" does not designate a universal (since all ravens are black).

Armstrong, a prominent advocate of immanent realism, places two further conditions on a predicate for it to designate a universal. Firstly, the predicate must apply in

virtue of a *genuine identity* in the particulars. Even if *yellow* and *raven* are both universals, the disjunctive predicate "yellow or a raven" does not pick out a universal, despite being instantiated, because there is no qualitative identity exemplified by a yellow submarine and a black raven. Secondly, Armstrong draws on a condition inspired by the Eleatic Stranger in Plato's *Sophist*, who makes the intriguing suggestion (known as the *Eleatic Principle*) that the mark of being is causal power. Armstrong requires that to be real, a universal must feature in causal laws. (Universals must do some work for their living.)

The Eleatic Principle gives natural science a crucial role in delineating the ultimate constituents of being. Interestingly, it also suggests an immanent realism that denies the instantiation condition. Michael Tooley has argued, within the same immanent realist framework, that there could be properties that play a genuine role in the causal order, because they enter into basic causal laws, but that could remain uninstantiated. If causal power is the mark of the real, it would be hard to deny them ontic standing, but if so, they are not *in rebus* universals—they exist independently of their exemplification by particulars. Clearly, there is wide scope for other, quite different versions of both transcendent and immanent realism about universals.

REALISM ABOUT UNIVERSALS, REDUCTIONISM ABOUT PARTICULARS

If one embraces realism about universals, then, for the sake of simplicity, it would be worthwhile exploring the reduction or particulars to universals. Every particular *X* comes along with a bundle (or a class) of properties, *B*(*X*)—the class of all the properties it has. Further, an object *X* has property *P* if and only if *P* is a member of *B*(*X*). This suggests that we embrace just one entity (the bundle) rather than two (the particular and its bundle). The bundle theory identifies a particular with the bundle of its properties.

This bundle theory faces problems analogous to the reduction of universals to classes of particulars—both too many classes and not enough classes. Firstly, there are too many classes. The class {golden, mountain} does not pick out any actual concrete particular—the golden mountain does not exist. This fact, however, may not be considered entirely undesirable. Meinong famously argued that metaphysics needs to accommodate data pertaining to the nonactual as well as to the actual. To explain the nonexistence of the golden mountain, the golden mountain must be an object with a specific nature, a nature that it possesses of necessity. If being golden and mountainous were contingent properties of the golden mountain, then who is to say that Kilimanjaro is not the golden mountain? The bundle theory thus dovetails nicely with this theory of possible objects.

The bundle theorist still owes us an account of the distinction between concrete existent particulars (Kilimanjaro) and merely possible particulars (the golden mountain). Meinong thought that it is their completeness that sets them apart. Kilimanjaro is complete—for every property, Kilimanjaro either has it or lacks it. The golden mountain is incomplete—it is a mountain, and it is made of gold, but for many properties (e.g., *more than 1 mile high*), it neither has the property nor lacks it. But this will not do—we could specify complete bundles of properties that do not correspond to any concrete particular.

Are there enough bundles of properties to accommodate all particulars, or does the bundle theory face a coextension problem? A bundle theory of particulars entails the Identity of indiscernibles (the converse of the indiscernibility of identicals): If *X* and *Y* are qualitatively identical (share all properties), then *X* and *Y* are numerically identical. This principle would be trivially true provided conditions such as *being identical to X* were genuine properties. But the bundle theorist cannot start with properties that presuppose antecedently given particulars. The bundling properties would have to be purely qualitative. But then it does seem possible for distinct particulars to share all their purely qualitative properties. (Quantum theory, for example, entails that it is possible for two bosons to share their fundamental quantum states—including the state corresponding to location.) That is incompatible with the bundle theory.

An essential property of a particular is a property without which that particular would not exist. It is controversial whether there are essential properties, and if so, which properties of any given particular are essential. However, there is widespread agreement that not every property of a particular is essential to it. At least some properties are such that an item could lose them without going out of existence. The bundle theory suffers an analogue of the necessary extension problem. Classes by their nature are necessary, eternal, and unchanging. So the bundle theory would appear to entail *super-essentialism*: that every property of a particular is essential to it; that if a particular lost a property, it would cease to be.

REJECTING REALISM ABOUT UNIVERSALS AND PARTICULARS

The space of possibilities is not restricted to reduction in one of two directions. Universals and particulars might

both be reduced to some third, more basic, kind of entity. One prominent example of such an approach is *trope theory*.

A trope (this patch of brownness, that instance of sweetness) is a particularized universal—a particular instantiation of a universal by a particular. Put like that, of course, tropes apparently presuppose both particulars and universals. They appear to be nonbasic entities. But it is a characteristic move in metaphysics to take as ontologically basic something that has hitherto been assumed to be derivative and reverse the ontological order.

Tropes have the advantage of incorporating both particularity and qualitative character in their nature and are thus promising building blocks. The proposal is that particulars and universals are both classes of tropes, albeit different kinds of classes: Particulars are classes of *co-located* tropes—tropes occupying the same space and time—and universals are classes of *exactly resembling* tropes. A particular *X* has property *P* just in case the class of co-located tropes that make up *X* overlaps with the class of resembling tropes that constitute *P*.

Trope theory has the advantages of simplicity and comprehensiveness. Further, it avoids the co-extension problem that besets extensionalism. No redness trope is identical to a roundness trope. So even if *redness* and *roundness* always go together (they are always co-located) the class of all roundness tropes is a distinct class from that of redness tropes.

Trope theory retains the necessary extension problem. Classes have their members by necessity, but a concrete particular does not have its properties by necessity. By identifying predication with the intersection of two classes, trope theory implies that all predications are necessary. Again, such problems may be avoidable by invoking possible worlds, but only at the cost of expanding the reduction base to something that makes the resulting reduction rather costly in terms of ontological resources.

THE FUTURE OF METAPHYSICS

The proposed schema for locating metaphysical theories is applicable in all of the various domains of the discipline, and the argument patterns for realism, antirealism, and determinationism bear important similarities across those domains. The basic data might involve claims about time, causation, possibility, the fundamental truths of arithmetic, mental states, spatial relations, value, morality, and so on. Theorists will take the data and lay out an ontology to explain them, or explain them away, either taking the surface ontology at face value, explaining it in terms of something more basic, or occasionally eliminating it altogether.

The modern metaphysician, aware of the underdetermination of theory by data, rarely expects or demands that the arguments conclusively establish any metaphysical proposal. Rather, the metaphysician will examine each metaphysical proposal on its explanatory merits, assessing first its explanatory adequacy with respect to the existing data, searching for new data to test and probe the proposal, and then turning to the more inherently contestable issues of theoretical elegance, economy, and overall coherence with other metaphysical theories. Inevitably, a considerable degree of fallibility and uncertainty remains. Still, that acknowledged, the future of metaphysics is no less secure than the future of science: Human beings can and will continue to probe the fundamental nature of the world right up to the limits of their cognitive abilities. Their doing so will, inescapably, implicate them in the enterprise of metaphysics.

See also Aristotle; Armstrong, David M.; Berkeley, George; Carnap, Rudolf; Chisholm, Roderick; Frege, Gottlob; Gödel, Kurt; Goodman, Nelson; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lewis, David; Mackie, John Leslie; Meinong, Alexius; Metaphysics, History of; Metaphysics, Nature of; Moore, George Edward; Newton, Isaac; Plato; Popper, Karl Raimund; Prior, Arthur Norman; Quine, Willard Van Orman; Russell, Bertrand Arthur William; Smart, John Jamieson Carswell; Spinoza, Benedict (Baruch) de; Strawson, Peter Frederick; Wittgenstein, Ludwig Josef Johann.

The Encyclopedia contains two additional general articles on this subject: *Metaphysics, History of*, and *Metaphysics, Nature of*. It also features the following articles: *Absolute, The*; *Apeiron/Peras*; *Appearance and Reality*; *Arche*; *Being*; *Categories*; *Causation: Metaphysical Issues*; *Chance*; *Chaos Theory*; *Continuity*; *Cosmos*; *Determinism, A Historical Survey*; *Dialectic*; *Emanationism*; *Essence and Existence*; *Eternal Return*; *Eternity*; *Hen/Polla*; *Idealism*; *Identity*; *Infinity in Theology and Metaphysics*; *Logos*; *Macrocosm and Microcosm*; *Materialism*; *Monad and Monadology*; *Monism and Pluralism*; *Naturalism*; *Nature, Philosophical Ideas of*; *Nomos and Phusis*; *Nothing*; *Nous*; *Ontology*; *Panpsychism*; *Personal Identity*; *Personalism*; *Persons*; *Pessimism and Optimism*; *Possibility*; *Relations, Internal and External*; *Solipsism*; *Substance and Attribute*; *Time*; *Unconscious*; *Universals, A Historical Survey*; *Vitalism*; *Voluntarism*; *Why*.

Differing schools of metaphysical thought are represented in the articles *Aristotelianism*; *Augustinianism*; *Cartesianism*; *Hegelianism*; *Neoplatonism*; *Ockhamism*; *Platonism and the Platonic Tradition*; *Scotism*; *Spinozism*; *Stoicism*; *Thomism*.

Bibliography

Aristotle. "Metaphysics." In *The Complete Works of Aristotle: The Revised Oxford Translation*, edited by J. Barnes, 1552–1728. Princeton, NJ: Princeton University Press, 1984.

Armstrong, D. M. *A Theory of Universals*. Cambridge, U.K.: Cambridge University Press, 1978.

Ayer, A. J. *Language, Truth, and Logic*. New York: Oxford University Press, 1936.

Berkeley, G. *Three Dialogues Between Hylas and Philonous*. Oxford: Oxford University Press, 1998.

Bradley, F. H. *Appearance and Reality*. London: Swan Sonnenschein, 1893.

Brentano, F. *Psychology from an Empirical Standpoint*, edited by O. Kraus. Translated by A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge, 1973.

Carnap, R. *Der Logische Aufbau der Welt*. Leipzig: Felix Meiner Verlag, 1928.

Carnap, R. *Meaning and Necessity*. Chicago: University of Chicago Press, 1947.

Carnap, R. "The Elimination of Metaphysics through Logical Analysis of Language." In *Logical Empiricism at its Peak: Schlick, Carnap, and Neurath*, edited by S. Sarkar, 10–33. New York: Garland, 1996.

Chisholm, R. M. *Person and Object: A Metaphysical Study*. La Salle: Open Court, 1976.

Churchland, P. M. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78 (1981): 67–90.

Davidson, D. "Mental Events." In *Experience and Theory*, edited by L. Foster and J. W. Swanson, 79–101. London: Duckworth, 1970.

Dennett, D. C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Duhem, P. M. M. *The Aim and Structure of Physical Theory*. Translated by P. Wiener. Princeton, NJ: Princeton University Press, 1954.

Feyerabend, P. "Mental Events and the Brain." *Journal of Philosophy* 40 (1963): 295–296.

Field, H. H. *Science without Numbers: A Defence of Nominalism*. Oxford: Blackwell, 1980.

Gödel, K. "Some Basic Theorems on the Foundations of Mathematics and Their Philosophical Implications." In *Collected Works. Vol. III*, edited by S. Feferman, et al., 304–323. Oxford: Oxford University Press, 1995.

Goodman, N. "A World of Individuals." In *The Problem of Universals*, edited by I. M. Bochenski, A. Church, and N. Goodman, 13–32. Notre Dame, IN: University of Notre Dame Press, 1956.

Goodman, N. "Predicates Without Properties." *Midwest Studies in Philosophy* 2 (1977): 212–213.

Hegel, G. W. F. *Phenomenology of Spirit*. Translated by A. V. Miller, Oxford: Oxford University Press, 1977.

Jackson, F. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1997.

Kant, I. *Prolegomena to Any Future Metaphysics*. Translated by L. W. Beck. Indianapolis: Bobbs-Merrill, 1950.

Kim, J. *Supervenience and Mind: Selected Philosophical Essays*. New York: Cambridge University Press, 1993.

Leibniz, G. W. *Monadology*. In *Monadology and Other Philosophical Essays*. Translated by P. Schrecker and A. M. Schrecker. New York: Bobbs-Merrill, 1965.

Lewis, D. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.

Lewis, D. *Parts of Classes*. Oxford: Blackwell, 1991.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth, NY: Penguin, 1977.

Maddy, P. *Realism in Mathematics*. Oxford: Oxford University Press, 1990.

Meinong, A. "On the Theory of Objects." In *Realism and the Background of Phenomenology*, edited by R. Chisholm, 76–117. Glencoe, IL: Free Press, 1960.

Montague, R. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven, CT: Yale University Press, 1974.

Moore, G. E. *Philosophical Studies*. New York: Harcourt, Brace, 1922.

Oddie, G. *Value, Reality, and Desire*. Oxford: Oxford University Press, 2005.

Parsons, T. *Nonexistent Objects*. New Haven, CT: Yale University Press, 1980.

Plato. *The Dialogues of Plato*. Translated and edited by B. Jowett. Oxford: Clarendon Press, 1953.

Popper, K. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge, 1963.

Post, J. F. *The Faces of Existence: An Essay in Nonreductive Metaphysics*. Ithaca: Cornell University Press, 1987.

Priest, G. *Beyond the Limits of Thought*. Oxford: Clarendon Press, 2002.

Prior, A. N. *Papers on Time and Tense*. Oxford: Clarendon Press, 1968.

Putnam, H. *Reason, Truth and History*. Cambridge, U.K.: Cambridge University Press, 1981.

Quine, W. v. O. *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953.

Rodriguez-Pereyra, G. *Resemblance Nominalism: A Solution to the Problem of Universals*. Oxford: Oxford University Press, 2002.

Ryle, G. *The Concept of Mind*. London: Hutchinson's University Library, 1949.

Russell, B. "On Denoting." *Mind* 14 (1905): 479–493.

Russell, B. *The Principles of Mathematics*. Cambridge, U.K.: Cambridge University Press, 1903.

Russell, B. *The Analysis of Matter*. London: Routledge, 1927.

Schlick, M. "Form and Content: An Introduction to Philosophical Thinking." *Moritz Schlick: Philosophical Papers. Vol. II*, edited by H. L. Mulder and B. F. B. van de Velde-Schlick. Translated by P. Heath. Dordrecht: Reidel, 1979.

Smart, J. J. C. *Philosophy and Scientific Realism*. London: Routledge & Kegan Paul, 1963.

Spinoza, B. "Ethics." In *The Collected Writings of Spinoza, Vol. 1*. Translated by E. Curley. Princeton, NJ: Princeton University Press, 1985.

Strawson, P. F. *Individuals, an Essay in Descriptive Metaphysics*. London: Methuen, 1959.

Tichy, P. "An Approach to Intensional Analysis." *Noûs* 5 (1971): 273–297.

Tooley, M. *Causation*. Oxford: Oxford University Press, 1987.

Tooley, M., ed. *Metaphysics. Vols. 1–5*. New York: Garland, 1999.

Williams, D. C. "The Elements of Being." *Review of Metaphysics* 7 (1953): 3–18; 171–192.

Van Cleve, J. "Three Versions of the Bundle Theory." *Philosophical Studies* 47 (1985): 95–107.

Wittgenstein, L. *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul, 1922.

Zalta, E. *Abstract Objects: An Introduction to Axiomatic Metaphysics*. Dordrecht, Holland: Reidel, 1983.

Graham Oddie (2005)

METAPHYSICS, HISTORY OF

The word *metaphysics* derives from the Greek *meta ta physika* (literally, "after the things of nature"), an expression used by Hellenistic and later commentators to refer to Aristotle's untitled group of texts that we still call the *Metaphysics.* Aristotle himself called the subject of these texts first philosophy, theology, or sometimes wisdom; the phrase *ta meta ta physika biblia* ("the books after the books on nature") is not used by Aristotle himself and was apparently introduced by the editors (traditionally by Andronicus of Rhodes in the first century BCE) who classified and cataloged his works. Later, classical and medieval philosophers took this title to mean that the subjects discussed in the *Metaphysics* came "after the things of nature" because they were further removed from sense perception and, therefore, more difficult to understand; they used Aristotle's frequent contrast of things "prior and better known to us" with things "prior and better known in themselves" to explain why the treatises on first philosophy should come "after the books on physics." In medieval and modern philosophy "metaphysics" has also been taken to mean the study of things transcending nature—that is, existing separately from nature and having more intrinsic reality and value than the things of nature—giving *meta* a philosophical meaning it did not have in classical Greek.

Especially since Immanuel Kant *metaphysics* has often meant a priori speculation on questions that cannot be answered by scientific observation and experiment. Popularly, "metaphysics" has meant anything abstruse and highly theoretical—a common eighteenth-century usage illustrated by David Hume's occasional use of *metaphysical* to mean "excessively subtle." The term has also been popularly associated with the spiritual, the religious, and even the occult. In modern philosophical usage *metaphysics* refers generally to the field of philosophy dealing with questions about the kinds of things there are and their modes of being. Its subject matter includes the concepts of existence, thing, property, event; the distinctions between particulars and universals, individuals and classes; the nature of relations, change, causation; and the nature of mind, matter, space, and time. In the eighteenth and nineteenth centuries *metaphysics* was used broadly to include questions about the reality of the external world, the existence of other minds, the possibility of a priori knowledge, and the nature of sensation, memory, abstraction, and so on. In present usage these questions are included in the study of epistemology.

THE CLASSICAL PERIOD

The history of metaphysics in Western philosophy (taking "metaphysics" in the contemporary sense) began with speculations by the Ionian cosmologists in the sixth century BCE about the origin of the physical universe, the matter or stuff from which it is made, and the laws or uniformities everywhere present in nature. Our knowledge of these early cosmologists comes mostly from Aristotle and other classical authors; the main figures were the Milesians (Thales, Anaximander, and Anaximenes), Pythagoras, and Heraclitus.

PARMENIDES. The beginning of metaphysics, however, is most conveniently dated from Parmenides (fl. c. 475 BCE), since some of the typical characteristics of metaphysics as a distinct philosophical inquiry are present in, or at least suggested by, his surviving writings. These characteristics are, first, the conception of philosophy as an attempt to understand the universe by means of a logical investigation that is a priori, appealing to meanings of terms rather than to the evidence of the senses. This method is in contrast to the method of natural science, which relies on sense perception. Second is a more or less explicit use of very general principles viewed as sufficient to arrive at a true account of reality. Such principles were, for example, noncontradiction and something like a principle of sufficient reason, which is expressed in Parmenides' poem: "Also, what necessity impelled it, if it did

spring from Nothing, to be produced later or earlier? Thus it must Be absolutely, or not at all." Philosophy was therefore conceived as a deductive science like mathematics. Third is the paradoxical contrast between apparent reality and true reality and the association of the truly real with singleness and unchangingness.

Of these features of Parmenides' writings, the first is fundamental; it can be taken as a defining characteristic of metaphysics. Like the natural scientist, the metaphysician gives an account of the universe; unlike the scientist, he does not base his account on observations and experiments, at least not on any special observations and experiments made for the purpose. His account is based primarily on analysis of concepts; if he does appeal to the evidence of the senses, he appeals to something generally familiar, not to new evidence he is adding to knowledge. Parmenides himself apparently believed he had done all that could be done by way of a philosophical account of the universe. His account consists in pointing to what he believed were the logical consequences of saying "It is." He dismissed everything else either as poetic imagery with no claim to truth or as empirical science; he indiscriminately referred to both as opinion. His position was not naive; it is not easy to see how a metaphysician can give an account of reality based on logic alone unless reality in some sense has the features of necessity and vacuous generality belonging to logical truths. And doctrines similar to Parmenides' logical monism have frequently reappeared in the history of metaphysics—for example, in Neoplatonism, in Benedict de Spinoza, and in nineteenth-century Hegelianism. There is more than a superficial resemblance between Parmenides' Being, the Neoplatonists' One, Spinoza's God or nature, and G. W. F. Hegel's Absolute as understood by a metaphysician like F. H. Bradley. Perhaps the underlying reasoning is that recognizing that metaphysics gives an account of the world based on analysis of concepts rather than on empirical evidence, these philosophers have felt that logic alone should be sufficient basis for making assertions about the world; since whatever is logically true is thought to be necessarily and always true, they have concluded that the world itself must be unchanging and in some sense necessarily what it is.

LATER PRE-SOCRATICS. Parmenides apparently believed he had said all that a metaphysician could say about the world. Accordingly, his followers Melissus and especially Zeno are more critical than constructive—a trait shown by many later metaphysicians who are more often concerned to demonstrate what they take to be logical failures in the ordinary or scientific understanding of reality than to give a positive account of reality. We learn from Plato's *Parmenides* that Zeno's paradoxes of motion were meant to support Parmenides' system by showing contradictions in the ordinary concept of change. (When does the arrow move? Not now, because at any given instant it is in one place and hence not moving; not at some other time, because if it is moving, it must be moving now.)

Parmenides' general effect, however, was to interest philosophers in following what seemed to be the logical implications of their assumptions. An example is Anaxagoras, who apparently argued from the assumption that reality is many and changing to the conclusion that the things we ordinarily call real are composed of unendingly smaller parts similar to the whole things, that "all things are together," that "everything contains a part of every other thing," and that although there are rearrangements of things, nothing is ever really created or destroyed. Like his contemporaries Empedocles and the atomists Leucippus and Democritus, Anaxagoras did rely on observation and experiments to give an account of nature, but the surviving fragments suggest that his cosmology was arrived at largely by a priori reasoning in the way Parmenides' was, although the resulting account of reality is the opposite of Parmenides' account. And in the same way that something like Parmenides' logical monism is repeated in Neoplatonism, in Spinoza, and in nineteenth-century Hegelianism, something like Anaxagoras's logical pluralism is repeated in Gottfried Wilhelm Leibniz's theory of monads and Bertrand Russell's logical atomism. The common feature of this kind of system is that on logical grounds reality is described as composed of elements viewed as the limit of an unending process of division; the least parts of things are, so to speak, real infinitesimals—things smaller or simpler than any given thing one can mention. The atomism of Leucippus, Democritus, and, later, Lucretius is, by contrast, primarily a physical theory. These thinkers believed that the existence of atoms can be shown empirically; their atoms have finite sizes and such recognizable physical properties as shape and motion and, perhaps, weight, and the theory anticipates Galileo Galilei and Isaac Newton rather than Leibniz and Russell.

PLATO. In Plato's *Phaedo* Socrates is made to say he once studied Anaxagoras but gave up this study and all empirical investigations of nature, deciding instead to "have recourse to conceptions and examine in them the truth of realities." Anaxagoras, Parmenides, and others had also had recourse to conceptions in contrast to the evidence of the senses; what is new in the *Phaedo* is the theory of

Ideas or Forms, which historians of philosophy sometimes ascribe to Plato (c. 427–347 BCE) and sometimes to Socrates himself. For Plato, at least, ideas exist independently of the things we see and touch; moreover, they are considered the source of existence of things we see and touch, somewhat as a man is the cause of his shadow or of his reflection in a mirror or a pool of water. Popularly, Plato's metaphysics means the theory of Ideas in this sense, and in this way the theory has had a great influence in the history of thought. Plato's own evaluation, however, was considerably more critical than that of many of his followers. The theory of Ideas in this form is presented in the *Phaedo* as a hypothesis that cannot be known to be true; in the *Parmenides* its logical weaknesses are pointed out; in the *Timaeus* it is used as part of a "probable" or "likely" cosmology. Nevertheless, Plato does consistently argue for the existence of mind or soul as a kind of entity distinct from, and in some sense prior to, physical objects. This thesis is developed, notably in the *Phaedo,* where the theory of Ideas is used as a step in proving the immortality of soul, in the *Phaedrus,* and in Book X of the *Laws.* In these contexts Plato argues that since bodies cannot move themselves (apparent self-motion is reduced to one part's moving another) whereas soul can, the ultimate source of observed motions must be soul or mind. In the *Laws* this argument is used to prove the existence of the gods, who are understood as sources of observed motions and changes in the visible universe.

Plato's technical contributions to metaphysics are contained in the difficult later dialogues, especially the *Parmenides* and *Sophist.* Both dialogues purport to be a criticism of Eleatic philosophy, by Parmenides himself in the *Parmenides* and by an "Eleatic stranger" in the *Sophist.* In the *Parmenides* Parmenides is represented as illustrating the method of dialectic by scrutinizing his own hypothesis that "the One exists" and deducing the logical consequences both of asserting and of denying this hypothesis. The point is that what follows depends on how the hypothesis is understood—in particular, on how one understands unity and existence. If, for example, unity is thought to be in no way compatible with plurality, a thing that has unity can hardly have anything else. Thus, it cannot have spatial extension, for it would then have a right and a left, an up and a down. The more straightforward *Sophist* classifies philosophers into materialists and idealists according to their criteria of reality. A general criterion of reality as power is suggested, and a number of concepts of equal generality with that of being are introduced and discussed—sameness, difference, rest, and motion. The apparent paradox in negation is explained by distinguishing absolute nonbeing (*A* does not exist) from relative nonbeing (*A* is non-*B*) or otherness and by distinguishing the existential is (*A* exists) from the is of predication (*A* is characterized by *B*). In the *Timaeus* the generic concepts are used in the mythical account of the construction of the physical universe by a godlike artisan using an ideal pattern as a blueprint.

ARISTOTLE. Aristotle (384–322 BCE) is indirectly the source of the term *metaphysics*; he is also the source of a systematic list of metaphysical issues, a technical language in which these issues are stated, and a metaphysical system that has had followers down to the present and has proved immensely fruitful. In part, the importance of this system has been in serving as an object of criticism, although this function has been served by Plato as much as by Aristotle and Aristotle himself illustrates Plato's importance as an object of criticism in the history of metaphysics.

The problems of "first philosophy," or metaphysics, listed by Aristotle in books Beta and Kappa of the *Metaphysics* are partly about metaphysics itself: Does its subject matter include all the basic concepts and assumptions of all the special sciences? Does it include the principles of logic? Is there metaphysical knowledge in contrast to opinion? These questions ask, in effect, whether metaphysics is a superscience proving the assumptions made by the special sciences and also the assumptions it itself uses—whether, in short, it is a logically self-contained body of knowledge contrasting with the logically incomplete special sciences. This concept of metaphysics was held, for example, by René Descartes, but on the whole Aristotle rejected this view. Metaphysics is less the capstone of a hierarchy of sciences than a discussion of problems left over by the special sciences. Physics, for example, assumes there is motion, but it is not part of the metaphysician's job as Aristotle saw it to prove this assumption; at most, he should explain it or defend it from criticism. Aristotle thought of metaphysics as explaining things we already know to be true rather than as giving reasons for the assumptions we make in the sciences and everyday life, thereby providing the underpinnings of science and common sense.

Some of the problems of metaphysics listed by Aristotle are questions about the kinds of things there are. In addition to physical objects perceived by the senses, do such abstractions as Plato's Ideas or the mathematician's numbers, points, lines, and so on also exist? Are all existing things particulars, or do universals like man or whiteness exist, too? Do particulars of the same kind have anything in common, and if so, what and how? Are phys-

ical objects something more than the material parts that compose them, and if so, what?

For Aristotle, however, the most fundamental questions of metaphysics concerned the concepts of being and unity. Are being and unity properties of things (since everything both is and is one thing), or are they entities or substances of some kind (as Parmenides seemed to have thought)? If being and unity are things in their own right, what kind of things are they? These questions are suggested by Plato's *Parmenides* and *Sophist.* Aristotle's answers are his most important contribution to metaphysics. In the *Sophist* Plato suggested a general definition of being as power but gave little by way of an explicit analysis of this sense of being, which does not correspond to the use of the word in ordinary language. Such an explicit analysis is the center of Aristotle's metaphysics; his contribution can be summarized as the view that although there are many ways in which things are and are one (and there are therefore many senses of being and unity) and although these ways are irreducibly distinct, they nevertheless depend on one basic kind of being. Being is neither an attribute nor a thing and cannot therefore be defined in the ways *triangular* or *horse* can be defined. But we can pick out a basic sense of being, illustrated in such statements as, "This is a horse" or "This is a man," and show how the other senses of being depend on it. "Being a horse," "being a man," and, in general, "being an *X*" in the basic sense of being means to have attributes and therefore to be a subject of thought and discourse without in turn being an attribute of something else; "being a horse" is not, for Aristotle, an attribute of some more basic subject of thought and discourse. Primarily, what there is, is this horse, this man, and so on when we are speaking of an individual; secondarily, what there is, is horse, man, and so on understood as species or kinds of things. Qualities, dates, locations, motions, relations, and the like are attributed to the things that exist in the basic sense; they themselves do not have independent existence and "are" only in a derivative and borrowed sense of being.

Aristotle's analysis of being is the heart of his metaphysics; it is not the whole of it or the part most stressed by his later followers. What is often referred to as Aristotle's metaphysics is his account of the universe. Roughly, it states that there are a large but finite number of things that for the most part (with exceptions such as the sun, the only thing of its kind, and biological "mistakes" resulting from mutation and crossbreeding) belong to definite kinds—for example, plant and animal species. In most cases the individual members of these kinds or classes are born and die, but the classes themselves do not change. Some things—for example, the stars—exist forever and apart from uniform motions do not change at all. There is an ultimate prime mover that is the source of all observed motion and change but is itself completely immaterial and therefore completely motionless and changeless. This set of ideas is in the *Metaphysics,* and the pluralism and some theory of natural kinds do follow from Aristotle's analysis of being. But the theory of prime movers and the Unmoved Mover is also in the *Physics* as a scientific—that is, demonstrable—account of the physical universe; it is not therefore a true part of his metaphysics, which is dialectical (arguing from common opinion and logic) rather than scientific.

The central chapters of the *Metaphysics* elucidate and defend the claim that such commonsense things as this horse, this man, and so on are the fundamental subjects of discourse. Aristotle upheld this claim against (1) the view that the ultimate material parts of things are the ultimate subjects of discourse (so that "This is a horse" would be understood as "These material elements have horselike attributes"); (2) the view that Platonic Ideas are the ultimate subject of discourse (where "This is a horse" is understood as "The horse is exemplified by these sensible qualities"); and (3) the view that the basic sense of being is illustrated in, for instance, "There is a horse in the barn"—the view according to which "there is" means "it is true that" or "it is a fact that." For Aristotle to be is to be an individual, and the being of a thing is primarily its nature or identifying features rather than the fact that it is. Aristotle hardly even recognized the sense of being involved in such sentences as "There are good men, and there are wicked men," which can be read as "Among all the things that are, some are—that is, have the identifying features of—good men; others are wicked men." Such sentences suggest that what exists primarily are featureless particulars, which can be referred to collectively as "the things that exist," not commonsense things.

In general, the question "What is being?" became for Aristotle "What is an individual?," a horse, a man, a house, and so on being understood as paradigms of an individual. And, positively, the central argument of the *Metaphysics* is that an individual is primarily the distinguishing features by which we identify and classify it. Aristotle himself believed that these classifications are learned through experience; he was a realist in the sense that he thought the groups and classes of things are there to be learned by observation and are not simply mental constructions. Therefore, there is a sense in which we learn empirically what being is. But metaphysics is not

itself an empirical study of being; Aristotle did not, for instance, think of metaphysics as a science of high-level generality describing the properties that all beings (individuals) have.

Aristotle's *Metaphysics* in its present form—and there is no reason to think it ever had a very different form—is barely readable in large stretches. Other parts read like outmoded astronomy; still other parts read like rather tedious lexicography. The devastating criticism of Plato is largely borrowed from Plato himself. However, the *Metaphysics* gives a surprisingly coherent set of answers to the questions it raises, and the questions themselves are those that metaphysicians still ask.

NEOPLATONISM. The Neoplatonists in the late classical period were metaphysicians of great power and originality. They were also of great importance in the development of metaphysics since they formed a link between ancient and medieval philosophy. The main figure of this movement, Plotinus (c. 204–270), associated metaphysics with mysticism and personal asceticism. The mystical and religious side of his philosophy was stressed by his disciple and editor Porphyry (c. 232–304), and such later Neoplatonists as Iamblichus and Proclus gave a further religious and even occult and superstitious emphasis to the movement. But the intellectual power of the movement is shown in as late a philosopher as Ancius Manlius Severinus Boethius (c. 480–524), and through Boethius Neoplatonism had a very strong influence on medieval philosophy and, therefore, indirectly on modern philosophy.

Plotinus. Plotinus's philosophy is a paradigm case of a metaphysical system according to one common conception of metaphysics. It asserts the unreality or half reality of the things of everyday experience; the illusory character of change, motion, and even space and time; the superior reality of soul or mind over matter. It conceives of goodness and intelligence as substantial things and stresses personal mysticism and an ascetic way of life. The line of thought by which Plotinus arrived at this position is not easy to follow, but, briefly, it seems to have been somewhat as follows. Whatever is, is one thing (even a collection of things is said to "be" only when counted as one thing—a collection); the answer to the question "What is being?," understood as a request for a description of being, is therefore unity or singleness. But unity or singleness cannot be described any further, although a direct, intuitive experience of it is in some sense possible. Since being is equivalent to unity and since things can have unity to a greater or lesser degree, we can speak of degrees of being. Although unity is itself ineffable, it does duplicate itself in a kind of descending series of things—in goodness and intelligence—in a lesser way in disembodied spirits, in a still lesser way in human souls, least of all in physical objects and their properties and relations. The emanation of successively less real things from unity is to be understood in a logical rather than a physical sense. Speaking accurately, unity or singleness (the One) is not a cause at all, although it can be described metaphorically, for example, as an inexhaustible fountain of being bringing existence to all the things that are by its continuous overflow. Plotinus's writings are full of these metaphors, but he recognized them as metaphors, and the underlying position is rigorously argued, granting the not implausible identification of being with unity or singleness.

Plotinus's line of thought begins with the assumption that being and unity are properties that things have—properties of utmost generality, to be sure, but still properties in the same way that black or being four-legged are properties of a horse. Combined with this seems to be the Platonic assumption that properties are not simply modifications of particulars or ways that particulars exist; properties are entities in their own right that particular things instance or exemplify. The first of these two assumptions is clearly made in the *Isagoge,* Porphyry's short introductory treatise on Aristotle's *Categories.* In Porphyry's account—and in this account he is presumably expressing a typically Neoplatonic point of view—the theory of categories or types of predication is a theory of kinds of predicates: genus, species, difference, property (that is, essential property), and accident. These kinds of predicates (the predicables) are distinguished from individuals. But even expressions designating individuals are predicates of a sort according to Porphyry; such expressions as "Socrates," "this man here," and "this thing here" are attributes, differing from the predicables because they are "only said of a single thing" whereas the predicables "are said of several things." The distinction is between attributes belonging to several things and attributes belonging to only one thing. But of individuals themselves, in contrast to attributes, nothing is said; they can apparently be characterized only indirectly, as the ultimate subjects of predication.

This account of predication makes the distinction between thing and property peripheral to metaphysics. The important distinction is between relatively less general and relatively more general attributes, culminating in the most general attributes, being and unity. Porphyry spoke of substance as "the most general genus" and in a

sense the only real genus, since unlike animal, for example, which is a genus relative to man but only a special case relative to "living thing," substance is not itself a special case of some higher genus. Neoplatonic metaphysics is largely an analysis, similar to Plato's *Parmenides,* of these ultimate genera; the main force of Plotinus's writings is the argument that the ultimate genera cannot be described in any ordinary way but are in some sense manifest in lower orders of being. Neoplatonism thus easily lends itself to religious interpretation; in the late classical world it actually was a theological system associated with a religious way of life competing with Christianity.

THE MIDDLE AGES

Porphyry's *Isagoge,* translated into Latin by Boethius in the sixth century, gave philosophers some basic tools and stimulated speculation on two questions in particular: (1) What is a thing considered just by itself, as a bare existent, apart from all its attributes? (2) Do attributes exist (or subsist) separately from human thought and discourse and from the things that are said to have attributes? The first question, implicit in Porphyry's account of predication, is roughly the problem of distinguishing essence from existence, what a thing is from the fact that it is. The second question (really, group of questions) was explicitly raised but not answered by Porphyry; it is the problem of universals much discussed throughout medieval philosophy.

For Aristotle the contrast between what a thing is and the fact that it is, is at best peripheral to metaphysics. Aristotle recognized that the question "Does *X* exist?" is distinct from "What is *X*?," but he attached no metaphysical importance to the distinction. Particular questions of the form "Does *X* exist?" are decided by sense perception or by proof; there is no general metaphysical question about the nature of existence ("thatness") in contrast to essence ("whatness"). The metaphysician is concerned with what things are rather than with their existence or nonexistence. Aristotle's position was that what things are—that is, their being—is primarily what is contained in their definitions; the definition of a thing describes its essence, which is equivalent to its species (the traits that identify it as the kind of thing it is) which is in turn identified with its genus, differentia, and essential properties. But when, as in Porphyry, genus (mammal), difference (solid-hoofed), species (horse), property (neighs), and accident (gray) are indiscriminately called attributes of the thing itself, it is natural to ask what it is that has these attributes or what it is that gives this collection of attributes an actual rather than a merely possible existence.

The problem of universals dominated metaphysics in the early Middle Ages; it was discussed by metaphysicians from Boethius in the sixth century to Roscelin and Peter Abelard in the twelfth century. The main philosophical tradition during this period was the Augustinian tradition, represented by Boethius himself, John Scotus Erigena (c. 810–c. 877), St. Anselm (1033–1109), William of Champeaux (d. c. 1120), St. Bonaventure (c. 1217–1274), and many others. This tradition favored realism; species and genera like horse and animal were thought to exist not only apart from human thought and discourse (epistemological realism) but also apart from particular horses and animals. Species and genera were regarded as paradigms, archetypes, or exemplars of particular things; as such, they exist in the mind of God and are used by him as models in creating nature. As in St. Augustine and Plato, the fundamental contention is that particulars cannot be recognized and identified as one of a general type unless we first have independent knowledge of the type; the inference is that these general types must exist apart from, and in some sense prior to, the particulars exemplifying them.

St. Anselm's proof of God's existence (anticipated by St. Augustine), has had an important history in its own right; it is also an illuminating example of Christian Platonism in the early Middle Ages. The argument cannot be appreciated apart from its context of religious meditation, but it can be picked out and studied (as it has been by philosophers to this day) as a kind of supreme test case of Platonic (or Neoplatonic) metaphysical assumptions. Briefly, the argument is that (1) we have a concept of a supreme being (a being "than which nothing greater can be conceived") so that (2) the Supreme Being "exists in the understanding." Since (3) it is greater to exist in reality than merely in the understanding, it is contradictory to say the Supreme Being exists only in the understanding; hence, we can infer that (4) the Supreme Being does exist in reality. Kant's objection seems decisive. The existence (as contrasted with the concept of existing) of the Supreme Being cannot be a part of our concept of the Supreme Being. If it were, our concept would be the Supreme Being, not its concept. But the argument seems inevitable if one assumes, as the Neoplatonists did, that existence is an attribute that things have and, in consequence of having it, are, as things are red in consequence of having the attribute redness. Combined with the assumption that attributes have an independent existence, this line of thought leads to the conclusion that existence or being is itself an existing thing; the existence of things in nature is thought of as being due to their receiving a part of the inexhaustible thing, being, some-

what as an illuminated object receives its light from a source of illumination. Furthermore, it seems to follow that existence must itself necessarily exist as an analytic consequence of what it is (just as "Redness is red" seems to state an analytic necessity). Given these assumptions, the Ontological Argument for God's existence, as Kant later called it, is at least a strong temptation; the argument has had a history identical with the history of logical monism in metaphysics, from Parmenides to Hegel and beyond, as well as a close association with Christian theology.

REVIVAL OF CLASSICAL PHILOSOPHY

Although the realism-nominalism controversy occupied philosophers in the eleventh and twelfth centuries, new ways of thinking in metaphysics were being prepared by translations of Greek and Arabic texts into Latin, especially translations of Aristotle and his Arabian commentators. In the early Middle Ages there was very little firsthand knowledge of the Greek philosophers. Plato's *Timaeus, Phaedo,* and *Meno* were known, but the important later dialogues, including *Parmenides* and *Sophist,* were not. The Greek texts had been preserved, however, and, especially after the capture of Constantinople by crusaders in 1204, were slowly recovered in the West. In the thirteenth century William of Moerbeke made a literal Latin translation of Proclus's *Commentary on the Parmenides*; the commentary contained the text of the *Parmenides* through the first hypothesis, thereby giving philosophers some firsthand knowledge of that important dialogue.

Aristotle was even less known and understood in the early Middle Ages. Only his logic, the text of *De Interpretatione,* and the other logical treatises in Neoplatonized versions through Boethius were known. As late as the thirteenth century, two Neoplatonic texts—the "Theology of Aristotle" (actually a compilation from Plotinus's *Enneads,* IV–VI) and the *Liber de Causis* (a work based on Proclus's *Elements of Theology*)—were wrongly attributed to Aristotle. However, Aristotle's writings had been translated into Syriac by Nestorian Christians in the fifth century and from Syriac into Arabic in the ninth century; Latin translations of Arabic texts were made in the twelfth century and directly from Greek texts by Robert Grosseteste and William of Moerbeke in the thirteenth century. By the end of the thirteenth century most of Aristotle was translated into Latin and was generally available to philosophers. In effect, Aristotle was a new philosopher who appeared on the scene and dominated it as if he were a contemporary; the *Metaphysics* was the stimulus for such metaphysicians as Albert the Great, St. Thomas Aquinas, John Duns Scotus, William of Ockham, and others in the thirteenth and fourteenth centuries.

THOMAS AQUINAS. Thomas Aquinas's metaphysics is an attempt to explain the distinctions between essence and existence, necessary and contingent existence, and particulars and universals, using the language and much of the metaphysical outlook of Aristotle. For Thomas commonsense things like horses and houses do exist in a literal and straightforward sense apart from human observers and also apart from God and paradigms of things in the mind of God. The existence of these commonsense things is not an attribute that they receive from outside; it is not like the light the earth receives from the sun. The existence of finite things in nature is an intrinsic act of existing that these things exercise. But Thomas also held that the ordinary things we experience exist contingently in the sense that their existing is not an analytic consequence of what they are; it is not something they do by nature. There must therefore be a cause (in a metaphysical, not a physical, sense of "cause") of their existence; this must be a necessary being, identified with God, who exists by his own nature. Contingent beings, like horses and houses, are obviously contingent because being composed of matter, their existence is finite—they begin to exist and cease to exist. Matter also accounts for the individuality of things; things that are identical insofar as what they are, or, in other words, things that have the same nature, are still different things because the matter of which they are composed is different. God, on the contrary, is immaterial and, hence, one and unchanging. Thomas, like the Neoplatonists, associated finitude, contingency, plurality, and change with matter. He differed from the Neoplatonists chiefly in his view that finite things—in particular, human persons—exist in their own right (by virtue of a delegated power, as it were) and do not merely participate in the existence of a higher order of being. In this view Thomas agreed with Christian theology and was close to Aristotle.

DUNS SCOTUS. John Duns Scotus (c. 1266–1308) seems to have agreed with Thomas that being is not an attribute or a thing in some sense shared by all the things said to be. On the other hand, he criticized Thomas's contrast of essence with existence, arguing that whatever we are aware of must be an essence in some sense, including even individuality or "thisness," which he treats as an attribute of individuals ("this horse here"), distinguishing them from indeterminate beings ("a horse" or "the horse" in general).

WILLIAM OF OCKHAM. William of Ockham (c. 1285–1349) held that general or indeterminate expressions like "a horse" or "the horse" do not correspond to general beings either in the mind or in reality but refer indifferently to individual horses. He was therefore conventionally called a nominalist in contrast to Duns Scotus, a realist. But William of Ockham's main point seems to be that logical distinctions between universal, particular, and singular are not distinctions between kinds of things—not an enumeration of what there is—but are, rather, ways of referring to the one and only one kind of thing that does exist—namely, the commonsense things we encounter in everyday experience. For this reason William was probably closer to Aristotle's own view than either Thomas or Duns Scotus; unlike them his explicit aim was to state Aristotle's original position as accurately as he could. But William's successors—notably, John of Mirecourt and Nicholas of Autrecourt—pushed William's views in a direction that anticipated Hume and even twentieth-century logical positivism. We can talk meaningfully only about what we are acquainted with through the senses, and we are acquainted only with particulars, so that all discourse about things refers ultimately only to particulars. The existence of a particular is never an analytic necessity or an analytic consequence of the existence of some other; hence, all meaningful statements about things are only probable.

DESCARTES TO KANT

DESCARTES. The revival of metaphysics in the seventeenth century begins with René Descartes (1596–1650), who has been traditionally considered the originator of modern philosophy. The ideas most commonly associated with Descartes are not original with him. In St. Augustine's writings can be found the *cogito ergo sum* argument and the view that our own existence is the ultimate certainty since we can be certain of it while the existence of all other things is in doubt. The argument that nothing less than God could have produced the idea of God in the human mind can also be found in St. Augustine. The Ontological Argument had a famous history in the Middle Ages, and the view that physical objects have only geometrical attributes of shape and motion was held by early Greek atomists. The concept of mind as a substantial thing more or less externally attached to the body is hardly original with Descartes. But to say this is to say only that Descartes used a good deal of material from old ruins in his work of "building from the foundation" in metaphysics in order "to establish a firm and abiding superstructure in the sciences."

Descartes was most original in his conception of philosophical method and philosophical truth. No metaphysical assertion is to be believed unless (1) it is understood with the kind of clarity and distinctness that mathematical propositions have and (2) its truth is either so intrinsically obvious that, like the postulates of geometry, it cannot be doubted or it is proved with the same rigor with which theorems are proved in geometry. Descartes's philosophy can be viewed in large part as an effort to reduce the second criterion to the first—that is, to show that at least in the case of metaphysical propositions, if we understand them clearly and distinctly, we are thereby certain of their truth. These claims made for his or any other metaphysical assertion were revolutionary and most influential. As Descartes and his followers understood them, they amounted to a demand that metaphysics be scientific, understanding by the word *scientific* being subject to a kind of rigorous intellectual discipline best illustrated in mathematics and the exact physical sciences.

SPINOZA. Benedict de Spinoza (1632–1677), following one interpretation of Descartes's demand for clarity and distinctness in metaphysics, thought of metaphysics as a deductive account of the universe to be developed from a few definitions—notably, the definition of substance as a being that requires nothing outside itself to be or to be conceived—and self-evident assumptions. His inferences are that there must logically be one and only one substance, uncreated and everlasting; there are an infinite number of attributes of the one substance, only two of which, thought and extension, are known to us; attributes are faces of the one substance—self-contained ways of describing it—rather than properties inhering in it the way we commonly think of colors as inhering in physical objects; the universe, described in terms of the attribute extension, is a mechanical system in which all happenings are links in a chain of physical causation; an equally complete causal determinism holds when the universe is conceived in terms of the attribute thought.

LEIBNIZ. Gottfried Wilhelm Leibniz (1646–1716) was also a follower of Descartes in the sense that he agreed with the demand for a rigorously scientific metaphysics and for clear and distinct ideas in contrast to scholastic verbiage. But while Leibniz agreed that metaphysical assertions are true if clearly and distinctly understood, he interpreted this to mean that metaphysical truths (and truths of reason generally, in contrast to contingent truths of fact) are logically necessary; their denial involves a self-contradiction. Leibniz understood clarity and dis-

tinctness in a logical rather than a psychological sense; for him "the true mark of a clear and distinct notion of an object is the means we have of knowing therein many truths by *a priori* proofs." And we know a truth by an a priori proof when "by the help of definitions or by the resolution of concepts" we "reduce" it to an explicit tautology of the form "*A* is *A*" or "*A* is not non-*A*."

Leibniz's metaphysical system is, in effect, an effort to get a clear and distinct idea of the universe in his own rather special sense of clarity and distinctness. And his technical writings in metaphysics consist largely of a series of somewhat different a priori proofs of a number of metaphysical assertions, including the following: There are an infinite number of substances, each of which is logically complete in that it contains in some sense all the properties it ever has exhibited or will exhibit; no two substances exhibit exactly the same properties ("identity of indiscernibles"); a complete description of any one substance would be a description of the entire universe "from a point of view"; space and time are relations among things, not things in their own right; the appearance of causal relations between things is illusory, reflecting God's deliberate prearrangement rather than any real influence exerted by one thing on another. In proving these assertions, Leibniz relied on a principle of sufficient reason stating, in effect, that there is always a rational explanation for a fact. But the principle of sufficient reason is not really a description of the universe for Leibniz. What it really expresses is the idea that in principle any truth can be given an a priori proof; the underlying thought is that when any statement is understood with perfect clarity and distinctness, it will be seen to be an explicit tautology.

LOCKE. Spinoza and Leibniz are usually grouped with Descartes as rationalists, as contrasted with British empiricists, represented in the seventeenth century by John Locke (1632–1704). But in an important way Locke, too, was a follower of Descartes; he was also mainly interested in replacing scholastic jargon with clear and distinct ideas and opening the way for the sciences. Locke's main contribution to metaphysics lies in his critical discussion of substance and essence. Descartes had laid it down as an indubitable common notion that "nothing is possessed of no attributes, properties, or qualities," so that "when we perceive any attribute, we therefore conclude that some existing thing or substance to which it may be attributed, is necessarily present." Locke did not deny that this is a valid inference; he does not question the distinction between thing and property. But he asked what we know (or, as he phrased it, "What is our idea") of a thing beyond its attributes, powers, and so forth. His answer was that we have no clear and distinct idea at all; we know only what the common notion itself says—namely, that if there are attributes, there must be something underneath that has them. We have no clear idea what is underneath or what "underneath" means in this context. We know only the attributes, powers, and so on (indiscriminately called qualities by Locke) of things, not the things in themselves.

Here, however, Locke was criticizing only the notion of substance as substratum underlying properties. And this is a concept of substance minimized by Aristotle and never stressed by metaphysicians. Thomas Hobbes, for example, argued that the accidents of body, such as shape or hardness, are the very "manner of our conception of body." To ask for a description of body apart from its accidents would be, for Hobbes, a senseless request. Locke's more important and original criticism concerns the notion of essence—the notion of what a thing is in contrast to what it is made of, how big it is, its location, its age, and the like. Locke argued at length that the distinction is a useless one; the question "What is *X*?" can be answered only by enumerating *X*'s observed properties, and (most important) we cannot see any logical necessity for the coexistence of just these and not some other combination of properties. We do not therefore have any knowledge of real essences except in cases where we ourselves construct the thing in question, as in mathematics. Locke reasoned, roughly, that we know the attributes and powers of things only through the simple sense impressions we have of them. Since, for the most part at least, there are no noticeable necessary connections between simple sense impressions, we cannot explain why things appear as they do but can only describe how they do appear. Locke never denied there is a reason for things' having just the attributes and powers they have and not some others, but he denied our ability ever to have clear and distinct ideas of these reasons. The effect of Locke's view is to deny the possibility of metaphysical knowledge when metaphysics is conceived of in the way Francis Bacon, for example, conceived of it, as a very general but still empirical and even experimental study of the formal causes of things, as distinguished from natural science, which studies material and efficient causes.

BERKELEY AND HUME. Locke never questioned the distinction between ideas of things and the qualities in things that cause ideas, and he thought we have at least a "relative and obscure" idea of a thing in contrast to its qualities. But George Berkeley (1685–1753) questioned both distinctions, partly on grounds of fact but more

especially on grounds of a general theory of meaning. For Berkeley the grammatical distinction between subject and predicate has no counterpart in a distinction between things and properties; we can talk meaningfully only about what we are acquainted with, and we are acquainted only with individual colors, sounds, tastes, and the like. Since these individual colors, sounds, and tastes have characteristics that are admittedly mental, such as pleasantness and painfulness, and are relative to the human observer in various ways, Berkeley concluded we can talk meaningfully only about mental entities or, as he called them, following the usage of Descartes and Locke, ideas in the mind. In this way Berkeley arrived at phenomenalism (things exist exactly as they appear to the senses) and idealism (things exist only as objects of conscious perception; their being consists in being perceived). Berkeley was not thoroughgoing in these positions; he thought it meaningful to talk about other minds and about God even though we cannot directly perceive such phenomena.

These qualifications, however, were swept aside in the thoroughgoing phenomenalism of David Hume (1711–1776). Hume criticized the notion of a mind as distinguished from the ideas said to be in the mind for the same reasons that Berkeley criticized the notion of matter. According to Hume, the notion of existence itself signifies nothing beyond a greater or less degree of force and vivacity attaching to sense impressions and mental images. Our beliefs in the continuous existence of physical objects and the presence of causal connections between them are explained as effects of habitual associations of ideas for which there is, strictly speaking, no evidence. Although Hume is usually and correctly called an empiricist in contrast to speculative metaphysicians like Leibniz or Spinoza, there is a sense in which he was as much a rationalist as his contemporary Christian Wolff. Hume assumed that the ultimate subject of thought and discourse must be something we are directly conscious of, that we are directly conscious only of individual sensations (or their more or less faint copies), and that whenever we can discriminate one sensation or feeling from another, these exist separately and hence count as different things. These assumptions amount to a theory of empiricism, but they are not themselves empirical assertions. Nor, on the other hand, are they necessary truths in Leibniz's sense—propositions whose denial involves a self-contradiction. In effect, they demonstrate how Hume understood Descartes's demand for clarity and distinctness in metaphysics and are analogous to Leibniz's principle of sufficient reason, which expressed his understanding of the same demand. For Leibniz clarity and distinctness meant, in the end, reduction to an explicit tautology; for Hume clarity and distinctness meant, in the end, reduction to directly verifiable assertions about sensations and feelings.

KANT. By the time of Hume's death, in 1776, the difficulties and ambiguities in Descartes's program for metaphysics were apparent. Cartesianism inspired both the speculative constructions of Spinoza, Nicolas Malebranche, Leibniz, and others and the critical and—at least, on the surface—increasingly skeptical philosophies of Locke, Berkeley, and Hume. This, at least, was the view taken by Immanuel Kant (1724–1804). It led Kant to ask whether metaphysics could be scientific—whether metaphysical knowledge is even possible and if not, how the questions that gave rise to metaphysics in the past could be answered. In discussing these problems, Kant made a very penetrating analysis of metaphysics as a discipline and a set of assertions and as a "human propensity"; Kant's contribution, apart from his own system, was to raise questions about what metaphysical assertions, as distinguished from scientific assertions, are, about the sense in which they claim truth, and about the grounds on which they are to be believed or disbelieved.

From Kant's point of view the history of metaphysics (insofar as metaphysics had claimed to be a science) had been a story of dogmatism versus skepticism. Dogmatists like Leibniz have held that metaphysics can, on the basis of purely logical or conceptual considerations, answer with absolute certainty questions about the origin of the universe, the existence of God, and the immortality of the soul. "Dogmatists," as Kant used the word, can be materialists, panpsychists, or dualists, monists or pluralists. What they share is a confidence that a metaphysician can give an account of the nature of reality using a priori reasoning. Skeptics, on the other hand, are empiricists; for them there are no universal and necessary truths of fact and reasoning alone, in contrast to observing and experimenting, is of no use whatsoever in answering questions about the existence or natures of things. For Kant this alternating dogmatism and skepticism was the effect of alternating overconfidence and lack of confidence in the abilities of the human mind. Accordingly, his critical philosophy is an effort to show what human knowledge is like and what its limits must necessarily be.

Dogmatic metaphysics in Kant's sense is not mere ad hoc speculation; it is an understandable and correctable misuse of basic concepts. The dogmatic metaphysician rightly sees that we actually use concepts like substance (in contrast to accidents) or causation (in contrast to

mere succession). He also correctly saw that we are a priori certain of such things as the irreversibility of time or the impossibility of two physical objects' occupying the same space. But he uncritically concluded that we have a power other than sense perception of knowing what things are like, whereas the true conclusion is that we ourselves determine in advance what any object of knowledge must be like. The questions we ask about things and the answers we look for are determined by our own a priori forms of perceiving (space and time) and of judging (every attribute must belong to some substance, every event must have some cause and so on). Mistaking these a priori forms of perceiving and judging for descriptions of things-in-themselves, the dogmatic metaphysician is led to speak of ultimate subjects and first causes. In Kant's view these speculations are misguided and even meaningless. But metaphysical ideas, such as an ultimate subject or a first cause, do have a regulative use in encouraging us never to be satisfied with what we actually know at any given time. And Kant did not infer that the beliefs that metaphysicians have tried to prove—beliefs in personal immortality or in the existence of God—are illusory. These beliefs are not like belief in perpetual motion machines; they can be justified and can even be supported by arguments—but by moral arguments, not speculative arguments. Dogmatic metaphysics can thus be explained and even in a sense vindicated. It cannot be taken seriously as a source of knowledge, however.

METAPHYSICS SINCE KANT

Kant's own metaphysical position was idealistic. Aristotle's categories reappear somewhat altered in Kant's philosophy as forms of judgment. The most immediate and obvious effect of Kant's thought can be seen in the idealistic systems of his younger German contemporaries and successors, Johann Gottlieb Fichte (1762–1814), Friedrich Schelling (1775–1854), Arthur Schopenhauer (1788–1860), and, above all, Georg Wilhelm Friedrich Hegel (1770–1831).

HEGEL. Among the idealists, however, it was Hegel whose metaphysical outlook has probably had more general intellectual influence than that of any other single recent philosopher. Kant's critical idealism assumes a clear-cut contrast between what is given in experience (sense impressions) and the forms we use to arrange and interpret what is given. In general, Kant assumed a clear distinction between what is directly perceived and what is inferred or constructed by the mind. Hegel's absolute idealism consists largely in denying this contrast; for him the underlying notion of a plurality of separately existing particulars, uniquely located in space and time (conceived as containers in which things are unambiguously placed), was a false, even a logically incoherent notion. He appears to have arrived at this conclusion from the assumptions that things-in-themselves cannot be distinguished meaningfully from things as we know them and that things as we know them gradually take shape in our consciousness and become defined only in contrast to other things. On this basis he concluded that all things shade off into their opposites and that the connections between things we establish in thought are as much a part of the things as their so-called inherent properties. Hegel was thus led to the monistic position that there is only one kind of substance and only one truly substantial entity. His idealism is an evolutionary pantheism in which the only self-subsistent reality is spirit; it contrasts not only with materialism in the traditional sense but with any metaphysical position associating reality with some kind of hard definiteness.

Outside of philosophy proper Hegel's influence was apparent mainly in inspiring a view of things as phases of a living and growing history; institutions, languages, ideas, even philosophies themselves, were seen as quasi-living and even quasi-personal phenomena whose histories were to be sympathetically grasped and appreciated rather than appraised by themselves on the basis of a priori standards. This widely held view has been encouraged by Hegel's absolute idealism, in which reality is associated with self-expression and all-inclusiveness, not with given things or facts. Within philosophy Hegel's influence can be seen in the many evolutionary idealisms of the nineteenth and early twentieth centuries. It can also be seen in the more rigorous and critical thought of Hegelians like F. H. Bradley (1846–1924) and J. M. E. McTaggart (1866–1925). Bradley in particular stressed the negative side of Hegelianism, finding logical antinomies in the ordinary concepts of things, properties, relations, causation, and space and time. McTaggart, on the other hand, attempted to rephrase Hegelianism as a clear and straightforward speculative system. This tradition is continued by such contemporary metaphysicians as Brand Blanshard.

METAPHYSICS AND PRAGMATISM. Largely through the influence of German idealism and especially of Hegel, metaphysics in the nineteenth century generally meant a priori cosmology and, in particular, an idealist cosmology contrasted and even opposed to the alleged mechanistic and materialist assumptions of science. Auguste Comte's positive—that is, nonmetaphysical—philosophy did not attack metaphysics as such; it attacked speculative philos-

ophy as a way of providing substitutes for religious beliefs. Popularly, metaphysics was associated with religion, idealism, and spiritualism and opposed to science, which was associated with empiricism and materialism. But this concept of metaphysics, although still popular, was only a temporary alignment in the history of metaphysics and was strongly challenged even in the nineteenth century.

A notable example is the American philosopher C. S. Peirce (1839–1914). Peirce was a Hegelian to the extent that he believed there are no self-identical particulars that can be unambiguously located or identified. Reality is indeterminate both in the sense that it is characterized by novelty and unpredictability and in the sense that things are not just what they are but shade off continuously into other things; reality is an evolutionary process that is in some sense rational. But for Peirce this outlook is required by reflection on experience and the sciences, metaphysics itself being an observational science whose job is "to study the most general features of reality and real objects" and whose backward condition is due chiefly to the fact that "its leading professors have been theologians." Science and experience force us to give up the concept of definite, unambiguous facts and fixed a priori assumptions; science is a community of inquirers sharing methods and a kind of moral and intellectual discipline rather than a body of knowledge or a set of assumptions (as Kant, for example, had thought). Metaphysics for Peirce was an attempt to describe how reality must seem to men imbued with science; reality is what will eventually be agreed on by the community of inquirers; general laws and relations among things are real since these, rather than particular facts, are the objects of scientific research. Peirce's concept of metaphysics influenced John Dewey (1859–1952), and largely through Dewey it has had considerable importance in recent American philosophy. Like Peirce, Dewey hoped metaphysics could be a descriptive account of generic traits exhibited in all experience.

LOGICAL POSITIVISM. The mainstream of metaphysics in the nineteenth and early twentieth centuries was idealistic; metaphysicians responded to Kant by constructing systems meant to extend or deepen Kant's critical idealism. But another response was to question dogmatic metaphysics more profoundly than Kant himself. This more radical questioning was begun by such nineteenth-century philosophers of science as Ernst Mach (1839–1916), who criticized the notion that general concepts of science (for example, force) described unobserved entities or that scientific laws are more than convenient formulas for summarizing observations.

This line of criticism has been most forcefully and systematically carried out by twentieth-century logical positivism. For the logical positivists metaphysics has a special meaning; an assertion is metaphysical if it purports to make a statement of fact but fails to do so—and therefore fails to have a meaning—since no observations count as evidence for or against it. This special use of metaphysics should be understood in the context of the belief of logical positivists that traditional questions of metaphysics do have a point, but a point that traditional formulations of the questions obscure. They are not questions about things at all but about language—in particular, about the types of words and sentences and the logical vocabulary needed to express the findings of the sciences.

The hope of some logical positivists was that if traditional metaphysical questions were translated into questions about the language of science, the answers would be immediately and clearly seen. If, for example, "Does nonbeing exist?" is phrased as "Are sentences of the form '*X* is not an *F*' ever true?," the answer is obviously "Yes." But it became increasingly clear that in the construction of languages expressing the findings of the sciences problems analogous to traditional metaphysical problems occur. For example, some positivists suggested that sentences such as "Two plus two equals four" owe their truth to linguistic usage rather than to a necessary connection between things, perceived by reason, as past metaphysicians often assumed. Critics pointed out, however, that since it is an empirical fact that we use language as we do, the substitution of "true by virtue of linguistic convention" for "necessary truth" threatens to make "Two plus two equals four" a merely empirical statement. Thus, a distinction is needed between what we merely do not say and what our language will not allow us to say. This does not, of course, mean that nothing was gained over traditional metaphysics, but it does mean that the achievement of logical positivism has been to elucidate or reconstruct traditional metaphysical issues rather than give a method for easily solving them. Accordingly, logical positivists now tend to accept metaphysics in its conventional sense, as the name of a legitimate part of philosophy, along with the special use of metaphysical to refer to pseudoinformative assertions that in reality are meaningless.

ORDINARY-LANGUAGE PHILOSOPHY. The logical positivists were strongly influenced by Bertrand Russell's

view that much of traditional metaphysics resulted from a superficial and hasty analysis of ordinary language as well as by the view of Russell and Peirce that past failures of metaphysicians were due to a narrowly restricted logic that prevented them from analyzing ordinary language correctly. The notion that traditional metaphysics resulted from a superficial understanding of ordinary language has been developed independently of logical positivism (although sometimes popularly confused with it) by Ludwig Wittgenstein, Gilbert Ryle, and a large number of contemporary British and American philosophers. Like the logical positivists the ordinary-language philosophers agree that traditional metaphysical questions are in some sense intelligible but need to be radically reformulated; unlike the positivists they are not concerned with rephrasing them as questions about the language of science. They want to show, rather, how metaphysical questions can be solved (or dissolved) by exhibiting the less obvious but essential presuppositions that give linguistic expressions the meanings they actually have in ordinary discourse. Positively, ordinary-language philosophers use linguistic analysis (for example, naming, referring, describing, and so on) to deal with traditional metaphysical issues, and like logical positivists they accept metaphysics in this positive sense as a legitimate area of philosophy.

PHENOMENOLOGY AND EXISTENTIALISM. Both logical positivism and ordinary-language philosophy could be viewed as extensions of Kant's criticism of dogmatic metaphysics; they both sharply contrast with Hegelianism and, in general, with the more or less speculative metaphysical systems inspired by Kant's idealism. A third major development in nineteenth-century and twentieth-century metaphysics, represented by phenomenologists and existentialists, agreed with Hegelians that metaphysics is not an observational science in any ordinary sense and also agreed with analytically minded philosophers that a priori reasoning cannot establish anything about the nature of reality. Accordingly, these philosophers sought new and unconventional ways of experiencing or encountering reality. This response is shown by more conventional metaphysicians like Henri Bergson (1859–1941), who stressed the inability of spatializing and static conceptual thinking to represent correctly the reality of immediate experience, especially its temporal flow, or by Alfred North Whitehead (1861–1947), who stressed imaginative feeling and emotion as a way of gaining access to the inner natures of things. Phenomenologists hold that common sense and science presuppose a more primitive experience that can be grasped by a deliberately naive description of how things actually appear to us; existentialists argue that the subject of metaphysics is a reality that cannot be described in an emotionally neutral way but is in some sense possessed or encountered in personal commitment to a cause or in facing the certainty of one's own death. Phenomenology and existentialism have been combined by systematic philosophers like Martin Heidegger and Jean-Paul Sartre, whose systems attempt to express an intuitive understanding of time, contingency, and particularity as these are experienced in human life.

PHILOSOPHICAL ANALYSIS. In the English-speaking world at least, the most original and important contributions to metaphysics at the present time come from analytic philosophers largely influenced by logical positivism or ordinary-language philosophy. These philosophers see the present situation in metaphysics somewhat as Aristotle did when he reviewed the history of metaphysics up to his own time. In a sense, Aristotle thought, everything had been said, but in a sense nothing had been said because the early philosophers were vague and inarticulate. Contemporary metaphysicians, however, are in a better position to review and analyze the history of their subject than was Aristotle, partly because the history itself is so much richer and partly because contemporary insights make the work of past metaphysicians more intelligible.

See also Albert the Great; Analysis, Philosophical; Anaximander; Anaximenes; Anselm, St.; Aristotle; Augustine, St.; Bergson, Henri; Berkeley, George; Blanshard, Brand; Boethius, Anicius Manlius Severinus; Bonaventure, St.; Bradley, Francis Herbert; Comte, Auguste; Descartes, René; Dewey, John; Duns Scotus, John; Erigena, John Scotus; Existentialism; Fichte, Johann Gottlieb; Galileo Galilei; Grosseteste, Robert; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Heraclitus of Ephesus; Hobbes, Thomas; Hume, David; Iamblichus; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Leucippus and Democritus; Locke, John; Logical Positivism; Mach, Ernst; McTaggart, John McTaggart Ellis; Melissus of Samos; Metaphysics, Nature of; Neoplatonism; Newton, Isaac; Ontological Argument for the Existence of God; Parmenides of Elea; Peirce, Charles Sanders; Phenomenology; Plato; Plotinus; Porphyry; Pragmatism; Proclus; Pythagoras and Pythagoreanism; Russell, Bertrand Arthur William; Ryle, Gilbert; Sartre, Jean-Paul; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Thales of Miletus; Thomas Aquinas, St.; Whitehead,

Alfred North; William of Champeaux; William of Moerbeke; William of Ockham; Wittgenstein, Ludwig Josef Johann; Wolff, Christian; Zeno of Elea.

Bibliography

GENERAL WORKS

Copleston, F. C. *A History of Philosophy,* 7 vols. London: Burns, Oates and Washbourne, 1946–. A careful, detailed work requiring close study; probably the most comprehensive and scholarly general history available. Includes excellent bibliographies and sections on metaphysics in chapters on individual philosophers.

De George, Richard T., ed. *Classical and Contemporary Metaphysics: A Source Book.* New York: Holt, Rinehart and Winston, 1962. A book of readings emphasizing contemporary authors but also including selections from classical texts.

Gilson, Étienne. *Being and Some Philosophers,* 2nd ed. Toronto: Pontifical Institute of Mediaeval Studies, 1952. A challenging essay on the history of metaphysics argued with great subtlety by an outstanding Roman Catholic philosopher.

Kaufmann, Walter, ed. *Philosophic Classics,* 2 vols. Englewood Cliffs, NJ: Prentice-Hall, 1994. A comprehensive sourcebook for the history of philosophy, with some excellent introductory essays.

Lovejoy, Arthur O. *The Great Chain of Being.* Cambridge, MA: Harvard University Press, 1936. Traces some main themes of metaphysics from the Greeks to the nineteenth century; a wide-ranging essay in the history of ideas rather than the conventional history of philosophy.

Smith, T. V., ed. *Philosophers Speak for Themselves,* 4 vols. Chicago: University of Chicago Press, 1956. A collection of readings from Thales to Kant.

Whitehead, A. N. *Adventures of Ideas.* New York: Macmillan, 1933. Suggestive and sometimes profound nontechnical essays on movements in science and philosophy by a great twentieth-century metaphysician.

PRE-SOCRATICS

Burnet, John. *Early Greek Philosophy,* 4th ed. New York, 1930. Lucid and authoritative; contains translations of some important fragments.

Freeman, Kathleen. *Ancilla to the Pre-Socratic Philosophers.* Cambridge, MA: Harvard University Press, 1948. English translations of the surviving pre-Socratic writings collected in Hermann Diel's monumental *Fragmente der Vorsokratiker.*

PLATO

Cornford, F. M. *Plato and Parmenides.* New York: Liberal Arts Press, 1957. Translations with running commentaries of Parmenides' *Way of Truth* and Plato's *Parmenides*; for advanced students.

Cornford, F. M. *Plato's Theory of Knowledge.* New York: Liberal Arts Press, 1957. Translations with running commentaries of Plato's *Theaetetus* and *Sophist.*

Lynch, William F. *An Approach to the Metaphysics of Plato through the Parmenides.* Washington, DC: Georgetown University Press, 1959. An attempt to see the *Parmenides* as a straightforward assertion of "basic positions in Platonic metaphysics."

Ryle, Gilbert. "Plato's *Parmenides.*" *Mind* 48 (1939): 129–151, 302–325. Article by a leading contemporary British ordinary-language philosopher; suggests a modern reading.

Vlastos, Gregory. "The Third Man Argument in the *Parmenides.*" *Philosophical Review* 63 (3) (1954): 319–349. Technical but clear and acute analysis of the third-man argument against the theory of ideas in the first half of the *Parmenides.*

ARISTOTLE

Aristotle. *Metaphysics.* Translated by Richard Hope. New York: Columbia University Press, 1952. With an analytical index of technical terms; a useful edition for advanced students. Index contains Greek terms with Latin equivalents.

Brumbaugh, Robert S. "Aristotle's Outline of the Problems of First Philosophy." *Review of Metaphysics* 7 (3) (1953). A brief but very helpful analysis of the organization of the *Metaphysics.*

Owens, Joseph. *The Doctrine of Being in the Aristotelian Metaphysics.* Toronto: Pontifical Institute of Mediaeval Studies, 1951. Highly technical but original and forcefully argued interpretation of the *Metaphysics.*

NEOPLATONISM

Bréhier, Émile. *The Philosophy of Plotinus.* Translated by Joseph Thomas. Chicago: University of Chicago Press, 1958. Readable sympathetic account of Plotinus and his school by a modern French historian of philosophy.

Porphyry. *Isagoge.* Translated by J. Tricot. Paris: J. Vrin, 1947. A translation into French of the complete text, with introduction and notes.

EARLY MIDDLE AGES

Anselm. *Proslogium and Monologium.* Translated by S. N. Deane. La Salle, IL: Open Court, 1962.

Malcolm, Norman. "Anselm's Ontological Arguments." *Philosophical Review* 69 (1) (1960): 41–62. A closely reasoned analysis of Anselm's argument and medieval and modern criticisms.

LATE MIDDLE AGES

Copleston, F. C. *Aquinas.* Harmondsworth, U.K.: Penguin, 1955. Clearly written, relatively simple introduction to Thomas; contains a useful bibliography.

Moody, Ernest A. *The Logic of William of Ockham.* New York: Sheed and Ward, 1935. A scholarly study of William of Ockham and medieval logic and metaphysics; difficult but searching and persuasive.

Thomas Aquinas. *Concerning Being and Essence.* Translated by George C. Leckie. New York: Appleton, 1937. Difficult but important treatise on a basic distinction in Thomistic metaphysics.

SEVENTEENTH CENTURY

Aaron, R. I. *John Locke.* London, 1955.

Balz, A. G. A. *Descartes and the Modern Mind.* New Haven, CT: Yale University Press, 1952.

Beck, Leslie J. *The Method of Descartes.* Oxford: Clarendon Press, 1952.

Descartes, René. *Selections.* Edited by R. M. Eaton. New York, 1927.

Hampshire, Stuart. *Spinoza.* Harmondsworth, U.K.: Penguin, 1987. An excellent exposition and criticism with a general concluding chapter on the nature of metaphysics.

Jackson, Reginald. "Locke's Distinction between Primary and Secondary Qualities." *Mind* 37 (1929): 56–76. Explains and criticizes Locke's contrast between ideas in the mind and qualities in things and between primary and secondary qualities.

Leibniz, Gottfried Wilhelm. *The Monadology and Other Philosophical Writings.* Edited by Robert Latta. London, 1948.

Leibniz, Gottfried Wilhelm. *New Essays concerning Human Understanding.* Translated, with notes, by Alfred G. Langley, 3rd ed. La Salle, IL, 1949. A standard edition containing some very interesting but little-known essays by Leibniz in an appendix.

Locke, John. *An Essay concerning Human Understanding.* New York: Dover, 1958.

Whitehead, A. N. *Science and the Modern World.* New York: Macmillan, 1925. Early chapters contain a brilliant critique of seventeenth-century and eighteenth-century science and philosophy.

EIGHTEENTH CENTURY

Kant, Immanuel. *Prolegomena to Any Future Metaphysics.* Edited by Lewis White Beck. New York: Liberal Arts Press, 1951. A good introductory text by a leading authority on Kant; contains a useful bibliography.

Smith, N. K. *Immanuel Kant's Critique of Pure Reason,* 2nd ed. New York, 1933. A classic translation of Kant's major work.

Warnock, G. J. *Berkeley.* Harmondsworth, U.K.: Penguin, 1953. An exceptionally good introductory work; highly critical of some of Berkeley's leading arguments.

METAPHYSICS SINCE KANT

Ayer, A. J., ed. *Logical Positivism.* Glencoe, IL: Free Press, 1959. A collection of the most important and influential papers in the logical positivist movement; includes an unusually complete bibliography covering the entire range of twentieth-century analytic philosophy.

Findlay, J. N. *Hegel.* New York: Humanities Press, 1964. A very clearheaded sympathetic exposition of Hegel's system.

Passmore, John. *A Hundred Years of Philosophy.* London: Duckworth, 1957. A useful and very well written survey of philosophy since 1860; emphasizes metaphysics and theory of knowledge in English-speaking countries.

Royce, Josiah. *The Spirit of Modern Philosophy.* Boston and New York, 1926. Lectures on post-Kantian philosophy, especially nineteenth-century German idealism.

White, Morton. *Toward Reunion in Philosophy.* Cambridge, MA: Harvard University Press, 1956. A forceful and lively essay on main issues in recent British and American philosophy; Part I is devoted to recent metaphysics.

Wolheim, Richard. *F. H. Bradley.* Harmondsworth, U.K.: Penguin, 1959. An introduction to Bradley; quite critical but closely and skillfully argued.

Roger Hancock (1967)

METAPHYSICS, HISTORY OF [ADDENDUM]

THE CRITIQUE OF METAPHYSICS

In the years just before and after World War II a decidedly negative attitude toward metaphysics pervaded the analytic tradition. Before the war the logical positivists appealed to their empiricist criterion of significance to conclude that taken at face value as claims about the nonlinguistic world, metaphysical statements are literally meaningless. After the war ordinary language philosophers were not much kinder in their assessment of metaphysical claims. Here, Ludwig Wittgenstein led the charge with his claim that metaphysical statements are nothing more than nonsense born of linguistic confusion; but even ordinary language philosophers who found the Wittgensteinian critique overblown thought defenders of traditional metaphysics naöve if not totally misguided. Then, in the space of just a single year, two books appeared that did much to soften these pervasive antimetaphysical prejudices: P. F. Strawson's *Individuals: An Essay in Descriptive Metaphysics* (1959) and Willard Van Orman Quine's *Word and Object* (1960). Although each of these books was written by a philosopher whose roots were squarely within one of the two traditions that had been so critical of metaphysics, both were avowedly metaphysical works, and both exerted enormous influence on succeeding generations of philosophers.

STRAWSON

For Strawson, metaphysics is an inquiry into the most general features of our thought about the world, what Strawson calls our conceptual framework. Revisionary metaphysicians find that framework philosophically problematic and seek to replace it with a superior framework; whereas descriptive metaphysicians have the more humble goal of describing the conceptual framework we actually employ. Strawson himself had been a leading figure in the ordinary language tradition that identified philosophy with conceptual analysis; but he denies that descriptive metaphysics is simply a form of conceptual analysis. It is both more general and more comprehensive than conceptual analysis, and it seeks to identify the presuppositions of the various uses of language that constitute the subject matter for conceptual analysis.

The topics that Strawson discusses in *Individuals* are those that provide the focus for traditional metaphysics: the individuation and persistence of particulars, the relationship between material bodies and the frameworks of

space and time, the mind-body problem, and the problem of universals. Nonetheless, Strawson's approach to these topics is colored by his ordinary language roots. He asks about the identification and reidentification of particulars, the attribution of psychological and physical predicates, and the underpinnings of the subject-predicate distinction. But not only is the methodology of *Individuals* rooted in the ordinary language tradition. Its substantive conclusions serve to vindicate the commonsense picture of the world that gets expressed in ordinary language; nor is this surprising since the book is supposed to be an exercise in descriptive metaphysics. Accordingly, the material bodies of everyday experience are taken to be the ontologically basic particulars; psychological and physical properties are construed as irreducibly different; the notion of a person is treated as a primitive concept not susceptible of any form of reductive analysis; and the distinction between particulars and the universals they instantiate is treated as ontologically fundamental.

QUINE

The metaphysical framework at work in Quine's *Word and Object* could hardly be more different. By Strawson's standards Quine is a revisionary metaphysician. In Quine's book the commonsense metaphysics Strawson defends gives way to an austere ontological scheme geared to accommodate the core insights of what Quine takes to be the most successful scientific theory: physics. Accordingly, we have the view that time is just another dimension along with the three spatial dimensions; familiar particulars are construed as space-time worms; we have a strictly materialist account of thought and experience; talk of meanings, properties, and propositions along with the appeal to the modal notions of necessity and possibility are rejected; and the only abstract entities we countenance are the classes or sets of the mathematician.

Furthermore, while Strawson pays close attention to the ways words function in ordinary language, Quine is an heir to the logical positivist tradition and employs the technical tools of formal logic in formulating and justifying his metaphysical theory. As Quine sees it, simply by endorsing the claims of physics one is committed to the metaphysical framework he defends. Here, he relies on an account of ontological commitment he developed in works before *Word and Object*. According to that account, to determine the ontological commitments associated with endorsing a certain body of discourse, one translates the sentences making up that body into the language of first-order logic. If we call the sentences resulting from that translation *S1 … Sn*, then we can say that in accepting the original body of discourse, one commits oneself to the existence of all those entities that must exist if *S1 … Sn* are to come out true.

So if, by this criterion, one discovers that a given statement commits one to the existence of entities of a certain sort, then, provided one accepts that statement, one is required to include entities of the relevant sort in one's ontological framework; or, better, one is so required unless one can show that the commitment is only apparent; and one succeeds in showing that if one can come up with a plausible paraphrase of the original statement that, by Quine's test, is innocent of any commitment to entities of the kind in question. The underlying theme of *Word and Object* is that there is no plausible paraphrase of the sentences making up physical theory that shows them to be free of the metaphysical commitments expressed in the ontology of *Word and Object*.

RECENT BRITISH METAPHYSICS

The work of Strawson and Quine led to a revival of traditional metaphysics. The change was gradual, and it tended to take different forms on the two sides of the Atlantic. In Britain the influence of Strawson's approach was especially strong. Strawson's view that metaphysics is concerned with the structure of our thought about the world led to a style of metaphysics where the emphasis is on our conceptual practices and the presuppositions of those practices. Given the centrality of the idea of conceptual structures in terms of which we talk and think about the world, it is not surprising that British metaphysicians over the past four decades or so have been deeply concerned with questions about the relationship between our thought and the world that thought is about. Pivotal here has been the opposition between what Michael Dummett (1978) calls realists and antirealists. Whereas Dummett's realists want to claim that there is a mind-independent world, correspondence to which makes our statements and beliefs true, his antirealists question the idea of a reality whose constitution is independent of our conceptual activities and the conceptual structures we bring to bear in inquiry, and they hold that what we call truth is some epistemic property like that of being supported by adequate evidence.

RECENT METAPHYSICS OUTSIDE BRITAIN

During the 1960s and 1970s metaphysical discussion outside Britain was heavily influenced by Quine. While they tended to endorse Quine's account of ontological com-

mitment, many philosophers from this period were uncomfortable with the austere metaphysical framework he had defended in *Word and Object*. A major area of concern was Quine's unwillingness to accept properties and propositions. He had argued that whereas sets have clear-cut identity conditions (a set α is identical with a set β just in case α and β have the same members), no such identity conditions are possible for properties and propositions. Critics such as Roderick M. Chisholm (1976) replied that we have no option but to accept abstract entities like these. The existence of properties, they said, is presupposed by our talk of similarity, by subject-predicate discourse, and by talk involving abstract singular terms like *wisdom*, *triangularity*, and *mankind*; and they argued that propositions are required to serve as the objects of our beliefs.

But in endorsing properties and propositions philosophers from this period found themselves confronted with important metaphysical questions. Familiar objects, we say, have properties, but what exactly is the relationship between an object and its properties? What is called the bundle theory provided one answer to this question. On this theory there is nothing more to an individual than the properties associated with it; familiar objects are just bundles of properties. But if that is so, it should be impossible for numerically different individuals to share all their properties. Critics of the bundle theory such as Gustav Bergmann (1967) and David Armstrong (1989) argued that since this is not impossible each familiar object incorporates a constituent over and above its properties, a constituent unique to that object. This individuating constituent was variously called a bare particular or a thin particular and was construed as the literal bearer of the properties copresent with it.

The notion of a proposition gave rise to other problems. Propositions, we think, are not just true or false; they can be necessarily true or necessarily false, contingently true or contingently false, and possibly true or possibly false. Now, Quine had notoriously rejected talk of modality. Modality, he said, is mired in obscurity. To make sense of modal notions, critics such as Saul Kripke (1972), David Lewis (1986), and Alvin Plantinga (1974) appealed to the Leibnizian notion of a possible world. The idea was, first, that our world (the actual world) is just one of many possible worlds and, second, that what is unique about modal discourse is that it takes the full range of possible worlds and not just the actual world as its subject matter. These theorists did not all agree about the nature and status of possible worlds, but they did agree in endorsing the Leibnizian idea that to say that a proposition is necessarily true is to say that it is true in all possible worlds and to say that it is possibly true is to say that it is true in some possible world. This approach to modality proved tremendously fruitful. Not only did the framework of possible worlds shed light on talk of propositional necessity and propositional possibility, but it proved helpful as well in clarifying a whole variety of otherwise puzzling phenomena like the distinction between essence and accident, the concept of meaning, counterfactual conditionals, the concept of causation, and the notion of a law of nature.

The influence of these possible worlds metaphysicians was felt throughout philosophy, and by the 1980s metaphysics had come back into its own. For metaphysicians trained in that decade and after, the positivist and ordinary language attacks on metaphysics were quaint episodes from a distant past. These younger metaphysicians were not in the least apologetic about their discipline. Indeed, they were anxious to develop and defend comprehensive metaphysical theories. The result has been a tremendously active period in which all of the topics on the traditional metaphysical agenda have come under debate. Questions about universals, the structure and individuation of ordinary objects, and possible worlds and modality continue to be discussed; but in recent years metaphysicians have dealt with a much broader range of questions including those about the nature of time and space-time, the nature of identity and existence, the existence and structure of events, persistence through time, material constitution, the nature of fictional entities, freedom of the will, causality, and the nature of the mental.

See also Metaphysics; Metaphysics, Nature of.

Bibliography

Armstrong, David. *Universals: An Opinionated Introduction*. Boulder, CO: Westview Press, 1989.

Bergmann, Gustav. *Realism: A Critique of Brentano and Meinong*. Madison: University of Wisconsin Press, 1967.

Chisholm, Roderick M. *Person and Object: A Metaphysical Study*. La Salle, IL: Open Court, 1976.

Dummett, Michael. *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, 1978.

Kim, Jaegwon, and Ernest Sosa. *Metaphysics: An Anthology*. Malden, MA: Blackwell, 1999.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1972.

Lewis, David. *On the Plurality of Worlds*. Oxford, U.K.: Blackwell, 1986.

Loux, Michael J. *Metaphysics: A Contemporary Introduction*. 2nd ed. New York: Routledge, 2001.

Loux, Michael J. *Substance and Attribute: A Study in Ontology*. Dordrecht, Netherlands: D. Reidel, 1978.

Loux, Michael J., and Dean W. Zimmerman. *Oxford Handbook of Metaphysics*. New York: Oxford University Press, 2003.

Plantinga, Alvin. *The Nature of Necessity*. Oxford, U.K.: Clarendon Press, 1974.

Quine, Willard Van Orman. *From a Logical Point of View: Nine Logico-philosophical Essays*. Cambridge, MA: Harvard University Press, 1953.

Quine, W. V. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Strawson, P. F. *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen, 1959.

Van Inwagen, Peter. *Material Beings*. Ithaca, NY: Cornell University Press, 1990.

Van Inwagen, Peter. *Metaphysics*. Boulder, CO: Westview Press, 1993.

Michael J. Loux (2005)

METAPHYSICS, NATURE OF

Almost everything in metaphysics is controversial, and it is therefore not surprising that there is little agreement among those who call themselves metaphysicians about what precisely it is that they are attempting. In beginning a discussion of the nature and validation of metaphysical arguments and theories, the best course we can follow is to list some of the standing preoccupations and ambitions of metaphysicians. For this purpose we need to make the assumption that there is a distinct class of metaphysical philosophers, a class into which such thinkers as Plato, Thomas Aquinas, René Descartes, Benedict de Spinoza, and G. W. F. Hegel would fall and from which purely critical or analytic philosophers like the later G. E. Moore would be excluded. It has to be admitted, however, that the line between metaphysical and nonmetaphysical philosophy is exceedingly hard to draw, for many metaphysicians from Plato on have been expert in the supposedly nonmetaphysical pursuit of analyzing or clarifying ideas, while few self-styled analysts have contrived to stick to pure analysis without the open or covert advocacy of a metaphysical point of view.

Setting these difficulties aside, we may note three main features of metaphysics as traditionally practiced. First, metaphysicians have constantly aspired to say what there is in the world or to determine the real nature of things; they have been preoccupied, that is, with the concepts of existence and reality. Their interest in these concepts springs from a double source: from the reflection that the surface show of things often misrepresents them, with the result that we are set the task of determining their real as opposed to their apparent constitution, and from the need to specify what ultimately different kinds of things there are in the world, a need that presses itself on our attention when we wonder whether, for example, minds or numbers are independent existents. The first of these tasks might seem to belong to the scientist rather than the philosopher, for science, too, makes constant use of the distinction between the apparent and the real; we shall indicate in the next paragraph why metaphysicians have not been ready to accept this proposal for lightening their labors.

Second, metaphysics has been commonly presented as the most fundamental and also the most comprehensive of inquiries. It claims to be fundamental because questions about what there is or about the ultimate nature of things underlie all particular inquiries. If you are to assess the results of mathematical investigations, for instance, you need to determine the ontological status of mathematical objects, and according to the theory, this is a task for the metaphysician. The claim of metaphysics to be comprehensive is more difficult to justify. One possible line of support for it, followed by Aristotle, is found in the reflection that questions about existence and reality, along with those about potential and actual being and about causation that are also raised by metaphysicians, cut across the boundaries of particular sciences and arise in connection with every sort of subject matter. Thus, metaphysics is comprehensive just because of its extreme generality. But there is another way in which the claim to comprehensiveness has been advanced. It has been customary to say that whereas sciences like physics and mathematics are departmental studies each of which deals only with a part or particular aspect of reality, metaphysics, by contrast, is concerned with the world as a whole. This explains why philosophers have been unwilling to accept the suggestion that scientists might be left to determine the true nature of things. A scientific theory purports to explain, for example, the real constitution of matter or the fundamental mechanisms of the human body but not to draw the distinction between appearance and reality in an entirely general way, not to tell us, to give an instance, whether matter is the ultimate reality, as materialists suppose, or whether it is itself a manifestation of spirit, as Hegel tried to argue.

This contrast between metaphysics and the particular sciences is sometimes developed in yet another way, again, as will be apparent, to the great advantage of metaphysics. It is said that inquiries in the individual sciences are carried out under assumptions it is the business of metaphysics to make explicit and either to justify or to correct. Metaphysics, by contrast, proceeds without assumptions and is thus fully self-critical where the par-

ticular sciences are in part credulous. This line of argument goes back to Plato, who tells us that mathematicians postulate the existence of "odd and even numbers" and "three kinds of angles," and implies that these "hypotheses," taken as "starting points" or "bases" in mathematics, could find their justification and thus lose their hypothetical character in the comprehensive "synoptic" study Plato called dialectic. The dialectician is a man who leaves nothing unquestioned, and just because of this the results of all other inquiries must be seen as no more than provisional; they await ratification or correction from the dialectician. The apparently arbitrary and obviously vague character of this suggestion has not prevented its having a continued appeal to philosophers. Even today, we sometimes hear it said that we need not be unduly disturbed by, for example, the findings of physiologists and psychologists, since the proponents of these sciences work under assumptions it is the business of philosophers to uncover and correct in the light of their knowledge of the whole man (for an argument on these lines see J. S. Haldane, *The Philosophy of a Biologist,* Oxford, 1935).

If metaphysics is to make good its claim to be uniquely self-critical, its propositions must be shown to be exempt from intellectual challenge as those of no other study are. Descartes, in fact, tried to offer such a demonstration. He argued first that such commonsense assertions as "There is a table under the window" were in every case open to theoretical doubt: However much I seemed to perceive a table, it might be that I was under perceptual illusion or was dreaming. Next, he maintained that even propositions whose truth appeared to be evident, such as those of mathematics, could not be accepted as necessarily in order. An evil demon could be deceiving me into thinking them clear and distinct when they did not really deserve this description. But matters were different when we came to the fundamental metaphysical truth "I think, therefore I am." This truth was such that in the very act of doubting it, one reaffirms it. To doubt is to think, and in thinking that I might not exist, I make clear that I do. Hence, there is at least one truth about whose correctness I could not be in error, and this is a truth of metaphysics. But Descartes was not content to stop at this point. He went on to argue that if I, a being with obvious limitations, certainly exist, then just as certainly there exists a perfect being whose nature is such that he would never deceive me into thinking that true which is not in fact so, once I have satisfied myself that it is by the test of clear and distinct perception. The effect of this move was to provide a guarantee for the findings of the sciences, which were otherwise open to "hyperbolical" doubt. We could henceforth be assured on metaphysical grounds that whatever was clearly and distinctly perceived was true. As for the propositions of metaphysics itself, their truth was guaranteed by their connection with the *cogito,* which, as we have seen, could not be intelligibly questioned.

The interest of these arguments for our present purpose lies not in their details but in the basic claims they involve. The propositions of metaphysics, according to Descartes, are intellectually impregnable, and in this respect they contrast not only with the beliefs of common sense but also with the pronouncements of the sciences, at least when these are considered apart from their metaphysical guarantee. But from where can they derive their unique certainty? The only possible answer is from their being the products of reason when that faculty is put to work in the fullest and freest way. The result will be that metaphysics is not only the most fundamental of studies; it is also one that relies for its results on the efforts of reason alone.

METAPHYSICS AND THE SUPERSENSIBLE

Thus far, we have observed three main features in the projected science of metaphysics. It claims to tell us what really exists or what the real nature of things is, it claims to be fundamental and comprehensive in a way in which no individual science is, and it claims to reach conclusions that are intellectually impregnable and thus possess a unique kind of certainty. Now, many critics of metaphysics have suggested that these claims could be justified only if metaphysics were a factual science providing us, on the strength of rational insight, with knowledge of things or aspects of reality that lie beyond the range of the senses. Nor is this view without support from practicing metaphysicians. Plato drew a contrast between "things seen" and "things unseen" and argued that only things unseen were proper objects of knowledge. From his time on there was a standing tendency to identify the province of the metaphysician with what was vaguely called the supersensible, or the realm of the intellect. Aristotle, for example, distinguished between sensible and insensible substance and assigned the investigation of insensible substance to "first philosophy," or metaphysics. Medieval and early modern philosophers thought of God, the "being of beings," as an entity without bodily extension or shape and for that reason considered him outside the province of the empirical sciences. More generally, it was widely believed that behind the phenomena that present themselves in everyday experience, there lie realities

whose existence and properties can be established only by use of the intellect and that can hence be described as noumena, or intelligible objects. In this view, the proper concern of metaphysics was to give us news about noumena.

From the eighteenth century on much ingenuity has been displayed in showing the untenability of this position. The idea that there might be a science that was at once factual and purely intellectual drew its firmest support from the example of mathematics. David Hume suggested, however, that the concern of the mathematician was not with matters of fact and existence but solely with "relations of ideas": His aim was only to make explicit what was already implicit in the premises from which he started. The propositions of mathematics were indeed necessary truths, but by the same token they gave no information about the world. If an inquiry was to pronounce on matters of fact, its method must be empirical, not conceptual, and this meant that its results could not possibly claim to be intellectually impregnable, for anything established on the strength of experience might need to be amended or even withdrawn in the light of further experience. There were no final empirical truths.

A natural reply to this is to argue that even if every factual inquiry must begin from experience, it need not necessarily terminate there. Why should not the metaphysician argue from the characteristics of things sensible to the existence and the nature of things supersensible, as, for instance, Thomas Aquinas and John Locke thought they could? Immanuel Kant was much concerned about the proper answer to this question. He allowed—and here he showed more sympathy with metaphysicians than empiricists then or now—that such concepts as cause and substance, which figure prominently in supposed inferences from the phenomenal to the noumenal, have a necessary character; in Kant's terminology they are a priori, as opposed to empirical, concepts. But he denied as stoutly as Hume that they can therefore be used to carry us beyond the range of possible experience. The question "What brought that about?" is a necessary question, one we cannot rationally refuse to ask, but the answer to such questions must always be sought within experience. If we try, as, for example, Descartes did, to maintain that there must be a First Cause, a necessary being entirely different from the contingent things with which we are familiar, we cease to attach any clear meaning to the concept of cause, for, as Hume saw, it is an essential part of the idea of cause that a cause precede its effect. We can talk about causes as long as we remain within the sphere of the temporal; once we step outside it, the concept loses its determinate character. And what is true of cause here is also true of substance and other metaphysical notions. We can give sense to the concept of substance if we understand it as the permanent that persists through change, but if we eliminate the reference to time, we are left with no more than the logical notion of that which is always a subject and never a predicate, an idea that in its pure form is too indeterminate to be put to metaphysical or, indeed, any other use.

Another attack on metaphysics as the supposed science of intelligible reality was made by the logical positivists. It is a mark of those propositions that belong to accredited sciences like mechanics or genetics, they argued, that we know in principle how to test them; we can see what difference it makes that they are true rather than false. But if a metaphysician comes along and tells us that what really exists is not trees or tables but, say, monads, what tests can we apply to determine the truth of his statement, and what difference does it make if it is true? By definition monads are entities that could never be encountered within experience, nor is their presence supposed to have particular empirical consequences like that of electrons and similar unobservables postulated by natural scientists. Thus, a metaphysical thesis will be compatible with any state of affairs whatsoever, just as the propositions of logic and mathematics are. But if this is so, how can it possibly be maintained that metaphysics gives us information about the world, even the unseen world? The news it purports to bring can only be news from nowhere.

These highly general refutations of a particular conception of metaphysics have seldom been found convincing by metaphysicians. One reason for this is that they fail to come to grips with individual metaphysical arguments, for example, with the *cogito.* Another is that they appear to prejudge the case against this sort of metaphysics. Why, for example, should it be supposed that a metaphysical thesis must make an empirical difference? Another cause of their failure to carry conviction, however, may be found in the fact that many metaphysicians have worked with a different concept of their subject, one that does not involve it in the claim that it provides information or rivals the empirical sciences. This conception will be considered below.

METAPHYSICS WITHOUT ONTOLOGY

We have already seen that metaphysicians have wanted to say both that their propositions possess a peculiar certainty and that they are significant as a purely analytic proposition is not. In Kantian terminology they pretend to the status of synthetic a priori truths. Now, many crit-

ics of metaphysics have made the assumption that a proposition could be synthetic a priori only if it at once stated a truth of fact and was established by conceptual means alone, a combination they regard as impossible. Facts must be established empirically; pure thinking can lead to the knowledge only of analytic truths. But if we look at Kant's alleged synthetic a priori judgments, particularly those he called principles of the understanding, we see that they make no claim to state facts, even very general facts. A principle like the principle of causality is not a very wide empirical truth, mysteriously known in a nonempirical way; it is, on the contrary, the expression of a rule of procedure that serves to tell us not what properties things have but how to interpret them. Kant supposed that principles of this sort had a special sort of necessity, though they did not logically compel; they owed this, he thought, to the fact that they are prescribed by the human mind as principles specifying what is to count as objective in our experience. Thus, we take it to be a feature of what is objectively there that no quality is present except in a determinate degree, that nothing ever goes entirely out of existence (all change is transformation), that nothing happens except for a reason, and so on.

Kant himself intended this doctrine to have limited application. He thought of the principles of the understanding as prescribing the form of the phenomenal world that we know by means of the senses and investigate in the natural sciences. In his view there were other aspects of experience, in particular the activities of the moral agent, in regard to which they had no legislative force. But it is possible to think of an extension of Kant's doctrine and imagine a set of principles that would prescribe the form not just of one department of experience, but of experience as a whole. A set of principles of this kind would tell us how to organize the data of our experience in such a way that we could give a unitary account of them; it would thus help us make sense of the scheme of things entire. Possessed of concepts of this sort, we could hope to resolve the apparent inconsistencies of science and common sense, together with the more serious conflicts between science and religion and science and morality. We should then be masters of an overall point of view enabling us to see things synoptically or have a set of ideas that would allow us to differentiate the real nature of the universe from its merely superficial aspects. We should, in short, be in possession of a metaphysics.

There can be no doubt that many of the classical metaphysical systems can be thought of as conforming to this schema. In the system of Aristotle, for instance, the key concepts are teleological, and their articulation is to be found in the doctrine of the four causes. It is axiomatic in Aristotle's thought that everything serves a purpose; Aristotle's ambition is to find the point of each phenomenon and thus specify its place in the articulation of the whole. He attempted to carry through his program not only at the biological level, the most obvious source of the concepts involved, but also above and below it—in moral, political, and social life, on the one hand, and in physical science, on the other. His success in these spheres is unequal, but that does not affect the general character of the enterprise.

The popular philosophy of materialism, again, can be seen as an attempt to make sense of the world as a whole on the basis of a distinctive set of first principles. The primary thought of the materialist might be expressed in the axiom that there is nothing that cannot be satisfactorily explained in natural terms; belief not merely in the competence, but also in the omnicompetence, of natural science is a prominent item in his credo. The materialist sees the world as a vast mechanism; whatever happens is the result of natural causes, and all other phenomena must be assessed and understood on this basis. Thus, the phenomena that characterize religious and moral life can be taken in psychological and social terms as things whose causes are ultimately natural, though scarcely in the terms favored by those who engage in them. Religion, as Sigmund Freud said, is an illusion but not an unintelligible illusion; science can account for it, as it can account for everything else.

Finally, Hegelianism made a conscious attempt to produce a metaphysics that constitutes an overall reading of experience. The central concept here is the concept of spirit; it is alleged that everything can be understood in terms of this concept once we take account of the fact that spirit cannot fulfill its potentialities except by working on and against something not itself—in Hegel's peculiar language, "its own other." Thus, we can make sense of the existence of a world of nature in this system; it is there to subserve the purposes of spirit. We can make sense of the social world, too, for many of the characteristics of mind are intelligible only when people are aware of one another and know that others are aware of them. Self-respect and self-contempt would be cases in point.

Each of the systems mentioned could be said to rest on a basic idea or intuition, an idea articulated in a series of concepts taken as definitions of reality and applied, with greater or less success, to the whole range of experience. To appreciate the force of such a system, we need to grasp the basic idea as well as understand the articulated concepts; we have to see the world as the metaphysician

in question saw it. The deviser of a metaphysical theory thus becomes a man with a vision of the scheme of things entire. It is important to add, however, that he is not merely a man with a vision, in which case he would be indistinguishable from a philosophical poet. He needs to work his vision out in a theory; he needs to argue his case both by adducing those facts that immediately support it and by explaining those that on the face of things do not.

It seems clear that most of the standard claims for metaphysics can be understood with this account of the matter. Since the first principles of a metaphysical system have prescriptive force, exactly as Kant's principles of the understanding had in regard to the world of nature, they can be properly thought to compel every rational thinker. Their certainty is not the certainty of logic, and yet it exceeds that of any individual statement of fact, for facts are descried only within a framework that these principles provide. Again, even if a system of this kind does not tell us precisely what there is, it nevertheless pronounces on the real character of the world as opposed to the surface show. According to the materialist, for instance, there seem to be features of experience that transcend the natural realm, but in the end it turns out that this is not so. Everything, including men's thoughts and actions, can be accounted for satisfactorily in natural terms. That a scheme of this kind is comprehensive, wider than that of any particular science, goes without saying; that it is fundamental because it is concerned with the coordination of ways of thinking in widely differing spheres is also obvious. True, there is no straightforward counterpart in this type of theory for the criticism by metaphysics of the assumptions of the particular sciences: Metaphysics not being a source of knowledge in itself, it cannot be claimed that other studies are dependent on it as, say, chemistry is dependent on physics. But this circumstance will not prevent this type of metaphysician from putting his own construction on the results of the sciences, as the example of Hegelianism shows. He may have no warrant to question such results, but all the same he may insist on interpreting them in his own way when he offers his reading of experience as a whole. Hegel was doubtless too brusque in his treatment of Isaac Newton and John Dalton, but it does not follow that the whole project for a philosophical treatment of natural phenomena is a mistake.

ARGUMENT AND TRUTH IN METAPHYSICS

If metaphysics answers the description given above, a description that would fit many if not quite all of the best-known metaphysical systems, two questions immediately arise. First, we may be asked what sort of a study metaphysics is in this account. Is it a priori or empirical, and to what sorts of argument does it appeal? Second, there is the question what criteria to use in choosing among metaphysical systems. Seeing that many systems are possible, are there any objective ways of deciding that one system embodies the true or the proper way to look at the world?

ARGUMENT. The answer to the first query is that metaphysics, according to this account, is neither a priori nor empirical, though it makes constant use of both deductive and probable reasoning. A metaphysician is concerned to advocate, articulate, and apply a set of basic interpretative principles, categorical principles we might call them, and principles of this kind cannot be grounded in either conceptual considerations or an appeal to empirical fact. They cannot be supported conceptually since no contradiction is involved in disputing them; they cannot be deduced from facts since they claim to apply with unrestricted validity, no matter what data turn up in experience. They may indeed be suggested by experience and commonly are, but that is not to say that they can be shown to be acceptable or unacceptable by simple empirical methods. Apart from anything else there are no absolutely neutral data to which we can appeal when supporting or attacking a metaphysical theory. For though it is the case that every metaphysician has the duty of explaining all the facts as he sees them, he also has the privilege of being able to decide what really is to count as fact. To see the importance of this we have only to reflect on the different views of religious phenomena taken by materialists and their opponents.

However, though it is true that a metaphysical theory on this account can be established neither deductively nor inductively, deductive and inductive argument both bulk large in metaphysical discussion. Like any other thinker the metaphysician is much concerned with consequences and consistency. He often wants to make the point that since p is true and p implies q; which in turn implies r, we are logically committed to r or to contending that since q is false and p implies q, p must also be false. The very fact that a metaphysician has a theory to put forward means that he must be preoccupied with the logical connections between the concepts that constitute his system. To say this, however, is not to deny his preoccupation with fact or with probable arguments. Unlike an empirical scientist he establishes no new facts, but all the same he has a double interest in fact. First, he is concerned, more than any specialized inquirer, to see similarities in widely different areas of fact, a process that is

relevant to both the formulation and the application of his theory and that involves him in much reasoning by analogy. Second, he needs to pay constant attention to the state of factual knowledge in working out and pressing home his central insight. He promises, after all, to make sense of all the data of experience, and he must consequently take continuous account of these data. The legend that metaphysicians are indifferent to fact has no foundation; on the contrary, they have a primary interest in facts of all sorts even though they do not originate any factual propositions. The extent to which advances in cybernetics have been discussed in recent years by philosophers interested in the truth of materialism affords an apt and striking illustration of this point.

TRUTH. We saw that one charge made against metaphysics as a doctrine of what there is was that no decisive considerations can be adduced either for or against such a theory; the monads of Gottfried Wilhelm Leibniz and the Forms of Plato make no empirical difference. In this respect are things any better in our revised form of metaphysics? It must be confessed that the initial appearance is not favorable. We have emphasized that the first principles of such a system are neither analytic nor empirical; the temptation to conclude that they must accordingly be no more than arbitrary prescriptions, representing a point of view taken up for no good reasons, is strong. And though we have also urged that metaphysicians of this sort have a special interest in fact, the force of that contention is considerably weakened by the admission that they claim the right to decide for themselves what really is fact. If we arm them with this veto—and it is hard to see how they could be refused it—the question of metaphysical truth seems wholly intractable.

It could be, however, that we are setting an impossible standard for metaphysics in requiring it to possess a decision procedure as clear-cut as those of mathematics and the natural sciences. One reason that we can get a straight answer about the acceptability of a theory in physics is that physics works on principles that it does not question (such as that every natural happening will have a sufficient natural explanation). In metaphysics, by contrast, we are concerned with the comparison and assessment of precisely this type of principle. As the widest and most general of all forms of thinking, metaphysics can appeal to no fixed criteria beyond itself except to the requirements of internal consistency that any theory must satisfy. Nor is it true that every reputable branch of knowledge possesses obvious and easily applicable decision procedures. If, for example, we compare metaphysics with history instead of physics, we may begin to see that there are areas of study where dispute and disagreement play a prominent part and that still can claim to proffer understanding and enlightenment. Once we pass beyond the mere ascertaining of fact, there are many histories written from many points of view and resting on many judgments about what is historically important; it is not really possible to hope for a final decision about which, if any, is correct or even about the relative merits of any two equally sophisticated interpretations. However, we do not conclude from this that history is a pointless pursuit rational men would do well to avoid. We realize that a study like history can enlarge the mind and educate the understanding even when it does not add to the sum of public knowledge.

A comparison with metaphysics that is in some respects even closer is provided if we consider the interpretation of a literary text. The data the literary critic confronts—I am thinking of someone who offers a reading of a controversial literary work like *Hamlet* or *Faust*—are "harder" than in the case of metaphysics, but this does not prevent the appearance of a wide variety of conflicting theories. And it happens that there are no accepted criteria for deciding among the various theories; all that each critic can do, in the last resort, is explain his way of looking at the text, marshal the points in its favor, and invite the reader to test the matter for himself. But we need not conclude from this that it will be a matter of luck or, perhaps, of psychology which theory will win the reader's approval. At the end of the day, he can be entirely convinced of the authenticity of one particular reading, and he can be persuaded that it offers more enlightenment, covers the central points more impressively, and does better justice to the evidence than its rivals. He may not be able to produce knockdown grounds in favor of his choice, but that is not to say that he has made it for no reason at all.

Metaphysical argument is like literary argument in that it reaches no apparent end; it is like it again in terminating, insofar as it ever does terminate, in an insight that is more personal than public. The old dream of a demonstrated metaphysics whose propositions were even more certain than those of mathematics could scarcely be further from realization. But it would be wrong on that account to think that the concepts of truth and falsity have no application in metaphysics. At the lowest estimate we can describe one system of metaphysics as more illuminating than another. We must, however, decide for ourselves what is really illuminating and what is not. As in the case of the humanities in general, we cannot just learn the truth from another.

CONTEMPORARY ANTIMETAPHYSICS

Theories that profess to deal with "the world as a whole," however they are meant to be taken, are today more often objects of suspicion than of interest, thanks to the influence of G. E. Moore and the later work of Ludwig Wittgenstein. Moore himself never attacked metaphysics explicitly, and indeed his early work, both in logic and in moral philosophy, showed pronounced metaphysical leanings of a generally Platonic kind. But the "Defence of Common Sense" with which he came to be most prominently associated was evolved as a counterblast to views put forward by contemporary metaphysical philosophers, views that, as Moore saw them, could be maintained only by someone prepared to disregard what he evidently knew to be true. When F. H. Bradley, for instance, argued that time is not real, Moore thought this an absurd paradox since the reality of time is taken for granted in any statement containing a temporal expression. If time is not real, it cannot be true that yesterday was Friday or that I had my breakfast before leaving for work. Moore's procedure here, which is to call the metaphysician's bluff by reminding him of what in an off-duty moment he will himself acknowledge that he knows, was generalized by some of his followers into an all-round exposé of metaphysics, which they represented as necessarily consisting of paradoxes and evident falsehoods. For this purpose the thesis that everything is material did not differ from its rival that everything is spirit; both were, when taken seriously, obviously false. There might be a point in maintaining such a thesis (it could be a revealing paradox, according to John Wisdom, or serve a deep-seated psychological purpose, according to Morris Lazerowitz), but in no sense could it express what was really the case.

Moore and his followers assume here that there can be only one correct description of a situation and that in matters like dating or temporal precedence it is known to all of us. It is not obvious that this view is correct, for it could be, as Bradley thought, that a description that was valid and serviceable at the commonsense level would need to be superseded when wider considerations were taken into account. One way of putting Bradley's view is to say that metaphysics claims to offer a conceptual scheme in terms of which we can give a description of the world that is ultimate and comprehensive, but that it also recognizes the existence of many subordinate and more limited schemes, each of which has its point in the characterization of appearances. The Bradleian doctrine of degrees of truth and reality is obviously relevant here, and it cannot be said that Moore gives it very serious consideration. But even if this point had to be granted, the respectability of metaphysics might still be in doubt, for the whole notion of an ultimate description of the world is itself suspect thanks to the work of Wittgenstein.

According to Wittgenstein, a principal source of philosophical error has been the idea that the primary function of language is to describe. The truth is, rather, that we engage in many different "language games," each of which serves its own purpose and each of which is authentic at its own level. There can be no question of ruling any such game out of court; the fact that it is played is sufficient evidence that it is appropriate. Nor are different sets of language users rivals; it could not be said, for instance, that physics gives a truer picture of the world than common sense or that the naïvetés of everyday moral language are corrected by the psychologist. If we keep these diverse languages apart, we see that each has its own point and utility. The idea of a finally correct language that would embrace and replace them all is clearly the height of absurdity, and, hence, metaphysics in its revised form is no more acceptable than was metaphysics in the shape of news from nowhere.

But this analysis, too, is built on questionable assumptions. First, is it really clear that language games or areas of linguistic activity are as distinct as Wittgenstein says they are? The point is by no means clear as far as the language games of science and common sense are concerned, for most scientists and many plain men think that the scientific account of the physical world gives a truer picture of it than that embodied in the ordinary man's everyday beliefs. Nor can we agree without further argument with the thesis that sufficient authentication is found for a language game when we note that it is played. There are, after all, games and games. In a form of game much played in the ancient world, elaborate formulas to appease the god of the sea were devised by those about to embark. As a result, a certain way of talking commanded a wide use and approval. But could that fact alone be invoked to show that it was legitimate? Surely, we should want to object that however much such language was used, its use could not be legitimate if in fact there was no god of the sea or if he exercised no influence on whether seafarers reached their destinations safely. To do this, however, is to make the propriety of a language game subject to the tenability of the factual assumptions on which it rests. Although this is not to maintain that the only use of language is to describe (which would be absurd), it is to claim a certain priority for the language game in which we say how things are.

Metaphysics as we have expounded it is concerned with resolving conceptual conflicts by finding a way of

speaking that will enable us to express the true nature of the world. If we possess such a way of speaking, we have a yardstick by which to measure the ultimate tenability, as opposed to the immediate use, of particular language games—the languages of religion, science, law, and so on. It is not self-evident that each of these is in order as it is, and though the fact that they are constantly used and understood is enough to show that they serve some purpose, it does not in itself show that they are suited for the purposes those who use them have in mind. These games are indeed played, but they could, for all that, be played on false pretenses. To decide whether they are, we must have recourse to metaphysics.

METAPHYSICS AS ANALYSIS

Even if the foregoing account of the nature of metaphysics were accepted as generally unobjectionable, there are many philosophers who would deny that it covers everything metaphysicians have attempted or are attempting to do. In particular, it fails to accommodate an activity pursued by many contemporary analytic and linguistic philosophers that has a clear affinity with the work of some of the classical metaphysicians. The classical metaphysicians were led to ask what there is partly because of puzzles about the status of numbers and qualities. Plato had produced arguments to show that these must be independently real, and Aristotle elaborated the doctrine of categories as an answer to them. Now, there are plainly parallels to this controversy in contemporary philosophy, both in the discussions among logicians about names and descriptions (which revive the ancient dispute about the relative priority of universals and particulars) and in the arguments about the relation of the mind and body that have recently been so prominent in British and American philosophy. What is notable about these issues, as opposed to those mentioned above, is that matters of fact appear to have no relevance to their solution. If we can solve them at all, we can solve them only by thinking.

This contrast is both genuine and important; there certainly are philosophical activities that are traditionally connected with metaphysics and that cannot be subsumed either under the schema given above or under that which it was meant to replace. These activities are in essence logical or analytic, and insofar as it is confined to them, metaphysics is indistinguishable from analysis. But there is no reason to confine metaphysics to such inquiries. That metaphysicians have been speculative theorists as well as ontologists in the restricted modern sense is almost too obvious to need mention; to decide, as some commentators do, that the speculation can be set aside as regrettable and the ontology played up is at best arbitrary. Nor is it true that we can make an entirely clear-cut distinction between the two. If we look at recent work on the mind-body problem, for instance, we see that much of it is indeed logical in a wide sense of that word but that considerations of substance also come in, for example, when we discuss the nature of consciousness or of thought bearing in mind the properties and possibilities of thinking machines. An all-important motive that impels men to persist with these questions is the need to take account once more of the claims of materialism against a background in which new scientific and technical discoveries seem to lend increased support to those claims. However fascinating logical problems may be, interest in them cannot be long sustained without some external stimulus. It is such a stimulus that metaphysics of the broad kind argued for above may be expected to provide.

See also Appearance and Reality; Aristotle; Being; Bradley, Francis Herbert; Categories; Descartes, René; Dialectic; Existence; Freud, Sigmund; Hegel, Georg Wilhelm Friedrich; Hegelianism; Hume, David; Kant, Immanuel; Language, Philosophy of; Leibniz, Gottfried Wilhelm; Locke, John; Logical Positivism; Materialism; Metaphysics, History of; Monad and Monadology; Moore, George Edward; Ontology; Plato; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Time; Wisdom, (Arthur) John Terence Dibben; Wittgenstein, Ludwig Josef Johann.

Bibliography

CLASSICAL WORKS

Aristotle. *Works,* Vol. VIII, *Metaphysics,* edited by J. A. Smith and W. D. Ross. Oxford, 1928. See especially Γ 1–2; E 1; Λ 1–2, 6. Compare also W. D. Ross, *Aristotle's Metaphysics,* Vol. I. Oxford, 1924. See pp. lxxvii ff.

Baumgarten, A. G. *Metaphysica.* Halle, 1739.

Bradley, F. H. *Appearance and Reality,* Oxford, 1893. See especially the preface and introduction.

Descartes, René. *Meditations on First Philosophy* (1641). In *The Philosophical Works of Descartes,* translated by E. S. Haldane and G. R. T. Ross, 2 vols. New York, 1931. See especially "Reply to Second Objections."

Hegel, Georg Wilhelm Friedrich. *Phenomenology of Mind* (1807). Translated by J. B. Baillie, 2nd ed. New York: Macmillan, 1931.

Hume, David. *Enquiry concerning Human Understanding.* London, 1748. There are many modern editions.

Hume, David. *A Treatise of Human Nature* (1739–1740), edited by L. A. Selby-Bigge. Oxford, 1888; 1941.

Kant, Immanuel. *Critique of Pure Reason* (1781). Translated by N. Kemp Smith. London, 1929.

Kant, Immanuel. "Inaugural Dissertation" (1770). In *Kant's Inaugural Dissertation and Early Writings on Space,* translated by J. Handyside. Chicago: Open Court, 1929.

Kant, Immanuel. *An Inquiry into the Distinctness of the Principles of Natural Theology and Morals* (1764). In *Kant's Critique of Practical Reason,* edited by L. W. Beck. Chicago, 1949.

Kant, Immanuel. *Prolegomena to Every Future Metaphysics* (1783). Translated by P. G. Lucas. Manchester, U.K., 1953.

Plato. *Republic.* Translated by A. D. Lindsay. London: Deat, 1976. See especially VI–VII.

Plato. *Sophist.* In *Plato's Theory of Knowledge,* edited by F. M. Cornford. London: K. Paul, Trench, Trubner, 1935. See especially 253ff.

Wolff, Christian. *Philosophia Prima Sive Ontologia.* Frankfurt, 1729; edited by J. Ecole, Hildesheim, 1962.

MODERN WORKS

Ayer, A. J. *Language, Truth and Logic.* London: Gollancz, 1936. Lively brief account of the logical positivist criticism of metaphysics. See also the introduction to the second edition (1945) for replies to objections.

Broad, C. D. "Critical and Speculative Philosophy." In *Contemporary British Philosophy,* edited by J. H. Muirhead, Vol. I. London: Allen and Unwin, 1924. An influential article.

Carnap, Rudolf. "The Elimination of Metaphysics through Logical Analysis of Language." In *Logical Positivism,* edited by A. J. Ayer. Glencoe, IL: Free Press, 1959. An important and influential article; originally published in 1932.

Carnap, Rudolf. *Philosophy and Logical Syntax.* London: Kegan Paul, 1935.

Collingwood, R. G. *An Essay on Metaphysics.* Oxford: Clarendon Press, 1940. Written in reaction to Ayer but also of independent interest.

Emmet, D. M. *The Nature of Metaphysical Thinking.* London: Macmillan, 1945.

Lazerowitz, Morris. *The Structure of Metaphysics.* London: Routledge and Paul, 1955. Mainly influenced by Moore.

Moore, G. E. "The Conception of Reality." In his *Philosophical Studies.* London: Routledge, 1922. Sharp criticism of Bradley.

Moore, G. E. "A Defence of Common Sense." In his *Philosophical Papers.* London: Allen and Unwin, 1959.

Moore, G. E. *Some Main Problems of Philosophy.* London: Allen and Unwin, 1953. Expounds at length the doctrine of "A Defence of Common Sense"; written in 1911.

Pears, D. F., ed. *The Nature of Metaphysics.* London: Macmillan, 1957. A series of modern discussions of metaphysics.

Walsh, W. H. *Metaphysics.* London, 1963.

Wittgenstein, Ludwig. *The Blue and Brown Books: Preliminary Studies for the Philosophical Investigations.* Oxford: Blackwell, 1958.

Wittgenstein, Ludwig. *Philosophical Investigations.* Oxford: Blackwell, 1953.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus.* London: Routledge, 1922.

W. H. Walsh (1967)

METAPHYSICS, NATURE OF [ADDENDUM]

What is metaphysics? An answer to this question requires a specification both of the *scope* of metaphysics—that is, of the nature of the questions that metaphysicians raise and attempt to answer—and of the *methods* that they employ in this enterprise.

THE SCOPE OF METAPHYSICS

As regards scope, a natural answer is that metaphysics is concerned with the investigation of the ultimate nature of reality, where this involves the attempt, first, to arrive at the most fundamental truths about what exists, and, second, to provide an account of the concepts that are involved in such fundamental truths. This characterization immediately gives rise to the question of the relation between metaphysics and science. The goal of physics, surely, is to arrive at the ultimate truth concerning the nature of the physical world. Similarly, the goal of psychology is to determine the ultimate nature of minds and mental states. How, then, do the sciences leave any room for the discipline of metaphysics?

This is a crucial question. But if one considers the issues that metaphysicians address, a clear answer will emerge. First of all, then, a central part of metaphysics involves offering accounts of concepts that are essential to scientific theories in general but of which no account is offered within any of the sciences themselves. These will include such concepts as those of particulars, properties, relations, persisting entities, events, states of affairs, causation, and laws of nature, and, with regard to these concepts, metaphysicians will ask, for example, whether causal relations logically supervene on noncausal states of affairs and whether laws of nature logically supervene upon the total history of the universe.

Second, philosophers attempt to establish necessary truths involving some of those concepts. Some of these possible necessary truths—such as the claim that any particular must have some intrinsic properties—may very well have no bearing upon scientific theories. Others, however, certainly do so. Thus, for example, the thesis that any particular must have some categorical properties implies that some current scientific theories are incomplete since they attribute propensities to objects without supplying any categorical basis. Or, more dramatically, other metaphysical theses—such as the much-discussed proposition that it is impossible for a cause to be either earlier than or simultaneous with its effect—are on a collision course with some scientific theories that have been

advanced. So, for example, this claim entails that tachyons cannot exist and that positrons cannot be electrons traveling backward in time. It also means that the General Theory of Relativity, in allowing for causal and temporal loops, is allowing for something that is logically impossible. Or, again, if a cause cannot be simultaneous with its effect, then the mathematical formulation of Newton's Second Law of Motion—$F = ma$—is not satisfactory since it fails to assign different times to the force and the acceleration that it causes.

Third, science is typically silent on questions that have no bearing on the experimental content and predictions of scientific theories. A vivid illustration here is provided by the philosophy of time. For, contrary to what Putnam and others have claimed, current scientific theories such as the Special Theory of Relativity do not settle the issue between tensed and tenseless accounts of the nature of time. Metaphysicians are, then, addressing a perfectly legitimate question when they ask whether a tensed view of time is right, or a tenseless view, and this is clearly not a question that physics attempts to answer.

Fourth, physics and the other sciences involve presuppositions for which they offer no justification. In particular, it is assumed that there is an external world and that it is a material world. Metaphysics, by contrast, makes no such assumption, and so treats it as a question to be investigated and, hopefully, answered whether there is a material world or whether, instead, the basic concrete particulars are mental entities so that some form of idealism is true.

Fifth, physics, in attempting to arrive at theories that will provide explanations of physical events, takes for granted the idea that the world of physical events is causally closed so that the only causes of physical events are other physical events. Our ordinary experience, on the other hand, appears to provide considerable support for the view that experiences involve qualitative properties, or qualia, that, in the first place, are not reducible to the fundamental entities, properties, and relations postulated in physics, and that, in the second place, appear to enter into the causation of some physical events. It is very natural to think, for example, that when a person sees something and says that it is red, that there was a property of qualitative redness that the person was aware of and that that property played a causal role in producing that person's utterance. Metaphysicians, accordingly, working in the philosophy of mind, view it as a controversial matter whether the causation of physical events involves only the entities, properties, relations, and states of affairs that are the stuff of physics. In addition, the idea that the world of physical events is causally closed rules out libertarian free will, and again, a metaphysician will insist, correctly, that until this issue is examined and settled, the assumption that physical events have only physical causes is not a justified assumption.

Sixth, the sciences rely upon induction in the form of the method of hypothesis or inference to the best explanation. The question of the justification of such methods—or of induction in general—is, of course, a question within epistemology. However, the answer to this epistemological issue may very well turn upon questions in metaphysics. So, for example, some philosophers have argued that, on the one hand, if laws are merely certain sorts of cosmic regularities, then one can never be justified in believing that any exceptionless, nonprobabilistic law obtains, and, on the other hand, that such beliefs can be justified given a metaphysically stronger conception of laws—such as the view that they are second-order relations between universals. A justification of the methods of science may depend, accordingly, upon the answers to important metaphysical questions.

Seventh, one of the crucial questions concerning the nature of reality is whether the natural world was brought into existence by God, or, at least, by some sort of immaterial being, possibly of a much more limited sort. Scientifically-based arguments have, of course, been offered both for and against the existence of an immaterial creator, but the evaluation of such arguments continues to be something that falls outside of the scope of science as presently practiced.

Finally, the sciences are concerned exclusively with the existence of contingent entities and states of affairs whereas metaphysics is not. For while questions about whether there are properties that are not reducible to those of physics, about whether the world of physical events is causally closed, about whether humans have libertarian free will, and, most would say, about whether God exists, are questions about contingent matters, metaphysics is also concerned about the existence of various things such that, if they do exist, it appears that their existence is necessary rather than contingent. Do numbers and other mathematical entities exist? Does the null set exist? Do other set-theoretical entities not involving any contingent entities exist? Is there a Platonic realm consisting of transcendent or uninstantiated universals? Do objective values exist—perhaps, as Plato thought—also in the same realm as transcendent universals? Is there a world containing intentional entities—such as concepts, propositions, or nonconcrete possible worlds? In conclusion, then, it seems clear that there are an enormous

number of very important issues that are concerned with the ultimate nature of reality and that do not fall within the scope of science.

THE METHODS OF METAPHYSICS

Metaphysical claims vary in their modal status: Some, if true, are contingent truths while others, if true, are necessary truths. One would expect, then, that quite different methods must be employed in these different cases. In fact, however, the variety is considerably greater than this suggests. Let us consider, then, some of the more important methods that philosophers use in the attempt to arrive at knowledge of metaphysical truths.

DIRECT ACQUAINTANCE WITH MENTAL ENTITIES. Consider disputes in present-day philosophy of mind concerning the existence and nature of qualitative properties of experiences, or qualia. Philosophers who affirm the existence of qualia appeal, for example, to logical possibilities of zombies and inverted spectra. Arguably, however, such appeals involve the idea that there are properties that one is directly aware of—properties that would be absent in the case of zombies and differently correlated with physicalistic properties in the case of inverted spectra. At bottom, accordingly, there seems to be an appeal to the idea of direct acquaintance with instances of properties and relations.

The idea of direct acquaintance is least controversial when invoked in support of properties and relations that can be completely given in experience. Many philosophers, however, maintain that one can also be directly acquainted with mental states that involve intentionality—such as thoughts, beliefs, preferences, emotions, and so on—while some philosopher claim that one can also be directly acquainted with a self that enjoys those various mental states.

DIRECT PERCEPTION OF NONMENTAL ENTITIES. If it exists, direct acquaintance provides one with noninferential knowledge—or, at least, noninferentially justified beliefs—concerning mental states of oneself. Many philosophers argue, however, that the scope of noninferential knowledge is not restricted to one's own current mental states. Thus it is claimed, for example, that one can have noninferentially justified beliefs about events that happened yesterday—which will therefore allow one to set aside Bertrand Russell's suggestion that perhaps the world came into being five minutes ago. Or, one can have noninferential knowledge about the existence of external, material objects, and so know that one is not a brain in vat and that idealism is not true. Or, one can be directly acquainted with objective moral values, such as the non-natural properties of George Edward Moore, or with mathematical entities, such as the natural numbers, or with supernatural minds, such as God.

None of these claims is, of course, uncontroversial. Indeed, some of them are highly contentious. The point is simply that in trying to get clear about what legitimate methods are available to the metaphysician, the idea of noninferential knowledge of contingent states of affairs—an idea often associated with such notions as direct awareness, direct acquaintance, and direct perception—deserves serious examination.

INDUCTIVE METHODS. However broad the scope of noninferential knowledge may be, it is surely true that many important metaphysical propositions concerning contingent matters of fact are such that they cannot be known in that way: They must, on the contrary, be justified on the basis of other justified beliefs. Consider, for example, the thesis that humans have immaterial, immortal souls, or the thesis that the mind is identical with the brain, or the thesis that the theoretical entities postulated by physics are real.

How do metaphysicians proceed in such cases? It is hard to see any alternative to the inductive methods employed within science where one employs such notions as hypotheticodeductive method, crucial experiments, and inference to the best explanation.

Thus, one possibility is to try to arrive at plausible entailments of the relevant proposition that can be experimentally tested. So, for example, the proposition that humans have immaterial minds would certainly seem to entail conclusions concerning what will happen in cases of brain damage. If this is so, one can then determine whether those predictions hold true. Here, as elsewhere, of course, if the predictions turn out to be false, one can modify the theory so that one has a theory that no longer has those entailments. But then considerations of simplicity and ad hocness, which are appealed to within science, will become relevant.

In some cases, one may not be able to construct an experimental test since one is dealing with theories that are experientially equivalent. Consider, for example, the problem of deciding between Berkeley's theory of reality and the view that there is a mind-independent, physical world. In such cases philosophers have sometimes been tempted to embrace the view that there are competing, interpretative, conceptual schemes between which there is no rational way of deciding. But here it is important to

notice that one can equally well have competing scientific theories that are observationally equivalent—a fact that does not mean that there cannot be rational grounds for preferring one theory to the other. Two theories may, for example, differ radically with regard to simplicity, and one may be able to show, within a sound, inductive logic, that the simplicity of a theory is directly related to the a priori probability that the theory is true.

Many metaphysical propositions, however, are not concerned with contingent matters of fact. What methods are available, then, when one is dealing with propositions that, if true, are necessary?

ANALYTIC DERIVATION. One fundamental method for establishing metaphysical truths that are necessary is by showing that they are analytically true statements, where this is a matter of showing that they follow from logical truths in the narrow sense via substitution in accordance with relevant definitions.

But how are the definitions to be assessed? Here there are at least two fundamental criteria: one positive and one negative. As regards the negative criterion, a definition must not be exposed to counterexamples, so a very important task in evaluating a definition is to see whether it is possible to construct counterexamples to the definition. If it appears to satisfy this negative criterion, then the next question is whether the definition enables one to derive what seem to be the fundamental necessary truths involving the concept in question.

THE SEARCH FOR TRUTHMAKERS. Another important technique that metaphysicians use in attempts to establish necessary truths is that of asking what sorts of facts or states of affairs could suffice to make relevant statements true. Thus, David Lewis (1986), for example, argued for the existence of a plurality of possible worlds by attempting to show that, on the one hand, such concrete worlds can serve as truthmakers for statements about what is logically possible, logically necessary, and logically impossible, and, on the other hand, that nothing else, including ersatz possible worlds, can do so. If this is right, and if, as is surely the case, at least some modal statements are true, then it follows that there is a plurality of concrete worlds.

Another illustration is provided by laws of nature. Thus, it is possible first of all to describe worlds that contain fewer and fewer instances of some basic law of nature that obtains in our world, and then, second, to argue that even if there were no instances, the law in question *could* still obtain. If this is so, then the truthmakers for nomological statements cannot be cosmic regularities, and other possibilities will have to be canvassed—such as states of affairs involving either dispositions that are never manifested or second-order relations between universals.

THE APPEAL TO INTUITIONS. A third important method that philosophers employ in attempting to arrive at necessary truths is that of appealing to intuitions. Where a metaphysical truth, if necessary, appears to be an analytic truth, the appeal to intuition would not seem to be a satisfactory terminus since it provides no account of why the proposition that seems to be necessarily true is true whereas an analytic derivation would do precisely that.

Many philosophers hold, however, that there are a priori necessary truths that are not analytic. So, for example, there are propositions concerning apparently simple, incompatible properties, such as the proposition that nothing can be both red and green at the same place at the same time. In addition, if ethical statements have cognitive content, then it is natural to think that there are basic moral statements that would be true in any possible world and thus which are necessary—such as the proposition that pain is intrinsically bad and the proposition that the killing of innocent persons is prima facie seriously wrong. But if this is right, then, if it can plausibly be argued that such propositions are not analytically true, there may be no alternative to the view that the truth of such propositions is known by means of some sort of direct, intellectual intuition, however uninformative such an account may seem.

See also Berkeley, George; Epistemology; Lewis, David; Metaphysics; Metaphysics, History of; Philosophy of Mind; Moore, George Edward; Putnam, Hilary; Russell, Bertrand Arthur William.

Bibliography

Armstrong, David M. *Universals and Scientific Realism*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1978.

Armstrong, David M. *A World of States of Affairs*. Cambridge, U.K.: Cambridge University Press, 1997.

Campbell, Keith. *Contemporary Metaphysics: An Introduction*. Encino, CA: Dickenson, 1976.

Carroll, John W. *Readings on Laws of Nature*. Pittsburgh, PA: University of Pittsburgh Press, 2004.

Jubien, Michael. *Contemporary Metaphysics*. Malden, MA: Blackwell, 1997.

Lewis, David. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.

Kim, Jaegwon, and Ernest Sosa, eds. *A Companion to Metaphysics*. Malden, MA: Blackwell, 1995.

Kim, Jaegwon, and Ernest Sosa, eds. *Metaphysics: An Anthology*. Malden, MA: Blackwell, 1999.

Loux, Michael, and Dean W. Zimmerman, eds. *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press, 2003.

Putnam, Hilary. "Time and Physical Geometry." *Journal of Philosophy*, 64 (8) (1967): 240–247.

Sosa, Ernest, and Michael Tooley, eds. *Causation*. Oxford: Oxford University Press, 1993.

van Inwagen, Peter, and Dean W. Zimmerman, eds. *Metaphysics: The Big Questions*. Malden, MA: Blackwell, 1998.

Michael Tooley (2005)

METEMPSYCHOSIS

See *Reincarnation*

METHOD IN PHILOSOPHY

See *Philosophy*

METHODOLOGY

See *Scientific Method*

MEYERSON, ÉMILE

(1859–1933)

Émile Meyerson, a French epistemologist and philosopher of science, was born in Lublin, Poland (at that time Russia). He was educated in Germany, where, after completing his classical studies, he studied chemistry under Robert Wilhelm Bunsen. In 1882 he settled in Paris; following a disappointing experience with industrial chemistry, he served as foreign editor of the Havas news agency and later as director of the Jewish Colonization Association for Europe and Asia Minor. After World War I he became a naturalized French citizen.

Meyerson never held an official teaching position. But a group of philosophers and other scholars, attracted by his celebrated erudition, formed an eager and attentive audience. He was especially well versed in the history of the sciences (chiefly, but not exclusively, the physicochemical sciences) from their origins to their most recent developments. His command of language, his clarity of thought, and his extraordinary capacity for work served him well. Both his writings and his person gave an impression of great robustness—"solid as a Roman wall," as André Lalande once remarked.

Meyerson's philosophy was offered not as a philosophy of nature but as a "philosophy of the intellect." He set himself the tasks of disentangling the principles that govern the advance of thought and of extracting from reason the kernel that constitutes the *intellectus ipse.* This search for the a priori, he held, this new critique of pure reason, should not itself be conducted in an a priori manner. It had to proceed empirically—not directly, through a psychological analysis of the activity of thought, but indirectly, through reflection on the products of thought. These products may be true or false, so long as they bear witness to a serious effort of the intellect. From this point of view, the history of the sciences provides unique documentation. Thus it is that, of Meyerson's three major works, the first (*Identité et réalité,* Paris, 1908) is almost exclusively epistemological; but in the second, *De l'explication dans les sciences* (Explanation in the sciences; 2 vols., Paris, 1921), and especially in the third, *Du cheminement de la pensée* (The ways of thought; 3 vols., Paris, 1931), the scope is widened to encompass the whole of knowledge. In the last two works it is shown that the mind works always and everywhere in the same fashion, and this catholicity of reason proves that it does indeed include a portion that is a priori.

Each of Meyerson's works begins with an attempt to dispel the positivist bias that weighed so heavily on his years of apprenticeship. Science requires the concept of thing; science searches for explanation. It is not content simply to bind together by laws the phenomena given us in sense experience in order only to predict and control them. Science tends to dissolve the qualitative datum—but only to reach behind it for a more lasting and more objective, substantial *real.* Science not only seeks to know the how, but also to understand the why. Its aim is speculative. Its theories are not merely edifices built of laws; they claim to reveal to us the innermost causes of things. Realism and causalism are two fundamental tendencies that, taken together, govern the entire activity of the scientist. For the scientist, "phenomenism" and "legalism," when he submits to them, are only provisional stages. His ambition is to get to the bottom of things, his ultimate purpose is an ontological one.

In what does explanation consist? It is at this point that the Meyersonian theory proper begins. In every domain, whether it be philosophy, science, or everyday life, to explain is to identify. Causality is nothing but a

form of logical identity. We understand a change only when it becomes evident to us that, at bottom, nothing has happened, that the entire effect was already present in the cause—or at least that the change has been reduced to the minimum, to a simple displacement. The old adage *causa aequat effectum,* mechanistic theories, and chemical equations all manifest this identifying tendency. As the Eleatic paradoxes attest, we are troubled even by change of place and by the mere passage of time. Reason is satisfied only to the degree that it succeeds in eliminating time. The principle of inertia, the reversibility of mechanical phenomena, the conservation of matter and energy, the permanence and immutability of the ultimate elements, show in what direction we insistently turn as we strive for intelligibility.

Yet in a world thus rigidly set, there still remains a qualitative diversity that is the source of new attempts at identification: the elimination of "secondary qualities," the explanation of apparent differences in terms of combinations of quite similar elements from which all but geometrical properties have been removed. Thus the world is fully intelligible to us only if we succeed in assimilating it, in the final analysis, to homogeneous space. Being, like becoming, tends to turn into its opposite when our reason seeks to explain it.

But reality resists this persistent will to identify. Carnot's principle defeats any hope of eliminating time. It proves that the irreversibility of the course of time is not a subjective illusion, that the future is not interchangeable with the past, in brief, that something really does happen. Furthermore, in denying sense qualities any place in the physical world, mechanism has not thereby made them disappear. The heterogeneity of the data of sense exists unexplained and indeed inexplicable from a mechanistic point of view. In addition, atomic discontinuity puts an obstacle in the way of geometrization. Reality rejects the identity to which reason would reduce it. The real is only partly intelligible; it contains elements that are irreducible, and hence irrational. It is in fact the presence of these irrational elements, contradicting the rationalist idealism of the philosophers, that can serve to define the real in opposition to the structures erected by our thought. Thus while reason may well move from success to success in the quest for identity that essentially motivates its activities, it can never win a definitive victory. In the end, it is condemned to defeat.

Indeed, how could matters be otherwise? There is something odd and almost absurd about this endeavor of reason, for its complete success would betoken its ultimate failure. To explain reality fully would amount precisely to denying it as real, to dissolving it into a motionless and undifferentiated space, that is, into nothingness. A perfect explanation of the world would end up in acosmism. And the conflict would be met with again even if the object studied were only an ideal one, as in the case of mathematical speculation. Reasoning, even that which is apparently formal, is never tautological. Thought, at work, advances; it does not just repeat interminably that *A* is *A*. Meyerson came to emphasize more and more reason's need for something diverse to assimilate, and he tended to define reason not so much by its end, identity, as by its activity, identifying. Reason is thus essentially divided against itself. This is the epistemological paradox.

Meyerson later extended these views to other domains, from scientific reason to philosophical reason, from the modern physicist to primitive man and the medieval thinker; but they were first suggested to him by reflection upon classical science. Have the revolutions in physics served to confirm or contradict them? In *La déduction relativiste* (Paris, 1925), Meyerson easily showed that relativity theory was inspired throughout by the same ideal of objectivization and geometrization. Like Parmenides's sphere or René Descartes's world, Albert Einstein's universe is resorbed into space. However, quantum physics, because it sets bounds to continuity and objectivity, contains something "unassimilable." Meyerson believed, nonetheless, that quantum theory, in the interpretation given it by the Copenhagen school, was a passing "aberration," and that as soon as the physicists recognized the possibility of doing so, they would hasten to return to traditional views—a conjecture that was in part subsequently verified.

If the detail is rich, the broad outlines of Meyerson's philosophy are simple and clear. It enjoyed great prestige about 1930. Since then, it has been somewhat overshadowed by the philosophy of the scientific theorists of the Copenhagen school, although Louis de Broglie retains the high estimate of it stated in his preface to Meyerson's *Essais.* Meyerson's philosophy has also been neglected because of the general shift of interest among contemporary philosophers from epistemological to existential problems.

See also Descartes, René; Einstein, Albert; French Philosophy; Identity; Lalande, André; Parmenides of Elea; Zeno of Elea.

Bibliography

Meyerson's writings include the small work *Réel et déterminisme dans la physique quantique* (Paris: Hermann,

1933) and a posthumously published collection, *Essais* (Paris: J. Vrin, 1936). An English translation of *Identité et réalité* by Kate Loewenberg appeared under the title *Identity and Reality* in London and New York in 1930.

See also André Lalande, "L'épistémologie de M. Meyerson," in *Revue philosophique de la France et de L'étranger* 96 (1922): 259–280; Léon Brunschvicg, "La philosophie d'Émile Meyerson," in *Revue de métaphysique et de morale* 33 (1926): 39–63; George Boas, *A Critical Analysis of the Philosophy of Émile Meyerson* (Baltimore: Johns Hopkins Press, 1930); Jacob Loewenberg, "Meyerson's Critique of Pure Reason," in *Philosophical Review* 41 (1932); Albert E. Blumberg, "Émile Meyerson's Critique of Positivism," in *Monist* 42 (1932): 60–79; André Metz, *Meyerson, une nouvelle philosophie de la connaissance,* 2nd ed. (Paris, 1934); Thomas R. Kelly, *Explanation and Reality in the Philosophy of Émile Meyerson* (Princeton, NJ: Princeton University Press, 1937), with bibliography; *Bulletin de la Société française de philosophie* (April 1961), an issue devoted to Meyerson in celebration of the centenary of his birth.

Robert Blanché (1967)
Translated by Albert E. Blumberg

MICHAEL SCOT

See *Scot, Michael*

MICROCOSM

See *Macrocosm and Microcosm*

MIDDLETON, CONYERS

(1683–1750)

Conyers Middleton was an English historian and clergyman; he entered Trinity College, Cambridge, in 1700. He took orders in the Church of England and became a fellow of his college, but he had to resign his fellowship at the time of his first marriage in 1710. He held various livings but never obtained any considerable preferment in the church. The course of Middleton's life unfortunately provides several grounds for questioning his integrity and ingenuousness.

Middleton's first major publication was *A Letter from Rome, showing an exact conformity between Popery and Paganism* (London, 1729). His theme was certainly not entirely original. It can, for instance, be traced to Part IV of Thomas Hobbes's *Leviathan* (1651), and there is even some suspicion of plagiarism at the expense of a little-known French treatise, *Conformité des cérémonies modernes avec les anciennes* (Leiden, 1667). What was remarkable was the force and skill with which Middleton traced the relics of the worship of Vesta in the cult of the Virgin and deployed passages from the Christian Fathers that excoriated as heathen such practices as the erecting of votive tablets or the use of holy water.

Daniel Waterland, in his *Scripture Vindicated* (London, 1731–1732), had attacked the deist Matthew Tindal's *Christianity as Old as the Creation* (London, 1730). In 1731 Middleton published an anonymous *Letter to Waterland,* in which he urged that it was unwise to insist on the literal truth of every sentence in the Bible, and in particular ridiculed bits of the book of Genesis. His authorship was discovered, and during the ensuing uproar the public orator of Cambridge was heard to cry for a book burning. Middleton next wrote a very profitable *Life of Cicero*; in this instance the charge of plagiarism seems to have been borne out.

After writing an *Introductory Discourse* (1747), Middleton published *A Free Enquiry into the Miraculous Powers, which are supposed to have subsisted in the Christian Church from the Earliest Ages, through several successive Centuries* (London, 1748). Coincidentally, David Hume's first *Enquiry,* containing the section "Of Miracles," which later became notorious, was published in the same year. Many years later, in *My Own Life* (London, 1777), Hume confessed his chagrin: "On my return from Italy, I had the mortification to find all England in a ferment, on account of Dr. Middleton's *Free Enquiry,* while my performance was entirely overlooked and neglected."

There was every reason to compare the two books, for the tendency of both was to undermine belief in the miraculous. But whereas Hume was raising methodological difficulties about the possibility of providing adequate historical proof of such occurrences, especially in a religious context, Middleton was concerned primarily with the historical evidence actually available. His argument was addressed in the first instance to those, including the great majority of educated Protestants, who believed both that the occurrence of miracles was a guarantee of religious truth and that the age of miracles was now past. This position was obviously precarious, for where precisely was the crucial dividing line to be drawn? Middleton directed his onslaught at this weak point. It was, as Leslie Stephen said, "incomparably the most effective of the whole deist controversy." Although Middleton himself never ventured to question the miracle stories of the New Testament, he attacked the credibility of similar accounts in the early Christian church. In a series of damaging quotations, he displayed the credulity of the Fathers, including some of the most respected, such as St.

Augustine, and even cited passages in which others seem to have been deliberately approving pious frauds. The impact of Middleton's attack would have been smaller on a position that was less inherently precarious. Arguments of this kind would not have been effective, for instance, with Protestant "enthusiasts" such as the Wesleys or with the Roman Catholics, who insisted that the age of miracles was not past. As a historian, Middleton displayed the faults characteristic of his period, particularly the naive view that stories must be either wholly and straightforwardly true or else just lies. His importance lies in the contributions he made toward undermining the arbitrary barriers between secular and sacred history.

See also Augustine, St.; Hobbes, Thomas; Hume, David; Miracles; Tindal, Matthew.

Bibliography

Apart from the works of Middleton mentioned in the text, see Sir Leslie Stephen's article on Middleton in the *Dictionary of National Biography* (London and New York, 1909), Vol. XIII. pp. 343–348, as well as his *History of English Thought in the Eighteenth Century* (3rd ed., New York: Putnam, 1902), Ch. 4.

Antony Flew (1967)

MIDDLETON, RICHARD OF

See *Richard of Mediavilla*

MIKHAILOVSKII, NIKOLAI KONSTANTINOVICH

(1842–1904)

Nikolai Konstantinovich Mikhailovskii (Mikhailovsky), the Russian philosopher, social thinker, and literary critic, was a theorist of Russian Populism and an exponent of a form of positivism first advanced by his contemporary, Pëtr Lavrov. Mikhailovskii was born near Meshchovsk, Russia, the son of a landowner of moderate means. After his parents' death, he was enrolled in the St. Petersburg Mining Institute in 1856. Expelled in 1861 for leading student protests against the government, he became a writer on social and literary topics for progressive St. Petersburg reviews. From 1869 to 1884 he edited *Otechestvennyye zapiski* (Annals of the fatherland), at that time the chief organ of Russian radicalism. Mikhailovskii was periodically banished from the capital by the tsarist authorities, but he sufficiently tempered the expression of his views to avoid imprisonment and permanent exile. He remained an influential radical spokesman until his death in St. Petersburg.

Mikhailovskii's humanistic, democratic outlook took shape early in his career, under the influence of John Stuart Mill, Pierre-Joseph Proudhon, and the Russian thinkers Aleksandr Herzen and Vissarion Belinskii. The most direct and extensive philosophical influence on Mikhailovskii was that of Lavrov, whose combination of an antimetaphysical positivism with an emphasis on the "subjective," moral demands of the human consciousness provided Mikhailovskii with his basic philosophical orientation. In his numerous philosophical essays, chief of which is *Chto takoe progress?* (What is progress?; 1869–1870), Mikhailovskii strongly developed the ethical foundation and the individualism of this orientation and defended it against the views of Herbert Spencer, Auguste Comte, Charles Darwin, and later against those of Karl Marx and Friedrich Engels.

In opposition to Spencer, Mikhailovskii argued that human progress cannot be understood "objectively," or nonteleologically, and that in general the phenomena of man's historical and social life can only be approached through a "subjective method" that takes into account the feelings and aims of the individual and makes moral evaluations. Mikhailovskii protested the stunting of the individual by the division of labor in modern industrial society, maintaining that the goal of progress should be a more homogeneous social order in which each individual would be able to develop his diverse abilities comprehensively and harmoniously. Against the social Darwinists he maintained that in human society a struggle for survival is neither inevitable nor desirable, and he asserted that as the division of labor was eliminated, economic competition would yield to cooperation. During the last quarter of the nineteenth century, Mikhailovskii was a leading exponent of Russian Populism—a form of agrarian socialism that emphasized the *obshchina*, or peasant village commune.

Like Comte, Mikhailovskii viewed historical progress as occurring in three stages. Adhering to the "subjective method," however, he distinguished these stages by reference to their teleology. In the objectively anthropocentric stage man sees himself as the end or purpose of nature. In the eccentric stage he still finds ends in nature but no longer regards himself as their unique focus. In the subjectively anthropocentric stage man finally realizes that ends or purposes do not inhere in nature but are pro-

duced by him; the individual dispenses with supernaturalism and metaphysics of every sort and relies on his own active energies for the promotion of his moral ideals.

Mikhailovskii's doctrines, and in particular his emphasis on the autonomous moral individual, brought him into sharp conflict with nascent Russian Marxism. In the 1890s his critiques of Marxism were extensively attacked by both Georgii Plekhanov and V. I. Lenin.

See also Belinskii, Vissarion Grigor'evich; Comte, Auguste; Darwin, Charles Robert; Engels, Friedrich; Herzen, Aleksandr Ivanovich; Lavrov, Pëtr Lavrovich; Lenin, Vladimir Il'ich; Marx, Karl; Marxist Philosophy; Mill, John Stuart; Plekhanov, Georgii Valentinovich; Positivism; Proudhon, Pierre-Joseph; Russian Philosophy; Spencer, Herbert.

Bibliography

WORKS BY MIKHAILOVSKII

Dostoevsky, A Cruel Talent. Translated by Spencer Cadmus. Ann Arbor, MI: Ardis, 1978.

Polnoe sobranie sochinenii (Complete works). 10 vols, edited by E. Kolosov. 4th ed. St. Petersburg, 1906–1914.

Selections from "What Is Progress?" and from other essays in *Russian Philosophy*, 3 vols. Edited by James M. Edie, James P. Scanlan, Mary-Barbara Zeldin, and George L. Kline. Chicago: Quadrangle, 1965.

Literaturnaia kritika i vospominaniia (Literary criticism and memoirs). Moscow: Iskusstvo, 1995.

WORKS ON MIKHAILOVSKII

Billington, James H. *Mikhailovsky and Russian Populism*. Oxford: Clarendon Press, 1958.

Kanevskaya, M. *N. K. Mikhailovsky's Criticism of Dostoevsky: The Cruel Critic*. Lewiston, NY: Edwin Mellen Press, 2001.

Zen'kovskii, V. V. *Istoriia Russkoi Filosofii*. 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy.* 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)

Bibliography updated by Vladimir Marchenkov (2005)

MIKI KIYOSHI

(1897–1945)

Miki Kiyoshi, a Japanese philosopher of history and leading intellectual in the stormy years before World War II, was born in Isseimura, Hyogo prefecture. He was a student of Nishida Kitarō and of Hatano Seiichi at Kyoto University. He developed an early interest in the philosophy of history and studied in Germany (1922–1924) under Heinrich Rickert and Martin Heidegger, absorbing also some socialist ideas. In 1927 he accepted a chair of philosophy at Hōsei University, Tokyo, but he had been rejected as a teacher by his alma mater for dubious reasons—he had a love affair with a widow, in his day a more than sufficient reason to be excluded from a state university. Feeling resentment, and moved by the social climate of the time, he became Japan's first spokesman for philosophical Marxism. His essays on historical materialism (1927–1930) created a stir in academic circles and in the general public. His Marxism, however, was strongly colored by Heidegger's *Anthropologie* and by Blaise Pascal's conception of man, two views he had studied as a youth. His later works are not at all Marxist. In 1930 he was briefly imprisoned for contributing money to leftist causes; as a result he had to give up his teaching career and make a living as a social critic. During the crucial years before World War II, as ultranationalism became pervasive, Miki at first held to liberal principles without compromise. In 1936, he joined the Shōwa Research Society, which was led by Prince Konoe Fumimaru and which strove to moderate though not to oppose the mounting militarist trend. As the Shōwa became more and more nationalistic, Miki, though liberal at heart, had to compromise. For opposing Japan's entry into World War II and for aiding prosecuted leftists, he was returned to prison toward the war's end, and there he died.

Miki's best works are *Rekishi tetsugaku* (Philosophy of history; Tokyo, 1932) and *Kōsōryoku no ronri* (The logic of the power of imagination; Tokyo, 1939). In the first work Miki's starting-point is the subjective existential and sensible experience of life. From this he proceeds to formulate the structure of "history-in-the-making." Fundamental experience of life, he says, creates selfhood, the historical subject that is the only maker of history, since in selfhood there are not subjective and objective factors, but only lived experience. *Kōsōryoku no ronri* reflects Miki's use of Immanuel Kant's *Einbildungskraft* ("imaginative power") as it was revived by Heidegger and also reveals the evolution of Miki's thought away from the logos as social rationality that dominated *Rekishi tetsugaku* and toward a major role for pathos, the subjective inspiration that in Japan led to ultranationalist feelings. Miki was perhaps hinting that rationality was losing ground to ultranationalist passion. At any rate, for Orientals, the logic of the imagination, with its creation of myths and of what Miki calls "forms" of technocultural systems, is said to have some advantages, such as artistic inventiveness and creativity, over conceptual knowledge and usual logic. Miki uses terms borrowed from his master Nishida, the originator of the Oriental "logic of field."

See also Hatano Seiichi; Heidegger, Martin; Historical Materialism; Japanese Philosophy; Kant, Immanuel; Marxist Philosophy; Nishida Kitarō; Pascal, Blaise; Rickert, Heinrich.

Bibliography

For works in Japanese, see *Miki Kiyoshi Choshaku-shū* (Miki Kiyoshi's collected works), 16 vols. (Tokyo, 1945–1951), and Miyagawa Tōru, *Miki Kiyoshi* (Tokyo, 1958). For works in English see Gino K. Piovesana, "Miki Kiyoshi: Representative Thinker of an Anguished Generation," in *Studies in Japanese Culture,* edited by Joseph Roggendorf (Tokyo: Sophia University, 1963), pp. 141–161.

Gino K. Piovesana, S.J. (1967)

MILETUS, SCHOOL OF

See *Pre-Socratic Philosophy*

MILHAUD, GASTON

(1858–1918)

Gaston Milhaud, a French philosopher, came to philosophy by way of mathematics, which he taught for nearly ten years in the lycées before becoming a professor of philosophy at the University of Montpellier. In 1909 he went to the University of Paris, where the chair of history of philosophy in its relationship to the sciences was created especially for him.

His courses on Antoine Augustin Cournot and Charles Renouvier were published (*Études sur Cournot,* Paris, 1927; *La philosophie de Charles Renouvier,* Paris, 1927). Under the influence of Paul Tannery, his works on the history of science were at first devoted to Greek science: *Leçons sur les origines de la science grecque* (Paris, 1893) and *Les philosophes géomètres de la Grèce* (Paris, 1900). Later they were extended to include modern science. Examples are *Études sur la pensée scientifique chez les Grecs et chez les modernes* (Paris, 1906); *Nouvelles Études sur l'histoire de la pensée scientifique* (Paris, 1911); and *Descartes savant* (published posthumously, Paris, 1923).

Milhaud was both a historian and an epistemologist. With Henri Poincaré, Pierre Duhem, and Édouard Le Roy he belongs to that group of French scholars who around 1900, following the path opened for them by Émile Boutroux, denounced scientific dogmatism, using as a basis the precise analysis of past and contemporary examples in history of science. They emphasized the role of spiritual initiative, and thus the element of contingency, in the construction of scientific theories. Milhaud himself generally avoided the dangerous words *convention* and *commodité* used by Le Roy and Poincaré. He spoke, rather, of free creations, of the activity of the mind, and of the spontaneity of reason (*Le rationnel,* Paris, 1898). In his thesis, *Essai sur les conditions et les limites de la certitude logique* (Paris, 1894), he maintained that certitude, which is founded on the principle of noncontradiction, is limited to the domain of pure mathematics. He believed that it was thus possible to establish a radical break between the realm of mathematical knowledge and the realm of knowledge of the real world.

However, almost immediately thereafter (2nd ed., 1897), he regretted having shown himself to be too much the logician: "I see today that even in the extreme example of absolute rigor dreamed of by the mathematician, the living and dynamic identity of thinking always takes precedence over the static immobility of the principle of identity." The fundamental concepts and principles of all sciences result from rational decisions that simultaneously transcend both experience and logic, in the sense that they are not determined by either external or internal necessities. Positivism is, therefore, outmoded. A "fourth stage" consists of the liberation of thought from the obstacles imposed on it by the dogmatism of Auguste Comte (*Le positivisme et le progrès de l'esprit,* Paris, 1902). Nonetheless, scientific contributions are not arbitrary, and they have a universal value, in that they have matured on a basis of fact and have gradually imposed themselves upon the mind as a network of relations in which logical exigencies are composed and harmonized with the demands of a practical and aesthetic order.

See also Boutroux, Émile; Comte, Auguste; Cournot, Antoine Augustin; Duhem, Pierre Maurice Marie; French Philosophy; Le Roy, Édouard; Mathematics, Foundations of; Philosophy of Science, History of; Poincaré, Jules Henri; Positivism; Renouvier, Charles Bernard.

Bibliography

For selections from Milhaud, see R. Poirier, *Philosophes et savants français,* Vol. II, *La philosophie de la science* (Paris, 1926), pp. 55–80. A. Nadal, "Gaston Milhaud," in *Revue d'histoire des sciences* 12 (1959): 1–14, has a bibliography.

Robert Blanché (1967)

MILL, JAMES
(1773–1836)

James Mill, a British historian, economist, psychologist, utilitarian philosopher, and father of John Stuart Mill, was born in Scotland but spent most of his adult life in London. His father was a shoemaker, but his mother was ambitious for James to get a good education and to rise to a higher rank in society. He attended the University of Edinburgh, supported by the patronage of Sir John Stuart (1759–1815), for whom John Stuart Mill was named. Mill distinguished himself as a Greek scholar, receiving his MA in 1794. He then studied divinity and was licensed to preach in 1797. He gave some sermons, but by this time he was an agnostic, basing his disbelief in a benevolent deity, according to his son, on the degree of evil in the universe. He did some tutoring in Scotland, but in 1802 he moved to London where he sought to make a living as a writer and editor. He contributed to a wide assortment of newspapers and journals, and, from 1803 to 1806, he edited the *St. James Chronicle* and the *Literary Journal*. The latter was an ambitious periodical that professed to give a summary view of all the leading departments of human knowledge. In 1805 he married Harriet Burrow, and their first child, born in 1806, was John Stuart Mill.

In 1808 Mill made the acquaintance of Jeremy Bentham (1748–1832), the founder of the utilitarian tradition in modern philosophy. Mill adopted Bentham's utilitarian philosophy and used it as the foundation for his writings on government, education, freedom of the press, and other topics. In 1806 Mill began an ambitious project: to write *The History of British India*, emphasizing the social conditions and movements rather than battles and rulers. This was not completed until 1818, but it immediately became the definitive work on the subject and led to Mill being offered a position at India House, from which the East India Company managed British interests in India. He rose to the position of head of the office and served there until his death.

Mill was not only a "disciple" of Bentham. He was a friend and for a time financially dependent on Bentham's support. At times he and his family lived in houses owned by Bentham, and he and his family spent several summers at Bentham's summer houses. On these summer visits Bentham depended on Mill to be his conversational companion. Mill also edited some of Bentham's writings.

One of Mill's life works, and that for which he is now most famous, is the education that he gave his son John Stuart Mill. From infancy John Stuart was tutored by his father, seven days per week, studying in the room where James was writing the *History of British India* and other articles to support the family. At the end of each day they would take a walk at which time John Stuart would report to his father what he had learned, and he was severely reprimanded if he had not gotten it right. At age three John Stuart was learning Greek from vocabulary cards; so he had already learned English. At age eight he began Latin. By the time that he was twelve he had read, in Greek and Latin, enormous tomes of classical literature, as reported in John Stuart Mill's *Autobiography* (1873).

James Mill was active in promoting the Benthamite philosophy in current politics. He was one of the founders of what came to be known as "philosophical radicalism," a force to the left of the two major parties, the Tories and Whigs. The group included such well-known persons as Francis Place (1771–1854), a successful tailor and organizer of London demonstrations by working people; David Ricardo (1772–1823), the economist, who was probably Mill's best friend; and John Austin (1790–1859), the utilitarian jurist. The radicals advocated extension of suffrage to all tax payers, if not universal suffrage; the secret ballot in elections; the removal of tariffs on imported grain and, in general, free trade; and other legislation for the benefit of the mercantile and working classes.

WRITINGS

Mill wrote on a wide variety of topics for a number of periodicals. These show the breadth of his interests and expertise. Subjects included money and exchange, Spanish America, China, General Francisco de Miranda (1750–1816), the East India Company, liberty of the press, Bentham's law reforms, education, prison discipline, slavery, and religious toleration. In 1805 he published a translation of C. F. Villers's *History of the Reformation*. In 1807 he wrote *Commerce Defended*, an answer to a book that claimed that Britain could be independent of commerce. He wrote a number of articles for the supplement to the fifth edition of the *Encyclopaedia Britannica*, which appeared from 1816 to 1823. Some of these articles were later published independently, the most important being those on "Jurisprudence," "Prisons," "Education," and "Government." Mill's *History of British India*, in three volumes, was finished and published in 1818. In 1821 Mill published *Elements of Political Economy*, which he intended as a "schoolbook" based on his teaching Ricardian economic theory to John Stuart Mill. From 1824 to 1826 he contributed to the *Westminster Review*, a periodical started as an organ of the Radicals to answer the *Quarterly Review* of the Tories and the *Edinburgh Review* of the Whigs. In 1829 appeared his

Analysis of the Phenomena of the Human Mind, in two volumes, putting forward his "associationist" psychological theories. His last major work was the *Fragment on Mackintosh*, published in 1835 after a delay caused by Sir James Mackintosh's (1765–1832) death. In it he presents his ethical views in opposition to those of Mackintosh.

PHILOSOPHY

Mill's philosophy is empiricist, assuming that all knowledge ultimately comes from sense experience, including muscular contractions and sensation from bodily organs. He believed that the inductive method, which had been fruitful in the physical sciences, would be equally effective in philosophy. In *Analysis of the Phenomena of the Human Mind* Mill, using the method of introspection, attempts to give a complete analysis of mental phenomena, resolving them into the primitive feelings from which they are derived by association. "Feeling," according to him, includes every phenomenon of the mind. One's experience is either a knowledge of feelings separately or a knowledge of the order in which they follow each other. Some philosophers had claimed that there are feelings not derived from sensations, but Mill thinks that this is a mistake. He follows David Hume in distinguishing between impressions and ideas. "Ideas" are copies of previous "impressions." Impressions, for Mill, are caused by the external world acting in some way on the mind. The philosopher can only classify the various modes in which they present themselves. One's consciousness reveals simply a series of "sensations" and "ideas." The mind is a stream of these phenomena. The connections of ideas are due to association in either "synchronous" or "successive" order.

When Mill turns to an analysis of sensations and ideas exciting to action, he again attempts to resolve them into simple laws. A desire is an idea of a pleasant sensation; an aversion, an idea of a painful sensation; each having tacit reference to a future time. One associates these pains and pleasures with their causes, coming to desire the causes, and one associates these with one's own actions as possible causes. In this theory of action Mill is a psychological hedonist, but he is not a psychological egoist, in one meaning of that term. Although the pleasure or pain is the agent's own pleasure or pain, it may be associated with the pleasure or pain of another person, such that one desires that person's pleasure or pain. This can even be generalized to a love for humanity, such that one has pleasure at the thought of anyone's pleasure. Thus, it can be possible to be motivated to seek the greatest happiness of everyone, the utilitarian criterion of right action. Mill held, however, that actions are right when they are foreseen to produce the greatest happiness, whether or not this is the motive of the action. But the motive to produce the greatest happiness is important in admiring or despising the character of the agent.

At the same time that Mill recognizes the possibility of altruistic action, of an agent finding pleasure in the sacrifice of his or her own good to the greater good of others, he does not rely on this motive in his political philosophy. He argues from the predominance of selfish interests in his arguments for representative democracy. In his article "Government" he starts from the utilitarian premise that the end of government, as of all conduct, is the greatest happiness. He claims that this can be achieved by assuring for all persons the greatest possible quantity of the produce of their labor. Thus, he defends property, if it reflects this objective. Government is people uniting to delegate to a few the power necessary for protecting this legitimate property. The difficult problems of government relate to the means of preventing these few from themselves having an interest contrary to that of the many. The key is representation. The community as a whole cannot desire its own misery, and, although it cannot act as a whole, it can act through representatives. If these representatives can be prevented by adequate checks from misusing their powers, good government is possible. He believes that responsible representation is possible if election is for brief periods, perhaps annual; by secret ballot; and if the right to vote is extensive enough to prevent the class of electors from having an interest contrary to the whole community. One problem that he addresses is that the people do not understand their own interests. His answer is that ignorance is curable, whereas government by a minority class is sure to be bad.

In *Fragment on Mackintosh* Mill engages in a polemic against a moral sense ethical theory, even one based on associationist psychology and a greatest happiness principle. Mackintosh agrees that the criterion of right and wrong is the greatest happiness, but he claims that the moral sense is a feeling produced by the contemplation of right and wrong that becomes an independent unit, no longer resolved into its origin. It becomes a particular faculty, necessary to discern right and wrong. On the contrary, Mill says that no particular faculty is necessary to discern utility. To say that conduct is right is the same thing as to say that it produces greatest happiness. If the moral sense orders conduct opposed to the general happiness, it is so far bad. If it never orders such conduct, then it is superfluous. Mackintosh uses the example of Fletcher of Saltoun to illustrate his point. Fletcher would have sac-

rificed his life to save his country, but would not do anything base to save his country. Mill attacks this. If you refuse to save your country because you think the means base, your morality is immoral. All general rules, he says, imply exception, but only when they conflict with the supreme rule. If a rule for increasing utility diminishes utility in a given case, then it must be broken in that case.

INFLUENCE

Mill was a significant contributor to the liberalism of nineteenth-century Britain. His articles calling for expansion of suffrage, freedom of the press, freedom of religion, free trade, abolition of slavery, state-supported education, and legal and prison reform no doubt had an influence on his contemporaries and the next generation. He was significant in popularizing Bentham's and Ricardo's views. His psychological theories were a foundation on which Alexander Bain and other psychologists sought to use associationism as one element in a more complete psychology. His most significant influence, however, was by way of his son, John Stuart Mill, who reflects, although he significantly revises, the philosophy of his father.

See also Austin, John; Bain, Alexander; Bentham, Jeremy; Democracy; Empiricism; Ethics, History of; Hume, David; Liberalism; Mill, John Stuart; Utilitarianism.

Bibliography

WORKS BY MILL

A Fragment on Mackintosh. London: Baldwin and Cradock, 1835.

Analysis of the Phenomena of the Human Mind (1829). 2 vols. New York: A. M. Kelley, 1967.

The Collected Works of James Mill. New York: Routledge, 1992a.

Political Writings, edited by Terence Ball. New York: Cambridge University Press, 1992b.

WORKS ABOUT MILL

Bain, Alexander. *James Mill: A Biography* (1882). New York: A. M. Kelley, 1967.

Halévy, Elie. *The Growth of Philosophic Radicalism*. Translated by Mary Morris. London: Faber and Gwyer, 1928.

Stephen, Leslie. *The English Utilitarians* (1900). 3 vols. New York: A. M. Kelley, 1968.

Henry R. West (2005)

MILL, JOHN STUART
(1806–1873)

John Stuart Mill, the English philosopher, economist, and administrator, was the most influential philosopher in the English-speaking world during the nineteenth century and is generally held to be one of the most profound and effective spokesmen for the liberal view of man and society. In the belief that men's opinions are the dominant influence on social and historical change, Mill tried to construct and to propagate a philosophical position that would be of positive assistance to the progress of scientific knowledge, individual freedom, and human happiness. Despite numerous flaws in his theories, he succeeded in providing an alternative to existing views on morals and politics and their foundations that was both specific and cohesive enough to give a markedly liberal tendency to social and political opinion, and also sufficiently tolerant and inclusive to gain it access to an extraordinarily large and diverse public. Mill cannot be ranked among the greatest of pure philosophers, either for his originality or for his synthesizing power. His work in logic, however, broke new ground and gave a badly needed impetus to the study of the subject, while his reformulations of classical British empiricism and Benthamite utilitarianism gave these positions a relevance and continuing vitality that they would not otherwise have had.

Although Mill's views on economics will not be discussed in the present article, an excellent summary of them is contained in the article on Mill by F. Y. Edgeworth in Palgrave's *Dictionary of Political Economy.*

LIFE

John Stuart Mill was born in London, the son of James and Harriet Burrow Mill. Outwardly his life was not eventful. He was educated by his father and never attended school, although for a short time he read law with John Austin. In 1823 he became a clerk in the East India Company, where his father was a high official, and worked there until 1858. Eventually he became chief of his department, a post involving considerable administrative responsibility. In 1831 he was introduced to Harriet Taylor, the wife of a successful merchant and mother of several children. Friendship between Mill and Mrs. Taylor rapidly developed into deep though Platonic love, and for the next twenty years they saw each other constantly, despite the increasing social isolation this involved. Mill was convinced that Mrs. Taylor was a great genius: He discussed all of his work with her and attributed to her an enormous influence on his thought. Her husband died in 1849, and three years later she married Mill. In 1858, while the Mills were on a tour of France, Harriet died in Avignon. Mill bought a house nearby so that he could always be near her grave.

In 1857 Mill had written a brilliant defense of the East India Company for the parliamentary debate on renewal of the company's charter. When renewal was not granted, Mill retired, refusing an offer of a position in the government as an official for Indian affairs. In 1865 he was invited to stand for election to Parliament as an independent member for Westminster. He accepted, and although he refused to campaign, contribute to expenses, or defend his views, he won, and served until the next election, in 1868, when he was defeated. Thereafter he spent his time alternately in London and in Avignon, admired and sought after by many, accessible to few. He died after a very brief illness, attended by his wife's daughter Helen, who had looked after him since her mother's death.

EDUCATION AND PHILOSOPHICAL RADICALISM. Until 1826 Mill's thought was completely controlled by his father. James Mill gave him one of the most formidable educations on record, starting him on Greek at the age of three and Latin at eight. By the age of fourteen he had read most of the major Greek and Latin classics, had made a wide survey of history, and had done intensive work in logic and mathematics. He had also been prepared for acceptance of the central tenets of philosophical radicalism, a set of economic, political, and philosophical views shared by the group of reformers who regarded Jeremy Bentham and James Mill as their intellectual leaders. When at the age of fifteen John Stuart Mill read Bentham's *Traité de législation,* it had the effect on him of a religious revelation. It crystallized his thoughts and fixed his aim in life—to be a reformer of the world. Guided by his father, he threw himself into the work of the radicals; he edited Bentham's manuscripts, conducted a discussion group, wrote letters to the press and articles critical of laws, judicial decisions, and parliamentary debates and actions.

DEPRESSION AND CHANGE OF VIEWS. Late in 1826, Mill suffered a sudden attack of intense depression, which lasted for many months. The attack led him to reconsider the doctrines in which he had been raised and to seek other than Benthamite sources of thought. He believed that his capacity for emotion had been unduly weakened by strenuous training in analytic thought, with the result that he could no longer care for anything at all. In the poetry of William Wordsworth he found something of a cure—an education of the feelings that helped to balance the education of intellect given to him by his father. In 1828 he met Gustave d'Eichthal, a French follower of Comte de Saint-Simon, who sent him an early essay by Auguste Comte and a great deal of Saint-Simonian literature. He also met John Sterling, a disciple of Samuel Taylor Coleridge. Mill came to admire both the Saint-Simonians and the Coleridgeans, and he attempted to incorporate into his own thinking what he took to be sound in their doctrines. In 1829 he published nothing at all, but by the following year he had reached a philosophical position that seemed to him far more adequate than the older Benthamism. He never again changed his philosophical views so radically.

COMTE AND SAINT-SIMON. The historical standpoint of the Saint-Simonians, as well as the appreciation of the value of old institutions emphasized by Coleridge, impressed Mill as important additions to Benthamism, which, he thought, simply neglected such factors. He accepted the outlines of the Saint-Simonian–Comtian philosophy of history, and particularly its theory that in social change there is an alternation between "critical" periods, in which society destroys outmoded forms of life and tends toward disintegration, and "organic" periods, in which new forms of common life are evolved and social cohesion is reestablished. He agreed also with the French view that in his own times society had come to the end of a critical period. From Coleridge he learned to think of the cultured class as the leader of opinion in a nation. He also came to believe that the problem he had in common with other intellectuals was that of assisting the world, and especially England, to emerge from the critical period and progress toward a new organic period. Unless this was done, he thought, the tendency toward disintegration might possibly grow too strong to be controlled.

Three important consequences followed from this. First, merely negative remarks upon institutions, laws, and political arrangements were no longer sufficient. Although much remained that needed to be changed, it was necessary now to replace what had been destroyed with something better. Second, the views of those who defended the old and outmoded could no longer be dismissed, in Benthamite fashion, as mere lies used in defense of vested interests. What is now outmoded must, at one stage of historical development, have served a valuable purpose; otherwise it could not have survived. Those who defend it are those who see the good still in it; hence we must seek for the truth in their views, and not merely reject the falsity. The particular vice plaguing social thought is not the tendency to make mistakes of fact or faulty inferences from facts, but the great ease with which data can be overlooked: in a word, one-sidedness. Hence, if we are to obtain sound social views, our greatest need is

for a complete survey of data, and this is possible to achieve only if we can appreciate the truth that our opponents have learned. For each man is naturally one-sided and can overcome this only by education and effort. Third, the tactics of a reformer must be adapted to the period in which he lives. In particular, during a critical period there is no point in promulgating an entire system: no one will listen, and the ideas will not serve to improve social cohesion. One must proceed cautiously, piecemeal, educating one's public as one goes. One must—especially in England, Mill held, where any appearance of system is abhorrent—confine oneself to particular issues, only slowly insinuating more general principles; or else work only from points on which there is general agreement, so as to avoid any shocking appearance of novelty.

This set of views dictated the program that Mill followed for the next twenty or more years. He did not abandon his early epistemology or ethical beliefs, but in developing them he always tried to emphasize their inclusiveness and their constructive power, rather than their critical and destructive powers. He refrained (with one major exception) from publishing a systematic account of his ideas, but wrote instead occasional essays dealing with fairly specific issues, in which he always tried to bring out the value of the books he was criticizing. (These tactics are largely responsible for the common view of Mill as a wavering, halfhearted, muddled thinker, appreciative of what others had to say but holding no clear opinions of his own.) He defended what he held to be sound views on philosophy, but he did not explicitly link these views together, except in his *System of Logic,* which was an entirely different case. Methods of investigation, Mill held, could be relatively neutral as regards political and moral opinion. Since these methods could be discovered from analysis of subjects like physical science, in which there was widespread agreement on results, there was a good chance of obtaining general agreement on the methods. The methods could thus serve as a cohesive, rather than a disruptive, social force.

THE *SYSTEM OF LOGIC*

Mill's *Logic* is in fact by no means neutral with regard to substantive issues. It is the first major installment of his comprehensive restatement of an empiricist and utilitarian position. It presents (sometimes, to be sure, only as "illustration") a fairly complete outline of what would now be called an "empiricist" epistemology, although Mill himself used "empiricist" in a deprecatory sense to mean "miscellaneous information," as contrasted with "scientific knowledge." It begins the attack on "intuitionism" that Mill carried on throughout his life, and it makes plain his belief that social planning and political action should rely primarily on scientific knowledge, not on authority, custom, revelation, or prescription. The *Logic* had a rapid and wide success. Adopted as a text first at Oxford and eventually at Cambridge, it was also read by many outside the universities, including workmen. Its success can be explained in part by its enormous superiority to any book then existing in the field, but credit must also be given to its clear and unmistakable relevance to social problems (and to religious questions: it was attacked as atheistic by some of its earliest reviewers).

With the publication of the *Logic,* Mill took a major step toward showing that the philosophy of experience, which had hitherto been identified primarily as a skeptical position, could offer at least as much in the way of constructive thinking as any other kind of view. His treatment of deductive inference was far more sympathetic to formal logic than that of previous empiricists; and by arguing that, with care, certainty could be attained even in inductive reasoning, he made it plain that empiricism was not committed to a Humean standpoint. Mill held that the philosophy of experience was more likely than any other to encourage the development of society along liberal lines. He therefore held that it was a matter of considerable importance to show that empiricism was a viable alternative to the less progressive views—notably, Scottish commonsense philosophy and German idealism—which were then dominant. The *Logic* succeeded in doing this.

The *Logic* is primarily a discussion of inferential knowledge and of the rules of inference. (The discussion of noninferential, or as Mill also called it, immediate or intuitive, knowledge belongs, in Mill's view, to metaphysics.) It contains six books. In the first two, Mill presented an empiricist theory of deductive inference, and, since mathematics is the chief deductive science, a discussion of the nature of the truth of mathematics, especially of its axioms. In Book III, Mill discussed induction, its grounds, its methods, and its results. Book IV, titled "Of Operations Subsidiary to Induction," contains chapters on observation and description, abstraction, naming, and classification. Book V is a discussion of fallacies. Book VI contains Mill's attempt to extend the methods of the physical sciences, as derived in Book III, to what were then called "moral sciences," that is, psychology and sociology. He argued for the possibility of a science of human nature and action, and assessed the value of the various methods for attaining it. He concluded with a chapter on

the logic of morality, discussing primarily the relation between rules for actions and the factual statements that serve as their foundations.

No adequate summary of the contents of the *Logic* can be given here, but some of Mill's leading views may be indicated.

DEDUCTIVE REASONING. Mill's argument in Book I of the *Logic* is intended to show the mistake of those who say that deductive inference (as found, for example, in the syllogism) is entirely useless because it involves a petitio principii, but at the same time to make it clear that deduction in general is never the source of new knowledge. Mill agreed that the conclusion of a syllogism may not contain more than is contained in the premises and that "no reasoning from generals to particulars can, as such, prove anything, since from a general principle we cannot infer any particulars, but those which the principle itself assumes as known."

It is useless to defend deduction by saying that it shows us what was "implicit" in our premises, unless we can go on to explain how something can be implicitly contained in what we already know. Mill's solution to this problem and his explanation of the value of rules of deduction rest on his view that "all inference is from particulars to particulars." When we reason "All men are mortal; Jones (not yet dead) is a man; so Jones is mortal," our real evidence for the assertion that Jones will die is our knowledge that Smith, Peters, Wilkins, and many other individuals who resemble Jones in many respects did die. We infer from their deaths to his. The general premise that all men are mortal is not itself our evidence. It is rather a note, or register, of the particular evidence on which the conclusion really depends, together with the prediction that what we have found in cases that we have already observed will also hold in similar cases not yet observed. The real inference, Mill thought, comes in constructing the general proposition on the basis of observation of particular cases. Deduction is to be understood as a way of interpreting the note that has been made of our previous inference. It is valuable because misinterpretation is very easy; but it no more gives us new information than do propositions that are true by definition. Such propositions, which Mill called "verbal," only pull out of a word what was previously put into it; and in the same way, a syllogism simply retrieves from a general proposition a particular one that was previously assumed to be in it. Since there is no real progress of thought in deduction, deductive inference is merely *apparent* inference. Induction is the only procedure that gives us nonverbal general propositions that go beyond what has actually been observed. Hence, only in induction do we make *real* inferences.

Mathematical knowledge is no exception to this. Taking geometry first, as the deductive science par excellence, Mill argued that its conclusions are necessary only in the sense that they necessarily follow from the premises from which they are deduced. But the premises themselves—ultimately, the axioms—are grounded on observation and are generalizations from what we have always experienced. (The definitions are in a somewhat different position, although an experiential element is involved in the belief that the entities they define, such as a geometric point or line, really exist.) That two straight lines do not enclose a surface is evident to us every time we look at two straight lines that intersect. The laws of psychology, operating on such experiential data, are sufficient to explain the production in us of the belief that such lines cannot possibly enclose a surface: hence we need not appeal to intuition or to some other nonexperiential source to explain the belief. Even the inconceivability of the denial of the axioms of geometry does not show, Mill argued, that they are not based on experience. For inconceivability is psychological, and the fact that we cannot think of something does not show that that thing cannot exist. Mill went on to offer an account of the way in which arithmetic and algebra are founded on experience. Here the essential point is that groups of four items, for example, may be rearranged into, or formed from, two groups of two items each, or a group of three items together with a group of one item. Seeing that this is always so, we come, through the operation of psychological laws, to believe that 2 + 2, or 3 + 1, *must* be the same as 4. Algebra is simply a more abstract extension of this sort of belief.

With these explanations Mill hoped to show how mathematics can yield propositions that are not merely verbal and that are certainly true of the world of experience, but that do not depend on any nonexperiential sources of knowledge. His account has never been accepted by philosophers as it stands, but there have been some attempts, among thinkers influenced by pragmatism, to work out a philosophy of mathematics along lines analogous to Mill's.

INDUCTIVE REASONING AND SCIENTIFIC EXPLANATION. In Mill's view, induction is clearly of central importance, since it is the only possible source of substantive general propositions. While the details of his theory are complicated, its main lines may be concisely

indicated. All methodical and critical induction rests on the fundamental principle of the uniformity of nature; namely, that what has happened once will happen again, if circumstances are sufficiently similar. Mill thought that this is a factual proposition that is itself derived by a primitive and natural process of induction: we first note a few limited regularities and predict that they will hold in the future. After our predictions come true, we spontaneously generalize, saying that since some events have been found to occur in repeating patterns, all events will be found to occur in repeating patterns. Belief in the uniformity of nature is thus derived from, and resolvable into, belief in the existence of less sweeping patterns of occurrences, or into particular causal laws. Mill defined "cause of a phenomenon" as "the antecedent, or concurrence of antecedents, on which it is invariably and unconditionally consequent." Like the "axiom" of the uniformity of nature, the principle that every occurrence has a cause is confirmed by all our experience. It is, in fact, simply a more precise way of stating the principle of the uniformity of nature. The hope of science is to formulate propositions about specific sequences of phenomena that can be relied on to the same degree as the law of causation. And the problem of methodical induction—which is the core of the problem of scientific reasoning—arises when it is discovered that the simplest method of induction (that of assembling positive instances of a sequence of phenomena and generalizing directly from them) often leads to general propositions that turn out to be false. We then seek ways of obtaining better results. The fundamental technique is to obtain evidence which will allow us to argue as follows: Either *A* is the cause of *a,* or else there are some events which have no cause; and since we are certain that every event has a cause, we may be certain that *A* causes *a.*

According to Mill, there are four inductive methods: the method of agreement, the method of difference, the method of residues, and the method of concomitant variations. He also discussed a combination of the first two, calling it the joint method of agreement and difference. We use the first two methods in this way. If we find that *A* under circumstances *BC* is followed by *abc,* while under circumstances *DE* it is followed by *ade,* then *A* cannot be the cause either of *bc* or of *de,* since they sometimes do not occur when *A* occurs (and hence by the definition of "cause," cannot be caused by it). But *a* occurs under both sets of conditions; hence it could be the effect of *A*: This illustrates the method of agreement. To ascertain if something other than *A* might be the cause of *a* we use the method of difference. Will *BC* without *A* be followed by *a*? If not, we have so far confirmed our view that *A* causes *a,* for, in the cases we have examined, *A* is always followed by *a* and *a* never occurs without being preceded by *A.* Hence, by the definition of "cause," *A* is, so far as our evidence goes, the cause, or part of the cause, of *a*—or else there are events without any regular cause.

Science does not rely upon induction and experiment alone. It is only infrequently, Mill thought, that we will find genuine causal laws, that is, absolutely invariable sequences. More frequently we will find regularities that hold as far as a limited experience shows but which, we have reason to believe, might well not hold under quite different circumstances. These "empirical laws" are not to be considered basic laws of nature. Much of the practical application of science depends on them, but we cannot claim to have truly scientific knowledge until we can deduce empirical laws from basic laws of nature, showing why the combination of circumstances and laws renders inevitable the limitations within which the empirical laws hold. This makes clear the aim of science: to discover laws of nature and empirical laws, and to connect them, in a deductive system, in such a way as to show how the unrestricted laws would give rise to the regularities reported by the empirical laws. The various sciences are differentiated by the ways in which these two types of laws must be discovered and connected. In some sciences it is possible to discover laws of nature directly, deduce what the empirical laws must be, and then proceed to verify the deductions by checking against experimental data. In others, empirical laws are discovered first, and laws of nature are presented as hypotheses to explain them. These alleged laws of nature are then tested by deducing further empirical laws from them and testing these deductions. In any science, however, explanation comes to an end when laws of nature are reached: These are simply ultimate facts that are to be accepted.

THE MORAL SCIENCES. In the last book of the *Logic,* Mill argued that the phenomena of individual or social human life are no exception to the law of causation, and that consequently it must be possible to determine what are the natural laws of human behavior. He investigated the various modes of inquiry used in the different physical sciences to determine which are most suited to this sort of investigation, and he sketched an outline of what a completed science of man will be. Here as elsewhere, Mill thought that "however complex the phenomena, all their sequences and coexistences result from the laws of the separate elements." Since the separate elements in this instance are men, it is the basic laws of psychology from which, when the science is completed, all the laws and regularities concerning social phenomena must be

deduced. Because of the enormous number of interacting elements, however, the complexity of social action is so great that no direct deduction of its regularities from basic psychological laws will be possible. In order to make this deduction it will be necessary first to construct a science of human character that will cover both the development of human character and the tendencies to action of different types of persons. From the laws of this science, which Mill called "ethology," we may hope eventually to get sociological laws. Even then, however, we will at best obtain statements of tendencies toward action, for the enormous number of factors involved in determining social action will not allow any more accurate predictions. Still, Mill held, "knowledge insufficient for prediction may be most valuable for guidance" in practical affairs. His chief interest lay in the possibility of obtaining scientific guidance for the direction of political decisions.

How far, then, had social science actually progressed? Mill thought that the basic laws of psychology were by then well established: they were the laws put forward by psychologists of the associationist school, among whom James Mill was preeminent. But the science of ethology, which John Stuart Mill had hoped to found himself, eluded him, and he gave up work on it shortly after he published the *System of Logic.* Although the absence of the intermediate laws that this science was designed to contribute made impossible the completion of sociology, Mill thought that at least one basic law of social change had been discovered and substantially proven: Comte's law of three stages. One element, Mill argued, is more important than any other single factor in causing change in society: "This is the state of the speculative faculties of mankind, including the nature of the beliefs which … they have arrived at concerning themselves and the world by which they are surrounded. … the order of human progression in all respects will mainly depend on the order of progression in the intellectual convictions of mankind." Comte had shown that opinion always passes through the same three phases. Men first try to understand their universe in theological terms, then in metaphysical terms, and finally in scientific or, as he called them, positive terms. He had also shown that correlated with these three stages of opinion are types of social organization, which change as opinions change. This generalization, for Mill, was enormously important to our understanding of history and to our practical decisions, and up to that time it was the sole example of a well-founded sociological law. But Mill had high hopes that, with work, much progress could be made in constructing a social science; and he looked forward to a time when "no important branch of human affairs will be any longer abandoned to empiricism and unscientific surmise."

EPISTEMOLOGY AND METAPHYSICS

With respect to metaphysics in the contemporary sense of systematic knowledge transcending experience, Mill claimed to have none; and his epistemology consists largely of an account of experiential knowledge in which he intended to show why nothing beyond such knowledge is either possible or necessary. Mill presented an empiricist theory of our knowledge of the external world and of persons which is equally free of the skepticism of David Hume and the theology of George Berkeley. He consequently covered quite thoroughly a good deal of the ground that was gone over again in the discussions among empiricists and logical positivists in the second and third decades of the twentieth century.

AIM AND METHOD. Mill held that we must know some things intuitively, without inference, if we know anything at all, and he rejected skepticism as failing to make a relevant distinction between knowledge and doubt ("In denying all knowledge it denies none"). For if all knowledge were inferential, there would be no firm starting point for inference, and we should be led into a vicious infinite regress of premises. But because whatever can be known only by intuition is beyond the realm of rational discussion and experimental test, such intuitive knowledge is not easily distinguished from dogmatic opinion. Hence, it was Mill's aim to reduce to an absolute minimum the number of points at which intuitions are required. In the *Logic* he argued that no intuitions are necessary for mathematics, logic, or the procedures of natural science. In the *Examination of Sir William Hamilton's Philosophy* (1865), he pursued these questions further and explicitly took up the questions he had claimed to avoid in the earlier work—especially those concerning the foundations and nature of our knowledge of bodies and of minds.

Mill argued that we cannot tell by intuition or by introspection what we know intuitively. In order to distinguish what is directly given to consciousness from what is there as a result of inference, we must try to investigate the *origins* of the present contents of our minds. And again, this cannot be done directly, because the minds of infants are not accessible to us. Hence, Mill concluded, "the original elements can only come to light as a residual phenomena, by a previous study of the modes of generation of the mental facts which are confessedly not original." This is the psychological method that was orig-

inated by John Locke. In using it, Mill attempted always to show how experience, acting in accordance with known laws of psychology, can explain all of our knowledge. If successful, such accounts make unnecessary (and therefore unwarranted, according to sound scientific methodology) any appeal to extraordinary faculties or to nonexperiential sources of knowledge.

MATTER AND MIND. Mill attempted to explain our belief in the existence of matter and in the existence of our own and other minds by using a psychological method. The "Psychological Theory of the Belief in an External World," as he called it, postulates first, a mind capable of expectation (that is, of forming the conception of possible sensations that would be felt if certain conditions were realized), and second, the psychological laws of association. The claim is that these two factors, operating on experienced sensations and reminiscences of them, would generate not only a belief in an external world but, in addition, a belief that this belief was immediate or intuitive. Mill argued first that by an external object we mean only something that exists whether it is thought of or not, that stays the same even if the sensations we get from it change, and that is common to many observers in a way that sensations are not. One's concept of the external world, Mill said, is made up only to a slight degree, at any moment, of actual sensations, but to a large degree of possible sensations—not of what I am sensing, but of what I would sense *if* I moved, or turned my head, and so forth. These possible sensations, moreover, are thought of as being in groups: numbers of them would be present if I did this, numbers of others if I did that. Contrasted with any particular actual sensation, these groups of possible sensations seem stable and permanent. Moreover, there is not very much regularity in the sequences of our actual sensations, but there is considerable regularity associated in our minds with the groups of possible sensations: We will regularly get this sensation following that one if we do this following that. Hence ideas of cause and power, which (as had been argued in the *Logic*) depend on regularity and succession, are associated with the groups of possible sensations, and not with the actual sensations. At this stage we begin to refer any actual sensation to some group of possible sensations, and even to think of the possibilities as the cause or root of the actual sensation. The groups of possibilities, having permanence and causal power, are so different from fleeting actual sensations that they come to be thought of as being altogether different from them. When it finally becomes clear that the permanent possibilities are publicly observable, we have a concept answering in all respects to our definition of externality. Hence, Mill said, matter "may be defined, a Permanent Possibility of Sensation"; this is all, he held, that the plain man believes matter to be, and indeed, Mill shared this belief. Mill's aim, however, was not so much to defend the belief, as to account for it. And his account, which appeals only to psychological laws known to operate in many other kinds of cases, is simpler than accounts that would make the belief in matter an original part of our mind or an intuitive belief: Consequently, he held, it is a better account.

Mill went on to ask how far a similar theory is adequate to account for mind. The theory will work, he thought, to a large extent, since we know nothing of our mind but its conscious manifestations, and since we know other minds only through inference from the similarities of other bodies and their actions to ours. But memory and expectation pose a fatal difficulty. They involve a belief in something beyond their own existence, and also the idea that I myself have had, or will have, the experience remembered or expected. Hence, if the mind is really a series of feelings, it is an extraordinary series, for it is one that is "aware of itself as a past and future." And if it is not this paradoxical series, it is something more than a series—but what that can be we have no idea. Mill concluded that at this point we are "face to face with that final inexplicability at which … we inevitably arrive when we reach ultimate facts," and all we can do is accept the facts as inexplicable. Hence, mind is not simply a permanent possibility of sensation.

Sensations and feelings—the data of experience—are, then, intuitively known; the fact of memory (a consequence of which Mill thought to be expectation) is also known directly; and the kind of link between past and present involved in memory (which Mill took to be the central inexplicable reality about the self) is known directly. Aside from these, there is only one additional inexplicable fact, and that is belief—the fact that there is a difference between contemplating, or imagining, or supposing, and actually believing. Mill rejected his father's analysis of belief, but could develop no adequate account of his own.

ETHICS

According to Mill, agreement on moral beliefs is the most important single factor making for cohesion in society, and where it is lacking society cannot be unified. In his own times he saw and recognized the significance of the first serious widespread breakdown of belief in the Christian moral scheme. He thought it a task of first importance to provide an alternative view of morality that

would be both acceptable to those who still clung, in part, to their older views, and capable of redirecting these older moral attitudes into newer paths. He was a utilitarian in ethics: that is, he held that an action is right if, and only if, it brings about a greater balance of good over bad consequences than any other act open to the agent, and he also believed that only pleasure is intrinsically good and only pain intrinsically bad. Bentham and James Mill had held a similar position, but John Stuart Mill modified their view in a number of ways, attempting always to show that utilitarianism need not be a narrow or selfish view and that it did not force one to rely, for social progress, purely on impersonal institutional arrangements and thereby compel one to leave human personality out of account. By arguing that the utilitarian could appreciate the wisdom embodied in traditional morality as well as offer rational criticism of it, and that he could also accept and account for the high value of self-sacrifice and could make the development and perfection of individual character the key obligation of morality, Mill sought to rebut the most frequent criticisms of the Benthamite morality and thereby make it more generally acceptable. Although his ethical writings (especially *Utilitarianism*) have been much criticized, they contain the most influential philosophical articulation of a liberal humanistic morality that was produced in the nineteenth century.

In his ethical writings, Mill pursued the attack on intuitionism that was so constant a feature of his other work. This issue is especially important with regard to moral problems. Intuitionism, he said in the *Autobiography,* is "the great intellectual support of false doctrines and bad institutions" because it enables "every inveterate belief and every intense feeling … to dispense with the obligation of justifying itself by reason.… There never was such an instrument devised for consecrating all deep-seated prejudices." The intuitionists supposed, Mill believed, that only their view could account for (1) the uniqueness of moral judgments, (2) the rapidity with which the plain man passes moral judgments, and (3) the authority to be given to commonsense moral judgments. To the first point, Mill answered with the theory that moral feelings may have unique properties, just as water has, and yet may still be derived, by a chemical compounding process, from simpler elements that do not have those properties. Hence, so far there is no need to say that these feelings are caused by unique intuitions. To the second point he replied that rapidity of judgment may be due to habit and training as well as to a faculty of intuition. And with regard to the third point, which is the crucial one, he argued that the utilitarian can give at least as good an account as the intuitionist of the authority of common sense in moral matters. Rules such as those that enjoin the telling of truth, the paying of debts, the keeping of promises, and so forth (Mill called these "secondary rules") were taken by him to indicate, not widespread intuitions, but the results of hundreds of years of experience of the consequences of actions. These rules, based on so much factual knowledge, are of considerable value in helping men to make correct decisions when time or data for a full calculation of the results in a particular case are lacking. The wisdom of the ages, thus embodied in the rules and precepts of commonsense morality, is an indispensable supplement to the limited knowledge and almost inevitable one-sidedness of any single person. It is for these reasons, utilitarians claim, that these rules and precepts have a certain cognitive authority. There is no need to appeal to a faculty of intuition to explain the authority, and therefore such an explanation is, from a scientific point of view, unwarranted.

Mill thus gave a prominent place to moral directives other than the utilitarian principle. But he was basically an act-utilitarian, believing that each particular obligation depends on the balance of pleasure and pain that would be produced by the act in question. The utilitarian principle is so abstract, Mill thought, that it is unlikely to be actually used, except in cases where two secondary rules come into conflict with each other. But it serves the invaluable function of providing a rational basis for the criticism of secondary rules (this is brought out especially well in the essay on justice, Ch. 5 of *Utilitarianism*), and there was no doubt in Mill's mind that there can never be a right act that contravenes the principle. This is true even with regard to the rule (to which Mill gave so much emphasis) dictating the development and perfection of individual character. It often seems that Mill placed more stress on individuality, or self-realization, than on general welfare, and critics frequently claim that he contradicted himself by saying that both of these constitute the sole highest good. But there is no contradiction in his views, for he held that self-development is the best way for an individual to work for the common good.

Mill's concern with the problem of free will sprang from his view of the importance of self-development. (He presented this view both in the *Logic* and in the *Examination of Hamilton.*) The doctrine of necessity, which he had been taught to believe, seemed to him to make a man a creature of his environment, and this doctrine depressed and disturbed him for many years. When he realized that the desire to improve oneself could be a powerful motive and that actions dictated by this desire, although not con-

travening the law of causation, are properly said to be due to oneself rather than to one's environment, he felt "as if an incubus had been raised off him." He thought that this view enabled him to make determinism compatible with his emphasis on the individual's responsibility for his own character.

Two aspects of Mill's *Utilitarianism* have been attacked more frequently than any others. The first is his attempt to broaden utilitarianism by making a distinction between *kinds* of pleasure, so that an act producing a smaller amount of a more valuable kind of pleasure might be obligatory, rather than an act producing a larger amount of a less valuable kind of pleasure. This line of reasoning has been said to involve him in flagrant contradictions, or else to be sheer nonsense.

The second aspect is his attempt to give some sort of reasoned support to the utilitarian principle itself, which led G. E. Moore to accuse him of committing the "naturalistic fallacy." Moore thought Mill was trying to give a conclusive proof of a first moral principle, but he was mistaken. Throughout his life, Mill consistently held that no such proof of the principle was possible, either deductively or inductively. There is, however, no agreement as to the manner in which Mill attempted, in the fourth chapter of *Utilitarianism,* to support his first principle so that he would not be open to the same reproach of dogmatism that he had made against the intuitionists. Mill's remarks here are extremely unclear. His problem arises because, while he insisted that there must be a factual basis for moral judgments, he held that moral judgments are different in kind from factual propositions and therefore cannot be strictly derived from them. Although he failed to solve this problem, he at least propounded it in precisely the form in which it has perplexed (not to say obsessed) recent moral philosophers.

SOCIAL AND POLITICAL PHILOSOPHY

Mill was more aware than were the older Benthamites of the importance of nonrational and noninstitutional factors to an understanding of society, and was consequently less disposed to rely on legal and governmental reforms for the improvement of it. He believed in democratic government, but he was convinced that it could not work well unless the citizens who lived under it were reasonably well educated, tolerant of opposing views, and willing to sacrifice some of their immediate interests for the good of society. He was profoundly worried about the tendency of democracies to suppress individuality and override minorities: Indeed, this, and not the problem of forcing those who control government to work for the interests of the people, seemed to him the crucial problem of his times. Hence, in his writings on social and political philosophy, his central concern was to show the importance of personal freedom and the development of strong individual character and to devise ways of encouraging their growth.

ECONOMIC THEORY. With regard to economic theory Mill at first supported a general policy of laissez-faire, but increasing awareness of the uselessness to the individual of political freedom without economic security and opportunity led him to reexamine his objections to socialism. By the end of his life he had come to think that as far as economic theory was concerned, socialism was acceptable. His reservations about it sprang from his fear that it would give overwhelming strength to the tendencies of the age toward suppression of individuality.

ON LIBERTY. Mill thought that his essay *On Liberty* was the most likely of all his works to be of enduring value. In it he maintained the view, which he had expressed as early as 1834, that "the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection." Mill argued for this view especially in regard to freedom of thought and discussion. "We can never be sure," he wrote, "that the opinion we are endeavoring to stifle is a false opinion; and if we were sure, stifling it would be an evil still": These are the lines of his defense, which rests ultimately on his assessment of the importance of sociological knowledge to the direction of social action and on his view of the peculiar difficulties in obtaining it. In the third chapter, Mill argued at length for the importance of "individuality," which, he held, comes from, or indeed is identical with, continued effort at self-development. Even eccentricity is better, he held, than massive uniformity of personality and the stagnation of society that would result from it. Mill's strong emphasis on this point stems from his conviction, here strongly influenced by Alexis de Tocqueville, that the chief danger of democracy is that of suppressing individual differences and of allowing no genuine development of minority opinion. Democratic tyranny would be far worse, he held, than aristocratic or despotic tyranny, since it would be far more effective in utilizing the most efficient of means of social control, the pressure of public opinion. Against this the only reliable safeguard would be the development of personalities strong enough to resist such pressures.

REPRESENTATIVE GOVERNMENT. In more specifically political matters the same concerns are evident. Mill

defended representative democracy, but not solely on the grounds used by the older Benthamites. Representative government, he held, is ideally the best form of government because it does more to encourage the growth and development of individuality than any other form of government. By leading people to participate in the processes of governing, representative government makes them more active, intelligent, and well rounded than even the best-intentioned of despotisms could. It thereby gives them vitally important moral training, by cultivating their public sympathies, strengthening their habit of looking at social questions from an impersonal point of view, and aiding their identification of personal interests with the interests of society. Care must be taken, however, to get a true democracy, one in which minorities as well as majorities are represented. For this reason Mill enthusiastically endorsed Thomas Hare's scheme of proportional representation. He also favored plural voting, which would allow educated and responsible persons to have more influence than the uneducated, by giving the former several votes. Mill's view of the function of the representative also shows his concern to get as much intelligence as possible into government. A properly educated constituency, he held, would be able and willing to select the best men available; and since those elected would be better informed and wiser on particular issues than the electorate, it would be absurd to bind the representatives to anything but a very general agreement with the beliefs and aims of the electors.

INDIVIDUALS AND SOCIETY. Mill is frequently criticized for overlooking the organic elements in society and for thinking of society as a mere aggregate of units in which each unit is what it is regardless of its membership in the whole. Mill certainly held this view as far as the most fundamental laws of psychology are concerned. But his view of individual character involves new considerations. Individuals, he held, are radically affected by their membership in society and inevitably formed by the customs, habits, morality, and beliefs of those who raise them. There is, however, no impersonal assurance, metaphysical or otherwise, that the individual will feel himself an organic member of any group. He will do so, Mill thought, only if he is educated to do so. Mill cannot be accused of underestimating the importance of ensuring that men are so educated, and it is not clear that an organic theory has anything better to offer on a practical level.

RELIGIOUS VIEWS

Mill maintained for the most part a determined silence on religious questions. Although he had written "On Nature" and "The Utility of Religion" by 1858, and although he lived during a period of increasingly free discussion of all possible religious subjects, he thought that the British public would not listen patiently to what he had to say on these questions and that he could not publish his views without alienating readers and losing public influence. And this, as he made quite clear in his correspondence with Auguste Comte, he was determined not to do. Despite his precautions, however, he was generally taken to be atheistic, and he was sometimes criticized for not openly stating the views that, so it seemed, he insinuated but did not defend. The consternation of his followers and the delight of his opponents was therefore considerable when it became apparent from the posthumously published *Three Essays on Religion* (1874) that Mill did not entirely condemn religious aspirations and hopes and even thought that there might be some faint possibility of the existence of rational support for a religious view of the world. Admirers felt betrayed, and religious critics proclaimed that Mill's secular education and materialistic position here issued in collapse and evident moral and intellectual bankruptcy.

GOODNESS OF GOD. Mill's most famous pronouncement on religion occurs not, however, in the *Three Essays,* but in the *Examination of Hamilton.* Discussing the use made by one of Hamilton's philosophical followers, Henry Mansel, of Hamilton's view that we cannot know the Absolute, Mill particularly criticized Mansel's theory that even the moral terms we apply to God do not mean what they mean when we apply them to men. Mill objected to this theory in the name of logic: If terms are not to be used in their usual sense, they ought not to be used at all. But, more strongly, he went on to say that a being, no matter how powerful, whose acts are not sanctioned by the highest human morality conceivable, is not deserving of worship. If Mill were convinced of the existence of such a being he would not worship him. "I will call no being good," Mill proclaimed, "who is not what I mean when I apply that epithet to my fellow creatures, and if such a being can sentence me to hell for not so calling him, to hell I will go."

NATURE. Of the *Three Essays,* the first two, at least, show no reversal or collapse of Mill's views. In "On Nature" Mill argued that the maxim "Follow Nature" is of no use as a guide to action. For "Nature" either means "everything that happens, good as well as bad," in which case it offers no guidance whatsoever; or it means "what happens without any human interference," and in that case the maxim is self-contradictory. Nature in the second

sense, Mill went on to argue, offers at least as much evil to our observation as good; it is rather a challenge to amendment than an ideal for imitation. From this, two conclusions follow. First, it is our job to improve nature, especially human nature; for it is only insofar as men have intervened to change things that the world has become civilized, safe, and happy, even to the limited extent that it has. Human virtues are not natural: They are preeminently the results of cultivation. Even justice is an artificial virtue, Mill said, and the idea of natural justice does not precede, but follows, it. Second, in view of the suffering and ugliness presented by much of the natural world, the only religious view that is at all tenable is one which holds that the deity is not omnipotent, that "the Principle of Good *cannot* at once and altogether subdue the powers of evil," and that, consequently, men should think of themselves as the far from useless helpers of a limited but benevolent God.

UTILITY OF RELIGION. In "The Utility of Religion," Mill argued that much of the social usefulness attributed to religion is actually due to the influence of a widely accepted and instilled moral code, and to the force of public opinion guided by that code. The belief in the supernatural origin of morality may once have helped it to gain acceptance, but is no longer needed, or indeed, even effectual, in maintaining this acceptance. The effect of religion on individuals springs largely from our need to have ideal conceptions that move us to action. "The essence of religion is the strong and earnest direction of the emotions and desires towards an ideal object, recognized as of the highest excellence, and as rightfully paramount over all selfish objects of desire." But a religion of humanity, Mill argued, can have this effect to an even greater extent than a supernatural religion. The religion of humanity would cultivate our unselfish feelings and would free us from any need for intellectual juggling or willful blindness with regard to its tenets, since it would rather point out than deny the evil in the world and urge us to work to remove it.

GOD. Thus, the first two essays of the *Three Essays* together suggest that the alternative to a supernatural religion is not simple acceptance of Nature, but the construction of an alternative way of living based on education and convention; and these themes are to be found throughout Mill's thought. The third essay, "Theism," drafted from 1868 to 1870, which assesses arguments in support of a supernatural religious view, seems to make more concessions to traditional religiosity than the other essays; but even these are slight. In this essay, Mill discussed the possibilities of rational support for supernatural beliefs. Dismissing all a priori reasoning, he found only the Argument from Design at all convincing, and this argument gives us at best "no more than a probability" that some intelligent creator of the world exists. For the same evidences that thus support the existence of a creator also go to show that he was not omnipotent and do not prove that he was omniscient. Mill suggested that we think of a limited deity faced with the independent existence of matter and force. To this picture of a Platonic demiurge, Mill thought we are entitled to add that benevolence may have been one (although surely not the only) moral attribute of the creator. But Mill emphasized strongly the importance of the work of man in improving the world. "If man had not the power," he said, "by the exercise of his own energies for the improvement both of himself and of his outward circumstances, to do for himself and other creatures vastly more than God had in the first instance done, the Being who called him into existence would deserve something very different from thanks at his hands."

IMMORTALITY AND MIRACLES. Mill argued that there is no evidence for the immortality of the soul and none against it. After a lengthy discussion of Hume's arguments on this point he found that roughly the same is true of miracles. But in each case he pointed out that there is room for *hope*: One may, if it is comforting and encouraging, hope that the soul is immortal and that the revelations attested by miracles are true. And it is this point more than any other in the essay that upset Mill's admirers. For while he concluded that the proper rational attitude to supernatural religion is skepticism rather than belief or positive disbelief and that "the whole domain of the supernatural is thus removed from the region of Belief into that of simple Hope," he also held that it may be valuable and justifiable to encourage religious hopes. This, he said, can be done without impairing the power of reason; and indulgence in such hopes may help some men to feel that life is more important and may strengthen their feelings for others. Furthermore, to construct a picture of a person of high moral excellence, such as Christ, and form the habit of seeking the approval of this person for one's acts, may aid that "real, though purely human, religion, which sometimes calls itself the Religion of Humanity, and sometimes that of Duty." Critics may wish to call these views objectionable, but in Mill at least they are not inconsistent. They hark back to his early discovery of the importance of cultivating the feelings and develop the further implications of his idea of the moral importance of educating the emotions. His

assessment of the degree to which scientific support can be given to a supernaturalist theory by evidences of design, low though it is, may seem far too high; but his interest in the theory of a limited deity with whom we must cooperate to bring about improvement in the world is hardly great enough or personal enough to lend credence to the accusations that he had undergone an emotional collapse.

See also Bentham, Jeremy; Bradley, Francis Herbert; British Philosophy; Causation; Coleridge, Samuel Taylor; Comte, Auguste; Empiricism; Hamilton, William; Liberty; Locke, John; Logic, History of; Mansel, Henry Longueville; Mill, James; Mill's Methods of Induction; Moore, George Edward; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Utilitarianism.

Bibliography

WORKS BY MILL

Mill's works have not yet been collected. Even the projected University of Toronto Press edition of his Works will probably not contain all of them. Mill's own *Bibliography,* edited by M. MacMinn, J. R. Hainds, and J. M. McCrimmon (Evanston, IL, 1945), is not quite complete.

Mill's books (all of which were published in London, unless otherwise noted) are as follows: *System of Logic,* 2 vols. (1843; 8th ed., 1872); *Essays on Some Unsettled Questions of Political Economy* (1844; written 1830–1831); *Principles of Political Economy,* 2 vols. (1848; 7th ed., 1871; variorum ed., W. J. Ashley, ed., 1909); *On Liberty* (1859); *Dissertations and Discussions,* periodical essays, 2 vols. (1859), 4 vols. (1875); *Considerations on Representative Government* (1861); *Utilitarianism,* reprinted from *Frasers Magazine,* 1861 (1863); *An Examination of Sir William Hamilton's Philosophy* (1865; 6th ed., 1889); *Auguste Comte and Positivism* (1865); *Subjection of Women* (1869; written in 1861); *Autobiography* (1873; more complete edition, J. J. Coss, ed., New York, 1924).

Among Mill's shorter writings of philosophical interest (most reprinted in *Dissertations and Discussions*) are the following: "Whately's Elements of Logic," *Westminster Review* (1828); "The Spirit of the Age," in the *Examiner* (1831), included in *The Spirit of the Age,* edited by F. Hayek (Chicago, 1942); "Prof. Sedgwick's Discourse" (1835); "Civilization" (1836); "Bentham" (1838); "Coleridge" (1840); "M. de Tocqueville on Democracy in America" (1840); "Bailey on Berkeley's Theory of Vision" (1842); "Michelet's History of France" (1844); "Dr. Whewell on Moral Philosophy" (1851); "Bain's Psychology" (1859); "Austin on Jurisprudence" (1863); "Plato" (1866); "Inaugural Address to the University of St. Andrews" (1867); "Berkeley's Life and Writings," *Fortnightly Review* (1871); "Grote's Aristotle" (1873); "Chapters on Socialism," *Fortnightly Review* (1879), reprinted as *Socialism,* edited by W. D. P. Bliss (Linden, MA, 1891).

Of Mill's literary essays, the best known are "What Is Poetry?" and "The Two Kinds of Poetry," in *Monthly Repository* (1833), reprinted in part in *Dissertations and Discourses* as "Thoughts on Poetry and Its Varieties."

WORKS ON MILL

Life

For Mill's life, see his *Autobiography*; F. E. Mineka, ed., *Earlier Letters,* 2 vols. (Toronto, 1963); H. S. R. Elliott, ed., *Letters,* 2 vols. (1910); J. Stillinger, ed., *Early Draft of John Stuart Mill's Autobiography* (Urbana: University of Illinois Press, 1961). See F. Hayek, ed., *John Stuart Mill and Harriet Taylor* (1951), for their correspondence. See also the standard M. St. John Packe, *The Life of John Stuart Mill* (1954); A. Bain, *John Stuart Mill* (1882); and W. L. Courtney, *Life of John Stuart Mill* (1886); H. O. Pappe, *John Stuart Mill and the Harriet Taylor Myth* (Melbourne, 1960); A. W. Levi, "The Writing of Mill's Autobiography," in *Ethics* 61 (1951).

Among many estimates of Mill's life and character are those by R. H. Hutton, reprinted in *Criticism on Contemporary Thought and Thinkers,* Vol. 1 (1894); J. Martineau, in *Essays* 3 (1891); J. Morley, in *Critical Miscellanies* 2 (1877); B. Russell, in *Proceedings of the British Academy* (1955); W. Ward, in *Men and Matters* (1914).

General Works

For general commentary on the thought of Mill see Sir Leslie Stephen, *English Utilitarians,* Vol. 3 (1900); R. P. Anschutz, *Philosophy of John Stuart Mill* (Oxford, 1953); Karl Britton, *John Stuart Mill* (1953).

Logic

See O. A. Kubitz, *Development of John Stuart Mill's System of Logic,* Illinois Studies in the Social Sciences, VIII (Urbana, IL, 1932); R. Jackson, *Deductive Logic of John Stuart Mill* (Oxford, 1941); W. Whewell, *Of Induction, with especial reference to Mr. J. Stuart Mill's System of Logic* (1849), and see E. A. Strong, "W. Whewell and John Stuart Mill," *Journal of the History of Ideas* (1955). Classic criticisms include: T. H. Green, "The Logic of John Stuart Mill," *Works,* Vol. II (1886); F. H. Bradley, *Principles of Logic* (Oxford, 1883), Bk. II, Part II, Chs. 1–3; W. S. Jevons, "John Stuart Mill's Philosophy Tested," reprinted in *Pure Logic* (1890).

Metaphysics

Among older studies of interest are: W. L. Courtney, *The Metaphysics of John Stuart Mill* (1879); C. M. Douglas, *John Stuart Mill: A Study of His Philosophy* (Edinburgh and London: Blackwood, 1895); J. McCosh, *An Examination of Mr. John Stuart Mill's Philosophy* (London and New York, 1866); and John Grote, *Exploratio Philosophica* (Cambridge, U.K.: Deighton Bell, 1865; 2 vols., 1900). Few recent discussions center explicitly on Mill.

Ethics and Utilitarianism

E. Halévy, *La formation du radicalisme philosophique,* 3 vols. (Paris: F. Alcan, 1901–1904), translated into English by Mary Morris as *Growth of Philosophic Radicalism* (London: Faber and Gwyer, 1928), is the basic study of the development of Benthamite doctrine; see also E. Albee, *History of English Utilitarianism* (1900) and J. Plamenatz, *The English Utilitarians* (Oxford: Blackwell, 1949). Especially valuable older critical works are John Grote, *Examination of the Utilitarian Philosophy* (Cambridge, U.K., 1870) and F. H. Bradley, *Ethical Studies* (Oxford, 1876), Ch. 3. Recent discussions start from the criticisms of G. E. Moore,

Principia Ethica (London: Cambridge University Press, 1903), Chs. 1 and 3. Compare J. Seth, "Alleged Fallacies in Mill's Utilitarianism," in *Philosophical Review* 17 (1908); E. W. Hall, "The 'Proof' of Utility in Bentham and Mill," in *Ethics* 9 (1949); J. O. Urmson, "Interpretation of the Moral Philosophy of John Stuart Mill," in *Philosophical Quarterly* 3 (1953). I. Berlin's lecture, "John Stuart Mill and the Ends of Life" (London, 1962), is more general.

Political Philosophy

See G. H. Sabine, *History of Political Theory,* 3rd ed. (New York: Holt, Rinehart and Winston, 1961); M. Cowling, *Mill and Liberalism* (Cambridge, 1963). J. F. Stephen, *Liberty, Equality, Fraternity* (1873) is an interesting early attack; others are summarized in J. C. Rees, *Mill and His Early Critics* (Leicester, U.K., 1956). B. Bosanquet, *Philosophical Theory of the State* (1899) and D. G. Ritchie, *Principles of State Interference* (1891) present representative criticism. J. H. Burns, "John Stuart Mill and Democracy," in *Political Studies* 5 (1957), traces the development of Mill's views. For criticisms of Mill's views on sociological method, see K. Popper, *Open Society and Its Enemies,* 2 vols. (London: Routledge, 1945), Ch. 14, and P. Winch, *Idea of a Social Science* (London: Routledge and Kegan Paul, 1958), especially Ch. 3.

J. B. Schneewind (1967)

MILL, JOHN STUART [ADDENDUM]

The most important development in John Stuart Mill scholarship of the past half century is the publication of the *Collected Works of John Stuart Mill* in thirty-three volumes (1963–1991), with John M. Robson as general editor. This is a monumental publication by the University of Toronto Press, which will provide data for Mill scholars in years to come. The seven volumes of letters and many volumes of essays, speeches, and journals, show that most of his writing was not on narrowly philosophical topics: Much of it was on concrete political issues of his day. There are four volumes of newspaper writings. The *Collected Works* also makes available all of the revisions in successive editions of his major works, such as *System of Logic* (Vols. VII–VIII) and *Principles of Political Economy* (Vols. II–III, and it makes available out-of-print works such as his *An Examination of Sir William Hamilton's Philosophy* (Vol. IX). The exhaustive index in the final volume enables scholars to find Mill's views on various topics scattered throughout his writings.

Another development is the publication of a periodical devoted to utilitarian studies. *The Mill Newsletter* began publication in 1965 by the University of Toronto Press under the editorship of John M. Robson. It carried long and short articles, news of new and forthcoming books and articles, and a continuing bibliography of works on Mill. In 1989 it merged with *The Bentham Newsletter* to become *Utilitas: A Journal of Utilitarian Studies*, now being published by Cambridge University Press. It has provided a vehicle for Mill scholarship including but not limited to his philosophy.

The most substantial studies of the totality of Mill's philosophy are *John Stuart Mill*, by John Skorupski (1989) and the collection of essays that he edited in *The Cambridge Companion to Mill* (1998). In the former, Skorupski gives a critical but sympathetic account of Mill's philosophy of language, his philosophy of science, his philosophy of mathematics, and his epistemology and metaphysics as well as a discussion of his moral and political philosophy. In *The Cambridge Companion to Mill*, the same areas are addressed by various contributors. Mill's radical empiricist theory of mathematical truth has been dismissed by most philosophers of mathematics since his time. But the essay by Philip Kitcher (Skorupski 1998, pp. 57–111) gives it a sympathetic interpretation.

The most widely read philosophical works of Mill continue to be his essays *Utilitarianism* (Mill 1963–1991, vol. X, p. 203–206) and *On Liberty* (Mill 1963–1991, vol. XVIII, p. 215–310). Debates concerning utilitarianism in the last half century, such as the distinction between act-utilitarianism and rule-utilitarianism and the plausibility of each, have included controversies over the interpretation and plausibility of Mill's position on these issues. Also, those attacking or defending liberalism have inevitably included references to Mill's essay as one of the most representative statements of the liberal position. With the development of feminist philosophy, his essay *The Subjection of Women* (Mill 1963–1991, vol. XXI, p. 259–348) has also received renewed attention as an early feminist statement, sometimes dismissed as the *liberal feminist* position, but sometimes defended against its critics.

Two controversial topics in Mill's utilitarianism continue to receive a focus of attention: his distinction between pleasures on grounds of superiority or inferiority of quality as well as quantity and his alleged *proof* of the principle of utility. In the early part of the twentieth century, the first of these was generally regarded as either inconsistent with his hedonism or as nonsense, and the second was regarded as a classic case of fallacious reasoning. In the last half century, these have been defended, although not always in the same ways. Some "friends" of Mill have tried to reduce the distinction of qualities to a quantitative distinction; others have insisted that Mill is

correct in recognizing the phenomenal diversity of pleasurable experiences. But even among the latter there is disagreement about whether Mill is correct in correlating the distinction with the distinctively human, as opposed to nonhuman animal, faculties and whether qualitatively distinct pleasures are consistently preferred by those who are qualified by experiences of both. Important works on these topics are found in books by Wendy Donner (1991) and Henry West (2004). Donner emphasizes that those qualified by experience to judge the qualities of pleasure are not simply those who have experienced different pleasures but those whose experience has been developed by education and enlightenment. Mill's *proof* has been the subject of numerous interpretations and controversy. It is no longer dismissed as a collection of fallacies, but whether it is a sound argument with plausible assumptions is still a matter of great debate. West defends it as a sound argument.

The consistency between Mill's apparently hedonistic utilitarianism and his essay *On Liberty* has been another topic of extensive discussion. Here again, more recent discussion has been more friendly to Mill but with differences in interpretation. Some commentators have claimed consistency for him by a reinterpretation of his utilitarianism to make it nonhedonistic, with a conception of happiness that essentially involves the free exercise of rational capacities. Others have seen in Mill's psychological assumptions, with a complex phenomenal account of pleasure, including *higher* and *lower* and the necessity for self-development as a necessary condition for the higher pleasures, a basis for consistency that remains hedonistic. Mill's *On Liberty* attempts to distinguish between conduct that concerns others and that concerns only oneself. Strictly construed, very little conduct concerns only oneself. Studies of *On Liberty* by C. L. Ten (1980), John Gray (1983, rev. ed. 1996), and J. C. Rees (1985) have reinterpreted the distinction in terms of conduct concerning the *interests* of self or others. Mill is seen to be holding the view that there is a right to liberty, which is a right to autonomy. There is controversy, however, over the substance of this right and also over the *harm principle* which limits it.

Whether Mill was a rule-utilitarian was one of the questions that generated the distinction between act-utilitarianism and rule-utilitarianism. The essay by J. O. Urmson (1953) interpreting Mill as a rule-utilitarian has been challenged and supported by citations from Mill texts both pro and con. A middle position, argued by Fred R. Berger (1984) and others, is that Mill endorsed a strategy for achieving the greatest happiness that was in practice rule-utilitarian but that Mill seemed to think that if all hidden utilities were taken into consideration, there would be no conflict between the two positions. Acts that violate useful rules weaken the rules and undermine the rule-abiding character of the agent. Acts that form part of a collection of acts that have bad consequences can theoretically be assigned a fraction of those bad consequences. Whether these moves are adequate to remove the conflict is suspect.

Perhaps most significant as a way of resolving the conflict in favor of a rule-utilitarian interpretation has been the attention drawn to the importance of sanctions in Mill's theory of morality. Most commentators make a distinction between act-utilitarianism as a criterion of right action and act-utilitarianism as a decision procedure for action. It is generally recognized that Mill rejected act-utilitarianism as a decision procedure in all cases, but some commentators, such as Roger Crisp (1997), still hold that he was an act-utilitarian with regard to the criterion of right action. Essays by David Lyons and by L. W. Sumter (in Cooper, et. al. 1979, 1–19 and 99–114), and in the study by West (2004) claim that Mill cannot be regarded as either an act-utilitarian or a rule-utilitarian but that his moral theory is more complex than either. Lyons's essays on various aspects of Mill's ethics are reprinted in *Rights, Welfare, and Mill's Moral Theory* (Lyons 1994).

In Chap. V of *Utilitarianism*, Mill has a theory of rights correlative to some but not all morally significant actions, and he restricts the morally obligatory to those actions for which punishment has utility; in *August Comte and Positivism*, (Mill 1963–1991, vol. X, p. 337–339), he clearly states a theory of morally meritorious action that goes beyond what is morally required. These would indicate that Mill's moral theory has a structure that is more complicated than any simple act- or rule-formulation.

Mill's contribution to the development of psychological theory is the subject of an important study by Fred Wilson (1990), who interprets Mill as a pioneer in turning psychology into an empirical science. A major study of Mill's economic theory is by Samuel Hollander (1985). Geoffrey Scarre has published a study of Mill's metaphysical views (1989).

Michael St. John Packe's *The Life of John Stuart Mill* (1954) continues to be the standard biography of the details of the Mill's life. But Nicholas Capaldi (2004) has recently written a thorough intellectual biography, arguing that Mill combined a Coleridgean/Germanic romanticism with his Benthamite Enlightenment heritage.

Mill continues to have his critics. H. J. McCloskey finds fault with nearly everything in Mill's philosophy (1971), and his liberalism has been attacked by Gertrude Himmelfarb (1974). McCloskey, however, does recognize that Mill's philosophy of language anticipated Ludwig Wittgenstein's notion of *family resemblances* in his rejection of Plato's essentialism.

See also Epistemology; Ethics; Logic, History of: Modern Logic: The Boolean Period: The Heritage of Kant and Mill; Metaphysics; Philosophy of Language; Plato; Social and Political Philosophy.

Bibliography

Alican, Necip Fikri. *Mill's Principle of Utility: A Defence of John Stuart Mill's Notorious Proof*. Amsterdam: Rodopi, 1994.

Anderson, Susan Leigh. *On Mill*. Belmont, CA: Wadsworth, 2000.

Berger, Fred R. *Happiness, Justice, and Freedom: The Moral and Political Philosophy of John Stuart Mill*. Berkeley: University of California Press, 1984.

Capaldi, Nicholas. *John Stuart Mill: A Biography*. Cambridge, U.K.: Cambridge University Press, 2004.

Cooper, Wesley E., Kai Nielsen, and Steven C. Patten, eds. *New Essays on John Stuart Mill and Utilitarianism. Canadian Journal of Philosophy* Supp. Vol. 5. Guelph, ON: Canadian Association for Publishing in Philosophy, 1979.

Crisp, Roger. *Routledge Philosophy Guidebook to J. S. Mill's "Utilitarianism"*. London: Routledge, 1997.

Donner, Wendy. *The Liberal Self: John Stuart Mill's Moral and Political Philosophy*. Ithaca, NY: Cornell University Press, 1991.

Duncan, Graeme Campbell. *Marx and Mill: Two Views of Social Conflict and Social Harmony*. Cambridge, U.K.: Cambridge University Press, 1973.

Gray, John. *Mill on Liberty: A Defence*. Rev. ed. London: Routledge and Kegan Paul, 1996.

Habibi, Don A. *John Stuart Mill and the Ethic of Human Growth*. Dordrecht: Kluwer, 2001.

Himmelfarb, Gertrude. *On Liberty and Liberalism: The Case of John Stuart Mill*. New York: Knopf, 1974.

Hollander, Samuel. *The Economics of J. S. Mill*. 2 vols. Toronto: University of Toronto Press, 1985.

Kitcher, Philip. "Mill, Mathematics, and the Naturalist Tradition." In *The Cambridge Companion to Mill*, edited by John Skorupski. Cambridge U.K.: Cambridge University Press, 1998.

Laine, Michael. *Bibliography of Works on John Stuart Mill*. Toronto: University of Toronto Press, 1982.

Lyons, David, *Rights, Welfare, and Mill's Moral Theory*. New York: Oxford University Press, 1994.

McCloskey, H. J. *John Stuart Mill: A Critical Study*. London: Macmillan, 1971.

Mill, John Stuart. *Collected Works of John Stuart Mill*. 33 vols., edited by J. M. Robson. Toronto: University of Toronto Press, 1963–1991.

The Mill Newsletter. Vols. 1–23. Toronto: University of Toronto Press, 1965–1989.

Morales, Maria. *Perfect Equality: John Stuart Mill on Well-Constituted Communities*. Latham, MD: Rowman & Littlefield, 1996.

Packe, Michael St. John. *The Life of John Stuart Mill*. London: Secker & Warburg, 1954.

Rees, J. C. *John Stuart Mill's "On Liberty"*. Oxford: Clarendon, 1985.

Riley, Jonathan. *Liberal Utilitarianism: A Social Choice Theory and J. S. Mill's Philosophy*. Cambridge, U.K.: Cambridge University Press, 1988.

Ryan, Alan. *J. S. Mill*. London: Routledge and Kegan Paul, 1974.

Ryan, Alan. *The Philosophy of John Stuart Mill*. 2nd ed. New York: Macmillan, 1988.

Scarre, Geoffrey. *Logic and Reality in the Philosophy of John Stuart Mill*. Dordrecht: Kluwer, 1989.

Schwartz, Pedro. *The New Political Economy of J. S. Mill*. London: Weidenfeld and Nicolson, 1973.

Semmel, Bernard. *John Stuart Mill and the Pursuit of Virtue*. New Haven: Yale University Press, 1984.

Skorupski, John, ed. *The Cambridge Companion to Mill*. Cambridge, U.K.: Cambridge University Press, 1998.

Skorupski, John. *John Stuart Mill*. London, New York: Routledge, 1989.

Ten, C. L. *Mill on Liberty*. Oxford: Oxford University Press, 1980.

Thomas, William. *Mill*. Oxford: Oxford University Press, 1985.

Thompson, Dennis F. *John Stuart Mill and Representative Government*. Princeton, NJ: Princeton University Press, 1976.

Urmson, J. O. "The Interpretation of the Moral Philosophy of J. S. Mill." *The Philosophical Quarterly* 3 (1953): 33–39.

Utilitas: A Journal of Utilitarian Studies 1– (1989–).

West, Henry R. *An Introduction to Mill's Utilitarian Ethics*. Cambridge, U.K.: Cambridge University Press, 2004.

Wilson, Fred. *Psychological Analysis and the Philosophy of John Stuart Mill*. Toronto: University of Toronto Press, 1990.

Henry R. West (1996, 2005)

MILLER, DICKINSON S.

(1868–1963)

Dickinson S. Miller was an American ethical philosopher and epistemologist who published both under his own name and under the pseudonym R. E. Hobart. He was born in Philadelphia and studied at the University of Pennsylvania, Clark University, the universities of Berlin and Halle, Hobart College, and Harvard University. He held a doctorate in philosophy from Halle and a D.Sc. from Hobart.

At Harvard, Miller was a student of William James, who became his longtime friend and with whom he often discussed and argued points of philosophy. James was

instrumental in getting Miller an appointment as associate professor of philosophy at Bryn Mawr College in 1893, the year after Miller's graduation from Harvard.

Miller left Bryn Mawr in 1898 to become first an instructor and then a professor of philosophy at Harvard. He subsequently joined the Columbia faculty, where he remained until the 1920s. He had also received a D.D. at Berkeley (California) Divinity School and in 1911 started to teach apologetics at the General Theological Seminary in New York City.

In his later days he lived for several years (1927–1932) close to his friend the critical realist Charles Augustus Strong, in Fiesole, near Florence, Italy. Strong appreciated Miller's company, especially because of Miller's neorealistic tendencies as opposed to Strong's different epistemological outlook. Their discussions were lively and interminable. George Santayana occasionally joined them, coming to Florence from Rome. Miller was a visitor during 1926 at the Vienna circle of logical positivists; although mostly a silent listener at the circle's sessions, he was an intensely interesting and challenging discussant in individual conversations. During his last twenty-five years he lived in Boston.

Miller's was an extremely penetrating and constructively critical mind. In a number of remarkable articles he addressed himself mainly to such topics as direct realism, the philosophy of mind, and also the controversy between William James and E. A. Singer on behaviorism. Especially interesting is "Is Consciousness 'A Type of Behavior'?" (1911), mainly about the "automatic sweetheart" puzzle. In 1951, Miller wrote "'Descartes' Myth' and Professor Ryle's Fallacy," a sharp critique of Gilbert Ryle's logical behaviorism. He also wrote on David Hume's views on causality and induction, on various topics in moral philosophy, and most notably, on the free-will–determinism issue. Miller's article provocatively titled "Free Will as Involving Determination and Inconceivable without It" (1934), published, for obscure reasons, under the name R. E. Hobart, has become a locus classicus of the free-will controversies. With remarkable lucidity and perspicacity Miller brought up to date the essentials of the point of view of Hume and J. S. Mill. He argued that once we realize the clear distinctions between causality and compulsion and between indeterminism and free will, the traditionally vexing problem disappears, and a fully adequate account of human freedom, responsibility, reward, and punishment can be given. Miller's views on religion and theology were extremely liberal and modern, close to the outlook of Unitarianism (in fact, he occasionally served as a Unitarian minister in the Boston area).

Miller's contributions to the epistemological controversies of his time may now seem a bit old-fashioned, but they are worthy of renewed attention because the same issues are still being debated, albeit in a different style and terminology.

See also Behaviorism; Epistemology; Ethics, History of; Hume, David; James, William; Logical Positivism; Mill, John Stuart; Ryle, Gilbert; Santayana, George.

Bibliography

WORKS BY MILLER

"Is Consciousness 'A Type of Behavior'?" *Journal of Philosophy* 8 (1911): 322–327.

"The Pleasure-Quality and the Pain-Quality Analysable, Not Ultimate." *Mind* 38 (1929): 215–218.

"Is There Not a Clear Solution of the Knowledge-Problem?" *Journal of Philosophy* 34 (1937): 701–712; 35 (1938): 561–572.

"An Event in Modern Philosophy." *Philosophical Review* 54 (1945): 592–606.

"Hume's Deathblow to Deductivism." *Journal of Philosophy* 46 (1949): 745–762.

"'Descartes' Myth' and Professor Ryle's Fallacy." *Journal of Philosophy* 48 (1951): 270–280.

Under the Name R. E. Hobart

"Hume without Scepticism." *Mind* 39 (1930).

"Free Will as Involving Determinism and Inconceivable without It." *Mind* 43 (1934): 1–27.

Herbert Feigl (1967)

MILLIKAN, RUTH GARRETT
(1933–)

Born December 19, 1933, and raised in Swarthmore, Pennsylvania, where her father taught physics, Millikan received her Ph.D. from Yale University in 1969. She began her career as a self-described "faculty housewife," raising four children before publishing her first book. Internationally recognized, Millikan has made significant contributions to philosophy of biology, animal cognition, philosophy of language, mind, and ontology. A unifying theme is the importance of the fact that humans are products of evolution. (Millikan's mother held a Ph.D. in paleontology—perhaps influencing Millikan's orientation to Darwinism.) A student of Wilfred Sellars, Millikan rejects epistemic *givens* and takes *meaning* talk to have the function of helping speakers bring their use into conformity with others; unlike other Sellarsians, Millikan

sees the sort of function that underwrites intentional content everywhere, not just in linguistic creatures. Her first book (Millikan 1984) is a detailed articulation of teleosemantics, a Darwinian account of both mental representations and language.

Millikan's work reaches far beyond her account of intentionality, as a small sample of her conclusions shows—among them: that dogs have perfectly good concepts, that some thoughts have two *directions of fit* at once, that understanding language is a form of direct perception. Difficult to summarize, Millikan's program can nonetheless be seen to be framed by three questions: In the philosophy of mind, What is it for one's *thoughts* to be of something?; in epistemology, What is it for one to *know what* one is thinking of?; and in metaphysics, What makes for the *objective samenesses* in the world that one's thoughts are of? Her interlocking answers form a picture of human cognition that challenges tradition on several scores, even as she seeks to defend tradition in the form of scientific realism and the correspondence theory of truth.

THOUGHT

What one's thoughts are of is, according to Millikan, determined by their historically selected function. All intentional items (bee dances, linguistic forms, perceptions, desires, fears, and so forth) have such *proper functions*, and what any particular intentional item is about, its content, is determined by such functions. (That individual words or token mental states have proper functions and that their content owes to proper functions are claims that have encountered vigorous opposition.) Specifically, a *proper function* of a feature *F* of an organism *O* is a task whose performance by earlier instances of *F* in other organisms of *O*'s kind in *O*'s lineage accounts for the proliferation of *F* in *O*'s kind here and now. Importantly, there are nonbiological cases of proper function—for example, customs, hammers, and nails—so the relevant notions of *task* and *lineage* must be understood broadly. The *content* of a representation type *R* is given by the connection between instances of *R* and worldly circumstances, recurrent exploitation of which by consumers of *R* has contributed to their proliferation over time.

What makes *mental representations*, such as thoughts, beliefs, and desires, distinct from other information-bearing items, such as bee dances? Mental representations are representations that "when they perform their proper function, their referents are identified" (Millikan 1984, p. 13). By *identified*, Millikan means that the referent is represented as being the same thing again. For example, Clarence's visual perception that a spider is crawling up his leg is an intentional state with a job to do, and such states exist in us because historically selected for performance of that job. The function of his thoughts is to coordinate information he already has about the spider with new information he is acquiring as well as with his subsequent action, trying to brush the self-same spider off his leg. For Clarence's thoughts to be of the spider, then, they must meet the additional requirement of functioning to create this sort coordination of information. The capacity to think of the same *as being the same*, or "coidentifying" (Millikan 2000), is an important accomplishment, distinctive of advanced cognition.

Millikan here joins company with P. F. Strawson and Gareth Evans in claiming that some form of reidentificatory capacity is necessary for thought about the objective world. Unlike Strawson or Evans, Millikan takes her insight about coidentification to have dramatic consequences for self-knowledge.

SELF-KNOWLEDGE

What sort of access do we have to our own thoughts? Millikan is a content externalist—just as the meanings of one's words are not settled by one's intentions, the content of one's thoughts are also determined by facts outside one's ken. To know what one is thinking of, then, is not an a priori matter. Some find this consequence troubling and seek to reconcile content externalism with first-person authority. But Millikan (1993) embraces this result, and argues that a still more radical conclusion follows from her functional account of cognition, namely, that *nothing* is epistemically "given" to thinkers. In particular, "meaning rationalism"—the doctrine that sameness and difference of meaning, univocity, and meaningfulness are all a priori accessible—is false. (It is a good question just who qualifies as a *meaning rationalist*—some argue, pace Millikan, that even Gottlob Frege not.) Millikan's rejection of meaning rationalism has several startling consequences: We can have no a priori access to logical possibility; there is nothing rationally wrong with believing contradictions; the validity of inferences is not an a priori property; and the very idea of a Fregean *mode of presentation* must be discarded. In short, like meaning, rationality ain't in the head.

Millikan's radical anti-individualism about meaning and rationality might be opposed by more moderate externalisms. And her attack on the very idea of modes of presentation meets with resistance from those who see a genuine explanatory role for modes, even within naturalistic accounts of the mind.

ONTOLOGY

Coidentification is the heart of thought because the goals of organisms are served by coidentifying. But the goals of organisms would only be thus served if there were genuine coidentifiables in the objective world. Millikan's ontology is decidedly realist. Her functional take on concepts has her carving nature at different joints than others might, however. For Millikan, empirical concepts are of *substances*, that is, *coidentifiables*. The category *substance* includes real kinds (e.g., *mouse*), individuals (e.g., *Mama*), event types (e.g., *breakfast again*), and numerous other stuffs and types (e.g., *ice, Starbucks Coffee House*). At an important level of abstraction, there is no genuine ontological distinction to be made among these things.

SUMMARY

In briefest summary: Millikan's program for understanding the nature of representation—which is to say, for understanding ourselves—is impressive for its combination of detail and scope.

See also Evans, Gareth; Frege, Gottlob; Philosophy of Biology; Philosophy of Language; Philosophy of Mind; Sellars, Wilfrid; Strawson, Peter Frederick.

Bibliography

WORKS BY MILLIKAN

Language, Thought and Other Biological Categories: New Foundations for Realism. Cambridge, MA: MIT Press, 1984.

White Queen Psychology: And Other Essays for Alice. Cambridge, MA: MIT Press, 1993.

"On Unclear and Indistinct Ideas." In *Philosophical Perspectives*, Vol. 8: *Logic and Language*. Edited by J. Tomberlin, pp. 75–110. Ascadero, CA: Ridgeview, 1994.

"Images of Identity." *Mind* 106 (423) (1997): 499–519.

On Clear and Confused Ideas: An Essay about Substance Concepts. Cambridge, U.K.: Cambridge University Press, 2000.

Varieties of Meaning. Cambridge, MA: MIT Press, 2002.

WORKS ABOUT MILLIKAN

Brandom, R. *Making It Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge, MA: Harvard University Press, 1994.

Burge, T., and C. Peacocke. "Our Entitlement to Self-Knowledge." *Proceedings of the Aristotelian Society* XCVI (1995): 91–158.

Evans, G. *The Varieties of Reference*. Oxford: Oxford University Press, 1982.

Fodor, J. "A Theory of Content II: The Theory." *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 1990, 65–77.

Lawlor, K. "Confused Thought and Modes of Presentation." *Philosophical Quarterly* 55 (218) (2005): 21–36.

Sainsbury, M. "Fregean Sense." In *Logica*. Edited by Timothy Childers, Petr Kolásr, and Vladimir Svoboda. pp. 261–276. Prague: Filosofia, 1997.

Strawson, P. F. *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen, 1959.

Strawson, P. F. *Subject and Predicate in Logic and Grammar*. London: Methuen, 1974.

Krista Lawlor (2005)

MILL'S METHODS OF INDUCTION

John Stuart Mill, in his *System of Logic* (Book III, Chapters 8–10), set forth and discussed five methods of experimental inquiry, calling them the method of agreement, the method of difference, the joint method of agreement and difference, the method of residues, and the method of concomitant variation. Mill maintained that these are the methods by which we both discover and demonstrate causal relationships, and that they are of fundamental importance in scientific investigation. Mill called these methods "eliminative methods of induction." In so doing, he was drawing an analogy with the elimination of terms in an algebraic equation—an analogy that is rather forced, except with respect to the various methods that are classed under the heading of method of difference. As will be demonstrated, it is perhaps best to use the term "eliminative methods" with reference to the elimination of rival candidates for the role of cause, which characterizes all these methods.

ILLUSTRATIONS OF THE METHODS

The general character of Mill's methods of experimental inquiry may be illustrated by examples of the two simplest ones, the methods of agreement and of difference. Mill's canon for the method of agreement is this: "If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree is the cause (or effect) of the given phenomenon."

For example, if a number of people who are suffering from a certain disease have all gone for a considerable time without fresh fruits or vegetables, but have in other respects had quite different diets, have lived in different conditions, belong to different races, and so on, so that the lack of fresh fruits and vegetables is the only feature common to all of them, then we can conclude that the lack of fresh fruits and vegetables is the cause of this particular disease.

Mill's canon for the method of difference is this: "If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon."

For example, if two exactly similar pieces of iron are heated in a charcoal-burning furnace and hammered into shape in exactly similar ways, except that the first is dipped into water after the final heating while the second is not, and the first is found to be harder than the second, then the dipping of iron into water while it is hot is the cause of such extra hardness—or at least an essential part of the cause, for the hammering, the charcoal fire, and so on may also be needed. For all this experiment shows, the dipping alone might not produce such extra hardness.

The method of agreement, then, picks out as the cause the one common feature in a number of otherwise different cases where the effect occurs; the method of difference picks out as the cause the one respect in which a case where the effect occurs differs from an otherwise exactly similar case where the effect does not occur. Both are intended to be methods of ampliative induction, that is, methods by which we can reason from a limited number of observed instances to a general causal relationship: The intended conclusion is that a certain disease is always produced by a lack of fresh fruits and vegetables, or that dipping iron into water while it is hot always hardens it, if it has been heated and hammered in a particular way. And the other three methods are intended to work in a similar manner.

These methods have been criticized on two main counts: First, it is alleged that they do not establish the conclusions intended, so that they are not methods of proof or conclusive demonstration; and second, that they are not useful as methods of discovery. Such criticisms have been used to support the general observation that these methods play no part, or only a very minor part, in the investigation of nature, and that scientific method requires a radically different description.

In order to estimate the force of such criticisms, and to determine the real value of the eliminative methods, Mill's formulation need not be discussed in detail. Instead, one need only determine what would be valid demonstrative methods corresponding to Mill's classes, and then consider whether such methods, or any approximations of them, have a place in either scientific or commonsense inquiry.

METHODS OF AGREEMENT AND OF DIFFERENCE

To avoid unnecessary complications, let us assume that the conclusion reached by any application of the method of agreement or of difference is to have the form "Such-and-such is a cause of such-and-such kind of event or phenomenon." For a formal study of these methods and the joint method we could regard a cause as a necessary and sufficient condition of the effect—or, in some cases, as a necessary condition only, or as a sufficient condition only—where to say that *X* is a necessary condition for *Y* is just to say that wherever *Y* is present, *X* is present, or briefly that all *Y* are *X*; and to say that *X* is a sufficient condition for *Y* is just to say that wherever *X* is present *Y* is present, or briefly that all *X* are *Y*.

In general we shall be looking for a condition that is both necessary and sufficient for the phenomenon, but there are variants of the methods in which we look for a condition that is merely necessary or merely sufficient. In practice, however, we are concerned with conditions that are not absolutely necessary or sufficient, but that are rather necessary and/or sufficient in relation to some *field*, that is, some set of background conditions, which may be specified more or less exactly. We are concerned, for example, not with the cause of a certain disease in general, but with what causes it in human beings living on the earth, breathing air, and so forth. Again, we are concerned not with the cause of hardness in general, but with that of a greater-than-normal hardness in iron in ordinary circumstances and at ordinary temperatures. The field in relation to which we look for a cause of a phenomenon must be such that the phenomenon sometimes occurs in that field and sometimes does not. We may assume that this field is constituted by the presence of certain qualities or at least of some general descriptive features, not by a specific location.

The *observation* that supports the conclusion is an observation of one or more instances in each of which various features are present or absent. An instance may be one in which the phenomenon in question occurs, which we may call a *positive instance*, or one in which the phenomenon does not occur, which we may call a *negative instance*.

To reason validly, however, from any such observation to a general causal conclusion, we require an additional general premise, an *assumption*. We must assume that there is some condition which, in relation to the field, is necessary and sufficient (or which is necessary, or which is sufficient) for the phenomenon, and also that this condition is to be found within a range of conditions

that is restricted in some way. For these methods fall within the general class of eliminative forms of reasoning, that is, arguments in which one possibility is confirmed or established by the elimination of some or all of its rivals. The assumption will state that there is a cause to be found and will limit the range of candidates for the role of cause; the task of the observation will be to rule out enough of the candidates initially admitted to allow some positive conclusion.

POSSIBLE CAUSES. It follows from the above that the assumption must indicate some limited (though not necessarily finite) set of what we may call *possible causes*. These are the factors (Mill calls them *circumstances* or *antecedents*) that, it is initially assumed, may be causally relevant to the phenomenon. Any possible cause, any factor that may be causally relevant in relation to the field in question, must, like the phenomenon itself, be something that sometimes occurs and sometimes does not occur within that field.

But are we to assume that a possible cause acts singly, if it acts at all? If the possible causes are *A, B, C*, etc., the phenomenon is *P*, and the field is *F*, are we to assume that the cause of *P* in *F* will be either *A* by itself or *B* by itself, and so on? Or are we to allow that it might be a conjunction, say *AC*, so that *P* occurs in *F* when and only when *A* and *C* are both present? Are we to allow that the necessary and sufficient condition might be a disjunction, say (*B* or *D*), so that *P* occurs in *F* whenever *B* occurs, and whenever *D* occurs, but only when one or other (or both) of these occurs? Again, are we to allow that what we have taken as possible causes may include counteracting causes, so that the actual cause of *P* in *F* may be, say, the absence of *C* (that is, the negation not-*C*, or $\overline{C}$) or perhaps $B\overline{C}$ so that *P* occurs in *F* when and only when *B* is present and *C* is absent at the same time?

There are in fact valid methods with assumptions of different sorts, from the most rigorous kind, which requires that the actual cause should be just one of the possible causes by itself, through those which progressively admit negations, conjunctions, and disjunctions of possible causes and combinations of these, to the least rigorous kind of assumption, which says merely that the actual cause is built up out of these possible causes in some way.

CLASSIFICATION OF THESE METHODS. There will be, then, not one method of agreement, one method of difference, and one joint method, but a series of variants of each. A complete survey could be made of all possible methods of these types, numbered as follows: A number from 1 to 8 before a decimal point will indicate the kind of assumption. Thus, it is assumed that there is an actual cause that is

(1) one of the possible causes;

(2) one of the possible causes or the negation of a possible cause;

(3) a possible cause or a conjunction of possible causes;

(4) a possible cause or a disjunction of possible causes;

(5) a possible cause or the negation of a possible cause, or a conjunction each of whose members is a possible cause or the negation of a possible cause;

(6) a possible cause, or the negation of a possible cause, or a disjunction each of whose members is a possible cause or the negation of a possible cause;

(7) a possible cause, or a conjunction of possible causes, or a disjunction each of whose members is a possible cause or a conjunction of possible causes;

(8) a possible cause, or the negation of a possible cause, or a conjunction each of whose members is a possible cause or the negation of one; or a disjunction each of whose members is a possible cause or the negation of one, or a conjunction each of whose members is a possible cause or a negation of one.

The first figure after the decimal point will indicate the sort of observation, as follows:

(1) a variant of the method of agreement;

(2) a variant of the method of difference;

(3) a variant of the joint method;

(4) a new but related method.

The second figure after the decimal point will mark further differences where necessary, but this figure will have no constant significance.

The complete survey cannot be given here, but a few selected variants will be considered, numbered in the manner set forth above.

POSITIVE METHOD OF AGREEMENT. Let us begin with an assumption of the first kind, that there is a necessary and sufficient condition *X* for *P* in *F*, that is, that for some *X* all *FP* are *X* and all *FX* are *P*, and *X* is identical

with one of the possible causes *A, B, C, D, E*. (It may be noted that a condition thus specified may sometimes not be what we would ordinarily regard as the cause of the phenomenon: We might rather say that it *contains* the real cause. However, in our present account we shall call such a condition the cause; it is explained below how the cause of a phenomenon may be progressively located with greater precision.)

We obtain a variant of the method of agreement (1.12) by combining with this assumption the following observation: A set of one or more positive instances such that one possible cause, say *A*, is present in each instance, but for every other possible cause there is an instance from which that cause is absent. This yields the conclusion that *A* is necessary and sufficient for *P* in *F*.

For example, the observation might be this:

	A	*B*	*C*	*D*	*E*
I_1	*p*	*a*	*p*	·	*a*
I_2	*p*	*p*	*a*	*a*	·

where *p* indicates that the possible cause is present, *a* that it is absent, and a dot that it may be either present or absent without affecting the result. I_1 and I_2 are positive instances: I_1 shows that neither *B* nor *E* is necessary for *P* in *F*, I_2 that neither *C* nor *D* is necessary, and hence, given the assumption, it follows that *A* is necessary and sufficient.

Since this reasoning eliminates candidates solely on the ground that they are not necessary, there is another variant (1.11) that assumes only that there is some necessary condition for *P* in *F* identical with one of the possible causes, and (with the same observation) concludes that *A* is a necessary condition for *P* in *F*.

Negative method of agreement. Besides the positive method of agreement, in which candidates are eliminated as not being necessary because they are absent from positive instances, there are corresponding variants of a negative method of agreement in which candidates are eliminated as not being sufficient because they are present in negative instances. This requires the following observation: A set of one or more negative instances such that one possible cause, say *A*, is absent from each instance, but for every other possible cause there is an instance in which it is present. For example:

	A	*B*	*C*	*D*	*E*
N_1	*a*	*p*	·	·	·
N_2	*a*	·	*p*	*p*	·
N_3	*a*	·	·	·	*p*

If the assumption was that one of the possible causes is sufficient for *P* in *F*, this observation would show (1.13) that *A* is sufficient, while if the assumption was that one of the possible causes is both necessary and sufficient, this observation would show (1.14) that *A* is necessary and sufficient.

METHOD OF DIFFERENCE. For the simplest variant of the method of difference (1.2) we need this observation: a positive instance I_1 and a negative instance N_1 such that of the possible causes present in I_1, one, say *A*, is absent from N_1, but the rest are present in N_1. For example:

	A	*B*	*C*	*D*	*E*
I_1	*p*	*p*	*p*	*a*	·
N_1	*a*	*p*	*p*	·	*p*

Here *D* is eliminated because it is absent from I_1, and hence not necessary, and *B, C*, and *E* are eliminated because they are present in N_1 and hence not sufficient. Hence, given the assumption that one of the possible causes is both necessary and sufficient for *P* in *F*, it follows that *A* is so. (Note that since it would not matter if, say, *E* were absent from I_1, the presence of the actual cause in I_1 need not be the only difference between the instances.) We may remark here that the method of difference, unlike some variants of the method of agreement, requires the assumption that there is some condition that is both necessary and sufficient for *P*. It is true, as we shall see later with variants 4.2 and 8.2, that the "cause" detected by this method is often not itself a necessary condition, or even a sufficient one; but the assumption needed is that *something* is both necessary and sufficient.

JOINT METHOD. The joint method may be interpreted as an indirect method of difference, that is, the job done by I_1 above may be shared among several positive instances, and the job done by N_1 among several negative instances. That is, we need (for 1.3) the following observation: a set S_i of one or more positive instances and a set S_n of one or more negative instances such that one of the possible causes, say *A*, is present throughout S_i and absent throughout S_n, but each of the other possible causes is either absent from at least one positive instance or present in at least one negative instance. Given that one of the possible causes is both necessary and sufficient, this yields the conclusion that *A* is so.

SIMPLE VARIANTS OF THESE METHODS. With an assumption of the second kind (that the requisite condi-

tion is either a possible cause or a negation of a possible cause) we need stronger observations. Thus, for variants of the positive method of agreement (2.11 and 2.12) we need this: two or more positive instances such that one possible cause (or negation), say *A*, is present in each instance, but for every other possible cause there is an instance in which it is present and an instance from which it is absent. This is needed to rule out, as candidates for the role of necessary (or both necessary and sufficient) condition, the negations of possible causes as well as the possible causes other than *A* themselves.

For the corresponding variant of the method of difference (2.2) we need this: a positive instance I_1 and a negative instance N_1 such that one possible cause (or negation), say *A*, is present in I_1 and absent from N_1, but each of the other possible causes is either present in both I_1 and N_1 or absent from both. For example:

	A	*B*	*C*	*D*	*E*
I_1	*p*	*p*	*a*	*a*	*p*
N_1	*a*	*p*	*a*	*a*	*p*

Since *B* is present in N_1, *B* is not sufficient for *P* in *F*; but since *B* is present in I_1, not-*B* is not necessary for *P* in *F*; thus neither *B* nor not-*B* can be both necessary and sufficient. Similarly, *C, D, E*, and their negations, and also not-*A*, are ruled out, and thus the necessary and sufficient condition must be *A* itself. This is the classic difference observation described by Mill, in which the only (possibly relevant) difference between the instances is the presence in I_1 of the factor identified as the actual cause; but we need this observation (as opposed to the weaker one of 1.2) only when we allow that the negation of a possible cause may be the actual cause.

The joint method needs, along with this weaker assumption, a similarly strengthened observation: That is, each of the possible causes other than *A* must be either present in both a positive and a negative instance or absent from both a positive and a negative instance, and then this variant (2.3) still yields the conclusion that *A* is both necessary and sufficient.

(What Mill and his followers describe as the joint method may be not this indirect method of difference, but rather a double method of agreement, in which a set of positive instances identifies a necessary condition and a set of negative instances identifies a sufficient condition. Such a combination is redundant with an assumption of either of the first two kinds, but not when the assumption is further relaxed.)

MORE COMPLEX VARIANTS. We consider next an assumption of the third kind, that the requisite condition is either a possible cause or a conjunction of possible causes. (This latter possibility seems to be at least part of what Mill meant by "an intermixture of effects.") This possibility does not affect the positive method of agreement, since if a conjunction is necessary, each of its conjuncts is necessary, and candidates can therefore be eliminated as before. But since the conjuncts in a necessary and sufficient condition may not severally be sufficient, the negative method of agreement as set forth above will not work. The observation of (1.13 or) 1.14 would now leave it open that, say, *BC* was the required (sufficient or) necessary and sufficient condition, for if *C* were absent from N_1 and *B* from N_2, then *BC* as a whole might still be sufficient: It would not be eliminated by either of these instances. This method now (in 3.14) needs a stronger observation, namely, a *single* negative instance N_1 in which one possible cause, say *A*, is absent, but *every* other possible cause is present. This will show that no possible cause or conjunction of possible causes that does not contain *A* is sufficient for *P* in *F*. But even this does not show that the requisite condition is *A* itself, but merely that it is either *A* itself or a conjunction in which *A* is a conjunct. We may express this by saying that the cause is ($A\underline{\ldots}$), where the dots indicate that other conjuncts may form part of the condition, and the dots are underlined, while *A* is not, to indicate that *A must* appear in the formula for the actual cause, but that other conjuncts may or may not appear.

The corresponding variant (3.2) of the method of difference needs only the observation of 1.2; but it, too, establishes only the less complete conclusion that ($A\underline{\ldots}$) is a necessary and sufficient condition of *P* in *F*. For while (in the example given for 1.2 above) *B, C, D*, and *E* singly are still eliminated as they were in 1.2, and any conjunctions such as *BC* which, being present in I_1, *might* be necessary, are eliminated because they are also present in N_1 and hence not sufficient, a conjunction such as *AB*, which contains *A*, is both present in I_1, and absent from N_1, and might therefore be both necessary and sufficient. Thus this assumption and this observation show only that *A* is, as Mill put it, "the cause, or an indispensable part of the cause." The full cause is represented by the formula ($A\underline{\ldots}$), provided that only possible causes that are present in I_1 can replace the dots.

In the corresponding variant of the joint method (3.3), we need a *single* negative instance instead of the set S_n, for the same reason as in 3.14, and the cause is specified only as ($A\underline{\ldots}$).

With an assumption of the fourth kind (that the requisite condition is either a possible cause or a disjunction of possible causes), the negative method of agreement (4.13 and 4.14) works as in 1.13 and 1.14, but the positive method of agreement is now seriously affected. For with the observation given for 1.12 above, the necessary and sufficient condition might be, say, (*B* or *C*), for this disjunction is present in both I_1 and I_2, though neither of its disjuncts is present in both. Thus the observation of 1.12 would leave the result quite undecided. We need (for 4.12) a much stronger observation, that is, a single positive instance in which *A* is present but all the other possible causes are absent together; but even this now shows only that the cause is (*A* or…). This assumption (that the cause may be a disjunction of possible causes) allows what Mill called a "plurality of causes," for each of the disjuncts is by itself a "cause" in the sense that it is a sufficient condition; and what we have just noted is the way in which this possibility undermines the use of the method of agreement.

The method of difference, on the other hand (4.2), still needs only the observation of 1.2; this eliminates all possible causes other than *A*, and all disjunctions that do not contain *A*, either as being not sufficient because they are present in N_1 or as not necessary because they are absent from I_1. The only disjunctions not eliminated are those that occur in I_1 but not in N_1, and these must contain *A*. Thus this observation, with this assumption, shows that a necessary and sufficient condition is (*A* or…), that is, either *A* itself or a disjunction containing *A*, where the other disjuncts are possible causes absent from N_1. This, of course, means that *A* itself, the factor thus picked out, may be only a sufficient condition for *P*.

The joint method with this assumption (4.3) needs a *single* positive instance, but can still use a set of negative instances and it specifies the cause as (*A* or…).

As the assumptions are relaxed further, the method of agreement requires stronger and stronger observations. For example, in 6.12, which is a variant of the positive method with an assumption allowing that the necessary and sufficient condition may be a disjunction of possible causes or negations, the observation needed is a set S_i of positive instances such that one possible cause, say *A*, is present in each, but that for every possible combination of the other possible causes and their negations there is an instance in which this combination is present (that is, if there are *n* other possible causes, we need 2^n different instances). This observation will climinate every disjunction that does not contain *A*, and will show that the requisite necessary and sufficient condition is (*A* or…), and hence that *A* itself is a sufficient condition for *P* in *F*. *A* corresponding variant of the negative method of agreement (5.14) shows that (*A*…) is a necessary and sufficient condition, and hence that *A* itself is necessary—a curious reversal of roles, because in the simplest variants, the positive method of agreement was used to detect a necessary condition and the negative one a sufficient condition.

In the method of difference, however, the observation of 1.2 (or, where negations are admitted, that of 2.2) continues to yield results, though the conclusions become less complete, that is, the cause is less and less completely specified. For example, in 8.2, where we assume that there is a necessary and sufficient condition for *P* in *F* which may be one of the possible causes, or a negation of one, or a conjunction of possible causes or negations, or a disjunction of possible causes or negations or of conjunctions of possible causes or negations—which in effect allows the actual condition to be built up out of the possible causes in any way—the observation of 2.2 establishes the conclusion that the requisite condition is (*A*… or…). that is to say, it is either *A* itself, or a conjunction containing *A*, or a disjunction in which one of the disjuncts is *A* itself or a conjunction containing *A*. Since any such disjunct in a necessary and sufficient condition is a sufficient condition, this observation, in which the presence of *A* in I_1 is the only possibly relevant difference between I_1 and N_1, shows even with the least rigorous kind of assumption that *A* is at least a necessary part of a sufficient condition for *P* in *F*—the sufficient condition being (*A*…).

The joint method, as an indirect method of difference, ceases to work once we allow both conjunctions and disjunctions; but a double method of agreement comes into its own with this eighth kind of assumption. In 8.12, as in 6.12, if there are *n* possible causes other than *A*, the set of 2^n positive instances with *A* present in each but with the other possible causes present and absent in all possible combinations will show that (*A* or…) is necessary and sufficient, and hence that *A* is sufficient. Similarly in 8.14, as in 5.14, the corresponding set of $2^{>n}$ negative instances will show that (*A*…) is necessary and sufficient and hence that *A* is necessary. Putting the two observations together, we could conclude that *A* is both necessary and sufficient.

A new method, similar in principle, can be stated as follows (8.4): If there are *n* possible causes in all, and we observe 2^n instances (positive or negative) which cover all possible combinations of possible causes and their negations, then the disjunction of all the conjunctions found in the positive instances is both necessary and sufficient

for *P* in *F*. For example, if there are only three possible causes, *A, B, C*,

A	*B*	*C*	*P*
p	*p*	*p*	*a*
p	*p*	*a*	*p*
p	*a*	*p*	*p*
p	*a*	*a*	*a*
a	*p*	*p*	*a*
a	*p*	*a*	*p*
a	*a*	*p*	*a*
a	*a*	*a*	*a*

and we have the observations listed in the accompanying table, then ($AB\overline{C}$ or $A\overline{B}C$ or $\overline{A}B\overline{C}$) is a necessary and sufficient condition for *P* in *F*. For if these are the only possibly relevant conditions, each combination of possible causes and negations for which *P* is present is sufficient for *P*, and these are the only sufficient conditions for *P*, since in all the relevantly different circumstances *P* is absent; but the disjunction of all the sufficient conditions must be both necessary and sufficient, on the assumption that there is some condition that is both necessary and sufficient.

MANY VALID METHODS. We thus find that while we must recognize very different variants of these methods according to the different kinds of assumptions that are used, and while the reasoning that validates the simplest variants fails when it is allowed that various negations and combinations of factors may constitute the actual cause, nevertheless there are valid demonstrative methods which use even the least rigorous form of assumption, that is, which assume only that there is some necessary and sufficient condition for *P* in *F*, made up in some way from a certain restricted set of possible causes. But with an assumption of this kind we must be content either to extract (by 8.2) a very incomplete conclusion from the classical difference observation or (by 8.12, 8.14, the combination of these two, or 8.4) to get more complete conclusions only from a large number of instances in which the possible causes are present or absent in systematically varied ways.

AN EXTENSION OF THE METHODS. An important extension of all these methods is the following: Since in every case the argument proceeds by eliminating certain candidates, it makes no difference if what is *not* eliminated is not a single possible cause but a cluster of possible causes which in our instances are always present together or absent together, the conclusion being just as we now have it, but with a symbol for the cluster replacing *A*. For example, if in 2.2 we have, say, both *A* and *B* present in I_1 and both absent from N_1, but each possible cause either present in both or absent from both, it follows that the cluster (*A,B*) is the cause in the sense that the actual cause lies somewhere within this cluster. A similar observation in 8.2 would show that either *A*, or *B*, or *AB*, or (*A* or *B*) is an indispensable part of a sufficient condition for *P* in *F*.

METHOD OF RESIDUES

The method of residues can be interpreted as a variant of the method of difference in which the negative instance is not observed but constructed on the basis of already known causal laws.

Suppose, for example, that a positive instance I_1 has been observed as follows:

	A	*B*	*C*	*D*	*E*
I_1	*p*	*p*	*a*	*p*	*a*

Now if we had, to combine with this, a negative instance N_1 in which *B* and *D* were present and *A, C*, and *E* absent, we could infer, according to the kind of assumption made, by 2.2 that *A* was the cause, or by 8.2 that ($A\underline{\ldots}$ $\underline{\text{or}\ldots}$) was the cause, and so on. But if previous inductive inquiries have already established laws from which it follows that given $\overline{A}B\overline{C}D\overline{E}$ in the field *F, P* would not result, there is no need to observe N_1; we already know all that N_1 could tell us, and so one of the above-mentioned conclusions follows from I_1 alone along with the appropriate assumption.

Again, if the effect or phenomenon in which we are interested can be quantitatively measured, we could reason as follows. Suppose that we observe a positive instance, say with the factors as in I_1 above, in which there is a quantity x_1 of the effect in question, while our previously established laws enable us to calculate that with the factors as in N_1 there would be a quantity x_2 of this effect; then we can regard the difference (x_1–x_2) as the phenomenon *P* which is present in I_1 but absent from N_1. With an assumption of kind (1) or (2) or (4) or (6)—that is, any assumption that does not allow conjunctive terms in the cause—we could conclude that the cause of *P* in this instance I_1 was *A* alone, and hence that *A* is a sufficient condition for *P* in *F*. With an assumption of kind (1) or (2) we could indeed infer that *A* is both necessary and sufficient, but with one of kind (4) or (6) we could con-

clude only that a necessary and sufficient condition is (*A* <u>or…</u>).

To make an assumption of any of these four kinds is to assume that the effects of whatever factors are actually relevant are merely additive, and this lets us conclude that the extra factor in I_1, namely *A*, by itself produces in relation to *F* the extra effect (x_1–x_2). But with an assumption of kind (3) or (5) or (7) or (8), which allows conjunctive terms, and hence what Mill calls an "intermixture of effects," we could only infer that the cause of (x_1–x_2) in this instance was (*A*<u>…</u>). With the other factors that were present in both I_1 and N_1, *A* was sufficient to produce this differential effect, but it does not follow that *A* is sufficient for this in relation to *F* as a whole. (Though Mill does not mention this, such a use of constructed instances along with some observed ones is in principle applicable to all the methods, not only to the method of difference in the way here outlined.)

METHOD OF CONCOMITANT VARIATION

The method of concomitant variation, like those already surveyed, is intended to be a form of ampliative induction; we want to argue from a covariation observed in some cases to a general rule of covariation covering unobserved cases also. To interpret this method we need a wider concept of cause than that which we have so far been using. *A* cause of *P* in the field *F* must now be taken, not as a necessary and sufficient condition, but as something on whose magnitude the magnitude of *P*, in *F*, functionally depends. For our present purpose this means only that there is some true lawlike proposition which, within *F*, relates the magnitude of the one item to that of the other. The *full cause*, in this sense, will be something on which, in *F*, the magnitude of *P* wholly depends, that is, the magnitude of *P* is uniquely determined by the magnitudes of the factors that constitute the full cause.

A full investigation of such a functional dependence would comprise two tasks: first, the identification of all the factors on which, in *F*, the magnitude of *P* depends, and second, the discovery of the way in which this magnitude depends on these factors. The completion of the first task would yield a mere list of terms, that of the second a mathematical formula. Only the first of these tasks can be performed by an eliminative method analogous to those already surveyed.

We should expect to find concomitant variation analogues of both the method of agreement and the method of difference, that is, ways of arguing to a causal relationship between *P* and, say, *A*, both from the observation of cases where *P* remains constant while *A* remains constant but all the other possibly relevant factors vary, and also from the observation of cases where *P* varies while *A* varies but all the other possibly relevant factors remain constant. And indeed there are methods of both kinds, but those of the second kind, the analogues of the method of difference, are more important.

As before, we need an assumption as well as an observation, but we have a choice between two different kinds of assumption. An assumption of the more rigorous kind would be that in *F* the magnitude of *P* wholly depends in some way on the magnitude of *X*, where *X* is identical with just one of the possible causes *A, B, C, D, E*. Given this, if we observe that over some period, or over some range of instances, *P* varies in magnitude while one of the possible causes, say *A*, also varies but all the other possible causes remain constant, we can argue that none of the possible causes other than *A* can be that on which the magnitude of *P* wholly depends, and thus conclude that *X* must be identical with *A*, that in *F* the magnitude of *P* depends wholly on that of *A*. (But *how* it depends, that is, what the functional law is, must be discovered by an investigation of some other sort.)

An assumption of the less rigorous kind would be that in *F* the magnitude of *P* wholly depends in some way on the magnitudes of one or more factors *X, X′, X″*, etc., where each of the actually relevant factors is identical with one of the possible causes *A, B, C, D, E*. Given this, if we again observe that *P* varies while, say, *A* varies but *B, C, D, E* remain constant, this does not now show that *B*, for example, cannot be identical with *X*, etc.; that is, it does not show that variations in *B* are causally irrelevant to *P*. All it shows is that the magnitude of *P* is not *wholly* dependent upon any set of factors that does not include *A*, for every such set has remained constant while *P* has varied. This leaves it open that the full cause of *P* in *F* might be *A* itself, or might be some set of factors, such as (*A,B,D*) which includes *A* and some of the others as well. All we know is that the list must include *A*. This observation and this assumption, then, show that a full cause of *P* in *F* is (A, <u>…</u>); that is, that *A* is an actually relevant factor and there may or may not be others. Repeated applications of this method could fill in other factors, but would not *close* the list. (And, as before, it is a further task, to be carried out by a different sort of investigation, to find *how* the magnitude of *P* depends on those of the factors thus shown to be actually relevant.)

To close the list, that is, to show that certain factors are actually irrelevant, we need to use an analogue of the method of agreement. If we assume, as before, that the

full cause of P in F is some set of factors (X, X', X'', etc.), but also that P is *responsive* to all these factors in the sense that for any variation, in, say, X while X', X'', etc. remain constant P will vary, and that X, X', X'', etc. are identical with some of the possible causes A, B, C, D, E, then if we observe that P remains constant while, say, A, C, D, and E remain constant but B varies, we can conclude that B is causally irrelevant, that none of the X's is identical with B.

USES AND APPLICATIONS OF THE ELIMINATIVE METHODS

We have so far been considering only whether there are demonstratively valid methods of this sort; but by stating more precisely what such methods involve, we may incidentally have removed some of the more obvious objections to the view that such methods can be applied in practice. Thus, by introducing the idea of a *field*, we have given these methods the more modest task of finding the cause of a phenomenon in relation to a field, not the ambitious one of finding conditions that are absolutely necessary and sufficient. By explicitly introducing the possible causes as well as the field, we have freed the user of the method of agreement from having to make the implausible claim that the user's instances have only one circumstance in common. Instead, the user has merely to claim that they have in common only one of the possible causes, while admitting that all the features that belong to the field, or that are constant throughout the field, will belong to all the instances, and that there may be other common features too, though not among those that he has initially judged to be possibly relevant.

Similarly, the user of the method of difference has only to claim that no *possibly relevant* feature other than the one he has picked as the cause is present in I_1 but not in N_1. Also, we have taken explicit account of the ways in which the possibilities of counteracting causes, a plurality of causes, an intermixture of effects, and so on, affect the working of the methods, and we have shown that even when these possibilities are admitted we can still validly draw conclusions, provided that we note explicitly the incompleteness of the conclusions that we are now able to draw (for example, by the method of difference) or the much greater complexity of the observations we need (for example, in variants of the method of agreement or method 8.4).

ELIMINATIVE METHODS AND INDUCTION. By making explicit the assumptions needed and by presenting the eliminative methods as deductively valid forms of argument, we have abandoned any pretense that methods such as these in themselves solve or remove the "problem of induction." Provided that the requisite observations can be made, the ultimate justification of any application of one of these methods of ampliative induction will depend on the justification of the assumption used; and, since this proposition is general in form, it will presumably have to be supported by some other kind of inductive, or at least nondeductive, reasoning. But we must here leave aside this question of ultimate justification.

ELIMINATIVE METHODS AND DETERMINISM. Some light, however, can be thrown on the suggestion frequently made that causal determinism is a presupposition of science. If these eliminative methods play some important part in scientific investigation, then it is noteworthy that they all require deterministic assumptions: They all work toward the identification of a cause of a given phenomenon by first assuming that there is some cause to be found for it. However, it has emerged that what we require is not a single universally applicable principle of causality, namely, that every event has a cause, but something at once weaker in some ways and stronger in other ways than such a principle. The principle assumed is that the particular phenomenon P in the chosen field F has a cause, but that a cause of P in F is to be found within a range of factors that is restricted in some way. We have also found that different concepts of a cause are required for concomitant variation and for the other methods. The complaint that the phrase "uniformity of nature" cannot be given a precise or useful meaning, incidentally, has been rebutted by finding in exactly what sense our methods have to assume that nature is uniform.

EMPLOYMENT OF THE METHODS. Such assumptions are in fact regularly made, both in investigations within our already developed body of knowledge and in our primitive or commonsense ways of finding out about the world. In both these sorts of inquiry we act on the supposition that any changes that occur are caused; they do not "just happen." In a developed science, the causal knowledge that we already have can limit narrowly the range of possibly relevant causal factors. It can tell us, for this particular phenomenon, what kinds of cause to be on the lookout for, and how to exclude or hold constant some possibly relevant factors while we study the effects of others.

In more elementary discoveries, we restrict the range of possibly relevant factors mainly by the expectation that the cause of any effect will be somewhere in the near spatiotemporal neighborhood of the effect. The possible

causes, then, will be features that occur variably within the field in question in the neighborhood of cases where the effect either occurs, or might have occurred, but does not.

USE OF METHOD OF DIFFERENCES. As an example of the above, singular causal sequences are detected primarily by the use of variants of the method of difference. Antoine-Henri Becquerel discovered that the radium he carried in a bottle in his pocket was the cause of a burn by noticing that the presence of the radium was the only possible relevant difference between the time when the inflammation developed and the earlier time when it did not, or between the part of his body where the inflammation appeared and the other parts.

Similar considerations tell us that a certain liquid turned this litmus paper red: The paper became red just after it was dipped in the liquid, and nothing else likely to be relevant happened just then. The situations before and after a change constitute our negative and positive instances respectively, and we may well be fairly confident that this is the only possibly relevant factor that has changed. We do not and need not draw up a list of possible causes, but by merely being on the lookout for other changes we can ensure that what would constitute a large number of possible causes (identified as such by their being in the spatiotemporal neighborhood) are the same in I_1 as in N_1.

Repeating the sequence—for example, dipping another similar piece of litmus paper into the liquid—confirms the view that the liquid caused the change of color. But it is not that in this case we are using the method of agreement; the repetition merely makes it less likely that any other change occurred to cause the change of color simultaneously with each of the two dippings, and this confirms our belief that the instances are what the use of the method of difference would require.

Since, in general, it will not be plausible to make an assumption more rigorous then one of kind (8), the conclusion thus established will only be that this individual sequence is an exemplification of a *gappy* causal law, of the form that (*A* <u>... or...</u>) is necessary and sufficient for *P* in *F*. But this is exactly what our ordinary singular causal statements mean: To say that this caused that says only that this was needed, perhaps in conjunction with other factors that were present, to produce the effect, and it leaves it open that other antecedents altogether (not present in this case) might produce the same effect.

General causal statements, such as "The eating of sweets causes dental decay," are to be interpreted similarly as asserting gappy causal laws. Anyone who says this would admit that the eating of sweets has this effect only in the presence of certain other conditions or in the absence of certain counteracting causes, and he would admit that things other than the eating of sweets might produce tooth decay. And such a gappy causal law can be established by the use of method 8.2, or the method of concomitant variation, or by statistical methods that can be understood as elaborations of these. Such general causal statements are, however, to be understood as asserting gappy causal laws, not mere statistical correlations: Anyone who uses such a statement is claiming that in principle the gaps could be filled in.

USE IN DISCOVERING EFFECTS. The use of the above methods is not confined to cases where we begin with a question of the form "What is the cause of so-and-so?" We may just as well begin by asking "What is the effect of so-and-so?"—for example, "What is the effect of applying a high voltage to electrodes in a vacuum tube?" But we are justified in claiming that what is observed to happen is an effect of this only if the requirements for the appropriate variant of the method of difference are fulfilled.

USE OF METHOD OF AGREEMENT. The simpler variants of the method of agreement can be used to establish a causal conclusion only in a case in which our previous knowledge narrowly restricts the possible causes and justifies the belief that they will operate singly. For example, if the character of a disease is such as to indicate that it is of bacterial origin, then the microorganism responsible may be identified through the discovery that only one species of microorganism not already known to be innocent is present in a number of cases of the disease. Otherwise, the observation of what seems to be the only common factor in a number of cases of a phenomenon can be used only very tentatively, to suggest a hypothesis that will need to be tested in some other way.

Where, however, we have a very large number of extremely diverse instances of some effect, and only one factor seems to be present in all of them, we may reason by what is in effect an approximation to method 8.12. The diverse instances cover at least a large selection of all the possible combinations of possibly relevant factors and their negations. Therefore it is probable that no condition not covered by the formula (*A* <u>or...</u>) is necessary, and hence, if there is a necessary and sufficient condition, (*A* <u>or ...</u>) is such, and hence *A* itself is a sufficient condition of the phenomenon.

Similarly, by an approximation to 8.14, we may reason that the one possibly relevant factor that is found to be absent in a large number of very diverse negative instances is probably a necessary condition of the phenomenon (that is, that its negation is a counteracting cause).

USE OF METHOD OF CONCOMITANT VARIATION. The method of concomitant variation, with statistical procedures that can be considered as elaborations of it, is used in a great many experimental investigations in which one possibly relevant factor is varied (everything else that might be relevant being held constant) to see whether there *is* a causal connection between that one factor and the effect in question. (Of course, what we regard as a single experiment may involve the variation of several factors, but still in such a way that the results will show the effects of varying each factor by itself: Such an experiment is merely a combination of several applications of concomitant variation.)

FURTHER USES. The "controlled experiment," in which a control case or control group is compared with an experimental case or experimental group, is again an application of the method of difference (or perhaps the method of residues, if we use the control case, along with already known laws, to tell us what would have happened in the experimental case if the supposed cause had not been introduced.)

An important use of these methods is in the progressive location of a cause. If we take "the drinking of wine" as a single possible cause, then an application of 8.2 may show that the drinking of wine causes intoxication: That is, this factor is *a* necessary element in *a* sufficient condition for this result. But we may then analyze this possible cause further and discover that several factors are included in this one item that we have named "the drinking of wine," and further experiments may show that only one of these factors was really necessary: The necessary element will then be more precisely specified. But the fact that this is always possible leaves it true that in relation to the earlier degree of analysis of factors, the drinking of wine was a necessary element in a sufficient condition, and the discovery of this (admittedly crude) causal law is correct as far as it goes and is an essential step on the way to the more accurate law that is based on a finer analysis of factors.

CRITICISM OF THE METHODS

The sort of example presented above helps to rebut one stock criticism of these methods, which is that they take for granted what is really the most important part of the procedure, namely, the discovery and analysis of factors. Any given application of one of these methods does presuppose some identification of possible causes, but it will not be completely vitiated by the fact that a finer analysis of factors is possible. Besides, the use of the methods themselves (particularly to discover singular causal sequences and hence the dispositional properties of particular things) is part of the procedure by which factors are further distinguished and classified. Also, the assumptions used, especially with regard to the range of possible causes allowed, are corrigible, and in conjunction with the methods they are self-correcting. A mistaken assumption is likely to lead, along with the observations, to contradictory conclusions, and when this happens we are forced to modify the assumption, in particular, to look further afield than we did at first for possibly relevant factors.

A fundamental and widely accepted objection to the claim that these methods form an important part of scientific method is that science is not concerned, or not much concerned, with causal relations in the sense in which these methods can discover them. It may be conceded that the formulation and confirmation of hypotheses and theories of the kind that constitute the greater part of a science such as physics is a scientific procedure quite different from the actual use of these methods. Even the discovery of a law of functional dependence is, as was noted, a task beyond what is achieved by our method of concomitant variation. It may also be conceded that many sciences are concerned largely with the simple discovery of new items and the tracing of processes rather than with causal relationships. Further, it was noted that these methods logically *cannot* be the whole of scientific procedure, since they require assumptions which they themselves cannot support.

In reply to this objection, however, it can be stressed, first, that a great deal of commonsense everyday knowledge, and also a great deal of knowledge in the more empirical sciences, is of causal relations of this sort, partly of singular causal sequences and partly of laws, especially of the incomplete or gappy form at which these methods characteristically arrive.

Second, it is largely such empirical causal relations that are explained by, and that support, the deeper theories and hypotheses of a developed science. But if they are to be used thus, they must be established independently.

Third, although descriptions of the eliminative methods of induction have often been associated with a kind of ground-floor empiricism that takes knowledge to

be wholly concerned with empirical relations between directly observable things, qualities, and processes, the methods themselves are not tied to this doctrine but can establish causal relations between entities that are indirectly observed. For example, as long as there is any way, direct or indirect, of determining when a magnetic field is present and when there is an electric current in a wire, the methods can establish the fact that such a current will produce a magnetic field.

Finally, even where such causal relations are not the main object of inquiry, in investigation we constantly make use of causal relations, especially of singular causal sequences. In measuring, say, a voltage, we are assuming that it was the connecting of the meter across those terminals that caused this deflection of its needle, and the precautions that ensure that this is really so are to be explained in terms of our methods.

In fact, these methods are constantly used, explicitly or implicitly, both to suggest causal hypotheses and to confirm them. One should not, of course, expect any methods of empirical inquiry to establish conclusions beyond all possibility of doubt or all need of refinement, but in using these methods we can frequently say at least this: We have reason to suppose that for an event of this kind in this field there is some cause, and if the cause is not such-and-such, we cannot see what else the cause might be.

See also Deduction; Determinism, A Historical Survey; Empiricism; Induction; Mill, John Stuart.

Bibliography

WORKS ON INDUCTION

The classical study of eliminative induction remains that of J. S. Mill, *A System of Logic* (London, 1843), Book III, Chs. 8–10. Mill acknowledges that his study owes much to John Herschell, *A Preliminary Discourse on the Study of Natural Philosophy* (London: Longman, Rees, Orme, Brown, and Green, 1831), Part II, Ch. 6, and both are fundamentally indebted to Francis Bacon, *Novum Organum* (London: Joannem Billium, 1620), Book II. Since Mill, the literature has become extensive, but mostly in textbooks rather than in original works on logic or philosophy. There have been many worthwhile treatments of eliminative induction that are far above the textbook level—notably those of John Venn, *Empirical Logic* (London, 1889), Ch. 17; Christoff von Sigwart, *Logic*, 2nd ed. (Freiburg, 1893), translated by Helen Dendy as *Logic* (London: Sonnenschein, 1895), Vol. II, Part II, Ch. 5; and H. W. B. Joseph, *An Introduction to Logic* (Oxford: Clarendon Press, 1906), Ch. 20. But there are only a small number of writers who, either by criticizing Mill or developing his account, have added something new and substantial to either the logic or the philosophy of eliminative induction.

CRITICISMS OF MILL'S METHODS

Mill's most important critics are William Whewell, *The Philosophy of Discovery* (London, 1860), Ch. 22; W. S. Jevons, *The Principles of Science* (London: Macmillan, 1874), Chs. 11, 19, and 23; F. H. Bradley, *The Principles of Logic* (London: K. Paul, Trench, 1883), Book II, Part II, Ch. 3; and M. R. Cohen and Ernest Nagel, *An Introduction to Logic and Scientific Method* (New York: Harcourt Brace, 1934), Ch. 13.

ELABORATIONS ON MILL'S METHODS

The main writers who have tried to develop Mill's ideas on the logical side are W. E. Johnson, *Logic* (Cambridge, U.K., 1924), Part II, Ch. 10; C. D. Broad, "The Principles of Demonstrative Induction" in *Mind* 39 (1930): 302–317 and 426–439; and G. H. von Wright, *A Treatise on Induction and Probability* (London: Routledge and Paul, 1951). Broad, following Johnson, undertakes a demonstrative reconstruction of Mill's methods and tries to extend eliminative methods to reasonings that terminate in quantitative laws. Von Wright's is the most thorough treatment so far published and studies the conditions under which "complete elimination" can be achieved even with what are here called the "less rigorous" kinds of assumptions. His account, however, seems somewhat unclear.

FURTHER STUDIES

The only major addition to the pure philosophy of induction is that of J. M. Keynes, *A Treatise on Probability* (London: Macmillan, 1921), Part III. Three more recent books that contain some discussion of it are J. O. Wisdom, *Foundations of Inference in Natural Science* (London: Methuen, 1952), Ch. 11; S. F. Barker, *Induction and Hypothesis* (Ithaca, NY: Cornell University Press, 1957), Ch. 3; and J. P. Day, *Inductive Probability* (London: Routledge and Paul, 1961), Sec. 5.

J. L. Mackie (1967)

MILTON, JOHN
(1608–1674)

John Milton, the English poet, author, and political writer, was born in London, the son of a prosperous scrivener. He was educated at St. Paul's School in London and Christ's College, Cambridge. After receiving an M.A. in 1632, he spent six years in study at his father's estate in Horton. In 1638 and 1639 he traveled to Italy, where he met Galileo Galilei, and on his return to London he found employment as a tutor. He wrote five pamphlets (1641–1642) attacking episcopacy, and his unhappy marriage in 1642 lent intensity to his subsequent tracts on divorce. In 1644 he published the tract *Of Education,* as well as *Areopagitica,* his famous attack on censorship of the press. His pamphlet justifying regicide, *Tenure of*

Kings and Magistrates (1649), probably brought him the post of secretary for foreign tongues to the Council of State. He wrote several defenses of the revolutionary government, but after 1652 total blindness forced him to withdraw gradually from public life. He turned to the completion of his theological treatise, *De Doctrina Christiana,* and his *History of Britain* and to the fulfillment of his poetic ambitions. Despite a brief return to public controversy in 1659 and 1660, Milton was treated leniently by the Restoration government. His epic, *Paradise Lost,* was published in 1667; *Samson Agonistes* and *Paradise Regained* appeared together, in one volume, in 1671. He died in 1674, survived by his third wife.

APPROACH AND METHOD

Milton was essentially a religious and ethical thinker, and his views are a striking blend of Christian humanism and Puritanism. The fullest statement of his position is *De Doctrina Christiana,* which was complete in all but certain details by 1660.

Milton believed that the Bible is divine revelation, plain and perspicuous in all things necessary to salvation. In matters of religion Scripture is the only outward rule or authority, and conscience, illuminated by the spirit of God, the only guide within. This scrupulous biblicism, however, is linked (as in Socinianism) with a strong emphasis on reason. Conscience, even when illuminated by the spirit, operates in rational terms rather than through mystical insight, so that "right reason" becomes the guide to Scripture. At the heart of this view, authorizing yet limiting the role of reason, is the doctrine that Scripture is an accommodation of God's will to the limited understanding of man. God has made in the Bible as full a revelation of himself as man is capable of receiving, and the safest approach is thus to form in the mind "such a conception of God, as shall correspond with his own delineation and representation of himself." This view eliminates speculations of a transcendental kind, reserving an area of mystery into which reason may not trespass; at the same time it encourages reason to assimilate biblical revelation to the categories of ethics. Thus, the theological treatise, like *Paradise Lost,* is a theodicy; its aim is to discover a view of God that is both worthy of him and consistent with revelation.

THEOLOGY

Milton's aim led him to some unorthodox conclusions, the most striking of which is his rejection of the doctrine of the Trinity. Embracing a loosely Arian position, he insisted on the unity of God and the consequent subordination of the Son and the Holy Spirit to the Father. The Son is the first of the creatures, and although he is the perfect image of the Father and even made of the same substance, he is not of one essence with the Father. The Spirit, a rather supernumerary figure, was created at a later date than the Son. Milton maintained that the doctrine of the Trinity is a purely manmade mystery, with no scriptural foundation; it defies logic and degrades our conception of deity.

There was a second deviation from orthodoxy in the direction of monism. Milton rejected the Augustinian doctrine of the creation of the world ex nihilo and presented a theory of creation *de Deo.* Drawing support from both Scripture and reason, he argued that the universe was made out of the substance of God. This view, he claimed, is not only more logical than the alternative position, but in its assertion of the goodness of matter it underlines more emphatically the benevolence of the creator. The same antiascetic impulse is present in Milton's theory of body and soul; he argued that the higher comprehends the lower, that spirit contains matter, and that the body should thus be seen not as the prison house of the soul but as integral to it: "The whole man is soul, and the soul man." From this conclusion two corollaries proceed: first, the human soul is not created immediately by God but is propagated from father to son in a natural order; second, the whole man dies, body and soul, and does not live again until the end of time. Milton's view of spirit and matter probably encouraged both his rejection of traditional Eucharistic theory and his radical endorsement of divorce and polygamy.

FREE WILL. The doctrines we have examined, which are departures from the main traditions of Christianity, were designed to avoid dualism and to make theology conform to the canons of logical thought. A second group of doctrines emerged as a defense of free will against Calvinism. Milton rejected the orthodox Calvinist view of predestination and reduced the decree of predestination to a general offer of salvation to all men who are willing to believe. Other Arminian views reinforced his conviction that man is free to pursue or refuse salvation. Milton wished to show that regeneration is a matter neither of faith nor of works but of works of faith. Faith, it is true, is a gift of God, but every man is given sufficient grace to put a saving faith within his reach. Finally, the object of a saving faith is God the Father rather than Christ, so that such a faith is possible beyond the bounds of the Christian religion.

ETHICS

The relation of the individual to the community absorbed Milton's attention during two decades of public controversy (1640–1660). His tracts, written in response to the disturbing events of the period, received force and direction from his lasting concern with liberty. Reason is "but choosing"; it is the power of ethical action, and man must therefore be free to choose between good and evil. Only by knowing evil and rejecting it can one become virtuous, for, as Milton remarked in *Areopagitica,* "That which purifies us is trial, and trial is by what is contrary." Prescriptive morality, enforced by church or state, prevents both the real understanding of truths already known and the discovery of new truths.

Milton defended the autonomy of reason by appealing from manmade authorities—positive law, canon law, custom, or tradition—to the law of nature. The work of John Selden probably encouraged him to develop a distinction between the primary law of nature, given to Adam at the creation, and the secondary law, the imperfect remnants of the primary law in fallen man. Secondary law allows for the "hardness of heart" that was introduced by the Fall and thus prescribes for such aspects of man's fallen state as war, servitude, divorce, and private property. In *De Doctrina Christiana,* however, Milton stressed the importance of the primary or unwritten law of nature that was "given originally to Adam, and of which a certain remnant, or imperfect illumination, still dwells in the hearts of all mankind; which, in the regenerate, under the influence of the Holy Spirit, is daily tending towards a renewal of its primitive brightness." This law teaches whatever is intrinsically good and agreeable to right reason, and in making it the final authority, Milton gave his ethic a religious orientation.

Thus, Milton's ethical position was that of the Christian humanist. Grace, he believed, comes to perfect nature, not to destroy it; by means of grace reason is illuminated and natural virtue sanctified. In this emphasis he resembled the Cambridge Platonists, writers like Benjamin Whichcote, John Smith, and Nathanael Culverwel, who sought to unify man's natural and religious experience by insisting that reason is "the candle of the Lord." Milton also resembled these philosophers in his habit of drawing upon Platonic writings, particularly on Plato's myths, in order to enrich his treatment of reason and the passions. Although his stress on the Bible prevented classical philosophy from making a direct contribution to his theology, Platonism nonetheless played a major and continuous part in shaping his ethical idealism.

The influence of Puritanism, as well as of humanism, led Milton to stress the importance of liberty. Believers are a "royal priesthood," and those who force the conscience of the individual are guilty of forcing the spirit of God. Central to Milton's conception of Christian liberty is the distinction between the Mosaic law, a law of bondage that extorts servile obedience through fear, and the Gospel, which offers a free, elective, and spiritual service based on man's filial relation to God. Spiritual regeneration, moreover, brings about a renewal of man's natural powers; the understanding is restored in large measure to its primitive clearness, the will to its primitive liberty. This strong emphasis on inner law led Milton to the antinomian view that Christ, by his life and death, abrogated the whole Mosaic law, the moral parts as well as the judicial and ceremonial parts. The sum of the law—love God and love your neighbor—remains and must be fulfilled by following the spirit, or the "internal scripture" (*De Doctrina Christiana,* I, xxvii). At this point, in spite of a continuing emphasis on reason, Milton had moved toward a position similar to the Quaker doctrine of inner light.

CHURCH AND STATE

Despite his early support of Presbyterianism, Milton soon came to believe that "*New Presbyter* is but *Old Priest* writ large." He defended the growth of religious sects on the ground that God requires unity of spirit rather than unity of doctrine, and he denied both the claim of the church to exercise secular power and that of the state to wield ecclesiastical power. His final view was that a particular church is a purely voluntary association of believers. Ministers should be elected by their congregations and supported by free offerings, and no ceremonial observances, such as the Sabbath, should be made obligatory. Despite his separation of the powers of church and state, however, Milton could not follow his more radical contemporaries in divorcing civil good from the good of religion. Although he denied the magistrate "compulsive" powers in matters of religion, he left him the "defensive" function of protecting Protestant Christianity from the threat of open "popery and idolatry."

Milton's view of the state varied in accordance with the changing conditions in which he was called upon to defend the revolutionary party. A basic line of his argument founds the state upon a social contract. Men are born free, but the effects of the Fall cause them to agree to a common league to bind one another from mutual injury. The people are thus the sovereign power in the state and have the right to revoke the power that they

have delegated. When it became apparent that the Puritan party represented a small part of the nation, Milton resorted to a further argument that was not entirely consistent with the social contract theory. The revolutionary party, he maintained, was guided by providence and consisted of those most worthy to rule and to interpret the good of the people. The minority must force the majority to be free.

POETRY

The themes and preoccupations of Milton's prose gain in power when expressed in the "more simple, sensuous, and passionate" language of poetry. All the major poems center on the theme of temptation and move toward a clarification of true heroism. Temptation works through passion, in its simplest form through sensuality and anger but more subtly through specious reasoning and the lure of evil means to good ends. The definition of true heroism involves the exposure of such false forms as the romantic sensuality of Comus in the early "Masque" (1634) or Satan's courage of despair in the late epics. *Paradise Lost,* which was written to justify God's ways to man by dramatizing man's freedom and responsibility, ends with Adam setting out to imitate the spiritual heroism of the Son of God—revealed to him in a vision—and thus to achieve a "paradise within" that will be "happier far" than the outward paradise he has lost. Samson, in *Samson Agonistes,* also achieves a victory over himself through suffering and discovers that freedom is enjoyed only in the service of God. *Paradise Regained,* which has as its subject the temptation of Jesus in the wilderness, presents Milton's final and most complete study of heroism. Avoiding the temptations to distrust and presumption, the Son rejects Satan's offers of worldly power and authority and realizes the spiritual sense in which he is Messiah.

ARTS AND SCIENCES

In his literary theory Milton emphasized the importance of genres and of decorum and urged the power of literature to create moral order in the individual and the society. (See his preface to Book II of *The Reason of Church Government,* the preface to *Samson Agonistes,* and the invocations to Books I, III, and IX of *Paradise Lost.*) His view of education (*Of Education*) was humanistic in its stress on languages and classical texts, its dislike of scholasticism, and its ethical aim. He showed no deep interest in the new science, and he used the traditional science in his poetry because it was for him a better source of metaphor. As a historian he had a critical sense of the value of evidence, but his view of history moved from millenarian optimism to the pessimism that informs the survey of history in the last two books of *Paradise Lost.*

See also Arius and Arianism; Culverwel, Nathanael; Determinism and Freedom; Galileo Galilei; Humanism; Liberty; Plato; Platonism and the Platonic Tradition; Smith, John; Socinianism; Whichcote, Benjamin.

Bibliography

WORKS BY MILTON

The definitive edition of the text is *The Works of John Milton,* edited by Frank Allen Patterson et al., 18 vols. (New York: Columbia University Press, 1931–1938), with an index, 2 vols. (New York, 1940). A supplement to the Patterson edition, *A Variorum Commentary on the Poems of John Milton,* with Merritt Y. Hughes as the general editor, was also published by Columbia University Press (New York, 1970).

There are editions of the poetry and selected prose by Frank Allen Patterson, *The Student's Milton* (New York: F. S. Crofts, 1930; rev. ed., 1933); Merritt Y. Hughes, *Complete Poems and Major Prose* (New York: Odyssey Press, 1957); and others. Editions of the poetry include those by James Holly Hanford, *Poems* (New York, 1937; 2nd ed., 1953); Harris Francis Fletcher, *Complete Poetical Works* (Boston, 1941); and Helen Darbishire, *Poetical Works,* 2 vols. (Oxford, 1952–1955).

Scholarly introductions provide continuous commentary on the prose in the closely annotated *Complete Prose Works,* edited by Donald M. Wolfe et al. (New Haven, CT: Yale University Press, 1953–1982), 8 volumes.

WORKS ON MILTON

General Criticism

Douglas Bush provides a penetrating review of Milton's life and works in *English Literature in the Earlier Seventeenth Century* (Oxford: Clarendon Press, 1945; rev. ed., 1962); a survey of the scholarship is found in James Holly Hanford, *A Milton Handbook,* 4th ed. (New York: F. S. Crofts, 1946). The standard biography is still David Masson's *The Life of John Milton,* 6 vols. (London: Macmillan, 1859–1880; rev. ed., with index, 1881–1896).

Milton's Thought

Pioneer work is found in Denis Saurat's stimulating if erratic *Milton: Man and Thinker* (London: Dent, 1925; rev. ed., 1944) and in the more literary *Milton* (London, 1930) by E. M. W. Tillyard. G. N. Conklin considers theological method in *Biblical Criticism and Heresy in Milton* (New York: King's Crown Press, 1949), and the growth and significance of Milton's theology are examined authoritatively by Maurice W. Kelley in *This Great Argument* (Princeton, NJ: Princeton University Press, 1941). On Milton's political and ethical views, see A. S. P. Woodhouse, "Milton, Puritanism and Liberty," in *University of Toronto Quarterly* 4 (1934–1935): 483–513; William Haller, *The Rise of Puritanism* (New York: Columbia University Press, 1938) and *Liberty and Reformation in the Puritan Revolution* (New York: Columbia

University Press, 1955); Arthur Barker, *Milton and the Puritan Dilemma, 1641–1660* (Toronto: University of Toronto Press, 1942; reprinted, 1955); Michael Fixler, *Milton and the Kingdoms of God* (London: Faber and Faber, 1964); Ernest Sirluck, "Milton's Political Thought: The First Cycle," in *Modern Philology* 61 (1964): 209–224.

Other aspects of Milton's thought are covered in Kester Svendsen, *Milton and Science* (Cambridge, MA: Harvard University Press, 1956); Howard Schultz, *Milton and Forbidden Knowledge* (New York: Modern Language Association of America, 1955); Walter C. Curry, *Milton's Ontology, Cosmogony and Physics* (Lexington: University of Kentucky Press, 1957); and Irene Samuel, *Plato and Milton* (Ithaca, NY: Cornell University Press, 1947). Further criticism is listed in a selective bibliography by Douglas Bush, op. cit.; in the bibliographies by David H. Stevens, *A Reference Guide to Milton, from 1800 to the Present Day* (Chicago: University of Chicago Press, 1930); Harris Francis Fletcher, *Contributions to a Milton Bibliography, 1800–1930* (Urbana: University of Illinois, 1931); and Calvin Huckabay, *John Milton: A Bibliographical Supplement, 1929–1957* (Pittsburgh: Duquesne University Press, 1960); and the annual bibliographies of *Studies in Philology* and *PMLA*.

H. R. MacCallum (1967)

MIMESIS

Mimesis has been a cardinal concept for those traditions of aesthetics, from antiquity to the present, that focus on the status and value of artistic representation. The semantics of the Greek term *mimēsis* cover much more than simple imitation; its senses include resemblance, dramatic impersonation, and other species of correspondence or likeness. The idea of mimesis came to designate the relationship between certain art forms (poetry, dance, music, painting/sculpture) and the aspects of reality they are capable of depicting or evoking. Although some strands of mimeticist thinking appeal to standards of verisimilitude and mirroring, it is mistaken to reduce all models of mimesis to a single canon of realism.

Plato's highly influential approach to mimesis is less straightforward than usually claimed. From *Cratylus* to *Laws*, he applies the language of mimesis to numerous relationships of ontological and/or semantic dependence (even, in *Timaeus*, e.g., 39e, the whole material universe's dependence on a divine prototype). Mimetic entities match, but never reproduce, their exemplars; the relationship can be construed as "qualitative," not "mathematical" (*Cratylus* 432). In representational art, moreover, those exemplars may be (partially) imaginative/fictive: witness, for example, the idealized painting that furnishes a metaphor for philosophy at *Republic* 472d. When, in *Republic* 10, Socrates notoriously critiques the mirror-like limitations of mimetic poetry and painting, locating artistic images at two removes from "the truth," his argument does not convict all mimesis of worthlessness but provocatively challenges lovers of art to identify a moral justification that transcends pleasure at merely simulated appearances (and the emotions they can excite). As *Sophist* 235d–6c, distinguishing *eikastic* (objective) from *phantastic* (viewer-dependent) mimesis, shows, Plato does not ascribe a uniform rationale to all artistic representation. At a psychological and cultural level, arguments such as *Republic* 392c–401a suggest that the impact of mimesis necessarily reflects the qualities of the supposed reality it projects.

Aristotle explicitly accepts that the contents of mimetic art, both musicopoetic and visual, can legitimately vary between the actual, the putative, and the ideal (*Poetics* 25). Regarding mimesis as an instinctual factor in the human need to model and understand the world, he embeds it in an anthropology of cognition that stretches from children's play to philosophy (*Poetics* 4). He also appreciates the powerful emotional effects of mimetic works on their audiences, a point equally illustrated by the *Poetics* and by the treatment of music as mimetic (i.e., affectively expressive) in *Politics* 8.5; for him, the passions, when well induced, are a medium of ethical judgment. Furthermore, Aristotle has a dual-aspect conception of mimesis that allows him to distinguish—more than Plato had done—between internal (work-centered) and external (truth-related) criteria of mimetic value. The resulting aesthetics is, importantly, neither formalist nor moralist.

Hellenistic and later Greek philosophers continued to grapple with epistemological and ethical issues raised by mimesis. Especially notable is Neoplatonism's ambivalent engagement with the concept; Plotinus, for instance, who discerned mimetic relationships hierarchically structuring all reality, disparaged much actual art yet allowed some artistic mimesis, *qua* creative intuition, to grasp the authentic forms of nature (*Enneads* 5.8.1). The legacy of this and other ancient versions of mimesis was revived in the Renaissance; it has remained a vital element in debates about the complex position of representational art between the poles of truth and fiction, realism, and imagination.

See also Art, Representation in.

Bibliography

Halliwell, Stephen. *The Aesthetics of Mimesis: Ancient Texts and Modern Problems*. Princeton, NJ: Princeton University Press, 2002.

Janaway, Christopher. *Images of Excellence: Plato's Critique of the Arts*. Oxford: Clarendon Press, 1995.
Sörbom, G. *Mimesis and Art*. Stockholm: Svenska Bokförlaget, 1966.

Stephen Halliwell (2005)

MINAGAWA KIEN
(1734–1807)

Minagawa Kien, a Japanese Confucianist, painter, and writer, was born in Kyoto. At the age of twenty-eight, having established himself as a Confucianist, he became the official scholar for Lord Matsudaira Nobumine. His literary skill made him an outstanding figure in Kyoto circles; he had a following of three thousand. For a Confucianist his life was unusually dissipated. His era was a time of moral decline, but this was eventually checked by several edicts. The 1790 edict against "heterodox doctrines" affected Minagawa and he reformed his habits, though his ideas did not change.

Minagawa's philosophical reputation has recently grown among Japanese philosophers because of his positivist approach to Confucian studies. He is considered an eclectic because he upheld neither the official Zhu Xi school of Neo-Confucianism nor the rival Wang Yangming school. Minagawa was analytic and positivist, which made him a kind of forerunner of Western philosophy in Japan. This assessment stems largely from two of Minagawa's works, *Ekigaku kaibutsu* (The learning of the book of changes on the discovery of things) and *Meichū rokkan* (Six chapters on categories).

Ekigaku kaibutsu starts from the Chinese classic *I Ching,* the "Book of Changes" or "Book of Divination," which despite its esoteric nature stimulated Minagawa and other Confucianists to make a study of celestial phenomena. *Ekigaku kaibutsu* clearly manifests his lifetime search for the nature of things. However, for him "things" are mainly human affairs seen from the ethicopolitical point of view, and their "discovery" or investigation is in relation to the ruling of the realm.

Meichū rokkan analyzes the origins of basic concepts or categories. Starting with words, Minagawa shows that they are abstract expressions of reality itself. He believes that we grasp reality objectively through its manifestation in words. This rather naive realist epistemology is an attempt to penetrate the nature of things without employing *ri,* Zhu Xi's abstract "principle," or the "innate knowledge" of Wang Yangming. Among Minagawa's categories, significant ones are learning or science (*gaku*) and wisdom (*tetsu*). Although he did not wholly grasp modern science or philosophy, he came very close.

Another topic of interest to Minagawa is the samurai class, which he criticizes in many of his writings. He hoped the samurai would survive as the intellectual and moral leaders of the ordinary people.

See also Chinese Philosophy; Japanese Philosophy; Positivism; Wang Yangming; Wisdom; Zhu Xi (Chu Hsi).

Bibliography

For Minagawa's works see *Nihon Tetsugaku Shisō Zensho* (Library of Japanese philosophical thought), edited by Saigusa Hiroto (Tokyo, 1957), Vol. I, pp. 109–119, and Saigusa Hiroto, *Nihon Yuibutsuronsha* (Japanese materialists; Tokyo, 1956), pp. 95–107.

Gino K. Piovesana, S.J. (1967)

MIND

See *Idealism; Mind-Body Problem; Other Minds; Personal Identity; Psychology; Reason; Thinking*

MIND AND MENTAL STATES IN BUDDHIST PHILOSOPHY

A fundamental idea of all nonmaterialist Indian schools of philosophy, whether orthodox ones that follow the Vedas or heterodox ones such as Buddhist and Jaina that do not, is the cultivation of mind and mental states. Techniques of *yoga* in Hindu tradition aim at attaining a conscious state in which ordinary mental activities, such as perception and imagination, are suspended. Classical yoga, as expounded by Patanjali's *Yogasutra* (Woods, 1927), is widely influential in the Hindu tradition.

ORTHODOX AND HETERODOX SCHOOLS

In Buddhism, *citta, mano*, and *vinnana* are three of the main terms to do with mind and mental states. These terms are highly nuanced but are roughly translatable as heart, mind, and consciousness, respectively. These are best understood as processes, not substances, and none are permanent. The *Majjhima Nikaya* (Middle length sayings), *Digha Nikaya* (Long discourses), *Samyutta Nikaya* (Kindred sayings), and *Anguttara Nikaya* (Gradual say-

ings) are the basic four collections of *suttas* (discourses) expounding the early Buddhist position, and *Vissudhimagga* (The path of purity) is a salient text.

Indian schools of philosophy include three heterodox (*nastika*) schools, which do not accept the Vedas as divine revelation. These three schools (Carvaka, Jainism, and Buddhism), each in their different ways, put more emphasis upon experience than revelation. The three schools represent a continuum on metaphysical matters from most materialistic (Carvaka) to least materialistic (Buddhism). Jainism at midpoint asserts a material, adhesive soul that gets darkened with negative karmic particles due to wrong actions such that many *jivas* (souls) cannot retain their natural luminosity.

According to ancient Indian materialism (Carvaka school), perception is the basic *pramana* (valid means of knowing), and accordingly, matter is the only reality because it alone is perceived. Here the soul is understood as a living body with the quality of consciousness. But how could materialists show that consciousness does not exist independently of body? Orthodox schools as well as the other two heterodox schools, Jainism and Buddhism, found materialistic reductionism of the mental to the physical unconvincing.

JAINISM. Jainism is especially well known for two doctrines: the view that all judgments of non-omniscient beings need to be qualified—that is, the "somehow view" (*syadvada*); and non-injury to sentient beings—that this, the nonviolence view (*ahimsa*). According to Jainism, consciousness is the essence of the *jiva*, and human consciousness is limited so that ordinary judgments of nonomniscient beings must be qualified by *syat* (somehow) to express conditional knowing. Only one of the *Tirthankaras*, that is, those who cross over to liberation, have omniscience in regard to salvific knowledge. In Jainism the *jiva* is self-luminous and illuminates other things, filling out the body like a radiant, eternal light within it. Jains believe that the *jiva* can attain complete freedom (*kaivalya*). When the *jiva* is in a state of ignorance or bondage, it is because its vision is obscured due to karmic particles adhering to it. So, although Jainism has a spiritual, ethical outlook that aspires to personal self-transformation, its metaphysics of the soul holds that the soul is material, of the shape of the body, and is afflicted by karmic particles. When these are thrown out of the *jiva* due to penance or good works, the *jiva* can see clearly. Harming living beings is one thing that causes karmic particles to cloud the soul's vision. In ethics, Jains think that the passions impeding liberation are anger, pride, infatuation, and greed. These sorts of passions bind the *jiva* to matter. Since there is consciousness in all parts of the body, the soul is coextensive with the body. Potentially, all souls are equal since all have the capacity for liberation (*kaivalya*).

BUDDHISM. Another of the heterodox schools, Buddhism, holds that right concentration of mind through four stages is the way to *nirvana* (enlightenment). The first stage is on reasoning and investigation regarding the truths; here there is the joy of pure thinking. The second stage of concentration is unruffled meditation, freedom from reasoning, and the arising of the joy of tranquillity. The third stage of concentration is detachment from even the joy of tranquillity; here there is indifference even to such joy and a feeling of bodily ease. The fourth stage of concentration is detachment from this bodily ease: At the fourth *jhana* (level of consciousness in meditation), there is perfect equanimity and the attainment of nirvana. At this level the psychic powers (*abhinna*) are said to develop. Overall, *sila*, *samadhi*, and *panna* (morality, concentration, and wisdom, respectively) form the essentials of the eight-fold noble path in Buddhism (right view, right intention, right speech, right action, right livelihood, right minfullness, right effort, right contemplation, right concentration). In Buddhism there is no permanent substance (*svabhava*) either in humankind or in deities, for experience shows that all things are impermanent, nonsubstantial, and unsatisfactory. The doctrine of *anatman* (no self, or nonsubstantiality) implies that there is no substance of a permanent, blissful, center of consciousness anywhere in the universe.

The doctrinal context of *jhana* is four noble truths: suffering, its arising, passing away, and the path to its passing away. The cessation of suffering occurs through meditation. The *jhanas* were instrumental in Buddha's enlightenment in that *jhanas* prepare one for higher insights (*abhinna*), are associated with liberating wisdom (*panna*), and are the spiritual endowment of the fully liberated person (*tathagata*). *Jhanas* have their own internal dynamic, contributing to purification and liberation of mind. In developing *jhanic* insight, one focuses on experience, eliminates ignorance, and achieves wisdom. There are really two systems: tranquillity and insight. The development of serenity or tranquillity meditation (*samatha bhavana*) is one system; the other is the development of insight meditation (*vipassana bhavana*) is the other. The former is also called development of concentration (*samadhi bhavana*); the latter is also called the development of wisdom meditation (*panna bhavana*). The practice of serenity meditation aims at developing a calm,

concentrated, unified state of consciousness to experience peace and wisdom. Insight meditation requires development of *samadhi*, and serenity is useful for this too, so the two systems work together. *Jhana* belongs inherently to the serenity side. Translation of *jhana* is difficult, with absorption coming closest. *Jhanas* involve total absorption in the object.

CONCEPTUAL STRUCTURES IN BUDDHISM

ORALITY AND MENTALITY. Oral tradition and group recitation of sutras marked the very beginnings of Buddhism of the Pali Nikayas (collections of *suttas* in different texts, e.g., *Majjhima Nikaya*). Despite the strong tradition of text, commentary, and subcommentary, Buddhism initially developed from oral tradition, as did Hinduism. In contrast with the European and North American preoccupation with journal articles and books as vehicles for intellectual debate, the power of the spoken word remains very much a part of Buddhism. This power of the spoken word can be seen, for example, in the Indo-Tibetan tradition of debate and the Sino-Japanese *kung-an* and *koan* traditions of perspectival shifts while becoming one with the *koan*.

It is clear is that Buddhism did not begin with manuscripts. It is not a religion as in the monotheistic (Judeo–Christian–Islamic) tradition but developed out of a forest tradition of meditation in which monks stayed in orchards, deer parks, mango groves, and forests, periodically reciting the words of the Buddha aloud in group recitation. Eventually, councils and canons of texts emerged. It was not so at first, and it is reasonable to believe that the authority of individual experience is at the heart of early Buddhism rather than hierarchy and the authority of promulgated texts.

MIRACLES OF INSTRUCTION, CONVERSION, AND MINDFULNESS. An unrepeatable event, violation of law of nature, and any extraordinary event are senses of *miracle* ordinarily recognized in Anglo-American philosophy of religion as a starting point for discussion. In Buddhism, the miracle of instruction is the starting point. Traditionally, one has to come and sit down by the side of the teacher. Texts show that *dhamma* (truth, doctrine) teaching sometimes includes a miracle, where conversion occurs and miracle becomes part of the experience of a Buddhist practitioner.

Oral recitation makes of oneself a holy scripture as the embodiment of truth: Truth is not so much a property of abstract disembodied proposition as it is embodied in the lives of those who practice Buddhism. Belief in the Buddha, the doctrine, and the Sangha (order of monks and nuns) is the recited *three refuges* formula for being Buddhist. Both confidence and knowledge are operative in Buddhism, both *belief in* and *belief that*. Buddhism did not emphasize authority of the guru or pundit but the authority of one's own experience, so there is no *blind faith*.

The baseless faith of the Brahmins is contrasted with the rational faith of the Buddhists. Brahmins are depicted as a string of blind people, each relying on the other but none of them seeing things as they really are. Buddhism is, by contrast, self-reliance, with several stages of confidence or faith. There is initial faith in coming to hear whether there is anything in the Buddhist doctrine, then there is path faith that is compatible with doubt and struggle, and then there is the achievement of a realized nonbacksliding faith; realized faith is the wisdom of knowing and seeing for oneself as things really are.

MIND AND MORALITY. By mind all things are made, all things are made by mind: Thus begins the *Dhammapada* (The path of purity), a popular Buddhist text. Morality is intimately connected with mentality on the Buddhist view, and intention is far more important than consequences in assessing *sila*, or morality. It would go too far to say that consequences are totally irrelevant to Buddhists: Following the first precept of harmlessness shows a concern with outcomes as well.

Buddhism defies categorization in Aristotelean, Utilitarian, and Kantian categories, not because of this conceptual confusion but because of its distinctive voice. Buddhism is most importantly about wisdom, not knowledge alone, and it is also about compassion, which is one of the ways to enlightenment. Although Mahayana Buddhism emphasized altruism and Theravada Buddhism had comparatively little to say about kindness and compassion, it is clear that there are Pali Canon texts that commend kindness, and value it as a means of attaining nirvana (Gombrich 1998). *Metta*, *karuna*, and *mudita* (loving kindness, compassion, and sympathy) are valued, ethically related mental states in even the earliest stratum of Buddhism, just as *priti* (joy) is a characteristic of Buddhist monks.

MEDITATION AND CONFIRMATION OF PRE-EXISTING BELIEFS. There is an epistemological basis for belief in propositions concerning *kamma* and *punabbhava* (rebirth; literally, "again becoming"). This emphasis on one's own experience extends even to epistemology,

where the *pramana* (valid means of knowing) of experience and, to a limited extent, inference based on experience, are emphasized instead of testimony, comparison, and divine revelation. The epistemological basis of belief in karma and rebirth is said in the texts and by modernist interpreters such as K. N. Jayatilleke (1963), K. N. Upadhyaya (1998), and D. J. Kalupahana (1992), to rest on meditational experience at the fourth *jhana*.

Some in Buddhism hold that knowing and seeing rebirth provides *empirical justification* for belief in karma and rebirth. These same thinkers believe that Buddhism has no metaphysics. However, first, it is dubious that memory, bodily continuity, or self-awareness will work as meaning conditions for the reidentification of the same person across lives. Second, metaphysics is not the same as speculation, and Buddhism can be antispeculative and still have metaphysical commitments to beliefs such as rebirth.

It is tempting to think of Buddhism as empiricism since it is described in the Pali texts as a *come and see* (*ehipassika*) doctrine, but while its claims may, in a weak sense, be experientially verifiable if true, they are not falsifiable if false. Hence they are not verifiable in a sufficiently robust sense to distinguish Buddhism from other path faiths and to count as *empirical verification*. What is at work, instead, is *experiential confirmation*. In addition, the mind and senses are not separated in Buddhism but are together the *six gateways* to knowledge so that there is no sharp cleavage between empiricism and rationalism, as there is in European and North American thought. All that can be had in Buddhism is experiential confirmation, as in the cases of other worldviews, such as that of Christianity. Psychological certainty is not identical to logical certainty. Experiential justification may be entirely convincing on a personal basis yet fall short of the objectivity involved in establishing the truth of observation, sentences that are testable and repeatable at will.

CONTINUITY, PERSONAL IDENTITY, AND *NAMARUPA*. The strength of a cord does not always depend on something running end to end, as in Buddhism where there is continuity of process but no speculative belief about a permanent substance underlying it all. In Buddhism, *vinnana* (consciousness) develops (rather than descends) in the womb in the rebirth process across lives. There is no one term that provides a link between lives in early Buddhism. Perhaps *sankhara* (dispositions) comes closest.

A view that superficially looks like the Buddhist one is Hume's *phenomenalist* view of the self. Here, the self is a bundle of perceptions. Hume famously says that all perceptions are distinct existences and that the mind never recognizes any necessary connections between these perceptions. However, one does not find exactly this view in Buddhism. Hume had a problem with combining the two assumptions about distinct existence of perceptions and no necessary connections, but early Buddhism's problem is not Hume's problem: To ask what keeps the perceptions of a person together in early Buddhism is to make what from an early Buddhist view is the unwarranted assumption of the distinct existence of perceptions.

Namarupa may be understood as that which appears (appearance or phenomenon) in its interrelationship with *nama*, or that which one uses to get a handle on an appearance (the concept). So *namarupa* is the reality formed by the unity of concept and phenomenon; it is conceptualized reality or the process of ordinary experiencing. Inadequate are "mind and body" or "name and form" as translations (Ross Rheat, in Potter: 1996 VII 45). It is evident that *namarupa* provides no evidence for substantialist mind-body dualism in early Buddhism. As Surendranath Dasgupta rightly observes (1922), matter and mind dualism and opposition are absent from Buddhism, Upanishads, and Samkhya schools of philosophy. Overall, Buddhism—which differs from Hume on the point of distinct existences—on the issue of *self*, is closer to Process philosophy than to either Humean empiricism or Cartesian rationalist dualism.

"LIFE AFTER DEATH": ETERNAL LIFE AND ENDLESS LIFE. In macro view the *punabbhava* rebirth realms, that is, humans, gods, animals, hungry ghosts, purgatory beings, and titans, may be viewed ontologically or psychologically. Viewed ontologically, in the Buddhist metaphysical view of the process of rebirth, the ordinary case is that one is reborn. There is also the extraordinary case of the *Tathagata* (the *thus gone* liberated one, e.g., Buddha Sakyamuni) who passes away in *parinibbana* (final enlightenment) having achieved *nibbana* (enlightenment) in this very life. Yet, no early Buddhist text gives a theory about what, if anything, happens after death in the case of the *Tathagata*. Afterlife views are regarded as speculative and discussing them not conducive to enlightenment. The antispeculative emphasis informs the Ten Speculative Questions (speculative questions that the Buddha would not commit to answering because they involve knowledge claims that go beyond experience) set aside by Buddha. The deathless (*amata*) may be viewed simply as the elimination of obsession, hate, and confusion in everyday life of the Buddhist practitioner.

Heaven (*devaloka*), the world of the gods, is simply another rebirth station. What is translated, *devaloka* is neither a permanent resting place nor a monotheist's beatific vision. From it some *devas* (the shining ones) may be reborn elsewhere, including as humans, before attaining final liberation.

The Buddhist goal is stopping the wheel of birth and death rather than attaining endless life. The emphasis is on attaining eternal life in the here and now by purifying ones heart and living well. In this conceptual scheme in which impermanence, nonsubstantiality, and suffering play key roles, the idea of striving after an immortality viewed as endless life would be not simply be unattainable but logically incoherent.

Accordingly, terms for mind and mental states in Buddhism are not terms for a permanent stuff or substance that is independent of conditions. Saying so does not deny continuity across lives. There is continuity without self-same substance. There is a stream of consciousness depending for its continuance on union of male and female, proper timing, and presence of *gandhabba* (cupid). Without these three conditions, there is no rebirth.

NIRVĀṆA. That Buddhist rebirth is not Hindu transmigration is evident from the anatta doctrine of Buddhism juxtaposed with the atman doctrine of Hinduism. At the level of meditation, there is considerable overlap of technique; however, such that an attempt to forge a complete disjunct between these two traditions will distort both history and practice. Buddha was born a Hindu and is considered by Hindus as an avatara of Vishnu. For polemical and practical purposes of building a Sangha, Buddhist texts routinely depict Buddhists triumphing over Jains and Brahmins in debate. So there is a distinctive Buddhist mentality such that Buddhism will never be rightly described as assimilable to Hinduism without remainder.

Early Buddhist texts are not perfectly consistent in the use of terms for the state of consciousness called enlightenment or being awake. However, a frequent finding is that *nibbana* (enlightenment) while alive is distinguished from *parinibbana* (final enlightenment) after death of a *Tathagata*. This distinction is subject to a range of textual emphases and resultant interpretations. The simplest, most clear way to draw the distinction is to say that enlightenment in life is the destruction of *raga, dosa*, and *moha* (obsession, hate, and confusion) in everyday life; that final enlightenment is death of one who has already been enlightened in life.

That dying is, but death is not, an experience in life is itself a conceptual truth. Hence, it is not logically possible to experience death and describe it, and there are no mental states to be ascribed to *the Tathagata after death*. Asked whether the Tathagata exists, does not exist, both, or neither, Buddha refused to assent to any of these. Buddha's silence shows that the matter of final enlightenment (*parinibbana*) is beyond experience.

See also Aristotle; Ayer, Alfred Jules; Brahman; Buddhist Epistemology; Cartesianism; Hume, David; Kant, Immanuel; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Meditation in Indian Philosophy; Mysticism, the Indian Tradition; Negation in Indian Philosophy; Philosophy of Language in India; Self in Indian Philosophy; Truth and Falsity in Indian Philosophy; Utilitarianism.

Bibliography

Carter, John Ross, trans. *The Dhammapada*. New York: Oxford University Press, 2000

Collins, Steven. *Selfless Persons: Imagery and Thought in Theravada Buddhism*. New York: Cambridge University Press, 1990.

Dasgupta, Surendranath. *A History of Indian Philosophy*. Vol. I. Cambridge, U.K.: Cambridge University Press, 1922. p. 95.

Gombrich, R. F. *Kindness and Compassion as Means to Nirvana*. Amsterdam: Royal Netherlands Academy of Arts and Sciences, 1998.

Griffiths, Paul J. *On Being Buddha: The Classical Doctrine of Buddhahood*. Albany: State University of New York Press, 1994.

Griffiths, Paul J. *On Being Mindless: Buddhist Meditation and the Mind-Body Problem*. Delhi: Sri Satguru, 1999.

Gunaratna, Henapola. *Path of Serenity and Insight*. Delhi: Motilal Banarsidass, 1985.

Hamilton, Sue. *Early Buddhism: A New Approach: The I of the Beholder*. Surrey, U.K.: Curzon Richmond, 2000.

Herman, Arthur L. *The Problem of Evil and Indian Thought*. Delhi: Motilal Banarsidass, 1976.

Hoffman, Frank J. *Rationality and Mind in Early Buddhism*. Delhi: Motilal Banarsidass, 1987.

Hoffman, Frank J., and Mahinda Deegalle, eds. *Pali Buddhism*. Richmond, U.K.: Curzon Press, 1996.

Hoffman, Frank J., and Godabarisha Mishra, eds. *Breaking Barriers: Essays in Asian and Comparative Philosophy in Honor of Ramakrishna Puligandla*. Fremont CA: Asian Humanities Press, 2003.

Horner, I. B. *Women in Primitive Buddhism: Laywomen and Alms Women*. Delhi: Motilal Banarsidas, 1999.

Jayatilleke, K. N. *Early Buddhist Theory of Knowledge*. London: Allen & Unwin, 1963.

Kalupahana, David J. *History of Buddhist Philosophy*. Honolulu: University of Hawaii Press, 1992.

Lamotte, Etienne. *History of Indian Buddhism: From the Origins to the Saka Era*. Translated from the French by Sara

Webb-Boin under the supervision of Jean Dantinne. Louvaine-la-Neuve: Universite Catholique de Louvain Institut Orientaliste, 1988.

Matilal, B. K. *Logic, Language, and Reality: Indian Philosophy and Contemporary Issues*. Delhi: Motilal Banarsidass, 1999.

Mohanty, J. N. *Classical Indian Philosophy*. Lanham: Rowman & Littlefield, 2000.

Potter, Karl, ed. *Encyclopedia of Indian Philosophy*. Vol. VII. Delhi: Motilal Banarsidass, 1996.

Puligandla, Ramakrishna. *An Encounter with Awareness*. Wheaton, IL: Theosophical Publishing House, 1981.

Smart, Ninian. *Doctrine and Argument in Indian Philosophy*. London: Allen & Unwin, 1964.

Upadhyaya, K. N. *Early Buddhism and the Bhagavad-Gita*. Delhi: Motilal Banarsidass,1998.

Warder, Anthony Kennedy. *Indian Buddhism*. Delhi: Motilal Banarsidass, 2000.

Frank J. Hoffman (2005)

MIND-BODY PROBLEM

In Genesis 3:19, God tells Adam, "dust thou art, and unto dust thou shalt return," reminding Adam that he was fashioned from the dust of the earth. Modern science tells us that the earth was formed from the dust of the sun and that we are composed of materials formed from star dust. We are, however, also possessed of mind: We can think, feel, and exercise our will—as did Eve when she ate the forbidden fruit of the tree of knowledge of good and evil. The ancient mind-body problem is how the mind or soul or spirit is united with the body. It has now been known for several centuries that our minds are related to our bodies via their relation to a certain bodily organ, the brain. The ancient problem led to the mind-brain problem: How are our minds related to our brains? Are they one thing or two? And if two, how are the two united? But the fundamental problem is: What is the place of mental phenomena in nature?

The doctrine that the soul is distinct from the body, existing prior to it and after bodily death, is found in the writings of Plato. (In the *Phaedo*, one argument of Socrates for immortality is that the soul is not made of parts, and so cannot come apart.) The Platonic idea of a soul independent of the body was embraced by Augustine of Hippo, a major figure in the development of the Christian doctrine of an immaterial, immortal soul. But as to how soul and body are united, Augustine could only marvel: "The manner in which spirits are united to bodies is altogether wonderful and transcends the understanding of men" (*On the City of God*, XXI, 10 Haldane 1994, p. 335).

René Descartes tried to lay the foundation for a science of nature according to which all bodies are located in a physical realm—a substance, *res extensa*, which pervades all of space— and all interactions among them are governed by mechanistic laws. But mind (*res cogitans*), he argued, lacks spatial extension (and even location at a spatial point) and so is not subject to the mechanistic laws of the physical realm, thus leaving the will free. Minds, moreover, are substances and so capable of existence independently of physical substance; thus, immortality of the mind is possible. Descartes argued that it is certain that he is his mind since doubt itself requires a doubter and thus a thinking subject, an *I*. And he argued that he is not his body since he can clearly and distinctly conceive of his existing without a body and that it is thus possible for him to exist disembodied.

He nevertheless also acknowledged in *Meditations on First Philosophy* (1641): "there is nothing nature teaches me more expressly, or more sensibly than that I have a body, which is ill disposed when I feel pain, which needs to eat and drink when I have feelings of hunger and thirst, etc. … I am joined to it very closely and indeed so compounded and intermingled with my body, that I form, as it were, a single whole with it" (Cottingham et. al. 1985, p. 59). On his view, what unites body and mind is causation, from body to mind (as in perception), and from mind to body (as in action), with the pineal gland in the brain being the primary locus of such interaction. In correspondence with Descartes, Princess Elisabeth of Bohemia pressed the issue of how states of, or changes in, a substance not in space could causally affect states of, or changes in, something in space, and declared such causal interaction too incredible to believe. Descartes was never able to provide a satisfactory answer to her how-question, and in a candid moment remarked: "It does not seem to me that the human mind is capable of conceiving quite distinctly and at the same time both the distinction between mind and body, and their union" (Kenny 1970, p. 142).

Nicholas Malebranche denied mind-body causal interaction, maintaining that God is the only causal agent (Nadler 1999). Were a certain a type of brain state *B* and mental state *M* to co-occur, then that would be because God, who is continually engaged in acts of creation of the world, only causes an instance of one of them when he causes an instance of the other; *B* and *M* would thus co-occur are a result of being dual-effects of God's acts of creation. This brand of *parallelism* is called *occasionalism*. Of course, if God is without spatial extension or location, then Elisabeth's how-question will recur for God's causal

interaction with the physical world. But it was thought that how-questions come to an end where the ways of God are concerned. Gottlieb Leibniz held a version of parallelism, *preestablished harmony*, according to which there is no causal interaction among substances, any regularities among them being the result of God's having actualized a world in which those regularities hold. And he held a kind of *idealism*, according to which all substances are monads, which have only states of perception and appetite (Sleigh 1999). Benedict (Baruch) Spinoza rejected Descartes's claim that the mind is a substance, arguing that only God or Nature (*Deus, sive Natura*) is capable of independent existence, and took all mentality and physicality to be different modes of God or Nature. On his view, a kind of pantheism, we are each finite modes of God or Nature, and our mind and body are identical modes though conceived of under two different kinds of attributes: bodily and mental (Garrett 1999). He thus held a kind of *dual-aspect theory*. Thomas Hobbes, an atheist, held a version of *materialism*, reminiscent of the ancient atomism of Democritus and Lucretius—Lucretius wrote of atoms moving in an infinite void—according to which all that exists is matter in motion (Gert 1999). He tried to show how mental processes are just mechanical brain processes, maintaining that thinking is just computation, thereby anticipating the computational view of mental processes prevalent in contemporary cognitive science.

There is something deeply commonsensical about Descartes's interactionism. It seems that bodily sensations such as aches, pains, itches, and tickles cause us to moan, wince, scratch, or laugh and do so by causing brain states that result in bodily movements. In deliberate action, we act on our desires, motives, and intentions in trying to carry out our purposes; and acting on them seems to involve their causing brain states, which cause our muscles to contract, and so our bodies to move, thereby affecting our environment. Perception of the environment seems to involve physical to mental causal transactions: What we perceive causes us to undergo a sense experience. Thus, when we see the scenes before our eyes, for instance, those scenes cause our visual experiences via their effects on our brains. Descartes's substance dualism, however, seems untenable.

But suppose that minds have not just temporal location but spatial location as well. (It is worthwhile pausing to note that according to the theory of general relativity, nothing can be in time without being in space.) Indeed, suppose that they are located where appropriately biologically functioning brains are but that they are nevertheless neither identical with brains nor composed of material particles, being entirely devoid of matter and lacking physical properties such as mass or charge. The spatiotemporal coincidence of minds and brains would be no violation of the principle that two physical objects cannot occupy exactly the same place at exactly the same time since, by hypothesis, minds are not physical objects. They are entirely disembodied even though they are spatiotemporally coincident with appropriately functioning brains. They are a kind of fundamental energy field coincident with such brains. On this conception might minds causally contribute to the animation of their coincident brains and the brains in turn causally influence them?

This sort of view was a subject of debate in the late nineteenth and early twentieth century, as were debates in biology concerning whether there are wholly immaterial entelechies that are spatiotemporally coincident with organisms and which generate a vital force that causally contributes to guiding the development of organisms and sustaining their integrity (McLaughlin 2003). This view of mentality offers no conception of the nature of minds beyond the negative one that they lack any physical properties save spatiotemporal location and the positive one that they are the seat of mental capacities and abilities, the bearers of mental properties, and what undergo mental change. No hint is offered as to how they could be the seat of mental capacities or abilities—of how such abilities and capacities could be exercised within them. No hint is offered as to what their operations might be, as they are entirely devoid of material constituents. Such matters must be taken as primitive; such how-questions are unanswerable.

Many philosophers have argued that to have a mind is not to bear a relation to an object (physical or otherwise) that is the mind but, rather, to have certain capacities and abilities, such as the capacity to think and to feel and the ability to will. We ourselves have these capacities and abilities. We ourselves are the bearers of mental properties, undergo mental events, and engage in mental activity. Moreover, we are embodied. It does not follow that we are identical with our bodies or some part of them such as our brains. A clay statue may fail to be identical with the lump of clay with which it is spatially coincident. They may fail to be identical because they have different temporal properties (perhaps the lump existed before being shaped into a statue) and because they have different modal properties (the lump can survive being squashed while the statue cannot). Rather, the lump may materially constitute the statue (Pea 1997). On a four-dimensionalist conception of objects, however, the lump

and statue are space-time worms that have spatiotemporal segments that are identical (Sider 2001). Perhaps we are materially constituted by our bodies (or brains) but fail to be identical with them since they, unlike us, lack mental properties. They may also have different temporal properties from us. If we could exist in a disembodied form after the death and disintegration of our bodies and their organs, then, of course, we are not identical with our bodies or brains.

But it is also true that we are not identical with our bodies or brains if they can continue to exist after we have ceased to exist. We may cease to exist at brain death; but at brain death, the brain still exists. Albert Einstein's brain was removed from his skull shortly after his death with the hope that it would yield insight into his prodigious intelligence. But if he ceased to exist upon the death of his brain, then he was not his brain; and it was not he who was removed from the skull of his corpse. Einstein with his famous equation $E=mc^2$, taught us that mass and energy are interconvertible. (Some contemporary New Age Spiritualists would tell us that Einstein's unique energy was released from the matter of his brain upon the expiration of his body, and so that he continues on decoupled from any body. Why any energy released would be Einstein is left entirely obscure, however; and the question of how his mentality was linked to his brain while it was carrying out its normal biological functions remains unanswered. Suffice it to note here that the study of matter-energy in space-time is the subject of physics. We will return to physics shortly.)

Our biologically functioning brains serve somehow as the basis of our capacities to think and to feel and of our volition. Another topic of debate in the late nineteenth and early twentieth centuries was whether, despite the nonexistence of any immaterial object that is the mind, the brain serves as only the causal basis of mental capacities and abilities. On this conception, when we exercise our mental capacities and abilities, mental events (and states) occur within our brains. But they are not identical with occurrences of any kinds of nonmental brain events such as physiological ones; and, indeed, mental events are linked to brain occurrences of other kinds only spatiotemporally and causally: They may accompany them and be causes or effects of them. Since, on this conception, mental events occur within the brain, it might be claimed that they thereby count as physical since the brain is a physical object. But that seems merely a verbal issue. The dualist will claim the important point is that types of mental events are not identical with any other types of brains events and that the only (relevant) relations that token mental events bear to tokens of other kinds of brain events are spatiotemporal and causal. The chief concern raised about this view was whether mental events exert any causal influence on other brain events.

Ewald Herring, in his 1870 lecture at the Imperial Academy of Sciences in Vienna, declared that physiologists should make "the unbroken causative continuity of all material processes an axiom of [their] system of investigation" (Butler 1910, pp. 64–65). He took this position on the grounds that, on the evidence, there seem to be no *gaps* in the physiological processes in the brain to be filled by mental events. The relationship between mental and physiological events, he maintained, should be left as a question for philosophy; brain physiologists can safely bracket it. The fact that there seem to be no gaps in physiological causal chains for mental events to occupy led Thomas Huxley (1874) to maintain that we (and other animals) are *conscious automata*: conscious events accompany certain physiological brain events as dual effects of other physiological events but are themselves causally inert. Trained as a medical doctor, William James (1890) appropriated the term *epiphenomena*, a medical term for symptoms of diseases, for mental phenomena that while caused, lack causal efficacy. James Ward (1903) coined the term *epiphenomenalism* for the view that mental phenomena have no causal effects. The claim that conscious phenomena are epiphenomena is, however, deeply perplexing. If they are, then our belief that we are in pain is never caused by our feeling of pain. And our experience of control over some of our bodily movements cannot give rise to our belief that we are in control of them, for that, too, would require mental causation.

During this period concern was also raised about whether mental causation would violate the law of conservation of energy. (Leibniz had argued earlier that Descartes was committed to minds affecting the motion of material particles in the pineal gland in violation of the conservation laws of momentum and kinetic energy; his mechanics, however, required contact forces, and was eclipsed by Isaac Newton's mechanics, which rejected that requirement [Woolhouse 1985, Papireau 2001].) One response made to the concern about conservation of energy is that causation may very well not require energy transfer; it does not, for instance, on a regularity theory of causation, according to which causation is subsumption under a law of nature, or on a conditional theory of causation, according to which one event causes another if, had the first not occurred, the second would not have occurred either (Broad 1925, ch. III).

Another point made in response was that the conservation of energy principle is silent about the causes of motion, stating only that energy must be conserved within the total system (Broad 1925, ch. III). (Given general relativity, it is mass-energy that is conserved within the total system.) Unlike on Descartes's conception of the mental, on the conception under consideration, mental events occur within the total system of space-time. Indeed, it seems logically possible that certain mental properties are fundamental force-generating properties, just as in classical mechanics the masses of bodies generate the gravitational force, and the electrical charges of bodies generate the electrostatic force.

Perhaps our will involves such a force. There could be a force that is exerted only when matters becomes so configured as to constitute a brain in which certain sorts of mental properties are realized, and that affects the behavior of material particles in ways that causally contribute to bodily behavior that we regard as being under the (partial) control of our volition. Perhaps, further, this *configurational force* is fundamental, affecting the behavior of bodies in ways unanticipated by laws governing matter at lower-levels of complexity. If so, then in the framework of classical mechanics, there would be a mental force law on a par with the inverse square laws—the law of gravity and Coulomb's law.

In the framework of nonrelativistic quantum mechanics, mental energy would contribute to determining the value of the Hamiltonian of Erwin Schröedinger's equation. Since mechanics is a branch of physics, it might be claimed that if mechanics has to take mental properties into account, then the properties would thereby count as physical. But the dualist would regard that as a merely verbal point and note that the important point is that mental properties would be fundamental, irreducible force-generating properties. It should be noted, however, that while such configurational forces could be accommodated within Newtonian mechanics and are compatible with Schröedinger's equation, the role of mental properties would by no means be straightforward on the view in question. By hypothesis, the configurational forces would be exerted only when certain enormously complex microstructural properties were realized by minute physical structures of portions of the brain. On the dualist hypothesis in question, mental properties are distinct from any microstructural properties—at most, accompanying them as a matter of fundamental law. But, then, mechanics would, arguably, have to advert only to the microstructural properties in question, taking them to be the configurational force-generating properties (McLaughlin 1992).

Another view discussed during the period in question is that every mental event is a physiological event but that mental properties are not physiological properties (Lewes 1985, Alexander 1920, Broad 1925). If mental events are physiological events, then they have causal effects. And the mistake made by theorists who found no gaps to be filled by mental events would be that they failed to realize that certain physiological events are mental events in that they fall under mental event types. This view faces the following issue: What is it about a physiological event in virtue of which it falls under a mental event type (or exemplifies a mental property)? Suppose that physiological event P falls under mental event type M and that physiological event P^* does not. It seems, then, that there must be some difference between P and P^* in virtue of which P is and P^* is not an event of type M. The issue is what that difference is. George Henry Lewes (1875) seems to have anticipated a functionalist answer of a kind sometimes given today (See Lewis 1966): He spoke of the role of the physiological event in the organism. But the most widely discussed answer during the period in question was that there are fundamental, irreducible laws of nature linking physiological properties with mental properties (Alexander 1920, Broad 1925).

Thus, the reason P is and P^* is not an instance of M is that P is an instance of a physiological event type that is linked via a fundamental noncausal law of nature to M while P^* falls under no such physiological type. Charles Dunbar Broad (1925–) called this view *emergent materialism*, and he called such laws of nature *transordinal laws*. (Transordinal laws were later denigrated as *nomological danglers* [Feigl 1950].) The guiding idea was that through the course of evolution, complex structures are formed that have genuinely new kinds of properties that are fundamental and thus irreducible. The emergent properties of wholes are linked to properties of their parts and relations among their parts only by fundamental laws. Emergent materialism is thus a kind of dual-aspect theory according to which the mental and physiological aspects of events are linked only by fundamental laws. On this view mental events are causes. But Broad raised the issue of whether they enter into causal relations only in virtue of their physiological properties and so not in virtue of their mental properties (Broad 1925, p. 473). If so, then emergent materialism is committed to a kind of property or type epiphenomenalism (McLaughlin 1989).

In the twentieth century science made truly momentous advances. The atomic theory of matter was vindi-

cated, a quantum mechanical explanation of chemical bonding was provided—dispelling the idea that there are fundamental chemical forces—and organic chemistry and molecular biology made giant strides leading to the demise of any form of vitalism. There seem to be no fundamental mental forces of nature, no mental energy on par with electromagnetic energy, no mental force fields. At least mechanics has as yet no need of such hypotheses. It is now thought that all the fundamental forces are ones that are exerted below the level of the atom: the gravitational force, the electromagnetic force, the weak force, and the strong force. There is some hope for unification, but no role is envisioned for the mental. Of course, current microphysics may well be false; there is at present no quantum theory of gravity. It is, moreover, at least logically possible that our current physics is profoundly mistaken and that the physics in fact true of our world is a kind of Cartesian physics in which mentality plays a fundamental role. But that seems just a fantasy. It is fairly widely assumed that whatever revisions lie ahead for physics, they will not substantially change the dialectic as concerns the mind-body problem.

The mind-body problem is fundamentally the problem of the place of mental phenomena in nature. Contemporary philosophical discussions of the mind-body problem typically proceed under the (often) tacit assumptions that: We are wholly constituted by atoms and more fundamental physical particles, all of which are ingredients of beings entirely devoid of mentality; any fundamental forces at work in us are also at work in many such beings; and that for any (caused) microphysical event *P*, there is a distinct microphysical event *P** that causally determines the objective probability of *P* (if determinism is true, that probability will be 1). The last—which, unlike the others, is often explicitly stated—is sometimes called *the closure of the microphysical* though it goes under other names as well.

Of course, one way of responding to the question of the place of an alleged mental phenomenon in nature is by denying that there actually is any such phenomenon. One can be an *eliminativist* about it. Most contemporary philosophers are eliminativists concerning not only nonspatial, immaterial minds, but also spatiotemporally located immaterial minds: They deny that there are any such things. And they do so for much the same reasons mentioned earlier. Moreover, most contemporary philosophers deny that there are sense data, essentially private mental objects of which only the subject can be aware. Nevertheless, most hold that there are mental properties, capacities, abilities, states, events, and processes. And discussion mainly focuses on their place in nature.

There are many unresolved questions. One central issue concerns the manner in which biologically functioning brains serve as a basis for our capacities to think and to feel and our ability to will: Are they merely a causal basis, or are they rather a constitutive basis? Other issues include whether freedom of the will is compatible with the manner in which they are, such a basis, and with the closure of the microphysical; whether there could be other kinds of material bases for mental capacities and abilities (e.g., silicon-based brains); and what the conditions for personal identity are given the fact of our material embodiment. And there are, as well, theological questions such as whether immortality may somehow be possible despite the fact of our material embodiment. (Might it be possible through the resurrection of the body?)

Among our mental capacities is the capacity to reflect on our own mental lives. Indeed, it is because we have such a capacity that we are able to formulate the mind-body problem. We are not only conscious (as are most kinds of animals), but self-conscious as well. The place in nature of our capacities for self-consciousness must be found. Engagement with the mind-body problem, moreover, requires theoretical reasoning. We form beliefs on the basis of others that provide reasons for them. And we engage in practical reasoning when we deliberate about courses of action (e.g., whether to finish reading the present article). Our capacities for theoretical and practical reasoning must also be located in nature.

The exercise of mental capacities and abilities involves mental states and events (including mental acts). The fundamental problem of the place of mental states and events in nature is that, on the one hand, they have or are instantiations of properties that seem sui generis, and on the other hand, they occur in space-time (arguably, within our skulls) and seem to enter into causal relations with other states and events, including microphysical ones (as, for example, when we deliberately move our bodies across the room with the result that physical particles in our bodies come to be on the other side of the room). The apparently sui generis properties primarily include those of intentionality and phenomenal consciousness.

Properties of intentionality divide into two broad kinds: modes of representation and representational contents. Beliefs, desires, hopes, and intentions, for example—so called *propositional attitudes*—are representational. They have an intentional (representational) mode—

belief, desire, hope, intention—and they have an intentional (representational) content, a content that is (semantically) satisfied or not, depending on the way the world is. States of phenomenal consciousness have phenomenal characters (*qualia*): It is like something for the subject of such a state to be in the state (Nagel 1974). States of phenomenal consciousness include bodily sensations, sense experiences, acts of mental imagery, felt emotions, and occurrent thoughts. Thus, for instance, it is like something for a subject to feel pain, or to visually experience red, or to visualize a sunset. Emotions such as fearing that *P* and being joyous that *P* have contents, and their characteristic manifestations in phenomenal conscious— feelings of fear and feelings of joy—have phenomenal characters. An occurrent thought such as thinking to oneself that it will rain tomorrow will have a representational content and a phenomenal character as well (even if not a distinctive, characteristic one). (Suffice it to note that the relationship between intentionality and phenomenal consciousness and whether one is primary are highly controversial issues.)

Many contemporary philosophers of mind are engaged in the project of trying to naturalize either intentional properties or phenomenal characters—that is to say, to locate them in nature conceived as fundamentally microphysical. It has been argued that such naturalization projects are doomed to failure where intentional properties are concerned because such properties are identifiable only by their place in a network of normative, rational relations and are thus irreducible, having *no echo* in the physical sciences (Davidson 1970). But even some philosophers who are optimistic about the prospects of naturalizing intentional properties maintain that the attempt to naturalize phenomenal consciousness may face insuperable difficulties. Huxley mused: “How it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djinn, when Aladdin rubbed his lamp” (Huxley 1986, p.193). Indeed, it has been claimed that “consciousness is what makes the mind-body problem really intractable” (Nagel 1974, p. 435); that it is “the hard nut of the mind-body problem” (McGinn 1989, p. 394); that it is “the hard part of the mind-body problem” (Strawson 1994, p. 93); and that phenomenal character poses “the hard problem” of consciousness (Chalmers 1996, p. xiii).

Some philosophers have maintained that the link between phenomenal characters and physicality is so mysterious that it is reasonable to hypothesize that the particles—the star dust—from which we are composed must have as yet undiscovered *protomental* properties, which, though their mode of combination somehow constitute phenomenal characters (James 1890; Nagel 1979). Physics, however, has as yet found no need of this *panpsychism* hypothesis. Moreover, if the protomental properties are not themselves phenomenal characters and are objective in nature, then the concern arises that their link with phenomenal characters would also be mysterious. In any case, so mysterious has the connection between phenomenal character and physicality seemed that some philosophers have maintained that we are cognitively closed to the sorts of concepts required for understanding the place of phenomenal characters in nature and thus that the matter transcends human understanding (McGinn 1989).

There are a variety of different naturalizing projects, and some are incompatible with others. However, there have been attempts to state a commitment shared by them all. One leading formulation of such a shared commitment is the following global supervenience thesis: Any minimal physical duplicate of our world is a duplicate *simpliciter* of it (Jackson 1998). A physical duplicate of our world (the actual world) is any possible world that is exactly like our world in every microphysical respect, in respect to its world-wide pattern of distribution of microphysical properties and relations, its world-wide pattern of distribution of microphysical objects, its microphysical laws of nature, and so on. A minimal physical duplicate of our world is any physical duplicate of it that contains nothing other than what is metaphysically required to be a physical duplicate of it. Proponents of different naturalizations programs will offer different explanations of why mental phenomena do not yield a counterexample to the supervenience thesis.

Philosophers, however, who maintain that mental properties of certain sorts are emergent properties, fundamental constituents of nature, linked to other properties only by contingent fundamental laws of nature, will deny the supervenience thesis. Since the laws in question (Broad’s *transordinal laws*) are contingent and fundamental, it is possible for them to fail to hold even though all of the actual microphysical laws of our world hold. Such philosophers are committed to there being a possible world that is a minimal physical duplicate of our world yet not a duplicate *simpliciter* of it because the world is devoid of the mental properties in question (or instantiations of them). For example, someone who holds that phenomenal characters are fundamental in nature will claim there is a possible world that is a minimal physical duplicate of our world yet, unlike our world, is devoid of phenomenal consciousness—*a zombie world* (Chalmers 1996)—and

thus not a duplicate *simpliciter* of our world. And, similarly, someone who held that intentional properties are fundamental will be committed to their being a possible world that is a minimal physical duplicate of our world but which fails to be a duplicate *simpliciter* of our world since it is devoid of intentionality.

Any world that is a minimal physical duplicate of our world will be one in which exactly the same microphysical causal transactions occur as do in our world. If either normative intentional properties or phenomenal characters yield counterexamples to the supervenience thesis, then such properties make no difference to what microphysical causal transactions occur in our world. And they could make a difference to whether certain causal transactions occur in our world only if those transactions fail to be implemented by microphysical ones. Such, it seems, are the facts of our world.

Whether intentionality and phenomenal consciousness can be naturalized—whether they can be located in nature conceived of as fundamentally microphysical—are the fundamental issues of the contemporary mind-body problem. These are issues of intensive, ongoing debate.

See also Augustine, St.; Broad, Charlie Dunbar; Cartesianism; Computationalism; Descartes, René; Dualism in the Philosophy of Mind; Einstein, Albert; Elisabeth, Princess of Bohemia; Functionalism; Hobbes, Thomas; Huxley, Thomas Henry; James, William; Leibniz, Gottfried Wilhelm; Leucippus and Democritus; Lucretius; Malebranche, Nicholus; Mental Causation; Newton, Isaac; Plato; Reductionism in Philosophy of Mind; Self-knowledge; Socrates; Spinoza, Benedict (Baruch) de; Supervenience.

Bibliography

Alexander, S. *Space, Time, and Deity, the Gifford Lectures at Glasgow, 1916–1918*. 2 vols. London: Macmillan, 1920.

Broad, C. D. *The Mind and Its Place in Nature*. New York: Harcourt Brace, 1925.

Butler, S. *Unconscious Memory*. London: A. C. Fifield, 1910.

Chalmers, D. J. *The Conscious Mind: In Search of Fundamental Theory*. New York: Oxford University Press, 1996.

Cottingham, J. "Descartes, Rene." In *Cambridge Dictionary of Philosophy*. Edited by R. Audi, 223–227. Cambridge, U.K.: Cambridge University Press, 1999.

Crane, T. *Elements of Mind*. Oxford: Oxford University Press, 2001.

Descartes, R. *The Philosophical Writings of Rene Descartes*, Vols. 1 and 2. Edited and translated by J. Cottingham, R. Stootoff, and D. Murdoch. Cambridge, U.K.: Cambridge University Press, 1985.

Davidson, D. "Mental Events." In *Experience and Theory*. Edited by L. Foster and J. W. Swanson. Amherst: University of Massachusetts Press, 1970.

Feigl, H. "The 'Mental' and the 'Physical.'" In *Concepts, Theories, and the Mind-Body Problem: Minnesota Studies in the Philosophy of Science Volume II*. Edited by H. Feigl, M. Scriven, and G. Maxwell, 370–497. Minneapolis: University of Minnesota Press, 1958.

Garrett, D. "Spinoza, Baruch." In *Cambridge Dictionary of Philosophy*. Edited by R. Audi, 870–874. Cambridge, U.K.: Cambridge University Press, 1999.

Gert, B. "Hobbes, Thomas." In *Cambridge Dictionary of Philosophy*. Edited by R. Audi, 386–390. Cambridge, U.K.: Cambridge University Press, 1999.

Haldane, J. "Medieval and Renaissance Philosophy of Mind." In *A Companion to Philosophy of Mind*. Edited by S. Guttenplan, 333–338. Oxford: Blackwell, 1994.

Huxely, T. H. "On the Hypothesis that Animals are Automata, and its History." *Fortnightly Review* 16 (1874): 555–580. Reprinted in T. H. Huxley, *Collected Essays, Vol. 1, Methods and Results*. 4th ed. London: Macmillan, 1904.

Huxley, T. H. *Lessons in Elementary Physics*. London: Macmillan, 1986.

Jackson, F. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.

James, W. *The Principles of Psychology*. Cambridge, MA: Harvard University Press, 1890.

Kim, J. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.

Kim, J. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press, 2005.

Kenny, A., ed. and trans. *Descartes: Philosophical Letters*. Oxford: Oxford University Press, 1970.

Lewes, G. H. *Problems of Life and Mind*. Vol. 2. London: Kegan Paul, Trench, Turbner, 1875.

Lewis, D. "An Argument for the Identity Theory." *Journal of Philosophy* 63 (1966): 17–25.

McCann, E. "Philosophy of Mind in the Seventeenth Century." In *A Companion to Philosophy of Mind*, edited by S. Guttenplan. Oxford: Blackwell, 1994.

McGinn, C. "Can We Solve the Mind-Body Problem?" *Mind* 98 (1989): 349–466.

McGinn, C. *The Character of Mind*. Oxford: Oxford University Press, 1982.

McLaughlin, B. P. "The Philosophy of Mind." In *Cambridge Dictionary of Philosophy*, edited by R. Audi, 684–694. Cambridge, U.K.: Cambridge University Press, 1999.

McLaughlin, B. P. "The Rise and Fall of British Emergentism." In *Emergence or Reduction*?, edited by A. Beckermann, A, H. Flohr, and J. Kim, 49–93. Berlin: Walter de Gruyter, 1992.

McLaughlin, B. P. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989): 209–235.

McLaughlin, B. P. "Vitalism and Emergence." In *Cambridge History of Philosophy 1870–1945*, edited by T. Baldwin, 631–639. Cambridge, U.K.: Cambridge University Press, 2003.

Nadler, S. "Malebranche, Nicolas." In *Cambridge Dictionary of Philosophy*, edited by R. Audi, 531–532. Cambridge, U.K.: Cambridge University Press, 1999.

Nagel, T. "Panpsychism." In *Mortal Questions*. Cambridge, U.K.: Cambridge, 1979.

Nagel, T. "What Is it like to be a Bat?" *Philosophical Review* LXXXIII (4) (1974): 435–450.

Papineau, D. "The Rise of Physicalism." In *Physicalism and Its Discontents*. Edited by C. Gillett and B. Loewer, 3–36. Cambridge, U.K.: Cambridge University Press, 2001.

Plato. *Phaedo*.

Rea, M., ed. *Material Constitution*. Latham, MD: Rowman & Littlefield, 1997.

Sider, T. *Four Dimensionalism*. Oxford: Oxford University Press, 2001.

Sleigh, R. C. "Leibniz, Gottfried Wilhelm." In *Cambridge Dictionary of Philosophy*, edited by R. Audi, 491–494. Cambridge, U.K.: Cambridge University Press, 1999.

Strawson, G. *Mental Reality*. Cambridge, MA: MIT Press, 1994.

Van Cleve, J. "Emergence vs. Panpsychism: Magic or Mind Dust?" *Philosophical Perspectives*. Vol. 4, edited by in J. E. Tomberlin, 215–226. Atascadero, CA: Ridgeview, 1990.

Ward, S. L. "The Conscious Automaton Theory." Lecture XII. *Naturalism or Agnosticism*, Vol. 234–264. London: Adam and Charles Black, 1896–1898.

Woolhouse, R. "Leibniz's Reaction to Cartesian Interaction." *Proceedings of the Aristotelian Society* 86 (1985): 69–82.

Brian P. McLaughlin (2005)

MIRACLES

The term "miracle," like the word *nice,* is often used to refer primarily to the responses of the user. In this usage, a miracle is merely some event that astounds the speaker, with perhaps some presumption that others will or should react to it in the same way; just as in the parallel case *nice* means simply "agreeable to me," with perhaps again some suggestion that all right-minded people will feel the same. But the senses of "miracle" that are of philosophical and methodological interest are stronger and less subjectively oriented. Although they include the idea that wonder is called for as at least part of the appropriate response, the crux as well as the ground for the wonder is that a miracle should consist in an overriding of the order of nature. A miracle is something that would never have happened had nature, as it were, been left to its own devices.

This idea of overriding is essential; however, it is certainly subject to various variations and additions. Some writers, for instance, insist that the word *miracle* should be used in such a way that it becomes necessarily true that a miracle can be worked only by God or by his specially deputed agents. Others even build into their very definition of *miracle* some reference to the purposes for which Authority is supposed to be prepared to consider making such an exception. Certainly, most theist theologians are also at great pains to maintain that a miraculous event could not properly be considered a violation, since it would not really represent any infringement, of the fundamental hierarchical order. "It is not against the principle of craftsmanship (*contra rationem artificii*) if a craftsman effects a change in his product, even after he has given it its first form" (Thomas Aquinas, *Summa contra Gentiles,* III, 100). But these very labors to show that and how such "violations" need involve no ultimate irregularity still admit and presuppose the essentially overriding character of the miraculous. There would be no point in trying to show in this way that a miracle must ultimately be no violation of regularity unless it were taken for granted that it apparently is such a violation.

This point is fundamental, and it needs to be stressed more heavily today than in the past. For in addition to the traditional theist reluctance to ascribe to the Deity anything savoring of unseemly irregularity, it is nowadays usual to encounter a certain shyness about any apparent repudiation of scientifically accepted modes of explanation. Thomas Aquinas, earlier in the chapter referred to above, gave a perfectly clear and unequivocal definition of *miracle* that makes no bones at all about the crux of the matter, namely, that "those things are properly called miracles which are done by divine agency beyond the order commonly observed in nature (*praeter ordinem communiter observatum in rebus*)." Again, in the twentieth century, Dr. Eric Mascall, remaining in the same forthright tradition, insisted in his article in *Chambers' Encyclopaedia* that the word *miracle* "signifies in Christian theology a striking interposition of divine power by which the operations of the ordinary course of nature are overruled, suspended, or modified."

MIRACLES AND NATURAL ORDER

To seize the fundamental point that a miracle is an event that violates the "ordinary course of nature" is to appreciate that the notion of a miracle is logically parasitical on the idea of an order to which such an event must constitute some sort of exception. This being so, a strong notion of the truly miraculous—a notion involving something more than the notions of the merely marvelous, the significant, or the surprising—can only be generated if there is first an equally strong conception of a natural order. The inevitable tension between the ideas of rule and of exception thus gives concepts of the miraculous an inherent instability. It is perhaps relevant to notice how this

tension has been felt in the history of ideas. Where there is as yet no strong conception of a natural order, there is little room for the idea of a genuinely miraculous event as distinct from the phenomenon of a prodigy, of a wonder, or of a divine sign. But once such a conception of a natural order has taken really firm root, there is a great reluctance to allow that miracles have in fact occurred or even to admit as legitimate a concept of the miraculous.

An interesting early case of this is provided by Benedict de Spinoza in his *Tractatus Theologico-Politicus,* in which he tried to reconcile his vision of a natural order (*Deus sive natura*) with an acceptance of the Bible as in some sense a privileged document. He did this partly by admitting the limitations of observatory powers of the men of biblical days, but mainly by urging that conventional interpreters of the Bible read far more miracles into it than it contains, because they constantly read poetic Hebrew idioms literally. Today, more and more theologians seem to be noticing the exact words used by the New Testament writers in describing the sorts of alleged events that, in more scientific ages, have been characterized (and perhaps dismissed) as miraculous. These words are τερατὰ ("wonders," or "prodigies"), δυναμεῖς ("powers"), σημεία ("signs"); and, particularly in St. Paul, χαρισματὰ ἰαμάτων ("graces of healing") and ἐνεργήματα δυνάμεων ("effects of powers"). None of these words seems to carry any entailments about the overriding of a natural order. On the other hand, once a really strong conception of natural order has arisen, its adherents tend to dismiss out of hand all stories of putative occurrences in the belief that if they allowed that these occurrences had taken place at all, they would have to admit them to have been miraculous. One may refer here to R. M. Grant's recent *Miracle and Natural Law in Graeco-Roman and Early Christian Thought* (Amsterdam, 1952) and to William E. H. Lecky's classic study *History of the Rise and Influence of Rationalism in Europe* (London, 1890). The former summarizes its own thesis as follows: "Credulity in antiquity varied inversely with the health of science and directly with the vigor of religion" (p. 41). This, however, was later qualified by the important observation that "at least in some respects Christians were far less credulous than their contemporaries, at least in the period before Augustine" (p. 120). Lecky traced a development in which stories of the ostensibly miraculous, from being accepted as a chief guarantee of the authenticity of the Christian revelation, become instead "a scandal, a stumbling block, and a difficulty" (Vol. I, p. 143). In the nineteenth century the radical biblical critic David Strauss announced in the introduction to his *Das Leben Jesu* (2 vols., Tübingen, 1835; translated by Mary Ann Evans as *Life of Jesus Critically Examined,* London, 1848), "We may summarily reject all miracles, prophecies, narratives of angels and demons, and the like, as simply impossible and irreconcilable with the known and universal laws which govern the course of events." And in the twentieth century there was even a bishop of the Church of England capable of saying of the author of Mark, "He was credulous inasmuch as the miracles, as they are narrated, cannot, in the light of our modern knowledge of the uniformity of nature, be accepted as historical facts" (F. W. Barnes, *The Rise of Christianity,* London and New York, 1947, p. 108).

DILEMMA OF HOLDING STRONG RULES WHILE ADMITTING EXCEPTIONS. The spokesman for the occurrence of the miraculous faces a dilemma that arises from the very essence of the concept he espouses. It is tempting, but wrong, for the believer in the miraculous to think that he can afford to gloat over any little local difficulties and embarrassments that may from time to time beset the forward march of science. But insofar as a miracle involves an alleged overriding of a law of nature, he too is committed to showing the subsistence of a natural order. Exceptions are logically dependent upon rules. Only insofar as it can be shown that there is an order does it begin to be possible to show that the order is occasionally overridden. The difficulty (perhaps an insoluble one) is to maintain simultaneously both the strong rules and the genuine exceptions to them. The oscillations in the history of thought are to be understood by reference to this tension (amounting perhaps to a contradiction) that is inherent in the concept of the miraculous, and it is on this same tension that the various logical and methodological problems also center.

LOGICAL AND METHODOLOGICAL PROBLEMS

It is with logical and methodological problems that we are primarily concerned. The classical, and by far the best, approach is by way of the notorious section X, "Of Miracles," in David Hume's *Enquiry concerning Human Understanding* (1748). This and Section XI of this *Enquiry,* both of which were parts of a single coordinated case, constitute Hume's answer to what was, in his day, the stock program of Christian apologetic. This program had two stages: the first was an attempt to establish the existence and certain minimal characteristics of God by appealing only to natural reason and experience, the second was an attempt to supplement this rather sketchy religion of nature with a more abundant revelation. This program,

in its characteristically eighteenth-century form, received its archetypal fulfillment in Archdeacon William Paley's *Natural Theology* (London, 1802) and also in his *Evidences of Christianity* (London, 1794). In the eighteenth-century form, the weight of the first part of the case was borne primarily by the Argument to Design. If from a watch we may infer a watchmaker, then the orderliness of the universe entitles us to infer, by parity of reasoning, a Maker of the universe. The second part of the case rested on the claim that there is ample historical evidence to show that the biblical miracles, including the crucial physical resurrection of Jesus bar Joseph, did in fact occur, and that this in turn proved the authenticity of the Christian revelation.

Paley's style of systematic rational apologetic has no doubt gone out of fashion, at least among Protestants. But Hume's challenges to the whole idea of a substantial natural theology and to the project of establishing the authenticity of any alleged revelation by proving that its claims have been supported by miracles are not, and are not likely to become, dead issues. For in 1870 the third session of the First Vatican Council defined as constitutive dogmas of the Roman Catholic religion both of the positions that Hume had challenged. The relevant passage of the canon dealing with the second reads, "If anyone shall say … that miracles can never be known for certain, or that the divine origin of the Christian religion cannot properly be proved by them: let him be cast out" (*si quis dixerit … aut miracula certo cognosci numquam posse nec iis divinam religionis christianae originem rite probari: anathema sit*; H. Denzinger, ed., *Enchiridion Symbolorum,* 29th ed., Sec. 1813, Freiburg im Breisgau, 1953).

PROBLEM OF SUPERNATURAL REVELATION. Hume's main contention was thus, in his own words, that "a miracle can never be proved so as to be the foundation of a system of religion." For him, all other questions about the miraculous were, officially at least, merely incidental to this basic tenet. He defined a "miracle" as "a transgression of a law of nature by a particular volition of the Deity, or by the interposition of some invisible agent." This definition has been attacked on various counts, but the criticism is misconceived, for two reasons. First, this was in fact the way in which the opponents whom Hume had in mind defined the term "miracle." Thus, Dr. Samuel Clarke, in his famous Boyle lectures (*The Works of Samuel Clarke,* Vol. II, London, 1738, p. 701), had defined "miracle" as "a work effected in a manner … different from the common and regular method of providence, by the interposition either of God himself, or of some intelligent agent superior to men." Second, if, as Clarke and the orthodox tradition would have it, the occurrence of a miracle is to serve "for the proof or evidence of some particular doctrine, or in attestation of the authority of some particular person," then surely a miracle must be conceived in this way. It is only and precisely insofar as it is something really transcendent—something, so to speak, that nature by herself could not contrive—that such an occurrence could force us to conclude that some supernatural power is being revealed.

In this context it would be worse than useless to appeal to revelation for criteria by which genuinely miraculous events may be identified, and thus distinguished from the unusual, the untoward, or the merely ordinary. For if the occurrence of a miracle is to serve as the endorsement of a revelation, then we have to find some means entirely independent of that revelation by which the endorsement itself may be recognized. Exactly the same point applies, of course, if, with what is now a rather fashionable school of apologetic, it is urged that miracles are not essentially overridings, but signs. If a sign is to signify to the unbeliever, then there must be some means independent of the doctrinal system itself by which the signs may be identified and read. As has been suggested already, there is much to be said for trying to interpret the records of τερατὰ and σημεία in the New Testament in terms of some notion of sign, rather than as miracle stories proper. But it is necessary to insist on two facts that seem to be often overlooked—namely, that part of the price that must be paid for this method of interpretation is the sacrifice of the use of these stories as independent evidence of the genuinely revelatory character of the doctrines; and that such a sacrifice presumably entails the rejection of at least one defined dogma of the Roman Catholic Church, and hence of the truth of Roman Catholicism as a theological system.

A similar but different point applies if a relativistic definition of "miracle" is adopted, as was done, for instance, by John Locke. In his *Discourse of Miracles* (written 1702, published posthumously), he defined the word *miracle* as "a sensible operation, which, being above the comprehension of the spectator, and in his opinion contrary to the established course of nature, is taken by him to be divine." It was also done, in a slightly different way, by St. Augustine, who insisted that "nature is the will of God" (*Dei voluntas rerum natura est*), and hence that "a portent is not contrary to nature, but contrary to our knowledge of nature" (*Portentum ergo fit non contra naturam, sed contra quam est nota natura; De Civitate Dei,* XXI, 8). To operate with a relativistic notion of this sort is

necessarily to be deprived of the possibility of arguing that a miracle is a miracle regardless of whatever anyone may happen to know or to believe about it, and hence to rob the attempt to base an apologetic on the occurrence of miracles of whatever initial plausibility it might otherwise possess. For the occurrence of events that are merely inexplicable *to us,* and *at present,* provides no good ground at all for believing that doctrines associated with these occurrences embody an authentic revelation of the transcendent. There is, of course, no particular reason why Locke himself should have been disturbed about this. The case of Augustine, however, is more interesting, for he is a recognized saint and one of the four great doctors of the church. And yet insofar as he held to a relativistic notion of a miracle, he was safeguarding the vital doctrine of the total dependence of the whole creation—but at the price of subverting a sort of apologetic which it has since become essential for Roman Catholics to believe in as a possibility.

PROBLEM OF IDENTIFYING AN EVENT AS MIRACULOUS. Up to this point it has been insisted that if the occurrence of a miracle is to serve—as Clarke and the orthodox tradition would have it—"for the proof or evidence of some particular doctrine, or in attestation of the authority of some particular person," then in a traditional sense, miracles must be conceived of as involving the overriding of some natural order that is at least partly autonomous. The importance of this crucial point is often overlooked. Another immediately consequential point, however, is overlooked perhaps even more often, namely, that if an occurrence that is miraculous in the traditional sense is to serve as evidence for anything, it must be possible to identify it as being miraculous. Furthermore, as was urged above, if its occurrence is to serve as an endorsement of some doctrinal system, the method of identification must be logically independent of that system. The difficulty of meeting this last requirement is often concealed by the acceptance of what seems, for many people, to be an almost unquestionable assumption. Protagonists of the supernatural, and opponents too, take it for granted that we all possess some natural (as opposed to revealed) way of knowing that and where the unassisted potentialities of nature (as opposed to a postulated supernature) are more restricted than the potentialities that, in fact, we find to be realized or realizable in the universe around us.

This is a very old and apparently very easy and tempting assumption. It can be found, for instance, in Cicero's *De Natura Deorum,* and hence presumably much earlier, in Cicero's Greek sources. Nevertheless, the assumption is entirely unwarranted. We simply do not have, and could not have, any natural (as opposed to revealed) criterion that enables us to say, when faced with something that is found to have actually happened, that here we have an achievement that nature, left to her own unaided devices, could never encompass. The natural scientist, confronted with some occurrence inconsistent with a proposition previously believed to express a law of nature, can find in this disturbing inconsistency no ground whatever for proclaiming that the particular law of nature has been supernaturally overridden. On the contrary, the new discovery is simply a reason for his conceding that he had previously been wrong in thinking that the proposition, thus confuted, did indeed express a true law; it is also a reason for his resolving to search again for the law that really does obtain. We certainly cannot say, on any natural (as opposed to revealed) grounds, that anything that actually happens is beyond the powers of unaided nature, any more than we can say that anything that any man has ever succeeded in doing transcends all merely human powers. For our evidence about the powers of nature in general, and of men in particular, is precisely and only everything that things and people do. For a scientist to insist that some recalcitrant fact constitutes an overriding of a still inviolably true law of nature is—to borrow Rudolf Carnap's mischievous analogy—as if a geographer were to maintain that the discrepancies between his maps and their objects show that there is something wrong with the territories concerned.

The insistence of the scientist, insofar as he is simply a scientist, on always seeking strictly universal laws is itself rooted in the fundamental object of the whole scientific quest: if scientists are to find comprehensive explanations, they must discover universal laws. A scientist's refusal to accept the idea that in any single case nature has been overridden by supernatural intervention is grounded partly on precisely the above-mentioned lack of any natural (as opposed to revealed) criterion for distinguishing natural from supernatural events, and partly on his commitment—which is chiefly what makes him a scientist—to continue always in the search for completely universal laws, and for more and more comprehensive theories. In view of this, it need be neither arbitrary nor irrational to insist on a definition of a "law of nature" such that the idea of a miracle as an exception to a law of nature is ruled out as self-contradictory.

The seductive but erroneous idea that we do possess some natural means for the identification of the supernatural is one that, in some respects, parallels the notion that it is logically possible to derive prescriptive norms

from knowledge of what is, in some purely descriptive sense, natural. In each case there are adherents for whom the division between natural and supernatural, or between natural and unnatural, is nothing but an incoherent muddle. Likewise, in each case there are others who, in support of their choice, are prepared to deploy some more or less elaborate structure of theoretical justification.

PROBLEM OF EVIDENCE. All of this argumentation, although both relevant and (in spirit at least) thoroughly Humean, has little in common with the line of argument Hume chose to develop in the section "Of Miracles." Although this line of argument is equally methodological, it treats the question of miracles as it arises in the field of history rather than as it might impinge upon natural science. Hume was primarily concerned not with the question of fact but with that of evidence. The problem was how the occurrence of a miracle could be proved, rather than whether any such events ever had occurred. Consequently, even if Hume was successful, the way would still remain clear for people to believe in miracles simply on faith. In his own mordant way, Hume himself was happy to allow for this, but he always insisted that "a wise man proportions his belief to the evidence."

This concentration on the evidential issue means that Hume's thesis, however offensively expressed, is nevertheless at bottom defensive. Hume hoped that he had discovered "a decisive argument … which must at least silence the most arrogant bigotry and superstition, and free us from their impertinent solicitations … an argument which … will … with the wise and learned, be an everlasting check to all kinds of superstitious delusion…." These words were very carefully chosen. The whole argument was directed to the wise—to those, that is, who insist on proportioning their belief to the evidence. It did not show that the substantive claims of the bigoted and superstitious are in fact false. It was intended to serve as a decisive check on any attempt to solicit the assent of rational men by producing proof of the occurrence of the miraculous. In particular, the object was to interdict the second movement of the standard apologetic attack as outlined above.

If for present purposes a certain amount of misguided psychologizing is ignored, the following would appear to be the gist of Hume's "everlasting check." There is, he remarked, "no species of reasoning more common, more useful, and even necessary to human life than that derived from the testimony of men and the reports of eye-witnesses and spectators." Yet all testimony must ultimately be subject to assessment by the supreme court of experience. Certainly there are, as Hume observed, "a number of circumstances to be taken into consideration in all judgments of this kind." Yet "the ultimate standard by which we determine all disputes … is always derived from experience and observation." (Of all people, Hume, as the author of that most famous paragraph in the *Treatise of Human Nature,* should have said not "*is,*" but "*ought*" always to be so derived.)

The weight of the testimony required must depend on the apparent credibility of the events reported. If the events are in some way marvelous and rare, then the testimony for them has to be treated with more circumspection than the witness to everyday occurrences. But supposing that the testimony is for events that, had they occurred, would have been genuinely miraculous: we are then confronted with a paradoxical dilemma, proof balanced against proof. However overwhelming the testimony might have appeared were it not being considered as evidence for a miracle, in this peculiar case the testimony must always be offset against a counterproof. In Hume's own words, "A miracle is a violation of the laws of nature; and as a firm and unalterable experience has established these laws, the proof against a miracle, from the very nature of the fact, is as entire as any argument from experience can possibly be imagined."

In the first part of section X, Hume argued generally from the concept of the miraculous—from, as he put it, "the very nature of the fact." In the second he deployed several more particular assertions about the corruptions to which testimony is liable, urging that such corruptions are exceptionally virulent where any religious issue is involved. He also added a further consideration relevant to any attempt "to prove a miracle and make it a just foundation for any … system of religion."

This consideration was expressed badly and was entangled in one or two inessential errors and confusions. But a letter makes clear Hume's intent. The point is that if the occurrence of some sort of miracle is to serve as a guarantee of the truth of a system of religion, then there must not have been any similar miracle under the auspices of a rival system, the truth of which would be incompatible with the truth of the first. Consequently, insofar as we are considering a miracle not as a putative bald fact but as a possible endorsement of the authenticity of a revelation, we have to throw into the balance against the testimony for the miracles of any one candidate revelation all the available testimony for all the miracle stories presented by all the rival systems that are inconsistent with the first. In its appeal to a necessary

conflict of evidence, this argument resembles the paradoxical dilemma expounded above.

MIRACLES AND THE PHILOSOPHY OF HISTORY

Enough already has been said to suggest that there is more to Hume's check than a trite insistence that since the occurrence of a miracle must be very improbable, it would have to be exceptionally well evidenced in order to be believed. C. S. Peirce was in possession of the vital clue (which he seems never to have exploited fully) when he remarked, "The whole of modern 'higher criticism' of ancient history in general, and of Biblical history in particular, is based upon the same logic that is used by Hume" (*Values in a Universe of Chance,* edited by P. P. Wiener, New York, 1958, pp. 292–293). When we follow this clue, it becomes obvious that Hume himself saw "the accounts of miracles and prodigies to be found in all history, sacred and profane" as presenting a methodological problem. This section on miracles constitutes the outer ring of Hume's defenses against the orthodox religious apologetic. But at the same time it is also part of his contribution to an understanding of the presuppositions and the limitations of critical history.

This fact seems not to have been appreciated as it should have been. There is, for instance, no reference to Hume's section "Of Miracles" in R. G. Collingwood's *The Idea of History* (Oxford, 1946); and neither Collingwood nor F. H. Bradley seems to have had any idea of the extent to which Bradley's own essay, "The Presuppositions of Critical History" (*Collected Papers,* Vol. I., Oxford, 1935), echoed arguments first developed by Hume. It is worthwhile to consider possible causes of this neglect. In part it is to be attributed to the insistence (at one time universal) on treating section X, "Of Miracles," as though it were a separate and disingenuous essay, irrelevantly inserted into the first *Enquiry* simply to cause scandal and thereby push up sales. This perverse and gratuitously offensive notion has misled interpreters to overlook some extremely relevant remarks in Part I of section VIII which concern the inescapably uniformitarian presuppositions of both the natural and the social sciences. Even those who have succeeded in appreciating section X as a very considerable piece of argumentation have been inclined to pigeonhole it as being a contribution to the philosophy of religion only. Certainly Hume's argument does, in the first instance, belong to the philosophy of religion; and this, of course, is how Hume presented it. Yet, as we have already seen, it also has a place in the philosophy of science. The fact that Hume appreciated this is perhaps suggested by his proposal that if, against all reasonable expectation, there were to be sufficient historical evidence to establish that the "miracle" of a universal eight-day eclipse had occurred in January 1600, "then our present philosophers [scientists], instead of doubting the fact, ought to receive it as certain; and ought to search for the causes whence it might be derived." It is surely significant that in this one context, and inconsistently with his own official definition of *miracle,* he spoke not of "a violation of the laws of nature," but rather, and more weakly, of "violations of the usual course of nature."

The same nodal argument which thus has a place in both the philosophy of religion and the philosophy of science belongs equally in the philosophy of history. For what Hume was contending (with certain lapses and hesitations) is that the criteria by which we must assess historical testimony, and the general presumptions which alone make it possible for us to interpret the detritus of the past as historical evidence, must inevitably rule out any possibility of establishing, upon purely historical grounds, that some genuinely miraculous event has indeed occurred. Hume concentrated on testimonial evidence because his conception of the historian, later illustrated in his own famous *History of England,* was of a judge assessing with judicious impartiality the testimony set before him. But the same Humean principles can be applied more widely to all forms of historical evidence.

The fundamental propositions are first, that the present detritus of the past cannot be interpreted as historical evidence at all, unless we presume that the same basic regularities obtained then as today; and second, that in trying his best to determine what actually happened, the historian must employ as criteria all his present knowledge, or presumed knowledge, of what is probable or improbable, possible or impossible. In his first work, the *Treatise of Human Nature* (II, iii, i), Hume had argued that it is only on such presumptions that we can justify the conclusion that ink marks on old pieces of paper constitute testimonial evidence. Early in the first *Enquiry,* in the first part of section VIII, he urged the inescapable importance of having such criteria. In a footnote to section X, he quoted with approval the reasoning of the famous physician De Sylva in the case of a Mlle. Thibaut: "It was impossible she could have been so ill as was proved by witnesses, because it was impossible she could, in so short a time, have recovered so perfectly as he found her."

FLAWS IN HUME'S ACCOUNT. Two very serious faults in Hume's presentation of his argument may obscure the

force and soundness of De Sylva's reasoning, as well as the fact that this sort of application of canons to evidence is absolutely essential to the very possibility of critical history.

The first fault is a rather wooden dogmatism of disbelief. For against all his own high, skeptical principles, Hume tended to take it for granted that what in his own day he and all his fellow men of sense firmly believed about the order of nature constituted not just humanly fallible opinion, but the incorrigible last word. He was thus betrayed into categorically dismissing as downright impossible certain reported phenomena that the later progress in the study of abnormal psychology and of psychosomatic medicine has since shown to have been perfectly possible. But the moral to be drawn from these lapses into dogmatism is not that Hume was mistaken in insisting that the critical historian must apply canons of possibility and probability to his evidence, but that he failed to appreciate that all such canons are themselves subject to criticism and correction.

The second major fault in Hume's treatment is both more serious and more excusable. He was unable to provide an adequate account of the logical character of a law of nature. Hence, he could not offer any sufficiently persuasive rationale for employing, as canons of exclusion in historical inquiry, propositions that express, or that are believed to express, such natural laws. The way may thus seem to be open for a historian who holds different presuppositions, yet still remains truly a historian, to endorse as veridical stories of events that, had they occurred, would have been truly miraculous. (For a sustained study of such attempts to have it both ways, see T. A. Roberts, *History and Christian Apologetic,* London, 1960.)

This problem of the logical nature of natural laws has, of course, many more aspects than those that immediately concern us here. But it is important first to emphasize that it is at least as much a problem for Hume's immediate opponents as for Hume. For it is his opponents who need a strong sense of "miracle," in which the miraculous can be distinguished from the merely marvelous. It is tempting, but entirely wrong, for the spokesman for the miraculous to think that he can afford to triumph over Hume's difficulties without being himself committed in any way to producing his own account of the character of laws of nature—an account that shall be more satisfactory as an analysis and yet, at the same time, consistent with the things the spokesman himself wants to say about the miraculous. His dilemma, to repeat, is that he needs to be able to accommodate simultaneously both the strong laws and the spectacular transgressions.

NOMOLOGICAL PROPOSITIONS. Casting back to the reasoning of De Sylva, it can now be seen that (and how) it constitutes a paradigm of critical history. For it is only and precisely by presuming that the laws that hold today held in the past and by employing as canons all our knowledge—or presumed knowledge—of what is probable or improbable, possible or impossible, that we can rationally interpret the detritus of the past as evidence and from it construct our account of what actually happened. But in this context, what is impossible is what is physically, as opposed to logically, impossible. And "physical impossibility" is, and surely has to be, defined in terms of inconsistency with a true law of nature. Or rather, since this sense of "impossible" is prior to the development of science proper, it might be said that what is physically impossible is whatever is inconsistent with a true nomological proposition.

Both causal propositions and those expressing laws of nature fall under the genus nomological. Although Hume himself concentrated on the causal species, what he said can easily be extended. In his view, when we say that *A* is the cause of *B,* the main thing we are saying is that *B*'s are constantly conjoined with *A*'s—never as a matter of fact *A* and not *B,* or, in modern terminology, *A* materially implies *B.* Of course, he went on, people think they are asserting not a mere constant conjunction, but some real connection, and in a way this is right. The fact is, according to Hume, that there *is* a connection, but that it is a psychological one: we have formed a habit of associating the idea of an *A* with the idea of a *B.*

Yet this account of causal propositions cannot be adequate. All causal propositions entail subjunctive conditionals. (A subjunctive conditional, appropriately enough, is a proposition of the form, "If it were … it would.") Thus, "*A*'s are the only things which cause *B*'s" entails "If *A* were not to occur (or to have occurred) *B* would not occur (or have occurred)." But no variation on the material implication theme, with or without benefit of associationist psychological speculation, can be made to entail any such subjunctive conditional. Furthermore, the same essential inadequacy afflicts any extension of a Humean analysis to cover nomologicals in general. For a nomological is, by the above definition, a contingent proposition that entails some contingent subjunctive conditional.

The essential difference between the contingent "All *X* are ϕ" and the equally contingent "Any *X* must be ϕ" is

that the former can be expressed as a material implication, "Not both *X* and not ϕ," whereas the latter cannot be so expressed, because it is a nomological, entailing such subjunctive conditionals as "If there were to have been an *X* (which in fact there was not) it would have been a ϕ." The nomological goes far beyond the statement of a mere conjunction of *X* and ϕ as a matter of fact. It asserts also a (contingent) connection between *X* and ϕ. For although the nomological is no more logically necessary than the corresponding material implication, it says not merely that, as it happens, a constant conjunction has been, is being, and will be maintained, but also that it would be and would have been maintained regardless of what anyone did or might have done. To assert the nomological is to assert that the conjunction is one that can be relied upon. It is for this reason that experimental evidence is so essential to our knowledge of nomologicals: the obvious and ultimately the only satisfactory test of the reliability of a law is to subject it to strains. It is for the same reason that a knowledge of nomologicals provides, at least in principle, a guarantee of repeatability. To say that the conjunction of *B*'s with *A*'s is reliable is to say that any time anyone likes to produce an *A* he will thereby bring about a *B*.

THE HISTORIAN'S APPROACH. In the light of the above discussion, we can again consider the question of historical evidence for the miraculous. The critical historian, confronted with some story of a miracle, will usually dismiss it out of hand, asking first only whether it can be used as evidence, not for the occurrence reported, but for something else. To justify his procedure he will have to appeal to precisely the principle Hume advanced: the "absolute impossibility or miraculous nature" of the events attested must, "in the eyes of all reasonable people … alone be regarded as a sufficient refutation." Our sole ground for characterizing the reported occurrence as miraculous is at the same time a sufficient reason for calling it physically impossible. Contrariwise, if ever we became able to say that some account of the ostensibly miraculous was indeed veridical, we can say it only because we now know that the occurrences reported were not miraculous at all.

OBJECTIONS TO THE HISTORIAN'S APPROACH. To this representation of the procedure of the critical historian there are two main objections. First, it will be argued that such an approach to what purports to be historical evidence for the miraculous is irrationally dogmatic, for in this instance the historian seems to be represented as dismissing all evidence that conflicts with his own fundamental prejudices and as defending a closed system in which his professional predilections are guaranteed against falsification by a "Heads-I-win: Tails-you-lose" argument. This is a very understandable objection. It is made more plausible by the regrettable fact that there have been, and still are, many historical writers whose actual procedures correspond rather too closely to this suggested representation. Also it is, of course, true that the dilemmas generated by the tension implicit in the concept of the miraculous must necessarily seem to their victims to have a "Heads-you-win: Tails-I-lose" aspect. Nevertheless, the critical historian is not committed to the sort of bigoted dogmatism the present objection attributes to him.

Nomological laws and reports of miracles. As Hume was insisting from first to last, the possibility of miracles is a matter of evidence and not of dogmatism. For, to proceed beyond Hume, the nomological proposition that provides the historian's canon of exclusion will be open and general and of the form "Any *X* must be ϕ." The proposition reporting the (alleged) occurrence of the miracle will be singular, particular, and in the past tense; it will have the form "This *X* on that particular occasion was not ϕ." Propositions of the first sort can in principle be tested at any time and in any place. Propositions of the second sort cannot any longer be tested directly. It is this that gives propositions of the first sort the vastly greater logical strength that justifies their use as criteria of rejection against the latter. It will indeed be only and precisely insofar as we have evidence sufficient to warrant our assertion of the general nomological that excludes the particular historical proposition that we shall have sufficient reason to claim that the event it reports would have been genuinely miraculous.

The logic of evidence. Suppose that in some particular case the evidence for a miracle appears extremely strong. Then perhaps the historian may ask himself whether the nomological proposition that precludes this event is after all true. It could, in principle at any rate, be further tested. If, as is possible, it were shown to be false after all, then perhaps the event so strongly evidenced did indeed occur. But by the same token, that event could now no longer be described as truly miraculous. This, surely, is what has happened in the case of so many of the reports of astonishing psychosomatic cures, which Hume himself, in his capacity as a historiographer, too rashly dismissed. (Consider, for example, his contemptuous rejection of the stories of faith healings by the Emperor Vespasian and of the many cures associated with the tomb of the Jansenist Abbé Paris, all in section X of his

first *Enquiry.*) Alternatively, the nomological proposition might survive even our further tests. Hume should be the last one to deny that it must remain always conceivable—logically, that is, as opposed to physically possible—that the event in question did in fact occur. Yet in this case, no matter how impressive the testimony might appear, the most favorable verdict that history could ever return must be the agnostic, and appropriately Scottish, "not proven."

Need for canons of evidence. The second objection to the above representation of the procedure of the critical historian suggests that there is something arbitrary or at least optional about the appeal to canons provided by some of our knowledge, or presumed knowledge, of what is probable or improbable, possible or impossible. Once again there is some ground for this objection. Certainly we can choose whether or not we will try to act as critical historians. But once that fundamental choice is made, there is nothing arbitrary and nothing optional about insisting on the employment of these canons. For the essential aim of the historian is to get as near as he can to a full knowledge of what actually happened, and why. To do this he must find and interpret evidence, for belief unsupported by evidence may be true, but it cannot constitute knowledge. Yet to interpret the detritus of the past as evidence, and to assess its value and bearing as such, we must have canons. And for a rational man, these canons can only be derived from the sum of his available knowledge, or presumed knowledge. It is not the insistence on the systematic employment of these always corrigible canons that is arbitrary; what is arbitrary is to pick and choose in the interests of your ideological predilections among the available mass of miracle stories, or to urge that it is (psychologically) impossible that these particular witnesses were lying or misinformed and hence that we must accept the fact that on this occasion the (biologically) impossible occurred. If one once departs in such arbitrary ways from these canons of critical history, then anything and everything goes. (For examples of precisely this sort of arbitrariness, see M. C. Perry, *The Easter Enigma,* London and New York, 1959.)

Possible justifications for belief in miracles. Nothing that has been said in this article decisively closes the door on faith. We have been concerned only with questions about the possibilities of having good reasons for belief in the miraculous. Again, nothing has been said to preclude the production of nonhistorical and nonscientific considerations that might, either by themselves or with the aid of historical or scientific evidence, justify our belief that certain miracles did indeed occur. Perhaps one might develop some defensible system of rational theology that would provide criteria both for identifying particular occurrences as miraculous and for separating the true miracle stories from the false. Hume tried to rule this out also, of course, in section XI of his *Enquiry,* and elsewhere. But it has been no part of our present task to examine arguments against natural theology. Finally, it is perfectly possible to develop a new concept and to apply to it the term "miracle." There is never anything to keep anyone from simply changing the subject.

See also Augustine, St.; Bradley, Francis Herbert; Carnap, Rudolf; Cicero, Marcus Tullius; Clarke, Samuel; Collingwood, Robin George; Eliot, George; Hume, David; Locke, John; Natural Law; Nature, Philosophical Ideas of; Paley, William; Peirce, Charles Sanders; Revelation; Spinoza, Benedict (Baruch) de; Strauss, David Friedrich; Teleological Argument for the Existence of God; Thomas Aquinas, St.

Bibliography

HISTORICAL STUDIES

Grant, R. M. *Miracle and Natural Law in Graeco-Roman and Early Christian Thought.* Amsterdam, 1952.

Lecky, W. E. H. *History of the Rise and Influence of Rationalism in Europe.* London, 1865. Chs. 1–3. These chapters carry on approximately from where Grant leaves off.

Stephen, Leslie. *English Thought in the Eighteenth Century,* 3rd ed. London, 1902. Vol. I. Contains the classic account of the entire controversy of which Hume's "Of Miracles" formed a part.

LOGICAL AND METHODOLOGICAL STUDIES

Augustine. *The City of God,* XXI, 8.

Hume, D. *An Enquiry concerning Human Understanding,* VIII–XI.

King-Farlow, J. "Miracles: Nowell-Smith's Analysis and Tillich's Phenomenology." *International Philosophical Quarterly* 2 (1962). Although this article appeared in a Jesuit journal, it seems to accept most of the sort of methodological criticism presented in the present article. The bibliography of recent Roman Catholic literature on the subject is especially valuable.

Lewis, C. S. *Miracles.* London and New York, 1947. Mozley (see below), Tennant (see below), and Lewis all give Anglican views, of which Tennant's is perhaps the most liberal and Lewis's the most conservative. The latter employs a theoretical structure of the type referred to in the discussion of the problem of identifying the supernatural in order to uphold the possibility of miracles. For a discussion and development of the ideas touched on in that section, see the exchanges between H. H. Dubs, A. Lunn, and P. Nowell-Smith in the *Hibbert Journal* (1950–1952).

Mill, J. S. *A System of Logic,* III, 4 and 25.

Mozley, J. B. *On Miracles.* London, 1865.

Roberts, T. A. *History and Christian Apologetic.* London: S.P.C.K., 1960.

Smith, G. D., ed. *The Teaching of the Catholic Church,* 2nd ed. London, 1952.

Spinoza, B. *Tractatus Theologico-Politicus,* VI.

Tennant, F. R. *Miracle and Its Philosophical Presuppositions.* Cambridge, U.K.: Cambridge University Press, 1925.

Thomas Aquinas. *Summa contra Gentiles,* III, 98–107.

ILLUSTRATIONS OF MAIN METHODOLOGICAL POINTS

Huxley, T. H. "Agnosticism and Christianity." In *Lectures and Essays.* London, 1902. An attack on Newman (see below).

Middleton, Conyers. *A Free Inquiry into the Miraculous Powers which are supposed to have subsisted in the Christian Church from the earliest ages through several successive centuries.* London: R. Manby and H.S. Cox, 1749. A pivotal work in the great eighteenth-century controversy described by Stephen.

Morrison, F. *Who Moved the Stone?* New York: Century, 1930. A gripping, if very unsophisticated, study of New Testament evidence.

Newman, J. H. "Essay on the Miracles Recorded in Ecclesiastical History." In *The Ecclesiastical History of M. L'Abbé Fleury.* Oxford, 1842. This essay by the future cardinal deserves to be read alongside Middleton's *Free Inquiry* (see above).

Perry, M. C. *The Easter Enigma.* London: Faber and Faber, 1959. An attempt to apply the findings of psychical research to the interpretation of the New Testament documents. It illustrates both an admirably undogmatic flexibility about what in fact is and is not possible and a certain inconsistency in applying the canons of critical history.

Schweitzer, Albert. *The Quest for the Historical Jesus, from Reimarus to Wrede.* London: Black, 1910. A classic history of a historiographical subject. It is interesting to note that the work contains no reference to Hume.

Thomson, J. M. *Miracles in the New Testament.* London: Edward Arnold, 1912. An important but neglected work by a historian who later achieved distinction in the field of secular history.

West, D. J. *Eleven Lourdes Miracles.* London: Duckworth, 1957. A study of the best evidence by a psychiatrist and sometime research officer of the Society for Psychical Research.

Antony Flew (1967)

MIRACLES [ADDENDUM]

Consistent with standard eighteenth-century accounts by Christian apologists, English deists, and skeptics like David Hume, a miracle is still usually thought of as an event with religious significance that is in some sense contrary to the laws of nature. There is no consensus about the best definition. Thus, in their investigations concerning the credibility of miracles, philosophers often have recourse to paradigmatic cases of purported miracles such as the resurrection of Jesus from the dead.

Most philosophers endorse a conception of miracles consistent with the possibility of a violation of the laws of nature. J. L. Mackie (1982) calls a miracle "a supernatural intrusion into the normally closed system that works in accordance with the laws of nature." Richard Swinburne (1970) holds that a miracle is "a non-repeatable counter-instance to a law of nature." A counterinstance to a given law of nature will either be a miracle or will require the formulation of some new alternative law. The second option will be unattractive to the degree that the new law fails to predict the phenomena supporting the original law and/or gives what most likely are false predictions, thus playing havoc with regularities of science. In that event, it would be better to postulate the occurrence of a miracle than to modify our formulae for the laws of nature.

A miracle may, in principle, be identifiable. But is belief in miracles ever epistemically justified? Hume argued that it is not, but his interpreters have disagreed about the specific nature and the overall success of his argument. The controversy centers on Hume's "general maxim" about testimony on behalf of a miracle, which states the following:

> That no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish; and even in that case there is a mutual destruction of arguments, and the superior only gives us an assurance suitable to that degree of force, which remains, after deducting the inferior. (Hume 1975, 115–116)

J. L. Mackie (1982) perceived a need for three specific improvements on Hume's argument: (1) decisions about the value of specific testimonial evidence must always be more provisional than Hume acknowledged; (2) a more accurate conception of inductive generalization is needed than what Hume assumed in his conception of a well-established law of nature; and (3) Hume misunderstood the potential exponential increase in probability conferred by multiple witnesses to an event. But, on balance, with these provisos in place, Hume's argument, Mackie concluded, succeeds in showing that the "intrinsic improbability" of a miracle is too great to be overturned by any degree of testimonial support.

More recently, several of Hume's interpreters have favored the translation of Hume's maxim into the language of the Bayesian probability calculus. This reformulation is desirable if there is to be a more precise exposition of the maxim and of its use in Hume's argu-

ment against miracles. Since Hume's informal statement of his maxim in English harbors ambiguities about conditional probabilities, this is easier said than done.

Let M be some miracle statement; let t(M) be testimony to the occurrence of M; let E be the evidence of "constant and uniform experience" confirming the laws of nature; and let K be background evidence. On a probabilistic reading, the first part of Hume's maxim mentions two probabilities—the probability of the event (M) that some testimony t(M) seeks to establish, and the probability of the falsehood of the testimony, t(M) & ~M—and indicates that the first probability must be greater than the second if the testimony is to make the miracle statement credible. For Bayesian purposes, this relation needs to be translated into a comparative probability statement reflecting conditional probabilities. Everything depends on how each probability statement is rendered.

Jordan Howard Sobel concludes that Hume's maxim, when rendered in the language of probability theory, furnishes the critic with a powerful means of arguing against miracles. One probabilistic reading he has favored is the following:

$$\Pr (M/E \,\&\, K) > \Pr [\sim M \,\&\, t(M))/E \,\&\, K]$$

as if the antecedent probability of the miracle given the laws of nature—Pr (M/E & K)—should count as the crucial thing. If this probability is less than 0.5, as it surely is, then the miracle statement lacks credibility.

Sobel (1987) has also proposed the following alternative reading of Hume's maxim:

$$\Pr [M \,\&\, t(M)/E \,\&\, K] > \Pr [\sim M \,\&\, t(M)/E \,\&\, K]$$

This version differs from the first in proposing an estimation of the prior probability of the *conjunction* of M & t(M) and in stipulating that this probability must be greater than the falsehood of the testimony.

John Earman (2000) identifies one general problem with Hume's maxim and objects to Sobel's rendition of it. First, Hume needs a maxim that specifies a *sufficient* condition for the actual occurrence of a miracle to be more likely than the falsehood of testimony for a miracle. But he specifies only a *necessary* condition.

Second, Earman argues that neither of the two probabilities alluded to in the maxim is captured in the right way by Sobel's translation. Both probabilities should be conditioned on t(M), the testimony to the miracle. This is because Hume is assuming a situation where the one seeking to determine the credibility of a miracle statement is aware of existing testimony to the occurrence of a miracle. The investigator is not concerned with the prior probability of testimony, or, for that matter, with the prior improbability of the miracle. The probability of the falsehood of the testimony should therefore be rendered

$$\Pr [\sim M/t(M) \,\&\, E \,\&\, K]$$

This makes sense of the plausible idea that the probability of falsehood is best determined by "the percentage of cases where no miracle occurs on occasions when the witness testifies to a miracle" (Earman 2000, 41). And the probability of the event some testimony seeks to establish should instead be represented as follows:

$$\Pr [M/t(M) \,\&\, E \,\&\, K]$$

The resulting probabilistic reading of Hume's maxim is as follows:

$$\Pr [M/t(M) \,\&\, E \,\&\, K] > \Pr [\sim M/t(M) \,\&\, E \,\&\, K]$$

This formula is tautologous and unexceptionable. It doesn't have the consequence that either Hume or Sobel intend; it does not deliver an a piori argument against the credibility of miracle statements. It turns out, Earman argues, that this rendition of the first part of Hume's maxim also spells trouble for the coherence of the second part, because it would then appear to be counseling an illicit double counting of countervailing factors in the evidence.

The upshot of Earman's analysis is that there is no way to deploy Hume's maxim in an argument against miracles without a careful exploration of the details of historical evidence concerning a miracle statement and testimonial evidence offered in support of it. This Hume failed to do. Earman himself stops short of carrying out this investigation and is, in effect, agnostic about the occurrence of miracles.

While there seems to be no a priori philosophical argument against the credibility of miracle statements, there are philosophical resources for estimating the value of historical evidence, such as that developed by New Testament historians for the resurrection (see, for example, N. T. Wright [2003]). Richard Swinburne (2003) goes further, estimating the value of background evidence for theism together with specific historical evidence for the resurrection of Jesus. The background evidence, he argues, provides good reason to expect a miracle or two, reducing the weight that must be carried by specific historical evidence in order to establish the actual occurrence of a miracle. In turn, the historical evidence tends

to confirm the theistic hypothesis because it indicates the fulfillment of the expectation generated by the background evidence and is itself additional evidence for the existence of God.

The philosophical study of the concept of miracle has led to a fruitful exploration of more general issues, such as the metaphysics of causation, the epistemology of testimonial evidence, and, in the philosophy of science, the proper conception of a law of nature.

See also Deism; Earman, John; Hume, David; Laws of Nature; Mackie, John Leslie.

Bibliography

Earman, John. *Hume's Abject Failure: The Argument against Miracles*. Oxford: Oxford University Press, 2000.

Geivett, R. Douglas, and Gary R. Habermas, eds. *In Defense of Miracles*. Downers Grove, IL: InterVarsity Press, 1997.

Hume, David. *Enquiries concerning Human Understanding and concerning the Principles of Morals*. Reprinted from the 1777 edition with Introduction and Analytical Index by L. A. Selby-Bigge; with text revised and notes by P. H. Nidditch. 3rd ed. Oxford: Clarendon Press, 1975.

Mackie, J. L. "Miracles and Testimony." Chapter 1 in *The Miracle of Theism: Arguments for and Against the Existence of God*. Oxford: Clarendon Press, 1982.

Sobel, J. H. "On the Evidence for Testimony for Miracles: A Bayesian Interpretation of David Hume's Analysis." *Philosophical Quarterly* 3 (1987): 166–86.

Sobel, J. H. "Hume's Theorem on Testimony Sufficient to Establish a Miracle." *Philosophical Quarterly* 41 (1991) 229–237.

Swinburne, Richard. *The Concept of Miracle*. London: Macmillan, 1970.

Swinburne, Richard. *The Resurrection of God Incarnate*. Oxford: Clarendon Press, 2003.

Wright, N. T. *The Resurrection of the Son of God*. Minneapolis, MN: Fortress Press, 2003.

R. Douglas Geivett (2005)

MIRANDOLA, COUNT GIOVANNI PICO DELLA

See *Pico Della Mirandola, Count Giovanni*

MIURA BAIEN

(1723–1789)

Miura Baien, a Japanese Confucianist who in the era of Tokugawa rule most closely approached Western philosophy, was born in Ōita prefecture on the island of Kyūshū. After the usual training in Chinese classics, Miura went to Nagasaki and learned astronomy, physics, medicine, and economics and developed a great admiration for Western experimental methods. This explains in part his rationalism in opposition to the general reliance on the authority of the classics. He devoted his life to scholarship, refusing several offers to serve feudal lords. To help the poor he organized a relief society based on communal principles. Miura's encyclopedic knowledge also included economics. In *Kagen* (The origin of price) he discussed currency like his contemporary Adam Smith. Miura wrote "if bad money finds wide circulation, good money will go into hiding," a statement similar, in words at least, to Gresham's law.

Miura's main philosophical works are three: *Gengo* (Abstruse words), an exposition of logic; *Zeigo* (Superfluous words), an exposition of the philosophy of nature; and *Kango* (Presumptuous words), an exposition of ethics. *Gengo* is highly esteemed as original because in it he expounds his ideas of *jori,* or the logic of "things" (an abstract concept covering everything). This logic is based not on ancient authority but on rational or experimental grounds. Miura built his logic according to the laws of nature and things. In these he saw a unity and order of antithetic natural elements. He called his dialectic *hankan gōitchi,* or "synthesis of the contraries." This dialectic is both a logical device and the inner reality of things. Things, which are always in the process of becoming, pass from unity to multiplicity and back again, through antithesis and synthesis. His merits as the forerunner of modern trends in science and philosophy notwithstanding, Miura had rather staid political and theological ideas. His criticism of Christianity, in *Samidare-shō,* focuses on the idea that a foreign religion that puts God before devotion to one's lord and one's father cannot be tolerated.

See also Chinese Philosophy; Japanese Philosophy; Nature, Philosophical Ideas of; Philosophy of Economics; Smith, Adam.

Bibliography

Miura's works are available in Japanese in *Baien zenshū* (The collected works of Miura Baien), 2 vols. (Tokyo, 1912). See also G. K. Piovesana, "Miura Baien, and His Dialectical and Political Ideas," in *Monumenta Nipponica* 20 (1965): 389–443, which contains a translation of Miura's letter "Answer to Taga Bokkyō." See also W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 489–497; N. S. Smith, "An Introduction to Some Japanese Economic Writings of the 18th Century," in *Transactions of the Asiatic Society of Japan,* 2nd series, 11 (1934): 80–88; and

L. Hurvitz, "The *Samidareshō,*" in *Monumenta Nipponica* 8 (1953): 289–326; 9 (1953): 330–356.

Gino K. Piovesana, S.J. (1967)

MODAL INTERPRETATION OF QUANTUM MECHANICS

The term *modal interpretation* is ambiguous. It is a proper name that refers to a number of particular interpretations of quantum mechanics. And it is a term that singles out a class of conceptually similar interpretations, which includes proposals that are not generally referred to as modal ones.

This ambiguity was already present when Bas C. van Fraassen coined the term in the 1970s by transposing the semantic analysis of modal logics to quantum logic. The resulting modal interpretation of quantum logic defined a class of interpretations of quantum mechanics, of which van Fraassen developed one instance in detail, called the Copenhagen modal interpretation. In the 1980s Simon Kochen and Dennis Dieks developed independently an interpretation of quantum mechanics that became known as *the* modal interpretation, turning the term into a proper name. In the 1990s further research produced new proposals, broadening attention to the class of modal interpretations.

The development of modal interpretations can be positioned as attempts to understand quantum mechanics as a theory according to which some but not all observables of physical systems have definite values. Quantum mechanics predicts the outcomes of measurements of observables pertaining to systems and is typically silent about whether these observables have values themselves. Attempts to add to quantum mechanics descriptions of systems in which all quantum-mechanical observables have values became deadlocked in the 1960s: Kochen and Ernst Specker's no-go theorem proved that such descriptions are inconsistent if these values have to comply to the same mathematical relations as the observables themselves; John S. Bell's inequalities showed that the descriptions easily lead to nonlocal phenomena at odds with relatively theory (Redhead 1987). Modal interpretations add descriptions to quantum mechanics according to which only a few *preferred observables* have values, and avoid in this way specifically the Kochen-Specker theorem.

A second common element is that modal interpretations do not ascribe one state to a system, as quantum mechanics does, but two: a dynamical state and a value state. By doing so another peculiarity of quantum mechanics is overcome, namely that states of systems evolve alternately by *two* mutually incompatible laws: the Schrödinger equation that yields smooth state evolution in between measurements, and the projection postulate that yields discontinuous evolution at measurements. In modal interpretations dynamical states of systems evolve with the Schrödinger equation only, and value states evolve typically discontinuously. A particular modal interpretation is now characterized by the value states it assigns to systems; value states fix the preferred definite-valued observables and their values.

Finally there is the claim that modal interpretations stay close to quantum mechanics. The dynamical states that modal interpretations assign can be taken as the states that quantum mechanics assigns, the only difference being that the former do not evolve by the projection postulate. Modal interpretations may thus be said to incorporate quantum mechanics instead of replacing it, as some hidden-variables theories do.

QUANTUM-MECHANICAL HILBERT-SPACE MATHEMATICS

In quantum mechanics the state and observables of a physical system are represented by mathematical entities defined on a Hilbert space associated with the system. A Hilbert space H contains vectors $|\psi\rangle$, and if it is an n-dimensional space, there exist sets $\{|e_1\rangle, |e_2\rangle, \dots |e_n\rangle\}$ of n vectors that are pair-wise orthogonal. Such a set is called a basis of the space, which means that any vector $|\psi\rangle$ in H can be decomposed as a weighted sum of the elements of the basis: $|\psi\rangle = \sum_i c_i |e_i\rangle$. The Hilbert space associated with two disjoint physical systems consists of the tensor product $H_1 \otimes H_2$ of the Hilbert spaces associated with the separate systems. If $\{|e_1\rangle, \dots |e_n\rangle\}$ is a basis of H_1 and $\{|f_1\rangle, \dots |f_m\rangle\}$ a basis of H_2, then any vector $|\Psi\rangle$part of $H_1 \otimes H_2$ can be decomposed as a sum $|\Psi\rangle = \sum_{i,j} C_{ij} |e_i\rangle \otimes |f_j\rangle$ (a double summation).

Linear operators A on a Hilbert space are linear mappings within that space. The operator that projects any vector on the vector $|\psi\rangle$ is called a projector and is written as $|\psi\rangle\langle\psi|$. In quantum mechanics the state of a system is represented by such a projector, or by a density projector W which is a complex sum $\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ of projectors. An observable pertaining to a system (e.g., its momentum or spin) is represented by a self-adjoint operator A. Self-adjoint operators and density operators can be decom-

posed in terms of their eigenvalues a_i and projectors on their pair-wise orthogonal eigenvectors $|a_i\rangle$, that is, $A=\sum_i a_i|a_i\rangle\langle a_i|$. (Complications due to degeneracies, phase factors, and infinities are ignored.)

PARTICULAR MODAL INTERPRETATIONS

In all interpretations named modal, the dynamical state of a system is represented by a density operator W on the system's Hilbert space. This dynamical state evolves with the Schrödinger equation and has the usual quantum-mechanical meaning in terms of measurement outcomes: If observable A is measured, its eigenvalue a_i is found with probability $p(a_i)=\langle a_i|W|a_i\rangle$.

The value state of a system is represented by a vector $|v\rangle$ and determines the values of observables by the rule: A has value a_i iff $|v\rangle$ is equal to the eigenvector $|a_i\rangle$ of A. This rule leaves many observables without values; a specific value state is an eigenvector of only a few operators, which then represent the preferred observables. Particular modal interpretations fix the value states of systems differently.

In van Fraassen's (1973, 1991) Copenhagen modal interpretation $|v\rangle$ is a vector in the support of the dynamical state (which implies that W can be written as a convex sum of $|v\rangle\langle v|$ and other projectors). Van Fraassen is more specific about value states after measurements. If an observable A of a system is measured, the dynamical state of the composite of system and measurement device may become $|\Psi\rangle\langle\Psi|$, with $|\Psi\rangle=\sum_i c_i|a_i\rangle\otimes|R_i\rangle$. The vectors $|a_i\rangle$ are eigenvectors of the measured observable, and the $|R_i\rangle$'s are eigenvectors of a device observable that represents the outcomes (the pointer readings). The value states after this measurement are, according to van Fraassen, with probability $|c_i|^2$ simultaneously given by $|a_i\rangle$ for the system and by $|R_i\rangle$ for the measurement device, respectively.

The decomposition $|\Psi\rangle=\sum_i c_i|a_i\rangle\otimes|R_i\rangle$ is mathematically special because it contains one summation (as said, a decomposition of a vector $|\Psi\rangle$ in a product space $H_1\otimes H_2$ relative to bases of the separate Hilbert spaces has usually a double summation). This special single-sum decomposition is called the bi-orthogonal decomposition of $|\Psi\rangle$, and a theorem (Schrödinger 1935) states that every vector $|\Psi\rangle$ in $H_1\otimes H_2$ determines exactly one basis $\{|e_1\rangle, \dots |e_n\rangle\}$ for H_1 and one basis $\{|f_1\rangle, \dots |f_m\rangle\}$ for H_2 for which its decomposition becomes such a bi-orthogonal decomposition.

Kochen (1985) and Dieks (1989) use this decomposition to define value states in their modal interpretation: If two disjoint systems have a composite dynamical state $|\Psi\rangle\langle\Psi|$ and the bi-orthogonal decomposition of the vector $|\Psi\rangle$ is $|\Psi\rangle=\sum_i c_i|e_i\rangle\otimes|f_i\rangle$, then the value states are with probability $|c_i|^2$ simultaneously $|e_i\rangle$ for the first system and $|f_i\rangle$ for the second. Kochen adds a perspectival twist to this proposal, absent in Dieks's earlier writing: For Kochen the first system witnesses the second to have value state $|f_i\rangle$ iff it has itself value state $|e_i\rangle$ (which is the case with probability $|c_i|^2$) and the second system then witnesses, conversely, the first to have value state $|e_i\rangle$.

The Kochen-Dieks proposal applies to two systems with a composite dynamical state represented by a projector $|\Psi\rangle\langle\Psi|$ only. The spectral modal interpretation by Pieter Vermaas and Dieks (1995) generalizes this proposal to n disjoint systems with an arbitrary composite dynamical state W. This composite state fixes the dynamical states of all subsystems. Let $W(x)$ be the dynamical state of the x-th system part of the composite and let it have an eigenvalue-eigenvector decomposition $W(x)=\sum_i w_i(x)|w_i(x)\rangle\langle w_i(x)|$. The value state of this x-th system is then $|w_i(x)\rangle$ with probability $w_i(x)$. Vermaas and Dieks gave, moreover, joint probabilities that the disjoint systems have simultaneously their value states $|w_i(1)\rangle$, $|w_j(2)\rangle$, etcetera.

In the spectral modal interpretation a composite system, say, system 1+2 composed of the disjoint systems 1 and 2, has an eigenvector $|w_k(1{+}2)\rangle$ of its dynamical state $W(1{+}2)$ as its value state. The atomic modal interpretation by Guido Bacciagaluppi and Michael Dickson (1999) fixes the value states of such composite systems differently. Bacciagaluppi and Dickson assume that there exists a set of disjoint atomic systems, for which the value states are determined similarly as in the spectral modal interpretation, and propose that the value states of composites of those atoms are tensor products of the value states of the atoms: the value state of the composite of atoms 1 and 2 is $|w_i(1)\rangle\otimes|w_j(2)\rangle$ iff the value states of the atoms are $|w_i(1)\rangle$ and $|w_j(2)\rangle$, respectively.

THE CLASS OF MODAL INTERPRETATIONS

The class of modal interpretations comprises those proposals according to which only a few observables have values, and that can be formulated in terms of dynamical and value states. The interpretations by Richard Healey (1989) and by Jeffrey Bub (1997) have this structure quite explicitly and are therefore often called modal ones (Healey's proposal has a number of similarities with the Kochen-Dieks proposal; in Bub's the value state of a system is an eigenvector of an observable fixed independ-

ently of the system's dynamical state). One may argue that David Bohm's mechanics (1952) is also a modal interpretation.

RESULTS

The development and application of modal interpretations have led to mixed results. The maximum set of observables that can have values by modal interpretations without falling prey to the Kochen-Specker theorem has been determined (Vermaas 1999). Bub and Rob Clifton showed that this set is the only one that satisfies a series of natural assumptions on descriptions of single systems (Bub, Clifton, and Goldstein 2000). The evolution of value states, which determines the description of systems over time, can be given (Bacciagaluppi and Dickson 1999). This evolution was, however, shown not to be Lorentz-covariant for the spectral and atomic modal interpretations and, to a lesser extent, for Bub's interpretation, revealing that the assumption that only a few quantum-mechanical observables have values, still may lead to problems with relatively theory (Dickson and Clifton 1998, Myrvold 2002).

Moreover, even though this assumption yields consistent descriptions of single systems, joint descriptions of systems were still proved to be problematic. First, it is commonly assumed in quantum mechanics that the observable of a system 1 represented by the operator A defined on H_1, and the observable of a composite system 1+2 represented by the operator $A_1 \otimes I_2$ on $H_1 \otimes H_2$ (I_2 is the identity operator on H_2) are one and the same observable. The Copenhagen, Kochen-Dieks, and spectral modal interpretations have the debatable consequence that these observables should be distinguished (Clifton 1996). Second, the spectral modal interpretation cannot give joint probabilities that systems 1, 2, … , and their composites, 1+2, … , have simultaneously their value states $|w_i(1)\rangle$, $|w_j(2)\rangle$, $|w_k(1+2)\rangle$, etcetera (Vermaas 1999, ch. 6).

These negative results motivated in part the formulation of the atomic modal interpretation but can also be avoided by adopting Kochen's perspectivalism, which implies that one accepts constraints on describing different systems simultaneously. Finally, the Kochen-Dieks, spectral, and atomic modal interpretations have problems with properly describing measurements, doubting their empirical adequacy. David Albert and Barry Loewer (1990) argued that after a measurement, the dynamical state of the system-device composite need not be $|\Psi\rangle\langle\Psi|$ with $|\Psi\rangle = \sum_i c_i |a_i\rangle \otimes |R_i\rangle$, and that the mentioned interpretations then need not yield descriptions in which the device displays an outcome (Bacciagaluppi and Hemmo 1996).

ASSESSMENT

These results allow critical conclusions about particular modal interpretations and raise doubts about the viability of the class of modal interpretations. Three remarks can be made about this assessment.

First, an evaluation of the results may depend on what one expects from interpretations. If interpretations are to provide descriptions that allow realist positions about quantum mechanics, the inability of, say, the spectral modal interpretation to give joint probabilities that systems have simultaneously value states, proves this interpretation problematic. But if interpretations, in line with van Fraassen's view, are to yield understanding of what quantum mechanics means, this inability of the spectral modal interpretation is an interesting conclusion about how quantum-mechanical descriptions of systems differ from those of other physical theories. The result that some modal interpretations may be empirical inadequate, is, however, fatal independently of one's expectations for interpretations.

Second, the set of particular modal interpretations that is analyzed so far does not exhaust the class of modal interpretations. Research therefore continues (e.g., Bene and Dieks 2002).

Third, these results are relevant to the project of interpreting quantum mechanics in general. Existing and new interpretations, modal or not, according to which only some observables have definite values, are constrained by the negative results and can now be assessed as such; and existing and new interpretations may benefit from the positive results about modal interpretations.

See also Bell, John, and Bell's Theorem; Bohm, David; Quantum Mechanics; Van Fraassen, Bas.

Bibliography

Albert, David Z., and Barry Loewer. "Wanted Dead or Alive: Two Attempts to Solve Schrödinger's Paradox." In *Proceedings of the 1990 Biennial Meeting of the Philosophy of Science Association*. Vol. 1., edited by Arthur Fine, Micky Forbes, and Linda Wessels. East Lansing, MI: Philosophy of Science Association, 1990.

Bacciagaluppi, Guido, and Michael Dickson. "Dynamics for Modal Interpretations." *Foundations of Physics* 29 (1999): 1165–1201.

Bacciagaluppi, Guido, and Meir Hemmo. "Modal Interpretations, Decoherence and Measurement." *Studies in History and Philosophy of Modern Physics* 27 (1996): 239–277.

Bene, Gyula, and Dennis Dieks. "A Perspectival Version of the Modal Interpretation of Quantum Mechanics and the

Origin of Macroscopic Behavior." *Foundations of Physics* 32 (2002): 645–671.

Bohm, David. "A Suggested Interpretation of Quantum Theory in Terms of 'Hidden Variables.'" *Physical Review* 85 (1952): 166–193.

Bub, Jeffrey. *Interpreting the Quantum World*. Cambridge, U.K.: Cambridge University Press, 1997.

Bub, Jeffrey, Rob Clifton, and Sheldon Goldstein. "Revised Proof of the Uniqueness Theorem for 'No Collapse' Interpretations of Quantum Mechanics." *Studies in History and Philosophy of Modern Physics* 31 (2000): 95–98.

Clifton, Rob. "The Properties of Modal Interpretations of Quantum Mechanics." *British Journal for the Philosophy of Science* 47 (1996): 371–398.

Dickson, Michael, and Rob Clifton. "Lorentz-Invariance in Modal Interpretations." In *The Western Ontario Series in Philosophy of Science*. Vol. 60, *The Modal Interpretation of Quantum Mechanics*, edited by Dennis Dieks and Pieter E. Vermaas. Dordrecht, Netherlands: Kluwer, 1998.

Dieks, Dennis. "Quantum Mechanics Without the Projection Postulate and Its Realistic Interpretation." *Foundations of Physics* 19 (1989): 1397–1423.

Healey, Richard A. *The Philosophy of Quantum Mechanics: An Interactive Interpretation*. Cambridge, U.K.: Cambridge University Press, 1989.

Kochen, Simon. "A New Interpretation of Quantum Mechanics." In *Symposium on the Foundations of Modern Physics*, edited by Pekka Lahti and Peter Mittelstaedt. Singapore: World Scientific, 1985.

Myrvold, Wayne C. "Modal Interpretations and Relativity." *Foundations of Physics* 32 (2002): 1773–1784.

Redhead, Michael. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press, 1987.

Schrödinger, Erwin. "Discussion of Probability Relations Between Separated Systems." *Proceedings of the Cambridge Philosophical Society* 31 (1935): 555–563.

van Fraassen, Bas C. *Quantum Mechanics: An Empiricist View*. Oxford: Clarendon Press, 1991.

van Fraassen, Bas C. "Semantic Analysis of Quantum Logic." In *The University of Western Ontario Series in Philosophy of Science*. Vol. 2, *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, edited by C. A. Hooker. Dordrecht, Netherlands: Reidel, 1973.

Vermaas, Pieter E. *A Philosopher's Understanding of Quantum Mechanics: Possibilities and Impossibilities of a Modal Interpretation*. Cambridge, U.K.: Cambridge University Press, 1999.

Vermaas, Pieter E., and Dennis Dieks. "The Modal Interpretation of Quantum Mechanics and Its Generalization to Density Operators." *Foundations of Physics* 25 (1995): 145–158.

Pieter E. Vermaas (2005)

MODALITY, PHILOSOPHY AND METAPHYSICS OF

Some things are true; some are false. Some are true, but might have been false. Some are true but could not have been false. Some are false, but might have been true. Some are false but could not have been true. Thus, there are at least these different modes of truth and falsity: necessity and possibility. A truth bearer—a proposition, a statement, or an interpreted sentence—is necessarily true if and only if (iff) it is not possible that it be false; it is possibly true iff it is not necessary that it be false. A contingency is what is possibly true as well as possibly false. The study of the ways in which truth and falsity interact with necessity and possibility is the subject of modal logics.

If there are modal distinctions to be made about truth bearers—because, say, while the sum of seven and five could not fail to be twelve, there might have been twelve planets orbiting the sun even though there are not—arguably there are modal distinctions to be made regarding the attributes objects possess. While Socrates could have failed to be snub-nosed, he could not have failed to be human or perhaps a person. Modality as it pertains to the bearers of truth is *de dicto* modality; modality that concerns the way in which an object possesses an attribute is *de re*. Conventionally, □ and ◊ express necessity and possibility, respectively.

KINDS OF NECESSITY

Necessities may be distinguished according to their scope or, perhaps, their subject matter. Some concern the limits of meaning and inference and are systematized by formal logical systems. Classical logicians maintain that meaning and inference are best understood in terms of a two-valued logic according to which truth bearers may take only the values of true or false and that exactly one of a truth bearer and its negation is true. These ideas are encoded by, though not strictly equivalent to, the laws of excluded middle, $P \vee \neg P$, and noncontradiction, $\neg(P \& \neg P)$. Where objects and their attributes are concerned, there are analogous laws: Each thing either has or lacks a given attribute and neither both has and lacks the same attribute, formally represented as $\forall x(Fx \vee \neg Fx)$ and $\forall x\neg(Fx \& \neg Fx)$. Some nonclassical logics have correspondingly different foundational laws. Such laws of logic are treated as necessary truths, though usually they are stated without any *de dicto* modal qualifier. Logical truths, the truths that follow validly from the axioms of logic, are the logical necessities. Whatever is consistent

with the laws of logic are the logical possibilities and if the falsity of something follows from the laws of logic, then it is a logical impossibility.

Disputes about which of the many logical systems is the single correct system that formalizes valid inference are, then, disputes about what the logical necessities are. Disputes about whether there is a single such correct system are disputes about whether there is a plurality of sets of logical necessities, one corresponding to each system of logic. Broader notions of logical necessity are sometimes given by adding linguistic or conceptual truths. Analytic or conceptual necessities are those that follow from the laws of logic plus the linguistic or conceptual truths. That all bachelors are unmarried is a favorite example of such an analytic truth.

Laws of nature and their associated counterfactual conditionals are often thought to be necessary in one sense and contingent in another. There are many true universal generalizations and many unbroken sequences of types of events. Not all of these, however, form part of some scientific theory; some generalizations and patterns are accidental. Those that are not accidental and constitute the fundamentals of a scientific theory are the laws of nature and whatever follows from those laws are the natural necessities. Whatever is consistent with the laws of nature are the natural possibilities. It is sometimes useful to make distinctions among the natural necessities and separate the physical, chemical, biological, psychological, and perhaps other necessities.

The natural necessities are not logical truths, however. Orthodoxy has it that logical truth is known *a priori*, without any specific experiences, while scientific truth is knowable only *a posteriori*, on the basis of experience. That empirical investigation is required for scientific knowledge is taken to show that the natural necessities are contingent; they might have been otherwise. Certainly, they are not analytic and cannot be known simply by reflecting on their contents. If one takes the most central laws of physics to be the axioms of physics, then what follows from those laws are the physical necessities. That the basic physical laws are required to infer the physical necessities demonstrates more conclusively that the physical necessities are not logical truths and, so, are not logical necessities. If laws of nature are necessary in some legitimate sense, then the meaning of $\Box$ when applied to laws of nature must differ from its meaning when it is applied to laws of logic.

Somewhat more controversial is the idea that there is an intermediate modality between the logical and the natural: the metaphysical. Like natural necessities, metaphysical necessities are not logical truths and yet, unlike the natural necessities, the central metaphysical principles are to be known *a priori* even if knowledge of some particular metaphysical necessities requires some empirical knowledge. While there is no contradiction in denying such metaphysical principles, their proponents maintain that they are, nevertheless, necessary and that they express limits on genuine possibilities for existence. Accordingly, logical possibilities that fail to be metaphysical possibilities would be merely formal possibilities.

Thus, if there are two distinct attributes that are essentially related, then while there is no contradiction in asserting that an object could possess the one without the other, it would be, strictly speaking, impossible for any object to possess the one without the other. There might be attributes such that if an object even possibly possesses it, then that object possesses that attribute essentially. Arguably, if something is a concrete object such as a brick, then it must be concrete and could not be an abstract object such as the power set of the real numbers. Those who embrace this intermediate modality think that metaphysical principles state nontrivially what is at least part of the essence of an object—that without which it could not be—and, so, they are known as essentialists or sometimes Aristotelian essentialists since Aristotle advocated a form essentialism.

While the philosophy of modality is dominated by controversies about whether there are the three modalities mentioned and, if so what their relations are, there are others of interest. If one takes up the general pattern used earlier and recognizes that a common way to characterize the content of a modality is to formulate a set of axioms and define a sense of $\Box$ so that it applies to all and only whatever follows from those axioms, then there are indefinitely many kinds of modality. For each formally characterized system of logic there is a candidate for logical necessity, not all of which are equivalent. Such a plurality cascades down through any modality that relies on logical consequence for its own characterization.

Among the most commonly discussed of the other modalities are the epistemic, doxastic, and moral necessities. Epistemic necessity can be thought of as whatever follows from what is known and the scope of the known can be specific to an individual or to a community. Doxastic necessity is what follows from what is believed. Moral necessities are the relevant moral obligations and duties. Whether any of the modalities mentioned is nothing but a special form of any of the others is a substantial question, but the only clear connection, given the way they have been characterized, is that all the nonlogical

modalities tacitly embed the logical. The axiom-theorem structure demands this, making each of them an extension of the logical and not a special case of it.

SOURCES OF NECESSITY

For any recognized kind of necessity there is the substantive question of what accounts for the fact that some truths, and not some others, are necessary in the relevant sense. As with other forms of discourse, there are deep philosophical questions about whether one has any knowledge involving modality. These questions have at least two forms. The first form is a standard challenge from the skeptic who does not deny, in this context, that there are necessary truths but who denies either that one has any knowledge of which truths are necessary or that nonskeptics are entitled to their knowledge claims according to their own standards of knowledge. The second form comes from the modal antirealist, one who either denies that there are any modal truths or who claims that modal truth is so closely bound to cognizers that statements involving modality lack significant objectivity, making them more like statements of taste or preference than statements of fact. Modal noncognitivists maintain that modal discourse is used not to make assertions but to, perhaps, express an attitude or a commitment toward some nonmodal truths. David Hume (1739) adopted a kind of noncognitivism about the relation of cause and effect. Simon Blackburn (1986) and Crispin Wright (1980, 1989) advanced contemporary defenses of versions of modal antirealism.

Realist interpretations of modal discourse treat some statements involving modality as true in some person-independent manner. If modal truth is to be a species of truth more generally then, thinks the realist about modal discourse, modal truth must concern sufficiently determinate and objective facts. This is the question of what grounds modal truth or, perhaps, what the truth conditions are for modal claims. Common suggestions have been that something is possible iff it is conceivable, iff it implies no contradiction, iff it is true in at least one mathematical model, or iff it is true in at least one possible world.

The success conditions for a theory of modality depend on the purpose of that theory. Any successful theory should be extensionally adequate; it should declare as necessary all and only what is necessary. Philosophical theories are often put forward as being more than merely extensionally adequate, sometimes because they are intended as linguistic or conceptual analyses. A conceptual analysis would state not only the appropriate biconditionals that fix the extension of *necessary*, it would do so in such a way that analyzed the meaning of that term. Successful conceptual analyses must not only be extensionally adequate, they must also be noncircular by avoiding the use of what is to be analyzed in the analysis or definition. Analyses of modal notions would need to avoid analyzing *necessary* in terms of any modal notions like *necessary*, *possible*, or their cognates. When empirical confirmation of the extensional adequacy of a theory is impossible to obtain, conceptual analyses are attractive. The conceptual analysis guarantees extensional adequacy because the analysis given means nothing more and nothing less than that for which it is a theory. Theories not intended as conceptual analyses must involve some other warrant for the thesis that the theory is extensionally adequate.

If proposed as an analysis, a conceivability theory faces difficulties regarding both extensional adequacy and circularity. If it is formulated in terms of conceivers who are not perfect conceivers, then extensional adequacy is not guaranteed; there may be possibilities of which no one is capable of conceiving or else conceivers may be capable of conceiving what is impossible without noticing that it is impossible. Formulating the theory in terms of what is merely conceivable, whether by an ideal or fallible conceiver, renders the analysis circular because the semantic analysis of *possible* is given in terms of what it is possible to conceive. Formulating it in terms of what is actually conceived by an omniscient being avoids this circularity but brings metaphysical commitments that few want to make on the basis of their philosophy of modality alone.

The logical positivists wished to maintain that logical and mathematical truths were necessary, but they resisted all substantive metaphysics as distinct from the ontologies of the sciences. Alfred J. Ayer (1936) developed a version of conventionalism about modality. By dividing propositions into the classes of those that concern ideas or concepts only and those that concern facts, Ayer maintained that only propositions regarding ideas were both necessary and knowable *a priori*. They make no claims about the empirical world and, so, are not subject to empirical falsification and are either necessarily true or necessarily false. These propositions are analytically true or analytically false according to Ayer because they are true or false due solely to the definitions or analyses of their constituent symbols, both logical and nonlogical. It is necessarily true that all bachelors are unmarried not because of the way the world is but because of what is meant by *all*, *bachelors*, *unmarried*, and tacitly, *if, then*.

Willard Van Orman Quine provided what has become the received critique of attempts to ground necessary truth and falsehood in the facts of language. In "Two Dogmas of Empiricism" (1951) Quine argued that there is no hard and fast distinction between propositions that are about the world and those that are not and, so, that no proposition is immune from refutation on partly empirical grounds. Thus, he argued that there is no interesting analytic-synthetic distinction on which the positivist program depends. In "Truth by Convention" (1948) he argued that stipulations regarding the meanings of expressions cannot be a general source of necessity, since at most they can transform obvious logical truths into more convenient but less obvious truths.

So, it is a logical truth that all unmarried males are unmarried and if bachelors just are, by definition, unmarried males, then the logical truth plus the definition of *bachelor* is sufficient for the truth of "all bachelors are unmarried." However, this transforming work of definitions requires something to begin with that is already necessarily true: the relevant logical truth. Linguistic conventions are unable to account for the necessity of the logical truths. Rudolf Carnap (1954) tried to solve this problem by avoiding the semantic foundation of meaning, thus avoiding Quine's critique, and by relying on syntactic facts of grammar and rules of logical proof. He understood logical truth as what is derivable from the null class of sentences.

While not relying on meanings, the standard problems regarding extensional adequacy and circularity arise. Standard understandings of logical systems have it that there are infinitely many sentences that may be derived from the null set, not all of which have been derived. Framing the theory in terms of what has actually been derived renders the account extensionally inadequate, while framing it in terms of facts of derivability renders it circular. A successful form of the linguistic theory of modality might retreat from the positivist's rejection of all metaphysics and appeal to facts about concepts or propositions in a Platonic Heaven of abstract objects. Alternatively, there could be a stipulation by a kind of ostention according to which *necessary* is stipulated to apply to some already established classes of truths, say logical, mathematical, and analytic truths. This would give one a kind of conventional basis for necessity, but not for the truth of what is by this convention called *necessary*. This account assumes that there are logical, mathematical, and analytic truths before the stipulation. While each account avoids the problems posed for Ayer's (1936) and Carnap's, they do not deliver what the positivists wanted: a general theory that demonstrates why logical, mathematical, and analytic truths are completely immune from empirical refutation while at the same time avoiding all metaphysics that they found philosophically distasteful. Alan Sidelle (1989) attempted to present a more defensible version of conventionalism.

POSSIBLE WORLDS AND MODAL LOGIC

Before and during the time that the positivists were developing their philosophical approach to modality and Quine (1948, 1951) was subjecting it to critical scrutiny, elementary first-order predicate logic was being extended with the use of modal operators, most famously by Clarence Irving Lewis (1918) and Lewis and Cooper Harold Langford (1932). Unlike the developments of nonmodal logics up to that time, about which there was widespread agreement that alternative axiom systems were equivalent, there were many inequivalent axiomatic systems of modal logic. Worse, standard first-order logics had been provided with mathematical semantic foundations from which the systems of proof could be shown to be adequate for proving all theorems of first-order logic and for never permitting the derivation of any nontheorems. Modal logics lacked a similar semantic framework. The many inequivalent systems made it impossible, on formal grounds alone, to determine which logic was the proper formalization of modal concepts that, in turn, caused some to wonder whether modal concepts were sufficiently respectable to be given systematic treatment.

Part of the difficulty arose because the modal expressions in formal languages, $\Box$ and $\Diamond$, were treated like the negation symbol, $\neg$. Thus, if P were a sentence of the formal language, $\neg P$, $\Box P$, and $\Diamond P$ would also be sentences of the language. Like negation, the modal operators could be used in quantified sentences of the language, so that if $\forall x Fx$ and $\exists x Fx$ were sentences of the language, $\Box\forall x Fx$, $\forall x\Box Fx$, $\Box\exists x Fx$, and $\exists x\Box Fx$ would be as well. The *de dicto* use of modality in $\Box\forall x Fx$ and $\Box\exists x Fx$ seemed innocent enough to those who were not convinced by Quine's (1951) critique of analyticity. More troublesome were the *de re* forms, $\forall x\Box Fx$ and $\exists x\Box Fx$. In stating that everything is necessarily or essentially F and that something is necessarily or essentially F, these sentences seem to make metaphysical claims, about which the positivists had succeeded in raising suspicion.

In 1963 Saul Kripke made prominent some developments in the semantics of modal logic. The central idea was to mimic an important aspect of the formal semantics for first-order logic. The mathematics of model the-

ory that had enabled logicians to define what it is for an argument to be formally valid involved appealing to a domain of objects, mathematical models, that were customarily thought to be abstract objects. In these models, one could define the extensions of predicates, intuitively the sets of objects that possessed the relevant attributes or that stood in the relevant relations to each other. Logical notions like validity could be defined in terms of these mathematical models.

Kripke and others saw that if this model-theoretic framework were extended, a similar formal semantics could be given for modal logics. Whereas standard models had concerned only everything that does exist, the extension of this approach was simply to take as the domain everything that exists not only in the actual world but also everything that exists in every possible world. The second key idea was to treat the modal operators like quantifiers. If □ was treated as ∀ and ◊ as ∃, then □P could be thought of as a expressing the claim that P is true in every possible world and ◊P could be thought of as expressing the claim that P is true in at least one such world, whether this world or not. A historical overview of developments of this general approach before Kripke's elegant presentation can be found in B. J. Copeland (1996).

POSSIBLE WORLDS AND METAPHYSICS

Those proposing this possible worlds semantics for modal logic thought of the structure quite abstractly. The suggestion to think of the main domain as the set of all possible worlds was merely a heuristic to illuminate the intuitive idea behind the abstract structure of the semantics. It was David Lewis (1973) who recommended taking this heuristic to have metaphysical significance. He argued that modal claims can be paraphrased with claims about possible worlds. Many agreed with this much, but resisted Lewis's genuine modal realism," according to which each world in this plurality was as robust and concrete as one thinks of this world. In some of these worlds there are donkeys that talk and in some there are blue swans. So, while those concerned with the semantics of modal logic were concerned with providing a formal mathematical structure according to which important logical notions like logical consequence could be precisely characterized, Lewis was concerned with the issue of the grounds for the truth values of modal claims. So, for Lewis, □P is true iff P is true in all the worlds; otherwise, not.

The formal apparatus involved an accessibility relation over this set of worlds and that relation could have variable extension. This permitted Lewis (1973) to assess counterfactual conditionals in terms of what happens not merely in some world or other, but what happens in close or sufficiently similar worlds. Thus, in some circumstances I could have done otherwise because in an appropriately similar world one similar to me does otherwise.

Lewis's (1973) genuine modal realism served as the focus of much discussion about the philosophy and metaphysics of modality, although the position has had relatively few adherents. The possible worlds theorist was able to take the mathematical results about modal logic and to find in them the grounds for modal truth. Initial discussions of the possible worlds framework, however, focused on reasons for thinking that while the framework should be adopted, Lewis's metaphysics of possible worlds should be resisted.

One serious problem for the genuine modal realist is epistemological. Suppose that there really is a plurality of concrete worlds and that it is facts about these worlds that make true or false one's modal assertions. How can this account of the truth conditions for modal claims be squared with the often-unstated starting point in the philosophy of modality: that one possesses some knowledge of modal truth that is not merely trivial? One thinks that one knows that there could be talking donkeys, blue swans, and many more things that do not actually exist. One also thinks one knows the truth of some counterfactual conditionals, such as that were the sun to cease to exist, then the earth would cool rapidly and that were a thin pane of ordinary glass to be struck by a flying rock, it would break. If the modal facts, however, really are facts about other worlds, how could one have gained any of this knowledge?

A second apparent problem is that the possible worlds framework looks ill suited to the task of philosophical analysis of modal idioms. If one says that ◊P is true iff P is true in every possible world, then the analysis certainly appears to be extensionally adequate, but at the cost of circularity. If one says, rather, that ◊P is true iff P is true in every world that there is, then obvious circularity is avoided at the cost of no longer exhibiting the extensionally adequacy of the analysis.

The epistemological problem was addressed by those who proposed accounts of the nature of possible worlds in terms of objects that, it was maintained, one already had reason to accept. Instead of thinking of truth in possible worlds as truth in or about concrete maximal spatiotemporal wholes, it was argued that truth in possible

worlds is really truth in maximal states of affairs (Plantinga 1974), truth in world stories—maximal consistent sets of propositions (Adams 1974), or truth about properties of a special kind—ways the world might have been (Stalnaker 1976). Each theory was actualist in that it recognized only objects that actually exist or, to use the vocabulary of possible worlds, each recognized only objects that exist in the actual world. To that extent each of these alternatives had the advantage of locating the ground for modal truths in this world and not another. That there was a useful solution to the general form of the epistemological problem posed for Lewis's (1973) genuine modal realism depends on whether the central feature of the problem was that the modally relevant facts inhabited or constituted worlds distinct from one's own.

Arguably, the central feature of the problem was that it was hard to wed the metaphysics of concrete worlds with plausible accounts of the nature of knowledge. Lewis's account of worlds permitted no physical or causal contact with features of other worlds. To avoid this general problem, some mutually favorable accounts of the natures of states of affairs, propositions, or properties on the one hand and knowledge on the other hand are required. To the extent that these entities are abstract and to the extent that abstract entities are not spatiotemporally or causally located, these actualist theories do not solve this epistemological problem. To the extent that spatiotemporal connectedness is not necessary for access to, say, propositions, then the genuine modal realist could, perhaps, take advantage of an alternative account of knowledge to avoid this particular problem.

Lewis (1986) recognized that his theory of modality could not serve as the basis for a proper analysis of modal notions, if he could not analyze the concept of a possible world. If he could not, possible truth would be analyzed in terms of possible worlds that, while involving some philosophical advance perhaps, does not constitute a full analysis of modal concepts in nonmodal terms. Lewis (1986) argued that each world is a maximal spatiotemporally connected whole; objects inhabit the same world when they spatiotemporally connect to each other. On the reasonably safe assumption that these spatiotemporal notions are not themselves modal, obvious circularity is avoided.

Extensional adequacy must still be secured. Lewis (1986) tries to secure it by somewhat contentious means. He appeals to a Humean principle of recombination to support the thesis that there are sufficiently many possible worlds. Recombination is the denial of necessary connections between distinct existences. So long as the objects occupy distinct spatiotemporal locations, anything could exist with anything else or, strictly speaking, a duplicate of anything could exist with a duplicate of anything else. This basis for plenitude is more contentious than was the avoidance of obvious circularity because it depends on the more controversial Humean principle. Essentialists reject that principle as do those who maintain that laws of nature are metaphysically necessary.

There may yet be some hidden circularity or other theoretical impropriety as argued by Scott A. Shalkowski (1994, 2004). Of course, if there is a plurality of concrete worlds in which sufficiently much of what one takes to be possible is true, then knowing this would be sufficient warrant for declaring that possible truth just is truth in some world or other. It is knowing that there is this match between one's apparent modal knowledge and the internal workings of the worlds in the plurality that is difficult to secure in a nonquestion-begging way. Were philosophical analysis sufficient to justify not only that there are possible worlds, but that they are concrete and sufficiently plentiful for the required correlation, then all would be well for the genuine modal realist. John Divers and Joseph Melia (2002), however, argued that analysis is inadequate to establish that there are sufficiently many worlds. The danger, then, is that the grounds for genuine modal realism as a full theory of modality are question-begging or else inadequate. Furthermore, they argue that the framework may not even be extensionally adequate because there may be no complete set of all possible worlds.

Some objections to genuine modal realism concerned whether the conditions it provides really are adequate to grounding the modal claims one thinks one is entitled to make. For example, one knows that in some instances one could have behaved otherwise than one did. Strictly speaking, though, I am a world-bound individual. I inhabit only this spatiotemporal whole and not another. However, it is what goes on in other worlds that is supposed to account for the fact that I could have behaved otherwise. I could have behaved otherwise because some world contains a counterpart of me that does, in a suitably similar situation, behave otherwise. This is Lewis's counterpart theory (1968, 1986).

Kripke (1972/1980) argued that counterpart theory is inadequate precisely because the modal claim under consideration concerns what I could do. How does what someone else somewhere else does make it the case that I could have followed that alternative course of action? That someone else in this space-time does something else

does not make it the case that I could have done the same, so someone else in another space-time seems no more relevant.

Though Kripke's (1972/1980) objection has intuitive appeal, arguably it is question-begging. Lewis (1986) develops counterpart theory so that the identity of individuals across worlds, transworld identity, just is a matter of having a counterpart in those other worlds. Just as there are philosophical issues about in what identity over time consists, there are philosophical issues about in what identity in modal contexts consists. According to some theories of identity over time, an object that lives for a hundred years is constituted by distinct temporal parts. There is, therefore, precedent for something like counterpart theory. What counts as a counterpart of an object in a distinct world is a matter of relevant similarity, where relevance is determined by, for example, the counterfactual conditional to be assessed. Similar remarks apply to Alvin Plantinga's (1974) objection from numerical identity.

D. M. Armstrong (1989) argued for a somewhat less ontologically ambitious modal realism: combinatorialism. According to combinatorialism possible worlds are recombinations of individuals, properties, and relations of the actual world. Like Lewis (1968), Armstrong relies on a Humean principle of recombination. A recombination of actual objects and actual properties and relations constitutes a nonactual possible world. One difference is that Lewis formulates his principle in terms of duplicates of objects, whereas Armstrong does so in terms of the objects themselves. Where Lewis has no need to countenance qualitatively indiscernible worlds as distinct, Armstrong does. That an object, a, is F and that another, b, is exactly like a except that it is G instead of F, provides for a recombination exactly like the actual world save that in this recombination it is b that is F and a that is G. This seems to involve a commitment to haecceitism, the view that there are nonqualitative differences between worlds. Though this seems to be a natural consequence of his basic combinatorial insight, Armstrong rejects haecceitism.

The are two important issues that confront the combinatorialist. First, some principled, nonmodal restriction on the principle of recombination must be given, since if there is no such restriction, impossible worlds will result and the theory will be extensionally inadequate. With no restriction, there is a recombination in which some object is both wholly red and wholly green, thus rendering it false that $\Diamond P$ is true iff P is true in at least one of the combinatorialist's worlds. Armstrong (1989) attempts restrictions that arise naturally from his own theory of universals in an attempt to solve this problem.

More significant is the problem of alien properties. It is plausible that there could be objects that possess properties that no actual object possesses and that cannot be constructed from any properties that actual objects possess. Unless one is prepared to claim that one's world is maximally qualitatively rich, this consequence is unwelcome. Those who, like Armstrong (1989), wish to acknowledge the existence only of properties that are exhibited, must concede that this is a feature of the theory, in spite of strong reasons to the contrary. For other than special pleading, what reason is there for thinking that this world does not stand to another world in the relation of relative-impoverishment with respect to properties as some simpler worlds stand in relation to this one? Those who adopt a more Platonistic theory of properties and recognize uninstantiated abstract properties avoid this problem of alien properties, but at the cost of needing to solve the epistemological problems regarding one's knowledge of properties rather than one's knowledge of the genuine modal realist's worlds.

FICTIONALISM AND MODALISM

One development that at least initially promises to retain the advantages of genuine realism without this epistemological trouble is modal fictionalism. Strictly speaking, while it is possible that there be talking donkeys, there are none. However, it is also literally true that according to the fiction of possible worlds, there are worlds in which there are talking donkeys. Gideon Rosen (1990) suggested that $\Diamond P$ is true iff according to the fiction of possible worlds, P or some appropriate paraphrase is true in some possible world. Possible worlds are taken to be useful fictions in the same way that scientists have found ideal gases and frictionless planes to be theoretically useful. Whether fictionalism gains any theoretical advantage over modal realism depends on the content of the operator "according to the fiction of possible worlds." It is natural to think that this should be interpreted as something like: "if the fiction of possible worlds were true, then," which is apparently modal.

Whether this is a problem for fictionalism is a matter of its point. If it is to possess all the advantages that Lewis (1986) claimed, specifically an account of all modal truth, then if the fictional operator is modal, fictionalism fails. The fictionalist also confronts a problem with incompleteness. No modal realist has given a complete specification of the contents of each world, so strictly speaking the modal fictionalist is confronted with truth value gaps

for the modal claims about which the modal realist has been silent. The realist can be content with this silence since the realist need not be committed to anyone being modally omniscient. That there are gaps in the fictionalist account is a departure from orthodoxy that must be warranted by significant argument.

Kit Fine (1977), Christopher Peacocke (1978), and Graeme Forbes (1989) suggested a modalist approach that rejects the call for a reductive theory of modality in nonmodal terms. If anything, the explanation goes in the opposite direction: something is true in a possible world iff it is possibly true. For the modalist, reality is irreducibly modal and this is exhibited by the attempts to translate the whole of the possible worlds theory into a modal language, expanding the basic modal language to include an actuality operator as well as indices for the operators to permit the tracking of modal contexts. For example, if one permits oneself w_1, to be a variable ranging over worlds, w^* to stand for the actual world, and E to be a two-place predicate by which one can express that an object exists in a world, then the possible worlds translation of "There could have been more things than there actually are" is:

$$\exists w_1[\forall w \forall x(Exw^* \rightarrow Exw_1) \ \& \ \exists y(\neg Eyw^*).$$

Where $\Diamond_1$ permits one to express a given possibility, A expresses actuality, and A_1 expresses what is actual in a specific possibility, the modalist translation for this is:

$$\Diamond_1\{[\Box \forall x(AEx \rightarrow A_1Ex)] \ \& \ \exists y \neg AEy\}.$$

Melia (1992) argued that the modalist translation is not a reduction of possible worlds discourse at all, but merely a notational variant of the possible worlds statement. Even if it is granted that one has a firm grasp on the modalist's basic modal and actuality operators, once the subscripts are added and one operator is placed within the scope of others, one has no intuitive grasp of their meanings in those contexts. The only way to understand them, indeed the way the modalist explains them, is by reference to the possible worlds semantics. Contrary to the modalist claims, this makes it appear as though possible worlds discourse, not the modalist's, is semantically basic and more perspicuous.

Some assumed that modalism is to be recommended only if it can reduce possible worlds discourse in modal terms. However, why, exactly, should the modalist provide translations of all possible worlds claims? What must be determined is what, if any, possible worlds claims are merely artifacts of the possible worlds framework. It is no reason to give up modalism if it cannot accommodate mere artifacts of the possible worlds framework that is, ultimately, rejected as a literal account of modal metaphysics. For instance, according to Lewis's (1968) developments, each world is as it is. It is not essential that a world be that way, but it is essential that that specific world be that particular way. Being that way is precisely what distinguishes that world from all others. Modalists are not bound to make this essentialist claim part of their theory. So long as the modalist can say all that one has either theory-neutral or modalist grounds for asserting, the failure to translate all the modal realist's claims should not count against modalism.

Though modalism is not wedded to essentialism, Fine (1977) argued not only that reality is irreducibly modal but also that *de re* modality is more basic than *de dicto*, defending the most general aspects of essentialism defended by Kripke (1972/1980) and Hilary Putnam (1975). These works brought essentialism back into philosophical discussion among analytic philosophers. Each was concerned with the semantics for proper names and natural kinds terms. Once necessity, analyticity, and *a priority* were clearly separated from each other, some essentialist theses—such as that the origins/genealogy are essential to some objects and that substances have their chemical constitution essentially—became more plausible. Fine extended this so that *de dicto* modality, concerned as it is with the necessary truth and falsehood of some propositions, is explained by *de re* modal facts about the natures of truth bearers or concepts or logical functions. This provides essentialism with an explanatory role so that if objects, whether concrete or abstract, have properties without which they would not be those very objects, then modal truth is on a par with nonmodal truth. Truth, whether modal or nonmodal, depends on being. The modalist simply maintains that being is irreducibly modal.

MODALITY AND METAPHYSICS

In the end, the philosophy and metaphysics of modality rests on metaphilosophical foundations. Many of the objections to the various positions have been piecemeal, showing that a theory has some consequence that is supposed to be intolerable. Lewis (1986) made quite clear that the case for genuine modal realism was a philosophical inference to the best explanation, not a single silver bullet–like argument. He claimed that when all things were considered his theory possessed the best balance of theoretical virtues and vices. Other theories might rely on less controversial ontologies or they may avoid some other counterintuitive consequences of modal realism.

Nevertheless, when all things are taken into account, Lewis thinks that his theory is the best package. Those willing to engage Lewis on his own terms must provide comparable details about the relative merits and difficulties of an alternative to properly undercut the warrant that Lewis thinks that he has given for modal realism.

An alternative is to question the appropriateness of inference to the best explanation in metaphysical contexts. That argument form is typically associated with contexts in which prior experience showed that one kind of event or fact—the activity of mice—explained another—the disappearance of cheese. When one confronts another instance of missing cheese and one has been unable to observe rodents, the inference to the activity of mice might well be appropriate. In metaphysics there is no analogue to prior experience. If the legitimacy of an argument form is not knowable *a priori*, some *a posteriori* basis is needed for thinking that the argument is appropriate to a given context of application. One knows that statistical inferences are appropriate under some conditions and not others because of what one knows from empirical investigation of the world. In the absence of some general reason to think that a metaphysical theory is more likely to be true when it is the conclusion of an inference to the best explanation, the application of an inference form may be warranted in some empirical contexts but unwarranted in metaphysical contexts. Thus, warrant for a specific theory of modality depends on deeper considerations about forms of argument appropriate to metaphysics.

See also Metaphysics; Modality and Language.

Bibliography

Adams, Robert Merrihew. "Primitive Thisness and Primitive Identity." *Journal of Philosophy* 76 (1) (1979): 5–26.

Adams, Robert Merrihew. "Theories of Actuality." *Noûs* 8 (1974): 211–231.

Armstrong, D. M. *A Combinatorial Theory of Possibility*. New York: Cambridge University Press, 1989.

Ayer, Alfred J. *Language, Truth, and Logic*. London: V. Gollancz, 1936.

Blackburn, Simon. "Morals and Modals." In *Fact, Science, and Morality*, edited by Graham Macdonald. Oxford, U.K.: Blackwell, 1986.

Brock, Stuart. "Modal Fictionalism: A Response to Rosen." *Mind* 102 (1993): 147–150.

Carnap, Rudolf. *The Logical Syntax of Language*. London: Routledge, 1954.

Chihara, Charles. *The Worlds of Possibility: Modal Realism and the Semantics of Modal Logic*. Oxford, U.K.: Clarendon Press, 1998.

Copeland, B. J., ed. *Logic and Reality: Essays on the Legacy of Arthur Prior*. Oxford, U.K.: Clarendon Press, 1996.

Divers, John. *Possible Worlds*. London: Routledge, 2002.

Divers, John, and Joseph Melia. "The Analytic Limits of Genuine Modal Realism." *Mind* 111 (2002): 15–36.

Fine, Kit. "Essence and Modality." In *Philosophical Perspectives. Vol. 8, Logic and Language*. Atascadero, CA: Ridgeview, 1994.

Fine. Kit. "Plantinga on the Reduction of Possibilist Discourse." In *Alvin Planting*, edited by James Tomberlin and Peter van Inwagen. Dordrecht, Netherlands: D. Reidel, 1985.

Fine, Kit. Postscript to *Worlds, Times, and Selves*. London: Duckworth, 1977.

Forbes, Graeme. *The Languages of Possibility*. Oxford, U.K.: Basil Blackwell, 1989.

Forbes, Graeme. *The Metaphysics of Modality*. Oxford, U.K.: Clarendon Press, 1985.

Forrest, Peter, and D. M. Armstrong. "An Argument against David Lewis' Theory of Possible Worlds." *Australasian Journal of Philosophy* 62 (1984): 164–168.

Gendler, Tamar-Szabo, and John Hawthorne, eds. *Conceivability and Possibility*. New York: Oxford University Press. 2002.

Hume, David. *A Treatise of Human Nature*. Bks. I and II. 2 vols. London, 1739. Bk. III, London, 1740. Modern editions are by L. A. Selby Bigge (Oxford, 1888 and 1896) and by A. D. Lindsay (London and New York, 1911, 2 vols).

Kripke, Saul A. "A Completeness Theorem in Modal Logic." *Journal of Symbolic Logic* 24 (1959): 1–14.

Kripke, Saul A. "Naming and Necessity." In *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman. Dordrecht, Netherlands: D. Reidel, 1972. Published with additions as *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Kripke, Saul A. "Semantic Analysis of Modal Logic." *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9 (1963): 67–96.

Kripke, Saul A. "Semantic Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94.

Lewis, Clarence Irving. *A Survey of Symbolic Logic*. Berkeley: University of California Press, 1918.

Lewis, Clarence Irving, and Cooper Harold Langford. *Symbolic Logic*. New York: Century, 1932.

Lewis, David. "Anselm and Actuality." *Noûs* 4 (1970): 175–188.

Lewis, David. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

Lewis, David. "Counterpart Theory and Quantified Modal Logic." *Journal of Philosophy* 65 (1968): 113–126.

Lewis, David. *On the Plurality of Worlds*. Oxford, U.K.: Blackwell, 1986.

Linsky, Leonard, ed. *Reference and Modality*. London: Oxford University Press, 1971.

Loux, Michael J., ed. *The Possible and the Actual*. Ithaca, NY: Cornell University Press, 1979.

Lycan, William G. "Armstrong's New Combinatorialist Theory of Modality." In *Ontology, Causality, and Mind: Essays in Honour of D. M. Armstrong*, edited by John Bacon, Keith Campbell, and Lloyd Reinhardt. New York: Cambridge University Press, 1993.

Lycan, William G. *Modality and Meaning*. Dordrecht, Netherlands: Kluwer Academic, 1994.

Marcus, Ruth Barcan, *Modalities: Philosophical Essays*. New York: Oxford University Press, 1993.

McGinn, Colin. "Modal Reality." In *Reduction, Time, and Reality*. New York: Cambridge University Press, 1981.

Melia, Joseph. "Against Modalism." *Philosophical Studies* 68 (1992): 35–56.

Pap, Arthur. *Semantics and Necessary Truth*. New Haven, CT: Yale University Press, 1958.

Peacocke, Christopher. "Necessity and Truth Theories." *Journal of Philosophical Logic* 7 (1978): 473–500.

Plantinga, Alvin. *The Nature of Necessity*. Oxford, U.K.: Clarendon Press, 1974.

Plantinga, Alvin. "Two Concepts of Modality: Modal Realism and Modal Reductionism." In *Philosophical Perspectives. Vol. 1, Metaphysics*, edited by James E. Tomberlin, 189–231. Atascadero, CA: Ridgeview, 1987.

Plantinga, Alvin. *Essays in the Metaphysics of Modality*. New York: Oxford University Press, 2003.

Prior, Arthur, and Kit Fine. *Worlds, Times, and Selves*. Amherst: University of Massachusetts Press, 1977.

Putnam, Hilary. "The Meaning of 'Meaning.'" *Minnesota Studies in the Philosophy of Science* 7 (1975): 131–193.

Quine, W. V. O. *From a Logical Point of View*. New York: Harper and Row, 1953.

Quine, W. V. O. "Truth by Convention." In *Philosophical Essays for A. N. Whitehead*, edited by O. H. Lee. New York: Longmans, 1948.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43.

Quine, W. V. O. *The Ways of Paradox*. New York: Random House, 1966.

Rosen, Gideon. "Modal Fictionalism." *Mind* 99 (1990): 327–354.

Rosen, Gideon. "A Problem for Fictionalism about Possible Worlds." *Analysis* 53 (1993): 71–81.

Shalkowski, Scott A. "Logic and Absolute Necessity." *Journal of Philosophy* 101 (2004): 55–82.

Shalkowski, Scott A. "The Ontological Ground of the Alethic Modality." *Philosophical Review* 103 (1994): 669–688.

Sidelle, Alan. *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Ithaca, NY: Cornell University Press, 1989.

Stalnaker, Robert. "Possible Worlds." *Noûs* 10 (1976): 65–75.

Van Inwagen, Peter. *Ontology, Identity, and Modality*. New York: Cambridge University Press, 2001.

Van Inwagen, Peter. "Two Concepts of Possible Worlds." *Midwest Studies in Philosophy* 11 (1986): 185–213.

Wright, Crispin. "Necessity, Caution, and Scepticism." *Proceedings of the Aristotelian Society Supplement* 63 (1989): 203–238.

Wright, Crispin. *Wittgenstein on the Foundations of Mathematics*. Cambridge, MA: Harvard University Press, 1980.

Scott A. Shalkowski (2005)

MODALITY AND LANGUAGE

See Appendix, Vol. 10

MODALITY AND QUANTIFICATION

Quantified modal logics combine quantifiers (∀ for *all*, and ∃, for *some*) with an intensional operator □ (for such expressions as 'necessarily' and 'Ralph believes that'). Quantifying into intensional contexts (or quantifying in, for short) occurs when a quantifier binds an open variable that lies within the scope of □, as in sentences with the form ∃□Fx. Systems of quantified modal logic (QML) routinely include formulas of this kind, but Willard Van Orman Quine (1963) famously argues that quantifying in is incoherent.

Here is a quick summary of his main line of reasoning. Consider (1)–(3), an apparent counterexample to the law of substitution for identity:

(1) 9 equals the number of planets

(2) Necessarily 9 is greater than 7

(3) Necessarily the number of planets is greater than 7

Although (3) is the result of the substituting 'the number of planets' for '9' in (2), and both (1) and (2) are true, (3) is presumably false. Quine calls term positions where substitution fails opaque contexts and argues that terms occupying them do not play their normal referring roles. Both '9' and 'the number of planets' refer to nine, so something other than the terms' referents must explain why the truth values of (2) and (3) differ. Presumably, the difference is in the manner of referring to or describing nine. Now note that the standard truth condition for ∃ says that (4) is true if and only if (iff) the open sentence (5) is true of some object:

(4) ∃x(necessarily, x is greater than 7)

(5) Necessarily, x is greater than 7

However, (5) results from putting 'x' for either '9' in (2) or 'the number of planets' in (3), and (2) and (3) were sensitive to the manner in which nine is described. Since 'x' does not describe anything at all, information needed to make sense of (5) being true of an object is now missing. As Quine puts it, what object is (5) true of? Presumably, it is nine, that is, the number of planets. However, the

number of planets appears not to satisfy (5), since (3) was false.

Arthur F. Smullyan (1948) was one of the first to respond to Quine's argument. He notes that when 'the number of planets' is translated away according to Bertrand Russell's theory of definite descriptions, (1)–(3) does not constitute a violation of the law of substitution (LS). On the analysis that matches the intuition that (3) is false, it is not possible to derive the translation of (3) from (2) and the translation of (1) in predicate logic, even given LS. If one adopts the position that any purported failure of substitution for an expression is a good reason to treat it as a definite description, then there are no terms in opaque contexts in the first place, and Quine's reasoning does not get off the ground. However, this solution, Quine notes, is limited to those cases where Russell's technique can be plausibly applied.

Alonzo Church (1943) and Rudolf Carnap (1947) propose a different tactic. Presuming that variables of quantification range over concepts rather than objects, Quine's complaint that satisfaction of (5) by an object is unintelligible does not apply. However, Quine finds quantification over concepts ontologically disreputable; and furthermore, citing an alternative treatment of quantification would not rebut an argument concluding the incoherence of quantifying in for quantification over objects, a result damaging enough to QML.

There are a number of different strategies for responding to Quine's objection in the case of quantifying over objects. One popular tactic, exemplified in David Kaplan's "Quantifying In" (1969), involves selecting a privileged class of terms (for Kaplan, the so-called vivid names). Although the truth values of (3) and (2) are sensitive to the ways nine is described, one argues that there is no corresponding indeterminacy in (5) because one of these ways is privileged. Presuming '9' is privileged, (2), and not (3), is used to resolve the status of (5). Since (2) is true, (5) is true of nine, and the fact that that (3) is false is irrelevant.

In note 3 of "Quantifying In" (1969) Kaplan suggests another way to circumvent Quine's objections to (5) without using privileged terms. The idea is (roughly) to revise the truth condition for ∃ so that ∃x(necessarily, x is greater than 7) is T iff some object satisfies the open sentence (6):

(6) x bears the property of being necessarily greater than seven

Since 'x' in (6) lies outside the scope of 'necessarily', substitution holds in this position, and Quine's worries no longer apply. (Something like this tactic is used by Quine himself in "Quantifiers and Propositional Attitudes" [1955] to analyze quantification into belief contexts.)

Kaplan's (1969) strategy is reflected in a solution implicit in the earliest published QML. The system (developed by Ruth Barcan Marcus [1946]) includes the axiom $\forall x \forall y(x = y \rightarrow \Box x = y)$, which is now known to correspond to the condition that variables are rigid designators, that is, they pick out the same object in every possible world. Under these circumstances (5) is equivalent to (6), and so (6) can be used to make sense of (5).

Kit Fine's "The Problem of De Re Modality" (1989) makes yet another contribution to the problem. Here a formal definition of satisfaction by objects for open sentences like (5) is provided in cases where 'necessarily' indicates logical or analytic necessity.

In "A Backward Look at Quine's Animadversions on Modalities" (1990) Marcus records how the force and variety of such criticisms of Quine's argument led him to a strategic retreat. He conceded that quantifying in is at least coherent, but raised a different objection. Quine perceived early on that attacks on his argument appear to pay a serious price. Appeals to privileged ways of describing things, to rigid designators, or to the cogency of (6) boil down to having to make sense of the idea that some objects bear necessary properties that other objects do not. Quine complains that this amounts to an unacceptable form of essentialism. What sense can it make to assert of an object itself (apart from any way of describing it) that it has necessary properties?

An influential response to this worry appears in the early pages of "Naming and Necessity" (1972), where Saul Kripke undermines Quine's presumption that it only makes sense to attribute necessary properties to an object under a description. Here, the focus shifts from brands of logical or analytical necessity, which were the main concern when Quine first wrote, to metaphysical or physical necessity. Kripke defends the view that objects in themselves do have essential properties. For example, molecules of water are necessarily composed of hydrogen and oxygen, because water just is H_2O.

Kripke and others rescued some brands of essentialism from the negative reputation it had when Quine first wrote. However, one need not respond to Quine by arguing for the coherence of a robust essentialism. In "Opacity" (1986) Kaplan argues that the essentialism produced by quantifying in is so weak as to be entirely innocuous. Terence Parsons, in "Essentialism and Quantified Modal

Logic" (1969), reports the technical result that sentences of QML that express a controversial essentialism will not be theorems, nor will they be derivable from any collection of premises expressing (nonmodal) facts.

Parsons (1969) and others point out that while quantifying in allows one to assert essentialist claims, this hardly qualifies as a reason for abandoning it. QML should provide an impartial framework for analyzing and evaluating argumentation on all philosophical positions, however misguided. That quantifying in provides resources to express (even the most obnoxious) essentialism is a point in its favor. In any case, Quine's complaint that QML is essentially essentialist amounts to a retraction of the view that quantifying in is (literally) incoherent, for if that were true, quantifying in would not entail essentialism, it would express nothing at all.

It is important to note that Quine's main argument against quantifying in would appear to apply equally well to expressions for propositional attitudes such as "Ralph believes that," for these also create opaque contexts. However, in the case of belief, the situation is different, since charges of essentialism are out of place. In "Intensions Revisited" (1981) Quine explores failings for belief that are analogs to essentialism for necessity.

Despite attacks on Quine's main argument, many still accept the conclusion that quantifying in is incoherent. Graeme Forbes (1996) notes that adherents of this view face a new puzzle, posed by strong intuitions in favor of the intelligibility of English sentences like those represented by ∃x(Ralph believes that x is a spy). So those adherents need an alternative analysis of the logical form of propositional attitude sentences that avoids quantifying in, one Forbes sets out to provide. A tension Quine faces here is that explanations placating intuitions that quantifying in is coherent for belief will provide tools that resolve his worries about necessity.

QML has come a long way in the sixty years since Quine first launched his attack on it. Possible worlds semantics has flourished, bringing a wealth of technical results. For example, soundness and completeness have been proven for a variety of systems that allow quantifying in but reject LS in modal contexts. Theorems are also available on exactly how and where essentialist features arise in QML (e.g., see Fine 1978, 1981). Though work in modal semantics employs ideas that are anathema to Quine (notably the notion of a possible object), it provides tools for better understanding worries about quantifying in. An interest in answering Quine's objections to QML has motivated many of these developments. So, oddly, Quine's legacy has enriched what he hoped to disinherit.

See also Modal Logic.

Bibliography

QUINE'S SEMINAL ARTICLES ON QUANTIFYING IN

"Notes on Existence and Necessity." *Journal of Philosophy* 40 (1943): 113–149.

"Quantifiers and Propositional Attitudes." In *Reference and Modality*, edited by Leonard Linsky, 101–111. New York: Oxford University Press, 1955.

"Reference and Modality." In *From a Logical Point of View*. New York: Harper and Row, 1963.

"Three Grades of Modal Involvement." In *The Ways of Paradox and Other Essays*. Cambridge, MA: Harvard University Press, 1966.

SOME OTHER IMPORTANT WORKS ON QUANTIFYING IN

Barcan, Ruth. "A Functional Calculus of First Order Based on Strict Implication." *Journal of Symbolic Logic* 2 (1946): 1–16.

Carnap, Rudolf. *Meaning and Necessity*. Chicago: University of Chicago Press, 1947.

Church, Alonzo. "Review of Quine." *Journal of Symbolic Logic* 8 (1943): 45–52.

Fine, Kit. "Model Theory for Modal Logic." *Journal of Philosophical Logic* 7 (1978): 125–156, 277–306; 10 (1981): 293–307.

Fine, Kit. "The Problem of De Re Modality." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press, New York, 1989.

Forbes, Graeme. "Substitutivity and the Coherence of Quantifying In." *Philosophical Review* 105 (1996): 337–372

Kaplan, David. "Opacity." In *The Philosophy of W. V. Quine*, edited by Lewis Edwin Hahn and Paul Arthur Schilpp, 229–289. LaSalle, IL: Open Court, 1986.

Kaplan, David. "Quantifying In." In *Words and Objections*, edited by Donald Davidson and Jaakko Hintikka. Dordrecht, Netherlands: D. Reidel, 1969.

Kripke, Saul. "Naming and Necessity." In *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman, 253–355. Dordrecht, Netherlands: D. Reidel, 1972.

Marcus, Ruth Barcan. "Essentialism in Modal Logic." *Noûs* 1 (1967): 91–96.

Marcus, Ruth Barcan. "A Backward Look at Quine's Animadversions on Modalities." In *Perspectives on Quine*, edited by Robert B. Barrett and Roger F. Gibson. Cambridge, MA: Blackwell, 1990.

Parsons, Terence. "Essentialism and Quantified Modal Logic." *Philosophical Review* 78 (1969): 35–52.

Quine, W. V. O. "Intensions Revisited." In *Theories and Things*. Cambridge, MA: Harvard University Press, 1981.

Smullyan, Arthur F. "Modality and Description." *Journal of Symbolic Logic* 13 (1948): 31–37.

James W. Garson (2005)

MODAL LOGIC

Traditionally, the modes implicit in modal logic are the modes of truth and ultimately the modes of being: necessary, possible, impossible, and contingent. While the study of the formal properties of those notions is still an important part of modal logic, other interpretations have been added over the years, such as temporal, epistemic, and deontic. Furthermore, more recently, other formal languages have been suggested, which, although not modal logic in a strict sense, are closely related to it, such as dynamic logic.

BRIEF HISTORY

Modern modal logic began in 1912 when Clarence Irving Lewis published a paper in *Mind*, in which he recommended that the logic of *Principia Mathematica* be supplemented with what he called intensional connectives. Among the latter was a binary connective of strict implication for which he introduced a new symbol, a "fishhook" to distinguish it from the "horseshoe" of the material conditional. Thus, $\phi \multimap \psi$ and $\phi \supset \psi$ would both be read "if ϕ then ψ," but Lewis specifically intended for the former to model the elusive notion of entailment. Other connectives were possibility, for which he used the symbol $\diamond$ (a diamond), and necessity, for which F. B. Fitch would later suggest $\Box$ (a box): thus, $\diamond\phi$ and $\Box\phi$ were read "it is possible that ϕ" and "it is necessary that ϕ," respectively. The interest in strict implication declined somewhat after it was discovered that there are paradoxes of strict implication in parallel with those of material implication in classical logic. Lewis's legacy was not lost, however. On the one hand, philosophers like Alan Ross Anderson and Nuel Belnap went on to develop logics of entailment and relevance, a tradition that has proved hardy. On the other hand, since necessity and possibility seem more interesting than strict implication—$\phi \multimap \psi$ is in any case analyzable as either $\Box(\phi \supset \psi)$ or $\neg \diamond (\phi \wedge \neg\psi)$—later logicians preferred to do their modal logic in terms of those concepts.

Lewis's original ambition was to find the logic of strict implication. Much to his surprise he later found himself confronted by a veritable embarrassment of riches: an ever increasing number of modal logics—not only his own famous quintuple of systems S1, S2, S3, S4, and S5 (his own tentative favorite was S3, the so-called Survey system) but also many, in fact, infinitely many others. Since he never translated his semantic intuitions into a formal structure, the differentiation between different proposals became a problem. Some help in this regard arrived in the form of the concept of a matrix, essentially a set of truth-values (usually but not necessarily finitely many) plus a truth-value table for each connective. This idea, which was due to Jan Łukasiewicz, was then generalized into the notion of an algebra (essentially a set with operators) and taken into modal logic by Alfred Tarski and his collaborators. The advent of algebraic logic revitalized modal logic. Two works from this period are particularly noteworthy. One was the first formal result in modal logic worth the name, J.C.C. McKinsey's algebraic characterization of S2 and a proof that it is decidable. The other was a paper in 1951 by Bjarni Jónsson and Tarski foreshadowing the next major development: the era of possible-worlds semantics.

Since the term *possible-worlds semantics* is today used pretty much synonymously with the term *Kripke semantics*, it is germane to ask: Who invented Kripke semantics? In fact, this question has been the object of much discussion, some heated. When the new semantics emerged at the end of the 1950s, Rudolf Carnap had laid the ground work; his states really played the role that possible worlds would later play, even if he only worked with descriptions of them. What Carnap did not have, and which turned out to make all the difference, was the accessibility relation (this concept is explained later on). The accessibility relation did appear in the Jónsson-Tarski paper mentioned earlier, and it now seems likely that Arthur Prior and C. A. Meredith had also discovered it in the early 1950s. But it was Saul A. Kripke, along with Stig Kanger and Jaakko Hintikka, who first published accounts in which the accessibility relation was a central concept and its versatility recognized. That Kripke's work overshadowed the work by Kanger and Hintikka and proved so much more influential than theirs is perhaps not surprising, given the clarity and mathematical maturity Kripke's papers and the systematic development of his theory.

After Kripke's early work followed a period of increasingly formal concern. Not surprisingly, the philosophers have focused on the philosophy of modal logic, including modal metaphysics, while the mathematicians have pursued the mathematics of modal logic, including model theory, algebra, and even category theory. Another significant development has been the expanding use of modal logic in theoretical computer science: with energy and inventiveness—but of course guided by their own interests—computer scientists have, within a short time, transformed modal logic.

SYNTAX

This entry only considers classical modal logic, that is, logic that extends classical logic. Historically, even though for a long time it is modal predicate logic that has been of particular interest to philosophers, propositional modal logic has received much more attention from formal logicians, probably because agreement on what constitutes a generally accepted conceptual framework for research was reached much earlier in the latter area.

PROPOSITIONAL LOGIC. To the set of the usual truth-functional connectives, add two new connectives: a box operator $\Box$ and a diamond operator $\Diamond$. After Tarski, a theory, in a technical sense, is a set of formulas (called theses of the logic) that contains all classical two-valued tautologies and is closed under *modus ponens* (if ϕ and $\phi \supset \psi$ are theses of the logic, then so is ψ). Similarly, a logic, in the technical sense used here, is a theory that is closed under uniform substitution (if $\phi(\chi/P)$ results from a formula ϕ by replacing all occurrences of a certain propositional letter P with a formula χ, then $\phi(\chi/P)$ is a thesis of the logic if ϕ is). A normal modal logic is a logic that contains as theses all instances of the schema $\Diamond\phi \equiv \neg\Box\neg\phi$ as well as of the so-called Kripke schema $\Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi)$ and, in addition, is closed under the rule of necessitation (if ϕ is a thesis, then so is $\Box\phi$). A great number of normal modal logics have been studied, many of them definable in terms of further schemata, for example,

(D) $\Box\phi \supset \Diamond\phi$,

(T) $\Box\phi \supset \phi$,

(4) $\Box\phi \supset \Box\Box\phi$,

(5) $\neg\Box\phi \supset \Box\neg\Box\phi$,

(G) $\Diamond\Box\phi \supset \Box\Diamond\phi$,

(H) $(\Diamond\phi \wedge \Diamond\psi) \supset (\Diamond(\phi \wedge \psi) \vee \Diamond(\phi \wedge \Diamond\psi) \vee \Diamond(\psi \wedge \Diamond\phi))$,

(W) $\Box(\Box\phi \supset \phi) \supset \Box\phi$.

To bring some order into the bewildering multiplicity of modal logics, E. J. Lemmon suggested $KX_1, \ldots, X_n$ as a code name for the smallest normal modal logic that contains all substitution instances of schemata $X_1, \ldots, X_n$. In this notation, one may identify K as the smallest normal logic, KT as the Gödel/Feys/von Wright logic, and KT4 and KT45 as the logics S4 and S5, respectively. The logics KD, KD4, and KD45, of special interest to deontic and doxastic logic, are sometimes called weak T, weak S4, and weak S5, respectively. The logics KT4G and KT4H are better known as S4.2 and S4.3, respectively, and the logic K4W as the Gödel/Löb logic GL. The set of all normal logics, ordered by set inclusion, forms a lattice of immense complexity, as do sets of more inclusive classes of nonnormal modal logics. The efforts to explore these structures continue but are increasingly a concern for mathematicians rather than for philosophers.

PREDICATE LOGIC. Modal predicate logic does not exhibit the relative orderliness or maturity of its propositional relative. Philosophical questions such as the proper treatment of individuals persist. Quantification, in particular into opaque contexts—that is, contexts within the scope of modal operators—has been a main problem, as evidenced by Quine's unrelenting criticism over a lifetime. A formal beginning was made by Ruth Barcan Marcus, after whom two central formulas have been named—the Barcan formula (BF) and the converse Barcan formula (CBF):

(BF) $\forall x\Box\phi \supset \Box\forall x\phi$,

(CBF) $\Box\forall x\phi \supset \forall x\Box\phi$

Other examples of formulas that were much discussed in early literature are

$$\forall x\forall y(x = y \supset \Box(x = y)),$$

$$\forall x\forall y(x \neq y \supset \Box(x \neq y)),$$

$$\mathrm{a} = \mathrm{b} \supset \Box(\mathrm{a} = \mathrm{b}),$$

$$\mathrm{a} \neq \mathrm{b} \supset \Box(\mathrm{a} \neq \mathrm{b}),$$

where a and b are individual constants. Various authors have held different views on which of these, if any, are valid. It would seem that to take a stand in such matters is to rely on implicit semantic ideas, however sketchy. It was accordingly an important step when at last, thanks to Kripke and others, formal semantics were articulated.

SEMANTICS

The development of modal logic, both material and formal, preceded in steps. Propositional logics were studied extensively before predicate logicians had been able to work out a generally accepted common ground. Till this day, the area of modal propositional logic is more definitive than the relatively more unsettled area of modal predicate logic.

PROPOSITIONAL LOGIC. The possible-worlds semantics, introduced by Kripke in the early 1960s, may be cast in the following form (which differs from Kripke's original formulation in terminology and, to some extent, in

substance). A frame is a pair (U, R), where U, the universe of discourse or simply the universe of the frame, is a nonempty set of elements that are often called possible worlds but that may more neutrally be called points, and R is a binary relation in U, called the accessibility relation or sometimes the alternativeness relation or even the alternative relation. If two points u, v of U are related by R (i.e., if $(u, v) \in R$), then one says that v is accessible from u or that v is an alternative to u. A valuation in U is a function V assigning to each propositional letter P a subset $V(\mathrm{P})$ of U. A model is a structure (U, R, V) where (U, R) is a frame and V is a valuation in U. Truth in modal logic is doubly relative: to a model and to a point in the model. Thus, if $\mathfrak{M} = (U, R, V)$ is a model, u a point in U and ϕ a formula, one may inductively define the notion of ϕ being true at u in $\mathfrak{M}$, schematically $u \vDash^{\mathfrak{M}} \phi$, as follows:

$$u \vDash^{\mathfrak{M}} \mathrm{P} \text{ iff } u \in V(\mathrm{P}), \text{ if P is a propositional letter;}$$

$$u \vDash^{\mathfrak{M}} \neg\phi \text{ iff not } u \vDash^{\mathfrak{M}} \phi,$$

$$u \vDash^{\mathfrak{M}} \phi \wedge \psi \text{ iff } u \vDash^{\mathfrak{M}} \phi \text{ and } u \vDash^{\mathfrak{M}} \psi,$$

$$u \vDash^{\mathfrak{M}} \phi \vee \psi \text{ iff } u \vDash^{\mathfrak{M}} \phi \text{ or } u \vDash^{\mathfrak{M}} \psi,$$

and similar conditions for other truth-functional connectives:

$$u \vDash^{\mathfrak{M}} \Box\phi \text{ iff, for all points } v, \text{ if } (u,v) \in R \text{ then } v \vDash^{\mathfrak{M}} \phi,$$

$$u \vDash^{\mathfrak{M}} \Diamond\phi \text{ iff there is some point } v \text{ such that } (u, v) \in R \text{ and } v \vDash^{\mathfrak{M}} \phi.$$

(Readers may note the roles played in this definition by R and V: The latter is needed to get the definition started, the former to evaluate formulas beginning with a modal operator; the truth-functional connectives are taken care of by the usual truth tables.) A formula is valid in a frame if it is true at every point in every model definable on that frame; and it is valid in a class of frames if it is valid in each one of the frames of the class.

There is a sense in which this semantics fits modal logic. The set of formulas that are valid in a given class of frames will always be a normal modal logic and can be called the logic determined by that class of frames. A logic is sound with respect to a class of frames if every thesis of the logic is valid in that class, and it is complete with respect to the class if every formula that is valid in the class is a thesis of the logic; hence a logic is determined by a class of frames if and only if it is both sound and complete with respect to that class. It is an interesting fact, and no doubt one reason for the popularity of Kripke semantics, that many of the logics defined in the philosophical literature are determined by simply defined classes of frames. For example, T, S4, and S5 are determined by the class of frames whose accessibility relations are reflexive, reflexive and transitive, and reflexive, symmetric, and transitive, respectively. Similarly, KD, KD4, and KD45 are determined by the class of all frames whose accessibility relations are serial, serial and transitive, and serial, transitive and euclidean, respectively. (A binary relation R is serial if, for every element u in its field there is some element v, not necessarily distinct from u, such that $(u, v) \in R$, euclidean if, for all elements u, v, w in its field, if $(u, v) \in R$ and $(u, w) \in R$ then $(v, w) \in R$.) At the extremes are the smallest normal modal logic K and the inconsistent logic, which are determined by, respectively, the class of all frames and the empty class of frames.

The way in which Kripke's semantics seems to fit modal logic led some authors, for example, Lemmon, to conjecture that all normal modal logics are complete, that is, determined by some class of frames. However, that the fit is less than perfect was proved in 1971 by Kit Fine and S. K. Thomason, who exhibited, independently of one another, instances of incomplete normal modal logics.

PREDICATE LOGIC. Among several possible versions of semantics for modal predicate logic, the following is essentially a modified version of Kripke's semantics for first-order modal logic from 1963. For simplicity, assume a formal language for predicate logic containing predicate letters and individual constants (but, for example, no descriptions or functional operators); thus, the terms of this language are individual variables or individual constants. To generalize the central concepts *frame* and *model* used in propositional modal logic, several new notions must be introduced. To begin with, besides a universe U of points (possible worlds) and an accessibility relation R, as before, one needs a nonempty set D of objects and a function E defined on U that takes values in the set of subsets of D. One can refer to D as the domain and to E as the existence function, to the elements of D as possible individuals and to the elements of E_u as individuals existing at u or individuals actual at u (where u is a point in U). Altogether, a structure (U, R, D, E), where U, R, D, E are as specified, is a frame. Next, one can say that I is an interpretation (in D with respect to U) if it is a family of functions I_u, where u ranges over U, such that I_u assigns a set of n-tuples of elements of D to each n-ary predicate letter and an element of D to each individual constant. If $\mathfrak{F} = (U,R,D,E)$ is a frame, then $\mathfrak{M} = (\mathfrak{F},I) = (U,R,D,E,I)$ is a model (on $\mathfrak{F}$) if I is an interpretation in D with respect to U.

The following observation shows the sense in which the present concept of model is a generalization over that of propositional semantics: Nullary predicate letters behave in the present setting as propositional letters do in the propositional case. To see this, let P be a nullary predicate letter. By the definition, the interpretation of P is a set of 0-tuples, hence $I_u(\mathrm{P})$ is either $\emptyset$ (the empty set) or $\{\emptyset\}$ (the singleton set whose only member is the one and only 0-tuple). If one arbitrarily identifies $\{\emptyset\}$ with truth and $\emptyset$ with falsity, one thereby also in effect identifies the set $\{u \in U : I_u(\mathrm{P}) = \{\emptyset\}\}$ with the proposition expressed by P in $\mathfrak{M}$. Thus, the interpretation plays a role in the predicate case similar to that of the valuation in the propositional case, albeit a much bigger role.

Besides all this, one needs yet another concept to define truth-conditions: something to take care of the quantifiers. An assignment (in a set D) is a function from the set of individual variables of one's formal language to D. Notice that if A is an assignment in D and x is a variable, then $A(x)$ is an element of D but perhaps not of E_u, if u is an arbitrary point in U. If $\mathfrak{M} = (U, R, D, E, I)$ is a model and A is an assignment in D, then the denotation of t in $\mathfrak{M}$ under A is a function $\|t\|^{\mathfrak{M}}_{A}$ defined on U as follows:

$$\|t\|^{\mathfrak{M}}_{A}(u) = \begin{cases} I_u(t), & \text{if } t \text{ is an individual constant,} \\ A(t), & \text{if } t \text{ is an individual variable.} \end{cases}$$

The truth of a formula ϕ in a model $\mathfrak{M}$ under an assignment A at a point u, in symbols $u \models^{\mathfrak{M}}_{A} \phi$, may now be defined:

$$u \models^{\mathfrak{M}}_{A} \mathrm{P}(t_0,\ldots,t_{n-1}) \text{ iff } (\|t_0\|^{\mathfrak{M}}_{u},\ldots,\|t_{n-1}\|^{\mathfrak{M}}_{u}) \in I_u(\mathrm{P}),$$
if P is an n-ary predicate letter,

$$u \models^{\mathfrak{M}}_{A} \forall x\, \phi \text{ iff } u \models^{\mathfrak{M}}_{B} \phi, \text{ for all assignments } B \text{ such that } B(x) \in E_u \text{ and, for all variables } y, \text{ if } x \neq y \text{ then } A(y) = B(y).$$

The remaining clauses of the definition (for the truth-functional connectives and the modal operators) are as before. In particular,

$$u \models^{\mathfrak{M}}_{A} \Box\phi \text{ iff, for all } v, \text{ if } (u,v) \in R \text{ then } v \models^{\mathfrak{M}}_{A} \phi.$$

As in the propositional case, one associates truth with models and validity with frames. Thus, one can say that a formula is true in a model if it is true under all assignments at all points in the model. By the same token, one can say that a formula is valid in a frame if it is true in all models on the frame.

Some object languages contain constant predicates besides predicate letters. Common examples of such predicates are the unary E (the existence predicate) and the binary = (the identity predicate) with corresponding truth-conditions:

$$u \models^{\mathfrak{M}}_{A} \mathrm{E}t \text{ iff } \|t\|^{\mathfrak{M}}_{A}(u) \in E_u, \text{ if } t \text{ is a term,}$$
$$u \models^{\mathfrak{M}}_{A} t = t' \text{ iff } \|t\|^{\mathfrak{M}}_{A}(u) = \|t'\|^{\mathfrak{M}}_{A}(u), \text{ if } t = t' \text{ are terms,}$$

The meaning of E and = depends neither on the interpretation I nor the assignment A; for this reason E and = may be called logical constants. Notice that if the identity predicate is available, the existence predicate is definable: Provided that t is distinct from x, if t is a variable, $\mathrm{E}(t) \equiv \exists x(x = t)$ is a valid schema.

The following remarks apply to this particular modeling. All instances of the Barcan formula (BF) are valid in all and only frames satisfying the condition of decreasing domains, that is,

for all u and v, if $(u, v) \in R$ then $E_u \supseteq E_v$.

Similarly, all instances of the converse Barcan formula (CBF) are valid in all and only frames that satisfy the condition of increasing domains, that is

for all u and v, if $(u, v) \in R$ then $E_u \subseteq E_v$.

Of the other predicate logical formulas discussed earlier, $\forall x \forall y(x = y \supset \Box(x = y))$ and $\forall x \forall y(x \neq y \supset \Box(x \neq y))$ are valid, while neither $\mathrm{a} = \mathrm{b} \supset \Box(\mathrm{a} = \mathrm{b})$ nor $\mathrm{a} \neq \mathrm{b} \supset \Box(\mathrm{a} \neq \mathrm{b})$ is valid. This reflects an important difference between how individual variables and individual constants are treated in this modeling: In spite of their name, the denotation of individual constants may vary from point to point in the universe, whereas the denotation of variables, their name notwithstanding, remains fixed throughout the universe. Here is obviously a niche to be filled! Suppose one introduces a new syntactic category of names and requires that the interpretation of a name n be constant over the universe of points; formally, $I_u(\mathrm{n}) = I_v(\mathrm{n})$, for all $u, v \in U$. Then, if m and n are any names, $\mathrm{m} = \mathrm{n} \supset \Box(\mathrm{m} = \mathrm{n})$ and $\mathrm{m} \neq \mathrm{n} \supset \Box(\mathrm{m} \neq \mathrm{n})$ are both valid. The proposed modification amounts to treating the elements of the new category of names as what is now known, after Kripke, as rigid designators.

Among other modelings for modal predicate logic, David Lewis's counterpart theory should be mentioned.

According to the Kripke paradigm, an individual may exist in more than one possible world (with respect to the formal modeling defined above, it is possible that E_u and E_v should overlap, in a model, even if $u \neq v$). For Lewis, however, each individual inhabits its own possible world; but it may have counterparts in other possible worlds. This approach has also been influential, both in philosophical and in mathematical quarters.

INTERPRETATIONS

The original interpretation of modal logic—the official interpretation, if one prefers—was of course the one that led to its construction: the interpretation in terms of necessity and possibility. But over time there have been many others.

THE ALETHIC INTERPRETATION. In formal philosophy, as in formal conceptual analysis generally, there is a constant interplay between intuition and formalism. Efforts to explicate pretheoretical notions lead to a formalism, for example, an axiom system in a formal language or a set theoretical modeling. Once a formalism is in place, it takes on a life of its own: Not only may it undergo a formal development but it can also be interpreted, sometimes in ways that are not foreseen. Reflections on such interpretations lead to refined, sometimes revised, intuitions. The latter in turn may inspire more sophisticated formalisms. And so it goes. The formalism described earlier in this entry is a product of such interplay, having arisen principally as a result of efforts to understand what Georg Henrik von Wright called the alethic modalities necessity and possibility. Not surprisingly, questions persist about to what extent this formalism is a successful explication of one's informal understanding of necessity and possibility.

Formal semantics for modal logic is, by itself, philosophically neutral. The elements of the universe of a modal logical frame, which from a formal point of view are just points in a logical space, must be given a substantial meaning by philosophers who wish to use them outside the realm of pure abstraction. In tense logic the points will be points of time, in epistemic logic perhaps epistemic situations, and so on. Under the alethic interpretation they are often referred to as possible worlds, an ordinary language word with no clear content. Indeed, the question as to what a possible world is has exercised philosophers since the beginning of the Kripke era. Answers—besides those rejecting the entire modal logical enterprise—have been numerous. Lewis argued for an extreme modal realism according to which possible worlds are concrete alternative universes existing in parallel with the actual world. Other philosophers, like Kripke, Alvin Plantinga, Robert Stalnaker, and David M. Armstrong also argued for one kind of modal realism or other but have taken them to be abstract entities. Still other philosophers regarded possible-worlds talk as a kind of convenient fiction or refer to linguistic conventions. The debate continues.

An exact and expressive formalism has the advantage that old informal questions falling within its range of interpretation can be addressed anew. One such question is the venerable distinction between *de dicto* and *de re*. To take Willard Van Orman Quine's well-worn example, consider the claim that the number of planets is necessarily greater than seven. Is it true? There seem to be two different ways of understanding this claim. To bring them out, one can translate them into an ad hoc, quasi-formal language:

(1) $\exists x((x = \text{the number of planets}) \wedge \Box(x > 7))$,

(2) $\Box\exists x((x = \text{the number of planets}) \wedge (x > 7))$.

Statement (1) is said to be *de re*, statement (2) *de dicto*. It may be argued that they say different things (presumably, most would agree that the former is true but that the latter is false). The former seems to "say of an object" (the *res*, the number of planets) that, by necessity, it has a certain property ("being greater than seven"). By contrast, the latter statement says that a certain statement is necessarily true (the *dictum*, namely, that the number of planets, whatever that number may be, is greater than seven). This example illustrates the important interaction between quantifying and modalizing: It is one thing to put a modal operator in front of a closed sentence, as in (2), it is another, arguably more problematic, to quantify into the scope of a modal operator, as in (1). The old topic of essences is obviously not far away.

Another distinction, which has been argued by Kripke, is that between logical modalities and metaphysical modalities (there may also be others, such as physical modalities). Logical necessity implies metaphysical necessity, but the converse is not true. For example, "Phosphorus is identical with Hesperus" (assuming the names *Phosphorus* and *Hesperus* are regarded as rigid designators) and "The chemical composition of water is H_2O" (again assuming that *water* and H_2O are rigid designators) have been offered as examples of statements that are metaphysically, but not logically, necessary.

The (epistemological) distinction between *a priori* and *a posteriori* also comes in here. In Kripke's theory, the two examples given in the preceding paragraph exemplify

statements that, although metaphysically necessary, are nevertheless *a posteriori*. By contrast, given certain assumptions, "The Paris meter is one meter long" may be an example of a statement that is true *a priori* but is not metaphysically necessary.

TWO EARLY MATHEMATICAL INTERPRETATIONS. In the 1930s two technical interpretations of modal logic were made by the two greatest logicians of the twentieth century. One was the so-called provability interpretation, due to Kurt Gödel, according to which $\Box\phi$ is interpreted as "ϕ is provable" or "ϕ is provable in *S*," where *S* is a certain formal system. This interpretation was never forgotten, but it attracted major attention only relatively recently. The other interpretation, due to Tarski, is in terms of topology: Let C and I denote the closure C*X* and the interior I*X*, respectively, of any subset *X* of a topological space *U*. Tarski noted that the closure operator and the interior operator behave in a way analogous to the way the possibility operator and the necessity operator behave in S4. For example, if ϕ and ψ correspond to *X* and *Y*, respectively, then the formulas $\Diamond(\phi \vee \psi) \equiv (\Diamond\phi \vee \Diamond\psi)$, $\Diamond\phi \equiv \Diamond\Diamond\phi$, $\phi \equiv \top$, and $\phi \equiv \bot$ correspond to the equations $C(X \cup Y) \equiv CX \cup CY$, $CX = CCX$, $X = U$ and $X = \emptyset$. More generally, Tarski proved that an equation in topological terms is true in all topological spaces if and only if the corresponding formula is a thesis of S4. Like Gödel's interpretation, Tarski's interpretation, which is related to the development of the theory of closure algebras, was seminal.

THE TEMPORAL INTERPRETATION. A long-standing interest in the work of early Greek logicians combined with a passion for modal logic led Arthur Norman Prior, in the 1950s, to the idea of a modal logic of time. He dubbed his creation tense logic since one of his original motivations was to throw light on the grammatical notion of tense. In the beginning Prior was led to study frames (U, R) in which *R* is a linear relation on *U* (i.e., reflexive, transitive, and connected). Under that interpretation, the interpretation of the modal operators $\Box$ and $\Diamond$ in effect becomes "always in the future" and "some time in the future." One focus for his early interest was the frame $(\mathbb{N}, \leq)$, where $\mathbb{N}$ is the set of natural numbers, which he associated with Diodorus Cronus. Trying to axiomatize the set of formulas valid in this frame—the Diodorean logic, as he called it—Prior successively made three conjectures. The first was that it is S4. This conjecture was disproved by Hintikka, who pointed out that all instances of the schema

(H) $(\Diamond\phi \wedge \Diamond\psi) \supset (\Diamond(\phi \wedge \psi) \vee \Diamond(\phi \wedge \Diamond\psi) \vee \Diamond(\psi \wedge \Diamond\phi))$

are theses of the Diodorean logic but not all of S4. Prior's response was the new conjecture that it is S4.3, that is, the logic whose Lemmon code is KT4H. However, Michael Anthony Eardley Dummett showed that all instances of the schema

(Dum) $\Box(\Box(\phi \supset \Box\phi) \supset \Box\phi) \supset (\Diamond\Box\phi \supset \Box\phi)$

are theses of the Diodorean logic but not all of S4.3. Prior's third conjecture was that the Diodorean logic is S4.3Dum. This final conjecture turned out to be correct, proved by R. A. Bull and, independently, by Kripke.

In general, Prior allowed the temporal ordering to be irreflexive. He also introduced operators for past time as well as for future time. Thus, the basic operators of tense-logic are the diamond operators F and P, with readings "it will be the case (some time in the future) that" and "it was the case (some time in the past) that," and the box operators G and H with the reading "always in the future" and "always in the past." Their truth-conditions in a frame $(U, <)$, where $<$ is at least a strict partial ordering (i.e., irreflexive and transitive), are:

$$u \vDash^{\mathfrak{M}} F\phi \text{ iff } v \vDash^{\mathfrak{M}} \phi, \text{ for some point } v \text{ such that } u \leq v,$$

$$u \vDash^{\mathfrak{M}} P\phi \text{ iff } v \vDash^{\mathfrak{M}} \phi, \text{ for some point } v \text{ such that } v \leq u.$$

$$u \vDash^{\mathfrak{M}} G\phi \text{ iff } v \vDash^{\mathfrak{M}} \phi, \text{ for all points } v \text{ such that } u \leq v,$$

$$u \vDash^{\mathfrak{M}} H\phi \text{ iff } v \vDash^{\mathfrak{M}} \phi, \text{ for all points } v \text{ such that } v \leq u.$$

Tense logic is in effect a kind of bimodal logic: It is natural to think of a tense-logical frame as a frame with two accessibility relations, one for the future and one for the past. What is special to tense logic is that those two relations are inverses of one another (and, consequently, all instances of the schemata $PG\phi \supset \phi$ and $FH\phi \supset \phi$ are valid).

The temporal operators mentioned are not the only ones possible. A particularly important pair of operators studied by Hans Kamp are SINCE and UNTIL:

$$u \vDash^{\mathfrak{M}} \phi \text{ SINCE } \theta \text{ iff there is some } w \in U \text{ such that } w < u \text{ and } w \vDash^{\mathfrak{M}} \theta \text{ and, for all } x \in U, \text{ if } w < x < u \text{ then } x \vDash^{\mathfrak{M}} \phi,$$

$$u \vDash^{\mathfrak{M}} \phi \text{ UNTIL } \theta \text{ iff there is some } w \in U \text{ such that } u < w \text{ and } w \vDash^{\mathfrak{M}} \theta \text{ and, for all } x \in U, \text{ if } u < x < w \text{ then } x \vDash^{\mathfrak{M}} \phi.$$

(In the literature, ϕ SINCE θ and ϕ UNTIL θ are often written $S(\theta, \phi)$ and $U(\theta, \phi)$, respectively.) Kamp proved that in certain contexts, for example, over $(\mathbb{R}, <)$ (where $\mathbb{R}$ is the set of reals and $<$ is the natural strict linear order) his

operators suffice for temporal completeness; that is, in those contexts, all operators corresponding to first-order conditions on the temporal relation can be defined in terms of SINCE and UNTIL and truth-functional connectives. But in general there is no temporal completeness in this sense.

Still another important tense-logical operator is NOW, which refers to a designated, fixed point of reference. A language involving that operator requires a somewhat modified truth-definition: Where before the definition is with respect to a model and a point, it will now be with respect to a model and two points, which one might call the current point and the point of reference—the former is variable, the latter is fixed throughout the definition. The clauses pertaining to the old operators, which only involve the current point, are obvious. The novel clause is

$$(u, t) \vDash^{\mathfrak{M}} \text{NOW}\ \phi \text{ iff } (t, t) \vDash^{\mathfrak{M}} \phi.$$

THE EPISTEMIC INTERPRETATION. The possibility of epistemic logic (the logic of knowledge) and doxastic logic (the logic of belief) was realized by von Wright, who coined the terms, but it was Hintikka who set the field going. Hintikka associated, with each agent *a*, two operators $\mathbf{K}_a$ and $\mathbf{B}_a$, reading "*a* knows that ϕ" for $\mathbf{K}_a\phi$ and "*a* believes that ϕ" for $\mathbf{B}_a\phi$. By the same token, the formal counterparts of "for all that *a* knows, ϕ" and "ϕ is consistent with everything *a* believes" are $\neg\mathbf{K}_a\neg\phi$ and $\neg\mathbf{B}_a\neg\phi$. Already Hintikka's new notation was useful. To know that someone Qs is not the same as knowing someone who Qs, but Hintikka's notation makes this patent—$\mathbf{K}_a\exists x\mathrm{Q}x$ has to mean something different from $\exists x\mathbf{K}_a\mathrm{Q}x$ (compare the distinction between *de dicto* and *de re* mentioned earlier). Discussion about logical relationships was also facilitated. For example, is it reasonable to regard the type (4) schema $\mathbf{K}_a\phi \supset \mathbf{K}_a\mathbf{K}_a\phi$ (positive introspection, the KK-thesis) and the type (5) schema $\neg\mathbf{K}_a\phi \supset \mathbf{K}_a\neg\mathbf{K}_a\phi$ (negative introspection) as valid for rational knowledge? (Hintikka's own inclination was to accept the former but reject the latter.) Another example of the applicability of Hintikka's logic was to the puzzle known after George Edward Moore as Moore's paradox. Suppose I am ignorant of the fact, say, that it is currently raining in Cambridge, England, but that I am sufficiently informed of my own beliefs to be aware of my ignorance. Then someone who knows me may say, truly, "It is raining, but you don't believe it." But, as observed by Moore, it would be distinctly odd of me to agree, saying, "Yes, it is raining, but I don't believe it." Hintikka accounts for the oddness by suggesting that a belief operator $\mathbf{B}_a$ must satisfy certain minimum conditions to count as an operator expressing rational belief. For example, it would be enough if the logic of $\mathbf{B}_a$ was at least as strong as the normal modal logic KD4, for in that logic a sentence $\phi \wedge \neg\mathbf{B}_a\phi$ may be consistent, but a sentence $\mathbf{B}_a(\phi \wedge \neg\mathbf{B}_a\phi)$ is always inconsistent (or, in Hintikka's terminology, doxastically indefensible).

Knowledge and belief about knowledge and belief has been an issue of late, of interest not only to philosophers but also to computer scientists and game theorists. It may be that everyone in a group of agents knows that ϕ, but this does not mean that ϕ is common knowledge in the group (a concept first studied by David Lewis); for that to be the case it is also required that everyone knows that everyone knows that ϕ, knows that everyone knows that everyone knows that ϕ, and so on. Interestingly, this concept can be axiomatized. If G is a nonempty, finite set of agents—for simplicity, assume that $G = \{1, \ldots, n\}$—write $\mathbf{E}_G\phi$ for "every member of G knows that ϕ" and $\mathbf{C}_G\phi$ for "it is common knowledge among the members of G that ϕ." Assuming that $\mathbf{K}_i$ is an S4-operator, for each $i \in G$, the logic of the two new operators may be characterized by requiring $\mathbf{C}_G$ also to be an S4-operator and adding the following conditions:

$$\mathbf{E}_G\phi \equiv (\mathbf{K}_1\phi \wedge \cdots \wedge \mathbf{K}_n\phi),$$

$$\mathbf{C}_G\phi \supset \mathbf{E}_G\phi,$$

$$(\phi \wedge \mathbf{C}_G(\phi \supset \mathbf{E}_G\phi)) \supset \mathbf{C}_G\phi.$$

THE DEONTIC INTERPRETATION. When von Wright published his seminal paper "Deontic Logic" in 1951, he in effect delivered a discipline just waiting to be born. The next decades saw a great number of papers and books written on this topic, but it is probably fair to say that the results are less definitive than those of several other subfields of modal logic. The basic idea is to study operators **O**, **P**, and **F** with the informal readings "it is obligatory that ϕ" for $\mathbf{O}\phi$, "it is permitted that ϕ" for $\mathbf{P}\phi$, and "it is forbidden that ϕ" for $\mathbf{F}\phi$. In so-called standard deontic logic (STD), **O** is treated as the box operator and **P** as the diamond operator of a normal logic; **F** may then be defined by a condition such as $\mathbf{F}\phi \equiv \mathbf{O}\neg\phi$ or $\mathbf{F}\phi \equiv \neg\mathbf{P}\phi$ (to be compared with the validities $\mathbf{P}\phi \equiv \neg\mathbf{O}\neg\phi$ and $\mathbf{O}\phi \equiv \neg\mathbf{P}\neg\phi$). STD—not a precise concept—provides the schema (D) $\mathbf{O}\phi \supset \mathbf{P}\phi$. One schema that for obvious reasons would be inappropriate in a deontic logic is (T), but weaker schemata such as $\mathbf{OO}\phi \supset \mathbf{O}\phi$ and $\mathbf{O}(\mathbf{O}\phi \supset \phi)$ are sometimes included in STD.

Efforts to apply STD to even fairly simple everyday situations will often fail, as shown by the existence of so-called paradoxes, a topic much discussed in the literature.

Best known among the latter are perhaps the paradoxes of William David Ross, Roderick Chisholm, and James W. Forrester. (Ross's paradox was originally formulated within the logic of imperatives, but it is equally relevant for deontic logic.) A person is under an obligation to see to it that (ϕ) a letter is posted. Should he or she do it by seeing to it that (ψ) the letter is burned? Since $\phi \supset (\phi \vee \psi)$ is a tautology, $\mathbf{O}\phi \supset \mathbf{O}(\phi \vee \psi)$ is a thesis of STD. Evidently, according to STD the person should see to it that the letter is posted or burned; Ross found this conclusion bizarre. In Chisholm's paradox there are two things A and B that you may or may not do: Whether (ϕ) you do A is logically independent of whether (ψ) you do B. On the one hand, it ought to be the case that you do B if you do A ($\mathbf{O}(\phi \supset \psi)$). On the other hand, if you do not do A, then neither ought you to do B ($\neg\phi \supset \neg\mathbf{O}\psi$). Furthermore, even though A is something you ought to do ($\mathbf{O}\phi$), you will not do it ($\neg\phi$). In STD this description of a situation, regrettable perhaps but otherwise unremarkable, leads to contradiction. Forrester's paradox is subtler: suppose there is something one must not do, but that if one nevertheless does it, then one should do it in such and such a way. Again, STD comes to grief.

Among the many problems still not resolved in modern deontic logic—Hector–Neri Castañeda's work and his distinction between propositions and practitions notwithstanding—is the age-old question about the relationship between Seinsollen (ought to be) and Tunsollen (ought to do). It is interesting that von Wright, the father of the discipline, originally had intended for his deontic operators to take as arguments, not propositions, but actions; he seems to have changed his mind for technical reasons. With the advent of dynamic logic, it is nowadays possible to reconsider this option.

OTHER INTERPRETATIONS. The techniques of modal logic have been applied to a number of other areas of philosophical interest: imperatives, action, preference, place, even questions. Many of the more interesting applications make use of several modalities. For example, Kanger's theory of rights, which builds on Wesley Newcomb Hohfeld's famous analysis, combines concepts from deontic logic and the logic of action.

EXTENSIONS OF MODAL LOGIC

CONDITIONAL LOGIC. The analysis of conditionals has occupied philosophers for generations. Not all the resulting analyses belong to the field of modal logic, but there is a natural sense in which the conditional logics of Robert Stalnaker and David Lewis may be seen as generalizations of classical modal logic. This is obvious if one employs a notation suggested by Brian Chellas: writing $[\phi]\psi$ and $\langle\phi\rangle\psi$ where Lewis had $\phi \mathrel{\Box\!\!\rightarrow} \psi$ ("if it were the case that ϕ, then it would be the case that ψ") and $\phi \mathrel{\Diamond\!\!\rightarrow} \psi$ ("if it were the case that ϕ, then it might be that ψ"), respectively. By this device, one moves from the language of traditional modal logic, where there is one box operator $\Box$, to a language in which there are as many box operators $[\phi]$ as there are well-formed formulas ϕ. Corresponding to the minimal normal modal logic K is the minimal normal conditional logic in which every box operator satisfies the Kripke schema and the rule of necessitation, and which is also closed under the rule of congruence (if θ^* is the result of replacing all occurrences of ϕ in θ by an occurrence of ψ, then $\theta \equiv \theta^*$ is a thesis if $\phi \equiv \psi$ is). Lewis's logic VC of counterfactuals is the smallest normal conditional logic that contains all instances of the schemata:

$$[\phi]\phi,$$

$$\langle\phi\rangle\psi \supset \langle\psi\rangle\top,$$

$$\phi \supset (\psi \supset [\phi]\psi),$$

$$\phi \supset ([\phi]\psi \supset \psi),$$

$$[\phi \wedge \psi]\theta \supset [\phi](\psi \supset \theta),$$

$$\langle\phi\rangle\psi \supset ([\phi](\psi \supset \theta) \supset [\phi \wedge \psi]\theta).$$

Stalnaker's logic is obtained by requiring that also all instances of the schema

$$\langle\phi\rangle\theta \equiv [\phi]\theta$$

be theses.

DYNAMIC LOGIC. Looking for a useful way to formalize reasoning about programs, Vaughan Pratt, a computer scientist, arrived at what is nowadays known as dynamic logic, a formalism similar to modal logic; in fact, dynamic logic may be viewed as a generalization of modal logic in the same way as Chellas-formulated conditional logic may be seen as a generalization of modal logic. With each program α Pratt associated a box operator $[\alpha]$ and a diamond operator $\langle\alpha\rangle$, reading $[\alpha]\phi$ as "after every terminating computation according to α, ϕ" and $\langle\alpha\rangle\phi$ as "after some terminating computation according to α, ϕ." The resulting logic, originally called the modal logic of programs, evidently contains two basic categories of expressions, terms (for programs), and formulas (for propositions). A further complication over modal logic is the existence of term operators for the so-called regular operations. Thus, if α and β are programs, then $\alpha + \beta$ is

the program consisting of α or β (the latter concept is of course of interest only in the context of nondeterministic automata) while $\alpha ; \beta$ is the program consisting of α immediately followed by β, and α^* is the program consisting of α some finite number of times, possibly 0 (again, of interest only in a nondeterministic context). Finally, Pratt allowed a test program: $?\phi$ is a program that, if run, fails if ϕ is false but otherwise returns to status quo. An axiomatization of PDL (propositional dynamic logic) is obtained by requiring each box operator $[\alpha]$ to be a normal modal operator and adding the following axiom schemata:

$$[\alpha + \beta]\phi \equiv ([\alpha]\phi \wedge [\beta]\phi),$$

$$[\alpha ;\beta]\phi \equiv [\alpha][\beta]\phi,$$

$$[\alpha^*]\phi \supset \phi,$$

$$[\alpha^*]\phi \supset [\alpha]\phi,$$

$$[\alpha^*]\phi \supset [\alpha^*][\alpha^*]\phi,$$

$$(\phi \wedge [\alpha^*](\phi \supset [\alpha]\phi)) \supset [\alpha^*]\phi,$$

$$[?\phi]\chi \equiv (\phi \supset \chi).$$

OTHER INTERPRETATIONS. Some of the generalizations of modal logic that have been made over the last few decades have an origin far from modal logic. Dynamic logic is one example that has already been mentioned. Another example is description logic, which is a family of formalisms used by computer scientists to represent knowledge that is already expressed in a certain regimented form; only after extensive work did those practitioners realize that what they were doing could be seen as a version of multimodal logic, that is, modal logic with several normal operators.

An example closer to ordinary modal logic is hybrid logic, a way of doing modal logic actually anticipated by Prior. Here, the object language of traditional modal logic is augmented by the introduction of concepts belonging to semantics, a device that can greatly increase the expressive strength of the formal language. One such augmentation is to allow a new category of syntactic objects, called nominals, a special set of propositional constants whose semantic interpretation is as singleton sets; in other words, nominals represent propositions that are true at exactly one point in the universe of a model. If i is a nominal and ϕ an ordinary formula, then $(\mathrm{i} \supset \phi) \vee (\mathrm{i} \supset \neg\phi)$ and $\Diamond(\mathrm{i} \wedge \phi) \supset \Box(\mathrm{i} \supset \phi)$ exemplify formulas valid in every frame. By contrast $\mathrm{i} \supset \Box\neg\mathrm{i}$ is an example of a formula valid in exactly the class of frames (U, R) in which R is irreflexive. This is a striking fact, for irreflexivity is notoriously not expressible in ordinary modal logic—the logic determined by the class of all frames with irreflexive accessibility relations is the same as the logic determined by the class of all frames, that is, K.

Like description logic, hybrid logic is actually a family of logics with different object languages. This proliferation of languages bears witness to the many different uses to which modal logic is nowadays being put. In this regard it is interesting to note a certain trade-off between more restrictive and more permissive options: in general, the more expressive a language is, the more endangered are desirable properties like completeness and decidability. Some philosophers may find the multifariousness of present-day computer science–driven modal logic bewildering. At any rate, we have come a long way from the beginning of modal logic when C. I. Lewis sought, and for a while thought he had found, the one and only logic of strict implication.

See also A Priori and A Posteriori; Armstrong, David M.; Carnap, Rudolf; Chisholm, Roderick; Diodorus Cronus; Dummett, Michael Anthony Eardley; Gödel, Kurt; Hintikka, Jaakko; Kripke, Saul; Lewis, Clarence Irving; Lewis, David; Logic, History of; Łukasiewicz, Jan; Marcus, Ruth Barcan; Mathematics, Foundations of; Modality, Philosophy and Metaphysics of; Moore, George Edward; Plantinga, Alvin; Prior, Arthur Norman; Provability Logic; Quine, Willard Van Orman; Ross, William David; Tarski, Alfred; Wright, Georg Henrik von.

Bibliography

ORIGINAL WORKS

Carnap, Rudolf. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press, 1947.

Hintikka, Jaakko. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca, NY: Cornell University Press, 1962.

Jónsson, Bjarni, and Alfred Tarski. "Boolean Algebras with Operators: Part I." *American Journal of Mathematics* 73 (1951): 891–939.

Kanger, Stig. "Provability in Logic." PhD diss., Stockholm University, 1957.

Kripke, Saul A. "Semantical Analysis of Modal Logic: I. Normal Modal Propositional Calculi." *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9 (1963): 67–96.

Kripke, Saul A. "Semantical Analysis of Modal Logic: II. Non-normal Modal Propositional Calculi." In *The Theory of Models*, edited by J. W. Addison, L. Henkin, and A. Tarski, 206–220. Amsterdam, Netherlands: North-Holland, 1965.

Kripke, Saul A. "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94.

Lewis, Clarence Irving. *A Survey of Symbolic Logic*. Berkeley: University of California Press, 1918.

Lewis, Clarence Irving, and Cooper Harold Langford. *Symbolic Logic*. New York: Century, 1932.

Lewis, David. *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

McKinsey, J. C. C. "A Solution of the Decision Problem for the Lewis Systems S2 and S4 with an Application to Topology." *Journal of Symbolic Logic* (1941): 117–134.

Montague, Richard. "Logical Necessity, Physical Necessity, Ethics, and Quantifiers." *Inquiry* 4 (1960): 259–269.

Prior, Arthur. *Past, Present, and Future*. Oxford, U.K.: Clarendon Press, 1967.

Prior, Arthur. *Time and Modality*. Oxford, U.K.: Clarendon Press, 1957.

von Wright, Georg Henrik. "Deontic Logic." *Mind* 60 (1951a): 1–15.

von Wright, Georg Henrik. *An Essay in Modal Logic*. Amsterdam, Netherlands: North-Holland, 1951b.

SURVEY ARTICLES

Blackburn, Patrick, Frank Wolter, and Johan van Benthem. *Handbook of Modal Logic*. Forthcoming.

Gabbay, D. M., and F. Guenthner. *Handbook of Philosophical Logic. Vol. 2, Extensions of Classical Logic*. Dordrecht, Netherlands: D. Reidel, 1984.

Gabbay, D. M., and F. Guenthner. *Handbook of Philosophical Logic*. Vols. 3, 4, and 7. 2nd ed. Dordrecht, Netherlands: Kluwer Academic, 2001–2002.

Gochet, P., and P. Gribomont. "Epistemic Logic." In *Handbook of the History and Philosophy of Logic*, edited by D. M. Gabbay and John Woods. Forthcoming.

Goldblatt, Robert. "Mathematical Modal Logic: A View of Its Evolution." *Journal of Applied Logic* 1 (2003): 309–392.

TEXTBOOKS

Chellas, Brian F. *Modal Logic: An Introduction*. New York: Cambridge University Press, 1980.

Fitting, Melvin, and Richard L. Mendelsohn. *First-Order Modal Logic*. Dordrecht, Netherlands: Kluwer Academic, 1998.

Goldblatt, Robert. *Logics of Time and Computation*. CSLI Lecture Notes. Vol. 7. Chicago: Chicago University Press, 1987.

Hughes, G. E., and M. J. Cresswell. *A New Introduction to Modal Logic*. New York: Routledge, 1996.

Krister Segerberg (2005)

MODEL THEORY

In 1954 Alfred Tarski proposed the name *theory of models* for the study of "mutual relations between sentences of formalized theories and mathematical [structures] in which these sentences hold." This definition hides a program that was to apply metamathematical results (particularly the Compactness Theorem of first-order logic) in what Abraham Robinson in 1950 had called "the development of actual mathematics." Anatolii I. Mal'tsev had launched this program in the Soviet Union in 1940, but communications were bad in this period and the program started afresh in the late 1940s with Tarski in the United States and Robinson in Britain. Mathematical model theory in the sense of this program has been remarkably successful, particularly in its applications to group theory and geometry, and it has far outgrown Tarski's initial definition of the theory of models.

Tarski's definition rested on the fact that one can use formal languages to define classes of structures. For mathematical applications it has turned out to be just as important that one can use formal languages to define sets and relations within a single structure. But at its base, model theory is more general even than this. Arguably it stands in the same relation to the traditional theory of definitions as modern proof theory stands to the traditional theory of syllogisms.

Most sentences are true in some contexts and false in others. If S is a sentence, then by an *interpretation* of S we mean a parcel of information about some possible context, which is enough to make S either true or false in that context. Suppose I is an interpretation of S. If I makes S true, we call I a *model* of S and we say that S is *true in I*. "Truth-in-a-model" is honest to goodness truth, no less than (say) being true at 3 o'clock.

The sentence S defines a class of interpretations, namely the class of its models. A simple example is the mathematical equation

$$x^2 + y^2 = 1$$

where x and y are variables ranging over real numbers. An interpretation of this equation consists of a pair of real numbers b,a where x is to name a and y to name b. Under this interpretation the sentence is true if and only if the point b,a lies on the circle C of radius 1 around the origin in the cartesian plane. So the circle C is the class of models of the equation. This example assumes that we have specified what form an interpretation of the equation should take. In concrete applications of model theory one begins with such a specification.

The sentence S can come from a natural language or a formal one. The range of information that might appear in interpretations is vast. They can specify the time of utterance, the time spoken of, the place, the speaker's identity, salient objects in the context (to give reference to "the previous owner", "the latter symbol", "Peter", etc.). They can also supply meanings for words that have none. But mathematical model theory concerns itself almost

entirely with interpretations of a kind called *structures*. A structure supplies a set of objects, called the *domain* or *universe* of the structure. Besides giving a domain, a structure interprets expressions by attaching them to elements of the domain, or to set-theoretic objects built up from elements of the domain. For example a mathematical model theorist, to interpret the sentence

The mome raths outgrabe.

would probably supply two sets X and Y, together with the information that X is the set of things that count as mome raths and Y is the set of things that count as having outgribben. This interpretation is a model of the sentence if and only if X is a subset of Y.

When the sentence S comes from a formal language of logic, one can describe precisely how the truth value of S depends on the sets or objects used to interpret symbols of S. Tarski's model-theoretic definition of truth and satisfaction is a paradigm for this kind of description. The model-theoretic truth definition was an adaptation of the truth definition that Tarski gave in 1933 for formal languages. In that earlier definition Tarski assumed that all symbols needing an interpretation already had one (in general a set-theoretic one), and so the definition was strictly not model-theoretic. But truth definitions that run along similar lines to Tarski's, for example the definitions of truth underlying Richard Montague's semantics for fragments of English, are called "model-theoretic"; probably the use of set theory and recursion on the complexity of formulas are the features that this name brings to mind.

As a discipline, model theory takes no stand at all on whether there are possible worlds or on what objects there are in the universe. If you believe in possible worlds you can study interpretations that involve possible worlds; if you don't, you probably won't. There are branches of model theory where one puts strong limits on the kinds of interpretation that are allowed: For example in *recursive model theory* the structures are built up from computable functions of natural numbers. But since structures are set-theoretic objects, most mathematical model theorists make free use of the axioms of Zermelo-Fraenkel set theory, including the Axiom of Choice.

One should distinguish between model theory and "mathematical modeling". Modeling a phenomenon usually involves constructing a formal theory rather than a set-theoretic structure. But there are overlaps. For example abstract state machines, introduced by the model theorist Yuri Gurevich, are set-theoretic structures used to model parallel computation. In another direction, papers in Morgan and Morrison discuss the relations between theories and structures in scientific research, with particular reference to physics and economics.

FIRST-ORDER MODEL THEORY

First-order model theory is the most developed part of model theory, and other parts of model theory tend to be generalizations or analogues of the first-order case. We begin with some preliminary definitions that rest on first-order logic.

DEFINING STRUCTURES, TRUTH, AND SATISFACTION. First we define signatures. A *signature* is a collection of symbols as follows:

(1) *Relation symbols*, usually

$$P, Q, R, R_0, R_1, R_2, \ldots.$$

(2) *Individual constant symbols*, or more briefly *constants*, usually

$$a, b, c, c_0, c_1, c_2, \ldots.$$

(3) *Function symbols*, usually symbols such as

$$F, G, H, F_0, F_1, F_2, \ldots.$$

Each relation symbol and each function symbol in a signature has an *arity*, which is a positive integer. If a symbol has arity n, we say that the symbol is *n-ary*. We normally require that no symbol occurs in more than one of these three kinds, and that no relation or function symbol occurs with more than one arity. We say that a signature σ is a *reduct* of a signature τ (and that τ is an *expansion* of σ) if every constant in σ is also a constant in τ, every relation symbol of σ is also a relation symbol in τ with the same arity, and likewise for the function symbols in σ.

Let σ be a signature. A *σ-structure* is an ordered pair $A = \langle \mathrm{dom}(A), f_A \rangle$ as follows:

dom(A) is a nonempty set, known as the *domain* of A.

f_A is a function whose domain is the set of symbols in the signature σ.

For each constant c of σ, $f_A(c)$ is an element of dom(A); we write this element as c_A.

For each relation symbol R of σ, $f_A(R)$ is an n-ary relation on dom(A), where n is the arity of R; we write this relation as R_A.

For each function symbol F of σ, $f_A(F)$ is an n-ary function $F_A : \mathrm{dom}(A)^n \to \mathrm{dom}(A)$, where n is the arity of F.

By a *structure* we mean a σ-structure for some signature σ.

If A is a τ-structure and σ is a reduct of τ then we can make A into a σ-structure by removing the symbols not in σ; the resulting σ-structure is written $A|\sigma$ and called a *reduct* of A. Likewise A is an *expansion* of $A|\sigma$.

By the *elements* of a structure A we mean the elements of dom(A). (For example a structure A and its reduct $A|\sigma$ have the same elements.) By the *cardinality* of A we mean the cardinality of dom(A).

For each signature σ there is a corresponding first-order language $L(\sigma)$ as in the entry "First-Order Logic". Since each first-order language L is of the form $L(\sigma)$ for a unique signature σ, we can also refer to σ-structures as L-*structures*. We borrow the following facts and definitions from the entry "First-Order Logic", under the assumption that L is a first-order language and A is an L-structure.

If ϕ is a sentence of L then ϕ is either true or false in A. If ϕ is true in A, we write $A \vDash \phi$ and we call A a *model* of ϕ. If ϕ is false in A we write $A \nvDash \phi$.

By a *theory* in L we mean a set T of sentences of L. By a *model* of T we mean a model of all the sentences in T. We say that T is *consistent* if T has a model; otherwise it is *inconsistent*.

Let T be a theory in L and ϕ a sentence of L. We say that ϕ is a *consequence* of T, and that T *entails* ϕ, in symbols

(1) $$T \vDash \phi,$$

if every L-structure that is a model of T is also a model of ϕ. The theory T is said to be *complete* if for every sentence ϕ of L, either ϕ or $\neg\phi$ is a consequence of T. The expression (1) is called a *sequent*; it is *valid* if T does entail ϕ.

We write $\phi(x_1, \dots, x_n)$ for a formula of L whose free variables are all among $x_1, \dots, x_n$. If $a_1, \dots, a_n$ are elements of A, we write

$$A \vDash \phi[a_1, \dots, a_n],$$

pronounced "a_1 to a_n satisfy ϕ in A", if ϕ is true in A when each free variable x_i is interpreted as a name of a_i. This notion can be defined set-theoretically without relying on the semantic notion "name of".

These fundamental facts and definitions allow us to use first-order sentences in order to define classes of structures, and to use first-order formulas in order to define classes of elements in structures.

DEFINING CLASSES OF STRUCTURES. We write Mod(T) for the class of all L-structures that are models of the theory T. If A is an L-structure, we write Th(A) for the set of all sentences ϕ of L which are true in A; Th(A) is known as the *complete first-order theory* of A. If **K** is a class of L-structures, we write Th(**K**) for the set of those sentences of L which are true in every structure in **K**. We say that two L-structures A and B are *elementarily equivalent*, in symbols $A \equiv B$, if Th(A) = Th(B). Elementary equivalence is an equivalence relation on the class of L-structures. We say that two theories S and T in L are *equivalent* if Mod (S) = Mod (T); this is an equivalence relation on the class of theories in L.

Theorem 1 *The notions* Mod *and* Th *are related as follows:*

1. *If* $T \subseteq U$ *then* $\mathrm{Mod}(T) \supseteq \mathrm{Mod}(U)$.
2. *If* $\mathbf{J} \subseteq \mathbf{K}$ *then* $\mathrm{Th}(\mathbf{J}) \supseteq \mathrm{Th}(\mathbf{K})$.
3. $\mathbf{K} \subseteq \mathrm{Mod}(\mathrm{Th}(\mathbf{K}))$ *and* $\mathrm{Th}(\mathbf{K}) = \mathrm{Th}(\mathrm{Mod}(\mathrm{Th}(\mathbf{K})))$.
4. $T \subseteq \mathrm{Th}(\mathrm{Mod}(T))$ *and* $\mathrm{Mod}(T) = \mathrm{Mod}(\mathrm{Th}(\mathrm{Mod}(T)))$.

These facts are all immediate from the definitions.

The theory Th(Mod(T)) is called the *deductive closure* of the theory T; it consists of all the consequences of T. A theory is said to be *deductively closed* if it is its own deductive closure. By 3 of Theorem 1, a theory is deductively closed if and only if it is of the form Th(**K**) for some class **K** of structures. (In some older literature, deductive closure was included in the definition of "theory".)

A class of structures of the form Mod($\{\phi\}$), where ϕ is a single sentence, is said to be *first-order definable*, or an EC class. A class of structures of the form Mod(T), where T is a theory, is said to be *first-order axiomatisable*, or *generalised first-order definable*, or an EC_Δ class. A class **K** of L-structures is said to be *closed under elementary equivalence* if every L-structure elementarily equivalent to a structure in **K** is also in **K**.

We pause for some examples.

Example 1: Equivalence relations. We use the signature with one binary relation symbol E; call this signature σ. We write Exy for $E(x,y)$. An equivalence relation is a σ-structure that is a model of the following finite theory, which we shall call T_{eq}:

$\forall x\, Exx$ (reflexive)
$\forall x \forall y\, (Exy \rightarrow Eyx)$ (symmetric)
$\forall x \forall y \forall z\, (Exy \wedge Eyz \rightarrow Exz)$ (transitive)

Example 2: Fields. The following example has been central in the development of model theory. We adopt a signature with constants 0 and 1, binary function symbols + and · and a 1-ary function symbol –. This signature is appropriate for talking about rings, so it is known as the *signature of rings*. We normally write $+(x,y)$, $\cdot(x,y)$ and $-(x)$ as $x + y$, xy and $-x$ respectively, and we use standard mathematical notation such as $x \neq y$ for $\neg(x = y)$. The *theory of fields*, T_f, consists of the following sentences:

1. $\forall x \forall y \forall z\, (x + (y + z) = (x + y) + z)$
2. $\forall x \forall y\, (x + y = y + x)$
3. $\forall x\, (x + 0 = x)$
4. $\forall x\, (x + -x = 0)$
5. $\forall x \forall y \forall z\, (x(yz) = (xy)z)$
6. $\forall x \forall y\, (xy = yx)$
7. $\forall x\, (x \cdot 1 = x)$
8. $\forall x \forall y \forall z\, (x(y + z) = xy + xz)$
9. $0 \neq 1$
10. $\forall x \exists y\, (x \neq 0 \rightarrow xy = 1)$

We write 2 for 1 + 1, 3 for 1 + 1 + 1 and so on. A field is said to be *of characteristic 0* if it is also a model of the infinitely many axioms

11. $n \neq 0$

where n is any positive integer. We write x^2 for the term xx, x^3 for xxx and so on. Let $t_n(x, y_1, \dots, y_n)$ be the term

$$x^n + y_1 x^{n-1} + y_2 x^{n-2} + \dots + y_{n-2} x^2 + y_{n-1} x + y_n.$$

A field is said to be *algebraically closed* if it is a model of the infinitely many axioms

12. $\forall y_1 \forall y_2 \dots \forall y_n \exists x\, (t_n(x, y_1, \dots, y_n) = 0)$

where n is any positive integer. The classes of fields, fields of characteristic 0 and algebraically closed fields were all well known before these axioms were written down as first-order sentences, and the first-order sentences say exactly the same as the earlier informal definitions of those classes.

In the light of our earlier definitions, several natural questions arise. For example:

> Question One. Is there an algorithm to determine whether any given sentence ϕ is a consequence of T_{eq}?
>
> Question Two. For which equivalence relations A is Mod(Th(A)) first-order definable?
>
> Question Three. If A and B are two algebraically closed fields of characteristic 0, how can we tell whether they are elementarily equivalent?
>
> Question Four. What is an example of a class **K** that is closed under elementary equivalence but not first-order axiomatisable?

We will return to these questions below.

The *infinite spectrum* of a class **K** is the class of infinite cardinals κ such that **K** contains a structure of cardinality κ. Questions about the possible infinite spectra of classes of the form Mod(T) were first raised by Leopold Löwenheim in 1915, and below we shall see some "Löwenheim-Skolem" theorems that describe these spectra.

DEFINING CLASSES OF ELEMENTS. The notions described above have analogues within a single structure. Suppose A is an L-structure. By an *n-tuple in A* we mean an ordered n-tuple $(a_1, \dots, a_n)$ of elements of A. We write $\Phi(x_1, \dots, x_n)$ for a set Φ of formulas of L of the form $\phi(x_1, \dots, x_n)$ (the same integer n for each formula). We say that an n-tuple $(a_1, \dots, a_n)$ in A *realises* Φ if

$$\text{for all } \phi \text{ in } \Phi, A \vDash \phi[a_1, \dots, a_n].$$

We write $\Phi(A^n)$ for the set of all n-tuples in A that realise Φ. If Φ contains just one formula ϕ, we write $\Phi(A^n)$ as $\phi(A^n)$ and we say that this set of n-tuples is (*first-order*) *definable without parameters*. The sets $\Phi(A^n)$ are said to be *infinitarily definable without parameters*, or $\bigwedge$-*definable without parameters*.

For the analogous notions of definability *with parameters* we allow the formulas ϕ to contain constants (in an expanded signature) to name some elements of A. For example if we are talking about the rational numbers in a signature whose symbols are < for the ordering and 0, 1 for the corresponding numbers, then the interval (0,1) of rational numbers strictly between 0 and 1 will be definable without parameters, the interval (3,4) will be definable with parameters, and the interval $(\sqrt{2}, \pi)$ will not be definable at all. When model theorists talk about definable sets, they sometimes mean with parameters and sometimes without; if in doubt you have to ask.

Let T be a theory in the first-order language L, and $\phi(x_1, \ldots, x_n)$ and $\psi(x_1, \ldots, x_n)$ formulas of L. We say that ϕ is *equivalent to* ψ *modulo* T if $\phi(A^n) = \psi(A^n)$ for every model A of T; this is equivalent to saying that the sentence

$$\forall x_1 \ldots \forall x_n (\phi \leftrightarrow \psi) \tag{2}$$

is a consequence of T. Likewise we say that ϕ is *equivalent to* ψ *in* the L-structure A if $\phi(A^n) = \psi(A^n)$; this is equivalent to saying that (2) is true in A.

If $\Phi(x_1, \ldots, x_n)$ is a set of formulas of L, we can ask whether there is an L-structure A in which $\Phi(A^n)$ is not empty. If the answer is Yes, we say that Φ is an *n-type*, or more briefly a *type*, and we say that the structure A *realises* the type. There may be other structures B for which $\Phi(B^n)$ is empty; these structures are said to *omit* the type. We say that the type Φ is *complete* if for every formula $\phi(x_1, \ldots, x_n)$ of L, exactly one of $\Phi \cup \{\phi\}$ and $\Phi \cup \{\neg\phi\}$ is a type.

For example let $\mathbb{N}$ be the 'natural number' structure whose domain is the set of natural numbers $\underline{0}, \underline{1}, \underline{2}, \ldots$, with symbols to express 0, 1, addition, multiplication and 'less than' <. Consider the infinite set $\Phi(x)$ of formulas

$$0<x,\ 1<x,\ 2<x,\ \ldots$$

The set $\Phi(x)$ is in fact a type, but it is clear that $\mathbb{N}$ omits this type; there are no infinite natural numbers. A natural question is:

> Question Five. Are there structures elementarily equivalent to $\mathbb{N}$ which realise this type?

The answers to Questions One to Five are not obvious. Many of the techniques of model theory were devised in order to answer just such questions. Historically the first three major techniques in this area were elimination of quantifiers, back-and-forth, and the Compactness Theorem. The next three sections discuss these.

ELIMINATION OF QUANTIFIERS

Thoralf Skolem, Charles Langford, and Alfred Tarski developed the method of elimination of quantifiers during the 1920s as a way of analyzing structures or classes of structures.

As the name indicates, the idea of elimination of quantifiers is to express as much as possible without using quantifiers. Let Φ be a set of formulas of a language L. Write Φ' for the smallest class of formulas of L such that (i) $\Phi \subseteq \Phi$', (ii) if ϕ is in Φ' then $\neg\phi$ is in Φ', and (iii) if ϕ and ψ are in Φ' then $(\phi\wedge\psi)$ and $(\phi\vee\psi)$ are in Φ'. The formulas in Φ' are called the *boolean combinations* of formulas in Φ. There can be quantifiers in the formulas in Φ, but when we form boolean combinations of them we add no more quantifiers.

Let L be a first-order language and **K** a class of L-structures. A successful elimination of quantifiers for **K** consists of the following items:

(i) a set T of sentences of L that are true in all structures in **K**;

(ii) a set Φ of formulas of L, called the *elimination set*;

(iii) a proof that if $\psi(x_1, \ldots, x_n)$ is any formula of L then ψ is equivalent modulo T to a boolean combination $\psi^*(x_1, \ldots, x_n)$ of formulas in the elimination set.

We carry out an elimination of quantifiers as follows. The class **K** of structures already determines the signature. We begin by choosing T to be—provisionally—a set of sentences that are clearly true in all structures in **K**; our exact choice could depend on aesthetic or pedagogic considerations. We launch (ii) by including all atomic formulas in the elimination set. From this point on, we aim to prove (iii) by induction on the number of occurrences of quantifiers in ψ, with a subinduction on the complexity. If $\psi^\star(x_1, \ldots, x_n)$ and $\chi^\star(x_1, \ldots, x_n)$ have been found, we can take $(\psi\wedge\chi)^\star$ to be $\psi^\star \wedge\chi^\star$, and likewise for other truth-functional combinations. We can take $(\forall x_n\psi)^\star$ to be $\neg\exists x_n \neg(\psi^\star)$. This leaves the case $\exists x_n\ \psi$, and this is where we "eliminate the quantifier".

We first put $\psi^\star$ into disjunctive normal form, and then we use the logical equivalence

$$\exists x\ (\phi_1 \vee\phi_2) \equiv (\exists x\ \phi_1 \vee \exists x\ \phi_2)$$

to reduce to the case of a formula $\exists x_n\ \theta$ where θ is a conjunction of formulas in the elimination set and negations of formulas in the elimination set. The next step depends on **K** and perhaps on our mathematical skill. If we can find a boolean combination ϕ of formulas in the elimination set, and a proof that ϕ is equivalent to $\exists x_n\ \theta$ modulo T, then this case is taken care of. Otherwise we have two options. First if we can find a suitable formula ϕ that is certainly equivalent to $\exists x_n\ \theta$ in all structures in **K** but we can't prove this equivalence from T then we can add the equivalence statement to T. Second, as a last resort, we can add $\exists x_n\ \theta$ to the elimination set. We hope to reach a point where we can prove (iii) for all formulas. When this point is reached the quantifier elimination proper is complete. If heaven favors us (and this is not guaranteed) by this stage we will also know which boolean combinations of sentences in the elimination set are true in all structures in **K**. Adding these to T gives a theory T' equivalent

to Th(**K**). With more good luck we may find that the reductions in (iii) allow us to construct an algorithm to determine whether any given sentence of L is a consequence of T', so that we have a decision procedure for Th(**K**).

Example 1 continued: Equivalence relations. For T we take the theory T_{eq} defining the class of equivalence relations. As a first attempt at the elimination set Φ we take all atomic formulas of L. There are two kinds of atomic formula, namely $(x = y)$ and Exy. Trial and error shows that for every positive integer n the formula $\chi_n(x)$ expressing "There are at least n elements that are in the same equivalence class as x" is not reducible to a boolean combination of atomic formulas; so we put χ_n in the set Φ too. Similarly we add to Φ all the sentences $\sigma_{m,n}$ expressing 'There are at least m equivalence classes containing at least n members each', where m and n are any positive integers, and the sentences $\theta_{m,n}$ expressing "There are at least m equivalence classes of size exactly n". It turns out that this is enough for an elimination set. There is an algorithm reducing each sentence to a boolean combination of sentences in Φ, and there is an algorithm determining which boolean combinations of sentences in Φ are consequences of T. Thus T is a decidable theory and we have an answer to Question One.

Example 3: The field $\mathbb{R}$ of real numbers. We take the signature to be the expansion of the signature of rings got by adding a binary relation symbol < for the ordering of the reals. (Without this added symbol we would need to put $\exists y\ (x = y^2)$ into the elimination set; with < this formula is equivalent to $(x = 0 \vee 0 < x)$.) Tarski showed that a set of axioms for Th($\mathbb{R}$) is given by the theory T_f of Example 2 together with an axiom saying that for every element r, either r or $-r$ is a square, an axiom saying that $r < s$ if and only if $r \neq s$ and $s - r$ has a square root, axioms saying that -1 is not a sum of squares, and the axioms 12 for odd positive integers n. It then came to light that these axioms define the class of real-closed fields. Tarski also gave a decision procedure for the set of consequences of this theory. The elimination set is interesting: it consists exactly of the atomic formulas. As a corollary, the subsets of $\mathbb{R}$ that are first-order definable with parameters consist of the finite unions of sets of the following kinds: singletons $\{a\}$, intervals $(a,b$ (the set of elements r with $a < r$ and $r < b$), intervals $(-\infty)$ (the set of all elements $< b$) and intervals (a,∞) (the set of all elements $>a$).

A structure A whose elements are linearly ordered by an ordering relation $<_A$, and for which the sets first-order definable with parameters are exactly the finite unions of singletons and intervals as in Example 3, is said to be *o-minimal*. The knowledge that a structure is o-minimal gives powerful information about the structure. Beginning with Alex Wilkie's demonstration in 1991 that the expansion of the field of real numbers with an exponentiation function x^y is o-minimal, many other o-minimal expansions of $\mathbb{R}$ have been found, and there is promise of deep applications in real function theory.

From around 1950 more powerful and algebraic methods were found that gave largely the same information as the method of elimination of quantifiers. But it remains one of the best methods for discovering decision procedures when the theory is decidable.

BACK-AND-FORTH

Suppose A and B are σ-structures, where σ is a signature. By a *partial isomorphism* from A to B we mean a function e from a subset X of dom(A) to dom(B) such that if $\phi(x_1, \ldots, x_n)$ is an atomic formula of L and $a_1, \ldots, a_n$ are any elements in X then

$$A \vDash \phi[a_1, \ldots, a_n] \text{ if and only if } B \vDash \phi[e(a_1), \ldots, e(a_n)].$$

If e is a partial isomorphism from A to B and the domain X of e is the whole of dom(A), we say that e is an *embedding* of A into B. If e is an embedding of A into B and every element of B is of the form $e(a)$ for some element a of A then we say that e is an *isomorphism* from A to B. We say that A is *isomorphic* to B, in symbols $A \cong B$, if there is an isomorphism from A to B. The relation $\cong$ is an equivalence relation on the class of L-structures, and its equivalence classes are called *isomorphism types*.

If A and B are isomorphic σ-structures, then A and B must be elementarily equivalent, $A \equiv B$. The definition of "partial isomorphism", and hence also the definition of $\cong$, are easily rewritten in ways that refer to the signature σ but not to any formula of the language $L(\sigma)$. In the 1950s one aim of research was to find an "algebraic" description of elementary equivalence that doesn't mention formulas either. Roland Fraïssé gave essentially the following answer, which is known as the *back-and-forth* method.

A *back-and-forth system* from A to B is a set I of partial isomorphisms from A to B such that

(a) I is not empty.

(b) If i is in I and a is an element of A then there are an element b of B and a partial isomorphism j in I such that

$$i \cup \{\langle a, b \rangle\} \subseteq j;$$

(c) If i is in I and b is an element of B, then there are an element a of A and a partial isomorphism j in I such that

$$i \cup \{ \langle a,b \rangle\} \subseteq j.$$

By a *finite relational signature* we mean a signature with only finitely many symbols, none of which are function symbols. (Constants are allowed.)

Theorem 2 *If there is a back-and-forth system I from A to B then A ≡ B. If A ≡ B then for every finite relational signature σ, there is a back-and-forth system from* $A|\sigma$ *to* $B|\sigma$.

Example 4: Dense linear orderings without endpoints. We adopt a signature σ with one binary relation symbol <, and we write $x < y$ for $<(x, y)$. A *dense linear ordering without endpoints* is a σ-structure that is a model of the following set of sentences:

1. $\forall x \neg(x < x)$	(irreflexive)
2. $\forall x\forall y\forall z\ (x < y \wedge y < z \rightarrow x < z)$	(transitive)
3. $\forall x\forall y\ (x < y \vee y < x \vee x = y)$	(linear)
4. $\forall x\forall y\exists z\ (x < y \rightarrow x < z \wedge z < y)$	(dense)
5. $\forall x\exists y\exists z\ (y < x\ \wedge x < z)$	(no endpoints)

We shall write this set of sentences as T_{dlo}.

Suppose A and B are dense linear orderings without endpoints. An *order-preserving partial map* from A to B is a function e from a finite set X of elements of A to the domain of B such that if the elements of X are $a_1, \ldots , a_n$ with

(3) $$a_1 <_A a_2 <_A \ldots <_A a_n$$

then

(4) $$e(a_1) <_B e(a_2) <_B \ldots <_B e(a_n).$$

Write $I(A,B)$ for the set of all order-preserving partial maps from A to B.

One can check from the definitions that every function in $I(A,B)$ is a partial isomorphism from A to B. Also $I(A,B)$ is a back-and-forth system from A to B. Suppose for example e is in I with domain $\{a_1, \ldots , a_n\}$ as in (3), and a is an element of A that is not in the domain of e. One possibility is that $a <_A a_1$. By sentence 5 of T_{dlo} there is an element b of B with $b <_B e(a_1)$; then $e \cup \{ \langle a,b \rangle\}$ is a function in I that extends e and has a in its domain. The other possibilities for a are similar, using sentences 4 and 5. The same argument, going from B to A, shows that if e is in I and b is an element of B then there is a function in I that extends e and takes some element of A to b.

By Theorem 2 it follows that $A \equiv B$; so any two dense linear orderings without endpoints are elementarily equivalent, and the theory T_{dlo} is complete.

We can say more. Suppose A and B both have countably many elements; list the elements of A as $a_0, a_1, \ldots$ and the elements of B as $b_0, b_1, \ldots$. Let e_0 be any function in $I(A,B)$. There is e_1 in $I(A,B)$ that extends e_0 and has a_0 in its domain and b_0 in its image. Then there is e_2 that extends e_1 and has a_1 in its domain and b_1 in its image; and so on through e_3, e_4, and so on. Finally define a function e by putting $e(a_i) = e_{i+1}(a_i)$ for each element a_i of A. By construction all elements of A are in the domain of e and all elements of B are in the image of e, and it follows that e is an isomorphism from A to B. We have proved a famous theorem of Cantor:

Theorem 3 *If* A *and* B *are countable dense linear orderings without endpoints, then* A *and* B *are isomorphic.*

There are many adaptations of Fraïssé's idea. One different presentation (though with the same content) uses the idea of a game between two players who take turns to choose elements from the structures A and B. The criteria for the second player to win can be set up so that this player has a winning strategy if and only if there is a back-and-forth system from A to B.

In another adaptation, we require that the domains of the functions in the back-and-forth system all have cardinality $\leq n$ for some positive integer n, dropping the requirements (b) and (c) when i has domain of size n. The existence of a back-and-forth system of this kind corresponds (as in Theorem 2) to the condition that A and B agree in all sentences with quantifier rank $\leq n$, in symbols $A \equiv_n B$. We omit the full definition of $\equiv_n$ here, but we note that any sentence with at most n occurrences of quantifiers in it has quantifier rank $\leq n$, and that in a finite relational signature there are only finitely many pairwise nonequivalent sentences of rank $\leq n$. It follows that a class **K** with finite relational signature is first-order definable if and only if it is closed under $\equiv_n$ for some n. This leads quickly to an answer to Question Two.

Theorem 4 *The equivalence relations A with* Mod(Th(A)) *first-order definable are precisely the finite ones.*

Back-and-forth methods are a model-theoretic generalisation of techniques developed in several areas of mathematics, notably in the study of linear orderings and abelian groups. They also adapt to some languages that are not first-order, and unlike much of first-order model theory, they work as well for finite structures as for infi-

nite ones. This has made them useful tools of theoretical computer science, for example in database theory.

One shouldn't come away from this section with the impression that proving elementary equivalence is a matter of finding a clever model-theoretic technique. Model-theoretic ideas can help to bring to the surface the place where work has to be done, but most proofs of elementary equivalence involve substantial mathematics. For example a problem that Tarski posed in the 1950s, namely whether all free groups with more than one generator are elementarily equivalent, resisted decades of efforts. About half a century after Tarski put the problem, a positive solution was announced by Zlil Sela; besides quantifier elimination, it used a range of techniques from different parts of group theory.

THE COMPACTNESS THEOREM

Almost everything in first-order model theory depends on the Compactness Theorem.

Theorem 5 (Compactness Theorem) *Let L be a first-order language and T a theory in L such that every finite subset of T has a model. Then T has a model.*

We sketch a proof using Hintikka sets as in the entry "First Order Logic." The proof needs a little set theory in the form of infinite cardinals and ordinals. (For the special case in which L has finite or countable signature, one needs only finite numbers.) Suppose the number of formulas of L is κ. We expand the language L to a language L^+ by adding κ new constants, the *witnesses*. Each of the clauses (H1)–(H6) in the definition of a Hintikka set describes a set of requirements on a Hintikka set; for example (H4) describes, for each formula $\phi(x)$ of L^+ and each equation $(s = t)$ where s and t are closed terms of L^+, the requirement that if $\phi(s)$ and $(s = t)$ are in the Hintikka set then $\phi(t)$ is in the Hintikka set. We list all these requirements as $(r_i : i < \kappa)$, in a list of order-type κ, arranging that each requirement appears as r_i for κ-many ordinals i.

Now we define a sequence of theories $(T_i : i \leq \kappa)$, by induction on i, in such a way that three properties hold:

(i) If $i < j \leq \kappa$ then T_i is a subset of T_j.

(ii) Each theory T_i has the property that every finite subset of T_i has a model.

(iii) For each $i < \kappa$ the number of sentences that are in T_{i+1} but not in T_i is finite.

The intention is that T_κ will be a Hintikka set.

We start by putting $T_0 = T$; this ensures that (ii) holds for T_0. If i is a limit ordinal then we take T_i to be $\cup_{j<i} T_j$; since (assuming (i)) every finite subset of T_i is already a subset of some T_j with $j < i$, this ensures that (ii) holds for T_i provided it already holds for each T_j with $j < i$.

Now for each ordinal $i < \kappa$ we define T_{i+1}, assuming that T_i has been defined, in such a way that requirement r_i will be met if it applies. (When r_i doesn't apply, we put $T_{i+1} = T_i$.) The details depend on r_i. We consider some typical cases.

Suppose r_i is the requirement (from (H1)) that if $(\phi \wedge \psi)$ is in the Hintikka set then so are ϕ and ψ. If this requirement applies, that is, if $(\phi \wedge \psi)$ is in T_i, then we take T_{i+1} to be $T_i \cup \{\phi, \psi\}$. It has to be checked that every finite subset of T_{i+1} has a model. Suppose U is a finite subset of T_{i+1}. Put $V = (U \cap T_i) \cup \{(\phi \wedge \psi)\}$. Then V is a finite subset of T_i, so by induction hypothesis it has a model, say A. Since A is a model of $(\phi \wedge \psi)$, it is also a model of ϕ and ψ, and hence it must be a model of U.

Suppose r_i is the requirement (also from (H1)) that if $\neg(\phi \wedge \psi)$ is in the Hintikka set then so is at least one of $\neg\phi$ and $\neg\psi$. If $\neg(\phi \wedge \psi)$ is in T_i then r_i applies. Put $S_1 = T_i \cup \{\neg\phi\}$ and $S_2 = T_i \cup \{\neg\psi\}$. If every finite subset of S_1 has a model then we put $T_{i+1} = S_1$. If not then there is some finite subset U of T_i such that $U \cup \{\neg\phi\}$ has no model. We claim that in this case every finite subset V of S_2 has a model. For consider $U \cup (V \cap T_i) \cup \{\neg(\phi \wedge \psi)\}$, which is a finite subset of T_i and hence has a model, say B, by induction hypothesis. Then B is a model of U and hence a model of ϕ; but B is also a model of $\neg(\phi \wedge \psi)$, so it must be a model of $\neg\psi$ and hence of S_2, as claimed. Hence in this case we can put $T_{i+1} = S_2$.

Suppose r_i is the requirement (from (H5)) that if $\exists x\, \phi(x)$ is in a Hintikka set then so is $\phi(c)$ for some constant c. Suppose that this applies, that is, that $\exists x\, \phi(x)$ is in T_i. By (iii) the number of witnesses used in sentences in T_i is less than κ, and so there must be at least one witness c not used yet. Choose such a c and put $T_{i+1} = T_i \cup \{\phi(c)\}$. Let U be a finite subset of T_{i+1}. Then $(U \cap T_i) \cup \{\exists x\, \phi\}$ is a finite subset of T_i, and so by induction hypothesis it has a model, say C. Since C is a model of $\exists x\, \phi$, there is an element a of C such that $C \vDash \phi[a]$. Let D be the same structure as C, except that $c_D = a$. Then since c appears nowhere in sentences of T_i, D is also a model of $U \cap T_i$. But by choice of c_D it is a model of $\phi(c)$ too, so it is a model of U.

Now suppose we have completed the definition of T_κ as described. Suppose $(\phi \wedge \psi)$ is in T_κ. Then $(\phi \wedge \psi)$ is already in T_i for some $i < \kappa$. Since the requirement referring to $(\phi \wedge \psi)$ is r_j for κ distinct ordinals j, it is r_j for some

$j > i$. So the requirement will have been met when T_{j+1} was defined, and hence T_κ meets the requirement. A similar argument for each of the requirements (H1)–(H6) shows that T_κ meets all these conditions for a Hintikka set. Condition (H7) holds because every finite subset of T_κ has a model. So T_κ is a Hintikka set, and by Metatheorem 16 in the entry "First-Order Logic," it has a model, say A. Since T is a subset of T_κ, A is a model of T, proving the theorem.

Now we can answer Question Five. Let L^+ be the first-order language of the structure $\mathbb{N}$, but with one extra constant c. Let T be $\mathbb{N}$ together with the infinitely many sentences

$$\underline{0} < c, \underline{1} < c, \underline{2} < c, \ldots$$

If U is any finite subset of T, then U includes at most finitely many of the sentences $\underline{n} < c$, and so we can choose a natural number m greater than any of the numbers n for which U mentions $\underline{n}$. Let $\mathbb{N}^+$ be the expansion of $\mathbb{N}$ got by putting

$$c^+_{\mathbb{N}} = m.$$

Then $\mathbb{N}^+$ is a model of U. It follows that every finite subset of T has a model, and hence by the Compactness Theorem there is a model A of the whole of T. Let B be the reduct of A to the language of $\mathbb{N}$. Then $B \equiv \mathbb{N}$ since T contains Th($\mathbb{N}$). But also B contains the element c_A which realizes the type consisting of all the formulas $\underline{n} < x$.

This argument illustrates the model-theoretic idea behind nonstandard analysis.

We can also answer Question Four. In any signature σ, let **K** be the class of finite structures. If A is a structure in **K** and B is a σ-structure elementarily equivalent to A, then A and hence also B are models of a sentence expressing "There are exactly n elements", for some finite n. So B is also in **K**. This shows that **K** is closed under elementary equivalence. But let τ be the expansion of σ got by adding infinitely many new constant symbols $c_0, c_1, \ldots$ and let T' be the theory consisting of all the sentences $(c_i \neq c_j)$ where $i < j$. Since every finite subset of Th(**K**) $\cup$ T' has a model (expanding a structure in **K**), the Compactness Theorem tells us that Th(**K**) $\cup$ T' has a model, and hence that Th(**K**) has an infinite model. Thus **K** is not first-order axiomatisable.

The general setting of our proof of the Compactness Theorem has many adaptations in model theory. A structure is built in a well-ordered sequence of steps, and we list in advance what feature of the structure has to be ensured at each step. Typical examples are the construction of models of a theory that omit certain types, the construction of "existentially closed" models of a theory, and the construction of "two-cardinal" models in which some definable parts are large but other definable parts are kept small.

SUBSTRUCTURES AND ELEMENTARY EMBEDDINGS

If X is a subset of Y then the *inclusion map* from X to Y is the function $i : X \to Y$ such that $i(x) = x$ for each element x of X. Let σ be a signature and A a σ-structure. We say that a σ-structure B is a *substructure* of A, and that A is an *extension* of B, in symbols $A \subseteq B$, if

- dom(B) is a subset of dom(A),
- the inclusion map from dom(A) to dom(B) is an embedding of A into B.

An embedding $e : A \to B$ between L-structures (for some first-order language L) is said to *preserve* a formula $\phi(x_1,\ldots,x_n)$ of L if

$$A \vDash \phi[a_1, \ldots ,a_n] \Rightarrow B \vDash \phi[e(a_1), \ldots ,e(a_n)]$$

for all elements $a_1, \ldots , a_n$ of A. We say that e is an *elementary embedding* if e preserves all formulas of L. If A is a substructure of B and the inclusion map is an elementary embedding, then we say that B is an *elementary extension* of A and that A is an *elementary substructure* of B, in symbols $A \preccurlyeq B$. Always $A \preccurlyeq A$. Also if $A \preccurlyeq B$ and $B \preccurlyeq C$ then $A \preccurlyeq C$. If $A \preccurlyeq B$ then $A \equiv B$.

Two important facts about elementary extension are:

Theorem 6 (*a*) *Let A be a substructure of the L-structure B such that*

> *for every formula $\phi(x_1, \ldots ,x_n)$ of L and all elements $a_1, \ldots , a_{n-1}$ of A such that $B \vDash \exists x_n \phi[a_1, \ldots , a_{n-1}]$ there is b in A such that $B \vDash \phi[a_1, \ldots , a_{n-1},b]$.*
>
> *Then* $A \preccurlyeq B$.

(*b*) (**Union of elementary chains**) *Suppose α is an ordinal and for every ordinal $i < \alpha$ an L-structure A_i is given, so that $A_i \preccurlyeq A_j$ whenever $i < j < \alpha$. Then writing A_α for the union of all the structures A_i, we have $A_i \preccurlyeq A_\alpha$ for all $i < \alpha$.*

Part (a) of Theorem 6 can be used to prove the following important result.

Theorem 7 (Downward Löwenheim-Skolem Theorem) *Let L be a first-order language, A an L-structure, X a set of elements of A, and κ an infinite cardinal number which is (i) at least as great as the number of sentences of L, (ii) at least as great as the cardinality of X and (iii) no*

greater than the cardinality of A. Then A has an elementary substructure of cardinality κ whose domain contains all the elements of X.

This is the result that creates the Skolem paradox. If the axioms of set theory have a model at all, then by this theorem they have a model of countable cardinality, although a sentence expressing "There are uncountably many real numbers" is true in the model!

Part (b) of Theorem 6 is useful for proving a similar result in the other direction. The argument after Theorem 5 adapts to show that every infinite structure has a proper elementary extension. By making repeated elementary extensions and using (b) to take unions at limit ordinals, we reach arbitrarily large elementary extensions.

Theorem 8 (Upward Löwenheim-Skolem Theorem) *Let L be a first-order language, A an infinite L-structure and κ a cardinal number which is at least as great as (i) the cardinality of A and (ii) the number of sentences of L. Then A has an elementary extension of cardinality κ.*

There is also a more algebraic construction that yields a proof of Theorem 8. It involves taking a cartesian product A^I of copies of the structure *A* and defining a homomorphic image in terms of an ultrafilter *D* on the set *I* indexing the copies. The resulting elementary extension A^I/D of *A* is called an *ultrapower* of *A*. Ultrapowers also yield a characterisation of ≡:

Theorem 9 *The following are equivalent, for any two L-structures A and B*:

(a) A ≡ B.

(b) There are a set I and an ultrafilter D on I such that A^I/D is isomorphic to B^I/D.

Many useful properties of ultrapowers spring from the fact that we can make them "highly saturated," that is to say, realizing many types. One can also use the Compactness Theorem and union of elementary chains to build highly saturated structures.

The upward and downward Löwenheim-Skolem Theorems led to a natural question about first-order theories: How far can a first-order theory restrict its models? Assuming that the theory *T* is in a countable language and has an infinite model, we know that it has a model in each infinite cardinality. So the tightest restriction possible is that in every infinite cardinality κ, *T* is κ-*categorical*, that is, it has a model of cardinality κ but all its models of cardinality κ are isomorphic to each other.

Michael Morley published the following theorem in 1965. Its main importance lies in its proof, which revolutionised the techniques of model theory and began the developments reported in the final part of this article.

Theorem 10 (Morley's Theorem) *Let L be a countable first-order language with infinite models, and T a theory in L. If T is κ-categorical for at least one uncountable cardinal κ then T is λ-categorical for all uncountable cardinals λ.*

A theory that is κ-categorical for all uncountable κ is said to be *uncountably categorical*. One major effect of Morley's Theorem was to switch attention from theories to the detailed construction of their models, and a mark of this is that the models of an uncountably categorical theory are now also said to be *uncountably categorical*. A theory that is κ-categorical in all infinite cardinalities κ is said to be *totally categorical*, and so are its models.

By linear algebra, the theory of infinite dimensional vector spaces over a given finite field is totally categorical. A well-known theorem of Ernst Steinitz says that any two algebraically closed fields of the same characteristic and transcendence degree are isomorphic, and it follows that the theory of algebraically closed fields of a given characteristic is uncountably but not totally categorical. An answer to Question Three follows as well. Suppose *A* and *B* are any two algebraically closed fields of the same characteristic. Choose a cardinal κ greater than the cardinalities of both *A* and *B*. By Theorem 8, *A* and *B* have elementary extensions A' and B' of cardinality κ. Then $A' \cong B'$ by uncountable categoricity, and hence $A \equiv A' \equiv B' \equiv B$.

INTERPOLATION AND DEFINABILITY

Let *L* be a first-order language containing a relation symbol *R*. Suppose ϕ is a sentence of *L*. We say that *R* is *upwards monotone* in ϕ if the following holds:

> If *A* and *B* are two *L*-structures which are identical except that $R_A \subset R_B$, and $A \models \phi$, then $B \models \phi$.

Likewise we say that *R* is *downwards monotone* in ϕ if the following holds:

> If *A* and *B* are two *L*-structures which are identical except that $R_A \subset R_B$, and $B \models \phi$, then $A \models \phi$.

In the late middle ages a relation symbol (a 'term' in the medieval terminology) was described as *undistributed* in a sentence if it was upwards monotone in the sentence, and *distributed* if it was downwards monotone in the sentence. For example one can symbolise 'All swans are white' as $\forall x\ (S(x) \rightarrow W(x))$; in this sentence *S* is distributed and *W* is undistributed. ("All swans are white" entails both of the sentences "All Bewick swans are white"

and "All swans are non-red". The medievals said that "swans" is distributed and "white" is undistributed in "All swans are white.")

There is a syntactic test for upwards and downwards monotonicity. We say that a formula ϕ is in *negation normal form* if $\rightarrow$ and $\leftrightarrow$ never occur in ϕ, and $\neg$ never occurs in ϕ except immediately in front of atomic formulas. Every formula is logically equivalent to one in negation normal form and with the same free variables. For example

(5) $$\forall x\,(S(x) \rightarrow W(x))$$

is logically equivalent to

(6) $$\forall x\,(\neg S(x) \vee W(x)).$$

When ϕ is in negation normal form, we say that an occurrence of a relation symbol R in ϕ is *negative* if it has $\neg$ immediately to the left of it, and *positive* otherwise. For example the occurrence of S in (6) is negative and the occurrence of W in (6) is positive. The next theorem is a straightforward consequence of the definition of satisfaction.

Theorem 11 *Let L be a first-order language, R a relation symbol of L and ϕ a sentence of L in negation normal form. If R has no negative occurrences in ϕ then R is upwards monotone in ϕ. If R has no positive occurrences in ϕ then R is downwards monotone in ϕ.*

Since upwards and downwards monotonicity clearly aren't affected when we pass between logically equivalent sentences, Theorem 11 confirms that S is downwards monotone and W is upwards monotone in (5).

Unlike Theorem 11, the next theorem is deep. It is known as *Lyndon's Interpolation Theorem*, after Roger Lyndon who published it in 1959.

Theorem 12 *Let L be a first-order language and ϕ, ψ sentences of L in negation normal form, such that ϕ entails ψ. Then there is a sentence θ of L in negation normal form such that*

- *ϕ entails θ and θ entails ψ,*
- *any relation symbol (apart from =) with a positive occurrence in θ has positive occurrences in both ϕ and ψ, and*
- *any relation symbol (apart from =) with a negative occurrence in θ has negative occurrences in both ϕ and ψ.*

The sentence θ in the theorem is called a *Lyndon interpolant.*

The following immediate consequence of Lyndon's Interpolation Theorem is called *Craig's Interpolation Theorem*. It was proved by William Craig before Lyndon's theorem was known.

Theorem 13 *Let L be a first-order language and ϕ, ψ sentences of L such that ϕ entails ψ. Then there is a sentence θ of L such that*

- *ϕ entails θ and θ entails ψ,*
- *any relation symbol (apart from =) that occurs in θ occurs in both ϕ and ψ.*

The sentence θ here is called the *Craig interpolant.*

We give two applications of these interpolation theorems.

LAWS OF DISTRIBUTION. A *syllogistic sentence* is a sentence of one of the forms $\forall x\,(R(x) \rightarrow S(x))$, $\forall x\,(R(x) \rightarrow \neg S(x))$, $\exists x\,(R(x) \wedge S(x))$ and $\exists x(R(x) \wedge \neg S(x))$, where R and S are different relation symbols. For each syllogistic sentence and each relation symbol in it, Theorem 11 tells us that the relation symbol is distributed or that it is undistributed. A *syllogism* is a sequent of the form $\phi, \psi \vDash \chi$ where each of ϕ, ψ and χ is a syllogistic sentence, three relation symbols are used, and each of them occurs in two sentences. For example one syllogism is

(7) $$\forall x(P(x) \rightarrow Q(x)),\ \forall x(R(x) \rightarrow Q(x)) \vDash \exists x(P(x) \wedge R(x)).$$

This syllogism happens to be invalid, but there are many examples of valid syllogisms.

Late medieval logicians looked for criteria to tell when a syllogism is valid. Two of their criteria were the following, known as the *laws of distribution*:

> If a relation symbol occurs in both premises, then it must be distributed in at least one of them. If a relation symbol occurs in a premise and the conclusion, and is undistributed in the premise, it must be undistributed in the conclusion.

Why do these criteria work? The answer is Lyndon's Interpolation Theorem. We illustrate with the invalid syllogism (7) above, which fails the first distribution law by having Q undistributed in both premises. The same recipe works for all cases.

Suppose for contradiction that (7) is valid. Then after a small rearrangement we have

$$\forall x(P(x) \rightarrow Q(x)) \vDash \forall x(R(x) \rightarrow Q(x)) \rightarrow \exists x(P(x) \wedge R(x)).$$

Convert the sentences to negation normal form, and let θ be a Lyndon interpolant for the resulting sequent. Since Q

occurs only positively on the left and only negatively on the right, it never occurs in θ. So we can introduce a new relation symbol Q', and we have a valid sequent

$$\forall x(P(x) \to Q'(x)) \vDash \theta.$$

Hence by combining the sequents again we infer that the sequent

$$\forall x(P(x) \to Q'(x)) \vDash \forall x(R(x) \to Q(x)) \to \exists x(P(x) \wedge R(x))$$

is valid, and hence so is

$$\forall x(P(x) \to Q'(x)), \forall x(R(x) \to Q(x)) \vDash \exists x(P(x) \wedge R(x)).$$

But it can't be because the two premises have no relation symbols in common and hence can't establish any nontrivial relationship between P and R.

The Port-Royal Logic of Arnauld and Nicole (1662) explains that (7) is invalid because Q "may be taken for two different parts of the same whole" (Rule I in their III.iii). This is vague and not properly justified, but one can see our argument above as a repair of the Port-Royal argument.

EXPLICIT AND IMPLICIT DEFINABILITY. Suppose L is a first-order language, one of whose symbols is a relation symbol R of arity n, and T is a theory in L. One can ask whether R is redundant in T, in the sense that in any model A of T the relation R_A is determined by the rest of A. Here are two different ways of making this notion of redundancy precise. (We write σ for the signature of the language L but with R removed.)

(a) T has a consequence of the form

(8) $\forall x_1 \ldots \forall x_n\, R(x_1, \ldots ,x_n) \leftrightarrow \phi(x_1, \ldots ,x_n))$

where ϕ is a formula of $L(\sigma)$.

(b) Whenever A and B are two L-structures that are models of T, and $A|\sigma = B|\sigma$, we have $A = B$.

When (a) holds we say that R is *explicitly definable* in T, and the sentence (8) is called an *explicit definition* of R in T. When (b) holds we say that R is *implicitly definable* in T.

It turns out that explicit definability and implicit definability are equivalent. (This is for first-order logic; part (b) in the theorem below fails for many other logics.)

Theorem 14 *Let L be a first-order language, R a relation symbol of L and T be a theory in L.*

(a) *If R is explicitly definable in T then R is implicitly definable in T.*

(b) *If R is implicitly definable in T then R is explicitly definable in T.*

Part (a) of the theorem, or more strictly its contrapositive, is known as *Padoa's method*, after Alessandro Padoa who was a researcher in Giuseppe Peano's school around 1900. The proof is straightforward.

Part (b) is called *Beth's Theorem*. It was proved by Evert Beth in 1953, but the following derivation from Craig's Interpolation Theorem is due to Craig. Assume that R is implicitly definable in T. Let T' be T but written with a new relation symbol R' in place of R, and let L^+ be L with R' added. Then the statement that R is implicitly definable in T implies the following:

Suppose an L^+-structure A is a model of $T \cup T'$. Then $R_A \subseteq R'_A$.

We can rewrite this as a sequent:

(9) $T \cup T' \vDash \forall x_1 \ldots \forall x_n\, (R(x_1, \ldots ,x_n) \to R'(x_1, \ldots ,x_n)).$

Add n new constants $c_1, \ldots , c_n$ to the language L^+. By (9), using Metatheorem 10 of the entry "First-Order Logic",

(10) $T \cup T' \vDash R(c_1, \ldots ,c_n) \to R'(c_1, \ldots ,c_n).$

Now by the Compactness Theorem there are finite subsets U, U' of T, T' respectively, such that

(11) $U \cup U' \vDash R(c_1, \ldots ,c_n) \to R'(c_1, \ldots ,c_n).$

Adding sentences to U and U' if necessary, we can suppose that U' is the same as U except that R is replaced by R'. Write ψ for the conjunction of the sentences in U, and ψ' for ψ with R replaced by R'. Then after some rearrangement (11) gives

(12) $\psi \wedge R(c_1, \ldots ,c_n) \vDash \psi' \to R'(c_1, \ldots ,c_n).$

Now apply Craig's Interpolation Theorem to find an interpolant θ such that the following sequents are valid:

(13) $\psi \wedge R(c_1, \ldots ,c_n) \vDash \theta,$

(14) $\theta \vDash \psi' \to R'(c_1, \ldots ,c_n).$

Since R occurs only on the left of (12) and R' occurs only on the right, neither symbol occurs in θ. So by (14) we have the following valid sequent:

(15) $\theta \vDash \psi \to R(c_1, \ldots ,c_n).$

By (13) and (15),

$$\psi \vDash R(c_1, \ldots ,c_n) \leftrightarrow \theta.$$

Now let ϕ be θ but with each constant c_i replaced by the variable x_i. (If necessary we first change the bound vari-

ables of θ so that no quantifier in θ binds any of these variables x_i.) Then by Metatheorem 10 again, we have

(16) $T \vDash \forall x_1 \ldots \forall x_n\,(R(x_1, \ldots ,x_n) \leftrightarrow \phi(x_1, \ldots ,x_n))$

as claimed.

There are many model-theoretic results that are close relatives of the examples above, either in their statements or in their proofs. For example a *preservation theorem* is a theorem stating that some syntactic condition is necessary and sufficient for a formula to be preserved under some algebraic operation. Here follows a typical preservation theorem. We say that a formula ϕ is a $\forall$ formula if ϕ has the form $\forall y_1 \ldots \forall y_n\, \psi$ where ψ is quantifier-free.

Theorem 15 (Łoś-Tarski Theorem) *Let L be a first-order language, $\phi(x_1, \ldots ,x_n)$ a formula of L and T a theory in L. Then the following are equivalent:*

(a) For every embedding $e : A \rightarrow B$ between models of T and all elements $a_1, \ldots , a_n$ of A,

$$B \vDash \phi[e(a_1), \ldots ,e(a_n)] \Rightarrow A \vDash \phi[a_1, \ldots ,a_n].$$

(b) ϕ is equivalent modulo T to a $\forall$ formula $\psi(x_1, \ldots , x_n)$.

EXTENSIONS OF FIRST-ORDER LOGIC

The structures that we defined above are sometimes called *first-order structures* because of their connection with first-order languages. There are other kinds of structure that have analogous connections with other kinds of formal language. Here are two important examples.

(a) Suppose L is a first-order language. Let I be a nonempty set carrying a binary relation R, and for each element of I let A_i be an L-structure. Then, given an element i of I and a sentence ϕ of L, we can ask whether $A_j \vDash \phi$ for all j such that Rij. We can introduce a sentence $\Box\phi$ which counts as true in A_i if and only if the answer is Yes. Indexed families of structures of this type appear in modal logic, temporal logic and some logics of action.

(b) We can consider structures with two domains, where the second domain is a set of subsets of the first domain. Structures of this kind appear in second-order logic, where we have first-order and second-order variables ranging respectively over the first domain and the second domain. They are also found in topological logics, where the second domain contains (for example) a base of open sets for a topology over the first domain.

Sortal structures and languages are a less drastic extension. A sortal signature lists a set of "sorts" and may put restrictions on the function symbols in terms of the sorts. Each sortal structure with this signature has a family of domains, one for each sort. The corresponding sortal language has separate variables for each sort. Sortal structures have some natural mathematical applications. For example a vector space involves a set of vectors and a set of scalars; we can multiply a vector by a scalar, but in general we can't multiply two vectors. So it is natural to work with one sort for vectors and another sort for scalars, and to restrict multiplication so that two vectors can't be multiplied.

If the only changes we make in passing from first-order to sortal are those just described, then the resulting languages behave very much as ordinary first-order languages, and the model theory of sortal structures and languages is hardly distinguishable from ordinary first-order model theory. If we put further restrictions the situation may change; for example if we require that the elements of one sort are exactly the sets of elements of some other sort, then we move into second-order logic.

But even for ordinary first-order structures we need not restrict ourselves to first-order languages. For example we can add to first-order logic

- quantifiers $Q_\kappa x$ that express 'There are at least κ elements x such that …';
- infinitary conjunctions of formulas $\bigwedge_{i<\kappa}\phi_i$ meaning 'ϕ_0 and ϕ_1 and …', and likewise infinitary disjunctions;
- transitive closure operators that express, given a formula $\phi(x,y)$, the property 'There is a finite sequence $a_1, \ldots , a_n$ such that $\phi(x,a_1)$ and $\phi(a_1,a_2)$ and $\phi(a_2,a_3)$ and … and $\phi(a_{n-1},a_n)$ and $\phi(a_n,y)$'.

For example the models of the infinitary disjunction

$$\bigvee_{i<\omega} \text{There are at most } i \text{ elements}$$

are exactly the finite structures. In section 5 we saw that there is no first-order sentence defining this class.

Some of these extensions of first-order logic, using first-order structures, have an elegant and well-developed model theory. But the general truth seems to be that none of them work as smoothly as first-order logic. In 1969 Per Lindström proved some theorems that capture this fact. He showed that if a logic $\mathscr{L}$ contains first-order logic and obeys some of the metatheorems that hold for first-order logic, then $\mathscr{L}$ must be equivalent to first-order logic, in the sense that for every sentence ϕ of $\mathscr{L}$ there is a first-order sentence with exactly the same models as ϕ. For example

Theorem 16 *Let $\mathscr{L}$ be a logic that contains first-order logic. Suppose that $\mathscr{L}$ has the properties:*

(a) Every sentence of $\mathscr{L}$ with infinite models has a model that is at most countable.

(b) If T is a set of sentences of $\mathscr{L}$ and every finite subset of T has a model then T has a model.

Then $\mathscr{L}$ is equivalent to first-order logic.

A fuller account would explain more precisely what is meant by a "logic that contains first-order logic".

STABILITY AND GEOMETRIC MODEL THEORY

Mathematical model theory has become a highly sophisticated branch of mathematics. The items below can be no more than pointers.

Morley's Theorem (Theorem 10) created a new paradigm for model theory. He made it possible to show that a purely model-theoretic condition—uncountable categoricity—on a theory *T* imposes some strong structural features on all the models of *T*. Each such model must contain a *strongly minimal set*, which is a set carrying a dependence relation that behaves like linear or algebraic dependence. In particular the strongly minimal set has a dimension in the same sense as a vector space, and the strongly minimal set is determined up to isomorphism by its dimension and *T*. (Steinitz' Theorem in section 6 is a special case of this fact, since every algebraically closed field is a strongly minimal set.) The rest of the model is very tightly constructed around the strongly minimal set. We can define a function assigning a "rank" to each set definable with parameters in the model; this *Morley rank* is a generalisation of Krull dimension and it allows a very detailed analysis of the model.

Much of the work following on from Morley's Theorem organised itself around one or other of two heuristic principles, known as *Shelah's dichotomy* and *Zilber's trichotomy*. Both of these principles rest on the fact that uncountably categorical theories are "good" in the sense that their classes of models are well-behaved. Both of them have been proved as theorems in certain cases.

For Saharon Shelah, "good" theories form one end of a scale from good to bad. There are several ways that a theory can be bad. One is that it has too many non-isomorphic models to allow any kind of cataloguing by invariants. Another is that it has models that are not isomorphic but are hard to tell apart. Shelah's policy is that at each point of the scale from good to bad, one should aim to maximise the difference between the theories on the "good" side and those on the "bad" side. On the "good" side one should aim to find as much good behaviour as possible, in terms of dependence relations, definability properties, rank functions, and so forth. On the "bad" side one should aim to construct intractable families of models, for example large families of models none of which is elementarily embeddable in any other. Though he applies this principle at all points of the scale, he also identified a *main gap* between the good side and the bad side, and when one speaks of "Shelah's dichotomy" one often has this particular gap in mind.

Shelah created a powerful body of techniques for handling models of theories towards the "good" end of the scale. Together with Morley's work it forms the bulk of *stability theory*. Shelah has also done a large amount of work towards eliminating the restriction to first-order theories. He has suggested several abstract frameworks, for example *excellent classes*, in which there is no counterpart of the Compactness Theorem but some techniques of stability theory still work.

Boris Zilber is more interested in exploiting the "goodness" of the good end of the scale. He is convinced that uncountably categorical structures should all be mathematically interesting, and in fact that they should all be equivalent, up to model-theoretic interpretation, to structures of interest in "classical" mathematics. So he set out to catalogue them, and in the early 1980s he pointed out a natural three-way division of uncountably categorical structures. The division rests on the dependence relation of the strongly minimal set. In the first place there are structures where this relation is "trivial". If it is not trivial, one looks at the lattice of closed sets under the dependence relation. The second class is where this lattice is modular, the third is where it is non-modular. So we have a division of uncountably categorical structures into trivial, modular, and non-modular. The trivial structures are now essentially all known. Modularity turns out to be a strong property, guaranteeing that the structure contains an infinite definable abelian group which exerts a controlling influence; broadly speaking, modular structures exhibit linear algebra.

Zilber conjectured that all non-modular uncountably categorical structures are (up to model-theoretic interpretation) algebraically closed fields. Several pieces of evidence pointed in this direction, notably (i) Angus Macintyre's observation in 1971 that an uncountably categorical field must be algebraically closed, and (ii) observations by Zilber himself and Greg Cherlin that uncountably categorical groups behave remarkably like algebraic groups over an algebraically closed field. *Zilber's trichotomy* is the

division into trivial, modular and non-modular, together with the conjecture that the non-modular structures are algebraically closed fields (up to interpretation). Zilber's trichotomy has been proved to hold for *Zariski geometries*; these are uncountably categorical structures that obey an axiomatisation of the Zariski topology. Ehud Hrushovski saw how to use this fact to solve some major open problems of diophantine geometry, for example proving the Mordell-Lang conjecture for function fields in all characteristics. (His proof in characteristic 0 has since been simplified by replacing the Zariski geometries by differential jet spaces). In 1998 Ya'acov Peterzil and Sergei Starchenko showed that a version of Zilber's trichotomy is true for o-minimal fields.

In 1989 Hrushovski found counterexamples to Zilber's conjecture: uncountably categorical non-modular structures containing no infinite field. At first Hrushovski's examples were mysterious. But Zilber was sure that they must have classical interest, and after some years he discovered structures of Hrushovski's type arising in complex analysis. He also pointed out a close link between Hrushovski's construction and Schanuel's Conjecture in number theory. At the same time Zilber gave examples from complex analysis to illustrate Shelah's excellent classes, thus bringing together two separate lines of research in model theory.

Through the 1990s a body of results converged to show that tools of stability theory are useful in contexts far outside those of 'good' first-order theories. In fact the complete theory Th(A) of a structure A can be largely irrelevant to the application of these tools to A. The need to translate classical descriptions into first-order sentences had always been a practical obstacle to integrating model theory with classical mathematics, and this need seemed to be receding in part. In this context Ludwig Faddeev, after reading the relevant papers presented to the International Congress of Mathematicians in Beijing in 2002, said at the closing ceremony

> Take for instance the sections of logic, number theory and algebra. The general underlining mathematical structures as well as language, used by speakers, were essentially identical.

See also Arnauld, Antoine; Computability Theory; Craig's Theorem; First-Order Logic; Hintikka, Jaakko; Infinitesimals; Infinity in Mathematics and Logic; Logic, History of, overview article; Modal Logic; Montague, Richard; Nicole, Pierre; Peano, Giuseppe; Second-Order Logic; Semantics; Set Theory; Tarski, Alfred.

Bibliography

Arnauld, Antoine, and Pierre Nicole. *La logique ou l'art de penser*. Paris: Gallimard, 1992.

Barwise, Jon, and Solomon Feferman. *Model-Theoretic Logics*. New York: Springer, 1985.

Bell, John L., and Alan B. Slomson. *Models and Ultraproducts*. Amsterdam: North-Holland, 1969.

Börger, Egon, and Robert Stärk. *Abstract State Machines: A Method for High-level Design and Analysis*. Berlin: Springer, 2003.

Buechler, Steven. *Essential Stability Theory*. Berlin: Springer, 1996.

Chang, Chen-Chung, and H. Jerome Keisler. *Model Theory*. 3rd ed. Amsterdam: North-Holland, 1990.

Doets, Kees. *Basic Model Theory*. Stanford, CA: CSLI Publications, 1996.

Dries, Lou van den. *Tame Topology and O-minimal Structures*. Cambridge, U.K.: Cambridge University Press, 1998.

Ebbinghaus, Heinz-Dieter, and Jörg Flum. *Finite Model Theory*. 2nd ed. Berlin: Springer, 1999.

Ebbinghaus, Heinz-Dieter, Jörg Flum, and Wolfgang Thomas. *Mathematical Logic*. New York: Springer, 1984.

Haskell, Deirdre, Anand Pillay, and Charles Steinhorn, eds. *Model Theory, Algebra, and Geometry*. Cambridge, U.K.: Cambridge University Press, 2000.

Hodges, Wilfrid. *Model Theory*. Cambridge, U.K.: Cambridge University Press, 1993.

Manzano, Maria. *Model Theory*. Oxford: Oxford University Press, 1999.

Marcja, Annalisa, and Carlo Toffalori. *A Guide to Classical and Modern Model Theory*. Dordrecht: Kluwer, 2003.

Marker, David. *Model Theory: An Introduction*. New York: Springer, 2002.

Morgan, Mary S. , and Margaret Morrison. *Models as Mediators*. Cambridge, U.K.: Cambridge University Press, 1999.

Pillay, Anand. *Geometric Stability Theory*. Oxford: Clarendon Press, 1996.

Robinson, Abraham. "On the Application of Symbolic Logic to Algebra." In *Proceedings of International Congress of Mathematicians. Cambridge, MA. 1950*. Reprinted in *Selected Papers of Abraham Robinson*. Vol. 1, *Model Theory and Algebra*, edited by H. Jerome Keisler, 3–11. Amsterdam: North-Holland, 1979.

Rothmaler, Philipp. *Introduction to Model Theory*. Amsterdam: Gordon and Breach, 2000.

Tarski, Alfred. "Contributions to the Theory of Models, I." *Indagationes Mathematicae* 16 (1954): 572–581.

Tarski, Alfred, and Robert L. Vaught. "Arithmetical Extensions of Relational Systems." *Compositio Mathematica* 13 (1957): 81–102.

Tatsien, Li, ed. *Proceedings of the International Congress of Mathematicians, Beijing 2002*. Vols. I, II. Beijing: Higher Education Press, 2002. Available from http//:math.ucdavis.edu/ICM2002.

Wilfrid Hodges (2005)

MODERNISM

Modernism was a movement in Catholic religious thought, and particularly in biblical criticism, that developed in the late nineteenth century and spent itself, as a distinctive movement, before World War I. It aimed at bringing Catholic traditions into closer accord with modern views in philosophy and in historical and other scholarship and with recent social and political views. Modernism ran parallel to liberal Protestantism; both tended to reject authority and rigid forms and, in their more extreme versions at least, to aspire to a kind of Christianized rationalism.

The kind of Christology and biblical exegesis undertaken in Germany by D. F. Strauss and in France by Ernest Renan, aided and encouraged by such philosophical currents as positivism and evolutionism, culminated in the late-nineteenth-century attempt to reconcile science with religion and historical criticism with belief. Renan's rejection of the supernatural, combined with his vague evolutionary religiosity, anticipated much that was to be written during the fifteen years following his death in 1892.

Modernism was represented in England by George Tyrrell, Friedrich von Hügel (a friend of Alfred Loisy), and Maude Petre; in Italy by Antonio Fogazzaro, Romolo Murri, and Salvatore Minocchi; and in Germany by Franz Xavier Kraus and Hermann Schnell. However, most of the controversy centered in France, on account of the writings and influence of Loisy, Édouard Le Roy, and Lucien Laberthonnière, who brought to their approach to religion the spirit of contemporary science and philosophy. Loisy, like Renan, rejected the supernatural and explained religion in terms of an immanent rather than a transcendent principle. Le Roy circumvented the difficulties inherent in Catholic dogmas by treating them as pragmatically true. Laberthonnière edited the *Annales de philosophie chrétienne,* a journal that was committed, according to its program, to a rationalistic interpretation of religion, recognizing "the duty to submit to reflection what we believe no less than what we do and think." The review's general policy favored the view that religion is progressively revealed, primitive revelation being only potentially complete. The maneuverings necessitated by the desire to reconcile faith and reason led to some inconsistency and self-contradiction.

From its inception, modernism was in constant trouble with the ecclesiastical authorities, but orthodoxy did not become militant until the accession of Pope Pius X in 1903. In 1907 the papal decree *Lamentabili Sane Exitu,* a collection of sixty-five condemned propositions aimed chiefly at Loisy, and the more general and philosophically grounded encyclical *Pascendi Dominici Gregis,* condemned the modernists' views. The requirement in 1910 that all clerics take the antimodernist oath, known as *Sacrorum Antistitum,* marked the end of the movement as such, although its spirit persisted and prospered in less rebellious forms.

See also Hügel, Baron Friedrich von; Laberthonnière, Lucien; Le Roy, Édouard; Loisy, Alfred; Positivism; Rationalism; Renan, Joseph Ernest; Strauss, David Friedrich.

Bibliography

Riviere, J. *Le modernisme dans l'église.* Paris, 1929.

Vidler, A. R. *The Modernist Movement in the Roman Church.* Cambridge, U.K.: Cambridge University Press, 1934.

Colin Smith (1967)

MODERNISM AND POSTMODERNISM

Modern philosophy is construed as beginning sometime in the Renaissance. A philosophy that seeks new foundations for knowledge was offered as an alternative to that provided by the ancient philosophers. Modern philosophy was presented as starting afresh from new beginnings—turning to nature directly (Francis Bacon), turning to the mind directly (René Descartes), turning to experience directly (Thomas Hobbes). The "quarrel between the ancients and the moderns" resulted from this basic disagreement as to the sources of philosophical knowledge.

Modern philosophy turned away from the past and toward the future, toward the advancement of knowledge, toward human understanding, and toward progress through method or through experience. With the break between the Continental rationalists (Descartes, Nicolas Malebranche, Gottfried Wilhelm Leibniz, and Benedict de Spinoza) and the British empiricists (Hobbes, John Locke, and David Hume) at the end of the eighteenth-century Enlightenment, a new formulation in modern philosophy was called for. Immanuel Kant brought together in his "critical" philosophy the commitments to the analytic exercise of the mind, on the one hand, and the empirical reception through the senses on the other. With Kant, modern philosophy combined the "transcendental unity of apperception" with the "manifold of expe-

rience." Modern philosophy was no longer based on a theory of representation—representation to the mind through reason or representation to the mind through experience—but on the linking of transcendental subjectivity and empirical objectivity. This "doublet," as Michel Foucault came to name it, accounted for a whole new way of philosophizing.

Modernism is distinguished from modern philosophy in that it is linked to certain movements in art and literature that began sometime around the end of the nineteenth century. While drawing upon some similar characteristics of "modern philosophy," modernism in art, literature, and philosophy involved novelty, break with tradition, progress, continuous development, knowledge derived either from the position of the subject or from claims to objectivity, and concomitantly the crisis in knowledge produced by this very dichotomy. Hence in modernism, at the same time that certain theories based knowledge on a centered, transcendental, interpreting subjectivity, and others based knowledge on certain, atomistic, analytic, empirical objectivity, the crisis in knowledge created a sense of uncertainty, paradox, incompleteness, inadequacy, emptiness, and void. Modernism in art and literature involved a shift away from the dichotomies of romanticism and realism to the stream of consciousness, lived and internal time-consciousness, transcendental subjectivity, narrated remembrance and awareness, portrayed speed, mechanisms, objects, and abstractions. Latent content was allowed to penetrate through the surfaces of manifest content. Understanding would have to delve more deeply than surfaces and mere appearances. A phenomenology would be needed in order to inventory the contents of consciousness (Edmund Husserl) or a psychoanalysis to delve the depths of what the mind was really thinking (Sigmund Freud), or a logical positivism would take the alternative tack by excluding all knowledge that cannot be verified logically and empirically (Bertrand Russell, early Ludwig Wittgenstein, A. J. Ayer). Modernism in philosophy involved at each stage the Kantian combination of the empirical and the transcendental, the objective and the subjective, the material and the intellectual—but each time measuring the doublet with weight on one side or the other.

The disintegration of modernism in philosophy was internal. The radical claims of logical positivism excluded all that was of value: metaphysics, aesthetics, axiology, and so forth. The rigorous science of transcendental phenomenology excluded the very existence of what it was investigating. The dualism of creative evolutionism left an irreparable dichotomy between lived experience and objective knowledge. The pragmatism of radical empiricism failed to provide a way to interpret the meanings of experience. The center of modernism in philosophy could not hold because its very foundations were in question. But attempts to retrieve it from itself by the turn to language—ordinary language, analytic philosophies of language, hermeneutics of language, semiologies of language—could not resolve the dilemmas of human existence. Modernism in philosophy faced the absurd, the ambiguous, and the dialectical. And it worked these theories to their limits.

In the mid-1960s philosophy came to look at its epistemological formations and to ask whether the humanisms and anthropologisms of modern philosophy had not circumscribed themselves. Maurice Merleau-Ponty's interrogations were reformulated in Foucault's archaeology of knowledge. The human sciences placed the optimisms and pessimisms of modern philosophy in question by circumventing the theory of "man." Knowledge formations were articulated in terms of multiple spaces of knowledge production and no longer according to a central source or position, or ego, or self, or subject, nor according to a multiplicity of sense-data, objective criteria, material evidence, or behaviors. Knowledge formations crossed disciplines and operated in multiple spaces where questions of structure, frame, margin, boundary, edge, limit, and so on would mark any discursive practice. In other words, knowledge was no longer produced from a center, foundation, ground, basis, identity, authority, or transcendental competency. Knowledge was dispersed, multiple, fragmented, and theoretically varied. Knowledge was no longer based on continuity, unity, totality, comprehensiveness, and consistency. Knowledge began to be understood in terms of discontinuity, difference, dissemination, and differends.

By the early 1970s postmodernism—a term that Daniel Bell used in connection with postindustrial society in the 1950s, that architects appealed to in the 1960s, and that art and literary historians invoked in the 1970s—had still not been invoked in connection with philosophy. Jacques Derrida's grammatology and theory of "difference" in 1967 (building upon Martin Heidegger's account of "the end of philosophy and the task of thinking") turned into a full-fledged deconstruction in the 1970s. Gilles Deleuze and Felix Guattari's notion of rhizomal thinking (as opposed to hierarchical, authorizing arborescent thinking) marked a move against psychoanalytic theories based on Oedipal authority and paternal insistence. Their idea of nomadism placed emphasis on

knowledge, experience, and relations that were not organized around a central concept. J. Kristeva's account of the revolution in poetic language marked the distinction between the semiotic and the symbolic. Where symbolic—scientific, theoretical, phallic, paternal—thinking had pervaded philosophy and science, Kristeva invoked the semiotic as the poetic, fluid, receptacle-like, maternal thinking that has been hidden in modern thought. Yet postmodern was hardly the term that was invoked to describe this kind of philosophizing. Correspondingly, the more restricted study of phenomenology and existentialism in philosophy gave way to the more multiple and diverse theories implicit in Continental philosophy: deconstruction, archaeology of knowledge, semanalysis, schizoanalysis, feminist theory, and so forth. Yet, while poststructuralism (in connection with Foucault, Derrida, Deleuze, Kristeva, et al.) was hailed as the successor to structuralism (Claude Lévi-Strauss, Roland Barthes, Jacques Lacan, Louis Althusser), and existential phenomenology (Heidegger, Jean-Paul Sartre, Merleau-Ponty, Simone de Beauvoir), postmodernism was still not a relevant category in philosophy until well into the 1980s. As time passed, postmodernism and postmodern thought came to take precedence over poststructuralism as the prevalent theoretical formulation.

Postmodern thought means the appeal to differences—differences in theories, differences in formulations, differences in identities. Postmodern thought rejects hierarchies and genealogies, continuities and progress, resolutions and overcomings (*Überwindungen*). Postmodern thought, in fact, cannot operate outside of the modern, for it is itself what can be called an "indecidable." The postmodern signals the end of modernity, but it operates at the same time necessarily within the modern. To claim that the postmodern is outside the modern is to identify it as other than the modern, but that which is outside or other reinscribes the identity of the modern and therefore the postmodern inscription within it. Hence the postmodern both marks places of difference within the modern and calls for an alternative to the modern. The postmodern in any case does not call for the destruction of the modern, not does it seek to deny the modern, since it is necessarily part of the modern.

The postmodern involves the question of the end or limit or margin of what is in question. History, man, knowledge, painting, writing, the modern—each is posed in terms of its end. The end is not a matter of termination or conclusion any more than a matter of goal and aspiration. The postmodern involves, as G. Vattimo notes, a *Verwindung* of modernity—a getting over, a convalescence, a recovering from modernity. This means that modernity is itself placed in question and no longer taken as an unquestioned given. The cracks and fissures in modernity, the places where modernity cannot be fully aware of itself, the moments of unpresentability in the modern—these are the concerns of postmodern thought. As J.-F. Lyotard has noted in his famous *The Postmodern Condition* (1984), the postmodern involves the presentation of the unpresentable in presentation itself—that is, in modernity, the concern was to present something new, something unheard of, something unique, something shocking, something unpresentable. The postmodern involves the presentation of the unpresentable in presentation itself—the formulation of the moments of unpresentability as they mark what is presented. Lyotard calls attention to the role of the "differend" as the place of conflict between two alternative positions. The differend does not belong to either side. It belongs only to the place between, to the gap between the two presentations on either side. This is the postmodern moment—such moments or events with which the modern is distinctively scarred and animated.

See also Ayer, Alfred Jules; Bacon, Francis; Barthes, Roland; Beauvoir, Simone de; Continental Philosophy; Deconstruction; Deleuze, Gilles; Derrida, Jacques; Descartes, René; Existentialism; Foucault, Michel; Freud, Sigmund; Heidegger, Martin; Hobbes, Thomas; Hume, David; Husserl, Edmund; Kant, Immanuel; Kristeva, Julia; Lacan, Jacques; Language; Leibniz, Gottfried Wilhelm; Locke, John; Logical Positivism; Lyotard, François; Malebranche, Nicolas; Merleau-Ponty, Maurice; Phenomenology; Postmodernism; Realism; Renaissance; Romanticism; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Self; Spinoza, Benedict (Baruch) de; Structuralism and Post-structuralism; Wittgenstein, Ludwig Josef Johann.

Bibliography

Derrida, J. *Margins of Philosophy.* Translated by A. Bass. Chicago: University of Chicago Press, 1982.

Foucault, M. *The Order of Things.* Translator anon. New York: Pantheon, 1970.

Kristeva, J. *Revolution in Poetic Language.* Translated by M. Waller. New York: Columbia University Press, 1984.

Lyotard, J.-F. *The Postmodern Condition.* Translated by G. Bennington and B. Massumi. Minneapolis: University of Minnesota Press, 1984.

Lyotard, J.-F. *Toward the Postmodern.* Translated by R. Harvey and M. Roberts. Atlantic Highlands, NJ: Humanities Press, 1993.

Natoli, J., and L. Hutcheon, eds. *A Postmodern Reader.* Albany: State University of New York Press, 1993.

Silverman, H. J., ed. *Postmodernism—Philosophy and the Arts.* New York: Routledge, 1990.

Silverman, H. J. *Textualities: Between Hermeneutics and Deconstruction.* New York: Routledge, 1994.

Vattimo, G. *The End of Modernity.* Translated by J. R. Snyder. Baltimore: Johns Hopkins University Press, 1988.

Hugh J. Silverman (1996)

MOIRA/TYCHĒ/ANANKĒ

All three Greek words denote causal powers that are beyond the reach of human control, and hence were often personified as goddesses.

The word "moira" means a share, part, or portion, and by derivation, the fate allotted to a person. In mythological contexts, it was personified either as a single goddess or, as in Hesiod's *Theogony* and in the myth of Plato's *Republic* X, as a group of three goddesses (Clotho, Lachesis, Atropos). Moira or the Moirai determine the fate of individuals by "spinning" the thread of one's life. The word "moira" sometimes euphemistically refers to death, as the fate of all humans. In other contexts it refers to one's rank or distinction or to the positive abilities allotted by the gods, such as poetic inspiration. In Stoic determinism, it is used in relation to universal fate (*heimarmenē*).

The noun "tychē" (fortune) is related to the verb "tynchano" (happen, befall). Tychē was taken to be the cause of chance events—events that one could not or did not calculate and that do not fit into a regular pattern. While moira determines one's course of life as a whole, tychē tends to be responsible for singular events of varying importance. The connotations of the word were originally more positive, but by Hellenistic times it regularly had the pejorative meaning of blind, impersonal, arbitrary chance. In philosophical contexts it is most often contrasted with rational choice and goal-driven action. Plato, in the *Laws* X, grouped tychē together with the mechanistic force of nature and opposed it to the rational, purposeful activity of a cosmic god. Aristotle, in the *Physics* II.5–7, classified tychē under spontaneity (*automaton*) and defined it as an accidental and indeterminable cause in the sphere of purposeful actions involving rational choice. In other words, tychē for Aristotle is the cause of events that might have been the outcome of rational human choice but in fact are not.

The word "anankē" originally referred to an external constraining force, and from this meaning it obtained the more abstract meaning of logical and physical necessity during the pre-Socratic period. It is often represented as the ultimate power with which even the gods must comply. In Parmenides' *Aletheia*, the personified Anankē guarantees that Being is unchangeable and immobile, and "holds [Being] in the bonds of a limit" (Diels and Kranz 1954, B8.30), while in the *Doxa* she keeps the starry heaven enchained (B11.6). In Empedocles' writings, Anankē's oracle sets the punishment of those who commit the ultimate sin of bloodshed (B115.1). In the myth of Plato's *Republic* X, Anankē is the mother of the three Moirai. Her function is primarily cosmological in that she holds a spindle whose movement stands for the celestial motions. In Plato's *Timaeus*, anankē is the regular but nonteleological causal force inherent in the physical realm. Insofar as the physical properties of the elements can be put into the service of the purposeful activity of reason, anankē becomes the auxiliary cause (*sunaition*) in teleological causation. Aristotle's distinction between simple and hypothetical necessity in the *Physics* II.9 shows clear traces of the conception of anankē in the *Timaeus*.

See also Aristotle; Being; Causation: Metaphysical Issues; Death; Empedocles; Parmenides of Elea; Plato; Pre-Socratic Philosophy.

Bibliography

Cornford, Francis Macdonald. *Plato's Cosmology*. London: Routledge and Kegan Paul, 1937.

Diels, Hermann, and Walther Kranz. *Die Fragmente der Vorsokratiker*. 7th, rev. ed. Berlin: Weidmann, 1954.

Dietrich, Bernard Clive. *Death, Fate, and the Gods: The Development of a Religious Idea in Greek Popular Belief and in Homer*. London: Athlone Press, 1965.

Schreckenberg, Heinz. *Ananke: Untersuchungen zur Geschichte des Wortgebrauchs*. Munich, Germany: C. H. Beck'sche Verlagsbuchhandlung, 1964.

Gábor Betegh (2005)

MOLESCHOTT, JACOB

(1822–1893)

Jacob Moleschott, a physiologist and philosopher often regarded as the founder of nineteenth-century materialism, was born in Holland. After studying at Heidelberg, Moleschott practiced medicine in Utrecht. He later became lecturer in physiology at Heidelberg. The controversial doctrines expressed in his book, *Der Kreislauf des Lebens* (The circuit of life; Mainz, 1852), and the materialistic tendencies of his teaching forced him to move to

Zürich. He later became professor of physiology at Rome, where his lectures were popular and his important research on diet earned him respect and many honors.

Materialism at that period was a philosophical trend with political, social, and scientific implications. The state-controlled German universities had produced an official philosophy (a watered-down Hegelianism) that was used as a defense against social reform and as a shield for religion or the spiritual life. Certain important scientists held conservative views about the role of science. The biologist Rudolf Virchow, for example, believed that all speculation about consciousness should be left to the church or even to the state. The German materialists, attempting to free scientific inquiry from such control, saw these conventional philosophical tendencies as obstructing intellectual and social progress.

PHILOSOPHIC MONISM

Moleschott's *Der Kreislauf des Lebens* went through many editions and helped to spur the materialist movement. The book was directed against Justus von Liebig's theologizing views as he had expressed them in his *Chemischen Briefen.* Liebig had especially objected to Moleschott's famous statement epitomizing materialist monism: "No thought without phosphorus." The German materialists of this period criticized dualists as being engaged in a system of philosophic double-entry bookkeeping.

Moleschott maintained, as did Ludwig Büchner, that force and matter were inseparable. Force cannot be viewed in an Aristotelian way, nor teleologically, nor as a vital force. It is not an entity separate from a material substratum, but is rather "one of its eternal indwelling properties." Matter cannot occur or be conceived without force, and vice versa; "A force unconnected with matter, hovering loose over matter, is an utterly empty conception."

Thus, any materialism attributing existence to matter independently of force was rejected. Moleschott maintained that to call his theory materialistic in this sense would be as wrong as to call it spiritualistic: "I myself was well aware that the whole conception might be converted, for since all matter is a bearer of force, endowed with force or penetrated with spirit, it would be just as correct to call it a spiritualistic conception." On the other hand, once the restriction of the term *material* to "dead matter" is given up, Moleschott appears materialistic indeed. He regarded the brain as the source of consciousness and emphasized physical conditions as the major determinants of human life. He was fascinated by circular processes, such as the miner digging lime phosphate from the earth, and the peasant later fertilizing his field with the same chemical. Life circulates through all parts of the world, and with life goes thought.

As was also typical of the materialists of the time, Moleschott emphasized the doctrine of the conservation of matter. This notion, he held, was discovered by the eighteenth-century encyclopedists. Recent science had confirmed it, and future science had to be built upon it. Chemistry is the basic science, and the solution to social questions depends on our discovering the proper way to distribute the matter with which thought and will are bound up. A rigid determinism was emphasized: "Natural law is the most stringent expression of necessity."

THEORY OF KNOWLEDGE

Moleschott inveighed against the Kantian thing-in-itself and emphasized the importance of what things could be known as rather than what they are alleged to be. All knowledge, he maintained, presupposes someone who knows and, thus, a relation between the object and the observer. The observer could be an insect or other creature; there is no restriction to man. All existence is by means of qualities; there is no quality that exists other than through a relation. In the case of a man's perceiving a tree, "it is just as necessary for the tree as for the man that it stands to him in a relation that manifests itself by the impression upon his eye."

Moleschott maintained a certain relativism, but also a certain objectivism: "Steel is hard as opposed to soft butter, ice is only cold to the warm hand, trees only green to a healthy eye." He argued that a vorticella with an eye having only a cornea must receive different representations of objects than a spider, which has a more complex eye with lenses. Yet, "Because an object is [exists] only through its relation to other objects, for instance, through its relation to the observer, because the knowledge of the object resolves itself into the knowledge of their relations, all my knowledge is an objective knowledge." Although there are difficulties in understanding Moleschott's doctrine here, it appears to have a strong family resemblance to recent objective relativism.

ETHICS

The German materialists were frequently criticized for promulgating doctrines subversive of received morality, especially theologically sanctioned morality. In general, they did protest against duty-centered, puritanical views of morality and adopted a kind of utilitarian hedonism. However, they did not advocate a continuing round of sensual pleasures. Moleschott argued that even a mis-

guided hedonism was socially less dangerous than some other views of morality: "The erroneous theory of seeking after pleasure will scarcely find half as many disciples, as the rule of priests of all shades had claimed unfortunate victims."

As was true of other contemporary materialistic theories, many of Moleschott's doctrines that once aroused immense wrath seem relatively mild today. His insistence that scientific inquiry is relevant to the solutions of many problems is now commonplace, but it caused shudders in the nineteenth century. The materialists' struggle against giving theological answers to scientific questions seems to have been largely successful.

See also Büchner, Ludwig; Encyclopédie; Hedonism; Materialism; Natural Law.

Bibliography

ADDITIONAL WORKS BY MOLESCHOTT

Die Physiologie der Nahrungsmittel. Darmstadt: C.W. Leske, 1850.

Physiologie des Stoffwechsels in Pflanzen und Thieren. Erlangen: F. Enke, 1851.

Lehre der Nahrungsmittel. Erlangen: F. Enke, 1853. Translated by Edward Bronner as *The Chemistry of Food and Diet.* London: Houlston and Stoneman, 1856.

Eine physiologische Sendung. Giessen: E. Roth, 1864.

Für meine Freunde. Lebenserinnerungen von Jacob Moleschott. Giessen: E. Roth, 1894.

WORKS ON MOLESCHOTT

Lange, Frederick A. *Geschichte des Materialismus,* 2 vols. Iserlohn, Germany, 1866. Translated by E. C. Thomas as *The History of Materialism,* 3 vols. London: n.p., 1877–1892. Gives an account of Moleschott.

Merz, John T. *A History of European Thought in the Nineteenth Century.* Edinburgh: Blackwood, 1903. Discusses Moleschott briefly and contemporary issues in considerable detail.

Rollo Handy (1967)

MOLINA, LUIS DE
(1535–1600)

Luis de Molina, S.J., was a central figure in the sixteenth-century renaissance of scholasticism on the Iberian peninsula. He was born in Cuenca, Spain, in 1535. At eighteen he entered the Jesuit order. He studied and later taught at Coimbra and Évora in Portugal. In 1583, he left his academic post to devote himself to writing. He spent the next fifteen years in Cuenca, Lisbon, and Évora. He died on October 12, 1600, shortly after being called to take a chair in moral theology at the newly established Jesuit University in Madrid.

Molina's best known work, *Liberi arbitrii cum gratiae donis, divina praescientia, providentia, praedestinatione et reprobatione concordia* (The compatibility of free will with the gift of grace, divine foreknowledge, providence, predestination, and reprobation) was first published at Lisbon in 1588; a second, expanded edition was published at Antwerp in 1595. He also authored a three volume commentary on Part One of St. Thomas's *Summa Theologiae*, titled *Commentaria in primam divi Thomae partem*, published at Cuenca in 1592. Although these works, especially the *Commentaria*, range broadly over theological and philosophical topics, critical attention focused on Molina's theory of *middle knowledge* (*scientia media*), which was formulated to reconcile God's comprehensive foreknowledge and providence with a strongly indeterministic conception of human free will.

According to the tradition shared by Molina and his rivals, at the moment of creation, God has perfect and infallible foreknowledge of everything that will happen in the created world. The tradition also maintains that God's knowledge is not like that of a passive observer. Rather, he specifically intends or knowingly permits everything that takes place, and he arranges created causes and exercises causal influence sufficient to bring about his creative plan to the last detail. God's foreknowledge, consequently, is to be explained in terms of his providence. He knows what will happen in the created world by his knowledge of his own decrees, together with his knowledge of what follows from those decrees, either directly or through the mediation of created causes. The fundamental difference between the positions of Molina and his adversaries lies in where they locate the main resources for God's providential foreknowledge. The Molinists emphasize the role of God's practical knowledge, his adversaries emphasize the role of his voluntary decrees.

The tradition distinguishes God's *prevolitional* knowledge, which he has independently of his will from his *postvolitional* knowledge, which depends on his free decrees. A majority of traditional philosophers and theologians maintain that God's knowledge of metaphysically necessary truths exhausts his prevolitional knowledge. On this view, God's knowledge of necessary truths (which Molina calls *natural knowledge*) is identified with his prevolitional knowledge, and his knowledge of contingent truths is identified with his postvolitional knowledge (which Molina calls *free knowledge*). Call this the standard view. It is also commonly held that proposi-

tions concerning what is metaphysically possible are themselves metaphysically necessary.

Consequently, God's knowledge of these propositions is part of his natural knowledge. According to the standard view, God is able to bring about any metaphysically possible state of affairs. God's creative activity can thus be described as (1) deciding which metaphysically possible states of affairs will be actual, and (2) making a causal contribution sufficient to actualize those states of affairs. It must be emphasized that, on the standard view, God's causal activity completely determines what is going to take place in the created world.

According to Molina, however, the *free* choices of a rational creature are not causally or logically necessitated either by God's causal activity or by the operation of created causes, including the beliefs, desires, character, and dispositions of the agent. For any free choice a rational creature makes in a fully specified set of circumstances, it is metaphysically possible that that creature makes a different choice in those very same circumstances. So, on Molina's view, God's natural knowledge of metaphysical possibilities, together with his knowledge of his own causal activity, cannot provide him with foreknowledge of the free choices that his creatures will make. Therefore he holds that an essential component of the theory of divine foreknowledge and providence is God's knowledge of a special class of propositions called *conditional future contingents*. These propositions concern what choices rational creatures *would* freely make in any of possible circumstances in which they may find themselves. Molina contends that God must have knowledge of these propositions *prior to* his creative decrees to exercise providence over the world, otherwise he would be unable to guarantee that his creation conforms to his providential design in all its detail.

Molina calls God's knowledge of conditional future contingents *middle knowledge*, because it stands between his natural knowledge of what is merely possible and his free knowledge of what is actually, though contingently, the case. Like his natural knowledge, but unlike his free knowledge, God's middle knowledge is prevolitional. Like his free knowledge, but unlike his natural knowledge, the objects of God's middle knowledge are contingent truths. According to Molina, then, God's providence and foreknowledge is a function of (1) his prevolitional natural knowledge of the possible arrangement of created causes, (2) his prevolitional middle knowledge of the contingent choices free creatures would make in each of these possible arrangements, and (3) his postvolitional free knowledge of the way in which he has decided to arrange created causes. This is how Molina reconciles God's providence and foreknowledge with his strongly indeterministic conception of freedom. In addition, Molina and his followers maintain that the theory of middle knowledge has fruitful applications in explaining a broad range of philosophical and theological issues such as the efficacy of grace, predestination and reprobation, petitionary prayer and prophecy.

Perhaps the weakest point in the Molinist theory is his explanation of *how* God can know what free creatures would do in various possible circumstances, given his strongly indeterministic conception of freedom. Critics maintain that there can be no basis for God's perfect and infallible knowledge of the choices that free creatures would make, given that these choices are not logically or causally determined by the activity of God or the operation of secondary causes. Unlike other defenders of middle knowledge (such as Suarez), Molina refuses to appeal to the determinate truth of conditional future contingents to explain God's knowledge of them. In fact, Molina follows Aristotle in maintaining that contingent propositions concerning the choice a free creature would make in specified circumstances do not have determinate truth prior to the creature making that choice in those circumstances.

Molina's explanation of God's knowledge of conditional future contingents involves what later came to be called *supercomprehension*. Given the indeterminacy of future contingent propositions, Molina believes that God's certain and infallible knowledge of them is due to the cognitive perfection of the knower. For Molina and his contemporaries, all of God's knowledge is ultimately grounded in his self-knowledge, either knowledge of his own essence or knowledge of his decrees. God's middle knowledge is grounded in his knowledge of his own essence, in which all possible creatures are eminently contained.

By perfectly comprehending his own essence, according to Molina, God is able to infallibly cognize the choices each possible creature would make in any possible circumstance in which they may find themselves, even though these choices are metaphysically indeterminate. Supercomprehension, on Molina's view, is a mode of cognition possible only for an infinite intellect with respect to finite creatures. Molina's readers, including those who defend middle knowledge (e.g., Suarez), are nearly unanimous in their rejection of the theory of supercomprehension. However, for Molina, the theory has the advantage over its competitors in explaining why God cannot have prevolitional knowledge of the choice he himself would make in various possible circumstances. Such knowledge, Molina believes, would destroy divine freedom.

The publication of Molina's *Concordia* aroused bitter controversy between the Molinists and the defenders of the standard view, primarily Domingo Bañez, Diego Alvarez, and other members of the influential Dominican order. The Dominicans accused the Molinists of undermining God's sovereignty over the created world by maintaining that God has no direct control over the choices of free creatures. The Molinists accused the Dominicans of destroying human freedom and making God morally responsible for sinful actions. The Vatican, anxious to avoid another divisive clash over the issues of grace and free will, called the factions to Rome to examine the matter. In 1597, Pope Clement VIII convened the *Congregatio de auxiliis*, and over the next ten years the Molinist position was scrutinized in eighty-five hearings and forty-seven debates. Initially things did not go well for the Molinists, and Molina died fearing that the censure of his views was imminent. However, the theory of middle knowledge ultimately escaped condemnation. In 1607, Pope Paul V closed the proceedings. He allowed both parties to continue teaching their doctrines and ordered the sides to refrain from accusing each other of contradicting the faith.

Though Molina's best known contributions are to speculative theology, he also authored a seven-volume treatise in moral and political philosophy entitled *De Justitia et jure* (published posthumously at Venice in 1614). This work discusses the source of legitimate political authority, the permissibility of slavery, and the justification of war, as well as economic issues such as taxation, free markets and monetary policy.

See also Báñez, Dominic; Foreknowledge and Freedom, Theological Problem of; Philosophy of Religion, History of; Scientia Media and Molinism; Suárez, Francisco; Thomas Aquinas, St.

Bibliography

WORKS BY MOLINA

Liberi Arbitrii cum Gratiae Donis, Divina Praescientia, Providentia, Praedestinatione et Reprobatione Concordia, edited by J. Rabeneck. Oña and Madrid: 1953.

On Divine Foreknowledge: Part IV of the Concordia. Translated by Alfred J. Freddoso. Ithaca, NY: Cornell, 1988.

Neue Molinaschriften, edited by Friedrich Stegmüller, in *Beiträge zur Geschichte der Philosophie und Theologie des Mittelalters*, Band 32. Münster, Germany: 1935.

WORKS ABOUT MOLINA

Adams, Robert M. "Middle Knowledge and the Problem of Evil." In *The Virtue of Faith and Other Essays in Philosophical Theology*, edited by Robert M. Adams. New York: Oxford University Press, 1987.

Bañez, Domingo, et al. *Apologia en defensa de la doctrina antigua y catholica por los maestros dominicos de la Provincia de España contra las afirmaciones contenidas en la Concordia de Luis de Molina sobre la gracia, praesciencia divina, providencia, predestinacion y reprobacion*. In *Domingo Bañez y las controversias sobre la gracia*, edited by Vincente Beltran de Heredia. Salamanca, Spain: 1968.

Flint, Thomas. "Two Accounts of Providence." In *Divine and Human Action*, edited by Thomas V. Morris. Ithaca, NY: Cornell University Press, 1988.

Freddoso, Alfred J. "Introduction." In *On Divine Foreknowledge: Part IV of the Concordia*, 1–81. Translated by Alfred J. Freddoso. Ithaca: Cornell University Press, 1988.

Opera Omnia, Tomus XI, edited by Carol Berton. Paris: Vivès, 1858.

Pegis, Anton. "Molina and Human Liberty." In *Jesuit Thinkers of the Renaissance*, edited by G. Smith. Milwaukee, WI: Marquette, 1939.

Michael V. Griffin (2005)

MOLINA GARMENDIA, ENRIQUE
(1871–1962)

Enrique Molina Garmendia, the Chilean spiritualist philosopher, was born at La Serena, Chile. After several years of practicing law and teaching on the faculty of the Liceo de Chillán, he became the first rector of the University of Concepción in 1919. He was one of the leading members of the generation of Latin American intellectuals who, under the influence of William James, Henri Bergson, and the French spiritualists, reacted against the positivism that had dominated the political and cultural life of Latin America for half a century.

Throughout the eleven books that he published between 1912 and 1952, Molina was basically concerned with philosophical anthropology and with offering "an interpretation of [the human spirit], acceptable even to the skeptics, formulating a consideration of the spiritual in human life where it is constructive and creative, and where it is involved with ethical exigencies" (*De lo espiritual en la vida humana*). This concern raised the problem of the nature of consciousness and its relation to being, as well as the problem of the origin and status of values in the natural order.

Rejecting both idealistic and materialistic ontologies, Molina maintained the priority of being over consciousness, although he noted that the emergence of the latter within natural processes indicates the potentiality for consciousness within being. Following the German philosopher Edmund Husserl, Molina declared that being

and consciousness are integrally united within experience. The priority of being "is affirmed, because it is first *lived* by consciousness as a totality of which consciousness forms a part" (*De lo espiritual en la vida humana*). Molina restated René Descartes's basic premise as "I think, therefore I exist *and Being exists*." An adequate conception of being must incorporate both the subjective and the objective poles of experience.

It is in man that spirit has become most fully actualized. Closely associated with consciousness, spirit is the locus of values and is characterized by the freedom that makes activity leading toward the realization of value possible. The realm of the spirit embraces all the realms that are the result of human creativity—morality, religion, the sciences, the arts, "all the work of enlightened intelligence." Spirit is that element within each of these realms which aspires to be, which strives to perfect itself and to go beyond itself. Reason is the highest structure of spirit. Through reason, the presence of being is recognized, mere automatic functioning of the organism is overcome, and the horizons of consciousness are opened to the possibilities for creative advance.

See also Bergson, Henri; Descartes, René; Husserl, Edmund; James, William; Latin American Philosophy; Philosophical Anthropology.

Bibliography

WORKS BY MOLINA

De lo espiritual en la vida humana (Concerning the spirit in human life). Concepción, 1936.

Confesión filosófica (Philosophical confession). Santiago: Nascimento, 1942.

Tragedia y realización del espíritu. Del sentido de la muerte y del sentido de la vida (Tragedy and realization of the spirit. The meaning of death and life). Santiago: Nascimento, 1952.

WORKS ON MOLINA

Enríquez Frodden, Edgardo. *Proyección del pensamiento y de la personalidad de Concepción* Chile: Universidad de Concepción, 1972.

Lipp, Solomon.*Three Chilean Thinkers*. Waterloo, ON: Wilfrid Laurier University Press, 1975.

Salvat Boloña, Pablo. *Visión del hombre y visión de América según Enrique Molina G*. Santiago, 1982.

Vidal Muñoz, Santiago. "Apuntes sobre la filosofía en Chile." *Cursos y conferencias* 48 (272) (1956): 39–60.

Fred Gillette Sturm (1967)
Bibliography updated by Michael Farmer and Vladimir Marchenkov (2005)

MONAD AND MONADOLOGY

The Greek term *μονάς*, from which the word *monad* is derived, means a "unit" or a "one." In Pythagorean writings it is the unity from which the entire number system, and therefore—as a consequence of the doctrine that "everything is number"—all things, are derived. Through Plato, who applied the Pythagorean term to the Ideas or Forms (*Philebus* V, 15B), it entered the tradition of Neoplatonism and Christian Platonism to mean a simple, irreducible, self-determining entity whose activity is the source of all composite beings. In this sense it was sometimes used to designate God as the simple source of all being and sometimes to signify the simplest irreducible entities in the created order out of whose harmonious action all existence is compounded.

A monadology is a metaphysical system that interprets the world as a harmonious unity encompassing a plurality of such self-determining simple entities. The term was first used in the early eighteenth century of the metaphysics of Gottfried Wilhelm Leibniz.

In its modern meaning since Leibniz, a monad is held to be (1) a simple, irreducible, and sometimes indestructible entity; and (2) the minimal unity into which the cosmos and all composite things in it can be resolved; yet (3) containing within itself, in contrast to material atoms, powers and relations of which it is itself the source. It is therefore conceived after the analogy of a mind or a *res cogitans* rather than a material substance. It is held to constitute, along with other monads, an all-inclusive unity or harmony of the cosmos as a whole.

A monadology may thus entail a theory of cosmic harmony, based upon a mathematical or scientific functionalism or upon a psychology of intersubjective relations, as well as a theory of relations, in which the relations constituting this cosmic harmony are brought into being through monadic action, although they do not affect the monads or organizations of monads that are the objects of the acts (Leibniz's perceptions and Alfred North Whitehead's prehensions are examples of such relations).

This intermonadic harmony may itself he regarded as a unity, or cosmic Monad, and this view may involve pantheism or a theistic theory of creation. The relation of the minimal monads to the supreme Monad is one of mirroring rather than being a part of; since the supreme Monad must itself be simple, each monad may be held to be a finite (unclear and indistinct) reflection of the attrib-

utes of the supreme Monad. (The metaphors of mirroring, of echoing, and of the infinite circle whose center is everywhere have commonly been used in monadologies.)

Monadologies may disagree in their fundamental categories. Monads are active substances and, therefore, also processes; Leibniz attempted, but with incomplete success, to unite a logical and a psychological analysis of the monad by applying the notions of intensionality and extensionality. The finite monads may be of a temporal nature; the cosmic order may be either eternal or temporal, or—as Whitehead and Charles Hartshorne held—both eternal and temporal. The finite monads themselves may be eternal changeless souls (John McTaggart). The cosmic harmony may be thought of as a divine Person or merely as the unitary society of monads.

In the history of modern monadologies, three conceptions have been operative: the Christian Platonist tradition of the soul as a simple substance possessing self-certainty in immediate unity (Augustine, *De Trinitate,* IX, 3; X, 9, 10); the Neoplatonic-Stoic conception of the One that is essentially represented in each of its parts; and a spiritualized form of atomism ultimately derived from this Neoplatonic-Stoic conception. The first tradition, mediated by Boethius, the Franciscans, and other medieval Platonists, became prominent in the seventeenth century in Francisco Suárez, René Descartes, and others. The second tradition emerged in the Renaissance in the concepts of the microcosm and macrocosm after a long history during which the Stoic doctrine of the Logos had been combined with the Neoplatonic theory of the One and the subordinate intelligences. This tradition involved the principle of plenitude, according to which the universe can achieve its maximal being only when God multiplies or reduplicates his nature in every created being. This principle was suggested by Meister Eckhart and explicated by Nicholas of Cusa in his doctrine of the coincidence of maximum and minimum in God. Giordano Bruno developed the principle of plenitude into a theory of material monads as spherical atoms that are spiritual reflections of the Divine Nature (*De triplice minimo et mensura … Libri quinque,* 1591; *De monade, numero, et figura Liber,* 1591).

Leibniz's concept of monad is variously ascribed to Bruno, Henry More, or Franciscus Mercurius van Helmont, all of whom had made use of the term. But the terms *Monas* and *monadica* appear in the early papers of Leibniz, written long before he had come to know any of these thinkers or had developed his mature metaphysics.

Leibniz's monadology involves a harmonious universe composed of an infinite number of monads, each of which was an infinite series of perceptive acts defined by a unique point of view or a unique law of series; each such law, in turn, was a particular finite combination of the perfections of God expressed in his creation. Leibniz presented a succinct but incomplete account of this system in his *Principles of Nature and of Grace* and the so-called *Monadology,* both written in 1714; he then devoted the last twenty years of his philosophical activity to a defense and amplification of his monadology through various papers and a vast correspondence. His system and that of Whitehead, who ascribed greater spontaneity and creativity to the monads and interpreted them as mind-like entities of limited duration, are the most detailed modern monadologies.

Trained in the Leibniz-Wolff tradition, Immanuel Kant wrote *Physical Monadology* in his precritical period (1756), in which the monads were treated as sources of motion in a Newtonian space. In the *Critique of Pure Reason* (1781), Kant called his second antinomy "the dialectic principle of monadology" (1st ed., p. 442). This antinomy is directed at the metaphysical claims for a monadology made by the Wolffian school. In their development of a realistic, spiritualistic metaphysics, Johann Friedrich Herbart, Hermann Lotze, and Gustav Theodor Fechner developed monadologies on a Kantian basis. In his third *Essai de critique générale* (Paris, 1859), and in *La nouvelle monadologie* (Paris, 1899), Charles Renouvier built a monadology upon his relativized interpretation of Kant, making the highest attainable harmony in "the best of all possible worlds" depend upon the freedom of human monads or persons. In contrast to this relativized monadism, Edmund Husserl, in his *Cartesian Meditations* (1929–1931), suggested a monadic completion of his transcendental phenomenology, describing a type of "indirect experience that possesses its own modes of verification" within one's own monadic experience and that also provides "the transcendental base" for an objective natural order; implied in this is a "sphere of monadological intersubjectivity." Other recent monadologies include Dietrich Mahnke's attempt to reconcile Leibniz's monadology with recent science and philosophy; H. Wildon Carr's *Theory of Monads* (London, 1922), influenced by the British personalistic tradition; and William Stern's hierarchical system of persons and things, inspired by Benedict de Spinoza, Fechner, and Lotze.

See also Augustine, St.; Boethius, Anicius Manlius Severinus; Bruno, Giordano; Descartes, René; Eckhart, Meister; Fechner, Gustav Theodor; Herbart, Johann Friedrich; Husserl, Edmund; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Macro-

cosm and Microcosm; McTaggart, John McTaggart Ellis; More, Henry; Neoplatonism; Nicholas of Cusa; Plato; Platonism and the Platonic Tradition; Renaissance; Renouvier, Charles Bernard; Spinoza, Benedict (Baruch) de; Stern, Louis William; Suárez, Francisco; Whitehead, Alfred North.

Bibliography

Discussions of monadologies are to be found in Heinz Heimsoeth, *Atom, Seele, Monade. Historische Ursprünge und Hintergründe von Kant's Antinomie der Teilung* (Wiesbaden, 1960). W. Cramer, in *Die Monade* (Stuttgart, 1954), begins with Kant's antinomy and treats him in terms of intermonadic relations. P. F. Strawson subjects a Leibnizian monadology to critical analysis in *Individuals: An Essay in Descriptive Metaphysics* (New York, 1963), pp. 114–133.

See also Dietrich Mahnke, *Eine Neue Monadologie* (*Kantstudien,* Erganzungsheft 39; Berlin: Reuther and Reichard, 1917) and *Unendliche Sphäre und Allmittelpunkt* (Halle: M. Niemeyer, 1937) by the same author; Edmund Husserl, *Cartesianische Meditationen und Pariser Vorträge,* edited by S. Strasser (The Hague, 1950), especially sections 55–56; William Stern, *Person und Sache,* 3 vols. (Leipzig: J. A. Barth, 1906–1924), Vol. I (Leipzig, 1906); and A. N. Whitehead, *Process and Reality* (New York: Macmillan, 1929).

L. E. Loemker (1967)

MONISM AND PLURALISM

How many things are there? Or how many kinds of thing? Monism is the doctrine that the answer to one or other of these questions is "Only one." Opposed to monism is the doctrine of pluralism, which is that there are many kinds of thing, or that there are many things. It will be apparent, on reflection, that this weaker form of pluralism, that there are many things, is quite consistent with the weaker form of monism, that there is only one *kind* of thing to which the many particular things belong. For instance, materialism, in the sense that everything existent is material, is a form of monism because it insists that all existent things are of a single *kind,* the material kind. Thus monism and pluralism, though opposed, do not always exclude each other.

A doctrine that might be regarded as a form of pluralism, possibly the most important form of it, is dualism, the belief that there are two things or two types of thing. In view of its importance, it will be treated below in a separate section.

MONISM

"Monism" is a name for a group of views in metaphysics that stress the oneness or unity of reality in some sense. It has been characteristic of monism, from the earliest times, to insist on the unity of things in time (their freedom from change) or in space (their indivisibility) or in quality (their undifferentiatedness). Such a view of the world is already found in a developed form in the pre-Socratic philosopher Parmenides and was nicknamed the "block universe" (by Thomas Davidson, a friend of William James), that is, the universe thought of as a single closed system of interlocking parts in which there is no genuine plurality and no room for alternative possibilities. Although this world view and similar ones are now classified as forms of monism, they may not have been seen as falling into a single category at all until the term *monism* had itself been invented. The term was coined by Christian Wolff (1679–1754), and he used it only in a narrow sense, applying it to the two opposite theories that everything is mental (idealism or mentalism) and that everything is material (materialism). The term was subsequently applied to a particular doctrine of the relation between mind and matter, namely, the theory of their absolute identity (the *Identitätsphilosophie* so often mentioned by William James). The main proponents of this doctrine were Friedrich Schelling and G. W. F. Hegel, although it actually originated with Benedict de Spinoza and is sometimes known as the double-aspect theory. It holds that mind and body are only modes of the same substance, and it is this substance to which they are both reducible, not one to the other. A more recent version of this theory is the "neutral monism" of William James, which Bertrand Russell at one time also adopted. On the other hand, it should be noted that the *Identitätsphilosophie* and neutral monism differ from the "identity theory," which is a form of materialism recently set forth by J. J. C. Smart, Herbert Feigl, and others. The identity theory holds that the mind is not some third thing, some "neutral stuff" like sensation, but is literally identical with the brain.

In the nineteenth century the word *monism* came to be given wider application and so to have a systematic ambiguity, that is, a consistent variation of meaning according to context. Since then any theory that tries to reduce all phenomena to a single principle, or to explain them by one principle, or to make statements about reality as a whole, has been labeled "monism." The ambiguity is not harmful, provided that theories about how many substances there are (substantival monism) are distinguished from theories about what *kinds* of substance exist

(attributive monism). This distinction also needs to be observed in the case of pluralism (see below).

Substantival and attributive monism are logically independent views, and the various possible combinations of attitude to these questions are actually found in the doctrines of major philosophers. Thus if by "substantival monism" we mean the theory that the apparent multiplicity of substances is really a manifestation of only a single substance in different states or from different points of view, then Spinoza, with his God-or-Nature, and Francis Herbert Bradley, with his Absolute, are typical substantival monists. Indeed, Part I of Spinoza's *Ethics* is the classic exposition of substantival monism, offering a proof that there can be only one self-subsistent and independent thing. But Spinoza rejected attributive monism, which maintains that all the substances that there are, whether one or many, are ultimately of a single kind. He believed in an infinity of real attributes. An opposite case is that of Gottfried Wilhelm Leibniz, who rejected substantival monism but accepted a monism of attributes, for in his philosophy all the monads are of one kind, being souls.

A further possible doctrine, that might be called partial monism, is the belief that even if there is more than one realm of being, there is only one substance within some particular realm. For example, René Descartes, who is the classic dualist insofar as he divides the world into the two realms of mind and matter, accepted partial monism about matter, which he treated as a unitary substance, while he rejected partial monism about minds.

If monism in one or other of these various senses keeps on turning up in quite diverse philosophical systems, that is not really surprising. A striving for unity in a world description, perhaps for the sake of easier comprehensibility and greater economy of explanation, perhaps resulting from the direct appeal of simplicity, is a perennial urge in human thought. Even a substantival pluralist, Leibniz for instance, usually maintains that the plurality of substances in his world do form a systematic unity "ideally" or when looked at from the viewpoint of an omniscient being. To many minds, a monistic theory is always the most attractive option if the obstacles to holding it can be removed.

DUALISM

Dualism is the position of those thinkers who find some radical and irreducible difference in the world, an insuperable gulf between two realms of being. Any philosophical system that divides the world into two categories or types of thing, or uses two ultimate principles of explanation, or insists that there are two substances or kinds of substance, is a form of dualism. (The same ambiguity is found here as with the other labels.) Even the presence of a cardinal though not all-embracing contrast in a philosophical system may justify calling it a dualism in a looser sense, as when we speak of the dualism of Plato, in whose works the world of flux presented to the senses is sharply contrasted with the world of Forms known by the intellect, or when we consider the corresponding dualism of phenomena and noumena in Immanuel Kant.

Although superficially dualism can be seen as a special case of pluralism, it should be clear from the foregoing that it has often been, so to speak, the expression of failed monism. Nor is it merely that monism has to many minds the attractiveness described earlier; the dualistic position is inherently unstable and puzzle-generating. Once we have divided the world into two—for example, into natural and supernatural, temporal and eternal, material and mental, particular and universal—we have on our hands the problem of the relation between the two resulting worlds. These bridging problems have bulked large in both ancient and modern philosophy. Even though dualism of mind and body, for instance, may be said to reflect the time-honored view of common sense and was adopted by philosophers at least as early as Anaxagoras, Descartes's version of it, with thinking substances operating mysteriously on bits of extended substance, set the problem for all subsequent philosophers until Gilbert Ryle, in *The Concept of Mind* (1949), dismissed it as a "category-mistake."

There may be thinkers for whom oppositions themselves have an attraction, just as triads certainly do for some others. If so, the series of opposites set up by the Pythagoreans may have had this motivation. Since, however, they reduced the two sets to two fundamental principles, the Limit and the Unlimited, they may have been forced by their mathematical discoveries to acknowledge a difference that blocked the way to monism. Whatever the correct interpretation in their case, it is plain that no philosopher would in advance adopt dualism as an ideal at which to *aim,* in creating his world picture.

What in fact drew attention to dualism as a type of theory was theology, where doctrines like Manichaeism, with its two ultimate principles of good and evil, or darkness and light, are found. Those who put forward such doctrines were labeled "dualists" by Thomas Hyde, writing in Latin about 1700. Later the term found its way into philosophy in various languages.

PLURALISM

If there is more than one kind of existent, why not any number instead of just two? The unsuccessful would-be monist may, through thinking in this way, lapse into pluralism. Others, like William James, may find they have a temperamental objection to monism, with its emphasis on the totality and its exclusion of individuality and quirkiness. Yet others may from the start see the world as having some kind of disconnectedness as an essential feature, without which motion, change, and free will, for example, would be impossible. The rejection of any form of monism of course entails adopting the corresponding pluralist viewpoint. There may, however, be different types of rejection. Pluralism may arise from the rejection of the metaphysical conception of the "block universe" or of the logical doctrine that all true statements are, in the last analysis, logically necessary. For if there are some truths of a merely contingent nature, the doctrine of internal relations, that all relations are grounded in the natures of the related terms, must be false, and this doctrine is fundamental to the idealist versions of monism. The case of Leibniz, who is often taken as a standard pluralist, does not illustrate this point, but an instance of this sort of conversion to pluralism is afforded by Russell, who writes of his early position, "I came to disbelieve Bradley's arguments against relations, and to distrust the logical bases of monism" (*The Philosophy of Bertrand Russell,* edited by P. A. Schilpp, Evanston, IL, 1944, pp. 11–12). Russell later adopted a full-blown pluralism associated with logic: For instance, "When I say that my logic is atomistic, I mean that I share the common-sense belief that there are many separate things" ("The Philosophy of Logical Atomism," 1918; reprinted in his *Logic and Knowledge,* New York, 1956, p. 178). Though this phase of Russell's philosophy is usually known as logical atomism, he also described it himself as "absolute pluralism." Even after abandoning logical atomism, Russell remained an enthusiastic pluralist; in 1931 he wrote of the proposition that the world is a unity, "the most fundamental of my intellectual beliefs is that this is rubbish. I think the universe is all spots and jumps, without unity, without continuity, without coherence or orderliness or any of the other properties that governesses love" (*The Scientific Outlook,* New York, 1931, p. 98).

See also Bradley, Francis Herbert; Categories; Descartes, René; Dualism in the Philosophy of Mind; Hegel, Georg Wilhelm Friedrich; James, William; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Mani and Manichaeism; Mind-Body Problem; Parmenides of Elea; Plato; Pluralism; Russell, Bertrand Arthur William; Ryle, Gilbert; Schelling, Friedrich Wilhelm Joseph von; Smart, John Jamieson Carswell; Spinoza, Benedict (Baruch) de; Wolff, Christian.

Bibliography

Helpful general discussions of monism, dualism, and pluralism are rather few in number. The only good general account of all three is A. M. Quinton, "Pluralism and Monism," in the *Encyclopaedia Britannica.* The best sources, though more difficult to use, are the actual works of the philosophers mentioned as proponents of the various doctrines.

MONISM

On monism see the works of philosophers named in the text, such as Parmenides, Spinoza, and Bradley. A useful discussion is C. E. M. Joad, "Monism in the Light of Recent Developments in Philosophy," in *PAS* 17 (1916–1917): 95–116. Now somewhat antiquated is A. Worsley, *Concepts of Monism* (London, 1907). A typical short account from the heyday of monism in British philosophy is A. E. Taylor, *Elements of Metaphysics* (London: Methuen, 1903), Chs. 2–3. Compare J. A. Smith, "The Issue between Monism and Pluralism," in *PAS* 26 (1925–1926): 1–24. See also Marvin Farber, "Types of Unity and the Problem of Monism," in *Philosophy and Phenomenological Research* 4 (1943–1944): 37–58, and postscript, ibid., 6 (1945–1946): 547–583; Raphael Demos, "Types of Unity According to Plato and Aristotle," ibid., 534–545; Abraham Edel, "Monism and Pluralism," in *Journal of Philosophy* 31 (21) (October 1934): 561–571; and Jonathan Bennett, "A Note on Descartes and Spinoza," in *Philosophical Review* 74 (3) (July 1965): 379–380. Such nineteenth-century works as Ernst Haeckel, *Der Monismus als Band zwischen Religion und Wissenschaft* (Bonn: E. Strauss, 1893; translated by J. Gilchrist as *Monism as Connecting Religion and Science,* London: A. and C. Black, 1895), are not now of much philosophical interest, for they are not about monism in general but are presentations of an outdated type of materialism.

DUALISM

On dualism see the main works of Descartes. The difficulties of the dualist position in general are well brought out by John Passmore in his *Philosophical Reasoning* (London: Duckworth, 1961), Ch. 3. See also Simone Pétrement, *Le dualisme chez Platon, les gnostiques, et les manichéens* (Paris: Presses Universitaires de France, 1947).

PLURALISM

The most readable book on pluralism and other theories is William James's *A Pluralistic Universe* (London: Longman, 1909). For further reading, there is James Ward, *The Realm of Ends, or Pluralism and Theism* (Cambridge, U.K.: Cambridge University Press, 1911). A dry but clear account is to be found in C. D. Broad, *The Mind and Its Place in Nature* (London: Kegan Paul, 1925), introduction. More difficult and technical but classic is G. E. Moore, "External and Internal Relations," in *PAS* 20 (1919–1920): 40–62, reprinted in his *Philosophical Studies* (New York: Harcourt Brace, 1922). Compare Bertrand Russell, "The Nature of Truth," in *Mind* 15 (1906): 528–533, reprinted as "The Monistic Theory of Truth," in Russell's *Philosophical Essays*

(London: Allen and Unwin, 1910). See also J. H. Muirhead, F. C. S. Schiller, and A. E. Taylor, "Why Pluralism?," in *PAS* 9 (1908–1909): 183–225; and P. Laner, *Pluralismus oder Monismus* (1905).

Roland Hall (1967)

MONTAGUE, RICHARD

(1930–1971)

Richard M. Montague, a logician who taught in the Philosophy Department at the University of California at Los Angeles from 1955 until his premature death in 1971, is probably best known for his contributions to linguistic semantics, although he also made important contributions to mathematical logic and philosophy.

Montague was born in 1930. He attended the University of California at Berkeley both as an undergraduate and a graduate student, concentrating not only in mathematics and philosophy, but in Semitic languages. Working with Alfred Tarski, he completed a doctoral dissertation in 1957 entitled "Contributions to the foundations of axiomatic set theory." By that time he had published a large number of papers in various areas of mathematical logic.

Montague's interests in mathematical logic were general and included set theory, proof theory, model theory, and abstract recursion theory. One early theme in his work in mathematical logic concerned the consequences of semantic reflection for axiomatic versions of set theory and other mathematical theories. That work has been widely cited and is still important.

The work for which Montague is best known was carried out late in his life (beginning with the 1968 publication of "Pragmatics") and dealt with the development of logics intended to serve as vehicles for the interpretation of natural language and the formalization of philosophy. From Tarski, Montague inherited the view that semantical theories could and should be formulated with mathematical precision. However, his project of applying Tarski's techniques to natural language seems to derive more naturally from the work of Rudolf Carnap and Alonzo Church.

Both Carnap and Church worked with a framework for logical formalization, which, although it was developed in connection with the language of mathematical theories, was clearly more broadly applicable. Carnap was mainly interested in using formalization as a tool for clarifying philosophy. Church considered what he called the "logistic method"—that is, the method of logical formalization developed in the first half of the twentieth century—to be applicable in a more general linguistic setting.

Carnap and Church both addressed a major obstacle standing in the way of generalizing Tarskian semantic theories to natural language—the problem of intensionality (which had already been raised by Gottlob Frege). Carnap explored how what are now called possible worlds could be used to model intensionality, while Church sought to formalize Frege's theory of sense and denotation. Influences of both can be seen in Montague's logical framework for interpreting natural language. Like Carnap, Montague appealed to possible worlds, and, like Church, he used higher-order logic. Montague's insight that a logic combining possible worlds with higher-order logic provided a flexible and powerful tool for natural language semantics proved to be fundamentally important.

All of Montague's publications concerning "Montague Grammar" are collected together in *Formal Philosophy: Selected Papers of Richard Montague* (1974). These papers develop the logical framework of "intensional logic." This is a higher-order logic involving on a system of types based on three primitive domains: entities (type e), possible worlds, and the two truth values T and F (type t). If σ and τ are types, then $<\sigma,\tau>$ is also a type and corresponds to the set of functions from the domain of σ to the domain of τ. Thus, for instance, $<e,t>$ is the type of functions from entities to truth values. If σ is a type, then $<s,\sigma>$ is also a type and corresponds to the set of functions from possible worlds to the domain of σ: $<s, <e,t>>$, then, is the type of intensions of sets of entities. For a book-length, systematic treatment of intensional logic, see Daniel Gallin's *Intensional and Higher-Order Logic* (1975).

A Montague grammar for a fragment of a language consists of a syntactic account of that fragment, which defines a set of syntactic structures showing how complex phrases are decomposed into components, and a semantic component that shows how a semantic value can be assigned to the structure given an assignment of values to the lexical items occurring in the structure. These values belong to the domains of a model of intensional logic. Intensional logic can serve as an intermediary in the mapping of syntactic structures to values, and as a vehicle for formulating postulates about the meanings of lexical items. This mapping conforms to a correspondence between grammatical categories like "Sentence" and "Noun-Phrase" and the types of intensional logic.

To see how the idea might work, consider the sentence "John wants a car." The noun phrase 'a car' has type <<e,t>,t>; it denotes the set of sets containing at least one car. (The insight that noun phrases denote sets of sets goes back to Frege's 1884 work, *The Foundations of Arithmetic*.) The verb 'wants' corresponds to a function that inputs the intension of a Noun-Phrase denotation and returns a function from entities to truth values. Give this function the intension of 'a car' and it returns a function saying of each entity whether that entity wants a car. The type of 'wants' is therefore <<s, <<e,t>,t>> , <e,t>>. Barbara Partee and Herman L. W. Hendriks provide a useful extended survey of Montague's semantic framework and its subsequent influences in their 1996 essay "Montague Grammar."

Montague himself saw intensional logic and his theory of language as a basis of formalizing philosophy, but the most important direct influence of his work was on the development of linguistic semantics, where its impact was enormous. Montague's semantic techniques can be associated with any generative syntactic framework; his syntactic approach has been less influential, outside of subsequent work in the categorical grammar framework. (See Jacobson 1996, for example.)

Although few philosophers would agree that the goal of formalizing philosophy is enabled by Montague's work, foundational questions raised by his approach have preoccupied and shaped subsequent work in analytic metaphysics and philosophy of language. Much of this influence is indirect, occurring through the work of David Lewis,who attended Montague's courses at UCLA and was influenced by his ideas.

Because of Montague's uncompromising emphasis on the technical dimension, his papers are difficult reading. But even now, they repay careful study. The linguistic papers and other philosophically relevant work were compiled in 1974 in *Formal Philosophy: Selected Papers of Richard Montague*. Further biographical information concerning Montague can be found in Anita and Solomon Feferman's biography of Tarski, *Alfred Tarski: Life and Logic* (2004).

See also Artificial and Natural Languages; Carnap, Rudolf; Church, Alonzo; Computability Theory; Frege, Gottlob; Lewis, David; Logic, History of: Modern Logic; Mathematics, Foundations of; Modal Logic; Model Theory; Proof Theory; Semantics, History of; Set Theory; Tarski, Alfred; Type Theory.

Bibliography

Carnap, Rudolf. *Meaning and Necessity*. 2nd ed. Chicago: University of Chicago Press, 1956. The first edition was published in 1947.

Church, Alonzo. "The Need for Abstract Entities in Semantic Analysis." *Proceedings of the American Academy of Arts and Sciences* 80 (1951): 100–112.

Church, Alonzo. *Introduction to Mathematical Logic*. Vol. 1. Princeton, NJ: University of Princeton Press, 1959.

Church, Alonzo. "Outline of a Revised Formulation of the Logic of Sense and Denotation (Part I)." *Nous*, 7 (1973): 24–33.

Church, Alonzo. "Outline of a Revised Formulation of the Logic of Sense and Denotation (Part II)." *Nous*, 8 (1974): 135–156.

Feferman, Anita Burdman, and Solomon Feferman. *Alfred Tarski: Life and Logic*. Cambridge, U.K.: Cambridge University Press, 2004.

Frege, Gottlob. *The Foundations of Arithmetic*. 2nd ed. Translated by J. L. Austin. Oxford, U.K.: Oxford University Press, 1953. Originally published in 1884.

Furth, Montgomery, C. C. Chang, and Alonzo Church. "Obituary of Richard M. Montague." Unpublished manuscript. Los Angeles: University of California at Los Angeles Department of Philosophy, 1971.

Gallin, Daniel. *Intensional and Higher-Order Logic*. Amsterdam: North-Holland, 1975.

Jacobson, Paulene. "The Syntax/Semantics Interface in Categorical Grammar." In *The Handbook of Contemporary Semantic Theory*, edited by Shalom Lappin, 89–116. Oxford, U.K.: Blackwell Publishers, 1996.

Montague, Richard. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven, CT: Yale University Press, 1974. Edited and with an introduction by Richmond H. Thomason.

Partee, Barbara H., with Herman L. W. Hendriks. "Montague Grammar." In *Handbook of Logic and Language*, edited by Johan van Benthem and Alice ter Meulen, 5–91. Amsterdam, Netherlands: Elsevier Science Publishers, 1996.

Richmond H. Thomason (2005)

MONTAGUE, WILLIAM PEPPERELL

(1873–1953)

William Pepperell Montague, an American realist philosopher, received his BA from Harvard in 1896, his MA the following year, and his PhD in 1898. He taught briefly at Radcliffe, Harvard, and the University of California. In 1903 he began teaching at Barnard and from 1907 to 1910 was an adjunct professor and a member of the Columbia University graduate faculty of philosophy. He became associate professor in 1910, professor in 1920, and was the Johnsonian professor of philosophy from 1920 to 1941. In 1928 he was Carnegie visiting professor

in Japan, Czechoslovakia, and Italy. He served as chairman of several delegations to the International Congress of Philosophy (1920, 1934, 1937) and as president of the eastern division of the American Philosophical Association in 1923.

REALISM

Montague advocated a frankly Platonic "subsistential realism." He called it a right-wing realism, in contrast with left-wing realism, whose adherents included the behaviorists, objective relativists, and—to some extent—pragmatists. At the turn of the twentieth century, the idealist claim that the object of knowledge was dependent on the knower and thus was "ideal" had come increasingly under attack in England and America. Montague, in "Professor Royce's Refutation of Realism" (1902), was one of the first to attack idealism by means of the realist theory of independence. This theory—that the object of knowledge is not dependent for its reality on the knowing relation—became one of the cardinal tenets of the New Realist movement, of which Montague was a charter member. However, by itself it was not enough to establish that the known is independent of the knower. It also had to be shown how a conscious, knowing organism could be in such a unique kind of rapport with events whose loci and dates were different from its own. Thus the central issue in epistemology for Montague was to establish the independence and the immanence of the object of knowledge.

Montague proposed his "subsistential realism" as a resolution of this issue. Subsistence included everything that could be made an object of discourse. The objects of knowledge then are subsistently real, that is, propositions and terms rather than commonsense objects, and as such they are directly present to mind (immanent), though independent of it. Montague thus brought the things of the earth into the realm of ideas by interpreting existence as a subclass of subsistence, hence also as a set of propositions.

With his idea of subsistent and existential propositions, Montague could distinguish nonveridical and unreal objects from the veridical and real. Existential propositions are the objects of true or real knowledge, and the "merely subsistent" propositions are the objects of false or unreal knowledge. Thus there is a tendency in Montague's thinking to identify the true, real, and existent on the one hand, and the false, unreal, and nonexistent on the other.

What, then, is the cause of error? Truth and falsity attach to our judgments, Montague said, because of their content, not because they are stated or believed. Error is the result of the selective action of sense perception and conception. He attributed error to these factors of the "personal equation" (as realists called the subjective aspect of knowledge) because he had said existential subsistent propositions cause themselves to be known in a way the "merely subsistent" cannot. But how can a proposition cause itself to be known?

The answer apparently was in the difference between the "merely subsistent" propositions and the existential subsistent propositions. Montague identified existential propositions with facts, and he described a fact as "something done," a fait accompli. But this was as far as he went.

ANIMISTIC MATERIALISM

Epistemology was secondary, however, to Montague's preoccupation with the psychophysical problem of the nature of mind and its relation to the body. Naturalistic monism, strongly supported by science, could not, Montague claimed, adequately account for such characteristics of mind as purpose, privacy, duration, and integration. Traditional dualism could account for them, but it was scientifically sterile in its reliance on concepts of spirit. Montague's answer, which he called "animistic materialism," was the hypothesis of a physical soul possessing all of the traits of mind although still physically describable.

Throughout his career, Montague considered the soul to be the only answer to the psychophysical problem. After proposing the idea of a substantial soul in his first published writing, Montague soon rejected it in favor of considering the soul as a new kind of energy, purely private, and internally observable as sensation. This "potential" energy comes into existence when and where the kinetic energy of a stimulus ceases to be externally observable as motion. Sensations (or consciousness) and their externally observable causes are thus qualitatively identical. The potentiality of the physical is the actuality of the psychical, and vice versa. Just as when successive twists are imposed upon a coiled spring there is left unobservable potential energy, so too the potential energies of sensations leave traces superposed on one another. These traces constitute the memory system and modify the organism's responses to later stimuli.

Thus, within the organism there arises a field of potential energy that is externally unobservable yet is causally effective upon the visible cerebral matrix; this inner organism possesses all the characteristics of mind. In Montague's relational dualism, therefore, mind and body are in radical contrast as relations but not as sub-

stances. The truths of psychophysical dualism were thus saved without departing from material categories. Montague in general maintained this materialistic dualism, yet at one point (in "A Realistic Theory of Truth and Error," 1912) he admitted to what he called a qualified panpsychism: Matter had something psychical about it.

RELIGIOUS VIEWS

Montague's "Promethean challenge to religion" (as he called it in *Belief Unbound*) was a challenge to authoritarianism, supernaturalism, and asceticism in religion. Montague denied what he termed the "pseudo creativeness" that idealism and pragmatism attribute to humans. Man has no transcendent power to legislate for nature, or to support infinite space and time by his consciousness. Realism instead gives to man an even greater responsibility of membership in the independent order of nature. Realism also adds to existent things the "quiet and infinitely great immensities of the realm of subsistence" where mind gains access to new and imperishable sources of joy and peace. Philosophy's one certainty is that ideals are eternal things, and the life that incarnates them attains an absolute value that time alone could not create and that death is powerless to destroy.

Ideals are not dependent on God's will. God is neither finite nor infinite in all things. He is infinite and eternal like the universe that is his body, all-perfect in himself and in his will but limited in power by that totality of actual and possible things which is within him yet not himself. God is to be loved because he is good, not because he is powerful.

Montague had a genuinely speculative and daring mind that explored not only the fields of philosophy but also such areas as time perception, mathematics, relativity theory, and quantum mechanics. At the beginning of Montague's career, philosophy suffered from what he called "internalism," a subjectivism sometimes carried to the point of solipsism, which, if it perhaps contained a grain of truth, was sterile. By the end of his life Montague feared that philosophy had gone to the other extreme. In "The Modern Distemper of Philosophy" (1951), he expressed his concern that it now suffered from an "externalism," a "distemper" that was eliminating important philosophical problems from discussion because they were insufficiently empirical.

See also Epistemology, History of; Idealism; New Realism; Platonism and the Platonic Tradition; Propositions; Realism.

Bibliography

Works by Montague include the following: "A Plea for Soul-Substance," in *Psychological Review* 6 (5) (September 1899): 457–476; "Professor Royce's Refutation of Realism," in *Philosophical Review* 2 (January 1902): 43–55; "A Realistic Theory of Truth and Error," in E. B. Holt and others, *The New Realism* (New York: Macmillan, 1912), pp. 251–300; *The Ways of Knowing; or The Methods of Philosophy* (New York: Macmillan, 1925), a good example of Montague's desire to save the truths in all philosophies; *Belief Unbound; a Promethean Religion for the Modern World* (New Haven, CT: Yale University Press, 1930); "Confessions of an Animistic Materialist," in *Contemporary American Philosophy*, edited by W. P. Montague and G. P. Adams, Vol. II (New York: Macmillan, 1930), pp. 135–158; *The Ways of Things: A Philosophy of Knowledge, Nature and Value* (New York: Prentice-Hall, 1940), the best single source for an overall view of Montague's philosophy; "The Human Soul and the Cosmic Mind," in *Mind* 54 (213) (January 1945): 50–64; *Great Visions of Philosophy: Varieties of Speculative Thought in the West from the Greeks to Bergson* (La Salle, IL: Open Court, 1950), Montague's Carus Lectures; and "The Modern Distemper of Philosophy," in *Journal of Philosophy* 48 (14) (1951): 429–435.

See also Helen Huss Parkhurst et al., "The Philosophic Creed of William Pepperell Montague," in *Journal of Philosophy* 52 (21) (1954): 593–637, which consists of articles on Montague and tributes to him by former colleagues and students.

Thomas Robischon (1967)

MONTAIGNE, MICHEL EYQUEM DE
(1533–1592)

Michel Eyquem De Montaigne, French essayist and skeptical philosopher, was born near Bordeaux. His father was an important merchant, and his mother belonged to a wealthy Spanish-Portuguese Jewish family that had fled to Toulouse. Montaigne was raised a Catholic and was given special training by his father, who would not allow him to hear any language other than Latin until he was six. At this time he was sent to the Collège de Guyenne at Bordeaux, where he studied with some of the leading humanistic teachers of the time, among them the learned Latin poet George Buchanan (1505–1582), who would later be arrested and charged by the Portuguese Inquisition for "judaizing" and skepticism. Montaigne also apparently studied at the University of Toulouse, a leading center of humanism and unorthodox religious ideas. For thirteen years he was a member of the *parlement* of Bordeaux and made several trips to Paris and the court seeking a more important position. His closest friend at this time was the stoic humanist and poet Étienne de La

Boétie (1530–1563). Montaigne's first significant writing was a letter describing La Boétie's death, published at the end of the latter's *Oeuvres* in 1570.

In 1568 Montaigne published his French translation of *Theologia Naturalis sive Liber Creaturarum* ("Natural Theology or the Book of Creatures") by Raimond Sebond (Raymond of Sabunde, d. 1436), a fifteenth-century Spanish theologian who had taught at Toulouse. In Montaigne's translation he somewhat modified Sebond's rationalistic claims that unaided human reason could comprehend the universe and establish the existence and nature of God. Montaigne also published La Boétie's works before retiring from public life in 1571. The following year he began writing his most important work, the *Essays*, a series of rambling, erudite, witty discussions on a variety of topics, serving as a self-portrait. The longest of the essays, the "Apology for Raimond Sebond," was written about 1576 while Montaigne was studying the recently rediscovered treasury of Greek skepticism—the works of Sextus Empiricus—and undergoing a personal skeptical crisis. He had mottoes from Sextus carved into the rafter beams of his study and adopted as his own motto, "Que sais-je?" ("What do I know?"). In 1580 the first two books of the *Essays* were published. Besides writing, Montaigne tried in vain during the 1570s to mediate between the Catholics and the Protestant leader, Henri of Navarre (later Henri IV).

In 1580 Montaigne went to Paris to present a copy of his *Essays* to the king; he then set out on a trip to Germany, Switzerland, and Italy, which he describes in his *Travel Journal*. The following year he was called back from Italy to become mayor of Bordeaux, a post he held for four years. He then added material to his earlier *Essays* and wrote a third volume of them; the complete edition was first published in 1588 in Paris. Montaigne went to Paris and probably negotiated on behalf of Henri of Navarre concerning his succession to the throne, his conversion to Catholicism, and the temporary settlement of the religious wars, which was later incorporated into the Edict of Nantes. Illness apparently prevented Montaigne from joining Henri IV's court, but he continued to revise his *Essays*. The final version was published posthumously in 1595.

"APOLOGY FOR RAIMOND SEBOND"

Montaigne's most important philosophical work, the "Apology for Raimond Sebond," had an enormous influence on the subsequent history of thought. A superbly written presentation of skepticism, it formulated a challenge that affected Descartes, Pierre Gassendi, Bacon, and many others and inspired monumental efforts to meet the challenge. The "Apology" gradually reveals a series of waves of doubt, continuously coupled with a new type of Christian fideism.

The essay begins with an account—probably not very accurate—of Montaigne's reasons for translating Sebond's *Theologia Naturalis*. Pierre Bunel, a Renaissance scholar, gave Montaigne's father a copy of the book, saying that it had saved him from Lutheranism. Long afterward, Montaigne's father asked his son to render it into French (from what Montaigne claimed was Spanish with Latin endings). After the translation appeared, Montaigne reported that some readers—mainly female—needed help in comprehending Sebond's contention that all the articles of the Christian faith could be established by reason. Two major objections to this thesis had been raised: the first held that Christianity should rest on faith rather than reason, and the second maintained that Sebond's reasons were not good ones. Montaigne purported to defend Sebond by showing that because all reasoning is unsound, Sebond's is no worse than anyone else's and, therefore, religion should rest on faith alone.

Montaigne held that people are vain, stupid, and immoral, and he pointed out that they and their achievements do not appear impressive when compared with animals and their abilities. The "noble savage" of the New World seemed to possess an admirable simplicity and ignorance that did not involve him in the intellectual, legal, political, and religious problems of the civilized European.

Montaigne suggested that our sole contact with the truth was due not to our intellect or reason, but rather to the grace of God; he agreed with St. Paul that ignorance is more useful than learning in acquiring truth. To show this, Montaigne examined the teachings of the ancient schools of philosophy and argued that those of the Pyrrhonists were the best and the most compatible with the Christian religion. All of the other philosophies were in conflict with one another, contained contradictions and absurdities, and relied on fallible human faculties and questionable premises to reach their conclusions. Only Pyrrhonists showed humans as naked and empty, portrayed their natural weaknesses, and by ridding them of their false or dubious opinions, left their minds a blank tablet, ready to receive whatever God might wish to write upon them. The modern Pyrrhonist would not be led into heresy, because he or she would accept no reasons or arguments that are open to question. In contrast to the Pyrrhonists, who suspended judgment on all matters, other philosophers offered their own opinions as genuine

truths. They thought that they had discovered the real nature of things and had measured the universe in terms of their own systems; they were only deceiving themselves.

In the later portions of the "Apology," Montaigne presented the Pyrrhonistic evidence that everything is dubious and that genuine knowledge must be gained either by experience or by reasoning. We do not, however, know the essence of what we experience (for example, the real nature of heat), and we do not even know the nature of our own faculties. We are constantly changing as our physical and emotional conditions alter, and the judgments we make and accept at one time, we find doubtful at another. Not only does this seem to happen to each of us, but it also appears to be the fate of humans in general. Each alleged scientific discovery is superseded by another, and what is thought true at one time is regarded as false or silly at another.

The new sciences of Copernicus and Paracelsus claimed that the ancient sciences of Aristotle, Ptolemy, and others were false. How could we know, Montaigne asked, that some future scientist would not make similar claims, on equally firm grounds, about these new discoveries? These same variations and disagreements occur in every area of human concern.

Montaigne then presented the more theoretical objections that Sextus Empiricus had raised about the possibility of gaining knowledge. All of our alleged knowledge, he argued, appears to come from sense experience, but perhaps we do not possess the requisite number of senses for gaining knowledge. Even if we do possess all of them, the information we gain through them is deceptive and uncertain. Illusions lead us to wonder when our senses are accurate. Dreams are often so similar to sense experiences that we cannot tell if sense experience itself is not really a dream. Each of our experiences differs from that of animals, from that of other human beings, and even from our other experiences; we cannot, therefore, know when to accept an experience as accurate. Such conditions as illness or drunkenness distort what we perceive. Perhaps normal experience itself is a kind of distortion.

In order to determine the accuracy of our experiences, we require a criterion. But we need some way of testing that criterion, and this requires a second criterion to establish how to test it, and so on. If reason is to be the judge of our experiences, then we need reasons to justify our reason, and so on, to infinity. Thus, if our ideas come from our sense experiences, we are hardly in a position to use our ideas to judge the nature of objects. Our experiences and our ideas tell us only how things seem to be, but not necessarily how they are in themselves. Trying to know reality, Montaigne concluded, is like trying to clutch water. We can deal with the world only in terms of appearances, unless and until God decides to enlighten us. In our present state, we can only try to follow nature, living as best we can.

INTENTIONS AND INFLUENCE

Montaigne questioned and cast doubt upon almost all of humankind's beliefs in philosophy, theology, science, religion, and morality, and criticized almost every superstition and accepted view. He insisted that he was merely showing the human inability to find truth by means of natural capacities and the human need to rely on faith as the sole access to truth. Montaigne's own portrayal of the human predicament succeeded in intensifying the doubts already produced by the religious crisis of the Reformation, the humanistic crisis of the Renaissance, and the philosophical-scientific crisis of revived Pyrrhonism. The three currents were fused into a massive and forceful onslaught in this "Apology." Montaigne's formulation of skepticism and the more didactic one of his disciple, Pierre Charron, provided the issues for seventeenth-century thought. Some, such as François de La Mothe Le Vayer, were to follow out the more destructive and anti-intellectual tendencies of Montaigne's doubt. Others, such as Marin Mersenne and Gassendi, were to formulate a mitigated skepticism that could accept its doubts while seeking information about the world of appearances. Still others, such as Bacon, Herbert of Cherbury, and Descartes, were to seek new philosophical systems to provide for human knowledge a basis impervious to Montaigne's doubts.

Some have seen Montaigne as a skeptic, questioning religion with everything else, and as the founder of the critical spirit of the Enlightenment. They have taken his fideism as a mask for his actual views and have portrayed him as a genuine freethinker and free spirit. Others have interpreted his fideism as an expression of his own resolution of his doubts. Although Montaigne lacked the religious fervor of Pascal, who regarded him as a skeptical nonbeliever, many of his contemporaries and later admirers took his skepticism as part of the Counter-Reformation, because it opposed the reasons and arguments of the Reformers by undermining the validity of all reasoning.

Montaigne played a vital role in the development of both Christian skeptical fideism and of the so-called *libertinage*, a later movement of critical freethinking that

preceded the Age of Reason. His views are compatible with both roles, in that his doubts neither imply nor contradict either a religious or an irreligious conclusion. He was probably mildly religious, accepting Catholicism in the light of the religious wars of his time. He apparently opposed fanaticism and wished for toleration of all sides, recognizing man as a fallible, limited creature struggling to live and comprehend with weak and uncertain capacities. Without God's assistance, man could only try to understand himself, guided by the past and the present. To understand himself and his situation would at least make him doubtful of radical proposals for solving everything, make him more tolerant, and—most important—make him capable of accepting himself and his fate. To philosophize, Montaigne said, was to learn to die.

See also Epistemology; Philosophy of Religion.

Bibliography

WORKS BY MONTAIGNE

The Essayes. Translated by John Florio. London, 1603. New York: Modern Library, 1933. The first English edition, probably known to Shakespeare and Francis Bacon.

The Essays. Translated by Jacob Zeitlin. 3 vols. New York: Knopf, 1934–1936. Good introduction and notes.

The Complete Essays of Montaigne. Translated by Donald M. Frame. Stanford, CA: Stanford University Press, 1958. The best modern translation, containing the essays, journals, and letters, plus an excellent introduction and annotated bibliography.

Œuvres complète. Paris: Gallimard, 1976.

Apologie de Raimond Sebond de la Theologia à la Théologie. Paris: Librairie Honoré Champion, 1990.

WORKS ABOUT MONTAIGNE

Allen, Don Cameron. *Doubt's Boundless Sea: Skepticism and Faith in the Renaissance*. Baltimore, MD: Johns Hopkins Press, 1964. Reasserts the irreligious interpretation of Montaigne. See chapter 3.

Boase, Alan M. *The Fortunes of Montaigne: A History of the Essays in France, 1580–1669*. London: Methuen, 1935. A study of Montaigne's impact.

Brahami, Frédéric. *Le scepticisme de Montaigne*. Paris: Presses universitaires de France, 1997.

Brunschwicg, Léon. *Descartes et Pascal, lecteurs de Montaigne*. New York: Brentano's, 1944. Shows the influence of Montaigne on both Descartes and Pascal.

Brush, Craig B. *Montaigne and Bayle: Variations on the Theme of Skepticism*. The Hague: Nijhoff, 1966. See chapter 3.

Busson, Henri. *Le Rationalisme dans la literature française de la renaissance (1533–1601)*. Paris: J. Vrin, 1957. Interprets Montaigne as a freethinker and as part of an irreligious, rationalistic milieu.

Frame, Donald M. *Montaigne: A Biography*. New York: Harcourt, Brace, 1965.

Frame, Donald M. *Montaigne's Discovery of Man: The Humanization of a Humanist*. New York: Columbia University Press, 1955. A study of the development of Montaigne's thought.

Frame, Donald M. "What Next in Montaigne Studies?" *French Review* 36 (1963): 577–587. A survey of the state of scholarship on Montaigne and an evaluation of various interpretations.

Giocanti, Sylvia. *Penser l'irrésolution: Montaigne, Pascal, La Mothe Le Vayer, Trois itinéraires sceptiques*. Paris: Honoré Champion éditeur, 2001.

Laursen, John Christian. *The Politics of Skepticism in the Ancients, Montaigne, Hume, and Kant*. Leiden, Netherlands: Brill, 1992. See chapters 4 and 5.

Limbrick, Elaine. "Was Montaigne Really a Pyrrhonian?" *Bibliothèque d'Humanisme et Renaissance* 39 (1977): 67–80.

Malvezin, Théophile *Michel de Montaigne, son origine, sa famille*. Bordeaux: C. Lafebvre, 1875. Contains much data about Montaigne's background and environment.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. Oxford, U.K.: Oxford University Press, 2003.

Popkin, Richard H. "Skepticism and the Counter-Reformation in France." *Archiv für Reformationsgeschichte* 51 (1960): 58–87. The role of Montaigne's skepticism in French Catholic theology of the time.

Strowski, Fortunat. *Montaigne*. 2nd ed. Paris: F. Alcan, 1931. The best-known scholarly modern French interpretation of Montaigne.

Villey, Pierre. *Les Sources et l'évolution des essays de Montaigne*. Paris: Hachette, 1908. Basic study of Montaigne's sources and the development of the *Essays*.

Villey, Pierre. *Montaigne devant la posterité*. Paris: Boivin, 1935. A study of how Montaigne has been interpreted.

Richard Popkin (1967, 2005)

MONTESQUIEU, BARON DE

(1689–1755)

The philosopher and political theorist Charles-Louis de Secondat, Baron de Montesquieu, afterward Baron de la Brède et de Montesquieu, was born at Labrède, near Bordeaux, in the year of the English revolutionary settlement that established the preeminence of Parliament. He was a follower of John Locke and the outstanding champion in France of the supposedly "English" notions of freedom, toleration, moderation, and constitutional government. He was also a pioneer in the philosophy of history and in the sociological approach to problems of politics and law. Honored in his own country, Montesquieu was even more revered in the English-speaking world. He described the constitution of England as "the mirror of liberty," and although his analysis of the English principles of government was generally considered defective by later historians, it was hailed as marvelously penetrating

by English readers of his own time. Charles Yorke, the future lord chancellor, told Montesquieu, "You have understood us better than we understand ourselves." Moreover, the founders of several new political societies, notably that of the United States, were profoundly affected by Montesquieu's teaching. Especially influential was his theory that the freedom of the individual could best be guaranteed by the division of the powers of the state between three distinct organs that could balance and check one another—a separation of powers Montesquieu, rightly or wrongly, believed to be characteristic of the English system.

Montesquieu belonged to the *noblesse de robe.* Part of his design in recommending the separation of powers in France was to elevate the French aristocracy to a position comparable to that of the English, for whereas Rousseau believed that political liberty could be achieved only in a democracy and Voltaire believed it could best be achieved by a philosopher-king, Montesquieu held that liberty was most secure where there was a potent aristocracy to limit the despotic tendency of both the monarch and the common people. He believed that the way to preserve freedom was to set "power against power."

No one wrote with greater eloquence against despotism than did Montesquieu, yet he was far from sharing the conventional liberal outlook of the eighteenth-century *philosophes.* He had all the conservatism characteristic of the landowner and the lawyer. In many respects he was positively reactionary; for instance, he wished to strengthen rather than diminish hereditary privileges. But like Edmund Burke, whom he influenced considerably, Montesquieu was able to reconcile his reforming and reactionary sentiments by insisting that he sought to restore old freedoms, not promote new ones. He argued that the centralizing monarchistic policy of Louis XIV had robbed Frenchmen of their ancient liberties and privileges. The only kind of revolution Montesquieu advocated was one that would give back to the French Estates—and to the nobility and the *parlements* in particular—the rights they had enjoyed before the seventeenth century. The actual French Revolution, which sought to enfranchise the bourgeoisie and the common people and to bring about a variety of other innovations, was far from the sort of change that Montesquieu had favored, although he inadvertently did help to inspire the events of 1789 and after.

Montesquieu's parents were not well off. He inherited his title and much of his wealth from an uncle who at the same time bequeathed him the office of *président à mortier* of the *parlement* at Bordeaux. About the same time his worldly position was further secured by a prudent marriage to a Protestant named Jeanne de Lartigue, who, although exceedingly plain in appearance, was heiress to a considerable fortune. Even so, Montesquieu remained an ambitious man, and, after twelve years as *président* in Bordeaux, he forsook his chateau and vineyards, to which he was deeply attached, and his wife, whom he loved perhaps rather less, to seek fame in Paris and to travel to other countries collecting material for his books. He was a success in the Paris salons, and although there seem to be no recorded examples of his wit in talking, he was celebrated as a conversationalist. He made friends with influential people and became the lover of the Marquise de Grave, among others. She inspired one of his early anonymous works, *Le temple de Gnide,* a mildly indecent erotic fantasy that was also a satire on the court of the infant Louis XV. After some difficulties Montesquieu was admitted to the French Academy in 1728.

He was on the whole a popular, but certainly not a generous, man. As a landowner he was most rigorous in the collection of even the smallest debts; at the same time he was slow to pay money he owed to others. In Paris he had a reputation for parsimony; more than one contemporary remarked that he "never ate at his own table." At his chateau, La Brède, English guests were struck by what they politely called the "plainness" of the fare, and Montesquieu even economized on the arrangements for the wedding of his daughter Denise. He once warned his grandson, "La fortune est un état et non pas un bien."

LES LETTRES PERSANES

Montesquieu made his name as a writer at the age of thirty-two with the publication of *Les lettres persanes* (1721). Presented in the guise of a series of letters sent from France by two Persian visitors, Usbek and Rica, and translated into French by Montesquieu, this book is a satirical attack on French values and institutions. It is written with great wit and skill. The Persian visitors begin by remarking on the strange customs of the French in such matters as cutting their hair and wearing wigs and reversing the Persian rule of giving trousers to women and skirts to men. They then proceed by degrees to express delicate amazement at the things the French choose to respect or hold sacred. They comment on the mixture of grossness and extravagance in the manners of Parisian society. Their sly digs at French politics are even more telling. They describe Louis XIV as a "magician" who "makes people kill one another even when they have no quarrel." The Persians also speak of "another conjuror who is called the Pope … who makes people believe that

three are only one, and that the bread one eats is not bread or that the wine one drinks is not wine, and a thousand other things of the same sort." The Spanish Inquisitors are described as a "cheerful species of dervishes who burnt to death people who disagreed with them on points of the utmost triviality." The revocation of the Edict of Nantes is likewise mocked, Louis XIV being said to have contrived "to increase the numbers of the faithful by diminishing the numbers of his subjects."

In the same book Montesquieu sought to establish two important principles of political theory—first, that all societies rest on the solidarity of interests and, second, that a free society can exist only on the basis of the general diffusion of civic virtue, as in the republics of antiquity.

Although Montesquieu attacked the manners of polite society in France, he did not fail to give *Les lettres persanes* a fashionable appeal. The two Persian travelers offer piquant descriptions of the pleasures of the harem and the sufferings of the women they have left behind them. Satire is nicely spiced with wit and the wit with impropriety, although this book is not quite so risqué as *Le temple de Gnide.* Montesquieu was said by Rutledge, one of his many admirers, to have "conquered his public like a lover; amusing it, flattering its taste, and proceeding thus step by step to the innermost sanctuary of its intelligence."

DE L'ESPRIT DES LOIS

Montesquieu's *Considérations sur les causes de la grandeur des Romains et de leur décadence* (1734), is a brilliantly written attempt to apply a scientific method to "historical understanding," to set forth—admittedly in a distinctly literary style—a sociological explanation of one phase of historical experience as a model for a new kind of positivistic history. This book is perhaps best read as a prolegomenon to Montesquieu's masterpiece, *De l'esprit des lois,* on which he worked for seventeen years.

De l'esprit des lois was first published in Geneva in 1748 against the advice of all the friends to whom Montesquieu had shown the manuscript. It was promptly placed on the Index, but it sold twenty-two editions in less than two years. It was a resounding success. Even so, it is a long, rambling, ill-arranged book that reflects the developments and changes in the author's point of view in the seventeen years he took to write it. But like *Les lettres persanes* and the *Considérations,* it is the work of an unmistakable master of French prose and of a man who knows how to entertain his readers as well as to instruct them.

By the *esprit des lois,* Montesquieu meant the raison d'être for laws, or the rational basis for their existence. Like Locke, he believed in natural law, but he was a much more thoroughgoing empiricist in his method than was Locke. Montesquieu believed that the way to learn about law was to look at the actual legal systems in operation in various states. Formal recognition of natural rights did not mean that men had positive rights. Mere a priori principles have little real value; it is important, he argued, to have the actual verifiable facts of the situations in which men find themselves.

Similarly, in his approach to the question of freedom, Montesquieu was less interested in abstract assertions of a general concept than in the concrete circumstances in which freedom had been or was being enjoyed. "Liberty," he wrote, "has its roots in the soil." He noted that freedom is more easily maintained in mountainous countries, such as Switzerland, than in fertile plains, and on islands, such as England, than on continents. Island and mountainous states find it easier to defend themselves from foreign invasion; in mountainous countries the very poverty of the soil encourages industry, frugality, and independence and so promotes individualism among the people. Another condition of freedom, he suggested, is that tranquility which comes from security. This can be enjoyed only where the constitution sets inviolable limits to the action of the state and where the law itself guarantees the rights of the individual.

Montesquieu always insisted that political liberty could never be absolute. "Freedom," he wrote, "is the right of doing whatever the laws permit." For example, he maintained that free trade did not mean that traders should do what they liked, for that would be to enslave the nation. Restrictions on traders were not necessarily restrictions on trade but might well be measures conducive to the liberty of all. Good laws were those that protected the common interest, and it was the mark of a free society that all the people be allowed to follow their own inclinations as long as they did not disobey the laws.

THE CONCEPT OF LAW

Montesquieu gives a rather bewildering definition of laws as "necessary relations," or "the relations which necessarily follow from the nature of things." Like most philosophers before David Hume, he failed to distinguish clearly between the normative laws of morals and the descriptive laws of science, but he was nevertheless conscious of having two tasks in seeking the raison d'être of laws. On the one hand, he was embarking on a sociological study of existing legal and political institutions, including the

institutions of positive law. Here Montesquieu the empiricist came to the front. On the other hand, Montesquieu the rationalist and the votary of natural law was seeking beyond his inductive generalizations for some general principles of justice and conduct, which he believed to be founded on reason.

> I first of all examined men, and I came to the conclusion that in the infinite diversity of their laws and customs they were not guided solely by their whims. I formulated principles, and I saw particular cases naturally fitting these principles: and thus I saw the histories of all nations as the consequence of these principles, with every particular law bound to another law and dependent on a further more general law.

At the highest level of abstraction, Montesquieu saw a uniform law—"Men have always been subject to the same passions"—but in various societies this higher natural law is expressed in differing systems of positive law. The systems differ because the external conditions differ. Montesquieu made much of the differences of climate and attempted to describe how different climates promote different customs, habits, economic arrangements, and religions. Much of political wisdom consists in adapting general principles to local circumstances. Solon was right to give people "the best laws they could bear."

The measure of relativism in Montesquieu affronted his friends among the *philosophes,* who believed in a kind of abstract universal individualism, but Montesquieu's method proved the more acceptable to social theorists of later generations. Émile Durkheim said it was Montesquieu who gave modern sociology both its method and its field of study. Montesquieu was ahead of his time in regarding social facts as valid objects of science, subject to laws like the rest of nature; he was also ahead of his time in seeing social facts as related parts of a whole, always to be judged in their specific contexts.

VIEWS ON RELIGION

Montesquieu resisted the notion that a "scientific" approach to problems of human conduct entailed determinism. He believed that God existed and that God had given men free will. "Could anything be more absurd," he asked, "than to pretend that a blind fatality could ever produce intelligent beings?" Assuredly, God had laid down the laws that govern the physical world, and "man, as a physical being, is, like all other bodies, governed by immutable laws." On the other hand, precisely because he is a rational, intelligent being, man is capable of transgressing certain laws to which he is subject. Some of the laws he transgresses are his own laws, namely positive laws, but governing the conduct of men are other laws antecedent to positive laws, and these are the general "relations of justice" or, in a more conventional term, natural law.

Montesquieu's attitude toward religion was very like that of Locke. He did not believe in more than a few simple dogmas about the existence of God and God's benevolence, but to that minimal creed he clung with the utmost assurance. On the other hand, Montesquieu grew to be much more cautious than Locke in his criticisms of religious institutions. In *Les lettres persanes,* Montesquieu did not hesitate to mock the Roman Catholic Church and clergy, but in later years he took care to avoid provocative utterances on the subject. In his biography of Montesquieu, Robert Shackleton gives an example of the philosopher's increasing wariness as revealed in successive drafts of the *Esprit des lois.* In the first draft of the chapter on religion, Montesquieu wrote, "Under moderate governments, men are more attached to morals and less to religion; in despotic countries, they are more attached to religion and less to morals." In the second draft Montesquieu introduced at the beginning of that sentence, "One might perhaps say that …." In the published version he cut out the remark altogether.

Much has been made of the fact that Montesquieu was reconciled to the Church of Rome on his deathbed. An Irish Jesuit named Bernard Routh got into the chateau at La Brède during Montesquieu's last illness, and in spite of the efforts of the Duchess d'Aiguillon to prevent him from "tormenting a dying man," the priest succeeded (or, at any rate, claimed to have succeeded) in leading the philosopher back to the path of devotion and repentance. The pope himself read Father Routh's account of Montesquieu's death "with the deepest reverence and ordered it to be circulated." Madame d'Aiguillon was able to rescue from the clutches of the Jesuits only one manuscript, that of the *Lettres persanes.* "I will sacrifice everything for the sake of reason and religion," Montesquieu had told the duchess, "but nothing to the Society of Jesus."

These dramatic scenes are perhaps less important to an understanding of Montesquieu's religious sentiments than is his behavior in less emotional times. He never asked his wife to give up her Protestantism, and he was always a fervent champion of religious toleration. At the same time, he remained on the best of terms with his several relations who were in holy orders in the Catholic Church. Besides, according to his "sociological" principle that every country had the religion its geographical and climatic conditions demanded, Montesquieu held that

Catholicism was the "right" religion for France, just as Anglicanism was the "right" religion for England. This is not to say that Montesquieu inwardly believed in more than a fraction of the teachings of the Catholic Church or that—until his deathbed repentance—the church regarded him as a true son. But he always detested atheism. To him the idea of a universe without God was *effroyable.* The concept of a loving creator played as prominent a part in his political theory as it did in that of Locke; indeed, whereas Locke had been content to see the church apart from the state, Montesquieu favored an alliance of organized religion with the government. In *Esprit des lois* he suggested that Christian principles, well engraved in the minds of the people, would be far more conducive to a good political order than either the monarchist notion of honor or the republican notion of civic virtue. Montesquieu was thus a deist in his heart and an Erastian in his politics.

See also Burke, Edmund; Durkheim, Émile; Locke, John; Philosophy of History; Political Philosophy, History of; Political Philosophy, Nature of; Rousseau, Jean-Jacques; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY MONTESQUIEU

Oeuvres de Montesquieu, 7 vols. Edited by E. Laboulaye. Paris, 1875–1879.

De l'esprit des lois, 2 vols. Edited by G. Truc. Paris, 1945.

Spirit of the Laws. Translated by Thomas Nugent. New York, 1949.

Oeuvres complètes, 3 vols. Edited by A. Masson. Paris, 1950–1955.

Considerations on the Causes of the Greatness of the Romans and their Decline. Translated by David Lowenthal. New York: Free Press, 1965.

WORKS ON MONTESQUIEU

Actes du congrès Montesquieu. Paris, 1956. Introduction by L. Desgraves.

André, Desiré. *Les écrits scientifiques de Montesquieu.* Paris, 1880.

Aron, Raymond. "Montesquieu." In *Main Currents in Sociological Thought,* Vol. I, translated by Richard Howard and Helen Weaver. New York: Basic, 1965.

Barrière, P. *Un grand provincial.* Bordeaux, 1946.

Berlin, Isaiah. "Montesquieu." In his *Against the Current: Essays in the History of Ideas,* edited by Henry Hardy. New York: Viking Press, 1980.

Cabeen, D. C. *Montesquieu: A Bibliography.* New York: New York Public Library, 1947.

Carrithers, David W., Michael A. Mosher, and Paul A. Rahe, eds. *Montesquieu's Science of Politics: Essays on "The Spirit of Laws."* Lanham, MD: Rowman & Littlefield, 2001.

Cotta, S. *Montesquieu e la scienza della societa.* Turin, 1953.

Dedieu, J. *Montesquieu, l'homme et l'oeuvre.* Paris, 1913.

Destutt de Tracy, Comte Antoine-Louise-Claude. *Commentary and Review of Montesquieu's Spirit of Laws.* Translated by Thomas Jefferson. Philadelphia: Burt Franklin, 1969.

Dodds, Muriel. *Les récits de voyages: Sources de l'Esprit des lois de Montesquieu.* Paris, 1929.

Durkheim, Émile. *Montesquieu et Rousseau.* Paris, 1953. Translated by Ralph Manheim as *Montesquieu and Rousseau.* Ann Arbor: University of Michigan Press, 1960.

Fletcher, F. T. H. *Montesquieu and English Politics.* London: Arnold, 1939.

Hulliung, Mark. *Montesquieu and the Old Regime.* Berkeley: University of California Press, 1976.

Manent, Pierre. *The City of Man.* Translated by Marc A. LePain. Princeton, NJ: Princeton University Press, 2000.

Pangle, Thomas L. *Montesquieu's Philosophy of Liberalism: A Commentary on the Spirit of the Laws.* Chicago: University of Chicago Press, 1973.

Richter, Melvin. *The Political Theory of Montesquieu.* Cambridge, U.K.: Cambridge University Press, 1977.

Shackleton, Robert. *Essays on Montesquieu and on the Enlightenment.* Edited by David Gilson and Martin Smith. Oxford: Voltaire Foundation at the Taylor Institution, 1988.

Shackleton, Robert. *Montesquieu: A Critical Biography.* London: Oxford University Press, 1961. The outstanding work on Montesquieu.

Shklar, Judith N. *Montesquieu.* Oxford: Oxford University Press, 1987.

Sorel, A. *Montesquieu.* Paris, 1887.

Maurice Cranston (1967)
Bibliography updated by Philip Reed (2005)

MONTGOMERY, EDMUND DUNCAN
(1835–1911)

Edmund Duncan Montgomery, a Scottish-American philosopher, anticipated in his "philosophy of vital organization" ideas of emergent evolution, the energetic nature of matter, and the pragmatic functioning of knowledge. Born in Edinburgh, he studied medicine in Germany in the 1850s, did research on cell pathology in London in the 1860s, and emigrated to America in 1870 with his sculptor wife, Elisabet Ney.

After a short-lived communitarian experiment at Thomasville, Georgia, the Montgomerys settled on Liendo Plantation, near Hempstead, Texas. There Montgomery wrote most of his philosophical articles and, in his later years, took an active role in community affairs. As chairman of the Waller County Democratic Party in the 1896 Bryan-McKinley campaign, he argued the dependence of political liberty upon economic reforms.

By 1867 Montgomery saw life as a power of certain compounds to reintegrate their chemical unity after damage, a power evolved by the inherent creativity of matter interacting in new combinations. He tested views of matter, mentality, selfhood, knowledge, and morality by this touchstone in over sixty articles in such journals as *Mind, Monist, Index, Open Court,* and the *International Journal of Ethics* and in five books. His major book was *Philosophical Problems in the Light of Vital Organization.*

Even inorganic compounds, Montgomery said, are inherently reactive, evolving in unpredictable ways by virtue of their peculiar composition and organization. Conservation of energy is thus wrongly viewed as requiring inertness of matter. Mentality is not dependent on a separate substance but is a capacity of certain complex organisms (chemical unities of a high order), heirs of evolution through foregone ages. Human knowledge and action are products of man's interplay with environment; they are instruments in preserving and enhancing well-being.

Some data of consciousness, such as kinesthetic and emotive states, seem to derive in each of us only from his own body, even though the body's activity thus perceived is in turn activated by outside stimuli. Others of our conscious states (such as visual data) are occasioned by features of either our own bodies or of external objects. Montgomery denied that this difference warrants the inference that there are two distinct kinds of substance, mental and material. All inferences from sensory data are conjectural. Data do not copy things but give "hieroglyphic signs" that permit discovery, prediction, and testing of natural relations among things.

Montgomery argued for a "naturalistic humanitarianism," a "religion of life," stressing ethical self-determination in a struggle against indifferent and hostile forces, to convey to the next generation a heritage nobler than the one received. Making common cause with those who wanted a religion and an ethic consistent with scientifically established knowledge, he added to classic criticisms of prevailing theologies and moral systems his own emphasis upon their failure to heed the full potentialities of men, the preeminent heirs of an evolution far from completed.

See also Consciousness.

Bibliography

WORKS BY MONTGOMERY

On the Formation of So-Called Cells in Animal Bodies. London, 1867.

Die Kant'sche Erkenntnislehre widerlegt vom Standpunkt der Empirie. Munich, 1871.

The Vitality and Organization of Protoplasm. Austin, TX: Gammel-Statesman, 1904.

Philosophical Problems in the Light of Vital Organization. New York and London: Putnam, 1907.

The Revelation of Present Experience. Boston: Sherman, French, 1910.

WORKS ON MONTGOMERY

A complete bibliography of writings by and about Montgomery before 1950 and an index of writings by and about him appear in Morris Keeton, *The Philosophy of Edmund Montgomery* (Dallas, TX: University Press in Dallas, 1950). The definitive biography is I. K. Stephens, *The Hermit Philosopher of Liendo* (Dallas, TX: Southern Methodist University Press, 1951).

Morris Keeton (1967)

MOORE, GEORGE EDWARD

(1873–1958)

George Edward Moore was born into moderately affluent circumstances in Upper Norwood (a suburb of London), the third son of D. Moore, M.D., and Henrietta Sturge Moore. The Sturges were prominent Quaker merchants and philanthropists. On his father's side there had been some tendency toward, and some prominence in, the practice of medicine.

Upon reaching eight, George Edward Moore commenced attendance at Dulwich College, a boarding and day school of excellent reputation located within walking distance of his home. In the ten years of his attendance there he acquired a thorough mastery of the classics. It was also at this time that he underwent a very painful experience. Having been converted around the age of twelve to "ultra-evangelism," he felt it his duty to preach the word of Jesus and to distribute religious tracts. He found these activities extremely repugnant and suffered much inward torment in carrying them out. This experience, which lasted two years or more, may account in some measure for his subsequent coolness to religious enthusiasms of any sort. Before leaving Dulwich College he was persuaded, through discussions with his eldest brother, the poet Thomas Sturge Moore, to adopt the view that was then known as "complete agnosticism." This seems to have been the view that there is no evidence in support of a belief in God's existence and almost as little in support of a belief in his nonexistence. So far as can be determined from his writings, Moore never departed from this view.

In 1892 Moore entered Trinity College, Cambridge, as a student in classics. At the beginning of his third year he changed his major concentration to philosophy and completed the moral science tripos in 1896. On the basis of a dissertation treating Immanuel Kant's ethics he was elected in 1898 a fellow for a term of six years. During the period 1898–1904 he carried on frequent and consequential discussions with Bertrand Russell, wrote *Principia Ethica,* presented several papers to the Aristotelian Society (to which he had been elected), and published a number of reviews and articles.

With the termination of his fellowship in 1904, Moore left Cambridge. Because of an inheritance he was still able to pursue his philosophical activities. He wrote articles, papers, and reviews, as well as the small volume *Ethics,* and gave a series of private lectures at Richmond. In 1911 he was invited to return to Cambridge as university lecturer. He lectured regularly at Cambridge from 1911 to 1925, first on philosophical psychology and later on metaphysics. In 1925 he succeeded James Ward as professor of mental philosophy and logic. His courses appear to have enjoyed a good deal of popularity among the more serious students of philosophy and had an immense influence upon the philosophizing going on in England at the time, as did his publications (notwithstanding that they consisted entirely of articles and papers).

In 1939, having reached the mandatory age of retirement, Moore gave up his professorship at Cambridge, though not his philosophical activities. These, with a few interruptions due to illness, he carried on to almost the very last years of his life, writing articles, editing his previous writings, working on problems, and holding discussions with friends and students. He died at Cambridge at eighty-five, survived by his wife, Dorothy Ely, whom he had married in 1916, and two sons, Nicholas, a poet, and Timothy.

Although Moore's life was extremely active in academic and philosophic spheres, it was almost without incident otherwise. Except for a brief sojourn in Germany in the summer of 1895, a somewhat longer stay in Scotland from around 1904 to 1908, and a couple of years spent during World War II lecturing in the United States, he resided entirely in England, mainly in or near Cambridge. His most noticeable personal trait appears to have been his intense and passionate absorption in philosophy. It is said, for example, that when discussing a question, whether with his professional peers or with a student, he gave himself wholly to the inquiry and viewed its progress with the constant fresh surprise of one considering a matter for the first time. Another trait that has been commented on was his lack of any intellectual pretensions (in spite of a formidable erudition) and an almost childlike naïveté concerning ordinary affairs.

Moore served as editor of the philosophical journal *Mind* from 1921 to 1947. The major honors that he received during his lifetime were the Litt.D. from Cambridge (1913), the honorary degree of LL.D. from the University of St. Andrews and election as a fellow of the British Academy (1918), and appointment to the Order of Merit (1951).

FORMATIVE PERIOD OF MOORE'S PHILOSOPHY

Moore's published philosophy falls into two distinct parts, divided by the year 1903. Although the writings published prior to 1903 are few and cover no more than five years, at least three different philosophical positions can be detected in them. In his first publication, a paper titled "In What Sense, if Any, Do Past and Future Time Exist?" (1897), Moore agreed wholly with F. H. Bradley. He argued that time does not exist, and he did so using Bradley's methods and premises, in particular the dogmas of internal relations and concrete universals and the principle that identifies reality with the absence of contradiction. When his conclusions, like the one that time does not exist, proved to outrage common sense, Moore was prepared to say that common sense is simply wrong, and he did so more than once.

One year later, in the essay "Freedom," Moore replaced Bradley with Kant as the philosopher with whom he was "in most agreement." What he agreed with most in Kant was the method of the transcendental exposition and the doctrine of synthetic necessary truths. He did not agree with the critical restrictions of Kant's philosophy or with what he took to be its psychological bias. He contended, for instance, that Kant was wrong in trying to conceive freedom in terms of the will (a psychological concept); freedom is rather to be understood and explained in terms of the idea of Transcendental Freedom, into which temporal relations do not enter. Thus, while accepting much of Kant's system and terminology, Moore continued to speculate in the critically unrestricted manner of the absolute idealists, maintaining that a reality transcending time and the senses is something that can be theoretically known and that must be theoretically known before the major problems of philosophy can be solved.

The next year, 1899, in the article "The Nature of Judgment," Moore adopted a third position. As part of his continuing attack upon psychologism in philosophy (an

attack he shared at the time with Russell), he proposed the doctrine, adumbrated in Thomas Reid, that mental acts and their objects are entirely separate existences. Applying this doctrine to Bradley's analysis of judgment, Moore concluded that the entire world—everything we can either think of or perceive with our senses—consists in qualitative universals, or what he called "adjectival concepts." These universals compose propositions, material objects, minds, and all other "complex objects." Not only do some universals (for example, red) exist through time, but some propositions also exist through time and are even objects of perception (for instance, the proposition that this book is red). Such universals and propositions are designated "empirical universals and propositions," as opposed to those that do not exist through time, such as the concepts *two* and *attribute,* which are called "a priori." This bizarre metaphysics, which might be termed "absolute realism" because according to it universals not only exist but, in fact, comprise everything that does exist, obviously repudiates all the major philosophical tenets to which Moore subscribed in his first essay: the dogmas of the nonreality of time, internal relations, concrete universals, and the transcendent monism that springs from them. Just as obviously it cannot be harmonized with the two-story world of phenomena and noumena that is attributed to Kant or with Kant's critical conclusions. Moore did, however, attempt to show that his realistic principles were compatible with, and even substantiated, Kant's method of transcendental exposition and distinction between a priori and empirical propositions and the doctrine of synthetic necessity. This Moore did by attempting to show that the possibility of a priori and empirical propositions, along with synthetic necessary truths, can be accounted for in terms of the realistic distinction between temporally existing (empirical) universals and nontemporal (a priori) universals and by shaping some of the arguments supporting this demonstration along the lines of a transcendental exposition. On the whole, though, the argumentation of "The Nature of Judgment," as well as of the articles and reviews that immediately followed (1899–1902), proceeds in the legislative, dogmatic manner of Bradley.

With this unstable amalgam of Bradley, Kant, and absolute realism, the first period of Moore's philosophizing came to a close. Marked by abrupt changes of doctrine, by either derivativeness (as in the first two positions adopted) or bizarreness (as in the third), it is recognizably an effort to find, rather than to express, a philosophy. It is therefore with some justice that these writings have been generally ignored by succeeding generations of philosophers, as they were ignored by Moore himself in his subsequent summations and compilations of his work. On the other hand, a complete understanding of Moore's later philosophy is difficult to arrive at without some familiarity with these earlier works. It will then be understood, for instance, that the charge sometimes leveled against Moore that he criticized the metaphysical theses of philosophers like Bradley piecemeal, without attempting to comprehend them fairly and in their entirety, is groundless. It will be understood, for instance, that in attacking items of Bradley's metaphysics Moore was attacking not only a system of thought with which he was thoroughly conversant but one to which he had himself once been most strongly attracted.

MOORE'S PHILOSOPHY PROPER

The system of philosophical thought and method that has come to be associated with Moore's name and that he was alone concerned to defend issued fully formed in the volume *Principia Ethica* and the essay "The Refutation of Idealism" in 1903. This is not to say that no alterations thenceforth took place in the body of Moore's philosophical doctrines and aims. They did. For example, with the passage of time Moore became increasingly concerned with eliminating from the world various entities, such as propositions, that his principles generate. The theory proposed in "The Refutation of Idealism," that we directly perceive material things, was replaced by a disjunction of theories respecting the relation between sense data and material things. And the note of philosophical optimism that expressed itself in *Principia Ethica* and "The Refutation of Idealism" in the view that solutions to the problems under discussion have either been completed in their pages or are on the brink of completion finally gave way to a note of philosophical pessimism and puzzlement. But in its main outlines what might be called Moore's philosophy proper was now permanently formed.

As will be seen in subsequent discussion, the tenets of this philosophy are largely based on the principle that sentences such as "I think of *X*" describe (*a*) mental acts and (*b*) objects related to but distinct from those acts. From 1903 until the late 1930s Moore almost invariably interpreted this principle realistically, and even after the late 1930s, when he was prepared to admit that the *esse* of sense data is *percipi,* this realist tendency continued to make itself felt in his philosophizing, especially with respect to universals. Moore's philosophy proper resembles, therefore, the absolute realism of "The Nature of Judgment." There exists, however, a fundamental metaphysical difference between the two positions. This differ-

ence lies in the fact that Moore's absolute realism of 1899 is reductionistic, being the view that everything can be resolved into qualitative universals, whereas the realism he enunciated in 1903 and afterward is, in intention at least, nonreductionistic. Thus, within the compass of things that are, Moore now included both particulars—for example, material things—and universals, and though he was not perfectly clear about just what a universal or a particular is, he wanted to maintain neither that universals can be resolved into particulars nor that particulars can be resolved into universals. His new view was that each sort of thing is what it is and nothing else (or, in the words of Bishop Butler, quoted on the frontispiece of *Principia Ethica,* "Everything is what it is, and not another thing").

The most striking and significant difference between Moore's philosophizing prior to 1903 and his philosophy proper lies not, however, in doctrine or even in the mechanics of method (though differences here are pronounced) but in the attitude and style of his philosophizing. These now project the familiar picture of Moore: the picture of a cautious and probing observer, attempting by the patient dissection and scrutiny of minute and hardly distinguishable objects to set straight the confused descriptions by philosophers of what is the case. This posture of Moore's lends to his philosophizing the appearance of a completely empirical inquiry whose conclusions represent only what is found or not found to be the case, as opposed to what is merely thought to be or not to be the case. It is in the solvent of this empiricist posture that Moore's initial philosophical optimism, as one might predict, evaporated into pessimism and puzzlement. For the principle from which it originated, that sentences such as "I perceive *X*" describe acts of mind and distinct objects, is itself something no amount of observation would seem to confirm or lend substance to.

In the first of the lectures that he delivered in 1910–1911, some forty years later published under the title *Some Main Problems of Philosophy,* Moore listed the main topics of philosophy as three. The first and primary aim of philosophy, he said, is to provide a metaphysical inventory of the universe, that is, "a general description of the *whole* of this universe, mentioning all the most important kinds of things which we *know* to be in it, considering how far it is likely that there are in it important kinds of things which we do not absolutely *know* to be in it." The second aim is epistemological: to classify the ways in which we can know things. The third topic of philosophy is ethics.

In "A Reply to My Critics," published in 1942, Moore again divided his philosophical discussion into three parts: ethics, theory of perception, and method. Although this alteration in the classification of topics indicates certain real alterations in Moore's interests and views, it will be convenient to treat his philosophy proper under the five heads mentioned: method, metaphysics, general epistemology, theory of perception, and ethics.

METHOD. By Moore's "method" will be understood the topics encompassed by the following: (1) The question: What did Moore believe he was doing in philosophizing, that is, what project did he think he was engaged in? (2) The question: How did he attempt to carry out this project? (3) Certain questions that are often raised in specific connection with Moore's method, such as: What is the role of common sense in his method? What is the role of analysis?

Moore's intentions. It has been suggested by some of his commentators that what Moore was trying to do was to analyze ordinary language, to defend common sense, or to recommend ways of speaking. As an answer to the question What was Moore *actually* doing? it is possible that one or all of these suggestions may be true. But it is clear that none of them describes what Moore believed he was doing.

Moore's conception of what he was doing originated in the following two principles, to which he consistently subscribed: the principle that sentences like "I think that *P*" and "I perceive *X*" designate acts of consciousness, on the one hand, and objects related to but distinct from those acts, on the other; and the principle that every object of consciousness is either a simple, in which case it is unanalyzable, or a complex, in which case it always possesses a definable essence in terms of which it is the sort of thing it is and not some other sort of thing. The first principle makes it appear as if there should be discoverable as the objects of consciousness a great many more kinds of entities and properties than persons ordinarily envisage, and these entities and properties should comprise, at least in part, what is objectively in the universe. When applied to these entities, the second principle makes it appear as if every complex object should be unequivocally reducible to simples. But this picture of things raises a question: If the constitution of the universe is both so determinate and so open to consciousness, why is it that there has been so much disagreement and confusion in the attempts of philosophers to describe it? And to this question the most obvious answer seems to be that past errors and confusion in philosophy have arisen

either from inattention on the part of philosophers to the objects of their consciousness or from a lack of clarity and preciseness in their statements and questions.

In fact, the two major concerns of Moore through the period 1903–1911 directly correspond to the above outline of subject matter. Primarily, Moore wished to determine what sorts of entities or properties fall within the province of his particular inquiry, for example, ethics, theory of perception; to classify these entities (where deemed necessary) as simples or complexes; and to analyze the essences of the complexes. Second, and always as a project subordinate to the first, he wished either to direct the reader's attention to the objects of consciousness that pertain to the inquiry at hand or to lay bare the ambiguities and unclarities of the terms customarily used by philosophers in conjunction with the inquiry at hand, and to supply "precising" definitions of the terms that he intended to use.

After the lectures of 1910–1911 an increasing concern with terminological questions was detectable in Moore's writing. This concern is traceable to an apparently growing conviction on his part (as well as on the part of his contemporaries) that the terminological sources of philosophical error and confusion are much more subtle, deeply rooted, and pervasive than he had originally thought and much more intimately connected with the logical grammar of ordinary language. In the last connection it is worth recalling that certain of Moore's contemporaries eventually decided that the root and cure of all philosophical problems lay in terminological confusion and clarification.

Moore never went so far as to assent to the last conclusion. He did, however, relinquish his earlier view that the primary concern of philosophy is to observe and delineate the entities objectively making up the universe. By 1940, when he composed his "Reply to My Critics," he described himself as engaged, not in the analysis of facts, but in the analysis of concepts. Although he was unclear about what the relation is between concepts, the entities objectively making up the universe, and verbal expressions, he appears to have thought that concepts are not only distinct from and (at least from their side) independent of their verbal expressions but also distinct from the entities objectively making up the universe (for otherwise, in analyzing concepts, he would be resolving philosophical doubts and questions in a way that he agreed that one cannot do and that he was not doing). But just what, then, are concepts according to Moore? In "A Reply to My Critics" he did not say. It is not improbable, however, that Moore had come full circle, back to something like Bradley's psychologically grounded view of concepts, which, ironically, served in "The Nature of Judgment" as the launching platform for Moore's philosophy of realism.

Moore's procedure. In much the same way that Moore's doctrine of mental acts and objects dictated his conception of what he was trying to do, it also dictated his conception of how to accomplish what he was trying to do. It is evident, for instance, that once sentences like "I think that *P*" and "I perceive *X*" are interpreted according to that doctrine, it must seem unjustified to argue in the legislative manner of Bradley, which Moore employed in "The Nature of Judgment" and the essays previous to it. If the objects of acts of judging, perceiving, and thinking are entities distinct from, and indeed independent of, those acts, then whatever we can learn about those objects must be by means of synthetic observations, not a priori thought. Moore throughout his philosophy proper adhered to this viewpoint. Where he conceived himself as primarily engaged in reporting, classifying, and analyzing the entities objectively constituting the universe, he assumed that he was basing his reports and analyses on observation. Where, as in "A Reply to My Critics," he conceived himself as engaged rather in analyzing concepts, it is evident that he thought of concepts as comprising some sort of object he was engaged in observing.

As was noted previously, this picture of philosophical inquiry suggests that philosophical questions have determinate and easy solutions that it might be expected all philosophers will agree on. Moore's explanation of this discrepancy between expectation and fact—that the disagreements and failures of philosophers stem either from a lack of attention to what is present to their consciousness or from terminological unclarities—suggests, in turn, that in order to be certain we are observing what we think we are we must make sure both that our attention is directed to the right objects and that we know the precise meanings of the terms we are employing in our thoughts.

It turns out, however, that even with this supplement observation fails to bring about the results that Moore anticipated or that his assumptions might have led him to anticipate. The answers to philosophical questions remain stubbornly shrouded in obscurity and disagreement. Moore was therefore compelled to add to his methods and procedures. In cases where he felt there was no conclusive answer to a question, he resorted to what might be termed the principle of weighted certainties. If, for instance, he felt that proposition *A* possessed more certainty than proposition *B*, or if he felt that he knew the

truth of *A* with more certainty than that of *B,* he would refuse to deny the truth of *A* on account of some argument based on *B.* In short, a lesser certainty (according to this principle) cannot rationally overturn a greater certainty per se (though a number of lesser certainties, cohering together, may). Moore also employed, in the same connection, the scholastic method of citing all the plausible arguments that can be advanced for or against a thesis in order to indicate its degree of credibility. And finally, in order to discredit a thesis (usually a thesis of skepticism), he employed either a *reductio ad absurdum* argument or what might be called a paradigm argument. He pointed out, for example, that the skeptic who maintains that we cannot know there are other persons is already contradicting himself by supposition in referring to the plural, *we.* Or he argued that if such-and-such is not an instance of knowing, then no one has ever known anything and there cannot be such a thing as knowing.

When these norms for evaluating philosophical conclusions are arranged in order of their indefeasibility, it would seem that where observation unequivocally reveals just what a thesis represents to be the case, according to Moore the thesis is indefeasible. Thus, Moore maintained that when we look at an inkwell we directly perceive a sense datum and that this claim is indefeasible in that observation unequivocally presents us with a sense datum. Where a thesis can be shown to contain an evident contradiction, according to Moore it is conclusively disproved. Thus, one can affirm with certainty that the skeptic who maintains that *we* cannot know other persons exist is wrong. Where the principle of weighted certainties or the method of citing plausible arguments has to be invoked, Moore would generally grant that answers are not conclusive or indefeasible, although there may be more to be said in favor of one answer than another. In certain cases, however, it would appear that the certainties or feelings of certainty (Moore rarely distinguished between the two) attaching themselves to a thesis are so absolute or overpowering that no denial of the thesis is either psychologically or rationally (in view of the principle of weighted certainties) possible.

Common sense. It is tempting, but wrong, to suppose that because Moore defended common sense, common sense constitutes a court of last appeal in his philosophy. Indeed, the very fact that he described himself as defending common sense indicates that it cannot.

In his works Moore used the term *common sense* to refer to two different, but related, things. He sometimes meant by it, he said, simply those beliefs that men universally or almost universally subscribe to at some particular epoch. At other times he meant either those beliefs that we are naturally inclined to hold or the propensity that issues in such beliefs.

Although there may exist a very intimate causal connection between these two forms of common sense, they are not one and the same thing. As the "universal" belief of men at a particular epoch, common sense can change, and Moore in fact argued that it can. As a natural tendency to believe something, common sense would not seem susceptible of change. It must be remarked, however, that Moore never explicitly drew the above distinction or attempted to "analyze" the notion of common sense beyond saying that it consists in the universal belief of men at a particular time. In practice, however, he would seem to have maintained that although both forms of common sense possess a certain amount of presumptive credibility, it is essentially as a natural tendency that common sense provides a foundation for philosophical conclusions. It does this in two ways. When we try to deny the latter form of common sense we find it virtually impossible to do so because what we naturally tend to believe keeps slipping into our assertions. We thus find ourselves contradicting ourselves by supposition, like the skeptic who says that *we* cannot know persons exist. On the other hand, what we naturally tend to believe will have attached to it some degree of certainty. This degree varies, it seems, from an absolute quantity, which makes dissent really impossible, to a quantity that only inhibits dissent. For example, Moore said he was naturally disposed to think that what he always saw directly when viewing a material thing was the surface, or part of the surface, of the material thing, but he finally decided it would be nonsense to maintain that he did.

Moore, then, defended common sense by showing that certain beliefs that we are naturally inclined to hold, and consequently that most men do hold, are supported by the principle of weighted certainties or by showing that the traditional counterclaims of skeptics are self-contradictory. He did not argue conversely that because a certain belief is a belief of common sense it is ipso facto indisputably true or need not be subjected to assessment.

Analysis. When Moore described himself as "analyzing," he conceived of himself as picking out and naming the essential constituents of complex objects. In his earlier works he viewed himself, when analyzing, as picking out and naming the essential constituents of various objective entities and facts; in his later works, as picking out and naming essential constituents of various complex concepts. In his reply to C. H. Langford in "A Reply to My

Critics," he explicitly denied that he ever engaged in the analysis of verbal expressions.

This last denial may not be disingenuous, but it is misleading. Moore maintained that the only proper meaning of the term *analyzing verbal expressions* is merely counting the letters in a sentence, noting the order of the letters, and so on. If this is true, then obviously Moore never engaged in analyzing verbal expressions, and just as obviously his denial that he did is trivial.

It may therefore be more significant to ask whether Moore engaged in linguistic analysis, where "linguistic analysis" is used as a technical term designating the following practices or inquiries: the determination of the meaning of a word or expression (not excepting the determination of its dictionary meaning); the determination of the various senses of a word or expression; the determination of the ordinary use of a word or expression; and the determination of discrepancies between the philosophical and ordinary uses of a word or expression. In all these senses of the technical term *linguistic analysis,* Moore, it is clear, engaged frequently in linguistic analysis. However, as was pointed out previously, he engaged in linguistic analysis never as an end in itself but always as an inquiry subordinate to the ascertainment of facts or the determination of the essential constituents of things or concepts.

METAPHYSICS. By the term *metaphysical* Moore sometimes meant to refer to nonnatural objects or qualities, that is, objects or qualities that are constituents of the universe but not of temporal events (or nature); sometimes he meant to refer to the sort of philosophical inquiry that concerns itself with the overall constitution of the universe. It is in the latter sense that the term *metaphysics* is being used here.

Although not without expressing some doubts on the matter, Moore inclined to the view that the things to be *found* in the universe are broadly of two sorts: those things that exist and those that simply *are* but do not exist. A third class of things consists of those that neither exist nor are; they simply are not. As Moore conceived of these categories, the main ontological division is between the things that are and those that are not. For the former, whether they exist or simply are, comprise the objective constituents of the universe and have equal claim to philosophical investigation. The latter are merely "chimeras" or "imaginary objects."

Moore suggested at least three ways of distinguishing between things that are and things that are not. First, the former possess the property of being; the latter do not. Second, borrowing from Russell's theory of descriptions, Moore claimed that whereas an object that *is* or possesses *being* can be the bearer of a name, imaginary objects can be described only by incomplete symbols. Thus, for example, "centaur" is not the name of anything (for there is nothing to bear the name), whereas "chair" is a name. Third, if a thing's *esse* is *percipi,* then it is an imaginary object and actually *is not.* There are only thoughts of centaurs, for example; there are not centaurs independent of our thoughts. Hence, centaurs are imaginary objects. Moore, however, discovered difficulties with the last description in that he thought it likely that the *esse* of acts of consciousness and sense data is *percipi,* and at the same time he did not want to say that acts of consciousness and sense data *are not.*

Where he did distinguish between mere being and existence (and in places he did not), Moore generally cited two grounds as the basis for the distinction. Sometimes he argued that whatever endures through parts of time exists; what does not endure through parts of time does not exist. He also sometimes argued that whatever can be an object of sensory perception exists. Although he never discussed the connection between these two criteria for existence, it seems from what he said on other matters that the temporal criterion states both a necessary and a sufficient condition for existence, whereas the sensory criterion states but a sufficient condition. For in Moore's system it is possible that material things are never the contents of sensory perception, but they are, par excellence, things that exist.

In addition to existence, being, and nonbeing, Moore treated at length and in detail the category of *reality.* Although painstakingly carried out, his thoughts on this subject possessed little overall coherence. In *Principia Ethica* he equated reality with existence; in the lectures of 1910–1911 he equated it simply with *being.* In the same lectures he referred to reality as a property; on the other hand, in *Philosophical Studies,* in the essay "The Conception of Reality" (1917), he denied that reality is a property. What he consistently maintained is expressed in his rejection of Bradley's view that reality possesses degrees and that the highest degree of reality is at an extreme remove from material things. Moore denied that reality possesses degrees. But if it does, he said, then he wanted to maintain, in opposition to Bradley, that material things possess the highest degree of it.

Within the category of *being* Moore distinguished between three kinds of objects: particulars, truths or facts, and universals. He generally, though not always, argued as if particulars may be divided into five sorts:

material things, sense data (for instance, *patches* of yellow), acts of consciousness, volumes of space, and intervals of time. He did not appear to think that the term *mind* refers to a particular substance in which acts of consciousness inhere. The theory he seemed to favor is that acts of consciousness are located in material bodies and are properties of material bodies and that the word *mind* stands for something like a logical construction from acts of consciousness. Truths or facts are the objects of true beliefs and comprise such things as mathematical equations—for example, 2 + 2 = 4—and the references of indicative sentences, such as "Tom stood to the left of Henry." Universals are again divisible into three sorts: relations, relational properties, and a third sort of universal that is neither a relation nor a relational property. Moore never provided an essential description of this third sort of universal, but he cited as clear-cut examples of it numbers and nonnatural qualities or objects, such as *good,* and as possible examples of it shades of color.

Of the three sorts of *being*—that is, particulars, facts, and universals—particulars alone exist; facts and universals merely *are*: This, at least, was Moore's view when he was prepared to grant that a significant distinction holds between existence and mere being. It was also his view that the only *substantial* things we are acquainted with are material bodies and acts of consciousness.

It should be remarked that the above inventory of the universe was not considered by Moore to be exhaustive. There may be things in the universe that we are in fact ignorant of or must even necessarily remain ignorant of. For example, Moore thought it is not impossible that God exists but found no evidence for maintaining that he does. Moore described himself as being certain, though, that all the things that have been mentioned as *being* or *existing* do constitute at least some of the constituents of the universe.

GENERAL EPISTEMOLOGY. Although a number of the topics that have been treated under the heading of Moore's methodology might as reasonably be considered under the heading of his general epistemology, and vice versa, under his methodology it was asked what Moore in his philosophizing was attempting to do and how he was attempting to achieve his aims, whereas under his epistemology these quite different questions are being asked: (1) What, according to Moore's philosophic account, does knowledge consist in? (2) Does knowledge, as so conceived, exist, and if it does, what is it knowledge of?

(1) What does knowledge consist in? Moore's basic metaphysical and methodological principles dictate that in order to discover what knowledge is, it is necessary to distinguish between the different senses (if there are different senses) of the verb "to know" and then to pick out and analyze the particular objects denoted by these senses of "to know" and the relations (if any) that hold between them.

Throughout his earlier writings and the lectures of 1910–1911, Moore was convinced that careful observation of facts and careful differentiation of terms provide us with the following results. First, every instance of cognition ultimately consists in an act of consciousness and, distinct from the latter, in an object. Second, an act of consciousness can exist only as long as the corresponding instance of cognition exists. Thus, when I cease to see a sense datum, my *seeing* of it ceases to exist. The object of cognition, however, may or may not exist after the act of consciousness to which it is related ceases. This is a matter to be decided by empirical considerations. Third, it is conceivable that an act of consciousness and its related object—for example, a sense datum—exist in two different locations. "It seems to me conceivable," wrote Moore in *Some Main Problems of Philosophy,* "that this whitish colour is really on the surface of the material envelope…. My seeing of it is in another place—somewhere within my body."

Reflecting this analysis of cognition and its objects, Moore thought that he could pick out four different ways of knowing and, corresponding to them, four different senses of the verb "to know." First and basic to an understanding of any other sense of "to know" is the sense in which "to know" stands for cases in which the relation between the object cognized and its correspondent act of consciousness is similar to or identical with the relation that a patch of color has to the consciousness of a person seeing that patch of color. This is knowledge by direct apprehension or knowledge by acquaintance. A second sense of "to know" represents cases in which the relation between the object cognized and the correspondent act of consciousness is similar to or identical with the relation that, for example, a hat on a table has to the act of consciousness of a person who is remembering that his hat was on the table. Thus, he knows that his hat was on the table, but neither the hat and table nor any sense data that were connected with the hat and table are directly present to his consciousness. This is knowledge by indirect apprehension. At least until 1911, Moore described himself as uncertain whether knowledge by indirect apprehension always necessitates direct apprehension of a proposition, by means of which, following Russell's theory of knowledge by description, one is made aware of the object indi-

rectly apprehended, but he was inclined to think it does. Third, there is a sense of "to know" that represents cases in which the following complex relation between acts of consciousness and objects holds: there is an act of consciousness; there is a proposition directly apprehended; this proposition is in fact true; we believe that it is true; and we believe that it is true because of some further relation or condition that it satisfies. What this further condition is Moore left undecided, though one might plausibly suppose that it had to do with conclusive evidence. In any event, Moore termed this way of knowing "knowledge proper." Last, and involving the previous senses of "to know," is that sense of "to know" in which we describe a person as knowing something, such as the multiplication table, even though he may not at the time be conscious of anything. We imply, in such cases, that the person in question has at some time known, in one of the other three senses of "to know," the multiplication table.

Moore also distinguished between what he termed "immediate knowledge" and "knowledge by direct apprehension." Immediate knowledge is a species of "knowledge proper." Thus, immediate knowledge is distinguished from knowledge by direct apprehension in that the latter does not require the presence of a proposition (for instance, I can directly apprehend sense data), whereas the former does. It is specifically the "kind of way in which you know a proposition to be true—really know it, not directly apprehend it—when you *do not* know any other proposition from which it follows" (*Some Main Problems of Philosophy*).

(2) Does knowledge exist? and of what things? Since Moore, purportedly on the basis of observation, resolved knowledge into a certain complex of objects, it is evident that knowledge, or "acts of knowing," exists in his view. The question of its existence becomes, indeed, a psychological or introspective question (it would seem) rather than an epistemological one.

In dealing with the question of what sorts of things are known, Moore generally, however, treated it as a de jure or epistemological, rather than a de facto or psychological, question. Thus, in defense of asserting that such-and-such a sort of thing can be known, he would sometimes appeal to the principle of weighted certainties (for example, he would ask, "Which is more certain—that I know that I am holding a pencil in my hand or that the principles of the skeptic are true?") and sometimes to paradigm arguments of the sort "If I do not know that *P*, then I can know nothing." In this connection, it is worth noting that Moore sometimes argued de jure that we know such-and-such a sort of thing exists although he was unable to discover by introspection the way in which we know it. For instance, he insisted that we know the existence of material things, such as the earth and our own body and other bodies like it, but he was unable to determine with any certainty in just what way we know their existence.

Moore claimed that in addition to the existence of material things, we know the existence of our own acts of consciousness and our own sense data, past events in our lives, the being of universals and nonnatural qualities or entities (such as good), the existence of other minds, synthetic necessary truths, and practically all matters of fact that are commonly thought to be known—for instance, that Caesar crossed the Rubicon, that Earth goes around the sun, and so on. Thus, in contrast to the skeptic, who traditionally maintains that the circumference of knowledge is much smaller than people ordinarily think, Moore appears to have maintained that it is much larger than people ordinarily think. For it is doubtful that people ordinarily think they know the existence of some things called sense data and acts of consciousness or the being of some things called nonnatural qualities or universals.

THEORY OF PERCEPTION. It is apparent that Moore's general epistemological principles and the premises that he operated with in his methodology enforce an empiricist approach to knowledge. They imply that all knowledge must finally be based on the observation of objects presented in experience. In three respects, however, Moore consistently parted company with traditional empiricists. He refused to limit the term *experience* to mean simply sensory experience. That is, he wanted to maintain that many sorts of objects other than those discovered by the senses are the objects of acts of consciousness—for example, timeless facts, relational universals, and nonnatural qualities. He also wanted to maintain (following Kant) that there are necessary synthetic truths and that we can apprehend these truths. And finally, he was never willing to reject what seemed to him a certain truth—for instance, that he was holding a pencil—because some less certainly true analyses or philosophical principles were incompatible with it. Thus, he consistently refused to acquiesce in the skeptical conclusions that traditional empiricism and indeed, it seems, his own empiricist principles tend to establish.

At the same time, these principles seem to have had two distinct effects on Moore's overall philosophizing. First, as time passed his interests converged on theory of perception and questions concerning our knowledge of an external, material world. Second, the skeptical conclu-

sions that empiricism appears to foster produced a constantly widening cleavage in his philosophy between what he wanted to assert preanalytically to be certainly true and what his analyses permitted him to assert to be certainly true. This ever-growing cleavage is nowhere more apparent than in his theory of perception.

In his essay of 1903, "The Refutation of Idealism," Moore maintained that material things can be directly apprehended and therefore can be known to exist with as much certainty as one's own acts of consciousness. Soon afterward, however, Moore was led to change his mind on this crucial point, apparently by what has sometimes been referred to as the argument from synthetic incompatibility. This argument assumes that the looks of things are the objects a person directly perceives, and then, because the looks of things change when the thing itself is not presumed to be changing, the argument concludes that what a person directly perceives is not the material thing or a part of its surface but some other kind of object that possibly exists only when he is perceiving it. This "other kind of object" is called by Moore a sense datum.

Moore had trouble in deciding just what a sense datum is: whether it is a particular or a universal, whether it is something like a color (in the case of visual sense data) or some other sort of thing. His final position on this question would seem to be that a visual sense datum is a *patch* of some color: The patch, which is a particular, is related to the color, which is a universal of the third sort (that is, it is neither a relation nor a relational property) in the way something is related to that which, in part, is spread over it.

The main problem concerning Moore in his theory of perception was not this, however, but the question of the relation between sense data and the material things to which they "belong." Although Moore concerned himself with this question in a series of remarkably closely reasoned essays, commencing with "The Status of Sense-Data" (1914) and concluding with his last published article, "Visual Sense-Data" (1958), he was never able to arrive at a definite or even a very plausible answer. Throughout most of these essays he presented three alternative theories as possibly true: phenomenalism, or what he termed the Mill-Russell theory—that is, the view that a material thing is simply a "logical construction" of sense data; some form of representational theory (varying from the theory that the relation between sense data and material things is an unanalyzable relation of "appearing" to causal theories resembling John Locke's); and the theory that visual sense data are identical with parts of the surfaces of material things. With all these alternatives he found grave difficulties and, indeed, was led in the end to dismiss the last as constituting, at least in most cases of perception, nonsense. But if we do not directly perceive material things or their surfaces (and Moore was willing to grant that perhaps we never do), and if by "material things" is meant nothing so Pickwickian as a logical construction of sense data (and Moore would have tended to agree that nothing so Pickwickian is meant), how can we possibly know that material things exist? Moore, in one of his last lectures, "Four Forms of Scepticism," suggested none too plausibly that we know their existence by analogical or inductive arguments.

ETHICS. As in the other branches of his philosophy, Moore was confident in his earlier works on ethics of the correctness and finality of the results he set forth; this confidence diminished constantly in the solvent of his empiricist methods of inquiry and was replaced in his later works by no more than tentative agreement with his earlier views. Also, as in the other branches of his philosophy proper, Moore's viewpoint toward both the proper method of ethical inquiry and the nature of the findings to be anticipated stemmed directly from his originally realist presuppositions.

Ethics, as Moore conceived of the discipline, takes the form of a partly definitional, partly descriptive science, resting on observation and induction. His theory is not, however, naturalistic. The fundamental object of observation for ethics, goodness, is a nonnatural quality or entity, according to Moore, and thus is one that neither exists through parts of time nor presents itself through sensory experience. On the other hand, his theory is not "metaphysical": it does not purport to *define* this fundamental entity or quality of ethics in terms of some *other* nonnatural entity or quality. Indeed, a main point of Moore's theory is that the fundamental entity of ethics cannot be defined at all and that any attempt to define it must commit what he termed "the naturalistic fallacy." This is essentially the fallacy that results from construing the "is" of attribution as an "is" of identity, and thus supposing, for example, that because pleasure is (attributive "is") good, good is identical with pleasure.

The fundamental object of ethics is the simple quality or entity good; being simple, good is unanalyzable and indefinable. One can only say that good = good. This is the outcome of the first and most basic inquiry any science of ethics must engage in, the answer to the question What is good?, where this ambiguous question is understood to ask for a definition.

A second important inquiry that the science of ethics undertakes is to determine what are the preeminent goods obtainable by men. Since the term *good* is here being used substantively (and not adjectively) to refer to complex wholes to which the quality or entity *good* attaches, definitions or analyses of such goods are possible, in the sense that the parts making up the wholes in question can be set forth. On the other hand, because the quality "good" is indefinable, it is not possible to determine which things are and which are not good analytically. This can be determined only by perceiving which wholes possess good, and to what degree or amount. Since they do not rest on any external evidence, such perceptions were termed by Moore "intuitions," and it is for this reason that his theory of ethics is sometimes called "intuitionistic." A further character of these perceptions is that when we perceive that a certain whole possesses in itself a certain amount of good, we perceive at the same time that any similar whole must possess in itself an equal amount of good. Thus propositions of the sort "Such-and-such possesses in itself such-and-such amount of good" or "Such-and-such is intrinsically good" express truths that are both synthetic and necessary.

The determination of what things are preeminently good is complicated by two factors. First, substantive goods are organic unities or wholes; that is, the good of a whole is not simply equal to the sum of the goods of its parts. This makes it impossible to determine what things are good and in what amount merely by determining previously the amount of good attaching itself to basic units of experience and adding up these units. Second, it is in fact difficult to separate, in our perceptions or intuitions, organic wholes from their consequences; hence, in assessing goods-in-themselves we are likely to include the good accruing to causal consequences of those wholes. In order to avoid the last sort of error, Moore proposed that we isolate the organic unity we are concerned with by imagining it as alone existing in the universe and then asking whether it is better that it exists or does not.

Applying this method to the question What are the preeminent goods obtainable by men? Moore maintained that "it is obvious that personal affection and aesthetic enjoyments include by far the greatest goods with which we are acquainted."

The third major inquiry of ethical science encompasses the questions of traditional casuistry: What are our duties? What is their order of precedence? What actions as a rule are right?, and so forth. The answers to all these questions are predicated, in Moore's system, on the assumption that unlike the term *good,* the terms *right, duty, virtue,* and so on are definable. They are all, in fact, definable in terms of good. When we say that a certain sort of action is right or our duty *we mean* that it is productive of the greatest amount of good in comparison with any possible alternative action. Thus, in determining duties and right actions we must not only determine what things are good in themselves but what causal effects actions will have, and this is an almost impossible task, except when conceived in rather short-term measures. As so conceived, Moore generally argued that the rights and duties enjoined and sanctioned in conventional morality are indeed just what the science of ethics shows to be our rights and duties.

CRITICISM OF MOORE'S PHILOSOPHY PROPER

Moore, in his last writings, confessed that he had not been a good answerer of questions, and if by a "good answerer to a philosophical question" is meant one who leaves the question settled or seemingly close to being settled, it is hard not to agree. In his ethics Moore provided simple, clear-cut answers to the problems and questions of traditional ethics, but their very simplicity (like saying the world is made of water) produces its own disbelief, and this disbelief is borne out by subsequent reflection. For example, if good is a simple objective quality of some sort, why should persons be concerned with maximizing it? In the other branch of inquiry with which he was primarily concerned, theory of perception, Moore failed even to provide clear-cut answers or decisions.

Again, if by "good philosophical answers" are meant answers that can be formed into a consistent system, it must be agreed that Moore is not a good answerer. In his philosophy there are a great many loose ends that he never tied together or attempted to tie together. For instance, he made no attempt to tie together his discussions of the two questions What is the relation of sense data (i.e., patches) to universals? and What is their relation to material things? In the same connection, Moore sometimes admitted that he was inclined to hold at one and the same time two incompatible views (as on the question whether the surfaces of material things are directly seen) and was unable to choose between them.

On the other hand, if a good philosophical answer is conceived as one that is closely reasoned and demands and instills close reasoning on the part of its auditor or reader, then Moore was a good answerer. Studying Moore, it can be fairly said, is like holding one's mind to a whetstone: A mind composed of good stuff is bound to be sharpened (and one of poor stuff to be dulled).

Further, if philosophy is conceived as an inquiry rather than a closed system, Moore was a good answerer. It is the essence of inquiry that every problem considered be freshly considered, that pat answers be abjured, that truth be placed ahead of remaining consistent or reaching conclusions, and that alternatives be given a hearing and their merits weighed. These are precisely the virtues of Moore's philosophizing.

A more serious objection that can be urged against Moore is that there are a certain number of philosophical prejudices that he adopted without question, but that he ought to have questioned. It is arguable, for instance, that he adopted without question the principle that there is something called an act of consciousness and something called an object of that act. Applied to the various topics of philosophy, this principle produces all sorts of obvious nonsense: a ridiculous proliferation of entities, and so on. Why, it may be asked, did Moore not seriously question this presupposition and remove it? And if he had, might he not have arrived at sound conclusions instead of the perplexity that he does in fact arrive at?

There is unquestionably a good deal of justice in this last objection. Yet, with some justice too, one may retort on Moore's behalf: "What other principle *seems* as certainly true as the above principle? Has some alternative assumption permitted philosophers to arrive at indisputably true conclusions? And if not, why should Moore not explore the resources of this principle, which seems true to him, just as other philosophers explore the resources of the principles they have accepted, which seem equally true to them?"

See also Being; Bradley, Francis Herbert; Common Sense; Consciousness; Definition Empiricism; Epistemology, History of; Error; Ethics, History of; Existence; Experience; Good, The; Idealism; Kant, Immanuel; Locke, John; Metaphysics; Paradigm-Case Argument; Pessimism and Optimism; Presupposing; Propositions, Judgments, Sentences, and Statements; Realism; Russell, Bertrand Arthur William; Sensa; Universals, A Historical Survey; Ward, James.

Bibliography

WORKS BY MOORE

Except in ethics, Moore's major published philosophical writings consist almost entirely of articles, papers (to be delivered), reviews, and compilations of articles and lectures. Exceptions would be his autobiography (a minor masterpiece in its genre) and "A Reply to My Critics," included in the collection of critical essays concerning Moore's philosophy titled *The Philosophy of G. E. Moore.* In this bibliography, the more important or influential articles of Moore's are noted in the compilations in which they occur.

Principia Ethica. London: Cambridge University Press, 1903.

Ethics. London: Williams and Norgate, 1912.

Philosophical Studies. London: Routledge, 1922. Collection of papers, including "The Refutation of Idealism" (1903), "The Status of Sense-Data" (1914), "The Conception of Reality" (1917), "Some Judgements of Perception" (1918), and "External and Internal Relations" (1919). It is in the last paper that the term *entailment* is first used and defined philosophically.

The Philosophy of G. E. Moore. Edited by P. A. Schilpp. Evanston, IL: Northwestern University Press, 1942; 2nd ed., New York, 1952. Contains "An Autobiography" and "A Reply to My Critics." In the second edition is the "Addendum to My 'Reply.'"

Some Main Problems of Philosophy. London: Allen and Unwin, 1953. Contains Moore's lectures of 1910 and 1911.

Philosophical Papers. London: Allen and Unwin, 1959. With an obituary notice by C. D. Broad. Includes "A Defense of Common Sense" (1923), "Is Existence a Predicate?" (1936), and "Proof of an External World" (1939).

The Commonplace Book, 1919–1953. Edited by Casimir Lewy. New York: Macmillan, 1962.

WORKS ON MOORE

Bridge, Ursula, ed. *W. B. Yeats and T. Sturge Moore, Their Correspondence, 1901–1937.* London, 1953. In this correspondence the two well-known poets (one of them Moore's brother) refer at some length to Moore's philosophy and some of Moore's comments on their interpretations of his philosophy. Although of little philosophical interest, their references provide an amusing picture of nonphilosophers trying desperately to understand Moore. Typical is Yeats's remark "I find your brother extraordinarily obscure."

Keynes, John Maynard. *Two Memoirs.* London: Hart-Davis, 1949. In the second memoir, "My Early Beliefs" (pp. 78–103), Keynes describes the members and discussions of the "Bloomsbury Club," c. 1903–1914. This is a fascinating, witty, and informative account of the tremendous influence Moore's *Principia Ethica* had on some of the finer and younger intellects of the early twentieth century in England; of their attempts, largely verbal, to put Moore's ethical theories into some sort of practice; and of Moore's role in the group, his method of verbal argument, and the "pure and passionate intensity" of his realistic "vision."

Malcolm, Norman. *Knowledge and Certainty.* Englewood Cliffs, NJ: Prentice Hall, 1963. "George Edward Moore," pp. 163ff. In the first part of this important essay Malcolm presents a penetrating and intimate description of Moore's character as a man and as a philosopher, based in large part on personal recollections and impressions. In the remaining parts he discusses the relationship of certain of Moore's "common-sense propositions" to the concept of common sense, to traditional philosophy, and to Wittgenstein's views on the proper role of philosophy with respect to ordinary language. Included is a philosophic evaluation of Moore's purported defense of common sense. This essay is notable not only for the light it sheds on some central aspects of Moore's philosophizing but for the original philosophizing that it

contains on the topics of ordinary language, the concept of common sense, and traditional skepticism concerning perception.

Passmore, John. "Moore and Russell." In *A Hundred Years of Philosophy.* London: Duckworth, 1957. Passmore presents a very searching account of Moore's earlier philosophy (especially as set forth in the 1899 essay "The Nature of Judgment") and his later views on the "analysis of meaning." Passmore also discusses, in an interesting and illuminating way, Moore's theory of sense data and his essays "The Refutation of Idealism" and "Proof of an External World."

Schilpp, P. A., ed. *The Philosophy of G. E. Moore* (see above). Contains critical essays on Moore's ethics by C. D. Broad, Charles L. Stevenson, William K. Frankena, H. J. Paton, Abraham Edel, and A. Campbell Garnett; on his theory of perception by O. K. Bouwsma, C. J. Ducasse, Paul Marhenke, and C. A. Mace; on what might broadly be called his method by Arthur E. Murphy, C. H. Langford, Norman Malcolm, Morris Lazerowitz, Alice Ambrose, John Wisdom, Richard McKeon, and V. J. McGill; and on his influence by L. Susan Stebbing. A number of the essays referred to, such as Bouwsma's "Moore's Theory of Sense-Data," are in their own right important contributions to the topics under discussion. But even when not intrinsically important, these essays constitute a particularly valuable commentary on Moore's philosophy in that Moore, in his "Reply," entered into several detailed discussions of their contents in an attempt to clarify his views. See especially his replies to the essays by Broad, Stevenson, Frankena, Bouwsma, Ducasse, and Langford.

Urmson, J. O. *Philosophical Analysis, Its Development between the Two World Wars.* Oxford: Clarendon Press, 1956. A penetrating work on the "analytic movement" that helps one place Moore in his later philosophical setting. Although only fragmentary references are made to Moore's philosophy, many of the points of view and many of the topics that Moore was concerned with and that influenced his own philosophizing (from 1910 on) are brought into the open and clarified.

Warnock, G. J. "G. E. Moore." In *English Philosophy since 1900.* London, 1958. This thin commentary propagates the thesis that Moore's philosophy can best be understood and appreciated through an understanding and appreciation of his temperament and character. The claim is made that Moore was a "man with no metaphysical quirks of temperament."

White, A. R. *G. E. Moore, a Critical Exposition.* Oxford, 1958. This work—the only English work of book length devoted exclusively to commentary on the philosophy of Moore—collects and collates most of the things Moore had to say on method, theory of ethics, and theory of perception.

John O. Nelson (1967)

MOORE, GEORGE EDWARD [ADDENDUM]

G. E. Moore's ethical writings, especially *Principia Ethica* of 1903, have long been regarded as philosophically revolutionary. In fact, Moore shared his main ethical views—nonnaturalism in metaethics and ideal consequentialism in normative ethics—with such late-nineteenth-century writers as Henry Sidgwick and Hastings Rashdall. But Moore defended these views with unusual vigor and so had a disproportionate influence on later moral philosophy.

Moore's nonnaturalism comprised two main theses. One was the realist thesis that moral judgments are objectively true or false; the other was the autonomy-of-ethics thesis that moral judgments are sui generis, neither reducible to nor derivable from nonmoral judgments such as scientific or metaphysical ones. Our knowledge of them must therefore derive from intuitive judgments of self-evidence.

Moore did not argue extensively for realism. Like others of his era, he took it largely for granted. But his argument for the autonomy of ethics has come to be known as the "open-question argument." If goodness were identical to pleasure, the claim that pleasure is good would be equivalent to the empty statement that pleasure is pleasure, which it plainly is not. Rather, whether pleasure is good is always an open question. Since this argument generalizes to all nonmoral properties, goodness cannot be identical to any such property. Some later philosophers challenged this argument against the "naturalistic fallacy"; others took it to support antirealist conclusions quite different from Moore's. But it remains a central argument for the irreducibility of moral claims.

Though these main theses were familiar, Moore did introduce two innovations. One was his view that the central irreducible moral property is *good* rather than *ought* or *right*; the other was that the intrinsic goodness can depend only on its intrinsic properties, apart from any relations to other states. It follows that to judge whether a given state is good, we must imagine a world containing only that state and ask whether such a world is good.

Moore's ideal consequentialism likewise comprised two theses. One was that right acts always produce the most good. The other was that there is a plurality of goods, all ideal in the sense that their being good does not depend on people's attitudes to them. (Moore thought that the naturalistic fallacy led philosophers to identify goodness with some one natural property and so to miss this plurality.) In *Principia Ethica* he held that one intrinsic good is beauty apart from any consciousness of it; another is vicious people's deserved pain. But the chief goods in this work were the admiring contemplation of beauty and personal love, which for Moore involved the

admiring contemplation of others' good qualities. In characterizing both goods, he used his "principle of organic unities," according to which the value of a whole need not equal the sum of the values of its parts. Beauty on its own has little value, and the contemplation of merely imagined beauty just moderate value, but the contemplation of real beauty has great value, more than the sum of the values of those components.

Principia Ethica was written with a self-confidence bordering on arrogance. Moore thought most previous moral philosophers had made crude conceptual errors, and that once those were exposed, the moral truth would be self-evident to all. This tone helped make his presentation of nonnaturalism the canonical one. As a result, twentieth-century metaethics can be seen as a sequence of reactions to his views. His substantive views about the good received less attention, but at the beginning of the twenty-first century, interest has revived in his claims about, for example, appropriate attitudes and organic unities. His moral philosophy is again alive as a whole.

See also Ethics, History of; Good, The; Intrinsic Value; Intuitionism, Ethical.

Bibliography

WORKS BY MOORE

Principia Ethica. Cambridge, U.K.: Cambridge University Press, 1903.

"The Conception of Intrinsic Value." In his *Philosophical Studies*. London: Routledge and Kegan Paul, 1912.

The Elements of Ethics, edited by Tom Regan. Philadelphia: Temple University Press, 1991.

WORKS ON MOORE

Baldwin, Thomas. *G. E. Moore*. London: Routledge, 1990. See especially chaps. 3–4.

Frankena, William K. "The Naturalistic Fallacy." *Mind* 48 (1939): 464–477.

Hurka, Thomas. "Moore in the Middle." *Ethics* 113 (2003): 599–628.

Hutchinson, Brian. *G. E. Moore's Ethical Theory: Resistance and Reconciliation*. Cambridge, U.K.: Cambridge University Press, 2001.

Regan, Donald H. "How to Be a Moorean." *Ethics* 113 (2003): 651–677.

Regan, Tom. *Bloomsbury's Prophet: G. E. Moore and the Development of His Moral Philosophy*. Philadelphia: Temple University Press, 1986.

Shaver, Robert. "*Principia* Then and Now." *Utilitas* 15 (2003): 261–278.

Thomas Hurka (2005)

MORAL ARGUMENTS FOR THE EXISTENCE OF GOD

From the time of Immanuel Kant to the present day, a great many attempts have been made to base arguments for God's existence not upon the mere fact that there is a world, nor on the general orderliness it manifests, but on a very special feature of that world—human moral experience. The popularity of moral arguments is not hard to understand. David Hume and Kant had produced powerful and apparently disabling criticisms of the traditional arguments of natural theology, criticisms that seemed decisive against any conceivable type of argument to God as the explanation of the world. Hume had no alternative theistic argument to offer and, insofar as theoretical reasoning is concerned, Kant had none either. The structure of Kant's ethical philosophy, however, accorded to "practical reason" privileges not shared by theoretical reason. If God was to retain any place in the Kantian system, the weight of apologetic had to be shifted from the theoretical to the practical, to exploring the implications of our moral situation. Between Kant's day and the middle of the twentieth century, skepticism about the theoretical arguments tended to deepen rather than to lighten; hence, there has been no lack of religious apologists following Kant's new "moral route" to God.

Another reason for the popularity of moral argument is religious rather than philosophical. Even if the argument to God as First Cause or "necessary being" were valid, these notions of deity can be more of an embarrassment than a help to the religious imagination. They present us with a divine object or superobject, whereas religion demands that God be primarily known as *person*. A moral argument offers hope of overcoming that external and thinglike character: It ensures that concepts of God will be, from the outset, personal concepts.

TYPICAL MORAL ARGUMENTS

Among the many varieties of moral argument, the following are both historically important and recurrent patterns. Several of them may be found in a single author.

First, if one understands moral rules as "commands," one may argue to the existence of a "commander." The commander cannot be the individual human moral agent, for what today I command myself to do, I can tomorrow command myself not to do. I can have absolute moral obligations only if a God exists to command them. Because I do have absolute moral obligations, it follows that God exists.

Second, a minor variant of this moral argument claims that if we recognize moral authority, we must ipso facto recognize the existence of God as alone able to confer that authority. We judge that the moral law retains its authoritativeness whether particular human wills are at any time actually accepting its rules and principles or not; therefore, the source of its authority must lie altogether outside those human wills.

Third, the notion of "moral law" itself is said to be incomplete without reference to God, for law implies "lawgiver," a divine legislator. Our very acknowledgment of a moral law, therefore, presupposes theism.

Fourth, it has been claimed that there is a remarkable degree of agreement among the moral judgments made by men in widely different cultures and historical periods. Many apparent disagreements can be attributed to differences in belief and thus held to be not fundamental. This impressive measure of agreement, it is argued, can be accounted for only on the supposition that God has written his law in the hearts of men.

Certain of the most interesting and influential moral arguments take as their premise some part of the content of the moral law itself. We are under moral obligation to perfect ourselves and to attain a "highest good" (summum bonum) that is manifestly unattainable in a life lived under the conditions we know here and now. We can, at best, make a start to a moral development that requires very different conditions for its completion. But since that complete development is demanded of us as duty, it must be attainable. God and immortality are thus presupposed in our actual moral experience.

ANALYSIS OF MORAL ARGUMENTS

Let us briefly consider each of the varieties of moral arguments again and attempt to estimate their strengths and weaknesses.

Of the moral commander argument we can pertinently ask: Is the notion of command basic to ethics? Certainly not in the sense of parade-ground commands, commands passively received and acted upon unreflectively. Such obedience is a long way from moral deliberation and judgment. An immature moral agent may see his duties as commands (parental, for example); but the mark of mature moral judgment is self-commitment to a policy on which one has deliberated. This policy may or may not be in harmony with someone's command; in any case, it does not owe its authenticity to its being commanded. "Here I stand," one may say; and this can express a settled resolution, one not to be made one day and rescinded the next.

Even if it were established that a celestial being unvaryingly commanded a certain policy as obligatory in an absolute sense, the unvaryingness of his command could not itself furnish the ground for the absoluteness of the obligation. For it is at least logically possible that this celestial being ought not to command unvaryingly what he does so command. If he commands what is right and obligatory, that is cause for thankfulness; but one could scarcely be thankful over a truth of logic. "Unvaryingly" must not equal "stubbornly" or "with chronic moral blindness"; these are unthinkable possibilities for Christian theism. But this does not affect the point being made: that absoluteness is not analyzable in terms of unvaryingness of command. Moreover, the Christian wishes to make one all-important moral judgment that could not possibly have its absoluteness reduced to commandedness by God—the judgment, namely, that God is morally perfect. But if a human being can make this moral judgment uncommanded, why can he not make others also?

Analogous criticisms can be made of the argument from the authority or authoritativeness of the moral law to the need for a divine source of authority. To put the main objection boldly: It is of the very nature of a fundamental moral judgment that it should be made on no authority but that of the agent who makes it. Certainly there are occasions when I may believe that another person has a superior measure of insight into the situation in which I have to act; I may then properly accept his judgment in lieu of my own. Yet if this is not to be a culpable moral abdication, I must have good grounds for trusting my temporary "authority": I must judge him to be morally reliable. But this is itself a moral judgment—one that I can make on nobody's authority but my own; or if on someone's authority, then this new person must be judged reliable on my own authority, and so on. A legitimate appeal to authority presupposes that autonomous moral judgments have already been made. Our argument held that we must postulate God as the authorizer of all our moral judgments—otherwise they would carry no authority; but we find, contrariwise, that God can play the role of authority only if we are able to make certain moral judgments without appealing to any external authority whatsoever.

The third version alleged that the notion of "moral law" is incomplete unless God is postulated as lawgiver. *Law,* however, is a word with many strands to its meaning; and it is only by failing to distinguish certain of the

strands that this can appear to be a plausible line of argument. It is perfectly intelligible to say that some person or group of persons has laid down positive laws, rules for a community, backed by penal sanctions. The existence of a developed body of such laws normally implies the existence of lawmakers or codifiers. It is quite another thing (and not really intelligible) to speak of anyone, human or divine, "laying down" the moral law itself. Laws, rules, and regulations are of the right logical type to be laid down in accordance with, or in conflict with, the moral law. But the moral law itself is not the sort of thing that needs to be, or that logically can be, laid down or promulgated by anyone. No conceivable story about men or gods could be taken, without absurdity, to describe the inauguration (or the annulling) of the moral law. Commands might be uttered, inscriptions miraculously appear; but it would never become a trivial or tautological question to ask of their content, "Is this in fact morally binding?" The distinctively moral authority of a rule or law does not lie in the prestige or power of its initiator, nor in the circumstances of its first recognition.

The argument from the convergence of moral codes is most often set forth in an objectivist ethical context. The existence of objective moral qualities "seen" to be there, or "intuited," by different moral agents in widely different places and times remains inexplicable unless we posit a God who creates and morally guides. It is less often noticed that the argument is perhaps stronger—certainly no weaker—if it is set forth in a subjectivist context instead. This was apparently noted by F. R. Tennant, who (in a conversation reported by R. B. Braithwaite) argued on the following lines. Failing the existence of any objective moral properties or moral relations, it is all the more remarkable that there should be such a measure of congruity among moral judgments or decisions: sufficiently remarkable to point the way, again, to divine activity. Yet this argument is not at all conclusive. The supernatural hypothesis that it puts forward is not the only hypothesis available to account for the data; and it has the disadvantage that it is not empirically confirmable or refutable. Powerful competitors would be arguments from the relative stability of basic human needs, desires, and aversions or from the pervasiveness of aggressive and social drives in the personality. These alone might well account for the actual agreements among moral judgments and would account for them without invoking the immensely problematic notion of divine causality.

PRESUPPOSITION OF THE HIGHEST GOOD

Our last group of arguments began its history in modern philosophy with a statement of Kant: "The idea of the highest good … cannot be realized by man himself … ; yet he discovers within himself the duty to work for this end. Hence he finds himself impelled to believe in the cooperation or management of a moral Ruler of the world, by means of which alone this goal can be reached (*Religion within the Limits of Reason Alone*).

Kant was not betraying the austerity or rigor of his moral philosophy; he was not offering religious inducements to moral behavior. He would have denied distinctively moral worth to someone whose "dutiful" actions were aimed at securing his own postmortem happiness. The emphasis in his argument is wholly on the intelligibility and rationality of the moral demand; it was inconceivable to him that the categorical imperative should be a mocking voice, laying obligations upon us and at the same time denying the environment in which alone the obligations could be fulfilled. (It has been claimed that Kant had abandoned these moral arguments by the time he wrote the *Opus Postumum,* but the contrary view has been argued more forcibly; see G. A. Schrader, "Kant's Presumed Repudiation … .")

The strength of Kant's moral argument is clearly dependent on the strength of his ethical theory as a whole. It is only because he saw moral judgment as the work of practical reason (not as a matter of emotive reactions or responses) that he was able to make plausible use of those judgments as a basis for theological demands. Any fundamental criticism of the Kantian ethic would ipso facto imperil the theology.

The argument is equally imperiled if we deny that we are under obligation to attain the highest good and our individual moral perfection, saying that we are obliged only to strive toward these unrealizable ends. We might indeed reverse Kant's argument as follows. From our observation of the world we conclude that the highest good and our moral perfection are unattainable; therefore, we can have no obligation to attain these but, at best, only an obligation to strive toward them. We can interpret them in Kant's own term, *regulatively,* as Kant himself sometimes did. (See John R. Silber, "Kant's Conception of the Highest Good … .")

The postulating of God and immortality is aimed at solving an antinomy—of making intelligible what, without the postulate, is inexplicable. But does the postulation of God in fact produce intelligibility, a lifting of mystery?

Or is there not so much mystery in the postulate itself that the final effect is a deepening, not a lightening, of perplexity? If independence, autonomy, and freedom are essential to a moral agent, that autonomy will presumably remain essential in a hereafter as well as in the here and now. But, if so, the postulation of God and immortality can by no means ensure that the ultimate moral goals will, in fact, be reached, even though it was precisely to ensure their attainment that the postulates were made.

Kant's theory of time as a "form of sensibility" makes it very dubious whether he could have spoken meaningfully of a continuing moral development and the attainment of the highest good in a hereafter. Granted that he disclaimed all theoretical insight into what such an existence would be like (this measure of agnosticism is part of the force of *postulate* as distinct from *demonstrate*), the notion of time still remains essential to Kant's moral argument. If we are unable to give meaning to it in that context, the argument cannot but suffer.

It is possible to reject some portions of Kant's detailed argumentation and yet to advance a moral argument of a definitely Kantian type. This was notably done in W. R. Sorley's *Moral Values and the Idea of God* and in A. E. Taylor's *The Faith of a Moralist.* Neither of these writers held the moral argument to be the sole and all-sufficient theistic proof, but they did believe that without it the case for theism is weak and dubious.

Sorley attempted first to show that "the moral order is an objectively valid order, that moral values belong to the nature of reality," and that "the history of the world-process is fitted to realise this order." If we were to assume that the goal of the world-process is the realizing of happiness, there would be the weightiest empirical evidence against us. With moral worth and goodness it is different. Conditions that work against happiness may work for, not against, the developing, trying, and testing of moral fiber. "The very imperfection of the world [is] an argument pointing to the theistic conclusion." There remains yet a gap between the claim that the universe works toward a moral purpose and the full claim that God exists: Sorley seeks to fill this gap by arguing that belief in God is presupposed by belief in an objective and "eternally valid" morality. If the moral law is eternally valid, and valid whether we recognize it or not, "how could this eternal validity stand alone, not embodied in matter and neither seen nor realized by finite minds, unless there were an eternal mind whose thought and will were therein expressed?"

One can readily agree that the world as we experience it is better adapted to be a vale of soul-making than a hedonistic holiday camp. Yet there are difficulties about even the soul-making view. Some human suffering (the unmerited suffering of young children, for instance) cannot always be treated plausibly as developing moral fiber, or as realizing any other moral value. The natural environment can figure as the destroyer of moral personality as well as its preserver and nourisher. Sorley's further argument, from the validity of the law to an eternal mind, surely contains a confusion of the logical and the psychological. Questions about validity and about truth are logically independent of questions about the propositions that are actually entertained in someone's mind. Whether or not there exists a person who says (or thinks) *p*, has no bearing on the truth of *p* or, if *p* is a moral principle, upon its bindingness or validity in the relevant sense.

A. E. Taylor saw the moral life not as a mere conforming to given static principles and rules but as directing the moral agent along certain paths of self-development. There is development within the moral ideal: "We discover tomorrow that today's ideal 'had more in it' than we had supposed." The goal transforms itself as we approach it. The further we pursue it, the less able we become to conceive the human good in purely this-world, secular terms. There is development also within our awareness of time. Purposeful, valuable activity produces an extension of our "conscious present"; it delivers us from the dullness of "one thing after another." The limiting case in this development would be well expressed by Boethius's account of eternity—"the complete and simultaneous fruition of a life without bounds."

"We may argue," Taylor then claimed, "from the existence of a function to the reality of an environment in which the function can find adequate exercise." But no view of the world, short of theism, can guarantee the completion of these directions of development that Taylor has described.

Whatever is decided about the validity of the argument as an argument, Taylor's *The Faith of a Moralist* is a lastingly impressive and eloquent account of a religiously oriented morality. On validity, however, some searching criticisms were made by C. D. Broad in his review of Taylor's work published in *Mind* (1931). Taylor had taken as his premises certain moral judgments and certain trends of development in our experience of value. He then had asked what these entailed; whatever they entailed was to be added to our true beliefs about the universe. Broad argued that, in order to avoid a vicious circle, we must be sure that our premises do not already covertly assume the theistic conclusion. We must know that we have these duties and aspirations without already presupposing God

and immortality. Only in this way could the existence of God and immortality be the conclusion of our argument. It is hard to be sure that these value judgments and aspirations are not the *consequence* of a prior theism. And a further point must be added: Only such a previously held theism, or cryptotheism, could entitle us to argue, with Taylor, "from the existence of a function to the reality of an environment in which [it has] adequate exercise." (Or, if this is true by definition of *function,* only such a theism can justify calling those value pursuits "functions.") Once again it might be added that the directions of moral development, although unrealizable in toto, could still be taken as targets for ever-nearer approximation. That they can be taken in this way, however, tells against Taylor's argument, for he wished to deny that we can be morally serious about these unless complete realization is possible.

Moral theories dominant in the mid-twentieth century did not tend to lead naturally into moral arguments for God. In Britain and the United States, at any rate, they were characteristically this-worldly. But exceptions do occur. Austin Farrer offered, if not a moral argument, then certainly a moral "persuasion" toward theism in the first chapter of *Faith and Logic,* "A Starting-Point for the Philosophical Examination of Theological Belief." His argument is that we are incontestably under an obligation to love our neighbor—that is, to hold him in highest regard; and that this is not impossible if our neighbor is a lovable person. If our neighbor occasionally lapses from lovability and from goodness, we may still manage to love his "normal" self, although it is temporarily obscured. If, however, he lapses chronically and grossly, how are we to love him? To love what he might be is now to love a fiction only; but it is persons, not fictions, that we ought to love. Farrer claimed Christianity provides a uniquely helpful way in which we can see the unlovable neighbor, admit his deficiencies, and yet succeed in loving him. In praying for and about our neighbor, we bring our view of him into relation with God's action—his action in creating our neighbor and his constant and costly redemptive action on our neighbor's behalf. Farrer insisted that, if these reflections help to give plausibility and impressiveness to the Christian view itself, they are not to be taken as a refurbishing of strong Kantian claims to establish God's existence.

Farrer appears to have assessed the capacity of this type of argument far more realistically than those who used it before him. If we judge that certain attitudes or evaluations are supremely worth realizing—for example, that "people ought to be held in the utmost regard"—then it is reasonable, even mandatory, to take up whatever stance will best further our task of realizing them. In our present example, we are required to meditate upon those reflections that uniquely put our neighbor in a regard-furthering light. Of course, provisos must be added. There must, for instance, be no logical incoherence in the description of the stance or of the context that furthers our neighborly love; otherwise, what we called the light or the stance might be in fact only a fugitive, quasi-aesthetic movement of feeling. To provide a point of entry to traditional Christianity, the stance must be capable of being expressed in a set of meaningful affirmations about reality. Another obvious proviso is that our premises must be sound. We must in fact be under obligation to hold our neighbor in highest regard, and all non-Christian ways of seeing our neighbor must be less helpful than the Christian way. It is particularly upon the second of these premises that, in a fuller discussion, argument necessarily would concentrate.

See also Broad, Charlie Dunbar; Ethics, History of; Hume, David; Immortality; Kant, Immanuel; Popular Arguments for the Existence of God; Presupposition; Sorley, William Ritchie; Taylor, Alfred Edward.

Bibliography

The chief sources for moral arguments to God are the following works of Kant: *Kritik der praktischen Vernunft* (Riga: J. F. Hartknoch, 1788), translated by T. K. Abbott (London and Dublin, 1879) and by L. W. Beck (Chicago, 1949), and *Die Religion innerhalb der Grenzen der blossen Vernunft* (Königsberg: Bey Friedrich Nicolovius, 1794), translated by T. M. Greene and H. H. Hudson as *Religion within the Limits of Reason Alone,* 2nd ed. (New York: Harper, 1960).

Post-Kantian works in which moral arguments to God play an important part are Austin Farrer, "A Starting-Point for the Philosophical Examination of Theological Belief," in *Faith and Logic* edited by B. G. Mitchell (New York, 1957); John Henry Newman, *A Grammar of Assent* (London, 1901), especially pp. 109–110; Hastings Rashdall, *The Theory of Good and Evil,* 2 vols. (Oxford: Clarendon, 1907), and *God and Man* (Oxford, 1930); W. R. Sorley, *Moral Values and the Idea of God* (Cambridge, U.K.: Cambridge University Press, 1918); and A. E. Taylor, *The Faith of a Moralist* (London: Macmillan, 1930), reviewed by C. D. Broad in *Mind* 40 (1931): 364–375.

Among contemporary studies of the Kantian argument are the following: H. P. Owen, *The Moral Argument for Christian Theism* (London: Allen and Unwin, 1965); G. A. Schrader, "Kant's Presumed Repudiation of the 'Moral Argument' in the Opus Postumum," in *Philosophy* 26 (1951): 228–241; John R. Silber, "Kant's Conception of the Highest Good as Immanent and Transcendent," in *Philosophical Review* 68 (1959): 469ff.; and W. H. Walsh, "Kant's Moral Theology," in *Proceedings of the British Academy* 49 (1963): 263–289.

Ronald W. Hepburn (1967)

MORAL ARGUMENTS FOR THE EXISTENCE OF GOD [ADDENDUM]

The moral argument purports to show that evidence from our moral experience supports the existence of God. From the 1970s onward, various forms of the moral argument have been developed by many philosophers. While defenders argue with increased sophistication, they also tend to make more modest claims about the force of the moral argument.

MORAL ARGUMENTS AS ABDUCTIVE AND CUMULATIVE ARGUMENTS

If the moral argument is construed as a deductive argument that moves from, say, the objectivity of moral values to the existence of God, then to rebut the argument, the critic needs to show only that the objectivity of moral values and the nonexistence of God are logically compatible. This is a relatively easy task. However, developments in epistemology and philosophy of science since the 1960s lead many to think that it is more realistic to look, in most areas of inquiry, for an abductive argument, an inference to the best explanation. We can formulate the moral argument too as an abductive argument, that is, argue that among diverse worldviews, the theistic worldview is the best explanation of, say, the objectivity of morals, especially in contrast with naturalism.

Since abductive arguments are by nature cumulative arguments, the force of an abductive moral argument will depend not on any single feature of morality, but rather on how well it can explain the whole gamut of moral experience, both its form and its content. For example, Louis Pojman argues, "Given the assumption of standard contemporary secular moral philosophy: I. The notion of *moral obligation* becomes seriously problematic; II. The notion of the *supremacy of morality* either becomes problematically analytic or it vanishes; III. The problem of morality and self-interest becomes insoluble; IV. The idea that human beings have intrinsic value ceases to make sense." Hence, "most contemporary secular ethical systems offer no hope of guiding human conduct, and should be abandoned" (1992b, p. 4).

In contrast, ethical systems that proceed from transcendent assumptions can offer resources unavailable to secular ethical systems. For example, the Christian tradition can appeal to a perfectly good, omnipotent God who created humanity in his image. Each person is endowed with a specific telos, which the individual must seek to realize. Within this framework, all humans have equal intrinsic worth, free will, and eternal destiny (see also Pojman 1991, 1992a). Morality consists of obeying God's commands, which promote human flourishing and are backed by rewards and punishments. All these provide a solid foundation for the existence of moral obligation and responsibility.

ARE MORAL TRUTHS ANALYTIC?

Not all theistic philosophers accept the moral argument. For example, Richard Swinburne (1974) believes that fundamental moral truths are necessary truths and do not need to be explained. Defenders have several ways to respond. First, they may flatly deny Swinburne's claim by pointing out that moral nihilism and moral relativism at least appear to be logically coherent positions. Second, even if many moral principles are necessary truths, it does not follow that they cannot be explained by more basic necessary truths about God's essential moral nature and logically necessary existence. Charles Taliaferro even suggests a cosmological-ethical argument that utilizes "the resources of a theistic metaphysic in providing a singular, comprehensive explanatory account of moral truths as well as other essential truths," in addition to explaining the existence of the cosmos. In this way, theism may exhibit "a marked simplicity and force missing from its competitors" (1996, p. 290).

Third, the realm of necessary moral truths appears to be mysterious and odd from a naturalistic perspective. It is puzzling why we should be aware of these truths and why moral consciousness features so prominently in human existence. Necessary moral truths by themselves do not have any power to endow agents with morality, and a naturalistic universe cannot have any causal interaction with these abstract truths. Why, then, should we suppose that a morally blind world would endow us with correct moral intuitions? The case is different with sense experience, whose reliability is to some extent tied with our survival. It is not without reason that most naturalists prefer moral skepticism. As John Mackie admits, "There can be a secular morality, not indeed as a system of objective values or prescriptions, but rather as something to be made [that is, invented]" (1977, p. 227; see also Harman 1984).

THE ODDITY OF MORAL OBLIGATION

For Mackie, another reason why morality is so odd is that moral claims are authoritative and objectively prescriptive: "Any wrong (possible) course of action would have not-to-be-doneness somehow built into it" (1977, p. 40).

George Mavrodes (1986) points out that naturalistic evolution may well produce creatures with moral sentiments conducive to survival. However, the existence of *actual* moral obligations appears to be strange, especially because moral obligations often come into conflict with self-interest.

Some atheists (David Gauthier [1986], Gregory S. Kavka [1984]) try to reconcile moral obligations with self-interest, and claim that in the long run it is in the best interests of everyone for every individual to act morally. The viability of this kind of contractual project depends on whether it can satisfactorily answer questions like these: Should one still be moral when in fact not everyone else will act morally? What about the moral free-rider? Do extremely powerful people really need to act morally? Why should we sacrifice our own interests for the benefits of people who cannot reciprocate, such as future generations, extremely marginalized people in one's own society, and people in distant countries? Is it rational to sacrifice one's life for the sake of morality?

THE MORAL GAP

A broadly Kantian moral argument continues to find defenders. Ronald Green (1978) starts from the question "Why should I be moral?" John Hare (1996) focuses on the gap between the demand that one act morally and our capacities to meet this demand, according to most moral theories. Since "ought" implies "can," the description of the moral life in this moral gap is incoherent. To resolve this incoherence, secular moralists either exaggerate our moral capacity, reduce the moral demand, or try to find some substitute for God to help bridge the gap. Hare criticizes many of these options and argues that the Christian doctrines of atonement and incorporation in Christ can solve the problem. Debates surround whether Hare's criticisms of the secular options are cogent and whether the Christian faith can really offer something that other options cannot (see Zagzebski 1999).

THE *EUTHYPHRO* DILEMMA

Atheistic philosophers such as Kai Nielsen (1990) and Michael Martin (2002) have produced sustained replies to the moral argument. They think that the *Euthyphro* dilemma (are morally good acts commanded by God because they are morally good, or are they morally good because they are commanded by God?) shows that morality has to be independent of God. If morality depends on God's command, then morality will be arbitrary, because God might command cruelty for its own sake. If one denies this possibility, one already commits to an independent standard of the good apart from God.

Some theists reply by saying that God's essential nature, from which the divine will flows, provides the ultimate standard of goodness, and this is neither independent of God nor arbitrary. Robert M. Adams has proposed a modified divine-command ethics that, as a postulate, equates being contrary to the commands of a *loving* God with the *essence* of being wrong (Adams 1987). This is not an analysis of the *meaning* of "right" and "wrong," because Adams grants that our moral practice gives us some basic understanding of morality apart from religion. However, it does not follow, he thinks, that, on the basis of this basic understanding, we can understand the *essence* of being wrong. (Someone who understands the meaning of "water," can further discover that the nature, or essence, of water is H_2O.) On Adams's view, the answer to the *Euthyphro* dilemma is that a *loving* God would not command cruelty for its own sake.

A CRITICAL DIALOGUE BETWEEN ETHICAL SYSTEMS

The success of the moral argument in the long run depends on the relative merits of the theistic and atheistic accounts of morality. (We should also include, say, Confucian ethics and Buddhist ethics among the contenders.) Adams (1999) has developed his theistic ethics into a comprehensive theory of the good and the right. Michael Moore (1996) and Michael Martin (2002) have used naturalistic moral realism (Brink 1989) to show that naturalistic ethics is superior and that theistic ethics is superfluous (see also Copan 2003, 2004). The moral argument does not appear to be conclusive. Its significance mainly lies in its contribution to a cumulative case for God's existence and its capacity to stimulate a lively debate on the implications of different worldviews on morality.

See also God, Concepts of; Mackie, John Leslie; Moral Skepticism.

Bibliography

Adams, Robert M. *Finite and Infinite Goods*. Oxford, U.K.: Oxford University Press, 1999.

Adams, Robert M. "*The Virtue of Faith*" *and Other Essays in Philosophical Theology*. Oxford, U.K.: Oxford University Press, 1987.

Brink, David O. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Copan, Paul. "The Moral Argument." In *The Rationality of Theism*, edited by Paul Copan and Paul K. Moser. London: Routledge, 2003.
Copan, Paul. "Morality and Meaning without God, Another Failed Attempt: A Review Essay on *Atheism, Morality, and Meaning*." *Philosophia Christi* 6 (2) (2004): 295–304.
Gauthier, David. *Morals by Agreement*. New York: Oxford University Press, 1986.
Green, Ronald. *Religious Reason: The Rational and Moral Basis of Religious Belief*. New York: Oxford University Press, 1978.
Hare, John. *The Moral Gap*. Oxford, U.K.: Oxford University Press, 1996.
Harman, Gilbert. "Is There a Single True Morality?" In *Morality, Reason, and Truth: New Essays on the Foundations of Ethics*, edited by David Copp and David Zimmerman. Totowa, NJ: Rowman and Allanheld, 1984.
Kavka, Gregory S. "The Reconciliation Project." In *Morality, Reason, and Truth: New Essays on the Foundations of Ethics*, edited by David Copp and David Zimmerman. Totowa, NJ: Rowman and Allanheld, 1984.
Mackie, John. *Ethics: Inventing Right and Wrong*. Harmondsworth, U.K.: Penguin Books, 1977.
Martin, Michael. *Atheism, Morality, and Meaning*. Amherst, NY: Prometheus, 2002.
Mavrodes, George I. "Religion and the Queerness of Morality." In *Rationality, Religious Belief, and Moral Commitment*, edited by Robert Audi and William J. Wainwright. Ithaca, NY: Cornell University Press, 1986.
Moore, Michael S. "Good without God." In *Natural Law, Liberalism, and Morality: Contemporary Essays*, edited by Robert P. George. Oxford, U.K.: Clarendon Press, 1996.
Nielsen, Kai. *Ethics without God*. Rev. ed. Amherst, NY: Prometheus, 1990.
Pojman, Louis P. "Are Human Rights Based on Equal Worth?" *Philosophy and Phenomenological Research* 52 (1992a): 605–622.
Pojman, Louis P. "A Critique of Contemporary Egalitarianism: A Christian Perspective." *Faith and Philosophy* 8 (1991): 481–504.
Pojman, Louis P. "Ethics: Religious and Secular." *Modern Schoolman* 70 (1992b): 1–30.
Swinburne, Richard. "Duty and the Will of God." *Canadian Journal of Philosophy* 4 (1974): 213–227.
Taliaferro, Charles. "God's Natural Laws." In *Natural Law, Liberalism, and Morality: Contemporary Essays*, edited by Robert P. George. Oxford, U.K.: Clarendon Press, 1996.
Zagzebski, Linda. "Review of *The Moral Gap*." *Philosophical Review* 108 (1999): 291–293.

Kai-man Kwan (2005)

MORAL DILEMMAS

The label *moral dilemma* is commonly applied to any difficult moral problem. Several introductory anthologies in ethics have been titled *Moral Dilemmas*, suggesting that all of the issues discussed therein are moral dilemmas, regardless of their structure, simply because they raise hard moral questions. Many people even talk about moral dilemmas when it is not clear whether or not morality is relevant at all.

Moral philosophers, in contrast, usually have in mind something more specific. Minimally, they count a situation as a moral dilemma only if one moral reason conflicts with another (moral or nonmoral) reason. Reasons conflict in a situation if the agent is not able in that situation to comply with all of the reasons. For example, if it is in Ann's interest to lie to a potential employer, then Ann's prudential reason to lie conflicts with Ann's moral reason not to lie. Similarly, moral reasons can conflict with religious reasons (as on one interpretation of the biblical story of Abraham being commanded by God to sacrifice his son, Isaac) or with aesthetic reasons (as on one understanding of Gauguin's decision to leave his family to pursue his art).

Moral philosophers normally restrict the class of moral dilemmas further to include only conflicts between one moral reason and another reason that is also moral in nature. In Plato's example, if Brad holds a weapon for a friend and promises to return it when that friend asks for it, then Brad has a moral reason to return it when the friend asks. But if Brad knows that this friend is going to use the weapon to commit a harmful crime, then Brad has a moral reason not to return the weapon to the friend (at least at that time).

Many philosophers would not classify this conflict as a moral dilemma because it is resolvable—the moral reasons against returning the weapon override the moral reasons in favor of returning the weapon, so overall Brad morally ought not to return the weapon, assuming that the harmful crime is serious enough. In contrast, even if moral dilemmas must be unresolvable, Carol is in a moral dilemma on this account if Carol has a moral reason to help the needy but can help only one of two equally needy people.

Some philosophers limit moral dilemmas even further to include only conflicts among certain kinds of moral reasons. A moral reason is a moral requirement just in case it would be morally wrong not to act on it without an adequate justification or excuse. Carol's moral reason to help a particular needy person, for example, is not a moral requirement if it would not be morally wrong for Carol to refuse to help this needy person (as long as Carol helps enough other needy people at other times). Then, if moral dilemmas are limited to unresolvable conflicts between moral requirements, Carol is not in a moral dilemma when she can help only one of two equally needy people. In contrast, if David can keep only one of

two conflicting promises, assuming that David has a moral requirement to keep his promises, then David is in a moral dilemma, even if moral dilemmas are defined as unresolvable conflicts of moral requirements.

Other moral theorists define moral dilemmas in different terms, for instance, as situations where every alternative is morally wrong. The term *wrong*, however, is unclear in this context. If an act is called morally wrong when, and only when, it violates a non-overridden moral requirement, then this definition reduces to the previous one. In contrast, if an act is called morally wrong only when it violates an overriding moral requirement, then this definition makes moral dilemmas obviously impossible. That obviousness suggests that philosophers who claim that moral dilemmas are possible do not use this strong definition of moral dilemmas. Instead, they seem to identify moral dilemmas with unresolvable moral requirement conflicts.

To show that a situation fits that definition, it is not enough to cite nonmoral facts, such as that the agent cannot do both acts or even that each act is necessary to fulfill a promise. The situation is not a moral dilemma unless there are moral requirements for conflicting alternatives and neither moral requirement overrides the other. In support of the claim that there is a real moral requirement on each side, philosophers who see the situation as a moral dilemma cite the counterfactual that it would be morally wrong not to choose a particular alternative if there were no moral reason to choose the conflicting alternative. They also often argue that moral requirements on each side provide the best explanation of why remorse (or guilt, but not just regret), an apology, compensation, or some other moral residue is appropriate after either choice.

In support of the claim that neither moral requirement overrides the other, philosophers who assert the possibility of moral dilemmas can argue that some situations are so symmetrical that neither moral requirement could override the other. A common symmetrical example is Sophie's choice between her two children when a Nazi guard threatens to kill her and both of her children if she does not pick one child to be killed. In nonsymmetrical cases, some philosophers also argue that conflicting moral requirements can be incomparable, in which case neither moral requirement overrides the other (although they are also not exactly equal).

Opponents who deny the possibility of (even resolvable) conflicts between moral requirements sometimes object that if one conflicting moral requirement overrides the other, then the other is no longer a moral requirement. This objection conflates overriding with cancellation. Like physical forces, moral requirements that are overridden by stronger moral requirements can still retain some moral force, as shown by their ability to justify remorse, apologies, compensation, and other forms of moral residue.

Another common objection to the possibility of moral dilemmas charges that, if neither moral requirement overrides the other, then the agent is morally permitted to choose either alternative and, hence, is not in a moral dilemma. However, if an act is not morally permitted only when it violates an overriding moral requirement, then the claim that both acts are morally permitted is compatible with the situation being a conflict between non-overridden moral requirements and, hence, a moral dilemma on the above definition. In contrast, if an act is not morally permitted when it violates a non-overridden moral requirement, then neither act is morally permitted in an irresolvable moral requirement conflict. Either way, the notion of permission does not rule out moral dilemmas.

Additional arguments against the possibility of moral dilemmas try to derive a contradiction from the definition of moral dilemmas. If the agent in a moral dilemma morally ought to adopt each alternative separately, then the agent morally ought to adopt both alternatives together, according to the agglomeration principle. If the agent morally ought to adopt both alternatives, then the agent must be able to adopt both alternatives, according to the principle that *ought* implies *can*. The agent cannot adopt both alternatives in a moral dilemma, by definition. Thus, the definition of moral dilemmas plus agglomeration and *ought* implies *can* imply a contradiction. Defenders of moral dilemmas respond by denying either agglomeration or *ought* implies *can*, or both.

Another formal argument applies a closure principle: An agent has a moral requirement not to do whatever prevents that agent from fulfilling a moral requirement. This closure principle implies that an agent in a moral dilemma has a moral requirement to adopt and also not to adopt each alternative. This is supposed to be absurd, because an agent cannot be required not to do what that agent is required to do. Defenders of moral dilemmas respond by denying either the closure principle or the claim that *required* implies *not required*, or both.

More arguments have been given against the possibility of moral dilemmas. Some philosophers claim that moral theories that yield moral dilemmas must be inconsistent or must fail to fulfill some purpose of moral theo-

ries, such as to prescribe particular decisions. Others argue that it would be unfair to blame or hold the agent responsible for failing to adopt one alternative when the agent adopted the other alternative in order to fulfill a non-overridden moral requirement. Defenders of moral dilemmas, of course, have responses to such arguments, but it remains controversial whether their responses are adequate.

See also Deontological Ethics; Duty; Moral Rules and Principles.

Bibliography

Gowans, Christopher W. *Innocence Lost: An Examination of Inescapable Wrongdoing*. New York; Oxford: Oxford University Press, 1994.

Gowans, Christopher W., ed. *Moral Dilemmas*. New York; Oxford: Oxford University Press, 1987.

Mason, G. E., ed. *Moral Dilemmas and Moral Theory*. New York: Oxford University Press, 1996.

Morton, Adam. *Disasters and Dilemmas: Strategies for Real-Life Decision Making*. Oxford; Cambridge, MA: Blackwell, 1991.

Sinnott-Armstrong, Walter. *Moral Dilemmas*. Oxford; New York: Blackwell, 1988.

Statman, Daniel. *Moral Dilemmas*. Amsterdam, Netherlands: Rodopi, 1995.

Zimmerman, Michael J. *The Concept of Moral Obligation*. New York: Cambridge University Press, 1996.

Walter Sinnott-Armstrong (2005)

MORAL EPISTEMOLOGY

It is easy to find examples of moral claims. People often say or write such things as: (a) Deliberate targeting of innocent civilians in war is wrong. (b) Women should get equal pay for equal work. I shall refer to the contents of moral claims as moral statements. I presuppose nothing controversial regarding the real nature of moral statements. The first two examples of moral statements are general, but many are particular, for example: (c) George Bush should not have invaded Iraq. (d) I ought to make a contribution to tsunami relief. Not all moral statements concern what is right or wrong, or what we should or should not do. Some concern our rights: (e) Everyone has a right to his or her own opinion. (f) The KKK has a right to adopt a highway just as any other group does. Other moral claims concern what is morally good or bad, what is virtuous or vicious, what is praiseworthy, when morally significant feelings such as guilt, remorse or gratitude, are appropriate, and so on. I hope this makes it sufficiently clear what I mean to count as moral statements.

Just as a person can be insincere in making a nonmoral claim, such as an ordinary factual claim, so a person can make a moral claim insincerely. It is not hard to imagine someone expressing agreement with others about some moral statement just to avoid confrontation, argument, or ridicule. In addition, we all recognize that what we say or write about morality does not exhaust our moral views, just as our factual beliefs can be more extensive than what we choose to make public. Let us therefore distinguish between people's moral judgments—that is, what they really think—and the public moral claims they make. I take no controversial stand regarding the nature of moral judgments.

A NARROW AND A BROAD UNDERSTANDING OF MORAL EPISTEMOLOGY

According to one traditional understanding, epistemology is the theory of knowledge. It is concerned with analyzing knowledge or specifying the conditions that must be satisfied for something to count as knowledge, with determining what we know and accounting for how we know it. Accordingly, moral epistemology would be concerned with moral knowledge. It would seek to determine whether any of our moral judgments count as knowledge and to provide an account of whatever moral knowledge we do have. Unfortunately this traditional understanding puts moral epistemology at risk of being a field with which many ethical theorists can have no substantial engagement.

Although there is a great deal of debate regarding the proper analysis of knowledge, nearly everyone agrees that for a person to know a statement (or proposition), that statement must be true. There is almost as wide agreement that a person must believe something in order to know it. In spite of the consensus that knowledge requires true belief, epistemologists do not work much on accounts of either truth or belief. They instead focus on figuring out what knowledge requires in addition to true belief and at understanding the precise nature of whatever else is required. Epistemologists do not agree about exactly what more is required for knowledge, but nearly all would accept that for a true belief to be knowledge it must be *good* in some yet to be specified but particularly epistemic sense. Epistemologists are, therefore, primarily concerned with understanding something normative or, more broadly, evaluative. When they attempt to determine whether we know something or how we might know it, they are engaged in an evaluative enterprise, seeking to address such questions as whether we ought to

hold the belief in question, whether we are justified or responsible or warranted in holding it, or simply whether the belief has some special positive epistemic status. Philosophically significant debates about skepticism regarding some type of belief rarely begin with the skeptics arguing that the beliefs are false. They typically charge that the beliefs are deficient in some other way—that they are unjustified or unwarranted—whereas nonskeptics try to show that the beliefs are legitimate or up to standard.

If we adopt the traditional knowledge-centered understanding of moral epistemology, many ethicists cannot take moral epistemology seriously; if they allow that there are any significant evaluative questions regarding moral judgments, they must take them to fall outside moral epistemology. One reason for this is that a great many ethical theorists accept some version of noncognitivism. This was the dominant metaethical position for a large part of the twentieth century, and it may still be the majority view. In spite of the apparent similarities between moral statements such as "Murder is wrong" and descriptive statements such as "The cat is black," noncognitivism holds that moral statements are not descriptive, that they do not state facts. Noncognitivists variously hold that moral statements instead do such things as vent emotions, state how one feels about certain actions and call upon others to feel the same way, make universal prescriptions, or express one's acceptance of norms.

Hence, according to noncognitivism, moral judgments blatantly fail to satisfy the most obvious necessary conditions for knowledge. No moral judgments are true for the simple reason that they are not the sort of thing that could possibly be true; like questions or commands, they are neither true nor false. We could put the point in other ways by saying that moral claims do not really make statements at all or that moral judgments do not have propositions—things that carry truth values—as their objects or contents. Hence, even when we sincerely make a moral claim, we are not really expressing a belief. If moral statements such as "Theft is wrong" are not descriptive but have some sort of noncognitive content—if they are, for example, ventings of emotion ("Theft: big time yucko!") or prescriptions ("Don't steal!")—then clearly their contents are not the sorts of things that one could possibly believe or, for that matter, disbelieve.

So noncognitivism entails the impossibility of moral knowledge. Regardless of how interesting the various versions of noncognitivism might be or how subtle and deep are the arguments that support them, no interesting normative epistemology is necessary to see this entailment. We need not get involved in any sort of epistemic evaluation of moral judgments to reach the skeptical conclusion. One need not do anything like reconstruct the evidence we have for our moral judgments and evaluate it to see how strong it is. One need not investigate the cognitive processes that produce moral judgments and attempt to determine how reliable they are. Since moral judgments just are not, according to noncognitivism, the sorts of things that could possibly be knowledge, there is no reason to get involved in the distinctive kind of evaluation of belief or judgment that is the special business of epistemology. Indeed, it would seem that epistemic evaluation of moral judgments could not really make any sense for a noncognitivist. Moral epistemology as an area of serious inquiry is left open only to cognitivists.

But of course this is not the way things are. Most people, regardless of their metaethical views, evaluate moral judgments, and they evaluate them in ways that seem no different from straightforward epistemic evaluations of ordinary factual judgments. They take some moral judgments to be epistemically better and others worse. People are dubious, for example, of moral judgments made on the basis of incomplete information or made when someone is tired or emotionally distraught, just as one would doubt factual judgments made in such circumstances.

We think we can at least sometimes provide reasons or evidence for or against moral judgments and that the reasons or evidence can be evaluated. We sometimes seek reasons for moral judgments we have already made, and at other times we try to find reasons that would allow us to make a moral judgment when we are unsure. In certain cases we ask others about their reasons for moral judgments and look askance upon their judgments if they can provide no adequate reasons. We are perfectly comfortable applying terms of epistemic evaluation such as *reasonable* and *unreasonable*, *rational* and *irrational*, *warranted* and *unwarranted*, or *justified* and *unjustifed* to moral judgments. When we apply these terms to moral judgments, it seems that we use them in the same way as when we apply them to other kinds of judgments.

There are two hard lines that affirm the restrictive understanding of epistemology as concerned exclusively with knowledge and accept that the conjunction of this conception with noncognitivism entails that the epistemic evaluation of moral judgments makes no sense. The one hard line concludes that epistemic evaluation of moral judgments, that is, moral epistemology, makes no sense. The other accepts the epistemic evaluation of moral judgments and rejects noncognitivism. I expect the first hard line approach would be more popular than the

second. But I prefer a third alternative. It maintains that epistemic evaluation of moral judgments makes perfectly good sense, as common practice suggests, and that most metaethical positions, including most versions of noncognitivism, can recognize this; it instead rejects the narrow understanding of epistemology. One advantage of this approach is that there are independent reasons for preferring a broader conception of epistemology.

We can extend the conception of epistemology, and specifically epistemic evaluation, in two ways. First, we should allow that epistemology is concerned with more than knowledge and its constituents. There are significant concepts of epistemic evaluation that do not figure in the analysis of knowledge. Some epistemologists account for knowledge in terms of reliable belief formation. Others disagree because a person's reliability may not be subjectively accessible to that person. They hold that knowing requires responsible belief, and that belief is irresponsible unless we have reason to think the belief is likely to be true. Others hold that to be known a belief must be properly based or grounded, while others hold that to be known a belief must be part of an extensive coherent system of beliefs. Yet others think a belief must be formed by a properly functioning cognitive mechanism. There are still more contenders: for example, those who analyze knowledge in terms of the exercise of intellectual virtues. Presumably at most one of these accounts provides a correct analysis of knowledge, but even those accounts that fail as analyses of knowledge may still succeed in identifying something that has epistemic value.

Whether reliable belief is necessary for knowledge, it is a good thing to be reliable in forming beliefs. The same holds for subjectively accessible reasons: It is clearly a good thing to have such reasons for a belief, regardless of whether they are necessary for knowledge—and so on for the various other evaluative characteristics of belief that have been put forward as necessary for knowledge. A strong case can be made that each is a real epistemic good. There are also concepts of epistemic evaluation that do not even seem to be required for knowledge. According to one account, rational beliefs are those that would stand up upon thorough reflection because they satisfy the believer's own deep epistemic standards. This is a highly subjective sense of rationality and therefore it is probably not required for knowledge. Nevertheless, it is epistemically good to have beliefs that satisfy one's own epistemic standards rather than beliefs that one would, upon careful consideration, regard as epistemically flawed. There are doubtless still more concepts of epistemic evaluation.

The second way to broaden epistemology is by abandoning the dominant monistic view of epistemic evaluation that regards truth as sole intrinsic epistemic good and all other epistemic goods as valuable because of some connection to truth such as being a means to true belief. There have been attempts to show that some features, such as coherence, make truth more likely, but these attempts have not met with much success. It has seemed obvious all along that other features, such as subjective rationality, do not make true belief objectively likely. We need not conclude that no such features are epistemically valuable. It is better to allow that some things we value epistemically do not make true belief likely. In the case of something like reliable belief, at least on some understandings, the connection with truth is obvious. But even here we should take a broader view, at least for moral judgments. As we have seen, noncognitivism entails that moral judgments have no truth values and hence cannot be reliable.

Nevertheless, it seems obvious that some moral judgments are more reliable than others. For example, moral judgments made by a person who is emotionally distraught or who has selfish interests at stake are less than reliable. Most noncognitivists can easily accept such seemingly obvious examples, since most draw some sort of distinction between correct and incorrect moral judgments. Hence a notion of reliability is available that is an extension of the familiar, truth-connected notion. It makes more sense to recognize judgments that are reliable in this extended sense as epistemically valuable than to think that we are making an epistemic evaluation when we criticize a factual belief because a person formed it, say, when in a rage, but some totally different kind of evaluation when we criticize a moral judgment for exactly the same reason.

EPISTEMIC EVALUATION OF MORAL JUDGMENTS

If the broad conception of moral epistemology is basically correct, we should not ask simply whether any moral judgments are known or justified. Recognizing that there are various significant concepts of epistemic evaluation, we should ask what, if any, positive epistemic statuses moral judgments might have and also whether moral judgments suffer from any epistemic flaws so severe that we should regard them with a robust skepticism—a skepticism that holds not merely that no moral statements are known, but that moral judgments are so flawed that it makes no sense to use them either in moral theorizing or as a guide to life and action.

Some moral judgments are bound to be epistemically flawed for straightforward reasons—for example, because they were formed by a person who was emotionally distraught or who stood to gain or lose depending upon the judgment, or because they were made on the basis of an incomplete or incorrect understanding of the facts of the case, or because the person judging feels unsure or has no stable opinion. We know that judgments like these run a significant risk of error, regardless of their content. Let us set aside such obviously flawed judgments and focus on those that are free of all such well recognized sources of error. Such moral judgments already have some positive epistemic status—they have managed to avoid some significant pitfalls. But this is not, perhaps, a very impressive status, so let us consider what more might be said on behalf of moral judgments.

Among the remaining moral judgments, we can distinguish between those formed or not formed on the basis of inference. We obviously cannot have formed all of our judgments by inferring them from other judgments. Some of our judgments must be noninferential. It might be that all moral judgments are inferred from nonmoral judgments, either immediately or by means of inferential chains that eventually terminate exclusively in nonmoral judgments. Certain ethicists have tried to ground moral judgments in something like this way, deriving them from theses regarding the meanings of moral terms in conjunction with purely empirical claims. But it seems highly unlikely that anything like this will work out, and near certain that the moral judgments of ordinary people are not grounded in this way. Ordinary people, and even philosophers when they are being ordinary, form many noninferential moral judgments, and when they do infer moral judgments, the inferences have moral premises that are, or eventually trace to, noninferential moral judgments. So let us focus on noninferential moral judgments.

Consider the widely shared judgment that it is wrong to cause animals suffering for no good reason. Those who share this general judgment will also make judgments regarding the wrongness of many particular cases of animal torture. It is certainly possible to reach the general judgment via inference or to infer the particular judgments from it. But it is also possible to make both judgments noninferentially. Even where the judgments are noninferential, it is quite obvious that they do not come from nowhere. We were taught to make such judgments as children. At some time or other when we were children, our parents or some other adult caught us, or perhaps a sibling or a friend tormenting some helpless small animal and scolded us. Maybe one incident was enough; maybe similar incidents were repeated, but eventually the lesson stuck.

Perhaps, then, our noninferential moral judgments get their epistemic status in the same way as our beliefs in other things we were taught as children. I believe that my maternal grandfather was killed in World War II, before I was born. This belief is noninferential, but it does not just pop into my head from I know not where. I know full well that it arises from testimony and memory. When I was a child my mother told me this. I believed her. Although I never received any objective confirmation of the belief—for example, by reading a letter from the War Department—neither did I encounter any reason to doubt what I was told. And I still remember what I was told. This is sufficient for my belief to have some fairly impressive epistemic credentials. My belief is rational or reasonable. You could say that I am epistemically responsible in believing. The belief coheres with other things I believe, although I would have to admit that most of the relevant beliefs are also things I remember being told by my parents. My mother has usually been reliable, and I know this to be so because in many cases what she told me has been borne out by the future course of experience. And I know my memory is fairly reliable as well, at least about things like this.

As good as all this is, however, it is still possible that my belief is seriously flawed. Suppose my grandfather mysteriously disappeared around the start of the war, and, although my mother knew he was involved with organized crime, she deceived herself into believing he had gone off to the war. When he didn't return at the end of the war, she came to believe he had been killed in action. Under this scenario, given the fact that my mother's original belief about her father's fate had little positive epistemic status—indeed, was flawed—my belief would be seriously flawed. It is significant that in such cases a testimonial belief can have a higher epistemic status than the belief of the testifier. Nevertheless, the epistemic status attainable by beliefs that are (solely) grounded in testimony and memory is constrained by the epistemic status of the testifier's belief. The epistemic status of memorial beliefs is similarly constrained by the status of the original belief.

Of course, the adults who taught us about morality when we were children probably did not fabricate their own moral views in some strange way. They were taught about morality by their parents, who were taught by theirs, and so on. This suggests a somewhat different problem: In the case of beliefs that have their source in

chains of testimony and memory involving a series of people, somewhere along the line someone must have formed the relevant beliefs in some other way. And if the beliefs that come later in the chain are to be free from significant epistemic defect, somewhere along the line some beliefs must have attained some fairly strong positive epistemic status in virtue of something other than testimony and memory.

In the case of historical beliefs, which presumably trace back to persons who witnessed the events in question, it might make sense to suppose that the original source beliefs had the requisite epistemic credentials. But in the case of moral judgments it is hard to credit such a view, unless one takes something like the biblical narrative of Moses quite literally and holds that all our moral judgments can be traced through a long chain of testimony and memory all the way back to Moses or some other prophet whose moral judgments came straight from God. My guess is that even many theists will find such a supposition incredible.

How, then, might noninferential moral judgments attain a significant positive epistemic status? Here is a possibility: We were also taught to make simple arithmetical judgments. I can well remember trying to memorize multiplication and division tables. But although testimony and memory are surely somehow involved in the arithmetical judgments we now make, these judgments do not get their epistemic status primarily from testimony and memory. Indeed, I doubt that our simple arithmetical judgments are even produced by memory and testimony any longer. Somewhere along the line, no doubt as a result of our training, we reached a point where we could simply see for ourselves that simple arithmetical propositions are true. Simple mathematical and logical propositions, and perhaps some few others, are special. Any person with the conceptual resources to really understand the propositions can simply see that they are true, or at least this is one venerable and still widely held view. Some ethicists have wanted to say that certain ethical statements are like this as well. So the first part of the current proposal is that although we were taught to make moral judgments when we were children, such judgments are no longer merely products of testimony and memory. Rather, when we understand and consider certain moral statements, they simply seem to us as though they are true, so we form the moral judgment. The second, explicitly epistemic part of the proposal is that such moral judgments have the same positive epistemic status as simple arithmetical judgments and come to have this status in the same way.

There are reasons for being suspicious that things are quite so simple. Before I explain why, here are a couple of terminological notes. Contemporary discussions of moral epistemology and methodology frequently are conducted in terms of considered moral judgments and moral intuitions. Considered moral judgments are typically characterized simply as noninferential moral judgments that are not subject to obvious sources of error. When we narrowed our focus to such judgments above, however, I did not refer to them as considered moral judgments because judgments formed through testimony are not inferential in their origin, and neither are memorial judgments. Nevertheless, moral judgments formed via testimony or memory are not considered moral judgments even if they have avoided the usual sources of error. In the first part of the proposal regarding moral judgments we have restricted our attention to judgments that are free of the usual sources of error and are not only noninferential but are also held simply because it seems to the believer that they are true. Such judgments are appropriately regarded as considered moral judgments.

The term *intuition* can be used in a stronger, epistemically loaded sense or a weaker, nonepistemic sense. In the weaker sense, intuitions are simply noninferential judgments that do not arise from any of the traditionally recognized sources of knowledge: Intuitions are not produced by sense perception, introspection, memory, or testimony. A person makes an intuitive judgment simply because the proposition seems true upon due consideration. Considered moral judgments are, therefore, a subset of moral intuitions, namely, those that have avoided obvious causes or error. Limiting ourselves to the first part of the current proposal regarding moral judgments, we could say these judgments are moral intuitions in the weak sense. There are various stronger concepts of intuition that add to the weak notion a claim to some positive epistemic status—often some strong status such as certainty or infallibility or incorrigibility. Critical discussions of intuitionism often assume a strong notion of intuition, most frequently one involving a very strong epistemic status. The second part of the current proposal takes moral judgments to be moral intuitions in a very strong sense.

I would like to consider two significant grounds for doubting that our considered moral judgments are epistemically similar to simple mathematical judgments. They also may seem to be grounds for doubting that considered moral judgments are intuitions in any strong sense and even that considered moral judgments could have a significant positive epistemic status. The first

ground for doubt is based on the fact that our considered moral judgments seem to be revisable; the second is based on the fact that there is considerable disagreement regarding these judgments.

Most people who reflect on their moral views encounter conflicts among their considered moral judgments. Many of us find certain moral principles intuitively obvious, particularly midlevel principles such as "It is right to keep one's promises" and "It is wrong to lie." One need not reflect very long to come up with cases where application of an intuitive principle produces a judgment at odds with our considered moral judgment regarding the case. (This is just what one does when arguing by counterexample.) Conflicts can also emerge if we make different intuitive moral judgments about different particular cases and there is no difference between the cases that we judge sufficient to justify our different moral judgments. When we encounter conflicts among our considered moral judgments, moral reflection obviously does not halt. We decide what to revise and move on. But the existence, or more properly, the frequency of such conflicts does seem to count against the claim that our considered moral judgments are epistemically similar to simple mathematical beliefs.

The problem is not that our intuitive judgments about simple logical and mathematical propositions could never come into conflict and can never be revised. There are mathematical propositions that seem intuitively obvious but lead to paradox—that is, they come into conflict with other intuitive mathematical propositions. In such cases we are led to revise some intuitive judgments. But such occurrences are the rare exception in mathematics and logic, however, and vastly more common with moral judgments. Hence, although we might get away with claiming that simple mathematical propositions can be seen to be true by anyone who adequately understands them, even though we are forced to allow that those who adequately understand are sometimes mistaken when they think they see something to be true, the parallel claim regarding considered moral judgments seems much less plausible.

People seem to disagree a lot about morality. Some of the differences might not constitute conflicts—that is, cases where the judgments are inconsistent. Some of the differences might arise from misunderstanding on the part of one or both parties, and some might not involve considered moral judgments. But even setting aside such disagreements, there are many real conflicts between the considered moral judgments of mature adults who fully understand. Not only does the existence and extent of these conflicts render untenable the claim that considered moral judgments have the same epistemic status as simple mathematical judgments, but it also could be taken to block the claim that considered moral judgments have any significant positive epistemic status. One reason is that, in many cases of conflict, the parties to the different sides come from different societies or cultures, a circumstance that seems to support the idea that moral judgments are some sort of social or cultural construct. They might then be reliable guides to the taboos or mores of the judge's own culture but not to anything more substantial or objective. When conflicts among considered moral judgments within cultures are added to the mix, we seem to have ample reason to doubt whether they are reliable guides to anything at all.

Actually, the fact that a single person's considered moral judgments can conflict and require revision is not a bad thing. Indeed, it is a fundamental element of the most influential approach to the construction and justification of moral theories. According to the method of reflective equilibrium, we should strive to mold our considered moral judgments and a set of moral principles that account for them into a coherent system via a series of mutual adjustments to principles and judgments, with revisions guided only by what seems most likely to be true upon due consideration. If only considered moral judgments and moral principles are involved, a narrow reflective equilibrium emerges. Inquirers should next strive to bring their judgments into a wide equilibrium, which also includes background theories and judgments—for example, views regarding the nature of persons or the role of morality in society. Once again, in the search for a wide reflective equilibrium, no type of judgment has a privileged status. Coherence is attained by a series of mutual adjustments.

The method of reflective equilibrium is an idealization of the kind of moral inquiry carried on by many philosophers and presumably by at least some reflective nonspecialists. We might, then, shift away from the considered moral judgments of ordinary people and ask about the epistemic status of the moral judgments we would hold if we brought our moral judgments into wide reflective equilibrium. It might be all but impossible for us ever to attain such equilibrium, but perhaps we can approach it ever more closely. The moral judgments a person holds in reflective equilibrium would have a number of epistemically good features: they would have been formed after careful and thorough reflection, they would not conflict with either the person's other moral judgments or any of the person's other beliefs, and they would

be part of a highly coherent system of beliefs and judgments. Moreover, we might hope that there would be fewer conflicts between the moral judgments of different people who had brought their beliefs into reflective equilibrium. One reason for this hope is that part of the method explicitly involves considering alternatives to one's own moral system.

Unfortunately, I fear we cannot expect that inquirers will converge upon a single system of moral judgments in wide reflective equilibrium. It is too easy to imagine people who begin with radically different moral perspectives being led to revise their judgments in different ways to overcome the conflicts internal to their own moral perspectives, and so at the end of their inquiries being led to accept very different, incompatible moral systems in reflective equilibrium. So questions about the reliability of moral judgments persist.

I will close by briefly describing one possible way of addressing such questions. Suppose that people differ in their capacities for making moral judgments. Suppose that this capacity needs to be developed through experience and possibly even training, but that it can also be corrupted. (For what it is worth, common sense strongly supports these suppositions.) Let's call a person with a well-developed capacity for moral judgment a competent moral judge. If two people with unequally developed capacities for moral judgment were to bring their moral judgments into reflective equilibrium, they would probably disagree to some extent. Such disagreement would not establish that the moral judgments of both inquirers were unreliable, however, for it might be that only one of the inquirers is a competent moral judge. Presumably the moral judgments the competent judge would make in reflective equilibrium would be quite reliable. Since the other person's moral judgments would also be in reflective equilibrium, it would not be possible to prove to that person that his or her judgments are unreliable or that the competent judge's moral judgments are reliable. But this would not change the fact that the competent judge's moral judgments would be reliable.

One might require that, in order for a person's moral judgments to have a significant positive epistemic status, a person must be able to prove that his or her moral judgments are reliable or that he or she is a competent moral judge. If this is right, then we will have to grant that even the moral judgments competent judges hold in reflective equilibrium have no significant positive epistemic status. We should note, however, that if similar requirements were imposed across the board, we would be forced to conclude that almost none of our beliefs or judgments have a significant positive epistemic status. On the other hand, if actually being reliable is sufficient for having a significant positive epistemic status, at least in conjunction with all the other epistemic goods we have identified, then it seems that the moral judgments competent judges would make in reflective equilibrium will have such a status. One might doubt whether there are any competent judges, but I do not think we know that there are not. So there is reason to hope that moral judgments can attain a strong positive epistemic status.

See also Metaethics; Meaning; Moral Skepticism; Noncognitivism; Rationality.

Bibliography

Alston, William. "Epistemic Desiderata." *Philosophy and Phenomenological Research* 53 (1993): 527–551.

Alston, William. "Level Confusions in Epistemology." In *Studies in Epistemology*. Vol. 5 of *Midwest Studies in Philosophy*, edited by Peter A. French, Theodore E. Uehling Jr., and Howard K. Wettstein, 135–150. Minneapolis: University of Minnesota Press, 1980.

Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knoweldge*. London: Routledge, 1998.

Audi, Robert. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton, NJ: Princeton University Press, 2004.

Ayer, A. J. *Language, Truth, and Logic*. London: V. Gollancz, 1938.

Blackburn, Simon. *Spreading the World*. Oxford: Clarendon Press, 1984.

BonJour, Lawrence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Brandt, Richard. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Brink, David. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Darwall, Stephen, Allan Gibbard, and Peter Railton. "Toward fin de siècle Ethics: Some Trends." *The Philosophical Review* 101 (1992): 115–189.

Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76 (1979): 256–282.

DePaul, Michael R. *Balance and Refinement: Beyond Coherentism in Moral Inquiry*. London: Routledge, 1993.

DePaul, Michael R. "Value Monism in Epistemology." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, edited by Matthias Steup. Oxford: Oxford University Press, 2001.

DePaul, Michael R., and William Ramsey. *Rethinking Intuition: The Psychology of Intuitions and Their Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield, 1998.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Gibbard, Allan. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press, 1990.

Goldman, Alvin. "What Is Justified Belief?" In *Justification and Knowledge*, edited by George Pappas. Dordrecht: D. Reidel, 1979.

Hare, R.M. *The Language of Moral*. Oxford: Clarendon Press, 1952.

Harman, Gilbert. *The Nature of Morality*. New York: Oxford University Press, 1977.

McDowell, John "Values and Secondary Qualities." In *Morality and Objectivity*, edited by Ted Honderich. London: Routledge & Kegan Paul, 1985.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1987.

Plantinga, Alvin. Warrant: *The Current Debate*. Oxford: Oxford University Press, 1993.

Prichard, H. A. "Does Moral Philosophy Rest on a Mistake?" *Mind* 21(1912): 21–37.

Rawls, John. "Outline of a Decision Procedure for Ethics." *The Philosophical Review* 60 (1951): 177–197.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Ross, W. D. *The Right and the Good*. Oxford: Oxford University Press, 1930.

Sinnott-Armstrong, Walter, and Mark Timmons, eds. *Moral Knowledge? New Readings in Moral Epistemology*. Oxford: Oxford University Press, 1996.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.

Sosa, Ernest. "Minimal Intuition." In *Rethinking Intuition: The Psychology of Intuitions and Their Role in Philosophical Inquiry*, edited by Michael DePaul and William Ramsey. Lanham, MD: Rowman and Littlefield, 1998.

Stevenson, Charles L. "The Emotive Meaning of Ethical Terms." *Mind* 46 (1937): 14–31.

Stich, Steven. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Wiggins, David. *Needs, Values, Truth: Essays in the Philosophy of Value*. Oxford: Basil Blackwell, 1987.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Zagzebski, Linda. "What Is Knowledge?" In *The Blackwell Guide to Epistemology*, edited by John Greco and Ernest Sosa. Oxford: Blackwell, 1999.

Michael R. DePaul (1996, 2005)

MORALITY

See *Ethics; Ethics and Morality*

MORAL LAW, THE

See *Kant, Immanuel*

MORAL MOTIVATION

See *Moral Psychology; Moral Sentiments; Virtue Ethics*

MORAL NATURALISM

See *Ethical Naturalism*

MORAL PHILOSOPHY

See *Ehtics, History of*

MORAL PRINCIPLES: THEIR JUSTIFICATION

The problem of how, if at all, we could set about justifying assertions about what we ought to do in various practical situations is one that has been the major concern of moral philosophers. Such basic questions are indeed endemic in most branches of philosophy. We ask not only if we can ever know what we ought to do but whether we can justify our claims to knowledge of an external world, how we can know the truth of statements about the past, or whether we can ever be sure of the existence of minds other than our own. But in ethics the problem seems more recalcitrant and, indeed, to many nonphilosophers at least, more real. For while skepticism about the existence of an external world or of other minds may seem difficult to refute, to most it is impossible to embrace, whereas skepticism about the possibility of claiming knowledge of any objective truths about what we ought to do is not so rare, either among men in general or those who would wish to characterize themselves as philosophers.

It is not, of course, surprising that this should be so. Ethical attitudes vary much more, from society to society and even between individuals, than do our beliefs about the external world or other people's feelings. The patent fact of ethical disagreement forces us to reexamine the bases of our moral beliefs. Furthermore, the disagreements we encounter concerning moral issues often seem to involve deep matters of principle that leave no common ground between the disputants. This is sometimes referred to as the problem of disagreement about ultimate moral principles. It is this problem—whether ultimate moral principles are susceptible of rational justification—that will be examined in this article.

Most philosophers would agree that the particular way in which a philosophical problem is formulated will make a great deal of difference to what solution is possible to it or, indeed, whether any solution is possible. It will be necessary therefore to set out in detail what is meant by a disagreement about ultimate moral principles and to defend this way of expressing the issue against certain objections before a possible solution is set out.

MORAL PRINCIPLES

A "man of principle" is sometimes thought of, with distaste, as a man who acts in accordance with a fixed set of rules, ignoring the complexities of the situation and failing to adapt his behavior to changing circumstances. The morality of principles and rules is sometimes contrasted with the morality of sensibility, which emphasizes such virtues as sympathy and integrity as against a rigid code of behavior. In either kind of morality, however, particular judgments will have to be made, based on a view of the situation in which the agent acts, and some factors in the situation will have to be regarded as reasons for acting in one way rather than another. There is, therefore, a more general sense of "moral principle," which can be regarded as common to both views, in which a moral principle indicates some factor that is generally relevant to what ought to be done.

Moral principles can then be regarded as statements picking out those factors of situations that can be appealed to as moral reasons. "Lying is wrong" suggests that the fact that a statement is known to be false is a reason for not making it to someone. "Adultery is wrong" suggests that the fact that someone is married is a reason for his refraining from sexual intercourse with any person who is not his spouse. And, again, "One ought to be kind" suggests that there are reasons for performing kind actions rather than unkind ones. Asserting a moral principle of this kind and denying the suggestion about reasons results in paradox. Thus, for example, if somebody says "Lying is wrong, but the fact that a statement constitutes a lie is no reason whatsoever for not making it," he seems to have taken back in the second half of his sentence what he asserted in the first.

If saying that someone ought to do something commits one to claiming that there is some fact in the situation that is a reason for doing the thing in question, then this reason must be subject to the requirement that reasons in general must satisfy: that anything that is a reason in any one case must be a reason in every case unless there are other special reasons for ignoring it. This applies to reasons generally, not just to moral reasons. For example, if the fact that it is raining is a reason for saying Smith will get wet, it is a reason for saying anyone else will unless there are some relevant differences in their cases, such as being indoors or carrying an umbrella. It is this that leads to the claim that moral principles must be universal, at least to the degree that they pick out factors that are universally relevant to what we ought to do, although not necessarily universally determining what we ought to do in every particular case. Thus it would seem that the correctness of the universal moral principle involved—or, in other words, that what is appealed to as a reason should indeed be a reason—is a necessary although not a sufficient condition of the correctness of the particular judgment about what ought to be done.

JUSTIFICATION OF MORAL JUDGMENTS. If the correctness of universal moral principles is a condition of the correctness of particular moral judgments, then obviously the first question we must ask in investigating how our particular moral judgments can be justified is, How can we justify claiming that certain moral principles are correct? There are, however, some objections to this way of treating the problem that must be considered.

It may be pointed out that value judgments in other areas do not seem to require justification by reference to some universally relevant factors. And if we are willing to allow that in other realms of value there are judgments that do not require to be backed by universal principles, why not in morals? For example, there are very considerable difficulties in representing judgments about the value of a work of art as being backed by or dependent on principles at all. It may be impossible, when we say some work of art is good, to indicate any feature the possession of which is bound to make any other work of art good. (One might be tempted to say that beauty is such a feature. But this is unconvincing because one is using the term either narrowly, in which case there are plenty of good works of art that one would never describe as beautiful, or so widely that it means only "good in the way that a work of art is good.") Surely, however, it must be agreed that the goodness of anything, including a work of art, depends on what qualities it has, however difficult it may be to say in a given case precisely what qualities it has that make it good. And in order to begin to justify the judgment that something is good, one must refer to its qualities; one cannot draw anyone's attention to the goodness itself. If it is proper to refer to these qualities to back one's claim that the object is good, then it is at least to the point to ask why something else, which has the same qualities, is not good. If such a question is to the point, it shows that we accept that the possession of certain qualities is

being put forward as a general reason for saying that the object is good.

Even if this is correct, however, it is clear that the features by virtue of which any given work of art is judged to be good tend to be many, complicated, and organically related. Although any feature pointed to in support of a judgment that a work of art is good must also be relevant to the criticism of other works of art, there may be in every other case many other relevant factors that alter the situation completely. The same thing might be claimed for moral cases. It may be said that every human situation is infinitely complicated, so that however many relevant features one may pick out in a particular case, there will always be a host of others that can be set against them. Such considerations would lead not so much to a denial of the universality of morally relevant features as to doubt about the utility of stating the problem in terms of principles. To this there are two answers.

First, it would be against common sense to claim, for example, that the wanton murder of children is not wrong. Even where other features that are regarded as morally relevant are also present—such as that one had promised one's old mother on her deathbed to try to exterminate the Jews—few would regard them as justifying child murder. So anyone who persists in claiming that it is always possible that such actions as child murder may be justified because of the complex character of every particular human situation is, at best, someone who has an unusual moral outlook, and this means that his very claim that every situation is so complicated that no general principles can be admitted is dependent on his having a different set of moral principles from most people's. So even to consider whether this objection is correct, we still have to ask which general principles are justifiable.

Second, we have already remarked that moral principles will be a necessary but not a sufficient condition of the correctness of our particular moral judgments. Although on their own they may never be sufficient to solving all moral problems, they will certainly be necessary to our having any moral problems at all. This may be illustrated in terms of a case mentioned by Jean-Paul Sartre. A young man has a dilemma. Should he join the French Resistance, or should he stay at home and look after his aging mother? Sartre points out that no rehearsal of general principles would ever serve to solve such a problem. This is no doubt true, but it does not show that the correctness of such principles is not relevant. For why is the young man worried about only those two possibilities? There are plenty of other things he could do. He could learn tightrope walking or set up as an ice-cream vendor or enlarge his earlobes with brass rings. But these are obviously of no importance, whereas looking after the old mother and joining the resistance are important. Why is Sartre's case serious and dramatic and the other suggestions frivolous and silly? Why does it matter what the young man does, to himself or to anyone? There can surely be no problem at all unless such things as joining the resistance (defending one's country) or looking after the old mother (kindness to a dependent) are morally relevant features of the situation—unless they are things that it is reasonable to consider in deciding what to do. And if there are morally relevant features in the situation, there are corresponding moral principles. If these principles are not correct (and, indeed, there are those who would question patriotic principles), then there is no problem, or at least not the same problem.

A different kind of objection can be disposed of very briefly. It is that as a matter of experience, we do not think in terms of principles. Rather, on particular occasions we simply know instinctively what is right. Now this may very well be true or perhaps true for a number of people. However, the question at issue is not a psychological one about the kind of process that goes on before a moral judgment is made; it is a philosophical one about how we may justify making the moral judgments we do make, by whatever psychological process we make them. Whatever goes on in the heads of mathematicians, it is still Euclid's proofs alone that can justify Euclid's theorems.

ULTIMATE MORAL PRINCIPLES

Moral principles in the sense adumbrated above will be of varying degrees of generality, and some will be held to be more fundamental than others. For example, the principle that one ought not to commit adultery may be defended on the ground that adultery is inimical to the stability of the family. In terms of reasons for acting, this can be put as follows. The fact that someone is married is held to be a reason for his refraining from sexual intercourse with anyone other than his spouse. But why is this a reason? Because, it might be said, *in fact* sexual infidelity is apt to break up the unity of the family. Such an argument would, of course, presuppose that the fact that something is apt to disrupt the family is a reason for avoiding it or, in other words, that one ought not to disrupt the family. Thus the principle "One ought not to commit adultery" would be regarded as less fundamental than the principle "One ought not to disrupt the unity of the family." In the process of trying to justify particular moral judgments, we will usually find ourselves trying to show that certain necessary conditions of their correct-

ness, our moral principles, have further necessary conditions in terms of more fundamental moral principles. The process will usually be much more complicated than I have represented it; in justifying a less fundamental moral principle, we will usually find a variety of more fundamental moral principles coming into play. But however complicated such a process may be, it is obvious that we cannot suppose it to go on forever. At some point we should reach some principles that we regard as the most fundamental. For example, we might want to say that we do not claim that one ought to be kind because this follows from some further principle; we ought to be kind because we ought, and that is an end to the matter. These we may call ultimate moral principles, and their correctness is a necessary condition of the correctness of all other moral judgments. Unless some such ultimate moral principles can be shown to be justifiable, no other moral judgments can be shown to be justifiable.

Some philosophers hold that this representation of the matter is utterly mistaken and, indeed, that it is precisely because of this "justificationist" view that so many philosophers despair of finding an answer and become ethical skeptics. If, it is argued, moral principles are regarded not as first premises from which a moral system is deduced but as conjectures that can be altered and amended by subsequent moral experience, we at least have a method of correcting our moral attitudes that will justify us in claiming that they are more or less rationally defensible. It will not be possible to do this view justice in a small space. It can only be said here that the major difficulty with this view is that the test of the moral principle is taken to be the particular judgments we are inclined to make, particular judgments that conflict with the supposed principle and thus refute it. But what is now the test of the correctness of the particular judgment? The suggested method would seem to be a way of finding out, by examining someone's particular judgments, what his moral principles are rather than a way of finding out which moral principles are correct. Furthermore, it has not been claimed in this article that moral principles are first premises from which whole moral systems can be deduced but only that moral principles are statements of relevant moral factors. Their correctness is a necessary, not a sufficient, condition of the correctness of moral judgments.

Nevertheless, the charge is certainly well founded that this way of setting out the problem is a most plausible invitation to ethical skepticism. For it would on the face of it appear that the very statement of the problem precludes its solution. If we look on more and more general moral principles as representing a regress of necessary conditions of the correctness of moral judgments, then either this regress is viciously infinite or there is a point at which it must stop. But any attempt to justify some principle as a stopping point would appear to start the whole process off again. To acquiesce in some stopping point would be to accept an ultimate principle and, it would seem, to accept that nothing further could be said in its justification. It looks then as if this way of putting the problem makes inevitable the conclusion that ultimate principles are unjustifiable.

AUTONOMY AND OBJECTIVITY OF MORAL PRINCIPLES

One way to put the problem is to regard it as a conflict between the autonomy and the objectivity of moral principles. The demand that ethics be regarded as autonomous originated with Immanuel Kant, in the view that an action is not moral unless it is determined by the agent's rational will rather than by something external to that will, such as a desire, or the will of another (a king, a friend, the state, God). Here the concern is with the determination of action, not directly with the determination or, rather, justification of moral judgment. The autonomy of moral principles, with which we are concerned, is not, however, entirely unconnected with Kant's sense of autonomy. It is the idea that a moral judgment can never depend for its correctness entirely on factors that are nonmoral; that is, that in the justification of any moral judgment one must have recourse to a moral principle, which must in turn be justified in terms of some more general moral principle and so on. In other words, a moral judgment or principle is never deducible from any set of premises that contain no moral judgment or principle.

The demand that morality be regarded as objective was also emphasized by Kant. A moral act for Kant was one that could be willed by an autonomous, rational will; its character as a moral act depended not on the particular nature or desires of the willing agent but on the nature of a rational will as such. For Kant a maxim is objective when it is valid for any rational being. Again, Kant's concern was with the determination of action rather than the justification of judgment. But once again our sense of objectivity is not unconnected with Kant's. When someone's judgment is stigmatized as subjective rather than objective, this means that some idiosyncratic factors such as the hopes and fears or special interests of the speaker have affected his judgment; an objective judgment, however, is one not affected by such idiosyncratic factors but

one that any reasonable and unbiased person would form in the circumstances. Obviously, we can speak of objective matters only in respect of matters that are publicly determinable, where we can talk of what would be judged by any reasonable and careful observer rather than what appears to be the case to some individual because of some peculiarities of his own. Thus, we might say with Kant that objectively true judgments are those that are "valid for all rational beings" rather than what merely seems to be so to certain individuals. The demand of objectivity in ethics may then be put at its most minimal as the demand that the truth of any moral judgment shall not depend on the peculiarities of the person making it but, rather, that it shall be determinable by any rational observer who is apprised of the facts. Its truth will not depend on the fact that it is judged so by some one person rather than another but on objective considerations.

The conflict between the demands of objectivity and autonomy is now not difficult to see. For how can ultimate principles, which cannot be based on any further considerations, be based on objective considerations? How can we claim that they are matters that are publicly determinable when it would seem that, if they were autonomous, no considerations beyond themselves would make their truth determinable at all?

Henry Sidgwick, impressed by the utilitarian moral system but despairing of the kinds of argument put forward by earlier utilitarians such as Jeremy Bentham and John Stuart Mill to justify their ultimate principle, substituted instead the doctrine of intuition, a doctrine that was accepted by many other philosophers who were very far from being utilitarians. It was thought that the problem of justification in ethics was parallel to similar problems in other fields of knowledge and that in each case one would find oneself with incorrigible starting points, truths known directly, without inference or the necessity or possibility of further justification. Thus, in our knowledge of the world we might be thought to begin with direct awareness of our experience; in mathematics, with the direct perception of mathematical relationships. In ethics we begin simply with the perception of universal ethical relationships, between what is right or fitting and certain states of affairs. Whatever the difficulties in this general epistemological theory, in ethics there is the additional difficulty that the commonsense roots of the problem of justification—the inescapable fact of disagreement on fundamental ethical matters—are untouched by the doctrine of intuitionism. The appeal to intuition in the face of this disagreement leaves no way of rationally resolving it.

TRANSCENDENTAL ARGUMENTS

It is possible, however, that an account of the justification of ultimate principles can be given that avoids both an infinite regress of justifying principles and any arbitrary stopping point. Kant's demands for autonomy and objectivity amount to the requirement that a morally good action be rationally chosen in accord with a law that is valid for all rational beings universally and that is determined by nothing beyond itself. The difficulties in making the demands of autonomy and objectivity compatible, so that this requirement becomes a feasible one, seem capable of only one kind of solution, which was the one adopted by Kant. If moral principles cannot be justified by considerations outside themselves yet must be regarded as objectively justifiable, then it seems that certain moral principles must somehow be demanded by the formal character of morality itself; certain rules must be required by any morality that is to satisfy the two demands.

Kant's particular solution has not seemed very satisfactory, but if a solution is to be found at all, it must be in the same direction. To put the point in more contemporary language, the only kind of solution that seems possible is one that shows that certain moral principles must be regarded as correct if moral discourse is to be possible at all, at least as an autonomous and objective form of practical discourse. An argument to this effect may be called a transcendental argument. If such arguments can be constructed, it should be easy to see how they solve the problem we have been considering. For a principle can be shown to be objectively true, without appealing to factors outside itself, if it can be shown that the form of discourse of which the principle is an example is impossible without presupposing the principle. That is, by showing that no one can claim to be using a form of autonomous, practical, and objective discourse unless he at the same time accepts the principle in question.

Three arguments of this kind can be advanced to establish three ultimate principles, which we may call the principles of impartiality, rational benevolence, and liberty. It is important that throughout it should be borne in mind that these arguments are intended to establish ultimate principles—that is, factors of the most general moral relevance, which will be necessary, but by no means sufficient, to establishing any correct moral theories, rules, or particular judgments. Even given that these arguments establish the ultimate principles of impartiality, rational benevolence, and liberty, there will still remain the difficult problem of their application in practice.

IMPARTIALITY. As far as we are concerned with a form of discourse in which we objectively judge actions right or wrong, so that a correct practical judgment is one that could in principle be reached by anybody, such judgments must be made in terms of features that the actions or the situations in which they are done possess and not on any other factors arbitrarily introduced by the person making the judgment. Thus, any feature picked out as relevant must be one that is always relevant unless there is some special explanation, for a feature that is relevant in one case and not in another, where there is no further difference, is one that is not relevant at all in any ordinary sense and forms no guide to action. It follows that any action that it is right or wrong for one person to do is right or wrong for every person to do unless there are some special factors present in the other cases. And from this demand of universality it follows, insofar as morality is practical, that one ought to act in accordance with it: What anyone ought to do in any given set of circumstances is what anyone else ought to do, as long as his case is not relevantly different, and anything one ought to do on any given occasion is what one ought to do on every occasion unless again there are factors present that are relevantly different. That one ought to treat similar cases similarly is obviously a general case of the particular requirement of justice toward men, that any form of treatment that is thought to be right for one man must be right for all others, unless the others are significantly different.

RATIONAL BENEVOLENCE. The principle of rational benevolence is that stated by Sidgwick, that one ought in action to consider the interests of all beings in the universe. That this is a most impractical injunction is important, but not fatal, for how in practical situations we may apply any ultimate principle is another, though admittedly difficult, question.

The principle may be justified as follows. The demand of objectivity is that what is right or wrong should be determinable at least in principle by all rational beings. This requires that moral discourse should be a form of public discourse, in which the relevance and force of any consideration is dependent on its content and not on the will or status of whoever puts it forward. That is, the remark of any rational being may be relevant to the question whether some action is right or wrong. The ideal of this form of discourse therefore requires that it should be possible for any rational being to participate in it as an interlocutor; if any is excluded arbitrarily then all may be, and the form of discourse as a public institution would be impossible. This does not mean that other forms of discourse may not be constructed in which certain possible interlocutors are excluded by fiat, but this would not then be the fully rational, autonomous, and objective form of discourse we require. A parallel may be found in scientific discourse. As far as it is objective, considerations must be dealt with on their merits and not in terms of the will or status of whoever puts them forward. If any arbitrary exclusion of possible interlocutors is made, then we do not have public objective scientific discourse but a sort of game in which arbitrarily selected players alone are entitled to make certain moves and in which what is determined in the outcome is who has won rather than what is true.

If moral discourse is to be public and objective, then it must allow for the participation of any possible rational interlocutor. Now let us define an interest as that which any rational being should seek for himself insofar as he considers the effects of his actions on himself and not on others except insofar as what affects others also affects him (for example, if it is rational for anyone to avoid pain, then it is in my interest to seek those actions that avoid pain to myself but not necessarily those that avoid pain to others except insofar as the pain of others causes pain to me or prevents my achieving some other end that it would be rational for me to choose for myself). Now it is by definition necessary that every rational being should seek his own interests as far as possible. It would be irrational for any being to participate in a form of discourse the practical effect of which would be to deny his interests; hence, it would be irrational for anyone to adopt moral discourse without further justification if from the beginning his interests were to be ruled out. But this means that anyone who wishes to adopt moral discourse must allow that any possible interlocutor must not have his interests ruled out of consideration from the beginning, and any rational being is in principle a possible interlocutor. It follows that as far as public objective moral discourse is to be possible, it is presupposed that what is determined by such means will not neglect the interests of any rational being—that is, that in deciding what I ought to do, or what anyone ought to do, the interests of all rational beings whatsoever must be taken into account.

LIBERTY. The principle of liberty is that one ought not to interfere, without special justification, in the chosen course of any rational being or impose on any rational being conditions that will prevent him from pursuing his chosen courses of action. Moral discourse is a form of discourse in which we try to guide action rationally. We try to determine action on the basis of a rational consid-

eration of the nature of the action and its context, not by some other means such as violence. Any interference with the chosen course of a rational being is a determination of his action by force or at least a limit imposed by force on the extent to which his actions may be rationally determined. Such interference must then be presupposed as absent in public objective practical discourse in which action is determined by reason, and hence in using such discourse, in participating in it as an institution, one is presupposing that one ought not to interfere by force, but only by rational persuasion, in the chosen course of any rational being.

The arguments given for these three principles are very much oversimplified, and it could not be claimed that they have the force of demonstrations. But enough has been said to show that the type of argument they represent is at least a possible one and hence that the apparent conflict between autonomy and objectivity is not a real one and that the problem of the justification of ultimate principles may not be insoluble.

COMPLETENESS AND APPLICATION OF PRINCIPLES. Two important problems remain. The first we may deal with briefly. It is one that was very important to Kant, with regard to both theoretical and practical principles. How can we be sure that we have achieved completeness in any list of principles? If ultimate principles can be established only by transcendental arguments, we have at least some clue to the answer to this problem; for the rest it might be argued that the problem is not so urgent as some have thought.

A transcendental argument is one that depends on an account of what is necessary to a given form of discourse; in ethics we are concerned with what is necessary to a form of discourse that is practical, universal, objective, and autonomous. We are, that is, dependent on a consideration of the formal characteristics of the form of discourse. This gives at least some negative criterion for deciding what principles may be justified as ultimate. Thus, it would be most implausible to suggest, for example, that "One ought not to drink alcoholic liquor on Sundays" could be justified as an ultimate moral principle. For it is reasonably obvious that no direct connection could be established between the purely formal characteristics of any form of discourse and such particular matters as are picked out by the concepts of the principle in question. Such a principle would have to be, if justifiable at all, one that would depend on matters beyond the purely formal characteristics of practical reason. It is always possible, however, though in this case surely a fantastic suggestion, that someone with sufficient ingenuity might show that some apparently low-level principle is in fact justifiable as an ultimate one by a transcendental argument. And this may disturb us, for how can we be sure that we are not failing to take account of such principles all the time? We should not, however, be much disturbed, for two reasons. First, if a principle is a necessary condition of the possibility of moral discourse, one would expect to find it as a pervasive explicit or implicit principle of most moral codes (allowing for the resources of human confusion), and this is true for the three principles—justice, benevolence, and liberty—we have mentioned. Second, when it is suggested that there is a reason for acting in one way rather than another, the suggestion requires justification, in the absence of which the suggestion may be reasonably ignored. The onus of proof is on anyone who suggests that a certain principle is correct; until such proof is at least suggested, the fear that there may be quite unknown principles, which are not generally accepted but which could, with sufficient ingenuity, be justified transcendentally, is an idle one.

The second difficulty that we face at this point is of the utmost importance; indeed, one might fairly say that out of it all the really important and difficult questions of substance in ethics arise. It is the problem of the application of these principles to particular situations, both in themselves and in relation to one another. Unless it is possible to show that these principles can be rationally applied, then no amount of rational demonstration of the ultimate principles will enable us to show that the particular moral judgments we make can be rationally justified.

In this article it has been argued that any account of how particular judgments about what ought to be done can be justified will need to examine principles that are necessary but not sufficient to justify particular judgments. These principles will pick out factors of general moral relevance, and the principles in turn will require justification. This may then require reference to more general principles, but some principles that are incapable of further justification will be reached in this way, and these we have called ultimate principles. It would seem that ultimate principles could never be justified objectively, but it is suggested that arguments that show them to be necessary if objective practical discourse is to be possible would justify them and that such arguments are possible. It is, however, emphasized that since ultimate principles are necessary but not sufficient to the justification of particular judgments, we have not by this suggestion solved the whole problem of how ethical disagreement can be rationally resolved. We have only

removed one ground for saying that they can never be rationally resolved.

See also Bentham, Jeremy; Ethical Objectivism; Impartiality; Intuition; Kant, Immanuel; Liberty; Mill, John Stuart; Moral Rules and Principles; Moral Sense; Rationality; Sartre, Jean-Paul; Sidgwick, Henry; Value and Valuation.

Bibliography

For an account of objectivity and autonomy, and an attempt to justify certain factors as ethically relevant from a consideration of the formal character of practical reason, see Immanuel Kant, *The Moral Law: Kant's Groundwork of the Metaphysic of Morals,* translated by H. J. Paton (London, 1964).

Attempts to justify ultimate principles are to be found in Jeremy Bentham's *Introduction to the Principles of Morals and Legislation* in *A Fragment on Government, and an Introduction to the Principles of Morals and Legislation,* edited by W. Harrison (Oxford, 1948); J. S. Mill's *Utilitarianism* in *Utilitarianism, Liberty, and Representative Government,* No. 482A in Everyman's Library; Henry Sidgwick's *Methods of Ethics* (London, 1907; reprinted 1962).

More recent attempts to justify some factors as of ultimate ethical relevance are to be found in Kurt Baier's *The Moral Point of View: A Rational Basis for Ethics* (Ithaca, NY: Cornell University Press, 1958), especially Chs. 7–8, and Baier's two articles in *Philosophical Studies,* 4 (1953)—"Good Reasons," 1–15, and "Proving a Moral Judgment," 33–44; J. N. Findlay's *Values and Intentions* (London: Allen and Unwin, 1961), his "Morality by Convention," in *Mind,* n.s., 53 (1944): 142–169, and his "Justification of Attitudes," in *Mind,* n.s., 63 (1954): 145–161; Philippa Foot's "Moral Arguments," in *Mind,* n.s., 67 (1958): 502–513, and her "Moral Beliefs," in *PAS,* n.s., 59 (1958–1959): 85–104; A. P. Griffiths's "Justifying Moral Principles," in *PAS,* n.s., 58 (1957–1958): 103–124; D. L. Pole's *The Conditions of Rational Inquiry* (London: University of London, Athlone Press, 1961), especially Ch. V; M. G. Singer's *Generalisation in Ethics* (London, 1963); S. E. Toulmin, *An Examination of the Place of Reason in Ethics* (Cambridge, U.K.: Cambridge University Press, 1950; reprinted 1958).

A. Phillips Griffiths (1967)

MORAL PSYCHOLOGY

Moral psychology is the area of scholarship that investigates the nature of psychological states that are associated with morality—states such as intentions, motives, the will, reason, moral emotions (such as guilt and shame), and moral beliefs and attitudes. The purview of moral psychology also includes associated concepts of virtue, character trait, and autonomy. It has generally been thought of as a descriptive enterprise rather than a normative one, though this is not always the case.

Traditionally we can see two different approaches to moral psychology. The first is the *a priori* approach to understanding moral psychology and the significance and function of psychological states. The second is the empirical approach that considers the evidence of their significance, function, and development. Both of these strands will take as their starting point common sense intuitions about how people think about morality, make moral decisions, and the circumstances under which they feel moral emotions. These intuitions may be based on a long history of observation of human behavior, or they may simply be the result of natural selection leading to similarity in thought which itself might be adaptive. Either way, common sense provides the baseline for research in moral psychology.

The *a priori* strand engages in conceptual analysis of the relevant psychological states and their connections. There is a debate, for example, about whether reason alone can motivate, or not. What *explains* our actions? Is it the case that when I give money to charity I do so simply because I believe it will help people who need help, or do I also need to desire to help them? This will engage us in a discussion of the distinction between belief and desire. A view, which can be traced back at least as far as David Hume, holds that beliefs are of matters of fact and can be true or false; desires, on the other hand, have no truth-value. And, it is desires that are essentially motivating.

Thus, whenever one wants to fully explain an action one needs to be able to identify the belief/desire combination that gives rise to it. But this seems to present a puzzle for moral action: often, morality requires us to act against our desires. I am required to keep my promises, even if I don't want to. But how can I keep my promises if I don't want to, when desire is necessary for action? So there is also a *normative* question that can be raised. Presumably I am giving money to charity because I think that it is a good thing to do. I accept the norms of giving. So, is it the case that if I think that giving to charity is good, it *necessarily* provides me a motivating reason for giving? Is there a necessary connection, or conceptual tie, between the normative reason (the recognition that giving is good) and my motivation to perform the action of giving? If I think there is, then I am an "internalist"; if, however, I do not believe that there is a necessary connection, then I am an "externalist." The acceptance of the norm, the recognition that giving to charity is a good thing, will then necessarily mean that I have at least a

weak desire to act on the reason. This could, of course, be defeasible.

But there are those who disagree. Those who are externalists, such as David Brink, argue that amoralists can recognize moral reasons—for example, the amoralist can recognize that it is good to give to charity—yet utterly fail to be moved by this recognition. Indeed, that is what it is to be an amoralist. They are defective not because they fail to see moral reasons as moral, but precisely because they recognize them and yet fail to be moved by them at all. Internalists argue that amoralists, when they articulate a belief that "*x* is good" and then fail to be moved, do not *really* believe what they have articulated. They are trying to make moral judgments, but they are failing to actually do so. Michael Smith also allows that such agents may be practically irrational.

A related feature of Hume's view of moral psychology is its commitment to the view that desire is a given. That is, one cannot reason oneself into a *basic* desire. One can reason about non-basic desires—for example, perhaps I would like to eat ice cream today. Then someone points out that ice cream really isn't very healthy. Since I would rather eat healthy food, I now no longer desire to eat the ice cream. But the desire to eat the ice cream is not basic. Rather, I would like to feel good—and once someone points out to me that a habit of eating ice cream will make me less likely to keep feeling good in the long run, that desire to eat ice cream falls away. But I have not been reasoned out of the basic desire. Indeed, it is its conflict with this desire that makes me ready to jettison the other.

But other writers disagree with this Humean conception of desire, and the reason/desire dichotomy. They believe that we can rationally reflect on basic desires and come to change them through the force of this rational reflection alone. For example, one might argue that desires, even some fundamental ones, are based in part on beliefs that we have. If I desire, for example, to avoid treating persons as means, it may be that I have this desire because I think that being respectful toward others requires this, and I believe, with good reason, that respecting others is obligatory. This desire could be basic in that it cannot be reduced to another desire. If this case is plausible, then we have a basic desire supported by reason.

One way to view this case is as that of a commitment one has. The desire to avoid treating others disrespectfully is more than just a strong basic desire that I have, which happens to be stronger than the other desires I have that might conflict with it. It is a commitment, a normative commitment that I have, and I have it for reasons that are motivating reasons. These reasons carry the desire to be respectful of others with them. Further, there are reasons for this desire having to do with my beliefs about, perhaps, what it is to be a flourishing human being. Presumably I could be argued out of the desire, then. A Humean might try to respond to this, however, by pointing out that any "argument" one would give would in turn depend upon some stronger desire for its force. Desires are not themselves true or false, but they can loosely be considered irrational if based on false beliefs. Beliefs exposed as false would then presumably lead to an alteration of the desire one had based on that belief. In the example that I cited above, then, the Humean would probably say that my desire to be respectful of others is based on the belief that this is good and obligatory—so that simply shows that I have a more basic desire to live up to my obligations.

The field of moral psychology also has a more empirical side. Aristotle believed that the observation of human beings could reveal to us what, for human beings, was *eudaimonia*. Thomas Hobbes believed that an astute observer of human nature would find support for psychological egoism. Charles Darwin believed that natural selection could account for the sorts of emotions that human beings feel, including the moral emotions. Data that psychologists have gathered about human behavior have influenced the way some think about morality. For example, the work of psychologist Carol Gilligan raised the issue of gender differences in approaches to thinking about moral problems, which in turn influenced writers in feminist ethics.

More recently, empirical psychological research has been brought into moral theory to shed light on a host of issues, ranging from the issue of what, exactly, goes on in a person's brain when she thinks about moral issues, to the issue of the innateness of our moral cognition, to the seemingly basic commitment human beings have to moral objectivity. There is also the extremely interesting and important issue of how natural selection has shaped our sense of morality and moral practices, as well as our moral intuitions. For example, Jesse Prinz has done work in comparative psychology that offers evidence against moral nativism. He believes that the evidence best supports the view that there is not even a minimal innate moral competence—instead it is culture that guides the formation of our moral capacities.

The work of Shaun Nichols draws on literature in developmental psychology to investigate the claim, widely argued in meta-ethics, that people are generally moral objectivists. That is, that people accept the view

that there are some true moral judgments, and when a moral judgment is true, it is non-relativistically true. Nichols points out that experiments in developmental psychology, though not at this point in time conclusive, point to the view that for persons, generally, moral objectivism is the "default position" when it comes to commonsense, or lay, meta-ethics.

There is also a trend in moral philosophy of exploring the significance of emotion in moral judgment. This has a counterpart in the psychological research. Joshua Greene and Jonathan Haidt refer to this as the "affective revolution." The interest in this area of psychological research was sparked by Antonio Damasio's work showing that good reasoners needed affect. When portions of the brain that regulate affect are damaged, agents do not perform very well on follow-through in practical reasoning tasks. The classic case, discussed by Damasio, is Phineas Gage. Gage was a railway worker who suffered damage to his frontal lobe in an accident in 1848. This caused an apparently extreme personality change that involved inappropriate emotional responses and a disposition to impulsive behavior. He became unreliable and untrustworthy. He was able to reason in the abstract but was not able to carry through. Affect thus at least seems crucial to effective moral motivation. This conclusion was supported by studies involving more recent cases of frontal lobe damage.

Greene's own work explores brain activity when persons consider moral dilemmas. He and his colleagues discovered that when personal dilemmas were presented to subjects—that is, situations in which those being harmed are close to the subject—there is far more brain activity in the emotional areas of the brain, and those areas of the brain underlying social cognition, than when the problem cases were impersonal. We do seem moved to help in personal cases to a greater extent than impersonal cases. This research supports what charitable organizations have long realized. To promote giving there is a need to make the plight of the suffering personal to potential givers—through photographs and letters, for example. Of course, this leaves untouched the question of what people ought to do. While it is true that our emotions are engaged more in these personal situations, that has no implications for what our obligations are in these cases. This is where we need normative ethics.

Still, this line of research supports the descriptive view that when we behave morally, or at least think about moral issues, in a way that has more motivating force, there is considerable engagement of our affective capacities. Further, when those affective capacities are impaired, we are left with agents whom we would describe as morally defective. Phineas Gage was widely considered to be a deadbeat after his accident. That is a moral judgment of his character, and the appropriateness of that judgment has something to do with the fact that he lacked the correct emotional responses, those appropriate for the circumstances in which he found himself.

Empirical psychological research has also influenced literature on virtue ethics. Virtue ethics is a type of normative ethical theory that bases moral evaluation on virtue concepts. The approach has been attacked for its failure to reflect psychological reality. For example, Gilbert Harman's work on virtues makes use of situationist literature in social psychology. He argues, citing situationist experiments, that there are no character traits. Rather, the best explanation for a person's behavior is his situation—so, if one would like a reliable way to predict behavior, one needs simply to look at the person's situation. Persons who are in a hurry will be less likely to help than persons who are not. Persons who smell fresh cookies baking are more likely to act benevolently than those who are not smelling the cookies, and so forth.

Thus, character traits need not be cited at all in reliable predictions or explanations. There is no reason to think they exist. Further, if there are no character traits, then there are no character traits that are virtues. It would follow then that virtue ethics is a non-viable normative ethical theory, since it assumes what does not in fact exist. There are no stable character traits, at least, no stable and robust moral character traits. John Doris has softened Harman's claim somewhat, also by bringing in evidence from empirical psychology. On Doris's view all that is warranted by the empirical data is the view that character traits are not "global"—that they are more narrowly prescribed and local than intuition would have it. Thus, there may not be a general robust trait of benevolence, but there may be a trait of "benevolence when one smells cookies" and "benevolence when one is not in a hurry," and so forth. Doris still views even this weaker position as a threat to virtue ethics since it cuts against the assumption that there are robust, global character traits. A virtue ethicist is free to respond that even if Doris is correct, virtue ethics may still offer a regulative ideal. After all, it is a theory of how we ought to be, not how we are.

Assuming, with common sense intuitions, that there are character traits that qualify as virtues, is there any particular psychology that characterizes moral virtue? Here we move away from use of evidence from experimental psychology and back to philosophical analysis of normative concepts that is, nevertheless, sensitive to our views of

psychological reality. In my own work I argue there is no special psychology that characterizes moral virtue, and that what counts as a moral virtue is characterized by externalities such as the consequences that the traits systematically produce. Other writers, such as Rosalind Hursthouse, disagree. Taking Aristotle as her inspiration, she holds that virtue states require that the agent have certain psychological states, such as a kind of practical wisdom that is needed for deliberating well about what to do—presumably, one needs to deliberate well in order to be a good person. Another writer who has attacked this moral psychology of the virtues is Nomy Arpaly, who argues that all that is needed is that the agent be responsive to the right sorts of reasons.

It is true that one thing that we hold people responsible for is their failure to be responsive to the right sorts of reasons. If one observes an agent acting with a callous disregard for the well-being of others, this can give rise to feelings of outrage. Thus, these failures of appropriate responsiveness can generate moral emotions that are indicative of our moral commitments. For example, we have a commitment to a norm of honesty. This norm is important to regulating our social interactions. In a person of reasonably good character, a failure to be honest will lead to feelings of remorse. Also, in a person of reasonably good character, seeing another behave dishonestly will give rise to a reactive attitude of outrage or resentment. When such feelings are appropriately felt, this may serve as good evidence that there has been a moral failure.

Reactive attitudes, then, can figure into accounts of moral responsibility and moral accountability. R. Jay Wallace, for example, has developed an account of what it is to hold someone responsible, morally—it is an attitudinal stance toward someone, a third-person stance that crucially involves reactive attitudes. If one holds someone responsible for having done something bad, then it is appropriate to feel something like resentment toward that person. Note that this is not a descriptive claim. It is true that normal persons do feel resentment under these circumstances. It is also the case that this indignation or resentment is *appropriate* when one has been wronged. Thus, though there is some disagreement over this, the sphere of moral psychology does involve an investigation of some normative issues having to do with the normative status of some of the mental states and character traits central to moral evaluation.

See also Egoism and Altriusm; Human Nature; Moral Motivation; Moral Sentiments; Sympathy and Empathy; Virtue and Vice; Virtue Ethics

Bibliography

Arpaly, Nomy. *Unprincipled Virtue*. Oxford and New York: Oxford University Press, 2003.

Damasio, Antonio. *Descartes' Error*. New York: Putnam, 1994.

Darwin, Charles. *The Descent of Man* (1888). Amherst, NY: Prometheus, 1998.

Doris, John. *Lack of Character*. Cambridge and New York: Cambridge University Press, 2002.

Driver, Julia. *Uneasy Virtue*. Cambridge and New York: Cambridge University Press, 2001.

Gilligan, Carol. *In a Different Voice*. Cambridge, MA: Harvard University Press, 1982.

Greene, J. D., et. al. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293 (2001): 2105–2108.

Greene, Joshua, and Jonathan Haidt. "How (and Where) Does Moral Judgment Work?" *TRENDS in Cognitive Sciences* 6 (December 2002): 517–523.

Harman, Gilbert. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99 (1999): 315–331.

Hume, David. *A Treatise of Human Nature*, edited by L. A. Selby-Bigge (1896), revised by P. H. Nidditch. Oxford: Clarendon Press, 1978.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford and New York: Oxford University Press, 1999.

Nichols, Shaun. *Sentimental Rules*. Oxford and New York: Oxford University Press, 2004.

Prinz, Jesse. "Against Moral Nativism." Manuscript.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.

Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.

Julia Driver (2005)

MORAL REALISM

Moral realism is a metaethical view committed to robust objectivity in ethics. No single description is likely to capture all realist views, but a reasonably accurate rule is to understand moral realism as the conjunction of three theses:

> The semantic thesis: The primary semantic role of moral predicates (such as "right" and "wrong") is to refer to moral properties (such as rightness and wrongness), so that moral statements (such as "honesty is good" and "slavery is unjust") purport to represent moral facts, and express propositions that are true or false (or approximately true, largely false, and so on).
>
> The alethic thesis: Some moral propositions are in fact true.
>
> The metaphysical thesis: Moral propositions are true when actions and other objects of moral assessment

have the relevant moral properties (so that the relevant moral facts obtain), where these facts and properties are robust: their metaphysical status, whatever it is, is not relevantly different from that of (certain types of) ordinary non-moral facts and properties.

To deny any one of these three theses is to embrace some form of moral irrealism. Many philosophers consider moral realism the default position because it appears best to capture many central features of ordinary moral thought: the assertoric surface character of ordinary moral discourse, the phenomenology of moral experience, our claim to have moral knowledge, and the possibility (and nature) of genuine moral error, progress, and disagreement even among sincere, open-minded, and well-informed people (Dancy 1986, Brink 1989, Shafer-Landau 2003).

The semantic thesis is (for better or worse) often associated with the related psychological thesis called cognitivism, according to which the primary role of moral judgments is to express beliefs. One form of irrealism, non-cognitivism, holds that their primary role is to express motivational "non-cognitive" states of mind, such as approving, prescribing, commending, or planning, but can assign moral predicates and judgments a secondary role of referring to (non-moral) properties and expressing (non-moral) beliefs (Copp 2001). How well realists can explain the reliable connection between moral judgment and moral motivation is a matter of some dispute (Smith 1994).

The alethic thesis says that some moral propositions are robustly true only if we combine it with the realist's metaphysical thesis. The irrealists' attitude to the alethic thesis depends on their conception of truth. Error theory accepts a robust reading of the semantic thesis but rejects the alethic thesis on this robust reading. It holds that ordinary moral thought presupposes the existence of robust moral facts and properties but is systematically in error: every moral judgment with existential import is mistaken because there are no robust moral facts to make any such judgment true (Mackie 1977). Non-cognitivist irrealists can accept a non-robust reading of the alethic thesis if they endorse minimalism about truth (but see Dreier 2004). This move may eventually earn them the right to speak of moral facts and truths, and to say all the same things that any morally decent person would say about what is right or wrong, good or bad, just or unjust, and so on, even though they reject the realist's metaphysical thesis (Blackburn 1993).

The metaphysical thesis is central to moral realism because realism is primarily a view about metaphysics, not about truth or semantics. It holds that moral facts and properties are not metaphysically inferior in kind to many ordinary sorts of non-moral facts and properties. What is it for a fact or property to be metaphysically robust, though? One sense in which ordinary non-moral properties are robust is that they enter into explanations of real phenomena; water has its surface properties because it is H_2O, for example. In this sense, the realist's metaphysical thesis says that moral properties enter into explanations of phenomena that irrealists would explain by other means (Dreier 2004). An irrealist might take the fact that one believes that inequality is unjust to consist in some such fact as that one has decided to include the reduction of inequality among one's aims. A realist might instead say it consists in standing in a certain belief-like relation to the properties of inequality and injustice. Likewise, the realist might say, whether such a belief is correct or mistaken is just a matter of whether the two properties are related as the belief represents them as being related. The realist's explanation of the assertoric features of ordinary moral discourse, possession of moral knowledge, and nature of moral disagreement would be analogous.

Understanding the metaphysical thesis as above affords one (albeit not the only) way of capturing many realists' conviction that ethics concerns objective matters of fact whose existence and nature are independent of anyone's sentiments, opinions, evidence, or theories about what is right or wrong, obligatory, permissible, or impermissible, good or bad, and so on. So understood, the thesis also classifies as irrealist any view according to which explanations of moral phenomena involve no essential reference to moral facts or properties, but only to such factors as our individual tastes, cultural or social conventions and agreements, basic human sentiments, or the beliefs or plans we would have if we were fully informed and rational. Thus ethical subjectivism, ethical relativism, projectivism, and most forms of constructivism in ethics rightly count as irrealist even though they accept the realist's semantic and alethic theses.

Disputes within the realist camp concern primarily the nature of moral facts and properties. Non-naturalist realists hold that moral properties are robust properties that are distinct from but supervene (see below) on natural properties (Moore 1903, Shafer-Landau 2003). Naturalist realists hold that moral properties are robust natural properties. Reductive naturalists hold that moral properties are identical to natural properties that we can represent in austerely non-moral terms (Railton 1986). Non-reductive naturalists hold that moral properties are an irreducible subclass of the class of natural properties,

which we may be unable to represent in austerely non-moral terms (Boyd 1988, Brink 1989).

Arguments for and against different forms of moral realism differ also depending on whether we take true statements of property identity to be analytic (true in virtue of the meanings of their constituent terms) or synthetic, and what we think qualifies a property as natural. If, for example, natural properties are just those that we can investigate empirically, then naturalism will hold that knowledge of any synthetic moral proposition is answerable to empirical evidence, whereas non-naturalism will hold that knowledge of some synthetic moral propositions is empirically indefeasible (Copp 2003, Shafer-Landau 2003). An issue for synthetic naturalists in particular is what determines the reference of moral predicates to the supposedly natural moral properties. Given their view of the matter, can they explain the intelligibility of such "open questions" as whether something that satisfies a given naturalistic non-moral predicate (such as "is pleasant") also satisfies a given moral predicate (such as "is good") (Moore 1903, Horgan and Timmons 1992)?

According to the supervenience argument against moral realism, we can distinguish between a weaker, true claim and a stronger, false claim about the supervenience of the moral on the natural. (*Supervenience* is a technical name for a relation of necessary covariance.) The alleged problem for the realist is that she cannot, but the irrealist can, explain why the weaker supervenience claim should be true, given that the stronger claim is false (Blackburn 1993). According to one clear version of the argument (Dreier 1992), the true claim is that it is analytically necessary that, for each moral property *M* that an object *O* has, there is a (possibly complex) natural property *N* that *O* has, and it is metaphysically necessary that *M* always accompanies *N*. The stronger, putatively false claim differs in saying that *M* always accompanies *N* as a matter of analytic necessity. (Variations of the argument concern predicates rather than properties and involve different types of necessity.) The objection is that if realists are committed to the thesis called "lack of entailment," according to which no set of non-moral naturalistic truths entails any particular moral truth, then they must admit (falsely) that it is possible for *M* sometimes not to accompany *N*.

Different forms of moral realism respond differently to the supervenience argument. Analytic naturalists may regard the argument as question-begging, for they deny that the stronger supervenience claim is false (Jackson 1998). Non-naturalists may accept a lack of analytical entailment but claim that duly specified sets of naturalistic truths metaphysically entail particular moral truths because the facts which the former concern exhaustively constitute (in some sense to be explained) the facts which the latter concern (Shafer-Landau 2003). Some synthetic naturalists may say that their theory explains why the weaker supervenience claim is true (since moral properties are natural ones), but entails that no set of non-moral naturalistic truths analytically entails any particular moral truth (since any connection between non-moral and moral truths is synthetic). Others may express doubts as to whether the relevant supervenience claims are formulated so as to make them both interesting and acceptable to synthetic naturalists.

According to the explanatory argument against moral realism, properties of a certain kind are metaphysically robust only if they make a distinctive contribution to our overall explanatory picture of the world (the "explanatory requirement"), but moral properties make no such contribution; therefore, moral properties are not metaphysically robust. A prominent version of this argument claims that mentioning moral properties such as wrongness makes no distinctive contribution to causal explanations of such occurrences as a person's indignation or her judgment "that's wrong" upon seeing some hoodlums set a cat on fire, above and beyond the contribution of mentioning the person's prior beliefs, aversions, and moral principles (Harman 1977). If so, the causal version of the explanatory requirement gives us good reason to deny that there are robust moral facts.

One realist response is to argue that the causal requirement is dubious; for if it is, then it would be no objection to moral realism if moral properties violated that requirement (Shafer-Landau 2003). Moral properties could still play non-causal explanatory roles. A very different response is to accept the causal requirement, but argue that mentioning moral properties can make a distinctive contribution to causal explanations of both intentional occurrences, such as moral judgments, and non-intentional ones (Sturgeon 1988, Brink 1989). On the latter score, one may argue that a person's kindness can cause her to help others or that injustice or oppression can provoke resistance, and that these properties can play such causal-explanatory roles only if they are real, and indeed natural, properties. Here the intricate question arises whether moral properties are epiphenomenal, in that they play no causal-explanatory role over and above the causal-explanatory role of the non-moral properties on which they supervene, or by which they are realized (Miller 2003, Sturgeon 2005).

What unites these debates about moral realism is the concern whether, and how, robust moral facts and properties enter into accounts of various phenomena that irrealists would explain by other means. One general moral may be that arguments in metaethics often are arguments about the best explanation of the phenomena in question. Other important debates between realists and irrealists and within the realist camp concern the rational authority of morality, the extent to which moral realism affords a rational basis for resolving moral disagreements, the existence of an internal connection between moral judgment and moral motivation and whether such "motivational internalism" would make moral properties metaphysically strange, and questions about moral methodology and moral epistemology, such as the place of ethics in a naturalistic worldview and the parity or continuity of ethics with empirical inquiry and the sciences.

See also Ethical Naturalism; Internalism and Externalism in Ethics; Intuitionism, Ethical; Metaethics; Moral Epistemology; Noncognitivism; Objectivity in Ethics; Rationalism in Ethics (Practical Reason Approaches).

Bibliography

Blackburn, Simon. *Essays in Quasi-Realism*. Oxford: Oxford University Press, 1993.

Boyd, Richard. "How to Be a Moral Realist." In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord. Ithaca, NY: Cornell University Press, 1988.

Brink, David O. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Copp, David. "Realist-Expressivism: A Neglected Option for Moral Realism." In *Moral Knowledge*, edited by Ellen Frankel Paul, Fred D. Miller Jr., and Jeffrey Paul. Cambridge, U.K.: Cambridge University Press, 2001.

Copp, David. "Why Naturalism?" *Ethical Theory and Moral Practice* 6 (2003): 179–200.

Dancy, Jonathan. "Two Conceptions of Moral Realism." *Proceedings of the Aristotelian Society*, suppl. 60 (1986): 167–187.

Dreier, James. "The Supervenience Argument against Moral Realism." *Southern Journal of Philosophy* 30 (1992): 13–38.

Dreier, James. "Meta-ethics and the Problem of Creeping Minimalism." *Philosophical Perspectives* 18 (2004): 23–44.

Harman, Gilbert. *The Nature of Morality*. Oxford: Oxford University Press, 1977.

Horgan, Terence, and Mark Timmons. "Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived." *Philosophical Papers* 21 (1992): 153–175.

Jackson, Frank. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

Miller, Alexander. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press, 2003.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Railton, Peter. "Moral Realism." *The Philosophical Review* 95 (1986): 163–207.

Shafer-Landau, Russ. *Moral Realism: A Defence*. Oxford: Oxford University Press, 2003.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.

Sturgeon, Nicholas L. "Moral Explanations." In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord. Ithaca, NY: Cornell University Press, 1988.

Sturgeon, Nicholas L. "Moral Explanations Defended." In *Contemporary Debates in Moral Philosophy*, edited by James Dreier. Oxford: Blackwell, 2005.

Sturgeon, Nicholas L. "What Difference Does It Make Whether Moral Realism Is True?" *Southern Journal of Philosophy*, suppl. 24 (1986): 115–142.

Wright, Crispin. "Realism, Antirealism, Irrealism, Quasi-Realism." *Midwest Studies in Philosophy* 12 (1988): 25–49.

Pekka Väyrynen (2005)

MORAL RESPONSIBILITY

See *Responsibility, Moral and Legal*

MORAL RULES AND PRINCIPLES

Normative rules and principles say what things are required or permitted or good or bad. In other words, normative rules and principles say what agents ought to do or what agents are allowed to do; or what deserves to be promoted, praised, or approved; or what deserves to be opposed, criticized, or disapproved. Moral rules or principles differ from normative ones of other kinds (such as rules or principles of law, etiquette, or clubs) in that moral rules or principles indicate what agents morally ought to do or are morally allowed to do, or what deserves moral praise and admiration.

Rules and principles are (to at least some extent) general—that is, they are about kinds of situations or about classes of cases, not about individual instances. So rules or principles are juxtaposed with judgments about a particular instance. The judgment that Martin Elginbrodde ought to feed his hamster at 8 a.m. on July 7, 2007, does not articulate a rule. Rather, it articulates a judgment about what a particular person should do on a particular occasion. Because rules and principles are about kinds of situations or classes of cases, rules or principles entail judgments about particular instances. The

principle that people ought to feed their pets entails that Miguel ought to feed his cat, that Janet ought to feed her dog, that Rahul ought to feed his bird, that Jo ought to feed his ferret, and so on for as many pet owners as there are.

Many philosophers have held that moral rules and principles must apply universally. What it is right for one person to do must be right for anyone else to do unless there is some morally relevant difference between the cases. This thought is reflected in the Golden Rule and serves as a cornerstone of the moral philosophies of Immanuel Kant and Richard M. Hare. But one important difference between Kant and Hare concerns the degree of detail and complexity they allow into moral principles. Kant thought moral principles had to be quite simple; Hare thought they could be highly detailed and complex as long as they were formulated in completely universal terms.

How stringent are moral rules and principles? Most people must take moral rules and principles to be very important—in particular, to generate very strong reasons for action. Otherwise, the degree of social cooperation and solidarity that moral rules and principles are supposed to provide is unlikely to be achieved. Some philosophers—for example, Ronald Dworkin (1977)—have held that moral rules can be more specific and less stringent than moral principles. A moral rule might be: "Be especially kind to your parents." A more general and stringent principle might be: "Be especially kind to your benefactors." In a case where a parent has not been a benefactor, for example, a father who always ignores the plight of his offspring, the rule "Be especially kind to your parents" might fade to nothing.

Admittedly, even the rule "Be especially kind to your benefactors" can be overridden. To take an extreme example, being kind to benefactors might conflict in some situation with saving many innocent lives. Suppose that for some reason one can either go to thank benefactors or devote the time to saving innocent lives, but not both of these things. With respect to such a case, the principle "Be especially kind to benefactors" seems morally less important than the principle "Prevent harm to others." Many other moral rules or principles are likewise capable of being outweighed or overridden in certain cases by other moral rules or principles.

Are there any rules or principles that always outweigh any opposing moral considerations? Consider the principles "Do not do what is morally wrong" and "Do what you morally ought to do." Such principles concern compliance with all-things-considered moral verdicts. These principles tell us to do whatever is, all things considered, morally required. They give us no indication which moral considerations win out over others to generate all-things-considered moral verdicts.

Are there any rules or principles that both provide information about what morality requires and always outweigh any opposing moral considerations? Two kinds of principles have been suggested. One of these kinds consists of moral principles outlawing evil purposes, such as "Do not, for its own sake, harm others" and "Do not, for its own sake, deceive others." The other kind consists of principles offered as the most general and basic principle of morality, such as Kant's "Act only on maxims that you can will to be universal laws" and the act-utilitarian's "Do whatever acts promote aggregate well-being."

There are other moral theories that put forward other foundational principles. For example, T. M. Scanlon's (1982) contractualist theory of morality claims that moral wrongness is determined by rules for the general regulation of behavior that no one could reasonably reject as the basis of informed, unforced, general agreement. Richard B. Brandt's (1967) rule-utilitarian theory holds that moral wrongness is determined by rules that have the highest expected impartial utility. Rosalind Hursthouse's (1999) virtue ethics holds that an act is wrong if it is one that would not be done by someone with a full set of the character traits that benefit others or the agent.

Some philosophers think that the theories just mentioned are mistaken to claim that morality is so unified. For example, pluralists such as William David Ross (1930) think that there is a plurality of basic moral principles that identify the features that count morally in favor of actions that have them (moral pros) and other features that count morally against the actions that have them (moral cons). These moral pros and cons are the appropriate inputs to moral assessment; a verdict about all-things-considered moral rightness or wrongness is the appropriate output. Rossian pluralists think that these moral principles (and thus the moral pros and cons that the principles identify) can conflict. For example, the fact that an act would benefit others counts in its favor, and the fact that an act would keep one's promise counts in its favor. Sometimes, however, keeping one's promise is not what would benefit others.

Rossian pluralists also think that the principles do not come in a strict hierarchy of importance that would resolve all the possible conflicts among them. This presents the question of what is the right thing to do when the Rossian principles conflict. Rossian pluralists hold that

which principle wins when there is conflict among them cannot be captured in a correct, informative, general principle. For example, a general principle that benefiting others always trumps keeping promises is not correct. Neither is a general principle that keeping promises always trumps benefiting others. Instead, in some situations it is right to keep a promise though one could benefit others more if one broke the promise, but in other situations it is right to break a promise if this is necessary in order to benefit others. So Rossian pluralists admit that moral verdicts about right and wrong cannot be systematized in correct informative general principles. They maintain that, when basic principles conflict, the right thing to do is a matter of judgment rather than a further principle. Still, Rossian pluralists think that moral principles have an important place, namely, in identifying the moral pros and cons.

Some philosophers think even principles about what counts as a moral pro or a moral con are incorrect. These philosophers are called moral particularists. Particularists hold that, for any feature of an action or its consequence that is a moral pro in one situation, that same feature might be a moral con in another situation. Whereas Rossians think that the fact that an act would benefit someone is always a reason in favor of the act, particularists think that, in some situations, the fact that an act would benefit someone is morally positive but in other situations it is morally negative. Wiping sweat from a torturer's brow, for example, would benefit the torturer but would not count in favor of the action. More generally, particularists maintain that features of actions can switch moral "polarity," depending on the context. Most will agree that one should try to help the person being tortured rather than wiping the torturer's brow. The question is how to explain what the inputs to that verdict are. Particularists say that the fact that wiping the torturer's brow would benefit him is no reason to do it, but rather, a reason against doing it.

On this issue, antiparticularists divide into two groups. Antiparticularists in one group say that the potential benefit to the torturer is massively outweighed by the importance of trying to help the person being tortured. But antiparticularists in this group hold that the fact that wiping the torturer's brow would benefit him counts at least a little bit in favor of wiping his brow. Antiparticularists in the other group agree with particularists that the fact that wiping the torturer's brow would benefit him is no moral reason to wipe his brow. Antiparticularists in this second group thus agree with particularists that the example about wiping the torturer's brow refutes the claim that benefiting someone is always a moral pro. But these antiparticularists oppose particularism by claiming there is some other feature that does always have the same moral polarity. For example, these antiparticularists might claim that any act with the feature of benefiting an innocent person has at least this in its moral favor. In other words, antiparticularists in this second group abandon the more general claim that benefiting a person is always a morally positive feature, but they insist on the somewhat less general claim that benefiting an innocent person is always a morally positive feature.

The debate over particularism is mostly about whether there are any correct informative general principles, either that specify all-things-considered moral rightness or that indicate which features always operate as moral pros or cons. Antiparticularists win the debate if they come up with correct informative general principles of one or both kinds. Particularists win if they show that every informative general principle put forward is incorrect.

The debate over particularism has other elements as well. On the one hand, particularists say that one can often see not only which features count in which way in a particular situation but also what is all-things-considered morally right in that situation. If particularists are right about that, the question is posed: What is the point of trying to formulate general principles if we can see which particular acts are right without them?

On the other hand, antiparticularists point out that we commonly take being unprincipled as a serious moral flaw. Why is being unprincipled such a moral flaw if acting on principles is not part of being moral? Furthermore, why does moral education start with learning rules and principles if these end up playing no role in determining moral rightness? And why does moral reasoning so often consist in comparing different cases if correct moral judgments are always about particular cases rather than about classes of cases or types of situations?

Particularists pose a challenge to the idea that principles play an essential role in morality. This challenge has forced other moral philosophers to be more specific about which principles they defend and about what roles they think principles must play. Rossian pluralists think correct informative principles are only about moral pros and cons. Many other philosophers—for example, utilitarians, Kantians, contractualists, and virtue ethicists—think that there is a correct informative general principle specifying a foundational principle of right and wrong,

yet there is persisting disagreement among them over what this principle is.

See also Deontological Ethics; Divine Command Theories of Ethics; Duty; Golden Rule; Dworkin, Ronald; Hare, Richard M.; Kant, Immanuel; Moral Dilemmas; Moral Principles: Their Justification; Rights; Ross, William David; Utilitarianism.

Bibliography

Blackburn, Simon. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press, 1998. See especially p. 281.

Brandt, Richard B. "Some Merits of One Form of Rule-Utilitarianism." In *University of Colorado Studies in Philosophy*, 39–65. Boulder: University of Colorado Press, 1967. Reprinted in Richard B. Brant, *Morality, Utilitarianism, and Rights*, 111–136. Cambridge, U.K.: Cambridge University Press, 1992.

Dancy, Jonathan. *Ethics without Principles*. Oxford, U.K.: Oxford University Press, 2004.

Dancy, Jonathan. *Moral Reasons*. Oxford, U.K.: Blackwell, 1993.

Dworkin, Ronald. *Taking Rights Seriously*. London: Duckworth, 1977.

Hare, R. M. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Oxford University Press, 1981.

Hare, R. M. "Objective Prescriptions." In *Naturalism and Normativity: Philosophical Issues*. Vol. 4., edited by E. Villanueva. Atascadero, CA: Ridgeview, 1993. Reprinted in R. M. Hare, *Objective Prescriptions and Other Essays*, 1–18. Oxford, U.K.: Oxford University Press, 1999.

Hooker, Brad, and Margaret Olivia Little, eds. *Moral Particularism*. Oxford, U.K.: Oxford University Press, 2000.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford, U.K.: Oxford University Press, 1999. See especially part 3, 198–228.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth, U.K.: Penguin, 1977. See especially chap. 4, 83–102.

Mackie, J. L. "The Three Stages of Universalization." In *Persons and Values: Collected Papers*. Vol. 2., edited by Joan Mackie and Penelope Mackie. Oxford, U.K.: Oxford University Press, 1985. See especially p. 178.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals* [1785]. Translated by H. J. Paton. New York: Harper & Row, 1964.

Ross, W. D. *The Right and the Good*. Oxford, U.K.: Oxford University Press, 1930.

Scanlon, T. M. "Contractualism and Utilitarianism." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams. Cambridge, U.K.: Cambridge University Press, 1982.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998. See especially chap. 5, 189–247.

Sidgwick, Henry. *Methods of Ethics*. 7th ed. London: Macmillan, 1907.

Brad Hooker (2005)

MORAL SENSE

In the first half of the eighteenth century certain British philosophers argued that the "moral sense" is the faculty by which we distinguish between moral right and wrong. The deliverances of this faculty are feelings or sentiments; hence, it is counted as a sense. Our observation of an instance of virtuous action is the occasion for a feeling of pleasure or satisfaction, which enables us to distinguish that action as virtuous. Similarly, our observation of an instance of vicious action is the occasion for a feeling of pain or uneasiness, which enables us to distinguish that action as vicious. The moral sense is also an influencing motive in our pursuit of virtue and our avoidance of vicious behavior, and it plays a part in our bestowal of praise and blame.

HISTORICAL BACKGROUND

Arguments for and against the moral sense take their character from the larger social and intellectual context in which they were advanced. The late seventeenth century and early eighteenth century in Europe saw the culmination of certain lines of thought that had their origin in earlier times. The Protestant insistence on individual conviction in purely religious matters had an effect on other areas of thought as well. The rejection of external authority as the guarantor of religious truth and the consequent reliance of each believer on his own inner light led to a full-blown theory of knowledge in which the different ways a person can know different kinds of subject matter were definitively cataloged. The way of knowing a given subject was appealed to as the foundation or guarantee of truth. The first account of this theory of knowledge was John Locke's *Essay concerning Human Understanding* in the late seventeenth century. The most comprehensive statement of it was the *Treatise of Human Nature* by David Hume in the eighteenth century. These developments in theory of knowledge were closely related to a growing interest in feelings and their expression. The new theory of knowledge was also closely connected with changes in beliefs about God's relation to the world. Speculations about the will of God were no longer a necessary preliminary to doing physics. When the notion of a physics without God met in men's minds with a resistance to religious authority in all matters, including morals, the problem was posed of the possibility of accounting for morality without an appeal to a divine source. But if morality is not founded on God's will, where is the foundation laid? In line with the new theory of knowledge, the most promising direction for a search appeared to be in human nature itself.

The first Englishman to search for the foundation of morals in human nature, Thomas Hobbes, returned with a brilliantly stated but outrageous report. He found that *good* and *evil* are relative to the person who uses these words; and when people are joined together in a commonwealth, then good and evil are subject to the determinations of the commonwealth. As for our motives for pursuing good and avoiding evil, they may be summed up as self-interest. Were it to our own interest to pursue what others, or the commonwealth, have designated as evil, we certainly would; but, for the most part, our appreciation of the convenience that follows from everyone's following the same rules and, at the worst, our fear of punishment on being caught deter us from the practice of evil.

Hobbes's unflattering picture of human nature and his relativistic account of morals, which he presented in *Leviathan,* are the ominous and ever-present background of all discussions of moral philosophy for the next hundred years. They called forth their contradictory counterpart in the writings of the third earl of Shaftesbury. Shaftesbury argued that Hobbes had made a shortsighted survey of human nature. There is benevolence in human nature, as well as selfishness; and indeed, if men were not originally endowed with a disposition to be sociable, the formation of a commonwealth would be impossible. Shaftesbury was the first to attribute to a moral sense our ability to distinguish between good and evil, virtue and vice. This sense, along with our natural affection for virtue, accounts for the possibility of morality. Shaftesbury, however, did not make clear how the possession of a moral sense enables us to avoid relativism in moral judgments; and indeed the specter of relativism must inevitably haunt the proponents of the moral sense.

DEVELOPMENT OF THE DOCTRINE

The systematic development of the doctrine of a moral sense was left to Shaftesbury's successors: first Francis Hutcheson and later Hume. Their first move was to fit the moral sense into the mainstream of eighteenth-century philosophy by finding a place for it in Locke's theory of knowledge. Looking into the human mind, Locke found that all knowledge consists of perceptions, which must arrive in the mind by one of two routes, either sensation or reflection. Whatever can be known must be accounted for as a perception; and whatever cannot be accounted for in this way is not knowledge. The proponents of the moral sense accounted for our knowledge of moral right and wrong as Lockean reflexive perceptions. When someone observes a given action or considers a certain character trait, these first perceptions are immediately followed by a secondary set of feelings of either pleasure or uneasiness, according to whether the action or character is virtuous or vicious. By consulting these secondary perceptions, we can make our moral judgments. The proponents of the moral sense were careful to point out that actions are not virtuous *because* they please. Rather we know them to be virtuous because we are pleased in a *certain manner.* Thus, moral pleasures and pains are distinctive feelings. Hume argued for the possibility of distinguishing different kinds of pleasure by pointing out, for example, that someone may be pleased both by a good musical composition and by a good bottle of wine, and their goodness is determined merely by the pleasure they give; but we do not say on that account that the wine is harmonious or the music of good flavor.

Besides accounting for our knowledge of right and wrong, the moral sense closes the gap between moral knowledge and moral behavior by providing a motive for moral behavior. Since moral knowledge consists of feelings of pleasure and uneasiness, the prospect of enjoying or avoiding these feelings is a sufficient motive for pursuing virtue and avoiding vice. If moral knowledge were not ultimately a matter of feelings, it would be possible for someone to know that a certain kind of action is virtuous but still have no motive for doing it. The moral sense also enables us to account for our approval and condemnation of actions and characters as following from our being pleased or pained by them.

CRITICISM

The moral sense was subjected to two sorts of criticism. The first sort was directed against supposed defects in the doctrine of the moral sense itself. The second sort of criticism advanced the claims of rival candidates for the title of moral faculty.

DEFECTS IN THE DOCTRINE. The bluntest form of the first sort of criticism was to interpret the proponents of the moral sense as talking about an extra organ of sense, "a moral nose" or "a moral ear." How acute they were to have discovered a new human organ which no one had noticed until they came along! Merely to mention the possibility was enough to show the nonexistence of such an organ and to render the doctrine of a moral sense laughable. Hutcheson was especially plagued with this kind of criticism. But he spoke of the moral sense as a determination of the mind, which left the way open for viewing the moral sense not as an organ but as a faculty that can be looked for only in the way memory or will can

be looked for. Hume's defenses against this criticism were somewhat better. He boldly asserted the principle that our acquaintance with our senses or faculties can never be anything but an acquaintance with their characteristic perceptions. Hence, he was justified in confining himself to talk of moral feelings and sentiments; and indeed, he never actually used the phrase "moral sense" in any argument but relegated it to a section title.

The next most severe criticism was to point out that although all men are said to be endowed with a moral sense, there is no universal agreement about moral right and wrong. Hutcheson turned aside this criticism by arguing that the moral sense may be inoperative or defective, just as human eyes may be. Hume added that differences in moral judgments may be attributed to differences in experience and education and to a failure to pass judgment from a disinterested point of view; and he hoped that by additional experience or by a greater effort to achieve disinterestedness moral disputants might be able to reach agreement.

But the critics of the moral sense thought that by far the most serious fault in the doctrine was its apparent foundation of the distinction between moral right and wrong on human nature itself. This opened the door to Hobbesian relativism: Whatever action pleases is virtuous, and whatever displeases, vicious. Actually, Hutcheson based the distinction between virtue and vice on the will of God, one step removed from human nature. It just so happens, he held, that God determined us to be pleased by benevolent actions; and when nothing interferes with the moral sense, we count benevolence a virtue and malevolence a vice. But, his opponents argued, to base the distinction between moral good and evil on God's will is no less arbitrary than to base it on human nature itself. If, by divine fiat, we count benevolence a virtue, we might very well have done the opposite, had God so pleased. What is more, the distinction between good and evil cannot possibly rest on God's will, for if good and evil have not some real character in themselves, what is there to determine his will in the first place?

Hume based the distinction between moral right and wrong directly on human nature—that is, our power to be pleased and displeased by different ways of acting—without an appeal to any divine determination of this power. But if there is to be a stability in the distinction between moral right and wrong, then there must be a consistency in human nature. This is no easy thing to show, for the slightest inspection of human affairs appears to tell against it. Yet Hume argued that, on balance, man is more of a social being than not. Indeed, this contention had always been strongly supported by proponents of the moral sense; but Hume added the refinement that man's very inclination to be social leads him to be pleased by those actions and character traits which tend to make society possible and to be displeased by those which tend to disrupt society. Thus, while the distinction between virtue and vice does indeed rest on human nature, it is not an arbitrary distinction. We do have a good reason for preferring one sort of action to another, namely the action's tendency to maintain society. Should someone ask, "And why should I prefer the maintenance of society to its destruction?" Hume had no answer in the form of a logical argument. He certainly recognized the possibility of someone's preferring the destruction of society over its maintenance; but on such a fundamental issue, he held, there can be no arguments pro or con, but only an appeal to feelings. Society exists because, as a matter of fact, by far the greater number of people have the kind of feelings that make it possible.

RIVAL MORAL FACULTIES. Another set of objections to the moral sense was advanced by those who argued that the faculty by which we discern moral right and wrong must be reason, or the understanding. The most notable members of this school were Samuel Clarke, John Balguy, and Richard Price. Their most characteristic doctrine was that moral right and wrong are unchanging and unchangeable and, thus, independent of any human, or even divine, determination. This school accepted Locke's theory of knowledge with the modification that the understanding is capable of originating new simple ideas for itself by considering those it gets by way of the two great avenues of sensation and reflection. Thus, according to Clarke, the understanding can discern a certain eternal fitness that things and actions bear to one another in their natures. He likened these moral discoveries to mathematical reasoning in which one discovers the consistency of certain concepts. The implication is that the absolute and immutable character of moral distinctions is such that they can be known only by reason. Therefore, the moral faculty could not possibly be a sense.

Hume endeavored to answer these arguments by pointing out that, strictly considered, reason is capable only of comparing ideas. Since moral knowledge is a sentiment or feeling that arises on the observation of an action or character trait, it is not the result of comparing ideas, and thus it cannot be a conclusion of reason. What is more, since our moral sentiments about certain actions may excite us to perform or to avoid these actions, it is even more doubtful that our moral knowledge comes from reason, for, according to Hume, the conclusions of

reason alone can never be an exciting motive to action. A person may know that a certain way of acting may have a certain result, but in order for him to act to achieve that result, he must first find it pleasing.

The moral sense and reason were not the only candidates for a moral faculty proposed at that time. Joseph Butler argued for conscience; and Adam Smith chose to argue for sympathy—which had also figured in Hume's moral philosophy—as the source of moral distinctions. Considering the arguments advanced on behalf of the different candidates for the moral faculty, one can see that the issue was never one that could be settled by empirical investigation. The search for a moral faculty had its origin in the general acceptance of a faculty psychology, supplemented with the Lockean assumption that the acts attributed to our mental faculties are to be accounted for as the occurrence of various sorts of perceptions. When one recognizes the ad hoc character of the conceptual framework in which the disputes over the nature of the moral faculty took place, one can see why there was no resolving them. When one no longer finds a need for a faculty psychology, the need to search for a moral faculty goes too.

The present-day moral philosopher no longer casts his study as an investigation of the deliverances of a moral faculty, but rather as a study of the logic of moral discourse. Despite their central preoccupation, the proponents of the moral sense have made a contribution to the development of modern moral philosophy. In particular, they contributed the points that morality assumes the value of society and is incomprehensible apart from this presupposition; that conduct must be judged by general rules; and that a general rule of definitive importance to morality is the injunction to act for the greatest good of the greatest number. But perhaps the most important contribution to moral philosophy by the proponents of the moral sense was their insistence that feeling has a place in morals and that to miss this fact is to omit a distinctive element in moral discourse.

See also Ethics, History of; Objectivity in Ethics.

Bibliography

For original statements of the moral sense doctrine, see Shaftesbury (Anthony Ashley Cooper), *An Inquiry concerning Virtue or Merit* (1699), to be found in *Characteristics of men, morals, opinions, times,* 3 vols. (London, 1711), Vol. II; Francis Hutcheson, *An Inquiry into the Original of our Ideas of Beauty and Virtue* (London, 1725) and *An Essay on the Nature and Conduct of the Passions and Affections with Illustrations upon the Moral Sense* (London, 1728); and David Hume, *A Treatise of Human Nature* (London, 1740), Book III.

For rival moral faculties, see, for reason, Samuel Clarke, *A Discourse concerning the Unchangeable Obligations of Natural Religion and the Truth and Certainty of the Christian Revelation* (London, 1706); John Balguy, *The Foundation of Moral Goodness* (London: John Pemberton, 1728); and Richard Price, *A Review of the Principal Questions and Difficulties in Morals* (London, 1758). For conscience, see Joseph Butler, *Fifteen Sermons upon Human Nature, or Man Considered as a Moral Agent* (London, 1726). For the rival moral faculty of sympathy, see Adam Smith, *Theory of Moral Sentiments* (London: Millar, 1759).

For modern studies, see James Bonar, *Moral Sense* (New York: Macmillan, 1930) and D. Daiches Raphael, *The Moral Sense* (London: Oxford University Press, 1947).

Elmer Sprague (1967)

MORAL SENTIMENTS

One's sentiments are the contents of one's sensed, or felt, experience—in contrast to the contents of simply one's thoughts. Whatever else they are, then, sentiments are affective phenomena. In common parlance, talk of sentiments refers alternatively to occurrent feelings, affective dispositions, and emotional attitudes taken toward people and objects. Moral sentiments, where the adjective *moral* is used in a descriptive sense, would then be some subset of these feelings, dispositions, and attitudes: those that are more or less intimately related to moral phenomena. Whether any of the moral sentiments thus understood are moral in a normative sense, that is, whether one morally may or should experience or express any of these sentiments in relevant circumstances, is a further question.

One problem that immediately confronts any philosophical account of moral sentiments is the question whether such affective phenomena in fact form a unified category. Affective responses vary widely with respect to their causes, phenomenology, duration, intentional objects (if any), and mode of expression, as well as their susceptibility to rational assessment and control. This variability is no less present in the case of that subset of affective phenomena related in some way to morals. Contrast, for example, rationally impervious and visceral disgust to resentment, a comparatively subdued attitude that arguably is a response fitting only to moral wrongs. Both disgust and resentment, however, are moral sentiments in the sense that people commonly experience these affective reactions in response to moral phenomena.

Just which phenomena one admits to the category of moral sentiments depends, of course, on the specific theory of the sentiments one accepts. Consideration of con-

temporary theories of the emotions is instructive here. Although such theories are quite varied, a common taxonomy distinguishes between cognitivist and noncognitivist theories of emotion. Cognitivist theories of emotion hold that emotions necessarily involve thoughts, beliefs, or judgments ascribing properties to their objects. Some cognitivists (Nussbaum 2001) identify emotions with evaluative judgments, for example, identifying fear with the evaluative judgment that the object of fear somehow threatens one's welfare or identifying one's resenting another's action with the judgment that the other wrongs one in so acting. Sentiments, understood as essentially affective phenomena, apparently play at best a peripheral role on some such theories of emotion.

Noncognitivist theories of emotion, in contrast, embrace a view of emotions as essentially felt experiences different in kind from thoughts beliefs or judgments. William James (1842-1910), famously identified emotions with the perception of bodily changes—or feelings—caused by external stimuli. Contemporary followers of James (Prinz 2004) have built on his emotional noncognitivism to avoid what they view as shortcomings of the cognitivist alternatives. Some noncognitivists object that emotions, unlike beliefs or judgments, are not properly subject to assessment in terms of truth or falsehood. Noncognitivists also object that cognitivist theories require that those subject to emotions possess a conceptual or propositional repertoire that obvious subjects of emotion—human infants and animals, for example—do not, in fact, possess. In response to such objections, some philosophers opt for mixed theories according to which emotions are some amalgam of cognition and affect (Oakley 1992).

Clarity about the correct theory of affective responses is a prerequisite for progress in the longstanding philosophical debate over the role of moral sentiments in moral agency. Philosophers have long debated the role of moral sentiments in, for example, (1) moral deliberation and judgment, (2) moral motivation, and (3) moral responsibility.

In examining the role of moral sentiments in moral deliberation and judgment, moral motivation, and moral responsibility, modern moral philosophers have been concerned especially with the role one should attribute to moral sensibility—generally understood as a capacity for experiencing, or disposition to experience, feelings, emotions, and attitudes that include guilt, resentment, respect, esteem, honor, pride, and shame—relative to the role of reason, understood as a cognitive capacity whose objects (e.g., thoughts or propositions) are amenable to evaluation in terms of truth or falsehood. Moral philosophers committed to treating moral judgments as bona fide judgments, whether by taking them to refer to causally explanatory moral properties or by regarding them as subject to similarly robust standards of truth and falsehood as is descriptive discourse, are often known as metaethical cognitivists. Metaethical noncognitivists, in contrast, deny that moral evaluations identify irreducibly moral properties or report truth-evaluable beliefs. The distinction between cognitivism and noncognitivism in one's metaethical theory is independent of the distinction between cognitivism and noncognitivism about emotions. However, differences among philosophers concerning the relative role of sentiment and reason in the moral domain reflect philosophical differences in the specific theory of sentiment, or emotion, they accept.

HISTORICAL CONTEXT

Although contemporary moral philosophers might be inclined to trace the term *moral sentiments* to developments in eighteenth-century British moral philosophy, philosophical interest in the affective aspects of one's moral experience is not limited to any specific epoch. Already in ancient Hellenistic philosophy, one finds a concern with the place of feelings, emotions, and affective attitudes generally in the constitution and care of the psyche, or soul.

For Plato (c. 429–347 BCE) and Aristotle (384–322 BCE), for example, human excellence (for them the subject of ethics) required that one's soul be properly constituted in the relation of its rational, desiderative, and appetitive parts—the latter comprising the domain of sentiments or emotions. For Plato, the proper constitution of the soul was an achievement of an upbringing where one's appetites (e.g., natural urges for food and sex), desires (e.g., aspirations for the goods of honor and victory), and rational judgments were in harmony. Absent a proper upbringing, the desiderative and appetitive parts of the soul were bound to prove unruly and psychically divisive, thereby making a good life unattainable. On such a view, arguably, all affects of the soul have ethical import, whether or not they have ethical content.

Aristotle further developed an account of the education of the soul where the parts concerned with feeling pleasures and pains functioned, at least in persons properly reared, as important guides for choosing and acting well. Affective dispositions of the soul could play this role for Aristotle because he understood the feeling part of the soul according to a perceptual model on which it could provide one with knowledge of the objects it perceived

through pleasure and pain, much as one's visual perceptions provide knowledge of the objects of sight. On this view, to possess virtues of character such as courage and justice consists in part in being disposed to experience the appropriate emotions in response to, respectively, fearful circumstances and the unfair distribution of goods.

The Stoic school of Hellenistic philosophers (of which Zeno of Citium [335–263 BCE] is an example) combined a rich cognitivist theory of the emotions as judgments concerning value with the prescription that the wise person, or sage, should ultimately expurgate himself of emotion altogether. The motivation for the prescription derived from the Stoic view that virtue is the only good. Other things of value (e.g., health, and wealth), while typically choice-worthy, are things about which the Stoic sage is properly indifferent. To the extent that emotions give importance to things other than virtue by judging them good, then, they implicate one in false judgments about the good. As such, emotions are antagonistic to reason, by whose power one should strive to eliminate them.

This antagonistic divide between sentiment and reason reappears in the early modern period, fueled by changes brought by the advance of Newtonian science, religious strife, and philosophy itself (e.g., John Locke's [1632–1704] philosophy of mind). It is in this, the early modern period, beset by changes that sustained doubt about the status of morality as a deliverance of revelation or of reason alone, that one first encounters the school of moral philosophers known as the *sentimentalists*: Anthony Ashley Cooper (the third Earl of Shaftesbury [1671–1713]) Frances Hutcheson (1694–1746), David Hume (1711–1776), and Adam Smith (1723–1790).

Shaftesbury is perhaps most often credited with having first used the phrase *moral sense*, defending it as a sense, quite literally, of moral right and wrong. According to Shaftesbury, the moral sense enables all persons to experience affections of approval or disapproval upon reflection on the first-order affections, or motives, of oneself and others. Judgments about what is morally right and morally wrong, as well as the motivations to action that they support, are on this view expressions of the reflective approval or disapproval that is fitting for one's and others' motives.

Hutcheson adapted Shaftesbury's theory of a moral sense that apprehended with approval virtuous motives, which Hutcheson understood as forms of benevolence, and responded with disapproval to the vicious. However, Hutcheson abandoned Shaftesbury's metaphysical views, among them the view that the immutable order of nature guarantees the fittingness of these moral affections for their objects. Hutcheson also viewed reason, as opposed to sentiment, as a purely theoretical and, so, motivationally inert faculty. These two features of Hutcheson's philosophy are echoed in the empiricist sentimentalism of Hume.

An admirer of Cicero (106–143 BCE) and Tacitus (c. 56–120 CE), Hume inverted the Stoic hierarchy of reason and sentiment when he announced that "Reason is, and ought only to be the slave of the passions" (Hume 1973, II. iii. 3, p. 415). To be sure, Hume's slogan conceals a more nuanced Humean view of moral evaluation and moral motivation. Humans are naturally constituted, on Hume's view, to feel certain passions, or sentiments, in response to certain causes. Reason, exclusively concerned as it is with matters of fact and relations of ideas, cannot oppose passion in the sense that reason cannot cause us to form moral beliefs or motivate us to act in the absence of some affective input. Hume nonetheless distinguishes between better and worse ways of forming evaluative beliefs and between better and worse motives. He does so by privileging moral assessments made from what he calls the common or general point of view, a point of view one succeeds in occupying when one evaluates motives or character traits in terms of their typical effects. Such evaluation proceeds not through the operations of a moral sense but through the influence of the general point of view on what Hume identifies as the mechanism of sympathy. For sympathetic creatures occupying the general point of view, moral evaluation consists in apprehending whether the motives or character traits being evaluated are immediately agreeable or useful to oneself or others. In this way, Hume concludes "Morality…is more properly felt than judg'd of " (Hume 1973, III. i. 2, p. 470).

Smith developed his brand of moral sentimentalism in his *The Theory of Moral Sentiments* (1759/1982). Smith's sentimentalism resembled Hume's in privileging sympathy as a psychological process by means of which one comes to take pleasure in virtue and be pained by vice. Significantly, however, the two differed in their conceptions of precisely how sympathy operates on the sentiments. Whereas Hume envisaged moral evaluation being made from an observer's point of view on the motives or character traits of others, Smith's sympathetic exercise required one to consider motives and character traits by projecting oneself into the point of view of the possessor or those affected by that point of view. By thus imagining oneself in another's situation, Smith maintained, one comes to share in the feelings that the other person experiences. Sympathy, understood as an imagi-

native capacity for fellow-feeling, in this way provides one a motive toward benevolence, according to Smith.

The antagonistic sentiment–reason divide that some early modern philosophers championed is transposed into a semantic key in the work of some twentieth-century Anglo-American philosophers. The work of metaethical noncognitivists such as Charles Stevenson (1908–1979) and Alfred Jules Ayer (1910–1989) especially influenced the development of contemporary philosophical debate over the respective roles of sentiment and reason in morals. Ayer notably argued for a distinction between cognitive meaning (possessed by descriptive statements and analytic statements) and emotive meaning (possessed by moral statements). On Ayer's view, moral utterances, appearances notwithstanding, serve not to assert facts but to express the emotions of the speaker. "In saying that a certain type of action is right or wrong," he wrote in *Language, Truth, and Logic* (1936/1952), "I am not making any factual statement, not even a statement about my own state of mind. I am merely expressing certain moral sentiments." Thus, on Ayer's view, ethical utterances are not candidates for assessment in terms of truth or falsity. Stevenson endorsed Ayer's so-called emotive view of moral language, though he was careful to stress the interconnections between descriptive and emotive meaning, as well as a meaning of *true* that might appropriately (albeit emotively) be applied to ethical statements. Caveats notwithstanding, Ayer's philosophy of language, combined with the view that the proper task of philosophy is analysis of cognitively meaningful language, issued in an era when the work of the moral philosopher was limited to that of metaethical reflection on moral language and phenomena.

CONTEMPORARY DEBATE

Although the days when Anglo-American moral philosophers limited their task to metaethical reflection have ended, the debate about the respective roles of sentiment and reason in moral deliberation and judgment, moral motivation, and moral responsibility lives on.

In the area of moral deliberation and judgment, the expressivists are the noncognitivist inheritors of emotivism. Like the emotivists, expressivists (e.g., Allan Gibbard and Simon Blackburn) distinguish between descriptive and evaluative discourse. Whereas emotivists such as Ayer hold that ethical utterances express the speaker's sentiments, however, the expressivists defend a more complicated account of the affective phenomena expressed in ethical utterance. They do so in an attempt to avoid now-familiar problems with emotivism; for, example, its difficulties accommodating ethical disagreement and explaining the behavior of ethical expressions in the embedded contexts and inferences common to moral deliberation.

More recently, a form of sentimentalism about ethical judgment has become popular that holds that ethical utterances that predicate some evaluative property *P* (such as, "murder is wrong") do not express emotion but, rather, express the speaker's endorsement of the having of certain emotions in response to property *P*. Following on the work of Gibbard, Daniel Jacobson and Justin D'Arms have begun to develop a form of *rational sentimentalism* that clarifies the kind of endorsement that is at issue in evaluative judgments generally. Making use of a generic relationship of *fittingness*, they offer an account of when one's moral sentiments are fitting to their objects, which distinguishes this question from other appraisals of the sentiments (such as prudential and moral appraisals of these responses). Correcting past philosophers' conflation of the claim that an emotion is fitting its object with the claim that it is morally appropriate to its object, their work promises to reinvigorate philosophical study of moral emotions that, while arguably fitting to their objects in certain circumstances, nonetheless have suffered neglect due to their perception as being somehow morally undesirable. Contempt and moralized disgust are examples.

Neosentimentalism about ethical judgment of the D'Arms/Jacobson variety is a form of emotional rationalism that bridges the metaethical noncognitivism/cognitivism divide. In some ways this neosentimentalism resembles the so-called sensibility theories of moral judgment espoused by metaethical cognitivists such as John McDowell. McDowell holds that moral sensibility functions much like perceptual ability: One's moral sensibility enables one to apprehend and form beliefs whose contents are irreducibly moral properties, much as one's visual perception allows one to apprehend color properties. Some contemporary moral philosophers rely on metaethical cognitivism, such as McDowell's, to urge a return to an Aristotelian view of moral sentiment that rejects an inherently antagonistic divide between reason and sentiment. On such an Aristotelian view, ethical deliberation and judgment primarily differ from deliberation and judgment about nonmoral phenomena, not in any metaphysical or epistemological peculiarities pertaining to their content but in the necessarily practical nature of their progeny.

In stressing the obviously practical nature of moral deliberation and judgment (that is, the way in which it engages with intention, action, and affect), however, the metaethical cognitivist risks unwittingly fueling the metaethical noncognitivist program. If one holds, with the metaethical noncognitivists, that propositional attitudes such as beliefs are motivationally impotent in themselves, then acknowledging the practical character of moral deliberation and judgment requires one to reject a view of moral deliberation and judgment as exclusively cognitive phenomena. The metaethical noncognitivist's rejection is complete in denying that they are even partly so, a rejection supported by the noncognitivist tendency to understand mental phenomena in terms exclusively of beliefs or desires.

One response to this denial proceeds from arguing that such a mental repertoire is impoverished in failing to admit that certain mental phenomena, emotions among them, may possess the representational character of beliefs while also possessing the motivational force of desires. In this way, a more nuanced moral psychology might advance contemporary debate.

Finally, it is worth noting that although the necessarily practical character of moral deliberation and judgment typically is raised as a challenge for the metaethical cognitivist, cognitivism in the theory of emotion—at least those versions that simply equate emotions with evaluative beliefs—also invites the question of how to account for the motivational potency of emotions.

As this brief taxonomy suggests, different theoretical commitments—whether in moral theory or the theory of emotions—support different conceptions of how sentiments figure in moral experience. These commitments also support different views concerning responsibility for one's moral emotions. If emotions are akin to urges and desires, to pleasures and pains, or to perceptions of some sort—with respect to which individuals arguably are passive—is it even intelligible to regard oneself and others as accountable for emotions? If, alternatively, emotions are judgments, are individuals thereby any closer to locating a form of control one exercises over them that would justify holding oneself and others accountable for them? Or should one challenge the assumption, as do some philosophers, that such control is necessary for justifying attributions of responsibility?

The philosopher P. F. Strawson famously argued that even should the metaphysical thesis of determinism hold true, individuals could not avoid holding themselves and others in general responsible for what he called the reactive attitudes (for example, gratitude, resentment, forgiveness, love, and hurt feelings). To be sure, he recognized, one often suspends these attitudes in special cases: the cases of children; the incapacitated. In the case of typical mature agents, however, susceptibility to the reactive attitudes is a condition of membership in a common humanity. On such a view, the theoretical question whether one possesses the freedom to control one's emotions is abandoned in favor of attending to the necessity of regarding oneself and others as responsible for emotions if one is to regard oneself and others as moral agents at all. The alternative, Strawson argued, is not a rational expurgation of such attitudes in deference to the determinist thesis but an objective stance toward oneself and others that amounts to viewing humans as perpetual patients, appropriate objects not of emotional engagement but of treatment. If Strawson is correct, philosophical interest in the moral sentiments is likely to continue to evade constraint to any single historical epoch, central as they are to moral personhood.

See also Aristotle; Ayer, Alfred Jules; Cicero, Marcus Tullius; Emotion; Emotive Theory of Ethics; Hellenistic Thought; Hume, David; Hutcheson, Francis; Locke, John; McDowell, John; Metaethics; Moral Sense; Newton, Isaac; Plato; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Shame; Smith, Adam; Stevenson, Charles L.; Strawson, Peter Frederick; Virtue Ethics; Zeno of Citium.

Bibliography

Aristotle. *Nicomachean Ethics*. Translated by Christopher Rowe. New York: Oxford University Press, 2002.

Ayer, A. J. *Language, Truth, and Logic*. London: V. Gollancz, 1936. Reprinted, New York: Dover, 1952.

Blackburn, Simon. *Ruling Passions*. New York: Oxford University Press, 1998.

D'Arms, Justin, and Daniel Jacobson. "The Moralistic Fallacy: On the 'Appropriateness' of Emotion." *Philosophy and Phenomenological Research* 61 (2000): 65–90.

D'Arms, Justin, and Daniel Jacobson. "Sentiment and Value." *Ethics* 110 (4) (2000): 722–748.

Deigh, John. "Cognitivism in the Theory of Emotions." *Ethics* 104 (4) (1994): 824–854.

De Sousa, Ronald. *The Rationality of Emotion*. Cambridge, MA: MIT Press, 1987.

Gibbard, Allan. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press, 1990.

Greenspan, Patricia. *Practical Guilt: Moral Dilemmas, Emotions, and Social Norms*. New York: Oxford University Press, 1995.

Griffiths, Paul E. *What Emotions Really Are*. Chicago: University of Chicago Press, 1997.

Hume, David. *A Treatise of Human Nature*, edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford, U.K.: Calrendon Press, 1973.

Hursthouse, Rosalind. "Virtue and the Emotions." In *On Virtue Ethics*. New York: Oxford University Press, 1999.

Hutcheson, Frances. *An Essay on the Nature and Conduct of Passions and Affections, with Illustrations on the Moral Sense*. London 1728. Reprinted, Aaron Garrett, ed. Indianapolis: Liberty Fund, 2002.

James, William. "What Is an Emotion?" *Mind* 9 (34) (1884): 188–205.

Kenny, Anthony. *Action, Emotion, and Will*. New York: Routledge and Kegan Paul, 1963.

Mason, Michelle. "Contempt as a Moral Attitude." *Ethics* 113 (1) (2003): 234–272.

McDowell, John. "Values and Secondary Qualities." In *Mind, Value, and Reality*, 131–150. Cambridge, MA: Harvard University Press, 1998.

Nussbaum, Martha. *Upheavals of Thought*. New York: Cambridge University Press, 2001.

Oakley, Justin. *Morality and the Emotions*. New York: Routledge, 1992.

Pitcher, George. "Emotion." *Mind* 74 (295) (1965): 326–346.

Prinz, Jesse J. "Embodied Emotions." In *Thinking About Feeling,* edited by Robert C. Solomon, 44–60.

Shaftesbury. *Characteristics of Men, Manners, Opinions, Times*, edited by Lawrence E. Klein. New York: Cambridge University Press, 2000.

Smith, Adam. *A Theory of Moral Sentiments*, edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Fund, 1982.

Solomon, Robert C. *The Passions: Emotions and the Meaning of Life*. Indianapolis: Hackett, 1993.

Solomon, Robert C., Ed.*Thinking about Feeling*. New York; Oxford University Press, 2004.

Stocker, Michael, and Elizabeth Hegeman. *Valuing Emotions*. New York: Cambridge University Press, 1992.

Strawson, P. F. "Freedom and Resentment." Reprinted in *Free Will*, edited by Gary Watson, 59–80. New York: Oxford University Press, 1982.

Taylor, Gabriele. *Pride, Shame, and Guilt: Emotions of Self-Assessment*. Oxford: Clarendon Press, 1975.

Velleman, J. David. "Love as a Moral Emotion." *Ethics* 109 (2) (1999): 338–374.

Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1996.

Williams, Bernard. "Morality and the Emotions." In *Problems of the Self: Philosophical Papers 1956–1972*, 207–229. New York: Cambridge University Press, 1973.

Williams, Bernard. *Shame and Necessity*. Berkeley: University of California Press, 1993.

Wollheim, Richard. *On the Emotions*. New Haven, CT: Yale University Press, 1999.

Michelle Mason (2005)

MORAL SKEPTICISM

The two main forms of skepticism about morality are skepticism about moral truths and skepticism about reasons to comply with moral considerations. These doctrines challenge the cognitive significance or rational authority of morality.

Skepticism about moral truths denies that there are—or that we can know that there are—true moral propositions (or facts) that entail that something has a moral attribute. This form of skepticism seems to imply that rational and informed agents would give moral claims no credence. It has been supported by a variety of arguments, including arguments about moral disagreement. One deep motivation for it is the difficulty of explaining the normativity or action-guiding nature of moral claims.

Noncognitivists attempt to explain the normativity of moral judgments by supposing that their function is to express states of the speaker and to affect behavior rather than to express propositions. Noncognitivists would agree that there are no true moral propositions, since they hold that moral claims do not express propositions. Yet they do not view moral claims as defective. According to noncognitivists, one who makes a claim, such as "Truthfulness is morally required." expresses a moral attitude or acceptance of a moral norm (Ayer, [1936] 1946; Gibbard, 1990; cf. Hume, [1739–1740] 1978).

Cognitivists object that our moral thinking cannot be understood except on the assumption that moral claims express propositions. To avoid skepticism, cognitivists must believe that there are moral properties that are sometimes exemplified. For if no moral property exists, or if none is exemplified, it follows that there are no moral requirements, no moral goods or bads, no moral virtues or vices. It may follow that there are no *honest* persons, for example, although there may be truthful persons.

A skeptic might hold that moral properties exist but that none is exemplified. This position seems implausible, however, for if there is the property of wrongness, it would be astonishing if nothing were ever wrong. Alternatively, a skeptic might argue that there are no moral properties. According to widely accepted views about propositions, however, the proposition that lying is wrong, for example, would attribute the property wrongness to acts of lying. The property would be a constituent of the proposition. Hence, if there are no moral properties, these views about propositions may lead to the conclusion that no proposition is expressed by sentences such as "Lying is wrong."

J. L. Mackie argued that there are no moral properties (1977). We conceive of moral properties as intrinsic; if an action is wrong, it is wrong "as it is in itself." But we

also conceive of moral properties as intrinsically action guiding; we can be motivated to act in an appropriate way simply by coming to know that an action would be wrong, regardless of any antecedent motivations. Yet, Mackie thought, it is not intelligible that it be intrinsic to an action's having an intrinsic property that the mere recognition that the action has the property could motivate a person. The idea of a moral property is not intelligible; moral properties would be metaphysically "queer."

Gilbert Harman (1977) argued for an epistemic version of skepticism about moral truths. He argued that there seems to be no good reason to affirm any moral proposition, for moral hypotheses are never part of the best explanation of any observation. There is always a better nonmoral explanation. The belief that there are true moral propositions is therefore unwarranted.

Skepticism about moral truth appears to have a life of its own in secular cultures, independent of skeptical arguments. Some people believe that moral truths are grounded in God's commands. A secular culture would tend to think, however, that all substantive facts are empirical and "natural." And natural facts do not seem to be normative in the way moral facts are normative. It is therefore difficult to see how a natural fact could be a moral fact.

The second skeptical doctrine is the thesis that there need be no reason to comply with moral considerations. According to this thesis, rational agents would not give attention to moral considerations, as such, in deciding how to live their lives. To be sure, we may desire to live morally, and this desire may give us a reason to live morally. Or we may find ourselves in a context in which living morally is in our interest. Yet these possibilities do not show that there is necessarily a reason to comply with moral considerations (Nielsen, 1974); they do not distinguish moral considerations from considerations of etiquette, for example.

Skepticism about compliance is typically motivated by the idea that morality can require actions that are not to the agent's advantage. Assuming that there are reasons for one to do something just in case it would be to one's advantage, this idea implies that there may be no reason to comply with morality.

The two main skeptical doctrines are closely linked, on certain ways of thinking. First, it may seem, we cannot be guaranteed to have reasons to comply with moral considerations unless there are moral truths of which we have knowledge. Second, a kind of "internalist" theory holds that moral facts are "constituted" by reasons. On this view there are no moral facts unless there are reasons of a relevant kind.

Internalist antiskeptical theories attempt to defeat both skeptical doctrines at once. Immanuel Kant held, in effect, that if a moral imperative corresponds to a truth, it does so in virtue of the fact that it would be complied with by any fully rational agent (Kant, [1785] 1981). "Externalist" theories attempt to deal with skepticism about moral truths independently from skepticism about compliance (Sturgeon, 1985). Those who believe that moral truths are grounded in God's commands may suppose, for example, that God necessarily gives us reasons to comply.

Philosophers who accept one of the skeptical doctrines typically try to defuse it. Skeptics about rational compliance may argue that people with normal psychologies invariably have reasons to comply with morality. Skeptics about moral truth may argue that there nevertheless are reasons to engage in the practice of judging things morally.

See also Ayer, Alfred Jules; Harman, Gilbert; Hume, David; Kant, Immanuel; Mackie, John Leslie; Metaethics; Moral Realism; Skepticism, History of.

Bibliography

Ayer, A. J. *Language, Truth, and Logic* (1936). London: Gollancz, 1946.

Copp, D. "Moral Skepticism." *Philosophical Studies* 62 (1991): 203–233.

Gibbard, A. *Wise Choices, Apt Feelings: A Theory of Normative Judgment.* Cambridge, MA: Harvard University Press, 1990.

Harman, G. *The Nature of Morality: An Introduction to Ethics.* New York: Oxford University Press, 1977.

Hume, D. *A Treatise of Human Nature* (1739–1740). Edited by P. H. Nidditch. Oxford: Clarendon Press, 1978.

Kant, I. *Grounding for the Metaphysics of Morals* (1785). Translated by James W. Ellington. Indianapolis: Hackett, 1981.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* Harmondsworth, U.K.: Penguin, 1977.

Nielsen, K. "Why Should I Be Moral?" In *Introductory Readings in Ethics,* edited by W. K. Frankena and J. T. Granrose. Englewood Cliffs, NJ: Prentice-Hall, 1974.

Nietzsche, F. *Basic Writings of Nietzsche.* Edited and translated by Walter Kaufmann. New York: Modern Library, 1968. See *The Genealogy of Morals* and *Beyond Good and Evil.*

Sturgeon, N. "Moral Explanations." In *Morality, Reason, and Truth,* edited by D. Copp and D. Zimmerman. Totowa, NJ: Rowman and Allanheld, 1985.

David Copp (1996)

MORE, HENRY
(1614–1687)

Henry More, the philosopher, poet, and Cambridge Platonist, was born at Grantham, Lincolnshire. His father, "a gentleman of fair estate and fortune," was a strict Calvinist but supported church and king against the Puritans. He introduced his son to Edmund Spenser's *Faerie Queene*, and Spenser's Platonism, allegorizing, and moral attitudes persist in More's own writings. At Eton, where More was educated, the religious atmosphere was latitudinarian; More abandoned the Calvinist doctrine of predestination without losing what he called "an inward sense of the divine presence." In December 1631 he entered Christ's College, Cambridge, where he was elected to a fellowship in 1639. He remained at Cambridge until his death, refusing preferments, except those he could pass on to such fellow Platonists as Edward Fowler and John Worthington. Unlike most of the Platonists he took no part in public affairs or in university administration. In *An Explanation of the Grand Mystery of Godliness* (1660) he defended what he called a "neutrality and cold indifference in public affairs."

When More entered Christ's College, it was split into three factions—the high church party, the Calvinistic Puritans, and the Medians, so called because they stood for a moderate church and had as their leader Joseph Mede, or Mead (1586–1638), author of *Clavis Apocalyptica* (1627), an allegorical interpretation of the Scriptures. More's tutor Robert Gell, whose *Remaines* were published in 1676, was a member of Mede's party; he emphasized even more strongly than Mede that salvation depended upon "good works," not on blind faith, and he shared Mede's fascination with demonology and Scriptural interpretation. More himself described Mede as an "incomparable interpreter of Prophecies," and in *The Grand Mystery of Godliness* defends his biblical interpretations against the criticisms of Hugo Grotius.

NEOPLATONISM

Developing a passion for philosophy, More read widely in Aristotle and the Scholastics. However, he became impatient with their failure, as he thought, to provide a satisfactory account of the relation between God and the individual self. He therefore turned to the Neoplatonists and to mystical writings, especially the *Theologia Germanica*, an anonymous fourteenth-century mystical handbook that Martin Luther republished in 1516. From the mystics and Neoplatonists More derived his belief that to acquire knowledge, one must first seek moral perfection and his definition of perfection as the process of becoming godlike by subduing egoism. More did not refer to Benjamin Whichcote, none of whose writings was published until just before More's death, but he told his biographer that 1637 was the date of his conversion to his "new way of thinking"; this was the year of Whichcote's appointment as Sunday lecturer at Trinity Church. More shared certain fundamental epistemological and metaphysical ideas with Ralph Cudworth. These were ultimately derived from Platonism, and how far Cudworth's formulation of them influenced More or vice versa is impossible to determine.

More's first philosophical writings were allegories in Spenser's manner, collected in 1647 as *Philosophical Poems*. They present a complicated world view in which the basic concepts of Neoplatonism are interpreted in Trinitarian terms. Christ is presented as a living demonstration that a human being can be wholly possessed by God, rather than as a Calvinistic redeemer. More's poems preach the lesson common to Cambridge Platonism that the life we live, not the creed we preach, is our path to salvation, but their obscure allegorical manner is quite remote from Whichcote's direct, epigrammatic style.

METAPHYSICS

In atmosphere the *Philosophical Poems* carry us back to the Renaissance. More saw Plato through the eyes of Plotinus and Plotinus through the eyes of Renaissance humanists such as Marsilio Ficino, who set out with the help of allegory to Christianize Neoplatonic metaphysics. Yet on December 11, 1648, More wrote the first of four Latin letters to René Descartes, in which he not only expressed the highest admiration for Descartes's work but added that Descartes's views "appear indeed to be my own—so entirely have my own thoughts run along the channels in which your fertile mind has anticipated me." Nor was this a merely transient enthusiasm. In the general preface to his *A Collection of Several Philosophical Writings* (1662), he still spoke with admiration of Descartes. Yet in the *Divine Dialogues* (1668) and even more severely in *Enchiridion Metaphysicum* (1671) More criticized "the superstitious admiration" for Descartes and alleged that his views led to atheism, a charge against which he had previously defended Descartes.

Not surprisingly, More's French critics accuse him of irresponsible fickleness. But if *Enchiridion Metaphysicum* is the first of More's writings to be officially an anti-Cartesian tract, the fact remains, as Descartes realized from the beginning but More only slowly, that More's leading ideas had always been in complete opposition to

Cartesianism. The central point in More's metaphysics as it is developed in *The Immortality of the Soul* (1659) and the metaphysical sections of *Divine Dialogues* and *Enchiridion Metaphysicum* is that extension is a characteristic of all substances and not, as Descartes had argued, a peculiarity of matter. Substances fall into two classes—spirits and material objects. Spirits are physically indivisible, can penetrate both other spirits and material objects, and can initiate motion; material objects are physically divisible, impenetrable, and capable of motion only when it has been communicated to them. But both spirits and material objects are extended. There are familiar objections to such an ontology; these concern, particularly, the compatibility of the two properties of being extended and being spiritual. In meeting these objections, More began by making two logical points. The first is that since we are never acquainted with essences but only with attributes, it is no objection to the extendedness of thinking beings that we "cannot see why" a being which thinks should also be extended. The second is that the intellectual separability of the properties of being extended and being spiritual is no proof of their incompatibility.

More's opponents have to show, he argued, that it is logically impossible for anything to be extended and yet to think. Most of the arguments that are supposed to establish this impossibility depend, according to More, upon the tacit identification of extension and materiality; the rest can be met by distinguishing between two forms of extension—metaphysical and physical. Metaphysical extension—pure space—is eternal, infinite, physically indivisible; physical extensions are finite, physically divisible, mutable. We can break up a particular cylinder, and we can easily imagine it not to exist, but we cannot take a piece out of space or imagine it not to exist. These properties it shares with God; indeed, space is an "obscure representation of the essence or essential presence of the divine being."

More came to see in Descartes the leader of what he calls the nullibists, who deny extension to spirits. And although Descartes had set out to defend God and immortality—this was one main reason why More approved of him—More finally concluded that nullibism is atheistic in tendency. For More the essential feature of the soul is that it initiates movement. To do this, however, it must be where body is. This is possible because unlike material objects spirits can penetrate both other spirits and material objects, contracting or expanding like Isaac Newton's "aether," as the occasion makes necessary. Thus, God, an individual mind, and a material object can all be present in the one place without losing their independence as substances. Spirit can be regarded, More argued, as a sort of fourth dimension; a body that contains a spirit has a certain "spissitude," or density of substances.

More's criticism of mechanical explanation is along the same general lines. At first, he had welcomed Descartes's mechanical explanations; by carrying ingenuity, so More thought, as far as it could be carried, they made it clear just what the limits of mechanical explanation were. But his conclusion is that mechanical explanation is never possible and that to suppose otherwise leads to atheism. (The emergence of Benedict de Spinoza from the Cartesian school encouraged More in this belief.)

A material object, he said, is nothing but a "congeries of physical monads"—that is, a collection of atomic particles. To explain how these particles are held together in solid objects, we have to introduce a nonmaterial, although spatial, spiritual agent. Equally, he argued, gravity is inexplicable in mechanical terms; mechanics—he meant, of course, Cartesian mechanics—cannot explain why a bullet once fired from a gun should ever return to Earth's surface. Even more obviously, the behavior of living organisms cannot be derived from a collection of particles.

Indeed, in order to explain any natural process, we have to refer to spirit as something additional to material particles; spirits are the true cause of all activity. This does not mean that all activity is the work of conscious rational beings. Spirit exists at various levels; "seminal forms," which are neither sensitive nor rational but are still capable of initiating motion, are responsible for actions at a level lower than animal feeling.

RELIGION AND ETHICS

More's metaphysical theories are not worked out in detail. His main interests, indeed, were religious rather than metaphysical: to defend Christianity against its three main enemies—namely, atheists, Roman Catholics, and "enthusiasts." *An Antidote against Atheism* (1653) reformulates the Ontological Argument but mainly relies upon anecdotes about animals to establish an Argument from Design and upon anecdotes about witches and apparitions to establish that spiritual forces are at work in the world. *Conjectura Cabbalistica* (1653), with the aid of the Jewish kabbalah, discerns Platonism and Cartesianism in Genesis; indeed, More expressed his regret that he had ever wasted his time on philosophy seeing that all fundamental truths are contained in the Bible. *A Brief Discourse of the Nature, Causes, Kinds and Cure of Enthusiasm* (1656) is directed against "enthusiasm," defined as "a full but false persuasion in a man that he is inspired."

More found the origin of enthusiasm in "melancholy"—that is, in a manic–depressive constitution. *The Grand Mystery of Godliness* defends the Cambridge Platonist concept of religion against Calvinists, atheists, and Roman Catholics alike; *An Antidote against Idolatry* (1674) attacks Roman Catholics. More had a special animosity against Quakers that increased in intensity when his disciple and admirer Anne Finch, Lady Conway, at whose home in Ragley, Warwickshire, he had been a frequent guest, became a convert to Quakerism.

More's *Enchiridion Ethicum* (1667), translated into English by Edward Southwell in 1690 with the appropriate title *An Account of Virtue*, was the most popular of More's writings in his own time but has since been neglected. It can be most succinctly described as a Christian version of Aristotle's *Nicomachean Ethics*, although the detail is influenced by Descartes's account of the passions and by mathematical ideals. (More set out a number of "moral axioms," which incorporate an ethical calculus.) Virtue, More argued, consists in pursuing what seems to be in accordance with right reason, but both our capacity to discover what actions accord with reason and our inclination toward those actions flow from a special "boniform" faculty. Reason itself cannot incite action; virtuous action can be instigated only by the passional side of our nature. The ultimate ground of all virtue is intellectual love. Thus, More hoped to weld the Christian doctrine of love and the Aristotelian doctrine of intellectual activity into a single ethical system.

INFLUENCE

More devoted the last seven years of his life to translating his English works into Latin in the hope of attracting wider interest on the Continent. They caught the attention of Gottfried Wilhelm Leibniz, but although he took an occasional phrase from More, he was interested in him mainly as a representative of the sort of view he particularly wished to avoid. In fact, More, the only one of the Cambridge Platonists to publish at all extensively, quite failed in what he conceived as his main task—to halt the advance of the mechanical worldview. More's metaphysics, however, had a considerable influence on Newton even if mathematicians, not metaphysicians, were Newton's principal masters. Newton did not refer explicitly to More—the Cambridge group almost never referred to one another—but the resemblances are conspicuous. Newton was taught mathematics at Grantham, More's birthplace, by a former pupil of More's; Newton's correspondence reveals that he and More stood close to one another.

See also Cambridge Platonists.

Bibliography

WORKS BY MORE

More's *Philosophical Poems* are reprinted in Alexander Balloch Grosart, *The Complete Poems of Henry More* (Blackburn, U.K, 1878). Geoffrey Bullough, *Philosophical Poems of Henry More* (Manchester, U.K.: University of Manchester, 1931), is a selection with a valuable introduction and notes. More's main philosophical writings are included in *A Collection of Several Philosophical Writings* (London, 1662) and his theological writings in *Theological Works* (London, 1708); the Latin version, *Opera Omnia* (London, 1675–1679), contains in addition a number of controversial pamphlets. Sections of the *Enchiridion Metaphysicum* were translated by Joseph Glanvill in his *Saducismus Triumphatus* (London, 1681) and are included in Flora Isabel MacKinnon, *Philosophical Writings of Henry More* (New York: Oxford University Press, 1925), with a useful bibliography and expository essays. Edward Southwell's translation of *Enchiridion Ethicum* has been reprinted by the Facsimile Text Society (New York, 1930). The correspondence with Descartes is partly translated by Leonora D. Cohen in *Annals of Science* 1 (1) (1936): 48–61, and is included in Geneviève (Rodis-)Lewis, *Correspondance avec Arnaud et Morus* (Paris, 1953).

WORKS ON MORE

For works on More see the bibliography under the "Cambridge Platonists" entry. See also Richard Ward, *The Life of Dr. H. More* (London, 1710); M. F. Howard has edited this life, but his introduction is not reliable (London, 1911). See also Marjorie Nicolson, ed., *The Conway Letters* (London: Oxford University Press, 1930), and Paul Russell Anderson, *Science in Defense of Liberal Religion* (New York and London: Putnam, 1933). For More's relation to Newton see Edwin Arthur Burtt, *The Metaphysical Foundations of Modern Physical Science* (London: Kegan Paul Trench and Trubner, 1925; rev. ed., 1950), and Alexandre Koyrė, *From the Closed World to the Infinite Universe* (Baltimore: Johns Hopkins Press, 1957). On the opposite side see Edward William Strong, *Procedures and Metaphysics* (Berkeley: University of California Press, 1936), and Stephen Edelston Toulmin, "Criticism in the History of Science: Newton on Absolute Space, Time and Motion," in *Philosophical Review* 68 (1) (1959): 1–30, and (2) (1959): 203–228.

Other Recommended Sources

Almond, Philip C. "The Journey of the Soul in Seventeenth-Century English Platonism." *History of European Ideas* 13 (6) (1991): 775–791.

Armstrong, Robert L. *Metaphysics and British Empiricism*. Lincoln: University of Nebraska Press, 1970.

Burnham, Frederic B. "The More-Vaughan Controversy: The Revolt against Philosophical Enthusiasm." *Journal of the History of Ideas* 35 (1974): 33–49.

Coudert, Allison. "A Cambridge Platonist's Kabbalist Nightmare." *Journal of the History of Ideas* 36 (1975): 633–652.

Daniel, Stephen. “Berkeley’s Pantheistic Discourse.” *International Journal for Philosophy of Religion* 49 (3) (2001): 179–194.

Gabbey, Alan. “Philosophia Cartesiana Triumphata: Henry More.” In *Problems of Cartesianism*, edited by Thomas M. Lennon, John M. Nicholas, and John W. Davis. Montreal: McGill Queens, 1982.

Hutton, Sarah. “Henry More and Anne Conway on Preexistence and Universal Salvation.” In *Mind Senior to the World*, edited by Marialuisa Baldi. Milano: FrancoAngeli, 1996.

Jacob, Alexander. “The Metaphysical Systems of Henry More and Isaac Newton.” *Philosophia Naturalis* 29 (1) (1992): 69–93.

Osler, Margaret J. “Triangulating Divine Will: Henry More, Robert Boyle, and Rene Descartes on God’s Relationship to the Creation.” In *Mind Senior to the World*, edited by Marialuisa Baldi. Milano: FrancoAngeli, 1996.

Patrides, C. A. *The Cambridge Platonists*. Cambridge, U.K.: Cambridge University Press, 1980.

Power, J. E. “Henry More and Isaac Newton on Absolute Space.” *Journal of the History of Ideas* 31 (1970): 289–296.

Reid, Jasper. “Henry More on Material and Spiritual Extension.” *Dialogue* 42 (3) (2003): 531–558.

Rogers, G. A. J. “Hobbes’s Hidden Influence.” In *Perspectives on Thomas Hobbes*, edited by G. A. J. Rogers and Alan Ryan. New York: Clarendon Oxford, 1988.

Sprague, Elmer. “Hume, Henry More and the Design Argument.” *Hume Studies* 14 (1988): 305–327.

Staudenbaur, C. A. “Galileo, Ficino, And Henry More’s Psychathanasia.” *Journal of the History of Ideas* 29 (1968): 565–578.

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

MORE, THOMAS

(1478–1535)

Sir Thomas More, later canonized St. Thomas More, was a lawyer and statesman rather than a philosopher. More was born the son of a London lawyer who later became a judge. He was educated at St. Anthony’s School and was appointed a page in the household of Archbishop (later Cardinal) Morton, who sent him to Canterbury Hall, Oxford, in the early 1490s. More left without a degree to study at New Inn and Lincoln’s Inn in London. His lectures dealt not only with law but also with St. Augustine’s *City of God*. He early composed various English poems and Latin epigrams that were not printed for years. However, a Latin translation of four Greek dialogues of Lucian appeared in 1506, and an English translation of the Latin life of his model, Giovanni Pico della Mirandola, in 1510. Increasingly involved in public affairs, More became a member of Parliament in 1504, beginning the career that led to the well-known events of his chancellorship and his martyrdom. By the time of the *Utopia* (1516), he had long since mastered Greek and enjoyed the friendship of such humanists as Desiderius Erasmus, Thomas Linacre, William Grocyn, John Colet, Cuthbert Tunstall, and St. John Fisher.

PHILOSOPHICAL ORIENTATION

With respect to his philosophy, Thomas More belonged very much to the early or Erasmian period of the English Renaissance in his emotional and intellectual attitudes—toleration of eclecticism, search for simplicity, stress on ethics, return to Greek sources, and desire for reform: social, political, educational, religious, and philosophical. These traits appear not only in his highly imaginative and durably significant creation, *Utopia*, but also in his most pertinent pronouncements in real life. The latter may be divided into two philosophical periods, roughly separated by the year 1521, the year of publication of Henry VIII’s *Defense of the Seven Sacraments* (*Assertio Septem Sacramentorum*), which More undertook to defend by his pseudonymous diatribe (1523) against Martin Luther’s strictures.

During his first period, in his justly famous letters to Martin Dorp (1515), to the University of Oxford (1518), and to a monk (1519–1520), More opted for a simplified logic, the study of all Aristotle’s works in Greek with their classical Greek commentaries, and the mastery of the Greek New Testament and Greek Fathers as well as the pagan classics in the original language. He praised the Aristotelian paraphrases of Jacques Lefèvre d’Étaples and, in a letter to Erasmus (May 26, 1520), expressed complete agreement with Juan Luis Vives’s *False Dialecticians* (*Pseudodialectici*). His attack on contemporary Schoolmen centered on their preoccupation with logic, the universals, and a mere fragment of the Aristotelian corpus.

In his second, controversial period, More rose to the defense of Thomas Aquinas and the scholastic theologians, whose doctrine he showed to agree with that of the earlier church. However, since the interest of these works, even of *A Dialogue of Comfort against Tribulation* (1534), is almost entirely theological, there is no need to dwell on them, except to point out that he held the common scholastic views on the mutual relationship, harmony, and assistance between reason and revelation, with philosophy as the propaedeutic to theology and as the handmaid of theology. This synthesis appears in a fundamental form even on the island of Utopia, where ethical norms are bolstered by religious truths and where the

true religion can prevail in an atmosphere of free and calm reasoning.

UTOPIA

Since *Utopia* is More's major, or at least most influential, writing, its philosophical elements will be discussed in detail.

BACKGROUND. Renaissance thinkers usually held that there were four great philosophical schools: Platonism, Aristotelianism, Stoicism, and Epicureanism, which differed mainly according to their opinions of the *summum bonum*. The Christianization of Aristotle was accomplished in the thirteenth and fourteenth centuries by the Schoolmen, and that of Plato in the fifteenth century by Marsilio Ficino and other humanists. Stoicism had found expression in almost boundless humanistic admiration for the writings of Seneca and especially Cicero before reaching definite formulation later in the Christian Stoicism of Justus Lipsius. It was therefore inevitable that humanistic attempts, if only rhetorical ones, should be made to Christianize Epicurus, too. The latter's rehabilitation had been much accelerated in the early fifteenth century by Ambrogio Traversari's Latin translation of his life by Diogenes Laërtius. Lorenzo Valla had set forth Epicurus's doctrine favorably in *De Voluptate ac de Vero Bono* (*Pleasure and the True Good*). Finally Erasmus undertook his thorough baptism in *De Contemptu Mundi* (*The Contempt of the World*, written c. 1490) and the colloquy *The Epicurean* (published 1533). In both these works, Erasmus manipulated the concept of pleasure and the principles of selection to establish a Christian Epicureanism.

EPICUREANISM IN *UTOPIA*. More's main sources for classical Epicureanism were undoubtedly the *Lives* of Diogenes Laërtius and the *De Finibus* of Cicero, with minor borrowings from Seneca, Quintilian, Lucian, and Aulus Gellius. The "Christian" modifications already introduced by such humanists as Lorenzo Valla and Erasmus should not be minimized. The preoccupation of Renaissance men with the problem of pleasure is evident from the many humanistic treatments of the subject, including that by Ficino. Consequently Epicurus and Epicureanism are here viewed not according to their historical reality but according to the light in which they appeared to Thomas More through his reading and conversation.

In spite of the great to-do in the *Utopia* about the philosophy of pleasure and in spite of the deliberate but superficial rejection of Stoicism, the emphasis on virtue and virtuous living is disproportionate, even extraordinary, and therefore suspicious. This respect for Stoicism also becomes explicit in the stress on the guidance of nature, the assumed existence of natural law, and the natural community of humankind.

There are several contacts between Utopian and Epicurean hedonism. The most evident, naturally, is the exaltation of pleasure as the *summum bonum*, to which all human activities, including the operations of the virtues, are directed and subordinated. But the term *pleasure* (*coluptas*) is so manipulated in the *Utopia* that it embraces everything from scratching an itch to enjoying eternal bliss with God. Like Epicurus, the Utopians hold to both kinds of pleasure: pleasure as a state and pleasure as motion. Hence health for them is a true pleasure. Like Epicurus, they belittle neither the joy arising from conferral of a benefit, nor the testimony of a good conscience as the reward for just deeds, nor the importance of mental pleasures. There is a common emphasis with Epicurus on the simple life, which in Utopia leads to the ridicule of false, unnatural delight in fine clothing, noble ancestry, glittering jewelry, gold and silver, gambling, and hunting. Perhaps the most important connection is the enunciation of the principles of selection; the single positive criterion is that a pleasure be natural—a criterion recognized as so obscure that it is delimited by three negative norms: that no pain follow the pleasure chosen, that no greater pleasure be lost, and that no social harm result.

DIVERGENCES FROM CLASSICAL EPICUREANISM. The departures from the postulates of classical Epicureanism are so radical that the Utopian philosophy in action can be labeled Epicurean, or even hedonistic, only in the broadest sense. For example, good Utopians must believe in the providence of God, the immortality of man's soul, and divine retribution in a future life. These Utopian principles are taken not from Epicurus but from More's great favorite, Plato, especially his *Laws*. Utopian ascetics, with their hope of reward in a future life, would be ridiculous to Epicurus. The Platonic origin of Utopian communism also is evident, for Epicurus thought that the holding of property in common by friends implied mutual mistrust. Minor points of divergence are the emphasis upon marriage (in contrast with its disapproval by Epicurus in spite of his traditional devotion to his parents), upon euthanasia (in comparison with Epicurus's denial of suicide even to the blind), and upon learning (Epicurus urged his disciples to fly from learning in the swiftest ship available). Utopians love their gardens, but for practical rather than philosophical purposes, so that,

surprisingly, no reference is made in *Utopia* to the connection between Epicurus and gardens.

RAPHAEL HYTHLODAEUS. The unconscious pull of Platonism and Stoicism, not to mention Christianity, is too great to allow a full-fledged Epicureanism in Utopia. This is perfectly consistent, however, with the engrossing character of the main narrator, Raphael Hythlodaeus, who is a philosopher by nature and profession and interjects mild expressions of disapproval of Utopian hedonism. He is unattached: His only commitment is to freedom, truth, and justice. Negligent in dress, he has divested himself of the cares of riches by giving his patrimony to his relatives. He now lives as he pleases (according to Cicero's definition of freedom), and he must speak his mind openly. In spite of being accused of too great speculativeness and idealism by Thomas More, he travels and searches for something quite practical: the good state and the good citizen. In this emphasis on the useful, and in his return to the sources (especially the Greek), Hythlodaeus is at one with the early English, as well as the northern, Renaissance. In his chosen field of philosophy, he finds nothing of value in Latin except Seneca and Cicero. But he is far from being narrow. The great books in Greek that he carries with him include Plato and Plutarch, as well as Aristotle and Theophrastus, dramatists, poets, historians—and Lucian. Devotion to Lucian undoubtedly helped to mark More's philosophical character as his friends saw him—as "another laughing Democritus." More's emphasis upon the Greek sources in medicine (Galen and Hippocrates) and science (Aristotle's *Meteorology*) makes him, in a sense, an unwitting scientific reactionary.

PLATO'S INFLUENCE. Of all the Greek authors, Plato is cited most frequently in the *Utopia* proper and in its preliminary materials. This is hardly surprising, since its true title may be translated as *The Best Order of Society* (*De Optimo Reipublicae Statu*). More is indebted, however, as much to Plato's *Laws* as to his *Republic*. His obvious but modified borrowings from Plato are dialogic form, but with a monologue in Book II; communism, which he broadens to embrace a whole nation, not merely an elite class; preeminence of learning, with transformation of the philosopher-king into the scholar-governor; the almost complete equality of men and women; and the connections between goodness and religion. The differences are radical: Utopia is a casteless democracy, not an aristocracy; and the family, not a ruling class with common wives and children, is the basic social and political unit. It is significant that More also briefly introduces the Aristotelian objections to communism of property.

PLEASURE AND THE BEST SOCIETY. It is a tribute to More's rhetoric (not philosophy) that the unwary reader is left under the impression that the Utopians espouse thoroughgoing hedonism. But this does not involve merely a humanistic *jeu d'esprit* or even a literary tour de force, for pleasure is related intimately to the main subject of the *Utopia*, the best society. The best society is one whose aim is the temporal well-being or happiness—or pleasure, as defined and described in Utopian terms—of all the citizens, not only of the rich or of the well-born. All are to share equally and equitably in all the good things—or pleasures—of this life and this world: food, clothes, houses, work, play, sleep, and education. More bridges the gap between Utopian philosophy and Utopian communism by the use of the basically Aristotelian phrase "the matter of pleasure" (*materia voluptatis*). Vital commodities (food, clothing, housing) constitute the pleasurable *matter*, which must be determined by a *form* (either private ownership or common possession). The Utopians have chosen communism, not private property, to bring the greatest pleasure to the whole nation. Only in this way will justice be introduced into an unjust society. In this at least theoretical espousal of communism, More agreed with Erasmus and many fellow humanists.

WEAKNESSES. On the debit side of the *Utopia* might be listed the deliberately static nature of this ideal society and the failure to recognize the individual person and his basic instincts, liberties, and even imperfections. The removal of all struggle and all insecurity would logically and psychologically lead to the prayer: "Give me something to desire."

INFLUENCE. The major influence of the *Utopia* lies not in its philosophic hedonism, with its concomitant communism, but in its establishment of a pattern for ideal commonwealths. Historically the type proliferated into a thousand different forms that can be found discussed in bibliographies and commentaries. In particular, the *Utopia* itself set an example for what might be termed the philosophical utopia that continued well into the eighteenth century. The most notable productions are Francis Bacon's *New Atlantis*, Tommaso Campanella's *City of the Sun*, and Samuel Johnson's *Rasselas*.

See also Aristotelianism; Aristotle; Augustine, St.; Bacon, Francis; Colet, John; Communism; Diogenes Laertius; Erasmus, Desiderius; Epicureanism and the Epicurean

School; Epicurus; Ficino, Marsilio; Galen; Hedonism; Hippocrates and the Hippocratic Corpus; Johnson, Samuel; Lipsius, Justus; Luther, Martin; Pico della Mirandola, Count Giovanni; Plato; Platonism and the Platonic Tradition; Pleasure; Plutarch of Chaeronea; Renaissance; Seneca, Lucius Annaeus; Stoicism; Theophrastus; Thomas Aquinas, St.; Universals, A Historical Survey; Utopias and Utopianism; Valla, Lorenzo; Vives, Juan Luis.

Bibliography

The best modern edition of More's Latin and English writings is the Yale edition: *The Complete Works of St. Thomas More*, R. S. Sylvester, executive ed. (New Haven, CT, 1963–1997), 21 volumes. Vol. IV, *Utopia* (New Haven, CT: Yale University Press, 1965), is edited by Edward Surtz and J. H. Hexter. Also see the Selected Works edition (New Haven: Yale University Press, 1961–), planned in seven volumes, of which two have appeared: *Selected Letters*, E. F. Rogers, ed. (1961), and *Utopia*, Edward Surtz, ed. (1964). E. F. Rogers had previously edited More's *Correspondence* (Princeton, NJ: Princeton University Press, 1947). Only the first two volumes of a contemplated seven-volume modern version of More's *English Works* (1557), edited by W. E. Campbell, were issued (London: Eyre and Spottiswoode, 1927–1931).

More's best biographers are the earliest: William Roper, *The Lyfe of Sir Thomas Moore*, edited by E. V. Hitchcock (London: Oxford University Press, H. Milford, 1935); and Nicholas Harpsfield, *The Life and Death of St. Thomas Moore*, edited by E. V. Hitchcock and R. W. Chambers (London: Oxford University Press, H. Milford, 1932). The best modern life is still R. W. Chambers, *Thomas More* (London, 1935), to be supplemented by E. E. Reynolds, *Saint Thomas More* (London: Burns and Oates, 1953).

Bibliographical data can be found in *St. Thomas More: A Preliminary Bibliography … to the Year 1750*, compiled by R. W. Gibson, with a bibliography of Utopiana compiled by R. W. Gibson and J. M. Patrick (New Haven, CT: Yale University Press, 1961). Also see F. and M. P. Sullivan, *Moreana, 1478–1945* (Kansas City, MO: Rockhurst College, 1945). In 1963, the international Amici Thomae Mori began publication of *Moreana: Bulletin Thomas More* (Angers).

Illuminating studies of the background can be found in W. E. Campbell, *Erasmus, Tyndale, and More* (London: Eyre and Spottiswoode, 1949); Fritz Caspari, *Humanism and the Social Order in Tudor England* (Chicago: University of Chicago Press, 1954); Pearl Hogrefe, *The Sir Thomas More Circle* (Urbana: University of Illinois Press, 1959); R. P. Adams, *The Better Part of Valor: More, Erasmus, Colet, and Vives on Humanism, War, and Peace* (Seattle: University of Washington Press, 1962); and especially G. Marc'hadour, *L'univers de Thomas More* (Paris: Vrin, 1963), corrected and supplemented currently in *Moreana*.

The principal interpretations of *Utopia* are those by Karl Kautsky, *Thomas More and His Utopia* (1888), translated by H. J. Stenning (reprinted, New York: Russell and Russell, 1959); H. W. Donner, *Introduction to Utopia* (London: Sidgwick and Jackson, 1945); Russell Ames, *Citizen Thomas More and His Utopia* (Princeton, NJ: Princeton University Press, 1949); J. H. Hexter. *More's "Utopia": The Biography of an Idea* (Princeton, NJ: Princeton University Press, 1952); Edward Surtz, *The Praise of Pleasure: Philosophy, Education, and Communism in More's "Utopia"* (Cambridge, MA: Harvard University Press, 1957); and Edward Surtz, *The Praise of Wisdom: A Commentary on the Religious and Moral Problems and Backgrounds of St. Thomas More's "Utopia"* (Chicago: Loyola University Press, 1957).

The fate of the utopia as a literary form can be followed in Richard Gerber, *Utopian Fantasy* (London: Routledge and Kegan Paul, 1955); J. O. Hertzler, *The History of Utopian Thought* (New York: Macmillan, 1923); and G. R. Negley and J. M. Patrick, eds., *The Quest for Utopia* (New York: Schuman, 1952).

See also Michael Jackson, "Imagined Republics: Machiavelli, Utopia, and *Utopia*," *Journal of Value Inquiry* (34[4] [2000]: 427–437; Anthony Kenny, *Thomas More* (Oxford: Oxford University Press, 1983).

Edward Surtz, S.J. (1967)

Bibliography updated by Tamra Frei (2005)

MORGAN, AUGUSTUS DE

See *De Morgan, Augustus*

MORGAN, C. LLOYD

(1852–1936)

C. Lloyd Morgan, an English biologist and philosopher, was born in London. His early education "was almost exclusively literary," but he later became attracted to scientific studies, attended the Royal School of Mines, and received a diploma in metallurgy. His deepest interest, however, was in the bearing of science on philosophical issues. This interest was given encouragement and direction by T. H. Huxley, under whom he studied biology. Henceforth, Morgan's vocation was to be that of an investigator of "borderland problems of life and mind" and the expositor of a philosophy of "emergent evolution." After teaching for five years at a small college near Cape Town, South Africa, he was appointed in 1884 to the chair of geology and zoology at University College, Bristol. When the college received a university charter in 1909, Morgan agreed to serve temporarily as its first vice-chancellor. At his own request, however, he resigned the next year and resumed his chair, now designated the chair of psychology and ethics. He retired in 1919. During his career at Bristol, Morgan devoted himself to the study of animal psychology and published such books as *Animal Life and Intelligence, Habit and Instinct, Animal Behavior,* and *Instinct and Experience.*

When he was elected a fellow of the Royal Society in 1899, he became the first person to be thus honored for scientific work in psychology. After his retirement he was invited to deliver the Gifford Lectures and used the occasion to expound his philosophical ideas, which subsequently appeared in *Emergent Evolution* and *Life, Mind, and Spirit.* Two other works, *Mind at the Crossways* and *The Emergence of Novelty,* contain elaborations of his position.

Morgan's psychological studies had a Darwinian background. Accepting the view that evolution is a continuous process, he sought to trace the development of mental characteristics in the world of living things. The focal point of his investigations was the behavior of those organisms that showed some capacity to learn from experience. He contended that the rudiments of intelligence are to be found wherever learning results from "the method of trial and error"—a phrase that he coined in 1894. Much of his experimental work was designed to show how this method is employed, even by relatively simple forms of life. Unlike his predecessors in animal psychology, Morgan was alert to the dangers of using casual reports of animal behavior, especially reports from untrained observers. He urged the importance of a methodological "law of parsimony," according to which we should never interpret what an animal does as the outcome of a higher psychical power if the action "can be interpreted as the outcome of the exercise of a power which stands lower in the psychological scale." Morgan's experiments usually were not strictly laboratory ones but involved artificially produced situations in the natural habitat of animals. His accurate and detailed observations of their behavior in these situations, however, gave comparative psychology a new scientific status.

The conceptual background of Morgan's work was neither mechanistic nor finalistic. He rejected the view that biological processes are to be understood in physicochemical terms and that physiology can give an adequate account of animal behavior. Radical behaviorism was likewise unacceptable to him. On the other hand, he rejected the view that teleology is operative throughout the living world and that even reflex action and instinctive responses must be explained teleologically.

In *Instinct and Experience* Morgan criticized Henri Bergson's teleological speculations. Morgan's own position, which he described as "naturalism," was that in all behavior there occurs an "unrestricted concomitance" of physical and psychical events. Hence, each behavior episode is susceptible of interpretation in both physiological and psychological terms. There are two stories to be told, each throwing light on the other, "but neither story as such *makes* the other what it is."

Philosophically, Morgan adopted the hypothesis that the twofold story was really about *one* natural order of events. Moreover, that one order of events has a progressive natural history designated by the word *evolution.* An adequate description of this process requires us to recognize that evolution has not been uniformly continuous, as Charles Darwin believed, but has involved from time to time major discontinuities or "critical turning points." These turning points are marked by the abrupt appearance of certain phenomena that Morgan called *emergents,* a term used by G. H. Lewes in 1874. An emergent (1) supervenes upon what already exists, (2) arises out of what already exists, (3) is something genuinely new in the history of the universe, (4) occurs in a manner that is unpredictable in principle since it conforms to no general laws, and (5) cannot be naturalistically explained but must be accepted "with natural piety." The successive emergents in the panorama of evolution mark stages of progress from lower to higher. Hence, Morgan followed Samuel Alexander in picturing the totality of nature as "a pyramidal scheme."

The full significance of emergent evolution cannot be grasped, however, as long as one remains at the level of "a philosophy based on the procedure sanctioned by the progress of scientific thought." It was essential, Morgan thought, to construct a metaphysical system within which the naturalistic version of evolution could be set. This system would formulate certain fundamental concepts and presuppositions by whose aid an "ultimate explanation" of the evolutionary process could be given. Nothing affirmed in this constructive scheme was to be at variance with science, but it would "complete the otherwise incomplete delivery of strictly scientific thought."

A necessary basic presupposition of the system Morgan proposed was the existence of a physical world that "is nowise dependent on being perceived or thought of by any human or sub-human mind." Since no conclusive proof of this contention had ever been given, it was simply "accepted under acknowledgment." Morgan then elaborated a psychophysiologically oriented theory of how organisms perceive the external world. Physical events exert an "advenient influence" on the sense receptors of organisms. By virtue of their psychical power, the organisms respond by referring the signs arising within the psychophysical system to regions of physical space in a process Morgan called "projicient reference." The result is an emergent object correlated with the external event in such a way as to be biologically useful to the organism.

Morgan's second presupposition was that the pyramid of emergent evolution is a hierarchy of kinds of relatedness. Four basic concepts are needed to unfold its consequences—stuff, substance, quality, and property. The ultimate stuff consists of psychophysical events, and the mode of their relatedness in a given system is that system's substance. Each system has intrinsic qualities grounded in its substance and extrinsic properties grounded in its relation to other systems. Besides the emergents there are *resultants,* or phenomena that are repetitive, predictable, and the source of quantitative continuity. Emergence generates progress in continuity, but through resultants there is continuity in progress.

The third presupposition that Morgan acknowledged was the universal correlation of physical and psychical events. He recognized a similarity between his system and that of Benedict de Spinoza in this respect, yet Morgan's view that "mind" is "a quality emergent at a high level of evolutionary advance" would have been quite unacceptable, or possibly unintelligible, to Spinoza. Even that from which mind in this sense emerges—the pervasive psychical correlate—is scarcely to be compared with a Spinozistic attribute.

The last presupposition introduced by Morgan affirmed that a directing activity, otherwise called "spirit" or "God," is manifested everywhere. Thus, "the whole course of events subsumed under evolution is the expression of God's purpose," which embraces all that has been and all that will be brought about in the course of evolutionary advance. This postulate can be neither proved nor disproved but only adopted to satisfy the need for an ultimate explanation of things.

Morgan's philosophy of evolution gave wide currency to the idea of emergence. Yet when compared with later discussions, his treatment of the idea lacks precision. He was not a close reasoner, and his speculative scheme was much less carefully worked out than that of Alexander, to whom he was indebted. A hostile critic might well question Morgan's policy of "acknowledging," rather than arguing for, important principles in his system. And, although he opposed Darwinism by insisting that evolution is "jumpy" and not continuous, each jump is, in Morgan's view of evolution, a mystery, unexplained and inexplicable except, perhaps, to God.

See also Alexander, Samuel; Animal Mind; Bergson, Henri; Darwinism; Emergence; Emergent Evolutionism; Huxley, Thomas Henry; Lewes, George Henry; Spinoza, Benedict (Baruch) de; Teleology.

Bibliography

WORKS BY MORGAN

Animal Life and Intelligence. London: Arnold, 1890.
An Introduction to Comparative Psychology. London: W. Scott, 1894.
Habit and Instinct. London and New York: Arnold, 1896.
Animal Behaviour. London: Arnold, 1900.
Instinct and Experience. New York: Macmillan, 1912.
Emergent Evolution. London: Williams and Norgate, 1923.
Life, Mind, and Spirit. London, 1926.
Mind at the Crossways. London: Williams and Norgate, 1929.
The Animal Mind. London: Arnold, 1930.
The Emergence of Novelty. London: Williams and Norgate, 1933.

WORKS ON MORGAN

McDougall, William. *Modern Materialism and Emergent Evolution.* New York: Van Nostrand, 1929.
MacKinnon, Flora I. "The Meaning of 'Emergent' in Lloyd Morgan's 'Emergent Evolution.'" *Mind* 33 (1924): 311–315.

T. A. Goudge (1967)

MORGAN, LEWIS HENRY
(1818–1881)

Lewis Henry Morgan was an American anthropologist and social philosopher. After graduating from Union College in 1840, he practiced law in Rochester, New York, from 1844 to 1864, but he devoted much of his time to anthropological research, which eventually became his exclusive interest. One of the most celebrated American scholars of his time, Morgan was elected a member of the National Academy of Sciences in 1875 and president of the American Association for the Advancement of Science in 1879. The results of his investigations into the life of various Indian tribes appeared in his *League of the Ho-dé-no-sau-nee or Iroquois* (Rochester, NY, 1851) and his later work, *Systems of Consanguinity and Affinity* (Washington, DC, 1871); these two books were hailed as pioneering achievements of the first order in the study of kinship systems by even the most outspoken of his critics.

Morgan's aim was not merely to describe how different civilizations had evolved; he wished to elicit from their history a general pattern of institutional progress. In his most ambitious work, *Ancient Society* (New York, 1877), Morgan sought to establish that human history falls into three main stages—savagery, barbarism, and civilization—and that each stage reflects a close correlation between economic and cultural achievements. Savagery was the period before pottery; barbarism was the ceramic era; civilization began with writing and the pho-

netic alphabet. The first two periods are further subdivided, and each subperiod is defined in terms of its characteristic technological innovations. The discovery of fire and the beginning of fishing, for example, are characteristic of the second subperiod of savagery, the invention of the bow and arrow of its third subperiod.

Although Morgan shared the view of his Swiss contemporary and fellow anthropologist Johann Jakob Bachofen that society had emerged from a state of primitive communism, and also accepted the Bachofen hypothesis of matrilineal descent, he had little interest in ancient myths and religions. His principal attention was focused on technological factors, kinship systems, and property systems, and their relations to social and political institutions. In spite of gaps and distortions, Morgan's account of the growth of civilization has been considered by so severe a critic of his ethnological theories as Robert H. Lowie to be a comprehensive scheme of cultural wholes far beyond anything attempted up to that time. Lowie has written, "Morgan's *Ancient Society* was a synthesis of sociological material that for the first time brought together material on Australian and American natives, on ancient Greece and Rome; and all this in an orderly arrangement prescribed by an evolutionary doctrine" (*The History of Ethnological Theory,* London, 1937, p. 56).

Moreover, *Ancient Society* speaks for a distinct social philosophy and philosophy of history. The collation and comparison of human institutions, inventions, and discoveries convinced Morgan of humankind's unity of origin, of the similarity of human wants in different societies at comparable stages of advancement, and of the uniformity in the operations of the human mind in similar conditions of society. He formed the view that the human race was "one in source, one in experience and one in progress" (*Ancient Society,* p. vi). The problem that preoccupied Morgan in his historical researches was the existence of social and economic inequality. He could not conceive that "a mere property career" was the final destiny of humankind. Man's obsession with private property, he felt, was only a transient stage of human civilization. For if it was not, it was bound to lead to society's self-destruction. If progress was to be the law of the future as it had been of the past, property would have to be diffused and if necessary controlled, so that "democracy in government, brotherhood in society, equality in rights and privileges, and universal education" would foreshadow the next higher plane of society, "to which experience, intelligence and knowledge are steadily tending" (*Ancient Society,* p. 552).

Morgan recognized that civilization could be aggressive as well as progressive. But his theory of social evolution has nothing in common with such imperialist notions as Rudyard Kipling's concept of the white man's burden. Progress, Morgan insisted, echoing Herder, is inherent in all cultures, civilized or not, and each has to advance along its own lines. Culture is a process, not an administrative imposition.

Although Morgan's theories were invoked by Karl Marx and by Friedrich Engels (notably in his *Origin of the Family, Private Property and the State*) in support of their interpretation of history, Morgan's social message bears only superficial similarities with Marxist doctrines. Nonetheless, the optimistic flavor of his evolutionism had a powerful appeal to social reformers. At the same time this very quality made it suspect to the uncommitted social scientist.

See also Bachofen, Johann Jakob; Culture and Civilization; Engels, Friedrich; Herder, Johann Gottfried; Marx, Karl; Philosophy of Social Sciences.

Bibliography

ADDITIONAL WORKS BY MORGAN

Diffusion against Centralization. Rochester, NY, 1852.

The American Beaver and His Works. Philadelphia: Lippincott, 1868.

Houses and House Life of the American Aborigines. Washington, DC: Government Printing Office, 1881.

Pioneers in American Anthropology: The Bandelier-Morgan Letters. Edited by Leslie A. White. Albuquerque: University of New Mexico Press, 1940.

WORKS ON MORGAN

Childe, V. Gordon. *Social Evolution.* New York: Schuman, 1951.

Colson, Elizabeth. *Tradition and Contract: The Problem of Order.* Chicago: Aldine, 1974.

Fortes, Meyer. *Kinship and the Social Order: The Legacy of Lewis Henry Morgan.* Chicago: Aldine, 1970.

Lowie, Robert H. "Evolution in Cultural Anthropology." *American Anthropologist* 48 (1946): 223–233.

Lowie, Robert H. *The History of Ethnological Theory.* New York: Farrar and Rinehart, 1937.

Resek, Carl. *Lewis Henry Morgan, American Scholar,* Chicago: University of Chicago Press, 1960. Contains a full bibliography.

Stern, Bernhard J. *Lewis Henry Morgan, Social Evolutionist.* Chicago: University of Chicago Press, 1931.

White, Leslie A. "Evolutionism in Cultural Anthropology." *American Anthropologist* 53 (1951): 11–18.

White, Leslie A. "Morgan's Attitude toward Religion and Science." *American Anthropologist* 46 (1944): 218–230.

Frederick M. Barnard (1967)
Bibliography updated by Philip Reed (2005)

MORGAN, THOMAS

(d. 1743)

Thomas Morgan, the Welsh deist, dissenting minister, doctor of medicine, freethinker, and religious controversialist, was born of a poor family but received a free education from the Reverend John Moore, a dissenter. Morgan was ordained in 1714 and became minister of Burton two years later and subsequently of Marlborough; in 1720 he was dismissed from this last post for his growing unorthodoxy. He then took up the study of medicine and produced several books on that subject—*Philosophical Principles of Medicine* (1725), *The Mechanical Practice of Physic* (1735), *Letter to Dr. Cheyne in defence of the "Mechanical Practice"* (1738).

Morgan is chiefly remembered, however, for his deistical tracts, or "Christian deistical," as he preferred to call them, in which he described himself as "M.D. and Moral Philosopher." *The Moral Philosopher, in a Dialogue between Philalethes, a Christian Deist, and Theophanes, a Christian Jew* (1737) is his major work. Controversy produced two further works under the same title, the second of 1739, subtitled "Being a farther Vindication of Moral Truth and Reason," and the third of 1740, subtitled "Superstition inconsistent with Theocracy." In 1741 he published *A Vindication of the Moral Philosopher; Against the False Accusations, Insults, and Personal Abuses, of Samuel Chandler, Late Bookseller and Minister of the Gospel.*

In general, Morgan was a rationalist espousing the five Common Notions of Lord Herbert of Cherbury. He was also one of the pioneers of historical criticism of the Bible, particularly of the Pentateuch, and was considerably influenced by John Toland and to some extent by Thomas Chubb. The latter's advocacy of free will, however, he strongly attacked in 1727 in *A Letter to Mr. Thomas Chubb, occasioned by his "Vindication of Human Nature"* and in 1728 in *A Defence of Natural and Revealed Religion.*

Morgan believed in the corruption of human nature and defended suicide for the "weary or satiated with living." His criticism of the Scriptures centered on the fact that so many different interpretations are possible and are accepted by so many different and sincere believers. Traditional religion, therefore, is not infallible but only probable, as is all history. Priestcraft, which instituted superstition, enthusiasm, and finally persecution, is the culprit for the erroneous notion of the infallibility of a catholic church. Reason and tolerance are the only cures.

See also Deism.

Bibliography

Additional works by Morgan include *A Collection of Tracts … occasioned by the late Trinitarian Controversy* (1725); *A Philosophical Dissertation upon Death. Composed for the Consolation of the Unhappy* (1732); *The History of Joseph Considered … by Philalethes* (1744). See also Sir Leslie Stephen's *History of English Thought in the Eighteenth Century* (London: Putnam, 1876; the paperback edition, 2 vols., New York: Harcourt Brace, 1963, follows the revised edition of 1902), and the general bibliography under the Deism entry.

Ernest Campbell Mossner (1967)

MORITZ, KARL PHILIPP

(1756–1793)

Karl Philipp Moritz, German novelist, man of letters, and aesthetician, was born to poor and radically Quietist (Protestant) parents. Moritz started his career as an apprentice hatmaker at the age of twelve and ended up as an intimate of Johann von Goethe, Friederich Schiller, and Johann Georg Herder, and as professor of archaeology and aesthetics at the Berlin academy of art as well as a member of the Prussian Academy of Sciences. A prolific writer, his works include the psychological novel *Anton Reiser* (1785–1790), a fictionalized account of his own passage from his narrow religious origins to the center of the German Enlightenment; the satirical novel *Andreas Hartknopf* (1786); a widely read account of *The Travels of a German in England in 1782* (1783); an *Essay toward a Practical Logic for Children* (1786); an English grammar for Germans (1784); as well as a work on German prosody (1786) and much more; and he edited the *Magazine for Empirical Psychology* from 1783 to 1793 as well as the *Monthly of the Academy of Arts and Sciences* in 1789 and 1790. But among philosophers, he is best known for the brief "Essay on the Unification of all Fine Arts and Sciences under the Concept of *That Which Is Perfect in Itself*" (1785) and the longer essay *On the Imaginative [bildende] Imitation of the Beautiful* (1788).

The first of these essays offers an early defense of the idea of *art for art's sake*. Moritz argues that an object is beautiful neither because it gratifies us nor because it is useful to us but because it possesses an entirely internal purposiveness that is so perfect that contemplation of it causes us to leave all our ordinary concerns behind: In such a moment of contemplation, "we sacrifice all of our individually limited existence to a kind of higher exis-

tence" (Moritz 1989, p. 11; Moritz 1993, vol. 2, p. 545). This position leads Moritz to the extreme conclusion that when one feels bad at seeing a play performed before an empty house, one shares the disappointment not of the playwright, actors, and producers but of the work of art itself.

Moritz's longer essay on the imitation of the beautiful is less radical and more deeply entrenched in longstanding traditions in aesthetics: Here the influence of neo-Platonism, Leibnizo-Wolffian aestheticians such as Alexander Gottlieb Baumgarten and Moses Mendelssohn (in spite of his criticism of Mendelssohn in the essay on the perfection of art), and Herder all become clear. Moritz argues that in properly imitating a beautiful work of art, one does not ape its outward appearance but, rather, strives to exercise one's own active powers in a way analogous to the exercise of the artist's powers that produced the object. At the same time, however, one seeks contemplation and repose in the experience of such an object. The apparent contradiction between these claims is resolved in Moritz's view that in contemplating the beauty of an object as a self-contained whole, one both experiences an intimation of the perfection of the cosmos as a whole and is also led to strive to transcend the limits of individuality and thereby to make one's own contribution to the perfection of that whole. Both passive and active relation to a beautiful work of art is thus a mirror of both passive and active relations to the perfection of the cosmos as a whole.

Although Moritz's name was not much mentioned by leading philosophers, his influence is clear. Kant surely knew Moritz's 1785 essay (it appeared in a number of the *Berlin Monthly* in which Kant also published an article), and his own concept of the *subjective purposiveness* of the experience of beauty may well have been intended as a corrective to Moritz's conception of the internal perfection of the work of art itself. There is no direct evidence that Kant knew Moritz's 1788 essay, but Kant's own distinction between being moved by the originality of a work of genius and merely aping its outward manner could certainly have come from Moritz. Moritz's analysis of one's both passive and active relation to beauty surely influenced Schiller's analysis of one's diverse drives with regard to beauty in his *Letters on the Aesthetic Education of Mankind*. And Moritz's idea that the contemplation of beauty allows one to transcend the limits of one's own individuality also anticipates a central theme of Schopenhauer's aesthetics. Moritz thus represents an important transition between the aesthetics of the mid-eighteenth century and classical German aesthetics.

See also Aesthetics, History of; Goethe, Johann Wolfgang Von; Herder, Johann Gottfried; Kant, Immanuel; Schopenhauer, Arthur.

Bibliography

WORKS BY MORITZ

Schriften zur Ästhetik und Poetik: Kritische Ausgabe, edited by Hans Joachim Schrimpf. Tübingen: Neimeyer, 1962.

Beiträge zur Ästhetik, edited by Hans Joachim Schrimpf and Hans Adler. Mainz: Dieterich, 1989.

Werke, edited by Horst Gunter. 3 vols. 2nd ed. Frankfurt am Main: 1993.

Anton Reiser: A Psychological Novel. Translated by Ritchie Robertson. New York: Penguin Books, 1997.

WORKS ON MORITZ

Boulby, Mark. *Karl Philipp Moritz: At the Fringe of Genius*. Toronto: University of Toronto Press, 1979.

Saine, Thomas P. *Die ästhetische Theodizee: Karl Philipp Moritz und die Ästhetik des 18. Jahrhunderts*. Munich: Finck, 1971.

Schrimpf, Hans Joachim. *Karl Philipp Moritz*. Stuttgart: Metzler, 1980.

Woodmansee, Martha. *The Author, Art, and the Market: Rereading the History of Aesthetics*. New York: Columbia University Press, 1994.

Paul Guyer (2005)

MOSCA, GAETANO
(1858–1941)

Gaetano Mosca, an Italian legal and political theorist and statesman, was born in Palermo. He was one of several social theorists, including Vilfredo Pareto and Robert Michels, who gave currency to the conception of ruling elites and their circulation as being the basic characteristic of politically organized societies. Mosca outlined his conception in *Sulla teorica dei governi e sul governo parlamentare* and elaborated it in his major work, *Elementi di scienza politica,* first published in 1895 and considerably expanded in the third edition, which appeared in 1923 (translated as *The Ruling Class*).

The *Elementi* ranges over a large number of problems in the philosophy of history and in the analysis of political organization and development. Mosca speculated about the stages of political and social development, the types of political and social systems, the role of moral forces and religions in political organization and change, the function of international and civil wars, the causes and types of revolutions, race and nationality, and the causal significance of economic factors. However, the

notion of the "political class," or "ruling class," is central to the *Elementi.*

Mosca asserted that every politically organized society of any degree of complexity is characterized by the existence of an organized minority that rules and a majority that is ruled. He rejected the Marxist position that the ruling class always derives from the organization of the economy. He held that in different types of societies, different qualities and functions characterize the members of the ruling class. In certain societies, warriors occupy a central role within the ruling class; in others, economic functions are important in determining membership; and still other societies have been characterized by a hereditary ruling class. In modern societies an important section of the ruling class is always the bureaucracy, the body of salaried officials professionally entrusted with the administration of the machinery of political, economic, and social life. (Mosca was particularly interested in the emergence of modern bureaucratic states and treated bureaucratic societies as one of the chief social types.)

It appears that Mosca loosely identified the ruling class with those who occupy the controlling or governing positions within the political organization of society. At times, however, he spoke as if the ruling class were a multiplicity of political, social, and economic elites, as when he wrote, for example, that "below the highest stratum of the ruling class there is, even in autocratic systems, another that is much more numerous and comprises all the capacities for leadership in the country." Without a ruling class, Mosca claimed, all forms of social organization would be impossible. He added that the democratic tendency—the tendency to replenish ruling classes from below—"is constantly at work with greater or less intensity in all human societies." Mosca, unlike Karl Marx, did not think of classes as necessarily conflicting social forces; nor did he think of the ruling class as always imposing its will on, and maintaining its distinctive class interests against, the rest of society.

He said that every organized political society has its "political formula," a doctrine or body of belief that legitimizes the political structure and the authority of the ruling class; there are, for example, the doctrines of divine right, and of democracy. It may often be the case that the power of the ruling class requires the use of force or violence; but Mosca thought that in stable, progressive, and flourishing societies the position of the ruling class may be founded on its intellectual and moral preeminence as well as on its care for the collective interests of the nation; the political formula that legitimizes the authority of the ruling class may be accepted by all members of the society.

In fact, in arguing that all developed societies are governed by a ruling class (and that the idea of democracy in the literal sense of government by the majority is an illusion) Mosca did not wish to imply that all societies are authoritarian or autocratic. Throughout the *Elementi* he argued strongly in support of a society marked by a high measure of what he called "juridical defence"—a society in which members of the ruling class are limited in their exercise of authority and power by moral codes that protect individual rights and liberties; a society that is pluralistic, or "open," in the sense that power is widely diffused throughout the community, and hence many different interests or social forces are able to express themselves within the political framework. Mosca was critical of parliamentary government in his early work, but later, especially in the material added to the 1923 edition of the *Elementi,* he spoke strongly of its merits; he saw it as the one form of organization able "to utilise almost all human values in the political and administrative departments of government, … [in which] the door has been left open to all elements in the governed classes to make their way into the ruling classes" (*The Ruling Class,* p. 389). Thus, although Mosca thought that recognition of the inevitable existence of the ruling class in any society was sufficient to destroy the illusions of democratic ideologies, his conclusions are not easy to distinguish from the standard doctrines of liberal-democratic political philosophy.

See also Marx, Karl; Michels, Robert; Pareto, Vilfredo; Philosophy of History; Social and Political Philosophy.

Bibliography

WORKS BY MOSCA

Sulla teorica dei governi e sul governo parlamentare. Turin: Loescher, 1884; 2nd ed., Rome, 1925.

Le costituzioni moderne. Palermo: Amenta, 1887.

Elementi di scienza politica. Rome, 1895; 2nd ed., Rome, 1896; 3rd ed., Turin, 1923; 4th ed., with preface by Benedetto Croce, Bari, 1947. There is a translation by Hannah D. Kahn of the third edition, edited and revised and with introduction by Arthur Livingstone, titled *The Ruling Class.* New York: McGraw-Hill, 1939.

Lezioni di storia delle istituzioni e delle dottrine politiche. Rome: Castellani, 1933. A revised, corrected, and retitled edition appeared in 1937 as *Storia delle dottrine politiche* (Bari, Italy: Laterza).

WORKS ON MOSCA

Bobbio, Norberto, "Liberalism Old and New," *Confluence* 5 (1956): 239–251.

Bottomore, T. B. *Elites and Society.* London, 1964.
Meisel, James H. *The Myth of the Ruling Class: Gaetano Mosca and the Elite.* Ann Arbor: University of Michigan, 1958.
Piane, Mario delle. *Gaetano Mosca: Classe politica e liberalismo.* Naples, 1952.
Piras, Quintino. *Battaglie liberali: Profili e discorsi di Benedetto Croce, Gaetano Mosca, Francesco Ruffini.* Novara, Italy, 1926.

P. H. Partridge (1967)

MOTION

The nature of motion and the philosophical problems surrounding it have been perennial issues in Western philosophy. Motion is a special case of change, and much discussion relevant to motion extends naturally to change in general (see Mortensen 2002).

Notable among the problems of motion are those provided by Zeno's paradoxes. Perhaps the hardest of these is the Arrow paradox. Consider an object in motion. At any instant of that motion, since it is an instant, the object makes no advance on its journey. But if it makes no advance in any instant of its journey, how can it make advance in all of them? The sum of a collection of nothings—even an infinite collection—is nothing. It would seem that it cannot move at all.

MOTION AND THE CALCULUS

Substantial progress concerning the topic of motion was made with the development of the calculus by Isaac Newton and Gottfried Wilhelm Leibniz in the seventeenth century. The velocity of an object at time t, $v(t)$ (with respect to a frame of reference), is given by the derivative of its spatial location, $x(t)$, with respect to time. That is, $v(t_0)$ is $dx(t)/\mathrm{d}t$, evaluated at t_0. An object is in motion at an instant if its velocity at that instant is nonzero; it is at rest if its velocity is zero.

The understanding of motion thus provided is, of course, parasitic on an understanding of the calculus itself and specifically on the notion of a derivative. In the eighteenth and early nineteenth centuries this depended on the notion of an infinitesimal; and infinitesimals behaved in a notoriously inconsistent fashion. Specifically, they were assumed to be nonzero (sometimes) and zero (sometimes).

HEGEL ON MOTION

Georg Wilhelm Friedrich Hegel, writing at the start of the nineteenth century, put the contradictory properties of the infinitesimal to the service of his dialectic. The continuous and the discrete are contradictory notions. There is, therefore, something that is their synthesis. This is a variable point: the infinitesimal. It has the property of being a point, so having zero extension, and being extended, so having nonzero extension.

This understanding allows him a particular view of the account of motion provided by the calculus. To be in motion at an instant is precisely to move an infinitesimal amount. Thus,

> [when a body is moving] there are three different places: the present place, the place about to be occupied and the place which has just been vacated; the vanishing of the dimension of time is paralyzed. But at the same time there is only *one* place, a universal of these places, which remains unchanged throughout all the changes [i.e., the variable point]; it is duration existing immediately in accordance with its notion, and as such it is motion. (Hegel 1970, p. 43)

That is, "Something moves not because at one moment of time it is here and at another there, but because at one and the same moment it is here and not here, because in this 'here' it at once is and is not" (Hegel 1969, p. 440). This provides Hegel with a simple solution to the Arrow paradox. The object advances on its journey because it does advance at each instant: It moves a tiny amount at each instant.

RUSSELL ON MOTION

Within fifty years Hegel's analysis of motion was rendered obsolete by new mathematical developments. Toward the end of the nineteenth century the notion of an infinitesimal disappeared from standard mathematics. This was because, through the work of Baron Augustin-Louis Cauchy, and particularly Karl Weierstrass, a different understanding of the derivative was developed. A derivative came to be understood simply as the limit of a certain ratio as some variable approaches a value. In particular, the velocity $v(t_0)$, that is, $dx(t)/dt$ as evaluated at t_0, came to be understood as the limit of $(x(t_0+\epsilon)-x(t_0))/\epsilon$ as ϵ approaches 0.

Therefore, the new interpretation of the calculus provided a different understanding of motion. This was spelled out by Bertrand Russell in *The Principles of Mathematics* as follows:

> [I]n consequence of the denial of the infinitesimal, and in consequence of the allied purely technical view of the derivative of a function, we must entirely reject the notion of a *state* of

> motion. Motion consists *merely* in the occupation of different places at different times.… There is no transition from place to place … no such thing as velocity except in the sense of a real number which is the limit of a certain set of quotients. (1938, p. 473)

The paradox of the Arrow can then be dismissed:

> In the case of motion, [Zeno's Arrow paradox] denies that there is such a thing as the *state* of motion. In the general case of a continuous variable, it may be taken as denying actual infinitesimals. For infinitesimals are an attempt to extend to the *values* of a variable the variability which belongs to it alone.… [The modern account of the variable has clarified this confusion, but] its absence in Zeno's day led him to suppose that continuous change was impossible without a state of change, which involves infinitesimals and the contradiction of a body's being where it is not. (Russell 1938, pp. 350–351)

PROBLEMS WITH THE ORTHODOX ACCOUNT

The view concerning motion expressed by Russell became the orthodox view of motion in the twentieth century. It is not without its problems, however. As Russell makes clear, according to this account there is no such thing as an intrinsic state of motion. That is, the instantaneous states of two objects, one in motion and one at rest at that instant, but at the same place, would be identical. Whether the object is in motion or at rest at that instant depends entirely on its states at neighboring instants. This is highly counterintuitive: Motion turns out to be a sequence (albeit a continuous one) of states that are indistinguishable from rest-states. There is no genuine flux. Motion occurs in much the same way as it appears to when successive stills in a cinema film are shown so fast that something seems to move. Indeed, one might call this the cinematic view of change. One way to bring home its oddity is as follows. Suppose that there is a particle that behaves as follows: At any time it exists simply at some place, but at any time it may disappear and reappear at some other place. Suppose that, by an accidental string of occurrences, the positions of the particle over a short period just happen to be a continuous function of time with a nonzero derivative. One would not, on this account, be inclined to say that the particle is in motion at each instant.

The cinematic account of change is not just counterintuitive. It has a number of other untoward consequences, as Russell himself notes (1938, p. 482). It is natural to take laws of nature to state causal relations between various quantities, such as velocity and its derivative, acceleration. Indeed, one normally takes it that the states of these quantities at a time are causal determinants of later states. If, in nature, there are no such things as these quantities, all this must be foregone—including the possibility of Laplacean determinism: the view that the intrinsic state of a system at any time determines its future states.

Further problems arise when one considers discontinuities of various kinds. Thus, suppose that an object is at rest before time t, and then starts to move with velocity *1*. That is, $x(t) = 0$ if $t<0$ and $x(t) = t$ if $t \geq 0$. The object has no velocity at $t = 0$ (since $x(t)$ has no derivative there), and a fortiori no acceleration. Still, it would seem that it ought to, if the motion is the result of an impulse applied to the object at $t = 0$. Worse: suppose that the object moves instantaneously at $t = 0$ to some other position where it is at rest; so $x(t) = 0$ if $t<0$ and $x(t) = 1$ if $t \geq 0$. If $t \neq 0$, the velocity of the particle is 0; and if $t = 0$, the velocity is undefined. Hence, the particle has changed places at $t = 0$, yet it has never been in motion!

Finally, and Russell's protestations to the contrary notwithstanding, it would appear that he has not so much solved the Arrow paradox as ignored it. He accepts that no progress is made on the journey in an instant, but simply insists that, nonetheless, progress is made in the whole journey. This is not a solution, it is what must be explained.

TOOLEY'S ACCOUNT

These and other objections were leveled against the Russellean account by Graham Priest (1985, 1987) and Michael Tooley (1988), each of whom offers an account of motion according to which velocity (relative to a frame of reference) is an instantaneous property of an object.

According to Tooley velocity is a theoretical (i.e., unobservable) property of an object that is causally efficacious in determining its behavior. Specifically, it is a quantity, $v(t)$, satisfying the equations:

$$x(t_1) = x(t_0) + {}_0\int^1 v(t)dt$$

$$m(t_1).v(t_1) = m(t_0).v(t_0) + {}_0\int^1 F(t)dt$$

where $m(t)$ is the inertial mass of the object at t and $F(t)$ is the force acting on it at that time. These, note, are the two key laws in (relativistic) kinematics involving velocity. The first relates velocity to position; the second to the forces acting. The crucial point is that, on Tooley's view,

these equations should be interpreted as stating relations between (instantaneous) physical quantities.

PRIEST'S ACCOUNT

Priest's account draws on Hegel. It does not resurrect Hegel's account of the categories; nor does it rehabilitate the notion of the infinitesimal. What it does do is take seriously the possibility that, at an instant, the position of a moving object may be spread out over a short (but non-infinitesimal) region. Because the object is in motion it may be impossible to localize it to any one position. This is called the spread hypothesis.

More specifically, let $x(t)$ be the locus of motion of an object, as it occurs in the laws of motion cited in the previous section. One can write r_t for the value of this function at t. For Russell, the state of the object at time t is characterized by the set of statements S_t = {'The object is at r_t'}∪ {'The object is not at r'; where $r \neq r_t$}. Given the spread hypothesis, one must suppose that there is an interval of times containing t, θ_t, such that the object is equally at $x(t')$ for all $t' \epsilon \theta_t$. The state of the object at t is therefore characterized by the set of all those statements in S_t' for $t' \epsilon \theta_t$. (What, exactly, θ_t is, is a matter to be determined by other consideration; possibly by nature itself. But it is not unnatural to suppose that the width of θ_t is proportional to $dx(t)/dt$ if this is defined.)

If $x(t')$ is constant for $t' \epsilon \theta_t$ (and, in particular, if θ_t contains just t), the state-description is identical to the Russellean state-description; in particular, it is consistent. But if $x(t')$ takes different values, r_1 and r_2, for $t' \epsilon \theta_t$, then it will be inconsistent: it will contain the statements that the object both is and is not at r_1 (and r_2).

To be in motion at an instant, then, according to this account, is to have an inconsistent state description at that instant. Objects in motion are at one place at one time, and another at another. But this is not sufficient. This would be equally true of an object at rest at each of these places. To be in motion at a time, an object must both be and not be at a place at that time.

THE ARROW AGAIN

If one is to have a theory according to which motion is an intrinsic property of an object, then the accounts of Tooley and Priest may not be the only ones; but they are the only two presently on offer. Therefore, it is natural to compare their relative merits.

One feature of Tooley's account, unlike Priest's, is that it is consistent. Priest's account (and Hegel's) presupposes that one can make sense of the possibility that the truth about a situation can be contradictory (dialetheism). It requires the use of a logic that is such that contradictions do not imply everything. One may take this to be a strong mark in Tooley's favor. Other objections against Priest can be found by consulting Tooley (1988). It appears that there are perfectly natural replies to these objections, but this is not the place to go into the matter.

On the other side, it is clear that Priest's account solves the Arrow paradox essentially as does Hegel's. The object, by occupying more than one point at an instant, does make progress during each instant, and so in the whole comprising them. Tooley's account would not appear to solve the paradox. It still leaves one with the fact that the object makes no progress during an instant of its journey. Russell, whether rightly or wrongly, took the problem to be solved by rejecting instantaneous states of motion. Even this step is not open to Tooley.

Doubtless, there is more to be said on these matters. Regardless, one thing is clear: Even after the development of the calculus, the theory of the limit, the understanding that it is possible to postulate unobservables in science, and even of paraconsistency, Zeno's paradox of the Arrow still haunts us.

See also Hegel, Georg Wilhelm Friedrich; Motion, A Historical Survey; Russell, Bertrand Arthur William; Zeno of Elea.

Bibliography

Boyer, Charles B. *The History of the Calculus and Its Conceptual Development*. New York: Dover, 1959.

Cajori, Florian. *A History of Mathematics*. 5th ed. New York: Chelsea, 1991.

Hegel, Georg W. F. *Hegel's Philosophy of Nature: Being Part Two of the Encyclopaedia of the Philosophical Sciences*. Translated by A. V. Miller. Oxford, U.K.: Clarendon Press, 1970. Originally published as *Encyclopädie der Philosophischen Wissenschaften im Grundrisse* (Heidelberg, Germany: Druk und Verlag von August Owald, 1827).

Hegel, Georg W. F. *Hegel's Science of Logic*. Translated by A. V. Miller. London: Allen and Unwin, 1969. Originally published as *Wissenschaft der Logic* (Nuremberg: Johann Leonhard Schrag, 1812–1816).

Mortensen, Chris. "Change." In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2002. Available at http://plato.stanford.edu/entries/change/.

Priest, Graham. *In Contradiction: A Study of the Transconsistent*. Dordrecht, Netherlands: Nijhoff, 1987.

Priest, Graham. "Inconsistencies in Motion." *American Philosophical Quarterly* 22 (1985): 339–346.

Russell, Bertrand. *The Principles of Mathematics*. New York: Norton, 1938.

Salmon, Wesley C., ed. *Zeno's Paradoxes*. Indianapolis, IN: Bobbs-Merrill, 1970.

Tooley, Michael. "In Defence of the Existence of States of Motion." *Philosophical Topics* 16 (1988): 225–250.

Graham Priest (2005)

MOTION, A HISTORICAL SURVEY

"Motion," or "movement," in its modern meaning, is change—or more precisely, change of the relative positions of bodies. The concept of motion thus involves the ideas of space and time. Kinematics, in the nineteenth century usually called "kinetics" or "phoronomics," is the science that deals exclusively with the geometrical and chronometrical aspects of motion, in contrast to dynamics, which considers force and mass in relation to motion. In medieval terminology, following Aristotelian tradition, "motion" (*motus* or *kinesis*) had a much wider significance, denoting any continuous change in quality, quantity, or place.

EARLY CONCEPTS OF MOTION

Ever since the beginning of philosophical speculation and scientific analysis, the concept of motion has played a predominant role in Western thought. Anaximander of Miletus (sixth century BCE) saw in motion an eternal agent of the cosmos. For Heraclitus motion was a cosmological principle underlying all physical reality (*panta rhei*, "everything is in perpetual flow"). Yet in spite of their insistence on the universality of motion, neither Anaximander nor Heraclitus seems to have inquired into the nature of motion itself. The Eleatics were probably the first to do so, when they discovered the contradiction inherent in the idea of motion and consequently denied the reality of motion, relegating its appearance to the realm of illusions and deceptions. A body, they argued, can move neither where it is nor where it is not; hence, reality is motionless and unchanging. Zeno's famous antinomies (Aristotle, *Physics* 239), such as the "Arrow" and "Achilles," seem to have been aimed, at least in part, at a refutation of the possibility of motion. On the other hand, for the atomists, such as Democritus and Leucippus, motion was a fundamental property of the atoms. All changes in nature were reduced to the movements of atoms in the void, and with the eternity and uncreatedness of the atoms their motion was eternal and uncreated; this motion itself, in the atomists' view, was not further analyzable. It remained a primary concept until Epicurus searched for a causal explanation. This (according to Lucretius) he thought to have found in weight, the cause of the downward movements of atoms, and in their little "swerves," by which he explained the otherwise incomprehensible collisions and redistributions of atoms without which physical processes could not be accounted for.

ARISTOTLE. In Aristotle's natural philosophy the concept of motion played a decisive role, since for him nature was the principle of movement or change: "We must understand what motion is; for, if we do not know this, neither do we understand what nature is" (*Physics* 200b12), a statement recurrent in Peripatetic philosophy under the motto *Ignato motu, ignatur natura* ("To be ignorant of motion is to be ignorant of nature"). For Aristotle, in contrast to his predecessors, motion raised a profound problem—not merely from the logical point of view. Expressing the deeply rooted metaphysical conviction of Western thought that motion is neither logically nor ontologically self-sufficient but requires an explanation, Aristotle contended that motion is neither in the causal, or genetic, nor in the ontological sense a primary concept. Causally, every motion originates in another motion; only animate organisms possess an inherent power to move. Hence his famous dictum *Omne quod movetur ab aliquo movetur* ("All things that move are moved by something else"). To avoid infinite regression and to find a satisfactory explanation of the existence of motion, Aristotle reduced the ultimate origin of all movements to an eternal mover who is himself unmoved. (*Physics* 258b). Ontologically, Aristotle derived motion from the basic notions of his metaphysics of substance and form by defining it as "the progress of the realizing of a potentiality *qua* potentiality" (*Physics* 201a10). Motion as the actualization of that which exists in potentiality may produce a substantial form (*generatio*), may change qualities (*alteratio*) and quantities (*augmentatio* or *diminutio*), or, finally, may be a change of place (*motus localis*). Although Aristotle did not reduce qualitative differences to quantitative relations of size and position, as did the atomists, his physics is essentially a physics of qualities. He did regard local motion as of a more fundamental character than the other kinds of motion (*Physics* 208a31); it is "the primary and most general case of passage and prior to all other categories of change" (*Physics* 260b22). Yet in spite of this preferential status, local motion for Aristotle is only a necessary concomitant of change, not, as the mechanistic physicists of the post-Newtonian era maintained, the essential and exclusive constituent of change.

In kinematics Aristotle distinguished between circular and rectilinear motion (*De Caelo* 268b17), the former, the more perfect, being the motion of the celestial bodies (*De Generatione et Corruptione* 338a18). Dynamically, motion is either natural or violate. Natural motion is circular for celestial and rectilinear for terrestrial objects; violate motion is the removal of a body from its natural place (*locus naturalis*) through the action of an external force.

ANCIENT AND MEDIEVAL CONCEPTS. Aristotle's kinematics, like his physics in general, was a qualitative science, incapable of providing a precise definition of such notions as velocity and acceleration. In fact, Greek mathematics, with its insistence on the illegitimacy of proportions or ratios between heterogeneous quantities, did not provide even the formal means of defining velocity as the ratio between distance and time; only topological, not metrical, determinations of motion could be formulated. Thus, Aristotle said that a body is quicker than another if it traverses equal spaces in less time or greater spaces in equal time (*Physics* 215a26). As related by Simplicius, Strato of Lampsacus, in a lost treatise "On Motion" (*De Motu*), was apparently the first to analyze in great detail these kinematic notions, in particular the concept of acceleration, although without trespassing the boundaries imposed by the Aristotelian conceptual scheme. The kinematics of uniform motion could be fully developed and rigorously formulated at least *in abstracto,* as exemplified by the treatise "The Motion of the Sphere" (300 BCE), written by the astronomer Autolycus of Pitane. Nevertheless, as far as is known, the earliest kinematicist to associate concrete numerical designations with velocities was Gerard of Brussels, in the thirteenth century (*Liber de Motu*).

The formulations of the basic concepts in the science of motion did not, however, evolve out of practical necessities, the study of simple machines, or other scientific or technical considerations; they were, rather, the outcome of a curious development that originated in connection with a purely philosophical, ontological, and even theological problem. The point of departure was the much discussed problem of the increase and decrease of qualities (*intensio et remissio formarum*), the question of how such qualities as warmness or blackness could vary in their intensities. Aristotle explicitly admitted (*Categories* 10b26) such alterations, but he also described such qualities as numbers (*Metaphysics* 1044a9) as immutable and unchangeable. One of the solutions, as listed by Simplicius, is that of Archytas, who suggested that every quality possesses a certain range of indeterminacy, or margin of variability (*platos*).

In Peter Lombard's "Books on the Sentences" (*Libri Quatuor Sententiarum,* c. 1150 CE) the same problem reappears in the realm of theology when it is asked, with reference to Scripture, how an intensification or diminution of the Holy Spirit or of the *caritas* is possible in man. Until well into the thirteenth century the Christian concept of *caritas* was par excellence the subject of discussions on the intension and remission of qualities and served as the standard example for intricate analyses of the notions of change and motion. One solution, advanced by Henry of Ghent in one of his *Quodlibeta,* referred in this connection explicitly to Archytas's previously mentioned conception of margin of variability, now termed the "latitude" (*latitudo*) of quality or change, a notion that was destined to play an important role in the foundation of classical kinematics.

GROWTH OF THE SCIENCE OF KINEMATICS

In order to understand the subsequent development of the concept of motion another problem that engaged the thirteenth century to a great extent must be mentioned, the question of what category change, or motion, belongs to. Aristotle was usually interpreted as having advocated an identification of *motus* with *terminus motus*—that is, viewing motion as an evolving process in the same category as the terminal, or the perfection, of this process. According to this view motion is a *forma fluens,* to use the terminology of Albert the Great, whereas the opposing view, which relates motion and its terminus to different categories, is the *fluxus formae* conception of motion. In the special case of local motion the *forma fluens* interpretation regards the process of motion as merely the continuous and gradual acquisition of the final *terminus motus,* just as the qualitative change of *nigrescere* (to become black) is merely the gradual acquisition of the *nigredo* (blackness). The concept of motion obtained its final and most radical formulation along these lines in the nominalistic statement of William of Ockham that motion is merely a name for the set of successive positions occupied by the mobile.

The nominalistic interpretation, often epitomized as *motus est mobile quod movetur,* met with considerable opposition, curiously enough among the Parisian terministic philosophers, such as Jean Buridan. One of the arguments for its rejection was undoubtedly its logical inapplicability to the motion of the outermost sphere, which, not further surrounded by any object, possessed

neither place nor space, according to the Aristotelian-scholastic theory of space; thus its motion clearly could not be interpreted as a set of successive positions. No wonder, then, that the *fluxus formae* interpretation of motion, which distinguished between the process, on the one hand, and the terminus or position (*locus*), on the other, and regarded motion as a specific quality inherent in the mobile, became predominant. Buridan, for example, defined motion, or *moveri,* as an inherent property in the mobile—*intrinsice aliter et aliter se habere*—and Blasius of Parma characterized local motion as a quality that is capable of gradual intensification or remission and is inherent in the moving object (*motus localis est qualitas gradualis intensibilis et remissibilis, mobili inhaerens subjective*).

Meanwhile the notorious *calculatores* of Merton College at Oxford, including Thomas Bradwardine, Richard Swineshead, and William Heytesbury, established their famous formalism of subjecting qualities of all kinds, but primarily the quality of *caritas,* to mathematical analysis and quantification. It was there, at Merton College, that the different trends converged. For motion, itself a quality according to the *fluxus formae* conception, soon became the favorite subject of mathematical description and took the place of *caritas* in these discussions. Employing the notion of latitude, the calculators analyzed the various possibilities of changes of motion and illustrated their theorems by graphical representations. Thus, through the conflux of various conceptual trends the foundations of modern kinematics were laid at Oxford: The concept of velocity was clarified by the introduction of the notion of instantaneous velocity, uniformly accelerated motion was unambiguously defined, the distance traversed by a body in uniformly accelerated motion was calculated, and, finally, a clear distinction between kinematics and dynamics was drawn. The results thus obtained seem, however, never to have been applied to any motions encountered in nature; they were, rather, a theory for the classification of possible motions.

The new knowledge soon spread to France, Germany, and Italy. Only Galileo Galilei, and possibly Dominic de Soto, applied these results to the study of specific natural phenomena, such as free fall. Since kinematic investigations formed the point of departure for the subsequent development of mechanics and physics in general, the analysis and clarification of the concept of motion may rightfully be regarded as of primary importance for the rise of modern science as a whole. With the establishment of a scientific kinematics the notion of motion also became purified from certain connotations that it carried from ancient times. Thus, according to the Aristotelian theory of motion the movement of any object presupposes the existence of an immobile body. Themistius, Averroes, and other commentators interpreted this statement as a proof of the immobility of Earth. In fact, for Averroes the immobility of the center was a necessary prerequisite not only for the motion of the spheres but also for the very spatiality of the outermost sphere (*caelum est in loco per centrum*). Not only was Earth unique as being the abode of man; its distinction was due also to the fact that it served as the basis for the localizability of the celestial spheres.

However, as soon as the *fluxus formae* conception characterized motion as a property inherent solely in the mobile, the Aristotelian presupposition of an immobile correlate lost its logical legitimacy. Celestial motions no longer needed to be conceived of as dependent on the immobility of Earth, and a severe obstacle to the Copernican doctrine could easily be removed.

RELATIVITY OF MOTION

It is a curious fact that the modern conception of motion, though historically and conceptually connected most intimately with the Copernican revolution, led to a partial reinstatement of the Aristotelian presupposition. Not the immobility but the existence of a correlate is the indispensable requirement for any physical significance of the concept of motion. For the relativization of the notion demands a body of reference. The question whether absolute motion, motion without reference to a physical object extraneous to the mobile, is a scientifically or philosophically meaningful conception or whether motion is only relative—that is, whether the statement "*A* moves" makes sense only if it means "*A* moves relative to *B*"—is the problem of the relativity of motion and has a long history of its own.

Aristotle's distinction between ordinary motion and motion *per accidens* may be regarded as the first implicit differentiation between absolute and relative rest, an idea further developed by Sextus Empiricus (*Adversus Mathematicos* 2, 55). The dynamical equivalence, under certain conditions, between relative rest and absolute rest was essential to the acceptance of the Copernican theory and, in fact, was explicitly stated by Nicolas Copernicus himself: *Inter motu ad eadem, non percipitur motus* (*De Revolutionibus Orbium Coelestium,* Nuremberg, 1583, Bk. 1, Ch. 3). It was further elaborated by Galileo (*Dialogo sopra i due massimi sistemi del monde,* second day) into what is now called the Galilean principle of relativity. René Descartes, fully aware of the implications of the relativity

of motion for the Copernican controversy, adopted a compromise position by distinguishing between "the common and vulgar conception of motion" as the passing of a body from one place to another and the "true or scientific conception" of motion as the transfer of matter from the vicinity of those bodies with which it was in immediate contact into the vicinity of other bodies (*Principia Philosophiae,* Part 2, Section 24). He thereby associated the relativity of true, or scientific, motion with the Aristotelian contiguity as the determinant of localization. Descartes is often credited with having been the first to enunciate explicitly the relativity of motion, and Gottfried Wilhelm Leibniz is cited as one of its most enthusiastic proponents.

For Isaac Newton and his doctrine of absolute space the notion of absolute motion was, of course, of physical significance, being "the translation of a body from one absolute place into another" (*Principles*). He defined relative motion, corresponding to the concept of relative space, as "the translation from one relative place into another." In spite of his professed adherence to Galileo's principle of relativity, Newton maintained the possibility of distinguishing absolute from relative motion by their "properties, causes and effects." His belief in the reality of absolute motion was based on his thesis that real forces create real motion. The reality of absolute motion, he argued, is manifested by the effects that such motions produce, for example, the appearance of centrifugal forces or effects. For Newton forces are metaphysical entities, and the motions they produce are therefore more than merely geometricotemporal or kinematic phenomena. Thus, rotation is an absolute motion, as he thought to have proved by an analysis of his famous pail experiment.

Apart from Christian Huygens, who from 1688 maintained the relativity of circular motion on physical grounds, and Leibniz, who rejected the Newtonian conception on philosophical grounds, it was primarily George Berkeley who treated the epistemological aspects of the problem (*Treatise concerning the Principles of Human Knowledge; De Motu*). He concluded:

> It does not appear to me, that there can be any motion other than *relative*: so that to conceive motion, there must be at least conceived two bodies, whereof the distance or position in regard to each other is varied. Hence if there was one only body in being, it could not possibly be moved. This seems evident, in that the idea I have of motion doth necessarily include relation.

However, in the eighteenth and early nineteenth centuries, primarily as a result of Leonhard Euler's justification of absolute motion on the basis of the principle of inertia (*Mechanica; Theoria Motus,* Secs. 84, 99) and Immanuel Kant's argumentations in his "Metaphysical Foundations of Natural Science" (*Metaphysische Anfangsgründe der Naturwissenschaft,* 1786), absolute motion was regarded by the majority of philosophers as a meaningful concept, not only in physics but also in philosophy. Toward the middle of the nineteenth century the situation changed. At first it was admitted that rotational motion is absolute but translational motion is relative (James Clerk Maxwell, P. G. Tait, H. Streinitz, L. Lange), and later all motion was regarded as relative. One of the most ardent proponents of the universal relativity of motion was Ernst Mach (*Die Mechanik in ihrer Entwicklung,* Leipzig, 1883); he refuted Newton's argument concerning the rise of centrifugal forces as evidence of the absolute nature of motion and explained it as an induction effect produced by the motion relative to the fixed stars. Whether Mach's conjecture can be corroborated rigorously is still a problem that engages modern research, especially in the theory of general relativity.

The question of the relativity of motion, initiated, as we have seen, by Descartes, gained increased importance, owing to the fact that the concept of motion became the basic element of physical explanation. In fact, it was Descartes's insistence on the exclusive admissibility of local motion that was decisive in this development. As is suggested in the *Principles of Philosophy* (Pt. 2, Sec. 23) and expounded in a letter to Marin Mersenne (1643), Descartes refused to attribute any reality to the so-called qualities of substances. The conception of such qualities, he contended, complicates and confuses rather than simplifies the explanation of physical phenomena in natural philosophy. In concluding such deliberations, Descartes declared local motion to be the only admissible element for physical explication. Descartes's rejection of the Aristotelian physics of qualities had a great appeal to philosophers (see, for example, Thomas Hobbes, *Elementorum Philosophiae Sectio Prima,* 1655; *De Corpore,* Sec. 8, Ch. 9) and was instrumental in the development of the mechanistic orientation of modern classical physics, which tried to reduce all natural phenomena to motions of masses in space.

Characteristic of this conception of classical physics is a statement by Maxwell: "When a physical phenomenon can be completely described as a change in the configuration and motion of a material system, the dynamical explanation of that phenomenon is said to be

complete" (*Scientific Papers,* Cambridge, U.K., 1890, Vol. 2, p. 418). The predominant role of the concept of motion in physical science poses a problem of great importance to philosophy. Why is it that all processes, laws, and formulas of physics—and modern physics is no exception—ultimately refer to motion, and why is it that even problems in statics, the science of equilibrium and absence of motion, are solved in terms of fictitious motions and virtual velocities? Is the answer to be found only in the historical circumstances, namely that kinematic investigations were the earliest successful approach to the establishment of a physical theory and that consequently forces were regarded as manifesting themselves only through motions? The answer probably lies in a vestige of ancient Eleatic philosophy that seems still to motivate our mode of thinking: A physical explanation of a natural phenomenon becomes more satisfactory the nearer it approaches the statement that nothing has happened. Motion, as Wilhelm Wundt pointed out, is the only conceivable process in which an object, so to speak, both changes and remains the same: It changes by assuming a different position relative to other objects; it remains the same by preserving its complete identity.

See also Albert the Great; Anaximander; Aristotle; Averroes; Berkeley, George; Bradwardine, Thomas; Buridan, John; Change; Copernicus, Nicolas; Descartes, René; Epicurus; Galileo Galilei; Henry of Ghent; Heraclitus of Ephesus; Heytesbury, William; Hobbes, Thomas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Leucippus and Democritus; Lucretius; Mach, Ernst; Maxwell, James Clerk; Mersenne, Marin; Motion; Newton, Isaac; Peripatetics; Peter Lombard; Philosophy of Physics; Relativity Theory; Sextus Empiricus; Soto, Dominic de; Space; Swineshead, Richard; Themistius; Time; Wundt, Wilhelm; Zeno of Elea.

Bibliography

PRIMARY SOURCES

Aristotle. *Works.* Edited by J. A. Smith and W. D. Ross. Oxford, 1908–1952. See especially the *Physics.*

Berkeley, George. *The Works of George Berkeley, Bishop of Cloyne,* Edited by A. A. Luce and T. E. Jessop, 9 vols. London, 1948–1957. See especially the *Treatise … of Human Knowledge* (1710), *De Motu* (1721), and *Siris* (1744).

Copernicus, Nicolas. *De Revolutionibus.* In *Opera Complete.* Warsaw: Officina Publica Libris Scientificis Edendis, 1873. Translated by C. G. Wallis in *Great Books of the Western World,* Vol. XVI. Chicago, 1952.

Descartes, René. *Philosophical Works.* Translated by E. S. Haldane and G. T. R. Ross, 2 vols. Cambridge, U.K.: Cambridge University Press, 1911–1912; 2nd ed., 1934. See especially the *Principia Philosophiae.*

Euler, Leonhard. *Opera Omnia,* 3 series. Leipzig: Teubner, 1911–1942; Lausanne, 1942–. See especially *Opera Mechanica et Astronomica.*

Galileo Galilei. *Dialogue concerning the Two Chief World Systems.* Translated by S. Drake. Berkeley: University of California Press, 1953.

Kant, Immanuel. *Sämtliche Werke.* Edited by the Preussiche Akademie der Wissenschaften, 22 vols. Berlin, 1902–1955.

Mach, Ernst. *Science of Mechanics.* Translated by T. C. McCormack. Chicago: Open Court, 1902.

Newton, Isaac. *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World.* Edited by Florian Cajori. Berkeley, CA, 1947.

Sextus Empiricus. *Philosophical Works,* 4 vols. Loeb Classical Library. Cambridge, MA, 1939–1957. See nos. 273, 291, 311, 382.

SECONDARY SOURCES

Lange, L. *Die geschichtliche Entwicklung des Bewegungsbegriffs.* Leipzig, 1886.

Wundt, Wilhelm. *Die physikalischen Axiome und ihre Beziehung zum Causalprincip.* Erlangen: F. Enke, 1886.

M. Jammer (1967)

MO TZU

See *Mozi*

MOUNIER, EMMANUEL

(1905–1950)

Emmanuel Mounier, the French personalist philosopher, was born in Grenoble. He studied philosophy from 1924 to 1927 in Grenoble and in Paris, where he was successful in the *agrégation* examination of 1928. After teaching philosophy in schools during 1931 and 1932, he collaborated with others in bringing out a work on the thought of Charles Péguy, whom Mounier as a Roman Catholic greatly admired. This collaboration was extended to plans for a review to carry on Péguy's work, and *Esprit* was launched in October 1932. Mounier continued to edit the review in the face of difficulties, not least of which was the feeling of some Catholics that his position was virtually Marxist. He taught at the French *lycée* in Brussels from 1933 to 1939. He was called up for military service on the outbreak of war and was demobilized shortly after the fall of France in 1940. Mounier contrived to continue the production of *Esprit* until August 1941, when the Vichy government banned it.

Suspected of subversive connections, he spent some months in prison in 1942, but was eventually acquitted and settled with his family, incognito, near Montélimar.

Mounier returned to Paris in 1945, and until his death he continued to produce books and a resuscitated *Esprit*, inspired by the times and his personalist response to them.

Mounier is the chief representative of the movement known as personalism. It is closely related, in the ideas it propounds, to existentialism. Personalism, however, is distinctively Christian and sees the personal "vocation" as seeking communication between unique persons, whereas existentialism is often divorced from religious belief, rejects the possibility of shared values, and is often strongly pessimistic concerning human relationships.

Mounier held that the person is entirely distinct from the political individual, who is "an abstract, legal, self-seeking entity, asserting his rights and presenting a mere caricature of the person." The person is "a spiritual being … subsisting by his adherence to a hierarchy of values freely adopted, assimilated, and lived through, thanks to a responsible commitment and a constant process of conversion."

The "unique vocation" of the person has little more specifiable content than Jean-Paul Sartre's "original project." Mounier, however, insisted on the distinctive character of legitimate commitment, which is both personalist and *communautaire*, or directed toward a fellowship of other persons. Man's chief task, Mounier wrote in *Qu'est-ce que le personnalisme?*, is not to master nature but increasingly to bring about communication leading to universal understanding.

Personalism is a natural product of the kind of French philosophy that has, since Maine de Biran, stressed the notion of a self that in some measure owes its being to an external reality which it apprehends or upon which it acts. Such thinking led Mounier to say that "as the philosopher who first shuts himself up within thought will never find a door leading to being, so he who first shuts himself up in the self will never find a path to others." Mounier criticized René Descartes, despite his modernity, for first adumbrating the solipsism that has since hung over modern man. In the economic field, bourgeois values "exalt the isolated individual and strengthen that economic and spiritual individualism" that still bedevils us. Mounier pointed the way from spiritually sterile self-absorption to the apprehension of reality in the form of not-self, particularly in the form of the other person with whom we communicate. The primitive experience of the person is the experience of the second person. The thou, including the we, precedes the I, or at least accompanies it. Mounier's objection to egoism was not only to economic individualism but also to its subtler forms, such as a fastidious withdrawal from modern vulgarity into the purity of the self. All true living is a transaction with the reality of the world and others in a process of mutual enrichment. There is no true inwardness that is not nourished by its interaction with an outer reality. "We must find our way out of our inwardness in order to sustain that inwardness."

See also Descartes, René; Egoism and Altruism; Existentialism; Maine de Biran; Marxist Philosophy; Personalism; Sartre, Jean-Paul.

Bibliography

WORKS BY MOUNIER

La pensée de Charles Péguy. Paris: Plon, 1931.

Révolution personnaliste et communautaire. Paris, 1935.

De la propriété capitaliste à la propriété humaine. Paris: Desclée de Brouwer, 1936.

Manifeste au service du personnalisme. Paris, 1936.

L'affrontement chrétien. Neuchâtel, 1944.

Introduction aux existentialismes. Paris, 1946.

Liberté sous condition. Paris, 1946.

Traité du caractère. Paris: Éditions du seuil, 1946.

Qu'est-ce que le personnalisme? Paris, 1947.

L'eveil de l'Afrique noire. Paris, 1948.

Le personnalisme. Paris, 1949.

Carnets de route. 3 vols. Paris, 1950–1953.

Oeuvres. 4 vols. Paris: Editions du Seuil, 1961–63.

TRANSLATIONS

A Personalist Manifesto. London, New York: Longmans, Green, 1938.

Be Not Afraid: Studies in Personalist Sociology. London: Rockliff, 1951.

Personalism. Notre Dame: University of Notre Dame Press, 1952.

The Spoil of the Violent. West Nyack, NY: Cross Currents, 1955.

The Character of Man. London: Rockliff, 1956.

WORKS ON MOUNIER

Amato, Joseph Anthony. *Mounier and Maritain: A French Catholic Understanding of the Modern World*. Tuscaloosa: University of Alabama Press, 1975.

Cantin, Eileen. *Mounier: A Personalist View of History*. New York: Paulist Press, 1973.

Copleston, Frederick. *Contemporary Philosophy*, 109–115. London: Burns and Oates, 1956.

Guissard, Lucien. *Mounier*. Paris, 1962.

Hellman, J. *Emmanuel Mounier and the New Catholic Left, 1930–1950*. Toronto: University of Toronto Press, 1981.

Lurol, G. *Mounier*. Paris: Édition universitaires, 1990.

Moix, Candide. *La pensée d'Emmanuel Mounier*. Paris: Éditions du Seuil, 1960.

Colin Smith (1967)

Bibliography updated by Thomas Nenon (2005)

MOZI

(c. 470–c. 391 BCE)

Mozi, also called Mo Di, was the founder of one of the classical systems of Chinese philosophy, Mohism, as well as of a religious community. After serving for a brief period as a civil servant, Mozi spent a number of years as a traveling counselor to feudal lords and princes, and, having never been given the opportunity to put his teachings into practice or the world in order, he had eventually to be contented with conducting a school and preparing his disciples for public office. He left a work consisting of seventy-one chapters, known as *The Mozi.* It is said that Mozi was at first a follower of Confucianism but later renounced it to found a system of thought of his own. He was critical of Confucianism for its emphasis on the codes of rituals and social elegance, which were to him burdensome and wasteful.

The rigoristic temperament of Mozi made him also a man who practiced what he preached. A chief concern for Mozi, for instance, was to reduce the recurrent military conflicts among the feudal states. There are records of his taking distant journeys to prevent the outbreak of impending wars. On one of his journeys, according to the record, he had to walk ten days and ten nights and tear off pieces of cloth from his garments to wrap up his sore feet.

A distinctive characteristic of Mozi's thought was his stress on methodology. He declared: "Some standard of judgment must be established. To make a proposition without regard for standard is similar to determining the directions of sunrise and sunset on a revolving potter's wheel." He attached great importance to the threefold test and the fourfold standard. The threefold test refers to the basis, the verifiability, and the applicability of a proposition. Explained in present-day language, this test is employed to examine a proposition for its compatibility with the best of the established conceptions, its consistency with experience, and its conduciveness to desirable ends when put into operation. The benefits resulting from the application of a proposition, the last part of the threefold test, are conceived in terms of the fourfold standard, namely, enrichment of the poor, increase of the population, removal of danger, and regulation of disorder. Mozi evidently would employ these tests and standards on all propositions without exception, and contemporary scholars have sometimes called him a pragmatist, and sometimes a utilitarian. There is a section of six chapters in *The Mozi* that has come to be spoken of as the section on Mohist logic. Most of the material contained therein has little utilitarian application, but it must have been written in Mozi's tradition, if not by his hand. This logical development is an outgrowth of Mozi's insistence on "standard of judgment" but is generally regarded as constituting a neo-Mohist movement.

A common problem that confronted all the thinkers of the classical age was how to bring order out of chaos. The system of feudalistic hierarchy instituted at the beginning of the Zhou dynasty had crumbled, the Period of Warring States (403–222 BCE) was setting in, and the people were living in suffering and bewilderment. By Mozi's diagnosis, the chaotic condition was brought about by selfishness and partiality. And the cure? "Partiality should be replaced by universality." Universal love is the keystone of Mozi's teaching. Mozi was dissatisfied with Confucianism for its gradation in benevolence, and he exhorted everyone to regard the welfare of others as he regarded his own. He was convinced that the practice of universal love would bring peace to the world and happiness to man, and he took pains to demonstrate that the principle of universal love was grounded simultaneously in its practicability on earth and its divine sanction from Heaven. Universal love for Mozi was at once the way of man and the way of God.

In contrast to most Chinese philosophers, Mozi spoke of Heaven with feeling and conviction; his conception of it was similar to the Western conception of God. The will of Heaven was to be obeyed by man and was to be the standard of human thought and action. Heaven loved all men, and it was the will of Heaven that men should love one another. Soon after Mozi's death the teacher's system became embodied in an organized church with a succession of elder masters and a considerable following.

As a religious congregation Mohism did not last long, but as a system of thought and teaching Mohism ranked with Confucianism for some two centuries as one of "the eminent schools of the day." Mohism was pushed into the background if not into complete oblivion by the ascendancy of Confucianism for the next two thousand years and was rediscovered only in the mid-twentieth century.

See also Chinese Philosophy; Heaven and Hell, Doctrines of; Logic, History of; Peace, War, and Philosophy; Scientific Method.

Bibliography

For the secondary literature on Mozi, see the bibliography in *Journal of Chinese Philosophy* 28 no. 1 & 2 (2001).

Fraser, Chris. "Mohism", *The Stanford Encyclopedia of Philosophy* (Winter 2002 Edition), Edward N. Zalta (ed.), URL: http://plato.stanford.edu/archives/win2002/entries/mohism/.

Lowe, Scott. *Mo Tzu's Religious Blueprint for a Chinese Utopia.* Lewiston, NY: Edwin Mellen, 1992.

Mei, Yi-pao, tr. *The Ethical and Political Works of Motse.* London: Probsthain, 1929; Westport, CT: Hyperion Press, 1973. Complete translation.

Mei, Yi-pao. *Mo-tse, the Neglected Rival of Confucius.* London: Probsthain, 1934; Westport, CT: Hyperion Press, 1973.

Tseu, Augustine. *The Moral Philosophy of Mozi.* Taibei: China Printing Limited, 1965.

Watson, Burton, tr. *Mo Tzu: Basic Writings.* New York: Columbia University Press, 1963. Selected translation.

Y. P. Mei (1967)
Bibliography updated by Huichieh Loy (2005)

MUKAMMAṢ, DAVID BEN MERWAN AL-

See *Muqammiṣ, David ben Merwan al-*

MULLĀ ṢADRĀ
(1571/1572–1640)

Mullā Ṣadrā is the name usually given to Muḥammad ibn Ibrāhīm Ṣadr al-Dīn Shīrāzī, the most outstanding of the later Muslim philosophers. (*Mulla* means teacher.) He is also known by the honorific title *Sadr al-muta'allihin,* "the foremost among the theosophers." Born in Shiraz into an aristocratic family, he received his early education in that city and his advanced training in Ispahan, the Safavid capital, where he studied with Mīr Dāmād and Bahā' al-Dīn ʿAmilī. After completing his formal education he retired to a village near Qum, where he spent ten years in asceticism and self-purification. Then, upon the demand of the Persian king, he returned to Shiraz as a professor in the school of Allāhwirdī Khān, where he taught and wrote for the rest of his life. He died in Basra on the return journey from his seventh pilgrimage to Mecca.

Mullā Ṣadrā wrote over fifty books, most of them after leaving his spiritual retreat. All his books are in Arabic except his "spiritual defense," the *Sih aṣl* (Three principles) and a few poems and letters, which are in Persian. His works can be classified into those dealing primarily with religion, such as his commentaries on the Qur'an and the *Uṣūl al-Kāfī* (Principles of Kāfī) of Kulainī, and those which deal mostly with philosophy and theosophy. In the latter category the most important is *Al-Ḥikmat al-mutaʿāliyah fi'l-asfār al-arbaʿah* (The exalted wisdom concerning the four journeys of the spirit), or simply *Asfār* (The journeys), a work of monumental proportions and the most advanced work on Islamic philosophy. Mullā Ṣadrā also wrote a large number of shorter treatises, such as *Al-Mashā ir* (The book of metaphysical penetrations), *Al-Shawāhid al-rubūbīyah* (Divine witnesses), and *Al-Ḥikmat al-ʿarshīyah* (The book of theophany inspired by the throne), which treat specific metaphysical and philosophical questions.

In Mullā Ṣadrā's work Muslim Peripatetic philosophy, especially that of Avicenna, the illuminationist theosophy of Shihāb al-Dīn Yahyā Suhrawardī, the gnostic doctrines of Muhyī al-Dīn ibn al-ʿArabī and certain themes of Muslim theology (*Kalām*) became unified in the background of Shīʿism and the teachings of the Shīʿite imams. The philosophy of Mullā Ṣadrā, however, is synthetic rather than eclectic, because out of these various threads he created a new intellectual perspective in which reason, revelation, and mystic vision are harmonized into a total, unified view of things.

Mullā Ṣadrā brought to fruition the attempt of Muslim thinkers from the beginning of the Middle Ages to harmonize religion and philosophy. In his thought the tenets of revelation, the dicta of reason, and the verities of gnosis discovered through illumination are all considered possible sources of knowledge and are blended together. His writings bridge discursive and intuitive knowledge by making the discoveries of reason the necessary background of spiritual knowledge, which is above reason without being irrational. Mullā Ṣadrā also revised many of the tenets of Peripatetic and illuminationist philosophy and established philosophy upon a set of principles, many of which were derived from Sufism, that had not been demonstrated as such and had not existed in philosophy before.

These principles include the unity, gradation, and principality of being, by which is meant that it is being rather than the quiddity or essence of things that is ultimately real. Moreover, being is inwardly unified as a single reality that possesses states and gradations. It is upon this principle that Mullā Ṣadrā built his "metaphysics of being." Another principle of his philosophy is the unity of the intellect, or intelligence, and the intelligible, of the knower and the known. At the moment of intellection the intellect becomes identified with the intelligible form of the object perceived. Thus, knowledge is intimately connected with being and affects the ontological state of the knower.

Mullā Ṣadrā also posited the principle of substantial motion. According to the previous Muslim philosophers and going back to Aristotle, motion is possibly only in the accidents of things, not in their substance. Mullā Ṣadrā thought that, on the contrary, motion implies an inner becoming within the substance of things and therefore a continuous development toward higher states of being (without in any way implying the modern theory of evolution).

Another important principle asserted by Mullā Ṣadrā is the catharsis and independence of the imaginative faculty from the body. There is an intermediate "imaginal world" (*mundus imaginalis*) not to be confused with the "imagination" of current usage. The human imagination is a microcosmic aspect of this cosmic imagination and it is precisely in this domain possessing a reality of its own that eschatological problems whose solution escaped earlier philosophers take place and can be understood. These and many other principles, some of whose roots are to be found in the writings of the earlier Sufis and philosophers, Mullā Ṣadrā systematized and developed to their full conclusion.

Mullā Ṣadrā had many students, of whom the most famous are Mullā Muhsin Faiḍ Kāshānī and ʿAbd al-Razzāq Lāhījī, who were among the leading Shīʿite thinkers. His disciples propagated his works and teachings in both Persia and India, and in fact he founded a school that has dominated the intellectual life of Persia for the past four centuries. It is, however, against his worldview that the founder of the Shaikhī movement, Shaikh Aḥmad Ahsāʾī, wrote his criticisms. The Bāb, the founder of Babism, also belongs to the current against Mullā Ṣadrā and should by no means be considered as a product of his school. The school of Mullā Ṣadrā is still alive in Iran today and is the most important traditional school of philosophy and theosophy there.

See also Aristotle; Avicenna; Ibn al-ʿArabī, Islamic Philosophy; Logic, History of; Peripatetics; Sufism; Suhrawardī, Shihāb al-Dīn Yaḥyā.

Bibliography

WORKS BY MULLĀ ṢADRĀ

Kitāb al-Ḥikmat al-ʿarshīyah (The book of theosophy inspired by the throne). Teheran, 1278 AH/1861 CE.

Al-Ḥikmat al-mutaʿālīyah fīʾl-asfār al-arbaʿah (The exalted wisdom concerning the four journeys of the spirit). Teheran, 1282 AH/1865 CE; new ed., edited by M. H. Ṭabāṭabāʾī, Teheran, 1378 AH–/1958 CE–.

Al-Shawāhid al-rubūbīyah (The divine witnesses). Teheran, 1286 AH/1869 CE.

Rasāʾil (Treatises). Teheran, 1302 AH/1884 CE.

Taʿlīqāt ʿalā ilāhīyāt Kitāb al-Shifā (Glosses upon the metaphysics of the book of the remedy of Ibn Sīnā). Teheran, 1303 AH/1885 CE.

Sharh al-hidāyat al-athīriyah (Commentary upon the guide of Athīr al-Dīn Abharī). Teheran, 1313 AH/1895 CE.

Kitāb al-mabdaʾ waʾl-maʿād (The book of origin and return). Teheran, 1314 AH/1896 CE.

Taʿlīqāt ʿalā Kitāb Ḥikmat al-ishrāq (Glosses upon the theosophy of the orient of light [of Suhrawardi]). Teheran, 1315 AH/1897 CE.

Asrār al-āyāt (Secret of Quranic verses). Teheran, 1319 AH/1901 CE.

Sih aṣl (The three principles). Edited by S. H. Nasr. Teheran, 1380 AH/1960 CE.

Kitāb al-mashāʿir. Edited and translated by Henry Corbin as *Le livre des pénétrations métaphysiques.* Teheran and Paris, 1964. Contains Arabic text, Persian version, and French translation.

Mafātīḥ al-ghaib (Keys to the hidden world). Teheran, n.d.

Sharḥ al-uṣūl min al-Kāfī (Commentary upon the principles of Kāfī). Teheran, n.d.

WORKS ON MULLĀ ṢADRĀ

Āshtiyānī, S. J. *Sharḥ-i ḥāl wa ārāy-i falsafi-i Mullā Ṣadrā* (The biography and philosophical views of Mullā Ṣadrā). Meshed, 1382 AH/1962 CE. In Persian.

Corbin, Henry. "La place de Mollā Ṣadrā Shīrāzī dans la philosophie iranienne." *Studia Islamica* 18 (1963).

Corbin, Henry. *Terre céleste et corps de resurrection: De l'Iran mazdéen à l'Iran shīʿite.* Paris: Buchet/Chastel, 1960.

Horten, Max. *Das philosophische System con Schirazi (1640) übersetz und erläutert.* Strassburg, 1913.

Langarūdī, Muḥammad Jaʿfar. *Sharḥ al-mashāʿir* (Commentary upon the book of metaphysical penetrations). Edited by S. J. Āshtiyānī. Meshed, 1384 AH/1964 CE.

Nasr, S. H. *Islamic Studies.* Beirut: Librairie du Liban, 1967.

Nasr, S. H., ed. *Mulla Ṣadra Commemoration Volume.* Teheran, 1380 AH/1960 CE.

Nasr, S. H. "Ṣadr al-Dīn Shīrāzī." In *A History of Islamic Philosophy,* Vol. II, edited by M. M. Sharif. Wiesbaden, 1966.

Sajjādī, S. J. *Muṣṭalaḥat-i falsafi-yi Ṣadr al-Dīn Shīrāzī mashbūr bi Mullā Ṣadrā* (The philosophical vocabulary of Ṣadr al-Dīn Shīrāzī). Teheran, 1380 AH/1960 CE. In Persian.

Seyyed Hossein Nasr (1967)

MULLĀ ṢADRĀ [ADDENDUM]

In developing his concept of existence Mullā Ṣadrā works against the backdrop of Shihāb al-Dīn Yahyā Suhrawardī's essentialist metaphysics on the one hand, and Avicenna's rather incomplete and occasionally imprecise remarks on being on the other. Suhrawardī had defended essence (*māhiyya*) as the sole reality and as the

proper subject matter of metaphysics. For him, existence (*wujūd*) is a common term and a secondary intelligible, shared by a multitude of objects without corresponding to any particular being. Suhrawardī proposed two objections against the primacy of existence. First, if existence is to be the real attribute of an essence in the extramental world, then this essence will have to have an existence of its own before receiving existence as an attribute. In this case existence will be the attribute of something that already exists. Second, if existence is to be the basis of reality, then it will have to exist before being such a basis. In this case this second existence will have to exist before serving as a basis for the first existence, and so on ad infinitum. Therefore, existence is a secondary intelligible posited by the mind, adding nothing to the concrete existence of quiddities.

Ṣadrā's response to this is based on a position he calls the primacy of existence (*aṣālat al-wujūd*). Instead of defining existence as a generic term and attribute, which things take on *a posteriori*, Ṣadrā construes it as that by which things are what they are. According to Ṣadrā one cannot say "existence exists" just as one cannot logically say "whiteness is white." When one talks about beings that exist, what one means is that things exist or simply are rather than they have existence. This means that the existence of something is its reality. According to Ṣadrā Suhrawardī's essentialism results from his failure to make a distinction between the concept and reality of existence. While existence as a concept is a secondary intelligible applicable to a multitude of objects, the reality of existence is such that it leaves no distance between the existence of something and its reality. Furthermore, Ṣadrā posits existence as the principle of both unity and difference. On the one hand, existence is that which makes things exist and, on the other, it is that which makes them what they are as a specific quiddity. Existence becomes delimited and multiplied by itself alone, displaying various modes of intensification and diminution. Ṣadrā explains this process with his central concept of the gradation of existence (*tashkīk al-wujūd*). In this gradational ontology essences are nothing but mental constructions produced by the human mind to denote the different particularizations of existence, which ultimately remain one and the same.

Ṣadrā's insistence on existence as the sole reality has far-reaching consequences for his epistemology. He defines knowledge as a mode of existence and relegates all cognition to the immediate perception of existence. In this view, to know something is to know its intelligible form. But since Ṣadrā takes intelligible forms to be various manifestations and self-delimitations of existence, one's epistemic access to things ought to be through the existence of what one knows. Furthermore, Ṣadrā's realist ontology considers intelligible forms ontologically more real and epistemologically more reliable than their material existence. The climax that one reaches through the unification of the intellect, the intelligible and the intellected is thus a mode of existential intensification and not a simple process of conceptual augmentation. These considerations lead Ṣadrā to develop a mystical theory of knowledge without totally jettisoning the traditional peripatetic noetics.

See also Avicenna; Essence and Existence; Mysticism, Nature and Assessment of; Peripatetics; Suhrawardī, Shihāb al-Dīn Yaḥyā.

Bibliography

WORKS BY ṢADRĀ

The Metaphysics of Mulla Ṣadrā. Translated by Parviz Morewedge. New York: Society for the Study of Islamic Philosophy and Science, 1992. Originally published under the title *Kitab al-Masha'ir*.

Tafsir al-qur'an al-karim. Qum: Bidar Press, 1366–1369/1987–1990. 7 vols. edited by Muhammad Khwajawi. Ṣadrā's incomplete commentary on the Qur'an written from a philosophical and mystical point of view.

Ma'ani al-alfaz al-mufradah min al-qur'an (The meanings of Qur'anic terms). A short exposition of certain Qur'anic terms and expressions.

Mutashabihat al-qur'an (Allegorical verses of the Qur'an). Ṣadrā's explanation of the Qur'anic verses that are called allegorical or metaphorical (*mutashabih*) by the Qur'an (Al-i 'Imran, 3:7) such as 'God's hand.' Published in Sayyid Jalal al-Din Ashtiyani, *Sih Risalah* (Mashhad: Mashhad University Press, 1352 (A. H. Lunar)).

Sharh usul al-kafi (Commentary on the *Usul al-kafi*). A commentary on the famous Shi'ite book of Hadith compiled by Abu Ja'far Muhammad ibn Ya'qub Kulayni (d. 329/940). Edited by Muhammad Khwajawi in 3 Vols. (Tehran: Mu'assasa-yi Mutala'at wa Tahqiqat-i Farhanghi, 1366 (A. H. Solar)).

Risalat al-Hashr (The treatise of resurrection). Muhammad Khwajawi has published a critical edition with Persian translation under the title *Risalat al-hashr ya kitab-i rastakhiz-i jahan* (Tehran: Intisharat-i Mawla, 1377 (A. H. Lunar; second edition)).

Huduth al-'alam (Temporal origination of the world). Ṣadrā's most important work on the temporal origination of the world. Printed in the *Rasa'il* (Qom: Maktab al-Mustafawi, 1302 (A. H. Lunar)) (first treatise).

Iksir al-'arifin fi ma'rifat tariq al-haqq wa'l-yaqin (The elixir of the gnostics for knowing the path of the truth and certainty). One of Ṣadrā's mystical works; Translated by William C. Chittick as *The Elixir of the Gnostics: A Parallel English-Arabic Text*. Provo, UT: Brigham Young University Press, 2003.

Kasr al-asnam al-jahiliyyah fi dhamm al-mutasawwifin (The demolition of the idols of ignorance in blaming those who pretend to be Sufis). Ṣadrā's attack on those who pretended to be Sufis and held excessive views during the Safawid period; edited by M. T. Danechepazuh (Tehran: 1340 (A. H. Solar)).

Kitab al-Masha'ir. Translated by P. Morewedge as *The Metaphysics of Mulla Ṣadrā*. Texas: Global Publications Associations, 1992. A summary of some key issues of Sadrean metaphysics.

Al-Mazahir al-ilahiyyah fi asrar al-'ulum al-kamaliyyah (Divine manifestations concerning the secrets of the sciences of perfection). One of Ṣadrā's major theological works synthesizing philosophical arguments with quotations from the Qur'an; edited with an introduction by Sayyid Muhammad Khamanei. Tehran: Bunyad-i Hikmat-i Islami-yi Ṣadrā, 1378.

Zad al-musafir (Provisions of the traveler) also known as *Zad al-salik* and *Ma'ad al-jismani*. A short treatise summarizing Ṣadrā's views on eschatology. Edited by S. J. Ashtiyani. Tehran: Mu'assa-yi Intisharat-i Amir Kabir, 1379 (A. H. Lunar).

Majmu'a-yi rasa'il-i falsafi-yi Sadr al-Muta'allihin, edited by Hamid Naji Isfahani. Tehran: Intisharat-i Hikmat, 1375 (A. H. Lunar). Contains a number of Ṣadrā's short treatises.

WORKS ABOUT ṢADRĀ

Açikgenç, Alparslan. *Being and Existence in Sadrā and Heidegger: A Comparative Ontology*. Kuala Lumpur: International Institute of Islamic Thought and Civilization, 1993.

Alawi, Hadi. *Nazariyyat al-Harakat al-Jawhariyyah 'ind al-Shirazi*. Beirut: Dar al-Tali'ah, 1983.

Horten, Max. *Die Gottesbeweise bei Shirazi*. Bonn, 1912.

Jambet, Christian. *Se rendre immortel: Traité de la résurrection*. Paris: Fata Morgana, 2000.

Kalin, Ibrahim. "An Annotated Bibliography of the Works of Mulla Sadra with a Brief Account of His Life." *Islamic Studies* 42 (1) (2003): 21–62.

Kalin, Ibrahim. "Between Physics and Metaphysics: Mulla Sadra on Nature and Motion." *Islam and Science* Vol. I, 2003, pp. 65–93.

Khajawi, Muhammad. *Du Sadraddin*. Tehran: Mawla, 1999.

Khajawi, Muhammad. *Lawami' al-'Arifin fi Ahwal Sadr al-Muta'allihin*. Tehran: Mawla, 1987.

Khurram-dashti, Nahid Baqiri, and Fatimah 'Asghari *Kitab shinasi-yi jami'-yi Mulla Sadra*. Tehran: Bunyad-i Hikmat-i Islami-yi Sadra, 1378/1999.

Morris, James Winston. *The Wisdom of the Throne: An Introduction to the Philosophy of Mulla Sadra*. Princeton, NJ: Princeton University Press, 1981.

Morris, Zaylan. *Revelation, Intellectual Intuition, and Reason in the Philosophy of Mulla Sadra: An Analysis of the al-Hikmat al-'Arshiyyah*. London: Routledge Curzon, 2003.

Musawi, Musa al-. *al-Jadid fi Falsafat Sadr al-Din al-Shirazi*. Baghdad: al-Dar al-'Arabiyyah li't-Tiba'ah, 1978.

Nasr, Seyyed Hossein. *Sadr al-Dīn al-Shīrāzī and His Transcendent Theosophy*. 2nd ed. Tehran: Institute for Humanities and Cultural Studies, 1997.

Peerwani, L. "Qur'anic Hermeneutics: The Views of Sadr al-Din Shirazi" in British Society for Middle East Studies Proceedings (1991), pp. 468–477.

Rahman, Fazlur. *The Philosophy of Mullā Sadrā*. Albany: SUNY Press, 1975.

Sajjadi, Ja'far. *Mustalahat-i Falsafi-yi sadr al-Din Shirazi*. Tehran: Danishgada-yi 'Ulum-u Ma'qul wa Manqul, 1961.

Ibrahim Kalin (2005)

MULTICULTURALISM

In many academic fields in the United States after 1970, multiculturalism has meant that members of historically disadvantaged nonwhite or minority racial and ethnic groups have distinctive knowledge and ways of knowing that ought to be incorporated into curricula and recognized in research. This idea has led to area studies programs and departments that concentrate on cultures from specific geographical locations, such as Africana or African American studies, Latino/a studies, Asian American studies, Native American studies, and more generally, American studies and ethnic studies. As well, new texts and different cultural perspectives have been incorporated into traditional fields in the humanities and social sciences.

Multiculturalists have advocated greater diversity and representation in the academic community, by increasing members of historically disadvantaged groups among faculty, staff, and students, and recognizing and addressing their distinctive intellectual and socially relevant interests. Multiculturalism has often been associated with identity politics, or advocacy of the interests of minority groups, by their members, in both national and local politics and representations of ideas and persons in specific institutional contexts. Multiculturalism has been opposed in academia, because it is believed to weaken traditional subject matter by minimizing the established canon and neglecting universal knowledge. This opposition has been largely from conservative white intellectuals, but not exclusively so. For example, sociologist Yehudi Webster has argued that multiculturalism deprives students of the opportunity to develop critical thinking skills, and philosopher Jason Hill has argued that in emphasizing the value of racial and ethnic identities, multiculturalism stifles individual creativity and shared cosmopolitanism.

MULTICULTURALISM IN THE U.S. PROFESSION OF PHILOSOPHY

The practice of academic philosophy in the United States has tended to be restricted to the work of English, French, German, and ancient Greek philosophers, with varied recognition of American philosophy or pragmatism. Advocates of multicultural inclusion have argued that philosophical inquiry has not been limited to the United States, Europe, and ancient Greece, but exists in intellectual traditions in China, India, Africa, and South America, as well as in the cultures of groups worldwide. Multicultural advocates therefore conclude that the canon of American academic philosophy ought to reflect more geographical diversity. Also, when Western European philosophy has presented itself as universal, the incorporation of multicultural philosophical perspectives would seem to imply that Western European philosophy itself is as local as philosophies from other parts of the world.

Such intellectual multiculturalism has been undertaken by a number of American philosophers since the end of the twentieth century; James Sterba (2002) has argued that there is a Western bias in ethics that can be corrected. In *Native Pragmatism* (2002), Scott Pratt identifies Native American perspectives in nineteenth and early twentieth century American philosophy, and tracks their transmission. Also, introductory anthologies have become more inclusive of African, Asian, and East Indian traditions—Max Hallman's (2003) collection, *Traversing Philosophical Boundaries*, is one example.

American philosophers have also addressed demographic multiculturalism, which aims to increase the racial and ethnic diversity of philosophers and resembles the kinds of multiculturalism in other fields that has been associated with identity politics. In 2003 the American Philosophical Association's (APA) Committee on Inclusiveness proposed that the APA Board consider, for possible approval by all APA members, the following statement on inclusiveness:

> The American Philosophical Association is committed to expanding and enhancing the inclusiveness of the profession by: (A) Increasing the numbers and respected presence of persons from groups that have historically been subjected to invidious discrimination. These groups include, but are not limited to, disabled persons; persons of African descent; American Indians; Asians and Asian Americans; Hispanics and Latinos/as; Jews; persons of Middle Eastern descent; multiracial persons; lesbian, gay, bisexual, and transgendered persons; women. (B) Recognizing and supporting the development of scholarly philosophical research, teaching, service, and professional activity pertaining to the concerns of these groups.

The APA Board and the profession of U.S. academic philosophers are likely to approve the statement on inclusiveness, although as the profession develops over the twenty-first century, both intellectual and demographic multiculturalism, and external political and social changes will probably result in its augmentation and revision. Still, many traditional philosophers have opposed multiculturalism, on the grounds that its distinctive knowledge and epistemologies are not contributions to the field of academic philosophy, but rather applications of philosophical methods to new subjects, or else simply not philosophical at all. There are also concerns about time constraints on courses resulting in superficial instruction of a variety of traditions, in place of more thorough investigation of one or two.

Nonetheless, the APA, which is the primary professional organization of U.S. academic philosophers, publishes biannual newsletters on: American Indians in philosophy; Asian and Asian American philosophers and philosophies; the black experience; and Hispanic/Latino issues in philosophy. Also, the APA Committee on Inclusiveness is a standing committee that includes APA special committees on: American Indians, Asian and Asian American philosophers and philosophies, blacks in philosophy, and Hispanics. All of these committees were formed to address the relatively small numbers of nonwhite philosophers, and the absence of strong professional support of multicultural writing and teaching in the field. In 2002, the number of nonwhite academic philosophers was lower than the 10 percent of nonwhites in the U.S. professoriate overall, a figure that had not changed since 1989, and the percentage of African American philosophers was less than the national 4.4 percent of the U.S. professoriate, half of whom were employed in traditional black colleges (see Wilson [2002] for the national figures).

As of 2005, multicultural scholarly work in philosophy has mainly focused on African American concerns, although in the late 1990s, Asian American, Native American, and Hispanic concerns began to appear in philosophy courses and publications. In addition, since the late 1970s, feminist philosophers have attempted to address issues raised by nonwhite women, partly in response to criticism by bell hooks, Elizabeth Spelman, and others,

that academic feminists were overly preoccupied with the problems of white middle class women.

This growing body of multicultural philosophical work is to some extent independent of the intellectual multiculturalism already mentioned, because it tends to be motivated by concerns about demographic inclusiveness and past oppression. It should be noted that the adjective *multicultural* does not always appear in multicultural scholarly work by philosophers, who are instead likely to use the terms *race* or *racial*, *black*, *African American*, *Asian American*, *Hispanic or Latino/a*, or *Native American* in their titles and within their work.

Still, the multicultural work of philosophers has often been multidisciplinary, with forays into anthropology, literature, sociology, law, the history of ideas, economics, and social theory. At the same time, multicultural philosophy has made use of traditionally analytic, continental (phenomenological and existentialist), and postmodern philosophical methodologies, sometimes combining different methodologies in the same texts. Much of the multicultural philosophical work is about race in U.S. society, and much of it is centered on social and individual problems or questions: Can affirmative action be morally justified? What is racism and how can it be remedied? What is racial identity? Are reparations for past oppression, such as slavery and the appropriation of indigenist lands, morally imperative? Does biology support ideas of human racial divisions in society?

Writings of historical figures have also been reexamined, for instance: David Hume and Immanuel Kant for their belief in the existence of hierarchies of human races; W. E. B. DuBois and Alain Locke for their ideas about racial identity; Frederick Douglass and Julia Ann Cooper for their contributions to theories of liberation; Frantz Fanon and Jean-Paul Sartre for ideas on individual freedom and authenticity and group emancipation. Overall, the subject of race in U.S. multicultural philosophy unifies into a set of logically connected concepts and subjects that scholars analyze from diverse starting premises, with considerable disagreement, albeit a common goal of increasing social justice for disadvantaged groups. At least three of these subjects merit closer examination in this context of multiculturalism in philosophy: the existence of biological race, racism, and affirmative action.

THE EXISTENCE OF BIOLOGICAL RACE

Whether or not human races exist as biological divisions of humankind has philosophical implications: If races are biologically real, then the social problems concerning race are matters of race relations; if races are not biologically real, then many social problems as well as much of the discourse about race must be understood by philosophers in terms of false beliefs that participants hold. That is, if biological races exist, then the philosophical discussion about race is in part a direct discussion about the world, whereas if biological races are fictional, then the philosophical discussion becomes a second-order discourse about what people believe. David Hume, Immanuel Kant, and other Enlightenment thinkers thought that the existence of human races was self-evident. During the time they wrote, the new sciences of biology and anthropology had begun to produce systems of classification that appeared to explain those physical differences among human groups, which were apparent in common sense.

By the mid-nineteenth century, human races were believed to be biological groups with common inherited physical, cultural, and psychic traits. American anthropologists were prominent proponents of natural human hierarchies, based on race and ultimately caused by racial essences, believed to be inherited in the blood. However, during the early twentieth century, anthropologist Franz Boas and his students Claude Lévi-Strauss, Margaret Meade, and Melville J. Herskovits established that history and culture were the causes of nonphysical racial differences. Subsequently, biological anthropologists came to agree that there were no general physical essences or even stable sets of particular traits shared by every member of any race. Blood types do not correspond to social racial groups. Mitochondrial DNA, used to track existing populations to ancestral groups in Asia, Africa, and Europe, has no relation to genes that determine inherited traits considered racial in society. And there is greater variation within races of those inherited racial traits than between any two races. In short, while biology confirms the existence of inherited physical traits that are considered to be racial in society, biology, according to some scholars, offers no support for a taxonomy of human races.

In the early 1990s, Kwame Anthony Appiah was the first U.S. philosopher to examine the lack of a scientific biological foundation for human races and he then argued that racial identities ought to be reconsidered. His work was taken up through controversial justifications for mixed race identity and more extensive philosophy of science analyses of how ideas of race are precluded by the findings and methodologies of biological anthropology, Mendelian heredity, and population genetics as of the late twentieth-century (Zack 1993 and 2002).

Yet, by 2005, most Americans continued to believe that human races are real physical divisions and that the

social taxonomy of three (or four or five) races can be verified within the biological sciences. Multicultural scholars in all fields and philosophers who begin their inquiries on the basis of common sense or received opinion, tend to concur with the public, although often for avowedly political motivations. Thus, Lucius Outlaw, writing in the tradition of W. E. B. Dubois and Alain Locke, advocates a conservation of ordinary ideas of race, with their biological connotations, for the sake of continued self-esteem and social justice for African Americans. Amy Gutman claims that retention of ideas of race is essential for identifying those groups who have been oppressed or discriminated against on the basis of their purported race, so that their members may be assisted toward equality of opportunity for success in society.

Furthermore, scholars of Latino philosophy such as Linda Alcoff, Jorge Gracia, Eduardo Mendieta, and Ophelia Schutte have included discussions of racism in their analyses of Hispanic and Latino ethnicity. This suggests that members of dominant white Northern European groups have sometimes viewed Hispanics and Latinos as a distinct race and that addressing discrimination associated with that view could include the construction of positive distinctive racial identities for Hispanics and Latinos. And even in a purely conceptual analysis, Michael O. Hardimon (2003) dismisses disputes about the scientific standing of race and their relation to the ordinary concept of race. Hardimon then asserts, "The ordinary concept of race is our concept. It is part of our discourses, our practices, our conceptual repertoire" (Hardimon 2003, p.438).

Similarly, in an Op-Ed piece in the *New York Times* on March 14, 2005, Armand Mare Leroi, an evolutionary developmental biologist, called upon scientists to resurrect notions of biological race in light of its cultural significance, citing the importance of the preservation of the Negritos, an ancient tribe on the Andaman Islands in the Indian Ocean. It is a paradox that while philosophical multiculturalism enables analysis of the biological foundations for race, multiculturalist beliefs about how to attain social justice are held to be incompatible with the results of such conceptual analyses, even though everything of social value that used to be called race can be captured by ideas of family heredity and culture.

RACISM

The term *racism* came into broad usage in the United States during the late 1960s and there has since been both implied and explicit disagreement about what racism is. The concept of racism is broader than its predecessors, bigotry, discrimination, intolerance, and prejudice, because it can refer to social conditions as well as intentions and attitudes of individuals. By the late twentieth century, there was a consensus in business, academia, politics, and public life generally that racism in individuals is morally wrong and that the practice of racism by representatives of institutions and organizations is unjust, as well as in violation of Title VII of the Civil Rights Act of 1964, which prohibits employment discrimination based on race, sex, national origin, or religion. (Title VI prohibits public access discrimination, which was relevant to the implementation of school desegregation and Title VIII was the first federal fair housing law).

Moral philosophers have traditionally posited justice as a cardinal social and individual virtue. Given the premise that racism is a kind of injustice against human beings based on their racial identities, the main philosophical argument has focused on whether the causes of racism and remedies for it are confined to individuals or can be understood as institutional. Because racism, as a wrong, requires remedies where it exists, the individual view focuses on psychological and educational remedies, whereas the institutional view supports progressive legal action and public policy. Both views are motivated by concepts of responsibility in the sense that both individuals and societies are believed to be accountable and subject to blame for wrongs they commit.

Some proponents of individual views of racism have worked within a Kantian moral tradition. J. L. A. Garcia (1997) has argued that racism is a kind of ill will or contempt in the hearts and minds of individuals, a lack of benevolence for which they are morally responsible. Racism in this sense may be present when others are not harmed by it and it may not be present when others are harmed in ways associated with their race.

Philosophers who study racism with a multidisciplinary approach are inclined to define racism institutionally, because historians, sociologists, and political scientists have provided many extended examples of behavior, traditions, and laws that explicitly or implicitly disadvantage members of nonwhite groups in comparison to whites. Slavery, segregation, and the status of African Americans according to many measures of demographic well-being are one set of examples; the failure of the U.S. government to honor its treaty obligations to Native Americans is a second; restrictions on nonwhite immigration are a third.

For all minority groups, evidence of institutional racism against them includes disproportionately higher rates of incarceration, poverty, and unemployment, and

disproportionately lower rates of income, family assets, advanced educational degrees, and presence in the political leadership class. Whereas most scholars in philosophy and other fields focus on institutional racism as a modern and postmodern phenomenon, several have drawn wider connections. Berel Lang (1997) claims that racism is historically prior to modern ideas of race and that metaphysical racism is a set of ideas and practices that can be attached to varied specific notions of race; Charles Mills (1997) argues that modern Western history has developed on the basis of a racial contract that places Europeans and Americans at the top of a hierarchy in which indigenist Americans, Africans, and Asians are oppressed and exploited.

In the context of American history, critical race theorists such as Derek Bell and Patricia Williams have argued that the American legal system is structurally racist, from the acceptance of slavery in the U.S. Constitution to the neglect of race-based disadvantage in laws presumed to be color blind. Finally, there is disagreement among philosophers about who the most disadvantaged or paradigm victims of racism are: Lewis Gordon (1995) has claimed that antiblack racism is more extreme than other forms, because of the historical association of darkness with sin in the Christian tradition; Native American philosophers refer to European conquest as a holocaust; Asian Americans claim group histories of exclusion in immigration law and exploitation as cheap labor.

There are also issues of whether nonwhites can be racist against whites and whether racism can be practiced by some members of the same race against others. Nonwhites can be individually racist against whites, although not institutionally because they do not have sufficient influence within major social institutions. Preferences for lighter skin color within nonwhite groups, as well as self-hatred on the grounds of nonwhite race would be examples of same-race racism.

AFFIRMATIVE ACTION

In 1965, according to U.S. Executive Order 11246, President Lyndon Johnson required that government officials take affirmative action (AA) to address the ongoing disproportionately low numbers of minorities directly and indirectly employed by the federal government. At that time, the concept of institutional racism was not widely accepted, but it was assumed that AA would override individual racism that could not otherwise be proved in hiring decisions. Arguments in favor of AA have been based on the value of minority role models, the justice of compensation and reparations for past wrongs, and the presumption that U.S. society has not ceased to disadvantage minorities on the grounds of race. Arguments against AA often proceed from the premise that minorities have gained formal and legal equality in the United States, to the claim that AA is unnecessary, and unjust because it penalizes otherwise deserving and innocent whites who are not responsible for past injustice.

In 2003 the U.S. Supreme Court's rulings in two University of Michigan cases offered a practical resolution of these disputes. In both *Grutter v. Bollinger* and *Gratz v. Bollinger*, the Court recognized the value of a diverse student body but used strict scrutiny in its rulings, determining whether two different forms of AA constituted a compelling government interest and were narrowly tailored to advance that interest. In *Grutter v. Bollinger*, the Court ruled in favor of the University of Michigan Law School's policy of considering the race of applicants holistically, as one factor among many; in *Gratz v. Bollinger*, the Court ruled against the University of Michigan's undergraduate admissions policy of uniformly giving the same number of points for minority racial identities. In its rulings in these and previous cases, the Court declared as unconstitutional, role model and compensation/reparation justifications for AA. However, in *Grutter v. Bollinger* the Court upheld the value of a critical mass of minority students, as opposed to tokenism. The Court's main justification for AA was the value of a racially integrated leadership class, which would in time make AA unnecessary.

Philosophers of race and racism are unlikely to accept judicial decisions on AA as the last word, because courts may revise or overturn previous rulings and legal reasoning has distinctive constraints, one of which is to assume that existing laws are effective. In the Michigan cases, the Supreme Court appeared to assume that formal legal equality guarantees equal opportunity. It therefore seemed to view AA as a strategy for achieving diversity, on the assumption of unequal ability, rather than a strategy for social justice. That is, the Court seemed to accept the fairness of admissions criteria that nonwhites disproportionately fail to meet and did not address the possibility of racism in the face of official race neutrality.

FURTHER ASPECTS OF MULTICULTURAL PHILOSOPHY

Whether multicultural academic philosophy will remain a distinct range of specializations or become part of the core curriculum is an open question. Africana Philosophy, which includes studies of race and racism, did become a recognized philosophical subfield by the 1990s.

Native American philosophers Anne Waters (2000) and V.P. Cordova (2000) have argued that Western philosophy has Christian religious foundations that are inimical to indigenist world views, a perspective that undermines beliefs in the universalism of traditional philosophy. Hispanic and Latin American philosophy has never been part of the recognized philosophical canon, although by 2000 it was clearly part of multicultural philosophy. Although Asian or Eastern philosophy has a long history as a distinctive body of knowledge, often addressed within comparative philosophy, the status and concerns of Asian Americans would be a new subject.

By 2005, the existence of philosophical multicultural research, published by academic book presses and journals, constituted a tradition capable of supporting graduate research and further professional scholarship, as well as the curricula of multicultural courses. Where studies of race have been limited to U.S. society, further work is likely to include international and world perspectives. Multidisciplinary approaches are likely to continue, drawing on studies in law, political science, sociology, public policy, and economics. Philosophy of science analyses of ideas of race in biological anthropology and population genetics could expand into ideas of race in the social sciences (for example, psychologist Roy Freedle [2003] has presented statistical data on standardized test scores, which indicate that minority students score higher than traditional white students on difficult questions). Feminist interest in racial differences among women adds a multicultural dimension to existing feminist philosophy. And finally, analyses of racism and its remedies are relevant to established work in the philosophy of education, as well as moral theory, ethics, and applied ethics.

In considering future directions for both intellectual and demographic multiculturalism in philosophy, and assessing progress at any given time, the subject itself suggests cross-national comparisons. In general, the extent of multiculturalism in philosophy seems to be more sensitive to external political, social, and demographic factors, than to purely intellectual interests in inclusion or exclusion. For example, as university subjects in the Soviet Union, *philosophy* referred to the work of Karl Marx and Vladimir Lenin, whereas Western philosophy, which was what Western Europeans and Americans called *philosophy*, was taught as a distinct and subsidiary subfield.

There have been two models of political and social pluralism that are relevant to multicultural philosophy, considered as an international subject. The assimilationist model encourages subordinate groups to achieve inclusion through their contributions to the common culture of dominant groups. The autonomous or diversity model advocates that subordinate groups participate in a shared but diverse common culture. American academic philosophy is becoming multicultural according to the diversity model.

In contrast, the trend in Great Britain has been to assimilate white women and minorities within the existing academic field of philosophy. Julian Baggini, editor of *The Philosophers' Magazine*, who interviewed sixteen leading British philosophers about their profession, observed in 2003 that participants in the main British philosophy conference were often all white, with a very small minority of women. Since the 1970s, Canada has had a strong multicultural political movement that has been reflected in its intellectual life. James Tully (1995) has examined how constitutionalism can coexist with diversity in *Strange Multiplicity*. And in *Multicultural Citizenship* (1995), Will Kymlicka argues that immigrant and indigenous groups in a multicultural society have disparate needs.

Developing parts of the world have perhaps been more interested in examining and constructing their own national and cultural intellectual perspectives, in a postcolonial era. Their work may be included in multicultural studies for Northern and Western audiences—for example, V. Y. Mudimbe's *Nations, Identities, Cultures*. But, multiculturalism for postcolonial critics is more likely to be a matter of deconstruction than inclusion. For example, in *Dislocating Cultures* (1997), Uma Narayan examines how British representations are an integral part of what is accepted as Indian culture and its products. Nevertheless, it is increasingly difficult to generalize about scholarly trends in multiculturalism. Chinese academics have launched *The Journal of Multicultural Discourses*, a forum for multicultural approaches to language, communities of discourse, cultural and literary criticism, and comparative studies, which will aim to be multidisciplinary across the social sciences and humanities, including philosophy.

See also Affirmative Action; Business Ethics; Enlightenment; Feminist Philosophy; Hume, David; Kant, Immanuel; Lenin, Vladimir Il'ich; Marx, Karl; Pluralism; Racism; Sartre, Jean-Paul; Toleration.

Bibliography

Alcoff, Linda, and Eduardo Mendieta. *Identities: Race, Class, Gender, and Nationality*. Malden, MA: Blackwell, 2003.

Appiah, K. Anthony. "'But Would That Still Be Me?': Notes on Gender, Race, Ethnicity as Sources of Identity." *The Journal of Philosophy* 77 (10) (1990): 493–499.

Appiah, K. Anthony, and Amy Gutman. *Color Conscious: The Political Morality of Race*. Princeton, NJ: Princeton University Press, 1993.

Baggini, Julian. "Two Philosophies, Separated by a Common Language." *The Chronicle of Higher Education* (2003).

Blackburn, Daniel G. "Why Race Is Not a Biological Concept." In *Race and Racism in Theory and Practice*, edited by Berel Lang. Lanham, MD: Rowman and Littlefield, 2000.

Boxill, Bernard. "The Morality of Reparation." *Social Theory and Practice* 2 (1) (1972): 113–122.

Cavalli-Sforza, Luigi Luca, and Francesco Cavalli-Sforza. *The Great Human Diaspora*. MA: Addison-Wesley, 1995.

The Civil Rights Project. *Reaffirming Diversity: A Legal Analysis of the University of Michigan Affirmative Action Cases: A Joint Statement of Constitutional Law Scholars*. Cambridge, MA: The Civil Rights Project at Harvard University, 2003. Available from http://www.civilrightsproject.harvard.edu.

Cohen, Carl, and James P. Sterba. *Affirmative Action and Racial Preferences: A Debate*. Oxford: Oxford University Press, 2003.

Cordova, V. P. "Exploring the Sources of Western Thought." In *Women of Color and Philosophy*, edited by Naomi Zack. Camden, MA: Blackwell Publishers, 2000.

Dubinin, N. P. "Race and Contemporary Genetics." In *Race, Science, and Society*, edited by L. Kuper. New York: Columbia University Press, 1965.

Eze, Emmanuel Chukwudi, ed. *Race and the Enlightenment: A Reader*. Cambridge, MA: Blackwell, 1997.

Freedle, Roy O. "Correcting the SAT's Ethic and Social-Class Bias: A Method for Reestimating SAT Scores." *Harvard Educational Review* 73 (1) (2003): 1–43.

Garcia, J. L. A. "Racism as a Model for Understanding Sexism." In *Race/Sex: Their Sameness, Difference, and Interplay*, edited by Naomi Zack. New York: Routledge, 1997.

Gordon, Lewis. *Bad Faith and Antiblack Racism*. Atlantic Highlands, NJ: Humanities Press, 1995.

Gracia, Jorge. *Hispanic/Latino Identity: A Philosophical Perspective*. Oxford: Blackwell, 2000.

Hallman, Max O., ed. *Traversing Philosophical Boundaries*. Belmont, CA: Wadsworth/Thompson Learning, 2003.

Hardimon, Michael O. "The Ordinary Concept of Race." *The Journal of Philosophy* 9 (2003): 437–455.

Hill, Jason. *On Becoming a Cosmopolitan: What It Means to Be a Human Being in the New Millennium*. Lanham, MD: Rowman and Littlefield, 2000.

Kymlicka, Will. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. New York; Oxford: Clarendon and Oxford University Press, 1995.

Lang, Berel. "Metaphysical Racism." In *Race/Sex: Their Sameness, Difference, and Interplay*, edited by Naomi Zack. New York: Routledge, 1997.

Leroi, Armand Marie. "A Family Tree in Every Gene." *New York Times*, March 12, 2004, p. A23.

Mills, Charles. *The Racial Contract*. Ithaca, NY: Cornell University Press, 1997.

Mudimbe, V. Y. *Nations, Identities, Cultures*. Durham, NC: Duke University Press, 1995.

Narayan, Uma. *Dislocating Cultures: Identities, Traditions, and Third-World Feminism (Thinking Gender)*. New York: Routledge, 1997.

Outlaw, Lucius, Jr. *On Race and Philosophy*. New York: Routledge, 1996.

Outlaw, Lucius, Jr. "2002–2003 Report of the American Philosophical Association Committee on Inclusiveness." Available from http://www.apa.udel.edu/apa/governance/committees/Inclusiveness/report2003.html.

Schutte, Ophelia. *Cultural Identity and Social Liberation in Latin American Thought*. Albany: State University of New York Press, 1993.

Spelman, Elizabeth V. *Inessential Woman: Problems of Exclusion in Feminist Thought*. Boston: Beacon Press, 1988.

Spratt, Scott. *Native Pragmatism: Rethinking the Roots of American Philosophy*. Bloomington: Indiana University Press, 2002.

Sterba, James P. *Three Challenges to Ethics: Environmentalism, Feminism, and Multiculturalism*. New York, Oxford: Oxford University Press, 2002.

Tully, James. *Strange Multiplicity*. New York; Cambridge, U.K.: Cambridge University Press, 1995.

Waters, Anne Schullherr. "That Alchemical Bering Strait Theory! or Introducing America's Indigenous Sovereign Nations Wolrdviews to Informal Logic Courses." In *Women of Color and Philosophy*, edited by Naomi Zack. Malden, MA: Blackwell, 2000.

Waters, Anne Schulherr, ed. *American Indian Thought: Philosophical Essays*. Malden, MA: Blackwell, 2003.

Webster, Yehudi. *Against the Multicultural Agenda: A Critical Thinking Alternative*. Westport, CT: Praeger, 1997.

Wilson, Robin. "A Kinder, Less Ambitious Professoriate." *The Chronicle of Higher Education* (2002): 10–11.

Young, Iris Marion. *Justice and the Politics of Difference*. Princeton, NJ: Princeton University Press, 1990.

Zack, Naomi. "American Mixed Race: Theoretical and Legal Issues." *Harvard Blackletter Law Journal* 17 (2001):33–46.

Zack, Naomi. *Philosophy of Science and Race*. New York: Routledge, 2002.

Zack, Naomi. *Race and Mixed Race*. Philadelphia: Temple University Press, 1993.

Zack, Naomi. "Race and Racial Discrimination." In *Oxford Handbook of Practical Ethics*, edited by Hugh Lafolette. Oxford: Oxford University Press, 2002.

Zack, Naomi. *Thinking About Race*. 2nd ed. Belmont, CA: Thompson Wadsworth, 2005.

Naomi Zack (2005)

MULTIPLE REALIZABILITY

Multiple realizability is a key issue in debates over the nature of mind and reduction in the sciences. The subject consists of two parts: multiplicity and realizability. "Multiplicity" designates variability in the mechanism and materials from which a particular type of thing can be made. "Realizability" designates a specific relation that exists when there is the stated variability.

REALIZABILITY

Apart from the broad folk notion of realization meaning that a thing is made real, philosophers apply several technical notions of realization to paradigm cases such as computational states realized by engineering states, minds realized by brains, and persons realized by their bodies. The technical notions fall into three broad traditions: mathematical, logico-semantic, and metaphysical.

The mathematical tradition equates realization with a form of mapping between objects. Generally speaking, *x* (mathematically) realizes *y* because elements of *y* map onto elements of *x*. The notion is useful for many purposes, for example, when constructing a formal model of a particular domain. However, since mapping extends to models as well as reality, it fails to distinguish between simulated versus genuine realizations. Heavenly stars can be mapped onto grains of sand, but grains of sand do not realize heavenly stars in any genuine sense. Similarly, the mental states described by a cognitive program can be mapped onto unthinking groups of things, but unthinking groups of things do not realize mental states in any genuine sense (Block 1978). Hence, to capture what is essential to genuine realization, William Lycan (1987) adds ideas about evolutionary function, while David Chalmers (1994) emphasizes facts about the causal structure of a system. To present Chalmers's idea and cast in terms of a computational model that informs the literature cited, a set of mental properties that constitute the cognitive program of a system is realized by a set of engineering properties possessed by that system if and only if (a) there is a one-to-one mapping between instances of the two sets of properties, and (b) the engineering involved has the causal structure to satisfy the computational state transitions required by the program.

The logico-semantic tradition translates realization into an interpretation of symbolic objects. Generally speaking, *x* (semantically) realizes *y* because *x* can be interpreted to meet the conditions for satisfying the term "*y*." Thus, logicians say that a set of objects is the realization of a formal language when the objects satisfy the predicates of that language (Tarski 1936/1956). Being a matter of semantic interpretation, might appear irrelevant to paradigm cases of realization whereby one thing (engineering or brains) generates or produces another thing (computation or minds). Yet Daniel Dennett (1978) addresses such cases by employing a method of agent interpretation, in effect turning the interpretation of symbols into an interpretation of rational symbol systems. Roughly, a set of mental properties that constitutes a system's cognitive program is realized by a set of engineering properties possessed by that system if and only if (a) the system's behavior supports an interpretation according to which instances of the computational properties are internal symbols involved in the operations of the system, and (b) it is rational for the system to possess those symbols and operations under the stated interpretation.

Finally, the metaphysical tradition views realization as a species of determination between objects. Generally speaking, *x* (metaphysically) realizes *y* because the properties of *x* determine the properties of *y*. Unlike other forms of determination, philosophers see a very close connection in paradigm cases of metaphysical realization. Regarding the particulars, some philosophers add that instances of realized and realizing properties occur at the same time, with the former composed out of the latter (Tye 1995). Regarding the properties, Stephen Yablo (1992) applies the notion of determinables and determinates by maintaining that a realized property stands to a realizing property as the determinable color red stands to its more determinate color scarlet. So human neurophysiology is a way of being a mind, like scarlet is a way of being red. In a different vein, Sydney Shoemaker (2001) employs metaphysical and set-theoretic notions by viewing the causal powers of a realized property as a subset of the causal powers of its realizing property. So mental abilities are a mere portion of the causal capacities of the appropriate engineering systems.

Many philosophers explain realization in terms of functionalism, the leading doctrine in the philosophy of mind. On this view, mental processes are understood by the functions they perform and not by the materials that realize the processes. On one popular version, each mental property is a higher-order property whose nature is defined as the possession of a lower-order physical property that plays an associated functional role. To present this idea in computational format, a set of mental properties that constitutes a system's cognitive program is realized by a set of engineering properties possessed by that system if and only if (a) the mental properties are higher-order properties that require lower-order physical properties to play their associated functional roles, and (b) the physical engineering properties of the system play the required functional roles.

MULTIPLE REALIZABILITY

Multiple realizability is a kind of variability in materials that philosophers call "property variability" or "compositional plasticity." Functionalists have this variability in mind when they observe that different physical properties

can play the same functional role in different individuals. Indeed, this observation is commonplace in computer science. Thus Alan Turing judged that the specific physical properties of an engineering system are unimportant for a theory of computation because the same computational function can be performed by systems with different engineering:

> Importance is often attached to the fact that modern digital computers are electrical, and that the nervous system is also electrical. Since Babbage's machine was not electrical, and since all digital computers are in a sense equivalent, we see that this use of electricity cannot be of theoretical importance. … If we wish to find [computational] similarities we should look rather for mathematical analogies of function. (Turing 1950, p. 439)

That is, while an instance of a given physical property may be sufficient to realize a computational property, as when the human brain computes addition, nevertheless that same physical property is not necessary. Other systems with quite different physical properties can compute addition: someone with a different neurophysiology, an artificial machine with a microprocessor, and so on. So the key to property variability is that sufficient conditions for the realization of higher-level properties are not necessary conditions.

More formally, property *G* is lawfully sufficient for property *F* if, as a matter of physical law, *F* is realized when *G* is realized. But *G* is not a necessary condition for *F* if *F* can be realized without *G*. For example, *G* is sufficient but not necessary for *F* if *F* is a computational function that can be realized on some occasion without the property *G* of having a human neural assembly but with the property *H* having an artificial microprocessor. To incorporate this idea into a formal definition in which *A* is a set of realized properties and *B* its realizing base:

> Property *F* in set *A* has variability with respect to set *B* if and only if there exist properties *G* and *H* in *B* such that
>
> (i) it is possible that *G* and *F* but not *H* are realized, and, as a matter of physical law, if *G* is realized then so is *F*;
>
> (ii) it is possible that *H* and *F* but not *G* are realized, and, as a matter of physical law, if *H* is realized then so is *F*; and
>
> (iii) there is no property *K* in set *B* such that, as a matter of physical law, *F* is realized if and only if *K* is realized. (Endicott 1994)

Clauses (i) and (ii) jointly express a minimal form of property variability, while the addition of clause (iii) expresses a form of deep property variability by guaranteeing that the variability of *F* with respect to *G* and *H* is not a superficial fact that masks an underlying common property, that is, ruling out any property in *B* that is lawfully coextensive with *F*.

Property variability also comes in degrees. Being a planet has many physico-chemical realizations (all possible minerals constituting large dense bodies in orbit), while being jade has only two such realizations (jadeite and nephrite). Accordingly, there is the project of explaining how variability arises and why. Dennett (1991) appeals to the forces of evolution, claiming that the brain developed variability in how it realizes cognitive functions to enhance the organism's ability to adapt to a changing environment. Robert Batterman (2000) offers a more general explanation based upon the notion of universality in physics, which concerns the procedure of finding similarities in behavior among physically diverse systems.

But however property variability is explained, it appears widespread. Neural plasticity is well documented (Johnson 1993). In particular, the brain is capable of compensatory plasticity, in which areas in the brain formerly dedicated to one cognitive task can, after injury or disease, become dedicated to another cognitive task (Rauschecker 1995). The brain is also capable of experience-dependent plasticity, in which the basic wiring of the brain is refined by an individual's sensory experience, creating individual differences in how the brain realizes mental functions (King 1999). At a more abstract level, functional properties are variable with respect to different physical properties, shapes can be shared by different kinds of matter, and the same spatial patterns can be discerned among physically distinct structures.

SUBSEQUENT DEBATE OVER IDENTITY AND REDUCTION

Hilary Putnam (1967/1975) and Jerry Fodor (1974/1981) developed an argument concerning special sciences such as psychology that was then extended by David Hull (1974) to the biological sciences. As a result of this argument, it became the dominant opinion among philosophers in the late-twentieth century that property variability supplies adequate evidence against type identity and physical reduction. Type identity is the theory that mental properties are identical with physical properties. And physical reductionism is the doctrine that all sci-

entific theories reduce to basic physical theories. Below is an outline of Putnam's and Fodor's multiple-realizability argument:

> (1) If a mental property *F* is identical with or reducible to a physical property *G*, then, as a matter of physical law, *F* is realized if and only if *G* is realized (they must be lawfully coextensive).
>
> Yet (2) this requirement that identical properties be lawfully coextensive is not met in cases where property variability applies, because *F* can be realized without *G*.
>
> So (3) mental property *F* is not identical with or reducible to physical property *G*.

Yet the issue is not settled. There are several responses, which divide into three main areas of discussion: variability, the notion of a property, and reduction versus identity.

VARIABILITY REEXAMINED. Jaegwon Kim (1972) challenges premise (2) by observing that physical differences between individuals who share the same psychology does not imply that no physical property is realized when and only when a given mental property is realized. In other words, the minimal form of property variability expressed by clauses (i) and (ii) does not imply the deep property variability captured by clause (iii) that rules out mental-physical identities. Moreover, Kim believes that the world reveals interlevel identities along with minimal property variability. For example, temperature is identical with mean kinetic energy in ideal gases, yet two aggregates of molecules with the same temperature will differ physically by having constituent molecules with different positions and directions. Accordingly, reductionists are optimistic that neuroscience will discover mental-physical identities, like the specialized Hubel-Wiesel cells, which detect edges in a visual field, or the identification of visual awareness with 40–70 Hz oscillations in the cortical system (Crick and Koch 1990/1997). Indeed, Patricia Churchland (1986) foresees that portions of psychology and neuroscience will coevolve to a point where they reductively converge because their methodologies are interdependent, as when neuroscientists employ psychofunctional criteria to identify brain structures, thereby establishing mental-physical correlations.

Antireductionists counter that, while mere physical differences do not guarantee that each mental property is not coextensive with some physical property, deep variability remains extremely plausible, given the functional nature of mental phenomenon and the actual record of how cognitive systems are built in a physically variable way. Consider again the case of computation. Having devised computational mechanisms that exhibit quite different engineering properties—from electrical charges passing through silicon pathways to light signals flashing across optical channels—scientists cannot point to a single necessary and sufficient physical condition for any computational function. So it seems unlikely that computation is like temperature in ideal gases, whose necessary and sufficient physical condition is mean kinetic energy. Moreover, antireductionists claim that neuroscientific discoveries only establish mental-physical correlations, not the coextensions that support property identity. Thus, various systems of computer vision carry out algorithms for edge detection, which shows that the activity of Hubel-Wiesel cells is sufficient but not necessary for that function. Furthermore, even if artificial systems are discounted and psychological theory is restricted to biological systems such as mammals, and even if neuroscience employs psychofunctional criteria to identify mammalian brain structures, those identifications must be compatible with compensatory and experience-dependent plasticity as well as any other physical variations that arise from evolution (Rosenberg 2001). This makes the identification of particular types of mental functioning with coextensive physical functioning unlikely.

RECONCEPTUALIZED PROPERTIES. Many reductionists challenge premise (2) in Putnam's and Fodor's argument by reconceptualizing the pertinent properties. On the side of the mental, David Lewis (1969) suggests that mental properties are lawfully coextensive with physical properties when the former are narrowly conceived species-specific properties. Thus, unlike pain per se, which might be realized in physically different ways across various species, pain *in human beings* may be lawfully coextensive with a human neurophysical property (see also Kim 1972, 1992/1993). On the side of the physical, Kim (1998) suggests that mental properties are lawfully coextensive with physical properties when the latter are broadly conceived disjunctive properties. Thus, the property of having pain is lawfully coextensive with the disjunctive property of having a particular human neural assembly *or* a particular extraterrestrial neural assembly *or*, and so on. Here the disjunctive property includes every possible realization of pain.

Yet, regarding species-specific mental properties, antireductionists counter that psychological theory also requires more general properties to explain cross-species generalizations. Moreover, they argue that even if theories

are restricted to species-specific properties, there remains the fact that variability occurs within a species and even the same individual over time (Horgan 2001). As for disjunctive physical properties, some critics deny that they exist, because they do not guarantee meaningful statements of similarity among objects or plausible statements about the causal powers of objects (Teller 1983). Others argue that disjunctive predicates do not always express natural kinds, yet projectible natural-kind predicates are needed for scientific prediction and explanation (Block 1997).

REDUCTION VERSUS IDENTITY. Finally, rather than cast doubt upon premise (2) in Putnam's and Fodor's argument, some philosophers promote views of reduction that do not require the identities at issue in premise (1). Granted, on the traditional account of scientific reduction associated with Ernest Nagel, one theory reduces to a more basic theory when the former can be deduced from the latter by means of connecting principles that express property identities. But there are other accounts that advertise no requirement concerning lawful coextensions which support intertheoretic property identities, including variations on approximate reduction (Paul Churchland 1979, Bickle 1998) and physicalist interpretations of functionalism (Kim 1998).

Critics counter that, among other problems, traditional connecting principles resurface within these alternatives (Endicott 1998, Marras 2002). Critics also add that, to the extent that any account avoids property identities, it is best understood as a model of scientific replacement, not reduction. In the end, philosophers have proposed many notions of reduction. But the fundamental metaphysical question remains: whether the properties of special sciences and physical sciences are identical or whether, because of multiple realizability, they fail to be identical.

See also Computationalism; Dennett, Daniel C.; Fodor, Jerry A.; Functionalism; Mind-Body Problem; Nagel, Ernest; Physicalism; Putnam, Hilary; Reduction; Reductionism in the Philosophy of Mind; Turing, Alan M.

Bibliography

REALIZABILITY

Block, Ned. "Troubles with Functionalism." In *Readings in Philosophy of Psychology*, edited by Ned Block. Vol. 1, 268–305. Cambridge, MA: Harvard University Press, 1978.

Chalmers, David. "On Implementing a Computation." *Minds and Machines* 4 (1994): 391–402.

Dennett, Daniel. Introduction. In his *Brainstorms: Philosophical Essays on Mind and Psychology*, xi–xxii. Cambridge, MA: MIT Press, 1978.

Lycan, William. *Consciousness*. Cambridge, MA: MIT Press, 1987.

Polger, Thomas. *Natural Minds*. Cambridge, MA: MIT Press, 2004.

Shoemaker, Sydney. "Realization and Mental Causation." In *Physicalism and Its Discontents*, edited by Carl Gillet and Barry Loewer, 74–98. Cambridge, U.K.: Cambridge University Press, 2001.

Tarski, Alfred. "On the Concept of Logical Consequence" (1936). In his *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, 409–420. Oxford: Clarendon Press, 1956.

Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press, 1995.

Yablo, Stephen. "Mental Causation." *Philosophical Review* 101 (1992): 245–280.

MULTIPLE REALIZABILITY

Batterman, Robert. "Multiple Realizability and Universality." *British Journal for the Philosophy of Science* 51 (2000): 115–145.

Dennett, Daniel. *Consciousness Explained*. Boston: Little, Brown, and Co., 1991.

Endicott, Ronald. "Constructival Plasticity." *Philosophical Studies* 74 (1994): 51–75.

Heil, John. "Multiple Realizability." *American Philosophical Quarterly* 36 (1999): 189–208.

Johnson, Mark. *Brain Development and Cognition: A Reader*. Oxford: Oxford University Press, 1993.

King, Andrew. "Sensory Experience and the Formation of a Computational Map of Auditory Space in the Brain." *BioEssays* 21 (1999): 900–911.

Rauschecker, Josef. "Compensatory Plasticity and Sensory Substitution in the Cerebral Cortex." *Trends in Neurosciences* 18 (1995): 36–43.

Turing, Alan. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433–460.

SUBSEQUENT DEBATE OVER IDENTITY AND REDUCTION

Bechtel, William, and Jennifer Mundale. "Multiple Realizability Revisited: Linking Cognitive and Neural States." *Philosophy of Science* 66 (1999): 175–207.

Bickle, John. *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press, 1998.

Block, Ned. "Anti-reductionism Slaps Back." *Philosophical Perspectives* 11 (1997): 107–132.

Churchland, Patricia. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press, 1986.

Churchland, Paul. *Scientific Realism and the Plasticity of Mind*. New York: Cambridge University Press, 1979.

Crick, Francis, and Christopher Koch. "Towards a Neurobiological Theory of Consciousness" (1990). In *The Nature of Consciousness*, edited by Ned Block, Owen Flanagan, and Guven Guzeldere, 277–292. Cambridge MA: MIT Press, 1997.

Endicott, Ronald. "Collapse of the New Wave." *Journal of Philosophy* 95 (1998): 53–72.

Endicott, Ronald. "Species-Specific Properties and More Narrow Reductive Strategies." *Erkenntnis* 38 (1993): 303–321.

Fodor, Jerry. "Special Sciences: Still Autonomous After All These Years," *Philosophical Perspectives* 11 (1997): 149–163.

Fodor, Jerry. "Special Sciences" (1974). In his *Representations: Philosophical Essays on the Foundations of Cognitive Science*, 127–145. Cambridge, MA: MIT Press, 1981.

Horgan, Terence. "Multiple Reference, Multiple Realization, and the Reduction of Mind." In *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*, edited by Gerhard Preyer and Frank Siebelt, 205–221. Landham, MD: Rowman and Littlefield, 2001.

Horgan, Terence. "Nonreductive Materialism and the Explanatory Autonomy of Psychology." In *Naturalism: A Critical Appraisal*, edited by Steven Wagner and Richard Warner, 295–320. Notre Dame, IN: University of Notre Dame Press, 1993.

Hull, David. *Philosophy of Biological Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

Kim, Jaegwon. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.

Kim, Jaegwon. "Multiple Realization and the Metaphysics of Reduction" (1992). In his *Supervenience and Mind: Selected Philosophical Essays*, 309–335. London: Cambridge University Press, 1993.

Kim, Jaegwon. "Phenomenal Properties, Psychophysical Laws, and the Identity Theory." *Monist* 56 (1972): 177–192.

Lewis, David. Review of *Art, Mind, and Religion*. *Journal of Philosophy* 66 (1969): 23–35.

Marras, Ausonio. "Kim on Reduction." *Erkenntnis* 57 (2002): 231–257.

Pereboom, Derk, and Hilary Kornblith. "The Metaphysics of Irreducibility." *Philosophical Studies* 63 (1991): 125–145.

Putnam, Hilary. "The Nature of Mental States" (1967). In *Mind, Language, and Reality*. Vol. 2 of *Philosophical Papers*, 429–440. New York: Cambridge University Press, 1975.

Rosenberg, Alex. "On Multiple Realization and the Special Sciences." *Journal of Philosophy* 98 (2001): 365–373.

Sober, Elliot. "The Multiple Realizability Argument against Reductionism." *Philosophy of Science* 66 (1999): 542–564.

Teller, Paul. "Comments on Kim's Paper." *Southern Journal of Philosophy* 22, suppl. (1983): 57–61.

Ronald Endicott (1996, 2005)

MUQAMMIṢ, DAVID BEN MERWAN AL-

David ben Merwan al-Muqammiṣ was one of the first medieval Jews to respond to the philosophical challenge of Muslim rationalism. Nothing about his life is known with any certainty, but he probably flourished in the early years of the tenth century. According to the account given by the tenth-century Karaite historian Kirkisani, David al-Muqammiṣ was a native of Raqqa, in Mesopotamia. Born into the Jewish faith, Kirkisani stated, al-Muqammiṣ became a Christian and then studied philosophy and theology at the well-known school of Nisibis, in Syria. Later, as reported by Kirkisani, he returned to Judaism but is supposed to have made good use of his Christian learning in his commentaries on Genesis and Ecclesiastes, which have been lost. In the latter part of the nineteenth century some quoted fragments of al-Muqammiṣ's philosophical work were discovered in Judah ben Barzilai's Hebrew "Commentary on the *Sefer Yezirah*" (early twelfth century). In addition, a substantial section of al-Muqammiṣ's major work, *ʿIshrūn maqālāt* (Twenty Chapters), in the original Arabic, was found by Abraham Harkavy in 1898 in the Russian Imperial Library at St. Petersburg, but it was never published.

This fragmentary and incomplete knowledge enables us to assert that al-Muqammiṣ's thought was deeply rationalistic, influenced in this direction by the Muʿtazilites (Arab theologians). His philosophy was, like theirs, generally cast in an Aristotelian mold, modified by some Neoplatonic elements. He shared with all Muslim philosophers a rigorous view of the divine unity; possibly it was the crystallization of this conviction that led to his rejection of Christianity and his return to Judaism. His discussion of the nature of the concept of unity as applied to God led him to distinguish between several ways of speaking about unity in ordinary language and to realize that none of these ways suggests what we mean in speaking of the unity of God, which is unique. More generally, al-Muqammiṣ argued, whenever we use the language of description we imply comparison and classification; however, God is incomparable and unclassifiable. Strictly, then, whether we speak of God in the language of the Bible or in that of philosophy, our language cannot be understood in any ordinary sense. If God is One, then each expression we use in speaking of him must be synonymous with every other expression. To use a variety of different expressions adds nothing, therefore, to our description of God. Al-Muqammiṣ suggested, however—anticipating Moses Maimonides in this suggestion—that although the different attributions add nothing positive, they do have the value of denying their antonyms.

In al-Muqammiṣ, then, we have the first suggestion in medieval Jewish philosophy of the theory of negative attributes. On other matters, such as the doctrine of rewards and punishments, al-Muqammiṣ seems to have had no difficulty in blending the traditional thought of the rabbis into his rational system.

See also Aristotelianism; Islamic Philosophy; Jewish Philosophy; Maimonides; Neoplatonism; Rationalism.

Bibliography

WORKS BY AL-MUQAMMIṢ

ʿIshrūn maqālāt (Twenty Chapters), edited and translated by Sarah Stroumsa as *Dawud ibn Marwan al-Muqammis's Twenty Chapters.* Leiden: Brill, 1989.

WORKS ON AL-MUQAMMIṢ

Ben-Shammai, Haggai. "Kalam in Medieval Jewish Philosophy." In *History of Jewish Philosophy,* edited by D. Frank and O. Leaman. London: Routledge, 1997.

Blau, Joseph L. *The Story of Jewish Philosophy.* New York: Random House, 1962.

Husik, Isaac. *History of Mediaeval Jewish Philosophy.* New York: Macmillan, 1916.

"Muqammis." In *Jewish Philosophy Reader,* edited by D. Frank, O. Leaman, and C. Manekin. London: Routledge, 2000.

Stroumsa, Sarah. "Saadya and Jewish Kalam." In *Cambridge Companion to Medieval Jewish Philosophy,* edited by D. Frank and O. Leaman. Cambridge, U.K.: Cambridge University Press, 2003.

Vajda, Georges. *Introduction à la pensée juive du moyen âge.* Paris: Vrin, 1947.

Vajda, Georges. "Le probleme de l'unité de Dieu d'après Dawud ibn Marwan al-Muqammis." In *Jewish Medieval and Renaissance Studies,* edited by A. Altmann. Cambridge, MA: Harvard University Press, 1967.

J. L. Blau (1967)
Bibliography updated by Oliver Leaman (2005)

MURDOCH, IRIS
(1919–1999)

Iris Murdoch is best known to the world as a novelist—she wrote twenty-six—but she was a tutor in philosophy at Oxford University from 1948 until 1963 and wrote several influential essays on moral philosophy in the 1950s and 1960s. Her collection of three such essays, *The Sovereignty of Good* (1970), remains her most influential work. Her most sustained philosophical work is *Metaphysics as a Guide to Morals* (1993), a sprawling work ranging over an extraordinary range of topics and also a difficult work not enjoying the impact on philosophy of her earlier work. Murdoch also wrote on literature, religion, and art. Her thought is a unique appropriation of Platonic, Freudian, and existentialist themes.

Murdoch's thought emerged from, and against, British moral philosophy of the 1950s and 1960s (which she calls "linguistic philosophy"), perhaps best represented by Richard Hare's *Language of Morals* (though Murdoch does not mention Hare by name). This school of thought held that the techniques of linguistic analysis could illuminate moral concepts while remaining neutral regarding substantive moral views.

In "Sovereignty of Good over Other Concepts," Murdoch rejects this distinction. "Moral philosophy can not avoid taking sides and would-be neutral philosophers merely take sides surreptitiously" (Murdoch 1970, p. 78). British philosophy, Murdoch says, suggests that the moral life does not present us with moral concerns of great depth or urgency. Its behaviorist proclivity, implying that morality resides only in outer behavior, does away with the substantial inner life of the mind and, by implication, any notion of moral vision.

Murdoch was initially attracted to Jean-Paul Sartre's existentialism (she had met Sartre briefly in 1945) as a philosophy that one could actually live by and also as a philosophy that subjects individual consciousness to philosophical scrutiny. (In 1953 she published the critical but appreciative study *Sartre: Romantic Rationalist.*) Yet she came to feel that Sartre's moral philosophy was quite similar to linguistic philosophy in its faulty conception of moral agency and the moral life, despite the enormous differences in aspiration and mood in the two schools of philosophy. The "existentialist/behaviorist" view, as she frequently refers to the two views, sees the self as a solitary will and sees the core of moral agency as lying in the exertion of the will at the moment of choice. This solitary moral agency operates in a shared world of evaluatively neutral facts, with freedom as a central value, and confers value through choices.

Murdoch regards this conception of moral agency as entirely faulty. The moral agent perceives the world as saturated with value, and one's choosings arise almost automatically from how one antecedently perceives situations. Moral activity is not confined to outward behavior; seeing other persons in a just and accurate manner is moral activity, even if one never performs actions affecting such persons. Therefore, moral life does not sporadically occur only at moments of choice, but is pervades throughout the agent's existence, shaping the perceptions that issue in action. We erect structures of value around us, generally without recognizing that we are doing so.

Murdoch also chastises British moral philosophy for failing to focus centrally on how agents can morally improve—a task that she understood primarily as gaining a clear grasp of the moral reality outside themselves. To characterize the psychic process by which this is accomplished, Murdoch appropriates the term "attention" from Simone Weil, a French philosopher of the 1930s and 1940s who exerted a strong influence on her. By attending to the outer world, the moral agent becomes open and receptive to a reality other than oneself in a way uncontaminated by personal needs, fantasies, illusions, and the

like. Murdoch sometimes speaks of attention as a kind of love and is critical of contemporary moral philosophy for leaving no room for love as a central moral notion.

Murdoch's conception of moral reality takes two somewhat distinct directions. The first is Platonic. (Plato is the philosopher Murdoch embraces most unambivalently.) On her Platonic conception, the ultimate moral reality is a transcendent Good, as she says in "On 'God' and 'Good'," a "single perfect transcendent non-representable and necessarily real object of attention"—a description that Murdoch draws partly from religion, though she explicitly rejects traditional theism (Murdoch 1970, p. 55). Murdoch thinks of the Good as something that can be contemplated, that exerts a kind of magnetic pull, and from which moral agents can draw a moral energy to overcome selfishness. She faults linguistic philosophy for discrediting metaphysics, which she sees as required for rendering the idea of the Good intelligible, an idea she develops further in *Metaphysics as a Guide to Morals*.

The second strand in Murdoch's conception of moral reality is particular other persons, especially those emotionally close to us, Murdoch's favored context for moral attention in her novels as well as her philosophy. "The fat, relentless ego" revealed by Freud, with its self-serving fantasies and illusions, presents daunting obstacles to apprehending moral reality. Murdoch is also pessimistic that by turning one's attention inward, one can identify and perhaps dispel one's particular psychic obstacles. Self-knowledge, she thinks, is largely a delusion.

Murdoch offers no systematic account of how to attain a state of attention, how to know the morally real, but she offers a few examples of things that can take us out of ourselves toward a reality external to us: art, natural beauty, prayer, a foreign language with its own logic, which cannot be distorted by personal wishes or fantasies. Her central example is art, especially literature. Good literature portrays human situations and human truth in an accessible form that provides readers a way to get outside themselves to a moral reality. Indeed, Murdoch sees the production of literature too as a moral task, a task in which authors must keep their own fantasies and illusions from distorting the creation of their characters. Murdoch's philosophy of art, inseparable from her moral philosophy, is developed in *The Fire and the Sun: Why Plato Banished the Artists* and in several essays.

Murdoch contributed to moral philosophy's greater attention to moral psychology (especially moral perception) since the 1970s, and she occasionally speaks of virtue. On the whole, however, Murdoch's work does not readily fit within any of the familiar schools of contemporary moral thought, and her insights and perspective remain a largely untapped resource and a formidable challenge to moral philosophy.

See also Moral Psychology; Sartre, Jean-Paul; Weil, Simone.

Bibliography

WORKS BY MURDOCH

Sartre: Romantic Rationalist. New Haven, CT: Yale University Press, 1953.

The Sovereignty of Good. London: Routledge and Kegan Paul, 1970.

The Fire and the Sun: Why Plato Banished the Artists. Oxford, U.K.: Clarendon Press, 1977.

Metaphysics as a Guide to Morals. New York: Penguin Books, 1993.

Existentialists and Mystics: Writings on Philosophy and Literature. New York: Penguin Books, 1998. A collection of all of Murdoch's nonfiction writings, including *The Sovereignty of Good*, but excluding *Metaphysics as a Guide to Morals* and *Sartre*.

WORKS ON MURDOCH

Antonaccio, Maria. *Picturing the Human: The Moral Thought of Iris Murdoch*. New York: Oxford University Press, 2000.

Antonaccio, Maria, and Schweiker, William, eds. *Iris Murdoch and the Search for Human Goodness*. Chicago: University of Chicago, 1996.

Blum, Lawrence. *Moral Perception and Particularity*. New York: Cambridge University Press, 1994. Several essays related to Murdoch's moral philosophy.

Broackes, Justin, ed. *Iris Murdoch, Philosopher*. New York: Oxford University Press, 2006.

Holland, Margaret. "Touching the Weights: Moral Perception and Attention." *International Philosophical Quarterly* 38 (3) (1998): 299–312.

Lawrence Blum (2005)

MURO KYŪSŌ
(1658–1734)

Muro Kyūsō was a Japanese Confucianist who was instrumental in defending the Zhu Xi school of Neo-Confucianism as the official learning of the Tokugawa government. Born in Edo (Tokyo), he was a pupil of Kinoshita Junan (1621–1698) in Kyoto. In 1711 he became, through the recommendation of the scholar-statesman Arai Hakuseki (1657–1725), the official scholar of the Tokugawa government. He was commissioned to compile the *Rikuyu engi-tai* (Outline of principles of Confucianism) that in 1724 became the standard textbook on Zhu Xi's doctrine for all official schools. Muro in

his early years was not a follower of the Zhu Xi school; as he tells us in his *Shundai zatsuwa* (Conversations at Surugadai), it was only at the age of forty, after a long period of doubt, that he embraced Zhu Xi's thought. The doctrine was then under heavy attack by such of the "ancient learning" scholars as Yamaga Sokō, Itō Jinsai, and Ogyū Sorai. Muro believed he had been chosen to defend the teaching of Zhu Xi, and to this task he dedicated the rest of his life with unsparing zeal.

Muro's ideas are not strikingly original, but they have the power of sincerity and conviction. Typical are his denunciations of hypocrisy, a trait not so uncommon among formalist Confucians, and his insistence upon virtue as springing from the inner self; two of his favorite maxims were "Be true to the self" and "The root of evil lies in the innermost recesses of the mind." His ideas on the Godhead bear a similarity to the Christian conception of the attributes of God. The deity (or deities) is omnipresent and omniscient. He stressed self-vigilance and the realization of heavenly reason in human life. The heavenly order was to be reflected in the social one, thus consolidating the immutability of Tokugawa society. His sense of the indebtedness (*gi*) and the gratitude (*on*) man owes to Heaven, the earthly lord, the parent, and the teacher was bound to foster obedience rather than self-assertiveness. Muro opposed the scholars of the "ancient learning" school, who, with others, supported the emperor; Muro stood solidly for the Tokugawa government. He was also critical of Buddhism and Shinto. But the tide was against him; especially in vain was his effort to preserve the ancient spirit of the samurai who more and more assimilated into the merchant class.

See also Chinese Philosophy; Itō Jinsai; Japanese Philosophy; Ogyū Sorai; Virtue and Vice; Yamaga Sokō; Zhu Xi (Chu Hsi).

Bibliography

Muro's *Rikuyu engi-tai* (Outline of principles of Confucianism) was published in Kyoto in 1722. His *Shundai zatsuwa* (Conversations at Surugadai) is available in *Nihon rinri ihen* (Library on Japanese ethics; Tokyo, 1903), edited by Inoue Tetsujirō, Vol. VII, pp. 81–122; it has been translated by G. W. Knox as "A Japanese Philosopher," in *Transactions of the Asiatic Society of Japan* 20, Part I (1893): 28–133. See also W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 433–442.

Gino K. Piovesana, S.J. (1967)

MURPHY, ARTHUR EDWARD

(1901–1962)

Arthur Edward Murphy, an American philosopher, was the creator of the phrase "objective relativism." Murphy was born in Ithaca, New York, and received his training in philosophy at the University of California (A.B. in 1923, Ph.D. in 1926). He taught successively at California, Chicago, Brown, Illinois, Cornell, Washington, and Texas; at the last four he was department chairman.

Murphy attracted attention at an early age with his article "Objective Relativism in Dewey and Whitehead" (1927). He argued that the writings of these two influential philosophers exhibited a convergence on a common doctrine, which reversed a tradition of treating "objects as primary, as substantives, and events as characters of objects." In contrast, for John Dewey and Alfred North Whitehead "the event is substantive and objects are characters of events. Thus relatedness, in all its complexity and interconnections, is made basic for the objective world." Murphy, himself, supported this doctrine, which had a vogue for a time.

In 1930, however, Murphy attacked Whitehead's *Process and Reality* in his article "The Development of Whitehead's Philosophy." In later writings he repeatedly charged both Dewey and Whitehead, among other metaphysicians, with attempting to prove by speculative metaphysics what would better be offered as sheer speculation, to be tested in appropriate contexts. Commenting on Dewey, he wrote: "What Mr. Dewey says about cognition is true of it as he defines it and false of it as more ordinarily understood" ("Dewey's Epistemology and Metaphysics," in *The Philosophy of John Dewey,* edited by P. A. Schilpp, p. 210, Evanston and Chicago, 1939).

Throughout his career Murphy maintained an acquaintance with philosophers of varied opinions. As a graduate student on a traveling fellowship, he explored the philosophical currents of Europe in 1924–1925, when realism was at its height. During the 1930s his work as book editor of the *Journal of Philosophy* gave him occasion to examine and to pass judgment on the purpose and achievements of his generation and the previous one.

Murphy spent the year 1937–1938 in England, and from his remarks it is apparent that he was directly influenced by Ludwig Wittgenstein through reading the *Blue Book,* as well as indirectly through Wittgenstein's colleagues in England. He grew increasingly dissatisfied with speculative metaphysics, as may be seen in his contribu-

tions to the Dewey, G. E. Moore, and Whitehead volumes of the Library of Living Philosophers. His disillusionment with his own creation, objective relativism, is reported in "What Happened to Objective Relativism." Yet, to the end, it was his opinion that the speculative philosophers had opened roads to "a better understanding of the values that are basic to human life" than had most of the so-called analytic philosophers.

Murphy's strong convictions on the importance of philosophy in a liberal education led him to expend a great deal of time on the work of the Commission on Philosophy in American Education of the American Philosophical Association. His opinions on this subject are to be found in the chapters that he contributed to *Philosophy in American Education* (1945) and in his own essays.

Much of his work illustrates his expressed intent "to write philosophy … with explicit reference to contemporary issues" (*The Uses of Reason,* p. 5). His early concern with epistemology and metaphysics changed to a dominating preoccupation with the uses of reason in ethical and social enterprises. His last twenty years were directed toward the working out of a systematic account of ethics. Sketches of this attempt appear in the chapter, "The Context of Moral Judgment," in *The Uses of Reason,* and in his essays. Murphy made good use of his powers of assimilation and criticism in examining the great moralists with a view to extracting and identifying points that must be taken account of in any subsequently defensible ethical theory. At his death, he left a long manuscript, *The Theory of Practical Reason,* which elaborates his Carus Lectures of 1955, originally known as "An Enquiry concerning Moral Understanding."

See also Dewey, John; Epistemology; Ethics, History of; Metaphysics; Moore, George Edward; Philosophy of Education, History of; Realism; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

Murphy's *The Uses of Reason* was published by Macmillan in New York in 1943. He contributed Chs. 2, 3, and 10 to *Philosophy in American Education* (New York: Harper, 1945). *Reason and the Common Good. Selected Essays of Arthur E. Murphy,* edited by W. H. Hay and M. G. Singer (Englewood Cliffs, NJ: Prentice Hall, 1963), includes a bibliography and the papers on objective relativism and Whitehead mentioned in this article. See also *The Theory of Practical Reason,* edited by A. I. Melden (La Salle, IL: Open Court, 1965).

William H. Hay (1967)

MUSIC, PHILOSOPHY OF

Since the veritable renaissance of aesthetics and philosophy of art in the 1960s, there has been a clear tendency to deal with the individual arts as presenting philosophical problems peculiar to themselves. This is not to say that philosophy of art in general has not also been pursued. Ambitious theories of art, attempting to encompass all of the fine arts in synoptic definitions, have occupied some of the best philosophical minds of the period, and brought much needed clarity and rigor to the discipline. But alongside of this more traditional, Socratic project there has flourished a busy community of philosophers exercising their analytic skills on the individual problems of arts such as literature, painting, dance, photography, cinema, drama, architecture, and, of course, music, the topic here.

MUSIC AND THE EMOTIONS

The oldest and most persistently scrutinized philosophical question with regard to music is the question of its emotive character. Plato expressed the view that music has the power to engender emotive states in the listener. Aristotle made the intriguing, though puzzling, suggestion that music "imitates" or represents the emotions. But we know little, if anything, about what their music sounded like. And without that knowledge we are at a loss to know what these philosophers were talking about, and consequently what they were really saying about it.

Modern speculation on this matter began at the beginning of the seventeenth century, when the inventors of opera began to speculate about music as the source of emotive expression in the newly minted dramatic form. But the problem did not take on the form in which contemporary aesthetics deals with it until, in the late eighteenth century, instrumental music emerged as a major musical genre and *the* major genre in the philosophy of music.

In the past seventy years, the question has taken a schematic form: What are we saying when we say "The music is *sad*"? Some answers have been that the music makes us sad; that the music expresses the composer's sadness; that the music somehow symbolizes or represents sadness; that the music possesses sadness as a perceptual quality, just as an apple possesses redness; some combination of the above; and finally, that the music just is *not* sad and it is nonsense to say that it is. The majority view, at the turn of the century, is that the emotive properties of music are perceived properties of it, although opinion is divided about whether it also *arouses* the emo-

tions it is expressive of. Those who argue against arousal (Peter Kivy, for example), argue that emotions are aroused in ordinary life by beliefs formed about states of affairs, which the appropriate emotions then take as intentional objects, and that music cannot provide the necessary conditions for such arousal, nor is there evidence that listeners are so aroused. In contrast, those who argue for arousal (Stephen Davies and Jerrold Levinson) maintain that because the emotions aroused by music are not full-blooded emotions, but close enough to be taken for them, music does indeed have the power to arouse emotions, though these emotions do not give rise to the normal behavioral responses of real-life emotions.

FORMALISM

While the topic of music and the emotions has perhaps been the most talked about in music aesthetics since time out of mind, it is arguable that the vital center of philosophy of music has been, since the end of the eighteenth century, the debate over musical formalism. Immanuel Kant seemed to entertain no doubt that pure instrumental music, "absolute music," as it came to be called, was a purely formal art (although he acknowledged its emotive aspect), and because it lacked ideational content, he was reluctant to consider it one of the fine arts at all.

Arthur Schopenhauer pretty much settled the issue in favor of absolute music as a fine art. He did so by considering music a *representational* art form, and thus an art form conforming to the eighteenth-century dogma of *mimesis* (imitation). But the cost was heavy, for the cumbersome metaphysical underpinnings of his theory would hardly be countenanced by philosophers with modern philosophical sensibilities.

The first full-blown formalist account of absolute music, that of Eduard Hanslick (1825–1904), followed not too long after. In musical aesthetics, formalism, as Hanslick construed it and as it continued to be construed until the 1980s, is the doctrine that absolute music, as an art object, must be considered a purely formal structure in sound, with no emotive significance at all. But when some writers came to see that the emotive properties of music could themselves be construed as perceptual properties of music, they saw that a formalism with emotive properties as part of the formal structure is, in spirit, a formalism as well. This view has come to be called "enhanced formalism."

As things stand at the beginning of the twenty-first century, there are those, particularly in historical musicology, who find even enhanced formalism too pallid, and views of absolute music as a "narrative without words" are surfacing in great profusion. What had seemed to many to be an issue firmly settled in favor of formalism has now become an issue very much in doubt.

MUSICAL UNDERSTANDING

Closely related to the concept of musical form is that of musical understanding. Whether or not one is a formalist, one has to assume that understanding the pure musical fabric is a prerequisite for understanding anything beyond the pure fabric—narrative content, for example. In other words, one must hear what one is listening to *as music* before one can hear it as a story *in* music.

It is generally agreed that understanding music is a matter of hearing it as a *connected* series of events that makes musical sense to the listener. How this basic musical understanding is to be recognized and construed are contentious questions. Furthermore, there is substantial disagreement about whether or not musical understanding requires knowing and attending to the large structural elements of musical compositions and the musical techniques that may govern the connections between events. This disagreement extends to whether or not knowledge of what is known in the trade as music theory has any relevance to the appreciation and enjoyment of absolute music. In the 1990s these questions were hotly disputed. In *Music in the Moment*, Jerrold Levinson maintained that normal listening requires attention merely to the connections between short segments of musical texture present to immediate perception, in what he calls "quasi-hearing." In the opposite camp, Peter Kivy, in *Music Alone* and elsewhere, has argued that music-theoretical knowledge, though not essential to minimal musical understanding, enlarges the intentional object of musical understanding, thus increasing by orders of magnitude the satisfaction of the musical experience.

REPRESENTATION

The question of whether instrumental music is capable of anything like pictorial representation is not high on the list of questions that philosophers of music at the beginning of the twenty-first century concern themselves with, although in the heyday of nineteenth-century Romanticism it was much discussed as a matter of "practical" music aesthetics and was closely associated with the issue of absolute versus program music. There are those who claim that music in principle cannot pictorially represent but can only imitate sounds, which is obviously a very different matter. Others maintain that there are instances of pictorial representation in music, although of a very minimal kind. Those committed to more or less elaborate

narrative interpretations of the canon of absolute music are committed, at least implicitly, to some more liberal view of music's representational capacities, although little philosophical light on the issue has been forthcoming from that quarter.

WORDS AND MUSIC

As questions in musical practice, how words are set to music and what role words and music play in this give-and-take enterprise have been argued vigorously, sometimes acrimoniously, since the last half of the sixteenth century, with opera as the major motivating force. Whether these are philosophical questions is debatable. Nonetheless, in the literature after 1990 those who *do* think of themselves as philosophers have shown an increase in interest in opera as an art form worthy of separate scrutiny. Among the issues raised have been whether opera is basically a musical form or a literary form with music, how we are rationally to understand a drama with characters who sing rather than speak, how drama can accommodate itself to musical form, how we are to understand, on rational grounds, the ubiquitous orchestral presence in the sung drama, and what capacity the music in opera has of "saying" things, beyond the capacity of the libretto to do so. These debates have blurred, in an intellectually healthy way, the boundaries between philosophy and various musical disciplines. At the same time, those outside both the philosophical and musical academic communities have made substantial contributions to the philosophical discourse.

Perhaps the central philosophical issue in the words-music debate is best revealed by the title that Joseph Kerman, a musicologist by trade, gave to his groundbreaking, widely admired book *Opera as Drama*. On Kerman's view, opera is to be viewed, at its best, as principally a form of drama, *dramma per musica*, in the venerable Latin phrase. Taking the opposite view, Peter Kivy, in *Osmin's Rage*, has put the emphasis, not on opera as *drama*, but rather on opera as *music*, drama-made-music, as he terms it.

THE WORK

Whatever one may think about the philosophical credentials of some of the questions that philosophers of music interest themselves in, the question of the ontological status of the musical work seems unequivocally philosophical. Who *else* but a philosopher, it might well be asked, would raise such a question, or be interested in the answer?

Musical ontology emerged in the 1960s in the form of two opposing answers to the question, What is a musical work? The term "art object" clearly suggests the kind of artwork that can, at least on first reflection, be identified with a physical object, locatable in space and time. But if the "object" in question is a musical work, it seems clear that it is not located anywhere. The *Mona Lisa* is in the Louvre. Where is Beethoven's *Fifth Symphony*?

Nevertheless, there are physical objects, broadly speaking, associated with musical works, namely their performances. One direction in which musical ontology went was the Platonic direction, taking musical works as universals or types, performances as their instances or tokens. The other direction, eschewing the specter of timeless, nonphysical Platonic entities, identified the musical work with the class of its performances. Both directions have problems, but the Platonic model, somewhat surprisingly, has been the one most exploited.

The major problem of musical Platonism has been the apparent conflict between two basic intuitions. Platonic entities are timeless, and hence cannot have come into being, whereas musical works do indeed come into being, are created, through the labor and inspiration of their composers. Platonists of the more doctrinaire kind have tried to argue that we can preserve our notion of composers as inspired, "creative" artists, in some sense or other, while biting the Platonic bullet and affirming that musical works are discovered rather than brought into being. Other, more moderate Platonists have opted for a kind of universal or type that comes into being in the composer's creative act but, in other respects, preserves the character of a Platonic universal or type so as to make the universal/particular or type/token distinction suitable for what they want to say about the relation between works and their performances. The latter approach seems to be more popular at the beginning of the twenty-first century, while the attempt to identify works with classes of their performances seems just about dead in the water.

THE PERFORMANCE

Since the most popular analyses of the musical work construe it as some kind of universal, with performances as the particulars, one would expect a substantial literature on musical performance. But until the late 1990s, this had not been so, it being assumed that performers and performances are philosophically transparent, presenting no conceptual puzzles. Then in the 1990s a movement in the *practical* world of performer and performance, the movement for so-called "historically authentic performances," began to generate considerable interest among philoso-

phers in the relation between performance and work, performer and composer. The historicist project in musicology, so long directed at establishing musical texts that are historically authentic, became, in the 1990s, directed as well at the historical authenticity of the musical performance of that authenticated text, the practical result being that more and more performances of music composed prior to the nineteenth century are attempts to reproduce, both physically and in interpretation, the kind of performance that the composer himself had in mind when he composed it.

After the turn of the century, philosophers began to cast an analytic eye on the concept of the historically authentic performance and on the aesthetic imperative that supposedly drives it. What is a historically authentic performance? One that reproduces a physical object or an intentional one? Does the integrity of the musical text require a historically authentic performance, or does the text survive an unabashedly modern one? Is the performer an artist in his own right, as tradition would have it, or is he the composer's machine? Is there an ideal performance of a work, and is it the historically authentic one? These questions have begun to generate articles and books of interest not only to the philosophical community but also to the musical community as well. Moreover, what the musical community has written about performance is now undergoing philosophical scrutiny. The results are not yet in.

THE REWARDS OF LISTENING

Finally, what contribution of value does the art of absolute music make to the human experience. What kind of satisfaction does it provide? Schopenhauer argued that since absolute music satisfies in the same manner as the other fine arts, which are unquestionably representational arts, absolute music too must be a representational art. He then cast about for an object that absolute music might represent, fixing on the metaphysical will—a result that few today would find plausible. Be that as it may, those who interpret the absolute-music canon in narrative terms are implicitly committed to Schopenhauer's general argument, if not to his conclusion about music's relation to the will. For the quest for stories in symphonies assumes that the satisfaction provided by such music requires an account, and since the satisfaction of temporal art forms lies in their storytelling capacity, the same must be true for the temporal art of absolute music. (Schopenhauer himself, however, does not carry his argument to this extreme.)

Formalists, of course, must find other sources for the value and satisfaction of absolute music. One answer, distinctly in the spirit of Schopenhauer, is that absolute music provides a kind of escape, a liberation from the world, from this veil of tears, into a world of pure sonic forms. The narrative and representational arts, anchored in this world as they are, cannot provide this liberation. Another answer simply rejects the question. There is no mystery about the satisfactions of absolute music. They lie simply in all the components of absolute music that music critics, analysts, and theorists talk about. It is obvious why these components please us. No further answer, it is claimed, is either needed or available.

Is the satisfaction of absolute music a mystery or a pseudomystery? Whatever the answer, absolute music, since the mid-1950s, has become a topic of intense interest in the philosophy of art, and the philosophy of music has become a recognized subdiscipline of the field. The interest shows no signs of abating.

See also Aesthetics, History of; Art, Definitions of; Art, Expression in; Art, Formalism in; Art, Ontology of; Art, Representation in; Art, Style and Genre in.

Bibliography

MUSIC AND THE EMOTIONS

Aristotle. *Politics*. In *The Complete Works of Aristotle: The Revised Oxford Translation*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984. Book 8, chapter 5.

Davies, Stephen. *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press, 1994. Chapters 4–6.

Hanslick, Eduard. *On the Musically Beautiful*. Translated by Geoffrey Payzant. Indianapolis, IN: Hackett, 1986. Chapters 1, 2, and 4.

Kivy, Peter. *The Corded Shell: Reflections on Musical Expression*. Princeton, NJ: Princeton University Press, 1980.

Langer, Susanne K. *Philosophy in a New Key*. New York: Mentor Books, 1959. Chapter 8.

Levinson, Jerrold. *Music, Art, and Metaphysics*. Ithaca, NY: Cornell University Press, 1990. Chapter 13.

Levinson, Jerrold. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996. Chapter 6.

Plato. *Republic*. In *The Collected Dialogues of Plato*. Princeton, NJ: Princeton University Press, 1963. Books 3 and 10.

Pratt, Carroll C. *The Meaning of Music*. New York: McGraw-Hill, 1932.

Strunk, Oliver, ed. *Source Readings in Music History: From Classical Antiquity through the Romantic Era*. New York: Norton, 1950. Section 8.

FORMALISM

Gurney, Edmund. *The Power of Sound*. London: Smith, Elder, 1880.

Hanslick, Eduard. *On the Musically Beautiful*. Translated by Geoffrey Payzant. Indianapolis, IN: Hackett, 1986. Chapter 3.

Kivy, Peter. *The Fine Art of Repetition*. Cambridge, U.K.: Cambridge University Press, 1993. Chapters 18 and 19.

Kivy, Peter. *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca, NY: Cornell University Press, 1990.

McClary, Susan. *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: University of Minnesota Press, 1991. Chapter 3.

Meyer, Leonard. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.

Newcomb, Anthony. "Once More 'Between Absolute and Program Music': Schumann's Second Symphony." *Nineteenth-Century Music* 7 (1984): 233–250.

MUSICAL UNDERSTANDING

Kivy, Peter. *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca, NY: Cornell University Press, 1990. Chapter 6.

Levinson, Jerrold. *Music in the Moment*. Ithaca, NY: Cornell University Press, 1997.

Tanner, Michael. "Understanding Music." *Proceedings of the Aristotelian Society*, supp. vol. 59.

REPRESENTATION

Kivy, Peter. *Sound and Semblance: Reflections on Musical Representation*. Princeton, NJ: Princeton University Press, 1984.

Robinson, Jenefer. "Music as a Representational Art." In *What Is Music?* edited by Philip Alperson. University Park: Pennsylvania State University Press, 1994. Chapter 6.

Scruton, Roger. "Representation in Music." *Philosophy* 51 (1976).

WORDS AND MUSIC

Kerman, Joseph. *Opera as Drama*. New York: Vintage Books, 1956.

Kivy, Peter. *Osmin's Rage: Philosophical Reflections on Opera, Drama, and Text*. Princeton, NJ: Princeton University Press, 1988.

Levinson, Jerrold. *The Pleasures of Aesthetics*. Ithaca, NY: Cornell University Press, 1996. Chapter 4.

Robinson, Paul. *Opera and Ideas from Mozart to Strauss*. New York: Harper and Row, 1985.

THE WORK

Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN: Bobbs-Merrill, 1968. Chapters 4 and 5.

Ingarden, Roman. *The Work of Music and the Problem of Its Identity*. Translated by Adam Czerniawski. Berkeley: University of California Press, 1986.

Kivy, Peter. *The Fine Art of Repetition*. Cambridge, U.K.: Cambridge University Press, 1993. Chapters 2–4.

Levinson, Jerrold. *Music, Art, and Metaphysics*. Ithaca, NY: Cornell University Press, 1990. Chapters 4 and 10.

Wollheim, Richard. *Art and Its Objects: An Introduction to Aesthetics*. New York: Harper and Row, 1968. Sections 1–39.

Wolterstorff, Nicholas. *Works and Worlds of Art*. Oxford: Clarendon Press, 1980. Part 2.

THE PERFORMANCE

Davies, Stephen. *Musical Works and Performances: A Philosophical Exploration*. Oxford: Clarendon Press, 2001.

Dipert, Randall R. "The Composer's Intentions: An Evaluation of Their Relevance for Performance." *Musical Quarterly* 66 (1980).

Godlovitch, Stan. *Musical Performance: A Philosophical Study*. London: Routledge, 1998.

Kivy, Peter. *Authenticities: Philosophical Reflections on Musical Performance*. Ithaca: Cornell University Press, 1995.

THE REWARDS OF LISTENING

Budd, Malcolm. *Values of Art: Pictures, Poetry, and Music*. London: Penguin Books, 1995.

Kivy, Peter. *Philosophies of Arts: An Essay in Differences*. Cambridge, U.K.: Cambridge University Press, 1997. Chapter 7.

Peter Kivy (2005)

MUSLIM PHILOSOPHY

See *Islamic Philosophy*

MUSONIUS RUFUS

(30–100 CE?)

Musonius Rufus belongs to a group of Roman Stoic thinkers that also includes Seneca and Marcus Aurelius. He was Epictetus's teacher. Only fragmentary accounts of his views, recorded by others, have survived (English translation in the edition by Cora Lutz).

Like other Stoics, Musonius rejects the distinction between theoretical and practical wisdom: philosophy is nothing else but to practice and put in good deeds what Stoic doctrine prescribes. All human beings have the potential to strive towards virtue. This view is anchored in a radically embedded concept of human nature: a human is a composite of soul and body and a member of the universe's community of gods and men, the so-called cosmopolis. Musonius reinforces this ontological embeddedness by emphasizing social responsibility in general, in existing communities of human beings.

Musonius is perhaps best known for his positive views on women (fragments 3 and 4): Both men and women have the same intellectual and moral capacities, and hence women should be educated in philosophy just as men are. But it is equally important that this stance has a social corollary in Musonius's highly positive assessment of marriage as a symmetrical and fully reciprocal relationship among equals that entails a union of soul as well as of body (fragments 12, 13 A and B, 14). Thus

Musonius represents a Stoicism that upgrades traditional relationships such as marriage to the level of philosophically inspired friendship between men.

The importance of social responsibility is also evident in Musonius's views on suicide. As fragment 29 states, "One who by living is of use to many has not the right to choose to die unless by dying he may be of use to more" (tr. Lutz). Hence the concern for others ought to be central in one's decision-making process.

Other themes in the preserved fragments reflect on the need for a king to be a philosopher, on the duties of parenthood, on curtailing one's bodily and material wants, and on patience with and forgiveness of people who have wronged one. Rudolf Hirzel (1895, 2: 239) dubbed Musonius "the Roman Socrates."

See also Epictetus; Stoicism.

Bibliography

Geytenbeek, A. C. van. *Musonius Rufus and Greek Diatribe*. Translated by B. L. Hijmans Jr. Assen, Netherlands: van Gorcum, 1963.

Hirzel, Rudolf. *Der Dialog, ein literarhistorischer Versuch.* 2 vols. Leipzig: S. Hirzel, 1895.

Lutz, Cora Elizabeth. *Musonius Rufus, "The Roman Socrates."* New Haven, CT: Yale University Press, 1947.

Nussbaum, Martha C. "The Incomplete Feminism of Musonius Rufus, Platonist, Stoic, and Roman." In *The Sleep of Reason. Erotic Experience and Sexual Ethics in Ancient Greece and Rome*, edited by Martha C. Nussbaum and Juha Sihvola. Chicago: University of Chicago Press, 2002. 283–326.

Reydams-Schils, Gretchen J. *The Roman Stoics: Self. Responsibility, and Affection.* Chicago: University of Chicago Press, 2005.

Gretchen J. Reydams-Schils (2005)

MYSTICISM, HISTORY OF

Mystical experience is a major form of religious experience, but it is hard to delineate by a simple definition for two main reasons. First, mystics often describe their experiences partly in terms of doctrines presupposed to be true, and there is no one set of doctrines invariably associated with mysticism. Some of the definitions of mysticism advanced by Western writers are quoted by W. R. Inge in his *Mysticism in Religion* (p. 25): "Mysticism is the immediate feeling of the unity of the self with God" (Otto Pfleiderer); "Mysticism is that attitude of mind in which all relations are swallowed up in the relation of the soul to God" (Edward Caird); "True mysticism is the consciousness that everything that we experience is an element and only an element in fact, i.e. that in being what it is, it is symbolic of something else" (Richard Nettleship). Quite clearly, such definitions import a religious and philosophical interpretation to the phenomenon of mysticism that would not be shared by all contemplatives. For instance, the Buddhist mystic, not believing in a personal God, would reject the first two of these definitions; and he might well be skeptical about the third—in what sense is the experience of nirvāṇa symbolic of something else?

Second, there is quite a difference between mystical experience and prophetic and, more generally, numinous experience, but it is not easy to bring out this phenomenological fact in a short definition. (A numinous experience is an experience of a dynamic external presence—described classically in Rudolf Otto's *The Idea of the Holy* as that of a *mysterium tremendum et fascinans,* an awe-inspiring and fascinating mystery.) Sidney Spencer says, for instance, "What is characteristic of the mystics is the claim which they make to an immediate contact with the Transcendent" (*Mysticism in World Religion,* p. 9). Such a definition includes under mysticism the experiences of the Old Testament prophets, those of Muḥammad, and the theophany described in the *Bhagavad-Gītā.* However, these differ so markedly from the interior illumination of such figures as Meister Eckhart, Teresa of Ávila, Śankara, and the Buddha that it is misleading to bracket the two kinds of experience. This article will explicitly exclude the prophetic and numinous experience, save where it becomes relevant to the experiences and doctrines of those properly called mystics. It is thus best to indicate what is meant by "mysticism" by referring to examples, such as Eckhart and the others cited above, and by sketching some of the important features of the type of experience in question without interpreting it doctrinally.

Generally, mystics as typified by Eckhart, Teresa of Ávila, Śankara, and the Buddha feel that their experience is somehow timeless, that it involves an apprehension of the transcendent (of some thing, state, or person lying beyond the realm of things), that it gives them bliss or serenity, and that it normally accrues upon a course of self-mastery and contemplation. These are certainly features of what has been called introvertive mysticism by W. T. Stace (*Mysticism and Philosophy,* p. 60). There are other experiences, however; those of extrovertive mysticism, where, according to R. C. Zaehner, one gains a kind of rapport with the world, or "panenhenic" feeling (*Mysticism Sacred and Profane,* Ch. 1). These neither coincide with prophetic experiences nor strictly with those of

introvertive mysticism, but since they sometimes occur in conjunction with the latter, it is convenient to treat them as mystical. Various abnormal mental states, such as those induced by mescaline, lysergic acid, and alcohol are sometimes considered mystical, but they are far enough removed from mainstream mysticism for it to be reasonable to neglect them here.

In the light of all this, we can distinguish various aspects of mysticism: The experiences themselves, the paths or systems of contemplative techniques often associated with them, and the doctrines that arise from mysticism or are affected by it. Also, such paranormal phenomena as levitation are sometimes ascribed to mystics, although they usually regard these as of secondary significance.

There is no single history of mysticism because some of the major religious traditions have been largely independent of one another. Further, there is no way of knowing the real origins of mysticism, since for such an intimate type of experience we must rely chiefly on written records and thus have no access to prehistoric mysticism. Studies of contemporary nonliterate cultures—in Africa, for instance—do not reveal the presence of much or any mysticism proper; for example, the religious experiences of the Nuer in the Sudan are more akin to those of Old Testament prophecy. It is thus convenient to confine attention to the main literate religious traditions: Indian religions (Hinduism, Buddhism, Jainism, Sikhism); Chinese and Japanese religions; and the Semitic faiths (Judaism, Christianity, and Islam). It may be noted that early Christian mysticism was influenced by Greek, notably Platonist, ideas.

THE INDIAN TRADITION

The mainstream of Indian mysticism centers on the practice of yoga, which in its general sense involves techniques of pacifying the mind and of attaining interior insight. Evidence from the pre-Aryan Indus Valley civilization indicates that it may have been practiced in the second millennium BCE or earlier. By contrast, the religion of the Aryans who settled in north India centered on sacrifice set in a polytheistic framework. As this ritual religion became more complex, questions arose concerning the inner meaning of the sacrificial rites. The *Upaniṣads* (the chief of which date from about 800 BCE to about 500 BCE) were in part concerned with extending and deepening sacrificial ideas in the quest for *vidyā,* or knowledge of sacred reality. Quasi-magical ideas surrounded this notion—for instance, that knowledge gives power over the thing known and that one can become identified with the thing known. At the same time, mystical ideas began to permeate religious thinking, notably the idea that through austerity and self-control one could attain a realization of one's eternal self. A confluence of these streams of religious thought resulted in the famous central identification expressed in the *Upaniṣads,* "That art thou"; the sacred reality embracing and sustaining the cosmos ("That") and the eternal self ("thou") are one. In brief, inner mystical knowledge brings a union with the Divine.

This union is described in various ways: "Just as a man embraced by his dear wife knows nothing at all, outside or inside, so does the eternal life-monad [*puruṣa*], embraced by the supreme spiritual Self, know nothing at all, outside or inside" (*Bṛhadāraṇyaka Upaniṣad,* IV. 3.21); "As rivers flow to their rest in the ocean and there leave behind them name and form, so the knower, liberated from name and form, reaches that divine Person beyond the beyond" (*Chāndogya Upaniṣad,* 6). Sometimes the lack of duality between the divine Being and the soul is stressed: "Where there is a duality, as it were, one sees another, tastes another, speaks to another.… But when everything has become one's own self then whom and how would one see? … The Self is not this, not that" (*Bṛhadāraṇyaka Upaniṣad,* IV. 5.15). Mystical consciousness is also said to be like a state beyond dreamless sleep. These passages hint at what is virtually universal throughout Indian yoga, the fact that the contemplative state in its highest form involves going beyond ordinary perceptions, mental images, and thoughts. It is thus not describable by the ordinary expressions for mental states. It is no doubt partly for this reason that the distinction between perceiver and perceived is not regarded as applicable, and so the contemplative who conceives himself as "seeing" *Brahman* (the divine Being) thinks of this as a kind of union with *Brahman.* By contrast, in atheistic systems of Indian religion, where there is nothing for the self to be identified *with,* the contemplative state is conceived in a rather different way.

Although identification between the self and *Brahman* is a central theme in Upaniṣadic religion, some of the writings, notably the *Kaṭha* and the *Śvetāśvatara Upaniṣads,* are more theistic in spirit and less inclined to speak in terms of identification. These differences of emphasis are partly the reason for the divergences in interpretation found in different types of Vedānta in the medieval period.

JAINISM AND YOGA. Jainism, Buddhism, and the tradition later formulated as classical yoga involved an atheistic or agnostic interpretation of mystical experience.

Jainism and classical yoga (the long-extinct Ājīvika school) were monadistic: They believed in an infinity of eternal life monads or souls, and the aim of the ascetic was to bring about the isolation of the soul from its material environment. Such an isolation would involve the cessation of reincarnation and thus final deliverance from suffering. Jainism, because it held that karma, the force determining people's situations as a result of their previous deeds, is a subtle form of matter, considered extreme *tapas* (austerity), which had the effect of annihilating this material force, the central means of liberation. Nevertheless, it seems that the Jain teacher Vardhamāna (known also as Mahāvīra), a contemporary of the Buddha, and his disciples claimed to attain a certain kind of higher state analogous to the experience of nirvāṇa in Buddhism. Thus, in Jain doctrine the life monad in its emancipated state gains omniscience, a concept reflecting the intense sense of insight accruing upon the contemplative experience.

BUDDHISM. The accounts of the Buddha's enlightenment—a crucial event in the history of Indian religion and likewise centrally important in the history of Indian mysticism—are elaborate and circumstantial. During the first night, the Buddha, seated under the bo tree, remembered the series of his former births; during the second, he acquired the "heavenly eye," which enabled him to view the entire world and the whole cyclical process of rebirth; during the third, he saw how the latter depended upon grasping and ignorance—if living beings were liberated from these, they would escape rebirth; and in the fourth, he attained supreme insight after going through the various stages of meditation (Sanskrit, *dhyāna*; Pāli, *jhāna*). In all this he gained supreme peace. No doubt the scriptural records are a formalized account, hardly based on the Buddha's autobiographical report, but they certainly point to the type of inner experience early Buddhism prized. Something can be learned from the *Theragāthā* and *Therīgāthā,* verses composed by monks and nuns and expressing the flavor of early Buddhist contemplative experience. These poems often show the sensitivity of the recluse to the beauty of nature:

> The peacocks shriek. Ah, the lovely crests
> and tails
> And the sweet sound of the blue-throated
> peacocks.
> The great grassy plain with water now
> Beneath the thunder-clouded sky.
> Your body is fresh; you are vigorous now and fit
> To test the teaching. Reach now for that
> saintly rapture,
> So bright, so pure, so hard to fathom,
> The highest, the eternal place.
>
> (*THERAGĀTHĀ* CLXVI)

The eternal place is, of course, nirvāṇa.

The achievement of inner peace and insight, as opposed to the use of complex psychological categories in explaining human nature, was given comparatively little doctrinal elaboration in early Buddhism because the Buddha apparently felt that the concepts of the transcendent state (nirvāṇa) and the cessation of rebirth through the perception or attainment of nirvāṇa were sufficient means of interpreting mystical experience. Certainly, he did not give the more elaborate type of interpretation found in the *Upaniṣads* and in theistic mysticism. It is clear, however, that the experience or experiences involved both the attainment of a marvelous serenity and a kind of knowledge or insight (something regarded as knowledge, given the presuppositions of the Buddhist mystical quest). Grasping and ignorance are dispelled by this peace and knowledge.

Buddhism rejected the doctrine of a plurality of eternal souls, but in a sense it can be seen as a transcendence of monadism, with the concept of the eternal soul replaced by that of the capacity to attain release. Thus early Indian mysticism is typically monadistic, except in the *Upaniṣads,* where the interior experience is related to the *Brahman* and where, therefore, the *Brahman-ātman* (self) equation is formulated. Only because the eternal self of the mystic is identified with the presupposedly single divine Being is the plurality of souls denied. The numinous religion of Brahmanism overlays that of the contemplative mysticism of yoga, and the mystical experience is interpreted in terms of union with the unitary divine Principle.

Mahāyāna Buddhism, from the first century BCE on, moved toward a more elaborate interpretation of the contemplative path. Nirvāṇa was identified with the Absolute, variously named Suchness (*tathatā*) and the void (*śūnya*). These terms served to bring out the ineffability and undifferentiated nature of ultimate reality, which in turn corresponded to the undifferentiated and "void" nature of the contemplative experience itself. The Absolute was also identified, from the standpoint of the ordinary worshipers, with the Truth Body of the Buddhas—the transcendent and essential aspect of buddhahood—and thus the mystical path involved being a bodhisattva (buddha-to-be). The distinctionless, nondual experience of ultimate reality, the goal of the path, was the achievement of identity with the Absolute, which

was equated with buddhahood. This is why the Mahāyāna path of contemplation was thought of as the path of bodhisattvahood, so that on his enlightenment the mystic would himself become a buddha.

As a preliminary, the aspirant practices individual worship (*pūjā*) of the celestial buddhas and bodhisattvas and can gain assurance from a living buddha that his aspiration to buddhahood will be fulfilled. He practices the perfections of the path, culminating in supreme wisdom or insight (*prajñā*).

There are three chief differences between Mahāyāna and Hīnayāna, now represented by the Theravāda (in Myanmar, Sri Lanka, and parts of Southeast Asia). First, the Mahāyāna stresses self-giving more strongly, so that the aspirant continually looks to the welfare of others; second, it is a path accessible to laymen as well as to monks; third, contemplation is supplemented by the use of sacramental and ritual practices, at least in certain phases of the Mahāyāna. Some of these practices, known as tantra, became well developed in the middle of the first millennium CE in both Hinduism and Buddhism and deeply affected the Buddhism of Tibet. It sometimes involved the ritual breaking of taboos (against meat-eating and against sexual intercourse outside marriage): Such a breaking of taboos was regarded as a means of testing and developing detachment. Coordinate with this type of Buddhism was a highly ritualistic use of sacred texts and recitations. The most outstanding figure of Tibetan mysticism was the poet and yogi Milarepa (1040–1123).

HINDUISM. The theistic religion implicit in some of the *Upaniṣads,* reinforced by popular cults and by an emphasis on *bhakti,* or loving adoration of God, led to a different valuation of mysticism in the *Bhagavad-Gītā.* The poem speaks of three paths to salvation: the way of knowledge (primarily contemplative knowledge), the way of works, and the way of devotion (*bhakti*). The three paths are stressed in different parts of the *Gītā,* but two significant lessons emerge. First, the pursuit of works (religious and moral duties) need not bind one to the world if they are performed in a spirit of self-surrender to God; the way of works should be seen in the light of the way of devotion. Second, the yogi who pursues knowledge (*jnana*) can become *Brahman* (VI.27). Elsewhere, however, *Brahman* is spoken of as part of God; the personal aspect of God is more important than his impersonal aspect. Thus the yogi, in pursuing a strictly contemplative path, can only unite himself with the lower, rather than the more important, aspect of the Lord's nature. This doctrine represented a higher evaluation of *bhakti* than of contemplative yoga. (It must be pointed out that traditional Indian commentators are divided on the question of what is the correct interpretation of the *Gītā.* However, there is little doubt that extraneous theological and philosophical presuppositions have played a large part in determining interpretations.)

The continued growth of devotional or *bhakti* religion led to a similar interpretation of mysticism during the medieval period. Thus, in the twelfth century Rāmānuja reversed the doctrinal priorities of Śankara (ninth century). Śankara's monism represented the most radical interpretation of the Upaniṣadic identity texts, asserting a numerical identity between the soul and the divine Being. While for Śankara the personal Lord was a lower manifestation of the Absolute, so that worship and devotion could be transcended when one had attained the apprehension of identity with Brahman, Rāmānuja, although recognizing identity as one religious goal, conceived it as an inferior form of release. The higher form was the vision of the personal God, in which the soul was in a state of loving dependence on the Lord. Both Madhva (thirteenth century) and the theistic Śaivite schools of Indian philosophy interpreted mystical experience in terms of union with God, but not a union involving the numerical identity of the soul and God. Thus, mystical experience was interpreted by reference to the duality of the soul and God implicit in the religion of *bhakti*: The worshiper has a strong sense of the majesty and glory of God, and thus of the difference between himself and the object of worship. Various analogies were used, including that of the marriage of the soul and God, since sexual love symbolizes the intimate union between the lover and the beloved while presupposing the difference between the two. This analogy tied in with the cult of Krishna: The legend of Krishna's amorous dalliance with the milkmaids was seen as an allegory of the relation between God and men's souls.

The interiorization of religion involved in both devotionalism and contemplation influenced Nānak (1469–1538), founder of the Sikh religion, who preached doctrines combining the anti-idolatrous monotheism of Islam and such characteristic Hindu ideas as reincarnation and karma.

There have been a number of outstanding contemplatives in modern Hinduism. Chief among them was Ramakrishna Paramahamsa (1834–1886), whose disciple Vivekananda (1862–1902) did much to popularize his teachings in both the East and the West; Vivekananda's organizing ability was chiefly responsible for the flourish-

ing state of the Ramakrishna movement, in which the contemplative life is geared to social service and also provides a pattern of living that can, according to the teachings of the movement, transcend the differences between the great living faiths. A twentieth-century mystic who tried to adapt traditional teachings to modern thought was Aurobindo. Contemplation and yoga, through the activities of numerous recluses, holy men, and gurus, continue to play a prominent role in Indian religion.

CHINA AND JAPAN

Chinese mysticism has two main sources, Daoism and Buddhism. A product of their interaction was Ch'ān, better known under its Japanese name, Zen. The teachings of Confucius were not much concerned with the contemplative quest for inner illumination, although certain mystical ideas were expressed in the *Book of Mencius* of the Confucian tradition. On the whole, however, early Confucianism was indifferent to the contemplative ideal.

DAOISM. The chief early mystical writing in China was the *Dao-de-jing,* traditionally ascribed to Laozi, who is thought to have been an older contemporary of Confucius. It is likely, however, not only that the book was later but also that it was the work of several men. The anthology expresses a roughly consistent viewpoint, one that, on the most natural account of it, has its roots in contemplation (although some commentators give it a nonmystical interpretation).

The Way, or Dao, referred to in the *Dao-de-jing* is both a principle underlying natural processes and a mode of life whereby the sage can gain identity or harmony with nature. Since nature acts spontaneously and effortlessly, the book claims that the sage likewise can be effective through inaction (*wu-wei*) and effortlessness. Thus, the pattern of life suggested is one of withdrawal and passivity. In these themes the *Dao-de-jing* reflects some of those found elsewhere in mystical literature: The sense of identification with the Principle *(li)* underlying the world and the need for an unworldly mode of existence. Because the attainment of harmony with Dao was seen as living in accord with nature, the Daoists reacted against what they considered the artificialities of social life and etiquette as practiced by the Confucians, and from the doctrine of *wu-wei* they derived political views not far from anarchism.

In practice the effortlessness of the Daoist contemplative was modified by the use of techniques of meditation, such as controlled breathing, analogous to those employed in Indian yoga. The Daoist aim of an immediate, intuitive, inner illumination was sufficiently close to the aim of Buddhist meditation for it to be natural that the two streams of religion should influence each other in the period after Buddhism's arrival in China, in the first century CE. In particular, it was during the sixth and following centuries that this interplay was most marked.

NEO-CONFUCIANISM. The success of Buddhism, which in part resulted, at least among intellectuals, from the subtlety of its metaphysical doctrines, was a factor in stimulating the so-called neo-Confucian revival, in which a metaphysics was elaborated to underpin the Confucian ethic.

One main phase of this revival was the growth of philosophical idealism, which owed something to mystical ideas. Thus, Lu Xiangshan (1139–1193) argued that there is a single underlying principle, *li,* that explains all things and is spiritual. Thus, he claimed, his mind and the universe were one. It followed that one can discover the truth by introspection.

Such an idealism was further developed by Wang Yangming (1472–1529), about whom a significant story is told. He and a friend were concerned about the method by which one should purify the mind, for Zhu Xi (1130–1200) had said that one should investigate the nature of things. Wang and his friend decided to contemplate a bamboo in the front courtyard but gave up after several days. It is notable that this attempt corresponds to one of the preliminary methods of Buddhist contemplation. Although unconvinced by such "external" contemplation, Wang nevertheless considered the interior quest—the purification of consciousness—important. He believed that through looking inward at one's own nature one could gain an intuitive knowledge of the whole of reality. It is said that while in banishment and living under poor and menial conditions, Wang had a mystical experience in which he realized this doctrine existentially. However, Wang was far from abandoning the traditional Confucian emphasis on ethical behavior; he did not advocate quietism and passivity but saw in mysticism a way of enhancing moral goodness. Inner illumination would shine through in active concern for others. However, in such neo-Confucianism the influence of Ch'ān Buddhism can be detected.

BUDDHISM. Ch'ān, or Zen, Buddhism embodies the most distinctive feature of both Chinese and Japanese mysticism, since it incorporated Daoist ideas into Buddhist mysticism. Other schools of Far Eastern Buddhism in varying ways carried on and developed the Buddhist

tradition and therefore incorporated Buddhist contemplative ideals. A powerful aspect of Far Eastern Buddhism was the success of the Pure Land school, which centered its teachings on the faith and devotion whereby the ordinary person could receive supernatural aid from the Buddha Amitābha and gain rebirth in the paradise of the Pure Land. With its stress on devotion and the efficacy of the Buddha's grace, this school tended to bypass contemplative mysticism and to focus religion upon worship.

JUDAISM

Although the Hebrew Bible contains virtually no expression of contemplative religion, mysticism developed within Judaism by the first century BCE. It centered mainly on the imagery of the *merkabah* (chariot), described in Ezekiel as a complex vision of the manifestation of divine power in the shape of supernatural beings riding on a mysterious four-wheeled chariot (Ezekiel 1). The Talmud indicates that some of the early rabbis practiced asceticism and self-purification as a preparation for a mystical "ascent into heaven." Philo Judaeus (fl. 20 BCE–40 CE) mentioned a community of Therapeutae near Alexandria who practiced a form of contemplative monasticism, and likewise mysticism may have been part of the Essene way of life. Philo himself was the greatest figure in these early phases of contemplative Judaism, although he was so deeply affected by Greek ideas that he is outside the mainstream of Jewish thought and piety. According to Philo, man, through his intellect, has an affinity with God; and through the contemplative life he can in principle attain a state where he can see God's essence. In accordance with Platonist and mystical ideas, Philo expounded a negative theology: God eludes the affirmations we try to make about him. Consequently, Philo's interpretation of Scripture was not at all literalistic, and he made lavish use of the allegorical method. He attempted, moreover, to show that the experiences of the prophets were mystical.

The most important period of Jewish mysticism was the Middle Ages. Beginning in the twelfth century there developed Hasidism, which made a lasting imprint on central European Judaism, and Kabbalism, mainly in Spain and southern France. The former takes its name from the term *Hasidim* ("devout ones"), a name originally applied to a movement of the second century BCE that was a forerunner of Pharisaism. Medieval Hasidism concentrated on the cultivation of the sense of divine presence. Modern Hasidism, dating from the eighteenth century, is more directly contemplative and is indebted to Kabbalism.

KABBALISM. Kabbalism centered on the esoteric teachings known as the Kabbalah, which found their chief expression in the *Zohar* ("splendor"), a work traditionally ascribed to the second century but actually dating from the thirteenth century or a little earlier, that conceives of God as the En-Sof, the "Endless" or "Infinite." In itself the En-Sof is qualityless, but there are ten ideal qualities, known as the *Sefirot,* that emanate from the Infinite—wisdom and power, for instance. These are used to explain the creation of the world. The cosmos that man inhabits, however, is the lowest sphere in which the *Sefirot* operate—a doctrine that expresses the way in which the perfect Infinite is far removed from the imperfect world we inhabit. The hierarchy of stages between God and the material world is reminiscent of Gnosticism. Nevertheless, the En-Sof, being infinite, does in some sense embrace lower forms of existence; and every entity in the universe reflects and interpenetrates everything else.

How is all this related to traditional Jewish teachings? According to the Kabbalah, the doctrine of interpenetration implies that lower events will stimulate corresponding activity from on high. The fall of Adam brought about a rupture in the cosmos; the Shekinah, or Divine Presence, became exiled from the En-Sof. No longer does the Presence pervade the whole world; it appears intermittently here and there—for instance, in ancient Israel—and has continued to be especially associated with the Jewish people. The aim of the pious should be to bring about a reunion of the En-Sof with the Shekinah. Since the human soul contains some of the *Sefirot,* the individual experience of such a reunion will have its cosmic effects and help to restore universal harmony. Consequently, the mystical life was given a dramatic and central place in the operations of the universe.

It will be apparent that some of these ideas, such as the ineffability of the En-Sof and the rather impersonal description of God, echo similar notions in Neoplatonism and other forms of mystical theology. Despite the unorthodoxy of much of their speculation, the Kabbalists continued the detailed observance of Jewish law, ascribing to it a mystical significance.

Isaac Luria. An important figure in the development of Kabbalism was Isaac Luria (1534–1572), of a Spanish Jewish family living in Palestine. He believed in reincarnation, which would give men ever fresh chances of living the pure life and would provide a framework for the punishment of those who had transgressed. Luria conceived of Adam as a universal being who before the Fall embraced the universe, then in an ideal state. With his fall, the material world was created, and the light of his

divine nature was fragmented into the sparks that illuminate the myriads of living souls. In the final consummation, all will be reunited. Asceticism and the practice of *kavannah*—concentrated devotion in all one's acts—were the means of purifying the soul. Social conditions may have helped the growth of such doctrines, for the emphasis on meekness, love, and a quiet interior life were well adapted to the unhappy outer circumstances of the Jewish people, and the Kabbalistic reinterpretation of the Messianic hope gave the contemplative a cosmic role.

MODERN HASIDISM. The founder of modern Hasidism was Israel Baal Shem-Tov (c. 1700–1760), who lived in Carpathia in eastern Europe. He gathered round him disciples who were devoted to the mystical life. His successor, Baer of Meseritz (1710–1772), was an energetic organizer and missionary who spread the movement among Jews throughout eastern Europe and the Ukraine. Stress was laid on the concept of the zaddik, or perfectly righteous man, through whom the favor of God is channeled. Only he can attain union with the divine Being; less perfect folk must find their spiritual development through his guidance. This doctrine is reminiscent of Hindu ideas about the guru as conveyor of illumination. In any event, Hasidism implied that the zaddik, rather than the rabbi or learned person, was the immediate source of authority. This gave Hasidic mysticism a popular following and organization, and the essential simplicity of its message—that salvation can be attained through prayer and pious acts—made it adaptable to the experience of people of no great sophistication or learning.

As elsewhere in the history of mysticism, antinomian tendencies made their appearance. Thus Sabbatianism, named after Sabbatai Zevi (1626–1676), a self-styled Messiah who preached apostasy from Judaism, made use of Kabbalistic ideas in order to justify the concept of the God-man who is "beyond good and evil," as in the teachings of Jacob Frank (c. 1726–1791).

Although the Hasidim often attacked official rabbinical teaching, the revival of Jewish learning in the nineteenth century paved the way for a reconciliation between orthodoxy and Hasidic piety, so that the latter still remains a force within the fabric of Jewish religion.

CHRISTIANITY

ORIGINS. As has been mentioned, there was little mysticism in the traditions of Judaism until the time of Christ, and there also seems to have been little in the experience of the earliest church. It is true that Paul underwent a powerful experience of being "caught up to the third heaven," which could have had a mystical character, although it is also reminiscent of certain prophetic experiences, such as those of Muḥammad. The origins of Christian mysticism can more plausibly be sought elsewhere, in the rise of monasticism and the influence of Neoplatonism. Some stimulus to such a development may also have been given by the existence of Gnostic sects both within and outside Christianity, from the end of the first century CE.

Gnosticism. Gnosticism—a term derived from the word *gnosis,* meaning knowledge, particularly the immediate inner knowledge of the divine Being—tended to be ascetic and esoteric. Its asceticism was expressed by the doctrine that matter is evil, so that liberation of the soul is achieved through withdrawal from the world. Because of the evil nature of the world, Gnostics frequently postulated a hierarchy of beings below God and concerned with the creation of the world. Thus God himself was not contaminated, so to speak, by direct contract with matter. Such a doctrine was heretical, for it did not square with the Christian doctrine of creation or with Christian attitudes to the world, but it was one factor in stimulating an orthodox asceticism and mysticism within Christianity.

Monasticism. Monasticism grew out of eremitic practices, mainly in Egypt. Famous among early hermits was Anthony the Great, whose asceticism became almost legendary. Early in the fourth century monasticism proper was established in Egypt, the key figure being Pachomius. Thereafter the movement spread rapidly in Egypt and the Eastern church. It was further organized by Basil the Great (c. 330–379), whose rule formed the basis of Orthodox monasticism. John Cassian (c. 360–c. 434) brought Egyptian-style monasticism to the West, founding two monasteries in the south of France. His rule underlay that of St. Benedict, who lived in the following century. The connection of monasticism with mysticism was a straightforward one, for a main rationale of monasticism was the cultivation of the spiritual life, whereby a foretaste of the beatitude of the blessed in heaven could be gained. Thus the ultimate destiny of man was seen in contemplative terms, and it was thought possible to anticipate this destiny by a regulated life withdrawn from the world.

Neoplatonism. Neoplatonism, which expressed a view of the world in part stemming from, and in part providing a rationale for, mystical experience, made a lasting imprint upon Christian contemplation. A sign of this was the composition of the Pseudo-Dionysian writings, which were ascribed to Dionysius the Areopagite, a convert of St. Paul, but really date from approximately the

beginning of the sixth century. These writings had a wide impact upon medieval mysticism. The negative theology expounded in them was not merely the result of logical difficulties involved in the ascription of ordinary predicates to God but, more importantly, was geared to the expression of the contemplative's inner experience of a "darkness clearer than light." Thus the mystical experience, being different from, and not expressible in terms of, perceptual and related forms of experience, seemed to imply that its object was likewise indescribable and therefore better conveyed by negations than by positive affirmations.

Neoplatonism also, of course, deeply influenced St. Augustine, and he has been a principal source of the notion, enshrined in monastic practice, that introvertive contemplation can give a foretaste of the heavenly life. Thus the highest state of Christian blessedness was increasingly identified with contemplation, and mysticism became the pattern after which eternal life was conceived.

EASTERN ORTHODOX MYSTICISM. The Pseudo-Dionysian writings also formed an important part of the fabric of Eastern Orthodox mysticism, for there were also features of the general theology of Orthodoxy that favored the contemplative ideal. John of Damascus, who in the eighth century summed up the work of the Cappadocian Fathers (fourth century), expressed in his writings a doctrine of deification that was both typical of and formative of Eastern Orthodox theology. Man was considered the connecting link between the visible and invisible worlds. He was created perfect but through the Fall lost his immortal, incorruptible, and passionless nature. A certain scope for free will remained, however. The image of God, although defaced, was not entirely lost. The restoration of man to the true end for which he was made—the contemplation of God—was effected through Christ's incarnation. Christ, by uniting the Godhead to human nature, restored that nature to its perfection; and by sharing in his perfect humanity, men also can be raised up and deified. In terms of Dionysian mysticism, this deification takes place through the illumination of the soul; its divinization, through the divine Light. Virtually throughout Eastern mysticism this imagery of light was to play a central part, and thus St. Simeon (949–1022), perhaps the most important of Eastern Orthodox mystics, identified the inner light with the glory emanating from God.

Hesychasm. Simeon was also a forerunner of the significant contemplative movement known as Hesychasm (from the Greek word *hesychos*, "quiet"), whose methods of training had some analogy to those found in Indian yoga.

The Hesychasts (eleventh–fourteenth centuries) held that their methods were conducive to the inner vision of the uncreated Light, identified with that which suffused Christ at the Transfiguration on Mount Tabor. This Light was conceived as emanating from God and was not to be identified with his essence, which is unknowable (this was a means of retaining orthodox teaching, by safeguarding mysticism from a full doctrine of union with, or knowledge of, God). Among the training methods used were breathing exercises and the continued repetition of the Jesus Prayer—"O Lord Jesus, Son of God, have mercy on me, a sinner." In a mysterious manner, the very repetition of the sacred name of Jesus was supposed to contain the divine power.

Gregorius Palamas (c. 1296–1359), the most noted and controversial exponent of Hesychasm, considered the Jesus Prayer as the central act of piety; and although the use of breathing techniques, which persisted until the eighteenth century, has been discontinued, the Jesus Prayer has survived as a characteristic part of Orthodox religion. Palamas and the Hesychasts were not, however, unopposed. Some opponents thought that the doctrine of the uncreated Light made a division within the Godhead—Palamas had even spoken of "divinities." Thus the attempt to soften the idea of mystical union by regarding it as identification not with the divine essence but with the divine illuminative energy, was criticized on the ground that it transferred the difficulty to another locus by introducing something like polytheism. Nevertheless, Hesychastic teaching came to be recognized officially, and the movement was the mainspring of medieval Orthodox contemplation.

ROMAN CATHOLICISM. The mystical life served to counterbalance the worldly tendencies that had permeated the early medieval church in the West. Pope Gregory the Great (c. 540–604) discovered in his own experience something that could be expressed in terms of the irradiation of the divine Light; and Gregory VII, elected pope in 1073, undertook extensive ecclesiastical and monastic reforms that were partly inspired by the intense cultivation of the personal and contemplative life he had discovered in the Cluniac movement—a monasticism whose rules and ideals emanated from the monastic center at Cluny in Burgundy.

The most important figure in monastic reform was Bernard of Clairvaux (1090–1153). Although he was

influenced by Augustine, his concerns were not primarily expressed in metaphysical language. He believed that in the mystical experience the soul is emptied and wholly lost in God, but he did not conceive this as an actual union with the Godhead. The soul and God remain distinct in substance, although they are joined by the "glue of love." Through man's love flowing up to God and through the downward movement of God's grace, the two become united. Bernard combined this intense mysticism with great powers of leadership and played a large part in the forward movement of the Cistercian order.

Other important mystics were Hugh and Richard of St. Victor, an Augustinian abbey in Paris in the twelfth century, and St. Bonaventure (c. 1217–1274) in the following century. St. Bonaventure evolved a theory of mysticism that set forth the three ways of the spiritual life: purgative, illuminative, and unitive. In the first stage, the individual purifies himself through meditation; in the second, he is illuminated by the divine mercy; in the third, he gains a continuing union with God through love. This love is nourished by concentrating upon God, to the exclusion of mutable things. Thus, Bonaventure's path typically followed that of introversion, while his theological doctrines leaned upon Augustine and Pseudo-Dionysius.

There were ways, however, in which mystical teachings, especially where they strongly emphasized the negative theology of Pseudo-Dionysius, could seem unorthodox. The work of Thomas Aquinas (1224?–1274), in excogitating a novel synthesis between Christian theology and Aristotelianism, accentuated differences of emphasis between some of the mystics and orthodox doctrine. Thus Meister Eckhart (c. 1260–1327/1328), a Dominican and therefore versed in Thomism, fell under condemnation.

The greatest of the German contemplatives, Eckhart spoke in ways that suggested not merely that there is an ontological distinction between the Godhead, which is beyond description, and the Trinity of describable Persons but also that it is possible for the contemplative to go "beyond God" in achieving identity with the Godhead. Despite his unorthodox language, Eckhart inspired a strong following, and the mysticism of Johannes Tauler (c. 1300–1361), Heinrich Suso (1295/1300–1366), Jan van Ruysbroeck (1293–1381), and the partly lay group known as the Friends of God in Germany, the Low Countries, and Switzerland owed much to him.

It was out of the Friends of God that the anonymous but famous mystical treatise, the *Theologia Germanica,* originated, stressing the abandonment of the soul to God. The corruption of the church and the disillusioning events of the Great Western Schism were motives for the Friends of God to attempt to revitalize faith through the inner life, and this sometimes involved a highly critical attitude toward ecclesiastical authority. It is worth noting, however, that the rather sudden flowering of mysticism in Germany during the fourteenth century owed much to the fact that in 1267 the Dominican friars had been charged by Pope Clement IV with the spiritual direction of the nuns in the numerous convents in the Rhineland. Hitherto they had frequently been without proper religious supervision.

Mysticism could lead in directions that seemed to be the reverse of Christian piety. The sect known as the Brethren of the Free Spirit, which dated from the early thirteenth century, believed that men are of the same substance as God: Every man is capable of becoming divine. It followed that when this divinization was achieved, a person could no longer sin, for God is sinless. Thus, whatever one did, it would not be a sin. Commandments and conventional tests of morality could no longer apply, and mysticism was therefore interpreted as justifying antinomianism. (Thus, it was not surprising that some of Eckhart's language, although not intended in this sense, could be regarded as dangerous—as when he said that God is beyond good.) Despite the efforts of the Inquisition, the Brethren of the Free Spirit spread, partly because they were able to organize themselves into a secret society.

The asceticism often associated with mystical religion may also be seen in another heretical movement of the twelfth and thirteenth centuries—the Albigensians or Cathari, found in southern France, northern Italy, and parts of Spain, who held doctrines close to those of Manichaeism.

The fourteenth century also saw a marked development of mysticism in England, as exemplified by the writings of Richard Rolle de Hampole (c. 1290–1349), who led the life of a hermit; the anonymous author of the famous *Cloud of Unknowing,* which was influenced by Pseudo-Dionysius; Julian of Norwich (c. 1340–1415); Walter Hilton (d. 1396), and others. On the whole, the temper of their mysticism was nonspeculative, and they emphasized the practical means of developing the inner life.

A movement closely related to the Friends of God was that of the Brethren of the Common Life, which was deeply influenced by Ruysbroeck. Its best-known fruit was the widely read *Imitation of Christ,* attributed to Thomas à Kempis. With its stress on practical love, it was well adapted to the needs of those who did not necessar-

ily feel the call to the cloister and was a means of giving mysticism a wider social impact. Similarly, Catherine of Siena (1347–1380) exhibited a dynamic concern for social and ecclesiastical service. She ministered to victims of the Black Death and played a part in the attempt to strengthen the ailing papacy, persuading Gregory XI to return from Avignon to Rome.

Catherine of Siena spoke vividly of mystical experience in terms of spiritual marriage, paralleling the symbolism whereby the church was looked on as the bride of Christ. Another woman mystic, Teresa of Ávila (1515–1582), gave further expression to this imagery. Her accounts of her own experiences in pursuing the contemplative life, in such works as *The Interior Castle* and in her autobiography, are valuable and sensitive sources for understanding the inner phenomena of mysticism.

Another important mystic who used the imagery of marriage was a younger contemporary of St. Teresa, John of the Cross (1542–1591). He gave detailed expression to the experience of the "dark night of the soul," an experience also recorded by Ruysbroeck and others. The mystic has, according to St. John, periods of despair in which he feels deserted by God. This he interprets as a means of purgation sent by God. The experience probably reflects the contrast between the bliss of union and the condition of striving for that bliss. It is not much written about in nontheistic mysticism, although Buddhist meditation involves the attempt to repress the feeling of bliss accruing on the attainment of higher states of consciousness, in order to obviate the depression liable to occur upon their cessation.

PROTESTANTISM. In one way, Protestantism provided a favorable milieu for mysticism, but in another and ultimately more important way, it provided an unfavorable one. The Protestant emphasis on personal experience of God could easily link up with the ideals of the contemplative life. Thus, the writings of the most famous Protestant mystic, Jakob Boehme (1575–1624), were widely diffused. Groups of followers known as the Behmenists flourished in England and were later absorbed in the Quaker movement, whose doctrine of the "inner light" was characteristically mystical. However, the type of experience that figured so centrally in early Protestantism and that has continued to be stressed in evangelical Christianity was that which gives the individual certitude of salvation. Such a "conversion" experience differs from the imageless rapture that is at the center of mystical religion. Moreover, Protestantism was organizationally unfavorable to the contemplative life, since this had flourished principally in monasteries and indeed had provided a main rationale for their existence. Protestantism could be puritanical, but it did not favor withdrawal from the world.

The antinomian tendencies exhibited by the Brethren of the Free Spirit in the thirteenth and fourteenth centuries were reproduced in various offshoots of Protestant mysticism, as in the movement known as the Ranters, who were strong in seventeenth-century England. Their doctrines were held by opponents to be pantheistic, but more correctly they believed in the essential divinity of all human beings. Since God cannot sin, neither can divinized men, however wrong their actions may look from the standpoint of conventional morality. This was another instance in the history of religion where mystical teachings, normally nurtured in the context of asceticism and unworldliness, were interpreted to justify the opposite. Other important mystics in the Protestant tradition were George Fox (1624–1691), the founder of Quakerism; William Law (1686–1761); and the eccentric poet William Blake (1757–1827).

Although contemplative writings have been less prominent in more recent times, there have been a number of striking mystics since 1850, among them the pseudonymous Lucie-Christine (1844–1908), whose experiences are recorded in her *Spiritual Journal*; the converted French army officer Charles de Foucauld (1858–1916), and the Indian Christian Sadhu Sundar Singh (1889–1929).

Moreover, there has been a renewed scholarly interest in mysticism, as seen in the writings of William James, Evelyn Underhill (1875–1941), and William Inge (1860–1954). Further stimulus to the study of mysticism has been provided by the increased interaction between Eastern religions and Christianity.

ISLAM

Early Islam was not especially conducive to mysticism, since its main spirit was that of the prophetic dynamism of Muḥammad's numinous experiences. Nevertheless, by the eighth century mysticism was developing within Islam. Greek philosophy had already made its impact on the Arabs and thus had opened the way to speculation about God that was partly contemplative. More important, the ex-Christians who had been absorbed into the faith in many Middle Eastern areas carried with them a respect for the ascetic life. Further, the culture of the Arabian desert had encountered the rich and sophisticated standard of living of the conquered, and this confrontation had induced tensions within Islam. Those who held to the older tradition were moved to accentuate the puri-

tanism of early Islam, and such asceticism accorded with the practice of contemplation. Moreover, it was possible for Muslims to interpret Muḥammad's prophetic experience in a mystical sense.

Muslim mysticism is generally known as Sufism. The word *Ṣūfī* is probably derived from the term *Ṣūf,* "undyed wool," which was the material of a garment worn as a sign of simplicity and austerity. Although complete world denial was scarcely in accord with Muḥammad's teachings, the world acceptance expressed in the struggle for power among his successors brought conformity with mere orthodoxy into disrepute among the pious. This represented an opportunity for the growth of an ascetic otherworldliness. Those who adopted the contemplative life could withdraw from politics and could harness self-mortification to the task of concentrating solely upon Allah.

The general structure of Islamic faith was adapted to the service of the inner life. The repetition of prayers enjoined by Islam could be extended from that normally required of the faithful until every moment could be spent in remembrance of God and adoration of him. Almsgiving, one of the seven "pillars of Islam," could be interpreted in terms of thoroughgoing self-denial. The whole of life could be seen as a pilgrimage to a spiritual Mecca. Although the earliest teachings of Islam had laid duties on the individual as a member of the community—conceived as a brotherhood—tendencies later developed that made religion essentially a matter for the individual alone.

The new asceticism was regarded primarily as a means toward inner illumination. Fear and obedience of God melted into a burning interior love of him that carried with it the hope that union with him might be gained through negation of the self. This interior knowledge was described in terms of light, and an important passage in the Qurʾan (Koran), the so-called Light Verse, was quoted as a backing for mysticism: "God is the light of the heavens and of the earth; His light is like a niche wherein there is a lamp, a lamp encased in glass, the glass as it were a glistening star." Also, the Sufis came to use the imagery of love as some Christian mystics did. An early example of this is to be found in the life and teachings of Dhū'l-Nūn (d. 861), an Egyptian influenced by Greek speculation.

HERETICAL ASPECTS. The knowledge prized by the Sufis was not the rational knowledge developed by the scholastic theologians (in Islam this meant mainly those who had come into contact with Greek philosophy); rather, it was the direct knowledge of Allah, or *maʿrifa.* This *maʿrifa* or gnosis was the crown of the Sufi path. However, the idea of direct acquaintance with God could have consequences that were scandalous to the orthodox.

Thus, Abū Yazīd of Bistam (d. 875) was so convinced of his identity with God in the experience of *maʿrifa* that he could say "Glory to me—how great is my majesty." This seemed like claiming divinity, which was blasphemous and strictly contrary to the orthodox opposition to any doctrine of incarnation. Abū Yazīd also put forth an idea destined to play a large part in subsequent Islamic mysticism—that of *fanāʾ,* the passing away and extinction of the empirical self, which follows self-control through asceticism and contemplative techniques. The "passing away" involved the loss of the consciousness of one's own individuality and helps to explain why the Sufis sometimes spoke in terms that suggested that they became merged or identified with God. As has been seen, similar ideas were expressed on occasion by Christian mystics such as Eckhart and are found in Hindu and Buddhist mysticism.

The most notable example of this trend was the experience of al-Hallāj (854–922) of Baghdad, who spoke as though he were an incarnation of the divine Being through mystical experience and consciously and overtly modeled himself upon Jesus. Such ideas were intolerable to the orthodox and he was (appropriately) crucified.

Although at first the Sufis operated individually, they later associated in loose groups. The elaboration of contemplative techniques and the trend toward celibacy (scarcely in accord with the spirit of the revealed law contained in the Qurʾan) brought about the creation of orders of Sufis who could work, and often live, together. It was common for such a group to be under the spiritual direction of a *shaykh* or *pīr,* and very often his residence would turn into a monastic community. The prestige of such holy men became great, and miraculous powers were ascribed to them. This prestige, combined with concepts clustering around *maʿrifa,* brought the ideal of the divine human and the cult of saints into Islam.

Persecution, as in the case of al-Hallāj, was no lasting answer to threats to orthodoxy; what was required was a synthesis between the new ideas and traditional theology that could harness Sufi piety to Qurʾanic ends. Al-Ghazālī (1058–1111) provided the most acceptable and influential solution to the problem. In his *The Revival of the Religious Sciences* he dealt with the question of how *fanāʾ* could most properly be interpreted. He held that the mystic, in experiencing the vision of God, is so overwhelmed that he imagines he is united with him. However, this is a sort of illusion, analogous to the belief of a person who

sees wine in a transparent glass and thinks that wine and glass are a single object. When the contemplative returns from the state of ecstasy ("drunkenness," as Ghazālī called it—metaphors of drinking were common in Sufi writings), he recognizes that there is a distinction between the soul and God. In such ways, Ghazālī tried to do justice both to the actual experience of the contemplative and to a religion's requirements of worship, which presupposes a dualism between the worshiper and the object of worship. Ghazālī stressed the way in which self-purification, as part of the Sufi path, follows penitence, which in turn depends on the recognition of the awe-inspiring majesty and holiness of Allah. Thus he tried to show that contemplation and orthodox religion go hand in hand. Hence, he also did not believe in a mysticism that involved withdrawal from the world. The mystic returns to ordinary life, revitalized by the dazzling vision of the divine Reality. Ghazālī's synthesis meant that henceforth Sufism had an accepted place within orthodox Islam, but contemplative and philosophical thought were not restricted.

PANTHEIST TENDENCIES. Notable among those who expressed a poetical and metaphysical Sufism was Ibn al-ʿArabī (1165–1240) of Spain. He influenced Dante Alighieri, who adopted the outline of Ibn al-ʿArabī's description of the ascent into heaven (combining astronomical theory and the story of Muḥammad's journey to heaven). His doctrines were pantheistic, and he considered human beings as offshoots of the divine essence that exist because of God's desire to be known; and in the realization of the divine Being, the contemplative reflects in his own person the structure of the universe. He also made use of the logos idea: The logos as the creative principle in the universe was identified with the spirit of Muḥammad. However, there are hints in Ibn al-ʿArabī's work that he considered himself superior to Muḥammad, having realized identity not with the logos but with the Godhead.

His voluminous writings, although regarded with distaste by the orthodox, were influential, especially in Persia, among such mystical poets as Jalāl ad-Din Rūmī (1207–1273) and Mawlana Nur ad-Din Jāmī (1414–1492). Rūmī, who founded one of the *darwīsh* orders (*darwīsh* literally means "mendicant," and is commonly transliterated *dervish*), also wrote poetry expressing the longing of the soul for its return to God. However, he was also keenly appreciative of the beauties of nature, and he saw in the ritual of the Mevlevi order, which he founded, with its solemn swirling dance to the sound of drum and pipe, a reflection of the movements of the planets and of nature in general.

It may be noted that some of the orders experimented with various external means of inducing ecstatic experiences, and the dance was one. (The term *dervish* should properly apply to all mendicant orders, and not just to the Mevlevi "dancing dervishes.")

Certain features of Sufi teaching are reminiscent of Indian mysticism, and it has been argued, although not conclusively, that there were borrowings from India. (See R. C. Zaehner, *Hindu and Muslim Mysticism,* on this question.) For instance, Abū Yazīd's language is similar to that of the *Upaniṣads*; and Ibn al-ʿArabī argued, with a logic like that of Śankara, that it is inappropriate to speak of *becoming* God through mystical experience, since one is already essentially identical with God—mystical realization involves no change of ontological status. Again, like nearly all Hindu theologians, Ibn al-ʿArabī treated hell as a purgatory, rather than as a place of everlasting punishment. Various similarities of this kind can probably best be explained not so much as borrowings but rather as reflections of similar patterns of experience and speculation.

MODERN SUFISM. In the modern period, Sufism has undergone a considerable decline, and the revitalization of Islam has come about through other forces—the puritanism of the Wahhābī, Pan-Arabism, and political advance. Sir Muhammad Iqbal (1877–1938), however, an important figure in Muslim modernism, was influenced by Sufi thought. Since he wished to distinguish sharply between religion and science—the former having to do with personal life—he found the interior quest of Sufism attractive.

See also Absolute, The; al-Ghazālī, Muhammad; Asceticism; Augustine, St.; Bernard of Clairvaux, St.; Blake, William; Boehme, Jakob; Bonaventure, St.; Buddhism; Buddhism—Schools: Chan and Zen; Caird, Edward; Chinese Philosophy; Confucius; Dante Alighieri; Eckhart, Meister; Gnosticism; Hinduism; Ibn al-ʿArabī; Illumination; Inge, William Ralph; Iqbal, Muhammad; Jainism; James, William; John of Damascus; John of the Cross, St.; Kabbalah; Laozi; Law, William; Lu Xiangshan; Mani and Manichaeism; Mysticism, Nature and Assessment of; Mysticism: The Indian Tradition; Neoplatonism; Nirvāṇa; Otto, Rudolf; Philo Judaeus; Pseudo-Dionysius; Rāmānuja; Ruysbroeck, Jan van; Śankara; Stace, Walter Terence; Sufism; Suso, Heinrich; Tauler, Johannes; Teresa of Ávila, St.; Thomas à Kempis; Thomas Aquinas, St.; Thomism; Wang Yangming; Yoga; Zhu Xi (Chu Hsi).

Bibliography

A good introduction to mysticism is Sidney Spencer, *Mysticism in World Religion* (Baltimore: Penguin, 1963). Evelyn Underhill's *Mysticism,* 6th ed. (London: Methuen, 1916), is a classic, although dated in some respects and confined largely to theistic mysticism. The same is true of William James, *Varieties of Religious Experience* (New York: Longmans Green, 1902). See also W. R. Inge, *Mysticism in Religion* (London, 1907); R. C. Zaehner, *Mysticism Sacred and Profane* (Oxford: Clarendon Press, 1957); W. T. Stace, *Mysticism and Philosophy* (Philadelphia: Lippincott, 1960); Steven Katz, ed., *Mysticism and Religious Traditions* (New York: Oxford University Press, 1983); Bernard McGinn, *The Presence of God,* 3 vols (New York: Crossroad, 1991–1998); Barry Windeatt, *English Mystics of the Middle Ages* (Cambridge, U.K.: Cambridge University Press, 1994); Geoffrey Parrinder, *Mysticism in the World's Religions* (New York: Oneworld, 1995); and Steven Katz, ed., *Mysticism and Sacred Scripture* (New York: Oxford University Press, 2000).

On Indian mysticism, see Edward Conze, *Buddhist Meditation* (New York: Dover, 2003); S. N. Dasgupta, *Hindu Mysticism* (London, 1927); D. T. Suzuki, *Mysticism: Christian and Buddhist* (London: Allen and Unwin, 1957); Rudolf Otto's classical *Mysticism East and West,* translated by B. L. Bracey and R. C. Payne (New York and London, 1957, in paperback); Ninian Smart, *Doctrine and Argument in Indian Philosophy* (London: Allen and Unwin, 1964), Ch. 10; Krishna Sivaraman, ed., *Hindu Spirituality: Vedas through Vedanta,* World Spirituality Series vol. 6 (New York: Crossroad, 1989); K.R. Sundararajan and Bithika Mukerji, eds., *Hindu Spirituality II: Postclassical and Modern,* World Spirituality Series vol. 7 (New York: Crossroad, 1997).

On Chinese and Japanese mysticism, see K. L. Reichelt, *Meditation and Piety in the Far East* (London: Lutterworth Press, 1953) and *Religion in Chinese Garment* (London: Lutterworth Press, 1951); Henri Maspero, *Le Taoisme* (Paris, 1950); Fung Yu-lan, *A History of Chinese Philosophy,* 2 vols. (Princeton, NJ: Princeton University Press, 1952–1953); and Julia Ching, *Mysticism and Kingship in China: The Heart of Chinese Wisdom* (Cambridge, U.K.: Cambridge University Press, 1997); Livia Kohn, *Early Chinese Mysticism: Philosophy and Soteriology in the Taoist Tradition* (Princeton, NJ: Princeton University Press, 1992). The *Dao-de-jing* has numerous English translations, ranging from that of Lionel Giles (London, 1911) to that of D. C. Lau (Baltimore: Penguin, 1963).

For mysticism in Judaism, see G. G. Scholem, *Major Trends in Jewish Mysticism* (London: Thames and Hudson, 1955); E. Müller, *History of Jewish Mysticism* (London, 1956); Martin Buber, *The Legend of the Baal-Shem,* translated by Maurice Freedman (London: Horowitz, 1956); Ira Chernus, *Mysticism in Rabbinic Judaism: Studies in the History of Midrash* (Berlin: Walter de Gruyter, 1982); and Arthur Green, *Keter: The Crown of God in Early Jewish Mysticism* (Princeton: Princeton University Press, 1997).

On Christian mysticism, see R. M. Grant, *Gnosticism and Early Christianity* (London: Oxford University Press, 1959); W. R. Inge, *The Philosophy of Plotinus,* 2 vols. (London: Longmans Green, 1918); A. J. Festugière, *Personal Religion among the Greeks* (Berkeley: University of California Press, 1954); Cuthbert Butler, *Western Mysticism,* 2nd ed. (London: Constable, 1922); Evelyn Underhill, *The Mystics of the Church* (London: James Clarke, 1925); Vladimir Losskii, *The Theology of The Eastern Church,* translated by the members of the Fellowship of St. Alban and St. Sergius (London: James Clarke, 1957); Jon Gregerson, *The Transfigured Cosmos* (New York: Ungar, 1960); Margaret Smith, *Studies in Early Mysticism in the Near and Middle East* (London: Sheldon Press, 1931); R. M. Jones, *Studies in Mystical Religion,* 3rd ed. (London: Macmillan, 1909); E. A. Peers, *Studies of the Spanish Mystics,* 3 vols. (London: Sheldon Press, 1927–1960); and Louis Bouyer, et al., *A History of Christian Spirituality,* 3 vols (New York: Seabury, 1969).

On mysticism in Islam, see A. J. Arberry, *Sufism—An Account of the Mystics of Islam,* 2nd impression (London, 1956); Alexander Knysh, *Islamic Mysticism: A Short History* (Leiden: Brill, 2000); R. A. Nicholson, *Studies in Islamic Mysticism* (Cambridge, U.K.: Cambridge University Press, 1921, reissued 1979); Annemarie Schimmel, *Mystical Dimensions of Islam* (Chapel Hill: University of North Carolina Press, 1975); A. J. Wensinck, *La pensée de Ghazzali* (Paris: Adrien-Maisonneuve, 1940); and R. C. Zaehner, *Hindu and Muslim Mysticism* (London: University of London, Athlone Press, 1960).

Ninian Smart (1967)

Bibliography updated by Christian B. Miller (2005)

MYSTICISM, NATURE AND ASSESSMENT OF

Attempts to define mystical experience have been as diversified and as conflicting as attempts to interpret and assess its significance. This is not surprising, for the language used to express and describe mystical experience is richly paradoxical, figurative, and poetical. Even if at times a mystic chooses what look like austere and precise metaphysical terms, this may be only an apparent concession to logic, for he will employ these terms in senses far from normal. Mystics have called the Godhead a sheer "Nothing" and yet the ground of all. They have affirmed simultaneously that the world is identical with God and that the world is not identical with God.

Some discriminations are possible, even if exact definition is not. Mystical experience is religious experience, in a broad but meaningful sense of "religious." It is sensed as revealing something about the totality of things, something of immense human importance at all times and places, and something upon which one's ultimate well-being or salvation wholly depends. More specifically, a mystical experience is not the act of acquiring religious or theological information but is often taken to be a confrontation or encounter with the divine source of the world's being and man's salvation. An experience is not held to be mystical if the divine power is apprehended as simply "over-against" one—wholly distinct and "other."

There must be a unifying vision, a sense that somehow all things are one and share a holy, divine, and single life, or that one's individual being merges into a "Universal Self," to be identified with God or the mystical One. Mystical experience then typically involves the intense and joyous realization of oneness with, or in, the divine, the sense that this divine One is comprehensive, all-embracing, in its being. Yet a mystical experience may be given much less theological interpretation than this description suggests. A mystic may have no belief whatever in a divine being and still experience a sense of overwhelming beatitude, of salvation, or of lost or transcended individuality.

Some mystical experiences occur only at the end of a lengthy, arduous religious discipline, an ascetic path; others occur spontaneously (like much nature-mystical experience); others are induced by drugs such as mescaline or take place during the course of mental illness.

An important distinction can be made between the extrovertive (outward-looking) and introvertive (inward-looking) types of mystical experience. In the first of these, the subject looks out upon the multiplicity of objects in the world and sees them transfigured into a living, numinous unity, their distinctness somehow obliterated. In nature mysticism, a form of extrovertive experience, the items of nature are not lost to consciousness; rather they are seen with unusual vividness and all as "workings of one mind, the features/Of the same face, blossoms upon one tree" (William Wordsworth, *The Prelude,* Book 6). In the introvertive type, the mystic becomes progressively less aware of his environment and of himself as a separate individual. He speaks of being merged in, identified with, dissolved into, the One. The subject-object distinction vanishes altogether. Some of the best-known mystics testify to experiences of both types, but the introvertive, being at the furthest remove from ordinary experience, is usually held to be the more developed of the two.

Although we can call mystical experience a kind of religious experience, we do not discover agreement among mystics about the nature and status of the mystical goal. Christian and Islamic mysticism, for example, interpret the experience theistically, although not with complete consistency; the Upanishads and Theravāda Buddhism are not theistic. Pantheist, monist, and agnostic interpretations have been offered, all with some prima facie plausibility.

ALTERNATIVE RELIGIOUS INTERPRETATIONS

The pantheist argues that mystical experience compels us to strip away anthropomorphic conceptions of deity and that although theism begins this work of refining, it stops long before it should. The theistic notion of God remains that of an infinite, supernatural individual. But apart from being intellectually unsatisfactory (infinity and individuality go awkwardly together), this picture contradicts the mystic's own experience, which is one not of an external face-to-face meeting with a deity but rather of merging with, and realizing one's own basic identity with, the mystical One. The theist has to set a great gulf between himself and his God; the mystic's experience testifies both to the existence of this gulf and, paradoxically, to its elimination. Brahman is both far and near.

Why have so many of the greatest Christian mystics used theistic language to describe their obviously intense mystical experiences? The pantheist will say that either they have simple-mindedly used the only religious terms they had been taught—despite their unsuitability—or else that the desire to conform to orthodox Christian dogma about God's transcendence has led them to muffle those parts of their individual experience that were opposed to it.

A pantheist interpretation claims that it alone does full justice to God's infinity and that its theology eliminates the last primitive remnants of deism. Since a mystical experience is a discovery, a realization, of what is eternally true, there need be no perplexing doctrines about special divine self-revelations and self-communications nor any interference with natural law. Accordingly, a mystical experience induced by drug or disease does not have to be judged illusory or demonic. In the determination of whether it is authentic or not, its causal circumstances are simply beside the point.

The theist, however, is not without a reply. He will reject the pantheist's conception of religious development. There has not been any general historical trend toward pantheism or monism in religion; and although early theisms were crudely anthropomorphic, this does not by itself entail that all personal language about God is equally false and crude. The doctrine of the Incarnation should teach the contrary—at least within Christendom.

Pantheism and monism, argues the theist, map only the lower slopes of the mystic's ascent. They are concerned with the preliminary purging of the senses and intellect; their raptures do not testify to an achieved union with God but only to what is perhaps an unusually fresh, innocent, and aesthetically intense awareness of the created world and its beauty. The mescaline-user and the temporarily psychotic, who make extravagant claims for their own identity with the mystical One, ought to—often do—think more humbly of their experiences once

normality returns. To the theist, the *unio mystica* is an objective that cannot be taken by assault; in the end, it is only the initiative, the grace, of God that bestows it. Causation does matter in this interpretation, and the inner, felt nature of the mystical experience cannot alone determine its authenticity.

PARADOXES OF RELIGIOUS INTERPRETATIONS

Short decisive arguments can hardly be invoked to settle the dispute between these interpretations of mystical experience. The experiences themselves seem able to bear either interpretation; the choice between pantheism and theism is a choice between two massive conceptual systems. Neither account can claim the merit of being free from internal difficulties both conceptual and religious. Theism has somehow to combine the notions that God is immeasurably "other" to man and, yet, that mystical union is possible. Pantheism identifies world and God while maintaining their distinctness; it denies that "God" is simply another way of saying "world."

Still more perplexingly, some mystics of great eminence speak the languages of both pantheism and theism. Meister Eckhart's writings give full-blooded examples of each, as do those of the Indian mystic Śankara. Even in the Upanishads, although Brahman is said to be beyond relation, featureless, unthinkable, it (or he) is acknowledged to have personal aspects.

No precise or determinate idea, no particularized image, is allowed to be adequate to the mystical One. Although the ontological status of God seems at times to be that of a numinous individual being, at other times all hints of such a status are repudiated. "Simple people," said Eckhart, "imagine that they should see God, as if He stood there and they here. That is not so." The Divine is a "desert," a "void," an "abyss," a "wheel rolling out of itself," a "stream flowing into itself."

Mystics will not always allow one even to say unequivocally that God exists. The pseudo-Dionysius, for example, denied that either the category of existence or of nonexistence applied to the Divine. These tensions and this indeterminateness—God is, or is not, a particular being, he is, or is not, an existent—can also be found in nonmystical theologies, but mysticism can enormously magnify them. Even Theravāda Buddhism contains deep-running paradox, despite its comparative reluctance to speculate at all. Attaining nirvāṇa, for instance, is like the extinguishing of a flame, yet nirvāṇa is not sheer simple extinction.

What attitude is it reasonable to adopt toward this display of tensions and antinomies? Four possibilities are worthy of serious discussion. (1) The paradoxes cannot be eliminated; they are to be taken literally and at their face value. Without paradox, we cannot speak of the mystic's experiences or of his God, but this is no argument against the truth of the mystic's claims. (2) The paradoxes are necessary in the same way that distortions of grammar and syntax are necessary to a poet attempting to say something that cannot be encompassed by ordinary language. They are not to be taken literally but are to be construed as analogies, hyperboles, metaphors, or oxymorons. (3) Since no logically coherent account of mystical vision seems attainable, it is more sensible to admit this fact and to believe the mystic's claim that his experience is ineffable and that all language falsifies it. We would now have a mysticism without a theology. A very high value could still be set upon mystical experience, but we should be reverently agnostic on all questions of interpretation. (4) The appearance of paradox in a piece of discourse is very often taken by philosophers as a reductio ad absurdum of its claims. (Compare the logician's story of the barber who shaves only those who do not shave themselves. When paradox arises over the question "Does the barber shave himself?," it is reasonable to infer that there logically cannot be a barber, so described.) Because the mystic says so many contradictory things about God, this demonstrates the logical impossibility of God's existence, so described. Criticisms charging illogicality can be supported by attempts to explain in naturalistic terms the mystical experiences themselves.

EVALUATION OF RESPONSES TO PARADOXES

Whether or not the paradoxes are finally to be judged literal and irreducible, we must clearly reject some of the speculations that are aimed at reducing their offense. For example, how God can be, but not by being an individual entity, is profoundly obscure. The mystery is not removed if we say that God is Being Itself or Being as such. Even if our ontology allowed such universals as "courage itself" or "blueness itself," we still could not meaningfully include Being Itself among their number; there is no characteristic named "being" that is common to all actual entities and that should figure in their complete description. "Being Itself" cannot logically refer to anything either particular or universal, divine or nondivine.

Similarly, if we are offended by the claim that God neither exists nor does not exist, we might try a familiar palliative and say that he is above being. Our concepts fail

to grasp him precisely for that reason. "Above being" carries echoes of "above the turmoil," "above suspicion," "above praise," with "above" indicating distance from and superiority to something. But in order to be "above," one must first of all be—and continue to be. "God is above being" really fails to satisfy the conditions under which any "above" sentence of this kind can have meaning. It can, of course, be given a sense if "being" here means finite and dependent being. But if God is superior to this sort of being, if he is infinite and independent, then that is a superiority of his nature, and to learn this about him gives us no help with the original paradox.

LITERAL VERSUS FIGURATIVE LANGUAGE. The paradoxes and enigmas may have to stand, but why not take them as poetical, metaphorical, or symbolic language? Against that suggestion, it may be argued that if the paradoxes are metaphors, it should be possible to translate them—at least roughly—into direct, nonmetaphorical language. The only language available to the mystic, however, seems to be a language of irreducible paradox.

This argument is not very powerful. There are nonmystical topics about which it is impossible to speak without metaphor, such as important topics within the philosophy of mind. The history of conceptions of the mind is, in many of its facets, the history of changing metaphors, myths, and analogies. To defend a parallel account of mystical discourse would be less of a scandal to reason and logic than to insist on the literal view.

The literalist will reply that there is, in fact, no scandal to reason. The laws of logic work admirably for every situation where multiplicity is present. In the mystic's unique case, all multiplicity has vanished and with it, therefore, the applicability of those laws. The mystic's discourse is about the One that has no other; it lies beyond the province of logic.

This leaves us with a discomforting worry. If logic is inapplicable to the mystic's discourse, does that not come very close to saying that discriminations cannot be made in this field between sense and nonsense, the sound and the unsound?

The literal approach must be, for a philosopher, a desperate measure, a last resort only. To treat it as anything else would be methodologically perverse. Apart from the difficulties of discrimination, where logic is inoperative, the approach demands an unshakable prior conviction that the mystic's paradoxes are to be taken at their face value as reports of veridical insights. Here there is much that can be challenged.

We refused to dismiss the figurative account for not being able to translate its metaphors, or to give literal equivalents for its symbols and analogies. Yet that inability is nonetheless an embarrassment to it. When the mystic says, "God is a desert"; "God is a blinding light"; "God is, and is not, identical with the world"; or "The mystical enlightenment is an absolute emptiness which is absolute fullness"; we are compelled to accept these metaphors and paradoxes on the faith—if we accept them at all—that they can be true in some inscrutable way of one and the same deity. This cannot be shown, although the mystic *feels* intensely that it is so. The skeptic complains that he cannot begin to see how such wildly incompatible predicates can refer to any one being, whereas he can understand with relative ease how they might, in fact, be the expression of some ecstatic inner experience of a quite noncognitive kind. He does not deny that some apparently incompatible predicates may be revealed as ultimately compatible. A psychoanalytic story can reveal how love and hate, desire and fear, can be harbored simultaneously by a person for a single object; the same can be true with conflicting analogies and metaphors. The last word of the mystic, however, is "ineffable"; he does not profess to have a reconciling story.

An objector might now suggest that it is easy enough to see how we could choose senses for the words *abyss, desert, light,* that would give us at least a glimmer of insight into their metaphorical reference to the same divine being. The words are rich enough in their connotations and implications, both near and remote. This is true, but it cannot be a key to all the paradoxes. Certain ones (like that of identity and difference between God and world) offer no scope at all for such imaginative siftings and surmisings—unless we paraphrase the mystic's claim so freely that he will disown our translation. "The world is, and is not, identical with God" does not mean to the pantheistic mystic that the world is godlike in some respects and not in others.

If a city were referred to as a desert, a trap, or a furnace, the selection of appropriate meanings for these words in their metaphorical use would be possible because of the knowledge of the given fixed point of reference: a city. However, the concept of city is ontologically stable and intelligible in a way that the concept of God is not. The mystic's paradoxical discourse is related ultimately to his basic assertions about God's metaphysical status; this makes his semantic situation enormously more complex and precarious. Once again, these reflections do not attempt to disprove the mystic's statements or even to show that they cannot be figurative as well as

semantically sound. If the mystic had independent grounds for believing in God, then one could readily accept the claim that he could speak about this God only in oblique language. Some mystics would say that they do have such independent grounds, but for others the mystical experiences themselves, reported in the language of paradox, furnish the grounds of belief. Here the risk of delusion is higher.

MYSTICAL EXPERIENCE AND AGNOSTICISM. "According to our scale of values," Rudolf Otto wrote, we shall consider the mystic's intuition "either a strange fantasy or a glimpse into the eternal relationships of things" (*Mysticism East and West,* p. 42). Need these be the only options? Might it not be possible to reject all the traditional interpretations of mystical experience but yet accord it very high intrinsic value? If the mystic cannot interpret his experience theologically without talking nonsense, it is then better for him not to attempt theology or metaphysics at all, lest he bring his experience itself into needless disrepute.

An approach of this kind would have strong sympathy with the agnostic elements of early Buddhism. Buddha taught the path to nirvāṇa but turned away any question about deities or the nature of a life hereafter. His emphasis was upon the moral quality of a life and upon attitudes toward life, death, suffering, and release from suffering. Mystical experience was attained in the course of a personal, practical discipline. It was understood as the culmination of such a discipline and given only the minimal theoretical interpretation. The lack of speculation did not, however, make the mystical experience unavailable to one who followed the Buddha's prescription for attaining it.

To insist that mysticism is possible without interpretation has the merit of avoiding unnecessary intellectual offense; it also allows us to admit as mystical the experiences of people outside both the theistic and monistic traditions but whose testimony, at the phenomenological level, shows great affinities with the mysticism of both traditions. Nevertheless, the mystical experiences of an agnostic are surely bound to differ in important respects from those of a Christian, a Buddhist, or a Muslim. The concepts used in interpretation help to determine the mystic's expectations of future experiences and to determine his map of the mystical path and the plotting of his position upon it. They shape the actual quality of his experience itself in a most intimate way. This does not imply that, but for the interpretative concepts, no experience could occur.

It may be feared that the theologically uninterpreted experience would tend to become a mere psychological curiosity, a luxury or consolation, isolated from all other parts of the subject's life. This can happen, but need not. Mystical experience basically involves a powerful urge toward the reconciliation, unification, and harmony of all with all, a feature that can readily be integrated with a moral outlook in which primacy is given to love. "Integrated," in fact, is really too weak a term; that moral ideal may receive its fullest and most splendid development in the mystical vision, and the moral agent gains a source of energy for the pursuit of the moral life.

These reflections may show, at least, that we cannot fairly assess the importance of mystical experience solely in terms of the interpretations that may be offered of it, whether speculatively pretentious or modest. An equally relevant question is what the mystic does with his experience, that is, what place he gives it in his total personal and moral existence. Evaluations based on this issue may often be at variance with those based upon a comparison of theories. A mystic may interpret elaborately and use his mystical experience as a mere refuge from responsibility, or he may be quite at a loss for interpretation, while recognizing in his experience the center and spring of a morally dedicated life.

OTHER PHILOSOPHICAL CRITICISMS

Our fourth type of response to the phenomena of mysticism was that offered by the radical philosophical critic, determined to call nonsense by its name, who takes the mystic's antinomies as a reductio ad absurdum of his claims. To those logical objections philosophers have added various epistemological and psychological difficulties.

THE PROBLEM OF OBJECTIVITY. The mystic (and we are no longer thinking of the agnostic mystic) normally claims that his experience is not only a way of being inwardly, subjectively moved, but also that it discloses the nature of reality, that it is a cognitive, objective experience. To support this he may appeal to the impressive convergences of testimony on fundamentals among mystics of different periods and parts of the world. The critic may contest this. In reports upon perceptual illusions, for instance, even unanimity does not remove their illusoriness.

That the experiences are disclosures about the entire universe in its ultimate nature may be an almost irresistible conclusion for the mystic. Nonetheless, it must involve interpretation of a demonstrably fallible kind. To

feel that the experience is revelatory is one thing; to judge confidently that it is so is quite another. A dream under nitrous oxide may strike the dreamer with the force of a satanic revelation, but on awakening and correlating the nightmare with the shock of tooth extraction, he may have little temptation to judge the experience as a genuine disclosure. The feeling of revealedness can attach itself with equal intensity to incompatible contents.

W. T. Stace has argued that mystical experience is neither objective nor subjective but that it transcends this distinction and is best classified as transsubjective. To be objective, an experience must be orderly and law governed; the criteria of subjective experience are disorderliness and incoherence. Mystical experience fits neither category. It is an experience of unity, untouched by plurality; and without plurality there can be neither order nor disorder.

This is an ingenious treatment, but it seems open to criticism at least on two points. First, the criterion of objectivity may be questioned. We may be quite properly convinced that certain phenomena are objective before we have assured ourselves of their orderliness, and they may indeed remain anomalous. The subjective events of dreams and fantasies are not disorderly, although the laws governing dreams are very different from those governing events in the public world. Second, we may wish to deny that mystical experience is, in fact, experience of a totally undifferentiated unity. There is, no doubt, a stage in which the mystic not only apprehends the world of plurality as issuing from a single divine source but sees that source and the world as a unity. Mystical experiences, however, cannot usefully be restricted to this one type. Perception of multiplicity does play a role, even if it is a subordinate one, in many other types. This is obviously so with extrovertive mystical experience in general, which is an experience not simply of oneness but of oneness in multiplicity. It is also apparent in the statement from Sri Aurobindo that "those who have … possessed the calm within can perceive always welling out from its silence the perennial supply of the energies which work in the universe" (*The Life Divine,* 1949, p. 28). The most favorable verdict we can pass upon claims to objectivity is "not proven."

EPISTEMOLOGICAL PROBLEMS. When we ask more particularly what sort of apprehension, what modes of knowing are involved in mysticism, the answers swell our fund of paradoxes. If one mystic claims to perceive the cosmic energies welling forth from the One, another denies that anything like perception takes place. St. John of the Cross speaks of a "supernatural knowledge and light" that is so completely "detached and removed from all intelligible forms, which are objects of the understanding, that it is neither perceived nor observed" (*The Ascent of Mount Carmel,* Vol. I, p. 123). Nor is mystical insight a purely intellectual act, for "the higher and more sublime the Divine light, the darker is it to our understanding." Union with God "transcends all knowledge." The difficulty is increased by the doctrine that in mystical experience the subject-object distinction breaks down, and with it, naturally enough, go all our thought models for cognitive activities. Faced with the risk of a complete failure in communication, the mystic usually resorts to a characteristic complex use of language. This works in part by negations ("not ordinary perception," "not simply emotion") and in part by descriptions of his religious situation as he interprets it in metaphysical and theological terms, enhanced with poetical imagery; God now dwells in him, or has "absorbed" him "in the embrace and abyss of His sweetness." It is easy to see why the mystic resorts to these forms of discourse and also why they offer little comfort to the epistemologist. For the interpretations assume precisely what is at issue: that mystical experiences are objective and reliably cognitive in nature.

Some critics maintain that the mystic's claim to "know" must at least be suspected of being spurious. When such expressions as "objectivity," "discovery," and "vision" are used in senses so radically far from normal and applied with obscure and idiosyncratic criteria, it is legitimate to ask whether some quite different (and noncognitive) thought model might give a more intelligible clue to what is being described.

For example, it is sometimes suggested that the mystic's language might be best understood not as a description of reality but as the expression of a state of mind. Certainly, some of the mystic's language is clearly emotive, and even when it seems to describe his "situation," as we have been using the word, this may still be an indirect expression of his state of mind. Instead of saying, "I have an oppressive, worried feeling," one may say, "I feel as if there were something terribly wrong." Instead of "I feel uneasy, insecure," he may say, "There is no sure footing; everything and everybody is working against me." Instead of "I have a feeling of unreality," he may say, "I am not real anymore." The use of such examples does not imply that the mystic is psychotic. Some psychotic experiences are mystical experiences, but it hardly follows that all mysticism is psychosis. The critic could confine himself to pointing out this disturbing parallel in the use of language: Both mystics and psychotics use situation-descrip-

tive language for what, in the latter case at any rate, is a serious misperception of one's situation, a projection of inner disturbances upon the outer world. Furthermore, the projection occurs, partly at least, because the disturbances are not understood for what they are, and there is a failure of insight.

In the mystic's defense, it must be pointed out that to analyze his experience as a state of mind is not necessarily to discredit it. States of mind can be—and normally are—elicited by objective states of affairs, properly interpreted. People do, on occasion, fall victim to real persecution; their fears and anxieties can be very well founded.

But decisiveness, either in criticism or defense, is once more not to be had. Of course one's fears can be well-founded, but a person who says he does not really exist any more *must* be deluded. Significantly, as soon as such remarks verge on the paradoxical, we cease to take them at their face value and treat them as certain signs of disorder.

CONTENT AND QUALITY OF MYSTICAL EXPERIENCE. We have been considering some epistemological and linguistic problems set by mysticism and some ways in which a philosophical critic can assault, although probably not overthrow, the mystic's claims. Of the central mystical experience, characterized by loss of individuality and dissolution in a limitless divine totality, little or nothing has been said from a philosophical or psychological viewpoint. How far could a naturalistic account of mystical experience cope with these central features? Or could justice be done to them only in a thoroughgoing mystical philosophy, reared upon the paradoxes themselves? Here a suggestion or two must suffice.

In the first place, the mystical experience is a vision of the world that is free, to a very unusual extent, from the interposition of concepts. Normal perception is closely linked to practical projects; we see the world in terms of our needs and desires and our intentions to manipulate it in various ways. Aesthetic experience provides a sharp contrast. One may succeed briefly in contemplating a pastoral landscape not in terms of land utilization or of the practical problems of traveling across it, but simply as colors, shapes, or volumes. Seen in this way, the landscape can be excitingly and startlingly different from its everyday utilitarian appearance. Mystical experience is even more disturbingly strange because it suspends the application of still more basic concepts and categories. "As long as a man has time and place and number and quantity and multiplicity, he is on the wrong track and God is far from him" (Meister Eckhart, *Sermons,* p. 202).

When concepts are withdrawn and fundamental distinctions obliterated, it is understandable that our ordinary sense of the limits and boundaries between thing and thing, person and person, should also temporarily disappear. In this we may have an important clue to the mystic's claims about the overcoming of finite individuality, the cessation of the subject-object relation, and mergings and meltings into the infinite. Because our normal sense of our powers and their limits is fostered by the utilitarian and practical view of the world, when that view is suppressed, there can come the sense of exhilarating expansion or liberation that is often described in the mystical literature.

Similarly, if the practical orientation is suspended and, with it, the related conceptual framework of normal experience, we may lose awareness of the passage of time. We are not demarcating event from event in the normal time-articulating manner. In introvertive mystical experience the awareness of space is also obliterated, for there is a still more thoroughgoing withdrawal from perception and even from sensation. The intensity and strangeness of mystical experience reinforce the effect of timelessness; the experience is dramatically discontinuous with the flow of events before and after and hence is felt as not belonging to it.

The mystic himself can afford to be sympathetic to many such naturalistic explanations. He can refuse to admit that they discredit his experience. They are simply (he will say "necessarily") incomplete, for they cannot account for the qualitatively unique tone of mystical feeling, and they do not disprove his claim that the object of mystical vision itself must elude the categories of naturalistic philosophy.

Mysticism can be upgraded or downgraded with bewildering ease through the choice of a metaphor or a simile; its paradoxes are unutterable truths or blatant contradictions; its clearest affinities are with trustworthy modes of knowing or with psychotic, delusory states of mind; of all human experience it is the most valuable or it is a psychological curiosity, fashioned by the unconscious from infantile materials. The excesses of these opposite poles are avoided in our remarks about an "agnostic" or "noninterpreting" mysticism, although this is perhaps more of a practical compromise than the germ of a full-fledged theory. It tries at least to stress the potential human importance of mystical experience—when yoked to moral vision—and it expresses the wishful thought that the paradoxes of mystical interpretation should not be altogether allowed to mask that importance.

Perhaps it is more advisable to reflect on the meaning of ineffability claims made by mystics within their contexts, and the complex ways of interaction between mystical experiences and mystical traditions (Katz 1992).

THE DEBATE OVER THEISTIC MYSTICISM

How we should classify different types of mysticism continues to be controversial. Some scholars do not regard theistic mysticism as a separate type. They argue that all mystical experiences have basically the same phenomenological content—the pure consciousness. Theistic mysticism is just the imposition of theistic interpretation on this core mystical experience.

However, R. C. Zaehner, William Wainwright, Stephen Payne, and Nelson Pike vigorously defend the distinctiveness of theistic mysticism. They appeal to the phenomenological data of Christian mysticism: God and the soul are said to be close, or in mutual embrace. The "language is radically dualistic" (Pike 1992, p. 108). Furthermore, the same mystic sometimes offers a theistic description and sometimes a monistic description. They seem to reflect differences in the content of the experiences. Moreover, the phenomenon of "spiritual sensations" can hardly be explained as the imposition of the Christian tradition.

Pike also argues that even if the theistic mystic may experience a monistic interval, the meaning of this experience should be determined with respect to the phenomenological context, which is a series of dualistic experiences of God. So it is legitimate to think that during a "monistic" interval, the spirit is simply "deluded by love into not noticing the difference between itself and God" (p. 156).

DRUG-INDUCED MYSTICISM

Mysticism can be induced by drugs. This kind of chemical mysticism has been made popular by Aldous Huxley, and confirmed by some empirical studies (Tisdale 1980, chap. 15). However, its philosophical significance is unclear. Some regard the drug-induced alternative states of consciousness as gateways to extra-mundane reality. Others think that chemical mysticism demonstrates that reductive explanations of mysticism are available. Both interpretations can be resisted. On the one hand, the skeptics argue that we cannot distinguish alternative states of consciousness from hallucinations.

On the other hand, some scholars contend that it has not been really established that drugs are sufficient to produce genuine mystical experiences. The experimental evidence only suggests that it can raise the likelihood and enhance the intensity of the experiences (Davis 1989, p. 220; Heaney 1973, p. 116; Vergote 1997, pp. 197ff). Even if drugs are causally sufficient to produce mystical experiences, it does not follow that they are unveridical. God may have laid down some psychophysical laws to the effect that whenever certain brain states are produced, a certain perception of the divine would be produced. There is no reason why those brain states cannot be caused by taking drugs. It has been argued that as long as the whole process is set up and upheld by God, such perception of God should be counted as veridical.

In any case, even if drug-induced mystical experiences are unveridical, it does not follow that non-drug-induced mystical experiences are also unveridical. What is shown is that on the experiential level, mystical experience can be faked. This is neither surprising nor uniquely true of mystical experience. Sense experiences can also be faked.

NEURAL SCIENCES AND MYSTICISM

Eugene d'Aquili, Andrew Newberg, and Vince Rause (2001) have proposed a neurophysiological theory of mysticism. They explain mystical states as the effect of "deafferentation"—the cutting off of neural input into various structures of the nervous system. As a result, an experience of "absolute unitary being" occurs. In similar ways, the theory proposes explanations of a continuum of mystical experiences, both theistic and non-theistic.

The theory of d'Aquili and Newberg is by no means proven at this stage. Moreover, they point out that " tracing spiritual experience to neurological behavior does not disprove its realness … both spiritual experiences and experiences of a more ordinary material nature are made real to the mind in the very same way—through the processing powers of the brain and the cognitive functions of the mind" (Newberg, d'Aquili, and Rause 2001, p. 37).

They also ask, " Why should the human brain, which evolved for the very pragmatic purpose of helping us survive, possess such an apparently impractical talent?" (Newberg, d'Aquili, and Rause 2001, p. 123). They in fact tend to think their biology of transcendence is congenial to religion. The neurophysiological theory by itself does not disprove the mystical experiences, just as psychophysical laws governing sense experiences would not disprove those experiences (Jerome Gellman 2001, p. 99). Of course, there are deep questions about naturalistic explanations of mysticism that deserve further exploration (Wainwright 1973; Yandell 1993, chaps. 6–7).

See also Agnosticism; Aurobindo Ghose; Being; Buddhism; Eckhart, Meister; Islamic Philosophy; John of the Cross, St.; Logical Paradoxes; Mysticism, History of; Mysticism: The Indian Tradition; Nirvāṇa; Otto, Rudolf; Pantheism; Pseudo-Dionysius; Religious Experience, Argument for the Existence of God; Religious Language; Śankara; Stace, Walter Terence.

Bibliography

PRIMARY SOURCES

Boehme, Jakob. *Works,* 4 vols. London, 1764–1781. An English translation.

Eckhart, Meister. *Selected Treatises and Sermons.* Translated by J. M. Clark and J. V. Skinner. London: Faber and Faber, 1958.

John of the Cross, St. *The Ascent of Mount Carmel.* London, 1943.

Progoff, Ira, ed. *The Cloud of Unknowing.* New York: Julian Press, 1957. An anonymous fourteenth-century treatise on contemplative prayer.

Ruysbroeck, Jan van. *The Chastising of God's Children. …* Edited by J. Bazire and E. Colledge. Oxford, 1957.

Underhill, Evelyn, ed. *The Cloud of Unknowing.* London, 1912. Contains introduction.

GENERAL STUDIES

Butler, C. *Western Mysticism.* London: Constable, 1922.

Hügel, F. von. *Eternal Life.* Edinburgh: Clark, 1912.

Hügel, F. von. *The Mystical Element of Religion.* London: Dent, 1908.

Inge, W. R. *Christian Mysticism.* London: Methuen, 1899.

Inge, W. R. *The Philosophy of Plotinus,* 2 vols. London: Longman, 1918.

Otto, R. *Das Heilige.* Breslau: Trewendt and Granier, 1917. Translated by J. W. Harvey as *The Idea of the Holy.* London: Oxford University Press, 1923.

Otto, R. *Mysticism East and West.* Translated by B. L. Bracey and R. C. Payne. New York: Macmillan, 1932.

Underhill, Evelyn. *Mysticism.* London: Methuen, 1911. Contains a substantial bibliography.

PSYCHOLOGY OF MYSTICISM

James, William. *The Varieties of Religious Experience.* London: Longman, 1902. Lectures XVI–XVII.

Leuba, J. H. *The Psychology of Religious Mysticism.* London: K. Paul, Trench, Trubner, 1925.

ACCOUNTS OF MYSTICISM

Smart, Ninian. *Reasons and Faiths.* London: Routledge and Paul, 1958. Particularly valuable for its discussions of various non-Christian as well as Christian forms of mystical and numinous religion.

Stace, W. T. *Mysticism and Philosophy.* Philadelphia: Lippincott, 1960.

Zaehner, R. C. *Mysticism Sacred and Profane.* Oxford: Clarendon Press, 1957.

PHILOSOPHICAL DISCUSSIONS

Glasgow, W. D. "Knowledge of God." *Philosophy* 32 (1957): 229ff.

Horsburgh, H. J. N. "The Claims of Religious Experience." *Australasian Journal of Philosophy* 35 (1957): 186ff.

Kennick, W. E. Review of Stace's *Mysticism and Philosophy. Philosophical Review* 71 (1962): 387–390.

Martin, C. B. *Religious Belief.* Ithaca, NY: Cornell University Press, 1959.

Ronald W. Hepburn (1967)

MYSTICISM, NATURE AND ASSESSMENT OF [ADDENDUM]

Since the 1960s, philosophical controversies concerning the nature of mysticism mainly surround the relationship between mysticism and language, and the typology of mysticism. Moreover, as standard empiricist epistemologies no longer dominate the scene, new types of epistemology, which grant mystical experiences much more evidential force, have been formulated.

MYSTICISM AND LANGUAGE

Concerning the relationship between mysticism and language, some believe that mysticism transcends language, as reflected in the claim that mysticism is essentially ineffable. Taken literally, this claim generates many paradoxes, and Keith Yandell (1993, chaps. 3–5) has made sharp criticisms of various versions of the ineffability thesis (Alston 1992, Matilal 1992).

At the other end of the spectrum, Steven Katz claims that mystical experiences are largely constructed out of the language provided by the mystics's conceptual framework and practice. His work has been largely responsible for the contextualist turn in the study of mysticism in the 1980s (Katz 1978, 1983). This kind of mystical constructivism has been fiercely contested, especially by Robert Forman (1990, 1998, 1999). He argues for the universality of the "Pure Consciousness Event," which is a purely nonconceptual state of consciousness without any intentional object, and that mystical constructivism cannot adequately explain mysticism's unpredicted and novel nature. Jess Hollenback (1996) provides cases of paranormal mystical experiences that "shatter the recipient's previous expectations" (p. 15). William Wainwright (1981) contends that while mystical experiences are shaped to some extent by the mystics's traditions, it does not follow that those experiences are entirely determined or created by those traditions.

It seems hazardous to make universal statements about the relationship between mysticism and language.

THE ASSESSMENT OF MYSTICISM AND THE DEMISE OF FOUNDATIONALISM

Since the 1980s, there is a revival of the argument from mystical experience. Richard Swinburne (1979) defends the "Principle of Credulity," which says we should trust our experiences unless there are special considerations to the contrary. William Alston has defended the rationality of mystical perception by propounding his "doxastic practice" approach. By "doxastic practice" Alston means a system of belief-forming mechanisms. His *Perceiving God*(1991) is an impressive work which argues that it is practically rational to regard all socially established doxastic practices as prima facie reliable. It is important to note that Alston requires those doxastic practices to have a significant degree of self-support, and an internal overrider system.

Alston's sophisticated argument has attracted a lot of criticisms (Fales 2004). Space does not permit detailed discussions of the debate. It is important to appreciate the significance of Alston's work (together with Swinburne, Yandell, and Gellman) as a new research project in epistemology. They are not only reviving natural theology, but also proposing a new approach that navigates between strong foundationalism and postmodern relativism. They admit our epistemic base is fallible but they advocate an attitude of prima facie trust to replace Cartesian doubt. While "trust without infallible proof" was formerly treated as irrational, they suggest that the spirit of rationality should instead be construed as "trust until shown otherwise by criticisms."

They maintain the emphasis on experience but try to break loose of the straightjacket of traditional empiricism by broadening the evidential base of experience. The basic rationale is that in the end we need to adopt an attitude of *basic trust* (i.e., a trust that cannot be non-circularly justified) toward our perceptual experiences. It would be unfair to grant this kind of basic trust to sense experiences alone while adopting skepticism toward other kinds of perceptual experiences. In the end, the epistemic assessment of mysticism will probably depend on the ability of this radically new epistemology to withstand objections. The controversy is still raging.

See also Religious Experience.

Bibliography

Alston, William. "Literal and Nonliteral in Reports of Mystical Experience." In *Mysticism and Language*, edited by Steven T. Katz. New York: Oxford University Press, 1992.

Alston, William. *Perceiving God: The Epistemology of Religious Experience*. Ithaca, NY, and London: Cornell University Press, 1991.

Davis, Caroline. *The Evidential Force of Religious Experience*. Oxford: Clarendon Press, 1989.

Fales, Evan. "Do Mystics See God?" In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. VanArragon. Oxford: Blackwell, 2004.

Forman, Robert K. C., ed. *The Innate Capacity: Mysticism, Psychology and Philosophy*. New York: Oxford University Press, 1998.

Forman, Robert K. C. *Mysticism, Mind, Consciousness*. Albany: State University of New York Press, 1999.

Forman, Robert K. C., ed. *The Problem of Pure Consciousness: Mysticism and Philosophy*. New York: Oxford University Press, 1990.

Gellman, Jerome. *Mystical Experience of God: A Philosophical Inquiry*. Aldershot, U.K.: Ashgate, 2001.

Heaney, John J., ed. *Psyche and Spirit*. New York: Paulist Press, 1973.

Hollenback, Jess Byron. *Mysticism: Experience, Response, and Empowerment*. University Park: The Pennsylvania State University Press, 1996.

Huxley, Aldous. *The Doors of Perception and Heaven and Hell*. London: Flamingo, 1994.

Katz, Steven T., ed. *Mysticism and Language*. New York: Oxford University Press, 1992.

Katz, Steven T., ed. *Mysticism and Philosophical Analysis*. New York: Oxford University Press, 1978.

Katz, Steven T., ed. *Mysticism and Religious Traditions*. New York: Oxford University Press, 1983.

Matilal, Bimal Krishna. "Mysticism and Ineffability: Some Issues of Logic and Language." In *Mysticism and Language*, edited by Steven T. Katz. New York: Oxford University Press, 1992.

Newberg, Andrew, Eugene D'Aquili, and Vince Rause. *Why God Won't Go Away: Brain Science and the Biology of Belief*. New York: Ballantine Books, 2001.

Payne, Stephen. *John of the Cross and the Cognitive Value of Mysticism: An Analysis of Sanjuanist Teaching and its Philosophical Implications for Contemporary Discussions of Mystical Experience*. Dordrecht, Netherlands: Kluwer, 1990.

Pike, Nelson. *Mystic Union: An Essay in the Phenomenology of Mysticism*. Ithaca, NY: Cornell University Press, 1992.

Swinburne, Richard. *The Existence of God*. Oxford: Clarendon Press, 1979.

Tisdale, John R., ed. *Growing Edges in the Psychology of Religion*. Chicago: Nelson-Hall, 1980.

Vergote, Antoine. *Religion, Belief and Unbelief: A Psychological Study*. Amsterdam: Rodopi, 1997.

Wainwright, William. *Mysticism*. Brighton: The Harvester Press, 1981.

Wainwright, William. "Natural Explanations and Religious Experience." *Ratio* 15 (1973): 98–101.

Yandell, Keith E. *The Epistemology of Religious Experience*. Cambridge, U.K.: Cambridge University Press, 1993.

Kai-man Kwan (2005)

MYTH

The relation between philosophy and mythology can be usefully set out under three main headings. There is first the period in Greek philosophy when philosophers wanted to discard and to criticize mythological modes of thought but when they were still so close to those modes of thought that mythology recurred in philosophical contexts. Then in modern thought there is the period from Giambattista Vico to Auguste Comte, when mythology was taken seriously as a clue to the primitive history of thought, and from the nineteenth century on, when there was a variety of systematic attempts at a science of mythology. Finally, there is the role of myth in modern irrationalisms.

To this scheme three objections may be made. The first is that in discussing the Greeks what is said will inevitably be conditioned by the writer's beliefs about what modern scientific approaches to mythology have yielded. Thus, the second section should precede the first. To this objection everything can be conceded except the conclusion, for it would be equally difficult to discuss the growth of the science of mythology before anything had been said about mythology itself.

A second objection might be that no initial definition of mythology has been offered. But here the danger is that by delineating the field of mythology too sharply, one biases one's account in favor of one sort of theory. And any definition broad enough to escape this charge would be either vague or a mere catalog.

The third objection would be that the Christian era until the time of Vico appears to be neglected by this schematism. For this there is good reason, however. In that era mythologies were predominantly treated as false theological accounts, rivals to the one true theological account, the Christian.

GREEK PHILOSOPHY

Greek myths, like those of other Mediterranean and Near Eastern cultures, include cosmogonies and accounts of great discoveries and inventions, such as that of fire; of the founding of cities; and of the ancestry of kings, in which relationships between gods and men are codified. In different stages of the mythology, such as in the distinction between the Olympian gods and the dark, chthonic deities, one can distinguish different social origins. From the time of Émile Durkheim and Jane Harrison anthropologists have stressed the function of myths as explanations of rituals that express the social consciousness of a group. In Greek society the public ritual continued to express the life of the community long after belief in gods had become questionable.

Greek philosophy only gradually separated itself from mythology. Personification, for example, was common in pre-Socratic philosophy, but at the same time rationalist criticism of mythology originated with writers like Xenophanes, who attacked anthropomorphic representation of the gods, and Euhemerus, who argued that myths were to be explained as stories about men who had been deified. Heraclitus attacked Homer and Hesiod for their dependence on myth.

PLATO. Plato used myths and allegories for a variety of purposes. Perceval Frutiger draws a distinction between myths properly so called and allegories, which, for example, lack the element of story; among allegories he would include the account of the Cave in the *Republic* or the noble lie about precious and base metals in the souls of different types of men. He divides myths in the full sense into those that function as allegories, those that function as genetic explanations, and those that function as other types of parascientific explanations. An example of allegorical myth is Diotima's account of the birth of Eros in the *Symposium*; among genetic explanations is the account of the creation in the *Timaeus*; and typical examples of what Frutiger calls parascientific are the accounts of a future life and of rewards and punishments for virtue and vice given in the *Republic, Gorgias, Phaedo,* and *Phaedrus.* Frutiger sees three features of Platonic myth as outstanding: the use of symbols, the freedom exhibited in the handling of the narrative, and what he pleasantly calls a prudent imprecision. The last is important. Plato uses myth where he wishes the precise extent of his own intellectual commitment to remain unclear. Thus, Plato's use of myth helps us to understand how the break with mythological thought forms involves the raising of sharp questions about truth and falsity which the mythological forms themselves are able to evade. This throws light on certain characteristics of mythology.

The subject matter of mythological narratives is no different from that of later philosophy and science; what differentiates myth from these is not merely its narrative form or its use of personification. It is, rather, that a myth is living or dead, not true or false. You cannot refute a myth because as soon as you treat it as refutable, you do not treat it as a myth but as a hypothesis or history. Myths that could not easily coexist if they were hypotheses or histories, as, for example, rival accounts of creation, can comfortably belong to the same body of mythology. There are often gradual processes of reconciliation and of

integration into a single narrative, but the discrepancies that give so much pleasure to the anthropologist are not discrepancies at all from the standpoint of the narrator.

Thus Plato, by falling back into myth, may be deliberately avoiding too direct an encounter not only with certain philosophical difficulties but also with rival religious traditions. For myth is not theology any more than it is hypothesis or history. Indeed, the dominance of theology in later religious thought and the insistence in the mystery religions and in Christianity on treating myth as theology are as responsible for the death of mythology as is any philosophical rationalism bred by the pre-Socratics and Plato. Of course, it was not only Greek mythology that was treated by Christianity in this way. Both Norse and Celtic mythology met the same fate, although they both survived in medieval literature as beliefs and not just as a source for tale telling.

MODERN THOUGHT

The first serious modern treatment of mythology occurs in Vico's *Scienza nuova.* In Vico's theory of history each period has its own unity and character, and periods succeed one another in a determinate order. The beginnings of civilization occur in "the age of the gods," when men live in families and center their lives around religion, marriage, and the burial of the dead; this period is followed by the "age of heroes," in which aristocratic states arise. Only then comes the "age of men," the age of democratic republics. By the third stage rational inquiry is established, but in the early stages poetry and myth express the vulgar wisdom of a people. Only from mythology can we discover the religion, morals, law, and social life of early society. Myths are not false narratives, nor are they allegories. They express the collective mentality of a given age.

Vico's treatment of myth is far closer to that of modern anthropology than is that of his immediate successors. The Enlightenment's belief in progress and attack on superstition produced an unsympathetic climate for such interests. Even Johann Gottfried Herder, whose sympathy was awakened by seeing in primitive poetry and song the spirit of the folk, was inclined to treat myths as pardonably false beliefs. In the nineteenth century this assumption underlay the first systematic attempts at a science of mythology, but there was also a new consciousness of the widespread prevalence of mythology and a wish to apply comparative methods.

In 1856 F. Max Müller published his *Comparative Mythology,* in which he tried to interpret mythologies by means of principles derived from philology. All Aryan languages are derived from Sanskrit, in which originally there were certain words named sun, sky, clouds, rain, and dawn. But language became diseased, the original meanings were lost, the words became treated as the names of divine beings, and what had been accounts of the sun ushering in the dawn and ending the reign of night were transformed into myths about battles between gods, heroic quests for gold, and the like. To understand a myth, asserted Müller, discover the etymology of the names.

Andrew Lang pointed out that rival philologists would give different etymological explanations of the same myth with apparently equal plausibility. Lang himself regarded myths as survivals of earlier social norms. The classical Greeks recount myths in which cannibalism and human sacrifice occur, although they practiced neither; however, among Polynesian and African peoples, of whom Lang's contemporaries were newly aware, just such customs and accompanying myths are found. In classical Greece the custom had vanished, but the myth remained. Or a nature myth may be found with its meaning plain in its Maori form today, whereas in its Greek version the story has been so changed that the original meaning has been lost. The anthropology Lang and his school used was that of E. B. Tylor, who himself criticized Müller's theorizing by showing how convincingly the nursery rhyme "Sing a Song of Sixpence" could be explained as a solar myth in Müller's terms.

RECURRENT THEMES AND COMPARATIVE METHODS. Lang took it for granted that the "same" myth could turn up both in Greece and in New Zealand. The modern collection of mythologies has emphasized nothing so much as the strikingly similar themes and stories that recur in widely different places and times. Myths of the creation of the world are widespread; myths of the creation of humankind occur everywhere. But even in detail myths resemble one another. Clyde Kluckhohn has written that he knows of no culture lacking myths of witchcraft in which were-animals move about at night; poisons can be magically introduced into the victim, causing illness and death; and there is some connection between incest and witchcraft. Rank has discussed the common myth pattern of a hero, born of noble parents, against whose birth an oracle warns his father, so that the child is left to die of exposure; the child is saved by shepherds or animals, grows up to return, perform great deeds, avenge himself, and finally be recognized. In the Far East, among the Navajo, and in Greece, as well as in many other places, we find this pattern. What is the explanation of its recur-

rence? We can distinguish three main types of explanation.

The first is psychoanalytic. Otto Rank, a Freudian, explains the hero as the ego of the child who rebels against his parents. His father, on whom the child's hate is projected, is pictured as exposing the child in a box on water. The box symbolizes the womb; the water, birth. The order of the story follows a sequence analogous to that of dreams in which natural events and symbols are combined in a single fantasy. The myth is the expression of all paranoid characters who hate the father who ousted them from the maternal love and care. Because such a character is widespread, the myth that expresses it is widespread, too; in general, it is the common biological, and, consequently, psychological, inheritance of humankind that underlies the common stock of mythology.

By contrast, the Jungian approach to mythology rests upon belief in a common human access to the collective unconscious. The individual continually finds himself giving expression to an archetypal symbolism that dominates not only the mythology but also much of the sophisticated literature of the world. The same myths recur in different times and places because all mythology has a common source. Modern man, who has overdeveloped the rational side of his nature, encounters in his dreams the same figures that appear in ancient and primitive mythology.

The difficulties in the Jungian account of mythology are difficulties that confront all Jungian theory. If the existence of the collective unconscious is a hypothesis designed to explain the recurrence of certain themes and symbols in myths and dreams, then it must be formulable in a way that is testable. But if such a hypothesis is to be testable, we must be able to deduce from it predictable consequences over and above the data it was originally formulated to explain. Yet no such consequences seem to follow from the hypothesis of the collective unconscious. It seems to be untestable; it certainly remains untested. As an explanation of the recurrence of mythological themes and symbols, it is also unnecessary, for there are simpler and less incoherent explanations.

Joseph Campbell has used the Jungian theory of archetypes to interpret the story *The Frog King,* one of the myths collected by the brothers Grimm. He sees the frog as a small-scale dragon whose outward ugliness conceals the depths of the unconscious, in which unrecognized and unknown treasure is to be found. The frog king summons the child to attain maturity and self-knowledge by exploration of the unconscious. Fortunately, we also have a Freudian interpretation of *The Frog King* by Ernest Jones according to which the frog is a symbol for the penis and the myth represents the child's overcoming disgust in approaching the sexual act. Müller had, of course, long before interpreted *The Frog King* as one more solar myth.

In the face of these rival interpretations the need for a criterion of correct interpretation is clearly urgent, and with this need goes the need for a criterion for deciding when two myths are and are not versions of the "same" myth. The first step toward providing such criteria is the collection and tentative classification of as many bodies of mythology as possible. The most interesting work here has been done by Kluckhohn, who has systematically established not only the recurrence of plots and characters but also the existence of constant tendencies within this recurrence. For example, we can discover cases where a myth is reinterpreted to fit a new cultural or social situation. Clearly, where we can distinguish the original from the reinterpreted version, we are in a stronger position to compare a myth with similar myths for other cultures. We can study and compare not merely one version of a myth with another but the development of one myth through a series of versions with the development of another; from this it is clear that even if we wish to stress certain psychological functions of myth (Kluckhohn has thrown light on Navajo mythmaking by showing how it exemplifies mechanisms of ego defense), it is only when we put myth into a social context that we are likely to understand what the nature of mythmaking and recounting is.

ANTHROPOLOGY. The work of Claude Lévi-Strauss is important not only because its treatment of myth does not abstract myths from the social and economic relationships of those who tell and hear them but also because by invoking a wider context he has been able to pick out hitherto unnoticed features of mythology. In *Totemism,* for example, Lévi-Strauss shows how a myth of the North American Ojibwa and a myth from Polynesian Tikopia both express relationships between nature and culture, between the species that provide food and the kinship system. In each case the myth helps to express both continuity and discontinuity in these relationships; both myths also stress that no direct and simple connection between the one type of relationship and the other is possible. The myths, as it were, warn anthropologists not to oversimplify.

If one did not notice the connection of these myths with foodstuffs and with kinship but simply abstracted the "story," one would certainly not necessarily conclude

that the Ojibwa myth and the Tikopia myth were the same myth. The resemblances between them appear fully only because Lévi-Strauss poses certain questions about the myths. These questions are formulated in the light of his general theory of kinship systems and invoke the notion of relationships that are specified in purely formal terms. Lévi-Strauss elsewhere has analyzed other myths with a view to showing that in their structure formal properties are both exhibited and implicitly commented upon. Perhaps not surprisingly, these formal properties parallel the formal properties exhibited by kinship systems and also parallel to some extent, much more surprisingly, the formal properties of certain linguistic structures.

What emerges from these studies is the thesis that myths incorporate and exhibit binary oppositions that are present in the structure of the society in which the myth was born. In the myth these oppositions are reconciled and overcome. The function of the myth is to render intellectually and socially tolerable what would otherwise be experienced as incoherence. The myth is a form in which society both understands and misunderstands its own structure. Thus, Lévi-Strauss gives a precise meaning to Vico's contention that "The fables of the gods are true histories of customs."

This judgment is perhaps inverted in the work of Lévi-Strauss's most important rival, Mircea Eliade. The customs of men, in Eliade's view, often turn out to be the expression of their beliefs about the gods. Thus, the behavior of shamans, who in a state of trance imitate animal sounds (birds' song, for example, among many peoples) is a reenactment and an attempt to restore man's primitive, paradisal, unfallen state in which he not only did not die or have to work but also communicated with the animals and lived in peace with them. Hence, Eliade concludes both that shamanism is part of the central religious tradition of humankind, stretching from primitive African myths to Christian theology, and that it is therefore not, as it first appears to be, an irrational phenomenon. Eliade distinguishes sharply between the particular cultural and social trappings that may surround a myth and what he calls the ideology behind the trappings that is exhibited in the myth itself. Thus, where Lévi-Strauss analyzes the content of a myth in terms of what is local and particular to a given society, Eliade wishes to relate the content to general human religious interests and as far as possible divorce it from the local and particular.

IRRATIONALISM. "Myths must be judged as a means of acting upon the present," said Georges Sorel in 1908. Sorel distinguishes those beliefs that it is appropriate to characterize in terms of truth and falsity and those it is appropriate to characterize in terms of effectiveness and ineffectiveness. A myth is essentially a belief about the future that embodies the deepest inclinations of some particular social group. The myth that Sorel himself wanted to propagate was the syndicalist project of a general strike. Other socialists treated their beliefs about the future as predictions; Sorel regards this as for the most part irrelevant. The only predicates in which he is interested are self-fulfilling ones.

Yet to regard beliefs about the future in this way is paradoxical. For example, when I try to propagate a myth, I am inviting people to believe. But insofar as I do this, I invite them to treat it as true rather than false and as susceptible to truth or falsity. It is difficult to resist the conclusion that anyone who holds a view like Sorel's will fall into a form of doublethink, treating the myth as true or false in certain situations but retreating into the assertion that questions of its truth or falsity are inappropriate in other situations. Certainly, just this kind of doublethink characterizes modern irrationalist mythmakers after Sorel. They wish to avoid hard questions that philosophers or social scientists might raise about their myths, but they also wish to claim some kind of truth for their utterances. Thus, we also get a concomitant doctrine of special kinds of truth or special criteria for truth—for example, in works as different as Alfred Rosenberg's *Myth of the Twentieth Century* and D. H. Lawrence's *The Plumed Serpent.* Rosenberg's version of Houston Stewart Chamberlain's amalgam of anti-Semitism, racism, and authoritarian German nationalism is, of course, utterly different in content and implications from Lawrence's appeal to "the dark gods" and his attempt to restore an imagination violated by the wrong kind of arid rationalism. However, the difficulty with all irrationalism is that the abandonment of the criteria of rationality leaves us defenseless before the most morally outrageous appeals to emotion. In such appeals the revival of myth has a key place.

See also Chamberlain, Houston Stewart; Comte, Auguste; Durkheim, Émile; Freud, Sigmund; Functionalism in Sociology; Heraclitus of Ephesus; Herder, Johann Gottfried; Homer; Irrationalism; Jung, Carl Gustav; Philosophical Anthropology; Plato; Pre-Socratic Philosophy; Sorel, Georges; Vico, Giambattista; Xenophanes of Colophon.

Bibliography

Eliade, Mircea. *Aspects du mythe.* Paris: Gallimard, 1963.

Frutiger, Perceval. *Les mythes de Platon: Étude philosophique et littéraire.* Paris: F. Alcan, 1930.

Hooke, S. H., ed. *Myth, Ritual, and Kingship.* Oxford: Clarendon Press, 1958.

Jones, Ernest. *Essays in Applied Psychoanalysis, Vol. II, Essays in Folklore, Anthropology and Religion.* London, 1951.

Jung, C. G., and Carl Kerényi. *Essays on a Science of Mythology.* Translated by R. F. C. Hull. New York: Pantheon, 1949.

Lang, Andrew. *Custom and Myth.* London, 1901.

Lévi-Strauss, Claude. *Structural Anthropology.* Translated by Claire Jacobson and Brooke Grundfest. New York: Basic, 1963.

Lévi-Strauss, Claude. *Totemism.* Translated by R. Needham. Boston: Beacon Press, 1963.

Müller, F. Max. *Chips from a German Workshop.* New York: Scribners, 1871.

Murray, Henry A., ed. *Myth and Mythmaking.* New York: G. Braziller, 1960.

Rank, Otto. *The Myth of the Birth of the Hero.* Translated by F. Robbins and S. E. Jellife. New York: R. Brunner, 1952.

Sebeok, Thomas A., ed. *Myth: A Symposium.* Bloomington: Indiana University Press, 1955.

Stewart, J. A. *The Myths of Plato.* London: Macmillan, 1905.

Vico, Giambattista. *The New Science.* Translated by T. G. Bergin and M. H. Fisch. New York, 1948; paperback ed., 1961.

Alasdair MacIntyre (1967)

MYTH [ADDENDUM]

As Alasdair MacIntyre says, some philosophers have treated myth, disparagingly, as the opposite of logos, as a nonrational form of understanding the world that either has been or should be displaced by science and reason. Others have agreed that myth is the opposite of logos but have consequently valorized it as a fruit of the primordial mind, a product of an archaic form of experience or mystical consciousness that the modern scientific mind, to its detriment, has lost. There is then a range of philosophical views of the relative value of myth, but philosophers have largely agreed with Ernst Cassirer in seeing myth as a quintessential product of pretheoretical consciousness and therefore as a foil for the scientific mentality of modern European civilization. Since 1967, however, this assumption has been problematized. The concept of myth has been deconstructed, and this deconstruction represents a double obstacle for any philosopher who wants to see in myths truths about the human condition.

The first obstacle arises as scholars realize the extent to which mythical accounts of the origins of the cosmos, of the gods, or of a people have been intimately tied to the social and historical context in which they are told. Far from being the ahistorical products of the unconscious or whimsical flights of speculation—"the wonderful song of the soul's high adventure," to quote Joseph Campbell—myths have typically served to legitimate a particular social order. A clear example is the story of Purusha in the Rig Veda, a story that inscribes the divisions of the caste system as a cosmic reality rather than as a human and hence contingent arrangement. Myths are therefore partisan, not apolitical. In Bruce Lincoln's (1999) slogan myths are "ideology in narrative form." A culture will typically have more than one cosmogony, some mythical accounts of origins will seek to justify the status quo, and rival accounts will seek to undermine it. In short myths typically have a legitimating function, and this fact is concealed by traditional philosophical approaches that ignore the myths' social and historical roots.

The second obstacle arises as scholars realize the extent to which the category of myth reflects the interests of those who employ it. To identify a particular story as a myth—identifying it as the product, for example, of pretheoretical consciousness—has operated to illustrate the superiority of certain ways of thinking over other ways of thinking and, sometimes explicitly, the superiority of certain cultures over other cultures. Thus, one can see that the category of myth is ideological. From this perspective the traditional account of the emergence of *mythos* and its struggle with and eventual defeat by *logos* is itself a myth, that is, a partisan, legitimating story that modern European philosophers tell of their own origins. *Myth* is in this sense therefore a construction of the scholar: myths are not discovered, they are invented, and philosophers who claim to find in myths *la pensée sauvage* tell us more about their own worldviews than they tell us about *les sauvages*.

A few Continental philosophers, such as Cassirer and Hans Blumenberg, explored the idea that myths play a role in the development of consciousness, but Anglophone philosophers were not especially interested. Modern philosophers of religion (who one might think would have a natural interest in myths) have tended to focus on religious "beliefs" deracinated from the oral and literary contexts from which they were drawn. They have also tended to avoid the study of any religion that is not monotheistic. When deconstructive arguments like those mentioned earlier are added to this aversion to the concrete, the result has been that myths have been left for social scientists to study. At the beginning of the twenty-first century there is almost no philosophical work being done on myth.

A FUTURE FOR THE PHILOSOPHICAL STUDY OF MYTHS

But this result is not inevitable. Two observations may point to a future for the philosophical study of myths. In the first place, even when it is not a philosopher studying the myth, philosophy is still present, because answers to philosophical questions are always already embedded in the theories of myth. Social scientific approaches to myths are not philosophically neutral. They inevitably embody a particular set of normative assumptions about what is real and not real, knowable and not knowable, and good and not good, and in this way theories of myth carry certain metaphysical, epistemological, and axiological presuppositions. That the study of myths is unavoidably "philosophy-laden" is perhaps seen most easily when one looks at how the theorist answers questions about rationality, for every theory of myth assumes a judgment regarding what is and is not rational to say. For example, when the Victorian anthropologist Edward Tylor proposed that myths were rational insofar as they originated in observations about natural phenomena, his empiricism was showing.

In the second place, that some scholars pursue questions about the social or political dimensions of myths does not preclude others from asking philosophical questions about the existential, phenomenological, metaphysical, or ethical dimensions of those same narratives. That a story serves ideological ends does not rule out the possibility that it might also house truths about the human condition. To argue otherwise is to collapse the questions of provenance with those of truth, the genetic fallacy. And granting that philosophers' use of the term *myth* has itself been ideological, the solution is not simply to switch the focus of reflection away from the narratives and onto the way that philosophers construct categories, but to practice philosophy self-consciously, self-reflexively, and without naïveté. Philosophers who work on culture should therefore become comfortable with working with historians, anthropologists, and others who deal with the contexts in which the myths have their sense, but they need not abandon the idea that philosophy has its own contribution to make.

In short, then, a philosophical contribution to the study of myths, though now moribund, waits on an appreciation, first, of the ways in which philosophical issues are woven into the theories at work in the social sciences and, second, of the ways in which philosophers of religion or of culture might broaden their studies to include narratives. The fact is that communities often tell stories that explain how the different forms of existence were established; stories that sanction a particular interpretation of history; stories that identify paradigmatic forms of proper behavior. Such stories can provide models of the lived world and of how best to operate within it, and philosophers can analyze and evaluate the truth and the rationality of these models. It can be expected that such stories will typically have an ideological function, but coming to terms with the interpretive and explanatory work of social scientists should strengthen and not eliminate a philosophy of myths.

See also Cassirer, Ernest; Hermeticism; Logos; MacIntyre, Alasdair.

Bibliography

Barbour, Ian G. *Myths, Models, and Paradigms: A Comparative Study in Science and Religion*. New York: Harper and Row, 1974.

Blumenberg, Hans. *Work on Myth*. Translated by Robert M. Wallace. Cambridge, MA: MIT Press, 1985.

Cassirer, Ernst. *The Philosophy of Symbolic Forms. Vol. 2: Mythical Thinking*. Translated by Ralph Manheim. New Haven, CT: Yale University Press, 1955.

Lincoln, Bruce. *Theorizing Myth: Narrative, Ideology, and Scholarship*. Chicago: University of Chicago Press, 1999.

Schilbrack, Kevin, ed. *Thinking through Myths: Philosophical Perspectives*. London: Routledge, 2002.

Kevin Schilbrack (2005)

N

NĀGĀRJUNA

(c. 150–250 CE)

Nāgārjuna is the first and most important philosopher of the Mahāyāna Buddhist tradition. His work is fundamental to all Mahāyāna philosophy and is widely discussed in the subsequent Buddhist literature of India, Tibet, and East Asia. His work has also attracted considerable attention in Europe and North America.

LIFE AND CONTEXT

Canonical hagiographies of Nāgārjuna report that he was born a Brahman in South India, became a Buddhist monk, and later adviser to a king of the Sātavāhana dynasty. He is credited with retrieving the *Prajñāpāramitā sūtras* from the undersea world of the *nāgās* to whom, according to legend, the Buddha had entrusted them for safekeeping. Given that Nāgārjuna probably lived at about the time that some of these texts were composed, it is possible that he was associated with their composition or dissemination. Nāgārjuna's philosophical work is grounded in the views articulated in these sūtras, and he develops a thorough exposition and defense of the central doctrine they articulate—that all phenomena are empty of essence. While Nāgārjuna's philosophical program, including his interpretation of emptiness and his doctrine of the two truths, is in many respects highly original, it is also in other respects continuous with early Buddhist accounts of the impermanence, interdependence, and selflessness of the person and of phenomena (Vélez 2005).

While there is disagreement regarding Nāgārjuna's dates and regarding the area of India in which he lived, a confluence of evidence, including Kumārajīva's biography and Joseph Walser's [(2004)] (2005) analysis of the context of the composition of *Ratnāvalī* (Jeweled Garland of Advice to the King)indicates that Nāgārjuna probably lived in the late second and early third centuries in the lower Krishna River Valley. If this is correct, Nāgārjuna was writing at a time when the Mahāyāna was a nascent movement, and his texts provide both the philosophical foundations for that movement and polemical defense of its doctrinal probity.

MAJOR WORKS

A large number of works are attributed to Nāgārjuna, including not only the philosophical works noted here, but also hymns, devotional poetry, and letters to royal patrons, as well as tantric and alchemical texts. It is likely that these latter were composed by another figure of the

same name, and that at least some of the devotional material ascribed to Nāgārjuna was not composed by the author of the philosophical texts that constitute the core of his corpus. The core texts, which are almost certainly composed by the same author, are *Mūlamadhyamakakārikā* (*Fundamental Verses on the Middle Way*), *Íunyatāsaptati* (*Seventy Verses on Emptiness*), *Yukti∑a∑†ika* (*Sixty Verses of Reasoning*), *Vigrahavyāvartanī* (*Replies to Objections*), *Ratnāvalī* (*Jeweled Garland of Advice to the King*), and *Vaidalyasūtra* (*Devastating Discourse*).

Of these, *Mūlamadhyamakakārikā* is the most important. The text comprises four hundred forty verses organized by Candrakīrti (c. 600–650 CE), in his commentary *Prasannapadā* (*Lucid Exposition*) into twenty-seven chapters. Nāgārjuna addresses a wide range of fundamental Buddhist categories and phenomena, arguing that each of them lacks essence. The text is terse and is difficult to interpret without a commentary, often considering opposing positions from non-Mahāyāna Buddhist schools and refuting them. Nāgārjuna relies almost exclusively on reductio ad absurdum arguments, arguing that any account of the essence of a phenomenon, or any account according to which something exists permanently, substantially, or independently, collapses into absurdity. As a consequence, he argues, all phenomena exist only interdependently, impermanently, and conventionally. Most importantly, the text identifies two truths: an ultimate truth—the emptiness of phenomena of any essence or substance; and a conventional truth—the empirical reality and interdependence of things, and argues that these two truths are mutually implicative.

Vigrahavyāvartanī is a reply to objections to *Mūlamadhyamakakārikā*, principally those of Nyāya philosophers. The first half of the text develops a series of objections, each to the effect that the doctrine that all phenomena are empty is self-refuting, on epistemological, logical or metaphysical grounds. In the second half, Nāgārjuna confronts each of these objections, demonstrating that each rests on a misunderstanding of emptiness—taking emptiness to be not essencelessness, but nonexistence. When emptiness is understood as interdependence, he argues, not only are none of these objections sound, but the alternative each proposes collapses into absurdity. This text is accompanied by a detailed and closely argued autocommentary.

Íunyatāsaptati and *Yukti∑a∑†ika* are each detailed verse explorations of specific themes raised in *Mūlamadhyamakakārikā. Íunyatāsaptati* addresses the relationship between the ultimate emptiness of phenomena and their conventional existence, arguing that the emptiness of phenomena does not undermine, but instead underwrites, their empirical reality. *Yukti∑a∑†ika* explores the sense in which Nāgārjuna's position constitutes a middle path, and characterizes the extremes between which it is a midpoint. One extreme is that of reification—the view that anything that exists does so in virtue of having some essence, that things remain in existence over time, and that anything that exists can in principle exist independently; the other is the extreme of nihilism—the view that because there is no essence, because all phenomena are impermanent and independent, nothing really exists at all. These extremes, Nāgārjuna argues, share the erroneous view that to exist is to exist substantially, independently and continuously, and that once this view is rejected the moderate view that things exist conventionally, dependently and impermanently is the only coherent metaphysical position. *Vaidalyasūtra* is a refutation of the foundationalist Nyāya epistemology, arguing that none of the kinds of foundations that school proposes for knowledge is in fact appropriately self-justifying and that none of their ontological categories is in fact basic.

Each of these texts is written in a technical vocabulary, in an academic style and involves arguments intended to be read by scholars. Each focuses on issues in metaphysics and epistemology. *Ratnāvalī*, while a closely argued philosophical text, is different. It is aimed at a lay audience, and is addressed to a royal patron. While it surveys Madhyamaka metaphysics, it also addresses topics in ethics, political philosophy and statecraft. Indeed, it is probably the first scholarly text on Mahāyāna ethics and the only Mahāyāna text on political philosophy. In *Ratnāvalī* Nāgārjuna explicitly grounds the Mahāyāna ethic that takes compassion as its foundation in the doctrine of emptiness, and defends a theory of statecraft according to which the ruler's obligations include a wide range of social welfare programs. The text is also sectarian, arguing in favor of the legitimacy of the Mahāyāna at a time when this movement and its texts would have been marginal and controversial, and appealing to the king for support for the monasteries.

PHILOSOPHICAL CONTRIBUTIONS

Nāgārjuna extends certain fundamental Buddhist doctrines to develop the metaphysics and epistemology distinctive of Madhyamaka. Five ideas deserve special attention: (1) the doctrine that all phenomena, including emptiness, are empty; (2) the doctrine of the two truths and the account of their relation to one another; (3) the deployment of both positive and negative tetralemmas;

(4) the claim that madhyamaka is not a philosophical position on a par with others, in that it is not an account of the nature of reality, but a refusal of all such accounts; and (5) the attack on epistemological foundationalism.

Nāgārjuna argues that all phenomena are empty of essence, of independence, of substance, in virtue of the fact that essence, independent existence and substance are incoherent. He argues that emptiness is not another essence, but rather is the complete absence of anything that could be an essence. Emptiness itself is just as empty, in just the same sense, as anything else. The emptiness of phenomena is, for Nāgārjuna, the ultimate truth about things—the truth found when the analysis of a thing is complete; this amounts to the fact that things are impermanent, interdependent, and have merely conventional, nominal identity conditions, but no basic nature. The conventional truth about things is the truth about them delivered by our ordinary faculties when used appropriately. But this is just the fact that things are impermanent, interdependent, and have conventional identity conditions. Hence the two truths, according to Nāgārjuna, are, from an ontological point of view, identical. Ultimate truth is therefore not a separate reality; conventional truth is not a veil of illusion. Rather, they are two aspects of one reality.

Nāgārjuna makes extensive use of the Buddhist tetralemma—the partition of logical space into affirmation, negation, both affirmation and negation and neither affirmation nor negation. His deployment is distinctive in that he presents both positive and negative forms of the tetralemma. From the perspective of conventional truth he argues, on the one hand, that we can say that there is a self (conventionally); that there is no self (ultimately); that there both is (conventionally) and is not (ultimately) a self; and that there neither is (ultimately) nor is not (conventionally) a self. On the other hand, from the ultimate point of view none of these can be asserted, as from that point of view there is only emptiness, which cannot be grasped discursively as it is, because discursive thought always involves reification and the mediation by universals. Hence, from the ultimate view there is neither a self, nor not a self, nor both nor neither.

Nāgārjuna asserts that he rejects all views, and that Madhyamaka is not a view. This assertion is variously interpreted by subsequent commentators. Candrakīrti's reading is the most straightforward: Many metaphysical positions are views about the fundamental nature of reality. Metaphysical disagreements are predicated on the view that there is a fundamental nature of reality, and reflect divergent views of what that nature is. Madhyamaka, Nāgārjuna argues, is the rejection of the coherence of the idea of a fundamental nature of reality. Hence it is not a metaphysical view in the sense that its rivals are.

According to many Indian philosophers, there are foundations of knowledge. Some argue that these are objects of knowledge; others that they are our means of gaining knowledge, such as perception or inference. Nāgārjuna argues that neither of these positions can be maintained: that objects of knowledge are only known in virtue of the employment of warranted means of obtaining knowledge, and that in turn these warranted means are only validated by the objects they deliver. Knowledge, such as the reality toward which it is directed, is hence groundless, interdependent, and conventionally constituted.

CANONICAL COMMENTARIES

Mūlamadhyamakakārikā is the subject of many commentaries in India, China and Tibet. The earliest is the *Akutobhayā*, whose authorship is not known. Some traditions regard it as Nāgārjuna's autocommentary, but because it cites the work of his immediate disciple óryadeva casts doubt on this attribution. Pingala's commentary (c. fourth century) exists only in a Chinese translation. Buddhapālita (fifth to sixth centuries) composed an important commentary, the *Buddhapālita*. Bhāvaviveka (sixth century) composed an extensive commentary *Prajñāpradipa* (*Lamp of Wisdom*) and subcommentary *Tarkajvāla* (*Blaze of Argument*). Bhāvaviveka offers extensive reconstructions of Nāgārjuna's arguments in line with the developments in Nyāya and Buddhist logic and takes issue with Buddhapālita's interpretation of the role of reductio argument in Madhyamaka methodology. Candrakīrtii (seventh century) in *Prasannapadā* defends Buddhapālita's reading against Bhāvaviveka's critique. His distinction between their respective understandings of Nāgārjuna's methodology and his account of the metaphysical implications of those understandings form the basis for the Tibetan distinction between the *svātantrika* (Tib: *rang rgyud pa*) and *prāsaṅgika* (Tib: *thal ʿgyur ba*) schools of Indian *madhyamaka* that has come subsequently to systematize much understanding of the diverse developments of Nāgārjuna's philosophy in India and Tibet. Candrakīrti also composed an extensive commentary on *Yukti∑a∑†ika*. Many commentaries on *Mūlamadhyamakakārikā* were composed in Tibet. The most extensive and influential is Tsong khapa's *rTsa she tik chen rigs p'ai rgya mtsho* (*Ocean of Reasoning: An Extensive Commentary on Mūlamadhyamakakārikā*) which com-

pares the Indian commentaries, defending Candrakīrti's reading.

TWENTIETH-CENTURY SCHOLARSHIP

The diversity of Western readings of Nāgārjuna's philosophical program is even greater than the diversity of Asian readings. Andrew Tuck (1994) notes that readings of Nāgārjuna in the West often follow fashions in Western philosophy and religious studies. He has been read as a mystic (Streng 1967), as a nihilist (Wood 1994), as a pragmatist (Kalupahana 1986), as an antirealist (Siderits 1988) and as a skeptic (Garfield 1995). There is also considerable debate concerning the degree to which Nāgārjuna argues cogently, and regarding whether his logic should be understood as akin to a European bivalent classical logic or as akin to a four-valued or paraconsistent logic (Robinson 1957, Hayes 1994, Garfield and Priest 2003).

See also Buddhism; Buddhism Schools: Madhyamika.

Bibliography

Battacharya, K., trans. *The Dialectical Method of Nāgārjuna*, edited by E. H. Johnston and A. Kunst. Delhi: Motilal Banarsidass, 1978.

Dunne, J., and S. McClintock. *The Precious Garland: An Epistle to a King—A Translation of Nagarjuna's Text from Sanskrit and Tibetan*. Boston: Wisdom Publications, 1997.

Garfield, J. *Fundamental Wisdom of the Middle Way: Nāgārjuna's Mūlamadhyamakakārikā*. New York: Oxford University Press, 1995.

Garfield, J., and G. Priest. "Nagarjuna and the Limits of Thought." *Philosophy East and West* 53 (2003): 1–21.

Hayes, R. "Nāgārjuna's Appeal." *Journal of Indian Philosophy* 22 (1994): 299–378.

Inada, K. *Nāgārjuna: A Translation of his Mūlamadhyamakakārikā with an Introductory Essay*. Tokyo: Hokuseido Press, 1993.

Kalupahana, D. *Nāgārjuna: The Philosophy of the Middle Way*. Albany: State University of New York Press, 1986.

Lindtner, C. *Master of Wisdom: Writings of the Buddhist Master Nāgārjuna*. Oakland, CA: Dharma Press, 1986.

Robinson, R. "Some Logical Aspects of Nāgārjuna's System." *Philosophy East and West* 6 (1957): 291–308.

Ruegg, D. S. *The Literature of the Madhyamaka School of Philosophy in India*. Wiesbaden, Germany: Harrassowitz, 1981.

Siderits, M. "Nāgārjuna as an Anti-Realist." *Journal of Indian Philosophy* 16 (1988): 311–325.

Sprung, M., and T. R. V. Murti. *Lucid Exposition of the Middle Way*. Boulder: Prajña Press, 1979.

Streng, F. *Emptiness: A Study in Religious Meaning*. Nashville: Abingdon Press, 1967.

Tsong khapa. *Ocean of Reasoning: A Great Commentary on Nagarjuna's Mūlamadhyamakakārikā*. Translated by G. N. Samten and J. L. Garfield. New York: Oxford University Press, [2005] 2006.

Tuck, A. *Comparative Philosophy and the Philosophy of Scholarship: On the Western Interpretation of Nāgārjuna*. New York: Oxford University Press, 1990.

Veléz, A. "Emptiness in the Pali Suttas and the Question of Nāgārjuna's Orthdoxy." *Philosophy East and West* 55 (4) (2005).

Walser, J. *Nāgārjuna in Context: Mahāyāna Buddhism and Early Indian Culture*. New York: Columbia University Press, 2005.

Wood, T. *Nāgārjunian Disputations: A Philosophical Journey Through an Indian Looking-Glass*. Honolulu: University of Hawaii Press, 1994.

Jay L Garfield (2005)

NAGEL, ERNEST

(1901–1985)

Ernest Nagel, the American philosopher of science, was born at Nove Mesto, Czechoslovakia and came to the United States at the age of ten, becoming naturalized in 1919. He was graduated from City College in 1923 and received an MA in mathematics from Columbia in 1925 and a PhD in philosophy in 1930. He served as the John Dewey professor of philosophy at Columbia University from 1955 to 1966, then took the position of university professor there until 1970, becoming emeritus in 1970. He expressed indebtedness to the teachings of Morris R. Cohen, John Dewey, and Frederick J. E. Woodbridge and to the writings of Charles S. Peirce, Bertrand Russell, and George Santayana.

PHILOSOPHY OF SCIENCE

Nagel belonged to the naturalist and logical empiricist movements, and he is primarily noted for his contributions to the philosophy of science. In 1934 he published, with Morris R. Cohen, *An Introduction to Logic and Scientific Method*. This noted text has been praised for its high level of rigor and for its enrichment of the traditional dry fare of logic with illustrations of the functions of logical principles in scientific method, in the natural and social sciences, and in law and history.

Nagel's book *The Structure of Science* is a unified and comprehensive distillation of many years of teaching and of his many publications on special aspects of scientific thought. It is the most complete exposition of Nagel's analysis of the nature of explanation, the logic of scientific inquiry, and the logical structure of the organization of scientific knowledge, and it illuminates the cardinal issues concerning the formation and the assessment of

explanation in physics and in the biological and social sciences.

Two other contributions by Nagel to logic and the philosophy of science are *Principles of the Theory of Probability* (1939) and *Gödel's Proof* (1958), written in collaboration with James R. Newman. These studies range over many issues, from the logic of probable inference to the basic conditions of the structure of formal systems.

GENERAL PHILOSOPHY

Two philosophical essays of a general scope by Nagel have been widely acclaimed. In "Logic without Ontology" Nagel defended a naturalistic interpretation of logic. He argued that logico-mathematical principles must be understood according to their functions in specific contexts, namely, in inquiries, and he criticized attempts to adduce an ontological ground or transcendent authority for the meaning, warrant, and necessary character of logical laws. Nagel had already repudiated his early view that logical principles "are inherently applicable because they are concerned with ontological traits of utmost generality" (*An Introduction to Logic and Scientific Method*, p. v). In "Logic without Ontology" he showed that the view that logic is ontologically determined or entails ontological commitments arises primarily from a failure to heed certain contextual and operational qualifications of the sense in which logical principles are supposed to possess "necessary truth."

In "Sovereign Reason" Nagel presented a penetrating critique, focused on the doctrine of internal relations, of Brand Blanshard's rational idealism. This critique exemplifies one of Nagel's strongest philosophical convictions and a main theme of "Logic without Ontology": Logical principles (and even pure Reason), just because they are analytic, are necessary but not sufficient instruments for acquiring knowledge or discovering truths about reality. The task of logic, according to Nagel, is to disclose the assumptions and clarify the methods on which responsible claims to knowledge are based and by which they are critically assessed. All claims to knowledge, even those most impressively supported by evidence and experiment, are subject to revision or rejection in the light of new advances in knowledge. This empiricist tenet led Nagel to accept contingency as a real trait of nature and fallibility as an inescapable feature of human inquiry.

SCIENCE AND SOCIETY

Nagel's technical interest in the logic and history of scientific knowledge did not prevent him from appreciating the social consequences and problems of science and technology in a democratic society. Much of his critical activity as a speaker, reviewer, and essayist was devoted to imparting a clearer understanding of the nature of science and to dispelling philosophical vagaries and bizarre notions concerning such matters as causality and indeterminism in physics; the alleged paradoxical character of abstract science or its utter disparity with common sense; the frequent claims that science is value-free, or metaphysically inspired, or mere codified sense data; and the revulsion or despair and the impassioned remedies that science has occasioned in some literary and theological circles.

MATERIALISM, DETERMINISM, AND ATHEISM

Nagel's philosophical naturalism led him to take a decisive stand on certain broad philosophical issues, notably materialism, determinism, and atheism. It has been charged that naturalists, being materialists, are unable to account for mental phenomena. Nagel replied, fully aware of the many senses of the word *materialism*, that naturalists are not materialists if materialism is taken to mean that such psychological predicates as "fear" or "feeling of beauty" logically entail or are reducible to physical terms such as *weight, length*, or *molecule*. Although he repudiated reductive materialism, Nagel held that mental events are aspects of and contingent on the organization of human bodies. Events, qualities, and processes are dependent on the organization of spatially and temporally located bodies. In this sense, naturalism is committed to materialism: Organized matter has a causal primacy in the order of nature. It follows that there can be no occult forces or disembodied spirits directing natural events and no personal immortality when bodily organizations disintegrate.

To assess the role of determinism in history and in ethical theory, Nagel formulated the meaning of *determinism* in natural science. A scientific theory is deterministic with respect to a set of properties when, given a specification of the set at any initial time, a unique set of the properties for any other time can be deduced by means of the theory. The theory might be a mechanical theory, and the sets of properties mechanical states. This theory might conceivably be of use in calculating the mechanical states of a human organism, but only its mechanical states. Whether other properties of the organism and its history were deterministic would remain an open empirical question. Nor would determinism in human history, if it were established, automatically empty moral endeavor and responsibility of significance. Which

modes of human experience and behavior, if any, are subject to deterministic theory remains an empirical question; and the sense in which these conditions might be characterized as "deterministic" remains an issue of analysis.

In several places, including his influential paper "The Causal Character of Modern Physical Theory," Nagel concerned himself with the philosophical implications of quantum theory. Like Albert Einstein and Max Planck, but unlike the majority of writers on the subject, Nagel denied that quantum theory has indeterministic consequences. He also showed in some detail how intellectual confusion thrives when distinctions of context and the relevance of theoretical language to specific contexts are ignored; for example, when "particle" in the context of Newtonian theory is transported into discussions of the uncertainty principle in modern physics. In another well-known essay, "Russell's Philosophy of Science," Nagel argued that the physical and physiological facts of perception do not require the abandonment of common sense in favor of the strange conclusions held by Russell and Arthur Stanley Eddington.

Nagel was one of the few naturalists to present a forthright statement of the naturalist critique of theism. His formulation of atheism is not couched as a sheer negation of theism but proceeds from a positive moral position according to which, while it is granted that there are inevitable tragic aspects of life, knowledge of life and nature is to be preferred to illusions. On matters of such supreme moment, the truth rather than fiction is the more fitting ideal of rational men.

Nagel did not, however, deny the value and authenticity of other than purely cognitive pursuits. He never argued that aesthetic qualities, ideals, suffering, and enjoyments are not genuine aspects of experience. On the contrary, he urged that naturalism, although obliged to render a competent account of scientific knowledge, also include in its scope a place for imagination, liberal values, and human wisdom.

See also Atheism; Determinism, A Historical Survey; Materialism; Philosophy of Science, History of.

Bibliography

WORKS BY NAGEL

On the Logic Measurement. New York, 1930.

An Introduction to Logic and Scientific Method. New York: Harcourt Brace, 1934. Written with Morris R. Cohen.

Principles of the Theory of Probability. Chicago: University of Chicago Press, 1939. Vol. I, No. 6, of *The International Encyclopedia of Unified Science*.

"Russell's Philosophy of Science." In *The Philosophy of Bertrand Russell*, edited by P. A. Schilpp. Evanston, IL: Open Court, 1944. Reprinted in *Sovereign Reason*.

"Logic without Ontology." In *Naturalism and the Human Spirit*, edited by Y. H. Krikorian. New York: Columbia University Press, 1944. Reprinted in *Logic without Metaphysics* (see below).

"Are Naturalists Materialists?" *Journal of Philosophy* 42 (1945): 515–530. Reprinted in *Logic without Metaphysics*.

"Sovereign Reason." In *Freedom and Experience. Essays Presented to Horace M. Kallen*, edited by Sidney Hook and Milton R. Konvitz. Ithaca, NY: Cornell University Press, 1947. Reprinted in *Sovereign Reason*.

"The Causal Character of Modern Physical Theory." In *Freedom and Reason*, edited by S. W. Baron, Ernest Nagel, and K. S. Pinson. Glencoe, IL: Free Press, 1951. Reprinted with revisions as part of Ch. 10 in *The Structure of Science*.

Sovereign Reason. Glencoe, IL: Free Press, 1954.

Logic without Metaphysics. Glencoe, IL: Free Press, 1956.

Gödel's Proof. New York: New York University Press, 1958. Written with James R. Newman.

"A Defense of Atheism." In *Basic Beliefs*, edited by J. E. Fairchild. New York: Sheridan House, 1959.

The Structure of Science. New York: Harcourt Brace, 1961.

Philosophy, Science, and Method: Essays in Honor of Ernest Nagel, edited by Sidney Morgenbesser, Patrick Suppes, and Morton Gabriel White. New York: St. Martin's, 1969.

With Sylvain Bromberger and Adolf Grünbaum, Adolf. *Observation and Theory in Science*. Baltimore, MD: Johns Hopkins, 1971.

Teleology Revisited and Other Essays in the Philosophy and History of Science. New York: Columbia University Press, 1979.

H. S. Thayer (1967)
Bibliography updated by Michael J. Farmer (2005)

NAGEL, THOMAS
(1937–)

Thomas Nagel has contributed to a wide spectrum of philosophical topics in ethical theory, moral psychology, applied ethics, and political theory, as well as to metaphysics and epistemology. His work is distinguished by its breadth, clarity, and acumen.

While there is not a single, narrowly defined theme running through all his work, Nagel has persistently engaged the problem of reconciling an objective view of reality with one's subjective, individual experience as a person. In his magisterial work, *The View from Nowhere*, Nagel writes: "This book is about a single problem: how to combine the perspective of a particular person inside the world with an objective view of that same world, the

person and his viewpoint included. It is a problem that faces every creature with the impulse and the capacity to transcend its particular point of view and to conceive of the world as a whole" (1986, p. 3). Nagel's defense of the legitimacy of both one's subjective perspective and an objective, nonindividual point of view, has been part of Nagel's resistance to philosophies that do away with either. So, in several books and many articles, Nagel has authored an influential critique of forms of physicalism that eliminate or do not take seriously the reality of subjective experience, and he has also been highly critical of philosophies that give way to skepticism because they grant excessive authority to subjectivity.

The Possibility of Altruism, his first book, argues that in an individual's recognition of goods and ills for him- or herself over time, there is an implicit recognition of the goods and ills that face other individuals. "In accepting goals or reasons myself I attach objective value to certain circumstances, not just value for myself" (1970, p. 85). In later work, Nagel refines the conviction that ethical and political theory needs to be comprehensively impartial and only comprised of agent-neutral reasons; these reasons are comprised of "what everyone ought to value, independently of its relation to himself" (1991, p. 40). Nagel allows that there is some tension between such an agent–neutral perspective and some of the values that have their place in specific, personal contexts. Nagel advocates an egalitarian social ideal (1991), while also recognizing that some goods are private and should be concealed from public surveillance and control (2002). Nagel's concern for the integrity of the individual pits him against overriding social engineering.

In philosophy of mind, Nagel is widely known for his essay "What Is It Like to Be a Bat?" (first published in *Philosophical Review* 1974, pp. 435–450, reprinted in *Mortal Questions* and widely anthologized). In this essay, Nagel identifies subjective, phenomenal experience as the central problem facing contemporary physicalism. He contends that a fully developed neurobiological, functional, materialist account of the human body would still leave out subjective experience (what it is like experientially to be conscious and undergo experiences), just as a fully developed neurobiological, functional, materialist account of a bat would still leave out what it is like to be a bat. In *What Does It All Mean?* Nagel employs the thought experiment of an inverted spectrum and other inverted sensations to exhibit the apparent contingency of the relationship between conscious, experiential states and functionalist, materialist ones. These are cases when the physicalist account of seeing some color or experiencing some taste is inverted, so that while the physicalist would conclude that one is having some taste, when it turns out one is having a quite different one. In *The View from Nowhere*, *Other Minds*, and elsewhere, Nagel opposes all philosophies of mind that fail to recognize the reality of subjective, lived experience.

Although Nagel's defense of the reality of phenomenal experiences and the apparent contingency of the mental-physical relation has seemed to some to lend credence to at least a modified form of dualism, Nagel himself holds that dualism can be avoided by developing a conceptual revision of one's current concept of the physical world and subjective experience. While philosophers do not yet possess this new world view, Nagel urges that future philosophical work be focused on conceiving of a single natural world that incorporates what one now sees as objective physical states and one's internal, mental subjectivity.

In his short book, *The Last Word*, Nagel offers an impassioned defense of reason as a reliable mode of inquiry, not subject to the objections of relativists, postmodernists, or contemporary pragmatists like Richard Rorty:

> Reason … can serve as a court of appeal not only against the received opinions and habits of our community but also against the peculiarities of our personal perspective. It is something each individual can find within himself, but at the same time it has universal authority. Reason provides, mysteriously, a way of distancing oneself from common opinion and received practices. … Whoever appeals to reason purports to discover a source of authority within himself that is not merely personal, or societal, but universal—and that should also persuade others who are willing to listen to it.
>
> (1997, pp. 2–3)

Nagel acknowledges the many ways in which one's reasoning may be impaired, but he nonetheless maintains the necessity of making recourse to reason in order to correct, however gradually, such impairments.

Nagel received a BA from Cornell University in 1958, a PhB from Oxford in 1960, and his PhD from Harvard in 1963. He has held academic appointments at the University of California at Berkeley, Princeton University, and New York University where he was appointed as University Professor in 2002. In addition to his specialized philosophical writing, Nagel has written on practical political and moral problems. For example, he has argued for a

highly restricted account of when and how a just war may be engaged.

See also Applied Ethics; Consciousness; Metaethics; Moral Psychology; Physicalism.

Bibliography

WORKS BY THOMAS NAGEL

The Possibility of Altruism. New York: Oxford University Press, 1970.

Mortal Questions. Cambridge, U.K.: Cambridge University Press, 1979.

The View from Nowhere. New York: Oxford University Press, 1986.

What Does It All Mean? New York: Oxford University Press. 1987.

Other Minds. New York: Oxford University Press, 1995.

The Last Word. New York: Oxford University Press, 1997.

Concealment and Exposure. New York: Oxford University Press, 2002.

The Myth of Ownership: Taxes and Justice. With Liam Murphy. New York: Oxford University Press, 2002.

WORKS ABOUT THOMAS NAGEL

Darwell, S. “The View from Nowhere.” *Ethics* 98 (1987): 137–157.

McGinn, C. “The View from Nowhere.” *Mind* 96 (1987): 263–272.

Charles Taliaferro (2005)

NAIGEON, JACQUES-ANDRÉ
(1738–1810)

Jacques-André Naigeon, a French writer, was an associate of Denis Diderot. Naigeon was not an original thinker; he became an editor, compiler, and commentator after having tried painting and sculpture, but he considered himself a philosopher and was proud of his classical erudition. A bibliophile, too, he accumulated one of the great collections of Greek and Latin classics of his time. Having been accepted into the group of Encyclopedists surrounding Baron d'Holbach, he became an aggressive atheist. He attached himself to Diderot as a disciple and tried to imitate his tone, his manner, and his ideas. Diderot in turn enjoyed Naigeon's wit and tolerated his bad temper, stiffness, and pedantry; Naigeon helped Diderot with the *salons* and the *Encyclopédie.* Naigeon later persuaded Diderot to make him his literary executor. He preserved and edited many of Diderot's manuscripts but did not publish others. He put out an incomplete edition of Diderot's works in 1798 and wrote a valuable but unfinished commentary on his life and writings, *Mémoires historiques et philosophiques sur la vie et les ouvrages de Diderot* (Paris, 1821). He also arranged the clandestine printing of several of Holbach's works in the Netherlands, and in 1770 published *Mélange de pièces sur la religion et la morale,* which contained some minor pieces by Holbach and other writers.

Naigeon edited the works of Seneca, completing the translation begun by N. La Grange and adding notes; he published it with Diderot's defense of Seneca, *Essai sur les régnes de Claude et de Néron* (Paris, 1778). A one-act musical comedy, *Les Chinois* (1756), is sometimes attributed to him, perhaps in collaboration with Charles-Nicolas Favart. His only “original” work was *Le militaire philosophe, ou Difficultés sur la religion, proposées au P. Mallebranche* (London and Amsterdam, 1768), which is based on an earlier anonymous manuscript and has a final chapter by Holbach. This dull work is of minor value as an example of dogmatic atheism and materialism, but it merely repeats the same ideas and arguments that had run throughout the radical writings of the entire century. Naigeon supports hatred of priests and the church with the doctrine of materialism and a naturalistic utilitarian morality. He denounces Christian ethics (asceticism, humility, etc.), demanding fulfillment of legitimate natural demands and a moral code based on social well-being. He points out contradictions in Christian ethics and doctrine, stressing its cruelty and its failure. He argues that Christian ethics leads to an inversion of the natural order of values, hence to intolerance, inhumanity, and crimes. Earth would be peaceful and happy if the idea of God were eliminated.

Naigeon continued this attack in his contributions to C. J. Panckoucke's *Encyclopédie méthodique.* This work consisted of separate dictionaries, and Naigeon edited the *Dictionnaire de la philosophie ancienne et moderne* (3 vols., Paris, 1791–1793), which was largely a compilation. In *Adresse à l'Assemblée nationale sur la liberté des opinions* (1790) he demanded absolute freedom of the press and again gave vent to his hatred of priests.

There are no studies on Naigeon, except in relation to his publication of Diderot's manuscripts, nor is any needed.

See also Diderot, Denis; Encyclopédie; Holbach, Paul-Henri Thiry, Baron d'; Seneca, Lucius Annaeus.

L. G. Crocker (1967)

NAIVE REALISM

See *Realism*

NAKAE TŌJU

(1608–1648)

Nakae Tōju, "the sage of Ōmi" (his native town in Shiga prefecture), the most respected Confucianist in the Tokugawa era, was an advocate of the Wang Yangming school. The ideas of Wang Yangming (in Japanese, Ōyōmei) were made known in Japan by the Zhu Xi scholar Fujiwara Seika (1561–1619), but only with Nakae did the Wang Yangming doctrine become a school of thought. The importance of this school lies in its impact on Japanese thinking and the nonconformists it produced. Its stress on *ryōchi* (literally, "good conscience"; more exactly, the innate knowledge that every man has from Heaven) favored the formation of strong individualists guided by the inner light of conscience without the formalistic restraints of Zhu Xi Confucianism. The cultivation of the mind combined with a stress on deeds rather than formal learning was another aspect of Nakae's teaching. His upright character showed in practice what it meant to be a Confucian sage, that is, almost a saint.

Nakae's intuitive and practical morality centering on filial piety had a great attraction for his pupils as well as for many later followers who for different reasons claimed him as their master. His outstanding followers were Kumazawa Banzan (1619–1691) and such men prominent in the nineteenth-century movement to restore the emperor as Ōshio Heihachirō, Yoshida Shōin, and Saigō Takamori. Kumazawa tried to persuade his master to leave the obscure village of Ogawa and enter the service of the lord of Okayama, but the humble Nakae shunned the proposal. In addition, Nakae's inclinations were ethico-religious rather than politico-economic, the characteristic of many of his followers. Nor was he a radical, although some of his admirers were.

Nakae strove for a middle way, mildly criticizing other points of view. He spoke of *ri,* the "principle," and *ki,* Zhu Xi's material force (which Nakae interpreted as matter-life), as two aspects of the "supreme ultimate." Nakae's terminology recalls the ancient Chinese sages and suggests Christian influence; *Jōtei,* the "Supreme Lord Above," he called "the absolute truth and the absolute spirit," and he ascribed almost personal attributes to this Being. Nakae also had pantheistic leanings, however, and he used anthropomorphic expressions to ally his *Jōtei* with Shinto deities. His moral ideas, though, are much more important than his cosmological views. Filial piety (*kō*) is the pivotal virtue, for him both the universe's moral power and its reason for being. Everyone, from the emperor to the most despised woman—Nakae being quite an equalitarian—was affected by filial piety, the creative force descending by degrees from Heaven. This virtue became in his late followers patriotism toward the emperor. Still, for Nakae, it was a cosmic and religious force not limited to one family or nation.

See also Japanese Philosophy; Kumazawa Banzan; Pantheism; Wang Yangming; Zhu Xi (Chu Hsi).

Bibliography

See *Tōju sensei zenshū* (The complete works of Nakae Tōju), 5 vols. (Tokyo: Iwanami Shoten, Shōwa, 1940). A secondary source in Japanese is Bitō Masahide, *Nihon hōken shisōshi kenkyū* (Studies on the history of feudal thought in Japan; Tokyo, 1961), pp. 136–216. See also G. M. Fisher, "The Life and Teaching of Nakae Toju," in *Transactions of the Asiatic Society of Japan* 36, Part 1 (1908): 24–94; W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 378–384, which offers selections, with an introduction.

Gino K. Piovesana, S.J. (1967)

NAMES, PROPER

See *Proper Names and Descriptions*

NAṢĪR AL-DĪN AL-ṬŪSĪ

(1201–1274)

Naṣīr al-Dīn al-Ṭūsī (1201–1274) is a Shiʿa Iranian author of some two hundred treatises in a number of disciplines, including philosophy, mathematics, astronomy, mysticism, and theology.

LIFE AND TIMES

Naṣīr al-Dīn was born in the city of Tus in the province of Khurasan in northwestern Iran, the first area to be devastated by the Mongolian invasion of the Middle East by Helagu Khan (1217–1265), grandson of Genghis Khan (1167?–1227). After completing his formal studies, al-Ṭūsī carried out research and publications under the patronage of various Ismaili rulers from 1227 until 1256, when he assisted the Ismaili ruler to surrender to Helegu

Khan, who employed al-Ṭūsī as his adviser until Helagu's death and then joined Abaq (1265–1282) until his own death. Al-Ṭūsī accompanied Helagu Khan in the Monogol attack on the last Sunni caliph in Baghdad, after which he built an observatory at Maragha in Azarbayijan in northwest Iran. There he spent the rest of his life in supervising innovations in astronomy and mathematics; in addition, he attracted the patronage of the Mongol ruler toward scientists, Shiʿa theologians, and writers on mysticism.

COSMOGONY AND ITS ETHICS

In formulating his views on the existence of God, al-Ṭūsī appeals to the Avicennan doctrine that God has no (external) cause; because entities are known by their causes, there cannot be any affirmative scientific type of knowledge (*ʿilm*) of God. In this light, one needs to note the Qurʾanic indication that the divine expresses creation in the language of command (*amr*) and in the logos of be/make (*kun*), which express the good intention of the creator as the paradigm of action. Here al-Ṭūsī proffers an Ismaʿili doctrine that the Imam is a physical incarnation, or an earthly instantiation of the divine goodwill. As a self-caused entity God must be a unity; and as a unity he can only create one entity, namely the Necessary Existent (*al-wajib al-wujud*), which has been equated with the First Intelligence (*nous*), from which the rest of the universe emanates in a series that has been represented by Neoplatonists as follows: After the Universal Soul emanates, the Individual Souls come forth and finally matter. Whereas Ibn Sīnā does not equate his Necessary Existent with the God of Islam, the major Ismaʿili theologian prior to al-Ṭūsī, Nāṣir Khosrow, explicitly states that God creates the Necessary Existent, from whom the rest of the universe then emanates. A Zoroastrian and a Nietzschian type of ethics is implied in al-Ṭūsī's cosmogony, where the good is associated with the good intention of the agent in the context of imitating the Imam.

THE THEODICY OF SOFT DETERMINISM

Al-Ṭūsī held that *free will*, *determinism*, and *indeterminism* are metalinguistic terms for explaining actions. A system is determined if the future can be predicted from a knowledge of all events and laws. When people are unaware of causes of behavior, free will is attributed to an agent, whose will corresponds with necessity—for example, a pregnant mother who wills the birth of her child. Having free will does not imply that the will is free and indeterminism is true. Total freedom is an intentional state of an agent that is achieved through knowledge of causes of events and one's "love"-receptivity to accept one's fate-role in the best of all possible worlds, as is exemplified by parents who graciously accept the facts of aging and welcome their children's well-deserved authority. In this tenor, al-Ṭūsī's system resembles Gottfried Leibniz's view of the best of all possible worlds. H. A. Wolfson notes that such a resemblance is due to Leibniz's copying Spinoza's theodicy, which in turn can be traced to the influence of Avicennan thought on Maimonides. Following Tolstoy's view that "free will is the essence of life, but it is an illusion," al-Ṭūsī holds that free will is an intentional concept. "Will per se," he states, "cannot be cause of any action in a mind-independent world" (pm: see *Metaphysics of Tusi*, p. 39–40). Al-Ṭūsī holds that to God, who is a unity, neither free will nor determinism applies, because an agent is free, if his or her will agrees with necessity (which implies a duality in the agent).

REFUTATION OF MATTER

Through a number of proofs al-Ṭūsī points out the incompatibility of the notion of the ultimate indivisible material substance of early Sunni theologians. Consider, for example, the following 4 by 4 arrangement of material substances:

A 0 0 0 0 B

0 0 0 0

0 0 0 0

C 0 0 0 0 D

Imagine a triangle, where hypotenuse is BC, the base is CD, and a side is BD. According to the atomic theory of homogeneous indivisible matters with no space between them, the base CD would equal the hypotenuse, which is BC. But this conclusion contradicts the Euclidean rule that a hypotenuse (BC) is longer then the base CD. Upholding the absoluteness of the Euclidean geometry, al-Ṭūsī uses this and seven other proofs to refute the material theory of substance.

THE APPLICATION OF PHILOSOPHICAL ANALYSIS TO DIFFERENT SENSES OF INFINITY

Al-Ṭūsī faces the following dilemma: As a philosopher he has to agree with Aristotle and Ibn Sīnā in holding that the "actual infinite" is not a legitimate notion, yet as a mathematician he needs to employ "infinity" in the the-

ory of numbers. Moreover, as a phenomenologist he had to use a continuum to explain perception and a continuum is often expressed by real numbers. In a clever manner that resembles R. Carnap's celebrated method of reconstructionalism and fits into the tradition of philosophical analysis, al-Ṭūsī proffers the following solution. He begins by distinguishing different senses of infinity in their application to various domains such as the "syntactical" realm, the actual world, the phenomenology of experiences such as perception, and the like.

INTENTIONAL MYSTICAL VIRTUES

Al-Ṭūsī wrote several texts on intentional analyses of the moral psychology of mystical experience. A number of investigators, such as Wilfred Madelung, hold that al-Ṭūsī's main purpose was to propose a practical experiential praxis of mysticism of the Shiʿa kind that was an alternative to the Sunni school of Ibn ʿArabi that had been advocated by Al-Qunawi.

See also Aristotle; Avicenna; Carnap, Rudolf; Determinism and Freedom; Ibn al-ʿArabī; Islamic Philosophy; Leibniz, Gottfried Wilhelm; Logos; Maimonides; Neoplatonism; Spinoza, Benedict (Baruch) de; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

Dabashi, H. "Khwajah Nasir al-Din al-Tusi: The Philosopher/Vizier and the Intellectual Climate of His Times." In *History of Islamic Philosophy*, edited by Sayyed H. Nasr and Oliver Leaman, 527–584. London: Routledge,1998.

Metaphysics of Tusi. Translated by Parviz Morewedge. New York: Institute of Global Cultural Studies, 1992.

Morewedge, P. "The Analysis of 'Substance' in Tusi's Logic and in the Ibn Sinian Tradition." In *Essays on Islamic Philosophy and Science*, edited by G. Hourani. Albany, NY: Suny Press, 1975.

Naṣīr al-Dīn al-Ṭūsī. *Contemplation and Action: The Spiritual Autobiography of a Muslim Scholar*, edited and translated by S. J. Badakhsahni. London: I. B. Tauris, 1998.

The Nasirean Ethics by Nasir al-Din Tusi. Translated by G. M. Wickens. London: George, Allen and Unwin, 1964.

Parviz Morewedge (2005)

NASR, SEYYED HOSSEIN

(1933–)

Seyyed Hossein Nasr is a Persian Islamic scholar and traditionalist philosopher. After receiving his primary school education in Iran, he was sent to the United States at the age of twelve and graduated from the Peddie School in New Jersey in 1950. He studied physics and mathematics at Massachusetts Institute of Technology (MIT) and received his doctorate from Harvard University in 1958 with specialization in Islamic cosmology and science. From 1958 until 1979, Nasr was professor of the history of science and philosophy at Tehran University where he became dean of the Faculty of Letters for some years. He also served as president of Aryamehr University in Iran. It was during these years in Iran that Nasr studied with such traditional philosophers as S. M. Kazim ʿAssar and S. M. Hossein Tabatabaʾi.

After the Iranian Revolution of 1979, Nasr migrated to the United States and taught at Temple University before joining the George Washington University in 1984. In 1981, Nasr gave the Gifford Lectures at the University of Edinburgh, which was published the same year as *Knowledge and the Sacred*. In 1999 he was chosen to be the first Muslim scholar to receive the Templeton Religion and Science Course Award. Most recently, a volume in the *Library of Living Philosophers Series* has been dedicated to him and his work.

As a prolific scholar and philosopher, Nasr has written extensively on topics as diverse as metaphysics and cosmology, tradition and modernity, Islamic science, comparative mysticism, Islamic art, interfaith dialogue, Sufism, and the environmental crisis. He is a prominent member of the traditionalist school of thought that includes such names as René Guénon, Ananda Coomoraswamy, and Frithjof Schuon. Nasr has played a key role in formulating and disseminating the ideas of the traditionalists on traditional metaphysics, sacred view of nature, and the critique of modern science. His *Knowledge and the Sacred*, his magnum opus in the field of philosophy and comparative religion, attempts to reconstruct traditional philosophy as an alternative to the modern worldview that Nasr describes as metaphysically blind and reductionist. Like the other traditionalists, Nasr places religion—or what Schuon calls *religio perennis*—at the heart of human history. A closely related term that permeates his work is perennial philosophy, which again points to the universality of tradition. In this view, tradition does not mean customs but signifies that primordial truth of divine origin that lies at the center of all cultures and religious traditions. Tradition is thus closely related to revelation and its articulation in philosophy, theology, mysticism, and sacred art.

Nasr's concept of traditional metaphysics is centered around a holistic and hierarchic view of reality. Saturated with traditional theocentrism, Nasr's view of metaphysics posits God or the One as the source, center, and end of all

there is. This principle takes on many different forms and formulations in different traditions but remains essentially the same. In keeping with the spirit of premodern philosophy, the spiritual has a higher ontological status over the material because the former is taken to reveal the divine and the latter to conceal it.The imagery of the great chain of being defines a good part of Nasr's metaphysical works. Nasr also attempts to create a holistic view of reality by showing the interrelatedness of the various levels and states of being.

Because every level of reality has its own meaning and place in the total economy of divine creation, none of them can be reduced to a lower order of reality nor the whole to one single element. According to Nasr, it is this teleological and hierarchic view of the universe that has prevented the premodern sciences of nature from slipping into reductionism and materialism. In addition to *Knowledge and the Sacred*, Nasr has provided a detailed analysis of these issues in his other works including *The Need for a Sacred Science* (1993) and *An Introduction to Islamic Cosmological Doctrines* (1964). In his major works on traditional metaphysics and cosmology, Nasr's main concern has been to revive *scientia sacra* (sacred science) by showing the underlying unity and interrelatedness of the transmitted, intellectual, and physical sciences under the umbrella of metaphysics.

Nasr sees all cultures and civilizations emanating from an essentially religious vision of the universe. This has led him to author a number of works on what he calls the "sacred view of the universe." From an ethical point of view, nature is seen as a sacred trust from God and from a metaphysical and theological point of view as *vestigia Dei* (signs of God; *ayāt Allah* in Arabic). This suggests that the order of nature has an essential telos, which makes it teleological, sacred, and intrinsically intelligible all at once. Nasr's lifelong interest in traditional and modern science can thus be seen as an extension of his view of metaphysics. In a number of works on *Islamic science*, a term Nasr has introduced to the field, he discussed the meaning of science within the context of the Islamic religious worldview and analyzed the achievements of Islamic scientific tradition in such fields as medicine, astronomy, mathematics, algebra, chemistry, physics, geography, and natural history.

Nasr's works on the relationship between religion, science, and the environmental crisis have had a longstanding impact in both the Islamic and European intellectual circles. His early work *The Encounter of Man and Nature: The Spiritual Crisis of Modern Man*, first appeared in 1968 and was one of the first books to predict the environmental crisis. The book is a philosophical critique of the modern conception of nature as inert matter. This is also the first book in which Nasr takes up the challenge of modern science and its secular outlook. The second important book to appear in this line of writings is *Religion and the Order of Nature* (1996) in which he gives an account of the rise of modern science, criticizes the secular and reductionist philosophies of nature, and presents the traditional religious view of cosmos and the human body as a viable alternative to modern scientism and reductionism.

An overall concern of Nasr's thought has been to define the fault lines of tradition and modernity. As a traditionalist philosopher, Nasr defines modernity as a distinct worldview based on the denial of the transcendent, and rejects it. He considers the environmental crisis, the modern culture of nihilism and skepticism, and the rise of scientific positivism and materialism a direct result of the various forms of modernism. Against the proponents of modernism in both the European and the Islamic world, Nasr calls for a revival of the Islamic intellectual tradition in particular and traditional thought in general to address the challenges of the modern world. His work on Islamic philosophy and Sufism has been instrumental in showing the relevance of this tradition for questions of immediate concern to the contemporary Muslim world.

See also Cosmology; Islamic Philosophy; Metaphysics; Nature, Philosophical Ideas of; Sufism.

Bibliography

WORKS BY NASR IN ENGLISH

An Introduction to Islamic Cosmological Doctrines. Cambridge, MA: Harvard University Press, 1964.

Three Muslim Sages. Cambridge, MA: Harvard University Press, 1964.

The Encounter of Man and Nature: The Spiritual Crisis of Modern Man. London: Allen and Unwin, 1968.

Science and Civilization in Islam. Cambridge, MA: Harvard University Press, 1968.

Sufi Essays. London: Allen and Unwin, 1971.

Islamic Spirituality. 2 vols. New York: Crossroads Publications, 1987–1992.

Knowledge and the Sacred. New York: Crossroad, 1981.

The Need for a Sacred Science. Albany: State University of New York Press, 1993.

The Islamic Intellectual Tradition in Persia. Edited by M. Aminrazavi. London: Curzon Press, 1994.

History of Islamic Philosophy, edited, with Oliver Leaman. London: Routledge, 1996.

Religion and the Order of Nature. Oxford: Oxford University Press, 1996.

The Philosophy of Seyyed Hossein Nasr, edited by L. Hahn, L. Stone, and R. Auxier. Chicago and La Salle, IL: Open Court, 2000.

Ibrahim Kalin (2005)

NATIONALISM

In defining the word *nationalism,* at least five senses can be identified: (1) a sentiment of loyalty to a nation (a variety of patriotism); (2) a propensity, as applied to policies, to consider exclusively the interests of one's own nation, especially in cases where these compete with the interests of other nations; (3) an attitude that attaches high importance to the distinctive characteristics of a nation and, therefore, (4) a doctrine that maintains that national culture should be preserved; and (5) a political and an anthropological theory that asserts that humankind is naturally divided into nations, that there are determinate criteria for identifying a nation and for recognizing its members, that each nation is entitled to an independent government of its own, that states are legitimate only if constituted in accordance with this principle, and that the world would be rightly organized, politically speaking, only if every nation formed a single state and every state consisted exclusively of the whole of one nation.

NATURE AND CRITERIA OF NATIONALITY

Nationalist doctrines and theories of the kinds referred to in (4) and (5) date from the end of the eighteenth century. Attachment to one's nation and the belief that, for instance, all Englishmen constitute an English nation are, no doubt, much older. Men have always had this kind of attachment to an in-group—whether tribe, city, or nation—and a corresponding awareness of (and perhaps hostility toward) nonmembers as foreigners. But what characterizes nations, distinguishing them from groups of other kinds?

THE NATION DEFINED BY THE STATE. A nation, wrote the French revolutionary ideologist the Abbé Sieyès in 1789, is "a union of individuals governed by *one* law, and represented by the same law-giving assembly." Thus conceived, a nation's unity and identity derive from political organization, and the state would thus be logically prior to the nation. This view was consistent with the individualist or atomistic interpretation of group phenomena of which John Locke was a typical exponent and which was characteristic of much of the social theorizing of the eighteenth-century Enlightenment. Writers like Denis Diderot and Marquis de Condorcet considered that individuals must be taken to concur in the setting up of a political order because (or insofar as) it is in their interests, several and collective. A public interest, thus created, is the ground of a duty to preserve and defend the order, and the state, as the subject of this interest, becomes a proper object of loyalty. Those sharing in such a common interest would constitute one people, or nation. This view of nationality is supported by the way in which, in ordinary speech, citizenship and nationality are interchangeable in many contexts. (This was once true of legal usage, too; however, many states now distinguish the rights and duties of a citizen from those of a national.) If, however, we do distinguish nationality from citizenship in ordinary speech, it is principally by narrowing citizenship to matters of political and legal status, whereas to determine nationality we take into account criteria like place of birth, parentage, language, and cultural tradition.

THE NATION DEFINED BY LANGUAGE AND CULTURE. The conception of nationality as language and culture became articulate, as an element in nationalist ideology, at the end of the eighteenth century, mainly through the work of German writers such as Johann Gottfried Herder, Novalis, Friedrich Schleiermacher, and Johann Gottlieb Fichte. Whereas for the French revolutionaries a nation was a group of individuals subject to a single political order, for the Germans nations were distinguished from one another by God and nature. Each had its peculiar character closely related to its common language. Since language is the vehicle of a tradition, preserving and transmitting sentiments, symbols, emotional associations, and myths, to share a native language is to share a common culture. "Every language," wrote Schleiermacher, "is a particular mode of thought, and what is cogitated in one language can never be repeated in the same way in another." This concept of nationality tended to be associated with a metaphysical doctrine that saw every nation as the expression of a spirit or idea, which in turn expressed a particular aspect of the divine image. The diversity of nations was a reflection of the diversity of reality, and each nation made its necessary contribution to the progress of humankind. Its members therefore had a moral duty to preserve and foster it. Thus, in reacting against the Francophile cosmopolitanism of the *Aufklärung* (German Enlightenment), the German cultural nationalist nevertheless continued to see the nations against the backcloth of humanity, each with a

role to play in what, in the end, was a drama of humankind.

As these writers saw it, a nation's existence did not depend on its members' choice or recognition; or, rather, because it formed their consciousness, they could hardly choose not to be members. If the German nation was a natural fact, it was because men reared in a German tradition would be essentially different from Englishmen or Frenchmen. Thus, a German who tried to ape the French inhibited the expression of his own nature and made do with what for him were artificial second bests.

THE NATION DEFINED BY COMMON HERITAGE. The conception of nationality as language and culture was challenged by Ernest Renan in the famous lecture *Qu'est-ce qu'une nation?* of 1882. It is a mistake, says Renan, to confuse nations with ethnographic and linguistic groups. Common racial origin, language, or religion, common economic interests, or the facts of geography are not sufficient to constitute a nation. There are nations like the Swiss, who do not share such characteristics, and there are linguistic groups like the English-speaking peoples, who do but who do not form a single nation. According to Renan, what constitutes a nation is the possession, first, of a common history, particularly of sufferings—of a store, that is, of common memories that are a source of common sympathy and pride. But it is important that some things be forgotten, too, for until old wounds have healed, the sense of sharing a common heroic tradition will be lacking. Thus, the second condition of nationality is a will to live together and to keep the common heritage alive. "To have done great things together, and the will to do more, these are the essential conditions for a people. … The existence of a nation is … a daily plebiscite."

Granted the importance of personal identification with a common tradition in the life of a nation, the metaphor of common memories does little, perhaps, to elucidate what gives a national tradition its unity and continuity. In the sense in which memory is important for individual self-knowledge and identity, individuals cannot remember what happened before they were born. Nor need their heroic ancestors stand in any generative relation to them. It is only in a figurative sense that a Frenchman could claim Joan of Arc for an ancestor. It is only because he is already a participant in a national tradition that he knows whom to call ancestor. Different situations call out different loyalties, and the ancestors a man acknowledges may differ accordingly. An American Jew of German descent might identify himself now with Thomas Jefferson, now with Judas Maccabaeus, now with Frederick the Great. Again, although men may share memories simply by having been present at the same event, to share a common history is not just to know the same historical facts; it is to identify with the same historic symbols, feel vicarious pride in the same achievements, and feel indignation at the same affronts. A Frenchman may *know* as much about Frederick as about Joan; it is because Joan is *his* and Frederick *theirs* that he is a Frenchman. A nation exists, then, where there is a group of individuals, attached in this way to a common body of symbols, who recognize one another as fellow members sharing similar attitudes to these symbols and who, because of this, feel a loyalty and concern for one another that they would not extend to outsiders. Linguistic, religious, or physiognomic features may have a part in determining who is so recognized, and the importance of any one of them may be different in different situations.

THE NATION DEFINED BY TERRITORY. A characteristic of nationality distinguishing it from most other kinds of group attachment is its relation to territory. For a group to have no special territorial affinity would not prevent one from calling it a sect, a family, or a social class. The idea of a homeland, however, seems essential to the idea of a nation. The true cosmopolitan has no place where he belongs. This illumines the close conceptual relation between nation and state, for a state is also territorially based and will admit nonmembers only on its own terms.

Where an area has a history of conflict among religious, linguistic, or racial groups each concentrated in a particular territory, the members of each will be conscious of themselves as a separate group with a history of supremacy or suffering associated with that territory; the characteristics that significantly differentiate the group from those around it will come to be thought of as those of people who belong to that territory, even when they are also found outside it. Any such group excluded from political power may be expected to aspire to independence and to want to settle in its own territory the terms on which power and prestige are enjoyed. There is, then, a wide range of features by which a national group might identify itself and its members. Which of them becomes the focus of nationality in any given case will depend on how the group has come to self-consciousness; that feature will very often correspond to the criterion by which it has been singled out as an object of oppression. Its homeland will be the territory in which the group so defined now predominates or predominated in some earlier period to which its common recollections go back.

THE NATION DEFINED BY COMMON AIM. However, because nationalism is so often a form of protest, the concept of the nation to which it is tied may depend as much on the definition of the out-group against which it is aimed as on the positive delineation of the in-group. In the twentieth century African and Asian nationalisms, for instance, relied heavily on the repudiation of white colonialism and on an aspiration to count as the white man's equal. However, on its own this cannot be enough to constitute a nation, for though the same sentiments are found throughout Black Africa, only a few Africans see themselves as a single nation aspiring to unity in a single state. Nationalism, in fact, can exist before the nation, as the aspiration of a European-trained elite aiming at native independence in a territory defined by an imperial power for administrative convenience, not by any native tradition or symbolic attachment. Having transformed a colony into a state, nationalists in countries such as Ghana must then create a nation. That states can be as important in making nations as nations can be in making states is borne out by the success of the United States. The failure of the Austro-Hungarian Empire to create a nation was the cause of its disintegration.

NATURE SELF-DETERMINITION

The twin sources of modern nationalist doctrine are the French conception of popular sovereignty and German romantic anthropological nationalism. In eighteenth-century political theory the attribution of sovereignty to the people instead of the monarch gave the people the right to determine its own mode of government. This implied no threat to the existing order of states and gave rise to no irredentisms in France and England, where the territorial boundaries of the self-conscious nation corresponded more or less with the established frontiers of the state and where the state itself was already a national symbol. In Germany and Italy, however, nationality spilled across frontiers. If the people, being sovereign, might choose the political order it wished and if "the people" was defined by nationality irrespective of existing states, then a national will to unity and independence was self-justifying even though it dismembered existing states, upset dynastic legitimacy, and sanctioned the invasion of one sovereign state by another in the interest of national liberation. The Italian nationalist Giuseppe Mazzini put the case in extreme terms, professing the belief that the political unity and independence of every nation within its natural boundaries was ordained by God. A characteristically more moderate view was stated by J. S. Mill in *Representative Government* (1861):

> Where the sentiment of nationality exists in any force, there is a *prima facie* case for uniting all the members of the nationality under the same government, and a government to themselves apart. This is merely saying that the question of government ought to be decided by the governed. One hardly knows what any division of the human race would be free to do if not to determine with which of the various collective bodies of human beings they would choose to associate themselves.

There are very great difficulties, however, in the notion of a right to self-determination, whether individual or collective. The idea of a state as an organization exercising authority over everyone within its boundaries is not compatible with the idea of conceding to each man a right to choose whether to give it his allegiance. Of course, everyone may have a right to some influence on how and by whom he will be governed. But this amounts to a right to participate in certain constitutional decision procedures that take the political framework for granted, not to a right to take or leave it as one likes. Nor is a collective right any easier. On the practical level no amount of fragmentation or partition could put every individual in an area like the Balkans into the right state.

A more fundamental problem, however, is to decide what constitutes a national group for the purpose of self-determination. In the name of national unity Ghanaian nationalists deny self-determination to the Ashanti as the Congolese denied it to Katanga. If Germans claim that all German-speaking people, as members of the German nation, ought to be included in Germany, would the principle of national self-determination leave so-called Germans abroad any choice in the matter? And if they demurred, would it be as Germans or as non-Germans? If as Germans, would this be compatible with the self-determination of the whole German people? Clearly, if nationality is to be judged by objective criteria like language, the principle of national self-determination would support irredentist expansion policies irrespective of the wishes of the subgroup concerned since the nation's will would presumably be more authoritatively expressed by the greater part than by the lesser. But if nationality is judged by subjective criteria, like a will to live under one government, repudiation by the subgroup would appear to be ground enough for saying that it was not part of the same nation after all. But a dissentient minority within that subgroup could then equally well claim a separate national identity and so on. If one accepts subjective cri-

teria for group self-determination, there is no reason for stopping short of individual self-determination.

The objective criteria, though often difficult to apply in actual cases, do provide clear principles for the proper constitution of states. However, they can claim no support from the individualist doctrine that political obligation must rest on consent. This principle has played its part in the history of nationalist doctrine. Immanuel Kant maintained that the principle of moral freedom and autonomy implied that men, as self-legislating members of the kingdom of ends, must impose political obligation upon themselves and that authority must derive from and be subject to the general will as expressed in law. Nationalists like the German political economist Adam Müller transformed the argument, however, by identifying the individual with the nation, insisting that the individual's permanent will was more truly expressed in the *Volksgeist,* or national spirit, than in any particular individual preference. Thus, the general will, which for Jean-Jacques Rousseau and Kant reconciled individual moral autonomy with political authority and obligation, became a way of denying the relevance of personal choice when it ran counter to the national spirit.

Early nineteenth-century nationalism was nevertheless liberal and humane in intention. Fichte and Mazzini would have argued that unless a nation was united in an independent sovereign state, its members, unable to command the respect of others as equals, would be lacking in dignity and self-respect. Much of the persuasive charm of nationalism in Africa and Asia has a similar source. Men of color repudiating white superiority feel that for their own self-respect they must be ruled by men of their own color and kind with whom they can identify and who will be received on equal footing by the leaders of other sovereign states.

However, the moral uncertainty out of which nationalism is born and which is perhaps its main justification, readily turns, once unity and independence has been won, into an aggressive assertiveness and national egoism, akin to what in France Charles Maurras called "integral nationalism," "the exclusive pursuit of national policies, the absolute maintenance of national integrity, and the steady increase of national power." The nation-state is no longer set in the context of a larger humanity; it is its own sufficient justification. Nationalism in this key is frankly irrationalist, delighting in the symbolic rhetoric of "blood and soil." Enormously important as it is for the historian and sociologist, it would be absurd to treat it as if it invited serious rational criticism.

See also Condorcet, Marquis de; Diderot, Denis; Enlightenment; Fichte, Johann Gottlieb; Herder, Johann Gottfried; Jefferson, Thomas; Kant, Immanuel; Loyalty; Mill, John Stuart; Novalis; Patriotism; Philosophical Anthropology; Racism; Renan, Joseph Ernest; Rousseau, Jean-Jacques; Schleiermacher, Friedrich Daniel Ernst; Self-Interest; Social and Political Philosophy; Sovereignty; State.

Bibliography

HISTORICAL STUDIES

Hayes, Carlton J. H. *Historical Evolution of Modern Nationalism.* Chicago, 1948.

Kohn, Hans. *The Idea of Nationalism.* New York: Macmillan, 1945. History of nationality and nationalism before 1789.

Kohn, Hans. *Nationalism: Its Meaning and History.* Princeton, NJ: Van Nostrand, 1955. Introductory study with readings and selected bibliography.

GENERAL STUDIES

Hertz, Frederick. *Nationality in History and Politics.* London: K. Paul, Trench, Trubner, 1944. Sociological and historical study of nationalist ideology.

Shafer, Boyd C. *Nationalism—Myth and Reality.* London: V. Gollancz, 1955. Has twenty-page bibliography.

Snyder, Louis L. *The Meaning of Nationalism.* New Brunswick, NJ: Rutgers University Press, 1954.

CLASSIC EXPOSITIONS

Fichte, J. G. *Reden an die deutsche Nation.* Berlin, 1808. Translated by R. F. Jones and G. H. Turnbull as *Addresses to the German Nation.* Chicago: Open Court, 1922.

Maurras, Charles. *Enquête sur la monarchie.* Paris: A. Fayard, 1924. Originally published in 1900.

Mazzini, Giuseppe. *Doveri dell'uomo.* London, 1860. Translated by Ella Noyes as "The Duties of Man," in *The Duties of Man and Other Essays, by Joseph Mazzini,* edited by Ernest Rhys. London and New York, 1907.

Reiss, H. S., ed. *The Political Thought of the German Romantics, 1793–1815.* Oxford: Blackwell, 1955. Contains readings from Fichte, Novalis, Adam Müller, Schleiermacher, and Friedrich Karl von Savigny; also has selected bibliographies.

Renan, Ernest. *Qu'est-ce qu'une Nation?* In *Oeuvres complètes de Ernest Renan,* Vol. I, edited by Henriette Psichari. Paris, 1947. Originally published in 1882.

CRITICAL STUDIES

Acton, John Emerich Edward. "Nationality." In *Essays on Freedom and Power by Lord Acton,* edited by Gertrude Himmelfarb. Glencoe, IL: Free Press, 1948. For criticism of Mazzini.

Cobban, Alfred. *National Self-Determination.* Chicago: University of Chicago Press, 1948.

Kedourie, Elie. *Nationalism.* London: Hutchinson, 1960; 2nd rev. ed., London and New York, 1961. Has notes for further reading; examines metaphysical foundations of nationalism and its influence on Europe and the Middle East.

Kohn, Hans. *Prophets and Peoples.* New York, 1957. Studies of J. S. Mill, Jules Michelet, Mazzini, Heinrich von Treitschke, Fëdor Dostoevsky.

Stanley I. Benn (1967)

NATIONALISM [ADDENDUM]

However it is characterized, nationalism is a phenomenon of central importance in the modern world because it reflects the special moral significance that most people in fact attach to their ties as members of a particular nation. All forms of nationalism share the view that it is right and good for some particular people, or all peoples, to promote a common national identity through appropriate institutions. Contemporary philosophers are increasingly concerned to evaluate the claims of nationalism. Are ties of nationality desirable? Do they generate special obligations among conationals that do not extend to others? Is national identity compatible with the rights of national minorities in a larger nation-state, and duties of global justice that are owed to distant peoples? Is nationalism compatible with standard liberal assumptions concerning the equal worth of all persons, and the impartiality required for justice? If liberalism, nationalism, and global justice come into conflict, which should give way to better accommodate the prior claims of the other? What separates a morality of nationalism from a politics of tribalism?

Such issues have come to the fore in the work of contemporary liberals, communitarians, multiculturalists, and cosmopolitans. They advance rival normative models of nationalism and justice within and without borders. Many liberals are drawn to a thin civic paradigm of nationality. Communitarians favor a thicker cultural paradigm closer to the historical experience of shared nationality. On the civic paradigm, the demands of impartial justice within and between nation-states are most secure when ties of nationality consist in nothing more than individuals' relations as equal citizens of one and the same political society. This paradigm grows out of republican traditions of thought that identify the nation or people, the sole legitimate source of sovereignty, with members of the state—individuals born and living within its political borders. The civic paradigm is exemplified by Rawls's seminal reinvention of contractarian liberalism in *A Theory of Justice* (1971). In this framework, political society is a system of social cooperation for mutual advantage, which can be ordered entirely on the basis of principles of justice and political ideals of freedom, equality, and fairness. Political society is just when it conforms to principles that free and equal persons would agree to under conditions that are impartial. Rawls's principles (equal civil and political liberties, equality of opportunity, and economic arrangements that either ensure material equality or use inequality to raise the material well-being of all) are embodied in his model of a just liberal-democratic welfare state.

Rawls argues that this model provides a stable, well-ordered nation-state on its own terms. It sustains the very sense of justice and mutual respect among persons as citizens that are necessary and sufficient to motivate them to support their political obligations, independently of any thicker ties of history, culture, religion, ethnicity, family, class, gender, and so on. For Rawls, justice is the primary virtue of human life in society. To this end, the civic paradigm of nationality constructs ties of nationality as the relations of equal citizens who recognize one another in a common allegiance to their shared political ideals and institutions. These ties are precisely what justice requires and all that justice requires by way of community. Of course the civic paradigm leaves ample room for the many cultural, ethnic, or religious attachments people may embrace. But for the Rawlsian liberal, the requirements of justice do not derive from any of these more particular ties. Furthermore, the requirements of social justice operate as background constraints on the permissible structure of all such ties, conforming them to persons' equal rights and duties as citizens. A civic nationalism justifies the politics of building just institutions and peoples bound together as citizens by their allegiance to shared political ideals.

COSMOPOLITAN LIBERALS

For cosmopolitan liberals such as C. R. Beitz, Thomas W. Pogge, and Brian M. Barry, John Rawls's contractarianism and the civic paradigm have a great advantage in that they lay the basis for impartial and egalitarian principles of global justice. On their cosmopolitan argument, ignoring particular ties of nationality, as well as other contingencies such as race or gender, is a natural extension of Rawlsian liberalism. It justifies a choice of principles of global justice that do for the poorest persons in the world what Rawls's principles are supposed to do for the worst off in any particular nation-state. The civic paradigm of nationality as political citizenship seems well suited to allow each person to be a citizen of a just world because it rests on political ideals that are supposed to be universal and impartial in scope. Cosmopolitan justice redistributes the

wealth and resources of the world and reconstructs international arrangements with the aim of gaining rough equality in all persons' individual liberties, standards of living, and political rights. These arrangements provide a background justice, which constrains the conduct of nation-states and reconciles individuals' rights and duties as citizens of a particular nation with their rights and duties as cosmopolitan citizens of the world. Whether or not cosmopolitan liberalism leaves room for any recognizable ties of nationality comes into question by communitarian nationalists.

COMMUNITARIANS

Communitarians such as Michael Sandel, Michael Walzer, Will Kymlicka, David Miller, and Charles Taylor provide the bases for an alternative cultural paradigm of nationality and nationalism. Communitarians argue that the civic paradigm is too thin to capture the bonds between persons that are both necessary and desirable for the robust experience of nationality. The ties of nationality are and ought to be the rich bonds of membership in a historic community marked by a shared societal culture and way of life. Such a culture involves a common language, time-honored customs, shared traditions, inherited institutions, agreed-on social meanings and values, and the exercise or aspirations of political autonomy in a certain geographic area. Nationalism is the process through which a societal culture of this sort is built up, enters into the identity of its members, and finds expression in political acts through which its members seek to create and preserve it as an independent entity. As such, national identity is not reducible to political ties of citizenship, or shared political ideals. Rather, a shared societal culture actively preserved by its members constitutes a people or nationality and this is precisely what a political society ought to recognize and protect. For communitarians, the existence of a nation in this rich culture and historical sense justifies rights of national self-determination, whether it is the right of a people to independent statehood (Walzer 1977) or more limited rights of self-determination within a multinational confederated state (Kymlicka 1989). The violation of these rights is taken to ground just war theory (Walzer 1977) and the right of a national minority to secede and form its own sovereign state (Buchanan 1991).

This paradigm of nationality as shared historical culture may be justified, and by implication challenged, in various ways. Some communitarians defend it on the basis of a philosophical conception of the self that holds that the self always gains its identity, purposes, and obligations from the particular community(ies) in which it is embedded (Sandel 1982). Aspects of this conception motivate the nationalist argument that the very rights and duties of citizenship, stressed by the civic paradigm, depend on the fact that citizens are already bound together by a common nationality and thick cultural ties to their compatriots (Kymlicka 1989, Miller 1995). Others argue that any knowledge of justice and framework of moral deliberation always depend on the intersubjective meanings and values shared by a particular political community (Walzer 1977). The cultural framework provided by national identity can be justified by the argument that such a framework is necessary to provide persons with their meaningful options in life. Without such options, people lack any genuine individual freedom and liberal equality (Kymlicka 1989). A shared national identity is also defended on the grounds that it provides the only appropriate basis of reciprocal recognition among peoples, encompassing both a respect for cultural difference and human commonality (Taylor 1994). The ties of national culture are defended as intrinsically valuable because of the special human virtues and goods they make possible, such as loyalty, courage, love of country among compatriots (Miller 1995). National identities provide persons with rich cultural self-images that probably cannot, and should not, be replaced by a bare image of oneself as part of humanity, a citizen of the world, or a disembodied impartial deliberator; so the communitarian argument goes.

The import of the cultural paradigm of nationality, and communitarian nationalism, critically depends on what kinds of national community one has in mind. In modern history, nationalism in its cultural communitarian form has often implied tribalism, and a virulent hostility, intolerance, or indifference to other peoples (Arendt 1948). More generally, what of illiberal or oppressive national communities that violate the rights or stigmatize the identities of some of their own members, or of nonmembers, outsiders, foreign peoples (Doppelt 1998, 1999, 2002)? In liberal theory, the moral right of national self-determination universally applies to all peoples and implies duties of every people to respect or even defend the rights of other peoples. In practice, what sort of national cultures and nationalism are compatible with the rights of all individuals and peoples? These concerns inspire recent debates among communitarian nationalists and liberals concerning the possibility of a liberal nationalism based on a cultural paradigm of national identity.

LIBERAL NATIONALISTS

Liberal nationalists seek to harness the special value of national belonging to liberal ideals so that each informs and constrains the other. For Kymlicka, one central problem for liberal nationalism arises because most nation-states contain several nationalities, involving majority and minority cultures. His multicultural liberalism employs the culture paradigm of nationality to justify group rights for national minorities in a larger nation-state, as a requirement of domestic justice. Such group rights to a limited measure of self-government, territorial sovereignty, and cultural autonomy empowers a national minority (French Canadians in Quebec; Native American bands on their tribal homelands) to preserve its historical community and national identity from assimilation to the dominant culture. The liberal nation-state should foster multinational cultures and institutions, as Canada, the United States, and other states in fact do. This multinationalism is supposed to be a liberal nationalism for two reasons. First, group rights are justified as necessary means to the fulfillment of liberal individualist ideals of freedom and equality for members of national minorities. Secondly, group rights to cultural autonomy are supposed to be fully compatible with the individual rights people possess as citizens of a liberal state. By this route, multiculturalism reconciles the cultural paradigm of nationality with the civic paradigm, and thus nationalism with liberalism. The citizen of a multinational state is supposed to combine a nationalist attachment to his or her own cultural community with a political attachment to the general rights and duties of citizenship.

NATIONALISM VS. LIBERALISM IN THE MULTICULTURAL MODEL

Tensions arise between nationalism and liberalism in the multicultural model if group rights protect minority cultures with some illiberal or oppressive practices, or these group rights are embraced by minorities to shield them from oppression by an illiberal majority culture. The societal cultures protected by group rights are not supposed to define national identities in essentialist terms that exclude people or discriminate against them, on the basis of alleged racial characteristics, blood, descent, and the like (Kymlicka 1989). The model may also be unstable to the extent that multiple national identities in one and the same nation-state may fail to sustain sufficient unify for domestic liberal justice (Doppelt 1998, 1999, 2001). Such instability can motivate the descent of multiculturalism into either secession, assimilation, or domination. Miller's model of liberal nationalism suggests that a weak or fragmented nation-state can be countered by providing unifying ties of national identity among all groups of citizens. It can do so by building a pluralistic national culture and identity, which is continually reshaped to include groups that have been oppressed, excluded, or marginalized. From this standpoint, the best liberal response to the existence of diverse ethnic and national minorities is not necessarily either group rights or cultural assimilation. To some extent, a liberal nationalism may require a democratic expansion of national culture and identity such that all groups can express and recognize themselves in it. All states engage in the construction of nationality through their activities in the spheres of law and public policy, immigration practices, public schooling, military service, political rhetoric, public ceremonies, and holidays, and so on. Criteria of liberal nation-building should be based on the extent to which a nation employs means that are consistent with democratic rights, and achieves results that embody an inclusive nationality (Kymlicka 1989).

FURTHER APPROACHES TO LIBERAL NATIONALISM

Rawls's turn in his later work to political liberalism and the law of peoples provides yet another approach to liberal nationalism. He develops a conception of political society with communitarian components that moves his liberalism from a civic to a more cultural paradigm of the nation, though not as thick as other communitarian nationalists. Rawls now grounds the liberal ideals underlying contractarian justice in our particular historical tradition of legal and political institutions and ideas. Our political tradition is expressed in canonical texts such as the Declaration of Independence, the U.S. Constitution, the Bill of Rights, and landmark legislation. It lives in all the ways their basic principles enter into our democratic institutions and political practices. Rawls's political liberalism responds to the perennial worry of liberal nations that agreement on the fundamentals of justice is blocked by the divergent comprehensive views of life held by people of different religions, ethical outlooks, national or ethnic identities, and so on. Yet by public reflection on the meaning of our shared political culture, people with divergent views of life can attain an overlapping consensus on the most basic ideals of American democratic citizenship, and thus on political justice. With this focus on building democratic institutions and a national political identity, Rawls appropriates the cultural paradigm of nationality to vindicate his account of domestic justice and civic nationalism. This challenges the liberal nation-

alists who require much thicker cultural ties of nationality to explain why the conception of free and equal persons built into democratic culture is not sufficient for national identity.

CHALLENGES TO LIBERAL NATIONALISM

The sharpest challenge to liberal nationalism is posed by cosmopolitan liberals who apply Rawls's principles of justice to generate the global egalitarian justice described above (Beitz 1983, Pogge 1994, and Barry 1999). Strong ties of nationality of the sort stressed by communitarians (liberal or not) support special obligations and affinities among conationals that either directly contradict their duties of cosmopolitan justice or weaken their motivation to comply with them. For example, the right of a people to national self-determination defended by communitarian nationalists is typically taken to include a right to control the national resources, wealth, and capital possessed by that people. But from the cosmopolitan standpoint, the distribution of wealth among nations may reflect a legacy of injustice or bad luck, and is, in any case, incompatible with the demands of global egalitarian justice. Nations have little or no independent moral standing except in the degree that their internal and external relations conform to the (Rawlsian) principles of justice that every person in the world could reasonably accept.

Liberal nationalists respond to this challenge by arguing either against the cosmopolitan conception of global justice or in favor of its compatibility with liberal nationalism. Walzer and Miller reject egalitarian cosmopolitanism because it fails to make sense of the ways persons' obligations to one another are typically rooted in a particular nation-state and not understood to extend to all humanity. Some philosophers underscore the ways common sense morality supports the dependence of persons' duties on the special associative ties of family, friendship, membership in a particular nationality, and the like (Scheffler 2001). Indeed, Rawls's turn to a conception of justice grounded in a people's own political traditions supports a nation-based account of international justice (his law of peoples) at odds with egalitarian cosmopolitanism. This tradition-based view of domestic justice allows that other peoples or nations with different, indeed illiberal political traditions can be just and decent; provided that they are well-ordered societies that respect the most basic rights of their own members (e.g., to physical security, the material means of life, etc.) and the rights of other just or decent peoples to self-determination. Liberal and illiberal but decent nations can reasonably agree to principles of international justice that imply some mutual rights and duties among peoples. These include duties of material assistance to burdened nations that need it to become well-ordered, decent, and just, but not duties of egalitarian global redistribution.

From this standpoint, cosmopolitan justice is unreasonable because it ignores that peoples with different cultural traditions, exercising their rights of national self-determination, can be expected to have different levels of economic development, standards of living, and criteria of domestic justice. As such, nations bear some responsibility for their standards of living and thus differences between them do not in themselves imply global injustice. So while both Miller and Rawls support duties of distributive global justice and the rights of all persons to the basic means of life, health, and subsistence, both argue that cosmopolitan egalitarian principles come at the price of people's national identities, responsibilities, and rights of self-determination. By this route, liberal nationalism is reconciled with a much less demanding account of global justice, and one more in tune with the communitarian view of the independent moral standing of nations.

Other liberal nationalists such as Kok-Chor Tan defend cosmopolitan egalitarianism and argue that properly understood, it is compatible with a limited form of nationalism. The debate between nationalists and cosmopolitans often lump together different claims concerning what is supposed to make them incompatible: conflicting institutional requirements, nation-state versus global state; conflicting identities or attachments, national identity versus citizen of the world; conflicting moralities, nationalist partiality versus liberal impartiality; and conflicting views of justice, special duties of justice among conationals versus duties to all humanity in cosmopolitan justice. Liberal nationalists who defend egalitarian cosmopolitanism take the issue of justice to be the one that is fundamental to their reconciliation. The basic strategy is to limit or circumscribe the special rights and duties individuals have as conationals, the right of nations to self-determination, and the imperatives of domestic justice, so that they conform to the morally prior requirements of global egalitarian justice. The thought is that even so, the rigorous constraints imposed by cosmopolitan justice leave some room for associative ties of nationality and special duties arising from them. Cosmopolitan justice constrains domestic justice, and how nations may exercise rights of national self-determination, but does not destroy the important space

they occupy. Such a vision of liberal justice is in principle neutral concerning institutional, cultural, and identity issues that are used to drive a wedge between nationalism and cosmopolitanism (Tan 2004). For example, cosmopolitan equality does not necessarily require one global government in place of national political institutions. Nor does people's acceptance of duties to support egalitarian global arrangements imply the abandonment of any national identity or partiality to one's compatriots in some respects.

For these reasons, a liberal nationalism may be compatible with cosmopolitan justice, at least in principle. In practice the dynamics of human motivation might raise anew the tensions between nationalism and cosmopolitanism, as conflicting attachments and identities. In any case, these debates enrich political philosophy by bringing issues of political culture, national identity, nationalism, and global justice into the heart of contemporary liberalism.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism; Social and Political Philosophy.

Bibliography

Arendt, Hannah. *The Origins of Totalitarianism*. London: Harvest, 1976. First published in 1948.

Barry, Brian. *Culture and Equality: An Egalitarian Critique of Multiculturalism*. Cambridge, U.K.: Polity Press, 2001.

Barry, Brian. "International Society from a Cosmopolitan Perspective." In *International Society: Diverse Ethical Perspectives*, edited by D. Mapel and T. Nardin. Princeton, NJ: Princeton University Press, 1998.

Barry, Brian. "Statism and Nationalism: A Cosmopolitan Critique." In *Global Justice*, edited by I. Shapiro and L. Brilmayer. New York: New York University Press, 1999.

Beitz, Charles R. "Cosmopolitan Ideal and National Sentiment." *Journal of Philosophy* 80 (10) (1983): 591–600.

Beitz, Charles R. *Political Theory and International Relations*. Princeton, NJ: Princeton University Press, 1979.

Buchanan, Allen E. *Secession*. Boulder, CO: Westview Press, 1991.

Doppelt, Gerald D. "Illiberal Cultures and Group Rights: A Critique of Multiculturalism in Kymlicka, Taylor, and Nussbaum." *Journal of Contemporary Legal Issues* 12 (2): (2002): 661–692.

Doppelt, Gerald D. "Is There a Multicultural Liberalism?" *Inquiry* 41 (1998): 223–238.

Doppelt, Gerald D. "Liberalism, Multiculturalism, and the Politics of Identity." In *Applied Ethics: Proceedings of the 21st International Wittgenstein Symposium Vienna*, edited by P. Kampits, K. Kokay, and A. Weiberg. Vienna: Verlag Holder-Pichler-Tempsky, 1999.

Habermas, J. "Citizenship and National Identity." *Praxis International* 12 (1992): 1–19.

Hurka, T. "The Justification of National Partiality." In *The Morality of Nationalism*, edited by R. McKim and J. McMahan. Oxford, U.K.: Oxford University Press, 1997.

Kymlicka, Will. *Liberalism, Community and Culture*. New York: Oxford University Press, 1989.

Kymlicka, Will. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford, U.K.: Oxford University Press, 1995.

Kymlicka, Will. *Politics in the Vernacular: Nationalism, Multiculturalism and Citizenship*. Oxford, U.K.: Oxford University Press, 2001a.

Kymlicka, Will. "Territorial Boundaries: A Liberal Egalitarian Perspective." In *Boundaries and Justice*, edited by D. Miller and S. Hashmi. Princeton, NJ: Princeton University Press, 2001b.

Margalit, A., and J. Raz. "National Self-Determination." *Journal of Philosophy* 87 (1990): 439–461.

Miller, David. *Citizenship and National Identity*. Cambridge, U.K.: Polity Press, 2000.

Miller, David. "The Ethical Significance of Nationality." *Ethics* 98 (4) (1988): 647–662.

Miller, David. "The Limits of Cosmopolitan Justice." In *International Society*, edited by D. Mapel and T. Nardin. Princeton, NJ: Princeton University Press, 1998.

Miller, David. *On Nationality*. Oxford, U.K.: Oxford University Press, 1995.

Moellendorf, D. *Cosmopolitan Justice*. Boulder, CO: Westview Press, 2002.

Pogge, Thomas W. "An Egalitarian Law of Peoples." *Philosophy and Public Affairs* 23 (3) (1994): 195–224.

Pogge, Thomas W. *Realizing Rawls*. Ithaca, NY: Cornell University Press, 1989.

Rawls, John. *The Law of Peoples: With, "The Idea of Public Reason Revisited."* Cambridge, MA: Harvard University Press, 1999.

Rawls, John. *Political Liberalism*. Paperback ed. New York: Columbia University Press, 1996. First published in 1993.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Sandel, Michael. *Liberalism and the Limits of Justice*. Cambridge, U.K.: Cambridge University Press, 1982.

Scheffler, Samuel. *Boundaries and Allegiances*. New York: Oxford University Press, 2001.

Tan, Kok-Chor. *Justice without Borders: Cosmopolitanism, Nationalism, and Patriotism*. Cambridge, U.K.: Cambridge University Press, 2004.

Taylor, Charles. "The Politics of Recognition." In *Multiculturalism: Examining the Politics of Recognition*, edited by Amy Gutmann, 52–56, and 58–60. Princeton, NJ: Princeton University Press, 1994.

Walzer, Michael. *Just and Unjust Wars*. New York: Basic Books, 1977.

Gerald Doppelt (2005)

NATIVISM, INNATISM

See *Innate Ideas; Innate Ideas, Nativism*

NATORP, PAUL
(1854–1924)

Paul Natorp was born in Düsseldorf and died in Marburg. Along with Hermann Cohen, he is known as one of the founders of the Marburg School of Neo-Kantianism. He studied history, philology, mathematics, and philosophy in Berlin, Bonn, and Strassburg. After completing his doctorate in history at the University of Strassburg in 1876, he went to Marburg where Cohen was working on the restoration of Kant's critical philosophy. In 1881 Natorp obtained his postdoctoral qualification with a thesis on the prehistory of criticism titled *Descartes' Erkenntnistheorie* (Descartes's theory of knowledge). He became an associate professor at the University of Marburg in 1885 and eventually a full professor of philosophy and pedagogy. In spite of being offered several chairs at other universities, Natorp remained in Marburg throughout his lifetime.

Despite their close relationship, Natorp cannot be seen as a genuine follower of Cohen, especially because of the explicitly historical foundation of his philosophy. Beside his interests in epistemology and the theory of science, both orientated on Cohen's logic, Natorp worked on problems in ethics, the philosophy of religion, philosophical psychology, and the philosophical foundations of pedagogy.

Three main periods of Natorp's philosophical work can be distinguished. During the earliest period, Natorp developed a methodical idealism that takes Kant and Cohen as its point of departure and is presented in *Die logischen Grundlagen der exakten Wissenschaften* (The logical foundations of the exact sciences, 1910) as well as in *Die Philosophie. Ihr Problem und ihre Probleme* (Philosophy. Its problem and its problems, 1911). Both books focus on the problem of definition in the natural sciences from an epistemological perspective. They take the mathematical approach as a paradigm for the object-creating function of consciousness. Natorp reduces the transcendental-logical analysis of the constitution of objects to the categories of scientific definition. Science, in this context, stands for the transcendental subject. This is the main characteristic of Natorp's philosophy: The sciences as *facts of reason* are taken as the only legitimate starting points of the *transcendental method* so that epistemology is understood as a theory of science.

Natorp attempted to find a basis for his methodical idealism not only through systematic thought but also through studies in the history of philosophy. His studies of Plato, as recorded in an extensive work on *Platons Ideenlehre* (Plato's theory of ideas, 1903), are a good example of his historically orientated method. The book is still discussed as an example of a strictly systematic view on the history of philosophy.

The second period in Natorp's work is introduced by the *Allgemeine Psychologie nach kritischer Methode* (General psychology according to a critical method, 1912). In this book, Natorp proposes philosophical psychology as a discipline that should be able to examine the transcendental constitution of objects through reference to the subject's concrete nexus of experiences. Natorp hereby added a genetic aspect to epistemology, which was adopted by several philosophers of his time and had a significant influence on Edmund Husserl's foundation of phenomenology. At the same time Natorp abandons the restriction of the transcendental analysis to the *fact of science* by enlarging the *factum* to a *fieri*, the *what is* to the *what is to be*, so that cognition is no longer taken as a mere fact but rather as a process within the subject.

The psychological method is *reconstructive* insofar as the psychologist analyses the cognitive *process* and goes back to the very origin of cognition in a step-by-step analysis. The subject's experience is taken as the original source of cognition, but it is not intuitively given (as the phenomenologist would put it). According to Natorp the experience of the subject can only be reconstructed in a genetically oriented epistemological process. Concrete experience inevitably has to be transformed into an abstract definition of experience. The correlative of the object—the subject—is *reconstructed* post hoc. However, the reconstruction itself is a cognitive act and, as such, bound to definition and the interruption of cognition in process. Reconstruction, like any other cognitive process, is an approximative approach to and a concretization of the object, except that it does not lead back to the natural object but to its correlate, namely, the subject. The individual subject is correlated to an individual object, and every individual *case* of definition corresponds to an individual *case* of cognition. Natorp's concept of subject here is still a very restrictive one.

In his later philosophy Natorp connected logic and psychology by transforming his understanding of cognition as a concretization of being into a *general logic* concerning the relation of objects that is integrated into an extensive metaphysical conception. In his *Vorlesungen über praktische Philosophie* (Lectures on practical philosophy, edited in 1925) as well as in the *Philosophische Systematik* (Philosophical systematics, edited in 1958). Natorp attempted to preserve the guiding themes from of his early thought in a transformed way that now incorpo-

rates an ontological perspective. Instead of analyzing the object in terms of a theory of knowledge or pursuing a psychological analysis of the subject, Natorp turns his attention to the correlation between subject and object. The unity of the object depends on the basic condition that there *is* something ("it is"—*estin*—whereby the "it" remains undefined). On the other hand the specific unity of the individual subject is the correlate of the defined object's unity (the specific "this one"—*tode ti*). Both parts of the correlation are connected by the process of categorical–ontological definition. This dialectic approach is elucidated in detail in Natorp's *Philosophische Systematik* and extended to a theory of categories, which is only faintly reminiscent of Kant. In contrast to the older type of logic (á la Cohen) in Neo-Kantianism, Natorp outlines the *it is* as the epitome of being, which is prior to any logical definition. In the end, Natorp argued for an ontological interpretation of the *origin* in transcendental philosophy.

Overall, Natorp's impact on the history of philosophy has been primarily indirect. His ideas were carried on above all by the youngest representative of the Marburg School, Ernst Cassirer. The logical motif of Natorp's earliest period had a great influence on Cassirer's philosophy of culture, which did influence the broader philosophical discussion. The early phenomenologists also referred to Natorp, even if these references were primarily critical, as when Husserl and Martin Heidegger, for example, tried to avoid Natorp's one-sided emphasis on transcendental logic. At present the general philosophical audience is becoming increasingly aware of the close relationship between Natorp's later thought and Heidegger's thinking of being. In some ways Natorp's philosophical development even illustrates the general development of transcendentalism during the twentieth century, which starts with a logic of pure cognition and ends with the problem of thinking, which is already and inevitably related to being.

See also Being; Cassirer, Ernst; Cohen, Hermann; Epistemology; Ethics; Heidegger, Martin; Husserl, Edmund; Neo-Kantianism; Phenomenology; Philosophy of Religion; Plato; Thinking.

Bibliography

WORKS BY NATORP

Please note: No English translations of these texts are available.

Forschungen zur Geschichte des Erkenntnisproblems im Altertum (Studies in the history of the problem of knowledge in antiquity). Berlin: Hertz, 1884 (later edition published by Hildesheim: Olms, 1965).

Einleitung in die Psychologie nach kritischer Methode (Introduction to psychology according to a critical method). Freiburg: Mohr, 1888.

Religion innerhalb der Grenzen der Humanität (Religion within the bounds of humanity). Tübingen: Mohr, 1894 (later edition published in 1908).

Sozialpädagogik. Theorie der Willensbildung auf der Grundlage der Gemeinschaft (Social pedagogy. A theory of the cultivation of the will on the basis of community). Stuttgart: Frommann, 1899 (later editions published in 1904, 1909, 1920, 1922).

Platons Ideenlehre (Plato's theory of ideas). Leipzig: Dürr, 1903 (later edition published in 1921 with a metacritical appendix and by Darmstadt: Wiss. Buchges, 1975).

Philosophie und Pädagogik. Untersuchungen auf ihrem Grenzgebiet (Philosophy and pedagogy. Studies in the area of their frontiers). Marburg: Elwert, 1909.

Die logischen Grundlagen der exakten Wissenschaften (The logical foundations of the exact sciences). Leipzig/Berlin: Teubner, 1910 (later editions published in 1921, 1923).

Philosophie. Ihr Problem und ihre Probleme (Philosophy. Its problem and its problems). Göttingen: Vandenhoeck & Ruprecht, 1911 (later editions published in 1918, 1921).

Allgemeine Psychologie nach kritischer Methode. First (and single) book: *Objekt und Methode der Psychologie* (General psychology according to a critical method). Tübingen: Mohr, 1912.

Deutscher Weltberuf. Geschichtsphilosophische Richtlinien (German calling for the world. Outlines of a philosophy of history). Jena: Diederichs, 1918.

Der Idealismus Pestalozzis (Pestalozzi's idealism). Leipzig: Verlag Felix Meiner, 1919.

Sozialidealismus. Neue Richtlinien sozialer Erziehung (Social idealism. New outlines for social education). Berlin: L. Springer, 1920 (later edition published in1922).

Individuum und Gemeinschaft (Individual and society). Jena: Diederichs, 1921.

Natorp, H., ed.*Vorlesungen über praktische Philosophie* (Lecture on practical philosophy). Erlangen: Verlag der philosophischen Akademie, 1925.

Natorp, H., ed. *Philosophische Systematik* (Philosophical systematics). Hamburg: Meiner, 1958.

SECONDARY LITERATURE

Holzhey, Helmut. *Cohen und Natorp*. 2 vols. Basel: Schwabe, 1986.

Jegelka, Norbert. *Paul Natorp: Philosophie, Pädagogik, Politik*. Würzburg: Königshausen & Neumann, 1992.

Lembeck, Karl-Heinz. *Platon in Marburg*. Würzburg: Königshausen & Neumann, 1994.

Sieg, Ulrich. *Aufstieg und Niedergang des Marburger Neukantianismus*. Würzburg: Königshausen & Neumann, 1994.

Karl-Heinz Lembeck (2005)

NATURALISM

Put most succinctly, metaphysical naturalism affirms that the natural world is the only real one, and that the human race is not separate from it, but belongs to it as a part. The term *naturalism* refers also to an aesthetic style in literature, drama, and painting, and in ethics, to the theory that the full meaning of value concepts such as good and evil can be spelled out using only terms from a natural, or factual vocabulary. These are not of concern here. What follows is a discussion of naturalism in metaphysics and epistemology.

Everyone has a rough working notion of what can happen in the course of nature, and is familiar with the idea that perhaps a transcendent or supernatural realm lies beyond nature, another world that may occasionally make contact with the everyday world by, for example, miraculous interventions. Yet the distinction between what is natural and what is not needs to be made with some care. As St. Thomas Aquinas pointed out, everything that happens is in some way natural. Thus, according to classical philosophical theology, God must always, of necessity, act in accordance with his own nature. So from the divine point of view, special miraculous intervention, or general providential guidance, lies entirely within the realm of what happens according to nature.

What is needed is the conception of a world all of whose normal workings count as natural, while whatever lies outside this limit does not. The space-time world, with its material constituents working according to the laws of cause and effect, seems a good place to start. The natural world is the world of space, time, matter, energy, and causality, and naturalism affirms that this natural world is the only one there is. Yet even here care is needed: In some modern interpretations of quantum theory, the so-called Many Worlds interpretations, this particular space-time world is not by any means the only one. What lies beyond this world are other spatio-temporal realms, inaccessible from this one, perhaps evolving under different laws, but equally a part of nature in its entirety. So naturalists must allow that nature comprises this spatio-temporal world together with all other realms required by the best scientific explanations of this one.

NATURALISM AS METHOD AND AS ONTOLOGY

Because specifying what nature is brings in reference to scientific explanation in this way, naturalism is sometimes regarded as a rule of method rather than a metaphysical doctrine. There is a natural method of inquiry, which consists in setting out to explain and understand the world by finding the natural causal processes by which natural objects come into being, produce their effects, and pass away. All genuine knowledge is of this natural, experimental kind; human beings, themselves part of the natural order, have no special insight or intuition that could provide a more direct path to knowledge. And the methods of the natural sciences, which are so successful, are these natural methods refined and made more systematic.

If naturalism is in this way a matter of method in inquiry, the natural world is the world revealed by the methods of the natural sciences. This does not, in itself, place many constraints on what sort of world that might be: One cannot tell in advance what the scientific method might reveal. Maybe it will uncover not just familiar items—ships and shoes and sealing wax, for instance—but fire-breathing dragons, the Fountain of Youth, or the Philosopher's Stone. Naturalism regarded as a method maintains that ontology should be developed a posteriori—whatever is vindicated by the sciences is acceptable, whatever is not, is not.

The attempts made during the twentieth century to establish the existence of the paranormal phenomena (telepathy, precognition, and telekinesis) illustrate this approach. The methods adopted were naturalistic methods, which in themselves set no limits to what can exist.

A more affirmative naturalism goes rather further: It claims not only that the scientific method provides the only sound basis for knowledge of reality, but also that it has already established that all nature has a physical basis. The fundamental causal network consists in chains of physical cause and physical effect, produced by the operation of physical forces. All realities have at least a physical nature of this kind, whatever else may prove to be true of them. This leaning toward a physical basis for everything has been encouraged by the development of more and more sophisticated instruments for probing the observable, tangible, and manipulable world of matter, and of increasingly successful physical theories to account for what is discovered.

Yet this tendency to regard physics and chemistry as the basic and comprehensive sciences does not in itself require a materialistic ontology. Physicalism is a particularly stringent version of naturalism. A physical basis for everything does not rule out other characteristics. It is possible to affirm naturalism while insisting that the higher faculties in humans and other animals cannot be given a physicalistic reduction, and nonmaterialistic naturalism avoids the difficulties that materialism has, for

example, in accounting for the intensional characteristics, such as linguistic meaning and psychological understanding.

THE CASE FOR NATURALISM

Bertrand Russell was once charged, as Hamlet had charged Horatio, that there are more things in heaven and earth than were dreamt of in his philosophy. He retorted that he preferred it to be that way, rather than the other way around. He thus expressed the naturalist attitude, which is imbued with the spirit of Ockham's Razor: Extravagance in ontology is to be avoided. We must recognize the reality of what most plainly exists, the familiar natural world in which we live and move, and have our being. Beyond that, one should be cautious. There is no compelling evidence, of any kind, that there is more to reality than the nature revealed by scientific investigation. So the rational position to adopt is the economical, minimalist one that there are no further realms.

THE ELEATIC ARGUMENT. The Eleatic Stranger in Plato's *Sophist* proposes that "Power is the mark of Being"—that the true test of reality is to be efficacious. That which is real makes a difference, changes things, has effects. The outcome of a serious and sustained inquiry into what actually passes this test, is naturalism. For whatever operates in such a way as to alter the course of nature belongs by that very fact to the causal network of the natural world. And whatever has no such impact has no claim to reality.

The Eleatic argument is perhaps even more powerful as a methodological one: The only way in which anything can call attention to itself, and so stake a claim to reality, is by having an effect, either directly, in perception, or indirectly, through the traces it leaves in instruments. Without any such impact, there can be no reason to suppose that the thing in question exists. And that which there is no reason to think exists, should have no place in any ontology.

This argument needs to be elaborated to cover purely theoretical reasons for admitting other realms—parallel universes, for example, or sets to underpin mathematics. It is then not so straightforward to exclude higher realms, with unmoved movers, divine providences, or guardian angels. Here the argument must be that, unlike the extra worlds of quantum theory, these other worlds have no essential link to the natural explanation of what occurs in this one.

THE SELF-CORRECTING VINDICATION. Naturalism should be adopted as the proper stance in philosophy, just because it is open to development. Wherever the current conception of the world of nature is inadequate, this deficiency is likely, sooner or later, to be revealed, for there will be unaccountable phenomena that need to be accounted for. Current explanatory resources having proved inadequate, they must be expanded. New entities, properties, or forces must be recognized. The ontology of naturalism will grow to whatever extent the facts require, no more, but no less. So naturalism will always be the best philosophical stance. To maintain this position, a naturalist must show that explanatory reasoning does not advance in this way from the natural to the supernatural.

These three lines of support for naturalism all rest on a negative base: the claim that there is no valid method of discovery beyond those used in the natural sciences. So a thorough naturalism must explore, and reject, a priori reasoning in natural theology, and the claims of religious experience to provide knowledge of a transcendent divinity. It must also argue that the hermeneutic method of some social sciences, and the empathy by which humans reach a commonsense understanding of one another, does not involve entities or processes beyond those revealed by naturalistic methods.

THE IMPLICATIONS OF NATURALISM

In general, naturalism and religion are at odds with one another. Most religions posit powerful and purposeful supernatural forces, responsible for creating the natural world, for shaping its progress, and for determining the destinies of its inhabitants. These beliefs are not compatible with the naturalistic outlook. This does not, however, preclude a religious attitude accompanying naturalism, involving feelings of awe and wonder toward the natural realm, and impulses to value and care for it. Nor does it rule out a pantheism such as Benedict de Spinoza's. Spinoza identified God with Nature, insisting, as naturalists do, that there is nothing beyond this law-governed world. The atheistic varieties of Buddhism, in which this world is the only one, and where law governs the world's unfolding, would also be naturalistic religions if they were to accord independent reality to the material realm.

Naturalism requires that religious experience, and in particular mystical experience, be given a reductionist interpretation. Such experiences are regarded as unusual states of mind that have their own causes and consequences within the natural world, but do not provide any contact with, or insight into, a supernatural realm.

METAPHYSICAL IDEALISM. Naturalism takes its cue from the natural sciences, and with the exception of some more fanciful interpretations of the measurement paradoxes in quantum theory, the sciences are resolutely realist about the material world. Realism maintains that the world of nature is as it is, irrespective of any human opinions about it. The natural world is not dependent on, or brought into being by human mind, will, or experience. As this is the working philosophy of the natural sciences, it is difficult to combine naturalism with metaphysical idealism, which implies that matter is in some way a function or aspect of mind.

A thorough-going phenomenalism, such as an atheistic version of George Berkeley's philosophy, might be thought to count as an idealistic naturalism in which every object of experience does indeed belong to a law-governed spatio-temporal world, but where to be spatio-temporal is to have a derivative status, with perceptual experiences as the basic elements out of which it is constructed.

However, such a view places the experiencing mind outside the world of nature, and this puts it in conflict with one of the most profound aspects of naturalism, the view that the human species enjoys no specially privileged position in the scheme of things. Naturalism implies that human beings share with all other beings a common status, as contingent, temporary configurations in the law-governed natural world. The human world is a part of the natural realm, not a distinct cultural sphere to be contrasted with it.

Realist naturalism takes the Earth and its living inhabitants as genuine independent realities, and by locating the human race within the natural world, can make progress toward explaining how it came into being, and how humans came to have the epistemic and cognitive capacities that they do. Not even the more objective post-Hegelian metaphysical idealisms can provide any basis for an explanation of how humans came to be as they are.

THE PROBLEM OF UNIVERSALS. Realists about universals—properties and relations—divide into the Platonists, who allow the real existence of properties even where there is nothing in this world that instantiates them, and Aristotelians, who admit the reality of instantiated universals only. Plato's heaven, a higher realm containing perfect patterns for the properties imperfectly realized here below, is clearly incompatible with naturalism, and it seems probable that unless they can be vindicated as being required for explanations of what happens in this world, no system that admits uninstantiated universals can be naturalistic. Nominalist accounts of properties do not face any problem so far as naturalism is concerned.

DETERMINISM. Although naturalism stresses that it is by natural processes, involving natural causes only, that anything at all occurs, it is not committed to an absolute determinism. If there are causes at work, they are natural ones, but there may not be a cause in every case. There has to be at least enough general order in the world for it to provide an environment suitable for life and consciousness, but that admits of exceptions, here and there, to every rule. Quantum theory is not fully deterministic, as its causal relations are probabilistic. Naturalism requires there to be at least as much causality and law in the world as the development of natural sciences calls for, but it does not require any more than that.

MATHEMATICS AND LOGIC. Naturalism is almost bound to take a reductionist view of the so-called abstract objects of mathematics and logic—their numbers, functions, and relations. For the number twenty-seven, or the square root of negative one, or the relation of contrariety seem to fail both the spatio-temporal location test, and the Eleatic causal power test for natural reality. W.V.O. Quine, who was very much of a naturalistic bent, found himself forced to accept the reality of sets as a foundation for mathematics, something essential for physics, which provides the best description of the world. So sets, although not themselves naturalistic beings, have a place in the best ontology. This is a departure from pure naturalism. Hartry Field (1980), among others, has attempted to develop a philosophy of mathematics that dispenses with numbers or other mathematical objects.

The situation with the objects of geometry seems less problematic. If space-time is taken realistically (not, as with G.W. Leibniz, as a mere system of relations among physical objects), then the objects of geometry (points, lines, shapes, and geometrical solids) can be given a naturalistic home as aspects or parts of space-time.

MODALITY

The natural world comprises not only objects and the properties they actually possess. It includes what might be, yet is not (natural possibility), and what not only is but must, in the course of nature, come to pass (natural necessity). To meet this situation, the properties that things now actually possess (categorical properties) must be distinguished from those that provide the basis on which things will change and develop (dispositional

properties or powers). Aristotle and the medieval Aristotelians such as Thomas Aquinas introduced potentiality (in contrast with *act*) to specify what an object is capable of—its range of possibilities. The modern version of this is the specification of an object's powers. The power to become *F* is a different property from being *F*, but it is itself a real categorical property, perhaps some feature of the underlying fine structure of the object that possesses it.

The powers that there are in the world determine and explain what is naturally possible, should they be exercised. And where they are exercised, the powers are bound to act as they do and produce their effects. This situation is therefore one of natural necessity.

Beyond natural possibility and necessity, however, lie that which is logically possible, even though ruled out by the laws of nature (such as a ball thrown into the air and just remaining there), and that which is, not just naturally, but logically necessary or impossible. Some philosophers treat possibility and necessity by introducing possible worlds, worlds in some way additional to the one actual world. This at least seems to be a departure from naturalism because additional, merely possible worlds do not belong in the same causal network with this world, and thus fail the Eleatic test on which naturalism insists.

Naturalism therefore seems to be committed to providing an account of the logical modalities that does not involve any special ontological commitments. The proposal that logical necessity is a reflection of language, of meaning and use, was an attempt to provide such an account. The linguistic theory has fallen out of favor; more recent accounts attempt to construct possible worlds from appropriately selected sets—sets of descriptions, or unactualized recombinations of elements from the actual world. These are accounts in terms of ersatz possible worlds. Provided a naturalistic account of sets can be given, such proposals would be naturalistic theories of necessity.

OBJECTIVE MORALITY

Morality is another problematic area for naturalism. The standard naturalistic characteristics are the contingent factual actualities, and these do not include in any straightforward way the values that objects or situations may have. The size and shape of an object enters into the natural causal nexus, but its goodness does not seem to. A naturalistic account of morality must find a place for good and evil, but not in the inherent structure of the world, as a fully objective moral realism does. Nor can naturalism ground moral law in the commands of a deity. It must explain right and wrong, good and evil, as arising in the nature, preferences, or reactions of people, and in the structure of the societies within which people live out their lives. Whether an account of morality along these lines can satisfactorily explain the authority and impersonal binding force that moral imperatives seem to possess, is perhaps the most difficult issue for naturalist theories of morality.

See also Aristotelianism; Berkeley, George; Determinism, A Historical Survey; Ethical Naturalism; Evil; Field, Hartry; Leibniz, Gottfried Wilhelm; Many Worlds/Many Minds Interpretation of Quantum Mechanics; Ontology, History of; Platonism and the Platonic Tradition; Quine, Willard Van Orman; Realism and Naturalism, Mathematical; Russell, Bertrand Arthur William; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Universals, A Historical Survey.

Bibliography

Armstrong, D. "Naturalism, Materialism, and First Philosophy." *Philosophia* 8 (1978): 261–276.
Craig, William Lane, and J. P. Moreland. *Naturalism, A Critical Analysis*. London: Routledge, 2000.
De Caro, M., and D. Macarthur, eds. *Naturalism in Question*. Cambridge, MA: Harvard University Press, 2004.
Field, Hartry. *Science Without Numbers*. Oxford: Blackwell, 1980.
French, Peter A, Theodore E. Uehling, and Howard K. Wettstein, eds. *Philosophical Naturalism*. Notre Dame, IN: University of Notre Dame Press, 1995.
Maxwell, Nicholas. *The Human World in The Physical Universe; Consciousness, Free Will, and Evolution*. Lanham, MD: Rowman & Littlefield, 2001.
Nielsen, Kai. *Naturalism and Religion*. Amherst, NY: Prometheus Books, 2001.
Papineau, D. *Philosophical Naturalism*. Oxford: Blackwell, 1993.
Quine, W. V. O. "Naturalism; Or, Living within One's Means." *Dialectica* 49 (1995): 251–261.
Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.
Rea, Michael. *World without Design: The Ontological Consequences of Naturalism*. Oxford: Oxford University Press, 2002.
Shook, John R., ed. *Pragmatic Naturalism and Realism*. Amherst, NY: Prometheus Books, 2003.
Villanueva, E., ed. *Naturalism and Normativity*. Atascadero, CA: Ridgeview, 1993.

Keith Campbell (2005)

NATURALISM IN ETHICS

See *Ethical Naturalism*

NATURALISTIC RECONSTRUCTIONS OF RELIGION

See *Religion, Naturalistic Reconstructions of*

NATURALIZED EPISTEMOLOGY

Naturalized epistemology is the proposal that the theory of knowledge bears a close relation to empirical studies of cognition. The proposal was first made by W. V. O. Quine in his influential article, "Epistemology Naturalized" (1969). Quine is usually interpreted as subscribing to *replacement naturalism*: epistemology is to be replaced with empirical psychology. But his proposal may well fall short of full replacement.

In his article, Quine distinguishes conceptual studies, which seek to clarify concepts by defining some of them in terms of others, from doctrinal studies, which attempt to establish laws by proving them. The conceptual studies, were they successful, would facilitate the doctrinal ones because clarifying concepts increases the chance that truths that would otherwise go unrecognized will come to be obvious or come to be perceived as "derivable from obvious truths" (p. 70). Quine allows that progress was made in conceptual studies when Jeremy Bentham suggested paraphrasing sentences about bodies in terms of sentences about sensory experience.

Unfortunately, the project of reducing talk of bodies to talk of sensory experience together with set theory, pursued by Rudolf Carnap in *The Logical Structure of the World* (1928/1967), did not come to fruition. Carnap's later attempts at a rational reconstruction of science abandoned the aim of providing equivalences that would enable us to eliminate the terms of science in favor of sensory terms, and so they did not legitimate science. Carnap's reduction failed, according to Quine, because scientific theories do not have observational consequences except in the presence of collateral scientific theories. In view of the failure of the reduction, Quine proposes that conceptual studies seeking to clarify terms be replaced by an empirical psychology that describes how science is related to experience: "If all we hope for is a reconstruction that links science to experience in explicit ways short of translation, then it would seem more sensible to settle for psychology" (p. 78). This is the first part of Quine's proposal that empirical psychology is to enter into epistemology: Empirical studies of the cognitive development of science are to succeed the earlier reductive conceptual studies.

Regarding the doctrinal studies, Quine notes that it has been clear since Hume's treatment of induction (1739/1978) that we cannot derive scientific theories from sensory observations. Moreover, Quine claims that scientific theories have consequences for sensory experience only in the presence of collateral science (the Duhem-Quine Thesis). So scientific theories are not supported by observation alone. Since support for any scientific theory depends in this sense on further science, there is no reason to persist in the Cartesian stricture that any reliance on empirical science to understand how science is related to observation is circular. And so, for Quine, there is no point in excluding empirical psychology from such an understanding. This is the second part of Quine's proposal that empirical psychology is to enter into epistemology.

But Quine's reasoning here can be challenged on two grounds. He infers from the permissibility of relying on collateral scientific theories to support a given scientific theory that it is permissible to rely on a specific scientific theory, psychology, to understand how the given scientific theory is related to observation. But psychology is generally not the collateral scientific theory on which, according to the Duhem-Quine Thesis, we are allowed to rely for support of a given scientific theory; and the argument from the permissibility of relying on a collateral theory to *support* a given scientific theory to the permissibility of relying on psychology to *understand* how the theory is related to observation is not clearly valid. The latter challenge raises the worry that, in moving from the issue of the support of the theory by observation to the issue of understanding how the theory is related to observation, Quine makes room for psychology, but only by changing the subject from the support of the theory to understanding the relation between theory and observation. This challenge does not, however, undermine Quine's argument if he does not propose a full replacement thesis but rather the idea that empirical psychology is to figure in the project of supporting scientific theory.

The naturalized epistemology that results from Quine's proposals thus has two parts. The conceptual studies that attempted to clarify concepts by reduction are to be replaced by a psychology that understands how science is related to observations. The doctrinal studies are also to be altered. Regarding the latter, most commentators have assumed that Quine intends that we replace normative epistemology with a descriptive psy-

chology of the cognitive development of theories. Quine's summary, however, leaves room for a normative as well as a descriptive enterprise:

> Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a physical human subject. This human subject is accorded a certain experimentally controlled input—certain patterns of irradiation in assorted frequencies, for instance—and in the fullness of time the subject delivers as output a description of the three-dimensional external world and its history. The relation between the meager input and the torrential output is a relation that we are prompted to study for somewhat the same reasons that always prompted epistemology; namely, in order to see how evidence relates to theory, and in what ways one's theory of nature transcends any available evidence. (p. 83)

This passage could be taken to propose the replacement of epistemology with psychology: "evidence" could have the descriptive meaning of observation rather than a normative meaning, and seeing how theory transcends evidence might be a descriptive enterprise.

But there are other interpretations that make better sense of Quine's argument in "Epistemology Naturalized." He might mean that we are to use psychology to judge the amount of support the observations provide for given scientific theories, but only in light of an assumed epistemology (that is, an account of what support amounts to) distinct from psychology. On this interpretation, the epistemology tells us how far beyond the observations a scientific theory may go before the observations no longer support the theory; the psychology measures how far beyond our observations the theory actually goes; and the epistemology and psychology combine to tell us whether the theory enjoys support. Alternatively, Quine might mean that we are to use psychology to judge how any suggested epistemology fares in light of whether our actual achievement meets its demands. On this alternative interpretation, the results of psychology constrain epistemology. The psychology measures how far beyond our observations our scientific theories go; and a suggested epistemology is rejected if the measured distance between our scientific theories and our observations exceeds the distance the epistemology sets as the threshold for support.

This second interpretation makes the best sense of the text. Psychology contributes to an account not merely of the causal but also of the support relation between observation and theory. Thus, there is continuity between the old task of supporting science by observations and the new task of accounting for the support relation between observation and theory. The interpretation responds to the charge that Quine's argument from the Duhem-Quine Thesis to the permissibility of relying on psychology changes the subject. And the interpretation is suggested by the fact that Quine assumes that the failure of the conceptual reduction of science to observations, or of the doctrinal derivation of science from observations, does not count decisively against a positive epistemic status for science. Without this assumption, Quine would have no reason to propose that epistemology should abandon reduction and derivation for psychology, rather than that we should terminate epistemology with the judgment that science lacks support because reductions and derivations fail despite being necessary for support. On the preferred interpretation, psychology enters after the failed conceptual reductions and doctrinal derivations, but the use of psychology is warranted only by the separate epistemological claim that the success of science is jeopardized by overshooting the observations, though the jeopardy is not so rigid that the failure of reduction and derivation entails skepticism.

Quine's "Epistemology Naturalized" led many philosophers to take seriously the relevance of psychology to epistemology. Although almost all interpreters read Quine as proposing to replace epistemology with psychology, few epistemologists follow Quine in embracing replacement. Many endorse instead a *conceptual naturalism*: our everyday epistemic concepts of knowledge, justified belief, or rational belief can be defined or clarified in naturalistic terms, where naturalistic terms are usually taken to be the terms of some respectable science, notably psychology. (On a more liberal view, naturalistic terms are simply those not patently normative.)

Most proponents of conceptual naturalism hold that our epistemic concepts are both normative and naturalistic. The motivations for conceptual naturalism are not often articulated, but they presumably include these: the concepts employed by our respectable sciences are our best-understood concepts and thus the best candidates for definitions in terms we can understand and also the best candidates for clarifying definitions; these are concepts with which we cannot now dispense in our intellectual lives, so for now these concepts are clearly available to provide definitions; and these are the concepts we have the best reason to believe succeed in referring to properties that are actually exemplified, so that knowledge

defined in terms of them does not turn out inadvertently to fail to obtain.

Perhaps the most popular version of conceptual naturalism has been *reliabilism*, according to which knowledge is true belief that results from a reliable belief-forming process—a process that tends to yield true beliefs (and similarly, justified belief is belief that results from a reliable process). Whether reliabilism is fully naturalistic depends on whether the notion of truth is naturalistic. Reliabilism has been most extensively developed by Alvin Goldman. In *Epistemology and Cognition* (1986), Goldman divides his epistemology into two parts. The first part is an analysis of epistemic concepts. Both knowledge and justified belief are defined in terms of reliability. Justified beliefs are (roughly) beliefs permitted by some right J-rule system. J-rules license certain cognitive processes, and "A J-rule system R is right if and only if R permits certain (basic) psychological processes, and the instantiation of these processes would result in a truth ratio of beliefs that meets some specified high threshold (greater than .50)" (p. 106).

This analysis is supported by intuitions in narrow reflective equilibrium. The second part of Goldman's enterprise is an attempt to discover which sorts of beliefs are justified given his analysis of justified belief. This would ideally lead to discovering a right system of J-rules, but Goldman regards such an effort as premature, since "Cognitive science is still groping its way toward the identification of basic processes" (p. 181). Instead, Goldman considers candidates for basic processes individually and attempts to discern their reliability or contribution to a high truth ratio. He examines perception, memory, deduction, probability judgments, judgments under uncertainty, and belief revision in light of the findings of cognitive science. It is fair to say that his review of the reliability of cognitive processes is the most detailed and comprehensive yet undertaken. This second part of Goldman's enterprise exemplifies *methodological naturalism*: that a significant part of epistemology is an inquiry into whether conditions of epistemic status are satisfied in light of empirical cognitive science. Quine, on the interpretation of his views suggested above, is a methodological naturalist in this sense.

Reliabilism is not the only proposed version of conceptual naturalism. Alvin Plantinga (1993) offers a *proper function theory* of knowledge that is conditionally naturalistic. According to the theory, knowledge is belief that results from the proper functioning of our cognitive faculties. This theory is naturalistic if proper functioning is naturalistic, although Plantinga denies that it is. More than one writer has noted, however, that most analyses of knowledge and justified belief that eschew the label "naturalism" nevertheless meet the requirements of conceptual naturalism just as well as reliabilism does (Foley 1994, Goldman 1994). For example, some coherence theories define justified belief in nonnormative and even naturalistic terms, such as consistency, mutual entailment, and the like. Despite this, versions of reliabilism differ from coherence theories in usually resulting from an inquiry motivated by the desire to define knowledge and justified belief in natural terms. Coherence theories do not usually result from a naturalistically motivated inquiry.

A view consistent with conceptual naturalism is *property naturalism*: the property of knowledge or justified belief is *identical*with certain natural properties—properties to which respectable science refers. Ruth Millikan (1984) offers a proper function account of knowledge along these lines, for which the relevant science is evolutionary biology. A view entailed by both conceptual and property naturalism is *supervenience naturalism*: epistemic properties *supervene* on natural properties. However, it has been noted (Foley 1994) that few epistemologists have wished to deny supervenience naturalism: Roderick Chisholm (1989) allows that justified belief supervenes on nonnormative properties, despite defining it in normative terms. Keith Lehrer (1997) is rare among epistemologists in denying that justified belief supervenes on nonnormative properties.

Within the category of conceptual and property naturalism, certain views are versions of what Goldman (1994) calls *substantive naturalism*, according to which the defining terms refer to natural processes or to relations between the subject's belief and the environment. Reliabilism would fit this label. But so would John Pollock's (1989) internalist version of naturalism based on classical artificial intelligence, and Paul Thagard's (1992) coherence theory, which understands the acceptability of a scientific theory as involving a connectionist mechanism. In an influential article, Philip Kitcher (1992) emphasizes a version of *psychologism* as central to naturalism: knowledge turns on the character of psychological belief-forming processes, as opposed to logical or statistical relations between evidence and belief.

Accounts of knowledge as involving or constituted by psychological processes like intuition and demonstration, and of justified belief as involving causal inference, were common in early modern philosophy. Louis Loeb (2002) argues that, in *A Treatise of Human Nature* (1739/1978), David Hume held a stability theory of justi-

fication: justified belief is the result of a belief-forming operation that tends to produce stable beliefs. C. S. Peirce is also commonly regarded as holding a stability theory in "The Fixation of Belief" (Schmitt 2002). But this psychologism was rejected in the late nineteenth and early twentieth centuries in tandem with the rejection of psychologism in semantical theory by Gottlob Frege. Recent naturalism, such as reliabilism and proper function theory, has brought psychological processes back to the fore in accounts of knowledge and justified belief. The chief ground for giving psychological processes a role in justification has been an attack on the "arguments on paper" thesis, which sees justification as turning merely on an evidential relation between the proposition believed and the evidence possessed by the subject (Goldman 1986, Kaplan 1994).

Related to the role of psychological processes is naturalist opposition to idealized epistemology, which derives norms or standards of justification from logic, probability theory, utility theory, or statistics, without attention to human limitations. For example, epistemologists have endorsed norms of rational belief-revision such as the following: we are to avoid contradictory beliefs, and we are to believe the proposition favored by the total evidence available to us. Hilary Kornblith (2002) cites reasons for doubting that we are generally able to guarantee that our beliefs are consistent. Goldman (1986) criticizes the claim that there must be a failure of rationality if one's belief fails to conform to the total available evidence; the fault may lie in one's access to memory rather than in one's reasoning, which is the focus of the evaluation of rationality.

Of course, the norm of avoiding contradiction or of conforming to the total available evidence could be understood as the qualified requirement that we are to avoid contradiction or conform to the total available evidence when we are able to do so. But if it turns out that we are rarely if ever able to satisfy to these norms, there is little plausibility to the view that rational belief-revision requires satisfying, or even being guided by, such norms. Again, the norm of avoiding contradiction could be understood as the requirement that we are to approximate as nearly as feasible (or as cost-effective) to avoiding contradiction. But if it turns out that we are far short of being able to approximate the goal, then it seems there is no such norm. The question concerns the content of epistemic norms and an associated issue of the methodology of identifying norms: Can we formulate norms in ignorance of contingent facts about our cognitive powers, protecting the norm from empirical disconfirmation by making it merely a requirement to approximate a goal, or must we craft norms under assumptions about human limitations that would best be empirically informed (Schmitt 2004)? The naturalistic methodology finds support in the theoretical point that the approximate idealizing view has no means of suppressing epistemic ideals that are intuitively plausible but so demanding that no norm should require approximating them to any degree.

A final issue within naturalism is methodological. Should we conduct epistemology by defining knowledge so as to explain the functions (biological, social, or cognitive) served by our *concept* of knowledge and *practices* of epistemic evaluation? Or should we instead identify knowledge with the real properties involved in states we label "knowledge," studying knowledge on the model of a natural kind like aluminum or frog, opening the possibility that knowledge diverges from the properties represented in our concept, and that it serves primary functions quite different from any suggested by the functions of our use of the concept?

An account of the first sort is offered by Edward Craig (1990), who defines knowledge so as to explain the functions served by our applications of the concept. He proposes that our concept has its content in virtue of serving the social-cognitive function of picking out good informants. Craig rests his conditions of knowledge on everyday observations of the function of our concept, but it would be possible to rely on scientific sociology in such a study. An account of the second sort is offered by Hilary Kornblith (2002). He argues, by appeal to studies of animal cognition, that animals possess knowledge, and he defends the view that human knowledge is no different in kind from animal knowledge. In effect he proposes that we infer the conditions of knowledge from the biological functions of the states we label "knowledge." As it happens, Craig's conditions of knowledge roughly coincide with Kornblith's: both are versions of reliabilism about knowledge. But Craig's methodology is incompatible with Kornblith's. For conditions of knowledge inferred from the social-cognitive functions of applying the concept of knowledge need not be coextensive with conditions inferred from the biological functions of knowledge itself. Nothing guarantees that the properties that humans ascribe in order to pick out good informants must be the properties that enable animals to survive in their habitats.

It is a further question whether, given Kornblith's approach and findings, knowledge turns out to be a natural kind—for example, in the sense of a homeostatic property cluster or a cluster of self-maintaining proper-

ties. If instances of knowledge are to be instances of a natural kind and knowledge is to be reliable belief, then every instance of knowledge must be a state of a cognitive system in which a variety of reliable processes (perhaps perceptual, memorial, and inferential processes) routinely support one another in producing knowledge. This would seem to be the weakest sense in which the properties essential to an instance of knowledge could be said to be self-maintaining. If so, the claim that knowledge is a natural kind entails two key assertions: that knowledge requires not merely a reliable process yielding the given belief but also that instances of knowledge are embedded in a nexus of reliable processes. We may wonder, however, whether there is any informative condition of embedding in such a nexus that holds for all instances of knowledge across species; if not, there is no general natural kind of knowledge.

See also Carnap, Rudolf; Chisholm, Roderick; Cognitive Science; Epistemology; Frege, Gottlob; Goldman, Alvin; Hume, David; Lehrer, Keith; Memory; Millikan, Ruth; Peirce, Charles Sanders; Perception; Plantinga, Alvin; Psychologism; Quine, Willard Van Orman; Reliabilism; Underdetermination Thesis, Duhem-Quine Thesis.

Bibliography

Carnap, Rudolf. *The Logical Structure of the World*. Translated by R. George. Berkeley and Los Angeles: University of California Press, 1967. First published in 1928.

Chisholm, Roderick. *Theory of Knowledge*. 3rd ed. Englewood Cliffs, N. J.: Prentice Hall, 1989.

Craig, Edward. *Knowledge and the State of Nature*. Oxford: Oxford University Press, 1990.

Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

Feldman, Richard. "Methodological Naturalism in Epistemology." In *The Blackwell Guide to Epistemology*, edited by John Greco and Ernest Sosa. Oxford: Basil Blackwell, 1999.

Foley, Richard. "Quine and Naturalized Epistemology." In *Midwest Studies in Philosophy*. Vol. 19, *Philosophical Naturalism*, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. South Bend, IN: University of Notre Dame Press, 1994.

Goldman, Alvin I. *Epistemology and Cognition*. Harvard University Press, 1986.

Goldman, Alvin I. "Naturalistic Epistemology and Reliabilism." In *Midwest Studies in Philosophy*. Vol. 19, *Philosophical Naturalism*, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. South Bend, IN: University of Notre Dame Press, 1994.

Hume, David. *A Treatise of Human Nature*. 2nd ed, edited by P. H. Nidditch. Oxford: Oxford University Press, 1978. First published in 1739.

Kaplan, Mark. "Epistemology Denatured." In *Midwest Studies in Philosophy*. Vol. 19, *Philosophical Naturalism*, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. South Bend, IN: University of Notre Dame Press, 1994.

Kitcher, Philip. "The Naturalists Return." *Philosophical Review* 101 (1992): 53–114.

Kornblith, Hilary. *Inductive Inference and Its Natural Ground*. Cambridge, Mass.: MIT Press, 1993.

Kornblith, Hilary. *Knowledge and Its Place in Nature*. Oxford: Oxford University Press, 2002.

Lehrer, Keith. *Self-Trust: A Study of Reason, Knowledge, and Autonomy*. Oxford: Oxford University Press, 1997.

Loeb, Louis. *Stability and Justification in Hume's Treatise*. Oxford: Oxford University Press, 2002.

Millikan, Ruth Garrett. "Naturalist Reflections on Knowledge." *Pacific Philosophical Quarterly* 65 (1984): 315–334.

Peirce, C. S. "The Fixation of Belief." In *The Essential Peirce: Selected Philosophical Writings*. Vol. 1 (1867–1893), edited by Nathan Houser and Christian Kloesel. Bloomington, IN: Indiana University Press, 1992.

Plantinga, Alvin. *Warrant and Proper Function*. Oxford: Oxford University Press, 1993.

Pollock, John. *How to Build a Person: A Prolegomenon*. Cambridge, MA: MIT Press, 1989.

Quine, W. V. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Quine, W. V. "The Nature of Natural Knowledge." In *Mind and Language*, edited by Samuel L. Guttenplan. Oxford: Oxford University Press, 1975.

Schmitt, Frederick F. "Justification and Consensus: The Peircean Approach," *Protosociology: An International Journal of Interdisciplinary Research* 16 (2002): 241–286.

Schmitt, Frederick F. "Epistemology and Cognitive Science." In *Handbook of Epistemology*, edited by Ilkka Niiniluoto, Matti Sintonen, and Jan Woleński. Dordrecht: Kluwer, 2004.

Shatz, David. "Skepticism and Naturalized Epistemology." In *Naturalism: A Critical Appraisal*. South Bend, IN: University of Notre Dame Press, 1993.

Thagard, Paul. *Conceptual Revolutions*. Princeton: Princeton University Press, 1992.

Frederick F. Schmitt (2005)

NATURALIZED PHILOSOPHY OF SCIENCE

Naturalization in the philosophy of science is related to projects for naturalization in other areas of philosophy, including ethics, the philosophy of language and mind, and, especially, epistemology. So there are some general features of naturalism shared by these different philosophical projects. Still, in each of these areas the impulse to naturalization has had different motivations and a distinctive history. Projects for naturalizing the philosophy

of science were advanced independently within the Vienna Circle by Otto Neurath and in the United States by John Dewey from roughly 1925 to 1945. A decade later a philosopher of science, Ernest Nagel, familiar with both Neurath and Dewey, defended a general philosophical naturalism in his presidential address to the American Philosophical Association. And in 1969 Willard Van Orman Quine published his influential article "Epistemology Naturalized." Nevertheless, interest in naturalization in the philosophy of science dates only from the 1980s. Three influences stand out. First, a growing dissatisfaction with logical empiricism and, more generally, with any philosophy of science conceived of as the logical or conceptual analysis of scientific and methodological concepts. Second, this dissatisfaction was in part sparked by a growing interest in the history of science, particularly as employed in Thomas S. Kuhn's 1962 book *The Structure of Scientific Revolutions*. Finally, beginning in the 1970s there was a challenge from a newly militant sociology of science claiming to provide the whole story of how science works.

In thinking about science, it is usual to distinguish between the process of doing science, scientific practice, and the product of that process, usually understood as scientific knowledge. The project of naturalization applies to both processes and products. The naturalist project for examining knowledge in various special fields rejects claims to special forms of logical and philosophical analysis, preferring to employ fundamentally the same tools used by the relevant scientists themselves. But philosophers may ask different questions than those that typically concern working scientists. For example, a philosopher of science may ask how the concept of causality in quantum mechanics differs from that in classical mechanics, or how the theories and methods of classical genetics differ from those of molecular genetics. The answers will be framed in terms that can be understood by both scientists and educated laypersons. No peculiarly philosophical concepts are required. This entry will focus on the naturalizing project for understanding the process of science, including methods for certifying particular knowledge claims.

BASIC FEATURES OF NATURALIZED PHILOSOPHY OF SCIENCE

In advancing a naturalized philosophy of science, one immediately rules out any philosophy of science invoking supernatural factors, which, however, occurs only in limited contexts. More generally, a naturalized philosophy of science rules out appeal to a priori principles, including the results of logical or conceptual analysis. Positively, a naturalized philosophy of science restricts its resources to those provided by the sciences themselves. So a naturalized philosophy of science becomes a kind of theoretical science of science. Even this minimal general characterization of naturalized philosophy of science raises several problems.

First, how could one justify ruling out the imposition of a priori principles, or even appeals to the supernatural, in the philosophy of science? This would seem itself to require an a priori argument, thus violating naturalism's own prohibition against the use of a priori principles. Second, given that the content of the sciences is continually changing, how can one specify just what counts as a resource for a naturalized philosophy of science? More simply, what counts as natural in either the philosophy of science or in the sciences themselves?

Both of these problems presume that naturalism is a thesis, indeed, a metaphysical thesis. Both problems vanish if, rather, naturalism is taken primarily as a methodological stance, a determination to employ only well-established scientific findings and methods, whatever they might be. Methodological naturalism, unlike metaphysical naturalism, can be defended simply in terms of past successes, first in physics and chemistry, but also especially in biology. Evolutionary theory and modern molecular genetics have pretty much demystified the phenomena of life. This provides a scientific reason for expecting that mental phenomena and even consciousness will some day be similarly demystified. Of course, this appeal to past scientific success to justify methodological naturalism strikes most nonnaturalistic philosophers of science as circular or regressive.

NATURALISM AND NORMATIVITY

The most common objection to the whole project of naturalized philosophy of science is that, based only on scientific findings, it can at most describe actual scientific practice; it cannot provide a normative basis for distinguishing good science from pseudoscience. Naturalism, it is often argued, leads straight to relativism. Naturalists point out that this objection assumes that there exists an extrascientific criterion for demarcating good science from pseudoscience. They argue, naturalistically, that the failure to find an agreed on criterion is good evidence that no such criterion exists. Still, it is a fact that scientists and others claim to distinguish good science from pretenders to that status. Naturalists need an account of the bases for such judgments.

The usual naturalist account is that the norms operative in science are all conditional norms of the general form: If the goal is G, use method M. The justification for such norms is itself empirical, consisting of evidence that employing M is a relatively reliable means of obtaining G. This reply itself raises several problems. One is the specification of the goal, or goals, of scientific inquiry. A second problem is the threatened regress of methods, since taking the determination of whether M is a reliable means to G as itself a goal of inquiry seems to require another method of inquiry whose reliability itself must be investigated.

REALISM VERSUS EMPIRICISM

Both naturalists and nonnaturalists argue that there is a single overarching goal to scientific inquiry throughout the history of modern science. Some make similar claims for scientific method. Proposed general goals include knowledge and truth, while proposed methods include "making use of evidence." These goals and methods are, however, so general as to be nearly vacuous. Surely one must ask: Knowledge (or truths) about what? What kind of evidence? How is evidence to be used?

Historians of science and historically oriented philosophers of science have identified at least two divergent general goals that have been pursued, often explicitly, by scientists since the seventeenth century. One is broadly empiricist while the other is broadly realist. Isaac Newton's professed refusal to "feign hypotheses" and injunctions only to make inductions from the phenomena are identified with empiricism. The nineteenth-century invocation of an aether to support electromagnetic radiation was an example of scientific realism. The later nineteenth-century debate between supporters of thermodynamics and supporters of statistical mechanics is seen as a dispute between empiricists and realists regarding the existence of atoms. In the twentieth century the weirdness of quantum physics (relative to classical physics) invited empiricist responses while molecular biology seemed uninhibitedly realist. Although most naturalists tended to argue for either an exclusively empiricist or realist understanding of science, the proper naturalist response seemed to reject the demand for a single goal for all of science as objectionably essentialist and to accept the historical diversity of goals as a natural part of science as a whole. Both empiricists and realists can be said to be seeking knowledge of the natural world rather than, say, spiritual enlightenment.

Returning to the threatened regress of methods, one question is whether or not avoiding an unacceptable relativism requires a method that can be justified a priori. Naturalists again argue that the failure of philosophers of science to agree on any such method is good evidence that no such method exists. More positively, it can be argued that the general pattern of inductive reasoning is fundamentally the same for higher-level claims about the effectiveness of various methods to deliver correct judgments at the object level as it is for object-level empirical claims themselves. There need be no regress of fundamentally different methods.

Nevertheless, naturalists tend to agree that, whatever the details of various methods for certifying scientific claims, there are no methods that can be employed without assuming that some empirical conditions obtain. There are no foundational methods any more than there are foundational empirical truths that can be known with certainty. To the extent that naturalists think this stance requires philosophical justification, that justification is usually sought in an appeal to some form of pragmatism.

NATURALISM AND PRAGMATISM

It is no accident that prominent naturalists of earlier generations embraced pragmatism. Naturalism needs a philosophical orientation that makes sense of its rejection of a priori metaphysical and epistemological principles. Pragmatism provides that orientation. The relevant pragmatist doctrine begins with the rejection of any view of knowledge that requires either deduction from a priori truths or induction from incorrigible sense experience. The positive doctrine is that one always begins from the current state of what is taken to be known. From that point, anything can be questioned and subjected to experimental tests, provided that there is some basis for doubt. But not everything can be questioned at once. Universal Cartesian doubt is ruled out. Thus, in place of a foundationist picture of knowledge of either rationalist or empiricist persuasion, one has claims to knowledge regulated by a method of motivated doubt and empirical investigation. It is this general method, not any particular claims, that matters for science.

A pragmatist orientation also fits well with a typical naturalist appeal to the evolutionary history of humans as providing an understanding of the origins of human knowledge. Evolutionary survival requires early humans to have had a serviceable understanding of the world around them, including other humans. Survival did not require having beliefs that one would now regard as true. Rather, it only required beliefs that made it possible to perform appropriate actions at appropriate times. Later, humans could develop methods for questioning and

improving earlier beliefs apart from their immediate application to particular courses of action.

Some naturalist philosophers of science argue that the development of modern science itself follows an evolutionary pattern and maybe even involves evolutionary-like mechanisms. Others disagree. This dispute takes place within a naturalistic framework and the answer does follow from that framework alone. It remains an empirical question within a naturalistic approach to the philosophy of science.

RESOURCES FOR A NATURALISTIC PHILOSOPHY OF SCIENCE

The purest statement of the logical empiricist approach to the philosophy of science was that the philosophy of science is the study of the logic of the language of science. This stance automatically put the focus of the philosophical study of science on the products of scientific activity rather than on the process of doing science. The naturalist project for the philosophy of science places greater emphasis on the practice of science. A programmatic formulation would be that a naturalized philosophy of science focuses on scientists as embodied agents practicing in a particular scientific culture. The question is what broadly scientific resources are to be employed in this study. Here, there is a diversity of opinion among those pursuing a naturalist program.

Following Kuhn and others, many philosophers of science study the activity of science using primarily historical concepts and methods. Sociologists of science, including historical sociologists partly inspired by Kuhn, invoke primarily historical and sociological categories, but differ among themselves as to which historical and sociological categories to employ. They mostly agree, however, on the desirability of there being a single unified sociological account. Some philosophers of science and cognitive scientists pursue the study of scientific practice as primarily a cognitive activity, borrowing concepts and methods from the cognitive sciences. A few philosophers and economists employ concepts from economics in their studies of science practice. Finally, feminist philosophers of science, for whom the idea of scientists as embodied and socially embedded is central, introduce concepts from feminist theory into the naturalized study of science.

Here again, the proper naturalistic response to this plurality of approaches would seem not to insist on a single unified approach, but to embrace a diversity of complementary approaches as appropriate for understanding a complex phenomenon such as science.

See also Confirmation Theory; Ethics, History of; Moore, George Edward; Naturalized Epistemology.

Bibliography

Boyd, Richard. "Scientific Realism and Naturalistic Epistemology." In *PSA 1980*. Vol. 2, edited by Peter D. Asquith and Ronald N. Giere. East Lansing, MI: Philosophy of Science Association, 1981.

Giere, Ronald N. "Philosophy of Science Naturalized." *Philosophy of Science* 52 (3) (1985): 331–356.

Godfrey-Smith, Peter. "Dewey on Naturalism, Realism, and Science." *Philosophy of Science* 69 (3) (2002): S25–S35.

Hankinson-Nelson, Lynn. "A Feminist Naturalized Philosophy of Science." *Synthese* 104 (3) (1995): 399–421.

Hooker, Clifford A. "Evolutionary Naturalist Realism: Circa 1985." In *A Realistic Theory of Science*, edited by Clifford A. Hooker. Albany: SUNY Press, 1987.

Hull, David L. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press, 1988.

Kitcher, Philip. "The Naturalists Return." *Philosophical Review* 101 (1) (1992): 53–114.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Laudan, Larry. "Progress or Rationality? The Prospects for Normative Naturalism." *American Philosophical Quarterly* 24 (1) (1987): 19–33.

Nagel, Ernest. "Naturalism Reconsidered." In *Logic without Metaphysics and Other Essays in the Philosophy of Science*, edited by Ernest Nagel. Glencoe, IL: Free Press, 1957.

Nersessian, Nancy. "The Cognitive Basis of Model-Based Reasoning in Science." In *The Cognitive Basis of Science*, edited by Peter Carruthers, Stephen Stich, and Michael Siegal. New York: Cambridge University Press, 2002.

Quine, W. V. O. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*, edited by W. V. O. Quine. New York: Columbia University Press, 1969.

Rosenberg, Alex. "A Field Guide to Recent Species of Naturalism." *British Journal for the Philosophy of Science* 47 (1) (1996): 1–30.

Solomon, Miriam. *Social Empiricism*. Cambridge, MA: MIT Press, 2001.

Thagard, Paul. *Conceptual Revolutions*. Princeton, NJ: Princeton University Press, 1992.

Uebel, Thomas E. "Neurath's Programme for Naturalistic Epistemology." *Studies in History and Philosophy of Science* 22 (4) (1991): 626–646.

Ronald N. Giere (2005)

NATURAL KINDS

Whether engaged in high-level scientific activity or in the ordinary business of living, we spend a great deal of our time sorting the objects we come across into kinds. Philosophers are concerned with the *kinds of kinds* into which we sort these objects, and with the principles that

distinguish one kind of kinds from another. One kind of kinds that has loomed large in recent philosophical discussions is that of so-called *natural* kinds. And one conception of natural kinds has dominated discussion in the contemporary philosophies of science, language, and mind, and this conception will concern us here. But first, some background.

Historical discussions of natural kinds (Ayers 1981) usually start with Aristotle and his conception of the *individuals* that are members of a kind of *substance* in virtue of the fact that they share a certain property (an *essence*) with all and only the other members of that kind. This essence can be specified in a *real definition* in terms of two of the five *predicables*: genus, species, difference, properties, and accidents. To give the most famous example, the *species* human being is part of the *genus* animal, and is distinguished from other animals by the *difference* rational; thus the essence of human beings is that they are rational animals. This essence determines the *properties* human beings possess (language, for example), although some members of the kind will also possess further properties that are not so determined, and these are the *accidents* (high intelligence, say, or lustrous skin).

In reaction against this Aristotelian vision, John Locke offered the distinction between real and nominal essences. Locke distinguished between the real essence ("the being of anything whereby it is what it is") and the nominal essence ("the abstract idea which the general, or sortal … name stands for") (1975 [1689]). He argued that when we use general terms, we refer to kinds whose definition can be given entirely in terms of their nominal essence. He maintained, first, that the members of a kind share a real essence in virtue of sharing some property concerning their microstructure, and, second, that because we lack "microscopical eyes," we can never know if an entity has this property or not. He then claimed that the features constitutive of the nominal essence of an entity are nonproblematically open to our view, and that as a result only these features are capable of ensuring that our use of a term refers to the kind in question.

This view rests upon some questionable assumptions. First, it is not obvious that our reference to an entity must be secured by features that are unproblematically open to our view. Second, it is not obvious that we cannot know that an entity possesses some microstructural feature simply because that feature is not observable. The modern view of natural kinds rejects both of these assumptions.

This modern view was inspired by the writings of Saul Kripke (1980) and Hilary Putnam (1975). The numerous advances in natural-scientific knowledge since Locke's time have greatly increased our sense that we are able to know about the microstructures of things, and these advances helped lead Kripke and Putnam to reject the second of Locke's epistemological assumptions. Far more radical, however, was their rejection of his first assumption. They insisted that the reference of a natural-kind term is secured by the real essence of the kind, even if no one has any idea what this essence is. A connection with the mental lives of those who use the relevant natural-kind term remains, but is secured insteadby the requirement that they use the term with the *intention* of referring to entities of the relevant kind. More specifically, when people learn the meaning of a natural-kind term, they are presented with a sample of the kind, and their competent use of the term is then (partly) a matter of their using it with the intention of referring to anything whose nature is the same as the relevant sample. This idea of a nature has clear Aristotelian resonances, and like Aristotle, Kripke and Putnam took the nature of an entity to be identical to its real essence.

Putnam analyzed the meaning of a natural-kind term into the following four components: a *syntactic marker* (the part of speech to which it belongs, obviously "noun"), a *semantic marker* (in the case of "water," this would be "liquid"), a *stereotype* (in effect, the nominal essence, the range of observable features commonly associated with the term; in this case, "colorless, tasteless liquid," for instance), and an *extension* (the things in the world determined by the real essence of the kind, whatever that may be). In the Kripke and Putnam picture, the stereotype provides guidelines for the use of the term, but it does not fix the reference of the term in a sentence containing it. Nonetheless, use of a term that is guided by the stereotype is still genuine use—partly in virtue of the intention to refer, and partly in virtue of what Putnam called "the division of linguistic labor" (the idea that a competent user of the term would defer to relevant experts on the matter of whether something actually is a member of the relevant kind).

This account of natural-kind terms has numerous advantages. It seems to provide an easy solution to the apparent problem of incommensurability, for example. Formulations of this problem start from an assumption characteristic of logical-empiricist accounts of scientific terms, namely, that the reference of a natural-kind term is fixed by the theoretically informed general beliefs of those who use that term. Consequently, when those beliefs change to a certain degree, so does the reference of the term. In the light of Thomas Kuhn's idea that science

undergoes massive revolutions in the theoretical beliefs of scientists, it seems, once this assumption is granted, that past scientists who used the term "electrons" were not speaking about the same things that present-day scientists speak about when they use this term. By insisting that reference is fixed by the real nature of kinds and not the transitory beliefs of scientists, the theory of Kripke and Putnam allows scientific terms to refer to the same entities over time even though the relevant beliefs of scientists change massively.

In more recent years, some philosophers have started to become suspicious of attempts to extend this account to all natural-kind terms (Dupré 1993). It is questionable how far the theory is capable of handling the kinds that biologists appear to speak of, for instance. Terms of ordinary language such as "frog," "toad," "rabbit," "hare," "onion," "garlic"—terms that one might assume are both natural kind terms and of relevance to biologists—are deployed by the latter in ways that radically diverge from how they are deployed by ordinary speakers. When ordinary speakers use these terms, it seems that their intention is not to refer to the putative real essence of, say, "garlic," but rather to something that serves a certain *function* ("garlic" refers to that which serves a certain culinary purpose, for instance). One obvious response at this point is to say that these terms refer not to natural kinds but to functional kinds (Wiggins 2001), and that their reference is fixed by some description available to the users of the term. This possibility of diverging intentions suggests that one kind term might be a natural-kind term among a group of scientists (given how they use it) and a functional-kind term among a group of lay persons (given how they use it).

There are interesting questions as to whether this account of natural-kind terms contravenes or accords with a Fregean view of meaning (Evans 1973). There are also questions about the exact role that an appeal to natural-kind terms should play in arguments for an externalist account of mental content (that the content of a mental state is determined by suitably 'external' features). In addition, if the arguments for externalism that rely on the Kripke and Putnam account of natural-kind terms are sound and if a term such as "garlic" denotes a natural kind on the lips of a scientist but a functional kind on the lips of a layperson, and if the reference of functional kind terms is fixed by a description, then we seem to be saddled with the idea that scientists have a greater number of broad, externally determined mental states than laypersons.

See also Aristotle; Essence and Existence; Kripke, Saul; Kuhn, Thomas; Laws of Nature; Locke, John; Meaning; Natural Law; Proper Names and Descriptions; Properties; Putnam, Hilary.

Bibliography

Ayers, Michael R. "Locke versus Aristotle on Natural Kinds." *Journal of Philosophy* 78 (1981): 247–272.

Dupré, John. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press, 1993.

Evans, Gareth. "The Causal Theory of Names." *Proceedings of the Aristotelian Society*, supp. vol. 47 (1973): 187–208.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Locke, John. *An Essay concerning Human Understanding*, edited by P. H. Nidditch. Oxford: Oxford University Press, 1975 (1689): bk. III, ch. 3, sec. 15, p. 417.

Putnam, Hilary. "The Meaning of 'Meaning.' " In his *Mind, Language, and Reality*. Vol. 2 of *Philosophical Papers*, 215–271. New York: Cambridge University Press, 1975.

Wiggins, David. *Sameness and Substance Renewed*. Cambridge, U.K.: Cambridge University Press, 2001.

Adrian Haddock (2005)
John Dupré (2005)

NATURAL LAW

Natural Law is a long-standing and widely influential theory in ethics and legal philosophy. Because of its long and varied history, and the diversity of definitions of the term "natural," it is somewhat difficult to summarize exactly what makes a position or methodology one of natural law—at least in such a way as to neatly include all the positions and methodologies that have gone by that name. In attempting to establish a broad set of characteristics such a theory would have to possess in order to be considered natural law, it is useful then to look at the historical development of paradigmatic theories, paying attention to David Hume's advice that when trying to understand a discourse that employs the concept of "nature," we must consider what the concept is contextually being opposed to, and "the opposition will always discover the sense, in which it is taken" (Hume 2000, p. 305, n.).

In general, we can say that the traditional notion of natural law has held to the following four propositions: (1) morality is ultimately real and objective and is not relative in its primary truths to culture, subjective taste, or social agreement; (2) morality is somehow grounded in human nature, which is a specific part of the general

order of nature, and is crucial for human happiness and flourishing; (3) the normative force and obligatoriness of morality is somehow the result of this grounding and may be understood using the terminology associated with a legal code; (4) the application of reason in examining human nature, and to some extent general nature, provides evidence for the specific content of our moral obligations.

Some theories, especially contemporary ones, may not clearly fit the pattern of this list. However, this speaks to a criticism that some recent "natural law" theories are not really natural law theories at all. It is in reference to the sort of positions specified above that such criticisms are made. There is also a problem in producing such a list as to whether reference should be made to God as a divine legislator of natural law. While the original and most traditional theories of natural law do rely on a theological foundation, it is characteristic of modern and contemporary versions that they do not, and therefore theism has not been listed as a basic proposition.

ANCIENT SOURCES

It is generally held that the first complete formulation of a natural law theory was a product of Stoic philosophers. It is also generally held, however, that classical Greek philosophers made significant conceptual contributions to what became natural law. Plato suggests the first, moral realist, tenet of traditional natural law theory in proposing his division of the Forms and appearances. In taking such a strong realist position, Plato provides material for the claim that goodness, or at least good order, is fundamentally real and our knowledge of it can be directly produced through reason. In dialogues such as *Gorgias*, *Protagoras*, and *Phaedrus*, Socrates defends a notion of objective truth and knowledge over the relativistic claims of sophists, which fits the natural law emphasis on moral realism. In the *Republic*, he analogizes the virtuous person to a healthy body and state, which fits the second proposition that morality is self-rewarding, tends toward happiness, and is the proper state of being. In the *Laws*, Plato touches upon the fourth proposition by referring to a law of nature forbidding homosexual sex as unnatural, appealing to animal behavior as evidence (836c–e).

Aristotle has an even stronger claim on influencing natural law, though his contribution is contested. One writer considers natural law his "principal legacy to Christian thought" (Hastings 2000, p. 465), whereas another believes that he "figures as a natural law thinker only ambiguously and not very helpfully" (Haakonssen 1992, p. 890). Howard P. Kainz (2004) points out that the passages in Aristotle commonly used to indicate support of natural law—"Universal law is the law of nature. For there really is, as every one to some extent divines, a natural justice and injustice that is binding on all men, even on those who have no association or covenant with each other" (pp. 6–8)—come from the *Rhetoric* (1373b5–1373b15), and are embedded in a section giving advice to lawyers on how to argue cases. Aristotle suggests using the rhetoric of natural law when "the written law tells against our case" but suggests that when "the written law supports our case" it is better to argue that "trying to be cleverer than the law is just what is forbidden by those codes of law that are accounted best" (Rhetoric, 1375a25–1375b25). But though Aristotle may not be as clearly a natural lawyer as some have thought, he does bequeath three important ideas that get taken up by natural law later on. First, in the *Physics*, Aristotle speaks at length concerning teleology—the notion that all natural objects have an end they are internally driven to fulfill (their telos) and that to understand a thing we must understand the end toward which it aims (194b15–199b30). Second, in the *Nicomachean Ethics*, Aristotle applies this principle to discover the end of human beings, arguing that humans, as natural, aim at some specific highest good for humans, which he defines as happiness—virtuous, rational, satisfactory activity (1097a15–1098a15). The teleology of natural objects and a complex virtuous happiness as the end of human beings will figure prominently in later natural law formulations, particularly those of Aquinas. Third, in the *Politics*, Aristotle argues that living in a political organization is entirely natural for humans. In fact, nature implants in us a social instinct and we can tell by the fact that humans are not individually self-sufficient that the purpose of the state is to produce well-being (1253a25–1253a35). States that work for this common well-being are genuine; states that do not are "perversions" (1279a25–1279b10).

It is commonly considered, however, that the first full-fledged description of natural law arises in Stoic philosophy. In general, Stoic philosophers were drawn to the idea that the universe is controlled by a perfectly rational and fateful principle called the logos, a concept prominent in Heraclitus's thought. The logos, as a rational principle that is creative, pervades all nature, and is reflected in human beings' ability to consciously reason and express logical relations in language, unites the metaphysical, the epistemological, and the ethical. As A. A. Long (1986) writes: "[I]t is clear that logos is something which can be heard, which serves to explain things, which is common to all" (p. 145). This unity is important for a view of reason as a law that connects nature, thought, and

morality. In ethics, Zeno of Citium and other Stoics advise us to accept the logos-determined activity of the universe as right and unchangeable. It is our moral obligation to live in accordance with nature and our nature includes the instinct for self-preservation and the possession of reason (Diogenes Laertius 1925, pp. 193–197). The mostly widely cited statement of Stoic natural law, however, comes from Cicero, who wrote:

> True law is right reason in agreement with nature; it is of universal application, unchanging and everlasting; it summons to duty by its commands, and averts from wrongdoing by its prohibitions. … It is a sin to try to alter this law, nor is it allowable to attempt to repeal any part of it, and it is impossible to abolish it entirely. We cannot be freed from its obligations by senate or people, and we need not look outside ourselves for an expounder or interpreter of it. And there will not be different laws at Rome and at Athens, or different laws now and in the future, but one eternal and unchangeable law will be valid for all nations and all times, and there will be one master and ruler, that is, God, over us all, for he is the author of this law, its promulgator, and its enforcing judge. Whoever is disobedient is fleeing from himself and denying his human nature, and by reason of this very fact he will suffer the worst penalties. (1928, p. 211)

In this passage, many of the traditional characteristics of natural law theory are asserted—the appeal to reason, natural ends, and universality, the lawlike features of obligation, commandment and punishment, the connection to human nature, our internal ability to determine natural law obligations through intuition, conscience, or acknowledgement of impulses, and the reliance on God as legislator. These aspects of natural law were subject to refinements and modifications at the hands of later thinkers including Roman jurists, such as Gaius, who focused on understanding natural law as the rational underpinning of positive law; Ulpian, who applied the natural law to all animals; and Gratian, who focused on natural law being spelled out as biblical commands (Kainz 2004).

MEDIEVAL SOURCES

With St. Thomas Aquinas (1225–1274), natural law reached a summary moment and was systematized and incorporated into the dominant Christian theological tradition of the West. Aquinas is so influential on the natural law tradition that his position is often seen as paradigmatic—a response that both limits the tradition and over-theologizes it.

Aquinas begins his discussion of the nature of law in the *Summa Theologiae* by defining law in general as "a rule and measure of acts, whereby man is induced to act or is restrained from acting" (Summa, Part 2, Part 1, Question 90, Answer 1), which is immediately "nothing else but a dictate of practical reason emanating from the ruler who governs a perfect community" (2.1.91.1). Proper laws always aim toward the general good and, following Aristotle, the goal of human life and thus the common good, is happiness (2.1.90.2). Aquinas then distinguishes between four types of law. Eternal law is the very idea of how things should be and has been intended in God's mind. This idea of how things should be according to God has "the nature of a law" (2.1.91.1). The natural law is essentially the way in which human beings, as rational beings, are positioned within this divinely designed order of things, directed toward fulfilling their nature in that order. Aquinas says:

> Wherefore, since all things subject to Divine providence are ruled and measured by the eternal law … it is evident that all things partake somewhat of the eternal law, in so far as, namely, from its being imprinted on them, they derive their respective inclinations to their proper acts and ends. Now, among all others, the rational creature is subject to Divine providence in the most excellent way.… Wherefore it has a share of the Eternal Reason, whereby it has a natural inclination to its proper act and end: and this participation of the eternal law in the rational creature is called the natural law.… It is therefore evident that the natural law is nothing else than the rational creature's participation of the eternal law. (2.1.91.2)

Aquinas adds the categories of human law (specific determinations of practical regulations) and divine law (scriptural revelations of certain specifics). It is the relationship between natural law and eternal law that is most important here, however. As Aquinas sees it, the natural law is the way in which humans participate in the eternal law, by fulfilling our natural ends in the created order which is itself the expression of the eternal idea of God. The natural law is "imprinted" on us so that we have certain inclinations toward our ends but we also have reason, which allows us to perceive and choose to follow the imprinted inclinations in the proper way. In this sense, Aquinas frames natural law as objective, grounded in human nature, dependent ultimately on God as the cre-

ator of its content, understood through reason and through observation of our own innate tendencies and capacities.

When it comes to laying out the actual rules that the natural law prescribes, the first general principle is "good is to be done and pursued, and evil is to be avoided" which is coupled with the principle that "good has the nature of an end, and evil, the nature of a contrary, hence it is that all those things to which man has a natural inclination, are naturally apprehended by reason as being good, and … their contraries as evil" (2.1.94.2). Aquinas then develops from these two principles other precepts, including the duty of self-preservation, procreation and education of offspring, seeking knowledge of God, living in society, and avoiding offending others. It is here that Aquinas begins a popular tradition among natural law theorists of laying out a set of necessary and basic human goods.

Aquinas goes on elsewhere to develop more specific rules dictated by the natural law, for example, famously outlawing masturbation, noncoital sex, and homosexual intercourse as "contrary to the natural order of the venereal act as becoming to the human race" (2.2.154.11). It is also largely on the basis of natural law reasoning that the teaching of the Roman Catholic Church rules out contraception as ever morally permissible (Hastings 2000, Catechism of the Catholic Church 1994).

It is important to realize, however that there is no simplistic equation of the natural with the moral in Aquinas. In addressing the question of whether the natural law can be changed, he distinguishes between adding to and subtracting from the requirements of the natural law and also between primary and secondary principles of the natural law. Adding to what the natural law requires is not by itself any problem "since many things for the benefit of human life have been added over and above the natural law, both by Divine law and by human laws" (2.1.94.5). Subtracting from what the natural law requires, however, depends on what level of principle we are considering. The primary principles, such as the first precept of pursuing good and avoiding evil and the immediately derivative precepts of self-preservation, and so on, cannot be changed at all. The secondary principles, however, which are "certain detailed proximate conclusions drawn from the first principles" may be changed "in some particular cases of rare occurrence, through some special cause hindering the observance of such precepts" (2.1.94.5).

With this added layer of complexity, it is incumbent upon people to use their reason and to attend to circumstances in order to determine what is and is not permissible according to the secondary principles. For example, in the pursuit of procreation, it might seem eminently natural for men to have multiple wives, yet the tradition of the church is for monogamy. How to decide this question? Aquinas argues that marriage has a primary end of producing and raising children, but also a secondary end of a social function within a community:

> Accordingly plurality of wives neither wholly destroys nor in any way hinders the first end of marriage, since one man is sufficient to get children of several wives. … But though it does not destroy the second end, it hinders it considerably for there cannot be peace in a family where several wives are joined to one husband, since one husband cannot suffice to satisfy the requisitions of several wives.… (3.suppl.65.1)

Thus according to a Thomistic reading of natural law, noncoital sex to the point of climax may never be permitted but a plurality of wives might be permitted if the material resources of the husband and culture made it workable.

Finally, for understanding the immense influence of Thomistic natural law, it is important to note that human law relies on natural law for its justification and authority. While human law may add various requirements in specifics (tax codes, civil regulations, etc.) it may not subtract from primary principles. Therefore, human laws are subject to a comparative test for their justification and authority. If they conflict with the natural law, they are not just, and not true law. Aquinas says:

> Now in human affairs a thing is said to be just, from being right, according to the rule of reason. But the first rule of reason is the law of nature, as is clear from what has been stated above. … Consequently every human law has just so much of the nature of law, as it is derived from the law of nature. But if in any point it deflects from the law of nature, it is no longer a law but a perversion of law (2.1.95.2).

As with Cicero years before, this idea makes it possible to judge human laws as unjust and nonobligatory, and opens the way for the possibility of just revolutions against unjust states and human laws.

Aquinas's analysis of natural law set the stage for an ongoing debate over the nature of the relationship between morality, God's will, and God's intellect. For Aquinas, the eternal law, which was expressed in material creation, was found in God's intellect, God's perfect rea-

son. As such, natural law was not simply an edict of God's will, as divine command theorists would argue, but rather was the automatic rational relationship between a created, purposeful order and the rational beings within that order. Presumably, if God had created a different type of purposeful world than he did, there would still automatically be a derived natural law that applied to that world as a function of reason, though its specific content would be different than the existing world. In this sense, God is bound by reason, and the natural law is the immediate rational product of created order. As Aquinas writes: "the natural law is something appointed by reason, just as a proposition is a work of reason" (Summa, Part 2, Part 1, Question 94, Answer 1).

One position, credited to Gregory of Rimini, took from this view that the natural law simply illuminated which actions and goals were intrinsically good and which were intrinsically evil. As such, the natural law "demonstrates" but is not literally a law in the sense of being legislated. As Francisco Suárez (1548–1617) encapsulates it in his influential *De legibus*, Gregory's position is "that the natural law is not a preceptive law … since it is not the indication of the will of some superior; but that, on the contrary, it is a law indicating what should be done, and what should be avoided, what of its own nature is intrinsically good and necessary, and what is intrinsically evil" (1944, p. 189). Another group of theologians called voluntarists, including to various degrees Bonaventure, Duns Scotus, and most prominently, William of Ockham, were defenders of the notion that the natural law was the product of God's will, not his intellect. As such, God could make the natural law, and thus morality, be anything he wished. Suárez writes: "This is the view one ascribes to William of Occam … inasmuch as he says that no act is wicked save in so far as it is forbidden by God and that there is no act incapable of becoming a good act if commanded by God" (p. 190).

Suárez himself, however, takes a middle course between the "intellectualist" and "voluntarist" positions, which he sees as being consistent with Aquinas. Suárez claims that the natural law not only demonstrates what is intrinsically good and evil but also "contains its own prohibition of evil and command of good" (p. 191). As indicating intrinsic good and evil, the natural law cannot be said to be simply willed by God. However, this does not mean that there is no divine command to follow the natural law on top of whatever rational obligation we might have to follow it. In fact, "it is revealed by the light of natural understanding, that God is offended by sins committed in contravention of the natural law, and that the judgments and the punishment of those sins pertain to Him" (p. 207). What this means is that although right reason can show us the intrinsic moral status of actions, and somehow produces some binding moral force, it is natural law's necessary connection to (but not identity with) the divine law that provides commanding obligation. Suárez writes:

> The binding force of the natural law constitutes a true obligation; and that obligation is a good in its own way, existing in point of fact; therefore, this same obligation must proceed from the divine will, which decrees that men shall be bound to obey that which right reason dictates. … Therefore, although the additional obligation imposed by the natural law is derived from the divine will, in so far as it is properly a preceptive obligation, nevertheless … that will presupposes a judgment as to the evil of falsehood, for example, or similar judgments (pp. 196-197; 199).

Suárez thus describes a natural law that is both morally independent of God's Will but always joined by willed legislation to follow it.

The concerns over the actual obligations implied by natural law made their way into important political and cultural disputes, including the formal debate between the theologians Juan Ginés de Sepúlveda (1494–1573) and Bartolomé de Las Casas (1474–1566) over the treatment of Native Americans by the Spanish kings. Sepúlveda, appealing to Aristotle (who claimed slavery was justified by nature in his Politics), Aquinas, and Augustine, argued that the Native Americans were "barbaric … ignorant, unreasoning … sunk in vice … cruel, and are of such character that, as nature teaches, they are to be governed by the will of others" concluding "that the Indians are obliged by the natural law to obey those who are outstanding in virtue … This is the natural order, which the eternal and divine law commands to be observed … " (Las Casas 1992, p. 11-12). Las Casas, defender of the natives, relies partially on natural law ideals by arguing that the leaders of a community are obligated to seek the common good and waging war does not seek that end, and also that the Indians are not unreasoning but instead have rational, though still incorrect, defenses of their barbaric practices. For the most part, however, he gives consequentialist arguments as to why war should not be waged, arguing that war will produce much more harm than good.

MODERN SOURCES

Hugo Grotius (1583–1645) is variously credited with being the "father of modern natural law," the "father of natural rights," and the "father of international law." While Grotius spends most of his writing analyzing the nature of international war and its adjudication, his appeal to natural law leads in several influential directions. First, Grotius rejects the skeptical view (typified by classical Greek opponent of natural law, Carneades [c. 214–129 BCE]) that humans and all animals are simply driven by self-interest and that therefore all laws have their source in individual expediency, which may change as conditions do. Instead, Grotius argues in his *Prolegomena to the Law of War and Peace*, humans have "an impelling desire for society, that is, for the social life—not of any and every sort, but peaceful, and organized … this social trend the Stoics called 'sociableness'" (*Prolegomena* 6). Grotius indicates that this innate sympathy and desire for peace is central: "This maintenance of the social order, which we have roughly sketched, and which is consonant with human intelligence, is the source of law properly so called. To this sphere of law belong the abstaining from that which is another's, the restoration to another of anything of his which we may have … the obligation to fulfill promises" (*Prolegomena*, pp. 8-9). In addition to sociableness, humans also have the rational power to discriminate between alternative actions and can choose what will actually "follow the direction of a well-tempered judgment, being neither led astray by fear or the allurement of immediate pleasure, nor carried away by rash impulse. Whatever is clearly at variance with such judgment is understood to be contrary also to the law of nature, that is, to the nature of man" (*Prolegomena*, p. 10).

Second, and largely because of this innate sociality and intelligence, Grotius claims that "what we have been saying would have a degree of validity even if we should conceded that which cannot be conceded without the utmost wickedness, that there is no God, or that the affairs of men are of no concern to him." (*Prolegomena*, p. 10). While Grotius was not the first to conceptually detach the natural law from God, his arguments lead to a significant shift in natural law language, making it easier to talk about natural law as intrinsically part of being human rather than something that reflects a divine idea. In fact, Grotius's later clarification on the importance of God's will—"the law of nature … proceeding as it does from the essential traits implanted in man, can nevertheless be rightly attributed to God because of his having willed that such traits exist in us"—ends up showcasing more the belief that human nature immediately provides the law, whatever the ultimate source of human nature (*Prolegomena*, p. 11). This move will permit the disconnection of God and natural morality, while making the source of obligation to follow the law a significant problem.

Third, the shift away from specifically religious natural law is made even more rhetorically available because of Grotius's development of the concept of natural rights. In *The Rights of War and Peace*, he first describes the term "right" as signifying what is just or at least not unjust, but then he goes on to say that "there is another signification of the word RIGHT … which relates directly to the person. In which sense, RIGHT is a moral quality annexed to the person, justly entitling him to possess some particular privilege, or to perform some particular act" (1901, p. 19). While the idea that individuals can possess moral qualities that produce privileges and impose duties on others has many conceptual problems, the upshot is that it allows for a discourse of human rights that steers clear of theological connections.

The emphasis on the social nature of human beings becomes central at this point, informing as it does both the content and general character of natural law. Some will agree with Grotius that humans have a natural sociability; some will argue that humans are naturally individualistic self-maximizers who are sociable only for practicality's sake. But the philosophical import of this talk is that even though modern philosophers will generally agree that there is a more or less fixed human nature and will continue to use the phrase "natural law," they may mean significantly different things by it.

For example, Thomas Hobbes (1588–1679) argues that nature has provided humans with certain set traits, including rough physical and intellectual equality. Out of this equality come roughly equal hopes of attaining the objects of desire and thus competition over goods, resources, and honor. With no limitations on such competition, violence ensues and a "war of every man against every man" arises. In analyzing a way out of this situation, Hobbes discusses "rights of nature," "laws of nature," and other phrases associated with the natural law tradition. Yet, when we read what Hobbes says about the character of natural laws, something seems to have changed. Hobbes says that

> a law of nature, (*lex naturalis*) is a precept, or general rule, found out by reason, by which a man is forbidden to do, that, which is destructive of his life. … and consequently it is a precept, or general rule of reason, that every man ought to endeavor peace … and when he cannot

> obtain it … seek and use all helps and advantages of war. … From this fundamental law of nature … is derived a second law; that a man be willing, when others are so too … to lay down this right to all things. (1988, pp. 86–87)

What appears to be happening here, in spite of some of the language used, is not that humans have a natural law moral obligation to seek peace, in the way Grotius might have envisioned, but rather that reason teaches us that our self-interest cannot be satisfied unless we agree with each other to give up some of our liberties and make social contracts. This means that the "law of nature" is not an objective moral obligation, but rather a pure practical realization of what we have to do in order to achieve our goals. Although it is tricky to try to use contemporary language here, it seems as if Hobbes's natural law is more about factual psychological principles and pragmatic planning. He agrees with traditional natural law theorists that we have a human nature and self-preservation is the first trait of that nature, but he sees the implications of that fact to have more to do with the satisfaction of desire than moral obligation.

This seems even clearer when Hobbes reductively defines human rights of nature as liberties to act and then defines liberties as merely "the absence of external impediments" (p. 86) and then later says that "where no covenant hath preceded, there hath no right been transferred, and every man has right to every thing; and consequently, no action can be unjust" (p. 95). Contrasting sharply with the traditional natural law claim that theft, for example, is immoral, Hobbes argues that theft only has meaning, and only becomes wrong, after social covenants are set up describing it as so. So here we see a case where the language of natural law is used but the substance is one of self-interested prudence. It is not surprising here that Hobbes's phrase, "state of nature," describes a dangerous environment that reason must be used to change. Our natural state is one of horror; our happy and peaceful state is one of artifice produced by reason.

Samuel Pufendorf (1632–1694) takes an approach both similar and somewhat more traditional. He agrees that humans are naturally self-interested and likely to engage in warlike activity to acquire the things they want. However, humans also seem to go beyond nature in excessive pursuit of the basics nature has provided them—lusting more than is necessary for procreation, seeking clothes more for show than for necessity, desiring tasty food far beyond what we need for nutrition (1991, p. 34). In a vein similar to Hobbes, Pufendorf writes:

> Man, then, is an animal with an intense concern for his own preservation … incapable of protection without the help of his fellows.… Equally, however, he is at the same time malicious, aggressive, easily provoked, and as willing as he is able to inflict harm on others. The conclusion is: in order to be safe, it is necessary for him to be sociable.… The laws of this sociality … are called natural laws. On this basis it is evident that the fundamental natural law is: every man ought to do as much as he can to cultivate and preserve sociality. (p. 35)

Here, though the term "natural law" and "ought" are used, they seem to be used prudentially, not as objective moral terms. However, Pufendorf recognizes, as did Suárez, this divide between self-preserving practicality and moral obligation and brings God back in to secure obligation. "Though these precepts have a clear utility, they get the force of law only upon the presuppositions that God exists and rules all things by His providence, and that He has enjoined the human race to observe as laws those dictates of reason which He has Himself promulgated by the force of the innate light. For otherwise though they might be observed for their utility, like the prescriptions doctors give to regulate health, they would not be laws" (p. 36). Thus, Pufendorf reverts to a modified form of divine command theory in order to fasten down the lawfulness of natural law.

John Locke (1632–1704), the most important social contract theorist after Hobbes, forms yet another subtle synthesis that ends up making natural law a moral constraint on the sorts of social contracts we can legitimately produce. As in that of Hobbes, in Locke's state of nature humans have the ability to do whatever they want but unlike Hobbes, they do not have the right to do whatever they want. Locke writes:

> Yet he has not liberty to destroy himself, or so much as any creature in his possession, but where some nobler use, than its bare preservation call for it. The state of nature has a law of nature to govern it, which obliges every one: And reason, which is that law, teaches all mankind, who will but consult it, that being all equal and independent, no one ought to harm another in his life, health, liberty or possessions. For men being all the workmanship of one omnipotent and infinitely wise Maker … they are his property … made to last during his, not another's pleasure. (Locke 1960, p. 271)

So we see here that even in the state of nature there is a natural law that provides a minimum moral code, namely, not to interfere with another's body, freedom, or property, and as Locke later lays out, the natural law also provides each person the authority to enforce and punish violations of this natural law (the abuse of which leads to the need to develop an unbiased state through social contract).

What is a bit uncertain here is the role of reason and God. In one sense, Locke says that reason is the natural law, which suggests a kind of prudential characterization, but he also says that it is the fact of our being the property of God that obliges us not to harm each other, which suggest a divine origin of obligation. However, it may be that reason teaches us first the moral principle that property is sacrosanct and that this principle is what informs us that as God's property we do not have the right to harm others. Locke also says that what makes a criminal is that he chooses to live by some other rule than reason, but then states that reason "is that measure God has set to the actions of men, for their mutual security" (Locke 1960, p. 272). In his constant appeal to reason for determining the specific obligations the natural law requires of us, however, Locke seems to work with the idea that reason both teaches us the content of moral truth instrumentally (we consult it), and is the natural law itself in some way.

With these sort of modifications, revisions, and perhaps even reversals, it is not surprising that natural law as a general ethical theory began to wane and by the eighteenth and nineteenth centuries, concerns about ethical theory shifted to debates among social contract theorists, skeptics, moral sense theorists, Kantians, and utilitarians. While the early social contract theorists still used the language of natural law, other philosophers clearly challenged the language and theory explicitly.

David Hume (1711–1776) famously maintained that it is a simple logical mistake to think you can "derive" a moral obligation from a biological or psychological fact (the is/ought distinction) and argued that because of the divergent definitions of the term "natural" that "nothing can be more unphilosophical than those systems which assert, that virtue is the same with what is natural, and vice with what is unnatural" (2000, pp. 302, 305).

Immanuel Kant (1724–1804) sought moral obligation in the realm of pure reason and repudiated any connection of actual contingent human psychology with moral truth. He argued in *Groundwork of the Metaphysics of Morals* that

> everyone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command 'thou shalt not lie' does not only hold for human beings, as if other rational beings did not have to heed it … ; that, therefore, the ground of obligation here must not be sought in the nature of the human being or in the circumstances of the world in which he is placed, but a priori in concepts of pure reason. (1997, pp. 2–3)

It is worth noting, however, that one of Kant's formulations of the categorical imperative is "Act as if the maxim of your action were to become by your will a universal law of nature" (p. 31). In spite of this phrasing, this is not natural law theory. What Kant is talking about is the understanding of a law of nature as a Newtonian universal regularity and is asking us to consider whether we could logically will our maxims to have such a universal character. He writes: "The universality of law in accordance with which effects take place constitutes what is properly called nature in the most general sense … that is, the existence of things insofar as it is determined in accordance with universal laws." (p. 31) and later comments that "We must be able to will that a maxim of our action become a universal law. … Some actions are so constituted that their maxim cannot even be thought without contradiction as a universal law of nature" (p. 33).

John Stuart Mill (1806–1873) criticized the entire project of trying to couple morality with nature, arguing that virtually all of our actions alter nature in some way and that an attempt to imitate nature would have us follow a guide of cruelty (1969, pp. 373–402). Of course, Mill here is arguing against the claim that we should look to nature in the large sense as a guide to behavior rather than specifically paying attention to the narrower concept of human nature (something he did pay attention to), which indicates how the concept of "nature" as a more narrow moral guide was being used by the 1800s.

Finally, John Austin (1790–1859), the founder of modern legal positivism, argued that law

> may be said to be a rule laid down for the guidance of an intelligent being by an intelligent being having power over him. … in the largest meaning which it has … the term law embraces the following objects:—Laws set by God to his human creatures, and laws set by men to men. The whole or a portion of the laws set by God to men is frequently styled the law of nature, or natural law: being, in truth, the only natural law

> of which it is possible to speak without a metaphor. … But, rejecting the appellation law of nature as ambiguous and misleading, I name those laws or rules … the Divine law, or the law of God. (2004, p. 24)

This command theory of law undermined the position that an obligation to act followed from anything other than sheer power and thus reduced natural law to nothing more than a confusing way of referring to divine command.

CONTEMPORARY SOURCES

In the twentieth century there was a revival of interest in natural law, as seen in the works of Jacques Maritain, Elizabeth Anscombe, Yves Simon, Ralph McInerny, Russell Hittinger, Robert George, Peter Geach, Anthony Kenny, and Alisdair McIntyre. In large part, the new attention to natural law was spurred by the Catholic Church's teachings on social and moral issues, including Pope Paul VI's encyclical letter *Humanae Vitae* (1968), which drew on Aquinas's moral theories to condemn artificial birth control. Prominent among the theological and philosophical defenders of the church's natural law teaching on contraception, abortion, homosexuality, and healthcare (though not necessarily following in the Thomistic tradition) were Germain Grisez and John Finnis. Grisez published an influential commentary on Aquinas's natural law system in 1965, which inspired John Finnis's work, culminating in *Natural Law and Natural Rights* (1980).

The heart of that book is Finnis's list of basic human goods, including life, knowledge, play, aesthetic experience, sociability (friendship), practical reasonableness (intelligently choosing and affecting one's own life), and religion (concern with transcendence) (1980, pp. 85–90). These are not moral goods, but more basically goods-for-us. It is our fundamental and self-evident awareness of these basic goods that creates moral choices for us—what are we to do? How are we to use our practical reasonableness to decide what to do? Finnis then attempts to use a "natural law method" of ethics, while still only using modern (presumably not natural law) terminology, to show purely through logic and other self-evident truths what we ought to do (p. 103). He argues for a set of basic requirements of practical reasonableness, which include a coherent, rational plan for our lives, no arbitrary preferences among either the basic goods or among persons, detachment and commitment, choosing efficient methods to achieve good, a limited attention to preference satisfaction (excluding such things as theft and murder), seeking the common good, following conscience, and perhaps the most controversial principle, "one should not choose to do any act which of itself does nothing but damage or impede a realization or participation of any one or more of the basic forms of human good." (p. 118).

It is this latter principle that Finnis believes rules out any consequentialist reasoning. Consequentialist ethics, he argues, is irrational because goods cannot possibly be measured, and therefore the ends never justify the means where the means includes damaging a basic good. Once he rules out consequentialism, the principle that a basic good cannot be impeded is "self-evident" and the moral rule can be summarized as "Do not choose directly against a basic value" (pp. 119, 123). This formulation of natural law begins with empirical claims about what things it is in our nature to value and then logically tries to come to our obligations. However, with no legislator to provide the traditional source of obligation (such as Suárez's and Pufendorf's God) there remains the question of whether this theory is actually a natural law theory. Finnis himself tells us that, like scientific laws, which are actually only metaphorically laws, "'Natural law'—the set of principles of practical reasonableness in ordering human life and human community—is only analogically law" (p. 280).

Finnis seems to think that reason by itself provides obligation, but it is not clear how this is supposed to occur. Reason can help us discover what desired ends we find in our psychological constitutions and can help us determine instrumentally how to achieve those ends, but how does reason create an *obligation* to pursue any end?

This question of whether their theory is properly called natural law theory also follows the most prominent twentieth century legal theorists. Lon Fuller (1964) describes a set of eight requirements that civil law must meet in order to be considered genuine law—requirements such as generality, noncontradictoriness, and nonretroactivity. In this, he is appealing to a set of objective conditions that one may subject civil laws to as a test for true lawfulness, but he emphasizes that this test is procedural rather than substantive (Bix 1996). Ronald Dworkin (1967, 1986) argues that principles of values always govern how we produce and interpret civil laws, and so there is no fundamental separation of the realms of law and morality, but this could be essentially a descriptive claim and does not imply that there is a self-evident objective moral order to which civil laws must adhere in order to provide obligations. It is perhaps primarily in the sense of providing opposition to legal positivism that these theories are classified as natural law theories.

CONNECTIONS TO OTHER ETHICAL THEORIES

While natural law is its own set of theories, the differences between it and other ethical theories are often exaggerated and oversimplified. There are significant connections and shared assumptions. For example, although Kant explicitly rejects appealing to empirical facts about human nature to determine the moral law, he begins his moral philosophy with a teleological principle widely held by natural law theorists, stating in *Groundwork* that "in the natural constitution of an organized being … we assume as a principle that there will be found in it no instrument for some end other than what is also most appropriate to that end and best adapted to it" (p. 8). Unlike natural lawyers, however, he concludes from this that the job of reason cannot be to produce happiness, because instinct would best accomplish that. Instead, reason's purpose is to produce a good will. Kant does connect nature and law through teleology though by claiming in *Idea for a Universal History from a Cosmopolitan Point of View* that "If we gave up this fundamental principle, we no longer have a lawful but an aimless course of nature" (1963, p. 13), and concluding that "The greatest problem for the human race, to the solution of which Nature drives man, Is the achievement of a universal civic society which administers law among men" (p. 16). In view of these commitments, it might be said that Kant shares with the Stoics a view of the metaphysical and epistemological aspects of the natural law, but not the essential moral aspects.

Mill, for all his criticisms of the use of the term "natural" in moral theory (Nature), is as quick as a natural law theorist to point to empirical facts about human psychology:

> The only proof capable of being given that an object is visible, is that people actually see it … the sole evidence it is possible to produce that anything is desirable, is that people do actually desire it … No reason can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable, desires his own happiness. (1998, p. 168)

He is quick also to appeal to our consciences as guides: "The internal sanction of duty … is one and the same—a feeling in our mind; a pain, more or less intense, attendant on violation of duty. … This feeling, when disinterested … is the essence of Conscience" (p. 161).

And of course, given the natural law emphasis on the pursuit of happiness, the importance of developing character traits which lend themselves to happiness and flourishing, the fundamental desire for self-preservation, the practical need to interact with others, and the ability to apprehend our obligations through internal self-observation, we see strong shared assumptions with virtue theory, social contract theory, and intuitionism.

PROBLEMS FOR NATURAL LAW

As seen through its historical development, the primary arguments for natural law have been that it is warranted theologically, that nature or human nature somehow imply that we should act in certain ways, that reason itself simply shows us the self-evident truth of natural law, and that it is necessarily practical that we act in certain ways given our nature. Criticisms have been leveled against these arguments and other aspects of natural law theory.

First, concerns about religion: If natural law theory relies on the existence of God, then proof of God must be forthcoming before we can move on to moral metaphysics—a complicated task. However, this point would only obviously apply to those versions of natural law which require God for moral obligation and some versions of natural law do not make this assumption. Problems do arise, though for relating natural law to divine command theory. For example, if, as Grotius argues, innate human traits have been directly willed in to us by God, then God's will is the source of moral obligation and thus natural law may be only a thin technical layer between human obligation and divine command theory. If, as Aquinas seems to think, some sort of natural law would proceed automatically from whatever world God created, irrespective of God's will, then this sort of moral relationship seems to be at least as fundamental and necessary as God—a point about which voluntarists are concerned.

Philip Quinn (2000), for example, actually emphasizes the divine command elements of Aquinas's thought, arguing that the *Summa*'s exoneration of Abraham in the sacrifice of Isaac story (Summa, Part 2, Part 1, Question 100, Answer 8, Reply 3) shows that Aquinas believed "the slaying of Isaac by Abraham, which would be wrong in the absence of the divine command, will not be wrong in its presence if Abraham obeys it" (p. 62). The issue is fundamentally about whether natural moral obligations are products of pure reason, and whether this implies there is some truth or reality that does not depend entirely on God.

Second, concerns about relativism: Just as voluntarist divine command theory is often seen as a type of moral relativism because God could (in some views at least)

have made anything a moral obligation, the apparent natural law assumption that morality depends on the actual contingent facts of biology and psychology seems to make morality relative to species (rather than culture or the individual, as traditional relativisms argue). This is not a practical problem for determining obligations when there is only one sapient species to consider, but for ethicists such as Kant, morality could only be said to be truly objective if it was necessary for all possible rational beings.

Third, concerns about is and ought: Hume pointed out that many attempts at moral philosophy make a near-imperceptible shift from the way things are to the way things should be—a move that logically requires connecting premises often not given (Hume 2000, p. 302). This criticism has been analyzed at great length (Hudson 1969). Natural law may be an attempt to breach the is/ought divide, but historically it often either does nothing to supply the connection, or supplies it arbitrarily, or tries to supply the connection simply by appealing to reason. It is unclear, however, how reason is supposed to produce moral obligation. It may be true, for example, that choosing a short-term pleasure over a long-term basic good interferes with comprehensive happiness, and thus may in one sense be called unreasonable or irrational. But this sense of "unreasonable" is more a matter of acknowledging empirical constraints on what will actually satisfy our desires, health, or continued existence rather than serving as any sort of logical proof of a moral obligation.

Instrumentally, reason can help us to satisfy the desires and inclinations we do in fact naturally have, but it is not clear how reason is supposed to indicate that we should try to satisfy them. There is nothing formally illogical about not satisfying desires we have or securing our own health and happiness. For versions of natural law that retain God as a moral lawmaker, this problem seems to be avoided because obligation can been seen in a positivist sense as legislated—but then this Ockamist or Austinian approach returns us to the problem of whether natural law simply reduces to divine command theory.

Fourth, concerns about the goodness of nature: There is the assumption in natural law that human nature is fundamentally good (even though flawed), which legitimates our appeal to it. This is an inheritance of Christian theology, even for those versions of natural law that argue for no dependence on God. Other explanations, less committed to design and eternal law formulations of the world's development, see aspects of human nature as more adventitious and thus less morally authoritative. Human traits are not necessarily here because they are supposed to be but because they survived. As a result, many inherent traits may be prone to producing what we think of as evil acts and ends. As Mill writes in "Nature":

> With regard to this particular hypothesis, that all natural impulses, all propensities sufficiently universal and sufficiently spontaneous to be capable of passing for instincts, must exist for good ends … this is of course true of the majority of them, for the species could not have continued to exist unless most of its inclinations had been directed to things needful or useful for its preservation. But unless the instincts can be reduced to a very small number indeed, it must be allowed that we have also bad instincts which it should be the aim of education not simply to regulate, but to extirpate. … among them one which they call destructiveness: an instinct to destroy for destruction's sake. I can conceive no good reason for preserving this. (p. 398)

Fifth, and related to the fourth, concerns about best explanation: One of the key purposes of natural law ethics, particularly in its modern versions, is to oppose the idea that there is no human nature, or that human nature is so widely divergent that no cultural or moral norms can be said to be better or worse than any other. In this sense, natural law is opposed to cultural moral relativism, behaviorist environmental determinism, and postmodern social constructivism. However, natural law is not the only theory that holds there is a human nature, that can produce a list of basic human goods, pays attention to biology and psychology, and opposes relativism. To some extent Rawlsian contractarianism does this, but in a way even more related to natural law, evolutionary ethics does as well.

Evolutionary ethics can take seriously the claim that the moral law is "written on our hearts" and that we only need our conscience to apprehend it. As Grotius defended the existence of the natural law by pointing to widespread regularities in moral beliefs (1957, pp. 25–26), evolutionary theorists defend the existence of an evolved moral sense, which explains cross-cultural similarity in moral emotions such as guilt and shame, and cross-culturally widespread moral restrictions on murder, betrayal, and sexual infidelity. But there is a difference.

Just as evolutionary theory covered much of the same territory as the argument from Design for the existence of God, but could explain both complexity and the existence of "imperfections" such as vestigial organs (having given up a perfect designer and therefore eliminating any expectation of perfect design), evolutionary ethics

can explain both the widespread facts of human cooperation and widespread selfish violations of moral norms (having given up a perfect moral inculcator and therefore eliminating any expectation of perfect moral inculcation).

Even for nonreligious versions of the natural law, there remains the idea that our consciences and innate natures are essentially good and trustworthy and thus have some difficulty explaining why warmongering, murder, lying, addiction, and rape are both so self-evidently bad and so persistent. It seems to some then that evolutionary ethics does a better job of explaining human moral nature and human immoral nature. Of course, evolutionary ethics is at heart descriptive, arguing that moral attitudes are simply what have been successful at replication over time and not that they represent any objective moral truth (anymore than our bodies reflect imperfectly some infallible objective body). This is indeed a disadvantage if one is in search of moral prescriptions, but evolutionary ethicists can attempt moral prescription as well, having at first glance no lesser or greater obstacle to overcome in moving from facts to obligations than natural law theorists (Rachels 2000).

CONTEMPORARY STANDING

Natural law theory is still active as an applied ethics (forming as it does the foundation of the Catholic Church's moral philosophy). It is also still active in some academic investigations, generating numerous titles each year in ethics and legal philosophy. It is safe to say, however, that it is a minority position in mainstream academic ethics, at least in its traditional form, and typically appeals mostly to ethicists of particular religious bents. However, the descendants (or perhaps distant cousins?) of natural law theory thrive in the form of natural rights or human rights theory, which form the backbone of much of the world's international moral discourse—particularly when criticizing a particular state's or culture's practices. Practically speaking, though, much of the rhetoric concerning natural law in its more explicit and narrow sense (in appeals to naturalness and unnaturalness) is spent on ethical issues of sexuality and reproduction, leading some critics to claim that debates over sexual morality are actually the last stand for popular traditional natural law appeals (Mohr 2005, pp. 122–123).

See also Anscombe, Gertrude Elizabeth Margaret; Aristotle; Augustine, St.; Austin, John; Bonaventure, St.; Carneades; Cicero, Marcus Tullius; Consequentialism; Duns Scotus, John; Dworkin, Ronald; Gregory of Rimini; Grotius, Hugo; Heraclitus of Ephesus; Hobbes, Thomas; Hume, David; Kant, Immanuel; Laws of Nature; Legal Positivism; Locke, John; Logos; Maritain, Jacques; Medieval Philosophy; Mill, John Stuart; Moral Realism; Peace, War, and Philosophy; Philosophy of Law, History of; Philosophy of Law, Problems of; Plato; Pufendorf, Samuel von; Rawls, John; Rights; Social Contract; Socrates; Sophists; Stoicism; Suárez, Francisco; Thomas Aquinas, St.; Thomism; William of Ockham; Zeno of Citium.

Bibliography

Aristotle. *Nichomachean Ethics*. Translated by W. D. Ross, revised by J. O. Urmson. In Barnes, *The Complete Works of Aristotle*, Vol. 2.

Aristotle. *Physics*. Translated by R. P. Hardie and R. K. Gaye. In Barnes, *The Complete Works of Aristotle*, Vol. 2.

Aristotle. *Politics*. Translated by B. Jowett. In Barnes, *The Complete Works of Aristotle*, Vol. 2.

Aristotle. *Rhetoric*. Translated by W. Rhys Roberts. In Barnes, *The Complete Works of Aristotle*, Vol. 2.

Austin, John. "A Positivist Conception of Law" In *Philosophy of Law*. 7th ed., edited by Joel Feinberg and Jules Coleman. Belmont, CA: Wadsworth/Thomson Learning, 2004.

Barnes, Jonathan, ed. *The Complete Works of Aristotle: The Revised Oxford Translation*. 2 vols. Princeton, NJ: Princeton University Press, 1984. A sixth printing with corrections published in 1995.

Bix, Brian. "Natural Law Theory." In *A Companion to Philosophy of Law and Legal Theory*, edited by Dennis Patterson. Oxford, U.K.: Blackwell, 1996.

Catechism of the Catholic Church. New York: Doubleday, 1994.

Cicero. *The Republic*. Translated by Clinton Walker Keyes. Loeb Classical Library. Cambridge, MA: Harvard University Press, 1928.

Copleston, Frederick. *A History of Philosophy*. Vol. 3, *Late Medieval and Renaissance Philosophy: Ockham, Francis Bacon, and the Beginning of the Modern World*. New York: Doubleday, 1993.

Diogenes Laertius. *Lives of Eminent Philosophers*. Translated by R. D. Hicks. Loeb Classical Library. Cambridge, MA: Harvard University Press, 1925.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Belknap Press of Harvard University Press, 1986.

Dworkin, Ronald. "The Model of Rules," *University of Chicago Law Review* 35 (14) (1967). Reprinted in *Philosophy of Law*. 7th ed., edited by Joel Feinberg and Jules Coleman. Belmont, CA: Wadsworth/Thomson Learning, 2004.

Finnis, John. *Natural Law and Natural Rights*. Oxford, U.K.: Clarendon Press, 1980.

Fuller, Lon. "Eight Ways to Fail to Make Law." In *Philosophy of Law*. 7th ed., edited by Joel Feinberg and Jules Coleman. Belmont, CA: Wadsworth/Thomson Learning, 2004.

Fuller, Lon. *The Morality of Law*. New Haven, CT: Yale University Press, 1964.

Grisez, Germain. *Christian Moral Principles: The Way of the Lord Jesus*. Chicago: Franciscan Herald Press, 1983.

Grisez, Germain. “The First Principle of Practical Reason: A Commentary on the Summa Theologiae 1–2, Question 94, Article 2.” *Natural Law Forum* 10 (1965): 168–196.

Grotius, Hugo. *Prolegomena to the Law of War and Peace*. Translated by Francis W. Kelsey. Indianapolis: Bobbs-Merrill, 1957.

Grotius, Hugo. *The Rights of War and Peace, Including the Law of Nature and of Nations*. New York: M. Walter Dunne, 1901.

Haakonssen, Knud. “Natural Law.” In *Encyclopedia of Ethics*. Vol. 2, edited by Lawrence C. Becker and Charlotte B. Becker. New York: Garland, 1992.

Hastings, Adrian. “Natural Law.” In *The Oxford Companion to Christian Thought*, edited by Adrian Hastings, Alistair Mason, and Hugh Pyper. Oxford, U.K.: Oxford University Press, 2000.

Hobbes, Thomas. *Leviathan*. Oxford, U.K.: Oxford University Press, 1988.

Hudson, W. D. *The Is-Ought Question: A Collection of Papers on the Central Problems in Moral Philosophy*. London: Macmillan, 1969.

Hume, David. *A Treatise of Human Nature*. Oxford Philosophical Texts Edition, edited by David Fate Norton and Mary J. Norton. Oxford, U.K.: Oxford University Press, 2000.

Kainz, Howard P. *Natural Law: An Introduction and Re-examination*. Chicago: Open Court, 2004.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. Cambridge, MA: Cambridge University Press, 1997.

Kant, Immanuel. “Idea for a Universal History from a Cosmopolitan Point of View.” In *On History*, edited by Lewis White Beck. New York: Macmillan, 1963.

Las Casas, Bartolomé de. *In Defense of the Indians*. Translated by Stafford Poole. Dekalb: Northern Illinois University Press, 1992.

Locke, John. *Two Treatises of Government*. Cambridge, MA: Cambridge University Press, 1960.

Long, A. A. *Hellenistic Philosophy: Stoics, Epicureans, Sceptics*. 2nd ed. Berkeley: University of California Press, 1986.

Mill, John Stuart. “Nature.” In *Essays on Ethics, Religion and Society*. Collected Works of John Stuart Mill, edited by J. M. Robson. Toronto: University of Toronto Press, 1969.

Mill, John Stuart. “Utilitarianism.” In *On Liberty and Other Essays*, edited by John Gray. Oxford, U.K.: Oxford University Press, 1998.

Mohr, Richard. *The Long Arc of Justice: Lesbian and Gay Marriage, Equality, and Rights*. New York: Columbia University Press, 2005.

Paul VI, Pope. *Humanae Vitae*. Papal Encyclical. Available from http://www.vatican.va/holy_father/paul_vi/encyclicals/documents/hf_p-vi_enc_25071968_humanae-vitae_en.html. 1968.

Plato. *Complete Works*, edited by John M. Cooper. Indianapolis: Hackett, 1997.

Pufendorf, Samuel. *On the Duty of Man and Citizen According to the Natural Law*. Cambridge, MA: Cambridge University Press, 1991.

Quinn, Philip L. “Divine Command Theory.” In *The Blackwell Guide to Ethical Theory*, edited by Hugh LaFollette. Oxford, U.K.: Blackwell, 2000.

Rachels, James. “Naturalism.” In *The Blackwell Guide to Ethical Theory*, edited by Hugh LaFollette. Oxford, U.K.: Blackwell Publishing, 2000.

Suárez, Francisco. *Selections from Three Works*. Translated by Gwladys L. Williams, Ammi Brown, John Waldron, and Henry Davis. Oxford, U.K.: Clarendon Press, 1944.

Thomas Aquinas, St. *Summa Theologica*. 2nd, rev. ed. Translated by the Fathers of the English Dominican Province, 1920. Available from http://www.newadvent.org/summa/. Online edition copyright 2003 by Kevin Knight (15 February 2005).

Patrick D. Hopkins (2005)

NATURAL SELECTION

See *Evolutionary Theory*

NATURE, PHILOSOPHICAL IDEAS OF

In its widest sense “nature” can mean “the totality of things,” all that would have to appear in an inventory of the universe. It can also refer to the laws and principles of structure by which the behavior of things may be explained. These two senses cannot be kept independent of each other at any sophisticated level of inquiry, for to state in any of the sciences what an entity is involves describing what it does, its patterns of activity or behavior, and the activity of its constituent elements, as far as they can be known and subsumed under laws.

In a particular philosophical context the sense in which nature is being used can be brought out most clearly by insisting upon the question “What is nature (or the natural) being contrasted with in this context?” In one group of cases the natural is contrasted with the artificial or conventional. This contrast requires some conception of how the object or organism would behave by reason of its immanent causality alone, the causal factors that are peculiar to that type of thing and make it whatever it is—a stone, a fish, or a man. The artificial and conventional are seen as interferences, modifying by an alien causality the characteristic patterns of behavior. In the sphere of human nature this distinction is at the center of an ancient and continuing controversy, for it is by no means easy—if, indeed, possible—to delineate a human nature free of interferences, left to itself. Organism and environment, individual and cultural climate, are in ceaseless interplay. An activity (like moral evaluation or social organization) that seems to some theorists on the “con-

vention side" of the boundary may be represented by others, with no less reason, as a development of natural potentialities. The controversy is further complicated by the intrusion of evaluative nuances in the distinction itself, so that the natural, for instance, may come to be more highly esteemed than the artificial and conventional, as the spontaneous or the basic is contrasted with the labored and derivative. The preference may be reversed, however; the natural can be taken as the mere raw material, the unfinished and preparatory, requiring artifice to complete and crown it.

In some contexts man is contrasted with nature; in others he is taken as part of nature. The difference is not trivially linguistic. To set man against nature is to emphasize his distinctiveness—his rationality, creativity, and freedom. But it may also support an unwarranted and distorting anthropocentricity. To count man as part and parcel of nature emphasizes the continuity of the human, animal, organic, and inorganic worlds and suggests that human behavior may be amenable to the same kinds of investigation that are effective in studying other domains of nature. Similarities as well as differences can be exaggerated, however, and overfacile generalizations can be made from the behavior, say, of rats to human behavior. Human distinctiveness and complexity may be overlooked in a tempting reductive analysis like that of behaviorism.

In still other contexts the natural world, man included, is contrasted with the supernatural. In part at least, the idea of the supernatural has tended to be constructed from allegedly miraculous events, events that, it is claimed, the power and laws of nature could not bring about. (There can be also an a priori element in the grounding of belief in the supernatural. Belief in a transcendent creator-God, who may be himself the subject of a priori proofs, implies the belief that nature's laws and processes can be overruled.)

It is anything but easy, however, to elaborate coherently the nature-supernature distinction. Crucial to it is the claim that we can distinguish what lies within the capacities of nature from what lies beyond them. Our knowledge of nature's powers and laws is itself derived from our experience and observation of events. What we judge to be possible depends upon what we have reason to believe actually occurs or has occurred. When we assemble the experiences out of which we are to construct these judgments about the possible, what shall we do with the happenings that, eventually, we wish to label miraculous? To exclude them would be to imply that we *already* know what nature's powers are, that there are criteria prior to experience by which we interpret our observations. But to include them makes it impossible for us to treat them later as miraculous exceptions to natural laws.

Certainly, it is not legitimate to move from saying, "This event is inexplicable in terms of our scientific knowledge of nature," to saying, "This event must be a supernatural intervention." The scientist is by no means committed to claiming that he has at any particular moment the concepts and theories adequate for every explanatory task. He is constantly revising and adding to these. We are not, therefore, forced to conclude that an event has a supernatural source on the grounds that it is inexplicable or anomalous in terms of present-day science. Indeed, it is only with the help of an independently established set of beliefs about God that one could plausibly interpret an event as supernatural. (See P. H. Nowell-Smith, "Miracles," in A. G. N. Flew and A. MacIntyre, eds., *New Essays in Philosophical Theology,* New York, 1955; and A. G. N. Flew, *Hume's Philosophy of Belief,* London, 1961.)

Although it has been implied above that God must be conceived in contradistinction to nature, this is true only if God is transcendent, not immanent (or, if immanent, then transcendent as well). In a pantheistic view if nature may be distinguished from God, it is only as different views or aspects of one and the same reality.

HISTORICAL TRANSFORMATIONS

The history of philosophical ideas of nature almost coincides with the history of philosophy itself. Where a philosophy is at all systematic, even if it is avowedly antimetaphysical, it cannot avoid stating or implying some interpretation of nature. This makes it impossible to compress the history of these interpretations into one entry. The comments that follow are thus no more than indications that the philosophers named made significant contributions to the development of the idea.

When the Ionian pre-Socratic philosophers asked, "What is nature?" they assumed that the question demanded an answer in terms of a primitive substance or substances out of which the world is constructed. One of the more reasonable answers was that of Anaximander, who claimed that the ultimate world stuff must be indeterminate and indefinite (apeiron) and could not be identified with familiar stuffs like water, air, and so on. But although plausible, Anaximander's answer was also unhelpful precisely because the apeiron lacked all determinateness and explanatory power. Far more fruitful was the Pythagorean concern not primarily with the question "What is nature made from?" but with "What is its struc-

ture?" where "structure" means geometrical form. We need to know only that the constituents of the world are able to receive mathematically describable form, and the way is opened for investigating how natural objects are related, in detail, to their underlying geometrical structure.

To Plato the possibility of knowledge of nature (or of the natures of things) rests on the intelligibility of the Forms that things imitate (or in which they participate). The creation story in the *Timaeus* (which came to have enormous influence) represents God and the Forms as distinct from each other, the spatiotemporal world—mutable nature—being created after the model of the eternally unchanging Forms. It is a world necessarily deficient in important respects; the very existence of time makes it unstable and incomplete. On the other hand, it is the product of a *divine* creativity. God in his goodness does not withhold being from anything that might exist, and thus nature displays his fecundity. Here is the initial statement of the vision of nature as a great chain, or ladder, of being.

Aristotle's Unmoved Mover stands to nature as its final or teleological cause, inspiring nature to imitate the divine activity as far as its various constituents are able. Particular things, therefore, are seen as striving to realize their appropriate forms, and in so doing, they realize their own natures. Underlying this view of nature is a clear analogy with biological growth.

To Christian thinkers the primary distinction has, of course, been between the underivative creativity of God and the derivativeness and dependence of nature. Augustine, for instance, contrasts the divine "first cause that causes all and is not caused itself" with "the other causes" (the world of nature) that "both cause and are caused" (created spirits) or are primarily passive effects, corporeal causes (*City of God* V, 9). This does not preclude a wider use in which mutable spatiotemporal nature is contrasted with divine nature, "the Nature which is immutable is called Creator" (*Epistolae,* 18, Sec. 2). In Thomas Aquinas, too, God can be called *natura naturans* and the contrast made with *natura naturata,* the creating contrasted with the created nature (*Summa Theologiae* IIa–IIae, 85, 6).

It was the Pythagorean-Platonic strand in philosophy of nature that furthered and came to dominate the rise of modern science. In Johannes Kepler, for example, nature appears as the realm of the quantitative, a realm amenable to mathematical study and, indeed, to more precise study than ancient philosophy ever demonstrated. Such a view of nature could coexist with a religious interpretation of things, for the mathematical structure could be taken as supplied and sustained by the mind of God.

Although in one way the growth of a mathematical science promised most impressively to unify nature by bringing widely diversified phenomena under laws, in another way it produced new problems about the relation of man to his world, problems that led to various dualisms—bifurcations of nature—such as René Descartes's. Those aspects of our experience that were not amenable to exact measurement were no longer to be identified with objectively real, accurately cognized features of the world. The measurable qualities were primary, the rest secondary, qualities—colors, sounds, tastes, and the like. Although materialist metaphysics boldly attempted (and still attempts) to reunite nature and man by describing the full range of his perceptual, moral, and imaginative life in terms of matter and motion, in a writer like Thomas Hobbes, for example, such explanations were only promissory notes. A great deal of development in physiology had to occur before the details of the mechanisms involved could be conjectured with any real plausibility.

Descartes gave the world of mind distinct ontological status alongside corporeal nature. Although this dualism saved mind from loss of reality or reduction to the nonmental, it introduced the problem, unsolvable in Cartesian terms, of how this bifurcated nature can yet be one, how the processes of mind and of matter can impinge on each other. The philosophies of nature in Benedict de Spinoza and Gottfried Wilhelm Leibniz both try strenuously to deal with this problem. Spinoza affirms a monistic and pantheistic position (*Deus sive natura*), but the dualism breaks out again in the inexplicable relation between extension and thought—a dualism not of substances but of attributes. In Leibniz's pluralist world the relation between material and mental aspects of monads is no more intelligible.

George Berkeley's account of nature involves a radical criticism and rejection of the notion of material substance. Our experience could, he argued, be explained simply in terms of minds and their ideas, including, crucially, the divine mind, in which the totality of sensible things exists.

In the philosophy of Immanuel Kant the burden of creativity further shifts to the human percipient. If we ask Kant why nature presents to us the persistent basic structure that it does present (such as the ubiquity of cause-effect relations and the spatiotemporal nature of all experience), his answer is that we are here dealing with the inescapable conditions for any experience of nature at

all because "the understanding is itself the source of the laws of nature" (*Critique of Pure Reason,* A 127). The natural world, in the sense of the totality of things, is not in Kant's view a given whole, not an object of knowledge; for instance, whether we try to show that the world is finite or infinite, our thought runs into an impasse.

In G. W. F. Hegel the dominant language is of development, nisus, toward the realization of Absolute Spirit, the end for which nature exists. Necessary transitions, logical rather than temporal, are made from level to level, from nature as inert matter with its externality to life, consciousness, the inwardness of spirit. Subsequent philosophies of nature, however, like those of Henri Bergson, Samuel Alexander, and A. N. Whitehead, were avowedly evolutionary, understandably so in an age that saw rapid development of the biological sciences, particularly biological evolutionary theory, and that had a new historical consciousness of human existence. Alexander saw the evolutionary process as the continuing "emergence" of the qualitatively new: God was to be conceived not as the initial creator or sustainer of nature but as the extrapolation of the evolutionary process to an ideal limit.

Theories involving a life force or other speculative, teleological accounts of nature have been strenuously opposed by various forms of materialism and antimetaphysical positivism.

USE OF ANALOGIES. Successive conceptions of nature (like conceptions of the state) can be seen as a procession of images or controlling analogies. Dominant in Greek cosmology, for instance, was the image of nature as suffused with life and intelligence, like a living and growing organism. At the opposite pole, as in some seventeenth- and eighteenth-century cosmologies, nature is pure machine, directed from without by the divine intelligence. Or, again, nature is neither permeated by mind nor is it a mechanism in the hand of its Mechanic; it is a self-transforming system, essentially temporal, whose development is best understood through the analogies of biological evolution or human history. To make explicit the guiding analogy is an important step in appraising an account of nature. For example, it is a standing temptation for a philosopher who is working out such an account to overextend an explanatory principle that is proving dramatically fruitful in some limited area of investigation to make it seem to cover nature as the totality of things and processes.

NATURE AS NORM

Corresponding to different philosophies of nature are markedly different answers to questions about the relation of nature to value: Can values be in any way derived from descriptions of nature? does nature set any norms for man? can appeals to nature and the natural properly settle moral or aesthetic perplexities? Various answers to these questions have been suggested in naturalistic ethical theories and in discussion of the naturalistic fallacy.

If, on the one hand, nature is seen as irreducibly complex, the theater not of a simple cosmic process but of countless and diverse processes, and if these processes have produced mind but are not themselves guided by intelligence, then there will be little plausibility in arguing directly from "natural" to "good" or "obligatory."

On the other hand, where nature is taken as created by a wholly good, wise, and omnipotent deity, to be natural is prima facie, to be *worthy* of being created by such a deity. But the existence of evil, however accounted for, makes the inference, even in this context, unreliable. The natural man may now be contrasted with the regenerate man, and "natural" thus come to have a depreciatory sense. Alternatively, the sinful can be held as unnatural—that is, as perverting the divinely appointed course of nature. The question "What *is* natural?" cannot now, however, be answered from a simple inspection of what actually happens in the world.

HISTORICAL EXAMPLES. The demand that we should follow nature occurs in a wide variety of ethical theories, not only in Christianity. It was against an ethic of following nature that J. S. Mill eloquently argued in his "Essay on Nature" (in *Three Essays*). To Mill nature means either (1) "the sum of all phenomena, together with the causes which produce them" or (2) those phenomena that take place "without the agency … of man." Which of these senses can be intended when someone is enjoined to follow nature or when some act is condemned as unnatural? In the first sense *every* action is natural; no ground is given for discrimination between alternative courses. But is the second sense more helpful? "For while human action cannot help conforming to Nature in the one meaning of the term, the very aim and object of action is to alter and improve Nature in the other meaning." Behind the injunction to follow nature lies a dim belief that "the general scheme of nature is a model for us to imitate." Look at nature in some detail, however. Its processes are quite indifferent to value and desert. "Nearly all the things which men are hanged or imprisoned for doing to one another, are nature's every day per-

formances." Even if it were true that some good ends were ultimately and obscurely served and realized by nature's processes, that would give no license to men to follow nature as a moral exemplar (to "torture because nature tortures," for example).

In any case, Mill argues, the presence of evil and indifference to value in nature cannot be reconciled with theistic claims about the omnipotence and perfect goodness of God. It is nonsense to argue that such a God has to bend to stubborn necessities since he "himself makes the necessity which he bends to."

With regard to *human* nature, as with nature at large, Mill's imperative is "not to follow but to amend it." Morality cannot be founded on instinct but on a strenuously achieved victory *over* instinct, as courage is a victory over fear. Similar views are found in T. H. Huxley and even, with important qualifications, in the later Sigmund Freud.

Philosophical views of nature can be relevant to problems of evaluation in much more complex ways than we have thus far noted. One's conception of how man is related to the rest of the natural world may help to determine—in conjunction with many other factors—one's sense of the importance or unimportance of human life, the roles judged reasonable and unreasonable for men to adopt. Here are some historical examples.

Did a geocentric astronomy give a uniquely privileged place to Earth and to humanity? The symbolism was ambiguous; to be in the center was certainly to be the focus of the cosmic drama of fall and redemption. "Man is but earth," said John Donne. "'Tis true; but earth is the centre" ("Sermon Preached at St. Paul's, Christmas Day, 1627"). Yet the center, the sublunary region, was nevertheless the humblest position, the realm of mutability, in contrast to the unchanging heavens. The shift to a heliocentric view was not, therefore, a catastrophic and disorienting demotion. It could be seen as an equally effective symbolic expression of creatureliness, Earth being placed in a proper subordination to the sun (for example, see Nicolas Copernicus and Kepler). "The sun, seated on his royal throne, [does] guide his family of planets" (Kepler, *De Revolutionibus,* Book I, Ch. 10).

A far more radical shift in sixteenth- and seventeenth-century cosmology was the move toward acceptance of the universe as infinite and with that the obliterating of a locatable center or circumference. But this view, which, in fact, had no effective *scientific* backing, was largely a late development of the metaphysical Platonic idea of God's infinite fecundity, a view that also guaranteed humanity a position of dignity in the ladder of being (see A. O. Lovejoy, *The Great Chain of Being,* Ch. 4). This well shows how (at least in a period of metaphysical confidence) the importance or unimportance of man has not been a matter of attempted inference from observations of nature alone.

The same point can also be illustrated from sixteenth- and seventeenth-century arguments about the alleged "cosmic fall." If nature is inclement and hostile, this is because nature participated in the effects of man's fall into sin. It follows that the proper, God-intended destiny of man cannot be found in this fallen nature; it must be discovered in the revealed word of God.

More generally, reference to man's place in nature, for instance to his physical minuteness, could be used to depreciate the quest for "worldly" glory as a preparation for spiritual discipline. "Who can be great," asked Drummond of Hawthornden, "on so small a Round as is this Earth?" And Blaise Pascal asked: "Qu'est ce qu'un homme dans l'infini?" ("What is a man in face of the infinite?"). The vastness of nature could equally well be taken as evidence of man's importance in God's eyes; for on independent theological grounds the whole of nature could be seen as primarily a dwelling place for man. As Pierre de la Primaudaye expressed it, "I cannot marvell enough at the excellencie of Man, for whom all these things were created and are maintained." Most of these arguments, with their ingredients capable of endless variation, assume that "in order to form a correct estimate of ourselves we must consider the results of the investigations … into the dimensions and distances of the spheres and stars" (Maimonides)—*mutatis mutandis* for later cosmologies.

In sharp contrast, at a time when there is little or no metaphysical and theological confidence and when deriving value judgments from statements of fact is deemed logically impossible, it is tempting to deny that accounts of nature can have any bearing on problems of value. F. P. Ramsey wrote: "My picture of the world is drawn in perspective, and not like a model to scale. The foreground is occupied by human beings, and the stars are all as small as threepenny bits" (*Foundations of Mathematics*). It is possible to make one's judgments about the value of human life independently of cosmic reflections and then to adopt an imaginative picture of the natural world that harmonizes rather than conflicts with that evaluation. There can be no logical or philosophical objections to that as long as one realizes exactly what is being done. Such an imaginative exercise, however, must be distinguished from a thoroughgoing anthropocentric philoso-

phy of nature, and Ramsey himself has been criticized for falling into exactly that (see J. J. C. Smart, *Philosophy and Scientific Realism,* New York, 1963, p. 25). For Ramsey went on to say: "I don't really believe in astronomy, except as a complicated description of human … and possibly animal sensation."

It is worth noting, finally, that arguments about aesthetic judgments have also relied on the vocabulary of *nature* and *natural* and relied on it in many differing and conflicting ways. Presenting or being true to nature has sometimes meant the faithful mirroring of the empirical world *or* the pursuit of the ideal type *or* the pursuit of the average type or a concern with whatever has not been modified by man (see A. O. Lovejoy, *Essays in the History of Ideas,* "Nature as Aesthetic Norm"). Works of art have been commended as sharing the characteristics of nature through being regularly patterned (compare to nature's mathematical intelligibility), through being rich in content, or through being austerely simple. To be natural can be to show spontaneity, to be unfettered by artificial rules, to reach toward the unspoiled and primitive. Where there is such extraordinary conflict of senses, only a scrutiny of the context can determine what criteria are being applied in any particular case, and a writer who is aware of this web of ambiguities in "natural" and "nature" may well decide to choose—wherever possible—words of greater precision and stability of meaning.

See also Aesthetic Judgment; Alexander, Samuel; Anaximander; Augustine, St.; Bergson, Henri; Berkeley, George; Copernicus, Nicolas; Cosmology; Descartes, René; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Kant, Immanuel; Kepler, Johannes; Laws of Nature; Leibniz, Gottfried Wilhelm; Lovejoy, Arthur Oncken; Maimonides; Mill, John Stuart; Natural Law; Pascal, Blaise; Plato; Ramsey, Frank Plumpton; Smart, John Jamieson Carswell; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Whitehead, Alfred North.

Bibliography

Because of the almost unlimited scope of the topic, references are for the most part confined to works mentioned in the text of the article or to which the article is in some general way indebted.

BACKGROUND TO GREEK PHILOSOPHIES OF NATURE

Aristotle. *Metaphysics.* Edited by W. D. Ross. rev. ed., 2 vols. Oxford, 1924. Δ (V), 4.

Bumet, John. *Greek Philosophy: Thales to Plato.* London, 1914; paperback ed., 1962.

Crombie, I. M. *An Examination of Plato's Doctrines.* 2 vols. New York: Humanities Press, 1962–1963. Vol. II, Ch. 2.

Guthrie, W. K. C. *A History of Greek Philosophy,* Vol. I. Cambridge, U.K.: Cambridge University Press, 1962.

Heinemann, F. *Nomos und Physis.* Basel: F. Reinhardt, 1945.

Kirk, G. S., and J. E. Raven. *The Presocratic Philosophers.* Cambridge, U.K.: Cambridge University Press, 1957.

Pohlenz, M. "Nomos und Physis." *Hermes* 81 (1953): 418–438.

Popper, Karl. *The Open Society and Its Enemies.* 2 vols. London: Routledge, 1945. Vol. I, especially Ch. 5, on nature and convention.

Ross, W. D. *Aristotle.* London, 1923; 2nd ed., 1930. Ch. 3.

WIDE-RANGING HISTORICAL INTERPRETATIONS

Burtt, E. A. *The Metaphysical Foundations of Modern Physical Science.* London: K. Paul, Trench, Trubner, 1925.

Collingwood, R. G. *The Idea of Nature.* Oxford: Clarendon Press, 1945.

Lovejoy, A. O. *Essays in the History of Ideas.* Baltimore: Johns Hopkins University Press, 1948. See "Nature as Aesthetic Norm."

Lovejoy, A. O. *The Great Chain of Being.* Cambridge, MA: Harvard University Press, 1936.

Lovejoy, A. O. and George Boas, eds. *Primitivism and Related Ideas in Antiquity.* Baltimore: Johns Hopkins University Press, 1935. See "Some Meanings of 'Nature.'"

OTHER STUDIES

Dewey, John. *Experience and Nature.* La Salle, IL: Open Court, 1925; paperback ed., New York, 1958.

Maritain, Jacques. *La philosophie de la nature.* Paris: Téqui, 1935.

Mill, John Stuart. *Three Essays on Religion.* London, 1874; reprinted, 1904.

Ramsey, F. P. *The Foundations of Mathematics.* London: K. Paul, Trench, Trubner, 1931; paperback ed., Paterson, NJ, 1960. See epilogue.

Russell, Bertrand. *Mysticism and Logic.* London: Allen and Unwin, 1917. See especially Ch. 3, "A Free Man's Worship."

Russell, Bertrand. *Why I Am Not a Christian, and Other Essays,* edited by Paul Edwards. New York: Simon and Schuster, 1957.

Sherrington, Charles. *Man on His Nature.* Cambridge, U.K.: Cambridge University Press, 1940.

Weizsäcker, C. F. von. *Die Geschichte der Natur.* Göttingen, 1948. Translated by F. D. Wieck as *The History of Nature.* London: Routledge, 1951.

Whitehead, Alfred North. *The Concept of Nature.* Cambridge, U.K.: Cambridge University Press, 1920.

Whitehead, Alfred North. *Nature and Life.* Cambridge, U.K.: Cambridge University Press, 1934.

Ronald W. Hepburn (1967)

NEGATION

Negation, or denial, is the opposite of affirmation. It may be something that somebody does ("I deny what you have said") or the answer "No" to a question, but its full expression is generally a sentence. One sentence or statement may be the negation or denial of another, or we may

call a statement simply a negation, or a negative statement, as opposed to an affirmative one, or affirmation. A negation in the last sense will contain some sign of negation, such as the "not" in "Grass is not pink" or "Not all leaves are green," the "no" in "No Christians are communists," or the phrase "it is not the case that" in "It is not the case that grass is pink." The negation of a sentence may simply be the same sentence with "it is not the case that" prefixed to it, or it may be some simpler form equivalent to this. For example, it might be said that "It is not the case that grass is pink" is negated or denied not only by "It is not the case that it is not the case that grass is pink" but also by the plain "Grass is pink" and that "If he has shut the door, it must have been open" is negated or denied by "He could have shut it even though it was already shut."

Contradictory negation, or contradiction, is the relation between statements that are exact opposites, in the sense that they can be neither true together nor false together—for example, "Some grass is brown" and "No grass is brown." Contrary negation, or contrariety, is the relation between extreme opposites (which may very well both be false)—for example, "No grass is brown" and "All grass is brown." Incompatibility is the relation between statements that cannot both be true, whether or not they stand at opposite ends of a scale ("This is black all over" is incompatible with "This is green all over" as well as with "This is white all over"). Incompatibles imply one another's denials (what is black all over is not green all over or white all over).

Some of these technical expressions apply to terms as well as to statements. The terms *black, green,* and *white,* for example, are incompatible; nothing can be more than one of these at once, at least not at the same time, at the same point, from the same angle, and so on. There are also "negative terms," usually formed by prefixing "non" or "not" to the corresponding positive term—for instance, *nonred, not-red.*

The concept of negation is closely related to that of falsehood, but they are not the same. Sometimes it is the negation that is true and the corresponding affirmation that is false. But in denying a statement, we implicitly or explicitly assert that the statement in question is false, though, of course, the assertion that something is false may itself be true.

There is also a connection between the concept of negation, especially as applied to terms, and that of otherness or diversity. What is not red is other than anything that is red, and what is other than anything that is red is not red. The class of things that are other than all the things included in a given class—that is, whatever exists besides the members of that class—constitutes the remainder or complement of the given class.

INTERNAL AND EXTERNAL NEGATION

When a proposition is complex, it is often important to distinguish the negation of the proposition as a whole ("external" negation) from propositions resulting from the negation of some component or components of it ("internal" negation). The Stoics noted, for example, that the contradictory denial of an implication "If *p*, then *q*" should not be formulated as "If *p*, then not-*q*" but as "Not (if *p*, then *q*)"—"That *p* does not imply that *q*." "If *p*, then *q*" and "If *p*, then not-*q*" are not even incompatible, although when they are both true, it follows that the component *p* (since it has contradictory consequences) must be false. Again "Not (*p* and *q*)," which is true as long as *p* and *q* are not true together, is not to be confused with "Not-*p* and not-*q*," which is true only if *p* and *q* are both false and is equivalent to "Neither *p* nor *q*"—that is, "Not (*p* or *q*)." "Either not-*p* or not-*q*" is similarly equivalent to "Not (*p* and *q*)." These relations between the internal and external negations of "and" and "or" statements are called De Morgan's laws, although they were well known to the medieval Scholastics long before the birth of the nineteenth-century logician Augustus De Morgan.

Some of the distinctions made in the preceding section are now commonly treated as special cases of external and internal negation. For instance, propositions with negative terms are thought of as involving the negation, not perhaps of internal propositions strictly so called, but of internal "propositional functions" ("open sentences")—for example, "Every non-*A* is a non-*B*" may be paraphrased as "For any *x*, if it is not the case that *x* is an *A*, then it is not the case that *x* is a *B*"; the difference between "No *A* is a *B*," the contrary opposite of "Every *A* is a *B*," and the contradictory opposite of the latter, "Some *A* is a *B*" or "Not every *A* is a *B*," is perhaps simply that between the internally negated form "For every *x*, if *x* is an *A*, then not (*x* is a *B*)" and the external negation "Not (for every *x*, if *x* is an *A*, then *x* is a *B*)." It is obviously possible to place a sign of negation either inside or outside a variety of other qualifying phrases; for example, we may distinguish "It will be the case that (it is not the case that *p*)" from "It is not the case that (it will be the case that *p*)" and "It is thought that (it is not the case that *p*)" from "It is not the case that (it is thought that *p*)."

By the use of open sentences all the varieties of negation are reduced to the placing of "not" or "it is not the case that" before some proposition or proposition like expression, the whole being either contained or not con-

tained within some wider propositional context. This reduction assumes that with the basic singular form "x is an A" or "x ϕ's" there is no real distinction between the internal negation "x is not an A" (or "x is a non-A") or "x does not ϕ" and the external negation "Not (x is an A)" or "Not (x ϕ's)." When the subject "x" is a bare "this," such an assumption is plausible, but when it is a singular description like "The present king of France," we must distinguish the internal negation "The present king of France is not bald" (which suggests that there is such a person) from the external negation "It is not the case that the present king of France is bald" (which would be true if there were no such person). The thesis that all forms of negation are reducible to a suitably placed "it is not the case that" can be maintained only if the last two cases have an implicit complexity and may be, respectively, paraphrased as "For some x, x is the sole present king of France, and it is not the case that x is bald" and "It is not the case that (for some x, x is the sole present king of France and is bald)."

POSITIVE PRESUPPOSITIONS

It is sometimes held that no negation can be bare or mere negation and that whenever anything is denied, some positive ground of denial is assumed, and something positive is even an intended part of what is asserted. It is trivially true that even in denials, such as that grass is pink, something is made out to be the case—namely, that it is not the case that grass is pink. But something more than this is usually intended by the contention.

One thing that could be meant is that every denial must concern something which, whatever else it is not, is itself and, indeed, simply is (exists). We have seen that some types of denial—"This is not a man" and "The man next door does not smoke" (also "Some men do not lie")—do assert or presuppose the existence of a subject of the denial. But this does not seem to be the case with all forms; for example, no existing subject seems to be involved when we say that there are no fairies. Or if this is taken to mean that among existing things no fairies are to be found (thus presupposing a body of "existing things"—of values for the bound variable x in "For no x is it the case that x is a fairy"), even this positive presupposition seems absent from "There could not be round squares."

It is also sometimes said that in denying that something is red, we at least assume that it is some other color (counting white, black, and gray as colors); in denying that something is square, we assume that it is some other shape. In general (to use the terminology of W. E. Johnson), in denying that something has a "determinate" form of some "determinable" quality, we assume that it has some other determinate form of it. Sometimes a distinction is made at this point between the predication of a negative term and the simple denial of a predication; for example, it is argued that in saying that a thing is nonblue, we do assume that it is some other color but we do not assume this in simply saying that it is not blue. Others contend that we assume that a thing is some other color even in simply denying that it is blue. All denial, it is said, is implicitly restricted to some universe of discourse; if we deny that something is blue or classify it as nonblue, it is assumed that we are considering only colored things.

Against the weaker form of the theory that the predication of a negative term has positive implications which the denial of a predication does not have, it may be objected that there is no more than a verbal difference between "x is a non-B" and "Not (x is a B)." Against the stronger form the objection is that it is perfectly proper to say that virtue is not blue simply on the ground that it is not the kind of thing that could have any color at all. We must always distinguish between what we say and our reasons for saying it (otherwise, there could be no inference at all, as premises and conclusion would coalesce), and there may be diverse reasons for saying exactly the same thing of different subjects—Jones's favorite flower is not blue because it is pink, and virtue is not blue because being an abstraction, it is not colored at all. But it is perfectly true of each of these subjects, and true in the same sense, that it is not blue.

It may be answered that "This flower is blue" and "Virtue is blue" fail to be true in profoundly different ways—the former because it is false, and the latter because it is meaningless, as meaningless as, for example, "Virtue is but" would be—and, further, whereas the denial of a false statement is true, the denial of a meaningless form of words (that is, the result of attaching a negation sign to it) is itself a meaningless form of words. To this, one possible reply (made by J. M. Shorter in "Meaning and Grammar") would be to deny that the negation of a meaningless form of words is meaningless; even "Virtue is not but" might be defended as true precisely because it is not only false, but also meaningless, to say that virtue is but. Less desperately, it could be argued that "Virtue is (is not) blue" is not on a par with "Virtue is (is not) but" since the former is at least a grammatically correct sentence while the latter does not even construe. Perhaps, however, the conception of grammar that suggests this distinction is a rather superficial one. Grammar

concerns what words go with what; it is not a set of commands directly fallen from heaven but reflects at least partly the feeling we already have for what does and what does not make sense. Perhaps we need only let this feeling lead us to slightly finer distinctions than the crude one between an adjective and a conjunction to see that "is (is not) blue" no more goes with "virtue" than "is (is not) but" goes with anything.

What is important is the line between falsehood (the negation of which is true) and nonsense (the negation of which is generally agreed to be only further nonsense), wherever this line be drawn. It is also important that what looks like true or false sense may on closer inspection turn out to be nonsense.

NEGATIVE FACTS

Many philosophers who have found negation a metaphysically embarrassing concept have expressed this embarrassment by denying that there are any negative facts. There are obviously negative as well as affirmative statements, but according to these philosophers, it is incredible that the nonlinguistic facts that make our statements true or false should include negative ones. (The linguistic fact that there are negative statements is, of course, not itself a negative, but a positive, fact.)

This question should not be confused with the question of whether there are objective falsehoods—that is, whether the universe contains such objects as the falsehood that Charles I died in his bed even if no one has ever believed or asserted this falsehood (whether there are falsehoods which are, as it were, waiting around to be asserted or believed, or even denied or disbelieved, just as there are facts waiting to be discovered and stated). For such objective falsehoods, if there were any, would not be facts—a fact is what is the case, not what is not the case. The present question is, rather, whether there are special facts that verify true negative statements, whether, for example, there is any such fact as the fact that Charles I did not die in his bed. There is nevertheless some connection between the two questions. For if there is any such language-independent and thought-independent fact as the fact that it is not the case that Charles 1 died in his bed, then, that Charles I died in his bed, which in itself is not a fact but a falsehood, would nevertheless seem to have some kind of existence "out there" as a constituent of this more complex object that is a fact.

In both cases, moreover, what deters the philosophers is partly the multiplicity of the objects involved. They cannot believe that there should be not only the fact that Charles I died on the scaffold but also, over and above that fact, the additional facts that he did not die in his bed, that he was not immortal, that he did not die by drowning, and, furthermore, the facts that he did not die in his bed of appendicitis, that he did not die in his bed of consumption, that he did not die by drowning in six minutes, that he did not die by drowning in six and a half minutes, and so on. This causes an embarrassment of the same sort as the idea that, over and above the fact that he died on the scaffold, there are "out there" the falsehoods that he died in his bed, that he was immortal, that he was drowned in six and a half minutes, and so on.

The most obvious way to reduce this excessive metaphysical population, and the one taken by Raphael Demos (one of the main opponents of negative facts), is to hold that what makes it false to say that Charles I died in his bed and true to say that he did not, false to say that he died by drowning and true to say that he did not, and similarly with all the other alternatives is simply the one positive fact that he died on the scaffold. Against this, however, it may be said that what is asserted by any true statement would seem to be some fact, and the true statement that Charles I did not die in his bed does not assert that he died on the scaffold (even if this is also true). It may be suggested that what the true statement asserts is that Charles died in some positive way that was incompatible with his dying in his bed. This suggestion has the disadvantage (*a*) that it only exchanges negative facts for facts that are vague and general in the way that assertions about something or other (but nothing in particular) are always vague and general and that philosophers who are uneasy about the former (because whatever is real must be particular and positive) are likely to be equally uneasy about the latter. The suggestion also presupposes (*b*) that there are facts of incompatibility—for example, the fact that Charles I's dying on the scaffold is incompatible with his dying in his bed and that these would seem, like straightforwardly negative facts, to contain objective falsehoods as constituents and would have the same dismaying multiplicity as negative facts or objective falsehoods do.

One way of answering objection (*b*) is to argue that the facts of incompatibility which explain the truth of negative statements never concern incompatibilities between propositions but always concern incompatibilities between qualities, like the incompatibility between red and blue or between one way of dying and another. This is to make a certain sort of internal negation the fundamental form in terms of which all other types of negation are to be defined. This eliminates the horde of positive falsehoods that are incompatible with the actual

positive facts in favor of a possibly smaller and anyway more acceptable horde of incompatible qualities, each capable in itself of qualifying a real object but unable to do so at the same time as the others. But although there is some plausibility in accounting for simple singular negations in this way (that is, in taking the simple "*x* is not *A*" to be true, when it is true, because *x* is something incompatible with being *A*), it is hard to deal similarly with the negations of more complex forms—for example, "Not everything is *A*" or "It is not the case that if *x* is *A*, then *y* is *B*."

Difficulties in dealing with more complex negations also arise with the suggestion that the facts that verify negative statements are facts not so much about incompatibility as about otherness. It is important to note that the otherness account cannot take quite the same form as the incompatibility one; although the fact that *x* is something incompatible with being red will suffice to verify "*x* is not red," "*x* is something other than red" will not, for *x* may be something other than red (for instance, round) and be red as well. The otherness account would have to claim that what verifies "*x* is not *A*" is the fact that *x* is other than everything that is *A*. This account, like the preceding one, seems to be applicable only to simple singular negation. However, if the complexities that can arise are capable of being listed, it might be possible to give a separate account of the negation of each kind of complexity. Thus, having said what the simple "*x* is not *A*" means, we may say that in forms like "Not (not-*p*)," "Not (*p* and *q*)," "Not (*p* or *q*)," "Not (everything ϕ's)," and "Not (something ϕ's)" (that is, "Not anything ϕ's"), the apparently external "not" is to be defined in terms of a comparatively internal "not" as follows:

Not (not-*p*) = *p*,
Not (*p* or *q*) = (Not-*p*) and (not-*q*),
Not (*p* and *q*) = (Not-*p*) or (not-*q*),
Not (for every *x*, *x* ϕ's) = For some *x*, not (*x* ϕ's),
Not (for some *x*, *x* ϕ's) = For every *x*, not (*x* ϕ's).

In any given complex formed in these ways the innermost negations—the only ones that remain when all the reductions have been performed—will be simple singular negations explainable as above in terms of otherness or incompatibility.

NEGATION, FACTS, AND FALSEHOOD

Another way of eliminating negative facts might be by defining negation in terms of disbelief or falsehood. Affirmative statements, we might say, express beliefs whereas negative ones express disbeliefs. Disbelief, however, is not just the absence of belief, and like belief it must have an object—it must be disbelief in something or disbelief of something—and it must be justified or unjustified; if justified, whatever justifies it must be either a negative fact or whatever we replace negative facts with when using some other and more objective method of dissolving them.

In terms of falsehood we might say that the contradictory negation of a statement is the statement that is true if the given one is false and false if the given one is true. This amounts to defining negation by means of its truth table, a course advocated by Ludwig Wittgenstein in the *Tractatus.* To this it may be objected that talk of the statement which is true when a given statement is false and false when it is true is legitimate only if we know that there is one and only one statement which meets these conditions, and this seems unlikely; for example, since "Oxford is the capital of Scotland" is false in any case, "Either Oxford is the capital of Scotland or grass is not green" is true if "Grass is green" is false and false if it is true, but what is stated by this complex does not seem to be simply the negation of "Grass is green." It may also be objected that statements are not simply true and false in themselves, as if truth and falsehood were simple properties requiring no further explanation. By the usual definition "Grass is green" is true if grass is green and false if it is not, but to say this is to define falsehood in terms of negation rather than vice versa.

Perhaps the whole problem about negative facts—and the problem about the objective falsehoods that would be parts of such facts if there were any—arise from thinking of facts (and falsehoods) too literally as objects or entities. It is not merely that there are no negative facts but, rather, that there are no facts. That is, expressions of the form "The fact that *p*" do not name objects, whether or not our "*p*" is negative in form. The word *fact* has meaning only as part of the phrase "it is a fact that" (that is, "it is the case that"), and "It is a fact that grass is (or is not) green" is just another way of saying the simple "Grass is (or is not) green." "There are negative facts" is true and, indeed, makes sense only if it means "For some *p,* it is not the case that *p.*" But in this sense it is true and metaphysically harmless; it does not mean that there are objects called "That *p*" which go through a performance called "not being the case," and still less does it mean that there are objects called "The not-being-the-case of that *p.*"

Even with this caution, however, one can sensibly inquire whether signs of negation are really indispensable—whether what we say when we use them cannot also be said, and more directly, without them—and whether signs of negation are not just convenient abbreviations

for complex forms into which no such signs enter. Putting the question in this way, modern logic has evolved other devices for eliminating negation besides the ones thus far mentioned, devices which are worth examining, even though they are a little technical, and which require some preliminary account of negation as the logician sees it.

LAWS OF NEGATION. Negation figures in formal logic primarily as the subject of certain laws, of which the best known are those of contradiction and excluded middle. The law of contradiction asserts that a statement and its direct denial cannot be true together ("Not both *p* and not-*p*") or, as applied to terms, that nothing can both be and not be the same thing at the same time ("Nothing is at once *A* and not-*A*"). The law of excluded middle asserts that a statement and its negation exhaust the possibilities—it is either the case that *p* or not the case that *p*—or, as applied to terms, that everything either is or is not some given thing—say, *A*. Each of these laws may be put in the form of an implication, or "if" statement; the law of contradiction then appears as "If *p,* then not not-*p,*" and the law of excluded middle as "If not not-*p,* then *p.*" Sometimes the combination of these two, "*p* if and only if not not-*p,*" is called the law of double negation.

Each of these laws involves a number of derived or related laws. From the law of contradiction it follows that what has contradictory consequences is false; if *p* implies *q* and also implies not-*q* (and so implies "*q* and not-*q*"), then not-*p.* From the law of excluded middle it follows that what is implied by both members of a contradictory pair is true; if *p* implies *q* and not-*p* equally implies *q,* then *q.* Again, because of the law of contradiction whatever implies its own denial is false, for if *p* implies not-*p,* it implies both *p* and not-*p* (since it certainly implies *p*) and thus cannot be true. This is the principle of *reductio ad absurdum.* To take an ancient example, if everything is true, then it is true (among other things) that not everything is true; hence, it cannot be the case that everything is true. Perhaps we can also argue that if it is a fact that there are no negative facts, then that is itself a negative fact; thus, it cannot be that there are no negative facts. Correspondingly, from the law of excluded middle it follows that whatever is implied by its own denial (that is, what we are compelled to affirm even when we try to deny it) is true. (The later Schoolmen called this the *consequentia mirabilis.*)

Another important law involving negation is the law of contraposition, or transposition, that if *p* implies *q,* then the denial of *q* implies the denial of *p* or, for terms, if every *A* is a *B,* then every non-*B* is a non-*A*. If this is combined with the first law of double negation ("If *p,* then not not-*p*"), we obtain "If *p* implies not-*q,* then *q* implies not-*p*"; if it is combined with the second law of double negation ("If not not-*p,* then *p*"), we obtain "If not-*p* implies *q,* then not-*q* implies *p,*" and with both we obtain "If not-*p* implies not-*q,* then *q* implies *p.*"

Many logicians have questioned the law of excluded middle and the laws associated with it. In particular, the intuitionist logic of L. E. J. Brouwer and Arend Heyting contains none of the laws "Either *p* or not-*p,*" "If not not-*p,* then *p,*" "If *p* implies *q* and not-*p* also implies *q,* then *q,*" "If not-*p* implies *p,* then *p,*" "If not-*p* implies *q* (not-*q*), then not-*q* (*q*) implies *p.*"

FORMAL DEFINITIONS OF NEGATION. The laws just discussed and many others figure in modern symbolic calculi as theorems derived by stated rules of inference from given axioms. Some of them, indeed, may themselves appear as axioms, different formulas being taken as axiomatic in different symbolic presentations. The symbols used, moreover, will be divisible into "primitive" symbols that are introduced without explanation and other symbols that are introduced by definition as abridgments of complexes involving other symbols. Which symbols are taken as primitive and which are defined will vary with the particular systematic presentation adopted.

Gottlob Frege, for example, took symbols corresponding to "if" and "not" as undefined and introduced the form "*p* or *q*" as a way of writing "If not-*p,* then *q*" ("Either I planted peas, or I planted beans" = "If I did not plant peas, I planted beans"). Bertrand Russell at one stage did the same, but he later took "not" and "or" as his primitives, defining "If *p,* then *q*" as "Either not-*p* or *q*" ("If you smoke, you'll get a cough" = "Either you won't smoke, or you'll get a cough") and "*p* and *q*" as "Not either not-*p* or not-*q.*" Other writers have defined all the other symbols in terms of "not" and "and." For example, they have defined "If *p,* then *q*" as "Not (*p* without *q*)"—that is, "Not (*p* and not-*q*)" and "*p* or *q*" as "Not both not-*p* and not-*q.*"

In all these examples the negation sign appears as one of the primitive or undefined symbols, but there are also systems in which this is not the case and in which "not" is defined in terms of something else. For example, Jean Nicod uses a single undefined stroke in such a way that "*p* | *q*" amounts to "Not both *p* and *q*" and "Not-*p*" is defined as "*p* | *p*" (Not both *p* and *p*). Russell sometimes attempts to avoid even the appearance of complexity in

his verbal rendering of Nicod's stroke by reading "*p* | *q*" as "*p* is incompatible with *q*," but this would ordinarily be understood as a little stronger than what is intended. We would not normally say that "London is the capital of England" was incompatible with "Berlin is the capital of France," but it is correct to say "London is the capital of England | Berlin is the capital of France," since the two components are not both true.

An earlier and more interesting device was that of C. S. Peirce, who defined negation as the implication of something false. This is not quite a definition of negation in terms of falsehood. Formally, what is meant is that we arbitrarily choose some false proposition—say, "The ancient Romans spoke Polish"—and introduce "Not-*p*" as an abbreviation for "If *p*, then the ancient Romans spoke Polish." It is also possible to take as our standard false proposition for this purpose a formula which itself has some logical significance. In his later years Peirce himself liked to use the proposition "For all *p*, *p*," which is, roughly, "Everything is true" (which was shown to be false in the previous section of this entry). In common speech we come close to defining "Not-*q*" as "If *q*, then for all *p*, *p*" when we say of something we wish to deny, "If you believe that, you would believe anything." A similar definition of "Not-*p*," used by Russell in his early writings, is "For all *q*, if *p*, then *q*." Starting in this way, it is possible to define all the symbols of logic in terms of "if" and the quantifier "for all *x*." Certain further technical devices make it possible to define both "if" and "for all *x*" in terms of a single operator that can be read as "For all *x*, if …, then …" or "If ever …, then …" (Russell's "formal implication," perhaps better called "universalized implication").

Given definitions of this type, the characteristic laws of negation fall into place as special cases of the characteristic laws of implication or of universality (or both). For instance, the law of transposition, "If (if *p*, then *q*), then (if not-*q*, then not-*p*)," expands to "If (if *p*, then *q*) then if (if *q*; then anything-at-all), then (if *p*, then anything-at-all)," which is just a special case of the law of syllogism, "If (if *p*, then *q*), then if (if *q*, then *r*), then (if *p*, then *r*)." Moreover, the peculiarities of the intuitionistic negation of Brouwer and Heyting turn out simply to reflect those of intuitionistic implication.

Intuitionistic logic, for example, contains the law "If *p* implies *q*, then if *p* also implies that *q* implies *r*, *p* implies *r*"; therefore, it contains the special case "If *p* implies *q*, then if *p* also implies that *q* implies the falsehood, then *p* implies the falsehood"—that is, "If *p* implies *q*, then if *p* also implies not-*q*, then not-*p*." But it does not contain the law "If *p* implies *r*, then if *p*'s implying *q* also implies *r*, then *r*" (this law, being verified by the usual truth-tables for "if" and "not," does appear in nonintuitionistic or classical implicational logic) and therefore does not contain the law "If *p* implies *r*, then if *p*'s implying the falsehood also implies *r*, then *r*" ("If *p* implies *r*, then if not-*p* also implies *r*, then *r*").

It is also possible in both intuitionistic and classical logic to separate those laws of negation which are (or may be represented as) merely special cases of laws of implication (as in the above examples) and those that reflect the special features of what a proposition is being said to imply when we negate it. For example, both versions of logic contain the law (1) "If *p*, then if also not-*p*, then anything-at-all." But neither logic contains as a law the implicational formula of which this would be (if they had it) a special case, "If *p*, then if *p* implies *r*, then anything-at-all." However, they do both have, quite naturally, (2) "If *p*, then if *p* implies that everything is true, then anything-at-all." To get (1), in other words, it is important not only that we should see "Not-*p*" as something of the form "If *p*, then *r*" but also as this particular thing, "If *p*, then everything is true." If we drop from intuitionistic logic those laws of negation which require attention to this more special point, we obtain the "minimal" calculus of I. Johannson ("Der Minimalkalkül," *Compositio Mathematica,* Vol. 4, 119–136).

TECHNICAL ELIMINATIONS OF NEGATION

Do the developments just sketched mean that we can dispense with negative facts by saying that the facts stated by true negative statements are ones that do not involve any special concept of negation but only (in one version) Nicod's stroke or (in the other) implication and universality? The suggestion, especially in its Peircean form, has its attractions. Peirce's definition would at least explain why negation is a proper subject of study for pure logicians. Logic studies universal rules of implication; even the purest logic must study whatever is involved in the very notions of implication and universality; and what Peirce means by negation is thus involved. Facts as to what is not the case are in this view only an instance of a more general type of complex fact without which logic would be impossible—namely, facts as to what leads to what.

Against this suggestion one might adduce the extreme artificiality and arbitrariness of these symbolic devices. Consider the fact that it is equally possible in a symbolic system to define "and" in terms of "or" and

"not" and "or" in terms of "and" and "not." Whatever this fact signifies, it cannot signify that "Not (not-*p* or not-*q*)" is the real meaning of "*p* and *q*" and that the very form "*p* or *q*" that is used in this explanation has for its real meaning "Not (not-*p* and not-*q*)." This procedure would obviously be circular, and for this reason we cannot, even symbolically, have both definitions in the same system. It is obvious that the form "or" cannot be both simple and unanalyzable and a complex built up out of "and" and "not"; at least, it can only be this by being used ambiguously and, similarly, *mutatis mutandis,* with "and." The systems with the different definitions are equivalent in the sense that, given suitably chosen axioms, the same formulas will appear in them as theorems, and the undefined "and" (or "or") and the defined one are equivalent in the sense of having the same truth tables. But if there is an intuitively simple meaning of the form "*p* and *q*," "and" in this sense simply does not appear (is not symbolized) in a system which has only "or" and "not" as its undefined symbols and introduces "*p* and *q*" as short for "Not (not-*p* or not-*q*)." Primitiveness in a convenient calculus is one thing; intuitive or conceptual simplicity, another. No one symbolic system, we may surmise, can express everything, and in any given system we can take whatever we please as undefined, even if its intuitive meaning is complex.

Turning now to the calculi in which "not" is defined, it is notoriously difficult to explain the meaning of Nicod's stroke except by saying that "*p* | *q*" means "Not both *p* and *q*" or that it means "Either not-*p* or not-*q*"; furthermore, the "not" that is introduced by defining "Not-*p*" as "*p* | *p*" cannot be the "not" which is used in this explanation, though for purposes of logical calculation it may serve just as well. It could similarly be said that the "if" which Peirce uses in his definition of "not" cannot be understood without a more primeval "not" being presupposed. For Peirce did not use "If *p,* then *q*" in the familiar sense in which it means that *q* would be a logical consequence of *p*; it is not true that whenever *p* happens not to be the case, it would logically follow from it that everything whatever is true. Even the colloquial "If you believe that, you would believe anything" is not said of anything we wish to deny but only of particularly outrageous items (things that not only are not, but also could not, be the case). What Peirce meant by "If *p,* then *q*," it might be said, can be explained only by saying that it means "Not at once *p* and not-*q*," and this explanation uses a "not" that cannot be derived from his definition because the definition presupposes that "not."

Additionally, it might be argued that our intuitions as to what is a construction from simpler conceptions and what is itself simple are not very reliable and that if a definition introduces new economies into a calculus and, still more, if it brings a new unity to a whole subject, this may well be a symptom that it also reveals what is conceptually fundamental. The treatment of "not being the case" as an extreme case of implication—as "implying too much," so to speak—does at least reflect something important about the relation between the two concepts. A proposition's implying something, having consequences, is like its taking a risk, and its not being the case is its having too strong consequences.

See also Brouwer, Luitzen Egbertus Jan; Correspondence Theory of Truth; De Morgan, Augustus; Frege, Gottlob; Logic, Traditional; Nothing; Peirce, Charles Sanders; Presupposition; Propositions, Judgments, Sentences, and Statements; Russell, Bertrand Arthur William; Stoicism.

Bibliography

For clear summaries of the stock problems see J. N. Keynes, *Formal Logic,* 4th ed. (London, 1906), Part I, Ch. 4; Part II, Ch. 3; and Appendix B. See also W. E. Johnson, *Logic* (Cambridge, U.K.: Cambridge University Press, 1921), Part I, Chs. 5 and 14.

For the special insights of certain major writers see C. S. Peirce, *Collected Papers,* edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks, 8 vols. (Cambridge, MA: Harvard University Press, 1931–1958), II.356, 378–380, 550, 593–600; III.381–384, 407–414. See also Peirce's article "Syllogism," in *Century Dictionary* (New York: Century, 1889–1901); Gottlob Frege, "Negation," in *Translations from the Philosophical Writings of Gottlob Frege,* edited by P. T. Geach and Max Black (Oxford: Blackwell, 1952); and Bertrand Russell, *Introduction to Mathematical Philosophy* (London: Allen and Unwin, 1919), Ch. 14; "The Philosophy of Logical Atomism," Lectures I–II, in *Logic and Knowledge* (London: Allen and Unwin, 1956); *An Inquiry into Meaning and Truth* (New York: Norton, 1940), Ch. 4; and *Human Knowledge, Its Scope and Limits* (New York: Simon and Schuster, 1948), Ch. 9.

See also Ludwig Wittgenstein, *Tractatus Logico-Philosophicus,* translated by D. F. Pears and B. F. McGuinness (London: Routledge and Paul, 1961), 1.12, 2.06, 4.06–4.1, 4.25–4.463, 5.254, 5.43–5.44, 5.451, 5.512, 5.5151, 6.1201–6.1203.

On Wittgenstein (and Russell and Frege) see also G. E. M. Anscombe, *An Introduction to Wittgenstein's Tractatus* (London: Hutchinson University Library, 1959), Chs. 1–4.

On falsehood and meaninglessness see J. M. Shorter's classic piece, "Meaning and Grammar," in *Australasian Journal of Philosophy* 34 (1956): 73–91.

Various philosophical problems concerning the nature of negation are discussed in the following works: F. H. Bradley, *Principles of Logic* (London: K. Paul, Trench, 1883; 2nd ed., London: Oxford University Press, 1922), Vol. I, Book 1, Ch. 3, and Terminal Essay 6; Bernard Bosanquet, *Logic or the Morphology of Knowledge,* 2 vols. (London: Clarendon Press, 1888), Vol. I, Ch. 7, Secs. 1–3, 5; Raphael Demos, "A Discussion of a Certain Type of Negative Proposition," *Mind* 24 (1917): 188ff; Ralph M. Eaton, *Symbolism and Truth* (Cambridge, MA: Harvard University Press, 1925); J. Cook Wilson, *Statement and Inference,* 2 vols. (Oxford: Clarendon Press, 1926), Vol. I, Ch. 12; J. D. Mabbott, Gilbert Ryle, and H. H. Price, symposium, "Negation," *PAS,* Supp., 9 (1929); F. P. Ramsey, "Facts and Propositions" (1927), in his *Foundations of Mathematics* (London: K. Paul, Trench, Trubner, 1931), pp. 138–155; A. J. Ayer, "Negation," *Journal of Philosophy* 49 (1952): 797–815, reprinted in his *Philosophical Essays* (London: Macmillan, 1954); Morris Lazerowitz, "Negative Terms," *Analysis* 12 (1951–1952): 51–66, reprinted in *Philosophy and Analysis,* edited by Margaret MacDonald (Oxford, 1954); R. L. Cartwright, "Negative Existentials," *Journal of Philosophy* 57 (1960): 629–639, reprinted in *Philosophy and Ordinary Language,* edited by Charles E. Caton (Urbana: University of Illinois Press, 1963); and Gerd Buchdahl, "The Problem of Negation," *Philosophy and Phenomenological Research* 22 (1961): 163–178.

A. N. Prior (1967)

NEGATION IN INDIAN PHILOSOPHY

From the early centuries CE onward, the philosophical traditions of ancient India produced theories of negation in a broad variety of contexts, dealing with such diverse issues as negative existentials, the referentiality of empty terms, and the laws of the excluded middle and double negation. Highly technical expositions of logical principles pertaining to negation can be found in particular, though not exclusively, in the literature of the so-called New Nyāya (*Navya-Nyāya*) as of approximately the tenth century (Ingalls 1951, Matilal 1968). Earlier theories are noteworthy especially for their reflections on the nature of absence and its knowledge, in other words, for addressing the issue of negative facts and negative knowledge. These theories developed on the background of an overarching discourse about instruments of knowledge (*pramāṇa*) that shaped philosophical debate from the first centuries CE onward throughout the first millennium and is one of the most distinctive traits of classical Indian philosophizing.

Modern research on Indian theories of negation is still at a preliminary stage, and source materials in some important areas are transmitted only in fragments. On the basis of what is currently known, the Vaiśeṣika, the Nyāya, and the Mīmāṃsā traditions of Indian philosophy, as well as the logico-epistemological branch of Buddhism, deserve to be highlighted for their theories of negative knowledge. The Vaiśeṣika, an early philosophy of nature that emerged during the first two centuries CE, is mainly concerned with comprehensive enumeration and identification of the constituents of the world. The Nyāya, which originated in an old debate tradition and is primarily interested in the method of proof, integrated the Vaiśeṣika's ontological foundations into its own set of logical and epistemological principles (Franco and Preisendanz 1998). The Mīmāṃsā, originally devoted mainly to the exegesis of the Veda, likewise took over Vaiśeṣika ontology, but with much more creative adaptation. Within the Mīmāṃsā, the views of Kumārila (early seventh century CE) about absence and its knowledge differ from those of Prabhākara, who may have been Kumārila's contemporary. The logico-epistemological branch of Buddhism has as its two main representatives Dignāga (late fifth/early sixth century) and Dharmakīirti (early seventh century), of whom the latter developed a succinct theory of negative knowledge, perhaps in critical response to Kumārila.

FORMS OF ABSENCE AND THEIR KNOWLEDGE IN VAIŚEṢIKA LITERATURE

In the *Vaiśeṣikasūtra* (*VS*), a compilation of often elliptic mnemonic sentences that gradually grew as of the first two centuries CE, we find disparate identifications of specific forms of absences and brief statements of how some of them are known. As interpreted by the earliest available commentary by Candrānanda (active between the sixth and tenth centuries), *VS* 9,1–5 present four varieties of absence: the prior absence of an effect in its cause (*prāgabhāva*), the posterior absence of a cause after its destruction (*pradhvaṃsābhāva*), the mutual absence (*anyonyābhāva*) as the mutual difference between two things like a cow and a horse, and the absolute absence (*atyantābhāva*) of, for example, a hare's horn. Further forms of absences, added in *VS* 9,8–11, were most likely inserted into the text at a later stage. *VS* 9,6–7 describe, again according to Candrānanda, how prior and posterior absence are known, but without specifying an instrument of knowledge.

According to the *Praśastapādabhāṣya* by Praśastapāda (early sixth century), which comes to represent classical Vaiśeṣika thought, absence is cognized through inference, but not through a separate instrument of knowledge, for just as an arisen effect is an inferential sign for the occurrence of its sufficient causes, so is the nonarisen effect an inferential sign for the nonoccurrence of its sufficient causes. Candramati, whose *Daśapadārthaśāstra* was most probably composed between 450 and 550 and is only preserved in Chinese translation and presents an idiosyncratic version of Vaiśeṣika, lists absence as a separate ontological category. Divided into five forms, it is the object of inference. In *Śrīdhara's Nyāyakandalī* (late tenth century), and in *Udayana's Kiraṇāvalī* (early eleventh century), absence is likewise accorded the status of a separate ontological category.

THE KNOWLEDGE OF ABSENCE IN *NYĀYASŪTRA*, *-BHĀṢYA*, AND *-VĀRTTIKA*

In the *Nyāyasūtra* (*NS*), the foundational text of the Nyāya tradition that was formed between the second and fifth centuries, the knowability of further forms of absences, over and above prior and posterior absence—mutual and absolute absence are not dealt with—is emphatically defended, on the basis of an example that Vātsyāyana's commentary *Nyāyabhāṣya* (late fifth century) explains as follows: With regards to a pile of marked and unmarked clothes, someone is told "get the unmarked clothes!" and then cognizes the absence of marks in some clothes (commentary on *NS* 2,2,8; Kellner 1997; for a different interpretation of this section from *NS*, compare Matilal 1968). Whereas these remarks can be read as an attempt to expand the scope of knowable absence, the beginning portion of the *Nyāyabhāṣya* addresses the knowability of absence from a general viewpoint. For Nyāya, knowing reality, that is, the "being such [of the sixteen cardinal principles of Nyāya]" (*tattva*), is required for attaining liberation from the cycle of rebirth. Reality is the existence of what exists and the nonexistence of what does not exist. Knowledge that something does not exist arises when, through a certain instrument of knowledge, something else is known to exist, based on the thought process "if this [absentee] existed here, it would have to be cognized just like this [actually existing thing]; because its cognition is absent, it does not exist." The instrument of knowledge that illuminates something existent also illuminates something nonexistent. In keeping with this line of thought, the subcommentator Uddyotakara (c. 550–610) specifies absence as an object of sensory perception in his *Nyāyavārttika*; this becomes the orthodox Nyāya position.

THE MĪMĀṂSAKA KUMĀRILA: A SEPARATE INSTRUMENT OF KNOWLEDGE FOR KNOWING ABSENCE

Both the Buddhist epistemologist Dharmakīirti and the Mīmāṃsaka Kumārila developed comprehensive and detailed theories about the knowledge of absences. But whereas Dharmakīirti appears to have found his way of formulating and addressing the knowledge of absence as a philosophical problem only gradually, in the course of his works *Pramāṇavārttika*, *Pramāṇaviniścaya*, and *Hetubindu* (Kellner 2003), Kumārila's conception of absence and its knowledge in his *Ślokavārttika* is already part and parcel of a general philosophical approach that John Taber (2001) dubs a theory of the unitary nature of substance. All features of a substance, while different from each other, are identical with the substance itself and indirectly with each other. Nonexistence is an integral building block of reality in that every real entity is existent as itself and nonexistent as everything else (Kellner 1996, 1997). Accordingly, nonexistence has the function of accounting for the unmixed character of real entities. Kumārila distinguishes the four types of absence that are later enumerated by Candrānanda while commenting on *Vaiśeṣikasūtra* 9,1–5. In keeping with the claim that an entity is nonexistent as something else, Kumārila describes all four types with the help of relational statements—a hare's head, for instance, is nonexistent as a horn-bearer, or a cow is nonexistent as a horse.

Though a part of every real entity, nonexistence is nevertheless separate from existence and requires an instrument of knowledge of its own. The five instruments of knowledge—perception, inference, verbal knowledge, analogy, and implication—are limited to grasping existence, whereas nonexistence is apprehended by the sixth instrument of knowledge called absence, an idea that in general must have been voiced already before Praśastapāda, as he rejected it. According to his commentators, Kumārila took it over from an earlier commentator on the *Mīmāṃsāsūtras* cited in the *Śabarabhāṣya* (early sixth century), but Kumārila's interpretation of this commentator's statements are heavily contested by the Prābhākara-Mīmāṃsakas.

As an instrument of knowledge, Kumārila's *absence* is the nonarising of the other five instruments. It can manifest itself either as the soul's (*ātman*) not being transformed into the knower of the absentee as existent, or as

the knowledge of nonexistence as a part of a real entity (on the latter alternative whose interpretation is problematic, see Kellner 1996, Taber 2001). Whether an entity is known as itself, or as not another, depends on the cognizing subject's intention; the respectively uncognized part always acts as a supporting factor. Kumārila strongly disagrees with the Nyāya view that absence is grasped by sensory perception; his main counterargument is that the five external senses are incapable of coming into contact (*sannikarṣa*) with absence, and Nyāya, after all, requires such contact for any sense perception. Among others, it is this argument that led later Nyāya philosophers like Jayanta (late ninth century) and Bhāsarvajña (tenth century) to revisit the role of contact in the definition of perception (for Jayanta, see Gillon 1997). In addition, Kumārila also argues against the theory that the absence of an object is known through an inference from the nonarising of the five other instruments of knowledge, mainly because this nonarising cannot have an established inferential connection with the absence of the object that any inference requires for being sound, and because the nonarising itself cannot be known—as the absence of arising, it would itself have to be inferred from a further nonarising of instruments of knowledge, and so forth.

In the *Ślokavārttika* and in his *Tantravārttika*, Kumārila applies this instrument of knowledge in arguments that reject entities that opponents assume to exist (Kellner 1996). After demonstrating that these cannot be known by any of the five other instruments, Kumārila concludes that they can only be known through absence, as a result of which they are nonexistent. Such types of arguments are aimed at, for instance, the emptiness (*śūnyatā*) of external reality of Buddhist idealism, a human author of the Vedas as propagated by Buddhists, and an omniscient human being that is, again, assumed by Buddhists. On the whole, Kumārila's theory of nonexistence and its knowledge seems to be geared to accounting for the nature of reality and to establishing philosophical and religious truths. Empirical knowledge of negative states of affairs in everyday life are at best a secondary concern.

DHARMAKĪRTI'S THEORY OF NEGATIVE ASCERTAINMENT THROUGH INFERENCE

Like other Buddhist philosophers before him, Dharmakīirti believed that absence cannot be an object of perception because perception arises from its particular object as a cause, bearing the object's shape; an absence, however, is devoid of any causal capacity. This belief also informs Dharmakīirti's rejection of absence as a separate instrument of knowledge, condensely articulated in *Pramāṇaviniścaya*, chapter 3, prose after verse 48, for any such instrument would have to be directly or indirectly caused by its object, and *absence* as an object lacks such a capacity.

Because for Dharmakīirti there is no further instrument of knowledge besides perception and inference, negative knowledge is for him the result of inference. While perception has direct and unmediated access to real particulars in a nonconceptual fashion, inference operates with properties and concepts that are superimposed on particulars in accordance with the practical function that these jointly fulfill, and in accordance with linguistic conventions. As a result, inferences that establish negative states of affairs, based on a special type of evidence called nonperception (*anupalabdhi*) that is exclusively reserved for this purpose, ultimately prove that something is suitable for being ascertained as, and in a second step verbally referred to or physically treated as absent. They do not in any way prove a real absence that might be given independently of being cognized.

Furthermore, such inferences are limited to ascertaining the absence of particular objects that, if they existed under given circumstances, would inevitably be perceived. For entities where such a necessary perceivedness cannot be ensured, either because they are intrinsically beyond the realm of perception or because the specific environmental conditions for their perception are incomplete, not perceiving them only establishes that we do not know that they exist, not that we know that they do not exist. A proper inference on the basis of the nonperception of a perceptible object is accordingly exemplified as "in this spot on the ground, a jar does not exist because, as an object that would necessarily be perceived if it existed here, it is not perceived." From this basic inferential structure, a variety of patterns are derived with the help of further relationships such as causality, extensional relations between genus and species, and factual incompatibility, as well as contrariety and contradiction between concepts.

In his further explication of the nonperception of perceptibles, Dharmakīirti works with the notion of an implicative negation (*paryudāsa*) developed in Sanskrit grammatical literature (Cardona 1967). When understood as expressing implicative negation, a negative nominal compound formed with the prefix *a(n)*-—here: *an-upalabdhi*—affirms a state of affairs other than the negated one. Nonperception is thus explicated as another perception, that is, as the perception of a specific object

other than the absentee—not perceiving an entity like a jar is nothing other than perceiving an empty spot on the ground.

In Dharmakīirti's earliest work, the *Pramāṇavārttika*, this claim is adopted because the alternative consideration of nonperception as the mere absence of a perception would result in specific antinomies, such as an infinite justificational regress. As an absence of a perception, nonperception itself would have to be established with the help of a further instance of nonperception, and so forth. Once nonperception is assumed to be the perception of another object, it can be established through the intrinsic self-awareness of that perception. In its most developed form in the *Hetubindu*, the absence of the absentee is likewise explained away as the presence of the perceived object, and the argumentation acquires a more reductive ontological flavor. In addition, the otherness of the absentee and the object perceived in its stead is narrowed down to one where, if both objects existed, they would have to mix within one perception. Prabhākara, the Mīmāṃsā philosopher who rejects Kumārila's separate instrument of knowledge, is credited with a similar view that identifies the nonperception of one object with the perception of another that lacks the absentee. However, as his statements in the *Bṛhatī* are highly elliptic, further details of his theory and its historical and theoretical relationship to Dharmakīirti's remain obscure.

Dharmakīirti's commentators contrast his account with that of his teacher, Īśvarasena (late sixth/early seventh century), whose works are lost. Īśvarasena is said to have understood nonperception as the simple absence of the absentee's perception, based on the notion of a simple negation (*prasajyapratiṣedha*), which, like that of implicative negation (*paryudāsa*), was developed in grammatical literature. As a counterpart to implicative negation, simple negation involves only the denial of an action—here: perception—and does not further imply the affirmation of a different state of affairs. It is not known whether Īśvarasena developed his theory of nonperception, which he is said to have assumed as a third instrument of knowledge besides perception and inference, merely to solve specific problems of the theory of inference, or whether he intended it as a general theory of negative knowledge.

See also Atomic Theory in Indian Philosophy; Brahman; Causation in Indian Philosophy; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Meditation in Indian Philosophy; Philosophy of Language in India; Self in Indian Philosophy; Truth and Falsity in Indian Philosophy; Universal Properties in Indian Philosophy.

Bibliography

Cardona, George. "Negation in Pāṇinian Rules." *Language* 43 (1) (1967): 34–56.

Chakrabarti, Arindam. *Denying Existence: The Logic, Epistemology, and Pragmatics of Negative Existentials and Fictional Discourse*. Dordrecht, Netherlands: Kluwer Academic, 1997.

Cox, Collett. "On the Possibility of a Nonexistent Object of Consciousness: Sarvāstivādin and Dārṣṭāntika Theories." *Journal of the International Association of Buddhist Studies* 11 (1) (1988): 31–87.

Franco, Eli, and Karin Preisendanz. "Nyāya-Vaiśeṣika." In *Routledge Encyclopedia of Philosophy*, edited by Edward Craig. London: Routledge, 1998.

Gillon, Brendan. "Negative Facts and Knowledge of Negative Facts." In *Relativism, Suffering, and Beyond*, edited by Purushottama Bilimoria and Jitendra Mohanty, 129–149. Delhi, India: Oxford University Press, 1997.

Gillon, Brendan. "Two Forms of Negation in Sanskrit." *Lokaprajñā* 1 (1) (1987): 81–89.

Ingalls, Daniel H. H. *Materials for the Study of Navya-Nyāya Logic*. Cambridge, MA: Harvard University Press, 1951.

Kellner, Birgit. "Integrating Negative Knowledge into *Pramāṇa* Theory: The Development of the *dṛśyānupalabdhi* in Dharmakīrti's Earlier Works." *Journal of Indian Philosophy* 31 (1–2) (2003): 121–159.

Kellner, Birgit. *Nichts bleibt nichts. Die buddhistische Zurückweisung von Kumārilas abhāvapramāṇa. Übersetzung und Interpretation von Śāntarakṣitas Tattvasaṅgraha vv. 1 647–1 690 mit Kamalaśīas Tattvasa:grahapañjikā sowie Ansätze und Arbeitshypothesen zur Geschichte negativer Erkenntnis in der indischen Philosophie*. Vienna, Austria: Arbeitskreis für Tibetische und Buddhistische Studien, Universität Wien, 1997. The second part of this German-language publication contains the relevant textual materials mentioned in this entry, with German translation and analysis.

Kellner, Birgit. "There Are No Pots in the Ślokavārttika: ŚV *abhāvapariccheda* 11 and Patterns of Negative Cognition in Indian Philosophy." *Journal of the Oriental Institute* (Baroda) 46 (3–4) (1996): 143–167. Though marked with a publication date of 1996, this article was written in 1999 and supersedes parts of Kellner 1997.

Matilal, Bimal Krishna. *The Navya-Nyāya Doctrine of Negation*. Cambridge, MA: Harvard University Press, 1968.

Matilal, Bimal Krisha. "Reference and Existence in Nyāya and Buddhist Logic." *Journal of Indian Philosophy* 1 (1970): 83–110.

Shaw, J. L. "The Nyāya on Cognition and Negation." *Journal of Indian Philosophy* 8 (1980): 279–302.

Staal, J. F. "Negation and the Law of Contradiction in Indian Thought: A Comparative Study." *Bulletin of the School of Oriental and African Studies* 25 (1) (1962): 52–71.

Taber, John. "Much Ado about Nothing: Kumārila, Śāntarakṣita, and Dharmakīrti on the Cognition of Non-being." *Journal of the American Oriental Society* 121 (1) (2001): 72–88. A review article of Kellner 1997.

Birgit Kellner (2005)

NELSON, LEONARD
(1882–1927)

Leonard Nelson, a German critical philosopher and the founder of the Neo-Friesian school, was born in Berlin. After studying mathematics and philosophy he qualified for teaching as a *Privatdozent* in the natural science division of the philosophical faculty at Göttingen in 1909. In 1919 he was appointed extraordinary professor.

THE CRITICAL SCHOOL

Nelson's philosophical work was concerned mainly with two problems: the establishment of a scientific foundation for philosophy by means of a critical method and the systematic development of philosophical ethics and philosophy of right and their consequences for education and politics.

Nelson's search for a strictly scientific foundation and development of philosophy soon led him to critical philosophy. Nelson took the *Critique of Pure Reason* to be a treatise on method and regarded the critical examination of the capacities of reason as its decisive achievement. Through this critique alone could philosophical concepts be clarified and philosophical judgments traced back to their sources in cognition. Therefore, Nelson undertook a close examination of the thought of Jakob Friedrich Fries (1773–1843), the one post-Kantian philosopher who had concentrated on Immanuel Kant's critical method, carried it further, and tried to clarify its vaguenesses and contradictions.

While Nelson was still a student, he began to collect Fries's writings. These were not easily available, for Fries was hardly known at that time; when he was mentioned at all in philosophical treatises, it was as the representative of an outmoded psychologism. In his own first works Nelson attempted to defend Fries against this reproach. Together with a few friends whom he had interested in Fries's philosophy, he began to publish a *neue Folge* (new series) of *Abhandlungen der Fries'schen Schule* in 1904—the same year in which he wrote his doctoral dissertation on Fries. A few years later he founded, together with these same friends, the Jakob-Friedrich-Fries-Gesellschaft to promote the methodical development of critical philosophy.

CRITICAL METHOD AND CRITIQUE OF REASON

In his own writings devoted to the critical method, Nelson distinguished between the critique of reason and two misinterpretations of it, transcendentalism and psychologism. The critique of reason was to prepare the grounds for a philosophical system and to give this system an assured scientific basis by means of a critical investigation of the faculty of cognition. Posing the problem in this way seems to require the critique of reason and the system of philosophy to be adapted to each other in such a way that either the critique of reason must be developed a priori as a philosophical discipline, because of the rational character of philosophy, or philosophy must be conceived as a branch of psychology, since the investigation of knowledge by means of the critique of reason belongs to psychology. Transcendentalism sacrifices the main methodical thesis of the critique of reason, that the highest abstractions of philosophy cannot be dogmatically postulated but must be derived from concrete investigation of the steps leading to knowledge. Psychologism fails to recognize the character of philosophical questions and answers, which is independent of psychological concepts.

Kant did not unequivocally answer the question whether the critique of reason should be developed as a science from inner experience of one's own knowledge or as a philosophical theory from a priori principles. His subjective approach, according to which philosophical abstractions should be introduced by a critique of the faculty of cognition, indicates the first interpretation, but in carrying out his investigations—and in the asserted parallelism between general and transcendental logic as well as in the demand for a transcendental proof of metaphysical principles—Kant tacitly assumed the second interpretation and interpreted the theorems of the critique as a priori judgments. Fries, who was mainly concerned with countering the contemporary tendency to develop Kant's teaching in the direction of transcendentalism, took the subjective approach and developed it consistently from inner experience, without, however, transforming philosophical questions and answers into psychological ones. The boundary between Fries's work and psychologism is not so clear, and for this reason most of his critics misunderstood his philosophy as a psychologistic system, albeit not a consistent one.

Nelson solved the problem that philosophy based on the critique of reason seemed necessarily to lead either to transcendentalism or to psychologism by proving that both tacitly assume that a basis of knowledge must consist of proving philosophical principles from theorems of the critique of reason. If the theorems of the critique and the foundations of the philosophical system were in fact related to each other in the same way that the premises and conclusions of logical problems are related, then

indeed the critique of reason and philosophy would have to be identical—that is, they would both have to be either empirical and psychological or rational and a priori. By investigating the problem of the critique of reason Nelson showed that and why this premise is mistaken: The critique serves to clarify one's understanding of the origin of philosophical notions and of their function in the human cognition of facts. Cognition is an activity of the self, motivated by sensual stimulation; data acquired by sensual stimulation are related to one another by cognition of the surrounding world. The function of the critique of reason is to demonstrate the connecting ideas in this process and the assumed criteria by which these ideas are applied by analyzing the concrete steps in cognition and to follow these connecting ideas back to their origin in the cognitive faculty by means of psychological theory; it is not its function to prove the objective validity of the principles in which these criteria are expressed. These principles themselves are of a philosophical rather than a psychological nature. They cannot be derived from the statements of the critique; indeed, since they are the basic assumptions of all perception, they cannot be derived from any judgments more valid than they are.

CRITIQUE OF REASON AND PHILOSOPHY. The connection between the critique of reason and the system of philosophy, according to this theory, is not one of logical proof; it is derived, rather, from "reason's faith in itself," as Fries put it, from the fact that all striving for knowledge assumes faith in the possibility of cognition. This faith is faith in reason, inasmuch as reason is the faculty of cognition instructed by the stimulation of the senses. This faith is maintained by the agreement of cognitions, but it cannot be further checked or justified by a comparison of cognitions with the object cognized. This sets an unsurpassable limit to the provability of cognitions. Nelson expressed this in his paper on the impossibility of the theory of knowledge, in which he understood the theory to be an attempt to investigate scientifically the objective validity of cognition. In contrast, the critique of reason should limit itself to investigating the direction in which faith in cognition is in fact turned.

In carrying out this investigation Fries and Nelson distinguished between indirect cognition, supported by some other claim to truth, and direct cognition, which simply claims the faith of reason and which therefore neither needs nor has any justification, even when it is obscure and enters consciousness only in its application as a criterion for the unity of sensually perceivable isolated cognition. Fries and Nelson, in agreement with Kant, considered the criteria which belong solely to reason to include the pure intuition of space and time and their metaphysical combinations according to the categories of substance, causality, and reciprocal action.

NATURAL PHILOSOPHY. Nelson's interpretation of cognition led him to the problem of a mathematical natural philosophy that had been sketched by Kant and further developed by Fries; this philosophy established a priori an "armament of hypotheses" for the empirical-inductive investigation of natural laws. It coincided in fact with the basic principles of classical mechanics and thereby came into conflict with modern physics. Nelson neither minimized this conflict nor confused it with problems of the principles of critical natural philosophy. He saw physics as being in the process of a radical changeover to modern theories, which had by no means yet been ordered into a conflict-free system comparable to that of classical physics. He was sure that every physical theory must go beyond the data provided by observation and experiment in developing concepts and making assertions. And he was convinced that the positivistic, antimetaphysical tendencies of contemporary physicists promoted a tacit and therefore uncritical metaphysics. Without himself being able to solve the conflict that had arisen within critical philosophy, he was convinced the progressive clarification of modern theories would lead back to a physics based on classical mechanics.

CRITICAL ETHICS

BASIC PRINCIPLES. Nelson systematically applied the critical method in his studies in practical philosophy—ethics in the broadest sense of the word, including philosophy of right and philosophically based educational and political theory. He added his own critique of practical reason to those of his predecessors. He developed his own processes, both for what he called abstraction (analysis of the assumptions underlying practical ethical value judgments) and for determining, by an empirical study of value judgments, "the interests of pure practical reason," that is, ethical demands put to the human will by reason itself. It is these interests that make value judgments possible. Nelson derived two basic ethical principles from these interests: the law of the balanced consideration of all interests affected by one's own deeds and the ideal of forming one's own life independently, according to the ideas of the true, the beautiful, and the good. These two principles were linked by the fact that, on the one hand, the law of balanced consideration, as a categorical imperative, determines the necessary limiting condition for the ideal value of human behavior; on the

other hand, the ideal of rational self-determination leads to the doctrine of the true interests of man and finds in these interests the standard for a balanced consideration of conflicting interests.

NELSON'S SYSTEM. From these two principles alone Nelson developed his system of philosophical ethics; he limited himself to such consequences as could be derived from these principles purely philosophically—without the addition of experience—but he attempted to grasp them completely and systematically. In this he was influenced, first, by his interest in systematically and strictly justifying the assumptions used in every single step and the logical connections of the concepts appearing in the principles and, second, by his interest in applying this practical science. The principles demonstrated are formal and permit determination of concrete ethical demands only through their application to given circumstances as justified by experience. But it is precisely this application of the principles to the world of experience that requires preparatory philosophical investigation if the application is to be guarded against hasty generalization of single results, in which changing circumstances are not taken into account, and against opportunistic adaptation to circumstances without regard for the practical consequences of ethical principles. In the system as a whole, ethics and philosophy of right appear side by side. Nelson distinguished between them according to different ways of applying the law of balanced consideration. As a categorical imperative, this law demands of the human will the balanced consideration of other persons' interests affected by its actions. By its content it determines the duties of the individual by the rights others have with regard to him; in this respect it is related to communal life and thereby provides a criterion for the value of a social order. Nelson defined this criterion as the concept of the state of right, by which he meant the condition of a society in which the interests of all members are protected against wrongful violation. Ethics, by this definition, is concerned with the duties of the individual; philosophy of right is concerned with the state of right. To each of these disciplines Nelson added another concerned with the conditions of realizing the values studied by them: philosophical pedagogics, as the theory of the education of man to the ethical good, and philosophical politics, as the theory of the realization of the state of right.

VALIDITY OF ETHICAL PRINCIPLES. The logically transparent construction of the entire system reveals clearly that the principles behind all further developments are strictly valid in all cases but can be applied only through full consideration of the concrete circumstances in each individual case; since they are objectively valid, they are not subject to arbitrary decisions and are valid even in cases where human insight and will fail to understand them; but they are justified only by reference to reason, which makes possible for each individual the autonomic recognition of these standards and the critical examination of their applications. Thus, the demands of equality for all before the law and of equality of rights are compatible with the demand to differentiate according to given circumstances; and the demands of force against injustice remain linked to those of freedom of criticism and of public justification for the legal necessity of certain coercive measures. Such coercive measures are particularly necessary when the freedom of man to form himself rationally within the framework of his own life is threatened; this freedom can be threatened because man's true need for it is at first obscure and can therefore be mistaken and suppressed.

Nature and chance. One conclusion appears again and again, determining the structure of the whole system. In each case it is a question of fighting with chance, to which the realization of the good is subject in nature. What happens in nature is, according to the laws of nature, dependent on the given circumstances and on the forces working through them, which are indifferent to ethical values: Under the laws of nature it is a matter of chance whether what should happen is in fact what happens or whether ethical demands are ignored. But what ethics demands should not be subject to chance but assured by the human will. Following this line of thought, Nelson derived the law of character in ethics, which demands from man the establishment of a basic willingness to fulfill his duty, by which he makes himself independent of given concrete circumstances; his inclinations and the influences on his will may or may not be in agreement with the commands of duty.

In the philosophy of right Nelson correspondingly finds certain postulates. These determine the forms of reciprocal action in society which alone assure just relations between individuals; among them are public justice, prosecutability, the law of contract, and the law of property. The transitions from ethics to pedagogy and from philosophy of right to politics are made in the same spirit. Education, among the many influences on man, should strengthen or create those elements that develop his capacity for good and oppose those that could weaken this capacity. Politics is concerned with the realization and securing of the state of right determined by the postulates of philosophy of right. This problem leads to the

postulation of a state seeking the rule of law and having the power to maintain itself against forces in society opposing the rule of law. A sufficiently powerful federation of states is necessary to regulate the legal relationships between states.

The same conclusion is reached in the last section of Nelson's *System der philosophischen Rechtslehre und Politik*. Here again, in a state of nature it is a matter of chance to what degree states realize the rule of law or violate its demands, unless men having insight into justice and moral will work to transform the existing state into a just state. These men must interfere in the struggle between social groups and parties and must themselves band together into a party. In this case, therefore, the ideal of a just state leads to that of a party working to achieve it.

FREEDOM AND NECESSITY. The conflict between natural necessity and man's freedom and responsibility impelled Nelson's thinking. Ethical standards are valid for human action in nature and are therefore directly relevant to two apparently mutually exclusive forms of legality: The theoretical form, according to which everything that happens in nature (including human behavior) is determined by natural laws working through the existing powers, and the practical one, which presents the human will with duties that can either be violated and ignored or become man's purpose.

Thus on the one hand Nelson insisted that demonstrated ethical standards be maintained without compromise and rejected the skeptical assumption that man, as a limited creature of nature, was incapable of maintaining them; this assumption he considered a sacrifice of known ethical truth, a mere excuse for those who were able but not willing. On the other hand, he expected the human will to act according to the strongest motivation of the moment, without any guarantee from nature that this motivation would direct man toward what is ethically required. For this reason he rejected any speculation that in a state of nature the good would pave its own way.

Within the framework of the critique, Nelson thoroughly examined the question of how man's freedom could be reconciled with this natural law. He sought the answer in the doctrine of transcendental idealism that human knowledge is limited to the understanding of relationships in the sphere of experience but cannot achieve absolute perception of reality itself. In the consciousness of his freedom, which is indissolubly bound to the knowledge of his responsibility, man relates himself by faith to the world of that which is real in itself and superior to the limitations of nature. Nelson unified the two points of view by connecting two results of his investigations of the critique of reason: the principle of the existence of pure practical reason, which as a direct moral interest makes moral insight and moral motivation possible, and the principle of the original obscurity of this interest, according to which it does not determine judgment and will by its very existence but rather requires enlightenment and is dependent on stimulation.

EDUCATION AND POLITICS. Concern with the realization of ethical requirements led Nelson beyond his philosophical work to practical undertakings, in which he gave primary emphasis to politics, particularly to political education.

Toward the end of World War I Nelson collected a circle of pupils and coworkers who were willing to undergo intensive education and discipline in preparation for the political duties imposed by ethics and philosophy of right. Together with these pupils he founded the Internationaler Jugendbund and in January 1926 developed his own political organization, the Internationaler Sozialistischer Kampf-Bund. In 1924 he opened a "country educational institution," Landerziehungsheim Walkemühle, directed by his coworker Minna Specht. Here youths and children were trained in a closely knit educational and working community for activity in the workers' movement, until the school was closed and appropriated by the National Socialists in 1933.

As a teacher and educator Nelson had a strong effect on his pupils. He led them by masterly Socratic discussions to a clarification and critical examination of their own convictions, and he required them to carry out what they had recognized as just and good in their actions with the same consistency that he demanded of himself. "Ethics is there in order to be applied."

See also Epistemology; Epistemology, History of; Ethics; Fries, Jakob Friedrich; Kant, Immanuel; Neo-Kantianism; Psychologism.

Bibliography

WORKS BY NELSON

"Die kritische Methode und das Verhältnis der Psychologic zur Philosophic." *Abhandlungen der Fries'schen Schule,* n.f., 1 (1) (1904): 1–88.

"Jakob Friedrich Fries und seine jungsten Kritiker." *Abhandlungen der Fries'schen Schule,* n.f., 1 (2) (1905): 233–319.

"Bermerkungen über die nicht-Euklidische Geometric und den Ursprung der mathematischen Gewissheit."

Abhandlungen der Fries'schen Schule, n.f., 1 (2) (1905): 373–392; 1 (3) (1906): 393–430.

"Inhalt und Gegenstand, Grund und Begründung. Zur Kontroverse über die kritische Methode." *Abhandlungen der Fries'schen Schule,* n.f., 2 (1) (1907): 33–73.

"Ist metaphysikfreie Naturwissenschaft möglich?" *Abhandlungen der Fries'schen Schule,* n.f., 2 (3) (1908): 241–299.

"Über das sogenannte Erkenntnisproblem." *Abhandlungen der Fries'schen Schule,* n.f., 2 (4) (1908): 413–850.

"Bemerkungen zu den Paradoxien von Russell und Burali-Forti." *Abhandlungen der Fries'schen Schule,* n.f., 2 (3) (1908): 301–334. Written with Kurt Grelling.

"Untersuchungen über die Entwicklungsgeschichte der Kantischen Erkenntnistheorie." *Abhandlungen der Fries'schen Schule,* n.f., 3 (1) (1909): 33–96.

"Die Unmöglichkeit der Erkenntnistheorie." *Abhandlungen der Fries'schen Schule,* n.f., 3 (4) (1912): 583–617.

"Die Theorie des wahren Interesses und ihre rechtliche und politische Bedeutung." *Abhandlungen der Fries'schen Schule,* n.f., 4 (2) (1913): 395–423.

"Die kritische Ethik bei Kant, Schiller und Fries." *Abhandlungen der Fries'schen Schule,* n.f., 4 (3) (1914): 483–691.

Die Rechtswissenschaft ohne Recht. Kritische Betrachtungen über die Grundlagen des Staats- und Völkerrechts, insbesondere über die Lehre von der Souveränität. Leipzig: Veit, 1917.

Die Reformation der Gesinnung durch Erziehung zum Selbstvertrauen. Gesammelte Aufsätze. Leipzig, 1917; 2nd enlarged ed., Leipzig, 1922.

Vorlesungen über die Grundlagen der Ethik. 3 vols. Leipzig: Veit, 1917–1932. Vol. II translated by Norbert Guterman as *System of Ethics.* New Haven, CT: Yale University Press, 1956.

Die Reformation der Philosophie durch die Kritik der Vernunft. Gesammelte Aufsätze. Leipzig, 1918.

Demokratie und Führerschaft. Leipzig, 1920.

Spuk, Einweihung in das Geheimnis der Wahrsagerkunst Oswald Spenglers. Leipzig: P. Reinhold, 1921.

"Kritische Philosophic und mathematische Axiomatik." *Unterrichtsblätter für Mathematik und Naturwissenschaft,* 34th year, (4 and 5) (1927): 108–115 and 136–142.

"Sittliche und religiöse Weltansicht." *XXVI Aasaner Studenten-Konferenz,* 7–25. Leipzig, 1922.

"Die Sokratische Methode." *Abhandlungen der Fries'schen Schule,* n.f., 5 (1) (1929): 21–78. Translated by Thomas K. Brown in *Socratic Method and Critical Philosophy.* New Haven, CT: Yale University Press, 1949. Selected essays.

Fortschritte und Rückschritte der Philosophie; von Hume und Kant bis Hegel und Fries. Frankfurt: Öffentliches Leben, 1962.

WORKS ON NELSON

Specht, Minna, and Willi Eichler, eds. *Leonard Nelson zum Gedächtnis.* Frankfurt, 1953. Contains essays.

Grete Henry-Hermann (1967)
Translated by Tessa Byck

NEMESIUS OF EMESA

(fl. c. 390)

Nemesius of Emesa was the author of a treatise, *De Natura Hominis* (On the nature of man), which is the earliest extant handbook of theological or philosophical "anthropology." All that is known of his life is that he was probably bishop of Emesa in Syria.

As a Christian, Nemesius viewed the Bible as his primary authority, but he derived the content of his work chiefly from Galen's *On the Use of the Parts of the Body,* which is superior to Nemesius's treatise both in thoroughness and originality; from Origen's *Commentary on Genesis*; and from some commentators on Aristotle, a few works by the Neoplatonist Porphyry, and doxographical materials. His subjects and sources can be outlined as follows: Ch. 1, man in the creation (Galen, Origen); Chs. 2–3, the soul and the body (doxographical, Porphyry, Galen); Chs. 4–5, the body and the elements (Galen); Chs. 6–14, the faculties of the soul, including human development, the senses, thought and memory, reason and speech (Galen, Porphyry); Chs. 15–28, the parts of the soul, the passions, and such matters as the nutritive and generative faculties and respiration (mostly Galen); Chs. 29–41, freedom, possibility, and fate (commentaries on Aristotle, Neoplatonists); Chs. 42–44, providence (in part ultimately from Posidonius, in part from Christian theologians).

In the last part of his book (Chs. 35ff.), Nemesius turns from minimizing the function of free will in human affairs (deliberation concerns only indifferent possibilities) to an elaborate attack upon the Stoic doctrine of fate and teaching about destiny. Utilizing Aristotle's distinction between voluntary and involuntary acts, he insists that men actually have free will, that its extent can be discovered (interrelated with the action of providence), and that it was given to mutable men so that they might become immutable. The work ends abruptly and seems to lack a conclusion.

Nemesius argued that the soul is an incorporeal being and is therefore immortal (in his opinion the latter point is also proved by the Bible). The problem of how it is united with the body is solved (Chs. 20–21) by following the Neoplatonist Ammonius. "Intelligibles" are capable of union with things adapted to receive them, but in such a union they remain confused and imperishable. The soul is "in a body" not locally but "in habitual relation of presence." From this analysis Nemesius turns in Ch. 22 to discuss the union of the divine Word with his manhood—as William Telfer points out, thus reversing

the usual patristic argument. Nemesius claims that the union in Christ is therefore not by "divine favor" but is "grounded in nature."

See also Aristotle; Galen; Neoplatonism; Origen; Philosophical Anthropology; Porphyry.

Bibliography

The Greek text, with Latin translation, of *De Natura Hominis* is in *Patrologia Graeca,* edited by J. P. Migne (Paris, 1857–1866), Vol. XL, Cols. 508–818. There is an English translation by William Telfer in *Cyril of Jerusalem and Nemesius of Emesa,* Vol. IV of the Library of Christian Classics (Philadelphia: Westminster Press, 1955).

For works on Nemesius, see Werner Jaeger, *Nemesios von Emesa* (Berlin: Weidmann, 1914); H. A. Koch, *Quellenuntersuchungen zu Nemesius von Emesa* (Berlin, 1921); Johannes Quasten, *Patrology* (Westminster, MD: Newman Press, 1960), Vol. III, pp. 351–355, which includes a full bibliography. See also the articles by E. Skard, "Nemesiosstudien," in *Symbolae Osloenses* 15–16 (1936): 23–43; 17 (1937): 9–25; 18 (1938): 31–41; 19 (1939): 46–56; 22 (1942): 40–48; and Skard's article "Nemesios," in *Realencyclopädie der classischen Altertumswissenschaft,* Supp., VII (Stuttgart, 1940): Cols. 562–566.

Robert M. Grant (1967)

NEO-KANTIANISM

"Neo-Kantianism" is a term used to designate a group of somewhat similar movements that prevailed in Germany between 1870 and 1920 but had little in common beyond a strong reaction against irrationalism and speculative naturalism and a conviction that philosophy could be a "science" only if it returned to the method and spirit of Immanuel Kant. These movements were the fulfillment of Kant's prophecy that in a hundred years his philosophy would come into its own.

Because of the complexity and internal tensions in Kant's philosophy, not all the Neo-Kantians brought the same message from the Sage of Königsberg, and the diversity of their teachings was as great as their quarrels were notorious. At the end of the nineteenth century the Neo-Kantians were as widely separated as the first-generation Kantians had been at its beginning, and the various Neo-Kantian movements developed in directions further characterized by such terms as Neo-Hegelian and Neo-Fichtean. But whereas G. W. F. Hegel, Friedrich Schelling, Johann Gottlieb Fichte, and others had used the words of Kant while being alien to their spirit, the Neo-Kantians were, on the whole, faithful to the spirit while being revisionists with respect to the letter. Attempting to legitimize their revisions by the *ipsissima verba* of Kant, they established the craft of "Kant-philology" and began an analysis of Kant's texts that had not been equaled in microscopic punctiliousness except in the exegesis of the Bible and of a few classical authors. Hans Vaihinger's immense commentary on the first seventy pages of the *Critique of Pure Reason* (*Commentar zu Kants "Kritik der reinen Vernunft,"* 2 vols., Berlin and Leipzig, 1881–1893) is an exemplar of this craft and industry.

Neo-Kantianism grew out of the peculiar social-cultural situation of German science and philosophy, and in turn it constituted a new academic situation with many characteristics of a long intellectual fad. Most of the groups of Neo-Kantians had their own journals—the *Philosophische Arbeiten* at Marburg, *Logos* at Heidelberg, the *Annalen der Philosophie und philosophischer Kritik* of Vaihinger, and the *Philosophische Abhandlungen* at Göttingen. (*Kant-Studien,* like the Kant *Gesellschaft,* was open to all.) Doctrines were known by the names of the universities where they originated; men entered and left the movement as if it were a church or political party; members of one school blocked the appointments and promotions of members of the others; eminent Kant scholars and philosophers who did not found their own schools or accommodate themselves to one of the established schools tended to be neglected as outsiders and contemned as amateurs. As many as seven distinct schools have been described by historians, but they do not agree on the programs, heresies, and bona fide membership of each school.

THE BEGINNINGS

So far as an intellectual movement can be said to have a beginning at a specific moment of time, Neo-Kantianism began with the publication at Stuttgart in 1865 of Otto Liebmann's *Kant und die Epigonen,* whose motto—"Back to Kant!"—has become famous. German philosophy was generally weak toward the middle of the nineteenth century; there was less interest in it, and less ability among its practitioners, than at perhaps any other time in modern German history. Earlier in the century, when Kant's philosophy had been submerged first in the great idealistic systems and then in those of nature-philosophy, there had been modest calls for a return to Kant (for instance, by I. H. Fichte, the son of J. G. Fichte, and by Ernst Reinhold, the son of K. L. Reinhold) as a means of escape from the kinds of philosophy that Kant would have held to be impossible and that seemed more and more to offer nothing of value to German cultural life as a counterbal-

ance to the materialism attendant upon the flourishing of natural science, technology, and national economy. However, in the decade preceding Liebmann's book there had been signs of change.

ZELLER AND FISCHER. Eduard Zeller (1814–1908), in his Heidelberg lecture, *Ueber Bedeutung und Aufgabe der Erkenntnistheorie* (published Heidelberg, 1862), called for a return to epistemology; and this, he spelled out explicitly, meant a return to Kant. Kuno Fischer (1824–1907), the greatest historian of philosophy at that time and the teacher of Liebmann, Johannes Volkelt, and Wilhelm Windelband, in 1860 published a monumental book on Kant (*Kants Leben und die Grundlagen seiner Lehre,* Mannheim and Heidelberg) that presented, in a form still useful although outmoded in details, a picture of Kant that could not but excite interest in and study of Kant. In 1865 Fischer initiated a great controversy with Adolf Trendelenburg on the proper interpretation of Kant's theory of space; this controversy mobilized most of the philosophical public in Germany on one side or the other, including Trendelenburg's pupil Hermann Cohen, who had hitherto concentrated mostly on Plato.

HELMHOLTZ AND LANGE. Two other men, Hermann von Helmholtz and F. A. Lange, almost simultaneously with Liebmann made their spiritual pilgrimage to Königsberg.

Helmholtz. Hermann von Helmholtz (1821–1894), then Germany's greatest scientist, had been arguing for years for a view whose origin he found in Kant. The doctrine of specific energies of sensory nerves had led him to a theory of the subjectivity of sensory qualities, which he regarded as signs of unknown objects interacting with our sense organs; he then extended this commonly held view to the conclusion that space itself is dependent upon our bodily constitution. This theory made it possible for Helmholtz to argue that there could be alternative spaces and geometries, each appropriate to a particular kind of nervous apparatus and necessary to the being so constituted, but none of them picturing the real structure of the world. Thus, while Helmholtz gave up Kant's theory of the unique status of Euclidean geometry, he held that his own theory of space was in keeping both with Kant's theory and with the most modern work in mathematics, physics, and physiology. Moreover, in his theory of unconscious inferences he accepted the Kantian theory that perception involves judgment. The guiding principle in such unconscious inference is the a priori principle of causation, which extends our knowledge no further than possible experience, but gives us the right to posit unknown causes of our sensations. Helmholtz vigorously rejected metaphysics but extolled philosophy as an ancilla to science. Both the strengths and the obvious weaknesses of Helmholtz's Kantianism were effective in making a return to Kant seem fruitful to science, for it meant that the greatest of German thinkers could be used on the side of science, against metaphysics.

Lange. The year 1866 saw the publication of Friedrich Albert Lange's *Geschichte des Materialismus* (Iserlohn and Leipzig; translated by E. C. Thomas as *History of Materialism,* 3 vols., London, 1877–1879). Lange, who was born in 1828 and died, while professor of philosophy at Marburg, in 1875, wrote his massive but readable book to point out the metaphysical mysteries and pretensions of materialism, which traditionally claimed to be only a courageous but unspeculative extension of the results of science into regions previously occupied only by theology and superstition. Like Helmholtz, Lange held that the sensible world is a product of the interaction between the human organism and an unknown reality. The world of experience is determined by this interaction, but the organism itself is only an object of experience, and it is to be understood by psychology and physiology. Causality, needed in all such sciences, is a mode of thought necessary to a mind constituted like ours; processes and principles of thought have physiological bases. Thus, materialism (although a phenomenal materialism, since matter itself is only a phenomenon) is the most likely truth about reality so far as it can be known. But what of Kant's intelligible world? Lange completely rejected Kant's teaching of the rational necessity of the structure of an intelligible but unknowable world; he held that our views of it are only products of poetic fancy (*Dichtung*). While Lange defended materialism as a doctrine of reality (phenomena) that serves as a bulwark against theology and metaphysics, he held that because knowledge is not man's whole goal, *Dichtung* is also important. "Man needs to supplement reality [about which materialism is the best truth we know] with an ideal world of his own creation," and this is a world of value "against which neither logic nor touch of hand nor sight of eye can prevail" (*History of Materialism,* Vol. III, pp. 342 and 347).

Two things stand out in the works of these precursors—if not direct progenitors—of Neo-Kantianism. Their Kantianism was exclusively theoretical, oriented entirely around the *Critique of Pure Reason* and neglectful or disdainful of Kant's practical philosophy. This puts them in the line of development of German positivism, a line that goes from them through Alois Riehl and the fic-

tionalist Hans Vaihinger to Ernst Mach and Moritz Schlick. Their Kantianism was also psychological and even physiological—the a priori elements they acknowledged were dependent upon the human constitution; the transcendental and logical aspects of Kant's work were neglected or rejected. In this respect they were followed by Hans Cornelius (1863–1947) and by Richard Hönigswald (1875–1947), a pupil of Riehl.

METAPHYSICAL NEO-KANTIANISM

Theoretical and physiological Kantianism was in the air when the twenty-five-year-old Liebmann published his manifesto. *Kant und die Epigonen* argued that Kant made one great mistake: believing in the existence of the thing-in-itself. This belief, however, was not an essential part of Kant's doctrine, but only a dogmatic residue that could be removed without damage to the rest of the system. However, Fichte, Schelling, Hegel, Jakob Friedrich Fries, Johann Friedrich Herbart, and Arthur Schopenhauer either did not recognize the belief that there is a thing-in-itself as an error (for instance, Schopenhauer) or, while recognizing it as an error, made analogous errors in their efforts to correct it (Fichte's transcendental ego is as unknowable and unthinkable as the thing-in-itself). The weaknesses thus introduced into their systems were fatal, since they depended upon a concept that Kant had only inadvertently admitted. Hence, none of them could be followed; one had to return to their common source, remove its error, and apply this improved Kantianism to present problems.

While Liebmann's first book showed remnants of a psychological interpretation of Kant, his next book, *Zur Analysis der Wirklichkeit* (Strasbourg, 1876) argued for a strictly transcendental "logic of facts" whose inspiration was as much Spinozistic as Kantian. In this book Liebmann stood close to the Marburg school, at least in his conclusions. However, in his later *Gedanken und Tatsachen* (2 vols., Strasbourg, 1882–1901) he admitted the need and argued for the possibility of a "critical metaphysics" as a "rigorous consideration of human views and hypotheses about the essence of things," growing out of "deep-rooted, ineradicable spiritual needs and intellectual duty" (ibid., 2nd ed., Vol. II, p. 113). His critical metaphysics makes hypotheses about the transcendent and the unknowable, but leaves open a field for value decisions that do not depend on claims to valid knowledge, but only on our wills as they are nurtured by culture. In this line of thought Liebmann seemed to draw closer to the Heidelberg school, but even in his earlier work there were anticipations of Windelband's famous analysis of the differences between historical and scientific knowledge.

RIEHL. Less openly metaphysical than Liebmann's was the realistic Neo-Kantianism of Alois Riehl (1844–1924). In contrast to Liebmann, Riehl insisted that Kant held to the real existence of things-in-themselves and that this concept is essential to Kant's—and to any sound—theory of knowledge. He asserted that Kant proved only that things-in-themselves cannot be known by pure reason, not that they are not known mediately in sense perception. Phenomena are simply their modes of appearance; they are not in a different ontological realm, but are merely actualizations of their Aristotelian potentialities in the context of a mind. The laws of the organization of phenomena are transcendentally (not psychologically) based on the activity of self-consciousness; their specific characteristics depend on the reality of that of which they are appearances. All knowledge is or can become scientific; philosophy is nothing but a theory of science; metaphysics is "an opiate of the mind."

Nevertheless, Riehl believed it both unavoidable and legitimate to reason hypothetically from phenomena to reality, for metaphysical hypotheses cannot be entirely excluded from science itself. He argued, for instance, for a double-aspect psychophysical theory of the relationship between mind and the world, for a partial duplication of phenomenal laws in the real world, and for complete determinism. The tone of his philosophy, however, was somewhat positivistic; he said he acknowledged "the metaphysical" but not "metaphysics." With *wissenschaftliche* (scientific) philosophy he contrasted *unwissenschaftliche* philosophy, or classical speculative metaphysics, which he rejected; and with both he contrasted *nichtwissenschaftliche* philosophy as a practical discipline for the realization of humanly created values (*Wertbegung* and *Geistesführung*). In his later life he was most concerned with the latter.

OTHER METAPHYSICAL INTERPRETATIONS. Another realistic metaphysical interpretation of Kant was given by the Kant philologist Erich Adickes (1866–1928) in his *Kants Lehre von der doppelten Affektion unseres Ich* (Tübingen, 1929).

Other attempts at "critical metaphysics" on a Kantian basis were made by Johannes Volkelt (1848–1930) and by Friedrich Paulsen (1846–1908). The former's *Kants Erkenntnistheorie* (Leipzig, 1879) and the latter's *Entwicklungsgeschichte der Kantischen Erkenntnistheorie* (Leipzig, 1875) tried to show that Kant himself was an idealistic

metaphysician *malgré lui.* Later works designed to bring out the metaphysics in Kant were by Max Wundt (*Kant als Metaphysiker,* Stuttgart, 1924), Heinz Heimsoeth (articles collected in *Studien zur Philosophie Immanuel Kants,* Cologne, 1956), and Gottfried Martin (*Kant, Ontologie und Wissenschaftslehre,* Cologne, 1951; translated by P. G. Lucas as *Kant's Metaphysics and Theory of Science,* Manchester, U.K., and New York, 1955). Martin Heidegger's *Kant und das Problem der Metaphysik* (Bonn, 1929; translated by J. S. Churchill as *Kant and the Problem of Metaphysics,* Bloomington, IN, 1962) presented an extreme form of this view but falls outside the scope of Neo-Kantian intentions.

MARBURG NEO-KANTIANISM

By the standards of recent philosophy Marburg Neo-Kantianism, or panlogistic transcendental philosophy, was no less metaphysical, but by the standards of the time its orientation around the "fact of science" seemed to make it at least antispeculative. In launching the journal of the Marburg school, Hermann Cohen and Paul Natorp wrote: "Whoever is bound to us stands with us on the foundation of the transcendental method.… Philosophy, to us, is bound to the fact of science, as this elaborates itself. Philosophy, therefore, to us is the theory of the principles of science and therewith of all culture" (*Philosophische Arbeiten,* Vol. I, No. 1, 1906).

HERMANN COHEN. Hermann Cohen (1842–1918), a younger colleague of Lange's at Marburg, rejected the naturalism he believed to be inherent in the Kantianism of Helmholtz, Lange, and Liebmann. They were wrong in thinking philosophy should begin with an analysis of consciousness and should show how conscious human beings apply concepts to the data of sensation in order to produce phenomenalistic world pictures that are distinguished from things as they are. The fact to be understood is not this highly dubious psychological process; the fact is science itself and, in ethics, it is not human motives and aspirations and feelings of duty but the fact of civil society under law as constructed in the science of jurisprudence. Kant himself had tried to understand "the fact of science and culture," but he failed to separate this fact from dubious psychological and phenomenological facts he seemed to be dealing with.

Logic for Cohen is not at all psychologistic; it is not even formal. The very notion of formal logic presupposes something not formal: data drawn from some other source, be it pure intuition or perception. Logic, as Cohen saw it, is the logic of knowledge, not the logic of empty thought; it is the logic of truth, in which any assertion gains its status as true solely by virtue of its systematic position in a body of universal laws that, in turn, require each other on methodological grounds. Thought, Cohen taught, accepts nothing as given and is not true of anything independent of it—certainly not of intuitional data, as Kant believed. Thought generates content as well as form, and the content of self-contained thought is reality itself as object and goal of knowledge. This extravagant panlogism was based on Cohen's ingenious interpretation of the history of the differential calculus, which he saw as the logic of mathematical physics. Not number and not observed motion, as Kant believed, are given as raw data to science; rather, the mathematical differential, which is not given at all but is created by thought, is the necessary device for the creation of nature as object of possible experience: "This mathematical generation of motion [by integration of the derivative] and thereby nature itself is the triumph of pure thinking" (*Logik der reinen Erkenntnis,* Berlin, 1902, p. 20). Through an interpretation of Kant's teachings concerning intensive magnitudes of sensations, Cohen saw in the method of the calculus a paradigm of the category of origin (*Ursprung*) and the logical process of production (*Erzeugung*) to which every fact owes its reality; that is, its position in a logically necessary scheme.

Through the work of thought on its own materials, Cohen believed he could dispense with all independent givens in knowledge. Nothing is given (*gegeben*); all is problematic (*aufgegeben*). Fact is that which is completely determined by thought. The thing-in-itself is not a thing at all. It does not exist, but is only a thought of a limit (*Grenzbegriff*) to our approach to a complete determination of things as they are; that is, as they would fully satisfy systematic thought.

Cohen's pupil Ernst Cassirer spoke of him as "one of the most resolute Platonists that has ever appeared in the history of philosophy." When Cohen said, for example, "Thinking itself produces what is to be held to be" (ibid., p. 67; cf. p. 402), he was not speaking of thought as a process in an individual. "Thought" is not the name of a process, but refers only to the corpus of the unending history of science. To be, then, is to be thought, but not to be thought in somebody's consciousness; to be thought means to be asserted under valid and immanent a priori principles that inescapably determine the unique structure of mathematical physics. Cohen was as much of a dogmatist as Kant himself with regard to the structure of science.

The original stages of Cohen's teachings are found in his three commentaries on Kant (*Kants Theorie der Erfahrung,* Berlin, 1871; *Kants Begrundung der Ethik,* Berlin, 1877; *Kants Begrundung der Aesthetik,* Berlin, 1889), one on each *Critique.* They are continuous criticisms of all of Kant's "givens"; for example, experience, intuition, categories, duty, things-in-themselves. The final stages are contained in his three systematic works (*Logik der reinen Erkenntnis,* Berlin, 1902; *Ethik des reinen Willens,* Berlin, 1904; *Aesthetik des reinen Gefühls,* 2 vols., Berlin, 1912), which parallel the three *Critiques.* At its midpoint Cohen's thought was close to the contemporary rejections of psychologism by Alexius Meinong and Edmund Husserl; at its end it would have taken only the "bathos of experience," to use Kant's words, to change it, in principle, into a kind of positivism or even historicism.

NATORP. The principal thinker among the second generation of Marburg Neo-Kantians was Paul Natorp (1854–1924). It fell to him to deal with the new developments in science (especially the theory of relativity, in his *Die logischen Grundlagen der exakten Wissenschaften,* Leipzig, 1910) by penetrating to a deeper level of methodology than Cohen could reach in his own work, which was largely restricted to classical mathematics and physics.

More important, it was Natorp's task to introduce the whole field of psychology into the body of knowledge considered and understood in Cohen's way, and thereby to fill the lacuna Cohen left between *Bewusstsein überhaupt* (consciousness in general, the "fact" of science) and the limited individual human consciousness. Natorp's *Einleitung in die Psychologie* (Freiburg, 1888) and his *Allgemeine Psychologie nach kritischer Methode* (Tübingen, 1912) attempted, first, to apply Cohen's transcendental method to psychology instead of leaving it exposed to the naturalistic methods of Cohen's and Natorp's rivals, such as Riehl. In this attempt Natorp came close to results like those of Wilhelm Dilthey without, he thought, having to draw his relativistic, skeptical, and historicistic conclusions. And, second, these books attempted to bridge the gap between the objective world of phenomena and the nonphenomenal, nonnatural self that possessed the knowledge of the phenomenal world. Cohen had moved so far from Kant toward Hegel that it was for him an almost insignificant accident that individual men and women know anything; *Bewusstheit* (known-ness), not *Bewusstsein* (consciousness), was important for him. Natorp had to undertake another almost Copernican revolution against objective panlogism without at the same time naturalizing the knowing subject, which would have led to relativism and skepticism.

He performed the first part of his task by the classical Kantian move of seeing empirical ego and empirical object as standing in a necessary correlation with each other, not as independent phenomena; the latter part he accomplished by insisting that the pure ego cannot be an object—it is as much a *Grenzbegriff* as the thing-in-itself. For Natorp the objective and the subjective were not two realms, either opposed to each other or one including the other. Rather, they were two directions of knowledge, objectification and subjectification, each starting from the same phenomenon and each employing the transcendental method of categorial constitution, resolution into *Ursprung* and *Erzeugung.* Just as Cohen's antipsychologistic panlogism had brought him close to Husserl's *Logische Untersuchungen,* Natorp's linking of psychology and panlogism brought him close to Husserl's *Ideen*; and it is easy to see how Nicolai Hartmann, Natorp's pupil, could move over into the phenomenological camp (J. Klein, "Hartmann und die Marburger Schule," in *Nicolai Hartmann, der Denker und sein Werk,* by Heinz Heimsoeth and Robert Heiss, Göttingen, 1952).

CASSIRER. The last great representative of Marburg Neo-Kantianism was Ernst Cassirer (1874–1945), whose works on the philosophy of science continued the line of argument initiated by Natorp and show some close resemblances to positivism. Cassirer's most important contribution, however, was to extend the Marburg conception of *Erzeugung* to the whole range of human culture (language, myth, art, religion, statecraft), ending not in panlogism but in "pansymbolism."

Other important Marburg Neo-Kantians were Rudolf Stammler (1856–1938) in the philosophy of law; Karl Vorländer (1860–1928), the historian of philosophy and the leading Kantian socialist (*Kant und der Sozialismus,* Berlin, 1900; *Kant und Marx,* Tübingen, 1911); Artur Buchenau (1879–1946), Albert Görland (1869–1952), and Arthur Liebert (1878–1946). A moderate form of Marburg Neo-Kantianism is represented in America by W. H. Werkmeister (*The Basis and Structure of Knowledge,* New York, 1948).

GÖTTINGEN NEO-KANTIANISM

In strong reaction against Marburg there arose, at the beginning of the twentieth century, the Neo-Friesian school in Göttingen, under the leadership of Leonard Nelson (1882–1927). Jakob Friedrich Fries (1773–1843) had interpreted Kant psychologically, not transcenden-

tally; in this he was followed by Jürgen Bona Meyer (1829–1897) in his *Kants Psychologie* (Berlin, 1870). Lange and Helmholtz were psychologistic in their Kantianism, taking the results of experimental psychology as having a bearing on the a priori. Nelson, on the contrary, professed to avoid psychologism and its attendant skepticism by using psychological introspection to discover the principles of experience in the spontaneity of reason; these principles could then be deduced (in the Kantian sense) from the analysis of experience into its necessary conditions. In this, Nelson developed the views of Fries, whom he defended against the accusation of psychologism, and opposed the psychological or physiological interpretations of the experimental and empirical psychologists.

Kant's transcendental deduction was regarded by Nelson as circular if it was meant as a proof; it began with the experience (science, mathematics, morality) it was meant to justify. The circle might have been broken by Kant's subjective deduction, but this was jettisoned in the second edition of the *Critique.* Nelson proposed to reestablish it, or rather to put his own deduction into its place. Upon introspection, we find principles we know immediately to be true and that we hold by a Cartesian-like "principle of the self-confidence of reason." The discovery of these self-evident principles is a psychological process; the principles, however, are not psychological but metaphysical in Kant's sense; that is, as a priori synthetic truths based on concepts, not on intuition. They are shown to be the same as those uncovered by a transcendental analysis of science and ordinary experience. (In ethics Nelson followed an analogous procedure.) In this way Nelson thought he could use psychology without falling prey to either naturalism or skepticism. A good example of his method is to be found in the well-known *Das Heilige* (Gotha, 1917; translated by J. W. Harvey as *The Idea of the Holy,* New York, 1958) by Nelson's colleague Rudolf Otto. Nelson never had the influence in Germany that was enjoyed by many other Neo-Kantians, although he was revered by many disciples in fields related to philosophy. There has recently been an increased interest in his work, and several English translations have appeared.

HEIDELBERG NEO-KANTIANISM

The Heidelberg school of Neo-Kantianism, led by Windelband and Heinrich Rickert, was not restricted to the University of Heidelberg, and is sometimes known as the Baden school or the Southwest German school of Neo-Kantianism. Wilhelm Windelband (1848–1915) was the most eminent historian of philosophy of his time, with the possible exception of Dilthey. Like Dilthey, he did not succeed in working out a complete system of philosophy, but certain of his ideas were decisive for the more systematic work of his followers in Heidelberg. His most characteristic doctrine was that the epistemological problem is really a problem in axiology; a judgment is known to be true not by comparison with an object (thing-in-itself) but by its conformity to an immediately experienced obligation to believe it. The teaching for which Windelband is chiefly remembered, however, was his distinction between natural and historical sciences as nomothetic and ideographic (law-giving and picturing the unique individual), respectively. The elaboration of these two points led to the systematic priority of axiological criteria to epistemological criteria, to the theory of the parallelism of norms and cultural consciousness, and to efforts to develop a Kantian categorization of historical and cultural experience.

RICKERT. The great system builder of the Heidelberg school was Heinrich Rickert (1863–1936), professor in Freiburg and then Windelband's successor in Heidelberg. Rickert, like Windelband, regarded judging as a form of valuing, truth being the value intended by this act. There are two realms of objects that may be judged; that is, that are objects of knowledge—the sensible world of science (about which Rickert accepted most of Kant's views) and an intelligible world of nonsensuous objects of experience that we know not by perception but by understanding (*Verstehen*). These latter are cultural objects (history, art, morality, institutions). Although not reducible to sense and thus not under the categories of nature, they are not metaphysical but are within experience and correspond, roughly, to Hegel's objective spirit. Both cultural objects and nature, as objects, require (in the Kantian manner) a correlative subject that cannot be objectified. This is "the third realm of being," which Rickert calls "pro-physical"; it is Kant's transcendental ego and Hegel's subjective spirit. There is a fourth realm of being, the metaphysical proper, which is only an object of faith (in the Kantian sense) and which we refer to in religion and in the transition from scientific philosophy to *Weltanschauung.*

By keeping the ethical "this side" of the division between the experiential and the metaphysical, Rickert was able to bring about a closer liaison between the theoretical and the practical than Kant had established. The primacy of practical reason does not, for Rickert, mark the supremacy of valuing over knowing, but signifies the valuational dimension of knowing itself. Autonomy is

thus the basis not only of ethics but also of thought even in science. Rickert criticized the Kantian conception of experience as too thin; not only nature, but also history, must be categorized out of the heterogeneous continuum of data, and from these categorizations arise the nomothetic and ideographic disciplines. In all these points Rickert was under the influence of both Fichte and Hegel, but his conceptual framework remained Kantian: a transcendental nonobjectifiable basis (realm 3) for experience (realms 1 and 2) and an unknown realm of objects of faith (realm 4).

OTHERS. Other important Heidelberg Neo-Kantians were Hugo Münsterberg (1863–1916), Jonas Cohn (1869–1947), Bruno Bauch (1877–1942; *Wahrheit, Wert und Wirklichkeit,* Leipzig, 1923), and Richard Kroner (*Von Kant bis Hegel,* 2 vols., Tübingen, 1921–1924). Kroner's *Kant's Weltanschauung* (Tübingen, 1914, translated by J. E. Smith, Chicago and Cambridge, U.K., 1956) is the only presentation in English of the characteristic Heidelberg interpretation of the historical Kant.

SOCIOLOGICAL NEO-KANTIANISM

Several philosophers close to *Lebensphilosophie* and concerned with the methodology of the *Geisteswissenschaften* were influenced by Kant's doctrine that we categorially construct the world of experience and that speculative metaphysics is impossible as science, but instead of having theories concerning the transcendental origin of the structural factors, they found the origin of the world of experience in the social situation. The most important of these philosophers were Wilhelm Dilthey (1833–1912), who is not usually characterized as a Neo-Kantian although Kantian elements are present in his thought, and Georg Simmel (1858–1918).

At various times Simmel took different attitudes toward, or at least emphasized different aspects of, Kantianism—the psychologistic and pragmatic, the transcendental, and the sociohistorical. He held that categories develop in the course of history, and that the structures of Hegel's objective spirit are historical products that cannot be taken ready-made for analysis in the Marburg manner. "[Even] the kind of science humanity has at any given moment depends upon the kind of humanity it is at that moment" (*Hauptprobleme der Philosophie,* Leipzig, 1910, Ch. 1). Because forms cannot be discerned except in the specific contents in which they appear, no categorial system is capable of structuring all experience. Different types of individuals have different styles for this structuring, and cultures are identified by their production of specific a priori forms for knowledge, the experience of values, and images of the world as a whole (systems of metaphysics).

Between the Heidelberg tradition and the Dilthey-Simmel position there were Max Weber (1864–1921) and Eduard Spranger. Neo-Kantian elements in the sociology of knowledge are especially clear in the works of Max Adler (*Das Soziologische in Kants Erkenntniskritik,* Vienna, 1924) and Karl Mannheim (1893–1947).

Windelband said, "To understand Kant means to go beyond Kant." Most of the philosophers dealt with here did go beyond Kant, and their later works contained little that was specifically Kantian. Even the movements as a whole were more explicitly Kantian in their early periods than in their later ones. All this was to be expected of active and creative minds and groups. By the end of World War I, Neo-Kantianism as an institution ceased to be a dominant force in German intellectual life, partly through the death of most of its leaders and partly through defection. Rapid changes in logic and natural science favored the more pragmatic systems of positivism in Berlin, Prague, and Vienna; the greater experiential resources of phenomenology favored the rival school in Freiburg, Munich, and Cologne; the German cultural crisis called for *Lebensphilosophie* and speculative metaphysics. None of these movements, however, was free of Kantian elements, which might not have been passed on to them but for the Neo-Kantians' rediscovery of Kant. Their Neo-Kantian heritage has given repeated confirmation of an aphorism attributed to Liebmann: "You can philosophize with Kant, or you can philosophize against Kant, but you cannot philosophize without Kant."

See also Cassirer, Ernst; Causation: Philosophy of Science; Cohen, Hermann; Dilthey, Wilhelm; Fichte, Johann Gottlieb; Fischer, Kuno; Fries, Jakob Friedrich; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Helmholtz, Hermann Ludwig von; Herbart, Johann Friedrich; Hönigswald, Richard; Husserl, Edmund; Irrationalism; Kant, Immanuel; Kantian Ethics; Lange, Friedrich Albert; Liebert, Arthur; Liebmann, Otto; Logical Knowledge; Mach, Ernst; Mannheim, Karl; Materialism; Meinong, Alexius; Natorp, Paul; Nelson, Leonard; Otto, Rudolf; Paulsen, Friedrich; Positivism; Psychologism; Rationalism in Ethics; Reinhold, Karl Leonhard; Rickert, Heinrich; Riehl, Alois; Schelling, Friedrich Wilhelm Joseph von; Schlick, Moritz; Schopenhauer, Arthur; Simmel, Georg; Spranger, (Franz Ernst) Eduard; Vaihinger, Hans; Weber, Max; Windelband, Wilhelm.

Bibliography

Studies of and works by individual Neo-Kantians are listed in the respective articles. There is very little material in English on Neo-Kantianism, but see Ernst Cassirer, "Neo-Kantianism," in *Encyclopaedia Britannica,* 14th ed. (1930), Vol. XVI, pp. 215–216; and R. B. Perry, *Philosophy of the Recent Past* (New York: Scribners, 1926), pp. 145–160. A complete history is being written by Mariano Campo; Vol. I of his *Schizzo storico della esegesi e critica kantiana* (Varese, 1959) covers the period up to about 1900. The most complete study, with excellent bibliographies, is K. Oesterreich in *Friedrich Überwegs Grundriss der Geschichte der Philosophie,* 12th ed. (Berlin, 1923), Vol. IV, pp. 410–483.

G. Lehmann reports the beginnings of the movement in "Kant im Spätidealismus und die Anfänge der neukantischen Bewegung," in *Zeitschrift für philosophische Forschung* 17 (1963): 438–457; see also his "Voraussetzungen und Grenzen der systematischen Kantinterpretation," in *Kant-Studien* 49 (1957): 364–388.

Good comparative studies of Neo-Kantianism are included in Wolfgang Ritzel, *Studien zum Wandeln der Kantauffassung* (Meisenheim, 1952) and H. Levy, *Die Hegel-Renaissance in der deutschen Philosophie* (Charlottenburg, Germany: R. Heise, 1927). Johannes Hessen, *Die Religionsphilosophie des Neukantianismus* (Freiburg: Herder, 1924) gives a Catholic criticism.

Authoritative presentations of two school programs are Paul Natorp, *Kant und die Marburger Schule* (Berlin, 1912; also in *Kant-Studien* 17 [1912]: 193–221) and Heinrich Rickert, *Die Heidelberger Tradition und Kants Kritizismus* (Berlin, 1934). The posthumously published (and incomplete) work by H. Dussort, *L'école de Marburg* (Paris, 1963) is excellent on the movement up through Cohen.

OTHER RECOMMENDED SOURCES

Adair Toteff, Christopher. "Neo-Kantianism: The German Idealism Movement." In *The Cambridge History of Philosophy 1870–1945*, edited by Thomas Baldwin. Cambridge, U.K.: Cambridge University Press, 2003.

Brandist, Craig. "Two Routes 'To Concreteness' in the Work of the Bakhtin Circle." *Journal of the History of Ideas* 63(3) (2002): 521–537.

Hallberg, Fred W. "Neo-Kantian Constraints on Legitimate Religious Beliefs." *American Journal of Theology and Philosophy* 16(3) (1995): 279–298.

Kröhnke, Klaus C. *The Rise of Neo-Kantianism: German Academic Philosophy Between Idealism and Positivism*, translated by R. J. Hollingdale. New York: Cambridge University Press, 1991.

Luthe, Rudolf. "The Development of the Concept of Concrete Subjectivity from Kant to Neo-Kantianism." *Journal of the British Society for Phenomenology* 13 (1982): 154–167.

Lewis White Beck (1967)

NEO-MANICHAEISM

See *Mani and Manichaeism*

NEOPLATONISM

GENERAL CHARACTERIZATION

Neoplatonism was the dominant philosophical current in late antiquity, and it had a lasting influence in the Middle Ages when it was adopted by Christian and Muslim thinkers. The term Neoplatonism was coined in the late eighteenth century and was used (in a rather pejorative sense) to distinguish authentic Platonism (as found in Plato's dialogues) from the later systematization and transformation(s) it underwent in the third through fifth centuries, starting with Plotinus.

By using the term Neoplatonism, historians of philosophy wanted to dissociate themselves from the perspective that for centuries had determined, if not distorted, the interpretation of Plato. Yet Plotinus would have been surprised if he had known he would once be called a Neoplatonist. He never intended to be anything other than a faithful interpreter of Plato's doctrines, coming, as he saw it, after centuries of neglect and distortion during which Stoicism and Aristotelianism had set the philosophical agenda, and true, that is, dogmatic, Platonism had, as it were, gone underground in order to survive. This is also how Augustine presents the history of the Platonic Academy in his *Against the Academics*: "Once the clouds of errors had been dispelled, Plato's face, which is the most pure and bright in philosophy, shone forth, above all in Plotinus. This Platonic philosopher is considered to be so similar to Plato that one could believe that they had lived together; but as there is so much time between them, one should think that Plato revived in him." (XVIII 41). One and a half centuries later, Proclus, in his *Platonic Theology* hails Plotinus and his followers Porphyry, Iamblichus, and all others following him, until his master Syrianus (d. 437CE), for having restored Platonism in its original splendor.

PLOTINUS'S RENEWAL OF PLATONISM. What then was so innovative in Plotinus's interpretation of Platonism to praise him so lavishly and to consider him as the founder of Neoplatonism? Plotinus came after two centuries of Platonic revival (in handbooks since Karl Praechter (1858–1933), this period is commonly called Middle Platonism). This does not mean that Plato had ever been neglected during the Hellenistic period. His dialogues, however, seem to discuss problems without arriving at a definite solution, they use dramatic scenery and mythological stories, and do not always provide concordant views. It may have seemed impossible to find in the works of Plato a systematic philosophy that could

compete with that of the Stoics. This could explain why a skeptic, nondogmatic interpretation of the dialogues prevailed for a long time. In the schools of the early Roman Empire, however, Plato was rediscovered as a dogmatic author, and Platonists attempted to systematize his views in handbooks and explain them in commentaries. Many innovations attributed to Plotinus are already present in the Platonists of the first centuries (such as Atticus, Alcinous, and Numenius of Apamea). Recent research has questioned the distinction between Middle and Neoplatonism and stressed once again the continuity of the Platonic tradition. In fact, the debate over the right interpretation of Plato's philosophy had already started in the Old Academy. Neoplatonism is in many respects a development of tendencies already present in the early school and even in the later dialogues of Plato himself as well as in his *unwritten doctrines*, in particular, in the speculations about the derivation of all beings from first principles. This continuity should not, however, make us underestimate the innovative character of Plotinus's philosophy.

The later tradition has always seen the doctrine of the three hypostases—Soul, Intellect, the One (or the Good)—as the most characteristic feature of Neoplatonism and has credited Plotinus with the first clear statement of this theory. Yet most elements of the doctrine are to be found in previous philosophers, as Plotinus himself admits, and, of course, in Plato's own work. With all Platonists, Plotinus strictly distinguishes the sensible from the intelligible realm. The sensible world is not a *hypostasis*, that is, it is not an independently subsisting reality, but depends for its being entirely on incorporeal principles that derive ultimately from the ideal Forms. Only what is incorporeal and intelligible can have hypostatic reality. Within this realm we have to distinguish between Soul, Intellect, and the One, which constitute an ascending series. This theory could strike one as a needless complication of reality and not as its explanation. From a Neoplatonic view, however, these three hypostases are essential steps in the ultimate explanation of all that exits.

Neoplatonism is, in fact, the most radical answer to the question that motivates Greek philosophy since Thales: What are the first principles of all things? To explain a complex reality such as this cosmos means to reduce it to the more simple elements from which it originates. To explain the multiple, Plotinus argues, is to reduce it to its ultimate principle of unity (*anagôgê eis hen*). Whatever exists, exists thanks to its unity. For without unity a thing has no essence, no being, falls apart: A house would no longer be a house but a mere heap of stones; a living being not an organism but flesh and bones; the soul not a soul but a bundle of emotions, memories, thoughts, and so on. Unity, then, is much more fundamental than *essence* or *form*. For being depends on being one. As Plotinus puts it, *being* is a *trace of the One*. Neoplatonism does not primarily offer a theory of being, an ontology as can be found in the Aristotelian metaphysics, but a doctrine of what is one and what ultimately explains unity and is therefore rather a henology. Proclus's *Elements of Theology* start with the proposition that "every multiplicity in some way participates in unity". It is not itself, however, the One, but a unified manifold, having unity as an attribute, and is therefore posterior to the One upon which it depends. For that reason no being can ultimately be explained by a principle of unity that is intrinsic to it. Unity that is participated in depends upon a transcendent principle of unity. Thus the living organism is one thanks to the soul giving life and unity to the body. The One must be identified with the Good, since it is the proper function of the One to hold together all things and maintain them in existence, which is also the function of the Good. For to hold a thing together and make it one is to give it its perfection and well-being whereas dispersion is the cause of its destruction and evil. Therefore, all things pursue unity as the good because they all strive to continue to exist and shun division as evil. Therefore, the One is to be identified with the Good, and the origin of the procession (*proodos*) of all things is also the end of their return (*epistrophê*).

In our search for an ultimate explanation, we will find always higher levels of unity until we arrive at the One itself. The whole sensible cosmos is one complex living organism wherein all things are connected in a chain of causes and linked by mutual sympathy, as the Stoics said. But what explains the unity and coherence of this world cannot itself be a material principle, such as the Stoic active principle, but has to be an incorporeal world soul. As Plato argued in the *Timaeus*, the soul is an intermediate between the sensible and the intelligible, the temporal and the eternal. But because it is incorporeal, the soul, at least the rational soul, is never entirely cut off from the intelligible world, not even when it is incarnated in a body. The soul, however, is not itself the origin of the specific forms and of the organic structure incorporated in this world. Whatever the soul (as *demiurge* or creative cause) conveys to this world derives from the ideal Forms contemplated by it. In fact, all production results from contemplation. If one subtracts from this sensible world matter, mass, spatial differences and time, coming to be, corruption and death and only understands what is

essential and eternal in it, one finds a wonderful organism, an articulated system of specific forms, eternal objects of thought. This is the intelligible world, true reality and divine Intellect, as one perfect science that comprehends in itself all being known in its essential structures. Although comprehending all forms eternally and at once, this self-thinking Intellect or Intelligible Being cannot be the ultimate explanation of the universe, as Aristotle thought. For it is characterized by the multiplicity of the Forms and by the duality of thinker and object of thought. This leads Plotinus to a provocative conclusion that seems to go against the grain of philosophy itself: "For thinking itself does not come first either in reality or in value, but is second and is what has come into being when the Good [already] existed." (V 6, 5, 5–6). This Good is, as Plato famously said, *beyond* (*epekeina*) thinking and being. It desires nothing, needs nothing. It is just One. Because it is nothing, it can be the origin of all things, not because it creates or produces them, but because they all come forth from its overflowing simplicity. Characteristic of Neoplatonism is this double transcendence: that of the Intelligible with respect to the sensible and that of the Good with respect to the Intelligible.

A SPIRITUAL EXPERIENCE. The amazing success of Neoplatonic philosophy, also beyond the limited circle of pagan philosophers, cannot be explained solely by elements of the doctrine. What made it so attractive was that it not only offered a theoretical understanding of reality, but also promised a way to ascend to the first principle of all, bringing the soul back to its own origin. Philosophy begins with the Delphic maxim *know thyself*, which is understood as an exhortation to return into thyself. "Go back into yourself and look," says Plotinus (I 6, 9, 7). This *epistrophê*, or return, of the soul upon itself is also the beginning of the return to the intellect and the One from which the soul proceeded. For within itself the soul does not only discover its own essence but also has access to the intelligible world to which it belongs essentially. Plotinus tells us of his personal experience: "Often I have woken up out of the body to my self and have entered into myself, going out from all other things. I have seen a beauty wonderfully great and felt assurance that then most of all I belonged to the better part; I have actually lived the best life and come to identity with the divine" (IV 8, 1, 1ff.). The truly wise person therefore "has already finished reasoning and turned to himself: all is within him" (VI 5, 12 17–18). The three hypostases, Soul, Intellect, the One do not solely exist *in nature*: We find them in ourselves, at least if we first discover that we are a *self*. Through a moral life we have to gather our self from the fragmentation of the daily needs of the body, which distract our attention toward the outside. We are more than souls taking care of our body. We belong to the intelligible world, or rather, each of us is the intelligible world, and in our deepest self, we are one, one with one another, one with the One cause of everything.

The different hypostases of reality are not just three levels of reality; they are different levels of spiritual existence, or different modes of being *self*. Neoplatonic philosophy is not just a theory about unity, for such a theory could never succeed on its own. It is an exhortation to find the one by becoming one and simple, eventually giving up reasoning and explanation, just being one, or even going beyond being, by reaching an *ecstatic experience*. This unification with the One is not an alien supplement, not a denial of philosophy, but a realization and radicalization of what always was the intention of philosophy: to reach the first principle; to overcome the distinction of knower and object known.

NEOPLATONISM: THE FULFILLMENT OF HELLENIC CULTURE. Neoplatonism is not just an effort to offer a comprehensive understanding of the Platonic doctrines scattered all over the dialogues. It also integrates within this Platonic perspective the whole philosophical tradition starting with Pythagoras. Aristotle himself is seen as essentially a Platonic thinker, at least if purified of the distortions of some later Peripatetics. Without a full knowledge of the Aristotelian logical writings and his treatise *On the Soul* it is not possible to understand the subtle Neoplatonic theory of knowledge. Aristotle's analyses of substance, matter and form, potency and act, quality and quantity, the different forms of causality provide the conceptual framework in which Plato's arguments are construed. To the Neoplatonists we owe the great commentaries on Aristotle, which made possible the reception of his philosophy by the medieval thinkers. When Neoplatonism took over the intellectual hegemony, after five centuries of being dominated by Stoicism, it also adopted many Stoics doctrines, in particular (part of) their ethics, and their views on providence and fate. Thus, they secured it an influence beyond antiquity. In short, Neoplatonism not only comes at the end of ancient philosophy, it integrates, in a way, the whole philosophical tradition in all its richness and diversity, making a synthesis of what had been for a very long time opposing schools.

In contrast to Plotinus, the later Neoplatonists became increasingly interested in the wisdom transmit-

ted through the ancient religious traditions, not only the Hellenic religion (as it was known through Homer and Hesiod (c. 700 BCE) between the Orphic revelations), but also the arcane doctrines and rituals of the barbarians, in particular, the Egyptians and Chaldaeans. Of particular interest for the later development of the school were the so-called *Chaldaean Oracles*. These oracles offer, in epic hexameters, a mythical theogony and cosmogony of Platonic inspiration. They are supposed to have been revealed by the gods to a certain Julian the Chaldaean and his son, the *theurgist* (c. 160–80 BCE). The term *the-urgy* (divine work) indicates certain ritual actions, which connect those who practice them with the gods. From Iamblichus onward, the *Chaldaean Oracles* gained a considerable authority comparable only to that of the sacred texts of Jews and Christians. This positive attitude toward the diverse religious traditions did not, however, include Christianity. Porphyry and Iamblichus wrote polemical treatises against the Christians and, following them, the emperor Julian, called the Apostate (331–363), even started persecuting them. They considered Christianity as a threat for the whole of Hellenic culture with its tradition of education, literature, religious practices, and philosophy. The intolerant attitude of the Christians made it impossible to integrate their views together with the other religious traditions in one comprehensive *Platonic theology*. The growing opposition against Christianity may explain why Neoplatonic philosophy itself, from Iamblichus onward, became increasingly *theological* in its project. The Christian authors liked to point to the contradictions within the pagan philosophical tradition. They perceived all schools to have divergent opinions, which would almost naturally lead to skepticism. In response to this the Neoplatonists made an attempt to systematize and reconcile the most diverse doctrines from an overall Platonic perspective, integrating in it all that was valuable in the mythological and religious traditions. Just like the Christians they had their own sacred books (which were wonderfully in agreement with Plato's wisdom), and their theurgical practices could be seen as a rival to the sacramental practices of the Christians aiming for the salvation of the soul.

At the end of antiquity, in particularly in the Athenian school, Neoplatonism had thus become the ideological justification of the old pagan culture wherein all the wisdom of the Hellenic tradition was integrated: the theology of Homer, Hesiod and Orpheus, Pythagoras, Parmenides, Plato himself, and also Aristotle and the Stoics.

HISTORICAL SURVEY

THE LEGACY OF PLOTINUS. Plotinus undoubtedly set off the Neoplatonic movement, though it is difficult to call him the founder of a school. His philosophy was in a way too *original*, too much linked to his own spiritual experience. Plotinus is provocative and daring in his expression, as he himself admits, as when he says that the soul is never fully distanced from the intellect. From a scholarly point of view, much in what he says remains unclear: How can the One be beyond all things and still be the *power of all things*; how can the One bring forth a multiplicity; what exactly is the role of the soul in the production of the World; and so on. Particularly challenging was Plotinus's philosophical appropriation of religion. The philosopher is the true priest who can ascend within himself to the divine principle of all. He has no need to go to temples, the gods "will come to him" (*Vita Plotini*, 10). Enough questions to stimulate further debate in the later school for over two centuries.

It would wrong, indeed, to see Neoplatonism as a unified movement: There was considerable divergence within the school, with conflicting interpretations of Plato; different views on essential points of the doctrine, such as the status of the One and the explanation of the procession of all things; the relation between the Intellect and the intelligible and the status of the Ideas; the role of the demiurge in the creation of the sensible world; the function of demons and other intermediary beings; the nature of the soul and its relation to the intelligible world; and above all, the role of theurgy. Nevertheless, all shared a common doctrine, the three hypostases: the transcendence of the One, the distinction between the sensible and the intelligible, the return upon the self as the origin and the end of philosophy.

The following survey shall sketch the main lines of the historical and institutional development of Neoplatonism, referring to the relevant entries in this Encyclopedia for more in-depth studies of major figures.

THE FIRST GENERATION AFTER PLOTINUS. After his arrival in Rome, Plotinus soon attracted to his lectures students and devotees who often belonged to the high Roman society. We are well informed about the intellectual climate in this close circle—about the texts that were read and the topics they discussed, about the interaction in the group—thanks to the *Life of Plotinus* written by his close disciple Porphyry as an introduction to his edition of the works of his master. As Porphyry tells us, Plotinus for a long time refused to write down his lectures. Only at the age of forty-nine, at the insistence of his students, did

he start scribbling down his arguments. It took Porphyry a great effort and a long time to make the texts ready for publication. The *Enneads*, as they were called (they consist of six groups of nine essays), were published about thirty years after the death of the master. This edition made the reputation of Plotinus and gave his thought a wide circulation beyond the circle of his immediate disciples. Soon a Latin adaptation of the work was made (probably a selection), which attracted enthusiastic readers among young intellectuals in Milan, as the example of Augustine shows. Porphyry also wrote a systematic introduction to Neoplatonic philosophy, the "Pathways to the Intelligible," making abundant use of material from Plotinus. Without the effort of Porphyry, the philosophy of Plotinus, this original individual, would never have had such an immense influence on the development of late antique and medieval thought. Porphyry defended *the harmony of Plato and Aristotle* (this is the title of one of his lost works) and contributed to the reception of Aristotle's works in the Neoplatonic curriculum as an introduction to the study of Plato. He wrote two commentaries on Aristotle's *Categories* and a short *Introduction* (*Eisagôgê*) to the study of categories, which soon gained the authority of an Aristotelian treatise.

In a famous treatise (the concluding part of which is known as Ennead II 9 [33]), Plotinus attacked some Gnostic Christians and defended the beauty of the Cosmos against their dualistic views. Porphyry in his *Against Christians* launches a direct attack against the Christians. This anti-Christian outlook would also be that of the later school. Despite his anti-Christian polemics, Porphyry has a great interest in the diverse religious traditions as a source of wisdom. He is the first philosopher to pay attention to the *Chaldaean Oracles* and is fascinated by the theurgical rituals as a means to achieve the *salvation of the soul* (that is, the return of the soul to God). But, maybe under the influence of Plotinus, he adopted a more intellectual interpretation of religion, which led him to question theurgy and other aspects of the Egyptian religion (for which he would be criticized by Iamblichus). Hence, Porphyry limits the efficacy of theurgical practices to the lower degrees of salvation (those concerned with the purification of the pneumatic body and the lower soul) while demanding strictly philosophical means for achieving the union with the One.

THE SYRIAN SCHOOL OF IAMBLICHUS. The Syrian Iamblichus stayed for some time as a student with Porphyry in Rome. He had, however, diverging views on many issues and did not hesitate to attack Porphyry in writing. Having returned to his native Syria at the end of the third century, he set up his own school at Apamea. While Porphyry's influence remained mostly limited to the Western part of the Empire (including the Latin tradition), Iamblichus left a definitive stamp on the development of Neoplatonism in the Greek world, both through his metaphysical speculations on the first principles and his passionate defense of theurgical practices. Whereas Porphyry, interpreting Plotinus, intended to see the One as the summit of the Intellect, Iamblichus emphasizes even more the transcendence of the first principle, putting the Ineffable even beyond the One. Within the intelligible realm, he further distinguishes the purely intelligible from the intellectual level. And whereas Porphyry, following Plotinus, identified the supreme part of the soul with the intellect, Iamblichus insists that the soul is a separate ontological entity, intermediate between the intelligible and the sensible and therefore lower than intellect. Situated between the soul and the intellectual gods, the classes of demons, angels, and heroes have an important mediating function. All this announces a tendency that will become dominant in the later development of the school: the introduction of ever more intermediaries in the procession from the One to the multiple to make the transition from one level to another less abrupt. It is also Iamblichus who introduces the distinction between a non-participated and a participated status of a principle (such as soul or intellect). He also develops the triadic schema of remaining, procession, and reversion and applied this and other structures to different ontological levels. Iamblichus seems to have developed all important principles that support the architecture of Neoplatonic metaphysics. He also deserves credit for having established the educational canon of Plato's dialogues as well as their reading order and for having developed the exegetical principles for the interpretation of Plato, the most important of which being the determination of the right scope or intention of a dialogue. Iamblichus also initiates the Pythagoreanizing trend in Neoplatonism. He considers Pythagoras as the real founder of the philosophical tradition in all of its branches and as the model of the philosophical life. Plato himself, so Iamblichus believes, was the most eminent exponent of that tradition. Iamblichus's Pythagorean leanings also explain the heavy emphasis on mathematics as the most universal science, having applications in all possible branches of philosophy, not only in physics, and astronomy, but also in ethics and theology. For his attempt to fuse Pythagoras and Plato into one mathematical–metaphysical system, Iamblichus could find inspiration in Neopythagorean authors of the first centuries CE, such as Nicomachus of Gerasa (c. 60–120 CE).

Even more important for the future development of the school is Iamblichus's novel attitude to religious rites. He could not agree with Porphyry's reserved rationalistic attitude toward religious practices and theurgy in particular, as is evident from his anonymous reply to the latter's *Letter to Anebo* (an Egyptian priest). Iamblichus's reply, since the Renaissance known under the title *On the Mysteries of the Egyptians*, is a comprehensive defense of religious practices, magic, and sacrifices:

> It is not thought that links the theurgists to the gods: for otherwise what should prevent the theoretical philosopher from enjoying a theurgic union with the gods? But this is not the case; theurgic union is attained only by the perfective operation of ineffable acts worthily performed, which are beyond all understanding, and by power of the unutterable symbols, which are intelligible only to the gods.

By thus insisting on the necessity of the practice of theurgic rites to accomplish the union with the gods, Iamblichus rejects, as E. R. Dodds notes, "the whole basis of the Plotinian intellectual mysticism" and "opens the door to all those superstitions of the lower culture which Plotinus had condemned in that noble apology for Hellenism, the treatise *Against the Gnostics*." (Dodds 1963, p. XX with quotation of *De myst*. II 11).

Some of Iamblichus's students devoted a lot of attention to the philosophical justification of magical and esoteric practices. They set up a school in Pergamum that seems to have gained some reputation when one of its students, Julian, became emperor. Julian drew upon Neoplatonic philosophy in his attempt to restore pagan rituals and traditions against the increasing influence of the Christians. Sallustius (fl. fourth century CE), who published a small introductory manual of Neoplatonic theology *On the Gods*, was probably a member of the same school.

THE ATHENIAN SCHOOL. The philosopher Plutarch of Athens (d. 432) gave a new inspiration to the Platonic Academy in Athens, which from then on adopted the philosophical style of Iamblichus. Although they no longer taught in the original building of the Academy, the successive heads of the school in Athens proudly considered themselves to be the "*diadochoi*," successors of Plato. Of Plutarch we have only indirect and fragmentary evidence. Proclus attributes to him an important role in the search for the right interpretation of the Parmenides. As a young student, he read with him Aristotle's treatise On the Soul and Plato's Phaedo. One would like to know how Plutarch attempted to reconcile the opposing views of Plato and Aristotle on the nature of the soul and its immortality, and on the origin of knowledge (anamnesis vs. abstraction).

After Plutarch's death in 432, Syrianus, a native from Alexandria, became the new head of the school. Of Syrianus we have only a commentary on some books of the *Metaphysics* in which he is often very critical of Aristotle. He recognizes Aristotle's great contribution in logic, ethics, and natural philosophy, even in theology. But, as he says, Aristotle's attack on the doctrine of the first principles of Pythagoras and Plato (an in particular, the doctrine of the Forms) is so unfair and shows so much misunderstanding that he felt compelled to defend the truth by showing Aristotle's arguments to be invalid (*In Metaph*. 80, 4-81, 14.)

When Syrianus died (c. 437), he was succeeded by Proclus who was born from a Lycean family still faithful to the old religion and had come from Alexandria to study philosophy in Athens. After a short term with Plutarch, Proclus continued his philosophical education under the guidance of Syrianus: "In less than two years Proclus read with him all of Aristotle's treatises on logic, ethics, politics, physics, and the theological science which surpasses them all. When Proclus was suitably educated through those studies which, so to speak, are a kind of preparatory initiation, or lesser mysteries, Syrianus led Proclus to Plato's mystagogy." (Marinus, *Life of Proclus*, §13).

Because of the loss of most of Syrianus's, work, it will never be possible to determine which ideas and doctrines Proclus inherited from his master and which ones he contributed himself. But it is evident that Syrianus had a profound influence on Proclus, as the latter gratefully acknowledges: "It is he who has granted us the privilege of partaking in the philosophy of Plato as a whole and who has communicated to us what he had received in secret from those senior to himself, and, above all, who joined us with himself as co-celebrants of the mystical truth of the divine principles." (*Theol. Plat*. I 1, p. 6.16-7.8 ed. Saffrey-Westerink, transl. J. Dillon). As is clear from this text, Proclus understands his Platonic education not just as a transmission of a philosophical doctrine but as a revelation of a mystical truth coming from the gods through Plato, and even as an initiation in a mystery cult and a participation in a ritual practice of life.

As we know from his biographer (and successor) Marinus (c. 440–c. 500), Proclus's whole life was devoted to teaching and writing. He wrote commentaries on the Platonic dialogues that were part of the Neoplatonic-

school curriculum. The course started with the reading of the *Alcibiades I*, a dialogue about self-knowledge, which was regarded as an introduction to philosophy. The curriculum culminated in the explanation of the two major dialogues of the Platonic corpus, which were considered to incorporate the whole of Plato's philosophy, namely the *Timaeus* (about the generation of the physical world), and the *Parmenides* (about the procession of all beings from the One). The commentaries of Proclus are masterpieces in their genre, as they not only offer a systematic interpretation of the text but also provide a wealth of information about the discussions within the Platonic tradition. In addition to his commentaries, Proclus owes his reputation to his two great syntheses of Neoplatonic philosophy, the *Elements of Theology* and the *Platonic Theology*.

In the *Elements of Theology*, Proclus demonstrates *in a geometrical way* the most fundamental theorems of the theological or metaphysical science as he understands it. The first part examines the fundamental principles that govern the structure of all reality, such as the relation between the One and the many; cause and effect; whole and parts; transcendence and participation; procession and reversion; continuity and discontinuity. In the second part he expounds the procession of the divine principles (henads, intellects, souls). The *Elements of Theology* is without doubt his most original work, not so much because of its content (which offers the standard doctrine of the Athenian school) but because of its extraordinary attempt to develop the entire Neoplatonic metaphysics from a set of axioms. It also had a tremendous influence, in particular through the Arabic adaptation that was made in the ninth century in the circle of Al-Kindi (805–873). In the middle of the twelfth century, this Arabic treatise was translated into Latin. The *Liber de Causis*, as it was named, circulated as the work of Aristotle and thus obtained a great authority in medieval scholasticism. The systematic character of the *Elements* and its rigorous method make it the best introduction for the student not only to Proclus's own thought but also to Neoplatonism in general.

Proclus was convinced that the truth about the gods had been revealed in many different ways—in obscure oracles, myths, and symbols. It was his ambition to prove the harmony between Plato and the other sources of divinely inspired wisdom, in particular, the *Chaldaean Oracles* and the Orphic poems. In his view only a genuinely philosophical approach could offer the conceptual framework for such a comprehensive interpretation. One finds such a framework in the *Parmenides* if one adopts the theological interpretation of this dialogue developed first by Syrianus. The *Platonic Theology*, written at the end of Proclus's life, is the perfect realization of this theological project—a pagan Summa of theology.

It is difficult to evaluate the originality of a thinker who, in most of his works, proclaims to be nothing but a faithful follower of his master Syrianus. But it is Proclus who put his mark on the subsequent development of Neoplatonism in Byzantine, Arabic, and Latin medieval thought. His huge influence—much greater than that of Plotinus—could extend itself mainly through two important indirect channels of transmission: the Arabic adaptation of the *Elements* in the *Liber de Causis*, and the Christianization of his Platonic theology by Dionysius the Areopagite. The latter author pretends to be, and was for centuries believed to have been, the Dionysius mentioned in the Acts of the Apostles who became Christian after the preaching of Saint Paul on the Areopagus (Acts 17:34). This *authorship* gave this work an almost apostolic authority both in Byzantium and in Latin Europe. Although the real identity of this author still remains unknown, he probably was a Syrian Christian who followed classes in Athens at the end of the fifth century (he may even have been a direct disciple of Proclus). In his works, and in particular in his treatise *On the Divine Names*, he expounded the Christian doctrine of the transcendent God, of the Trinity, and of creation and incarnation in terms of Proclus, eliminating references to the pagan religion and substituting the Christian sacred writing for the *Chaldaean Oracles*.

Among Proclus's fellow students under Syrianus were Hermias, who would return to his hometown Alexandria and start teaching there, and Domninus of Larissa (c. 420–480), who had a predominantly mathematical interest and was criticized by Proclus for his unorthodox interpretation of Plato.

On the further history of the Platonic school in Athens at the turn of the fifth century, inside information is provided by Damascius, the last head of the school, in his *Life of Isidore* (Isidore [fifth century] was his predecessor). Thanks to his energetic reforms and inspiring teaching, the Academy would revive one last time. Damascius is known, among other things, for his commentaries on the *Philebus* and the *Parmenides*, but above all things, for his treatise *On the First Principles* (*De principiis*). This work concludes a period of a thousand years of philosophical speculation on the first causes. Damascius has no ambition to develop a system that would surpass that of his predecessors. His own thought is primarily aporetic: He raises critical questions in the

margin of the doctrine of the principles as it had been developed in the Neoplatonic tradition and confronts it with all sorts of difficulties. When he risks a solution—and on many issues he can be very original (for instance, his doctrine on time)—he again calls it into question by raising new aporias. The most fundamental aporia is discussed at the beginning. Is the first principle itself a part of the whole of which it is the principle? The first, it seems, is neither principle nor cause nor does it fit in any other category used to explain relations between beings: It is an ineffable *nothing* we have to postulate beyond the one whole. This ineffable is even beyond the One, which is the first principle of all things. More than any other Platonic philosopher, Damascius is aware of the precarious nature of all rational discourse when dealing with questions that go beyond the limits of what can be experienced. About the first principles we can only speak by making use of analogies and *indications*. His sharp critical mind does not, however, lead him to skepticism. If a philosophical explanation remains tentative and fragile, there is also the mythological tradition and religious practice, to which Damascius remains very devoted. In many ways his work is a wonderful swan song of pagan Hellenism.

The renaissance of the Academy under Damascius may have been one of the reasons for its closing by a decree of the emperor Justinian (c. 482–565) in 529. The decree is one of the multiple measures of the emperor against pagans: They were formally excluded from all official positions, including teaching. According to the historian Agathias (536–582), Damascius, together with Simplicius, Priscianus the Lydian, and other philosophers went into exile at the court of King Chosroes (?–579) in Persia. After two years Chosroes concluded a peace treaty with Justinian, which contained a clause about the exiled philosophers: "They were free to return to their country and live quietly by themselves without being compelled to accept any belief against their conviction or to renounce the creed of their fathers" (Agathias, II, 28–32 ed. Keydell, transl. Westerink). Whether they returned to Athens or Alexandria or stayed in other places remains uncertain.

ALEXANDRIAN SCHOOL. Alexandria had always been a city with a dynamic intellectual life, and it remained so in late antiquity though Christian theological debates now dominated the scene and church authorities set restrictions to the teaching of pagan philosophy. A notorious case, symbolic of the changing times, is the lynching of Hypatia in 415 by a Christian mob. Educated by her father Theon (335–405), Hypatia had become an outstanding mathematician. What her philosophical interests were are unknown, but among her admiring disciples was Synesius (c. 370–414), author of *On Dreams* of Neoplatonic inspiration, who also shows an interest in the *Chaldaean Oracles* even after he had become a Christian bishop.

The first to introduce Neoplatonic philosophy in Alexandria was Hierocles (c.400–460 CE), who studied in Athens with Plutarch. He is the author of a commentary on the *Golden Verses* of Pythagoras and a treatise *On Providence*. In the introduction of the latter work, he criticizes "all those who try to break up the unanimity of Plato and Aristotle". Thanks to his master Plutarch, he was educated in a tradition that harmonizes the thought of both great philosophers and goes back to Ammonius (c. 175–243 CE), who was teacher of Plotinus in Alexandria: "This man, Hierocles says, was the first to bring the teachings of Plato and Aristotle into one and the same view and to transmit a philosophy without factions to all his students." This hermeneutical approach—different from the more polemical attitude to Aristotle of Syrianus and Proclus—would be continued in Alexandria by the following generations of philosophers and find its magnificent expression in the great commentaries on Aristotle of Simplicius.

The leading Neoplatonic philosopher in Alexandria was another Ammonius (c. 440–526) who had come from Athens with his father Hermias. In his youth Ammonius followed courses with Proclus, and he would adopt the basic principles of the latter's Neoplatonic synthesis. Of Ammonius, however, we possess only commentaries on Aristotle, one of which he wrote himself (on *De Interpretatione*), others of which were published in the form of lecture notes by his students. Since most of his teaching was devoted to the explanation of Aristotle's logic and (meta-)physics, the typical Neoplatonic doctrines (the three hypostases; the procession of all things from the One; the structure of the intelligible world; the ascent and mystical union of the Soul) are rarely discussed and explained. Had we also had Ammonius's commentaries on Plato, the picture might have been somewhat different. But it may also be the case that Ammonius intentionally avoided controversial subjects as he noticed the growing number of Christian students in his audience. The Alexandrian School was a much more open system of education than the Athenian Academy, which had in its last phase become somewhat of an esoteric group. However, from the extant texts, it emerges that Ammonius had more interest in explaining the structure of the physical world than in elucidating the architecture of the intelligible world.

Scholars have often said that the Alexandrian School represents a different kind of Platonism from that of Athens:

> In Athens the speculative, mystical, theurgic, and religious elements predominated; and that school remained to the end a stronghold of paganism. In Alexandria scholarly interests and a noncommittal exegesis of texts prevailed. The Platonism that the Alexandrian School professed was in some respects closer than that of the Athenian School to the pre-Plotinian version; thus, the doctrine of the ineffable One and the mystic union with it had no prominent place.… Thus, the "baptizing" of Greek philosophy—including the stress on those parts of the Aristotelian philosophy that were metaphysically neutral—so often considered characteristic of the medieval period, was to a certain extent anticipated in Alexandria; after the Arab conquest it was perhaps replaced by 'Islamizing."

Thus writes Ph. Merlan in the first edition of this Encyclopedia, following the views of Praechter. Recent studies (in particular, by Ilsetraut Hadot), however, tend to minimize the differences between the two schools. There were indeed very close relations, even family relations, between the members of both schools, and there was a lively intellectual exchange. All members were educated in the same tradition. The fact that some doctrines are less prominent in the extant works of the Alexandrians can be explained by the fact that only their work on Aristotle have come down to us. A close reading of the works of the Alexandrian philosophers shows that they had fundamentally the same views on the most important metaphysical issues (such as the distinction between the demiurge and the absolute One) as their colleagues in Athens. And yet it cannot be denied that there are important differences between the two schools and that the view of Praechter and Merlan contains some truth. First, as noticed, there is the harmonizing, not polemical, approach to Aristotle. One may even go so far to say that the Alexandrians were primarily interested in presenting a Platonized Aristotle. Second, though Hadot may be right in denying that the Alexandrian thinkers return to a pre-Plotinian form of Platonism, they tend to simplify considerably the highly complicated Proclean system. Third, the philosophy of the school of Ammonius is less connected with openly pagan beliefs. The project of a comprehensive Platonic theology seems to be alien to them. According to Damascius (who speaks about it with contempt), Ammonius had concluded a pact with the patriarch Athanasius. We do not know what concessions he made to preserve the freedom of teaching in the school. Maybe he promised not to discuss certain doctrines contrary to Christian faith, such as the eternity of the world or the preexistence and reincarnation of the soul.

Two of the most famous students of Ammonius deserve special mention: John Philoponus and Simplicius. The latter is rightly famous for his voluminous commentaries on the *Physics*, the *De Caelo*, and the *Categories* (the commentary *On the Soul* is not his work but probably of his colleague Priscianus), which still are of great use to any interpreter of Aristotle. Simplicius attended Ammonius's courses on Aristotle, but he mentions also Damascius as his teacher. This double education situates him somehow halfway between Alexandria and Athens. He is well acquainted with Damascius's metaphysical speculations (on the procession of all things, on time and place), but never forgets the first intention of his work, which is to offer a faithful elucidation of the views of Aristotle in a Neoplatonic perspective. His commentaries also contain rich historical and doxographical information on the Presocratics (of whom he preserves many fragments), on Stoic philosophy, and on the later developments of the Peripatetic and Platonic school. He also quotes long sections from Plato's *Dialogues* and misses no opportunity to demonstrate that there is no contradiction between Plato and Aristotle in doctrinal matters. When Aristotle does seem to attack his master, so Simplicius argues, his critique only concerns the manner in which Plato expresses his views. For Plato often uses a narrative form and a metaphorical language, which, if taken literally, may lead the reader to erroneous views. To defend the harmony of Plato and Aristotle was for Simplicius also of great strategic importance in his controversy with the Christian Philoponus. The latter liked to exploit the oppositions within the philosophical tradition in order to undermine it.

Philoponus was one of the brightest students of Ammonius. He published several of his lecture courses and continued to comment on Aristotle in the manner of his master. What sets Philoponus apart from the other members of the school, however, is the publication of a treatise against Proclus in which he attacked, from an overtly Christian point of view, the doctrine of the eternity of the world. Philoponus attempts to prove that the world had a temporal origin and that this was also the authentic doctrine of Plato in the *Timaeus*. In the later versions of his commentaries on Aristotle, he adopts the same polemical attitude whenever he finds Aristotle in

contradiction with the Christian understanding of creation.

The publication of *Against Proclus* in 517 must have provoked quite a scandal in the school, where Philoponus was one of the leading figures. Scholars have advanced many solutions to explain his sudden change from a Neoplatonic to a Christian philosophy. The fact that the publication of the polemical treatise coincides with the closing of the Academy in Athens and with other hostile measures that were taken against pagan philosophers may provide a useful clue. By publishing his book against Proclus, Philoponus probably wanted to distance himself from the allegedly pagan elements in Neoplatonic philosophy. In his later work he is only engaged in theological discussions.

Simplicius says he never met Philoponus and speaks of this *newcomer*, not really a *philosopher*, with utter disdain. In his commentary on the *De Caelo*, he came to the defense of Aristotle and of the old Hellenic pagan view of the cosmos as an everlasting, wonderful expression of the intelligible world.

The successor of Ammonius as head of the school was Olympiodorus (c. 500–565). We also have some of his commentaries on Plato, which show that he did not consider himself to be a Christian. For he continued to defend, though with caution and without offending his audience, some views that belonged to the pagan tradition. He upheld polytheism by explaining the lower gods as powers of the first God rather than as many gods.

Olympiodorus's two pupils, Elias and David, who lectured on Aristotle's *Organon*, certainly were Christians though their belief does not really have an impact on their teaching. The last teacher in the school was Stephanus, who became professor at the newly founded academy in Constantinople (in 610). The transfer of the school (and its library) to Constantinople may explain why so many works of pagan Neoplatonists have survived.

LATIN NEOPLATONISM. In the western part of the empire, too, we find authors who were influenced by Neoplatonic ideas. Since they all wrote in Latin, they would have a determinative influence on the formation of Medieval Platonism. There are, of course, Christian thinkers, such as Ambrose (c. 339–397), Marius Victorinus (c. 280–365), and above all Augustine, who all considered Plato closer to Christian faith than any other philosopher. Yet besides them there also was a small group of authors who continued to practice philosophy in the old tradition. Even if they were Christians, their beliefs had almost no effect on their arguments (contrary to what we see happen in Augustine). A good example is Calcidius (late fourth century), who translated and commented the *Timaeus* and followed Porphyry in many of his interpretations. His work had an immense success in the early Middle Ages. The same is true for the *Commentary on the "Dream of Scipio"* by Macrobius (c. 400), who quotes also from Plotinus and Porphyry. Also Martianus Capella (early fifth century), author of the much read *On the Marriage of Philology and Mercury* that offers an allegorical introduction to the seven liberal arts and makes them part of the philosophical wisdom, shows a thorough acquaintance with the Platonism of late antiquity. Last but not the least is Boethius who is undoubtedly a Christian (as his theological work shows). Yet in his practice of philosophy, he does not allow Christian arguments to interfere directly. He is author of the celebrated *Consolation of Philosophy*, which is profoundly Neoplatonic in its argument, but he also wrote translations and commentaries on Aristotle. It was his ambition to translate and comment on all of Plato's and Aristotle's works and to demonstrate that they are in agreement on fundamental questions. This program situates him in the tradition of Alexandria, with which he was well acquainted. He shows also to be familiar with the works of Porphyry and Plotinus.

EPILOGUE: CHRISTIAN NEOPLATONISM

As we have seen Neoplatonic philosophy from the beginning took a very polemical attitude toward Christianity. Plotinus attacked some Gnostic Christians in his entourage; Porphyry wrote a vehement attack against the Christians, as did Iamblichus and Julian. The latter even used the Neoplatonic philosophy in his policy of restoration of paganism. In the Athenian School Neoplatonism became the ideology of pagan religion in its multiple guises. When Christianity became the dominant religion, philosophers had to be more cautious and could only make indirect criticism. Proclus and Damascius just ignored Christian thought and looked down with contempt upon the Christian establishment. The Christian authors, of course, attacked paganism, but were, on the other hand, surprisingly positive toward Neoplatonism, which they considered to be the philosophy that came closest to the Christian Weltanschauung. This is the case for Augustine in the west and for Gregory of Nyssa and Gregory of Nazianzus in the East. The reasons for this fascination are manifold: the other-worldness of Neoplatonism; the emphasis on the transcendence of God; the

nondualistic doctrine of creation (procession); the spiritual antimaterialistic interpretation of the world; the immortality of the soul; the access to the divine through the soul's return upon itself. The differences were no less evident, in particular, the doctrine of incarnation, personal providence, and the belief in resurrection. Whereas Christian thinkers were often deeply influenced by Neoplatonic thought, the pagan philosophers, on the contrary, showed no influence from Christianity: They absolutely ignored it. There was no interaction between Neoplatonism and Christianity, only a strong influence in one direction.

The integration of Neoplatonic arguments in the explanation of the Christian wisdom give rise to original speculations about creation, the world, the place of humankind, and the relation of soul-body. Some scholars may argue that this Christian appropriation of Neoplatonism is a betrayal of the original spirit of philosophy. But this transformation is in itself a wonderful testimony to the creativity of Neoplatonic thinking. Take the concept of the *self*, which in Neoplatonism gained a much greater richness than ever before in Greek philosophy. Augustine took over the notion of self-reflexivity but gave it an incredible concrete existential richness, making it a leitmotif of his autobiography (*Confessions*). Another example is eschatology. According to the Neoplatonic view, the procession and return are constitutive movements of each being in relation to its cause. Christian thinkers historicized this process: At the beginning of time, all things proceeded from God and will return to Him at the end of time. This interpretation made it possible to give a meaning to history and even to the contingent events of human life.

Thanks to this creative modification, Neoplatonism had a continuing and expanding influence after the death of the pagan intellectual culture. Already prior to Justinian's decision to close the school of Athens, the pagan philosophical tradition had become a rather marginal phenomenon in late antique civilization. Its practitioners were an esoteric group of intellectuals, nostalgic for the past glories of Hellenic culture, practicing magical rituals, and praying to old gods. Pagan Neoplatonism had become an ideology at the service of a disappearing civilization. Once this philosophy became integrated in the Christian culture, and later in the Muslim world, it gained a new importance, which Plotinus could never have foreseen.

See also Alcinous; Aristotle; Augustine, St.; Boethius, Anicius Manlius Severinus; Damascius; Gregory of Nazianzus; Gregory of Nyssa; Hellenistic Thought; Homer; Iamblichus; Liber de Causis; Medieval Philosophy; Metaphysics; Numenius of Apamea; Parmenides of Elea; Peripatetics; Philoponus, John; Plato; Plotinus; Porphyry; Proclus; Pseudo-Dionysius; Pythagoras and Pythagoreanism; Simplicius; Stoicism; Thales of Miletus.

Bibliography

For a bibliographical survey of recent studies, see Steel, Carlos, and Christoph Helmig. "Neue Forschungen zum Neuplatonismus (1995–2003)." *Allgemeine Zeitschrift für Philosophie* 29 (2004): 145–162; 225–247.

GENERAL

Armstrong, Arthur H., ed. *The Cambridge History of Later Greek and Early Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1967.

Beierwaltes, Werner. *Platonismus in Christentum*. 2nd ed. Frankfurt am Mainz: Klostermann, 2002.

Cleary, John, ed. *The Perennial Tradition of Neoplatonism*. Leuven: Leuven University Press, 1997.

Dillon, John. *The Golden Chain: Studies in the Development of Platonism and Christianity*. Aldershot, U.K.: Variorum, 1990.

Dillon, John. *The Great Tradition: further Studies in the Development of Platonism and Early Christianity*. Aldershot, U.K.: Variorum, 1997.

Dillon, John, and Lloyd P. Gerson, eds. *Neoplatonic Philosophy. Introductory Readings*. Indianapolis, IN: Hackett, 2004.

Dörrie, Heinrich, ed. *Karl Praechter. Kleine Schriften*. New York: Georg Olms, 1973.

Dörrie, Heinrich, and Mathias Baltes, eds. *Der Platonismus in der Antike*. 6 vols. Stuttgart: Frommann-Holzboog, 1987–2002.

Gersh, Stephen. *Middle Platonism and Neoplatonism. The Latin Tradition*. 2 vols. Notre Dame, IN: University of Notre Dame Press, 1986.

Haase, Wolfgang, ed. *Aufstieg und Niedergang der römischen Welt*. Vol. II/ 36/1–2 and II/36/7. Berlin: de Gruyter, 1987–1994.

Lloyd, A. C. *The Anatomy of Neoplatonism*. Oxford: Clarendon Press, 1990.

O'Meara, Dominic. *Pythagoras Revived, Mathematics and Philosophy in Late Antiquity*. Oxford: Clarendon, 1989.

Merlan, Philip. "Alexandrian School." In *Encyclopedia of Philosophy*. 1st edition. New York: Macmillan, 1967.

Saffrey, Henri Dominique. *Le néoplatonisme après Plotin*. Paris: Vrin, 1992.

Saffrey, Henri Dominique. *Recherches sur le néoplatonisme après Plotin*. Paris: Vrin, 1990.

Smith, Andrew. *Philosophy in Late Antiquity*. London: Routledge, 2004.

Sorabji, Richard, ed. *Aristotle Transformed*. London: Duckworth, 1990.

Sorabji, Richard, ed. *The Philosophy of the Commentators, 200–600 AD: A Sourcebook in Three Volumes*. London: Duckworth, 2004.

Wallis, R.T. *Neoplatonism*. London: Duckworth, 1972.

SPECIAL

On individual authors, see the more than fifty volumes in the series *Ancient Commentators on Aristotle*, edited by R. Sorabji). London: Duckworth. Excellent biographical notes in Goulet, Richard, ed. *Dictionnaire des philosophes antiques*, 4 vols. Paris: Editions du Centre national de la recherche scientifique, 1994– 2003.

Athanassiadi, Polymnia, ed. *Damascius. The Philosophical History*. Athens: Apameia 1993 (on the Athenian school after Proclus).

Armstrong, Arthur H. *Plotin. Enneads*. VII vol. Harvard University Press 1966–1988.

Dodds, E. R. *Proclus. The Elements of Theology*. 2nd ed. Oxford: Clarendon Press, 1963.

Hadot, Ilsetraut. *Le problème du néoplatonisme Alexandrin. Hiéroclès et Simplicius*. Paris: Etudes augustiniennes, 1978.

Hadot, Ilsetraut, ed. *Simplicius, sa vie, son oeuvre, sa survie*. Berlin: de Gruyter, 1987.

Saffrey, Henri Dominique, and Alain-Philippe Segonds, eds. *Marinus. Proclus ou sur le bonheur*. Paris: Les belles lettres, 2001. On the life of Proclus and the school of Athens.

Schibli, Hermann, ed. *Hierocles of Alexandria*. Oxford: Oxford University Press, 2002.

Steel, Carlos. *The Changing Self. A Study on the Soul in later Neoplatonism: Iamblichus, Damascius and Priscianus*. Brussels: Academie, 1978.

Thiel, R. *Simplikios und das Ende der neuplatonischen Schule in Athen*. Stuttgart: Steiner, 1999.

Carlos Steel (2005)

NEOPLATONISM [ADDENDUM]

When Islam took over the Middle East it came into contact with a flourishing local culture heavily influenced by Greek thought. As far as philosophy was concerned, neoplatonism was the leading approach. For example, many of the most important neoplatonists such as Plotinus, Porphyry, and Proclus had studied in Alexandria, a city conquered by the Muslims in 642. A number of key texts became important when translated into Arabic. These were the *Theology of Aristotle*, in fact mainly parts of Plotinus's *Enneads* and the *Liber de causis*, based on Proclus's *Elements of Theology*. Also popular among philosophers were the extensive commentaries on Aristotle by Alexander of Aphrodisias, Themistius, and others—commentators imbued with the values of neoplatonism to some extent. One significant aspect of neoplatonism was the idea that Plato and Aristotle did not differ much on important issues, together with the doctrine of emanation and a cosmology that has the world being produced out of one being or principle. The translation project that transmitted Greek manuscripts into Arabic introduced a good many of these ideas and doctrines into the Islamic world, and so philosophy in the sense of *falsafa* or Peripatetic philosophy became identified in the first few centuries with neoplatonic philosophy.

THE MAIN DOCTRINES

The emphasis on the unity of the creator may well have found a welcoming reception by Muslim thinkers, and it is certainly there strongly in Islamic neoplatonism. One of the central issues is how there came to be many things in existence when really there exists only one absolute being or principle. An explanation is that the One thinks and through thinking brings other things into existence, because once it thinks it realises that it is a thinking thing, and this brings about a mental bifurcation in its unity, a bifurcation that leads to the production of a range of beings that exist either closer or more distantly from it. The more perfect and abstract they are, the closer they are, the less perfect and the more material are more distant.

Another issue was how God related to the world. If God is identified with the One, then the usual account is that he creates the world by emanation, not production. God thinks about himself and through a variety of stages other things are brought into existence, but it would be an interference with God's perfection were he to know about any of these lesser things. The only thing he should think about is himself, and so the world comes about as an indirect effect of this form of thought. An implication of this is that the world is eternal, because God has always existed, and so has always thought about himself. He did not suddenly start thinking, since it is part of his essence to think. Because God is eternal, his thinking must be eternal, and whatever stems from it eternal also. As can be seen, these are all doctrines that do not fit neatly within the framework of a religion such as Islam. The Qurʾan suggests, although does not explicitly state, that God created the world at a particular time, when he wanted to, and it states that he knows everything that goes on in the world. The indirect account of creation as emanation in neoplatonism seems different from the understanding of creation in the Qurʾan.

THE MAIN PHILOSOPHERS

The first Islamic philosopher to construct a thoroughly neoplatonic philosophy was al-Fārābī, and he led the way to Ibn Sīnā (Avicenna), who produced the most developed such theory. They both described emanation as consisting of ten intellects that link the Necessary Being or One with our world, where the active intellect (often

identified with the moon) is the highest level of thought that we can attain. The political implications of the theory are important too. Those who can attain the active intellect are the appropriate rulers, and prophets are those who are able to think at the level of the active intellect, or come into contact with it at least occasionally. This enables them to understand the organization of the world because the active intellect is the most abstract form of thought that human beings can attain, and once it is combined with the facts we observe in the world the prophet can easily predict what is going to happen. For one thing, the organization of the world, according to Ibn Sīnā, is in terms of necessity, so the pattern of existence is something that may be understood rationally by an advanced thinker.

ATTACKS ON NEOPLATONISM

Neoplatonism came under attack by Muḥammad al-Ghazālī, who criticized it in his "Refutation of Philosophy" both for being heretical and also for being invalid philosophically. He picked out in particular the theses that God cannot know individual things, that the world is eternal and that bodily resurrection is inconceivable. The latter follows from neoplatonism due to its prioritization of the soul over the body, and the principle that the material aspects of human beings are not important enough to survive death. The account of immortality in the Qurʾan is clearly material, and the idea that only souls survive death does not seem to fit it. God would not know individual things because he has no sense machinery and he is separated from the everyday activities of this world. Yet as al-Ghazālī argues, how can he punish and reward us on the day of judgement if he has no idea what we do in this world? He rightly points to a range of ideas that really give God little to do, whereas the God of the Qurʾan is directly involved in our everyday affairs.

But these are theological points, and al-Ghazālī also uses the arguments of his opponents to refute them. He tries to disprove the whole neoplatonic apparatus, importing God's will to keep nature in operation as a unified system instead of necessity. He argues in particular that causal necessity is only an idea we have and we could easily think of different connections, or no connections at all, between familiar causes and effects. This really does threaten the whole neoplatonic system, because this involves necessary connections between events, so that when one thing occurs, something else has to occur also. Al-Ghazālī makes a lot of use of imagination here, using thought experiments to try to show that the putative necessary connections are not necessary at all. When Ibn Rushd (Averroes) responded to his attack in his "Refutation of the Refutation" he was fighting with one hand tied behind his back, because Ibn Rushd disapproved of many of the neoplatonic principles as incompatible with the thought of Aristotle, where his main allegiance lay. Ibn Rushd was able to discern many of the divergences between Aristotle and neoplatonism, but in order to defend philosophy as such he was obliged to defend neoplatonism, because this was the main form of philosophy in the Islamic world at that time. Islamic neoplatonism also had a considerable effect on Ismaʿili thought, and on ishrāqī (illuminationist) thought. The esoteric Brethren of Purity (Ikhwān al-Safaʾ) were thoroughly imbued with neoplatonic ideas, although often not very orthodox ones.

DECLINE OF ISLAMIC NEOPLATONISM

Neoplatonism also came under attack by the mystics in Islam who saw its limited access to God as a significant problem. The highest we can come to God is to come into contact with the active intellect, a range of abstract thinking that is really a long way from God. Mystics tend to advocate a much closer connection to God and criticized neoplatonists for their view on this. However, they could use aspects of the theory to explain different levels of reality and their interconnections, although these had to be suitably reinterpreted of course along Sufi lines. Similarly some ishrāqī thinkers replaced the language of the levels of intelligences and worlds with levels of illumination, while at the same time arguing against neoplatonism itself. Neoplatonic philosophy went into a serious decline in the Arab world after the twelfth century, but interest in it continued up to now in the Persian cultural sphere, because its contribution to ishrāqī and Sufi thought was acknowledged and respected.

See also Alexander of Aphrodisias; al-Fārābī; al-Ghazālī, Muhammad; Aristotle; Averroes; Avicenna; Islamic Philosophy; Mysticism, History of; Plato; Plotinus; Porphyry; Proclus; Sufism; Themistius.

Bibliography

Fakhry, Majid. *Al-Fārābī: Founder of Islamic Neoplatonism; His Life, Works and Influence*. Oxford: OneWorld, 2002.

Leaman, Oliver. *Brief Introduction to Islamic Philosophy*. Oxford: Polity, 1999.

Leaman, Oliver. *Introduction to Classical Islamic Philosophy*. Cambridge, U.K.: Cambridge University Press, 2003.

Nanji, Azim. "Ismaʿili Philosophy." In *History of Islamic Philosophy*, edited by S. H. Nasr and O. Leaman. London: Routledge, 1996.

Netton, Ian. *Allāh Transcendent: Studies in the Structure and Semiotics of Islamic Philosophy Theology and Cosmology*. London: Routledge, 1989.

Oliver Leaman (2005)

NEO-PYTHAGOREANSIM

See *Pythagoras and Pythagoreanism*

NEO-SCHOLASTICISM

See *Scotism; Thomism*

NEO-THOMISM

See *Thomism*

NEUMANN, JOHN VON

(1903–1957)

American mathematician, physicist, and economist John von Neumann was born in Budapest, Hungary. He showed an early precocity in mathematics and was privately tutored in the subject; his first paper was written before he was eighteen. He studied at the universities of Berlin, Zürich, and Budapest and received his doctorate in mathematics from Budapest in 1926, almost simultaneously with an undergraduate degree in chemistry from Zürich. After serving as *Privatdozent* at Berlin, he accepted a visiting professorship at Princeton in 1930. Following three years there, he became a professor of mathematics at the Institute for Advanced Study, a position that he held for the rest of his life. In 1955 he was appointed one of the commissioners of the U.S. Atomic Energy Commission, on which he served brilliantly until his death.

Von Neumann made fundamental contributions to mathematics, physics, and economics. Furthermore, these contributions were not disjointed and separate but arise from a common point of view regarding these fields.

Mathematics was always closest to his heart, and it is the field to which he contributed the most. His earliest significant work was in mathematical logic and set theory, topics that occupied him from 1925 to 1929. His accomplishments were of two sorts; they concerned the axiomatics of set theory and David Hilbert's proof theory. In both of these subjects he obtained results of extraordinary importance. He became the first to set up an axiomatic system of set theory that satisfied the two conditions of allowing the development of the theory of the whole series of cardinal numbers and employing axioms that are finite in number and are expressible in the lower calculus of functions. This work contained a full classification of the significance of the axioms with regard to the elimination of the paradoxes. With regard to Hilbert's proof theory, von Neumann clarified the concept of a formal system considerably.

His work on the theory of Hilbert space and operators on that space was probably stimulated by what he had done on rigorous foundations for quantum theory. Essentially, von Neumann demonstrated that the ideas originally introduced by Hilbert are capable of constituting an adequate basis for the physical consideration of quantum theory and that there is no need for the introduction of new mathematical schemes for these physical theories. Von Neumann's papers on these subjects constitute about one-third of his printed work and have stimulated extensive research by other mathematicians.

Von Neumann was one of the founders of the theory of games; since the publication of von Neumann's first paper in 1928 it has become an important combinational theory, applied and developed with continuing vigor. Von Neumann's first paper contains rigorous definitions of the concepts of pure strategy (a complete plan, formulated prior to the contest, that makes all necessary decisions in advance) and of mixed strategy (the use of a chance device to pick the strategy for each contest). The central theorem in this theory, the minimax theorem, was not only enunciated and proved by von Neumann but in his hands became a powerful tool for obtaining new methods for combinatorial problems.

A decade after this fundamental paper was written, von Neumann began a collaboration with Oskar Morgenstern that led to *The Theory of Games and Economic Behavior,* a book that has decisively affected the entire subject of operations research.

Von Neumann's principal interest in his later years was in the possibilities and theory of the computing machine. He not only conceived the concept of the so-called stored program computer in 1944 but he made three other signal contributions. First, he recognized the importance of computing machines for mathematics, physics, economics, and industrial and military problems; second, he translated this insight into active sponsorship of a machine (it was called Johniac by his collaborators) that served as a model for several impor-

tant computers; third, he was one of the authors of a series of papers that provided a theoretical basis for the logical organization and functioning of computers. These papers set out the complete notion of the flow diagram and contained the genesis of many programming techniques.

See also Computing Machines; Decision Theory; Game Theory; Hilbert, David; Mathematics, Foundations of; Proof Theory; Quantum Mechanics; Set Theory.

Bibliography

Much of von Neumann's work, with the exception of certain still-classified reports or papers that are essentially duplicates of other works, is published in *John von Neumann, Collected Works,* 6 vols. (New York: Macmillan, 1961–1963). Other important books by von Neumann are *Continuous Geometry,* 2 vols. (Princeton, NJ: Institute for Advance Study, 1936–1937); *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press, 1955); *The Computer and the Brain* (New Haven, CT: Yale University Press, 1958). For a survey of von Neumann's life and work, see the memorial volume *John von Neumann, 1903–1957, Bulletin of the American Mathematical Society* 64 (3:2) (May 1958).

Herman H. Goldstine (1967)

NEURATH, OTTO

(1882–1945)

Otto Neurath, an Austrian sociologist and philosopher, was one of the originators of logical empiricism and an independent Marxist socialist. A man of great vitality, intelligence, and good humor, Neurath was a polymath and an energetic organizer of academic, educational, and economic affairs. His major work was in sociology, economic and social planning, scientific method, and visual education, this last especially by means of an international language of simplified pictures ("isotypes"), but he was also interested in the history of science, political and moral theory, economic history, and statistical theory and was engaged in recurrent efforts to create a new encyclopedism.

ECONOMIC AND COMPARATIVE HISTORY

Neurath's first article, published in 1904, was "Geldzins im Altertum" (Commercial interest in antiquity), and in 1909 he published a popular history of the economic systems of classical Greece and Rome, *Antike Wirtschaftsgeschichte* (Leipzig, 1909), which he supplemented by shorter studies of ancient economic thought. His historical interests then turned to physical science. A little-known paper of 1915, "Prinzipielles zur Geschichte der Optik," compared the ideas on optics of Isaac Newton, René Descartes, Nicolas Malebranche, Francesco Maria Grimaldi, Christian Huygens, Thomas Young, Augustin-Jean Fresnel, Jean-Baptiste Biot, and Étienne-Louis Malus with respect to their conceptual images of periodicity, polarization interference, and Huygens's principle of continuity of centers of force.

Neurath generalized the logic of this analysis to compare systems of hypotheses by a procedure that selects basic notions to be calculated and then enumerates all theories that may be constructed from permutations of these notions. The simple view that theories of light may be divided into wave theories and corpuscular theories is replaced by a more accurate, complex, and systematically clear historical development. To Neurath this use of basic explanatory notions, which are sometimes images and sometimes abstractions, illustrated the value of philosophical understanding for the historian of natural and social science. Neurath's own philosophical understanding anticipated later reliance on alternative sets of epistemologically basic sentences in the structural elucidation of scientific theories.

In 1916, Neurath wrote a general paper on classification, "Zur Klassifikation von Hypothesensystemen," and elaborated on this topic in his monographs *Empirische Soziologie* (1931) and *Foundations of the Social Sciences* (1944). Classification by hypotheses seemed to Neurath to be a principal method for comparative studies of theories and explanations and a crucial tool for rational understanding of cross-cultural phenomena.

ECONOMIC PLANNING, WAR, AND SOCIALISM

During 1919, Neurath served in the Central Planning Office of the Social Democratic government of Bavaria and of its successor, the short-lived Bavarian Soviet Republic. Although he was a civil servant and not a party man, he was imprisoned when the Communist regime was overthrown; upon his release in 1920 he went to Vienna. He there took up again an earlier career as a publicist for socialist economics by efforts on behalf of a socialist conception of civic education, moral and religious reform, and individual responsibility. With Josef Popper-Lynkeus, Neurath was one of the first socialists to call for a centrally planned, rational economy based on Marxist concepts but deriving its policy recommendations from welfare goals and a statistical analysis of the

production and distribution of goods and of standards of living.

Less clear to Neurath than equitable distribution of wealth was how a community spirit could be developed while the workers themselves were still overwhelmed by the established culture and the habits of the competitive capitalist order. Nevertheless, he fused his hypotheses about social-economic planning with a moral optimism about the acceptance by the workers of enlightened and rational attitudes toward all life's problems. Neurath's *Lebensgestaltung und Klassenkampf* poignantly tried to teach the reader about a transformed way of life in which he could realistically experience something of the peaceful and cooperative future at least in his private life, and at the same time come to sober realization of the obstacles placed by exploitative society in the way of a rich inner life and good personal relations as well as in the way of the transformation of society by rational socioeconomic planning.

EMPIRICAL SOCIOLOGY

In the 1930s and early 1940s, between the publication of the two monographs on sociology (the *Empirische Soziologie* and the *Foundations*), Neurath published several smaller papers on sociological topics. The most important were "Soziologie im Physikalismus," a physicalist restatement of sociological theories and problems, "Soziologische Prognosen," on social-historical predictions, and "Inventory of the Standard of Living," on the problem of making a rational calculation of the standard of living.

To make sociology scientific, Neurath urged the use of a physicalist language in which all the possible empirical statements would be descriptive of space-time things and properties; this was, roughly, a demand for behaviorism in social theory. He believed that this social behaviorism carried out Karl Marx's claim that historical materialism was empirical, starting from the factual situation of real men in objective circumstances and basing theories upon hypotheses which are free of wishful or evaluative assumptions. Human beings, streets, religious books, prisons, gestures can be so described, and they may be grouped in accord with physicalist theoretical systems. Happiness and suffering, too, may be described empirically, even in a manner similar to a mechanical description of space-time entities. But man, in some situations, dominates the lawlike mechanism of the natural environment. In Neurath's typical formulation: Formerly when there was a swamp and man, man disappeared; nowadays the swamp disappears.

But the language of mechanism is laden with myth and metaphysical presuppositions, and Neurath tried to eliminate all impure or careless terminology. Just as he would ban metaphysics as a misuse of unverifiable but grammatically correct word-signs, so he wished to forbid social theorists to use words that carry multiple meanings and assumptions; he himself never used the word *capital*. Sociological descriptions demand arguments over the entire range of environmental and causal science; biological, geological, ethnological, and chemical statements must join social, psychological, economic, legal, and other statements of purely human reference. Hence it would be useful to invent an empirical language suited to all the sciences, one that avoids descriptive distinctions that are the result of mere linguistic convention. Neurath hoped that empirical sociology might be formulated with clear and univocal physicalist predicates. However, we start with inexact "clots," with indistinct and unanalyzed evidence, and we must tolerate and even carefully devise a correspondingly rich vocabulary which is also amenable to analysis of regularities and at times to the creation of a calculus.

Neurath often wrote of an essential uncertainty in all scientific description and predication, of the probabilistic nature of learning from experience. Historians should explain the present from knowledge of portions of the past, but to predict the future with precision is beyond us. There are too many variables; at least some of these are unknown, and the greater the anticipated change, the less our scientific assurance about its realization. We may, in Neurath's view, strive to construct a future state of affairs, but whether we feel hesitant or confident, we have in sociological lawlike historical statements no rational ground for predictions that are certain. Moreover, some predictive statements, notably self-fulfilling or hortatory prophecies, are codeterminant; they carry causal weight which disturbs their subject matter. Other predictions seem impossible on their face. How should a nation that could not invent the wheel predict the invention of the wheel? Others are too complex. Will painters in misty regions paint misty pictures or, just because of the mistiness, sunny ones? Neurath carried out this analysis of pseudorational certainty throughout his work, using it with a moral force. Decisions cannot be replaced by calculation or by reasoning—not in practical life, and not in scientific work.

SCIENTIFIC METHOD

Physicalism was developed mainly by Neurath and Rudolf Carnap. It may be seen as Neurath's attempt to

express, in epistemological terms, the materialist (objective) foundation of knowledge, since the persistent recognition he gave to the natural fact of socially intersubjective agreement was a principal source of his antiphenomenalist role within the Vienna circle. Despite Neurath's insistence on a sharp distinction between scientific and metaphysical expressions by means of criteria for empirical meaning, it was his view that intersubjective agreement provides approximate unanimity about the grounds for judgments, not for meanings. By use of a physicalist language, skeptical inquirers display and share a common standard for confirmation. Physicalism had the further merit for Neurath that it was a linguistic doctrine which overcame any systematic mutual incomprehension of special disciplines not by reduction to the special discipline of physics but by a doctrine of reference to the generalized physics of public space-time states (in the human macroscale).

Neurath freely admitted that this doctrine was a hypothesis; the world was assumed to be unified, a causal network whose multiplicity of descriptions should tend toward a unified language that includes the social, biological, and physical sciences. Moreover, as an analysis of the process of scientific knowledge, physicalism programmatically explicated (for any special science) the relations among the physiology and social psychology of sensuous perception, the physics of experimental and measurement technology, and the known scientific or commonsense entities. Neurath saw physicalism as the further hypothesis that the world is knowable in principle everywhere and throughout. Finally, in "Protokollsätze" (1932) Neurath represented physicalism as providing a sophisticated revision of the doctrine of atomic bits of knowledge, conveyed by individual reports, or "basic sentences," also known at the time as "protocol sentences," by demanding that they, too, be intersubjective and, however psychologically certain, logically tentative and empirically testable. Indeed, the truth of protocol sentences was attributable to their cohering role in a theory (or system of theories) to which empirical evidence gave confirmatory evidence, and consequently the possibility existed that a conflict between a particular protocol statement and a theoretical statement of more complex form and function might, by choice and for convenience, be resolved by discarding the protocol. Neurath found his early analysis of alternative hypothesis systems and their fact-fitting auxiliary statements borne out within this empirical conventionalist interpretation of the physicalist basis.

VISUAL EDUCATION

Both the union of scholars and ordinary workers and the overcoming of national and linguistic divisions were in Neurath's mind when he began to develop his "Vienna method" of visual education. In rudiment, he used an invariant and self-explanatory pictorial sign for a given thing, so as to give quick information, unencumbered by irrelevancies and easily remembered. Neurath's maxims were simple: He who knows what best to omit is the best teacher; to remember simplified pictures is better than to forget accurate figures.

UNITY OF SCIENCE AND ENCYCLOPEDISM

Neurath was the principal organizer of several related philosophical enterprises. By 1929 the regular but informal Thursday meetings of philosophers and scientists who met for discussion with Moritz Schlick in Vienna had gathered sufficient force to produce a noted manifesto of a scientific world conception, signed by Neurath, Hans Hahn, and Carnap although it was largely Neurath's work. This led in the same year to the first of a series of international congresses for scientific philosophy. Neurath's stress upon the unification of the sciences by means of a unifying language, unity of method, and interdisciplinary dialogue led him to plan the *International Encyclopedia of Unified Science*, edited by himself, Carnap, and Charles Morris as the principal effort of the new Institute for the Unity of Science (founded in the Hague in 1936 and later removed to Boston, Massachusetts), directed chiefly by Philipp Frank. The first two introductory volumes appeared in parts, but even these were still incomplete nearly two decades after Neurath's death. Only the Institute for Visual Education (Isotype) continued with vigor after 1945, directed by Neurath's colleague and third wife, Marie Reidemeister Neurath.

See also Basic Statements; Behaviorism; Carnap, Rudolf; Descartes, René; Historical Materialism; Logical Positivism; Malebranche, Nicolas; Marx, Karl; Marxist Philosophy; Newton, Isaac; Physicalism; Popper-Lynkeus, Josef; Schlick, Moritz.

Bibliography

WORKS BY NEURATH

"Geldzins im Altertum." *Plutus* 29 (1904): 569–573.

"Prinzipielles zur Geschichte der Optik." *Archiv für die Geschichte der Naturwissenschaft und die Technik* 5 (1915): 371–389.

"Zur Klassifikation von Hypothesensystemen." *Jahrbuch des Philosophischen Gesellschafts an der Universität Wien*, 27ff. Liepzig, 1916.
Durch die Kriegswirtschaft zur Naturalwirtschaft. Munich: Callwey, 1919.
Lebensgestaltung und Klassenkampf. Berlin: Laub, 1928.
Wissenschaftliche Weltauffassung: Der Wiener Kreis. Vienna: Arthur Wolf, 1929. Written jointly by Neurath, Carnap, and Hans Hahn.
"Wege der wissenschaftlichen Weltauffassung." *Erkenntnis* 1 (1930): 106–125.
Empirische Soziologie: Der wissenschaftliche Gehalt der Geschichte und Nationalökonomie. Vienna: J. Springer, 1931.
"Soziologie im Physikalismus." *Erkenntnis* 2 (1931): 393–431.
"Protokollsätze." *Erkenntnis* 3 (1932): 204–214. Translated into English by Frederick Schick (who is credited as "George") in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.
"Soziologische Prognosen." *Erkenntnis* 6 (5/6) (1936): 398–405.
"Inventory of the Standard of Living." *Zeitschrift für Sozialforschung* 6 (1937): 140–151.
Foundations of the Social Sciences. Chicago: University of Chicago Press, 1944. This is Vol. 2, No. 1, of the *International Encyclopedia of Unified Science*.
Philosophical Papers, 1913–1946. Edited by Marie Neurath and R. S. Cohen. Dordrecht, Holland, and Boston: D. Riedel, 1983. With a bibliography.

OTHER RECOMMENDED WORKS

Cartwright, Nancy. *Otto Neurath: Philosophy between Science and Politics*. New York:; Cambridge University Press, 1996.
Hofmann-Grüneberg, Frank. *Radikal-empiristische Wahrheitstheorie: Eine Studie über Otto Neurath, den Wiener Kreis und das Wahrheitsproblem*. Wien: Hölder-Pichler-Tempsky, 1988.
International Encyclopedia of Unified Science Vol. II, no. 1. Chicago, University of Chicago, 1970.
Koppelberg, Dirk. *Die Aufhebung der analytischen Philosophie: Quine als Synthese von Carnap und Neurath*. Frankfurt am Main: Suhrkamp, 1987.
Logical Empiricism at Its Peak: Schlick, Carnap, and Neurath. Edited by Sahotra Sarkar. New York: Garland, 1996.
Morris, Charles W. *Otto Neurath and the Unity of Science Movement: A Collection of Material Commemorating Otto Neurath's Place in the Unity of Science Movement*. Jerusalem, 1966.
Nemeth, Elisabeth. *Otto Neurath und der Wiener Kreis: Revolutionäre Wissenschaftlichkeit als politischer Anspruch*. Frankfurt am Main; New York: Campus, 1981.
Nemeth, Elisabeth, and Friedrich Stadler. *Encyclopedia and Utopia: The Life and Work of Otto Neurath (1882–1945)*. Dordrecht; Boston: Kluwer, 1996.
Neurath, Otto. *Empiricism and Sociology*. Edited by Marie Neurath and R. S. Cohen. Boston: Reidel, 1973.
Neurath, Otto. *Gesammelte philosophische und methodologische Schriften*. Wien: Hölder-Pichler-Tempsky, 1981.
Neurath, Otto. *Philosophical Papers, 1913–1946*. Edited by Marie Neurath and R. S. Cohen. Dordrecht, Holland; Boston: D. Riedel; Hingham, MA: Kluwer, 1983.
Neurath, Otto, ed. *Unified Science: The Vienna Circle Monograph Series Originally Edited by Otto Neurath, Now in an English Edition*. Edited by Brian McGuinness. Dordrecht, Holland; Boston: D. Reidel; Norwell, MA: Kluwer Academic, 1987.
Neurath, Paul, and Elisabeth Nemeth. *Otto Neurath, oder, Die Einheit von Wissenschaft und Gesellschaft*. Wien: Böhlau, 1994.
Neurath, Gramsci, Williams: Theorien der Arbeiterkultur und ihre Wirkung. Hamburg: Argument-Verlag, 1993.
Schlick und Neurath—Ein Symposion. Wien, Amsterdam: Rodopi, 1982.
Uebel, Thomas E. *Overcoming Logical Positivism from Within: The Emergence of Neurath's Naturalism in the Vienna Circle's Protocol Sentence Debate*. Amsterdam; Atlanta: Rodopi, 1992.
Uebel, Thomas E., ed. *Rediscovering the Forgotten Vienna Circle: Austrian Studies on Otto Neurath and the Vienna Circle*. Dordrecht; Boston: Kluwer, 1991.
Uebel, Thomas E. *Vernunftkritik und Wissenschaft: Otto Neurath und der ersten Wiener Kreis*. Wien: Springer, 2000
Zolo, Danilo. *Reflexive Epistemology: The Philosophical Legacy of Otto Neurath*. Dordrecht; Boston: Kluwer, 1989.

Michael J. Farmer (2005)
Robert S. Cohen (1967)

NEUROSCIENCE

Neuroscience is the scientific study of nervous tissue, activity, organization, systems, and interactions. It is paradigmatically interdisciplinary, currently including biophysics, organic and biochemistry, molecular through evolutionary biology, anatomy and physiology, ethology, neuropsychology, and the cognitive and information sciences. Investigators include basic scientists and clinicians. During the late twentieth century, neuroscience underwent enormous growth. Quantitative information available on the Society for Neuroscience's Web site speaks to this. Beginning in 1970 with 500 members, at last count (summer 2004) the Society boasts more than 34,000 members worldwide. More than 30,000 registrants attended the 2004 annual meeting, where more than 14,000 posters and oral presentations were delivered. There are now more than 300 graduate training programs worldwide in neuroscience. With its increasing academic influence and its obvious connection with philosophy's perennial mind-body problem, it was inevitable that philosophers would begin taking serious interest.

Academic philosophy's systematic interest might be dated to 1986, the year that Patricia Churchland's *Neurophilosophy* appeared. She boldly proclaimed that "nothing is more obvious than that philosophers of mind could profit from knowing at least something of what there is to know about how the brain works" (p. 4). Her book presented what was then textbook neuroscience, contex-

tualized by developments in postlogical empiricist philosophy of science. It set the stage for much neurophilosophy and philosophy of neuroscience that followed, especially the branch of neuroscience that philosophers attended to (cognitive neuroscience). This entry will present some neuroscientific techniques and results that have attracted philosophers' attention. In the interest of pedagogy, the emphasis will be on the scientific details. It will close with a section describing another field of contemporary neuroscience that unfortunately has captured less philosophical attention, followed by a more detailed discussion of implications for mind-brain reductionism. Space limitations preclude a comprehensive survey and the bibliography is limited, both in number of entries and primarily to textbook sources and review articles (all containing extensive references to the primary scientific literature, however). This is befitting an encyclopedia entry, but philosophers who are interested in acquiring a serious understanding of actual neuroscience are urged not to stop with these sources. There is no shortcut around delving into the primary literature. Superficial neuroscience still serves too often in straw arguments in the philosophy of mind.

Ideally this entry would also include work on pain processing, especially on the two types of pain circuits (rapidly conducting Ad and slowly conducting C fibers) and the different pain qualities carried by each; the neural mechanisms of dream sleep, especially endogenously produced activity in sensory regions; the discovery of mirror neurons in primate brains that are active when the subject performs a specific motor task and when the subject observes a cohort performing that task; the sea change in computational neuroscience during the 1990s, away from abstract network modeling (inspired by early successes of "connectionist" artificial intelligence) and toward compartmental modeling, where the patch of neural membrane and its ion-specific conductance capacities become the basic units of analysis; and the neurobiology and behavioral genetics of schizophrenia (as elaborated in numerous publications by Kenneth Schaffner). Philosophers have argued for implications from each. But choices were necessary.

FUNCTIONAL NEUROIMAGING

Functional neuroimaging provides a window into the active, healthy brain. Results from two imaging techniques have dominated philosophers' attention: positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). PET is based on radioactive decay of positrons (positively charged electrons). Subjects are injected with water (or sugars) labeled with a radioactive, positron-emitting isotope (such as oxygen-15, whose nuclei are manufactured to contain the normal eight protons but only seven neutrons). During the minute following injection, radioactive water accumulates in biological tissues in amounts directly proportional to local blood flow. Positrons leave the nuclei of the unstable, radioactive atoms and travel only a short distance through biological tissue (at most a few millimeters).

After losing their kinetic energy positrons are attracted to negatively charged electrons. This collision annihilates both and the resulting energy manifests in two photons traveling 180o away from the annihilation site. These photons exit the tissue being imaged and are detected by radiation detectors arranged in coincidence circuits (the "PET camera"). Photons arriving simultaneously at opposing detectors are counted and these counts are converted into an image that reflects the relative number of annihilation collisions localized to a given region. A single ring of coincident detectors can only image a single "slice" through the tissue; but modern PET cameras contain multiple rings and so can image multiple parallel "slices" simultaneously. Powerful algorithms and computer graphics can reconstruct the functional images in any desired orientation. Color codes are typically used to denote intensity of activity.

By subtracting images generated during a carefully selected control task from those generated during an experimental task, PET generates a picture of the location and intensity of activity specific to performing the experimental task. These are the colorful images published in PET studies. But what PET measures directly is localized blood flow to a small region of biological tissue. The activity interpretation exploits the known (and independently verified) positive correlation between increased local blood flow and increased cellular activity in that region.

fMRI—more precisely, Blood Oxygenation Level Dependent (BOLD-) fMRI—also exploits the established correlation between localized blood flow changes and cellular activity in tiny neural regions. But to measure these changes, it takes advantage of the different properties of oxygen-bearing and deoxygenated hemoglobin in a strong magnetic field. Oxygenated hemoglobin is more prevalent in the bloodstream in regions of high cellular activity. The metabolic demands of highly active neurons and glial cells generate signals to blood vessels to increase blood flow to the region (the "hemodynamic response"). The resulting supply exceeds the cells' capacity to remove oxygen from hemoglobin. As of 2004, these different

magnetic properties can be measured and localized in fMRI scanners approved for human use to less than one millimeter. Stronger magnetic fields generate more precise measurements and localizations. Algorithms and graphics capabilities comparable to PET technology reconstruct “slices” through the imaged tissue at any desired orientation. By normalizing and contrasting BOLD signals across experimental and carefully selected control tasks, experimenters can image activity location and intensity specific to the experimental task. A variety of postprocessing techniques are employed to account for the potentially variable hemodynamic delays between neural activity generated by task performance and increased blood flow.

A handful of functional neuroimaging studies (mostly older ones from the early days of PET!) recur in philosophical discussions. (All of the studies discussed below are also discussed and referenced in Michael Posner and Marcus Raichle’s popular book, *Images of Mind*, 1997.) One still sees reference to Per Roland and his colleagues’ regional cerebral blood flow studies from the mid-1980s. Their subjects performed a number of cognitive tasks, including verbalizations, arithmetical calculations, and a complicated memory imagery task involving walking familiar streets and making a system of turns while reporting landmarks visualized along the way. The memory imaging task produced increased blood flow bilaterally to regions in the parietal and temporal lobes—regions that lesion data from human neurological patients had previously revealed to be involved in mental imagery. Stephen Kosslyn’s work on mental imagery using neuroimaging techniques, especially work reported in his *Image and Brain* (1994), is also discussed often by philosophers in debates about the structure of cognitive representations. Much of Kosslyn’s work demonstrates that the same neural regions are activated when subjects form a visual mental image and when they visually perceive a similar stimulus. He has demonstrated these effects as far back in the visual processing pathways as primary visual cortex (V1). They hold for locations containing neurons known to specialize for the size of perceived stimuli and for stimuli viewed from typical or atypical perspectives.

Much philosophical attention on functional neuroimaging focuses on its implications for localization hypotheses of cognitive functions. Steve Petersen and his colleagues’ studies on language processing and use from the late 1980s are still cited and discussed. They employed PET and a hierarchical experimental design that enabled them to separate activations generated by passively viewing words, passively listening to words, speaking words viewed or heard, and generating semantically related words to those viewed or heard. Different tasks in this hierarchy produced PET activation increases in different neural regions, suggesting to some the localization of different tasks involved in language processing, including word perception, speech production, and semantic access. Localization arguments and their scientific grounding in functional neuroimaging studies have been challenged, notably by William Uttal in *The New Phrenology* (2001).

A handful of functional neuroimaging studies on attention rose to philosophical prominence with growing interest in consciousness. A popular example uses the Stroop task to induce conflict. Color words are presented visually in either compatible or incompatible print colors (e.g., compatible: “red” printed in red; incompatible: “red” printed in green). Subjects are asked to name the color of the print. Behaviorally, as measured by errors and response time, subjects find incompatible conditions much harder. Some psychologists have argued that incompatible conditions require conscious effort to inhibit saying the color word. José Pardo and his colleagues in the early 1990s found strong activation effects specific to the (forebrain) anterior cingulate gyrus when compatible PET activation results were subtracted from incompatible ones. These results are consistent with behavioral data from patients with anterior cingulate lesions and lend empirical support to earlier speculations about the neural components of an executive atttentional control network.

CLINICAL NEUROPSYCHOLOGY AND NEUROLOGY

Philosophers have long taken interest in the behavioral effects of brain damage and disease. (Bryan Kolb and Ian Whishaw’s *Fundamentals of Human Neuropsychology*, 2003, is an excellent textbook that includes discussions of topics covered in this section and extensive references to the primary scientific literature.) Commissurotomy (“split brain” surgery) is one contemporary example. To treat otherwise intractable epilepsy, neurosurgeons in the early 1960s revived a surgical technique of cutting a patient’s corpus callosum. The corpus callosum is a huge bundle of axon fibers that connect homologous regions of the left and right cortical hemispheres.

The procedure was clinically successful with a minimum of apparent behavioral effects, until Roger Sperry and his collaborators (Michael Gazzaniga, Joseph Bogen) applied more sophisticated tests. They discovered that

these patients had lost the capacity of their two cerebral hemispheres to communicate directly with each other. Owing to the segregation and crossing of axon projections from sensory receptor organs to relay neurons in the thalamus and sensory cortex, experimenters could direct, for example, different visual stimuli to the left and right cortical hemispheres. If one then asked the subject to pick up an object related to the visual display with his or her left hand, the subject would pick up an object related to the visual display in his or her right hemisphere. (As with sensation, the motor system also crosses over: Right motor cortex controls left side movement and vice versa.) If one then asked that subject to explain verbally why he or she was holding that object (and the subject was among the roughly 85 percent of humans with speech localized to the left hemisphere), the subject indicated no awareness in his or her verbal response of the display presented to the right visual hemisphere and instead confabulated a verbal account that related the chosen object to the left hemisphere's visual display. The variety and number of similar results led to speculations about two seats of conscious awareness and control in a single human brain, and subsequent philosophical reflections about the unity of self (or lack thereof).

Blindsight refers to preserved visual capacities following damage to visual cortex. Such damage produces a scotoma (a "blind spot") at circumscribed locations in the patient's visual field. Despite no conscious awareness of visual stimuli presented there, these patients nevertheless display some impressive visual abilities when prompted to guess about stimuli presented in their scotoma, including pointing accurately to visual stimulus location, detecting movement, and discriminating shapes (and in a few cases, colors). Their performances far exceed chance. As reviewed in Lawrence Weiskrantz's *Consciousness Lost and Found* (1998), experimental work over the past three decades has mostly confirmed early results and has introduced controls to address methodological criticisms of the early studies. Blindsight has figured into philosophical discussions of the nature of visual consciousness and the location of its neurobiological mechanisms, as well as epistemological discussions about accurate perceptual judgments and the purported necessity of awareness.

Denial symptoms are the opposite of blindsight. Blindness denial (Anton's syndrome) can result from cortically induced blindness and renders patients functionally blind by all objective tests and measures; yet these patients vehemently claim that they can see. Paralysis denial can result from damage to motor cortex and renders patients functionally paralyzed on the side of their bodies opposite the damage; yet these patients vehemently deny that they are paralyzed. Many patients generate spontaneous confabulations (e.g., "it is dark in this room," "I have bad arthritis in my left shoulder—it hurts to move my left arm") to explain their failures on simple behavioral measures. Numerous controls are standard in neurological assessment to rule out cases of confusion or persistent stubbornness to accept or admit the deficit. Some philosophers and neurologists have argued from these clinical details toward revisions of our commonsense conceptions of awareness, conscious control, and the initiation of behavior. Vilayanur Ramachandran and Susan Blakeslee's popular book, *Phantoms of the Brain* (1998), is a good example, with elaborate discussions of clinical cases and a good bibliography to primary sources.

Contralateral neglect ("hemineglect") is a condition whereby patients ignore the side of their body and the world opposite the side of damage to parietal cortex. (Typically the damage is to right hemisphere, producing left side neglect.) The neglect invades all sensory modalities, is sometimes accompanied by denial and confabulation (to the point of patients denying that their neglected limbs even belong to them), and even invades memories and images. A famous study from the late 1970s by neurologist Edoardo Bisiach and his colleagues asked recent stroke patients demonstrating neglect symptoms to remember a famous square in Milan from one vantage point and to describe all objects they remembered. They were then asked to visualize the square from the opposite vantage point and describe the objects remembered. In both cases, they described objects only on their nonneglected sides—meaning that they described a different set of objects from the separate vantage points. Hemineglect appears to be an awareness deficit. If the only available objects for patients to attend are on the neglected side, they can attend to them. But when objects are present on the nonneglected side, they seem to lose all awareness of the opposite space. Philosophers working on consciousness, awareness, their brain mechanisms, and on body awareness and body-in-space representations have appealed to neglect data.

THE BINDING PROBLEM

Conscious experiences are present to us as unified wholes. Visual object perception provides rich examples. In ordinary circumstances I see a football zooming toward me, not separately brown color, oblong shape, in motion (speed, trajectory) toward me. Yet each of these visual qualities is extracted by neuronal activity in spa-

tially separated areas. Separate neural pathways respond to qualities that characterize a perceived object's identity (the ventral or "what" stream through inferior temporal cortex) and its location, motion, and my actions toward it (the dorsal or "where/how" stream through posterior parietal cortex). Neurons specialized for specific aspects of the visual stimulus are at distinct locations within each pathway. Seeing an object requires neuronal activity in spatially separated regions and there is no evidence for "grandmother" neurons further downstream onto which all of these active neurons project. This is the "binding problem." How is activity in these spatially separated regions bound together to become active as a unit and so produce a unified visual percept? And given that an object seen is often also heard, felt, or smelled simultaneously, and that these multimodal perceptual experiences are also unified in conscious experience, we actually confront a set of binding problems. (Neuropsychologist Ann Treisman's 1996 review article is an excellent introduction.)

Throughout the 1990s a variety of "temporal synchronicity" solutions were popular. These held that binding results from induced synchronous activity in specific neurons in the separate pathways and processing areas. The discovery of a robust "40 Hz oscillation pattern" across the mammalian cortex during wakeful attention and rapid eye movement (REM, "dreaming") sleep inspired this approach. Feedforward and reciprocal feedback anatomical projections between sensory modality-specific and nonspecific neuron clusters ("nuclei") in the thalamus and sensory cortex provided a biologically plausible hypothesis for how temporal synchronicity might be induced.

However, problems quickly surfaced. It is notoriously difficult to determine the "binding window," the time interval during which the spatially separated processing must occur. Are mechanisms sensitive to temporally coherent discharges tied to the full length of activated neuronal discharges, making the binding window up to several hundred milliseconds? If so, then because distinct and changing stimuli clutter the visual field continuously over this long an interval, how do we successfully bind together the right combination of features? Is activity onset or rise time of discharge the relevant temporal feature? If so, this leads to difficulties when we consider the variable latencies of activity in different areas of modality-specific sensory pathways. Latency differences exist all the way back to activity in sensory receptor cells: Hair cells at different locations on the cochlea and photoreceptors at different locations on the retina respond at slightly different times to a single auditory or visual stimulus. Moving up both auditory and visual processing streams, the temporal differences at which information about different aspects of a single stimulus reaches later points can be tens of milliseconds. Somehow, a temporal synchronicity binding mechanism must compute these processing time differences. (The problem of latency differences is exacerbated when we consider multimodal—for example, visual-auditory binding mechanisms.)

These biological details suggest the need for neural regions where temporal information converges (to carry out the latency computations); but now temporal synchronicity solutions confront a similar problem to the one that sunk purely spatial solutions—no solid evidence for such convergence sites. Temporal synchronicity solutions are less popular now. But the binding problem continues to attract philosophers' attention due to its obvious connections with consciousness and brain mechanisms. Rodolfo Llinás and Patricia Churchland's *The Mind-Brain Continuum* (1996) is a good edited volume that was published at the time that these debates about binding and temporal synchronicity were raging.

MOLECULAR AND CELLULAR COGNITION

The reader might have noticed that most examples of neuroscientific work that has attracted philosophers' attention are dated. This is not necessarily a bad thing. Philosophical reflection on scientific results depends on their scientific credibility and that takes time to establish. However, this limitation risks missing important new developments and changing foundational assumptions in a rapidly developing science. The lessons philosophers draw might then be dated as well. There is evidence that "foundational" change has occurred recently in neuroscience, having to do with the increasing impact of molecular biology.

More than a decade ago neurobiologists Eric Kandel, James Schwartz, and Thomas Jessell, in the third edition of their textbook, *Principles of Neural Science* (1991), wrote that "the goal of neural science is to understand the mind: how we perceive, move, think, and remember. In the previous editions of this book, we stressed that important aspects of behavior could be explained at the level of individual nerve cells. ... Now it is possible to address these questions directly on the molecular level" (p. xii). With the publication of the text's fourth edition (2000), and after another decade of cellular and molecular investigations, these same authors announce mind-to-molecules "linkages" as accomplished scientific results:

> This book … describes how neural science is attempting to link molecules to mind—how proteins responsible for the activities of individual nerve cells are related to the complexity of neural processes. Today it is possible to link the molecular dynamics of individual nerve cells to representations of perceptual and motor acts in the brain and to relate these internal mechanisms to observable behavior. (p. 3–4)

These are heady claims, backed up by more than 1,400 pages of textbook evidence drawn from a huge scientific literature. Yet to read much philosophical discussion of neuroscience, one would not even know that this work and attitude exists—much less that it constitutes the current mainstream of the discipline. (This mountain of supporting evidence also refutes the pitying lament so often uttered by philosophers and cognitive scientists: "If we only knew more about how the brain works …" We do.)

Much of this research is congealing around a field dubbed "molecular and cellular cognition." According to the Molecular and Cellular Cognition Society's Web site, the field's stated goal is to discover "explanations of cognitive processes that integrate molecular, cellular, and behavioral mechanisms, literally bridging genes and cognition." The field emerged in the early 1990s, after gene engineering techniques were introduced into mammalian neurobiology to generate knockout and transgenic rodents for behavioral studies. Memory has been a principal research focus, with an emphasis on consolidation (the transformation of labile, easily disrupted short-term memories into stable, enduring long-term forms) and on hippocampus-based memories that neuropsychologists call "declarative" or "explicit." This field's methodology is ruthlessly reductive. Its basic experimental strategy is to intervene into cellular or intracellular molecular pathways and then track their effects in the behaving animal using standard tests borrowed from experimental psychology for the phenomenon under investigation. (So despite the new molecular-genetic techniques for intervening directly at increasingly lower levels of biological processes, the basic experimental logic remains interestingly similar to that of classical lesioning and pharmacological studies.)

At last count, more than sixty molecules have been implicated in the molecular mechanisms of mammalian long-term potentiation (LTP), an activity-dependent form of synaptic plasticity with memorylike features. However, a few figure prominently and have been targets of bioengineered mutations and subsequent behavioral study in declarative memory consolidation tasks. Cyclic adenosine monophosphate (cAMP) is a product of adenosine triphosphate (ATP) conversion into energy to drive cellular metabolism and activity. cAMP is the classic "second messenger" of molecular biology, functioning as an intracellular signal for effects elsewhere in the cell. When available in high quantities in active neurons it binds to the regulatory subunits of protein kinase A (PKA) molecules, freeing the catalytic PKA subunits. In high enough quantities, the latter translocate back to the neuron's nucleus, where they phosphorylate cAMP response element binding proteins (CREB), a family of gene transcriptional enhancers and repressions that turn on or inhibit new gene expression and protein synthesis.

Specific targets of phosphorylated CREB transcriptional enhancers include genes coding for regulatory proteins that keep PKA molecules in their active state and effector proteins that resculpt the structure of active synapses, keeping those synapses potentiated to presynaptic activity for days to weeks. Numerous features of LTP have made it an attractive theoretical mechanism for memory consolidation for years; results from molecular and cellular cognition have finally lent experimental backing to this decades-old speculation.

Alcino Silva's group has used mice with a targeted mutation of the CREB gene on a variety of short- and long-term memory tasks, including the Morris water maze task, a combined environment-conditioned stimulus fear conditioning task, and a social recognition memory task. These mice do not synthesize the CREB molecules required for long-lasting "late" LTP (L-LTP), although they have all the molecules necessary for shorter-lasting "early" LTP (E-LTP). Eric Kandel's group has developed PKA regulatory subunit transgenic mice that overexpress those molecules in specific neural regions. When activity-driven cAMP molecules release PKA catalytic subunits, an abundance of regulatory subunits are available to block PKA catalytic subunit translocation to the neuron's nucleus (in the regions of the brain where the transgene is expressed). This effect halts the gene expression and protein synthesis necessary for L-LTP. If the molecular mechanisms of L-LTP are those of memory consolidation, then Silva's CREB enhancer mutants and Kandel's PKA regulatory transgenics should be intact in short-term memory tasks but impaired in their long-term form. These are exactly their published experimental results. Kandel's results are especially compelling because the transgenic mice acquire long-term memories on tasks that involve activity in brain regions where the transgene is not expressed—tasks they learn

simultaneously with the long-term memory tasks on which they fail. This suggests that the deficit is not sensory, motor, or attentional, but instead is specific to memory consolidation.

New results from molecular and cellular cognition are reported in virtually every issue of journals such as *Cell*, *Neuron*, *Journal of Neuroscience*, *Journal of Neurophysiology*, and *Nature Neuroscience*. However, they have yet to creep into philosophical awareness. This is unfortunate for at least two reasons. First, this is mainstream neuroscience at the turn of the twenty-first century, employing techniques common to the bulk of the discipline's practitioners (especially compared to the number of cognitive neuroscientists). Second, this work is reductionistic, especially compared to higher-level neuroscience. Philosophers who limit their attention to the latter not only come away with a mistaken impression of what constitutes state-of-the-art neuroscience; they also miss the reductionist attitude that informs the mainstream. This carries problems especially for philosophy of mind. These implications are serious enough to motivate fuller discussion in the final section.

PHILOSOPHICAL IMPLICATIONS: REDUCTION REVISITED

When presenting important neuroscientific findings above, some philosophical implications were mentioned. In this final section, implications for reductionism will be discussed in more detail. Philosophical attention to neuroscience began with this concern. Reduction occupied an entire chapter in Patricia Churchland's ground breaking *Neurophilosophy* (1986). Other concerns emerged as philosophers engaged neuroscience, but reduction remains central to neurophilosophy—as witnessed by its prominent treatment in the first single-authored, introductory neurophilosophy textbook (Churchland 2002). Unfortunately, the term "reduction" is less univocal than it once was, and its philosophical treatments and discussions remain frustratingly abstract and distant from actual scientific practice. These features cast suspicion on assessments of psychoneural reductionism's philosophical potential. Might closer attention to mainstream (cellular and molecular) neuroscience rectify this?

Philosophical discussions of reduction were clearest and most fruitful when *intertheoretic reduction* was their explicit concern. This treatment goes back most prominently to Ernest Nagel's classic *The Structure of Science* (1961, ch. 11). According to Nagel, reduction is deduction—of the reduced theory, characterized syntactically as a set of propositions, with the reducing theory serving as premises. In interesting scientific reductions, the reduced theory contains descriptive terms that don't occur in the reducing, so the premises of the derivation must also contain bridging principles or correspondence rules. Typically these principles were treated as material biconditionals (although Nagel explicitly permitted material conditionals) containing terms from the two theoretical vocabularies. In interesting scientific cases, the reducing theory also often corrects the reduced. On Nagel's account, this feature is handled by introducing premises expressing counterfactual limiting assumptions and boundary conditions on the application of the reducing theory.

Both of these features came under serious philosophical criticism, many of which resulted from attempts by philosophers to apply Nagel's account to increasingly better described cases from the history of science (including classical equilibrium thermodynamics to statistical mechanics and the kinetic theory of gases, Nagel's own detailed example). Led by Thomas Kuhn, Paul Feyerabend, Kenneth Schaffner, Lawrence Sklar, Robert Causey, and Clifford Hooker, philosophers of science proposed alternatives to Nagel's conditions. Patrick Suppes even proposed scraping the entire syntactical view of theories and replacing it with a semantic view—theories as sets of models sharing set-theoretic or category-theoretic features. Intertheoretic reduction then turns into a mapping of these sets into one another in light of a variety of constraints and conditions.

One problem with applying these detailed accounts from the philosophy of science to philosophy of mind is that neither neuroscience nor psychology seems to provide robust enough theories. Most theories of intertheoretic reduction require a complete account of lower level phenomena in terms of laws, generalizations, or their model-theoretic counterparts. But even in the best cellular and molecular neuroscience, as in cell and molecular biology in general, few (if any) explanations are framed in terms of laws or generalizations. Many interactions are known to occur with predictable regularity and have both theoretical and experimental justification; but biochemistry hasn't even provided molecular biology with a general (and hence generalization-governed) account of how proteins assume their tertiary configurations. Molecular biologists know much about how specific molecules interact in specific contexts, but few explanatory generalizations are found in experimental reports, review articles, or textbooks; and the few that are found do not by themselves yield extensive predictions or explanations of lower level interactions. Finally, real molecular neuro-

science does not provide what some law-based accounts of scientific theory structure require. Its explanations do not specify how molecular biological entities interact in all possible circumstances. In light of these mismatches, intertheoretic reduction looks like a naive account of actual scientific practice.

Furthermore, its philosophical successor, *functional reductive explanation*, fares no better. According to this view, whose prominent advocates include Jaegwon Kim, David Chalmers, and Joseph Levine, a reductive explanation of a higher-level phenomenon is a two-step process. Step 1 requires a functional characterization of the phenomenon, in terms of its principal causes and effects. Step 2 involves the empirical, scientific search for the lower level processes, events, or mechanisms that realize this functional characterization. The reductive explanation of water by aggregates of H_2O molecules is a commonly cited example. Scientists characterize the causal roles of water and its basic properties, like its boiling point at sea level; and empirical research reveals that aggregates of H_2O molecules, with their physical and chemical properties and dynamics, provide the underlying mechanisms for those causes and effects. (This account of reductive explanation is often employed by critics of mind-brain reductionism. Many philosophical champions of the qualitative features of consciousness insist that no reductive explanation of them should be expected, because any attempt to functionalize these features will fail to capture their qualitative essence. Hence Step 1 of their potential reductive explanation cannot be achieved.)

It is not illuminating—quite the reverse, in fact—to force the actual details of state-of-the-art "molecular and cellular cognition" into this format. No procedures that typically occur in these experiments are serious candidates for Step 1 functionalization. And the empirical searches for mechanisms typically focus on finding specific *divergences* from control group behavior in experimental protocols that are commonly used to study the cognitive phenomenon whose neurobiological reduction is at issue. The key step in these experiments is the intervention step, where techniques of cell and molecular biology are used to manipulate increasingly lower levels of biological organization in living, behaving organisms. Animals receiving the intervention—be it cellular, pharmacological, or a bioengineered mutation—are compared to control animals on a variety of behavioral tests to find specific, narrow behavioral deficits.

These experiments are designed to leave most behaviors intact. For only then do experimenters claim to have found a "reduction," an "explanation," or a "mechanism" of cognition. To force this experimental practice into the common philosophical model of functional reductive explanation occludes the subtlety of choosing which cellular or molecular pathways to intervene into, the exquisiteness of the invention techniques employed, and the specificity of the measured behavioral effects when these experiments are successful. Good philosophical accounts of a scientific practice should illuminate, not obscure, these types of features-in-practice. Any consequences drawn about "psychoneural reduction" from an account that obscures them should be treated with suspicion.

This problem is beginning to look like one of imposing borrowed philosophical ideals onto actual scientific practice. Based on prior epistemological or metaphysical commitments, many philosophers approach the neuroscientific literature with preconceptions about "what reduction has to be." When they fail to find their relation obtaining, they either deny that psychoneural reduction is on offer or redescribe actual cases so that these at least approximate it. Both responses are objectionable. The first drives philosophy of mind continuously farther away from mainstream neuroscience, which grew increasingly reductionistic in the last two decades of the twentieth century. The second keeps borrowed philosophical ideals alive when their actual value grows increasingly questionable, and engenders criticisms of "reductionism" based on "better knowledge of the actual scientific details." A better approach within the philosophy of neuroscience might be to articulate the actual practices of reductionistic neuroscientists—the ones whose work contributes to the "mind-to-molecular-pathways-linkages" expressed in the quote cited above by Kandel, Schwartz, and Jessell. The result will be an account of real reduction in real reductionist neuroscience. One could then ask the different question of whether these practices and their results serve the philosophical purposes that reductionism claimed to serve.

It is still too early in this metascientific investigation to know the answer to the last question. But careful examination of the experimental work described toward the end of the previous section above shows that the dominant reductionistic methodology involves *intervening* into cellular or molecular processes and then *tracking* the behavioral effects in the living animal using standard tests drawn from experimental psychology. Often much *in vitro* experimental work must be done first to discover where these interventions are best placed and which intervention techniques are best suited for the task. Cellular physiology still contributes intervention techniques

such as cortical microstimulation; pharmacology still contributes a variety of drugs and delivery systems. During the last decade of the twentieth century, transcranial magnetic stimulation developed more precise techniques for delivering a circumscribed magnetic field to increasingly precise neuronal targets. And molecular biology and biotechnology provided powerful techniques for gene manipulations, enabling experimenters to develop targeted gene knockouts and to insert transgenes to inhibit or exacerbate specific protein synthesis. Attached to appropriate promoter regions (base pair sequences in the genetic material that control the onset of gene expression), transgenic expression and subsequent protein synthesis can be limited to increasingly localized neuron populations.

Armed with these cellular and molecular intervention techniques, and coupled with detailed neuroanatomical knowledge about cell circuits leading ultimately to motor neurons and the muscle fibers they innervate, neuroscientists can make increasingly accurate predictions of behavioral effects on a variety of experimental tasks. Successful experimental results yield the conclusion that the specific cognitive phenomenon, "operationalized" using the behavioral tests employed, reduces to the cellular or molecular processes intervened into, within the neurons comprising the circuits leading ultimately to the musculature. Appeals to "higher level" neuroscientific concepts and resources no longer appear in the resulting explanations. One reads in this scientific literature about contributions to "a molecular biology of cognition, to "bridges linking genes and behavior," and to explanations "of cognitive processes that integrate molecular, cellular and behavioral mechanisms." Within "molecular and cellular cognition," resources from cognitive neuroscience play essential heuristic roles. But once they have served their purposes to yield new "intervene molecularly and track behaviorally" results, they fall away from the discipline's best available account of cognition's neural mechanisms. Philosophers (and many cognitive scientists) might not recognize these scientific practices and results, but that reaction reflects nothing more than their lack of familiarity with ongoing neuroscientific practice. This methodology is central to mainstream reductionistic neuroscience at the turn of the twentieth century. If one wishes to rail against "psychoneural reductionism," one should at least rail against the actual practices and results of real reductionistic neuroscience—not against preconceived assumptions about what those practices and results "have to be."

This final point raises the intriguing question of whether neuroscience as a whole is univocal about the nature of reduction. More than likely it is not. Midway through the first decade of the twenty-first century, neuroscience is a remarkable interdisciplinary melding of different experimental techniques, methodological hunches, and interpretive assumptions. Molecular biology revolutionized the discipline in the late twentieth century, but so did new tools for functional brain imaging. Dynamical systems mathematics, applied initially to analyze artificial neural networks, provided fruitful new formal resources. Neuroscience's traditional core disciplines, neuroanatomy and electrophysiology, have enjoyed continual refinement. Rigorous neurological and neuropsychological assessment continue to develop. With so many questions being pursued—and philosophers would do well to compare attendance at their annual professional meetings with the more than 30,000 registrants at the 2004 Society for Neuroscience annual meeting—and so many techniques pitched at so many different levels of brain organization, it would be astonishing if "reduction"s meant the same thing across this discipline. Perhaps disagreements within philosophy about the neuroscientific plausibility of "psychoneural reduction" result more from philosophers latching onto different uses of this notion across neuroscience, rather than from ignorance or mistaken analysis. Sorting through these notions and discovering which neuroscientific practices employ each is one way that philosophers could contribute to ongoing neuroscientific development, instead of serving as mere sideline spectators or "science journalists."

See also Kim, Jaegwon; Kuhn, Thomas; Memory; Mind-Body Problem; Nagel, Ernest; Philosophy of Biology; Philosophy of Mind; Reductionism in the Philosophy of Mind.

Bibliography

Churchland, Patricia. *Brain-Wise: Studies in Neurophilosophy*. Cambridge, MA: MIT Press, 2002.

Churchland, Patricia. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press, 1986.

Kandel, Eric R., James S. Schwartz, and Thomas J. Jessell, eds. *Principles of Neural Science*. 3rd ed. New York: McGraw-Hill, 1991.

Kandel, Eric R., James S. Schwartz, and Thomas J. Jessell, eds. *Principles of Neural Science*. 4th ed. New York: McGraw-Hill, 2000.

Kolb, Bryan, and Ian Q. Whishaw. *Fundamentals of Human Neuropsychology*. 5th ed. New York: Bedford, Freeman, Worth, 2003.

Kosslyn, Stephen M. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press, 1994.

Llinás. Rodolfo, and Patricia Churchland, eds. *The Mind-Brain Continuum: Sensory Processes*. Cambridge, MA: MIT Press, 1996.

Molecular and Cellular Cognition Society. Available from http://www.silvalab.com/approachesf0.htm.

Nagel, Ernest. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace and World, 1961.

Posner, Michael I., and Marcus E. Raichle. *Images of Mind*. New York: Scientific American Library, 1997.

Ramachandran, Vilayanur S., Susan Blakeslee. *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: William Morrow, 1998.

Society for Neuroscience. Available from http://apu.sfn.org/.

Treisman, Ann. "The Binding Problem." *Current Opinion in Neurobiology* 6 (1996): 171–178.

Uttal, William R. *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge, MA: MIT Press, 2001.

Weiskrantz, Lawrence. *Consciousness Lost and Found: A Neuropsychological Exploration*. New York: Oxford University Press, 1998.

John Bickle (2005)

NEWCOMB'S PROBLEM

See *Decision Theory*

NEW ENGLAND TRANSCENDENTALISM

The New England transcendentalists were an influential but decidedly heterogeneous group of young writers, critics, philosophers, theologians, and social reformers whose activities centered in and around Concord, Massachusetts, from about 1836 to 1860. Insofar as they can be considered to have subscribed to a common body of doctrine, their leader and spokesman was Ralph Waldo Emerson (1803–1882). Apart from Platonism and Unitarian Christianity, the chief formative intellectual influence on the group was German idealism. It was not, however, the dense and difficult epistemological works of Immanuel Kant, Johann Gottlieb Fichte, Friedrich Schelling, and G. W. F. Hegel that primarily attracted the transcendentalists; although nearly all had made some attempt to read the German philosophers, very few had persevered to the point of mastering them. Rather, it was the more personalized and poetic expressions of Johann Wolfgang von Goethe, Novalis, William Wordsworth, Samuel Taylor Coleridge, and Thomas Carlyle, together with the belletristic expositions of Mme. de Staël's *De l'Allemagne* (New York, 1814) and Victor Cousin's *Introduction à l'histoire de la philosophie* (English translation, Boston, 1832) that provided Emerson and his disciples with whatever philosophical nourishment they possessed. Thus, far from being in any strict sense a primarily philosophical movement, New England transcendentalism was first and foremost a literary phenomenon. It was a passionate outcry on the part of a number of brilliant and highly articulate young Americans who had become so intoxicated with the spirit of European romanticism that they could no longer tolerate the narrow rationalism, pietism, and conservatism of their fathers.

After Emerson and Henry David Thoreau (1817–1862), the more important early transcendentalists were William Ellery Channing (1780–1842)—"Dr. Channing," as Emerson called him—distinguished clergyman and social reformer, leader of the Unitarian revolt against Calvinism; Amos Bronson Alcott (1799–1888), mystic, educationalist, and reformer; George Ripley (1802–1880), Germanist, disciple of François Marie Charles Fourier, and one of the founders of the Brook Farm community and of the *Dial* (the chief transcendentalist periodical); Orestes Augustus Brownson (1803–1876), journalist and clergyman whose lifelong attempt to reconcile religious conviction with radical views about social reform led him to embrace, in turn, nearly every available variety of Christianity from Presbyterianism to Catholicism; Frederic Henry Hedge (1805–1890), scholar, authority on German philosophy, founder in 1836 of the informal Transcendental Club for "exchange of thought among those interested in the new views in philosophy, theology and literature"; Margaret Fuller (1810–1850), literary critic, political radical, feminist, author of *Woman in the Nineteenth Century* (1845), and first editor of the *Dial* (1840–1844); Theodore Parker (1810–1860), dissenting Unitarian preacher and abolitionist whose ordination discourse, "The Transient and Permanent in Christianity" (delivered in Boston in 1841), denied the necessity of believing in biblical inspiration and in miracles and led Emerson to nickname him the Savonarola of transcendentalism; Jones Very (1813–1880), poet and eccentric; James Freeman Clarke (1810–1888), Unitarian minister and religious pamphleteer; and Christopher Pearse Cranch (1813–1892), minister, painter, critic, and poet. Among the later transcendentalists were John Weis (1818–1879), Samuel Longfellow (1819–1892), J. E. Cabot (1821–1903), O. B. Frothingham (1822–1895), and Moncure D. Conway (1832–1907). It is debatable whether Nathaniel Hawthorne should be counted as a transcendentalist, but

it is certain that, with other major imaginative writers like James Russell Lowell, John Greenleaf Whittier, Henry Wadsworth Longfellow, and Walt Whitman, Hawthorne owed much to his contact with transcendentalist modes of thought and feeling.

THE NATURE OF TRANSCENDENTALISM

"What is popularly called Transcendentalism among us," Emerson explained to a Boston audience in 1842, "is Idealism; Idealism as it appears in 1842" ("The Transcendentalist"). Yet we must add that it was a form of idealism that included and frequently confused the technical or epistemological idealism of the post-Kantian philosophers and the more vaguely understood "idealism"—in the sense of romantic aspirationism—of Wordsworth's "Intimations" ode and Novalis's *Fragmente.* The term *transcendental* was derived, Emerson claimed, from the use made of it by Kant, who had demonstrated that there was "a very important class of ideas, or imperative forms, which did not come by experience, but through which experience was acquired; that these were intuitions [*sic*] of the mind itself"; and that Kant had called them "Transcendental forms." This somewhat subjective exposition (contrast, for example, *Critique of Pure Reason,* B 25, A 11–12) led Emerson to conclude that consequently "whatever belongs to the class of intuitive thought, is popularly called at the present day *Transcendental.*" Here, of course, the word *intuitive* is being employed in its most general sense, quite dissociated from any philosophical use, so that Emerson could immediately go on lamely to characterize the "Transcendentalist" as one who displays a predominant "tendency to respect [his] intuitions."

The failure on the part of the movement's leader to give any really informative definition of transcendentalism is nevertheless instructive. Because of their intellectual eclecticism and avowed individualism, their subjective fads and eccentricities, and, above all, their wide range of activities, which embraced almost every aspect of American cultural life in the mid-nineteenth century, any attempt to express the outlook of the New England transcendentalists in a single formula is bound to fail. O. B. Frothingham was certainly right when he admitted that transcendentalism was not a systematic theory of life but something more like a state of mind, "an enthusiasm, a wave of sentiment, a breath of mind that caught up such as were prepared to receive it, elated them, transported them, and passed on—no man knowing whither it went."

In a clear sense, however, the transcendentalists were the inheritors of certain forms of sensibility already well developed within the European romantic movement: a vague yet exalting conception of the godlike nature of the human spirit and an insistence on the authority of individual conscience; a related respect for the significance and autonomy of every facet of human experience within the organic totality of life; a consequent eschewal of all forms of metaphysical dualism, reductivism, and positivism; nature conceived not as a vast machine demanding impersonal manipulation but as an organism, a symbol and analogue of mind, and a moral educator for the poet who can read her hieroglyphics; a sophisticated understanding of the uses of history in self-culture; in general, the placing of imagination over reason, creativity above theory, action higher than contemplation, and a marked tendency to see the spontaneous activity of the creative artist as the ultimate achievement of civilization—these were the more pervasive principles shared by all thinkers of the New England school. Yet if "idealism," or, better still, "romanticism," serves roughly to denote the genus of transcendentalism, it is important to determine the specific characteristics of the American version.

AMERICAN CHARACTERISTICS

American transcendentalism differed from its European counterparts in at least two important ways. First, unlike most forms of European idealism in the nineteenth century, transcendentalism was not simply closely allied with contemporary theological speculation and debate but arose directly out of it. The majority of its original adherents, including Channing, Emerson, Parker, Ripley, and Cranch, were, or had been, Unitarian clergymen, and from the point of view of cultural history the advent of transcendentalism must be seen as the final liberation of the American religious consciousness from the narrow Calvinism that Unitarianism had already done much to ameliorate. This is not, however, to imply that transcendentalism was primarily a movement within the Christian church. For its outcome, as the works of Emerson and Thoreau, for example, amply testify, was essentially secular and humanist in the widest sense.

Second, the later inception of romantic idealism in the United States led its exponents to less fluctuating and at the same time less radical programs of social reform. If the typical German or English romantic began with an enthusiasm for the ideals of the French Revolution, became disillusioned by the Terror, and ended his career a conservative, Emerson's disciples felt the outcome of the Revolution as something more distant and, in any case,

European. Their social philosophy was the natural outcome of their reactions to the very different American scene. The majority of transcendentalists never wavered in their active opposition to slavery, imperialism, bureaucratization, and cultural philistinism; yet, partly because the United States had already achieved a democracy and partly because Western expansion kept economic conditions relatively good, the transcendentalists were not incited to the more extreme forms of political protest characteristic of such European inheritors of idealism as Karl Marx and Pierre-Joseph Proudhon.

See also Brownson, Orestes Augustus; Carlyle, Thomas; Channing, William Ellery; Coleridge, Samuel Taylor; Cousin, Victor; Emerson, Ralph Waldo; Fichte, Johann Gottlieb; Fourier, François Marie Charles; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Idealism; Kant, Immanuel; Marx, Karl; Neo-Kantianism; Novalis; Parker, Theodore; Platonism and the Platonic Tradition; Proudhon, Pierre-Joseph; Schelling, Friedrich Wilhelm Joseph von; Staël-Holstein, Anne Louise Germaine Necker, Baronne de; Thoreau, Henry David.

Bibliography

Perry Miller's two collections, *The Transcendentalists: An Anthology* (Cambridge, MA: Harvard University Press, 1950) and *The American Transcendentalists: Their Prose and Poetry* (Garden City, NY: Doubleday, 1957), provide excellent selections of transcendentalist writings; they also contain bibliographies. O. B. Frothingham, *Transcendentalism in New England: A History* (New York: Putnam, 1876), is still the best intellectual history of the movement.

Michael Moran (1967)

NEW ENGLAND TRANSCENDENTALISM [ADDENDUM]

The transcendentalist departure from Unitarianism was bolstered by the Biblical criticism of Johann Gottfried von Herder, who suggested in *The Spirit of Hebrew Poetry* (1782/1833) both that the Bible is a human poetic construction, and that works just as authoritative can still be written. This was precisely Emerson's standpoint at the opening of *Nature* (1836), where he asked, "Why should not we have a poetry and philosophy of insight and not of tradition, and a religion by revelation to us, and not the history of theirs?" (1971–, 1:7). In his controversial "Divinity School Address" (1838), Emerson urged Harvard graduates to find redemption in the "Soul," not in an "eastern monarchy of a Christianity" that proceeded "as if God were dead" (1971–, 1: 82, 84).

In "Experience" (1844), Emerson developed most fully and creatively the Kantian idea that there are forms through which we acquire experience. Stating that the universe "inevitably wear[s] our color," Emerson developed a categoreal scheme that he called "the Lords of Life"—including "Temperament," "Surface," "Succession," "Surprise," and "Illusion." Against this background he set out an epistemology of moods, according to which moods are like beads strung on the iron wire of temperament, each showing "only what lies in its focus" (1971–, 3: 30). Emerson stated in "Circles" that "our moods do not believe in each other" (1971–, 2: 182)—a statement showing that moods contain beliefs and at the same time indicating their radically inconsistent outlooks.

Emerson's ethical thought centered on "self-reliance," which is both a positive search for the best in oneself—our "unattained but attainable self," as he put it in "History" (1971–, 2: 5)—and, in its negative moment, an "aversion" to "conformity." Emerson characterized society as "in conspiracy against the manhood of every one of its members" (1971–, 2: 29)—a conspiracy all too effective in producing individuals who "skulk" and "sneak" through their lives, or gather together like "bugs" and "spawn." Emerson's critique was thus directed not so much at specific actions as at a manner of living. He gave an existentialist twist to a passage from René Descartes's *Meditations* when he wrote, "Man is timid and apologetic; he is no longer upright; he dares not say 'I think,' 'I am,' but quotes some saint or sage" (1971–, 2: 38). For Emerson, as for his contemporary Søren Kierkegaard, thinking and existing are not just given; they are risky ventures. Emerson's heroes manifest a sense of command and overflowing worth, as well as a tendency toward spontaneity and whim. Friendships of such heroes are alliances of "large formidable natures, mutually beheld, mutually feared" (1971–, 2: 123).

Henry David Thoreau, in *Walden* (1854/1989), produced a work of ethical and political philosophy that, like Plato's *Republic*, considers the necessities of life. On the basis of his "experiment" of living at Walden Pond for two and a half years, Thoreau concluded that he can survive for a year on six weeks of labor. This left him time to "own" the landscape by sitting in it, sound the depths of the pond, watch the spring come in, talk with the occasional visitor, and, more generally, "improve the nick of time." Guided by the Greek and Roman philosophy he

read as an undergraduate at Harvard College and by his readings in Indian and Chinese thought, Thoreau understood philosophy as the search for "a life of simplicity, independence, magnanimity, and trust." In this sense, he observed, "there are nowadays professors of philosophy, but not philosophers" (p. 14).

In the "Economy" chapter of *Walden*, Thoreau considered human life as a precious commodity: "The cost of a thing is the amount of what I will call life which is required to be exchanged for it, immediately, or in the long run." He concluded that people pay a high cost for the lives they lead, that their lives are modes of strange "penance," and that a "stereotyped but unconscious despair is concealed even under what are called the games and amusements of mankind" (p. 8).

Although he portrayed himself variously as growing beans, peering through the ice of the pond, walking and sitting and "suddenly finding himself neighbor to the birds," the main outcome of Thoreau's time at Walden Pond was the book in which he recorded his life there, a book that, in the chapter "Reading," offered a theory of itself. Thoreau contrasted with the "classics" of every great culture a popular series of books called "Little Reading": books, as he put it, that "we have to stand on tiptoe to read and devote our most alert and wakeful hours to." After he finished *Walden*, Thoreau began to think of his immense journal as just such a book, perhaps even closer to nature, with "each page … written in its own season & out of doors" (1993, p. 67).

Thoreau's "Resistance to Civil Government" (1849) was a response to his night in jail for not paying the poll tax, and served as a source for the nonviolent resistance practiced by Mahatma Gandhi and Martin Luther King Jr. Thoreau argued that the citizen has no duty to align his conscience with the state, and a responsibility to oppose its immoral actions. He wrote, "I cannot for an instant recognize that political organization as my government which is the slave's government also" (1973, p. 67). The country could rid itself of slavery, he argued, if large numbers of people refused to pay their taxes and were willing to go to jail. Later, as Thoreau and Emerson became more agitated about slavery, Thoreau supported violence to end it. In "A Plea for Captain John Brown" (1859), he stated, "A man has a perfect right to interfere by force with the slaveholder, in order to rescue the slave" (1973, p. 132).

Margaret Fuller's death in a shipwreck in 1849 deprived the transcendentalists of a powerful journalist and feminist writer. In *Woman in the Nineteenth Century* (1845), a revision of her essay "The Great Lawsuit" (1843), she maintained that masculinity and femininity are intertwined, that there is "no wholly masculine man, no purely feminine woman." Women's free self-development, she argued, is necessary for the renovation of society, including marriage. "Union," she wrote, "is only possible to those who are units" (Myerson 2000, pp. 418, 419).

INFLUENCES ON PHILOSOPHY

Friedrich Nietzsche read Emerson at three critical points in his life, transcribed passages from Emerson's essays in his journals, and wrote, "*Emerson*.—Never have I felt so much at home in a book, and in *my* home" (Goodman 1997, p. 160). Emerson's ideas about nobility, history, friendship, overcoming self-inertia, and self-reliance presage Nietzsche's *Untimely Meditations* and *Thus Spoke Zarathustra*. A sentence from Emerson's "History" is the epigraph to the first edition of Nietzsche's *Gay Science*: "To the poet, to the philosopher, to the saint, all things are friendly and sacred, all events profitable, all days holy, all men divine" (Emerson 1971–, 2: 8).

In the United States, Emerson's stress on action and the future, his humanistic or Kantian portrayal of the role of the self in forming the world, and his focus on the individual chimed with central emphases of William James's pragmatism. John Dewey considered Emerson "the one philosopher of the New World fit to have his name uttered in the same breath with that of Plato," and found in his writings an anticipation of his view that ideals are present in our "immediate experience." Emerson and Thoreau are central to Stanley Cavell's investigations of "reading," "aversive thinking," and "moral perfectionism," and to his related discussions of Martin Heidegger, Friedrich Nietzsche, and Ludwig Wittgenstein in *The Senses of Walden* (1981), *Emerson's Transcendental Etudes* (2003), and other works.

See also Cavell, Stanley; Conscience; Descartes, René; Dewey, John; Emerson, Ralph Waldo; Emotion; Heidegger, Martin; Herder, Johann Gottfried; James, William; Kierkegaard, Søren Aabye; King, Martin Luther; Neo-Kantianism; Nietzsche, Friedrich; Plato; Pragmatism; Thoreau, Henry David; Wittgenstein, Ludwig Josef Johann.

Bibliography

PRIMARY SOURCES

Emerson, Ralph Waldo. *The Collected Works of Ralph Waldo Emerson*, edited by Robert B. Spiller, et al. Cambridge, MA: Harvard University Press, 1971–.

Emerson, Ralph Waldo. *Emerson's Antislavery Writings*, edited by Joel Myerson and Len Gougeon. New Haven, CT: Yale University Press, 1995.

Emerson, Ralph Waldo. *The Journals and Miscellaneous Notebooks of Ralph Waldo Emerson*, edited by William H. Gilman, et al. Cambridge, MA: Harvard University Press, 1960–.

Fuller, Margaret. *Margaret Fuller, Critic: Writings from the New-York Tribune, 1844–1846*, edited by Judith Mattson Bean and Joel Myerson. New York: Columbia University Press, 2000.

Fuller, Margaret. *"These Sad but Glorious Days": Dispatches from Europe, 1846–1850*, edited by Larry J. Reynolds and Susan Belasco Smith. New Haven, CT: Yale University Press, 1991.

Myerson, Joel, ed. *Transcendentalism: A Reader*. New York: Oxford University Press, 2000.

Thoreau, Henry David. *Journal*, edited by John C. Broderick, Elizabeth Hall Witherell, et al. Princeton, NJ: Princeton University Press, 1984–.

Thoreau, Henry David. *Reform Papers*, edited by Wendell Glick. Princeton, NJ: Princeton University Press, 1973.

Thoreau, Henry David. *Walden* (1854). Princeton, NJ: Princeton University Press, 1989.

Thoreau, Henry David. *A Year in Thoreau's Journal: 1851*, edited by H. Daniel Peck. New York: Penguin Classics, 1993.

SECONDARY SOURCES

Baumgarten, Eduard. "Mitteilungen und Bermerkungen über den Einfluss Emersons auf Nietzsche." *Jahrbuch für Amerikastudien* 1 (1956): 93–152.

Cameron, Sharon. *Writing Nature*. New York: Oxford University Press, 1985.

Cavell, Stanley. *Emerson's Transcendental Etudes*. Stanford, CA: Stanford University Press, 2003.

Cavell, Stanley. *The Senses of Walden*. Exp. ed. San Francisco: North Point Press, 1981.

Conant, James. "Nietzsche's Perfectionism: A Reading of Schopenhauer as Educator." In *Nietzsche's Postmoralism*, edited by Richard Schacht. Cambridge, U.K.: Cambridge University Press, 2000.

Goodman, Russell B. *American Philosophy and the Romantic Tradition*. Cambridge, U.K.: Cambridge University Press, 1990.

Goodman, Russell B. "The Colors of the Spirit: Emerson and Thoreau on Nature and the Self." In *Nature in American Philosophy*, edited by Jean De Groot, 1–18. Washington, DC: Catholic University of America Press, 2004.

Goodman, Russell B. "Moral Perfectionism and Democracy in Emerson and Nietzsche." *ESQ: A Journal of the American Renaissance* 43 (1997): 159–180.

Packer, B. L. "The Transcendentalists." In *The Cambridge History of American Literature*, edited by Sacvan Bercovitch, 2: 329–604. Cambridge, U.K.: Cambridge University Press, 1995.

Richardson, Robert D., Jr. *Emerson: The Mind on Fire*. Berkeley: University of California Press, 1995.

Versluis, Arthur. *American Transcendentalism and Asian Religions*. New York: Oxford University Press, 1993.

Von Frank, Albert J. *The Trials of Anthony Burns: Freedom and Slavery in Emerson's Boston*. Cambridge, MA: Harvard University Press, 1998.

Russell B. Goodman (2005)

NEWMAN, JOHN HENRY

(1801–1890)

John Henry Newman, an English philosopher of religion and cardinal of the Roman Catholic Church, was born in London, the son of a banker (later a brewer) who gave his children a love of music and literature. The young Newman was thoroughly familiar with the writings of both the romantic poets and the English deists. Raised as an Anglican, he underwent a deep religious experience when he was fifteen, and thenceforth he was strongly convinced of God's interior presence and providence. The mottoes chosen by Newman at this time foreshadowed his religious quest and interest in development: "Holiness rather than peace," and "Growth the only evidence of life."

He matriculated in 1816 at Trinity College, Oxford, where he read strenuously in the classics and mathematics. A fellowship at Oriel College at Oxford won him entrance to its common room, which proverbially "stank of logic." In 1824 Newman took holy orders.

The Oriel noetics, led by Richard Whately, gave Newman a taste for cool logical analysis of religious problems. His greatest influence at Oxford was exerted in company with Richard Froude, John Keble, and Edward B. Pusey. The Oxford movement sought to revive a living, full sense of the church and tradition through a series of incisive *Tracts for the Times* (1833–1841), culminating in Newman's *Tract 90,* which earned him an official censure. Newman's historical research in the Church Fathers and his theory of development in Christian doctrine eventually convinced him that the ideal of an Anglican *via media* was illusory. In 1845 he was received into the Roman Catholic Church, in 1847 he was ordained, and in 1848 he established the Birmingham Oratory as a center for those who shared his aspirations.

Newman struggled futilely during the years 1851–1858 to succeed as rector of the new Catholic University of Ireland, but political forces were too strong for him. Out of this defeat, however, came his main educational work, *The Idea of a University* (1852, 1859), which looked forward to a new synthesis of scientific, humanistic, and theological studies. Newman's strongly felt defense of his religious integrity and conversion expressed in his *Apologia Pro Vita Sua* (1864) restored his

rapport with educated readers in England. It also cleared the path for the presentation of his basic philosophical views on knowledge and his defense of the reasonable character of the act of religious faith. Newman regarded his *Essay in Aid of a Grammar of Assent* (1870) as his way of discharging an intellectual debt to his generation and to religious seekers of every age. In recognition of his distinguished service to the church, Pope Leo XIII created him a cardinal in 1879. Even in his last years, Newman kept up an active interest in questions of science, biblical criticism, and religious beliefs.

Newman belongs in the tradition of British churchmen who have contributed to philosophical thought. This he did in the course of dealing with certain problems of a religious and theological nature. He was well read in such Enlightenment sources as David Hume, Voltaire, and Thomas Paine and had an early awareness of the modern philosophical difficulties propounded against Christianity. Under pressure from such critics, Newman felt obliged to sift the grounds for his own adherence to theism and the Christian faith. He made a close study of the rationalistic apologetic used by William Paley and by Whately in defense of the existence of God and the basic articles of the Christian creed. Although Newman appreciated their search for rigor, he remained unconvinced by their particular way of achieving it. Their formalism remained completely impersonal and abstract, leaving out of account the process whereby the individual mind comes to see the import of an argument and gives its assent to the statements under discussion. Newman found a much more realistic account of mental operations in the analyses of inquiry made by three sources: Aristotle (especially in the *Nicomachean Ethics*), the Greek Fathers, and Joseph Butler. These sources all stressed the importance of probable reasoning and analogy, especially in cases involving contingent realities and moral questions. Somewhat to his surprise, Newman also discovered a similar stress in Francis Bacon, Isaac Newton, and the Newtonians as soon as they faced the problem of relating their formal structures to concrete nature.

FORMAL AND INFORMAL REASONING

Groping during his Oxford years for a way of stating the difference between the sequence of logical steps and the path of the mind in discovery, Newman came to the distinction between formal and informal reasoning. In mathematics and formal logic, the regulative principle is furnished by the formal relations among the elements of the argument and the internal consequence of steps. The relations can be stated in a general way without taking into account the difficulties that individual minds may have in following the formal entailments. From the logicomathematical standpoint, questions about our way of grasping the proof are either deemed irrelevant or assigned to the psychological order. Newman accepted this position insofar as it was meant to preserve the integrity of the standpoint of formal reasoning and the rigor of its deductive method. But he was unable to accept Whately's rationalistic conclusion that nothing more is ever required for establishing a doctrine than to exhibit its conformity with a pattern of formal reasoning. If a statement asserts something about existent things and if we are invited to accept this assertion, then something more is involved than the application of a general pattern of formal argument. The particular ways of backing the argument must be considered, and they must be considered by individual minds called upon to weigh their agreement with the world we experience.

When Newman himself tried to set down in the *Apologia Pro Vita Sua* the stages in his religious journey toward Catholicism, he found further evidence of his contention that the grounds and stages of argument in concrete matters cannot be fully formalized. He did not regard religious inquiry as being peculiar in this respect, but rather as agreeing with the common human condition of informal reasoning. The religious inquirer uses his mind in much the same way as does the jurist, the historian, and the biologist: All share in a common pattern of inquiry that demands a distinctive and responsible use of intelligence moving in a region somewhere between formalism and psychologism. A prominent task of Newman's main philosophical book, *An Essay in Aid of a Grammar of Assent,* was to explore the middle ground of inference that eludes complete formalization and yet achieves results capable of surviving the formal tests. In a general way, he described this region as a concrete personal mode of reasoning, which he customarily divided into natural and informal inference.

"CONCRETE" REASONING

The reasoning is called "concrete" as an indication of its ultimate terms of reference and control. Newman was strongly convinced that ours is a world of individual unit things, each of which has its unique nature and history. There is sufficient likeness among individuals to permit comparison and general statements, but there is no real identity and hence no completely general way of following the logical rules to establish our statements about them. In the study of individual entities, a gap eventually opens between general rules and concrete matters of fact.

It cannot be closed by carrying on some further manipulation of the formal procedures in logic, and one is forced to bring into play the personal discernment of the living mind working upon what it experiences. Man's reasoning becomes concrete in response to this situation.

When he inquires about concrete existents, each man assumes personal responsibility for the conduct of his own understanding. Although he cannot violate the logical system or the pattern of the language, he must determine issues that cannot be settled solely in their formal terms. In the ordinary course of life, one does not stop to reflect upon the methodological issues involved but plunges directly into the particular matters at hand. Newman refers to this unreflective and implicit sort of concrete thinking as a natural mode of inference, one that is not burdened by any second-level questioning about the kind of use being made of the mind. Every person is faced with practical decisions and moral choices that require a personal assessment of the circumstances and particular means and end in view. There is a point at which even a great military leader cannot rely solely upon the rules of strategy and his formal conception of warfare; he must place all these aids at the service of his personal estimate of a particular military situation in order to make a responsible decision. He is directly engaged in concrete reasoning in the natural mode of inference.

Yet Newman did not restrict concrete reasoning to conditions of great practical stress, where reflection on one's method is a luxury that cannot be indulged. He recognized the pattern of concrete intelligence in the judgments made by the historian, the art critic, the jurist, and the scientist. Here there is often an opportunity for attending to the problem of method. In the degree that individuals who make these judgments reflect upon their procedures and make an explicit theme of them, they are involved in what Newman calls concrete reasoning in the informal mode of inference. The concrete uses of intelligence are now thematized and critically controlled. The reasoning is informal insofar as it deals with questions that cannot be settled by appealing simply to the formal logical rules, but still it is a quite deliberate and reflective way of reasoning. Informal reasoning is required by our world of particulars, but this world does not prevent us from reflecting upon the way in which we explore and interpret it.

THE ILLATIVE SENSE

Newman proposed the theory of the illative sense to account for the certitude that may be attached to informal judgments. Here he was not trying to burden the mind with a new and esoteric faculty but sought instead to account for a definite feature of our intellectual activity. Hence he remarked that illative sense is only a grand name for designating a very ordinary way of using the mind.

A distinction is needed between certainty and certitude. Newman regarded certainty as a formally determinable quality of propositions and assigned its study to the logician. Newman's own interest centers upon certitude as a quality of the mind when it is engaged in concrete reasoning of both the natural and the informal sort. Concrete reasoning yields certitude when it enables us to recognize and affirm the truth of some proposition. Certitude is not achieved, as the rationalists maintain, through an impersonal coercion of the mind by the force of the formal elements contained in it. In all reasoning, but especially in concrete inference, certitude consists in an active response of the mind to the weight and tenor of the argument, a living recognition of the meaning and the truth of the proposition that states some findings. Furthermore, this certitudinal apprehension of the truth of the proposition is an inalienably individual act. I come to grasp the import of an argument; I see the bearing of the evidence; I give my assent to the proposition as true.

For my warrant in accepting the proposition, I cannot fall back exclusively upon the general canons of logic and the common structure of the language. Although Newman recognized their indispensable contribution by way of opposition to sentimentalism in thought, he believed that in the final analysis these elements cannot settle issues about the concretely existent. The illative sense refers to the type of operation of the human mind as it engages in concrete reasoning, reaches a conclusion of inference, and determines whether to give its certitudinal assent to the inferred proposition about a concrete reality:

> The sole and final judgment on the validity of an inference in concrete matter is committed to the personal action of the ratiocinative faculty, the perfection or virtue of which I have called the Illative Sense. … It is the mind that reasons, and that controls its own reasonings, not any technical apparatus of words and propositions. This power of judging and concluding, when in its perfection, I call the Illative Sense. (*Grammar,* Ch. 9)

Thus when Newman claimed to be developing a theory of the mind more empirical than John Locke's, he instanced this functional analysis of the illative sense.

The illative use of the mind is observable not only in the concluding act of an inference in concrete issues but also at the outset and along the way of the reasoning. Newman pointed out the need for a personal use of intelligence—especially in creative work as done, for example, by Newton or Edward Gibbon—in order to suggest the governing hypothesis, to gauge the strength of some particular stage in the inquiry, and to discern the bearing of many outlying investigations upon the main problem. We seek to conduct ourselves responsibly in all these operations, and the term *illative sense* refers to the intellectual mastery or perfection that an individual develops for inquiries in some concrete field. It comes close to the Aristotelian habit of prudence or practical wisdom, except that it can reach into the speculative order and attain certitude there. Newman added that despite a similar pattern of concrete logic for different fields, the personal mastery cannot simply be transferred from one area to another. A man may give us good grounds for trusting his judgment in military affairs or biological questions, whereas he may be utterly lacking in sagacity in respect to political legislation.

Newman did not isolate religious inquiry from other concrete uses of intelligence but required it to conform to the common requirements of concrete inquiry. The religious person is not concerned solely with abstract and general issues but seeks the truth about the reality of God, the person of Christ, the complex life of the church, and the individual soul's response to them all. These matters belong in the region of concrete existence and thus impose their own requirements upon the searcher's mind. The interested individual cannot do justice to the issues if he confines himself to what can be ascertained exclusively from the use of formal reasoning. Such a restriction is bound to lead to a noncommittal attitude, not because of the religious issues as such but because of the failure to make use of the concrete reasoning required by the situation.

PROBABILITY AND ASSENT

At this juncture, however, Newman was confronted with a strong objection propounded by William Froude (brother of Richard Froude) and other members of the Victorian scientific community. They noted Newman's statement in the *Apologia* about his agreement with Joseph Butler that probability is the guide of our life. In addition they noted the function assigned by Newman to the illative sense of discerning the convergence of probabilities among several strands of argument. To Froude, it seemed that the unavoidable result is that Newman's way of concrete reasoning can yield nothing higher than a probable conclusion, which is essentially open to constant revision. This falls considerably short of the certitude claimed by Newman for the act of religious faith.

Newman's treatment of this difficulty constitutes another major topic in the *Grammar of Assent.* Indeed, the book's title derives from his wrestling with this issue, as recorded in the following entry in his journal. "At last, when I was up at Glion over the Lake of Geneva, it struck me 'You are wrong in beginning with certitude—certitude is only a kind of assent—you should begin with contrasting assent and inference.' On that hint I spoke, finding it a key to my own ideas" (*Journal,* August 11, 1865). In fixing upon assent as something different from inference, Newman was able to clarify his position with respect to Froude's objection. His terminology was geared to the earlier, Lockean era in British empiricism, but the thrust of his argument concerns the relationship between religious faith and what Charles Peirce was already calling the ideal of scientific fallibilism.

Newman felt that at least one difficulty rested upon a linguistic confusion. His critics treated probability as a trait belonging to propositions and arguments, in which respect they contrasted it with the certainty of propositions. But just as he considered certitude a quality of the mind, so Newman viewed probability as a relationship involving the mind in an existential situation, rather than as a relationship among propositions in an argument. In Newman's conception, reasoning is probable to the extent that it is nonformal. Whenever inference is carried on in a context other than that of formal logic and mathematics, it is probable in the sense of not being governed by the intention of yielding a logicomathematical sort of proof. So understood, the probable is not contrasted with the demonstrative and the certain as such, but rather with the formal kind of demonstration and the abstract kind of certitude. Whenever the mind is inquiring about a concrete matter of fact, it is engaged in probable reasoning. This means that we are adapting our investigation to the conditions of particular existents, not that we are seeking only a weaker form of evidence and consequence in our reasoning. Thus probability, as understood by Newman, does not exclude certitude of assent but permits it to be achieved in matters pertaining to the concrete world and its connections in being.

Historically, Newman had to face Locke's restriction of probability to those inevident relations among ideas that permit neither intuitive nor demonstrative knowledge. Locke also held that belief is an act of assent that cannot rise above the probability of the inference leading

to it and hence cannot enjoy the certainty of intuition or demonstration. Newman had two grounds of disagreement with this teaching. First, there is no general rule necessarily subsuming religious assent under Lockean probability. Whether there is certitude in an act of religious faith cannot be settled by general stipulation about the meaning of probability and the judgment of belief. There must be a direct examination of the particular case and its grounds for claiming something about the order of concrete fact. Second, the act of assent is no mere shadow or reduplication of the conclusion of the inferential process. Using J. S. Mill's canons of induction, Newman sought to show the distinctive nature of assent as an act of the mind that remains irreducible to either the formal conclusion of an inference or to its psychological correlate in the act of concluding. We always conclude in a referential and conditional way, in view of what the premises state. But assent is made directly to the proposition as true; hence assent intends the certitudinal acceptance of the proposition in itself as being a true one. Newman made an extensive analysis of such expressions as "half assent, "conditional assent," and "hesitating assent." These describe circumstances surrounding the assent or features of the content to which assent is given rather than the act of assent itself.

The drift of Newman's reply to Locke and Froude is fairly clear. The sort of probability that he accepts as a guide and about which the illative sense must make an appraisal consists in a relation of the human mind to concrete modes of being. We follow the way of probability when we adapt our analysis to the concrete particulars and make a personal appraisal of the particular evidence. Our concrete personal thinking does not always attain certitude, but there is no a priori reason drawn from the definition of assent and probability that prevents us in principle from attaining it. Furthermore, there remains a difference in structure and intention between the inferential process and the act of assent. The revisability attaching to the former, especially in scientific inquiries, does not prevent the achievement of assent with certitude in some concrete instances. Newman's defense of the certitude in the act of religious faith depends upon keeping inference and assent distinct, as well as upon interpreting probability in terms of his theory of concrete reasoning.

NOTIONAL AND REAL ASSENT

Within the order of assent itself, Newman distinguished between notional and real assent. His view cannot be understood if it is taken as implying an opposition in principle between these modes of assent, or as assigning all the intellectual worth to real assent. The distinction is a functional one, arising from Newman's study of the interpretative operations of the mind. In assenting to a proposition, we can intend to accept the statement itself as true or to accept the real thing intended by the statement. A notional assent is one made to the truth of the proposition itself, whereas a real assent is one made to the reality itself intended by the proposition. Thus one may give a notional assent to God in terms of some abstract divine attributes and also give a real assent to God considered as a personal being who cares for one as an individual person. This is a matter of interpretation on the part of the mind that is considering the statement. In the case of purely ideal inquiries, a notional assent is sufficient. But we live in a translinguistic world, and our questions reach out to the community of real existents, especially to other persons. Here, the mind's notional assent must be integrated with, and further perfected by, a real assent to the very realities under investigation.

For Newman, the fully appropriate intellectual response to our human situation is unavoidably a complex one, involving both notional and real assents. Taken by itself, the way of real assent is intense but unclarified. We need to engage in both formal and informal inference, weighing the evidence carefully and arriving at a careful act of notional assent. Inference and notional assent are indispensable elements in human cognition; otherwise we could not weigh the pertinent evidence on an issue, do justice to the difficulties, or formulate the theoretical findings with cool precision of statement. Thus Newman assigned a large role to the modes of formal and informal inference and to notional assent in the total composition of human knowledge.

But he also insisted upon the need for directly relating the mind to individual existents. The act of real assent achieves our intellectual orientation toward the domain of concrete existents and their values. It does so by furnishing a concrete image of the individual being under consideration and by establishing the relevance of that imaged reality to the inquirer's own personal life. Real assent does not necessarily ensure action, but it does furnish a necessary condition for our practical responses by directing our mind toward the real existent, grasped in an image that can appeal to our passions and will.

There is a strongly theistic motive behind Newman's insistence upon blending inference, notional assent, and real assent. Humankind's relationship to God is not yet one of direct vision; hence we must engage in inference. Since theistic inquiry concerns a real existent, it is not enough to employ formal inference, even though its

resources must be used to analyze and test our arguments. A concrete personal mode of reasoning is also required in order to proportion our inquiry as fully as possible to the situation of man's search after the truth about God. Our aim must be the complex one of attaining some definite and well-grounded propositions to which we can legitimately give our notional assent, and also of forming a concrete image of the personal, morally good, and providential God to whose reality we can then give our real assent and practical attachment.

CONSCIENCE AND THE MORAL LIFE

Newman's final philosophical problem in the *Grammar* was to describe the area where he personally could realize this synthesis of intellectual acts bearing on the being of God. He readily admitted that there are many ways to God and that many natural informants lead us to him: the way of causality and purpose, the meaning of human existence and history, and the import of our moral life. As a reader of Hume and a contemporary of Charles Darwin, however, Newman refused to grant independent value to the design argument, which he regarded as a supplementary way of looking at nature on the part of those who already accept God on other grounds. To reach the transcendent, personal God, Newman examined the witness of our moral life, for this is a personal region where relations with other persons are best established. It is here that we have the experience of conscience, of being under command to do and not to do, of being responsible to a just and caring person who transcends our human reality but does so in a way that keeps him personally concerned about our conduct. Conscience as a commanding act discloses the full human situation of our responsibility toward the good God.

Three features of the living command of conscience recommend it to Newman as the best way of achieving real as well as notional assent to God: its intentional character, its personal significance, and its practical ordination. The dictate of conscience by its very structure refers the conscientious man beyond himself, pointing him toward the reality of the supreme lawgiver and judge of his moral actions. This is not a purely abstract orienting of our mind but involves a concrete image of God as our concerned father. Another advantage of the way of conscience is that the moral relationship in which it consists is personal in both poles of reference. Conscience engages me precisely as a personal self; hence it enables me to give a real assent to God as a morally concerned person. Finally, the acts of conscience relate us to the personal God in a concrete way that leads to moral and religious actions. Hence the approach to God from conscience encourages us to assent to the truth about God not only notionally but really, not only in respect to our propositions but also in respect to the personal, provident reality of God himself as the practical goal of our knowledge and love.

As a reader of Hume and Mill, Newman was very sensitive to the naturalistic criticism based upon physical and moral evil in our world. He suggested that the moral problem of theism be treated within a moral context. One cannot pose an objection to theism on moral grounds and then rule out the conditions that would permit theism to present its moral type of interpretation. Real assent to God as the lord of conscience furnishes a frame of reference for wrestling with evil and discerning his providential presence. A mind that is carefully formed upon the theistic implications of conscience "interprets what it sees around it by this previous inward teaching, as the true key of that maze of vast complicated disorder; and thus it gains a more and more consistent and luminous vision of God from the most unpromising materials. Thus conscience is a connecting principle between the creature and his Creator" (*Grammar,* Ch. 5). Whereas the naturalistic critic appeals to the vast disorder as an antecedent reason for withholding our assent from God, Newman asks us to secure first of all the inward principle of interpretation provided by the personal and moral relation of men to the lord of conscience. The work of this principle is not to soften or gloss over the power of evil, but to bring in the other considerations concerning God and moral man that will enable us to understand and work with hope against physical and moral evil in our world.

HISTORICAL DEVELOPMENT AND SOCIAL PRINCIPLES

Like other nineteenth-century thinkers, Newman was dissatisfied with the older empiricism's emphasis on the solitary and static individual perceiver. Hence he widened his horizon to include the social, developmental, and historical aspects of human experience. His *Essay on the Development of Christian Doctrine* (1845) opens with a chapter on the general nature and kinds of development among ideas. Here Newman explores the logic of those social ideals that grip the minds of men and account for developments in their beliefs and institutions.

For Newman, two questions are of prime importance in understanding the social growth of ideas and institutions: Why do certain ideas display themselves only through historical development? What pattern is com-

mon to diverse sorts of developing social principles? As an answer to the first question, Newman points to the interpretative activity of many minds as they are engaged in judging, relating, evaluating, and dealing practically with our complex world. There are some meanings that can be worked out only in this gradual social way. Historically important ideas are those that contain many facets and require the interpretative activity of many minds, testing and developing them over many years. "Ordinarily an idea is not brought home to the intellect as objective except through this variety; like bodily substances, which are not apprehended except under the clothing of their properties and results, and which admit of being walked around, and surveyed on opposite sides, and in different perspectives, and in contrary lights, in evidence of their reality" (*Development,* Ch. 1). We can grasp the intentional structure of basic human meanings only through studying their various perspectives, forcing them to enter the battlefield of critical discussion, and sometimes embodying them in visible, powerful social institutions.

Newman also suggested that there is a common pattern of development that has certain traits distinguishing a healthy growth from a sickly one. His seven criteria for genuine development are preservation of the type of principle that is socially influential, continuity of these principles, their capacity for assimilation of new data, their logical sequence in organizing a complex social process, their anticipation of their own future, conservation of their past achievements, and their chronic vigor. He deliberately illustrated these criteria by showing their development in kingdoms, economic policies, religious convictions, scientific hypotheses, and philosophical theories. Although the entire analysis is applied ultimately to the theological question of development among Christian doctrines, Newman's comparative use of empirical materials indicates the wider significance of his study of the dynamics of human thought and institutional forms. He himself, in fact, makes an explicit application of this theory of development to the ideas of civilization, the political constitution, and the university.

THE UNIVERSITY

Newman's effort at interpreting the Western ideal of the university in the context of his theory of development is revealed in *The Idea of a University.* He was more keenly aware than most of his contemporaries that the crucial decisions affecting the course of cultural development were being made within the university. It was replacing the episcopal palace, the banking house, and the parliamentary floor as the real center for determining the long-range direction of human history. Newman looked for a fresh synthesis of tradition and originality in the university community. The task of such a community is to educate men for the world by gradually introducing them to the full complexity of our humanistic, scientific, and religious interpretations. This it should try to do by cultivating an understanding of the various methods and ways of knowing, along with an awareness of their differences, limitations, and possibilities for unification.

As a Catholic churchman, Newman devoted the bulk of his writings to problems raised by the Christian faith and its practical institutions, especially as they are brought into close relation with modern humanistic and scientific ideas. His contributions to these issues might be considered as a sustained effort at education that draws its strength from both Christianity and the other components in the university ideal.

See also Aristotle; Bacon, Francis; Butler, Joseph; Darwin, Charles Robert; Enlightenment; Hume, David; Locke, John; Mill, John Stuart; Newton, Isaac; Paine, Thomas; Paley, William; Peirce, Charles Sanders; Propositions, Judgments, Sentences, and Statements; Religion; Religion and Morality; Voltaire, François-Marie Arouet de; Whately, Richard.

Bibliography

WORKS BY NEWMAN

The following are modern editions of Newman's works: *Two Essays on Biblical and on Ecclesiastical Miracles* (New York, 1924); *An Essay in Aid of a Grammar of Assent,* edited by C. F. Harrold (New York: Longmans, Green, 1947); *The Idea of a University,* edited by C. F. Harrold (New York: Longmans, Green, 1947); *Essays and Sketches,* 3 vols., edited by C. F. Harrold (New York: Longmans, Green, 1948); *Apologia Pro Vita Sua,* edited by C. F. Harrold (New York: Longmans, Green, 1947); *An Essay on the Development of Christian Doctrine,* edited by C. F. Harrold (New York: Longmans, Green, 1949); *Sermons and Discourses,* 2 vols., edited by C. F. Harrold (New York: Longmans, Green, 1949); and *The Letters and Diaries of John Henry Newman,* edited by C. S. Dessain (New York: T. Nelson, 1961–; to be published in 30 vols.).

WORKS ON NEWMAN

Benard, E. D. *A Preface to Newman's Theology.* St. Louis: B. Herder, 1945.

Boekraad, A. J. *The Argument from Conscience to the Existence of God according to J. H. Newman.* Louvain: Nauwelaerts, 1961.

Boekraad, A. J. *The Personal Conquest of Truth according to J. H. Newman.* Louvain: Nauwelaerts, 1955.

Bouyer, Louis. *Newman: His Life and Spirituality.* New York: P.J. Kenedy, 1958.

Collins, James. *Philosophical Readings in Cardinal Newman.* Chicago: Regnery, 1961. A presentation of Newman's philosophical views in terms of his theory of personal knowledge, his concrete inference to God, his notion of social development, and his account of the relationship between faith and reason.

Culler, A. D. *The Imperial Intellect: A Study of Newman's Educational Ideal.* New Haven, CT: Yale University Press, 1955. An interpretation of the liberal and traditional strains in Newman's mind and educational outlook.

Walgrave, J.-H. *Newman the Theologian.* New York: Sheed and Ward, 1960. An advanced analysis of the theory of development and the psychological aspect of Newman's approach.

Ward, Wilfrid. *The Life of John Henry Cardinal Newman.* 2 vols. New York: Longmans, Green, 1912. A reliable, standard life, with many letters and documents.

James Collins (1967)

NEWMAN, JOHN HENRY [ADDENDUM]

Since 1967, the publication of new primary source material has generated an expanding resource pool for secondary scholarship on Newman, particularly with the appearance of 24 new volumes to complete the thirty-one volume collection of Newman's *Letters and Diaries*. In addition, two volumes of Newman's *Theological Notebook* (1970), two volumes of his *Theological Papers* (on Faith and Certainty [1976], and on Biblical Inspiration and Infallibility [1979]), and an annotated bibliography of his Tract and Pamphlet Collection (1984) have been published. A new critical edition of the *Grammar of Assent* was produced by Ian Ker in 1985, and new editions of several of Newman's works appeared: *Oxford University Sermons* (1970), *Apologia Pro Vita Sua* (1993), and *Arians* (2001). The celebration in 1990 of the centenary of Newman's death was the occasion for two new biographies by Ian Ker (1989) and Sheridan Gilley (1990). Moreover, the journal *International Cardinal Newman-Studien* (known until 1987 as *Newman-Studien*) continues to appear annually. The result of these increased resources has been a wide variety of secondary literature documenting Newman's contributions to classical themes, as well as the opening up of some new directions in scholarship.

Of particular relevance to philosophy is the continuing discussion of Newman's understanding of the relation between faith and reason and the relation between faith and doubt. Debates that locate Newman in the history of responses to skepticism (including Wittgensteinian responses) continue about the plausibility of Newman's claims that there are no degrees of assent, that assent (including the reflex assent of certitude) is an act of the will, that indubitability (the absence of "reasonable" doubt) can be achieved through convergent, nondemonstrative reasoning, and that certitude is indefectible. In particular, the period from 1969 to 1980 saw increased attention to a debate about whether Newman was a "volitionalist" (aligned with people like René Descartes and Søren Kierkegaard)—that is, whether assent was an act of the will distinguished from and following on the reasoning process, according to a "logic of decision." While there continue to be advocates of Newman's volitionalism, this debate opened up a new direction for research—namely, the theme of Newman and rhetoric. In addition to three book-length studies of Newman as a rhetorician, in the sense of classical rhetoric, three new studies of his preaching appeared. A collection of essays on romanticism and rhetoric in Newman's thought was complemented by the beginning of significant discussion of the role of imagination in Newman's proposals concerning concrete reasoning and the illative sense.

Theological interest in Newman's thought has resulted in works on his ecclesiology, and the topics of liturgy and revelation. Another interesting new direction in Newman studies has been an increased emphasis on spirituality. Although there were earlier works on Newman's spirituality, such as Hilda Graef's *The Spirituality of John Henry Newman* (1968), the late 1980s and early 1990s saw the publication of three additional works on Newman's spirituality, his "spiritual theology" and Newman's teaching on "Christian holiness." Perhaps this increased interest in spirituality is related to the initiation of the process of beatification and canonization of Newman begun by the Roman Catholic Church in 1980; in 1991 the first official step in that process was taken when Pope John Paul II declared Newman "Venerable."

While there has been no notable book-length feminist study of Newman's thought, there has been some interest in Newman's relation to women (Joyce Sugg, *Ever Yours Affly: John Henry Newman and His Female Circle*, 1996), as well as the influence of Mariology (Philip Boyce, *Mary: The Virgin Mary in the Life and Writings of John Henry Newman*, 2001).

Finally, in addition to publications in church history, in which Newman is related to the Oxford Movement and to Modernism, the centenary celebration of Newman's death brought about a number of retrospectives in the form of edited volumes of essays by specialists, for example, Ian Ker and Alan Hill's 1990 *Newman After a Hundred Years*. There followed a decade of increased interest

in Newman, including two collections of interdisciplinary studies in which scholars consider Newman from the perspectives of literature, history, and education (edited by Magill, 1993 and 1994).

See also Descartes, René; Doubt; Faith; Kierkegaard, Søren Aabye; Modernism; Reason; Skepticism; Volition; Wittgenstein, Ludwig Josef Johann.

Bibliography

Biemer, Günter, and Heinrich Fries. *Christliche Heiligkeit Als Lehre Und Praxis Nach John Henry Newman—Newman's Teaching on Christian Holiness*. Sigmaringendorf, Germany: Regio Verlag Glock und Lutz, 1988.

Blehl, Vincent Ferrer. *The White Stone: The Spiritual Theology of John Henry Newman*. Petersham, MA: St. Bede's, 1994.

Boyce, Philip, ed. *Mary: The Virgin Mary in the Life and Writings of John Henry Newman*. Grand Rapids, MI: Eerdmans, 2001.

Britt, John. *John Henry Newman's Rhetoric: Becoming a Discriminating Reader*. New York: Peter Lang, 1989.

Coulson, John. *Religion and Imagination: "In Aid of a Grammar of Assent."* Oxford: Clarendon Press, 1981.

Ferreira, M. Jamie. *Doubt and Religious Commitment: The Role of the Will in Newman's Thought*. Oxford: Clarendon Press, 1980.

Gilley, Sheridan. *Newman and His Age*. London: Darton Longman and Todd, 1990.

Graef, Hilda. *God and Myself: The Spirituality of John Henry Newman*. New York: Hawthorn, 1968.

Jost, Walter. *Rhetorical Thought in John Henry Newman*. Columbia: University of South Carolina Press, 1989.

Ker, Ian. *The Achievement of John Henry Newman*. Notre Dame, IN: University of Notre Dame Press, 1990.

Ker, Ian. *Healing the Wound of Humanity: The Spirituality of John Henry Newman*. London: Darton, 1993.

Ker, Ian, and Alan Hill, eds. *Newman after a Hundred Years*. Oxford: Oxford University Press, 1990.

Kerr, Fergus, and David Nicholls. *John Henry Newman: Reason, Rhetoric and Romanticism*. Carbondale: Southern Illinois University Press, 1991.

Magill, Gerard. *Discourse and Context: An Interdisciplinary Study of John Henry Newman*. Carbondale: Southern Illinois University Press, 1993.

Magill, Gerard. *Personality and Belief: Interdisciplinary Essays on John Henry Newman*. Lanham, MD: University Press of America, 1994.

Pailin, David A. *The Way to Faith: An Examination of Newman's "Grammar of Assent" as a Response to the Search for Certainty in Faith*. London: Epworth Press, 1969.

Sugg, Joyce. *Ever Yours Affly: John Henry Newman and His Female Circle*. Leominster, U.K.: Gracewing, 1996.

Tolhurst, James. *The Church—A Communion in the Preaching and Thought of John Henry Newman*. Leominster: Fowler Wright, 1988.

Whalen, David M. *The Consolation of Rhetoric: John Henry Newman and the Realism of Personalist Thought*. San Francisco: International Scholars Publications, 1994.

Willi, Peter. *Sünde Und Bekehrung in Den Predigten Und Tagebüchern John Henry Newmans*. St Ottilien: Eos, 1992.

M. Jamie Ferreira (2005)

NEW REALISM

"New Realism" arose at the turn of the twentieth century in opposition to the Idealist doctrines that the known or perceived object is dependent for its existence on the act of knowing and that the immediately perceived object is a state of the perceiving mind. The Austrian philosophers Franz Brentano and Alexius Meinong first enunciated the cardinal tenet of this new realism: that what the mind knows or perceives exists independently of the acts of knowing and perceiving. Developing mainly as a polemic against Idealism, this new realism was represented prior to 1900 in England in the works of such men as John Cook Wilson, Thomas Case, H. W. B. Joseph, and H. A. Prichard. Similar realist polemics were taking place in Sweden and Italy.

In America the movement known as New Realism dates from the critical writings of William P. Montague and Ralph Barton Perry in 1901 and 1902. Their immediate aim was to refute Josiah Royce's "refutation" of realism, which he had based on the claim that the knower and the known could not be independent of each other and still be related. The movement took definite form when Montague and Perry were joined by four others in a statement of a New Realist program ("The Program and First Platform of Six Realists") in 1910.

In England, New Realism took explicit form in the works of T. P. Nunn, Bertrand Russell, and G. E. Moore. In both America and England, New Realists asserted the independence of consciousness and its object, but serious differences soon appeared between the two groups and between individuals within each group. The differences were particularly noticeable in their statements about the nature of consciousness and of its object, and of the relation between them. Moore claimed that the act of consciousness included both a nonmental, independent object and a transparent, or "diaphanous," mental act of consciousness. He agreed with Brentano and Meinong that consciousness involved awareness in the form of an act of intending something other than itself. To have an idea, to perceive or be aware at all, is already to be beyond consciousness and to be confronted by an independent object. American New Realists, on the other hand, took their view of consciousness from William James. While he, too, described consciousness as a relation, James

denied that there was anything uniquely mental or psychic about it at all, and associated consciousness rather with the behavioral responses or functions of the organism.

But there were also differences between Moore, Nunn, and Russell. Nunn argued that both primary and secondary qualities not only exist as they are perceived, but also are really *in* their objects, whether perceived or not. He even argued that pain is something independent of mind, with which mind may come into various relations. In this he was closer to the American New Realism of Perry and E. B. Holt. Russell was influenced by Nunn's view, but his New Realism took a frankly Platonic turn that brought it closer to the New Realism of Montague. Russell's Realism, however, was soon significantly altered. Another variant of English New Realism, perhaps more a development from it than a version of it, was Samuel Alexander's. It, too, resembled American New Realism.

AMERICAN NEW REALISM

Although American, English, and, to a lesser extent, European New Realists influenced one another, it was among the Americans that New Realism flourished, particularly as a movement. Their aim was to produce an account of how a real object could be present in consciousness and knowledge and still be independent of that relation, and they sought to do this without a dualistic separation of knower and known. "The independence of the immanent" was their manifesto. Their first platform statement consisted of six lists of doctrines that had been discussed at length, revised, and agreed to by all, and that all thought were consistent. The lists were signed by Holt and Perry at Harvard, Walter T. Marvin at Rutgers, Montague and Walter B. Pitkin at Columbia, and Edward C. Spaulding at Princeton.

At a Philosophical Association meeting in 1909, five of these six had found themselves in agreement against a common foe that still spoke with authority and was listened to with deference: Idealism. Pitkin and Montague are credited with the idea of translating their agreement into an articulate statement, and papers soon began circulating. F. J. E. Woodbridge at Columbia gave encouragement, although he declined an invitation to join. Montague, in "Confessions of an Animistic Materialist," described E. B. McGilvary, Morris R. Cohen, J. E. Boodin, J. Lowenberg, and Douglas C. Macintosh as "unofficial" New Realists. Believing that philosophic disagreements were the result chiefly of a lack of precision and uniformity in the use of words, plus a lack of planned cooperation in research, the original six banded together in the hope of revealing the genuine philosophic disagreements that were more than mere differences of personal opinion. They hoped thereby to open the way to the solution of genuine philosophic disputes. They called for a new alliance between philosophy and science and formulated a statement of principles and doctrines, a program of constructive work with a method based on these, and an agreed-upon system of axioms, methods, hypotheses, and facts.

In 1912 they published their cooperative volume, *The New Realism; Cooperative Studies in Philosophy*. Although they were still preoccupied with polemics, the six authors hoped to go beyond criticism to produce a complete philosophy that would play a major part in human thought. They saw themselves as proponents of a doctrine concerning the relation between the knowing process and the thing known. They described their most urgent problem (one that had not been resolved by naive realism, dualism, or subjectivism) as how to give an adequate account of "the facts of relativity" in the knowing process from a Realist point of view; how, in other words, to reconcile the apparently hopeless disagreement of the world presented in immediate experience with the true or corrected system of objects in whose independent reality they believed. While New Realism succeeded in showing the fatal weaknesses in dualistic answers to this problem, it nonetheless failed to provide an adequate answer of its own.

THE "FACTS OF RELATIVITY." New Realism faced the above problem not just because Idealism had failed to resolve it but also because Idealism had made it impossible to ignore these "facts of relativity." Thus, any attempt by New Realists to return to the naïveté of earlier doctrines of realism, to a primitive notion that nothing intervenes between subject and object (particularly nothing attributable to the subject), was out of the question. Equally closed to them was any recourse to a Lockean or Cartesian dualism that, they thought, never escaped the subject's own mental states. The third traditional answer to the problem, subjectivism, was also impossible. Of the three approaches, subjectivism was most often the object of criticism by New Realists, and they identified it as the fatal doctrine of Idealism. They saw it as an illicit argument from the "egocentric predicament," an argument based on the difficulty of conceiving known things to exist independently of their being known. New Realists refuted Idealism by refuting this argument; but then it became their turn to reconcile the facts of relativity, of which the predicament was one, with their theory of the

independent existence, or reality, of objects of consciousness and knowledge.

New Realist writings thus were largely devoted to such facts of relativity as illusion, error, secondary qualities, and—later—choosing, valuing, meaning or intending, and purposing. The New Realists also thought that Idealism had gone too far in its view of the subject's role. However, if Idealism went too far in that direction, New Realism went too far in the opposite direction; its polemical theory of independence could not be reconciled with the facts of relativity. This in turn provoked such reactions as Critical Realism, Perspective Realism, and Objective Relativism.

Chief among the positive aspects of the doctrines of the New Realists was what they called the "emancipation of metaphysics from epistemology," the result of their theory of independence. Contrary to the Idealist claim that knowing was the universal condition of being and hence constitutive of it, the New Realists argued that knowing and being were independent. This, Perry showed, did not mean they were therefore unrelated, as Royce had argued, but simply that there was not the particular relation of dependence between them. Dependence is a special type of relation in which the dependent element contains, implies, or is exclusively caused or implied by that on which it is dependent. Between knowing and being, therefore, it was possible for there to be relations both of independence (external relations) and of dependence (internal relations). In holding out this possibility against the Idealist claim that all relations are internal, New Realism became identified with a theory of external relations.

In "immediate and intimate connection" with this theory was the doctrine that the content of knowledge is numerically identical with the thing known; things, when consciousness is had of them, become contents of consciousness, thus figuring both in the external world and in "the manifold which introspection reveals." This view was very close to James's Neutral Monism, but only Holt worked out its fullest implications. The theory of numerical identity soon became the target of critics of New Realism, and it was difficult to determine whether, and to what extent, any New Realist other than Holt maintained it. Yet for a time, at least, it was said to be fundamental to New Realism. If there was a numerical identity between consciousness and its contents, then the "things" of thought would have to be given full ontological status along with the "things" of sense. This the New Realists claimed to do in their volume. They said they were Platonic Realists in granting this status to subsistents as well as existents. Here, again, a belief held by all in the beginning became in the end the belief of but a few, notably Montague and Spaulding.

THE EGOCENTRIC PREDICAMENT. The facts of relativity haunted New Realism throughout the life of the movement. That the New Realists ultimately failed in their professed aim of doing justice to these facts was in part the result of their constant polemical concern with asserting their doctrine of independence against Idealism and in part the result of their failure to recognize some possibly constitutive elements within the knowing relation. One such fact was the egocentric predicament, described by Perry as the fact that the "extent to which knowledge conditions any situation in which it is present cannot be discovered by the simple and conclusive method of direct elimination" ("The Ego-Centric Predicament"). Perry thought this was merely a methodological difficulty, one faced by all philosophers. Idealism had used it to argue that since it was impossible to discover anything that is, when discovered, undiscovered by someone, therefore it is impossible to discover anything that is not thought. The argument, Perry contended, rested on a confusion between "everything which is known, is *known*," and "everything which *is*, is known."

Perry concluded that the predicament could not be used to support either Idealism or Realism. Idealists could not use it as an argument for dependence, or internal relations, and New Realists could not use it as an argument for independence, or external relations. But while exposing its illicit use, New Realists did not offer a convincing way out of the predicament. As a test for the dependence or independence of any element in consciousness, Perry proposed that insofar as the element was deducible from anything other than consciousness, it was independent. To be dependent, or subjective, the element would have to be exclusively determined by consciousness. However, it was pointed out, the predicament would prevent us, by the very test Perry proposed, from reaching an object that we could be sure was independent of consciousness, for we would be using consciousness (deduction) in order to get to it.

Spaulding maintained that New Realism had provided a solution to the predicament and that this solution was its most important doctrine. He argued that any sort of analysis purporting to discover—and not merely create—what is *there* would be impossible if it did not presuppose a Realist position; that is, presuppose relatedness with independence. Even a theory that argued against the Realist position would have to take that position toward

the very state of affairs it described, assuming that it was a genuine state of affairs, not one created, altered, or modified by virtue of the knowing relation. Every philosopher, knowingly or not, solves the predicament by the Realist attitude he assumes toward his subject. But the question remained: What warrant do we have for such an assumption?

Pitkin attempted to support the doctrine of external relations by refuting the assertion that biology provided evidence for the internalist view. On the contrary, he argued, biology supports the externalist view through the discovery that organic parts do not depend upon the whole in which they naturally occur; and an organic whole does not depend upon its individual parts for its total specific organic character.

Beyond this, and apart from showing that independence did not rule out relatedness, the New Realists did not demonstrate how the knowing relation was external and independent, nor did they show how the facts of relativity were to be reconciled with externality and independence. In their cooperative volume they had refused to recognize ultimate immediacies, or any nonrelational or indefinable entities other than the simples in which they claimed analysis terminates. Their view that the knowing relation was external required such simples, or "neutral entities," that would maintain their identity no matter what relations they entered into. But it was never clear why analysis had to stop where the New Realists said it did—usually with the simples of mathematics and logic. Nor was it clear whether these simples were the product of their analysis or a genuine discovery by it.

EPISTEMOLOGY AND ONTOLOGY. In its constructive phase, New Realism proposed an epistemological monism and an ontological pluralism. James had argued that consciousness was not a substantive entity, and Moore similarly argued that it was diaphanous and transparent. In both cases, consciousness of something was viewed as a direct, unmediated, immanent affair. All content of consciousness, with the exception of Moore's psychical, diaphanous element, was thus objective in the sense that it consisted of objects in the real, external world. This was New Realism's epistemological monism: Thought and its object are numerically the same.

Its ontology was pluralistic, however: Some elements of the object would not be found in the consciousness of that object. Any elements in consciousness not found in the object would give consciousness a constitutive role beyond mere selection or grouping. The problem was to account for all of the "facts of relativity" through the selective and grouping function of consciousness without jeopardizing the New Realist theory of immanence that asserted that it was the "real" objects of the external world that were present in consciousness.

There were two principal positions taken on this matter among New Realists. Montague called them the left and right wings of New Realism. One was Neutral Monism, developed by Holt and, to a lesser extent, by Perry, but eventually abandoned by both. The other was a Platonic Realism developed by Montague into what he called Subsistential Realism.

Holt and Perry. Neutral Monism derived from James's idea of "pure experience." Pure experience was pure because it was uncontaminated by such distinctions as "object," "content," "subject," or "knower and known." It was "neutral" in terms of these distinctions; such distinctions could only be made later in terms of the relations between portions of pure experience. A "thing" could be said to be one portion of pure experience that was represented by another portion. A "thought" could be said to be one portion of pure experience that represented another portion. The dualisms of "inner" and "outer," mind and body, thus were undercut. All such distinctions were a matter of relations between bits of pure experience, but these relations had to be external. Hence, "mental," "nonmental," "real," "external," and "physical," are accidental features. New Realists thus were driven back to a realm of indefinable simples that come into and go out of various relations but never change their original identities. Where could such a realm be found? And what could these simples be?

Where James thought they were bits of pure experience (and may have been working toward an identification of experience with nature), Holt and Perry, influenced by developments in mathematics and symbolic logic, found these entities in a mathematical-logical realm of "being." It was a realm of entities having no definition or identity: neutral entities. These entities were similar to the simples that the New Realists had said analysis ultimately discloses. What we call consciousness is a grouping of these entities resulting from the selective (although not constitutive) response of the nervous system. This explanation enabled Holt and Perry to maintain the New Realist claim that consciousness and its objects were identical: Error and illusory experiences were no less objective or real than veridical experience. However, it failed to give an account of the difference between objects grouped and objects not grouped by consciousness. And it was still no easier to give an account

of the organism's response to objects that were spatially or temporally distant.

Although he espoused Neutral Monism in his early years, Perry never went as far as Holt. He admitted that error and other nonveridical experiences were cases of "mis-taking" entities for something other than what they are. In a later development he identified this mis-taking as an anticipation or expectation of an event that does not, when acted on or verified, occur as expected. By this time, however, Perry had departed from the New Realist theories of independence and immanence.

Spaulding and Montague. Spaulding also identified error as a mis-taking, but he described it as a case of taking something to be existential that was only "subsistential." This mis-taking was the only subjective feature in consciousness. Therefore, he concluded, illusory objects and errors are objective and real because both the existential and subsistential are objective and real. It is the taking of a thing to be what it is not that is the psychic or subjective element in consciousness, and the problem of error—why error occurs—is one for psychologists and not for philosophers. Along with Pitkin, Spaulding also took a behaviorist view of consciousness, describing its objects as nonspatial projections or dimensions of spatial objects resulting from the interaction of organism and environment.

The second major attempt to formulate a New Realist epistemology and ontology consistent with the doctrines of independence and immanence was developed furthest by W. P. Montague, the only one of the New Realists who argued for uniquely mental, subjective elements in knowledge and experience. While admitting this was dualism, he insisted it was not the psychophysical dualism rejected by New Realism. He invoked a realm of subsistents, identifying them as propositions of which existential propositions, and hence existence, were a part. Error was a case of mis-taking the "merely" subsistential to be an existent as well.

CRITIQUES OF NEW REALISM. All of these attempted solutions raised the question of whether New Realism's epistemology, based on an independently real object immanent in experience, could coexist with its view that the real object was part of the commonsense world. When the independence of the object of knowledge was emphasized, the facts of relativity were slighted, but the object could more easily be identified with commonsense objects. On the other hand, when immanence of the object was emphasized, it tended to lose its commonsense quality, becoming instead a neutral entity, or subsistent, or simple, supposedly disclosed by a rather sophisticated analysis. At the same time, however, the facts of relativity could more easily be taken into account. The former emphasis moved in the direction of dualism; the latter in the direction of monism.

Criticisms of New Realism in the second decade of the twentieth century were concerned mainly with showing that the organism intervenes in a considerably less naive way than the New Realists had thought and that their theories of external relations, independence, and immanence did not adequately account for what was given in knowledge and experience. Describing New Realism as the first phase of the "revolt against dualism," A. O. Lovejoy said its constructive program argued that since nothing "mental" could be admitted without leading to subjectivism and skepticism, therefore no content could be held to be psychically generated or dependent upon percipient functions. New Realism was left with things in a purely external relation to consciousness, or at best a bare and sterile awareness of them. In rejecting all mediated knowledge, he argued, New Realism could only hold the position that all content of experience must be identical with reality; everything before or "to" mind or consciousness was "objective." When this claim collided with the manifestly disparate content of nonveridical experience, an objective but "subsistent" content was said to be directly present or immanent; or, alternatively, this content was said to be no less objective than veridical content because it was at bottom ("neutrally") the same as it. But, Lovejoy concluded, this was little more than what the earlier naive, or commonsense, realism had said.

Although the New Realists hoped to produce other collections of studies, and although their discussions continued through 1914, according to Perry disagreements that had been subordinated and only imperfectly concealed, divergence of interests, and the ambition of each to write his own book soon divided them. As a movement, New Realism was soon displaced by the second major realist movement of the twentieth century, Critical Realism, which also developed and published a platform and joint program.

See also Alexander, Samuel; Brentano, Franz; Cohen, Morris Raphael; Critical Realism; Holt, Edwin Bissell; Idealism; James, William; Lovejoy, Arthur Oncken; McGilvary, Evander Bradley; Meinong, Alexius; Montague, William Pepperell; Moore, George Edward; Perry, Ralph Barton; Realism; Royce, Josiah; Russell, Bertrand Arthur William; Woodbridge, Frederick James Eugene.

Bibliography

WORKS BY NEW REALISTS

Holt, Edwin B. et al. "The Program and First Platform of Six Realists." *Journal of Philosophy* 7 (July 21, 1910): 393–401. Reprinted in their cooperative volume *The New Realism; Cooperative Studies in Philosophy*. New York: Macmillan, 1912.

Montague, William P. "Confessions of an Animistic Materialist." In *Contemporary American Philosophy*, edited by W. P. Montague and G. P. Adams. New York: Macmillan, 1930. Vol. II, 135–158. Pieces by other New Realists, "official" and otherwise, will be found in both volumes of this work.

Montague, William P. "Professor Royce's Refutation of Realism." *Philosophical Review* 11 (January 1902): 43–55.

Moore, G. E. "The Refutation of Idealism." *Mind*, n.s. 12 (1903): 442–453. Reprinted in Moore's *Philosophical Studies*. London: Routledge, 1922.

Perry, Ralph B. "The Ego-Centric Predicament." *Journal of Philosophy* 7 (1) (1910): 5–14. Reprinted in *The Development of American Philosophy*, edited by W. G. Muelder and L. Sears. Boston: Houghton Mifflin, 1940; 2nd ed., 1960.

Perry, Ralph B. "Professor Royce's Refutation of Realism and Pluralism." *Monist* 12 (1901–1902): 446–458.

Perry, Ralph B. "William Pepperell Montague and the New Realists." In "William Pepperell Montague." *Journal of Philosophy* 51 (21) (October 14, 1954): 593–637.

Spaulding, Edward G. *The New Rationalism*. New York: Holt, 1918.

WORKS ON NEW REALISM

Bowman, Lars. *Criticism and Construction in the Philosophy of the American New Realism*. Stockholm, 1955. One of the two extant studies devoted entirely to American New Realism, this work is mainly expository, using the tools of modem philosophical analysis in order to determine the central doctrines of New Realism. It includes one of the better short bibliographies.

Chisholm, Roderick M. *Realism and the Background of Phenomenology*. Glencoe, IL: Free Press, 1960. A collection of readings, with probably the best bibliography of the works of all major Realists and their critics.

Harlow, Victor. *A Bibliography and Genetic Study of American Realism*. Oklahoma City: Harlow, 1931. One of the best separate bibliographies of American Realism.

Hasan, Syed Zafarul. *Realism: An Attempt to Trace Its Origins and Development in Its Chief Representations*. Cambridge, U.K., 1928. An extensive treatment of New Realism (particularly of E. B. Holt) that includes a useful bibliography.

Hill, Thomas E. *Contemporary Theories of Knowledge*. New York: Ronald Press, 1961. Although limited to theories of knowledge, this work includes an extensive critical treatment of American New Realists, centering on Perry ("polemical"), Holt ("radical"), and Montague ("conservative").

James, William. "Does Consciousness Exist?" *Journal of Philosophy* 1 (18) (September 1, 1904). Reprinted in James's *Essays in Radical Empiricism*. New York: Longmans Green, 1938.

Kremer, René. *Le néo-réalisme américain*. Paris: Alcan, 1920. Devoted exclusively to New Realism.

Lapan, Arthur. "The Significance of James' 'Essay.'" PhD diss., Columbia University, New York, 1936.

Lovejoy, Arthur O. *The Revolt against Dualism: An Inquiry concerning the Existence of Ideas*. La Salle IL: Open Court, 1930. Lovejoy was probably the most persistent and incisive critic of New Realism.

Morris, Charles W. *Six Theories of Mind*. Chicago: University of Chicago Press, 1932. A discussion of the relationships of the various realisms, Pragmatism, and Objective Relativism.

Passmore, John. *A Hundred Years of Philosophy*. London: Duckworth, 1957. The chapter on New Realism is mainly concerned with English New Realism, most particularly Alexander's. Selected bibliography.

Piller, Christian. "The New Realism in Ethics." In *The Cambridge History of Philosophy 1870–1945*, edited by Thomas Baldwin. Cambridge, U.K.: Cambridge University Press, 2003.

Royce, Josiah. *The World and the Individual*. New York, 1912.

Schneider, Herbert W. *Sources of Contemporary Philosophical Realism in America*. Indianapolis: Bobbs-Merrill, 1964.

Werkmeister, W. H. *A History of Philosophical Ideas in America*. New York: Ronald Press, 1949. A good survey of the development of New Realism, including a detailed account of the polemical exchanges beginning with early New Realist statements in 1907.

OTHER RECOMMENDED TITLES

Almeder, Robert. *Blind Realism*. Lanham, MD: Rowman & Littlefield, 1992.

Alston, William. *A Realist Conception of Truth*. Ithaca, NY: Cornell University Press, 1996.

Baldwin, Thomas. "Ethical Non-Naturalism" In *Exercises in Analysis*, edited by Ian Hacking. Cambridge, U.K.: Cambridge University Press, 1985.

Devitt, Michael. *Realism and Truth*. Oxford: Blackwell, 1984.

Moser, Paul K. "Beyond Realism and Idealism." *Philosophia* 23 (1994): 271–288.

O'Connor, David. *The Metaphysics of G. E. Moore*. Boston: Reidel, 1982.

O'Leary-Hawthorne, John. "Anti-Realism, Before and After Moore." *History of Philosophy Quarterly* 12 (1995): 443–467.

Putnam, Hilary. *The Many Faces of Realism*. La Salle, IL: Open Court, 1987.

Skolimowski, Henryk. *Polish Analytical Philosophy: A Survey and Comparison with British Analytical Philosophy*. New York: Humanities Press, 1967.

Soames, Scott. *Philosophical Analysis in the Twentieth Century.* Vol. 1: *The Dawn of Analysis*. Princeton, NJ: Princeton University Press, 2003.

Stroud, Barry. *The Significance of Philosophical Skepticism*. Oxford: Clarendon Press, 1984.

Thomas Robischon (1967)
Bibliography updated by Benjamin Fiedor (2005)

NEWTON, ISAAC
(1642–1727)

Isaac Newton formulated the theory of universal gravity, was an inventor of the calculus, and made major discoveries in optics. He has long been regarded as, perhaps, the greatest scientist and as one of the greatest mathematicians ever to have lived. More recently, philosophers have begun to appreciate the extent to which Newton's remarks on scientific method illuminate the seminal contribution he made, especially in his *Principia*, to the transformation of natural philosophy into the physical sciences as we know them today. We now know, also, that Newton put at least as much effort into alchemy and theology as he did into his celebrated contributions to mathematics and science.

LIFE

Newton entered Trinity College Cambridge in 1661. In what has come to be called his *annus mirabilis*, he spent much of 1665 and 1666 at his family home in Woolsthorp while the university was closed because of the plague. This time at home was part of an extraordinarily productive period of intense effort concentrated on mathematics and natural philosophy. The binomial theorem and the fundamentals of the calculus are among the important new results in mathematics he obtained during this period. In natural philosophy he developed mechanics, including an analysis of circular motion. During this period he, also, conducted optical experiments that led to his account of white light and colors. In 1667 Newton became a fellow of Trinity College at Cambridge University.

In 1669 he became Lucasian Professor of Mathematics, presumably through the recommendation of Isaac Barrow (1630–1677), the first Lucasian Professor. It was Barrow who, in late 1671, delivered the reflecting telescope Newton had designed and built to the Royal Society of London. This led to Newton's being offered a fellowship in the Royal Society and to the publication in the Society's *Philosophical Transactions* of his account of white light and colors in 1672. This paper occasioned considerable debate. In that debate Newton began to articulate what he called his "experimental philosophy," which sharply distinguishes experimentally established results from conjectured hypotheses. By the late 1670s Newton withdrew from correspondence in natural philosophy.

In late 1679 Robert Hooke (1635–1703), who had recently become secretary of the Royal Society, wrote to encourage Newton to resume his public participation in natural philosophy. In this letter he invited Newton to use his mathematical methods to determine the trajectory a body would follow under a combination of inertial motion and an inverse-square force directed toward a center. In August 1684 a visit by Edmund Halley (1656–1742), who later became the Astronomer Royal, convinced Newton of the importance of the relation he had established between elliptical orbits and inverse square centripetal forces. By November Newton had sent Halley a small but revolutionary treatise, *De Motu*. An extraordinarily intense effort by Newton transformed this small treatise into his masterpiece, the *Principia*. It was published in 1687. Halley, who appreciated the importance of what Newton had achieved, oversaw the printing and paid for it out of his own pocket.

In 1689 and again in 1701, Newton was elected to represent Cambridge University in Parliament. He was made warden of the mint for England in 1696. By 1698 he had successfully carried out a major recoinage for the English economy. In 1699 he became master of the mint. In 1699 Newton also became an associate member of the French Academy of Sciences. He resigned his professorship at Cambridge in 1701. In 1703 he became president of the Royal Society of London, a post that, along with that of master of the mint, he held until his death. He was knighted in 1705.

In 1704 Newton published the first edition of his *Opticks*. It included two earlier mathematical papers as supplements, one of which was his first publication on the calculus. Newton's long delay in publishing his work led to his priority dispute with Gottfried Wilhelm Leibniz (1646-1716) over the invention of the calculus. This dispute extended from the mid-1690s until after Leibniz's death and came to focus on differences over natural philosophy as well as the calculus priority claims.

The second edition of *Principia* was published in 1713, after four years of effort under the able guidance of its editor Roger Cotes (1682–1716). The third edition was published in 1726. Conspicuous ways in which these two differ from the first edition appear to be responses to objections by Christian Huygens (1629–1695), Leibniz, and others. Some claims that had been called *Hypotheses* at the beginning of Book 3 in the first edition became, with changes and additions, *Regulae Philosophandi*, and others, such as Kepler's area and 3/2 power rules, became *Phaenomena*. The famous General Scholium clarifying what Newton took to be the proper practice of natural philosophy was added at the end.

The Latin editions of the *Optics* in 1706 and 1717 included queries that shed further light on his "experimental philosophy," as does his attack on Leibniz in his "Account of the Book Entitled *Commercium Epistolicum*" published anonymously in 1715. It ends as follows: "And must Experimental Philosophy be exploded as *miraculous* and *absurd*, because it asserts nothing more than can be proved by experiments, and we cannot yet prove by Experiments that all the Phaenomena in Nature can be solved by meer Mechanical Causes?" (1715, p. 224).

THE EXPERIMENTAL PHILOSOPHY IN THE LIGHT AND COLORS DEBATE

Newton's response to Hooke in the debate over his light and colors paper is a good illustration of his experimental philosophy. In that paper Newton claimed that his experiments conclusively established that the phenomenon of the oblong shape of the image of sunlight shined through a round hole and refracted through a prism is caused by sunlight's being made up of rays that are refracted different amounts by the prism. (Newton's reflecting telescope was designed to avoid problems caused by such differential refraction by using mirrors instead of lenses.)

Hooke interpreted Newton as claiming that the experiments established a corpuscular theory of light and argued that his own wave hypothesis could account for the results equally well. Newton responded by pointing out that the hypothesis that light is a body was put forward only as a conjecture suggested by the experiments, and not as part of what he claimed to have been established by them.

> But I knew, that the *Properties*, which I declar'd of *Light*, were in some measure capable of being explicated not only by that, but by many other Mechanical *Hypotheses*. And therefore I chose to decline them all, and to speak of *Light* in *general* terms, considering it abstractly, as something or other propagated in every way in streight lines from luminous bodies, without determining, what that thing is (1958, pp. 118–119).

Newton went on to outline how Hooke's wave hypothesis, as well as several other mechanical hypotheses, could explain the properties of differential refraction of different kinds of light he had concluded from the experiments.

In other contributions to the debate, Newton outlined how, according to his experimental philosophy, diligently establishing properties of things by experiment is to take precedence over framing hypotheses to explain them. He also made clear that the propositions he regarded as conclusively established by experiment were, nevertheless, subject to correction based on detailed criticism of the experimental reasoning establishing them or on further experimental results challenging them.

MATHEMATICS

Newton's mathematical papers include substantial discoveries in algebra, pure and analytic geometry, as well as his extensive work on the calculus and infinite series. His results on converging series allowed mathematicians to treat such infinite series as legitimate alternative forms of the functions they represented. These results also provided the basis for his approach to the calculus. In 1669 Newton first allowed one of his manuscripts on the calculus to circulate.

The basic mathematics of the *Principia* is not the calculus but a new form of synthetic geometry incorporating limits. Newton's lemmas on first and last ratios, which open Book 1, show that this alternative geometrical approach can recover many of the basic elementary results of the calculus. The need to rely on geometrical figures, however, makes this approach less able to facilitate more complex calculations made accessible by algebraic manipulation in the symbolic calculus.

STUDIES IN ALCHEMY, THEOLOGY, AND CHRONOLOGY

Newton's alchemical work may well have contributed to a corpuscular theory of matter that may have informed his scientific thinking; however, like his conjectured corpuscular account of light, such a theory of matter was not something Newton claimed to have established.

His extensive notes on his alchemical work indicate a number of elaborate chemical experiments carried out from the mid-1670s until 1693. These display Newton's great discipline as an experimenter. The reported results, however, appear to include nothing that would have altered the course of chemistry had they become public at the time.

Newton first became preoccupied with theology in the early 1670s, probably in response to the requirement that he accept ordination to retain his Trinity fellowship. (He was granted a dispensation in 1675.) By 1673 he had rejected the doctrine of the Trinity and concluded that Christianity had become a false religion through a corruption of the scriptures in the fourth and fifth centuries. He returned to these studies and to work on chronology

and prophecies in subsequent decades, especially in the last years of his life. During his lifetime he conveyed his radical views to only a few. But, two such manuscripts were published within a few years of his death.

Recent investigations of the alchemical and theological writings suggest that Newton's natural philosophy was to be part of a larger investigation that would look through nature to see God. This may have helped him to free himself from the restraints of the mechanical philosophy. Newton's intense religious faith was no impediment, and may well have aided, his extraordinarily successful applications of his experimental philosophy in pursuit of empirically establishing scientific knowledge. Moreover, Newton's efforts at scientific understanding of nature did not prevent his efforts to inform his faith by the study of scripture.

SPACE, TIME, AND THE LAWS OF MOTION

Newton's distinction between absolute (or true) and relative (or apparent) motion are based on his laws of motion, which he described as "accepted by mathematicians and confirmed by experiments of many kinds" ([1687] 1999, p. 424). His distinctions between absolute and relative space and time, which have been such salient targets of criticism by philosophers, are mostly designed to accommodate this primary distinction between true and merely relative motions. Newton was aware of the empirical difficulties raised by such distinctions: "It is certainly very difficult to find out the true motions of individual bodies and actually to differentiate them from apparent motions, because the parts of that immovable space in which bodies move make no impression on the senses" (p. 414).

The *Principia*'s title, *Mathematical Principles of Natural Philosophy*, refers to the propositions of Books 1 and 2 that Newton demonstrated from his laws of motion. These provide his resources for addressing this difficulty: "But in what follows, a fuller explanation will be given of how to determine true motions from their causes, effects, and apparent differences, and conversely, of how to determine from motions whether true or apparent, their causes and effects. For this was the purpose for which I composed the following treatise" (p. 415). In Book 3 Newton shows how the calculation of centripetal forces and masses of central bodies from orbital motions around them can determine the center of mass of the planetary system. This calculation picks out the sun-centered Keplerian system as approximately true and the corresponding earth-centered Tychonic system as wildly inconsistent with the measured masses.

Such inconsistencies among the measured forces and masses indicate a failure to be dealing with true motions. For Newton, the adequacy of his appeal to absolute space, time, and motion was an empirical issue to be decided by the long term development and application of a science of motion.

INFERENCES FROM PHENOMENA AND RULES OF NATURAL PHILOSOPHY

The propositions of Books 1 and 2 are powerful resources for establishing conclusions about forces from phenomena of motion. For example, propositions 1 and 2 together establish that Kepler's area rule holds if and only if the force acting on the moving body is centripetal. A corollary adds that the rate at which areas are swept out be radii from the center increases just in case the net force is off-center in the direction of motion, and decreases just in case it is off-center in the opposite direction. These systematic dependencies make the constancy of the areal rate measure the centripetal direction of the force. Similar systematic dependencies are involved in the inferences to the inverse-square variation of orbital centripetal forces from Kepler's 3/2 power rule and from the absence of orbital precession.

Newton was not the first to exploit such theoretical dependencies to draw inferences from phenomena. Huygens had used his laws of pendulums to measure the acceleration of gravity from the lengths and periods of pendulums. But, Newton turned the technique into a general way of using theory mediated measurements to do empirical science.

The rules of reasoning strengthen the inferences that can be drawn from measurements by phenomena.(See Scientific Method) The first two rules, for example, endorse the inference identifying the force holding the moon in orbit with terrestrial gravity on the basis of the moon-test, which shows that the length of a seconds pendulum at the surface of the earth and the centripetal acceleration of the moon's orbit can count as agreeing measurements of a single earth centered inverse-square acceleration field.

The third rule supports the inference that all bodies gravitate toward each planet with weights proportional to their masses. Newton argues that terrestrial pendulum experiments and the moon-test show this for gravitation toward the earth. Similarly, the harmonic laws for orbits

about them show this for gravitation toward Saturn, Jupiter, and the sun. In addition, the agreement between the accelerations of Jupiter and its satellites toward the sun, as well as between those of Saturn and its satellites and those of the earth and its moon toward the sun also show this for weight toward the sun. All these count as phenomena giving agreeing measurements of the equality of the ratios of weight to mass for all bodies at any equal distances from the sun or any planet.

The fourth rule authorizes the practice of treating propositions appropriately supported by reasoning from phenomena as either "exactly or very nearly true notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions" (p. 796). It was added in the third edition to justify treating universal gravity as an established scientific fact, notwithstanding complaints that it was unintelligible in the absence of an explanation of how it results from mechanical action by contact. This rule and the related discussion of hypotheses in the General Scholium most distinguish Newton's experimental philosophy from the mechanical philosophy of his critics.

GRAVITY AS A UNIVERSAL FORCE OF INTERACTION

The systematic dependencies via which the basic inverse-square forces are measured by Keplerian phenomena are one-body idealizations. Universal gravity entails interactions among bodies, producing perturbations that require corrections to the Keplerian phenomena. Such corrections can count as higher-order phenomena that carry information that can be exploited to develop successively more accurate approximations.

The *Principia* includes a successful treatment of two-body interactions and some limited results on three-body interactions including Newton's account of the variational inequality in the lunar orbit. Applications of calculus facilitated by the use of Leibniz's notation by such figures as Leonard Euler (1707–1783), Jean Le Rond d'Alembert (1717–1783), and Alexis-Claude Clairaut (1713–1765) led to successful Newtonian treatments of more complex interactions. By the mid-1700s such successes in the treatments of the shape of the earth, the precession of the equinoxes, the lunar precession and motions of comets had led to the virtual abandonment of vortex theories as serious rivals. By the end of that century, the monumental treatise on celestial mechanics by Pierre Simon de Laplace (1749–1835), with his successful treatment of the long recalcitrant great inequality in Jupiter-Saturn motions as a periodic perturbation, led to general acceptance of a Newtonian metaphysics of bodies interacting under deterministic laws.

Newtonian treatments of perturbations do more than provide the required corrections to Keplerian phenomena. They also show that Newton's original measurements of inverse-square centripetal forces continue to hold to high approximation in the presence of perturbations. Interactions with other bodies account for the precessions of all the planets except Mercury. The zero residuals in these precessions are agreeing measurements of the inverse-square variation of gravity toward the sun.

Even in the case of Mercury the famous forty-three seconds of arc per century residual in its precession yields -2.00000016 as the measure of the exponent, instead of the exact -2 measured for the other planets. That such a small discrepancy came to be a problem at all testifies to the extraordinary high level to which Newton's theory of gravity had realized a standard of empirical success. On this standard of empirical success, a theory succeeds by having its parameters be accurately measured by the phenomena it purports to explain.

In 1915, Einstein discovered that his theory of general relativity explains the missing forty-three seconds. The success of this explanation depends on the capacity of general relativity to also account for the additional precession of about 530 seconds per century explained by Newtonian perturbations of Mercury's orbit. This requires that Newton's theory count as an appropriate approximation for explaining that part of the phenomenon of Mercury's orbital precession.

Einstein's great excitement over this discovery is appropriate because it showed that his theory of general relativity did better than Newton's theory of universal gravitation by Newton's own standard of empirical success. There was and is no need to appeal to additional or different standards to count general relativity as better supported. The subsequent development of testing frameworks for general relativity continues to be guided by the same standard. Newton's methodology of successive approximations supported by the empirical success of theory mediated measurement accommodates, even, the radical conceptual transformation from Newton's metaphysics of bodies under forces of interaction to Einstein's conception of gravity as given by the geodesic structure of curved space-time.

See also Classical Mechanics, Philosophy of; Space.

Bibliography

WORKS BY NEWTON

"An Account of the Book Entituled *Comerciumm Epistolicum*." *Philosophical Transactions* 29 (342) (1715): 173–224.

Observations upon the Prophecies of Daniel, and the Apocalypse of St. John. In *Sir Isaac Newton's Daniel and the Apocalypse: With an Introductory Study … of Unbelief, of Miracles and Prophecy*, by William Whitla. London: J. Murray, 1922. Reprinted in *Sir Isaac Newton's Daniel and the Apocalypse with an Introductory Study of Unbelief of Miracles and Prophecy*, London, 1733.

Opticks, or, A Treatise of the Reflections, Refractions, Inflections & Colours of Light. Based on the 4th. ed., 1730. New York: Dover, 1952.

Isaac Newton's Papers and Letters on Natural Philosophy and Related Documents, edited by I. B. Cohen and assisted by R. E. Schofield. Cambridge, MA: Harvard University Press, 1958. 2nd, rev. ed., 1978. Contains the publications in the dispute on light and colors as they appeared in the *Philosophical Transactions* of the Royal Society.

The Correspondence of Isaac Newton, edited by H. W. Turnbull, A. Scott, A. R. Hall, and L. Tilling. 7 vols. Cambridge, U.K.: Cambridge University Press, 1969–1977.

Unpublished Scientific Papers of Isaac Newton: A Selection from the Portsmouth Collection in the University Library, Cambridge, edited by A. R. Hall and M. B. Hall. Cambridge, U.K.: Cambridge University Press, 1962. Reprint, 1978.

The Background to Newton's Principia:. A Study of Newton's Dynamical Researches in the Years 1664–84, edited by J. W. Herivel. Oxford, U.K.: Clarendon Press, 1965.

The Mathematical Papers of Isaac Newton, edited by D. T. Whiteside. 8 vols. Cambridge, U.K.: Cambridge University Press, 1967–1981.

Isaac Newton's Philosophiae Naturalis Principia Mathematica. 3rd ed., 1726, edited by A. Koyré and I. B. Cohen. 2 vols. Cambridge, MA: Harvard University Press, 1972. Contains variant readings.

Certain Philosophical Questions: Newton's Trinity Notebook, edited by J. E. McGuire and M. Tamny. Cambridge, U.K.: Cambridge University Press, 1983. Reprint, 1985.

The Optical Papers of Isaac Newton. Vol. 1, edited by A. E. Shapiro. Cambridge, U.K.: Cambridge University Press, 1984.

Isaac Newton's Mathematical Principles of Natural Philosophy. Translated by I. B. Cohen, and A. Whitman. Los Angeles: University of California Press, 1999.

WORKS ABOUT NEWTON

General

Cohen, I. B. "Isaac Newton." In *Dictionary of Scientific Biography*, vol. 10, 42–103. New York: Scribners, 1974.

Cohen, I. B. *The Newtonian Revolution*. Cambridge, U.K.: Cambridge University Press, 1980. 2nd rev. ed., 1995.

Cohen, I. B., and G. E. Smith G.E. *The Cambridge Companion to Newton*. Cambridge, U.K.: Cambridge University Press, 2002.

Gjertsen, D. *The Newton Handbook*. London: Routledge & Kegan Paul, 1986. Contains a complete listing of Newton's work and an extensive bibliography of secondary sources.

Whiteside, D. T. "The Prehistory of the *Principia* from 1664–1686." In *Notes and Records of The Royal Society of London*, vol. 45, 11–61. London: Royal Society of London, 1991.

Wilson, C. "The Newtonian Achievement in Astronomy." In *Planetary Astronomy from the Renaissance to the Rise of Astrophysics*, edited by R. Taton and C. Wilson. Cambridge, U.K.: Cambridge University Press, 1989.

Life

Gleick, J. *Isaac Newton*. New York: Pantheon Books, 2003.

Hall, A. R. *Isaac Newton: Adventurer in Thought*. Oxford, U.K.: Oxford University Press, 1992.

Westfall, R. S. *Never at Rest: A Biography of Isaac Newton* Cambridge, U.K.: Cambridge University Press, 1980. An abridged version, *The Life of Isaac Newton*, was published in 1993.

Further References

Bricker, P., and R. I. G. Hughes, eds. *Philosophical Perspectives on Newtonian Science*. Cambridge, MA: MIT Press, 1990. Among the many excellent collections inspired by the three-hundredth anniversary of the first edition of *Principia*, this one is especially focused on philosophy.

Earman, J. *World Enough and Space-Time Absolute verses Relational Theories of Space and Time*. Cambridge, MA: MIT Press, 1989.

Harper, W. L. "Howard Stein on Isaac Newton: Beyond Hypotheses?" In *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, edited by David Malament, Chicago: Open Court, 2002.

Koyré, A. *Newtonian Studies*. Chicago: University of Chicago Press, 1968.

Palter, R., ed. *The Annus Mirabilis of Sir Isaac Newton, 1666–1966*. Cambridge, MA: MIT Press, 1971. This classic collection contains papers still relevant to most topics.

Smith G. " From the Phenomenon of the Ellipse to an Inverse-Square Force: Why Not?" In *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, edited by David Malament, Chicago: Open Court, 2002.

Stein, H. "From the Phenomena of Motions to the Forces of Nature: Hypothesis or Deduction?" *PSA 90* 2 (1991): 209–222.

Theerman, Paul, and Adele F. Seeff, eds. *Action and Reaction: Proceedings of a Symposium to Commemorate the Tercentenary of Netwton's Principia*. Newark: University of Delaware Press, 1993.

William L. Harper (2005)

NICHOLAS OF CUSA

(1401–1464)

The theologian, philosopher, and mathematician Nicholas of Cusa, also known as Nicholas Kryfts or Krebs, was born at Kues on the Moselle River between Trier and Koblenz. After attending the school of the Brothers of

the Common Life in Deventer, Holland, he studied philosophy at Heidelberg (1416), canon law at Padua (1417–1423), and theology at Cologne (1425). Nicholas received a doctorate in canon law in 1423. About 1426 he gave legal assistance to Cardinal Orsini, papal legate to Germany. At about the same time began his lifelong interest in collecting classical and medieval manuscripts. Among his notable discoveries were twelve lost comedies of Plautus. He took an active part in the Council of Basel, first as a lawyer of Count von Manderscheid and later as a member of the deputation *De Fide*. Nicholas's *De Concordantia Catholica*, a vast program for reform of the church and the empire, supported the conciliar theory of the supremacy of the council over the pope. Later, disillusioned by the council's failure to reform the church, he abandoned the conciliar theory and supported the papal cause.

Nicholas carried out several missions for the pope in an effort to unify and reform the church. He was a member of the commission sent to Constantinople to negotiate with the Eastern church for reunion with Rome, which was temporarily effected at the Council of Florence (1439). In 1450 Nicholas was sent to Germany as a legate to carry out church reforms. He was created a cardinal in 1448 and appointed bishop of Brixen (Bressanone) in 1450. He died in Todi, Umbria.

KNOWLEDGE

According to Nicholas, a man is wise only if he is aware of the limits of the mind in knowing the truth. Knowledge is learned ignorance (*docta ignorantia*). Endowed with a natural desire for truth, humans seek it through rational inquiry, which is a movement of the reason from something presupposed as certain to a conclusion that is still in doubt. Reasoning involves a relating or comparing of conclusion with premises. The greater the distance between them, the more difficult and uncertain is the conclusion. If the distance is infinite, the mind never reaches its goal, for there is no relation or proportion between the finite and infinite. Hence, the mind cannot know the infinite. The infinite is an absolute, and the absolute cannot be known by means of relations or comparisons.

Accordingly, the mind cannot comprehend the infinite God. By rational investigation we can draw ever nearer to him but cannot reach him. The case is the same with any truth, for every truth is an absolute, not admitting of degrees. Since reason proceeds by steps, relating conclusion to premises, it is relational and hence never arrives at absolute truth. According to Nicholas, "our intellect, which is not the truth, never grasps the truth with such precision that it could not be comprehended with infinitely greater precision" (*De Docta Ignorantia* I, 3). As a polygon inscribed in a circle increases in number of sides but never becomes a circle, so the mind approximates to truth but never coincides with it.

Thus, knowledge at best is conjecture (*coniectura*). This is no mere guess or supposition that may or may not be true; it is an assertion that is true as far as it goes, although it does not completely measure up to its object. Reason is like an eye that looks at a face from different and even from opposite positions. Each view of the face is true, but it is partial and relative. No one view, nor all taken together, coincides with the face. Similarly, human reason knows a simple and indivisible truth piecemeal and through opposing views, with the result that it never adequately measures up to it.

The weakness of human reason was evident to Nicholas because its primary rule is the principle of noncontradiction, which states that contradictories cannot be simultaneously true of the same object. He insisted that there is a "coincidence of opposites" (*coincidentia oppositorum*) in reality, especially in the infinite God. He criticized the Aristotelians for insisting on the principle of noncontradiction and stubbornly refusing to admit the compatibility of contradictories in reality. It takes almost a miracle, he complained, to get them to admit this; and yet without this admission the ascent of mystical theology is impossible.

Nicholas preferred the Neoplatonists to the Aristotelian philosophers because they recognized in humans a power of knowing superior to reason which they called intellect (*intellectus*). This was a faculty of intuition or intelligence by which we rise above the principle of noncontradiction and see the unity and coincidence of opposites in reality. He found this faculty best described and most fruitfully cultivated by the Christian Neoplatonists, especially St. Augustine, Boethius, Pseudo-Dionysius, St. Anselm, the School of Chartres, St. Bonaventure, and Meister Eckhart. Following their tradition, he constantly strove to see unity and simplicity where the Aristotelians could see only plurality and contradiction. He frequently expressed his views in symbols and analogies, often mathematical in character, because the rational language of demonstration is appropriate to the processes of reason but not to the simple views of the intellect.

GOD

Nicholas was most concerned with showing the coincidence of opposites in God. God is the absolute maximum

or infinite being, in the sense that he has the fullness of perfection. There is nothing outside him to oppose him or to limit him. He is the all. He is also the maximum, but not in the sense of the supreme degree in a series. As infinite being he does not enter into relation or proportion with finite beings. As the absolute, he excludes all degrees. If we say he is the maximum, we can also say he is the minimum. He is at once all extremes, the absolute maximum as well as the absolute minimum. In short, in God, the infinite being, all opposition is reconciled in perfect unity.

The coincidence of the maximum and minimum in infinity is illustrated by mathematical figures. For example, imagine a circle with a finite diameter. As the size of the circle is increased, the curvature of the circumference decreases. When the diameter is infinite, the circumference is an absolutely straight line. Thus, in infinity the maximum of straightness is identical with the minimum of curvature. Or, to put it another way, an infinite circle is identical with a straight line.

Nicholas offered several a priori proofs for the existence of the absolute maximum, or God. The first argued that the finite is inconceivable without the infinite. What is finite and limited has a beginning and an end, so that there must be a being to which it owes its existence and in which it will have its end. This being is either finite or infinite. If it is finite, then it has its beginning and end in another being. This leads either to an infinite series of actually existing finite beings, which is impossible, or to an infinite being which is the beginning and end of all finite beings. Consequently, it is absolutely necessary that there be an infinite being, or absolute maximum.

The second proof argued that the absolute truth about the absolute maximum can be stated in three propositions: It either is or is not. It is and it is not. It neither is nor is not. These exhaust all the possibilities, so that one of them must be the absolute truth. Hence there is an absolute truth, and this is what is meant by the absolute maximum.

As the absolute maximum, God contains all things; he is their "enfolding" (*complicatio*). He is also their "unfolding" (*explicatio*) because they come forth from him. Creatures add nothing to the divine reality; they are simply limited and partial appearances of it. As a face reproduces itself more or less perfectly in a number of mirrors, so God reflects himself in various ways in his creatures. In this case, however, there are no mirrors.

God transcends the universe but is also immanent in it, as a face is present in its mirrored images. Each creature is also present in every other, as each image exists in every other. Thus, as Anaxagoras said, everything is in everything else. Gottfried Wilhelm Leibniz recalled this doctrine of Nicholas's in his *Monadology* when showing that each monad mirrors every other.

Like all medieval Platonists, Nicholas upheld the reality of universal forms. According to him, the most universal of all created forms is the form of the universe, called the Soul of the World. This form embraces in its unity all lower forms, such as those of genera and species. These lower forms are "contractions" of the form of the universe; they are the universe existing in a limited way. They exist in the universe, and it in turn exists in a limited way in them. Individuals are further contractions of universal forms—for example, Socrates is a contraction of the form of humanity. The universe as a whole is a contraction of the infinite God. Thus, all things exist in a unified manner in the universe, and the universe in turn exists in the unity of God. Oppositions and contradictions that appear on the level of individuals and lower universal forms are reconciled in the unity of the universe and ultimately in the unity of God.

COSMOLOGY

Since the universe mirrors God, it too must be a maximum—not the absolute maximum, to be sure, but the relative maximum, for it contains everything that exists except God. Nicholas denied that the universe is positively infinite; only God, in his view, could be described in these terms. But he asserted that the universe has no circumference and consequently that it is boundless or undetermined—a revolutionary notion in cosmology. (See Alexandre Koyré, *From the Closed World to the Infinite Universe*, Baltimore, 1957.) Just as the universe has no circumference, said Nicholas, so it has no fixed center. The earth is not at the center of the universe, nor is it absolutely at rest. Like everything else it moves in space with a motion that is not absolute but is relative to the observer.

Nicholas of Cusa's cosmology in some respects broke with the Ptolemaic and Aristotelian cosmological views of the Middle Ages and anticipated those of modern times. He was above all concerned with denying the absolute oppositions in the world of Ptolemy and Aristotle. In Nicholas's world there was no center opposed to its circumference, no maximum movement of the spheres opposed to the fixity of Earth, no movement of bodies in absolutely opposed directions, such as up and down. Nicholas also denied that the heavenly bodies are com-

posed of a substance different from that of sublunar bodies.

Nicholas extended his principle of the coincidence of opposites to religion. In his irenical work *On the Peace of Faith*, while maintaining the superiority of Christianity over other religions, he tried to reconcile their differences. Beneath their oppositions and contradictions he believed there is a fundamental unity and harmony, which, when it is recognized by all, will be the basis of universal peace.

In a century of social, political, and religious unrest, Nicholas revitalized Neoplatonism as the most effective answer to the needs of his time. His thought was firmly rooted in the philosophy of Proclus and Christian medieval Neoplatonism and was opposed to the Aristotelianism that had prevailed in western Europe since the thirteenth century. It was also highly original and expressed in a language abounding in symbolism and paradox. Nicholas of Cusa had many of the traits of the Renaissance person: love of classical antiquity, all-encompassing curiosity, optimism, cultivation of literary style, critical spirit, preoccupation with the individual, and love of mathematics and science. His works were widely read for several centuries, and they influenced the philosophy of the Renaissance and of early modern times.

See also Anaxagoras of Clazomenae; Anselm, St.; Aristotelianism; Aristotle; Augustine, St.; Boethius, Anicius Manlius Severinus; Bonaventure, St.; Chartres, School of; Eckhart, Meister; Infinity in Theology and Metaphysics; Leibniz, Gottfried Wilhelm; Medieval Philosophy; Neoplatonism; Platonism and the Platonic Tradition; Proclus; Pseudo-Dionysius; Renaissance; Socrates; Universals, A Historical Survey.

Bibliography

Nicholas's religious-political works are *De Concordantia Catholica* (1433–1434); *De Pace Fidei* (1453); and *Cribratio Alchorani* (1461).

His theological-philosophical works are *De Docta Ignorantia* and *De Coniecturis* (1440); *De Deo Abscondito* (1444); *De Quaerendo Deum* (1445); *De Genesi* (1447); *Apologia Doctae Ignorantiae* (1449); *Idiotae Libri* (containing *De Sapientia, De Mente*, and *De Staticis Experimentis*; 1450); *De Visione Dei* (1453); *De Beryllo* (1458); *De Possest* (1460); *Tetralogus de Non Aliud* (1462); *De Venatione Sapientiae* (1463); *De Ludo Globi* (1463); and *De Apice Theoriae* (1463).

His mathematical-scientific works are *De Staticis Experimentis* (1450); *De Transmutationibus Geometricis* (1450); *De Mathematicis Complementis* (1453); and *De Mathematica Perfectione* (1458).

Editions of Nicholas's works are *Opera* (Basel, 1565); *Opera*, 3 vols. (Paris, 1514; reprinted Frankfurt am Main, 1962); *Nicolaus von Cues, Texte seiner philosophischen Schriften*, edited by A. Petzelt, Vol. 1 (Stuttgart: Kohlhammer, 1949); *Opera Omnia*, 14 vols. (Leipzig and Hamburg, 1932–1959); and *De Pace Fidei*, edited by R. Klibansky and H. Bascour (London: Warburg Institute, 1956).

Translations include *The Vision of God*, translated by E. G. Salter (New York: Dutton, 1928); *The Idiot*, translated by W. R. Dennes (San Francisco, 1940); *Of Learned Ignorance*, translated by G. Heron (London: Routledge and K. Paul, 1954); and *Unity and Reform; Selected Writings of Nicholas de Cusa*, edited by J. P. Dolan (Notre Dame, IN: University of Notre Dame Press, 1962).

Literature on Nicholas includes P. Duhem, "Nicholas de Cues et Léonard de Vinci." in *Études sur Léonard de Vinci*, Vol. II (Paris: Hermann, 1909); E. Vansteenberghe, *Le Cardinal Nicolas de Cues* (Paris: Champion, 1920); P. Rotta, *Il cardinale Nicolò di Cusa, la vita ed il pensiero* (Milan: Società editrice, "Vita e pensiero," 1928); and *Nicolò Cusano* (Milan: Bocca, 1942); A. Posch, *Die "Concordantia Catholica" des Nikolaus von Cusa* (Paderborn, 1930); H. Bett, *Nicholas of Cusa* (London: Methuen, 1932); M. de Gandillac, *La philosophie de Nicolas de Cues* (Paris: Aubier, Éditions Montaigne, 1941); J. Koch, *Nikolaus von Cues und seine Umwelt* (Heidelberg: Carl Winter, 1948); P. Mennicken, *Nikolaus von Kues* (Trier: Cusanus-Verl, 1950); Étienne Gilson, *Les métamorphoses de la cité de Dieu* (Paris: J. Vrin, 1952); K. H. Volkmann-Schluck, *Nicolaus Cusanus* (Frankfurt am Main: Klostermann, 1957); E. Meuthen, *Die letzten Jahre des Nikolaus von Kues* (Cologne: Westdeutscher, 1958); G. Santinello, *Il pensiero di Nicolò Cusano nella sua prospettiva estetica* (Padua: Liviana, 1958); G. Heinz-Mohr, *Unitas Christiana* (Trier, 1958); E. Zellinger, *Cusanus-Konkordanz* (Munich: Hueber, 1960); P. E. Sigmund, *Nicholas of Cusa and Medieval Political Thought* (Cambridge, MA: Harvard University Press, 1963); and Ernst Cassirer, *Individuum und Kosmos in der Philosophie der Renaissance* (Leipzig: Teubner, 1927), translated by Mario Domandi as *The Individual and the Cosmos in Renaissance Philosophy* (Oxford: Blackwell, 1963).

See also Louis Dupre and Nancy Hudson, "Nicholas of Cusa," in *A Companion to Philosophy in the Middle Ages*, edited by Jorge J. E. Gracia (Malden, MA: Blackwell, 2003); Jasper Hopkins, "Nicholas of Cusa (1401–1464): First Modern Philosopher?" *Midwest Studies in Philosophy* (26[2002]: 13–29; John Longeway, "Nicolas of Cusa and Man's Knowledge of God," *Philosophy Research Archives* (13[1987–1988]: 289–313.

Armand A. Maurer (1967)
Bibliography updated by Tamra Frei (2005)

NICHOLAS OF ORESME

See *Oresme, Nicholas*

NICOLAI, CHRISTIAN FRIEDRICH
(1733–1811)

Christian Friedrich Nicolai, a German publisher, editor, and author, was born in Berlin and studied there and at a Pietist institution in Halle, but he never attended a university. Nicolai spent three years as a business apprentice in Frankfurt an der Oder. Upon his father's death in 1752, he took over the family bookstore, managing it—except for a short period—until his death and expanding it into a very successful and lucrative publishing house. He became a close friend of G. E. Lessing and of Moses Mendelssohn, and was active in Berlin intellectual life. He edited the *Bibliothek der schönen Wissenschaften und freien Künste* (Library of aesthetics and fine arts) from 1757 to 1758, the *Literaturbriefe* (Letters on literature) from 1759 to 1765, and the *Allgemeine deutsche Bibliothek* (Universal German library) from 1765 on. The last-mentioned journal became the most famous German literary review of its time and was widely influential in theology as well.

Nicolai's own works, like those of many Enlightenment figures, were largely higher journalism consisting mainly in forceful and lively attacks on contemporary intellectual and literary personalities and trends. His *Briefe, den jetzigen Zustand der Schönen Wissenschaften betreffend* (Letters on the state of the arts; Berlin, 1755) were directed against the influential literary critic J. C. Gottsched. His philosophical novel *Sebaldus Nothanker* (3 vols., Berlin, 1773–1776) was an attack on certain reactionary circles in Halle. In various articles in his journals he attacked J. G. Hamann, Johann Caspar Lavater, Christian Garve, and others. He quarreled with J. G. Herder and F. H. Jacobi. The novels *Daniel Säuberlich* (Berlin, 1777–1778) and *Die Freunden des jungen Werthers* (Berlin, 1775) were parodies of Johann Wolfgang von Goethe, Johann Gottfried Herder, G. A. Bürger (author of the ballad *Lenore*), and the *Sturm und Drang.* He attacked Catholicism as a source of superstition and Jesuitism; and, although he was himself a member of the Order of the Enlightened (*Illuminaten*) and of the Freemasons, he accused both of being secret instruments of the Jesuits (which resulted in his forced resignation). In the philosophical novels *Geschichte eines dicken Mannes* (The story of a fat man; 2 vols., Berlin, 1794) and *Sempronius Gundibert* (Berlin, 1798) and in other works, he accused Immanuel Kant and his school and Johann Gottlieb Fichte of being crypto-Catholics. His *Vertraute Briefe von Adelheid B. an ihre Freundin Julie S.* (Confidential letters from Adelaide B. to her friend Julie S.; Berlin, 1799) was directed against Friedrich Schleiermacher.

Nicolai wrote many other works, notably a large work devoted to the economic, cultural, social, and religious life in Germany and Switzerland, *Beschreibung einer Reise durch Deutschland und die Schweiz im Jahre 1781* (Description of a journey through Germany and Switzerland in 1781; 12 vols., Berlin, 1783–1796). Although Nicolai was awarded an honorary doctorate by the Helmstedt Theological Seminary in 1799 and was made a corresponding member of the Academy of St. Petersburg in 1804, his hostility toward the most influential persons of his time and his lack of understanding of the new critical philosophy and of romanticism led to a negative evaluation of his work by his leading contemporaries and by the following generation.

Nevertheless, Nicolai was one of the most typical representatives of "popular philosophy." Basing his theories on common sense, he avoided abstract thought and complex speculation and favored useful and easy knowledge. He opposed orthodoxy, intolerance, enthusiasm, mysticism, and secret machinations. He attacked the scholastic Wolffian philosophy; the newer critical and idealistic philosophies; Protestantism, both orthodox and mystical, and Catholicism; secret societies; Gottsched's classicism in literature as well as the glorification of the peasant by J. H. Voss and Bürger; *Sturm und Drang*; and early romanticism. He considered them all to be reactionary and pernicious, and his writings were full of misunderstandings, misrepresentations, and exaggerations.

His religious views incorporated his rejection of intellectualism, dogmatism, and mysticism. He held that religion and science should not be confused. Orthodox religion corrupted morality and tended toward an obnoxious hierarchical system. He denied original sin and eternal damnation and accepted the doctrines of free will and of the immortality of the soul. Religion should be based on the individual conscience and not on revelation—on common sense and not on enthusiasm.

According to Nicolai, religion and morality are not the same. Morality is based on social sense and experience; religion is a feeling for God's goodness and providence as mirrored in the goodness and beauty of the Creation. Although Nicolai was a deist himself, he did not believe that a purely natural religion would suffice for the common people, and therefore he refused to reject publicly the Christian tradition.

Nicolai was influenced in aesthetics by the classicists Nicolas Boileau and Jean Baptiste Dubos and by the Swiss

critics J. J. Bodmer, J. J. Breitinger, and J. G. Sulzer. He tried to find a middle ground between the classical doctrine of the imitation of nature and the newer stress on the imagination. He opposed the classical ideal of literature as deduced from a set of rules, the sentimental school of literature, and the *Sturm und Drang* emphasis on intuitive genius. He held that poetry should be simple and reasonable and designed chiefly for moral improvement.

See also Aesthetics, History of; Boileau, Nicolas; Common Sense; DuBos, Abbe Jean Baptiste; Enlightenment; Fichte, Johann Gottlieb; Garve, Christian; Goethe, Johann Wolfgang von; Gottsched, Johann Christoph; Hamann, Johann Georg; Herder, Johann Gottfried; Jacobi, Friedrich Heinrich; Kant, Immanuel; Lavater, Johann Kaspar; Lessing, Gotthold Ephraim; Mendelssohn, Moses; Religion and Morality; Schleiermacher, Friedrich Daniel Ernst; Sulzer, Johann Georg.

Bibliography

ADDITIONAL WORKS BY NICOLAI

Ueber meine gelehrte Bildung. Berlin, 1799.

Philosophische Abhandlungen. Berlin, 1808.

WORKS ON NICOLAI

Aner, K. *Der Aufklärer Friedrich Nicolai.* Giessen: A. Töpelmann, 1912.

Fichte, J. G. *Fr. Nicolais Leben und sonderbare Meinungen.* Tübingen, 1801.

Meyer, Friedrich. *Friedrich Nicolai.* Leipzig: Bücherstube, 1938.

Ost, G. *Fr. Nicolais Allgemeine Deutsche Bibliothek.* Berlin, 1928.

Philips, F. C. A. *Nicolais literarische Bestrebungen.* The Hague, 1926.

Sommerfeld, M. *Nicolai und der Sturm und Drang.* Halle, 1921.

Strauss, Walter. *Nicolai und die kritische Philosophie.* Stuttgart, 1927.

Giorgio Tonelli (1967)

NICOLAS OF AUTRECOURT

(c. 1300–after 1350)

Nicolas of Autrecourt, also called Nicolaus de Ultracuria, was a leading anti-Aristotelian philosopher of the fourteenth century. The condemnation of extreme Aristotelianism at Paris in 1277 was probably responsible for the critical tendencies in many fourteenth-century philosophers and theologians. An extreme form of this critical tendency is to be found in the writings and lectures of Nicolas of Autrecourt. He was at the Sorbonne as early as 1328, lectured on the *Sentences* at Paris, and in 1340 was summoned by the Roman Curia to answer charges of heresy and error. His trial was interrupted when Pope Benedict XII died, and was resumed under Pope Clement VI by Cardinal Curty. In 1346 the trial was concluded, Nicolas was forced to recant many of his published statements, his works were publicly burned, and he was declared unworthy of advancement and unworthy to continue teaching. We last hear of him as a deacon at the cathedral of Metz in 1350.

His literary remains consist of (1) two complete letters to the Franciscan Bernard of Arezzo, a reply to a certain Giles (whose letter to Nicolas is also extant), and the fragments of seven other letters to Bernard of Arezzo; (2) a theological discussion concerning the increase of cognitive powers; and (3) the "universal tractate of Master Nicolas of Autrecourt for seeing whether the statements of the Peripatetics are demonstrative" (usually called *Exigit Ordo Executionis* from its *incipit*), which survives in a single manuscript that breaks off toward the end.

The continuing research on fourteenth-century thought will probably show that many other Schoolmen of the period expressed doctrines similar to those of Nicolas. In fact, similar doctrines have already been found in Robert Holkot and John of Mirecourt on epistemological issues, and in Henry of Harclay, Gerard Odo, and some others on atomism and the constitution of the continuum. Nevertheless, there is some reason to attribute to Nicolas a considerable measure of originality and of persistent thought. For one thing, his contemporary John of Mirecourt attributes to Nicolas the proof that causal connections cannot be demonstrated. This may mean merely that Mirecourt was making an acknowledgement to a colleague and was unaware that similar doctrines were taught at Oxford. But there must be some significance in the fact that Nicolas was singled out for attack by the decrees of the Paris faculty in 1339 and 1340 and was one of those summoned to the Curia in 1340.

The main historical origin of Nicolas's skeptical and critical views about the extent of natural knowledge was undoubtedly the prominence given to the article of the Creed "I believe in one God, Father Omnipotent, Maker of heaven and earth, …" after the condemnation of 1277. As the theologians of the fourteenth century interpreted this article, it meant that God can accomplish anything the doing of which involves no logical contradiction. Now, the miracles of the Old Testament and New Testament are incompatible with the doctrines of Aristotle and his strict interpreters, especially Averroes, in ways that touch directly on the point. Whereas Aristotle denies the

possibility of accidents without substrata, the Eucharist involves the supernatural existence of the accidents of bread and wine after the substance no longer exists (that is, after the substance of bread and wine has been converted into the body and blood of Christ when the priest consecrates the Host). Again, whereas Aristotle had held that effects inevitably arise from their causes unless there is some natural impediment, the episode of the three Israelites who were not consumed in the fiery furnace involves the miraculous interruption of the natural effects of causes where there is no impediment. Consideration of these and like cases led theologians to the following result: The common course of nature can, without logical absurdity, be interrupted by divine power. Hence, the relation of causes and effects or of substances and their accidents is not logically necessitated.

CERTITUDE, SUBSTANCE, AND CAUSE

Nicolas of Autrecourt must have begun his reflections from the consideration of the theological doctrine just mentioned. He maintained that, excepting the certitude of faith, there is but one kind of certitude and this certitude depends on the principle of contradiction: Contradictories cannot be simultaneously true. Nothing is prior to this principle and it is the ultimate basis of all certitude. This certitude is absolute and no power can alter it. It has no degrees and all certitude is reducible to it. Thus, all reasoning by syllogism depends on the principle of contradiction. In every implication (*consequentia*) that is reducible to the principle of contradiction either immediately or by a number of intermediate steps, the consequent of the implication and the antecedent (or a part of the antecedent) are really identical. Otherwise it would not be evident that the antecedent is inconsistent with the denial of the consequent. From all this Nicolas derives the following result: From the fact that one thing is known to exist it cannot be inferred with an evidence reducible to that of the principle of contradiction that another thing exists. Neither the existence nor the nonexistence of one thing can be evidently inferred from the existence or nonexistence of any other thing.

The consequences of this discovery, Nicolas thought, were enough to destroy the whole intellectual enterprise of the Schools. Not only is it impossible that the existence of effects entails the existence of causes, but there is no way to have any evident knowledge of any substance other than one's own soul starting from the objects of sensation or of inner experience. Things apparent to the senses are not substances, and therefore substance cannot be evidently inferred from sensibly appearing objects. Hence the existence of material substances or of other spiritual creatures cannot be inferred with certitude from the evidence of the senses. But this is not all. In one sense of "probable," there is not even a probability that there are any substances. For, in the sense in which the probable is what happens frequently, we can say, for example: When I in the past put my hand toward a fire, it was warmed; it is now probable that if I put my hand toward a fire, it will be warmed. But since there has never been (and could never be) a conjunction in my experience between any appearance and a substance, there is no appearance that renders the existence of a substance so much as *probable* in this sense of the word.

Some of Nicolas's critics urged that substance is deducible from appearances and that causes are deducible from their effects. But he replied that all such deductions depend upon descriptions of appearances and effects that, implicitly or explicitly, contain reference to substances or causes. The deductions from such descriptions are perfectly valid, but nothing in experience or in our stock of self-evident propositions provides the slightest evidence that anything corresponds to such descriptions. In a word, every attempt to prove the existence of substances or causes from appearances or effects begs the question. This point was made in other philosophical writings both before and after Nicolas. The Muslim theologian Mohammad al-Ghazālī, in his *Tahāfut al-Falāsifah* (Incoherence of the Philosophers; see *Averroes' Tahafut al-Tahafut,* edited and translated by Simon van den Bergh, London, 1954, pp. 329–333), pointed out that logically guaranteed inferences concerning causes depend on the description and definitions of terms and so, in a sense, are mainly verbal arguments. Nicolas could not have had access to this work because the relevant sections were not translated until sometime later. David Hume's negative critique of belief in causation and belief in substance parallels that of Nicolas very closely, but Hume had no possible access to the writings of Nicolas because these were not discovered until the nineteenth and twentieth centuries in the Bibliothèque Nationale and the Bodleian Library.

CRITIQUE OF ARISTOTLE

The purpose of Nicolas's critique of Aristotle and his followers is set forth in the prologue to his *Exigit Ordo Executionis.* He tells us that he read the works of Aristotle and his commentator Averroes and discovered that the demonstrations of their doctrines were defective, that arguments for the opposite of these doctrines can be found that are more plausible than arguments for them.

(The word *plausible* here is intended to translate the Latin word *probabilis* because, in this usage, it does not mean "frequent" but "plausible.") Moreover, men have spent their entire lives studying Aristotle to no avail while neglecting the good of the community. Men would live better lives and contribute to the common good, in matters religious and moral, if only they knew that very little certitude about things can be learned from natural appearances and that what little can be learned can be obtained in a short while, provided men attend to things rather than the treatises of Aristotle and Averroes. In a word, the intellectual culture of Nicolas's age is condemned as largely vain; and the purpose of his criticism is simply to show this in detail. This is not to say that Nicolas is opposed to empirical investigation, but it would be a mistake to see in his attack on Aristotle an interest in empirical investigation such as we find in the promoters of natural science in the sixteenth and seventeenth centuries.

The criticism of Aristotle as set forth in the *Exigit* has an aspect not indicated in his controversy with Bernard of Arezzo. In the letters to Bernard he declared that nothing that is said about infrasensible reality is even probable. In another sense of probability, introduced in the *Exigit* (but one of the accepted senses of the term in the Middle Ages and derived, in fact, from Aristotle), a proposition or opinion is probable if there are arguments in its favor that, although inconclusive, would be approved by an impartial judge. In this sense, a proposition or opinion has a probability that varies as our information increases. Accordingly, Nicolas begins with a conception that is accepted by his adversaries: The principle that the Good exists in our minds as a kind of measure for evaluating things. According to this, we may assume that the things in the universe are so arranged that whatever is good exists and whatever is bad does not exist. Since there is no way of demonstrating that things exist in a certain arrangement, we are obliged to depend on the principle of the Good in order to determine what is probably the case. Following this principle we can suppose that (1) all things in the universe are mutually connected so that one thing exists for the sake of another (like Aristotle's view that all things are ordered to one ultimate end, that is, God; cf. Aristotle, *Metaphysics* 1075a15ff.); (2) there is systematic subordination of all things to a single end so that nothing exists that does not somehow contribute to the good of the entire universe; (3) the universe, so conceived, must be at all times equally perfect.

ATOMISM

From the above, Nicolas concludes that any particular thing that now exists has always existed and will always exist. For whatever now exists, exists for the good of the whole, and because this whole is always and everywhere equally perfect, all its parts must always exist. Hence, on the principle of the Good, every ultimate entity in the universe is eternal.

The eternity of things is obviously incompatible with Aristotle's thought, in which the generation and corruption of substances and their accidents is an essential feature. Here Nicolas is content to show that all the Aristotelian arguments to prove the occurrence of generation, corruption, or other kinds of change are inconclusive. For example, we cannot prove conclusively that sensible qualities cease to exist. The only method of proving this is to argue that a quality ceases to exist because it no longer appears to us, and this is obviously inconclusive. Hence, Nicolas argues, the atomic theory in its most radical form is more plausible than Aristotle's nonatomistic theory of change. The *appearance* of change can be accounted for in terms of the aggregation and separation of atomic particles.

There is much of interest in the finer details of Nicolas's atomism, particularly in his defense of indivisible minima as the ultimate constituents of the continuum, his defense of the vacuum, and his theory of motion. But here he is by no means original. His theory of the nature of motion, for example, is taken over from William of Ockham, and his views about indivisibles owe much to other fourteenth-century Scholastics. Moreover, there are radical deficiencies in his views on these subjects. Nicolas also adopted the radical Ockhamist thesis that relations are reducible to their terms, so that there are no extracognitive referents to our relational concepts. The denial of extracognitive relations is mistaken, and this part of Nicolas's speculations suffers from this error.

The *Exigit* also develops a theory of knowledge in terms of which whatever appears to be the case is the case, that is, that the objects of cognition are all in some way real. Nicolas also develops a positive theory of causation, and there is a related theory of eternal recurrence. Whether he derived this from Stoic sources is not clear.

INFLUENCE AND IMPORTANCE

The skeptical and critical views, as well as Nicolas's probabilistic defense of atomism, produced some responses among his contemporaries and successors. Albert of Saxony, Jean Buridan, and others replied to his critical views

on causation and substance, and Thomas of Strasbourg discussed his atomism. Many references to his views on the nature of propositions occur in later fourteenth-century theologians. Moreover, although Nicolas's views were formally condemned by the Curia in 1346, at the end of the century Cardinal Pierre d'Ailly not only adopted many of these views but also wrote that "many things were condemned against [Nicolas] because of envy which were later publicly stated in the schools."

The importance of Nicolas of Autrecourt in the history of thought can best be summarized as follows: He was a radical representative of an increasing tendency in fourteenth-century thought to reject the idea that any of the principles of natural theology admit of demonstration, and he thus contributed to the decline of the authority of Aristotle. Although some of his reflections are both important and valid, they seem not to have had any direct effect on the development of philosophy in early modern times. From one point of view, he and some of his contemporaries achieved a clarity about the nature of beliefs in causation and substance that was neither equaled nor surpassed until the eighteenth century in the writings of Hume.

See also Ailly, Pierre d'; Albert of Saxony; al-Ghazālī, Muhammad; Aristotelianism; Aristotle; Atomism; Averroes; Buridan, John; Henry of Harclay; Holkot, Robert; Hume, David; John of Mirecourt; Medieval Philosophy; William of Ockham.

Bibliography

Editions of Nicolas of Autrecourt's writings are found in J. Lappe's "Nicolaus von Autrecourt," in *Beiträge zur Geschichte der Philosophie des Mittelalters* 6 (2) (Münster, 1908), and J. Reginald O'Donnell's "Nicholas of Autrecourt," in *Medieval Studies* 1 (1939): 179–280, which contains an edition of the *Exigit.*

A study of Nicolas's work is found in Lappe's article. Other studies have been made by P. Vignaux, "Nicolas d'Autrecourt," in *Dictionnaire de théologie catholique,* Vol. XI (Paris, 1931); J. Reginald O'Donnell, "The Philosophy of Nicholas of Autrecourt and His Appraisal of Aristotle," in *Medieval Studies* 4 (1942): 97–125; J. R. Weinberg, *Nicolaus of Autrecourt* (Princeton, NJ: Princeton University Press for University of Cincinnati, 1948); Mario del Pra, *Nicola di Autrecourt* (Milan: Fratelli Bocca, 1951); and V. Zoubov, "Nicolas iz Otrekura i Drevnegrecheskie Atomisti," in *Trudi Instituta Istorii Estestvoznaniia i Tekhniki* (SSSR Akademiia Nauk) 10 (1956): 338–383.

Julius R. Weinberg (1967)

NICOLAS OF AUTRECOURT [ADDENDUM]

Documentation about Autrecourt's life is scarce. His date of birth is now placed sometime between 1295–1298. He came from the diocese of Verdun and attended the arts faculty at Paris. He also held a degree in civil law, which he must have obtained outside of Paris. His membership in the Collège de Sorbonne places Autrecourt back in Paris in the 1330s as a student in theology. He died in 1369, either on July 16 or 17.

Over the last two decades, it has become apparent that the study of Autrecourt's thought has been wrongly placed in the larger context of putatively skeptical tendencies in scholastic thought and the battle against Ockhamism at the University of Paris in the years 1339–1347. In his *Universal Treatise* (*Exigit ordo*), which originated at the arts faculty during the years 1333–1335, he defends the Aristotelian thesis that our sensory experiences are reliable—that what appears really is, and that what appears to be true really is true (*Metaphysics* IV, 5). He finds this view more plausible than its opposite, namely that the intellect is incapable of certitude.

In his *Letters*, Autrecourt attacks the "Academics" or ancient Skeptics. Yet, at the same time, he challenges the prevailing Aristotelian tradition, in particular of substance-accident structure of reality and the principle of causality. This view is the result of his stance that all *evident* knowledge (with the exception of the certitude of faith) must be reducible to the principle of noncontradiction (*primum principium*). This outlook was developed in his correspondence with a Master Giles (of Feno?) and his two extant letters to the Franciscan theologian Bernard of Arezzo, which must have been written sometime between October 1335 and June 1336. These exchanges hark to a previous discussion between Aurtecourt and Bernard of Arezzo at their inaugural lectures (*Principia*) on the *Sentences* about the validity of Aristotle's principle of noncontradiction.

Bibliography

TRANSLATIONS OF AUTRECOURT'S WRITINGS

Nicolas d'Autrecourt, Correspondance, articles condamnés; texte latin établi par L.M. de Rijk; introduction, traduction et notes par Christophe Grellard. Paris: Vrin, 2001

Nicholas of Autrecourt: His Correspondence with Master Giles and Bernard of Arezzo: English translation by L. M. de Rijk. Leiden: E. J. Brill, 1994.

Nicholas of Autrecourt: The Universal Treatise. English translation by Leonard A. Kennedy, Richard E. Arnold, and

Arthur E. Millward; introduction by Leonard A. Kennedy. Milwaukee: Marquette University Press, 1971.

Nicolaus von Autrecourt: Briefe. German translation by R. Imbach and D. Perler. Hamburg: Meiner, 1988.

WORKS ON AUTRECOURT

Dutton, B. D. "Nicholas of Autrecourt and William of Ockham on Atomism, Nominalism, and the Ontology of Motion." *Medieval Philosophy and Theology* 5 (1996): 63–85.

Kaluza, Z. "Éternité du monde et incorruptibilité des choses dans l'Exigit ordo de Nicolas d'Autrecourt." In *Tempus, aevum, aeternitas. La concettualizzazione del tempo nel pensiero tardomedievale*, edited by G. Alliney and L. Cova, 207–240. Florence: L.S. Olschki, 2000.

Kaluza, Z. *Histoire littéraire de la France*. Vol. 42, *Nicolas d'Autrecourt. Ami de la vérité*. Paris: Histoire littéraire de la France, 1995.

Kaluza, Z. "Les catégories dans l'*Exigit ordo*. Étude de l'ontologie formelle de Nicolas d'Autrecourt." *Studie Mediewistyczne* XXXIII (1998): 97–124.

Scott, T. K., "Nicholas of Autrecourt, Buridan, and Ockhamism." *Journal of the History of Philosophy* 9 (1971): 15–41.

Tachau, K. H. *Vision and Certitude in the Age of Ockham: Optics, Epistemology, and the Foundations of Semantics, 1250–1345*. Leiden: Brill Publishers, 1988.

Thijssen, J. M. M. H. *Censure and Heresy at the University of Paris, 1200–1400*. Philadelphia: University of Pennsylvania Press, 1998.

Thijssen, J. M. M. H. "John Buridan and Nicholas of Autrecourt on Causality and Induction." *Traditio* 43 (1987): 237–255.

Thijssen, J. M. M. H. "The Quest for Certain Knowledge in the Fourteenth Century: Nicholas of Autrecourt against the Academics." In *Ancient Scepticism and the Sceptical Tradition*, edited by J. Sihvola, 199–223. Helsinki: Societas Philosophica Fennica, 2000.

Thijssen, J. M. M. H. "The 'Semantic Articles' of Autrecourt's Condemnation." *Archives d'histoire doctrinale et littéraire du moyen âge* 65 (1990): 155–175.

Zupko, J., "Buridan and Skepticism." *Journal of the History of Philosophy* 31 (1993): 191–221.

Johannes M. M. H. Thijssen (2005)

NICOLE, PIERRE

(1625–1695)

Nicole Pierre was born in Chartres, the son of Jean Nicole, a member of the Parlement de Paris. In 1642 he began his studies in philosophy in Paris, where he received his Master of Arts in 1644. Subsequently, he studied theology with Alphonse Le Moine and Jacques Sainte-Beauve, and under the direction of the latter he started an intensive consideration of the theological writings of St. Augustine. During this time Nicole became involved in the activities of the reformist convent of Port-Royal des Champs through his aunt, Marie des Anges Suireau, who was for a short time the abbess there. Nicole taught in the *petite écoles* attached to Port-Royal, where one of his students was Jean Racine, the future poet. After receiving his Bachelor of Arts degree in 1649, he withdrew to Port-Royal, becoming one of the solitaires associated with the convent.

During the 1650s Nicole went against the French theological and political establishment in defending the theological orthodoxy of the *Augustinus* of Cornelius Jansenius, the late theologian and bishop. He joined his fellow solitaire Antoine Arnauld and other Port-Royalists in protesting the papal bulls in the 1650s that attributed to this work four heretical propositions and one false proposition concerning sin, free will, and grace. The controversy that derived from this protest was such that when he returned to Paris in 1654, Nicole was forced to take the assumed name of M. de Rosny.

In 1658, during a tour in the German territories, he translated the *Provinciales* (1656–1657) into Latin, using the pseudonym Guillaume Wendrock. This work, written by the brilliant Port-Royalist Blaise Pascal, was a popular satirical critique of Jesuit moral theology. Nicole also defended both the *Augustinus* and Port-Royal throughout the 1660s, when Louis XIV exerted considerable pressure on the members of the convent to bring them into conformity with official church policy. During this time, in 1662, he published with Arnauld, under the pseudonym of Sieur le Bon, the first of what was to be six editions of the *Logique ou l'art de penser*. This work reflects the teaching at the petite écoles at Port-Royal before their disbandment by Louis XIV in 1660. This work combines an Augustinian distinction between a theology grounded in trust of authority and a philosophy grounded in trust of natural reason with René Descartes's rejection of radical Pyrrhonian skepticism and his metaphysical conclusion that mind as a thinking thing is a substance really distinct from body as an extended thing. Nonetheless, Nicole was never as enthusiastic about the new Cartesian philosophy as his coauthor, Arnauld, was. In several letters published in his four-volume *Essais de morale* (vol. 2, 1679) Nicole emphasized the weakness of human reason and the inability of the Cartesians to offer more than probable conclusions. This sort of emphasis was in line with the skepticism concerning the new philosophy reflected in the views of Port-Royal solitaires such as Le Maistre de Sacy and Louis-Paul du Vaucel. Such skepticism belies the claim of the Calvinist Pierre Jurieu that "the theologians of Port-Royal are as attached to Cartesianism as they

are to Christianity" (*La politique du clergé de France* [Cologne, 1681], 107).

The Peace of the Church that Pope Clement IX established in 1669 with the help of Louis XIV brought about a decade-long cessation of hostilities against the Jansenists. During this period Arnauld and Nicole devoted themselves to their three-volume *La perpétuité de la foy*, in which they defended the Catholic doctrine that Christ is "physically present" in the Eucharist against the view of the Calvinist minister JeanClaude that Christ has a merely "spiritual presence" in this sacrament. Nicole and Arnauld also condemned the attempt of the French Benedictine Robert Desgabets to defend the view in Descartes's unpublished correspondence that the physical presence of Christ involves merely the union of His soul with the matter of the Eucharistic elements. The anonymous publication of this defense in the *Conisdérations sur l'état present* (1671) was one of the triggers of the official campaign against Cartesianism in France during the 1670s.

The Peace of the Church officially ended with Louis XIV's banishment of Nicole and Arnauld, along with other Port-Royalist sympathizers, to the Spanish Netherlands (now Belgium) in 1679. In contrast to Arnauld and the other Port-Royalists, however, Nicole was eager to reconcile himself with the French authorities, and negotiations with the bishop of Paris, François de Harlay de Champvallon, allowed him to return to Paris in 1683. After this return, he further revised his *Essais de morale* and attacked in print the views of the Calvinists. Nicole also attempted (unsuccessfully) to moderate the tone of the increasing bitter philosophical and theological debate during the 1680s and early 1690s that pitted Arnauld against the French Cartesian Nicolas Malebranche.

In the 1690s Nicole also became embroiled in his own dispute with Arnauld over Nicole's view that God grants us a "general grace" that involves at least an implicit knowledge of moral truth. Appealing to the Cartesian doctrine of the transparency of the mind, Arnauld objected to any knowledge of moral truth that does not involve explicit awareness. The response to this line of objection in Nicole and his defenders, including the Louvain theologian Gommaire Huygens and the French Benedictine François Lamy, invoked the purported implication in Augustine that we see truths in God by means of divine illumination that we do not grasp completely. The case of this dispute serves to further illustrate the complexities of the relations between Augustinianism and Cartesianism during the seventeenth century.

During the 1690s, Nicole also found himself opposed to Lamy over the "quietist" doctrine of the French Cardinal François de Fénelon that we are to have a "pure love" of God that involves no concern for the self. Whereas Lamy defended Fénelon, Nicole joined the French Bishop Jacques-Bénigne Bossuet in arguing for the conclusion, which Rome later endorsed, that quietism is heretical. Soon after this dispute, Nicole suffered a stroke, and he died in Paris on November 16, 1695, a little over a year after Arnauld's death.

See also Arnauld, Antoine.

Bibliography

Arnauld, Antoine, and Pierre Nicole. *Logic, or, The Art of Thinking,* edited by Jill Vance Buroker. New York: Cambridge University Press, 1996.

WORKS BY NICOLE

De l'éducation d'un prince [later, vol. 2 of *Essais de morale*]. Paris: Savreaux, 1670.

Die Philosophie des 17.Jahrhunderts. Band 2, Frankreich und Niederlande. Basil: Schwabe, 1993.

Essais de morale, contenus en divers traitez sur plusieurs devoirs importans. 4 vols. Paris: Savreux (vol. 1) 1671, and Desprez (vol. 2) 1679; (vol. 3) 1675; (vol. 4) 1678.

Traité de la grâce générale. 2 vols. N.p.: n.p., 1715.

Oeuvres. 25 vols. Paris: n.p., 1765–1782.

WORKS ON NICOLE

James, Edward D. *Pierre Nicole, Jansenist and Humanist: A Study of His Thought*. The Hague: Nijhoff, 1972.

[Rodis-] Lewis, Geneviève. "Augustinianisme et cartésianisme à Port-Royal." In *Descartes et le cartésianisme hollandaise*, edited by E. J. Dijksterhuis et al., 131–182. Paris: Presses Universitaires de France, 1950.

Solère, Jean-Luc. "Arnauld versus Nicole: A Medieval Dispute." In *Interpreting Arnauld*, edited by Elmar J. Kremer, 127–146. Toronto: University of Toronto Press, 1996.

Tad M. Schmaltz (2005)

NIEBUHR, REINHOLD
(1892–1971)

Reinhold Niebuhr was eminent in two fields. One was social action and analysis of current social problems; the other was the interpretation of the Christian faith. This entry will concentrate on his religious and ethical thinking.

Niebuhr was born in Wright City, Missouri. His father was Gustave Niebuhr, a minister in the Evangelical Synod of the Lutheran Church, who came to the United States when he was seventeen years old. His mother was

the daughter of the Reverend Edward Jacob Hosto, a second-generation German American of the same religious sect. Niebuhr studied at Elmhurst College, Eden Theological Seminary, and Yale University. He was ordained in 1915 and was pastor at the Bethel Evangelical Church of Detroit until 1928. He was then appointed professor at the Union Theological Seminary in New York, where he taught until 1960, when he became professor emeritus.

RELIGIOUS VIEWS

The central theme of Niebuhr's religious teaching can be stated as follows: A divine, forgiving, and timeless love "beyond history" gives meaning to human life. Nothing actually operating in human history can ever be sufficiently dominant over sinful pride and sensuality to deliver men from despair, although men attempt to conceal reality with optimistic illusions. But if we look beyond the temporal process to transcendent being, we find, through faith, a forgiving and perfect love that gives to human life a grandeur beyond the reach of despair and a zeal beyond the reach of apathy. This love from beyond history has been revealed to us in Jesus Christ. We know it is from beyond history because in history this kind of love, called agape, is ineffective before the powers that rule this world. It is futile and meaningless except when, as in the Christian faith, it reveals the ultimate purpose of our existence by an evaluation that transcends history.

SIN AND ANXIETY. Sin arises from anxiety, although anxiety is not sinful in itself. Man is rendered anxious by criticizing himself and his world, by recognizing his own limitations and the contingencies of his existence, and by imagining a life infinitely better than what actually is.

Anxiety would not lead to sin if we brought it under control by trusting ourselves to God's forgiving love and ultimate power. But instead of this, we seek to bring anxiety under control by pretending to have power or knowledge or virtue or special favors from God, which we do not have. This pretense leads to pride, cruelty, and injustice. Or we seek to escape anxiety by dulling the awareness of it with sensuality. All this is sin because it is a turning away from God to a self-centered existence. Sin thus induced is not inevitable, but it is universal. Also Niebuhr obscurely suggested that sin was in the world before men became sinners, this prehuman sin being symbolized by Satan.

In this predicament we have two alternatives. We may trust ourselves along with the whole of human history to God's forgiving love. The other alternative is twofold: to sink into annihilating despair or to conceal our predicament with illusions that render our condition even more desperate in the end. If we take the first alternative, we live not only for whatever love can be attained in history but also and primarily for the divine love beyond history. In this way the whole of history takes on meaning. Otherwise we have only glimpses of meaning in developments occurring here and there but no meaning for the whole of history.

TRANSCENDENCE. Themes continuously recurrent throughout Niebuhr's writing are transcendence, freedom, reason, and love. Niebuhr's language often suggests that by "transcendence" he means the timeless ideal of perfect love. But for Niebuhr this love is not merely an ideal. It is a God who loves, yet is beyond time, cause, and world.

Self-transcendence is a central theme in Niebuhr's thought. If this merely meant that the self can change into a better self, the meaning would be obvious. But Niebuhr seems to mean that the self, while never escaping finitude in one dimension, does somehow, in another dimension, transcend time and causation and self. It does this by surveying past and future and by self-criticism. But to survey past and future is to be aware of one's involvement in time; and in self-criticism the self in retrospect is criticized by the present self; and this criticizing self may in turn be criticized by the self at a later time. Niebuhr would seem to be wrong, therefore, in claiming that in self-criticism the self can transcend time and causation.

FREEDOM AND REASON. Niebuhr affirmed human freedom by paradox: Man is both bound and free, both limited and limitless; he is, and yet is not, involved in the flux of nature and time. As spirit he "stands outside" time, nature, world, and self, yet is involved in them. Freed of paradox, these affirmations assert that humankind is free in the dimension of spirit but not in the dimension of natural existence. The human spirit transcends the self, time, and nature because the individual can know himself as an object, can judge himself to be a sinner, can survey past and future. "The ultimate proof that the human spirit is free is its recognition that its will is not free" (*Nature and Destiny of Man,* Vol. I, p. 258).

Niebuhr would seem to be making contradictory statements. The self is not free if only the "spirit" transcending the self is free. The critical comment made above on his concept of transcendence would apply here also.

Reason is an instrument, says Niebuhr, which can be used for either good or evil. One evil use of reason is to impose rational coherence upon reality and to reject as unreal what cannot be fitted into that coherence. But Niebuhr is mistaken in thinking that one who insists on subjecting every affirmed belief to the tests of reason is thereby claiming that reason comprehends all reality. To the contrary, such a person fully admits that unknown reality extends beyond his knowledge; but he refuses to conceal his ignorance by superimposing religious beliefs where knowledge cannot reach. Niebuhr defended such beliefs because they relieve anxiety by providing courage and hope.

Another sinful use of reason, says Niebuhr, is to make it the basis of a false security, thus turning away from the one sure ground of security, which is a belief beyond the tests of reason, namely, that God in forgiving love will overrule all evil "at the end of history." Here again the question arises: Is true security to be found in beliefs exempt from the tests of reason or is it to be found by rejecting such beliefs and recognizing the unknown without concealing it beneath beliefs that cannot be rationally defended?

On the other hand, Niebuhr used to the full his own magnificent powers of rational intelligence in dealing with problems arising in the temporal process of human existence. He completely accepts the powers of reason in dealing with such problems. For him reason has the further use of demonstrating its own incapacity for dealing with those religious beliefs that Niebuhr affirms while admitting that they cannot be rationally defended.

In June Bingham's book *Courage to Change* (p. 224) she reports that Niebuhr wrote to a friend that he (Niebuhr) adhered to the religious pragmatism of William James. He validates Christian belief, when it cannot be rationally defended, by the courage, hope, peace, zeal, love, sense of being forgiven, and other psychological effects resulting when these beliefs are affirmed. Niebuhr identified these psychological effects as the grace bestowed upon us by God when we affirm these beliefs with the total self. Thus are we assured that we are loved and forgiven by God while we are yet sinners. Niebuhr also affirmed that beyond all the incoherence of our existence and beyond all our rational powers to know there is an all-comprehending and perfect coherence that somehow overcomes and absorbs all the manifest incoherences that we experience.

LOVE. Niebuhr distinguished three kinds of love: heedless love (agape), which seeks nothing in return; mutual love; and calculating love. Heedless love is God's way of loving; and human beings by God's grace may have it to some degree. Since it seeks nothing in return, it cannot have the intention of awakening responsive love, although this may be its unintended result. Suffering endured with intention to awaken responsive love would be calculating love. Hence God's suffering love in Christ is not to awaken responsive love, although this may be its unintended result; but the intention is to protect God's righteousness in forgiving sin because forgiveness without atonement would be condoning sin.

POLITICAL VIEWS

In making political judgments, the individual is inevitably biased by the social position and historical process in which he finds his security and personal identity. No one can be entirely free of this bias, but its distortions are reduced by a faith that finds its ultimate security not in any social position or historical process but in the God of love and mercy who rules supreme over the whole course of history, determining its final outcome as no plan or purpose of man can ever do. Such a faith in God's power and forgiveness enables one to practice "Christian realism," whereby one is able to see the evil in the self and in the historical process with which the self is identified, as well as the depth of evil in all of human life. Political judgment can then be more free of the illusions generated by false pride, on the one hand, and by despair, on the other.

Justice requires the coercions of government to support moral demands; and the power of opposing parties must be equalized if one is not to be subordinated unjustly to the interests of the other. Also, to have justice, freedom to criticize is required. Justice serves love by providing the social conditions required for the practice of love. Love is the final norm but cannot by itself guide political action, because every project set forth in the name of love amid the contests for political power is infected with self-interest whereby the needs of others are falsely identified with those of self.

With his highly developed rational powers and critical intelligence, Niebuhr sharply distinguished between problems subject to rational treatment and religious beliefs that cannot be rationally defended. This gives us what at times seems to be two Niebuhrs: One, the naturalist struggling with the problems of our existence with all the tools of human reason; the other, the mystic upholding a superstructure of religious belief beyond the tests of reason. Whether one of these, or both, will prevail in the course of history, only time can tell. However, the

impact of Niebuhr's thought and action on our civilization will continue in one form or another for a long time.

See also Determinism in History; James, William; Love; Philosophy of History; Philosophy of Religion, History of.

Bibliography

WORKS BY NIEBUHR

Does Civilization Need Religion? A Study of the Social Resources and Limitations of Modern Life. New York: Macmillan, 1928.

Leaves from the Notebook of a Tamed Cynic. Chicago: Willett, Clark and Colby, 1929; new edition, Hamden, CT: Shoe String Press, 1955. A diary of experiences when pastor in Detroit.

Moral Man and Immoral Society: A Study in Ethics and Politics. New York: Scribners, 1932.

Reflections on the End of an Era. New York: Scribners, 1934.

An Interpretation of Christian Ethics. New York: Harper, 1935.

Do the State and Nation Belong to God or the Devil? London: Student Christian Movement Press, 1937.

Beyond Tragedy, Essays on the Christian Interpretation of History. New York: Scribners, 1937. Sermons.

Europe's Catastrophe and the Christian Faith. London: Nisbet, 1940.

Christianity and Power Politics. New York: Scribners, 1940.

The Nature and Destiny of Man: A Christian Interpretation. 2 vols. New York: Scribners, 1941–1943. The most complete statement of his thought.

The Children of Light and the Children of Darkness: A Vindication of Democracy and a Critique of Its Traditional Defense. New York: Scribners, 1944.

Discerning the Signs of the Times, Sermons for Today and Tomorrow. New York: Scribners, 1946. Sermons.

Faith and History: A Comparison of Christian and Modern Views of History. New York: Scribners, 1949.

The Illusion of World Government. Whitestone, NY: Graphics Group, 1949.

Christian Realism and Political Problems, Essays on Political, Social, Ethical, and Theological Themes. New York: Scribners, 1953.

The Self and the Dramas of History. New York: Scribners, 1955.

Pious and Secular America. New York: Scribners, 1958.

The Structure of Nations & Empires. New York: Scribners, 1959.

WORKS ON NIEBUHR

Bingham, June. *Courage to Change.* New York: Scribners, 1961.

Fox, Richard. *Reinhold Niebuhr: A Biography.* New York: Pantheon, 1985.

Gilkey, Langdon. *On Niebuhr: A Theological Study.* Chicago: University of Chicago Press, 2001.

Kegley, Charles, and Robert Bretall, eds. *Reinhold Niebuhr: His Religious, Social and Political Thought.* New York: Macmillan, 1956.

Lovin, Robin. *Reinhold Niebuhr and Christian Realism.* Cambridge, U.K.: Cambridge University Press, 1995.

Lovin, Robin. "Reinhold Niebuhr in Contemporary Scholarship: A Review Essay." *Journal of Religious Ethics* 31 (2003): 489–505.

Richards, Priscilla. *Annotated Bibliography of Reinhold Niebuhr's Works.* Madison, WI: American Theological Library Association, 1984.

Scott, Nathan, ed. *Legacy of Reinhold Niebuhr.* Chicago: University of Chicago Press, 1975.

Warren, Heather. *Theologians of a New World Order: Reinhold Niebuhr and the Christian Realists.* New York: Oxford University Press, 1997.

Henry Nelson Wieman (1967)
Bibliography updated by Christian B. Miller (2005)

NIETZSCHE, FRIEDRICH
(1844–1900)

Although trained as a philologist, Friedrich Nietzsche has been among the philosophers most influential upon European and North American culture and philosophy during the twentieth century. While he has always had an audience among writers, artists, and Germanists, through the first half of the twentieth century—and especially among philosophers—Nietzsche was read and discussed primarily by German philosophers, including Martin Heidegger, Karl Jaspers, and Karl Löwith. His criticisms of traditional philosophical positions, along with his often metaphorical and hyperbolic writing style, led to his being taken much less seriously by English-language philosophers. And Nietzsche's political views and the posthumous appropriation—many would argue misappropriation—of some of his ideas by thinkers associated with fascism and National Socialism (Nazism) led initially to a hostile response to his works among many British and French readers.

By the early 1960s, however, Nietzsche's fortunes had begun to change considerably. Anointed along with Marx and Freud as one of the three "masters of suspicion," Nietzsche's philosophical works found enthusiastic readers among those coming of age philosophically in the 1960s, and this—along with a new critical edition of his works and several generations of scholarly explication and analysis—resulted in Nietzsche being among the most widely read and known of Western philosophers by the end of the twentieth century.

BIOGRAPHY

Nietzsche was born October 15, 1844, in Röcken, a small village in Prussian Saxony, on the birthday of King Friedrich Wilhelm IV of Prussia, after whom he was

named by his father Karl Ludwig, 31, and his mother Franziska (née Oehler), 18. His father, as well as both of his grandfathers, were Lutheran ministers. In 1846, Nietzsche's sister Elisabeth was born, and two years later, his brother Joseph was born. The following years were difficult ones: in 1848, Nietzsche's father became seriously ill; he died on July 30, 1849, of what was diagnosed as "softening of the brain" (a frequent diagnostic notation for tertiary syphilis). The following year, Nietzsche's younger brother died; and in April 1850, Nietzsche's mother moved the household—which now included her two young children, as well as Nietzsche's paternal grandmother and her two sisters—to Naumberg, a much larger town of 15,000 people.

In 1858, Nietzsche was offered free admission to Pforta, the most prestigious high school in Germany, located only a few miles from Naumberg. He was an excellent student and graduated in 1864 with a thesis in Latin on the Greek poet Theognis. After graduation, he registered at the University of Bonn as a theology student, but quickly changed his focus to philology, as Bonn's department had a distinguished reputation grounded on the work of two professors: Otto Jahn (1813–1869) and Friedrich Wilhelm Ritschl (1806–1876). There were, however, deep personal and professional disagreements between the two and when Ritschl decided to leave for the University at Leipzig, Nietzsche followed him there in 1865 and registered as a student of classical philology. Nietzsche soon became Ritschl's star pupil, and he was invited by Ritschl to publish an essay on Theognis in *Das Rheinische Museum für Philologie*, which Ritschl edited. In addition to his work in philology, writing essays on Diogenes Laertius and Democritus, among others, three other events took place in Leipzig that would profoundly influence the rest of Nietzsche's life: his discovery of Schopenhauer's *Die Welt als Wille und Vorstellung* (*The World as Will and Representation*) in 1865, of F. A. Lange's *Geschichte des Materialismus* (*History of Materialism*) in 1866, and in 1868, his meeting Richard Wagner, with whom he shared a love of music, of Schopenhauer, and a hope for the revitalization of European culture.

When a position at the University of Basel appeared in 1869, Ritschl gave an extraordinary recommendation for Nietzsche, who had not yet written a doctoral thesis, and Nietzsche was appointed to the Chair of Classical Philology at Basel in 1869 at the age of twenty-four. The University of Leipzig proceeded to confer the doctorate without either thesis or examination, and Nietzsche moved to Basel in April 1869. Basel offered him not only a university appointment but also easy access to the Wagner residence at Tribschen, which allowed Nietzsche to develop a close relationship with both Wagner and his wife Cosima, the daughter of Franz Liszt. While at Basel, Nietzsche lectured on Homer, Hesiod, Plato, Aristotle, the pre-Socratics, Diogenes Laertius, and classical rhetoric. He was becoming increasingly disengaged from philology, however, and spent much of his time working on the texts of ancient Greek and Roman philosophy and thinking about broad cultural issues. These two features can be seen in his first book, *The Birth of Tragedy* (1872), which merged philosophical reflection with philological interpretation as it sought to frame Wagnerian opera as a way to recuperate what European culture had lost since the demise of ancient Greek tragedy. While Nietzsche thought his work would revolutionize the discipline of philology, it was poorly received and all but destroyed his professional standing as an academic philologist.

During the 1870s in Basel, Nietzsche became increasingly uncomfortable with Wagner and the Wagner circle at Tribschen and Bayreuth. While there is no question that *The Birth of Tragedy* proclaims Wagner's world-historical importance as a cultural phenomenon, *Richard Wagner in Bayreuth*, the fourth of his *Untimely Meditations*, is much more ambivalent. By 1878, Nietzsche had had enough of Wagner and among the reasons he offers subsequently to explain his break with Wagner are Wagner's turn to Christianity in *Parsifal* and his support for and association with political anti-Semitism. In 1879, Nietzsche resigned his chair at Basel because of the increasing severity of his health problems, and over the next ten years, he lived in several places in Europe, including Sils Maria, Switzerland, and Genoa and Turin, Italy. During these ten years, Nietzsche wrote ten books, living off a modest pension from the university, and he was plagued by constant and severe health problems. He suffered a total mental breakdown in Turin in January 1889, and after a brief stay at the psychiatric clinic run by Dr. Otto Binswanger in Jena, he spent the remaining years of his life under the care of his mother and then his sister until his death in Weimar on August 25, 1900.

No account of Nietzsche's life can avoid his health and his madness. Beginning in childhood, his health was poor. He was plagued by headaches that, as young as nine, kept him from school, and by age twelve, his eyes began to cause him serious problems. Throughout his life, his work habits were affected by the migraines that forced him to remain in darkened rooms, gastrointestinal problems, and limited eyesight that made reading at times painful and at times impossible. Not surprisingly, the themes of sickness, convalescence, and health, both

metaphorically and literally, hold a central place in his philosophical reflections.

The question of his madness has been a focus of attention and speculation almost from its outbreak. What is clear is that on the morning of January 3, 1889, Nietzsche saw a horse being beaten by its coachman on a street in Turin, embraced the animal, and then collapsed. In the few days preceding and following this event, he sent letters to Jacob Burckhardt, Peter Gast, George Brandes, Cosima Wagner, and August Strindberg, among others, that, while at moments lucid and beautiful, are also clearly not the writings of a sane individual. While there has been much speculation as to the cause of Nietzsche's insanity, there is no conclusive evidence to support either of the two most common hypotheses: that he inherited syphilitic dementia from his father or he caught syphilis from prostitutes in a Leipzig brothel during his time as a student there. Recently, new research carried out by Dr. Leonard Sax, director of the Montgomery Center for Research in Child Development in Maryland and published in the *Journal of Medical Biography*, suggests that Nietzsche's symptomatology is consistent with cancer of the brain and in fact is not consistent with syphilis (based on the number of years Nietzsche remained alive following his breakdown). The syphilis story, it appears, can be traced to a book written by psychiatrist Wilhelm Lange-Eichbaum in 1946, *Nietzsche: Krankheit und Wirkung*, that sought to discredit Nietzsche, and this story was then adopted as fact by intellectuals who shared Lange-Eichbaum's politically motivated desire to destroy Nietzsche's reputation.

WRITINGS

During the sixteen years of Nietzsche's productive life, he wrote eighteen books in addition to leaving an extensive correspondence and several thousand pages of unpublished writings. While there are some minor differences in the way his works are periodized by scholars, his writings tend to be divided into three periods: his early more scholarly, philological work written while teaching in Basel from 1872–76; his aphoristic texts, written between 1878–1882; and his mature works, which begin with *Thus Spoke Zarathustra* in 1883 and continue until his last works in 1888.

THE BASEL WRITINGS. Nietzsche's early works, written while a professor of classical philology at the University of Basel, include *The Birth of Tragedy out of the Spirit of Music*, and the four *Untimely Meditations*: *Richard Strauss, Confessor and Writer*; *On the Use and Disadvantage of History for Life*; *Schopenhauer as Educator*; and *Richard Wagner in Bayreuth*. In addition to these published works, there are several unpublished works from this period that have attracted scholarly attention, the most important of which are the essays "On Truth and Lies in an Extra-moral Sense," "Homer's Contest," and "Philosophy in the Tragic Age of the Greeks."

First published in 1872, *The Birth of Tragedy* offers a theory of tragedy, a theory of art, and a proposal for cultural renewal. A second edition, published in 1886 with a new preface titled "Attempt at a Self-Criticism," and a new subtitle, "Hellenism or Pessimism," takes note of Nietzsche's move away from the Schopenhauerian sensibilities that marked this text by highlighting the opposition between Greek cheerfulness and Schopenhauerian pessimism. *The Birth* opens with Nietzsche's distinction between the Apollonian and Dionysian, which designates both forces of nature and basic artistic impulses. As forces of nature, the Apollonian names the principle of individuation that gives form to the chaos by isolating and distinguishing between things, whereas the Dionysian names the primal unity of all things in an endless play of forces of becoming. As artistic impulses, the Apollonian marks the world of beautiful illusions, whereas the Dionysian marks the sensual world of rapturous frenzy. Sculpture is the purest Apollonian art as a transfiguration of the real into a beautiful, illusory image, whereas music is the purest Dionysian art insofar as music is the process of change itself, with nothing that endures but the whole that survives each individual note's destroying what has come before it.

Nietzsche argues concerning Greek culture that when faced with the absurdity and horrible and terrifying aspects of existence, the Apollonian and Dionysian denote two opposing tendencies of human nature: to cover existence with beautiful illusions or to plunge into the absurdity and horror of existence and affirm it, as such, as a world of continual creation and destruction. From this comes his thesis about tragedy: Attic Tragedy—Sophocles and Aeschylus; Oedipus and Prometheus—manifests the pinnacle of Greek art as the perfect union of Dionysian joy and Apollonian illusion: It reflects both the Greek tragic wisdom that by accepting destruction as part of the great world-game, the tragic hero masters the cruelty of fate, and reveals the tragic Dionysian wisdom that the human spirit will not be broken by the pains and hardships of existence. This is the "metaphysical comfort" that tragedy leaves one with: "that life, despite all the changes in appearances, is at bottom indestructibly powerful and pleasurable" (§ 7). This tragic insight, which

gave birth to Attic Tragedy, was, according to Nietzsche, destroyed by Socrates and his tragedian spokesman Euripides, for whom in order to be beautiful, everything had to be intelligible. Much of Nietzsche's *Birth* is spent analyzing the death of tragedy at the hands of Socrates and Euripides, and the anticipation of its rebirth in Wagnerian opera.

Nietzsche's four *Untimely Meditations* were published between 1873 and 1876. Originally planned as a series of thirteen volumes of cultural criticism, Nietzsche only published four (though he completed a substantial amount of work on a fifth volume on academic philology, "Wir Philologen"). In *David Strauss, the Confessor and Writer* (1873), Nietzsche criticizes Strauss, a Hegelian and author of *The Life of Jesus* (1835) and the then (1870s) popular work *The Old and New Faith*, for his smugness and the ease with which he dispenses with Christian doctrine. Strauss is also treated as representative of German popular culture, pleased with itself and its cultural "superiority" following Prussia's victory in the Franco-Prussian war, and Nietzsche spends much of the text challenging the *Bildungsphilister* or "cultural philistines" who mistake their "popular" culture for "genuine" culture. Because of Strauss's popularity, this was one of Nietzsche's most popular works, which although often critically reviewed was widely read.

On the Use and Disadvantage of History for Life (1873) has been the most widely discussed of the four meditations, although it was the least successful in its day. Taking as his critical foil Eduard von Hartmann's *Philosophy of the Unconscious* (1869), Nietzsche challenges the neo-Hegelian historicist tendency to valorize the present as the goal toward which history had been teleologically directed. While attacking the high value placed upon history in contemporary German culture and education, Nietzsche offers his tripartite account of historical scholarship—antiquarian, monumental, and critical—and offers an early version of what later became his genealogical method of examining the past in order to better understand the present.

Schopenhauer as Educator (1874), which Nietzsche later came to realize should have been called "Nietzsche as Educator," offers an early account of the exemplary individual engaged in a project of self-perfection. One finds relatively little comment in this text about Schopenhauer's philosophical views, about which Nietzsche had, by the time of its writing, come to question. Instead, one finds Nietzsche discussing Schopenhauer as an exemplary philosopher who willingly suffers in pursuit of the truth. It is, then, not Schopenhauer's philosophy but the Schopenhauerian image of man that educates, and Nietzsche's third meditation is one of his most personal books in providing several comments that describe the exemplary individual that Nietzsche himself wanted to become.

That *Richard Wagner in Bayreuth* (1876) came to be published at all is due largely to Nietzsche's friend Heinrich Köselitz ("Peter Gast," 1854–1918). Begun in 1874, Nietzsche's adoration of Wagner began to fade in 1874–75 and he abandoned the project in 1875. Gast read the unfinished manuscript early in 1876 and persuaded Nietzsche first to complete the manuscript as a gift to Wagner for his birthday (May 22), and Nietzsche subsequently decided to publish the volume as the fourth *Untimely Meditation*, presenting it to Wagner in August during the first festival at Bayreuth. Although on the surface an homage to Wagner, with its liberal quotation and paraphrase from Wagner's own writings, the text also suggests that Wagner and his circle may themselves be "cultural philistines" who are failing to live up to the cultural and aesthetic ideals that Wagner's writings proposed. While important in terms of understanding Nietzsche's ambivalence toward Wagner during this period, and offering several insightful comments on art, culture, language, and science, this volume stands as perhaps Nietzsche's least popular and least read work.

In addition to these five published works, Nietzsche also left a number of unpublished essays and fragments from this period. Of these, three are of particular significance: "On Truth and Lies in an Extra-Moral Sense" (1873), in which he offers a tropological account of the origins of knowledge as grounded in the fundamental human drive toward the formation of metaphors; "Homer's Contest" (1872), in which he discusses the role of the *agon* or competition in Greek culture and democracy; and "Philosophy in the Tragic Age of the Greeks" (1873), in which he offers some of his most sustained commentary on the major pre-Socratic philosophers, including Heraclitus, Parmenides, Anaximander, and Anaxagoras.

APHORISTIC TEXTS. Between 1878 and 1882, Nietzsche wrote five works that, on the back cover of the final one, he noted as having a common goal: "to erect a new image and ideal of the free spirit." Motivated in part by his dissatisfaction with Wagner, he turned in these works against art, but more importantly, these works display a sympathy toward science as a legitimate source of truth and knowledge that has led some to refer to the works of this middle period as Nietzsche's "positivistic" works.

These works also shared a common style, that of the aphorism, which Nietzsche adopts in part as a way to mark his antipathy to the German philosophical tradition (Kant, Hegel) and his sympathy to French moral psychologists such as La Rochefoucauld, Montaigne, and Chamfort, whose aphoristic works he was then reading with his new friend Paul Rée (1849–1901).

In each of his aphoristic works, although themselves divided into chapters or parts, Nietzsche numbers his paragraphs sequentially from beginning to end. Some of these paragraphs are several pages long, and others are as short as a single sentence. The first of these works was *Human, All Too Human* (1878). Dedicated to Voltaire on the centenary of his death and subtitled "A Book for Free Spirits," it surveys a full range of philosophical topics, including metaphysics, epistemology, morality, religion, science, art and literature, culture, society, the family, and the state. In addition to being a public announcement of his break with Wagner, this volume also marked a break with the style of his earlier writings, and the multiplicity of authorial voices that speak through the 638 aphorisms are the first published expression of Nietzsche's perspectivist approach. *Human, All Too Human* was followed by two sequels, *Mixed Opinions and Maxims* (1879) and *The Wanderer and His Shadow* (1880), which each offer a collection of aphorisms on a variety of topics that have no apparent organizational structure, and were subsequently published together in 1886 as Volume Two of *Human, All Too Human*.

Unlike his earlier aphoristic works, *Daybreak: Thoughts on the Prejudices of Morality* (1881) remains relatively focused on the single topic of morality and the various themes that moral theorists typically address: moral judgment, moral psychology, moral values, the emotions, the virtues, and so on. It is an important text because it offers an early version of his critique of morality that anticipates many of the ideas that will receive extensive discussion in Nietzsche's later works, especially as concerns the origins of morality in general and some of the Western philosophical and religious traditions' privileged moral values in particular. In *Human, All Too Human*, one glimpses Nietzsche's first explorations into a naturalistic approach to ethics; in *Daybreak*, one finds Nietzsche much more committed to the idea that our moral values have their genesis in our biological and psychological needs.

The Gay Science (1882, 1887) is clearly the most significant work of this middle period, both in bringing to completion the series devoted to the free spirit and in being the text in which Nietzsche first formulates two of his most famous themes: the death of God (§125, "The Madman") and the eternal recurrence (§341: "The Greatest Weight"). While sharing the aphoristic style with the other works of this period, *The Gay Science* stands out in terms of its consistency with the themes that will be expressed in his subsequent writings. It stands out as well in terms of the internal coherence between aphorisms: Where the organization among the various aphorisms in his preceding four books often seems unclear if not nonexistent, there is often in *The Gay Science* a development from the topic of one aphorism to the next that rewards a careful attention to their sequence.

A case in point is the last three sections of Part Four—the last three sections of the first edition—in which Nietzsche moves from "The Dying Socrates" (§340), where Socrates, on his deathbed, discloses his true belief that existence is a disease; to "The Greatest Weight" (§341), in which Nietzsche first introduces the eternal recurrence through the voice of a demon, echoing Socrates's *daimon*, and suggests that contrary to Socrates's judgment, life might be affirmed; to *Incipit Tragoedia* (§342; "The Tragedy Begins"), which is identical to the first section of the Prologue of Nietzsche's next book, *Thus Spoke Zarathustra*, thus introducing Zarathustra as a teacher with an alternative to the moral teachings of Socrates, Kant, and Christianity. In 1887, Nietzsche published a second edition of *The Gay Science*, now with a new preface, an appendix of "Songs of Prince Vogelfrei," and a fifth book that offers some of Nietzsche's most sophisticated reflections on questions of language, consciousness, science, morality, religion, and art. Although appended to this earlier work, the fifth book really belongs to Nietzsche's "mature" period, in which he has fully committed to the perspectivist and constructivist accounts of knowledge.

MATURE PERIOD: TRANSVALUATION OF ALL VALUES. The texts of Nietzsche's mature period, written from 1883 to 1888, include those for which Nietzsche as a philosopher is best known: *Thus Spoke Zarathustra, Beyond Good and Evil*, and *On the Genealogy of Morals*. In addition to these works, he also wrote five books in 1888: two books on Wagner—*The Case of Wagner* and *Nietzsche contra Wagner—Twilight of the Idols, The Antichrist*, and *Ecce Homo*, an autobiography and appraisal of his works, which was published posthumously in 1908.

In *Thus Spoke Zarathustra*, Nietzsche offers the fictional narrative of Zarathustra, his image of the yes-saying spirit, who offers an alternative to the messages of the New Testament. Intentionally parodying the Gospels and,

to some extent, the life of Jesus, Zarathustra opens by taking note of the death of God and subsequently offers his alternative teachings concerning the transvaluation of all values in which the values of this world, the body, self-overcoming, and creativity are all affirmed. Within the beautiful prose of this work, one can find all of Nietzsche's major themes discussed and, in particular, three of Nietzsche's most well-known themes find their primary expressions among his published works here: the *Übermensch* or overhuman (man is something to be overcome), the eternal recurrence (standing at the gateway of the *moment*—the present—two paths confront human beings, one forward in time, one backward, each infinite. And then each person must ask him- or herself: Must not all things that *can* happen *have already* happened and will they not continue to happen? Is not everyone entangled in a complex causal network that cannot be changed and that recurs eternally, in the identical form?), and the will to power (the metaphysical principle that animates all life).

While *Thus Spoke Zarathustra* was the work that first attracted attention to Nietzsche as a philosopher, and it had a profound influence on the existentialist interpretation of Nietzsche's philosophy, it is on the basis of his next two books, *Beyond Good and Evil* and *On the Genealogy of Morals*, that Nietzsche's reputation as a major philosopher resides. In the nine chapters of *Beyond Good and Evil*, Nietzsche offers his clearest criticisms of many central themes in the history of philosophy (free will, the Cartesian ego, the representational model of knowledge, idealism, realism, reason vs. instinct, Kant's transcendental philosophy). He also offers some of his most striking criticisms of religion, of morality (§260 first introduces the distinction between master morality and slave morality), of nationalism, and provides his clearest expression of a philosophy of power (§13: "A living thing seeks above all to discharge its strength.").

Beyond Good and Evil also offers Nietzsche's most sustained defense of perspectivism and his most serious questioning of the value of truth. The text opens with a preface that places truth, aligned with Plato, Christianity ("Platonism for the people"), and dogmatism, in contrast to perspective, and from there moves in Part One—"On the Prejudices of Philosophers"—to question the value of truth as well as the value of many of the central ideas, presumed to be true, of past philosophers, including Plato's Forms, Kant's thing-in-itself, Descartes's ego, and Schopenhauer's will. Throughout his analysis, Nietzsche suggests that the question that should be asked, when considering these philosophical articles of faith is not "Are they true?" but "Why is belief in their truth necessary?"

On the Genealogy of Morals offers Nietzsche's most sustained and powerful account of the origin and value of morality. The work itself unfolds in three carefully constructed essays. In the first, Nietzsche distinguishes between two moral frameworks: the noble morality that is based on distinguishing "good and bad," and the slave morality that makes judgments of "good and evil." The central idea of this first essay, Nietzsche writes, is his discovery of the birth of Christianity out of the slave's spirit of *ressentiment*. The second essay traces the moral concept guilt (*Schuld*) back to its origins in the economic relation of creditor and debtor, and offers an interpretation of the psychology of conscience, not as the voice of God in man, but as the instinct of cruelty that turns back on itself after it can no longer discharge itself externally.

In the third essay, Nietzsche inquires into the meaning of the ascetic ideal and, following an examination of the appearances of the ascetic ideal in philosophy, religion, art, morality, and science, discovers that the ascetic ideal is the harmful ideal par excellence. But the third essay also argues that the ascetic ideal has performed an essential, preservative function in that even though what the ascetic ideal has willed, throughout its long history, has in fact been imaginary (i.e., it has willed "nothing"), through its willing of nothingness, the will itself—that is, the ability to will—was saved. Nietzsche's genealogy of the ascetic will reveals that this will to nothingness, in the form of willing God or willing truth, while an aversion and hostility to life, was still a will that has preserved itself and has driven the deployment of *reactive* forces that is the history of the ascetic ideal. He offers, however, only tantalizing suggestions of a counter-will, a will to power that would no longer be a will to truth but would allow for the deployment of *active* forces that would make possible the overcoming of nihilism that has resulted from two thousand years of ascetic willing.

In 1888, the last year of his productive life, Nietzsche composed five short books. The first, *The Case of Wagner*, is Nietzsche's most sustained criticism of Wagner, and offers as well several insightful comments on art. Nietzsche describes *Twilight of the Idols* in letters on September 12 and 14, 1888, to his friends Peter Gast, Paul Deussen (1845–1919), and Franz Overbeck (1837–1905) as a "summary of my essential philosophical heterodoxies" (*Nietzsche Briefwechsel*), and this short text does indeed offer something of a survey of his basic themes while displaying his stylistic mastery, evidenced well in the title's play on Wagner's 1876 opera *Göttendämmerung*

(Nietzsche's *Götzen-Dämmerung* spoofing Wagner's "Twilight of the Gods"). Among the most interesting sections are his discussions of Socrates ("The Problem of Socrates") Kantian rationalism ("'Reason' in Philosophy"), philosophy ("Four Great Errors"), the influence of religion on morality ("Morality as anti-Nature"), and his highly condensed, six sentence history of Western philosophy and religion ("How the 'Real World' at last Became a Myth: History of an Error"), in which he moves from Plato to Christianity to Kant to positivism to the death of God and Nietzsche's own contributions of the free spirit and Zarathustra.

The Antichrist, which when published Nietzsche conceived, as he noted in the preface to *Twilight*, as the first volume of a longer work to be titled *Transvaluation of All Values*, is Nietzsche's most aggressive critique of Pauline Christianity. *Ecce Homo*, while completed in 1888, was withheld from publication by his sister Elizabeth until 1908. In it, Nietzsche offers a hyperbolic autobiographical and literary self-appraisal that only recently, with the increased attention to Nietzsche's writing style, has attracted the serious philosophical attention it deserves. Nietzsche's final published work, *Nietzsche Contra Wagner*, was dated Christmas 1888, less than two weeks before his collapse. Nietzsche's shortest work, he here reproduces with some minor emendations a selection of his earlier criticisms concerning Richard Wagner, thus making clear that the prosecution of Wagner in *The Case of Wagner* was not a late motif that Nietzsche arrived at only following Wagner's death.

No discussion of Nietzsche's work can fail to take account of his unpublished *Nachlass* of 1883 to 1888, in part because his sister published *The Will to Power*—a relatively small (slightly more than ten percent) and highly edited selection of these notes, first as approximately 400 sections in 1901, and in a second, expanded edition of 1067 sections in 1906—as if it had been a text written by Nietzsche himself. There is no doubt that for several years Nietzsche considered publishing a major work with this title, but there is equally no doubt that he definitively abandoned this project well before his collapse. As a consequence, claims made by Elisabeth and others as to this work being Nietzsche's *magnum opus* clearly cannot be sustained.

Heidegger's claim that *The Will to Power*, by which Heidegger meant the entire 1883 to 1888 *Nachlass* and not just Elisabeth's edition, contained the essence of Nietzsche's philosophizing is a more difficult claim to refute, especially as it relates as much to Heidegger's own desire to situate Nietzsche as the culminating figure in the history of metaphysics. What is clear is that many of Nietzsche's comments on his so-called major themes—most importantly, the eternal recurrence, will to power, and the *Übermensch*—are found primarily in these unpublished notes and, were one to discount the unpublished notes as well as Nietzsche's fictionalized account in *Thus Spoke Zarathustra*, it would be difficult to justify any of these three themes as being a significant part of Nietzsche's published prose works. That said, there is much of interest in these published notes for the philosopher as well as the Nietzsche scholar. While some passages are rough, or simply notes to himself for future work, or ideas and thought-experiments that he played with and chose, quite consciously, not to publish, others may well be ideas that he was still actively working on when his productive life ended.

Of particular note in this regard are his comments on scientists and scientific texts, especially biological texts, that he was reading in the mid- to late-1880s. Nietzsche was during this period reading as much if not more in scientific texts than philosophical texts, and while his biologistic account of life makes its way into some passages in *Beyond Good and Evil* and elsewhere, the best evidence of his thinking on these issues remains to be read in the unpublished notes of the *Nachlass*.

INFLUENCE

Walter Kaufmann opened and closed his article on Nietzsche in the first edition of the *Encyclopedia of Philosophy* with allusions to Nietzsche's influence upon modern philosophy and literature. Yet Kaufmann could scarcely have imagined the explosion of interest in Nietzsche's works, particularly in philosophical circles, that began in the mid-sixties and still continues. Kaufmann's bibliography, a perspectival review to be sure, lists only two secondary works on Nietzsche written in English—his own *Nietzsche: Philosopher, Psychologist, Antichrist* (1950) and George A. Morgan's *What Nietzsche Means* (1941). But since 1967, almost two thousand volumes focused primarily on Nietzsche—more than half of them in English—have appeared in English, French, and German, and perhaps ten times that number of essays, articles, or book chapters have been published.

Charting the expanding horizons of Nietzsche's influence quickly becomes a sociological study of the dominant motifs of late twentieth-century culture, and surveying the influence within the narrower field of philosophical inquiry is equally complex. There may in fact be no philosopher whose works admit less happily to a canonical or consensus interpretation, a claim sup-

ported by the staggering diversity of interpretations of Nietzsche's philosophy that have appeared since 1967. Nevertheless, some general observations can be made concerning the range of these new interpretations.

One can locate at least three primary factors in the increased philosophical attention to Nietzsche over the past forty years. First is the tremendous influence of Martin Heidegger's reading of Nietzsche. Published in Germany in 1960, translated into French in 1962 and into English between 1979 and 1987, Heidegger's overarching interpretation of Nietzsche as the culminating figure in the history of metaphysics inspired an enormous range of exegetical and critical response while leading several generations of philosophers and philosophy students back to read or re-read Nietzsche's texts.

A second reason for the increased attention by philosophers to Nietzsche can be located in the discovery of a "new Nietzsche" that emerged in conjunction with the rise of recent French philosophy. While most widely associated with Jacques Derrida and the deconstructionist attention to questions of textuality and the styles of philosophical discourse, Nietzsche's inclusion, along with Marx and Freud, as one of the three "masters of suspicion," and his importance in the philosophical works of Michel Foucault and Gilles Deleuze, have shown him to be an intellectual influence on much of what is called poststructuralist thought. And, as in the case of Heidegger, the popularity of poststructuralist French thought brought with it a renewed interest—among literary critics and theorists, historians, political theorists, and philosophers—in Nietzsche's thinking.

The third reason for the increased attention to Nietzsche concerns the transformation of philosophy within the anglo-American tradition. In the 1960s, Kaufmann's text, along with Arthur Danto's *Nietzsche as Philosopher* (1965), had first to justify Nietzsche as a philosopher whose ideas warranted serious philosophical consideration. As the scope of English-language philosophy has broadened, a distinctly anglo-American tradition of Nietzsche interpretation has appeared which is informed by the questions of ethics, metaphysics, and epistemology that occupy analytically trained philosophers.

This entry concludes with a brief survey of some of the main issues that have emerged in recent Nietzsche scholarship. To be sure, there is still much work offering interpretations of the classical Nietzschean themes: will to power, eternal recurrence, *Übermensch*, nihilism, perspectivism, and so on. But other issues have appeared as well. For example, an attention to questions of texts and textuality has played a role in much of the recent literature. It has become increasingly common to distinguish between Nietzsche's published texts and his unpublished notes, especially as concerns themes whose primary expression is to be found in the "book" constructed by his literary executors after his death and titled *The Will to Power*. One also finds an increasing tendency to read Nietzsche's texts *as* texts, following their internal development as opposed to simply viewing these texts as collections of remarks from which one can pick and choose the comments relevant to one's own argument. A third theme emerging from the recent interest in textuality is an attention to the various styles of Nietzsche's philosophical prose, in other words, an attention to his use of metaphor, to the literary character of much of his writing (in particular, *Thus Spoke Zarathustra*), to the different genre of writing (aphorism, essay, polemic, poem, etc.), and to other issues characterized collectively as the "question of style."

A second range of topics within the recent Nietzsche literature addresses some of the classic questions of philosophy: Does Nietzsche have a "theory of truth"? Does he have a "theory of knowledge"? An "ontology"? Is Nietzsche a metaphysician in the way that Heidegger defines metaphysics? Is Nietzsche an ethical naturalist? Within these questions, a topic that continues to draw attention is the issue of self-reference; in other words, when Nietzsche makes claims (about truth, reality, being, subjectivity, etc.), do these claims refer or apply to or hold true for his own philosophical conclusions? The most obvious case where the question of self-reference arises concerns the question of truth and interpretation: if Nietzsche claims that "there is no Truth," or that "everything is an interpretation," are these claims put forward as "true"? If they are, then they appear to contradict themselves; but if they are not true, then why should we be interested in them? The issue has been extended beyond the confines of epistemology, however, and one finds discussions of the eternal recurrence or the *Übermensch* or the ascetic ideal in terms of the question of self-reference.

A third and final set of issues that warrants noting is the extension of Nietzschean themes into new areas not discussed, or only hinted at, in the earlier Nietzsche scholarship. Among the most important topics producing much recent scholarship are Nietzsche's influence on postmodernism, his position on "woman" and his relevance for feminism, and his political philosophy and impact on twentieth-century political and social movements.

"Some are born posthumously," Nietzsche wrote in 1888. "One day my name will be associated with the

memory of something tremendous," he claimed in *Ecce Homo*, at the beginning of a chapter titled "Why I am a Destiny?" One hundred years later, these remarks appear prophetic, and at the beginning of the twenty-first century, it would be difficult to find a philosopher whose influence on matters philosophical and cultural exceeds that of Nietzsche.

See also Anaxagoras of Clazomenae; Anaximander; Aristotle; Burckhardt, Jakob; Danto, Arthur; Deleuze, Gilles; Derrida, Jacques; Descartes, René; Diogenes Laertius; Existentialism; Foucault, Michel; Freud, Sigmund; Hartmann, Eduard von; Heidegger, Martin; Heraclitus of Ephesus; Homer; Jaspers, Karl; Kant, Immanuel; La Rochefoucauld, Duc François de; Leucippus and Democritus; Marx, Karl; Montaigne, Michel Eyquem de; Parmenides of Elea; Plato; Pre-Socratic Philosophy; Schopenhauer, Arthur; Voltaire, François-Marie Arouet de.

Bibliography

NIETZSCHE'S PUBLISHED WORKS AND SELECTED ENGLISH TRANSLATIONS

The definitive editions of Nietzsche's works as well as his letters and biography are the following:

Nietzsche Werke. Kritische Gesamtausgabe, edited by Giorgio Colli and Mazzino Montinari. Berlin: Walter de Gruyter, 1967ff. An English translation of the slightly abridged German critical edition *Kritische Studienausgabe* (Berlin: Walter de Gruyter, 1980) was begun under the General Editorship of Ernst Behler for Stanford University Press. The General Editorship was subsequently taken over by Bernd Magnus, and now is under the control of Alan D. Schrift and Daniel W. Conway.

Nietzsche Briefwechsel: Kritische Gesamtausgabe, edited by Giorgio Colli and Mazzino Montinari. Berlin: Walter de Gruyter, 1975ff.

Janz, Curt Paul. *Friedrich Nietzsche: Biographie*, 3 vols. Munich: C. Hanser Verlag, 1978–1979.

There are many translations available of Nietzsche's works. What follows are the full German titles, with year of publication, and the best available English translations:

Die Geburt der Tragödie aus dem Geiste der Musik (1872)

The Birth of Tragedy. Translated by Walter Kaufmann. New York: Random House, 1967.

The Birth of Tragedy and Other Writings, edited by Raymond Geuss and Ronald Speirs. Translated by Ronald Speirs. Cambridge, U.K.: Cambridge University Press, 1999.

The Birth of Tragedy. Translated by Douglas Smith. Oxford: Oxford University Press, 2000.

Unzeitgemässe Betrachtungen: I. *David Strauss, der Bekenner und Schriftsteller* (1873); II. *Vom Nutzen und Nachteil der Historie für das Leben* (1873); III. *Schopenhauer als Erzieher* (1874); IV. *Richard Wagner in Bayreuth* (1876)

Untimely Meditations. Translated by R. J. Hollingdale. Cambridge, U.K.: Cambridge University Press, 1983.

Unmodern Observations, edited by William Arrowsmith. Translated by Herbert Golder, Gary Brown, and William Arrowsmith. New Haven, CT: Yale University Press, 1990. (Includes a translation of *Wir Philologen* [We classicists]).

Unfashionable Observations. Translated by Richard T. Gray. Stanford, CA: Stanford University Press, 1995.

Menschliches, Allzumenschliches: Ein Buch für freie Geister (1878); *Menschliches, Allzumenschliches* Vol. II; *Vermischte Meinungen und Spruche* (1879); *Der Wanderer und sein Schatten* (1880)

Human, All Too Human: A Book For Free Spirits. Translated by R. J. Hollingdale. Cambridge, U.K.: Cambridge University Press, 1986. 2nd edition, 1996.

Human, All Too Human. Vol. One. Translated by Gary Handwerk. Stanford, CA: Stanford University Press, 1997.

Morgenröte: Gedanken über die moralischen Vorurtheile (1881)

Daybreak: Thoughts on the Prejudices of Morality. Translated by R. J. Hollingdale. Cambridge, U.K.: Cambridge University Press, 1982. 2nd edition, edited by Maudemarie Clark and Brian Leiter, 1997.

Die fröhliche Wissenschaft ("la gaya scienza") (1882, 1887)

The Gay Science. Translated by Walter Kaufman. New York: Vintage, 1974.

The Gay Science. Translated by Josefine Nauckhoff. Cambridge, U.K.: Cambridge University Press, 2001.

Also sprach Zarathustra: Eine Buch für Alle und Keinen (1883–1885)

Thus Spoke Zarathustra. Translated by Walter Kaufmann. New York: Viking Press, 1954.

Thus Spoke Zarathustra. Translated by R. J. Hollingdale. Harmondsworth, U.K.: Penguin, 1961.

Jenseits von Gut und Böse: Vorspiel einer Philosophie der Zukunft (1886)

Beyond Good and Evil: Prelude to a Philosophy of the Future. Translated by Walter Kaufmann. New York: Random House, 1966.

Beyond Good and Evil. Translated by Marion Faber and Robert C. Holub. Oxford: Oxford University Press, 1998.

Beyond Good and Evil. Translated by Judith Norman. Cambridge, U.K.: Cambridge University Press, 2002.

Zur Genealogie der Moral: Eine Streitschrift (1887)

On the Genealogy of Morals. Translated by Walter Kaufmann and R. J. Hollingdale. New York: Random House, 1967.

On the Genealogy of Morality and Other Writings, edited by Keith Ansell-Pearson and Carol Diethe. Translated by Carol Diethe. Cambridge, U.K.: Cambridge University Press, 1994.

On the Genealogy of Morals. Translated by Douglas Smith. Oxford: Oxford University Press, 1996.

On the Genealogy of Morality: A Polemic. Translated by Maudemarie Clark and Alan J. Swensen. Indianapolis: Hackett, 1998.

Götzendämmerung: Oder wie man mit dem Hammer philosophiert (1888)

Twilight of the Idols. Translated by Walter Kaufmann. In *The Portable Nietzsche*. New York: Viking Press, 1954.

Twilight of the Idols. Translated by R. J. Hollingdale. Harmondsworth, U.K.: Penguin, 1968.

Twilight of the Idols: or How to Philosophize with a Hammer. Translated by Duncan Large. Oxford: Oxford University Press, 1998.

Der Antichrist (1888)

The Antichrist. Translated by Walter Kaufmann. In *The Portable Nietzsche*. New York: Viking Press, 1954.

The Anti-Christ. Translated by R. J. Hollingdale. Harmondsworth, U.K.: Penguin, 1968.

Nietzsche contra Wagner (1888)

Nietzsche Contra Wagner. Translated by Walter Kaufmann. In *The Portable Nietzsche*. New York: Viking Press, 1954.

Der Fall Wagner (1888)

The Case of Wagner. Translated by Walter Kaufmann. New York: Random House, 1967.

Ecce Homo: Wie man wird, was man ist (1888)

Ecce Homo. Translated by Walter Kaufmann. New York: Random House, 1967.

Ecce Homo: How One Becomes What One Is. Translated by Duncan Large. Oxford: Oxford University Press, 2004.

Der Wille zur Macht (1883–1888)

The Will to Power. Translated by Walter Kaufmann and R. J. Hollingdale. New York: Random House, 1967.

SECONDARY WORKS

Abel, Gunter. *Nietzsche: Die Dynamik der Willen zur Macht und die ewige Wiederkehr*. Berlin: Walter de Gruyter, 1984.

Allison, David, ed. *The New Nietzsche*. New York: Dell, 1979.

Allison, David. *Reading the New Nietzsche*. Lanham, MD: Rowman and Littlefield, 2001.

Ansell-Pearson, Keith. *An Introduction to Nietzsche as Political Thinker: The Perfect Nihilist*. Cambridge, U.K.: Cambridge University Press, 1994.

Ansell-Pearson, Keith. *Nietzsche Contra Rousseau: A Study of Nietzsche's Moral and Political Thought*. Cambridge, U.K.: Cambridge University Press, 1991.

Babich, Babette E. *Nietzsche's Philosophy of Science: Reflecting Science on the Ground of Art and Life*. Albany: State University of New York Press, 1994.

Bataille, Georges. *Sur Nietzsche*. Paris: Gallimard, 1945. Translated by Bruce Boone as *On Nietzsche*. New York: Paragon House, 1992.

Blondel, Eric. *Nietzsche, le corps et la culture: La Philosophie comme généalogie*. Paris: Presses universitaires de France, 1986. Translated by Sean Hand as *Nietzsche, the Body and Culture*. Stanford, CA: Stanford University Press, 1991.

Breazeale, Daniel, ed. *Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870's*. Translated by Daniel Breazeale. Atlantic Highlands, NJ: Humanities Press, 1979.

Clark, Maudemarie. *Nietzsche on Truth and Philosophy*. Cambridge, U.K.: Cambridge University Press, 1990.

Conway, Daniel W. *Nietzsche's Dangerous Game: Philosophy in the Twilight of the Idols*. Cambridge, U.K.: Cambridge University Press, 1997.

Cox, Christoph. *Nietzsche: Naturalism and Interpretation*. Berkeley: University of California Press, 2000

Danto, Arthur C. *Nietzsche as Philosopher*. New York: Macmillan, 1965.

Deleuze, Gilles. *Nietzsche et la philosophie*. Paris: Presses universitaires de France, 1962. Translated by Hugh Tomlinson as *Nietzsche and Philosophy*. New York: Columbia University Press, 1983.

Derrida, Jacques. *Eperons: Les Styles de Nietzsche*. Paris: Flammarion, 1977. Translated by Barbara Harlow as *Spurs: Nietzsche's Styles*. Chicago: University of Chicago Press, 1978.

Fink, Eugen. *Nietzsches Philosophie*. Stuttgart, Germany: Kohlhammer, 1968. Translated by Goetz Richter as *Nietzsche's Philosophy*. London: Continuum, 2003.

Gooding-Williams, Robert. *Zarathustra's Dionysian Modernism*. Stanford, CA: Stanford University Press, 2001.

Granier, Jean. *Le problème de la vérité dans la philosophie de Nietzsche*. Paris: Éditions du Seuil, 1966.

Hatab, Lawrence J. *A Nietzschean Defense of Democracy: An Experiment in Postmodern Politics*. Chicago: Open Court, 1995.

Hayman, Ronald. *Nietzsche: A Critical Life*. Oxford: Oxford University Press, 1980.

Heidegger, Martin. *Nietzsche. Band I-II*. Pfullingen, Germany: Neske, 1961. *Vol I: The Will to Power as Art*. Translated by David Farrell Krell. New York: Harper & Row, 1979. *Vol II: The Eternal Recurrence of the Same*. Translated by David Farrell Krell. San Francisco: Harper & Row, 1984. *Vol. III: The Will to Power as Knowledge and as Metaphysics*. Translated by Joan Stambaugh, David Farrell Krell, and Frank A. Capuzzi. San Francisco: Harper & Row, 1987. *Vol. IV: Nihilism*. Translated by Frank A. Capuzzi. New York: Harper & Row, 1982.

Hollingdale, R. J. *Nietzsche: The Man and His Philosophy*. London: Routledge and Kegan Paul, 1965.

Jaspers, Karl. *Nietzsche: Einführung in das Verständnis seines Philosophierens*. Berlin: Walter de Gruyter, 1936. Translated by Charles F. Wallraff and Frederick J. Schmidtz as *Nietzsche: An Introduction to the Understanding of his Philosophical Activity.* Tucson: University of Arizona Press, 1965.

Kaufmann, Walter. *Nietzsche: Philosopher, Psychologist, Antichrist*. Princeton, NJ: Princeton University Press, 1950.

Klossowski, Pierre. *Nietzsche et la cercle vicieux*. Paris: Mercure de France, 1969. Translated by Daniel W. Smith as *Nietzsche and the Vicious Circle*. Chicago: University of Chicago Press, 1997.

Kofman, Sarah. *Nietzsche et la métaphore*. Paris: Payot, 1972. Translated by Duncan Large as *Nietzsche and Metaphor*. Stanford, CA: Stanford University Press, 1993.

Lampert, Laurence. *Nietzsche's Teaching: An Interpretation of Thus Spoke Zarathustra*. New Haven, CT: Yale University Press, 1986.

Lampert, Laurence. *Nietzsche's Task: An Interpretation of Beyond Good and Evil*. New Haven, CT: Yale University Press, 2001.

Löwith, Karl. *Nietzsches Philosophie der ewigen Wiederkehr des Gleichen*. Hamburg: Meiner, 1978. Translated by J. Harvey Lomax as *Nietzsche's Philosophy of the Eternal Recurrence of the Same*. Berkeley, CA: University of California Press, 1997.

Magnus, Bernd. *Nietzsche's Existential Imperative*. Bloomington: Indiana University Press, 1978.

Magnus, Bernd, et al. *Nietzsche's Case: Philosophy And/As Literature*. New York: Routledge, 1993.

Magnus, Bernd, and Kathleen M. Higgins, eds. *The Cambridge Companion to Nietzsche*. Cambridge, U.K.: Cambridge University Press, 1996.

Müller-Lauter, Wolfgang *Nietzsche: Seine Philosophie der Gegensätze und die Gegensätze seiner Philosophie*. Berlin: Walter de Gruyter, 1971. Translated by David J. Parent as *Nietzsche: His Philosophy of Contradictions and the Contradictions of His Philosophy*. Urbana: University of Illinois Press, 1999.

Nehamas, Alexander. *Nietzsche: Life as Literature*. Cambridge, MA: Harvard University Press, 1985.

Oliver, Kelly, and Marilyn Pearsall, eds. *Feminist Interpretations of Friedrich Nietzsche*. University Park: Pennsylvania State University Press, 1998.

Parkes, Graham. *Composing the Soul: Reaches of Nietzsche's Psychology*. Chicago: University of Chicago Press, 1994.

Richardson, John. *Nietzsche's New Darwinism*. Oxford: Oxford University Press, 2004.

Richardson, John. *Nietzsche's System*. Oxford: Oxford University Press, 1996.

Roberts, Tyler T. *Contesting Spirit: Nietzsche, Affirmation, Religion*. Princeton, NJ: Princeton University Press, 1998.

Schaberg, William H. *The Nietzsche Canon: A Publication History and Bibliography*. Chicago: University of Chicago Press, 1995.

Schacht, Richard. *Making Sense of Nietzsche: Reflections Timely and Untimely*. Urbana: University of Illinois Press, 1995.

Schacht, Richard. *Nietzsche*. London: Routledge and Kegan Paul, 1983.

Schacht, Richard, ed. *Nietzsche, Genealogy, Morality: Essays on Nietzsche's Genealogy of Morals*. Berkeley: University of California Press, 1994.

Schrift, Alan D. *Nietzsche and the Question of Interpretation: Between Hermeneutics and Deconstruction*. New York: Routledge, 1990.

Schrift, Alan D. *Nietzsche's French Legacy: A Genealogy of Poststructuralism*. New York: Routledge, 1995.

Schrift, Alan D., ed. *Why Nietzsche Still? Reflections on Drama, Culture, and Politics*. Berkeley: University of California Press, 2000.

Schutte, Ofelia. *Beyond Nihilism: Nietzsche Without Masks*. Chicago: University of Chicago Press, 1984.

Sedgwick, Peter R., ed. *Nietzsche: A Critical Reader*. Oxford: Blackwell, 1995.

Shapiro, Gary. *Alcyone: Nietzsche on Gifts, Noise and Women*. Albany: State University of New York Press, 1991.

Shapiro, Gary. *Nietzschean Narratives*. Bloomington: Indiana University Press, 1989.

Small, Robin. *Nietzsche in Context*. Aldershot, England: Ashgate, 2001.

Solomon, Robert C., and Kathleen M. Higgins. *Reading Nietzsche*. New York: Oxford University Press, 1988.

Staten, Henry. *Nietzsche's Voice*. Ithaca, NY: Cornell University Press, 1990.

Strong, Tracy B. *Friedrich Nietzsche and the Politics of Transfiguration*. Berkeley: University of California Press, 1975. Rev. ed. Urbana: University of Illinois Press, 2000.

Thiele, Leslie Paul. *Friedrich Nietzsche and the Politics of the Soul: A Study of Heroic Individualism*. Princeton, NJ: Princeton University Press, 1990.

Tongeren, Paul van. *Reinterpreting Modern Culture: An Introduction to Friedrich Nietzsche's Philosophy*. West Lafayette, IN: Purdue University Press, 2000.

Warren, Mark. *Nietzsche and Political Thought*. Cambridge, MA: MIT Press, 1988.

Alan D. Schrift (2005)

NIHILISM

The term *nihilism* appears to have been coined in Russia sometime in the second quarter of the nineteenth century. It was not, however, widely used until after the appearance of Ivan Turgenev's highly successful novel *Fathers and Sons* in the early 1860s. The central character, Bazarov, a young man under the influence of the "most advanced ideas" of his time, bore proudly what most other people of the same period called the bitter name of nihilist. Unlike such real-life counterparts as Dmitri Pisarev, Nikolai Dobrolyubov, and Nikolai Chernyshevskii, who also bore the label, Bazarov's interests were largely apolitical; however, he shared with these historical personalities disdain for tradition and authority, great faith in reason, commitment to a materialist philosophy like that of Ludwig Büchner, and an ardent desire to see radical changes in contemporary society.

An extreme statement by Pisarev of the nihilist position as it developed in the late 1850s and 1860s in Russia is frequently quoted: "Here is the ultimatum of our camp: what can be smashed should be smashed; what will stand the blow is good; what will fly into smithereens is rubbish; at any rate, hit out right and left—there will and can be no harm from it" (quoted in Avrahm Yarmolinsky, *Road to Revolution,* p. 120). Bazarov echoes this idea, though a bit feebly, when he accepts a description of nihilism as a matter of "just cursing."

Use of the term spread rapidly throughout Europe and the Americas. As it did, the term lost most of its anarchistic and revolutionary flavor, ceasing to evoke the image of a political program or even an intellectual movement. It did not, however, gain in precision or clarity. On the one hand, the term is widely used to denote the doctrine that moral norms or standards cannot be justified by rational argument. On the other hand, it is widely used to denote a mood of despair over the emptiness or triviality of human existence. This double meaning appears to derive from the fact that the term was often employed in the nineteenth century by the religiously oriented as a club against atheists, atheists being regarded as ipso facto nihilists in both senses. The atheist, it was held, would not feel bound by moral norms; consequently, he would tend to be callous or selfish, even criminal. At the

same time he would lose the sense that life has meaning and therefore tend toward despair and suicide.

ATHEISM

There are many literary prototypes of the atheist-nihilist. The most famous are Ivan in Fëdor Dostoevsky's *Brothers Karamazov* and Kirilov in Dostoevsky's *The Possessed.* It was into Ivan's mouth that Dostoevsky put the words, "If God does not exist, everything is permitted." And Dostoevsky made it clear that it was Ivan's atheism that led him to acquiesce to his father's murder. Kirilov was made to argue that if God does not exist, the most meaningful reality in life is individual freedom and that the supreme expression of individual freedom is suicide.

Friedrich Nietzsche was the first great philosopher—and still the only one—to make extensive use of the term *nihilism.* He was also one of the first atheists to dispute the existence of a necessary link between atheism and nihilism. He recognized, however, that as a matter of historical fact, atheism was ushering in an age of nihilism. "One interpretation of existence has been overthrown," Nietzsche said, "but since it was held to be *the* interpretation, it seems as though there were no meaning in existence at all, as though everything were in vain" (*Complete Works,* Edinburgh and London, 1901–1911, Vol. XIV, p. 480). Albert Camus later dealt with this historical fact at some length in *The Rebel* (1951).

The tendency to associate nihilism with atheism continues to the present. It is to be found, for instance, in a work by Helmut Thielicke titled *Nihilism,* which first appeared in 1950. During the course of the twentieth century, however, the image of the nihilist changed, with a corresponding change in the analysis of nihilism's causes and consequences. Professor Hermann Wein of the University of Göttingen wrote, for instance, that the members of the younger generation of his time tended to think of the nihilist not as a cynical or despairing atheist but as a robotlike conformist. For them nihilism is caused not so much by atheism as by industrialization and social pressures, and its typical consequences are not selfishness or suicide but indifference, ironical detachment, or sheer bafflement. The literary prototypes are not the romantic heroes of Dostoevsky but the more prosaic and impersonal heroes of Robert Musil's *Man without Qualities* (first volumes published 1931–1933) or Franz Kafka's *The Trial* (1925).

MORAL SKEPTICISM

If by nihilism one means a disbelief in the possibility of justifying moral judgments in some rational way and if philosophers reflect the intellectual climate of the times in which they live, then our age is truly nihilistic. At no period in Western history, with the possible exception of the Hellenistic age, have so many philosophers regarded moral statements as somehow arbitrary. For many Continental philosophers, especially the atheistic existentialists, moral values are products of free choice—that is, of uncaused, unmotivated, and nonrational decisions. The most notable statement of this view is in *Being and Nothingness* (1943) by Jean-Paul Sartre. In England and America, most philosophers tend to the view known as emotivism, according to which moral statements are ultimately and essentially products of pure social conditioning or brute feeling. The most noted, though not the most extreme, representatives of this position are A. J. Ayer and Charles Stevenson.

It is impossible to state here with reasonable detail and accuracy the positions so summarily described in the last paragraph, much less to discuss their logical merits. For an understanding of nihilism, however, it is important to note how these positions relate to the ideas of those to whom nihilism of this kind is anathema. As already indicated, the most vociferous antinihilists were originally theologians, like Dostoevsky, who feared that disbelief in God would lead to selfishness and crime. If, they argued, there is no divine lawgiver, each man will tend to become a law unto himself. If God does not exist to choose for the individual, the individual will assume the former prerogative of God and choose for himself. For these antinihilists the principal enemy would have been Sartre. The later antinihilists, however, tend to save their fire for the emotivists, whom they accuse of sanctioning moral indifference and mindless conformity. If all moral codes are essentially matters of feeling and social pressure, then no one would be better or worse than another. The wise man, like the Sophists of Plato's day, would simply adjust as best he could to the code of the society in which he happened to be living. John Dewey's fervid insistence upon critical individual intelligence as the prime agent of social and moral reconstruction places him squarely in the second group of antinihilists.

Whether belief in atheistic existentialism or emotivism does in fact have the kinds of consequences suggested above is not at issue here. The point is simply that antinihilists of the older variety do not regard conventional morality, especially in its other-regarding aspects, as adequately justified unless it has a cosmic or divine sanction, whereas more contemporary antinihilists do not regard any moral code as adequately justified unless there is some standard or touchstone more universal than

pure feeling or social pressure to which it may be shown to conform. The pertinent question here is whether the antinihilists have a good case for these views.

It would appear that the demand for justification of conventional moral rules by appeal to a divine or cosmic power cannot be logically admitted without abandoning widespread and deeply felt notions about the nature of moral justification. If the higher power that presumably legitimizes our moral code is by definition good and just, an appeal to that power would involve us in a vicious circle. How would we know that that power was good and just unless there were some purely human ideas about the good and the just to which we felt entitled independently of that power's sanction? If, on the other hand, the presumed higher power is not by definition good or just, if, for instance, it were defined merely as a creator and sustainer of life, by what right could we appeal to it to legitimize our moral views? Might or power, even the power to create and sustain life, is not to be confused with right or legitimacy.

The demand that moral codes be justified by more universal standards than pure feeling or social dictate is, on the contrary, much more consonant with widespread, intuitive notions about the nature of moral justification. If social pressure is taken as the touchstone of morality, we once again court a confusion between might and right; if feeling is taken as the touchstone, we must apparently abandon not only the notion of a universal morality, feelings being notoriously fluctuating and individual, but also the notion that one of the functions of morality is to refine, direct, and control individual feelings. It may, of course, be the case that there is no universal morality and that whatever power morality possesses must derive from individual feeling and social conditioning alone. It would be surprising, however, if even the emotivists did not experience a certain chagrin that the truth in ethical theory should be so contrary to human hopes.

MEANINGLESS OF LIFE

Passing to the second meaning of the term *nihilism,* we find that the pertinent questions are less logical or technically philosophical than psychological or sociological. There are two questions here, corresponding to the two forms of antinihilism. Is it true that a loss of faith in God or cosmic purposes produces a sense of despair over the emptiness and triviality of life, consequently stimulating selfishness and callousness? Is it true that industrialization and conformist social pressures have trivialized life in a similar way, causing us to adopt an attitude of ironic detachment? A negative answer to these questions would appear to fly in the face of most contemporary social criticism and analysis as well as the testimony of most contemporary literature.

It is doubtful, however, whether a simple yes would be a proper response to the first question. When it is assumed that humankind needs a sense of divine or cosmic purpose in order to lead a rich and morally wholesome life, one is generalizing far beyond the evidence. The most that the evidence can be made to support is that relatively large numbers of people in certain societies at certain times have felt this need. No one who has read, for instance, Lev Tolstoy's account of his religious crisis in middle age could doubt the depth of his despair or the reality of his need for a vital relationship to an eternal being. One can reasonably doubt, however, whether that need and despair spring from universal and firmly rooted human aspirations. Some psychologists regard Tolstoy's conversion crisis as a symptom of involutional melancholia, and there are many who believe it to be a consequence of Tolstoy's social position as a member of Russia's decaying aristocracy.

Bertrand Russell went through a similar crisis earlier in life. He not only survived that crisis without reverting to faith in God or cosmic purpose; he also survived it, as his essay "A Free Man's Worship" (1902) attests, by deliberately espousing a world outlook that emphasizes the finitude and cosmic isolation of humankind. And no one who is familiar with the facts of his life would dare to suggest that the later Russell was less morally earnest than the young believer or less wholeheartedly and happily engaged in the process of living.

Those who attribute the nihilistic malaise of our time to industrialization and conformity are less vulnerable to the charge of overgeneralization. This is not because they limit their analysis to a given historical epoch, for they, too, are making an implicit generalization about universal human needs. Their point is that all people need, if they are to be whole and healthy, the sense that they can by a unique and personal effort contribute to the social process and that society will appreciate and reward this individual effort. This generalization is less vulnerable than the first simply because there is more evidence for it. Novels and biographies, ethnographic reports and individual clinical histories, not to mention commonsense attitudes of most men in all societies at all historical periods, tend to support it. And the issue raised by nihilism in this sense of the term is one of the great unresolved political and social problems of the twentieth and twenty-first centuries. Whether philosophers in their

professional capacity are competent to contribute to its solution is a question we shall not attempt to answer here.

See also Atheism; Ayer, Alfred Jules; Camus, Albert; Chernyshevskii, Nikolai Gavrilovich; Dewey, John; Dostoevsky, Fyodor Mikhailovich; Kafka, Franz; Life, Meaning and Value of; Moral Skepticism; Nietzsche, Friedrich; Pessimism and Optimism; Pisarev, Dmitri Ivanovich; Russell, Bertrand Arthur William; Russian Philosophy; Sartre, Jean-Paul; Stevenson, Charles L.

Bibliography

On Russian nihilism of the 1850s and 1860s, see Avrahm Yarmolinsky, *Road to Revolution* (New York: Macmillan, 1959), and Ivan Turgenev, *Fathers and Sons* (New York, 1958).

For interpretations of nihilism reflecting a theological or religious point of view, see Helmut Thielicke, *Nihilism* (New York: Harper, 1961); Ernst Benz, *Westlicher und östlicher Nihilismus* (Stuttgart: Evangelisches, 1948); Nikolai Berdyaev, *Sinn und Shicksal des russischen Kommunismus* (1937); and Lester G. Crocker, *An Age of Crisis* (Baltimore: Johns Hopkins Press, 1959).

For emotivism see A. J. Ayer, *Language, Truth and Logic,* 2nd rev. ed. (New York: Dover, 1952); Charles Stevenson, *Ethics and Language* (New Haven, CT: Yale University Press, 1944); and Ingemar Hedenius, "On Law and Morals," in *Journal of Philosophy* 56 (1959): 117–125.

For existentialist nihilism see Jean-Paul Sartre, *Being and Nothingness* (New York: Philosophical Library, 1956).

See also Friedrich Nietzsche, *Will to Power,* in *Complete Works,* edited by Oscar Levy (New York: Russell and Russell, 1964); Albert Camus, *The Rebel* (New York: Knopf, 1954); Hermann Wein, "Discussion on Nihilism," in *Universitas* 6 (2): 173–182; C. F. von Weizsäcker, *The History of Nature* (Chicago: University of Chicago Press, 1949), pp. 71ff.; Bertrand Russell, "A Free Man's Worship," in his *Mysticism and Logic* (London: Allen and Unwin, 1917), and "What I Believe," in his *Why I Am Not a Christian* (New York: Simon and Schuster, 1957), Ch. 3; Lev Tolstoy, "My Confession" and "What Is to Be Done," in *Complete Works* (New York, 1898), Vol. VII; and Hermann Rauschning, *The Revolution of Nihilism* (London, 1939).

Robert G. Olson (1967)

NIRVĀṆA

Nirvāṇa is the ultimate goal of Buddhist practice, although there has been disagreement among Buddhists concerning its nature and the means of attaining it. The word derives from a Sanskrit verbal root meaning "to blow" and a prefix meaning "out." The underlying meaning of the word is traditionally explained as expressing one of two metaphors. The first is that the term means the act of blowing out or extinguishing, as of a flame. The second is that it means the act of being cooled down, as by a breeze. The two metaphors have in common the notion of fire or heat as a source of pain that is alleviated by a breeze. So the principal characteristic of nirvāṇa is relief from pain and the prevention of future pain through the eradication of its root causes. It is, in other words, the permanent release from the conditions that make pain possible, both physical pain and forms of psychological suffering such as sadness, grief, despondency, melancholy, frustration, and anxiety. Traditionally nirvāṇa is said to occur in two stages: the extinction of the causes of rebirth and the end of rebirth itself. For ease of exposition, the latter will be discussed first.

NIRVĀṆA AS THE END OF REBIRTH

The Buddhist doctrine of nirvāṇa arose in the context of a view of the world that was common throughout India at the time when Buddhism was founded, in the sixth century BCE. According to that view, the world is both beginningless and endless and constantly changing. Among the many kinds of change in this world are the various stages undergone by a living being, or, more properly, an individual continuum of conscious. Such a being is born, matures, decays and eventually dies. When a living being dies, it does not cease to exist; rather it is transformed into another living being that also undergoes birth, maturity, decay and death. The cycle of rebirths that any given being undergoes is beginningless. The doctrine of Buddhism, and of many other systems of thought in ancient India, asserts that the cycle can, however, come to an end, provided that the conditions that keep the cycle going are eliminated. The name that is given to the end of the cycle of rebirths for any given continuum of consciousness is final nirvāṇa. It is described in Buddhist texts as the cessation of the process of being reborn into any kind of existence in any realm in the cosmos. Since all kinds of existence are at least potentially painful, the only way of eliminating the very possibility of experiencing physical or psychological pain is to stop existing altogether.

NIRVĀṆA AS THE EXTINCTION OF THE CAUSES OF REBIRTH

According to Buddhist doctrine, the ultimate cause of rebirth is simply the desire to continue existing. When a deity or human being or animal dies wishing that life could continue, life does continue. The consciousness of the dying person then finds itself associated with a different body, which may or may not belong to the same bio-

logical species as the body that has just died. The type of body with which the continuing consciousness finds itself associated is determined by the overall mentality of the consciousness continuum at the time of the death of the previous body. What all rebirth has in common is the desire to continue existing, and this is the consequence of delusion, a fundamental misunderstanding about the real nature of existence. The real nature of existence is that every existing thing is characterized by impermanence. Because of this impermanence, nothing that anyone experiences endures, and therefore nothing, however pleasant it may be, can be a source of enduring satisfaction. Because every satisfactory experience comes to an end, it is ultimately disappointing and unsatisfactory.

These two characteristics of existence, impermanence and disappointment, give rise to a third feature of existence, namely, that no existing thing is part of an abiding self, and nothing can ever be owned. The delusions that fuel the desire to continue existing, therefore, are the erroneous beliefs that anything can be permanent, satisfactory, and either part of oneself or a potential piece of property that one can own. Nirvāṇa, then, is the elimination of those delusions by understanding existence as it really is. This correct understanding is called awakening or enlightenment. All of Buddhist doctrine and practice, then, can be seen as a process of working toward the state of enlightenment that makes final nirvāṇa possible. Enlightenment is therefore described as a name that is given to the absence of specific delusions, in the same way that final nirvāṇa is a name given to the absence of further rebirth.

STAGES LEADING TO NIRVĀṆA

According to most schools of Buddhism, the path to enlightenment is incremental. One does not rid oneself of all delusion at once, because delusion itself is part of a complex mentality that consists of various vices that are caused by and that in turn reinforce the habit of having a naive and superficial perspective on one's experience. Although the specific manifestations of superficiality and its attendant vices differ for every individual, there is said to be a general pattern of how progress to enlightenment is made.

To understand the stages along the way to nirvāṇa, it is helpful to know that Buddhist tradition enumerates ten mental habits that obstruct peace of mind. They are

(1) the opinion that complex objects are real,

(2) suspicion or intense doubt,

(3) abiding by rules and vows for the sole purpose of gaining merit for oneself,

(4) desire for sensual pleasure,

(5) malevolence,

(6) passion for material things and for material forms of existence,

(7) passion for spiritual or nonmaterial things, such as meditative states, and for nonmaterial forms of existence,

(8) conceit, which is explained as the habit of constantly comparing and measuring oneself against others,

(9) agitation or excitement,

(10) misconception or ignorance, which includes any kind of failure to see things as they really are.

The first stage on the path to nirvāṇa is reached when the first three of these obstacles have been eliminated, and it is claimed that all three of these first three are eliminated at the same time, since the second and third are effects of the first. This first stage is also said to be reached more easily when one keeps good company, that is, the company of others who have reached at least the first stage. For this reason, much of Buddhist practice centers on maintaining a community of men and women who are helping one another strive for nirvāṇa. Although much of that struggle requires personal effort and a thorough knowledge of one's own mentality, the individual's efforts are said to be nearly impossible without the support of a community of like-minded people.

The second and third stages of the path to nirvāṇa are reached when one reduces and then eliminates the fourth and fifth obstacles. The final goal, nirvāṇa itself, is reached when one has eliminated all ten obstacles, and especially the passions for both material and spiritual states of being. A person who has attained nirvāṇa is called an *arhant* (feminine *arhatī*), which literally means "a person worthy of admiration." The *arhant* is someone who has all the characteristics of a buddha and differs from a buddha only in having required instruction to achieve nirvāṇa, whereas a buddha achieves nirvāṇa without ever having been taught how to attain it. All people who have reached any of these four stages that culminates in arhanthood are collectively known as nobles (*ārya*), and the path to nirvāṇa is known as the noble path or the path of the nobles.

KNOWLEDGE OF NIRVĀṆA

According to a formula often repeated in canonical sources, when a person attains nirvāṇa, then he or she knows "what needed to be done has been done, and I shall never again be reborn in any realm in any form." This raises the difficult question of how one can know an absence and especially a future absence. Obviously, one cannot directly experience an absence, nor can one directly experience anything that takes place, or fails to take place, in the future. The knowledge of the absence of one's future rebirths, then, must be an inference of some kind. Even the knowledge that in the present life there will never again arise the ten obstacles enumerated above must be an inference of some kind. Just how such an inference might work and what kind of inference is involved occupied the attention of Buddhist scholastics from the time of Dharmakīrti on.

THE SPECIAL ONTOLOGICAL STATUS OF NIRVĀṆA

According to Buddhist teachings, all conditioned things are impermanent, because the conditions upon which something depends can disappear, and when they disappear, so does anything that depends on them. nirvāṇa, however, is said to be a permanent achievement. If it is permanent, then it cannot be conditioned. From these considerations one of two possibilities follow. Either all things are conditioned, in which case nirvāṇa cannot be a thing at all, or nirvāṇa is an exception to the otherwise universal rule that all things are conditioned. Both of these possibilities have had their advocates among Buddhist scholastics. Those who regarded nirvāṇa as an unconditioned thing came to characterize nirvāṇa as a permanent entity that is constantly lucid and blissful, or as a state of being aware of a permanently lucid and blissful and essentially transcendent reality.

LATER DOCTRINAL DEVELOPMENTS: NONABIDING NIRVĀṆA AND HAPPY REALMS

Several centuries after the founding of the Buddhist community, a movement arose that placed an emphasis on a kind of virtuoso known as a bodhisattva. The term itself originally referred to a person who was dedicated to becoming a buddha and thus referred to the Buddha Gautama (the founder of Buddhism) in his previous lives. In an extension of that original meaning, the term *bodhisattva* came to be applied to anyone who had come to realize that suffering is present in all realms of the universe, that the vast majority of sentient beings lack the capacity to achieve nirvāṇa on their own strengths, and that they therefore require the help of someone dedicated to helping others attain nirvāṇa. A bodhisattva is a person who not only realizes all that but also vows not to attain final nirvāṇa until all other sentient beings have also attained it. Texts dealing with the bodhisattva ideal say that a bodhisattva may either postpone his or her own attainment of nirvāṇa until others have attained it or, preferably, may attain nirvāṇa and then renounce it in order to remain among sentient beings in need of help. Attaining nirvāṇa and then renouncing it is said to be preferable because the bodhisattva who does this already knows the way and can therefore better show others. This nirvāṇa that the bodhisattva attains and then renounces is called nonabiding nirvāṇa.

Another doctrine that began with the realization that most beings are incapable of attaining nirvāṇa through their own discipline alone was the myth that some buddhas have attained final nirvāṇa only after establishing special realms in which there are no environmental obstacles to tranquility. In such a realm, known as a happy land (*sukhavatī bhūmi*) or, following Chinese translations of the Sanskrit term, a pure land, all who abide there are surrounded by inspirational teachings. Even the babbling of brooks and the chirping of birds are discourses on virtue. In the absence of a painful external environment and in the presence of incessant sermons, the residents of the happy lands quickly attain final nirvāṇa. Two mythological buddhas who are said to have established happy lands are Amitābha and Akshobhya, the former of whom became the focus of an extensive cult in China and East Asia.

See also Buddhism; Mysticism, History of; Reincarnation.

Bibliography

Collins, Steven. *Nirvāṇa and Other Buddhist Felicities.* Cambridge, U.K.: Cambridge University Press, 1998.

Harvey, Peter. *An Introduction to Buddhism*. Cambridge, U.K.: Cambridge University Press, 1990.

Harvey, Peter. *The Selfless Mind: Personality, Consciousness and Nirvāṇa in Early Buddhism*. Surry, U.K.: Curzon Press, 1995.

Williams, Paul. *Mahāyāna Buddhism: The Doctrinal Foundations*. London and New York: Routledge, 1989.

Richard P. Hayes (2005)

NISHI AMANE
(1829–1897)

Nishi Amane, the pioneer in bringing Western philosophy to Japan, was born in Tsuwano, Shimane prefecture.

After the usual Confucian training he went to Edo (Tokyo) for further studies and was attached to Bansho Torishirabe-sho (Center for the Investigation of Western Books). In 1862 he was sent with other promising Japanese to Holland to study Western law and military science. In Holland his interest in philosophy was reawakened, and with his friend Tsuda Masamichi he became acquainted with the positivism of Auguste Comte, the utilitarianism of J. S. Mill, and Immanuel Kant's *On Eternal Peace.* He returned to Japan in 1865 and was appointed to the Kaisei School in Edo, where the government of the shogun requested him to translate books on law. After the Meiji restoration, Nishi was put in charge of educational matters for the Ministry of Military Affairs. At this time he also wrote most of his philosophical books. He became a member of the *Meirokusha,* the group of leading intellectuals of the time, who advocated Western culture and mores. Nishi was several times president of the Tokyo Academy. He was made a baron and was appointed to the upper chamber of the legislature, the House of Peers, in 1890.

Nishi's importance as the "father" of Western philosophy in Japan lies in the new terminology he created—from his Japanese term for philosophy, *tetsugaku,* to his various translations—and in the original works that established a new tradition of speculative thinking. His positivist bent is revealed in *Reikon ichigenron* (Monism of the soul), one of his earlier works. More famous are his panoramic treatments of Western learning and philosophy in *Hyakugaku renkon* (Encyclopedia; written in 1874), a kind of philosophical or cultural dictionary, and *Haykuichi shinron* (A new theory on the many doctrines; written in 1874). In these Nishi prefers Mill's inductive method to Comte's positivism. In 1874 Nishi also wrote *Chichi keimō* (Logic, an introduction), the first of its genre in Japan. His utilitarian ethics is clearly manifested in "Jinsei sampō-setsu" (The three treasures theory of man's life), which appeared in the *Meiroku Journal* in 1875. He replaced Confucian ethics with a quest for the three treasures: health, wealth, and knowledge.

As a translator Nishi has to his credit Mill's *Utilitarianism* and a work titled *Mental Philosophy* by Joseph Haven, an American philosopher influenced by Scottish realism.

In later life Nishi became more conservative in his view of Western ideas, an attitude consonant with the country's post-1886 reaction against ultra-Westernization. As a director of a teacher's college, Shihan Gakkō, he proposed a combination of East and West in ethics; but in the last analysis he remains an expositor of Western philosophy who never really tried to combine East and West in his thought and writing.

See also Comte, Auguste; Japanese Philosophy; Kant, Immanuel; Mill, John Stuart; Mill's Methods of Induction; Positivism; Utilitarianism.

Bibliography

Nishi's collected works are available in two editions. One is *Nishi Amane zenshū* (The complete works of Nishi Amane), 3 vols., edited by Okube Toshiaki (Tokyo, 1960–). Vol. I contains Nishi's philosophical works. The other edition is *Nishi tetsugaku chosakushū* (Collected philosophical works of Nishi Amane), edited by Asō Yoshiteru (Tokyo, 1933).

Studies of Nishi can be found in *Japanese Thought in the Meiji Era,* edited by M. Kōsaka, translated into English by D. Abosch (Tokyo: Pan-Pacific Press, 1958), pp. 99–113; Gino K. Piovesana, *Recent Japanese Philosophical Thought, 1862–1962* (Tokyo: Enderle Bookstore, 1963), pp. 5–18; and R. F. Hackett, "Nishi Amane, A Tokugawa-Meiji Bureaucrat," in *Journal of Asian Studies* 18 (2) (1959): 213–225.

Gino K. Piovesana, S.J. (1967)

NISHIDA, KITARŌ
(1870–1945)

One of modern Japan's most prominent philosophers, Nishida was born in the village of Unoke, located on the Japanese Sea near Kanazawa, which was the capital of the Ishikawa prefecture. He attended the Prefecture Gymnasium in Kanazawa, where he began a lifelong friendship with Teitaro (Daisetz) Suzuki. He then enrolled at the University of Tokyo, choosing philosophy over mathematics, in which he was quite gifted, and studied Western philosophy there from 1891 until 1894. After completing his studies with a thesis on David Hume, Nishida returned to his home, married, and devoted himself intensely for about ten years after 1897 to the practice of Zen.

In 1899 he was appointed as a teacher at the Forth Senior High School (previously the Prefecture Gymnasium) in Kanazawa, where he taught logic, ethics, psychology, and German until 1909. During this period, which Nishida would later characterize as the best of his life, he laid the solid and fertile groundwork for his subsequent philosophical work, a groundwork based on the unusual combination of Western philosophy and Zen. Each day he faithfully practiced Zen meditation and sitting exercises (*Zazen*), but he also worked through the main texts of Western philosophy from Plato and Aristotle through to Henri-Louis Bergson, William James,

Heinrich Rickert, and Alexius Meinong. His own philosophy was to emerge out of the seemingly impossible combination of these two parallel directions.

In 1910 he received an appointment as assistant professor for ethics at the University of Kyoto. The following year his first work titled *An Inquiry into the Good* appeared, in which a philosophy of pure experience is developed. It is the first monumental philosophical work in Japan according to the full sense of the word *philosophy* as it was imported into Japan from the West. In 1913 he was named full professor for the philosophy of religion and then in 1914 full professor in philosophy. This was the beginning of what has come to be known as the Nishida period in the philosophy department at the University of Kyoto, during which a philosophical community arose around him with both high scholarly standards and close personal attachments. Hajime Tanabe, later Nishida's successor, and Tetsurō Watuji were recruited by him. Similarly successful students arose under his tutelage, including Kiyoshi Miki and Keiji Nishitani. The philosophical department flourished during this period and became a significant factor in the intellectual and academic life of modern Japan. This circle of scholars came to be known as the Kyoto School. He retired from the university in 1928.

His family life during this time, however, was difficult and painful, as he recalled on the occasion of his retirement: "For ten years, I have pursued my scholarly work while faced with continually unbearable, unfortunate circumstances in my family, which has been very difficult for me." In 1920 he lost his beloved first son, in 1925 he lost his wife, who had been bedridden at home for six years as the result of a serious stroke. One of his daughters suffered with tuberculosis for several years. Two others were hospitalized for acute typhoid fever, one of whom never completely recovered. Earlier, during the Kanazawa period, he had already lost his brother and two young daughters. In several philosophical essays, Nishida wrote, "What impels one towards philosophy is the sorrow and pain of human life." Not wonder that there is something rather than nothing, not methodical doubt as a means to achieve certainty, but rather the fate of human life as a whole on earth motivated Nishida to pursue philosophy. Nishida's basic question is, "What is the structure of the actual world into which we are born, in which we labor, and in which we die? What is our self in this actual world?"

Nishida's concern is not only the life-world, not only the historical world, but also the world of life and death. Nishida is not concerning solely with the self that lives in the world, but rather the whole self that is born, lives, and dies. Sorrowful, painful events in human life tear open the world. This tear or rift opens up a window and gives access to the profundity of the world. Nishida says that "Grasping the common everydayness of our lives most profoundly leads to the most profound philosophizing." This profundity is nothing other than the profundity of everyday life. Nishida speaks of eschatological everydayness.

In the middle of the painful sorrows of his life, Nishida could say: "The ground of my heart, infinitely deep, will not be reached by all of the waves of joy and cares." For Nishida profundity or depth was experienced profundity. In his calligraphic work, for which he also counts as an artist, Nishida expresses a beautiful power rising out of this profundity. In spite of the difficult circumstances he faced in his life, he worked continuously every day. Even in the year in which he retired, he published five essays, including *Predicative Logic*, *The Place Wherein One Sees Oneself and the Place of Consciousness*, and *The Intelligible World*. His creative powers were sustained up to the end of his life, whereby the pathway for his thinking did not get any easier.

After his retirement Nishida spent half of each year in Kyoto and half in Kamakura at the seashore. He said, "I love the sea. There is something infinite that is suspended and moves in the sea." One student characterized Nishida's philosophy as a philosophy of the sea. His boyhood friend Suzuki also lived in Kamakura after he had returned from the United States so there the two of them often met for conversations in the space between Zen and philosophy, and Nishida attributed much in his philosophy to the influence of Suzuki. After his second marriage in 1931, Nishida's family situation was much better, but his concerns over Japan's worsening internal political situation and its external policies became increasingly grave. In 1939 the Second World War began and in 1941 the Pacific War with the United States began, which plummeted Japan on the war to its catastrophic defeat in August of 1945.

In May of 1945, Nishida wrote to Suzuki regarding the impending defeat: "Things are happening as we always feared they would. A state that is based on military power will perish by military power." As he was intensely searching for the possibility of a world culture that could unite humanity in the newly unified world that was to come after the world war, Nishida died on June 7, 1945, on account of an acute kidney infection. On his desk lay the unfinished manuscript of an essay *Concerning My Logic*. Nishida worked up until the last day of his life.

During the year before he died, he wrote articles titled *Concerning A Philosophy of Religion Governed by the Pre-established Harmony*, *Life*, *Philosophical Foundations of Mathematics*, and *The Logic of Place and the Religious World View*.

HIS WORK

In 1926 the essay *Place* (*Basho*) appeared. Concerning that article, Nishida wrote, "It seems to me that I attained my final standpoint with the notion of 'place'." Nishida's philosophy can in a real sense be characterized as a philosophy of place. The basic idea behind the notion of place is: Everything that is, is located in a place. *Being* means *being in*. The proposition "S is P" means in truth that "S is in P." Nishida states one time simply and concretely, "Place is where we are located." *Place* for Nishida, then, corresponds to what Martin Heidegger called *world* as a component of being in the world. For Nishida, place consists both of the place of being and the place of the absolute nothing in the sense that the place of being is surrounded by the place of absolute nothing. The place of being as the place of limited disclosedness is located within the place of nothing as the unlimited disclosedness, infinite openness. *Place* thereby has a twofold disclosedness for us. Those of us who find ourselves in a place find ourselves not only in a world, but also in the unlimited openness that surrounds the world, a view that is different from Heidegger's. Nishida explicitly discusses the *we* as something that is located in a place in his essay *I and Thou* (1932).

According to Nishida, *I and thou* means that I am what I am in that I am nothingness in the unlimited openness, and conversely, *I and thou* means that you are what you are in that you are in the nothingness of unlimited openness. Nishida views the relationship differently than Martin Buber in that the I-thou is an aspect, the face-to-face aspect of the full reality of what is located in a place, a reality that consists in the fact that this one single individual and this other single individual are both in contradiction and in unity based on the abyss of the absolute nothingness where there is neither *I* nor *thou*. The basic traits of the notion of place according to Nishida can only be understood in correspondence to an originary pure experience because the notion of place is developed out of this experience.

Nishida's philosophy of pure experience arose through an original and radical encounter between West and East. There is a qualitative divide between the thinking of Western philosophy and the nonthinking in Zen. This rift inside Nishida himself, where both philosophy and Zen coexisted, threatened to rip him apart, but instead it came to serve as a magnetic field in which philosophy and Zen actually touched and permeated each other. This is where Nishida's philosophy was born, a philosophy of another beginning. For a philosophy of pure experience, it is crucial to explain everything through the fact that the only real reality is pure experience. In attempting to explain everything within a single context, Nishida orients himself on Western philosophy; his pure experience, however, comes from Zen.

Pure experience is not a monadic substance-like foundational entity, but rather an original occurrence of experiencing, an event like the following: "In the moment of seeing, of hearing, still without reflections such as 'I see flowers' and without judgments like 'These flowers are red,' in this moment of momentary seeing or hearing, there is neither subject nor object." This immediately experiencing experience occurs as the ground of the truly real reality because in immediate seeing and hearing the undifferentiatedness that obtains before splitting into difference is at work. Here, a direct connection between the empirical and the metaphysical is revealed in a unique way. For Nishida, the metaphysical does not disclose itself beyond experience but rather within experience, that is, within the immediately experiencing experience. Nishida sees the origin of the true self in pure experience because in it shackles of the ego are shattered. The empirical, the metaphysical, and the existential are integrated here prior to their differentiation.

Human experience, which is usually encountered as constrained or shackled inside the subject/object framework, breaks through this framework into the unlimited openness through the originary event of pure experience as immediate seeing and hearing. Pure experience is then articulated within the subject-object framework, but now not as a constraining frame, but rather as a projective ladder into openness. The place for the self-articulation of experience is now in the subject-object framework within the infinite openness. This is then the equivalent to the place of being within the place of the absolute nothingness. From this perspective, pure experience articulates itself as the originary unified whole, sometimes from the subjective side, but not as a subject, and sometimes from the objective side, but not as an object. Illustrating the differentiation of pure experience in dynamic relationships is what we mean by explaining everything.

Nishida actually does present these explanations. Explanation, however, is work that takes place at the level of reflection. How are pure experience and explanation related? To answer this question, the standpoint of

pure experience turns into the standpoint of self-consciousness/self-awareness, which unites intuition and reflection within itself. Here, once again, the question of the place of self-consciousness is fundamental. This kind of awareness is more than self-consciousness because the limited place in the unlimited place of openness, this duality of place, is mirrored in the limited place that arises as the focal point of self-consciousness that is transparent to itself for the unlimited openness. Self-consciousness/self-awareness says, "I am I, in not being I" instead of simply "I am I."

The dynamic connection of "pure experience, self-consciousness or self-awareness, and place" (Basho) serves as the basis for further philosophical deliberations that Nishida carried out in the areas of art, history, society, the state, practical philosophy, the study of experience, mathematics, physics, and others areas in which he showed over and again how they are all permeated by this fundamental constellation. In the course of thinking that does not always proceed smoothly, Nishida tried out some unique categories such as active intuition, historical body, absolutely contradictory self-identity, and converse parallel to name just a few.

Nishida's thinking proceeds from a new beginning. Its basic category is place instead of substance, God, or the modern notion of an absolute (transcendental) subject. Logic as the logic of a contradictory self-identity, or rather the self-identity of the self-contradictory (the logic of place) instead of a logic of identity; the unity of the contradictory subject-object on the basis of something-before-the-split instead of a subject-object schema; reason as something that is active in intuition or rather acts as intuition instead of one side of a regional, qualitative distinction between sense and reason—all of these things arise out of pure experience. If global philosophy is going to take into account non-Western cultural traditions, Nishida's philosophy needs to be discussed within the horizon of that philosophy.

See also Buddhism; Phenomenology.

Bibliography

PRIMARY WORKS

Fundamental Problems of Philosophy: The World of Action and the Dialectical World. Translated by D. A. Dilworth. Tokyo: Sophia University Press, 1970.

Art and Morality. Translated by D.H. Dilworth and V. H. Viglielmo. Honolulu: University of Hawaii Press, 1973.

Intelligibility and the Philosophy of Nothingness. Translated by R. Schinzinger. Westport: Greenwood Press, 1973.

"The Logic of Topos and the Religious World View." Translated by M. Yusa. *The Eastern Buddhist* 19 (1986) and 20 (1987).

Intuition and Reflection in Self-Consciousness. Translated by V.H. Viglielmo with Y. Tekeuchi and J. S. O'Leary. Albany: SUNY Press, 1987.

Xenshu (Collected Works). 19 vols. 4th ed. Tokyo: Iwanami, 1987–1989.

An Inquiry into the Good. Translated by M. Abe and C. Ives. New Haven, CT: Yale University Press, 1990.

SECONDARY WORKS

Heisig, James W. *Philosophy of Nothingness: An Essay on the Kyoto School*. Honolulu: University of Hawaii Press, 2001.

Nishitani, Keiji. *Nishida Kitaro*. Translated by S. Yananoto and J.W. Heisig. Berkeley: University of California Press, 1991.

Wargo, Robert J. J. *The Logic of Nothingness: A Study of Nishida Kitaro*. Honolulu: University of Hawaii Press, 2005.

Yusa, Michiko. *Zen and Philosophy: An Intellectual Biography of Nishida Kitaro*. Honolulu: University of Hawaii Press, 2002.

Ueda Shizuteru (2005)

NJEGOŠ

See *Petrović-Njegoš, Petar*

NOMINALISM

See *Goodman, Nelson; Hobbes, Thomas; Quine, Willard Van Orman; Roscelin; Universals, A Historical Survey*

NOMINALISM, MODERN

In its main contemporary sense, nominalism is the thesis that abstract entities do not exist. Equivalently, it is the thesis that everything that does exist is a concrete object. Since there is no generally accepted account of the abstract-concrete distinction, and since it remains genuinely unclear how certain (putative) entities are to be classified, the content of modern nominalism is to some degree unsettled. Certain consequences of the view are, however, tolerably clear. For example, it is widely agreed that the objects of pure mathematics—numbers, sets, functions, abstract geometrical spaces, and so on—are to be classified as abstract. It is also widely agreed that certain objects of metaphysics and semantics—propositions, meanings, properties and relations, and so on—must be abstract if they exist at all. Modern nominalists thus commit themselves to rejecting these paradigmatic abstract entities and hence to rejecting any scientific, mathematical, or philosophical theory according to which such

things exist. In this sense nominalism is standardly opposed to platonism (or, less commonly, antinominalism).

EARLY HISTORY

The first significant philosophical system in the modern period to insist on the existence of abstract objects is due to Gottlob Frege. Frege (1980) held that the truths of pure mathematics concern a domain of mind-independent abstract entities. Frege (1984) further held that any adequate account of thought and language must allow that meaningful linguistic expressions are associated, not simply with concrete worldly items, but also with senses (*Sinne*), and that for various reasons these linguistic senses must exist in a "third realm," distinct both from the realm of subjective mental items and the realm of sensible, concrete things. Frege's vigorous defense of platonism in semantics and the philosophy of mathematics forms the background for the emergence of modern nominalism in the 1920s.

The Warsaw school of logicians centered around Stanisław Leśniewski and Tadeusz Kotarbiński set itself the task of reconstructing modern logic and mathematics along nominalistic lines. Kotarbiński's reism, for example, was a methodological position according to which, wherever possible, statements that apparently concern abstract entities (e.g., "Bonds of brotherhood unite Orsetes and Electra") are to be replaced by statements that concern only concrete entities and parts thereof (e.g., "Orestes is Electra's brother") (Kotarbiński 1955). The principal motivation for the program was to prevent scientific work in these areas from becoming embroiled in ancient metaphysical and epistemological controversies. The nominalistic project was introduced into Anglophone philosophy by W. V. Quine, who first encountered it in conversations with Leśniewski and Alfred Tarski in 1933. Quine's main positive contribution to the program was the seminal 1947 manifesto "Steps Toward a Constructive Nominalism," coauthored with Nelson Goodman. Quine soon abandoned nominalism in favor of a moderate and distinctive form of platonism. It may nonetheless be said that all subsequent discussion of modern nominalism in the Anglophone tradition derives directly from this paper.

MOTIVATIONS FOR NOMINALISM

In "Steps Toward a Constructive Nominalism" Goodman and Quine defend their rejection of abstract entities by invoking "a philosophical intuition that cannot be justified by appeal to anything more ultimate" (1947: 97). In subsequent years philosophers have sought to provide a more explicit motivation for the view.

OCCAM'S RAZOR. According to a slogan associated with the tradition of medieval nominalism, "entities are not to be multiplied beyond necessity." Some modern nominalists appeal to this principle in motivating their position. These writers typically concede that existing scientific and mathematical theories entail the existence of abstract entities and are therefore nominalistically unacceptable. They maintain, however, that it is possible to produce nominalistically adequate versions of, or surrogates for, these theories, and so to "dispense" with abstract objects. Occam's razor is then invoked to argue that when such parsimonious surrogates are available, it is rational to reject the standard platonistic theories and to embrace the surrogates instead.

Much of the constructive work in the nominalist tradition consists in providing nominalistic surrogates for existing theories. Roughly speaking, a nominalistic surrogate T_N for a platonistic theory T_P is a theory whose quantifiers range only over concrete objects, but which is nonetheless fit to do much of the same theoretical or explanatory work as the original. For example, standard formalizations of physical theories involve quantifiers that range over both concrete physical entities (particles, fields, points, and regions of space-time, etc.) and mathematical entities (real numbers, vectors, functions, etc.) A nominalistic alternative to (say) classical electrodynamics would be a theory whose quantifiers range only over concrete objects, but whose predictive and explanatory power exactly matched that of the standard platonistic formulations.

Nominalistic surrogates for standard theories have been developed in a number of domains (Field 1980, Hodes 1984, Chihara 1990, Balaguer 1998). However, the significance of these reconstructive programs is open to doubt for several reasons. For example, while the nominalistic surrogates do indeed typically posit fewer entities than the platonistic originals, they are typically inferior to the originals in other respects. In some cases they require a substantial extension of the extensional first-order logic that suffices for platonistically formulated science. In most cases the nominalistic theory is significantly less perspicuous and flexible than its platonistic counterpart.

One may therefore concede that other things being equal, nominalistic theories are to be preferred on grounds of parsimony, while insisting that since other things are not equal, Occam's razor has no clear application. A more profound challenge is directed at the razor

itself. Contemporary philosophers who cite ontological parsimony as a basis for theory choice often suppose that the principle derives its authority from its role in the sciences. But as critics have pointed out (Burgess and Rosen 1997), there is scant evidence that scientists accept the principle in its most general form. Scientists may be concerned to minimize the number of physical mechanisms or fundamental laws in the theories they accept. But working scientists and mathematicians have shown no interest in reducing the number of abstract entities posited by the mathematical theories they invoke. To the contrary, in mathematics and mathematical physics there is some concern to maximize the range of mathematical objects and structures (Maddy 1997). If this is correct, then proponents of the Occamist case for nominalism must maintain that the impulse to ontological parsimony to which they appeal is not a principle of scientific methodology, but an independently compelling philosophical principle.

THE ACCESS PROBLEM. The most widely cited ground for nominalism derives from Paul Benacerraf (1973). Benacerraf notes that since abstract mathematical objects are causally inert and therefore incapable of affecting our senses, even indirectly, there is a question as to how one might come to know that they exist. Benacerraf invokes the *causal theory of knowledge*, originally proposed by Alvin Goldman (1967) for other purposes, according to which, roughly, a person S knows that p only if S stands in some suitable causal relation to the objects with which p is concerned. This principle entails that true claims about abstract objects cannot be known to be true, even if they are true. And while this does not entail that there are no abstract entities, it does entail that platonism is unstable in the following sense: Proponents of a Platonistic theory must concede that they cannot know whether the theory they accept is true. However, as critics were quick to point out the causal theory of knowledge on which Benacerraf relies is objectionable on other grounds (Steiner 1975). In the subsequent debate nominalists rarely invoke this or any other detailed theory of knowledge. Instead, they maintain that the causal inefficacy of the abstract leaves our access to the abstract domain an utter mystery. Since it is clearly desirable to avoid such mysteries, this provides a motivation for pursuing, and perhaps also for accepting, nominalistic alternatives to standard theories.

THE DISPENSABILITY ARGUMENT. Hartry H. Field (1980, 1989) provides a number of motivations for nominalism that do not depend on the causal theory of knowledge. Field begins with a question for the platonist: What reason might one have for believing the claims of standard mathematics? If one has reason to believe the axioms, then one might acquire reason to believe the theorems by constructing proofs. So the question becomes: What reason might one have for believing the axioms of standard mathematics? Since the axioms involve substantial existential claims, it is hard to see how they could be known *a priori* (but see Wright 1983, Hale 1988). And since these claims concern causally inert abstract entities, it seems clear that they cannot be verified directly by observation or experiment. Field thus concludes that the only reason one can have for believing the axioms is that they play an indispensable role in one or another well-confirmed scientific theory. Earlier writers (Quine 1960, Putnam 1971) defended platonism in this way. For example, Hilary Putnam (1971) notes that since the laws of physics are standardly formulated in mathematical terms, someone who denies the existence of (say) real numbers is not in a position to formulate, much less to employ, even the most elementary laws of physics. Quine and Putnam thus offer the following indispensability argument for platonism:

> (1) One is justified in believing that abstract objects exist if, but only if, theories that entail the existence of such objects are indispensable for scientific purposes.
>
> (2) Standard mathematics entails the existence of abstract objects.
>
> (3) Standard mathematics is indispensable for scientific purposes.
>
> (4) Therefore, one is justified in believing that abstract objects exist.

Field rejects premise (3), thereby turning the argument on its head. He argues that in certain cases it is possible to produce reasonably attractive nominalistic versions of standard platonistic theories: versions in which the only objects posited are material bodies and space-time regions. Field maintains that to the extent that such nominalistic surrogates are available, they establish that abstract objects are dispensable for scientific purposes. The construction of such surrogates thus undercuts the only reason one might have had for believing in abstract objects, and so provides a roundabout motivation for nominalism.

Field concedes that the nominalistic alternatives he constructs are in certain respects inferior to the standard platonistic theories on which they are based. They are

typically unwieldy and imperspicuous: Derivations are typically longer and harder to follow. Field concedes that it would be unreasonable for working scientists to use these nominalistic theories for most purposes and hence that platonistic theories are indispensable in practice. His central claim is that they are nonetheless dispensable in principle and that for the purposes of the Quine-Putnam challenge dispensability in principle is what matters.

One distinctive ingredient in Field's view is a demonstration that scientists who accept only the nominalistic physics that Field constructs are nonetheless entitled to use platonistic mathematics in the course of their work. This claim is supported by a formal result. Let T_P be a standard platonistic theory, and let T_N be a nominalistic surrogate for T_P constructed according to Field's method. It may then be shown (with certain important qualifications) that for any nominalistic statement S—that is, any statement whose quantifiers are restricted to concrete entities—S is a theorem of T_P if and only if S is a theorem of T_N. This conservative extension theorem supports the claim that a theorist who accepts T_N may legitimately employ the full mathematical resources of T_P for the purpose of deriving nominalistic claims about the concrete world (for a discussion on this, see Shapiro 1983, Burgess and Rosen 1997). Such theorists may then legitimately regard the mathematical apparatus of T_P as a useful fiction in which they indulge for various practical purposes. Field's version of nominalism is thus a form of fictionalism about mathematical objects.

Field's work has provoked an intense critical response (Irvine 1990). Field himself notes that his procedures for nominalizing platonistic theories are inapplicable to an important class of theories, including Albert Einstein's general theory of relativity and quantum mechanics, and hence that it remains an open question whether platonistic theories are dispensable even in principle for the purposes of contemporary physics (compare Balaguer 1996). Others wonder why Platonistic theories that are indispensable in practice should not provide one with adequate grounds for believing in the abstract objects they posit. Perhaps the most fundamental philosophical response to Field's approach calls into question premise (1) of the indispensability argument, which is also a crucial premise in Field's positive defense of nominalism. In effect, the premise asserts that abstract objects have the status of theoretical entities, in the sense that one acquires reason for believing in them only when the assumption of their existence is required for some urgent scientific purpose.

Against this, critics maintain that some propositions about abstract objects—for example, the claim that there is a number between 3 and 5, or the claim that Jane Austen wrote six novels—are perfectly ordinary claims. Anyone who has learned basic arithmetic can supply a reason for believing that there is a number between 3 and 5 (Parson 1986), and anyone who knows how to use the library can verify that Austen wrote six novels. It is a presupposition of the debate between Field and proponents of the indispensability argument that these relatively nontheoretical justifications for platonistic claims are inadequate. But this claim may be challenged. If one's ordinary reasons for believing platonistic claims are good enough, then the fact that such claims are dispensable for certain theoretical purposes has no immediate bearing on the debate over nominalism.

REVOLUTIONARY VERSUS HERMENEUTIC NOMINALISM

In the nominalist tradition that runs from Goodman and Quine (1947) to Field (1980), it is generally conceded that since standard mathematics entails the existence of abstract objects, the nominalist must supply an alternative to standard mathematics, both pure and applied. This alternative might take the form of a genuinely novel formulation, as with Field's nominalistic version of Newtonian gravitational theory. But it may also take the form of a reinterpretation of existing theories. On this approach the nominalist proceeds by supplying a revisionary account of the meanings of mathematical statements. For example, the nominalist may maintain that while existential arithmetical statements like "There is a number between 3 and 5" in fact affirm the existence of abstract entities, they should be reinterpreted as claims about (say) concrete numeral inscriptions. In either case the nominalist must argue for a revision in accepted science and mathematics. Nominalist programs of this sort have thus been labeled *revolutionary* (Burgess 1983).

Revolutionary nominalism is contrasted with hermeneutic nominalism. Hermeneutic nominalists maintain that it is a mistake to interpret ordinary mathematics as involving claims about abstract objects in the first place. They might maintain, for example, that as they are ordinarily understood, existential claims like "There is a number between 3 and 5" are in fact claims about concrete numeral inscriptions and hence that such claims might be true even if there were no abstract entities. On this sort of account nominalism requires no revision in settled doctrine.

The most straightforward version of hermeneutic nominalism would maintain that abstract singular terms like *3* and *the cosine function* denote particular concrete objects. Claims of this sort are rarely plausible, however, and so proposals in this domain are typically more complex. For example, Geoffrey Hellman (1989) proposes that a statement S in the language of arithmetic is true if and only if a certain modal condition holds: (a) there might have been an infinite sequence of objects satisfying the axioms of arithmetic, and (b), if there had been such a sequence, a certain structural condition derived from S would have been true of it. Hellman then argues that since this sort of modal claim might be true even if there are in fact no abstract objects, the original mathematical claim is nominalistically acceptable, appearances to the contrary notwithstanding.

There are two main objections to hermeneutic proposals of this sort. The first notes that since such claims are ultimately claims in empirical linguistics—they are claims about the meanings of ordinary mathematical statements—they require empirical support and that in the relevant cases no such support has been forthcoming (Burgess 1983). The second notes that even if hermeneutic nominalists' semantic claims were tenable, it is not clear that they would serve their purpose. Unlike their revolutionary counterparts, hermeneutic nominalists do not deny the claims of standard mathematics. But these claims include *existence theorems*: assertions of the form "There exists a number n such that …" Hermeneutic nominalists must therefore allow that these ordinary existence claims are true and hence that by their own lights, numbers and the like exist. On the face of it, however, this claim is incompatible with their nominalism (Alston 1958, Burgess and Rosen 1997; see also Stanley 2001).

CONTEMPORARY FICTIONALISM

As they are usually understood, the programs of revolutionary and hermeneutic nominalism both require detailed constructive work. Theorists proceed by constructing an autonomous, independently intelligible nominalistic theory T_N, which is then used either to replace or to interpret the original (apparently) platonistic theory, T_P. The development of a suitable theory T_N is typically a nontrivial task, which in many cases requires a profound analysis of the original.

However, some nominalists maintain that detailed constructions of this sort are unnecessary. Easy fictionalism, as the approach is sometimes called, holds that even in the absence of an autonomous nominalistic alternative, nominalists may make free use of standard mathematics and of other platonistic theories without thereby committing themselves to the existence of abstract objects.

Consider for example the claim (S): (S) the mass (in grams) of $A = 3.6$. On its face (S) asserts that the object *A* stands in a certain relation to a number. The claim is literally true only if two conditions are satisfied: on the concrete side, the object *A* must have a certain intrinsic property—a property for which one may have no standard name that does not invoke a relation to numbers; and on the abstract side, the number 3.6 must exist. To maintain the literal truth of (S) is thus to maintain that abstract objects exist. But consider the claim that things are, in all concrete respects, as if (S) were true. The suggestion is that this claim says just what (S) says about the intrinsic configuration of the concrete world, while making no claim whatsoever about the existence of abstract entities.

Easy fictionalists propose that as a matter of convenience one routinely pretends that abstract objects of various sorts exist and that one conveys information about the concrete world by endorsing theories that purport to affirm relations between concrete things and abstract things. Their suggestion is that in "endorsing" these theories, one commits oneself only to the nominalistically acceptable claim that things are, in all concrete respects, as if one's theories are true.

Easy fictionalism comes in a number of varieties. It may be put forward as a hermeneutic proposal, describing the attitude that scientists and mathematicians normally adopt toward their own claims about abstract entities (Yablo 2001). More commonly, it is put forward as a revolutionary proposal. Here, the suggestion is that in light of the arguments in favor of nominalism, it would be rational (or at least, rationally permissible) to adopt a fictionalist attitude toward discourse about abstract objects (Balaguer 1998, Rosen 2001). The main challenge for easy fictionalism is to provide a clear account of the central idiom, "Things are, in all concrete respects, as if S were true," or perhaps, "According to the fiction of mathematical objects, S." The most natural account involves a counterfactual conditional. To say that things are in all concrete respects as if S were true is to say that if there were abstract objects (and the concrete world were just as it is in all intrinsic respects), then S would be true. But counterfactuals of this sort are problematic. It is widely held that the existence of abstract objects could not possibly be a contingent matter (Hale and Wright 1992; compare Field 1993). And if this is right, then by nominalists' own lights, such conditionals involve a necessarily false

antecedent. A second challenge for the approach is to provide an account of pure mathematics, where the aim of the discourse is not simply to provide information about the configuration of the concrete world.

See also Realism and Naturalism, Mathematical.

Bibliography

Alston, William. "Ontological Commitments." *Philosophical Studies* 9 (1958): 8–17.

Balaguer, Mark. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press, 1998.

Balaguer, Mark. "Towards a Nominalization of Quantum Mechanics." *Mind* 105 (1996): 209–226.

Benacerraf, Paul. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–680.

Burgess, John P. "Epistemology and Nominalism." In *Physicalism in Mathematics*, edited by Andrew D. Irvine, 1–15. Dordrecht, Netherlands: Kluwer Academic, 1990.

Burgess, John P. "Why I Am Not a Nominalist." *Notre Dame Journal of Formal Logic* 24 (1983): 93–105

Burgess, John P., and Gideon Rosen. *A Subject with No Object: Strategies for Nominalistic Interpretation of Mathematics*. New York: Oxford University Press, 1997.

Chihara, Charles S. *Constructibility and Mathematical Existence*. New York: Oxford University Press, 1990.

Field, Hartry H. "The Conceptual Contingency of Mathematical Objects." *Mind* 102 (1993): 285–299.

Field, Hartry H. *Realism, Mathematics, and Modality*. New York: Blackwell, 1989.

Field, Hartry H. *Science without Numbers: A Defence of Nominalism*. Princeton, NJ: Princeton University Press, 1980.

Frege, Gottlob. *The Foundations of Arithmetic: A Logico-mathematical Enquiry into the Concept of Number*. 2nd ed. Translated by J. L. Austin. Evanston, IL: Northwestern University Press, 1980. This was originally published in German under the title *Der Grundlagen der Arithmetik* in 1884.

Frege, Gottlob. "Thoughts." In *Collected Papers on Mathematics, Logic, and Philosophy*, translated by Max Black and edited by Brian McGuiness. New York: Blackwell, 1984. This essay was originally published in German under the title "Der Gedanke" in 1918.

Goldman, Alvin. "A Causal Theory of Knowing." *Journal of Philosophy* 64 (1967): 357–372.

Goodman, Nelson, and W. V. Quine. "Steps toward a Constructive Nominalism." *Journal of Symbolic Logic* 12 (1947): 105–122.

Hale, Bob. *Abstract Objects*. New York: Blackwell, 1988.

Hale, Bob, and Crispin Wright. "Nominalism and the Contingency of Abstract Objects." *Journal of Philosophy* 89 (1992): 111–135.

Hellman, Geoffrey. *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. New York: Oxford University Press, 1989.

Hodes, Harold. "Logicism and the Ontological Commitments of Arithmetic." *Journal of Philosophy* 81 (3) (March 1984): 123–149.

Irvine, Andrew D., ed. *Physicalism in Mathematics*. Dordrecht, Netherlands: Kluwer Academic, 1990.

Kotarbiński, Tadeusz. "The Fundamental Ideas of Pansomatism." *Mind* 64 (1955): 488–500. Originally published in Polish in 1935.

Maddy, Penelope. *Naturalism in Mathematics*. New York: Oxford University Press, 1997.

Parson, Charles. "Quine on the Philosophy of Mathematics." In *The Philosophy of W. V. Quine*, edited by Lewis Edwin Hahn and Paul Arthur Schilpp. La Salle, IL: Open Court Press, 1986.

Putnam, Hilary. *Philosophy of Logic*. New York: Harper and Row, 1971.

Quine, W. V. *Word and Object*. Cambridge, MA: Cambridge Technology Press of the Massachusetts Institute of Technology, 1960.

Rosen, Gideon. "Nominalism, Naturalism, Epistemic Relativism." *Philosophical Perspectives* 15 (2001): 69–91.

Shapiro, Stewart. "Conservativeness and Incompleteness." *Journal of Philosophy* 80 (1983): 521–531.

Stanley, Jason. "Hermeneutic Fictionalism." In *Midwest Studies in Philosophy. No. 25: Figurative Language*, edited by H. Wettstein and P. French, 36–71. New York: Blackwell, 2001.

Steiner, Mark. *Mathematical Knowledge*. Ithaca, NY: Cornell University Press, 1975.

Wright, Crispin. *Frege's Conception of Numbers as Objects*. Aberdeen, Scotland: Aberdeen University Press, 1983.

Yablo, Stephen. "Go Figure: A Path through Fictionalism." In *Midwest Studies in Philosophy. No. 25: Figurative Language*, edited by H. Wettstein and P. French, 77–102. New York: Blackwell, 2001.

Gideon Rosen (2005)

NOMOS AND PHUSIS

Phusis is the ancient Greek word for "nature," cognate with the verb "to grow" (*phuein*); as in English, it can be used both for the natural world as a whole and for the "nature" (i.e., the essential or intrinsic characteristics) of any particular thing, which it has "by nature" (*phusei*). *Nomos* encompasses both law and unwritten, traditional social convention. The contrast between the two concepts is central to ancient sophistic thought, with roots in the pre-Socratic inquiry into the underlying natures of things.

For the Sophists, nomos and phusis are polar terms, roughly equivalent (respectively) to the socially constructed and the universally, objectively given. The contrast was most strikingly applied in relation to justice. Antiphon's *On Truth* argues that justice is a matter of nomos, and nomos and phusis conflict; one should observe the requirements of justice when there are witnesses, but follow the dictates of nature otherwise. By "nature," Antiphon seems to understand what is physio-

logically given to all humans (Greeks and barbarians alike). By following it one gains what is advantageous to one's existence: life, pleasure, and freedom. In Plato's *Gorgias*, Callicles argues, with an appeal to animal behavior, that it is a matter of "justice according to nature," as opposed to convention, for the strong to prey upon the weak.

However, the same conceptual framework, including the assumption that nature represents an authoritative norm, could be used to support the opposite stance. The *Anonymous Iamblichi* argues that law and justice should be obeyed as having "kingly rule" among human beings—a rule established by human nature itself. So the nomos-phusis contrast was a framework for discussion rather than a theory in itself. It allowed for fruitful debate as to where the testimony of nature might be observed, what guidance it could provide, and how the norms of law and morality might relate to it.

Far from being restricted to justice, nomos-phusis is best understood as a catch phrase for the general sophistic inquiry into the institutions of human society. Thus various Sophists seem to have applied the concepts to slavery, gender roles, language, and religion. For instance, the *Sisyphus* fragment (by either Critias or Euripides) argues that religion was invented by ancient sages as a device for social control, implying that the gods exist only by convention. The contrast could even be extended to questions of general epistemology. Democritus (usually classed as a pre-Socratic, but associated by sources with Protagoras) summed up his atomism by claiming that sensory properties, such as colors and tastes, are merely conventional; in reality there are only atoms and the void. Here, conventional seems to be tantamount to mind-dependent, or merely apparent.

The adoption of nature as a normative standard is the most powerful legacy of sophistic thought. Plato and Aristotle both constructed their ethics and politics around their understanding of human nature, and took this to be in harmony with the nature of the cosmos and the divine. Later, Epicureans and Stoics both argued that the good life is one lived in accordance with nature (*kata phusin*), which they explicated by invoking animal behavior in the "cradle argument." But these philosophers differed widely in their treatment of nomos, and the nomos-phusis polarity as such faded from prominence after the Sophists.

See also Antiphon; Protagoras of Abdera; Sophists.

Bibliography

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 3. Cambridge, U.K.: Cambridge University Press, 1969.

Heinimann, F. *Nomos und Physis*. Basel: F. Reinhardt, 1945.

Kahn, Charles. "The Origins of Social Contract Theory in the Fifth Century B.C." In *The Sophists and their Legacy*, edited by G. B. Kerferd. Wiesbaden: Steiner, 1981.

Kerferd, G. B. *The Sophistic Movement*. Cambridge, U.K.: Cambridge University Press, 1981.

Rachel Barney (2005)

NONCOGNITIVISM

Noncognitivists (or nondescriptivists) hold that the function of normative judgments is not, or not primarily, to describe or state facts and that because of this, these judgments lack a truth-value. A strong form of ethical nondescriptivism says that moral judgments have no descriptive function, but weaker forms say only that their nondescriptive function is primary or dominant.

Differing accounts of the nondescriptive function of moral language generate a variety of nondescriptivisms. Moral judgments have been said to express emotions, feelings, attitudes, or stances; and they have been characterized as tools for performing other nondescriptive tasks such as commanding, requesting, endorsing, or commending. A. J. Ayer, whose position is called emotivism, said that "ethical terms" express emotions or feelings and that they "are calculated also to arouse feelings, and so to stimulate action" (1952, p. 108). C. L. Stevenson, whose metaethical theory is called noncognitivism, argued that the major use of "ethical statements" is dynamic rather than fact stating. They are not, he said, primarily used to describe interests or attitudes but rather to change or intensify attitudes and to influence behavior. What Stevenson called the emotive meaning of ethical terms makes this dynamic use possible and also explains why ethical judgments, unlike factual ones, are capable of moving us to action.

From the thought that moral judgments are exclamations and disguised commands Ayer concluded that they "have no objective validity whatever" and that "it is impossible to dispute about questions of value" (1952, p. 110). Stevenson tried to show that there is a place for ethical arguments, but he did not go beyond the claim that a reason is "relevant" when it is likely to influence some attitude. This means, at least to the critics of Stevenson, that the relation between the premises and the conclusion of an ethical argument is psychological rather than logi-

cal and that there is no clear distinction between ethical argument and propaganda.

Both Ayer and Stevenson were in the positivist tradition, but by the 1950s an interest in ordinary language also led increasing numbers of analytic philosophers to nondescriptivism. These thinkers acknowledged that moral language can be used descriptively, but they insisted that its "primary" (basic, fundamental) use is to perform any of a number of nondescriptive speech acts. R. M. Hare argued that the primary function of the word *good* is to commend and that when we commend anything "it is always in order, at least indirectly, to guide choices, our own or other people's, now or in the future" (1952, p. 127). Words such as *right* and *ought* are used for giving advice or, as he said, for prescribing. According to Hare, the claim that something is good has both descriptive and prescriptive meaning. The descriptive meaning of the word *good* changes as it is applied to different things, but the prescriptive meaning remains constant because *good* is invariably used to commend. This is why the prescriptive meaning is primary.

Hare described his own position as nondescriptivism, but he was more positive than Ayer and Stevenson about the role and value of logic in ethical arguments. Moral judgments, he said, are a subclass of "prescriptive" rather than "descriptive" language—they are "universalizable prescriptions." Unlike attempts to persuade or to influence attitudes, a judgment that something is good or right is a prescription that is complete in itself, even if no change is brought about in the hearer's attitudes or behavior. Hare believed that there could be logical relations among prescriptive judgments, even commands; and he developed a logic of prescriptive discourse to account for those relations. In the end he concluded that while we can argue logically about what to do, a complete justification of a moral decision will always require the adoption, without justification, of some basic principle or principles as a part of a freely chosen "way of life."

P. H. Nowell-Smith offered a form of nondescriptivism he called multifunctionalism. He said that evaluative language is used "to express tastes and preferences, to express decisions and choices, to criticize, grade, and evaluate, to advise, admonish, warn, persuade and dissuade, to praise, encourage, and reprove, to promulgate and draw attention to rules; and doubtless for other purposes also" (1954, p. 98). Though his position is more complex than Hare's, he does agree that "the central activities for which moral language is used are choosing and advising others to choose" (p. 11).

After the contributions of Ayer, Stevenson, Hare, Nowell-Smith, and others, nondescriptivism was neglected as interest in applied ethics flourished and as those who did think about metaethics developed naturalistic forms of descriptivism. The new naturalists conceded that normative language has nondescriptive functions, but they then pointed out how those functions are compatible with simultaneous descriptive intent and therefore with the possibility of evaluating normative pronouncements in terms of truth and falsity. In the 1980s interest in metaethics was stimulated by new forms of nondescriptivism developed by Simon Blackburn and Allan Gibbard. The dominant issue at that time, however, was the dispute between moral (or ethical) realists and antirealists. Nondescriptivists are more likely to be antirealists, and descriptivists are more likely to be realists, but there are complications.

Formerly, both intuitionists and naturalists were descriptivists. Intuitionists identified moral facts with nonnatural facts, and naturalists identified moral facts with natural facts. If one who believes that moral facts are natural facts can be said to be a moral realist, then both naturalists and intuitionists were moral realists and were in a position to say that moral judgments are true when they correctly describe some natural or nonnatural reality. But there is a way to combine descriptivism with antirealism and another way to combine nondescriptivism with at least the practices of the realist. J. L. Mackie develops a descriptivist account of much normative language, but he argues that judgments of moral obligation, which are thought to be both objective and prescriptive, and judgments of "intrinsic" value are always false. One who says that something is "good in itself" is always speaking falsely because nothing is good in itself.

Both Blackburn and Mackie begin with a Humean projectivism according to which the normativity we think we discover in nature is projected onto a value-free world by us. When we see and are moved by cruelty to the bull, we objectify our negative attitude, and promote it too, by saying that bullfighting is wrong. Projectivists are antirealists. Mackie combines his antirealism with descriptivism and takes this to result in an error theory. Blackburn begins with antirealism, adds his version of nondescriptivism or "expressivism," and emerges with what he calls quasi realism, the idea that the linguistic practices of the realist—saying that bullfighting is really wrong, for example—are perfectly in order and that no error is made. One of his main concerns is to defend this quasi realism by showing how we "earn the right" to "practice, think, worry, assert, and argue" as though

moral commitments are true in some straightforward way (1984, p. 257).

Blackburn's view is that we do not describe reality correctly or incorrectly when we make moral claims—we express "stances." He characterizes a stance as a "conative state or pressure on choice and action" but admits that we could also call this an attitude. But whatever we call it, "its function is to mediate the move from features of a situation to a reaction, which in the appropriate circumstances will mean choice" (1993, p. 168).

Gibbard also defends a nondescriptivist or "expressivist" account of normative judgments. Normative judgments, he says, take the form of saying that some act, belief, or feeling is "rational," or "makes sense." The point of making such a judgment is not to describe something, not to attribute a property to it, but "to express one's acceptance of norms that permit it" (1990, p. 7). A norm, according to Gibbard, is "a linguistically encoded precept," and the capacity to be motivated by norms "evolved because of the advantages of coordination and planning through language" (p. 57). There are norms of many kinds, but when we say that what someone did was morally wrong, we are expressing and endorsing norms that govern feelings of guilt by the agent and of anger by others.

Three arguments are traditionally deployed against nondescriptivists. According to the grammatical argument, since moral judgments are phrased in ordinary indicative sentences, there is a prima facie reason to treat them as statements and to treat those who make them as attempting to make statements. Nondescriptivists will reply that here the grammar is misleading, but they can then be asked to explain why this should be so. There is also a logical argument against nondescriptivism. If moral judgments lack a truth-value, then it is impossible for them to play a role in truth-functional constructions (implication, conjunction, and negation, for example) and in arguments. It is also difficult to know how they are to be interpreted when they occur embedded in complex constructions such as statements of belief and doubt. According to what has been called the phenomenological argument, not only do moral claims look and behave like descriptive utterances, they "feel" like them too. When we claim that something is good or right, we do not seem, even to ourselves, to be merely expressing ourselves or ordering others to do things. Nondescriptivists will try to explain why these judgments have this distinctive feel, but descriptivists will insist that the feeling is important data that cannot easily be explained away.

Starting with Ayer, each nondescriptivist has been forced to develop some reply to these, as well as to other, difficulties. Blackburn, for example, responds to the logical argument by developing an expressivist account of truth. He wants to show how it makes sense to claim moral truth even if there are no moral facts and even if our moral claims are no more than expressions of stances or attitudes. Gibbard sketches a solution to the embedding problem that exploits the idea that when we make a normative statement we are expressing a state of mind that consists in "ruling out various combinations of normative systems with factual possibilities." He develops a formalism that allows him to use this idea to account for "the logical relations that hold among normative statements" (1990, p. 99).

Owing to the work of Blackburn and Gibbard, nondescriptivism is alive and well, but its prospects are uncertain because it is truly difficult to develop convincing and definitive answers to the objections from grammar, logic, and phenomenology. Furthermore, nondescriptivism needs a fact/value distinction, and this is something about which philosophers have become increasingly nervous. The early descriptivists tried to reduce values to facts, or they accepted the fact/value distinction and then relegated values to a philosophically insignificant pragmatic limbo. Since then there has been a tendency to argue that many statements that appear to be safely descriptive must be understood to have nondescriptive elements. Nondescriptivists now point out that even if the line between facts and values is blurred or moved, we can still draw an important distinction between assertions and expressions. This claim, however, will continue to be challenged by those who are impressed by the descriptive nature of norms or the normative nature of descriptions.

See also Applied Ethics; Ayer, Alfred Jules; Hare, Richard M.; Mackie, John Leslie; Metaethics; Moral Realism; Projectivism; Stevenson, Charles L.

Bibliography

Ayer, A. J. *Language, Truth, and Logic.* New York: Dover, 1952.

Blackburn, S. "How to Be an Ethical Anti-realist." In *Essays in Quasi-Realism.* New York: Oxford University Press, 1993.

Blackburn, S. *Spreading the Word: Groundings in the Philosophy of Language.* Oxford: Clarendon Press, 1984.

Blackburn, S. "Wise Feelings, Apt Reading." *Ethics* 102 (1992): 342–356.

Darwall, S., A. Gibbard, and P. Railton. "Toward *Fin de Siècle* Ethics: Some Trends." *Philosophical Review* 101 (1992): 115–189.

Gibbard, A. *Wise Choices, Apt Feelings: A Theory of Normative Judgment.* Cambridge, MA: Harvard University Press, 1990.

Hare, R. M. *Freedom and Reason.* Oxford: Clarendon Press, 1963.

Hare, R. M. *The Language of Morals.* Oxford: Clarendon Press, 1952.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* Harmondsworth, U.K.: Penguin, 1977.

Nowell-Smith, P. *Ethics.* London: Penguin, 1954.

Stevenson, C. L. "The Emotive Meaning of Ethical Terms." In *Facts and Values.* New Haven, CT: Yale University Press, 1963.

Stevenson, C. L. *Ethics and Language.* New Haven, CT: Yale University Press, 1944.

Urmson, J. O. *The Emotive Theory of Ethics.* London: Hutchinson, 1968.

Richard Garner (1996)

NONDESCRIPTIVISM

See *Noncognitivism*

NONEXISTENT OBJECT, NONBEING

We think and talk about things that do not exist—or so it seems. We say that Santa Claus lives on the North Pole and that unicorns are white. We admire Sherlock Holmes or judge him to be more clever than J. Edgar Hoover. People search for the Northwest Passage and the Fountain of Youth. They dream about lottery winnings and fear disasters that do not materialize. A childless couple hopes for a daughter. So, according to Alexius Meinong and others, there are things that do not exist. Even to deny that Santa Claus or the Fountain of Youth exists, we must be able, it seems, to identify what it is whose existence we are denying.

Bertrand Russell's rejection of this line of thought is well known. Sentences containing expressions that appear to denote nonexistents are to be paraphrased, in accordance with his theory of descriptions, with ones that do not. Russell shifted the emphasis from thoughts and other intentional attitudes that appear to have nonexistents as objects to the language in which those thoughts and attitudes are expressed. Many later analytic philosophers shared with Russell a distaste for what they saw as Meinong's bloated universe. Even those who rejected his theory of descriptions often assumed that apparent references to nonexistents can somehow be paraphrased away. But there have been few serious attempts since Russell's to show how this can be done, and the task has proven to be much more difficult than it once seemed. Since 1970 several very different sophisticated realist theories were developed, some of which claim not that there are nonexistent objects, but that the entities in question exist (Woods, Van Inwagen, Parsons, Routley, Wolterstorff, Thomasson). These have been countered by a new generation of antirealist theories, many of them based on notions of pretense or make-believe (Evans, Walton, Yablo, Kroon; also see Currie).

Many discussions after 1970 have focused especially or primarily on one variety of purported nonexistents: characters and other objects in fiction and mythology. Posits of failed scientific theories (Vulcan, ether, phlogiston), sought after marvels and wished-for children, the golden mountain, the round square, and the present King of France are often treated along the way, although the issues they involve are not entirely analogous to those concerning fictions. Fictions are in some ways especially compelling, and also especially puzzling. We speak easily and elaborately about fictional characters as though they were ordinary people, describing Sherlock Holmes as a detective who lives on Baker Street, speculating about Hamlet's motivations, and recounting the amorous adventures of the various Don Juans. Yet when pressed in certain ways, we readily deny that there are such things as fictions. Parents assure their frightened children that there really are not any goblins or monsters or ghosts like the ones in storybooks, and they confess to having lied about Santa Claus.

LITERALISM

What sorts of things are nonexistents, if there are such? Some take descriptions of Sherlock Holmes as a man and a detective at face value, understanding characters to be people, to possess the same kinds of ordinary properties that real people do, and to differ only in lacking existence. The golden mountain is, literally, golden and a mountain, according to Meinong, and the wished-for child is a child. Such *literalists*, as Kit Fine (1982) calls them, usually accept that, unlike existing objects, most nonexistents are *incomplete* (Holmes neither has a mole on his back nor lacks one) and some are impossible (fictional time travelers, the round square).

Literalism threatens to get out of hand, at least as far as fictions are concerned. Not only do we readily describe Holmes as a person and a detective, we are also prepared to say, in much the same spirit (that is, speaking within the story), that he and other characters exist. Macbeth's dagger may be a mere figment of his imagination, but

Macbeth himself is real; he exists. (We do not, in a comparable spirit, describe the childless couple's wished-for child or the golden mountain as existing.) In general, we are prepared to assert what we take to be true in a story, or *fictional*. It is fictional both that Holmes is a person and that he exists. If Holmes is literally a person, it is awkward to deny that he literally exists. But fictional statements that do not involve fictitious particulars—statements such as "There are ghosts" and "Julius Caesar was warned of the ides of March" understood literally and straightforwardly, obviously may fail to be true. Why should "Holmes is a person," not to mention "Holmes exists," be different?

The most obvious alternative to literalism, in the case of fictions anyway, is to treat statements like "Holmes is a person and a detective" as elliptical, as short for "It is fictional (true in the story) that Holmes is a person and a detective." Of course, we still seem to have an entity on our hands—Holmes. He (or it) is not literally a person or a detective, but he is such that it is fictional that he possesses these attributes. And it is fictional that he exists, which does not have to mean that he literally exists. Holmes may possess other kinds of properties as well, ones that do not consist in something being fictional of him: He is a fictional character, and (on some accounts) was created by Conan Doyle, and is admired by millions of readers. Abandoning literalism in this way removes the embarrassment of an incomplete Holmes, and we need not worry about running into inconsistent fictional objects. Holmes is not such that fictionally he has a mole on his back, nor is he such that fictionally he does not have a mole on his back, even though, it is fictional that he either does or does not have one there. A character may, be such that it is fictional both that she is both *P* and that she is not *P*, but that does not mean that the character *really* does possess incompatible properties.

In an alternative to this strategy, developed by Edward Zalta (1988), Holmes is not a person in the sense that J. Edgar Hoover is; he does not *exemplify* personhood. Unlike Hoover, Holmes exemplifies properties such that of being a fictional character. Holmes bears a different relation, which Zalta calls *encoding*, to personhood, to being a detective, and to the other properties attributed to him in the Sherlock Holmes stories. J. Edgar Hoover, by contrast, is not the kind of thing that encodes properties.

ABSTRACT OBJECT THEORIES

If Holmes is not a person, what is he? What sort of thing is fictionally a person (or encodes personhood)? Realists who are not literalists usually understand fictions to be abstract entities of one sort or another, and to have whatever ontological standing the abstract entities in question do. Some take properties like being a person and a detective to *constitute* (rather than characterize) fictions, and so identify Holmes with the class of properties or conjunction of properties attributed to him in the stories. Some construe fictions as abstractions of other sorts: "theoretical entities of literary discourse" (Van Inwagen 1977), "kinds" (Wolterstorff), or "abstract artifacts" (Thomasson 1999). Zalta (1988) has fictions exemplifying abstractness. Different abstract-object theories give different answers to a battery of tricky questions about the identity and individuation of fictions and other nonexistents or nonactuals. Are they Platonic entities that *are* (some even say "exist") necessarily and eternally (Parsons 1980), or are they created when, for example, the relevant story is written or when they are thought about (Van Inwagen 1977, Thomasson 1999)? Do they cease to be if the story is destroyed and forgotten? If characters in different unrelated stories happen to have exactly the same characteristics attributed to them, are they identical? Are undifferentiated characters in a single fiction distinct from one another (the individual sheep in a fictional flock, for instance if nothing is said about any of them apart from the others)? Can the same character appear in more than one story if the characteristics attributed to it in each of them are not exactly the same? If so, by virtue of what are the characters identical?

The apparent fact that readers admire Holmes, or care about characters in stories, poses an awkward challenge for abstract-object theories. Do readers admire and care about abstract entities, be they properties or classes or theoretical entities or abstract artifacts? This is certainly not how readers themselves think of their experiences. It hardly helps to claim that Holmes is a person in the sense that he "encodes" personhood. He belongs to an ontological category fundamentally different from that of the usual objects of admiration—Mahatma Ghandi, Abraham Lincoln—which *exemplify* personhood and do not encode any properties at all.

The antiliteralist might deny that people do, literally, admire or care about Holmes or Willy Loman or Desdemona, just as he denies that it is literally true that they are persons (or that they exemplify personhood). But then what *is* the reader's relation to them? Does the reader *imagine* admiring or caring about these abstractions, or is it true in an extended fiction that he admires them? Does the reader imagine of an abstract object that it is a person, one that he admires and cares about? That would be quite an imaginative feat!

Similar worries arise simply with the notion of fictionality and infect purported nonexistents of other kinds, as well as fictions. Fictional propositions are commonly characterized as propositions that appreciators or readers are to imagine, or ones that works of fiction invite them to imagine. If it is fictional that Holmes is a person, readers are to imagine of this abstract object that it is a person (or to imagine of something that encodes personhood that it exemplifies personhood instead). To imagine this would be to imagine a blatant impossibility. If a wished-for daughter is not actually a daughter or a person but an abstract entity, does the childless couple wish, futilely, of this abstraction, that it is a daughter and a person?

PRETENSE THEORIES

Pretense or make-believe theories return to a more intuitive understanding of statements like "Holmes is a detective," without embracing literalism. The speaker pretends to refer to an ordinary existing person and to attribute to him, in the ordinary way, the ordinary property of being a detective. *Within the scope of the pretense*, everything is normal. Yet nothing is actually referred to, and what is said, understood literally, is not true. This is pretense, yet with a serious purpose. The speaker does actually assert something by engaging in the pretense, very likely something about the Sherlock Holmes stories. Pretense theorists need to give some account of what is asserted, though it may be asking too much to expect an exact literal paraphrase. Part of the point of speaking in pretense to make a serious assertion may be to express something that is difficult or impossible to express literally (Yablo 1998).

Some philosophers find pretense accounts of "Holmes is a detective," "Hamlet hesitated," and the like plausible, but draw a sharp line between these statements and statements such as "Holmes is a fictional character" and "Holmes is smarter than any real detective." In the Sherlock Holmes stories it is presumably fictional that Holmes is a detective; readers are to imagine that this is so. In saying "Holmes is a detective," speakers are playing along with the fiction, pretending to assert of a person they refer to as "Holmes" that he is a detective. But it is not fictional in the stories that Holmes is a fictional character. So, it is claimed, to say "Holmes is a fictional character" is not to play along with the fiction; the speaker must really be referring to something by means of "Holmes," not just pretending to, and attributing to the thing referred to the property of being a fictional character.

This line is not a sharp one, however, and in any case, it is not to be drawn in the place indicated. People often speak with tongue more or less evidently in cheek when what they are expressing is not fictional in an established work of fiction, not what a recognized work of fiction prescribes or invites them to imagine. We play along with established fictions in special or unusual or unauthorized ways, altering or extending them in various directions, in order to make serious points by engaging in pretense. Sometimes we improvise new fictions. The commentator who remarks that the Hardy Boys, still living at home and attending Bayport High, have turned 75, and that their publisher now equips them with cell phones, is speaking in pretense, although what he pretends to assert is not fictional in any of the stories or the series as a whole. He is, in effect, observing that the Hardy Boys stories have been published for 75 years, and that it is fictional in some of them that the brothers use cell phones. In explaining the tenets of a discredited scientific theory, we may convert it into a fiction, speaking as though we accept it as true. An example: "Vulcan is a planet in our solar system between Uranus and Neptune."

Pretense theorists propose to understand other kinds of apparent references to fictional objects and nonexistents as merely pretended, or at least as less than straightforwardly literal. Evaluating such proposals is not easy. Apparent references to wished-for children, failed scientific posits, and claims of existence and nonexistence often lack any apparent tongue-in-cheek flavor. But pretending, like other psychological states and processes, need not be explicit, conscious, or open to introspection. That people are engaging in pretense may be the conclusion of inferences to the best explanation. Moreover, there is room for adjusting or refining the notion of pretense, or replacing it with something weaker. Some pretense theorists prefer to characterize speakers as merely *making as if* referring to something.

In the end, what matters is the success of one or another variant of the pretense theory as a whole and how it compares to its competitors.

See also Existence; Fictionalism; Meinong, Alexius; Realism; Russell, Bertrand Arthur William.

Bibliography

Azzouni, Jody. *Deflating Existential Consequence: A Case for Nominalism*. Oxford, U.K.: Oxford University Press, 2004.

Currie, Gregory. *The Nature of Fiction*. Cambridge, U.K.: Cambridge University Press, 1990.

Donnellan, Keith. "Speaking of Nothing." *Philosophical Review* 83 (1974): 3–31.

Evans, Gareth. *The Varieties of Reference*. Oxford, U.K.: Oxford University Press, 1982.

Everett, Anthony, and Hofweber, Thomas, eds. *Empty Names, Fiction, and the Puzzles of Non-existence*. Stanford, CA: CSLI Publications, 2000.

Fine, Kit. "The Problem of Non-existence. I: Internalism." *Topoi* 1 (1982): 97–140.

Howell, Robert. "Fictional Objects: How They Are and How They Aren't." *Poetics* 8 (1979): 129–177.

Ingarden, Roman. *The Literary Work of Art: An Investigation on the Borderlines of Ontology, Logic, and Theory of Literature*. Translated by George G. Grabowicz. Evanston, IL: Northwestern University Press, 1973.

Kroon, Fred. "Belief about Nothing in Particular." In *Fictionalist Approaches to Metaphysics*, edited by Mark E. Kalderon. Oxford, U.K.: Oxford University Press, 2005.

Meinong, Alexius. "The Theory of Objects." In *Realism and the Background of Phenomenology*, edited by Roderick Chisholm. New York: Free Press, 1960.

Parsons, Terrance. *Non-existent Objects*. New Haven, CT: Yale University Press, 1980.

Quine, W. V. O. "On What There Is." In his *From a Logical Point of View*. New York: Harper, 1953.

Récanati, François. *Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation*. Cambridge, MA: MIT Press, 2000.

Routley, Richard. *Exploring Meinong's Jungle and Beyond: An Investigation of Noneism and the Theory of Items*. Canberra, Australia: Research School of Social Sciences, Australian National University, 1980.

Russell, Bertrand. "On Denoting." *Mind*, n.s. 14 (1905): 479–493.

Thomasson, Amie L. *Fiction and Metaphysics*. Cambridge, U.K.: Cambridge University Press, 1999.

Van Inwagen, Peter. "Creatures of Fiction." *American Philosophical Quarterly* 14 (1977): 299–308.

Voltolini, Alberto, ed. *Do Ficta Follow Fiction?* Special issue, *Dialectica* 57 (2003).

Walton, Kendall. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press, 1990.

Wolterstorff, Nicholas. *Works and Worlds of Art*. Oxford, U.K.: Clarendon Press, 1980.

Woods John. *The Logic of Fiction: A Philosophical Sounding of Deviant Logic.* The Hague: Mouton, 1974.

Yablo, Stephen. "Does Ontology Rest on a Mistake?" *Proceedings of the Aristotelian Society*, supp. vol. 72 (1998): 229–263.

Zalta, Edward N. *Intensional Logic and the Metaphysics of Intentionality*. Cambridge, MA: MIT Press, 1988.

Kendall L. Walton (1996, 2005)

NON-LOCALITY

LOCAL PHYSICAL MAGNITUDES

Non-locality, as the term suggests, is best approached via the notion of locality. As it will be seen, the notion of locality as it appears in physics has several components, but the foundational component is that of a local state in space-time. If one conceives of space-time as a four-dimensional (or eleven- or twenty-six-dimensional) manifold, one can think of covering the manifold with overlapping open neighborhoods, such that every point is contained within at least one neighborhood. One can also imagine indefinitely shrinking the size of the neighborhoods and indefinitely increasing their number. For any particular neighborhood, the intuitive notion of a neighborhood-local state is a physical state that depends only on what is inside the neighborhood. To get a more formal handle on this, a necessary condition for a neighborhood-local state is that the values of quantities for such a state put no constraints on the values of neighborhood-local states for any nonoverlapping neighborhoods. By this criterion, familiar physical properties, like the locations and velocities of particles or the values of electric fields, are neighborhood-local, while global physical properties, like the total charge of the universe, are not. (It is tempting to try to take this notion to the limit, where the neighborhoods become punctate, but this leads to many technical problems that are unrelated to the basic notion.)

Classical physics has many neighborhood-local quantities: for example, mass and charge densities, field strengths, velocities and accelerations, and the relativistic space-time metric. Anything represented by a tensor in a classical theory will be, by this account, neighborhood-local. Indeed, in classical physics it appears that all non-neighborhood-local quantities, such as the total charge of the universe, are functions of the neighborhood-local ones in the following sense: Cover the space-time manifold with open neighborhoods in any way one likes and specify the neighborhood-local quantities in each neighborhood and the neighborhood-local quantities in all intersections of neighborhoods, and one will thereby fix the value of the global quantities. Given the charge in every little patch, and in the intersections of all the little patches, the total charge of the universe follows.

Physics textbooks do not typically present the notion of neighborhood-locality in this way: They rather get at it via an account of coordinatizing the manifold. Rather than demanding a single, global coordinate system that completely covers a manifold (which in many cases will not exist), one is rather required only to break up the manifold into overlapping neighborhoods (each of which is topologically simple) and to coordinatize each neighborhood. The coordinatization of each neighborhood is called a chart, and a collection of charts for neighbor-

hoods that cover the manifold is called an atlas. In addition, one is required to specify how the coordinates assigned to a point in one chart are related to the coordinates assigned by any other chart in which the point occurs. That is, one is required to specify how the different coordinate systems relate to one another where they overlap. The assumption that all the physics is ultimately neighborhood-local is then essentially the supposition that physical states can be assigned to each charted neighborhood such that the total physical state of the universe is determined by the information in the atlas. One can say that in such a case the total physics is neighborhood-local.

The neighborhood-locality of physics accepts the physical reality of many global properties, such as the total charge. It also accepts the physical reality of more subtle global properties. Consider, for example, a cylinder and a Möbius strip. In a certain sense, a cylinder and a Möbius strip can be made to match locally: each can be divided into overlapping neighborhoods such that every neighborhood of the Möbius strip is exactly like the corresponding neighborhood of the cylinder. In this sense, the twist in the Möbius strip is not located anywhere in particular: it is a global rather than local feature of the space. Nonetheless, one could tell from an atlas whether one was dealing with a cylinder or a Möbius strip. Begin, for example, by drawing an F on one chart. The chart contains enough information to determine how the F could move rigidly in the neighborhood covered by the chart. So one could move it into a region that overlaps another chart, and the functions relating the chart would show how the F shows up in the new region. Continuing in this way, one could determine from the information in the atlas what the result of any rigid motion of the F would be. On a Möbius strip, some such motion will bring the F back to the original neighborhood mirror-reflected, while on a cylinder this can never happen.

The notion of neighborhood-locality is therefore quite broad: all of classical physics and relativity theory (both special and general) count as neighborhood-local in this way. One's commonsense picture of the world is also neighborhood-local. Albert Einstein powerfully expressed the notion of neighborhood-locality this way:

> It is … characteristic of … physical objects that they are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects "are situated in different parts of space." Unless one makes this kind of assumption about the existence (the "being-thus") of objects which are far apart from one another in space—which stems in the first place from everyday thinking—physical thinking in the familiar sense would not be possible. … This principle has been carried to extremes in the field theory by localizing the elementary objects on which it is based and which exist independently of each other, as well as the elementary laws which have been postulated for it, in the infinitely small (four-dimensional) elements of space.
>
> (in Born 1971, p. 170)

If one reads "situated in different parts of space" as "situated in nonoverlapping neighborhoods," and understands the "existence (the being-thus)" as the demand that the physical state defined on one neighborhood puts no constraint on the physical state in a nonoverlapping neighborhood, one sees that Einstein is expressing the same idea.

Suppose that physics is neighborhood-local in the sense that the physical information provided in any atlas is complete (determines all the physical properties of the universe). This appears to be a mild constraint, seeing as it takes in all of classical physics and relativity. It is hard to see, in fact, how the postulate of neighborhood locality puts any real empirical constraint on a theory: Could not any set of phenomena be accounted for by a neighborhood-local physics? As it will be seen, this is correct: To get an empirical constraint one will have to add on to neighborhood-locality in this sense. However, the postulate of neighborhood-locality does do something: It implies that, for any region in space-time, there is something that counts as the physical state of that region. Recall the twin requirements: The physical state in any neighborhood should not put any constraints on the physical state in a nonoverlapping neighborhood and the totality of physical states in an atlas (including appropriate information about overlapping charts) should determine the total physical state of the universe. Meeting these requirements demands that many well-defined quantities cannot count as local. For example, although the center of mass of the solar system is, one may suppose, always located at some particular point in space, it does not count as a part of the local physical state of that space. For taking a small neighborhood that contains that point, one cannot specify that the center of mass of the solar system occupies that point without thereby constraining the physical state of the nonoverlapping neighborhoods that contain the sun and planets.

If the requirement of neighborhood-locality is so mild, why was Einstein concerned with it? Because, taken at face value, the quantum theory rejects neighborhood-locality. Consider a pair of particles in the singlet state

$$1/\sqrt{2}|x\text{–up}>_R|x\text{–down}>_L - 1/\sqrt{2}|x\text{–down}>_R|x\text{–up}>_L$$

where the particle on the right and the particle on the left are far apart, in different regions of space. Then one's atlas could contain a neighborhood that includes particle R but not particle L, and a nonoverlapping neighborhood that contains L but not R. And the requirement of neighborhood-locality would then demand the existence of some physical state that can be assigned to (the neighborhood containing) R that puts no constraint on the state of L, and a state that can be assigned to L that puts no constraint on the state of R, such that from these two local states the singlet state for the pair can be recovered.

The singlet state itself cannot be used for this purpose: It makes reference to both particles and requires for its existence the existence of both particles. There is a well-defined state that quantum mechanics associates with particle R alone: It is called the reduced state for R from the singlet state. The reduced state supplies enough information to make quantum-mechanical (probabilistic) predictions for the result of any experiment carried out on R alone. There is a similar reduced state for L. These states are, mathematically speaking, mixed quantum mechanical states.

Why, then, can one not take the reduced state for R to be its neighborhood-local state, and the reduced state for L to be its neighborhood-local state, and do the physics using these? The reason is because different joint quantum mechanical states for the pair of particles give rise to exactly the same pair of reduced states for R and L, and these different joint states make different predictions for some measurements that involve both particles. For example, the singlet state is mathematically distinct from the m = 0 triplet state

$$1/\sqrt{2}|x\text{–up}>_R|x\text{–up}>_L + 1/\sqrt{2}|x\text{–down}>_R|x\text{–down}>_L.$$

Furthermore, the m = 0 triplet state makes different predictions for the pair: If one measures the spin of both particles in the *x*-direction, the singlet state predicts that the outcomes on the two sides will be different, while the m = 0 triplet state predicts they will come out the same. Even so, the reduced states for R and L that can be derived from these are identical (they both predict a 50 percent chance for the measurement of *x*-spin to be up). So one cannot use the joint state as a neighborhood-local state, and one cannot use the reduced states as neighborhood-local states (and recover the full physical state of the pair from the atlas), and quantum mechanics provides no other states one can use.

What Einstein saw was that quantum mechanics is not neighborhood-local on account of the entanglement of states for spatially separated systems. And since Einstein thought that physics must be neighborhood-local, he thought quantum mechanics must not be giving one a complete account of the physical states of things.

NON-LOCALITY AND EXPERIMENT

So far, all one has is a remark about the formalism of quantum mechanics, not about the empirical predictions of quantum mechanics. However, Einstein saw that the peculiar entanglement of quantum-mechanical states forced another kind of non-locality on the standard quantum mechanical accounts of experiments.

Consider a pair of separated particles in the singlet state. Given only that state, the quantum formalism permits no definite predictions about the outcome of an *x*-spin measurement on either side: For each individual particle, quantum mechanics assigns a 50 percent probability for each possible outcome. If the quantum description is complete and leaves no physical facts about the particle out of account, then these probabilities must reflect objective indeterminacy in nature: Nothing in the universe determines which outcome will occur. Nonetheless, as Einstein saw, quantum theory does make a perfectly definite prediction: Whatever the outcome of the experiments on the two particles, the results for the pair will be opposite—one will yield *x*-spin up and the other *x*-spin down (in the m = 0 triplet state, the results are instead guaranteed to be the same). So the question is: If nothing in the whole universe determines what the result of measuring the particle on the right will be, and if the particle on the left can be arbitrarily far away, what could possibly ensure that the outcome on the left will be the opposite of that on the right?

In the standard quantum formalism, this correlation between the outcomes is secured by the collapse of the wave function: when the particle on the right displays, for example, *x*-spin up, then the overall quantum state for the pair suddenly changes from $1/\sqrt{2}|x\text{–up}>_R|x\text{–down}>_L - 1/\sqrt{2}|x\text{–down}>_R|x\text{–up}>_L$ to $|x\text{–up}>_R|x\text{–down}>_L$. Because of the non-locality of the wave function, this change is a change not only in the physical state of particle R but a change in the state of particle L as well. When particle R displays *x*-spin up, particle L changes from a state of

indefinite *x*-spin to a state of definite *x*-spin down. It is by this "spooky action-at-a-distance" that the standard quantum interpretation manages to secure the correlation in spins between distant particles, neither of which is initially in a definite spin state.

There are several ways in which the wave collapse is "spooky." One is that is it unmediated: The measurement on the right influences the state on the left without the aid of any particles or waves traveling between the two sides. However, more important, the collapse is instantaneous: Even if there were mediating particles or waves, they would have to travel faster than light. This last property seems to contradict the theory of relativity. Einstein rejected the quantum theory because of this feature. He saw that, in this particular case, the spooky action-at-a-distance is not required by the empirical phenomena: The perfect correlations can be easily explained in a neighborhood-local physics without resorting to any direct causal connection between the two sides. One need only suppose (as the quantum theory does not) that the results of the spin measurements are predetermined by the local state of each electron and that the electrons are created in states in which they are disposed to give the opposite outcomes to all spin measurements.

Putting Einstein's two requirements together, one can now specify what it is for a theory to be simply local: First, all the fundamental physical properties of the theory should be neighborhood-local, and second, no physical influences in the theory should be allowed to propagate faster than light. (One could also add that causal connections between events should be mediated by continuous processes, but that is not needed for the sequel.) Einstein's argument against quantum theory as complete is that taking it to be complete requires that one treat the physics as non-local, even though the phenomena do not force non-locality on the theory. Einstein thought it perverse to insist that the theory is complete instead of trying to supersede it by a local theory that recovered all the same empirical predictions.

A local theory can be either deterministic or indeterministic. In a deterministic theory, every event is determined by the physical state that precedes it, and in a local deterministic theory, those determining factors cannot be so far away that it would require a superluminal influence for them to have their effect. Putting these together, it follows that in a local deterministic theory, every event is determined by the neighborhood-local state on its past light cone.

In an indeterministic local theory, an event need not be determined by the physical state of its past light cone, but the probability for the event will be. Furthermore, nothing outside the past light cone can have any influence on the event. That is, conditionalizing an event on the state of its past light cone should yield a probability that is screened off from any further information about events at space-like separation. (The probability will not be screened off from events in the future light cone, which can be effects of the event in question.) So positing that a theory is local is not the same as positing that it is deterministic, but it puts definite mathematical constraints on the nature of any local theory, whether deterministic or indeterministic. What Einstein had argued, in the 1935 Einstein, Boris Podolsky, and Nathan Rosen (EPR) paper, was that quantum mechanics is an indeterministic, non-local theory, but the sorts of correlations he discussed admit a deterministic, local explanation. And despite Einstein's oft-cited remarks about God's gambling habits, it was the spooky action-at-a-distance, the non-locality, that was the focus of his criticism in the EPR paper. As it turns out, if one is to recover the perfect EPR correlations with a local theory, it must also be a local deterministic theory (otherwise the correlations will not be perfect), but recovering determinism is not the main issue.

BELL'S THEOREM AND LOCALITY

What Einstein did not realize is that although the perfect correlations he discussed can be recovered by a local theory, the full range of quantum mechanical predictions cannot be recovered by any local theory. This was proven in 1964 by John Bell. Bell demonstrated that the predictions of any local theory, deterministic or indeterministic, must satisfy a certain statistical constraint called Bell's inequality. Furthermore, the predictions of the quantum theory violate that inequality, and the violations have been experimentally confirmed in the laboratory. So the non-locality of quantum theory is not just an artifact of the quantum formalism: It is a physical aspect of nature.

Although in principle a neighborhood-local theory could predict violations of Bell's inequality (by use of neighborhood-local items that travel faster than light), the only presently existing accounts of physical non-locality employ the quantum wave function, which is not a neighborhood-local object. The role of the wave function differs from interpretation to interpretation, but in every case it is the wave function that secures the violation of locality and the superluminal physical connection between the distant particles.

It is a first-order technical problem to reconcile the non-locality of quantum theory with the space-time structure postulated by the theory of relativity. The sim-

plest way to construct a non-local theory is to add a preferred foliation of space-time to the relativistic picture, thereby violating the spirit of relativistic physics. Such a foliation also allows for a straightforward causal account of the phenomena: Intraction with one of the particles is the cause of a change of behavior in the other. It is, however, feasible (although quite tricky) to construct theories that achieve non-locality but employ only the relativistic space-time structure. In these cases, it appears that standard causal locutions cannot the recovered: there is a real physical connection between space-like separated events, but one cannot identify one of the events particularly as a cause and the other as an effect.

See also Bell, John, and Bell's Theorem; Einstein, Albert; Philosophy of Physics; Quantum Mechanics; Relativity Theory; Space; Time.

Bibliography

Born, M. *The Born-Einstein Letters.* Translated by I. Born. New York: Walker, 1971.

The Collected Papers of Albert Einstein. Vols. 1–6, 8. Translated by Anna Beck. Princeton, NJ: Princeton University Press, 1987–1998.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777–780.

Maudlin, Tim *Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics*. Oxford, U.K.: Basil Blackwell, 1994.

Tim Maudlin (2005)

NON-MONOTONIC LOGIC

Modern symbolic logic was developed beginning in the latter part of the nineteenth century for the purpose of formalizing mathematical reasoning, in particular that process by which mathematicians arrive at conclusions on the basis of a small number of distinct basic principles. This kind of reasoning is characterized by a particular type of cogency: The conclusions are not merely *probable* or *plausible* on the basis of whatever evidential support the basic principles might provide, but *certain* and *indubitable*. In particular mathematical reasoning enjoys a property referred to as *monotonicity* by modern logicians: if a conclusion follows from given premises *A*, *B*, *C*, … then it also follows from any larger set of premises, as long as the original premises *A*, *B*, *C*, …are included.

By contrast in many instances of ordinary or everyday reasoning, people arrive at conclusions only tentatively, based on partial or incomplete information, reserving the right to retract those conclusions should they learn new facts. Such reasoning is often called defeasible or non-monotonic, precisely because the set of accepted conclusions can become smaller when the set of premises is expanded.

Taxonomies provide a rich source of examples of defeasible reasoning (but they are not by any means the only source). Suppose for instance that you are told that Stellaluna is a mammal. It is then natural to infer that Stellaluna does not fly, because mammals by and large are not capable of flight. But upon learning that Stellaluna is a bat, such a conclusion is retracted in favor of its opposite. In turn even the new conclusion can be retracted upon learning that Stellaluna is a baby bat and so on, in complex retraction patterns that seem to cry out for systematization.

The aim of non-monotonic logic is precisely that of providing such a systematization. There is, in fact, no one thing which is called "non-monotonic logic," but rather a family of different formalisms, with different mathematical properties and degrees of material adequacy, that aim to capture and represent such patterns of defeasible reasoning.

A broad class of non-monotonic formalisms can be characterized as "consistency-based" approaches. The name is derived from the fact that while all non-monotonic formalisms deal with conflicts between new facts and tentative conclusions in the same way (the facts win and the conclusions are retracted), some of these formalisms also allow for potential conflicts between the tentative conclusions themselves (and then they might differ as to the way this second kind of conflicts are handled).

Non-monotonic inheritance networks provide a consistency-based formalism developed for the purpose of representing taxonomies. A non-monotonic inheritance network is a collection of nodes (each associated with a particular taxonomic category) and directed links between nodes, representing the subsumption relation between categories. Suppose for instance that you are told by a reliable (but fallible) source that Nixon is both a Quaker and a Republican, and that while Quakers by and large are pacifists, Republicans are not. The network corresponding to this situation is given below:

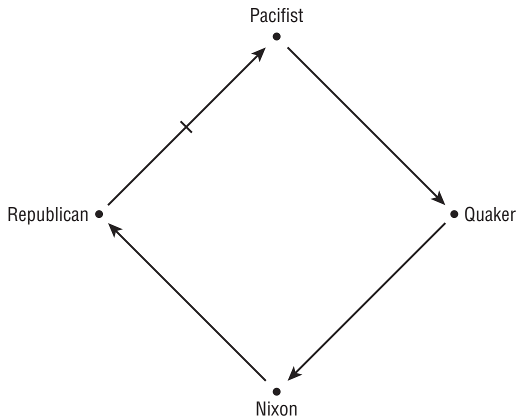

Obviously here we have a conflict between the two potential conclusions that Nixon both is and is not a pacifist. Steps need to be taken to maintain consistency. We will not go into detail here, but in general one can take a *credulous* approach and endorse one or the other conclusion, or one can take a *skeptical approach* and in the presence of conflict refrain from endorsing either conclusion.

Sometimes, special considerations such as *specificity* can be brought to bear on the resolution of conflicts in other inheritance networks. In the Stellaluna example above for instance one wants to conclude that bats fly (because information about bats is more specific than information about mammals) but that Stellaluna does not (because information about baby bats is more specific than information about bats). A network representing the situation is given below:

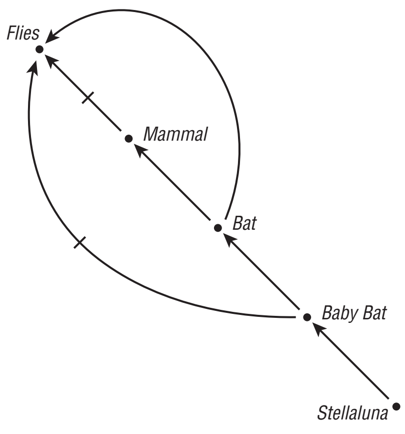

Inheritance networks are not well suited to deal with complex information (e.g., disjunctive or conjunctive statements). For this reason a more expressive formalism, *default logic* was developed. The basic representation formalism of default logic is the *default inference rule*, a rule of the form $A : B / C$, whose intended interpretation is that if A is known, and we have no reason to reject B (i.e., B is *consistent* with our knowledge base), then we can conclude C. Default logic provides a way for the consistency condition to be satisfied both *before* and *after* the default rule is applied.

Among the approaches to non-monotonic logic that are not consistency based, one needs to mention *circumscription*, which is based on the idea that many instances of defeasible reasoning have to do with the *minimization* of certain predicates, particularly those representing the set of *exceptions* to a given generalization. Circumscription uses the expressive power of second-order logic to ensure that any generalization has as few exceptions as possible. So, for instance, in the absence of information to the effect that bats are exceptional mammals, one would conclude that they do not fly, but when that information is adjoined to our knowledge base, circumscription immediately accounts for the exception.

See also Computationalism; Logic, History of: Modern Logic; Mathematics, Foundations of.

Bibliography

Antonelli, Aldo. " Non-monotonic Logic." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Available from http://plato.stanford.edu/archives/sum2003/entries/logic-nonmonotonic/.

Antonelli, Aldo. " Logic." In *The Blackwell Guide to the Philosophy of Computing and Information*, 263–275. Blackwell, 2004.

Gabbay, Dov, Christopher Hogger, and John Alan Robinson, eds. *Handbook of Logic in Artificial Intelligence and Logic Programming*. Vol. 3. New York: Oxford University Press, 1994.

Ginsberg, Matthew, ed. *Readings in Nonmonotonic Reasoning*. Los Altos, CA: Morgan Kauffman, 1987.

G. Aldo Antonelli (2005)

NONNATURALISM

See *Ethics, History of; Ethics, Problems of; Moore, George Edward*

NONREDUCTIVE PHYSICALISM

Beginning the 1960s Hilary Putnam, Jerry Fodor, and Richard Boyd, among others, developed a type of materialism that denies reductionist claims. In this view, explanations, natural kinds, and properties in psychology do not reduce to counterparts in more basic sciences, such as

neurophysiology or physics (Putnam 1967, 1974; Fodor 1974; Boyd 1980a). Nevertheless, all token psychological entities—states, processes, and faculties—are either identical with (Fodor 1974) or just wholly constituted of (Boyd 1980a) physical entities, ultimately out of token entities over which microphysics quantifies. This view was soon widely endorsed and since then has persisted as an attractive alternative to reductionist and eliminativist forms of materialism. Reductionists, notably Jaegwon Kim, have raised a series of serious objections to this position, to which nonreductivists have responded, thereby developing the view more thoroughly.

IRREDUCIBILITY, MULTIPLE REALIZABILITY, AND EXPLANATION

In his early argument for nonreductive materialism, Putnam adduces the phenomenon of multiple realizability as its main justification (Putnam 1967). Kinds or types of mental states can be realized by many kinds of neurophysiological states, and perhaps by many kinds of non-neurophysiological states, and for this reason they do not reduce to kinds of neurophysiological states. Multiple realizability also has a key role in Fodor's more general argument against reductionism in the special sciences (Fodor 1974). Consider a law in some special science:

$$S_1x \text{ causes } S_2x$$

where S_1 and S_2 are natural kind-predicates in that science. A standard model for reduction requires that every kind featured in this law be identified with a kind in the reducing science, by way of bridge principles. Bridge principles might translate kind-predicates in one science into those of a more basic one, or they might specify a metaphysical relation, such as *being identical with* or *being a necessary and sufficient condition for*, between the kinds of one science and those of the reducing science. But in some cases, Fodor contends, the sort of bridge principle required for reducibility will not be available. If kinds in psychology, for instance, are multiply realizable in an indefinite variety of ways at the neurophysiological level, purported bridge principles for relating psychological to neurophysiological kinds will involve open-ended disjunctions. These purported bridge principles will be of the form:

$$P_1 = N_1 \text{ v } N_2 \text{ v } N_3 \ldots$$

which states that a certain psychological state, P_1, is identical with an open-ended disjunction of neurophysiological states, N_1 v N_2 v N_3 … , or

$$P_1 \leftrightarrow N_1 \text{ v } N_2 \text{ v } N_3 \ldots.$$

which states that a certain psychological state is necessary and sufficient for an open-ended disjunction of neurophysiological states. Fodor argues that because open-ended disjunctions of kinds in neurophysiology are not natural neurophysiological kinds, psychological kinds cannot be reduced to neurophysiological kinds. Fodor's reason for denying that such disjunctions are not natural kinds is that they cannot appear in laws, and they cannot appear in laws because "laws" involving such disjunctions are not explanatory. Such "laws" are not explanatory because they do not satisfy our interests in explanation. Fodor's argument for irreducibility, then, appeals to the fact that purported explanations for psychological phenomena are unsatisfying when couched in terms of open-ended disjunctions.

One reductionist reply is that these open-ended disjunctions nevertheless constitute genuine laws and explanations, even if they fail to meet certain subjective requirements. If only we were capable of taking in more information at once, we wouldn't have any trouble regarding open-ended disjunctive "laws" as genuine laws (Jaworski 2002). That people fail to find laws satisfying when they contain open-ended disjunctions may simply show a failing on our part, rather than a failing of the putative laws. This standard argument for nonreductive materialism appears to rely on a certain formal prescription for laws and explanations—that they cannot contain disjunctive properties, or at least not wildly disjunctive properties.

But even if the formal argument fails, multiple realizability can still sustain an important component of nonreductive materialism. In general, whether or not a property is multiply realizable can indicate the level at which it should be classified. Is the kind *corkscrew* a kind of steel thing? No, for it also has a possible aluminum realization. Is the kind *believing that cats are nearby* a neural kind of thing? If mental states are also realizable in silicon, then no. Multiple realizability might then provide the key to precluding classification of mental states as essentially neural, or as essentially classified at some lower level yet.

Kim argues that multiple realizability might fail to undermine reductionism for a different reason. He contends that a higher-level property is precisely as projectible as the disjunction that expresses its multiply realizable character at a more basic level, and thus a generalization involving such disjunctive properties is just as lawlike as the higher-level generalization that it was

meant to reduce (Kim 1992). The reason is that a higher-level property is nomically equivalent to such a disjunctive property. Nomic equivalence might be defined in this way: properties F and G are nomically equivalent if they are coextensive in all possible worlds compatible with the laws of nature. If Kim is right, then Fodor's formal argument does not appear to be sound, for it relies on the possibility that generalizations involving a higher-level property be lawlike whereas those involving the corresponding disjunctive property are not. But furthermore, Kim contends that wildly disjunctive properties are not projectible, and hence higher-level properties that are nomically equivalent to such properties are not projectible either. As a result, such higher-level properties cannot figure into laws, and they are not genuinely scientific kinds.

The example of a disjunctive property Kim adduces to make his point is *being jade*. "Jade" is a category that comprises two mineralogical kinds, *jadeite* and *nephrite*, and hence *being jade* is the same property as *being either jadeite or nephrite*. As a result, *being jade* will not be projectible. But in reply, *being jade* might turn out to be projectible despite its underlying complexity. Ned Block points out that all samples of jade share certain appearance properties, similarities that give rise to a certain degree of projectibility (Block 1997). More generally, properties that are multiply realizable can yet be projectible with respect to properties of selection, learning, and design. Because there are typically only a few ways in which entities of a particular higher-level type can be designed and produced, one can expect relatively broad similarities among these things that would render corresponding higher-level properties significantly projectible (Antony and Levine 1997).

Thus the heterogeneity of the possible realizations of a property is compatible with their having significant features in common, features that will sustain the projectibility of the property to some degree or other. This point is consistent with Kim's claim that a higher-level property is precisely as projectible as the disjunctive property that comprises all of its possible realizations. One should not conclude from the heterogeneity of the possible realizations of a higher-level property that there is no feature that can undergird its projectibility—in fact, of both the higher-level property and of the disjunctive property that comprises all of its possible realizations. Indeed, the projectibility-sustaining feature of a kind could be a characteristic that is significantly homogeneous across its heterogeneous realizations, one that might instantiate a unitary causal power at the level of description of the kind (Pereboom 2002).

FUNCTIONALISM AND MENTAL CAUSATION

By way of objecting to Kim's reductionism, Block asks: "What is common to the pains of dogs and people (and all other species) in virtue of which they are pains?" (Block 1980, pp. 178–179). In reply to this concern, Kim points out that nonreductive materialists typically argue from a functionalist perspective, and that functionalists characterize mental states solely in terms of purely relational features of those states. Functionalism identifies mental state types with type-level dispositions to cause mental states and behavioral outputs given perceptual inputs and mental states—with the understanding that these dispositions are purely relational: that they are to be analyzed in terms of causal relations to perceptual inputs, behavioral outputs, and other mental states, and no intrinsic mental components. Functionalists claim that what all pains would have in common, by virtue of which they are all pains, is a pattern of such relations described by some functional specification. Kim then argues that in providing an answer to Block's question, the local reductionist—the one who opts for species- or structure-specific reductionism—is no worse off than the functionalist. Both are committed to the claim that there is no nonrelational or intrinsic property of pain that all pains have in common, and both can specify only shared relational properties (Kim 1992).

Kim implies that a functional specification does not provide a genuinely satisfactory answer to Block's question (Kim 1999). On the nonreductive view, if M is a mental property and B is its neural or microphysical base, then realizers for M can be found in B (at the level of B). This position allows that nondisjunctive realizing properties might be found in B for individual species- or structure-types—as long as there is no well-behaved (not wildly disjunctive) property in B that realizes every possible instance of M. The nonreductive materialist claims that none of this entails a genuine reduction of M to properties in B. As Kim assumes, the standard strategy for preserving M as meeting these specifications is to envision M as a functional mental property. But in Kim's view, the problem with the functionalist picture is that the causal powers of any instance of M will be causal powers in the physical base—they will not, at the token level, be irreducibly mental causal powers (Kim 1992, Block 1990). Hence functionalism cannot preserve the view that there exist causal powers that are in the last

analysis irreducibly mental, and it is thus incompatible with a genuinely robust nonreductive materialism about the mental. Furthermore, Kim points out that given the genuine multiple realizability of the property M, the causal powers of the realizers of M in B will exhibit significant causal and nomological diversity, and for this reason the causal powers of M will exhibit such diversity. Thus, in his estimation, M will be unfit to figure in laws, and is thereby disqualified as a useful scientific property. He concludes that the functionalist model cannot protect m as a property with a role in scientific laws and explanations.

However, there is available a nonfunctionalist account of these higher-level powers that nevertheless remains nonreductive (Pereboom 1991, 2002). Functionalists typically maintain that the causal powers that have a role in explaining the dispositional features of mental states are nondispositional properties of their realization bases. For example, many suppose that nondispositional neural properties, which instantiate neural causal powers, would serve to explain why being pinched causes wincing behavior. But if these causal powers are all nonmental, a robust sort of nonreductive materialist account of the mental is precluded, for then none of the causal powers would be essentially mental themselves. By contrast, the nonreductivist might endorse intrinsic mental properties that instantiate specifically mental causal powers (Pereboom 1991, 2002; Van Gulick 1993). Such a view would be incompatible with functionalism. It need not deny that there exist functional mental properties, or, more generally, relational properties of mental states, but it would endorse nonfunctional mental properties that, by virtue of the causal powers they instantiate, play an important part in explaining dispositional features of mental state types.

Consider the example of a ball piston engine, the most recent version of the rotary internal combustion engine, which has a specific internal structural configuration. Characteristic of this engine is its having parts with particular shapes and rigidities, and these parts must be arranged in a particular way. These features are manifestly not functional relations that such an engine stands in; rather, they constitute intrinsic characteristics of this type of engine. At the same time, these characteristics are multiply realizable. The parts of the engine can be made of material of different sorts—as long as the material can yield, for example, the required shapes and rigidities. The ball piston engine, then, has nonfunctionalist intrinsic structural properties that instantiate its causal powers, but nevertheless admit distinct realizations.

Similarly, it might be that the heterogeneous physical realizations of the dog's and the human's belief *that cats are nearby* exhibit a structure of a single type that is intrinsic to this kind of mental state, a structure that instantiates the causal powers of this belief. This structure may be more abstract than any specific sort of neural structure, given that it can be realized in distinct sorts of neural systems (Boyd 1999). Perhaps this same structure can be realized in a silicon-based electronic system, and such a system could then also have the belief. Imagine a silicon system that replicates the capacities of and interconnections among neurons in a human brain as closely as possible, and suppose this system is excited to mimic as nearly as possible what happens when a human being has this belief about cats. It is possible that this silicon state would realize the same belief, and have a structure that, conceived at a certain level of abstraction, is similar enough to the structure of the ordinary neural system for both to count as examples of the same type of structure. In this case and more generally, one does not seem forced to retreat to mere functional resemblance prior to investigating whether the relevant similarities extend to intrinsic properties.

EXPLANATORY EXCLUSION

According to nonreductive materialism, an event such as Jerry's feeding the cat (M2) will have a psychological explanation in terms of a complex of mental states—beliefs and desires he has (M1). Each of M1 and M2 will be wholly constituted of microphysical events (P1 and P2 respectively), and there will be a microphysical explanation of P2 in terms of P1. The explanation of M2 by M1 will not reduce to the explanation of P2 by P1. Underlying the irreducibility of this explanation is that M1 is not type-identical with P1, and that M2 is not type-identical with P2.

This picture gives rise to a pressing question: What is the relationship between the microphysical and psychological explanations for M2? In particular, given that both sorts of explanation refer to causal powers, what is the relationship between the causal powers to which the microphysical explanation appeals and those to which the psychological explanation appeals? Here is where Kim's challenge from causal or explanatory exclusion enters in (Kim 1987, 1998). If a microphysical account yields a causal explanation of the microphysical constitution of M2, then it will also provide a causal explanation of M2 itself. How might there also be a distinct psychological causal explanation of this action? Kim argues that it is implausible that the psychological explanation appeals to

causal powers sufficient for the event to occur, and at the same time the microphysical explanation appeals to distinct causal powers also sufficient for the event to occur, as a result of which the event is overdetermined. It is also implausible that each of these distinct sets of causal powers yields a partial cause of the event, and that each by itself would be insufficient for the event to occur.

By the solution to this problem that Kim develops, real causal powers exist at the microphysical level, and so the microphysical explanations refer to real microphysical causal powers. Only if psychological explanations in some sense reduce to microphysical explanations does it turn out that the psychological explanations also appeal to real causal powers—these causal powers will then ultimately be microphysical. Psychological explanations that do not reduce to microphysical explanations will fail to refer to causal powers, and thus will have some diminished status—such explanations might express regularities without at the same time referring to causal powers. This strategy solves the exclusion problem because if the causal powers to which the psychological explanation appeals are identical with those to which the microphysical explanation appeals, then there will be no genuine competition between explanations, and if the psychological explanations do not refer to causal powers at all, there will be no competition either. However, this solution, which Kim believes is the only possible solution to the problem he raises, would rule out any nonreductive view about mental causal powers.

Various proposals have been advanced in the name of nonreductive materialism according to which mental properties are causally relevant or causally explanatory, without being causally efficacious as mental properties. Such views, like Kim's, claim that all causal efficacy is nonmental (for example, Jackson and Pettit 1990). As Kim points out, these proposals do not amount to a robust sort of nonreductive materialism, which would preserve the claim that mental properties, as mental properties, are causally efficacious (Kim 1998).

What sort of response might the advocate of the robust view provide? First, in Kim's conception, any token causal powers of a higher-level property at a time will be identical with some token (micro)physical causal powers. There would be no token causal powers distinct from token microphysical causal powers, and this would preclude any robust nonreductive materialism. Higher-level kinds and explanations would at best group token microphysical causal powers in a way that does not correspond to the classifications of microphysics itself (Kim 1998, Horgan 1997). Such a classification might be of value for prediction, but there would remain no sense in which there exist causal powers that are not microphysical.

However, is token mental state M identical with P, its actual token microphysical realization base? Suppose that M is realized by a complex neural state N. It is possible for M to be realized differently only in that a few neural pathways are used that are token-distinct from those actually engaged. One need not rule at this point on whether the actual neural realization N is token-identical with this alternative—it might well be. But it is evident that this alternative neural realization is itself realized by a microphysical state P* that is token-distinct from P. It is therefore possible for M to be realized by a microphysical state not identical with P, and thus M is not identical with P. But furthermore, this reflection would also undermine a token-identity claim for mental causal powers—should they exist—and their underlying microphysical causal powers. For supposing that the token microphysical realization of M had been different, its token microphysical causal powers would also have been different. Consequently, there is good reason to suppose that any token mental causal powers of M would not be identical with the token microphysical causal powers of its realization (Boyd 1980a, Pereboom and Kornblith 1991, Pereboom 2002).

On this conception, a token mental state would have the mental causal powers it does ultimately by virtue of the token microphysical states of which it is constituted (setting aside any fundamentally relational causal powers). For this reason it makes sense to say that token mental causal powers are wholly constituted by token microphysical causal powers. More generally, the causal powers of a token of kind F are constituted of the causal powers of a token of kind G just in case the token of kind F has the causal powers it does by virtue of its being constituted of a token of kind G.

And now, just as no competition between explanations arises in the case of reduction and identity, competition also does not arise in the case of mere constitution. For if the token of a higher-level causal power is currently wholly constituted by a complex of microphysical causal powers, there are two sets of causal powers at play that are constituted from precisely the same material (supposing that the most basic microphysical entities are constituted of themselves), and in this sense we might say that these powers *coincide constitutionally*. That they now coincide in this way might give rise to the thought that these causal powers are token-identical, but, as has been shown, there is a substantial argument that they are not. And because it is possible for there to be wholly constitutionally coin-

ciding causal powers that are not even token-identical, it is possible that there be two causal explanations for one event that do not exclude each other and at the same time do not reduce to a single explanation (Pereboom 2002).

If identity and not just constitutional coincidence were necessary for explanatory noncompetition, then there would be features required for noncompetition that identity has and current constitutional coincidence does not. The candidate features would be constitutional coincidence at all other times, and constitutional coincidence at all other possible worlds, even now. But it is difficult to see how the token causal powers' constitutional noncoincidence at some past time, or at some future time, or their merely possible constitutional noncoincidence even now would result in explanatory competition, whereas actual current constitutional coincidence in absence of any features of this sort (i.e., identity) would guarantee noncompetition.

Imagine that a person's current token mental state M actually constitutionally coincides with token microphysical state P. Now assume with Kim that if M were identical with P, and if their causal powers were identical, there would be no explanatory competition. Then if mere constitutional coincidence without identity resulted in explanatory competition, that would have to be because at some time in the past or in the future, or at some other possible world even now, M and P and their causal powers are constitutionally noncoincident. Suppose that M would still exist even if a few neural pathways in its neural realization were token-distinct from what they actually are. These neural changes would render M's microphysical realization base distinct from P, and thus M and P would be constitutionally noncoincident in some other possible world, and, similarly, *mutatis mutandis* (that is, the necessary changes having been made) for their causal powers. How could a possibility of this sort introduce explanatory competition? It would appear that actual current constitutional coincidence alone is relevant to securing noncompetition, and thus for this purpose constitutional coincidence without identity would serve as well as identity. Consequently, it would appear that available to the nonreductivist is a solution to the exclusion problem no less adequate than Kim's own.

THE THREAT OF EMERGENTISM

Kim contends that nonreductive materialism is committed to emergentism (sometimes called *strong* emergentism, which he thinks is a radical and implausible view. In his analysis, emergentism claims a distinction between two sorts of higher-level properties, *resultant* and *emergent*, that arise from the basal conditions of physical systems (Kim 1999). The basal conditions of a physical system comprise (i) the basic particles that constitute the physical system, (ii) all the intrinsic properties of these particles, and (iii) the relations that configure these particles into a structure. The higher-level properties that are merely resultant are simply and straightforwardly calculated and theoretically predictable from the facts about its basal conditions—which presumably include the laws that govern the basal conditions—whereas those that are emergent cannot be calculated and predicted. Theoretical predictability contrasts with inductive predictability. Having regularly witnessed that an emergent property is realized by particular basal conditions, we would be able to predict this relationship, but this sort of inductive predictability is not at issue. Rather, according to emergentism, knowledge of the basal conditions alone, no matter how complete, does not suffice to yield a prediction of an emergent property.

Emergentism also endorses downward causation; it claims that higher-level states can have lower-level effects. Emergentism about the mental asserts that mental events can cause microphysical events. Plausibly, nonreductive materialism also countenances downward causation of this sort—M1 causes M2, but because M2 is wholly constituted of P2, M1 also causes P2. Kim thinks that by virtue of endorsing this sort of downward causation, nonreductive materialism is committed to emergentism.

However, the nonreductive view's allowing for downward causation is not by itself sufficient to render it emergentist. Endorsement of downward causation would indeed be radical if it also specified that mental properties could effect changes in the laws that govern the microphysical level independently of any emergent properties (call them the *ordinary* microphysical laws). Supposing that M1 were such an emergent mental property, M1 could cause P2 in such a way that P2 is no longer governed by the ordinary microphysical laws, but instead by laws that take into account the special characteristics of the emergent properties, or no laws at all. But nothing essential to nonreductive materialism entails this radical variety of downward causation (Pereboom 2002).

We might suppose that the capacity for altering the ordinary microphysical laws is what provides emergent properties with their distinctive nature. And this potentially explains why such properties would not be predictable from the microphysical base together with these ordinary laws. Information about the ordinary laws and the microphysical base might be insufficient to predict the law-altering behavior of the higher-level property. But

there is no feature of the nonreductive model per se that renders higher-level properties any less theoretically predictable than they would be on a reductive model. In each model, holding relational conditions fixed, a particular set of basal conditions will necessitate the same unique higher-level properties. The nonreductivist is no more committed to some factor that threatens theoretical predictability, such as the capacity of higher-level properties to alter the ordinary microphysical laws, than is the reductionist.

Arguably, therefore, nonreductive materialism can respond effectively to the most serious arguments made against it over the last forty years, and as a result, it remains a viable position about the nature of the mental.

See also Functionalism; Mind-Body Problem; Multiple Realizability; Physicalism.

Bibliography

Antony, L., and J. Levine. "Reduction with Autonomy." In *Philosophical Perspectives*. Vol. 11, *Mind, Causation, and World*, 83–105. Oxford: Blackwell, 1997.

Block, N. "Anti-Reductionism Slaps Back." In *Philosophical Perspectives*. Vol. 11, *Mind, Causation, and World*, 107–132. Oxford: Blackwell, 1997.

Block, N. "Can the Mind Change the World?" In *Meaning and Method: Essays in Honor of Hilary Putnam*, edited by George Boolos, 137–170. Cambridge, U.K.: Cambridge University Press, 1990.

Block, N. "Introduction: What is Functionalism?" In *Readings in the Philosophy of Psychology*, vol. 1, edited by Block, 178–179. Cambridge, MA: Harvard University Press, 1980.

Boyd, R. "Kinds, Complexity, and Multiple Realization." *Philosophical Studies* 95 (1999): 67–98.

Boyd, R. "Materialism Without Reductionism." In *Readings in the Philosophy of Psychology*. Vol. 1, edited by N. Block, 67–106. Cambridge, MA: Harvard University Press, 1980a.

Boyd, R. "Scientific Realism and Naturalistic Epistemology." In *Proceedings of the Philosophy of Science Association*. Vol. 2. East Lansing, MI: Philosophy of Science Association, 1980b.

Clapp. L. "Disjunctive Properties: Multiple Realizations." *Journal of Philosophy* 98 (2001): 111–136.

Fodor, J. "Special Sciences: Still Autonomous After All These Years." In *Philosophical Perspectives*. Vol. 11, *Mind, Causation, and World*, 149–163. Oxford: Blackwell, 1997.

Fodor, J. "Special Sciences (or: The Disunity of Science as a Working Hypothesis)." *Synthese* 28 (1974): 97–115.

Heil, J. "Multiple Realizability." *American Philosophical Quarterly* 36 (1999): 189–208.

Horgan, T. "Kim on Mental Causation and Causal Exclusion." In *Philosophical Perspectives*. Vol. 11, *Mind, Causation, and World*, 165–184. Oxford: Blackwell, 1997.

Jackson, F., and P. Pettit. "Program Explanation: A General Perspective." *Analysis* 50 (1990): 107–117.

Jaworski, W. "Multiple Realizability, Explanation, and the Disjunctive Move." *Philosophical Studies* 108: (2002): 298–308.

Kim, J. "Making Sense of Emergence." *Philosophical Studies* 95 (1999): 3–36.

Kim, J. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.

Kim, J. "Multiple Realizability and the Metaphysics of Reduction." *Philosophy and Phenomenological Research* 52 (1992): 1–26.

Kim, J. "The Myth of Nonreductive Materialism." *Proceedings and Addresses of the American Philosophical Association* 63 (1989): 31–47.

Kim, J. "Phenomenal Properties, Psychophysical Laws, and the Identity Theory." *Monist* 56: (1972): 177–192.

Kitcher, P. S. "1953 and All That: A Tale of Two Sciences." *Philosophical Review* 93 (1984): 335–373.

Pereboom, D., and H. Kornblith. "The Metaphysics of Irreducibility." *Philosophical Studies* 63 (1991): 125–145.

Pereboom, D. "Robust Nonreductive Materialism." *Journal of Philosophy* 99 (2002): 499–531.

Pereboom, D. "Why a Scientific Realist Cannot Be a Functionalist." *Synthese* 88 (1991): 341–358.

Putnam, H. "Language and Reality." In his *Philosophical Papers*, vol. 2, 272–290. Cambridge, MA: Harvard University Press, 1975. This paper was delivered as a Machette Lecture at Princeton University, May 22, 1974.

Putnam, H. "The Nature of Mental States." In his *Philosophical Papers*. vol. 2, 429–440. Cambridge, MA: Harvard University Press, 1975. First published as "Psychological Predicates," in *Art, Mind, and Religion*, edited by W. H. Capitan and D. D. Merill, 37–48 (Pittsburgh: Pittsburgh University Press, 1967).

Shapiro, L. "Multiple Realizations." *Journal of Philosophy* 97 (2000): 635–654.

Sober, E. "The Multiple Realizability Argument Against Reductionism." *Philosophy of Science* 66 (1999): 542–564.

Van Gulick, R. "Who's in Charge Here? And Who's Doing All the Work?" In *Mental Causation*, edited by J. Heil and A. Mele, 233–256. Oxford: Oxford University Press, 1993.

Yablo, S. "Mental Causation." *Philosophical Review* 101 (1992): 245–280.

Derk Pereboom (2005)

NON-TRUTH-CONDITIONAL MEANING

There are two dominant approaches to semantics. One sees the task of semantics as to provide a systematic account of the truth conditions of (actual and potential) sentence uses. The other assumes that a use of a sentence expresses a statement (proposition, thought—terminology varies here), a statement being the sort of thing that can be asserted and believed, and also the sort of thing that, as a representation of how the world is, can be

assessed as true or false. The task of semantics, on this view, is systematically to spell out how sentence uses are associated with statements.

While the aims of the two types of theories are different, they are related. A use of a sentence to make a statement is, after all, presumably true (or false) in virtue of the truth (or falsity) of the statement made. Hence, to assign statements to sentence uses is to assign those uses truth conditions. Thus both approaches give pride of place in semantics to an account of how sentence uses come to be true or false.

No one thinks that giving an account of truth conditions or of what statements say, for a language, says all there is to say about conventional meanings of expressions in the language, though exactly what more there might be is a matter of controversy. Here are the main candidates for what might be left out of such accounts.

MOOD AND FORCE

The theories just discussed aim to illuminate what is going on when one uses the sentences of a language to make assertions, to commit to the truth of a claim. But, of course, we can do much more than make assertions with our sentences, and some aspects of conventional meaning are obviously keyed to doing things other than asserting. Examples are grammatical and phonological forms associated with questioning, ordering, and exclaiming. It is a fact about conventional meaning if anything is, that subject/auxiliary inversion is used to question in French, German, and English, that prefixing a declarative sentence in English with "if only" signals a wish, that sentences such as "Yuck!" and "Damn it!" express attitudes that are not to be evaluated as true or false. One task not discharged by truth-conditional or statement semantics, then, is detailing when and how linguistic forms have as part of their conventional meaning the task of signaling that a particular sort of speech act (asserting, questioning, promising, warning, expressing disgust, etc.) is being performed.

One might question the extent to which this is more than just an appendicle to truth-conditional or statement semantics. One might say that interjections like "Grody!" and "Awesome!" are elliptical for truth bearers ("That is grotesque!" "That's awesome!") uttered with a particular force. Whether or not this is so, the interjections do not combine with connectives and the range of sentences in the language to produce complex sentences; their meanings, if different from that of declaratives, would thus seem to be walled off from other aspects of meaning. There seems to be a rather small catalog of devices, like auxiliary inversion and the subjunctive, to indicate force; such devices, furthermore, do not seem to be iterable, as constructions that contribute to truth conditions are. While one can disjoin a negation, then enclose the result inside the consequent of a conditional, etc., force indicators *seem* by and large to exclude one another (one cannot, for example, turn the optative "would that he were gone" into a question). Furthermore, it is not clear that any particularly novel sort of meaning is required in an account of the meanings of, for example, orders and questions. One might suspect that in some sense the content of the declarative "You will sit" and of the imperative "Sit!" are the same, the difference lying only in the force of their utterance. Perhaps questions have a slightly novel meaning. For example, it is often suggested that the meaning of "Who will sit?" is something like the set of (contextually relevant or possible) answers to it. But this makes the meaning of a question just a set of statements.

J. L. Austin once claimed that a good deal of natural-language vocabulary has meanings whose job is to signal that one is, and is only, performing a (nonassertive) speech act. For example, on Austin's view, to utter "I promise to meet you at 5:00" is not to assert anything, but to make a promise. Austin (1962) gives a lengthy catalog of verbs (part of) whose conventional meaning, he claims, is to signal (when used in the first-person present) that a particular speech act is being performed, representative examples being "acquit," "nominate," "bet," "toast," and "concede." He suggests that the number of such verbs contained in English is "of the third order of the power of 10."

There are arguably many expressions whose purpose is in part or in whole to signal that, whatever else the speaker might be doing, he is performing a particular nonassertive speech act, though exactly which expressions do this is a matter of controversy. "Just between you and me" (as in "Just between you and me, the provost hates the president") might be a conventional means to warn or ask one's audience not to divulge the information imparted by the rest of the sentence. Racial slurs are, inter alia, conventional means of insulting and displaying contempt for their targets, as are the merely obscene or insulting things we may call someone in the course of commenting on them. Presumably, though, to utter something like "That jerk Smith is at the door" is to say something true or false, depending (only) on whether Smith is at the door.

CONVENTIONAL IMPLICATURE

Grice (1967/1989) drew a distinction between what the use of a sentence "strictly" says and what it implies. Both what is said and what is implicated are statements. Indeed, what a sentence use says, in Grice's sense, seems to be the statement that a semantic theory (of the second sort discussed above) aims to assign to the use. According to Grice, it is what a sentence strictly says, and *only* what it strictly says, that is relevant to the question of whether the use of the sentence is true.

What, then, is the role of what is implicated by the use of a sentence? Some such implication is a one-off affair, as when one says, "There's an umbrella in the closet," expecting one's auditor to work out that rain is in the offing. Implication of this sort exploits facts obvious to all—for example, that speakers generally try to say helpful and relevant things—to efficiently convey information; it allows us to convey much more than our words literally mean.

Grice distinguished this sort of implication—*conversational implicature*, as he called it—from cases in which "the conventional meaning of the words used … determine[s] what is implicated, besides helping to determine what is said" (Grice 1967/1989, p. 25). Grice's examples were the words "therefore" and "but." In uttering "*A*; therefore *B*," Grice claimed, I say that *A*, say that *B*, commit myself to *B*'s following from *A*, but I have not "*said* (in the favored sense)" that *B* follows from *A*: "I do not want to say that my utterance … would be, *strictly speaking*, false should the consequence in question [fail] to hold" (Grice 1967/1989, pp. 25–26). In uttering "He is *F* but *G*," one speaks truly, Grice said, just in case the relevant individual is *F* and *G*, though one clearly conveys some sort of contrast between being *F* and being *G*. To use "therefore" or "but" is to commit to these implications. Since the implications are carried by the very words used, they are not one-off conversational implicatures but *conventional implicatures*.

A rather large class of expressions have been said to give rise to conventional implicatures. Karttunen and Peters (1979) suggest that words and constructions often said to give rise to presuppositions in fact give rise to conventional implicature. Here are some examples, with the word purportedly carrying the conventional implicature italicized and the implicature roughly indicated in parentheses:

Even John understands it. (John is unlikely to understand it.)

Martin *still* loves her. (Martin loved her in the past.)

Jed *failed* to pass. (Jed tried to pass.)

Other examples of purported conventional implicatures are nonrestrictive relative clauses and appositives. "Martina, a yogi, hunts bears" commits the speaker to Martina's being a yogi, but arguably would be true even if she is not one, so long as she does hunt bears.

It is controversial whether there is such a thing as conventional implicature. Bach (1999) argued that a complete report of Bob's utterance of "Even Mo likes Jo" is given with "Bob said that even Mo likes Jo"; simply saying, "Bob said that Mo likes Jo" is not giving a complete report. Since "that even Mo likes Jo" is here specifying what Bob said, Bach concluded, part of what Bob's utterance says must be (something like the claim) that Mo's liking Jo is unexpected. But if that is part of what is said, then the utterance is true only if it is unexpected that Mo likes Jo. According to Bach, this sort of argument shows that pretty much every expression alleged to carry a conventional implicature in fact does not.

It is not clear that this argument succeeds in showing that conventional implicatures are a fiction. "What is said (by utterance *u*)," as used by Grice, is a technical term. The phrase and its cousins have an everyday use as well. It is not at all clear that Grice assumed that if an utterance would naturally and correctly be reported as saying that *p*, then *p* must be part of what it says *in the technical sense*. We are, after all, pretty loose in how we report indirect speech.

One might hold that conventional implicatures are just as much *said by* a use of a sentence as anything, but have properties and relations to sentence uses that make it worthwhile to distinguish them from other claims literally made by sentence uses. Christopher Potts (2005) distinguished what he called "at issue" claims made by a sentence use (roughly, what Grice had in mind by "what is said") from conventional implicatures. (However, Potts's view, unlike Grice's, is apparently that conventional implicatures are relevant to truth conditions. He takes conventional implicatures to be "entailments," and holds that sentences carrying such implicatures can typically be paraphrased by conjunctions, one conjunct of which is the implicature.)

For Potts, one putative difference between conventional implicature and at-issue content is that even when a speaker embeds an expression carrying a conventional implicature, the speaker becomes committed to the implicature; this is not so with at-issue content. To see the point, consider "Bob, a linguist, likes clams," where the at-issue content is that Bob likes clams and the conventional

implicature is that Bob is a linguist. When one embeds the sentence under negation or an attitude verb (as in "It is false that Bob, a linguist, likes clams," or "Mary said that Bob, a linguist, likes clams"), use of the resulting sentence seems to commit the user to the conventional implicature, but not to the at-issue content.

A conventional implication is like a presupposition in this regard. Potts argues that conventional implicatures are not presuppositions, since false conventional implicatures and false presuppositions have different effects. When a presupposition of a sentence is false, the assertion of its at-issue content is unfelicitous, perhaps without truth value; this is not so with conventional implicature. In the case of conventional implicature, that Bob is not a linguist does not impugn or cast doubt on the claim that Bob likes clams. Knowledge that the presupposition of "It was Bob who stole the book (namely, that someone stole it) is false makes the assertion that it was Bob who did it unacceptable.

NONPROPOSITIONAL MEANING

What is conventionally implicated has truth conditions. A non-truth-conditional conventional implicature does not enter into the truth conditions of the use of a sentence; its truth or falsity is not relevant to the truth or falsity of the sentence use implicating it. Other alleged sorts of non-truth-conditional meanings, however, are non-truth-conditional in the sense that they simply are not the sort of thing that can be true or false—they are, as it is sometimes said, not truth-apt.

One (alleged) example of such a meaning is presented by those who hold that linguistic meaning, or an aspect thereof, is to be identified with one or another psychological role associated with an expression. It has been proposed that the meaning of a sentence as used by a particular speaker is or involves one or more of: its inferential role (reflected by the speaker's dispositions to make inferences from and to the sentence), its evidential role (reflected in what observations and experiences incline the speaker to accept or reject the sentence), and its probabilistic role (the function that sends a sentence *S* and a collection *C* of sentences to the subjective probability the speaker would assign *S* if he held all of *C* true). (Developments of such views are in Boer and Lycan 1986, Field 1977, Sellars 1954.) None of these things can sensibly be evaluated for truth or falsity. Those who champion such psychological accounts of meaning often hold that meaning is a two-factor affair, the other factor being truth-conditional. Typically, though not invariably, the two factors are held to be independent.

In part, the appeal of adding psychological role to truth conditions in an account of meaning is that it seems to reflect a genuine tension in our pretheoretic conception of meaning. Consider Putnam's fantasy (in 1975) that there is a Twin Earth as much like Earth as possible, save that something other than H_2O, call it XYZ, plays the role that H_2O plays on Earth: XYZ has all the sensible properties of H_2O; it is XYZ, not H_2O, that fills the seas, that people drink and wash with, etc. Putnam holds, and many concur, that "water" means different things on Earth and on Twin Earth, for here it refers to H_2O, while there it refers to XYZ. But many think that in some very important sense the word has the same meaning in both places, for someone transported to Twin Earth who was innocent of chemistry, it is felt, would not mean anything different by "water" there than he means here. If there are two factors to the meaning of "water"—a truth-conditional one (which varies between Earth and Twin Earth) and a psychological one (which is constant), both intuitions are partially vindicated.

A different kind of nonpropositional meaning is what is sometimes called "expressive" meaning. The idea of such meaning has its roots in the work of emotivists like A. J. Ayer and Charles Stevenson. According to Ayer, the role of ethical discourse is completely noncognitive. Utterances of sentences such as "Stealing is wrong" and "Friendship is good" are not assertions and do not express beliefs. Rather, they are expressions of attitudes of approval or disapproval. Uttering "Stealing is wrong" is doing the sort of thing one does when one shouts "Down with stealing!" or accompanies utterance of the word "stealing" with a disapproving shake of the head. Stevenson's somewhat more sophisticated take on such sentences is that uttering them both expresses a distinctive sort of approval and exhorts (or at least attempts to bring) the audience to share this approval.

Sentences whose role is clearly exhausted by the expression of attitude—"Boo!" "Liver—yuck!" "Damn!"—are not candidates for combining with connectives and quantifiers to form larger sentences. "If liver—yuck, then I won't make dinner" does not have a meaning, for it is not even a sentence. But sentences such as "Stealing is bad" quite obviously do combine with connectives and other sentences, and the results certainly do seem to be meaningful. It seems incumbent on any account of semantics to explain what the meaning of a sentence such as "Stealing is bad only if it causes pain to someone."

Geach (1965), expanding on points in Frege (1918/1952), objects that the emotivist cannot make any

sense of the use of normative vocabulary in complex sentences, of embedded uses, as is sometimes said. Someone who utters "If failing Mary will make her sad, you shouldn't do it" need not be expressing disapproval of anything. Even if there is a way around this—one might invoke some sort of "conditional disapproval"—emotivist views make the fact that we give normative arguments an utter mystery. The argument "Borrowing and not returning something is bad; if that is bad, so is stealing; so stealing is bad" is valid—its conclusion follows from its premises. But it seems to be nonsense to think that a feeling of disapproval for stealing *follows* from a feeling of disapproval for borrowing and not returning and whatever attitude might be associated with the conditional above. "Following from," after all, is a relation normally defined in terms of preserving truth. But if this makes no sense, the idea that the argument is valid makes no sense in emotivist terms.

These considerations, incidentally, bear on the view of Austin mentioned above. The argument "If I promise to meet you, I will meet you; I promise to meet you; so I will meet you" seems obviously valid. But there is a sort of ambiguity, on Austin's view, in "I promise to meet you." Embedded in the antecedent of a conditional, it presumably does nothing but express the *statement* that its user promises to meet the addressee. Unembedded, it apparently does not do this, as one, in uttering the sentence, does not *assert* that one promises, on Austin's view; one simply promises. It thus seems like the sense of "I promise to meet you" varies across the two premises of the argument, and thus the argument is not valid.

Expressivists such as Simon Blackburn and Alan Gibbard have recently tried to respond to this sort of objection, giving accounts that (more or less) agree with the emotivist line about simple sentences like "Hooking up is good" and attempting to derive therefrom meanings for complex sentences in which normative vocabulary occurs. Blackburn (1993) agrees with the emotivist that sentences like "Stealing is bad" express motivational states such as attitudes of disapproval. But he aspires to give an account of the meanings of the full range of uses of normative vocabulary, including such sentences as "Mary believes that stealing is bad" and "It's true that stealing is bad." The account is to be one that systematically assigns, to complex sentences, complex attitudes—typically in one or another way compounded out of the attitudes expressed by simple sentences. The sentence "If borrowing and not returning something is bad, then so is stealing it," for example, expresses a commitment to either tolerating borrowing and not returning, or to disapproving stealing. Such a view would allow us to characterize validity in terms of preservation of commitment—an argument is valid just in case it is impossible to fulfill the commitments associated with premises without fulfilling those associated with the conclusion.

Gibbard (1992, a recast of 1990) suggested that normative sentences—not just sentences from morality, but sentences about what is or is not rational—absorb their meanings holistically from their relations to "immediate motivations," that is, to the states one expresses if one thinks to oneself "Do/Don't do that now!" The idea, roughly put, is that just as complex statements get their truth-conditional content from their inferential relations to sentences expressing observations, so normative statements absorb their content from inferential relations to sentences expressing immediate motivations. Gibbard suggests that the meaning of a normative sentence (including complex combinations of normative and nonnormative elements) can be represented as a set of "factual-normative" worlds, which are pairs of possible worlds and systems of norms. The idea, again roughly, is this. A simple factual statement holds at world *w* and norm *n* if it is true there. A simple normative statement such as "That is bad" (whose connection with "Don't do that!" is obvious) holds at *w* and *n* provided that *n* forbids the act referred to. With this as a basis, one can use standard techniques to assign sets of factual-normative worlds to compound sentences.

One might argue with Blackburn and Gibbard about the details of their approaches, worrying, for example, that Blackburn helps himself without justification to the idea that there is a distinctive sort of moral disapproval. Yet it would seem that something along the lines of Blackburn's or Gibbard's story *must* be correct. Here is why.

Forget about claims about morality, rationality, or other obviously normative concerns. Think instead about what is going on when we talk about talk that obviously aspires to be true or false—about what happens when one person says "Jo is bald" and another says "That's not true," or when someone says "The sentence on the board isn't true." It seems obvious that such talk can get it right without being true. If the sentence on the board is a liar sentence, one thing that we *know* about it is that it is not true. We can, after all, *prove* that it is not. But paradox ensues if we take this thing we know—that the sentence is not true—to be true. After all, if what we know—that the sentence is not true—is true, then, since the sentence *says* just that—that it is not true—what the sentence says is true. So what we know is false. But one cannot know something that is false. Similarly, if vague predicates are

neither true nor false of their borderline cases (and surely this is the most plausible thing to say about them) and Jo is borderline-bald, then while the person who utters "Jo is bald" says something, what he says is *not* true. But if it is *true* that the sentence is not true, then (since what is not true is false), "Jo is bald" must be false. But since Jo is borderline bald, "Jo is bald" cannot be false either.

What should we make of this? Well, for one thing, when we say, referring to the liar, "That is not true," we should not be understood as *asserting* something, that is, committing to its truth. Rather, we are performing the sui generis speech act of *denial*, where (roughly put) denying a potential truth bearer is the appropriate thing to do if it is *not* true ("not" being used here to deny). This sort of thing applies quite generally to uses of other logical connectives. Sometimes, for example, when someone utters "*A* if and only if (iff) *B*," they are to be understood as asserting the material equivalence of *A* and *B*. But when we say things with the form " '*S*' is true iff *S*" and *S* happens to be a liar sentence, we are not to be understood as *asserting* anything. Rather, we are performing an act that is apt if the claims connected by "iff" have the same (perhaps non-truth-conditional) status.

When we utter sentences, we perform different sorts of speech acts. Sometimes we assert, sometimes we deny, sometimes we perform the sort of act just mentioned. And when we perform such acts, we incur various commitments. For example, assertion commits us to the truth of what is asserted; denial of a potential truth bearer commits us to the nontruth thereof. Sentence-compounding devices, at least on some occasions, contribute not to *sense*, by (for example) expressing truth-functional negation, but to *force*. In the case of "not," for example, one sometimes signals that one is denying, where to deny *S* is to commit to the inaptness of whatever commitment is associated with uttering *S*.

Think of the simplest sentences of one's language as vehicles for performing speech acts, each such act involving its own distinctive kind of commitment, each commitment having its own conditions of appropriateness and inappropriateness. Annexing words like "not" and "if" to sentences yields (when the connectives signal force) sentences that are vehicles for performing speech acts with their own distinctive kinds of commitments, their own aptness conditions. Compounding sentences with several connectives playing the role of force indicators produces a sentence that can serve as a vehicle for performing a complex speech act determined by the meanings of the constituent sentences and the force-indicating meanings of the connectives. Uttering "If *S* is a liar sentence, then it is not true," for example, performs an apt speech act if it is apt either to deny that *S* is a liar sentence or to deny that *S* is true.

Beyond an account of sense or reference, a theory of meaning for a language—at least one component of such a theory—must tell the story of how the acts and commitments associated with the parts of a complex sentence determine the act for which the complex sentence is a vehicle, the commitments one incurs with the act, and the aptness conditions of such commitments. (For the beginning of such a story, see Richard 2006.) Such a story generalizes the sort of ideas Blackburn had. With such a story, one can see that logical validity, in its most basic sense, is preservation of commitment: An argument is valid provided that whenever the commitments associated with the premises are apt, so are those associated with the conclusion.

It was mentioned above that there was something importantly right about Gibbard's and Blackburn's accounts of normative discourse. What is important and surely right is not their view of the nature of the acts performed and commitments incurred in normative utterances. Perhaps those accounts are on the right track, because normative discourse is expressive, not truth-apt. Perhaps they are wrong, and normative discourse is no less truth-evaluable than a stock-price quotation. What is important is the insight that validity (and the other properties we associate with rational discourse) are not the exclusive property of truth-conditional discourse. Sometimes meaning and validity are to be explained in terms of truth conditions. But this is not the only case—it is but a special case.

See also Meaning.

Bibliography

Austin, J. L. *How to Do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press, 1962.

Ayer, A. J. *Language, Truth, and Logic*. London: Victor Gollancz, 1936.

Bach, Kent. "The Myth of Conventional Implicature." *Linguistics and Philosophy* 22 (4) (1999): 367–421.

Blackburn, Simon. *Essays in Quasi-Realism*. New York: Oxford University Press, 1993.

Boer, Stephen, and William Lycan. *Knowing Who*. Cambridge, MA: MIT Press, 1986.

Field, Hartry. "Logic, Meaning, and Conceptual Role." *Journal of Philosophy* 74 (1977): 379–409.

Frege, Gottlob. "Negation" (1918). In *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black. Oxford, U.K.: Blackwell, 1952.

Geach, Peter. "Assertion." *Philosophical Review* 74 (1965): 449–465.

Gibbard, Alan. "Reply to Blackburn, Carson, Hill, and Railton." *Philosophy and Phenomenological Research* 52 (4) (1992): 969–980.

Gibbard, Alan. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press, 1990.

Grice, H. P. "Logic and Conversation" (1967). In his *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press, 1989.

Karttunen, Lauri, and Stanley Peters. "Conventional Implicature." In *Presupposition*, edited by Choon-Kyu Oh and David Dinneen. New York: Academic Press, 1979.

Potts, Christopher. *The Logic of Conventional Implicatures*. Oxford, U.K.: Oxford University Press, 2005.

Putnam, Hilary. "The Meaning of 'Meaning.'" In his *Mind, Language, and Reality*. New York: Cambridge University Press, 1975.

Richard, Mark. *Beside Truth*. Oxford: Oxford University Press, 2006.

Sellars, Wilfred. "Some Remarks on Language Games." *Philosophy of Science* 21 (1954): 204–228.

Stevenson, Charles. *Ethics and Language*. New Haven, CT: Yale University Press, 1944.

Mark Richard (2005)

NONVIOLENCE

See *Violence*

NORMATIVE ETHICS

See *Ethics, History of; Normativity*

NORRIS, JOHN

(1657–1711)

John Norris, the English philosopher and disciple of Nicolas Malebranche, was associated with the Cambridge Platonists. Norris was born in Collingbourne-Kingston, Wiltshire. His father was a clergyman and at that time a Puritan. Educated at Winchester and at Exeter College, Oxford, which he entered in 1676, Norris was appointed a fellow of All Souls in 1680. During his nine years at All Souls, he was ordained (1684) and began to write, mostly in a Platonic vein and often in verse. In 1683 he published *Tractatus adversus Reprobationis absolutae Decretum*, in which he attacked the Calvinist doctrine of predestination. His Platonism and anti-Calvinism naturally attracted Norris to the Cambridge Platonists; in 1684 he began to correspond with Henry More and Damaris Cudworth, the daughter of Ralph Cudworth.

The philosophical essays included in *Poems and Discourses* (1684)—renamed *A Collection of Miscellanies* in the 1687 and subsequent editions—could, indeed, have been written by a Cambridge Platonist. Their main argument is that since truth is by its nature eternal and immutable, it must relate ideas which are also eternal and immutable; this condition, according to Norris, can be fulfilled only by ideas which are "in the mind of God"—that is, manifestations of God's essence. Thus, the existence of God is deducible from the very nature of truth; the atheist is involved in a self-contradictory skepticism.

In Norris's *The Theory and Regulation of Love* (1688)—for all that Norris dedicated it to the former Damaris Cudworth, now Lady Masham, and included as an appendix his correspondence with More—the influence of Malebranche began to predominate. At first, it reinforced rather than weakened Norris's sympathy with Cambridge Platonism. Norris followed Malebranche in distinguishing two kinds of love—desire, which seeks to unify itself with the good it pursues, and benevolence, which seeks good for others. But, as also in *Reason and Religion* (1689), Norris explicitly rejected Malebranche's view that the only proper object of desire is God. The objects of desire, Norris said, form a hierarchy—God, the good of the community, intellectual pleasures, and sensual pleasures are all in some measure good. God is the highest but not the only good.

In 1689, Norris married and resigned his fellowship to become rector of Newton St. Loe in Somerset. In his *Reflections on the Conduct of Human Life* (1690), addressed to Lady Masham and intended as an admonition to her, he condemned the life he had lived at Oxford on the ground that he had interested himself in public affairs and in intellectual pursuits; in the future he proposed to dedicate himself in retirement to the "moral improvement of my mind and the regulation of my life." This is Malebranche's, not the Cambridge Platonists', ideal of conduct; even the pursuit of knowledge is conceived of as a worldly enticement.

In 1691, as a result of John Locke's influence, Norris became rector of Bemerton, near Salisbury, where he died on February 5, 1711. He did not win the approval of his Cambridge Platonist bishop, Gilbert Burnet, who would certainly not have appreciated Norris's attack on toleration in *The Charge of Schism continued* (1691). Norris's *Discourse concerning the Measures of Divine Love* (*Practical Discourses*, Vol. III, 1693) and *Letters concerning the Love of God* (1695) reveal the complete disciple of Malebranche; we ought, Norris now said, to love nobody but God. Substantially reversing Immanuel Kant's dictum, he

argued that we should treat other human beings as means—occasions of happiness to us—and never as ends. Lady Masham was naturally indignant; in her anonymous *Discourse concerning the Love of God* (1696), a reply to Norris, she argued that men are "made for a sociable life" and should love their fellow men in the same way they love God.

THOUGHT

Norris's metaphysical views, sketched in *Reason and Religion*, are set out in detail in his *Essay towards the Theory of the Ideal and the Sensible World* (Vol. I, 1701; Vol. II, 1704), which fully justifies his nickname "the English Malebranche." Yet the argument of the first volume of the *Essay* would still entitle Norris to be described as a Platonist—or as a Thomist or an Augustinian. Plato, the "Platonic father" Augustine, Francisco Suárez, and Thomas Aquinas all taught, he tried to show, the same lesson as Malebranche—that knowledge is of the eternal and, therefore, of God.

In the second volume, however, when Norris came to consider in more detail how our knowledge of "the world of sense" is related to our knowledge of "the intelligible world," his break with the Platonist tradition, arising out of his allegiance to Malebranche, is at once apparent. It is true that when he did (mildly) criticize Malebranche, it is on the Platonic ground that his theory of the imagination allows too much to sensation; Malebranche's phrase "We see all things in God," he also thought, might suggest to the careless reader that sensation is our analogue for knowledge. "Divine ideas," Norris preferred to say, "are the immediate objects of our thought in the perception of things." But these are minor reformulations. Of much greater significance is the fact that he agreed with the Cartesians that "the world is a great mechanism and goes like a clock" and even accepted, although with some little hesitation, the Cartesian doctrine of animal mechanism. He did not even bother to refer to the Platonist theory of "plastic powers" or to More's criticism of René Descartes's extension-thought dualism. He is a Platonist only where Malebranche is a Platonist—for example, in his rejection of the Thomas-Locke account of abstraction.

Norris's philosophy might properly be described, in the phrase commonly applied to Benedict de Spinoza, as "God-intoxicated." God, for him as for Malebranche, is the efficient cause of all happenings, the only good, the only object of knowledge. We know God directly; everything else is known by way of our apprehension of God's nature as revealed in the ideas that emanate from him. Norris could not explain, he confessed, how spiritual ideas can represent a material world; the material world is, indeed, an embarrassment to him, fading into the empty concept of "that which occasions our apprehensions" that George Berkeley criticized. He was so concerned to leave nothing lovable in the world, nothing that could be a source of happiness to us, that he reduced it to a nonentity; it exists only as something to be shunned. The relation between our mind and God's is left in equal obscurity.

In 1692 Locke and Norris quarreled on a matter involving Lady Masham; Locke came to be very impatient with Norris's views, which probably provoked his *Examination of Malebranche* (first published in *Posthumous Works*, edited by Peter King, London, 1706); he directly criticized Norris in an essay first published in *A Collection of Several Pieces of Mr. John Locke* (1720). In general, Locke thought of Norris as a completely reactionary thinker.

Other of Norris's works deserving mention are *An Account of Reason and faith in relation to the Mysteries of Christianity* (1697), in which he argued—in reply to John Toland's deistic *Christianity not Mysterious* (1696)—that it is not unreasonable to believe the incomprehensible, and *A Philosophical Discourse concerning the Natural Immortality of the Soul* (1708), which makes use of Platonic-scholastic arguments against Henry Dodwell's *Epistolary Discourse proving … that the Soul is naturally Mortal* (1706). Many of his works, although not *The Ideal World*, were extremely popular, but it is usually impossible to distinguish his influence from Malebranche's. One of the least original of philosophers, he nevertheless displays considerable powers of criticism and exposition. He had a direct influence on Arthur Collier.

See also Cambridge Platonists.

Bibliography

There is no modern edition of Norris. Norris brought together several of his minor works, including *Reason and Religion*, as *Treatises upon Several Subjects* (London, 1697). For his *Poems* see the edition by Alexander Balloch Grosart in the *Fuller Worthies' Library*, Vol. III (Blackburn, U.K., 1871), pp. 147–348; John Wesley included an abbreviated version of *Treatise on Christian Prudence* and *Reflections upon the Conduct of Human Life* in his *Christian Library*, Vol. XXX (London, 1827).

See also Frederick James Powicke, *A Dissertation on John Norris* (London, 1894); Ernest Trafford Campagnac, *The Cambridge Platonists* (Oxford: Clarendon Press, 1901); Flora Isabel MacKinnon, *The Philosophy of John Norris of Bemerton, Philosophical Monographs*, No. 2 of the *Psychological Review* (October 1910); John Henry Muirhead, *The Platonic Tradition in Anglo-Saxon Philosophy* (London,

1931); John K. Ryan, "John Norris, a Seventeenth Century Thomist," in *New Scholasticism* 14 (2) (1940): 109–145; Charlotte Johnston, "Locke's *Examination of Malebranche* and John Norris," in *Journal of the History of Ideas* 19 (4) (1958): 551–558; Richard Acworth, *The Philosophy of John Norris of Bemerton (1657–1712)* (Hildesheim, NY: Olms, 1979); Richard Acworth, "Locke's First Reply to John Norris," *Locke Newsletter* 2 (1971): 7–11; and Charles J. McCracken, *Malebranche and British Philosophy* (Oxford: Clarendon Press, 1983).

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

NOTHING

"Nothing" is an awe-inspiring yet essentially undigested concept, highly esteemed by writers of a mystical or existentialist tendency, but by most others regarded with anxiety, nausea, or panic. Nobody seems to know how to deal with it (he would, of course), and plain persons generally are reported to have little difficulty in saying, seeing, hearing, and doing nothing. Philosophers, however, have never felt easy on the matter. Ever since Parmenides laid it down that it is impossible to speak of what is not, broke his own rule in the act of stating it, and deduced himself into a world where all that ever happened was nothing, the impression has persisted that the narrow path between sense and nonsense on this subject is a difficult one to tread and that altogether the less said of it the better.

This escape, however, is not so easy as it looks. Plato, in pursuing it, reversed the Parmenidean dictum by insisting, in effect, that anything a philosopher *can* find to talk about must somehow be there to be discussed, and so let loose upon the world that unseemly rabble of centaurs and unicorns, carnivorous cows, republican monarchs and wife-burdened bachelors, which has plagued ontology from that day to this. Nothing (of which they are all aliases) can apparently get rid of these absurdities, but for fairly obvious reasons has not been invited to do so. Logic has attempted the task, but with sadly limited success. Of some, though not all, nonentities, even a logician knows that they do not exist, since their properties defy the law of contradiction; the remainder, however, are not so readily dismissed. Whatever Bertrand Russell may have said of it, the harmless if unnecessary unicorn cannot be driven out of logic as it can out of zoology, unless by desperate measures that exclude all manner of reputable entities as well. Such remedies have been attempted, and their effects are worse than the disease. Russell himself, in eliminating the present king of France, inadvertently deposed the present queen of England. W. V. Quine, the sorcerer's apprentice, contrived to liquidate both Pegasus and President Harry Truman in the same fell swoop. The old logicians, who allowed all entities subsistence while conceding existence, as wanted, to an accredited selection of them, at least brought a certain tolerant inefficiency to their task. Of the new it can only be said that *solitudinem faciunt et pacem appellant*—they make a desert and call it peace. Whole realms of being have been abolished without warning, at the mere nonquantifying of a variable. The poetry of Earth has been parsed out of existence—and what has become of its prose? There is little need for an answer. Writers to whom nothing is sacred, and who accordingly stop thereat, have no occasion for surprise on finding, at the end of their operations, that nothing is all they have left.

The logicians, of course, will have nothing of all this. Nothing, they say, is not a thing, nor is it the name of anything, being merely a short way of saying of anything that it is not something else. *Nothing* means "not-anything"; appearances to the contrary are due merely to the error of supposing that a grammatical subject must necessarily be a name. Asked, however, to prove that nothing is *not* the name of anything, they fall back on the claim that nothing *is* the name of anything (since according to them there are no names anyway). Those who can make nothing of such an argument are welcome to the attempt. When logic falls out with itself, honest men come into their own, and it will take more than this to persuade them that there are not better cures for this particular headache than the old and now discredited method of cutting off the patient's head.

The friends of nothing may be divided into two distinct though not exclusive classes: the know-nothings, who claim a phenomenological acquaintance with nothing in particular, and the fear-nothings, who, believing, with Macbeth, that "nothing is but what is not," are thereby launched into dialectical encounter with nullity in general. For the first, nothing, so far from being a mere grammatical illusion, is a genuine, even positive, feature of experience. We are all familiar with, and have a vocabulary for, holes and gaps, lacks and losses, absences, silences, impalpabilities, insipidities, and the like. Voids and vacancies of one sort or another are sought after, dealt in and advertised in the newspapers. And what are these, it is asked, but perceived fragments of nothingness, experiential blanks, which command, nonetheless, their share of attention and therefore deserve recognition?

Jean-Paul Sartre, for one, has given currency to such arguments, and so, in effect, have the upholders of "nega-

tive facts"—an improvident sect, whose refrigerators are full of nonexistent butter and cheese, absentee elephants and so on, which they claim to detect therein. If existence indeed precedes essence, there is certainly reason of a sort for maintaining that nonexistence is also anterior to, and not a mere product of, the essentially parasitic activity of negation; that the nothing precedes the not. But, verbal refutations apart, the short answer to this view, as given, for instance, by Henri Bergson, is that these are but petty and partial nothings, themselves parasitic on what already exists. Absence is a mere privation, and a privation of something at that. A hole is always a hole *in* something: take away the thing, and the hole goes too; more precisely, it is replaced by a bigger if not better hole, itself relative to its surroundings, and so tributary to something else. Nothing, in short, is given only in relation to what is, and even the idea of nothing requires a thinker to sustain it. If we want to encounter it *an sich,* we have to try harder than that.

Better things, or rather nothings, are promised on the alternative theory, whereby it is argued, so to speak, not that holes are in things but that things are in holes or, more generally, that *everything* (and everybody) is in a hole. To be anything (or anybody) is to be bounded, hemmed in, defined, and separated by a circumambient frame of vacuity, and what is true of the individual is equally true of the collective. The universe at large is fringed with nothingness, from which indeed (how else?) it must have been created, if created it was; and its beginning and end, like that of all change within it, must similarly be viewed as a passage from one nothing to another, with an interlude of being in between. Such thoughts, or others like them, have haunted the speculations of nullophile metaphysicians from Pythagoras to Blaise Pascal and from G. W. F. Hegel and his followers to Martin Heidegger, Paul Tillich and Sartre. Being and nonbeing, as they see it, are complementary notions, dialectically entwined, and of equal status and importance; although Heidegger alone has extended their symmetry to the point of equipping *Das Nichts* with a correlative (if nugatory) activity of nothing, or nihilating, whereby it produces *Angst* in its votaries and untimely hilarity in those, such as Rudolf Carnap and A. J. Ayer, who have difficulty in parsing *nothing* as a present participle of the verb "to noth."

Nothing, whether it noths or not, and whether or not the being of anything entails it, clearly does not entail that anything should be. Like Benedict de Spinoza's substance, it is *causa sui*; nothing (except more of the same) can come of it; *ex nihilo, nihil fit.* That conceded, it remains a question to some why anything, rather than nothing, should exist. This is either the deepest conundrum in metaphysics or the most childish, and though many must have felt the force of it at one time or another, it is equally common to conclude, on reflection, that it is no question at all. The hypothesis of theism may be said to take it seriously and to offer a provisional answer. The alternative is to argue that the dilemma is self-resolved in the mere possibility of stating it. If nothing whatsoever existed, there would be no problem and no answer, and the anxieties even of existential philosophers would be permanently laid to rest. Since they are not, there is evidently *nothing to worry about.* But that itself should be enough to keep an existentialist happy. Unless the solution be, as some have suspected, that it is not nothing that has been worrying them, but they who have been worrying it.

See also Atheism; Ayer, Alfred Jules; Bergson, Henri; Carnap, Rudolf; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Logic, History of; Nihilism; Parmenides of Elea; Plato; Quine, Willard Van Orman; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Spinoza, Benedict (Baruch) de; Tillich, Paul.

Bibliography

Modern writers who have had something to say about nothing include:

Barrett, William. *Irrational Man.* Garden City, NY: Doubleday, 1958.

Bergson, Henri. *L'évolution créatrice.* Paris: F. Alcan, 1907. Translated by Arthur Mitchell as *Creative Evolution.* London: Macmillan, 1911.

Carnap, Rudolf. "The Elimination of Metaphysics." In *Logical Positivism,* edited by A. J. Ayer, 69–73. Glencoe, IL: Free Press, 1959.

Edwards, Paul. "Professor Tillich's Confusions." *Mind,* n.s., 74 (1965): 192–214.

Findlay, J. N. *Meinong's Theory of Objects and Values.* 2nd ed. Oxford: Clarendon Press, 1963.

Heidegger, Martin. *Einführung in die Metaphysik.* Tübingen: Niemeyer, 1953. Translated by Ralph Manheim as *An Introduction to Metaphysics.* New Haven, CT: Yale University Press, 1959.

Heidegger, Martin. *Sein und Zeit.* Halle: Niemeyer, 1927. Translated by John Macquarrie and Edward Robinson as *Being and Time.* New York: Harper, 1962.

Heidegger, Martin. *Was ist Metaphysik?* Bonn: Cohen, 1929; 4th ed., Frankfurt, 1943. Translated by R. F. C. Hull and Alan Crick as "What Is Metaphysics?," in *Existence and Being,* edited by W. Brock. Chicago: Regnery, 1949.

Lazerowitz, Morris. *Structure of Metaphysics.* London: Routledge and Paul, 1955.

Munitz, M. K. *Mystery of Existence.* New York: Appleton-Century-Crofts, 1965.

Prior, A. N. "Non-entities." In *Analytical Philosophy I,* edited by R. J. Butler. New York: Barnes and Noble, 1962.
Quine, W. V. *From a Logical Point of View.* Cambridge, MA: Harvard University Press, 1953.
Russell, Bertrand. "On Denoting." *Mind,* n.s., 14 (1905): 479–493.
Sartre, Jean-Paul. *L'être et le néant.* Paris: Gallimard, 1943. Translated by Hazel E. Barnes as *Being and Nothingness.* London: Methuen, 1957.
Taylor, Richard. "Negative Things." *Journal of Philosophy* 49 (13) (1952): 433–448.
Tillich, Paul. *The Courage to Be.* New Haven, CT: Yale University Press, 1952.
Toms, Eric. *Being, Negation and Logic.* Oxford: Blackwell, 1962.

P. L. Heath (1967)

NOUNS, MASS AND COUNT

Many languages mark a grammatical distinction that is commonly referred to as the "mass/count-distinction"; for example, the distinction between the occurrences of "hair" as a mass-noun in "There is hair in my soup," on the one hand, and its occurrences as a singular and plural count-noun in "There is a hair in my soup" or "There are hairs in my soup," on the other. Awareness of this linguistic contrast may, in the Western tradition, date as far back as the pre-Socratics, Plato, and Aristotle; in modern times, however, the first explicit formulation of it is usually credited to Otto Jespersen (1924).

1. THE PROBLEM OF CLASSIFICATION

Almost every aspect of the mass/count-distinction is unclear and contested, including the question of how it is to be drawn:

> **The Problem of Classification:**
>
> (i) Between what sorts of entities is the mass/count-distinction to be drawn?
>
> (ii) By means of what sorts of criteria is the mass/count-distinction to be drawn?

What underlies question (i), for one thing, is a certain ambivalence as to whether the contrast concerns uses or occurrences of expressions or expressions themselves (and, if the former, we face the further question as to what a "use" or an "occurrence" of an expression really is; that is, how, for example, occurrences contrast with types and tokens of expressions). (In what follows, for reasons of convenience, we will speak of both uses or occurrences as well as of expressions themselves as being mass or count.) Moreover, question (i) also encompasses the issue of whether the contrast in question can be properly drawn only with respect to nouns and noun-phrases or whether it can be sensibly extended to other categories, such as adjectives (e.g., with "red" on the mass-side and "circular" on the count-side) as well as verbs and verb-phrases (e.g., with atelic activity-verbs such as "run in circles" being classified as mass and telic achievement- or accomplishment-verbs such as "recognize" or "grow up" being classified as count; see Hoepelman 1976, Taylor 1977, Mourelatos 1978).

Question (ii), on the other hand, asks whether the distinction in question is best drawn, for example, by means of syntactic, morphological, semantic, or pragmatic criteria. To illustrate—restricting ourselves, as is customary, to the category of nouns and noun-phrases, and to such purely syntactic criteria (exhibited overtly in English) as the admissibility of plural-morphology as well as the licensing of "bare" (i.e., unquantified) occurrences or particular kinds of determiners and quantifiers (e.g., "much" versus "many")—we arrive at the following sort of classification:

(M) *Mass*: "air," "water," "mud," "sand," "dust," "snow," "gravel," "asparagus," "traffic," …
(C) *Count*: "beach," "cloud," "chair," "piece of furniture," "virus," "bacteria," "sheep," "university," "hurricane," "football game," …
(D) *Dual-Use*: "hair," "chicken," "carrot," "apple," "cloth," "pain," "disease," …

The nouns in the first list permit "bare" occurrences (as in "Water is wet"); they do not, in their use as mass-nouns, permit pluralization; and they can occur together with such quantifiers as "much" or "little" (as in "much air" and "little air"). The nouns in the second list do not permit (singular) "bare" occurrences (as in "*Beach is sandy"); but they can, in their use as count-nouns, be accompanied by plural morphology; and they are found together with such quantifiers as "many" or "few" (as in "many beaches" and "few beaches"). The nouns in the third list standardly have both sorts of occurrences. A list of this kind, however, masks several potential sources of trouble, which an adequate treatment of the problem of classification would need to address.

AMBIGUITIES. First, some grammatical contexts are at least at first sight ambiguous, in that the most obvious syntactic criteria such as those just cited do not by them-

selves clearly differentiate a given noun-occurrence as mass or count: examples include the occurrences of "lamb," "apple," and "fish" in "Mary had a little lamb," "The apple in the dessert is moldy," and "Fish floated in the water"; the occurrence of "home" in "at home"; as well as the occurrence of "tape" in such compound expressions as "tape recorder."

TRICKY CASES. Secondly, while the syntactic criteria mentioned above involving plural morphology and quantification do speak to most of the following cases, we may wonder whether they do not in fact misclassify at least some of them:

Collective Mass: "furniture," "jewelry," "silverware," "clothing," …
Collective Plural: "spaghetti," "groceries," "news," "clothes," …
Collective Singular: "crew," "crowd," "mob," "committee," …
Irregular Plural: "scissors," "pants," "tweezers," "goggles," …
Proper Names: "Bertrand Russell," "the Holy Roman Empire," "the sixties," …

Thus, we may feel, for example, that "clothing" and "clothes" are sufficiently similar in their semantic contribution that they should be classified together, even though one occurs standardly as a mass-noun in English, whereas the other standardly occurs as an invariably plural count-noun.

ABSTRACT NOUNS. Thirdly, the syntactic criteria mentioned above also apply to nouns and noun-phrases whose denotations are either abstract or at least not straightforwardly concrete, such as the following:

Abstract Mass: "knowledge," "evidence," "poetry," "money," "information," …
Abstract Count: "belief," "mistake," "rendition," "symphony," "discovery," …
Abstract Dual-Use: "logic," "truth," "justification," "science," "theory," …

It has, however, been questioned whether the mass/count-distinction can be sensibly drawn for such nouns and noun-phrases, possibly because the semantic and ontological vocabulary, which will feature prominently below, may not easily extend to their case.

NEW USES FOR OLD NOUNS. Fourthly, it should be noted that the examples given so far attest only to the way in which these nouns are currently and standardly used in English. However, it is relatively straightforward to introduce new uses for old nouns, or even to use a noun in a nonstandard way without much setup. For example, the noun "email" has effortlessly acquired a count-use, even though it was initially used only as a mass-noun; moreover, the use of "car" in "A BMW 300-series is not much car for the money," while deliberately nonstandard, is, as far as issues of grammar are concerned, not completely out of the question. Thus, the mass/count-distinction cannot be viewed as written in stone even within a particular language; expressions can change their status, if speakers of the language, for whatever reasons, so desire.

CROSS-LINGUISTIC VARIATION. Finally, there is considerable cross-linguistic variation in how particular languages pattern with respect to the mass/count-distinction. For one thing, specific nouns that belong to different languages but intuitively have the same meaning can be classified as mass in one language and count in another; for example, the German word for hair ("Haare") is, except for poetic contexts (such as "Rapunzel, let your hair down!"), standardly used only as a singular or plural count-noun, whereas the English noun "hair" standardly has both mass- and count-uses. Furthermore, different languages can differ in how they mark the mass/count-distinction or, indeed, in whether they do so in any obviously visible way at all. In this context, it has been observed that Asian classifier-languages such as Mandarin Chinese and Japanese are of special interest, because they require that every noun be preceded by a classifier reminiscent of the sort of "reference-dividing" relations we observe in English primarily in connection with mass-nouns and plural count-nouns ("basket of," "bouquet of," "bucket of," …). This has motivated some writers, such as R. Sharvy (1978) to speculate that perhaps all nouns are at bottom mass not only in these overt classifier-languages, but across the board, on the theory that such classifiers may be present covertly in every language.

2. THE PROBLEM OF LOGICAL FORM

While consideration of the problem of classification is often regarded only as a means to an end—namely, as a way of clarifying the nature of the subject-matter beyond the clear cases—its importance should not be underestimated, especially given its role in deciding whether or not a specific, more or less tricky, case should be viewed as a counterexample to a particular analysis. Most of the attention surrounding the mass/count-distinction, however, has been focused on the question of what (if any) its semantic and ontological significance might be. Thus, the

mass/count-distinction, more so perhaps than any other comparable issue, has provided fertile soil on which to debate questions concerning our most central semantic notions—those of meaning and truth, reference, and quantification—as well as ontological questions concerning the basic categories of what there is; and therein, surely, lies its central interest for linguists and philosophers. Among the wealth of semantic issues that are debated in this connection, the following may be singled out as particularly prominent.

Semantic Role

(iii) What is the semantic role played by mass-nouns and count-nouns?

At least as far as singular count-nouns are concerned, this question is thought to have a straightforward answer; in fact, traditional accounts of meaning, truth, reference, and quantification, with their frequent appeals to the predicate-calculus and the apparatus of set-theory, seem to be in many ways specifically tailored to the semantic needs of singular count-nouns. Such nouns are typically analyzed as playing the semantic role of a predicate whose extension consists of objects, each of which (or so it seems) could at least in principle be referred to as a such-and-such (for some appropriate substantival phrase). These objects, in turn, are thought to compose the domain of values over which variables and quantifiers are interpreted as ranging; and they are taken to enter into set-theoretic relationships with one another.

Mass-nouns and plural count-nouns, on the other hand, have for a variety of reasons resisted straightforward assimilation into this familiar vocabulary. The former in particular have appeared puzzling, for one thing, because they seem to lead, in W. V. O. Quine's words, a "semantic double-life of sorts" (1960, p. 97), in some of their occurrences (e.g., "Snow is white") apparently playing the role of a name or singular term, in others (e.g., "Most snow is white") that of a predicate or general term. This appearance of a "semantic double-life" led Quine to conclude that mass-nouns can play both roles, that of a name and that of a predicate, depending on their position within the statement (see also Ter Meulen [1981] for another version of what may be called the "mixed view"). Others have thought it necessary to choose between these two semantic categories, by defending either a version of the "name view" or the "predicate view." (For examples of the name view, see Parsons 1970, Moravcsik 1973, Bunt 1979, 1985, Chierchia 1982, Link 1983, Lønning 1987, and Zimmerman 1995; for examples of the predicate view, see Burge 1972, and Koslicki 1999; as well as, arguably, Cartwright 1963, 1965, 1970; Montague 1973; Pelletier 1974; Bennett 1977; Sharvy 1980; Roeper 1983; Pelletier and Schubert 1989; and Higginbotham 1994; though some of these writers are difficult to place.)

Finally, an influential attitude toward the apparently schizophrenic semantic behavior of mass-nouns has also been to detect here a category that resists this sort of classification into either name or predicate, because it harks back somehow to a more "primitive," "pre-individuative," "pre-reference-dividing," "merely feature-placing," "non-objectual," "pre-particular level of thought," one which predates the dichotomy of singular term and general term (see especially Strawson 1953–1954, Quine 1960, Evans 1975, and Laycock 1972, 1975, 1989, 1998 for expressions of this attitude). It is not obvious, however, what to make of this somewhat ambivalent sentiment, because apparently the mode of expression associated with the use of mass-nouns fits comfortably into our present usage and we do not currently inhabit this supposed "archaic" time.

As argued convincingly in Burge (1972), all three views—the mixed view, the name view, and the predicate view—give rise to potential difficulties. The mixed view has trouble capturing inferences which turn on the common semantic core apparently shared by both namelike and predicative occurrences of mass-nouns (e.g., "Snow is white; this stuff is snow; therefore, this stuff is white"). The name view, on the other hand, is forced to invoke an arguably question-begging "reference-dividing" relation, of the form "is a … of" (e.g., "is a quantity of"), to account for those cases in which mass-nouns play an apparently predicative role (e.g., "most snow," on this view, becomes something along the lines of "most quantities of snow"). Moreover, as noted in Koslicki (1999), the supposed evidence for the name view (and, hence, for one half of the mixed view, as well) is shaky to begin with, because it is drawn from the class of so-called generic sentences; but genericity is not a phenomenon peculiar to mass-nouns and is exhibited to an equal extent by singular and plural count-nouns.

Finally, the predicate view, given our familiar way of thinking about predication as involving domains of objects, threatens to do away completely with the intuitive contrast between the different kinds of noun-occurrences. Whether this threatened obliteration should be taken as cause for alarm, however, depends in part on one's reaction to the kind of skeptical attitude displayed in Burge (1972), according to which the mass/count-distinction seems ultimately to be a pragmatic phenomenon, the grammatical manifestation of the contrast between cases in which, for whatever reasons, standards (though not necessarily clear ones) are already available

for what is to count as a such-and-such (for some appropriate substantival phrase) and cases in which there has not been any comparable pressure to clarify or supplement our current practice.

This skeptical outlook takes the linguistic distinction in itself to be a relatively superficial phenomenon, at least from the point of view of semantics and ontology, though there might be a good deal of interest to be said about it, for example, from the perspective of epistemology, philosophy of science, philosophy of mathematics, and psychology especially concerning our practices of counting and measuring (see for example Frege 1884, Carnap 1926, Carey 1985, 1994, Xu 1997). Some of the considerations raised above in Section 1, especially the striking heterogeneity of class of expressions at issue noted in (b) and (c), as well as the flexibility of current usage and the cross-linguistic variation noted in (d) and (e), might in fact be thought to count as prima facie evidence in favor of such a skeptical approach.

In addition to the apparent "semantic double-life" that has been ascribed to mass-nouns by writers such as Quine, this mode of expression has also seemed to pose special challenges with respect to the following question:

Mass-Logic and Mass-Quantification:

(iv) How do mass-nouns behave under quantification and in combination with logical connectives such as negation, disjunction, and others?

As R. Sharvy (1980), P. Roeper (1983), J. T. Lønning (1987), and J. Higginbotham (1994) in particular have discussed in detail, it seems that such statements as "The hot coffee did not disappear" or "All phosphorus is either red or black" cannot be understood straightforwardly in terms of quantification over quantities of coffee or phosphorus and in terms of such set-theoretic notions as membership, subset, union, intersection or complement. For example, it has been argued that "All phosphorus is either red or black" does not mean the same as "Every quantity of phosphorus is either red or black," because, of those quantities of phosphorus that include both red phosphorus and black phosphorus, it is neither true to say that they are red nor that they are black (Roeper 1983, p. 254). Statements of this kind have been taken to provide motivation for thinking that, as in the case of predication, our familiar approach to quantification and other logical operations, as involving domains of objects that can be interpreted as standing in set-theoretic relations to one another, does not do justice to the semantic properties of mass-nouns and the system of determiners that accompanies them.

The suspected failure of the traditional apparatus to yield a fully general logic has commonly been traced to a certain combination of mereological characteristics exhibited by mass-nouns (or their denotations, or the concepts expressed by them). Thus, from the beginning, writers have been struck because not only do sums of, say, mud yield more mud (as of course do sums of, say, people), but because divisions of mud generally (i.e., with the exception of small and not readily accessible parts) also yield more mud (see, for example, Leonard and Goodman 1940, Goodman 1951, Quine 1960, Burge 1972, Laycock 1972, Cheng 1973, Bunt 1979, 1985, Ter Meulen 1981, Roeper 1983, Simons 1987, Higginbotham 1994, and Zimmerman 1995). The first of these properties is known as "cumulativity," the second as "distributivity," and their conjunction is often called "homogeneity"; the semantic relevance (if any) of "parts that are too small" (Quine 1960, p. 98) has given rise to what is known as the "problem of minimal parts."

Moreover, while divisions of mud into more mud, as we now know from empirical inquiry, cannot go on forever, it has been said that, at the very least, it is not part of the meaning of the term "mud" that there are atoms of mud, in the mereological sense of "atom" (i.e., quantities of mud that have no proper parts that are themselves mud), while apparently it does follow from the meaning of such terms as "person" or "people," or at least from the fact that they are standardly used as count-nouns, that their extensions do consist of such atoms, with each single person counting as one of them.

Thus, if these observations are correct, they would lead to the following tripartite division: (i) singular count-nouns are neither cumulative nor distributive, but they are atomic; (ii) plural count-nouns are cumulative and atomic, but not distributive; and (iii) mass-nouns are homogeneous (i.e., both cumulative and distributive), but nonatomic (i.e., uncommitted as between the properties of atomicity and full-fledged atomlessness). And where there are no atoms, so it has seemed to many writers, there set-theoretic operations and the associated approaches to quantification can take no hold; instead, nonatomic, algebraically characterizable systems (such as Boolean algebra or lattice theory) have seemed more appropriate in light of the semantic peculiarities of mass-nouns (see especially Cartwright 1963, for the first fully developed, but unpublished, algebraic account; later analyses in the same style include Bunt 1979, 1985, Roeper 1983, Link 1983, Simons 1987, Landman 1991, and Higginbotham 1994).

Despite the popularity of this style of approach, however, it is at least debatable, first, whether mass-nouns in fact are homogeneous, given the problem of minimal parts; and, secondly, whether the question of atomicity can in fact carry the semantic weight ascribed to it, given that, for example, we can without difficulty refer to something as a building, even when the object in question has proper parts that are themselves buildings (see Koslicki 1999 for a skeptical voice). Also relevant in this connection is the debate in contemporary metaphysics concerning the so-called "problem of the many" (see, e.g., Unger 1981), which concerns the question of whether each region of space-time occupied by something we would ordinarily refer to as, say, "one person" is in fact occupied by indefinitely many numerically distinct, but largely overlapping, persons: however exactly this debate in metaphysics ought to be resolved, at the very least we cannot accuse the philosophers involved in it of not being competent speakers of English!

3. OTHER PURPORTED DIFFERENCES

In addition to the apparent mereological differences as well as the purported differences in semantic role just cited, the following considerations are frequently also thought to bear some relevance to the mass/count-distinction.

CONSTITUTION AND THE (ALLEGED) "STUFF"/"THING" DICHOTOMY. Exaggerated emphasis on a relatively small class of examples, such as "mud" versus "chair," has led to the idea that the linguistic mass/count-distinction maps straightforwardly onto an alleged metaphysical distinction between "stuff" and "things." A related misconception is that the denotations of mass-nouns constitute the denotations of count-nouns, because it is thought that mass-nouns denote "stuff" and count-nouns denote "things," and that the former constitutes the latter. Whatever exactly the notion of "stuff" comes to, however, it is simply not true that the constitution-relation connects mass- and count-noun denotations in this one-directional way (because, for example, particular virtues may constitute someone's virtue and particular pieces of furniture constitute furniture).

Moreover, as it stands, allusions to the notion of "stuff" are, in the absence of further elucidation, not particularly helpful. According to our ordinary usage, the term, "stuff," is employed in an extremely wide and varied range of contexts and is, in fact, often intersubstitutable with the term, "thing," as in "the stuff/things you've written," "the stuff/things in your attic," and so on. Thus, unless it can be clarified, for example, whether such mass-noun denotations as asparagus, trash, jewelry or traffic should be considered "stuff," and whether such count-noun denotations as clouds, bacteria or viruses should not be considered "stuff," and, if so, why, this notion is simply too hazy to be of much theoretical use. Moreover, given the heterogeneity of the class of expressions at issue, the flexibility of current usage and the cross-linguistic variation noted in considerations (b) through (e) of Section 1, it is highly questionable whether any single metaphysical distinction can be found to underlie this linguistic contrast.

SHAPE-, STRUCTURE- AND SPACE-OCCUPANCY PROPERTIES. Relatedly, one often finds the mass/count-distinction described as involving a contrast between "units" that are "discrete," "delineated," and "definite," have a "certain shape" or "precise limits," on the one hand, and something that is more "undifferentiated," "continuous," "nondelineated," or "unstructured," on the other hand (see for example Pelletier 1991, Jespersen 1924 for representative formulations). It is difficult to tease apart how much of this vocabulary is intended to be understood epistemically (as terms such as "definite" and "precise" intimate) and how much of it is to be understood metaphysically; in either case, however, it is difficult to discern here anything more than what is already contained in either consideration (a) above or consideration (c) below.

DIVIDED REFERENCE/CRITERIA OF IDENTITY AND INDIVIDUATION. The mass/count-distinction is almost universally conceived of as involving a contrast between expressions that "carry within themselves" criteria of identity and individuation and ones that fail to supply at least one or possibly both sorts of criteria. Thus, Quine famously remarks that, while "shoe," "pair of shoes," and "footwear" all range over the same "scattered stuff," they differ in that the first two "divide their reference" in different ways and the third not at all (1960, p. 91); and P. F. Strawson comments, equally notoriously, that "the general question of the criteria of distinctness and identity of individual instances of snow or gold cannot be raised or, if raised, be satisfactorily answered," because, in his view, "we have to wait until we know whether we are talking of veins, pieces or quantities of gold, or of falls, drifts or expanses of snow" (Strawson 1953–1954, p. 242; see also Laycock 1972, pp. 31–32).

However, as Helen Cartwright has argued forcefully in a series of early papers (especially Cartwright 1965, 1970), if "individuation" is what goes on when a noun has a paradigmatically predicative occurrence (e.g., one that

appears next to such determiners such as “all,” “some,” “most,” “the,” “this,” “much,” and “little”), then the mass/count-distinction does not point to a general contrast in whether an expression “individuates,” only arguably in how it does so; moreover, the question of identity is an equally moot point, because, as Cartwright points out, there are as many clear or tricky cases on the count-side as on the mass-side (e.g., compare “word” with “work,” to use Cartwright’s example). Finally, considerations that turn on the phenomenon of change over time, as when we speak for example of something’s being the same water from one time to another, even while the water in question is slowly evaporating, also fail to isolate a feature that is peculiar to the denotations of any one class of expressions (see Laycock 1972, 1975, 1989, 1998).

COUNTING AND MEASURING. Finally, we come to a more promising area to explore in connection with the mass/count-distinction, namely the distinction between counting and measuring, that is, the distinction, on the one hand, between the practice of counting and measuring, and that between what we count and what we measure, that is, the subject-matter to which these practice are directed, on the other hand (see for example Parsons [1970] and Cartwright [1975a] for discussion of amounts and measures of amounts). Simply put, the contrast in this area is taken to be the following: whereas mass-noun denotations can only be measured, count-noun denotations can also be counted: thus, in the former case, only the vocabulary of amounts and measures of amount is appropriate, whereas the latter also admits of the apparatus of number and cardinality.

However, even in this area, matters are less clear than is often supposed. For, as it stands, the contrast between what we can and cannot measure really only marks off the sorts of magnitudes discussed by the physicist (e.g., temperature, mass, velocity, distance, and the like) from those entities which, in some way, exhibit these magnitudes; and while it is true that such magnitudes tend to be referred to by means of mass-nouns, the class of mass-nouns is of course thought to be much wider than simply what is encompassed by these magnitude-denoting terms. The area of counting as well is still radically under-explored, at least from the point of view of philosophy, though much interesting work has been done on the subject by psychologists (see for example Carey [1985, 1994] and the references cited therein). If counting involves, as Frege would put it, an association between a concept and a cardinal number, then the key question that arises in this context is just the question G. Frege himself was concerned to answer in Section 54 of the *Grundlagen,* namely what sorts of requirements must be met by a concept to admit association with number (for discussion, see for example Geach 1962, Dummett 1973, Koslicki 1997, Blanchette 1999). If what has been suggested in the previous paragraph is correct and no general contrast exists between mass- and count-nouns at least in whether they provide criteria of individuation and identity, then the answer to Frege’s question concerning counting must lie elsewhere; and what this answer is, it is fair to say, is still an open question.

IV. CONCLUSION

As sobering as we might find this outcome to be, it may be that, at the end of the day, the only absolutely general and incontestable truism that can be stated in connection with the mass/count-distinction is that a true statement containing a singular or plural count-noun, as in “There is a hair in my soup” or “There are hairs in my soup,” insures the presence of either exactly one whole hair, or exactly two whole hairs, and so forth, whatever precisely this comes to in metaphysical terms; whereas a true statement of the form “There is hair in my soup” is compatible with there not being exactly one whole hair, or exactly two whole hairs, and so forth, because what is present may be parts of hairs or sums of parts of hairs or sums of hairs. And while this truth-conditional difference, stated in this stark and austere form, without the usual accompaniment of highly metaphorical and generally unhelpful vocabulary, might at first glance strike us as entirely trivial, its semantic and ontological significance, as can be gleaned among other things from the sorts of inferences that are licensed by it, should not be underestimated. Even if hair, perhaps, is no more “stufflike” than hairs, there is still an interesting story to be told as to what makes something one whole hair, or, for that matter, one whole anything (see Fine 1994, 1999, Harte 2002).

See also Aristotle; Frege, Gottlob; Plato; Pre-Socratic Philosophy; Proper Names and Descriptions; Properties; Quine, Willard Van Orman; Semantics; Strawson, Peter Frederick.

Bibliography

Bennett, M. “Mass Nouns and Mass Terms in Montague Grammar.” In *Linguistics, Philosophy, and Montague Grammar*, In edited by S. Davis and M. Mithun, 263–285. Austin: University of Texas Press, 1977.

Blanchette, P. “Relative Identity and Cardinality.” *Canadian Journal of Philosophy* 29 (1999): 205–224.

Bunt, H. C. “Ensembles and the Formal Semantic Properties of Mass Terms.” In *Mass Terms: Some Philosophical Problems*,

edited by Francis Jeffry Pelletier, 249–277. Dordrecht, Netherlands: D. Reidel, 1979.

Bunt, H. C. *Mass Terms and Model-Theoretic Semantics*. Cambridge, U.K.: Cambridge University Press, 1985.

Burge, T. "Truth and Mass Terms." *Journal of Philosophy* 69 (1972): 263–282.

Carey, S. *Conceptual Change in Childhood*. Cambridge, MA: MIT Press, 1985.

Carey, S. "Does Learning a Language Require the Child to Reconceptualize the World?" *Lingua* 92 (1994): 143–167.

Carnap, R. "Physikalische Begriffsbildung." In *Wissen und Werken* 39. Karlsruhe, Germany: G. Braun, 1926.

Cartwright, H. M. "Amounts and Measures of Amounts." *Nous* 9 (1975a): 143–164.

Cartwright, H. M. *Classes, Quantities and Non-Singular Reference*. Unpublished Ph.D. diss. University of Michigan, Ann Arbor, 1963.

Cartwright, H. M. "Heraclitus and the Bath Water." *Philosophical Review* 74 (1965): 466–485.

Cartwright, H. M. "Quantities." *Philosophical Review* 79 (1970): 25–42.

Cartwright, H. M. "Some Remarks About Mass Nouns and Plurality." *Synthese* 31 (1975b): 395–410.

Cheng, C.-Y. "Comments on Moravcsik's Paper." In *Approaches To Natural Language*, edited by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes, 286–288. Dordrecht, Netherlands: D. Reidel, 1973.

Chierchia, G. "On Plural and Mass Nominals." *Proceedings of the West Coast Conference on Formal Linguistics* 1 (1982): 243–255.

Dummett, M. *Frege: Philosophy of Language*. Cambridge, MA: Harvard University Press, 1973.

Evans, G. "Identity and Predication." *Journal of Philosophy* 72 (1975): 343–363.

Fine, K. "Compounds and Aggregates." *Nous* 28 (1994): 137–158.

Fine, K. "Things and Their Parts." *Midwest Studies in Philosophy* 23 (1999): 61–74.

Frege, G. *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*. Translated by J. L. Austin. Evanston, IL: Northwestern University Press, 1980.

Geach, P. T. *Reference and Generality*. Ithaca, NY: Cornell University Press, 1962.

Gillon, B. S. "Towards a Common Semantics for English Count and Mass Nouns." *Linguistics and Philosophy* 15 (1992): 597–639.

Goodman, N. *The Structure of Appearance*. Dordrecht, Netherlands: D. Reidel, 1951.

Harte, V. *Plato on Parts and Wholes: The Metaphysics of Structure*. Oxford, U.K.: Clarendon Press, 2002.

Higginbotham, J. "Mass and Count Quantifiers." *Linguistics and Philosophy* 17 (1994): 447–480.

Hoepelman, J. "Mass Nouns and Aspect, or: Why We Can't Eat Gingercake in an Hour." *Amsterdam Papers in Formal Grammar* 1, edited by J. Groenendijk and M. Stokhoff, 132–153. 1976.

Jespersen, Otto. *The Philosophy of Grammar*. London: Allen and Unwin, 1924.

Koslicki, K. "Isolation and Non-Arbitrary Division: Frege's Two Criteria for Counting." *Synthese* 112 (1997): 403–430.

Koslicki, K. "The Semantics of Mass-Predicates." *Nous* 33 (1999): 46–91.

Koslicki, K. *Talk about Stuffs and Things: The Logic of Mass and Count Nouns*. Unpublished Ph.D. diss. Cambridge, MA: MIT, 1995.

Landman, F. *Structures for Semantics*. Dordrecht, Netherlands: Kluwer, 1991.

Laycock, H. "Matter and Objecthood Disentangled." *Dialogue* 28 (1989): 17–21.

Laycock, H. "Some Questions of Ontology." *Philosophical Review* 81 (1972): 3–42.

Laycock, H. "Theories of Matter." *Synthese* 31 (1975): 411–442.

Laycock, H. "Words without Objects." *Principia* 2 (1998): 147–182.

Leonard, H., and N. Goodman. "The Calculus of Individuals and Its Uses." *Journal of Symbolic Logic* 5 (1940): 45–55.

Link, G. "The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach." In *Meaning, Use, and Interpretation of Language*, edited by R. Bäuerle, C. Schwarze, and A. Stechow, 302–323. Berlin: Walter de Gruyter, 1983.

Lønning, J. T. "Mass Terms and Quantification." *Linguistics and Philosophy* 10 (1987): 1–52.

Montague, R. "Comments on Moravcsik's Paper." In *Approaches To Natural Language*, edited by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes, 289–294. Dordrecht, Netherlands: D. Reidel, 1973.

Moravcsik, J. "Mass Terms in English." In *Approaches To Natural Language*, edited by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes, 263–285. Dordrecht, Netherlands: D. Reidel, 1973.

Mourelatos, A. "Events, Processes and States." *Linguistics and Philosophy* 2 (1978): 415–434.

Parsons, T. "An Analysis of Mass Terms and Amount Terms." *Foundations of Language* 6 (1970): 362–388.

Pelletier, F. J. "Mass Terms." In *Handbook of Metaphysics and Ontology*, edited by H. Burkhardt and B. Smith, 1, 495–499. Munich: Philosophia Verlag, 1991.

Pelletier, F. J., ed. *Mass Terms: Some Philosophical Problems*. Dordrecht, Netherlands: D. Reidel, 1979.

Pelletier, F. J. "Non-Singular Reference: Some Preliminaries." *Philosophia* 5 (1975): 451–465.

Pelletier, F. J. "On Some Proposals for the Semantics of Mass Terms." *Journal of Philosophical Logic* 3 (1974): 87–108.

Pelletier, F. J., and L. K. Schubert. "Mass Expressions." In *Handbook of Philosophical Logic*. Vol. 4, edited by D. Gabbay and F. Guenthner, 327–407. Dordrecht, Netherlands: D. Reidel, 1989.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Roeper, P. "Semantics for Mass Terms with Quantifiers." *Nous* 17 (1983): 251–265.

Schubert, L., and F. J. Pelletier. "Problems in the Representation of the Logical Form of Generics, Plurals and Mass Nouns." In *New Directions in Semantics*, edited by E. LePore, 385–451. London: Academic Press, 1987.

Sharvy, R. "Maybe English Has No Count Nouns: Notes on Chinese Semantics." *Studies in Language* 2 (1978): 345–365.

Sharvy, R. "A More General Theory of Definite Descriptions." *Philsophical Review* 89 (1980): 607–624.

Simons, P. *Parts: A Study in Ontology*. Oxford, U.K.: Clarendon Press, 1987.

Strawson, P. F. "Particular and General." *Proceedings of the Aristotelian Society* 54 (1953–1954): 233–260.

Taylor, B. "Tense and Continuity." *Linguistics and Philosophy* 1 (1977): 199–220.

Ter Meulen, A. "An Intensional Logic for Mass Terms." *Philosophical Studies* 40 (1981): 105–125.

Ter Meulen, A. *Substance, Quantities and Individuals*. PhD Thesis, Stanford University, 1980.

Unger, P. "The Problem of the Many." In *Studies in Epistemology*, edited by P. French, T. Uehling, and H. Wettstein, 411–468. Midwest Studies in Philosophy 5. Minneapolis: University of Minnesota Press, 1981.

Ware, R. X. "Some Bits and Pieces." *Synthese* 31 (1975): 379–393.

Xu, Fei. "From Lot's Wife to a Pillar of Salt: Evidence that *Physical Object* is a Sortal Concept." *Mind and Language* 12 (1997): 365–392.

Zimmerman, D. W. "Theories of Masses and Problems of Constitution." *Philosophical Review* 104 (1995): 53–110.

Kathrin Koslicki (2005)

NOUS

Nous is most likely derived from the root *snu*, meaning "to sniff." Homer uses *nous* to mark the realization or understanding of a situation or state of affairs. *Nous* penetrates beyond the surface features of a situation and reveals the underlying truth of the matter. It is not divorced from perception and its most primitive function is that of apprehending or "smelling" danger. In Homer *nous* is also linked to the visualization of a plan of action that is immediately prompted by the awareness of a situation possessing emotional impact.

In Parmenides *nous* maintains its Homeric function as that which reveals ultimate truth. However, it also serves as the source of logical reasoning. In Parmenides *nous* is divorced from perception and it is best understood to mean "thought" or "intellect." In accordance with his rather austere ontology, Parmenides may well hold that that which exists is also that which thinks (i.e., no thing that exists fails to be a thing that thinks).

Anaxagoras treats *nous* as a mass term, like water or air (as opposed to a count term, like man or leaf). He appears to treat *nous*, not as "intellect," but as "reason" or "the virtue of rationality." *Nous*, for Anaxagoras, is the ultimate source of order and motion in the cosmos. By both initiating and governing a vortex, *nous* brings order to an otherwise static primordial chaos. Anaxagoras asserts that *nous* is the lightest and purest thing. In so doing, he may well be attempting to articulate the idea that *nous* is an immaterial substance.

Plato incorporates elements from Parmenides, Homer, and Anaxagoras into his treatment of *nous*. First, following Parmenides, Plato considers *nous* to be an intellectual faculty that is wholly divorced from perception. Second, following Homer, Plato considers *nous* to be a source of insight or intuition. Still, for Plato, intuition is a nonempirically based grasp of unchanging and eternal truth. Finally, following Anaxagoras, Plato considers *nous* to be the source of order and motion in the cosmos. *Nous*, as rationality itself, is the substance that orders the heavens for the sake of the best. It is the cause of regular celestial motion and it is the cause of rationality in humans.

Aristotle, in his treatment of *nous*, displays acute awareness of views advanced by his predecessors. First, Aristotle takes *nous* to be a source of insight. *Nous* is a grasp of the salient features of a situation, but it is also a grasp of universal scientific principles. *Nous*, even in its later role, is not divorced from perception. It is the grasp of principles that are acquired by induction from perceived cases. Second, Aristotle uses *nous* to mean "intellect." He asserts that one's *nous* is separate from the body. In so doing, Aristotle is likely to be advancing the view that human intellect is an immaterial faculty. Finally, Aristotle's God, the Prime Mover, is *nous*. It is a separately existing and fully actualized rationality. This *nous* is the chief cause of motion, order, and goodness in the cosmos.

See also Anaxagoras of Clazomenae; Aristotle; Homer; Parmenides of Elea; Perception; Plato; Thinking.

Bibliography

Lesher, James H. "The Meaning of *Nous* in the Posterior Analytics." *Phronesis* 18 (1973): 44–68.

Menn, Stephen. *Plato on God as Nous*. Carbondale: Southern Illinois University Press, 1995.

Von Fritz, Kurt. "*Noos* and *Noein* in the Homeric Poems." *Classical Philology* 38 (1943): 79–93.

Von Fritz, Kurt. "*Nous*, *Noein*, and Their Derivatives in Pre-Socratic Philosophy (excluding Anaxagoras)." *Classical Philology* 40 (1945): 223–242 and 41 (1946): 12–34.

John E. Sisko (2005)

NOVALIS
(1772–1801)

Novalis was the pseudonym of Friedrich Leopold Freiherr von Hardenberg, the lyric poet and leader of the

early German romanticists. Novalis was born of Pietistic parents on the family estate, Oberwiederstedt, in Saxony. In preparation for a civil service career, he studied jurisprudence, philosophy, chemistry, and mathematics at Jena, Leipzig, and finally at Wittenberg, where he completed his studies in 1794. In Jena, Novalis came under the influence of Johann Wolfgang von Goethe, Friedrich Schiller, and especially Johann Gottlieb Fichte. Soon afterward he became friendly with Friedrich and August Wilhelm von Schlegel, Ludwig Tieck, Friedrich von Schelling, and Johann Wilhelm Ritter. While apprenticed to a local official in Tennstedt, Novalis became engaged to thirteen-year-old Sophie von Kühn in 1795. Her death in 1797 reinforced his romantic mysticism and culminated in a poetic transfiguration of his loss, in which his love and his desire to follow her into death are mingled (*Hymnen an die Nacht,* first published in 1800). From 1796 on, Novalis worked in the administration of the Saxon salt works at Weissenfels. From 1797 to 1799 he studied mining at Freiburg, where he became engaged to Julie von Charpentier. He died at Weissenfels.

With Friedrich Schlegel, Novalis is the most characteristic spokesman of early romanticism. In opposition to the ideals of the Enlightenment and early classicism he presented his vision of the romantic life. In his novelistic fragment *Heinrich von Ofterdingen,* which was written in opposition to Goethe's *Wilhelm Meister,* he furnished the age with a poetic description of the poet. The self-consciousness implicit in such an undertaking is characteristic of Novalis. Thinking about his own situation, the poet tries to answer the more general question of the destiny of humankind; the poet is a seer who leads man home. The homelessness presupposed in this theme is also manifest in Novalis's characterization of the modern age as fragmented. By contrast, according to Novalis's idealized picture, the Middle Ages was a time of unity.

These ideas are further developed in *Die Christenheit oder Europa* (1799), an essay on the history of Western civilization, in which Novalis attacks the Protestant Reformation and the Enlightenment for having destroyed medieval unity. Also, he proposes that the most important reason for the homelessness of man is simply that he is a finite being. To be finite is to be in search of the infinite, which can be recovered in the depths of the human soul, a concept which develops ideas derived from Fichte's *Wissenschaftslehre.* Meaning, being, and truth are identified with the absolute ego. When the adept in *Die Lehrlinge zu Sais* (1798) lifts the veil of Isis that hides the meaning of human existence, he discovers only his true self. At the same time, this discovery is an escape from all that separates man from nature and from others.

The poet, through knowledge of his true self, is intuitively able to grasp the meaning of the world, which is veiled by mechanistic explanations, and to reveal this meaning to others. Poetry is an attempt to draw away the veil of the finite, which hides the mysterious meaning of everything. It thus has an apparently negative effect. The claims of the finite must be destroyed for the sake of the infinite. Romantic irony negates the ordinary significance of things and paves the way for a magic transformation of reality. Novalis's magic idealism may be described as an esoteric game in which relationships are suggested that may seem fantastic but are designed to reveal a higher meaning. The best example of this is *Heinrich von Ofterdingen,* in which past and present, fairy tale and everyday reality, mingle in such a way that the reader loses his bearings. This loss liberates his imagination. The world reveals its meaning when it is transformed into something man has freely chosen, and the opposition between man and nature is thereby overcome. Salvation lies in the godlike freedom of the artist.

Meaning escapes adequate conceptualization; it can only be hinted at. Fragment and aphorism (*Blütenstaub,* published in 1798) lend themselves particularly well to this purpose, as they point to meanings beyond themselves which must remain unstated. The romantic's refusal to mediate between the finite and the infinite, his assertion that there is no relationship between mere facts and transcendent meanings, makes it impossible to give any definite content to that reality which is said to be the goal of man's search. The movement toward salvation becomes indistinguishable from a flight into nothingness. Thus, in his *Hymnen an die Nacht* Novalis celebrates the night, in which all polarities are reconciled, and opposes it to more shallow day—a theme taken up by Arthur Schopenhauer, Friedrich Nietzsche, and their more recent followers.

See also Enlightenment; Fichte, Johann Gottlieb; Goethe, Johann Wolfgang von; Nietzsche, Friedrich; Reformation; Romanticism; Schelling, Friedrich Wilhelm Joseph von; Schiller, Friedrich; Schlegel, Friedrich von; Schopenhauer, Arthur.

Bibliography

WORKS BY NOVALIS

German Editions

Editions of Novalis's collected works are Ludwig Tieck and Friedrich Schlegel, eds., 2 vols. (Berlin, 1802); Paul

Kluckhohn and Richard Samuel, eds., 4 vols. (Leipzig, 1929; 2nd ed., Stuttgart, 1960–); C. Seeling, ed., 5 vols. (Zürich: Bühl, 1945); and E. Wasmuth, ed., 4 vols. (Heidelberg, 1953–1957).

For Novalis's philosophical works, see *Das philosophische Werke,* Richard Samuel, ed., Vol. I (Stuttgart, 1965), the first volume of a projected four-volume critical edition.

English Translations

The Devotional Songs of Novalis. Edited by Bernard Pick. Chicago, 1910. Also contains German text.

Henry of Ofterdingen: A Romance. Cambridge, MA: Owen, 1842.

Hymns to the Night and Other Selected Writings. Translated by C. E. Passage. New York, 1960.

The Novices of Sais. Translated by R. Manheim. New York, 1949.

WORKS ON NOVALIS

Biser, E. *Abstieg und Auferstehung, Die geistige Welt in Novalis Hymnen an die Nacht.* Heidelberg, 1954.

Carlyle, Thomas. "Novalis." In *Critical and Miscellaneous Essays,* in *Works,* edited by H. D. Trail. London, 1896–1899; New York, 1896–1901.

Dilthey, W. *Das Erlebnis und die Dichtung: Lessing, Goethe, Novalis, Hölderlin,* 3rd ed. Leipzig, 1910.

Friedell, E. *Novalis als Philosoph.* Munich, 1904.

Kuhn, H. "Poetische Synthesis oder ein kritischer Versuch über romantische Philosophic und Poesie aus Novalis Fragmenten." *Zeitschrift für Philosophische Forschung* (1950–1951): 161–178; 358–384.

Küpper, P. *Die Zeit als Erlebnis des Novalis.* Cologne: Böhlau, 1959.

Rehm, W. *Orpheus, Der Dichter und die Toten.* Düsseldorf: L. Schwann, 1950.

Karsten Harries (1967)

NOZICK, ROBERT

(1938–2002)

Robert Nozick was born in Brooklyn, New York, graduated from Columbia University in 1959, and received a PhD from Princeton University in 1963. After stints at Princeton University and the Rockefeller University, Nozick went to Harvard University in 1969, at age thirty, as full professor. There he was named Arthur Kingsley Porter Professor of Philosophy in 1985, then Joseph Pellegrino University Professor in 1998. He was a fellow of the American Academy of Arts and Sciences and served as president of the Eastern Division of the American Philosophical Association.

Nozick and his Harvard colleague John Rawls were the giants of twentieth-century political philosophy. Where Rawls stuck to one task, elaborating and defending his magisterial *Theory of Justice*, Nozick was notably restless and interested in everything. He once said, "I didn't want to spend my life writing *Son of Anarchy, State, and Utopia*, *Return of the Son*, and so on" (Socratic Puzzles 1997, p. 2). In an age of subspecialization, the range of Nozick's contributions is shocking.

POLITICAL PHILOSOPHY

In *A Theory of Justice*, Rawls described himself as working toward a theory of pure procedural justice. He proposed as a test of distributive justice that inequalities are just only if they offer the greatest possible benefit to the worst-off. In *Anarchy, State, and Utopia* (1974), Nozick's departure was to develop a *genuinely* procedural theory, aimed at no particular end state. Indeed, Nozick's product was less a theory of just distribution than a theory of just transfer. A transfer from one person to another is truly just, according to Nozick, if truly voluntary.

Nozick's argument sometimes is said to lack foundations, to merely postulate rights. More charitably, Nozick's bold claims about rights are his conclusions rather than his premises. Starting from Rawls's foundation—individuals are separate and may not be sacrificed for others—Nozick, in the process arguing for this premise, carries it to its logical conclusion. Part 1 of *Anarchy, State, and Utopia* argues that a world where persons are respected as separate entities within a minimal state is a *possible* world. Part 3 argues that this is an *attractive* world. Part 2 argues that a world where our separateness is *not* taken to its logical conclusion—not taken to culminate in some more or less literal interpretation of Rawls's call for the "most extensive system of liberty compatible with like liberty for all" (1971, p. 302)—is neither attractive nor just.

In one of the century's more influential philosophical examples, Nozick asks us to suppose that we are in a situation as perfectly just and equal as we can imagine. Then someone offers Wilt Chamberlain a dollar for the privilege of watching him play basketball. Before we know it, thousands of people happily are paying Wilt a dollar each every time he puts on a show. Wilt gets rich. The distribution is no longer equal, but no one is complaining. Nozick's question: If we assume for argument's sake that justice is a pattern of equality achievable at a given moment, what happens if we achieve the ideal? Must we then prohibit everything—consuming, creating, trading, giving—that upsets perfect equality? Recent egalitarian work is an evolving response to the problem Nozick's story revealed. In part due to Nozick's argument, egalitarians at the beginning of the twenty-first century realize that any equality worthy of aspiring to will focus

less on equality as a time-slice property of economic distribution of wealth and more on how people are treated: how they are rewarded for their contributions and *enabled* over time to make contributions worth rewarding.

METAPHYSICS AND EPISTEMOLOGY

Nozick's last book, *Invariances* (2001), spans a range of topics including truth, objectivity, and consciousness, and his second book, *Philosophical Explanations* (1981), offers fresh ideas on free will, personal identity, and knowledge. For example, philosophers for millennia had analyzed knowledge as justified true belief. That is, *S* knows that *p* just in case *p* is true, *S* believes that *p*, and *S*'s belief is justified. Since 1963, though, philosophy had been reeling from Edmund Gettier's refutation of this seemingly straightforward analysis. Nozick's response is among the most creative. The problem with justification, as Gettier construed it, is that a belief can be justified, in virtue of coinciding with the facts, without being properly *sensitive* to the facts. Nozick, instead of refining or supplementing the justification condition, *replaced* it with a pair of *tracking* conditions:

> If it were not true that *p*, *S* would not believe that *p*.
>
> If it were true that *p*, *S* would believe that *p*.

DECISION THEORY

Nozick's *Socratic Puzzles* (1997), a collection of essays, includes his essay "Newcomb's Problem and Two Principles of Choice." In it Nozick introduced a class of puzzles for prevailing formulas for maximizing expected utility. For example, the devout go to heaven, according to John Calvin, but why? *Because* they are devout? If so, expected utility would suggest that we ought to be devout. Or because of predetermined grace, a side effect of which is an urge to be devout? In this second case, since it is more fun not to be devout, expected utility would suggest that we ought not to be devout. The crucial issue is not whether the outcome is probabilistically linked to one's action but whether it is *affected* by one's action. Therefore, rational choice cannot be entirely captured by any probabilistic formula. Even at its most formulaic, rational choice would have to begin with the problem of choosing a formula to govern subsequent choices, soothe choosing begins prior to having the formula. The chosen formula will be a way of processing information not only about probabilities and utilities but also about causal connections between actions and outcomes. Nozick's essay spawned hundreds of responses.

NONCOERCIVE PHILOSOPHY

One of Nozick's biggest contributions to philosophy was to reflect on, and poke fun at, the competitiveness of philosophical discourse. Nozick returned to this theme in the introductions to each of his major works; it was the only topic that occupied Nozick continuously. "Philosophical training molds arguers. … A philosophical argument is an attempt to get someone to believe something, whether he wants to believe it or not. … To argue with someone is to attempt to push him around verbally. … Perhaps philosophers need arguments so powerful they set up reverberations in the brain: if the person refuses to accept the conclusion, he *dies*" (1981, p. 4). Nozick's remarks on the ideal of "coercive philosophy" led to a generation of self-deprecating humor in seminars across the United States and eventually to a widespread relaxing of what had been a more confrontational, less cooperative disciplinary style.

See also Calvin, John; Decision Theory; Justice; Personal Identity; Rawls, John; Rights.

Bibliography

WORKS BY NOZICK

Anarchy, State, and Utopia. New York: Basic Books, 1974.

Philosophical Explanations. Cambridge, MA: Harvard University Press, 1981.

The Examined Life. New York: Simon and Schuster, 1989.

The Nature of Rationality. Princeton, NJ: Princeton University Press, 1993.

Socratic Puzzles. Cambridge, MA: Harvard University Press, 1997.

Invariances: The Structure of the Objective World. Cambridge, MA: Harvard University Press, 2001.

WORKS ON NOZICK

Feser, Edward. *On Nozick*. Toronto: Thomson Wadsworth, 2004. An accessible discussion of *Anarchy, State, and Utopia*.

Lacey, A. R. *Robert Nozick*. Princeton, NJ: Princeton University Press, 2001. Discusses the range of Nozick's thought.

Paul, Ellen, Jeffrey Paul, and Fred D. Miller Jr., eds. *Natural Rights Liberalism from Locke to Nozick: Essays in Honor of Robert Nozick*. New York: Cambridge University Press, 2005.

Paul, Jeffrey, ed. *Reading Nozick*. Oxford, U.K.: Blackwell, 1982. A collection of responses to *Anarchy, State, and Utopia*.

Rawls, John. *A Theory of Justice*. Cambridge, MA: The Belknap Press of Harvard University Press, 1971.

Schmidtz, David. *Robert Nozick*. New York: Cambridge University Press, 2002. A collection of essays by specialists in the various fields to which Nozick contributed.

Schmidtz, David, and Sarah Wright. "What Nozick Did for Decision Theory." In *The American Philosophers*, Peter French and Howard K. Wettstein, eds. Oxford, U.K.: Blackwell, 2004.

Wolff, Jonathan. *Robert Nozick: Property, Justice, and the Minimal State*. Stanford, CA: Stanford University Press, 1991. The first book-length critique of *Anarchy, State, and Utopia*.

David Schmidtz (2005)

NUMBER

Numbers are central to science. They underlie what Galileo Galilei and Isaac Newton called the primary properties of things, the properties that can be measured (John Locke listed these as number, motion and rest, size, figure, and impenetrability). These underlie secondary properties (like colors and musical harmonies and discords), which in turn underlie the tertiary properties, like beauty, which make life worth living.

The centrality of numbers to science indirectly confers on them philosophical significance, but they have also played a direct role in metaphysics. Plato's theory of universals begins from the problem of the One over Many. Behind the superficial diversity of things in the world, it is often the case that there is one thing that many numerically distinct individuals share in common. For instance, when one doubles the length of the string on a lyre or the length of a column of air in a flute, the note it sounds is always lowered by the same musical interval, an octave. The things that distinct individuals share in common are called universals, and "Platonism" is used as a name for a broad and loose family of theories that affirm the existence of universals.

The existence of numbers has always been central to the history of Platonism, from ancient times to the present. In the nineteenth and twentieth centuries foundational work in the philosophy of mathematics, especially by Gottlob Frege and Bertrand Russell, affirmed the existence of numbers. Following them, Willard Van Orman Quine affirmed the existence of numbers, and he rightly called this doctrine "Platonism". Quine argued that it is reasonable to believe in the existence of numbers because numbers are central to mathematics, which in turn is central to science. This reason for believing in numbers is close to the guiding Pythagorean and Platonist idea that to understand the world we must find the unified mathematical patterns that lie behind the diversities of appearances.

EARLY HISTORY OF NUMBERS

The history of numbers in India, China, and elsewhere is deep and diverse, but it is still not properly understood. In ancient and modern histories of ideas in Europe, the origin of geometry was traditionally traced to ancient Egypt; and relatively sophisticated advances in arithmetic and algebra have been recognized as having emerged in Mesopotamia; and both these sources entered European traditions through ancient Greece. Knowledge of this ancient history is improving, but it is still incomplete.

The early mathematical advances of ancient Greece are better known, though even here the evidence is sparse. Almost no written records survive from the Pythagorean oral traditions before Plato. What survives from before Euclid's *Elements* consists in little more than hints in Plato and Aristotle.

Euclid's *Elements*, the first systematic presentation of geometry and arithmetic, is magnificent, but little is known of its sources and motivations. It is relatively apparent, however, that some of his theorems consist in translations of algebraic results, known in Mesopotamia, into geometric counterparts. For instance, an algebraic thesis, like $(a + b)^2 = (a^2 + 2ab + b^2)$, would become a theorem concerned with the division of a square into two smaller squares and two rectangles. For some reason, the mathematicians of Plato's Academy emphasized geometry rather than arithmetic, and arithmetic was subsumed under geometry.

PROLIFERATION OF KINDS OF NUMBERS

Besides the whole numbers (or natural numbers) the Greeks also recognized relationships of ratio between numbers. For example, the numbers 9 and 6 stand in the same ratio as 3 to 2, and one can call this ratio (3:2). This same relationship of ratio that holds between any two numbers will also hold between two possible geometrical lengths.

However, among the relationships of ratio that hold between various magnitudes, as for instance between lengths of lines, there are some that do not hold between any two whole numbers. Plato and Aristotle allude, many times, to a proof that no ratio between whole numbers will match the relationship of proportion that holds between the diagonal and the side of a perfect square. This fact would now be expressed by saying that $\sqrt{2}$ is an irrational number, which means that there are no whole numbers a and b such that $a / b = \sqrt{2}$.

The ancient Greeks thought of ratios among lines as forming a domain distinct from the domain of numbers. Numbers consisted simply of whole numbers. As the centuries advanced, the term *number* gradually expanded to

include the entire domain of what are now called the positive real numbers. This domain includes all the irrational numbers, such as $\sqrt{2}$ and π, that can be represented as nonterminating decimals (such as 3.1415926 …). This domain includes, as a subdomain, the rational numbers, which are the ratios that hold between whole numbers; and, as a smaller subdomain, the integers, which correspond to just the rational ratios to the unit measure. The domain of number did not initially, however, include the number zero or the negative numbers.

Over several centuries the domain of things that were included as numbers expanded to include zero and negative numbers. First, there was an expansion to include a symbol "0" that was at first to be thought of not as signifying any number, but just as a place holder in the system of Arabic notation that is used today. In the notation "12," the "1" is placed in the second column from the right, and this means that it signifies one group of ten. Take 2 away from 12 and the result is written "10" with the "0" not referring to anything at all, but just serving to keep the "1" in the second column, so that it continues to signify one group of ten.

As time went by, however, the symbol "0" did come to be thought of as standing for something that might be called "the number zero", trusting that there was some suitable thing for this symbol to refer to. It was only gradually that any clear conception began to arise of what kind of thing this number zero might be.

Likewise, negative numbers began as notation that did not refer to any extra numbers, but just told one what to do with ordinary, positive whole numbers. With time, however, this notation came to be thought of as referring to new numbers, and eventually a conception emerged about what kinds of things these new numbers might be.

There was also a tentative expansion, with deep philosophical misgivings, to include what are now known as imaginary and complex numbers. Briefly, the imaginary number *i*—assuming there is such a thing, and calling it a number—is defined to be that mathematical object that is such that the ratio of 1 to it is the same as the ratio of it to minus 1. That is, $i / 1 = -1 / i$, so that $i^2 = -1$. Complex numbers consist of all the numbers that can be obtained from *i* by taking multiples of it and adding the result to other numbers.

There was also a tentative expansion, with deep philosophical misgivings, to include infinitesimal magnitudes. These extra entities seemed to be indispensable in the new mathematical theory of physical magnitudes like velocity and acceleration, invented by Newton and Gottfried Wilhelm Leibniz and referred to as "calculus" or, in its most general form, "analysis".

Despite the immense success of the calculus in science, the concept of an infinitesimal—a magnitude greater than zero, but less than any finite magnitude—was viewed with some suspicion. In the nineteenth century, through the work of Augustine-Louis Cauchy and Karl Weierstrass, the concept of an infinitesimal was replaced by the concept of the limit of a sequence of numbers. An infinite sequence of numbers $s_1, s_2, s_3, \ldots$ approaches the limit l if the difference between l and s_n can be made as small as one likes by taking sufficiently large values of n. That is, given any positive number d, no matter how small, there is some number N such that the difference between l and s_n is less than d, for every n (N.

Using this concept, the nineteenth-century mathematicians showed how the concepts of continuity, convergence, differential, and integral could all be precisely defined. In this way, it was shown how talk of infinitesimals could be dispensed with entirely.

A further nineteenth-century development was the introduction by Georg Cantor of the concept of a transfinite number. The transfinite numbers can be thought of as measuring the size of infinite sets. Cantor introduced the symbol $\aleph_0$ (pronounced "aleph-null") for the number measuring the size of the set of all positive whole numbers and the symbol c for the transfinite number measuring the size of the set of all real numbers. By a simple, yet ingenious argument (the celebrated diagonal argument), Cantor was able to show that there are more real numbers than whole numbers: $c > \aleph_0$.

Cantor proved that there are always more subsets of a given set than elements of that set (so there are more sets of natural numbers than natural numbers for example). Hence, given any transfinite number measuring the size of an infinite set, there is a larger transfinite number, which measures the size of the set of all subsets of that set.

Cantor developed a transfinite arithmetic for these new numbers, showing how operations corresponding to addition and exponentiation could be defined for them. Again, the new numbers were viewed initially with the deepest suspicion by the mathematical community.

FREGE AND THE PARADOXES

The work of the nineteenth-century mathematicians had begun a reverse process of defining one kind of number in terms of simpler kinds. The complex numbers, it had been shown, could be defined as pairs of real numbers (like the *x*-*y* coordinates of Cartesian geometry) along

with special rules for adding and multiplying these pairs. The real numbers had been shown by Julius Dedekind and Cauchy to be definable as infinite sequences or sets of rational numbers, while the rational numbers themselves can be identified with sets of pairs of natural numbers.

What of the natural numbers themselves? Frege's work can be seen as an attempt to complete this reverse process of rigorization by providing a firm foundation for the fundamental theory of the natural numbers. Dedekind and Giuseppe Peano had independently specified some simple axioms for that theory (called number theory or arithmetic). However, Frege wanted to answer the questions: What are the natural numbers? How may they be defined? The Dedekind-Peano axioms specify the laws governing the numbers, but do not provide a definition of them.

Imagine if one thought that the number of soldiers in an army was one of the army's most significant properties. One might then think of whole numbers as properties of aggregates. However, as Frege pointed out, if one points to the things on a desk and asks how many there are, one has not yet asked a complete question. There may be two decks of cards; and if so, then there are also 106 cards; and there are a great many molecules; and so on.

This suggests that number is a property of properties. The property of "being a deck of cards on the table" has the property of having one instance; the property of "being a card on the table" has the property of having 106 instances; and so on. The property of "being a unicorn" has the property of having no instances. That higher-order property, the property of having no instances, might aptly be called the number zero.

Consider, then, the theory that the number 2 is a property of a property, namely the property of having two instances, that the number 3 is the property of having three instances, and so on. Frege turned decisively aside from this theory. He argued that numbers could not be universals or concepts, but had to be objects.

For Frege, the fundamental kind of expression used to ascribe numbers to things are expressions like "the number of cards on the desk" or "the number of planets in the solar system." The expression "the number of Fs" is a singular term, purporting to pick out an object, in just the same way as "the brother of John" is a singular term, purporting to pick out a certain individual. So for Frege, ascriptions of number depend for their truth on the existence of objects, which are the referents of expressions of the form "the number of Fs."

Is it legitimate to suppose that given any general term F, there is also an object corresponding to the "the number of Fs"? Frege held that it is legitimate to speak of objects of a certain kind, provided there is a criterion of identity for them. What is the criterion of identity for numbers? The answer is given by the following principle, known as Hume's principle: "The number of Fs = the number of Gs if and only if (iff) there is a one-to-one correspondence between the Fs and the Gs."

A one-to-one correspondence is a relation that pairs each F with exactly one G and each G with exactly one F. So, for example, the number of knives on the table is equal to the number of forks provided that each fork can be paired with a unique knife and each knife with a unique fork.

Frege demonstrated that all the Dedekind-Peano axioms for number theory can be proved from Hume's principle alone, given appropriate definitions that he devised; a fact now known as Frege's theorem.

Frege attempted to go further by giving an explicit definition of "the number of Fs," from which Hume's principle itself could be proved. He defined "the number of Fs" as the set of all properties that can be put in one-to-one correspondence with the Fs. That is, the number *n* is identified with the extension of the second-order property of having *n*-members.

This was a disaster. The principle concerning sets that Frege appealed to in his derivation of Hume's principle states that every predicate has an extension. The extension of a predicate is the set of all (and only) those objects that satisfy the predicate. As Russell's paradox shows, however, this principle is inconsistent. If every predicate has an extension, then the predicate "is not a member of itself" has an extension, which would be the set of all (and only) the objects that are not members of themselves. Call this set R. It follows that R is a member of R iff R is not a member of R, a contradiction. Frege's logical system had turned out to be inconsistent. This was the first of a number of paradoxes of set theory that were to have a formative influence on subsequent work in the foundations of mathematics.

There were varying responses to Russell's paradox. Russell and Alfred North Whitehead took one approach: the theory of types. Ernst Zermelo and others took a different approach: that of axiomatic set theory. Given the now standard axioms for set theory, the Frege-Russell definition of the numbers will not work; the assumption that there is a nonempty set of all three-membered sets, for example, leads to a contradiction. A different

approach to the definition of the numbers is required. Instead of taking the numeral n to refer to the set of all n-membered sets, it can be taken to refer to some particular, paradigm example of an n-membered set.

John von Neumann provided an effective sequence of paradigm n-membered sets. The number zero is the paradigm zero-membered set: the empty set, $\emptyset$. The number 1 is the set whose only member is zero. The number 2 is the paradigm two-membered set whose members are 0 and 1. And in general, each number n is the n-membered set whose members consist of all and only the whole numbers from 0 up to $(n - 1)$.

One can then say that there are n members of a particular set iff that set can be placed into a one-to-one correlation with the paradigm n-membered set. For instance, there are two decks of cards on the table iff the members of the set of decks of cards on the table can be placed into a one-to-one correspondence with the members of the paradigm two-membered set $\{\emptyset, \{\emptyset\}\}$.

PHILOSOPHIES OF NUMBER

Philosophical accounts of number (and mathematics more generally) can be divided into two broad categories: realist and antirealist.

A realist about number holds that statements concerning numbers are objectively true or false. On this view, statements such as "there are nine planets in the solar system," "there are infinitely many prime numbers," "34957 + 70764 = 105621," or "every even number greater than two is the sum of two primes" (Goldbach's conjecture), say something that is objectively either true or false, even if no one knows which it is. In addition, the realist claims that some such statements are indeed true. That is, the realist typically accepts as true most, or all, of accepted mathematics.

By contrast, an antirealist denies one or both of the two realist claims. That is, the antirealist will deny that there is an objective fact of the matter about the truth value of all statements concerning number or that all currently accepted mathematical statements concerning number are actually true.

THE ARGUMENT FOR PLATONISM

Platonism, as that term is used in modern philosophy of mathematics, is the view that mathematics is the study of an objective realm of independently existing objects. In addition, the platonist holds that these objects are abstract, rather than physical objects. A physical object is something that (if it exists) has a location in space and time, can undergo changes of state, and can interact causally with other spatiotemporally located objects. Cups and saucers, stars and planets, plants and animals, and atoms and photons are all examples of physical objects. By contrast, an abstract object is something that (if it exists) lacks some or all of these properties. Abstract objects have no location in space and time, they have no state and no history, and they do not interact causally with other objects.

The main philosophical argument for platonism in modern philosophy proceeds as follows. Many statements of arithmetic appear to make existential claims. For example, the statement "there is a prime number greater than three" asserts the existence of an object having certain properties. Since many such arithmetical statements are true, it follows that numbers exist. This does not yet show that numbers must be abstract, but various arguments can be given against the alternatives. For example, every physical object has a location in space and exists for a certain time. Numbers have neither of these properties. Then again, there are infinitely many numbers, but perhaps a finite number of physical objects. It follows that numbers, if they exist, must be nonphysical, abstract objects.

The argument for platonism can be summarized as follows:

> P1. Arithmetical sentences express statements that are objectively true or false
>
> P2. Some arithmetical statements are true
>
> P3. Arithmetical statements quantify over certain objects (numbers)

Therefore:

> C1: Numbers exist.

However:

> P4: Numbers, if they exist, must be abstract (nonphysical, nonmental) objects.

Therefore:

> C: Numbers are abstract objects.

THE EPISTEMOLOGICAL PROBLEM

The central problem facing a platonist philosophy of number is epistemological. Abstract objects cannot be directly perceived, nor can they have any effects on objects or processes that can be directly perceived. How then is it possible for us to know anything at all about such objects? The causal isolation of abstract objects

appears to make them unknowable. Hence either we have no mathematical knowledge, or platonism is not the correct view of mathematics.

Before Frege, philosophers had postulated that human beings have some kind of direct cognitive access to mathematical objects, through perception or some rational faculty analogous to perception, or (for Immanuel Kant) through an *a priori* intuition or construction.

According to Frege our our only access to numbers is through our knowledge of the truth-values of arithmetical statements. Certain sentences of our language contain terms standing for numbers and quantifiers that range over numbers. If it can be shown that some of those sentences are true (and Frege hoped to show that they are logically true), then we will have explained how we can know about numbers, even though we have no direct perceptual or causal contact with them.

If the reduction of arithmetic to logic could be carried out, our knowledge of numbers would have been shown to be based on our knowledge of the truth of the basic laws of logic. Frege thought there was no real problem about how we know that the laws of logic are true; we can just see that they are. Explaining the psychological mechanisms that give us this ability is outside the scope of philosophy and can be left to the psychologists. The discovery of the paradoxes ruined this comfortable picture, showing that we have no infallible insight into the fundamental truths of logic after all. The reduction of arithmetic to set theory does not resolve this problem. Our knowledge of the basic laws of set theory cannot be any more certain or secure than our knowledge of the fundamental laws of arithmetic.

REALIST ALTERNATIVES TO PLATONISM

In view of the epistemological problem for platonism, many philosophers of mathematics have sought to avoid the conclusion that numbers are abstract objects. However, if that conclusion should be rejected, the argument for platonism given earlier must be unsound. Alternatives to platonism can be usefully classified according to which of the premises of that argument are rejected.

An obvious point at which that argument might be attacked is at premise P4. That is, one could deny that numbers, if they exist, must be abstract. Along these lines are various attempts to provide a physicalist account of mathematical objects such as numbers and sets. Such accounts differ from platonism only in denying that numbers are entirely nonphysical and abstract. The payoff is epistemological. If, for example, numbers are properties or relations that can be instantiated by ordinary physical objects, then some basic knowledge of numbers could be acquired by ordinary perception.

Another realist alternative is to accept P1 and P2, but deny P3; the claim that mathematical statements quantify over a domain of special objects of some kind. One strategy is to think of arithmetic as the theory, not of a special realm of objects, but of a certain pattern or structure. In the case of arithmetic the structure in question is that shared by any infinite progression of objects (also called an ω-sequence) in which (1) there is a unique first element and (2) for any given element there is a distinct, unique next element in the sequence, called the successor of the given element.

According to one variety of structuralism, the truths of arithmetic are simply those that hold in every system of objects that form an ω-sequence. An equation such as $2 + 1 = 3$ is interpreted as elliptical for the generalization; "If S is any system of objects that form an ω-sequence, then the successor of the successor of the first element of S added to the successor of the first element of S is equal to the successor of the successor of the successor of the first element of S."

Structuralism is often motivated by a certain ontological problem for platonism. According to the platonist, sets and numbers are abstract objects. What is the relationship between them? We have already described one way in which the natural numbers can be defined as sets. This is the definition of the natural numbers as the von Neumann numbers: $0 = \emptyset$, $1 = \{0\}$, $2 = \{0,1\}$, and $3 = \{0,1,2\}$ and in general, $N+1 = \{0, 1, ..., N\}$.This is not the only possible set-theoretic definition of the natural numbers, however. Zermelo, for example, defined the sequence as follows: $0 = \emptyset$, $1 = \{0\}$, $2 = \{1\}$, and $3 = \{2\}$ and in general, $N+1 = \{N\}$.

From a purely mathematical point of view the definitions seem equally valid, since they both validate exactly the same theorems of arithmetic. However, the two definitions are certainly not equivalent, since they identify some numbers with distinct sets; on von Neumann's definition $2 = \{\emptyset, \{\emptyset\}\}$ (a set with two members), while on Zermelo's definition $2 = \{\{\emptyset\}\}$ (a set with just one member).

From a platonist perspective there is something puzzling about this. If numbers are independently existing objects, then there must be a fact of the matter about which set, if any, the number 2 is identical with. It cannot

be that there are two equally correct definitions of the number 2 that identify it with different sets, but this is exactly what one seems to have in the case of the von Neumann and Zermelo definitions: two equally correct accounts of the number 2 that assign it to distinct sets.

A generalization of this line of argument yields the conclusion that numbers cannot be objects of any kind. Any definition of numbers in terms of particular mathematical objects of some other kind is arbitrary, in the sense that equivalent, but distinct alternative definitions will always be available. However, if, as the platonist holds, numbers are objects, there must be a fact of the matter about which objects the numbers really are. So there must be something wrong with platonism.

From a structuralist perspective, however, this kind of ontological relativity is readily explicable. For on that account, arithmetic is not concerned with a domain of specific objects, but only with what holds good in all ω-sequences. The sequences of sets defined by von Neumann and Zermelo are both examples of ω-sequences, so both systems have the required structure. Any system of objects (sets or otherwise) having the same structure will do just as well, for arithmetic is just the theory of the properties shared by all ω-sequences.

CONVENTIONALISM

A different approach, long popular with empiricists, is to say that mathematics is concerned not with objects, but with relations between concepts. A good example is the account of mathematics associated with the philosophical movement known as logical positivism, which had its heyday in the 1930s and 1940s. According to the positivists, the truths of logic and mathematics are alike in being analytic, by which they meant that they are true solely in virtue of the meanings of the symbols they contain, meanings that are established by linguistic convention.

On this view, $2 + 2 = 4$ is true because of the stipulations we have laid down governing the use of the symbols "2," "4," "+," and "=." As such it is completely without empirical content and this explains the irrelevance of empirical evidence to mathematics. No fact about the world can contradict the statement $2 + 2 = 4$, because its truth does not depend on facts about the world, but only on facts about what the mathematical symbols occurring in it mean. What our symbols mean is a matter of arbitrary linguistic convention. We can simply stipulate that our symbols are going to have certain meanings and then the truth of various statements involving them will follow. The truths of arithmetic on this view are records of the stipulations we have laid down governing the use of the arithmetical symbols.

Largely as a result of criticisms developed by Quine and others, conventionalism is no longer widely accepted. One difficulty is that even if it were possible simply to stipulate that the terms of a mathematical theory are to be assigned whatever meaning makes all the axioms turn out true, the stipulation will backfire if the axioms are inconsistent. Whether the axioms are consistent or not is itself a mathematical fact which is independent of our stipulations and conventions. If so, then not all mathematical facts can be purely conventional or true in virtue of meaning.

The specific objections to conventionalism are however, less significant than the alternative account of the epistemology of mathematics developed by Quine, which if correct, would undermine the main epistemological motivation for conventionalism. Quine's alternative account is described in the final section of this article.

NOMINALISM

Nominalism is the philosophical thesis that there are no abstract objects that is, everything that exists is a concrete, physical particular. In interpreting mathematics and science then, the nominalist has two options. One option is to say that despite appearances, mathematical and scientific theories do not involve reference to abstract objects after all. The other option is to say that they do and are therefore literally false. The nominalist may then seek to provide a positive account of mathematical and scientific theories, showing how they can be reformulated so as to avoid any reference to abstract objects. The result would be an error theory of science and mathematics.

This second approach is the one taken by the nominalist philosopher Hartry Field, who has attempted to demonstrate that reference to abstract objects can be eliminated from science by showing how nominalistic versions of physical theories might be constructed: versions which do not presuppose the existence of abstract objects such as numbers or functions. The interested reader is referred to the bibliography for further details of the construction and the philosophical debate surrounding it.

FORMALISM

This type of nominalist antirealism concerning numbers and other abstract objects consists in a denial only of the second premise (P2) of the argument for platonism given earlier. On such a view, mathematical sentences express

statements which can be true or false, but it is argued that there is no good reason for thinking that any mathematical statements involving quantification over abstract objects are literally true.

More radical types of antirealism deny the first premise (P1) of the argument for platonism. One way of denying that premise is to say that mathematical sentences do not express statements that could be true or false at all, objectively or otherwise. This is the approach taken by the formalist account of mathematics. On this view, mathematics is not a body of statements that can be true or false. Instead, mathematics is thought of as analogous to a game, like chess. It is a game played with symbols according to certain rules.

The most sophisticated version of the formalist account of mathematics is that proposed by the mathematician David Hilbert in the 1920s and early 1930s. According to Hilbert mathematics has a meaningful part and a purely formal part. The meaningful part consists of finitary statements. These are decidable statements concerning only perceptible concrete symbols, such as the numerals 0, S0, SS0, SSS0. The purely formal component consists of ideal statements, statements that involve unbounded quantification over infinite domains such as the natural numbers. All such statements are strictly meaningless, according to Hilbert. Their introduction into mathematical theories was to be justified on purely instrumental grounds. They provide the mathematician with an extremely powerful, but in principle dispensable, means of proving facts about the real finitary subject matter of mathematics.

Hilbert's program was to show, using only finitary methods, that the introduction of such ideal statements into arithmetic could never lead to any false finitary statement becoming derivable. This is equivalent to proving using only finitary methods that classical arithmetic is consistent. He hoped to establish the same result for set theory, thereby establishing that the threat of inconsistency implied by the paradoxes could be guaranteed not to arise there either. "No one," wrote Hilbert, "shall drive us out from the paradise that Cantor has created for us" (Benacerraf and Putnam [1983], p. 191).

There is a fairly broad consensus that Kurt Gödel's second incompleteness theorem shows that Hilbert's program is unachievable, even at the level of arithmetic. Let T be any standard formal system for arithmetic. Suppose there was a finitary consistency proof for T. Then that proof could be formalized as a derivation in T of a formula expressing the consistency of T. However, by Gödel's second incompleteness theorem, no consistent formal system for arithmetic can contain such a derivation. It follows that the goal of Hilbert's program is unachievable for arithmetic and so also for set theory.

INTUITIONISM

A different response to the paradoxes, current at the time Hilbert was writing, was the intuitionist account of mathematics, proposed by the mathematician Luitzen Egbertus Jan Brouwer and developed by Arend Heyting. Intuitionism can be thought of as denying the first premise of the argument for platonism by claiming that although mathematics does constitute a body of statements that can be true or false, the truth or falsity of a mathematical statement is not independent of human beings.

The platonist thinks of the natural numbers as an infinite domain of objects that exist independently of human thought and that make arithmetical statements objectively true or false. By contrast, intuitionists such as Brouwer and Heyting think of the natural numbers as mental constructions, objects that are created by the human mind. On this view, what makes a mathematical statement true or false is not the existence of objects that are independent of human beings, but the existence of a certain kind of mental construction, a proof (though not a proof in a formal system).

This conception of mathematical truth led the intuitionists to reject the law of excluded middle, as applied to mathematics. That is, they denied the universal validity of the logical schema "Either A or not-A." For on the intuitionist view, a mathematical conjecture for which neither proof nor disproof has yet been constructed, cannot be said to be either true or false.

Although paradox may be avoided in the intuitionistic reconstruction of mathematics, many contemporary philosophers would reject it. One reason is the apparent truncation of classical mathematics necessitated by intuitionism; many theorems of classical analysis and set theory are false when interpreted intuitionistically. A deeper reason may be a distrust of the reforming nature of the intuitionism. The role of philosophy, it is thought, should be to provide an account or interpretation of mathematics as it actually is, not to reformulate or remake mathematics in a new image.

An exception to this general trend is Michael Dummett. A widely accepted philosophical thesis has it that the meaning of a statement is given by its truth conditions. Dummett argues that this is empty, unless accompanied by a substantive account of truth; an account

which goes beyond the mere equivalence of 'P is true' with P. But meaning cannot be explained in terms of a concept of truth according to which truth is something that may apply to a statement quite independently of whether it is possible to know that it does, for then our knowledge of the meaning of a statement would not always be capable of being made manifest by publicly observable behaviour—a condition which is necessary for meaning to be communicable. Instead, meaning must be explained in terms of verification conditions; to know the meaning of a mathematical statement is to know what would count as a proof of it. Thus the argument leads to a version of intuitionism; to say that a mathematical statement is true is to say that we have a proof, while to say that it is false is to say that we have a disproof.

Dummett's argument depends only on very general considerations concerning the communicability of meaning. If valid, the argument would apply to statements of any kind whatsoever and not just to mathematical ones. For example, it would be a consequence that perfectly ordinary statements about the past which can no longer be verified or refuted could not be considered either true or false, unless it could be shown that there is some special feature of our use of such statements which makes a verification transcendent account of their meaning possible.

THE INDISPENSABILITY ARGUMENT

The epistemological objection to platonism is one aspect of a more general problem for empiricism. Mathematics appears to be highly non-empirical, in both its subject-matter and its methodology. Empirical evidence does not appear relevant to mathematics. However if, as the empiricist asserts, all our knowledge is ultimately empirical, there seems to be no good reason for thinking that mathematics is true at all. The logical positivist's claim that mathematics is analytic, or true only in virtue of meaning, was an attempt to solve this problem by showing how mathematical statements could be true, though independent of all empirical evidence.

In a now classic series of papers written in the late 1940s and early 1950s, Quine launched a major critique of this conventionalist solution to the problem, while also developing a significant alternative account of the structure of empirical knowledge and the place of mathematics within that structure. Quine argued that the mistake made by earlier empiricists was to think that individual statements can be tested empirically in isolation from each other. Instead, it must be recognized that our scientific beliefs form an interlocked system or web that "faces the tribunal of experience as a corporate body." (Quine, 'Two Dogmas of Empiricism', in 'From a logical point of view', p. 41). Mathematical beliefs form an indispensable part of this system and are therefore justified to the extent that they contribute to the goals of scientific prediction and explanation. In Quine's view the mathematics used in a successfully confirmed scientific theory is confirmed along with the rest of that theory. Mathematical objects, like numbers and sets, are theoretical posits, epistemologically on a par with electrons and photons.

Quine draws a further conclusion. If mathematics can be supported by empirical evidence, it can also be undermined by it. Our mathematical beliefs are open to empirical falsification and revision, in just the same way as our scientific beliefs. The illusion of a difference between mathematical and other scientific statements is generated, according to Quine, by pragmatic considerations. We are far more reluctant to revise the mathematical and logical components of our scientific theories because these are so deeply embedded in the system of total science that altering them would result in a major restructuring of the entire system. But if the result of such a restructuring was an overall simplification or improvement in the total system of science, then it would be perfectly rationally justified.

Quine's epistemology is significant because it provides a solution, consistent with empiricism, to the epistemological problem for platonism. Quine can accept all the premises in the argument for platonism given earlier. Mathematical statements can be taken to refer, as they appear to refer, to abstract objects such as numbers and sets. But the epistemological problem is resolved. Numbers and sets cannot be perceived, either directly or indirectly, but their utility in enabling us to predict and explain the world provides us with all the justification for their existence we need or could ever be entitled to.

The indispensability of mathematics in science has two aspects: one emphasized by Quine, the other by Hilary Putnam. Quine argues from the indispensability of mathematics in the derivation of the observation statements that confirm or disconfirm scientific theories and hypotheses. Putnam emphases a different aspect of the indispensability of mathematics in science. Mathematics is used in science, not only in deriving predictions from theories but also in formulating the empirical hypotheses of those theories. Consider Boyle's law, for example, which states the relationship between the pressure, temperature, and volume of a fixed quantity of gas enclosed in some container. The law states that the pressure of the

gas is equal to a constant multiplied by the temperature of the gas, divided by the volume.

$$P = kT / V$$

Pressure, temperature, and volume are all numerical quantities. The pressure of the gas in kilopascals at a certain time is a real number, as are the volume in cubic centimeters and the temperature in degrees Celsius at a time. The law states that a certain mathematical relationship always holds between these real numbers. Boyle's law is therefore just as much committed to the existence of real numbers and functions as it is to the existence of gases. In this way, realism about physical theory leads to realism about mathematical objects such as numbers.

The Quine-Putnam argument allows for the empirical justification not only of the often highly specific mathematical statements (such as numerical equations) used to derive predictions from a theory but also of any more general mathematical statements, such as set-theoretic axioms, which imply them. Boyle's law can be derived from the more fundamental laws of thermodynamics. Hence any empirical confirmation of Boyle's law accrues also to the thermodynamic laws used to derive it. In just the same way, since arithmetic can be reduced to set theory, the numerical equations used to derive predictions from Boyle's law can be derived from the axioms of set theory. Hence any empirical confirmation of those equations accrues also to the axioms of set theory. In this way, it might be hoped that a great deal of even abstract mathematics can be justified by means of the Quine-Putnam argument.

See also Frege, Gottlob; Mathematics, Foundations of; Platonism and the Platonic Tradition; Quine, Willard Van Orman; Russell, Bertrand Arthur William.

Bibliography

On the historical development of the mathematics of number, see Morris Kline, *Mathematical Thought from Ancient to Modern Times* (New York: Oxford University Press, 1972).

For Euclid's *Elements*, the standard reference is T. L Heath, *The Thirteen Books of Euclid's Elements* (New York: Dover Publications, 1956).

Collections of mathematical papers by Frege, Cantor, Gödel, Zermelo, and others can be found in D. E. Smith, ed. *A Source Book in Mathematics* (New York: Dover Publications, 1959) and J. van Heijenoort, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931* (Cambridge, MA: Harvard University Press, 1967).

For the philosophy of mathematics, the standard collection of papers is Paul Benacerraf and Hilary Putnam, eds. *Philosophy of Mathematics: Selected Readings* (2nd ed. New York: Cambridge University Press, 1983). Also useful is W. D. Hart, ed. *The Philosophy of Mathematics* (New York: Oxford University Press, 1996).

Frege's classic *Die Grundlagen der Arithmetic* (Breslau, Poland: Koebner, 1884) is translated by J. L Austin as *The Foundations of Arithmetic* (Oxford, U.K.: Basil Blackwell, 1953).

For philosophical expositions of Frege's work, see Michael Dummett, *Frege Philosophy of Mathematics* (London: Duckworth, 1991) and Crispin Wright, *Frege's Conception of Numbers as Objects* (Aberdeen, Scotland: Aberdeen University Press, 1983). W. Demopoulos, ed. *Frege's Philosophy of Mathematics* (Cambridge, MA: Harvard University Press, 1995) provides an excellent collection of papers on the technical aspects of Frege's work.

On the epistemological problem for platonism, see Paul Benacerraf, "Mathematical Truth," *Journal of Philosophy* 70 (1973): 661–679.

For a survey of work on physicalist accounts of mathematical, see A. D. Irvine, ed. *Physicalism in Mathematics*. Dordrecht, Netherlands: Kluwer Academic, 1990.

On structuralism and the ontological problem for platonism, see Paul Benacerraf, "What Numbers Could Not Be," *Philosophical Review* 74 (1965): 47–73; Hilary Putnam, "Mathematics without Foundations," *Journal of Philosophy* 64 (1965): 5–22; and Charles Parsons, "The Structuralist View of Mathematical Objects," *Synthese* 90 (1990): 303–346.

On the conventionalist account of mathematics and logic, see A. J. Ayer, *Language, Truth, and Logic* (London: V. Gollancz, 1946); Rudolf Carnap, "Empiricism, Semantics, and Ontology" (1956) and Carl Gustav Hempel, "On the Nature of Mathematical Truth," *American Mathematical Monthly* 52 (1954), both reprinted in Paul Benacerraf and Hilary Putnam, eds. *Philosophy of Mathematics: Selected Readings* (2nd ed. New York: Cambridge University Press, 1983).

Quine's critiques of conventionalism can be found in "Truth by Convention" and "Carnap and Logical Truth" in Quine, *Ways of Paradox and Other Essays* (Cambridge, MA: Harvard University Press, 1976).

For the indispensability argument, see Quine, "Two Dogmas of Empiricism" and "On What There Is" in his *From a Logical Point of View* (2nd ed. Cambridge, MA: Harvard University Press, 1980). See also Quine, *Word and Object* (Cambridge, MA: MIT Press, 1960).

Putnam's exposition of Quine's argument is to be found in Hilary Putnam, *Philosophy of Logic* (London: Allen and Unwin, 1971).

On Field's program, see Hartry Field, *Science without Numbers: A Defence of Nominalism* (Princeton, NJ: Princeton University Press, 1980). For a critical analysis, see Stewart Shapiro, "Conservativeness and Incompleteness," *Journal of Philosophy* 80 (1983): 521–531. A. D. Irvine, ed. *Physicalism in Mathematics* (Dordrecht, Netherlands: Kluwer Academic, 1990) also contains a number of relevant articles on nominalism.

David Hilbert's formalist program is outlined in a readable way in his "On the Infinite" (1926), which is reprinted in Paul Benacerraf and Hilary Putnam, eds. *Philosophy of Mathematics: Selected Readings* (2nd ed. New York: Cambridge University Press, 1983, pp. 83–201). A scholarly exposition of this and other foundational programs can be

found in Marcus Giaquinto, *The Search for Certainty: A Philosophical Account of Foundations of Mathematics* (Oxford, U.K.: Clarendon Press, 2002).

On intuitionism, see the papers by Brouwer and Heyting in Paul Benacerraf and Hilary Putnam, eds. *Philosophy of Mathematics: Selected Readings* (2nd ed. New York: Cambridge University Press, 1982). Michael Dummett's *Elements of Intuitionism* (Oxford, U.K.: Clarendon Press, 1977) is a standard reference work on intuitionistic mathematics and philosophy. Dummett's own argument for an intuitionistic interpretation of mathematics can be found in "The Philosophical Basis of Intuitionistic Logic" in his *Truth and Other Enigmas* (London: Duckworth, 1975).

John Bigelow (2005)
Sam Butchart (2005)

NUMENIUS OF APAMEA

Numenius of Apamea, the second-century Greek philosopher perhaps best known for his description of Plato as an Atticizing Moses, was a precursor of Plotinus and Neoplatonism and also had affinities with Gnosticism and the Hermetic tradition. Of his life practically nothing is known, and even the approximate dates of his birth and death are uncertain. Since his description of Plato is quoted by Clement of Alexandria (*Stromateis* i, 22.93), he cannot have survived much later than 200 CE, while the latest writers cited in the fragments of his works belong to the time of Nero (37–68 CE). He may have been of non-Greek origin, and his name, like that of Porphyry, may have been a Greek translation of a Semitic original. Our sources commonly describe him as a Pythagorean, but Iamblichus and Proclus call him a Platonist, which comes to much the same thing in an age when Plato was considered a disciple of Pythagoras. Certainly Numenius is best grouped with such Middle Platonists as Albinus. His work was based primarily upon exegesis of Plato and presents a systematization of Plato's thought with a dualist emphasis. It is possible that he had some knowledge of Christianity, but what is truly remarkable is his knowledge of Judaism. It has been suggested that he himself was a Jew, but this is far from certain. What is clear is that he sought to go back before Plato and Pythagoras to the teachings of the ancient East, the Brahmins, the Jews, the Magi, and the Egyptians. In this respect there are links with the Hermetic books and with the *prisca theologia* of such Renaissance writers as Marsilio Ficino and Giovanni Pico della Mirandola, although scholars differ as to the extent to which Numenius's philosophy was actually influenced by Oriental ideas and the extent to which it was purely Greek.

A notable feature of his thought is his doctrine of the Demiurge. He postulates two opposed principles, God and matter, the monad and the dyad, but whereas the Pythagoreans adhered to monism by making the dyad emanate from the monad, Numenius developed a dualistic theory. Matter is evil, and the supreme God can therefore have no contact with it; hence the need for a second god, the Demiurge, who is of dual nature, an *anima mundi* related both to God and to matter (cf. the Philonic Logos). There are also two souls in the world, one good and one evil, and two souls in man, a rational and an irrational; and the only escape from this dualism is by deliverance from the prison of the body. Astrological elements in Numenius's anthropology suggest an attempt to give astrology a rational basis.

Numenius is important for his influence on later Neoplatonists, although some of his views were to be rejected by them. The allegation that Plotinus merely plagiarized Numenius prompted Plotinus's disciple Amelius to write a book pointing out the differences between them (Porphyry, *Vita Plotini* 17). The hierarchy of three gods, for example, appears to be similar to Plotinus's hierarchy of being, but the three entities in each case do not correspond exactly in detail. Moreover, Plotinus rejected Numenius's dualistic and Gnosticizing tendencies.

See also Platonism and the Platonic Tradition; Pythagoras and Pythagoreanism.

Bibliography

FRAGMENTS

des Places, Édouard, ed. and tr. *Numénius, Fragments.* Paris: Belles Lettres, 1973. (Includes French translation.)

STUDIES

Armstrong, A. H. *The Architecture of the Intelligible Universe in the Philosophy of Plotinus.* Cambridge, U.K.: Cambridge University Press, 1940.

Beutler, R. "Numenios." In Pauly-Wissowa, *Real-Encyclopaedie.* Supp. VII (1940), pp. 664ff.

Burnyeat, Myles F. "Platonism in the Bible: Numenius of Apamea on *Exodus* and Eternity." In *Metaphysics, Soul, and Ethics. Themes from the Work of Richard Sorabji,* edited by Ricardo Salles. Oxford: Oxford University Press, 2004.

Dillon, John. *The Middle Platonists. A Study of Platonism, 80 B.C. to A.D. 220,* 2nd ed. London: Duckworth, 1996.

Dodds, Eric R. (1960). "Numenius and Ammonius." In *Les sources de Plotin = Entretiens sur l'Antiquité Classique.* Vol. 5. Geneva: Fondation Hardt, 1957.

Festugière, A. J. *La révélation d'Hermes Trismégiste.* Vol. III: *Les doctrines de l'âme.* Paris: Lecoffre, 1953. Vol. IV: *Le dieu inconnu.* Paris: Lecoffre, 1954.

Frede, Michael. "Numenius." In *Aufstieg und Niedergang der römischen Welt,* Vol. II.36.2, edited by H. Temporini and W. Haase. Berlin: de Gruyter, 1987.

Leemans, E. A. *Studie over den Wijsgeer Numenius van Apamea.* Mémoires de l'Académie Royale de Belgique. Vol. 37. Brussels, 1937. Includes fragments and testimonia.

Puech, H. C. *Mélanges Bidez.* Vol. II. Brussels, 1934.

Vogel, C. J. de. *Greek Philosophy.* Vol. III, 421ff. Leiden: Brill, 1959.

R. McL. Wilson (1967)
Bibliography updated by G. R. Boys-Stones (2005)

NUSSBAUM, MARTHA
(1947–)

Martha Nussbaum has contributed to ethics, political theory, classics, philosophy of mind, legal theory, educational theory, public policy, and gender studies. Educated at New York University (BA, 1969) and Harvard University (MA, 1971; PhD, 1975), she has taught at Harvard, Brown University, Oxford University, and the University of Chicago.

Nussbaum's work ranges widely, but she has consistently returned to such themes as: the nature of emotion and its role in philosophical argument, the extension and application of the "capabilities approach" in the theory of justice, the role of philosophical argument and reflection in the public sphere, and the relationship between philosophy and art and literature. Her work can be helpfully characterized as a sustained critique of Platonism. *The Fragility of Goodness* (1986), her first major book, argued that the Platonic view of the good life marks "an aspiration to rational self-sufficiency through the 'trapping' and 'binding' of unreliable features of the world." Such self-sufficiency omits "a kind of human worth that is inseparable from vulnerability, an excellence that is in its nature other-related and social, a rationality whose nature it is *not* to attempt to seize, hold, trap, and control, in whose values openness, receptivity, and wonder play an important part" (pp. 19–20).

Nussbaum has consistently defended the latter. Against the Platonic-Christian view that transcendent Good or God is at the heart of morality, she advances her own comprehensive, Aristotelian-Kantian-Jewish view that religion highlights the largely autonomous, primary domain of human moral effort. The highest moral paradigms are not such figures as the saints or Gandhi, but those who, like Nehru, found the good life in human finitude and limitation. For Nussbaum, rigorist or ascetic moralism, whether in Gandhi or Plato, betrays a violence toward the self that may undermine morality and compassion.

Upheavals of Thought: The Intelligence of Emotions (2001) develops the moral psychology that figures in Nussbaum's ethical and political work. The Platonic ascent of love is criticized for having the lover climb to such heights as to be beyond compassion and human need, beyond even altruistic contact with actual human beings. Christian and Romantic views fail in the same way, and can reinforce developmental tendencies positively inimical to morality—childhood emotions of shame, disgust, and envy. Nussbaum works out a highly qualified "neo-Stoic" view of the emotions, according to which "once one has formed attachments to unstable things not fully under one's control, once one has made these part of one's notion of one's flourishing, one has emotions of a background kind toward them—on my view, judgments that acknowledge their enormous worth—that persist in the fabric of one's life, and are crucial to the explanation of one's actions" (p. 71). Thus, emotions are a type of evaluative judgment, construed in a way broad enough to allow that nonhuman animals and infants, who lack propositional thought, can also be said to have emotions. And they have a narrative structure, found in one's life history. Acknowledging one's neediness, however, and representing the world from the personal point of view and with considerable ambivalence, the emotions so characterized pose problems for moral and political theories stressing mutual respect, dignity, and concern for others.

Nussbaum's account of such emotions as compassion, shame, and disgust, which also receive extended treatment in her *Hiding from Humanity* (2004), is vital for understanding her political philosophy, which draws heavily on Aristotle, Immanuel Kant, John Stuart Mill, Karl Marx, John Rawls, and Amartya Sen. She defends a broadly Rawlsian political liberalism that frames an account of human flourishing adapted to the demands of liberal political theory, respecting the reasonable plurality of views of the good life to be found in the modern world. Her collaboration with Sen, beginning with *The Quality of Life* (1993), has yielded a critique of conventional economic measures of human welfare and pointed up the virtues of instead measuring people's capabilities, what they are capable of doing or being across central areas of human life. Her aim has been to bring her Aristotelianism into harmony with the capabilities approach, adapted to serve as a form of political liberalism that could also undergird the type of universalistic critique required by feminism.

Nussbaum's development of the capabilities approach in connection with feminism has led her to introduce more Kantian and Millian elements into her arguments and to emphasize the recognition of human dignity as a core feature of political liberalism. *Sex and Social Justice* (1999) and *Women and Human Development* (2000) develop the capabilities theory as the philosophical groundwork for basic constitutional standards, applicable to all governments, defining the minimal requirements of respect for human dignity. These works provide a highly developed account of the central human capabilities—life, bodily health, bodily integrity, senses, imagination and thought, emotions, practical reason, affiliation, concern for nature and other species, play, and political and material control over one's environment—and articulate the political liberal demand that all citizens must, as a requirement of justice, enjoy a basic threshold level of each of these capabilities. Her focus on the injustices confronting women, gays, and lesbians, and others suffering from insidious forms of oppression, has widened to cover problems of international justice and justice with respect to nonhuman animals.

Nussbaum has also paid special attention to education. *Cultivating Humanity* (1997) argues for an education (inspired by Plato's earlier, truly Socratic dialogues) that would awaken students to self-scrutiny and to their capabilities for love and imagination. Promoting a greater role for such philosophical reflection in public life has been one of Nussbaum's chief priorities.

See also Aristotelianism; Aristotle; Feminism and the History of Philosophy; Feminist Philosophy; Justice; Kant, Immanuel; Marx, Karl; Mill, John Stuart; Plato; Platonism and the Platonic Tradition; Rawls, John; Sen, Amartya K.; Women in the History of Philosophy.

Bibliography

WORKS BY MARTHA NUSSBAUM

The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy. New York: Cambridge University Press, 1986.

The Quality of Life, edited by Martha Nussbaum and Amartya Sen. New York: Oxford University Press, 1993.

The Therapy of Desire: Theory and Practice in Hellenistic Ethics. Princeton, NJ: Princeton University Press, 1994.

Cultivating Humanity: A Classical Defense of Reform in Liberal Education. Cambridge, MA: Harvard University Press, 1997.

Sex and Social Justice. New York: Oxford University Press, 1999.

Women and Human Development: The Capabilities Approach. New York: Cambridge University Press, 2000.

Upheavals of Thought: The Intelligence of Emotions. New York: Cambridge University Press, 2001.

Hiding from Humanity: Disgust, Shame, and the Law. Princeton, NJ: Princeton University Press, 2004.

Frontiers of Justice: Disability, Nationality, Species Membership. Cambridge, MA: Harvard University Press, 2005.

WORKS ABOUT MARTHA NUSSBAUM

Goodin, R., and D. Parker, eds. "Symposium on Martha Nussbaum's Political Philosophy." *Ethics* 111 (October 2000). Comprehensive essays by L. Antony, R. Arneson, H. Charlesworth, and R. Mulgan, with responses by Nussbaum.

R. Barton Schultz (2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

OAKESHOTT – PRESUPPOSITION

7 volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON
GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

ENCYCLOPEDIA OF
PHILOSOPHY
2nd edition

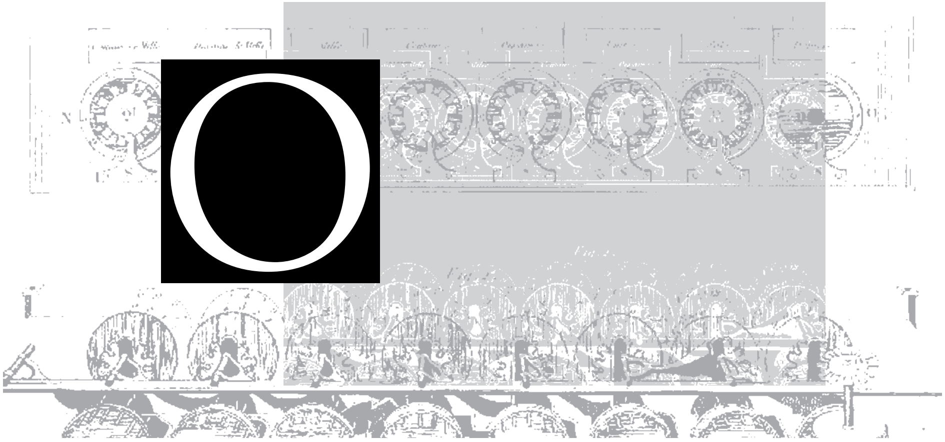

OAKESHOTT, MICHAEL

(1901–1990)

Michael Oakeshott, a wide-ranging thinker mostly known for his work in social and political philosophy, was born in Chelsfield, Kent, on December 11, 1901. Oakeshott read history at Gonville and Caius College, Cambridge, and graduated in 1923. He returned as a fellow in 1925. In 1940 he enlisted in the British Army and served with "Phantom," an intelligence unit that worked on artillery spotting. In 1949 he went to Oxford as a fellow of Nuffield College and in 1951 he was appointed to the chair of political science at the London School of Economics. He retired in 1969, but continued to be active from his retirement home in Acton, Dorset, where he died on December 18, 1990.

EXPERIENCE AND ITS MODES

Experience and Its Modes (1933) was Oakeshott's first major work. In the book Oakeshott creates some of the major distinctions that mark his social/political philosophy. The most important concerns experience itself. Influenced by the holism of Plato and Hegel (especially the *Phenomenology of Spirit*) and the idealism of Francis Bradley (*Appearance and Reality*), Oakeshott posits that "experience is a single whole, within which modifications may be distinguished, but which admits of no final or absolute division; and that experience everywhere, not merely is inseparable from thought, but is itself a form of thought" (1933, p. 10). Within the unity of experience people attempt to make sense of it via interpretative devices such as "history research," "scientific experimentation," and "practical reasoning." But all of these paths will ultimately fail. This is demonstrated by a relentless skepticism. The futile interpretative modes rely upon a false understanding of the primacy of Enlightenment-style rationalism. Instead, the agent finds herself in the midst of her own reflections and poetic imaginings. This agent-centered construction creates a tension in a world of other minds. The result is a necessary travail to reconcile one's own experience with that of others. This process is necessary to make social existence coherent.

Along with this amalgam of skeptical idealism Oakeshott posits freedom:

> The starting place of doing is a state of reflective consciousness, namely, the agent's own understanding of his situation, what it means to him. And, of course, it is no less *his* situation even though it may be a concern with what he understands to be the situation of another or of oth-

> ers. … And it is in this respect of this starting-place in an understood contingent situation that the agent in conduct may be said to be "free." (1975, p. 37)

Freedom is thus one of the properties of consciousness that allows the interpretative awareness of consciousness to develop.

Because freedom is a precondition of people's experience of the world, it is vain for totalitarian dictators to endeavor to suppress it. To do so would mean that the dictator tries to suppress an aspect of human nature that underlies the possibility of human experience. It just can't happen. Freedom will exhibit itself in one form or another. This is not a teleological expression of human nature but rather an indication that people will interpret and respond to what life presents them. This is a concrete and practical vision. Though some may be drawn to the modes to make sense of it all (a vain endeavor), the primary imperative (á la Berkeley) is first to accommodate the primary data of experience as it presents itself: "And no matter how far we go with it, we shall not easily forget the sweet delight which lies in the empty kisses of abstraction" (1933, p. 356).

RATIONALISM IN POLITICS

The essays in *Rationalism in Politics* (1991) form the core of Oakeshott's social/political thought. In the title essay Oakeshott extends some of the concepts of his earlier work to critique Enlightenment rationalism as a device that is serviceable for guiding social and political thinking. He proclaims this Hobbesian skepticism of rationalism as a useful tool for politics in language that is reminiscent of Aristotle (*EN* I.1).

> Every science, every art, every practical activity requiring skill of any sort, indeed every human activity whatsoever, involves knowledge. And universally, this knowledge is of two sorts. … The first sort of knowledge I will call technical knowledge or knowledge of technique. … The second sort of knowledge I will call practical, because it exists only in use, is not reflective and (unlike technique) cannot be formulated into rules. (1991, p. 12)

This essay then goes on to evaluate these two aspects of reason with a critique of traditional accounts that aspire to make rationalism a transcendent tool. Instead, Oakeshott insists, reason is merely the handmaiden of free holistic experience.

In "The Tower of Babel" Oakeshott sets out a Hegelian understanding of the existing community and its proper influence on the individual. Two sorts of morality are posited: The first represents the existing moral community (akin to the German *Sittlichkeit*). The second is a philosophical critique that may alter the first. Alan Donagan contends that Oakeshott (like Hegel) misses the force of deontological commands by favoring the *Sittlichkeit* over *Moralität*. By being biased toward experience, as such, Donagan believes that fundamental principles that supercede morality are not given their due. The mere existence of the second (philosophical) form of morality is not adequate. This much resembles the Kant-Hegel debate on the proper place of experience in evaluating the moral community. Oakeshott's position of affirming the existing moral community puts him into the camp of political conservatism. How much one is to make of this is still a subject of critical debate.

"The Voice of Poetry in the Conversation of Mankind" is another key essay in the collection that proclaims an aesthetics that is disinterested. It is not for the sake of instruction nor is it a conscious imitation of nature. "The poet does not recognize and record natural or conventional correspondencies or use them to 'explore reality'; he does not invoke equivalencies, he makes images" (1991, p. 528). In this way, the work of art is for the sake of the pleasurable contemplation of images. In some ways Oakeshott's aesthetic stance is reminiscent of Schiller and some readings of Kant. It is consistent with the holism standpoint that was established in *Experience and its Modes*.

CONCLUSION

Michael Oakeshott may be best known as a conservative political writer in the tradition of Hobbes. However, as the comments above suggest, he is more than that. He grounds his thinking in a comprehensive epistemological theory that also supports other explorations (such as aesthetics, history, and education). To evaluate his work, it is important to view Oakeshott within this larger context.

See also Aristotle; Berkeley, George; Bradley, Francis Herbert; Enlightenment; Epistemology, History of; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Idealism; Kant, Immanuel; Plato; Rationalism; Social and Political Philosophy.

Bibliography

MAJOR WORKS BY OAKESHOTT

Experience and Its Modes. Cambridge, U.K.: Cambridge University Press, 1933. Reprinted in 1966; paperback edition 1986.

The Social and Political Doctrines of Contemporary Europe. Cambridge, U.K.: Cambridge University Press, 1939.

Rationalism in Politics and Other Essays. London: Methuen, 1962. Revised by Timothy Fuller. Indianapolis: Liberty Press, 1991.

On Human Conduct. Oxford: Clarendon Press, 1975. Paper ed., Oxford: Clarendon Press, 1990.

On History and Other Essays. Oxford: Basil Blackwell, 1983.

The Voice of Liberal Learning: Michael Oakeshott on Education, edited by Timothy Fuller. New Haven, CT: Yale University Press, 1989.

"Introduction" to Thomas Hobbes, *Leviathan*. Oxford: Basil Blackwell, 1946. Included in *Rationalism in Politics*, 1991.

A complete bibliography to 1968 is available in *Politics and Experience: Essays Presented to Professor Michael Oakeshott on the Occasion of his Retirement*, edited by Preston King and B. C. Parekh. (Cambridge, U.K.: Cambridge University Press, 1968). An online bibliography is available from the Michael Oakeshott Association (available from www.michael-oakeshott-association.org).

WORKS ABOUT OAKESHOTT

Auspitz, Josiah Lee. "Individuality, Civility, and Theory: The Philosophical Imagination of Michael Oakeshott." *Political Theory* 4 (3) (1976): 261–294.

Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1977.

Flathman, Richard E. *The Practice of Political Authority*. Chicago: University of Chicago Press, 1980.

Franco, Paul. *The Political Philosophy of Michael Oakeshott*. New Haven, CT: Yale University Press, 1990.

Fuller, Timothy. "Authority and the Individual in Civil Association: Oakeshott, Flathman and Yves Simon." In *NOMOS* 29: *Authority Revisited*, edited by Roland Pennock and John Chapman, 131–151. New York: New York University Press, 1987.

Grant, Robert. *Oakeshott*. London: Claridge Press, 1990.

King, Preston, and B. C. Parekh, eds. *Politics and Experience: Essays Presented to Professor Michael Oakeshott on the Occasion of His Retirement*. Cambridge, U.K.: Cambridge University Press, 1968.

Letwin, Shirley Robin. "Morality and the Law." *Ratio Juris* 2 (1) (March 1989): 55–65.

Minogue, Kenneth Minogue. "Michael Oakeshott." In *Contemporary Political Philosophers*, edited by Zbigniew Pelczynski and John Gray, 120–146. New York: St. Martin's Press, 1984.

Nardin, Terry. *The Philosophy of Michael Oakeshott*. University Park, PA: Penn State University Press, 2001.

Podoksik, Efriam. "The Voice of Poetry in the Thought of Michael Oakeshott." *Journal of the History of Ideas* 63 (4) (2002): 717–733.

Wells, Harwell. "The Philosophical Michael Oakeshott." *Journal of the History of Ideas* 55 (1) (1994): 129–145.

Worthington, Glen. "Michael Oakeshott and the City of God." *Political Theory* 28 (3) (2000): 377–398.

Michael Boylan (2005)

OBJECTIVISM

See *Objectivity in Ethics*

OBJECTIVITY IN ETHICS

What objectivity in ethics is depends, in part, on what ethics is. On the narrowest understanding, ethics consists in judgments about moral constraints, which govern a person's treatment of other people, as such. On the broadest understanding, ethics includes all normative judgments, which say which responses one ought to have, and all evaluative judgments, which assess people and things against standards, as good or bad, beautiful or ugly, and so on. While it may seem strained to interpret "ethics" so broadly, many of the questions about the objectivity of ethics in the narrow sense apply to normative and evaluative judgments in general.

In one sense, what is objective is what is so independently of one's particular attitudes or position. But this idea can be specified in different ways. In one sense, a particular ethical judgment is objective if and only if it is correct, where this is an evaluation of the judgment itself, not of how it is formed or sustained. If ethical judgments are beliefs, then it is natural to think that they are correct if and only if they are true. Scholars might call this *objectivity as truth*. But ethical judgments might be correct in some way other than being true. Immanuel Kant held that some ethical judgments are correct, even though ethical judgments are commands, which cannot be true or false. Scholars might call this more inclusive conception *objectivity as correctness*.

In another sense, a particular ethical judgment is objective if and only if it is formed and sustained in response to factors that tend to make such judgments correct. An ethical judgment is objective in this sense if it results from the judger's responsible assessment of the relevant ethical considerations, not unduly influenced by his or her desires, emotions, or affiliations. Scholars might call this *objectivity as justification*.

A different kind of objectivity, described by Thomas Nagel (1979), is possessed, in the first instance, not by particular judgments themselves, but instead by what those judgments are about. Something has objective

value, in this sense, if it gives everyone reason to respond to it in the same way, regardless of his or her relation to it. For example, human suffering gives everyone reason to do what he or she can to alleviate it. Scholars might call this *objectivity as impersonality* and the associated values "impersonal" or "neutral" values. They contrast with things of "personal" or "relative" value, which give persons who stand in special relations to them reason to respond to them in special ways. For example, a child's suffering gives that child's parent more pressing reason to alleviate it than it gives others. There is a tendency, Nagel (1986) observed, to assimilate impersonality with justification and correctness, which misleadingly suggests that judgments of personal value, such as that a parent has reason to care specially for his or her child, are necessarily biased or false.

So far this entry has been considering the objectivity of particular ethical judgments and their contents. But some ask whether ethics as a whole, the sum of humankind's actual or possible ethical judgments taken together, is objective. Vaguely put, the question is whether ethical judgments are answerable to anything independent of them.

One might interpret this question as asking, "Is there an ethical reality?" where this "reality" is what ethical judgments would be answerable to. This question can be construed, in turn, as asking, "Are there ethical entities existing out there, in the world?" But this may be a tendentious formulation. What makes judgments distinctively ethical is not that they are about entities of a distinctive kind, which might exist somewhere, but instead that they predicate properties of a distinctive kind. What the question "Is there an ethical reality?" more plausibly asks is, "Do things actually have ethical properties?" And this seems to boil down to the questions "Are some actual or possible ethical beliefs, which predicate ethical properties of things, true? Can it be so that something is good or bad, right or wrong?" This is objectivity as truth, generalized to the domain as a whole. Note that in order for ethics to be objective in this sense, it is not enough that ethical judgments be either true or false. The "error theory" that J. L. Mackie (1977) proposed, which denies this kind of objectivity to ethics, asserts that all ethical judgments are false because they all contain a mistaken presupposition that something's having an ethical property is something that can be so.

Those who deny that ethical judgments are beliefs may still affirm that they can be correct, in some way other than being true. There are right and wrong answers to ethical questions, they may say, even if there is no ethical reality that makes them right or wrong. They affirm objectivity as correctness generalized to the domain as a whole.

In another sense, ethics is objective if some actual or possible ethical judgments are or could be justified. This is objectivity as justification generalized. If ethics lacks justification, it does not follow that it lacks correctness. The fact that no ethical beliefs are justified, for example, does not mean that no ethical beliefs are true. But it may seem to have similar practical implications. Even if one's ethical beliefs might be true, one has no reason to treat them as true.

In still another sense, ethics is objective if it does not "depend on" one's psychology. Scholars might call this *objectivity as mind independence*. Since the claim that ethics is mind independent is just the denial of the claim that ethics is mind dependent, the way to come to terms with the former is to come to terms with the latter. To understand what it might mean to deny that ethics "depends on" one's psychology, in other words, one needs first to understand what it might mean to assert it. It cannot be to assert that ethical judgments depend on one's psychology. This is a truism; all judgments are psychological phenomena. Nor can it be to assert that the things about which one makes ethical judgments depend on one's psychology. No one denies that some ethical judgments can be about psychological states, such as intentions to harm others.

A more promising interpretation of the idea that ethics "depends on" one's psychology—of what is denied by the claim that ethics is objective in the present sense—is that ethical judgments predicate some property involving human psychology. An extension of this idea, which scholars might call *mind dependence of properties*, might capture the sense in which noncognitivism represents ethics as mind dependent. According to noncognitivism, ethical judgments only appear to predicate properties of things, while they in fact only express the judger's decisions or feelings regarding those things. Noncognitivists, therefore, will not agree that ethical judgments predicate psychological properties. But they may say something that approximates this: that in place of predicating properties, ethical judgments express judgers' psychological states.

Another possible interpretation of the idea that ethics "depends on" one's psychology, which scholars might call *mind dependence of correctness*, is that what makes ethical judgments correct, when they are, is something about one's psychology. The mind dependence of ethical properties entails the mind dependence of ethical

correctness. If ethical judgments predicate psychological properties, then what makes those judgments true or false are psychological facts. But one might deny that ethical judgments predicate properties, while still holding, first, that they can be correct and, second, that their correctness is mind dependent. A Kantian theory might claim that ethical judgments do not predicate special ethical properties of actions, but instead command that they be done. But it might hold first that these commands can be correct and, second, that what makes them correct is something about the human will.

A natural way of spelling out the thought that ethical properties are mind dependent, which David Lewis (1989) explored in his work, is dispositionalism. Dispositionalism holds that what it is for something to have an ethical property (to be good, say) just is for it to be the case that subjects in certain conditions would respond to it in a certain way (such as by approving of or desiring it). One reservation about dispositionalism is whether the relevant response can be specified without appealing to the ethical property at issue. If approving of or desiring something consists in believing it to be good, for example, then dispositionalism appears to be circular.

Another reservation is that dispositionalism seems to imply, implausibly, that the extension of ethical properties varies with dispositions to respond, so that if the relevant subjects in the relevant conditions were not to approve of, say, kindness, it would no longer be good. One proposal to overcome this reservation, considered by David Wiggins (1998), is to identify actual dispositions as the relevant dispositions. If dispositions in the actual world are held fixed, then the extension of goodness does not vary across possible worlds, even ones in which dispositions vary. Does this mean, however, that as the identity of the actual world varies, the extension of goodness also varies? If so, then, as Lewis (1989) and Christopher Peacocke (2004) observed, the source of the original reservation seems only to have been relocated. If not, then, as Barry Stroud (2000) argued, it is unclear in what sense goodness is still being said to "depend" on dispositions. The dispositions that are held fixed are held fixed, it seems, simply because they are responsive to goodness.

Dispositionalism, it is sometimes said, is compatible with the correctness—indeed the truth—of ethical judgments. According to dispositionalism, the judgment that something is good is true if and only if subjects in the relevant conditions would approve of it. It might be said, however, that dispositionalism does not allow ethics to be correct in a more thoroughgoing sense. Although dispositionalism holds that judgments about the relevant responses can be correct, it also holds that there is no sense in which the responses themselves can be correct.

Some theories attempt to make mind dependence hospitable to a more thoroughgoing kind of correctness. John McDowell (1985) and Wiggins (1998) suggested that the relevant responses can be "merited" by their objects, and they proposed that what it is for something to be have an ethical property is, in part, for it to "merit" a certain response. In what way, then, are ethical properties still mind dependent? It is a necessary truth about any property that something has that property only if it "merits" a certain response: at very least, the judgment that it has that property. Perhaps the claim is that while this may be a necessary truth about every property, it is not an essential truth about every property. It is not part of "what it is" for something to have a shape property, for example, that it merits a response, whereas it is part of "what it is" for something to have an ethical property.

Kantians also argue for a mind dependence that is hospitable to a more thoroughgoing kind of correctness than dispositionalism allows. What makes an ethical judgment correct, according to Christine Korsgaard (1996), is that endorsing that judgment is constitutive of rational, reflective agency. Thus, the correctness of ethical judgments depends not on contingent tendencies of particular minds, as dispositionalism supposes, but instead on the necessary structure of a mind that is capable of asking ethical questions at all.

So much for what it might mean to assert or deny that ethics, as a whole, is objective. Why might one assert or deny it? Some have thought that ethics could be correct if and only if God laid down ethical laws. There are laws only where there is a lawgiver, the reasoning may go, and mortal lawgivers can establish only conventional laws. Therefore, God alone can establish ethical laws. Do all laws, however, require a lawgiver? Perhaps ethical laws, like logical laws, are not chosen by anyone. Moreover, it is unclear whether God could choose all ethical laws, for reasons given in the *Euthyphro* of Plato. If God chose certain ethical laws without regard for their goodness, then those laws would appear to be arbitrary, which it seems ethical laws cannot be. If instead God chose certain ethical laws because they were good, then God would appear to have been responding to prior and independent ethical laws, which he did not choose.

Others are anxious to deny that ethical judgments can be correct because they wish to justify tolerance of different ethical judgments. It is true that if no ethical judgment is correct, then one cannot ground one's intolerance of differing judgments on the claim that one's own

judgments are correct. However, this shows only that there is a false premise in one argument for intolerance. It does not provide any positive justification for an ethical principle of tolerance. Moreover, to justify such an ethical requirement would seem to amount to establishing the correctness of at least one ethical judgment. So it is not clear whether the denial that ethical judgments can be correct is even compatible with the attempt to justify an ethical principle of tolerance.

A more prevalent concern among contemporary academic philosophers is that the objectivity of ethical judgments is incompatible with the apparent link between making an ethical judgment and being motivated to act accordingly. For example, Mackie (1977) denied that ethical judgments can be true, on the grounds that they presuppose "queer" properties: properties such that when someone believes that an object possesses one, he or she necessarily is moved in a particular way. Perhaps what is "queer" here, however, is the unqualified claim that making an ethical judgment entails being motivated to act accordingly. More plausible, as Michael Smith (1994) and Korsgaard (1986) argued in their works, is the thesis that making an ethical judgment entails being motivated, insofar as one is not irrational, to act accordingly. Smith and Korsgaard appeared to believe, however, that this revised thesis can be explained only if the content or correctness of ethical judgments is in a way mind dependent: dependent not on the tendencies of particular contingent minds, but instead on the structure or content of ideally rational psychology.

Other philosophers are impressed by disagreement in ethics. Ethical disagreement alone, however, does not entail that ethical judgments cannot be correct, any more than scientific disagreement entails that scientific judgments cannot be correct. The thought may be—as Mackie (1977), for example, seemed to pursue it—that ethical disagreement is in some way different from other kinds of disagreement, and that this difference is evidence that ethical judgments are explained by something other than their subject matter, or that ethics cannot settle the questions that it asks. As this entry will discuss, however, these claims—that ethics can be given an "unmasking explanation" and that it cannot resolve its own questions—may seem plausible even in the absence of actual disagreement.

Still other philosophers, such as Gilbert Harman (1977), Bernard Williams (1985), and Crispin Wright (1992), doubted that ethics can be objective, on the grounds that its subject matter does not provide causal explanations. That an action was wrong, for example, does not seem to explain why anything that followed took place.

While causal powers might be required by a stipulated sense of "objectivity," it is not immediately obvious how they are relevant to objectivity intuitively understood as answerability to something independent of judgment. To be sure, some judgments are about causal powers, and so the possession of such powers is straightforwardly relevant to the correctness of such judgments. If celestial events have no influence on the fates of men, for example, then astrological beliefs are false. But as Ronald Dworkin (1996) and T.M. Scanlon (2003) noted, ethics does not purport to make judgments about causal powers. So whether ethical properties possess such powers does not seem to be similarly relevant to the correctness of ethical judgments.

What seems more plausibly relevant to objectivity is the power of the subject matter of ethics to explain, specifically, ethical judgments. If ethical beliefs, for example, are explained by something other than their putative subject matter—if, as Stroud (2000) put it, an "unmasking explanation" can be given of ethics—then it may seem that ethical beliefs are not suitably responsive to their subject matter. And if ethical beliefs are not suitably responsive to their subject matter, then they are not justified. Moreover, an unmasking explanation may be reason to doubt that ethical beliefs are true: to conclude that ethics, as a whole, is a kind of illusion. Such is the upshot of more familiar unmasking explanations of beliefs about, for example, ghosts and desert oases.

Dworkin (1996) and Scanlon (2003) questioned the assumption that beliefs can be suitably responsive to a subject matter, and hence justified, only if they are causally explained by it. Mathematical beliefs, by analogy, seem to be justified without being caused by their subject matter. Stroud (2000) doubted that an unmasking explanation of ethics can even be given. He argued that one cannot recognize ethical beliefs—the *explanandum*—without accepting some ethical claims, which the "unmasking" *explanans* was supposed to avoid.

A final concern, as Wiggins (1995) and Scanlon (2003) have suggested, is simply that ethics may seem unable to settle any, or enough, of the questions it asks. It may seem, for example, that no argument could settle whether lying to one's friend to spare her feelings in a certain kind of situation is the right thing to do. Here there seems to be a sharp contrast with mathematics, which is able to settle many of the questions it asks. The failure of ethical argument might suggest that ethical judgments cannot be justified: that we lack sufficient reason to hold them. Or it might suggest that ethical judgments cannot be correct: that the subject matter of ethics does not con-

strain unique answers to the questions that can be asked about it.

This is a "first-order" or "substantive" doubt, which arises within ethical thought itself, about the prospects of its success. It is often distinguished from "second-order" or "metaethical" doubts, such as those raised by Mackie (1977) and Harman (1977), which are supposed neither to be based on, nor to imply anything, about the prospects of "internal" ethical argument. Dworkin (1996) doubted that this distinction can be sustained, concluding that purportedly "second-order" positions about the objectivity of ethics are, if they are intelligible at all, simply substantive positions within ethics.

See also Error Theory of Ethics; Ethical Naturalism; Ethical Relativism; Ethical Subjectivism; Metaethics; Moral Principles: Their Justification; Moral Realism; Noncognitivism; Rationalism in Ethics (Practical Reason Approaches); Response-Dependence Theories.

Bibliography

Darwall, Stephen, Allan Gibbard, and Peter Railton. "Toward *Fin de Siècle* Ethics: Some Trends." In *Moral Discourse and Practice*. Oxford: Oxford University Press, 1997.

Dworkin, Ronald. "Objectivity and Truth: You'd Better Believe It." *Philosophy and Public Affairs* 25 (2) (1996): 87–139.

Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics*. Oxford: Oxford University Press, 1977.

Korsgaard, Christine. "Skepticism about Practical Reason." *Journal of Philosophy* 83 (1) (1986): 5–25.

Korsgaard, Christine. *The Sources of Normativity*. Cambridge, U.K.: Cambridge University Press, 1996.

Lewis, David. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* Supp. Vol. 63 (1989): 113–137.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth, Middlesex: Penguin, 1977.

McDowell, John. "Values and Secondary Qualities." In *Morality and Objectivity*, edited by Ted Honderich. London: Routledge & Kegan Paul, 1985.

Nagel, Thomas. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press, 1979.

Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.

Peacocke, Christopher. "Moral Rationalism." *Journal of Philosophy* 101 (10) (2004): 499–526.

Scanlon, T. M. "Metaphysics and Morals." *Proceedings and Addresses of the American Philosophical Association* 77 (2) (2003): 7–22.

Shafer-Landau, Russ. *Whatever Happened to Good and Evil?* Oxford: Oxford University Press, 2004.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.

Stroud, Barry. *The Quest for Reality: Subjectivism and the Metaphysics of Color*. New York: Oxford University Press, 2000.

Wiggins, David. "Objective and Subjective in Ethics, with Two Postscripts about Truth." *Ratio* 8 (3) (1995): 243–258.

Wiggins, David. "A Sensible Subjectivism?" In *Needs, Values, Truth*. 3rd ed. Oxford: Oxford University Press, 1998.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Wright, Crispin. *Truth and Objectivity*. Cambridge, MA: Harvard University Press, 1992.

Niko Kolodny (2005)

OCCASIONALISM

See *Cartesianism*; *Geulincx, Arnold*; *Malebranche, Nicolas*

OCKHAM, WILLIAM OF

See *William of Ockham*

OCKHAMISM

"Ockhamism" is a term used by some historians of medieval philosophy to characterize the critical and skeptical attitude toward natural theology and traditional metaphysics that became prevalent in the fourteenth century and is ascribed to the influence of William of Ockham (c. 1285–1349). There is little historical basis for speaking of an Ockhamist school, since Ockham had scarcely any avowed disciples; nor was the critical attitude toward natural theology initiated by him, although his logical criteria of demonstration and evidence undoubtedly gave it a powerful implementation. With these reservations one may, in a general sense, attach Ockham's name to the movement of thought that, in the fourteenth century, closed out the medieval enterprise of synthesizing Aristotelian philosophy with Christian theology and initiated new lines of development that led toward the scientific empiricism of the seventeenth century. The Ockhamist or nominalist movement was known in the fourteenth and fifteenth centuries as the "modern way" (*via moderna*), and was contrasted with the "old way" (*via antiqua*) associated with thirteenth-century Scholasticism.

One may distinguish two main phases of this movement of fourteenth-century thought. The first phase, occurring between 1330 and 1350, was marked by the rapid spread of Ockham's doctrines and method among the theologians and philosophers teaching at the universities of Oxford and Paris, where Ockham's logical techniques were used in criticism of the older scholastic

tradition. The second phase, less directly associated with Ockham's own teachings, commenced around 1350 and involved what may be described as a reconstruction of philosophy, and of theology as well, on foundations compatible with Ockham's empiricism and nominalism.

CRITIQUE OF SCHOLASTICISM

The influence of Ockham's logic and of his nominalistic critique of the thirteenth-century metaphysical syntheses of philosophy and theology was exhibited at Oxford in the work of Adam Wodeham (d. 1349), a Franciscan who had studied with Ockham, and of Robert Holkot (d. 1349), a Dominican theologian who lectured at Oxford around 1330 and later taught at Cambridge. Holkot was an outspoken nominalist who minced no words in stating that theology is not a science and that its doctrines can in no way be demonstrated or even comprehended by human reason. Christian dogma, for Holkot, was accepted by an act of will, on the authority of the church.

Thomas Bradwardine (c. 1290–1349) reacted against what he regarded as a new Pelagianism embodied in the Ockhamist interpretation of revealed theology, but he used Ockham's logical techniques to draw deterministic consequences from the doctrine of divine omnipotence, invoking the authority of Augustine for his views. Other Oxford teachers influenced by Ockham, and particularly by his logical methods, included Richard Swineshead ("the Calculator"), John Dumbleton, William Heytesbury, and Richard Billingham.

THE "MODERN WAY"

It was at Paris, more than at Oxford, that Ockham's influence led, after an initial resistance, to establishment of a relatively stable, and in some respects scientifically fruitful, philosophical school that endured and spread through central Europe in the late fourteenth and early fifteenth centuries.

One of the first Parisian theologians to embrace Ockham's doctrines was John of Mirecourt, a Cistercian monk who lectured on Peter Lombard's *Sentences* in 1344–1345. His skeptical treatment of the arguments of traditional theology led to a condemnation by the theological faculty at Paris of articles taken from his lectures. In many respects Mirecourt's positions resembled those of Holkot, by whom he may have been influenced.

Another victim of disciplinary action by the authorities of the University of Paris was Nicolas of Autrecourt, who was condemned to burn publicly, in November 1347, his letters to Bernard of Arezzo and his treatise *Exigit ordo executionis*. Nicolas, reacting to the Ockhamist thesis that God, by his absolute power, could cause an intuitive cognition of a nonexistent object, or could cause sensible qualities to exist without any substance being qualified by them, held that the only things of which man can have certain knowledge are the qualities perceived by his five senses, the acts or affections of his own mind, and those propositions logically evident by the principle of contradiction. From this he argued that we have no ground for belief in substances or for making inferences on the basis of causal relations, and he asserted that the whole philosophy of Aristotle is a fictitious construction devoid of any evidence or even of probability, since it rests on the assumption of substances and of causal necessities that are neither logically nor empirically evident. Preferring certainty to the Ockhamist "hypothesis of nature," Nicolas turned Ockham's critique of metaphysical necessity against Ockham's own empiricism and was rebuked by John Buridan for demanding absolute evidence, or logical necessity, in a domain of inquiry in which only conditional evidence based on the assumption of a common course of nature is appropriate.

In the hands of Buridan, a teacher on the faculty of arts at Paris, Ockham's logic, theory of knowledge, and nominalistic ontology were made the basis of a natural philosophy or physics of empirical type, within which Buridan developed the impetus theory of projectile motion and gravitational acceleration and subjected the assumptions of Aristotelian physics and cosmology to critical analysis in terms of empirical criteria of evidence. Buridan's reconstruction of natural philosophy as a positive and empirically based science of observable phenomena undermined the Aristotelian tradition and provided some of the main starting points for the development of modern mechanics in the seventeenth century.

At the same time a theologian of Paris, Gregory of Rimini (d. 1358), who became general of the order of Augustinian Hermits, made a constructive use of Ockhamist methods and doctrines in a theological synthesis of nominalism and Augustinianism; although he took issue with both Ockham and Buridan on some issues of metaphysics, the later Scholastics regarded him as a modern theologian of the nominalist group.

Natural philosophy, as distinguished from theology, was dominated by the moderately Ockhamist tradition established at Paris by Buridan, developed by Albert of Saxony and Nicholas of Oresme, and carried to the new universities of central Europe by Albert, Marsilius of Inghen, Henry of Hainbuch, and Henry of Oyta. A document drawn up by the faculty of the University of

Cologne in 1425 speaks of the period of preeminence of the *via moderna* as the century of Buridan (*saeculum Buridani*), indicating that the Ockhamism of the later fourteenth century had become associated with Buridan and his followers more than with Ockham.

RELIGIOUS INFLUENCE

The Ockhamist divorce of Christian theology from Aristotelian metaphysics, with the corresponding emphasis on religious faith and the tradition of the Church Fathers as foundation of Christian doctrine, was reflected in the popular religious movement associated with the school of Deventer and the *devotio moderna* and in the criticisms of the scholastic methods of theological disputation and argument made by Jean de Gerson at the end of the fourteenth century. Gabriel Biel (c. 1410–1495) was the last influential theologian of the Ockhamist school, and in his work the influence of Gerson, Gregory of Rimini, Holkot, and of Ockham himself brought together the diverse strands of this nominalist tradition in a doctrine with strong religious emphasis.

Ockhamism, as a well-developed philosophical and religious tradition, was submerged by the Reformation and the Counter-Reformation, as well as by the humanist revolt against the medieval cultural tradition. However, its leading ideas, in the liberation of both the Christian faith and the scientific investigation of nature from dogmatic Aristotelianism, remained operative outside the schools and bore fruit in the seventeenth and eighteenth centuries.

See also Buridan, John; Gregory of Rimini; John of Mirecourt; Nicolas of Autrecourt; William of Ockham.

Bibliography

HISTORICAL SURVEYS

Ashworth, E. J. *Language and Logic in the Post-Medieval Period*. Dordrecht: Reidel, 1974.

Courtenay, W. J. "The Reception of Ockham's Thought at the University of Paris." In *Preuve et raisons à l'Université de Paris*, edited by Z. Kaluza and P. Vignaux, Paris: Vrin, 1984, pp. 43–64.

Courtenay, W. J. "Was There an Ockhamist School?" In *Philosophy and Learning: Universities in the Middle Ages*, edited by M. J. F. M. Hoenen, J. H. J. Schneider, and G. Wieland, Leiden: Brill, 1995, pp. 263–92.

Courtenay, W. J., and K. Tachau. "Ockham, Ockhamists and the English-German Nation at Paris, 1339–1341." *History of Universities* 2 (1982): 53–96.

Klima, G. "The Medieval Problem of Universals." *The Stanford Encyclopedia of Philosophy (Winter 2004 Edition)*, edited by Edward N. Zalta. <http://plato.stanford.edu/archives/win2004/entries/universals-medieval/>.

Nuchelmans, G. *Late-Scholastic and Humanist Theories of the Proposition*. Amsterdam: North Holland, 1980.

Scott, T. K. "Nicholas of Autrecourt, Buridan and Ockhamism." In *Journal of the History of Philosophy* 9 (1971): 15–41.

Thijssen, J. M. M. H. *Censure and Heresy at the University of Paris, 1200–1400*. Philadelphia: University of Pennsylvania Press, 1998.

Zimmermann, A., ed. *Antiqui und Moderni: Traditionsbewusstsein und Fortschtrittsbewusstsein im späten Mittelalter*. Berlin: de Gruyter, 1974.

OCKHAM

Adams, M. M. *William Ockham*. 2 vols. Notre Dame, IN: University of Notre Dame Press, 1987; 2nd ed., 1989.

Panaccio, C., ed. *The Cambridge Companion to Ockham*. New York: Cambridge University Press, 1999.

Panaccio, C. *Le discours intérieur: De platon à Guillaume d'Ockham*. Paris: Éditions de Seuil, 1999.

Panaccio, C. *Les mots, les concepts et les choses. Le sémantique de Guillaume d'Occam et le nominalisme d'aujourd'hui*. Montréal-Paris: Bellarmin-Vrin, 1991.

Panaccio, C. *Ockham on Concepts*. Hampshire/Burlington: Ashgate, 2004.

Panaccio, C. "Some Epistemological Implications of the Burley-Ockham Dispute." *Franciscan Studies* 35 (1975): 212–22.

Panaccio, C. "Synonymy and Equivocation in Ockham's Mental Language." *Journal of the History of Philosophy* 18 (1980): 9–22.

Panaccio, C. "Three Versions of Ockham's Reductionist Program." *Franciscan Studies* 56 (1998): 335–46.

Spade, P. V. "Ockham's Distinctions between Absolute and Connotative Terms." *Vivarium* 13 (1975): 55–76.

BURIDAN

Friedmann, R. L., and S. Ebbesen, eds. *John Buridan and Beyond: Topics in the Language Sciences 1300–1700*. Copenhagen: The Royal Danish Academy of Sciences and Letters, 2004.

Klima, G. "Buridan's Logic and the Ontology of Modes." In *Medieval Analyses in Language and Cognition*, edited by R. L. Friedmann and S. Ebbesen, Copenhagen: The Royal Danish Academy of Sciences and Letters, 1999, pp. 473–495.

Klima, G. "The Essentialist Nominalism of John Buridan." *The Review of Metaphysics*, 58 (2005): 301–315.

Klima, G. "Latin as a Formal Language: Outlines of a Buridanian Semantics." *Cahiers de l'Institut du Moyen-Age Grec et Latin*. Copenhagen, 61 (1991): 78–106.

Normore, C. "Buridan's Ontology." In *How Things Are: Studies in Predication and the History and Philosophy of Science*, edited by J. Bogen and J. E. McGuire, Dordrecht: Reidel, 1985, pp. 189–203.

Pinborg, J., ed. *The Logic of John Buridan*. Copenhagen: Museum Tusculanum, 1976.

Thijssen, J. M. M. H., and Jack Zupko, eds. *The Metaphysics and Natural Philosophy of John Buridan*. Leiden: E. J. Brill, 2001.

Zupko, J. "Buridan and Skepticism." *Journal of the History of Philosophy* 31 (1993): 191–221.

Zupko, J. *John Buridan: Portrait of a 14th-Century Arts Master*. Notre Dame, IN: University of Notre Dame Press, 2003.

NICHOLAS OF AUTRECOURT

Grellard, C. *Croire et savior: Les principes de la connaissance selon Nicolas d'Autrecourt*. Paris: Vrin, 2005.

Kaluza, Z. "Nicolas d'Autrecourt. Ami de la vérité." In *Histoire littéraire de la France*, vol. 42, fasc. 1. Paris, 1995.

Scott, T. K. "Nicholas of Autrecourt, Buridan, and Ockhamism." *Journal of the History of Philosophy* 9 (1971): 15–41.

Tachau, K. H. *Vision and Certitude in the Age of Ockham. Optics, Epistemology and the Foundations of Semantics, 1250–1345*. Leiden: Brill, 1988.

Thijssen, J. M. M. H. "John Buridan and Nicholas of Autrecourt on Causality and Induction." *Traditio* 43 (1987): 237–255.

Thijssen, J. M. M. H. "The Quest for Certain Knowledge in the Fourteenth Century: Nicholas of Autrecourt against the Academics." In *Ancient Scepticism and the Sceptical Tradition*, edited by J. Sihvola, Helsinki: Societas Philosophica Fennica, 2000, pp. 199–223 (*Acta Philosophica Fennica*, 66).

Thijssen, J. M. M. H. "The 'Semantic Articles' of Autrecourt's Condemnation." *Archives d'histoire doctrinale et littéraire du moyen âge*, 65 (1990): 155–175.

Zupko, J. "How It Played in the rue de Fouarre: The Reception of Adam Wodeham's Theory of the *Complexe Significabile* in the Arts Faculty at Paris in the Mid-Fourteenth Century." *Franciscan Studies* 54 (1994–97): 211–225.

Ernest A. Moody (1967)
Bibliography updated by Gyula Klima (2005)

OGYŪ SORAI

(1666–1728)

Ogyū Sorai, or Butsu, was a Japanese Confucianist of the *kogakuha* ("school of ancient learning"), and famous as a political thinker. Ogyū was born in Edo (Tokyo). He was a gifted pupil and soon mastered classical Chinese; the classical style is characteristic of his writings. Proud by nature, Ogyū distinguished himself in the defense of official Zhu Xi Neo-Confucianism in polemics against Itō Jinsai. In 1716, however, his views changed, and in *Bendō* (Defining the way) and *Bemmei* (Definitions of terms) he supports most of Itō's ideas. All of Ogyū's other works were inspired by the ancient sages in accord with the maxim "back to antiquity," a maxim applicable to many of his innovations. These innovations were expressed in *Taiheisaku* (A policy for great peace) and *Seidan* (Discourses on government). Ogyū's cosmological views differ little from Itō's; Ogyū, too, rejects the dichotomy of *ri,* the principle, and *ki,* the material energy.

Ogyū holds a positivist and historicist conception of the Way (*dō*); it became for him the factual order of society, with its positive laws and institutions. He rightly points out how Confucius stressed the societal implications of the Way. Ogyū goes much further, excluding personal ethics until only "rites," that is, propriety and social behavior, combined with obedience to the government, remain. In this sense he comes very close to the Chinese Legalists in utilitarian ethics. Although he was apparently inspired by Xunzi c. 295–c. 238 BCE), he does not mention the name. For Ogyū, human nature cannot be much corrected; in this only social institutions are of any use. The sole meaning of "humaneness" is the giving of peace and prosperity to the people, and "virtue" is the virtue of the ruler in discerning able men. His political and economic ideas have little in common with Confucian moralizing. Government is a practical technique (*jutsu*), and the economy is not based on thrift but on sound social policies. He was against the idea of fanatic loyalty to the lord and advocated some social mobility, believing that the lower samurai but not the common people should be allowed to improve their status.

Ogyū's views of history are distinguished by the same practical approach. The founder of a dynasty plays a great role because of the public institutions he has to establish, yet rulers often fall because of the difficulty of preventing economic decline. Living under the Tokugawa shogunate, Ogyū rejected even the nominal sovereignty of the emperor (an opinion his best pupil, Dazai Shundai [1680–1747], concurred in). Shintoism for Ogyū was an invention of Yoshida Kanetomo (1435–1511). Ogyū's stand in favor of the Tokugawa government and his rejection of Shintoism explain why he was not repressed for his daring ideas and anti-Zhu Xi doctrine.

See also Chinese Philosophy; Itō Jinsai; Japanese Philosophy; Xunzi; Zhu Xi (Chu Hsi).

Bibliography

The principal works of Ogyū Sorai can be found in several collections. See *Nihon rinri ihen* (Library on Japanese ethics), edited by Inoue Tetsujirō (Tokyo: Ikuseikai, 1902), Vol. VI, pp. 11–203; *Nihon keizai taiten* (Classics on Japanese economics), edited by Takimoto Seiichi (Tokyo: Shishi Shuppansha, 1928), Vol. IX, pp. 3–375; *Ogyū Soraishū* (Collected works of Ogyū Sorai; Tokyo, 1937).

Two secondary sources in English are J. R. McEvan, *The Political Writings of Ogyū Sorai* (Cambridge, U.K.: Cambridge University Press, 1962) and W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of*

Japanese Tradition (New York: Columbia University Press, 1958), pp. 342–343, 422–433.

Gino K. Piovesana, S.J. (1967)

OKEN, LORENZ

(1779–1855)

Lorenz Oken, a German biologist and philosopher, was born at Bohlsbach, Baden. He was graduated from the faculty of medicine at Freiburg in 1804 and obtained his first professorship in medicine at Jena in 1807. Oken left Jena in 1819 because as editor of the liberal periodical *Isis* he had incurred the disfavor of the authorities. He traveled in Germany and France, lectured at the University of Basel in 1821 and 1822, and after a brief appointment at the University of Munich he became professor of physiology in Zürich, where he remained until his death.

After a few years in Jena, Oken was asked to transfer from medicine to philosophy. Yet ten years later, in his second term at Basel, he was listed as professor of medicine only, with no reference to philosophy. These changes reflect Oken's development and the superseding of romantic nature philosophy by a more objective study of natural phenomena. Under the influence of Friedrich von Schelling and the thinkers of the romantic school, Oken's imagination—rather than a genuine philosophical bent—swept him on to his own version of philosophy of identity. If in his time Oken was thought to be a greater philosopher than even Schelling, it was because he had a much wider knowledge of the natural sciences to illustrate and support his metaphysics. His most significant book in this connection is the *Lehrbuch der Naturphilosophie* (*Elements of Physiophilosophy*). This work aroused great interest, especially among the New England transcendentalists. Oken tried to establish a correspondence between mathematical structures and nature, and between metaphysical essences and nature. Fond of Pythagorean mysticism, he argued that all life is cast in the mold of mathematical symbols. Zero is nothingness and the infinite at the same time. The evolution of positive and negative numbers out of zero is the counterpart of a descending and ascending order of things—the descent being from matter (heavenly bodies, rocks, minerals, etc.) to some primeval mucus, while the ascent is from this mucus, seminated by infusoria and helped along by galvanism, through the whole scale of plant and animal life to man.

Metaphysically, zero is God. The disintegration of matter to mucus and the evolution of living beings illustrate God's desire to manifest himself in nature—when he comes to man, he meets himself; man is a god created by God. Theogony turns into hylogeny, the creation of matter. By the same token, all that exists is embedded in and permeated by an everlasting stream of vitality—pantheism and vitalism combine in Oken's view of the universe and its parts.

A poet in science, Ralph Waldo Emerson called Oken admiringly. The appropriateness of this remark is underlined by Oken the physiologist, who regarded man as an assembly of all the sense organs and other bodily parts developed along the ascending path; and by Oken the psychologist, who saw all animals as contributing to the psychology of the crowning organism, man. Mollusks gave man prudence and caution; from the snails man received seriousness and dignity; courage and nobility came from the insects; and the fish brought him the dowry of memory. Oken as a scientist with imagination may have had his merits, but as a philosopher he was unable to raise thought from the level of matter, chemistry, physiology, and cosmogony to a level of creative independence. Mind for Oken was merely a mirror in which God and nature could behold themselves.

In his less poetic moods, Oken came close to being a modern scientist. He held, with Johann Wolfgang von Goethe but independently of him, that the cephalic bones are a repetition of the vertebrae, and he was not far from establishing the cellular structure of living organisms. His publications after *Physiophilosophy-Lehrbuch der Naturgeschichte* and *Allgemeine Naturgeschichte für alle Stände*—reverted to the method of his earlier works: close observation and faithful description. If in Oken's days the natural sciences had to extricate themselves from preconceived mystical notions wrongly called philosophy, they beg today to be understood again in some wider context. The wheel has come full circle, as it must according to Oken's belief in the alternating processes of dynamic expansion and nostalgic reduction to a state of absolute quietness, a belief reminiscent of Friedrich Nietzsche's eternal recurrence of the same. The difference is that for Oken the fascination of this unending spectacle ended where Nietzsche's interest in it began, with the arrival of man and the search for values.

See also Emerson, Ralph Waldo; Goethe, Johann Wolfgang von; New England Transcendentalism; Nietzsche, Friedrich; Realism and Naturalism, Mathematical; Schelling, Friedrich Wilhelm Joseph von; Structuralism, Mathematical; Value and Valuation.

Bibliography

PRINCIPAL WORKS BY OKEN

Lehrbuch der Naturphilosophie, 3 vols. Jena: F. Frommann, 1809–1811; improved and final ed., 1843. Translated by Alfred Tulk as *Elements of Physiophilosophy.* London: Ray Society, 1847.

Lehrbuch der Naturgeschichte, 5 vols. Leipzig and Jena, 1813–1826.

Allgemeine Naturgeschichte für alle Stände, 13 vols. Stuttgart: Hoffman, 1833–1841.

WORKS ON OKEN

Ecker, Alexander. *Lorenz Oken. Eine biographische Skizze.* Stuttgart, 1880.

Hübner, Georg Wilhelm. *Okens Naturphilosophie prinzipiell und kritisch bearbeitet.* Borna-Leipzig, 1909.

Schuster, Julius. *Oken. Der Mann und sein Werk.* Berlin, 1922.

Hermann Boeschenstein (1967)

OLIVI, PETER JOHN

(1248–1298)

Peter John Olivi was one of the most original philosophers of the late thirteenth century. Despite the influence his ideas had in the Middle Ages and in the formation of the early modern thought, his own writings have been studied little. The Council of Vienne (1311–1312) and Pope John XXII (in 1326) condemned some of his views, and after this his works (most of which have survived in the Vatican library) remained mostly in obscurity. His innovative ideas on the philosophy of history, on Aristotelian metaphysics, and especially on human freedom were developed by other philosophers whose texts had a more constant and wider circulation (e.g., John Duns Scotus, William Ockham, and Peter Aureol).

As a twelve-year-old youth in 1261, Olivi entered the Franciscan order and thereby also one of the best educational systems of the time. From 1267 to 1272 he studied in Paris with St. Bonaventure and other famous thinkers. Possibly because of arrogant opinions, he did not receive a doctorate. Nevertheless, he moved on to teach at different Franciscan schools in southern France. After some of his views were condemned in 1283, he withdrew from such duties. He was rehabilitated in 1288 with the help of his former teacher, Cardinal Matthew of Aquasparta, and taught in Florence for two years before returning to Montpellier and later Narbonne, where he stayed until his death on March 14, 1298.

Readers of Olivi's works have often noted that Olivi had a very distinctive writing style. Though his works clearly belong to the genres of medieval academic writing, they contain a very personal tone that seems to spring from Olivi's intimate experiential touch to philosophical thinking. Olivi clearly had a liking for arguments, and often he refrained from making a determinate solution, although he did not hesitate to take strong stances on some very controversial issues. In general, his habits of thought have a surprisingly modern feel.

SOCIAL PHILOSOPHY

Olivi's most important innovations in social philosophy are related to the Franciscan ideal of poverty. In his commentary on the Apocalypse and already in the early *Questions on Evangelical Perfection* he formulated a theory of how the Franciscans used the necessities of life without having property in them (*usus pauper*). The theory differs in its detail to what John Duns Scotus and William Ockham presented later, but the crucial philosophical innovations can be found already in Olivi's works.

The idea of subjective right is often connected to early modern political philosophy, but it was developed already in the discussions concerning Franciscan poverty. Olivi's view concerning rights differed from the Aristotelian orthodoxy of the time, for according to him the natural order does not imply rights. Rather, they must be constituted by an act of a free will. This view becomes clear in his theory of property acquisition and of political power. Though Olivi taught for the Franciscans absolute obedience to the superiors, he qualified that the power of the superiors must accord with the purpose of the power. This makes obedience in fact an issue that each person must weigh in his or her own conscience.

Olivi was a theologian, and he wrote many biblical commentaries, often with an apocalyptic message. He also had a historical view of the Church as a changing institution. He has often been understood as claiming that the Antichrist will be a pope.

HUMAN FREEDOM

The human free will is a topic that receives a large share of what can be called Olivi's main philosophical work, the commentary on Peter of Lombards *Sentences*. Some of Olivi's strongest anti-Aristotelian formulations come form this context. Like apparently all the texts where he explicitly opposes Aristotelian thought, it was written soon after the bishop Etienne Tempier's condemnation of 1277 against 219 more or less Aristotelian theses. Olivi showed no knowledge of the documents of the condemnation themselves, but attacked the Aristotelian positions and apparently also Thomas Aquinas's views quite openly.

According to Olivi's main argument for the freedom the human will, the ground for human social practices like friendship and gratitude, and even personhood, would collapse if human beings denied the freedom of the will. In Olivi's view, free choice is a real possibility open for all mentally healthy adult humans in their normal condition. Unlike the animals, humans can make choices self-reflexively as their own choices. Olivi discussed the Aristotelian practical syllogism and accepted that humans consider rationally what would be the best course of action given a certain end. But even after this consideration, humans remain free to follow the best course of action or to do something else. Also, the human will is always free to posit a new ultimate end. In Olivi's example, if one hates one's enemy and reasons the best way to harm the person, one remains free not to inflict harm, or even to begin loving the person for his or her own sake. Every human has an almost infinite moral worth based on such freedom, and as a free agent can be treated as a person.

METAPHYSICS

Olivi's ontological view of the human soul was rejected by the fourteenth-century Church as too dualist. He was understood to have claimed that the soul is not the form of the body, though his point was subtler. According to his metaphysics, all individuals consist of matter and form. However, he distinguished two kinds of matter: corporeal and spiritual. The human soul informs matter of both kinds, but the intellectual soul does not inform any corporeal matter. The human soul is thus a form of the corporeal body only in respect to its sensitive part. Thus, Olivi accepted the Aristotelian metaphysics of form and matter, but thought that the human intellectual soul is a full individual capable of existence and activity even without the body. This tradition of thought was continued by later Franciscans like Scotus and Ockham, although they gave up the idea of spiritual matter and with it also the universality of the form-matter metaphysics, making the intellectual soul an immaterial substance. In this way, Olivi's theory can be seen as direct predecessor of René Descartes's seventeenth-century dualist view.

In the philosophy of mind, Olivi's most important starting point was that the mind is active and the corporeal bodies are passive. He described sensory perception in terms of an intentional relation where the mind comports to the world, thus rejecting the standard Aristotelian model that the corporeal things act upon the cognitive systems. Olivi also developed a relatively elaborated theory of the self and human self-understanding.

Olivi was a well-educated intellectual working in a way similar to his contemporaries. In most of the topics he treated he refrained from putting forward a full theory. Rather, he aimed at deeper, though incomplete, understanding on the complexity of the problems, and called for recognition of the imperfections of the human reasoning capacities. Olivi did not oppose rational thought, but he saw its limits. Much of his philosophical originality lies in the way he strove for a rationally un-Aristotelian way of thinking at a time in which basic university education was based on Aristotle's texts.

See also Aristotelianism; Bonaventure, St.; Descartes, René; Determinism and Freedom; Duns Scotus, John; Matthew of Acquasparta; Medieval Philosophy; Peter Aureol; Peter Lombard; Philosophy of History; Philosophy of Mind; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS BY PETER JOHN OLIVI

For a listing of Olivi's writings, see *Archivum Franciscanum historicum* 91 (3–4) (1998).

Quaestiones in secundum librum Sententiarum. Bibliotheca Franciscana Scholastica 4–6. Edited by B. Jansen. Quaracchi: Collegium S. Bonaventurae, 1922–1926.

Quodlibeta quinque. Edited by S. Defraia. Grottaferrata: Editiones Collegii S. Bonaventurae ad Claras Aquas, 2002.

WORKS ABOUT PETER JOHN OLIVI

Boureau, A., and S. Piron, eds. *Pierre de Jean Olivi (1248–1298)*. Paris: Vrin, 1999.

Burr, David. *Olivi and Franciscan Poverty: The Origins of the Usus Pauper Controversy*. Philadelphia: University of Pennsylvania Press, 1989.

Lagerlund, H., and M. Yrjönsuuri, eds. *Emotions and Choice from Boethius to Descartes*. Kluwer: Dordrecht, 2002.

Pasnau, Robert. *Theories of Cognition in the Later Middle Ages*. Cambridge, U.K.: Cambridge University Press, 1997.

Putallaz, François-Xavier. *Insolente liberté. Controverses et condamnations au XIIIe siècle*. Paris: Cerf, 1995.

Yrjönsuuri, Mikko (2005)

OMAN, JOHN WOOD
(1860–1939)

John Wood Oman, the philosopher of religion and theologian, was a Scotsman from the Orkney Islands. After being educated at Edinburgh and Heidelberg universities and serving for seventeen years in a rural pastorate in

Northumberland, he taught for twenty-eight years at Westminster College, Cambridge, the seminary of the English Presbyterian Church. The chief influence on his developing thought was that of Friedrich Schleiermacher, whose *Reden* Oman translated into English.

In the massive *The Natural and the Supernatural* (1931) Oman portrays the root of religion as man's immediate sense of the Supernatural. The primary religious awareness is not inferential but is, in words that Oman used to describe the similar conception of Schleiermacher, "intuition of reality, an intercourse between a universe, present always in all its meaning, and a spirit, responding with all its understanding" (p. 36). By the Supernatural, Oman does not mean the mysterious, the uncanny, or the miraculous but a larger environment than physical nature, "a special kind of environment, which has its own particular sanctions" (p. 23), through commerce with which man receives his characteristically human degree of independence within his natural environment.

The Supernatural is variously conceived in different types of religion, as is the character of the redemption that the supernatural makes possible. In primitive religion redemption is found by seeking the Supernatural in nature as an animistic force indefinitely many and yet vaguely one. In polytheism the Supernatural consists of individual spirits that rule different parts of nature, and redemption means the managing of nature through its many divine masters. Cosmic pantheism accepts nature in its wholeness as the Supernatural, while the acosmic mysticism of India wholly excludes nature from the Supernatural, as illusion. Religions of the ceremonial-legal type, such as priestly Judaism and Islam, divide the Natural into a sacred realm and a secular realm, cultivating the sacred or religious while leaving the secular outside the sphere of redemption. Finally, for the prophetic monotheism of the Hebrew prophets and of Christianity redemption is reconciliation to the Natural by finding within it the purpose of the one personal Supernatural. To be reconciled to God is to accept all the experiences of one's life as of God's appointing, and one's duties as divine commands. Thus, prophetic religion is intensely practical and this-worldly. Speaking of its Old Testament representatives, Oman says, "What determines their faith is not a theory of the Supernatural, but an attitude towards the Natural, as a sphere in which a victory of deeper meaning than the visible and of more abiding purpose than the fleeting can be won" (p. 448).

Oman emphasizes that knowledge of our environment, whether the natural or the Supernatural, does not consist in the mere registering of "impacts" but always consists in a perception of "meaning." In order to become aware of our environment, we must rightly interpret its impingements upon us. "Thus knowledge is not knowledge as an effect of an unknown external cause, but is knowledge as we so interpret that our meaning is the actual meaning of our environment" (p. 175). In this interpretative process, the mind exercises a degree of freedom. That degree is established by the individual frontiers of each mind, which are largely controlled from within and across which the meaning of the environment can pass only as a meaning recognized by the individual.

The Supernatural presents itself to the human mind with the quality of the sacred or of absolute worth. To be aware of the Supernatural is to recognize some sacred value that lays an absolute claim upon us, even if in the early stages of man's dealings with the Supernatural this is only an irrational taboo. Religion is "essentially a dealing with an unseen environment of absolute worth, which demands worship" (p. 23). This recognition of and allegiance to the sacred frees man from the dominance of his physical surroundings: "He obtained firm footing to deal with his environment the moment he regarded anything as sacred, because he could say 'No' and was no longer its mere creature" (p. 85).

While man's sense of the Supernatural gives him a fixed point amid the evanescent and a degree of freedom in relation to the natural, he can gain this only by exercise of his own freedom. For "The peculiarity of the supernatural environment is that we cannot enter it except as we see and choose it as our own" (p. 309).

Oman makes no use of the attempted logical coercion of the traditional theistic proofs. He does not try to establish the truth of religion independently of religious experience. Rather he starts from the fact of the religious man's awareness of a larger supernatural environment, in terms of which he lives, and argues that this awareness has no greater need or possibility of philosophical justification than has our awareness of the natural environment. "Among Western thinkers from [René] Descartes onwards, attempts have been made to prove the existence of a material world by other evidence than the way it environs us, but the result was no more reassuring for the reality of the natural world than for the reality of the supernatural" (p. 51).

The same basic standpoint is evident in Oman's contributions to doctrinal theology, especially his *Grace and Personality* (1919). Oman was the first of a series of twentieth-century Christian thinkers—such as Karl Heim, Emil Brunner, H. H. Farmer, and John Macmurray—to

treat as a normative principle of his theology the insight that God is the supremely personal reality, that his dealings with men take place in the personal realm, and that the great central Christian terms—revelation, faith, grace, sin, reconciliation—are to be understood as part of the language of personal relationship and are perverted when construed in nonpersonal ways. Oman taught that religious truths are not infallibilities declared authoritatively from heaven but claim acceptance only because they irresistibly impress our minds as true, and that God seeks our trust only by showing himself to be trustworthy.

There are in Oman's works the elements of a religious philosophy that might well appeal to many today because it is consistently empiricist, being based upon what is given in human experience. However, it is often expressed in Oman's pages on a higher level of generality, and with less detailed precision, than has become customary since he wrote, and there is therefore scope for the development of these same themes in more contemporary terms.

See also Brunner, Emil; Descartes, René; Heim, Karl; Religion, Naturalistic Reconstructions of; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

WORKS BY OMAN

Vision and Authority. London: Hodder and Stoughton, 1902; 2nd ed., 1928.

The Problem of Faith and Freedom in the Last Two Centuries. London: Hodder and Stoughton, 1906.

The Church and the Divine Order. London: Hodder and Stoughton, 1911.

The War and Its Issues. Cambridge, U.K.: Cambridge University Press, 1916.

Grace and Personality. Cambridge, U.K.: Cambridge University Press, 1919; 4th ed., 1931.

The Paradox of the World: Sermons. Cambridge, U.K.: Cambridge University Press, 1921.

The Book of Revelation. Cambridge, U.K.: Cambridge University Press, 1923.

The Office of the Ministry. London: Student Christian Movement Press, 1928.

The Text of Revelation. Cambridge, U.K.: Cambridge University Press, 1928.

The Natural and the Supernatural. Cambridge, U.K.: Cambridge University Press, 1931.

Concerning the Ministry. London: Student Christian Movement Press, 1936.

Honest Religion. Cambridge, U.K.: Cambridge University Press, 1941.

A Dialogue with God. London: James Clarke, 1951. Sermons.

WORKS ON OMAN

Bevons, Stephen. *John Oman and His Doctrine of God.* Cambridge, U.K.: Cambridge University Press, 1992.

Healey, F. G. *Religion and Reality: The Theology of John Oman.* Edinburgh: Oliver and Boyd, 1965.

Hood, Adam. *Baillie, Oman, and Macmurray: Experience and Religious Belief.* Burlington: Ashgate, 2003.

Langford, Thomas. *In Search of Foundations: English Theology 1900–1920.* Nashville: Abington Press, 1969.

Langford, Thomas. "The Theological Methodology of John Oman and H. H. Farmer." *Religious Studies* 1 (1965): 229–240.

Tennant, F. R. "John Wood Oman, 1860–1939." *Proceedings of the British Academy* 25 (1939): 333–338.

John Hick (1967)
Bibliography updated by Christian B. Miller (2005)

ONTOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD

The Ontological Argument for the existence of God was first propounded by Anselm (c. 1033–1109), abbot of Bee and later archbishop of Canterbury, in his *Proslogion* (Chs. 2–4) and in his *Reply* to a contemporary critic.

He begins (*Proslogion* 2) with the concept of God as "something than which nothing greater can be conceived" (*aliquid quo nihil maius cogitari possit*, and other equivalent formulations). It is clear that by "greater" Anselm means "more perfect." (Sometimes he uses *melius*, "better," instead of *maius*, "greater": for instance, *Proslogion* 14 and 18.) Since we have this idea, it follows that "Something than which nothing greater can be conceived" at least exists in our minds (*in intellectu*) as an object of thought. The question is whether it also exists in extramental reality (*in re*). Anselm argues that it must so exist, since otherwise we should be able to conceive of something greater than that than which nothing greater can be conceived—which is absurd. Therefore "Something than which nothing greater can be conceived" must exist in reality.

In *Proslogion* 3 Anselm adds that "Something than which nothing greater can be conceived" exists in the truest and greatest way (*verissime et maxime esse*); for whereas anything else can be conceived not to exist (and thus exists only contingently), "Something than which nothing greater can be conceived" cannot be conceived not to exist (and thus exists necessarily). For that which cannot be conceived not to exist is greater than that which *can* be conceived not to exist, and therefore only

that which cannot be conceived not to exist is adequate to the notion of "Something a greater than which cannot be conceived."

Anselm explains (in his *Responsio*) that by a being which cannot be conceived not to exist he means one that is eternal in the sense of having no beginning or end and always existing as a whole, that is, not in successive phases. He argues that if such a being can be conceived, it must also exist. For the idea of an eternal being that has either ceased to exist or has not yet come into existence is self-contradictory; the notion of eternal existence excludes both of these possibilities. This latter argument has been revived and developed in our own day (see below).

Many of the earliest manuscripts of the *Proslogion* contain a contemporary criticism (attributed in two of the manuscripts to one Gaunilo of Marmoutier) together with Anselm's reply. The criticism, summed up in the analogy of the island, is directed against Anselm's argument as presented in *Proslogion* 2. Gaunilo sets up what he supposes to be a parallel ontological argument for the existence of an island more perfect than any known island: such an island must exist, since otherwise it would be less perfect than any known island, and this would be a contradiction. In reply Anselm develops the reasoning of *Proslogion* 3. His argument cannot be applied to islands or to anything else whose nonexistence is conceivable, for whatever can be conceived not to exist is *eo ipso* less than "Something than which nothing greater can be conceived." Only from this latter notion can we (according to Anselm) deduce that there must be something corresponding to it in reality.

Perhaps the most valuable feature of Anselm's argument is its formulation of the Christian concept of God. Augustine (*De Libero Arbitrio* II, 6, 14) had used the definition of God as one "than whom there is nothing superior." The Ontological Argument could not be based upon this notion, for although it is true by definition that the most perfect being that there is, exists, there is no guarantee that this being is God, in the sense of the proper object of man's worship. Anselm, however, does not define God as the most perfect being that there is but as a being than whom no more perfect is even conceivable. This represents the final development of the monotheistic conception. God is the most adequate conceivable object of worship; there is no possibility of another reality beyond him to which he is inferior or subordinate and which would thus be an even more worthy recipient of man's devotion. Thus metaphysical ultimacy and moral ultimacy coincide; one cannot ask of the most perfect conceivable being, as one can of a first cause, necessary being, unmoved mover, or designer of the world (supposing such to exist) whether men ought to worship him. Here the religious exigencies that move from polytheism through henotheism to ethical monotheism reach their logical terminus. And the credit belongs to Anselm for having first formulated this central core of the ultimate concept of deity.

DESCARTES'S ARGUMENT

Anselm's argument was rejected by Thomas Aquinas in favor of the Cosmological Argument and as a consequence was largely neglected during the remainder of the medieval period. It was, however, again brought into prominence by René Descartes in the seventeenth century, and most subsequent discussions have been based upon Descartes's formulation. Descartes made explicit the presupposition of the argument that existence is an attribute or predicate which, like other predicates, a given *x* can meaningfully be said to have or to lack. He claims that just as the idea of a triangle necessarily includes among the defining attributes of a triangle that of having its three internal angles equal to two right angles, so the idea of a supremely perfect being (a different formula from Anselm's) necessarily includes the attribute of existence. Consequently we can no more think, without contradiction, of a supremely perfect being which lacks existence than of a triangle which lacks three sides.

Descartes considers the following objection: From the fact that in order to be a triangle a figure must have three sides it does not follow that there actually are any triangles; and likewise in the case of the concept of a supremely perfect being. His reply is that whereas the notion, or essence, of a triangle does not include the attribute of existence that of a supremely perfect being does, and that therefore in this special case we are entitled to infer existence from a concept.

KANT'S CRITICISM

Descartes's version of the Ontological Argument had some important contemporary critics—for example, Pierre Gassendi and Johannes Caterus (Johan de Kater)—but the classic criticism is that of Immanuel Kant. This moves on two levels. First, leaving the argument's presuppositions for the moment unchallenged, he grants the analytic connection that Descartes had affirmed between the concept of God and that of existence. In the proposition "A perfect being exists" we cannot without contradiction affirm the subject and reject the predicate. But, he points out, we can without contra-

diction elect not to affirm the subject together with its predicate. We can reject as a whole the complex concept of an existing all-perfect being.

Second, however, Kant rejects the assumption that existence is a real predicate. If it were a real, and not merely a grammatical, predicate, it would be able to form part of the definition of God, and it could then be an analytic truth that God exists. But existential propositions (propositions asserting existence) are always synthetic, always true or false as a matter of fact rather than as a matter of definition. Whether any specified kind of thing exists can be determined only by the tests of experience. The function of "is" or "exists" is not to add to the content of a concept but to posit an object answering to a concept. Thus, the real contains no more than the possible (a hundred real dollars are the same in number as a hundred imagined ones); the difference is that in the one case the concept does and in the other case it does not correspond to something in reality.

RUSSELL'S ANALYSIS. Essentially the same point—so far as it affects the Ontological Argument—was made in the twentieth century by Bertrand Russell in his theory of descriptions. This involves an analysis of positive and negative existential propositions, according to which to affirm that *x*'s exist is to affirm that there are objects answering to the description "*x*," and to deny that *x*'s exist is to deny that there are any such objects. The function of "exists" is thus to assert the instantiation of a given concept. "Cows exist" is not a statement about cows, to the effect that they have the attribute of existing, but about the concept or description "cow," to the effect that it has instances. If this is so, then the proper theological question is not whether a perfect being, in order to be perfect, must together with its other attributes have the attribute of existence but whether the concept of an (existing) perfect being has an instance. This question cannot be determined a priori, as the Ontological Argument professes to do, by inspection of the concept of God. The nature of thought on the one hand and of the extramental world on the other, and of the difference between them, is such that there can be no valid inference from the thought of a given kind of being to the conclusion that there is in fact a being of that kind. This is the fundamental logical objection to the Ontological Argument.

HEGELIAN USE OF THE ARGUMENT

Prior to Kant, the Ontological Argument had been used by Benedict de Spinoza and Gottfried Wilhelm Leibniz. Since Kant, the form of it that he discussed has remained under the heavy cloud of his criticism. However, G. W. F. Hegel and his school put the argument to a somewhat different use. As Hegel himself expressed it, "In the case of the finite, existence does not correspond to the Notion (*Begriffe*). On the other hand, in the case of the Infinite, which is determined within itself, the reality must correspond to the Notion (*Begriffe*); this is the Idea (*Idee*), the unity of subject and object" (*Vorlesungen über die Philosophie der Religion*, Vol. II, p. 479). Otherwise stated, Being itself, or the Absolute, is the presupposition of all existence and all thought. If finite beings exist, Being exists; when beings think, Being comes to self-consciousness; and in the reasoning of the Ontological Argument, finite thinking is conscious of its own ultimate ground, the reality of which it cannot rationally deny.

The defect of this argument is that its conclusion is either trivial or excessively unclear. It is trivial if the reality of Being is synonymous with the existence of the sum of finite beings; but on the other hand, it is so unclear as to be scarcely interesting if Being is regarded as a metaphysical quantity whose distinction from the sum of finite beings cannot be explicated.

The use of the argument in early twentieth-century French "reflexive" philosophy (see bibliography) has affinities with the Hegelian use.

CONTEMPORARY DISCUSSIONS

Discussion of the Ontological Argument has continued throughout the modern period and is perhaps as active today as at any time in the past. For there is perennial fascination in a piece of reasoning that employs such fundamental concepts, operates so subtly with them, and professes to demonstrate so momentous a conclusion.

Among theologians, attempts have been made to maintain the value of the argument, not as a proof of God's existence but as an exploration of the Christian understanding of God. Thus, Karl Barth regards the proof as an unfolding of the significance of God's revelation of himself as One whom the believer is prohibited from thinking as less than the highest conceivable reality. On this view Anselm's argument does not seek to convert the atheist but rather to lead an already formed Christian faith into a deeper understanding of its object. Again, Paul Tillich treated the theistic proofs as expressions of the question of God that is implied in our human finitude. They analyze different aspects of the human situation, showing how it points to God. Thus, the Ontological Argument "shows that an awareness of the infinite is included in man's awareness of finitude." This is in effect a Hegelian use of the argument.

HARTSHORNE AND MALCOLM. At the same time, some contemporary philosophers—especially Charles Hartshorne and Norman Malcolm—revived the second argument, or second form of the argument, found in Anselm's *Proslogion* (3) and in his *Responsio* to Gaunilo. As they have reconstructed it, this argument starts from the premise that the concept of God as eternal, self-existent being is such that the question whether God exists cannot be a contingent question but must be one of logical necessity or impossibility. A being who exists, but of whom it is conceivable that he might not have existed, would be less than God; for only a being whose existence is necessary rather than contingent can be that than which nothing greater is conceivable. But if such a necessary being does not exist, it must be a necessary rather than a contingent fact that he does not exist. Thus God's existence is either logically necessary or logically impossible. However, it has not been shown to be impossible—that is, the concept of such a being has not been shown to be self-contradictory—and therefore we must conclude that God necessarily exists.

Hartshorne formalizes the argument as follows:

(1) $q \rightarrow Nq$	"Anslem's principle": perfection could not exist contingently
(2) $Nq \vee {\sim}Nq$	Excluded middle
(3) ${\sim}Nq \rightarrow N{\sim}Nq$	Form of Becker's postulate: modal status is always necessary.
(4) $Nq \vee N{\sim}Nq$	Inference from (2, 3)
(5) $N{\sim}Nq \rightarrow N{\sim}q$	Inference from (1): the necessary falsity of the consequent implies that of the antecedent (modal form of *modus tollens*)
(6) $Nq \vee N{\sim}q$	Inference from (4, 5)
(7) ${\sim}N{\sim}q$	Intuitive postulate (of conclusion from other theistic arguments): perfection is not impossible
(8) Nq	Inference from (6, 7)
(9) $Nq \rightarrow q$	Modal axiom
(10) q	Inference from (8, 9)

In this formalization q stands for $(\exists x)Px$ ("There is a perfect being" or "Perfection exists"); N means "analytic or L-true, true by necessity of the meanings of the terms employed"; and $\rightarrow$ signifies strict implication.

CRITICISM. The above argument seems to depend upon a confusion of two different concepts of necessary being. The distinction involved is important for the elucidation of the idea of God and represents one of the points at which study of the Ontological Argument can be fruitful even though the argument itself fails. The two concepts are those of logical necessity and ontological or factual necessity. In modern philosophy, logical necessity is a concept that applies only to propositions; a proposition is logically necessary if it is true in virtue of the meanings of the terms composing it. And it is a basic empiricist principle that existential propositions cannot be logically necessary. In other words, whether or not a given kind of entity exists is a question of experiential fact and not of the rules of language. On this view, the notion of a logically necessary being is inadmissible, for it would mean that the existential proposition "God exists" is logically true or true by definition. Anselm's principle, however, which is used as the first premise of Hartshorne's argument, was not that God is a logically necessary being (in this modern sense) but that God is an ontologically or factually necessary being, For, as noted above, Anselm was explicit that by a being whose nonexistence is inconceivable he meant a being who exists without beginning or end and always as a whole. (This is virtually the scholastic notion of *aseity*, from *a se esse*, "self-existence," that is, eternal and independent existence.) Interpreting "For God to exist is for him to exist necessarily" (prop. 1) in this way, we can validly infer from it that God's existence is ontologically either necessary or impossible (prop. 6). For if an eternal being exists, he cannot, compatibly with the concept of him as eternal, cease to exist: thus his existence is necessary. And if such a being does not exist, he cannot, compatibly with the concept of him as eternal, come to exist: thus his existence is impossible.

However, it does not follow from this that an eternal being in fact exists but only that if such a being exists, his existence is ontologically necessary, and that if no such being exists, it is impossible for one to exist. Hartshorne's argument can advance from proposition 6 to its conclusion only by assuming at this point that it has been established that the existence of God is (not, or not only, ontologically but) *logically* necessary or impossible. He can then rule out the latter alternative (prop. 7), and conclude that God necessarily exists (prop. 8) and hence that he exists (prop. 10). Thus, in propositions 1–6 "necessary" means "ontologically necessary"; in propositions 6–10 it means "logically necessary"; and proposition 6 itself is the point at which the confusion occurs. (The same illicit shift between the notions of ontological and logical necessity can be observed in Malcolm's version of the argument.)

The conclusion to be drawn is that the Ontological Argument, considered as an attempted logical demon-

stration of the existence of God, fails. In both of the forms that are found in Anselm, and which are still matters of discussion today, the flaw in the argument is that while it establishes that the concept of God involves the idea of God's existence, and indeed of God's necessary (in the sense of eternal) existence, it cannot take the further step of establishing that this concept of an eternally existent being is exemplified in reality.

See also Anselm, St.

Bibliography

For the general history of the Ontological Argument, see M. Esser, *Der ontologische Gottesbeweis und seine Geschichte* (Bonn, 1905); J. Kopper, *Reflexion und Raisonnement im ontologischen Gottesbeweis* (Cologne, 1962); and C. C. J. Webb, "Anselm's Ontological Argument for the Existence of God," in *PAS* 3 (2) (1896): 25–43. On its alleged origin in Plato, see J. Moreau, "L'argument ontologique" dans le *Phédon*," in *Revue philosophique de la France et de l'étranger* 137 (1947): 320–343, and J. Prescott Johnson, "The Ontological Argument in Plato," in *Personalist* (1963): 24–34. The main passage in Augustine which may have influenced Anselm (in relation to the Ontological Argument) is *De Libero Arbitrio*, Book II, translated by J. H. S. Burleigh as "On Free Will," in *Augustine: Earlier Writings*, edited by J. H. S. Burleigh (London and Philadelphia, 1953).

Anselm's version of the argument occurs in *Proslogion*, Chs. 2–4, and in his exchange with Gaunilo. The Latin text is in Dom F. S. Schmitt's critical edition of the *Opera Omnia*, Vol. I (Edinburgh, 1945), pp. 101–104 and 125–139. The best English translations of the *Proslogion* as a whole are those of Eugene R. Fairweather in *A Scholastic Miscellany: Anselm to Ockham* (Philadelphia: Westminster Press, 1956); A. C. Pegis in *The Wisdom of Catholicism* (New York: Random House, 1949); M. J. Charlesworth, *St. Anselm's "Proslogion" with "A Reply on Behalf of the Fool" by Gaunilo and "The Author's Reply to Gaunilo,"* with an introduction and philosophical commentary by Charlesworth (Oxford: Clarendon Press, 1965); and A. C. McGill in *The Many Faced Argument*, edited by John Hick and A. C. McGill (New York: Macmillan, 1967), where McGill, as well as translating *Proslogion* 2–4, has translated Anselm's exchange with Gaunilo, arranging the texts so as to present Anselm's replies in relation to the specific criticisms to which they refer. The translations by S. N. Deane in *St. Anselm: Basic Writings*, 2nd ed. (La Salle, IL, 1962) include, in addition to the *Proslogion*, Gaunilo's *On Behalf of the Fool* and Anselm's *Reply*. Anselm's dialogue *De Veritate* is also relevant; there is an excellent English translation by Richard McKeon in *Selections from Medieval Philosophers*, Vol. I (New York: Scribners, 1929).

On the interpretation of Anselm, the most important older writings are Charles Filliatre, *La philosophie de saint Anselme* (Paris, 1920); Alexandre Koyré, *L'idée de Dieu dans la philosophie de S. Anselme* (Paris, 1923); Karl Barth, *Fides Quaerens Intellectum: Anselms Beweis der Existenz Gottes in Zusammenhang seines theologischen Programms* (Zürich, 1930; 2nd ed., 1958), translated by Ian Robertson as *Anselm: Fides Quaerens Intellectum* (Richmond, VA: John Knox Press, 1960); A. Stolz, "Zur Theologie Anselms im Proslogion," in *Catholica* (1933): 1–21, translated by A. C. McGill in *The Many Faced Argument* (see above); Étienne Gilson, "Sens et nature de l'argument de saint Anselme," in *Archives d'histoire doctrinale et littéraire du moyen âge* 9 (1934): 5–51; and Adolf Kolping, *Anselms Proslogion—Beweis der Existenz Gottes im Zusammenhang seines Spekulativen Programms: Fides Quaerens Intellectum* (Bonn, 1939).

For more recent historical studies, see *Spicilegium Beccense*, Proceedings of the Congrès International du IX^{e} Centenaire de l'arrivée d'Anselme au Bec (Paris: J. Vrin, 1959).

For the view that there are two different arguments in *Proslogion*, see Charles Hartshorne, introduction to *St. Anselm: Basic Writings* (see above); "What Did Anselm Discover?," in *Union Theological Seminary Quarterly Review* (1962): 213–222; and *The Logic of Perfection* (La Salle, IL: Open Court, 1962), Ch. 2. See also Norman Malcolm, "Anselm's Ontological Arguments," in *Philosophical Review* 69 (1 (1960): 41–62, reprinted in Malcolm's *Knowledge and Certainty* (Englewood Cliffs, NJ: Prentice Hall, 1963). St. Thomas's criticism of the argument occurs in his *Summa Contra Gentiles* I, 10 and 11, and *Summa Theologiae* I, 2, 1.

For the argument after Anselm, see O. Paschen, *Der ontologische Gottesbeweis in der Scholastik* (Aachen, 1903); John Duns Scotus, *De Primo Principio*, translated by E. Roche (St. Bonaventure, NY: Franciscan Institute, 1949); I. Hislop, "St. Thomas and the Ontological Argument," in *Contemplations Presented to the Dominican Tertiaries of Glasgow* (London, 1949), pp. 32–38; P. L. M. Puech, "Duns Scotus et l'argument de saint Anselme," in *Nos Cahiers* (1937): 183–199; A. Runze, *Der ontologische Gottesbeweis, kritische Darstellung seiner Geschichte seit Anselm bis auf die Gegenwart* (Halle, 1882); O. Herrlin, *The Ontological Proof in Thomistic and Kantian Interpretation* (Uppsala: Lundquistska bokhandeln, 1950); Robert Miller, "The Ontological Argument in St. Anselm and Descartes," in *Modern Schoolman* (1954–1955): 341–349, and (1955–1956): 31–38; and Alexandre Koyré, *Essai sur l'idée de Dieu et les preuves de son existence chez Descartes* (Paris: E. Leroux, 1922).

Descartes's own version of the argument occurs in his *Meditations* (V, but see also III). Spinoza uses the argument in his *Ethics*, Part I, Props. 7–11. Leibniz's version is found in his *New Essays concerning Human Understanding*, Book IV, Ch. 10 and Appendix X, and in his *Monadology*, Secs. 44–45. Kant's criticism of the argument (in its Cartesian form) occurs in his *Critique of Pure Reason*, Transcendental Dialectic, Book II, Ch. iii, Sec. 4.

The Hegelian use of the argument is to be found in Hegel's *Vorlesungen über die Philosophic der Religion*, edited by P. Marheineke (Berlin, 1832), Appendix: "Beweise für das Dasein Gottes," Vol. II, pp. 466–483, translated by E. B. Speirs and J. B. Sanderson as *Lectures on the Philosophy of Religion* (London: K. Paul, Trench, Trubner, 1895), Vol. III. pp. 347–367; and in *Vorlesungen über die Geschichte der Philosophie*, edited by K. L. Michelet (Berlin, 1836), Vol. III, pp. 164–169, translated by E. S. Haldane and F. H. Simson as *Lectures on the History of Philosophy* (London, 1896), Vol. III, pp. 62–67; Edward Caird, "Anselm's Argument for the Being of God," in *Journal of Theological Studies* 1 (1899):

23–39; E. E. Harris, *Revelation through Reason* (New Haven, CT: Yale University Press, 1958); Paul Tillich, "The Two Types of Philosophy of Religion," in his *Theology of Culture* (New York: Oxford University Press, 1959); and R. G. Collingwood, *An Essay on Philosophical Method* (Oxford: Clarendon Press, 1933). Collingwood's use of the argument was criticized by Gilbert Ryle in "Mr. Collingwood and the Ontological Argument," in *Mind*, n.s., 44 (174) (1935): 137–151. There is a reply to Ryle by E. E: Harris in *Mind*, n.s., 45 (180) (1936): 474–481), and a rejoinder by Ryle in *Mind*, n.s., 46 (181) (1937): 53–57. This discussion is reprinted in *The Many Faced Argument* (see above).

For the argument in recent French "reflexive" philosophy, see M. Blondel, *L'action* (Paris, 1893 and 1950), *L'étre et les ètres* (Paris, 1935), and *La pensée* (Paris, 1934 and 1948); J. Paliard, *Intuition et réflexion* (Paris, 1925); and A. Forest, "L'argument de saint Anselme dans la philosophie réflexive," in *Spicilegium Beccense*, pp. 273–294, translated by A. C. McGill in *The Many Faced Argument* (see above).

Some important contemporary philosophical discussions are C. D. Broad, "Arguments for the Existence of God," in *Journal of Theological Studies* 40 (157) (January 1939); J. N. Findlay, "Can God's Existence Be Disproved?," in *Mind*, n.s., 57 (226) (1948): 176–183, reprinted in *New Essays in Philosophical Theology* edited by A. G. N. Flew and Alasdair MacIntyre (London, 1958); Nicholas Rescher, "The Ontological Proof Revisited," in *Australasian Journal of Philosophy* 37 (2) (1959): 138–148; William P. Alston, "The Ontological Argument Revisited," in *Philosophical Review* 69 (4) (1960): 452–474; Norman Malcolm, "Anselm's Ontological Argument" (see above); Charles Hartshorne, *The Logic of Perfection* (see above); Jerome Shaffer, "Existence, Predication and the Ontological Argument," in *Mind*, n.s., 71 (283) (1962): 307–325, reprinted in *The Many Faced Argument* (see above); and a discussion by R. E. Allen, Raziel Abelson, Terence Penelhum, Alvin Plantinga, and Paul Henle in *Philosophical Review* 70 (1) (1961): 56–109.

On the notions of "existence" and "necessary existence," see Bertrand Russell, "The Philosophy of Logical Atomism" (1918), in his *Logic and Knowledge*, edited by R. C. Marsh (London: Allen and Unwin, 1956); G. E. Moore, "Is Existence a Predicate?," in *PAS*, Supp., 15 (1936), reprinted in *Logic and Language*, 2d series, edited by A. G. N. Flew (Oxford, 1959); William Kneale, "Is Existence a Predicate?," in *PAS*, Supp., 15 (1936), reprinted in *Readings in Philosophical Analysis*, edited by Herbert Feigl and Wilfrid Sellars (New York: Appleton-Century-Crofts, 1949); George Nakhnikian and W. C. Salmon, "'Exists' as a Predicate," in *Philosophical Review* 66 (4) (1957): 535–542; John Hick, "God as Necessary Being," in *Journal of Philosophy* 57 (22–23) (1960): 725–734; and Terence Penelhum, "Divine Necessity," in *Mind*, n.s., 69 (274) (1960): 175–186.

John Hick (1967)

ONTOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD [ADDENDUM]

Work on the ontological argument since 1970 has been mainly concerned with the so-called modal ontological argument for the existence of a perfect being.

THE CONCEPT OF A PERFECT BEING

Descartes defined a (supremely) perfect being as a being that possesses all perfections. But if a property F is a perfection, it would seem that a being that is F but might not have been F falls short of perfection. Hence a better definition of a perfect being would be as follows: a being that has all perfections and could not have lacked any perfection—a perfect being is a being that has all perfections *essentially* (has all perfections in every possible world in which it exists).

THE LOGICAL VALIDITY OF THE MODAL ONTOLOGICAL ARGUMENT

The argument has two premises: (1) A perfect being is possible (exists in some possible world); (2) Necessary existence (existence in every possible world) is a perfection.

Plantinga (1974) has shown that the existence of a perfect being is logically deducible from these two premises. (The proof presupposes the strongest system of modal reasoning, S5. [For more discussion on S5, see the entry "Modal Logic."] Here we assume without argument that a modal argument that is valid in no weaker system than S5 is not objectionable on that ground. For a contrary view, see Salmon [1989].) Suppose a perfect being exists in some possible world *w* [premise (1)]. This being is necessarily existent in *w* [premise (2)], and must therefore exist in every possible world, for if there were some world in which it did not exist, it would not be *necessarily* existent in *w*. This being has in *w* all perfections [premise (1)]. It must therefore have all perfections in every possible world in which it exists (that is, in every possible world), for if there were some world in which it existed but failed to have all perfections, it would not have all perfections *essentially* in *w*. This being therefore exists in the actual world and in every other possible world, and has all perfections in the actual world and in every other possible world. It is therefore necessarily existent in the actual world (if it were not necessarily existent in the actual world, there would be some world in which it did

not exist) and has all perfections essentially in the actual world (if it did not have all perfections essentially in the actual world, there would be some world in which it did not have all perfections). That is to say— there exists a perfect being.

THE PREMISES OF THE ARGUMENT

The conclusion of the argument follows (in S5) from its two premises. But are the two premises true? Critics of the argument are typically willing to grant premise (2) but see no reason to accept premise (1).

Plantinga has conceded that there seems to be no way to demonstrate the possibility of a perfect being. (And he recognizes that one may not simply presume that a concept is possible in the absence of a demonstration of its impossibility. So to presume can in fact lead one into contradiction, because there are pairs of concepts, neither of which can be *shown* to be impossible and at least one of which must *be* impossible. If it cannot be shown that a perfect being is impossible, the concept of a perfect being and the concept of a being who knows that there is no perfect being are such a pair.) Plantinga contends, however, that it is not irrational to believe that a perfect being is possible (just as it is not irrational to believe that a private language is possible or that free will is possible: a philosopher who believes in the possibility of these things is not *ipso facto* irrational). He further contends that it is not irrational to believe the demonstrated logical consequences of things that are not rational to believe, and that it is therefore not irrational to believe that there is a perfect being. He concludes that although the modal ontological argument is not a proof that a perfect being exists, its logical validity in effect constitutes a proof that it is not irrational to believe that a perfect being exists. This conclusion has been disputed by van Inwagen (1977).

GÖDEL'S POSSIBILITY PROOF

The most important recent attempt to prove that a perfect being is possible occurs in a brief note (unpublished in his lifetime) by Kurt Gödel ("Ontological Proof" in Fefferman, ed. [1995]). The argument (slightly modified) is this: Necessary existence and the "essentialization" of every other perfection (having that perfection essentially) are all positive properties, and any set of positive properties is consistent or possible because the set of all positive properties is possible. This last statement is a consequence of two "axioms": (1) The set of all positive properties is closed under entailment; and (2) If a property is positive, its negation is not positive. (A set of properties entails the property F if it is impossible for something to have all the properties in that set and to lack F. A set of properties is closed under entailment if it contains every property entailed by any of its subsets.)

PROOF. Suppose that the set of all positive properties is impossible or inconsistent. We show that this entails a contradiction. Since an impossible set of properties entails any property, the only set of properties that is both impossible and closed under entailment is the set of all properties: the set of all positive properties is the set of all properties. But the negation of a positive property is not a positive property: the set of all positive properties is not the set of all properties.

Unfortunately, Gödel's attempts to explain the idea of a positive property are compressed and cryptic. They leave the reader with no reason to suppose that there is a set of properties such that (1) necessary existence and the essentialization of every other perfection are members of that set, (2) that set is closed under entailment, and (3) if a property is member of that set, its negation is not. The modal ontological argument therefore remains inconclusive.

See also Degrees of Perfection, Argument for the Existence of God; Descartes, René; Gödel, Kurt; Gödel's Theorem; Modal Logic; Plantinga, Alvin.

Bibliography

Feferman, Solomon; John W. Dawson Jr., Warren Goldfarb, Charles Parsons, Robert N. Solovay, eds. *Kurt Gödel: Unpublished Essays and Lectures Collected Work*. Vol. 2 of *Collected Works*. Oxford: Oxford University Press, 1995.

Gödel, Kurt. "Ontological Proof." In Feferman *et al.*, 303–304.

Plantinga, Alvin. *God, Freedom, and Evil*. New York: Harper & Row, 1974.

Plantinga, Alvin. *The Nature of Necessity*. Oxford: Clarendon Press, 1974.

Salmon, Nathan. "The Logic of What Might Have Been." *The Philosophical Review* 97 (1989): 3–34.

van Inwagen, Peter. *Metaphysics*. 2nd ed. Boulder, CO: Westview, 2002.

van Inwagen, Peter. "Ontological Arguments." *Noûs* 11 (1977): 375–395.

Peter van Inwagen (2005)

ONTOLOGY

Ontology is the most general science or study of Being, Existence, or Reality. An informal use of the term signifies what, in general terms, a philosopher considers the world to contain. Thus it is said that Descartes proposed a dual-

ist ontology, or that there were no gods in d'Holbach's ontology. But in its more formal meaning, ontology is the aspect of metaphysics aiming to characterize Reality by identifying all its essential categories and setting forth the relations among them.

BEING QUA BEING

Existence, as the most comprehensive category of all, should embrace members with the least in common. Nevertheless, Western philosophy long sought some substantive common content present in anything just in virtue of its existence. The history of these attempts to identify the common character of being qua being is not encouraging.

In *The Sophist*, Plato's Eleatic Stranger proposes that a role in the world's causal network is the necessary and sufficient condition for existence, that "Power is the mark of Being." This idea has had some currency in the twentieth century, particularly in the work of David Lewis (1986) and D. M. Armstrong (1978, 1989, 1997). This *Eleatic principle* is an attractive test for reality in the natural world, for whatever is real in nature should be able to make a difference. It might be necessary to weaken the requirement and admit a passive space-time that provides the arena within which the active beings exert themselves. Even so, the Eleatic principle seems to be at best a contingent aspect of the world because there seems to be no impossibility involved in the idea of a completely inert being. It also begs the question against abstract entities such as numbers, or geometric points, or sets, which, if they exist, lie outside the causal nexus.

For Samuel Alexander (1920), to be is to be the exclusive occupant of a volume of space-time. This rules out not only abstract entities, but even a field theory of the natural world, for force fields occupy regions of space-time, yet do not exclude one another.

J. M. E. McTaggart (1921–1927) argued that the mark of being is to stand in a determining correspondence with all of one's infinite parts. A determining correspondence ensures that from a sufficient description of anything, a sufficient description of any of its parts can be derived. This requirement implies that space, the natural world, and most of the contents of minds are unreal. From this consequence the conclusion to be drawn is that McTaggart's proposed mark of being is excessively demanding.

The problem of a substantive content for being qua being is reflected in the idiosyncratic behavior of the verb "to exist." Consider singular negatives: "Aristotle does not speak Spanish" is true because the predicate "does not speak Spanish" applies to the item referred to by the subject term. But "Pegasus does not exist" cannot be true because its predicate applies to the item referred to by the subject term. If the subject term refers to anything, that item exists, which would make the whole statement false.

Kant famously declared that existence is not a property, and this view has become widely accepted. The modern logic that descends from Gottlob Frege and the *Principia Mathematica* (1910–1913) of Alfred North Whitehead and Bertrand Russell replaces all expressions using "exists" with others using "There are." Thus, "Lions exist" becomes "There are lions," while "Dragons do not exist" becomes "There are no dragons."

In technical terms, this process replaces any existence claim with one using a quantifier ranging over a domain (the world), so that to exist becomes a matter not of possessing the special property *existence*, but of possessing some other, ordinary, properties. The determination to restate all claims to existence or nonexistence with "There are …" and "There are no …" is expressed in W. V. Quine's dictum: "To be is to be the value of a variable."

If existence is not a property, it cannot be a perfection. This undercuts those versions of the ontological argument for the existence of God that rely on existence being among the perfections. A recent response has been to argue that, even if existence is not a property, necessary existence is (Plantinga 1974, 1975; van Inwagen 1993).

REALITY AND ACTUALITY

Is existence all there is, or should we recognize categories even broader that that of Being? In Plato, and even earlier, is to be found the distinction between Reality (What is) and Appearance (What is not nothing, yet only seems to Be). Aristotle distinguishes the fully existent (Being), from that which is still in formation (Becoming). These distinctions are perhaps best seen as advocating different grades of reality within the one category of Being.

Aristotle also distinguishes the fully Real (Act) from that which may be (Potency). This distinction is the forerunner of a strong strand in ontology that recognizes possible worlds in addition to the actual world, the one we inhabit. In the Neoplatonists, and again in Alexius Meinong, the realm of the existent is augmented by that of the subsistent, which encompasses what does not exist although it might have done so, such as golden mountains.

A full-scale ontology of this kind, in which the realm of Essence is wider than that of Existence, was presented

by James K. Feibleman in 1951. In the work of Richard Sylvan (1980), this is extended even further. In Sylvan's system, the individual variables range over not only the actual and the possible, but the impossible as well.

POSSIBLE WORLDS. Gottfried Wilhelm Leibniz was the first to make systematic use of the idea that all the possibilities can be regarded as forming worlds—each a complete internally consistent realm that may combine some elements matching the actual world with others in which it differs. The actual world is one of the possible worlds, distinguished from all others by the fact that none of its elements is merely possible. If one is able to refer to possible worlds, it is easy to define *necessary beings*, otherwise so difficult to characterize, as those present in all possible worlds (see below).

MODAL REALISM. Possible worlds make available explanations of causal powers, of counterfactual conditionals, of unexercised dispositions, and of real uninstantiated properties. Such advantages led David Lewis (1986) to embrace modal realism, which affirms the literal reality of all possible worlds.

Other philosophers, while appreciating these advantages, have balked at the apparently infinite expansion of the ontology that this requires. This has led to accounts of ersatz possible worlds: Rudolf Carnap and others have proposed that a possible world is a maximally consistent set of sentences. Armstrong and others have developed Wittgenstein's idea that a possible world is a nonactual recombination of the elements of this world. Peter Lopston (2001) advances a reductive realism, which expands the kind of property assigned in the actual world to include might-have-had features. The success of these approaches is subject to continuing controversy.

MANY WORLDS IN QUANTUM THEORY. The notion that the world we live in is not the only one has also been canvassed recently in the interpretation of some otherwise baffling paradoxes in quantum physics. On these accounts the world is not a single unified entity, but one subject to continual bifurcation, a process that generates an ever-increasing number of worlds. Many-world views of this kind are in an important way different from modal realism: all these quantum worlds are supposed to be actual but mutually inaccessible.

THE CATEGORIES OF BEING

The principal task of ontology is to furnish an inventory of the categories, the most general divisions of Reality. The most important of these are:

SUBSTANCES. An individual or particular substance is an object, a thing in its own right. Common everyday things, such as bricks and bedsteads, provide a model for the category of substance. Substances are required to have several basic features, although it is not clear that these features are compatible with one another.

Particularity and individuality. A substance is both a particular and an individual; not just some duck or other, but this very duck. An object is of the *kind* it is (a duck) on account of its properties. But if these properties are universals, shared by many particulars, they cannot themselves confer particularity. Some philosophers, most notoriously Locke, proposed a constituent of substances that would perform this role, a *substratum* that would confer both particularity and individuality. A substratum would be a bare particular, an item inherently particular and individual, yet without any other feature. It is difficult to see how such bare particulars could be distinguished from one another, but if bare particulars are all exactly alike, how could any one of them individualize its own substance? More generally, bare particulars conflict with Aristotle's dictum that the minimum of being, the least thing there can be, is a "this-such," a particular having a property.

Another proposal is that substances are individuated by their location. Locations—space-time points and regions—are themselves unique particulars; if they can have primitive particularity, that raises the question why other particulars require a substratum or other particularizer. There are other difficulties with location also: Location will not individuate force fields or other physical entities that do not monopolize their space. It fails also for any items of an immaterial kind.

Either individuality—and hence particularity—are primitive, or there are bare particulars, or each substance has a special property, known as *haecceity* or *thisness*, which can bestow particularity and individuality on its bearer. For a discussion see chapter fifteen of John Heil's *From an Ontological Point of View* (2003).

Indivisibility. Individual substances must be distinguished from compounds, so a single substance must be indivisible, in the sense that it has no parts that are themselves substances. This disqualifies ordinary things as individual substances. This simplicity requirement is much emphasized in Aquinas's doctrine of God. It leads

in Leibniz to the monadology, and in Roger Joseph Boscovich to the doctrine of material points.

Persistence. Substances are distinguished from their properties by a capacity to persist, that is, to retain their identity through at least some changes. A fire truck can change in color, yet remain the fire truck it always has been. The ordinary compound substances of everyday life have some persistence, but cannot survive all changes. A fire truck dismantled and scrapped is no longer a fire truck. Complete persistence belongs only to the fundamental substances.

Independence. Any substance could be the only thing in existence. If this independence is interpreted causally, no ordinary object is a substance, for they are all brought into being, and hence depend for their existence on their causes. Space-time and its fields might qualify, yet even these depend, in theistic systems, on the creative activity of God. So in Thomism, God is the substance par excellence, but the natural world includes created substances, dependent on God, but otherwise existent in their own right. Spinoza, insisting on absolute independence, concluded that there can be only one substance, the all-embracing totality, God-or-Nature.

If one takes the independence of substances in a logical, rather than a causal sense, a substance is anything that, in principle, could stand alone. This was David Hume's requirement, and anything meeting it is a Humean substance. For compounds, the requirement is that the thing, including all its parts, could exist alone. This requirement is much less rigorous than causal independence and requires no persistence.

No-substance theories. There have been attempts to dispense with substances. Russell has proposed that an ordinary concrete object is no more than a bundle of all its properties. There is always an issue over what it is that binds the bundle. Moreover, as the properties are universals, this theory implies that no two things can be exactly alike.

In Donald Williams's version of the bundle theory (1966), the properties are particular instances or tropes (see below). This avoids any problem with the possibility of there being two exactly resembling objects, but it requires that all members of the bundle be "compresent"—all at the same place in space-time. There are difficulties in treating a space-time location as just one further trope in the bundle, but if it is given special treatment it becomes a substantializing substratum.

Russell also advocated an event ontology as a no-substance view. He used "event" for the occurrence of a property at a place and a time; such events are not happenings, but states of affairs (see below). He proposed that ordinary substances, and their more fundamental parts, are sequences of clusters of such events.

The basic elements in these ontologies may not be simple or indivisible, and they lack persistence. Nevertheless, these states of affairs or events are Humean substances. Indeed, unless there is nothing at all, something must be a Humean substance, and in that sense, any *no substance* theory must fail.

PROPERTIES AND RELATIONS. Properties are the intrinsic features or characteristics of things, which belong to them considered singly. Relations, involving two or more terms, are the ways in which things stand to one another. In many respects, properties and relations can be treated together.

Properties as universals. Properties are usually thought of as universals that can characterize indefinitely many instances. There is but one Eiffel Tower, but the tower's height, weight, and iron constitution are features it has in common with many other things. The Problem of Universals is the problem of explaining how any one real entity could possibly exist, fully and completely, in many different instances. This problem has attracted three different proposed solutions: nominalism, conceptualism, and realism. Nominalism and conceptualism both deny that properties are genuinely universal. According to nominalism, the only element common to all iron things is that they can all be described using the predicate "iron," or all are members of the class *iron things*, or all resemble some typical iron objects. According to conceptualism, the universal element consists in an impulse of our minds to group several things together. These reductive theories have had adherents since the time of Plato and were particularly prevalent among the British Empiricists and their descendents. Nominalism and conceptualism were explicitly challenged by Russell in *Problems of Philosophy* (1912). The most thorough case against such views is presented in D. M. Armstrong, *Universals and Scientific Realism* (1978).

Realism regarding universals is at least as old as Plato. His theory of Forms presents a thoroughgoing realism that accords to genuine properties both a real existence, in a realm of their own, and a status superior to any this-worldly instantiations of them there may be. The Forms exist *ante rem*—that is, whether or not they are instantiated. The traditional account of Aristotle ascribes to him a modified realism, according to which properties are real, and universal, but can exist only *in rebus*, as the

properties of concrete instances. Here one encounters again his view that the least that is "apt for being" is a this-such, a union of particular with universal.

Realism has always faced two principal objections. First, that it is uneconomical, especially in its Platonist form. The question of economy is a current issue in the philosophy of science, as it at least appears that our best physical and chemical theories involve uninstantiated properties. The second objection is that it can provide no coherent account of the link between a property and the substance that bears it, the *inherence* relation. Inherence cannot be a normal relation, for then it is just one further universal standing in need of an inherence link to its terms, the substance and the original property. But if it is not a relation in the ordinary sense, what is it? The problem with inherence lends support to versions of realism in which properties are particulars.

Properties as particulars. Even if the property *iron* is universal, the particular case of *being iron* that occurs in the Eiffel Tower belongs to the tower alone and is as particular as the tower itself. Trope theory, as developed first by Donald Williams, treats the instance not as a dependent entity arising from the instantiation of a universal, but as a Humean substance in its own right.

When this approach is coupled with a bundle or compresence account of ordinary many-featured substances, the problem of any inherence relation disappears. There is a further significant economy, for there is no need for a separate category of substance. These possibilities are explored further in Keith Campbell's book *Abstract Particulars* (1990).

Relations. When Russell reanimated the realism debate he accorded to relations a status fully equal to that of inherent properties. Indeed, it was his reflections on the role of relations in the foundations of mathematics and of logic that led him to his realism. Armstrong's realism takes the same form.

There is, nevertheless, a long tradition that accords primacy to the intrinsic properties. Aristotle held that relations are "the least of the things that are"; Hobbes and others held that the existence of relations depends on a mental act of comparison; and Leibniz's view was that every relation has its foundation in an intrinsic feature of one or both of its terms. This reductive program is expounded in Campbell (1990).

Relations do seem to be dependent in the sense that they must have substances as their terms, and these substances must have intrinsic properties. So unless there are intrinsic properties there can be no relations, but not vice versa. Bundle theories of ordinary things concern only the intrinsic properties. To include relations in the bundles leads to problems over where to assign the relations, and this in turn induces a tendency towards a monism such as Francis Herbert Bradley's, in which ordinary substances are absorbed into a single all-embracing totality.

Powers. Some properties, such as *square*, seem to belong to how an object is. Others, such as *being a solvent*, seem to refer to what an object can do. This is the distinction between categorical and dispositional properties. One line of thought takes up the Eleatic principle, and identifies real properties as those that confer on their bearer a disposition to act or to be acted upon. Such dispositions are *powers*; a metaphysic of powers is set forth in George Molnar's *Powers* (2003) and in Brian Ellis's *Scientific Essentialism* (2001).

COMPLEXES. *Substance* and *property* are basic categories. In combination, they can provide a richer ontology.

States of affairs. A basic state of affairs consists in a particular having a property, or in two (or more) particulars standing in a relation. A single property inhering in a single particular is a minimal "this-such." Wittgenstein's *Tractatus Logico-Philosophicus* (1921) presented an ontology in which the world is composed of minimal relational states of affairs: those that actually obtain being facts, those merely possible being the remaining states of affairs. These themes—that the basic categories only ever occur in combination, and that these combinations constitute reality—are taken up in D. M. Armstrong's *A World of States of Affairs* (1997).

Events and processes. A state of affairs is static. To account for the dynamic aspects of the world requires an account of change. This can be done by using sequences of states of affairs: stability consists in successive states of affairs closely resembling one another, whereas change consists in the states of affairs at one time being replaced by others systematically different. An event is a single change, involving a pair of states of affairs; a process is a more complex series of events.

Whitehead, in *Process and Reality* (1929) accorded priority to the dynamic; all apparently persisting substances are actually slowly evolving processes. The status of space-time is still controversial. It may be a Humean substance; however, some accounts of matter assign it a place as a process, a sequence of complex, changing relations between particulars.

ABSTRACT OBJECTS

Human thought, particularly in mathematics and logic, seems to involve entities that have no apparent place in the natural spatiotemporal world, and no causal role. To admit such items challenges the principle of economy, yet successful reductions are difficult to accomplish.

NUMBERS AND SETS. Because all numbers can be represented in set theory, there is no need to admit both sets and numbers. Russell had proposed to eliminate sets in favor of propositional functions, but this proved impossible for more than a fragment of mathematics (Goodman and Quine 1947, Quine 1969). Because the variables of set theory have sets as their values, and to be is to be the value of a variable, we are committed to their reality—this is Platonism about sets and numbers. The most important work in attempting to avoid Platonism is Hartry Field's (1980, 1989).

GEOMETRICAL OBJECTS. Unlike anything in the natural world, the objects of geometry—Euclidian cubes, for example—are thought of as perfect, changeless, timeless, and without any physical causal powers. Moreover, there are geometries, and corresponding geometrical objects, with many more dimensions than this world has. A geometrical space can be divided and subdivided into an infinity of different shapes of different sizes. Platonism in geometry thus involves an infinite expansion in ontology.

One approach to this issue is to consider geometrical objects as abstracted objects, that is, objects taken from a context. On this view, every cube is just a particular spatial fragment of space-time and every triangle a fragment of one of space-time's spatial surfaces. One problem with this is that not all shapes will be available. If our space-time is nowhere perfectly Euclidean, there will be no real Euclidean cubes. We can treat these nonexistent objects as imaginary variations on the actually existing ones, and geometries that quantify over such things, as not literally true.

LOGIC. The philosophy of logic makes reference to propositions, operators, functions, and inferences. These are abstract entities, related to reasoning in much the same way as numbers are related to counting and measuring. The problems and prospects of a reductive treatment of them are also parallel.

NECESSARY BEINGS

Ordinary things are usually held to exist contingently; that is, they do exist, but might not have. Had our world's initial conditions or laws of nature been different, there would have been a different group of contingent beings. But some things seem to be immune from the vagaries of cause and chance; being outside the causal net, they cannot be brought into being and cannot be destroyed. These are "necessary beings." If Platonism is correct regarding any of the abstract objects, these will be necessary beings, even, paradoxically, the null class.

For Aristotle, anything that exists through an infinite time is necessary because he held that over infinite time every possibility would at some point be actualized. For Plotinus, any divine being would be outside time, and as such could not change, could not cease to exist, and thus would be A necessary being. For Aquinas, God's necessity derives from his simplicity: God's essence and his existence are identical; in this way he is a kind of being that must exist. For Spinoza, every genuine substance is *causa sui*, containing within itself the sufficient explanation for its own being, and thus it can guarantee its own existence under all possible conditions.

Duns Scotus, then Descartes, linked necessary being with logic: A necessary being is one, the denial of whose existence would be self-contradictory. "Real"—i.e., existing—"beans do not exist" is a self-contradiction, but only trivially because existence has been inserted into the definition of the subject. This does not make beans necessary beings. If existence is not inserted into the subject term's definition, it is doubtful whether any denial of existence would be a self-contradiction. The best discussion of necessary being is in Alvin Plantinga (1974, 1975).

See also Metaphysics.

Bibliography

Adams, R. M. "Primitive Thisness and Primitive Identity." *The Journal of Philosophy* 76 (1979): 5–26.

Alexander, Samuel. *Space, Time, and Deity: The Gifford Lectures at Glasgow*. 2 vols. London: Macmillan, 1920.

Armstrong, D. M. *A Combinatorial Theory of Possibility*. Cambridge, U.K.: Cambridge University Press, 1989.

Armstrong, D. M. *Universals and Scientific Realism*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1978.

Armstrong, D. M. *A World of States of Affairs*. Cambridge, U.K.: Cambridge University Press, 1997.

Campbell, Keith. *Abstract Particulars*. Oxford: Blackwell, 1990.

Ellis, Brian. *Scientific Essentialism*. Cambridge, U.K.: Cambridge University Press, 2001.

Feibleman, James K. *Ontology*. Baltimore: Johns Hopkins Press, 1951.

Field, Hartry. *Realism, Mathematics and Modality*. Oxford: Blackwell, 1989.

Field, Hartry. *Science without Numbers: A Defence of Nominalism*. Oxford: Blackwell, 1980.

Forbes, Graeme. *The Metaphysics of Modality*. Oxford: Clarendon Press, 1985.

Goodman, N., and W. V. Quine. "Steps Towards a Constructive Nominalism." *Journal of Symbolic Logic* 12 (1947): 105–122.

Heil, John. *From an Ontological Point of View*. Oxford: Clarendon Press, 2003.

Lewis, David K. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.

Loptson, Peter. *Reality: Fundamental Topics in Metaphysics*. Toronto: University of Toronto Press, 2001.

McTaggart, J. M. E. *The Nature of Existence*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1921–1927.

Molnar, George. *Powers: A Study in Metaphysics*. New York: Oxford University Press, 2003.

Plantinga, Alvin. *God, Freedom, and Evil*. London: Allen and Unwin, 1975.

Plantinga, Alvin. *The Nature of Necessity*. Oxford: Clarendon Press, 1974.

Quine, W. V. *From a Logical Point of View: 9 Logico-Philosophical Essays*. 2nd ed. Cambridge, MA: Harvard University Press, 1961.

Quine, W. V. *Set Theory and Its Logic*. Cambridge, MA: Belknap Press of Harvard University, 1969.

Russell, Bertrand. *The Analysis of Matter*. London: Kegan Paul, Trench and Trubner, 1927. Event ontology.

Russell, Bertrand. *An Inquiry into Meaning and Truth*. London: Allen and Unwin, 1940. Substances as bundles of universals.

Russell, Bertrand. *The Problems of Philosophy*. London: Oxford University Press, 1912. Realism about universals

Sylvan, Richard. *Exploring Meinong's Jungle and Beyond*. Canberra: Australian National University, 1980.

Van Inwagen, Peter. *Metaphysics*. Boulder, CO: Westview Press, 1993.

Whitehead, A. N. *Process and Reality: An Essay in Cosmology*. Cambridge, U.K.: Cambridge University Press, 1929.

Williams, Donald C. *Principles of Empirical Realism: Philosophical Essays*. Springfield, IL: Thomas, 1966.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. London: Routledge and Kegan Paul, 1921.

Keith Campbell (2005)

ONTOLOGY, HISTORY OF

The term *ontologia* was coined by scholastic writers in the seventeenth century. Rudolf Goclenius, who mentioned the word in 1636, may have been the first user, but the term was such a natural Latin coinage and began to appear so regularly that disputes about priority are pointless. Some writers, such as Abraham Calovius, used it interchangeably with *metaphysica*; others used it as the name of a subdivision of metaphysics. Johannes Clauberg (1622–1665), a Cartesian, coined instead the term *ontosophia*. By the time of Jean-Baptiste Duhamel (1624–1706), ontology was clearly distinguished from natural theology. The other subdivisions of metaphysics are cosmology and psychology, from which ontology is also distinguished. Thus, *ontologia* as a philosophical term of art was already in existence when it was finally canonized by Christian Wolff (1679–1754) and Alexander Gottlieb Baumgarten (1714–1762).

WOLFF

For the authors mentioned above, the subject matter of ontology was being as such. "Being" was understood univocally, as having one single sense. Ontology can therefore claim as ancestors John Duns Scotus and William of Ockham, rather than Thomas Aquinas. In the case of Wolff himself, Gottfried Wilhelm Leibniz was a stronger influence than scholasticism, but in his *Philosophia Prima Sive Ontologia,* Wolff refers explicitly to Francisco Suárez. According to Wolff, the method of ontology is deductive. The fundamental principle applying to all that is, is the principle of noncontradiction, which holds that it is a property of being itself that no being can both have and not have a given characteristic at one and the same time. From this, Wolff believed, follows the principle of sufficient reason, namely, that in all cases there must be some sufficient reason to explain why any being exists rather than does not exist. The universe is a collection of beings each of which has an essence that the intellect is capable of grasping as a clear and distinct idea. The principle of sufficient reason is invoked to explain why some essences have had existence conferred on them and others have not. The truths about beings that are deduced from indubitable first principles are all necessary truths. Thus, ontology has nothing to do with the contingent order of the world.

The influence of late scholasticism (or of what Étienne Gilson calls "essentialism") on rationalist metaphysics was repaid in kind, for the division of metaphysics into ontology, cosmology, and psychology found its way back into scholastic manuals, where it has persisted until very recently. Along with this division, there persisted the view that being constitutes an independent subject matter over and above the subject matter of the special sciences. The persistence of this view is perhaps to be explained by cultural rather than by intellectual factors. In the eighteenth and nineteenth centuries scholasticism was isolated in seminaries until Pope Leo XIII guided Thomism back into intellectual debate. Only in this way was scholasticism able to avoid the nemesis (in the form of Immanuel Kant) that awaited rationalist metaphysics.

KANT

In the written announcement of lectures given from 1765 to 1766, Kant treated ontology as a subdivision of metaphysics that included rational psychology but was distinguished, in his case, from empirical psychology, cosmology, and what he called the "science of God and the world": "Then in *ontology* I discuss the more general properties of things, the difference between spiritual and material beings." But when Kant came to write the *Critique of Pure Reason,* he settled matters with ontology once and for all. The two key passages are the discussion of the second antinomy of pure reason and the refutation of the ontological argument. Wolff had argued a priori that the world is composed of simple substances, themselves neither perceived nor possessing extension or shape, and each of them different, and that physical objects are composite, collections of such substances. In the second antinomy the thesis is that "every composite substance in the world consists of simple parts, and nothing exists anywhere that is not either simple or composed of simple parts"; and the proof that Kant presented is effectively Wolffian. But he presented an equally powerful proof for the antithesis, namely, that "no composite thing in the world consists of simple parts, and there exists nothing simple anywhere." In exposing the shared fallacy of both proofs, Kant made it impossible ever again to accept ontology as a deductive body of necessary truths that is akin to geometry in form but has being as its subject matter. His analysis of existence in his refutation of the Ontological Proof is a counterpart to this.

Since Kant, the most influential use of the term *ontology* outside scholastic manuals has been in the writings of Martin Heidegger and W. V. Quine. Both have been greeted by scholastic writers as engaged in essentially the same enterprise as they themselves, Father D. A. Drennen taking this view of Heidegger, and Father I. M. Bocheński of Quine.

HEIDEGGER

In regard to Heidegger's ontology, Father Drennen is perhaps partly correct. Heidegger wished to explain what character being must have if human consciousness is to be what it is. He began by quarreling with the principle of sufficient reason in its Leibniz-Wolff form. This, he said, is an inadequate starting point for ontology because the question "Why is there something rather than nothing?" presupposes that we already know what being and nothing are. Heidegger treated "Being" and "Nothing" as the names of contrasted and opposed powers whose existence is presupposed in all our judgments. In negative judgments, for example, to speak of what is not the case is implicitly to refer to Nothing. Heidegger's ontology, however, was not deductive or even systematic in form. It proceeds at times by the exegesis of poetry or of the more aphoristic fragments of the pre-Socratic philosophers and is thus very different from scholastic ontology.

QUINE

In the case of Quine, the name *ontology* has been in fact given to a quite different set of preoccupations. Quine has been concerned with two closely allied questions: To the existence of what kind of thing does belief in a given theory commit us? And what are the relations between intensional and extensional logic? His answer to the first question is that to be is to be the value of a variable: We have to admit the existence of that range of possible entities for which names could occur as values for those variables without which we could not state our beliefs. His answer to the second question is that intensional logics and extensional logics involve the admission not merely of different but of incompatible types of entity. "Both sorts of entity can be accommodated in the same logic only with the help of restrictions such as Church's, which serve to keep them from mixing, and this is very nearly a matter of two separate logics with a universe for each" (*From a Logical Point of View,* p. 157).

It is clear that Quine's logical preoccupations are in fact relevant to Wolff and the scholastics only in that an understanding of Quine's inquiries would preclude one from trying to construct a deductive ontology in the mode of Suárez or Wolff.

See also Baumgarten, Alexander Gottlieb; Church, Alonzo; Clauberg, Johannes; Cosmology; Gilson, Étienne Henry; Heidegger, Martin; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Ontology; Psychology; Quine, Willard Van Orman; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Wolff, Christian.

Bibliography

PRIMARY SOURCES

Baumgarten, Alexander Gottlieb. *Metaphysica.* Halle, 1740.

Clauberg, Johannes. *Opera Omnia,* edited by J. T. Schalbruch. 2 vols., 281. Amsterdam, 1691.

Duhamel, Jean-Baptiste. *De Consensu Veteris et Novae Philosophiae.* Paris, 1663.

Duns Scotus, John. *Opera Omnia.* 12 vols. Paris, 1891–1895. Vol. III, *Quaestiones Subtillissimae Super Libros Metaphysicorum Aristotelis.*

Heidegger, Martin. *Being and Time.* Translated by John Macquarrie and Edward Robinson. New York: Harper, 1962.

Heidegger, Martin. *Existence and Being.* Translated by D. Scott, R. Hall, and A. Crick. Chicago: Regnery, 1949.

Kant, Immanuel. *Critique of Pure Reason.* Translated by Norman Kemp Smith. London: Macmillan, 1929.

Quine, Willard Van Orman. *From a Logical Point of View.* Cambridge, MA: Harvard University Press, 1953.

Wolff, Christian. *Philosophia Prima Sive Ontologia.* Frankfurt and Leipzig, 1729.

Wolff, Christian. *Philosophia Rationalis, Sive Logica Methodo Scientifica Pertractata et ad Usum Scientiarum Atque Vitae Aptata.* Frankfurt and Leipzig, 1728.

SECONDARY SOURCES

Bocheński, I. M. *Philosophy—An Introduction.* Dordrecht, Netherlands, 1962.

Drennen, D. A., ed. *Modern Introduction to Metaphysics.* New York: Free Press of Glencoe, 1962.

Ferrater Mora, José. "On the Early History of Ontology." *Philosophy and Phenomenological Research* 24 (1963): 36–47.

Geach, P. T., A. J. Ayer, and W. V. Quine. "Symposium: On What There Is." *PAS,* Supp., 25 (1951): 125–160.

Gilson, Étienne. *Being and Some Philosophers.* Toronto: Pontifical Institute of Mediaeval Studies, 1949.

Martin, Gottfried. *Immanuel Kant: Ontologie und Wissenschaftstheorie.* Cologne: University of Cologne, 1951. Translated by P. G. Lucas as *Kant's Metaphysics and Theory of Science.* Manchester, U.K., 1955.

Marx, Werner. *Heidegger und die Tradition.* Stuttgart: W. Kohlhammer, 1961.

Owens, Joseph. *The Doctrine of Being in the Aristotelian Metaphysics.* Toronto: Pontifical Institute of Mediaeval Studies, 1951.

Alasdair MacIntyre (1967)

OPERATIONALISM

"Operationalism" is a program that aims at linking all scientific concepts to experimental procedures and at cleansing science of operationally undefinable terms, which it regards as being devoid of empirical meaning. Scientists adopted the operational approach to their subject before the principles of operationalism were made articulate. Operationalist theory was erected not on the basis of independent philosophical considerations but upon what was already implicit in the working practice of scientists. P. W. Bridgman, the Nobel Prize–winning physicist who is commonly regarded as the founder of operationalism, emphasized this point when he said, "it must be remembered that the operational point of view suggested itself from the observation of physicists in action" ("The Present State of Operationalism," in *The Validation of Scientific Theories,* edited by Philipp Frank, Boston, 1956, p. 79).

A fairly nontechnical illustration of the kinds of development in science in which one can discern an implicit operational point of view is the manner in which physicists treated the concept of physical length. In the nineteenth century it was discovered that Euclid's geometry was not logically unique and that other geometries based on different axioms were not necessarily internally inconsistent. The question was raised about the nature of physical space. Do lines and figures in physical space obey the theorems of Euclid?

At first sight this seems a perfectly sensible question to which there must be a definite answer. Even today some amount of sophistication is required to ask whether we have a clear notion of what could be done to find out whether space has a certain set of properties. Unless we can give an affirmative answer to this question, we should not take it for granted either that space has or that it lacks certain geometrical properties. By the end of the nineteenth century, however, scientists had accepted the view that if we cannot devise operations that would disclose whether or not space was Euclidean, then no definite geometrical properties can be assigned to space at all.

It is clear that in order to determine the geometrical properties of physical figures we must be able to compare distances. If we are unable to say whether distance *AB* is greater, smaller, or equal to distance *CD,* where *AB* and *CD* do not lie alongside one another, then we cannot even begin to investigate the geometrical nature of space. We take it for granted, however, that in order to compare distances we need a rigid measuring rod, that is, a rod which can be relied upon not to change in length while being transported from place to place. But the question whether the lengths of transported rods are preserved cannot be settled unless we presuppose the possession of some other standard of measurement to which these rods could be compared, but it is agreed that the sole standard of length is a rigid rod. Thus, there are no rigid rods except by fiat, and distances consequently cannot be spoken of as being objectively equal or unequal to one another, and the nature of space cannot uniquely be determined. From an operational point of view, therefore, space has no intrinsic metric, and it is a matter of convention whether we say space obeys this or that set of geometrical axioms.

THE OPERATIONALIST THESIS

Although the idea that physical entities, processes, and properties do not have an independent existence transcending the operations through which we may ascertain their presence or absence played an influential role in the thoughts of scientists before the 1920s, it was not until

1927 that Bridgman, in his celebrated *Logic of Modern Physics,* stated operationalism as an explicit program, made an articulate case for it, and undertook extensive operational analyses of the foundations of numerous physical concepts.

Bridgman soon had to retreat from his first extreme statement of operationalism. He had maintained that every scientifically meaningful concept must be capable of full definition in terms of performable physical operations and that a scientific concept is nothing more than the set of operations entering into its definition. The untenability of this view was quickly noticed—for example, by L. J. Russell, who in 1928 pointed out that in science one often speaks of certain operations as being better than others and that one cannot do so except in relation to something existing over and above them. Moreover, useful physical concepts do not as a rule lend themselves to an exhaustive definition. Any connection they have with instrumental operations may be loose and indirect: statements in which the concepts appear may, in the context of a set of other statements (but not on their own), entail statements describing physical operations. Consequently, in his later writings Bridgman freely permitted "paper and pencil operations," by which he meant mathematical and logical maneuverings with the aid of which no more is required of a concept than that it should be "indirectly making connection with instrumental operations."

It is not hard to see how by taking as one's model a physical concept like the length of a body one arrives at Bridgman's original position. But suppose someone objected that the stepping-off procedure carried out by measuring rods is not the only way to compute the length of a body. We may, for example, define it equally well in terms of the result obtained by timing the body's oscillation when it is allowed to swing as a pendulum and by using the well-known equation connecting the length with the period of oscillation. Length, after all, may enter into all sorts of relationships with other physical parameters, some of which we perhaps have not yet discovered.

To this objection it would have been replied that there is a fundamental difference between the ways in which the two sets of operations are related to the concept of length. The length of a body is "synonymous" with the number of times one can lay a rigid standard of length alongside it; when we speak of the length of a body we mean no more nor less than the number obtained through the stepping-off procedure performed by a measuring stick. When, however, we time a pendulum and then make the appropriate calculations, we merely measure length indirectly, via the relationship of length to other physical parameters. The second approach does not define length but rather inserts the already defined concept of length into an equation accepted as representing a genuine physical relationship.

It is much more difficult to maintain this distinction in the case of such concepts as temperature. One way to give an operational definition of temperature is in terms of measurements made by a mercury thermometer; another way is in terms of measurements made by a platinum-wire thermometer. The first way relies on the theory that the length of bodies varies with temperature; the second, on the theory that electrical resistance varies with temperature. It is easy to see that the concept of temperature is no more than partially interpreted through each of these, and doubtless other, sets of operations to which it is linked by relevant theories. This same position has become generally adopted toward all physical concepts.

We may thus distinguish three stages in scientific theorizing. In the first, preoperational stage, the universe was thought to contain many things and processes that transcend our theories about them and the operations and manipulations through which we may catch a glimpse of them in the mirror of experience. In the second, "naive" operational stage, the other extreme was taken, and all the terms of science were regarded as no more than abbreviations for our experimental results. In the third stage, scientific terms are still not regarded as standing for things and processes having an independent existence of their own, but the meaning of scientific terms is given by a more or less elaborate system of empirical theories in which the terms appear, together with the observations on which the theories embodying the terms are grounded. It is recognized that the concepts of science can never be fully grasped as long as the theories which contain them are open to further development.

The three stages in scientific theorizing are perhaps more dramatically accentuated in psychology than in the physical sciences. Until the early twentieth century the prevailing view was that psychology is a unique discipline dealing with a very special class of events, processes, and entities: the constituents of the realm of consciousness, to which no one but the experiencing individual has access. Although this realm is out of the reach of objective public operations and experimentations, many theorists regarded it as real—indeed, as more real than anything else—and believed that it should be studied by a unique method, introspection.

The radical behaviorism that replaced this mentalistic psychology is a form of naive operationalism and is based on the tenet that psychology is the study not of mental events, processes, or entities but of behavior. Psychologists were not to be concerned with publicly unobservable phenomena, and introspection—at best a private method of inquiry—was completely outlawed.

Today, in the third stage, sensations, images, and thought processes are no longer regarded as beyond the reach of scientists. They are studied through overt behavior, just as in physics nonobservables are studied indirectly through what is observed. The situation in psychology is very much like that in physics. That which is conceptualized need not be completely defined in terms of operations, although it must make contact with the world of public experience.

OPERATIONALISM AND VERIFICATIONISM

Operationalism is a movement within the philosophy of science. It is instructive to study its development in conjunction with a parallel movement in general philosophy: logical positivism, or logical empiricism. Central to logical positivism is the principle of verifiability, according to which any statement that is not a tautology must be verifiable or else is meaningless. It was thought that through the extensive employment of this principle it would be possible to show that many of the traditional unsolved problems of philosophy could be dealt with by demonstrating that they are simply meaningless. It was soon found, however, that the principle as originally conceived would get rid not only of troublesome problems but also of much useful discourse. The principle consequently underwent a number of revisions in rapid succession.

Rudolf Carnap's paper "The Methodological Character of Theoretical Concepts" embodies all the significant revisions. Carnap clearly exhibits a desire not to prescribe what should be regarded as meaningful from some metascientific or philosophical point of view but rather to describe what is commonly and usefully regarded as empirically meaningful. As mentioned earlier, operationalism from the beginning sought to explicate an approach already implied in the work of practicing scientists. Whereas verificationists previously tried to embrace all human discourse, they now, like the operationalists, confine their attempts to designing a criterion that will faithfully reflect what is meaningful discourse within empirical science. It has been realized that meanings are contextual and that one is therefore not to inquire whether a given sentence or word has or lacks meaning by itself but rather whether it has or lacks meaning relative to a specified system of theoretical, observational, and mixed statements.

A third important change, also clearly enunciated for the first time in Carnap's paper, is the departure from the original policy of inquiring directly into the meaningfulness of whole sentences. Instead, like the operationalists, Carnap deals with individual terms. He distinguishes between logical and empirical terms and also between observational-empirical and theoretical-empirical terms. Theoretical-empirical terms are not admitted into empirical discourse unless they can be shown to be anchored in observation. They need not be completely defined observationally, but a sentence must be constructible that, in conjunction with other sentences, logically implies that certain observations take place. A theoretical-empirical term is then regarded as having passed the test of empirical meaningfulness. The empirical significance of a sentence is now made dependent on the possession of significance by the terms it contains: Any syntactically well-formed sentence in which every term is significant (that is, is either a logical, an observational-empirical, or a theoretical-empirical term which has passed the test of empirical meaningfulness) is itself significant in the context of the group of sentences forming our system of science.

The only issue that divides operationalism from logical positivism is that operationalism seems to associate meaningfulness with linkability to experimental activities, whereas the principle of verifiability is satisfied if an expression is anchored to mere passive observation. However, this particular requirement of operationalism can safely be discarded, leading to a complete merger of these two contemporary offshoots of empiricism.

CRITICISM

Even in its present form, operationalism has not gone uncriticized. The chief complaint is that in the course of weakening its demands in order to accommodate highly theoretical but useful terms that would otherwise have been excluded from science, it has become so watered down as to lose all significance. Operationalism, according to its critics, says nothing we did not know all along. Even in a discipline less precise than physics—for example, in the social sciences—and in a period when standards of rigor had not reached their present stringency, if anyone had advanced a theory employing concepts which had no bearing whatsoever on observables, his theory would have been rejected. It is admitted that operationalism as originally conceived did have practical impact;

there are concrete results, especially in psychology, whose production was motivated by the naive operationalistic distrust of anything remote from experience—for example, results obtained in the investigation of subaudible speech. Psychologists came to this area of inquiry chiefly through their search for objective, nonmentalistic alternatives to thought processes. But now, with the liberalization of the criterion for empirical significance—so the complaint goes—when all that is stipulated is that no term qualifies for membership in the vocabulary of science unless it is in some way connected to the universe of operations, observables, and experience, the principle of operationalism is merely platitudinous.

In attempting to reply to this, we must not forget that the scope of operationalism is not confined to the weeding out from scientific vocabulary of terms devoid of empirical significance. Once we have adopted the operational point of view, we have formed in our own minds a particular image of the nature of scientific concepts, which colors our expectations and influences in all sorts of ways our practical approach and methodology.

The world of experience and observation was at one time looked upon as containing mere dim reflections of the world that is conceptualized in physics and whose real existence was on a transcendental plane ultimately beyond our reach. Admittedly, that which is without any observable manifestations whatsoever, which, so to speak, casts no shadow onto the plane of experience, would never have been considered as being of any use to science. Nevertheless, it is not unimportant whether we regard our operations as capturing at most the shadows of the furniture of the universe or as dealing with the furniture itself. Objects totally dissimilar in substance and even in size and shape may under particular circumstances cast identical shadows. Therefore, from the similarity of shadows one cannot infer a similarity in the corresponding objects or even that these objects always cast similar shadows. Similarly, so long as we regard as mere reflections the observations to which physical concepts are linked, the finding of resemblances between some of them will not give rise to the expectation that they resemble in all particulars. On adopting the operational point of view, on the other hand, we think we are looking not at reflected shadows but at the very entities and processes that are conceptualized in science, and our attitude changes accordingly.

To give an illustrative example, the properties of gravitational force and the laws governing it had been exhaustively investigated in the seventeenth and eighteenth centuries. Electromagnetic forces were comparative newcomers in science. Were they to be expected to behave like mechanical forces? There are excellent grounds for saying no: the sources from which electromagnetic forces arise, the systems with which they are associated, and the means by which they are generated are totally different from those involving mechanical forces. However, operationalists tend to see in the product of mass and acceleration (that is, in the measure of force) the very substance of force, although others might see in it no more than force's most immediately apparent reflection. Indeed, as soon as it was observed that electromagnetic phenomena are accompanied by the forcelike effect of accelerating masses, it was taken for granted that they are fully governed by all the laws of Newtonian mechanics, even though the latter was developed to deal with an effect of totally different origin.

An important aim of operationalism besides the practical one is philosophical. For philosophical purposes, it is far from sufficient to state generally that every empirically significant term must somehow be linked to observables—one must precisely articulate the nature of this link and construct in full detail a criterion of meaningfulness. Therefore, many concepts in the various sciences were analyzed in detail in order to clarify the exact role instrumental operations and observations play in the definition or explication of them. Believers in the ultimate formalizability of empirical significance hoped that the results would be generalized and expressible in a philosophically satisfactory way. It is, however, by no means clear that such work has been entirely successful. In fact, some philosophers are of the opinion that such efforts are altogether in vain and that although when faced with any individual term we are able quite easily to judge whether it is empirically significant, we shall never succeed in explicating the general criterion distinguishing meaningful from meaningless utterances.

There is thus unquestionably much scope for operationally clarifying basic concepts. The skeptic might try to show that just as there are no formal criteria by which to distinguish a fertile from a sterile theory, so there is no criterion by which to distinguish the empirically significant from the meaningless. One who believes that the contact empirical concepts must make with operations or experience in general can be precisely formalized might try to show that if our demands are modest enough and we do not expect the criterion of empirical significance to provide guidance for future scientific research, there are in principle no obstacles in the way of such formalization. Their next step would be to execute this formalization in a manner that would stand up to all criticism.

See also Bridgman, Percy William; Carnap, Rudolf; Logical Positivism; Scientific Theories; Verifiability Principle.

Bibliography

Bridgman, P. W. *The Logic of Modern Physics.* New York: Macmillan, 1927.

Bridgman, P. W. *The Nature of Physical Theory.* Princeton, NJ: Princeton University Press, 1936.

Bridgman, P. W. *The Nature of Some of Our Physical Concepts.* New York: Philosophical Library, 1952. Many individual concepts carefully analyzed.

Bridgman, P. W. *The Nature of Thermodynamics.* Cambridge, MA: Harvard University Press, 1941. Carefully analyzes many individual concepts.

Bridgman, P. W. *Reflections of a Physicist.* New York: Philosophical Library, 1950. A collection of papers.

Bures, C. E. "Operationism, Construction, and Inference." *Journal of Philosophy* 37 (1940): 393–401.

Carnap, Rudolf. "The Methodological Character of Theoretical Concepts." In *Minnesota Studies in the Philosophy of Science,* Vol. 1, edited by Herbert Feigl and Michael Scriven, 38–76. Minneapolis: University of Minnesota Press, 1956.

Carnap, Rudolf. "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–471, and 4 (1937): 1–40.

Crissman, P. "The Operational Definition of Concepts." *Psychological Review* 46 (1939).

Dingle, Herbert. "A Theory of Measurement." *British Journal for the Philosophy of Science* 1 (1950). An extreme operationalist viewpoint.

Feigl, Herbert. "Operationism and Scientific Method." *Psychological Review* 52 (1945). A lucid and fair assessment of operationalism. Written as a contribution to "Symposium on Operationism," presented in the same issue.

Frank, Philipp. *Modern Science and Its Philosophy.* Cambridge, MA: Harvard University Press, 1949. On p. 44 the work of Bridgman is likened to that of Carnap.

Frank, Philipp, ed. *The Validation of Scientific Theories.* Boston: Beacon Press, 1956. Various writers assess the significance of operationalism in the 1950s.

Hearnshaw, L. J. "Psychology and Operationalism." *Australasian Journal of Psychology and Philosophy* 18 (1941).

Hempel, C. G. "A Logical Appraisal of Operationalism." *Scientific Monthly* 79 (1954): 215–220. Reprinted with modifications in his *Aspects of Scientific Explanation,* 123–133. New York, 1965.

Lindsay, L. B. "A Critique of Operationalism in Science." *Philosophy of Science* 4 (1937). Important and fair criticism of operationalism.

Margenau, Hans. *The Nature of Physical Theory.* New York, 1952. On p. 232, expresses impatience with operationalism.

Pap, Arthur. "Are Physical Magnitudes Operationally Definable?" In *Measurements, Definitions, and Theories,* edited by C. West Churchman. New York, 1959. Argues for the abolition of the demand for active operations.

Peters, Richard. "Observationalism in Psychology." *Mind* 68 (1959).

Reichenbach, Hans. *Experience and Prediction.* Chicago: University of Chicago Press, 1938.

Russell, L. J. Review of Bridgman's *The Logic of Modern Physics. Mind* 47 (1938).

Schlesinger, Georg. *Method in the Physical Sciences.* London: Routledge and Paul, 1963. Chapter 4 is devoted to a discussion of the practical scope of operationalism.

Skinner, B. F. "Behaviorism at Fifty." *Science* (1963).

Skinner, B. F. "The Operational Analysis of Psychological Terms." *Psychological Review* 52 (1945): 270–277.

Spence, K. W. "The Postulates and Methods of 'Behaviorism.'" *Psychological Review* 55 (1948): 67–78.

Stevens, S. S. "The Operational Basis of Psychology." *American Journal of Psychology* 46 (1935).

Stevens, S. S. "The Operational Definition of Psychological Concepts." *Psychological Review* 42 (1935). A major spokesman for operationalism in psychology.

G. Schlesinger (1967)

OPTIMISM AND PESSIMISM

See *Pessimism and Optimisim*

ORDER

See *Chaos Theory*; *Measurement and Measurement Theory*; *Teleological Argument for the Existence of God*

ORESME, NICOLE
(c. 1320–1382)

Nicole (Nicholas) Oresme was a Master of Arts and Theology at the University of Paris, royal counsellor, translator into French of Aristotle's works, and bishop of Lisieux. Of humble origin, he was admitted in the College of Navarre in 1348, where he became Grand Master in 1356, after having obtained the license of Master of Theology. He was born in Normandy probably no later than 1320, in a village near Caen (Allemagne, today Fleury-sur-Orne). His ecclesiastical career depended on his university teaching as well as on his connections with the royal court. The first benefice was granted by Pope Clement VI in 1342, in reply to a supplication list of the University of Paris in order to obtain support for master and students (Oresme is recorded as master); the election to the bishop's chair of Lisieux in 1377 was Charles V's (1364–1380) reward for Oresme's translations of Aristotle's works, made by royal request. His main ecclesiastical functions were in Normandy, a region with high strategic

importance during the wars between France and England. He was appointed canon of Rouen Cathedral in 1362, and two years later he was chosen as dean. He reduced, but did not cut short, his connections with the university and with the royal court in Paris. In 1370 he disputed at the university a quodlibetal question; in 1375 he was charged, together with Simon Fréron and Richard Barbe, to find out if Marsilius of Padua's *Defensor Pacis* had been translated into French. Oresme translated and commented upon Aristotle's *Ethics* (*Le livre de ethiques d'Aristote*), *Politics* (*Les politiques*), *Economics* (*Le livre de yconomique d'Aristote*), and *De Caelo* (*Le livre du ciel et du monde*). He wrote also in French an elementary treatise on astronomy (*Livre de l'éspere*), and a treatise against the astrologers (*Livre de divinacions*). He died on July 11, 1382, in Lisieux.

His commentaries on Aristotle's physical writings (*Physics*, *On the heaven*, *On coming to be and passing away*, *On the soul,* and *Methereologics*), as well as his treatises (*Ad pauca respicientes*, *De proportionibus proportionum*, *De commensurabilitate motuum caeli*, *De configurationibus qualitatum*) bear witness to his prevailing scientific interests, and above all to his conviction of the importance of using mathematics in dealing with physical problems (qualitative changes, motion, duration). In his commentaries, Oresme discusses the main philosophical issues debated at the University of Paris after the dissemination of William of Ockham's works and the condemnations of John of Mirecourt (1347) and Nicolas of Autrécourt (1348).

THE SUBJECT OF HUMAN KNOWLEDGE AND THE CERTITUDE OF PHYSICAL SCIENCE

Oresme offered rather original solutions to two very important problems traditionally discussed in the opening questions of medieval commentaries on the physical writings of Aristotle: the subject of human knowledge, and the degree of certitude of physical science. Concerning the first, Oresme rejects the reductionist view, usually attributed to William Ockham, according to which human knowledge concerns exclusively the conclusion of a syllogism, as well as the claim that it deals with singular objects. He believes that human knowledge concerns properly what can be expressed through a proposition (*complexe significabile*) rather than through a single term.

On the certitude of physical science, Oresme shares the common position, strongly attacked by Nicolas d'Autrécourt, according to which it does not need the highest degree of certitude typical of mathematics and metaphysics. The convenience of having recourse to mathematics in physical inquiries, however, permits one in some way to extend to physics this highest degree of certitude.

The possibility of applying mathematics to physics is warranted either by widening the field of physical inquiries to a hypothetical, non contradictory state of things, or by assuming the geometrical model of perspective in explaining physical actions like heating. The extension of imaginary cases to physical inquiries actually increases the potential of physics, whose limits coincide with the law of noncontradiction. In his *Quaestiones de spera* (q. 2), Oresme explicitly upholds the use of mathematical fictions (*imaginationes*), like points and lines, in physics, stating that in astronomy (and in the so called *scientiae mediae*) truth can not be reached without the aid of mathematics and geometry (he quotes for this solution the authority of Aristotle's *De coelo*).

The plurality of worlds and the daily rotation of the earth on its axis while the heavens remain stationary—two of the topics to which Oresme owes his celebrity among historians of science since Pierre Duhem—are such hypothetical cases. Oresme amply discussed the possibility of such hypotheses, concluding always in favor of the traditional view. The relativity of motion is a central issue in the astronomical hypothesis of the earth's daily rotation; Oresme's position concerning the nature of motion is an original attempt to maintain an absolute notion.

MATHEMATICS AND PHYSICS

One of Oresme's major contributions to natural philosophy is his solution concerning the "intension and remission of qualities"—that is the variation of intensity of qualities, motion, velocity, and every kind of successive thing. *De configurationibus qualitatum* opens by confirming the utility of making recourse to mathematics in physical inquiries: Intensities of qualities can be easily measured by representing them through geometrical figures, whose one line represents the subject where the quality is distributed (*extensio*), on which there are perpendicularly erected lines representing the intensities of the quality (*intensio*). The line connecting the higher points of the intensities (*linea summitatis*) can immediately inform us about the type of change (uniform, uniformly difform, difform).

Oresme avails himself of this method of graphing the varying of intensities of qualities and motions in order to explain the diversity of actions of physical agents, and also of human passions, occult virtues, aesthetic problems, and magical operations. In his effort to reduce uniformly difform types of variation to uniform ones, Oresme proposes

a geometrical demonstration of the so called mean-speed theorem (the distances traversed by two moving objects, the former moving uniformly/difformly and the latter uniformly with the mean speed of the former, is the same). Galileo used an analogous geometric demonstration for freely falling bodies in his *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*.

Oresme adhered to Thomas Bradwardine's solution, according to which velocity depends on a proportional change of the force as well as of the resistance. In order to double velocity, it is not enough to double force or to halve resistance, but the square of the proportion between force and resistance must also be obtained.

In *De proportionibus proportionum* III, prop.10, Oresme resorts to mathematics to argue for the high degree of probability of the incommensurability of any two unknown ratios: "because if many unknown ratios are proposed it is most probable that any one would be incommensurable to any other" (E. Grant's translation, p. 247). He proposes a similar argument in *De commensurabilitate* to support the incommensurability of heavenly circular motions in order to invalidate astrological predictions based on planetary conjunctions, which would be unpredictable.

MODI RERUM

Oresme's *Physics* commentary contains an original philosophical doctrine concerning the nature of motion, place, and time, and more generally the ontology of natural things. Evidently dissatisfied by the two opposing solutions—the reductionist, inspired by Ockham, according to which motion is nothing different than the moving object; and the realist, according to which motion is a quality inherent to the moving object—Oresme proposed to consider motion, as well as place, time, and other continuous natural things, as complex objects or events rather than as simple qualities and properties. To do that he availed himself also of semantical tools like the meaning of the proposition (*complexe significabile*). Oresme was convinced that his solution was able to avoid some ontological problems in natural philosophy: He explicitly quotes intension and remission of qualitative forms, with qualities considered as *modi* of the substance and not accidental properties inhering to the substance.

See also Aristotle; Bradwardine, Thomas; Duhem, Pierre Maurice Marie; Galileo Galilei; John of Mirecourt; Marsilius of Padua; Mathematics, Foundations of; Medieval Philosophy; Nicolas of Autrecourt; William of Ockham.

Bibliography

On Oresme's biography see: F. Neveux, "Nicole Oresme et le clergé normand du XIVe siècle." In *Autour de Nicole Oresme*, pp. 9–36, edited by J. Quillet. Paris: Vrin, 1990; W. J. Courtenay, "The Early Career of Nicole Oresme," *ISIS* 91 (2000), pp. 542–548.

French translations and commentaries: *Le livre de ethiques d'Aristote*, A. D. Menut, ed. (New York: G. E. Stechert, 1940); *Le livre de yconomique d'Aristote*, A. D. Menut ed. with English transl., in *Transactions of the American Philosophical Society*, N. S. Vol. 47, Part 2 (Philadelphia: American Philosophical Society, 1957); *Le livre de Politiques d'Aristote*, A. D. Menut, ed., in *Transactions of the American Philosophical Society*, N. S. Vol. 60, Part 6 (Philadelphia: American Philosophical Society, 1970); *Le livre du ciel et du monde*, A. D. Menut and A. J. Denomy, eds. with English transl. (Madison: Univ. of Wisconsin Press, 1968); *Le traité des monnaies* (French transl. of *De mutationibus monetarum*), ed. L. Wolowski, *Petit traité de la première invention des monnaies de Nicolas Oresme*, Paris, Guillaumin 1864 (Slatkine Reprints, 1976).

French treatises: *Traité de l'espere*, ed. J. V. Myers (Syracuse Univ., 1940), and L. M. McCarthy (Toronto Univ., 1943); *Le livre de divinacions*, G. W. Coopland ed. (Cambridge, MA: Harvard Univ. Press, 1952).

Latin commentaries and other writings in form of university questions: *Quaestiones super Geometriam Euclidis*, H. L. L. Busard, ed., (Leiden: Brill, 1961); *Quaestiones de sensu*, J. Agrimi, ed. (Firenze, 1983); *Expositio et Quaestiones in Aristotelis "De anima,"* B. Patar ed. (Louvain-Paris: Éditions de l'Institut Supérieur de Philosophie-Éditions Peeters, 1995); *Quaestiones super De generatione et corruptione*, S. Caroti ed. (München: Bayerische Akademie der Wissenschaften, 1996); *Quaestiones super Physicam*, S. Kirschner ed. (Stuttgart: Franz Steiner, 1997), the complete edition is in preparation; *Quaestio contra divinatores horoscopios*, S. Caroti, ed. in *Archives d'histoire doctrinale et littéraire du Moyen Âge*, 43, 1976, pp. 201–310; *Quodlibeta*, B. Hansen (Toronto: Pontifical Institute of Mediaeval Studies, 1985, partial ed.); *Quaestiones super de coelo*, C. Kren ed. (Univ. of Wisconsin, 1965); *Quaestiones super de spera*, G. Droppers ed. (Univ. of Wisconsin, 1966); *Quaestiones super quatuor libros Meteororum*, S. C. McCluskey ed. (Univ. of Wisconsin, 1974, partial ed.); *De visione stellarum*, D. E. Burton ed. (Indiana Univ., 2000); *Questio utrum aliqua res videatur*, J. B. Watson ed. (Harvard Univ., 1973).

Not yet identified is his commentary on the *Sentences*; the *De perfectione specierum* to which Oresme alludes in his *De configurationibus* (I, 20) is probably part of this commentary. Other theological writings: *Tractatus de communicatione idiomatum*, E. Borchert, ed. In *Beiträge zur Geschichte der Philosophie und Theologie des Mittelalters*, 35, München: Aschendorff, 1940, H. 4–5; *Determinatio facta in resumpta in domo Navarrae*, ms.

Latin treatises: *Algorismus proportionum*, M. Curtze ed. (Berlin: Calvary, 1868) and E. Grant ed. (Univ. of Wisconsin, 1957, partial ed.); *De proportionibus proportionum*, *Ad pauca respicientes*, E. Grant ed. with English transl. (Madison: Univ. of Wisconsin Press, 1966); *De configurationibus qualitatum et motuum*, M. Clagett ed. with English transl. (Madison: Univ.

of Wisconsin Press, 1968); *De commensurabilitate vel incommensurabilitate motuum celi*, E. Grant ed. with English transl. (Madison: Univ. of Wisconsin Press, 1971); *Tractatus contra iudiciarios astronomos*, H. Pruckner ed. (Leipzig-Berlin: Teubner, 1933) and G.W. Coopland ed. with English transl. (Cambridge, MA: Harvard Univ. Press, 1952); *Tractatus de mutationibus monetarum*, C. Johnson, ed. with English transl. (London: T. Nelson, 1956).

In addition to the pioneer studies of P. Duhem, *Le système du monde*, 10 vols. (Paris: Herman, 1913–1959, Vols. IV, VII, VIII, and IX) and A. Maier, *Studien zur Naturphilosophie der Spätscholastik*, 5 vols. (Rome: Edizioni di Storia e Letteraturea, 1949–1958), where Oresme's role in the history of science was highly appreciated, and to the volume of O. Pedersen's *Nicole Oresme og hans natufilosofiske System* (Copenhagen: Munksgaard, 1956, with a French summary), much new material can be found in the introductions to the editions of his works (above all in Clagett's introduction to *De configurationibus*). Two conferences have been dedicated to Oresme: *Nicolas Oresme. Tradition et innovation chez un intellectuel du XIVe siècle*, P. Souffrin, A. Ph. Segonds, eds. (Padova-Paris: Programma e Editori-Les Belles Lettres, 1988, with a very important contribution of H. Hugonnard-Roche, *Modalités et argumentation chez Nicole Oresme*) and *Autour de Nicole Oresme*, J. Quillet ed. (Paris: Vrin, 1990). The n. 3 (2000) of the review *Oriens-Occidens. Sciences, mathématiques et philosophie de l'Antiquité à l'Âge classique* is partly dedicated to Nicole Oresme's *Physics*. New material is also in the proceedings of the Parma conference: *Quia inter doctores est magna dissensio. Les débats de philosophie naturelle à Paris au XIV*[e] *siècle*, S. Caroti, J. Celeyrette eds. (Florence: Olschki, 2004).

Stefano Caroti (2005)

ORGANISMIC BIOLOGY

The term *organismalism* was coined by the zoologist W. E. Ritter in 1919 to describe the theory that, in his words, "the organism in its totality is as essential to an explanation of its elements as its elements are to an explanation of the organism." Subsequent writers have largely replaced *organismal* with the more euphonious *organismic* as a title for this theory, for the many variations on its main theme, and for some subordinate but supporting doctrines concerning the teleological and historical character of organisms.

Ritter regards Aristotle as the founder and most distinguished exponent of the organismic theory. But Aristotle is also claimed as the father of vitalism, a view that organismic biologists in general reject. In fact, there is considerable affinity between the two schools. They both agree that the methods of the physical sciences are applicable to the study of organisms but insist that these methods cannot tell the whole story; they agree that the "form" of the single whole organism is in some sense a factor in embryological development, animal behavior, reproduction, and physiology; and they both insist on the propriety of a teleological point of view. On all of these points, Aristotle not only agrees but presents, in his own terminology, careful and persuasive arguments in their favor. But organismic biology and vitalism differ in one fundamental respect: The latter holds (and the former denies) that the characteristic features of organic activity—all of which fall under the heading of "regulation"—are caused by the presence in the organism of a nonphysical but substantial entity. There are different interpretations of Aristotle (which we cannot examine here) on the question of whether he believes there are such vital entities. In this writer's view, Aristotle is clearly a vitalist.

The affinity between vitalism and organismic biology is more than an accident. In the history of biology it is difficult to disentangle vitalistic and organismic strands, since both schools are concerned with the same sorts of problems and speak the same sort of language. The distinction between them was drawn clearly only in the twentieth century. Organismic biology may be described as an attempt to achieve the aims of the murky organismic-vitalistic tradition, without appeal to vital entities.

The writings of contemporary organismic biologists present a number of difficulties for a philosophical commentator. The position of organismic biology is usually stated in a vocabulary that plays little or no theoretical role in the working language of biology. For example, "whole," "unity," "integrity," "part," "form," "principle," "understanding," and "significance" all occur frequently in their works. Now any biologist will use these terms occasionally in the course of his professional writing, just because they are perfectly good words in the English language. But they are not technical expressions; they are not, in ordinary usage, laden with biological theory; and they are trouble-free only when employed in contexts that make clear their function as items in the common language. The organismic biologist, however, makes them bear a heavy burden in the description of the nature of living organisms. And many, but by no means all, organismic biologists also assign a great deal of weight to some rather mysterious formulas. Here are a few: "The whole acts as a causal unit … on its own parts" (W. E. Agar); "The living body and its physiological environment form an organic whole, the parts of which cannot be understood in separation from one another" (J. S. Haldane); "No part of any organism can be rightly interpreted except as part of an individual organism" (W. E. Ritter). And here are a few more that are characteristic but not direct quotations: "The organic whole is greater than the

sum of its parts"; "Knowledge of the goal of an animal's behavior is necessary for understanding its significance"; "Biological theory should be autonomous, with concepts and laws of its own." These formulas may be termed "mysterious" because, according to their most natural interpretations (as will be argued), they are all the barest of truisms.

Two additional points should be mentioned. Organismic biologists have employed some of the more obscure technical conceptions of speculative philosophy, such as "formal cause," "emergence," "hormic," "telic," and so on. And since their writings are a minority report on biological phenomena, organismic biologists are often polemical, engaging in denunciations of other biologists—"mechanists," "elementalists," and "reductionists"—whose positions they leave just as obscure as their own. For all of these reasons, an account of the organismic position that aims at answering the questions likely to be raised by philosophers of science involves elements of reconstruction and interpretation. Thus, a fuller description of the position and an interpretation designed to do justice both to the letter and spirit of the organismic tradition follows.

THE POSITION OF ORGANISMIC BIOLOGY

All organismic biologists hold that there is a gulf between organic and inorganic phenomena in one or more of the following respects.

ORGANIC UNITY. Organic systems are so organized that the activities of the whole cannot be understood as the sum of the activities of the parts. All members of the school agree on this point. As the term *organismic* implies, the most important example of such wholes is the single organism, but there are others, such as cells, organs, colonies, and some populations.

J. H. Woodger, whose *Biological Principles* is the most careful and extensive exposition of organismic biology, explains the conception of organic unity in the following way. Consider a system *W* that is *totally* composed of physicochemical parts—elementary particles, for example. The activities of these parts are described by the laws of physics. These particles may be the sole constituents of other systems (for example, molecules) which also totally compose *W* and which exhibit, in addition to activities described by the laws of physics, other activities described by the laws of chemistry. Molecules may similarly be the sole constituents of other systems, which are in turn the constituents …, up to the whole system *W*. In Woodger's terminology, *W* exhibits a series of "levels of organization." The parts of *W* belong to a particular level, its physical parts to the physical level, its chemical parts to the chemical level, and so on. System *W* constitutes a perfect "hierarchy" of parts from levels 0 (zero) to *n* (a finite number), if 0-level parts are the sole constituents of all 1-level parts, and if every part at each level *i* (any given level) except the 0-level is totally composed of parts at level *i*–1.

Woodger points out that organisms are not perfect hierarchies, since some parts of the organism at an *i*–level may have parts at the *i*–2 level, while the *i*–2 parts are *not* organized into *i*–1 parts (for instance, blood has cellular *and* chemical but noncellular parts). Nevertheless, he contends, organisms approximate to a hierarchical organization. If we ignore deviations from the perfect hierarchy, we may let *W* represent a whole organism, and we may say that its 0-level parts are physical parts. Now this analysis permits us to say that the organism is composed totally of physical parts. Perhaps some philosophical materialists would be content with this thesis; at any rate, if it is true, it rules out vitalism. But it is false that the organism is composed *only* of physical parts, for there are parts at higher levels of organization. It is Woodger's contention, and a general thesis of organismic biology, that the laws which determine the behavior of the parts at a given level of organization are silent about some aspects of the behavior of the parts at the higher levels. To use an extreme example, the laws of quantum physics have nothing to say on the question of why honeybees kill their drones. According to Woodger, it is necessary to study the relations between the relata at *each* level of organization. In order to understand the behavior of cells during morphogenesis, for example, we must develop a theory of cell relations and not be content, for example, with only a theory of the relations between molecules.

DETERMINING FEATURES OF THE WHOLE. The parts of organic wholes not only exhibit patterns of behavior in virtue of their relations to other parts at the same level of organization, but in addition, *some* of the features of the parts at a given level are determined by the pattern of organization at *higher* (and, of course, at lower) levels of organization. This is the general form of the special thesis that the properties of the whole determine the properties of the part; and it seems to have the methodological consequence that a theory of the elements at a given level could not be complete without a theory of the elements at the higher levels. Woodger puts the point this way: the parts of organisms must be studied in situ, for we cannot learn how they would behave in situ by studying them in isolation.

TELEOLOGICAL BEHAVIOR OF ORGANISMS. One kind of activity, which is a consequence of organization at a level higher than that of the organism's physical parts, is directive or teleological behavior. Directiveness is an aspect of organisms that is shown in their physiology, in the behavior of individual animals, and in the social systems of some animals; and an account of directiveness is not only legitimate but necessary. E. S. Russell argues that since directiveness (processes aimed at the production and maintenance of organic unities) is a fact, then a physiological process, or piece of animal behavior, cannot be understood until we understand its function or its goal.

INTERPRETATION OF ORGANISMIC BIOLOGY

It was remarked above that if we give the slogans of organismic biology their most direct interpretations, they are nothing more than truisms. Consider, for example, the statement that the whole (if it is an organic unity) is more than the sum of its parts. This looks like a simple warning against the fallacy of composition: we are being warned, for example, that from the premise "No part of a bird can fly" we cannot infer "No whole bird can fly." No weighty volume is required to convince us that a whole may have numberless properties that its parts lack. Of course, there are other possible interpretations of the slogan. It might be taken to mean, especially in the form "The behavior of the whole is more than the sum of the behavior of its parts," that no description of the behavior of the parts could be a description of the behavior of the whole. So far from being a truism, this is obviously false. Finally, it might be taken to mean something like the following. Employing an analysis of Ernest Nagel, we might say that the behavior B of a system S is more than the sum of the behavior $b_1, b_2, \cdots, b_n$ of its parts $s_1, s_2, \cdots, s_n$, with respect to an antecedently specified theory T, if (1) B is an instance of a law L; (2) L is not part of T: (3) the laws in T describe $s_1, s_2, \cdots, s_n$ in such a way that they explain $b_1, b_2, \cdots, b_n$; and (4) L is not deducible from a description of $s_1, s_2, \cdots, s_n$ together with laws in T. An important point to notice here is that B can be identical with events $b_1, b_2, \cdots, b_n$, and yet the law of which B is an instance is not derivable from the laws of which $b_1, b_2, \cdots, b_n$ are instances.

This account makes the "more than" relation relative to a body of theory. Relative to existing physical and chemical theories, it is true (but perhaps not a truism) that much organic activity is more than the sum of the physical and chemical activities of its parts. The thesis that there are cases of higher-level behavior that will remain greater than the sum of the behavior of its physical parts, for all possible physical theories, is the doctrine of emergence, which many organismic biologists believe to be true. But it is essential to note two points—first, that the thesis is dubious and unproved, and second, that one can be an organismic biologist without believing it (L. von Bertalanffy is an example).

Let us now look at two more formulas of the organismic biologists. Woodger holds that an organic part, such as a cell, has properties in the organism that it does not have in isolation from the organism. This, too, is a truism: An excised eye lacks the property of contributing to the sight of its former owner. Now if we add, as Woodger does, that the properties of the part in the whole could not be uncovered by studying the part outside the whole, the thesis reduces to the thesis of emergence. And certainly, one of the commonest scientific procedures consists in predicting the behavior of a part in a system that has not yet been studied, although this prediction is assuredly made on the basis of knowledge gained by studying the part—not in "isolation," but as a part of other systems. For instance, the behavior of an electron in a cathode ray tube allows us to predict the electron's behavior in a cyclotron.

Finally, we may consider E. S. Russell's remark that understanding the significance of an animal's behavior requires understanding its goal. This, at least on Russell's interpretation, is a truism, for he connects the notion of a goal with the notion of adaptive value for the animal and identifies "significance" with adaptive value.

Omitting specific discussion of the other formulas cited, the general point is clear: Organismic biology seems to collapse either into doctrines that are not controversial or into unclarified, unproved, and dubious assertions about emergence, unpredictability, and irreducibility. Nevertheless, organismic biology is an important and valuable movement, for the following reasons.

First, organismic biology is perfectly correct in pointing out that there are levels of organization above the chemical level which exhibit laws of behavior that are not exhibited at lower levels (for example, molecules do not sting other molecules to death). Higher-level behavior can be treated without reference to behavior at lower levels, which means that the biologist can (and indeed does) construct concepts that are tailored to the description of higher-level behavior. The principles at the higher levels must be formulated before the question of their reducibility to lower level principles can even be considered. A biochemical geneticist is not only a biochemist; he is also a geneticist, because he is involved in elucidating

the processes involved in the sort of gross biological phenomena studied by Gregor Mendel.

Second, the insistence of organismic biologists on the importance of functional analysis is well founded. Focusing on the biological ends of physiological and behavioral processes provides the only means for developing the conceptual schemes that are needed in morphology, ethology, evolution theory, and other branches of biology. This point is developed in detail in Morton Beckner's *Biological Way of Thought.*

Third, although organismic biology is a set of truisms, it is none the worse for being so. The trouble with truisms is their great number: there are so many that we easily overlook, sometimes systematically, some of the most important ones. Even though in fact many biologists agree with the organismic position, they will say that they disagree. This leads to the position (generally deleterious in the sciences) of the scientist's doing one thing and describing it as if he were doing something else.

To sum up, organismic biology is to be interpreted as a series of methodological proposals, based on certain very general features of the organism—namely, the existence in the organism of levels of organization with the biological ends of maintenance and reproduction. These features are sufficient to justify "a free, autonomous biology, with concepts and laws of its own," whether or not the higher levels are ultimately reducible to the lower ones.

See also Aristotle; Bertalanffy, Ludwig von; Philosophy of Biology; Teleology; Vitalism; Woodger, Joseph Henry.

Bibliography

Agar, W. E. *A Contribution to the Theory of Living Organisms.* 2nd ed. Melbourne, 1951.

Beckner, Morton. *The Biological Way of Thought.* New York: Columbia University Press, 1959.

Bertalanffy, Ludwig von. *Modern Theories of Development.* London: Oxford University Press, 1933.

Haldane, J. S. *Mechanism, Life, and Personality,* 2nd ed. New York: Dutton, 1923.

Lillie, R. S. *General Biology and Philosophy of Organism.* Chicago: University of Chicago Press, 1945.

Nagel, Ernest. "Mechanistic Explanation and Organismic Biology." In *Philosophy and Phenomenological Research* 2 (1951): 327–338.

Ritter, W. E. *The Unity of the Organism.* 2 vols. Boston: R.G. Badger, 1919.

Russell, E. S. *The Behaviour of Animals.* 2nd ed. London: E. Arnold, 1938.

Woodger, J. H. *Biological Principles.* London, 1948.

Morton O. Beckner (1967)

ORIGEN

(c. 185–253)

Origen, the Christian theologian and exegete of the Bible, was the foremost member of the catechetical school at Alexandria. Born of Christian parents in Alexandria, he was made head of a Christian school there in 204. He taught until 231, when conflict with the bishop forced him to leave for Caesarea in Palestine, where he taught until his death. He apparently heard lectures by Ammonius Saccas, founder of Neoplatonism, although he regarded philosophy as essentially preparatory to theology in the same way that other studies were prerequisite to philosophy itself. However, the influence of philosophy (primarily Platonic but also Stoic) on his thought was highly significant; it can be observed much more clearly in his presuppositions and arguments than in explicit quotations, which are relatively unusual except in the apologetic treatise *Contra Celsum*. The most important of his voluminous writings are *De Principiis*, a treatise on first principles and the earliest extant Christian systematic theology; the treatise *On Prayer*; and *Contra Celsum*.

DE PRINCIPIIS

A relatively early work, *De Principiis* begins with the statement that apostolic doctrine, as found in the New Testament, is incomplete because the apostles intentionally left some matters untouched for the sake of their spiritual successors. Origen devotes the first book to a consideration of the spiritual hierarchy consisting of the Father, who acts on all beings; the Logos (Word or Reason), who acts upon rational beings; the Spirit, who acts upon those rational beings who are sanctified, and the angels. The second book deals with the material world. Man, created because the angels fell, is a preexistent fallen spirit in a material body. After Adam's transgression came redemption by the incarnate Logos; later there will be resurrection, the last judgment, and the life of all men restored to spiritual bodies (a succession of other worlds may follow as it has gone before). The third book discusses freedom, characteristic of creatures but not of the Creator. When a soul is in a body, it can struggle for victory, helped by angels and hindered by demons. Since it possesses free will, it is capable of choosing the good. After a brief summary, Origen turns in the fourth book to an explanation of how the Scriptures can be shown to have various levels

of meaning. Like man himself, they have flesh (literal meaning), soul (moral meaning), and spirit (allegorical-spiritual meaning). The exegetical difficulties in Scripture were placed there by their ultimate author, God, in the way that similar obstacles to faith were placed in the cosmos so that man could use his mind.

Origen's work, written in Greek, is extant only in fragments (Book IV is almost entire). The Latin version by Tyrannius Rufinus was severely criticized by St. Jerome on the ground that it lacks unorthodox passages that were in the original, but it has come to be regarded more favorably by modern scholars. The title *De Principiis* has parallels in second-century philosophy, as do many of the subjects Origen discusses; his approach, however, seems to be essentially Christian.

ON PRAYER

In *On Prayer*, written later in his life, Origen discusses prayer in general (Chs. 3–17) and the Lord's Prayer in particular (Chs. 18–30). The principal problem is that presented by prayer to an omniscient God who has foreordained everything. Once again, Origen insists upon God's gift of free will; the primary purpose of prayer is not petition as such but sharing in the life of God. Origen classifies prayer as petition, adoration (only of the Father), supplication, and thanksgiving. In each case he emphasizes—as do contemporary middle Platonists—the spiritual attitude of the one who prays.

CONTRA CELSUM

The late apologetic treatise against Celsus, written in 248, reveals the extent to which Origen was able to argue on grounds shared by his philosophical opponents; there is actually a wide measure of agreement between him and Celsus. Both are opposed to anthropomorphism, to idolatry, and to any crudely literal theology. Origen, however, consistently defends Christianity as he sees it and does not hesitate to attack philosophies and philosophers.

ORIGEN AND PHILOSOPHY

The precise extent of Origen's debt to philosophy was discussed in antiquity; the Neoplatonist Porphyry claimed (according to Eusebius, *Historia Ecclesiastica* VI, 19, 8) that Origen drew upon Plato, Numenius, Cronius, Apollophanes, Longinus, Moderatus, Nicomachus, Chaeremon the Stoic, and Cornutus. Since Origen does refer to many of these writers, whose names occur in Porphyry's description of the Neoplatonic curriculum, Porphyry may be attempting to demonstrate both the extent and the correctness of Origen's Neoplatonism. The systems and works of various philosophers—except for the "atheists"—were studied thoroughly in Origen's school. Origen himself often made use of philosophical dictionaries for the definitions of various terms, but he also studied the writings of the philosophers themselves, not only those of Plato and the Platonists but also those of the Stoics and, occasionally, the Peripatetics.

It is sometimes claimed that there were two Origens, one a pupil of Ammonius Saccas and the other the Christian theologian. It is more likely that both aspects were combined within one person, the first Christian to be a genuinely philosophical theologian.

See also Celsus; Eusebius; Neoplatonism; Numenius of Apamea; Patristic Philosophy; Peripatetics; Platonism and the Platonic Tradition; Porphyry; Stoicism.

Bibliography

TEXTS AND TRANSLATIONS

De la Rue, C. *Opera Omnia*. 4 vols. Paris, 1733–1759, reprinted in J. P. Migne, *Patrologia Graeca* XI–XVII. Paris, 1857–1866.

Critical editions of individual works in *Die griechischen christlichen Schriftsteller der ersten drei Jahrhunderte*. Leipzig: Hinrichs: Hinrichs, 1897–1941; Berlin and Leipzig, 1953; Berlin: Akademie-Verlag: Akademie-Verlag, 1954–.

English translations in *Ante-Nicene Christian Library*. Edinburgh: T. and T. Clark, 1864–.

STUDIES

Benjamins, Hendrik S. *Eingeordnete Freiheit: Freiheit und Vorsehung bei Origene*. Leiden: E.J. Brill, 1994.

Berner, Ulrich. *Origenes*. Darmstatt: Wissenschaftliche Buchgesellschaft, 1981.

Bienert, W. and U. Kühneweg, ed. *Origeniana septima: Origenes in den Auseinandersetzungen des 4. Jahrhunderts*. Leuven: Leuven University Press, 1999.

Brox, Norbert. "Spiritualität und Orthodoxie: Zum Konflikt des Origenes mit der Geschichte des Dogmas," in Ernst Dassmann and K. Suso Frank, eds, *Pietas: Festschrift für Bernhard Kötting = Jahrbuch für Antike und Christentum*. Ergänzungsband 8 (1980): 140–54.

Clark, Elizabeth A. *The Origenist Controversy: The Cultural Construction of an Early Christian Debate*. Princeton, NJ: Princeton University Press, 1992.

Crouzel, Henri. *Origène*. Paris: Lethielleux, 1985.

Crouzel, Henri. *Origène et la philosophie*. Paris, 1962.

Crouzel, Henri. *Origène et Plotin: Comparaisons doctrinales*. Paris: Aubier, 1991.

Dively, Elizabeth. *The Soul and Spirit of Scripture within Origen's Exegesis*. Leiden, 2005. Boston: Brill, 2005.

Edards, Mark. *Origen Against Plato*. Aldershot: Ashgate, 2002.

Hällström, Gunnar af. *Charismatic Succession: A Study on Origen's Concept of Prophecy*. Helsinki, 1985.

Hällström, Gunnar af. *Fides Simpliciorum according to Origen of Alexandria*. Helsinki: Societas Scientiarum Fennica, 1984.

Harl, Marguerite. *Origène et la fonction révélatrice du Verbe incarné*. Paris, 1958.
Jonas, Hans. *Gnosis und spätantiker Geist, vol. 2, Von der Mythologie zur mystischen Philosophie*. Edited by Kurt Rudolph, reprint, Göttingen: Vandenhoeck & Ruprecht, 1993.
Kannengiesser, Charles, ed. *Origen: His World and His Legacy*. Notre Dame, IN: University of Notre Dame Press, 1988.
Lange, Nicholas R. M. de. *Origen and the Jews: Studies in Jewish–Christian Relations in Third-Century Palestine*. Cambridge, 1976.
Lubac, Henri de. *Histoire et Esprit: L'intelligence d'Écriture d'après Origène*. Paris: Aubier, 1950.
Neuschäfer, Bernhardt. *Origenes als Philologe. Schweizerische Beiträge zur Altertumswissemchaft* 18. 2 vols. Basel, 1987.
Peri, Vittorio. *Omelie origeniane sui Salmi*. Vatican City: Biblioteca apostolica vaticana, 1980.
Stroumsa, Guy G. *Hidden Wisdom: Esoteric Traditions and the Roots of Christian Mysticism*. Leiden: E.J. Brill, 1996.
Strutwolf, Holger. *Gnosis als System: Zur Rezeption der valentinianischen Gnosis bei Origenes = Forschungen zur Kirchen- and Dogmengeschichte* 56. Göttingen: Vandenhoeck & Ruprecht, 1993.
Torjesen, Karen Jo. *Hermeneutical Procedure and Theological Method in Origen's Exegesis*. Berlin: De Gruyter, 1986.
Trigg, Joseph W. "The Angel of Great Counsel: Christ and the Angelic Hierarchy in Origen's Theology." *Journal of Theological Studies* n.s. 42 (1991): 35–51.
Trigg, Joseph W. *Origen: The Bible and Philosophy in the Third-Century Church*. Atlanta: J. Knox, 1983.

Robert M. Grant (1967)
Bibliography updated by Scott Carson (2005)

OROBIO DE CASTRO, ISAAC

(c. 1617–1687)

Isaac Orobio de Castro was born Baltazar Orobio de Castro in Braganza, Portugal. He grew up among crypto-Jews who were trying to preserve some of their heritage in the face of the Spanish Inquisition. He became an important Spanish doctor and a professor of metaphysics. He was arrested by the Inquisition for secretly practicing Judaism. After being tortured and tried, he was released. He then fled Spain for France, where he became professor of pharmacy at Toulouse (c. 1660). Finally, deciding to abandon living as a Christian, he moved to Holland, where in 1662 he changed his name from Baltazar to Isaac and became one of the leading intellectual figures and a medical practitioner in the Spanish-Portuguese Jewish community in Amsterdam. When he arrived in the Jewish community, he learned that there had been trouble about a former classmate of his from Spain, Juan de Prado. Prado was apparently involved with the young Spinoza and they were both charged with various heresies. Orobio wrote an answer, now lost, to one of Prado's works and against a work by Prado's son. Prado and his son held that the law of nature takes precedence over the law of Moses, and Orobio criticized their deism.

Orobio also wrote a metaphysical defense of his religion, based on mainly Spanish-Catholic Scholastic works and an answer to Alonso de Cepeda. His most famous works are an extremely rationalistic and Scholastic answer to Spinoza in geometrical form, *Certamen Philosophicum Propugnatum Veritatis Divinae ac Naturalis* (1684), which was published with Fénelon's *Demonstration de l'existence de Dieu*. The *Certamen* is the only critique of Spinoza by any member of the Jewish community that has survived and was considered one of the most important criticisms of Spinoza at the time.

Orbio engaged in a dialogue with one of the liberal Protestant leaders in the Netherlands, Philip van Limborch. They debated the truth of the Christian religion in 1687. This was a public debate where John Locke was present. The debate was published by Limborch under the title *Amica Collatio cum Erudito Judaeo* (1687) just after his opponent died, and Locke wrote a long review of it. Limborch met Orobio in Amsterdam in the 1680s and was much affected by his report of the Inquisition, which, through Limborch's *Historia Inquisitionis*, became for the next two centuries the best-known study of Inquisitorial investigation and torture methods. Orobio's most important anti-Christian work was *Prevenciones divinas contra la vana idolatria de las gentes*. He did not publish it because, as he explains in the note written in his own hand, he did not want to cause scandal, but he sent it to the Jesuits in Brussels, who liked it very much. It was published in French under the title *Israel vengé* (1770) by Baron d'Holbach. This work was used as important ammunition by French atheists against Christianity.

Through his works, Orobio de Castro showed an extremely acute understanding of metaphysics, using his knowledge of Spanish Scholasticism to buttress his religion against freethinkers and liberal and orthodox Christians. Some of his arguments against the doctrine of the Trinity are close to Spinoza's arguments against the plurality of substance.

See also Jewish Philosophy; Metaphysics.

Bibliography

WORKS BY OROBIO DE CASTRO

Additional works include *La observancia de la divina ley de Mosseh* (Coimbra, 1925), which has a preface by Moses

Bensabat Amzalak. There is also an English translation of the Limborch-Orobio debate (Farnborough, Gregg, 1969).

WORKS ABOUT OROBIO DE CASTRO

Angelis, Enrico de. "Crisi di conscienza fra I seicentisti per il metodo geometrico." *Annali della Scuola Normale Superiore di Pisa, lettere, storia e filosofia*, 2nd ser., 31 (1962): 253–271.

Carvalho, Joaquim de. *Orobio de Castro e o espinosismo*. In *Memorias da Academia das Ciências de Lisboa, Classe de letras*. Vol. 2. Lisbon: Academia das Ciências de Lisbon, 1937. A study on Orobio de Castro that includes the 1721 Spanish translation of Certamen Philosophicum.

Cohen, Monique-Lise, Ruth Attias, and Elie Szapiro. *Isaac Orobio de Castro et son temps: 1617–1687: juif portugais d'Espagne, de Toulouse et d'Europe*. Toulouse, France: Bibliotheque Municipale de Toulouse, 1992.

Kaplan, Yosef. *From Christianity to Judaism: The Story of Isaac Orobio de Castro*. Translated by Raphael Loewe. Oxford: Oxford University Press, 1989.

Révah, I. S. *Spinoza et Juan de Prado*. Paris: Mouton, 1959.

Richard Popkin (1967, 2005)

ORPHISM

"Orphism" is a modern term attached to two connected phenomena of Greek religion. The first is a body of traditional poetry, possibly from as early as the seventh century BCE, ascribed to a mythical singer called Orpheus and containing an account of the creation of the world and of the afterlife of the soul, its judgment and punishment for sins on Earth, and its final reincarnation in another living body. The second is the way of life adopted by those who accepted the truth of these writings, such truths being regarded with as much respect as the revelations in the traditional Greek "mysteries" at Eleusis and elsewhere.

CONTENTS OF ORPHIC WRITINGS

A number of fragments of the Orphic poems have survived, some of which belong to the poems as they were known in Athens in the fifth and fourth centuries BCE. However, these writings, in the manner of popular poetry, were constantly growing by accretion, and they seem to have become a general compendium of poetical accounts of theogony, cosmogony, and the soul's nature and fate. The contents of the poems as they existed in the fifth and fourth centuries BCE must be derived mainly from evidence in contemporary literature and, to a certain extent, in painting and sculpture.

ORPHEUS. It was in Greek art and literature of the sixth century BCE that Orpheus first appeared as a famous singer. The tradition that Orpheus sang while Musaeus wrote down his master's songs may reflect the moment of transition from oral to written literature—which probably occurred in the second half of the seventh century BCE—and this may be the time when these songs were composed.

To the poets of classical Greece, Orpheus was the singer possessed of supernatural powers. As such, he was enrolled among the Argonauts. According to an Alexandrian poet, Orpheus soothed his quarreling companions by singing to them of the creation of the world and of the dynasties of the gods. Euripides wrote of Orpheus's special connection with the underworld. A Naples bas-relief, executed at the end of the fifth century BCE, depicts his attempt to bring back his wife Eurydice from the dead. A little earlier in the same century, Polygnotus executed his famous picture of the underworld in which Orpheus was shown lyre in hand, amidst a group of legendary musicians.

It seems likely that this figure of Orpheus reflected the existing body of Orphic poetry, that his traits in fact represent its contents—a theogony which is an account of creation and a description of the underworld and of the soul's fate there.

THEOGONY. Plato's quotation of passages from an Orphic poem (in the *Cratylus* and *Philebus*) and Isocrates' description (in the *Busiris*) of what Orpheus wrote about suggest an Orphic theogony very like the one which is preserved as the work of Hesiod, the eighth-century BCE oral poet. From much later writers (Athenagoras, of the second century CE, and Damascius, of the fifth century CE) we learn of Orphic theogonies that contain non-Hesiodic elements—the cosmic egg and the creator Phanes. Since Phanes seems to be identifiable with the figure Eros that appears, together with the cosmic egg, in a cosmogony related in Aristophanes' fifth-century play *Birds,* both elements may accordingly be regarded as ancient. Three Orphic fragments joined by Otto Kern, which present a picture of the universe, may also be early since this picture of the universe, may also be early, since this picture bears a marked resemblance to Plato's image of the universe in the myth of Er at the end of the *Republic.* According to these fragments, the heaven, the earth, the sea, and the "signs with which the heaven is ringed" are abound round with a bond of Aether.

AFTERLIFE OF THE SOUL. Whereas Hesiod's *Theogony* contained a description of the underworld, inserted nominally in connection with the story of Zeus's overthrow of the Titans, this possibly traditional element was

developed in the Orphic poems into a detailed account of the soul's fate after death, its judgment and its reincarnation. Plato, throughout his writings, plainly drew on an account of the soul's late which he had read about in Orphic literature. In the *Gorgias* (493B) he refers to "one of the wise, who holds that the body is a tomb" and he also reports the story that the soul of an uninitiated man is like a sieve: In Hades the uninitiated is most miserable, being doomed to an eternity of filling sieves with water, by means of other sieves. Quoting the same story in the *Republic* (363D), he speaks of Musaeus and Eumolpus enlarging on the rewards of the righteous in the other world, and he also speaks of other who "when they have sung the praises of justice in that strain … proceed to plunge the sinners and unrighteous men into a pool of mud in the world below, and set them to fetch water in a sieve." In the *Phaedo* (69E) he says that "the man who reaches Hades without experiencing initiation will lie in mud, whereas the initiated when he gets there will dwell with the gods." In the *Cratylus* (420B) Plato attributes specifically to the Orphic poets the theory that the body is the tomb of the soul. Two surviving Orphic fragments (Kern Fr. 222) speak of the differing fates of the just and the unjust in the afterlife, and several (Kern Fr. 223ff.) deal with the rebirth of the soul in various forms. Plato must certainly have been referring to Orphic poems when he said in the *Meno* (81A) that among others "Pindar and many another poet who is divinely inspired … say that the soul of man is immortal, and at one time comes to an end, which is called dying, and at another is reborn, but never perishes. Consequently a man ought to live his life in the utmost holiness."

THE ORPHIC LIFE

For those who believed the eschatological dogma contained in the Orphic poems, there followed certain consequences for the conduct of life.

PROHIBITIONS. *Adikia,* injustice against any living creature, had to be strictly avoided. In Euripides' *Hippolytus* the diet "of food without soul," which was required of followers of Orpheus, is mentioned. Herodotus referred to the Orphic practice, which was also Pythagorean, of avoiding the use of wool (robbed from sheep) in burial. Men who observed these scruples might be described as living as "Orphic life," in the words of Plato in the *Laws.*

INITIATIONS. Proclus spoke of those who were initiated under Orpheus's patronage with Dionysus or Kore (in the case of the latter, at Eleusis). In Euripides' play *Rhesus,* Orpheus's amanuensis Musaeus is an Athenian, and Orpheus himself is closely connected with the Eleusinian initiations. It is certainly to these initiations that Aristophanes referred in the play *Frogs* when a character says, "Orpheus taught *teletai* [initiations] and abstinence from killing."

Evidently, the Orphic initiation had an essentially written character. Euripides referred to the person who observes Orphic scruples as "honoring the smoke of many writings." Plato mentioned "a mass of books" of Orpheus and Musaeus. Later writers contrasted this written initiation with the visual revelation at Eleusis, as when Pausanias wrote, "Whoever has seen an initiation at Eleusis or read the writings called Orphic knows what I mean." The Orphic literature seems to have borne the same relation to visual and oral instruction as a correspondence course bears to "live" teaching, and it appears to have been freely available.

Initiation into the mysteries was supposed to give a revelation of truth that would enable men to reach the next world in a state of guiltlessness. Plato reported that mendicant seers, who "frequented the doors of the rich," capitalized on this belief by offering cities and individuals the means of purification from sins committed. Among these are no doubt to be reckoned the *Orpheotelestai,* of whom Theophrastus spoke.

SIGNIFICANCE OF ORPHISM

Was Orphism, then, either a philosophy or a religion? It certainly was not a philosophical system, since in had no developed doctrine—merely a mythical account, derived from the popular oral poetry of the past, of the nature of the universe and of the afterlife of the soul. The philosophical importance of the Orphic literature lies in its influence, first of Pythagoras and Empedocles and then on Plato.

Pythagoras seems to have taken over the Orphic stories so completely that they could be referred to by Aristotle as Pythagorean stories, and earlier, Ion of Chios could say that Pythagoras had fathered his writings on Orpheus. The immortality and transmigration of the soul is the one doctrine which can certainly be attributed to the earliest Pythagorean society; Plato spoke of a Pythagorean way of life, based, as we know from other sources, on ritual prescriptions designed to ensure the purity and blamelessness of the soul.

Empedocles, who lived in Sicily in the fifth century BCE, exhibited a similar belief in the soul's immortality and transmigration.

In the *Symposium* Plato does not appear to believe in the soul's immortality, but in the *Meno* he accepts the preexistence and survival of the soul on the authority of "divinely inspired poets," among whom Orpheus in certainly to be reckoned. This doctrine became a cornerstone of Plato's entire metaphysical system.

Orphism was not in itself a religion, although it was closely related to the initiations at Eleusis and elsewhere, which were perhaps the most striking religious manifestations of classical Greece. The Orphic element was, however, merely a traditional poetical account that provided the eschatological dogma that was the basis for certain observances to the described as a way of life. The religious depth of this way of should not be exaggerated. There were no organized rituals, religious communities, or priesthood. In the sense in which we ordinarily use the word *religion* in the study of the ancient world, Orphism was not a religion.

See also Aristotle; Empedocles; Plato; Proclus; Pythagoras and Pythagoreanism.

Bibliography

Guthrie, W. K. C. *Orpheus and Greek Religion.* London: Methuen, 1935.

Kern, Otto. *Orphicorum Fragmenta.* Berlin, 1922; reprinted 1963.

Linforth, Ivan M. *Arts of Orpheus.* Berkeley: University of California Press, 1941.

Nilsson, M. P. "Early Orphism and Kindred Religious Movements." *Harvard Theological Review* 28 (1935): 181–230.

John Morrison (1967)

ORPHISM [ADDENDUM]

A number of archaeological discoveries in the second half of the twentieth century have considerably supplemented our knowledge about Orphism.

OLBIA BONE PLATES

The bone plates, found in Olbia on the Black Sea and dated to the fifth century BCE, probably functioned as tokens for those who received initiation. The inscription on one of them suggests that initiates could identify themselves as Orphics, even if the initiation did not necessarily imply any radical reform in their lifestyles.

GOLD LEAVES

The gold leaves are tiny inscribed gold strips buried with the dead containing instructions on what to do and what to say in the underworld. Gold leaves were found in Southern Italy, in Thessaly, and on Crete. Though there was a period of skepticism, newly found specimens make it likely that gold leaves were used by Orphic initiates.

THE DERVENI PAPYRUS

The Derveni papyrus was found in 1962 in a small sepulchral site near Thessalonica. The text was probably composed in the first half of the fourth century BCE, and its author might have been one of the Orphic initiates (*orpheotelestai*) that Plato and Theophrastus talked about. The first part of the text develops a rationalizing explanation of ritual acts, and quotes Heraclitus' fragments B3 and B94 in such a way that suggests that these fragments originally formed one sentence. In the second part the author interprets verses from a poem he attributes to Orpheus, some of which we know from other Orphic theogonies. The poem focuses on an episode when Zeus swallows all existing beings, so that for a moment everything is contained in him. Zeus then brings them back to light, and the story continues with the birth of new gods. This allows the poet to say, "Zeus is the head, Zeus is middle, and from Zeus all things get their being" (frag. 14.2, Bernabé). This episode expresses in the language of myth some central concerns of the pre-Socratic philosophers, such as the one/many problem and the question of the ultimate source of everything. The commentator interprets the poem allegorically, claiming that it propounds a cosmological theory. He argues that the different divine names in the poem designate the different cosmic functions of a unique god who created the present world order from primordial chaos. This unique god is called Mind (Nous) and is identified with the element air. The commentator's interpretation is heavily influenced by Anaxagoras and, to a lesser extent, Archelaus and Diogenes of Apollonia.

See also Anaxagoras of Clazomenae; Diogenes of Apollonia.

Bibliography

Bernabé, Alberto. "Orphisme et présocratiques: Bilan et perspectives d'un dialogue complexe." In *Qu'est-ce que la philosophie présocratique*, edited by André Laks and Claire Louguet. Villeneuve-d'Ascq, France: Presses universitaires du Septentrion, 2002.

Betegh, Gábor. *The Derveni Papyrus: Cosmology, Theology, and Interpretation*. Cambridge, U.K.: Cambridge University Press, 2004.

Gábor Betegh (2005)

ORTEGA Y GASSET, JOSÉ

(1883–1955)

José Ortega y Gasset, the Spanish essayist and philosopher, was born in Madrid of a patrician family. He was educated at a Jesuit college near Málaga and at the University of Madrid, where he received a doctorate in philosophy in 1904. Ortega spent the next five years at German universities in Berlin and Leipzig and at the University of Marburg, where he became a disciple of the neo-Kantian philosopher Hermann Cohen. Appointed professor of metaphysics at the University of Madrid in 1910, he taught there until the outbreak of the Spanish Civil War in 1936. During those years he was also active as a journalist and as a politician. In 1923 he founded the *Revista de occidente,* a review and series of books that was instrumental in bringing Spain into touch with Western, and particularly German, thought. Ortega's work as editor and publisher, as a contribution toward "leveling the Pyrenees" that isolated Spain from contemporary culture, ranks high among his achievements.

Ortega led the republican intellectual opposition under the dictatorship of Primo de Rivera (1923–1930), and he played a part in the overthrow of King Alfonso XIII in 1931. Elected deputy for the province of León in the constituent assembly of the second Spanish republic, he was the leader of a parliamentary group of intellectuals known as *La agrupación al servicio de la república* (In the service of the republic) and was named civil governor of Madrid. This political commitment obliged him to leave Spain at the outbreak of the Civil War, and he spent years of exile in Argentina and western Europe. He settled in Portugal in 1945 and began to make visits to Spain. In 1948 he returned to Madrid, where, with Julián Marías, he founded the Institute of Humanities, at which he lectured. By the time of his death, Ortega was the acknowledged head of the most productive school of thinkers Spain had known for three centuries, and he had placed philosophy in Spain beyond the reach, not of opposition and criticism, but of the centuries-old reproach that it was un-Spanish or antinational and therefore either a foreign affectation or a subversive danger.

WRITINGS AND STYLE

Ortega was a prolific writer. His numerous volumes consist mostly of essays and newspaper or magazine articles of general cultural interest. He wrote fewer strictly philosophical works; his vast influence on Spanish philosophy was exercised chiefly through his teaching.

All of Ortega's works are written in magnificent prose. He wrote in a clear, masculine style, and his mastery of Castilian has seldom been surpassed. On the other hand, he had a tendency to be wordy and to be content with literary brilliance and striking metaphor when argument and explanation were crucial.

Ortega's literary gifts had other, more important consequences. He used them to create a philosophical style and technical vocabulary in a tongue that until then had lacked models for philosophical writing and words for many modern concepts. But his literary virtuosity disarmed criticism in much of the Spanish-speaking world, so that his followers have often confounded philosophy with fine writing and emotional declamation.

RATIO-VITALISM

Ortega called his philosophy the "metaphysics of vital reason," or "ratio-vitalism." By metaphysics he meant the quest for an ultimate or radical reality in which all else was rooted and from which every particular being derived its measure of reality. He found this ultimate reality in Life, a word that he first used in a biological sense, like the vitalists, but which soon came to mean "my life" and "your life"—the career and destiny of an individual in a given society and at a certain point in history. In his first philosophical book, *Meditaciones del Quijote* (1914), Ortega sought to go beyond the opposition of idealism (which, he claimed, asserted the ontological priority of the self) and realism (which asserted the priority of the things the self knows). He asserted that in truth self and things were constitutive of each other, each needing the other in order to exist. The sole reality was the self-with-things: *Yo soy yo y mi circunstancia* (I am I and my circumstances). The things around me, he said in the *Meditaciones,* "are the other half of my personality." The experience-matrix comprising self and things is not simply one of coexistence, because the self acts on things and realizes itself in so doing. This activity is life, the dynamic interaction of mutually dependent self and things in the course of which the self carries out a mission of self-fulfillment.

PERSPECTIVISM

Ortega called his theory of knowledge "perspectivism." The world can be known only from a specific point of view. There is no possibility of transcending one's relative perspectives through absolute or impartial knowledge. "The definitive being of the world is neither mind nor matter nor any determinate thing but a perspective." Each perspective is unique, irreplaceable, and necessary, and all are equally true: "The only false perspective is the one that claims to be the one and only perspective." Ortega joined perspectivism to his notion of life as comprising the matrix self-with-things in the declaration, "Each life is a point of view on to the universe."

REASON AND LIFE

Although the *Meditaciones* seemed to place Ortega in the vitalist tradition, he dissociated himself from its antirationalism. Rather, just as he reconciled idealism and realism, he proposed to reconcile rationalism and vitalism. He agreed with the vitalists to "dethrone Reason," to dismiss abstract reason and bring it back to its rightful role as "only a form and function of Life." Yet Ortega stressed so strongly the rationality of the élan vital at the human level and underscored so firmly man's dependence on reason as an instrument for coping with life that he appeared to enthrone reason again beneath a vitalist disguise. He used the terms "Life" and "Vitality" to describe man's restless search for knowledge, understanding, and spiritual satisfaction, which others would have called "intelligence" or "practical reason." In fact, Ortega seemed to identify vitality and reason: Thus, in *En torno a Galileo* (1933), he wrote, "Living means being forced to reason out our inexorable circumstances." Therefore, ratio-vitalism was more rationalism than vitalism, and Ortega's thought was far removed from the irrationalist, romantic vitalism that flourished after World War I.

EXISTENTIALISM

Later, when Ortega appeared to have joined the existentialists (or, as he would have said, was joined by them), his insistence on the role of reason in the existential predicament gave his theories a distinctive color and allowed him to pour scorn on the sentimentalism of French existentialism. Ortega's dissociation from vitalism became complete when he took account of "the historical horizons of human life"—that is, of the social and cultural conditions of vitality in humankind. He gradually came to prefer the term "historical reason" to "vital reason." Life for Ortega now meant not biological vitality but "one man's life," and the vocation of the self was now conceived as what it must do with things—a mission of self-realization. This is the language of existentialism, and Ortega spoke it with a rare eloquence.

> Man does not have a nature, but a history.… Man is no thing, but a drama.… His life is something that has to be chosen, made up as he goes along, and a man consists in that choice and invention. Each man is the novelist of himself, and though he may choose between being an original writer and a plagiarist, he cannot escape choosing.… He is condemned to be free.… Freedom is not an activity exercised by an entity that already possessed a fixed being before and apart from that activity. Being free means … being able to be something else than what one is and not being able to settle down once and for all in any determined nature.… Unlike all the other things in the universe which have a pre-fixed being given to them, man is the only and almost inconceivable reality that exists without having an irrevocably pre-fixed being.… It is not only in economics but also in metaphysics that man must earn his living [*ganarse la vida,* win his life]. (*Historia como sistema*)

Each man has one best choice, and this is his imperative vocation or mission. "'Mission' means the awareness that each man has of his most authentic self which he is called upon to realize. The idea of mission is a constitutive ingredient of the human condition.… The being of man is at one and the same time natural and extranatural, a sort of ontological centaur" (*Obras completas,* Vol. V, pp. 209, 334). Ortega's moral theory thus derives directly from his anthropology; and indeed it is difficult, as with other existentialists, to separate his metaphysics from his anthropology and ethics. The moral life is the authentic one, the one that stays faithful to a life project or vocation; the immoral life is to abandon oneself to transient, outside influences, to drift instead of realizing a personal destiny. The choice of one personality out of the various possible personalities engages the whole of a man's reasoning powers and requires perpetual lucidity and concentration. This helps to explain Ortega's emphasis on the rationality of the élan vital at the human level. It is by intelligent reckoning with his circumstances that a man gains his being and becomes himself. Reasoned choice is constitutive of human personality.

SOCIAL THEORY

Life is always a problem, an insecurity, a "shipwreck," not only for the individual but for societies too. The desper-

ate measures society takes to struggle against perpetual foundering constitute human culture. It was Ortega's social theory, set forth in *La rebelión de las masas* in 1930 (*The Revolt of the Masses,* New York, 1931), that first brought him international recognition. Ortega started from the belief that culture is radically insecure and that a constant effort is required to prevent it from lapsing into barbarism and torpor. That effort is beyond most men, who can merely contribute to it by accepting the leadership of a liberal aristocracy, which does most of humanity's works. The fact that men have no essence or fixed nature but each must choose himself implies their inequality. "Because the being of man is not given to him but is a pure imaginary possibility, the human species is of an instability and variability that make it incomparable with animal species. Men are enormously unequal, in spite of what the egalitarians of the last two centuries affirmed and of what old-fashioned folk of this century go on affirming" (*Meditación de la técnica,* p. 42).

Ortega distinguished interindividual from social relations. In the former, which include love and friendship, individuals behave as rational and responsible persons, whereas in social relationships, which include customs, laws, and the state, we encounter the irrational and impersonal, the imposed and anonymous. The resulting contrast of man and people (*El hombre y la gente*), of the individual and the collectivity, betrayed Ortega's aristocratic distrust of democracy and contemporary mass society. There is no collective soul, he said, because "society, the collectivity, is the great soulless one, because it is humanity naturalized, mechanized and as if mineralized." Everything that is social or collective is subhuman, intermediate between genuine humanity and nature; it is a "quasi-nature." Nevertheless, social relationships have their uses; they make other people's behavior predictable, they carry on inherited traditions, and by automatizing part of our lives, they set us free for creation in the important interindividual sphere. These gains of socialization need constant defense, for men's antisocial drives are never vanquished. Society is neither spontaneous nor self-perpetuating. It has to be invented and reinvented by a minority that, however, must be able to procure the cooperation of the masses. The elite is essential to any society; by proposing a project for collective living, it founds the community and then governs and directs it.

The masses are incapable of framing a project, for they live without plan or effort. When they revolt and claim to govern themselves, society is threatened with dissolution. Ortega thought this was happening in twentieth-century democracies, whether totalitarian, communist, or parliamentary. Nationalism was exhausted as a collective project, and the next plan had to be supranational. Ortega favored the "Europeanization of Spain" in a supranational entity governed by an irreligious intellectual elite. Catholicism was to be extirpated, but gradually and cautiously, with a first stage of "liberal religion" leading toward the secular state.

The sensitive intellectual would have as little as possible to do with governing, for it was inevitably degrading. "There is no political health when the government functions without the active cooperation of majorities. Perhaps this is why politics seems to me a second-class occupation" (*Invertebrate Spain,* p. 201).

ARISTOCRATIC LOGIC

The notion of an aristocracy of talents is the key to Ortega's logic. In *Ideas y creencias* ("Ideas and Beliefs," in *Obras completas,* Vol. V, pp. 377–489), he claimed that ideas are the personal creation of the thinking minority, while the mass lazily accepts plain commonsense beliefs that in reality are vulgar ruling opinions imposed by "a diffuse authoritarianism." The archetype of mob belief is empiricism, or as Ortega called it, "sensualism." Sensualism is a reliance on the evidence of the senses, on self-evident truisms, on experiments in science or on documents in history. Philosophy since Parmenides has been a reaction against the vulgar prejudice in favor of the senses. "Against the *doxa* of belief in the senses, philosophy is, constitutionally and not accidentally, *paradox*" (*La idea de principio,* p. 285).

These views were developed with remarkable vigor in his unfinished, posthumously published magnum opus, *La idea de principio en Leibniz y la evolución de la teoría deductiva* (Buenos Aires, 1958; *The Idea of Principle in Leibniz and the Evolution of Deductive Theory,* New York, 1971). He assailed every form of the belief that principles or axioms can be founded on sensible intuition, taking Aristotle as the first representative of this belief and following its transmission through the Stoics and Scholastics. Such a belief, Ortega declared, is "idiot," "plebeian"; it results from a mental derangement akin to catalepsy, in that it entails sitting bemused before brute reality instead of thinking creatively. The only principles available to us, he held, are posed arbitrarily by the mind. They are assumptions that cannot be proved to the satisfaction of the senses, but "prove themselves" by allowing the deduction of a coherent corpus of propositions. This is the advance of post-Cartesian thought over traditional realism. "Modern philosophy no longer begins with

Being but with Thought" (*La idea de principio,* p. 263). The only proof modern philosophy knows is theoretical use: If axioms or methods give good results, there is no more to be said.

> Principles can only come from the understanding itself as it is before and apart from any acquaintance with sensible things. From these purely intellectual principles may be deduced consequences that form a whole world of intellectual determinations, that is, of ideal objects. … The activity of knowing used to seem to consist in an effort to reflect, mirror, or copy in our mind the world of real things, but it turns out to be just the opposite, namely, the invention, construction, or fabrication of an unreal world. (*La idea de principio,* p. 394)

Since he considered this idealist logic a characteristically aristocratic attitude, Ortega thought it significant that Plato and René Descartes, the two men who did most to construct it, were of noble blood. In contrast, the empiricism of Aristotle was popular, vulgar, "demagogic." "It is the criteriology of Sancho Panza. Faith in the senses is a traditional dogma, a public institution established by the irresponsible and anonymous opinion of the People, the collectivity" (*La idea de principio,* p. 286). Even the principle of contradiction, "that dogma of ontological sensualism," was a mere commonplace of the collective mind, unsupported by reasons and anything but self-evident. Aristotle had failed to prove the principle of contradiction, that *A* could not both be and not be *X,* and Immanuel Kant's transcendental deduction of it had no force. Ortega was not seeking to dispense with that principle but to argue that it could not be proven. Logic is a calculus tested by coherence, not an abstraction from sensible experience. Principles are assumptions that are useful for particular purposes.

Philosophy, science, and mathematics are "pure exact fantasy" based on principles that are arbitrary conventions. They are phantasmagoria, not far removed from poetry. They are the creation of an aristocracy of intellect that reveals the characteristics of all aristocracies: playfulness, lack of seriousness, and love of sport and games. Ortega meant quite literally that logic and science were games played according to strict but perfectly gratuitous rules by a minority that seeks to escape the tedium, vulgarity, and deadly seriousness of the world of beliefs. We never really believe in science or philosophy; they remain "mere ideas" to play with, and they are always somewhat spectral and unserious compared with the visceral faith we put into beliefs. Theory, like any fantasy, is by definition always revocable. Therefore, we ought to play at philosophy, jovially and without pathos, with the mock seriousness required to "obey the rules of the game."

See also Aristotle; Descartes, René; Existentialism; Idealism; Kant, Immanuel; Marías, Julián; Parmenides of Elea; Plato; Rationalism; Realism; Vitalism.

Bibliography

WORKS BY ORTEGA

Meditaciones del Quijote. Madrid: Ediciones de la Residencia de Estudiantes, 1914. Translated by E. Rugg and D. Marin as *Meditations on Quixote.* New York: Norton, 1961.

España invertebrada. Madrid: Calpe, 1922. Translated by M. Adams as *Invertebrate Spain.* New York: Norton, 1937.

El tema de nuestro tiempo. Madrid: Calpe, 1923. Translated by James Cleugh as *The Modern Theme.* London: C.W. Daniel, 1931.

La deshumanización del arte. Madrid: Revista de Occidente, 1925. Translated by W. Trask as *The Dehumanizatian of Art.* New York, 1956.

En torno a Galileo. Madrid, 1933. Translated by M. Adams as *Man and Crisis.* New York: Norton, 1958.

Meditación de la técnica. Madrid, 1933.

Estudios sobre el amor. Madrid, 1939. Translated by T. Talbot as *On Love.* New York: Meridian, 1957.

Del imperio romano. Madrid, 1940–1941. Translated by H. Weyl as *Concord and Liberty.* New York: Norton, 1946.

Historia como sistema. Madrid: Revista de Occidente, 1941. Translated by H. Weyl, E. Clark, and W. Atkinson as *Toward a Philosophy of History.* New York: Norton, 1941. 2nd ed., *History as a System.* New York: Norton, 1961.

El hombre y la gente. Madrid: Revista de Occidente, 1957. Translated by W. Trask as *Man and People.* New York: Norton, 1959.

Que es filosofia? Madrid: Revista de Occidente, 1957. Translated by M. Adams as *What Is Philosophy?* New York; Norton, 1960.

Obras ineditas. 7 vols. Madrid and Buenos Aires, 1957–1961. Contains the later and posthumously published works.

The Idea of Principle in Leibniz and the Evolution of Deductive Theory. Translated by Mildred Adams. New York: Norton, 1971.

Obras completas. Vols. 1–12. Madrid: Alianza Editorial, 1983.

The Revolt of the Masses. Translated by Anthony Kerrigan; edited by Kenneth Moore. Notre Dame, IN: University of Notre Dame Press, 1985.

WORKS ON ORTEGA

Borel, J. *Ratson et vie chez Ortega y Casset.* Neuchâtel, 1959.

Cascalès, Charles. *L'humanisme d'Ortega y Gasset.* Paris, 1956.

Ceplecha, C. *The Historical Thought of José Ortega y Gasset.* New Haven, CT, 1957.

Dobson, Andrew. *An Introduction to the Politics and Philosophy of José Ortega y Gasset.* New York, 1989.

Donoso, A., and H. C. Raley. *José Ortega y Gasset: A Bibliography of Secondary Sources.* Bowling Green, OH: Philosophy Documentation Center, 1986.

Guy, Alain. *Ortega y Gasset, critique d'Aristote.* Paris, 1963.

Holmes, Oliver W. *Human Reality and the Social World: Ortega's Philosophy of History.* Amherst: University of Massachusetts Press, 1975.

Marías, Julián. *La escuela de Madrid.* Buenos Aires, 1959.

Marías, Julián. *Historia de la filosofia.* Madrid, 1941.

Marías, Julián. *Ortega, circunstancia y vocación.* Madrid, 1960.

Marías, Julián. *Ortega y tres antípodas.* Buenos Aires, 1950.

Ouimette, Victor. *José Ortega y Gasset.* Boston: Twayne, 1982.

Ramírez, Santiago. *La filosofia de Ortega y Gasset.* Barcelona, 1958.

Ramírez, Santiago. *Un orteguismo católico.* Salamanca, 1958.

Ramírez, Santiago. *La zona de seguridad.* Madrid, 1959.

Rodríguez Huéscar, Antonio. *José Ortega y Gasset's Metaphysical Innovation.* Trans. and ed. by Jorge García-Gómez. Albany: SUNY Press, 1995.

Silver, Philip W. *Ortega as Phenomenologist: The Genesis of "Mediations on Quixote."* New York: Columbia University Press, 1978.

Weintraub, Karl J. "Ortega y Gassét." In *Visions of Culture.* Chicago: Chicago University Press, 1966.

Neil McInnes (1967)
Bibliography updated by Philip Reed (2005)

OSTWALD, WILHELM

(1853–1932)

Wilhelm Ostwald was a German chemist, philosopher, and historian of science whose main scientific achievement was his pioneer work in physical chemistry, particularly in electrochemistry. With J. H. van't Hoff he founded the *Zeitschrift für physikalische Chemie* in 1887. He was awarded the Nobel Prize in chemistry in 1909.

ENERGETISM

Ostwald's philosophical outlook, known as energetism or energetic monism, was strongly influenced by his scientific background and by the state of physical science at the end of the nineteenth century. In particular, the first and second laws of thermodynamics—the law of conservation of energy and the law of entropy—decisively influenced his thought. Ostwald claimed that energy is the substrate of all phenomena and that all observable changes can be interpreted as transformations of one kind of energy into another. This claim was based on both epistemological and physical considerations. Ostwald pointed out that we never perceive anything but energy, or more accurately, differences in energy. One never perceives a material substance itself, but only its energetic interaction with his own organism.

In an argument similar to a classical argument of René Descartes's, Ostwald showed that even impenetrability, which, according to mechanists, is the constitutive feature of matter, is a mere sensory quality that is perceived only when there is a difference in kinetic energy between a piece of matter and one's own organism. No sensation of hardness would arise if a piece of matter which one tried to touch retreated at the same velocity with which his finger moved toward it. Ostwald interpreted all aspects of matter in terms of energy: Mass is the capacity of kinetic energy; occupancy of space is "volume-energy"; gravity is energy of distance. Thus, matter is nothing but a "spatially ordered group of various energies" which do not require any material substrate. Material substance belongs with caloric, phlogiston, and electric and magnetic fluids in the category of discarded and useless fictions. Ostwald prophesied that ether too would soon disappear from science, as the increasing difficulties in constructing a satisfactory model of it indicated.

This difficulty was for Ostwald only one symptom of mechanism's general failure to provide a satisfactory explanation of physical phenomena. He even doubted the usefulness of kinetic explanations of thermal phenomena, although the mechanical theory of heat had been extremely successful. The atom itself was for Ostwald only a convenient methodological fiction, which he refused to reify. (Only around 1908, under the growing pressure of new experimental confirmations of the discontinuous structure of matter, did he modify his view.)

The ubiquity and constancy of energy make it "the most general substance," and the conservation of energy underlies the validity of the law of causation. The succession of cause and effect is nothing but the transformation of one form of energy into another, the total amount of energy remaining constant. The law of conservation of energy guarantees the quantitative equality of cause and effect; and the direction of transformations is determined by the law of entropy, according to which all forms of energy are being gradually transformed into heat. Ostwald rejected all attempts to limit the application of the law of entropy; opposition to applying it to the whole of cosmic history was, in his view, nothing but emotional reluctance to accept the eventual death of civilization and even of humankind. The mechanistic view, which regards all processes as in principle reversible, fails to account for the irreversibility of time embodied in the law of entropy.

Ostwald belonged to a generation of philosophers of science that included Ernst Mach, Pierre Duhem, and J. B. Stallo, who were acutely aware of the limitations of mechanistic explanations. They overlooked the power and fruitfulness of mechanical and particularly of corpuscular models even on the molecular level, and atomic physics was not yet advanced enough to show the inadequacy of corpuscular models of subatomic phenomena. When this inadequacy became apparent, the crisis of the traditional scheme proved to be far more profound than Ostwald expected. While claiming to reduce all manifestations of matter to energy, he still retained mass, the basic concept of mechanism, under the disguised form of "capacity of energy." He anticipated the later relativistic fusion of mass and energy only in a hazy and qualitative way.

In this respect Ostwald can be compared with Herbert Spencer, with whom he shared other ideas: the substantialization of energy, the deduction of the causal law from the law of conservation of energy, an energetist approach to social science and ethics, and a determinist monistic metaphysics disguised by positivistic and agnostic formulas. Ostwald, however, lacked Spencer's philosophical sophistication; this is especially visible in his approach to the mind-body problem. Ostwald believed that he had refuted materialism by identifying consciousness with neural energy; he did not realize that his view was only a variant of physicalism. Like Ernst Haeckel, whom he greatly respected, Ostwald believed that his view was identical with Benedict de Spinoza's double-aspect theory, but this is not true. The haziness of Ostwald's monism invited criticism from antagonistic camps; Hans Driesch called it disguised materialism, and V. I. Lenin denounced it as "sheer idealism."

Ostwald devoted much time to propagating his views on monism. He founded the pantheistically oriented League of German Monists in 1906, and in 1911 he began to publish the series Monist Sunday Sermons (*Monistische Sonntagspredigten*).

ETHICS AND SOCIAL THOUGHT

Ostwald regarded the law of entropy as the basis for the theory of values. What we term *mind* or *consciousness* is nothing but a form of neural energy and is subject to the same laws as other forms of energy. In a temporally reversible world the concept of value would be meaningless, whereas it acquires a precise scientific meaning in the framework of energetism. Evolutionary advance consists in the fact that increased coordination between increasingly specialized organs results in increased efficiency of the organism and a minimum waste of energy. The same law—increased coordination resulting in maximum efficiency—determines the progress of civilization. Immanuel Kant's categorical imperative should be replaced by the "energetic imperative": "Do not waste your energy." Ostwald's applications of his energetic imperative to social thought were even more ambiguous than his views on the mind-body problem. Prior to 1914 Ostwald regarded war and conflict as a wasting of energy, and he favored internationalism and pacifism. But during World War I he justified his militant nationalism by claiming that the organization, efficiency, and minimum waste of energy of the German state represented the highest existing evolutionary form of human society.

HISTORY OF SCIENCE

In history of science Ostwald deserves credit for editing *Ostwalds Klassiker der exacten Wissenschaften,* a series of reprints of important scientific writings. His own classification of creative scientific minds into "classics" and "romantics," however, is probably oversimplified although interesting. Ostwald also founded and edited the journal *Annalen der Philosophie* (1901–1921).

See also Causation: Philosophy of Science; Chemistry, Philosophy of; Descartes, René; Duhem, Pierre Maurice Marie; Energy; Haeckel, Ernst Heinrich; Lenin, Vladimir Il'ich; Mach, Ernst; Materialism; Mind-Body Problem; Nationalism; Philosophy of Science, History of; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY OSTWALD

Die Überwindung des wissenschaftlichen Materialismus. Leipzig: Veit, 1895.

Vorlesungen über die Naturphilosophie. Leipzig, 1895; 2nd ed., Leipzig, 1902.

Individuality and Immortality. Boston: Houghton Mifflin, 1906. The Ingersoll Lecture.

Grundrisse der Naturphilosophie. Leipzig, 1908.

Die energetische Grundlagen der Kulturwissenschaften. Leipzig, 1909.

Der energetische Imperativ. Leipzig: Akademische Verlagsgesellschaft, 1912.

Monism as the Goal of Civilization. Hamburg, 1913.

Die Philosophie der Werte. Leipzig: A. Kröner, 1913.

Lebenslinien. 3 vols. Leipzig, 1926–1927. Autobiography.

Wissenschaft und Gottesglaube. Edited by F. Herneck. Leipzig, 1960.

WORKS ON OSTWALD

Adler, F. W. *Die Metaphysik in der Ostwaldschen Energetik.* 1905.

Delbos, Victor. *Une théorie allemande de culture: W. Ostwald et sa philosophie.* Paris, 1916.

Driesch, Hans. *Naturbegriffe und Natururteile.* Leipzig: Wilhelm Engelmann, 1904.

Duhern, Pierre. *L'évolution de la mécanique.* Paris: A. Joanin, 1903. Esp. p. 179.

Lasswitz, K. "Die moderne Energetik in ihrer Bedeutung für die Erkenntniskritik." *Philosophische Monatshefte* 39: 1–30, 177–197.

Lenin, V. I. *Materialism i Empirio-krititsizm.* Moscow: Zveno, 1909. Translated by David Kirtko and Sidney Hook as *Materialism and Empirio-Criticism.* London: Lawrence, 1927. Esp. Ch. 5

Meyerson, Émile. *L'identité et réalité.* 5th ed. Paris, 1951. Esp. Ch. 10. Translated from the 1908 edition by Kate Loewenberg as *Identity and Reality.* London, 1930.

Rey, Abel. *La théorie de la physique chez les physiciens contemporains.* 2nd ed. Paris, 1923.

Rolla, A. *La filosofia energetica.* Turin, 1908.

Schnehen, Wilhelm von. *Energetische Weltanschauung.* Leipzig, 1908.

Milič Čapek (1967)

OTHER MINDS

The question of how all of us know that there are other beings besides themselves who have thoughts, feelings, and other mental attributes has been widely discussed, especially among analytic philosophers in the English-speaking world. At least three of the most influential German philosophers—namely, Edmund Husserl, Max Scheler, and Martin Heidegger—have also dealt with this problem. The problem of "other minds" becomes a serious and difficult one because the traditional and most obvious solution to it, the argument from analogy, is open to grave objections. At the present time it would seem that a majority of the philosophers who have concerned themselves with the question consider the traditional solution—that our belief in other minds can be adequately justified by an analogical argument—at least inadequate, if not radically and unremediably defective.

ARGUMENT FROM ANALOGY

In general terms to argue by analogy is to argue on the principle that if a given phenomenon *A* has been found to be associated with another phenomenon *B*, then any phenomenon similar to *A* is very likely to be associated with a phenomenon similar to *B*. In the particular case of other minds, it is said, I observe that there is an association between my mental states, on the one hand, and my behavior and the physical state of my body, on the other. I then notice that there are other bodies similar to mine and that they exhibit behavior similar to my own. I am justified, therefore, in concluding by analogy that mental states like the ones I experience are associated with those other bodies in the same way that my mental states are associated with my body. I notice, for example, that when I have a pain in my tooth, it is likely to be decayed and that I am likely to groan, complain, and hold my jaw. Observing another body like my own that has a decayed tooth and behaves as my body behaves when I have a toothache, I conclude that this body, like mine, is the body of a being that has a toothache.

OBJECTIONS TO THE ANALOGY ARGUMENT. The first and least radical objection to the argument from analogy is that it does not establish its conclusion with an adequate degree of certainty. The argument, it is said, would be relatively strong if the correlation of the mental and the physical was observed to hold in a large and varied collection of instances before it was concluded that it also held in other similar cases. But this is not so. If I use the argument from analogy, I have only one case, my own, as a basis for my inference. Moreover, the characteristics and behavior of the other bodies vary markedly from my own. How can I be sure that the differences between myself and others are not associated with the presence of mental attributes in my own case and with the absence of them in other cases?

The other difficulties in the argument from analogy concern two features of that argument—first, that it is logically impossible to check up on the correctness of the conclusion of the argument and, second, that the argument's validity implies that one must learn from one's own case alone what it is to have a mental attribute. Let us elaborate a little on each of these points.

In the case of a normal analogical argument, it makes good sense to suppose that one might check up directly on the conclusion of the argument; in principle one could always dispense with reasoning by analogy, even though this may not be practicable in some cases. Of course, one who says that we know of the existence of other minds by analogy must deny that we can check up on our conclusion in some more direct way, for if we could, the argument by analogy with ourselves could be dispensed with. It also seems that he cannot say that our inability to check up is merely a practical matter. Such checking up cannot consist in making further observations of a person's behavior and body; this we can often do sufficiently well in practice. It would have to consist in some other opera-

tion that we cannot in fact perform but which we can conceive of ourselves performing; perhaps it would be something like telepathy.

But aside from any difficulty in making clear sense of the notion of telepathy, why should telepathy be regarded as a more direct way of checking up than ordinary observation of behavior? Indeed, it seems that one's grounds for thinking that one has telepathic knowledge of another person's state of mind must include the knowledge that what one seemed to know telepathically generally correlates well with what one knows as a result of ordinary observation. The same would also seem to apply to any other extraordinary but conceivable way of knowing about another's mental state. Granted, then, that the supporter of the argument from analogy must hold that the impossibility of checking the conclusion more directly is not any variety of empirical impossibility, why is this held to destroy the argument? Perhaps there is a difference here between this argument and other valid analogical arguments, but why does this difference make this argument unacceptable? The answer given is that this difference renders the conclusion of the argument senseless. What can the phrase "He is in pain" mean to me if no conceivable observation I could make would show that it was true or false, if I have no criterion for its truth, and if I have no idea of what would count for or against it? It will not do to say that the sentence means that he has the same as I have when I am in pain, for, again, what counts as being the same here?

The other main difficulty in the analogical argument centers, as we have said, on the necessity, implied by that argument, for each of us to learn from his own case alone what it is to have a mental attribute. Two arguments have been advanced to show that this is impossible.

According to the first, which derives from Ludwig Wittgenstein, the analogical argument requires that one be able to pick out something (for example, a pain or a state of anger) and thereafter to identify it, when it recurs, as a pain or a state of anger. The trouble is, however, that this account leaves no room for a distinction between a correct and an incorrect identification. Behavioral and other checks are ruled out, leaving no conceivable means of deciding whether a mistake has been made. But a distinction between a correct and a mistaken identification is surely essential to the very notion of identification itself. In this way the analogical argument, which requires that we be able to make correct identifications of our inner states, also deprives the notion of identification of any meaning.

The second argument, which has been advanced by P. F. Strawson, is more complex. According to him, the idea of a predicate involves the idea of a range of individuals to which that predicate can be significantly applied. In the case of mental attributes, this range includes both oneself and others; one cannot have the notion of a mental attribute unless one has a notion of oneself and a notion of another. Since the notion of oneself is the notion of a subject of mental and other attributes, one cannot have the notion of oneself without the notion of some mental attributes. Therefore, one cannot have a notion of oneself without also having the notion of another subject of mental attributes. This notion, however, can be possessed only if one knows how to ascribe mental attributes to such subjects. Hence, until one knows how to do this, one has no notion either of oneself or of another. But the argument from analogy requires that one should first have a notion of oneself, of one's own case, and then discover how to ascribe mental attributes to others by arguing analogically from correlations that are found to hold in one's own case. A person without a notion of his own case could indeed argue analogically. He could find that pain was to be expected when a certain body (his own, as *we* say) was branded with a hot iron. He could infer that there would also be a pain when another similar body was similarly affected. But he would soon find out that he was mistaken in this conclusion, for he would detect no pain when the hot iron was applied to any body other than his own.

DEFENSES OF ANALOGY ARGUMENT. Some persistent attempts (especially by A. J. Ayer) have been made to defend the argument from analogy against the charges laid against it. To counter the charge of weakness, the following suggestions have been made. Emphasis has been laid upon the special feature of the argument from analogy—that people can speak and that their descriptions of their mental states are very like those I would give of some of my own. This, it is claimed, is something more telling than a mere similarity of behavior. Against this it is pointed out that speech can be regarded as something understood by the speaker only if it is accompanied by the appropriate nonverbal behavior.

Another defense is that conclusions drawn analogically from behavioristic similarities are powerfully reinforced by like conclusions drawn by arguments based on similarities in the state of the nervous system. This consideration hardly meets the main complaint—namely, that I base my inference on one case only, my own.

According to a rather more convincing attempt to meet this complaint, no more can be asked of any method of inference than that I be able to test its conclusion more directly in some cases and that when I do so, the conclusion usually turns out to be correct. The argument from analogy satisfies this test. I can suppose that there are, as there seem to be, other people besides myself and that these people argue analogically that I have certain thoughts and feelings. I can check on these imagined inferences and find that their conclusions are generally true. Whether these inferences are in fact made is neither here nor there; I can see that the method would work if it were used. Nor need I be worried because I can check only those cases in which the conclusion is about myself. In all or most inferences there will be a restricted class of cases that I can check up on. It is, for instance, logically impossible that I should make a direct check on a change of color that occurred where I could not observe it. But it would be a mistake to argue that any analogical argument that a color change had occurred was weak because it was based upon one sort of case only—the sort that I was able to observe. Why should it make a difference to the strength of the other minds argument that the relevant class of case is my own mental states as opposed to what I myself observe?

An argument similar to this one can also be used to rebut the charge that there is no conceivable means of checking up on the conclusion of the argument from analogy. There are in fact some cases in which I can make a check—namely, those cases that concern myself. Moreover, although it is logically impossible for me to be some other person and hence to make a direct check on that other person's mental states, this is unimportant, for it is never logically impossible that I should check on the truth of a psychological statement when the subject is referred to by a descriptive phrase, even though that description fits someone other than myself. It is logically impossible, perhaps, that I should be Robinson, but it is not logically impossible that I should now be the man flying a certain aircraft, even though Robinson is in fact that man. Moreover, it is claimed, when I make a statement about Robinson, what is stated is, in effect, that someone who answers to such and such a description has had such and such an experience. To this it has been objected that the only interpretation of this claim that yields the desired conclusion is untrue, namely, the interpretation that "Robinson has a pain" means the same thing as some sentence of the form "The so and so has a pain." However, this objection clearly fails to settle the matter, as can be seen by considering the following statements:

(1) The man sitting in this chair is angry.

(2) Robinson is the man sitting in this chair.

(3) Robinson is angry.

Statement (1) cannot be said to be unintelligible to me on the ground that I, not being the man in question, cannot check up directly, for it is conceivable that I might have been sitting in the chair; statement (2) can also be checked on by me; statement (3) follows from (1) and (2). It is surely quite implausible to hold that statement (3) is unintelligible to me, whereas statements (1) and (2) are not.

There is, however, another possible difficulty in the argument from analogy that is usually not at all clearly distinguished from the one just considered—namely, that it is in principle impossible for more than one person to check directly on the conclusion. It is often said that publicity is the essential requirement. But does this mean that it must be logically possible for each person to make the check, or is it the more stringent requirement that it be possible for everyone, or at least more than one, to do so? If the latter, then the difficulty has not been overcome. Equally it has not been shown clearly why publicity should be required in the more, rather than in the less, stringent form.

This brings us to the reasons given for holding that one cannot understand psychological predicates from one's own case alone, which is a requirement of the argument from analogy. One of these reasons, as we have seen, is that there is no sense in the idea of an identification that is subject to no check, where there is no criterion of correctness. This view has been questioned on two grounds. Strawson has argued that a criterion of correctness is not needed in all cases of identification, and according to Ayer, an identification of a sensation can be satisfactorily checked, without recourse to anything publicly observable, by means of other private sensations.

OTHER SOLUTIONS TO THE PROBLEM

BEHAVIORISM. Assuming that the argument from analogy is unacceptable, the most obvious alternative is to adopt some form of that variety of behaviorism according to which all psychological expressions can be fully understood in terms of behavior. If behaviorism is correct, there is clearly no room or need for the argument from analogy. In ascribing a pain to someone, for example, one is asserting something that is in principle subject to a public check—something about the way the individual is behaving, about how he would behave in certain circumstances, about what the circumstances in fact are,

or the like. There is no need to make any inference from the publicly observable to something radically different.

This is not the place for a general discussion of behaviorism. Any objection to a given form of behaviorism will, of course, be an objection to that form of behaviorism as a solution to the problem of other minds. There is, however, one difficulty that has given rise to a number of closely related attempts to deal with the problem—namely, that it is implausible to give a behavioristic account of some first-person psychological statements. When, for example, I say that I have a terrible pain, I do not say this on the basis of observation of my own behavior and the circumstances in which I am placed. Nor am I speculating about how I would behave in other, hypothetical circumstances.

This difficulty has become of central importance for many philosophers who are impressed by some or all of the arguments that purport to refute the argument from analogy. They regard such arguments as showing, not only that this argument fails, but, more positively, that the connection between mental states, on the one hand, and behavior and circumstances, on the other, is logical or conceptual, not contingent. What is needed to remove the difficulty about our knowledge of other minds, it is thought, is to clear away the obstacles that prevent us from seeing clearly that this connection *is* a conceptual one. The primary obstacle in this instance is the peculiar nature of first-person psychological statements. It is this obstacle that prevents us from wholeheartedly accepting the true view and that makes us always hark back to the picture of mental states as objects to which the owner has privileged access.

There are at least two points involved here. First, if my own statements about my mental states are not about private happenings to which only I have access and if they are not about my behavior either, then what account *is* to be given of them? Second, the statement "I am in pain," made by me, contradicts the statement "He is not in pain," made about me by someone else. If one admits that the former is not about my behavior, how can one avoid the conclusion that the latter also is not about my behavior? But if the latter is not about my behavior, how can it be maintained that the connection between my pain and my behavior is a logical one?

WITTGENSTEIN. In dealing with the question "How do words refer to sensations?" Wittgenstein suggested, "Here is one possibility: words are connected with the primitive, the natural, expressions of sensation and used in their place" (*Philosophical Investigations*, Sec. 244). This suggestion, which is not elaborated much by Wittgenstein, has sometimes been treated as an attempt to deal with the first point stated above and has had certain merits ascribed to it—for example, by Norman Malcolm. It explains how the utterance of a first-person psychological statement can have importance for us; such an utterance has the importance that natural expressions of sensation and emotion have. It is also said to explain certain features of the logic of psychological statements, the absurdity of someone's concluding that he has a pain from the observation of his own behavior, and the impossibility of someone's being mistaken about whether he has a pain or of wondering whether he has a pain. However, whatever its merits, this stress on the likeness of first-person sensation statements to natural expressions of emotion and sensation merely sharpens the second of the difficulties noted above—namely, that "I am in pain" can contradict "He is not in pain." It even makes it hard to see how the former can be a statement at all; a cry of pain is not a statement.

This difficulty is obviously insuperable for one who, unlike Wittgenstein, adopts the extreme position that apart from being verbal and learned responses, first-person sensation statements are exactly like natural expressions of sensation. Wittgenstein, however, appears to hold that a statement like "My leg hurts" is never in all respects like a cry of pain but is sometimes more like it and sometimes less, depending on the context of utterance. There seem to be three main likenesses that he wanted to stress in all first-person present-tense expressions of sensation and in many such expressions of emotion—namely, (1) the impossibility of these expressions being mistakenly uttered; (2) the possibility of their being insincere or pretended; and (3) the fact that such statements can justifiably be made without a basis of self-observation. The problem that arises in formulating a successful defense of his views is showing how a statement that bears the above likenesses to a cry of pain can yet be different enough to contradict another statement for which the criteria of truth lie in the realm of the publicly observable—that is, in the behavior of the speaker.

It cannot be said that Wittgenstein himself made a serious attempt to cope with this difficulty. Others have made the attempt, but no attempt has been very convincing. The second and third points of likeness present no great difficulty (see Douglas Gasking, "Avowals"). Any statement can be made insincerely, and there are many nonautobiographical statements that a person can justifiably make without observing that the criteria for their truth are satisfied. For example, some people can tell you

that a certain note is middle C without first carrying out the tests that determine whether it has the appropriate frequency. For such statements to be justified, it is necessary only that those who make them are usually right in such cases.

Alleged incorrigibility. The first difficulty, which arises from the alleged incorrigibility (as it is termed) of first-person present-tense statements, is not so easily disposed of. The most hopeful approach—indeed, the only approach—is to exploit the fact that the natural expressions of sensation and emotion can be feigned. An insincere groan is akin to a lie, and a lie is a false statement. Perhaps a verbal expression can reasonably be called false if it is insincere and true if it is sincere, the distinction between sincerity and insincerity being a matter of the behavior of the speaker. In this way a plausible account could be given of how something very like a groan could also in some ways be like a statement and be regarded as such. The incorrigibility of such statements would then be accounted for.

But this is not enough; it does not explain how such a "statement" can be the contradictory of another statement that is logically connected with statements about the behavior of the maker of the "statement." For (1) "I have pain," said by me about myself, is the contradictory of (2) "I have not a pain," said by me about myself. Therefore, since (3) "He has a pain," said about me by someone else, is also the contradictory of (2), (1) and (3) must both be the same statement. Consequently, if (3) is logically connected with certain behavioral statements, (1) must also have these connections. This makes it difficult to see how (1) can be incorrigible. If I can be mistaken about my own behavior, as is the case, and if there is a logical connection between my pain and my behavior, then, it would seem, I can be mistaken about my pain. This difficulty is not overcome by assimilating the truth of a first-person pain statement to the sincerity of a groan. For (4) "I am sincere in saying I have a pain," said by me about myself, is the same statement as (5) "He is sincere in saying he has a pain," said about me by someone else. Therefore, if (5) is logically connected with statements about my behavior, so is (4), and, if (4) is so connected, it must, it seems, be corrigible. For to claim sincerely that *p* is to think that *p* when one makes the claim, and to claim insincerely that *p* is to think that not-*p* when one makes the claim. If (4) is corrigible, then someone might think he is sincere in claiming he has a pain when in fact he is insincere—that is to say, he might think that he thinks that he has a pain, although in fact he thinks that he has not a pain. If, however, one cannot be mistaken about one's own pain, then to think that one thinks one has a pain is to think one has a pain, and to think one has not a pain is not to have a pain. It follows that if (4) is corrigible, someone might think that he has a pain although, in fact, he has not a pain. In short, if (4) is corrigible and (1) is not, then (1) is corrigible.

There are apparently only two ways out of these difficulties that do not involve abandoning the thesis of the incorrigibility of first-person psychological statements and thus ceasing to attach much value to the assimilation of such statements to natural expressions of emotion and sensation. One might deny that (1) and (3) are the same statement, or one might maintain that although (1) is logically connected with behavioral statements about which I can be mistaken, yet I cannot be mistaken about (1). The first of these alternatives would involve finding a satisfactory explanation of why I cannot assert the same thing that someone else does when he asserts (3). The second would require an account of the notion of a logical connection that would allow for the existence of statements that, when made by myself, are incorrigible, but which are logically connected with other statements that, when made by myself, are not incorrigible.

In fact it has been argued by some that there are no psychological statements that are incorrigible and that the problem we have just been discussing is therefore an unnecessary one. It seems to be quite true that there are some ways in which one can be mistaken when one says one has, say, a pain. But the matter has not yet been clarified sufficiently for anyone to be justified in saying with confidence that this renders the problem unnecessary. Even if first-person present-tense pain statements are corrigible, this does not show that they are corrigible in all the ways that other statements are corrigible. Nor has it been shown convincingly that they are corrigible in such a way as to obviate any difficulty that may arise from the fact that "I have a pain," said by me, contradicts "He has a pain," said about me.

In addition to the above objections to Wittgenstein's views on the subject of psychological statements, there is another one that is of a less definite character and to which Wittgenstein himself alludes when he puts into the mouth of an imaginary objector such words as "and yet you again and again reach the conclusion that the sensation itself is a *nothing*" (*Philosophical Investigations*, Sec. 304). He protests, of course, that this is not the sort of impression he wishes to create and that it arises from his "setting his face against the picture of the inner process." Nevertheless, it cannot be said that he altogether succeeds in dispelling this impression. His problem might indeed

be described in just these terms—to set his face against the inner process picture without creating the impression that he wishes to deny the existence of sensations. It does not seem that he succeeds in this.

P. F. STRAWSON. It is perhaps Wittgenstein's failure that in part gives rise to another attack on the problem—namely, that of P. F. Strawson. Strawson, like Wittgenstein, is convinced that the argument from analogy is mistaken and that skepticism about other minds is senseless or at least empty and pointless. Like Wittgenstein, he holds that the relation of the behavior of other people to their mental states is not contingent: "the behavior-criteria one goes on [in assigning P-predicates—that is, psychological predicates] are not just signs of the presence of what is meant by the P-predicate, but are criteria of a logically adequate kind for the ascription of the P-predicate" (*Individuals*, p. 106).

In spite of this he is out of sympathy with Wittgenstein in many ways. He considers that the assimilation of first-person present-tense psychological statements to the natural expressions of sensation and emotion "obscures the facts and is needless" (*Individuals*, p. 107). He is unconvinced by Wittgenstein's reasoning against the idea of a private language that might serve as a basis for the argument from analogy. He sees little difficulty in the notion of a person's inventing for himself a private language in which he has names for his sensations even when such sensations have no outward expressions: "He might simply be struck by the recurrence of a certain sensation and get into the habit of making a certain mark in a different place every time it occurred" (*Individuals*, p. 85). Nor does he consider the notion of a person's continuing to exist in a disembodied state as logically absurd (*Individuals*, pp. 115–116). He accuses Wittgenstein of hostility to the idea of what is not observed and of a "a prejudice against the inner" ("Critical Notice," p. 91).

All these criticisms of Wittgenstein suggest that Strawson holds the view that the connection between behavior and mental states is, after all, a contingent one. But this, as we have seen, is not so. How, then, does Strawson reconcile these apparently conflicting aspects of his thought? His line of thought appears to be approximately that general agreement in judgment is necessary before it is possible to have a common language. Such general agreement exists about, for example, "what it looks like here," and this agreement makes possible our common impersonal language of, for example, color. There is no such general agreement about "whether or not 'it's painful here,' " and there is thus no possibility of a common impersonal pain language. However, there is something available (namely, pain behavior) on which general agreement is possible, and if we are therefore to have a common pain language, we must each ascribe pain to others on the basis of their behavior. In this way a common personal language becomes possible.

In discussing Strawson's thought, it is crucial to emphasize that until a person decides to ascribe pains to others on the basis of their behavior, he has not got and cannot have our concept of pain, for part of that concept is that a pain is something that someone possesses. Nevertheless, he can have a concept (or perhaps something more rudimentary than a full-fledged concept) that is akin to our concept of pain but does not involve the idea of something that is had or possessed by either himself or others.

Perhaps this can be made more intelligible by considering a conceivable though unlikely case, that of a young child who has not yet got our concept of pain but is on the way to getting it. When he falls and knocks his head or scrapes his knee, he says, "It hurts." He has learned this sentence, perhaps as a replacement for natural cries of pain, and he uses it to get picked up and otherwise comforted. However, when his twin brother or a brick falls off the table, and the child is asked, "Does it hurt?" he replies, "No." Nevertheless, he cannot be said to mean by "It hurts" what is meant by "It hurts me," even though he says the former only when the latter is true, for he attaches no sense to "Does it hurt John?," as opposed to "Does it hurt me?" Nor, with regard to what he calls hurting, does he see any difference between John and a brick. If John says, "It hurts," when he himself is feeling all right, he regards what John says as simply untrue. In order for this child to make the transition to the concept of pain as something that either he or someone else has, he must learn to say, "It hurts John," when John bumps his head and cries and to say, "It hurts me," when formerly he said only, "It hurts." Until this linguistic convention is acquired, the child cannot be said to have the concept of pain as a property of persons at all, not even as a property of himself.

Thus, the argument from analogy breaks down because it assumes not only that a person can have a private language but that this language contains our concept of pain (ascribed pain). But such a language could contain at best only a concept of what we may call unascribed pain. The connection between unascribed pain and my behavior is a contingent one, but the connection between behavior and ascribed pain is not. We can see now why Strawson says, "I have argued that such a … 'justification'

[of our beliefs about others] is impossible, that the demand for it cannot be coherently stated" (*Individuals*, p. 112). To talk about other people's pains at all is to accept and use the concept of ascribed pain, and it is an integral part of this concept that behavior shows any person whether that concept applies to other people.

Criticisms of Strawson. Strawson's views are open to some of the criticisms that have been directed against opinions that are the same as his own. In addition, Ayer has directed a number of criticisms specifically against Strawson's positions, asserting that his notion of logical adequacy is obscure and arguing that this obscurity is irremediable. It is certainly true that Strawson does not make the notion of logical adequacy as clear as he might, but Ayer's reasons for thinking that this obscurity could not be remedied are themselves inconclusive. Ayer's other main criticism is directed against Strawson's reason for holding that neither the argument from analogy nor the philosophical skepticism that arises from this argument can be stated coherently. This criticism is based on a failure properly to understand Strawson's position, which in turn leads to the mistaken idea that Strawson cannot allow for the existence of someone with the concept of a person "who was invariably mistaken in ascribing states of consciousness to others" (*The Concept of a Person and Other Essays*, p. 106).

There is nothing in Strawson's position to prevent him from holding that analogy is used in the ascription of states of consciousness to others; the only thing that he rules out is analogical *argument* of the traditional pattern. To understand this, let us use the words "upain" and "utickle" for the concepts of unascribed pains and tickles. According to Strawson, in order to pass from these concepts to those of (ascribed) pains and tickles, I must adopt verbal rules according to which I say "I have a pain" when there is a "upain" and "He has a pain" when another body exhibits certain behavior, and so on. But what sort of behavior, and so forth? There is no reason that Strawson's answer should not be along some such lines as "behavior, etc., that is like the behavior, etc., that this body (i.e., mine) exhibits when there is a upain." In accepting such a rule, I am not *arguing* by analogy. Now, I can adopt such a rule and thus have the concept of a person, but I can still fail to realize that all the objects I regard as persons are in fact unlike myself in ways that I have not noticed.

Ayer describes an imaginary child who is brought up and taught to speak by lifelike robots and who never meets real people. He argues, quite correctly, that this child would have the concept of a person and yet always be mistaken when he ascribes mental attributes to anything. But no consequences fatal to Strawson's views follow from this. The child has adopted the verbal rule whose acceptance, according to Strawson, is necessary for the possession of the concept of a person. The child mistakenly thinks that the robots are persons because he believes that they are much more like himself than in fact they are. This gives no ground for the skeptical conclusion that I may here and now be mistaken in my belief that there are other people besides myself. If one accepts Strawson's position, such skepticism need be justified only if what I think to be other people are a great deal less like me in behavior, etc., than I take them to be. If there is a doubt left here, according to Strawson it can have nothing very specifically to do with other minds. The basis of Ayer's misunderstanding is his mistaken belief that Strawson "infers that any attempt to justify the belief that there are other persons by relying on the premiss that one knows oneself to be a person would be circular; the premiss would already assume what the argument is supposed to prove" (ibid., p. 104). But Strawson's objection to the argument from analogy is not that it is a circular argument. According to him, the trouble is that the argument both uses the concept of a person and rejects the verbal rule that is a necessary part of that concept, namely, the rule that mental attributes are to be ascribed to things on the basis of their behavior, and so on.

JOHN WISDOM'S VIEWS. Finally, something should be said of John Wisdom's very important work on this problem. It is quite impossible to summarize Wisdom's contribution as another solution to the problem of other minds. This impossibility is inherent in his views about philosophy and in the method he used in conformity with these views. All that can be done here is to give some idea of what is to be found in his writings on the problem of other minds by sketching his method of dealing with it.

Wisdom was much influenced by Wittgenstein, especially in regard to the idea that the treatment of a philosophical problem is in some ways like the treatment of an illness. Such a problem or puzzle is a symptom of deep-seated intellectual disorder that consists in a persistent tendency to think about a certain area of thought and language in accordance with a misleading and partially inappropriate model. The puzzle is dissipated when one is "cured" of this tendency. Inattention, however, is not the only remedy, nor is the taking of drugs. The only "cure" available to a philosopher qua philosopher is a certain form of insight. The misleading model that distorts one's thinking is largely an unconscious one. Insight and freedom from its grip are obtained by bringing it into the

open, by making quite clear in detail how our thought is governed by it, and by giving us a proper view of the nature of, for example, our knowledge of other minds.

Thus, Wisdom's first aim is to induce and sharpen philosophical perplexity by showing how it arises precisely out of the sort of position that is at first sight the most attractive to us. For example, the most natural answer to the question about other minds is the traditional one. But it is from this answer and the way of thinking that goes with it that philosophical skepticism most easily arises. Skepticism is satisfactorily removed only when we are brought to see that knowing about other minds is not altogether like other ways of knowing that are by analogy and that it need not be. It might be thought that the aim of a philosopher should be to find a correct model that does not mislead. But according to Wisdom, this is not so. Although every statement has its own logic, the logic of every statement is in some degree like that of every other. We cannot usefully create a limited set of pigeonholes into one of which goes our knowledge of other minds along with, say, our knowledge of the past, while our knowledge of any theoretical entity goes into another. The matter cannot come to this sort of a conclusion. There will be important differences that will make inappropriate any such pigeonhole, as well as the likenesses that make it possible. To get a true grasp of the nature of our knowledge of other minds, it is necessary to make a very large number of detailed comparisons between the various ways in which we know or might know things and between the logic of various types of statements. Only then will we see psychological statements and the ways in which we know of the existence of other people's thoughts and feelings in all their idiosyncrasies and in all their similarities to other statements and to other ways of knowing things. Until this is done, we cannot be entirely freed from our tendency to see things as they are not.

As may be deduced, Wisdom's writings about other minds are almost as much about induction, the past, perception, philosophy of science, and so on as they are about other minds. He used his method with subtlety, inventiveness, and imagination. Many points made by later writers on the problem of other minds are little more than elaborations or oversimplifications of points already made by Wisdom.

See also Private Language Problem.

Bibliography

ENGLISH-SPEAKING PHILOSOPHERS

Ayer, A. J. *The Concept of a Person and Other Essays*. London: Macmillan, 1963. Ch. 4. Criticizes Strawson's views.

Ayer, A. J. *The Problem of Knowledge*. London: Macmillan, 1956. Ch. 5. Defends the argument from analogy.

Bilgrami, Akeel. "Dummett, Realism, and Other Minds." In *The Philosophy of Michael Dummett*, edited by Brian McGuinness. Dordrecht: Kluwer, 1994.

Brewer, Bill. "Emotion and Other Minds." In *Understanding Emotions: Mind and Morals*, edited by Peter Goldie. Brookfield, U.K.: Ashgate, 2002.

Buford, T. O., ed. *Essays on Other Minds*. Chicago: University of Illinois Press, 1970.

Dennett, Daniel C. "Beliefs about Beliefs." *Behavioral and Brain Sciences* 1 (1978): 568–570.

Dretske, Fred. "Animal Minds." *Philosophic Exchange* 31 (2000–2001): 21–33.

Fodor, Jerry. "A Theory of the Child's Theory of Mind." *Cognition* 44 (1992): 283–296.

Foley, Richard. *Intellectual Trust in Oneself and Others*. New York: Cambridge University Press, 2001.

Gasking, Douglas. "Avowals." In *Analytical Philosophy*, edited by R. J. Butler, 154–169. Oxford, 1962.

Ginet, Carl. "Plantinga and the Philosophy of Mind." In *Alvin Plantinga*, edited by James E. Tomberlin. Dordrecht: Reidel, 1985.

Goldman, A. "In Defense of the Simulation Theory." *Mind and Language* 7 (1992): 104–119.

Heal, Jane. "Other Minds, Rationality, and Analogy." *Proceedings of the Aristotelian Society*, supplement 74 (2000): 1–19.

Heal, Jane. "Understanding Other Minds from the Inside." In *Current Issues in Philosophy of Mind*, edited by Anthony O'Hear. New York: Cambridge University Press, 1998.

Hill, Christopher. "On Getting to Know Others." *Philosophical Topics* 13 (1985): 257–266.

Leslie, A. M. "Pretense and Representation: The Origins of 'Theory of Mind.'" *Psychological Review* 94 (1987): 412–426.

Malcolm, Norman. "Knowledge of Other Minds." *Journal of Philosophy* 55 (1958): 969–978. A radical criticism of the argument from analogy and the traditional viewpoint.

Malcolm, Norman. "Wittgenstein's Philosophical Investigations." *Philosophical Review* 63 (4) (1954): 530–559. Defends Wittgenstein's views on private languages.

McDowell, John. "Criteria, Defeasibility, and Knowledge." *Proceedings of the British Academy* 68 (1982): 455–479.

McGinn, Colin. "What Is the Problem of Other Minds?" *Proceedings of the Aristotelian Society*, supplement 58 (1984): 119–137.

Mill, J. S. *An Examination of Sir William Hamilton's Philosophy*. London: Longmans Green, 1865. Ch. 12. Contains a straightforward version of the argument from analogy.

Nichols, Shaun. "Mindreading and the Cognitive Architecture underlying Altruistic Motivation." *Mind and Language* 16 (2001): 425–455.

Sober, Elliot. "Evolution and the Problem of Other Minds." *Journal of Philosophy* 97 (2000): 365–386.

Strawson, P. F. "Critical Notice of Wittgenstein's Philosophical Investigations." *Mind* 63 (249) (1954): 70–99. A criticism of Wittgenstein's views on private languages.

Strawson, P. F. *Individuals*. London: Methuen, 1959. Ch. 3.

Wisdom, John. "Other Minds." *Mind* 49 (196): 369–402; 50 (197): 1–22; 50 (198): 97–122; 50 (199): 209–242; 50 (200) 313–329; 51 (201): 1–18 (1940–1942).

Wisdom, John. "Other Minds." *Logic and Reality, PAS*, Supp. 20 (1946): 122–147.

Wisdom, John. *Other Minds*. Oxford: Blackwell, 1965. Reprints all the above.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Blackwell, 1953.

GERMAN PHILOSOPHERS

Husserl, Edmund. *Cartesianische Meditationen. Husserliana*, Vol. I. The Hague, 1950. Meditation V. Translated by Dorion Cairns as *Cartesian Meditations; An Introduction to Phenomenology*. The Hague, 1960.

Scheler, Max Ferdinand. *Zur Phänomenologie und Theorie der Sympathiegefühle und von Liebe und Hass*. Halle, 1913. Appendix.

J. M. Shorter (1967)
Bibliography updated by Benjamin Fiedor (2005)

OTTO, RUDOLF

(1869–1937)

Rudolf Otto, the German theologian, was born at Peine in Hanover. He studied at Erlangen and Göttingen, where he became a *Privatdozent* in systematic theology in 1897. In 1904 Otto was appointed professor of systematic theology at Göttingen. He accepted similar posts at Breslau in 1914 and at Marburg in 1917, where he remained until his death. In addition to his philosophical work, Otto published works on Christ, on Indian religious thought and its relation to Christianity, and on various theological topics.

RELIGIOUS FEELING AND RELIGIOUS KNOWLEDGE

Otto's most significant philosophic contribution is to be found in his discussion of religious feeling and religious knowledge—a discussion that begins with his earliest work and culminates in *The Idea of the Holy.*

In *Naturalism and Religion* (1904) Otto discusses the relation of religion to a naturalism that demands that everything be explained on the basis of mathematical-mechanical laws, thus excluding the beyond, purpose, and mystery, which are essential to religion.

COGNITIVE CLAIMS OF RELIGION. Religion makes certain claims—that the world is conditioned and dependent, that there is a providence, that there is a side other than that which appears to us. These claims are not put forward as poetry but as truths. They cannot, however, be justified by, nor derived from, a consideration of nature in any straightforward sense. Reason may show that science does not conflict with these claims and even that science is unable to consider their truth-value. Reason may also point out hints in nature that suggest that these claims are true; reason cannot, however, justify them. These truths differ in kind from those of science and common sense and have their own grounds—the heart and conscience, feeling and intuition. Correlations can be made between various feelings, on the one hand, and religious claims, on the other. Corresponding to the claim that the world is conditioned and dependent is the feeling of the dependence and conditionally of all things. The claim that there is a providence, or teleological order, in things implies that certain value judgments are true and these value judgments rest on feeling and intuition. Corresponding to the claim that there is a beyond is piety—a feeling and intuition, which is bound up with our experience of the beautiful and the mysterious, that there is a reality behind appearances.

RELIGIOUS FEELINGS AND INTUITIONS

In *Naturalism and Religion* it is not entirely clear just what these feelings and intuitions are. Otto sometimes talks of them as if they were feelings in a straightforward sense. At other times he talks of them as if they were half-formulated judgments that carry with them an inescapable sense of conviction, and at still other times he talks of them as if they were cognitive experiences in somewhat the same way that visual experiences are cognitive.

CATEGORIES AND IDEAS. The notion of religious feelings and intuitions receives a more complete treatment in *The Philosophy of Religion Based on Kant and Fries* (1909), in which Otto follows the position of Jakob Friedrich Fries. We have an immediate knowledge of reality, the noumenal world, which shows itself in "feelings of truth." These feelings can be brought to full consciousness as ideas. An idea is a concept that can be applied to reality. When temporally schematized, the categories of theoretical reason can be applied to appearances and can also, when schematized by the principle of completeness (a principle based on reason's "perception and knowledge"

that real existence is necessary, one, and complete), be applied to reality itself. A category thus schematized is an idea. These ideas are essentially negative. In effect, they exclude certain characteristics—temporality, contingency, and so on—from reality.

In the case of the practical reason the "feeling of truth" cannot be completely conceptualized. Practical reason does, however, derive the idea of reality as "the reign of purpose" from the principle of the dignity of the person that underlies the concept of duty. The idea is again presumably negative.

The negative judgments obtained through applying the ideas of theoretical and practical reason to reality must be supplemented by positive knowledge, which is gained through feelings or perceptions that cannot be adequately expressed although they can be communicated. These feelings, or perceptions, again seem to be, simultaneously, feelings in an ordinary sense, the ability to make judgments according to criteria that cannot themselves be formulated, and a direct perception of an objective existence—in this case, reality. Otto distinguishes between the feeling of beauty and of the sublime, on the one hand, and religious feelings, on the other. Although the discussion is somewhat obscure, it would seem that all three of these feelings either directly or indirectly disclose reality.

NUMINOUS FEELINGS

In *The Idea of the Holy* (1917), Otto attempts to make a clear distinction between numinous, or religious, feelings and feelings that might be confused with them, such as the feeling of the sublime. Numinous feelings have two primary aspects—a feeling of religious dread and a feeling of religious fascination. The closest analogue to religious dread, or awe, is the feeling of uncanniness—the feeling one has when the hair on the back of one's neck rises, the shudder or terror on hearing a ghost story, the dread of haunted places. The feeling of fascination by, attraction to, and prizing of the object that arouses the feeling in question creates both the desire to approach the object and the feeling that one possesses no value when considered in relation to the fascinating and prized object.

Otto's attempt to describe the various feelings must be distinguished from his theory about numinous feelings. Numinous feelings are unique; they cannot be analyzed as a complex of such nonnuminous feelings as love, fear, horror, a feeling of sublimity, and so on. Second, the capacity for numinous feelings is unexplainable; although the capacity may appear in the world only when certain conditions are fulfilled, the conditions do not constitute an adequate explanation of the capacity in question.

Numinous feelings are also cognitive. Two claims are made at this point. First, the feelings are the source of the concept of the numinous—the concept of something that is both a value and an objective reality. The numinous feelings are also cognitive in the sense that they are like visual experiences. They have "immediate and primary reference to an object outside the self"—the numinous quality or object, which is an object of numinous feelings in somewhat the same way that visible objects and qualities might be said to be the object of visual experiences.

INTERPRETATIONS. The relation between these two claims is not clear. At least two interpretations are possible. The first interpretation makes central the claim that numinous feelings disclose the numinous object. The encounter with the numinous object through numinous experiences gives rise to the concept of the numinous in much the same way that encounters with objects and qualities through visual experiences are thought to give rise to the concepts of those objects and qualities. The concept of the numinous is, then, a posteriori in the sense that it is derived from the experience of an object or quality. It is, however, a priori in the sense that it is not derived from any sense experience. In this interpretation the feeling is the source of the concept only in the sense that it discloses the object of the concept, the encounter with the object producing the concept of the object.

In the second interpretation the feeling gives rise to both the concept and the disclosure of the numinous object, yet it is not the encounter with the numinous that gives rise to the concept of the numinous. Rather, the feeling furnishes the concept in much the same way that Immanuel Kant's theoretical reason furnishes the various a priori categories. The concept of the numinous is, then, a priori in a standard sense. The feeling does more than this, however. The feeling that furnishes the concept also discloses the object to which the concept applies. How are these two functions of numinous feelings related? Neither the concept nor the object is, it would seem, given in isolation. Rather, the object is given through the concept or as structured by the concept. The two are given together although one is not derived from the other. In either interpretation Otto makes the claim that feeling puts us in contact with, discloses, is an awareness of, intuits something outside ourselves. In this respect feeling is like visual and auditory experiences. It has an objective referent whether this is structured by an a priori concept or

whether it simply gives rise to a concept. Unfortunately, the difficulties involved in this claim are not discussed. Obvious disanalogies with ordinary perception (the absence of tests for "mis-seeing," the fact that no sense organ is tied to numinous experiences, the fact that nonpsychological predictions cannot be based on numinous experiences in the way in which they can be based on visual experiences, and so on) are ignored.

THE NUMEN

Otto calls the object of numinous feelings the numen, something that is both value and object but which can be only indirectly characterized by means of "ideograms"—that is, by designating properties which would appropriately call forth a feeling response analogous to that evoked in the encounter with the numen. For example, the encounter with the numen evokes religious dread. This is analogous to fear. Accordingly, we indicate the property of the numen that arouses religious dread by *wrath,* a term that refers to a property which often produces fear. In addition to this, however, we can and should "schematize" the numen by means of such rational concepts as goodness, completeness, necessity, and substantiality. That is, concepts of this sort may be predicated of the numen. The resulting judgment is synthetic a priori. It may be suggested that the cash value of the last claim is that we just "see" the connection to be appropriate if we possess numinous feelings.

THE HOLY

When the concept of the numinous and the schematizing concepts are brought together in this way, we have the "complex category of the 'holy' itself." The category is a priori in the sense that (1) the connection between the notion of the numinous and the schematizing concepts is a priori, (2) the concept of the numinous is a priori in that although it arises "amid the sensory data … of the natural world, … it does not arise out of them," and (3) the schematizing concepts are a priori.

The last claim is difficult to maintain, however, for Otto's examples of the schematizing concepts seem to make this impossible. It could perhaps be argued that schematizing concepts such as completeness, necessity, substantiality, and goodness are a priori. Otto also wishes to say, however, that the concepts of love, mercy, and moral will can function as concepts that schematize various aspects of the numinous. It is difficult to maintain that a concept such as love is a priori. What Otto maintains is that although "love" as applied to the numen and "love" as applied in ordinary situations have the same content, their form differs. When referred to the numen, the term is taken absolutely; when it is applied in ordinary situations, it is not. Otto seems to mean that love in the ordinary sense admits of degrees that can be arranged on a scale. The love of the numen is the limit of this scale. Since the limit (whatever this might be) is not given to us in sense experience, we may call it a priori.

RELIGIOUS FEELINGS AND THE NUMEN

We can now explicate more fully the role that religious or numinous feelings play in religious knowledge. They disclose the numen to us. They are the source of the concept of the numinous. Finally, they appear to warrant the synthetic a priori judgments that link the schematizing concepts to the concept of the numinous.

The relation between the account presented in *The Philosophy of Religion* and *The Idea of the Holy* is, I think, clear. The ideas have become the "Idea of the Holy" (which breaks down into the concept of the numinous and the schematizing concepts), reality has become the numen, and feelings and intuitions have become numinous feelings.

AUTONOMY OF THE SPIRIT

Another theme, although less philosophically interesting, is of central concern to Otto himself—the autonomy of the spirit and of the spirit's religious capacities. In asserting that the spirit is autonomous, Otto is claiming that the laws of the spirit are fundamentally different from those of the natural world. In effect, they are the prescriptive laws of logic and ethics (and of religion?) rather than the descriptive laws of physics and psychology. Insofar as a spirit determines itself by prescriptive laws, it is free. Otto is further claiming that spirit is the source of concepts, principles, intuitions, and valuations that cannot be derived from sense experience. And, finally, he is claiming that although spirit develops under the influence of external stimuli, it is something unique in its own right. Spirit cannot be explained by, nor can its occurrence be predicted on, the basis of a consideration of sense experience alone. Spirit and its operations "emerge" under certain conditions but are not explained by these conditions.

See also A Priori and A Posteriori; Fries, Jakob Friedrich; Kant, Immanuel; Mysticism, Nature and Assessment of; Naturalism.

Bibliography

WORKS BY OTTO

Naturalistische und religiöse Weltansicht. Tübingen: Mohr, 1904. Translated by J. A. Thomson and M. R. Thomson as *Naturalism and Religion.* New York: Putnam, 1907.

Goethe und Darwin, Darwinismus und Religion. Göttingen: Vandenhoeck and Ruprecht, 1909. *Darwinismus und Religion* translated by S. G. Cole and E. M. Austin as "Darwinism and Religion" in *Crozer Quarterly* 8 (1931): 147–161.

Kantisch-Fries'sche Religionsphilosophie und ihre Anwendung auf die Theologie. Tübingen: Mohr, 1909. Translated by E. B. Dicker as *The Philosophy of Religion Based on Kant and Fries.* London: Williams and Norgate, 1931.

Das Heilige; über das Irrationale in der Idee des Gottlichen und sein Verhältnis zum Rationalen. Breslau: Trewendt and Granier, 1917; 25th ed. Munich, 1936. The later editions contain additional material. Translated by J. W. Harvey as *The Idea of the Holy: An Inquiry into the Nonrational Factor in the Idea of the Divine and Its Relation to the Rational.* New York, 1958.

Aufsätze das Numinose betreffend. Gotha: Klotz, 1923.

West-Östliche Mystik. Vergleich und Unterscheidung zur Wesensdeutung. Gotha: Klotz, 1926. Translated by B. L. Bracey and R. C. Payne as *Mysticism, East and West. A Comparative Analysis of the Nature of Mysticism.* New York: Macmillan, 1932.

Das Gefühl des Überweltlichen (Sensus Numinis). Munich, 1932. The first part of the fifth and sixth editions of the *Aufsätze* with some added material.

Sunde und Urshuld und andere Aufsätze zur Theologie. Munich, 1932. The second part of the fifth and sixth editions of the *Aufsätze* with some added material.

Religious Essays. A Supplement to the "Idea of the Holy," by Rudolf Otto. Translated by B. Lunn. London: Oxford University Press, 1931. This consists primarily of translations of essays found in the *Aufsätze* and the two preceding works.

Freiheit und Notwendigkeit, Ein Gespräch mit Nicolai Hartmann über Autonomie und Theonomie der Werte. Tübingen: Mohr, 1940.

WORKS ON OTTO

Almond, Philip. *Rudolf Otto: An Introduction to his Philosophical Theology.* Chapel Hill: University of North Carolina Press, 1984.

Ballard, Steven. *Rudolf Otto and the Synthesis of the Rational and the Non-Rational in the Idea of the Holy.* New York: Peter Lang, 2000.

Davison, R. F. *Rudolf Otto's Interpretation of Religion.* Princeton, NJ: Princeton University Press, 1947.

Feigel, F. C. *"Das Heilige." Eine kritische Abhandlung über Rudolf Otto's gleichnamiges Buch.* Haarlem, Netherlands, 1929.

Gooch, Todd. *The Numinous and Modernity: An Interpretation of Rudolf Otto's Philosophy of Religion.* New York: de Gruyter, 2000.

Haubold, W. *Die Bedeutung der Religionsgeschichte für die Theologie Rudolf Ottos.* Leipzig, 1940.

Moore, J. M. *Theories of Religious Experience with Special Reference to James, Otto and Bergson.* New York: Round Table Press, 1938.

Poland, Lynn. "The Idea of the Holy and the History of the Sublime." *Journal of Religion* 72 (1992): 175–197.

Proudfoot, Wayne. *Religious Experience.* Berkeley: University of California Press, 1985.

Raphael, Melissa. *Rudolf Otto and the Concept of Holiness.* Oxford: Clarendon Press, 1997.

Schlamm, Leon. "Numinous Experience and Religious Language." *Religious Studies* 28 (1992): 533–551.

Schlamm, Leon. "Rudolf Otto and Mystical Experience." *Religious Studies* 27 (1991): 389–398.

Siegfried, T. *Grundfragen der Theologie bei Rudolf Otto.* Gotha, 1931.

Sommer, J. W. E. *Der heilige Gott und der Gott der Gnade bei Rudolf Otto.* Frankfurt, 1950.

William J. Wainwright (1967)

Bibliography updated by Christian B. Miller (2005)

OUSIA

In classical Greek philosophy, *ousia* (a noun derived from the present participle of the Greek verb "to be") most often expresses one or another of four closely connected concepts: (1) what something is in itself, its being or essence; (2) an entity which is what it is, at least with respect to essential attributes, on its own and without dependence on any more fundamental entity of another type outside itself (in Plato's middle dialogues, the forms; for Aristotle, substance; for the Stoics, the material substrate); (3) for Plato, being as opposed to becoming; and (4) for the Stoics in some instances, existence as opposed to nonexistence. Depending on the context, *ousia* may be translated as "being," "essence," "reality," or "substance."

Employed in ordinary Greek to speak of a person's wealth and possessions, the word *ousia* was put to philosophical use by Plato in his early dialogue *Euthyphro* to state a requirement on definitions. Asked what piety is, Euthyphro answers that it is what is loved by all the gods. Socrates responds with a clear statement of concept (1), saying that Euthyphro has mentioned merely something that qualifies piety externally and has failed to give the *ousia* of piety (11a4–b1), what it is in itself that leads the gods to love it.

The transition from concept (1) to concept (2) occurs most clearly in the *Phaedo*, a dialogue of Plato's middle period. There the character Socrates introduces several forms, including the just itself and the beautiful itself (65d4–8), and speaks of them as the *ousia* of other things (65d13), in the sense that other things become just

or beautiful, for example, only by participation in, or dependence on, the corresponding form (101c3–4). Each such form is an *ousia* according to concept (2) (76d9, cf. 77a2), a being or reality (78d1) that is always the same and unchanging (78d1–7), an object of thought rather than sensation.

In the *Republic* a similar picture obtains, but there the character Socrates speaks of the forms collectively as *ousia*, with the exception of the form of the good (VI, 509b8–9), and contrasts this invariant, unqualified, and cognitively reliable being first with the many sensible things, which can appear, for example, beautiful in one respect but ugly in another (V, 479c7, cf. 479b6–d1), and then with the collective becoming and decaying of these sensibles (VI, 485b21). This use of the word *ousia* to express concept (3), being as opposed to becoming, is frequent in book VII, where the study of the mathematical sciences serves to lead the prospective philosopher-rulers to turn away from becoming and toward being (VII, 525b5, cf. 525c6, 526e6, 534a3). This strong distinction in the *Republic* between being and becoming has been questioned by some scholars. In any case, it is considerably attenuated in some of Plato's later dialogues, including the *Philebus*, where the character Socrates asserts "Every process of generation … takes place for the sake of some particular being [*ousias tinas hekastēs*]" (54c2–3).

In the *Categories*, Aristotle uses the word *ousia* occasionally in the concept (1) sense of essence (e.g., at 1a1–2), but at the center of the discussion in the *Categories* is concept (2), and *ousia* in this sense becomes a technical term rendered by most translators as "substance." Moving even further from the view of the *Republic* than Plato does in his later dialogues, Aristotle argues that *ousia* in sense (2) belongs primarily and most of all to sensible entities like a particular human or a particular horse (2a11–14), since these "primary substances" (2a35) are substrates, or ontological subjects, not only of their own essential attributes but also, differently, of inherents from other categories, such as a certain quality or a certain quantity, that happen to be "in" them at one time or another (2a34–b5). He concludes that everything else under discussion in the *Categories*, including the species and genera of primary substances (called "secondary substances" at 2a14) as well as all the inherents in other categories, depend on primary substances for their being, in the sense that without primary substances, none of the others could be (2b5–6). (For an even stronger claim that all depends on substance, the focal or referential theory of the meaning of "being" [Gk. "to on," the participle], see *Metaphysics*, IV, 1003b5–10; cf. Devereux, pp. 220, 232.)

Aristotle's other extended discussion of *ousia* (*Metaphysics*, VII, VIII) accepts the view of the *Categories* that particular animals and plants fall under *ousiai* in sense (2) (VII, 1028b8–10). But book VII, having brought in the distinction between matter and form introduced in the *Physics* (190b1–191a22) to explain the coming-to-be and passing-away of particular sensible substances, subsequently regarded as composites of matter and form, says that such composite sensible substances are "posterior" to both matter and form (1029a30–32). It then argues at length for the thesis that form is primary substance (1037a5–7 and 1037a27–30, cf. 1032b1–2). This thesis raises two important questions. How does the thesis fit with Aristotle's position in the *Categories* that entities like particular horses and particular humans are primary substances? And is the primary substance the form of the species, which, though not a universal (1038b1–16), is nevertheless present in all the particular members of that species, or is it the particular form of a particular member of the species, unique to it and not present in any other member? These issues have been much debated since the 1950s, but in the 1980s and 1990s the weight of scholarly opinion shifted somewhat toward the particular-forms view, even as the widespread assumption that *Metaphysics* VII–VIII is a later work than the *Categories* came into question. (On these issues, see both Frede and Wedin; for a different view, see Loux.) The thesis that form is primary substance opens up the possibility of an inquiry, promised in book VII (1028b27–33), as to whether there can be any substance entirely separate from matter. This inquiry, carried out in book XII, leads Aristotle to conclude that there are not only eternal material substances (e.g., the planets, on his view) but also eternal immaterial substances (1071b4–5), including Aristotle's god, the first unmoved mover whose ceaseless thinking upon thinking (1072b1–30) inspires the movement of the outer sphere of fixed stars (1073a23–30).

Among the Stoics, by contrast, *ousia* in sense (2) is the single material substrate of all things, considered in abstraction from all qualities and relations depending on it (Calcidius, see Long and Hedley, Vol. 1, p. 269–270; for the Stoics' debt here to Plato, *Timaeus* 50a5–c6, see Menn, p. 216). Some Stoics also use the word *ousia* in sense (4), existence as opposed to nonexistence, to distinguish objects of thought that exist, objects that are peculiarly qualified portions of the material substrate *ousia*, for example, a particular horse, from objects of thought that are purely fictional and do not exist, for example, a centaur (Seneca, see Long and Hedley, Vol. 1, p.162).

See also Aristotle; Essence and Existence; Plato; Stoicism; Substance and Attribute.

Bibliography

Aristotle. *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Devereux, Daniel. "The Primacy of *Ousia*: Aristotle's Debt to Plato." In *Platonic Investigations*, edited by Dominic J. O'Meara. Washington, DC: Catholic University Press, 1984.

Frede, Michael. "Individuals in Aristotle." In his *Essays in Ancient Philosophy*. Minneapolis: University of Minnesota Press, 1987.

Long, Anthony A., and David N. Sedley. *The Hellenistic Philosophers*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1987. This book includes both Calcidius's *Plato's Timaeus: Translation And Commentary*, sections 292–293, p. 269–270; and Seneca's *Letters*, 58.13–15.

Loux, Michael J. *Primary Ousia: An Essay on Aristotle's "Metaphysics" Z and H*. Ithaca, NY: Cornell University Press, 1991.

Menn, Stephen. "The Stoic Theory of Categories." In *Oxford Studies in Ancient Philosophy*. Vol. 17, edited by David Sedley. Oxford, U.K.: Oxford University Press, 1999

Plato *Plato: Complete Works*, edited by John Cooper. Indianapolis, IN: Hackett, 1997.

Wedin, Michael. *Aristotle's Theory of Substance: The "Categories" and "Metaphysics" Zeta*. Oxford, U.K.: Oxford University Press, 2000.

John Driscoll (2005)

OWEN, G. E. L

1922–1982

Gwilym Ellis Lane Owen was a major force in the post–World War II upsurge of analytically oriented philosophical work on ancient philosophy. The author of articles of enduring value, the subject of much discussion and controversy, many of them among the classics of the philosophical study of pre-Socratic philosophy, Plato, and Aristotle, he was concerned principally with the logic of argument, metaphysics, and philosophy of language; he had no substantive interests in ethics, political theory, or aesthetics. He understood the ancient philosophers as engaged in conceptual investigations of live philosophical interest. Raised in a Welsh family in Portsmouth, England, he matriculated at Corpus Christi College, Oxford, in 1940, completing his bachelor of arts degree in 1948, after war service in the Pacific arena. In 1950 he received a bachelor of philosophy degree under Gilbert Ryle's supervision, with an epoch-making thesis on logic, philosophy of language, and metaphysics in Plato's *Theaetetus*, *Parmenides*, *Sophist*, *Statesman*, and *Philebus*. Its main ideas formed the basis of his influential, though controversial, first publication, "The Place of the *Timaeus* in Plato's Dialogues."

After postdoctoral research at the University of Durham, Owen returned to Oxford in 1953 as university lecturer in ancient philosophy (from 1958, also nontutorial fellow of Corpus Christi), university reader (1957), and professor of ancient philosophy as first incumbent of that chair (1963). In 1966 he went to Harvard as professor of philosophy and the classics to direct a new PhD program in classical philosophy. In 1973 he returned to Great Britain as Laurence Professor of Ancient Philosophy in the Classics Faculty at Cambridge, and as fellow of King's College, where he remained until his early death in 1982.

Owen's year-long Oxford lectures on pre-Socratic philosophy, and his courses and seminars on Plato and Aristotle throughout his career, were famously exhilarating, challenging, and fast-paced explorations of central texts and topics in the study of ancient philosophy. A remarkably high percentage of the leading ancient philosophers of the next generation learned their craft and drew their initial inspiration from these classes. More than any of his contemporaries, Owen's example and personal influence shaped the growth and expansion in the philosophical study of ancient philosophy in the late twentieth century.

More than half of Owen's published work concerned Aristotle primarily, but his work on Plato and the pre-Socratic philosophers Parmenides and Zeno of Elea was equally ground breaking. He rejected the traditional idea that Plato's *Timaeus*—with its conception of the physical world as a "copy" drawn by a creator god from his intellectual vision of Forms existing in a separate nonphysical realm—was the culmination and permanent legacy of Plato's work in metaphysics. Rather, he read the dialectical and logical investigations of the *Parmenides* and *Sophist*, and others of what under his influence came to be referred to simply as the "late" dialogues, as containing deeper and more adequate reflections on issues of being and not-being, unity and multiplicity, becoming and change.

Confused ideas about these issues had motivated the "middle-period" theory of Forms, of *Symposium*, *Phaedo*, *Republic*, and *Timaeus*. Owen argued that *Timaeus* was in fact composed, not, as traditionally assumed, toward the end of Plato's life, but rather as a premature copestone to the middle-period theory, which was to be undermined and reconsidered in the "late" dialogues. His influential essays, "Notes on Ryle's Plato" and "Plato on Not-Being," dealing respectively with *Parmenides* and *Sophist*, cast new light on these intriguing but very obscure works, and

spearheaded a generation of subsequent scholarly and philosophical work on them. His essays "Eleatic Questions," "Zeno and the Mathematicians," and "Plato and Parmenides on the Timeless Present" had a similar effect on studies of Parmenides and Zeno.

Owen's work on Aristotle concentrated on logic, methodology, physics, and metaphysics, but included one provocative paper on "Aristotelian Pleasures." This investigates Aristotle's idea that pleasure is to be conceived not as a passive experience but is itself an activity. Owen advanced the challenging thesis that Aristotle's two discussions of pleasure in *Nicomachean Ethics* VII and X have interestingly divergent conceptions of the relationship between the activity that pleasure itself is and whatever one takes pleasure in. In "Logic and Metaphysics in some Earlier Works of Aristotle" he paid careful attention to logical and philosophical details in some of Aristotle's earliest works and showed that the then popular picture of Aristotle's development (due to Werner Jaeger) was unacceptable. Far from only gradually freeing himself from a committed belief in a universal science of being, gained through the knowledge of middle-period Platonic Forms, Aristotle began by rejecting both the existence of such Forms and the possibility of any universal science of being.

It was only much later, with the employment of what Owen called a theory of "focal meaning" for *being*, that Aristotle could reconcile himself to any general science of being, or metaphysics. It was, however, the being of Aristotelian substances, not Platonic Forms, which provided the linchpin and focus of that science. In "The Platonism of Aristotle" and "Particular and General," he carried this analysis forward, finding in the middle books of Aristotle's *Metaphysics* an avowed sympathy with Plato's general metaphysical program—with Aristotelian forms, not Platonic Forms, at the center of the enterprise. Other well-known papers proposed an influential analysis of the "appearances" that Aristotle notoriously made the basis for the use of dialectical inquiry in physics, ethics, and other areas of philosophy ("*Tithenai ta phainomena*"), and argued that in his theory of categories Aristotle countenanced nonrepeatable individuals only in the category of substance. In other categories the "individuals" were such things as specific, narrowest *shades* of colors, not color-instances possessed uniquely by individual substances ("Inherence"). His paper "Aristotle on Time" also generated much discussion.

Owen was a moving force for the founding in 1957 of the Symposium Aristotelicum, a triennial select meeting of British, European, and North American scholars for concentrated joint study of a single Aristotelian text or topic. These meetings have done much to bring the diverse national traditions of Aristotelian scholarship into mutual communication. Several of Owen's articles originally appeared in the Symposium's triennial volumes. Many of his papers were reprinted in collections too numerous to list. After his death, they were all published together in 1986 (as *Collected Papers*); details of the original and other prior publications can be found there.

See also Aristotle; Parmenides of Elea; Philosophy of Language; Plato; Pre-Socratic Philosophy; Ryle, Gilbert; Zeno of Elea.

Bibliography

WORKS BY G. E. L. OWEN

"The Place of the *Timaeus* in Plato's Dialogues." *Classical Quarterly* NS 3 (1953): 79–95.

"Zeno and the Mathematicians." *Proceedings of the Aristotelian Society* 58 (1957–1958): 199–222.

"Eleatic Questions." *Classical Quarterly* NS 10 (1960): 84–102.

"Logic and Metaphysics in some Earlier Works of Aristotle." In *Aristotle and Plato in the Mid-Fourth Century*, edited by Ingemar Düring and G. E. L. Owen, 163–190. Göteborg, Sweden: Elanders Boktryckeri, 1960.

"*Tithenai ta phainomena*." In *Aristote et les problèmes de méthode*, edited by S. Mansion, 83–103. Louvain, Belgium: Publications Universitaires de Louvain, 1961.

"Inherence." *Phronesis* 10 (1965): 97–105.

"Plato and Parmenides on the Timeless Present." *The Monist* 50 (1966): 317–340.

"The Platonism of Aristotle." *Proceedings of the British Academy* 51 (1966): 125–150.

"Notes on Ryle's Plato." In *Ryle*, edited by O. P. Wood and George Pitcher, 341–372. Garden City, NY: Doubleday, 1970.

"Plato on Not-Being." In *Plato I*, edited by G. Vlastos, 223–267. Garden City, NY: Doubleday, 1970.

"Aristotelian Pleasures." *Proceedings of the Aristotelian Society* 72 (1971–1972): 135–152.

"Particular and General." *Proceedings of the Aristotelian Society* 79 (1978–1979): 1–21.

Logic, Science, and Dialectic: Collected Papers in Greek Philosophy. Edited by Martha C. Nussbaum. London: Duckworth, 1986.

John M. Cooper (2005)

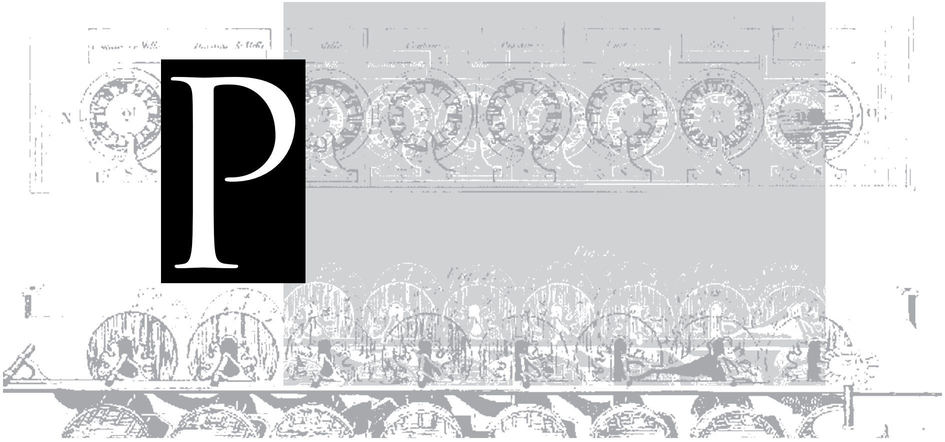

P

PACIFISM

"Pacifism" is moral opposition to war. The concept embraces a wide range of positions from an absolute prohibition of all use of force against persons to a selective and pragmatic rejection of particular forms of such force under varying circumstances. Pacifists vary on their moral grounds for rejecting war and on their commitments to varieties of nonviolence.

Etymologically, pacifism comes from the Latin *pax, pacis,* "peace" (originally "compact") + *facere,* "to make," and literally means "peacemaking." Often, pacifism is incorrectly identified as passivism, which derives from the Latin *passivus,* "suffering," and means being inert or inactive, suffering acceptance. Pacifists may be passivists but often are activists, choosing nonviolent means to resolve conflict and achieve personal and social goals.

Pacifism consists of two parts: the moral opposition to war and the commitment to cooperative social and national conduct based on agreement. Beyond the mere absence of war, peace is a condition of group order arising from within by cooperation among participants rather than order imposed from outside by domination by others. Pacifism's opposition to war is much more frequently reflected in philosophic literature than is its active creation of peace.

Moral opposition to war is discussed across the history of Western philosophy. While early considerations of the morality of war can be found in ancient Greek texts (e.g., Plato, *Republic,* Book IV, 469c–471c), more thorough treatments are much later—notably from Desiderius Erasmus in the sixteenth century and Immanuel Kant in the late eighteenth. Adin Ballou articulated pragmatic pacifism in the mid-nineteenth century, and William James explored pacifist philosophy in the early twentieth. Arguments for pacifism tend to focus on the evils of war, including human suffering—especially of innocents—and moral degradation of participants as well as the uncontrollability of modern warfare.

The case for pacifism varies with the form of pacifism being put forth. Absolute pacifism, the view that it is wrong under all circumstances to use force against persons, may rest on one interpretation of Kant's categorical imperative, on Mohandas Gandhi's Satyagraha (truth force), on Martin Luther King Jr.'s notion of Christian love, or on other moral bases. Weaker forms of pacifism may rest on interpretations of these same principles or on other grounds. Epistemological pacifists stress the impossibility of knowing sufficiently to warrant taking lives,

while pragmatic pacifists trace the empirical history of war to emphasize failures in achieving the ends that were to justify carnage. Nuclear pacifists focus on the projected effects of thermonuclear exchange, and ecological pacifists consider the effects of modern war on ecosystems.

See also Erasmus, Desiderius; James, William; Just War Theory; Kant, Immanuel; King, Martin Luther; Love; Peace, War, and Philosophy; Plato; Russell, Bertrand Arthur William; Social and Political Philosophy; Violence.

Bibliography

Ballou, A. "Christian Non-Resistance." In *Nonviolence in America: A Documentary History,* edited by S. Lynd. Indianapolis: Bobbs-Merrill, 1966.

Cady, D. L. *From Warism to Pacifism: A Moral Continuum.* Philadelphia: Temple University Press, 1989.

Cromartie, Michael, ed. *Peace Betrayed?: Essays on Pacifism and Politics.* Washington, DC: Ethics and Public Policy Center, 1990.

Erasmus, D. *Complaint of Peace.* 1517.

Erasmus, D. *Praise of Folly.* 1512.

Gandhi, M. K. *Non-violent Resistance.* Edited by B. Kumarappa. New York: Schocken, 1951.

Holmes, Robert L. *On War and Morality.* Princeton, NJ: Princeton University Press, 1989.

Holmes, Robert L. "Pacifism for Nonpacifists." In *Social and Political Philosophy,* edited by James P. Sterba. New York: Routledge, 2001.

James, W. "The Moral Equivalent of War" (1910). In *War and Morality,* edited by R. Wasserstrom. Belmont, CA: Wadsworth, 1970.

Kant, I. *Perpetual Peace* (1795). Edited and translated by L. W. Beck. New York: Liberal Arts Press, 1957.

King, M. L., Jr. *A Testament of Hope: The Essential Writings of Martin Luther King Jr.* Edited by J. M. Washington. San Francisco: Harper and Row, 1986.

Ruddick, S. *Maternal Thinking: Toward a Feminist Peace Politics.* Boston: Beacon Press, 1989.

Sharp, Gene. *Power and Struggle,* part 1 of *The Politics of Nonviolent Action.* Boston: P. Sargent, 1973.

Teichman, J. *Pacifism and the Just War.* Oxford: Blackwell, 1986.

Tolstoy, L. *The Kingdom of God and Peace Essays* (1909), 2nd ed. Oxford, 1951.

Duane L. Cady (1996)
Bibliography updated by Philip Reed (2005)

PAIN

There is no consistent philosophical view concerning the nature of pain, how to understand it, or what an understanding of pain might mean for philosophy of mind. Just about every conceivable position concerning the nature of pain is held by some leading thinker. Each of these positions has become grist for someone's mill in arguing either that pain is a paradigm instance of a conscious state or that pain is a special case and should not be included in any general theory of consciousness.

PHILOSOPHICAL VIEWS OF PAIN

Some philosophers and psychologists hold that pain is completely subjective: Either it is essentially private and completely mysterious, or it does not correlate with any biological markers but is completely nonmysterious. The International Association for the Study of Pain (IASP), the formal organization charged with defining pain, has articulated a paradigm subjective view. They write: "Pain is always subjective. . . . Many people report pain in the absence of tissue damage or any pathophysiological cause; usually this happens for psychological reasons. There is usually no way to distinguish their experience from that due to tissue damage if we take the subjective report. . . . [Pain] . . . is always a psychological state" (1986).

However, if one holds that pain does not correlate in some way with some sort of bodily state or event, one becomes a dualist. If pain just is a private experience, and that experience has no consistent underlying physical cause or correlate, then any interesting connection between the mind and the body over pain is lost.

Philosophers can eschew dualism by retreating to so-called token-token identity theory. Every experience in some creature is correlated with—identical to—some event or other in that creature's brain. And every experience in some other creature is correlated with—identical to—some event or other in that creature's brain. If the subjectivists are right, then there is no identifiable neural activity that is the same across all experiences of a type of pain. There is no brain correlate for the type "having a migraine headache," for example. Generic headache experiences occur only at a level of abstraction above brain activity—namely, in the mind and its cognitive states.

However, if philosophers deny type-type identity for larger brain structures across organisms, then they are also denying any hope of discovering mind-brain connections. For mental event-physical state correlations taken one at a time are all a robust token-token identity theory allows.

At the same time, scientists do believe that there are areas in the brain dedicated to pain processing, just as there are other areas dedicated to vision, audition, touch,

and so forth. They believe that these areas are basically the same across humans, despite individual variation. Thus, even though a strict type-type identity might fail for particular sensory experiences, it still underlies views of our sensory systems taken as a whole. Types in science are allowed some play in them. They have to, or else there would be no mechanism by which to pick out any sort of cognitive processing in the brain at all.

All these lessons are missed by proponents of the subjective view, for they identify pain with the experience of pain and then explicitly deny that that experience has any correlation with any particular bodily reaction. But insofar as they want to be materialists interested in a scientific understanding of pain, they will have to permit generalizations connecting something in the body with the sensation of pain (see Hardcastle 1999).

Other philosophers and neurophysiologists argue that pain is completely objective; it is either intrinsic to the injured body part, a functional state, a set of behavioral reactions, or a type of perception. Pain is something that can be measured in bodies or behavior. As such, its connection to mentality, to sensations of pain, is secondary at best. Humans might recognize pain in terms of how it feels—the skin burns, for example. But, according to objective views that take pain as intrinsic to the injured body part, the pain itself is in the tissue. Hence, beliefs or judgments about the condition of the tissue are derivative—that is, pain is inferred from peripheral nociceptive or pain information (Annad and Craig 1996, Derbyshire 1996).

Similarly, if pain is understood as a type of perceptual process, then it works no differently than vision or olfaction. Animals receive some sort of perceptual input on their transducers, manipulate that information in their brains, and then use that manipulated information to alter motor reactions and other mental states. Part of the manipulated information might come into conscious awareness, but that sensation would constitute only a subset of what is meant by pain processing. According to this view, conscious experiences of pain, the damaged tissue itself, and the bodily and emotional reactions are all fundamental to pain processing. Each is one component in a larger process. Working together, these components take pressure, temperature, and chemical readings of tissues and use this information to track what is happening in bodies (Wall and Melzack 1989).

In these cases and most other instances of the objective view, pain is something entirely physical. Prima facie, it appears that the states or processes identified with pain could occur without any awareness of them at all. Most objective views of pain have the unintuitive consequence of divorcing pain from sensations of pain or making the mental events associated with pain processing secondary to and dependent upon the pain processing itself.

There are a few objectivist philosophers who hold that pain is not a purely physical event. Instead, it is something like an attitudinal relation. Pain requires both a bodily state and then cognition over that state. Pain itself is the attitude, the belief, regarding one's bodily condition. This approach gets around the intuitive difficulties of the objective views by identifying pain with the consequent mental state. "Pain" then just refers to the mental event associated with pain processing. According to this view, there is pain processing and then pain proper.

CENTRAL PHILOSOPHICAL ISSUES

There are three large philosophical difficulties in defending any of the theories about pain processing outlined above: the problem of mental causation, the problem of naturalizing content, and the threat of eliminativism.

The difficulty with mental causation is roughly as follows. If one drops a hammer on one's foot and subsequently experiences pain, *that experience* is the proximal cause of one's writhing, cursing, and gnashing of teeth. Dropping a hammer on one's foot leads to pain behavior only if it causes in one the sensation of pain and the belief that one is in pain. If one were unconscious or otherwise oblivious to one's surroundings, then one could not sense any pain, nor could one believe that one were in pain. One could manifest no pain-related behavior either.

On the other hand, a neurophysiological view of the hammer-dropping incident seems be able to explain exactly the same events without appealing to mentality or any sort of psychological entities at all. Neurophysiologists might talk about how the intense pressure of the hammer head on a foot stimulates various nerve endings and thus causes action potentials to travel up a leg to a spinal column, where other nerves are then stimulated to fire. These nerves transmit the firing pattern to other nerves, and so it goes until nerves that cause muscles to contract are likewise stimulated and one gets the writhing, wincing, and teeth-gnashing behavior. Why doesn't the possibility of this sort of more precise, purely physical explanation rule out the higher-level, more general mental account? Or why doesn't it make the mental account nothing more than a placeholder until the details of our central nervous system get figured out? As long as one is persuaded by reductionism, then pain provides an exemplar case for why psychological explanations appear so tricky.

There is some evidence that depression is related to pain processing. One view is that untreatable chronic pain causes depression, which in turn increases the sensations of pain. This is a (grossly oversimplified) mentalistic explanation of how a mood causally interacts with other psychological states. At the same time, we know that depression is correlated with a decrease in the neurotransmitter serotonin. Persons suffering from just an imbalance of a neurotransmitter and sensations of pain are some neural state or other, then it seems that the relation between depression and pain should be explained in terms of neurotransmitters affecting neural activity. In this case, the mentalistic explanation is just a stand-in until all the more basic neurphysiological details are revealed.

Mental events causing other mental events seems to be a natural part of the explanatory world. At the same time, accounts of mental causation appear to be nothing over and above a sloppy characterization of more fine-grained and little understood physical details. The difficulty for those who would like to keep the mind intact as an explanatory unit is explicating how it is that mental causation has a legitimate place in an understanding of the universe above and beyond being a surrogate for the real causal story.

Though most philosophers of mind treat mental causation separately from issues concerning reference, explaining the causal powers of the mind really piggybacks on the problem of naturalizing content. What makes the question of mental causality peculiar is that the content of the mental states is relevant to their efficacy. One winces and nurses one's foot because one's corresponding mental states are about one's foot. If they were about something else, then one would most likely be doing something else. To explain exactly how it is that mental events cause other things, philosophers are first going to have to explain how it is they refer. That is, to justify privileging a mentalistic explanation of sensations and beliefs over a lower-level physicalistic one of neuronal firing patterns or ionic flow, first philosophers have to have a clear grasp on what it means to have mental events with content, since their content is what is causally relevant to subsequent behavior.

The question about the power of the content of beliefs and other mental states is quite important to understanding pain processing (Gamsa 1994). What one is thinking and believing about the world strongly influences how much pain one feels. Athletes intently focusing on their game can break large bones and not even notice it. But the same athletes, alone in their living rooms, will writhe on the floor if they stub their toes. Chronic pain patients can be trained to diminish their sensation of pain by changing their focus of attention and their beliefs about death and disease. Those suffering congenital indifference to pain often lead short and unpleasant lives both because they can't sense painful stimuli but also because they cannot form appropriate beliefs about the meaning of the vague tinglings they do feel. How pain feels depends to a large extent on the current doxastic milieu. Hence, understanding pain is going to require understanding what beliefs and desires (and other mental states) are and how they refer.

One implication of current scientific theories of pain is that folk ways of describing pains are inadequate and people would be better off eliminating the descriptors from everyday practices (Dennett 1978). The claim is that folkways of talking about pain comprise a rough and ready theory of pain. This theory assumes that pains are identical to the sensations of pain and that the word *pain* can capture the essence of that sensation. From the perspective of some objective views of pain, both assumptions are dubious. Pain processing is enormously complicated, and sensations of pain form only a tiny subset of what these processors do. But even if one focuses exclusively on sensations, the most important to folkways of being, the folk theory is still inadequate. Words to express all the dimensions of pain experiences simply do not exist. The descriptors used are either metaphorical or nonexistent. The folk theory of pain needs to be replaced by something commensurate with the phenomenology.

Consider that not only can the sensory, affective, and cognitive dimensions of pain be distinguished phenomenologically, but they can also be manipulated independently of one another. Mammals can feel a shooting pain in their legs but not suffer in the least from it; they can be in agony from pain without feeling any particular sensation localized to any part of their bodies. Philosophers could just decide by fiat that *pain* is going to refer to the localized sensations, or they could just decide that *pain* is going to refer to the suffering. But either way they do violence to folk notions of pain, which require that a single simple sense datum both seem to occur in some place and be unpleasant.

In response to these sorts of claims, some have argued that folk views of pain do not constitute a theory in any meaningful sense. Some believe that certain introspective facts are known indubitably. Pain is touted as one of those things. Perhaps there are some sensory states, like pain, about which people have special first-person apprehension; no inference of judgment is required.

However, it is quite easy to demonstrate that introspective knowledge of pain can be mistaken. If one burns one's hand by touching something hot, one jerks one's hand away from the heat source. This is a reflex action; the nociceptive information travels up the arm to the spinal column and then back down again. It takes about 20 to 40 msec from stimulus to behavior. The information also travels up the spinal column to the brain. One feels the burn as well. Unlike the reflex movement, this processing is more complicated and takes about 200 to 500 msec from stimulus to percept, a full order of magnitude longer.

Nevertheless, if one introspectively reports on what the incident feels like, one says that one moved one's hand away after one felt the pain; feeling pain initiated the motor sequence. For whatever reason, brains backdate pain sensations so that they seem causally relevant to reflex behavior. But clearly the effect is not caused after it occurs, so the introspective report has to be wrong. There is not any special, first-person knowledge of pains. Whatever knowledge is had is embedded and informed by a conceptual framework of the brains' devising. Despite protests to the contrary, pain experiences have all the earmarks of being at least prototheoretical in nature.

Other detractors point out that even if a completed science of pain does not use folk terms for pain, that would not imply that those sorts of mental states do not exist; they just would not be referred to in scientific discourse. The notion of pain would be analogous to ideas about tables and chairs, germs and gems, and birthday presents and birthday cake. These are perfectly legitimate terms. Science just does not use them. Being cultural artifacts of one stripe or another, they do not refer to things about which there are laws. There might not be a mental science or laws about pains, but folk psychology could still be used as it is now, in everyday explanations of behavior.

There is something undoubtedly right about this charge. In many ways, pain experiences are environmentally determined. Puppies raised without ever experiencing pain and without ever seeing any other dog in pain will exhibit no pain behavior. They will repeatedly sniff a lighted match without fear and then show no reaction when burned. Children learn both pain behaviors and the emotional concomitants to pain from the reactions of others around them. Expressions of pain and reports of sensation and experience are significantly different across cultures. Most of pain experiences and expressions are socially relative, a cultural artifact of sorts.

However, social relativity is not enough to show that folkways of understanding pain are adequate. Different cultures have different experiences; they also have different ways of understanding these experiences. Nevertheless, the burden falls on the folk psychologist to demonstrate how folk theories of pain are actually successful. This work has not just begun.

THE ETHICS OF PAIN TREATMENT

One of the most hotly debated subjects in pediatric care concerns whether infants are insensitive to pain (cf. Lawson 1988). The presumption historically has been that because young infants are not conscious, they cannot sense pain. As a result, analgesics and anesthesias are rarely used, even in the most invasive of procedures.

At first, this presumption of insensitivity is curious because infants' reactions to painful stimuli are well documented. Even premature neonates exhibit stress responses, hormonal fluctuations, and slowed recovery to painful interventions. In fact, the afferent nociceptive system is up and running by twenty-nine weeks of gestation, even though the pain inhibitory systems do not come on line until later. If anything, infants should be more sensitive to pain than adults. At least, by all indications, infants are sensitive to pain in some sense or other.

However, the question for many doctors is whether infants are aware of their pain. Some argue that unless neonates can consciously apprehend pain, then any sort of response they give to noxious stimuli are merely reflexes. Hence, there is no reason to treat infants' pain because the infants cannot feel anything.

Suppose they are right, even though there is much that goes on in brains that is neither conscious nor mere reflex. It is still the case that infants react to pain, both behaviorally and physiologically, that these reactions can be modified with relatively simple treatments, and that treating pain has an impact on recovery. Early exposure to pain, whether remembered or not, affects later experiences of and reactions to pain by altering the developmental course of the nervous system. Infants, like other newborn animals, learn to attach particular meanings or emotions or importance to particular experiences in virtue of what is associated with those experiences. This sort of behavioral malleability is very important if an organism is going to survive in a complex environment. Consequently, manipulating early experiences can have drastic effects later on, as animal studies show. Merely by changing the smells associated with suckling, scientists can alter adult sexual behavior in male rats, for example. Similar changes occur with pain processing in young

infants. Nociceptive stimuli increase the size of the somatic receptive fields for neurons sensitive to pain and help maintain dendritic connections that would otherwise be eliminated over time. Perhaps, as some believe, chronic pain and hypersensitivity can result from early acute pain episodes, given how the neural receptors change. Early pain experiences have been shown to influence later personality and temperament. Something as common as circumcision can have lasting effects on pain sensitivity if done without anesthesia.

Given the impact early pain processing can have on later development, doctors have every reason to prevent infant pain, even if it feels dissimilar to an adult's, even if it feels like nothing at all to the infant. Whether infants consciously experience pain—and whether they are aware of some noxious stimulus or their own suffering—is a red herring. Available evidence converges around the idea that infants process pain, though perhaps not in the same way adults do. This processing has an impact on current behavior and later development. Because this influence is generally negative, insofar as we are able to prevent or alleviate some of their pain, we should.

See also Qualia.

Bibliography

Anand, K. J. S., and K. S. Craig. "New Perspectives on the Definition of Pain." *Pain* 67 (1996): 3–6.

Deberyshire, S. W. G. "Comment on Editorial by Anand and Craig." *Pain* 67 (1996): 210–211.

Dennett, D. C. "Why You Can't Make a Computer That Feels Pain." *Synthese* 38 (1978): 449.

Gamsa, A. "The Role of Psychological Factors in Chronic Pain, I and II." *Pain* 57 (1994): 5–29.

Hardcastle, V. G. *The Myth of Pain*. Cambridge, MA: The MIT Press, 1999.

International Association for the Study of Pain (IASP). Subcommittee on Classification. "Pain Terms: A Current List with Definitions and Notes on Usage." *Pain* (supplement) 3 (1986): 217.

Lawson, J. "Pain in the Neonate and Fetus." *New England Journal of Medicine* 318 (1988): 1, 398.

Wall, P. D., and R. Melzack, eds. *Textbook of Pain*. 2nd ed. New York: Churchill Livingstone, 1989.

Valerie Gray Hardcastle (2005)

PAIN, ETHICAL SIGNIFICANCE OF

Pain is a paradigm of an intrinsically bad mental state: It is an experience that is harmful to those who undergo it and makes their life go worse. Virtually all moral theories recognize norms to assist those who suffer from pain and to avoid inflicting unnecessary pain on others, though there is some disagreement about the source of these norms, their exact content, and their scope. The moral status of the pain of animals, for instance, remains a matter of controversy.

Pain has ethical significance when it is understood as an affective experience that is unpleasant or disliked in itself. Thus understood, pain belongs to a family of distinct but overlapping evaluative notions such as distress and suffering. The word "pain," however, is also used to refer to a type of bodily sensation typically associated with damage to body tissue. We normally find such sensations unpleasant, but when they are unaccompanied by an affective response (as reported by patients after frontal lobotomy) or when they are very mild, they are not experienced as unpleasant and no longer have this ethical significance. Furthermore, many hurtful experiences, both physical (nausea, electric shock) and mental (fear, regret) have a negative affective dimension without possessing the specific sensory quality common to cuts and burns. It is thus only pain in the broader, affective sense that is of direct interest to ethics.

The experience of pain is bad in itself but pain is also associated with other ills. Physical pain often accompanies bodily injury, and pain generally tends to incapacitate agents. It is important to distinguish the intrinsic badness of pain from these further harms. We also need to distinguish the badness of pain from a range of goods in which pain can play a part. Pain is instrumentally good insofar as it alerts us to bodily injury, for example. Many regard the painful aspect of just punishment as good, and some view pain as a necessary condition for the development of moral character and spiritual growth, for example. In all of these cases, however, pain can still be said to retain its badness for the agent. Thus pain justly inflicted on those who deserve it counts as punishment, and as good overall, only because it is also bad in itself for the offender. Other cases, such as masochism and the pain of grief, are harder to interpret.

Pain is often contrasted with hedonic states of positive value, such as pleasure and enjoyment. It should not be assumed, however, that pain and pleasure are simple contraries, since the occurrence or prospect of pain appears to have a different moral status, and to give reasons of greater force and urgency, than the occurrence or prospect of pleasure of equal intensity.

Pain also raises questions of ascription and measurement. It is often thought that subjects' sincere reports

about their own pain are authoritative. There are also objective, largely behavioral criteria for ascribing pain. These used to be our exclusive means of detecting pain in animals and infants. These first- and third-person criteria seem ill-equipped, however, to deal with some of the cases reported by doctors and scientists, such as frontal lobotomy and hypnosis. The increased availability of devices that can directly detect the neural correlates of pain may present further challenges to our everyday practice of ascribing and assessing pain.

See also Happiness; Hedonism; Intrinsic Value; Pleasure.

Bibliography

Cassell, Eric J. *The Nature of Suffering and the Goals of Medicine*. New York: Oxford University Press, 1991.

Hare, Richard M. "Pain and Evil." In his *Essays on the Moral Concepts*. London: Macmillan, 1972.

Mayerfeld, Jamie. *Suffering and Moral Responsibility*. Oxford, U.K.: Oxford University Press, 1999.

Pitcher, George. "The Awfulness of Pain." *Journal of Philosophy* 68 (1970): 481–492.

Wall, Patrick. *Pain: The Science of Suffering*. London: Weidenfeld and Nicolson, 1999.

Guy Kahane (2005)

PAINE, THOMAS

(1737–1809)

Thomas Paine, the author, deist, and American revolutionary leader, was born at Thetford, Norfolk, in England. After an inconspicuous start in life as corset maker and customs officer, Paine emigrated at the age of thirty-seven from England to Philadelphia, carrying a letter of recommendation from Benjamin Franklin. Caught up almost immediately in the turmoil of the developing revolution, Paine published *Common Sense* (January 1776), the first public appeal for American independence as well as the pioneer enunciation of the diplomatic doctrine of avoiding European entanglements. In addition to attacking hereditary aristocracy, Paine expounded the theory that government and society are distinct entities and are not to be confounded, a theory also developed by Jean-Jacques Rousseau and later by William Godwin.

During subsequent stages of the American Revolution, Paine wrote a number of influential newspaper essays, including a famous series, the *Crisis,* concerned with particular political, economic, and military issues. In order to extend his reputation to Europe, Paine wrote the *Letter to the abbé Raynal, on the Affairs of North America* (1783), refuting among other concepts of the French *philosophes,* the assertion that the Revolution concerned only economic issues and had no moral foundation. A confident affirmation of the idea of progress was incorporated in Paine's notions that the circle of civilization was soon to be completed and that commerce and science had already combined to improve the world to the point where there no longer existed a need to make war for profit.

After the American victory, Paine proceeded to France to seek financial support for an iron bridge of his own invention, once again carrying letters of recommendation from Franklin. In January 1790 he began a work defending Lafayette and the principles of the revolution that had broken out in France, a work that he later converted to an attack on Edmund Burke's highly critical *Reflections on the French Revolution.* The resulting treatise, *The Rights of Man* (Part I, 1791; Part II, 1792), gave a solid theoretical basis to the contingent appeals of Paine's American journalism. Affirming that government should be founded on reason rather than on tradition or precedent, Paine argued that democracy—a society in which all men have equal rights and in which leadership depends upon talent and wisdom—is superior to aristocracy. Although his political principles resemble those of John Locke, Paine later maintained that they were based entirely on his own reasoning and that he had never read the works of the English philosopher.

As a result of his republican writings, Paine was made an honorary citizen of France and in September 1792 he was elected to the French National Convention, taking his seat later that month.

Disturbed by the dogmatic atheism of the French revolutionary leaders, Paine began a treatise on religion, *The Age of Reason,* ostensibly a defense of deism but primarily an attack on Christianity. In Part I (1794), he rejected all forms of supernatural revelation in favor of the religion of nature, elevating, as he put it, reason and scientific observation over the three modes of superstition in Christianity: mystery, miracle, and prophecy. In Part II (1795), Paine continued to praise "the Perfection of the Deity," even though he exposed the abuses of Christianity with such vehemence that he brought upon himself the inaccurate accusation of opposing religion itself.

Although Paine dismissed the miracles of Christianity, he was later ready to believe that providence intervened in his own life. The story is incredible, but it reflects Paine's egoism. Because of his moderate policies in the Convention, particularly in an appeal to save Louis XVI

from the guillotine, he was dismissed from the Convention and incarcerated in Luxembourg Prison. On his return to America, Paine explained that the cell doors of prisoners destined for execution were customarily marked with a number, and he argued that divine providence had protected him by causing his jailer to place the fatal number by mistake on the inside of his door so that it could not be seen the next morning.

One must turn to Paine's minor works to discover the positive side of his deism. His proof of the existence of God (in "A Discourse at the Society of Theophilanthropists") adopts essentially the same reasoning that Isaac Newton had used in a series of letters to an Anglican clergyman, Richard Bentley. Since the laws of mechanics, the argument runs, cannot explain the origin of motion, there must have been an external first cause to give the planets their original rotation. Paine stressed the concept of the plurality of worlds and assumed absolute moral laws. In "Private Thoughts on a Future State," he expressed a faith in an immortality strikingly different from that of most deists. The good people, he believed, would be happy in another world; the wicked would be punished; and those in between—the indifferent ones—would be "dropped entirely." Although contending that religion should be a private affair between each man and his creator, he insisted that no rational mind could logically reconcile new science and old Christianity.

Unable to adjust to French political life under Napoleon Bonaparte, Paine returned to America in 1802, where he was welcomed by liberal Jeffersonians but excoriated by most Federalists. Although he contributed extensively to newspapers under his revolutionary pseudonym of "Common Sense," he failed to regain his earlier influence and died in obscurity.

Paine, as much as any thinker of his age, was obsessed with the notion of the order and uniformity of nature, and he delighted in establishing parallels between one branch of learning and another. He believed that the fundamental laws of nature operative in religion, natural science, and politics were clear, simple, and within the reach of the average man. He developed no epistemology as such but combined a type of Quaker inner light with deistic reason. The fundamental weakness of his system—a weakness shared by most deists—is that he nowhere took up the problem of evil. Although he lavishly praised God for the regularity of the universe, the only suffering he noticed is that caused by social injustice.

Yet even though Paine was more influential as an agitator than as a theorist, he certainly understood and upheld the ideals of the Enlightenment and deserves to be ranked as one of America's outstanding *philosophes.*

See also Deism; Democracy; Egoism and Altruism; Enlightenment; Evil, The Problem of; Franklin, Benjamin; Godwin, William; Locke, John; Newton, Isaac; Political Philosophy, History of; Rousseau, Jean-Jacques.

Bibliography

WORKS BY PAINE

Writings of Thomas Paine. 4 vols, edited by Moncure D. Conway. New York. 1894–1896.

Complete Writings of Thomas Paine. 2 vols, edited by Philip S. Foner. New York: Citadel Press, 1945.

WORKS ON PAINE

Aldridge, Alfred Owen. *Man of Reason: The Life of Thomas Paine.* New York, 1959.

Ayer, A. J. *Thomas Paine.* Chicago: University of Chicago Press, 1988.

Claeys, Gregory. *Thomas Paine: Social and Political Thought.* London: Routledge, 1989.

Conway, Moncure D. *Life of Thomas Paine.* 2 vols. New York: Putnam, 1892.

Dyck, Ian, ed. *Citizen of the World: Essays on Thomas Paine.* New York: St. Martin's Press, 1988.

Foner, Eric. *Tom Paine and Revolutionary America.* Oxford: Oxford University Press, 1976.

Russell, Bertrand. "The Fate of Thomas Paine." In *Why I Am Not a Christian.* New York: Simon and Schuster, 1957.

Alfred Owen Aldridge (1967)

Bibliography updated by Christian B. Miller (2005)

PALÁGYI, MENYHERT

(1859–1924)

Menyhert (or Melchior) Palágyi, a scientist, literary critic, and philosopher, was born in Paks in west central Hungary. He studied science at Budapest, but his main activity there was as a literary critic. After 1900 he spent much time in Germany, studying informally with philosophers in many places. For a time he held a readership in physics and mathematics in Kolozsvár, Hungary (now Cluj-Napoca, Romania). He had little contact with Hungarian philosophers, however, and eventually returned to Germany, where he died in Darmstadt.

Throughout Palágyi's philosophical works, psychological doctrines and speculations on theoretical physics are mingled with his main interest in epistemology. He interpreted and criticized the then new theory of relativity from the point of view of epistemology, and episte-

mology from the point of view of his psychological theory. As he expressed his views in response to the new developments in these fields, he became somewhat lost in their transitional stages, and the fact that he criticized them from his own particular standpoint hindered his understanding of them. The central dominating idea throughout his works is a broadly Hegelian principle of polarity. It asserts an interdependence of opposites, a sort of cooperative unity, and it was applied by Palágyi with no apparent consistency and even more liberally than Hegelian dialectics would be. Palágyi was a monist who held a curious version of the denial of the distinction between the a priori and a posteriori.

His most purely philosophical work is *Der Streit der Psychologisten und Formalisten in der modernen Logik* (Leipzig, 1902). In it, among other things, he criticized Edmund Husserl for "tearing" logic away from psychology and "submerging" it in mathematics, and for his "ideal meaning" and his distinction between real and ideal laws. (Husserl himself reviewed this book in *Zeitschrift für Psychologie und Physik des Sinnesorgane* 31 [1903].) In the same year Palágyi wrote his *Die Logik auf dem Scheidewege* ("Logic at the Crossroads," Berlin and Leipzig, 1903). In these works Palágyi's main concern was not, despite his criticisms of Husserl, a return to psychologism but his principle of polarity. In his psychology, in fact, he tried to rescue from psychologism that which he termed "mental" (even though he only obscurely described the term). The source of all error is to mistake what is mental for what is merely vital (and, in the spirit of "polarity," what is vital for what is merely mental). He distinguished between mechanical and vital processes and consciousness. The mechanical is publicly observable, and the vital indirectly observable, but consciousness escapes observation by the methods applicable to the other processes: consciousness "punctuates" the vital process and is discontinuous. (He nevertheless explicitly affirmed the unity of the self, although it is doubtful how he could maintain this.) Our knowledge depends on the speed of these punctuations. God is the limiting case who grasps the whole time process instantaneously; for him all punctuations are one. This led Palágyi to such metaphysical claims as that our knowledge catches eternity in the fleeting moment, which is both temporal and eternal.

At the base of this theory of perception was his notion of imagined movement. Touch being the basic sense, all perception depends on our ability to trace the object in the imagination. He mistakenly supported this view by reference to the Kantian role of imagination in perception. His theoretical physics, in which his main interest was our perception of space-time (space-time being a unity in polarity), can best be understood if approached through this theory of perception.

See also A Priori and A Posteriori; Epistemology; Hegelianism; Husserl, Edmund; Imagination; Philosophy of Physics; Psychologism; Relativity Theory; Touch.

Bibliography

Ausgewahlte Werke, 3 vols. (Leipzig, 1924–1925), contains Palágyi's most important works.

Works on Palágyi are Werner Deubel, "Die Philosophie und Weltmechanik von Melchior Palágyi," in *Preussische Jahrbuch* 203 (1926): 329–356, with complete bibliography; W. R. Boyce Gibson, "The Philosophy of Melchior Palágyi," in *Journal of Philosophical Studies* 3 (1928): 15–28, 158–172; L. W. Schneider, *Der erste Periode in philosophischen Schaffen M. Palágyis* (Würzburg, 1942); and A. Wurmb, *Darstellung und Kritik der logische Grundbegriffe der Naturphilosophie Melchior Palágyis* (Leipzig, 1931).

Julius Kovesi (1967)

PALEY, WILLIAM

(1743–1805)

William Paley was an English theologian and moral philosopher. His father, William, was vicar of Helpston, Northamptonshire, and a minor canon of Peterborough; he later became headmaster of Giggleswick grammar school, where the younger Paley was educated. Paley entered Christ's College, Cambridge, in 1759, where he studied mathematics and became a senior wrangler. After an interlude of school teaching, he was elected a fellow of his college in 1766 and was ordained a priest in the established church in 1767. He taught at Cambridge for nine years, leaving the university only on his marriage. He held successively a number of different offices in the church, rising to be the archdeacon of Carlisle. Paley was the author of three books, one on morals and two defending Christian belief, all of which were widely read and accepted as textbooks. As late as 1831, Charles Darwin, studying for his BA examination at Cambridge, had to "get up" Paley's *A View of the Evidences of Christianity, The Principles of Moral and Political Philosophy,* and *Natural Theology.* The *Moral and Political Philosophy* contains Paley's famous satire on property, in which he describes the plight of a flock of pigeons in which private property is permitted. Although he immediately proceeds to list the advantages of a system of private property, his satire is savage ("the weakest perhaps, and worst pigeon of the

flock" controls and wastes all the grain as he pleases), and Paley's friends are said to have assured him (correctly) that the publication of the passage would cost him a bishopric. It did earn him the nickname "Pigeon Paley."

Paley's *The Principles of Moral and Political Philosophy* (London, 1785) is a handbook on the duties and obligations of civil life rather than a philosophical treatise. The subtlety of the work may be gauged by its opening sentence: "Moral philosophy, Morality, Ethics, Casuistry, Natural Law, mean all the same thing; namely, that science which teaches men their duty and the reasons of it." Paley's definition of duty follows from his theological utilitarianism. The nature of the human frame implies that it is God's will for us to be happy in this life as well as in the next. Virtue is doing good to humankind, in obedience to the will of God and for the sake of everlasting happiness. Allegiance to God's will and a desire for everlasting happiness are sufficient grounds for moral obligation. Paley offers this account of moral obligation after finding that such obligation follows from the command of a superior, which is made persuasive by the prospect of a reward.

We may discover the will of God by consulting either Scripture or "the light of nature," both of which lead to the same conclusion. The will of God with regard to any action may be found by inquiring into its "tendency to promote or diminish the general happiness." We should carry out those actions that promote the general happiness and avoid those which diminish it. Promoting the general happiness requires paying attention to the general consequences of our actions. Paley offers a rule for assessing general consequences that resembles Immanuel Kant's categorical imperative: "The general consequence of any action may be estimated by asking what would be the consequence if the same sort of actions were generally permitted."

Paley believed that no special faculty is required to enable us to have moral knowledge. Thus he dismissed the views of those who have argued that morality requires either a moral sense, or an intuitive perception of right and wrong, or any other innate or instinctive capacity. All that is required for the foundation of morality is that each man has the wit to see that certain actions are beneficial to himself. Then the sentiment of approbation that naturally arises when these actions benefit him will continue to accompany his perception of these actions when they benefit someone else. Thus the custom of approving certain actions is begun, and children, who learn everything by imitating their elders, carry it on.

The bulk of the *Principles* is a detailed discussion of our duties to others, to ourselves, and to God. The final part is an outline of the elements of political knowledge. The wide acclaim accorded Paley's work is said to have stirred Jeremy Bentham to bring out his own version of the utilitarian doctrine in *Introduction to the Principles of Morals and Legislation* (1789).

Paley is the author of two theological works with the word *evidence* in their titles. The first, *A View of the Evidences of Christianity* (2 vols., London, 1794), is an essay in apologetics. The second, *Natural Theology; or, Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature* (London, 1802), is, as its title implies, an essay on natural theology. The books, which are similar in tone (they are both presented as judicious, lawyerlike statements of a case) doubtless owe much to Paley's lifelong interest in trials and the art of advocacy.

A View of the Evidences of Christianity demonstrates what can be said on behalf of Christian belief by an appeal to the behavior of the earliest Christians. Paley asks his readers to grant the possibility that God should have destined his human creation for a future state and that he should acquaint human beings with their destiny. If these possibilities are granted, then the need for miracles is clear, for they are the certification of revelation. The credibility of the Christian revelation hangs, therefore, on the issue of whether its miracles are genuine.

It is Paley's claim that the miracles on which Christianity is based (including those of the Old Testament) are genuine; and that indeed the only genuine miracles are those of Christianity (including its Jewish origins). Paley accepts David Hume's contention that the believability of Christianity rests ultimately on the reliability of the testimony of the earliest Christians, but he rejects Hume's thesis that no testimony for a miracle can ever be relied on because such testimony goes against universal experience. He argues that universal experience is too strong a test. By definition, miracles must be exceptions to universal experience or they would not be miracles. The real issue is whether there is a test for the reliability of witnesses who report an event that necessarily only they could have experienced. Paley finds such a test in our observation of whether the person who reports a miracle will cling to his report at the risk of his comfort, his happiness, and even his life. According to Paley, the original witnesses of the Christian miracles pass this test, since they labored and suffered "in attestation of the accounts which they delivered, and solely in consequence of their belief of these accounts."

Paley's hospitality for miracles is not quite so broad as we might at first think. The miraculous event must be in support of a revelation that is important to human happiness. Mere wonders are thus ruled out; and Paley also holds out against any event that may be resolved into a false perception and against any report that is guilty of exaggeration. But even after setting these limits, Paley maintains that a significant core of miracles stands as the guaranty of the Christian revelation. But the acceptance of these miracles must finally rest on the steadfastness of the original Christians; and the weakness of Paley's argument can be seen when we consider its close resemblance to a lawyer's defending his client by calling for the testimony of none but character witnesses. *A View of the Evidences of Christianity* had a huge success, and the bishops made Paley a prebendary of St. Pancras in the Cathedral of St. Paul's and the subdean of Lincoln.

In his *Natural Theology,* Paley appeals to a number of natural phenomena to establish the existence of a god. He states his argument at the very outset, and the remainder of the work is a train of examples illustrating that argument. The line of the argument runs as follows. If I found a stone while crossing a heath, and if I "were asked how the stone came to be there, I might possibly answer, that, for any thing I knew to the contrary, it had lain there forever; nor would it perhaps be very easy to show the absurdity of this answer. But suppose I had found a watch upon the ground, and it should be enquired how the watch happened to be in that place, I should hardly think of the answer which I had before given, that, for anything I knew, the watch might have always been there. Yet why should not this answer serve for the watch, as well as for the stone?" Paley answers, "For this reason, and for no other, viz. that when we come to inspect the watch, we perceive (what we could not discover in the stone) that its several parts are framed and put together for a purpose"—that is, to tell the time. The care with which the parts have been made and the fineness of their adjustment can have only one implication, namely, that the watch must have had a maker who understood its construction and who designed it for the use for which it is fitted. The conclusion would not be weakened if we had never seen a watch being made or could not conceive of how to make one. Nor would it be weakened if there were parts of the watch whose purpose we could not understand, or even if we could not ascertain whether these parts had some effect in the general purpose of the watch. Nor should we be satisfied if we were told either that the existence of the watch is to be explained by a principle of order which exists in things and disposes the parts of the watch into their present form and situation, or that the watch is the result of the laws of "metallic nature." Finally, we should be surprised to hear that the mechanism of the watch is no proof of contrivance, but "only a motive to induce the mind to think so." In short, where there is mechanism, instrumentality, or contrivance, there must have been an intelligence who designed and made the machine, the instrument, the contrivance.

Paley then turns to nature with this argument in hand and, in his own words, applies it to adduce evidences of the existence of God. The bones and muscles of human beings, animals, and their insect equivalents, are of special interest to Paley, for the fitting together of joints and the adaptation of muscles are mechanisms that imply most forcefully a designing intelligence. The chemical side of physiology does not interest him much, for chemical action does not suggest the work of a divine mechanic. But Kiell's *Anatomy* is ransacked for appropriate examples, and the hare's backbone is picked apart at the end of the meat course to show the finesse of divine contrivance. The example that most interests Paley, and to which he often returns, is the eye, in its various parts and in the combination of these parts and their adaptation to function as an instrument of sight. As he remarks, he offers many examples of natural mechanism, but a single instance, the eye alone, should suffice to convince us of the existence of the divine intelligence that designed it.

The evidence drawn from nature, in addition to establishing the existence of God, permits us to infer certain of his characteristics. Because God has a mind, he must be a person. That there is a single intelligence at work is shown by the uniformity of the divine plan, as it is applied to all parts of the world. Finally, God's goodness is shown both by the fact that most contrivances are beneficial and by the fact that pleasure has been made an animal sensation.

At bottom, Paley's argument rests on his original decision to regard certain parts of nature as mechanisms or contrivances. If this decision is unquestioned, then his argument takes a long stride toward plausibility. Everything depends, however, on whether the human eye, for example, is analogous to a machine, and if so, how far this analogy takes us in the inference of other characteristics that the analogy might imply. These questions are raised and examined with devastating effect by Hume in the *Dialogues concerning Natural Religion,* a work published a quarter of a century before Paley's *Natural Theology.* It is to be regretted that Paley does not meet Hume's arguments head-on in the *Natural Theology,* in the same way that he meets Hume squarely on the issue of the believ-

ability of miracles in *A View of the Evidences of Christianity.*

See also Bentham, Jeremy; Darwin, Charles Robert; Hume, David; Kant, Immanuel; Miracles; Moral Sense; Revelation; Teleological Argument for the Existence of God; Utilitarianism.

Bibliography

WORKS BY PALEY

William Paley's collected writings are in *Works,* 8 vols. (London, 1805–1808). See also his *Natural Theology,* edited by Frederick Ferré (Indianapolis: Bobbs-Merrill, 1963).

WORKS ON PALEY

Barker, Ernest. "Paley and his Political Philosophy." In *Traditions of Civility.* Cambridge, U.K.: Cambridge University Press, 1948.

Clarke, Martin. *Paley: Evidences for the Man.* Toronto: University of Toronto Press, 1974.

LeMahieu, Dan. *The Mind of William Paley.* Lincoln: University of Nebraska Press, 1976.

Nuovo, Victor. "Rethinking Paley." *Synthese* 91 (1992): 29–51.

Oppy, Graham. "Paley's Argument for Design." *Philo* 5 (2002): 161–173.

Schneewind, Jerome. *Sidgwick's Ethics and Victorian Moral Philosophy.* Oxford: Clarendon Press, 1977.

Sweet, William. "Paley, Whately, and 'Enlightenment Evidentialism.'" *International Journal for Philosophy of Religion* 45 (1999): 143–166.

Elmer Sprague (1967)
Bibliography updated by Christian B. Miller (2005)

PALMER, ELIHU
(1764–1806)

Elihu Palmer was a radical spokesman for the Age of Reason and Revolution in America, who along with Thomas Paine and Ethan Allen gave expression to the ideals of deism and republicanism. Born in Canterbury, Connecticut, Palmer was graduated from Dartmouth in 1787. Originally a minister, he was persecuted for his extreme religious views and forced to flee the pulpit. In 1793 he was admitted to the bar. Blinded by disease, he spent the last years of his life defending deism. He edited the deistic weekly journal *Prospect, or View of the Moral World* and helped to organize the Deistical Society in New York.

Palmer's religious radicalism stemmed from his reaction to Calvinism. He rejected the doctrine of original sin as well as the idea of a punitive and arbitrary divine being. This reaction developed into a militant anti-Christianity and anticlericalism. Palmer rejected the claims of divine revelation, miracles, and prophesies, and he accused the Bible of inconsistency, contradiction, and vagueness. Not only did he deny the divinity of Christ, but he considered Jesus, Moses, and Muḥammad indecent and immoral and Christian salvation absurd and irrational. He attacked organized and institutionalized religion for its hypocrisy and self-interest.

Like other deists, Palmer defended a religion of nature, in which the order and harmony of the universe is believed to proclaim the existence of one supreme being, the divine creator. Palmer maintained that evil is not inherent in man or in nature but is due to corrupt social institutions and to defective human knowledge, which can both be corrected. He had boundless faith and optimism in reason, science, and education, believing that man possesses the capacities for intellectual and moral progress. In place of the traditional religious depreciation of human ability and dignity, he proposed a humanistic ethics. With others of this period, he held an empiricist epistemology, locating the source of all knowledge in sensation, and he was sympathetic to scientific and materialistic philosophy. Palmer was an ardent supporter of liberty and republicanism and saw in the American Revolution the inception of a new era for humanity.

See also Deism; Paine, Thomas; Progress, The Idea of; Republicanism.

Bibliography

Palmer's works include *Political Miscellany* (New York, 1793); *The Examiners Examined: Being a Defence of the Age of Reason* (New York: Printed for the author and sold by L. Wayland and J. Fellows, 1794); *An Enquiry Relative to the Moral and Political Improvement of the Human Species* (New York: John Crookes, 1797); *The Political Happiness of Nations: An Oration* (New York: n.p., 1800); *Principles of Nature: Or, a Development of the Moral Causes of Happiness and Misery among the Human Species* (New York, 1801); *Prospect, or View of the Moral World for the Year 1804* (New York: E. Palmer, 1803–1805), which he edited; and *Posthumous Pieces* (London, 1826).

For literature on Palmer, see G. Adolf Koch, *Republican Religion: The American Revolution and the Cult of Reason* (New York: Holt, 1933).

Paul Kurtz (1967)

PANAETIUS OF RHODES
(c. 185–110 BCE)

Panaetius of Rhodes was a pupil of Diogenes of Babylon and Antipater of Tarsus, both heads of the Stoic school in

Athens, and he succeeded Antipater as scholarch in 129. Little is known about his life though it is clear that he spent considerable time in Rome and in the circle of P. Cornelius Scipio Aemilianus. None of his writings survive, but traces of his importance do.

First, isolated testimony from antiquity reveals that Panaetius was especially willing to disagree with earlier Stoics about central matters of doctrine. He rejected the Stoic belief in divination, and against the earlier account that the cosmos would be consumed periodically in flames, he insisted that the world is everlasting. He maintained that virtue is not sufficient for happiness, since health, some resources, and strength are also necessary, and he divided virtues into the contemplative and practical, which sits uneasily with traditional Stoic intellectualism.

These examples suggest that Panaetius was keen to incorporate more Platonic and especially Aristotelian doctrines into his Stoicism, and many ancient sources directly attest to this desire. This feature of Panaetius's philosophy links him to his pupil Posidonius, the polymath who showed similar willingness to infuse pre-Stoic ideas into his Stoicism. Together, Panaetius and Posidonius have been taken to epitomize Middle Stoicism, which stands between early Greek Stoicism and later Roman Stoicism, but this periodization is of limited utility because there are more than three ancient Stoicisms. Nevertheless, the affinities between Panaetius and Posidonius distinguish them from most other Stoics. Their broadly shared approach is also linked to the syncretizing philosophy of the first century BCE that is typified by Antiochus of Ascalon. Such thought has been disparaged as *eclectic*, but there is nothing unworthy in the attempt to produce a well-grounded synthesis of a rich and varied philosophical tradition.

The second trace of Panaetius is due to Cicero, who has characters call Panaetius "a great and extremely learned man" (*Leg* III 14) and "chief among the Stoics" (*Acad* II 107). Cicero based the first two books of his *On Duties* (*De Officiis* on Panaetius's *On Duty or Appropriate Action* (*Peri tou kathêkontos*), and this makes Panaetius influential since, as Henry Sidgwick notes: "There is probably no ancient treatise which has done more than [Cicero's] *De Officiis* to communicate a knowledge of ancient morality to medieval and modern Europe" (Sidgwick 1902, p. 95).

Among the prominent features of *De Officiis* that are likely due to Panaetius, the following three are especially important. First, Cicero notes that anyone who is beneficent must choose his beneficiaries carefully, and he insists that one should help some people more just because one stands in a naturally closer relationship with them. He develops the point by suggesting a hierarchy of natural relationships, from the closest (marriage) to the most remote (the relationship one shares with all other human beings). The later Stoic Hierocles imagines the hierarchy as a series of concentric circles, but Cicero's version of the probably Panaetian idea that one's duties of beneficence are tied to certain relational facts independent of how one feels about those relationships has proven enormously influential.

Second, after identifying the traditional virtue of temperance or moderation with seemliness (*decorum*), Cicero insists that to display *decorum*, one must act in accordance with all of one's roles (*personae*). So, one must consider not only the role that all human beings share in common but also the particular role one has on account of one's peculiar natural talents. Additionally, one must consider the role that fortune assigns by giving one power, wealth, standing, and their opposites, and one must consider the demands of the role one chooses by taking up a particular career. With this schema, Cicero, no doubt inspired by Panaetius, takes the traditional Stoic concern to act appropriately in the particular circumstances, and he incorporates special attention to the ways in which social roles and individual talents matter to the circumstances.

Third, Cicero spends much of *De Officiis* II providing advice about how to pursue honor or glory. Earlier Stoics generally agreed that although honor might be useful, it has no intrinsic attraction. Cicero rejects that view in favor of a more Platonic line, according to which humans are naturally drawn to honor. Because the honorable is dependent upon what other people honor, this line generally ties one's pursuit of natural aims to the values of others in one's society. It also represents an especially concrete way in which the Panaetian approach of Cicero's *De Officiis* moves away from the paradoxical excellences of the early Stoics' sage and closer to the virtues of Roman politicians.

There is a final trace of Panaetius's importance, for he seems to be central to the eventual diffusion of Stoic thought. Most obviously, as a member of the Scipionic Circle, Panaetius helped to spread Stoicism in Rome. More speculatively, one might think that he contributed decisively to the decentralization of the Stoic school. There is no record that Panaetius had a successor as head of the Stoic school in Athens. His student Posidonius attracted pupils not to Athens but to Rhodes, which, curiously enough, was Panaetius's but not Posidonius's

hometown. Did Panaetius arrange to have the school leave Athens? Did he otherwise let it die? Whatever his intentions, later Stoics studied and taught in a variety of places around the Mediterranean, and Stoicism continued to seep into a broad array of intellectual currents.

See also Antiochus of Ascalon; Aristotle; Cicero, Marcus Tullius; Diogenes Laertius; Plato; Posidonius; Sidgwick, Henry; Stoicism.

Bibliography

Cicero, Marcus Tullius. *Acadamica*. Hildesheim, Gg. Olma, 1966.

Cicero, Marcus Tullius. *On Duties* (*De Officiis*), edited by M. T. Griffin and E. M. Atkins. Cambridge, U.K.: Cambridge University Press, 1991.

Cicero, Marcus Tullius. *On the Commonwealth; and, On the Laws*, edited by James E.G. Zetzel. New York: Cambridge University Press, 1999.

Sidgwick, Henry. *Outlines of the History of Ethics for English Readers*. London: Macmillan, 1902.

TEXTS

Alesse, F. *Panezio di Rodi e la Tradizione Stoica*. Naples: Bibliopolis, 1994.

Straaten, M. van, ed. *Panaetii Rhodii Fragmenta*. Leiden: E. J. Brill, 1952.

Straaten, M. van. *Panétius*. Amsterdam: H. J. Paris, 1946.

STUDIES

Gill, Christopher. "Panaetius on the Virtue of Being Yourself." In *Images and Ideologies: Self-Definition in the Hellenistic World*, edited by A. Bulloch, et al. Berkeley: University of California Press, 1993, pp. 330–353.

Gill, Christopher. "Personhood and Personality: The Four-Personae Theory in Cicero, *De Officiis*." In *Oxford Studies in Ancient Philosophy*. Vol. VI. Oxford: Oxford University Press, 1988, pp. 169–199.

Eric Brown (2005)

PANENTHEISM

See *Emanationism; Krause, Karl Christian Friedrich*

PANNENBERG, WOLFHART

(1928–)

The thought of Wolfhart Pannenberg follows in the tradition of twentieth-century German systematic theology in replying to the secularizing nature of post-Enlightenment thought. Pannenberg's writings, however, unlike those of his near contemporaries, most notably Karl Barth and Rudolf Bultmann, do not reject the characteristic intellectual developments of Enlightenment thought. Rather, Pannenberg seeks to incorporate many of the key components of the Enlightenment into his comprehensive theological world view. Born in 1928, Pannenberg began his education as the University of Berlin. In 1950 he studied theology under Barth in Basle, and in 1951 he moved to Heidelberg where he completed his doctoral studies on the doctrine of predestination in Duns Scotus. Following this, he took up a teaching post at Heidelberg, later becoming Professor of Systematic Theology successively at Wuppertal, Mainz, and finally, in 1968, Munich.

Pannenberg's philosophical development was transformed by what he has described as an "intellectual conversion" to Christianity. This conversion, which was driven by his reading of philosophical as well as theological texts in his youth, has had two important influences on the development of his thought. First, Pannenberg's initial concerns are not with the Church and ecclesial theology. Instead, his thought centers on the role of religious experience on the individual within a created world defined by God. This *anthropological* aspect to Pannenberg's thought lies at the heart of his theological and philosophical system. Second, Pannenberg has been more receptive than many of his contemporaries in understanding and the developments in secular philosophical thought. Through all his writings, Pannenberg argues that many of the problems of modern secular thought can be resolved if God is reestablished as the defining principle of all creation. Pannenberg's most profound contribution to this debate has been through his dialogue with the secular aspects of critical history and latterly with the philosophy of science.

THE ANTHROPOLOGY OF RELIGIOUS EXPERIENCE

The starting point of Pannenberg's thought is his anthropological account of religious belief. Pannenberg's thought is based on the belief that God can be found naturally and freely within all aspects of human experience. This anthropological approach comes out most clearly in Pannenberg's 1983 work *Anthropology in Theological Perspective*. His main impetus in approaching theological questions in this manner is to address directly the implicit atheism of much post-Enlightenment thought. Pannenberg argues that the philosophical atheism of the Left-Hegelians, especially Ludwig Feuerbach, is in essence misguided anthropology. The philosophical atheism of Feuerbach defines God as merely the creation of the his-

torically developing human mind. Pannenberg takes issue with this, arguing that it crucially misinterprets the place and role of God in human thought. By concentrating on the social and cultural uses of religious forms and structures, Pannenberg argues that the Left-Hegelians were able to dismiss these as constructs of the alienated human mind. Therefore Feuerbach, in particular, was able to collapse theology into anthropology, asserting the form of the divine as God simply a construct of the human mind (Pannenberg 1973, p. 87).

To counter this powerful philosophical criticism of theology and religion, Pannenberg argues that we must consider humans in the first instance without recourse to religious categories or structures. He argues that such an approach is a necessary part of thinking about religion in the post-Enlightenment world, because the Enlightenment moved humans away from the traditional structures and forms of religious belief. Consequently, Pannenberg argues, we must look for God in all parts of human experience, not simply those that are exclusively religious. This approach, which he characteristically describes as coming to God "from below," places Pannenberg in opposition to the theology of Barth. Barth's solution to the dilemma presented by philosophical atheism was to stress God as "Wholly Other," inaccessible to man accept through the initiative of Jesus Christ.

Pannenberg argues that it is self-contradictory to talk of God in a manner that makes him completely inaccessible to humans. If God is the creative force of all creation, he must be accessible to people in all parts of creation. In the first instance one is able to come to this realization, Pannenberg argues, through a process of self-examination. By carrying out this anthropological enquiry, Pannenberg believes that people are able to recognize in themselves transcendent categories such as imagination that draw the human mind above and beyond a simple, mundane corporeal existence. It is through grasping this natural sense of transcendence that the human mind first comes to comprehend the existence of God. In doing this, Pannenberg is not rejecting traditional theological forms; rather he argues that the natural human desire to comprehend the divine is driven by very real human characteristics that God places in the human mind. Pannenberg's anthropology of religious experience places him between the philosophical atheism of the Left-Hegelians and the Christian supremacy of Barth, stressing the real existence of the divine in all parts of the created world, a world in which humans are intimately and definitively involved.

HISTORY AS REVELATION

Pannenberg's primary contribution to the philosophy of religion has been in his attempts to build on this anthropological position to show the unity of human history with the experience God. Pannenberg's work on this subject is, in the first instance, a reaction to post-Enlightenment critical history. It is also defined in reaction to the rejection of historicism as a category within theology by Barth and, in particular, Bultmann. Pannenberg rejects the belief that historical research, even in areas such as the historical Jesus, do not provide any theological insight. Pannenberg argues that if God is the author of creation, he must be discernible in all parts of creation. Therefore to stress the eschatological and a historic nature of Christ as Bultmann does, is to remove God from the created world that is, by definition historical in form (Pannenberg 1970, p. 87).

The culmination of this work was the publication in 1961 of *Revelation as History*. In this collection of essays, which Pannenberg edited and contributed to, Pannenberg argues that theology, correctly understood, can reconcile the Hegelian understanding of history as the self-disclosure of the *Absolute* with twentieth-century developments in secular critical history. Pannenberg believes he is able to reconcile these two opposing understandings of history by stressing what he believes to be the defining principle of the human history: the desire to comprehend oneself within the created world in which we live. This essentially dialectical understanding of history, Pannenberg argues, underpins the subject areas, method, and approach of secular, critical history. At the most basic level, he argues, the modern secular historian makes judgments about the place and role of actions and events on history. Through this intellectual judgment the historian is implicitly assuming, Pannenberg's argues, that human history has a fundamental source and purpose. Consequently, the narrowly defined terms of critical history always assume, even at the most basic level, the existence of a suprahistorical intellectual structure. No historical person or event can define this structure; this can only be achieved by God who transcends and encompasses all history within himself. Pannenberg therefore believes one can reconcile theology with history if one accepts that they are different methods of understanding the self-disclosure of God within history. Therefore when we engage with the historical world in any way we are, by definition, understanding something of God's revelation to the world.

The Hegelian basis of this argument is clear; however, Pannenberg differs crucially from Hegel in two key

components with his argument. First, looking back to his anthropology, Pannenberg asserts a narrower understanding of human reason than the version of reason we find in Hegel. This allows Pannenberg to retain a greater critical distance between the rational nature of God and ability of human reason to comprehend form and nature of God. Second, Pannenberg argues that although God reveals himself to humankind through the process of history this is, unlike in Hegel, not a necessary, but rather a contingent relationship. This more orthodox understanding of the human faculties and of God's relation to creation allows Pannenberg to reclaim something of the Hegelian understanding of universal history from the Left-Hegelian conflation of the God of universal history into anthropology.

This historicism has, inevitably, created new problems that Pannenberg's thought has not fully answered. Most importantly, Pannenberg's view of the contingent nature of God to human history opens up the problem of how to account for the existence of evil in a divinely ordained world. Pannenberg has countered, and to a limited extent answered this criticism by stressing that one has to understand the positive nature of human endeavor and action before one can understand the perversions. That is, we can only understand why humans turn from God if we first know how we are defined in relationship to God in the first instance (Tupper 1973).

CONCLUSION

The culmination of Pannenberg's intellectual output came with the publication of his three-volume *Systematic Theology* between 1988 and 1993. In this work, which completes the intellectual process begun in his earliest writings, Pannenberg argues that the pursuit of truth, the fundamental object of theology, can only come about within a rigorous and thoroughgoing philosophical framework. Through this framework Pannenberg has argued that it is possible to reconcile scientific research to theology in much the same way as he argues the critical history can be brought into the theological understanding of universal history. By stressing the systematically metaphysical form of theology, Pannenberg argues that theology can save science from intellectual narcissism by providing the overarching structure of truth within which the specific insights of scientific research can be comprehended. Although perhaps not as influential as his writings on theology and history, this engagement with modern science highlights the refreshing willingness, identifiable in all Pannenberg's work, to enter into dialogue with those intellectual disciplines of the post-Enlightenment world that sit outside the traditional corpus of religious and theological thought.

See also Barth, Karl; Bultmann, Rudolf; Duns Scotus, John; Enlightenment; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Hegelianism; Historicism; Philosophy of Religion; Philosophy of Science, History of; Philosophy of Science, Problems of.

Bibliography

WORKS BY PANNENBERG

Jesus—God and Man. Translated by Lewis L. Wilkins and Duane Priebe. London: SCM Press, 1968.

Revelation as History, edited by Wolfhart Pannenberg. Translated by David Granskou. New York: Macmillan, 1968.

What Is Man? Translated by Duane A. Priebe. Philadelphia: Fortress Press, 1970.

Basic Questions in Theology. 3 vols. Translated by George H. Kelm. London: SCM Press, 1970–1973.

Theology and the Philosophy of Science. Translated by Francis McDonagh. London: Darton, Longman and Todd, 1976.

Anthropology in Theological Perspective. Translated by Matthew J. O'Connell. Philadelphia: Westminster, 1985.

Christianity in a Secularized World. Translated by John Bowden. London: SCM, 1989.

Systematic Theology. 3 vols. Translated by Geoffrey W. Bromiley. Edinburgh, U.K.: T and T Clark, 1991–1998.

WORKS ABOUT PANNENBERG

Galloway, Allan D. *Wolfhart Pannenberg*. London: Allen and Unwin, 1973.

Tupper, Frank E. *The Theology of Wolfhart Pannenberg*. Philadelphia: Westminster Press, 1973. Contains a postscript by Wolfhart Pannenberg.

Benjamin Carter (2005)

PANPSYCHISM

"Panpsychism" is the theory according to which all objects in the universe, not only human beings and animals but also plants and even objects we usually classify as "inanimate," have an "inner" or "psychological" being. The German philosopher and psychologist G. T. Fechner wrote:

> I stood once on a hot summer's day beside a pool and contemplated a water-lily which had spread its leaves evenly over the water and with an open blossom was basking in the sunlight. How exceptionally fortunate, thought I, must this lily be which above basks in the sunlight and below is plunged in the water—if only it might be capable of feeling the sun and the bath. And

> why not? I asked myself. It seemed to me that nature surely would not have built a creature so beautiful, and so carefully designed for such conditions, merely to be an object of idle observation. … I was inclined to think that nature had built it thus in order that all the pleasure which can be derived from bathing at once in sunlight and in water might be enjoyed by one creature in the fullest measure. (*Religion of a Scientist,* pp. 176–177)

To many readers this may seem to be merely charming poetry, but Fechner was writing in defense of a philosophical theory for which he argued with great passion and resourcefulness. "Where we see inorganic Nature seemingly dead," wrote the American panpsychist Josiah Royce, "there is, in fact, conscious life, just as surely as there is any Being present in Nature at all" (*The World and the Individual,* second series, p, 240). "All motion of matter in space," in the words of Hermann Lotze, "may be explained as a natural expression of the inner states of beings that seek or avoid one another with a feeling of their need.… The whole of the world of sense … is but the veil of an infinite realm of mental life" (*Microcosmus,* Vol. I, p. 363).

PANPSYCHISM AND RELATED DOCTRINES

Although panpsychism seems incredible to most people at the present time, it has been endorsed in one way or another by many eminent thinkers in antiquity as well as in recent times. Among those who were either outright panpsychists or who inclined to a position of this kind, in addition to Fechner, Royce, and Lotze one may count Thales, Anaximenes, Empedocles, several of the Stoics, Plotinus and Simplicius; numerous Italian and German Renaissance philosophers (including Paracelsus, Girolamo Cardano, Bernardino Telesio, Giordano Bruno, and Tommaso Campanella); G. W. Leibniz, F. W. J. von Schelling, Arthur Schopenhauer, Antonio Rosmini, W. K. Clifford, Harald Høffding, C. B. Renouvier, Eduard von Hartmann, and Wilhelm Wundt; the German freethinkers Ernst Haeckel, Wilhelm Bölsche, and Bruno Wille; C. A. Strong, Erich Adickes, Erich Becher, Alfred Fouillée, C. S. Peirce, and F. C. S. Schiller; and, in the twentieth century, A. N. Whitehead, Samuel Alexander, Bernardino Varisco, Paul Haeberlin, Aloys Wenzel, Charles Hartshorne, and the biologists Pierre Teilhard de Chardin, C. H. Waddington, Sewall Wright, and W. E. Agar.

Few panpsychists, writing in recent years, would make the claim that their position can be proven, but they do assert that the available evidence favors their theory or at the very least enables it to be a serious contender. According to Fechner, it is the best, clearest, most natural, and most beautiful account of the facts of the universe. According to Schiller, who was both a pragmatist and a panpsychist, the doctrine "renders the operation of things more comprehensible" and also enables us to "act upon them more successfully" (*Studies in Humanism,* p. 443). Similarly, Whitehead, after quoting a passage in which Francis Bacon declared his belief that "all bodies whatsoever, though they have no sense … yet have perception," claims that this line of thought "expresses a more fundamental truth than do the materialistic concepts which were then being shaped as adequate for physics" (*Science and the Modern World,* p. 56). Agar, who was a follower of Whitehead's, conceded that there can be "no coercive demonstration" of the truth or falsehood of panpsychism, but it "leads to a more consistent and satisfying world picture than any of the alternatives"; and, unlike these alternatives, panpsychism is not committed to the paradoxical view that "the mental factor … made its appearance out of the blue at some date in the world's history" (*The Theory of the Living Organism,* pp. 109–110).

Modern panpsychists have been quite aware that their theory ran counter to what Fechner's distinguished follower Friedrich Paulsen called "the obstinate dogmatism of popular opinion and of the physical conception of the universe" (*Introduction to Philosophy,* p. 93). This obstinacy they attributed to the prevalence of the "night-view" of the universe—an outlook natural in a mechanized civilization in which people are incapable of noticing and appreciating anything that cannot become the subject of measurement and calculation. In arguing for panpsychism, Fechner and Paulsen (among others) believed that they were counteracting a pernicious tendency in modern life, not merely defending a philosophical viewpoint. Fechner conceived of himself as "awakening a sleeping world" (*Religion of a Scientist,* p. 130) and frequently appealed to his readers to "meet nature with new eyes" (p. 211). Whether plants have souls is not, in the opinion of these writers, an idle or trivial question but on the contrary has a "broader bearing," and its answer decides many other questions and indeed determines one's "whole outlook upon nature" (Fechner, op. cit., p. 163). It is only by accepting panpsychism that a modern man (who finds it impossible to believe in the claims of traditional religion) can escape the distressing implications of materialism.

Unlike Fechner and Paulsen, Lotze supported the traditional religious doctrines of a personal, immaterial deity and a substantival, immortal soul; and hence he did not claim that we had to embrace panpsychism in order to avoid materialism. Lotze also repeatedly insisted, quite unlike Royce and Schiller, that we must not introduce panpsychism into science. Nevertheless he, too, greatly emphasized the emotional benefits accruing from the acceptance of panpsychism. Although science may and should set aside all reference to the "pervading animation of the universe," the "aesthetic view of Nature may lawfully fill out the sum of what exists." If we are panpsychists we no longer "look on one part of the cosmos as but a blind and lifeless instrument for the ends of another," but, on the contrary, find "beneath the unruffled surface of matter, behind the rigid and regular repetitions of its working, … the warmth of a hidden mental activity." Lotze was particularly concerned to vindicate "the fullness of animated life" in such lowly things as "the dust trodden by our feet [and] the prosaic texture of the cloth that forms our clothing." Dust, Lotze declares, is "dust only to him whom it inconveniences," and he asks us to remember that human beings who are "confined" in a low social position, in which the outflow of intellectual energy is greatly impeded, are not by any means deprived of their "high destiny." If in the case of such "oppressed fragments of humanity," of "this dust of the spiritual world," we may yet affirm a divine origin and a celestial goal, then we have far less reason to deny an inner life to physical dust particles; uncomely as these "may appear to us in their accumulations, they at least everywhere and without shortcoming perform the actions permitted to them by the universal order" (*Microcosmus,* Vol. I, pp. 361–363).

HYLOZOISM. Panpsychism is related to but not identical with hylozoism. "Hylozoism" is sometimes defined as the view that matter is "intrinsically" active and in this sense is primarily opposed to the view of philosophers, like Plato and George Berkeley, who asserted that matter is "essentially" inert or passive. More frequently, it refers to the theory that all objects in the universe are in some literal sense alive. Any panpsychist who endorses the usual view that mind implies life would automatically be a hylozoist in the latter sense, but the converse does not hold. In fact most panpsychists have been quite ready to have themselves labeled hylozoists, but there are some exceptions, of whom Schopenhauer is perhaps the most famous. According to Schopenhauer, all objects have an inner nature that he calls "will," but although this will may be described as psychic or mental, it is not necessarily a form of life. "I am the first," Schopenhauer wrote, "who has asserted that a *will* must be attributed to all that is lifeless and inorganic. For, with me, the will is not, as has hitherto been assumed, an accident of cognition and therefore of life; but life itself is manifestation of will" (*On the Will in Nature,* p. 309).

William James is responsible for some terminological confusion that should be cleared up before we go any further. In several of his later writings James strongly supported a theory he stated in the following words: "there is a continuum of cosmic consciousness, against which our several minds plunge as into a mother-sea or reservoir. … we with our lives are like islands in the sea, or like trees in the forest" (*Memories and Studies,* p. 204). Not only psychical research, he held, but also metaphysical philosophy and speculative biology are led in their own ways to look with favor on some such "panpsychist view of the universe as this." Elsewhere he remarks that the evidence from normal and abnormal psychology, from religious experience and from psychical research combine to establish a "formidable probability in favor of a general view of the world almost identical with Fechner's" (*Varieties of Religious Experience,* p. 311). It is true that Fechner held to a theory of a cosmic reservoir of consciousness, regarding God as the universal consciousness in which all lesser souls are contained, but it was not the acceptance of this theory that made him a panpsychist, and James himself was *not* a panpsychist. He nowhere maintained that plants and inanimate objects have an inner psychic life, and it is not easy to see how the reservoir theory by itself logically implies panpsychism.

WORLD SOUL. It should also be pointed out that the theory of the "world soul" is not identical with and does not necessarily follow from panpsychism. A number of panpsychists have in fact maintained the existence of a world soul, and they regarded it as a natural extension of panpsychism. Thus, Fechner in his *Zend-Avesta* (Vol. I, p. 179) concluded that "the earth is a creature … , a unitary whole in form and substance, in purpose and effect … and self-sufficient in its individuality." It is related to our human body as "the whole tree is to a single twig, a permanent body to a perishable, small organ." "Nothing," in the words of Zeno the Stoic (as approvingly quoted by Cicero), that "is destitute itself of life and reason, can generate a being possessed of life and reason; but the world does generate beings possessed of life and reason; the world therefore is not itself destitute of life and reason" (*On the Nature of the Gods,* Bk. II, Sec. VIII). In a very similar vein Paulsen argues that Earth, since it "produces all living and animated beings and harbors them as parts

of its life," may itself be plausibly regarded as "alive and animated." Only the person who is "not open to the inner life of things" will find it difficult to regard Earth as a unitary organism with an inner life as well as a body (*Introduction to Philosophy,* p. 108). To demand to be shown the eyes and ears, the mouth and digestive system, the skin and hair, the arms and legs, the nervous system and the brain of Earth is quite improper. Unlike an animal, Earth does not need a mouth and a stomach because it does not have to take in substances from outside. An animal pursues its prey and in turn attempts to escape its pursuers, and hence it needs eyes and ears, but Earth is not a pursuer and is also not pursued. An animal needs a brain and nerves in order to regulate its movements in response to its environment, but Earth moves around without any such aid. Much like Fechner, Paulsen concludes that "it has regulated its relations to the external world in the most beautiful and becoming manner." "Please do not," he adds, slightly hurt by the irreverent objections of some critics, "please do not ask it to do what is contrary to its nature and cosmical position" (ibid.). This elevated idea of Earth soul has not won general acceptance among panpsychists. Charles Hartshorne, a twentieth-century panpsychist who, like Fechner, is a friend of religion, pays tribute to the "eloquence" of Fechner's account but questions whether "the advances of science since his time have served to confirm" his view. While it may be plausible to regard an electron as "a rudimentary organism," the larger systems that Fechner and Paulsen dealt with so enthusiastically "seem to contemporary knowledge rather too loosely integrated to be accepted as sentient subjects." A tree, it seems plausible to argue, has less unity than one of its own cells, and, similarly, Earth has less unity than the animals which inhabit it ("Panpsychism," p. 447). Hartshorne, as just observed, is a religious thinker, but there have also been atheistic and agnostic panpsychists, and there is no doubt that they would dismiss the theory of the world soul as quite absurd and as an illegitimate extension of panpsychism.

DEGREES OF CONSCIOUSNESS. There is one other terminological confusion against which we should be on guard. Rudolf Eisler, in the article on panpsychism in his *Wörterbuch der Philosophischen Begriffe,* first supplies the definition that we have adopted here and that is the one generally accepted. Later, however, he remarks that many panpsychists merely assert that all matter has a "disposition towards the psychological"—that is, that they ascribe to inorganic things no more than a "hypothetical" or low-grade mentality. Now, panpsychists have indeed generally emphasized that there are degrees of "mentality" or "soul life" and that the mentality or psychic nature of inanimate objects is of an exceedingly simple order, but a low degree or level of mentality must be distinguished from "hypothetical mentality" or the capacity to become the subject of mental activities. To qualify as a panpsychist a person must claim that all bodies actually have an inner or psychological nature or aspect. That all matter is potentially the subject of mental activities or characteristics is something that many other philosophers, including not a few materialists, would concede. To say that a stone is made of elements which, when suitably combined, form an entity that thinks and feels is not the same thing as to say that the stone itself has an inner, psychological being.

Royce is a notable exception to the statement that panpsychists regard the psychic character of inorganic bodies as much lower than that of human beings or animals. He thought that the difference was mainly one of speed and that the "fluent" nature of the inner life of inorganic systems tends to go unnoticed because of its "very vast slowness." To this he added, however, that slowness does not mean "a lower type of consciousness" (*The World and the Individual,* second series, pp. 226–227).

NAIVE AND CRITICAL PANPSYCHISM. Eisler distinguishes between "naive" and "critical" panpsychism—by the former he means the animism of primitive peoples and of children, by the latter he means panpsychist theories that are supported by arguments. In this article we are, of course, concerned exclusively with the "critical" or philosophical variety of panpsychism. Most critical panpsychists would probably endorse Agar's judgment that although primitive animism was "in its analogical way of thinking basically sound," it was also "full of errors" and "ludicrously mistaken in detail" (*The Theory of the Living Organism,* p. 109).

It should be observed that some philosophical panpsychists are not consistently "critical" in the sense just indicated. Thus, while offering elaborate arguments and conceding quite explicitly on numerous occasions that the inner psychic processes of plants and inanimate objects are not given to us in immediate experience but have to be inferred, both Schopenhauer and Fechner occasionally take the opposite position. In a remarkable passage, Schopenhauer tells us that if we consider various inanimate objects "attentively," we shall observe (among many other things) the "strong and unceasing impulse with which the waters hurry to the ocean, [the] persistency with which the magnet turns ever to the North Pole, [the] readiness with which iron flies to the magnet,

[the] eagerness with which the electric poles seek to be reunited, and which, just like human desire, is increased by obstacles [as well as] the choice with which bodies repel and attract each other, combine and separate, when they are set free in a fluid state, and emancipated from the bonds of rigidity." Furthermore, if we attend to the way in which a load "hampers our body by its gravitation towards the earth," we shall "feel directly [that it] unceasingly presses and strains [our body] in pursuit of its one tendency." This passage is taken from the early first volume of *Die Welt als Wille und Vorstellung* (Bk. II, Sec. 23). His later work *Über den Willen in der Natur* consists largely of lists of scientific facts "proving" Schopenhauer's assorted philosophical theories, including his panpsychism. Here we are told to "*look* attentively at a torrent dashing headlong over rocks," whose "boisterous vehemence" can arise only from an "exertion of strength" (p. 308). As for the celestial bodies, if we observe them carefully we shall see that they "play with each other, betray mutual inclination, exchange as it were amorous glances, yet never allow themselves to come into rude contact" (p. 305). Fechner, a milder man than Schopenhauer and more interested in plants than in boisterous torrents or burdensome loads, records experiences in which "the very soul of the plant stood visibly before me," in which he "saw" not only a special "outward clarity" of the flowers but also "the inward light" that in all likelihood caused the outer appearance (op. cit., pp. 211–212).

To see what is at issue between panpsychists and their opponents, it is important to point out that passages such as these are aberrations. It may indeed be held that in addition to the more familiar properties, to which philosophers refer as the primary and secondary qualities, physical objects possess a further set of qualities that are not noticed by observers who lack certain gifts or a suitable training. Such a view need not be mystical and has been plausibly defended in the case of the so-called tertiary qualities, especially those of artistic productions and performances. However, the initial definitions of "soul," "psychic," and "inner," or of any of the other terms used by panpsychists in statements of their position, preclude them from adopting a position of this kind. The "soul," the "inner" nature of an object, its "mental side" is by definition—a definition to which the panpsychists subscribe—something private that only the object itself can experience or observe. Hence, even if one grants that panpsychists possess gifts of which other mortals are deprived, these cannot possibly be the means of directly perceiving the inner qualities or states of any object external to the observer. Moreover, the great majority of panpsychists, including Schopenhauer and Fechner, do not, in their more considered presentations, claim any special faculty for themselves that the opponents of panpsychism supposedly lack. On the contrary, it is implied that, starting from certain generally accessible facts, sound reasoning will lead a person to a panpsychist conclusion.

ARGUMENTS FOR PANPSYCHISM

The arguments for panpsychism may be conveniently grouped according to whether they presuppose the acceptance of a particular metaphysical system or some controversial epistemological theory or whether they are or purport to be of an empirical or inductive character. Some of the arguments of Leibniz and Royce are based on their respective versions of metaphysical idealism, and some of the arguments of Schopenhauer and Paulsen presuppose a Kantian theory of knowledge. It is impossible to evaluate any such arguments without getting involved in an appraisal of their particular metaphysical or epistemological framework, and we shall therefore confine our discussion to arguments of the other kind. It is perhaps worth noting in this connection that, especially since the mid-1800s, many panpsychists have regarded themselves as opponents of metaphysics, or, if they did not object to being labeled metaphysicians, they took care to add that theirs was an "inductive," not a speculative, variety of metaphysics. Fechner in particular prided himself on dispensing altogether with "a priori constructions," and he was a leading figure, along with von Hartmann and Wundt, in a movement to renounce any claim to a special philosophical method distinct from the method employed in the natural sciences. The only method that, on his view, could lead to a tenable theory about the universe as a whole was "generalization by induction and analogy, and the rational combination of the common elements gathered from different areas," as he observes in *Zend-Avesta.* Furthermore, even some of the panpsychists who were also speculative metaphysicians appealed to empirical considerations. They thought that panpsychism could be supported in different ways that were logically independent of one another. Royce was one of the philosophers who adopted this approach. Insisting that his "Idealistic Theory of Being … furnishes a deep warrant" for panpsychism, he nevertheless regarded panpsychism as also resting on "a merely empirical basis" (op. cit., p. 213). "Wholly apart from any more metaphysical consideration of the deeper nature of Reality," certain empirical facts suggest panpsychism as the conclusion of "a rough induction." In this connection, the

theory should be treated as a "hypothesis for further testing" (ibid., pp. 223–224).

GENETIC ARGUMENTS. The arguments that have been most widely urged in defense of panpsychism, and which go back at least as far as Telesio and Campanella, rely, in one way or another, on the assumption that mental facts can be causally explained only in terms of other mental facts. Philosophers who have arrived at a parallelistic answer to the body-mind problem have been specially prone to endorse such arguments, but these can be stated independently of any commitment to parallelism. It is perhaps interesting to note in passing that many early champions of Darwinism (for example, Clifford in England and Haeckel and L. Büchner in Germany) were attracted by reasoning of this kind, although they were frequently repelled by the analogical arguments considered later in the present article. We shall here examine two such genetic arguments—one advanced by Paulsen, the other by a twentieth-century British scientist.

How, asks Paulsen, did soul life originate? Modern biology assumes, quite rightly in Paulsen's opinion, that organic life had a beginning on Earth and that the "first creations" arose from inorganic matter. The question then arises how "psychic life" came into being. "Is the first feeling in the first protoplasmic particle something absolutely new, something that did not exist before in any form, of which not the slightest trace was to be found previously?" (*Introduction to Philosophy,* pp. 99–100). To suppose that the first feeling in the first protoplasmic particle was something "absolutely new" would, however, imply a "creation out of nothing," which would be totally at variance with the basic (and well-founded) principles of science. You might as well, Paulsen remarks, ask the natural scientist "to believe that the protoplasmic particle itself was created out of nothing." The natural scientist rightly assumes that natural bodies arise from preexisting elements. These enter into new and more complicated combinations, and as a result the bodies are capable of performing "new and astonishing functions." Why does the natural scientist "not make the same natural assumption" in the case of the inner psychic processes as well? Why does he not say that "an inner life was already present in germ (*keimhaft*) in the elements, and that it developed into higher forms?"

It is not easy to appraise this line of reasoning because of the vagueness of the expression "absolutely new." As Ernest Nagel and others have pointed out, it is frequently not at all clear whether two processes or occurrences are to be counted as different instances of the same property or as different properties—whether they are or not usually depends on the purpose of the particular investigation. Furthermore, what may be "absolutely new," in the sense of not being predictable from certain initial conditions in conjunction with a certain set of laws, may at the same time not be absolutely new in the sense of being predictable from these initial conditions together with a different set of laws. However, let us assume that in a given case all parties agree that if at a moment T_1 the features of a system were of a certain kind and if at a subsequent moment T_2 they were of a certain different kind, something "absolutely new" came into being at T_2. More specifically, let us assume that the conditions at T_1 do not include any mental fact but that at T_2 they include "the first feeling" in the first protoplasmic particle. Now, according to Paulsen's argument, anybody who supposes that this is the kind of thing that actually happened—and a person who accepts certain scientific facts while rejecting panpsychism has to suppose that this is what happened—is committed to the view that something came from nothing. But to suppose that something came from nothing is unscientific and absurd.

There is a simple answer to this. By saying that something must always come from something and cannot come from nothing, we may mean either (1) that every phenomenon or event has a cause or (2) the scholastic principle that any property residing in an effect must also have been present in its cause. If we suppose that at time T_1 there was no mental fact in the universe while at a later time T_2 the first feeling occurred in a protoplasmic particle, we would indeed be violating proposition (2), but we would not at all be violating proposition (1). Yet if anything can here be regarded as "unscientific" or "absurd" it would be exceptions to (1). For reasons explained earlier, it is not easy to state (2) or its denial with any precision, but, in the most familiar sense of "new," experience seems to show that there are any number of effects possessing new properties—properties not present in the cause. The very course of evolution, to which Paulsen and other proponents of the genetic argument appeal, provides a multitude of illustrations of this. At any rate, an opponent of panpsychism would deny proposition (2) and would insist that such a denial is in no way unempirical or unscientific. To assume the opposite without further ado would surely be to beg one of the basic questions at issue.

Let us now consider a more recent version of a genetic argument: "Something must go on in the simplest inanimate things," writes the distinguished British geneticist C. H. Waddington, "which can be described in the same language as would be used to describe our self-

awareness" (*The Nature of Life,* p. 121). It is true, he continues, that we know nothing of its nature, but the conclusion is forced on us by the "demands of logic and the application of evolutionary theory" (p. 122). Waddington's argument opens with the declaration that the phenomenon of self-awareness is a "basic mystery." This is so because awareness "can never be constructed theoretically out of our present fundamental scientific concepts, since these contain no element which has any similarity in kind with self-consciousness." But self-awareness undoubtedly exists, and hence we must infer that the mode we experience "evolved from simple forms which are experienced by non-human things." It is not difficult to accept this conclusion as far as animals like dogs and cats are concerned. But, Waddington proceeds, we cannot stop there if we take the theory of evolution seriously. According to the initial premise it is inconceivable that self-awareness "originated from anything which did not share something in common with it and possessed only those qualities which can be objectively observed from outside." Hence, we are forced to conclude that "even in the simplest inanimate things there is something which belongs to the same realm of being as self-awareness." Waddington's argument is not overtly based, as Paulsen's was, on the contention that somebody who accepts evolution but rejects panpsychism is committed to the absurd proposition that something comes from nothing. According to Waddington such a person would be committed to the view that self-awareness is not a mystery—that is, that it is explicable in physical terms—and this Waddington takes to be plainly false.

In reply it should be pointed out that Waddington appears to use the word *explanation* in two very different senses in the course of his argument. Sometimes when we ask for the explanation of a phenomenon we are looking for an account of its makeup, of how its parts are related and how they work. We use the word *explanation* in this sense when we want to have the nature of a car or a clock or perhaps a human eye explained to us. At other times, and more frequently, in asking for the explanation of a phenomenon we are looking for its cause. It is not easy to see why awareness should be said to be a "mystery" just because it cannot, in the first sense of "explanation," be explained in physical terms (this betrays a strange materialistic bias that regards a phenomenon as properly explicable, in the first sense, only if it is something material—one wonders why physical objects are not equally mysterious, since they cannot be explained in terms of predicates that are applicable only to mental states). But waiving this point—allowing, that is, that awareness cannot be adequately characterized by the kinds of predicates usually applied to material objects and that this makes awareness incapable of explanation in the first of the two senses distinguished, none of this implies that awareness cannot be explained, in the second sense of the word, in terms of purely physical factors. Avoiding the word *explanation,* the point can be expressed very simply: Granting that awareness is not a physical phenomenon, it does not follow that it cannot be produced by conditions that are purely physical. When the matter is put in this way, it becomes clear that we are back to the difficulty besetting Paulsen's form of the argument. Waddington's argument does not, aside from the acceptance of the evolutionary theory, depend merely on the admission that awareness is not a physical phenomenon, that it "cannot be constructed" out of physical concepts: It also depends on the maxim that any property of the effect must also be present in the cause. We have already mentioned reasons for rejecting this principle, but perhaps it is worth adding that in the context of the body-mind relationship it seems particularly implausible. Brain tumors and other damage to the body, to give some very obvious examples, lead to all kinds of psychological states, but we do not for this reason refuse to regard them as explanations of the latter.

ANALOGICAL ARGUMENTS. The second set of arguments commonly employed by panpsychists, independently of any metaphysical system, purport to be of an analogical kind. Here the more systematic panpsychists usually proceed in two steps: The first consists in arguing that plants are in "essential" respects so much like animals that one cannot consistently attribute a psychic or soul life to animals but refuse it to plants; it is then maintained that the borderline between animate and inanimate objects is not sharp and that a careful examination of inanimate objects reveals them to have many impressive likenesses to animals and plants, indicating the existence of inner psychic being there also.

Plants manifest many of the same vital processes that are found in animals: nutrition, growth, reproduction, and many more. Like animals, plants are born and also die. Moreover, it is simply not true that plants lack the power of spontaneous movement that we observe in animals. "Does not the plant," asks Paulsen, "turn its buds and leaves to the light, does it not send its roots where it finds nourishment, and its tendrils where it finds support? Does it not close up its petals at night or when it rains, and does it not open them in sunshine?" If there is so great a "correspondence" between the visible processes, why should there not be a similar correspondence in "the invisible processes"? (op. cit., pp. 96–97). If it is argued that these analogies are too vague and trifling, because

plants have neither a brain nor a nervous system, the answer is surely that there are animals that also lack brains and nervous systems. Fechner was particularly concerned to exhibit the weakness of this counterargument. He observes that if we remove the strings of a piano or a violin it becomes impossible to obtain any harmonic sounds from these instruments. If somebody concluded from this that the presence of strings is essential to the production of musical tones, he would be completely mistaken, because there are many instruments, like flutes and trombones, with which we can produce musical sounds although they have no strings; but this argument would be not one whit worse than that of the critic of panpsychism.

There are, to be sure, differences between plants and animals, and these a panpsychist has no wish to deny, but, according to Paulsen, they "may be conceived as indicating a difference in inner life also" rather than the absence of any inner processes. The differences indicate "that plants possess a peculiar inclination to receptivity and a decentralized extensity, whereas the psychical life of the animal shows more spontaneity and centralized intensity" (ibid., p. 98). Fechner is even more specific and compares the difference in psychical life between animals and plants to the difference in the psychology of men and women. Elsewhere he compares the former difference to that between the emotions of travelers and those who are "homebodies," between the pleasures associated with "running hither and thither" and those accompanying a "quiet and sedentary sphere of endeavor" (*Religion of a Scientist,* pp. 178–179). Paulsen adds, however, it does not really matter what we think about the details of the inner processes, since all such attempts at conceiving the nature of the psychic life of plants are "at best feeble." It should be remembered that we do not really fare any better if we try to "interpret" the psychical life of animals, especially that of the lower species. We know very little, Paulsen remarks, "about the inner experiences of a jelly-fish or the feelings of a caterpillar or a butterfly."

When we come to inanimate objects, Paulsen continues, the first thing to note is that organic and inorganic bodies must not be regarded as belonging to two separate worlds. There is constant interaction between them. They are composed of the same ingredients and acted on by the same forces. If this were all, however, the analogy would not be strong enough. It would be objected that unlike animals and plants, objects like stones are lifeless and rigid, that they lack all spontaneous activity. This opinion, Paulsen argues, is totally mistaken and is based on the Aristotelian-scholastic theory, taken over by materialistic scientists of the eighteenth and nineteenth centuries, that matter is inherently and absolutely passive. This theory, whether in its original or in its modern atomistic form, is quite untenable. In fact a stone is not an "absolutely dead and rigid body" and devoid of "inner impulses." Modern physics has discarded such a view. Its molecules and atoms are "forms of the greatest inner complexity and mobility." Not only are the constituents of an apparently rigid object like a stone in continuous motion, but the entire system is "in constant interaction with its immediate surroundings as well as with the remotest system of fixed stars" (pp. 101–102). In the light of this it is not only not absurd but quite plausible to conclude that "corresponding to this wonderful play of physical forces and movements" there is a system of inner psychic processes "analogous to that which accompanies the working of the parts in an organic body." We thus arrive, on the basis of scientific evidence, at a view substantially like that of Empedocles that "love and hate form the motive forces in all things"—not, to be sure, quite as we know them in ourselves, but nevertheless in a form that is "at bottom similar" to these human emotions.

It is natural to object to such arguments that the analogies are altogether inconclusive. It is true that there are certain similarities between, say, a stone and a human body, but there are also all kinds of differences. Paulsen assures us that the similarities are "essential," but if "essential" here means that, as far as the inference to an inner psychic process is concerned, the similarities count and the differences do not, that they are relevant whereas the differences are irrelevant, one may well ask how Paulsen knows this. Surely no proposition has been or could have been established to the effect that inner physical movement is always and necessarily connected with psychic activity. Any such general proposition is precisely what the opponent of panpsychism would deny or question. Furthermore, leaving aside any discussion of whether those who regard matter as "active" and those who maintain it to be "passive" are engaging in a factual dispute (so that one party could be said to be right and the other wrong), it must be emphasized that in rejecting panpsychism one is in no way committed to the view that matter is devoid of "inner activity." The view that matter has no inner psychic aspect in no way precludes the admission of inner physical processes such as those postulated by modern physical theory.

These criticisms, however, do not go far enough. They assume, what seems very doubtful, that the arguments under discussion are of a genuinely empirical

character. In this connection it is pertinent to raise the question what the universe would have to be like so that there would be no evidence for panpsychism, or, more strongly, so that the evidence would clearly favor the opposite position. We saw that Paulsen considered the fact that human bodies and inanimate objects are composed of the same elements to be evidence for his position. He also regarded the internal movements of the particles of apparently stationary objects as evidence of their inner life. But suppose that stones and human bodies were not composed of the same elements; would this constitute evidence against panpsychism or would it at least deprive panpsychism of evidence that is at present supporting it? Suppose that electrons were not buzzing inside the stone; would this show or would it be any kind of evidence for the view that the stone does not have a psychic life? From the writings of panpsychists it seems probable that the answer to these questions would be in the negative: If the elements of stones were quite different from those of human bodies, it might be an indication that the psychic processes in stones are even more different in detail from those of human beings, and if the internal constituents of the stones were not in constant motion it might indicate a more restful psychic life, but it would not indicate that no psychic life at all is going on. If this is an accurate presentation of the panpsychist position, it shows that the analogical arguments we have been considering are not genuinely empirical, that the facts pointed to are not, in any accepted sense, evidence for the conclusion. This is a far stronger criticism than the claim that the analogies are weak or the arguments inconclusive.

IS PANPSYCHISM AN INTELLIGIBLE DOCTRINE?

Some contemporary philosophers who have given more thought to the conditions of meaningful discourse than was customary in previous times are inclined to dismiss panpsychism not as false or unproven but as unintelligible. Thus, in his *Philosophical Investigations* Ludwig Wittgenstein raises the question "Could one imagine a stone's having consciousness?" and comments that if anyone can imagine this, it would merely amount to "image-mongery" (Sec. 390, p. 119 e). Such image-mongery, Wittgenstein seems to imply, would not show at all that in attributing consciousness to a stone one is making an intelligible statement. It would probably be pointless to try to "prove" that panpsychism is a meaningless doctrine. Any such attempt is liable to involve one in an elaborate and inconclusive defense of some controversial meaning criterion. However, it may be of some interest to explain more fully, without intending to settle anything, why not a few contemporary philosophers would maintain that the panpsychists do not succeed in asserting any new facts and in the end merely urge certain pictures on us.

To this end let us first consider the following imaginary disputes about the "inner" nature of a tennis ball. *A* holds the common view that the ball is made of rubber and not of living tissue, while *B* holds the unusual opinion that if we were to examine the inside of the tennis ball under a powerful microscope we would find a brain, a nervous system, and other physiological structures usually associated with consciousness. Furthermore, *B* maintains that if we listened very attentively to what goes on while tennis balls are in their can we would hear one ball whispering to the other, "My brother, be careful—don't let them hit you too hard; if you roll into a bush on the other side of the fence you may spend the rest of your days in blissful peace." There is genuine empirical disagreement between *A* and *B* and, as far as we know, *A* would be right if the ball or balls in question are of the familiar kind. Let us next suppose that *C,* after reading Paulsen and Waddington, becomes converted to panpsychism and starts saying such things as "the tennis ball is not a mere body—it has an inner psychic life, it is moved by love and hate, although not love and hate quite as we know them in human beings." To an uncritical outsider it may at first appear, chiefly because of the images one associates with the word *inner,* that *C,* like *B,* is asserting the existence of strange goings on inside the ball, never suspected by the ordinary man or the physicist. In fact, however, if *C* is a philosophical panpsychist, he will not expect to find a brain or a nervous system or any kind of living tissue inside the ball, and he will disclaim any such assertion. Nor will he expect that tennis balls whisper gentle warnings to one another when they are alone. If he should start serving less forcefully in order to avoid hurting the ball, a professional panpsychist would undoubtedly advise him not to be silly, explaining that although their lives are governed by love and hate, balls do not get hurt in any sense that need concern a sympathetic human being. In other words, *C* does not disagree with *A* about what would be found inside the ball or about the ball's behavior while it is in the can, and he is also not treating the ball any differently from the way *A* does—or at any rate no different treatment is logically implied by his opinion that the ball has an inner psychic life. *B* really contradicts *A* and, at least in the case of the balls we all know, he is quite certainly mistaken. *C* is not mistaken, but one begins to wonder whether he is asserting any

facts not allowed for in the ordinary, nonpanpsychist view of the ball. A semantically sensitive observer might comment that ordinary people (and uncritical philosophers) are apt to suppose that they understand well enough what panpsychism asserts and that they proceed to dismiss it as silly or incredible (that is, as plainly false) because they regard panpsychism as a theory like *B*'s unusual opinion about the tennis ball. In fact, panpsychism is not like *B*'s opinion but like *C*'s, and the appropriate criticism seems to be not that it is a false theory but that one does not really know what, if anything, has been asserted.

SCHILLER. Let us now turn to the procedure of an actual panpsychist to see the full relevance of the preceding reflections. F. C. S. Schiller argued that inanimate objects, contrary to the usual opinion, take notice of other inanimate objects, as well as of human beings. "Inanimate objects," he wrote, "are responsive to each other and modify their behavior accordingly. A stone is not indifferent to other stones" (*Logic for Use,* p. 447). Nor are stones indifferent to human beings: "In a very real sense," he wrote elsewhere, "a stone must be said to know us and to respond to our manipulation" (*Studies in Humanism,* p. 443). It is "as true of stones as of men" that if you treat them differently they behave differently (*Logic for Use,* p. 447). It must be emphasized, however, that the responsiveness, the nonindifference, of stones is not quite what we mean when we talk about the responsiveness and nonindifference of human beings. How does a stone exhibit its nonindifference to other stones? Very simply: in being gravitationally attracted to them (ibid.). Nor are we "recognized" by the stone "in our whole nature." It does not "apprehend us as spiritual beings," but this does not mean that the stone takes no note whatever of our existence. "It is aware of us and affected by us on the plane on which its own existence is passed." In the physical world we and stones share, "'awareness' can apparently be shown by being hard and heavy and colored and space-filling, and so forth. And all these things the stone is and recognizes in other bodies" (*Studies in Humanism,* p. 442). The stone "faithfully exercises" all its physical functions: "it gravitates and resists pressure, and obstructs ether vibrations, etc., and makes itself respected as such a body. And it treats us as if of a like nature with itself, on the level of its understanding, i.e., as bodies to which it is attracted inversely as the square of the distance, moderately hard and capable of being hit." The stone does not indeed "know or care" whether a human being gets hurt by it; but in those operations that are of "interest" to the stone, as, for example, in house building, "it plays its part and responds according to the measure of its capacity." What is true of stones, Schiller continues, is also true of atoms and electrons, if they really exist. Just as the stone responds only "after its fashion," so atoms and electrons also know us "after their fashion." They know us not as human beings but "as whirling mazes of atoms and electrons like themselves." We treat stones and atoms as "inanimate" because of "their immense spiritual remoteness from us" and "perhaps" also because of "our inability to understand them" (ibid., pp. 442, 444).

Some of his readers, Schiller realizes, will "cry" that the views just reported amount to "sheer hylozoism," but he does not regard this as any reason for concern. "What," he answers, "if it is hylozoism or, still better, panpsychism, so long as it really brings out a genuine analogy," and this, he is convinced, it does. "The analogy is helpful so long as it really renders the operations of things more comprehensible to us, and interprets facts which had seemed mysterious" (ibid., p. 443). Schiller illustrates his claim by considering the chemical phenomenon of catalytic action. It had "seemed mysterious" and "hard to understand" (presumably prior to the publication of Schiller's "humanistic" panpsychism), that two bodies *A* and *B* may have a strong affinity for each other and yet refuse to combine until the merest trace of a third substance *C* is introduced, which sets up an interaction between *A* and *B* without producing an alteration in *C* itself. But, asks Schiller, "is not this strangely suggestive of the idea that *A* and *B* did not know each other until they were introduced by *C,* and then liked each other so well that *C* was left out in the cold?" To this he adds—and here surely not even the most hostile critic would disagree—that "more such analogies and possibilities will probably be found if they are looked for." Nevertheless, panpsychism does not merely render the operation of things more comprehensible. It has a further virtue, to which Schiller alludes later in the same discussion: "The alien world which seemed so remote and so rigid to an inert contemplation, the reality which seemed so intractable to an aimless and fruitless speculation, grows plastic in this way to our intelligent manipulations" (ibid., p. 444).

Perhaps the most striking features of Schiller's presentation are the constant modifications or retractions of what at first appear truly remarkable assertions. Inanimate objects are "responsive to each other," but not the way in which human beings or animals are—they are responsive in being gravitationally attracted by other inanimate objects. The stone is "aware of us," but not, of course, in the sense in which human beings are aware—it is aware on "*its* plane"; the stone "recognizes" other bod-

ies and is "interested" in operations such as house building, but "on the level of *its* understanding"; it "plays its part," but "according to the measure of *its* capacity"; atoms and electrons know us no less than we know them, but "after *their* fashion." It is not, perhaps, unfair to say that Schiller takes away with one hand what he gives with the other, and it may be questioned whether anything remains. When one is told that the stone is aware of us one reacts with astonishment and is apt to suppose that a statement has been made that contradicts what an ordinary nonpanpsychist believes; but this turns out to be more than doubtful since the stone's awareness, on its plane, seems to consist simply in being hard, heavy, space-filling, and colored. The stone makes itself respected and is interested in operations like house building, but in its own fashion, and this consists in gravitating, resisting pressure, and all the usual characteristics of stones, which are not questioned by those who do not subscribe to panpsychism. Schiller plainly believed that the panpsychist asserts (if he has not in fact discovered) facts about stones and atoms that are denied by, or whose existence is unknown to, the ordinary person and the materialist. He evidently did not believe that it was just a question of using words in different senses. But, if so, what are the facts he asserts and his opponents deny? Schiller's qualifications remind one of a song in the musical *Kiss Me, Kate* in which a lighthearted lady sings of her numerous and constantly changing amorous involvements, adding at the end of each verse, "But I'm always true to you, darling, in my fashion; yes, I'm always true to you, darling, in my way." How does the stone's awareness in its own way differ from what other people would refer to as absence of awareness?

EMPIRICAL PRETENSIONS OF PANPSYCHISTS. Even if one is disinclined to go so far as to dismiss panpsychism as meaningless, there is surely good reason to dispute the empirical and pragmatic pretensions of certain panpsychists. We saw that Royce regarded panpsychism (among other things) as a hypothesis "to be tested," but unfortunately he did not tell us anything about the way or ways in which this was to be done. Royce did indeed guard himself by maintaining that the mental processes in physical systems occur over "extremely august" temporal spans (*The World and the Individual,* second series, p. 226), so that a human being would be unable to detect a process of this kind. However, making the fullest allowance for this qualification and granting ourselves or some imaginary observer the "august" time span required by Royce's "hypothesis," this would still not do, since Royce omitted to inform us what such an observer should look for.

Schiller, it will be remembered, assured us that as a result of accepting panpsychism the previously "remote" and "rigid" reality "grows plastic … to our manipulations." But he did not explain how and where these happy transformations would take place. Is a bricklayer who has been converted to panpsychism going to lay bricks more efficiently? Does a tennis player's game improve if he becomes a disciple of Schiller? No, but perhaps the chemist will find catalytic action more comprehensible, and "more such analogies and possibilities" will make other "intractable" processes less "mysterious." Regrettably, the opinion that panpsychism makes any of these phenomena easier to understand is the result of a confusion that hinges on an ambiguity in "comprehensible" and related expressions. Sometimes we attempt to make phenomena or correlations of events more comprehensible. In this sense, a phenomenon (for example, a certain disease or a plane crash) is comprehended or understood if its cause is discovered, and a correlation or a law becomes comprehensible if it is subsumed under a wider law (if, for example, the administration of a certain drug has in many cases been followed by the cure of a given condition, the correlation becomes comprehensible if we determine what it is about the drug that has this effect; and this is another way of saying that we subsume the correlation under a law). But at other times when we talk about making something comprehensible, we are concerned with explaining the meaning of theories or statements, not with the explanation of phenomena or of correlations. Unlike the first, this kind of problem may be regarded as pedagogical, and here all kinds of analogies may be helpful that do not or need not shed any light on the causes of the phenomena dealt with in the statements we are trying to make more comprehensible. It cannot, of course, be denied that an analogy such as the one Schiller offers may well make catalysts more comprehensible in this pedagogical sense—it may, for example, help schoolchildren to understand what a chemist is talking about. It is equally clear that such an analogy does absolutely nothing to make catalytic action more comprehensible in the earlier sense we mentioned, and it was surely in this sense that Schiller claimed panpsychism to make things less mysterious and easier to understand. It is difficult to believe that either Schiller or any other champion of panpsychism would be satisfied to have the theory regarded as no more than a pedagogical device in the teaching of natural science.

See also Alexander, Samuel; Anaximenes; Berkeley, George; Bruno, Giordano; Campanella, Tommaso; Cicero, Marcus Tullius; Clifford, William Kingdon; Empedocles; Fechner, Gustav Theodor; Fouillée, Alfred; Haeckel, Ernst Heinrich; Hartmann, Eduard von; Høffding, Harald; James, William; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Macrocosm and Microcosm; Materialism; Nagel, Ernest; Pantheism; Paracelsus; Paulsen, Friedrich; Peirce, Charles Sanders; Plato; Plotinus; Renouvier, Charles Bernard; Rosmini-Serbati, Antonio; Royce, Josiah; Schelling, Friedrich Wilhelm Joseph von; Schiller, Ferdinand Canning Scott; Schopenhauer, Arthur; Simplicius; Teilhard de Chardin, Pierre; Telesio, Bernardino; Thales of Miletus; Varisco, Bernardino; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann; Wundt, Wilhelm.

Bibliography

The fullest systematic defenses of panpsychism since the mid-1800s are found in the writings of Paulsen, Fechner, Lotze, and Royce. Paulsen's arguments are presented in his very influential *Einleitung in die Philosophie* (21st ed., Stuttgart and Berlin, 1909), translated by F. Thilly as *Introduction to Philosophy* (2nd American ed., New York, 1906, with a preface by William James). Fechner's main writings on the subject are *Nanna: oder über das Seelenleben der Pflanzen* (3rd ed., Leipzig, 1903) and *Zend-Avesta: oder über die Dinge des Jenseits* (2nd ed., Hamburg: L. Voss, 1906). There is an English translation of selections from Fechner's works by W. Lowrie titled *Religion of a Scientist* (New York: Pantheon, 1946). Fechner's ideas are discussed in some detail in G. Stanley Hall, *Founders of Modern Psychology* (New York: Appleton, 1912); G. F. Stout, *God and Nature,* edited by A. K. Stout (Cambridge, U.K., 1952); Otto Külpe, *Die Philosophie der Gegenwart in Deutschland* (Leipzig: Teubner, 1902), translated by M. L. Patrick and G. T. W. Patrick as *Philosophy of the Present in Germany* (London: G. Allen, 1913); and G. Murphy, "A Brief Interpretation of Fechner," in *Psyche* 7 (1926): 75–80. Although Wilhelm Wundt condemned Fechner's speculations about the souls of the stars and Earth as "a fantastic dream," he himself concluded that mental life can arise only out of conditions that are themselves mental (*System der Philosophie,* Leipzig, 1889). Lotze's defense of panpsychism is contained in Vol. I of *Mikrokosmus* (Leipzig, 1856–1864), translated by E. Hamilton and E. E. C. Jones as *Microcosmus* (New York, 1890). Royce's panpsychism is presented in Lecture V of *The World and the Individual,* second series (New York: Macmillan, 1901). The American neorealist W. P. Montague, a student of Royce, relates how he "jumped with almost tearful gratitude" at Royce's "hypothesis about the varying time-spans in nature." He regarded this "hypothesis" as "a new and challenging contribution to the great panpsychist tradition," as "a clear and great thought" that "might even be true" (*The Ways of Things,* London, 1940, p. 669). Montague referred to his own position as "animistic materialism," and he is sometimes classified as a panpsychist, but in fact it is very doubtful whether his animism implies panpsychism as we have here defined it.

Little was said in this article about A. N. Whitehead, probably the most distinguished champion of panpsychism in the twentieth century, chiefly because his views on the subject could not have been discussed without consideration of other features of his difficult system. Whitehead would have disagreed with many other panpsychists about the "units" that are to be regarded as the bearers of psychic life. These, he held, are not stars or stones but the events out of which stars and stones are constituted and that Whitehead calls "occasions." His views are presented in *Science and the Modern World* (New York: Macmillan, 1925), *Process and Reality* (New York: Macmillan, 1929), and, most fully, in "Nature Alive," Lecture 8 of *Modes of Thought* (New York: Macmillan, 1938). Panpsychistic views strongly influenced by Whitehead are put forward in Charles Hartshorne, *Beyond Humanism* (Chicago: Willett Clark, 1937) and *Man's Vision of God* (Chicago: L Willett Clark, 1941), and in W. E. Agar, *The Theory of the Living Organism* (Melbourne, 1943). Samuel Alexander, whose metaphysical position has many similarities to Whitehead's, also expresses views akin to panpsychism in his British Academy lecture "The Basis of Realism," reprinted in *Realism and the Background of Phenomenology* edited by R. M. Chisholm (Glencoe, IL: Free Press, 1960).

Of works by earlier panpsychists, special mention should be made of G. W. Leibniz, *Monadology* (various editions), and Arthur Schopenhauer, *Die Welt als Wille und Vorstellung,* 3 vols. (Leipzig, 1818), translated by R. B. Haldane and J. Kemp as *The World as Will and Idea* (London: Trubner, 1883), as well as his *Über den Willen in der Natur* (Frankfurt, 1836), translated by K. Hillebrand as *On the Will in Nature* (London, 1889).

Giordano Bruno's panpsychist views are presented in the second dialogue of *De la causa, Principio e uno*; for translations see Sidney Greenberg's *The Infinite in Giordano Bruno* (New York: King's Crown Press, 1950) and Jack Lindsay's version in *Cause, Principle and Unity* (New York: International, 1964). The works by Telesio and Campanella in which their panpsychism is expounded are not available in English. There is a very clear summary of their arguments in Harald Høffding, *A History of Modern Philosophy,* Vol. I (London, 1908). The texts of the pre-Socratics, some of whom were hylozoists rather than panpsychists, are available in English translation in G. S. Kirk and J. E. Raven, *The Presocratic Philosophers* (Cambridge, U.K.: Cambridge University Press, 1957). Because of his remarks about the "plastic nature in the universe" in *The True Intellectual System of the Universe,* Ralph Cudworth is described as a panpsychist in various reference works, but it is doubtful that this classification is accurate. Cudworth appears to have postulated the "plastic nature" for living things only and he should be labeled a "vitalist" in a sense in which this theory does not automatically imply panpsychism. C. B. Renouvier's panpsychism, which is in many ways similar to that of Leibniz, is expounded in several of his works, most fully in *La nouvelle monadologie* (Paris: A. Colin, 1899). Eduard von Hartmann advocates the view that even atoms possess an unconscious will in *Grundriss der Naturphilosophie,* Vol. II of *System der Philosophie im Grundriss* (Bad Sachsa im Harz, 1907). Benedict de Spinoza

and Henri Bergson were not listed as panpsychists in the text because there is some doubt as to how some of their remarks are to be interpreted. In Spinoza's case there is at least one passage (*Ethics,* Pt. II, Note 2, Prop. XIII) supporting such a classification. Similarly, some of the remarks in "Summary and Conclusions," in Bergson's *Matter and Memory* (London: Allen and Unwin, 1910), may be construed as an endorsement of panpsychism.

C. H. Waddington's genetic argument is presented in *The Nature of Life* (London: Allen and Unwin, 1961). W. K. Clifford advocates very similar arguments in his essays "Body and Mind" and "On the Nature of Things-in-Themselves," in *Lectures and Essays,* Vol. II (London, 1903). The American critical realist C. A. Strong also employs genetic arguments in support of panpsychism in *The Origin of Consciousness* (London: Macmillan, 1918). Sewall Wright, a distinguished contemporary biologist, defends panpsychism on scientific grounds in "Gene and Organism," in the *American Naturalist* 87 (1953). Hackel's views are found in *Natürliche Schöpfungsgeschichte* (4th ed., Berlin, 1892), translated by E. Ray Lankester as *The History of Creation* (London, 1892), and in *Zellseelen und Seelenzellen* (Leipzig, 1909). Panpsychism is also defended on the basis of an appeal to continuity in nature in Harald Høffding, *Outlines of Psychology* (London, 1919). Høffding, however, is rather more diffident than the other writers mentioned in this paragraph. Schiller's defenses of panpsychism are contained in his *Studies in Humanism* (London: Macmillan, 1907) and *Logic for Use* (London: G. Bell, 1929). There is a full discussion of William James's views on panpsychism and various related theories in W. T. Bush, "William James and Panpsychism," in *Columbia University Studies in the History of Ideas,* Vol. II (New York, 1925).

A defense of the scholastic doctrine that an effect cannot possess any perfection which is not found in its cause is contained in G. H. Joyce, *Principles of Natural Theology* (London: Longmans, Green, 1923), Ch. 3. The question of what may be meant by the claim that an effect contains a "new" property is discussed in Arthur O. Lovejoy, "The Meanings of 'Emergence' and Its Modes," in *Proceedings of the Sixth International Congress of Philosophy* (New York, 1927); Ernest Nagel, *The Structure of Science* (New York: Harcourt Brace, 1961); and Arthur Pap, *An Introduction to the Philosophy of Science* (New York: Free Press of Glencoe, 1962). Certain contemporary arguments about the alleged causal inexplicability of human actions, similar to the genetic arguments by Paulsen and Waddington, are examined in Bernard Berofsky, "Determinism and the Concept of a Person," in *Journal of Philosophy* 61 (1964): 461–475.

General surveys of panpsychism are found in A. Rau, *Der moderne Panpsychismus* (Berlin, 1901), and Charles Hartshorne, "Panpsychism," in *A History of Philosophical Systems,* edited by V. T. A. Ferm (New York: Philosophical Library, 1950). Almost all extended discussions of panpsychism occur in the works of writers who accept the theory or who are at least sympathetic to it. One of the few highly critical discussions is contained in Alois Riehl, *Zur Einführung in die Philosophie der Gegenwart* (Leipzig: Teubner, 1903). Eisler's article on panpsychism in his *Wörterbuch der Philosophische Begriffe* (4th ed., Berlin: E. S. Mittler, 1929) contains a very elaborate list of panpsychists and their writings.

Paul Edwards (1967)

PANTHEISM

"Pantheism" is a doctrine that usually occurs in a religious and philosophical context in which there are already tolerably clear conceptions of God and of the universe and the question has arisen how these two conceptions are related. It is, of course, easy to read pantheistic doctrines back into unsophisticated texts in which the concept of the divine remains unclarified, but it is wise to be skeptical about the value of such a reading. Some commentators have confidently ascribed pantheistic views to the Eleatics simply because they assert that what is, is one. But even if one considers Xenophanes, the most plausible candidate for such an ascription, it is clear that considerable care must be exercised. Thales and Anaximenes had some idea of objects in the world being infused with a divine power or substance that conferred life and movement. Xenophanes took over this idea and added to it a critique of Homeric and Hesiodic polytheism, attacking both their anthropomorphism and the immorality in which they involved the gods; his own consequent view of deity remains mysterious, however. Aristotle said that Xenophanes "with his eye on the whole world said that the One was god," but he also complained that Xenophanes "made nothing clear." It seems likely that Xenophanes, like other early Greek thinkers, did not distinguish clearly between asserting that an object was divine and asserting that a divine power informed the object's movement.

A failure by commentators themselves to observe this distinction makes it misleadingly easy to present both earlier pre-Socratic and later Stoic philosophers as recruits to the ranks of pantheism. But even Marcus Aurelius, the only notable thinker among them who can plausibly be represented as a pantheist, when he addressed the Universe itself as a deity did not clearly address it in the sense of all that is rather than in the sense of some principle of order that informs all that is.

VEDIC PANTHEISM

As in Greek thought, the approach to pantheism in Indian thought is a systematic critique of polytheism. Although there are also conceptions of a god who reigns as the highest deity—Indra at one time held this position—what emerged with the growth of theological

reflection was the notion of Brahman. Brahman is the single, infinite reality, indefinable and unchanging, behind the illusory changing world of perceived material objects. The equation of plurality and change with imperfection is an assumption of the Vedanta teachings. From it there is drawn a proof of the illusory character of the material world, as well as of its imperfection. Were the material world real, it must, being neither self-existent nor eternal, have originated from Brahman. But if Brahman were such that from within it what is multifarious, changing, and therefore imperfect could arise, then Brahman would be imperfect. And what is imperfect cannot be Brahman.

We take the illusory for the real because our knowledge is itself tainted with imperfections. Our ordinary knowledge is such that knower and known, subject and object, are distinct. But to know Brahman would be for subject and object to become identical; it would be to attain a knowledge in which all distinctions were abolished and in which what is known would therefore be inexpressible. Two features of the pantheism of the Vedanta scholars deserve comment. The first is the affinity between their logical doctrines and those of F. H. Bradley, whose treatment of the realm of appearance is precisely parallel to the Vedanta treatment of the realm of illusion (*māyā*); Bradley's Absolute resembles Brahman chiefly in that both must be characterized negatively. As with Bradley's doctrine, the natural objection to Vedanta pantheism is to ask how, if Brahman is perfect and unchangeable, even the illusions of finitude, multiplicity, and change can have arisen. The Vedanta doctrine's answer is circular: Ignorance (lack of enlightenment) creates illusion. But it is, of course, illusion that fosters the many forms of ignorance.

Yet if the explanation of illusion is unsatisfactory, at least the cure for it is clear; the Vedanta doctrine is above all practical in its intentions. It will be noteworthy in the discussion of other and later pantheisms how often pantheism is linked to doctrines of mystical and contemplative practice. The separateness of the divine and the human, upon which monotheists insist, raises sharply the problem of how man can ever attain true unity with the divine. Those contemplative and mystical experiences, common to many religions, for whose description the language of a union between human and divine seems peculiarly appropriate—at least to those who have enjoyed these experiences—for that very reason create problems for a monotheistic theology, problems that have often been partly resolved by an approach to pantheistic formulations. It is at least plausible to argue that the essence of the Vedanta doctrine lies in its elucidation of mystical experience rather than in any use of metaphysical argument for purely intellectual ends.

WESTERN PANTHEISM TO SPINOZA

The pantheism of the Vedanta argues that because God is All and One, what is many is therefore illusory and unreal. The characteristic pantheism and near pantheism of the European Middle Ages proceeded, by contrast, from the view that because God alone truly is, all that is must in some sense be God, or at least a manifestation of God. Insofar as this view implies a notion of true being at the top of a scale of degrees of being, its ancestry is Platonic or Neoplatonic. It would be difficult to call Neoplatonism itself pantheistic because although it views the material world as an emanation from the divine, the fallen and radically imperfect and undivine character of that world is always emphasized.

ERIGENA AND AVERROES. However, the translation of Neoplatonic themes of emanation into Christian terms by John Scotus Erigena (c. 810–c. 877) resulted in *De Divisione Naturae,* which was condemned as heretical precisely because of its break with monotheism. It might be argued that Erigena does not seem to be wholly pantheistic in that he did not treat every aspect of nature as part of the divine in the same way and to the same degree. This would be misleading, however, for on this criterion no thinker could ever be judged a pantheist.

According to Erigena the whole, *natura,* is composed of four species of being: that which creates and is not created, that which is created and creates, that which is created and does not create, and that which is not created and does not create. The first is God as creator; the last, God as that into which all created beings have returned. The second and third are the created universe, which is in process of passing from God in his first form to God in his last form. Erigena wrote as if each class of beings belongs to a different period in a historical unfolding, but he also treated this as a misleading but necessary form of expression. *Natura* is eternal; the whole process is eternally present; and everything is a *theophania,* a manifestation of God.

Pope Honorius III condemned *De Divisione Naturae* in 1225 as "pullulating with worms of heretical perversity," and much earlier Erigena's other work had been described by the Council of Valence (855) as "Irish porridge" and "the devil's invention." Clearly, part of what perturbed them was Erigena's ability to interpret in a pantheistic sense both the biblical doctrine of creation

and the biblical notion of a time when God shall be all in all.

A similar problem arose for the Islamic interpreter of Aristotle, Ibn Rushd (Averroes), whose discussions of the relation of human to divine intelligence aroused suspicion of pantheism and whose assertions of fidelity to the Qur'an did not save him from condemnation. A Christian Aristotelian such as Meister Eckhart, the Dominican mystic, was also condemned. Both Eckhart and Johannes Tauler spoke of God and man in terms of a mutual dependence that implies a fundamental unity including both. However, in every medieval case after Erigena the imputation of pantheism is at best inconclusive. Only since the sixteenth century has genuine pantheism become a recurrent European phenomena.

BRUNO. Giordano Bruno (1548–1600) was an explicitly anti-Christian pantheist. He conceived of God as the immanent cause or goal of nature, distinct from each finite particular only because he includes them all within his own being. The divine life that informs everything also informs the human mind and soul, and the soul is immortal because it is part of the divine. Since God is not distinct from the world, he can have no particular providential intentions. Since all events are equally ruled by divine law, miracles cannot occur. Whatever happens, happens in accordance with law, and our freedom consists in identifying ourselves with the course of things. The Bible, according to Bruno, insofar as it errs on these points, is simply false.

BOEHME. Jakob Boehme (1575–1624) was a shoemaker, a mystic, and a Lutheran whose wish to remain within the church was shown by the fact that to the end he received the sacraments. The pantheism of Erigena or Bruno was founded upon a view that the universe must necessarily be a single all-inclusive system if it is to be intelligible. Their pantheism derived from their ideal of explanation. Boehme, by contrast, claimed that he was merely recording what he has learned from an inward mystical illumination. He saw the foundation of all things in the divine *Ungrund,* in which the triad of Everything, Nothing, and the Divine Agony that results from their encounter produces out of itself a procession of less ultimate triads which constitute the natural and human world. Boehme made no distinction between nature and spirit, for he saw nature as entirely the manifestation of spirit. It is not at all clear in what sense the propositions that Boehme advanced can have been the record of vision; it is clear that both in claiming authority for his vision and in the content of his doctrine he was bound to encounter, as he did, the condemnation of the Lutheran clergy.

SPINOZA. Benedict de Spinoza's pantheism had at least three sources: his ideal of human felicity, his concept of explanation, and his notion of the degrees of human knowledge. His explicit aim was to discover a good that would be independent of all the ordinary contingencies of chance and misfortune. Only that which is capable of completely filling and occupying the mind can be the supreme good in Spinoza's sense. The only knowledge that could satisfy these requirements would be the knowledge that the mind is part of the total system of nature and is at one with it when recognizing that everything is as it must be. Felicity is the knowledge of necessity, for if the mind can accept the necessity of its own place in the whole ordering of things, there will be room neither for rebellion nor for complaint. Thus, from the outset Spinoza's characterization of the supreme good required that his philosophy exhibit the whole universe as a single connected system.

So it is with his concept of explanation. To explain anything is to demonstrate that it cannot be other than it is. To demonstrate this entails laying bare the place of what is to be explained within a total system. Spinoza made no distinction between contingent causal connections and necessary logical connections. A deductive system in which every proposition follows from a set of initial axioms, postulates, and definitions mirrors the structure of the universe, in which every finite mode of existence exemplifies the pattern of order that derives from the single substance, *Deus, sive natura* (God, or nature). There can be only one substance, not a multiplicity of substances, for Spinoza so defined the notion of substance that the relation of a property to the substance of which it is a property is necessary, and therefore intelligible and explicable; however, the relation of one substance to another must be external and contingent, and therefore unintelligible and inexplicable. But for Spinoza it is unintelligible that what is unintelligible should be thought to exist. Hence, there can be only one substance; "God" and "Nature" could not be the names of two distinct and independent substances.

It follows that God cannot be said to be the creator of nature, except in a sense quite other than that of Christian or Jewish orthodoxy. Spinoza did distinguish between nature as active (*natura naturans*) and nature as passive product (*natura naturata*), and insofar as he identified God with nature as creative and self-sustaining rather than with nature as passive, he could speak of God as the

immanent cause of the world. But this is quite different from the orthodox conception of divine efficient causality. Also, in Spinoza's view, there can be no divine providential intentions for particular agents and there can be no miracles. What, then, of the Bible?

Spinoza regarded the Bible as an expression of truth in the only mode in which the ordinary, unreflective, irrational man is able to believe it or be guided by it. Such men need images, for their knowledge is of the confused kind that does not rise to the rational and scientific explanation of phenomena, let alone to that *scientia intuitiva* (intuitive knowledge) by which the mind grasps the whole necessity of things and becomes identical with the *infinita idea Dei* (infinite idea of God). Freed from all those passions that dominated his actions so long as he did not grasp them intellectually, man is moved only by a fully conscious awareness of his place in the whole system. It is this awareness that Spinoza also identified as the intellectual love of God.

In using theological language to characterize both nature and the good of human life, Spinoza was not concealing an ultimately materialistic and atheistic standpoint. He believed that all the key predicates by which divinity is ascribed apply to the entire system of things, for it is infinite, at once the uncaused *causa sui* and *causa omnium* (cause of itself and cause of everything) and eternal. Even if Spinoza's attitude to the Bible was that it veils the truth, he believed that it *is* the truth that it veils. He considered his doctrine basically identical with both that of the ancient Hebrew writers and that of St. Paul. This did not save him from condemnation by the synagogue in his lifetime, let alone from condemnation by the church afterward.

GERMAN PANTHEISM

Erigena, Bruno, Boehme, Spinoza—each of these, no matter how much he may have made use of material drawn from earlier philosophical or religious writing, was a thinker who was independent of his specifically pantheist predecessors and who revived pantheism by his own critical reflections upon monotheism. It was only in the eighteenth century that something like a specifically pantheist tradition emerged. The word *pantheist* was first used in 1705 by John Toland in his *Socinianism truly stated.* Toland's hostile critic, J. Fay, used the word *pantheism* in 1709 and it speedily became common. With the increased questioning of Christianity, accompanied by an unwillingness to adopt atheistic positions, pantheism became an important doctrine, first for Johann Wolfgang von Goethe and Gotthold Ephraim Lessing, both of whom were influenced by Spinoza, then for Friedrich Schleiermacher, and finally for Johann Gottlieb Fichte, Friedrich von Schelling, and Georg Wilhelm Friedrich Hegel.

GOETHE AND LESSING. Goethe's aim was to discover a mode of theological thinking, rather than a theology, with which he could embrace both what he took to be the pagan attitude to nature and the redemptive values of Christianity. Suspicious as he was of Christian asceticism, he also recognized a distinctive Christian understanding of human possibility, and his various utterances about Christianity cannot be rendered consistent even by the greatest scholarly ingenuity. In the formulas of pantheism, which he was able to interpret in the sense that he wished precisely because he failed to understand Spinoza correctly, Goethe found a theology that enabled him both to identify the divine with the natural and to separate them. The infinite creativity Goethe ascribed to nature is what he took to be divine; but while the seeds of a consistent doctrine can be discerned in this aspect of Goethe's writings, it would be wrong to deny that part of pantheism's attraction for him was that it seemed to license his will to be inconsistent.

Lessing, by contrast, was consistent. He found the kernel of truth in all religions in a neutral version of Spinozism, which allowed him to see Judaism, Christianity, and Islam as distorted versions of the same truth, distorted because they confuse the historical trappings with the metaphysical essence.

SCHLEIERMACHER. Schleiermacher's quite different preoccupation was to make religion acceptable to the cultured unbelievers of his own time. The core of religion, on his view, is the sense of absolute dependence; to that on which we are absolutely dependent he gave a variety of names and titles, speaking of God in both monotheistic and pantheistic terms. However, he committed himself to pantheism by asserting that it is the Totality that is divine.

FICHTE. It is clear from Goethe, Lessing, and Schleiermacher that Spinoza's writing had become a major text for philosophical theology, but for these writers he was an inspiration rather than a precise source. With the advent of German idealism, the attempt to criticize the deductive form of Spinoza's reasoning while preserving the pantheistic content became a major theme of German philosophy. Nowhere is this more evident than in Fichte's writing, in which God and the universe are identified because the world is nothing but the material through which the Ego realizes its infinite moral vocation, and the

divine is nothing but the moral order that includes both world and Ego. The divine cannot be personal and cannot have been the external creator of the world. Fichte poured scorn on the unintelligibility of the orthodox doctrine of creation ex nihilo (out of nothing). He distinguished sharply between the genuinely metaphysical and the merely historical elements in Christianity. It is the theology of the Johannine Gospel that he treated as the expression of the metaphysical, and to this he gave a pantheistic sense.

SCHELLING AND HEGEL. Schelling's pantheism was cruder than Fichte's—according to him, all distinctions disappear in the ultimate nature of things. The divine is identified with this ultimate distinctionless merging of nature and spirit, a unity more fundamental than any of the differences of the merely empirical world.

Hegel was subtler and more philosophically interesting than either Fichte or Schelling. Like Boehme and Schleiermacher, he remained within orthodox Protestant Christianity, claiming to be engaged in the interpretation rather than the revision of its dogmatic formulas. The Hegelian Absolute Idea preexists its finite manifestations logically but not temporally, and it receives its full embodiment only at the end of history, when it is incorporated in a social and moral order fully conscious of its own nature and of its place in history. This phase of self-consciousness is already reached at the level of thought in Hegel's *Logic.* But the Absolute Idea has no existence apart from or over and above its actual and possible manifestations in nature and history. Hence, the divine is the Totality.

After Hegel pantheism was less in vogue. The critique of Christianity became more radical, atheism became a more acceptable alternative, and Spinoza dominated the intellectual scene far less. In England a poetic pantheism appeared in Percy Bysshe Shelley and William Wordsworth, but in Shelley it coexisted with something much closer to atheism and in Wordsworth with a Christianity that displaced it. In any case, the intellectual resources of such a pantheism were so meager that it is not surprising that it did not survive in the nineteenth century.

CRITICISMS OF PANTHEISM

Pantheism essentially involves two assertions: that everything that exists constitutes a unity and that this all-inclusive unity is divine. What could be meant by the assertion that everything that exists constitutes a unity? It is first and most clearly not a unity derived from membership of the same class, the view that seems to have been taken by Boehme. "There is no class of all that is," wrote Aristotle. Why not? Because existence is not a genus. To say that something exists is not to classify it at all. When Boehme asserted that the universe includes both existence and nonexistence, he both anticipated a long tradition that culminated in Martin Heidegger and remained unintelligible. The notion of *a* unity that includes all that exists—or even all that exists and all that does not exist—is a notion devoid of content. What could be unitary in such an ostensible collection?

The unity might be of another kind, however. In Spinoza the unity of the universe is a logical unity, with every particular item deducible from the general nature of things. There is a single deductive web of explanation—there are not sciences; there is science. About such an alleged unity two points must be made. First, the contingent aspect of nature is entirely omitted. Even a total description of the universe in which every part of the description was logically related to some other part or parts (assuming for the moment such a description to be conceivable) would still leave us with the question whether the universe was as it was described; and if it was as it was described, this truth would be a contingent truth that could not be included in the description itself and that could stand in no internal conceptual relationship to the description. The fact of existence would remain irreducibly contingent. Second, the actual development of the sciences does not accord with Spinoza's ideal. The forms of explanation are not all the same; the logical structure of Darwinian evolutionary theory must be distinguished from the logical structure of quantum mechanics. Thus, the kind of unity ascribed by Spinoza to the universe seems to be lacking.

In Fichte and Hegel the unity ascribed to the universe is one of an overall purpose manifest in the pattern of events, as that pattern is discovered by the agent in his social and moral life. In order for this assertion to be meaningful it must be construed, at least in part, in empirical terms; in Fichte's case as a hypothesis about moral development, in Hegel's case as a hypothesis about historical development. Neither hypothesis appears to be vindicated by the facts.

Suppose, however, that a unity of some kind, inclusive of all that is, could be discovered. In virtue of what might the pantheist claim that it was divine? The infinity and the eternity of the universe have often been the predicates that seemed to entail its divinity, but the sense in which the universe is infinite and eternal is surely not that in which the traditional religions have ascribed these

predicates to a god. What is clear is that pantheism as a theology has a source, independent of its metaphysics, in a widespread capacity for awe and wonder in the face both of natural phenomena and of the apparent totality of things. It is at least in part because pantheist metaphysics provides a vocabulary that appears more adequate than any other for the expression of these emotions that pantheism has shown such historical capacity for survival. But this does not, of course, give any warrant for believing pantheism to be true.

See also Aristotle; Averroes; Boehme, Jakob; Bradley, Francis Herbert; Brahman; Bruno, Giordano; Darwinism; Eckhart, Meister; Erigena, John Scotus; Eternity; Fichte, Johann Gottlieb; Goethe, Johann Wolfgang von; God, Concepts of; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Indian Philosophy; Infinity in Theology and Metaphysics; Jacobi, Friedrich Heinrich; Lessing, Gotthold Ephraim; Marcus Aurelius Antoninus; Neoplatonism; Pantheismusstreit; Schelling, Friedrich Wilhelm Joseph von; Schleiermacher, Friedrich Daniel Ernst; Shelley, Percy Bysshe; Spinoza, Benedict (Baruch) de; Tauler, Johannes; Toland, John; Xenophanes of Colophon.

Bibliography

Boehme, Jakob. *Works.* Edited by C. J. Barber. London, 1909–.

Bruno, Giordano. “Concerning the Cause, Principle, and One.” Translated by Sidney Greenberg in *The Infinite in Giordano Bruno.* New York: King’s Crown Press, 1950.

Bruno, Giordano. “On the Infinite Universe and Worlds.” Translated by Dorothea W. Singer in *Giordano Bruno: His Life and Thought.* New York: Schuman, 1950.

Eckhart. *Meister Eckhart.* Edited, with introduction, by O. Karrer. Munich, 1926.

Erigena, John Scotus. *Opera.* In *Patrologia Latina,* edited by J. P. Migne. Paris, 1844–1864. Vol. 122.

Fichte, J. G. *Die Schrifte zu J. G. Fichte’s Atheismusstreit.* Edited by H. Lindau. Munich, 1912.

Fichte, J. G. *The Science of Knowledge.* Translated by A. E. Kroeger. Philadelphia, 1868.

Flint, Robert. *Antitheistic Theories.* London, 1878. Baird lectures.

Hegel, G. W. F. *Lectures on the Philosophy of Religion.* Translated by E. B. Speirs and J. B. Sanderson, 3 vols. London: K. Paul, Trench, Trubner, 1895.

Schelling, Friedrich. *Werke.* Edited by M. Schröter. 8 vols. Munich: Beck, 1927–1956.

Schleiermacher, Friedrich. *On Religion: Speeches to Its Cultured Despisers.* Translated by J. W. Oman. London: K. Paul, Trench, Trubner, 1893.

Spinoza, Benedict de. *The Chief Works.* Translated by R. H. M. Elwes, 2 vols. New York: Dover, 1951.

Alasdair MacIntyre (1967)

PANTHEISMUSSTREIT

Pantheismusstreit or the pantheism controversy, came to the attention of the public in 1785 when Friedrich Heinrich Jacobi published *Ueber die Lehre des Spinoza,* his correspondence with Moses Mendelssohn concerning Gotthold Ephraim Lessing’s late Spinozist phase. Other prominent writers, including Immanuel Kant, Johann Gottfried Herder, Johann Wolfgang von Goethe, Johann Kaspar Lavater, and Johann Georg Hamann, became involved in this dispute, which led to an objective reappraisal of Spinozism. The first important reaction to Benedict de Spinoza’s influence in Germany had been Gottfried Wilhelm Leibniz’s *Theodicy* (1710). At the time of the pantheism controversy, the distorted image of Spinoza, the “satanic atheist,” was definitely destroyed. This image had been created by Pierre Bayle and cultivated in Germany by Theophil Gottlieb Spitzel (1639–1691), Johann Christophorus Sturm (1635–1703), Johann Konrad Dippel (c. 1672–1734), and Christian K. Kortholt (1633–1694), whose *De Tribus Impostoribus Liber* (1680) had attacked Herbert of Cherbury, Thomas Hobbes, and Spinoza as “impostors.”

INCEPTION OF THE CONTROVERSY

Jacobi’s book constituted one stage in the struggle waged by the supporters of Hamann (whose sentimentalist faith Jacobi attempted to combine with Kant’s critical philosophy) against the religious rationalism of the Berlin Enlightenment, whose proponents were grouped around Friedrich Christian Nicolai and the *Berlinische Monatsschrift.* In his *Golgotha und Scheblimini* (1784), Hamann had attacked the theistic rationalism of Mendelssohn’s *Jerusalem* (1783). A work prized by Kant, Herder, Mirabeau, and Christian Garve, *Jerusalem* was directed against state-imposed creeds and religions of revelation.

Jacobi’s hasty publication of his correspondence with Mendelssohn, too, was indirectly inspired by Hamann. The latter informed Jacobi on June 29, 1785, that the first part of Mendelssohn’s *Morgenstunden* was already being printed. Wrongly suspecting that Mendelssohn had mentioned their controversy over Lessing in this work, Jacobi committed a dual breach of trust. To his *Ueber die Lehre des Spinoza* he appended anonymously a fragment from Goethe’s unpublished “Prometheus” (1774) that Jacobi had shown Lessing during a conversation at Wolfenbüttel on July 7, 1780.

It was this conversation that served as the starting point and focus of the pantheism controversy. To the report of this conversation Jacobi added a digest of an

argument with Mendelssohn that had ensued from a report by Elise Reimarus (February 1783) to the effect that Mendelssohn was busy with a work on Lessing. Through her, Jacobi led Mendelssohn to believe that "Lessing had been a Spinozist" but had never admitted it to his friend Mendelssohn because the latter had never taken seriously a relevant hint concerning the Spinozist purport of Paragraph 73 of Lessing's *Erziehung des Menschengeschlechts.* Mendelssohn, through Elise Reimarus, then addressed precise questions to Jacobi regarding the character of Lessing's alleged Spinozism. He considered it unlikely that, one, Lessing had been a Spinozist and that, two, he would have remained silent about it to a friend of many years' standing (Mendelssohn) while confiding it to the first stranger that had come along (Jacobi). Mendelssohn suggested courteously that perhaps Lessing, as was his nature, had made in jest certain paradoxical statements to Jacobi. However, if Jacobi could conclusively demonstrate Lessing's Spinozism, then, Mendelssohn allowed, he would have to give precedence to the truth in the work he planned to write about his friend.

In his reply of November 4, 1783, Jacobi again gave details of his conversations with Lessing. But in so doing, he misjudged his own situation. It was obvious that Lessing, tired of hearing Spinoza treated "like a dead dog," had been attempting to provoke Jacobi into a refutation of Spinozism. Jacobi, however, had declared himself helpless against the geometrical reasoning of Spinoza, which seemed unanswerable to him. Although he rejected Spinoza's "fatalism" and the concept of a God who created without insight and without will, he could find no counterarguments. To this Lessing had replied, "I note that you would like to have your will free; I do not crave free will." Lessing characterized the tendency to give thought the precedence over other life forces as a human prejudice. He asked Jacobi whether he thought he could derive the concept of an extramundane rationally creative deity from Leibniz. "I fear," Lessing added, "that Leibniz himself was fundamentally a Spinozist." He recalled "a passage in Leibniz where it is said of God that he himself is in a state of everlasting expansion and contraction, and that this constitutes the creation and existence of the world." Hard-pressed by the logic of Lessing as well as that of Spinoza, which "admits of no cause of things separate from the world," Jacobi saved himself by a leap into a sentimentalist faith in the God of Christianity who orders the world teleologically. With unconcealed irony, Lessing remarked that such a leap of faith ending up in a somersault was something he could no longer exact of his "old legs and heavy head." Derisively, he professed to find agreements with his own system even in Charles Bonnet's *Palingénésie,* which Lavater—without the author's permission—had translated and had dedicated to Mendelssohn in an ill-fated attempt at proselytizing. Lessing also claimed to discern "obvious Spinozism" in Frans Hemsterhuis's *Aristée.* Jacobi himself believed he recognized in the disputed Paragraph 73 of Lessing's *Erziehung des Menschengeschlechts* his Spinozist interpretation of Christ as reality (*natura naturata*) and of God as the infinite substance (*natura naturans*).

Seven months after his reply to Mendelssohn (June 1784), Jacobi learned from Elise Reimarus that Mendelssohn had put aside his *Lessing* "in order first to venture a round with the Spinozists or 'all-in-one'rs.'" In August of that year, Mendelssohn wrote his *Erinnerungen* and sent them to Jacobi without, however, publishing them at that time. (They first appeared in 1786 in *Moses Mendelssohn an die Freunde Lessings,* pp. 36–56). In the *Erinnerungen* Mendelssohn marshaled rationalistic arguments against Spinoza and again expressed his disbelief in Lessing's Spinozism. He dealt sarcastically with Jacobi's "honorable retreat under the flag of faith" as a device necessary for Christian philosophers; Mendelssohn's own religion, on the other hand, allowed him to "raise doubts on grounds of reason" and did not dictate to him "any belief in eternal verities." Mendelssohn left unanswered Jacobi's *Lettre à M. Hemsterhuis,* a copy of which the author had sent him on September 5, 1784. But he notified his correspondent once again that pantheism would indeed come under discussion in the first part of the *Morgenstunden,* although their mutual correspondence would be disregarded. Mendelssohn requested that Jacobi delay publishing his "counterrecollections" until after the publication of the *Morgenstunden.*

Jacobi again sent Mendelssohn an exposition of Spinozism, in forty-four paragraphs, which ended in an enthusiastic identification of Christian faith, love, and—surprisingly—knowledge (in the sense of knowledge of nature). Mendelssohn, astonished at Jacobi's proselytizing zeal, called on Reimarus to act as arbiter in the matter of the controversy over Lessing. Reimarus counseled silence about the whole affair so as not to dishonor the memory of Lessing. Still another exegesis of Spinozism by Jacobi in six paragraphs began with the traditional thesis: "Spinozism is atheism."

Despite Mendelssohn's renewed assurances to Elise Reimarus on May 24, 1785, that he would not make use of his correspondence with Jacobi, the latter with an utter lack of consideration published the letters on August 28, 1785. Jacobi's account reads like an exorcism of the mag-

netic powers of Spinozism, whereas Mendelssohn's concern in the controversy was only to clear Lessing of the charge of Spinozism and to contrast his own religion of reason with Jacobi's visionary religion of sentiment, as well as to polemicize against Spinoza with Wolffian arguments. Mendelssohn's main proof for the existence of a rational God (in Part I of the *Morgenstunden*) was that all that is real must first be thought as real by some being, hence there exists an infinite intellect.

RESULTS OF THE CONTROVERSY

The pantheism controversy spread to wider circles of German intellectual life with the anonymous publication in 1786 of *Die Resultate der Jacobi'schen und Mendelssohn'schen Philosophie* by Thomas W. Wizenman, a young follower of Hamann and a Pietist, who had been induced by Jacobi to read Spinoza. Wizenman, under the guise of a disinterested spectator, openly took Jacobi's side. As Kant later revealed it, Wizenman launched into an *argumentum ad hominem* against Mendelssohn, attempting to destroy deism with atheism, and atheism with deism. For the fideist Wizenman, it was impossible to demonstrate the existence or the nonexistence of God and his relationships to the world. He tried to define the concept of reason in such a fashion that the rationality of a belief in revealed religion would proceed from this definition, once historical evidence of the revelation was at hand.

Compelled by Wizenman's publication to express an opinion, Kant in "Was heisst: sich im Denken orientieren?" (*Berlinische Monatsschrift,* October 1786) rejected both Jacobi's sentimentalist faith and Mendelssohn's rationalist faith as subjective views that conceal in themselves the danger of fanaticism. As in the later *Critique of Judgment* (Paragraph 80), Kant declared that pantheism did not provide a teleological explanation of things, so in the *Monatsschrift* article he defended himself against the reproach that his *Critique of Pure Reason* had promoted Spinozism: "Spinozism speaks of thoughts that themselves think and thus of an accidental thing that still at the same time exists for itself as subject—a concept that is not to be found at all in the human understanding and cannot be brought into it." Kant disapproved of Mendelssohn's attempt to reduce the quarrel of freedom of will versus determinism to a matter of pure logomachy (*Einige Bermerkungen zu Jakobis Prüfung der Mendelssohnschen Morgenstunden,* Leipzig, 1786).

More important than the polemics of the pantheism controversy were its effects on Herder and Goethe and later on Friedrich Schleiermacher, Friedrich von Schelling, and G. W. F. Hegel. Herder, in his five conversations titled *Gott* (1787), deplored Spinoza's terminological dependence on René Descartes, but he accepted Spinoza's concept of God, whom he regarded as the primal power from which all other powers derive. Thus in his own way he came close to the concept of the primal phenomenon that Goethe, as a metaphysical philosopher of nature, was seeking to investigate.

Goethe himself had reread Spinoza in January 1785 and had found in him the foundations for his own holistic or antimechanistic, anti-Newtonian concept of the universe. He had already, on June 4, 1785, objected to Jacobi: "You acknowledge the highest reality, which is the basis of Spinozism, on which all else rests, from which all else flows. He does not prove the Being of God, Being *is* God. And if for this reason others scold Spinoza for being an atheist, I should like to name him and praise him as *theissimum,* indeed, *christianissimum.*" On October 21 of the same year, Goethe sharply attacked Jacobi's play on the word *believe* as the behavior of a "faith-sophist," admonished him to apply himself to "clarity and distinctness of expression," and admitted "that while by nature I do not share Spinoza's mode of conception, if I had to cite a book that, more than any I know, agrees most fully with my own conception, I should have to name the *Ethics.*" On May 5, 1786, he expressed his disagreement with Jacobi:

> I cling more and more firmly to the reverence for God of the atheist [Spinoza] … and I cede to you [Christians] all that your religion enjoins and must enjoin … When you say that one can only *believe* in God … then I say to you that I lay great weight on *looking and seeing* and when Spinoza, speaking of *scientia intuitiva,* says *Hoc cognoscendi genus procedit ab adaequata idea essentiae formalis quorundam Dei attributorum ad adaequatam cognitionem essentiae rerum* [This manner of knowing moves from the adequate idea of the formal essence of some attributes of God to the adequate knowledge of the essence of things], these words give me courage to devote my entire life to the contemplation of the things that I can reach and of whose *essentia formali* I can hope to fashion an adequate idea

Just as Goethe, who, inspired by the pantheism controversy to make a study of Spinoza, became conscious of his own holism while reading the *Ethics,* so pantheism, thanks to its contact with Spinozism, progressed from its traditional manifestation as Neoplatonic emanation to a concept of evolution, which in Hegel's philosophy (and in

the twentieth century, that of Henri Bergson) entails the development of the Absolute in and with the world.

See also Hamann, Johann Georg; Jacobi, Friedrich Heinrich; Mendelssohn, Moses; Spinoza, Benedict (Baruch) de.

Bibliography

Jacobi, Friedrich Heinrich. *Ueber die Lehre des Spinoza in Briefen an den Herrn Moses Mendelssohn.* Breslau, 1785. 2nd ed., revised and enlarged, Breslau: G. Löwe, 1789.

Mendelssohn, Moses. *An die Freunde Lessings.* Berlin: C. Voss, 1786.

Mendelssohn, Moses. *Morgenstunden, oder über das Daseyn Gottes.* Berlin: C. Voss, 1786.

Scholz, H. *Die Hauptschriften zum Pantheismusstreit zwischen Jacobi und Mendelssohn.* Berlin: Reuther and Reichard, 1916.

Kurt Weinberg (1967)
Translated by Albert E. Blumberg

PAPINI, GIOVANNI
(1881–1956)

Giovanni Papini, an Italian pragmatist philosopher and literary figure, was born in Florence into a family of modest means and had no formal education. Papini described himself in his *Un uomo finito* (Florence, 1913; translated by Virginia Page as *Failure; Un Uomo Finito,* New York, 1924), a book that was frankly and painfully biographical, as self-taught, urged on by an insatiable curiosity and a burning desire to investigate the various forms of knowledge. He quickly made a name for himself in Italian culture at the beginning of the twentieth century with his attack on the then prevailing positivist philosophy of Roberto Ardigò and his support of nationalistic tendencies and opposition to the ideals of democracy. He became a close friend of Giuseppe Prezzolini and other young writers who advocated doing away with the old oligarchies and giving a new impetus to the spiritual life of the country. The fruit of this collaboration was the birth in 1903 of *Leonardo,* a nonconformist review that published the most important contemporary thinkers. They chose Friedrich Nietzsche, Henri Bergson, William James, and F. C. S. Schiller as their exemplars and leaders, but the interests of the *Leonardo* group embraced the avant-garde currents in art and literature as well.

In his writings, later gathered together in a book titled *Pragmatismo* (Milan, 1913), Papini defined the essential aspects of his thought. His is a kind of magic pragmatism, markedly different from the logical and scientific pragmatism of C. S. Peirce. This pragmatism rejects the positivists' agnosticism concerning issues that go beyond experience; that metaphysical problems lack meaning does not indicate a lack in our intellectual capabilities but rather how very human the nature of knowing is. Instead of striving for definitive explanations in the manner of the traditional philosophies, the pragmatist is concerned with the methods and instruments that aid in defining the various forms of knowledge and activity. He does not believe in absolute principles or immutable truths; neither does he stop at mere description and generalization of the facts of experience. His aim is to develop laws and predictions, with the sole purpose of increasing the power of man over nature. No metaphysical hypothesis, observed Papini, is more valuable than another, and none can be recognized as true. On the contrary, the pragmatist viewpoint is one of maximum freedom and advocates a plurality of attitudes. Papini's celebrated definition of pragmatism was praised and quoted by William James:

> Pragmatism is a *corridor theory,* a corridor of a great hotel where there are 100 doors that open onto 100 rooms. In one there is a faldstool and a kneeling man who wants to regain his faith, in another a writing-desk and a man who wants to kill every metaphysic, in a third a laboratory and a man who wants to find new vantage points on the future. (*Pragmatismo,* p. 82)

Papini's *Leonardo* period, with neo-Hegelians such as Benedetto Croce and Giovanni Gentile aiding the attack on positivism, terminated in 1906. But this was only the beginning of a painful intellectual journey in which Papini sought, without success, to give form and coherence to his thought. He participated in the battle of ideas of *La voce,* directed by his friend Prezzolini; then he broke away and in 1911, in collaboration with Giovanni Amendola, directed a review with a strong moral bent, *L'anima*; and finally he founded *Lacerba,* an avant-garde journal violently opposed to the prevailing order of things. In the meantime, his literary output was enriched by numerous works, including *Il crepuscolo dei filosofi* (The twilight of the philosophers; Milan, 1906), *La cultura italiana* (Florence, 1906), written in collaboration with Prezzolini, and *L'altra metà* (The other half; Ancona, 1912). In addition to these books, a great number of articles testify to his zeal and his cultural interests. In this period Papini drew further away from the idealism gaining popularity in Italy, intensified his dissent with the school of Croce, and supported the futurist movement in accordance with his rebellion against traditional aesthetic rules.

Papini strongly favored Italian intervention in World War I because he saw the war as a decisive conflict between the old and the new. However, the war led him to a reassessment of Christian values and to embrace the works of the fathers of the church, and in particular those of St. Augustine. He regarded Augustine, to whom he devoted a book (*S. Agustino,* Florence, 1929), as a defender of the faith, an uncompromising polemicist, and an unsurpassable model of humanity reaching out toward the divine. Papini's activity did not diminish after his religious "conversion," but gradually became less and less concerned with philosophical matters, and concentrated instead on literary and scholarly subjects. Stricken by a disease that deprived him almost completely of the use of his senses but left his mind as active as ever, Papini bore up bravely until his death.

See also Ardigò, Roberto; Augustine, St.; Bergson, Henri; Croce, Benedetto; Gentile, Giovanni; James, William; Nietzsche, Friedrich; Peirce, Charles Sanders; Pragmatism; Schiller, Ferdinand Canning Scott.

Bibliography

See Michele Federico Sciacca, *Il secolo XX,* 2nd ed. (Milan, 1947), Vol. I, pp. 22–25; Eugenio Garin, *Cronache di filosofia* (Bari: G. Laterza, 1955); Antonio Santucci, *Il Pragmatismo in Italia* (Bologna: Mulino, 1963).

Antonio Santucci (1967)
Translated by Robert M. Connolly

PARACELSUS

(1493–1541)

Paracelsus was the pseudonym of Philippus Aureolus Theophrastus Bombastus (Baumastus) von Hohenheim, the reformer of medicine and pharmacology, chemist, philosopher, iconoclast, and writer. If he himself assumed this name, it could signify "higher than high," or "higher than Hohenheim," a jibe at his illegitimate paternal grandfather. Born in Einsiedeln, Switzerland, where his father practiced medicine, Paracelsus later lived at Villach in Carinthia (Austria), a center of mining, smelting, and alchemy—metal lores that were to occupy him for the rest of his life. From the age of fifteen his life was migratory. After medical studies at various German and Austrian universities, he seems to have completed his doctorate in 1515 at Ferrara under a faculty that was Scotist, Platonist, and humanist.

For the next eleven years, Paracelsus traveled throughout Europe, jeopardizing his authority as a physician by practicing surgery (then a craft, not a learned profession) in the army of Charles V and by experimental prescriptions. He visited spas, analyzed the waters, treated by hypnosis, and sometimes alleviated pain with laudanum. At Salzburg he narrowly escaped execution for participating in a peasants' revolt. When, in 1526, he settled at Strasbourg to establish himself in medical practice, he was famous as an object of superstitious distrust. But his spectacular cure of the printer Johann Froben quickly led to friendships with such men as Desiderius Erasmus and Oecolampadius and an appointment—against the will of the faculty—as medical lecturer at the University of Basel.

His eminence was short-lived. Lectures in German (rather than Latin), rejection of the canonical theory of Avicenna and Galen, denunciation of the apothecaries, and a public burning of the works of Avicenna were topped by the death of Froben. Those whose vested interests had been threatened tricked Paracelsus into behavior that could justify dismissal and arrest.

From 1528 until his death, his life was once again nomadic. Unkept promises and unstable patronage led him to Colmar, Nuremberg, Saint Gall, Villach, Vienna, and finally to Salzburg, where he died, probably of cancer, perhaps of metal poisoning.

Among his medical innovations were chemical urinalysis; a biochemical theory of digestion; chemical therapy; antisepsis of wounds; the use of laudanum, ether (without awareness of its anesthetic properties), and mercury for syphilis; and the combining of the apothecary's and surgeon's arts in the profession of medicine.

Paracelsus's numerous books are mostly variants on the theme of man (the microcosm) in relation to nature (the macrocosm). The most important are *Archidoxis* (c. 1524); the treatises on syphilis (c. 1529); *Opus Paragranum* (c. 1529); *Opus Paramirum* (c. 1530); *Philosophia Sagax* (c. 1536); and *Labyrinthus Medicorum* (1538).

Paracelsian philosophy was both traditional and new. Its medieval elements are traceable to alchemy and Kabbalism, which are branches of a trunk rooted in Hellenistic Neoplatonism, the Corpus Hermeticum, and Gnosticism. These occult lores shared the concept of creation through corruption; the axiom "That which is above is one with that which is below"; belief in a bisexual, homogeneous, hylozoic universe; a cyclic theory of time; and an animism approximating pantheism.

A mystery religion of life rather than merely of gold, medieval alchemy employed Semitic and Greco-Roman

mythology as a screen against the unenlightened and as a vehicle of private communication for adepts. Although Paracelsus counted himself an adept, he abandoned the tradition of reserve and discarded most of the mythological symbolism. Unlike his predecessors, he wrote to clarify. He explained that alchemy's real desideratum was the secret of life.

Like Kabbalists and alchemists, Paracelsus believed in the theory that decay is the beginning of all birth. Nature emerges through separations: First, prime matter separates out of ultimate matter (also called *Yliaster* or *Mysterium Magnum*), which is eternal and paradoxically immaterial. "The first was with God … that is *ultima materia*; this *ultima materia* He made into *prime matter* … that is a seed and the seed is the element of water [fluid]." God spins ultimate matter out of himself. This yields, by separation, the prime matter of individual objects, a watery matrix, perpetually spawning nature, perpetually resolvable back into ultimate matter. Human creativeness in art, alchemy, or pharmacology repeats the primal act. The human demiurge, like God, separates rather than combines.

The Paracelsian theory of time resembles that of Plotinus. Time is qualitative change: growth, transition—even fate. Given the basic concept of cyclic generation and decay, Paracelsian time would be for the material cosmos a cycle of becoming. But there are two orders of time: force time (within) and growing time (without). Like the Paracelsian concept of "prime matter" in relation to "ultimate matter," this theory of time is essentially dualist.

"Above" and "below" are substantially the same: "Heaven is man and man is heaven, and all men together are the one heaven," but microcosm and macrocosm are contained by membranes or partitions.

Paracelsus rejected the concept of humors as governed by planets and substituted a chemical theory of humors as properties: salt, sweet, bitter, and sour. He retained the medieval alchemistic variant of the four elements and a quintessence, the fifth element, that is life. He tended to treat fire as less elementary than the combustible principle, sulfur. Medieval alchemy had stressed the sexual polarity of two elements, fire (identified with the male principle) and water (identified with the female principle), and contrasted flame with flow and sulfur with mercury. Paracelsus reinterpreted these as principles rather than as elements and added a third principle, salt. These are properties or states—combustible, fluid or vaporous, and solid; each confers on matter its structure, corporality, and function. As constituents of ultimate matter, these are absolutes; as components of nature, they are infinitely variable in all sensuously discernible properties. Every natural object has its own sulfur, salt, and mercury, as well as its own quintessence.

Absolute life comes from *Ens Seminis*, the cosmic protoplasm. *Ens Astrale* is to the microcosm (man) as the firmament is to the macrocosm (nature). It can sustain or poison from within, as a toxic atmosphere can poison sea water and fish. *Ens Veneni* is the poison from without. Nature lives by dying; life eats life. Man may eat the flesh of an animal whose food would poison him, but within every living body there is an alchemist that selects what is food for that body. *Ens Naturale* is the bodily harmony of the chemical humors. *Ens Spirituale* has its equivalent in what psychiatry calls the psyche. Against the common belief of his day, Paracelsus argued that madness was not demonic possession and that evil dreams were not intercourse with incubi or succubi. Mind produces diseases both in itself and its own body or in another mind or body through hypnosis, fetishism, or demonstrable ill will. Most diseases are positive evils, but there is *Ens Dei*, God's will, which no doctor can circumvent.

Although accused by Erasmus of dualist heresy because of the importance he gave primal matter and because he described illness as intrinsically evil, Paracelsus died in the Church of Rome, and his burial place became a shrine.

See also Avicenna; Erasmus, Desiderius; Galen; Gnosticism; Kabbalah; Macrocosm and Microcosm; Neoplatonism; Pantheism; Plotinus; Time.

Bibliography

WORKS BY PARACELSUS

Opera Omnia. 12 vols, edited by John Huser. Basel, 1589–1591. The original German text.

Opera Omnia. 3 vols, edited by F. Bitiskius. Geneva, 1658. In Latin.

Sämtliche Werke. 15 vols, edited by Karl Sudhoff and E. Matthiessen. Munich, 1922–1933. In German; the standard critical edition.

Four Treatises of Theophrastus von Hohenheim, edited by Henry Sigerist, C. Lilian Temkin, George Rosen, and Gregory Zilboorg. Baltimore: Johns Hopkins Press, 1941.

Selected Writings, edited by Jolande Jacobi; translated by Norbert Guterman. New York: Pantheon, 1951 and 1958. Contains an introduction by the editor. Excellent.

WORKS ON PARACELSUS

Browning, Robert. *Paracelsus*. London, 1835.

Dear, Peter. *Revolutionizing the Sciences: European Knowledge and Its Ambitions, 1500–1700*. Princeton, NJ: Princeton University Press, 2001.

Debus, Allen G. *The Chemical Philosophy: Paracelsian Science and Medicine in the Sixteenth and Seventeenth Centuries*. New York: Science History Publications, 1977.

Debus, Allen G. "The Paracelsian Compromise in Elizabethan England." *Ambix* 8 (June 1960): 71–97.

Donne, John. *Ignatius His Conclave*. London, 1613.

Koyre, Alexandre. "Paracelsus (1493–1541)." *Graduate Faculty Philosophy Journal* 24 (1) (2003): 169–208.

Pachter, Henry M. *Magic into Science*. New York: Schuman, 1951. Represents Paracelsus as a proto-Faust; readable.

Pagel, Walter. "Paracelsus and the Neoplatonic and Gnostic Tradition." *Ambix* 8 (October 1960): 125–166.

Pagel, Walter. *Paracelsus. An Introduction to Philosophical Medicine in the Era of the Renaissance*. New York: Karger, 1958. Excellent.

Pagel, Walter. "The Prime Matter of Paracelsus." *Ambix* 9 (October 1961):, 117–135.

Stillman, John Maxson. *Paracelsus*. London: Open Court, 1920. Emphasis on science.

Stoddart, Anna M. *The Life of Paracelsus*. London: Murray, 1911. Browning's interpretation.

Webster, Charles. *From Paracelsus to Newton: Magic and the Making of Modern Science*. Cambridge, U.K.: Cambridge University Press, 1982.

Weeks, Andrew. *Paracelsus: Speculative Theory and the Crisis of the Early Reformation.* Albany: State University of New York Press, 1997.

Linda Van Norden (1967)
Bibliography updated by Tamra Frei (2005)

PARACONSISTENT LOGICS

The driving thought of paraconsistency is that there are situations in which information, or legal, scientific, or philosophical principles (and so on) are inconsistent, but in which people want to draw conclusions in a sensible fashion. Clearly, if one uses a logical consequence relation in which contradictions imply everything—that is, in which $A, \neg A \vdash B$, for all A and B—this is not possible: a person would have to conclude everything (*triviality*). This motivates the definition of a paraconsistent logic. The principle of inference that contradictions entail everything is called *explosion* (or *ex falso quodlibet sequitur*). A paraconsistent logic is one in which explosion is not valid.

Paraconsistent logics are not new. As Aristotle (*An. Pr.* 63^b31–64^a16) points out, syllogistic is paraconsistent. The idea that explosion is a correct principle of inference seems to have arisen in the twelfth century, with the discovery of the following simple argument. Suppose that $\neg A$; then $\neg A \vee B$. But now suppose that A as well. Then B follows by the disjunctive syllogism ($A, \neg A \vee B \vdash B$). Explosion and the disjunctive syllogism had variable fortunes in later Medieval logic. A common move was to distinguish two notions of validity: one (*material*) for which they held; and one (*formal*) for which they do not. All this was forgotten after the Middle Ages. But since the early twentieth century, the hegemony of Frege/Russell (classical) logic, according to which explosion is valid, has ensured the orthodoxy of the principle.

Modern formal paraconsistent logics started to appear in the second half of the twentieth century. Amongst the earliest paraconsistent logics were those proposed by Stanisław Jaśkowski (1948) and Newton da Costa (1963). The paraconsistent possibilities of the relevant logic of Alan Anderson and Nuel Belnap (1960s) was also soon recognized. By the end of the twentieth century there were many paraconsistent logics with well-defined semantics and proof theories.

In the semantics of most paraconsistent logics, validity is defined in terms of the preservation of truth-in-an-interpretation. It must therefore be possible to have interpretations where A and $\neg A$ are both true. There are several ways of achieving this end. One is to take truth to be truth-at-a-world in a world-semantics for modal logic (as in Jaśkowski's system D_2, "discussive logic"). In this case, the inference of adjunction ($A, B \vdash A \& B$) will fail, giving rise to a nonadjunctive paraconsistent logic. Another possibility is to graft a non-truth-functional negation on to some positive logic (as in the da Costa *C*-systems). The truth value of $\neg A$ is not determined by that of A; both may then be true. This gives so-called "positive-plus" paraconsistent logics. A third possibility is to employ a many-valued logic in which some designated truth value, v, is a fixed point for negation. That is, if the value of A is v, the value of $\neg A$ is also v. v may be the value *both true and false*, as in Graham Priest's *LP*, or the value 0.5 where the semantics has the real numbers between 0 and 1 as truth values. The way that negation is handled in relevant logic also has the same effect.

In nearly all paraconsistent logics, there are ways of recapturing the full force of classical reasoning. Thus, in discursive logic, if the premises are conjoined then they have all of their classical consequences. Da Costa suggested augmenting the language with an operator, °, such that, intuitively, $A°$ expresses the consistency of A. The classical negation of A can then be expressed by $\neg A \& A°$. A different way was suggested by Diderik Batens. Consistency-ordering is defined on interpretations, such that classical interpretations (and only those) come out as the most consistent. A notion of validity is then defined according to which an inference is valid iff (meaning "if and only if") the conclusion holds in all those interpreta-

tions which are as consistent as possible, given only that the premises hold in them. This gives a nonmonotonic notion of consequence according to which the consequences of a consistent set of sentences are just their classical consequences. (Batens developed the idea into a whole family of nonmonotonic logics with interesting properties, Adaptive Logics.)

Paraconsistent logics have many applications. They can be used as the inference engine for a computational database, where the data may not be reliable, or used to analyze the reasoning of inconsistent theories in the history of science—such as the original infinitesimal calculus or Bohr's theory of the atom. (The inconsistency of each of these was acknowledged in their times.) The same also holds true for the inconsistent but nontrivial theories that paraconsistent logic makes possible, including various mathematical theories. One can be interested in these because they have an intrinsically elegant structure, are instrumentally useful, and are good approximations to the truth. None of this requires one to suppose that the inconsistent theories may be true.

The view that some contradictions are true is dialeth(e)ism (a di/aletheia being a true statement of the form $A \& \neg A$). Unless a dialetheist takes everything to be true (not an attractive view!), they also require a paraconsistent logic. Though there have been dialetheists—such as Hegel—in the history of European philosophy, dialetheism is a strongly heterodox view because it flies in the face of the Law of Noncontradiction. The construction of contemporary paraconsistent logics has given the view a new lease of life. In particular, beginning in the 1970s, it was advocated by Priest and Richard Sylvan (né Routley).

Modern dialetheists argue for their view by appealing to certain features of motion, inconsistent systems of norms, and various other considerations. A major appeal has always been to the paradoxes of self-reference, such as the Liar and Russell's paradox (and related phenomena such as Gödel's incompleteness theorem). The paradoxical arguments are what they appear to be: arguments establishing that certain contradictions are true. In particular, a dialetheist can subscribe to the principles which generate these paradoxes: the unrestricted *T*-schema for truth ("*A*" is true iff *A*) and the unrestricted comprehension principle for sets (for any condition, there is a set comprising all and only those things satisfying that condition). In particular, it is possible to construct inconsistent but nontrivial theories containing these principles. Not all paraconsistent logics are suitable for this enterprise, however. In this context, any logic which endorses the principle of contraction $(A\rightarrow (A\rightarrow B) \vdash A\rightarrow B)$ gives rise to triviality, in the form of Curry paradoxes. Such logics include the da Costa *C* logics and the stronger relevant logics.

See also Logic, History of; Logic, Non-Classical; Relevance (Relevant) Logics.

Bibliography

Brown, Bryson. "Paraconsistent Logic: Preservationist Variations." In *Handbook of the History of Logic*. Vol. 7, edited by Dov M. Gabbay and John Woods. Amsterdam, Holland: Elsevier, forthcoming.

Batens, Diderik. "Inconsistency-Adaptive Logics." In *Logic at Work: Essays Dedicated to the Memory of Helena Rasiowa*, edited by E. Ortowska, 445–472. Heidelberg, Germany: Physica Verlag, 1999.

Batens, Diderik, Chris Mortensen, Graham Priest, and Jean-Paul Van-Bendegem, eds. *Frontiers of Paraconsistent Logic*. Baldock, U.K.: Research Studies Press, 2000.

Carnielli, Walter, A., Marcelo E. Coniglio, and João Marcos. "Logics of Formal Inconsistency." In *Handbook of Philosophical Logic*. Vol. 12, edited by Dov M. Gabbay and Franz Guenthner. Dordrecht, Holland: Kluwer, forthcoming.

Mortensen, Chris. *Inconsistent Mathematics*. Dordrecht, Holland: Kluwer, 1995.

Priest, Graham. *In Contradiction: A Study of the Transconsistent*. 2nd ed. Dordrecht, Holland: Martinus Nijhoff, 1987. Oxford: Oxford University Press, forthcoming.

Priest, Graham. "Paraconsistent Logic." In *Handbook of Philosophical Logic*. 2nd ed., vol. 6, edited by Dov M. Gabbay and Franz Guenthner, 287–393. Dordrecht, Holland: Reidel, 2002.

Priest, Graham, Richard Routley, and Jean Norman, eds. *Paraconsistent Logic: Essays on the Inconsistent*. Munich, Germany: Philosophia Verlag, 1989.

Graham Priest (2005)

PARADIGM-CASE ARGUMENT

"Paradigm-case argument" is a form of argument against philosophical skepticism found in contemporary analytic philosophy. It counters doubt about whether any of some class of things exists by attempting to point out paradigm cases, clear and indisputable instances. A distinguishing feature of the argument is the contention that certain facts about language entail the existence of paradigm cases. This claim, however, has been disputed in recent years, and the future status of the argument depends upon whether it can be upheld.

The paradigm-case argument has been used against a wide range of skeptical positions. A typical example is doubt about our ability to perceive directly material objects. Such doubt can be raised by reflection upon the physiological and physical facts about perception. For example, since seeing involves the transmission of light waves to our eyes and these waves are what immediately affects our eyes, it may appear that we are mistaken in thinking that we see objects. If anything, we should say that we see light waves. The fact that it takes a certain amount of time for light to travel from an object to our eyes lends support to this. How can we see something unless we see it as it is at the present moment? While considerations such as these show how skepticism can arise, one striking fact about the paradigm-case argument is that if it is valid, the skeptic can be refuted directly without the necessity of examining in detail the reasons behind his position.

The first step in the argument is to make the skepticism bear on particular cases. If we cannot perceive material objects, then, presumably, we cannot see the table we are working on or the pen with which we write. Next, a situation is sketched in which, ordinarily, no one would hesitate to affirm just the opposite. If the light is excellent, our eyes open, our sight unimpaired, the table directly before us, and so on, then we should ordinarily have no qualms about stating that we see a table.

The argument would be weak if it relied merely on the fact that people would ordinarily have no doubts in such situations, for it does not follow from this that they state the truth. But the argument claims something more for the kind of situations it describes. It holds that they are indisputably examples of seeing a table because of their relationship to the meaning of the expression "seeing a table." Typically, this relationship is brought out by saying that such a situation is just what we call "seeing a table" or that it is just the sort of circumstances in which one might teach someone the meaning of the expression "seeing a table." Generalizing and taking the strongest interpretation of the force of these remarks, one might ask: "If this *is* just what we call *X*, then in saying that it is *X*, how can we fail to state the truth? If this *is* a situation in which we might teach the meaning of *X*, then how can it fail to be a case of *X*?" In denying that anyone ever sees a table, the skeptic seems to be placed in the position of refusing to apply the expression "seeing a table" to the very situation to which that expression refers.

If the skeptic concedes that the situation presented is an instance of that which he doubted to exist, then he admits defeat. But if, despite what has been said, he will not concede this, the final stage of the argument poses a dilemma. When the skeptic wonders whether we ever really see such things as tables, we naturally understand the words he uses in their usual sense. By "usual sense" is meant no more than what we should have understood by his words *see* and *table* if, instead, he were describing some scene he had witnessed. But how can his words be construed in this way when he refuses to use them of a typical situation in which their usual meaning might be taught and which is just what we ordinarily call "seeing a table"? On the other hand, if the skeptic claims some different or novel meaning for his words, the original shock of his skeptical conclusion is blunted. For in some special sense of the words, it may be true that we never see tables. In fact, what often happens is that the skeptical position maintains its plausibility only through an unnoticed fluctuation between the usual sense of the key expressions and some special sense. The paradigm-case argument may serve to bring out into the open the fact that an unusual meaning must be looked for.

FURTHER APPLICATIONS

Other examples of philosophical doubt to which the paradigm-case argument has been applied include skepticism about the validity of inductive reasoning, about man's free will, about the possibility of knowledge concerning empirical facts generally, and about the reality of the past. In many cases these skeptical positions are founded entirely on a priori considerations, and their stand is not merely that, as a matter of fact, there are no instances of some class of things, but that, as a matter of logical necessity, there could not be any. Philosophers who have argued that we can never genuinely know anything about the empirical world, for example, have almost invariably thought such knowledge a logical impossibility. Their reason is often the supposed impossibility of complete verification of any empirical assertion about the world. But this they take to be a necessary truth following from the fact that there are an infinite number of possible observations and investigations relevant to any such assertion. Similarly, the impossibility of justifying inductive reasoning (that which goes from examined cases to a general conclusion or from past instances to a prediction) has been held on the grounds that there is a logical obstacle in the way of all attempts at justification.

Against such a priori skepticism the argument need not produce an actual paradigm case. The mere fact that a hypothetical case can be described is sufficient. This in part accounts for the fact that philosophers who have employed the argument in practice do not bother to

describe an actual occurrence. So, for example, one writer, in using the argument to refute skepticism about induction, asks us to imagine that "the observed confirmatory instances for the theory of gravitation were a million or ten million times as extensive as they now are" (Paul Edwards, "Bertrand Russell's Doubts about Induction," p. 65). By its very statement this is only a hypothetical case. But the skeptic about induction cannot admit that if this were to happen, we should *then* be justified in accepting the law of gravitation, because if justification were a logical impossibility, no paradigm case of justified inductive inference would even be conceivable.

But not all philosophical skepticism is completely a priori. Doubts about the human ability to choose among genuine alternatives is often supported, for example, by citing the success of the behavioral sciences and arguing that they will eventually be able to describe and predict human actions through causal laws. Here the philosopher appears to argue from empirical premises. But here, also, the descriptions of paradigm cases offered to the skeptic have usually been hypothetical. A writer, for example, who pointed to a marriage where there has been no pressure and the like placed on the two people as a paradigm case of choosing freely would not feel compelled to prove the existence of some actual marriage fitting this description.

The reason why a purely hypothetical instance can be given even where the skepticism is based on empirical premises is that there is a sense in which the skeptic does not deny the existence of paradigm cases. In this example he would not, for instance, dispute the frequent occurrence of the sort of marriage described. And he would be prepared to admit that in such cases the appearances are in favor of a free choice having been exercised. But, he thinks, the other considerations provided by his skeptical argument show that, in fact, it is doubtful or impossible that such an occurrence should be an instance of genuinely free choice. This is why the appeal to the connection between such situations and the meaning of, in this example, the expression "free choice" is the vital step in the paradigm-case argument. It is that which, if anything, shows that whatever the skeptical argument, these circumstances *must* be counted as instances of free choice.

BACKGROUND

The idea that philosophy cannot cast doubt on the applications ordinarily made of everyday expressions is not a new one. It can be seen, for example, in George Berkeley's refusal to draw skeptical consequences from his radical thesis that nothing exists apart from the mind. He did not conclude that we are mistaken in talking of material objects such as trees and tables; instead, he attempted to show how his thesis could be used to analyze the meaning of statements about these things. Everyday language succeeds in saying something true about the world; the only question for him was, *What* does it say?

But what is perhaps novel is the erection of this idea into an explicit philosophical argument. And this is largely the product of what has been called the "revolution in philosophy," which began in England shortly before World War II and which has subsequently dominated much of Anglo American philosophy. The possibility of defeating skepticism by reference to particular cases, however, was already present some time before this in the many essays on the subject, dating from the first decade of the twentieth century, by G. E. Moore.

G. E. MOORE. Moore thought of his opposition to skepticism in any form as a defense of common sense. The statements of common sense that he wished to defend were of two kinds: such context-free statements as "Earth has existed for many years" and such context-bound statements as "Here is a human hand" and "This is a pencil." Moore held that he knew with certainty the truth of statements of both kinds. Any skeptical argument, therefore, which entailed that he did not or could not know them must be mistaken. To his critics this has seemed a strange sort of defense of common sense, for how can one defend a position merely by reaffirming it? In answering this, some writers have suggested that Moore was implicitly using the paradigm-case argument. While it is difficult to interpret Moore's affirmation of context-free statements in this way, the suggestion is quite plausible, for example, when we find him attacking skepticism about the existence of material objects by holding up his hand and saying that it is quite certain that this is a human hand and that at least one material object therefore exists ("Proof of an External World," pp. 145–146).

Moore himself, however, apparently saw his procedure in a different light. He thought of it as a challenge to the skeptic: Which is more certain, the (usually esoteric) premises of your argument or the commonsense statements that you are compelled to deny? Moore also pointed out that whereas the skeptic has an argument that leads to the denial of some commonsense statement, a counterargument can be constructed using the commonsense statement as a premise and the denial of the skeptical reasons as a conclusion. The question then seems to resolve into who has the more certain premises. And in this conflict common sense surely seems to be on

firmer ground. In an examination of four assumptions from which Bertrand Russell had drawn skeptical conclusions, for example, Moore ends by saying: "I cannot help answering: 'It seems to me more certain that I *do* know that this is a pencil and that you are conscious, than that any single one of these four assumptions is true, let alone all four'" ("Four Forms of Scepticism," p. 226). And at a much earlier time he wrote: "I think the fact that, if [David] Hume's principles were true, I could not know of the existence of this pencil is a *reductio ad absurdum* of those principles" (*Some Main Problems of Philosophy*, p. 120).

In this interpretation of his procedure, Moore defends common sense as the more certainly true view of the world. The paradigm-case argument, in contrast, appeals to language to show that skepticism conflicts with the facts about the use of expressions needed to state it. Although Moore pointed to the importance of particular cases, it is necessary to look at the ideas that have subsequently come to the forefront of Anglo American philosophy to see why a connection with language should be thought relevant.

WITTGENSTEIN. Of central importance are the views of Ludwig Wittgenstein, whose work has heavily influenced many of those who have used the paradigm-case argument. (It is, however, debatable whether Wittgenstein himself employed the argument.) One of his central contentions, in opposition to his own earlier work, the *Tractatus Logico-Philosophicus*, was that while rules can be formulated for language, it is a mistake to view the particular uses of language as deriving their correctness from being in accord with rules. Rather, the fact that those who speak the language agree that *this* is the correct thing to say here and *that* incorrect there shows what the rules are. If anything, this agreement in judgment about particular cases is primitive. So, in the notes he dictated to some of his students in 1933–1934 (subsequently known as the *Blue Book*), Wittgenstein said, "It is part of the grammar of the word 'chair' that *this* is what we call 'to sit on a chair.'" It would be a mistake to take it as a consequence of such remarks that if the users of a language agree in calling *this* an example of *X*, then, in the sense which the expression has in their language, this *must* be a case of *X*. Such a principle would indeed immediately yield the validity of the paradigm-case argument.

But there is an obvious objection that an example will illustrate. There was a time, perhaps, when all agreed in calling Earth flat, although it was not. They were in agreement, but they were all mistaken. This, however, is a situation in which people were relying upon certain evidence that proved misleading. And in holding that there is a connection between the situations in which we should use a description and the meaning, or "grammar," of the description, Wittgenstein was probably thinking of circumstances in which we are not relying on evidence. It was one of his important ideas that where it makes sense to speak of having evidence that something is so, it must be (logically) possible to get beyond mere evidence.

Thus, while we may sometimes have evidence that someone is sitting in a chair (from, for example, a report that he is), Wittgenstein would argue that when we are standing in a well-lit room looking at the person so seated, it would be a mistake to suppose we then have mere evidence. This idea runs directly counter to long traditions in philosophy. For philosophers, even those who are not skeptics, have most often held that one gets beyond evidence only in a very small class of statements—in general, only first-person, singular, present-tense assertions about one's own mental life. It appears reasonably certain, however, that some such general claim as Wittgenstein's must be substantiated before the paradigm-case argument can be declared valid, because a paradigm case of, for example, a free choice must be one in which there is *more* than just good evidence that a free choice has been made. Otherwise, the skeptical reasons may be sufficient to show that the evidence is misleading.

Whether Wittgenstein's view, if correct, is sufficient to show the validity of the paradigm-case argument is another question. It will depend, for example, upon whether a situation in which we have got beyond mere evidence is also one in which we cannot be mistaken.

It is important to note that the idea that we must be able to get beyond evidence presupposes that we are dealing with a concept free from logical inconsistency. We cannot, for example, ever be confronted with a round square or a genuine trisection of an angle. But a priori skepticism is based on a "proof" that a certain concept could have no instantiation because there would be some inconsistency in supposing it did. The paradigm-case argument, if it is to be generally employed, may need a proof of its own that no expression in everyday use can turn out to designate a self-inconsistent idea. While this has sometimes been held, more needs to be said about it. It seems impossible that anyone should prove, for example, that the idea of a table is self-inconsistent, but it is not so implausible to suppose that someone might show that the idea of a time machine or of transmigration of souls, which are ordinary expressions in the sense intended, contain contradictions. And is it beyond doubt that the

concept of a free choice, for example, is logically irreproachable? Moreover, if it were to be demonstrated independently that no expression in ordinary language can designate a self-inconsistent idea, this would be sufficient by itself to discredit any a priori skepticism concerned with such expressions and would render the subsequent use of a paradigm-case argument superfluous.

There is a further difficulty in supposing Wittgenstein's view—that what we say in particular circumstances is determinant of what we mean—to entail the validity of the paradigm-case argument. This arises from the fact that particular cases can be related to the meaning of an expression without necessarily being paradigm cases.

This may be brought out by an illustration. Suppose someone doubts the existence of elephants. Very likely the surest way to convince him of his mistake would be to show him the elephants at a zoo or circus. That we call *these* elephants shows something about the meaning of the word *elephant*. If the skeptic about elephants sees no connection between what he has been shown and the existence of elephants, we have grounds for suspecting that he does not know what the word *elephant* means. But the connection need not be that having seen these things, he must admit that elephants exist. All he must admit is that these things have the appearance of elephants (see Wittgenstein, *Philosophical Investigations*, paragraph 354). If he maintains, for example, "These certainly look like elephants, but I am sure that they are in reality camels with false noses and padding," he has acknowledged a connection between what he has been shown and the meaning of the word *elephant*. His skepticism, however, remains.

At this time it is an open question whether the important general ideas about the connection of language to particular cases that have fostered the use of the paradigm-case argument also entail its validity.

CRITICISM AND VARIATIONS

Critics of the paradigm-case argument have questioned the legitimacy of the move from "This is just what we call *X*" to "Thus, it is a genuine case of *X*." Some reasons for doubt about this transition have already been mentioned. It should be pointed out, however, that there are times when the transition is legitimate, although the paradigm-case argument can draw no comfort from this fact.

Suppose, for example, that someone doubted that there are any bachelors but admitted that there are unmarried males of marriageable age. We might naturally say to him, "But this is just what we call 'being a bachelor.'" Here, however, the doubter has no reply (other than to question whether this *is* how the word is used) because *this* refers to a description that logically entails "being a bachelor." In the paradigm-case argument, however, especially where the case is actually pointed out instead of described, no such entailment is normally claimed.

If there is not an entailment, however, then there seems room for the skeptic to maneuver. How can one hold that no matter what the skeptic's reasons may be, he must admit *this* as an instance of what he doubted to exist? Faced with such difficulties, some proponents of the paradigm-case argument have placed restrictions on its use. They have said that it is valid only for expressions designating concepts that must be taught ostensively—that is, taught through examples. Philosophers have often held, for example, that color words can be taught only in this fashion. The usual reason given is that the concept of a particular color is simple and that its meaning cannot be captured by a verbal definition. Hence, it must be taught by pointing out things that are of that color. When the paradigm-case argument is confined to such concepts, a special reason is supplied for why there must be indisputable instances. If there were not (or had never been) any red objects, how could the concept get into the language?

The appeal to what must be taught ostensively is frequently presented as if it were merely an elucidation of the force of the paradigm-case argument. But it seems, instead, to be a separate and distinct form of argument. There is, for example, no need to describe or point out particular circumstances. The conclusion that there are instances of, for example, red objects is drawn directly from the premise that the concept can be taught only ostensively. There would, perhaps, be point in calling this form of argument by a different name.

ARGUMENT FROM OSTENSIVE TEACHING. Whether such an argument is valid against a skeptic will depend upon several questions that have yet to be conclusively answered. First, are there any concepts that can be taught only ostensively? Is it logically impossible for someone to have the concept of, for example, redness without having obtained it through ostensive teaching? Second, even if a concept must be taught through such methods, must there be exemplifications of the concept? It seems possible, for example, to teach someone the meaning of "is red" by using objects that merely appear to be red as long as this fact is concealed from the student. Third, even if

the answer in the above cases is affirmative, are the important concepts that give rise to skepticism of the required kind? Is the concept of choosing freely, for example, one that can be taught only by such methods?

Sometimes it is said that the paradigm-case argument need be confined only to those concepts that can be taught ostensively. When this is done, no conclusion can be immediately drawn about the existence of cases falling under the concept. The concept of a unicorn could be taught ostensively if only there were such a creature, but as things stand, it never has been. What, then, is the value of such a restriction? The idea seems to be that if a concept can be taught ostensively, then there must be conceivable circumstances, at any rate, in which something falls under the concept—those circumstances in which it could be taught in this fashion. Such an argument, in general, has force only against an a priori skeptic. But it is possible that the circumstances in which, it is claimed, the concept could be taught ostensively actually occur and that the skeptic may not wish to dispute their existence. It might be urged, for example, that the concept of acting freely can be taught ostensively in circumstances which the skeptic about freedom would have to admit do occur. Some of the same problems about ostensive teaching arise for this kind of argument as for the previous one.

EVALUATIVE CONCEPTS. Still another restriction on the use of the paradigm-case argument has been proposed by some writers. J. O. Urmson questions the legitimacy of applying it to evaluative expressions such as "good (inductive) reasons" ("Some Questions concerning Validity"). His point is that the use of evaluative expressions has a dimension that the use of purely classificatory expressions lacks. Evaluative expressions not only sort out things and situations but also signify approval or condemnation. The skeptic, therefore, may be willing to grant that there are differences between what we call, for example, "good inductive reasons" and "bad inductive reasons" and that he has said nothing to show that these differences are not exemplified. But he may question whether these differences support our approval of the one and our rejection of the other. Thus, to take Urmson's analogy, he may grant a difference between what we call "good apples" and what we call "bad apples" but urge that our standards are faulty. How can pointing out that *this* is just what we call a "good apple," he may ask, show that we would not do better to approve of some other kind?

TWO SORTS OF SKEPTICISM. Urmson's point, if valid, appears to have many consequences. The dispute concerning whether we can exercise genuine freedom of choice about our own actions does not seem on the surface to be a dispute involving evaluative concepts. Philosophers, however, have been particularly uneasy about the use of the paradigm-case argument in this area, in contrast, for example, to its employment against skepticism about the existence or perception of material objects. The explanation may be that there are two sorts of skepticism involved. It may be that the skeptic about human freedom is not, in fact, denying that many of the ordinary relevant expressions mark genuine distinctions but, rather, querying the purpose to which we put these distinctions. In contrast, the skeptic about the existence of material objects does appear to deny that there is, for example, a distinction between a material object and the mere appearance of one.

We contrast seeing material objects with seeing hallucinatory or imaginary objects. By describing circumstances in which we ordinarily are in no doubt about which member of these distinctions is present, the paradigm-case argument may be construed as pointing out that the everyday expressions do, after all, serve a function. The fact that we do make these contrasts in practice and, more importantly, that we generally agree in our judgments shows that some genuine distinction is being made. Moreover, the skeptic does not usually dispute the fact that we can independently reach agreement about particular cases. Thus, it might be said to him, "Whatever your arguments to show that we never see material objects, for example, after we have looked at them and debated them, there will still be that difference between what we have called 'a real object' and what we have called 'hallucinations,' 'illusions,' or 'imaginary objects.' We shall still need to mark that distinction and so return to our usual way of describing things."

While this seems quite powerful against, for example, skepticism about the perception of material objects, the same sort of explanation of the paradigm-case argument is not so convincing when tried out on disputes about evaluative terms or the existence of genuinely free choices. The trouble may be that although the skeptic's arguments cannot destroy the correctness of contrasting what we should call cases of freely choosing from those we should not, his argument may still destroy what we thought to be the point of making the distinction. To say that a choice was free often involves the ascription of responsibility and the possibility of praise and blame. We behave differently toward persons who have made a free choice than we do toward those who have been coerced. If we knew all our "choices" to be the product of prior conditioning or hereditary traits—a possibility that

appears often to generate skepticism about our freedom—would we still be on solid ground in behaving differently toward those who have made a "free choice"? Although we could continue to make the same distinctions we do now as far as classification goes, we might think that to call certain choices "free" would have a hollow ring.

Whatever the ultimate verdict on the paradigm-case argument as a refutation of skepticism, there can be no doubt that its use in recent philosophy has generated very important questions about the relationship of language to the world.

See also A Priori and A Posteriori; Common Sense; Induction; Knowledge, A Priori; Moore, George Edward; Philosophy of Language; Russell, Bertrand Arthur William; Skepticism, History of; Wittgenstein, Ludwig Josef Johann.

Bibliography

BACKGROUND

Austin, J. L. "Other Minds." In *Essays on Logic and Language*, edited by Antony Flew. 2nd series. Oxford: Blackwell, 1953.

Chappell, V. C. "Malcolm on Moore." *Mind* 70 (1961): 417–425.

Malcolm, Norman. "George Edward Moore." In his *Knowledge and Certainty*. Englewood Cliffs, NJ: Prentice Hall, 1963.

Malcolm, Norman. "Moore and Ordinary Language." In *The Philosophy of G. E. Moore*, edited by Paul A. Schilpp. 2nd ed. New York, 1952.

Moore, George Edward. *Philosophical Papers*. New York: Macmillan, 1959. See especially "A Defence of Common Sense," "Proof of an External World," "Four Forms of Scepticism."

Moore, George Edward. *Some Main Problems of Philosophy*. London: Allen and Unwin, 1953. See especially Chs. 1, 5–6.

White, A. R. *G. E. Moore*. Oxford: Blackwell, 1958.

Wisdom, John. "Philosophical Perplexity." In his *Philosophy and Psycho-analysis*. Oxford: Blackwell, 1953.

Wittgenstein, Ludwig. *The Blue and Brown Books*. Oxford: Blackwell, 1958.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G. E. M. Anscombe. New York: Macmillan, 1953.

APPLICATIONS AND CRITICAL DISCUSSIONS

Alexander, H. G. "More about the Paradigm Case Argument." *Analysis* 18 (1958): 117–120.

Beattie, Catherine. "The Paradigm Case Argument: Its Use and Abuse in Education." *Journal of Philosophy of Education* 15 (1981): 77–86.

Black, Max. "Paradigm Cases and Evaluative Words." *Dialectica* 27 (1973): 262–272.

Black, Max. "Making Something Happen." In *Determinism and Freedom*, edited by Sidney Hook. New York: New York University Press, 1958. Application to freedom of the will and causation.

Bouwsma, O. K. "Descartes' Evil Genius." *Philosophical Review* 58 (1949): 141–151. Application to skepticism about the external world.

Butchvarov, Panayot. "Knowledge of Meanings and Knowledge of the World." *Philosophy* 39 (1964): 145–160.

Danto, Arthur C. "The Paradigm Case Argument and the Free-Will Problem." *Ethics* 69 (1959): 120–124. Critical.

Edwards, Paul. "Bertrand Russell's Doubts about Induction." In *Essays on Logic and Language*, edited by Antony Flew. 1st series. Oxford: Blackwell, 1951. Application to skepticism about induction and a sympathetic analysis.

Eveling, H. S., and G. O. M. Leith. "When to Use the Paradigm Case Argument." *Analysis* 18 (1958): 150–152.

Findlay, J. N. "Time: A Treatment of Some Puzzles." Ibid. Application to skepticism about the passage of time.

Flew, Antony. "'Farewell to the Paradigm-Case Argument': A Comment." *Analysis* 18 (1957): 34–40. Defense against criticism by Watkins.

Flew, Antony. "Philosophy and Language." In *Essays in Conceptual Analysis*, edited by Antony Flew. London: Macmillan, 1956. Sympathetic.

Hanfling, Oswald. "What's Wrong with the Paradigm Case Argument?" *PAS* 91 (1991): 21–38.

Harre, R. "Tautologies and the Paradigm Case Argument." *Analysis* 18 (1958): 94–96.

MacIntyre, A. C. "Determinism." *Mind* 66 (1957): 28–41. Contains criticism of application made to freedom of the will.

Mackie, J. L. *Contemporary Linguistic Philosophy—Its Strength and Weakness*. Dunedin, New Zealand: University of Otago, 1956.

Malcolm, Norman. "Moore and Ordinary Language." In *The Philosophy of Bertrand Russell,* edited by Paul A. Schlipp. Sympathetic analysis and several applications.

Nagel, Ernest. "Russell's Philosophy of Science." In *The Philosophy of Bertrand Russell*, edited by Paul A. Schilpp. Evanston, IL: Open Court, 1944. Application to several of Russell's views.

Passmore, John. *Philosophical Reasoning*. London: Duckworth, 1961. See Ch. 6 for a critical analysis.

Richman, Robert J. "On the Argument of the Paradigm Case." *Australasian Journal of Philosophy* 39 (1961): 75–81. Critical.

Richman, Robert J. "Still More on the Argument of the Paradigm Case." *Australasian Journal of Philosophy* 40 (1962): 204–207.

Sosa, E. "The Paradigm Case Argument: Necessary, Causal, or Normative." *Methodos* 15 (1963): 253–273.

Stebbing, L. Susan. *Philosophy and the Physicists*. London: Methuen, 1937. See Ch. 3 for application to skepticism about the properties of material objects.

Stroud, Barry. *The Significance of Philosophical Skepticism*. Oxford: Oxford University Press, 1984.

Urmson, J. O. "Some Questions concerning Validity." In *Essays in Conceptual Analysis*, edited by Antony Flew. London: Macmillan, 1956. Critical of application to evaluative concepts.

Watkins, J. W. N. "Farewell to the Paradigm-Case Argument." *Analysis* 18 (1957): 25–33. Critical.

Watkins, J. W. N. "A Reply to Professor Flew's Comment." *Analysis* 18 (1957): 41–42.

Will, F. L. "Will the Future Be Like the Past?" In *Essays on Logic and Language*, edited by Antony Flew. 2nd series. Oxford: Blackwell, 1953. Application to skepticism about induction.

Williams, C. J. F. "More on the Paradigm Case Argument." *Australasian Journal of Philosophy* 39 (1961): 276–278. Defense of the argument against criticisms of Richman.

Keith S. Donnellan (1967)
Bibliography updated by Benjamin Fiedor (2005)

PARADOXES

See *Logical Paradoxes; Zeno of Elea*

PARANORMAL PHENOMENA

See *Parapsychology; Precognition*

PARAPSYCHOLOGY

Parapsychology is the modern name for what used to be called psychical research. The word is usually used in a narrow sense, as scientifically based research, but sometimes it is used more broadly to cover the whole range of the occult. The term *psi* is often used as a briefer equivalent. Psi phenomena are paranormal, that is, beyond the range of what is considered to be part of the ordinary world.

The Society for Psychical Research (SPR) was established in England in 1882 and is given credit for organizing systematic research in the English-speaking Western world. Many of its founders were distinguished intellectuals who were themselves spiritualists and interested in immortality. The American Society for Psychical Research was founded soon after. Though some of the earlier researchers did scientific studies, they more often conducted other kinds of investigations of psi, investigating ostensible cases of dramatic psi, and frequently working with mediums.

THE MODERN SCIENTIFIC ERA

The modern scientific era in parapsychology is usually credited to Joseph Banks Rhine, who established the first university laboratory devoted exclusively to experimental research on psi. In 1957 Rhine and others organized the Parapsychological Association, which twenty years later became, over much opposition, an affiliate member of the American Association for the Advancement of Science. Some parapsychologists, however, insist that the pursuit of psi by scientific methods is ill advised and advocate a return to the more traditional types of psychical research.

Parapsychologists work primarily on a common core of phenomena that include telepathy (mind-mind), clairvoyance (mind-matter, now called remote viewing), psychokinesis (PK; mental action on matter), precognition and retrocognition (direct awareness of future or past mental or material states), and often survival issues (disembodied existence and reincarnation). More specialized topics might include hauntings and apparitions, séances, poltergeists, dowsing, psychic healing, and near-death and out-of-the-body experiences, but probably not astrology, the Bermuda Triangle, UFOs, past-life regression, and alien abductions.

Psi research is commonly defined as the study of things and processes that go beyond the commonly accepted ways of interaction in the world. Parapsychology is unusual in that what it studies is defined primarily in negative terms. For example, extrasensory perception (ESP) is defined in terms of gathering information not by sensory means. Moreover, parapsychologists typically admit that they lack agreement on what psi is or how it operates, and some parapsychologists prefer to speak of their field as the study of a limited range of "anomalies," refusing to make positive claims that psi is an actual power of some sort.

IMPLICATIONS OF PSI

Clearly, the existence of psi would have enormous implications for Western philosophy, not only by extending the range of commonly accepted ways of interacting with the world but also by reinforcing dualistic and idealistic worldviews that have hitherto supported their critiques of science on non-psi grounds, that is, on the alleged failure of the dominant materialist paradigm to allow proper room for consciousness, including qualia, volitions, intentions, and logical reasoning. If it were shown that psi exists, the foundations of modern Western metaphysics would be shaken, most would say, overthrown.

C. D. Broad (1953) formulated the issue in terms of what he called "the basic limiting principles" of Western thought, which he said were justified either by self-evidence or by overwhelming and uniformly favorable empirical evidence. These principles, abbreviated, are that causation always works forward by acting through a continuous chain of events linking cause and effect, that mind acts on matter only through its own brain, that

mental activity is dependent on the brain, and that knowledge is acquired only through present sensations or communications. It is clear that psi would challenge all these principles and thus threaten the Western worldview.

Judged in terms of these limiting principles, psi is not only *para*normal but *anti*normal. Telepathy and clairvoyance imply that minds act directly on remote minds or material objects, bypassing the brain and violating the rule of continuous links in causation. PK would also violate that chain, providing mental action directly on remote objects, and both precognition and retrocognition would violate both forward causation and how knowledge is acquired. There is no question that psi is magical, judged by our basic limiting principles.

Some parapsychologists attempt to minimize the conflict between psi and the Western worldview by claiming that psi should be understood in terms of psychological laws rather than laws of physics by assimilating psi to such psychological connections as association or by pointing to altered states such as dream states or hypnosis that seem to facilitate psi production. They point out that psi fits in well with idealism, panpsychism, and typically Eastern philosophies that tend to understand nature in terms of mind rather than by understanding mind in terms of nature as in Western views. Also, some dualists point out that the mind-brain dualism itself violates the modern scientific paradigm and claim that ESP and PK (but not precognition) can be assimilated to mind-brain interaction, as an extended application of the powers that the mind uses to interact with its own brain (Dilley 1988).

PROBLEMS IN PSI RESEARCH

Parapsychologists are hampered in their research by the lack of a common body of theory as well as by not knowing how to produce psi on demand or predict how it will behave once it occurs. Without any firm basis for understanding psi, it is difficult to test alternative hypotheses. Few believe that psi can be controlled consciously, and some believe that psi is actually resistant to demonstration, sometimes called the *shyness effect*. Moreover, successful psi production seems to be related to belief in psi. Even when psi is produced successfully, investigators do not know for sure whether psi is coming from the subjects of an experiment, from the experimenter, from defects in the experimental design, or even from fraud. Skeptics point to an additional problem about psi that arouses their suspicion, that psi does not seem to affect ordinary experiments in physics laboratories or enable psychics to win steadily at casinos.

Critics of psi research claim that replication requirements demanded by modern science have not yet been met and that experimenters have not yet devised protocols that will guarantee positive results and can be obtained by independent investigators. Defenders of psi sometimes accept this charge, but reply that the unpredictability of psi prevents replication in the strong sense and that multiple demonstrations of psi by well-run experiments should constitute acceptable scientific evidence. Psi researchers continue to try to understand psi in the hopes of learning how to control it but progress has been disappointing, considering that more than a century has passed since the founding of the SPR.

Parapsychologists are unanimous that psi is incompatible with present materialism. They accept a wide range of metaphysical theories. There are a few, a vigorous minority, who think that psi can be reconciled with current science by massive revisions in the concepts of Western science. They point to various modifications proposed by physicists that could result in fitting psi into a revised physics. As has been already mentioned, some parapsychologists have turned to idealism, panpsychism, or various kinds of Eastern philosophy that better accommodate psi.

By far the more prevalent view is that psi should be understood in terms of metaphysical dualism, that ESP and PK are just extraordinary extensions of the powers that the mind uses to interact with its own brain. Opinions are divided whether telepathy is a third power, using unconscious levels of mind to connect conscious minds, or whether so-called telepathic phenomena can be reduced to ordinary mind-brain interactions. Henri Bergson once suggested that minds might be potentially omniscient and able to influence every object in the universe, but that brains limit the activity of psi to what is biologically and socially more useful. Both Broad and H. H. Price have made use of this model to explain why psi occurs only seldom.

CONTROVERSIES ABOUT PSI

Controversies over proper methods to be used in parapsychology also divide parapsychologists. Many parapsychologists believe that stories and anecdotes cannot be trusted and that the only reliable way to establish the existence of psi is by using the scientific method, while others believe that careful examination of anecdotes and other subjective reports can show the existence of psi and worry that the use of the scientific method stifles psi production.

Why cannot anecdotal evidence and the testimony of personal experience or the results of the kinds of investigations of early researchers be trusted? Such evidence has often turned out to be highly unreliable. Standards of evidence were often weak and many of even the strongest apparently evidential cases have been exposed as fraudulent or careless. It is commonly acknowledged that the history of psi research has been troubled by fraud, and some studies thought exceptionally thorough (such as those done by Samuel G. Soal) have been exposed as fraudulent. On the positive side, John Beloff (1993) presents a reasonably cautious survey of the case for psi, covering many important researchers and their subjects.

Those readers interested in the history of fraud should consult Paul Kurtz's *A Skeptic's Handbook of Parapsychology* (1985), which has ten chapters devoted to fraud, as well as George P. Hansen's "Deception by Subjects in Psi Research" (1990), which offers an extensive analysis of fraud. Faced with the problem of doubts about nonexperimental evidence, many parapsychologists have devoted themselves to gathering evidence for psi that will meet contemporary standards for scientific evidence and much has been accomplished since the 1990s.

THE CASE FOR PSI

Does psi exist? Opinions are widely variant. Popular opinion polls indicate widespread belief in psi in the general population. A poll of parapsychologists attending a Parapsychological Association meeting showed more than a 90 percent favorable response to the claim that psi exists, and more than an 80 percent favorable response to precognition. A poll of college faculty in 1979 produced a wide gap between humanities and arts faculty and psychologists. More than two-thirds of arts and humanities faculty answered affirmatively that psi was an established fact, a proportion similar to that of the general population, but only one-third of psychologists held that opinion. A poll of elite scientists in 1984 showed that only 4 percent thought that ESP was an established fact, with 25 percent thinking that ESP was a likely possibility. A large number expressed no opinion, but 10 percent thought that ESP was an impossibility.

There is no consensus on the existence of psi. Even some parapsychologists have become discouraged, either leaving the field or continuing to function as parapsychologists even though they do not believe in psi. Western philosophers and psychologists tend to reject psi, believing that a combination of fraud, careless investigation, gullibility, and wishful thinking (such as the wish for immortality) can account for the continued belief in psi. On the contrary, psi believers sometimes claim that skeptics reject psi because psi powers are intrinsically threatening and that the existence of psi would overthrow the reigning paradigm in Western thought. More cautious people on both sides claim that there are interesting cases that suggest psi and that there is evidence that supports the existence of psi, but that the case for psi is not yet conclusive.

However true it might be that earlier investigations failed to meet modern experimental standards, parapsychologists overwhelmingly claim that the available evidence is virtually conclusive and have claimed to provide evidence that meets even the most scrupulous standards. However, it is fair to say that the best that those recent experimental findings have provided is evidence of a low level of psi, that psi cannot yet be demonstrated on demand, and that psi still cannot be produced reliably or consistently by independent investigators.

However, for the first time in the history of psi research, it is possible that psi researchers can produce the kind of evidence that will be regarded by knowledgeable skeptics as constituting scientific evidence. In particular, there are three major lines of ongoing research efforts that prove interesting and that have been analyzed carefully by skeptics: studies using the Ganzfeld procedure, remote viewing experiments, and experiments involving efforts to affect random number generators. Many of the results of these studies are discussed by K. Ramakrishnan Rao (2001). Some of the skeptics most conversant with psi research have been impressed with these results but still have reservations. Interested persons should consult Ray Hyman and Charles Honorton (1986), Daryl J. Bem and Honorton (1994), and Hyman (1989). James E. Alcock (1990) presents a number of reservations about the scientific case for psi in general, including specific criticisms of remote viewing and random number generation studies. There is agreement on both sides that they need to be at least open to persuasion and that continuing studies are needed, especially studies done in independent laboratories. It should also be pointed out that studies subsequent to those referred to earlier have not been conclusive.

Many parapsychologists would argue that there is convincing evidence for psi in studies that seem to provide evidence for survival of bodily death. The best candidates for evidence are cited in the literature on "cross-correspondences" gathered by members of the SPR almost a century ago, and studies of well over two thousand putative "reincarnation cases." Gardner Murphy (1979) does a careful analysis of cross-

correspondences, and Stevenson (1987) cites some of the best cases for reincarnation. Hoyt L Edge et al. (1986) provide a careful presentation of issues related to survival. Paul Edwards (1996) offers one of the most comprehensive general attacks on reincarnation evidence, as well as on the character and competence of reincarnation scholars in general; however, be forewarned that he is known for his ad hominem attacks.

The cross-correspondences involved mediums whose trance writings and utterances were purported to be communicated by Frederic Myers, a classics scholar and one of the founders of the SPR, and by other deceased persons. The material lent itself to the interpretation that Myers was attempting to communicate to researchers, through different mediums who were separated by time and place, using bits of information and images that could be put together to provide a coherent set of references to the same classical myth. The case for survival was weakened by the fact that some of the participants were themselves classicists and might have inadvertently produced the data telepathically, and led some to prefer the "super-psi" hypothesis, so-named because the power of psi required to explain the phenomena surpasses any degree of psi that is reinforced by the experimental literature.

A recent development in survival research since the 1990s is the use of combination locks, set by believers who hope to use mediums to communicate the combinations that will open the locks. In one case so far, a lock has been opened by use of computer techniques, and survival researchers are putting their hopes on more sophisticated encryptions. As of 2005, no lock has been successfully opened by the proposed methods of disclosure.

SOURCES OF PSI

Generally, Western philosophers have been skeptical of psi, but there have been many who have vigorously defended it. Prominent among them are Henri Bergson, Charlie Dunbar Broad, C. J. Ducasse, James Hyslop, William James, C. W. K. Mundle, H. H. Price, Robert Almeder, Robert Brier, Stephen E. Braude, Hoyt L. Edge, and David Ray Griffin. There have also been some defenders among psychologists, prominently John Beloff, Irvin Child, Alan Gauld, Harvey J. Irwin, Gardner Murphy, William McDougall, and Charles Tart. Among the knowledgeable skeptical psychologists are James E. Alcock, Ray Hyman, and Charles E. M. Hansel, as well as Susan Blackmore and Richard Wiseman, who are former pro-psi proponents. The best-known anti-psi philosophers who have worked on the psi literature are Paul Edwards, Antony Flew, and Paul Kurtz. Also, the unclassifiable Martin Gardner is firmly among the unconvinced.

Besides the invaluable Proceedings of the SPR, there are a number of journals devoted entirely to psi phenomena, such as *Journal of Psychical Research*, *Journal of the American Society of Psychical Research*, *Journal of Parapsychology*, *European Journal of Parapsychology*, and *International Journal of Parapsychology*, all of which are reliable sources of the best in psi. Of note is the *Journal of Scientific Exploration*, which sometimes reports on psi topics and is generally pro-psi. The multidisciplinary *Journal of Consciousness Studies* sometimes gives coverage to psi and related issues. Two other journals deserve special mention: the nicely balanced but short-lived *Skeptical Zetetic* and the *Skeptical Inquirer*, which claims to maintain an open mind but is widely regarded as being openly hostile to psi in all of its forms. There are also journals, too numerous to mention, that are devoted to more limited phenomena usually included in parapsychology.

See also Bergson, Henri; Broad, Charlie Dunbar; Consciousness; Dualism in the Philosophy of Mind; Ducasse, Curt John; Idealism; Immortality; James, William; Materialism; Panpsychism; Philosophy of Mind; Precognition; Qualia; Reincarnation; Volition.

Bibliography

BOOKS AND ARTICLES ADVOCATING PSI

Almeder, Robert. *Death and Personal Survival: The Evidence for Life after Death*. Lanham, MD: Rowman & Littlefield, 1992.

Beloff, John. *Parapsychology: A Concise History*. New York: St. Martin's Press, 1993.

Beloff, John. *The Relentless Question: Reflections on the Paranormal*. Jefferson, NC: McFarland, 1990.

Bem, Daryl J., and C. Honorton. "Does Psi Exist? Replicable Evidence for an Anomalous Process of Information Transfer." *Psychological Bulletin* 115 (1) (1994): 4–18.

Broad, C. D. *Religion, Philosophy, and Psychical Research: Selected Essays*. London: Routledge, 1953.

Broad, C. D. *Lectures on Psychical Research: Incorporating the Perrott Lectures Given in Cambridge University in 1959 and 1960*. London: Routledge and Kegan Paul, 1962.

Dilley, Frank. "Mind-Brain Interaction and Psi." *Southern Journal of Philosophy* 26 (4) (January 1988): 469–480.

Edge, Hoyt L., et al. *Foundations of Parapsychology: Exploring the Boundaries of Human Capability*. Boston: Routledge, 1986.

Griffin, David Ray. *Parapsychology, Philosophy, and Spirituality: A Postmodern Exploration*. Albany: SUNY Press, 1997.

Irwin, H. J. *An Introduction to Parapsychology*. 2nd ed. Jefferson, NC: McFarland, 1994.

Murphy, Gardner. *Challenge of Psychical Research: A Primer of Parapsychology*. Westport, CT: Greenwood Press, 1979.

Rao, K. Ramakrishnan, ed. *Basic Research in Parapsychology*. 2nd ed. Jefferson, NC: McFarland, 2001.

Stevenson, Ian. *Children Who Remember Previous Lives: A Question of Reincarnation*. Charlottesville: University Press of Virginia, 1987.

Wolman, Benjamin E., ed. *Handbook of Parapsychology*. Jefferson, NC: McFarland, 1986.

BOOKS AND ARTICLES REJECTING PSI

Alcock, James E. *Science and Supernature: A Critical Appraisal of Parapsychology*. Buffalo, NY: Prometheus Books, 1990.

Edwards, Paul. *Reincarnation: A Critical Examination*. Amherst, NY: Prometheus Books, 1996.

Gardner, Martin. *Science: Good, Bad, and Bogus*. Buffalo, NY: Prometheus Books, 1981.

Hansel, C. E. M. *ESP and Parapsychology: A Critical Reevaluation*. Buffalo, NY: Prometheus Books, 1980.

Hansen, George P. "CSICOP and the Skeptics: An Overview." *Journal of the American Society for Psychical Research* 84 (January 1990): 25–80. Originally titled "Deception by Subjects in Psi Research."

Hyman, Ray. *The Elusive Quarry: A Scientific Appraisal of Psychical Research*. Buffalo, NY: Prometheus Books, 1989.

Hyman, Ray, and Charles Honorton. "A Joint Communiqué: The Psi Ganzfeld Controversy." *Journal of Parapsychology* 50 (1986): 351–364.

Kurtz, Paul, ed. *A Skeptic's Handbook of Parapsychology*. Buffalo, NY: Prometheus Books, 1985.

COLLECTIONS OF PHILOSOPHICAL ARTICLES ON PARAPSYCHOLOGY

Flew, Antony, ed. *Readings in the Philosophical Problems of Parapsychology*. Buffalo, NY: Prometheus Books, 1987.

French, Peter A., ed. *Philosophers in Wonderland: Philosophy and Psychical Research*. St. Paul, MN: Llewellyn, 1975.

Ludwig, Jan, ed. *Philosophy and Parapsychology*. Buffalo, NY: Prometheus Books, 1978.

Thakur, Shivesh C., ed. *Philosophy and Psychical Research*. London: Allen and Unwin, 1976.

Wheatley, James M. O., and Hoyt L. Edge, eds. *Philosophical Dimensions of Parapsychology*. Spring Field, IL: Charles C Thomas, 1976.

Frank B. Dilley (2005)

PARETO, VILFREDO

(1848–1923)

Vilfredo Pareto, the Italian economist, sociologist, and philosopher, was born in Paris, where his father, the Marchese di Pareto, a supporter of Mazzini, was living as a refugee. In 1858 the family returned to Italy, where Pareto received a mixed mathematical and classical secondary education. In 1870 he graduated with a degree in engineering from the Turin Istituto Politecnico. He embarked on a career with the Italian railways and soon became a director. He was deeply, though ambivalently, influenced by his father's involvement in radical politics. Throughout his life Pareto believed in the superiority of liberal free trade, but his disillusionment with the economic protectionism of the Italian government developed into a fierce hatred of the political and social side of liberal ideology, which he thought had resulted in indefensible economic policies. This hatred led Pareto into intemperate attacks on the government, which retaliated by banning his lectures, and Pareto was eventually forced to abandon his career in government service. At about this time he became acquainted with the mathematical economist Léon Walras, professor at Lausanne. In 1893 Pareto was appointed lecturer at Lausanne, and he succeeded to Walras's chair the following year. He lived in Switzerland for the rest of his life, eschewing political activity until Benito Mussolini's advent to power in 1922. The Fascists acknowledged a large debt to Pareto's writings and conferred numerous honors on him, but since he died after only one year of the Fascist regime, his considered attitude to it must be a matter of conjecture.

LOGICAL AND NONLOGICAL CONDUCT

Pareto's social thought was largely conditioned by his reactions to contemporary political developments in Italy. He claimed to provide an impartial presentation and explanation of the facts of social existence without commitment to any particular sectional interest. In fact, however, his writings constitute a violently polemical defense of economic liberalism and political and social authoritarianism. This gulf between his professions and his practice is ironically in tune with his skepticism about the extent of men's understanding of their own behavior. In his economic writings, *Cours d'économie politique* (2 vols., Lausanne, 1896–1897) and *Manuel d'économie politique* (Paris, 1909), he tried to prove mathematically that the system of free trade provides maximum social benefit. In *Les systèmes socialistes,* (2 vols., Paris, 1902), he attempted to refute the claims of socialism that it provided a superior solution to economic problems. But if the logical case for economic liberalism was as overwhelming as it seemed to Pareto, he had to show why it was not generally practiced. This led him from economics to sociology and to the distinction between logical and nonlogical conduct, which constitutes one of his most distinctive contributions to sociological theory.

Pareto introduced this distinction in the course of a discussion of the nature of a scientific sociology. His conception of "logico-experimental" science was largely Baconian, and his methodological desiderata for a scientific sociology were that all its concepts should have

strictly controlled empirical reference; that all its theories should be subject to rigorous experimental or observational control; and that all its inferences should follow with strict logic from the data. He set himself to show how these norms should be applied in the sociological investigation of the ideas and systems of thought current in a given society, which, because they bear "the image of social activity," are an important part of the sociologist's data. Pareto thought it important not to accept such ideas and theories at their holders' valuations but to ask two questions about them: (1) Are their explanatory claims justified by logico-experimental standards? (2) Why are they accepted, and what are the social consequences of this acceptance? The question of acceptance became particularly pressing for Pareto in the case of widely held theories that did not seem to measure up to logico-experimental criteria. He thus regarded the logical critique of sophistries as only a prolegomenon, although a necessary one, to the real problems of sociology.

Many of Pareto's own criticisms of sophistries, especially of those committed by his political opponents, are extremely cogent and witty. However, his general account of the distinction between sound explanation and sophistry is less satisfactory. He held that an action was logical if it was performed by the agent with the intention of achieving an empirically identifiable end, if it actually tended to result in the achievement of that end, and if the agent had sound logico-experimental grounds for expecting this end to result. He designated as nonlogical any action that failed to measure up to any of these diverse criteria, and proceeded to classify what seemed the most characteristic ways in which this failure could occur.

Pareto regarded economic activity directed at maximizing profit, clearheaded Machiavellian political activity, and scientific work as the three most important types of logical conduct. But he left largely unasked most of the fundamental philosophical questions to which such an account gives rise. In particular, unlike his contemporary Émile Durkheim, he did not investigate the possibility that established forms of social behavior are themselves presupposed by the concepts most fundamental to his account—concepts such as "empirical reference," "respect for logic," and "setting oneself an end." Pareto's important insight, however, contained in his idea of "nonlogical conduct," that there are many forms of activity concerning which it makes no sense to ask what reasons people have for performing them, could naturally have led to such an investigation, had Pareto been more of a philosopher and less of a brilliant political pamphleteer. His failure to press this line of inquiry impeded him from maintaining a clear distinction between nonlogical and illogical actions, and what he claimed to be a dispassionate account of the nature of social life became a massive polemical indictment of alleged human folly. It is also one of the roots of his uncritical acceptance of science as the mother and guardian of logic, notwithstanding his repeated attacks on worshipers of "the Goddess Science."

RESIDUES AND DERIVATIONS

If the reasons offered by men for many of their own actions are not logically compelling, a different kind of explanation seems to be needed. To find this explanation Pareto undertook a wide-ranging, but unsystematic and biased, historical and comparative survey of human social behavior. In the course of it, he claimed to detect a contrast between kinds of conduct that constantly recur with very little variation and those that are highly diverse and changeable. The former he labeled "residues," the latter "derivations." The variable elements, or derivations, prove to be the theories with which people attempt to justify their residues. The alleged persistence of the same residue, even after the agent's abandonment of the derivation that had been supposed to justify it, gave Pareto an additional reason for claiming that the derivation was not the real explanation of the existence of the residue.

This theory has obvious affinities with Karl Marx's concept of "ideology," with Sigmund Freud's "rationalization" (although Pareto seems to have been ignorant of Freud's work), and with Durkheim's "collective sentiments." Unlike these writers, however, Pareto offered no systematic account of why men have recourse to derivations, contenting himself with the observation that among the residues is to be found a tendency of men "to paint a varnish of logic over their conduct."

The theory of residues is similarly incomplete. His most consistently held view seems to have been that the residues are constants and must be accepted as brute facts. At times he said that they were determined by certain congenital psychological "sentiments," although he failed clearly to distinguish these from the residues themselves. Nor did he explain how sentiments differ from the "interests" that he supposed to underlie logical economic activities. At other times he suggested that residues change as a result of social conditions. "A number of traits observable in the Jews of our time, and which are ordinarily ascribed to race," he wrote, "are mere manifestations of residues produced by long centuries of oppression." Moreover, in his Machiavellian advice to statesmen to reinforce in their subjects those residues that are politically advantageous to themselves, by means of

propaganda in favor of suitable derivations, Pareto even implied that derivations could influence residues. Such difficulties stemmed largely from Pareto's failure to face the philosophical questions about the nature of logic that his theories should have led him to ask.

ELITES AND THE CYCLE OF HISTORY

The two classes of residues most important for Pareto's sociological theory were combinations and persistence of aggregates. Men dominated by combinations are the innovating, risk-taking experimenters, the "foxes," linked by Pareto with the economic class of speculators. At the other extreme are the "lions," dominated by persistence of aggregates, wedded to the status quo and willing to use force in its defense. These are to be found among the *rentier* class. Pareto thought that all societies are ruled by elites, composed of those naturally most able in the various forms of social activity. The balance between combinations and persistence of aggregates in the elites and the lower social strata respectively determines the general character of a society. Inconsistently with his insistence on the nonlogical character of value judgments, Pareto thought there was an objective distinction between healthy and decadent social states, a distinction strongly influenced by his own attachment to free trade and political authoritarianism. Elites must be enterprising and innovative but also ready to use force in defense of their authority. However, the latter propensity tends to hinder the "circulation of the elites," leading to an accumulation of ability among the masses. Alternatively, the former tendency may degenerate into a flabby humanitarianism that weakens authority. In either case, a revolution results, leading to government by new elites. Pareto's belief in the constant repetition of this process led him to a cyclical view of history.

See also Decision Theory; Durkheim, Émile; Freud, Sigmund; Marx, Karl; Philosophy of Economics; Sociology of Knowledge.

Bibliography

WORKS BY PARETO

Trattato di sociologia generale, 2 vols. Florence: Barbera, 1916. 2nd ed., 3 vols. Florence, 1923. Translated by A. Bongiorno and A. Livingston as *The Mind and Society,* 4 vols. London; Cape, 1935.

Sociological Writings. Translated by Derick Mirfin. New York: Praeger, 1966.

WORKS ON PARETO

Aron, Raymond. "Vilfredo Pareto." In *Main Currents in Sociological Thought,* Vol. 2. Translated by Richard Howard and Helen Weaver. New York: Basic, 1967.

Borkenau, Franz. *Pareto.* London: Chapman and Hall, 1936.

Bruni, Luigino. *Vilfredo Pareto and the Birth of Modern Microeconomics.* Cheltenham, U.K.: Edward Elgar, 2002.

Burnham, James. *The Machiavellians.* New York: John Day, 1943.

Curtis, C. P., and G. C. Homans. *An Introduction to Pareto: His Sociology.* New York: Knopf, 1934.

Parsons, Talcott. *The Structure of Social Action.* New York: McGraw-Hill, 1937.

Powers, Charles H. *Vilfredo Pareto.* Newbury Park, CA: Sage, 1987.

Sica, Alan. *Weber, Irrationality, and Social Order.* Berkeley: University of California Press, 1988.

Winch, Peter. "The Mind and Society." In *The Idea of a Social Science,* 2nd ed. Atlantic Highlands, NJ: Humanities Press, 1990.

Peter Winch (1967)
Bibliography updated by Philip Reed (2005)

PARFIT, DEREK

(1942–)

Derek Parfit is senior research fellow of All Souls College; a regular visiting professor at Harvard, New York University, and Rutgers; and a fellow of both the British Academy and the American Academy of Arts and Sciences.

Born in China and educated in England at the Dragon School and Eton, Parfit took his degree in modern history at Oxford University and later turned to philosophy. He is legendary as a mentor and for his acute monograph-length criticisms of manuscripts, as well as for his important contributions to ethics, practical reasoning, and metaphysics. Parfit is widely regarded as one of the most important contemporary philosophers.

Along with John Rawls's *A Theory of Justice*, Parfit's magnum opus, *Reasons and Persons*, helped turn ethics from a moribund and peripheral subject that largely focused on the meanings of moral terms into a vibrant and central philosophical topic. Brimming with ingenious examples, powerful arguments, and startling conclusions, it has significantly shaped the philosophical agenda, introducing into discussion a host of new topics, examples, and terminology.

In Part One, Parfit discusses the ways in which theories about morality and rationality can be self-defeating and also makes claims about rational irrationality, blameless wrongdoing, imperceptible harms and benefits, harmless torturers, and other mistakes in moral mathe-

matics. Part Two defends a theory of individual rationality, the Critical Present-aim Theory, which rejects both purely desire-based instrumental theories and a purely self-interested or egoistic theory. Parfit offers a new outlook on the old question of whether morality must lose out in a conflict with prudence or rational egoism. Parfit notes that rational egoism is a hybrid position, neutral with respect to time but partial with respect to persons. Correspondingly, it can be challenged from one direction by morality, which is neutral with respect to both persons and time, and from the other direction by a present-aim theory, which is partial with respect to both persons and time. Parfit suggests that rational egoism rests on an unstable middle ground that requires a firm distinction between persons and time that is metaphysically dubious. Of additional interest are Parfit's insights regarding the rationality of attitudes to time and time's passage.

In Part Three, Parfit propounds a reductionist account of personal identity, somewhat like the Buddhist *no-self view*. Appealing to a dazzling array of so-called puzzle cases involving hypothetical fission, fusion, and branch lines of different *selves* or *person-stages*, Parfit challenges widely held beliefs about the nature and importance of personal identity. Most assume that there is a *deep, further fact* that constitutes personal identity, a fact that must be all or nothing and that matters greatly in rational and moral deliberations. On Parfit's view, while the logic of identity is all or nothing, the relations that constitute personal identity over time are matters of degree, and sometimes there may be no answer to the question of whether a future self will be me. What matters in survival are physical and psychological continuities with the *right* kind of cause, where the right kind of cause, he provocatively suggests, might be any cause.

Part Four presents a host of puzzles and paradoxes regarding future generations. The Non-Identity Problem is raised by the fact that any choice between two social or economic policies will affect who it is who will later live. Even if one's choice between two such policies would greatly lower the quality of life of future generations, this choice may not be worse for any of the people who would later live since if one had chosen the other policy, these people would never have existed. Parfit here challenges the deeply held view that moral arguments should appeal to the interests of all of the affected people. Parfit argues that it is hard to avoid what he calls the Repugnant Conclusion, or the view that compared with the existence of billions of people whose quality of life is very high, it would be in itself better if there existed some much larger number of people whose lives would be barely worth living. Parfit also presents the Mere Addition Paradox, in which various plausible assumptions are shown to lead to a contradiction. These arguments profoundly challenge deep beliefs about moral and practical reasoning.

At the time of the writing of this entry, Parfit was completing a second book *Climbing the Mountain* that will be about Kant's ethics, contractualism, and consequentialism. In discussing Kant's Formula of Humanity, Parfit argues that although one should not *regard* other people merely as a means, whether one is acting wrongly never depends on whether one is treating people merely as a means. Parfit defends Kant's claim that one must never treat people in ways to which they could not rationally consent. He then argues that if one revises Kant's Formula of Universal Law and appeals to a view about rationality and reasons that is not desire based but value based, Kant's formula can provide the best version of contractualism.

On the standard moral map, there are two main kinds of systematic moral theory. One kind is consequentialist, with utilitarian theories as the best-known examples. The other kind is Kantian theories and various forms of contractualism, which are often presented as the main systematic alternative to all forms of consequentialism. This map, Parfit argues, should be redrawn. Of the different ways of thinking about morality, it is Kantian and contractualist theories that do most to support consequentialism. Kantians, contractualists, and consequentialists ought to conclude that, in John Stuart Mill's metaphor, they have been climbing the same mountain on different sides.

Parfit also argues that Kantian and contractualist theories should take less ambitious forms. These theories should be presented not as accounts of wrongness or of moral reasoning but as claiming to describe a higher-level property that can make acts wrong, under which ordinary wrong-making properties can be subsumed. There are, moreover, several kinds of wrongness; and the most important questions are not about wrongness, but about reasons.

Parfit believes that the best way to respond to skepticism about the possibility of ethical progress is to make some. Perhaps as much as any philosopher in the last 100 years, he has done so.

See also Ethics; Kant, Immanuel; Metaphysics; Rawls, John; Thinking.

Bibliography

Rawls, J. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

WORKS BY PARFIT

Reasons and Persons. Oxford: Oxford University Press, 1984.

"Overpopulation and the Quality of Life." In *Applied Ethics*, edited by Peter Singer. Pp. 145–154. Oxford: Oxford University Press, 1986.

"Equality or Priority?" Delivered as the Lindley Lecture at the University of Kansas, November 21, 1991. Copyright by the Department of Philosophy, University of Kansas, 1995. Reprinted in *The Ideal of Equality*, edited by Matthew Clayton and Andrew Williams, pp. 81–126 (New York: St. Martin's Press, 2000).

"The Unimportance of Identity." In *Identity*, edited by H. Harris, pp. 13–45. Oxford: Oxford University Press, 1995.

"Reasons and Motivation." *Proceedings of the Aristotelian Society* (supplement). 1997, pp. 99–130.

"Why Anything? Why This?" *The London Review of Books* (January 22; February 5, 1998): 24–27; 22–25.

"Experiences, Subjects, and Conceptual Schemes." *Philosophical Topics* 26 (1 and 2) (1999): 217–270.

"Rationality and Reasons." In *Exploring Practical Philosophy*, edited by Dan Egonsson, et al, pp. 17–39. Burlington, VT: Ashgate, 2001.

"Justifiability to Others." In *On What We Owe to Others*, edited by Philip Stratton-Lake, pp. 67–89. Oxford: Blackwell, 2004.

"Kant's Arguments for His Formula of Universal Law." In *The Egalitarian Conscience, Essays in Honor of G. A. Cohen*, edited by Christine Sypnowich, Oxford University Press (in press).

"Persons, Bodies, and Human Beings." In *Contemporary Debates in Metaphysics*, edited by John Hawthorne, Dean Zimmerman, and Theodore Sider. Oxford: Blackwell (in press).

Works about Parfit

Barry, B., et al. "Symposium on Derek Parfit's *Reasons and Persons*" *Ethics* 96 (1986): 703–872.

Dancy, J., ed. *Reading Parfit*. Oxford: Blackwell, 1997.

Temkin, L. "Intransitivity and the Mere Addition Paradox." *Philosophy and Public Affairs* 16 (2) (1987): 138–187.

Larry Temkin (2005)

PARKER, THEODORE

(1810–1860)

Theodore Parker, an American theologian and social reformer, was the grandson of Captain John Parker, who led the Lexington minutemen. Theodore Parker was born in Lexington, Massachusetts, and, except for scattered months of formal schooling during the winter, was almost entirely self-taught. Although unable to afford tuition, he was allowed to take the Harvard examinations, and in 1834 he was admitted to the Harvard Divinity School. He was ordained minister of a small parish in West Roxbury, Massachusetts, in 1837. In 1845, after he had become a controversial figure and commanded a large audience, his supporters created the 28th Congregational Society in Boston and later rented the Boston Music Hall, where Parker preached to one of the largest congregations in the country. He became equally famous as a scholar, preacher, theologian, and reformer. Parker died in Florence, Italy.

In his religious thought Parker's radicalism was partly instinctive and partly the result of environmental influences. In an autobiographical essay completed just before his death, Parker remembered how he had been taught as a boy to respect the voice of conscience as the "voice of God in the soul of man" and encouraged to develop a spirit of free inquiry "in all directions." His religious upbringing was extremely liberal, and when he entered upon his formal theological studies, he had not only rejected the doctrine of the Trinity but was already suspicious of the validity of miracles and the "infallible, verbal inspiration of the whole Bible." Profiting by the encouragement of the liberal Unitarian professors at Harvard, he began an intensive study of the Bible that ultimately led him to a knowledge of twenty languages and did much to confirm his earlier suspicions regarding biblical authority.

As a young minister Parker was a great admirer of William Ellery Channing and Ralph Waldo Emerson. He responded to Emerson's Divinity School Address with enthusiasm and was an anonymous contributor to the polemical pamphlet war that followed.

Parker's own religious philosophy was strongly influenced by Immanuel Kant and by the critical studies of such biblical scholars as Wilhelm Martin DeWette and theologians such as David Friedrich Strauss and Ferdinand Christian Baur. Academic study and his own religious experience convinced him that the foundation of religion was based on "great primal intuitions of nature that depend on no logical process of demonstration." The three most important were the intuition of God, the intuition of morality, and the intuition of immortality. Basing his theology on these facts of consciousness, Parker emphasized the infinite perfection of God and the perfectibility of man.

His ideas first received wide publicity in 1841, when he delivered an ordination sermon titled "The Transient and the Permanent in Christianity." In this sermon Parker contrasted the transiency of theology and Scripture with the permanence of the great moral truths of Christianity,

truths that depended for their validity not on the authority of Christ but on the voice of God in the human heart. Parker spoke as a Unitarian minister, but the reception he received from organized Unitarianism was as wrathful as Channing's reception had been at the hands of the Calvinists twenty years earlier. As his more conservative followers faded away, Parker developed his radical ideas at greater length in a series of lectures he published in 1842 as *A Discourse of Matters Pertaining to Religion.* The following year he published his own edition and translation of DeWette's critical study of the Old Testament, *Beiträge zur Einleitung in das Alte Testament.*

Emerson referred to Parker as "our Savonarola," and Parker's essay on transcendentalism is one of the clearest expressions that we have of the American rejection of the empirical philosophy of the Enlightenment. Modern scholarship has established, however, that Parker's transcendentalism was not identical with Emerson's, for Parker relied less completely on intuition and more on the critical study of history and theology.

Parker's extraordinary capacity for sustained scholarly endeavor was almost matched by his capacity for action. The "Absolute Religion" he advocated required the application of religious truth to social problems, and Parker often preached on such subjects as crime, poverty, temperance, and prostitution. Long before the proponents of the social gospel, Parker recognized the power of organized evil in the world and sought to marshal religious sentiment against it. He was inevitably drawn into abolitionism. A friend of Wendell Phillips and William Lloyd Garrison, he helped to lead the resistance to the Fugitive Slave Law in Boston and was a supporter of John Brown before Harper's Ferry.

Parker traveled widely on lecture tours, making about one hundred appearances a year during the last decade of his life. His influence on the public mind was at its peak just before his death.

See also Channing, William Ellery; Consciousness; Emerson, Ralph Waldo; Enlightenment; Intuition; Kant, Immanuel; Neo-Kantianism; New England Transcendentalism; Religion and Morality; Strauss, David Friedrich.

Bibliography

Parker's work is collected in *Theodore Parker's Works,* edited by Frances P. Cobbe, 14 vols. (London, 1863–1870). A centenary edition was published by the American Unitarian Association, 15 vols. (Boston, 1907–1911). Henry Steele Commager has edited *Theodore Parker: An Anthology* (Boston: Beacon Press, 1960).

Biographical studies include Henry Steele Commager, *Theodore Parker, Yankee Crusader* (Boston: Little Brown, 1936), John Dirk, *The Critical Theology of Theodore Parker* (New York: Columbia University Press, 1948), and John Weiss, *Life and Correspondence of Theodore Parker,* 2 vols. (New York: Appleton, 1864).

Irving H. Bartlett (1967)

PARMENIDES OF ELEA

(born c. 515 BCE)

Parmenides of Elea, the most original and important philosopher before Socrates, was born c. 515 BCE. He changed the course of Greek cosmology and had an even more important effect upon metaphysics and epistemology. He was the first to focus attention on the central problem of Greek metaphysics—What is the nature of real being?—and he established a frame of reference within which the discussion was to be conducted. The closely related problem of knowledge, which to a great extent dominated philosophy in the fifth and fourth centuries, was raised at once by his contrast between the Way of Truth and the Way of Seeming. His influence can be found in Empedocles, Anaxagoras, and the atomists; it is strong in most of Plato's work, particularly in the vitally important dialogues *Parmenides, Theaetetus,* and *Sophist.*

Plato in his dialogue *Parmenides* describes a meeting in Athens of Parmenides, Zeno, and Socrates. Parmenides was then about 65, Zeno about 40, and Socrates "very young." Though the meeting is probably fictitious, there is no reason why the ages should be unrealistic. Since Socrates died in 399, when he was about 70, and since he was old enough in Plato's dialogue to talk philosophy with Parmenides, the meeting would have to be dated about 450, making Parmenides' birth about 515. An alternative dating (Diogenes Laërtius, *Lives* IX, 23, probably from Apollodorus's *Chronica*) puts his birth about 25 years earlier, but this can be explained away.

Plato's remark (*Sophist* 242D) that the Eleatic school stems from Xenophanes is not to be taken seriously. Parmenides founded the school in the Phocaean colony of Elea in southern Italy, and its only other noteworthy members were his pupils Zeno and Melissus (the tradition that the atomist Leucippus was from Elea is probably false).

WRITINGS

The work of Parmenides is not extant as a whole. Plato and Aristotle quote a line or two; from later writers, par-

ticularly Sextus Empiricus and Simplicius, about 150 lines can be recovered. Parmenides wrote in hexameter verse. All the fragments seem to come from a single work, which may have been called *On Nature*; it is unlikely to have been very long, and the fragments may amount to as much as a third of the whole. The survival of a long consecutive passage of more than sixty lines (Fr. 8) is of the greatest importance; it is the earliest example of an extended philosophical argument.

The poem begins with a description of the poet's journey to the home of a goddess, who welcomes him kindly and tells him that he is to learn "both the unshakeable heart of well-rounded Truth, and the beliefs of mortals, in which there is no true reliability" (Fr. 1). The rest of the poem consists of the speech of the goddess in which she fulfills these two promises.

The interpretation of Parmenides is thoroughly controversial, and a short article cannot do more than offer one possible account, with a brief mention of the more important and plausible variants. In the interests of brevity many expressions of doubt have been omitted.

THE PROEM

Sextus Empiricus (*Adversus Mathematicos,* VII, 111ff.) quotes 32 lines that he asserts to be the beginning of Parmenides' *On Nature* (Fr. 1). The poet describes his journey in a chariot, drawn by mares that know the way and escorted by the Daughters of the Sun. The Sun Maidens come from the Halls of Night and unveil themselves when they come into daylight. There is a gateway on the paths of Night and Day, with great doors of which the goddess Justice holds the key. The Sun Maidens persuade Justice to open the gates for themselves and Parmenides, and they pass through. "The goddess" welcomes him kindly as a mortal man in divine company, shakes his hand, and sets his mind at rest by telling him that it is right and just that he should have taken this road. He must now learn both the truth and the unreliable beliefs of mortals.

Although few examples of contemporary poetry have survived for comparison, it is safe to say that this proem is a mixture of tradition and innovation. The "journey" of the poet is a literary figure closely paralleled in an ode by Pindar (*Olympian* 6). There, as for Parmenides, the journey is an image of the course of the song; the poet rides in a chariot, a gate has to be opened, the team knows the way, and the road is notably direct. The route followed by Parmenides' chariot, although straight and swift, is impossible to chart. The details are vague. What is clear is that the whole journey is nowhere on earth, but in the heavens, and that it begins in the realm of darkness and ends in the realm of light. This imagery is confirmed by other indications—the escort of Sun Maidens and their unveiling.

It can hardly be doubted that the journey symbolizes progress from ignorance to knowledge on a heroic or even cosmic scale. The epic verse form signifies a deliberately heroic context, for earlier philosophers probably wrote in prose (though Parmenides may also have chosen verse as being more memorable). Parmenides' journey in search of knowledge must recall Odysseus's journey to Hades (*Odyssey* XI) to get directions from Teiresias to guide him on his way home. The location of Parmenides' journey recalls the magic regions of this part of the *Odyssey,* where in one place dawn follows immediately upon nightfall because "the ways of night and day are close together" (X, 86) and where in another place there is no daylight at all, since Night envelops everything (XI, 19). There may also be reminiscences of the journey of Phaethon in the chariot of the Sun.

Sextus, after quoting Fragment 1, gives a detailed allegorical interpretation of it, and in this he has been followed by some modern scholars. But this is wrong; it is impossible to trace a consistent allegory, and in any case detailed allegory was a later invention.

The identity of the goddess is puzzling. The wording of the proem itself suggests that she is the same as the goddess Justice who holds the keys of the gates; in a later fragment, however, she speaks of Justice in the third person (possibly even in Fr. 1.28; certainly in Fr. 8.14). It may be that Parmenides left the identification intentionally vague. Simplicius does not mention the goddess at all but introduces his quotations as if the first person referred to Parmenides himself. The Neoplatonists appear to have called her "the nymph Hypsipyle" (that is, High Gate; Proclus, "Commentary on the *Parmenides*" Book IV, Ch. 34).

It is probably wrong to say that in his proem Parmenides is setting himself up as a mystic or that he is claiming to have received a divine revelation. If mysticism entails some privileged access to truth through nonrational means, then Parmenides was no mystic. The fragments show that he argued for his conclusions; his goddess tells him to use his reason to assess her words (Fr. 7.5). A single visionary experience is ruled out by the opening of the proem, in which the tenses show that the journey is a repeated one—perhaps repeated every time the poem is recited. Unless the claim of every poet to be inspired by the Muses is itself a claim to a divine revelation, this seems to be an inappropriate description of Parmenides' experience.

At the time of its composition, the proem was probably understood as a claim that the poet had something of great importance to say. The course of his divinely inspired song was a path that led to the light of knowledge. By making Justice responsible for opening the gate for him, he claimed that this was a right and proper path for him to follow and, therefore, a path that led to truth. By putting the whole of his doctrine into the mouth of a goddess, he claimed objectivity for it; it was not beyond criticism, since the goddess instructed him to judge it by reason, but it was not to be regarded as a merely personal statement by Parmenides.

THE THREE WAYS

The goddess begins by telling Parmenides what are the only possible ways of inquiry. She describes three ways, produces reasons for ruling out two of them, and insists on the remaining one as the only correct one.

First two ways are stated, each being defined by a conjunctive proposition. The first is "that it *is,* and cannot not be; this is the way of Persuasion, for she is the attendant of Truth." The second is "that it *is not,* and must necessarily not be, this I tell you is a way of total ignorance" (Fr. 2).

The literal meaning of Parmenides' Greek in these propositions is hard to see. The verb "to be" is used in the existential sense. He uses it in the third person present indicative without any subject expressed. Some interpreters say that there is no subject to be understood; however, without any subject the sentence is incomplete, and no doubt the impersonal subject "it" is to be regarded as contained in the verb, as it often is. What this "it" refers to has to be derived from the rest of the argument and will be discussed shortly.

Immediately after the statement of the first two ways, the second way is ruled out on the ground that it is impossible to know or to utter what does not exist: "Whatever is for thinking and saying *must* exist; for it can exist [literally, 'is for being'], whereas nothing cannot" (Fr. 6). The line of thought seems to be that the object of thought *can* exist, and since "nothing" cannot exist, the object of thought cannot be nothing. But it must either exist or be nothing; hence, it *must* exist. The basic premises then are that "nothing" is nonexistent (presumably regarded as tautological) and that the object of thought *can* exist (that is, it is possible to think of something).

Parmenides makes it quite plain, by the use of inferential particles, that there *is* an argument in this passage (though this has been denied) along the lines described. It is therefore legitimate to fill in the basic proposition of the Way of Truth ("it is") from the grounds on which it is based. The unexpressed subject of this proposition must be "the object of thought or knowledge" (this is convincingly shown by G. E. L. Owen, "Eleatic Questions"). The Way of Truth will therefore show what can be said of a thing if it is to be a proper object of thought; the first step is to assert that it must *be,* that it should not *be* is unthinkable. Subsequently, the subject is referred to as *τὸ ἐόν* ("that which is," "what is real," "what exists").

After ruling out the second way, the goddess continues with a warning against a third way, the way followed by mortal men, who wander about senselessly, knowing nothing and getting nowhere. Their characteristic error is that they have made up their minds that "to be and not to be is the same and not the same" (Fr. 6). The third way can be identified with "the beliefs of mortals" mentioned at the end of the proem and discussed in detail in the main body of Parmenides' work, after the Way of Truth (this identification is often denied). Mortals treat existence and nonexistence as the same in that they attach them both to the same objects by supposing that things sometimes exist and sometimes do not (that is, that there is change) and by supposing that some things exist that contain less of being than others and therefore contain some nonexistence (that is, that there is difference). They treat them as not the same in that they suppose they have different meanings. The language in which the censured doctrine is expressed is reminiscent of Heraclitus, but Heraclitus is certainly not the only mortal who suffers from Parmenides' lash here.

The third way is ruled out by pointing to an alleged contradiction in it. It asserts that "things that are not, are" (Fr. 7). From the arguments of the recommended way, described later, it would appear that what is objectionable in the third way is its assumption of intermediate degrees of existence, of things that exist at one time but not another, at one place but not another, or in one way but not another. Ordinary habits of speech and the data of sense perception would lead a man along this path; the goddess gives a warning to "judge by *reason* the hard-hitting refutation that I have uttered."

THE WAY OF TRUTH

The Way of Truth has now been shown by elimination to be the right way. The long Fragment 8 proceeds to make deductions from the basic proposition that "it" (the object of thought and knowledge if the analysis given above is correct) "exists and must exist."

Its first property is that it is ungenerated and indestructible. It cannot have come into being out of what does not exist since what does not exist is absolutely unthinkable and since there would, moreover, be no explanation of why it grew out of nothing at one time rather than another. There is no growth of what exists (and no decay either, but Parmenides offers no separate argument for that); hence, "either it is or it is not" (Fr. 8.16)—and that decision has already been made. It *is,* as a whole, entirely.

Since there is no growth or decay of what exists Parmenides argues that no distinctions can be made within it. There are no degrees of being—differences of density, for instance; the whole is full of continuous being. What exists is single, indivisible, and homogeneous. Here Parmenides apparently moves from the temporal continuity of being to its spatial uniformity; in the same way Melissus, his pupil, argues for the absence of a beginning or end in time and then assumes the absence of a beginning or end in space (Melissus, Frs. 3–4).

Next follows an assertion that since there is no generation or destruction, there is no motion or change in what exists. This argument is expanded by Melissus (Fr. 7). Any form of change or rearrangement implies the destruction of a state of affairs that exists and the generation of one that does not exist. Thus, Parmenides concludes that what exists "remains the same, in the same … held fast in the bonds of limit by the power of Necessity" (Fr. 8.29). It already is whatever it can be. Motion, as a species of change, is apparently denied by the same argument.

The last section of the Way of Truth is particularly difficult. Parmenides repeats his assertion that there is no not-being and there are no different degrees of being; what exists is equal to itself everywhere and reaches its limits everywhere. From this he concludes that it is "perfect from every angle, equally matched from the middle in every way, like the mass of a well-rounded ball" (Fr. 8.42–44). There is no agreement among modern scholars as to whether this is a literal assertion that what exists is a sphere (a view held by John Burnet and F. M. Cornford) or only a simile indicating that it is like a ball in some respect other than shape (a view held by H. Fränkel and Owen). The latter view seems more probable. Parmenides' stress lies on the qualitative completeness, or perfection, of what exists, not on its spatial extension. The point of the simile might be put like this: As a ball is equally poised about its center so that it would make no difference which direction you took if you examined it from the center outward, so what exists is all the same from any center.

THE WAY OF SEEING

Having completed her account "about truth," Parmenides' goddess fulfills her promise to describe mortal beliefs. Only about forty lines survive from this part of the poem. The fundamental difference from the Way of Truth is made clear at the outset: Mortals give names to two forms, and that is where they are wrong, for what exists is single. They assume the existence of two opposites, Fire and Night, probably characterized in terms of sensible opposites such as hot–cold, light–dark, light–heavy, soft–hard. Using these two forms as elements, the Way of Seeming apparently offered a detailed account of the origin of the stars, sun, moon, earth and all the things on the earth "as far as the parts of animals" (Simplicius, *In de Caelo* 559.25), some embryology, sense perception, and doubtless other things. The details are unimportant (though Parmenides is credited with the first assertion that the morning star is identical with the evening star, according to Diogenes Laërtius, *Lives* IX, 23); the interesting and puzzling thing is that he should have added a cosmogony to the Way of Truth at all. Modern scholars differ about his intention.

Eduard Zeller took the cosmogony to be an account of the beliefs of Parmenides' contemporaries; Burnet called it "a sketch of contemporary Pythagorean cosmology." However, there is no evidence for this. Such a review would seem to be pointless, and in antiquity the cosmogony was recognized as Parmenides' own. One can ignore the suggestion that it represents those of his early beliefs that were later superseded. The discussion now turns on this point: Is the Way of Seeming granted relative validity as a sort of second best, or is it wholly rejected? If it is wholly rejected, why did Parmenides write it?

Recently, the first view has been defended as follows by, for example, W. J. Verdenius, Gregory Vlastos, Hans Schwabl, and W. R. Chalmers. The goddess in the prologue promised that Parmenides would learn about mortal beliefs as well as truth and would hardly have done so if they had no validity at all. Unless the phenomenal world is granted some degree of reality, the philosopher himself, the learner of truth, appears to be condemned to nonexistence; however, the mind, described in physical terms in the Way of Seeming (Fr. 16), is the faculty that grasps what is real in the Way of Truth. Moreover, some of the language of the Way of Seeming deliberately echoes that of the Way of Truth. The two opposites, Fire and

Night, transgress the canons of truth by being distinguished from each other, but they are each described as self-identical and as containing no nonexistence, like the real being of the Way of Truth (Frs. 8.57–59, 9.4). Later writers in antiquity, notably Aristotle (*Metaphysics* A5, 986b27–34), took Parmenides to be yielding to the necessity of providing his own account of the phenomenal world. For reason, Aristotle said, there was just one being, but for sense perception more than one. Others have argued that the Way of Truth is the way an immortal looks at the world sub specie aeternitatis, whereas the Way of Seeming is the way mortals see the same world in time. Many variations on these themes have been suggested.

The contrary view, defended recently in differing forms by Owen, A. A. Long, and Leonardo Taran, has more justification in the text of Parmenides. The goddess makes it clear enough that the Way of Seeming is wholly unreliable (Frs. 1.30, 8.52) and that the Way of Truth leaves no room whatsoever for intermediate degrees of reality. The text itself contains a statement of the intention: "Thus no judgment of mortals can ever overtake you" (Fr. 8.61; the metaphor is from chariot racing). Although this is ambiguous, the likeliest sense is that Parmenides is equipped by the Way of Seeming to defeat any mortal opinion about the phenomenal world. All descriptions of the phenomenal world presuppose that difference is real, but the Way of Truth has shown that what exists is single and undifferentiated. The transition to the Way of Seeming is made by pointing to the fundamental mistake in assuming even the minimum of differentiation in reality—that is, in assuming that two forms of what exists can be distinguished (Fr. 8.53–54). Once this assumption is made, a plausible description of the phenomenal world can be offered, but anyone who has followed Parmenides thus far will recognize the fundamental fallacy in even the most plausible description. This explanation is more consistent with the later history of Eleaticism, for Zeno and Melissus showed no interest in positive cosmology.

PARMENIDES AND GREEK PHILOSOPHY

There is general agreement that Parmenides followed the Milesians, Heraclitus, and Pythagoras and preceded Empedocles, Anaxagoras, and the atomists (the thesis of K. Reinhardt that Heraclitus answered Parmenides has been generally rejected). Ancient tradition credits him with a Pythagorean teacher, Ameinias (Diogenes Laërtius, *Lives* IX, 21). It is often said that the rigorous deductive method of the Way of Truth was learned from the mathematicians, who at that time in Italy were likely to be Pythagoreans, but the truth is that too little is known of the mathematics of the time to allow this to be more than a guess.

In general, the relevance of Parmenides to earlier philosophy is fairly clear, though there is room for doubt about his attitude toward individual men. (Various scholars have found in the text attacks on Anaximander, Anaximenes, Heraclitus, and the Pythagorean school.) All previous systems had assumed the reality of change in the physical world and attempted to explain it. Thales, Anaximander, and Anaximenes held that the world evolved from a simpler state into a more complex one. Anaximander's view was that different substances ("the opposites") grew out of a primitive undifferentiated "indefinite"; Anaximenes gave a more precise description of the manner of differentiation and said that the original substance, air, turned into other substances by rarefaction and condensation. Heraclitus apparently abandoned the idea of an original simple state, asserting that everything in the world is always changing—"an ever-living fire." In somewhat less materialistic language the Pythagoreans produced a cosmogony based on the imposition of limit upon the unlimited. Parmenides' critique was equally damaging to all of these theories, since his argument, if accepted, condemned all difference as illusory.

It is often said that Parmenides' attack on the reality of the physical world depends on his confusion of two senses of the verb "to be"—the existential and copulative. It cannot logically be true that a subject *is* and at the same time *is not* (existentially); from this Parmenides is supposed to have concluded that it cannot be true that a subject *is* black and at the same time *is not* white and hence that all differentiation is impossible. The surviving text does not bear this out. Parmenides' premise (and his fundamental fallacy) was, rather, that "what is not" is absolutely unthinkable and unknowable. Every change would involve the passage of what is into what is not, and hence every attempt to describe a change would involve the use of an unintelligible expression, "what is not."

The argument of the Way of Truth is metaphysical and would apply to any subject matter whatsoever; it is false to suppose that it applied only to Pythagorean cosmogony or only to the materialist cosmogonies of the Ionians. But that Parmenides' primary intention was to criticize the earlier cosmogonists seems clear from the addition of the Way of Seeming to the Way of Truth. His own Real Being was certainly not a ball of matter, as Burnet and others thought. On the other hand, it was not

something to which spatial terms were wholly inapplicable. It filled the whole of space and thus was in some sense a competitor of other accounts of the cosmos. The main effects of his work, too, were on cosmology.

The error of Parmenides' ways was not seen immediately, perhaps not until Plato's *Sophist.* Their immediate effect was to produce theories that attempted to save the natural world from unreality without transgressing Parmenides' logical canons. In brief, they produced theories of elements. Empedocles envisaged a cosmos made of the four elements that were later made standard by Aristotle—earth, water, air, and fire. He satisfied some of Parmenides' criteria by making his elements unchangeable and homogeneous. What he refused to accept from Parmenides was that difference was impossible without diminution of reality; his four elements were asserted to be different from one another yet equally real. He explained apparent change as the rearrangement in space of the unchanging elements. Anaxagoras went further to meet Parmenides by asserting that all natural substances, not just a privileged four, were elementary and unchangeable. The atomists responded in a different way; they accepted that no qualitative difference is possible but rescued the phenomenal world by asserting that "what is not" exists in the form of void—that is, as empty space separating pieces of real being from each other. (The equation of void with "what is not" is sometimes attributed to Parmenides himself, but it was probably first made by his follower Melissus, who explicitly denied its existence in his Fragment 7.)

Plato inherited from Parmenides the belief that the object of knowledge must exist and must be found by the mind and not by the senses. He agreed that the object of knowledge is not something abstracted from the data of sense perception but a being of a different and superior order. He differed, however, in that he allowed the sensible world to have an intermediate status, as the object of "belief," rather than no status at all (*Republic* 477B and elsewhere). He differed more significantly, too, in that he reimported plurality into the real and knowable by distinguishing different senses of "not-being" (*Sophist* 237B ff. and 257B ff.).

See also Anaxagoras of Clazomenae; Anaximenes; Aristotle; Atomism; Cosmology; Diogenes Laertius; Empedocles; Epistemology; Heraclitus of Ephesus; Leucippus and Democritus; Melissus of Samos; Metaphysics; Neoplatonism; Nothing; Plato; Pythagoras and Pythagoreanism; Quantum Mechanics; Sextus Empiricus; Simplicius; Socrates; Space; Thales of Miletus; Xenophanes of Colophon; Zeno of Elea.

Bibliography

Fragments and ancient testimonia are in *Fragmente der Vorsokratiker,* edited by H. Diels and W. Kranz, 10th ed. (Berlin, 1961), the standard collection. Fragments with English translation and commentary are in G. S. Kirk and J. E. Raven, *The Presocratic Philosophers* (Cambridge, U.K.: Cambridge University Press, 1957), but the commentary is rather inadequate. They are also found in Leonardo Taran, *Parmenides* (Princeton, NJ: Princeton University Press, 1965), and in Italian, with a long bibliography, in Mario Untersteiner, *Parmenide* (Florence: Nuova Italia, 1958).

The most important recent studies are H. Fränkel, "Parmenidesstudien," in his *Wege und Formen frühgriechischen Denkens* (Munich, 1955), and G. E. L. Owen, "Eleatic Questions," in *Classical Quarterly,* n.s., 10 (1960): 84–102.

Other studies are H. Diels, *Parmenides Lehrgedicht* (Berlin: G. Reimer, 1897); John Burnet, *Early Greek Philosophy* (London, 1892); K. Reinhardt, *Parmenides und die Geschichte der griechischen Philosophie* (Bonn, 1916); W. Kranz, "Über Aufbau und Bedeutung des Parmenideischen Gedichtes," in *Sitzungsberichte der Deutschen Akademie der Wissenschaften zu Berlin* 47 (1916): 1158–1176; Eduard Zeller, *Die Philosophie der Griechen,* 6th ed., Vol. I (Leipzig, 1919), Ch. 1; F. M. Cornford, "Parmenides' Two Ways," in *Classical Quarterly* 27 (1933): 97–111, and *Plato and Parmenides* (London: K. Paul, Trench, Trubner, 1939); G. Calogero, *Studi sul eleatismo* (Rome, 1932); and Harold Cherniss, *Aristotle's Criticism of Presocratic Philosophy* (Baltimore: Johns Hopkins Press, 1935).

Some more recent studies are W. J. Verdenius, *Parmenides: Some Comments* (Groningen, Netherlands, 1942); Olof Gigon, *Der Ursprung der griechischen Philosophie von Hesiod bis Parmenides* (Basel, 1945); Gregory Vlastos, "Parmenides' Theory of Knowledge," in *Transactions of the American Philological Association* 77 (1946): 66–77; Hans Schwabl, "Sein und Doxa bei Parmenides," in *Wiener Studien* 66 (1953): 50–75; C. M. Bowra, *Problems in Greek Poetry* (Oxford, 1953); Eric A. Havelock, "Parmenides and Odysseus," and Leonard Woodbury, "Parmenides on Names," in *Harvard Studies in Classical Philology* 63 (1958); W. R. Chalmers, "Parmenides and the Beliefs of Mortals," in *Phronesis* 5 (1960): 5–22; A. A. Long, "The Principles of Parmenides' Cosmogony," in *Phronesis* 8 (1963): 90–107; J. Mansfeld, *Die Offenbarung des Parmenides und die menschliche Welt* (Assen: Van Gorcum, 1964); and W. K. C. Guthrie, *A History of Greek Philosophy,* Vol. II (Cambridge, U.K.: Cambridge University Press, 1965).

David J. Furley (1967)

PARMENIDES OF ELEA [ADDENDUM]

David Furley's original entry remains an exemplary introduction to Parmenides' thought. Since its publication, philosophers have focused on the character of the routes of inquiry that the goddess lays out in the poem, suggest-

ing different interpretations of the subjectless is (or *esti*), and of the nature of *to eon*, the subject of inquiry. In addition, scholars have continued to study the *Proem* (the opening lines of the poem) and the *Doxa* (the goddesses' statement of mortal opinion), but there is no consensus about either.

Newer studies emphasize the undoubted influences of Homer and Hesiod (fl. c. 800 BCE) as models for Parmenides' language and poetic images, while others recognize the continuity of Parmenides' thought with that of his predecessors. For example, Xenophanes of Colophon questions whether human knowledge is possible: In the absence of divine warrant or intercession, how can human beings of limited experience achieve genuine understanding? Parmenides' analysis of the unchanging nature of the object of genuine thought and inquiry, and his use of a goddess who nevertheless uses arguments and demands that her hearer evaluate her claims (DK 28 B7.5) can be seen as an attempt to defend the possibility of human knowledge and explore its limits. Some scholars suggest that this account of Parmenides is too rationalistic, but the consensus remains that he is part of a philosophical tradition that continues in Plato, Aristotle, and later Greek thought.

Reading Parmenides as exploring the nature of inquiry and the proper object of understanding and knowledge, many scholars are more willing to countenance forms of "to be" in Parmenides that are not primarily existential. Attention has been paid to predicative, veridical, and fused predicative-existential notions of being, and it is likely that some sort of hybrid account best captures Parmenides' meaning. What-is (*to eon*) must exist, but existence is not Parmenides' primary concern. Rather, the object of genuine thought must be or have an essence (predicative), and must be what is the case (veridical). What is not (or lacks an essence) cannot be real. As such it cannot be an object of understanding. Contrary to mortal thinkers, Parmenides denies that coming-to-be and other sorts of change are real or can be attributed to what is real. The arguments of fragment 8 show that only what is wholly of a single kind, unchangingly and perfectly what it is, can be real. Such an entity (*eon*) is a unity, admitting none of what is not, and so can be grasped completely by thought.

There is no doubt that Parmenides claims that what-is is one. The question is the sort of unity or monism to which Parmenides is committed. Some scholars challenge the interpretation (going back to Plato) that Parmenides advocated numerical monism in the same sense as Melissus of Samos, who asserted the reality of only one thing. On the alternative account, although whatever there is must be one, more than one thing may be real. Stronger and weaker versions of this view have been taken. It can be argued that numerical pluralism is consistent with Parmenides' views of the unified nature of what-is, although Parmenides himself does not specify how many basic entities there are.

The role of the *Doxa* section of the poem remains a problem, especially if one follows many scholars in rejecting the view that mortals err by positing what does not exist or by supposing that there is a plurality of real things. There is no general agreement, and some modern interpretations accept the more traditional view, found in Furley's entry, that no cosmological account can be acceptable. Another suggestion is that, although the sensible world of change and becoming described in the *Doxa* is not the world of genuine reality, the cosmology of the *Doxa* nonetheless succeeds because it gives a true account and explanation of the unreal world of appearances. Or the *Doxa* might be intended as a lesson, offering a model cosmological account with a problem at its heart (the commitment to genuinely real opposite forms) that shows what must be avoided in an adequate account of how things are.

A further focus of study has been the positive importance of Parmenides' arguments for later philosophers (the later pre-Socratics and Sophists as well as Plato). This has led to a new appreciation of the Parmenidean basis for pluralistic and atomistic pre-Socratic theories and for the foundations of Plato's thought. In addition, scholars explore differences of theory and argument strategy among Parmenides, Zeno of Elea and Melissus, controverting the traditional interpretation that lumps them together as maintaining a single "Eleatic position."

See also Aristotle; Homer; Melissus of Samos; Plato; Pre-Socratic Philosophy; Sophists; Xenophanes of Colophon; Zeno of Elea.

Bibliography

Barnes, Jonathan. "Parmenides and the Eleatic One." *Archiv für Geschichte der Philosophie* 61 (1979): 1–21.

Caston, Victor, and Daniel W. Graham, eds. *Presocratic Philosophy: Essays in Honor of A. P. D. Mourelatos*. Burlington, VT: Ashgate, 2002

Coxon, A. H. *The Fragments of Parmenides*. Dover, NH: Van Gorcum, 1986.

Curd, Patricia. *The Legacy of Parmenides*. Rev. ed. Las Vegas, NV: Parmenides Publishing, 2004.

Gallop, David. *Parmenides of Elea: Fragments*. Toronto: University of Toronto Press, 1984.

Kahn, Charles H. "The Thesis of Parmenides." *Review of Metaphysics* 22 (1969/70): 700–724.

Lesher, J. H. "Parmenides' Critique of Thinking: The poludêris elenchos of Fragment 7." *Oxford Studies in Ancient Philosophy* 2 (1984): 1–30.

Long, A. A. "Parmenides on Thinking Being." In *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 12, edited by J. J. Cleary and W.C. Wians. Lanham: University Press of America, 1996.

McKirahan, Richard. *Philosophy Before Socrates*. Indianapolis: Hackett, 1994.

Mourelatos, A. P. D., ed. *The Pre-Socratics: A Collection of Critical Essays*. Princeton, NJ: Princeton University Press, 1993.

Mourelatos, A. P. D. *The Route of Parmenides*. New Haven, CT: Yale University Press, 1971.

Schofield, M. "Did Parmenides Discover Eternity?" *Archiv für Geschichte der Philosophie* 52 (1970): 113–135

Solmsen, F. "The 'Eleatic One' in Melissus." *Mededelingen der Koninklijke*. Nederlandse Akademie van Wetenschappen, Afd. Letterkunds, Nieuwe Reeks 32/8 (1969): 221–233; reprinted in Solmsen, Kleine Schriften III, 137–149.

Patricia Curd (2005)

PASCAL, BLAISE

(1623–1662)

Blaise Pascal was a French mathematician, physicist, inventor, philosopher, and theologian. He was born in Clermont in Auvergne, the son of a minor noble who was a government official. Pascal's mother died in 1626. In 1631 the family moved to Paris but fled in 1638 because of the father's opposition to the fiscal regulations of Richelieu. The next year Pascal's younger sister, Jacqueline, successfully acted in a children's play performed for Richelieu and thus gained a pardon for her father, who then became the royal tax commissioner at Rouen.

MATHEMATICS AND PHYSICS

Pascal was a prodigy, privately educated by his father, who was an excellent mathematician. His father wanted his son to have a good humanistic background before he learned mathematics and science, but at the age of twelve, Pascal discovered by himself the principles of geometry. When his father realized this, he abandoned his original plan for his son's education and encouraged his mathematical development. While still a teenager, Pascal published important mathematical and scientific papers and was a young prodigy in the Parisian intellectual circles. His father and he became members of a scientific discussion group organized by Father Marin Mersenne. There he would have met a wide range of people, probably including Thomas Hobbes, Descartes, and others. At sixteen, Pascal wrote his first major work, *Essai pour les coniques* (published in 1640), which his sister reported was "considered so great an intellectual achievement that people said that they had seen nothing as mighty since the time of Archimedes." In 1642 Pascal invented the calculating machine, originally designed to help his father in his tax work. This machine was one of the first applied achievements of the "new science." Pascal's writings on the calculating machine from 1645 to 1652 indicate the inordinate difficulties of putting theory into practice, the wide divergence between the levels of metallurgical and mathematical skill, and the monumental importance of this early contribution to the industrial revolution.

For the rest of his life Pascal continued to make major mathematical contributions in probability theory, number theory, and geometry. Although he gave up serious concern with mathematical problems after his religious conversion in 1654, a notable analysis of the nature of the cycloid grew out of a night's insomnia in 1658. Pascal's important work in the philosophy of mathematics, *L'esprit géométrique*, was probably written in 1657 and 1658 as a preface to a textbook in geometry for the Jansenist school at Port-Royal.

THE VACUUM

In 1646 Pascal learned of Evangelista Torricelli's (1608–1647) experiment with a barometer, which involved placing a tube of mercury upside down in a bowl of mercury. Having successfully repeated the experiment, Pascal asked himself what kept some of the mercury suspended in the tube and what was in the space above the column of mercury in the tube. Many scientists believed that the pressure of the outside atmosphere was responsible for holding up the column of mercury, but they had no proof. All agreed that the space at the top of the tube contained some kind of rarefied and invisible matter; hence, no vacuum. In 1647 Pascal published *Experiences nouvelles touchant le vide*, a summary of a series of experiments with variously shaped and sized tubes and different liquids, in which he set forth the basic laws about how much water and how much mercury could be supported by air pressure and about how large a siphon had to be to function. He also sketched out the reasons why a genuine vacuum could and did exist above the column of mercury or other liquid supported in the barometer.

Father Estienne Noel, rector of the Collège de Clermont in Paris, challenged Pascal, insisting that nature abhors a vacuum and therefore would not allow one to exist; thus, the alleged empty space created in Pascal's

experiments actually contained a special kind of matter. Pascal's reply, in which he gave the conditions for judging a hypothesis, is one of the clearest statements on scientific method made during the seventeenth century. Pascal asserted that a hypothesis could be disproved if one could elicit either a contradiction or a conclusion counter to fact from the affirmation of the hypothesis. However, if all the facts fit the hypothesis or follow from it, this merely shows the hypothesis is probable or possible. "In order to show that a hypothesis is evident, it does not suffice that all the phenomena follow from it; instead, if it leads to something contrary to a single one of the phenomena, that suffices to establish its falsity." Pascal showed that Noel's and Aristotle's hypothesis that there is no vacuum is false because conclusions contrary to experimentally established facts follow from it, whereas his own theory of a genuine vacuum is a possible or probable explanation of the facts in question.

In 1648 Pascal's brother-in-law performed the experiment of carrying a barometer up a mountain. This established the change in the level of the column of mercury. Pascal checked the results at various heights on a church tower in Paris. He then declared that these results established

> that Nature has no abhorrence of a vacuum, that she makes no effort to avoid it; that all the effects that are ascribed to this horror are due to the weight and pressure of air; … and that, due to not knowing this, people have deliberately invented that imaginary horror of a vacuum, in order to account for them.

Combining his ingeniously derived experimental data with a clear analysis of the possible explanatory hypotheses, Pascal arrived at one of the major achievements of seventeenth-century science. His theory of the vacuum and air pressure played an important role in the development of the mechanical theory of nature and the elimination of some of nature's alleged occult qualities and personal characteristics. The preface to the *Traité du vide* (which is all that has survived of the *Traité*) contains a defense of the new science and a discussion of the nature of scientific progress. In the study of nature, Pascal insisted that respect for authority should not take precedence over reasoning or experience (in theology, however, he maintained that it should). The secrets of nature, he said, are hidden from us, and although it is always active, we do not always discover its effects. In the course of time, through experience and understanding, we come to learn more about the natural world. Hence, as more data are accumulated, we should expect to find previously accepted hypotheses replaced by newer ones. Our conclusions about nature are always limited by the amount of experience gathered up to now. In time we seek for truths in terms of our experience and comprehension. What is sought for may be unchanging, but the results of the quest are the variable developments that constitute the history of science. Thus, there is no reason for preferring the ancient scientific views of Aristotle or anyone else to the latest achievements of scientific reasoning, based on the most recent data.

PASCAL AND JANSENISM

Pascal's mathematical and scientific accomplishments are among the most important of his time, but his religious and philosophical views have overshadowed them. His writings in religion and philosophy grew out of his involvement with the Jansenist movement. In 1646, after his father was injured, two Jansenists came to take care of him. The whole family, including Blaise, became interested in and involved with this Catholic reform movement, with his sister Jacqueline, becoming a nun at Port-Royal de Paris. From 1652 to 1654, Pascal turned away from religious interests, spending his time mainly with libertine friends who were gamblers, womanizers, and probably freethinkers. Pascal often visited his sister at Port-Royal, indicating to her that he had a great contempt for the world and people but that he did not feel drawn to God. However, after a traumatic experience crossing the Pont Neuf in Paris during a storm, Pascal had a religious conversion. He recorded this religious experience in *Le Mémorial* as "certitude, certitude, feeling, joy, peace." A year later, in 1655, with the encouragement of his sister, he made his first retreat at Port-Royal-des-Champs. Thereafter, Pascal objected vehemently to the philosophy of Descartes, unfavorably contrasting the God of the philosophers—namely, Descartes's God—with the God of Abraham, Isaac, and Jacob.

In January 1655 Pascal went to Port-Royal-des-Champs, the order of the two Port-Royal convents, for a two-week retreat. There a famous discussion with the Jansenist theologian, Isaac Le Maistre de Saci, took place, published in the *Entretien avec M. de Saci*. This text indicates that Pascal had already formulated many of the views later developed in the *Pensées*. During the next several months, Pascal often visited the two Port-Royal convents. On one of these visits Pascal met Antoine Arnauld, the leading Jansenist philosopher and theologian, who was about to be condemned by the Sorbonne for his views. In *Lettres provinciales*, a series of eighteen letters published in 1656 and 1657, Pascal defended Arnauld and

satirized his Jesuit opponents and their theological and moral view. These letters, published under the pseudonym Louis de Montalte, were probably the cooperative work of Pascal, Arnauld, and Pierre Nicole, though they were principally by Pascal. One of the great French literary masterpieces, the *Lettres provinciales* mercilessly ridicules the casuistry of various Jesuit moralists for what Pascal considered their lax, inconsistent, and unchristian views and defends Jansenism against charges of heresy. The arguments of various sixteenth-century and seventeenth-century scholastics are torn apart, and the charges against the Jansenists rebutted in a dazzling display of wit, irony, abuse, argument, and literary brilliance. Nevertheless, the *Lettres provinciales* was placed on the Index in 1657, and shortly thereafter the Jansenist movement was condemned by the pope. In 1661 the schools at Port-Royal were closed, and the nuns and solitaires had to sign a submission to the church.

Until 1659 Pascal worked on a wide variety of subjects defending Jansenism, composing his *Écrits sur la grâce*, *De l'esprit géométrique*, *De l'art de persuader*, and the works on the cycloid and preparing his *Apologie de la religion chrétienne*, the unfinished work posthumously published as the *Pensées*. In 1659, seriously ill, Pascal practically stopped writing. In 1660 he was somewhat better and wrote his *Trois discours sur la condition des grands*. The next year, after the suppression of Jansenism and the death of Jacqueline, Pascal wrote his final work on Jansenism, *Écrit sur la signature du formulaire*, urging the Port-Royalists not to give in. He then withdrew from all further controversy. His last achievement, illustrating another side of his genius, was the invention of a large carriage with many seats and the inauguration of what was in effect the first bus line, carrying passengers from one part of Paris to another for a fixed fare. One of his motives was to gain money to give to the poor, because he had already disposed of almost all his worldly possessions. Much of his will is devoted to bequeathing portions of his bus revenues to various hospitals.

PHILOSOPHY OF MATHEMATICS AND SCIENCE

Pascal left unpublished his two most important philosophical works, the *Pensées* and *De l'esprit géométrique*. *De l'esprit géométrique* was first published in the eighteenth century. In it Pascal dealt with the problem of the method for discovering truths. The ideal method, he declared, would be one which defined all of the terms employed and demonstrated all propositions from already established truths, but this is impossible, because the basic terms to be defined presuppose others to explain their meaning, and the fundamental propositions to be proved presuppose still others. Thus, it is impossible to reach first terms and principles. Instead, we find primitive terms that admit of no further definitions that clarify them and principles that are so clear that nothing clearer can be found to aid in proving them. "From which it seems that men are naturally and unalterably powerless to deal with any science whatsoever in an absolutely perfected manner."

Given this state of affairs, geometrical procedure is the most perfect known to humankind—a balanced one in which those things that are clear and known to everyone are not defined and everything else is defined, and in which those propositions known by all are assumed and other propositions are derived from them. Pascal insisted that this did not mean either that human beings could know by natural means that the premises of geometry were really true or that the fundamental concepts were thoroughly understood. Rather, the geometrical method provided the greatest certitude attainable by use of our limited capacities. Essentially, it developed an axiomatic system in which, from primitive terms and axioms, a set of propositions could be logically derived. Such a set would be true if the axioms were true.

In the companion piece to *L'esprit géométrique*, *De l'art de persuader*, Pascal explained how we come to be convinced of first principles and of conclusions from them. Conclusions are explained via the geometrical method. The problem of first principles raises a basic point for Pascal's theory of knowledge that is developed in the *Pensées*. Our reason and understanding can only work out axiom systems. Because we cannot prove the first principles, we can always cast skeptical doubts upon their truths, no matter how certain they may appear to us at various times. We can overcome this constant tendency toward skepticism (which also occurs in scientific research, because we can never know the secrets of nature but only plausible and as yet unrefuted hypotheses about the world) only by recognizing that principles are gained through instinct and revelation. This recognition requires admitting the importance of feelings and of submission to God in the quest for truth.

RELIGION

Pascal left the *Pensées* unfinished, with many notes of varying sizes pinned together. The first editors copied all the materials exactly as Pascal left them but published only those portions that they felt were completed, organizing them as they saw fit. Later editors assumed that the

Pensées was a collection of fragments, left in a disordered state by their author, and that each editor could arrange the fragments as he wished. Victor Cousin in 1842 pointed out that only selections of the *Pensées*, often somewhat embellished by the various editors, existed in print, and he urged a definitive edition based on the manuscripts in the Bibliothèque Nationale. One of these, the *Recueil original*, consists of the fragments in Pascal's own handwriting, pasted on large sheets of paper. For the next century editors used this manuscript for varying presentations of the text. In the 1930s and 1940s Zacharie Tourneur and Louis Lafuma established that the *Recueil* was pasted together after Pascal's death and that another manuscript, a copy by one of Pascal's relatives, represented the actual state of the work as organized and partially completed by the author. This led to Lafuma's definitive edition in 1952, which radically changed the order of the fragments, finally presenting the development of the themes in the *Pensées* as Pascal had intended them to be read.

THE HUMAN CONDITION. In the Lafuma edition the initial sections, "Order," "Vanity," "Misery," "Boredom," and "Causes of Effects," all portray the human condition by showing humankind's ways of dealing with and reacting to the ordinary world. The sixth and seventh sections turn to the core of humankind's philosophical problem—how to find truth and happiness. If humans are miserable, vain creatures, unable by their own resources to find first truths from which to derive others, they have to realize that "we know truth not only by reason but more so by the heart. It is in this latter way that we know first principles, and it is in vain that reason, which plays no part in this, tries to combat them" (Lafuma 1952, p. 110; Brunschvicg, p. 292). The principles of geometry are known instinctively by the heart, and reason employs these principles to establish theorems. Both heart and reason yield results that are certain, but by different routes, and it would be ridiculous to require proofs of the heart's instincts and intuitions or intuitive knowledge of what is proved. The inability of reason to establish first principles serves to humiliate reason but not to undermine our certainty. The realization of the limitations of reason helps us, Pascal declared, to recognize our wretchedness, and the greatness of humankind is that people alone are capable of such a recognition.

The climax of this attempt to show the ultimate nonrational foundation of our knowledge of first principles comes in the next section, "Contradictions." In a famous passage on skepticism (131 and 434) Pascal began by pointing out that the strongest contention of the Pyrrhonists was that we have no assurance of the truth of any first principles apart from faith and revelation except that we feel them within us. This natural feeling is no convincing proof of their truth, because apart from faith we cannot tell whether humans were created by a good God, an evil demon, or by chance. The truth-value of the principles depends upon their source. Pascal then explored the depths of complete skepticism and showed that if one had no assurance or any principles, one could be certain of nothing; but at the same time one could not even become a complete skeptic.

> What then will man do in this state? Will he doubt everything? Will he doubt whether he is awake? Whether he is being pinched, whether he is being burned, will he doubt that he doubts, will he doubt that he exists?
>
> We cannot go so far as that; and I set it forth as a fact that there has never been a complete perfect Pyrrhonist. Nature sustains our feeble reason and prevents it from raving to that extent. …
>
> What kind of a chimera then is man? What novelty, what monster, what chaos, what subject of contradictions, what prodigy? Judge of all things, imbecile worm of the earth, depository of truth, sink of uncertainty and error, glory and scum of the world.
>
> Who will unravel this tangle? Certainly it surpasses dogmatism and Pyrrhonism; and all human philosophy. …
>
> Nature confounds the Pyrrhonists and reason confounds the dogmatists. …
>
> Know then, proud man, what a paradox you are to yourself. Humble yourself, weak reason. Silence yourself, foolish nature, learn that man infinitely surpasses man, and hear from your master your real state which you do not know.
>
> Hear God.

The problem of knowledge thus becomes, for Pascal, a religious one. Only through submission to God and through acceptance of his revelation can we gain completely certain knowledge. The greatest achievements in science and mathematics rest on a fundamental uncertainty, because the basic principles employed, known through instinct and intuition, are open to question. Skeptical probing can only reveal the human predicament in its fullest and prepare us to submit and accept a religious foundation of knowledge.

The *Pensées* then proceeds to show how humans try to avoid recognizing their situation through diversion and philosophy. Philosophy can only lead us continually to skepticism, from which we are saved by our own intuitive knowledge of truth. We seek for happiness but cannot find it apart from religion. Pascal then tried to show in the famous wager argument (418 and 233) that it is not unreasonable to believe in God. God, he argued, is infinitely incomprehensible to us. But either God exists or he does not exist, and we are unable to tell which alternative is true. However, both our present lives and our possible future lives may well be greatly affected by the alternative we accept. Hence, Pascal contended, because eternal life and happiness is a possible result of one choice (if God does exist) and because nothing is lost if we are wrong about the other choice (if God does not exist and we choose to believe that he does), then the reasonable gamble, given what may be at stake, is to choose the theistic alternative. The person who remains an unbeliever is taking an infinitely unreasonable risk just because he or she does not know which alternative is true. Pascal's dialectic in his religious apologetics prods people to realize that there is not enough evidence to confirm the religious hypothesis and not enough to reject it. So, a person in his or her fallen state chooses on moral characters rather than philosophical ones.

Pascal is not just presenting the problem of human knowledge in philosophical terms. As he once explained to his fellow members at Port-Royal, what he was working on as the culminating statement of his views was "an apology for the Christian religion." The *Pensées* are either this apology or reflect a good deal of its content or design. The skeptical problems and the skeptical attitude are part of the apologetic project. But Pascal does not see skepticism as leading to religious knowledge or religious truth, but more as neutralizing man's rational impulses. Pascal was not following the route of Michel Eyquem De Montaigne, Pierre Charron, and Francois de La Mothe Le Vayer. He was using their skeptical weapons to combat the dogmatists and to make the skeptics aware of the religious dimension. Pascal did not see skepticism as leading to the relaxed, tranquil view of the ancient Pyrrhonists, but rather to a sharpened and heightened desperation. The desire to know could not be satisfied by human rational faculties but there was a necessity to know.

What Pascal contributes to the skeptical discussion is what José Maia Neto (1995) has called the "Christianization of Pyrrhonism." The Christianization of Pyrrhonism is seen in Pascal's description of people's state without God. This state, theologically, is what has happened to humankind in the Fall. Humans in this condition can find no security through reasoning or the use of their faculties, *and* they can unfortunately realize the desperation of their situation. They still have a glimmer or afterglow of the prelapsarian state of affairs but are unable to reach it. Pascal tried to show how belief can be achieved by curbing the passions, submitting to God, and using reason as a means of realizing that true religion is beyond reason and is known only through Jesus. We are suspended between two infinities, the infinitely small (the void) and the infinitely great (the Divine). Reason exposes our plight to us. Our desire for truth and happiness makes us see the futility of science, mathematics, and human philosophy as ways of finding the answers humans seeks.

THE CHRISTIAN RELIGION. The later sections of the *Pensées* are devoted to apologetics, arguing that the Christian religion is the true religion. From historical data, moral precepts, miracles, and the fulfillment of prophecies, Pascal argued that the Bible is the source of true religious knowledge. He contended that the Old Testament foretold Christ's coming and the Jewish rejection of him. Using the recently rediscovered Spanish antiSemitic classic by Raymundus Martinus, *Pugio Fidei*, Pascal took material from many Jewish sources to claim that "God used the blindness of the Jewish people for the benefit of the elect" (469 and 577) and that "if the Jews had been completely converted by Jesus Christ, we would not have had any but suspect witnesses. And if they had been exterminated, we would not have had any at all" (592 and 750). The apologetic argument, Pascal admitted, was not logically decisive but only persuasive. The real problem was to *be* a Christian, and here reason could not help. Humans could submit, but they still desperately required God's Grace.

> The prophecies, the miracles themselves, and the proofs of our religion are not of such a nature that it could be said they are absolutely convincing, but they are also of such a kind that it cannot be said that it would be unreasonable to believe them. Thus there is evidence and obscurity to enlighten some and confuse others, but the evidence is such that it surpasses or at least equals the evidence to the contrary, so that it is not reason that can determine men not to follow it, and thus this can only be as a result of lust or malice of heart…[so] that it appears that in those who follow it [religion], it is grace and not reason which makes them follow it, and that in

> those who shun it, it is lust and not reason that makes them shun it. (835 and 564)

Pascal's views hardly constitute an organized system. Most of his works are fragmentary, and he apparently made no effort to put the fragments together. His career first as a mathematical prodigy, then as a student of physics and finally as a religious thinker made continuous intellectual development difficult. From the vantage point of his fideistic religious views his mathematical and scientific efforts appeared to him as of small significance. Throughout the *Pensées* Pascal tried to characterize the role and limits of mathematical and scientific achievements, in keeping with what he himself had accomplished. But his religious views were essentially antiphilosophical. Among philosophical views he found skepticism the most congenial insofar as it revealed most clearly "the misery of man without God" and prepared men for faith and grace.

Pascal's religious concerns have overshadowed his other contributions and as a result his impact has been mainly on thinkers concerned with religious subjects. In recent years Pascal has been studied seriously by existentialists because of his brilliant portrayal of the human condition, and he has often been compared with Kierkegaard, especially in terms of his antiphilosophical and fideistic statement of Christianity. Pascal's works on scientific method and the philosophy of mathematics have tended to be neglected, but in these areas he was one of the clearest and most advanced thinkers of his age. His many-sided genius and his unequaled command of the French language make him one of the most inspiring and thought-provoking of writers. Pascal fills a major place in the history of ideas both for his work in mathematics, physics, and philosophy of science and for his insights into human nature and his analysis of Christianity.

See also Epistemology; Jansenism; Philosophy of Religion; Philosophy of Science, History of.

Bibliography

WORKS BY PASCAL

Le manuscrit des pensées de Pascal (1662), edited by Louis Lafuma. Paris: Libraries Associés, 1962. A photoreproduction of the manuscript text with a preface by Jean Guitton.

Oeuvres completes, edited by Jean Mesnard. Paris: Desclée de Brouwer, 1992.

Oeuvres completes, edited by Louis Lafuma. Paris: Editions du Seuil, 1963. Preface by Henri Gouhier.

Oeuvres de Blaise Pascal: publiées suivant l'ordre chronologique, avec documents complémentaires, introductions et notes, edited by Léon Brunschvicg, P. Boutroux, and F. Gazier. 14 vols. Paris: Hachette et cie, 1904–1914.

Pensées sur la religion et sur quelques autres sujets. 3rd ed. Paris: Delmas, 1960.

English Translations

Great Shorter Works of Pascal. Translated and with an introduction by Émile Cailliet and John C. Blankenagel. Philadelphia: Westminster Press, 1948.

Pensées: The Provincial Letters. Translated by W. F. Trotter and Thomas McCrie. New York: Modern Library, 1941.

Pensées. Translated by W. F. Trotter. New York: E. P. Dutton, 1958. Introduction by T. S. Eliot.

Pensées. Translated with an introduction by Martin Turnell. London: Harvill, 1962.

The Physical Treatises of Pascal: The Equilibrium of Liquids and the Weight of the Mass of Air. Translated by I. H. B. Spiers and A. G. H. Spiers. Introduction and notes by Frederick Barry. New York: Columbia University Press, 1937.

WORKS ABOUT PASCAL

Abercrombie, Nigel. *Saint Augustine and French Classical Thought*. Oxford: Clarendon Press, 1938.

Bishop, Morris. *Pascal: The Life of Genius*. New York: Reynal and Hitchcock, 1936.

Bishop, Morris, and Sister Marie Louis Hubert. "Pascal Bibliography." In *A Critical Bibliography of French Literature*. Vol. 3, edited by Nathan Edelman. Syracuse, NY: Syracuse University Press, 1961.

Bremond, Henri. *Histoire littéraire du sentiment religieux en France*. 12 vols. Paris: Bloud et Gay, 1916–1933. See vol. 4 (1921) for Pascal.

Brunschvicg, Léon. *Descartes et Pascal, lecteurs de Montaigne*. New York: Brentano's, 1944.

Brunschvicg, Léon. *Le génie de Pascal*. Paris: Hachette, 1924.

Busson, Henri. *La pensée religieuse française de Charron à Pascal*. Paris: J. Vrin, 1933.

Busson, Henri. *La religion des classiques (1660–1685)*. Paris: Presses Universitaires de France, 1948.

Cahiers de Royaumont. *Blaise Pascal, l'homme et l'oeuvre*. Paris: Editions de Minuit, 1956.

Cailliet, Émile. *Pascal: The Emergence of Genius*. 2nd ed. New York: Harper, 1961.

Chestov, Léon [Leo Isakovich Shestov]. *La nuit de Gethsémani; Essai sur la philosophie de Pascal*. Paris: Grasset, 1923.

Goldmann, Lucien. *Le Dieu caché*. Paris: Gallimard, 1955. Translated by Philip Thody as *The Hidden God: A Study of Tragic Vision in the Pensées of Pascal and the Tragedies of Racine*. New York: Humanities Press, 1964.

Humbert, Pierre. *Cet effrayant génie … : L'oeuvre scientifique de Blaise Pascal*. Paris: A. Michel, 1947.

Jovy, Ernest. *Études pascaliennes*. 9 vols. Paris: J. Vrin, 1927–1936.

Julien-Eymard d'Angers [Charles Cheshenau]. *Pascal et ses précurseurs*. Paris: Nouvelles Éditions latines, 1954.

Lafuma, Louis. *Histoire des Pensées de Pascal (1656–1952)*. Paris: Éditions du Luxembourg, 1954.

Laporte, Jean. *Le Coeur et la raison selon Pascal*. Paris: Elzévir, 1950.

Maia Neto, Jose. *The Christianization of Pyrrhonism: Scepticism and Faith in Pascal, Kierkegaard, and Shestov*. Dordrecht, Netherlands: Kluwer, 1995.

Mesnard, Jean. *Pascal, l'homme et l'oeuvre*. Paris: Boivin, 1951. Translated by G. S. Fraser as *Pascal: His Life and Works*. London: Harvill, 1952.

Strowski, Fortunat. *Pascal et son temps*. 3 vols. Paris: Plon-Mourrit et cie, 1907.

WORKS ON JANSENISM

Abercrombie, Nigel. *The Origins of Jansenism*. Oxford: Clarendon Press, 1936.

Orcibal, Jean. *Les origines du Jansénisme*. 5 vols. Louvain: Bureaux de la Revue, 1947–1962.

Sainte-Beuve, Charles A. *Port-Royal*. 7 vols. Paris: Eugéne Renduel, 1840–1859. 3-vol. ed., Paris: Gallimard, 1953–1955.

Richard Popkin (1967, 2005)

PASTORE, VALENTINO ANNIBALE

(1868–1956)

Valentino Annibale Pastore, an Italian philosopher and logician, was born at Orbassano (Teramo), Italy. He educated himself in literary studies, and then obtained a degree in letters from the University of Turin, under Arturo Graf, with a thesis on *La vita delle forme letterarie* (The life of literary forms), which was published at Turin in 1892. Pastore then turned to philosophy and was influenced by Hegelianism through the teachings of Pasquale d'Ercole. At the same time he was influenced by such scientists as Friedrich Kiesow, A. Garbasso, and Giuseppe Peano. In 1903 he published in Turin his thesis in philosophy, *Sopra le teorie della scienza: logica, matematica, fisica* (On the theories of science: logic, mathematics, physics). In 1911 he began teaching theoretical philosophy at Turin, where he was full professor from 1921 until 1939 and where he instituted a laboratory of experimental logic.

Pastore's thesis was published in the same year in which Benedetto Croce's *La critica* appeared and in which irrationalism burst out in Italy in diverse forms—as a revolt against positivism, as a rebirth of idealism, as an expression of the "bankruptcy of science." Having been educated in an environment in which Hegelianism was not ignored but was linked with the point of view of classical positivism, Pastore became aware of the impossibility of separating the sciences (mathematical and natural) from philosophy, or of substituting the sciences for philosophy. In the first case, if philosophy were severed from the conditions that render it possible and nourish it, it would become empty and would wither; in the second case, the sciences themselves would eventually lose consciousness of their relationships, their fundamental rationale, and their methods and goals. Pastore therefore sought to assess the meaning of scientific knowledge and of its logical procedures.

Turning his attention to logical problems in particular, Pastore was at first drawn toward Bertrand Russell's thesis of the identity of logic and mathematics, as is shown in *Logica formale e dedotta dalla considerazione dei modelli meccanici* (Formal logic deduced by the consideration of mechanical models; Turin, 1906) and *Sillogismo e proporzione* (Syllogism and proportion; Turin, 1910). His principal work of this period, *Il problema della causalità, con particolare riguardo alla teoria del metodo sperimentale* (The problem of causality, with particular attention to the theory of experimental method; 2 vols., Turin, 1921), which deals with causality, shows his systematic effort to single out the mutual relationship between scientific investigation and philosophical research. Pastore examined three aspects of causality—experience, science, and philosophy—and distinguished and analyzed the idea of cause, the concept of the causal relation, and the principle of causality.

After 1922, Pastore's interests were still focused on scientific knowledge, but he clarified his conception of philosophy as the study of "pure thought," as "not of that which is common to all particular systems, by being inherent in each one, but of that which results from all the particular systems, even though not being inherent in each one." From this conception he evolved his idea of a "general logic" whose basis lies "outside of particular logical systems." Around 1936, assisted by Ludovico Geymonat, he investigated the "logic of strengthening" as a "theory of primal systems," that is, as a search for "the process of construction of the most elementary forms of thinking and of their relationships," by means of a distinction between logic as logicality (general presystematic logic) and logic as a particular system, joining, as he himself said, "the deduction of the discourse (*D*) with the logical intuition of the universe (*U*)." Pastore did not seek to reach a demonstration of intuitive principles, nor to propose an ontological intuition, but rather to establish the laws of the relationship between *D* and *U*, between the analysis of the discourse and a synthetic vision of the universe.

In the final phase of his work Pastore's concern with the sense of mystery became marked ("logic has always two allies at its side: sadness and mystery"). In the light of this concern he examined and discussed both the existen-

tialist movements and the historical materialism of Karl Marx and V. I. Lenin.

See also Croce, Benedetto; Experience; Hegelianism; Irrationalism; Lenin, Vladimir Il'ich; Logic Machines; Marx, Karl; Peano, Giuseppe; Positivism; Russell, Bertrand Arthur William.

Bibliography

Additional works by Pastore are *Il solipsismo* (Turin, 1923); *La logica del potenziamento* (Naples, 1936); *Logica sperimentale* (Naples, 1939); *L'acrisia di Kant* (Padua: CEDAM, 1940); "Il mio pensiero filosofico," in *Filosofi italiani contemporanei,* edited by M. F. Sciacca (Como, 1944), pp. 333–349; *La filosofia di Lenin* (Milan: G. Bolla, 1946); and *La volontà dell'assurdo. Storia e crisi dell'esistenzialismo* (Milan: G. Bolla, 1948).

Works on Pastore are Carmelo Ottaviano, "La 'logica del potenziamento' della scuola di Torino," in *Logos* 17 (1934): 277–289; Francesco Crestano, "Intorno alla logica del potenziamento e alla logica dei comportamenti," in *Archivio di filosofia,* 5 (1935): 322–331 (and in *Idee e concetti,* Milan, 1939); and Filippo Selvaggi, *Dalla filosofia alla tecnica: La logica del potenziamento* (Rome: Apud aedes Universitatis Gregorianae, 1947), with bibliography.

Eugenio Garin (1967)
Translated by Robert M. Connolly

PATER, WALTER HORATIO
(1839–1894)

Walter Horatio Pater, an English essayist and critic, lived mainly in Oxford, where he read classics at Queens College and later became a fellow of Brasenose. He was a central figure of and inspiration for English fin de siècle art and art criticism and a profound influence on Oscar Wilde. He is of importance in philosophical aesthetics for his association with and championing of the *l'art pour l'art* doctrine of his age and for his insistence on "aesthetic criticism" of literature and the fine arts, stressing the subjective sensitivity of the critic and his power to paint evocative pictures of moments of intense experience in finely wrought, decorative prose. He is important in general philosophical history for his aphoristic but consistent statements that a relativist position was the only appropriate position for the modern temperament.

In the course of his career he proposed a highly personal conception of Platonism (*Plato and Platonism,* New York and London, 1893), playing down the immutable aspect of the theory of forms and emphasizing the imaginative sweep of Plato's more informal thinking. Pater maintained that moral values and moral standards were relative to the achievements and conditions of an age. Although he was formerly a Christian, he did not believe that Christian revelation had a privileged status, and he stressed the anthropological interpretation and psychological significance of all religious ritual. His tendency to ethical relativism, his inclination to praise goodness for its beauty, and his attitude toward religion as an aesthetically satisfying experience without final commitment made him many enemies in Oxford. The Paterian temperament was identified with aestheticism, or the hedonistic enjoyment of the intensely lived moment of beauty, the "exquisite passion," regardless of formal and moral standpoints. He was blamed for much of the moral eccentricity and artistic preciousness and pretentiousness of his followers, who deliberately courted decadence. However, he himself led a rather carefully balanced, withdrawn life, to which the famous sentence from the conclusion to *The Renaissance,* "To live always with this hard, gemlike flame, to maintain this ecstasy, is success in life," can be applied only with some difficulty.

In his *Imaginary Portraits* (London, 1887), Pater developed the genre of imaginative presentation of personalities embodying certain philosophies of life. His novel, *Marius the Epicurean* (London, 1885), regarded by many as his major work, is one such imaginary portrait on a large scale, picturing the religious development of a highly civilized, aesthetically sensitive agnostic at the time of Marcus Aurelius and probably indicating Pater's own attitude toward religion.

Pater's importance for English letters might be said to lie largely in his having cultivated the essay form to a high level of competence combined with elegance, making a fine art out of deliberate abstention from judgment, out of tentativeness and the impressionistic recording of subjective states of mind. His best criticism occurs in the collection *The Renaissance,* in his essay on Samuel Taylor Coleridge in *Appreciations* (London, 1889), and in the essay on style (appended to *Appreciations*).

Pater understood the "historical method" to be the attempt to understand artistic phenomena in relation to the conditions that produced them and to commend them to the sympathetic imagination of the reader. Unlike Matthew Arnold, who had contrasted personal and historical assessment with the "real" assessment of art, Pater did not believe in any fully objective standards but only in the completely honest account of personal impressions against the background of historical relativity. While ostensibly agreeing with Arnold that one must see the object "as it really is," he insisted that this can be done only on the basis of knowing one's own impressions

"as they really are." The critic needs a certain kind of temperament, the power of being deeply moved by the presence of beautiful objects. Pater acknowledged no distinction here between beautiful things in and apart from art. Yet he offered some fine insights into the autonomy and interdependence of the various arts, especially in the implications of his much-quoted passage from the essay "The School of Giorgione" in *The Renaissance*: "All art aspires constantly towards the condition of music." In the preceding paragraph of the essay, Pater wrote that each art has "its own specific order of impressions, and an untranslatable charm." Yet each art form, as art, needs the complete fusion of matter and form that music exemplifies in its purity.

See also Aesthetic Judgment; Aesthetic Qualities; Aesthetics, History of; Arnold, Matthew; Coleridge, Samuel Taylor; Marcus Aurelius Antoninus; Plato; Platonism and the Platonic Tradition; Wilde, Oscar Fingal O'Flahertie Wills.

Bibliography

PRINCIPAL WORKS BY PATER

Essays from the "Guardian." London, 1897. Published uniformly, may be regarded as Vol. IX of the works.

Works, 8 vols. London, 1900–1901.

WORKS ON PATER

Benson, A. C. *Walter Pater.* London: Macmillan, 1906.

Cecil, Lord David. *Walter Pater.* Cambridge, U.K.: Cambridge University Press, 1955.

Eliot, T. S. "The Place of Pater." In *The Eighteen-Eighties,* edited by Walter de la Mare. Cambridge, U.K.: Cambridge University Press, 1930.

Gaunt, William. *The Aesthetic Adventure.* London: J. Cape, 1945.

Greenslet, Ferris. *Walter Pater.* London: William Heinemann, 1904.

Hough, Graham. *The Last Romantics.* London: Duckworth, 1949.

Iser, Wolfgang. *Walter Pater.* Tübingen: Niemeyer, 1960.

Eva Schaper (1967)

PATERNALISM

The term *paternalism* has long been in currency among moral and political philosophers, but its circulation became much wider, and its definitions much more precise, following the widely read debate over "the legal enforcement of morality" between Patrick Devlin (*The Enforcement of Morals,* 1965) and H. L. A. Hart (*Law, Liberty, and Morality,* 1963). Hart had endorsed the liberal doctrine of J. S. Mill, that the only legitimate reason for state interference with the liberty of one person is to prevent him from harming other persons. Mill was especially emphatic in denying that the actor's "own good, either physical or moral," is ever an adequate reason for interference or criminal prohibition ([1859], 1985, p. 9). What Mill denied in this passage is precisely what came to be called "legal paternalism" in the writings of his followers, including Hart nearly a century later. Thus, paternalism was regarded as a thoroughly unacceptable view by nineteenth-century liberals.

PHYSICAL AND MORAL

In his exchange with Devlin, however, Hart conceded that a certain amount of physical paternalism could be accepted by twentieth-century liberals, here departing from Mill who, he wrote, "carried his protests against paternalism to lengths that may now appear to us as fantastic" (Hart 1963, p. 32). He cited, for example, Mill's criticism of restrictions on the sale of drugs. Devlin then responded by drawing a distinction between "physical paternalism," which protects people from physical harm that could be caused by their own voluntary conduct, and "moral paternalism," which offers similar protection against "moral harm" of the actor's own causing. Devlin could see no consistent way in which the physical paternalist like Hart could avoid commitment to moral paternalism, for if it is the prevention of harm that justifies prohibition in the one case, why not use state power to prevent an equal amount of harm, though of a different kind, in the other case? Similarly, Devlin concluded, there is no relevant difference between criminalization meant to prevent moral harm and criminal prohibitions meant to "enforce the moral law as such." The view that "enforcement of morality," quite apart from harm prevention, is a valid reason for criminal prohibitions is widely called "legal moralism." It is anathema to liberals.

One way in which liberals sometimes defend themselves from Devlin's argument is by maintaining that Devlin's moves from physical to moral harm and from preventing moral harm to "enforcing the moral law" do not follow logically. One liberal critic, Joel Feinberg (1986), even goes so far as to deny, in the teeth of the immense combined authority of Plato and Aristotle to the contrary, that "moral harm" is a coherent concept.

HARD AND SOFT

A distinction is commonly made between hard (or strong) paternalism and soft (or weak) paternalism. Hard paternalism justifies the forcible prevention of some dan-

gerous but self-regarding activities even when those activities are done in a fully voluntary (i.e., free and informed) way. Soft paternalism, on the other hand, permits individuals or the state to prevent self-regarding dangerous behavior only when it is substantially nonvoluntary or when temporary intervention is necessary to establish whether it is voluntary or not.

Most soft paternalists are liberals strongly opposed to paternalism. Most of them, when they think of the paternalism they oppose, think of what is here called hard paternalism. Therefore they would prefer to go by the name of soft antipaternalists. The term *hard antipaternalism* could be reserved for the totally uncompromising liberal who would oppose interference even with some choices known to be involuntary, and with temporary compulsory intervention that is only for the purpose of determining whether the intended conduct truly is voluntary, and even with the imposition of compulsory education about risks or state-administered tests to assess the dangerous actor's understanding of the risks, with licenses required for self-regarding dangerous behavior, like mountain climbing. Clarity would be improved if philosophers would speak of paternalism only when what is meant is hard paternalism, justifying prohibition even of wholly voluntary self-regarding conduct, when dangerous. Then soft and hard antipaternalism would be the names of a moderate and extreme liberalism, respectively.

The controversy over paternalism in the criminal law is genuine and difficult. Those who are strongly opposed to paternalism find it not only mistaken but arrogant and demeaning. It is very difficult to reconcile it with even a minimal conception of personal autonomy (rightful self-government) when it proclaims that state officials may rightfully intervene even against my protests to "correct" my choices, and this on the ground that they know what is good for me better than I do myself. But if we reject paternalism altogether, we seem to fly in the face both of common sense and of long-established customs and laws. The state, for example, does not accept "consent" as a justification for mayhem or homicide. Similarly, the law of contracts will not validate certain agreements even though they are voluntary on both sides—when, for example, they are usurious or bigamous. One would be hard put to accept these traditional state-created disabilities without abandoning one's opposition to paternalism. But if we continue our adherence to paternalism, we may discover that in other areas paternalism justifies too much, the flat-out prohibition, for example, of whiskey, cigarettes, and fried foods, which tend to be bad for people too, whether they know it or not.

MEDICAL CONTEXTS

Writers on medical ethics confront paternalism at every turn, often in human contexts that are less familiar to those whose interest is primarily focused on criminal law. Those characteristic social situations have led to some forms of ethical analysis supplementary to those that prevail among the critics and defenders of "legal paternalism." For example, not all of the moral problems raised by paternalism in medical settings are problems for legislators drafting mandatory rules or other governmental officials such as judges or police officers. Moreover, paternalism is not exclusively a criterion for the legitimacy of coercion. Sometimes what is at issue is some other practice that normally has high moral costs, most notably deception rather than coercion, as in false but comforting statements to frightened patients or the unacknowledged or mendacious use of placebos. Sometimes a medical provider may have to decide whether to tell a "white lie" to his patient, not for the sake of her health, but rather as a way of preventing her from experiencing intense despair in her final hours about a matter having no direct connection with medical treatment. In a hypothetical case invented by C. M. Culver and B. Gert (1976, p. 46), a woman on her deathbed asks her physician how her son is doing, and the doctor replies that he is doing well even though he knows that "the son has just been killed trying to escape from jail after having been indicted [a fact unknown to his mother] for multiple rape and murder." An opponent of (hard) paternalism would probably consider the doctor's mendacity to be a violation of the patient's autonomy. A medical paternalist would probably argue that the truthful alternative in this case would be cruel to the point of indecency. They might both be right.

PROS AND CONS

Problems involving paternalism in medical contexts are quite diverse. They include not only truth-telling cases but also suicide attempts, requests for euthanasia, and the use of human volunteers in dangerous experiments. The paternalist position in these conflicts is that protecting volunteers or patients from harm and promoting their benefit should take precedence over respecting their autonomy by permitting them to act freely on their well-informed choices in matters that are almost exclusively self-regarding.

T. L. Beauchamp (1977) and Beauchamp and J. F. Childress (1979) in their influential works rejected hard paternalism nearly categorically, emphasizing that to overturn the deliberate choices of adult human beings

that affect only them, or only them clearly and directly, is to deny that their lives really belong to them. The apparent exceptions—cases in which commonsense morality would seem to justify interference with the patient's voluntary choice—invariably turn out to be cases in which that choice is not fully voluntary after all; that is, the patient or volunteer subject had not been adequately informed about the risks he would be accepting, or he was not perfectly free of coercive influences, or some other condition, such as infancy, drug intoxication, high fever, rage, or depression, had diminished his capacity to act rationally. To restrict his liberties in such circumstances, or to motivate him by telling him a lie, would be to interfere with actions that are not fully voluntary in the first place. To interfere with dangerous self-regarding but less-than-voluntary behavior can be justified by soft paternalism (that is by soft *and* hard antipaternalism). Another example illustrates the point. "If we see a normally calm person who we know has been experimenting with hard drugs, go into a sudden frenzy, and seize a butcher knife with the clear intention of cutting his own throat, then [of course!] we have the right to interfere. In so doing we will not be interfering with his real self or blocking his real will.… His drug-deluded self is not his 'real self,' and his frenzied desire is not his 'real choice,' so we may defend him against these internal threats to his autonomous self, which is quite another thing than throttling that autonomous self with external coercion" (Feinberg 1986, p. 14). Interference on this ground is no more paternalistic than interference designed to protect an individual from an attack by some berserk assailant. Paternalists have been quick to point out, however, that this example, and others like it, hardly fit the more usual examples of risky choice making.

Writing from the practical point of view, and a philosophical position more friendly to paternalism, Culver and Gert (1982), in response to Beauchamp, point out that many crucial questions remain for the soft antipaternalist analysis. Most of these stem from the vagueness of the distinction between voluntary and nonvoluntary. Culver and Gert remind us that voluntariness is usually a matter of degree with no conveniently placed bright lines to guide us. In this respect it resembles the concept of harm (which is also crucially involved in hard paternalists' calculations) and the degree of violation of a moral rule, like that forbidding telling lies, or that condemning coercion, and even the degree to which the overruled choices of, say, a patient, are purely self-regarding—another essential variable.

Culver and Gert, however, do not endorse the hard paternalistic position without limit. Rather, they hold that some (hard) paternalistic interventions are justified, and some are not, but reject the unqualified antipaternalism of Beauchamp and Childress, which denies that (hard) paternalistic prohibitions and interferences are ever justified, and the unqualified paternalism of many utilitarian writers, which holds that *all* paternalistic behavior is justified, except that which will be counterproductive in the long run.

See also Aristotle; Bioethics; Hart, Herbert Lionel Adolphus; Liberty; Mill, John Stuart; Plato.

Bibliography

Arneson, Richard. "Paternalism, Utility and Fairness." *Revue Internationale de Philosophie* 170: 409–423.

Beauchamp, T. L. "Paternalism and Bio-behavioral Control." *Monist* 60 (1977): 62–80.

Beauchamp, T. L., and J. F. Childress. *Principles of Biomedical Ethics.* New York: Oxford University Press, 1979.

Brennan, Samantha. "Paternalism and Rights." *Canadian Journal of Philosophy* 24 (1994).

Brock, D. "Paternalism and Promoting the Good." In *Paternalism,* edited by R. Sartorius. Minneapolis: University of Minnesota Press, 1983. This is one of the leading statements of a qualified utilitarian theory of paternalism.

Culver, C. M., and B. Gert. "Paternalistic Behavior." *Philosophy and Public Affairs* 6 (1976): 45–47.

Culver, C. M., and B. Gert. *Philosophy in Medicine.* New York: Oxford University Press, 1982.

Devlin, P. *The Enforcement of Morals.* London: Oxford University Press, 1965.

Dworkin, G. "Paternalism." *Monist* 56 (1972): 64–84. An influential early article that helped shape twenty years of discussion.

Dworkin, G. "Paternalism: Some Second Thoughts." In *Paternalism,* edited by R. Sartorius, 105–113. Minneapolis: University of Minnesota Press, 1983.

Dworkin, Gerald. "Moral Paternalism." *Law and Philosophy.* 2005.

Dworkin, Gerald. "Paternalism." *The Stanford Encyclopedia of Philosophy* (Winter 2002), edited by Edward N. Zalta.

Dworkin, Ronald. "Equality and The Good Life." In *Sovereign Virtue,* 268–270. Cambridge, MA: Harvard University Press, 2000.

Faden, R. R., and T. L. Beauchamp. *A History and Theory of Informed Consent.* New York: Oxford University Press, 1986. The definitive work on its subject.

Feinberg, J. *The Moral Limit of the Criminal Law.* Vol. 3: *Harm to Self.* New York: Oxford University Press, 1986.

Freeman, Samuel. "Liberalism, Inalienability, and Rights of Drug Use." In *Drugs and the Limits of Liberalism,* edited by Pablo DeGrieff. Ithaca, NY: Cornell University Press, 1999.

Goldman, Alan. "The Refutation of Medical Paternalism." *The Moral Foundations of Professional Ethics.* Towata: Rowman & Littlefield.

Hart, H. L. A. *Law, Liberty, and Morality.* London: Oxford University Press, 1963.

Husak, Douglas N. "Legal Paternalism." In *The Oxford Handbook of Practical Ethics,* edited by Hugh LaFollette. New York: Oxford University Press, 2003.

Kleinig, J. *Paternalism.* Totowa, NJ: Rowman and Allanheld, 1983.

Mill, J. S. *On Liberty* (1859). Indianapolis, 1985.

Sartorius, R., ed. *Paternalism.* Minneapolis: University of Minnesota Press, 1983. This excellent collection includes, in addition to the selections by Brock and Dworkin already cited, fifteen useful articles and a superb bibliography.

Shiffrin, Seana Valentine. "Paternalism, Unconscionability Doctrine, and Accommodation." *Philosophy and Public Affairs* 29 (Summer 2000): 205–250.

Sunstein, Cass, and Richard Thaler. "Libertarian Paternalism Is Not an Oxymoron." *University of Chicago Law Review* 70 (Fall 2003): 1166.

VanDe Veer, D. *Paternalistic Intervention: The Moral Bounds of Intervention.* Princeton, NJ: Princeton University Press, 1986.

Joel Feinberg (1996)
Bibliography updated by Gerald Dworkin (2005)

PATRIOTISM

The various current meanings of the term *patriotism* emerged during and after the early modern period. In the tradition of republicanism, patriotism is the citizens' commitment to or love for their shared political freedom and the institutions that sustain it. This commitment manifests itself in civic activity on behalf of the political commonwealth and its members. In this tradition, *patriotism* is often synonymous with *public spiritedness*. In the nineteenth century, patriotism was increasingly interpreted in a different, nationalist manner, and patriotism and nationalism are nowadays often equated.

An effective tool for mobilizing popular support for national policy, including military aggression and other forms of national aggrandizement, patriotism is often regarded as implying the glorification of war and imperialism. Opponents of such policies have also acted in the name of patriotism, however. Therefore, two preliminary questions require answers before any further discussion of the ethical implications of patriotism can proceed. First, the *patria*, the object of patriotic loyalty or activity, needs to be specified. According to some, this is the constellation of political institutions one finds oneself in. According to others, it is one's cultural or linguistic community (nation), one's country, the physical environment in which one was born or with which one identifies, or a combination of these. Accordingly, one can distinguish between different kinds of patriotism, such as "constitutional patriotism" and "nationalist patriotism." Second, while all agree that patriotism is a certain attitude, there is disagreement as to its precise nature. If it is an attitude of loyalty, what does loyalty require? Does patriotism require a certain sentiment, such as love or enthusiasm? Or should it primarily be understood as a social practice, and if so, what type of practice?

The belief in the superiority of one's own *patria* and a concomitant disdain for others is not a necessary element of the concept of patriotism. Nor does patriotism require that one refrain from criticizing one's *patria*. Indeed, criticism of governmental policy is often presented as patriotic since it aims at improving the *patria*.

The contemporary philosophical discussion of patriotism focuses on its relation to cosmopolitanism, as one aspect of the more general debate about particularism and universalism. Many authors in the republican tradition have argued emphatically that patriotism and cosmopolitanism are compatible, even that patriotism is a step toward cosmopolitanism, as it widens the individual's scope of concern beyond that of the family and so prepares one for the wider community of humankind. Others, however, especially defenders of patriotism in the nationalist tradition (as well as many defenders of cosmopolitanism), have seen an irreconcilable tension between (nationalist) patriotism and cosmopolitanism, on the grounds that cosmopolitanism would (rightly or wrongly, depending on which side one is on) prohibit favoring one's own national group over the rest of humankind.

Any defense of patriotism should address the question where justified special care or commitment ends and unjust parochialism begins. Thus, the philosophical debate over patriotism takes place in the context of the debate over "special obligations." Here the question is whether patriotism is prohibited (e.g., as necessarily jingoistic, as violating a moral standard of impartiality), permissible (and if so, under what conditions), or a duty (e.g., as a necessary condition for a well-functioning polity, or as a special obligation toward one's fellow citizens). Clearly, the specific answer one gives to this question depends on both one's particular conception of patriotism and one's underlying moral theory.

See also Loyalty; Nationalism.

Bibliography

Kleingeld, Pauline. "Kantian Patriotism." *Philosophy and Public Affairs* 29 (2000): 313–341.

Nussbaum, Martha C., et al. *For Love of Country: Debating the Limits of Patriotism*, edited by Joshua Cohen. Boston: Beacon Press, 1996.

Primoratz, Igor, ed. *Patriotism*. Amherst, NY: Humanity Books, 2002.

Viroli, Maurizio. *For Love of Country: An Essay on Patriotism and Nationalism*. Oxford: Clarendon Press, 1995.

Pauline Kleingeld (2005)

PATRISTIC PHILOSOPHY

"Patristic philosophy" is the term used to refer to the philosophical presuppositions, motifs, and structures in the writings of the early Christian apologists and Church Fathers. These writers were essentially theologians rather than philosophers, for their starting point lay in God and his self-revelation. Their use of philosophy can be divided into three periods: (1) the beginnings (roughly the first and second centuries CE), in which ideas derived from Platonism, Stoicism, and (to a lesser extent) Skepticism were employed chiefly for apologetic purposes, largely under the influence of Hellenistic Judaism; (2) the early Alexandrian period, during which Middle Platonism and Stoicism were dominant, especially in the thought of Clement and Origen; and (3) the development of Christian Neoplatonism, first under the influence of Porphyry and later under that of Proclus. The influence of Philo of Alexandria may have been felt during the first period and certainly was an important factor in the second.

BEGINNINGS

THE NEW TESTAMENT. In the New Testament, as in the Apocrypha (for example, in the Wisdom of Solomon), there are ideas that are at least latently philosophical. As early as Paul's first letter to the Corinthians (8:6), the Christian faith was being formulated with the use of prepositions that in Greek philosophy indicated causal relations. For Christians there was "one God the Father, from whom is everything and for whom are we, and one Lord Jesus Christ, through whom is everything and through whom are we." The Father was thus represented as the first and final causes (see Romans 11:36, a doxology), the Lord as the instrumental cause. Such an analysis was presumably derived from Hellenistic Judaism; Philo spoke thus concerning God and the Logos. In Romans 1:19–21 Paul discussed the primal knowledge of God's eternal power and deity, which he revealed by means of what he created. Men capable of receiving revelation knew God but turned away to worship the creation instead of the Creator (Romans 1:25; cf. Philo, *De Opificio Mundi*, Bk. 7). The theme of a revelation implicit in the structure of the created world is further developed in sermons ascribed to Paul in Acts 14:15–17 and 17:22–31 (the setting of the latter sermon contains reminiscences of the charges brought against Socrates and other philosophers at Athens), and in Colossians 1:15–20 the causal functions of Christ are further elaborated. The idea of the Logos, or creative Word of God, in John 1:1–14 is not necessarily philosophical either in its origin, which is probably not Philo, or in its expression. Later Christian theologians, however, interpreted it as philosophical, thus creating a bridge between Christianity and philosophy. These later theologians may perhaps have relied on Philo.

SECOND-CENTURY CHRISTIANITY. In the apocryphal *Preaching of Peter*, God is described by means of adjectives clearly philosophical in origin. God is uncontained, without needs, incomprehensible, eternal, imperishable, and invisible. These negative adjectives reflect ideas current not only in the Platonism of the time but also in Hellenistic Judaism. They are close to later Gnostic developments, and it has been suggested that both are derived from a rather fully developed doctrine of God current in early second-century Christianity. This view is confirmed by what Ignatius of Antioch (early second century) says of Christ as God and man: "the timeless, the invisible who for us was visible, the intangible, the impassible who for us was passible" (*Polycarpi* 3.2). Ignatius is obviously employing current language about God to describe Christ. About 140 the doctrine was more fully expressed in the *Apology* of Aristides (Ch. 1). God is the unmoved mover and ruler of the universe, for "everything that sets in motion is more powerful than what is moved, and what rules is more powerful than what is ruled." God is eternal, without beginning (what begins also ends) or end (what ends is destructible); he is therefore ungenerated, uncreated, immutable, and immortal. He has no defects or needs; he is not contained or measurable but contains all; he is immobile (he could not move from one place to another); and he is positively Wisdom and wholly Mind. According to Philo and others, God has no name, form, or parts.

A problem arose when such negative attributes were combined with traditional Jewish and Christian ideas about God as the Creator active in history. Basilides, a Christian Gnostic, tried to avoid any kind of analogical statement by arguing that the doctrine of emanation would make God spiderlike, whereas the doctrine of creation would make him anthropomorphic. Basilides claimed instead that originally there was absolutely noth-

ing, and then the nonexistent God made, so to speak, a nonexistent universe out of the nonexistent. Like certain Middle Platonists, Basilides held that God was completely transcendent, since "the universe cannot speak of him or contain him in thought"; he cannot even be called ineffable.

Christian thinkers, however, were generally less audaciously speculative. The apologist Justin Martyr (c. 160) wrote an account of conversion from Platonic religious philosophy to Christian truth. Justin had experienced the teaching offered by Stoics, Peripatetics, Pythagoreans, and Platonists but had little insight into any but the last. While a novice in Platonism he encountered a Christian who—apparently with Peripatetic arguments—demolished his defenses of the innate immortality of the soul and its reminiscence of the eternal world. After his conversion Justin continued to quote from Plato's dialogues (which in his view were partly based on the Old Testament), although his position was now fully eclectic: "Whatever has been said well by anyone belongs to us" (*Apologies*, Bk. 2, Ch. 13). He criticized the Stoic doctrines about fate and the *ekpyrosis* (destruction of the cosmos by fire) but expressed his admiration not only for Heraclitus and Socrates but also for the first-century Stoic moralist Musonius Rufus. Justin's disciple Tatian was much less friendly to philosophers, although he tried to create a theology largely Platonic in inspiration. His incidental reference to "the God who suffered" suggests that at a crucial point he had to rely on paradox.

The writings of the later apologists show that philosophy continued to influence theology. In the *Legatio* of Athenagoras (c. 178), there is an important attempt to demonstrate the oneness of God and consequently an approach toward a doctrine of the Trinity. In another treatise the logical necessity of corporeal resurrection is upheld on grounds that are largely Peripatetic. About the same time, Theophilus of Antioch set forth the doctrine that God is known only through his activities, to which his attributes and appellations refer; God is without beginning because uncreated, immutable because immortal. The word *theos* is derived from verbs referring to his creative acts. His invisibility is explained by analogies to the soul, a pilot, the sun, and a king. God is not "contained" but is the locus of the universe. He is known only through his Logos, originally existing within him as reason (*endiathetos*), then expressed as word (*prophorikos*) at creation.

Philosophical ideas influenced not only the apologists but other Christians as well. Irenaeus of Lyons (c. 185) was no philosopher, but in five passages he accepted a description of God originally derived from Xenophanes, "seeing entirely, knowing entirely, hearing entirely" (Fr. 24 in *Fragmente der Vorsokratiker*, edited by H. Diels and W. Kranz, 10th ed., Berlin, 1961) and amplified it, ascribing it both to "religious men" and to "the Scriptures." In three instances he added the Platonic phrase "the source of all good things."

During the crucial second century, then, Christian theologians generally shared their doctrine of God with Platonists. Their doctrine of the Logos resembled that of the Stoics, although Christian theologians believed in one Logos (as in Philo) rather than many. They used Skeptical arguments against the pagan gods. Their ethical teaching was often close to that of the Roman Stoa as represented by Musonius (and Epictetus). Like non-Christians of various schools, they tended to believe that there had once been a unified religious philosophy, Oriental in origin, from which later philosophers had deviated. This first philosophy, it was thought, had been based on the inspiration of the divine Logos or on borrowing from Moses, or on both. The views of the Christian theologians were thus close to the kind of Hellenistic Judaism represented by Philo. Few writers took up the philosophical problems presented by the Incarnation; several of them do not even mention Jesus.

THE CHRISTIAN PLATONISTS OF ALEXANDRIA

In the cultural center of Alexandria, Christian philosophical theology came into its own, first in the writings of Clement of Alexandria (late second century) and later in the fuller treatment of Origen. The rather disdainful attitude of both writers toward "simpler believers" illustrates the tension between traditional and philosophical theology in their time. Philosophy was often viewed elsewhere as a seedbed of heresy; such was the case at Rome with Hippolytus and at Carthage with Tertullian, even though both these writers used philosophical definitions and arguments. Clement and Origen made use of the writings of Philo and other Hellenistic Jews, although both were directly acquainted with most of the works of Plato, some Middle Platonic writings, a few Aristotelian treatises, and a great deal of Stoic literature. Clement's learning was both broader and more superficial than Origen's. His philosophical ideas apparently developed away from the boldness of his semi-Gnostic *Hypotyposes* (now lost) toward the greater caution reflected in the *Stromata*, in which philosophy became the handmaid of a theology traditional in essence if not always in expression.

The principal points at which the influence of philosophy is obvious are the doctrine of transcendence of God and the ideal world, analysis of the divine nature of Christ, divine impassibility as a model for human conduct, and Platonic and Stoic ethical conceptions. Following Philo, Clement made use of the allegorical method in order to relate his theology to the Bible. He was the head of a private philosophical school, training pupils to become Christian Gnostics. In later times he was far less influential than Origen, head of an authorized church school first at Alexandria and later at Caesarea. The ideas of both teachers, however, continued to create theological ferment as late as the sixth century.

LATER PATRISTIC PHILOSOPHY

We can hardly view Eusebius of Caesarea as a philosopher, but in the writings of the Cappadocian Fathers (especially Gregory of Nazianzus and Gregory of Nyssa) technical philosophical arguments are frequently adapted for theological use, as they are throughout the patristic period. During the fourth century the attack upon Christianity by Porphyry was largely forgotten (a new attack was produced by the emperor Julian), and the logical rigor of his eclectic Neoplatonism was viewed as supporting theology. Extensive quotations from Porphyry and his master Plotinus appear in Eusebius's writings as well as in the later treatise *Against Julian* by Cyril of Alexandria. Toward the end of the fourth century, a faintly Christianized Neoplatonism appeared in the West in the commentary on Plato's *Timaeus* by a certain Calcidius, who relied primarily on Porphyry. Before being baptized, Marius Victorinus had translated one of Porphyry's works into Latin; he made frequent use of Porphyry's teaching in his later treatises *On the Trinity*. Both Ambrose and Augustine were deeply influenced by Porphyry, whose writings paved the way for Augustine's conversion. In the late fifth century the ideal world of the Neoplatonist Proclus was Christianized in the influential writings ascribed to Dionysius the Areopagite.

See also Apologists; Origen.

Bibliography

For critical textual editions, the following two series:

Corpus Scriptorum Ecclesiasticorum Latinorum. Vienna: Kommission zur Herausgabe des Corpus der lateinischen Kirchenväter, 1864–present.

Corpus Christianorum Series Latina. Turnhout, Belgium: Brepols Publishers, 1953–present.

For two series with translations of many complete works:

The Fathers of the Church. Washington, DC: Catholic University of America Press, 1947–present.

Roberts, Alexander, and James Donaldson, eds. *The AnteNicene Fathers: Translations of the Writings of the Fathers down to A.D. 325*. Buffalo: Christian Literature Publishing. Many of these translations have been superseded, but there is some material available in English only in these volumes.

For a one-volume sampler of texts in translation that will serve the interests of many readers:

Bettenson, Henry, ed. and tr. *The Early Christian Fathers*. London: Oxford University Press, 1956.

For general studies, see the following three works.

Wolfson, Harry Austryn. *Studies in the History of Philosophy and Religion*. Vol. I. Cambridge, MA: Harvard University Press, 1973.

Armstrong, A. H., ed. *The Cambridge History of Later Greek and Early Medieval Philosophy*. Cambridge, U.K.: Cambridge University Press, 1970.

Gersh, Stephen. *Middle Platonism and Neoplatonism: The Latin Tradition*. 2 vols. Notre Dame, IN: University of Notre Dame Press, 1986. Vol. II is more relevant to the subject.

For recent studies on individual figures or themes mentioned in the original article:

Barnard, Leslie W. *Athenagoras: A Study in Second Century Christian Apologetic*. Paris: Beauchesne, 1972.

Barnard, Leslie W. *Justin Martyr: His Life and Thought*. London: Cambridge University Press, 1967.

Cerrato, J. A. *Hippolytus between East and West: The Commentaries and the Provenance of the Corpus*. Oxford: Oxford University Press, 2002.

Cyril of Alexandria. *Letters*. 2 vols. Translated by John I. McEnerney. Washington, DC: Catholic University of America Press, 1987.

Cyril of Alexandria. *Select Letters*, edited and translated by Lionel R. Wickham. Oxford: Clarendon Press, 1983.

Goodenough, Erwin R. *The Theology of Justin Martyr: An Investigation into the Conceptions of Early Christian Literature and Its Hellenistic and Judaistic Influences*. Amsterdam: Philo Press, 1968.

Grant, Robert M. *Irenaeus of Lyons*. London: Routledge, 1997.

Hadot, Pierre. *Marius Victorinus: recherches sur sa vie et ses oeuvres*. Paris: Études augustiniennes, 1971.

Joly, Robert. *Christianisme et philosophie: études sur Justin et les apologistes grecs du deuxième siècle*. Bruxelles: Éditions de l'Université de Bruxelles, 1973.

Minns, Denis. *Irenaeus*. Washington, DC: Georgetown University Press, 1994.

Moorhead, John. *Ambrose: Church and Society in the Late Roman World*. London: Longman, 1999.

Morino, Claudio. *Church and State in the Teaching of St. Ambrose*. Translated by M. Joseph Costelloe. Washington, DC: Catholic University of America Press, 1969.

Osborn, Eric. *Irenaeus of Lyons*. New York: Cambridge University Press, 2001.

Osborne, Catherine. *Rethinking Early Greek Philosophy: Hippolytus of Rome and the Presocratics*. Ithaca, N.Y.: Cornell University Press, 1987.

Pelikan, Jaroslav. *Christianity and Classical Culture: The Metamorphosis of Natural Theology in the Christian*

Encounter with Hellenism. New Haven, CT: Yale University Press, 1993.

Ramsey, Boniface. *Ambrose*. London: Routledge, 1997.

Vallée, Gérard. *A Study in AntiGnostic Polemics: Irenaeus, Hippolytus, and Epiphanius*. Waterloo, Ont., Canada: Wilfrid University Press, 1981.

Wilken, Robert L. *Judaism and the Early Christian Mind: A Study of Cyril of Alexandria's Exegesis and Theology*. New Haven, CT: Yale University Press, 1971.

Robert M. Grant (1967)
Bibliography updated by William E. Mann (2005)

PATRIZI, FRANCESCO
(1529–1597)

Francesco Patrizi, also known as Patritius, was a vigorous defender of Platonism and an unremitting foe of Aristotelianism. He was versatile even for his time, being at once philosopher, mathematician, historian, soldier, and literary critic. Born in Dalmatia, he studied at Padua (Francesco Robertelli was a teacher-friend) and Venice. Having been an early and avid reader of Marsilio Ficino's *Theologia Platonica,* he turned from careers in business and in medicine to develop further his interest in Platonism.

After some years in France, Spain, and Cyprus in the service of various noblemen, Patrizi was in 1578 appointed by Duke Alfonso II as professor of Platonic philosophy at the University of Ferrara—which, with Florence and Pisa, was an important center of Platonism in Italy. In 1592 he was called to the University of Rome by Pope Clement VIII. He considered the privilege of expounding Platonism at Rome his crowning achievement, and he held that position until his death.

Although intellectual activity was his chief concern, Patrizi also showed interest in practical matters: He offered means for diverting a river threatening Ferrara, and presented plans for improving military strategy against the Turks and naval plans against the British.

In 1553 Patrizi's *Discorso* on types of poetic inspiration appeared, followed by his dialogues on history (1560). After visiting France, Spain, and Cyprus, he published *Discussiones Peripateticae* (1581), which violently attacked Aristotelianism. His achievement dates largely from his appointment at Ferrara, although correspondence with Telesio (1572) indicates an earlier interest in the study of nature. In *Della Poetica* (1586), he produced the first modern study of literary history, which also was an attack on Aristotle's *Poetics.* In 1587 there appeared several polemics defending his friend Orazio Ariosto against Torquato Tasso and Jacopo Mazzoni and upholding Patrizi's Platonic view of art as transcendental against their Aristotelian theory of poetry as imitation.

Patrizi's chief philosophical work, *Nova de Universis Philosophia* (1591), contained four parts: *Panaugia,* on light; *Panarchia,* on first principles; *Pampsychia,* on souls; and *Pancosmia,* on mathematics and natural science. Dedicated to Gregory XIV, who had been a fellow student at Padua, its aims were the linking of Christianity with the teachings of Zoroaster, Hermes, and Orpheus; the derivation of the world from God through emanation; and the insistence on a quantitative study of nature. His last work was *Paralleli Militari* (1594).

Patrizi's metaphysics of light is suggestive of Ibn Gabirol and Robert Grosseteste, and places him in the company of Geronimo Cardano and Bernardino Telesio. Defending the cognitive value of mathematics (as did Nicholas of Cusa), Patrizi helped to establish the subsequent priority of space over matter in the study of nature. His doctrines, fanciful yet impressive, failed (as did those of Giordano Bruno and Telesio), for want of an adequate method, to overthrow the well-entrenched Aristotelians. The decisive attack came only in the seventeenth century, when Galileo Galilei and others postulated a new physics of quantities that was related to astronomy and was based on experiments and calculations.

See also Aristotelianism; Bruno, Giordano; Ficino, Marsilio; Galileo Galilei; Grosseteste, Robert; Ibn Gabirol, Solomon ben Judah; Matter; Nicholas of Cusa; Platonism and the Platonic Tradition; Space; Telesio, Bernardino.

Bibliography

WORKS BY PATRIZI

Discorso. 1553.

Della Historia. Venice, 1560.

Discussiones Peripateticae. Basel, 1581.

Della Poetica. Ferrara: V. Baldini, Stampator Ducale, 1586.

Nova de Universis Philosophia. Ferrara, 1591.

Paralleli Militari. Rome: L. Zannetti, 1594.

WORKS ON PATRIZI

Brickman, B. *An Introduction to Francesco Patrizi's* Nova de Universis Philosophia. New York, 1941.

Kristeller, P. O. *Renaissance Thought: The Classic, Scholastic, and Humanistic Strains.* New York, 1961.

Robb, N. *Neoplatonism of the Italian Renaissance.* London: Allen and Unwin, 1935.

Salata, F. "Nel terzo centenario della morte di F. Patrizi." *Atti e memorie della Società Istriana di Archeologia e Storia Patria* 12 (1897): 455–484.

Jason L. Saunders (1967)

PAULER, AKOS

(1876–1933)

The Hungarian philosopher Akos Pauler, son of an archivist and historian and grandson of a professor of law, grew up in an intellectual and bookish environment. Even before he matriculated, he published his first article in the scholarly journal *Bölcseleti Folyoirat* in 1893. It was a defense of metaphysics against positivism—metaphysics starts from what is given and goes back to that without which the given cannot be thought. This is, in germ, Pauler's "reductive method" (as against induction and deduction), which became his main preoccupation in later life. However, influenced by his university professor Imre Pauer, he was first a positivist for about a decade. After obtaining his doctorate at Budapest in 1898, he spent a year at Leipzig and another at the Sorbonne. In 1902 Pauler became *Privatdozent* at Budapest and, in 1906, lecturer in ethics on the faculty of law at Pozsony (Bratislava). His departure from positivism seems to have started during this period, since his work on ethics published in 1907 at Budapest, *Az Etikai Megismerés,* is close to the Kantianism of Heinrich Rickert. In 1912 Pauler became professor of philosophy at Kolozsvár, and from 1915 he occupied the chair of philosophy at Budapest.

Most expositions of Pauler mention his division of philosophy into five parts—logic, ethics, metaphysics, aesthetics, and ideology—presented in the first seven paragraphs of his *Bevezetés a Filozofiaba* (Introduction to philosophy; Budapest, 1920; revised 3rd ed., Budapest, 1933). However, it will be sufficient to discuss only his logic and metaphysics.

For Pauler, logic is the most important part of philosophy, which is not surprising in view of his broad notions of logic, the scope and nature of which can be seen from his four "laws of logic"—the law of identity: "Everything is identical only with itself," from which follow the laws of contradiction and excluded middle; the law of connection: "Everything is connected with other things," which includes the law of sufficient reason; the law of classification: "Everything can be classified," which includes the *dictum de omni et nullo*; and the law of correlativity: "There is nothing relative without an absolute." Only the first three laws, in a slightly different version, are found in earlier works. The fourth law was added in the "Introduction to Philosophy."

Pauler's metaphysics is a combination of Aristotelian and Leibnizian elements, but by the end of his life it had moved toward Platonism and Neoplatonism. A substance is a center of self-activity based on intention or wish (*vágy*); the body is a manifestation of this activity. The interaction of substances not only proves their plurality but also provides the unity of the world. Since all change is from potentiality to actuality, the whole world process is a self-realization and self-liberation. All substances strive toward the first principle of their development, the principle of self-liberation, which is the Absolute. Moreover, substances exist insofar as they strive toward the Absolute. At first, God was described as something other than the Absolute, but Pauler later developed this Absolute into a theistic concept. He also introduced the Platonic *anamnesis* and the Augustinian *illuminatio* into his theory of knowledge.

Toward the end of his life he seems to have identified his reductive method with the Platonic dialectic, and his reductive method ultimately leads us to the notion of Good. He also criticized Aristotle for having misunderstood Plato. According to Pauler, Aristotle was mistaken in assuming that the Ideas are in the field of reality. They are, in fact, in the field of validity; that is, we do not come to them in the search for new entities, but in the search for those presuppositions without which we cannot think validly. We do this not by induction or deduction but by reduction.

See also Absolute, The; Aristotle; Logic, History of; Metaphysics; Neoplatonism; Platonism and the Platonic Tradition; Positivism.

Bibliography

There is a full bibliography of Pauler's published and unpublished works in *Pauler Akos Emlékkönyv* (Budapest, 1933), a special number (No. 6) of the publications of the Hungarian Philosophical Association that is devoted to Pauler.

See also J. Somogyi, "Die Philosophic Akos Paulers," in *Kantstudien* 30 (1925): 180–188, and C. Carbonara, "Akos von Pauler e la logica della filosofia dei valori," in *Logos* 3 (1931).

German translations of Pauler's works are *Grundlagen der Philosophie* (Berlin: W. de Gruyter, 1925) and *Logik* (Berlin and Leipzig: W. de Gruyter, 1929).

Julius Kovesi (1967)

PAULING, LINUS

(1901–1994)

Linus Pauling was a chemist, peace campaigner, and double Nobel Laureate who played a central role in two great unifying projects of twentieth-century science.

Born in Oregon, U.S.A., in 1901, Pauling worked his way through college, receiving a BS in chemistry from Oregon Agricultural College in 1922. There he read papers on valence by physical chemists G. N. Lewis and Irving Langmuir, sparking his interest in the theory of chemical structure and bonding. He moved to California Institute of Technology (Caltech) for doctoral work on X-ray studies of inorganic crystal structures and had published twelve papers by the time he graduated in 1925. In 1926 he traveled to Europe on a Guggenheim postdoctoral fellowship, visiting Munich and other centers of the new quantum mechanics. On his return from Europe, Pauling resumed his work on X-ray crystallography, developing what he later called his *chemical intuition* about possible crystal structures. He also set about applying quantum mechanics to chemistry. Simultaneously with physicist John Clarke Slater, he developed physicists Walter Heitler and Fritz London's 1927 work on the hydrogen molecule to explain the structure of polyatomic molecules. The resulting valence-bond approach to molecular quantum mechanics, which modeled observed molecular structures as *resonance hybrids* of classical structures, faced competition from the *molecular-orbital* approach. The early success of the valence-bond approach is largely due to Pauling's advocacy, his developing intuitive visual representations to accompany his theoretical work, and his publication of the enormously influential *Nature of the Chemical Bond* (1939), which brought together his many contributions to structural chemistry.

Despite this central role in unifying the sciences, Pauling was no reductionist. He regarded his application of quantum mechanics to chemistry as a synthesis of physical theory with independent principles of chemical structure.

Pauling's second great unifying project was the chemical understanding of biologically important molecules. From the 1930s onward, he applied the X-ray and electron-diffraction methods, used earlier on inorganic crystals, to the structure of peptides and proteins, including hemoglobin. Subsequently, Pauling studied the molecular basis of the immune system and identified the first molecular disease—sickle-cell anemia. Pauling's work was also influential in James Watson and Francis Crick's proposal of a double-helix structure for DNA in 1953, though Pauling denied having participated in a *race* to discover the structure of the molecule.

Pauling was a controversialist in science and in politics: Though he publicly defended Japanese internees, he supported U.S. entry in to the Second World War and was active scientifically in the war effort, earning a Presidential Medal of Merit in 1948. During the cold war, however, he became increasingly involved in campaigning for nuclear disarmament and for a test-ban treaty on both political and scientific grounds. This, and his defense of blacklisted scientists, led to interest from the Federal Bureau of Investigation and the denial of a passport in the early 1950s. A passport was forthcoming, however, when Pauling was awarded the 1954 Nobel Prize in Chemistry for his work on the chemical bond and his contributions to the understanding of the structure of proteins. His political campaigning also earned him a second Nobel Prize (in Peace) in 1962.

Pauling left Caltech in 1964, partly as a result of his high political profile, spending the next decade at the Santa Barbara Center for the Study of Democratic Institutions (1964–1967), the University of California at San Diego (1967–1969), and Stanford University (1969–1973). On retirement from there, he cofounded the Linus Pauling Institute of Science and Medicine in Palo Alto, California, from where he continued his popular, though scientifically controversial, advocacy of high doses of vitamin C to improve health and to slow down aging. He remained active in research until nearly the end of his life.

See also Chemistry, Philosophy of; Peace, War, and Philosophy; Quantum Mechanics; Social and Political Philosophy.

Bibliography

Hager, Thomas. *Force of Nature: The Life of Linus Pauling*. New York: Simon and Schuster, 1995.

Mason, Stephen. "The Science and Humanism of Linus Pauling (1901–1994)." *Chemical Society Reviews* (1997): 29–39.

Pauling, Linus. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals; An Introduction to Modern Structural Chemistry*. 3rd ed. Ithaca, NY: Cornell University Press, 1960.

Pauling, Linus. "Fifty Years of Progress in Structural Chemistry and Molecular Biology." *Daedalus* 99 (4) (Fall 1970): 988–1014.

Robin Findlay Hendry (2005)

PAUL OF VENICE

(1369–1429)

Paolo Nicoletto Veneto joined the Hermits of St. Augustine as a boy and later taught at the Augustinian convent and the University of Padua for most of his life. The

order's *Register* lists him at the *Studium* in Oxford from 1390 to 1393 where he studied theology but not logic, as often believed. Briefly, he served as prior general of the Augustinian order and later as ambassador to Cracow, Poland. In 1420, he was implicated in sedition against the Venetian Republic, was banished, and spent his last years in Siena and Perugia.

More than twenty works, extant in some 270 manuscripts, are attributed to him, but Paul's authorship of some of those works is questionable. His popular *Logica Parva* transmitted elementary Oxford logic to Italy. His *Lectura super librum Posteriorum Analyticorum* and *Summa Naturalium* were similarly important for conveying the Oxford style of scientific investigation to Italy. Judged by the number of manuscripts, other works had less influence, for example, *Lectura super librum de Anima*.

The *Logica Magna*, a gigantic work (200 folios) attributed to Paul, exists in only one manuscript and two fragments. This encyclopedic album covers most topics of scholastic logic that were disputed at Oxford in the last half of the fourteenth century. Its author undoubtedly took part in those debates that occurred while Paul was yet unborn or still a youth. With few exceptions, inconsistencies of doctrine, rules, and examples between *Logica Magna* and *Logica Parva*, as well as other factors, make it highly unlikely that they were written by the same person.

Logica Parva contains the core of scholastic logic that remained resilient against Humanist criticism well into the modern world. Focusing on logical form, it distinguishes between the logical signs (e.g., of affirmation/negation, of quantification, of conjunction, disjunction and implication) and nonlogical signs (ordinary nouns and verbs) of a language. Next, it gives inference rules (*consequentia*) keyed to the logical signs. Finally, it supports a truth-conditional concept of truth in which the truth of a sentence is decidable in virtue of its logical form. Material supposition serves as a quotational device within a meta language where any sentence of the object language can be quoted. Translation is understood as the substitution of one sentence for another in virtue of their common logical form and comparable nonlogical terms.

Paul of Venice organized and conveyed Oxford learning to Italy in the early fifteenth century. Humanists who urged a return to classical Latin usage and condemned the *barbari britanni* undoubtedly had works like his in mind, but few humanists read or understood them. Lorenzo Valla's *Dialectica* criticizes the *logica vetus* of Boethius but ignores the *logica moderna*. J. L. Vives rejects *sophismata* as a pedagogical method in *Adversus pseudodialecticos* but retains Scholastic concepts under a neoclassical nomenclature in *De artibus*.

See also Augustinianism; Boethius, Anicius Manlius Severinus; Humanism; Logic, History of; Philosophy of Science, History of.

Bibliography

PRIMARY SOURCES

Logica Magna. Tractatus de suppositionibus, edited and translated by A. R. Perreiah. St. Bonaventure NY: Franciscan Institute, 1971.

Logica Magna. Part II, Fascicule 6: *Tractatus de Veritate et Falsitate Propositionis, Tractatus de SignificatoPropositionis*, edited by Francesco Del Punta, translated by Marilyn McCord Adams, 1978; Part I, Fascicule 1: *Tractatus de Terminis*, edited and translated by Norman Kretzmann, 1979; Part I, Fascicule 7: *Tractatus de scire et dubitare*, edited and translated by Patricia Clark, 1981; Part II, Fascicule 8: *Tractatus de Obligationibus*, edited and translated by E. Jennifer Ashworth, 1988; Part II, Fascicule 3: *Tractatus de Hypotheticis*, edited and translated by Alexander Broadie, 1990; Part II, Fascicule 4: *Capitula de Conditionali et de Rationali*, edited and translated by G. E. Hughes, 1990; Part I, Fascicule 8: *Tractatus de necessitate et contingentia futurorum*, edited and translated by C. J. F. Williams, 1991. Oxford: Oxford University Press.

MODERN EDITIONS

Ruello, F., ed. *Paulus Venetus, Super Primum Sententiarum Johannis de Ripa Lecturae Abbreviatio*. Firenze: Leo S. Olschki, 1980.

Perreiah, A. R. trans. *Paulus Venetus Logica Parva*, (an English translation of the 1472 edition with introductory essay and notes). Munich: Philosophia Verlag, 1984.

Perreiah, A. R., ed. *Paulus Venetus Logica Parva, First Critical Edition from the Manuscripts with Introduction and Commentary*. Leiden, NY: E. J. Brill, 2002. Bibliography, pp. 301–310.

SECONDARY SOURCES

Conti, Alessandro D. *Esistenza e Verita: Forme e strutture del reale in Paolo Veneto e nel pensiero filosofico del tardo medioevo*. Roma: Istituto Storico Italiano per il Medio Evo, 1996. Bibliography, pp. 301–316.

Perreiah, A. R. *Paul of Venice: A Bibliographical Guide*. Bowling Green, OH: Philosophy Documentation Center, 1986.

Alan R. Perreiah (2005)

PAULSEN, FRIEDRICH
(1846–1908)

Friedrich Paulsen, a German philosopher and educational theorist, was born in the village of Langenhorn, Schleswig-Holstein, to a farming family descended from generations of seamen of the North Frisian Islands. In his

autobiography Paulsen described his early life in detail, attributing to it the firm moral character and concern for people that marked his later work in philosophy and education. After attending the Altona Gymnasium, he entered the university at Erlangen in 1867. The following year he went to the University of Berlin, where a reading of F. A. Lange's *History of Materialism* and participation in Adolf Trendelenburg's seminar on Aristotle induced him to abandon theology for philosophy. After studies in Berlin, Bonn, and Kiel, Paulsen taught at Berlin. The professorship of philosophy to which he later succeeded there was, due to his own interests and the needs of the university, expanded to include pedagogy.

Philosophy could not, for Paulsen, be detached from the moral and cultural issues of private and public life, and the needs of the general public determined both the language and the content of his teaching and writing. Although far from negligent of the critical problems of theoretical and practical philosophy, he always tested the validity of their solutions by common sense and the public well-being. His collection of essays and addresses *Zur Ethik und Politik* (1905) shows the range of his interests and his public concern. Although he was temperate and reasonable, his efforts to distinguish good from evil in contemporary political and social life subjected him to political attack and involved him in public controversy.

Although Paulsen influenced all levels of German education, his published works deal chiefly with German universities and preparatory schools. His *Geschichte des gelehrten Unterrichts auf den deutschen Schulen und Universitäten* (1885) pioneered in the history of higher education and aroused wide controversy, helping to effect a liberalization of preuniversity education.

Paulsen usually described his philosophical position as idealistic monism but sometimes described it as pantheism. Participating in the revival of Immanuel Kant and Aristotle in the second half of the nineteenth century, Paulsen found in both an epistemological realism, an emphasis upon practical reason over theoretical reason, and a teleological metaphysics. His own position was formulated in opposition to the two extremes of a rigid Christian orthodoxy and scientific materialism. Irrational supernaturalism and mechanistic naturalism are the enemies in his two textbooks, *System der Ethik* (1889) and *Einleitung in die Philosophie* (1892), and in his *Philosophia Militans* (Berlin, 1901). He rejected Christian supernaturalism because of its dualism in theoretical philosophy and its legalism and rigorism in practical philosophy. Materialism was discarded because its denial both of human freedom and of the reality of purposes is offensive to man's ethical demands.

Paulsen's two textbooks were addressed not merely to students but to the thoughtful layman. Simply written with many concrete applications and references to contemporary ethical and social problems, they appeared in many editions in German and in translation and set a pattern for introductory textbooks and courses in philosophy for at least four decades. In them Paulsen formulated his method as (1) analysis of problems and the construction of possible solutions, (2) a survey of the historical development of philosophical thought on each problem, and (3) a choice of the solution most coherent with an inclusive world view.

This method brought Paulsen close to a pragmatic and personalistic viewpoint. In his ethics he supported a modern utilitarianism or eudaemonism that repudiated the hedonism of the British school, replacing it with the goal of human welfare and an objective perfection of the ends of life. The good life is thus grounded in the will, not in feeling. In determining the valid ends of conduct, the individual must be guided by the historical tradition, which may be trusted ultimately to destroy evil and to bring about the survival of the good. Book I of the *System,* devoted to such historical evaluation, is still a most useful introduction to the history of ethics. Paulsen stressed the distinctions between the ascetic ethical ideals of early Christianity and the humanism of classical Greece, but he regarded as necessary the modern effort to reconcile them.

Ethical thought involves the problems of evil, of freedom, and of God. Evil is justified in a monistic world, because by overcoming evil we find the way and the will to the good. Although human freedom is real, it is never a motiveless freedom of action. The psychological theory of freedom is correct in finding the ground of free action in the human will or in man's determining his conduct through deliberation and resolution. The metaphysical theory of freedom, which denies that there are causes of the will, must itself be denied. Morality, in its historical development of responsibility and a sense of duty, comes to require a higher will with a right to command and thus provides an argument for the existence of a deity who is also implicit in the evolutionary account of nature.

In such later ethical writings as the article "Ethik" in Paul Hinneberg's *Systematische Philosophie* (Berlin and Leipzig, 1907), Paulsen moved closer to G. W. F. Hegel by introducing an "objective will" as the manifestation in the social forms of life of a universal reason to which individual conscience is a cognitive response. Paulsen held

that the principles of ethics are rational in the sense that they arise from the conditions of life. They need not determine one's metaphysics, but teleological ethics demands an evolutionary teleology in which the purpose of nature is fulfilled in human reason.

Paulsen's *Introduction to Philosophy* was devoted to metaphysical and epistemological questions. In it he is led to monism by the Lotzean argument from finite interaction, by E. Hartmann's vitalism and energism, and by a creative vitalistic interpretation of evolution. His solution to the mind-body problem is a theory of panpsychistic parallelism, showing the influence of Benedict de Spinoza and Gustav Theodor Fechner. Mind and body are distinct aspects of a unified "All-One," a mental process of which history and nature are the two series of "modifications." This identity is affirmed of God in relationship to nature and to history. Science is limited to the phenomenalistic aspect of nature. Although God enters into interaction with lesser spirits, the concept of personality must be purged of its human limitations before it can be ascribed to God, who is to be thought of rather as a superpersonal source of energy and reason in nature and man.

See also Aristotle; Fechner, Gustav Theodor; Hartmann, Eduard von; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Lange, Friedrich Albert; Materialism; Neo-Kantianism; Panpsychism; Pantheism; Spinoza, Benedict (Baruch) de; Utilitarianism; Vitalism.

Bibliography

WORKS BY PAULSEN

Geschichte des gelehrten Unterrichts auf den deutschen Schulen und Universitäten vom Ausgang des Mittelalters bis zur Gegenwart. Leipzig, 1885. 3rd ed., 2 vols., Berlin, 1919–1920, translated by E. D. Perry as *The German Universities: Their Character and Historical Development.* New York: Macmillan, 1895.

System der Ethik mit einem Umriss des Staats und Gesellschaftslehre. Berlin, 1889. Translated from the 4th German edition by Frank Thilly as *A System of Ethics.* New York: Scribners, 1899.

Einleitung in die Philosophie. Berlin, 1892. Translated from the 3rd German edition by Frank Thilly as *Introduction to Philosophy,* New York: Holt, 1895.

Immanuel Kant: Sein Leben und seine Lehre. Stuttgart: Frommanns, 1899. Translated from the 3rd German edition by J. E. Creighton and A. Lefevre as *Immanuel Kant: His Life and Doctrine.* New York: Scribners, 1902. Paulsen's most important and influential historical study. Interprets Kant as a realist.

Philosophia Militans. Gegen Klericalismus und Naturalismus. Berlin: Reuther and Reichard, 1901.

Zur Ethik und Politik: Gesammelte Vorträge und Aufsätze, 2 vols. Berlin, 1905.

Aus meinem Leben. Jugenderinnerungen. Jena: Diederichs, 1909. Translated by Theodor Lorenz as *Friedrich Paulsen: An Autobiography.* New York: Columbia University Press, 1938. Gives details of Paulsen's academic career and political influence as well as bibliographical data. Part II of this work, never published in German, was translated from Paulsen's manuscript.

WORKS ON PAULSEN

Schulte-Hibbert, B. *Die Philosophie Friedrich Paulsens.* Berlin, 1914.

Speck, Johannes. *Friedrich Paulsen. Sein Leben und sein Werk.* Langensalza, Germany, 1926.

Spranger, Eduard, ed. *Gesammelte pädagogische Abhandlungen.* Stuttgart: J.G. Cotta, 1912. Contains a complete bibliography.

L. E. Loemker (1967)

PAVLOV, IVAN PETROVICH

(1849–1936)

Ivan Petrovich Pavlov, the Russian physiologist and originator of conditioned-reflex method and theory, was born the eldest son of a priest in Riazan'. After home tutoring, church school, and theological seminary (where he read G. H. Lewes's *Physiology)*, he entered the University of St. Petersburg, where I. F. Tsyon confirmed his physiological interests. At the Military Medical Academy, as assistant to Tsyon and later to S. P. Botkin, the experimental pharmacologist, he excelled in surgery and in experimental physiological research, which he continued in Botkin's laboratory after qualifying as an approved physician in 1879. In 1881 he married a fellow student, and despite desperate financial struggles, he received his MD in 1883 with a dissertation on the heart's innervation. With a traveling fellowship, he worked in Leipzig with Karl Ludwig and in Breslau with Rudolf Heidenhain; he returned to Botkin's laboratory in 1886 to continue research on nervous control of circulation and digestion. In 1888 he discovered the secretory nerves of the pancreas, and the following year he wrote on "sham feeding" and gastric "psychic secretion" (at sight of food).

In 1890 he became professor of pharmacology at the Military Medical Academy and director of the physiological department of the new Institute of Experimental Medicine donated by the prince of Oldenburg. In 1895 he was named professor of physiology at the Military Medical Academy, although the rector, Pashutin, delayed confirmation of the appointment till 1897. *The Work of the Digestive Glands* (1897), which reported the research that

won Pavlov the Nobel Prize in physiology in 1904, was widely translated. Next he investigated salivary "psychic" secretion, devising a neat surgical technique to enable collection and measurement of the saliva of dogs. Reflex salivation was measured upon ingestion (natural stimulus) and sight ("psychic" stimulus) of food, and also upon application, to hungry dogs before feeding, of artificial ("conditioned") stimuli—visual, auditory, olfactory, and tactile. The "conditioned reflex"—a term coined by I. E. Tolochinov—was thus a simple unit of acquired behavior, as involuntary as salivation itself; its formation, persistence, and disappearance followed rules that Pavlov elucidated in meticulous experiment for more than thirty years, gradually constructing a neurophysiological theory of behavior and learning. Pavlov's work attracted pupils and collaborators, produced a plentiful literature, and continued without significant interruption through World War I and the Russian Revolution.

A reflex theory of behavior accorded well with Marxist dialectical materialism, and Pavlov's researches received governmental encouragement and financial support. Pavlov was never a Marxist or a communist; he resigned his professorship in 1924 in protest against anticlerical discrimination at the academy, but he continued to enjoy state support, including new laboratories, and official foreign-language publication; his research village, Koltushy, was even renamed Pavlovo. When conditioned-reflex theory was extended to human behavior, Pavlovian doctrine became the Soviet Union's official "psychology," basic to psychiatry, pedagogy, industrial research, and other fields ranging from criminal reeducation to space exploration.

Pavlov's collected lectures appeared in English, French, and German translations in the 1920s, with a further volume, *Conditioned Reflexes and Psychiatry*, in 1941. He observed that a conditioned reflex might comprise excitation (secretory or motor) or inhibition, both processes located in the cerebral cortex. Concentration and irradiation of excitation, enabling discrimination and generalization of response, followed laws of induction, conceived as resembling ionic polarization, with excitation and inhibition spreading wavelike over a largely unspecialized cortex. Specialization occurred in the analyzers, or cortical receptor areas (visual, auditory, etc.), which sorted stimulus signals and regulated responses.

Pavlov found that for permanence a conditioned reflex required reinforcement with the unconditioned stimulus. Disturbance of an already established temporal or spatial pattern of stimuli, including excessive requirement of discrimination, produced disordered responses in the three successive phases of (*a*) equalization of response to all stimuli, (*b*) paradoxical responses, and (*c*) ultraparadoxical responses, involving reversal of positive and inhibitory responses. Ultimate derangement ("neurosis") was behavioral breakdown in uncontrolled excitement or complete inhibition, depending upon the type of the nervous system. An increasing preponderance of inhibition was evident in the progression from (*a*) controlled activity, to (*b*) delayed activity, corresponding to deliberation or thought, to (*c*) hypnotic states with concentrated activity bounded by general inhibition, to (*d*) sleep considered as generalized inhibition. Nervous systems were classified as strong excitable, weak inhibitable, and two central "balanced" types, lively and stolid, analogous to the "Hippocratic temperaments," choleric, melancholic, sanguine, and phlegmatic, respectively. Conditioned reflexes were most stable in the two more inhibited types of dog (and probably of humans).

From 1928 until his death Pavlov surveyed human psychology and psychiatry, drawing bold analogies between psychiatric syndromes and the reactions of dogs to experimental laboratory situations. Manic-depressive psychosis was viewed as an excitation-inhibition disorder and paranoia as a pathologically persistent excitatory process in a circumscribed cortical area. Later work by others has shown the value of conditioning theory for a "how" explanation and for an empirical treatment for certain phobias and compulsions, but Pavlov's formulations, without direct experimental or adequate clinical basis, are subjective intuitions clothed in pseudophysiological vocabulary. His experimental observations were objective and sound, and his apparently prosaic method allowed repeatable exact measurement, although what else was being measured by measuring saliva remains unclear. When he wrote of "reflexes" of freedom and slavery in dogs and humans, or of an animal's "strong" or "weak" cortex, or of ripples of excitation or inhibition, he failed to recognize the subjective nature of his interpretations. Insight was hindered by his premature oversimplification and an increasingly militant materialist monism.

Pavlov's was the principal and most developed of the several physiopsychologies of his time. His priority was disputed by V. M. Bekhterev, a neurologist whose "reflexology" of "associated reflexes" was developed simultaneously although independently in the same academy; Pavlov undoubtedly published first, however. Pavlov yielded experimental priority to the American E. L. Thorndike and admitted the theoretical influence of I. M. Sechenov, a former professor of physiology in St. Peters-

burg, whom Pavlov styled "father of Russian physiology." Sechenov's *Reflexes of the Brain* (1863, in *Selected Physiological and Psychological Worts*, Moscow, 1952–1956) followed his studies in Berlin, where Wilhelm Griesinger taught a psychology of temperamental types and psychic reflexes that was philosophically based upon Arthur Schopenhauer and René Descartes (*Mental Pathology and Therapeutics*, Berlin, 1845 and 1861; translated by C. L. Robertson and J. Rutherford, London, 1867).

Pavlov's influence continues paramount in Russia. Elsewhere it is an important component in behavior theory and therapy, but with a strong admixture of Bekhterev and John B. Watson in practical techniques and a preponderance of C. L. Hull's learning theory in vocabulary.

See also Behaviorism; Descartes, René; Dialectical Materialism; Induction; Marxist Philosophy; Schopenhauer, Arthur.

Bibliography

WORKS BY PAVLOV

The Work of the Digestive Glands. Translated by W. H. Thompson. London, 1902.

Conditioned Reflexes; an Investigation of the Physiological Activity of the Cerebral Cortex. Translated and edited by G. V. Anrep. London: Oxford University Press, 1927; New York: Dover, 1960. The best English translation of the early papers.

Lectures on Conditioned Reflexes. Vol. I: *Twenty-five Years of Objective Study of the Higher Nervous Activity (Behavior) of Animals*. Translated and edited by W. Horsley Gantt, with G. Volborth. New York: International, 1928. Vol. II: *Conditioned Reflexes and Psychiatry*. Translated and edited by W. Horsley Gantt. New York: International, 1941.

Selected Work. Translated by S. Belsky, edited by J. Gibbons under the supervision of K. S. Koshtoyants. Moscow: Foreign Languages, 1955.

Experimental Psychology and Other Essays. New York: Philosophical Library, 1957; London, 1958. A reprint of *Lectures on Conditioned Reflexes*, omitting Pavlov's patriotic and political speeches.

Essays in Psychology and Psychiatry. New York, 1962. A selection from *Experimental Psychology and Other Essays* in paperback.

Psychopathology and Psychiatry; Selected Works. Compiled by Y. Popov and L. Rokhin, translated by D. Myshne and S. Belsky. Moscow: Foreign Languages, 1962.

Conditioned Reflexes, Lectures on Conditioned Reflexes, and *Selected Works* contain some of the same papers in different translations. *Lectures on Conditioned Reflexes* and *Selected Works* have some later "psychiatric" papers, and *Selected Works* has some early papers on circulation and digestion preceding Pavlov's work on conditioned reflexes. *Psychopathology and Psychiatry* is selected entirely from his "psychiatric" work.

WORKS ON PAVLOV

Asratyan, E. A. *I. P. Pavlov, His Life and Work*. English ed., Moscow: Foreign Languages, 1953. An early and approved biography.

Babkin, B. P. *Pavlov: A Biography*. Chicago: University of Chicago Press, 1949 and 1960; London: Gollancz, 1951.

Cuny, H. *Ivan Pavlov: The Man and His Theories*. London: Souvenir Press, 1964. Brief, semipopular outline of Pavlov's life and work.

Frolov, Y. P. *Pavlov and His School—The Theory of Conditioned Reflexes*. Translated by C. P. Dutt. New York: Oxford University Press, 1937.

Gray, J. A., ed. and tr. *Pavlov's Typology*. Oxford and New York: Pergamon Press, 1964. Includes a detailed authoritative survey of this field by B. M. Teplov.

Platonov, K. I. *The Word as a Physiological and Therapeutic Factor—The Theory and Practice of Psychotherapy according to I. P. Pavlov*. Translated by D. A. Myshne. Moscow: Foreign Languages, 1959. This work applying Pavlov's teachings openly acknowledges a debt to Bekhterev's.

Scientific Session on the Physiological Teachings of I. P. Pavlov. Moscow: Foreign Languages, 1951.

Simon, B., ed. *Psychology in the Soviet Union*. Stanford, CA: Stanford University Press, 1957.

Todes, P. *Ivan Pavlov: Exploring the Animal Machine*. New York: Oxford University Press, 2000.

Todes, P. *Pavlov's Physiology Factory: Experiment, Interpretation, Laboratory Enterprise*. Baltimore, MD: Johns Hopkins University Press, 2002.

Wells, H. K. *Pavlov and Freud*. 2 vols. New York: International, 1956–1960; reprinted 1962.

J. D. Uytman (1967)

Bibliography updated by Vladimir Marchenkov (2005)

PEACE, WAR, AND PHILOSOPHY

Speculation about war and peace as conditions of interstate relations has tended to divide thinkers into two groups—those who regard war as inevitable, perhaps even desirable, and those who consider it an evil capable of being replaced by lasting peace through good will or improved social arrangements. The first group is sometimes described as "realist" and the second as "idealist," but these terms have the drawback that such idealist philosophers (in the ontological sense) as Plato and Georg Wilhelm Friedrich Hegel often accept war as a permanent condition of human existence. It is therefore proposed here simply to call the first group "conservatives" and the second "abolitionists," though a wide spectrum of opinion clearly exists within each subdivision.

THE CONSERVATIVE TRADITION

THE GREEKS. Ancient Greek thought commonly accepted war between the city-states themselves and between Greeks and "barbarians" as part of the order of nature. The Greek gods were a warlike breed who had come to power after a brutal struggle with the Titans. Ares was one of their leading figures, but the goddess of peace, Irene, was merely a subordinate deity attendant on the great gods. A view of war widely prevalent in Greece was that of Heraclitus of Ephesus. War, Heraclitus taught, was the "father of all and king of all," and it was through war that the present condition of humankind, some men free and some enslaved, had evolved. If strife between the warring elements in nature were abolished, nothing could exist; "all things," according to Heraclitus, "come into being and pass away through strife."

It was not until the later phases of the war between Athens and Sparta (431–404 BCE) that a pacifist note unusual in the Greek world was struck in such works as Euripides' *The Trojan Women* (performed in 415 BCE) and Aristophanes' *Lysistrata* (411 BCE). Even so, the conclusion drawn by Plato from the Peloponnesian War was that the state must be organized for violent survival in an unruly world. Plato's *Republic* is, in effect, a design for a military community on the Spartan model. Plato does, however, distinguish between war among Greeks and war between Greeks and outsiders; the former, according to the *Republic,* is to be legally regulated whereas any excess is permissible in the latter.

CHRISTIANITY AND NATURAL LAW. The conservative acceptance of war as a fact of life was also basic to the intellectual attitudes of the Roman Republic and Empire and was sustained during the Middle Ages, when Catholic writers wrestled with the problem of the conditions on which ecclesiastical approval could be given to the wars of secular monarchs. St. Thomas Aquinas in *Summa Theologiae* (Question 40), while claiming that peace was the greatest aim toward which man should strive in fulfillment of his natural ends, nevertheless placed on monarchs the duty to defend the state. Similarly, Dante contended in *De Monarchia* that "peace was the target at which all shafts were sped" but that it was to be attained by the imposition of a world law, if necessary by force, issuing from a revived Roman Empire. The legacy of Christian teaching that had the most lasting influence, however, concerned the application of natural law, strongly tinged by Christian ethics, to the conduct of war.

The Spanish Jesuit theologian Francisco Suárez held that war is not intrinsically evil and that just wars may be waged. Suárez defined three conditions of legitimate war. It must be waged by lawful authority—that is, by the supreme sovereign; the cause of making war must be just, and other means of achieving justice must be lacking; and war must be conducted and peace imposed with moderation. A similar view was taken by Hugo Grotius, who held that far from war's being a breakdown of the law of nations, it is, in fact, a condition of life to which law is as applicable as it is to the conditions of peace. War, Grotius argued in his *De Iure Belli ac Pacis Libri Tres* (1625), should not be fought except for the enforcement of rights and, when fought, should be waged only within the bounds of law and good faith. This conception survives in the assumption behind such twentieth-century international organizations as the League of Nations and the United Nations that only wars fought on behalf of international interests, such as the maintenance of world peace, are just.

THE ADVENT OF NATIONALISM. In the era of European secular nationalism following the Renaissance the idea of war as a necessary or desirable institution strengthened. The Italian city-states of the Renaissance, whose diplomatic practice formed the model for the early European national states, were continually at war with one another; these were, however, limited conflicts that aroused no great indignation among philosophers. A typically acquiescent view of war was that of Sir Thomas More in his *Utopia* (1518). The Utopians have a pragmatic, not particularly heroic idea of war, which they regard as a normal event; war is to be fought as economically and safely as possible when one's lands are invaded or one's allies are oppressed.

A more profound view was that of the Florentine statesman and writer Niccolò Machiavelli. Like all conservatives, Machiavelli assumed that armed conflict was part of the human lot not because man was evil—Machiavelli was inclined to regard man as weak and stupid rather than evil—but because of the activity of malign fate (*fortuna*), which is always forcing man to arm himself against adversity. Machiavelli, unlike Heraclitus, held out no hope that war raised man to a higher plane; the prince is condemned to seek victory in war merely in order to survive in the hostile world. In peace a ruler should not sit with hands folded but should always be improving his state's military power against the day of adversity.

At the same time the formation of great national states in England and France was forcing men to speculate on the justification of government, especially since the acceptance of the papacy as the ultimate and sacred

authority had been considerably weakened. The concept of a "state of nature" in which men exist without a common superior and in a state of internecine war was introduced to help explain the growth and functions of government. Thomas Hobbes explained in his *Leviathan* (1651) that war is not the act of fighting but the disposition *to fight* that exists where there is no common superior to ensure that violence shall not be permitted. Only through the establishment of a commonwealth—that is, a superior law-enforcing agency to which all men are subject—can peace and civilization be ensured. Hobbes did not regard the state of nature as a historical condition that had occurred in the past; he inferred that such would be man's state if the commonwealth did not exist.

John Locke differed from Hobbes in holding that there were natural rights in the state of nature that it was government's function, after its establishment, to protect; hence, war was not a universal condition in the state of nature but occurred only when force was exercised without right. For Locke there was an intrinsic difference between war waged for natural rights and war waged without this sanction. For Hobbes war in the state of nature, as well as war between sovereign states, could be neither right nor wrong since these categories exist only within the commonwealth. Benedict de Spinoza shared Hobbes's view of the inevitability of war where men are without a common government, but, like Locke, he could not reconcile himself to the total absence of morality or law in the state of nature. The Hobbesian argument has nevertheless been of immense importance in shaping modern Western man's attitude toward war and peace. It is that peace is the result of man's determination, deriving from fear of death and the wish for what Hobbes called "commodious living," to create an overriding government. Hobbes did not make clear whether he thought that man could sustain peace in his international relations, but it is clear that, unlike Locke, he considered that nothing short of a world state with a monopoly of power over the nations would suffice to ensure such peace.

Before the Napoleonic Wars, however, war, owing to its limited scale, could not be regarded as the decisive factor in the health or illness of nations. But with the Messianic fervor unleashed by the French Revolution, all Europe appeared to be caught up in revolt against the existing order, internal and external, and the expansion of national wealth showed for the first time the potentialities of nationalistic wars for good or evil. It was in the aftermath of the revolution that the more extreme conservative attitude toward war came into its own in certain countries and war began to be thought of as a positive principle of national regeneration. Germany in particular fostered these views, possibly because that country entered the struggle for national ascendancy somewhat late so that its militarism was proportionately more intense.

Hegel is well known for his conception of history as a struggle of opposites from which a synthesis emerges that transcends the two original conflicting forces. For Hegel the national state was the means by which the Idea realized itself in history. Since the Idea can materialize itself only if the state is allowed to live out its predetermined functions, it follows that the individual's life has no meaning except insofar as it serves the state's ends and that no principle is left by which the relations between states can be subject to moral criteria. Hegel had no patience with the notion of a league of nations for the establishment of permanent peace because he believed war was the catalyst through which history unfolded its purpose. Man must accept war or stagnate.

Arthur Schopenhauer rejected Hegel's idea of the state as the divine expression of justice. For him the state exists because there is injustice; the state is needed to protect man against the effects of his own egotism. In turn, man's egotism and his generally evil nature are a reflection of the dissonances of the Will that for Schopenhauer lies behind the world's realities. Under these conditions war is inevitable, but Schopenhauer, unlike Hegel, did not see war as a progressive factor in history but as a result of the immaturity and weakness of the masses and the love of luxury and power of their strong-willed leaders. Schopenhauer saw no hope of lasting peace.

THE MILITARISTS. Friedrich Nietzsche may be judged as an extreme representative of the romantic cult of war and as marking the transition to modern totalitarian militarism. Nietzsche was capable of deploring the wastefulness of war; however, in his fully mature writings, *Thus Spake Zarathustra* (1892) and *The Will to Power* (first published in 1901), he glorified war and the dangerous life. The phrase "a good war hallows every cause" (*Thus Spake Zarathustra*), may be taken as typical of this attitude. For Nietzsche's supermen war is a natural activity, the supreme witness to their superior quality; they should never succumb to the "slave morality" of Christianity, with its accent on humility, submissiveness, and turning the other cheek.

In the teaching of Heinrich von Treitschke the functions of the state were unlimited, as was the individual's duty to submit to its commands. The state's first duty was to maintain its power in its relations with other states and

to maintain law within its own borders; its second duty was the conduct of war, the crucible in which the elements in a state's greatness are fused. The hope of a world state or permanent peace is vain; the Aryan race can only keep by the sword what it has won by the sword. Treitschke admitted that the cost of war had risen steeply and, hence, that wars should be shorter and less frequent. But this did not affect the basic axiom that war is the "one remedy for an ailing nation."

Treitschke's ideas were absorbed by the German military writer Friedrich von Bernhardi, who used them to foster the militantly nationalist mood in which Germany entered World War I. In *Germany and the Next War,* Bernhardi repeated the basic notions of Treitschke: War is the process by which the truly civilized nations express their strength and vitality, life is an unending struggle for survival, war is an instrument in biological evolution. And Bernhardi drew on other conservative writers: Heraclitus; Frederick the Great, whose writings represented war as bringing out man's finest qualities; and Karl von Clausewitz, who described the nation's place in the world as a function of the interplay between its national character and its military tradition.

The conservative-militarist tradition, with its racist overtones, was inherited by the German Nazi and Italian fascist writers of the interwar period, though these added little to the work of their forebears. More recently, the advent of nuclear weapons has made nonsense of the glorification of war, though belief in its inevitability is still not uncommon. Almost the only considerable section of contemporary opinion that believes that national survival after nuclear war is conceivable is that of the Chinese communists. Even they, however, are careful to insist that they would never initiate a nuclear war, and it is, moreover, a feature of all communist thought that the final global victory of communism will remove all cause of war. Communists therefore differ from the conservatives we have considered in that although they regard war as contingent (or perhaps inevitable) in a capitalist system, they have no doubt that permanent peace is attainable under communism.

THE ABOLITIONISTS

THE PREMODERN AGE. As we have seen, the ancient Greeks (and the same may be said of the writers of the Roman world) were not distinguished for protests against war, though the Stoics of the Roman Empire preached a cosmopolitanism that assumed the oneness of all humankind, making war between its members an affront. When Stoicism was embraced by the Roman emperors, however, it lost its pacifist element, and the same may be said for the early Christian doctrine of nonviolence. Also, during the Middle Ages the fact that the papacy was both the supreme fount of church doctrine and a temporal power of considerable military strength ruled out complete pacifism as a church doctrine.

The outstanding opponent of war during the Renaissance was the great humanist Desiderius Erasmus, though it is incorrect to speak of him as an absolute pacifist. In his *Anti-polemus, or the Plea of Reason, Religion and Humanity against War* (1510), Erasmus argued that every man's duty was to spare no pains to put an end to war. War was directly opposed to every purpose for which Erasmus conceived man to have been created; man is born not for destruction but for love, friendship, and service to his fellow men.

PROJECTS FOR EUROPEAN PEACE. During the seventeenth century speculation in Europe about the possibility of permanent peace began to develop, stimulated by growing international commerce and the desire to bind Europe together in a final effort to expel the Turks. This anti-Muslim aim had already figured prominently in the plan for the unification of Europe designed by Pierre Dubois in *De Recuperatione Terre Sancte* (1305–1307) and in the celebrated proposal for a federation of Christian princes that George of Poděbrad, king of Bohemia, had presented to his fellow monarchs in 1461. The seventeenth-century proposals were immensely varied, ranging from utterly Utopian ideas to some that might have achieved realization as limited international alliances. Some were limited to Western Europe, others included all Europe, and some embraced the whole Christian world. "The Grand Design" (1620–1635), probably compiled by the duke of Sully, the chief minister of Henry IV of France, and *Some Reasons for an European State* (1710) by John Bellers both proposed to divide Europe into provinces of roughly equal size under a common government. A few schemes, such as Emeric Crucé's *The New Cyneas, or Discourse of the Occasion and Means to Establish a General Peace and the Liberty of Commerce throughout the World* (1623), aimed at the formation of a single world state with all the races and religions under its jurisdiction. In these plans provision was generally made for some form of representative government. William Penn in *An Essay towards the Present and Future Peace of Europe* (1693) contemplated annual European parliaments; the Abbé de Saint-Pierre in *A Project for Settling an Everlasting Peace in Europe* (1713) preferred a perpetual congress in order to reflect the viewpoints of the states in his European federation; Crucé called for world assemblies. These

confederations were chiefly advocated as defenses of peace, though other aims were also mentioned; Henry IV and the duke of Sully, for instance, had in mind, besides European peace, wars against the Muscovites and Turks and the weakening of the Hapsburgs as the preliminary steps to uniting Europe under French hegemony.

In the eighteenth century these peace plans were given a new lease of life with the French and German Enlightenment. Jean-Jacques Rousseau took the peace project of the Abbé de Saint-Pierre and applied it to the Europe of his own day in *A Project of Perpetual Peace* (1761), with the insistence that unless the proposed central authority was powerful enough to overawe all the constituent states, the proposal would fail. Rousseau recommended the plan to governments on the ground that a single European authority strong enough to enforce peace would also ensure internal stability in the constituent states. He admitted, however, that governments were probably too shortsighted to appreciate the merits of the plan. A similar project of European confederation was that of Immanuel Kant, titled *Eternal Peace* (1795). Kant's recipe is notable for its claim that the maintenance of peace requires the achievement of constitutional government by the states.

NINETEENTH-CENTURY PEACE MOVEMENTS. The nineteenth century was even more prolific in its plans for organizing the nations to ensure peace. In Europe and the United States there arose strong unofficial peace movements that urged the creation of agencies for the arbitration of interstate differences and the equitable settlement of political issues, together with the strengthening and codification of international law. In the atmosphere of harmony that followed the Congress of Vienna the Great Powers of Europe met regularly to deal with threats to peace, while such functional organizations as the European river commissions and the Universal Postal Union (1875) dealt quietly with matters of practical concern to the nations. The hope of a permanent international assembly that might develop into a world legislature was held out at the Hague conferences of 1899 and 1907, and it seemed likely that the growing stake of nations in peaceful intercourse would soon render war obsolete.

The English utilitarians, such as Jeremy Bentham, James Mill, and John Stuart Mill, provided much of the theoretical background of the peace movements. They contended that war was an anachronistic encumbrance on a free society, benefiting no one but aristocrats and professional soldiers. Richard Cobden voiced the commercial classes' distaste for war in his pamphlet *Russia* (1836). Herbert Spencer, an extreme exponent of laissez-faire society, denounced war in his *Social Statics* (1851) as an outcome of excessive government authority; with the functions of government reduced and individual liberty restored, all reason for war would disappear. This liberal, economic case for peace culminated in the striking claim by Norman Angell in *The Great Illusion* (1908) that war had become so destructive of all economic values that nations would never again engage in it.

PACIFICISM AND INTERNATIONALISM. World War I disastrously falsified Angell's prophecy; nevertheless, it reinforced the conviction of liberal-minded people that war was an absolute evil and that the creation of expedients to keep the peace, such as the League of Nations and collective security, was the most urgent task of the twentieth century. A strong cleavage now became apparent between absolute pacifists—for example, H. M. Swanwick, Gerald Heard, Aldous Huxley—and those who supported "just" wars fought under the league's aegis—for example, Gilbert Murray, Lord Cecil of Chelwood, P. J. Noel-Baker. Few of the abolitionists, however, considered a world federation necessary to ensure permanent peace. John Dewey, for instance, argued in the 1920s that it would be sufficient for states to agree to declare war illegal and to prosecute countries that resorted to it as criminals.

The advent of World War II and the invention of nuclear weapons, followed by the failure of the great powers to act unanimously in the United Nations Security Council, raised the question whether the abolitionists' aim can be attained short of the total surrender of national sovereignty. One curious effect of the nuclear stalemate has been to drive many abolitionists into the somewhat conservative belief that peace must be kept by the maintenance of a military balance between the two world camps. Others, like John Strachey in *On the Prevention of War* (London, 1962), contend that the two superpowers must go beyond this and exercise a kind of condominium over the rest of the world.

The outstanding British philosopher Bertrand Russell continued to believe that the rational conviction of the utter futility of nuclear war can in itself maintain peace provided that the realities of thermonuclear war are widely enough publicized (*Common Sense and Nuclear Warfare,* London, 1959). As a long-term measure, however, Russell saw no alternative to a world state, which must in the first instance be imposed by one nation or group of nations; only after the world authority has been in power for a century or so will it feel confident enough

to base its power on consent rather than force (*New Hopes for a Changing World,* London, 1951, p. 77). It is not clear, however, whether Russell really wished to pay the price of global despotism in return for peace; elsewhere, he wrote that a new war would be preferable to a universal communist empire (Robert E. Egner and Lester E. Denonn, eds., *The Basic Writings of Bertrand Russell,* London, 1961, p. 691). Here, in essence, is the issue facing the abolitionist in the nuclear age; whether war is a greater or lesser evil than the imposition on himself and his nation of hostile values which the present anarchic world, with its attendant threat of war, allows him to keep at a distance.

See also Bentham, Jeremy; Dante Alighieri; Dewey, John; Enlightenment; Erasmus, Desiderius; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Heraclitus of Ephesus; Hobbes, Thomas; Just War Theory; Kant, Immanuel; Locke, John; Machiavelli, Niccolò; Mill, James; Mill, John Stuart; More, Thomas; Nationalism; Nietzsche, Friedrich; Pacifism; Plato; Renaissance; Rousseau, Jean-Jacques; Russell, Bertrand Arthur William; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Stoicism; Suárez, Francisco; Thomas Aquinas, St.; Violence.

Bibliography

The following are sources for some of the views discussed above.

CONSERVATIVES

Bernhardi, Friedrich von. *Germany and the Next War.* Translated by Allen H. Powles. London: E. Arnold, 1914.

Clausewitz, Karl von. *On War.* Translated by J. J. Graham. London: K. Paul, Trench, Trubner, 1940.

Guthrie, W. K. C. *A History of Greek Philosophy,* Vol. I. Cambridge, U.K.: Cambridge University Press, 1962. Ch. 7 contains quotations from Heraclitus.

Hegel, Georg Wilhelm Friedrich. *Hegel's Philosophy of Right.* Translated by T. M. Knox. Oxford: Clarendon Press, 1942. Especially pp. 209–223.

Hobbes, Thomas. *Leviathan.* Edited by Michael Oakeshott. Oxford: Blackwell, 1960. Especially Ch. 13.

Schopenhauer, Arthur. *The World as Will and Idea.* Translated by R. B. Haldane and J. Kemp, 3 vols. London, 1909. Especially Vol. III, Ch. 46.

Spinoza, Benedict de. *The Political Works.* Edited by A. G. Wernham. Oxford: Clarendon Press, 1958. Especially p. 23.

Thomas Aquinas. *Selected Political Writings.* Edited by A. Passerin d'Entrèves. Oxford: Blackwell, 1959. Especially pp. 159–161.

Treitschke, Heinrich von. *Politics.* Translated by Blanche Dugdale and Torbende Bille, 2 vols. London, 1916. Especially Vol. I, pp. 61–70.

ABOLITIONISTS

Dewey, John. *Characters and Events,* 2 vols. New York: Holt, 1929. Vol. II, p. 670.

Kant, Immanuel. *Eternal Peace.* Translated by W. Hastie. Boston, 1944.

Rousseau, Jean-Jacques. *A Project of Perpetual Peace.* Translated by Edith M. Nuttall. London: R. Cobden-Sanderson, 1927.

SECONDARY SOURCES

Adams, Robert P. *The Better Part of Valor: More, Erasmus, Colet and Vives, on Humanism, War, and Peace.* Seattle: University of Washington Press, 1962.

Aron, Raymond. *Paix et guerre entre les nations.* Paris: Calmann-Lévy, 1962. A sociological and historical inquiry into the conditions in international relations that make for war and peace, together with an assessment of proposals for maintaining international equilibrium.

Beales, A. C. F. *The History of Peace.* London, 1931. A survey of movements, predominantly unofficial, for the promotion of world peace since the creation of the first "peace societies" in 1815. The book includes some useful chapters on nineteenth-and early-twentieth-century ideas on the maintenance of peace.

Hemleben, S. J. *Plans for World Peace through Six Centuries.* Chicago: University of Chicago Press, 1943. A handy abstract of famous plans for the maintenance of peace since the late Middle Ages.

McDonald, L. C. *Western Political Theory in the Modern World.* New York, 1962. A useful survey of modern Western political ideas, including thinking on war and peace; the book deals with both secular trends and individual thinkers.

Meinecke, Friedrich. *Machiavellism: The Doctrine of Raison d'État and Its Place in Modern History.* Translated by Douglas Scott. London, 1957. A translation of *Die Idee der Staatsräson in der neueren Geschichte* (1924), a treatise on the perennial conflict between the power impulse in human nature and the search for a higher ethical rule in political relations.

Stawell, F. M. *The Growth of International Thought.* London: Butterworth, 1929. A concise history of pacifism and internationalism.

Vagts, Alfred. *A History of Militarism.* New York: Norton, 1937. Mainly a study of the ideas and practices of the eighteenth- and nineteenth-century European social movements that sought to make military men the dominant power in the state.

F. S. Northedge (1967)

PEACE, WAR, AND PHILOSOPHY [ADDENDUM]

The nuclear threat that preoccupied Bertrand Russell receded into the background during the Vietnam War. After that war's end in 1975, the risk of a nuclear con-

frontation between the superpowers again became a major concern.

This issue came in for sustained moral analysis in Douglas Lackey's *Moral Principles and Nuclear Weapons* (1984) and Steven P. Lee's *Morality, Prudence, and Nuclear Weapons* (1993). But considerable philosophical interest focused more narrowly on so-called paradoxes of nuclear deterrence. Herman Kahn had spoken of "rationality of irrationality" strategies in *On Thermonuclear War* (1960). The question was whether it is rational to threaten to do the irrational (wage all-out nuclear war). The strategy of Mutual Assured Destruction (MAD) seemed to presuppose that it is, since it rested on the threat of massive retaliation in the event of a major nuclear first-strike. The moral version of the paradox, explored by Gregory S. Kavka in *Moral Paradoxes of Nuclear Deterrence* (1987), concerns whether it is moral to threaten to do the immoral (wage all-out nuclear war). This interest, and concern with the nuclear issue generally, waned with the end of the cold war following the collapse of the Soviet Union in the early 1990s.

Meanwhile, the 1960s saw a resurgence of interest in the just war theory, with both religious and secular attention to the doctrine extending into the twenty-first century. The religious approach had both Roman Catholic and Protestant advocates. Theologian Paul Ramsey set the tone for the Protestant approach in *War and the Christian Conscience: How Shall Modern War Be Conducted Justly?* (1961) and *The Just War: Force and Political Responsibility* (1968). The American Catholic Bishops detailed the Catholic position in their 1983 pastoral letter, *The Challenge of Peace: God's Promise and Our Responsibility*. An influential secular contribution appeared with political scientist Michael Walzer's *Just and Unjust Wars* (1977), in which a Hobbesian approach to political theory was adapted to the moral assessment of war. Philosophers quickly took up the issue of just war, particularly after the 1991 Persian Gulf War in which the United States under President George H. W. Bush expressly invoked the just war theory in defense of the U.S.-led war to drive Iraq out of Kuwait.

Just war theorists include both conservatives and abolitionists. Some regard war as virtually inevitable. They seek to ensure that it is undertaken only when justified and that its destructiveness is minimized. Others believe that war may eventually be done away with, but in the meantime, they believe, the moral criteria justifying resort to war and its conduct must be followed.

Set apart from just war theorists are pacifists, who believe that war, at least in the modern world, cannot be justified morally. Duane L. Cady provides a conceptual analysis of pacifism in *From Warism to Pacifism: A Moral Continuum* (1989). While theoretically one could be a "just war pacifist," holding that the just war theory contains the correct criteria for morally assessing war but maintaining that those criteria are never in fact met, most pacifists believe that just war criteria are inadequate, and that even if they are satisfied, they do not suffice to justify war. In particular, they reject the resort to the principle of double effect that would justify the foreseeable killing of innocents so long as it is not intentional and other conditions are met.

With the resurgence of feminism in the 1960s, some feminist philosophers took up the issue of war. While rarely strict pacifists, they tended to be abolitionists and to argue that war is a manifestation of patriarchal society and can be done away with only with the transformation of that society into one of gender equality. In particular, many of them, such as Sara Ruddick in *Maternal Thinking: Toward a Politics of Peace* (1989), see the key to a new way of thinking about war and violence in the distinctive experiences of women, particularly in mothering and caregiving.

RIGHTS AND SOVEREIGNTY

As the modern nation-state system began forming in the seventeenth century, the notion of the equality of states and their right to be free of interference in their internal affairs by other states eventually became the recognized (though not always honored) norm. The treatment of persons within a state's own borders was generally considered its own business. Toward the end of the twentieth century, there was wider acceptance of the idea that states could violate the sovereignty of other states if necessary to prevent crimes such as genocide and massacres of individuals within those states' borders. In the 1990s, genocide in Rwanda and so-called ethnic cleansing in the former Yugoslavia commanded particular attention in this regard. This presented philosophers and experts on international law with a challenge to show either that, properly understood, international law already allows such actions or that it could be modified to make room for them.

Thus, the world government that Russell proposed presents a challenge to state sovereignty from one direction, threatening to eliminate the plurality of independent sovereign states. The idea of unilateral military intervention for humanitarian reasons presents a challenge from a different direction, retaining the plurality of

states but making their sovereignty conditional upon their honoring of human rights.

Additionally, so-called low intensity conflicts and the rise of terrorism brought conceptual issues to the forefront. With the declared wars characteristic of the first half of the twentieth century receding into the past, even standard war itself, in the sense of vast armies arrayed against one another, may be phasing out. In its place, the twenty-first century is seeing terrorism, violence, guerrilla warfare, and flexible, far-reaching military actions, such as by the United States in Afghanistan and Iraq. Whether these represent new forms of war or a twenty-first-century substitute for war is a conceptual issue that philosophers and international lawyers have yet to decide. The challenge of peace, in any event, is to find nonviolent ways of dealing with the conflicts leading to these various forms of violence.

See also Feminist Philosophy; Just War Theory; Pacifism; Rights; Russell, Bertrand Arthur William; Sovereignty; Terrorism; Violence.

Bibliography

Cady, Duane L. *From Warism to Pacifism: A Moral Continuum*. Philadelphia: Temple University Press, 1989.

Chatterjee, Deen K., and Don E. Scheid, eds. *Ethics and Foreign Intervention*. Cambridge, U.K.: Cambridge University Press, 2003.

Cohen, Avner, and Steven Lee, eds. *Nuclear Weapons and the Future of Humanity: The Fundamental Questions*. Totowa, NJ: Rowman & Allenheld, 1986.

Christopher, Paul. *The Ethics of War & Peace: An Introduction to Legal and Moral Issues*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2004.

Elshtain, Jean Bethke, ed. *Just War Theory*. New York: New York University Press, 1992.

Elshtain, Jean Bethke, and Sheila Tobias, ed. *Women, Militarism, and War: Essays in History, Politics, and Social Theory*. Savage, MD: Rowman & Littlefield, 1990.

Gallie, W. B. *Philosophers of Peace and War: Kant, Clausewitz, Marx, Engels and Tolstoy*. Cambridge, U.K.: Cambridge University Press, 1978.

Gray, J. Glenn. *The Warriors: Reflections on Men in Battle*. New York: Harper & Row, 1967.

Holmes, Robert L. *On War and Morality*. Princeton, NJ: Princeton University Press, 1989.

Johnson, James Turner. *Morality and Contemporary Warfare*. New Haven, CT: Yale University Press, 1999.

Kahn, Herman. *On Thermonuclear War*. Princeton, NJ: Princeton University Press, 1960.

Kavka, Gregory S. *Moral Paradoxes of Nuclear Deterrence*. Cambridge, U.K.: Cambridge University Press, 1987.

Lackey, Douglas. *Moral Principles and Nuclear Weapons*. Totowa, NJ: Rowman & Allanheld, 1984.

Lee, Steven P. *Morality, Prudence, and Nuclear Weapons*. Cambridge, U.K.: Cambridge University Press, 1993.

Norman, Richard. *Ethics, Killing, and War*. Cambridge, U.K.: Cambridge University Press, 1995.

Ramsey, Paul. *The Just War: Force and Political Responsibility*. New York: Scribner, 1968.

Ramsey, Paul. *War and the Christian Conscience: How Shall Modern War Be Conducted Justly?* Durham, NC: Duke University Press, 1961.

Ruddick, Sara. *Maternal Thinking: Toward a Politics of Peace*. Boston: Beacon Press, 1989.

Sterba, James P., ed. *The Ethics of War and Nuclear Deterrence*. Belmont, CA: Wadsworth, 1985.

Walzer, Michael. *Just and Unjust Wars*. New York: Basic Books, 1977.

Wasserstrom, Richard A., ed. *War and Morality*. Belmont, CA: Wadsworth, 1970.

Robert L. Holmes (2005)

PEANO, GIUSEPPE

(1858–1932)

Giuseppe Peano, an Italian mathematician and logician, was a professor of mathematics at the University of Turin from 1890 to 1932 and also taught at the military academy in Turin from 1886 to 1901. In 1891 he founded the *Rivista di matematica,* which was later also published in French (*Revue de mathématique*) and in Interlingua (an international language developed from Latino sine flexione, an auxiliary language based on Latin), which Peano propounded in 1903. In 1898 Peano acquired a small printing establishment in Turin, and he soon became an accomplished printer; his skill seems to have been of help to him in the process of simplifying logico-mathematical symbolism.

Peano's contributions to mathematics include the first statement of vector calculus (*Elementi di calcolo geometrico,* Turin, 1891) and the first example of integration by successive approximations within the theory of ordinary differential equations; with the single hypothesis that the data were continuous he proved the existence of the integrals of such equations. He submitted to rigorous criticism the foundations of arithmetic, of projective geometry, and of the general theory of sets. Peano's postulates (1899) were a set of five postulates for the arithmetic of natural numbers that allowed arithmetic to be constructed as a hypothetical-deductive system. In 1882 Peano first arrived at the principle that rigorous language can be separated from ordinary language both within and without mathematics. As Bertrand Russell wrote, Peano's method "extended the region of mathematical precision

backwards towards regions which had been given over to philosophical disagreement" ("My Mental Development," in *The Philosophy of Bertrand Russell,* Paul A. Schilpp, ed., Evanston, IL, 1951, p. 11).

In 1890 Peano introduced the use of iota and inverted iota to distinguish a one-member class from its member, which permitted him to overcome previous confusion between ∈ ("being a member of"), ⊃ ("contained in"), and = ("equal to"). In general, Peano showed the importance of distinguishing the properties of a class from those of the individuals of that class, a need shown, for example, by his "sophism" (actually, a paralogism): "Peter and Paul are apostles; the apostles are twelve; therefore Peter and Paul are twelve."

Peano's work in mathematical logic is to be found in his "Formulario completo," which includes, among other items, the well-known *Formulaire de mathématiques,* a compendium of mathematics derived from a set of postulates by means of a new notation. The "Formulario," in its encyclopedic, high-level approach, anticipated the thorough expositions of Bourbakism. In using a notation at least as rigorous as those of C. S. Peirce and Gottlob Frege, and more comprehensive and expedient than theirs, Peano's work marked a transition from the old algebra of logic to contemporary methods. His notation is still partially in use, mainly through its adoption by Russell and A. N. Whitehead in *Principia Mathematica.*

After 1913 Peano ceased to follow developments in symbolic logic. He regarded as artificial Russell's interpretation of numbers as classes of classes. Peano made several hints concerning the need for analyzing the relation of formal language to ordinary language, but he was not himself interested in undertaking such analysis. A philosophical interpretation of some of Peano's techniques is to be found in the work of his pupil Giovanni Vailati, who pointed out the general importance of Peano's discoveries concerning recursiveness, implicit definitions, and the theory of postulates. The "Formulario completo," however, still offers suggestions for research.

See also Computability Theory; Frege, Gottlob; Logic, History of: Modern Logic; Mathematics, Foundations of; Peirce, Charles Sanders; Russell, Bertrand Arthur William; Vailati, Giovanni; Whitehead, Alfred North.

Bibliography

WORKS BY PEANO

Opere scelte. 3 vols, edited by Ugo Cassina. Rome: Edizioni Cremonese, 1957–1959. Includes much of what is referred to as the "Formulario completo": The 5 vols. below, plus minor publications of 1888–1913.

Formulaire de mathématiques. Vol. I, Turin: Bocca Frères, Ch. Clausen, 1894; Vol. II, in 3 parts, Turin, 1897–1899; Vol. III, Paris, 1901.

Formulaire mathématique. Vol. IV. Turin, 1903.

Formulario mathematico. Vol. V. Turin, 1905–1908. In Latino sine flexione.

WORKS ON PEANO

Cassina, Ugo. "L'oeuvre philosophique de Giuseppe Peano." *Revue de métaphysique et de morale* 40 (4) (1933): 481–491.

Cassina, Ugo. "Vita et opera de Giuseppe Peano." *Schola et Vita* 7 (3) (1932): 117–148. In Interlingua.

Couturat, Louis. "La logique mathématique de M. Peano." *Revue de métaphysique et de morale* 7 (4) (1899): 616–646.

Jourdain, P. E. B. "Giuseppe Peano." *Quarterly Journal of Pure and Applied Mathematics* 43 (1912): 270–314.

Russell, Bertrand. *The Principles of Mathematics.* Cambridge, U.K.: Cambridge University Press, 1903.

Stamm, E. "Józef Peano." *Wiadomości Matematyczne* 36 (1933): 1–56.

Terracini, Alessandro, ed. *In memoria di Giuseppe Peano.* Cuneo, 1955. Essays by various authors.

Ferruccio Rossi-Landi (1967)

PEARSON, KARL

(1857–1936)

Karl Pearson, a British scientist and philosopher of science, was born in London. He studied mathematics at King's College, Cambridge, where he became acquainted with James Clerk Maxwell, Sir George Stokes, and Isaac Todhunter and developed an interest in history, religion, and philosophy. He became a fellow of his college in 1880 and also studied law at Heidelberg and Berlin. Although he was called to the bar in 1881, he never practiced law. In 1884, at the age of twenty-seven, he was appointed to the chair of applied mathematics and mechanics at University College, London, a post that he held until 1911. For part of this time he also held a lectureship in geometry at Gresham College, London, where he developed his ideas in the philosophy of science for a popular audience. Through his friend Francis Galton he became interested in statistical problems in the biological sciences, helped to lay the foundations of modern statistical theory and biometry, and, in 1901, with Galton and Weldon, founded the journal *Biometrika.* In 1896 he was elected a fellow of the Royal Society and in 1911 he was appointed to the new chair of eugenics at University College. Pearson was an enthusiastic socialist and humanist. He retired in 1933 and died three years later.

Pearson published many scientific papers, as well as essays on most of the subjects in which he was interested. His philosophical work is contained mainly in *The Grammar of Science* (1892) and *The Ethic of Freethought* (1888), a collection of essays and lectures. He is usually regarded as an important early figure in modern positivism, but his contribution in this field has perhaps been overrated. Much of his work derives from that of Ernst Mach.

He accepted and developed Mach's sensationalist, antimetaphysical standpoint, but he was not afraid to talk with approval of "a sound idealism" replacing "the crude materialism" of earlier physics. His concern was to emphasize the social background of science and to urge that good citizenship demanded the application of the scientific habit of mind to everyday living. He appears to have regarded this as a large part of the justification of scientific activity, but he also held that science "justifies itself in its methods." Like Mach he dwelt on "the unity of science," which depends upon its method rather than upon its material. This method, based as it is upon verification, rules out metaphysics. The metaphysician is a poet, who does no harm so long as he is recognized as such, but he is often taken to be something more. According to Pearson, an acceptable moral theory is more likely to develop from the experiments of the biologist than from the speculations of the philosopher.

He saw scientific laws as brief formulas representing complex relationships between many phenomena. Their "discovery" is the work of a creative but disciplined imagination; they are products of the human mind. Following Lloyd Morgan, he said that an external object is a construct; that is, "a combination of immediate with past or stored sense-impressions." He asserted, mysteriously and unsatisfactorily, that the distinction between real objects and imaginary ones is that only the real objects depend upon immediate sense impressions.

A fundamental distinction in his work is that between perception, the "physical association" of stored sense impressions, and conception, their "mental association." This appears to mean that perception is merely the copresentation of impressions, while conception is the "recognition" of relations. But the physical and the psychical differ only in degree, not in kind, because both physics and psychology deal with relations between sense impressions, although from different standpoints. On the whole, human brains work in the same way, and thus one receives the same sense impressions and forms the same constructs as another. This ensures the universal validity of science. The field of study of the various sciences is, in fact, immediate sense impressions; these are the phenomena that scientific laws relate, so that "the field of science is much more consciousness than an external world." The consciousness of others is established by an argument from analogy.

We tend to project our sense impressions and to regard them as existing externally to and independently of ourselves, but this is a mistake. The distinction between external and internal is arbitrary and no more than a practical convenience. It is based on distinguishing between classes of sense impressions, not between sense impressions and something else. We cannot assert the existence of causes of sense impressions, but Pearson wanted to leave open the possibility of such existents. He therefore used the term *sensation* in an unusual way: Sensation is "that of which the only knowable side is sense-impression." This is intended to express agnosticism about the causation of sense impressions while allowing him to say, "The outer world is for science a world of sensations, and sensation is known to us only as sense-impression."

Some scientific concepts are not of immediate sense impressions; for instance, atom and molecule. There are just two possibilities: Scientists may regard the atom as real and thus capable of being a direct sense impression, or as ideal and thus merely a "mental conception assisting them in formulating laws." In contrast, a metaphysical conception is of what is both real and independent of sense perception.

Pearson concluded that science is not explanatory but merely descriptive. For instance, Isaac Newton's law of gravitation is a description in the simplest possible terms of a wide range of phenomena; that is, of the "routine" of our perceptions. To talk of it as ruling nature is to confuse other senses of "law" with the scientific sense. Causal statements are records of regular sequences in past experience and cannot assert any necessity in them. Using Humean arguments, Pearson held that forces, because they are not discoverable in sense experience, cannot be regarded as causes. "Force" is but a name hiding our ignorance of the explanation of motion. The idea of necessity is appropriate only to relations between conceptions, not to relations between perceptions. Prediction and knowledge are possible only because we find repetition in our sense impressions. Even so, our knowledge is only probable and should, strictly speaking, be called "belief."

The whole of science involves the distinction between the perceptual and the conceptual. Scientific concepts generally are ideal limits of concepts originating in perception. This is especially obvious in the mathematical treatment of the world. Empirical space and time

are "modes of perception." Space is "a mental expression for the fact that the perceptive faculty has separated coexisting sense impressions into groups of associated impressions"; time indicates "the progression of perceptions at a position in space." Neither space nor time is infinite or infinitely divisible, since each must be limited by our powers of perception and discrimination. Conceptual space and time, and the space and time of mathematics, are idealizations of their empirical counterparts and do not suffer from their limitations.

The aim of science is to construct conceptual models of the universe, devices to assist us in describing the correlation and sequence of phenomena. The failure to recognize this has led scientists to accept definitions of force, mass, atom, and—in the biological sciences—life that are riddled with metaphysical obscurities. Much of Pearson's philosophical writing consists in the empiricist elucidation of these fundamental concepts, in an attempt to remove these obscurities.

See also Belief; Mach, Ernst; Maxwell, James Clerk; Morgan, C. Lloyd; Newton, Isaac; Philosophy of Science, History of; Positivism; Scientific Method; Space; Time.

Bibliography

Pearson's main philosophical work is *The Grammar of Science* (London, 1892). The second edition (1900) contained two new chapters. The third (1911) contained only the first eight chapters (physical sciences) of the first two editions but had a new chapter on causation and a new final chapter on modern physical ideas, written largely by E. Cunningham. The Everyman edition (London, 1937) contains a more detailed account of the various editions.

Other works by Pearson are *The Ethic of Freethought, a Selection of Essays and Lectures* (London: Unwin, 1888; and London: A. and C. Black, 1901); *The Chances of Death and Other Studies in Evolution* (London: Arnold, 1897), a volume of essays and lectures; *National Life from the Standpoint of Science* (London: A. and C. Black, 1901); and *The Life, Letters and Labours of Francis Galton,* 3 vols. (Cambridge, U.K.: Cambridge University Press, 1914–1930).

Pearson edited and completed Isaac Todhunter, *A History of the Theory of Elasticity and of the Strength of Materials from Galilei to the Present Time,* 2 vols. Cambridge, U.K.: Cambridge University Press, 1886–1893), and W. K. Clifford, *Common Sense of the Exact Sciences* (New York: Appleton, 1885), for which he wrote the chapter "Position" and much of "Quantity" and "Motion."

Works on Pearson include V. I. Lenin, *Materialism and Empirio-Criticism,* translated by A. Finchberg (Moscow, 1937); G. M. Morant, *A Bibliography of the Statistical and Other Writings of Karl Pearson* (London: Biometrika Office, University College, 1939); E. S. Pearson, "Karl Pearson, an Appreciation of Some Aspects of His Life and Work," in *Biometrika* 27 (1936): 193–257, and 29 (1937): 161–248; C. S. Peirce, *Collected Papers* (Cambridge, MA: Harvard University Press, 1931–1958), passim, but especially Vol. VIII, which contains a long review of *The Grammar of Science*; and G. U. Yule and L. N. G. Filon, "Karl Pearson," in *Obituary Notices of Fellows of the Royal Society,* Vol. II (London, 1936–1938), pp. 73–110.

For reviews of *The Grammar of Science,* see those by C. G. K. (probably C. G. Knott) in *Nature* 46 (1892): 97–99, with replies by Pearson on pp. 199 and 247; by F. A. D. (of 2nd ed.) in *Nature* 62 (1900): 49–50; by E. A. Singer Jr. (of 2nd ed.) in *Philosophical Review* 9 (1900): 448–450; and an unsigned review in *Mind,* n.s., 1 (1892): 429–430.

There are numerous casual references to Pearson's views in books on the philosophy of science but few detailed discussions.

Peter Alexander (1967)

PECKHAM, JOHN

(c. 1225–1292)

John Peckham, or Pecham, the English philosopher and theologian, and defender of Augustinian doctrines, was born in Patcham, near Brighton, Sussex. Educated at the monastery at Lewes, he continued his studies at Oxford and Paris, and sometime during the 1250s he joined the Franciscan friars at Oxford. Subsequently he became a master of theology in Paris in 1269 and returned to Oxford in 1272. Peckham was provincial of the English Franciscans from 1275 to 1277 and then lectured at the papal court for two years. In 1279 he was appointed archbishop of Canterbury and held this office until his death.

Peckham's philosophical career represents a concentrated effort to counteract the growing allegiance to Aristotle through a return to the thought of Augustine. There seems little doubt that he was motivated to take this stand by the Lenten sermons of St. Bonaventure, who in the late 1260s had alerted his friars to the growth of heterodox Aristotelianism—which was apparent, for example, in the work of Siger of Brabant. Peckham did not reject all philosophy that stemmed from Greek and Arabic sources—as a matter of fact, he systematically used Aristotelian terminology—but his approach was a highly selective use of non-Christian philosophers to the extent that their works could be made to harmonize with the thought of Augustine. Among the disciples of Peckham who perpetuated this attitude were Matthew of Acquasparta, Roger Marston, and, later, Vital du Four.

Peckham's theory of knowledge shows the persistence of a special type of apriorism in the Franciscan school of this period. Clues to this apriorism are to be

found in the *Summa* of Alexander of Hales, which taught that the human intellect is incapable of a satisfactory a posteriori analysis of the first principles or of the most basic "perceptibles," such as time and space. Similarly, Augustine said: "If we both see that which you say to be true, and both see that which I say to be true, where, I ask you, do we see it? Neither I in you, nor you in me, but both in the unchangeable Truth itself, which is above our minds" (*Confessions* XII.24). Peckham concludes that more is required for the operation of the intellect than mere sensation that "contacts" accidents but does not reach the essence of things.

Even granting the intellect's power of abstracting essences, Peckham says that the mind does this either knowingly or unknowingly. If knowingly, then the mind knows before abstracting, and hence it is useless to abstract. If unknowingly, then the mind is at the mercy of chance and can hardly be called an intellect at all. Consequently, the intellect is not a passive Aristotelian tabula rasa, but a beam moving outward and casting its light on things. However, this explanation is not sufficient because in matters of intellectual knowledge, certitude, and evidence, man must be assisted by a divine illumination—a divine active intellect—in addition to his own human active intellect. This assistance by divine illumination is not a direct vision of God or an infusion of ideas. Rather, it is an assistance over and above that given by God as the conserving cause of all that exists. Its purpose is to guarantee necessity and certitude (considered irrevocably unobtainable through sensation) for our knowledge.

In the realm of natural theology, there was one key axiom that pervaded Franciscan philosophical circles in Peckham's time—that creatures are entirely dependent upon the First Cause with regard both to the fact of existing and to their ability to act. From this it follows that whatever causal powers a creature may possess are ontologically delegated to it by the First Cause. The important corollary of this principle is that the First Cause can bypass the agency of the creature and intervene to produce the effect immediately. Peckham invokes this principle to some extent in the illumination theory of knowledge. He also uses it to defend the autonomous existibility of prime matter without any form against the contrary opinion of Thomas Aquinas.

Peckham also took rather strong exception to Thomas's opinion that no single thing ever has more than one form. All medieval philosophers were agreed that the First Cause was pure form and that prime matter was completely formless. Against Thomas, Peckham and his confreres held that in each thing there are many forms, or at least many grades of one form. The dispute soon fossilized into two schools—the Dominicans and the Franciscans—and as often as not their arguments generated more heat than light. In any case, Peckham held that in humanity there are several forms—vegetative, sensitive, and rational—in a gradated order that cooperates toward the good and unity of the being as a whole.

John Peckham's career represents a sincere effort to perpetuate and to update the doctrines of Augustine. He suffered much distress as archbishop of Canterbury when, as a stubborn defender of Augustine, he incurred the wrath of the equally stubborn Dominican defenders of Thomas.

Many of the points that were merely hinted at in Peckham's philosophy were taken up by his disciples and elaborated in full-length treatises. A final judgment of this English Franciscan must await the publication of many of his works that are still in manuscript.

See also Alexander of Hales; Aristotelianism; Aristotle; Augustine, St.; Augustinianism; Bonaventure, St.; Marston, Roger; Matthew of Acquasparta; Siger of Brabant; Thomas Aquinas, St.

Bibliography

WORKS BY PECKHAM

Tractatus Pauperis. Caps. 1–6, edited by A. Van den Wyngaert, Paris, 1925, 3–72; caps. 7–9, edited by F. Delorme, in *Studi Francescani*, series 3, 4 (1932): 47–62, 164–193; caps. 10, 16, edited by A. G. Little, in *British Society of Franciscan Studies* 2: 27–55, 63–87; caps. 11–14, edited by F. Delorme, in *Collectanea Franciscana* 14 (1944): 90–117; caps. 14, edited by F. Delorme, in *Fr. Richardi de Mediavilla Quaes. Disp. de Privileg. Martini Papae IV* 79–88, Quaracchi, 1925.

Tractatus de anima. Edited by G. Melani. Biblioteca di Studi Franciscani. Florence, 1948.

Perspectiva communis. Edited by D. Lindberg. Madison: University of Wisconsin Press, 1970.

Tractatus de perspectiva. Edited by D. Lindberg. St. Bonaventure, NY: Franciscan Institute, St. Bonaventure University, 1972.

De numeris misticis. Edited by B. Hughes. *Archivum Franciscanum Historicum* 78 (1985): 3–28, 333–383.

Quodlibeta Quatuor. Edited by G. Etzkorn and F. Delorme. Bibliotheca Franciscana Scholastica XXV. Grottaferrata, 1989.

Quaestiones Disputatae. Edited by G. Etzkorn, L. Oliger, H. Spettmann, I. Brady, and V. Potter. Bibliotheca Franciscana Scholastica XXVIII. Grottaferrata, 2002.

Tractatus de sphaera. Edited by B. MacLaren. PhD diss. Eastern Kentucky University.

WORKS ON PECKHAM

Boureau, A. *Théologie, science et censure au XIIIe siècle. Le cas de Jean Peckham*. Paris: Belles Lettres, 1999.

Callebaut, André. "Jean Peckham O.F.M. et l'augustinisme." *Archivum Franciscanum Historicum* 18 (1925): 441–472.

Crowley, Theodore. "John Peckham O.F.M., Archbishop of Canterbury, versus the New Aristotelianism." *Bulletin of the John Rylands Library* 33 (1951): 242–255.

Douie, Decima L. *Archbishop Pecham*. Oxford: Clarendon Press, 1952. Includes a good bibliography.

Ehrle, Franz. "J. Peckham über den Kampf des Augustinismus und Aristotelismus in der zweiten Hälfte des 13 Jahrhunderts." *Zeitschrift für katholosche Theologie* 13 (1889): 172–193.

Etzkorn, G. "John Pecham, O.F.M.: A Career of Controversy." In *Monks, Nuns and Friars in Mediaeval Society*, edited by E. King, J. Schaefer, and W. Wadley, 71–82. Sewanee, TN: University of the South, 1989.

Etzkorn, G. "Révision dans l'ordre des Quodlibets de Jean Pecham." *Bulletin de Philosophie Médiévale* 19 (1977): 65.

Spettmann, Hieronymus. "Die Psychologie des Johannes Pecham." In Vol. XX of *Beiträge zur Geschichte der Philosophie des Mittelaters*, 1–102. Münster, 1919.

Teetaert, A. "Peckham Jean." In *Dictionaire de Théologie Catholique*, XII/1, 1933, col. 100–140. One of the best overall articles on Peckham.

Girard J. Etzkorn (1967, 2005)

PEIRCE, CHARLES SANDERS
(1839–1914)

Charles Sanders Peirce, the American philosopher, physicist, and mathematician and the founder of pragmatism, was born in Cambridge, Massachusetts. His father, Benjamin Peirce, was the leading American mathematician of the time and Perkins professor of mathematics and astronomy at Harvard. Young Charles was born and bred a scientist, and from his earliest years he showed great promise in mathematics and the physical sciences. He attended Harvard, graduated in 1859, and subsequently studied at the Lawrence Scientific School, from which he received his degree in chemistry summa cum laude in 1863.

During the next fifteen years, Peirce simultaneously pursued several distinct careers. He worked as an astronomer at the Harvard Observatory, where he did pioneer work in photometric research. He also worked as a physicist for the U.S. Coast and Geodetic Survey, of which his father was superintendent, and achieved some distinction for his discovery of hitherto undetected errors in pendulum experiments used to determine the force of gravity. And he worked, more or less privately, at philosophy and logic, steadily publishing works on these subjects from 1866 on. By 1879 he had achieved sufficient stature in these last two fields to be appointed lecturer in logic at the newly organized Johns Hopkins University in Baltimore, Maryland. He remained at Johns Hopkins from 1879 until 1884, meanwhile continuing to work for the Coast and Geodetic Survey—a connection that he sustained until 1891. In 1887, after having inherited some money, he retired to Milford, Pennsylvania, where he lived in relative isolation until his death. Peirce was twice married—in 1862 to Harriet Melusina Fay, whom he divorced in 1883, and in 1883 to Juliette Froissy, who survived him. He had no children.

PHILOSOPHICAL ORIENTATION

Peirce was a systematic philosopher of great breadth, and his writings cover almost all fields of philosophy. His greatest contributions were in the field of logic, but he wrote extensively on epistemology, scientific method, semiotics, metaphysics, cosmology, ontology, and mathematics, and less extensively on ethics, aesthetics, history, phenomenology, and religion. Since Peirce's views underwent considerable change as he grew older, it is not possible to speak of his philosophy as a single system: Rather, he formulated several systems, each of which represents a different phase in his development. These different systems, however, deal with the same problems and embody the same fundamental concept of philosophy.

Peirce came to philosophy as a student of Immanuel Kant, from whom he had acquired the architectonic theory of philosophy. In brief, this theory holds that the domain of knowledge can be so characterized that general assertions can be proven true of all possible knowledge; the theory also holds that it is the dependence of all knowledge upon logic that makes such a characterization possible. Accordingly, the doctrine holds that it is possible to derive from logic the fundamental categories and principles that form the basis of all that can ever be known. In formulating this theory, Kant assumed that logic was a completed, unchanging science. But Peirce was one of that group of men, including George Boole, Augustus De Morgan, Gottlob Frege, and others, who revolutionized logic and prepared the way for A. N. Whitehead and Bertrand Russell's *Principia Mathematica.* Hence, for Peirce, logic was a growing, changing subject, and as it changed, so, according to the architectonic theory, Peirce's philosophy had to change with it. Thus the major shifts in Peirce's system are correlated with his major discoveries in logic and reflect the modifications that he thought those discoveries entailed. In the following expo-

sition, Peirce's work will therefore be dealt with chronologically, and each system will be treated in order.

THE FIRST SYSTEM, 1859–1861

Peirce's first system is a form of extreme post-Kantian idealism. The sources of this idealism are not known: Whether he evolved it himself or derived it from some other source, such as Emersonian transcendentalism, cannot now be determined. What is clear is that by 1857 he was seeking to combine the Transcendental Analytic with Platonic idealism.

CATEGORIES. From Kant's doctrine of the Transcendental Sciences, Peirce derived a threefold ontological classification of all there is into matter (the object of cosmology), mind (the object of psychology), and God (the object of theology). Peirce referred to these three categories as the It (the sense world), the Thou (the mental world), and the I (the abstract world), respectively; and it was from these pronouns that he subsequently derived the names Firstness, Secondness, and Thirdness, by which he usually called his categories.

Having divided all there is into these three categories, Peirce's problem was then to define the relations among them. Specifically, the problem of knowledge as it appears in the first system is how the ideas in the mind of God can be known by human minds. Peirce thought he had found the solution to this problem in the Kantian principle that all phenomena and all concepts—all that can be before the mind—are representations, for he understood this to imply that the ideas in the mind of God, which Peirce conceived as Platonic archetypes, are first given a material embodiment in the form of the objects of our experience and are then derived by us from those objects by abstraction. So Peirce took the Transcendental Analytic to be a description of this process: The synthesis in intuition is the synthesis of the divine idea (already present in an unconscious form within the soul) with "the matter of sensation" to form the empirical object which is also, by virtue of the divine idea, the transcendental object; and the concept is derived by abstraction from the object given in intuition. But when it came to explaining just how the Kantian categories served to effect so un-Kantian a synthesis as that demanded by his own semiotic idealism, Peirce found himself in grave difficulties, and after struggling with the problem for some time he was forced to conclude that the Kantian table of categories was simply inadequate.

TRANSITIONAL PERIOD: STUDY OF LOGIC

According to the architectonic principle, the inadequacy of the table of categories implies the inadequacy of Kant's logical classification of propositions. In 1862, therefore, Peirce began the serious study of logic, and he naturally turned to the Scholastics for instruction. Although he began his study in the belief that the fundamental problem was the classification of propositions, he soon learned from John Duns Scotus that the classification of arguments, or forms of inference, was more fundamental, since the significance of propositions depends upon the role they play in inference. He was therefore led to investigate the irreducible forms of inference, and so to study Kant's famous paper "The Mistaken Subtlety of the Four Syllogistic Figures," in which Kant argued that all inference is reducible to Barbara or to a combination of Barbara and immediate inference. In the "Memoranda concerning the Aristotelian Syllogism," which he published in 1866, Peirce showed that Kant's argument is invalid, for the syllogism by which the reduction of the second and third figures is made is itself in the figure from which the reduction is being made. Peirce therefore concluded that the first three figures are irreducible. Moreover, Peirce noted that if the first figure is defined as the deduction of a conclusion from a major and a minor premise, then the second figure can be described as the inference of the major from the minor and conclusion and the third figure as the inference of the minor from the major and conclusion. Accordingly, Peirce held that the first figure is purely deductive, the second figure inductive, and the third figure hypothetical.

For Peirce this discovery had great importance. His previous belief in the existence of synthetic a priori propositions had rested on the two doctrines, derived from Kant, that all thought involves inference and that all inference is in Barbara. Granting these doctrines, it is clear that the major premises must be innate in the mind. But with the discovery of the role of hypothesis and induction, all synthetic propositions can be regarded as inferred, and so the problem shifts to the process of synthetic inference and to scientific inquiry.

At about the same time that he discovered the irreducibility of the three figures, Peirce made another important discovery in logic—namely, that the copula can be interpreted as the sign relation. This view, which was probably derived from the scholastic theory of supposition, enabled him to regard all propositions as instances of a single fundamental relation, and the analysis was quickly extended to inferences also by treating the

conclusion as a sign that is determined by the premises to represent the same state of affairs that they themselves represented. Such a result was thoroughly in line with Peirce's early semiotic idealism, and it meant that the fundamental logical relation from which the categories must be derived is signhood.

THE SECOND SYSTEM, 1866–1870

In 1867 Peirce published a paper titled "On a New List of Categories," in which he attempted to solve the problem of relating his three ontological categories of mind, matter, and God.

THE SIGN RELATION. Starting from Kant's position that knowledge occurs only when the manifold is reduced to the unity of a proposition, Peirce asked what that unity consisted in. Since he conceived the proposition in subject-predicate form, this is equivalent to asking how the predicate is applied to the subject. On the basis of the reduction of the copula to signhood, Peirce argued that the predicate is applied to the subject by being made to stand for the same object for which the subject stands. Thus a proposition would be impossible without reference to some object. But how does the predicate come to stand for this object? Only, Peirce held, by being interpreted as standing for it by some interpreting representation, or mind, so that no proposition is possible unless such an interpretant also exists. And how does the mind make this interpretation? Only, Peirce held, by the sign's representing its object in some respect, that is, by referring to some attribute of the object. Hence, propositions would be impossible if there were no pure abstract attributes embodied in the object to form the basis of comparison among them. So his argument, in essence, was that all synthesis involves the sign relation, that the sign relation consists in a sign standing for something to someone in some respect, and therefore that unless there are things, minds, and abstractions, there is no knowledge. But since the pure abstract attribute is the Platonic Form in the mind of God, what Peirce was really arguing is that without his three ontological categories signhood would be impossible.

Aspects of reference. In the "New List," Peirce did not present his categories directly as ontological classes; rather, he began with the problem of unifying the manifold by joining the predicate to the subject through the sign relation and then analyzed signhood into the three aspects of reference: reference to abstraction, reference to an object, and reference to an interpretant. These three aspects are then made the basis for a systematic classification of signs according to the prominence given to each reference, and this mode of classification is applied to terms, propositions, and arguments. In the case of arguments, Peirce rederived the division into hypothesis, induction, and deduction, thus presenting the three forms of syllogistic as consequences of his analysis of signs.

Logic, however, is not the only science of signs; indeed, it is but one of three, each of which studies a particular aspect of the subject. The first is speculative grammar, which studies the relation of signs to the abstraction; the second is logic, which investigates the relation of signs to their objects; and the third is speculative rhetoric, which investigates the reference of signs to their interpretants. Peirce could therefore derive his three ontological categories by abstraction from the three references of signs, but he had to show further how we can know the objects referred to and whether or not they are real. For these purposes he needed a theory of cognition and a theory of reality.

COGNITION. Peirce stated his new theories of cognition and reality in three articles published in 1868 in the *Journal of Speculative Philosophy.* These papers simply develop the implications of the "New List." Since the reference of a sign to its object is established by its being predicated of another sign which already refers to that object, and since the predication exists only because there is an interpreting sign that so interprets it, it is clear that the series of signs is doubly infinite. Peirce accepted this conclusion and asserted that there is neither a first nor a last cognition. While this doctrine appears bizarre, it has a clear purpose. What Peirce was trying to avoid was the classic dilemma of the empiricist who, having tracked cognition back to an original impression of sense, finds himself completely unable to prove the accuracy of that first impression.

Peirce held that if we examine what actually occurs in cognition, we find the process to be something like the following. In the flood of sensory stimuli that pours in upon us, we detect certain relations that lead us to segregate some stimuli and to interpret these as having a common referent. We do not know what the first such stimulus having that referent may have been, and the question is meaningless, since it is only after many stimuli have occurred that we note their relations. As experience progresses and we acquire more relevant stimuli, we further conceptualize this referent, and in time we acquire a progressively more and more complete and precise idea of it. But our knowledge is never fully complete, so that

this process of learning and inquiry is endless. It is true that once we have a relatively detailed concept of the referent, we assume that the object antedated our experience of it and in fact caused that experience; epistemologically, however, it is the experience that comes first and the notion of the object that comes later. The object, then, is a hypothesis designed to give coherence to our experience, and this hypothesis is derived by hypothetical and inductive reasoning; hence, the process of cognition can be fully described by the three forms of inference. Moreover, it follows that the object must be as we conceive it, since it is only as we conceive it that it is postulated at all, and therefore there can be no such thing as an incognizable cause of cognition, for the postulate that an object exists is warranted only by the coherence it gives to experience. Accordingly, whatever is, is cognizable.

REALITY. The above theory of cognition leads at once to a theory of reality. The object is real, Peirce held, only if as the number of cognitions goes to infinity, the concept of the object tends to a limiting form. It follows, therefore, that although the object is not independent of being thought (since it is only as it is thought that it exists at all), it is nevertheless independent of the thought of any particular man and represents what would be agreed upon by an ideal community of investigators if inquiry were to go on forever.

Many empiricists would agree with Peirce that if the object is real, then if inquiry does go on forever, our hypotheses will converge to a final true description. But few would follow him in holding that the object is real because inquiry converges. What Peirce was attempting to do in this instance was to propound a doctrine that was at once phenomenalistic and realistic. To do this, he had to give a phenomenal definition of reality that would compromise neither the inexhaustibility of the real nor the particularity of the phenomenal, and the infinite series of cognitions seemed to do just that. But could Peirce prove that the infinite series is convergent? In 1868 he thought he could do this by means of an argument that purported to show that the concept of a universe in which induction and hypothesis would not lead to agreement was self-contradictory. When he subsequently discovered that this argument was fallacious, his theory of reality had to be substantially revised.

Universals. Peirce's theory that reality consists in the convergence of inquiry led to a further consequence. For it follows that the real object must be as we conceive it to be, and since, as the "New List" showed, the predicate of a judgment is always general, it further follows that universals are real. On this basis Peirce declared himself a scholastic realist of the moderate, or Scotist, school. The claim is misleading, for whereas the scholastic doctrine rests on the assertion that the universal in the mind and the individual out of the mind have a common nature, Peirce's argument rests on the fact that no cognition is wholly determinate—that is, that there is no true individual, and that therefore everything is to some degree general. Peirce's "realism" was thoroughly idealistic throughout.

THE THIRD SYSTEM, 1870–1884

By 1870 Peirce had propounded, in outline at least, an architectonic philosophy based upon the principles that all cognition involves the sign relation; that the sign relation involves three classes of referents; and that these referents are real and can be adequately known by scientific inquiry. But this theory depended upon logical doctrines that Peirce was forced to abandon when he discovered the logic of relations.

The logic of relations. The first work on the new logic had been done by Augustus De Morgan, but little progress was made with the subject until Peirce entered the field in 1870. It was in this area that Peirce made his greatest contributions to logic, and it is no exaggeration to say that it was he who created the modern logic of relations. Philosophically these new discoveries in logic had important consequences, for the logic of relations forced Peirce to abandon the subject-predicate theory of the proposition that underlies the "New List," and so required that he overhaul his basic position. Probably the most notable revisions directly attributable to the new logic are the doctrines of pragmatism and the doubt-belief theory of inquiry.

THE DOUBT-BELIEF THEORY OF INQUIRY. Peirce formulated the doubt-belief theory in 1873, but it was first published in a series of six papers in *Popular Science Monthly* in 1877 and 1878. These papers do not constitute a rejection of the earlier theory of cognition; rather, they elaborate the earlier theory and set it in the context of biological evolution.

Any organism that is to survive, Peirce held, must develop habits of behavior that are adequate to satisfy its needs. Such habits are rules of behavior that prescribe how we should act under given conditions in order to achieve a particular experiential result. Now such habits, when thoroughly adopted, Peirce called beliefs. Since to possess beliefs is to know how to satisfy one's wants, belief is a pleasant state: Doubt, or the absence of belief, is an

unpleasant state, since one is then uncertain how to act and is unable to attain the desired goals. The organism will therefore seek to escape from doubt and to find belief. The process by which the organism goes from doubt to belief Peirce defined as inquiry. Clearly, there are various methods of inquiry, and the most satisfactory method will be that which leads most surely to the establishment of stable belief—that is, to beliefs that will stand in the long run.

PRAGMATISM. From the standpoint of the inquiring organism, a belief concerning a particular object is significant because it permits the organism to predict what experiences it will have if it acts toward the object in a given way. Recalling Kant's use of the term *pragmatic,* namely, "contingent belief, which yet forms the ground for the actual employment of means to certain actions, I entitle *pragmatic belief*" (*Critique of Pure Reason,* A 824, B 852), Peirce propounded what he called the pragmatic theory of meaning, which asserts that what the concept of an object means is simply the set of all habits involving that object. This doctrine involves a major change in Peirce's thinking, and one that is directly due to the logic of relations.

Prior to 1870, Peirce conceived the meaning of a term as the embodied abstraction that it connotes. The meaning of the concept of an object is therefore the same abstraction that is the essence of the object. But once relations were admitted as propositional constituents coordinate with quality, it became possible to conceive the object not only in terms of indwelling qualities but also in terms of relations among its states and with other objects—that is, in terms of its behavior. Accordingly, instead of regarding the behavior of the object as determined by its qualitative essence, the behavior itself may now be regarded as the essence. The meaning of the concept of an object may therefore be given by the set of laws completely specifying the behavior of the object under all conditions. These laws are conditional statements relating test conditions to phenomenal results, and such laws, considered as governing behavior, are habits relating action to experiential effects. Hence, the principle of pragmatism asserts that the concept of the object is synonymous with the set of all such conditionals. Since actual synonymy is asserted, it follows that the concept of a real object can be completely translated into phenomenal terms, but only, it should be noted, into dispositionally phenomenal terms—a point that was to cause Peirce considerable trouble.

Pragmatism: A theory of meaning. Pragmatism is Peirce's most famous philosophical doctrine, although it was made famous by William James rather than by Peirce. As Peirce defined it, pragmatism is purely a theory of meaning—not of truth. Moreover, it is a theory of meaning that combines two rather distinct emphases. First, Peirce intended pragmatism to be a principle of scientific definition. By permitting the translation of a concept into phenomenal results that are observable under stated test conditions, the principle legitimizes the use of theoretical constructs in science and thus does much to clarify the nature and status of scientific theory and proof. But when Peirce chose to call the doctrine pragmatism and insisted that the concept must be translatable into "practical effects," the choice of Kantian terminology was not accidental. Peirce was also stressing the utilitarian aspect of science and of all knowledge—that is, the fact that significance lies in the relation to ends desired. Peirce drew no distinction between these two aspects of pragmatism: For him they formed a single doctrine.

Scientific method. Taken together, pragmatism and the doubt-belief theory imply that the stable beliefs sought by inquiry are in fact the laws of science. The problem of finding the best method of inquiry therefore becomes that of the justification of scientific method, which in Peirce's terms means the justification of induction and hypothesis. Although Peirce formally presented this justification in terms of the operating characteristics of the procedures, he admitted that the relative frequency with which inductive and hypothetical inferences lead to the truth cannot be calculated; hence, our assurance that synthetic inference does ultimately lead to truth comes from the fact that inquiry will converge to a limiting result that is true by definition. Thus, in this instance Peirce admitted that the convergence of inquiry to a final opinion cannot be proven but must be assumed, and since his definition of reality rests upon the convergence of inquiry, this is equivalent to saying that the existence of the real is improvable and must be assumed. But even as an assumption the doctrine presents problems, for it amounts to saying that if inquiry were to go on forever it would converge, and thus involves fundamental questions concerning counterfactuals.

Counterfactuals. The problem of counterfactuals is central to Peirce's philosophy, and his failure to solve it was one of the chief reasons that his system of the 1870s had to be rejected. Pragmatism requires that the concept of a real object be wholly translatable into a set of conditionals relating test conditions to observations. But then it would seem that the concept of the real object is devoid

of content: That is, if the concept of the real object is synonymous with the set of conditionals, each of which is purely phenomenal, then the assertion of reality adds nothing to which a nominalist might object. Peirce, however, did not regard the concept of reality as vacuous; he argued that the conditionals are asserted to be true always, whether actually under test or not. The real, therefore, is a permanent possibility of sensation—not merely a series of sensations. But this leads directly to the counterfactual problem, or the equivalent problem of real possibility. Peirce's theory requires that there be real possible sensations—an assertion that is not only unprovable but pragmatically meaningless, since possible sensations are pragmatically equivalent to actual sensations. Thus, far from proving phenomenalism realistic, Peirce found his position reduced to a subjectivism that was the exact antithesis of the scholastic realism he had hoped to establish.

THE FOURTH SYSTEM, 1885–1914

During the years he spent at Johns Hopkins, Peirce was extremely productive in the field of logic. He further developed and extended the calculus of relations and applied it to problems in mathematics. He also clarified and revised his theory of synthetic inference, began the study of the Cantor set theory, and in 1885, with the help of his student, O. H. Mitchell, discovered quantification—a discovery in which Frege had anticipated him by six years. These new developments in logic, together with the rather serious difficulties in his own philosophical position that had become apparent by the end of the 1870s, led Peirce to attempt a radical reformulation of his position in 1885. This reformulation involved a complete revision of the categories, the theory of cognition, and the theory of reality.

THE CATEGORIES. In the 1885 version of the categories, Peirce distinguished sharply between their formal and material aspects. Formally considered, the categories (Firstness, Secondness, and Thirdness) are simply three classes of relations—monadic, dyadic, and triadic. Moreover, Peirce held that these classes are irreducible and that all higher relations (quartic, quintic, etc.) are reducible to some combination of these three. The irreducibility of monadic and dyadic relations is generally admitted. The irreducibility of triadic relations is argued on the ground that all combinatorial relations are triadic, since they involve a relation between two elements and a resulting whole. Granting this, it follows that triadic relations are irreducible, because analysis could only resolve them into components and a combinatorial relation, and that combinatorial relation would itself be triadic. But once the notions of element and combination are given, relations of more than three correlates are easily generated, and so all higher relations may be regarded as being constructed from the three basic types.

Among triadic relations Peirce distinguished pure and degenerate species. A pure triadic relation is one in which no two of the correlates would be related without the third. His example of such a relation is signhood, for the sign relates object and interpretant, the interpretant relates sign and object, and the object, by establishing the identity of the extensional domain, relates sign and interpretant. Since Peirce held that all thought is in the form of signs, it follows that all thought is irreducibly triadic, which is another way of stating the Kantian doctrine that all thought is synthetic.

Since a monadic relation is a one-place predicate, the material aspect of Firstness must be qualitative, and Peirce therefore called it quality; what he meant by this term in 1885, however, was not the embodied abstraction that he had described in 1867. Quality now refers not to a concept but to a phenomenal suchness that is the immediate, nonconceptual given of sensation. In the 1885 version, not the concept red, but that suchness of an object that leads us to classify it as red, is a quality.

Peirce called the material aspect of Secondness *haecceity,* a term derived from Duns Scotus's *haecceitas,* meaning "thisness." As experienced, haecceity is known as shock or brute resistance: Peirce described it as an immediately given, nonconceptual experience of dyadic opposition or "upagainstness." The fact that the experience implies the dynamic interaction of two things, and is therefore dyadic in structure, permits it to qualify as the material aspect of Secondness. For Duns Scotus, haecceity was the principle of individuation, and Peirce accepted this meaning: Only individual things have haecceity. It was apparently the discovery of quantification theory that led Peirce to this formulation, for in the variable of quantification theory he found a sign capable of referring directly to an object without describing it, and "thisness" was intended as that property of the object by virtue of which such a reference can be made.

The material aspect of Thirdness is less clearly defined than that of the other two categories. Peirce described it as combination, or mediation, where the latter term signifies either connection or means-ends relations among things. Signhood may also be regarded as part of the material aspect of Thirdness, and so too may generality, since the general constitutes a connection among particulars. Clearly, what Peirce was describing in

this instance has much less the character of the immediately given than is the case for the other two categories. The reason is that Peirce not only regarded all thought as triadic—he also regarded all pure triads as conceptual. The material aspect of Thirdness is therefore the experience of thought or rationality. One of Peirce's problems was to explain just how so immaterial a thing can be perceived.

COGNITION. The revision of the categories raised some important problems in regard to cognition. Not only did Peirce have the problem of demonstrating how Thirdness can be perceived, but he also had the problem of explaining how quality and haecceity could be perceived. For in his earlier writings on cognition, Peirce had explicitly denied the existence of first impressions of sense of precisely the sort that he now introduced as the material aspects of his first two categories. Moreover, a further set of problems relating to cognition arose from the doubt-belief theory itself. For in that theory, logic, both deductive and synthetic, is treated as a method whereby an inquiring organism seeks belief. The status of logic, therefore, is that of a useful but contingent means to a sought end—contingent both upon our seeking this particular end, which is a characteristic of the present evolutionary state, and upon our choosing the most efficient of the several available means. Thus, in the doubt-belief theory, logic loses that necessary relation to all possible knowledge that is asserted by the architectonic theory and required to prove the universality of the categories.

Classification of knowledge. Throughout the 1890s Peirce labored at the problem of reconstructing the architectonic theory. Since the architectonic theory presupposes a classification of knowledge into two classes—logic, and all other knowledge—Peirce's problem was to develop this classification so as to ensure the universality of the categories, while at the same time not contradicting his theory of inquiry. The final system of classification was not attained until 1902. In that system, Peirce divided knowledge into practical (or applied) and theoretical sciences, and then further subdivided the theoretical sciences into sciences of discovery and sciences of review (the latter merely summarizing the findings of the sciences of discovery). The major portion of the classification thus deals with the sciences of discovery. The classification is by presupposition.

The first science is mathematics, which Peirce regarded as presupposed by all others. Mathematics is divided into three branches: mathematics of logic, mathematics of discrete series, and mathematics of continua. It is to the mathematics of logic that Peirce assigned the threefold classification of relations that constitutes the formal aspect of the categories. Next after (and presupposing) mathematics comes philosophy, which Peirce divided into phenomenology, normative science, and metaphysics. Phenomenology, which here appeared in Peirce's writing for the first time, is defined as the study of all that can be before the mind, but in practice, it is devoted to proving that all phenomenal experience is resolvable into three factors, which are the material aspects of the three categories. Thus Peirce sought to show that his categories, in both their formal and material aspects, are presupposed by all other knowledge.

Normative science has three divisions: aesthetics, ethics, and logic. In this classification logic appears explicitly as the science of how we ought to reason in order to obtain our objectives—whatever they may be. Thus the contingent and utilitarian aspect of logic, first brought out by the doubt-belief theory, is here made central. But reasoning as we ought is only one aspect of acting as we ought, which is the proper subject of ethics: Hence, logic presupposes the science of ethics, or the science of how conduct should be regulated to attain our ends. But what our conduct ought to be depends on our aims, and these Peirce held to be the subject of aesthetics, which is the science of what is desirable in and of itself. Hence Peirce subscribed to an aesthetic theory of goodness and made the good and the beautiful coincide.

Following and presupposing philosophy is idioscopy, which Peirce subdivided into the physical and psychical sciences. Each division is further subdivided to yield what we would ordinarily regard as the physical, biological, and social sciences. All domains of science thus fall within the classification, and so depend upon the categories. The classification thus serves the purpose of preserving the architectonic while ensuring the normative role of logic.

Perception. Peirce's determination to preserve both the universality and phenomenal observability of the categories as well as the normative character of logic is evident in the theory of percepts and perceptual judgments that he propounded at this time. According to Peirce, physiology and psychology tell us that our percepts are synthesized from the myriad neural stimuli that assail us from without. Of these neural stimuli themselves and of the process of synthesis we are entirely unaware; the earliest step in cognition of which we are at all conscious is the percept. But we cannot really be said to know the percept; what we know is a perceptual judgment, which is a proposition telling us what the nonlinguistic percept was. The perceptual judgment, such as "red patch here now," is

a hypothesis that explains the percept, but it is a peculiar hypothesis, since it is immediate and indubitable. Even if the perceptual judgment is immediately followed by a contradictory perceptual judgment, still that second perceptual judgment relates to a later percept, and it remains indubitable that my first and now forever vanished percept was truly red. Perceptual judgments, therefore, form the real starting point in knowledge and must be taken as the ultimate evidence statements.

Peirce described the processes of synthesis that precede and lead to the perceptual judgment as unconscious inference. Their inferential character is defended, here as in his earlier writings, by an argument that identifies the psychological processes of association with the forms of inferences. But since these processes are unconscious, they are beyond our control and thus are not subject to logical criticism—for logical criticism, being normative, is applicable only to voluntary and controllable behavior. On the other hand, conscious inferences, such as the processes whereby we derive knowledge from the perceptual judgments, are thoroughly subject to logical criticism. Accordingly, Peirce could hold both that there is no first impression of sense and that the object (percept) is given to us by a synthesis in intuition. He could further hold that our knowledge has a definite starting point in propositions that give direct reports of phenomenal observation and that whatever is asserted in those judgments of perception must be accepted as given. Thus, in the theory of percepts and perceptual judgments, Peirce tried to reconcile his denial of first impressions with his doctrine of direct phenomenal contact with the world.

On the basis of this theory, Peirce held that the material aspects of all three categories are empirically observable. Quality and haecceity are argued to be directly observable aspects of the percept. But so, too, according to Peirce, is Thirdness, for what is asserted in the perceptual judgment is necessarily true, and the perceptual judgment, being a proposition, has a predicate that is general. Since the generality is given in the perceptual judgment, and since criticism cannot go behind the perceptual judgment, this generality must be regarded as given in perception, and hence as being observable. Thus, by phenomenological analysis, all the categories can be shown to be present in experience.

REALITY. In the course of his study of the logic of relations, Peirce noted that the analysis of certain relations leads to an infinite regress. Thus the relation "in the relation *R* to" must itself be related to its subjects by the same relation, for example, "in the relation 'in the relation *R* to' to," and so on. Such relations, which can be analyzed only into relations of the same sort, Peirce called continuous relations, since they fit the definition of the continuum as that of which every part is of the same nature as the whole. They are, according to Peirce's theory, pure triadic relations; therefore their irreducibility follows from the irreducibility of Thirdness. Moreover, since every relation must be related to its subjects by some such relation, Peirce drew the conclusion that all relations involve a continuous relation.

Continua. During the 1880s, Peirce had become acquainted with Georg Cantor's work on set theory, which bears directly on the problem of continuity. Recognizing at once the great importance of Cantor's work for both logic and mathematics, Peirce undertook the study of the foundations of mathematics and attempted to construct his own theory of cardinal and ordinal numbers. Peirce's papers on this subject are highly technical, and only the briefest summary of them can be given here. In developing his theory of cardinal numbers, Peirce discovered a form of the paradox of the greatest cardinal. His efforts to solve this paradox led him to the erroneous conclusion that the series of transfinite cardinals is only countably infinite and has an upper limit that is the power of the linear continuum. It follows that if the continuum consisted of discrete elements, then there would exist a greatest cardinal, and to avoid this conclusion he held the continuum to be a "potential" set consisting of possible points. Accordingly, although subsets of any multitude may be actualized from the continuum, nevertheless, not all of the possible points are actualizable, since if they were, we should have a greatest cardinal and hence a contradiction. Peirce believed that by such arguments he had established that whatever is truly continuous involves unactualized possibility; hence the problem of the existence of real possibility, which he had found insoluble in the 1870s, was now reduced to that of the reality of continuity. Peirce used the arguments of Zeno in an attempt to prove that space and time must be truly continuous in his (Peirce's) sense, and he went on to argue that continuous relations are truly continuous both intensively and extensively. In defining the continuum as that of which every part is the same sort as the whole, Peirce was brought to the conclusion that real relations, and so real laws, are in some sense continua.

Synechism. The doctrine that the world contains real continua Peirce called synechism. He regarded this as his most important philosophical doctrine and preferred to have his whole philosophy called by this name. He also asserted that it was a modern form of scholastic realism.

Scholastic or not, it is certainly realistic, for it holds that the external referents of true laws are real continua which, since they involve unactualized possibilities, contain real generality. To support this doctrine, Peirce had to define an ontology that would explain what those referents might be. Peirce was no stranger to such an enterprise. He began his work in philosophy in the 1850s, with the doctrine of the three ontological categories, and although he subsequently redefined the categories several times in less ontological fashion, he never forgot the question of what realities lay behind his categories. It is therefore not surprising that following the 1885 revision of the categories, Peirce returned to the problem of ontology, and this soon led him to propound an evolutionary cosmology.

EVOLUTIONARY COSMOLOGY. Peirce had several reasons for formulating an evolutionary cosmology in the 1890s. Not only did synechism require a clarification of his ontological commitments, but he was also impelled toward such a formulation by problems arising within the theory of cognition. First, the doubt-belief theory, by imbedding inquiry within an evolutionary context, made the utility of scientific method relative to a particular evolutionary adaptation, the permanence of which is by no means guaranteed and must therefore be investigated.

A second reason for Peirce's formulating an evolutionary cosmology in the 1890s springs from his doctrine of critical common sense. Like all students of scientific method, Peirce was perplexed by the problem of how we discover true hypotheses. Considering the infinity of possible false hypotheses, it is evident that not even Peirce's theory of synthetic inference could account for the remarkable frequency with which we do, in fact, find a true explanatory hypothesis. Utilizing the evolutionary doctrines current at the time (including the inheritance of acquired characteristics), Peirce argued that the human mind must possess some innate adaptation that enables us to guess the correct laws of nature more readily than pure chance would allow. Such an adaptation would mean that true hypotheses appear to us peculiarly simple and natural. According to Peirce, it follows, then, that judgments of common sense, conceived through the mechanism of the inheritance of acquired characteristics as quasi-instinctual beliefs that have been built up through centuries of experience, should have a greater probability of being true than have parvenu doctrines. But this probability is at best low, so that commonsense judgments cannot be accepted without critical analysis and careful test. Thus Peirce's doctrine of common sense is thoroughly critical: Common sense is to be regarded as a likely source of true hypotheses, but no hypothesis is to be accepted without empirical validation. But in terms of the doubt-belief theory, this doctrine leads to a serious problem. Should the course of evolution alter significantly, our innate adaptation, which has proven so useful in the past, would become positively harmful, since it would direct us to seek explanations in terms of an adaptation that no longer obtains. Accordingly, it becomes a question of considerable moment to inquire what the future course of evolution will be.

The continuous external referent. In the doubt-belief theory, Peirce had formulated the principle that a law, which he conceived as governing the behavior of an organism, is a habit. Now a habit, considered as a psychological entity, is a connection among feeling states and actions, and this connection, Peirce held, must consist in an actual substantive continuity among them. Peirce based this assertion on a variety of arguments, including the felt continuity of mental phenomena (the impossibility of memory without continuous connection between past and present) and certain arguments drawn from the behavior of protoplasm under stimulation. It was therefore Peirce's doctrine that habit, considered as a psychological entity, is a continuum corresponding to a law that is conceived as governing behavior. To find continuous external referents for all laws, Peirce asserted that the universe is itself a living organism possessed of feelings and habits and that our laws of nature describe the habits of the universe. Thus, after 1885, the subjective idealism of Peirce's early writings became an extreme form of objective idealism.

Knowledge, feeling, volition. From the position that the universe is an organism, it follows that all our experience of the external world must be describable as experience of some state or behavior of this organism. But the possible forms of experience are defined by the material aspects of the categories, while Peirce took the possible components of mind to be defined by the traditional division into knowledge, volition, and feeling. He had already identified knowledge with belief-habit and made it the correspondent of law, or Thirdness. He now identified feeling as the correspondent of Firstness and volition as the correspondent of Secondness. But the doctrine asserts more than mere correspondence, for Peirce seeks to account for the fact that all our experience can be classified by the categories, and his explanation for this fact is that what is for the cosmic organism feeling, volition, and belief is experienced by the individual as Firstness, Secondness, and Thirdness.

Chaos and order. The habits created through inquiry are, objectively viewed, laws of behavior. What then,

according to Peirce, is doubt, or the absence of belief? In the state of doubt, there will be feeling, but no habit and no order—hence, objectively viewed, the state of doubt will appear as purely random or chance behavior. Thus, objective orderliness or randomness corresponds to states of the universe in which habit is either strong or weak. The irritation of doubt is redefined as an intense consciousness associated with states of unordered feeling; as order or habit increases, the intensity of consciousness declines until, in the case in which virtually complete regularity has been established, it is so low as to be all but undetectable. Mind that is so hidebound with habit we regard as dead matter.

When the doubt-belief theory is applied to the organic universe itself, the result is an evolutionary cosmology. In the beginning, Peirce held, there is nothing but an undifferentiated continuum of pure feeling wholly without order—a primal chaos. From this starting point, the universe evolves by means of the development of habits. We have here the typical Spencerian passage from homogeneity to heterogeneity, but without benefit of Herbert Spencer's mechanical model. In the course of time, the universe becomes ever more orderly—but at any given time its habits remain less than perfectly regular and there are still areas requiring the further fixation of belief.

This cosmology is the basis for Peirce's doctrine of tychism—that there is absolute chance in the universe. For as law is the objective manifestation of habit, so chance is the objective manifestation of lack of habit; hence the primal undifferentiated continuum of feeling is literally a world of pure chance. Evolution constantly diminishes the amount of objective chance in the universe, but only in the limit does it wholly disappear. At any given time, some chance remains, and the laws of nature are not yet wholly exact.

Pragmatism and universal evolution. The doubt-belief theory describes inquiry as an attempt to escape the irritation of doubt. But it is hardly proper to say that the universe seeks to escape from doubt, and some better motive is required. The state toward which the universe is evolving is, according to Peirce's theory, one of complete order. Since such a state involves the complete subjection of feeling and action to belief, Peirce regarded it as the realization of rationality in the concrete, or, in his terms, of "concrete reasonableness." But it is also a state of maximum beauty, for Peirce's aesthetic is a coherence theory of beauty. Accordingly, the normative theory of inquiry may be brought to bear in explaining the evolutionary process. The end sought is concrete reasonableness; the means, supplied by ethics, is the regulation of conduct by this aim. In the area of inquiry, this implies the discovery of those laws necessary to regulate behavior. Thus pragmatism, or pragmaticism, as Peirce renamed his doctrine after 1905 in order to distinguish it from James's, also serves the cause of evolution, for in translating the concept into a set of habits we discover the practical effects of the object—that is, how our conduct is affected. It remains for scientific inquiry, then, to discover the truth or falsity of potential habits and hence to fix belief. Thus the course of universal evolution and our modes of inquiry must remain ever in harmony, for the objective logic of evolution is identical with the logic of discovery. All nature works by a common process to a common end, and the duty of the individual man is to aid that process by devoting himself to scientific inquiry.

See also Boole, George; Cantor, Georg; Categories; Chance; Common Sense; Counterfactuals; De Morgan, Augustus; Duns Scotus, John; Frege, Gottlob; Idealism; Induction; James, William; Kant, Immanuel; Logic, History of; Mathematics, Foundations of; Pragmatism; Realism; Russell, Bertrand Arthur William; Scientific Method; Scotism; Universals, A Historical Survey; Whitehead, Alfred North.

Bibliography

WORKS BY PEIRCE

The Collected Papers of Charles Sanders Peirce, Vols. I–VI, edited by Charles Hartshorne and Paul Weiss, Cambridge, MA: Harvard University Press, 1931–1935; Vols. VII–VIII, edited by Arthur Burks, Cambridge, MA: Harvard University Press, 1958. This is the basic published collection of Peirce's writings. (The usual method of citation to these volumes is by volume number, followed by a decimal point and the paragraph number—for example, 3.456.)

Charles S. Peirce's Letters to Lady Welby. Edited by Irwin Leib. New Haven, CT: Whitlock's, 1953. These letters, written between 1903 and 1911, are largely devoted to the theory of signs and contain some of Peirce's best writings on that subject.

WORKS ON PEIRCE

Buchler, Justus. *Charles Peirce's Empiricism.* New York: Harcourt Brace, 1939. An incisive study of Peirce's more empirical doctrines, with particular emphasis on pragmatism and common-sensism.

Feibleman, James. *An Introduction to Peirce's Philosophy, Interpreted as a System.* New York: Harper, 1946. A broad but superficial survey.

Gallie, W. B. *Peirce and Pragmatism.* London: Penguin, 1952. A thoughtful book devoted chiefly to Peirce's pragmatism.

Goudge, Thomas A. *The Thought of C. S. Peirce.* Toronto: University of Toronto Press, 1950. Goudge holds that Peirce's work contains two contradictory positions, which he

calls naturalism and transcendentalism. The book is an exposition of this thesis and of its implications.

Lewis, Clarence I. *A Survey of Symbolic Logic.* Berkeley: University of California Press, 1918. Ch. 1, Sec. 7. This is still the best essay on Peirce's work in logic.

Moore, Edward C., and Richard S. Robin, eds. *Studies in the Philosophy of Charles Sanders Peirce, Second Series.* Amherst: University of Massachusetts Press, 1964.

Murphey, Murray G. *The Development of Peirce's Philosophy.* Cambridge, MA: Harvard University Press, 1961. An attempt to interpret Peirce's work chronologically and systematically through the architectonic principle.

Thompson, Manley. *The Pragmatic Philosophy of C. S. Peirce.* Chicago: University of Chicago Press, 1953. A thoughtful and systematic study of Peirce's pragmatism and related problems.

Weiss, Paul. "Charles Sanders Peirce." In *Dictionary of American Biography.* New York, 1934. Vol. XIV. A very fine biographical article on Peirce.

Wiener, Philip, and Frederic Harold Young, eds. *Studies in the Philosophy of Charles Sanders Peirce.* Cambridge, MA: Harvard University Press, 1952. This collection of essays on Peirce's philosophy is extremely uneven: it contains some excellent articles and some very poor ones. The papers by Savan, Thompson, Fisch and Cope, and Weiss are particularly good.

Murray G. Murphey (1967)

PEIRCE, CHARLES SANDERS [ADDENDUM]

Charles Sanders Peirce, one of America's most original philosophers, produced a body of work remarkable for its scope and enduring relevance. For many years Peirce's principal contributions to mainstream philosophy were in logic and philosophy of science, but changes in the philosophic terrain since 1967 have brought new areas of his thought to prominence. The resurgence of interest in pragmatism, due in large measure to its promotion by Richard Rorty, and the adoption of Peirce by the Frankfurt School as the philosopher who may hold the key to the problem of modernity, have brought attention to Peirce's unique brand of pragmatism and to his philosophy of signs. Outside of philosophy, the active interdisciplinary field of semiotics that began in Chicago with Charles Morris acknowledges Peirce as the founder of modern sign theory.

Peirce was a late child of the enlightenment, a staunch believer in the universal applicability of mathematics and in the continuous growth of knowledge through sustained inquiry. He was a diligent student of the history of science and understood that the advancement of knowledge is crucially linked to nondeductive (inductive and abductive) reasoning and shared experimental methods. He was convinced that a prerequisite for successful experimentation is an external world resistant to actions arising from misconceptions of it. These views led Peirce to an anti-Cartesian epistemology rooted in perceptual experience and committed to fallibilism and the repudiation of deductive foundationalism. Peirce generalized his view of the advancement of science to all forms of learning from experience, and he concluded that all meaningful conceptions are necessarily related to experiential expectations (conceived consequences). This is the epistemological motivation for his meaning-focused pragmatism (pragmaticism).

Sometimes Peirce is said to have equated truth with settled belief, but that applies only when belief is settled as the result of a steadfast application of scientific method. Other methods for overcoming doubt and settling belief, such as the a priori method or the methods of tenacity and authority, while not without some advantages, do not provide grounds for confidence that truth will be reached. Even the sustained application of scientific method can never issue in a guarantee that inquiry has "stormed the citadel of truth." Truth is always relative to propositions and is, therefore, grounded in the conventionality of symbolism (for propositions can only be expressed symbolically). The true represents the real precisely insofar as inquiry forces beliefs to yield to the dictates of an independent reality, but the "correspondence" of truth and reality that is hoped for at the end of inquiry is at best an ideal limit; we can never be certain that we have reached the truth. This is Peirce's fallibilism. It is typical of Peirce's philosophy that truth and reality are correlates in a triadic relation, where the mediating relate involves a community of inquirers (interpreters).

Peirce believed that the key to intelligence of any kind is sign action (which is always goal directed), and he formulated an elaborate semiotic theory to facilitate the analysis and classification of signs. Peirce's division of signs into icons, indexes, and symbols is his best-known semiotic bequest—although his distinction between tones, tokens, and types is also widely used—but these are only two of many triads that permeate his philosophy. Peirce held that minds are sign systems and thoughts are sign actions, and it is not too far-fetched to say that the mission of his semiotic is similar to that of modern-day cognitive science. Peirce's epistemological shift from a focus on ideas to signs marks him as a forerunner, if not a founder, of philosophy's so-called linguistic turn and, also, of the modern—and postmodern—emphasis on textualism. Peirce's triadic theory of signs distinguishes

semiotics from semiology, a generally dyadic theory of signs stemming from the work of Ferdinand de Saussure. Recently there have been attempts to reconcile these two approaches.

Current interest in Peirce's thought extends over most of philosophy. Peirce's graphical logic (his existential graphs) is used as a basis for computational linguistics. The recent move away from logicism has led to renewed interest in Peirce's philosophy of logic, according to which logic is not the epistemic foundation for mathematics. The rehabilitation of systematic and speculative thought has attracted attention to Peirce's evolutionary cosmology, which holds that the principal constituents of the universe are chance, law, and habit formation. Peirce insisted that change is really operative in nature (his tychism), that continuity, in general, prevails (his synechism), and that love or sympathy has a real influence on the course of events (his agapism). He contributed America's most original and thoroughgoing phenomenology (his phaneroscopy), and he advanced unique views on religion and on the significance of sentiment and instinct. He stressed the importance of the existent and the individual while, at the same time, admiring the ideal and insisting that rationality is rooted in the social. Peirce's intellectual legacy is a rich system of thought that helps organize and unify a broad array of issues in modern philosophy.

See also Chance; Classical Foundationalism; Cognitive Science; Enlightenment; Logic, History of; Philosophy of Science, History of; Pragmatism; Rorty, Richard; Truth.

Bibliography

WORKS BY PEIRCE

The New Elements of Mathematics by Charles S. Peirce, 4 vols. Edited by C. Eisele. The Hague: Mouton, 1976.

Complete Published Works Including Selected Secondary Material (microfiche edition). Edited by K. L. Ketner et al. Greenwich, CT, 1977. A companion bibliography is also available: *A Comprehensive Bibliography of the Published Works of Charles Sanders Peirce with a Bibliography of Secondary Studies.* Edited by K. L. Ketner. Greenwich, CT: Johnson Associates, 1977. Rev. ed., Bowling Green, OH, 1986.

Writings of Charles S. Peirce: A Chronological Edition. Edited by the Peirce Edition Project. Bloomington: Indiana University Press, 1982–.

Historical Perspectives on Peirce's Logic of Science: A History of Science, 2 vols. Edited by C. Eisele. New York: Mouton, 1985.

Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898. Edited by K. L. Ketner. Cambridge, MA: Harvard University Press, 1992.

The Essential Peirce: Selected Philosophical Writings, 2 vols. Edited by N. Houser and C. Kloesel. Bloomington: Indiana University Press, 1992–1998.

WORKS ON PEIRCE

Apel, K.-O. *Charles S. Peirce: From Pragmatism to Pragmaticism.* Translated by J. M. Krois. Amherst: University of Massachusetts Press, 1981.

Brent, J. *Charles Sanders Peirce: A Life.* Bloomington: Indiana University Press, 1993.

Burch, R. W. *A Peircean Reduction Thesis.* Lubbock: Texas Tech University Press, 1991.

Delaney, C. F. *Science, Knowledge, and Mind: A Study in the Philosophy of C. S. Peirce.* Notre Dame, IN: University of Notre Dame Press, 1993.

Eisele, C. *Studies in the Scientific and Mathematical Philosophy of Charles S. Peirce.* Edited by R. M. Martin. The Hague, 1979.

Esposito, J. L. *Evolutionary Metaphysics: The Development of Peirce's Theory of Categories.* Athens: Ohio University Press, 1980.

Fisch, M. H. *Peirce, Semeiotic, and Pragmatism.* Edited by K. L. Ketner and C. J. W. Kloesel. Bloomington: Indiana University Press, 1986.

Freeman, E., ed. *The Relevance of Charles Peirce.* La Salle, IL: Hegeler Institute, 1983.

Hausman, C. R. *Charles S. Peirce's Evolutionary Philosophy.* Cambridge, U.K.: Cambridge University Press, 1993.

Hookway, C. *Peirce.* London: Routledge and Kegan Paul, 1985.

Houser, N., D. D. Roberts, and J. Van Evra, eds. *Studies in the Logic of Charles S. Peirce.* Bloomington: Indiana University Press, 1996.

Ketner, K. L., ed. *Peirce and Contemporary Thought: Philosophical Inquiries.* New York: Fordham University Press, 1995.

Murphey, M. G. *The Development of Peirce's Philosophy.* Cambridge, MA: Harvard University Press, 1961.

Raposa, M. L. *Peirce's Philosophy of Religion.* Bloomington: Indiana University Press, 1989.

Roberts, D. D. *The Existential Graphs of Charles S. Peirce.* The Hague: Mouton, 1973.

Nathan Houser (1996)

PELAGIUS AND PELAGIANISM

Pelagius was a spiritual adviser to Christian aristocrats in Rome around the turn of the fifth century CE. In a commentary on the Pauline epistles, a treatise *On Nature*, and other writings, he sought to bolster Christian asceticism by opposing Manichaean determinism and affirming human capacity to progress toward moral perfection. His moral character and theological insights attracted followers who defended and developed his teachings.

Opposition to Pelagius and his followers began to intensify after Alaric's sack of Rome forced them to emigrate. In 411 one of Pelagius's protégées, Caelestius, sought ordination to the priesthood in Carthage and instead was condemned for his views on the nature and effect of Adam's sin. In his defense, Caelestius appealed to the teachings of a priest named Rufinus, whom Caelestius had heard oppose the notion of inherited sin. Pelagius himself traveled quickly through North Africa to Palestine where his teaching aroused Jerome's ire. In 415 Pelagius was called to defend himself before the bishop of Jerusalem and again before an episcopal synod at Diospolis, both of which acquitted him.

Indignant at these acquittals, Augustine—who had already written several anti-Pelagian treatises—led the literary and ecclesiastical attack on Pelagianism. Following conciliar, papal, and imperial condemnations in 418, Pelagius and Caelestius largely disappear from the historical record. Nineteen Italian bishops refused to subscribe to the papal proscription; among them was Julian of Eclanum, who wrote several lengthy polemical treatises, fragments of which survive embedded in Augustine's refutations. The judgment that Pelagian teachings were heretical was upheld by the ecumenical council at Ephesus in 431.

Modern scholarship has emphasized the importance of distinguishing between Pelagianism as a historical movement and Pelagianism as a theological system, the latter caricaturing the former. From the viewpoint of Christian orthodoxy, Pelagianism has often been construed as the heretical mirror image of Augustine's theology. Whereas Augustine defended established practices and doctrines such as infant baptism and original sin, Pelagianism controverted these and other traditions with novel heretical teachings that have been characterized as naturalistic, Stoic, and even godless. The theological tradition also canonized Augustine's characterization of Pelagians as *enemies of grace*, thereby implying that they deliberately denied grace, or at least reduced it to God's provision of the law and free will. Moreover, Pelagianism is accused of vainly overemphasizing the capacity of human free will. According to Augustine's full-blown predestinarian scheme, even the faith with which fallen human beings respond to God's gracious offer of salvation is itself a gift from God, given to some and withheld from others. As the opposite, Pelagianism implies an overconfidence that human nature is uncorrupt and possesses sufficient resources to attain moral perfection and eternal salvation solely by its own efforts without assistance from God's grace.

Like any caricature, this portrait of Pelagianism contains true features but distorts them by exaggerating some details and omitting others. The identification and subsequent scholarly analysis of additional Pelagian writings have revealed that Pelagian tenets are more nuanced than the prevailing stereotype suggests. Pelagius and his followers did not intentionally oppose Christian orthodoxy. Quite the contrary, they not only contrasted their teachings with the heresies of Arianism, Manichaeism, Origenism, and Jovinian, but also hurled countercharges of novelty and heterodoxy back at their opponents. As an historical movement, Pelagianism encompassed a diverse group of individuals who differed on a number of practical and theological issues but united in opposition to moral laxity and theological determinism. The defining characteristic of Pelagianism was not a negative denial of grace but, rather, the positive affirmation that it was possible (at least theoretically) for human beings to live sinlessly. If human beings *ought* to avoid sin—and most Pelagians considered this proposition a scriptural imperative—then human beings must be *able* to avoid sin.

Philosophical questions about freedom, responsibility, and justice were prominent in the Pelagian controversy but always in relation to theological concerns. For example, both Pelagius and Augustine strove to balance human free will and divine grace. Pelagius affirmed grace not only as God's creation of human free will and God's revelation through the law and through Christ, but also as the remission of sins in baptism and even as a constant help to free will, although Augustine dismissed the sincerity of the latter conception. Conversely, Augustine affirmed free will but apart from grace limited its scope in fallen humanity to choosing among evils. While Augustine accused his opponents of emphasizing free will to the extent that they denied any role for God's active grace, the Pelagians argued that Augustine's understanding of grace amounted to a determinism that eliminated free will.

The Pelagians defined sin as an act of will, not a substantial defect of nature; hence sin must be avoidable, and conversely that which cannot be avoided cannot be sin. Thus, when human beings choose to sin, they bear moral responsibility for their own actions and cannot blame God, the Devil, or even a vitiated nature. Consequently, the Pelagians understood the effect of Adam's sin as imitation of sinful habits rather than inheritance of a sinful nature, and most of them affirmed infant baptism, denying only that its function was to cleanse the newborn of inherited sin. Moreover, they argued that the inevitability and substantiality of original sin made God responsible for evil. For Julian, Augustine's teaching that the guilt of

Adam's sin was transmitted to each human being at conception also implied that marriage and reproduction were tainted by evil, therein betraying Augustine's lingering affinity with Manichaeism.

Finally, both sides in the Pelagian controversy refused to embrace theological positions that appeared to impugn divine justice. If sin were unavoidable, the Pelagians argued, it would be unjust for God to demand sinlessness and then to condemn human beings for sinning. Similarly, they saw injustice in the notion that God would condemn infants not for acts of their own volition but merely for inherited sin. Indeed, any god who would impute to one person the sins of another would be unjust. Augustine countered that a just God could not abide the suffering of infants unless these miseries were somehow deserved as a result of original sin, which rendered all humanity liable to God's just condemnation. Augustine posited that even God's sovereign choice to save some and not others, though an inscrutable mystery, could not be unjust.

See also Arius and Arianism; Augustine, St.; Augustinianism; Determinism, Theological; Freedom; Justice; Mani and Manichaeism; Origen; Philosophy of Religion; Religion; Responsibility, Moral and Legal.

Bibliography

WORKS BY PELAGIUS

De divina lege. In "Patrologia Latina." Vol. 30, cols. 106–116, edited by J. P. Migne, Paris: 1865 (included here among works falsely attributed to Jerome.

De natura. Fragments collected in "Pelagius' Schrift *De natura*: Rekonstruktion und Analyse." Winrich A. Löhr. *Recherches augustiniennes* 31 (1999): 235–294.

De virginitate. In *Sulpicii Severi Libri qui supersunt*, edited by C. Halm. "Corpus Scriptorum Ecclesiasticorum Latinorum," Vol. 1, 225–250. Vindobonae: C. Geroldi filium, 1866 (included here with the works of Sulpicius Severus).

De vita Christiana. In "Patrologia Latina." Vol. 40, cols. 1031–1046, edited by J. P. Migne, Paris: 1887 (here attributed to Augustine).

Epistola ad Celantiam. In *Sancti Eusebii Hieronymi Epistulae*, edited by Isidorus Hilberg. "Corpus Scriptorum Ecclesiasticorum Latinorum," Vol. 56.1, 329–356. Vindobonae: Verlag der Österreichischen Akademie der Wissenschaften, 1996 (here attributed to Jerome).

Epistola ad Demetriadem. In "Patrologia Latina." Vol. 30, cols. 16–46, edited by J. P. Migne, Paris: 1865 (included here among works falsely attributed to Jerome).

Expositiones XIII Epistularum Pauli. In *Pelagius's Expositions of Thirteen Epistles of St. Paul*, edited by Alexander Souter. Cambridge, U.K.: Cambridge University Press, 1922–1931 (Latin edition and commentary; 3 vols.).

Libellus fidei. In "Patrologia Latina." Vol. 48, cols. 488–491, edited by J. P. Migne. Paris: 1862.

OTHER PELAGIAN WRITINGS

Bohlin, Torgny. *Die Théologie des Pelagius und ihre Genesis*. Uppsala University Arsskrift 9. Uppsala: A.-B. Lundequist, 1957.

Caelestius. *Liber de 13 capitula*. Fragments collected in *Rom und Pelagius: Die theologische Position der Römischen Bishöfe im pelagianischen Streit in den Jahren 411–432*, edited by Otto Wermelinger, 297–299. Stuttgart: Anton Hiersemann, 1975.

Caspari, C. P., ed. *Briefe, Abhandlungen, und Predigten aus den zwei letzten Jahrhunderten des kirchlichen Alterthums und dem Anfang des Mittelalters*. (1890). Reprinted Brussels: Culture and Civilization, 1964 (Latin edition of six works by an anonymous Pelagian, with commentary).

Julian of Eclanum. *Ad Florum*. Fragments embedded in Augustine's *Contra Julianum opus imperfectum*, Books I–III, *Sancti Aureli Augustini Opera*, Sec. 8, Part. 4, edited by E. Kalinka and M. Zelzer. "Corpus Scriptorum Ecclesiasticorum Latinorum." Vol. 85.1. Vindobonae: Hoelder-Pichler-Tempsky, 1974; and Books IV–VI, "Patrologia Latina," Vol. 45, edited by J. P. Migne. Paris: 1865.

Julian of Eclanum. *Iuliani Aeclanensis. Expositio libri Iob; Tractatus prophetarum Osee, Iohel et Amos; accedunt operum deperditorum fragmenta post Albertum Bruckner denuo collecta aucta ordinate*. "Corpus Christianorum." Series Latina, Vol. 88, edited by Lucas De Coninck. Turnholt: Brepols, 1977 (Latin edition of two exegetical works and collected fragments of Julian's other writings).

Rufinus. *Liber de fide*. In *Rufini presbyteri Liber de fide: A Critical Text and Translation with Introduction and Commentary*, edited by Mary W. Miller. Washington, DC: Catholic University of America Press, 1964.

Fragments of Caelestius's *Definitiones* and Pelagius's *Pro libero arbitrio, Eclogarum*, and *Epistola ad Innocentium* are collected in "Patrologia Latina," Vol. 48, col. 255–698, edited by J. P. Migne. Paris: 1862 (also includes a discussion of Pelagianism).

SECONDARY STUDIES

Bonner, Gerald. *Augustine and Modern Research on Pelagianism*. Villanova: PA: Villanova University Press, 1972. Good overview of influential scholarship and its conclusions.

Bonner, Gerald. "Pelagianism and Augustine." *Augustinian Studies* 23 (1992): pp. 33–51, and 24 (1993): pp. 27–47. Updated summary and observations on trends in Pelagian research.

Brown, Peter. "The Patrons of Pelagius: The Roman Aristocracy between East and West." In *Religion and Society in the Age of St. Augustine*. London: Faber and Faber, 1972, pp. 208–226

Brown, Peter. "Pelagius and his Supporters: Aims and Environment." In *Religion and Society in the Age of St. Augustine*. London: Faber and Faber, 1972, pp. 183–207.

Clark, Elizabeth A. "From Origenism to Pelagianism." In *The Origenist Controversy: The Cultural Construction of an Early Christian Debate*. Princeton: Princeton University Press, 1992, pp. 194–244.

Evans, Robert F. *Pelagius: Inquiries and Reappraisals*. NY: Seabury Press, 1968. Highlights Jerome's role in the Pelagian controversy.

Ferguson, John. *Pelagius: A Historical and Theological Study*. Cambridge: W. Heffer & Sons, 1956.

Lamberigts, Mathijs. "Julian of Aeclanum: A Plea for a Good Creator." *Augustiniana* 38 (1988): 5–24.

Morris, John. "Pelagian Literature." *Journal of Theological Studies* 16 (1965): 26–60.

Nuvolone, Flavio G., and Aimé Solignac. "Pélage et pélagianisme." In *Dictionnaire de spiritualité, ascétique et mystique, doctrine et histoire*, Vol. 12B, cols. 2889–2942. Paris: Beauchesne, 1986. Excellent comprehensive survey of Pelagian literature and its authors and attribution, with extensive bibliography.

Plinval, Georges de. *Pélage. Ses écrits, sa vie et sa réforme*. Lausanne: Librairie Payot, 1943. Dated in its attribution of most extant Pelagian writings to Pelagius, but influenced Pelagian scholarship for decades.

Rees, B. R. *The Letters of Pelagius and His Followers*. Woodbridge, Suffolk: Boydell Press, 1991. Translations of a number of Pelagian letters and ascetic treatises.

Rees, B. R. *Pelagius: A Reluctant Heretic*. Woodbridge, Suffolk: The Boydell Press, 1988.

Michael R. Rackett (2005)

PERAS

See *Apeiron/Peras*

PERCEPTION

The term *perception* may be used generally for mental apprehension, but in philosophy it is now normally restricted to sense perception—to the discovery, by means of the senses, of the existence and properties of the external world. Philosophers have been concerned with the analysis of perception—that is, the study of its nature and of the processes involved in it—and with its epistemological value—that is, how far, if at all, it can be regarded as a source of knowledge about the world. Their answers to these closely interrelated questions have been formulated in various theories: the commonsense theory and other kinds of direct realism, the representative or causal theory, critical realism, the sense-datum theory, and phenomenalism. This entry will be devoted to the main features of perception that underlie the various theories and that have raised philosophical problems and controversy. It will discuss both the initial evidence that may be analyzed without recourse to scientific findings and the causal and psychological process revealed by scientific investigation.

INITIAL EVIDENCE AND ANALYSIS

REFLECTIVE EXAMINATION. As percipients we are all familiar with perception, and so the first evidence should come from reflection on our own experience. The following points may thus be made about perception.

First, it is awareness of the external world—of material objects, to use a technical term for physical objects in general, animals, plants, and human beings insofar as they are perceptible (their bodies, in fact). The main characteristics of such objects are that they are external, independent of the percipient, and public, meaning that many people can perceive them at once. Perception, in being the awareness of such objects, may be contrasted with imagery, bodily sensations, or having dreams.

Second, perception is, or seems to be, intuitive—immediate and normally undoubting, a direct face-to-face confrontation with the object in sight or a direct contact in touch. Nor are we normally conscious of any processes of reasoning or interpretation in it. On the rare occasions when we reason or we have doubts about what an object is, the reasoning or doubts are about the identity or character of something already perceived—for instance, a rectangular red object or something white on the hillside.

Third, perception is variable in quality and accuracy; we may fail to notice something, to see clearly, to hear distinctly, and so forth. Three types of variation may be involved: variations in attention, in what we notice or discriminate; variations in quality or distinctness (for instance, where there is nearsightedness or fog); and variations in liability to err—we may misidentify what we perceive or mistake its qualities.

Fourth, perception nevertheless normally gives us knowledge of material objects and properties. With a few fairly obvious tests, like touching and looking closely, or using the evidence of other percipients, we can establish certainty or else correct the first sight or hearing.

Fifth, perception often issues in some judgment or assertion (to others or perhaps only to oneself)—for example, "There is a green fly on the roses" or "Here's the milkman"—but it may not.

ILLUSIONS. Illusions, comprising illusions proper, hallucinations, and cases of the relativity of perception, have traditionally been the most important origin of the major problems of perception. The two main claims of the argument from illusion are (1) illusions show that perception is never absolutely certain, that tests are never final, and (2) the appearances we are aware of in illusions, especially

hallucinations, cannot be identified with the real properties of objects and must therefore be private objects of awareness, or sensa (indeed, all perception involves awareness of sensa that in correct, or veridical, perception belong to the object or correspond to its properties).

The first claim was long thought to rule out perceiving as a source of knowledge; instead, one had to turn to pure reason, or rational intuition, which was held to provide mathematical knowledge. But, since the absolute certainty of mathematics came to be generally ascribed to its ultimately analytic, or even tautological, character, the tendency now would be to stress the negligibility of the possibility of error in tested perception and to use a different standard of certainty and knowledge concerning matters of fact, one that allows perceptual statements to qualify.

The second claim, concerning the existence of sensa, is vital in that almost all theories of perception either found their analyses on it (as does the sense-datum theory) or seek to controvert it or explain it away (as does commonsense realism). The seeds of this conflict already lie in the results of the reflective examination. Insofar as perceiving seems to vary in quality and accuracy, it is easy to say that in illusions we merely see the object looking different from what it is. But if perception is a direct intuitive confrontation, the illusory appearance must be a genuine existent, perceived as it really is, a sensum in fact; "looking different from what it is" must be interpreted as "presenting sensa different from the standard ones." In any case, some phenomena—for example, the integration of hallucinatory images with a perceived background—are difficult to explain without supposing awareness of private sensa in all perception, and almost all the phenomena require scientific and psychological findings for their full explanation, thus pointing beyond this initial evidence.

PERCEPTUAL CONSCIOUSNESS AND PERCEIVING. The occurrence of illusions may lead to ambiguity in the use of *perceive* and allied terms. Thus, in double vision a man may be conscious of two bottles where there is only one. Do we say "He perceived two bottles" or "He perceived one bottle"? Each alternative has been adopted philosophically, and to avoid ambiguity, it is safer to distinguish (1) "*X* is perceptually conscious of *Y*" from (2) "*Y* is present to *X*'s senses (or light from it is acting on his sense organs)" and use "*X* perceives *Y*" only when both are meant.

This recommendation is claimed to have the further advantage of enabling us to discuss as perceptual consciousness the state of mind (or mental act) occurring in both veridical perceiving and illusions. Perceptual consciousness of, for example, a dagger might occur when only a stick was present or even, as in Macbeth's case, when nothing was there. The notion of such consciousness as a common factor in perceiving a real dagger, in having hallucinations of one, and in mistaking something else for one fits in best with dualist theories, such as the sense-datum theory (especially H. H. Price's version) or critical realism, since it suggests that the contents of such consciousness differ from the external object perceived. Direct realists are suspicious of it; for them having hallucinations is something (imagery, perhaps) quite different from normal perception, even if confused with it, whereas in illusions they want to stress that one is perceiving the real object present—seeing a stick *as* a dagger or the round table *as* elliptical.

But even if perceiving a round table as round or in perspective as elliptical is taken as immediate confrontation needing no further analysis, seeing a stick as a dagger (or a piece of wax as a tomato or a bush in a fog as a man) can hardly be equally simple and immediate. In such cases and in hallucinations one has to admit that one seems to see an object quite different from that present to the senses. This can fairly be described as perceptual consciousness of the (ostensible) object (dagger, wax, or man) and distinguished in analysis from actually perceiving an object (dagger, wax, or man). And in view of the subjective similarity it is but a short step to suppose that perceptual consciousness of *X* also occurs in perceiving *X* as *X*, the difference between illusory and veridical perception of an *X* lying not in this common consciousness but in whether *X* is present and acting on the sense organs. Any philosophy of perception should analyze this perceptual consciousness and explain how it may occur without the presence of the corresponding object.

ANALYSES OF PERCEPTUAL CONSCIOUSNESS. Three major analyses are integral parts of the theories of perception mentioned above. First is the traditional notion that perceiving—that is, perceptual consciousness—is the interpretation of sensations as properties of external objects. Second, the sense-datum theory claims that perceptual consciousness is taking for granted that the sense datum one is sensing belongs to a material object. Third, the analysis of the critical realists, though stated as an analysis of perceiving, amounts to saying that perceptual consciousness is taking an intuited datum or character complex to characterize an external object.

The essential difficulty in these three analyses is that they contradict introspective evidence by splitting up perceptual consciousness into the awareness of some private data, recognizable by analysis as such, and the act of interpreting them as, or taking them to be, objects or object properties. In experience there is no such core of sensing or intuiting data distinguishable from the consciousness of a material object, even if only subsequently; still less is there any passage of the mind from awareness of sensation as such to object perception. And if some critical realists are less liable to this difficulty because they do not treat their data or character complexes as existents readily distinguishable from material objects, they do this only at the expense of obscurity or disagreement as to what the data are. Attempts at a remedy must be postponed until the psychological processes in perception are considered.

A fourth analysis is the idealist claim that all perceiving is judging, which is really an analysis of perceptual consciousness but is easier to follow if stated in terms of perceiving. It is that perceiving (perceptual consciousness) consists in making a judgment, which has an implicit sensory basis, about the real existence of an object or property. Thus, perceiving a tomato on a plate or perceiving that the dog has hurt its leg are the sensorily grounded judgments "There is a tomato on a plate" or "The dog has hurt its leg." The "perceiving that" description of a perception (for example, "He saw that the dog was hurt") certainly seems to suggest judgment, though the form may be misleading and may only be for emphasis of the feature noticed. But the main reasons for this analysis are (*a*) that perceiving is true (veridical) or false (erroneous) and only judgments or assertions can be true or false and (*b*) that perception is more than just sense experience, for we identify and interpret what is given (that is, it involves inference from implicit data, and the conclusion of such an inference must be a judgment).

One may object that truth characterizes what is asserted, not the asserting—the judgments, in the sense of propositions, to which perceiving may lead, but not the act of perceiving itself. Perceiving may be proper, correct, clear, or accurate, but not true or false. Many other things we do may be done correctly and be liable to errors without being forms of judging, such as playing the piano, playing games, tying knots. False judgment is not the only form of error. Also, the idealist doctrine that all perceiving is judging is open to the general objection to the first three analyses above, particularly because the nature of the implicit data or sensory grounds is very obscure. Attempts to elucidate it—for example, Brand Blanshard's—turn them into sensa. Furthermore, the term *judgment* suggests something intellectual, explicit, and considered, with consciousness of the evidence for the assertion—conditions inappropriate to much perception. Also, we may correct a faulty judgment on learning the truth, but such knowledge does not enable us to correct illusory perceptions; we still see the mirage, and the railroad tracks still appear to meet in the distance.

Fifth, there is a causal analysis of perceptual consciousness—namely, that it is inferring that one's sensa are caused by an external object. This may be associated with representative realism but is not essential to it; representative realism's main thesis is that the sensa and the consciousness are externally caused by objects that the sensa "represent." One may accept this thesis along with any of the analyses of perceptual consciousness—the causal inference it involves is subsequent to the perceiving, and so is a claim about perception. The difficulties of supposing that perceptual consciousness *consists in* such an inference from effect to cause are that (*a*) we are not conscious of such an inference; (*b*) if we started only with private sensa, any inference to external causes would be too difficult and complex to be automatic and unconscious—it would *have* to be conscious; and (*c*) it leads to paradoxes, such as that children, being ignorant of the supposed causation of perception, cannot therefore perceive or be perceptually conscious of anything.

CONCEPTUAL ANALYSIS OF PERCEIVING. A rather different approach to perceiving is adopted by those who advocate conceptual analysis—that is, a close study of the ordinary meaning (or use) of expressions. This analysis is naturally associated with commonsense realism, for ordinary language tends to reflect ordinary views on perception or at least what once were such views. Such analysis, however, may well indicate features of perception that are not normally realized and so supplement or even correct reflective examination of an introspective kind.

Much attention along these lines has been directed to the categorization or classification of perceiving. Previous philosophers have referred to perception in various ways: as an act, even an operation, as a process, and as a mental state. None of these is satisfactory. "Act," at least as activity or operation, suggests listening or watching rather than just hearing or seeing; "state" and "process" suggest something long-term, and "process," like "activity," suggests something open to public observation—yet whereas one may observe *X* looking at *Y* one cannot observe *X* seeing *Y*. (One can perhaps claim that the best description of perception is "mental act," which would put per-

ceiving in a special category with realizing, noticing, deciding and so on, but mental acts as such are suspect to these philosophers.) One suggestion is that perceiving is simply having an experience, but this neglects the active side of recognizing and identifying involved in it. A more popular suggestion is that perceiving is a skill or art, or, rather, since seeing *X* or hearing *Y* occur at a definite time, perceiving is the exercise of a skill.

Oddly enough, the evidence for this is not linguistic. We may speak of a skilled observer, one who can direct and coordinate a series of perceptions, but not a skilled perceiver; we do not say that *X* is an expert at the art of seeing or hearing things. Rather, this suggestion is based on the fact that perceiving can be improved by learning and experience, so that one recognizes things easily, avoids mistakes, or can make allowances for such factors as distance. Although this may occur to one on reflection, however, its full and precise extent has been established only by psychological investigation. As soon as one seeks out this and other psychological evidence about perceiving or even asks how one learns by and exploits experience in perceiving, one is carried far beyond language and conceptual analysis to a scientific study of the subject. Also, to maintain that perceiving is the exercise of a skill brings one back to the suggestion that it is an operation or activity.

More striking perhaps was the earlier claim of Gilbert Ryle that "perceiving" is an achievement verb, like "finding" or "winning," and indicates the scoring of an investigational success. This means that perceiving is not an activity or process, though it may be the successful termination of the activity of looking for something; it is instantaneous, not something that takes time or can be observed. Ryle's aim was to attack representative realism and its associated dualisms of mind and body, sensa and object, by claiming that (*a*) perceiving, usually thought to be a private mental activity because it is not an overt one, is not an activity at all and thus provides no evidence of a mental world and that (*b*) since it is not a process, perceiving is not the final stage or effect of a process, particularly not of the causal process from object to person. Hence, there is no need to suppose that science proves that perceiving is awareness of private sensa.

These are not very convincing arguments. As to the first, winning or scoring involves some activity such as kicking a ball. Likewise, perceiving involves experiences of colors or sounds and the psychological processes discussed below; these are normally claimed to be mental. The second is a non sequitur—instantaneous success may be the end and result of a causal process. Thus, scoring and finding may be observed and may be the result of a process or series of activities; other conditions may also be required but do not rule out their being effects. More generally, if perceiving is an achievement, what are misperception, illusion, failure to see properly, a casual glance? An analysis of perceiving must take these into account and not apply only to veridical perception. Ryle also failed to show how perceiving is related to the causal processes representative realism emphasizes. Thus, if instantaneous, perceiving can no longer be the relation across time and space that direct realism would need to claim in view of the factually verified time lag, the time taken by the causal transmission from a distant object. Indeed, contrary to Ryle's intention, a dualist interpretation of his claim is possible. Perceptual consciousness is instantaneous; when it is also successful, that is, when its content corresponds to the properties of the object causing it, it is perceiving; when unsuccessful, it is misperception or illusion.

THE CAUSAL PROCESSES IN PERCEPTION

THE CAUSAL CHAINS. The causal processes involved in perception form causal chains from the external object to the percipient's brain. In sight a complex system of light waves, sometimes emitted by the object but normally a differential reflection of light from the object's surface, travels from the object to the percipient's eyes. This system is diversified in intensity and wavelength according to the shape, brightness, and color of the object surface and, on striking the eyes, is focused so that an image of the object is cast (upside down) on each retina. Each retina has a mosaic of more than 120 million receptors, which are activated by the light cast on them in this image. The light causes chemical changes in the receptors; these changes, in turn, cause electrical impulses to pass along the nerve fibers that lead from the receptors to one of the two visual receiving areas of the brain. The impulses set up activity there and in certain other association areas; this done, the person then sees the object. More than one million such fibers form the optic nerve from each eye, and each fiber consists of a succession of cells that are made to conduct by a chain reaction; the resultant impulses can be picked up and reproduced on a cathode-ray tube.

In hearing, a pattern of sound waves is emitted or reflected from the object and strikes the eardrum; this causes vibrations to be transmitted through a series of bones to the liquid filling in the inner ear, thereby setting up vibrations in the basilar membrane of the cochlea

according to the frequency (waves per second, corresponding to pitch) and intensity of the sound. The receptors in the cochlea then transmit electrical impulses along the nerve fibers to another receiving area in the brain. (These impulses are not at the frequency of the incoming sound.)

In smell and taste there is a chemical stimulation of receptors in the nose and tongue by particles of the substance perceived, and the receptors, in turn, send neural impulses to another area of the brain. For touch, the brain is linked to receptors all over the skin, some of which respond to the pressure of direct contact with the object, some to heat, and some to cold (or, rather, to rate of change in skin temperature). Other receptors in the skin and the body respond to a wide range of stimuli by transmitting to the brain impulses that ultimately cause a sensation of pain. There are also other senses—for example, a kinesthetic sense by which receptors in the muscles send impulses to the brain so that the position of the limbs is sensed or unconscious adjustments are made to guide and make efficient voluntary movement. There are also receptors in the vestibule and semicircular canals behind the ear that assist balance and give us information about head position.

The chain process (object-[waves]-receptor-nerve impulses-brain activity) is a necessary condition of perception of an external object, for if it is interrupted by damage to the sense organ, no perception occurs. It is not a fully sufficient condition in that other areas of the brain must be suitably active so that the person is conscious and minimally attentive—that is, not wholly absorbed in thought. The interesting question is whether or how far the chain process is necessary and sufficient for perceptual consciousness of an object, granted conditions of consciousness and attention. At least the brain activity is clearly necessary, but theoretically one might insert stimulation at some point on the chain and thus cause experiences the same as those that would normally be attributed to the external object. This apparently happens naturally in illusions and hallucinations, including phantom limbs, and electrical stimulation of the appropriate areas of the brain may cause sensations of color, smell, or touch. (The sensations are not like the contents of perceptual consciousness of objects, but this difference may be due to the comparative crudity of the artificial stimulation by an electrode; also, activity in the association areas is necessary for normal perception.) Thus, it seems probable that suitable activity in the nervous system is a necessary and sufficient condition of perceptual consciousness, though it may be that some kind of external stimulation, even one quite unlike the object perceived, is required to trigger it.

TIME LAG. Causal processes take time. In the case of distant objects this is marked. Thus, because sound waves travel much more slowly than light waves, the flash of some distant gunfire or explosion may be seen appreciably before the sound is heard. Even at its great speed light takes eight minutes to reach us from the sun and four years and four months from the nearest star. Consequently, we may well be "seeing" a star long after it has disintegrated, for the perceptual consciousness occurs at approximately the time of the arrival of the star's light on Earth. But as time is required for the sense organ to be activated and for the nerve impulses to travel to and spread in the brain, there is a slight but variable time lag in all perception; an accurate estimate is not possible but the delay is probably of the order of one-tenth of a second for nearby objects.

UNIFORMITY OF NERVE IMPULSES. One surprising fact is that the nerve impulses are of a similar type for all the senses. All that travels along any nerve from any receptor to the brain is a sequence of such impulses varying normally between 10 and 100 per second. The frequency variation is, in fact, a mark of intensity; the stronger the stimulus, the more impulses per second. Consequently, what distinguishes causation of an experience of sound from that of smell or an experience of a high pitched sound from that of a low one is not the impulse itself but the connections of the nerve fibers excited and conducting—where they start in the sense organ and where they end in the brain. (Though if one imagines a cross section across a bundle of nerve fibers, the pattern of some conducting and some not conducting can be regarded as a changing code.) Thus, excitation by nerve impulses of one tiny portion of the brain results in awareness of a loud shrill sound, excitation of another in awareness of a blue line. Various areas of the body are mapped in the brain, a group of receptors (or sometimes an individual receptor) in the skin and tissues corresponding to each point in the cerebral receiving area. Similarly, the retinal image is reproduced point by point in the brain, though with each half reproduced in a separate area, duplicated there, and distorted. Again, a strip of brain tissue is activated at different points according to the frequency of the sounds heard, as if it were a keyboard.

COMPLEXITIES. There are nevertheless many complexities in the system, only a few of which we can mention

here. The nerve connections are intricate, with feedback fibers from the brain to the incoming sensory fibers and cross connections between the sensory fibers; in the grouping of receptors and in the brain there is summation—several nerves join one that conducts only when all or most of them do. (In fact, neurologists constantly use such terms as *selecting, integrating, summating*, and *coding*.)

Binocular vision involves retinal disparity (a slight difference in the images cast on the two retinas) and the operation of two visual receiving areas reached by crossed-over nerve fibers so connected that the left-hand receiving area receives the signals coming from the right half of each retina and the right-hand area receives those of the left half. As a result we somehow normally see one object with depth and solidity rather than two two-dimensional ones.

Constant small eye movements are necessary for vision, with a shifting of the retinal image and of the resultant pattern of impulses in the fibers of the optic nerve, yet the object is seen as steady.

Most of the impulses reaching the brain from the eye come from a small portion of the retina (the fovea) that has relatively many receptors giving great distinctness; for exact vision the image is focused on the fovea by eye movement.

Color vision is particularly complex, and its mechanism is disputed. All the colors we know can be produced by suitable mixtures of red (long-wave), green (medium-wave), and blue (short-wave) light. White light can be formed by an appropriate combination of three colors or even of two widely separated ones. (Light, or "spectral" colors, mix differently from paint colors.) The simplest theory is that there are receptors in the eye reacting to each primary light color (red, green, and blue) and the brain, by summating the three color inputs, is enabled to cause the final color sensation. Thus, grass looks green because it absorbs red and blue light but reflects green, and a buttercup is yellow because it absorbs blue but reflects green and red, which combine to produce yellow. There are many difficulties in this theory. For instance, no receptors for blue can be positively located in the eye, only for red and green ones; the light from a green surface actually contains a mixture of wave lengths, with green predominating; the light wave lengths cover the spectrum of all the colors of the rainbow; the brightness and purity of the color also affect its hue. A final theory must therefore be very complicated.

The auditory receiving area gets impulses from both ears. This enables us to locate the source of a sound. If a sound is to the left, then sound waves reaching the left ear differ in phase (that is, timing of the wave crests) and in intensity from those reaching the right ear. The brain apparently combines the different inputs so that the location is done unconsciously and we just hear the sound as if it came from a certain direction.

LIMITATIONS OF THE SENSES. Radiant energy is known to range from short cosmic rays to long radio waves, but the eye responds only to visual light, which is a narrow band occupying about one-seventieth of the whole range. Even then we cannot distinguish light of different polarizations, as bees and birds apparently can, or see very small objects or fine structures. Similarly, in hearing we can distinguish only waves between 20 and 20,000 cycles per second; dogs, cats and rats can hear higher notes. Our sense of smell is obviously very inefficient compared with that of most other animals. Hence, though we can extend our range of observation by microscopes, infrared or X-ray photographs, radiotelescopes, and so on, it is clear that our senses themselves are very limited as a direct source of knowledge of the external world.

THE CAUSAL ARGUMENT. The causal argument maintains that the existence and character of these causal processes refute direct realism and force the adoption of a dualist position. Perception of an external object cannot be the direct contact or immediate confrontation it seems to be, since it requires this causal chain from object to the percipient's brain and is prevented if that is interrupted—for instance, if the optic nerve is cut or one of the small bones in the ear does not move properly. In this sense directness or immediacy must mean no intermediary and no possibility of interruption. The causal chain suggests that perceptual consciousness and its objects are generated, or brought into being, by the causal process, presumably by its last stage, the brain activity. In other words, insofar as perceptual consciousness is intuitive, it is awareness of some content or object quite distinct from the external object.

This suggestion is supported by various points. First, the time lag—perceptual consciousness may occur after the external object has disappeared or moved, so its content cannot be identified with the object. Second, the possibility of perceptual consciousness without any external object at all or without one at all similar seems confirmed by the production of sensations by stimulation of the cerebral cortex and seems actualized in hallucinations.

Third, the enormous complexity involved shows that the subjective simplicity of perception is illusory, at least insofar as a relation to an external object is concerned. Fourth, illusions and the relativity of perception are often explicable in terms of the causal processes. Unless the contents of perceptual consciousness are generated and conditioned by the causal process, one has to attribute bizarre and contradictory properties to the external object. Fifth, the simplicity and uniformity of nervous impulses show that they cannot transmit all the various secondary qualities that make up objects as we know them. These qualities must thus characterize contents of consciousness generated by the causal process. (This point is supported by such other limitations as the purely mechanical transmission through the bones of the ear.) Hence, it follows from the fourth and fifth points that one must abandon the other assumption of direct realism—that even when we are not perceiving them, objects continue to exist with the exact qualities we normally observe in them.

There is a good deal of resistance to these conclusions. One obstacle is that they seem to require a self-refuting type of representative realism. This fear is unjustified. It must also be noted that granted the dualist conclusion that the causal process generates the sensory experience whose content is (numerically) different from the external object, the nature of that experience and its content is still open. It may be that the awareness is of sensa, or it may be a full-fledged perceptual consciousness of percepts or ostensible objects. One is not even forced to adopt a mind-body dualism, though it is normally thought that sensa or percepts are mental. One might claim that though apparently distinct objects, they are in fact only the contents of sense experience, not existing apart from the sensing of them (adverbial analysis), and that they and the brain activity are two aspects of the reaction of the organism or person as a whole. This would mean that sensa are only a correlated aspect of brain activity, not effects of it, though still conditioned by the rest of the chain. In this way one might bypass one of the notorious difficulties of ordinary dualism—the unique and obscure causal relation supposed to exist between material brain and immaterial mind.

Sometimes, however, the opposition takes the form of denying the relevance of the scientific evidence to philosophy; it tells us only what the causes of perceiving are, not what perceiving itself is. Philosophers must investigate the latter and leave the causal processes to the scientist. But scientists normally hold that these processes require the adoption of representative realism, thus giving them philosophical relevance; also, those philosophers who wish to concentrate on the nature of perception alone usually come up with some answer (the sense-datum analysis or a view that perceiving is the exercise of a skill or an investigational success) that is compatible with or even supports a dualist interpretation of the causal processes. But, above all, to achieve full understanding of anything so vital as perception, one must consider its causes and conditions, particularly as their study has traditionally been claimed to transform our concepts of perceiving itself and of our knowledge of the external world.

THE PSYCHOLOGICAL PROCESSES

It is clear from experimental psychology that perceptual consciousness involves a whole range of adjustments and selective or quasi-interpretative processes. The main evidence for this lies in differences between what psychologists often call the phenomenal properties of an object (those we are perceptually conscious of) and its stimulus properties. In this context the stimulus is the pattern of light rays from the object striking the eye, of sound waves striking the ear, or of heat or pressure from touching the object. The stimulus properties are those that we should observe in the stimulus (such as shape, color, pitch) could we observe it directly and in itself. This is difficult to achieve, and in fact the evidence of cameras, tape recorders, and other instruments is used, plus knowledge of the nature of the object and reasoning from the laws of perspective or of physics generally. The difference between the two kinds of property is presumed to be the result of modifications by the percipient.

ATTENTION AND SELECTION. It is a simple fact of experience that the quality and accuracy of perception vary with our attention. We often look inattentively and fail to notice pronounced features of a scene, yet we may carefully observe and thus notice unexpected details—a mark on the wallpaper, a printer's error, a wrong note in a recorded symphony. From the evidence of other people, from photographs, and from other means there may be no doubt that these features appeared all the time in the stimulus properties even when we were unconscious of them. Besides confirming this, psychologists have shown how greatly what we do or do not notice depends on habits of attention or interests, on often unconscious "priming" or "set." A mother will hear her baby cry but not notice much louder noises; an architect may notice features of buildings, and a boy notices makes of cars, both being oblivious to much else. Thus, perceptual consciousness is very selective, and this selection is usually

largely unconscious, though voluntary attention can greatly modify it. One special case of voluntary attention is of importance—"perceptual reduction" or "phenomenological observation," where we concentrate on the sensible qualities of what we perceive and not, as is usual, on the identification of the object concerned. An artist must do this when he has to paint a scene, and this kind of observation may reveal all sorts of previously unnoticed details of color, shape, and so on. It is open to question whether this kind of reduction reveals an element present in all perception—namely, sensing—or whether, and this is more plausible, it is simply a special kind of perception of external objects not found in normal perceiving.

ERRORS AND ENRICHMENT. Some errors in perception can be attributed to psychological factors—misidentifications because of careless observation, seeing what one expects to see rather than what is actually present, thinking that one hears the expected visitor coming when no one is there, and the like. These point to a common characteristic of perception and one apparent only when it goes wrong—the enrichment of perception by imagery and thought. Many psychological experiments have been concerned with this. For example, vague or ambiguous stimuli (pictures or sounds) are presented to different groups of people who see or hear them as definite objects or words, and the direction in which they are thus unconsciously supplemented or altered can be shown to be caused by suggestion or by the interests, emotional state, or physical state of the person. Another kind of case is the divergence between several eyewitnesses' accounts of an incident, which may all differ from a filmed record. Again, blind spots or other visual defects are often not apparent to the subject, who unconsciously fills in the gap (this happens to us all if we look with one eye, for there is a blind spot where the optic nerve leaves the retina). Extreme cases are hallucinations where the apparition is integrated with the background or casts shadows. Unnoticed supplementation by imagery, which is admittedly private and mental, seems strong evidence for the dualist claim that the contents of perceptual consciousness are similarly private and must be distinguished from object properties.

LEARNING AND CUES. Our perception is clearly affected by learning and experience. Identification and discrimination afford obvious examples; one can learn to identify objects seen or photographed from unusual angles, to detect animals in natural camouflage, to distinguish different birds' songs. Driving a car involves perceptions of distance and relative speeds, perceptions that are acquired by experience. Psychological investigation has shown the role of learning to be far greater than this. Perception of spatial relations generally depends to a large extent on learning (normally unconscious and in childhood) to harmonize sight and touch and to use various cues. This is shown by various experiments, such as those with distorted rooms or inverted spectacles, and by the evidence of blind men who recover their sight. Among the various cues used for perception of distance and of solidity are shadows, aerial and linear perspective, parallax (or relative movement), and the interposition of objects. These assist binocular vision and enable us to see depth even with one eye.

FIGURE-GROUND AND GESTALT. In perception our immediate consciousness is of an organized or structured whole. Some shape or feature stands out and is seen as the figure against a background, and if discrete units such as dots are presented, we see them as grouped or patterned in some way. This characteristic of experience has been particularly stressed by Gestalt psychologists, who produced much experimental evidence to show that we see wholes or structures (Gestalten, literally, "forms") and that perception develops by discriminating these in and from a background and not by synthesis of atomic elements or point sensations first perceived separately. Such organization of the visual field, though little affected by learning, is nevertheless largely the result of processes in the percipient himself. The clearest evidence of this comes from the reversals, or "alternating illusions," where the stimulus (picture or succession of sounds) is constant but is perceived differently at different times; thus, sometimes one pattern or shape stands out as the figure, sometimes another. Examples are the goblet that may appear as two faces in profile, Edwin Boring's wife–mother-in-law figure, Ludwig Wittgenstein's duck-rabbit, and the staircase that seems to be seen now from above, now from below.

PERCEPTION OF MOTION. Perception of motion was closely investigated by the Gestalt psychologists, who drew attention to the Phi phenomenon, which is the impression of movement between adjacent stationary stimuli that are activated in succession. This underlies the consciousness of movement on a motion picture or television screen and is used in illuminated advertisements in which if groups of lights are successively switched on for a brief time, one is perceptually conscious of a moving figure or even of words moving along. Intermittent illumination may also make moving objects appear stationary. Thus, when illuminated by the flashing light of a

stroboscope, a moving crank in a machine may, if the flashing is properly adjusted, be seen as stationary and examined for defects; if there is a slight maladjustment of the flashing, it may seem to rotate slowly backward like the wheels of coaches in Western films. There is a clear distinction in these cases between the properties of the stimulus and the contents of consciousness. Figure-ground effects also occur in movement perception, such as when the moon seems to sail through the clouds or when one's stationary train seems to move if an adjacent one starts.

OBJECT CONSTANCY. The widespread phenomenon of object constancy in perception differs from the above in that the phenomenal properties of an object tend to remain constant or nearly so even though the stimulus properties vary considerably. Thus, when we look at a round object—for example, a dish—from an angle, it often still looks round and not elliptical, although by the laws of perspective the stimulus (light-ray pattern) or retinal image is elliptical, as would appear on a photograph taken from the percipient's viewpoint. (This causes complication in stating the argument from illusion and perspective realism.) Only if the angle is very marked does the dish look elliptical. ("Look" here refers to the sensible quality, not to what we judge to be the object's shape.) Similarly with size, brightness, and color—a man looks much the same size at ten yards' distance as at five even though the image cast on the retina is half as high in the former case; a white patch in the shade reflects less light than a dark one in bright sunlight, but it still looks white; a white patch in a yellowish light still looks white although it is reflecting yellowish light (one may be surprised by color photographs taken in the evening, for the camera cannot adapt itself to the yellower light).

In general, over a range of varied stimuli we tend to see something corresponding to the property of the object or at least some compromise between this and the stimulus property. Experiments show that this constancy depends not on knowledge of the object but on the visibility of its background, for if the background is cut off by a screen so that only the object is visible, constancy does not hold and the stimulus property is seen. It is as though we made unconscious allowance for distance, angle of sight, and illumination as revealed by the whole scene. But this is not a learned or intelligent adjustment; children and even chickens or fish apparently see things with constancy, though to some extent it can be counteracted by adopting a stimulus attitude (trying to see the stimulus property).

PHILOSOPHICAL SIGNIFICANCE. The existence of these many complex processes that underlie perceptual consciousness and affect its content reinforces the causal argument by making even more incredible the direct-realist notion of perceiving as a straightforward direct confrontation with the actual properties of objects. If perception were a simple intuitive awareness of such properties, there would be no place in it for variations in quality; for the effects of interests, priming, and learning; and for the use of cues for enrichment by or integration with supplementary imagery, for constancy adjustments (especially where they produce a compromise between object and stimulus properties), for changing figure-ground effects, or for the Phi phenomenon.

The range of these processes is far greater than that which would be compatible with the usual analyses of perceptual consciousness—namely, that it is the interpretation of sensations (or inference from implicit grounds) or the taking for granted that a sensed datum belongs to an object. These views were mainly influenced by the possibility of error in perception, particularly in identification, although they took some account of the use of cues and of the role of learning. But they seem inadequate to cover the part played by attention and unconscious selection or by such organization adjustments as figure-ground, grouping, object constancy, or the Phi phenomenon, whereas some of the imaginative supplementation goes far beyond what can be called interpreting a datum. It is sometimes claimed that these adjustments are interpretations. But this is implausible, for they seem little affected by learning and are not intelligent since lower animals make them. Nor can many of the illusions or adjustments be overcome by knowledge of the facts or by conscious interpretation; where some counteraction is possible, as in object constancy, it is very difficult, and for most people the presence or absence of screens in experiments is compelling in its effect.

The final objection to such analyses concerns the alleged pure sensory data; interpreting or taking for granted, insofar as we are aware of it, is of something we are conscious of as distinct and external and which is thus already the effect of many of these processes. Normally, however, perceptual consciousness seems intuitive—that is, without interpretation and quite unanalyzable; except in perceptual reduction its content almost always consists of ostensible objects. All the same, psychological evidence shows that there is a range of subjective processes. The only answer seems to be a genetic hypothesis, not an analysis into elements. Perceptual consciousness is introspectively a whole but must be supposed to be a product

of a range of selective, supplementary, integrative or organizational, and quasi-interpretative processes acting on a supposed basic sentience. But—and this is the point—both processes and sentience are unconscious and so may plausibly be regarded as cerebral activities or adjustments of the nervous system. However, since we cannot as yet give any precise neurological statement of these processes, we have to describe them as if they were conscious, basing the description on the difference between the input to the senses and the finished product, but this product (perceptual consciousness) does not reveal within itself the processes that may be supposed to form it.

The suggestion that perceptual consciousness is the product of many unconscious processes is controversial, and any general conclusions about perception are bound to be personal. Hence, the main attention in this entry has been on the facts that have to be taken into account in any fully adequate view of perception, and the reader is also referred to the statement of the various theories here and in other related entries. In this way one has the material for assessing the general view here adopted—namely, that the causal and psychological processes essential to perception, as well as its liability to illusion, require abandonment of direct realism for a dualist position. One must distinguish perceptual consciousness, whose content or objects are subjective and private to the percipient, from perception that occurs when this perceptual consciousness is caused by an external object with properties corresponding to its content. But one must not confuse this dualism with the traditional representative realism, which is only a variant of it, some form of critical realism being superior; the sense-datum theory's dualism of sense data and objects (perceptual consciousness is not thus analyzable, and its content consists of ostensible material objects); or the Cartesian mind-body dualism (it is possible also to adapt this view of perception to a double-aspect account of mind and body).

See also Illusions; Phenomenalism; Primary and Secondary Qualities; Realism; Sensa.

Bibliography

GENERAL INTRODUCTIONS

Perception is discussed, though usually less extensively than in the related entries in this encyclopedia, by most introductory books on philosophy. For example, see Bertrand Russell, *Problems of Philosophy* (London: Williams and Norgate, 1912) and *Outline of Philosophy* (London: Allen and Unwin, 1927); Charles Harold Whiteley, *Introduction to Metaphysics* (London, 1955); John Hospers, *Introduction to Philosophical Analysis* (New York: Prentice-Hall, 1953; London: Routledge and Paul, 1956); and Arthur Pap, *Elements of Analytic Philosophy* (New York: Macmillan, 1949). A clear and useful outline with emphasis on the associated epistemological problems is A. J. Ayer's *The Problem of Knowledge* (London: Macmillan, 1956). A more detailed introductory treatment is given by R. J. Hirst, *The Problems of Perception* (London: Allen and Unwin, 1959), Chs. I–VI; the later chapters of this book develop a more advanced treatment of perceptual consciousness and the scientific evidence on the lines adumbrated here.

PERCEPTUAL CONSCIOUSNESS

For analyses of perception in the sense of perceptual consciousness, see H. H. Price, *Perception* (London: Methuen, 1954) on the sense-datum theory; Durant Drake and others, *Essays in Critical Realism* (London: Macmillan, 1920), and Roy Wood Sellars, *The Philosophy of Physical Realism* (New York: Macmillan, 1932), on critical realism; Brand Blanshard, *The Nature of Thought*, Vol. I (London: Allen and Unwin, 1939), on idealism; and Roderick M. Chisholm, *Perceiving: A Philosophical Study* (Ithaca, NY: Cornell University Press, 1957), which is a causal analysis with a general consideration of perceiving and epistemological questions. Roderick Firth, "Sense-Data and the Percept Theory," in *Mind* 58 (232) (1949): 434–465, and 59 (233) (1950): 35–36, is critical of any such analyses.

CONCEPTUAL ANALYSIS

For the approach by conceptual analysis see "On Seeing and Hearing" by Winston H. F. Barnes, "Identification and Existence" by Stuart Hampshire, and "Sensation" by Gilbert Ryle in *Contemporary British Philosophy*, Vol. III, edited by H. D. Lewis (London: Allen and Unwin, 1956); Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), Ch. 7, and *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954); Anthony M. Quinton, "The Problem of Perception," in *Mind* 64 (253) (1955): 28–51; and D. W. Hamlyn's discussion of psychological theories, *The Psychology of Perception* (London: Routledge and Paul, 1957). John Langshaw Austin's *Sense and Sensibilia* (Oxford: Clarendon Press, 1962) is an outstanding defense of common sense against the argument from illusion, but unfortunately he does not discuss the scientific evidence. Another direct-realist study is D. M. Armstrong's *Perception and the Physical World* (New York: Humanities Press, 1961).

SCIENTIFIC BACKGROUND

Useful elementary accounts of the causal and psychological processes are given in such textbooks as Ernest Ropiequet Hilgard, *Introduction to Psychology* (New York: Harcourt Brace, 1953); David Krech and Richard Stanley Crutchfield, *Elements of Psychology*, 3rd ed. (New York: Knopf, 1962); and Robert Sessions Woodworth and D. G. Marquis, *Psychology* (New York: Holt, 1947; London: Methuen, 1949). Rather more detail, with a philosophical discussion as well, is given in George McCreath Wyburn, Ralph William Pickford, and R. J. Hirst, *Human Senses and Perception* (Edinburgh: Oliver and Boyd, 1964). Two significant but unconventional psychological books that raise philosophical questions of interest to the advanced student are James Jerome Gibson, *The Perception of the Visual World* (Boston: Houghton Mifflin, 1950), and Friedrich August von Hayek,

The Sensory Order (Chicago: University of Chicago Press, 1952; London: Routledge, 1953).

INTERPRETATIONS OF SCIENTIFIC EVIDENCE

See the works by Price, Chisholm, Hirst and the critical realists mentioned above and Arthur O. Lovejoy, *The Revolt against Dualism* (La Salle, IL: Open Court, 1930), also a critical realist. C. D. Broad's *Perception, Physics and Reality* (Cambridge, U.K.: Cambridge University Press, 1914) and *Scientific Thought* (London: Kegan Paul, 1923), sense-datum analyses, are very technical at times. William Pepperell Montague, *The Ways of Knowing* (New York: Macmillan, 1925), gives a general comparison of the different realist (and other) views. See also the more sophisticated types of representative realism—namely, John Raymond Smythies, *Analysis of Perception* (London: Routledge and Paul, 1956), or Bertrand Russell, *Analysis of Matter* (London: Kegan Paul, 1927) and *Human Knowledge* (London: Allen and Unwin, 1948). Of Russell's two books, the first is often difficult, and the second is more for the general reader.

OTHER RECOMMENDED TITLES

Alston, William P. *The Reliability of Sense Perception*. Ithaca, NY: Cornell University Press, 1993.

Armstrong, D. M. *Belief, Truth, and Knowledge*. Cambridge, U.K.: Cambridge University Press, 1973.

Ayer, A. J. *The Foundations of Empirical Knowledge*. London: Macmillan, 1940.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Brewer, Bill. *Perception and Reason*. Oxford: Clarendon Press, 2002.

Burnyeat, Myles. "Conflicting Appearances." *Proceedings of the British Academy* 65 (1979): 69–111.

Dancy, J., ed. *Perceptual Knowledge*. Oxford: Oxford University Press, 1988.

Dicker, Georges. *Perceptual Knowledge*. Dordrecht: Reidel, 1980.

Dretske, Fred. *Seeing and Knowing*. London: Routledge and Kegan Paul, 1969.

Fales, Evan. *A Defense of the Given*. Lanham, MD: Rowman & Littlefield, 1996.

Fumerton, Richard. *Metaphysical and Epistemological Problems of Perception*. Lincoln: University of Nebraska Press, 1985.

Goldman, Alvin. "Perceptual Knowledge and Discrimination." *Journal of Philosophy* 73 (1976): 771–791.

Grice, H. P. "The Causal Theory of Perception." *PAS*, Supp. 35 (1961): 121–168.

Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.

Perkins, M. *Sensing the World*. Indianapolis: Hackett, 1983.

Sellars, Wilfrid. *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Press, 1991.

R. J. Hirst (1967)

Bibliography updated by Benjamin Fiedor (2005)

PERCEPTION, CONTEMPORARY VIEWS

Philosophical accounts of perception aim to give a coherent and systematic account of the nature of our sensory experiences. Philosophical accounts differ from scientific ones, which aim at explaining how the specific mechanisms of perception work. Philosophers are interested in general features that are common to anything that we might reasonably call perception, abstracting away from the specific mechanisms by which we perceive the world. Contemporary theorists of perception have proposed theories aimed at addressing a number of questions about perception, including the following: What accounts for the distinctive feel of our sensory experiences? Is perception a representational state with specific content (like desires and beliefs)? Is perception a "direct" awareness of the world? How does perception make possible beliefs and thoughts about the world? How do perceptions serve as reasons for belief, making possible knowledge of the world?

APPEARANCE, REALITY, AND PHENOMENAL CHARACTER

One main source of philosophical puzzlement that has persisted since ancient times is the distinction between appearance and reality. To see the distinction, consider an example in which you see a ripe tomato sitting on a well-lit table. Assuming your eyesight is good, the tomato will appear a certain way to you; for example, it may appear red and round. This is a case of what we will call veridical perception. The tomato appears red and round to you, and in reality it is that way. It is, of course, also possible to misperceive, in which case the way things appear will not match the way they are. For example, if the tomato is in unusual lighting, it might appear to be purple rather than red. Likewise, if you are wearing shape-distorting glasses, the tomato might appear to be tall and skinny rather than short and plump, as it really is. These are cases of illusion, which involve objects appearing to you to have properties other than the ones that they have in reality. A second kind of misperception, distinct from illusion, is hallucination. Hallucinations are experiences in which it appears to you as if an object with certain properties is present, when in reality you are not in perceptual contact with any such object. For example, it might appear to you as if there is a red and round tomato before you when in fact there is no object there at all.

One problem that the possibility of misperception raises is epistemic and has to do with whether we are able

to know things about the world. Skeptics about knowledge of the external world have held that, in order for you to have knowledge of the world, you must be able to rule out the possibility that you are now misperceiving. But, these skeptics claim, there are certain possibilities of radical misperception that you cannot properly rule out—for example, you can't rule out the possibility that you are right now dreaming, or the possibility that you are really a brain in a vat being fed experiences as of the external world by an evil superscientist who is directly stimulating your brain. Defenders of the common sense idea that we have perceptual knowledge attempt to reply to the skeptic's challenge.

As we will see, the possibility of misperception also provides a challenge for metaphysical accounts of the nature of perceptual appearances. The challenge arises in part because giving a theory of the nature of appearances requires accounting for what is sometimes called the phenomenal character of experience, or, more simply, the phenomenology of experience. The phenomenal character of a cognitive episode is, in Thomas Nagel's famous phrase, "what it is like" to undergo it. A feature unique to conscious states is that there is something it is like to be in them. There is, for example, a way it is like for one to see a tomato.

The phenomenal character of perceptual experiences seems to be a crucial part of what distinguishes such experiences from other conscious mental events such as occurrent thoughts, desires, and beliefs. For example, what it is like for you to think about a tomato that is in front of you with your eyes closed will be very different from what it is like to open your eyes and *see* the tomato. Seeing a tomato has a sensory, visual phenomenology that merely thinking about a tomato lacks. Although nonperceptual mental states like beliefs and desires arguably have a phenomenal character (for example, there is presumably something it is like for you to think about mathematical sums while in a sensory-deprivation tank with no perceptual experience at all), the phenomenal character of perceptual awareness is distinctive.

It may be that not all perceptions are conscious and so have a phenomenal character. It is common in psychology to distinguish between unconscious and conscious perceptions, and there is a growing psychological literature suggesting that much of the perceptual information that guides our actions is not conscious. (A good introduction to the psychological evidence is in Melvin Goodale's *Sight Unseen*.) It is a question of considerable philosophical interest what unconscious perception is and how to distinguish it from conscious perception. Nevertheless, we will focus here on theories of perception that seek to give an account of conscious perceptual experience.

There are several aspects of the phenomenal character of perceptual experience that philosophers have thought need to be reflected in a philosophical account of perception. First, there are differences in phenomenal character at the level of the different sense modalities. For example, what it is like to see a tomato is different from what it is like to taste, touch, or smell it. Each mode of perceptual awareness—vision, taste, touch, smell, and audition—has its own distinctive sensory phenomenal character.

Second, there are similarities and differences in phenomenal character at the level of experiences within a sensory modality. For example, a tomato might appear to be red, another might appear to be green, and a third might appear to be very similar in color to the red one. Philosophers are also interested in the way that experiences from different perspectives give rise to differences in phenomenal character, even when there is no change in the way objects appear to be. For example, looking at the tomato from different angles or from nearer or farther away yields differences in the appearances, even though all of these experiences are arguably veridical perceptions and there is no change in the way the object appears to be. When viewed from close up, the tomato in a meaningful sense "appears larger" than when one looks at it from afar. Or, to take another example, it may be that part of the surface of a tomato "appears white" owing to the way the light is reflecting off its surface, even though in another sense the tomato appears uniformly red. These observations suggest a distinction between what we will call constant and perspectival modes of appearance talk. In the constant mode, saying that "an object appears so-and-so" implies that if you are not subject to an illusion, then the object *is* so-and-so. But this is not the case in the perspectival mode; it can be the case that "an object appears so-and-so" and that you are not subject to an illusion, while *not* being the case that the object really is so-and-so. For example, when you see the highlight on the tomato, it is correct to say that patch of the tomato "appears white" in the perspectival sense of appearance talk, but also correct to say that it "appears (to be) red" in the constant mode.

Since what we want in an account of perceptual experience is an account of perceptual consciousness, a correct theory of sense perception must be phenomenally adequate; it must do justice to the phenomenal character of experience.

ARGUMENTS FROM ILLUSION

A problem that divides philosophers of perception is how to account for the phenomenal character of experience while at the same time explaining the possibility of misperception. To see how the problem arises, consider a theory of perceptual experience that some philosophers have dubbed Naive Realism. As its name suggests, Naive Realism tries to take what is seen as a prereflective account of perception and use it as a philosophical theory of perception. According to them, perceptual consciousness is, in its fundamental nature, a relation of direct awareness between a perceiver and public objects and their properties. Moreover, it is these properties and objects of which one is aware that explain the phenomenal character of experience. Consider again our case of a tomato's visually appearing to you to be red and round. What explains the phenomenology of such an experience? The Naive Realist thinks that common sense is clear about what explains this: It is the tomato itself and the qualities of it presented to awareness that constitute what it is like to see the tomato. To explain what it is like to have a perceptual experience, we simply need to describe the objects that appear to you and their properties of which you are aware.

One challenge for Naive Realism is to explain differences in the phenomenal character of appearances described in the perspectival mode. For example, we saw that in the perspectival mode it is correct to say that the tomato viewed from afar appears smaller than when viewed from close up. Yet in both viewings of the tomato, it seems reasonable to suppose that you veridically perceive the size of the tomato, a property of the tomato that does not change. (This is why it is correct to say in the constant mode that whether the tomato is viewed from up close or from afar it appears to be the same size, say, roughly the size of your closed fist.) It seems, then, that what explains the difference in phenomenal character of these two viewings is not a property of the tomato, as the Naive Realist supposes.

Perhaps an even more difficult problem for the Naive Realist arises from the possibility of misperception. When you see the ripe tomato and your experience is veridical, you are in a perceptual state that we can describe by saying that "it appears to you as if there is something red and round before you." But it seems entirely possible for the very same type of state described in this way (complete with its distinctive phenomenal character) to occur as part of illusory or hallucinatory experience. For example, if you were wearing shape- and color-distorting glasses, it might be that what is in reality a tall, oblong, purple thing looks to you just like a plump, red tomato. This illusory experience might have the same phenomenal character as your veridical perception of a red, round tomato. The problem for the Naive Realist is that it *cannot be* in this case that the real color and shape properties of the thing you are seeing are what explain what it is like to see the object. The thing you are seeing is tall and purple, whereas your experience is as of something red and round.

The possibility of hallucination raises an exactly similar problem for Naive Realists. Consider a case in which you have a hallucination of a tomato when there is not one anywhere nearby. To fill out the case a bit, we might imagine that a futuristic superscientist stimulates your visual cortex in just the same way that it is stimulated when you see a tomato and thereby produces in you an experience that is every bit as vivid as a veridical perception of a tomato. If this were the case, it obviously can't be true that what explains the phenomenal character of your experience is a direct awareness of a real tomato that is red and round. In the case as described, there is not even a tomato there!

Although we have been focusing on specific visual examples involving seeing tomatoes, there is nothing special about our choice of examples. For *any* veridical perception that we could describe as one in which "it perceptually appears to you as if such-and-such is the case," it seems possible for you to be in a state with the very same phenomenal character that is an illusion or a hallucination. The problem for the Naive Realist is that they don't have the resources to explain the phenomenal character of these states, since their account at best only explains the phenomenal character of veridical perceptions.

The considerations here are related to a family of arguments that were commonly referred to in the twentieth century as "the argument from illusion." As might be apparent from our discussion, we can actually distinguish among arguments from perspective, illusion, and hallucination, depending on which of these phenomena is under consideration. Further on we will consider how different theorists propose to answer these problems, including responses on behalf of those who want to defend Naive Realism from the objections.

SENSE-DATA THEORY

One historically important answer to the problems of perspective, illusion, and hallucination is that of the Sense-Data Theory. The theory is not as commonly held among contemporary theorists as it was among philoso-

phers in the early twentieth century (such as G. E. Moore, H. H. Price, and C. D. Broad), but it still has a few defenders today (for example, Howard Robinson). According to Sense-Data Theorists, perception involves an immediate awareness of mental "Sense-Data," which are taken to be objects such that awareness of them fully determines their existence and nature.

Its proponents offer the Sense-Data Theory as the best explanation of the perspectival character of appearances, and of the possibility of misperception. Consider again our example of seeing the tomato. We saw that one challenge to Naive Realists is to answer questions about the perspectival character of appearances like this one: Why is it that looking at a tomato up close results in an experience that can be described in the perspectival mode as one in which the tomato appears larger than it does when you are standing far away from it? The Sense-Data Theorist will answer that this is because in the former case you are aware of a sense datum that *really is* larger than the sense datum you are aware of when you look at the tomato from afar. An advocate of the arguments from illusion and hallucination against Naive Realist might also ask this question: How is it, then, that a state with a single phenomenal character—for example, a state in which it seems to you as if there is a red, round tomato before you—could occur either in a veridical perception or in a hallucination or in an illusion? The Sense-Data Theorist's answer is that the veridical perception, illusion, and hallucination all involve your being directly aware of sense data with the same properties, for example, sense data that are red and round.

According to the Sense-Data Theory, one is aware of objects and properties in the world only indirectly, in virtue of a more direct awareness of sense data and their properties. One challenge for the Sense-Data Theorist is to explain how sense data must be related to the world in order for one to perceive the world (albeit indirectly). For example, a Sense-Data Theorist owes us an answer to the following question: What makes it the case when you veridically perceive a tomato that being directly aware of a red, round sense datum counts as perceiving the real-world tomato? One possible reply would be that in order to perceive the tomato, you must be aware of a sense datum that has properties that resemble (or are isomorphic to) the properties of the tomato in the world. But this cannot be quite right. You can perceive a tomato even when your experience is a radical illusion such that you misperceive all of the tomatoes properties.

For example, if you look at the tomato through shape and color distorting lenses, the sense datum of which you are aware will not match the tomato in any of its shape or color properties (for example, the sense datum might be purple and tall while the tomato is short and red). But it might still be true that you see the tomato, even though you misperceive its properties. A second reply on behalf of the Sense-Data Theorist might be that you see the tomato if and only if the tomato causes the sense data of which you are aware. This proposal faces the problem that there are many different causes of the sense datum that don't count as things that you see. For example, the image on your retina is one of the causes of your perceptual experience (and its properties even seem to resemble the qualities of the sense data of which you are aware). But you do not see the images on your retina. Only eye doctors who are looking inside your eyes see retinal images. It seems that an object must cause an experience in the "right way" in order for the subject to perceive the object. It is a difficult problem, though, to say what this right way is.

Sense-data theories have been subject to many other objections. Arguments from illusion to the existence of sense-data theories have been criticized on grounds that they illicitly rely on a general principle of the following form: If it appears to you as if something has a certain property, then you are aware of something that really does have that property. Relying on this claim is sometimes referred to as the "sense-datum fallacy." The assumption has been thought to lead to absurd conclusions, like the conclusion that when an antique vase appears ancient and cracked to me, there is a sense datum that really is ancient and cracked. However, this conclusion might be blocked by restricting the properties mentioned in the principle to perceptible properties, such as color and shape. Moreover, the arguments from illusion, hallucination, and perspective should perhaps best be thought of as inferences to the best explanation. On this way of construing the arguments, the Sense-Data Theorist claims that postulating sense data offers the best explanation of the possibility of phenomenally identical illusions and hallucinations, and offers the best account of the perspectival nature of experience.

Other common objections to Sense-Data Theory allege that the view leads to skepticism, setting up a problematic epistemic "veil of perception" between the world and us, or that sense data are not scientifically respectable because they do not seem to be the sorts of things that fit easily into a physical picture of the world. In recent years, perhaps the most common objection to sense-data theories arises from a point about the phenomenal character of experience. Philosophers such as Gilbert Harman (1990/1997) and Michael Tye (1995) have claimed that

there is a tension between Sense-Data Theory and what is sometimes metaphorically referred to as the "transparent" or "diaphanous" nature of experience. The idea that experience is transparent is the idea that perception, and in particular visual perception, seems on the face of it to be a direct presentation of objects and properties as they are in themselves, and does not seem to involve an awareness of subjective properties and objects that represent objects in the world, as the Sense-Data Theory suggests. In perception, it is argued, we seem to be aware only of public properties and objects. For example, philosophers who think experience is transparent will say that when you see the tomato and reflect on your experience, the only properties that you will seem to be aware of are the public properties of the tomato. As rendered by the transparency metaphor, experience doesn't seem to be an opaque object that we know to be related in some way to the external world, as we might expect if the Sense-Data Theory were true; rather, experience seems "transparent," and the world and its properties (metaphorically speaking) *shine* through it.

INTENTIONALIST THEORIES

Philosophers such as Gilbert Harman (1990/1997), Michael Tye (1995), and Fred Dretske (1995), have suggested that by treating experience as an intentional state we can account for the transparency of experience while agreeing with the Sense-Data Theorist that there is a common kind of state involved in veridical perception that could also occur in illusion or hallucination. Intentional states are those with representational contents that can be correct or incorrect. A familiar example is belief. To believe that there is a tomato on the table, for example, is to be in a state that has a representational content—namely the content *There is a tomato on the table*. This content can be correct or incorrect depending on whether there is in fact a tomato on the table.

Intentionalists claim that experience is like belief in being a state that represents the world as being some way or other, and they hold that the representational content of experience fully explains its phenomenal character. (Sometimes this claim of Intentionalists is put in terms of what is called a "supervenience claim": phenomenal properties supervene on intentional content, i.e., there can be no change in phenomenal qualities without a change in the intentional content of experience.) When it appears to you as if there is a tomato before you, for instance, you are in a state that represents certain properties typical of tomatoes (for example, being round, red, and so on). According to Intentionalists, the way in which the world is represented explains the phenomenal character of the experience. Moreover, the same experience could occur in a misperception. The experience is correct if there really is a tomato with those properties before me. It is illusory if there is an object there, but it isn't red or round. The experience is hallucinatory if there is no object there at all.

Intentionalists accommodate the transparency of perceptual experience by claiming that, even though perceptual experience involves a state that represents, introspection is open only to the properties and objects represented by the experience, all of which are taken by Intentionalist theorists to be external properties and objects. The way objects are represented in perceptual experiences is consequently not like the way in which objects are represented when one looks at a photograph of them. When one looks at a photograph of one's grandmother, one is aware of some of the features on the film in virtue of which the photograph represents Grandma (for example, the colors and shapes on the surface of the film). Perceptual experience is more like conceptual thought, at least thoughts that do not employ mental imagery. When one thinks about one's grandmother (supposing one doesn't use a bit of mental imagery to do so), one is not aware of the properties in virtue of which one's thought is about one's grandmother. One is simply aware of the represented object, one's grandmother. Likewise, according to Intentionalists, when one *sees* one's grandmother, one is not aware of the properties in virtue of which one's experience is representing grandmother; one is only aware of what is represented—Grandmother and her properties.

Some early versions of Intentionalism claimed that perception is not merely *similar* to belief, it is in fact a kind of belief. (This was, for example, David Armstrong's view in his book *Perception and the Physical World* [1961]) However, such a view faces serious objections. A noncontroversial way of showing that experiences are not beliefs is to note that experiences are not revisable in light of counterevidence in the way that beliefs are. For example, one might believe that one's current experience is illusory or hallucinatory. If one has good enough reason to believe this, one can fail to believe the evidence of one's senses, even though the perceptual experience, complete with its phenomenal character, will remain intact.

A related question that arises for those who hold that perception is not a kind of belief is whether experience is like belief insofar as it essentially involves a deployment of concepts. Some philosophers of perception have propounded Conceptualism, the view that every sensory ele-

ment of perception involves an exercise of concepts by the perceiver. Conceptualism is often held on the ground that the only way that a state can serve as a reason for belief is if the state is conceptual through and through. Conceptualism is defended in this way by Bill Brewer (1999) and John McDowell (1994), although both argue for the position in the context of defending Disjunctivism (a view explained below) rather than Intentionalism. Some theorists object to Conceptualism on the grounds that animals or small children can perceive the world even though they lack concepts that would allow them to form beliefs about the world. Others object to Conceptualism on the grounds that the fine-grained phenomenal character of experience suggests that experience has "nonconceptual content." These philosophers suggest that the complexity and specificity of the properties and objects that you see in a single glance outstrip your conceptual capacity to form conceptual thoughts about these objects and properties.

Several potential objections to Intentionalist theories have been raised in the philosophical literature. One challenge for Intentionalists is the same as a challenge raised above for Sense-Data Theorists, namely to give an account as to how an object must be related to perceptual experience in order for the experience to be a perception of the object. It has seemed to most Intentionalists that the answer to this question involves an object's causing the experience in "the right sort of way." (For example, your experience as of a tomato must be caused in the right way by a tomato in order for you to see a tomato.) But it is difficult to say what this "right sort of way" is.

Quite a few philosophers have objected that Intentionalism lacks the resources to explain what is distinctively sensory about the phenomenal character of experiences. This general objection is pressed in a variety of ways. Some philosophers (such as Christopher Peacocke 2001) have challenged Intentionalists to provide an account of facts about appearances described in the perspectival mode, such as the way the tomato appears smaller when one moves further away from it. Other theorists attack the alleged transparency of experience by citing examples of what they claim are experiences that do not seem to be about public objects or properties. In some examples of perceptual experience, these philosophers claim, we seem to be aware of objects or properties that are essentially private and depend on our awareness of them. Proposed examples include experiences involving afterimages, double vision, blurred vision, and the "inner light show" that one experiences when one shuts one's eyes tightly.

Still other philosophers have objected that Intentionalists cannot explain the difference in phenomenal character between perception and other intentional states such as thinking. Earlier it was suggested that the phenomenal character of seeing a tomato is very different from merely thinking about the tomato. But both seem to be intentional states, and it seems that they might have the very same content—for example, the content *There is a red and round tomato on the table*. A challenge for the Intentionalist is to explain the difference between these two states. Some Intentionalists have suggested that the difference can be explained because perceptual experience is nonconceptual and plays a distinctive role in relation to beliefs and desires. A related challenge for Intentionalists is to distinguish between the phenomenal character of experiences in different modalities. For example, one can both feel the roundness of a tomato and also see the roundness. These states both represent the same property, the roundness of the tomato, so the Intentionalist might seem to be committed to thinking that the phenomenal character is the same. But of course the phenomenal character of the states is quite different.

Those who find the foregoing objections to Intentionalism compelling might still hold on to the idea that perception is an intentional state and that the content of the state *in part* explains the phenomenal character of experience. They will hold, however, that something in addition to the intentional content is required in order to account for the distinctively sensory phenomenal character of experience. Some philosophers (for example, Timothy Crane 1992) have suggested that in order to explain the phenomenal character of experience fully, we need to appeal not only to intentional contents but also to *modes of presentation* of those contents. For example, to explain the phenomenal character of your seeing the tomato we need to mention not only that you are in a state with the content that there is a red tomato before you, but also that this content is presented *visually*, rather than, say, tactilely. Others, such as Ned Block, suggest that we need to appeal to nonintentional properties of experience, sometimes called "qualia" in order to fully account for the phenomenal character of experience. This alternative is, in fact, consistent with the Sense-Data Theory. It is possible to develop a view according to which the perception of the tomato has an intentional content (for example, the content *There is something red and round before you*) that partly explains the phenomenal character of the experience, while also arguing for the need to postulate an awareness of a mental sense datum with certain properties in order to give a complete explanation of the phenomenal character of experience. (This seems to be a

view held by Christopher Peacocke in his book *Sense and Content*, though he speaks of awareness of "visual fields" rather than sense data.)

DISJUNCTIVISM

In recent years there has been a resurgence of attempts to defend Naive Realism by giving what is called a disjunctive account of experiences. Disjunctivists challenge the claim that for any veridical perceptual state of a subject (seeing a ripe tomato, for example), an event of the very same kind, individuated by its phenomenal character, could occur in a misperception. As stated earlier, one can describe the state of seeing the tomato as one in which "it appears to you as if there is something red and round before you," and this state can occur either in veridical perception, illusion, or hallucination. According to Disjunctivists, the state that we describe in this way is not a unified kind. The most that can be said about it is that this it is *either* (1) a state in which you are veridically perceiving a red and round tomato (in which case you are directly aware of the tomato and its properties) *or* (2) a state in which you are having a hallucination or an illusion that is indistinguishable from a veridical perception as of a tomato.

One might complain that so far, this is no theory at all, but at best a promise of one. The theory does not tell us anything, for example, about the phenomenal character of hallucinatory experiences. Disjunctivists, one might think, owe us an account of the phenomenal character of the "bad" side of the disjunct that involves hallucinatory experience. Many Disjunctivists resist the call to give a robust account of the phenomenal character of hallucinatory experience. For instance, Michael Martin, in "The Limits of Self Awareness" 2004, gives a purely epistemic characterization of hallucination. According to him, the most that can be said about the nature of hallucination is that it is indistinguishable from a genuine perception. For example, in the case where an advanced neuroscientist stimulates your visual cortex in exactly the way it is stimulated when you veridically perceive a tomato, Martin will say that the most fundamental thing we can say to explain the nature of this state is that this is a state such that you can't know purely on the basis of the experience that it isn't a genuine perception of a tomato. Many theorists, though, will think that the obvious explanation as to *why* your hallucination of a tomato can't be distinguished from a veridical perception is that the hallucination has a phenomenal character of a kind that requires a substantive metaphysical explanation—for example, the sort of explanation that Sense-Data Theorists and Intentionalists give.

Other Disjunctivists have made some tentative proposals for what accounts for the phenomenal character of hallucinatory states. Harold Langsam (1997), for example, says that it is possible to develop a theory according to which it is the physical regions of space around the subject where the object appears to be that are the relata of hallucination, and William Alston (1999) has suggested in passing that hallucination may involve an awareness of mental images. Such theorists face what might seem to be embarrassing questions that challenge their theoretical disunity. Given that their account of hallucinatory states fully explains the phenomenal character of experience, why not apply that same explanation to the case of veridical perception? Isn't it explanatory profligacy to rely on a disjunctive account when a unified one is available?

In response, Disjunctivists might counter that the explanatory cost of having an ununified view is well worth paying because alternative accounts of perception are subject to fatal flaws. In fact, a typical strategy of Disjunctivists has been to try to show that alternative theories of perception face insurmountable difficulties, leading to skepticism or making it mysterious how it is possible to think about the external world, or failing to do justice to the phenomenal character of experience.

See also Alston, William P.; Armstrong, David M.; Broad, Charlie Dunbar; Dretske, Fred; Harman, Gilbert; Illusions; McDowell, John; Moore, George Edward; Nagel, Thomas; Realism; Sensa.

Bibliography

Alston, William P. "Back to the Theory of Appearing." *Philosophical Perspectives* 13 (1999): 181–203

Armstrong, D. M. *Perception and the Physical World*. London: Routledge, 1961.

Block, Ned. "Mental Paint and Mental Latex." *Philosophical Issues* 7 (1996): 19–49.

Brewer, Bill. *Perception and Reason*. Oxford: Oxford University Press, 1999.

Crane, Tim. *The Contents of Experience*. New York: Cambridge University Press, 1992.

Dretske, Fred I. *Naturalizing the Mind*.Cambridge, MA: MIT Press, 1995.

Foster, John. *The Nature of Perception*. New York: Oxford University Press, 2000.

Goodale, Melvin, David Milner, and A. D. Milner. *Sight Unseen: An Exploration of Conscious and Unconscious Vision*. Oxford: Oxford University Press, 2004.

Harman, Gilbert. "The Intrinsic Quality of Experience." In *The Nature of Consciousness: Philosophical Debates*, edited by N.

J. Block, O. J. Flanagan, and G. Guzeldere. Cambridge, MA: MIT Press, 1997. First published in 1990.
Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.
Johnston, Mark. "The Obscure Object of Hallucination." *Philosophical Studies* 120 (2004): 113—183.
Langsam, Harold. "The Theory of Appearing Defended." *Philosophical Studies* 87 (1997): 33–59.
Martin, Michael G. F. "The Limits of Self-Awareness." *Philosophical Studies*. 120 (2004): 37–89.
McDowell, John. *Mind and World*. Cambridge, MA: Harvard University Press, 1994.
Peacocke, Christopher. "Does Perception Have a Nonconceptual Content?" *Journal of Philosophy* 98 (2001): 239–264.
Robinson, Howard. *Perception*. London; New York: Routledge, 1994.
Shoemaker, Sydney. "Introspection and Phenomenal Character." *Philosophical Topics* 28 (2000): 247–273.
Smith, A. D. *The Problem of Perception*. Cambridge, MA: Harvard University Press, 2002.
Swartz, Robert J. *Perceiving, Sensing, and Knowing*. Garden City, NY: Anchor Books, 1965.
Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind, Representation, and Mind*. Cambridge, MA: MIT Press, 1995.

Michael Pace (2005)

PERCEPTUAL CERTAINTY

See *Illusions*

PERCEPTUAL CONSCIOUSNESS

See *Perception*

PERFECTION

The concept of "perfection" has two closely allied and often overlapping meanings. First, it means "completeness," "wholeness," or "integrity": *X* is perfect when he (or it) is free from all deficiencies. Second, it means the achievement of an end or a goal. This meaning emerges most clearly from the connection between the Greek words *teleios* ("perfect") and *telos* ("end" or "goal"). An entity is perfect (to use Aristotelian terms) when it has achieved its goal by actualizing its potentialities and realizing its specific form. Bringing these two meanings together, one would say that a thing is complete or entire when it has fulfilled its nature and thereby reached its "end." The concept is best examined first under its religious, and second under its moral, aspect.

DIVINE PERFECTION

It has not always been believed that God (or, more generally, "the divine") is perfect. Thus, the deities of the Homeric pantheon were both ontologically and morally deficient. They differed from men only in being "deathless" (*athanatoi*). But in Christian theology the perfection of God has always been affirmed by orthodox writers. In St. Anselm's celebrated definition, God is *id quo nihil maius cogitari possit* ("that than which nothing greater can be conceived"). St. Thomas Aquinas later maintained that since God is self-existent, he must be infinite (or limitless) in intelligence, goodness, and power. He also claimed, in the fourth of his five Ways, to prove the existence of God as absolute perfection from the limited degrees of perfection in creatures. Thomists hold that by the "analogy of proportionality" we can attribute to God "in a more eminent way" (*eminentiori modo*) every "pure" perfection that exists in creatures (that is, every perfection that is capable of preexisting in an infinitely spiritual degree).

Those who hold this view of God's infinity must face two questions that have continually perplexed Christian philosophers. First, can we intelligibly assert that all perfections coexist infinitely in a single being? Thus, can God be both infinitely just and infinitely merciful? Second, if God is both infinitely powerful and infinitely good, how can we explain the presence of evil in the world?

MORAL PERFECTION

Ever since men began to reflect on the moral life, they have been aware of some perfect ideal of character and conduct toward which they must strive. Thus, in the Greco-Roman world the Stoics wrote copiously of the "perfect" (*teleios*) man. In their view perfection consisted in the subjugation of the passions to reason (*logos*) in a state of "self-sufficiency" (*autarkeia*). Sometimes they regarded moral virtue as the imitation of divine perfection, and sometimes they held out a human figure (especially Socrates) as the model of excellence; but more often they wrote abstractly of their ideal "wise man."

There can be no doubt that Jesus required moral perfection of those who would follow him. Thus, in the Sermon on the Mount, he told his disciples, "You, therefore, must be perfect, as your heavenly Father is perfect" (Matthew 5:48). In saying this Jesus reaffirmed the Old Testament, in which the Jews, as the people of the covenant, are required to be perfect (or "holy") by obedi-

ence to the law (*Torah*) which embodies God's will and reflects his character. The above-mentioned verses (Matthew 5:38–47) show that love, especially love of one's enemies, is the element in divine perfection that disciples are to imitate. Jesus' moral perfectionism was further expressed in his demands for complete inward purity (Matthew 5:21–22, 27–28) and self-renunciation (Mark 8:34–38).

Inevitably, theologians have affirmed that moral perfection is the goal of the Christian life. In the New Testament epistles perfection has three main characteristics. First, the norm of perfection is Christ himself, as the Incarnation of God. Second, the essence of perfection is love—the divine love revealed in Christ and made available to believers through the Spirit. Thus, St. Paul, having listed several virtues, wrote, "And above all these put on love, which is the bond of perfectness" (Colossians 3:14). Third, perfection is corporate. Thus, the author of Ephesians looks forward to the time when "we all attain to the unity of the faith and of the knowledge of the Son of God, to perfect manhood, to the measure of the stature of the fulness of Christ" (4:13). Postbiblical theologians (for example, St. Augustine and Thomas Aquinas) continued to give primacy to love, by which all the natural virtues are supernaturally perfected.

Two comments on this Christian scheme are relevant. First, as early as St. Ambrose there emerged a distinction between the basic "precepts" according to which all Christians were expected to live and the "counsels of perfection" that only a few ("the religious") could follow. This distinction, which persisted throughout the Middle Ages, was based on such texts as Matthew 19:16–22 and could be plausibly represented as an attempt to combine adherence to Christ's absolute demands with a realistic attitude toward the spiritual capacities of the average Christian in a secular occupation. But it was rejected by the Reformers, and with special vehemence by Martin Luther.

Second, although some Christians have held that it is possible to achieve perfection (that is, sinlessness) in this life, the majority have held that the strength of original sin makes this impossible. Moreover, many biblical texts (particularly I John 1:8–10) imply the Lutheran view that all Christians remain throughout their mortal lives *simul justi et peccatores* ("at the same time justified and sinners"). From a purely philosophical standpoint Immanuel Kant held that since the moral law requires holiness, and since we cannot achieve it in this life, we must postulate another life in which an infinite progress toward it will be possible (*Critique of Practical Reason,* translated by T. K. Abbott, London, 1909, p. 218).

Finally, if we take human perfection in its widest sense to mean an ideal that satisfies man's deepest needs or fulfills his "true" being, we can see clear points of similarity between Christian and non-Christian systems. Thus, although humanists, Buddhists, and Christians have in common many virtues that they regard as normative, they put them in differing contexts. These virtues are practiced by the humanist as self-sufficient ends, by the Buddhist as means of entrance to nirvana, and by the Christian as both the outcome of present faith in God and a preparation for a future vision of him "face to face."

See also Anselm, St.; Augustine, St.; Degrees of Perfection, Argument for the Existence of God; Thomas Aquinas, St.; Virtue and Vice.

Bibliography

Flew, R. N. *The Idea of Perfection in Christian Theology.* Oxford: Oxford University Press, 1934.

Kirk, Kenneth E. *The Vision of God.* London: Longmans, Green, 1931.

Niebuhr, Reinhold. *An Interpretation of Christian Ethics.* London: Harper, 1935. An important examination of Jesus' perfectionism and the problems that it raises.

Saunders, Kenneth. *The Ideals of East and West.* Cambridge, U.K.: Cambridge University Press, 1934.

H. P. Owen (1967)

PERFECTION, DEGREES OF

See *Degrees of Perfection, Argument for the Existence of God*

PERFORMATIVE THEORY OF TRUTH

Until relatively recently, it was taken for granted by all philosophers who wrote on the subject of truth, regardless of their differences on other matters, that words such as *true* and *false* were descriptive expressions. This presupposition has been challenged by P. F. Strawson, who developed the theory that "true" is primarily used as a performative expression. A performative utterance may be understood by considering a paradigm case: "I promise." To say "I promise" is not to make a statement about my promising but simply to promise. To use a performative expression is not to make a statement but to perform

an action. Strawson, in his essay "Truth," holds that to say that a statement is true is not to make a statement about a statement but to perform the act of agreeing with, accepting, or endorsing a statement. When one says "It's true that it's raining," one asserts no more than "It's raining." The function of "It's true that" is to agree with, accept, or endorse the statement that it's raining.

Strawson's performative analysis of "true" was conceived as a supplement to F. P. Ramsey's assertive redundancy, or "No Truth," theory of truth. Ramsey claimed that to say that a proposition is true means no more than to assert the proposition itself. "It is true that Caesar was murdered" means no more than "Caesar was murdered." "It is false that Caesar was murdered" means no more than "Caesar was not murdered." According to this view, "true" has no independent assertive meaning, and the traditional notion of truth as a property or relation is misguided. Ramsey suggested that "true" is used for purposes of emphasis or style, or to indicate the position of a statement in an argument.

CRITICISM OF SEMANTIC THEORY

Strawson set himself the positive task of explaining the use of "true" in ordinary language and criticizing the metalinguistic or semantic theory of truth, which has an affinity with Ramsey's view. Philosophers such as Rudolf Carnap, who hold the metalinguistic position, agree with Ramsey that to say that an assertion is true is not to make a further assertion. However, these philosophers claim that truth is a metalinguistic property of sentences, which means that to say that a statement is true is to make a statement about a sentence of a given language. According to this thesis, the statement that it's true that it's raining should, strictly speaking, be written: "'It's raining' is true in English."

Strawson argues that translation practice shows the metalinguistic thesis to be false. He points out that a translator would not handle a truth declaration as if it were a sentence description. Consider the manner in which a translator would handle a case where it is perfectly clear that one really is speaking about an English sentence:

(1) "It's raining" is a grammatical English sentence.

Suppose a translator wanted to translate (1) into a different language. He would retain the constituent "It's raining" in its original English, in order to show that (1) is a description of an English sentence. But consider

(2) It's true that it's raining.

There would be no hesitation in translating the whole statement, including the constituent "It's raining." This shows that (2) is not, as the metalinguistic thesis claims, a description of an English sentence. Hence, "true" is not a metalinguistic predicate.

Philosophers who maintain that "true" is a descriptive expression have been misled by grammatical form. "True" is a grammatical predicate, but it is not used to talk about anything. Strawson compares "true" with "Ditto." *A* makes an assertion. *B* says "Ditto." Insofar as *B* talks about or asserts anything, he talks about or asserts what *A* talked about or asserted. *A*'s assertion is the occasion for the use of "Ditto," but because "Ditto" is not composed of a grammatical subject and predicate, one is not tempted to think that in uttering "Ditto" *B* is making an additional statement.

The parallel with "Ditto" illuminates the tie between statements and "true." The making of a statement is the occasion for, but not the subject of, a truth declaration. "True" has no statement-making role. To say that a statement is true is to perform the act of agreeing with, accepting, endorsing, admitting, confirming, or granting that statement. Such expressions as "I grant …," "I confirm …," and "Yes" are perfectly capable of substituting for "The statement is true."

EXPRESSIVE USE OF "TRUE"

While Strawson emphasizes the performative role of "true," he also calls attention to another kind of use, which he calls expressive. This use is often found in sentences beginning "So, it's true that …," "Is it true that …," and "If it's true that …." In these utterances, "true" functions like the adverb "really," to express surprise, doubt, astonishment, or disbelief. However, "true" has only an expressive function in these utterances. It does not contribute, in either its expressive or its performative role, to the assertive meaning of what is said. Thus Strawson's thesis remains compatible with Ramsey's view. "True" does not change the assertive meaning of a statement. It has no statement-making role.

RESOLUTION OF "LIAR" PARADOX

The performative analysis of a truth declaration enabled Strawson to offer an original resolution of a well-known paradox that arises when one says:

(3) What I am now saying is false.

If (3) is true, then it is false; and if it is false, then it is true. Hence, we arrive at a paradox whose resolution has been one of the achievements of the metalinguistic analy-

sis of "true." According to this analysis, (3) is read in the following manner:

(3*a*) The object-language statement I am making now is false.

Since (3*a*) no longer refers to itself, the contradictory consequences disappear. Strawson dispenses with the metalinguistic solution and dissolves the paradox in a manner consistent with his own analysis of "true." To utter (3) is like saying "Ditto" when no one has spoken. It is not to make a statement but, rather, to produce a pointless utterance. Since (3) is not a statement, it is not a statement that implies its own denial. Hence, the paradox disappears without the necessity for metalinguistic machinery.

OBJECTIONS TO STRAWSON'S ANALYSIS

Strawson does not distinguish a truth declaration from such expressions as "I grant …," "I accept …," "I concede …," "I admit …," "I insist …," "Yes …," or "Ditto." It should be noted, however, that there are differences between using these expressions and saying that a statement is true. Expressions such as "I grant …," "I concede …," "I accept …," "I admit …," and "I insist …" suggest a "me versus you" background. They underline the act performed as *mine*. This is not the role of "That's true." Moreover, one should distinguish between expressions like "Yes," which simply register bare assent, and "The statement is true." If asked whether I agree with Smith's statement, I may say, "Yes, but my opinion isn't worth very much; I haven't studied the evidence." However, to say "His statement is *true*, but my opinion isn't worth very much; I haven't studied the evidence" sounds unnatural. "True," unlike "Yes," has the force of adequate evidence.

GEACH'S CRITICISM. P. T. Geach offered the following criticism of Strawson's analysis of "true" ("Ascriptivism," p. 233). Consider arguments of this pattern.

If *x* is true, then *p*;
x is true;
Ergo p.

Strawson claims that the second premise, "*x* is true," should be analyzed as an agreeing performance. However, it cannot be claimed that in the hypothetical premise "If *x* is true, then *p*," the constituent "*x* is true" is an agreeing performance. If I say, "If *x* is true, then *p*," I am not agreeing with or accepting *x*. Hence, the explanation of "true" in the hypothetical premise must differ from its explanation in the second categorical premise. However, if the explanation of "true" changes from one premise to another, the argument would be invalid, since the fallacy of equivocation has been committed. However, the argument is clearly valid. Hence, Strawson's analysis of "true," which implies that a different explanation is required for occurrences of "true" in hypothetical and categorical statements, must be wrong.

Geach's criticism, however, appears to rest on a misunderstanding of the behavior of performatives in logical arguments. Take a clear case of a performative, "I promise to help you." Now consider the following argument.

If I promise to help you, then I'm a fool;I promise to help you; *Ergo* I'm a fool.

There is a performative occurrence of "I promise" in the second premise, but not in the first. When I say, "If I promise to help you, then I'm a fool," I am not promising to help you. Hence, the use of "I promise" in the first hypothetical premise requires an explanation that differs from the explanation of "I promise" in the second hypothetical premise, yet the argument remains perfectly valid. A fallacy of equivocation is not committed simply because an expression has a performative use in one premise of a logical argument and a nonperformative use in another.

Occurrences of "true" in hypotheticals do not fit a performative analysis, but it must be remembered that while Strawson emphasizes the performative use, he does not claim that this is the whole story. The nonperformative use of "true" in hypothetical statements may be considered to fall under what Strawson calls the expressive use. What is the difference between the following statements?

(4) If Khrushchev's statement is true, there are no missile bases in Cuba.

(5) Khrushchev's statement implies there are no missile bases in Cuba.

While (4) and (5) have the same assertive meaning, (4) suggests that Khrushchev's statement is in doubt. Hence "true" in (4) contributes only to the expressive quality of the statement. Since "true" in (4) has only an expressive function, but not a statement-making role, (4) does not constitute an exception to Strawson's analysis.

"BLIND" USES OF "TRUE." An interesting challenge to Strawson's position is found in "blind" uses of "true." This use of "true" is exemplified when a person applies "true" to a statement without knowing what the statement is. For example, suppose a man says, "Everything the pope says is true." Presumably he does not know every state-

ment the pope has made. It cannot, therefore, be claimed that he is making the statements made by the pope. One cannot substitute the pope's statements for "Everything the pope says is true" without a change in meaning. Hence, "Everything the pope says is true" does not, as Strawson claims, have the same assertive meaning as the pope's statements. The notion, which Strawson takes over from Ramsey, that a truth declaration has the same assertive meaning as the statement dubbed true, does not hold for blind uses of "true."

It may be argued that the speaker is blindly endorsing all the pope's statements. In that case, "Everything the pope says is true" would be analyzed as a performative use of "true" which falls outside the range of Ramsey's thesis. But this analysis could not be maintained for blind uses like "I hope that what Jones says will be true." The speaker is plainly not endorsing what Jones will say. Moreover, since "true" in this case does not function like the adverb "really," it cannot be maintained that "I hope that what Jones will say is true" exemplifies an expressive use of "true" either. Hence, neither Strawson's nor Ramsey's position seems to hold up for blind uses of "true."

Strawson, however, has analyzed blind uses of "true" in what he takes to be a Ramsey-like method. In his later paper, "A Problem about Truth—A Reply to Mr. Warnock," Strawson shifts from his original position and grants that "at least part of what anyone does who says that a statement is true is to make a statement about a statement" (p. 69). This is a departure from his earlier view that "true" has no statement-making role. For the blind truth declaration "Everything the pope says is true," Strawson would offer the following Ramsey-like paraphrase: "Things are as the pope says they are." According to Strawson, this paraphrase is a statement about the pope's statements, but it also conforms to the spirit of Ramsey's view. Presumably, Strawson considers this analysis to be a Ramsey-like analysis because "true" is eliminated from the paraphrase. It must be remembered, however, that Ramsey held "true" to be eliminable because "true" is a "superfluous addition" to a statement ("Facts and Propositions," p. 17). Hence, one can always substitute *P* for "*P* is true" without loss of assertive meaning. While Strawson has eliminated "true" from "Everything the pope says is true" in the paraphrase "Things are as the pope says they are," he has not fulfilled Ramsey's claim that "true" is superfluous. A philosopher who holds the correspondence theory of truth can also eliminate "true" by substituting "Everything the pope says corresponds to the facts" for "Everything the pope says is true." However, this surely would not be a Ramsey-type elimination. Since "true" is not a superfluous addition to a blind truth declaration, it does not seem that blind uses can be paraphrased in the spirit of Ramsey.

See also Carnap, Rudolf; Paradigm-Case Argument; Performative Utterances; Pragmatism; Ramsey, Frank Plumpton; Semantics, History of; Strawson, Peter Frederick; Truth.

Bibliography

Austin, J. L. *How to Do Things with Words*. Cambridge, MA: Harvard University Press, 1962.

Bach, Kent, and R. M. Harnish. "How Performatives Really Work." *Linguistics and Philosophy* 15 (1992): 93–110.

Ezorsky, Gertrude. "Truth in Context." *Journal of Philosophy* 60 (5) (February 28, 1963).

Geach, P. T. "Ascriptivism." *Philosophical Review* 69 (April 1960).

Horwich, P., ed. *Theories of Truth*. New York: Dartmouth, 1994.

Kincade, J. "On the Performatory Theory of Truth." *Mind* 67 (1958).

Kirkham, Richard. *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press, 1992.

Miller, S. R. "Performatives." *Philosophical Studies* 45 (1984): 247–260.

Ramsey, F. P. "Facts and Propositions." *PAS*, Supp. 7 (1927). Reprinted in *Truth*, edited by George Pitcher, 16. Contemporary Perspectives in Philosophy Series. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Recanati, F. *Meaning and Force: The Pragmatics and Philosophy of Performative Utterances*. Cambridge, U.K.: Cambridge University Press, 1987.

Searle, J. "How Performatives Work." *Linguistics and Philosophy* 12 (1989): 535–558.

Searle, J. "Meaning and Speech Acts." *Philosophical Review* 71 (1962).

Sinnott-Armstrong, Walter. "The Truth of Performatives." *International Journal of Philosophical Studies* 2 (1) (1994): 99–107.

Strawson, P. F. "A Problem about Truth—A Reply to Mr. Warnock." In *Truth*, edited by George Pitcher, 68 (see above).

Strawson, P. F. "Truth." *Analysis* 9 (6) (1949). Reprinted in *Philosophy and Analysis*, edited by Margaret MacDonald, 260. Oxford, 1955.

Strawson, P. F., and J. L. Austin. "Truth." *PAS*, Supp. 24 (1950).

Walsh, W. H. "A Note on Truth." *Mind* 61 (1952).

Warnock, G. J. "A Problem about Truth." In *Truth*, edited by George Pitcher, 54 (see above).

White, Michael. "A Suggestion Regarding the Semantical Analysis of Performatives." *Dialectica* 30 (1976): 117–134.

Ziff, Paul. *Semantic Analysis*, 118. Ithaca, NY: Cornell University Press, 1960.

Gertrude Ezorsky (1967)
Bibliography updated by Benjamin Fiedor (2005)

PERFORMATIVE UTTERANCES

At the beginning of *How to Do Things with Words* (1962), John Langshaw Austin challenged the common assumption that "the business of [a declarative sentence] can only be to 'describe' some state of affairs, or to 'state some fact'" (p. 1). Obviously, that is not the business of interrogative and imperative sentences, but Austin argued that even certain declarative sentences are typically used to do something other than make statements. For example, an employer can fire someone by saying "You're fired," and an employee can quit by saying "I quit." In uttering such a sentence, one is not merely saying what one is doing, one is actually doing it. Such a sentence has a remarkable property: To utter it is (typically) to perform an act of the very sort named by its main verb.

It does seem remarkable that you can do something just by saying what you are doing. Most types of acts are not like that. You cannot stand on your head by saying that you are standing on your head, and you cannot convince someone that you love them by saying that you are convincing them that you love them. Yet in the right circumstances you can fire someone or quit a job just by uttering the right sort of sentence. How is this possible, and what sorts of acts can be performed in this way? Does this phenomenon of *performativity* require a special explanation, perhaps involving some kind of convention, or it is just a special case of something more general?

EXPLICIT PERFORMATIVE UTTERANCES

Austin (1961) dubbed *performative* such verbs as "promise," "apologize," "request," "fire," and "quit." Performative sentences are generally in the first-person singular with their main, performative verb in the simple present tense, active voice. So, for example, you can promise to attend by saying "I promise to attend" (but not by saying "I promised to attend" or "She promises to attend"), and you can apologize by saying "I apologize" (but not by saying "I apologized" or "She apologizes"). The word "hereby" may be inserted before the performative verb, thereby indicating that this utterance is the vehicle of the performance of the act named by the verb. Some performative sentences are in the first-person plural ("We guarantee your safety"), the second-person singular or plural ("You are advised to get vaccinated"), or the impersonal passive ("Smoking is prohibited)." Occasionally the performative verb is in the present progressive, as in "I'm warning you to stay away" and "I'm asking you for the last time to clean up your room." Because utterances of performative sentences are characteristically performances of acts of the very sort named by their main verbs, Austin called them "explicit performative utterances," or simply "performatives."

Notice that such acts as promising, apologizing, and requesting, which Austin called "illocutionary acts," can be performed without using a performative sentence, hence without making explicit what one is doing. For example, one can promise by saying "You can count on me to … ," apologize by saying "I'm sorry," and request by saying "I'd like you to …" This raises the question whether performativity, although involving the use of a special sort of sentence, requires a special explanation. In this regard note also that performative sentences do not have to be used performatively and obviously are not so used when they are embedded in larger linguistic contexts. For example, saying "If I promise to take you to the play, will you quit nagging me?" is not to make a promise, and saying "I apologize only if I feel guilty" is not to apologize.

PERFORMATIVES AND CONVENTIONS

It is generally accepted that linguistic meaning is a matter of convention. So to that extent every utterance is conventional, insofar it is made with linguistic means. However, it might seem, as it did to Austin, that performatives are conventional in a more specific way and that this explains their performativity. If so, then, for example, an utterance of "I promise to …" amounts to a promise because, and only because, there is a convention, or what John Searle (1969) called a "constitutive rule," to the effect that an utterance of such a sentence counts as a promise. That is, roughly, it counts as such only because it is generally recognized to count as such. This view seems plausible as regards certain institution-bound performatives, where a specific form of words is designated, and often required, for the performance of an act of a certain sort. For instance, uttering the words "I pronounce you husband and wife" counts (in the requisite circumstances) as the act of marrying a couple; uttering "The jury finds the defendant guilty" counts as finding the defendant guilty (convicting the defendant); and uttering "(I) double" counts as doubling in bridge. Indeed, in institutional contexts there are often designated expressions that, though not performative in form, have the same effect, such as an umpire's "Out," a legislator's "Nay," or a judge's "Overruled." Of course, these specialized performatives and other designated forms of words have to be uttered by the appropriate person in the appropriate circumstances, but

the relevant convention provides for this. Not just anyone can adjourn a meeting, sentence a convicted criminal, or christen a ship, and not just under any circumstances (with his "doctrine of the Infelicities" Austin classified the various ways in which such utterances can go wrong as "flaws," "hitches," and "abuses" [1962, pp. 12–38]). So it does seem that in institutional cases performativity is a matter of convention: A certain person's uttering a certain form of words in a certain context plays a certain official role because, and only because, it is generally recognized as so doing.

However, as P. F. Strawson (1964) contended, Austin was overly impressed with institution-bound cases. In such cases there are specific, socially recognized circumstances in which a person with specific, socially recognized authority may perform an act of a certain sort by uttering words of a certain form in order to effect, or officially affect, institutional states of affairs (see Bach and Harnish 1979, ch. 6). Ordinary performative utterances, on the other hand, are not bound to particular institutional contexts. Like most illocutionary acts, Strawson argued, they involve an intention not to conform to an institutional convention but to communicate something to an audience. An utterance counts as a promise, an apology, or a request because, and only because, the speaker intends it to count as such and the audience, recognizing that intention, regards it as such. To be sure, it is only under certain circumstances that a speaker will make such an utterance with such an intention and his audience will so regard it, but this is not in virtue of any convention.

It might be suggested, as it was by Jerrold Katz (1977), that performativity is explained not by social conventions but by linguistic ones. Perhaps there is some distinctive feature of the meaning of performative verbs that explains how one can perform an act of the very sort named by the verb by uttering a performative sentence containing that verb. However, this suggestion loses its plausibility when one takes into account a range of linguistic data beyond the simple performatives considered so far. In particular, there are what Bruce Fraser (1975) called "hedged performatives," which philosophers have largely overlooked, such as "I can promise you . . . ," "I must ask you . . . ," and "I would like to invite you . . ." Utterances of such sentences standardly have performative effect, but the meanings of the sentences themselves are not inherently performative. This is clear because without contradicting myself I could say "I can promise you, but I won't," "I must ask you, but if I did, my wife would never forgive me," or "I would like to invite you, but I can't." In each of these cases I would not be performing an act of the type in question but would merely be telling you that I am able to promise, that I am required to ask you, or that I would like to invite you. In addition, there are other sorts of sentences that, unlike hedged performatives, do not even contain performative verbs but which are standardly used in the same kind of way: "It would be nice if you . . ." to request, "Why don't you . . . ?" to advise, "Do you know . . . ?" to ask for information, "I'm sorry" to apologize, and "I wouldn't do that" to warn. Clearly these standard uses are not predictable from their linguistic meanings alone.

The variety of forms of sentences that are standardly used to perform acts of the same types as those accomplished by explicit performative utterances suggests that performativity is not a matter of convention, whether social or linguistic. Performativity requires no special explanation. Rather, its explanation belongs to the general theory of speech acts (see Searle 1989 and Bach and Harnish 1992 for two contrasting accounts). Performative sentences are just one kind among various kinds of sentences that are standardly used to perform types of illocutionary acts not predictable from their meanings alone (see Bach and Harnish 1979, ch. 10). Performativity is a pragmatic phenomenon not a semantic one, a matter of language use rather than linguistic meaning. The standardization of performative and other forms of sentences for uses not predictable from their meanings does not show that they are governed by special conventions but merely that there is a practice of using sentences of certain forms in certain ways. The claim that they are conventional falsely entails that an utterance of a certain form of words would not have the force it has unless it is generally recognized to count as such. The claim that they are merely standardized for these special uses requires something less. Standardization merely streamlines the inference the hearer must make to identify the speech act being performed; it creates the illusion of conventionality where there is really but a pragmatic regularity. (For further discussion of these issues, see Reimer 1995, Bach 1995, and Harnish 1997).

PERFORMATIVES AND STATEMENTS

When introducing the notion of performatives, Austin contrasted them with utterances like "I state that … ," "I claim that … ," and "I predict that …" These explicit *constatives* are like utterances of ordinary declarative sentences in that they "describe some state of affairs, or to state some fact," which Austin denied that performatives do. Yet he came to realize that explicit constatives are rel-

evantly similar to explicit performatives: Their main verbs also make explicit the type of act being performed. After all, an assertion or a prediction is made with "I assert …" or "I predict …" in just the same way that a promise or a request is made with "I promise…" or "I request …" Accordingly, what makes explicit performatives distinctive is not what the speaker does but that the speaker makes explicit what he or she is doing.

Austin also came to realize that what can be done explicitly without a performative can also be done without making explicit the type of act being performed. In the latter part of *How to Do Things with Words* he developed the distinction between *locutionary* and *illocutionary* acts, which effectively superseded the distinction between constative and performative utterances. Locutionary acts are acts of saying something, and illocutionary acts are performed in the act of saying something. This distinction applies not only to promises, requests, and apologies, but also to statements and the like (Austin retained the term "constative" for them).

For example, in uttering "I promise to be there" and thereby explicitly saying that one promises to be there or in uttering merely "I will be there" and thereby just saying that one will be there, one can promise to be there. Similarly, in uttering "I state that Mars has two moons" and thereby explicitly saying that one states that Mars has two moons or in uttering merely "Mars has two moons" and thereby just saying that Mars has two moons, one can state that Mars has two moons. Note that stating is distinct from saying. In the right circumstances, one might say that Mars has two moons but state, albeit figuratively, that a certain belligerent person has two obsequious functionaries. In general, a speaker need not make explicit what he or she is doing in order to do it. Explicit performatives do have a distinctive self-referential character, but that does not mean that their illocutionary force requires special explanation. Indeed, if the successful "performance of an illocutionary act involves the securing of uptake" (Austin 1962, p. 116), then if anything it should be easier for an explicit performative to succeed, precisely because the speaker is saying what he or she is doing.

One remaining question concerns whether performatives are statements too (see Bach 1975), contrary to Austin's insistence that making explicit "is not the same as stating or describing" (1962, p. 61). When he introduced the category of explicit performative utterances, he claimed that even though they are utterances of declarative sentences, they are not cases of making statements and are not descriptive. However, this does not seem right, for the simple reason that the verbs in performative sentences can be modified, as in "I gladly promise … ," "I sincerely apologize … ," and "I reluctantly request …" This strongly suggests that a speaker of such a sentence would be making a statement. The speaker would be describing himself or herself, as promising gladly, apologizing sincerely, or requesting reluctantly.

Performatives have even been described as "self-verifying" (originally by Lemmon 1962 and more recently by Johansson 2003). Clearly they are self-referential, in that if one utters a performative sentence and uses it performatively, one is making explicit what one is thereby doing. But to describe them as self-verifying is to claim that they make themselves true. This seems right, but notice that a performative is not self-verifying in the way that an utterance of, for example, "I am speaking" or "I am alive" is self-verifying. It is not the bare fact of the utterance that, given its content, makes it true. Suppose I utter "I hereby apologize" and thereby apologize. It is true that I am thereby apologizing, but what makes this true is that I am using the sentence to perform the illocutionary act of apologizing. In that way, it is self-verifying.

Does this self-referential, self-verifying character help explain performativity, is it just a curious feature of explicit performative utterances, or what? As Searle (1989) has argued, the performativity of performative utterances does not depend on their being self-verifying. That gets things backwards: they are self-verifying statements because of their performativity. However, as Kent Bach and Robert Harnish (1992) have argued, their character as statements plays a key role in the speaker's being able to communicate to his audience what he is doing, precisely because he is using a performative to make explicit what he is doing.

See also Austin, John Langshaw; Pragmatics.

Bibliography

Austin, J. L. *How to Do Things with Words*. Cambridge, MA: Harvard University Press, 1962.

Austin, J. L. "Performative Utterances." In his *Collected Papers*, edited by J. O. Urmson and G. J. Warnock. Oxford: Oxford University Press, 1961.

Bach, Kent. "Conventionalization vs. Standardization." *Linguistics and Philosophy* 18 (1995): 677–686.

Bach, Kent. "Performatives Are Statements Too." *Philosophical Studies* 28 (1975): 229–236.

Bach, Kent, and Robert M. Harnish. "How Performatives Really Work." *Linguistics and Philosophy* 15 (1992): 93–110.

Bach, Kent, and Robert M. Harnish. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press, 1979.

Fraser, Bruce. "Hedged Performatives." In *Syntax and Semantics*. Vol. 3, edited by Peter Cole and Jerry Morgan. New York: Academic Press, 1975.

Ginet, Carl. "Performativity." *Linguistics and Philosophy* 3 (1979): 245–265.

Grice, H. P. "Meaning." *Philosophical Review* 66 (1957): 377–388.

Harnish, Robert. M. "Performatives and Standardization: A Progress Report." *Linguistische Berichte* 8 (1997): 161–175.

Johansson, Ingvar. "Performatives and Antiperformatives." *Linguistics and Philosophy* 26 (2003): 661–702.

Katz, Jerrold J. *Propositional Structure and Illocutionary Force*. New York: Crowell, 1977.

Lemmon, E. J. "Sentences Verifiable by their Use." *Analysis* 12 (1962): 86–89.

Recanati, François. *Meaning and Force: The Pragmatics of Performative Utterances*. Cambridge, U.K.: Cambridge University Press, 1987.

Reimer, Marga. "Performative Utterances: A Reply to Bach and Harnish." *Linguistics and Philosophy* 18 (1995): 655–675.

Searle, John R. "How Performatives Work." *Linguistics and Philosophy* 12 (1989): 535–558.

Searle, John R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, U.K.: Cambridge University Press, 1969.

Searle, John R. "A Taxonomy of Illocutionary Acts." In *Language, Mind, and Knowledge*, edited by Keith Gunderson. Minneapolis: University of Minnesota Press, 1975.

Strawson, P. F. "Intention and Convention in Speech Acts." *Philosophical Review* 73 (1964): 439–460.

Kent Bach (2005)

PERGAMUM, SCHOOL OF

See *Neoplatonism*

PERIPATETICS

The original meaning of the word *peripatos* was "a covered walking place." The house that Theophrastus provided for the school of Aristotle contained such a *peripatos*. This yielded a proper name for the school itself—the Peripatos—and its members came to be known as "those from the Peripatos" or "Peripatetics." This derivation should be preferred to that previously current, according to which the term "Peripatetic" referred to a method of teaching while walking about, known to have been used by Protagoras, for example, and assumed to have been adopted by Aristotle. Although this view goes back to Hermippus at the end of the third century BCE, it is now generally regarded as a mistaken inference, based on nothing more than the name itself.

The history of the Peripatetics can be divided into two periods—that immediately following the death of Aristotle and that following the revival of interest in Aristotelian studies resulting from the edition of the treatises by Andronicus of Rhodes in the time of Marcus Tullius Cicero or a little later. When Theophrastus became president of the school in the year before Aristotle's death, he continued to show an interest in virtually the whole range of Aristotelian studies. But whereas it is now generally supposed that Aristotle retained a keen interest in metaphysical questions to the end of his life, it was the shift of emphasis away from Platonic otherworldliness to the phenomena of the world around us, a subject also found in Aristotle, which seems to have attracted Theophrastus most. Strato, Theophrastus's successor, made important developments in physical theory, transforming Aristotle's doctrine into a fairly full-blooded materialism. But after Strato's death about 269 BCE, his successors became almost exclusively concerned with questions about the content of the good life and the way to reach it, with questions of rhetoric, and with the distinctively Hellenistic interest in anecdote, gossip, and scandal. Many of the specifically Aristotelian doctrines were abandoned, and the school had become very much the same as a number of others in Athens by the end of the second century BCE.

The reasons for this disintegration are uncertain. It may be that the concentration of interest upon empirical questions discouraged speculation. Empiricism as such, however, has interested philosophers intensely at other periods of history. Some have supposed that the disintegration was part of a philosophic failure of nerve characteristic of the Hellenistic age as a whole. But this view of the Hellenistic age is probably incorrect, and in any case such a failure of nerve clearly applied less to Stoics, Epicureans, and Skeptics of the period than it did to the Peripatetics. Thus, their fate would remain unexplained.

It may be that the history of the Aristotelian writings had something to do with what happened to the Peripatetics. According to the well-known story, on Theophrastus' death his copies of Aristotle's writings went to Neleus of Scepsis in the Troad (Asia Minor). In one extreme view this meant that the Peripatetics in Athens thereafter had access only to the published works of Aristotle—namely, the dialogues. In fact, there seem to have been copies of at least some of the treatises available in Alexandria, in Rhodes, and probably in Athens throughout the Hellenistic period. They do not appear to have been much studied in the Peripatos, however, where knowledge of Aristotle came primarily from the writings of Theophrastus when not from the dialogues. Indeed, in

a sense the school of Aristotle might more correctly be called the school of Theophrastus. The weakness of its links with Aristotle's own thought may explain its relative failure in philosophy.

Andronicus of Rhodes wrote a special study on the order of Aristotle's works and published an edition of the treatises in the order in which they have survived to us. His edition is the source of all subsequent ones. Andronicus is sometimes dated as early as 70 BCE, but as Cicero never refers to his edition, it may not have been published until after Cicero's death in 43 BCE. Andronicus initiated a revival in Aristotelian studies, and the Peripatos flourished at least down to the time of Alexander of Aphrodisias (about 200 CE). Among those influenced by this revival were the geographer Ptolemy and the physician Galen. Alexander wrote important commentaries on the main Aristotelian treatises, and the tradition of writing such commentaries continued into the Byzantine period through such scholars as Themistius, Ammonius, and Simplicius, who must be classed as Platonists rather than as Aristotelians. All the commentators treated Aristotle's writings as a systematic corpus, and from the start all were influenced in varying degrees by both Stoic and Platonist doctrines.

The general approach, apart from certain unintended distortions, was intensely conservative. From time to time modifications of interest were proposed, however. The successor of Andronicus, Boëthius of Sidon (who is not to be confused with the earlier Stoic of the same name), rejected the doctrine that the universal is prior by nature to the particular and would not grant to form the title of primary substance. In so doing, he took a big step in the direction of medieval nominalism. The pseudo-Aristotelian treatise *De Mundo* is often regarded as a product of this period. It culminates in a theology in which a transcendent deity maintains order in the cosmos by the exercise of an undefined power, and in a general way the work has affinities with both Stoic writers like Posidonius and Neoplatonists. It seems, however, to imitate the Aristotle of the dialogues rather than the treatises, and it may antedate the edition of Andronicus.

See also Alexander of Aphrodisias; Aristotelianism; Aristotle; Cicero, Marcus Tullius; Empiricism; Epicureanism and the Epicurean School; Galen; Hellenistic Thought; Neoplatonism; Platonism and the Platonic Tradition; Posidonius; Protagoras of Abdera; Simplicius; Stoicism; Strato and Stratonism; Themistius; Theophrastus.

Bibliography

The earlier Peripatetics, fragments and testimonia, are in *Die Schule des Aristoteles*, edited by F. Wehrli, 10 parts (Basel, 1944–1959). See also P. Moraux, *Les listes anciennes des ouvrages d'Aristote* (Louvain, Belgium, 1951); I. Düring, *Aristotle in the Ancient Biographical Tradition* (Goteborg, Sweden, 1957), Part III, Chs. XVII and XVIII, and Part IV; and C. O. Brink, "Peripatos," in *Realencyclopädie der classischen Altertumswissenschaft*, edited by A. Pauly and G. Wissowa, Supp. Vol. VII (Stuttgart, 1940). See also the Aristotelian commentators in *Commentaria in Aristotelem Graeca*, 23 vols. and 3 supp. vols. (Berlin, 1882–1909). *De Mundo* is translated by D. J. Furley with Aristotle's *On Sophistical Refutations*; translated by E. S. Forster (Cambridge, MA, and London, 1955).

G. B. Kerferd (1967)

PERRY, RALPH BARTON

(1876–1957)

Ralph Barton Perry, the American realist philosopher, was born in Poultney, Vermont. He attended Princeton University, where he received his B.A. in 1896; he received his M.A. from Harvard in 1897 and his Ph.D. in 1899. For a brief period he taught at Williams and Smith colleges. From 1902 to 1946 he taught at Harvard, where, after 1930, he was the Edgar Pierce professor of philosophy. He was Hyde lecturer at various French universities during the year 1921–1922. In 1920 he was elected president of the eastern division of the American Philosophical Association, and he served as Gifford lecturer from 1946 to 1948.

Perry was the author of some two hundred essays and two dozen books, in addition to countless lectures and letters to newspapers, and he was considered the chief living authority on William James. Perry believed that a comprehensiveness of view is philosophy's contribution to human wisdom; in his own work he willingly risked inaccuracy to range over every province of science, art, philosophy, and religion. He insisted on the merit of this venture, insofar as it was an attempt to achieve systematic unity in a field that would otherwise be divided between experts who were unaware of one another's achievements.

REACTION AGAINST IDEALISM

As an early polemicist against idealism, Perry claimed that the relationship of the world to the mind is an accidental or subordinate aspect of the world. He argued that the relationship of knowing the world is not like the relationship of owning an object. An object owned becomes

in a sense a part of the owner, whereas the world, although it lends itself to being known, does not thereby become entirely a part of the knower. It is not exhaustively defined by the relationship of being known. This claim became one of the basic tenets of what Perry and five other young American philosophers formulated as New Realism in their cooperative volume *New Realism* (1912). They argued that the world is real and independent of mind, and that it is directly present or "immanent" to the mind in knowledge and consciousness. Together these tenets formed their "cardinal principle"—the "independence of the immanent."

In his article "The Ego-centric Predicament" (1910), Perry had shown how this "predicament" had been used illicitly to argue for idealism. The idealist argument begins with the predicament that "it is impossible for me to discover anything which is, when I discover it, undiscovered by me," and concludes that "it is impossible to discover anything that is not thought." The idealist, Perry claimed, has confused the statement that "everything which is known, is *known*" with the claim that "everything which *is,* is known." Perry maintained that the predicament was simply methodological: the extent to which knowledge conditions any situation in which it is present cannot be discovered by the simple and conclusive method of direct elimination.

Perry did not deny that this predicament presents a real difficulty, but he did deny that it argues either for idealism or realism. He never suggested what could be done to overcome the difficulty, but he did not think there were other than methodological implications in it. Instead, Perry argued that the objects of knowledge and experience are independent of egocentricity. "Independence" here refers not to a particular kind of relation but rather to the absence of one. Perry defined it as nondependence. The independent object may be related or not, provided that it is not related in the way the dependent object is. The independent object can be related to consciousness, or mind, but not be dependent on that relationship for its existence.

However, as Perry developed his position (in *Present Philosophical Tendencies,* 1925), it turned out that independent objects of knowledge are not the real independent objects of the commonsense world but "neutral entities" indifferent to both the subjective and the physical (or objective) relations in experience. They do not exist in any place; they exist only in the logical sense, as either a class or members of a class. They are therefore preeminently independent of consciousness. The propositions of logic and mathematics are typical of such entities, and Perry contended that analysis of such propositions reveals neither a knowing relation nor reference to a knower.

In taking this position, Perry had adopted James's neutral monism, and although he eventually abandoned it, he continued to describe his own philosophy as, among other things, "neutralism." Perry's move away from neutral monism and New Realism is best seen in his two works on value theory, *General Theory of Value* (1926) and *Realms of Value* (1954). The first work sets forth Perry's theory of the generic nature of value, while the second details the varieties and types of this value as they appear in the major human institutions, or "realms of value."

THEORY OF VALUE

Believing that value was neither unanalyzable nor purely emotive, Perry formulated his well-known definition, "Any object, whatever it be, acquires value when any interest, whatever it be, is taken in it." Value is that which attaches to any object of any interest. Interest is defined as that which is characteristic of the motor-affective life, namely, instinct, desire, feeling, will, and all their states, acts, and attitudes. A thing is an object of interest when its being expected induces actions that anticipate its realization or nonrealization. Interested action is thus actively selective, tentative, instrumental, prospective, and fallible.

According to Perry, this theory did not conflict with the "independence of the immanent," because the latter, being restricted to knowledge, did not demand that values be conceived as independent. Yet Perry's theory included a cognitive element in all value or interest. Cognition gives the interest its object, Perry said, and the character of the object of interest is essentially the same as that of the object of cognition. The "mediating judgment" in interest and cognition is expectation and belief, and without belief there would be no basis for truth and error. All interest is characterized by expectancy, but it differs from cognition in that it also includes being for or against, favoring or disfavoring, the expected. Since both interest and cognition have this element of expecting something and being prepared to cope with it, expectancy is the key to understanding both.

Because expectancy looks forward and does not disclose itself except through a train of subsequent events, the object of interest and of cognition can be conceived of only as an ideal or "problematic" object, possessing the ambiguity or dual possibility of truth and error. This object is "internal" to the act or cognition and must be distinguished from its "external" referent, that which con-

firms or fails to confirm the expectation of the problematic object. Expectation is the meaning of an object.

Perry pointed out that during the process by which a sensory stimulus leads to an eventual sense perception, not only muscles and nerves, but attitudes, meanings, and interpretations are oriented toward the stimulus. Thus, when the ear is assailed by a stimulus, the organism listens toward the source and acts, or prepares to act, both upon that source and upon its context. At this point a conversion takes place: one hears the sound there and then perceives it as a bell having further characteristics. Thus, a stimulus touches off a reaction, and then the stimulus is superseded by thought, which now has an object, although the original stimulus has ceased to exist. The stimulus has been converted into an object; the sound has been converted into a bell, or in other words, into what it means, what is expected of it. This is the "perceptual object," that part of the total surrounding field to which the organism alerts itself, embracing what is expected of the sensory object.

This object is characterized both by meaning—that is, by what the organism expects of it—and by being part of the surroundings. When Perry went on to describe its status further, his monistic bias became apparent. He maintained that if the ideal object is not somehow present in nature, it would be impossible to affirm that nature is as it is "represented" in the finished product of scientific inquiry. If the logical and mathematical structures of knowledge are to be true of nature, they must be *in* nature; the laws of nature reign in the realm of nature and not in the realm of natural science, which discovers them.

MORAL VALUE. Having offered his theory of value, Perry went on to show in what sense we can say one value is "better" or "worse" than another. This too, he thought, called for a definition—that is, a descriptive account of the meaning of "better" and "worse." For Perry, that meant a description of those conditions that would enable us to say with justification that one object of an interest was better (or worse) than another.

The key to this problem of value was integration or harmony of interests. To integrate or harmonize interests is to remove from them such qualities as independence, irrelevance, dissimilarity, opposition, indifference, antagonism, or incompatibility. Harmony in place of conflict is Perry's *summum bonum.* Morality takes the conflict of interests as its point of departure. What Perry called the moralization of life—the harmonizing of interests for the sake of the interests harmonized—is effected through "reflective agreement" between the personal and the social will. "Harmonious happiness" is justified by its provision for the several interests that it harmonizes. Ought and obligation, then, are not moral ultimates but are justified by the good end.

That Perry's moral criterion was an absolute in an otherwise nonabsolutistic theory did not occur to him. However, he did assert that the criterion must agree with human nature and the circumstances of human life in such a way that men can adopt it and be governed by it. It must also possess qualifications for being accepted in lieu of other standards. Perry thought his concept of harmony, in its appeal to each knower's will, did possess universality because it embraced all interests—that is, that it was to some extent applicable to everybody's interest.

The adequacy of Perry's theory rests therefore on his assumption that for all men "better" signifies a greater inclusiveness and harmony of values. Perry was by no means unaware of the need for social arrangements that would render the interests of individuals mutually innocent and cooperative. Almost half of his books were devoted to some aspect of this problem, and they were often written in response to the problems facing his country at the time. He brought to all of them his standard of harmonious happiness, or reflective agreement, a "creed of inclusiveness" that excluded only hatred and personal aggrandizement.

See also Ethics, History of; Idealism; James, William; Monism and Pluralism; New Realism; Realism; Value and Valuation.

Bibliography

WORKS BY PERRY

"The Ego-centric Predicament." In *The Development of American Philosophy,* edited by W. G. Muelder and L. Sears. Boston and New York: Houghton Mifflin, 1940. Reprinted from original article in *Journal of Philosophy* 7 (1) (1910): 5–14.

"A Realistic Theory of Independence." In *The New Realism,* by E. B. Holt et al., 99–151. New York: Macmillan, 1912.

The Present Conflict of Ideals; A Study of the Philosophical Background of the World War. New York: Longmans, Green, 1918.

Present Philosophical Tendencies; A Critical Survey of Naturalism, Idealism, Pragmatism and Realism Together with a Synopsis of the Philosophy of William James. New York, 1925.

General Theory of Value; Its Meaning and Basic Principles Construed in Terms of Interest. New York: Longmans, Green, 1926; reissued in 1950.

Philosophy of the Recent Past; An Outline of European and American Philosophy since 1860. New York and Chicago: Scribners, 1926.

"Realism in Retrospect." In *Contemporary American Philosophy,* edited by G. P. Adams and W. P. Montague, Vol. II, 187–209. New York: Macmillan, 1930.

The Thought and Character of William James, as Revealed in Unpublished Correspondence and Notes, Together with His Published Writings, 2 vols. Boston: Little Brown, 1935. A Pulitzer Prize biography.

Our Side Is Right. Cambridge, MA: Harvard University Press, 1942.

Puritanism and Democracy. New York: Vanguard Press, 1944. Best single statement of Perry's social and political philosophy.

The Citizen Decides: A Guide to Responsible Thinking in Time of Crisis. Bloomington: Indiana University Press, 1951.

Realms of Value; A Critique of Human Civilization. Cambridge, MA: Harvard University Press, 1954.

WORKS ON PERRY

Boman, Lars. *Criticism and Construction in the Philosophy of the American New Realism.* Stockholm, 1955. Mainly an exposition of Perry and other New Realists that utilizes the tools of modern philosophical analysis.

Harlow, Victor. *A Bibliography and Genetic Study of American Realism.* Oklahoma City: Harlow, 1931.

Hill, Thomas English. *Contemporary Theories of Knowledge.* New York: Ronald Press, 1961. Critical discussion of Perry as a "polemical" New Realist.

Thomas Robischon (1967)

PERSISTENCE

Smith is reading an open book that was shut this morning. At least it certainly seems like the same book he placed closed on his nightstand and opened to read this evening. But, then again, nothing can be both shut and not shut, Smith's book being included among those things that cannot violate G.W. Leibniz's law. So, no matter that common sense dictates that Smith's shut book did not blink out of existence to be instantaneously replaced by an open book, perhaps it is a different book after all.

Very roughly, that is the start of the problem of persistence—an initial worry about how an object can persist through a change in its properties. It is a problem that may seem easily dismissed until we identify its source in some of our basic metaphysical commitments and recognize the costs that accompany any way of addressing it. The understanding of the problem of persistence expressed below was developed alongside and informed by Sally Haslanger's work (Haslanger 2003).

THE INITIAL WORRY

We can sharpen the initial worry about books and other ordinary objects that persist through change by noting that it emerges from the conjunction of three core metaphysical theses.

THREE CORE METAPHYSICAL THESES

CONSISTENCY: Nothing can have incompatible properties.
CHANGE: Change involves incompatible properties.
PERSISTENCE: Objects persist through change.

The core theses express firmly held intuitions that most metaphysicians would agree are central to a coherent theory of how ordinary things—books, rocks, Smith, and even ourselves—exist and persist in the world. But, a commitment to any two of the theses seems to implicitly deny the remainder. Suppose PERSISTENCE and CHANGE are true, that some objects persist through change that involves incompatible properties. For instance, consider the book that Smith removed from his nightstand to read that was shut, and though open now, remains the same book. If we also assume CONSISTENCY is true, then nothing can have the incompatible properties of being shut and being open (given that a book is open if, and only if, it is not shut). Thus, it seems that the shut book from Smith's nightstand must be distinct from the open book in his hands. But that denies that the book persisted in the first place.

A careful reader will note that the contradiction was not precisely forced; nevertheless there is a significant tension that at least threatens contradiction. One strategy for responding to this worry is to bypass it by rejecting CONSISTENCY, PERSISTENCE, or CHANGE. A second strategy is to resolve the tension by first identifying its source and then clarifying or modifying our ideas to remove that source.

DISMISSING THE INITIAL WORRY

There are three options in pursuing the straightforward strategy of dismissing the initial worry about persistence by denying one of the core theses, none of which is promising. First, we might contend that something can both have and not have a property (forfeiting CONSISTENCY). However, such a move entails rejecting the law of noncontradiction, Aristotle's "most certain of all principles" according to which "the same attribute cannot at the same time belong and not belong to the same subject in the same respect" (Barnes 1984. Aristotle's *Metaphysics* IV.3.1005b1.17). But, countenancing contradictions to find a noncontradictory account of persistence makes no sense (though someone like Donald Baxter, 2001, might disagree). Indeed, such a drastic step may allow for something to both have and not have the property of persisting.

Second, we might adopt the position that change either does not happen or does not involve incompatible properties (forfeiting CHANGE). Here, we could deny

change altogether, perhaps accepting Parmenides's picture of a static, monolithic reality in which "what is is ungenerable and imperishable, a whole of a single kind, and unshaking and complete" (Curd 1998, p. 68). Or, we could hold that change occurs without involving incompatible properties. But change just does involve either something being *F* and something becoming not-*F*, or something being not-*F* and something becoming *F*. Sacrificing our minimal metaphysical commitments about how change works amounts to change nihilism. This strategy avoids contradiction at a very high metaphysical cost.

Finally, we could argue that nothing persists (forfeiting PERSISTENCE). Heraclitus told us: "You could not step twice into the same rivers; for other waters are ever flowing on to you" (EpistemeLinks.com 2005, Heraclitus of Ephesus, *On the Universe*, fragment 41). We might go along with him, agreeing that: "Nothing endures but change," giving us a metaphysics that does not include persisting objects, but merely flowing processes (EpistemeLinks.com 2005, Diogenes Laertius, *Lives of Eminent Philosophers*, Bk. IX, sec. 8). Such persistence nihilism is again a move at odds with strong intuitions and a range of metaphysical theories.

Thus, the strategy of dismissing one of the core theses leaves us without an intuitively tenable account of how ordinary things—Smith, books, rocks—exist and persist in the world. This motivates the search for an account of persistence that genuinely addresses the worry by reconciling the core theses.

FINDING THE SOURCE

The second way of dealing with the initial worry is to get much clearer about the source of the problem and then seek remedies by revising our ideas in a way that avoids the problem by attacking the source directly. Our understanding of CONSISTENCY needs to remain intact unless we allow contradictions, which is off the table here. However, CHANGE and PERSISTENCE leave room for interpretation. For instance, they leave open what counts as persistence, change, and incompatible properties being involved in change.

Modifying our understandings of these phenomena can ease the tension among the core theses. In our everyday understanding of the world, we assume that persisting objects survive the gain and loss of some simply instantiated properties. The following three aspects of this understanding are central to grasping why philosophical issues arise with persistence.

CHANGE AS ALTERATION. An object alters by gaining or losing properties. More precisely, an object alters if, and only if, it is numerically identical to objects that have different properties at different times. In our everyday understanding of the world, objects change by altering, and plenty of ordinary objects alter. Smith's book that was shut and Smith's book that is open is a single book that has the properties of being shut and open at different times. When Smith opened his book, the shut book did not wink out of existence exactly when an open book happened to blink into existence right into his hands. Rather, Smith's book was shut and then open—it altered as Smith turned to his bookmarked page.

PERSISTENCE AS SURVIVAL. An object survives if it has more than a momentary existence. More precisely, an object survives if, and only if, it is numerically identical to something that exists at a different time. In our everyday understanding of the world, objects persist by surviving, and plenty of ordinary objects survive. Consider the book Smith placed on his nightstand last evening that went untouched until this evening, and the book he removed from his nightstand this evening. The book that Smith put down last evening is the very same book that he picked up this evening. Although a day older, it is numerically identical to the book Smith read the prior evening—it survived the day spent on his nightstand.

INVOLVING INCOMPATIBLE PROPERTIES AS JUST HAVING INCOMPATIBLE PROPERTIES. An object just has a property if, and only if, it simply instantiates (*Fx*) that property. That is, an object just has a property if, and only if, no extrinsic facts are relevant to the truth of the proposition that the object has that property. In our everyday understanding, ordinary objects just have incompatible properties sometimes, regardless of how the rest of the world is. David Lewis brings out the intuitiveness of this when he writes: "When I sit I'm bent, when I stand I'm straight. When I change my shape, that isn't a matter of my changing relationships to other things, or my relationship to other changing things. I do the changing, all by myself. Or so it seems" (Lewis 1999, p. 187).

Like Lewis being straight, with respect to Smith's book, we tacitly hold that nothing beyond his book matters to its being shut—that there is a primitive, non-relational bond between the book and the property of being shut. If it is not open, Smith's book just has the property of being shut, regardless of its relation to the nightstand it rests upon at 7:00 a.m. We can capture these key aspects of our everyday understanding in terms of how objects persist through change with three additional theses.

THREE EVERYDAY METAPHYSICAL THESES

ALTERATION: If an object changes, then the object existing before the change and the numerically identical object existing after the change are the proper subjects of the incompatible properties involved in the change.

SURVIVAL: If an object persists through change, then the object existing before the change is numerically identical to the one existing after the change.

ATEMPORAL INSTANTIATION: If an object is the proper subject of a property, then (i) the object has that property, and (ii) facts about time and tense are irrelevant to the truth of the proposition that the object has that property.

ALTERATION constrains how things change. SURVIVAL constrains how things persist. ATEMPORAL INSTANTIATION constrains how incompatible properties are involved in change. Making our everyday understanding explicit is useful because it allows us to see that: (1) This understanding conjoined with the three core theses forces a contradiction; and (2) reconciling the core theses requires denying or revising some part of our everyday understanding. The following argument demonstrates both points. In it, we suppose that Smith opens the book that had been resting shut on his nightstand.

AN ARGUMENT AGAINST OUR EVERYDAY UNDERSTANDING

What follows are three assumptions about the book that capture the three core metaphysical theses: (1) It is not the case that the book is shut and the book is open (captures CONSISTENCY); (2) the book persists through change (captures PERSISTENCE); (3) the book changes in a way that involves the incompatible properties of being shut and being open (captures CHANGE).

The following steps draw on the three everyday metaphysical theses: (4) The book existing before the change is numerically identical to the book existing after the change (SURVIVAL and step two); (5) the book is the proper subject of being shut and being open (ALTERATION, steps three and four); (6) the book is shut and the book is open (ATEMPORAL INSTANTIATION and step five). From these six steps, a contradiction arises as steps (1) and (6) cannot both be true. One can conclude, then, that given the truth of the core metaphysical theses, something within the everyday metaphysical theses is false.

This argument can be run for any ordinary object that persists through change. Thus, to address, rather than dismiss, the initial worry, one of the three everyday theses must be revised or forfeited. The problem is to do so while striking a balance between respecting our intuitions and achieving philosophical success. Such is the strategy of three broad approaches to persistence below. Each blocks step (6) in its own way and thereby achieves a consistent view. But, given the nature of the problem demonstrated above, each solution will obviously face trade-offs in terms of intuitive appeal.

ADDRESSING THE WORRY

Perdurantism, exdurantism, and endurantism are each accounts of persistence that retain a commitment to the core metaphysical theses, but give up part of our everyday understanding of how things such as Smith, books, and rocks persist and change in our world. The first two accounts are built on a metaphysics of temporal parts, whereas the third depends on a metaphysics of enduring things.

METAPHYSICS OF TEMPORAL PARTS AND PERSISTENCE Ordinary objects have spatial parts. Perhaps they also have modal parts, dependent parts, abstract parts, or logical parts, among others. The metaphysics of temporal parts (MTP) leaves that open. The particular claim MTP makes is that objects have temporal parts. These temporal parts, time slices, or stages exist only at a moment. So, on a view consistent with MTP, multiple momentary book stages could exist—a shut-book stage, a distinct open-book stage, and so on. Perdurantism and exdurantism rely on the temporal stages of MTP to explain the persistence of ordinary objects.

Perdurantism Perdurantists take change over time to be analogous to change over space. Just as color changes across the surface of a canvas when different spatial parts of the canvas have incompatible colors, so the color of a lemon changes across the time as it ripens when different temporal parts—a distinct green stage and a distinct yellow stage—have incompatible colors. In both cases, change consists in distinct parts of an object having incompatible properties.

On this view, ordinary objects are space-time worms composed of distinct momentary stages. So, just like a taut rope extends through space, it also extends through time. For, as a fusion of its temporal stages, it has parts in the past, present, and future. An object that is a space-time worm is only partially present at any one moment because its different stages exist at different times.

The perdurantist ontology makes the three core and two everyday metaphysical theses co-realizable. An object changes when distinct stages of a single space-time worm just have the incompatible properties involved in change

(CHANGE and ATEMPORAL INSTANTIATION). It survives a change in virtue of the space-time worm that exists at the times that its distinct stages exist (PERSISTENCE and SURVIVAL). And, because distinct stages bear the incompatible properties rather than a single object, there is no one thing that has incompatible properties (CONSISTENCY).

For instance, Smith's book changes because its stages just have the incompatible properties of being shut and being open. The book survives this change because it is numerically identical to the space-time worm constituted by its stages. Finally, no contradiction arises because distinct stages of the book have the incompatible properties, rather than Smith's book as a whole.

However, perdurantism requires us to sacrifice change as alteration. ALTERATION entails that change occurs only if one and the same thing has a property and then lacks the property. It entails that the book changes only if it and something numerically identical to it have the incompatible properties of being open and shut. But, perdurantists hold that distinct proper parts of a space-time worm book bear the incompatible properties—the shut-book stage and the open-book stage. So, there is no one thing that has incompatible properties—indeed that is how perdurantism avoids contradiction. By blocking step (5) in the argument above, perdurantists also block (6). Yet, in gaining a coherent account of persistence, perdurantists accept an account on which change is merely a succession of momentary stages that have incompatible properties.

Exdurantism Exdurantists or stage theorists take identity over time to be analogous to identity between possible worlds. To see this, assume that an actual sill-length window swag could be a floor-length swag in virtue of a floor-length counterpart in some possible world. Analogously, exdurantists assume that Smith's now open book was shut in virtue of a closed book counterpart resting on his nightstand in the past. In both cases, distinct objects (the sill-length swag and its floor-length counterpart, the present open book and its earlier shut counterpart) have incompatible properties.

On this view, an ordinary object is a single momentary stage that extends through space, but not through time, and that has temporal counterpart stages. Any object that is a single stage is wholly present at exactly and only the moment it exists.

The exdurantist ontology makes the three core and one everyday metaphysical theses co-realizable. An object changes when it and a counterpart stage just have the incompatible properties involved in change (CHANGE and ATEMPORAL INSTANTIATION). It persists when it and its temporal counterpart exist at different times (PERSISTENCE). And, because distinct stages bear the incompatible properties rather than a single object, there is no one thing that has incompatible properties (CONSISTENCY).

For instance, the change in Smith's book involves incompatible properties because his book just has the property of being open and a counterpart stage just has the property of being shut. Smith's book persists through this change in virtue of standing in a counterpart relation to a stage from a different time in the actual world. Finally, because no single thing is open and is shut (rather, distinct stages are), no contradiction arises.

Notice that according to exdurantism, the object that changes and persists just has one of the incompatible properties—Smith's book, the entire book, just is open. In contrast, according to perdurantism the object that changes and persists never just has either of the incompatible properties—Smith's book is never just open or shut. Exdurantism thus fares a bit better intuitively on this point, for when we look at Smith and see him reading an open book, we think his book is open, not some other object that is merely part of his book.

However, exdurantism pays for this metaphysical perk elsewhere. Exdurance precludes the possibility of persistence as survival, for no ordinary objects survive. SURVIVAL entails that a persisting object exist both before and after it changes. It entails that if Smith's book persists, then the shut book on the nightstand is numerically identical to the open book in Smith's hands. But, exdurantists maintain that no book is numerically identical to both the earlier open stage and the later shut stage. At best, a persisting object continues (in some sense) in virtue of a succession of distinct momentary stages bearing the relevant counterpart relations to each other. But, an earlier and a later stage in such a succession are no more one and the same object than the first and third links in a five-link chain are one and the same link. Thus, given the ontology in which ordinary objects are all momentary stages, nothing exists that could survive change.

Moreover, because it shares the strategy of using MTP to explain persistence with perdurantism, exdurantism also forfeits ALTERATION. As above, there is no one object that loses one property and gains another. Instead, distinct objects bear the incompatible properties—Smith's open book and a shut stage to which it stands in a counterpart relation. But, the costs of exdurantism do return a benefit—giving up both SURVIVAL and ALTERATION blocks both (4) and (5) without which (6) does not follow. Of course, to recoup these costs, exdurantists may try to retain some form of ALTERATION or SURVIVAL by revising our notion of

existence. They could hold momentary stages derivatively exist across time in virtue of counterpart relations to other stages that exist at different times. Clearly, the burden of proof would fall on an exdurantist to prove that derivative existence *just is* existence.

To sum up, both perdurantists and exdurantists endorse MTP. They maintain a commitment to the three core theses by using temporal parts to bypass the contradiction that arises by simply predicating incompatible properties to a single object. Both approaches conflict with change as alteration—so neither can hold simply that the book is open and the book is shut, rather distinct stages have these properties. Ultimately, though, the views differ in metaphysical costs. Perdurantists may maintain that persisting objects survive change because they attribute incompatible properties to different parts of a single space-time worm. Exdurantists must deny SURVIVAL because they attribute incompatible properties to distinct ordinary objects.

METAPHYSICS OF ENDURING THINGS AND PERSISTENCE According to the metaphysics of enduring things (MET), some objects endure. To claim that some objects endure is to claim that in some cases a numerically identical object is wholly present at different times. This claim states the minimal metaphysical commitments that distinguish the ontologies of MET from MTP.

MET and MTP agree that ordinary objects have spatial parts, and that they may have modal parts, dependent parts, abstract parts, or logical parts, among others. MET also leaves open whether any objects have temporal parts.

However, although it permits stages, MET requires the existence of some objects that fall outside the ontologies of perdurantists or exdurantists. For, an enduring object is wholly present at different times and neither a space-time worm nor a single momentary stage can be wholly present at different times.

Endurantism relies on MET's enduring objects to explain how ordinary objects can be altered and survive change. These objects are the key resource that perdurantism and exdurantism lack by being grounded in MTP.

Endurantism Endurantists hold that ordinary objects persist through change by enduring. In doing so, they take identity over time to be numerical identity between objects wholly present at different times. They take change over time to be the instantiation of incompatible properties by numerically identical objects at different times. So, arguably they hold the most intuitive understanding of change over time as a phenomenon that is nothing more than one and the same object gaining and losing properties across time.

On a basic endurantist view, ordinary objects are enduring things. For example, an endurantist would hold that as an ordinary object, a book is not constituted by stages because it is wholly present at different moments. Thus, an ordinary, enduring book would be distinct from any sort of space-time worm or single momentary stage or counterpart stage that may or may not also exist.

The endurantist ontology makes the three core and two everyday metaphysical theses co-realizable. An object changes by altering because, in some sense, it has the incompatible properties involved in change (CHANGE and ALTERATION). It survives a change in virtue of the single enduring object that has those properties in some sense at different times (PERSISTENCE and SURVIVAL). Finally, although a wholly present ordinary object in some sense has incompatible properties, it does not just have those properties. Rather, facts external to an ordinary object concerning time or tense mediate the instantiation of incompatible properties. There are a variety of ways to mediate the instantiation. For instance, given a pair of incompatible properties and an object that has them in some sense, an endurantist could hold that the object has one property now and had the other property earlier. With various forms of mediated instantiation, the endurantist avoids contradiction (CONSISTENCY).

For instance, Smith's book changes because it has incompatible properties in some sense—his book is open but that very book was shut. Smith's book survives this change in virtue of being numerically identical to the book at the time it is open and the book at the time it was shut. Finally, because no single thing is open and is shut (rather, the book is open and was shut), the position remains consistent.

The important move of adopting temporally mediated property instantiation—instantiation mediated by time or tense—allows endurantists to hold that an ordinary object can be wholly present both before and after a change in spite of its having incompatible properties. This is why the view allows for the survival and alteration of objects so easily.

However, endurantism faces its own metaphysical cost—it requires us to give up the idea that an object just has the properties in virtue of which it changes. ATEMPORAL INSTANTIATION entails that there be a primitive bond unmediated by time or tense between the object and the relevant properties. Perdurantists and exdurantists preserve this bond because on their views distinct objects

just have the incompatible properties. MTP allows them to say, without contradiction, that one book stage just is open and a distinct book stage just is shut (*Fx* and *not-Fy*). In contrast, without stages as a resource, to preserve that bond the endurantist would have to say that the book just is open and that the book just is shut (*Fx* and *not-Fx*)—a flat contradiction. Instead, the view of change involving objects just having incompatible properties is replaced by one in which objects have incompatible properties in some sense mediated by time or tense (*F* is *x* and *F* was *not-x*). This is how endurantism directly blocks step (6) in the argument above.

In contrast to the sacrifices of the MTP theorists that include losing robust notions of survival and alteration, giving up primitive instantiation in favor of mediated instantiation may be appealing. But there are repercussions.

First, temporal concerns intuitively seem irrelevant to whether an object has those intrinsic properties in virtue of which it can change. Smith's green eyes, the position of his nightstand, and, likewise, the time of day all seem to be matters outside of the metaphysical status of Smith's book in terms of whether it is open or shut.

Second, those concerned about Bradley's regress may worry about relying on mediated property instantiation to explain persistence. Some take the position that primitive bonds are required to block the regress. The endurantist strategy rules out the possibility of such bonds holding between persisting objects and the properties involved in change.

Third, it obscures how the properties involved in change are incompatible. An enduring object has the properties of being *F* and not being *F* involved in change in a way that does not generate contradiction because, in some sense, they can be co-instantiated. For instance, if Smith's book is shut-in-the-morning and open-in-the-evening, this looks no more problematic than Smith's book being rectangular and red. Thus, with any kind of mediated instantiation, the endurantist will need to explain the incompatibility of the relevant properties. For, without incompatibility between the properties, change itself becomes questionable.

Various strands of endurantism handle these worries more or less well, depending in large part on how they mediate property instantiation. Possible methods include: time indexed properties (*x* is *F*-at-*t*), time relative predicate relations (*x* is-at-*t F*), relations with times as arguments (*x* is *F* at *t*), adverbial accounts (*x* is *F t-ly*), temporal context sensitivity (obtains at *t* (*x* is *F*)), and tense (*x* was *F*).

CONCLUSION

Perdurantism, exdurantism, and endurantism share the virtue of allowing us to maintain a commitment to the core theses of CONSISTENCY, PERSISTENCE, and CHANGE. Each does so by offering an account of persistence through change on which no single object just has the incompatible properties involved in change—whether it is because distinct objects just have those properties or a single object has them in a mediated way. Though they differ in particular metaphysical costs and benefits, this common feature is why they succeed in addressing rather than dismissing the initial worries with persistence.

At this point, the real problem with persistence is not deciding whether things persist—but rather explaining how they persist. The challenge today is to choose well among the metaphysical costs of reconciling the core theses so as to yield a coherent, useful theory that still respects our intuitions. The heart of the current persistence debate revolves around which view does the best job. Thus, it is worth remarking very briefly on three metaphysical concerns that provide, or seem to provide, reasons for favoring one approach to persistence over another.

First is the metaphysics of time. Eternalism, presentism, and the growing block view are among the main alternative accounts of the nature of time. Their different commitments regarding the reality of times make these views incompatible. The eternalist claims that all times exist, the presentist argues that only the present exists, and the growing block theorist holds that the past and the present exist, but not the future. The truth of eternalism or presentism or the growing block view would help choose between accounts of persistence if, as some have suggested, MTP entails either eternalism or the growing block view, or MET entails presentism. However, recent work on persistence suggests that MTP or MET can incorporate eternalism, presentism, or the growing block view, though perhaps not with equal ease.

Second is a concern with how propositions about the past, present, and future have truth values. At issue, is whether the *is* of predication is irreducibly tensed (serious tensing) or the *is* is timeless (surface tensing) in the logical structure of propositions. Some have thought this issue will help decide among approaches to persistence because they believe that endurantists must use serious tensing. However, though endurantists must use some form of mediated property instantiation, it need not be a form that depends on tensing.

Third is an issue about how temporary intrinsic properties must be instantiated. Intuitively, an intrinsic

property of an object is one that the object has simply by virtue of being itself. Temporary intrinsics are intrinsic properties that an object has only temporarily. Above, being bent and being straight are temporary intrinsic properties of Lewis. Real change occurs when an object has, in some sense, incompatible temporary intrinsic properties at different times. Thus, any tenable account of persistence will need to explain how objects have temporary intrinsic properties.

Now, many hold the view that there must be a primitive bond between an object and its temporary intrinsic properties, that objects just have them. If so, then endurantism is not a viable account of persistence. For, endurantism achieves consistency only by insisting on some form of mediated property instantiation. However, among the many forms of mediated property instantiation, some mesh better than others with our intuitions and theoretical commitments regarding temporary intrinsics. So there is room for endurantists to come up with a reasonable account of temporary intrinsics when they devise an alternative to **ATEMPORAL INSTANTIATION**.

See also Aristotle; Identity; Lewis, David; Metaphysics; Parmenides of Elea; Time.

Bibliography

Armstrong, David. "Identity Through Time." In *Time and Cause*, edited by P. van Inwagen. Dordrecht: Reidel, 1980.

Balashov, Yuri. "Enduring and Perduring Objects in Minkowski Space-Time." *Philosophical Studies* 99 (2000): 129–166.

Barnes, Jonathan. *The Complete Works of Aristotle: The Revised Oxford Translation*. Princeton, NJ: Princeton University Press, 1984.

Baxter, Donald. "Loose Identity and Becoming Something Else." *Noûs* 35 (2001): 592–601.

Bradley, F. H. *Appearance and Reality*, 2nd ed. Oxford: Oxford University Press, 1897.

Brogaard, Berit. "Presentist Four-Dimensionalism." *Monist* 83 (2000): 341–356.

Carter, William and H.S. Hestevold. "On Passage and Persistence." *American Philosophical Quarterly* 31(1994): 269–283.

Cartwright, Richard. "Scattered Objects." In *Analysis and Metaphysics*, edited by Keith Lehrer. Dordrecht: D. Reidel, 1975.

Chisholm, Roderick. "Identity Through Time." In *Language, Belief, and Metaphysics*, edited by H. Keifer and M. Munitz. Albany: State University of New York Press, 1970.

Craig, William Lane. "McTaggart's Paradox and the Problem of Temporary Intrinsics." *Analysis* 58 (1998): 122–127.

Curd, Patricia. *Legacy of Parmenides: Eleatic Monism and Later Presocratic Thought*. Princeton, NJ: Princeton University Press, 1998.

EpistemeLinks.com. "Quotations for philosopher Heraclitus," www.epistemelinks.com/Main/Quotations.aspx?PhilCode=Hera, 2005.

Forbes, Graham. "Is there a Problem about Persistence?" *Aristotelian Society* 61 (1987): 137–155.

Gallois, André. *Occasions of Identity: The Metaphysics of Persistence, Change, and Sameness*. New York: Oxford University Press, 1998.

Haslanger, Sally. "Humean Supervenience and Enduring Things." *Australasian Journal of Philosophy* 72 (1994): 339–359.

Haslanger, Sally. "Persistence through Time." In *The Oxford Handbook of Metaphysics*, edited by Michael J. Loux and Dean W. Zimmerman. Oxford: Oxford University Press, 2003.

Haslanger, Sally. "Persistence, Change, and Explanation." *Philosophical Studies* 56 (1989): 1–28.

Hawley, Katherine. *How Things Persist*. Oxford: Oxford University Press, 2001.

Hawley, Katherine. "Persistence and Non-Supervenient Relations." *Mind* 108 (1999): 53–67.

Heller, Mark. "Varieties of Four-Dimensionalism." *Australasian Journal of Philosophy* 71 (1993): 47–59.

Hinchliff, Mark. "The Puzzle of Change." In *Philosophical Perspectives, 10, Metaphysics*, edited by James Tomberlin. Cambridge, MA: Blackwell, 1996.

Hirsch, Eli. *The Concept of Identity*. Oxford: Oxford University Press, 1982.

Johnston, Mark. "Is There a Problem about Persistence?" *Aristotelian Society* 61 (1987): 107–135.

Lewis, David. *On the Plurality of Worlds*. Oxford: Basil Blackwell, 1986.

Lewis, David. "Rearrangement of Particles: Reply to Lowe." In *Papers in Metaphysics and Epistemology*. Cambridge, U.K.: Cambridge University Press, 1999.

Lewis, David. "Tensing the Copula." *Mind* 111 (2002): 1–13.

Lewis, David. "Zimmerman and the Spinning Sphere." *Australasian Journal of Philosophy* 77 (1999): 209–212.

Lombard, Lawrence. "On the Alleged Incompatibility of Presentism and Temporal Parts." *Philosophia* 27 (1999): 253–260.

Lowe, E. J. "Lewis on Perdurance versus Endurance." *Analysis* 47 (1987): 152–154.

Lowe, E. J. "The Problem of Intrinsic Change: Rejoinder to Lewis." *Analysis* 48 (1988): 72–77.

Lowe, E. J. "Substantial Change and Spatiotemporal Coincidence." *Ratio* 16 (2003): 140–160.

Ludlow, Peter. *Semantics, Time, and Tense*. Cambridge, MA: MIT Press, 1999.

Markosian, Ned. "A Defense of Presentism." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman. Oxford: Oxford University Press, 2003.

Markosian, Ned. "The 3D/4D Controversy and Non-Present Objects." *Philosophical Papers* 23 (1994): 243–249.

McTaggart, J. M. E. *The Nature of Existence*, vol. 2. Cambridge, U.K.: Cambridge University Press, 1927.

Mellor, D. H. *Real Time*. Cambridge, U.K.: Cambridge University Press, 1981.

Merricks, Trenton. "Endurance and Indiscernibility." *Journal of Philosophy* 91 (1994): 165–184.

Merricks, Trenton. "Persistence, Parts, and Presentism." *Noûs* 33 (1999): 421–438.

Myro, George. "Identity and Time." In *The Philosophical Grounds of Rationality*, edited by Richard Grandy and Richard Warner. New York: Clarendon Press, 1986.

Oaklander, L. Nathan. "Temporal Passage and Temporal Parts." *Noûs* 26 (1992): 79–84.

Oderberg, David. *The Metaphysics of Identity over Time*. London; New York: Macmillan, 1993.

Parsons, Josh. "Must a Four-Dimensionalist Believe in Temporal Parts?" *The Monist* 83 (2000): 399–418.

Prior, A. N. *Papers on Time and Tense*. London: Oxford University Press, 1968.

Prior, A. N. *Past, Present and Future*. Oxford: Clarendon Press, 1967.

Prior, A. N. "Thank Goodness That's Over." *Philosophy* 34 (1959): 12–17.

Quine, W. V. O. "Identity, Ostension, and Hypostasis." In *From a Logical Point of View*, 2nd ed. Evanston: Harper and Row, 1963.

Sider, Theodore. "Presentism and Ontological Commitment." *Journal of Philosophy* 96 (1999): 325–347.

Sider, Theodore. *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford: Oxford University Press, 2001.

Simons, Peter. "Continuants and Occurrents." *Proceedings of the Aristotelian Society* 74 (2000): 59–75.

Simons, Peter. "How to Exist at Time When You Have No Temporal Parts." *The Monist* 83 (2000): 419–436.

Taylor, Richard. "Spatial and Temporal Analogies and the Concept of Identity." *Journal of Philosophy* 52 (1955): 599–612.

Thomson, Judith Jarvis. "Parthood and Identity across Time." *Journal of Philosophy* 80 (1983): 201–220.

van Inwagen, Peter. "Four-Dimensional Objects." *Noûs* 24 (1990): 245–255.

van Inwagen, Peter. "Temporal Parts and Identity across Time." *Monist* 83 (2000): 437–459.

Varzi, Achille. "Perdurantism, Universalism, and Quantifiers." *Australasian Journal of Philosophy* 81 (2003): 208–215.

Wasserman, Ryan. "The Argument from Temporary Intrinsics." *Australasian Journal of Philosophy* 81(2003): 413–419.

Zimmerman, Dean W. "One Really Big Liquid Sphere: Reply to Lewis." *Australasian Journal of Philosophy* 77 (1999): 213–215.

Zimmerman, Dean W. "Temporary Intrinsics and Presentism." In *Metaphysics: The Big Questions*. Cambridge, MA: Blackwell, 1998.

Zimmerman, Dean W. "Temporal Parts and Supervenient Causation: The Incompatibility of Two Humean Doctrines." *Australasian Journal of Philosophy* 76 (1998): 265–288.

Roxanne Marie Kurtz (2005)

PERSONAL IDENTITY

One of the commonest of daily experiences is that of recognizing our friends. A less common, though still fairly familiar, experience is the decision that a certain person is or is not the person he claims to be. The problem of personal identity is that of clarifying the principles behind these indispensable processes of reidentification. To reidentify someone is to say or imply that in spite of a lapse of time and the changes it may have wrought, the person before us now is the same as the person we knew before. When are we justified in saying such a thing, and when are we not?

THE BASIC PROBLEMS

Some philosophers have said that we are never justified, because sameness and change are, in themselves, incompatible. They have argued that it is almost paradoxical to say that something has changed and yet is still the same. There is nothing special about the case of persons in this connection, except, of course, that we might, as persons ourselves, be expected to be more concerned about this case or to have access to some of the facts needed to deal with it. One set of such facts is the private set of thoughts, feelings, and images that each of us has, and such philosophers as David Hume have emphasized how constant and rapid are the changes in them with which our identity has to contend. The problem generated by this alleged paradox will be referred to as the problem of the unity of a person through change or, more briefly, as the problem of unity.

Most discussions of personal identity, however, have taken it for granted that sameness and change are, at least, often compatible and have concentrated on the conditions under which reidentification of persons can take place. What enables us to say, in spite of the changes wrought by time, that person *A*, before us now, *is* the person *B* whom we formerly knew and that person *C*, also before us now, is not?

The problem of the conditions for reidentifying persons should be distinguished from the problem of individuating persons. To individuate among a class of beings is to pick out one from another; to reidentify a member of a class of beings is to recognize him as the same as someone known at an earlier time. It is, of course, unlikely that these two notions can be kept separate, since, on the one hand, one has to be able to pick out a being from among his contemporaries before one is able to identify him with a past member of his class (which, in turn, had to be picked out) and since, on the other hand, it is hard to see how a being that exists in our world of time and change can be picked out, at least in the deeper sense of being recognized, without being picked out as a being with a certain history. It is not accidental that the word *identify* can sometimes mean the one procedure

and sometimes the other. This article will be concerned directly only with the problem of the reidentification of persons, which will be called the problem of criteria.

It has had two main competing answers. One is that the criterion of the identity of a person is the identity of the body that he has—that it is either a necessary or a sufficient condition of saying correctly that this person before us is Smith that the body this person before us has is the body that Smith had. The other answer is that the criterion of the identity of a person is the set of memories he has—that it is either a necessary or a sufficient condition of saying correctly that this person before us is Smith that he should have memories of doing Smith's actions or of having Smith's experiences.

It is clear that in practice we settle problems of identification in both ways. But we can still ask of each one whether it is necessary or sufficient; we can ask whether each is independent of the other; and we can ask whether one is more fundamental than the other. It is in connection with these questions that we find what are usually called puzzle cases. These are stories, sometimes true but usually imaginary, which are thought to contain prima facie conflicts between the two criteria. In deciding how the conflict is to be resolved, it is thought that we show the order of priority of the two criteria. For instance, there are the cases of ostensible "bodily transfer," like that of the cobbler and the prince mentioned by John Locke. In this story what physically seems to be a cobbler wakes one morning with all the apparent memories of a prince, with no knowledge of shoe mending, and with disgust at his present sordid surroundings. We might make the story harder by imagining that at the same time what looks like the prince wakes up in the royal palace with cobbler memories. In a story like this, persons seem to recall actions and events associated with a body other than the one they now have. Should we say that they are the persons their supposed memories suggest they are or the persons they physically seem to be? To decide this entails deciding on the relative importance of the two criteria of identity.

RELATED ISSUES

The two problems I have distinguished are bound to and do overlap in the literature. The difficulty and importance of the question of personal identity, however, are greatly increased by the fact that it lies at the point of intersection of several major lines of philosophical inquiry.

INFLUENCE OF DUALISM. The problem of personal identity has traditionally been raised in a dualist context. Those who have discussed it have been greatly influenced by the picture of a person as composed of two entities—body and mind—which are only contingently related to each other. This has restricted the problem of unity so that it has become the problem of how one can be justified in attributing unity to the mind. This looks much harder than the problem in its more comprehensive form, since the thoughts, feelings, and images a person has are far less stable than is his body and since it is, to say the least, not easy to find what Hume calls "the bond that unites" them. Failing to find it, a philosopher may resort to a doctrine of spiritual substance and say that within each person there is some central component that preserves his identity because it never changes as his thoughts and feelings do; the philosopher must then decide whether this component can be detected by introspection or is unknowable. If he rejects this doctrine, as Hume did, he may give way to complete skepticism about identity.

SELF-KNOWLEDGE. The second issue with which the problem is involved is the relation between the knowledge a person has of himself and the knowledge that others have of him. There are a great many facts about a person that others can learn, it is often said, only by inference but to which he himself seems to have direct and privileged access. The usual examples are facts about his present thoughts, feelings, and intentions. But it looks as though something similar may be true about the past. Although others may have to ascertain whether I am a certain previously known person or did a certain past action by reference to external records or to my observable appearance, I seem to know this directly, in memory. This bears on the puzzle stories. It seems absurd, if we imagine ourselves as one of the participants in these tales, to suggest that someone else might know better than we who we are. If this is really absurd, the puzzles have to be settled in favor of memory; if it is not, we have to explain our natural tendency to want to settle them this way.

IMMORTALITY. A third connected issue is the possibility of survival. If the unity of a person is necessarily connected with the continuance of his body through time, then it is logically impossible for a person to survive the death of his body. If bodily identity is a necessary criterion of personal identity, then even if it could be shown that some nonphysical characteristics of a person continued after his bodily death, the person himself would not have been shown to have survived any more than (to use

Antony Flew's example) he would have been if it had been possible to preserve his appendix in a bottle. On the other hand, if bodily identity is not a necessary criterion of personal identity, perhaps bodily death is merely one major event in a person's history and not the end of him. And if the fundamental criterion of identity is memory, it would seem to follow that a person might be known, at least to himself, to have survived death because he continued to have memories in his disembodied state.

MORAL CONSIDERATIONS. The concept of a person has moral connections. Problems of reidentification arise in practice largely when we have to decide questions of right or responsibility, such as right to inheritances or responsibility for crimes. Identity is a necessary though not a sufficient condition of someone's being accorded rights or being made to shoulder penalties. This applies in the afterlife, too. Only if beings who exist after our death can be identified with us can they rightly be held heirs to our merit or blame. A theory of personal identity must take this fact properly into account.

THE "SELF"

One result of these wider connections has been an unfortunate technical restriction on the language in which personal identity has come to be discussed. It has been referred to as the problem of the self. This word is sometimes used to mean the whole series of a person's inner mental states and sometimes, more restrictedly, the spiritual substance to which the philosopher says they belong. The use of the word *self,* however, has the effect of confining the question to the unity of the mind and of preventing the answer from relying on the temporal persistence of the body. This has made the unity problem seem intractable, especially when the fleetingness of mental images, feelings, and the like is contrasted with the temporal persistence their owner needs in order even to engage in the relatively lengthy processes of dreaming, reasoning, or scrutinizing the external world. This article therefore avoids a terminology that has ruled out one line of solution ab initio by making it impossible to endow the owner of mental processes with physical characteristics.

By far the most important classical discussions are those of Locke and Hume, and it is therefore useful to begin consideration of the problem of personal identity by reference to their attempts to solve it.

LOCKE

INCOMPLETENESS OF THE CONCEPT OF IDENTITY. Locke began his discussion of identity in Chapter 27 of the *Essay concerning Human Understanding* by pointing out a vital fact that others, including Hume, have since neglected. The concept of identity has to be joined to some substantive notion like that of a tree or a person in order to have any use at all. What makes us say that a given entity is the same depends on what sort of entity it is. This implies an answer to the unity problem—an entity of any sort can remain the same throughout its changes provided that the changes that take place in it are characteristic of entities of that sort and are allowed for in their concept. Over the years a tree can double its size and remain the same tree since this sort of change is characteristic of trees and is allowed for in the concept of a tree. It cannot, however, sprout wings and fly or burn to ashes and still remain a tree, for changes of these kinds are not allowed for. This being so, no hidden substance is necessary for the retention of its identity since there is no need of the unchanging character that this is said to provide. The same is true, presumably, of persons, and all that seems to remain is the much harder question of what changes are allowed for in this concept—the problem of the criteria of identity. Locke characteristically failed, however, to follow through the implications of his own insight. Although he saw the inutility of the concept of substance, he still retained it and led himself into some confusions.

These confusions are partly engendered by his apparent assumption that is it possible to find one single criterion of identity for each sort of being. Our concepts are not as tidy as this. When the assumption is brought to bear on the very untidy concept of a person, the result is a distortion of the concept's logical character. This takes the form of a supposed distinction between "person" and "man."

"MAN" AND "PERSON." A man, according to Locke, is a certain sort of living organism whose identity depends on its biological organization. On the other hand, he defined a person as "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places; which it does only by that consciousness which is inseparable from thinking and essential to it." Further, "as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person." To sever the two notions in this way is a radical departure from ordinary usage, in which the two words are often interchangeable. Locke admitted this, without, however, seeing that the admission conceded that his account must be inaccurate as a description of the two "ideas." Of course, there is a point in the division; behind

it lies the recognition that there are two criteria of identity for persons. This Locke tried to accommodate to his belief that for each sort of entity there is one criterion only, by arguing that there are two distinct concepts, each of which has its own unique criterion, rather than one concept with two criteria. But Locke was not trying merely to be tidy; more important is the motive supplied in his claim that "person" is what he called a "forensic" term. A person is a morally responsible agent. It is clear that to establish by physical evidence that the man before us in the dock is the one who did the deed is not sufficient to show that he should suffer the penalty (though it is surely sufficient to show that no one else should, unless he instigated or compelled the deed). Locke wanted to mark this fact by a special restriction on the notion of a person, so that to state that someone is the same person who did the deed is to imply accountability without room for more (or much more) dispute. He thought it obvious that what makes people accountable for their actions is their ability to recognize them as their own. This seems to mean two things: first, an awareness of what one is doing when one is doing it and, second, an ability to remember having done it. Hence, he said that the criterion for the identity of persons, as distinct from men, is consciousness, a concept intended to embrace both awareness and memory. The fact that the same *man* is before us does not mean that the same person is, since the *man* may not be conscious of having done the deed in question and if the *man* is not conscious of having done it, then the *person* did *not* do it. Here Locke brought in the puzzle cases:

> Should the soul of a prince, carrying with it the consciousness of the prince's past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, everyone sees he would be the same *person* with the prince, accountable only for the prince's actions…. Had I the same consciousness that I saw the ark and Noah's flood, as that I saw an overflowing of the Thames last winter, I could no more doubt that I who write this now, that saw the Thames overflowed last winter, and that viewed the flood at the general deluge, was the same *self* … than that I who write this am the same *myself* now whilst I write … that I was yesterday.

Locke was misconstruing the facts to which he draws our attention. Even granting that only persons are accountable, persons are still men (for men are accountable). We may be morally right in making the memory of crimes a condition for punishment, but memory does not thereby become the sole criterion of identity, for physical presence at the crime is also a condition of responsibility for it. Both the criteria are used together, and the most Locke has shown is that the satisfaction of only one is not, for moral purposes, enough; he has not shown that each serves a different concept. One is tempted to sever them only because of the puzzle stories. These, however, do not represent the conditions under which our concepts have been evolved but, rather, imaginary new conditions that might force us into the decision to change them. As things now stand, we have one complex concept, represented variously by words like "person," "man," or "human being" and embedded in the specific notions of cobbler, prince, beggar, or thief. This concept has two complementary criteria of identity. If we allow ourselves to be forced to say that there are two concepts, each with one criterion, we are saying that our criteria here and now allow us to hold that the memory of a crime, even without physical presence, is enough to establish responsibility for it.

There is a possible Lockean reply to this. It is to say that when a person remembers his deeds but clearly does not have the body that performed those deeds, the deeds can nevertheless still be his because he may have done them in a previous body and have inherited another since. The same person will then no longer be the same man. This cannot be evaluated until we have considered the puzzle cases at some length. For the present let us turn to Locke's attempt to make memory the single necessary and sufficient criterion of personal identity. If this attempt is successful, his treatment of the puzzles is made highly plausible; if not, it becomes highly suspect.

IDENTITY AS MEMORY. That there is a big difficulty in the problem of identity as memory was clear to Joseph Butler and has recently been very skillfully argued by Antony Flew. Locke wished to say that Smith is the same person who did or witnessed *X* if, and only if, Smith has the memory of doing or witnessing *X*. But this is unclear. The verb *remember* and its cognates have a strong and a weak sense. In the strong sense, to say that someone remembers something is to imply the correctness of his recollection (at least in all but minor details). To say in this idiom that someone's recollection is erroneous is to say that he does not really remember, but only seems to. In the weak sense, to say someone remembers something is merely to say that he sincerely claims to remember it (in the strong sense). In the weak sense, memories can be mistaken. Now, it is clear that even though we do pay special attention to what people claim to remember when settling questions of identity, the fact that someone claims to remember doing or witnessing something does

not show that he did it or witnessed it. Even though sincere, he might be mistaken. Thus, to say that Smith is the same person if he has the memory of *X* must, it seems, mean that he has to remember *X* in the strong sense of "remember." But here a twofold difficulty arises.

How are we to decide between a genuine and an apparent memory in any given case? The candidate's inner conviction is unreliable. We seem to have to resort to more than the memory claims themselves. And the critical evidence would seem to have to be evidence of the person's physical presence at the scenes he describes. This suggests that the memory criterion is not self-sufficient, as Locke says it is, for in order to know that it is satisfied on a given occasion, we seem to have to use the bodily criterion first.

Apart from this it is much too stringent to restrict personal identity to cases where a person can actually recall his past actions or experiences. People forget. Therefore, we must alter our wording. Smith, we have to say, is the same person who did or witnessed *X* if, and only if, he could remember it. But what does "could" mean here? Taken in a practical sense, it seems too strong, for this would imply that if Smith did do or witness *X,* there is some actual set of procedures that, if we applied them, would enable him to recall it. But this may not be so; even psychoanalysts fail. If, on the other hand, "could" is not given this sort of sense, it is hard to see what its use here contributes, unless it is merely another way of saying that Smith is the one who did or witnessed *X* if, and only if, he is the person to whom the application of procedures designed to induce recollection is appropriate. Unfortunately, this is either straightforwardly untrue (since before we discovered who did or saw *X,* it would be appropriate to apply such procedures to all likely candidates, not just to Smith) or merely a concealed way of saying that Smith is actually the person who did it, so that *no one other than he could* remember it. Thus, the concept of memory seems, in this argument, to presuppose that of personal identity, rather than the reverse.

These arguments show that Locke was mistaken in trying to define personal identity in terms of memory because such a definition is necessarily circular. In at least this sense Butler was correct when he said that memory presupposes, and does not constitute, personal identity. Some philosophers have gone on to say that memory is not a criterion of identity for persons at all, since, they say, we cannot know whether someone's apparent memories are real without knowing by physical means that he is the person who was involved in the events he recalls. But this, it will be argued later, is also a self-defeating move. For the present it can be seen that Locke was undoubtedly wrong in holding that memory could be the *sole* criterion of identity for persons.

SPIRITUAL SUBSTANCE. A great deal of the argument of Locke's chapter is designed to reconcile his preference for memory with his doctrine of spiritual substance. The doctrine of spiritual substance is inherited from his view that some doctrine of substance is necessary to account for the fact that the qualities of an object cohere. This is presumably intended to account for their exhibiting a permanent ownership through time, as well as their belonging together in one region of space. Yet Locke denied that we have any knowledge of what substance is like, since our knowledge is restricted to the qualities of things. In the case of persons the doctrine is one of spiritual rather than material substance (whatever the difference between two unknowns may be). But it is clear that nothing whose character is totally unknown can be detectable by the senses or by introspection, so that the doctrine of substance, as Locke held it, cannot provide any answer to the problem of criteria. No one could be said to be applying a concept on the basis of facts to which he has no access. An intractable problem now arises. Granting for the moment that memory is the sole criterion of identity, what is the relation between this fact and that of the existence of the underlying substance? Is it not possible that the application of the memory criterion might lead us to ascribe identity when this was not metaphysically backed by the continuance of one substance? If this should happen, would we have made a mistake?

The most straightforward answer is the paradoxical one of saying that the memory criterion is merely a guide for making identity judgments and that their ultimate metaphysical justification must forever elude us—which would mean that we could never be more than roughly sure we were punishing the right people for crimes. Locke sought to soften this by two devices. One was to sever "substance" and "person" in the same way that he severed "man" and "person" and to insist that only persons are bearers of responsibility, the concept of substance being obscure and irrelevant. The difficulty with this is that it leaves the doctrine of substance without any connection to those entities whose unity it was supposed to explain. The other device was to say that it is the "more probable opinion" that the consciousness that makes for personal identity is "annexed to" one immaterial substance rather than a plurality and to found the faith in its not being otherwise on the goodness of the Deity "who, as far as the happiness or misery of any of his sensible creatures is

concerned in it, will not, by a fatal error of theirs, transfer from one to another that consciousness which draws reward or punishment with it."

But these are no more than devices and have to be used only if we represent the identity of persons as composed of one kind of fact yet recognized through another. For Locke himself, in his early comments on the varying criteria of identity for objects of different kinds, has provided us with a demonstration of the total inutility of the doctrines of substance. We do not need them to account for our ascriptions of identity through change; these rest upon our noticing characteristic patterns of sequence in things. But these patterns do not just supply the criteria for ascribing continuance. They are also the reasons for our doing so at all. In other words, the answer to the unity question lies in the same facts that yield the answer to the criteria question. The invention of substance was intended to explain a practice whose explanation Locke had himself provided in another way. That he did not draw the moral and altogether abandon this invention may in part be the result of his having inherited it from others and in part the result of the incompleteness of his account of the criteria of personal identity.

In Locke, then, we find: one answer to the unity problem in terms of substance and another in terms of the objects' characteristic patterns of change, which renders the first answer unnecessary; a clear recognition of the connection between problems of practical identification and moral responsibility, which is exaggerated to the point of caricature by the separation of the concepts of a person and a man; an unambiguous claim for the priority of the memory criterion of identity for persons, which seems on superficial examination to lead to circularities; and an introduction of the puzzle cases to force a decision in favor of the last claim. With the lessons of Locke's insights and errors behind us, we turn to Hume.

HUME

In Hume's famous section on personal identity (*Treatise of Human Nature,* Book I, Part IV, Sec. 6), we find a treatment of the topic that is, as would be expected, more polished and consistent than that of Locke. But since it is also radically defective, its very tidiness makes it less fertile. It has had a baffling effect on generations of readers because of Hume's ability to destroy metaphysical palliative solutions to problems without uncovering the confusions that give rise to them. This, in turn, issues in a paralyzing skepticism that rendered Hume even less capable than Locke of reaching a clear understanding of the conceptual structure he examines.

Hume began by attacking the spiritual-substance solution to the problem of unity, as it appears in the claim that there is a unique and simple "self" that each person is able to detect within himself. He argued with effective simplicity that he was unable to detect it in himself. He was accordingly forced to conclude that the belief in personal identity, since it lacks this justification, is erroneous. People are "nothing but a bundle or collection of different perceptions" in a constant state of change—for perceptions are all that Hume *could* detect in himself. In this situation all that a philosopher can do is examine how it is that men (himself included) "suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives." This psychological objective Hume tried to attain by uncovering a basic conceptual confusion that he claimed we all fall into. We fail, he said, to distinguish properly between two things—the "idea of an object, that remains invariable and uninterrupted thro' a supposed variation of time" (which is the prototype of identity) and the "idea of several different objects existing in succession, and connected together by a close relation" (which is as good an example as any other of diversity). We confuse these two ideas because of the mental laziness that makes us content with their superficial similarity. Strictly ("to an accurate view"), change destroys identity, but we are easily beguiled into overlooking that change has occurred. Once launched upon this convenient path of error, the mind is led further and further along it by certain recurring facts—it is easier for us to overlook than notice gradual changes, changes that are characteristic of certain objects, and changes that occur according to certain smooth and regular patterns, and so we choose to overlook them. Everyone is prone to this error, which therefore acquires the dubious sanction of custom. Sooner or later, however, philosophers arrive on the scene and notice the recurrent paradox in which men have thus involved their thinking. They see both that we do ascribe identity to changing things and that we have no apparent ground for doing this. The result is that since they cannot *find* such a ground, they *invent* one. Hence, the metaphysical fancies of substance and the self. But these are hollow solutions; there is no discernible bond uniting a person, though there are sufficient interrelationships between his thoughts, feelings, and memories to explain why we erroneously ascribe unity to him. Hume had no consolation to offer us in this alleged predicament other than his usual one: Even though philosophical constructions cannot justify custom, philosophical criticism cannot dislodge it. For philosophical reasonings have power only in the study, not at the backgammon table.

SAMENESS AND CHANGE. Given the premise that Hume shared with the philosophers of substance, his conclusions follow only too clearly. This premise is that there is indeed a paradox in ascribing both change and identity to the same subject, since to ascribe change is to deny that we *have* the same subject. To agree to this is to deny that there can be any genuine solution to the problem of unity and to show that even a substance solution is at best a palliative—and a misleading one. But this is a very odd premise to concede without a battle. It has the extreme, language-destroying consequence that no predicates that cannot be simultaneously ascribed to one subject can be ascribed to a subject at two different times. If it is a mere matter of custom that we violate this principle, at least the custom is indispensable. Surely, much argument is needed to show that the custom is paradoxical and the principle necessary. And there is very little argument in Hume to this effect. His account of the fundamental confusion he claimed to have detected is made plausible only by its vagueness. It looks reasonable to say there is a contrast between one continuing object and a succession of related objects, but this is so only if "object" is tacitly replaced by the same noun in each case. There is a contrast between one continuing note and a succession of related notes (and who would confuse one with the other?) but not between one continuing tune and a succession of related notes. It is by means of the second sort of arrangement, not the first, that we incorporate change into our language. In order to understand the unreality of the contrast that Hume was foisting upon himself, one has merely to recall Locke's principle that "same" is an incomplete term that functions only in conjunction with substantives. There are some conjunctions that would yield the contrast—"same note" and "succession of different notes" is obviously one. It is equally obvious that "same tune" and "succession of different notes" is not one. Thus, Hume was wrong to look for the source of the contrast, when it does exist, in the concepts of identity and diversity considered alone. The concepts do not operate alone and yield his conflict only in those cases where they are joined with the right substantives. In most cases it does not exist, because most substantive concepts (including that of a person) are designed to incorporate changes.

There is, of course, one sense of the words *same* and *identical* in which sameness and change are incompatible. This is the sense of "same" in which, if applied to two distinct things, it means "alike" and, if applied to one thing at different times, it means "unaltered." This we might call the comparative sense of the word. It is to be distinguished from the numerical sense, in which two things said to be the same are said not to be two, but one. Clearly, one thing cannot be said to be both changed and the same if the comparative sense is intended, but this is not the sense we intend when we wonder whether we are entitled to consider someone the same throughout changes. Once this is noted, we can easily see that there is no need to assume that "to an accurate view" an object has to be the same in the comparative sense to remain the same in the numerical sense. If this is missed, a sense of paradox will be only too easy to sustain.

On the other hand, our concepts do not allow all kinds of change indiscriminately. How much is allowed depends on the concept in question. A man can change in more ways before he is destroyed than a chair can. To know what alterations are and are not allowed is to know, among other things, what the criteria of identity are for the class of entities grouped under the concept in question. These matters may not always be easy to settle precisely. We may not be in a position to say whether we have the same things on certain occasions. When the roof is removed, does the house still exist, or are we left with something else? If the walls are torn down and rebuilt, do we have the same house or another? Sometimes the only answer at such a point is a decision on the scope of the concept. But for general purposes usage over the years has provided us with rough and ready conventions that (this is a truism) language-users know.

Hume was aware of this fact, but the logic of his position forced him to misrepresent it. Instead of presenting us with some general indications of the sorts of change that tend to be allowed under concepts (changes that are gradual, small, functionally absorbable into the whole, and so on), he claimed to present us with the factors which, in his view, beguile us most regularly into the habit of ignoring the changes taking place in objects right under our noses. But these factors (which do not at all conceal the changing character of our world from us) are the same ones that appear without this disguise in a correct account of the situation. It is from a detailed knowledge of the very facts he outlined that we derive the criteria for those very identity judgments that he declared to be always unjustified. This is not the first or last time a philosopher has drawn our attention to facts supposed to support one theory when they in fact support another.

Similar considerations apply to what Hume said about the creation of substance doctrines. It is probably true that philosophers have invented these in order to answer the unity problem, and it is, of course, a merit in Hume that he saw that there is no independent evidence for the truth of such doctrines. But he did not see that the

primary objection to them is not that they cannot be shown to be true, but that they are unnecessary. They are invoked to soften a paradox that does not exist. There is no contradiction between saying a thing or person has changed and remains that same thing or person if the changes are characteristic of that sort of thing or of persons. If there is no paradox here, there is no need of any metaphysical postulate to conceal it. If Hume had seen this, he would not have tried to render more palatable the skepticism to which he was led by rejecting the doctrine of substance, for such skepticism could arise only if the doctrine were thought to be both false *and* necessary. But it is only false. The substantialists do not vindicate the ordinary language-user, and Hume does not convict him. Both have misdescribed what he is doing.

PERSONS. In the specific comments that Hume made about the identity of persons, he was clearly working, as was Locke, in the restricted framework in which "person" means "mind." Only thus can we read his statement that people are nothing but bundles of perceptions. The restriction makes him exaggerate for skeptical purposes the discontinuity he claimed to have discovered in the life histories of persons—a discontinuity that does not exist if we include the history of each person's body as well as that of his mind.

But this error hides a deeper one. There is a curious unreality about Hume's discussion of whether we can observe any real bond between the perceptions of a person. This question cannot, of course, be raised unless we can already distinguish between one person and another. Hume, that is to say, was asking whether there is any uniting bond among those perceptions that belong to one person. But why should this question puzzle him if he can already distinguish between those perceptions that belong to one person and those that belong to another? It is at least likely that those features of persons that enable us to distinguish one from another (to individuate) at any one time should also enable us to reidentify people after lapses of time. Yet these features are, and have to be, largely physical ones. For each of us can have (or perceive) only his own perceptions, and without the recognition of the bodies of others, there would be no question of the ownership of perceptions other than one's own ever arising (or, therefore, of the ownership of one's own). In asking his question, Hume was assuming that the perceptions persons are alleged to consist of are somehow known to be in parallel strings, so that the only question remaining is what unites those perceptions that belong on any one string. But if, as he saw, there is no clear psychical factor uniting them, it might still be true that whatever determines their belonging to a particular string also serves to join them together along it. And this, after all, is part of what the body does. His puzzle arises in the form that baffled him only if we first differentiate persons from one another on the basis of their bodies and then, forgetting that we have done it this way, look for some substitute for this principle among the contents of the mind. The principle that the question throws into doubt has to be assumed for the question to be raised.

In Hume, therefore, we find a dismissal of metaphysical construction and an awareness of the general characteristics of the complex facts out of which we forge our criteria of identity. These, however, are rendered completely sterile by the skeptical use to which Hume had to put them. The skepticism is, in turn, the result of a rationalistic oversimplification of the notion of identity that prevented Hume from discovering the muddle at the heart of the unity puzzle and of the dualistic framework of thought within which he worked.

SOME INTERIM CONCLUSIONS

We can now draw some conclusions from this investigation of the two main classical discussions of self-identity. The first is that the problem of the unity of persons is a spurious problem that rests upon two errors concerning the idea of identity. One of the errors is the failure to take enough note of the distinction between comparative and numerical identity. The other is the failure to note that the concept of numerical identity works in harness with substantive class concepts that provide those who know how to employ them with rules for making correct identity judgments on entities within their classes.

The second conclusion is that the concept of spiritual substance is not only unverifiable (as Hume saw) but also unnecessary (as Locke saw and Hume did not).

The third conclusion is that the unity problem has acquired a specious appearance of difficulty because of a tacit restriction placed by philosophers on the concept of a person. Since only the psychical components of the person are considered, a picture of change and discontinuity is conjured up that makes the fictitious contrast between identity and change seem even more alarming.

This leads naturally into the fourth conclusion—that it is salutary to remind ourselves that our actual concept of a person is of a psychophysical being. Hence, talk of the criteria of identity for purely psychical beings is not talk of the concept of a person that we actually have. How far they would differ has yet to be decided, but we must at least begin by asking what the actual criteria for embod-

ied human beings are. Here we must bear in mind the apparent circularity of the view that memory is the sole criterion for the identity of human beings. The examination of Locke suggested that in order to apply it some covert reference to the identity of the body has to be made. We must first examine this suggestion with some care.

We shall begin by trying to clarify further the notion of a criterion. It will then be argued that the bodily criterion of identity is in certain important ways more fundamental than the memory criterion in present discourse, although memory is still properly called a criterion in spite of Locke's failure. It could not, however, be the sole criterion. We shall finally consider the puzzles and argue that although they present us with some difficult conceptual decisions, they would not *necessitate* a change of convention in favor of memory, although this is a *possible* response to them. An attempt will be made to show that the response, if made, is innocuous, so that the puzzles are devoid of the wide implications philosophers have thought them to have.

CRITERIA

Thus far, two things have been meant in calling bodily identity and memory criteria of personal identity. One is that it is by reference to one or the other of these facts about people that questions of identity are usually settled. The other is that practical knowledge of how to settle these questions in these ways is a necessary part of having the concept of a person. More needs to be said than this.

There are two areas where the notion of a criterion has been of special concern in recent philosophy. One is the problem of the knowledge of the mental life of other persons. It has been said by some, following Ludwig Wittgenstein, that we can have this knowledge because people's behavior is able to supply us with criteria for saying correctly that they have certain mental states. The other is the problem of the relationship between judgments of fact and evaluative judgments. It has been said by J. O. Urmson, R. M. Hare, and others that certain facts about things or people serve as criteria for evaluating them as good or bad. In both these cases the relationship the word *criterion* names is thought to be tighter than an inductive one and yet looser than a deductive one. In this discussion the word will not be used in this sense, since the relationship between bodily identity and memory, on the one hand, and personal identity, on the other, seems to be closer than this; it seems, in fact, to be straightforwardly deductive. In the discussion of Locke we saw that saying someone remembers something in the strong sense entails that it forms a part of his life history. It is now claimed that if a person before us has the body that Smith used to have, it follows that he is Smith.

Two comments are necessary. First, this does not commit us to any view about how we know that the criteria are satisfied. To explain how we discover that this man really remembers or really has Smith's body, it might be necessary to use the notion of a criterion in some other, weaker sense—to say, for example, that a certain accumulation of evidence left no more room for reasonable doubt on the matter. But this is another issue. Second, an objection has to be countered. It might be objected that if the relationship between memory or bodily identity and personal identity is deductive, then the criteria are sterile and unusable. For, the argument might go, if either of these facts entailed that this was the same person, we would have to know independently who it was before we could be sure the criterion was satisfied. (This is the objection mentioned in the case of Locke.) This is not a genuine difficulty, but it is instructive. The reason for introducing it can only be the doctrine that if one proposition, *P*, entails another, *Q*, then it is impossible to know *P* without first knowing *Q*. But this is only a dogma that has to be tested against the facts, which do not bear it out.

The difficulty can teach us, however, that the standard objection to Locke is too simple. Even though the fact that memory entails personal identity prevents us from defining one in terms of the other without a circle, it is still possible that we may sometimes know that a person remembers without having previously checked on his identity. If this were not so, then memory could not serve as a criterion, for it is an additional part of the notion of a criterion, as all philosophers have used the term, that it can be applied. I shall shortly argue that this knowledge is possible.

BODILY IDENTITY. Some philosophers have said that the bodily criterion is not a criterion at all because there are some occasions in which we find human bodies that are not persons—that is, dead bodies or bodies that are biologically alive but incapable of exhibiting personality. But my thesis is that bodily identity is a sufficient criterion for reidentifying persons and by hypothesis these are not persons. If we are asking whether *X* before us, who is a person, is the same as Smith whom we once knew, who was a person, it is a sufficient condition of an affirmative answer to know that *X*'s body is Smith's body.

A more serious-looking argument against bodily identity comes from the puzzles. It might be said that

when we use the bodily criterion, we are covertly assuming that there has not been any bodily transfer. This raises an important point of method: How are the puzzles to be treated? We shall treat them as cases of proposed conceptual innovation, as if those who invent them do so to make us imagine circumstances that would force us to change our conceptual habits and rely on one criterion alone, even though we now use two. I have argued that in using two criteria, we have not faced the sorts of problems the puzzles present. If this is right, then no proviso against them can be embodied in our present thinking, even covertly. (If anyone considers that such contingencies *are* already provided for, then what is said below about the puzzles can probably be transposed into the key needed to examine his view of what sort of provision we make.)

There are several ways in which the bodily criterion is more fundamental than the memory criterion. In the present thesis these statements should seem like truisms.

Although both criteria are sufficient, only bodily identity is necessary. "This is the person who fired the shot" is entailed equally by "This person has the body of the person who fired the shot" and "This person remembers firing the shot"; but although the third statement entails the first statement, it does not entail the second.

The bodily criterion is more extensive. It is a matter of chance that men remember the tracts of their lives that they do remember rather than those that they do not, and we can apply the memory criterion only when there are memories to use. But in a clear sense the bodily criterion can always be used, for the body is present whenever the person is.

The bodily criterion is more varied. There are more ways in which we can determine whether a person is physically the same as someone than there are ways of determining whether his recollections are genuine. There are blood tests, fingerprints, photographs, the testimony of witnesses, and much else. Of course, a candidate's memory claims can be used to support this evidence, just as physical evidence can be used to support memory claims. The resort to physical tests when the memory claims are in doubt, however, is much more nearly inevitable than the resort to memory claims when physical evidence is inconclusive, since there are so many ways of adding to the physical evidence and it is free from the nagging thought that there is more than one way of coming by information about the past.

These examples are enough to show that we should regard overconfident readings of the puzzle cases with some suspicion, since the normal order of priority between our criteria is not what these readings suggest that it is.

MEMORY. It has already been suggested that even though Locke was mistaken in thinking that he could define personal identity in terms of memory, it does not follow that he was wrong to think of memory as a genuine criterion of personal identity. It might be possible to know that someone remembered without first ascertaining in another way who he is. But if this is possible, it has to coexist with the fact that when men's memory claims are in doubt, decision hinges for the most part on physical tests.

One way of trying to relate these two is to say that when we accept a memory claim unchecked, as we often do, we are relying on an inductive connection between the memory claims of a person and the events he refers to. We have found, that is, that this man's memory claims are usually true or, perhaps, that most people's are usually true. We now accept his word on this basis. Sydney Shoemaker has argued persuasively that this is too simple. He has claimed that it is a logical truth that memory claims are usually true, not an inductive one. Following are his arguments: (1) If someone frequently said with sincerity that he remembered events that did not occur, we would be justified in concluding that he did not know how to use the word *remember.* (2) If a child learning the language were to behave in this way with the word *remember* or one of its cognates, we would tell him that he had not learned how to use it. (3) If we were translating an unknown language and were inclined to translate certain expressions in it as memory expressions, our decision whether to do so would have to hinge in part upon the truth or falsity of the statements beginning with those expressions; if they were generally false, we could not translate them in this way.

If these arguments are accepted, it should probably be added that in order to understand memory claims at all, we must be able to recognize cases of genuine memory, so that there must be *some* such cases and also that just as lies and false promises must be in the minority to succeed, so must insincere or mistaken memory claims. These arguments appear enough to refute any generalized skepticism about memory, unless the skeptic is prepared to deny that our language has those features on which these arguments depend—that its users are generally successful in communicating by means of it and that it is learned and not instinctive. We shall not investigate how far it is correct to regard something established by this

sort of argument as a "necessary truth." Although the arguments do depend upon features of language that might be argued to be contingent ones, it is still clear that the conclusions are not straightforwardly inductive, and for this reason I shall allow the label to stick.

It is, then, a necessary truth that memory claims are usually true, from which it follows that they can usually be relied upon in practice. But this does not tell us whether any given memory claim is true. The situation here is, rather, that we are justified in accepting someone's memory claims unless there is some reason to doubt them. Only when there is such a reason do we need to check them. It is this that enables memory to serve as a criterion of identity.

But this is a far cry from Locke's theory that memory is the sole criterion. The very facts that show it to be a criterion at all show that it could not be the only one. We must be able to use the distinction between true and false memory claims (even to learn memory language), and this means we must have at our disposal a way or ways of checking the claims that are made. This implies, of course, that we must be able to discover whether the speaker was, indeed, present at that which he describes. Thus, the availability of the bodily criterion of identity is a necessary condition of our having made the distinction between genuine and false memories, even though it often must, from our previous arguments, be in order not to resort to it but to accept memory claims at their face value. Memory is thus a criterion of identity, but it is absurd to suggest it could be the only one, for without the ability to use another we would lack the ability to use it.

This bears out the view that the bodily criterion is more fundamental. There are arguments in Shoemaker, however, which suggest that just as the memory criterion depends on the bodily criterion in the way we have seen, a similar dependence exists the other way. There is a dependence the other way, but it is not a parallel one. The dependence is one found in all cognitive procedures. Unless people had memories, they could not know past facts. If they did not know past facts, they could not know past facts about themselves or other persons, for we have to depend on either our own recollections or those of other witnesses to learn about the past of a human body. At some point memory testimony has to be accepted without further question, and to accept someone's testimony is to accept that he was indeed a witness to some past event. This is true and supplies us with one more argument to show that memory claims must usually be correct, but it does not establish parity between the two criteria because it does not show that in dealing with a problem of reidentification, it is impossible in theory to dispense with the memory claims of the candidate himself. This is possible, however, and is one of the reasons for the greater importance of the bodily criterion.

In spite of this many philosophers have accorded memory greater weight than the bodily criterion. This seems to be a result of what I shall call the "internality" of memory. In remembering, a person seems to have direct, rather than inferential, access to his own past, to know past facts about himself from the inside. This view of memory is reinforced by the fact that most people would admit to having quasi-perceptual experiences in the form of mental imagery when they remember. Most readers unhesitatingly follow the writers of bodily-transfer stories in assuming them to be intelligible—for how could someone who had systematic recollections of this kind be proved wrong about his own identity by outsiders?

This attitude is not shaken as much as it should be by the fact that in ordinary unsystematic cases we frequently find that even the most vivid recollections are illusory. This is presumably because of the traditional picture of memory as some sort of introspective contemplation of imagery. But what brings memory into the public arena and enables us to use it as a criterion of identity is not this or that sort of private experience but the claims made as a result of it. Indeed, the memory claims of those who deny having memory images are as negotiable in common speech as those of the rest of us. If someone were to claim that he remembered an event and if we were able to determine that he had indeed witnessed it, could give us correct information about it, and could not have come by this information through later research or hearsay, there could be no doubt that he did remember it. The presence or absence, vividness or faintness, of his private images would be of no interest.

It is nevertheless characteristic that when people remember, they have images. If it were not, it is hard to see how the traditional picture of memory could have gained currency. It is true that memory claims are corrigible public claims to knowledge about the past and true that those who make them usually seem to have memory images. It is the first claim that explains why memory has the status it has as a criterion of personal identity. It is the second claim that helps us to understand why some have thought it more fundamental a criterion than it is. For although the subject's unique possession of his images does not confer immunity on the claims he makes, it may have much to do with the fact that he makes them. And it is easy to imagine cases where someone has such experiences and makes the memory claims that they character-

istically engender only to find out afterward that these claims are unfounded. This is common enough. It is an easy extension of this to imagine situations in which the events described by such a person did in fact take place, but in the presence of a human body other than the one he has. We then have a typical philosopher's puzzle case. In such a situation characteristic image-laden experiences might take place, and the customary memory claims might be uttered, yet the contextual conditions surrounding correct memory claims would not exist. To allow in some such cases that the speaker really does remember is to change the meaning of this word, but the characteristic intimacy and feeling of conviction that such inner experiences engender might hide this fact from those imagining such examples.

BODILY TRANSFER

It is now time to look at the puzzles. There are, however, a great variety of these, and without deliberate restriction it is impossible to produce any example of the intricate conceptual decisions involved in them. We shall accordingly leave aside puzzle stories of persons who seem to vanish and reappear or who seem to be reincarnations of someone dead and keep to the case of apparent bodily transfer. What is said here is probably comparable to what could be said in these other cases.

Let us take a story in which the servants in a royal palace waken a person who looks as if he is the prince but who evinces complete bewilderment at his surroundings, utters memory claims befitting a cobbler, is astonished on looking into the mirror, and so on. At the same time a man who looks as if he is the cobbler produces princely reactions and memory claims and demands to be returned to the royal palace. What should we say?

B. A. O. Williams has pointed out that the puzzle cases are harder to state in detail than is usually thought. Are we really able to imagine a person with the cobbler's memories (which will include some acquired skills and personality traits) and the prince's body? I shall ignore this complication, though in fact it tends to support what I shall argue to be the best solution.

The first thing to notice about such a puzzle is that it is puzzling. We are torn two ways over it, as we would expect to be if we have two criteria in apparent conflict. On reflection, however, it is more puzzling because if what I have said above is correct, the bodily criterion is the more fundamental of the two, so that the priorities in present practice would lead one to expect that the puzzle should be settled in its favor. Yet those such as Locke, who invent these stories, take it for granted that our temptations are to settle it in favor of memory. And as far as their judgment of the temptations of most readers goes, they seem to be right. Any answer to the puzzle must take both sides of this paradox into account and try to reconcile them.

PRIORITY OF BODILY CRITERION. Let us first consider the recommendation that our cobbler-prince episode should make us abandon the bodily criterion in favor of the memory criterion.

Put in this bald way, the proposal is absurd. We have already seen reason to say that memory could not be the sole criterion for the identity of persons because using it requires the availability of another. But this, although true, is far too brusque a reaction to the puzzle, which could be used to argue a more modest proposal—to *weaken* the bodily criterion in certain circumstances.

The advocate of bodily transfer could begin his case by making certain admissions and could then say that they do not destroy the case for it. The admissions would be these.

First, in order to set up any case at all, we have to have someone who now makes memory claims that fit a body other than the one he now has. This requires that he should be reidentifiable as the same throughout the period during which he utters the claims. The claims have to be systematic in the circumstances, so the period has to be considerable. For such reidentification the criterion of bodily identity would be necessary.

Second, in order to set up any case at all, we have to know that there was actually a person in the past about whose life these memory claims seem to be accurate reports and that all the claims fit the life history of the *same* person in the past, who *was* the person the claimant now says he *is.* This can be known only if in the past we were able to reidentify that person over the period of his life. This requires the past availability of the bodily criterion.

But when these admissions are made, the advocate of bodily transfer need go no further; he can hold his ground here and say that bodily transfer is still possible. If we had a case where the memory claims of the man who seemed to be the cobbler systematically fitted the past of the prince and vice versa, these claims could be checked up on in detail. And they would be found, ex hypothesi, to fit a past human body; the only difference from normal would be that the body that they fitted was not the body uttering them. Yet the past of the body uttering them would itself be taken care of by a systematic set of memory claims now uttered by that body which they

did fit. In such circumstances it surely would be wholly natural to say that the two men had exchanged bodies.

In spite of much recent writing on the puzzles, there seems to be no satisfactory demonstration that the change in convention that would follow on our saying a transfer had occurred would lead to absurdities. It is therefore a possibility. If we make this decision, we would be forced to so weaken the bodily criterion that we were entitled to infer from its being the same body to its being the same person only if there were no (systematic) memory claims which pointed to its being another person. This would place the two criteria in a position of relative parity, for the memory criterion would hold in normal circumstances subject to bodily checks and the bodily criterion would hold except in those abnormal cases where there were detailed and systematic memory claims that conflicted with the normal reading of the bodily evidence.

Having allowed this, we must now emphasize two things. One is that other readings of these cases could be made, as will shortly be argued. The other is that even the adoption of the bodily-transfer reading of them does not have the exciting implications most have thought.

We have already seen that it lends no support to the view that memory either is or ever could be the sole criterion of identity for persons.

It also does nothing to support the suggestion that people could exist with no bodies at all or to give concrete meaning to the common picture of bodily transfer as someone's *going out of* one body *into* another.

Transfer cases, even if allowed, could only be exceptional. If they were not, we would have a world in which the procedures for applying memory concepts would be much more complex than they now are, and virtually impossible to learn. I do not think we could come to learn memory language if the basic use of the word *remember* were one in which it could refer not only to the past of the body uttering it but also to the past of another body (which, in turn, it could be allowed to "fit" only if it were certain that there were no other systematic memory claims to fit the same period available from that body itself). A concept as epistemologically fundamental as that of memory has to be more easily come by than it would be in this sort of world. But granted that it is simpler and has been learned in more straightforward ways, as at present, then it could be stretched to subsequently cover the exceptional cases.

The conclusion is, therefore, that although the logical possibility of bodily transfer has to be admitted, the implications are small and the wisdom of this particular change in our conventions is not self-evident.

ABANDONING THE MEMORY CRITERION. We shall now consider the reverse suggestion—that in the face of such a puzzle we abandon the memory criterion and keep the bodily criterion.

It is not immediately obvious what could be meant by this. If it means that we should ignore the memory claims of candidates for reidentification, this is something we could do in any case; the point at issue is the status of those claims when they are considered. If it means that we should reject memory claims that clash with the bodily facts, then this is something we do already and no change in conventions is implied in it. It must mean that we disallow the inference from "He remembers *X*" to "*X* formed part of his life history." But the difficulty here is that in order even to gather the bodily facts, we need to learn about the pasts of others, we have to use either our own memories or those of witnesses, and checks on one set of memories, as we saw earlier, require reliance on other sets. So a change of convention here must allow for the continuance of this reliance.

It seems possible to allow for it in only one way—to continue to say that memory claims are generally correct accounts of past actions or events but to add that these actions or events may have formed part of the life of a person other than the one now making memory claims about them. People, in other words, would be allowed to recall events in the lives of others. Two comments may be made here.

For reasons that would parallel those in the previous section, it seems that cases where people *did* recall events in the lives of others would have to be rare.

Suppose that in spite of his protestations *X* was just admitted to be the prince because he has the prince's body. He now says, "But I remember mending the shoes last night." Suppose *X* finally gives in and concedes that he must be the prince although it is still agreed on all sides that the cobbler did mend the shoes last night. *X* cannot just say, "Oh, I really remember the cobbler's mending the shoes, not myself." This will not do because it fails to distinguish between the new, special case in which one person remembers the deeds of another without having done them (or even having been present) and the familiar case in which one person remembers another's deeds through having witnessed those deeds. It is the second case that would be conveyed by a sentence like "I remember the cobbler's mending the shoes." I am not sure how far this difficulty could be removed by ver-

bal adjustment, but it is at the minimum an inconvenience under the new convention.

The conclusion is as before.

DENIAL THAT ONE CRITERION IS SATISFIED. There would thus seem to be two possible alternative conceptual changes that we could make, each of which would weaken a familiar inference and each of which would be awkward, though not demonstrably impossible. As a matter of fact, however, we already have at our disposal a much simpler device for dealing with such puzzles. Instead of pretending to abandon or to alter one criterion, we can refuse to allow that one of them is satisfied. This need not be thought of as merely a temporary device. If we were to come across odd examples of pieces of iron that did not obey the lodestone but seemed otherwise to satisfy tests for being iron, we could postpone conceptual change for some time by insisting either that the tests had not been properly administered or that it was not really a lodestone. Such moves would become irrational only if maintained in the face of repeated examples. It is hard to admit that the point of irrationality could ever be reached in the present case.

There are clearly only two moves of this sort here. We can deny that it is really the same body, on the grounds that the memory claims it utters fit another, or we can deny that it is really the case that the speaker remembers, because it was not the body before us that was present. Note that neither move involves denying a criterion as the term is being used here. It merely involves refusing to accept that one criterion is satisfied in those cases where accepting that both were satisfied would land us in direct contradiction. There seem to be insuperable obstacles in adopting the first move. For one thing, it would require us, in the case of human bodies, to adopt standards of reidentification that differ from those we accept in the case of all other physical objects. (And if we disregarded this and insisted on behavioral or memory criteria for the identity of human bodies, we would destroy the distinction between a human body and a person.) For another, we would find ourselves led straight into an absurdity. Note again that we are retaining the bodily criterion while making this move. If what is known to be spatiotemporally continuous with the prince's body utters cobblerlike memory claims and if for this reason we say that it is not really the prince's body, we are not able to go on to say that it is, instead, the cobbler's body; for, by hypothesis, it is not spatiotemporally continuous with the cobbler's body and is therefore *not* the same physical object as that body. Thus, it is nobody's body at all, which is absurd. Hence, we are not able to make the move of denying that it is the body it seems to be. But there is nothing to prevent us from making the other move—of saying that unless the bodily facts at least coincide with the memory claims a person utters, then these claims are false, however closely they fit the past of someone else. This would merely be the determined application to special cases of a procedure we now follow.

We could not, of course, stop there, for we would have to explain how the person came to forget his own past and have so much accurate information about another's. Heroic hypotheses of retrocognitive clairvoyance would have to be brought forward to deal with such strange things. Such hypotheses would have to explain how it was that a person could have information about someone else's past in a manner so phenomenologically similar to the way in which he normally remembered his own. But no greater heroism would be called for here than would be called for by accepting that one person could exchange bodies or memories with another—for the second idea would require much the same sort of hypothesis as the one I have mentioned, and the first would make it puzzling that people should remember their own pasts. Of course, each would introduce a difficult conceptual change.

PUZZLE CASES BECOMING COMMON. But would we not be forced into a conceptual change if such cases became common? For once, the complexities of our problem make it easier to deal with and enable us to give a negative answer. This can be understood from two sides. It has already been argued that either of the possible conceptual changes would require that the cases of bodily transfer or memory exchange be rare; otherwise, we would not have the memory concepts we do have. Yet in order even to state the problem, we must use memory concepts. From the other side, we have to remember that if we were to adopt the device recommended, then in cases like the one in our story we would say of the characters not that they remembered but that they "retrocognized." If such a convention were adopted, however, it would become the appropriate language for the persons to use in such situations. For what makes our problem is what makes the memory criterion possible—the occurrence of memory claims. These are made in public memory language. If the public language changed so that the inappropriateness of a standardly worded memory claim for such circumstances became generally recognized, then the persons themselves, on discovering that the bodily facts did not fit, would not say that they remembered but that they "retrocognized." Thus, by the time the cases

became common, they would cease to exist in the logically puzzling form, because they would cease to be heralded by claims to remember. Pieces of iron do not talk; people are different, and the very data of the puzzles would change if the cases occurred frequently.

PRIMACY OF MEMORY. The solution has now to contend with the fact that we do feel a genuine compulsion to read these puzzle stories in some way that favors memory and to say that the claimant himself must know who he is better than others ever could. There are two reasons for this compulsion. One derives from the internality of memory, the other from psychophysical dualism.

On the internality of memory it is enough to repeat that although it is people's public memory claims that relate to decisions about their identity, such claims seem to be made for the most part when people have had characteristic image-laden experiences. Many philosophers consider these to be more closely related to the logic of remembering than they really are, and since the privacy of imagery places reports of it in an epistemologically privileged position, this privilege is erroneously thought to extend to memory claims—overlooking the fact that memory claims are not reports of imagery. When a person imagines himself being involved in a puzzle story, he supplies himself with vivid and systematic imagery to occasion memory claims that do not fit his present body, and he forgets that the persistence and vividness of the memory could not override the impact of the public physical checks that are a necessary part of the conventions governing memory claims.

As for the theoretical dualism that lies behind so many arguments about personal identity, it has here been argued that however we read them, the puzzles do nothing to support dualism. But the investigation of them has been conditioned in many cases by dualist preconceptions.

Shoemaker correctly remarked that the concept of bodily transfer is compatible with a behavioristic view of the mind, for one might mean, when saying that the cobbler and prince had exchanged bodies, that in the case of each person his distinctive behavior patterns (including his memory claims and behavior) were to be found in a body other than the one in which they used to be found. This is true, but if this solution to the puzzles were urged upon us in conjunction with an overtly behaviorist view of personality, it seems plain that there would be no special obviousness in or compulsion toward this solution as opposed to the others, even though it would still be a possible one. The reason that we all feel some degree of compulsion toward accepting the bodily-transfer solution is that dualist preconceptions intrude themselves when we investigate the stories. It is taken for granted that we have an independently clear concept, with recognized criteria of identity, of a soul, spirit, or mind, which can be thought of as having a purely contingent relationship to the body, which it may abandon in favor of another body. (Locke's phraseology in introducing the puzzle is to the point: "Should the soul of a prince … enter and inform the body of a cobbler…".) The only available criterion for such a purely psychical being is presumably memory, but we have already seen that it cannot be self-sufficient in the way it would have to be for us to conceive such an entity independently. Yet this is necessary to justify otherwise vacuous talk about such an entity's entering one body, leaving another, and the rest. Anyone feeling impelled to read the puzzles in favor of memory is probably making covert use of this illegitimate picture.

An important objection could now be raised. It might be said that even though much reflection has been infused with a dualist theory, this is a linguistic fact of life that philosophers must accept without complaint, for all language-users, not just philosophers, tend to be dualists. Thus, all language-users, if faced with the puzzles, would tend to opt for the memory solution. If so, how can a philosopher cavil at this solution? For what we *should* say is usually to be determined by a decision as to what we *would* say.

This raises the difficult general question of how to react to a misleading theory that has filtered into ordinary discourse. In the present case we could argue as follows. Philosophers such as Gilbert Ryle have exposed many errors and confusions in traditional dualism. But they have spoiled their own case by representing themselves as champions of the common man against the professional philosopher. It is easy enough to show that nonphilosophers are dualists, too. However, the common man is a dualist in the same sense in which the philosopher is one—when he interprets his own thinking about mental qualities and conduct. What the antidualist arguments show is that laymen misconstrue in their interpretative moments the utterances and thoughts that they engage in in their day-to-day existence. (We could say that all of us are occasionally philosophers, when we think about our ordinary mental concepts, but most of us are bad philosophers because we misinterpret them.) These common theoretical misconstructions, though inconsistent with our daily use of such concepts, are usually harmless because of the merciful logical dispensation that allows us to make good sense with our concepts while talking non-

sense about them. Occasionally, however, the prolonged continuance of the misguided theory can infect the practice. One such occasion is the present one, where the tacit appeal to the illegitimate concept of an independently identifiable psychical entity exerts a compulsion upon the reader of a puzzle story to interpret that story as a case of bodily transfer. Here it seems legitimate to replace bad theory by better and to argue against taking this solution for granted. The memory solution the dualist reading implies is at best one competitor among others, and one is led to think it is required only by our use of concepts on more normal occasions if one has misunderstood those occasions.

CONCLUSIONS

Of the two problems distinguished at the outset, this article has tried to show that the first, the unity problem, is spurious, since the paradox on which it rests is only apparent. The criteria problem admits of no such clear-cut solution, since it is clear on examination that both the bodily criterion and the memory criterion are ineluctable components of our concept of a person. The bodily criterion is more fundamental, but the memory criterion is, in its own way, indispensable because of the basic epistemological status of memory itself. This is one of the many facets of the irreducibly psychophysical nature of persons. One important result of this conclusion is that it is absurd to consider memory as the sole necessary or sufficient condition of identity. Thus, it would not even seem possible to construct a coherent concept of an independently identifiable bodiless person of whose identity memory would be the sole criterion. It would seem to follow that disembodied survival is logically absurd. It is impossible to decide here whether the doctrine of bodily resurrection fares better. Our examination shows that the puzzle stories can at most embody situations in which the relationships between the two criteria could be altered by conceptual decision. They could not embody situations in which either could be abandoned in favor of the other.

See also Butler, Joseph; Dualism in the Philosophy of Mind; Hare, Richard M.; Hume, David; Identity; Immortality; Locke, John; Memory; Persons; Reincarnation; Self; Self-Knowledge; Shoemaker, Sydney; Williams, Bernard.

Bibliography

The literature devoted explicitly to the problem of personal identity is fairly small, but the amount that is devoted to related questions is immense. These works were made direct use of in the article: John Locke, *Essay concerning Human Understanding,* edited by Campbell Fraser (Oxford: Clarendon Press, 1894), Book 2, Ch. 27; David Hume, *A Treatise of Human Nature,* edited by L. A. Selby-Bigge (Oxford: Clarendon Press, 1896), Book I, Part 4, Sec. 6; Joseph Butler, "Of Personal Identity," appendix to *The Analogy of Religion,* edited by W. E. Gladstone (Oxford, 1897), Vol. I, pp. 385ff.; Antony Flew, "Locke and the Problem of Personal Identity," *Philosophy* 26 (1951): 53–68; Terence Penelhum, "Hume on Personal Identity," *Philosophical Review* 64 (1955): 571–589, and "Personal Identity, Memory, and Survival," *Journal of Philosophy,* 56 (1959): 882–903; B. A. O. Williams, "Personal Identity and Individuation," *PAS* 57 (1956–1957): 229ff.; Sydney Shoemaker, "Personal Identity and Memory," *Journal of Philosophy* 56 (1959): 868–882.

The most important classical discussion that is not discussed in this article is Thomas Reid, *Essays on the Intellectual Powers of Man* (Edinburgh, 1785), Essay III, especially Chs. 4 and 6. Reid has some admirable criticisms of Locke, but he is too wedded to the concept of substance to see that Locke's departures from common conceptual practice are not remedied by the use of it.

Another classic discussion is Immanuel Kant, *Critique of Pure Reason,* translated by Norman Kemp Smith (London: Macmillan, 1929), pp. 341ff.

There are some recent books whose discussions repay close study. See C. D. Broad, *The Mind and Its Place in Nature* (London, 1937), Sec. E, pp. 553ff. For stimulating arguments in favor of the notion of a substantial self, see C. A. Campbell, *On Selfhood and Godhood* (London: Allen and Unwin, 1957). Risierei Frondizi, *The Nature of the Self* (New Haven, CT: Yale University Press, 1953), contains interesting discussions of Locke and Hume, but its positive discussion seems to be vitiated by the restrictions of the terminology in its title. P. A. Minkus, *Philosophy of the Person* (Oxford: Blackwell, 1960), is obscure to a degree but has the only extended discussion of Reid. See also Sydney Shoemaker, *Self-Knowledge and Self-Identity* (Ithaca, NY: Cornell University Press, 1963), and his entry "Memory" in this encyclopedia.

The following articles take positions that radically differ from the arguments of the present article. H. P. Grice, "Personal Identity," *Mind* 50 (1941): 330–350, argues that the "self" is a logical construction consisting of experiences linked conceptually by memory. An authoritative presentation of the Kantian thesis that perceptual acts require a persisting subject or owner is found in H. J. Paton, "Self-Identity," in his *In Defence of Reason* (London, 1951). J. R. Jones, "The Self in Sensory Cognition," *Mind* 58 (1949): 40–61, attempts to dispense with the notion of a subject of perceptual acts. This paper generated an exchange on the concept of the self between its author and Antony Flew; see Antony Flew, "Selves," *Mind* 58 (1949): 355–358, and J. R. Jones, "Selves: A Reply to Mr. Flew," *Mind* 59 (1950): 233–236. This article has not been able to deal with the detail of the arguments presented in these papers but would hold that each in its own way is handicapped by the restrictions placed on the discussion of personal identity by Hume and by the use of the terminology of the "self." On the perplexities surrounding the notion of a purely mental subject, rather than the psychophysical person, as the owner of mental acts and events, see Ch. 6 of Gilbert Ryle's *The Concept of the*

Mind (London: Hutchinson, 1949), and Ch. 3 of P. F. Strawson's *Individuals* (London: Methuen, 1959). Both of these books have strongly influenced this article.

Other helpful recent treatments are C. B. Martin, *Religious Belief* (Ithaca, NY: Cornell University Press, 1959), Ch. 6, and A. M. Quinton, "The Soul," *Journal of Philosophy* 59 (1962): 393–409.

Terence Penelhum (1967)

PERSONAL IDENTITY [ADDENDUM]

At the center of the debate about personal identity since the 1970s has been the work of Derek Parfit, whose ideas, first published in his article "Personal Identity" (1971) and then extended and elaborated in his monumental *Reasons and Persons* (1984, part 3), revitalized and to some extent transformed the topic. The following discussion explains how this has come about and relates Parfit's ideas to those of other influential writers on personal identity from the 1960s on, in particular Bernard Arthur Owen Williams (1973), Sydney Shoemaker (1970, 1985, 1999), Robert Nozick (1981), Roderick M. Chisholm (1976), David Wiggins (1967, 1980, 1996), and Richard G. Swinburne (1973–1974). Since the 1990s the debate about personal identity has come to be focused on the correctness of the animalist view, the view that we are animals and that our identity conditions are entirely biological. This view is defended by a number of authors including Paul F. Snowdon (1991), Peter van Inwagen (1990), and Eric T. Olson (1997). Once again, a knowledge of Parfit's views is essential to understanding the arrival of animalism on the philosophical scene and assessing the plausibility of the animalist's position.

THE REDUPLICATION ARGUMENT

The starting point for the development of Parfit's ideas was provided by Williams in "Personal Identity and Individuation" (1973), in which he puts forward his famous reduplication argument, intended as an objection to any account of personal identity that entailed the possibility of reincarnation. Any such account, he argues, would have to make personal identity consist in psychological links between the later reincarnation claimant and the original person. But no such account could rule out the possibility of a situation in which there were two equally good "candidates" for identity with an earlier person, two people bearing just the same psychological links to the earlier person. But since two people cannot be identical with one person, no such account can provide a sufficient condition of personal identity.

A consensus quickly emerged, however, among other writers on personal identity, that the significance of Williams's argument was greater than he had seen. Though Williams himself remained recalcitrant, others saw that his argument consequently challenged, not just any account of personal identity that allowed for such possibilities as reincarnation, which involves a radical separation of personal identity from bodily identity, but any account of personal identity that proposed as a sufficient condition of personal identity a conceivably duplicable relation—that is, a relation that could conceivably take a one-many form. The result of this was to focus attention on the principle underlying Williams's argument, called the "only x and y rule" by Wiggins (1967, 1980) in his discussion of the reduplication argument, which emphasized the generality of the argument. The correct formulation of this principle is difficult, but roughly speaking it asserts that the question whether later x is the same person as earlier y can depend only on facts about x and y and the relationship between them, and no facts about any other individuals can be relevant to whether x is y. Otherwise put, what this principle asserts is that whether later x is identical with earlier y can depend only on the intrinsic relationship between them; it cannot be determined extrinsically.

RESPONSES TO THE REDUPLICATION ARGUMENT

One way to respond to the reduplication argument while retaining the only x and y rule is to question the logic of the argument. According to Williams in a reduplication situation the rival candidates for identity with the original person must be new existents, identical neither with him or her nor with one another. But it is possible, as argued by several writers, including John Perry (1972) and David Lewis (1976), to reject this description of the reduplication situation. It must be accepted that the post-fission rivals are distinct people, but it is possible, according to these philosophers, to reject the view that they are new existents; rather, they have existed all along, but have only become spatially distinct with the fission. There are various versions of this view. Their common element is the multiple occupancy thesis, that what makes it the case that two people existing at one time are two may be facts about what is the case at other times. This implies that we cannot know for certain how many people exist at a certain time without knowing the future.

This response to Williams allows the retention of an account of personal identity, which allows the possibility of reincarnation, while accepting the only x and y rule. However, a simpler, and more popular, response to Williams is simply to reject the only x and y rule and to elaborate an account of personal identity that explicitly packs into its sufficient condition the constraint that x is identical with y only if there is no third candidate z who can be considered a better or equally good candidate for identity with y. Such an account of personal identity, in terms of psychological continuity, is elaborated by Shoemaker in "Persons and Their Pasts" (1970), in which he also fashions the important concept of quasi memory as a way of responding to the objection that a vicious circle must be involved in explaining personal identity in terms of, possibly among other things, memory. Another sophisticated development of the best candidate approach is contained in Nozick's *Philosophical Explanations* (1981).

IDENTITY AND SURVIVAL

But the straightforward rejection of the only x and y rule is implausible, unless some account of its attractiveness is given. It is at this point that Parfit's ideas become relevant. In response to Williams's argument Parfit (1971, 1984) proposes that identity does not matter in survival. What does matter is a relation of psychological connectedness-cum-continuity that does conform to the only x and y rule, but it seems plausible that identity obeys the only x and y rule only because we mistakenly identify this relation with identity.

The contention that identity does not matter in survival, which is Parfit's most discussed claim, is one component of the reductionist view of personal identity he recommends, according to which facts about personal identity are not facts over and above other facts, as facts about nations are not facts over and above facts about people and their relations. Another component is that there need be no answer to a question of personal identity: Personal identity may in some cases be indeterminate. In addition, Parfit holds that there are no facts about personal identity other than facts about mental states, their relations to one another, and their relations to physical bodies and the happenings therein. Persons are not "separately existing" entities, and a complete description of reality could be wholly impersonal.

Of these three components of the reductionist view the first is the most obscure. What Parfit means by it, however, is that we do not have among our basic concerns a desire for our own continued existence and well-being. Insofar as we are concerned about these our concern is derivative from a concern for those future people (in the actual world, contingently, ourselves) linked by certain relations of psychological continuity and connectedness to ourselves as we are now. It is because we do not appreciate that this is the structure of our basic concerns that we are tempted to think that the only x and y rule is correct. The contention that personal identity may be indeterminate is a more straightforward claim. What Parfit has in mind is that in at least some of the puzzle cases described in the literature on personal identity our concepts, suited as they are in the first place to our actual circumstances, have no determinate application. Whether such indeterminacy is to be regarded as due merely to vagueness in language or to vagueness in the world is, however, a debatable point (for the argument that it must be regarded as due merely to vagueness in language, see Evans 1978). Parfit's third contention, that facts about personal identity are nothing over and above facts about the relations of mental states, indicates the Humean influence on his views.

RESPONSES TO PARFIT

Opponents of the reductionist view are described by Parfit as nonreductionists or as proponents of the simple view. According to this view personal identity is an unanalyzable datum. One such nonreductionist is Chisholm (1976), whose work is perhaps the most careful working out of such a view in the literature. Chisholm defends the simple view as the development of the views on personal identity by Bishop Butler (1897) and Thomas Reid (1941). Personal identity is what it is and not another thing, and it is identity in a strict and philosophical sense. Another philosopher who defends the simple view, and does so in conscious opposition to Parfit, is Swinburne (1973–1974). Swinburne emphasizes in particular the difficulty of making sense of the idea that one's own personal identity may be indeterminate and in doing so draws on arguments from Williams (1970).

These philosophers reject the whole Parfitian reductionist package. But the elements of the package are, arguably, separable. Or, at least, so some philosophers think. Thus, Shoemaker (1985) rejects the Parfitian claim that persons are reducible to their experiences in any sort of Humean way but accepts both that identity does not matter in survival and that personal identity can be indeterminate. Again, Lewis (1976) rejects Parfit's claim that identity does not matter in survival and the best candidate approach that it supports, while accepting that personal identity can be indeterminate.

Parfit's reductionist thesis about personal identity is not easy to assess or respond to. But, just as no philosopher writing on personal identity can afford to ignore the work of John Locke or David Hume, the same is true for Parfit. It can now be said that no other philosopher of the last century has had such an impact on the debate about personal identity. And Parfit's influence continues to affect the twenty-first-century debate, most notably by indicating how there is philosophical space for the animalist position.

ANIMALISM

The animalist thesis is that we—you and I and any other readers of this entry—are animals of a certain kind, that is, human beings, members of the species Homo sapiens. The thesis is not that all persons are animals. The possibility of persons that are not animals, but gods, angels, or inorganic robots is allowed. But the animalist does insist that we are human animals and as such have the persistence conditions of human animals. The second claim made by the animalist is that such persistence conditions involve no form of psychological continuity whatsoever and are entirely biological (a compromise position defended by Wiggins [1996] and McDowell [1997] is that we are animals, but our persistence conditions are neither wholly psychological nor wholly biological).

According to the animalist, then, things of different kinds can be persons, and the persistence conditions of an entity that is a person will depend on the kind of person it is. Hence, there are no necessary and sufficient conditions for personal identity as such, as sought for by the proponents of the psychological continuity approach. In that sense, there is no problem of personal identity. *Person* does not name a sort of substance but is merely a functional term, like *genius* or *prophet*, and is applicable to any thing with certain capacities (thought and reflective self-consciousness).

The main objection to the animalist thesis is that it does not accord with the intuition that transplantation of a cerebrum from one head to another with consequent transference of psychology (as in Shoemaker's Brown-Brownson case [1970]) will preserve the identity of the person. It is the transplant intuition that makes plausible—independently of a dualist metaphysics—psychological continuity accounts of personal identity.

The animalist, however, has a response to this argument. And it is at this point that Parfit's ideas become relevant. The transplant intuition is mistaken, the animalist can say (Olson 1997), and only seems to be attractive to us because the cerebrum recipient (Brownson) is the Parfitian survivor of the cerebrum donor (Brown) (for example, stands to the former in those relations of psychological continuity and connectedness that constitute what matters in survival), and we mistakenly believe that identity is what matters in survival. So we are led to believe that the cerebrum recipient is the same person as the cerebrum donor. Indeed, it may even be correct to say that the cerebrum recipient is the same person as the cerebrum donor, because we may use the phrase *same person* in ordinary speech not to express strict identity but to only imply Parfitian survival (Olson 1997).

DIFFICULTIES FOR ANIMALISM

The animalists can explain away the attractiveness of the transplant intuition in this way, of course, only if they endorse Parfit's thesis that identity is not what matters in survival. Moreover, since Parfit's thesis is controversial, the animalists must either endorse Parfit's own argument for it, or substitute another if they are to employ it with intellectual integrity. However, Parfit's own argument for his thesis, which appeals to cases of reduplication, involves rejection of the only x and y rule and the acceptance of a best candidate account of personal identity. But it seems difficult to accept that for natural biological organisms like human beings the only x and y rule must be rejected and a best candidate account endorsed.

However this may be, animalism has thus brought us back to the debate over the reduplication argument initiated by Williams (1973) and further explored by Lewis (1976) and Shoemaker (1970).

Shoemaker (1999) also points out that to reject the transplant in the way just described the animalist must in fact make a more radical divide between what matters and personal identity than Parfit himself. For it is consistent with Parfit's thesis that if a future person is my Parfitian survivor then he is literally identical with me unless fission or fusion or some other circumstance obtains, which precludes literal identity on logical grounds. But to explain the transplant intuition away by appeal to the distinction between what matters and personal identity, the animalist must reject this proposition. Again, the animalist must reject the proposition that if I exist at a future time I am then one of my present self's Parfitian survivors.

THE TOO MANY MINDS OBJECTION

These are difficulties for the animalists. However, their contention is that their opponents have still greater difficulties. The chief positive argument for their position given by animalists is the Too Many Minds Objection or,

as Olson (1997) calls it, the Problem of the Thinking Animal (see also Snowdon 1991).

The basic structure of the Too Many Minds Objection is straightforward. If I am not a human animal, then as I sit here writing this so does another thinking intelligent being with reason and reflection. For human animals are surely thinking things and if the human animal I presently coincide with (but am not identical with) lacks what it takes to think, then so do I (we share our brain, nervous system, and whole past history since I can satisfy Locke's definition of a person). So, if animalism is false, there are at least two rational beings within my skin, a person and an animal, and I am never alone. But the animal I share my skin with is not a person—there are not two persons here (so Locke's definition is wrong). However, since it shares the entire material basis for my thinking, it shares my thoughts, so it thinks that it is a person. But then how do I know that I am the person and not the animal thinking wrongly that I am a person (no doubt I have my reasons, but so does my animal, and since it is not a person, they must be insufficient)?

The defense of the animalist position is thus that to reject it involves an absurdly inflated ontology (I am never alone) and an outrageous skepticism (I cannot ever know that I am a person). And, if that were not enough, its rejection also undermines the formulation of the very problem its opponents seek to solve, since human animals are rational, intelligent beings, that is, Lockean persons, and yet must be denied to be persons in the sense the debate concerns. So whatever answer the opponent of animalism gives to the question of personal identity, it cannot be an answer to the Lockean question it was originally advertised as an answer to. In fact, there can be no answer to that question.

Three responses to this argument exist. The first is to say that we are human animals, but that our persistence conditions are partly psychological, so that the transplant intuition can be endorsed (Wiggins 1996, McDowell 1997). The second response is to deny that human animals can think (Shoemaker 1999) because a certain sort of persistence condition is necessary for being a thinker. The third response is to accept that human animals think and that we are never alone, but to deny that this involves the absurdities or the loss of the problem of personal identity that the animalist suggests. The concept of a person relevant to the debate, it has to be said, is not that which Locke explicitly defines, but that of the self, the object of first-person reference, and a distinction is needed between the concept of the thinker of "I"-thoughts (which applies both to the person and the animal, and the object of self-reference (which applies only to the person) (Noonan 1998).

CONCLUSION

Which, if any, of these responses to the Too Many Minds Objection can be accepted is a matter of current controversy. But even if they are all rejected, the animalist still faces challenges.

One of the most powerful is that the animalist's position is itself vulnerable to the Too Many Minds Objection. One way of arguing this is to defend (with Shoemaker 1999) the contention that the animalist must recognize something that coincides throughout its life with the animal, but outlasts it, the entity Shoemaker calls its "corpse to be." Another way of arguing that the animalist faces the Too Many Minds Objection is to suggest that he or she cannot, but must, accommodate indeterminacy in human personal identity over time without acknowledging coinciding thinkers unless he or she can hold that such indeterminacy is in the world rather than in language.

Whether or not these ways of arguing for the vulnerability of animalism to the Too Many Minds Objection are ultimately acceptable, it is clear that at the beginning of the twenty-first century the debate over personal identity is as lively and unsettled as ever. It is also becoming evident that its final resolution must turn on wider issues of ontology and philosophical logic.

See also Philosophy of Mind.

Bibliography

Butler, Joseph. *The Analogy of Religion: Natural and Revealed*. London: Henry G. Bohn, 1897.

Chisholm, Roderick M. *Person and Object: A Metaphysical Study*. La Salle, IL: Open Court, 1976.

Evans, Gareth. "Vague Objects." *Analysis* 38 (1978): 208.

Lewis, David. "Survival and Identity." In *The Identities of Persons*, edited by Amélie Oksenberg Rorty, 17–40. Berkeley: University of California Press, 1976.

McDowell, John. "Reductionism and the First Person." In *Reading Parfit*, edited by Jonathan Dancy, 230–250. Malden, MA: Blackwell, 1997.

Noonan, Harold. "Animalism versus Lockeanism: A Current Controversy." *Philosophical Quarterly* 48 (1998): 302–318.

Nozick, Robert. *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.

Olson, Eric T. *The Human Animal: Personal Identity without Psychology*. New York: Oxford University Press, 1997.

Parfit, Derek. "Personal Identity." *Philosophical Review* 80 (1971): 3–27.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Clarendon Press, 1984.

Perry, John. "Can the Self Divide?" *Journal of Philosophy* 69 (1972): 463–488.
Reid, Thomas. *Essays on the Intellectual Powers of Man*, edited by A. D. Woozley. London: Macmillan, 1941.
Shoemaker, Sydney. "Critical Notice of Reasons and Persons." *Mind* 94 (375) (1985): 443–453.
Shoemaker, Sydney. "Eric Olson: The Human Animal." *Noûs* 33 (1999): 496–504.
Shoemaker, Sydney. "Persons and Their Pasts." *American Philosophical Quarterly* 7 (4) (1970): 269–285.
Snowdon, Paul F. "Personal Identity and Brain Transplants." In *Human Beings*, edited by David Cockburn. New York: Cambridge University Press, 1991.
Swinburne, Richard G. "Personal Identity." *Proceedings of the Aristotelian Society* 74 (1973–1974): 231–247.
Van Inwagen, Peter. *Material Beings*. Ithaca, NY: Cornell University Press, 1990.
Wiggins, David. *Identity and Spatio-temporal Continuity*. Oxford, U.K.: Blackwell, 1967.
Wiggins, David. "Replies." In *Essays for David Wiggins: Identity, Truth, and Value*, edited by Sabina Lovibond and S. G. Williams. Cambridge, MA: Blackwell, 1996.
Wiggins, David. *Sameness and Substance*. Cambridge, MA: Harvard University Press, 1980.
Williams, Bernard Arthur Owen. "Personal Identity and Individuation." In *Problems of the Self: Philosophical Papers, 1956–1972*. Cambridge, U.K.: Cambridge University Press, 1973. Originally appeared in the journal *Proceedings of the Aristotelian Society* in 1956–1957.
Williams, Bernard Arthur Owen. "The Self and the Future." *Philosophical Review* 79(1970): 161–80.

Harold W. Noonan (1996, 2005)

PERSONALISM

"Personalism" is a philosophical perspective or system for which person is the ontological ultimate and for which personality is thus the fundamental explanatory principle. Explicitly developed in the twentieth century, personalism in its historical antecedents and its dominant themes has close affiliations with and affinities to other (mainly idealist) systems that are not strictly personalist. This article will concentrate on American personalism, although the movement is not only American; there are and have been advocates of personalism or closely related positions in Europe, Great Britain, Latin America, and the Orient.

BACKGROUND OF THE TERM

The term *person* comes from the Latin word *persona,* meaning mask and/or actor. It came to refer to a role and to a man's dignity in relation to other men. This usage is reinforced by theological language for which *persona* is the Latin equivalent of the Greek *hypostasis* (standing under) and for which both *persona* and *hypostasis* are closely related to *ousia* (substance). These associations foreshadow the ultimacy that personalism attaches to personality, both in value (a person is identified with his dignity) and in being (person is substance). On this basis we can understand the importance that personalists have attached to Ancius Manlius Severinus Boethius's definition of person as an individual substance of a rational nature (*Persona est naturae rationabilis individua substantia*). The effect of the modern critique of the concept of substance on the definition of person will be considered later.

In comparison with *persona,* the term *personalism* is relatively recent. Walt Whitman and Bronson Alcott both used the term in the 1860s; early in the twentieth century it was adopted and applied more systematically. In France, Charles Renouvier wrote *Le personnalisme* in 1903; in Germany, William Stern developed critical personalism in *Person und Sache* (1906). In the United States, Mary Whiton Calkins began to use the term in 1907 and Borden Parker Bowne adopted it the following year. Bowne said of himself, "I am a Personalist, the first of the clan in any thorough-going sense." About this time, personal idealism established itself in England. Shortly thereafter, Neo-Scholastic (and hence, more realistic) versions of personalism emerged, especially in France.

HISTORICAL ANTECEDENTS

The historical antecedents of these personalistic philosophies are so pervasive and for the most part so well-known that they need not be discussed in detail here. A. C. Knudson supplies abundant historical background in *The Philosophy of Personalism* (1927). In general, personalism has been decisively influenced by both the Greek metaphysical and the biblical religious motifs of the dominant Western theological tradition. With the notable exception of J. M. E. McTaggart's atheistic personalism, personalism in virtually all its forms has been integrally connected with theism. Nevertheless, it has usually considered itself a system defensible on philosophical grounds and not one based merely on theological presuppositions.

Recognition of the dominant historical influences on personalism would not, therefore, be complete without mention of several modern philosophers. Following René Descartes, the primacy and indubitableness of personal experience and its identification as mental substance have exercised a decisive influence on nearly all forms of personalism. The Cartesian principle is apparent in Edgar Sheffield Brightman's definition: "A person … is a com-

plex unity of consciousness, which identifies itself with its past self in memory, determines itself by its freedom, is purposive and value-seeking, private yet communicating, and potentially rational" (in *A History of Philosophical Systems,* edited by V. Ferm, p. 341).

Gottfried Wilhelm Leibniz is sometimes spoken of as the founder of personalism. His doctrine that all reality is composed of monads (psychic entities) without remainder and that monads are essentially centers of activity has been particularly influential on idealistic personalists of pluralistic and panpsychistic types.

The influence of George Berkeley converged with that of Leibniz in providing an impetus to idealistic personalism. Material substance is reinterpreted as the "language" of the Divine Person. Further reinforcement for this theme is found in Immanuel Kant's doctrines of the phenomenality of the sense world and the primacy of the practical reason. It is only in the personal world of the practical (moral) reason that one has access to the noumenal. This Kantian direction has had enormous influence on what might be called ethical personalism.

G. W. F. Hegel was the single most important influence in the development of absolute idealism (absolutistic personalism). His emphasis on dialectical movement toward wholeness, on the concrete universal, and on the ultimacy of spirit has had a decided influence on other forms of idealistic personalism, notably that of Brightman.

One thinker who does not compare with the foregoing figures in eminence deserves to be mentioned because of his influence on such American personalists as Bowne and G. T. Ladd. He is Hermann Lotze, whose main work is *Mikrokosmus* (1856–1858).

TYPES OF PERSONALISM

In characterizing more precisely the systematic position of personalism, it will be helpful to distinguish two major forms: realistic personalism and idealistic personalism. The former can best be understood in the context of supernaturalism or traditional metaphysical realism, and the latter in terms of metaphysical idealism.

REALISTIC PERSONALISM. For realistic personalists, personality is the fundamental being. That is, ultimate reality is a spiritual, supernatural being. There is also, however, a natural order of nonmental being, which although created by God is not intrinsically spiritual or personal. Many Neo-Scholastics, for example J. Maritain, E. Gilson, and E. Mounier, identify themselves as personalists in the realistic sense. In fact, realistic personalism has been developing with remarkable vitality both in Europe and America in conjunction with the resurgence of Catholic theological thought. There are, however, some realistic personalists who do not stand in the scholastic tradition; among them may be mentioned N. Berdyaev, J. B. Pratt, D. C. Macintosh, Georgia Harkness, and A. C. Garnett.

IDEALISTIC PERSONALISM. Excluding Platonism and Kantianism, there are three main types of idealism: absolute idealism, panpsychistic idealism, and personal (pluralistic) idealism. Although there are no neat lines of demarcation separating these types, oversimplification can in this case be illuminating.

(1) Absolute idealism (or absolutistic personalism) is the view that reality is one absolute mind, spirit, or person. All finite beings, however otherwise designated (for example, as physical things, logical entities, or human beings), literally participate in this absolute being; they *are* ontologically by virtue of their being manifestations or activities of the absolute mind. Since this is so distinctive a philosophical tradition, it receives full treatment elsewhere. Representative thinkers who have either had influence on or association with other personalistic positions are Edward Caird, T. H. Green, Josiah Royce, A. E. Taylor, Mary W. Calkins, and W. E. Hocking. With reservations, C. A. Campbell, Brand Blanshard, Paul Tillich, and Gabriel Marcel may also be included here.

Absolute idealism has not commended itself to personal idealism, which, in opposing complete immanence or monism, is closer to realistic personalism and related theistic positions.

(2) For panpsychistic idealism, Leibniz's monadology is the paradigm. Reality is a hierarchy of psychic beings (monads) determined by the degree of consciousness possessed by any monad. The supreme monad (God) has created all other monads in preestablished harmony. Panpsychism has been developed in various ways by James Ward, F. R. Tennant, H. W. Carr, A. N. Whitehead, and Charles Hartshorne.

In many respects, panpsychistic idealism may be considered to be continuous with personal idealism. Although personal idealists do not deny the possibility that there are more grades of self or mind than the human and the divine, they tend to believe that panpsychists have not adequately resolved the tension between pluralistic and monistic strains in their position.

(3) Personal idealism is usually considered the most typical form of personalism. It is idealistic: all reality is

personal. It is pluralistic: reality is a society of persons. It is theistic: God is the ultimate person and, as such, is the ground of all being and the creator of finite persons. Henceforth *personalism* will be used to mean personal idealism.

SYSTEMATIC THEMES

Among the first generation of American exponents of personalism the most significant were George Holmes Howison (1834–1916) and Borden Parker Bowne (1847–1910).

In the 1860s Howison was a member of the St. Louis Philosophical Society. The discussion of Hegelian idealism, to which this group devoted so much of its time, led Howison to reject what he considered the submerging of the finite individual in the Absolute.

His basic metaphysical position is stated categorically: "All existence is either (1) the existence of *minds,* or (2) the existence of *the items and order of their experience*; all the existences known as 'material' consisting in certain of these experiences, with an order organized by the self-active forms of consciousness that in their unity constitute the substantial being of a mind, in distinction from its phenomenal life" (in J. W. Buckham and G. M. Stratton, eds., *George Holmes Howison,* p. 128). Howison's unswerving pluralism led him not only to reject pantheism but also to deny creation. "These many minds … have no origin at all—no source in *time* whatever. There is nothing at all, prior to them, out of which their being arises… . They simply *are,* and together constitute the eternal order" (ibid., p. 129). Howison's "eternal republic" is reminiscent of Royce's community.

Bowne taught philosophy at Boston University from 1876 until his death. Berkeley, Kant, and Lotze were the major influences on his thought. Like Howison, Bowne was a pluralistic idealist, but unlike Howison, he was explicitly theistic. The Divine Person is not only the creator of finite selves or persons but is also the "world ground," whose "self-directing intelligent agency" shows itself in the order and continuity of the phenomenal world.

Bowne's famous chapter in *Personalism* on "The Failure of Impersonalism" expresses his basic polemic against Hegelian absolutism, Herbert Spencer's evolutionism, associationism, and materialism. At the same time, he fought just as hard against fundamentalism and dogmatic supernaturalism. Through his influence on many generations of students at the Boston University School of Theology, he contributed decisively to liberalizing the leadership of the Methodist Church.

Three of Bowne's students were the leading exponents of personalism in the period following World War I. Albert C. Knudson (1873–1953) continued the personalist tradition in theological context at Boston University School of Theology. Ralph Tyler Flewelling (1871–1960) developed the School of Philosophy of the University of Southern California and also founded and edited the journal the *Personalist.*

Edgar Sheffield Brightman (1884–1953), the most important of Bowne's students, taught at Boston University from 1919 until his death. Brightman, a creative and original thinker, developed a comprehensive and coherent personalistic system.

Brightman espoused an epistemological dualism of "the shining present" (or "situation-experienced") and "the illuminating absent" (or "situation-believed-in"). Immediate experience is the inescapable starting point, but experience always refers beyond itself (self-transcendence). The possibility of reference is found in the activity of the mind in knowing; the adequacy of reference is determined by the criterion of coherence. Maximum coherence in interpreting experience is maximum truth. In his emphasis on the tentativeness and testing of hypotheses, Brightman is empirical; in his emphasis on system and inclusive order, he is rationalistic.

In metaphysics, Brightman maintained that "everything that exists [or subsists] is in, of, or for a mind on some level." He defined personalism as "the hypothesis that all being is either a personal experient (a complex unity of consciousness) or some phase or aspect of one or more such experients" (*Person and Reality,* p. 135). The natural world is understood as an order within or as a function of the mind of God. Finite persons are created by the uncreated Person. Human persons are, therefore, centers of intrinsic value.

Brightman might be called a value empiricist. His *Moral Laws* (1933), which has not received the attention it deserves, works out an impressive ethical theory. In his philosophy of religion values have a central place. The value dimension of human experience provides the evidence of a religious dimension of reality. Hence, generically, God is the source and conserver of values.

The most distinctive aspect of Brightman's thought is his revision of the traditional idea of God. He argued that if we are to take personality seriously as the basic explanatory model, then we must accept a temporalist view of God. If God is personal, he is omnitemporal, not

timeless. Brightman also argued that the traditional conception of divine omnipotence could not be maintained without seriously qualifying the divine goodness. His penetrating consideration of evil, suffering, and death led him to conclude that the will of God is limited by nonrational conditions (the Given) within the divine nature that are neither created nor approved by that will. God maintains constant and growing—although never complete—control of the Given. Some personalists, including L. Harold DeWolf, prefer to follow Bowne's more traditional view of God's eternity and omnipotence. Others, like Peter A. Bertocci, find in Brightman's revisions the conditions of an intelligible and cogent theism.

CURRENT DEVELOPMENTS

In recent years, personalism may seem to have been eclipsed by the rise of existential and analytic philosophies. However, many of the doctrines and motifs of personalism have been or are being appropriated and elaborated by other positions. Existentialism and the phenomenological movement have turned to the exploration of personal existence in ways that will be gratifying to most personalists. This movement should be particularly fruitful for personalists since it grapples in new ways with the relation of the body to the person, a problem that has caused a long-standing ambiguity in personalistic thought. Both realistic and idealistic personalists have stumbled over this problem. Phenomenological investigations may therefore provide an impetus for new conceptions of personality.

The analytic concentration on language also contributes to an improved understanding of personal symbolizing and communication, and the renewed interest in philosophy of mind, stimulated by recent psychological theories, again provides material that is important in the development of personalist thought. Personalists would seem to have an advantage in being willing to risk a systematic conception of the total person that would combine surface experience (sense) and depth dimension (value).

Among the large number of Brightman's students who have been developing various facets of personalistic thought, the best known is Bertocci, Brightman's successor as Borden Parker Bowne professor of philosophy at Boston University. Other contemporary personalists also continue to demonstrate that personalism can be a viable alternative among persistent philosophical perspectives.

See also Absolute, The; Berdyaev, Nikolai Aleksandrovich; Berkeley, George; Blanshard, Brand; Bowne, Borden Parker; Brightman, Edgar Sheffield; Caird, Edward; Descartes, René; Existentialism; Gilson, Étienne Henry; God, Concepts of; Green, Thomas Hill; Hegel, Georg Wilhelm Friedrich; Hocking, William Ernest; Howison, George Holmes; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Marcel, Gabriel; Maritain, Jacques; McTaggart, John McTaggart Ellis; Mounier, Emmanuel; Panpsychism; Platonism and the Platonic Tradition; Renouvier, Charles Bernard; Royce, Josiah; Taylor, Alfred Edward; Tennant, Frederick Robert; Tillich, Paul; Whitehead, Alfred North.

Bibliography

GENERAL WORKS

Brightman, E. S. "Personalism (Including Personal Idealism)." In *A History of Philosophical Systems,* edited by V. Ferm, 340–352. New York: Philosophical Library, 1950.

Flewelling, R. T. "Personalism." In *Twentieth Century Philosophy,* edited by D. D. Runes, 323–341. New York: Greenwood, 1968.

Knudson, A. C. *The Philosophy of Personalism.* New York: Abingdon Press, 1927.

AMERICA

Beck, Robert N. *The Meaning of Americanism.* New York: Philosophical Library, 1956.

Bertocci, P. A. *Introduction to the Philosophy of Religion.* New York: Prentice-Hall, 1951.

Bertocci, P. A., and R. M. Millard. *Personality and the Good.* New York: David McKay, 1963.

Bowne, B. P. *Metaphysics.* New York, 1898.

Bowne, B. P. *Personalism.* Boston, 1908.

Bowne, B. P. *Theism.* New York, 1902.

Brightman, E. S. *Person and Reality.* New York: David McKay, 1958. Edited after Brightman's death by P. A. Bertocci; contains a selected bibliography of Brightman's writings compiled by J. E. Newhall.

Brightman, E. S. *A Philosophy of Religion.* New York: Prentice-Hall, 1940.

Buckham, J. W. *The Inner World.* New York: Harper, 1941.

Buckham, J. W., and G. M. Stratton, eds. *George Holmes Howison, Philosopher and Teacher.* Berkeley: University of California Press, 1934.

Čapek, Milič. *The Philosophical Impact of Contemporary Physics.* Princeton, NJ: Van Nostrand, 1961.

DeWolf, L. H. *The Religious Revolt against Reason.* New York: Harper, 1949.

Flewelling, R. T. *Creative Personality.* New York: Macmillan, 1926.

Flewelling, R. T. *The Person or the Significance of Man.* Los Angeles, 1952.

Howison, G. H. *The Limits of Evolution.* New York: Macmillan, 1901.

Muelder, W. G. *Foundations of the Responsible Society.* New York: Abingdon Press, 1959.

Munk, Arthur W. *History and God.* New York: Ronald Press, 1952.

White, H. V. *Truth and the Person in Christian Theology.* New York: Oxford University Press, 1963.

Closely related positions are developed in the works of J. E. Boodin, J. S. Moore, D. S. Robinson, J. S. Bixler, and C. Hartshorne.

ENGLAND

Carr, H. W. *Cogitans Cogitata.* London: Favil Press, 1930.

Carr, H. W. *The Unique Status of Man.* London, 1928.

Oman, J. *Grace and Personality.* Cambridge, U.K., 1917.

Sturt, H., ed. *Personal Idealism.* New York: Macmillan, 1902.

Webb, C. C. J. *God and Personality.* London, 1919.

Note also the writings of H. Rashdall, W. R. Sorley, and F. C. S. Schiller.

GERMANY

Eucken, Rudolf. *Die Einheit des Geisteslebens in Bewusstsein und Tat der Menschheit.* Leipzig, 1888.

Lotze, H. *Mikrokosmus.* Leipzig, 1856–1858.

Stern, William. *Person und Sache.* Leipzig, 1906.

Note also the writings of Max Scheler.

FRANCE

Brunner, A. *La personne incarnée.* Paris: Beauchesne, 1947.

Lahbari, M. A. *De l'être à la personne: Essai de personnalisme réaliste.* Paris, 1954.

Maritain, J. *The Person and the Common Good.* London: Bles, 1948.

Mounier, E. *A Personalist Manifesto.* Paris, 1936.

Mounier, E. *Le personnalisme.* Paris: Presses Universitaires de France, 1950.

Nedoncelle, M. *Vers une philosophie de l'amour et de la personne.* Paris, 1957.

Ravaisson, F. *De l'habitude.* Paris, 1933.

Renouvier, C. *Le personnalisme.* Paris, 1903.

Note also the writings of Henri Bergson.

ADDITIONAL WORKS

Berdyaev, N. *The Destiny of Man.* Translated by N. Duddington. London: Centenary, 1937.

Buber, Martin. *I and Thou.* Edinburgh: Clark, 1937.

Romero, Francisco. *La filosofia de la persona.* Buenos Aires: Talleres Gráficos de la Editorial Radio Revista, 1938.

Stefanini, L. *Personalismo filosofico.* Rome, 1954.

John H. Lavely (1967)

PERSONS

INTRODUCTION

This entry is on personhood and the general philosophical question that will be treated is: What is a person? Common use of the term *person* makes reference to adult human beings. Typical examples of sentences in which this term is used are: "Descartes is the person most responsible for inaugurating the Modern Period in Western thought"; "No person can be President of the United States unless he/she was born in the United States"; and "Human fetuses may be considered persons." As the controversial last example should make clear, the term person is not used exclusively to refer to adult human beings. In much of the literature on persons, the term is used in a non-species-specific way. Many authors take *human being* to be a term of biology and leave the definition to science. Given that, here is a restatement of the initial question: What must a being be like to be a person?

There are many categories into which the term person fits. People refer to social persons, moral persons, metaphysical persons, legal persons, religious persons, and so on. While no one category of personhood can be considered the correct category, philosophers have tended to concentrate on either the metaphysical or the moral aspects of personhood. After a few words on the other categories, the metaphysical and moral notions of person will be the primary focus of the present entry.

The principal use of the concept of a person in the Christian community is that of God's personhood. This comes out most clearly in the tradition where the Holy Trinity is referred to as "three persons in one God." Although the concept of the Holy Trinity defies comprehension for many, one of the ideas spawned by this is that there is some way humans are like God, which is that they are both persons. Aquinas affirms that the term person applies to God as well as to human beings, though it does not apply in the same way. His definition of person is "a subsistent individual of a rational nature" (Aquinas 1945, p. 290). As applied to humans, Aquinas takes his lead from the use of person as one who is dignified, of high standing (in the community). He says that each individual of a rational nature is a person. However, since the dignity of God is greater than every other dignity, therefore, person applies preeminently to God. It is perhaps obvious that Aquinas is applying cultural as well as metaphysical attributes in his definition of the term.

As used in the legal sense, person refers to any being, object, or organization that has standing before the law. Perhaps the most enlightening example in the literature of law is that corporations are persons in the legal sense. This is because corporations have legal rights and responsibilities (some have also argued that corporations should be considered moral persons with moral rights and responsibilities). Legal rights would include equal protection, freedom of the press, due process, and so on, all of which can certainly be applied to corporations. Some

legal findings have not, however, extended full personhood to corporations, denying the following: pleading the Fifth Amendment in order to avoid self-incrimination, and Fourth Amendment rights of protection of persons.

Other interesting cases in the *legal persons* category are those of the fetus and the newborn. While these beings are protected under the law and, therefore, may be claimed to be legal persons, many philosophers have taken the position that fetuses are not persons in the moral sense of this term. Michael Tooley (1983) has argued that late-term fetuses and even newborns are not persons in the moral sense of this term. Tooley takes the side of caution here with newborns and says that since our knowledge of their development is limited, we need to agree on some cut-off point or other; he settles for a week, after which we can with clear conscience consider the newborn a person.

The social person is not so clearly defined, it seems, as persons of the other categories. The general framework for someone being a person in the social sense is for that being/person to be recognized as a person by those who are recognized as persons within the social community. Here, thoughts run to some of the ideas of Richard Rorty (1979, 1982), who takes the view that persons will be decided upon and not discovered. This is a provocative, and for some a rather radical, view, leaning toward relativism (though this is denied by Rorty) because if someone or some group in a society is judged by the society to be nonpersons, and if personhood is a matter of decision and not discovery, then said someone or the members of said group are, in fact, simply not persons. Ultimately, Rorty's position is that the concept of personhood is something that has been, and is still being, worked out in the *conversation* that is the history of the world.

METAPHYSICAL CONSIDERATIONS

This section is devoted to the metaphysical aspects of the concept of personhood.

CONDITIONS FOR PERSONHOOD. Over the centuries, necessary and sufficient conditions for personhood have been laid out by various philosophers. John Locke is usually the starting place for any serious philosophical study of the concept of personhood because he seems to be the first to make explicit what he meant by the term. He writes that a person "is a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does by that consciousness, which is inseparable from thinking, and as it seems to me essential to it." (Locke 1975, p. 335) Although Locke is here working on the idea of personal identity, there are at least three important concepts he introduces that would seem indispensable conditions of personhood proper, namely, reason, a first-person perspective, and consciousness. These characteristics of personhood arise in virtually all of the literature on the topic.

There is also the sense in which Locke uses person as a legal (forensic) term that may be useful to consider. Again, Locke is working on the issue of personal identity; however, what he says is important for thinking about persons in both the metaphysical and moral senses of the term. He writes:

> *Person*, as I take it, is the name for this *self*. Where-ever a Man finds, what he calls *himself*, there I think another may say is the same *Person*. It is a Forensick Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery. This personality extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable.
>
> (LOCKE 1975, P. 346)

While it can easily be seen that Locke is here referring to concern and accountability in the legal sense, the reference to happiness and misery may naturally lead one to contemplate what it means to be a person in the moral sense of the term. The section "Moral Considerations" herein will be devoted to this discussion.

P. F. STRAWSON'S THEORY OF PERSON. What was at one time *the* dominant paradigm on persons is the British philosopher P. F. Strawson's theory. While there are moral overtones, his is primarily a metaphysical theory. He gives the following definition: "the concept of a type of entity such that *both* predicates ascribing states of consciousness *and* predicates ascribing corporeal characteristics, a physical situation, etc., are equally applicable to a single individual of that single type" (Strawson 1963, pp. 101–102).

Strawson argues that a person is not some sort of compound of two different kinds of substance: (1) a pure consciousness/ego, and (2) a corporeal entity. These exist together in one being, according to Strawson. He is doubtful that there could even be such a thing as a pure consciousness existing on its own, devoid of any connection with a "physical situation." When he says that a person is not an "animated body" or an "embodied anima," he is here speaking to the idea that person refers to an

individual who must be analyzed as a unified individual of whom both types of predicates can be ascribed.

The predicates referred to here are as follows: *M-predicates*, on the one hand, are applicable to material bodies, to which there is no question of applying states of consciousness. Examples are: "is in the park," "is blue," "is flat." *P-predicates*, on the other hand, are all other predicates ascribed to persons. These are various, says Strawson. His examples are: "is smiling," "is going for a walk," "is in pain," "thinking hard," "believes in God" (Strawson 1963, pp. 104).

It is interesting to note that some P-predicates imply the having of consciousness by the subject of reference. Strawson's example is *posted a letter*. One consequence of this is that, theoretically, there are ways to tell when to ascribe P-predicates to others as well as to oneself. That is, there will often be indicators of the presence of P-predicates. What are they? One cannot just argue from one's own case. Strawson holds that one can ascribe a P-predicate to oneself only if one can apply it to others. On many occasions, one ascribes P-predicates to others on the basis of observing their behavior. He is not saying that others' behavior is a *sign* that P-predicates may be ascribed but, rather, that the criteria of observed behavior is logically adequate for the ascription of P-predicates. Further, some P-predicates one ascribes to oneself are not ascribed by using self-observation. This would seem to call into question the adequacy of Strawson's criteria for ascribing P-predicates in which he says that the same criteria for ascribing P-predicates to others must be/is adequate for ascribing P-predicates to oneself.

His conclusion on this point is that the character of P-predicates is such that one uses behavior criteria for ascribing to others and both behavior and nonbehavior criteria for ascribing to oneself. For him, to have the concept of a person is to be a "self-ascriber" as well as an "other-ascriber" of P-predicates.

THE CONSTITUTION VIEW. Lynne Rudder Baker is a leading proponent of this theory of personhood. In her closely argued book *Persons and Bodies* (2000), Baker tells us that while persons are constituted by their body, a person and a person's body are not identical. Her definition of *constitution* amounts to this: Where *x* constitutes *y* at time *t*, *x*, and *y* must be spatially coincident; *x* must be in a circumstance where *y*'s primary-kind property can be realized (where a primary-kind property is the property or characteristic an individual has by virtue of the kind of thing it is; for example: Secretariat's primary-kind property is that of being a horse); it is necessary that if anything (*z*) has some property at *t* that is *z*'s primary-kind property and if *z* is in a favorable circumstance to have *y*'s primary-kind property, then there is some individual *u* such that *u* has *y*'s primary-kind property at *t* and *u* is spatially coincident with *z* at *t*; it is possible that: *x* exists at *t* and there is no individual *w* such that *w* at *t* has *y*'s primary-kind property and is spatially coincident with *x*; *y* being immaterial implies that *x* is immaterial. Recall here that Baker is setting up her definition of what it means to be a person and hence has in mind (at least) what is usually taken as a clear example of a person, to wit, the adult human being, with a physical body.

A principal theme in Baker is that of the nonidentity of the person and the person's body. She draws an analogy between a thing and that of which it is constituted, and a person and that which a person is constituted, by using the example of Michelangelo's work of art *David* and the material of which it is constituted. Baker claims that the marble (called Piece) is not identical with *David*. Part of the argument runs as follows: If *David* and Piece are identical, then there is no property had by one and not had by the other. Piece has the property of being able to exist in a world without art whereas *David* (having as its primary-kind property that of being a statue, a work of art) does not have this property. Hence, constitution does not entail identity. (This is a very lean version of Baker's argument and the reader is advised to study the longer work for important details.)

This much said, Baker goes on to distinguish the person from the person's body (as that of which the person is constituted). Her argument hinges on the fact that the body (*qua* body) fails to possess what can be called the person-making property, that is, possession of a *first-person perspective*. The first-person perspective quite simply is the perspective by which one is/becomes conscious of oneself as oneself. Baker distinguishes two grades of the first-person perspective. An example of the *weak grade* would be referenced by someone uttering "I am 6 foot, 2 inches tall." The person (P) who utters this sentence is thought to have the ability to distinguish P from others. However, this is only half of what a full-on first-person perspective can be, according to Baker. If P utters the sentence "I wish I were 6 foot, 2 inches tall," this indicates that P sees not only that P is distinct from others, but also that P sees P as P. Following Castañeda, Baker uses the asterisk/star on the pronoun indicating first-person perspective to indicate as much. Hence, the sentence uttered would be written as "I wish I* were 6 foot, 2 inches tall."

To restate the important conclusion, the upshot of all of this is that since a person's body cannot take the first-person perspective, and since a person is a being who does or has the capacity to take the first-person perspective, a person's body and a person are not identical.

According to Baker, the first-person perspective underlies all versions of what it means to be a person, which rely on self-consciousness as the person-making characteristic. One example of a self-consciousness-based theory of personhood is one that Tooley (1983) writes about. On his interpretation, a being is self-consciousness to the extent that it is in possession of a concept of a self as a continuing subject of experiences and other mental states, is such an entity itself, and believes that it is itself such an entity. Tooley's important analysis of this, and other concepts, will be treated in the next section because Tooley's program revolves around the concept of personhood in the moral sense.

OTHER SUGGESTED CONDITIONS FOR PERSONHOOD. One of the most widely considered conditions for personhood is freedom of the will. A unique and pivotal contribution to this subject comes from Harry Frankfurt (1971), who argues that freedom of the will, in the guise of what he calls "second-order volitions," is a sufficient condition for personhood.

Consider an individual who smokes a pipe and is addicted to pipe smoking. A "first-order desire" here might be the bare desire for the sensation of filling one's lungs with smoke from the tobacco burning inside the pipe bowl. There may also be other, associated first-order desires, such as the desire for sensing the aroma present when one is filling the bowl; the feeling and taste of the pipe stem on one's lips, teeth and tongue; and so on. This bare, first-order desire to smoke can take the propositional form "R desires to *x*."

A "second-order desire" is to be construed as a desire referring to the first-order desire. For example, where R desires to smoke but also has the desire to not desire to smoke (say, for health reasons), the desire to not desire to smoke is a second-order desire. In a situation where R experiences both desires but is moved by and acts on the second-order desire, Frankfurt says that R's second-order desire is the *effective* desire. Frankfurt understands this as R wanting R's second-order desire to be R's will. In this case, where the second-order desire comes to be R's will, Frankfurt terms this a "second-order *volition*," which he says is a sufficient condition for personhood. In Frankfurt's terms, a "wanton" (W) is someone who doesn't care about W's will, which is clearly not the case for R. Wantons have first-order desires but are not persons because they have no second-order volitions (albeit it is possible that they have second-order desires). Freedom of the will amounts simply to making one's second-order volition(s) one's will.

A chief benefit, according to Frankfurt, of this interpretation of freedom of the will is that it implies moral responsibility for the actions that R takes when acting on R's second-order volitions. Where R has the will R wants to have, and acts on this will, R is taken to be morally responsible for the actions R commits.

Another important contributor to the literature on persons is Daniel Dennett, who makes a distinction between metaphysical persons ("roughly, the notion of an intelligent, conscious, feeling agent") and moral persons ("roughly, the notion of an agent who is accountable, who has both rights and responsibilities") (Dennett 1976, p. 176). Though Dennett focuses for the most part on the conditions for metaphysical personhood, he does say that the concept of a person is "inescapably normative." Shy of drawing the conclusion that *the* set of necessary and sufficient conditions for personhood will never be fully articulated, he does lend some voice to a few of the conditions he considers necessary.

The six conditions Dennett delineates are: consciousness (being the subject of intentional predicates); rationality; being the object of a certain attitude (having a *personal* attitude taken toward one); the ability to reciprocate this attitude; verbal communication; self-consciousness. According to Dennett, to be rational is just to be Intentional, and to be Intentional is just to be the object of a certain attitude. These three conditions, says Dennett, are themselves necessary, though not sufficient, for the ability to reciprocate the personal attitude, which is itself necessary but not sufficient for the capacity for verbal communication, which is itself a necessary, though not sufficient, condition for self-consciousness, which is itself a necessary condition for moral personhood.

Some would say Dennett's last word is overly skeptical. Not only does he not believe the set of sufficient conditions for personhood will ever be known, and not only are the chosen conditions in some sense arbitrary, and not only is it sometimes impossible to recognize just who are persons, when problems of moral responsibility arise, "we cannot even tell in our own cases if we are persons."

MORAL CONSIDERATIONS

This section is devoted to the moral aspects of the concept of personhood. One important aspect of the topic of

personhood is the use of person in a moral sense. The central question, that is, What is a person? can be translated into the question, What must a being be like to have moral rights (and moral responsibilities)? Setting off "and moral responsibilities" in parentheses here is meant to highlight the problem of assigning moral responsibilities to such beings as human infants; many, if not all, nonhuman animals; and perhaps those humans who are, say, in the late stages of Alzheimer's disease. While there are many who argue that these *persons* have moral rights, there is scant literature proclaiming their having moral responsibilities. This suggests a further question about moral personhood, to wit, whether a person can be the bearer of rights but not responsibilities.

MICHAEL TOOLEY'S THEORY. Tooley writes: "The question of what beings it is seriously wrong to destroy is one of the central questions of ethics." The question covers human as well as nonhuman beings. It applies to human fetuses, newborns, the mentally/cognitively challenged, the criminally insane, sociopaths, and those in the throes of diseases that impair brain activity. It also covers dogs, cats, giraffes, dolphins, whales, chimpanzees, gorillas, trout, sharks, trees, birds, and alligators. The question is distinctly *not* kind-, type-, or species-specific.

While the final goal in Tooley's work on the concept of personhood appears to be discovering whether abortion and infanticide are morally permissible, his work is distinctively metaphysical. He seems to believe that a person may be defined as a being who possesses at least moral rights (and perhaps moral responsibilities), and he sees that the analysis of these concepts requires laying out the concepts closely associated with these. However, Tooley has certain other questions in mind as he analyzes various conditions for personhood. Take the example of *rationality* as a suggested condition for personhood. He asks whether a being could rightly be thought a person who lacked the capacity for rationality. On the heels of this is the pointed question about whether it would be seriously wrong to destroy a being who was rational (staying with the example). It is this question that places his work squarely in the area of the moral aspects of personhood rather than the metaphysical. Or, if one prefers, any analysis of the moral aspects of personhood will automatically require metaphysical analysis as well.

Tooley runs through many of the suggested conditions for personhood, analyzing them in terms of whether they are necessary and/or sufficient conditions. Four of these suggestions are that a person is: (1) a subject of nonmomentary interests; (2) an entity that possesses rationality; (3) an entity that is capable of action; (4) an entity that possesses self-consciousness.

A brief sketch of Tooley's treatment of these conditions is as follows: As a subject of nonmomentary interests, an individual will have the capacity for a host of desires, the total set being in some sense "unified." While Tooley is not identifying interests with desires, he is making the case that desires may be inferred from interests. This is as it should be when interpreting interests in such a way that the subject can be said to *be interested*, as in "Don is interested in astronomy." However, it is more difficult to make sense of the idea of interests here when the meaning of *interest* has to do with what is in an individual's benefit, as characterized by the sentence "As an astronomer, it would be in Don's interest to study mathematics." While the former meaning of interest, allowing the inference to desires, would not seem to have the relevant moral sense, Tooley brings in moral significance by associating this concept of interest with the representation of the item of interest in consciousness. In the end, Tooley says that persons may be identified with "entities that have desires that are interrelated in such a way that the entities can be viewed as subjects of nonmomentary interests."

As to whether a being in possession of rationality is a person, Tooley takes the view that the relevant sort of rationality to be discussed has to do with what is called *agency*, where an agent is an enduring substance of a mental nature, with the capacity for deliberative reason-based action. Rightly claiming that there is little disagreement that this sort of rationality is insufficient for personhood, he argues that neither is it necessary. Though Tooley does not believe it plausible that rationality necessitates personhood, he does allow that any being who is rational and possesses nonmomentary interests is a person. Even the addition of a relevant form of free will, or the capacity for rational deliberation, is not enough to make rationality itself a necessary condition for personhood.

Tooley's third suggested condition for personhood is that of having the capacity for action. The name for anyone capable of action is *agent*, and Tooley claims there is little disagreement whether being an agent is a sufficient condition for being a person; it is. It is not, however, a necessary condition, according to Tooley. One important concern he brings up here is that if agency involves what is called a libertarian free will, then if universal determinism should turn out to be true, even normal adult human beings would not be persons. Tooley's reasoning on this is that even if it should be the case that all events are deter-

mined, that fact would not lead to the conclusion that it is not seriously wrong to destroy a normal adult human being. But now, on an account of agency that does not necessarily involve the possession of free will, Tooley presumes that the agent will possess nonmomentary interests. Since these sorts of interests have already been argued to be unnecessary to confer personhood, adding these to agency will not have the result of necessitating personhood on an agent so characterized.

The last suggested condition for personhood analyzed by Tooley is self-consciousness, which he argues is neither necessary nor sufficient for personhood. It is not necessary because there could be an individual who was aware of a continuing self but not in possession of this awareness *qua* individual continuing self. Self-consciousness is not a sufficient condition, according to Tooley, because it is conceivable that some individual may well be self-conscious but not be a subject of either momentary or continuing interests. For all this, however, it appears that Tooley would agree with the general consensus that it would be seriously wrong to destroy such an individual.

OTHER AREAS, OTHER CONCERNS. The area of medical ethics has produced by far the greatest amount of work on the concept of personhood. And within this field, the question of the status of the fetus has generated the most debate. The issue here is whether or not a fetus is a person in the moral sense of that term, that is, whether the fetus has a right to life. As is clear, this is but one issue in the abortion debate; yet it has generated as many books and papers as any topic in contemporary moral philosophy. The question of the moral status of the fetus characteristically revolves around discussions as to whether the fetus possesses any of the suggested conditions for personhood. Early term fetuses, whose brains have not developed sufficiently for, say, consciousness and rationality, are widely agreed to be nonpersons (with the notable exception that the religious contingent—specifically Roman Catholics—will not accept this conclusion, arguing that a fetus is a person from the moment of conception). A great controversy still surrounds mid- and late-term fetuses because it is simply unclear what their capacities are, and it appears an important question whether these individuals are more or less like nonhuman animals usually denied personhood.

Another interesting debate centers on the fetus being a *potential* person. The issue is whether a being who is going to be a person in the natural course of events should be treated as a person prior to becoming what it will be. One of the considerations that makes this question so significant is that there seems to be little relevant difference between a very-late-term fetus and a newborn infant. If such a fetus is not a person, that is, fails to possess self-consciousness, rationality, free will, and so on, then it would appear that the newborn is not, either. But this conclusion is one very few people have been willing to draw. (Tooley's work on potential personhood, in *Abortion and Infanticide*, is crucial reading.) A significant point made by some people on this topic is that the infant, upon birth, becomes a member of the specific community into which it is born whereas the fetus is not yet a member. It is somehow thought that having seen, held, and fed the infant are *attachment factors* leading to the community seeing the infant as a person. Such is not the case with even a late-term fetus.

Another question one can ask is whether people who commit heinous crimes lose their status as persons in the moral sense. This sort of case brings out clearly a distinction between the legal and moral senses of the concept of personhood. Under the law, a murderer/rapist can, in certain circumstances, retain the right to life (that is, not be sentenced to death). One argument many opponents of the death penalty have used is the following: premise 1: the individual sentenced to death under the law has a moral right to life, premise 2: no law can abridge a moral right, conclusion: the death penalty violates an individual's moral right to life. It is easy to see how this argument might be run if one accepts the conditions for personhood outlined above, to wit, self-consciousness, rationality, the ability for complex communication, free will, and so on. The committing of atrocious crimes would not appear incompatible with the agent possessing these characteristics.

However, if other necessary conditions are added to the list, such as the concern for others and respect for persons, it is more difficult to see how anyone could commit such crimes and at the same time maintain this person also had respect for others. Where the moral sense of person is defined as "a being with moral rights and responsibilities," the way would be open to argue that the death penalty *is* morally permissible. From this perspective, the conditions of personhood have significant practical impact.

Finally, the issue of animal rights has become one of the most widely debated issues of our time. Opponents argue, to a person, that nonhuman animals are nonpersons, though no one this writer is aware of argues that therefore we can treat nonhuman animals anyway we want (such as causing unnecessary pain). Proponents

sometimes argue that many nonhuman animals display characteristics matching a fair number of the suggested conditions for personhood. For example, some will say the neighbor's dog is conscious, displays rational behavior, can engage in fairly complex communication, and has a large measure of free will. This is to say that these animals possess very important characteristics thought to be relevant for designating adult humans as persons. Unless people will assent to some form of speciesism, they say, people must admit that these animals need to be treated as persons. This is at least sufficient, it is believed, to make it seriously wrong to harm the animal.

An interesting topic in animal rights, where the concern is whether nonhuman animals are, or should be, considered persons, is the question whether persons, in the moral sense, are beings who do have both moral rights as well as moral responsibilities. It is never argued that the neighbor's dog has moral responsibilities. This being the case, proponents of animal rights are never proponents of animal responsibilities. Even if there are cases where a person seemingly has a right without there existing a corresponding responsibility, it remains an open question whether these cases speak to the essential issues regarding the questions of personhood.

See also Abortion; Baker, Lynne Rudder; Dennett, Daniel C.; Frankfurt, Harry; Locke, John; Rights; Strawson, Peter F.; Thomas Aquinas, St.

Bibliography

Adams, E. M. "The Concept of a Person." *Southern Journal of Philosophy*, XXII(4), 1985: 403–412.

Aquinas, Thomas. *The Summa Theologica*, Piana text. In *The Basic Writings of Saint Thomas Aquinas*, edited by Anton C. Pegis. New York: Random House, 1945.

Armstrong, Susan, and Richard G. Botzler, eds. *Animal Ethics Reader*. New York: Routledge, 2003.

Baker, Lynne Rudder. *Persons and Bodies: A Constitution View*. Cambridge, UK, Cambridge University Press, 2000.

Beauchamp, Tom L. "The Failure of Theories of Personhood." *Kennedy Institute of Ethics Journal* 9 (4) (1999): 309–324.

Becker, Gerhold K., ed. *The Moral Status of Persons*. Atlanta: Rodopi Press, 2000.

Becker, Lawrence C. "Human Being: The Boundaries of the Concept." *Philosophy and Public Affairs* 4 (4) (1975): 334–359.

Bertocci, Peter. "The Essence of a Person." *Monist* 61 (1978): 28–41.

Brandt, Richard. "The Morality of Abortion." *Monist* 56 (1972): 503–526.

Burke, Michael B. "Copper Statues and Pieces of Copper: A Challenge to the Standard Account." *Analysis* 52 (1) (1992): 12–17.

Callahan, Daniel. *Abortion: Law, Choice and Morality*. New York: Macmillan, 1970.

Callahan, Sydney. "The Moral Duty to the Unborn and Its Significance." In *The Silent Subject*, edited by Brad Stern. Westport, CT: Greenwood, 1996.

Carter, W. R. "Once and Future Persons." *American Philosophical Quarterly* 17 (1) (1980).

Castañeda, Hector-Neri. "He: A Study in the Logic of Self-Consciousness." *Ratio* 8 (1966): 130–157.

Chisolm, Roderick M. *Person and Object: A Metaphysical Study*. LaSalle, IL: Open Court, 1976.

Dennett, Daniel. "Conditions of Personhood." In *Identities of Persons*, edited by A. O. Rorty. Berkeley: University of California Press, 1976.

Duska, Ronald "On Confusing Human Beings and Persons." *Proceedings of the American Catholic Philosophical Association*, LIX (1984): 158–165.

English, Jane. "Abortion and the Concept of a Person." *Canadian Journal of Philosophy* 5 (2) (1975): 233–243.

FitzPatrick, William J. "Totipotency and the Moral Status of Embryos: New Problems for an Old Argument." *Journal of Social Philosophy* 35 (1) (2004): 108–122.

Fletcher, Joseph. *Humanhood: Essays in Biomedical Ethics*. Amherst, NY: Prometheus Books, 1979.

Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*, LXVIII (1) (1971): 5–20.

French, Peter. "The Corporation as a Moral Person." *American Philosophical Quarterly* 16 (1979): 207–215.

French, Peter. "Kinds and Persons." *Philosophy and Phenomenological Research* 44 (2) (1983): 241–254.

Garrett, Jan-Edward. "Persons, Kinds, and Corporations: An Aristotelian View." *Philosophy and Phenomenological Research* 49 (1988): 361–381.

Goodman, Michael F., ed. *What Is a Person?* Clifton, NJ: Humana Press, 1988.

Kitcher, Patricia. "Natural Kinds and Unnatural Persons." *Philosophy* 54 (1979): 541–547.

Langford, Glenn. "Persons as Necessarily Social." *Journal for the Theory of Social Behavior* 8 (1978): 263–283.

Lewis, David. "Counterparts of Persons and Bodies." *Philosophical Papers*. Vol. 1, 47–54. New York: Oxford University Press, 1983.

Locke, John. *An Essay concerning Human Understanding*, edited by P. H. Nidditch. Oxford, U.K.: Oxford University Press, 1975.

Lomasky, Loren E. "Being a Person—Does It Matter?" *Philosophical Topics* 12 (3) (1982): 139–152.

Meldon, A. I. *Rights and Persons*. Berkeley: University of California Press, 1977.

Morris, Herbert. "Persons and Punishment." *Monist* 52 (4) (1968): 475–501.

Noonan, John T. Jr. "An Almost Absolute Value in History." In *The Morality of Abortion*, edited by John T. Noonan Jr. Cambridge, MA: Harvard University Press, 1970.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Clarendon Press, 1984.

Regan, Tom. *The Case for Animal Rights*. Berkeley: University of California Press, 1983.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Rorty, Richard. "The World Well Lost." In *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press, 1982, 3–18.

Rosenkrantz, Gary S. "Reflections on the Ontological Status of Persons." *Philosophy and Phenomenological Research* 65 (2) (2002): 389–393.

Scott, G. E. *Moral Personhood: An Essay in the Philosophy of Moral Psychology*. Albany: State University of New York Press, 1990.

Singer, Peter. *Animal Liberation: A New Ethics for Our Treatment of Animals*. London: Cape, 1976.

Singer, Peter. *Practical Ethics*. Cambridge, U.K.: Cambridge University Press, 1979.

Strawson, P. F. *Individuals: An Essay in Descriptive Metaphysics*. Garden City, NY: Doubleday, 1963.

Tooley, Michael. *Abortion and Infanticide*. Oxford, U.K.: Clarendon Press, 1983.

Weiss, Roslyn. "The Perils of Personhood." *Ethics* 89 (1978): 66–75.

Werhane, Patricia H. *Persons, Rights, and Corporations*. Englewood Cliffs, NJ: Prentice Hall, 1985.

Young, Frederic C. "On Dennett's Conditions of Personhood." *Auslegung* 6 (1979): 161–177.

Michael F. Goodman (2005)

PERSPECTIVE REALISM

See *Realism*

PESSIMISM AND OPTIMISM

"Pessimism" and its opposite, "optimism," are only secondarily philosophical theories or convictions; primarily they are personal opinions or attitudes, often widely prevalent, about the relative evil or goodness of the world or of men's experience of the world. As such they vary with the temperaments and value experiences of individuals, and with cultural situations far more than with philosophical traditions.

Both pessimism and optimism in the above sense may be reactions to experiences that vary in scope and content. Four types of reactions or judgments may be distinguished: (1) psychological or anthropological (involving judgments about the dominance of evil or good in one's own experience or in human experience generally); (2) physicalistic (judging the physical world to be dominantly evil or good); (3) historicistic (based on appraisals of the evil or goodness of a historical or cultural period or of the forces and institutions that determine history); and (4) universal, or cosmic (involving judgments about the dominance of evil or good in the universe as a whole).

Since the issue of the goodness or evil of human life involves belief in beneficent or malevolent forces upon which man's well-being is dependent, optimism and pessimism are prominent aspects of religious beliefs, and these beliefs may involve many or all of the above types of judgments.

Philosophical pessimism and optimism result from the critical analysis and clarification of judgments of the dominance of good or evil, an evaluation of the experiences upon which these judgments are based, and the presentation of reasons to justify or refute such statements. There is widespread doubt whether the terms *optimism* and *pessimism* are sufficiently precise for philosophical purposes and also whether optimistic and pessimistic beliefs are philosophically justifiable. This article will be concerned chiefly with philosophical formulations and arguments for optimism and pessimism with some reference to their manifestations in religion.

Optimistic and pessimistic attitudes and theories are much older than the terms used to describe them. The term *optimisme* was first used in the Jesuit journal *Mémoires de Trévoux* in 1737 to designate Gottfried Wilhelm Leibniz's doctrine (which appears in his *Théodicée* and in other of his philosophical writings) that this is the best of all possible worlds. Leibniz himself used the term *optimum* in a technical sense that applied to the unique maximal or minimal instance of an infinite class of possibilities, and he held that this principle of the optimum was applied by God in the creation of the world. *Optimisme* was admitted by the French Academy to its dictionary in 1762. The first known appearance of the term *optimism* in English was in 1759, also in reference to the system of Leibniz. *Pessimism* came into general use only in the nineteenth century, although its first known appearance in English was in 1795 in one of Samuel Taylor Coleridge's letters.

The superlative form of the Latin adjectives *optimus* and *pessimus* is not generally justified by any form of philosophical optimism or pessimism. It is true that Leibniz defended an optimal position in the formula "the best of all possible worlds," but this use of the superlative did not prevent his acknowledging the existence of much evil—indeed, the necessity of evil in all finite existence. Similarly, Arthur Schopenhauer affirmed that this is the worst of all possible worlds, but his chief philosophic concern was with finding a way of salvation from the evil of the world through art, a morality of sympathy, and philosophic and religious contemplation. The most thor-

oughgoing philosophical pessimist of the nineteenth century, Eduard von Hartmann, held that this is the best of all possible worlds; yet evil necessarily outweighs good in it, and it would be better if there were no world at all.

The philosophical issues might better have been served by the comparative forms "meliorism" and "pejorism" ("betterism" and "worsism"). Although the verb forms "meliorate" and "pejorate" did appear in the sixteenth and seventeenth centuries, respectively, "pejorism" has found no acceptance, while "meliorism" has been used, following William James, to express the view that although the world is a mixture of good and evil, it can be bettered by man's moral efforts to improve it.

RELIGIOUS AND PHILOSOPHICAL ISSUES

Optimism and pessimism are thus relative terms; the former theory undertakes to give philosophical reasons for assuming that in whatever horizon or context is involved, good preponderates over evil, while the latter theory attempts to show that evil preponderates over good. The arguments in each case may be efforts to generalize from experiences of good and evil, or they may, and usually do, also involve a priori factors, basic definitions, and theological or metaphysical doctrines.

EMPIRICISM AND RATIONALISM. A primary consideration in discussing optimism and pessimism is the definition and criteria of good and evil. Empiricists have generally adopted a hedonistic definition of good, and hedonism has frequently ended in pessimism: The universe seems not to be constituted to provide man with more pleasure than pain. But it has proved difficult to reduce normative judgments of value to the psychological measures of pleasure and pain, joy and sorrow, or satisfaction and dissatisfaction. Other criteria are also involved—for example, the conservation or destruction of life, the progress or decay of cultural institutions and values, human freedom and bondage (in various senses), and the just control of power.

While empiricism shows an inclination toward pessimism (and skepticism), rationalism operates with normative principles that have an affinity with affirmations of the identity of reality and goodness. Nevertheless, exponents of hedonism are driven to recognize qualitative distinctions between pleasures and pains and the complex interplay of pleasures and pains that makes possible greater goods, while beneath the most rational and optimistic systems of modern thought lurks the shadow of fear, if not of despair. Leibniz wrote during a period of devastating European wars and intended his thought to serve as the foundation for a European culture that would protect Europe against the threat of a new barbarism. Voltaire, Edward Gibbon, and Pierre Maupertuis expressed the same fears, and in America, Benjamin Franklin, Thomas Jefferson, Alexander Hamilton, and John Adams had forebodings of the dangers of revolution and the collapse into barbarism that might follow a failure to establish a sound political order.

RELIGION. Religion involves both optimistic and pessimistic aspects. Since the essence of religion is salvation from evil, an optimistic element is essential to it; yet not all individuals or groups are saved. The magical component in religion is optimistic, since it promises success in the achievement of desired values; yet the failure of religious rites or prayers is common enough to support pessimism. Salvation is postponed to a future life, and the present world is viewed as a vale of tears, or as the historical conflict between good and evil, or as a source of desires to be resisted, or as an illusory order that possesses no substance. Yet in all religion there is also a joyous world-affirming element that expresses itself in community life and mystical or prophetic exaltation. Eschatological religions combine pessimism about a temporal world that is destined to end with joyous optimism about the new life that will follow.

METAPHYSICS. If hedonistic criteria of good and evil are a common source of pessimism, those systems of thought that hold to an ultimate identity of existence and value are the mainstay of optimism. Two philosophical convictions in particular have supported optimistic convictions in Western thought. One rests upon the Platonic and Aristotelian ideal of the perfectibility of man. It regards all the powers of man as capable of control and harmonization (without great resistance from senses and impulses). The other is metaphysical but has the same sources. Regarding the universe as a hierarchy of being and goodness, ordered from infinite perfection though all levels of particularization to the total formlessness of matter, or mere potentiality, it finds all evil and error to consist in a negation or privation of being.

Other traditions also have a bearing upon optimism and pessimism. Efforts to interpret the universe as normatively indifferent (traditional materialism, for example) usually end in pessimism. Dualisms of various kinds, on the other hand, whether they distinguish between cosmic powers of good and evil or between a real order of value and a phenomenal order of fact, tend to end in optimism.

SCIENCE. Finally, natural science has presented considerations that affect the problem of optimism and pessimism. Fires, earthquakes, floods, storms, diseases, and, ultimately, death have always been regarded as evil because they interfere with human purposes and hopes. But the theory of natural selection and the second law of thermodynamics, which has been held to imply an end to the universe at a finite time in the future, have put the issue of the destructiveness of natural powers, animate and inanimate, on a more objective basis by casting serious doubts upon the possibility and the goodness of evolution and progress.

HISTORY OF PESSIMISM AND OPTIMISM

RELIGIOUS PESSIMISM. Religion is relevant to the problem of optimism and pessimism insofar as it offers salvation to men, evokes attitudes of world-affirmation and world-renunciation, and involves beliefs about the place of man and his hopes in the world. In this sense Schopenhauer was justified in calling religion the metaphysics of the people. Most religions combine a certain joyous response to divine grace with a sense of anguish and guilt at man's failures. Most advanced religions reflect a deeply rooted intuition of natural and historical evils and of the human limitations to which man is subject.

Indian thought. When the Brahmanic tradition in India emerged from the earlier Vedic religious forms, it partly concealed an underlying pessimism with the doctrine of maya—namely, that the world in which man suffers is a world of illusion, and release follows from recognizing this and the supplementary truth that man's true nature is one with the Brahman. This Brahmanic tradition was supplemented by a popular polytheistic religion that combined an easy tolerance of the diversity of natural delights and griefs with a singleness of purpose in carrying out those disciplines (whether physical, moral, intellectual, or mystical) that assure the self of its ultimate release and redemption. The fatalistic doctrine of the eternal cycle of rebirth, together with the doctrine of karma, intensifies a mood of pessimism, since this cosmic law of justice sentences most men to relive the deceptions of life again and again.

This element of pessimism implicit in Hinduism became the driving force of Buddhism in its various forms. The fourfold truth revealed to Gautama under the bo tree begins with the misery of human existence, caused by desire, and offers as salvation only the renunciation of desire and the attainment of that state of negation which is the highest bliss.

Western religions. As the Eastern religions show, the religious source of pessimism is to be found in the emergence of man's self-consciousness at a level at which he feels his isolation and estrangement in a world in which sickness, suffering, and death interfere with, and ultimately nullify, his hopes for a desired future. This mood showed itself in early Babylonian and Egyptian literature, as well as in the Hebrew Scriptures and in the Greek conception of life as being lived in the shadow of a fate (*moira*) from which death itself fails to offer a complete escape. Homer, although generally healthy-minded, judged that "there is nothing more wretched than man, of all things that breathe and are" (*Iliad* XXIV, 446ff.), and Sophocles wrote, in *Oedipus at Colonus,* "Not to be born is the most to be desired; but having seen the light, the next best is to go whence one came as soon as may be." In the Old Testament, the books of Job and Ecclesiastes reflect the same struggle with the meaninglessness of life.

However, the Judeo-Christian tradition is generally regarded as being optimistic. It applied a theistic view of Providence first to the history of a "chosen people" and then more universally to the moral interpretation of human history and of divine justice. The meaning of history is the redemption of God's people and, more generally, the Kingdom of God or the Reign of Grace. Moreover, although the Hebrews had only a vague conception of life after death, Christianity offered the assurance of a blessed life—an assurance based neither upon a concept of strict justice, as in karma, nor upon works, but on divine Grace.

However, much Christian eschatology has condemned the present world to destruction and the people in it to judgment and condemnation. The division of people into saints and sinners has often comforted those conscious of their sainthood but has not generally strengthened the ideal of a great community of love. Doctrines of original sin and predestination of the damned, of apocalyptic horrors terminating history, and of the complete alienation of man from the world (the despair of life) have been a part of the Christian tradition and have been revived in our own time, when the consciousness of guilt and of alienation has been reinforced by the secular study of modern man.

Thus, most religion, in different contexts, emphasizes both good and evil in man, the universe, and history.

ANCIENT PHILOSOPHICAL VIEWS. The Greeks, whose thought turned about the polarities of matter and form, impulse and reason, power and justice, freedom and order, and the transient and the permanent in experience,

came to conclusions that have influenced all later discussions of the problem of good and evil in Western culture. When Friedrich Nietzsche condemned Socrates for making the Apollonian mood supreme in Greek art and thought, he attributed to him a type of serene intellectualistic optimism that has formed much of Western culture, particularly through its elaboration and systematization by Plato and Aristotle, who by ultimately identifying existence and value and supporting the ideal of rational perfectibility provided the philosophical grounds for Western optimism. But Plato was not so one-sidedly optimistic as Neoplatonism later became. The *Republic,* for instance, recognizes the possibility for man and society to attain justice and happiness, but it imposes harsh conditions for their attainment and is pessimistic about their ever being achieved by more than a select few.

In Hellenistic and Roman thought the nature of evil was a persistent problem that was shared by Epicureans, Stoics, Skeptics, and eclectics. Skepticism is often regarded as the intellectualistic counterpart of pessimism, but it has also often been the basis for an optimistic fideism. Although Epicureans and Stoics answered the question of the nature of evil differently, both the qualified hedonism of the one and the rejection of all external goods and emphasis upon self-sufficiency of the other tended to support a cultured tranquillity of contented, sometimes even grateful, acceptance. Both denied the evil of death, and the Stoics denied the evil of pain as well. While the Stoics relied upon determinism, and the Epicureans upon indeterminism, both denied that the gods were in any way connected with, or cognizant of, man's good. From Plutarch's *De Stoicorum Repugnantiis* (first century CE) to Vanini's *Amphitheatrum Aeternae Providentiae* (1615), the Stoics were charged with attributing evil to divine Providence, while the Epicureans grounded their conception of the contentment of the wise man upon his freedom from interference by the gods.

The decline and fall of Rome brought to consciousness a new dimension of pessimism—the despair evoked by the collapse of a historical order that had claimed eternity and universality. The relativity of good and evil to historical change provided the individual with a mode of adjustment to the evils of social and institutional decline. St. Augustine's great adaptation of Platonism to a Christian solution to this problem has been the source not only of most later religious optimism, but also of the great theodicies of the West, from the medieval and Renaissance Platonists to Leibniz and G. W. F. Hegel.

EARLY MODERN VIEWS. The Middle Ages have often been regarded as having been clouded with pessimism (they provided Hegel with the cultural type that he described as "the unhappy consciousness"), while the Renaissance and seventeenth century have been regarded as comparatively optimistic, culminating in baroque exuberance. But recent scholarship views the medieval and Renaissance periods as a cultural continuity moving toward "modernity." In the face of a deep concern for the physical, social, and moral evils of Europe, intellects in both periods were engaged in a concerted effort to lay a rational Christian foundation for human happiness and harmony. While the political and social conditions varied, and the ideal of transformation changed from an eschatological revolution to continuous progress, Greek and Roman intellectual traditions continued to limit the philosophical effort to synthesize science, moral rationalism, and religious faith. Science and technology, nationalism, new ideals of individual freedom and toleration, and contact with new lands and cultures shifted and enlarged the scope of inquiry and intensified the problems, but the differences between Peter Abelard, Thomas Aquinas, John Duns Scotus, and William of Ockham on the one hand, and René Descartes, Benedict de Spinoza, Francis Bacon, and John Locke on the other are far more superficial than the continuity of their problems and their tradition.

Seventeenth-century discussions of the dominance of good or evil were affected by the new perspectives on human life that evolved in the Renaissance—notably, the emphasis upon individualism; the conflict about the nature of human freedom; the problem of the control of political power, which resulted from the collapse of the medieval synthesis and the multiplication of small states; and the ideal of a rule of reason, strengthened by the successful combination of mathematics and experimentation in the scientific mastery of nature.

Developments in psychology. The discussion of optimism and pessimism was affected by two developments in psychological thought: Galen's doctrine of the four humors was applied to man's reactions to good and evil, and there was a wide recognition of the role of the affections and appetites in human life. A comparison of Albrecht Dürer's famous engraving of Melancholia (1514) with Robert Burton's *Anatomy of Melancholy* (1621) is revealing. In Dürer's time the dominance of the melancholy humor was held to be the source of contemplation and therefore of mathematical and other forms of learning; Burton treated melancholia as pathological and analyzed its types, causes, and cures. Unfortunately, there is no work analogous to Burton's erudite essay that deals

with the dominance of the opposing humor, the sanguinary. But the use of the humors to explain pessimism and optimism initiated a long tradition of distinctions that includes the Earl of Shaftesbury's and Jean-Jacques Rousseau's theories of the natural affections, the *Weltschmerz* and *Weltfreude* of the German romantics, and after Schopenhauer, the psychoanalytic classifications of Sigmund Freud and Alfred Adler and the psychological typologies of worldviews by William James, Wilhelm Dilthey, Max Scheler, and others.

A closely related trend was the growing recognition of the role of the affections in determining human attitudes and conduct. The third book of Luis Vives's work on the mind (*De Anima et Vita Libri Tres,* 1538) was an important source for later attempts by such thinkers as Descartes, Spinoza, and Thomas Hobbes to explain human actions in terms of feeling and desires. In Hobbes the result was a pessimistic theory of human nature; in Michel Eyquem de Montaigne, Blaise Pascal, and thinkers of the libertine tradition, it was a relativization of human ends that undermined the absoluteness of goods and evils; but in the thinking of Vives himself and in the rationalistic tradition of the seventeenth century (for example, Descartes, Spinoza, and Leibniz), an idealistic optimism resulted from the doctrine that the affections are docile and readily moldable into socially constructive attitudes.

Politics and history. The problem of power (particularly political power) and its responsiveness to reason was a second noteworthy development affecting the estimation of good and evil. Machiavelli had formulated the fundamental theory of a *raison d'état* in a way that provided pragmatic support for the principle of the divine right of rulers. The series of disastrous wars that swept over Europe, however, intensified a mood of eschatological expectation and heightened the fear or hope of revolution and an overthrow of the existing order. The transfer of the eschatological hope from an afterlife to the temporal world, and the resulting faith in human progress, were the result primarily of the increase of scientific and technological knowledge and the wider expansion of faith in reason. Hobbes entirely restricted his realistic definition of justice as the power of the strongest to the limits of the present historical order, thus secularizing St. Augustine's pessimistic appraisal of the City of Man and providing a modern ancestry for pessimistic interpretations of history.

Rationalism. From the metaphysical point of view, however, the rationalistic tradition of the seventeenth century may be regarded as optimistic; it constituted an effort to bring the real into harmony with the ideal or the normative. This effort concentrated on the law of nature and on the individual's relation to the absolute source of power and wisdom. In Descartes, human passions are regarded as supporting the ideal of *generosité* and *honnêteté*; in Spinoza, actuality is generalized into possibility, and passive affections are shown to be imperfect but corrigible through active affections; in Leibniz, truths of fact are held to be grounded in truths of reason, if we could only completely analyze the former. This optimistic doctrine of reality is supported in these thinkers by the conviction that evil is finitude or limitation and that as our ideas move from confusedness, indistinctness, and inadequacy toward clarity, distinctness, and adequacy, the goodness of the world and of our life is brought to light in an absolutely convincing way. Not all thinkers, of course, accepted this optimistic metaphysical resolution of the problem. Pascal was driven by his perception of the finiteness of man and the terror in which this finiteness involves him to a philosophy in which the heart, not the intellect, provides knowledge about ultimate reality. Pierre Bayle had recourse to a combination of skepticism and Manichaean dualism, while Locke was attracted on the one hand to libertinism, pluralism, and toleration, and on the other hand to arguments for faith in a determining divine Providence.

LEIBNIZ AND THE ENLIGHTENMENT. Gottfried Wilhelm Leibniz (1646–1716) is generally regarded as the outstanding modern philosophical optimist. His *Théodicée* (1710) is a prolonged argument for the rationality of Christian faith, the reasonableness of creation, and the view that this is the best of all possible worlds. The argument of this work is supported by a large body of writings that aimed at a *philosophia perennis* (a synthesis of the truth in all of the classical systems of thought) as well as a harmonious ordering of scientific, philosophical, and theological truth. This philosophical system, in turn, was intended to serve as the ethical basis for the great Leibnizian projects for engaging the leaders of Europe in the restoration of peace through the advancement of science and technology, the reform of the law, the perfection of logical and mathematical tools of learning and a universal encyclopedia, the reuniting of the churches, and the Christian conquest of the pagan parts of the world. Thus, Leibniz's optimism, although grounded on one of the most remarkable philosophical systems of Western thought, was also ideological; it aimed at concerted action in a variety of related fields, and in this sense it presupposed a deep sensitivity to the existing evils that were to be overcome.

In general, Leibniz's argument is that the man of good will (*homo honestatis*) should find his greatest happiness ("toute la joie dont un mortel est capable") in the recognition that in spite of its glaring evils this is the best of all possible worlds, because its creation involved the fullest possible realization of the divine attributes. He should also recognize that there prevails in the world a divine harmony that requires evil not only for the full manifestation of the infinite greatness of the world's Creator but also in order that this evil may contribute to a greater good than would otherwise be possible. The conception of evil involved in this argument combines three theories: the privative theory (supported by Leibniz's essentialist metaphysics) that the complete notion or law of every individual monadic series is a finite combination (erected by God) of its own simple perfections; a legalistic moral theory somewhat inconsistent with this, according to which justice requires retribution for man's sins and compensation for man's suffering; and an aesthetic theory that finds limited evil necessary (like the dark parts of a painting) for the perception of a more complete and inclusive good. Leibniz's defense of God is brilliant, and the many editions through which his *Théodicée* passed in the original French and in Latin and German translations produced an extensive following on the Continent and even in England, where it may have influenced the optimistic thought of Lord Bolingbroke, Alexander Pope, and others. Yet his argument is defective, most notably in his failure successfully to reconcile human freedom and responsibility with the determinism of the divine creation, and in his general inclination to explain what is in terms of what ought to be. Many readers have agreed with Jean Guitton (*Pascal et Leibniz,* Paris, 1951, p. 121) that "one would have to change very little to transform this supreme joy (in the supreme goodness of things) into a radical despair."

Deism. The optimism of the eighteenth century, influenced by Leibniz's defense of God rather than by his more subtle metaphysics, was deistic, and much of its thought followed the five creedal points of Lord Herbert of Cherbury, who asserted an instinctive faith in the law of nature that dictates belief in one God, a divine order of justice, a moral imperative, individual immortality subject to a system of rewards and punishments, and a condemnation of "enthusiasm" as divisive and disruptive of true religion. The spirit of deism was activistic, sometimes revolutionary, and intent upon scientific progress and the dissipation of superstition. In this sense it was optimistic.

Maupertuis. The eighteenth century was also the breeding ground of modern pessimism. Voltaire's shocked reaction to the Lisbon earthquake and his satirical attack on the Leibnizian formula in *Candide* stimulated the change in mood, but even more significant was the influence of Pierre-Louis Moreau de Maupertuis (1698–1759), to whom both the utilitarian Jeremy Bentham and the philosophical pessimist Eduard von Hartmann were indebted for their conception of a "balance of pain and pleasure." In his *Essai de philosophie morale* (1749), Maupertuis proposed a measure of good and evil in terms of *plaisir* and *peine.* (The French terms, their English equivalents *pleasure* and *pain,* and the German words *Lust* and *Unlust* have somewhat different psychological connotations that must here be ignored.) Maupertuis defined these terms functionally: *Plaisir* is any "perception" that the soul prefers to experience rather than not to experience; *peine* is the opposite. An examination of life in terms of moments of pleasure and pain, Maupertuis concluded, shows in a frightening way how preponderant pain is. Life is a constant wish to change one's perceptions in order to achieve fulfillment and to see the intervening times destroyed (*anéantir*). But if God were to abolish these intervening periods from even the longest life, only a few hours would remain. "In the usual life the sum of evil is greater than the sum of well-being."

KANT. If the optimism of the Enlightenment found the goodness of creation revealed both in nature and in historical progress, the decline of this tradition and the growth of a new pessimism grounded in the romantic movement may be traced in the thought of Immanuel Kant. The *Versuch einiger Betrachungen über den Optimismus,* written in 1759, argued for the Leibnizian "best of all possible worlds" in two steps: first, there must be one possible world that is the best, and second, it is necessary that this existing world is that best of all possible worlds. Kant urged the faith that each human being, recognizing "that the whole is the best and everything is good for the sake of the whole," should find his small place in this world. But in his critical period, after 1781, he found the fact of evil decisive in invalidating the Teleological Argument and recognized a "radical evil" in man that prevents him from exercising the good will and doing his duty. In the short paper of 1785, *Muthmasslicher Anfang der Menschengeschichte,* Kant could only advise maintaining one's courage in the face of life's tribulations.

ROMANTICISM AND IDEALISM. The shift in attitude noted above deepened into the pronounced pessimism of the romantics, many of whose writings reflect a feeling of

overwhelming anguish at man's situation in the world. Johann Wolfgang von Goethe's early works (especially the *Sorrows of Young Werther*) reveal this *Weltschmerz,* as do the works of Heinrich Heine, Lord Byron, and Giacomo Leopardi. However, the German idealist philosophers struggled against it through various forms of voluntarism—a voluntarism that encompassed the cosmos in Johann Gottlieb Fichte, was involved in history through great individuals in Hegel, and developed into a theory of emerging personal creativity in the context of chaos in Schelling's philosophy of freedom. Thus, Eduard von Hartmann and Olga Plümacher were unjust to the influence of this *Weltschmerz* when they excluded it from consideration as a form of philosophical pessimism. In a real sense it anticipated, and was the historical forerunner of, the twentieth-century irrationalist philosophies and philosophies of despair.

SCHOPENHAUER AND VON HARTMANN. The greatest philosophical protagonist of the pessimistic tradition is, of course, Arthur Schopenhauer (1788–1860), who gave expression to it in the context of the Kantian distinction between a phenomenal nature and a real intelligible world in which the moral will and an interpersonal society of willing beings are primary. Schopenhauer interpreted the realm of phenomena as "illusory" and as the result of human conceptualization; the real world is irrational will-to-live, known intuitively through man's perception of his own nature. To discover this world is to recognize the ultimate and inescapable evil of existence.

Man's life, Schopenhauer held, is permanently condemned to be in bondage to the will-to-live. As the Indian thinkers discovered, the essential nature of every human life is desire, and this desire is never stilled, since even its satisfaction results in increased desire or ennui. The world as will, therefore, is unmitigated evil; good is illusory, but man, by his very nature as an intelligent, feeling animal, and facing inevitable death, is driven beyond this illusion to discover his own plight. This is therefore the worst of all possible worlds, since there is no good in it. The only escape is through renouncing will, but only the great artists, thinkers, and prophets are capable of doing this—and only in a finite and impermanent degree. There is, however, an ethics involved in this pessimism; it is the ethics of sympathy and amelioration of the suffering of one's fellows.

Von Hartmann. Eduard von Hartmann found Schopenhauer's pessimism to be the ultimate expression of a romantic *Weltschmerz* in which a sense of guilt over the quest for pleasure was implicit. Although he adhered generally to Schopenhauer's metaphysics (supplementing the will, however, with a parallel order of ideas, both will and ideas having their seat in the unconscious), he modified his own theory of conflict in nature by stressing the purposiveness of every individualized act of will. He also rejected the Darwinian theory of change through struggle and survival in favor of a theory of evolutionary creativity in which new forms arise in the germplasm of the old. In contrast to Schopenhauer's pessimism, von Hartmann claimed that his was a "powerful, energetic pessimism, filled with the joy of action," whose historical antecedent is to be found in Kant, not Maupertuis. This is not the worst of all possible worlds; the logical element (that is, the ideas) ensures that the world is a best possible world. Yet it would be better if there were no world at all, and this is in truth the end to which the universal will, spatialized, and individualized through the particularizations of intellect, is driving—the total negation of all will through the fulfillment of its purposes.

Although von Hartmann argued that his metaphysical system of the unconscious would be valid without his pessimism, it is apparent that the converse is not the case: his pessimism rests directly upon his metaphysics of the unconscious. Yet he supported his pessimism by a comprehensive examination of empirical arguments from neurology, psychology, and the history of culture. The optimistic illusion takes form in three stages: the belief first, that happiness is attainable in the present world; second, that there will be a future otherworldly life in which the good will be attained; and third, that the surplus of happiness will be achieved sometime in this world's future history. The transition from each stage of optimism to the next already involves a surrender of hope. Von Hartmann's refutation of optimism is not merely negative but consists of a constructive argument for three corresponding levels of pessimism, which he labeled empirical, transcendental, and metaphysical respectively. Transcendental pessimism involves the denial of life after death, a conclusion von Hartmann undertook to prove through a metaphysical argument for the inseparability of body and mind. Metaphysical pessimism is supported a priori by the inevitability of misery in a world of will individuated by ideas and by the total lack of feeling of the will after all existents have ceased to be. It is also shown, however, by the finiteness and ultimate failure of all the values of human life—particularly the ethical, religious, and aesthetic values.

It is in his argument for empirical or eudemonistic pessimism that von Hartmann showed his greatest skill in penetrating human motives and the interaction between

pleasure and displeasure in human action. Twelve arguments, cumulative in force, were offered for the preponderance of pain over pleasure. On the simplest level, the growing fatigue induced by nervous processes diminishes the effort to retain pleasures, and as the fatigue grows, it increases the resistance to pleasure. Moreover, most pleasure is merely the negative kind that results from the cessation of positive unpleasantness or pain; thus, it can in no way equal the unpleasantness that it terminates. Displeasure coerces consciousness in a way that pleasure cannot, since pleasure must consciously be sought and discovered and occurs only when there is conscious motivation or desire for it. In shared experiences of pleasure the sense of solidarity and sympathy may momentarily intensify that pleasure, but this intense pleasure is correspondingly sooner exhausted than unintensified pleasure. In shared suffering or displeasure this sympathetic response may also occur, but it is overbalanced by callous and egoistic reactions. Moreover, history shows that as cultures advance in sensitivity and refinement, this overbalance of suffering increases proportionally. Such arguments, von Hartmann held, conclusively establish an excess of *Unlust* that confirms eudemonistic pessimism.

In his late work on the history and foundation of pessimism (2nd ed., 1892), von Hartmann modified his theory through an analysis of the different measures of value (*Wertmassstäbe*), of which pleasure is only one, the others being purposiveness, beauty, morality, and religiosity. These independent measures of value in themselves point to an optimistic view of life. Thus, he now called his thought a "eudemonological pessimism" but a "teleologico-evolutionary optimism"; yet the new measures are themselves not unmixed with the subjective feeling dimension, so that we must conclude that the overall balance of pleasure in the world is negative.

Von Hartmann's influence. Unlike Schopenhauer's pessimism, which was slow in gaining acceptance, von Hartmann's *Die Philosophie des Unbewussten* (Berlin, 1869; 9th ed. translated by W. C. Coupland as *The Philosophy of the Unconscious,* 3 vols., London, 1884) met with an immediate favorable response because of the changing intellectual and cultural mood of the last half of the century. The worst effects of the industrial revolution had become too conspicuous to be overlooked; colonialism involved nations in guilt; utopian reforms frequently ended in disillusionment; socialism shifted from its philanthropic to its "scientific" stage (von Hartmann himself was one of the early critics of social democracy); Darwinism intensified the perception of suffering and struggle in animate nature; and the romantic mood collapsed into a new naturalism according to which man was held in bondage to social forces and unconscious powers beyond his control. Novelists such as Charles Dickens, whose early works radiated Mr. Pickwick's cheerful vision of life, turned to the wretchedness of life and the irreducible evil of actual educational, penal, and political systems. Nathaniel Hawthorne and Herman Melville in America and Thomas Hardy in England reflected different aspects of this pessimistic movement, which mounted in strength until it developed into the fin de siècle mood of disillusionment, mortification, and decadence described and criticized by Cesare Lombroso, Max Nordau, and others.

Several of von Hartmann's followers carried his pessimism to the limit of nihilism. Julius Bahnsen (1830–1881) analyzed the "dominance of the offended spirit" (*das angekränkelte Gemüth*) that is split by hate, malcontent, and horror, and Philipp Mainländer (pseudonym of Philipp Batz, 1841–1876) pushed pessimism to its ultimate conclusion in total annihilation. In his *Philosophie der Erlösung* (2 vols., Berlin, 1876–1886) Mainländer held that the will to annihilation (*Vernichtungswille*) is included in the nature of every individual being, inorganic as well as organic, and that the ethics of the individual is egoistic and implies virginity and suicide as means of world salvation (that is, annihilation).

Von Hartmann's pessimism, although more critical and balanced than Schopenhauer's, also received extensive philosophical criticism. James Sully in England, Johannes Volkelt, Johannes Rehmke, Hermann Lotze, and Gustav Fechner in Germany, the spiritualists in France, and William James and others in America replied in terms of a more positive voluntarism or a more positive theory of value, thus laying the basis for a restoration of constructive liberalism in the twentieth century.

NIETZSCHE. The influence of Schopenhauer upon Friedrich Nietzsche was described by the latter in detail and is well known. He agreed with Schopenhauer's view that life is filled with suffering and a preponderance of evil, but rejected his ethics of resignation and of sympathy that was based upon it, as he also came to reject the metaphysical doctrine of will upon which it rested. Instead, Nietzsche's doctrine of the Dionysian man, or the superman, demanded a vigorous affirmation of life and power that would transcend both the "weakness doctrines of optimism" and tragedy as "the art of metaphysical comfort." In his "Versuch einer Selbstkritik" (1886; English translation in *The Philosophy of Nietzsche,* Modern Library edition, New York, pp. 934–946) Nietzsche corrected his earlier romantic reliance upon the ideal of

"a pessimism of strength" that he found in Greek tragedy (*The Birth of Tragedy*), replacing it with an affirmation of man's powers of joyous creativity—the "laughter of Zarathustra." Although Nietzsche's ideal of a life "beyond good and evil" is ambiguous and easy to misread, he clearly transcended traditional conceptions of pessimism and optimism, pressing from the conceptual to the realm of personal living and valuing. His superman is a mixture of the rejection of accepted contemporary values, a rigorous discipline of the self in loneliness, and the joy of creativity and the hope of a new aristocracy of creative individuals.

Nietzsche's criticism of modern culture as nihilistic is beyond pessimism in the same sense that his ethics is beyond good and evil. Abstract theories of the balance of good and evil fall far short of reflecting the plight and the opportunity of modern man, upon whose will to power the civilization of the future must rest.

SANTAYANA AND FREUD. Two thinkers who differed greatly in their theoretical and practical approaches to human problems, George Santayana and Sigmund Freud, developed pessimistic theories that were similar in important respects to the pessimism of Schopenhauer. (Freud arrived at his pessimism independently and did not read Schopenhauer until late in life.)

Santayana found in metaphysical matter what Schopenhauer found in will—the ultimate ground of all permanence, power, and life and therefore the ultimate ground of the tragedy that is involved in man's efforts to live the life of reason and spirit. Through concrete personal vision Santayana transcended the old debate between optimism and pessimism. Unlike Nietzsche, he found his personal resolution of the problem of evil not in the egocentric ideal of the superman but in an ideal of stoic acceptance and self-sufficiency.

In Freud's work the libido and, later, the id play a role similar to that of the will in Schopenhauer's system. The failure to gratify the impulses emanating from the id produces basic dislocations in the "libido economy" and thus leads to suffering and illness. In *Das Unbehagen in der Kultur* (Vienna, 1930 [1929]; translated by Joan Riviere as *Civilization and Its Discontents,* London, 1930) Freud traces human suffering to three sources—the superior power of nature, the decay and death of our own bodies, and the shortcomings of social relations and institutions. Of these, the first two are insurmountable, and the third inevitably results in unhappiness and alienation from man's culture. Moral judgments are merely "the effort to support illusions with argument." The illusory world of subjective imagination and thought sometimes offers successful sublimations and corrections, but the ultimate way to soundness can be found only (if at all) by a return to the natural and cultural roots of our being through psychoanalytical techniques. In an earlier work, *Die Zukunft einer Illusion* (Vienna, 1927; translated by W. D. Robson-Scott as *The Future of an Illusion,* London and New York, 1928), Freud held out much hope for this ideal through the elimination of religion, which he saw as likely to accompany the progress of science.

THE TWENTIETH CENTURY. In the twentieth century, with its dislocation and destruction of human life and values, the tremendous potentialities of its technological advances, its moral and cultural uncertainties, and its rifts in the texture of human society, the problem of optimism and pessimism shifted from an attempt to determine the relative goodness and badness of the world to an attempt to face the plight of modern man—his situation and his powers and resources for achieving good. This is a shift from conceptual modes of assessing the goodness of man, nature, and the universe to cautious nominalistic and phenomenological analysis of the individual.

It is true that a moralistic optimism has found strong defense and influence through the work of William James and John Dewey, while Alfred North Whitehead and others have offered metaphysical support of rationality, creativity, and the discovery of values in general. On the other hand, Bertrand Russell, in "A Free Man's Worship" (1903), gave moving expression to a naturalistic pessimism that regards man's existence in an indifferent universe as brief and without meaning, yet exhorts him to resist these natural powers with all the force of a living and vigorous faith in himself and in the powers of man. More commonly, the prevailing temper is to ignore the natural order as being neutral toward good and evil, and to show concern rather for the human person as a self-conscious being cast in a given historical situation. Man's natural environment, which John Dewey (in agreement with Hegel) found to be an aspect of the situation in which man is to achieve his freedom, is now taken by many as an aspect of the situation into which man is "thrown," but which he transcends in his capacity as insular self-consciousness, will, decision maker, or confronter of the divine.

Existentialism is the final expression of the inverted romantic spirit that began with Schopenhauer. Rousseau's attack on civilization is broadened and shifted: it is not just civilization that debases man; the entire situation in which *Dasein* finds itself forces upon it

a sense of aloneness, alienation, and despair. But this is not pessimism; conceptual theory is irrelevant. The person's response must be "existential," taking the form of a blind affirmation of will or a surrender to a confrontation (whether with Christ or communism). Such a response is beyond optimism as well. According to the existentialist, no theory of the goodness of the world is relevant, but only unreasoning hope. Although the works of Martin Heidegger and Jean-Paul Sartre are replete with themes that evoke reactions of pessimism and optimism, they significantly avoid raising the old issues concerning the relative predominance of good and evil in the world. Gabriel Marcel has eloquently made the distinction between optimism and hope in *Homo Viator* (Paris, 1944, Ch. 2). The more completely irrationalistic followers of the existentialist movement (Jean Genêt, for example) push this rejection of the traditional philosophical issue further into an ultimate reversal of good and evil and a doctrine of redemption through evil.

Although optimism and pessimism are terms that are useful in expressing fundamental human attitudes toward the universe or toward certain aspects of it, they have an ambiguity and relativity that makes them useless for a valid philosophical analysis. The question of the relative amounts of good or evil in human life and its environment is too involved to be resolved with existing philosophical tools. The dominant movements in contemporary philosophy prefer to describe and analyze the human situation more carefully in order to achieve greater understanding of the elements involved in it. That this must be done in cooperation with psychology and the natural and social sciences seems obvious; yet there are distinctively philosophical issues involved (some of which are very old) that are receiving more fruitful analysis with recent philosophical techniques. Until the basic concepts involved in a philosophical anthropology have received such analysis, the terms *optimism* and *pessimism* might wisely be avoided.

Among analytic philosophies, the empirical and positivistic trend that brushes aside all metaphysical and ethical issues as unphilosophical offers little help in this undertaking, although the old issue of a pleasure-pain balance may be regarded as an important attempt to meet analytical and empirical requirements of method. On the other hand, contemporary linguistic analysis is seeking firm ground for some of the ethical and axiological terms upon which discussions of good and evil must be based. But the analytic movement has been cautious in moving toward the metaphysical decisions upon which the resolution of these complex problems depend. It may be conjectured that when the present interest in analytic and phenomenological exploration develops into a bolder metaphysical phase, the terms *optimism* and *pessimism* may survive as descriptions of dominant human attitudes, but they may be superseded as philosophical theories by more adequate and more complex conceptual formulations of the meaning of human life and history.

See also Abelard, Peter; Adler, Alfred; Analysis, Philosophical; Aristotle; Augustine, St.; Bacon, Francis; Bayle, Pierre; Bolingbroke, Henry St. John; Brahman; Buddhism; Coleridge, Samuel Taylor; Darwinism; Descartes, René; Determinism and Indeterminism; Dewey, John; Dilthey, Wilhelm; Duns Scotus, John; Empiricism; Enlightenment; Evil, The Problem of; Existentialism; Fechner, Gustav Theodor; Fichte, Johann Gottlieb; Franklin, Benjamin; Freud, Sigmund; Galen; Gibbon, Edward; Goethe, Johann Wolfgang von; Hartmann, Eduard von; Hedonism; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Herbert of Cherbury; Hobbes, Thomas; Homer; Idealism; James, William; Jefferson, Thomas; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Leopardi, Count Giacomo; Life, Meaning and Value of; Locke, John; Lotze, Rudolf Hermann; Machiavelli, Niccolò; Maupertuis, Pierre-Louis Moreau de; Montaigne, Michel Eyquem de; Nietzsche, Friedrich; Nihilism; Pascal, Blaise; Plato; Platonism and the Platonic Tradition; Pope, Alexander; Rationalism; Rehmke, Johannes; Renaissance; Romanticism; Rousseau, Jean-Jacques; Russell, Bertrand Arthur William; Santayana, George; Sartre, Jean-Paul; Scheler, Max; Schopenhauer, Arthur; Skepticism, History of; Socrates; Spinoza, Benedict (Baruch) de; Teleological Argument for the Existence of God; Thomas Aquinas, St.; Vives, Juan Luis; Voltaire, François-Marie Arouet de; Whitehead, Alfred North; William of Ockham.

Bibliography

HISTORY OF PESSIMISM AND OPTIMISM

Billisch, Friedrich. *Das Problem des Übels in der Philosophie des Abendlandes,* 3 vols. Vienna: A. Sexl, 1959.

Diels, Hermann. *Der antike Pessimismus.* Berlin, 1921.

Hartmann, Eduard von. *Zur Geschichte und Begründung des Pessimismus,* 2nd ed. Leipzig, 1892.

Plümacher, Olga. *Der Pessimismus in Vergangenheit und Gegenwart. Geschichtliches und Kritisches.* Heidelberg, 1884.

Siwek, Paul. "Optimism in Philosophy" and "Pessimism in Philosophy," in *New Scholasticism* 23 (1948): 239–297, 417–439.

Sully, James. *Pessimism: A History and a Criticism.* London: n.p., 1877.

Tsanoff, Radoslav A. *The Nature of Evil.* New York: Macmillan, 1931.

Vyverberg, Henry. *Historical Pessimism in the French Enlightenment.* Cambridge, MA: Harvard University Press, 1958.

NINETEENTH-CENTURY VIEWS

Bailey, Robert B. *Sociology Faces Pessimism: A Study of European Sociological Thought amidst a Fading Optimism.* The Hague: Nijhoff, 1958.

Caro, Elme Marie. *Le pessimisme au XIXe siècle: Leopardi—Schopenhauer—Hartmann.* Paris: Hachette, 1878.

Copleston, Frederick. *Arthur Schopenhauer: Philosopher of Pessimism.* London, 1946.

Dorner, August. *Pessimismus: Nietzsche und Naturalismus, mit besonderer Beziehung auf die Religion.* Leipzig: F. Eckardt, 1911.

Gass, Wilhelm. *Optimismus und Pessimismus.* Berlin: G. Reimer, 1876.

Nordau, Max. *Entartung.* Berlin: C. Duncker, 1892. Translated from the second German edition as *Degeneration.* New York: Appleton, 1895.

Petraschek, Karl. *Die Logik des Unbewussten.* Munich, 1920. Vol. II especially.

Petraschek, Karl. *Die Rechtsphilosophie des Pessimismus.* Munich, 1929.

CONTEMPORARY ISSUES

Marcel, Gabriel. *Homo Viator.* Paris: Aubier, 1945. Translated by Emma Craufurd as *Homo Viator.* Chicago, 1951.

Marcuse, Ludwig. *Pessimismus: ein Stadium der Reife.* Hamburg: Rowohlt, 1953.

Unamuno, Miguel de. *Des sentimiento trágico de la vida en los hombres y en los pueblos.* Madrid, 1913. Translated by J. E. Crawford Flitch as *The Tragic Sense of Life in Men and in Peoples.* London, 1921.

L. E. Loemker (1967)

PESTALOZZI, JOHANN HEINRICH
(1746–1827)

Johann Heinrich Pestalozzi was a Swiss educator whose views profoundly affected the history and philosophy of education. Pestalozzi's father, a clergyman in Zürich, then the most lively center of awakening German culture and literature, died when his son was six years old. Pestalozzi's profound piety, the desire to love and to be loved, his compassion for suffering—and his extreme sensitivity and awkwardness in dealing with the practical affairs of life—were due largely to the exclusive upbringing of his pious mother.

After graduating from the Collegium Humanitatis (a secondary school), he turned to agriculture and experimented at his newly acquired farm, the Neuhof, with a school for the children of the neighboring farmers that was to combine elementary education with practical work. The Neuhof enterprise was a failure, financially as well as educationally, but it brought him the insights that determined his later educational, social, and religious theory and practice. These insights are jotted down in aphoristic style in *Die Abendstunde eines Einsiedlers* (Evening hour of a hermit; 1780), one of those astounding works of sudden illumination which we sometimes find in the lives of men of rare genius.

As a young man, Pestalozzi sympathized with a liberal student movement which was considered subversive by the patrician government of Zürich. He also sympathized actively with the Swiss and French revolutions at the end of the eighteenth century but was soon disappointed in the development of both.

In 1789 he took over the education of the desolate children of the town of Stans, which had been the scene of a battle between the French and the Swiss and had been badly ransacked by the French victors. Later he founded schools at Burgdorf and Münchenbuchsee, and finally at Iverdon on the shore of the Lake of Neuchâtel, attracting increasingly the attention of reform-minded men and women all over Europe. "Pestalozzianism," as a method of education that emphasized the importance of individual differences and the stimulation of the child's self-activity as against mere rote learning, was transferred also to the United States and resulted, about 1860, in a thorough reorganization of its elementary schools.

Like John Amos Comenius (whom he mentions, without being influenced by him), Pestalozzi was able to fuse his Christ-centered piety with a romantic concept of nature. First impressed by Jean-Jacques Rousseau, whose ideas he later rejected, he used the term *nature* as synonymous with all that is genuine, authentic, and free from artificiality. He regarded it as the function of education, as of all other social activities, to find the "organic" or "elemental" principles by which the inherent talent of every person could be developed to his fullest individuality, or to his "truth." His concept of truth, therefore, does not aim at logical universality; rather, it is, to use a modern term, *existential.*

A person can be educated toward maturity only if he has been allowed to sense in his earliest infancy and under the care of his mother and his family the vital element in all human relations, altruistic love. And he can safely pass over to his next developmental stage only if he has fully mastered the experiences and tasks of the preceding stage, if the whole of his personality has been formed by the "education of the heart, the hand and the

mind," if the things he has learned have become really his own and have aroused a sense of commitment, and if, finally, he discovers the vertical line, his personal relation to God, without which all relations between man and man, man and nature, and man and knowledge remain empty and meaningless.

According to Pestalozzi, it is the curse of modern civilization that its hasty and primarily verbal education does not give man enough time for the process of *Anschauung*, a term perhaps best translated as "internalized apperception," or as dwelling on the meaning and challenge of an impression. Thus modern civilization leads a person more and more away from his deeper self into a tangle of self-perceptions, of useless, if not dangerous, knowledge, and of false ambitions, which will make him unhappy.

As in many similar cases, Pestalozzi's fame as an educator has prevented the scholarly world from recognizing the full scope and depth of his interests. Besides a few and often inadequate accounts, little attention has been paid to Pestalozzi as a man of passionate concern for social justice and for new forms of religious education which were intentionally prevented by corrupt ecclesiastical institutions.

Nor has his essay "Meine Nachforschungen über den Gang der Natur in der Entwicklung des Menschengeschlechtes" (On the path of nature in the history of mankind) received sufficient attention, although it is profounder and more realistic than the contemplations on human progress by the Marquis de Condorcet, Anne Robert Jacques Turgot, and other philosophers of the Enlightenment. According to Pestalozzi, the development of the human race is reflected in the life of every person. Each of us has in himself the primitive, the social, and the ethical human. Injustice, therefore, will remain, although we may profit from the experiences of earlier generations. But the state of moral freedom will be achieved by only a few chosen individuals, and they (in this sentence he refers to his own life) will hardly find a niche in the house of humankind.

See also Philosophy of Education, History of.

Bibliography

WORKS BY PESTALOZZI

German Editions

Sämtliche Werke, edited by A. Buchenau et al. Berlin: W. de Gruyter, 1927–. Critical edition, not yet completed.

Gesammelte Werke in zehn Bänden, edited by Emilie Bosshart et al. Zürich, 1944–1947.

Werke, edited by Paul Baumgartner. Zürich: Rotapfel, 1944–1949.

Translations

Leonard and Gertrude. Translated by Eva Channing. Boston, 1885.

How Gertrude Teaches Her Children, edited by Ebenezer Cooke, and translated by L. E. Holland and Francis Turner. London, 1894.

Pestalozzi's Main Writings, edited by J. A. Green. New York, 1912.

WORKS ON PESTALOZZI AND SELECTIONS

Anderson, L. F., ed. *Pestalozzi*. New York: McGraw-Hill, 1931.

Gutek, Gerald Lee. *Pestalozzi and Education*. New York: Random House, 1968.

Silber, Käte. *Pestalozzi: The Man and His Work*. 3rd ed. New York: Schocken Books, 1973.

Ulich, Robert. *History of Educational Thought*, 258–270. New York: American Book, 1950.

Ulich, Robert. *Three Thousand Years of Educational Wisdom*, 480–507. Cambridge, MA, 1959.

Robert Ulich (1967)

Bibliography updated by Michael J. Farmer (2005)

PETER AUREOL

(c. 1275/1280–1322)

Peter Aureol (or Petrus Aureolus, Petrus Aureoli, Peter Oriole, etc.), the French Franciscan philosopher and theologian called "Doctor Facundus," was born near Gourdon, Lot, between 1275 and 1280 and died in 1322. He entered the Franciscan order before 1300, probably at Gourdon, and was assigned to the province of Aquitaine. In 1304, Peter was at Paris, but whether he studied under John Duns Scotus is uncertain. His first work was *Tractatus de Paupertate* (1311). In 1312 he was lector at the *studium generale* at Bologna, where he composed his only purely philosophical work, the unfinished *Tractatus de Principiis Naturae* in four books. From 1314 to 1315, as lector at Toulouse, he wrote the original and influential tract *De Conceptione B. M. V.* and the *Repercussorium* against certain opponents of the tract. From 1313 to 1316, probably also at Toulouse, he composed his extensive *Scriptum Super I Sententiarum,* dedicated to John XXII. At the Chapter General of Naples in 1316, Peter was nominated to lecture on the *Sentences* at Paris. The newly elected general of the order, Michael of Cesena, who had just finished his own *Sentences* at Paris, gave his consent as required although Peter openly opposed him. Peter lectured at Paris from 1316 to 1318; his *Reportata,* formerly called "the first redaction," is now believed to belong to this period. In a letter dated July 14, 1318, John XXII asked the chancellor of Paris to grant Peter the licentiate.

Peter is later mentioned (November 13, 1318) as among the master regents. For the next two years he taught Scripture at Paris while composing his often-published *Compendium Sensus Litteralis Totius Scripturae* (1319). At the end of 1320, Peter became provincial of Aquitaine but was nominated archbishop of Aix and consecrated by the pope himself in 1321. He died either at Avignon or at Aix.

Although Peter's doctrines have never been thoroughly studied, he has long been regarded as a highly critical thinker who often discarded as useless philosophical theories of his time—for example, he rejected contemporary opinions on the cosmic influence of the intelligences. In particular, he criticized many theories of Thomas Wylton and Hervaeus Natalis. He often attacked Duns Scotus, yet he also frequently followed and defended him.

Peter's own philosophical system is characterized by skeptical and empirical traits. In epistemology he supported a form of conceptualism—a doctrine midway between the realism of the great Scholastics and the nominalism of William of Ockham—in which the intelligible species is not merely the *medium quo* but itself the immediate object of our knowledge. Universal concepts have some psychic reality but no objective foundation; any principle of individuation is thus rendered superfluous. Knowledge of the individual, because of its high degree of clarity and truth, is to be preferred to knowledge of the universal. In keeping with the principle of economy often called Ockham's razor, the constitutive elements of beings are to be limited, so that without extremely cogent reasons we should not accept a plurality of "realities" in a thing. In other philosophical fields Peter had many theories of his own. He defended the existence of neutral propositions, neither true nor false, and this led him to think that God cannot know with certainty future contingent events. Peter emphasized that man's knowledge of God is largely dependent upon the psychological dispositions of the individual; moreover, ontologically there is no common ground of being between men and God. In cosmology Peter had his own opinions on the plurality of forms, the notion of an infinite, the subjectivity of time, and the meaning of movement. He thus bears witness to the fact that there was no dogmatic uniformity in medieval Scholasticism.

See also Duns Scotus, John; Empiricism; Epistemology, History of; Medieval Philosophy; Skepticism, History of; Universals, A Historical Survey; William of Ockham.

Bibliography

Peter's *Tractatus de Paupertate* was edited in *Firmamenta Trium Ordinum B. P. N. Francisci* (Paris, 1511), Part IV, folio 116r–129r. The *Tractatus de Principiis Naturae* is preserved only in manuscript. *De Conceptione B. M. V.* and the *Repercussorium* were edited at Quaracchi, Italy, in 1904. The *Sententiarum* was edited in Rome in 1516; the critical edition by E. M. Buytaert, in 2 vols. (to date), (St. Bonaventure, NY, 1953–1956), includes the prologue and Book 1 (Distinctions 1–8), and the difficult question of the double redaction of the *Sentences* is discussed in the introduction. The *Reportata* was edited in 2 vols. (Rome, 1596–1605). The *Compendium* was edited by P. Seeboeck (Quaracchi, 1896). Peter's other works include *Compendiosa Expositio Evangelis Joannis,* Friedrich Stegmüller, ed., in *Franziskanische Studien* 33 (1951): 207–219; *Recommendatio et Divisio S. Scripturae; Commentariorum in Isiam*; one *Quodlibet* (1320) of 16 questions (edited Rome, 1605); and unedited questions and sermons.

SECONDARY SOURCES

Barth, T. Article on Peter in *Lexikon für Theologie und Kirche.* 2nd ed., Vol. VII (1963), p. 350.

Beumer, J. "Der Augustinismus in der theologische Erkenntnislehre des P. A." *Franziskanische Studien* 36 (1954): 137–171.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages,* 476–480, 777–779. New York: Random House, 1955.

Maier, Anneliese. Article on Peter in *Enciclopedia cattolica.* Vol. II (1949), 409–411.

Maier, Anneliese. "Literarhistorische Notizen über P. A." *Gregorianum* 29 (1948): 213–229.

Pelster, Franz. "Estudios sobre la transmisión manuscrita de algunas obras de P. A." *Estudios eclesiasticos* 9 (1930): 462–479, and 10 (1931): 449–474.

Pelster, Franz. "Zur ersten Polemik gegen Aureoli." *Franciscan Studies* 15 (1955): 30–47.

Pelster, Franz. "Zur Überlieferung des Quodlibet und anderer Schriften des P. A." *Franciscan Studies* 14 (1954): 392–411.

Stegmüller, Friedrich. *Repertorium Biblicum.* Vol. IV, notes 6415–6422.

Stegmüller, Friedrich. *Repertorium Comment. in Sententiae Petri Lombardi.* 2 vols. Würzburg, 1947. Vol. I, notes 314–318, 657–663.

Teetaert, A. "Pierre Auriol." In *Dictionnaire de théologie catholique.* Vol. XII. Paris, 1935. Cols. 1810–1881.

A. Emmen, O.F.M. (1967)

PETER AUREOL [ADDENDUM]

Peter Aureol (Petrus Aureolus, Petrus Aureoli, Peter Auriol, Peter Oriole), French Franciscan philosopher and theologian called "Doctor Facundus," was born near Gourdon, Lot. He entered the Franciscan order before 1300 and was assigned to the province of Aquitaine. In 1304, Aureol was at Paris, but whether he studied under

John Duns Scotus there is uncertain. His first work was *Tractatus de Paupertate* (1311). In 1312 he was lector at the studium generale at Bologna where he composed his only purely philosophical work, the unfinished *Tractatus de Principiis Naturae*. From 1314 to 1316, as lector at Toulouse, he wrote the original and influential tract *De Conceptione B. M. V.* and the *Repercussorium* against certain opponents of the tract. Probably in his Bologna and Toulouse period, Aureol was composing his extensive *Scriptum super Primum Sententiarum*; the work was substantially completed by late 1316 and dedicated to Pope John XXII. At the Chapter General of Naples in 1316, Aureol was nominated to lecture on the *Sentences* at Paris. The newly elected general of the order, Michael of Cesena, who had just finished his own *Sentences* at Paris, gave his consent as required, even though Aureol had openly opposed him.

Aureol lectured at Paris from 1316 to 1318; several extant commentaries on books I–IV of the *Sentences* are probably related to the lectures held in this period, but the relationship between the various versions is still not entirely clear [see, though, Nielsen (2002) and Schabel (2000)]. In a letter dated July 14, 1318, John XXII asked the chancellor of Paris to grant Aureol the licentiate. Aureol is later mentioned (November 13, 1318) as among the regent masters. For the next two years, he taught Scripture at Paris while composing his often-published *Compendium Sensus Litteralis Totius Scripturae* (1319) and holding at least one Quodlibetal disputation (1320). At the end of 1320, Aureol became provincial of Aquitaine but was nominated archbishop of Aix-en-Provence and was consecrated by the pope himself in 1321. He died either at Avignon or Aix.

Aureol is a perceptive critic of the views of earlier thinkers, frequently using the thought of Thomas Aquinas, Henry of Ghent, and Duns Scotus, to name but a few, as a springboard for arriving at his own opinion on the matter at hand. Aureol's views are often innovative, and some of them provoked heated reaction from contemporaries such as Hervaeus Natalis and Thomas Wylton, as well as important later thinkers such as William of Ockham, Gregory of Rimini, and John Capreolus. Aureol's thought influenced the scholastic discussion into the seventeenth century.

Aureol holds that there is no principle of individuation since only individuals exist in extramental reality. This is the foundation of Aureol's conceptualism inasmuch as it entails that all universality is a product of mental activity. Thus, Aureol rejects both the strict realism of Plato and the more moderate realism of the thirteenth century. Nevertheless, Aureol insists that our universal concepts have direct foundations in the really existing individuals in the world. All individuals have certain essential features; these features are proper to the individual (they are in no way universal), yet essential features in individuals of the same natural kind (e.g., rationality in each human being) are so similar that they cause any intellect to form the same universal concept. Which universal concept an individual someone actually forms (e.g., genus or species) depends on how closely that person wills to focus the intellect on the object of cognition. Concepts for Aureol are the products of intellectual acts, and, in one of his most idiosyncratic views, he argues that this product is numerically identical with the object of cognition, merely in another *mode of being* which Aureol calls *apparent* or *intentional* being (the being the object has in virtue of its being perceived). Aureol argues along similar lines for sense perception, and behind these views is his belief in the fundamental activity of cognitive powers: They place the object of cognition in another mode of being.

Aureol wants to ensure that his philosophical and theological explanations do not jeopardize human free will, and this comes to light in his ideas on predestination and particularly on future contingents and divine foreknowledge. In the latter areas, Aureol holds that future-tensed propositions can be neither determinately true nor determinately false but have to be neutral with regard to truth–value because otherwise everything would be determined and there would be no free will. Moreover, since for Aureol immutability is equivalent to necessity, if God knows in a determinate fashion future events as future, this knowledge will be subject to God's immutability, and hence it, and the events it describes, would be necessary. Thus, Aureol claims that God understands the future, not as future, but *indistantly* and as abstracted from all time. Aureol's view was revived at the University of Leuven in the fifteenth century and created a European-wide debate of such gravity that in 1474 the pope condemned aspects of the view.

In his epistemology, Aureol stresses the psychological experience of perception. Thus, in his interpretation of the important later-medieval distinction between intuitive and abstractive cognition, the difference between these two ways in which cognitive faculties form representations is phenomenological: Intuitive cognition appears as clear and immediate (like sight) while abstractive cognition appears discursive and mediate (like imagination). This same emphasis on psychology is found in Aureol's ideas on the foundation of knowledge, proposi-

tions known through themselves (*propositiones per se notae*): For Aureol, these propositions are characterized by being known suddenly (i.e., imperceptibly quickly) and without the aid of a teacher.

In metaphysics, Aureol adopts Duns Scotus's view that the concept of being is univocal between God and creatures and between substance and accident, but he modifies it to avoid some of the problems he sees with Duns Scotus's ideas. For Aureol, the concept of being is a totally indeterminate concept having no explicit content of its own; any intellectual acquaintance, no matter how weak, can be the basis for the formation of the concept of being. This position in turn had consequences for Aureol's view of metaphysics as a science since he holds that the subject of metaphysics is being as such. Aureol's pronounced voluntarism is in line with the Franciscan tradition, as is his view that theology is a practical (as opposed to a speculative) science, but his description of theology as *declarative* (as opposed to deductive or scientific) is quite unusual. Aureol also has distinctive views on the categories (especially on relations), on the ontology of accidents, and on infinity.

See also Capreolus, John; Duns Scotus, John; Epistemology; Gregory of Rimini; Henry of Ghent; Hervaeus Natalis; Metaphysics; Plato; Phenomenological Psychology; Thomas Aquinas, St.; William of Ockham.

Bibliography

WORKS

Tractatus de Paupertate (1311) was printed in *Firmamenta Trium Ordinum B. P. N. Francisci*, Part IV, Paris, 1511, fols. 116r–129r. *Tractatus de Principiis Naturae* (1312) is preserved only in manuscript. *De Conceptione B. M. V*. and the *Repercussorium* (1314–1316) were printed at Quaracchi, Italy, 1904 (*Bibliotheca Franciscana Scholastica medii aevi*, 3). *Scriptum super Primum Sententiarum* (1316) was printed in Rome, 1596; E. M. Buytaert published two volumes of a superior edition comprising the Prologue and Book 1 (Distinctions 1–8), St. Bonaventure, NY: 1952–1956. *Compendium Sensus Litteralis Totius Scripturae* (1319) was edited by P. Seeboeck in Quaracchi, 1896. Versions of books II–IV of *Sentences* commentary, along with *Quodlibet* (1320) of sixteen questions, were published in two volumes in Rome, 1605. Other works include *Compendiosa Expositio Evangelis Joannis*, edited by Friedrich Stegmüller, in *Franziskanische Studien*, Vol. 33, 1951, 207–219; *Recommendatio et Divisio S. Scripturae*; and unedited questions and sermons.

SECONDARY SOURCES

Bolyard, Charles. "Knowing Naturaliter: Auriol's Propositional Foundations." *Vivarium* 38 (2000): 162–76.

Friedman, Russell L. "Peter Auriol on Intentions and Essential Predication." In *Medieval Analyses in Language and Cognition: Acts of the Symposium. The Copenhagen School of Medieval Philosophy*, edited by Sten Ebbesen and Russell L. Friedman, 415–430. Copenhagen: Royal Danish Academy of Sciences and Letters, 1999.

Friedman, Russell L., admin. "The Peter Auriol Home Page." Available from http://www.igl.ku.dk/~russ/auriol.html. (This site includes information on Auriol's life and works, as well as extensive bibliography and text editions).

Halverson, James L. *Peter Aureol on Predestination: A Challenge to Late Medieval Thought*. Leiden: E. J. Brill, 1998.

Nielsen, Lauge Olaf. "Peter Auriol's Way with Words. The Genesis of Peter Auriol's Commentaries on Peter Lombard's First and Fourth Books of the *Sentences*." In *Mediaeval Commentaries on the* Sentences *of Peter Lombard*, edited by G. R. Evans 149–219. Leiden: E. J. Brill, 2002.

Pickavé, Martin. "Metaphysics as First Science: The Case of Peter Auriol." *Documenti e studi sulla tradizione filosofica medievale* 15 (2004): 487–516.

Schabel, Chris. *Theology at Paris 1316–1345. Peter Auriol and the Problem of Divine Foreknowledge and Future Contingents*. Aldershot, U.K.: Ashgate, 2000.

Tachau, Katherine H. *Vision and Certitude in the Age of Ockham: Optics, Epistemology, and the Foundation of Semantics, 1250–1345*. Leiden: E. J. Brill, 1988.

A. Emmen and Russell L. Friedman (2005)

PETER DAMIAN

(1007–1072)

Peter Damian, one of the greatest churchmen of the eleventh century, was born in Ravenna. After studying and teaching the liberal arts in several Italian cities, he joined a community of hermits at Fonte Avellana, near Gubbio, in Umbria (c. 1035), and became prior about 1040. He was soon called from the monastic life, however, to become an active leader in the growing movement of ecclesiastical reform. He became cardinal bishop of Ostia in 1057 and was sent on papal missions to Milan (1059), France and Florence (1063), Germany (1069), and Ravenna (1072). He died at Faenza.

Damian's attitude toward the humanistic culture of his time was ambiguous. Although he was a fine Latin stylist in both prose and verse, and a master of argument, he nevertheless belittled both grammar and dialectic. He argued, for example, that the study of grammar had begun badly when the devil taught Adam and Eve to decline *deus* in the plural (Genesis 3:5, "Ye shall be as gods"). As for dialectic, it could be nothing more than the "handmaid" (*ancilla*) of theology, and its usefulness even in that office was strictly limited.

The ascetic tradition of disdain for the world (*contemptus saeculi*), stemming from early Christian opposi-

tion to the naturalism and hedonism of pagan culture, dominated Damian's life and his pastoral care of others. His hostility to literary and logical studies was rooted in the conviction that the true purpose of human existence is to be found in the contemplation of God. Because he believed that religious communities should be nurseries of contemplatives, he was especially critical of the pursuit of secular studies by monks.

The intellectual conflicts of the age confirmed Damian in his opposition to dialectic. Theologians skilled in elementary Aristotelian logic were applying their analytical methods to major Christian doctrines, with more or less destructive results. While some defenders of orthodoxy responded to this challenge by attempting to formulate a rational apologetic for Catholic dogma, others (including Damian) were convinced that the pretensions of the dialecticians must be countered by unequivocal condemnation.

Peter Damian's most radical critique of human reason appeared in his major theological work, *De Divina Omnipotentia.* Here he argues not only that Christian dogma, being based on divine revelation, is beyond the range of rational demonstration but also that the norms of human rationality need not apply to the content of dogma. Indeed, his fundamental theological principle excluded any reasonable assurance that human experience as a whole could be orderly and intelligible. For Damian, the entire created order depends simply on the omnipotent will of God, which can even alter the course of past history.

See also Aristotelianism; Asceticism; Hedonism; Logic, History of: Medieval (European) Logic; Naturalism; Reason.

Bibliography

WORKS BY PETER DAMIAN

Opera omnia. In *Patrologia Latina,* edited by J.-P. Migne. Vols. 144 and 145. Paris: Vivès, 1867.

Selected Writings on the Spiritual Life. Translated by P. McNulty. London: Farber and Farber, 1959.

Book of Gomorrah. Translated by J. P. Payer. Waterloo, ON: Wilfrid Laurier University Press, 1982.

Letters. Vols. 1–4, translated by O. J. Blum; Vol. 5, translated by O. J. Blum and I. V. Resnick. Washington, DC: Catholic University of America Press, 1989–2004.

WORKS ON PETER DAMIAN

Endres, J. A. *Petrus Damiani und die Weltliche Wissenschaft.* Munich, 1910.

Gaskin, Richard. "Peter Damian on Divine Power and the Contingency of the Past." *British Journal for the History of Philosophy* 5 (1997): 229–247.

Holopainen, Toivo J. *Dialectic and Theology in the 11th Century.* Leiden: Brill, 1996.

Jestice, Phyllis G. "Peter Damian against the Reformers." In *The Joy of Learning and the Love of God. Studies in Honor of Jean Leclercq,* edited by E. R. Elder, 67–94. Kalamazoo, MI: Cistercian, 1995.

Eugene R. Fairweather (1967)
Bibliography updated by Jonathan J. Sanford (2005)

PETER LOMBARD

(c. 1095–1160)

Peter Lombard, the theologian and bishop of Paris, was born at Lumellogno, Lombardy. He was elected bishop in 1159 and died the next year in Paris.

Born of a Longobard family (hence his "surname"), Peter probably studied at Bologna. He went to France about 1134, first to Rheims and then to Paris, where he soon became a teacher at the school of Notre Dame. By 1142 he was known as a "celebrated theologian," and in the same year Gerhoh of Reichersberg mentions his gloss on St. Paul, which had been preceded by a commentary on the Psalms (both works were soon adopted as the standard Scripture gloss). His fame rests chiefly on his *Book of Sentences* (*Libri Quatuor Sententiarum*), finished in 1157 or 1158.

THE "SENTENCES"

The fruit of Peter Lombard's patristic studies, scholastic lectures, and long familiarity with theological literature and problems was the *Book of Sentences.* After a classical prologue, it treats of the Trinity and the divine attributes, of creation and sin, of the Incarnation and the life of grace and virtues, of the sacraments and Last Things. It seems to have received certain retouching and additions at the hands of the author before it was published in final form. Since it surpassed all other *summae* of the twelfth century in clarity of thought and didactic practicality, as well as in the range of its subject matter, it soon acquired great popularity. After 1222, when Alexander of Hales used it as the basis of his own theological course, it obtained official standing at Paris and other medieval universities; all candidates in theology were required to comment on it as preparation for the doctorate.

The work is basically a compilation, with numerous citations of the "sentences" of the Fathers and generous and often literal borrowings from near contemporaries: Anselm of Laon, Peter Abelard's *Theology,* the anonymous *Summa Sententiarum,* Hugh of St. Victor's *De*

Sacramentis Fidei Christianae, the *Decretum* of Gratian, and the *Glossa Ordinaria.* Not all Peter Lombard's opinions found acceptance: Lists of his positions not commonly accepted abound in medieval manuscripts. However, this did not lessen the work's influence in shaping scholastic method and thought for four or more centuries. Scholastic theology flourished within the framework of the *Sentences* but also suffered from the defects and limitations of this work. Because Peter Lombard failed to treat certain questions, such as the nature and constitution of the church, the role of Christ's resurrection in the economy of salvation, and certain other aspects of Christology, these subjects were not developed in the scholastic period.

THE SCHOLASTIC METHOD

Despite his overt criticism of dialectics, Peter Lombard was largely responsible for introducing the scholastic method into the schools. Anselm of Laon (d. 1117) and his school had begun a more systematic approach to the questions of theology as a result of the growth of dialectics in the eleventh century. This approach was perfected by Peter Abelard, whose *Theologia Scholarium* is a reasoned study of theological doctrine, and whose *Sic et Non* is a vast assemblage of scriptural, patristic, and canonical material used in arguing for and against specific questions. In the prologue of the latter work, Abelard proposed principles for the reconciliation of opposing texts by semantic analysis, the authentication of texts, possible changes of opinion on the part of an author, and so on. Although critical of Abelard on many doctrinal positions, Peter Lombard was thoroughly influenced by his method of contrasting authorities and arguments, interpreting their meaning, analyzing words, and drawing conclusions. As this method passed to the great Scholastics of the thirteenth through fifteenth centuries, it eventually led to the neglect of Scripture as the core of theological studies. Roger Bacon was to complain in 1267 that a "fourth sin" of contemporary theologians was their use of a *Summa magistralis,* the *Sentences,* in place of the Bible as the text of the faculty of theology.

DOCTRINES

To dismiss Peter Lombard, as some authors have done, as primarily an unoriginal compiler almost completely lacking any philosophical foundations, and of historical importance only through the popularity his work attained, is not exactly a just judgment. Certainly Peter did not possess the deep speculative mind of, for example, his contemporary Gilbert of Poitiers or the dialectical keenness of Abelard. He made no pretense of being a philosopher, whatever he may have known of philosophical tradition. Rather, his work seems consciously to exclude the speculations of philosophy and to be primarily, if not exclusively, a work of theology based on Scripture and the doctrines of the Church Fathers. Peter Lombard was undoubtedly a compiler, yet a compiler who was master of his sources and of his own thought. Often enough, his doctrinal importance emerges only when his teachings are examined against the background of his times.

On the nature of God, for example, Peter Lombard is much more precise than the anonymous *Summa Sententiarum.* While the latter is inclined to speak of the divine essence or substance, the *Sentences,* following Augustine, makes it clear that, properly speaking, "substance" should not be predicated of the divine nature because it carries the connotation of accidents; rather, "essence," in the sense of absolute and total "beingness," or subsistent "being" (*esse*), is the proper name of God. From this Peter Lombard deduces the corollary that immutability is primary among the divine attributes. From God's immutability follows his simplicity, in marked contrast to the multiplicity which in one form or another characterizes all created beings. If other attributes are predicated of God—that he is strong or wise or just—these imply no division, composition, or distinction which would militate against his absolute self-identity. Hence, while God knows all things in one perfect, unchanging act of knowledge, things do not thereby exist in God in such a way that they share his essence. Here, however, Peter Lombard provides but the barest minimum on a question that was to receive much attention in the late thirteenth and early fourteenth centuries, the being of intelligibles.

When the creation of the world is considered in the first pages of Book II, Peter seems to react against the loquacity and daring speculation of some contemporary theologians in explaining Genesis; to all appearances, he deliberately avoids the teachings of the School of Chartres and follows Augustine's exegesis of the hexaemeron (through the *Glossa Ordinaria*), the *Summa Sententiarum,* and Hugh of St. Victor. His thought hesitates between the literal interpretation of the six days and the possibility of a simultaneous creation; although inclined to hold to the letter of the Scripture, Peter Lombard leaves the way open to the position that creation was a single act and that matter later developed according to the capacities implanted in it. Far less attention is given to the nature of man and the soul than to the purpose of man's creation and his dignity as the image of God. With a cer-

tain vehemence Peter insists on creation rather than emanation or traducianism to explain the origin of the soul. The powers of soul on the levels of sense, reason, and free will are considered almost exclusively in their relation to divine grace.

The same disregard for philosophical questions characterizes Peter's moral doctrine, which is based far less on simply rational standards of human nature or of law than on man's natural dignity as the image of God, the supernatural gift of grace, and the indwelling of the Spirit. Unlike Abelard, whose moral doctrine is man-centered in the Aristotelian tradition, Peter Lombard proposes an ethic based on God, with likeness to God as the goal of ethics and human life. If, as a theologian, he emphasizes man's absolute need of grace for virtuous acts, he lays equal stress on man's ability, under grace, to do good despite the weaknesses of human nature. The result is a moral doctrine that is far more positive than negative in character, an ethic of dignity.

See also Abelard, Peter; Alexander of Hales; Aristotelianism; Augustine, St.; Bacon, Roger; Chartres, School of; Dialectic; Gilbert of Poitiers; Medieval Philosophy; Patristic Philosophy; Saint Victor, School of.

Bibliography

WORKS BY PETER LOMBARD

"Gloss on the Psalms" (c. 1135–1137) may be found in *Patrologia Latina,* edited by J. P. Migne (Paris, 1844–1864), Vol. 191, pp. 55–1296; "Gloss on the Epistles of St. Paul" (1139–1141), ibid., pp. 1297–1696 and Vol. 192, pp. 9–520. Some twenty-nine sermons published under the name of Hildebert of Lavardin are contained in *Patrologia Latina,* Vol. 171, pp. 339–964. The *Libri Sententiarum* is available in many old editions; a critical edition was published at Quaracchi (Florence) in 1916, and a new edition was prepared by Editiones Collegii S. Bonaventure, Rome, in 1971–1981.

WORKS ON PETER LOMBARD

Among the important articles in *Miscellanea Lombardiana* (Novara: Istituto Geografico De Agostini, 1957) are L. Ott, "Pietro Lombardo: Personalità e opera," pp. 11–23; S. Vanni Rovighi, "Pier Lombardo e la filosofia medioevale," pp. 25–32; R. Busa, "La filosofia di Pier Lombardo," pp. 33–44; Stanley J. Curtis, "Peter Lombard, a Pioneer in Educational Method," pp. 265–273; and A. Gambaro, "Piero Lombardo e la civiltà del suo secolo," pp. 391–402.

Many articles of interest also appeared in the now defunct review *Pier Lombardo* between 1957 and 1962. Among them are E. Bertola, "La dottrina della creazione nel *Liber Sententiarum* di Piero Lombardo" 1 (1) (1957): 27–44; E. Bertola, "La dottrina lombardiana dell'anima nella storia delle dottrine psicologiche del XII secolo" 3 (1) (1959): 3–18. G. De Lorenzi, "La filosofia di Pier Lombardo nei Quattro Libri delle Sentenze" 4 (1960): 19–34; C. Fabro, "Attualità di Pietro Lombardo," ibid., 61–73; and I. Brady, "A New Edition of the Book of Sentences" 5 (3 and 4) (1961): 1–8. The Brady article is a sort of prospectus of the forthcoming Quaracchi edition of the *Sentences.* All of *Miscellanea* and *Pier Lombardo* are of interest.

See also P. Delhaye, *Pierre Lombard, sa vie, ses oeuvres, sa morale* (Montreal and Paris, 1961).

Ignatius Brady, O.F.M. (1967)

PETER LOMBARD [ADDENDUM]

Although Father Ignatius C. Brady's entry from the first edition remains authoritative, significant progress has since been made in research on Peter Lombard. Most importantly, the new edition of the *Book of Sentences* to which Brady referred has become available in two volumes (Brady 1971–1981). Each of the two volumes contains an introduction, with detailed treatment of Lombard's life and works.

Brady's original entry requires two factual corrections. The first concerns the *Summa Sententiarum*, an important source of the *Book of Sentences*. This *Summa Sententiarum* has been identified as the work of Otto of Lucca—an identification that Brady himself accepted in a later publication (see Gastaldelli 1980, Brady 1986). Secondly, due to Lombard's indebtedness to the *Summa Sententiarum*, it now appears likely that he studied at Lucca, rather than at Bologna.

Brady spoke of the need to study Lombard against the background of his times, so that it might become possible to understand the superiority of the *Book of Sentences* by comparison with similar twelfth-century works. This task is addressed by Colish (1994).

The *Book of Sentences* was one of the most influential texts in medieval philosophy and theology. For recent research on the tradition of commentaries on the *Sentences*, see Evans (2002). Finally, for a concise introduction to the *Sentences*, see Rosemann (2004).

See also Medieval Philosophy.

Bibliography

Brady, Ignatius C., ed. *Magistri Petri Lombardi Sententiae in IV libris distinctae*. 2 vols. Grottaferrata, Italy: Editiones Collegii S. Bonaventurae Ad Claras Aquas, 1971–1981.

Brady, Ignatius. "Pierre Lombard." In *Dictionnaire de spiritualité*, vol. 12, *1604–1612*, edited by Marcel Viller. Paris: Beauchesne, 1986.

Colish, Marcia L. *Peter Lombard*. 2 vols. Leiden, Netherlands: E. J. Brill, 1994

Evans, G. R., ed. *Mediaeval Commentaries on the Sentences of Peter Lombard: Current Research*. vol. 1. Leiden, Netherlands: E. J. Brill, 2002.

Gastaldelli, Ferruccio. "La Summa Sententiarum di Ottone da Lucca. Conclusione di un dibattito secolare." *Salesianum* 42 (1980): 537–546.

Rosemann, Philipp W. *Peter Lombard*. New York: Oxford University Press, 2004.

Philipp W. Rosemann (2005)

PETER OF SPAIN

(13th century)

Many medieval authors are referred to by the name of Peter of Spain. One Peter of Spain is the author of a standard textbook on logic, *Tractatus* (Tracts), a work that became widely known as *Summule logicales* (Sum of logic) by *magistri Petri Hispani* and that would enjoy great renown in Europe for centuries to come. This work is typical of the manuals that gradually started to emerge within the context of twelfth- and thirteenth-century teaching practices.

With regard to the identity of this Peter of Spain, matters are rather complicated. Already in the Middle Ages there existed two traditions. One ascribed the *Tractatus* to a member of the Dominican Order (Black Friars), the other to the Portuguese secular priest who in 1276 became pope under the name of John XXI. The latter identification was favored until the latest research, done in the late 1990s, showed that most likely the author of the *Tractatus* was not John XXI but a Spanish Dominican, whose identity still remains unknown.

The *Tractatus* are believed to have been written between 1230 and 1245. Another work that has been attributed to the same author is on syncategorematic words (*Syncategoreumata*), probably written some time between 1235 and 1245. Besides these two introductory tracts on logic, there are other works written by a Peter of Spain, namely, a famous medical work titled *Thesaurus pauperum* and fourteen other works on medicine. A Peter of Spain also wrote *Scientia libri de anima* and commentaries on Aristotle's *De anima*, *De morte et vita*, and *De sensu et sensato*, as well as commentaries on works by Dionysius the Pseudo-Areopagite. In the manuscripts all these works are ascribed to Pope John XXI. In the late twelfth century another Peter of Spain (in modern times called Petrus Hispanus non-papa) compiled a textbook on grammar, *Summa "Absoluta cuiuslibet"*.

The author of the tracts on logic appears to be particularly interested in matters of ontology, in dealing with which he takes a realistic approach. Every common noun signifies a universal nature and can stand for anything sharing that nature. In sentences of the form A is B, in which A and B are common nouns, the copula *is* signifies some composition that includes the extremes (subject and predicate) A and B, and always expresses a qualified mode of being (*esse quodammodo*). Such a composition usually applies to a state of affairs possessing being in the absolute sense (*esse simpliciter*), as in "Man is an animal," but if the subject refers to a fictitious entity, for example, in "A chimera is a nonbeing," being should be understood as being in a qualified sense (*ens quodammodo*).

In the first example the expression *man*, in line with Peter's ontological stand, stands for the universal nature of manhood. Therefore, the expression is necessarily true, even if no man exists. Logical necessity, then, is based on ontological necessity, or, in other words, the necessity of propositions is founded on the necessity of the things spoken about. Necessity is associated with different types of things, like the relationships between certain concepts (such as genera and species) signifying them. Another type of necessity is found in mathematical entities. In logical argument it is important to distinguish sharply between (timeless) necessary being and being-at-a-certain-time. So an inference like "A man is necessarily an animal; therefore Socrates (who is a man) is necessarily an animal" is not valid, because a transition is made from necessary being to a being at a certain time. For Peter, the notion of necessity ultimately refers to a necessary state of affairs in reality, something that is, and must always be, the case.

Peter's account of the use of the consequential "if," in which he explains consequence in terms of causality, shows a similar connection between language and the domain of reality. Like the majority of his contemporaries, Peter has to deal with the famous question "whether from the impossible anything follows" (*utrum ex impossibili sequatur quidlibet*). According to him, the notion of impossibility can be taken in two ways, namely, either (1) absolute impossibility, which amounts to being-nothing, or (2) an impossible state of affairs, the objective content, that is, of expressions containing incompatible concepts, like in "A man is an ass." Indeed, from the latter type of impossibility something (but not anything) can follow, for example, the true conclusion "Therefore a man is an animal." From absolute impossibilities, such as the one present in "You know that you are a stone," nothing can be correctly inferred, and so any-

thing follows. To be able to make a correct inference, the antecedent should be a "something" (*res*), not a "nothing."

See also Aristotelianism; Medieval Philosophy.

Bibliography

WORKS BY PETER OF SPAIN

Syncategoreumata. First Critical Edition with an Introduction and Indexes by L.M. de Rijk. Translated by Joke Spruyt. Leiden, Netherlands: Brill, 1992.

Tractatus: Called afterwards Summule logicales. First Critical Edition from the Manuscripts with an Introduction by L.M. de Rijk, Assen: Van Gorcum, 1972.

WORKS ABOUT PETER OF SPAIN

d'Ors, Angel. "Petrus Hispanus O.P., Auctor Summularum." *Vivarium* 35 (1) (1997): 21–71.

d'Ors, Angel. "Petrus Hispanus O.P., Auctor Summularum (II): Further Documents and Problems." *Vivarium* 39 (2) (2001): 209–254.

d'Ors, Angel. "Petrus Hispanus O.P., Auctor Summularum (III) 'Petrus Alfonsi' or 'Petrus Ferrandi'?" *Vivarium* 41 (2) (2003): 249–303.

de Rijk, L. M. "On the Genuine Text of Peter of Spain's *Summule logicales* I. General Problems concerning Possible Interpolations in the Manuscripts." *Vivarium* 6 (1) (1968): 1–34.

Spruyt, Joke, trans. *Peter of Spain on Composition and Negation*. Nijmegen, Netherlands: Ingenium, 1989.

Spruyt, Joke. "Thirteenth-Century Discussions on Modal Terms." *Vivarium* 32 (2) (1994): 196–226.

Spruyt, Joke. "Thirteenth-Century Positions on the Rule 'Ex impossibili sequitur quidlibet.'" In *Argumentationstheorie*, edited by Klaus Jacobi, 161–193. Leiden, Netherlands: Brill, 1993.

Joke Spruyt (2005)

PETRARCH

(1304–1374)

Petrarch, or Francesco Petrarca, the Italian humanist, poet, and scholar, was born in Arezzo into an exiled Florentine family. He was taken to Avignon in 1312, and there he spent most of his life until 1353, except for a period as a student of law at Montpellier and Bologna and several long journeys to Italy. After 1353 he lived in Italy, mainly in Milan, Venice, and Padua; he died in Arquà near Padua. Petrarch held several ecclesiastical benefices and also enjoyed the patronage of the Colonna and the Visconti.

Petrarch's fame rests first on his Italian poems and second on his work as a scholar and Latin writer. His Latin writings include poems, orations, invectives, historical works, a large body of letters, and a few moral treatises. Among the treatises we may mention especially *De Remediis Utriusque Fortunae* (On the remedies of good and bad fortune; 1366), *De Secreto Conflictu Curarum Mearum,* better known as *Secretum* (On the secret conflict of my worries; completed before 1358), *De Vita Solitaria* (On the solitary life; 1356), and *De Sui Ipsius et Multorum Ignorantia* (On his own and many other people's ignorance; 1367).

Petrarch was no philosopher in the technical sense, and even his treatises on moral subjects are loosely written and lack a firm structure or method. Much of his thought consists of tendencies and aspirations rather than of developed ideas or doctrines, and it is inextricably linked with his learning, reading, tastes, and feelings. Nevertheless, it would be wrong to underestimate Petrarch's impact on the history of Western thought. He was the first great representative of Renaissance humanism, if not its founder; as a poet, scholar, and personality, he had a vast reputation during his lifetime and for several subsequent centuries. In many ways he set the pattern for the taste, outlook, and range of interests that determined the thought of Renaissance humanism down to the sixteenth century. Petrarch was regarded, by himself and by his contemporaries, not only as a poet, orator, and historian but also as a moral philosopher, and many of his attitudes were to receive from some of his successors the intellectual and philosophical substance which they seem to lack in Petrarch's own work.

One important aspect of Petrarch's thought that was to be developed by many later humanists was his hostility toward Scholasticism—that is, the university learning of the later Middle Ages. He attacked astrology as well as logic and jurisprudence and dedicated entire works to criticizing the physicians and the Aristotelian philosophers. These attacks, though sweeping and suggestive, are highly personal and subjective and rarely enter into specific issues or arguments. When Petrarch rejects the authority of Aristotle or of his Arabic commentator Averroes, he does so from personal dislike, not from objective grounds; when he criticizes such theories as the eternity of the world, the attainment of perfect happiness during the present life, or the so-called theory of the double truth (that is, of the separate validity of Aristotelian philosophy and of Christian theology), his main argument is that these doctrines are contrary to the Christian religion.

Yet the positive value that Petrarch opposed to medieval science was neither a new science nor mere religious faith but the study of classical antiquity. All his life

Petrarch was an avid reader of the ancient Latin writers; he copied, collected, and annotated their works and tried to correct their texts and appropriate their style and ideas. He felt a strong nostalgia for the political greatness of the Roman Republic and Empire, and the hope to restore this greatness was the central political idea that guided him in his dealings with the pope and the emperor, with the Roman revolutionary Cola di Rienzo, and with the various Italian governments of his time.

Of the ancient Latin writers, Cicero and Seneca were among Petrarch's favorites. His polemic against dialectic and other branches of scholastic learning and his emphasis on moral problems seem to be modeled after the more moderate skepticism which Seneca expresses in his *Moral Epistles* with reference to the subtle dialectic of the older Stoics. To Seneca, Petrarch owes his taste for moral declamation and the Stoic notions that appear in his writings—the conflict between virtue and fortune, the contrast between reason and the four basic passions, and the close link between virtue and happiness. Even greater is Petrarch's enthusiasm for Cicero, to whom he owes the form of the dialogue and much of his information on Greek philosophy. We might even say that Petrarch and other humanists owe to their imitation of Cicero and Seneca not only the elegance of their style, but also the elusive and at times superficial manner of their reasoning.

Petrarch could not fail to notice the numerous references to Greek sources in the writings of his favorite Roman authors. He made an attempt to learn Greek, and although he did not progress far enough to read the ancient Greek writers in the original, his awareness of Greek philosophy and literature did affect his outlook and orientation. He owned a Greek manuscript of Plato and read the *Timaeus* and *Phaedo,* which were available to him in Latin translations. He also gathered information on Plato in Cicero and other Roman authors and cited some Platonic doctrines. However, more important than these occasional references to specific theories is Petrarch's general conviction that Plato was the greatest of all philosophers, greater than Aristotle, who had been the chief authority of the later medieval thinkers. "Plato is praised by the greater men, whereas Aristotle is praised by the greater number." In his *Triumph of Fame,* Petrarch places Plato before Aristotle, and his lines appear to be a conscious correction of the praise Dante had given to the "master of those who know." Petrarch's Platonism was a program rather than a doctrine, but it pointed the way to later humanist translations of Plato and to the Platonist thought of the Florentine Academy.

Petrarch assigned second place to Aristotle, but he was far from holding him in contempt. He knew especially Aristotle's *Ethics,* and he repeatedly suggested that the original Aristotle may be superior to his medieval translators and commentators. Petrarch thus pointed the way to a new attitude toward Aristotle that was to take shape in the fifteenth and sixteenth centuries. Aristotle was to be studied in the original Greek text and in the company of other Greek philosophers and writers; his medieval Latin translations were to be replaced by new humanist translations, and his medieval Arabic and Latin commentators were to give way to the ancient Greek commentators and to those modern Renaissance interpreters who were able to read and understand Aristotle in his original text. Thus, Petrarch was the prophet of Renaissance Aristotelianism, as he had been of Renaissance Platonism.

Although Petrarch opposed the classical authors to the medieval tradition, he was by no means completely detached from his immediate past. Christian faith and piety occupy a central position in his thought and writings, and there is no reason to doubt his sincerity. Whenever a conflict between religion and ancient philosophy might arise, he is ready to stand by the teachings of the former. The *Secretum,* in which Petrarch subjects his most intimate feelings and actions to religious scrutiny, is a thoroughly Christian work, and his treatise *De Remediis Utriusque Fortunae* is equally Christian, even specifically medieval. His treatise *De Otio Religioso* (On the leisure of the monks) belongs to the ascetic tradition, and even Petrarch's polemic against Scholasticism in the name of a genuine and simple religion continues or resumes that strand of medieval religious thought which found expression in Peter Damian and St. Bernard. In his treatise on his ignorance, Petrarch goes so far as to oppose his own piety to the supposedly irreligious views of his scholastic opponents. This shows that it was at least possible to reject Scholasticism and remain a convinced Christian, and to reconcile classical learning with religious faith.

In accordance with this attitude, Petrarch liked to read the early Christian writers, especially the Church Fathers, along with the pagan classics but without the company of the scholastic theologians. His favorite Christian author was St. Augustine, who occupies a position of unique importance in his thought and work. Aside from numerous quotations scattered in Petrarch's writings, it is sufficient to mention two notable instances. Petrarch's *Secretum* takes the form of a dialogue between the author and St. Augustine, who thus assumes the role of a spiritual guide or of the author's conscience. And in the

famous letter in which Petrarch describes climbing Mont Ventoux, he expresses his feelings by a quotation on which his eyes chanced to fall in his copy of Augustine's *Confessions:* "And men go to admire the high mountains, the vast floods of the sea, the huge streams of the rivers, the circumference of the ocean, and the revolutions of the stars—and desert themselves" (*Confessions x,* 8, 15).

Besides these and a few other general attitudes, there is at least one theoretical problem on which Petrarch formulates views akin to those of many later humanists. He keeps asserting that man and his problems should be the main object and concern of thought and philosophy. This is also the justification he gives for his emphasis on moral philosophy, and when he criticizes the scholastic science of his Aristotelian opponents, it is chiefly on the grounds that they raise useless questions and forget the most important problem, the human soul. This is also the gist of the words with which Petrarch describes his feelings when he had reached the top of Mont Ventoux. The words are Petrarch's, and they express his own ideas, but they are characteristically interwoven with quotations from Augustine and Seneca.

Petrarch expresses for the first time that emphasis on man which was to receive eloquent developments in the treatises of later humanists and to be given a metaphysical and cosmological foundation in the works of Marsilio Ficino and Giovanni Pico della Mirandola. This is the reason that the humanists were to adopt the name "humanities" (*studia humanitatis*) for their studies—to indicate their significance for man and his problems. Yet behind Petrarch's tendency to set moral doctrine against natural science, there are also echoes of Seneca and St. Augustine and of Cicero's statement that Socrates had brought philosophy down from heaven to Earth. When Petrarch speaks of man and his soul, he refers at the same time to the blessed life and eternal salvation, adding a distinctly Christian overtone to his moral and human preoccupation. He thus comes to link the knowledge of man and the knowledge of God in a distinctly Augustinian fashion and also to discuss an important problem of scholastic philosophy that had its root in Augustine: the question of whether the will or the intellect is superior. In discussing this scholastic problem, Petrarch follows the Augustinian tradition, as other humanists and Platonists were to do after him, in deciding the question in favor of the will.

Petrarch, the great poet, writer, and scholar, is clearly an ambiguous and transitional figure when judged by his role in the history of philosophical thought. His thought consists in aspirations rather than developed ideas, but these aspirations were developed by later thinkers and were eventually transformed into more elaborate ideas. His intellectual program may be summed up in the formula that he uses once in the treatise on his ignorance: Platonic wisdom, Christian dogma, Ciceronian eloquence. His classical culture, his Christian faith, and his attack against Scholasticism all have a personal, and in a way modern, quality. At the same time everything he says is pervaded by his classical sources and often by residual traces of medieval thought. In this respect, as in many others, Petrarch is a typical representative of his age and of the humanist movement. He did not merely anticipate later Renaissance developments because he was unusually talented or perceptive; he also had an active share in bringing them about, because of the enormous prestige he enjoyed among his contemporaries and immediate successors.

See also Aristotelianism; Aristotle; Augustine, St.; Augustinianism; Averroes; Bernard of Clairvaux, St.; Cicero, Marcus Tullius; Dante Alighieri; Dialectic; Dogma; Florentine Academy; Humanism; Medieval Philosophy; Patristic Philosophy; Peter Damian; Pico della Mirandola, Count Giovanni; Plato; Platonism and the Platonic Tradition; Renaissance; Seneca, Lucius Annaeus; Stoicism.

Bibliography

WORKS BY PETRARCH

Petrarch's Italian poems have been printed in numerous editions and translations; see also Roberto Weiss, *Un inedito Petrarchesco* (Rome, 1950). Of the *Edizione nazionale* of his collected works only six volumes have appeared, containing his poem *Africa,* a part of his letters—*Le familiari,* edited by V. Rossi and U. Bosco, 4 vols. (Florence: Sansoni, 1933–1942)—and the *Rerum Memorandarum Libri,* edited by Giuseppe Billanovich (Florence: Sansoni, 1943). See also K. Burdach, *Aus Petrarcas aeltestem deutschen Schuelerkreise* (Berlin, 1929); *Petrarcas "Buch ohne Namen" und die päpstliche Kurie,* edited by P. Piur (Halle, Germany: Niemeyer, 1925); and *Petrarcas Briefwechsel mit deutschen Zeitgenossen,* edited by P. Piur (Berlin: Weidmann, 1933).

The collection of *Prose,* edited by G. Martellotti et al. (Milan and Naples, 1955), contains the *Secretum, De Vita Solitaria,* and selections from the invectives and other treatises. *Le traité De Sui Ipsius et Multorum Ignorantia,* edited by L. M. Capelli (Paris, 1906), is the only complete modern edition of this important treatise. For many other Latin works of Petrarch the old edition of his works, *Opera* (Basel, 1581), must still be used. See also *Scritti inediti,* edited by A. Hortis (Trieste, 1874).

English translations are available for the *Secret,* translated by William H. Draper (London, 1911); *The Life of Solitude,* translated by Jacob Zeitlin (Urbana, IL, 1924); *On His Own Ignorance,* translated by H. Nachod, who added the letter on

the ascent of Mont Ventoux and excellent notes, in *The Renaissance Philosophy of Man,* edited by Ernst Cassirer, Paul Oskar Kristeller, and John H. Randall Jr. (Chicago: University of Chicago Press, 1948), pp. 36–133; the *Testament,* translated by Theodor E. Mommsen (Ithaca, NY, 1957); and for many letters—*Petrarch, the First Modern Scholar and Man of Letters,* 2nd ed., translated by James Harvey Robinson (New York, 1907), *Petrarch's Letters to Classical Authors,* translated by Mario E. Cosenza (Chicago: University of Chicago Press, 1910), and *Petrarch at Vaucluse,* translated by Ernest H. Wilkins (Chicago: University of Chicago Press, 1958).

WORKS ON PETRARCH

From the vast literature on Petrarch only a few works can be mentioned; for a bibliography, see N. Sapegno, *Il trecento* (Milan, 1948). For Petrarch's life and works see Edward H. R. Tatham, *Francesco Petrarca,* 2 vols. (London, 1925–1926); U. Bosco, *Petrarca* (Turin, 1946); Morris Bishop, *Petrarch and His World* (Bloomington: Indiana University Press, 1963); and, above all, numerous books and articles by Ernest H. Wilkins: *Studies in the Life and Works of Petrarch* (Cambridge, MA: Mediaeval Academy of America, 1955), *Petrarch's Eight Years in Milan* (Cambridge, MA: Mediaeval Academy of America, 1958), *Petrarch's Later Years* (Cambridge, MA: Mediaeval Academy of America, 1959), *Petrarch's Correspondence* (Padua, 1960), and *Life of Petrarch* (Chicago: University of Chicago, 1961).

For Petrarch as a scholar see Pierre de Nolhac, *Pétrarque et l'humanisme,* 2 vols., 2nd ed. (Paris, 1907), and "de Patrum et Medii Aevi Scriptorum Codicibus in Bibliotheca Petrarcae Olim Collectis," in *Revue des bibliothèques* 2 (1892): 241–279; numerous studies by Giuseppe Billanovich, especially *Petrarca letterato,* Vol. I, *Lo scrittoio del Petrarca* (Rome, 1947), and "Petrarch and the Textual Tradition of Livy," in *Journal of the Warburg and Courtauld Institutes* 14 (1951): 137–208. See also J. H. Whitfield, *Petrarch and the Renascence* (Oxford: Blackwell, 1943).

For Petrarch's political thought see Theodor E. Mommsen, *Medieval and Renaissance Studies,* edited by Eugene F. Rice (Ithaca, NY: Cornell University Press, 1959); Aldo S. Bernardo, *Petrarch, Scipio and the Africa* (Baltimore: Johns Hopkins Press, 1962); Jules Alan Wein, *Petrarch's Politics* (unpublished thesis, Columbia University, 1960); and Mario E. Cosenza, *Petrarch and the Revolution of Cola di Rienzo* (Chicago, 1913).

For Petrarch's religious and philosophical ideas see Armando Carlini, *Il pensiero filosofico religioso di Francesco Petrarca* (Iesi, Italy, 1904); Elena Razzoli, *Agostinismo e religiosità del Petrarca* (Milan, 1937); P. P. Gerosa, *L'umanesimo agostiniano del Petrarca* (Turin, 1927); K. Heitmann, *Fortuna und Virtus* (Cologne, 1958); William Granger Ryan, *Humanism and Religion in Petrarch* (unpublished thesis, Columbia University, 1950); and N. Iliescu, *Il canzoniere petrarchesco e Sant'Agostino* (Rome, 1962).

Paul Oskar Kristeller (1967)

PETRONIEVIĆ, BRANISLAV
(1875–1954)

Branislav Petronievć, a Yugoslav philosopher and paleontologist, was born in Sovljak, Serbia. He taught as a professor of philosophy at the University of Belgrade and was a member of the Serbian Academy of Science and Arts. In paleontology, Petronievć was the first to distinguish between the genera *Archaeopteryx* and *Archaeornis*; he also discovered new characteristics of the genera *Tritylodon* and *Moeritherium.*

Petronievć systematically treated many problems, both in pure philosophy and in scientific methodology. He considered himself a "born metaphysician" and devoted himself to constructing his own metaphysical system. But, although original, it grew out of the nineteenth-century empirical metaphysics of Hermann Lotze, Eduard von Hartmann, and Petronievć's teacher, Johannes Volkelt.

Petronievć's epistemological theory of empiriorationalism claimed that all contents of consciousness are absolutely real in the same sense as things per se. Thus there can be no absolute or immanent or transcendental illusion. Petronievć rejected phenomenalism also, specifically Immanuel Kant's. He claimed that an analysis of directly given empirical contents of consciousness shows that there are qualitatively simple evidences of experience, the "givenness of something"—the givenness of simple sensuous qualities as basic correlates of the laws of thought. Thought and being are identical, and apodictic knowledge of being itself is possible.

In his main philosophical work, *Principien der Metaphysik,* Petronievć claimed that the basic task of metaphysics is to explain the structure of the "world of multitude, diversity, and change" as the "pre-evidence" of the directly given empirical and transcendental reality. According to Petronievć, the world is a manifold of "discrete points of being" and of quality, of will, and so on. The world as a manifold is possible only because the real points of being are separated by real "acts of negation," which determine the qualities of being and without which being would be absolutely homogeneous. Petronievć regarded the principle of negation as "the absolute principle of the world," of both being and thought; only on the basis of this principle can the diversity and multiplicity of the world be deduced and explained. On similar grounds Petronievć considered the principle of sufficient reason the fundamental law of true knowledge.

Petronievć synthesized Benedict de Spinoza's monism and Gottfried Wilhelm Leibniz's monadological pluralism in his monopluralism. His original and pro-

found "hypermetaphysical" teachings on the origin and development of the qualitative and quantitative manifoldness of the world have yet to be studied and evaluated. His views on real space and real time, which he regarded as discreta rather than abstract continua, deserve special attention. He constructed a new geometry of real discrete space.

Petronievic's view was essentially idealistic, since he held that absolutely unconscious atoms are impossible and that the soul, which is immortal, is a conscious monad.

Petronievic upheld an ethical theory of transcendental optimism and free will. He devoted a number of studies to aesthetics, particularly in the work of the Yugoslav poet Petar II Petrović-Njegoš and of Lev Tolstoy.

Among his most notable contributions to the logical foundations of mathematics are his work on typical geometries, on the problem of the finitude or infinitude of space, the three-bodies problem, on differential quotients, and on mathematical induction. In psychology he developed theories about the observation of the transparent and on the depth and observation of compound colors. In the history of science his most notable works were on the methodology of Isaac Newton's discovery of the law of gravitation, on Johann Gottfried Galle's and Urbain-Jean-Joseph Leverrier's discovery of Neptune, and on Dmitri Mendeleev's discovery of the periodic system of elements.

See also Consciousness; Geometry; Hartmann, Eduard von; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Mathematics, Foundations of; Monism and Pluralism; Newton, Isaac; Phenomenalism; Spinoza, Benedict (Baruch) de; Tolstoy, Lev (Leo) Nikolaevich.

Bibliography

WORKS BY PETRONIEVIĆ

Der ontologische Beweis für das Dasein des Absoluten. Leipzig, 1897.

Der Satz vom Grunde. Belgrade: Staatsdruckerei, 1898.

Prinzipien der Erkenntnislehre. Berlin, 1900.

Prinzipien der Metaphysik, 2 vols. Heidelberg, 1904–1911.

Die typischen Geometrien und das Unendliche. Heidelberg, 1907.

L'évolution universelle. Paris, 1921.

Résumé des travaux philosophiques et scientifiques de Branislav Petronievic. Academie Royal Serbe, Bulletin de l'Academie des Lettres, No. 2. Belgrade, 1937.

WORKS ON PETRONIEVIĆ

Spomenica Branislav Petronijević. SAN No. 13. Belgrade, 1957. Articles on Petronievic by various authors.

Bogdan Šešić (1967)

PETROVIĆ-NJEGOŠ, PETAR
(1813–1851)

Petar Petrović-Njegoš, Prince Petar II of Montenegro, was born in the village of Njegusi near Cetinje. As the government of Montenegro was then a theocracy, Njegoš, who ruled from 1830 to 1851, had to act as high priest, much against his own views and wishes. He was religious by conviction, but opposed to any religious fanaticism or formalities. By setting up a number of civil and cultural institutions, he transformed Montenegro from a tribal to a modern state.

Njegoš was one of the greatest Yugoslav poets. His principal works are *Slobodijada* (Ode to liberty), *Gorski Vijenac* (The mountain wreath), *Luča Mikrokozma* (The ray of the microcosm), *Šćepan Mali* (Schepan the small), and a number of minor poems, the best of which is the reflective poem *Misao* (The thought). His main themes were man's destiny, marked by struggle and suffering, and freedom, which he understood as partly the struggle for national liberty. The elaboration of these themes led Njegoš to many philosophical thoughts and meditations. Being predominantly a poet, he presented these thoughts in poetic images and visions. The philosophical conception implicit in these images is a Platonic dualism. God and matter are coeternal. Mind and body are opposed principles both ontologically and axiologically. Mind originates in heaven, whereas body belongs to the "realm of decay." The body is "the physical shackles of the soul"; passions "lay man below the beast," whereas mind makes him "equal to immortals." In *Luča Mikrokozma* Njegoš interpreted the union of mind and body as a consequence of sin and the Fall. The first man, Adam, was once pure spirit, but he joined Satan in his rebellion against God, although he soon repented. He was then "clad in a body" and cast upon Earth, which was created by God as a place of expiation after man's sin. Thus, Njegoš's Adam, unlike John Milton's or the Adam of official church doctrine, sinned prior to his bodily creation.

Luča Mikrokozma can be seen as providing metaphysical and religious reasons for the inevitability of suffering. *Gorski Vijenac* is a mighty hymn to the national struggle for liberation and to the struggle against evil in general. To justify this struggle Njegoš elaborated a dynamic and basically dialectical conception of the world. The world is made up of opposed and dangerous forces at permanent war. Through this struggle, order emerges out of chaotic disorder, and spiritual power triumphs over great confusion. Struggle and suffering are not mere evils but have a positive, creative aspect as well. The spark appears only

after the flint is struck hard, and the soul that has endured temptations "nourishes the body with internal fire." Heroism is the master of evil, and human life has an aim only if it contributes to the realization of liberty, honor, and dignity. Njegoš's ethics were essentially derived from his people and, in turn, had a powerful influence on them in all the trying moments of their history.

See also Dialectic; Dualism in the Philosophy of Mind; Milton, John; Mind-Body Problem; Platonism and the Platonic Tradition.

Bibliography

WORKS BY PETROVIĆ-NJEGOŠ

Cjelokupna Djela, 9 vols. Belgrade, 1951–1955.

WORKS ON PETROVIĆ-NJEGOŠ

Djilas, Milovan. *Legenda o Njegošu.* Belgrade, 1952.

Latković, Vido. *Petar II Petrović Njegoš.* Belgrade, 1963.

Petronijević, Branislav. *Filozofija u "Gorskom Vijencu" i "Luči Mikrokozma."* Belgrade, 1924.

Šmaus, A. *Njegoševa "Luča Mikrokozma."* Belgrade, 1927.

Velimirović, Nikolaj. *Religija Njegoševa.* Belgrade, 1921.

Vuko Pavićević (1967)

PETZOLDT, JOSEPH
(1862–1929)

Joseph Petzoldt, a German empiriocritical philosopher, was born at Altenburg and taught mathematics and natural science at a Gymnasium in Spandau. In 1904 he became *Privatdozent* at the Technische Hochschule in Berlin-Charlottenburg, and in 1922 he was named associate professor. For a number of years he was chairman of the Gesellschaft für positivistische Philosophie.

Petzoldt was indebted to Ernst Mach's positivism, to the immanence philosophy of Wilhelm Schuppe, and above all to the empiriocriticism of Richard Avenarius. Petzoldt presented Avenarius's difficult philosophy in a popular form and developed it independently. For example, he offered a psychological explanation of the "narrowness," and therewith the unity, of consciousness; he tried to demonstrate the unlimited validity of psychophysical parallelism; and he analyzed ethical and aesthetic values and proposed a theory of the ethical and aesthetic permanence, or maximum stability, of humankind. According to this theory, all evolutionary processes end in states of permanence. Hence, human evolution is also heading toward a state of complete stability and toward the marking out of defining forms of permanence, that is, of invariably repeatable, fixed components of mental acts. The most basic feature of all the goals of our thought and creative work is permanence or durability—the realization of ever recurrent, repeatedly used ways of acting and the establishment of enduring forms amidst the profusion of particular configurations. An example of this is the tendency of thought toward stability, the striving for a stable conceptual system.

Petzoldt called his philosophy a "relativistic positivism." According to this view, both causality and substantiality are untenable and unnecessary categories, and the difference between the mental and the physical reduces to a difference in the "mode of interpretation." Petzoldt, like Avenarius, held that the concept of cause should be replaced by the mathematical concept to functional dependence, or uniqueness of coordination. According to Petzoldt, the causal relation is fully exhausted in a "law of uniqueness," which holds that for every process, the elements that exclusively determine it should be specified. Because there is thus nothing in the real world corresponding to the "animistic" concept of cause, this concept should be eliminated. The demand for a causal explanation that goes beyond the complete and simplest description of processes rests on misunderstandings; such an explanation is in principle unrealizable and is therefore meaningless.

The concept of substance, according to Petzoldt, originates from a need for stability in thinking. There are no absolute substances but only relatively constant complexes of sensory qualities. Since all properties hold good only relative to a subject, the idea of an absolute, nonrelative being should be discarded, and with it the category of substance. There is no "world-in-itself"; there is only a "world-for-us," whose elements are sensations, even though "things" are to be thought of as "continuing to exist" even when we do not perceive them. The world-for-us is apprehended as being mental insofar as it is perceived and as being physical insofar as it is known as a correlation of elements. That which is ultimately "given" is thus neither mental nor physical, neither immaterial nor material, neither "internal" nor "external," neither thing-in-itself nor phenomenon. These antitheses are merely relatively valid limiting concepts, intelligible only in their interrelation: they are formed only subsequent to, and on the basis of, the primordial unitary experience. Petzoldt's conception resembled Bertrand Russell's neutral monism.

Petzoldt's philosophy culminated in an evolutionary naturalism. "Man is not a permanence type, but an organism in a state of very active development; yet, like all

other organisms and like self-developing systems generally, he is headed toward a form of permanence" (*Einführung in die Philosophie der reinen Erfahrung* [Introduction to the philosophy of pure experience], Vol. 2, p. 3). Just as organic evolution tends toward the production of permanence states and "man's brain approaches more and more a form of permanence," the spiritual and intellectual evolution of man likewise tends to permanence states. We strive for the completion of science, for the perfection of social institutions and customs by a progressive adjustment of national and social differences, and for the fulfillment of art through "emphasis on the typical and essential in the phenomena."

The goal of ethics is that in all that we do and think we help to realize the future permanence state that flows from the nature of man and his environment (p. 206). This is the state of maximum utilization of powers, and hence of maximum stability, toward which all evolution strives. Each of us must risk everything "in order to perfect his personality in accordance with the nature and extent of his abilities and to place himself entirely at the service of human society" (p. 212).

See also Avenarius, Richard; Ethics, History of; Evolutionary Theory; Mach, Ernst; Positivism; Schuppe, Ernst Julius Wilhelm.

Bibliography

Works by Petzoldt include *Maxima, Minima und Oekonomie* (Altenburg, 1891); *Einführung in die Philosophic der reinen Erfahrung,* 2 vols. (Leipzig, 1900–1904); *Das Weltproblem vom Standpunkte des relativistischen Positivismus aus historisch-kritisch dargestellt* (Leipzig, 1906); *Die Stellung der Relativitütstheorie in der geistigen Entwicklung der Menschheit* (Dresden, 1921); and *Das natürliche Höhenziel der menschheitlichen Entwicklung* (Leipzig, 1927).

An article on Petzoldt is by Christian Herrmann, "Nachruf: Joseph Petzoldt," in *Kant-Studien* 34 (1929): 508–510.

Franz Austeda (1967)
Translated by Albert E. Blumberg

PFÄNDER, ALEXANDER
(1871–1941)

Alexander Pfänder, a German philosopher and phenomenologist, was born in Iserlohn. In 1891 he began his studies at the University of Munich, where he came under the influence of Theodor Lipps. With the publication of the *Phänomenologie des Wollens: Eine psychologische Analyse* (Phenomenology of willing: a psychological analysis; 1900) he joined the philosophical faculty in Munich, where he remained for the rest of his life. In 1904 he came into contact with Edmund Husserl. Though the two of them had much in common in their phenomenological orientation and accordingly had great respect for each other, Pfänder was the leader of the phenomenological circle in Munich, which was distinct from the one that Husserl led in Göttingen and later in Freiburg. Under Pfänder's influence, the Munich phenomenologists were especially wary of the transcendental turn and its concomitant idealism that Husserl put forward in his *Ideas Pertaining to a Pure Phenomenology and Phenomenological Philosophy* (1913). In Pfänder's later years, in which he suffered from protracted ill health, he worked toward the development of an understanding psychology and elaborated on his concept of phenomenology and phenomenological philosophy in his lectures. His most outstanding contributions, however, are to be found in his specialized treatment of volitional, emotional, and intellectual phenomena.

Pfänder had embarked on phenomenological investigations already in the late nineteenth century, before the publication of Husserl's *Logical Investigations* (1900–1901). The fruit of these investigations, namely *Phenomenology of Willing*, is thus a noteworthy achievement as a phenomenological work that came about independently of Husserl. Here, Pfänder is concerned with volitional phenomena in particular, but the work encompasses important considerations of method. The method that Pfänder employs is explicitly a descriptive one and thereby excludes any attempt to explain the phenomena under consideration in terms of cause and effect. At the same time this descriptive method avoids the sort of metaphysical speculation about willing such as what had been put forward by Arthur Schopenhauer in the nineteenth century. It is also important to note that Pfänder's phenomenology is not an introspective endeavor of the sort in which Lipps was engaged. His insistence that introspection is in fact retrospection is rather reminiscent of Franz Brentano, as is Pfänder's description of phenomena by means of an analysis into elements. His emphasis on the experienced ego throughout his analyses, however, is no doubt an aspect of his phenomenology that he drew from Lipps.

According to Pfänder volition always involves not only a presentation of the willed object but also an attention relief in which the object is made prominent against the background of others. Moreover, "willing" is used in a broad sense to designate striving, but also in a narrower sense that is closer to the one of ordinary language. While

he does not dismiss the possibility of pleasure as the goal of willing, he does not find this to be the case in all instances of willing. Moreover, his analyses of willing are guided by the observation that one can will only that which one believes to be possible. Accordingly, the experience of the volitional sphere involves considerations of other aspects of consciousness.

In 1913 and 1916 Pfänder turned his attention to the emotional rather than the volitional sphere of mental life, albeit with the conviction that the two are closely related. The articles that he published in these years for Husserl's phenomenological *Yearbook* are particularly concerned with sentiments (*Gesinnungen*) insofar as they are directed toward persons, places, animals, and so on, either positively or negatively, as when one speaks of someone being "well disposed" or "ill disposed" toward this or that. When there occurs a stirring of sentiment, this is an actual as opposed to a virtual or habitual sentiment. In each case the sentiment is something between a subject and an object and involves a centrifugal direction and streaming from the subject to the object. Moreover, the sentiment is either friendly or hostile toward the object in question. Sentiments can also be divided into genuine and spurious ones. The latter are exemplified by how one is disposed toward the characters in a theatrical performance. The rich array of analyses Pfänder employs in his investigations of sentiments was meant to be a contribution to the foundation of ethics and pedagogy.

Pfänder's *Logik* (Logic; 1921) should not be read as a logic textbook and certainly not as a logic in the technical sense that prevails in the current understanding of this term. Still, this work is of considerable interest as a philosophy of logic. Though his analyses of volitional and emotional phenomena are by and large focused on the acts of willing and feelings, *Logic* is primarily concerned with the correlates of intellectual acts. Pfänder calls these correlates thoughts (*Gedanken*), which are comparable to the meanings (*Bedeutungen*) that Husserl identifies as the subject matter of pure logic in the *Logical Investigations*, except that Pfänder conceives of thoughts as products of thinking. Moreover, Pfänder acknowledges not only special correlates of thinking but also a host of other objective correlates that are produced in a social context. In this sense *Logic*, like the works of other Munich phenomenologists (especially Adolf Reinach), opens up a new domain in the objective sphere for phenomenological investigation. Logic, Pfänder maintains, is particularly concerned with a class of thoughts known as judgments (*Urteile*). These are peculiar insofar as they involve a claim to truth and refer to states of affairs (*Sachverhalte*), which are made focal in Pfänder's reflections on rules of inference as well as on the laws of identity, noncontradiction, the excluded middle, and sufficient reason.

See also Husserl, Edmund; Lipps, Theodor; Phenomenology.

Bibliography

Spiegelberg, Herbert, in collaboration with Karl Schuhmann. *The Phenomenological Movement: A Historical Introduction*, 3rd ed. The Hague: Nijhoff, 1982, pp. 170–191.

WORKS BY PFÄNDER

Phänomenologie des Wollens: Eine psychologische Analyse. Leipzig: Barth, 1900.

"Zur Psychologie der Gesinnungen." *Jahrbuch für Philosophie und phänomenologische Forschung* 1 (1913): 325–404; 3 (1916): 1–125.

Logik, 3rd ed. Tübingen: Max Niemeyer, 1963.

WORKS ON PFÄNDER

Spiegelberg, Herbert, and Eberhard Avé-Lallement, eds. *Pfänder-Studien*. The Hague: Nijhoff, 1982.

Robin D. Rollinger (2005)

PHANTASIA

The Greek word *phantasia* is usually translated "imagination." However, in Greek thought the word always retains a connection with the verb *phainomai*, "I appear." It can be used to refer both to the psychological capacity to receive, interpret, and even produce appearances and to those appearances themselves.

Plato has little to say about *phantasia* as such, although in *Sophist* 264a he describes it as "a blend of perception and judgement (*doxa*)." Elsewhere, in *Timaeus* 70eff., in a strange passage that locates parts of the soul in particular parts of the body, he describes the liver as functioning like a mirror that reflects images coming from the rational part of the soul, suggesting a link between imagination, dreams, and inspired prophecy.

Aristotle gives *phantasia* a specific place in his psychology, between perception and thought. In *De anima* 3.3 he offers an account of *phantasia* that includes mental images, dreams, and hallucinations. For Aristotle *phantasia* is based on sense-perception and plays a crucial role in animal movement and desire, as he explains in *De anima* 3.9 and in the *De motu animalium*.

In Hellenistic philosophy the term *phantasia* is most commonly used to refer not to the capacity to receive or interpret appearances but to those appearances them-

selves. Both the Epicureans and the Stoics use the word to refer to the impressions we receive through our senses. The Stoics developed a distinctive theory of the *katalēptikē phantasia* or "cognitive impression," an impression that was self-evidently certain and therefore, they believed, offered, the criterion of truth and a secure basis for knowledge.

In later Greek thought the concept of *phantasia* is developed in a number of different ways. Literary critics, such as Longinus in *On the Sublime* 15.1, used it of a writer's capacity to visualize what he is describing and to recreate such visualization in the audience. In the second century CE, Philostratus, rather unusually, contrasts *phantasia* with *mimesis*, distinguishing between the ability of a sculptor like Phidias to portray gods he had never seen and the technique of copying, or imitation, employed by lesser artists. The link between imagination, dreams, and inspired prophecy suggested in Plato's *Timaeus* was developed by a number of later thinkers such as Plutarch (*De Pythiae oraculis* 397c, *De defectu oraculorum* 431bff.), Synesius (*De insomniis* chs. 4, 5, and 6) and Iamblichus (*De mysteriis* 3.2.3 and 3.14).

The Neoplatonists took over Aristotle's concept of *phantasia* along with the rest of his psychology but developed it in ways of their own. Plotinus in *Ennead* 4.3.30–31 suggests that there are two "image-making powers," one that receives images from sense-perception, and one that receives images from the intellect. The idea that imagination can receive images from the intellect is used by later Neoplatonists in connection with mathematics. Proclus, for example, in his commentary on Euclid, expounds the idea that when we are doing geometry, the figures about which we are thinking are "projections" in the imagination of innate intelligible principles.

See also Aristotle; Epicureanism and the Epicurean School; Imagination; Plato; Plotinus; Proclus; Stoicism.

Bibliography

Blumenthal, H. J. *Aristotle and Neoplatonism in Late Antiquity: Interpretations of the* De anima. London: Duckworth, 1996. See especially chapter 10.

Bundy, Murray Wright. *The Theory of Imagination in Classical and Medieval Thought*. Urbana: University of Illinois Press, 1927.

Cocking, J. M. *Imagination A Study in the History of Ideas*, edited by Penelope Murray. London: Routledge, 1991.

O'Meara, Dominic J. *Pythagoras Revived: Mathematics and Philosophy in Late Antiquity*. Oxford: Clarendon Press, 1989.

Watson, Gerard. *Phantasia in Classical Thought*. Galway: Galway University Press, 1988.

Anne Sheppard (2005)

PHENOMENALISM

Most philosophers have been led by the argument from illusion, by the causal argument, or by the introspective analysis advocated in the sense-datum theory to conclude that our immediate awareness in perception is not, as direct, or commonsense, realism claims, of material objects (of distinct, external physical entities perceptible by different persons at once) but of sensa (private, transitory, probably mental existents that may also be called sensations, sense data, ideas, representations, or impressions). Once this position is adopted, a serious difficulty arises concerning the nature and status of material objects. Representative realism claims that they exist external to us and cause the sensa or representations that correspond to them. The notorious difficulty of this view is that if all our direct awareness is concerned with the alleged effects, or sensa, how do we ever find out that material objects exist as their causes or what characteristics they possess? The theory seems to make material objects unobserved, and indeed unintelligible, causes of our perception. Although representative realism, especially in modern versions, tries to deal with this difficulty, it is still widely felt to be unsatisfactory. Therefore, alternative attempts have been made to deal with the problem of the nature of material objects. One such approach, which may loosely be called phenomenalist, is to reduce material objects to sensa, that is, to explain them as consisting solely of sensa or as being primarily groups or patterns of them. This approach results in slightly varying views, and when the term *phenomenalism* is used, reference is very often intended only to what we here call linguistic phenomenalism.

To introduce these variants of phenomenalism, we may consider one central problem that faces any attempt to reduce material objects to sensa, namely, the fragmentariness of perception. Any material object is believed to exist for long periods when it is not observed—for example, the furniture in an empty room, the beams in the roof, and so on—and some objects, such as rocks in Antarctica or under the ocean, may never have been observed. Yet when they are not observed, material objects cause no sensa, have no sensa belonging to them or constituting them. Hence, if material objects are reduced to actual sensa and consist only of them, they must cease to exist when unobserved, and those never observed must never have existed. Worse still, the material objects in a room must apparently come into and go out of existence as one looks at or away from them—the blinking of a human eye can destroy or create them. This seems such an intolerable paradox that George Berkeley,

though tempted to say that material objects are simply collections of ideas, had to introduce God as their continuing basis or cause. True phenomenalism, however, can no more allow unobserved divine causes than unobserved material ones.

PROBLEM OF FRAGMENTARINESS

Several approaches to the problem of fragmentariness may be taken.

HUME. One might accept fragmentary existence, though saying it is no insuperable paradox: Objects are no more than groups or patterns of sensa, but owing to the regularity with which the same or similar series of sensa occur, we imaginatively fill in the gaps and falsely suppose that continuously enduring objects exist. This was David Hume's official view. One may say that just as a tune may bridge various pauses when no sound occurs and thus be a pattern of sounds with intervening gaps, so an object may be a group or pattern of sensa and gaps. Nevertheless, the theory is incredible and is only on the fringe of the phenomenalist group of theories. For one thing, it is difficult to see why sensa recur in groups or patterns if nothing exists in between; the existence of some continuant basis or focus of them seems a far simpler and more plausible hypothesis than what would be a series of unexplained coincidences.

SENSIBILIA. Hume himself toyed with the supposition that impressions might exist unobserved—that the gaps might be filled with unsensed sensa—and if H. H. Price is right, Hume should have developed this as his official theory. Such a development was explicitly formulated by Bertrand Russell in his *Mysticism and Logic*, where he gave the name "sensibilia" (singular, sensibile) to these "objects that have the same metaphysical and physical status as sense-data without necessarily being data to any mind." Russell regarded sensibilia as the ultimate constituents of matter; thus, objects consist of systems of sensed sensibilia (that is, sensa) and unsensed ones.

However, he soon abandoned this position, which seems untenable on two main grounds. First, it cannot explain the causal processes in perceiving. How does the sensing of sensibilia bring sensa into being? The evidence of the causal processes and of the conditioning of perception by the state of the nervous system and sense organs suggests that sensa are "generated," that is, brought into being, by events in the brain; this seems incompatible with the view that they existed as sensibilia before they were sensed. Second, what evidence is there of the existence or the nature of sensibilia? One cannot observe that such entities fill gaps between actual sensa; they are just as obscure and hypothetical as the unobserved material objects of representative realism and, in fact, introduce the very difficulty that they were intended to avoid.

FACTUAL PHENOMENALISM. Factual phenomenalism attempts to fill the gap between actual sensa with possible ones by defining material objects as groups of actual and possible sensa. This view was originated by J. S. Mill, who held that matter consists of "groups of permanent possibilities of sensation." Unfortunately, this theory also leaves quite obscure what possible sensa could be and adds the further implausibility that the gap-filling entities are purely possibilities and not actualities at all. If taken strictly, this should mean that nothing actually fills the gaps. To say that something, for instance, an accident, is possible implies that it is not actual, though it might be claimed that a possible *X* is an actual *Y*; for instance, the possible winner of a race is an actual horse, in which case once again matter will consist largely of unknown and unobservable entities. The view is also open to many of the objections to phenomenalism stated below.

LINGUISTIC PHENOMENALISM. Linguistic phenomenalism sees the basic problem before it in a different light, as one not of stating the constituents of matter but of elucidating the concept of a material object, of defining it in terms of sensa; and it seeks to achieve this not by formal definition but by a "definition in use," that is, by providing translations of statements about material objects into equivalent sets of statements about sensa. Thus, it is intended to show that what is meant by talking about tables, chairs, or similar objects can be expressed solely by talking about sensa; sometimes this is expressed by saying that material objects are logical constructions out of sensa. The underlying position is, in essence, that of Hume—that all we know to exist are sensa occurring in various patterns or sequences—but one main difference lies in the claim that these regular relationships between sensa are not something to be supplemented by imagination but are actually what we indirectly refer to by talking of material objects. Such objects are in fact coordinating concepts, devices that enable us to group and correlate our sense experiences, to identify and to refer to patterns in them.

The other main difference from Hume's position is in the linguistic presentation, the attempt to elucidate the concept by translation into a set of equivalent statements. This is in accordance with the linguistic approach contemporary with the heyday of phenomenalism, and it was

held that statements about material objects and statements about sensa are simply two different ways of describing the same set of facts (facts that really concern sense experiences, their patterns and sequences). However, the sets of sensum statements not only are translations but also have a special form. Insofar as the object is observed, they are all categorical, but when it is unobserved, they are hypothetical. Thus, "I see a book on the table" is equivalent to "I have sensa *XYZ*," where *XYZ* might stand for "of a rectangular, red, solid-seeming shape on a flat brown expanse." However, "There is a book on the table in the next room" is equivalent to "If you were in the next room, you would have sensa *XYZ*." This introduces the notion of possibility that was not in Hume and that factual phenomenalism expresses so implausibly. It has the great advantage of expressing the possibility of sensa in the hypothetical form of the statement without suggesting that possible sensa are somehow constituents, perhaps the sole ones, of actual objects. Also of interest is that this approach was anticipated but not developed by Berkeley (*Principles*, Sees. 3 and 58), and occurs in places in J. S. Mill.

The result is an ingenious theory that transforms the problem of producing a viable alternative to representative realism. If successful, it would be an enormous theoretical economy; it would enable the facts of experience to be accounted for solely in terms of one type of existent, sensa, without any need to go beyond them and postulate other orders of material existence behind them. Indeed, it could further claim to be neutral between the sense-datum and adverbial analyses of sensing, for one could, as Alfred Jules Ayer did, translate material-object statements into statements about "sense contents," a term used to describe how we sense but not to refer to separate entities.

This version of phenomenalism achieved great popularity from about 1930 to 1950, particularly because it was associated with (1) logical positivism and operationalism, the meaning of material-object statements being held to lie in their mode of verification, that is, in the sensum statements that verify them; (2) Russell's analysis of abstract terms, for instance, that space is not an entity but a logical construction out of observations and measurements; (3) a way of dealing with unobserved entities in physics, namely, that electron statements are equivalent to, are logical constructions out of, sets of statements about physicists' observations. However, in the last two cases the data for the construction are prima-facie observations of material objects, and the construction is thus at a different level. Furthermore, the third case gains plausibility from the fact that electrons are agreed to be unobservable; but no such unobservability belongs to tables and chairs.

DIFFICULTIES IN PHENOMENALISM

Because of its merits, linguistic phenomenalism became the dominant version of phenomenalism (so much so that the qualification "linguistic" may seem pedantic). All the same, many difficulties soon appeared in it and defied ingenious, almost desperately ingenious, attempts to deal with them. Further, the theory presupposed that our direct awareness is entirely of private sensa; consequently, it has suffered from the recent revival in direct realism. Without questioning that presupposition, we shall now consider the general difficulties in the theory.

LACK OF EQUIVALENCE. The original aim of linguistic phenomenalism was to give a fully equivalent translation of a material-object statement into sets of sensum statements, thus proving that it meant no more than is meant by a series of such statements. For various reasons this seems impossible. In the first place, according to the basic supposition of the sense-datum theory that is shared by phenomenalism, there is a different sensum for every different look, sound, feel, or other appearance of a material object. When a dish looks elliptical, one sensum belonging to it is obtained; when it looks round, another one is obtained; when it is felt, yet another; and so on. When one considers all the different points of view from which the dish can be seen and can look different, and then adds all the variations possible for the other senses and for other conditions of lighting and such, it would seem that the number of sensa belonging to the dish, and therefore the number of sensum statements necessary to produce a full analysis or translation of "There is a dish on the table," would be very great. Sometimes it is said that the number would be infinite because the different points of view are infinite in number; but this is dubious, for owing to object constancy, a slight change in point of view would not necessarily mean a different sensum.

At any rate, the list of sensum statements would have to be far longer than can be achieved in practice. Furthermore, if the analysis is really to be adequate, it must be systematic: The sets of sensum statements must be so ordered as to show something of the patterns or correlations that justify the material-object concept; but far from doing this, phenomenalists usually give up after one or two of the sensum statements have been formulated.

Equivalence has also been denied on the ground of difference in form. The original material-object statement is a categorical one, clearly stating that something

actually exists. However, the translation is a series of hypothetical statements, and even when the apodoses of these describe experiences, their normal function seems to be either to avoid asserting actual existence (or occurrence) or to convey something quite different, such as a promise or a warning—"If you touch that, you will get burned." Indeed, "If you go to the next room, you will see a book on the table" may function as a request or a suggestion that the person go there. Worse still, in the counterfactual statements that form the translation offered about past events, actual existence is denied by implication. Thus, "Pterodactyls lived in the Mesozoic era" would probably be translated "If an observer had been present in the Mesozoic era, he would have had pterodactyl-like sensa." However, there was no observer at that time—in fact, no human beings at all—and no sensa as we know them. Thus, the assertion of actual existence is replaced in the alleged translation by assertions about what might have happened but did not.

Another bar to the claim of equivalence is that there is not full mutual entailment of original and translation. On the one hand, there might be some illusion or hallucination in which the sensum statements would be true and the material-object statement false: All the red booklike sensa might be present, and yet the object might be a box covered and shaped to look and feel like a book. This can, no doubt, be ruled out in practice by getting enough sensa, especially those resulting from such tests as opening the book, but it is doubtful how far results of such tests are really part of the meaning of the material-object statement and are therefore true features of the translation. On the other hand, the material-object statement might be true and the sensory ones false. There might be a book on the table, and yet you might not get sensa of it—the light might fail, you might be taken ill suddenly or be careless and inattentive, the book might be covered by other objects, and so on. There is a large range of conditions that would have to be stated to ensure the truth of the sensum statement. This is particularly true if the object is a small one: "There is a needle in this haystack." If you looked, would you get the needlelike sensa?

IMPURITY OF ANALYSIS. A troublesome practical difficulty facing phenomenalists is that it is impossible to specify more than a few sensa without recourse to material-object language (and not always then). Since in considering a book, the formula "sensa of a rectangular, red, solid-seeming shape on a flat brown expanse" would not differentiate the book from, say, a chocolate box, the temptation is to say "a red, rectangular, booklike sensum." But then one no longer has a translation, and the analysis is impure; it is like saying that in French *cheval* means an animal of a *cheval*-like nature. Most phenomenalists succumb to this temptation and blame it on the poverty of language, which was designed for speaking about material objects; they say, not very convincingly, that they could invent a proper terminology for describing sensa accurately but that it would take too long.

Another type of impurity in phenomenalistic analyses lies in the protases of the hypotheticals, where reference is normally made to observers and landmarks, for example, "If you go to the next room, you will get sensa *XYZ*." Even if only your body is a material object, you are at least not a sensum; and similarly, the room is physical and material. Thus, such a hypothetical statement is not a pure sensum statement. Even giving directions by compass points, for example, "If you look north …," would seem to involve some dependence on material objects, such as the sun or a compass. Ayer suggested an ingenious way out of this difficulty: Instead of mentioning the observer and others, you describe the available sensa of the room or location, thus getting "Given sensa *ABC*, then sensa *XYZ* are obtainable," where *ABC* are "interior-of-roomlike sensa" and *XYZ* are "booklike sensa." (This also slightly mitigates the difficulty about standing conditions mentioned with respect to mutual entailment: If roomlike visual data are given, at least there is light enough to see large objects.) But once again, specifying the roomlike data without mentioning the room, though perhaps theoretically possible, presents great practical difficulties that no one has tried to surmount. Nevertheless, this second impurity problem has at least been reduced to the first one.

PUBLICITY AND PERSISTENCE OF OBJECTS. In view of the great difficulties facing any attempt at a fully equivalent and pure translation, the phenomenalist may modify his aims. He may say that by producing a few sentences of the translation and by using such short cuts as "booklike sensa" he can show the form a full analysis would take; he can give a schema or blueprint of it sufficient to show that a material-object statement really means no more than a set of sensum statements and to reveal the kind of relation between sensa that justifies the material-object concept. Others would argue that this is to abandon the real aim of phenomenalism: Unless one produces a fully equivalent translation, one cannot be sure that there is not some characteristic of material objects that cannot be rendered in terms of sensa. This objection is supported by drawing attention to several features of the ordinary concept of a material object that seem particularly resistant to phenomenalist analysis.

The first of these are the publicity of material objects (the fact that many people can perceive them at once) and their persistence or relative durability. Sensa are private and transitory, so how can statements about them convey the meaning of statements about objects? A phenomenalist answer would be that all we mean or are entitled to mean by saying that an object is perceived by two people at the same time is that they simultaneously sense similar sensa. This can be formulated as: Observer *A* has sensa *XYZ* at time *t*; observer *B* has sensa *X′Y′Z′* also at time *t*; and both sets of sensa are located similarly with respect to other background sensa. The analysis can be supported by saying that when *B* senses visual and tactile data describable as data of his touching the object, then *A* gets visual data describable as data of *B* touching it. As to the persistence of objects, all this amounts to is that sequences of similar data recur. In development of this point, Hume claimed that it involves constancy (recurrence of the same data each time you look) and coherence (sequences of data changing in an orderly manner); Ayer, however, put most emphasis on the recurrence of reversible series of data, as when you look round the room and then back again.

But these answers are inadequate for the following reasons.

(1) They make the analysis impure by reference to observers: The whole point in the publicity of material objects is that two observers have similar sensa, as opposed to a case of double vision, where one person has two sets of sensa; in the persistence of material objects it is that one observer has the recurrent or reversible series of sensa. (Actually, the best evidence of persistence would be that *A* sees the object during the gaps in *B*'s observation of it, for which mention of observers is clearly essential.)

(2) A more fundamental objection is that the assertion of the publicity and persistence of material objects is meant to convey more than the assertion of sets of sensa: One is maintaining, first, that a public object exists as the focus of two persons' perceptions and, second, that such an object continues to exist during the gaps between series of perceptions. ("Focus" here means either a common object of both perceptions, as in direct realism, or the common cause of the different sensa, as in representative realism.) It might be objected that this is simply putting forward an alternative to phenomenalism, but it seems fair to say that something like this realist claim is what we mean by a material object. Without the notion of focus or continuant, the agreement of different people's sensa or the recurrent sequences of one person's sensa are incredible series of coincidences. Why, for example, are such agreements so common in perception of objects but so rare in pains or dreams or imagery? Surely because there is something besides the sense experiences responsible for the agreement, namely, a common object or cause.

(3) Furthermore, the fragmentariness of our perception of an object is closely correlated with our own actions, as are Ayer's reversible series. If sensa of a table are replaced by sensa of the view outside the window, we must have moved our head and have looked out of the window; if we get sensa of the interior of the room after an hour's gap, we must have dozed off or have gone out and returned. This seems to show that the sensa are caused by continuing objects, the room and furniture; since the fragmentariness of our observation of these objects is explained by our actions, we do not have to assume that the objects are fragmentary as well—indeed, if they were, we should find them and their sensa appearing and disappearing without any action on our part, like the Cheshire Cat in *Alice in Wonderland*.

CAUSAL PROPERTIES AND PROCESSES. Any material object is thought to possess and to exercise many causal properties (its various powers to affect other objects by heat, propulsion, impact, pressure, chemical or electrical properties), and the concept of such an object may be claimed to involve them. They are so important that for many philosophers (for example, Price) they form the main stumbling block to the acceptance of phenomenalism, at least of the factual kind. Not only are these causal properties regularly exercised when the object is unobserved (fire still boils the water when the cook is not looking, beams still support the floor and roof even when quite hidden, and so on) but the properties and processes involved in the causation of perception—the events in the eyes and nerves of percipients—are also rarely if ever observed, and then only by scientists with special equipment. Thus, one may often perceive or experience the effects of unobserved causal properties; hence, actual sensa may be causally dependent on what are only possible ones—which is absurd.

Followers of linguistic phenomenalism may claim to avoid this. The observed movement of the hands of a

clock caused by unseen works inside it, for example, is not a case of actual sensa due to possible ones. What one should say, rather, is that sensa of hands moving are sensed, and if one were to get sensa of the back of a clock with the cover removed, one would get sensa of cog wheels and shafts moving; or, more generally, given sensa of the effect, then if certain other sensa occur, sensa of the cause would also occur—S_e, and if S_x, then S_c. It must be noted that such an analysis presupposes the Humean, or regularity, view of causation, in which all that a causal relation amounts to is that the "effect" has been observed regularly to follow the "cause" (*C* causes *E* means whenever *C*, then *E*)—any conviction that the effect is brought about by some force in the cause that compels it to happen is mere superstition or is to be explained psychologically as the projection of our feeling of expectancy. However, this analysis will not satisfy those who maintain other theories of causation.

But even granting the regularity view, there is a special difficulty for phenomenalism. Presumably the "ifs" in the phenomenalist analyses are equivalent to "whenever" and themselves state regularities; whenever the floor board is taken up, one sees the beams supporting the floor. Hence, if causal relationship means no more than regularity or constant conjunction, the formula "S_e , and if S_x , then S_c" amounts to "S_e , and whenever S_x , then S_c" or "S_e , and S_x causes S_c ." However, this expresses a causal relationship different from the original one; it concerns *X* and *C* rather than *C* and *E*, and, more important, expresses a relation between sensa, suggesting that one lot of sensa causes another. Indeed, this last conclusion must follow if nothing but sensory experiences exist. Thus, "The beam supports the floor" becomes "If (whenever) you have under-floor sensa, you have beam sensa," and hence, "Under-floor sensa cause beam sensa"—which is far from the original. (This point applies with greater force to the causation of perceptions; the causal properties of the percipient's nervous system must be expressed in terms of the sensa of some other person entirely—namely, the physiologist, who can observe them.)

It has been objected that all this is unfair; the causal language belongs only to material-object language, and causal relations are between material objects and events, while in the sensum language and analysis they are expressed as equivalent correlations. However, according to the regularity view of causation there is no reason why the relevant sensa, which are events and are regularly correlated, should not be causally connected. Hence, the difficulty illustrated by "under-floor sensa cause beam sensa" still stands; it suggests that causal connections are more than relations of sensa, and thus that phenomenalism is false.

Quite apart from this special difficulty, the proposed analyses of causal properties are open to the general difficulties of the phenomenalist account of the existence of objects. There is a similar impurity, particularly with respect to the causation of sense experiences, analysis of which involves reference to different observers. Equivalent translation is even more clearly ruled out: Since causal properties involve other objects as well as the object analyzed, they are more complex than such simple, sensible ones as color or shape and thus require a longer and more intricate set of sensum statements for their analysis. They also produce their effect only when a whole range of standing conditions holds, for instance, the spring will not drive the clock if the bearings are clogged with dirt. All these conditions would have to be specified for the mutual entailment of a causal material-object statement and a set of sensum statements.

See also Perception; Sensa.

Bibliography

STATEMENTS AND DEFENSES

In addition to Hume's own writings, of special interest are attempts to modernize and improve on Hume: Henry Habberley Price, *Hume's Theory of the External World* (Oxford: Clarendon Press, 1940), and Alfred Jules Ayer, *The Foundations of Empirical Knowledge* (New York: Macmillan, 1940).

On sensibilism, see Bertrand Russell, *Mysticism and Logic* (London: Allen and Unwin, 1918), and Price's *Hume's Theory of the External World*. Russell's theory developed from the phenomenalistic views he put forward in *Our Knowledge of the External World* (London: Allen and Unwin, 1914), which were criticized by C. D. Broad in "Phenomenalism," in *PAS* 15 (1914–1915): 227–251.

Factual phenomenalism is expounded by John Stuart Mill in *An Examination of Sir William Hamilton's Philosophy* (London, 1872), Ch. 11 and appendix to Ch. 12. H. H. Price, *Perception* (London: Methuen, 1932), is sympathetic, though Price finally abandons factual phenomenalism. For allied views, see Karl Pearson, *The Grammar of Science*, 3rd ed. (London: A. and C. Black, 1911), or Ernst Mach, *Contributions to the Analysis of Sensations* (Chicago, 1897; London, 1900).

Linguistic phenomenalism, including criticisms of earlier variations, may be found in A. J. Ayer, *The Foundations of Empirical Knowledge* (see above), and "Phenomenalism," in *PAS* 47 (1946–1947): 163–196, reprinted in his *Philosophical Essays* (London: Macmillan, 1952); however, his views in *The Problem of Knowledge* (London: Macmillan, 1956) involve some recantation. Also useful is D. G. C. MacNabb, "Phenomenalism," in *PAS* 41 (1940–1941): 67–90. Compare the sophisticated version by Clarence Irving Lewis in his *Analysis of Knowledge and Valuation* (La Salle, IL: Open

Court, 1947)—comments on this in Roderick M. Chisholm, "The Problem of Empiricism," in *Journal of Philosophy* 45 (19) (1948): 512–517; a reply by Lewis, ibid., 517–524—and Roderick Firth, "Radical Empiricism and Perceptual Relativity," in *Philosophical Review* 59 (1950): 164–183 and 319–331.

GENERAL CRITICISMS AND DISCUSSIONS

Introductory surveys are given by John Hospers, *Introduction to Philosophical Analysis* (New York: Prentice-Hall, 1953; London: Routledge and Paul, 1956), and Charles Harold Whiteley, *Introduction to Metaphysics* (London, 1955). There is a general discussion by A. C. Ewing, R. I. Aaron, and D. G. C. MacNabb in the symposium "The Causal Argument for Physical Objects," in *PAS*, Supp. 19 (1945): 32–100.

More advanced and definitely critical are Rodney Julian Hirst, *The Problems of Perception* (London: Allen and Unwin, 1959); Alfred Cyril Ewing, *Idealism, a Critical Survey* (London: Methuen, 1934); David Malet Armstrong, *Perception and the Physical World* (London: Routledge and Kegan Paul, 1961); J. J. C. Smart, *Philosophy and Scientific Realism* (New York: Humanities Press, 1963); Wilfrid Sellars, *Science, Perception and Reality* (New York: Humanities Press, 1963), Ch. 3; Richard Bevan Braithwaite, "Propositions about Material Objects," in *PAS* 38 (1937–1938): 269–290; R. I. Aaron, "How May Phenomenalism Be Refuted?" in *PAS* 39 (1938–1939): 167–184; George Frederick Stout, "Phenomenalism," ibid., 1–18; Isaiah Berlin, "Empirical Propositions and Hypothetical Statements," in *Mind* 59 (235) (1950): 289–312; W. F. R. Hardie, "The Paradox of Phenomenalism," in *PAS* 46 (1945–1946): 127–154; and Paul Marhenke, "Phenomenalism," in *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950).

OTHER RECOMMENDED TITLES

Adams, Robert Merrihew. "Phenomenalism and Corporeal Substance in Leibniz." *Midwest Studies in Philosophy* 8 (1983): 217–258.

Ayer, A. J. *Language, Truth and Logic*. 2nd ed. New York: Dover, 1946.

Brandom, Robert. "Pragmatism, Phenomenalism, and Truth Talk." *Midwest Studies in Philosophy* 12 (1988): 75–93.

Fumerton, Richard. *Metaphysical and Epistemological Problems of Perception*. Lincoln: University of Nebraska Press, 1985.

Graff, Delia. "Phenomenal Continua and the Sorites." *Mind* 110 (2001): 905–935.

Jolley, Nicholas. "Leibniz and Phenomenalism." *Studia Leibnitiana* 18 (1986): 38–51.

Kamooneh, Kaveh. "Hume: A Supernaturalist and a Phenomenalist." *Philosophical Inquiry* 24 (2002): 95–102.

Moser, Paul K. "Beyond Realism and Idealism." *Philosophia* 23 (1994): 271–288.

Sosa, Ernest. "Epistemology Today: A Perspective in Retrospect." *Philosophical Studies* 40 (1981): 309–332.

Tye, Michael. "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion." *Mind* 108 (1999): 705–725.

Tye, Michael. "A Theory of Phenomenal Concepts." In *Minds and Persons: Royal Institute of Philosophy Supplement: 53*, edited by Anthony O'Hear. Cambridge, U.K.: Cambridge University Press, 2003.

Van Cleve, James. "C. I. Lewis' Defense of Phenomenalism." *Philosophy and Phenomenological Research* 41 (1981): 325–332.

Williamson, Timothy. *Identity and Discrimination*. Cambridge, MA: Blackwell, 1990.

R. J. Hirst (1967)
Bibliography updated by Benjamin Fiedor (2005)

PHENOMENOLOGICAL PSYCHOLOGY

"Phenomenological psychology" departs from empirical psychology by suspending naturalistic assumptions about human consciousness and by adopting a unique method, namely the phenomenological reduction, as a means of access to consciousness. Furthermore, its aim as a science is to reveal essential features of consciousness, eidetic structures, that hold for consciousness in general. Within the reduction, the focus can either be mundane, that is, directed to the mental as a region within itself, or transcendental, that is, directed to consciousness as the unique region within which all other forms of objectivity are constituted. When phenomenological psychology proceeds as an eidetic science, any results it may obtain will hold for any possible existing consciousness, but it cannot make any assertions about which of the possibilities it identifies are instantiated factually, since it must suspend all judgments about empirical facts. Phenomenological psychology reveals that mental life is intentional and at bottom temporal, and that it constitutes itself as a complicated, yet unified web of intentional relationships. This has led it to be closely associated with Gestalt theories. The task of phenomenological psychology is to reveal the various strata of mental life including both its active and passive elements, to exhibit the essential relationships among them, and to show how the complex and abstract levels are constituted out of simpler and more basic simple elements of consciousness.

In his contribution on phenomenology composed for the *Encyclopaedia Britannica* in 1928, Edmund Husserl introduced phenomenological psychology as a propaedeutic to transcendental phenomenology in general. Through the investigation of pure subjective consciousness, its forms and genesis, along with those of its correlative intentional objects, phenomenological psychology can provide the material for transcendental phenomenology. Phenomenological psychology makes clear that the starting point for phenomenology is consciousness as it presents itself to pure reflection. However, tran-

scendental phenomenology proceeds one step further by bracketing out any necessary relationship to consciousness as a worldly phenomenon belonging to humans or any other animate beings, and by investigating the very nature of consciousness in general. Transcendental phenomenology is thus nothing other than a consequence of the universal epoché that belongs to the meaning of the transcendental question concerning the ultimate basis for cognition and its objects in general. From this perspective, the instantiation of consciousness in human and other animals is merely one example that can provide the point of departure for a change in attitude that leads to the notion of a pure transcendental consciousness in which all intentionalities, including the intention of oneself as an existing individual consciousness, are constituted.

See also Consciousness in Phenomenology; Husserl, Edmund.

Bibliography

Husserl, E. "Philosophie als strenge Wissenschaft." *Husserliana,* Vol. XXV, edited by T. Nenon and H.-R. Sepp. Dordrecht: Kluwer Academic, 1987. Translated by Q. Lauer as "Philosophy as a Rigorous Science" in *Phenomenology and the Crisis of Philosophy.* New York: Harper, 1965. Introduces his rejection of naturalistic approaches to the study of consciousness.

Husserl, E. *Phänomenologische Psychologie. Husserliana.* Vol. IX, edited by W. Biemel. The Hague, 1962. Translated by J. Scanlon as *Phenomenological Psychology.* The Hague: Nijhoff, 1977. Provides detailed analyses illustrating Husserl's general methodology and many specific results. Contains all four drafts of Husserl's article for the *Encyclopaedia Britannica,* and the subsequent Amsterdam lecture, which was based upon that article.

Gurwitsch, A. "Husserl's Conception of Phenomenological Psychology." *Review of Metaphysics* 19 (1965–1966): 689–727.

Kockelmans, J. J. *Edmund Husserl's Phenomenology.* West Lafayette, IN: Purdue University Press, 1994. An introduction to Husserl's mature thinking through a careful and extensive commentary on the *Encyclopaedia* article.

Thomas Nenon (1996)

PHENOMENOLOGY

"Phenomenology" is a term that has been used in as many widely varying senses in modern philosophy as has the term that names the subject matter of this science, "phenomena."

Johann Heinrich Lambert, a German philosopher contemporary with Immanuel Kant, first spoke of a discipline that he called "phenomenology" in his *Neues Organon* (Leipzig, 1764). He took "phenomenon" to refer to the illusory features of human experience and hence defined phenomenology as the "theory of illusion." Kant himself used "phenomenology" only twice, but he gave a new and broader sense to "phenomenon" that, in turn, resulted in a redefinition of "phenomenology." Kant distinguished objects and events as they appear in our experience from objects and events as they are in themselves, independently of the forms imposed on them by our cognitive faculties. The former he called "phenomena"; the latter, "noumena," or "things-in-themselves." All we can ever know, Kant thought, are phenomena.

The next generation of philosophers, notably G. W. F. Hegel, was at great pains to show that this was a mistake. Hegel's first major work, *Phenomenology of the Spirit* (1807), traced the development of Spirit (or Mind) through various stages, in which it apprehends itself as phenomenon, to the point of full development, where it is aware of itself as it is in itself—as noumenon. Phenomenology is the science in which we come to know mind as it is in itself through the study of the ways in which it appears to us.

In the middle of the nineteenth century, the definition of "phenomenon" was further extended until it became synonymous with "fact" or "whatever is observed to be the case." As a consequence, "phenomenology" acquired the meaning that it possesses most frequently in contemporary uses—a purely descriptive study of any given subject matter. In this sense, Sir William Hamilton, in his *Lectures on Metaphysics* (1858), spoke of phenomenology as a purely descriptive study of mind. Similar was Eduard von Hartmann's use of the word in the title of his book *Phenomenology of Moral Consciousness* (1878), which had as its task a complete description of moral consciousness. When the American philosopher C. S. Peirce used the term *phenomenology*, he had in mind not only a descriptive study of all that is observed to be real but also of whatever is before the mind—perceptions of the real, illusory perceptions, imaginations, or dreams. It was the task of phenomenology to develop a list of categories embracing whatever can be included in the widest possible meaning of "to be." Peirce introduced this sense of the term in 1902.

The changes described so far are all due to extensions of the meaning of "phenomenon," but phenomenology, the science of phenomena in these different senses, remained one field of study among others, having a rela-

tion to philosophy as a whole comparable to those of logic, ethics, and aesthetics. Frequently it was recommended as a descriptive study that was to precede any attempt to provide explanations of the phenomena. But since Edmund Husserl employed the term in the early 1900s, it has become the name of a way of doing philosophy—by using the phenomenological method. For the phenomenologists, who regard their method as the only correct way of proceeding in philosophy, phenomenology is therefore the best and perhaps the only legitimate way of philosophizing today. For other philosophers, phenomenology is one school or movement in philosophy today. At the same time, however, the older sense of the term persists. "Phenomenology" is therefore used in two distinct senses. In its wider sense it refers to any descriptive study of a given subject. In the narrower sense it is the name of a philosophical movement. This entry will deal with phenomenology in the second sense.

THE MOVEMENT AND ITS ORIGINS

"Phenomenology" became the name of a school of philosophy whose first members were found in several German universities in the years before World War I, notably at Göttingen and Munich. Between 1913 and 1930 this group published a series of volumes of phenomenological studies titled *Jahrbuch für Philosophie und phänomenologische Forschung*, whose editor in chief was Husserl, the most original and most influential thinker of the group. Most of the better-known members of the phenomenological movement—Moritz Geiger, Alexander Pfänder, Max Scheler, and Oscar Becker—were coeditors, at least for a time. Martin Heidegger was another coeditor, but he cannot be counted among the phenomenologists without serious qualifications. Other major figures in the movement were Adolf Reinach and Hedwig Conrad-Martius.

The contributions to the *Jahrbuch* ranged from Husserl's writings about the foundations of phenomenology, to essays in the philosophy of mind and Scheler's major work on ethics, to pieces on the nature of analytic judgments and the paradoxes in set theory. As the interests of the various phenomenologists differed, so did their conceptions of phenomenology. These disagreements emerged only gradually, as Husserl developed the theory of the phenomenological method further and encountered a progressively more critical reception among his fellow phenomenologists. At the outset, there was general agreement that phenomenology was to be descriptive and that it was to describe phenomena by means of direct awareness (*Anschauung*). It is best to begin to clarify these terms by showing what they could, but do not, mean.

DESCRIPTION

The terms *descriptive, phenomenon*, and *direct awareness* all suggest that phenomenology is here used in its wider sense as a purely descriptive science of observable phenomena. But this wider sense of the term does not include what for the phenomenologists is the most important feature of phenomenology—that it is a nonempirical science. From the very beginnings of the phenomenological movement, when the conception of phenomenology was otherwise still quite vague, there was general agreement that phenomenology does not describe empirically observable matters of fact. Insisting on this, the early phenomenologists took a stand in opposition to philosophical views then in vogue.

Kant had distinguished three kinds of statements: empirical statements, statements true by definition (which he called "analytic"), and a third kind that he called "synthetic a priori." After being temporarily eclipsed by the German idealism of the early nineteenth century, Kant found many vigorous adherents in the later decades of that century. But there were also many philosophers who found Kant's account of the third type of statement—the statements that are neither empirical nor analytic—profoundly unsatisfactory and who, instead of attempting to supply an alternative account, rejected the tripartite classification altogether. This was done, for instance, by the German positivists Ernst Mach and Richard Avenarius, who insisted that there are no nonempirical statements that are not analytic. Of equal, if not greater, importance were those philosophers who regarded all statements as empirical. Analytic statements seemed to them clearly to rest on "the artful manipulation of language" (Mill's phrase), and they thought it therefore implausible that the statements of logic and/or mathematics should be analytic, that they should be true, and, more important, that they should be applicable to objects of everyday experience and science merely by virtue of an arbitrary choice of definitions. Accordingly, John Stuart Mill in England and Christoph Sigwart in Germany, among others, sought to show that statements in logic and mathematics are no less empirical than statements in the sciences.

In the case of logic, the most plausible argument for such a view begins with the observation that logic deals with correct and incorrect thinking. Thinking is a mental or psychological activity and must, therefore, be studied in psychology just as any other mental or psychological

activity. It seems to follow, then, that logic is either a special field within empirical psychology or a practical discipline whose theoretical foundations are supplied by empirical psychology. In the former case, the relation of logic to psychology is comparable to that of learning theory or abnormal psychology to psychology as a whole. In the latter view, logic is related to psychology as surveying is to geometry or accounting to arithmetic.

OPPOSITION TO PSYCHOLOGISM. The phenomenologists were not the first to question the identification of logical with psychological statements—a view they called "psychologism." But while some other philosophers had approached the issue by distinguishing logic from psychology in terms of the distinction between theoretical and practical disciplines, the phenomenologists attacked the identification of logical with psychological statements on the grounds that the latter are empirical statements and the former are not. The most sustained and painstaking critique of psychologism is contained in the first volume of Husserl's *Logische Untersuchungen* (Logical Investigations; Halle, 1900–1901), and the arguments in that book served as a first rallying point for phenomenologists.

Husserl's attack on psychologism had a special edge to it because his *Philosophie der Arithmetik* (Philosophy of Arithmetic; Vol. I, Halle, 1891; the projected second volume was never published) had been a frankly psychologistic account of arithmetic. His change of heart was in part occasioned by a controversy with the German mathematician and philosopher of mathematics Gottlob Frege, in which Frege had insisted that a sharp line be drawn between psychological statements, on the one hand, and logical and/or mathematical ones, on the other.

Husserl devoted an entire book to the detailed examination and refutation of every variety of psychologistic doctrine, taking careful account of each view and trying to show its inadequacy. Underlying all his arguments, however, were a few general principles to which he appealed again and again in the course of his discussion:

(1) Psychology deals with facts; therefore its statements are empirical. It has not, until now, produced any precise scientific laws, and its generalizations are vague. The rules of logic, on the other hand, are precise. Hence, psychological generalizations can neither be identical with logical laws nor be premises from which they may be derived.

(2) Empirical statements are probable, at best, for there is always a real possibility that further evidence will show them to be false. Logical truths are necessary truths. A logical principle such as *modus ponens* ("Given that 'If *p*, then *q*' is true and that '*p*' is true, '*q*' is true") is not probable; it is necessarily valid.

(3) Closely connected with (1) and (2) is the argument that empirical generalizations rest on induction; they are derived from a number of individual cases. This is not true of logical rules.

Both (2) and (3) are supported by pointing out that where there is a conflict between a logical principle and an empirical generalization, the logical principle will always emerge victorious because necessary truth is not to be refuted by a probable statement and logical truth cannot be shown to be false by an inductive generalization.

(4) The empirical generalizations of psychology produce, at best, causal laws, and logical principles are not causal laws. Premises and conclusions of an argument are not related as cause and effect; the truth of a conclusion is not the effect of the truth of the premises. Causal relations hold between events, and events happen at definite times in definite places. But the premises of an argument do not "happen," nor does the conclusion; they are either true or false. In a valid argument the truth of the conclusion "follows" from the premises; it is not the effect of events called premises.

(5) Empirical laws imply matters of fact; logical rules do not. Since empirical laws are, presumably, derived from the observation of particulars, the existence of such particulars in some place and at some time can be inferred from the truth of the empirical law. *Modus ponens*, on the other hand, does not imply that there exists, in a particular place and at a particular time, a pair of statements of the form "If *p*, then *q*" and "*p*." Nor are any corresponding facts implied by any other logical law. This point is sometimes stated in a phrase, borrowed from Gottfried Wilhelm Leibniz, that empirical laws are true only for this actual world; logical laws are true "for all possible worlds."

The upshot of these arguments is that logical and empirical statements differ in kind. Logical statements are precise, necessarily true, and not derived inductively from particulars. They are, or give rise to, logical rules, not causal laws, and they do not imply matters of fact. Empirical statements, on the other hand, are vague, probably

(but not necessarily) true, and based on inductive generalizations. They are, or give rise to, causal laws and imply the existence of matters of fact. Quite clearly, in the refutation of psychologism, the decisive argument, for Husserl, consisted in showing that there are two kinds of statements: empirical and nonempirical. Phenomenological statements are to be nonempirical.

To deny that phenomenological statements are empirical is to deny that their truth or falsity depends on sensory observation. But if not on sensory observation, on what does their truth depend? Some philosophers might be inclined to say that phenomenological statements are analytic. Insofar as only those statements are analytic that are true by virtue of explicit definition of terms, phenomenologists deny that their statements are analytic. We shall have abundant evidence that they are right in this, for phenomenological statements are not true by virtue of stipulation of meaning. But insofar as "analytic" is used in some other sense, it is not helpful either to assert or to deny that phenomenological statements are analytic; the meaning of the term *analytic* is much debated in contemporary philosophy and has therefore become extremely obscure. It is more profitable to ask the phenomenologists about the truth conditions of their statements. Their preliminary answer to this question consists in introducing the term *phenomenon* by saying that phenomenological statements are true if they accurately describe phenomena. This answer, however, remains merely a verbal maneuver unless *phenomenon* can be shown to have a clear and definite meaning.

PHENOMENA

We have seen that *phenomenon* is a technical philosophical term that different philosophers have used in very different senses. The phenomenologists sometimes say that "phenomenon" is their name for whatever appears to us in "immediate experience." By "immediate experience" they do not mean sensory observations that have not been interpreted or classified under general concepts ("raw sense data"). Like many other contemporary philosophers, the phenomenologists are not at all sure that there are for us any sensory observations that are not interpreted or classified under general concepts. The appeal to phenomena or to immediate experience is therefore not an appeal to simple, uninterpreted data of sensory experience. Furthermore, the appeal to phenomena does not presuppose the existence of a special class of objects called "phenomena." The phenomenologists do not claim to have discovered that besides all the kinds of entities found in this world (physical objects, thoughts, numbers, feelings, poems, etc.) there is one other class, phenomena. Any object is a phenomenon if looked at or considered in a particular way. This particular way of looking at all kinds of objects is recommended in the slogan "Zu den Sachen!"

Literally translated, this slogan means "To the things!" where "things" must be taken in the widest possible sense to embrace all possible kinds of objects. Like other slogans, moreover, this one gains its force from having more than one meaning. If a German says to someone, "Zur Sache!" he is exhorting him, as we would say, "to get down to business." "Zu den Sachen!" admonishes one to get down to the proper business of the philosopher by examining and describing all kinds of objects in the particular way that reveals them as phenomena.

This explication of "phenomenon" is, so far, circular. To clarify what is meant by that term, we must therefore explain what alternative ways of doing philosophy are excluded by telling us to examine and describe phenomena. We must explain the polemical import of the slogan "Zu den Sachen!" Once this is done, we must pursue the concept of phenomenon further by attempting to clarify the nature of the examination and description that shows all kinds of objects as phenomena.

OPPOSITION TO REDUCTIONISM. The polemical import of "Zu den Sachen!" is readily made clear. In it the phenomenologists expressed their opposition to all reductionism, or, as Reinach called them, "nothing-but philosophies." Such philosophies are couched in sentences like "Logical laws are nothing but psychological laws," "Moral laws are nothing but the expressions of the mores of a given society," and "Aesthetic judgments are nothing but expressions of personal taste." To oppose all views of this sort would seem dogmatic. Some "nothing-but" statements may be false, but perhaps others are true; and one would think that each would have to be examined on its merits rather than be rejected summarily as an example of reductionism. However, the phenomenologists did not attack these "nothing-but" views on the grounds that they are false but on the grounds that the philosophers who held them, held them for the wrong kinds of reasons.

Psychologism, which is just one example of reductionism, did not assert that logical laws are nothing but psychological laws in the light of a thorough examination of the nature of logical laws that proved that they are identical with psychological ones. If someone challenged the psychologistic philosopher's views, he was not invited to examine for himself the nature of logical laws and to

discover that they did not differ from those in psychology. Instead, he was given an argument from which it followed that logical laws "must" be psychological ones. Psychologistic assertions about logical and psychological laws do not result from an examination of laws in logic and psychology but are the logical consequences of certain more general assumptions. These assumptions themselves are not examined but are taken as self-evident.

Reductionism as attacked by the phenomenologist is the outcome of accepting certain statements that have not been examined carefully. If the implications of these assumptions are shown to conflict with facts about the world, the reductionist does not, the phenomenologists say, reexamine his original assumptions. Instead, he redefines the terms used to describe the facts about the world in such a way that the contradictions between these descriptions of facts and the implications of the original assumptions disappear. The redefinitions necessitated by the conflict between assumptions and facts are expressed in the "nothing-but" statements.

Opposition to phenomenalism. An example of a specific reductionist view attacked by the phenomenologists will clarify the process. David Hume's empiricism was attacked for its phenomenalism, that is, for its view that physical objects, as well as human beings, are no more than collections of their observable properties. ("Phenomenology" must not be confused with "phenomenalism.") "Observable properties" in this context refers exclusively to sensory qualities like shape, color, sound, etc. This view of Hume's did not issue from a careful examination of the nature of physical objects. Instead, it was a product of his psychological theories about the origin and meaning of concepts and words. Hume held that all concepts are either derived directly from sensory experience or are complex collections of such concepts. He regarded it as a consequence of this view that all concepts refer either to sensory qualities like shape, color, and sound or to complex collections of these. He also thought that all nouns are the names of concepts. It follows from this that all nouns naming physical objects refer to concepts that can be completely analyzed into simple concepts referring to sensory qualities. Hence physical objects—what is named by physical object nouns—are no more than complex collections of sensory qualities. However, this view is not supported by a careful examination of physical objects themselves but follows from, and hence "must" be true in the light of, Hume's psychology and views on the meanings of words.

Opposition to psychological atomism. Another target of the antireductionist polemic was the then popular attempt by philosophical psychologists like Wilhelm Wundt to define consciousness as a set of contents—sensations, feelings, affects—on which operations—association and apperception—are performed. This view was not the product of careful examination and description of the series of phenomena that we call consciousness but was a logical consequence of more general assumptions about the world. It missed, the phenomenologists maintained, the essential characteristic of consciousness that they, following Franz Brentano, called "intentionality."

Opposition to scientism. Also objectionable was the so-called scientism of the positivists Mach and Avenarius. Scientism regarded scientific statements as premises in philosophical arguments such that the truth of statements in philosophy depends on the truth of scientific statements. This view was a direct consequence of two assumptions: that all statements are either empirical or analytic, and that all empirical statements are, at least ideally, statements in science. Given these assumptions, there is a choice between restricting philosophy to the practice of logic, in which statements are often thought to be analytic, or saying that philosophical truths are empirical. If we choose the latter alternative, philosophical statements "must" have scientific premises.

But this conclusion, phenomenologists held, was drawn without paying careful attention to actual and possible functions of philosophy, which, they held, is independent of science. In this they were not motivated by any hostility toward science; on the contrary, their aim was to establish philosophy as a "rigorous science" by means of the phenomenological method. Husserl had discussed this aim at some length in his article "Philosophie als strenge Wissenschaft" ("Philosophy as Rigorous Science," in *Logos*, Vol. I, 1910–1911, 289–341). This phenomenological and rigorously scientific philosophy was expected to provide the foundations for the existing sciences by providing clear explications of the concepts that the sciences use but do not themselves explicate. For instance, the definition of number, in which Reinach was interested, was considered a task for phenomenology. Husserl was concerned with clarifying epistemological terms such as *meaning* and *truth*. So conceived, phenomenology had to be independent of the existing sciences because it was to explicate the concepts and procedures presupposed by them. To consider philosophy a branch or subsidiary of existing science was one more example of "nothing-but" philosophy.

Presuppositionless inquiry. Here it must be asked whether philosophers must not make certain assumptions. We cannot, it would seem, show that all statements

are true by reference to the truth of other statements; some we must merely assume to be true. But phenomenologists are unconvinced by this sort of argumentation. Statements in phenomenology are not true because certain other statements are true; they are true because they describe the phenomena correctly. In order to achieve true description, the phenomenologist must resist the temptation to make assumptions and, afterward, to define his terms in such a way as to make the descriptions of facts consistent with the assumptions and what must be inferred from them. The phenomenologist does not frame theories; he merely examines and then describes phenomena as they present themselves to his unprejudiced view. Having no theoretical commitment and only one practical one—to examine all phenomena carefully and to take none of them as familiar or understood until they have been carefully explicated and described—the phenomenologist says that his science is descriptive and that it is presuppositionless.

This obviously does not mean that at any given time the phenomenologist may not be operating with certain unexamined assumptions—this can always happen. The claim of presuppositionlessness expresses the resolution to eschew all unexamined assumptions and the belief that such assumptions are unnecessary; No statement must be taken as true without examination. Phenomenology does not need any true but unexamined premises; the truth of all its premises can be tested by examining the phenomena.

This sheds some light on the second, affirmative sense of the slogan "Zu den Sachen!"—an exhortation to examine phenomena and to make them the sole touchstone of the truth of philosophical statements. But the precise import of this exhortation remains unclear until the meaning of "phenomenon" has been explicated, so this is a pressing question. It is also a question fraught with particular difficulties. Phenomena, as was stated, are those aspects of objects of every kind that are revealed by a particular way of looking at objects. The phenomenal aspects of objects are not revealed by ordinary empirical observation but only by looking at them as phenomena. The meaning of "phenomenon," on the other hand, cannot merely be stipulated in analytic statements. Hence, explications of "phenomenon" must result from using the phenomenological method and must be couched in phenomenological statements. But what these statements are cannot be made clear until it is clear what a phenomenon is, nor do we know what the phenomenological method is until we know what a phenomenon is.

"**METHODOLOGICAL CIRCLE**." The entire phenomenological enterprise is involved in a circle that can be called the "methodological circle." This methodological circle does not differ formally from the circle involved in any kind of logical investigation where the rules of inference, for instance, which the completed investigation hopes to formulate and justify must be employed during the course of the investigation itself so that its result, the logical rules, is the product of the application of the rules to themselves. The existence of this circle does not prove that logic is an impossible or unjustifiable discipline, nor does its presence in phenomenology support an analogous argument against it.

The occurrence of this circle should, however, put one on his guard against taking for completed analyses statements made by phenomenologists that are, in fact, merely gropings toward and anticipations of what phenomenology, its method, and the completed theory of method will be like in some indefinitely remote future. Phenomenology does not exist as a set of doctrines but at best as a method—and this method is to be developed by applying phenomenology to itself. Hence, even the phenomenological method is still in the process of being clarified, properly described, and elaborated; it is, at least to date, quite incomplete.

Husserl liked to refer to himself as a "perpetual beginner," an expression that meant several things to him. In one of its senses, it expressed what was just said about phenomenology: It is a method that can only be progressively developed by applying it to itself. Accordingly, most of Husserl's published works are discussions of the phenomenological method. This has sometimes been taken as a symptom of an excessive fondness for writing manifestoes, but discussions of phenomenological method are not of the nature of manifestoes prior to doing phenomenology, nor are they propaedeutics. Only while doing phenomenology can we clarify its method. To write about it was, in Husserl's case, to do phenomenology.

THE INTUITION OF ESSENCES

The preceding discussion has brought to light three properties of phenomenological statements:

(1) Phenomenological statements are nonempirical.

(2) Phenomenological statements are descriptive.

(3) Phenomenological statements describe phenomena.

These leave the task of making clear what phenomena are, a matter of disagreement among phenomenologists:

Most of the schisms within the phenomenological movement originate in disagreements about the set of conditions necessary for anything to be a phenomenon. We shall examine a variety of conditions proposed, beginning with the most simple and proceeding to more complex sets as the simpler ones turn out to be incomplete. The criterion of completeness for this set of necessary conditions is that any set of conditions required for anything to be a phenomenon must at least be consistent with the first requirement for phenomenological statements—that they be nonempirical. Hence, the set of conditions laid down for anything to be a phenomenon must clearly rule out any possibility that phenomena can be described in empirical statements.

The simplest specification of phenomenon, given by some early phenomenologists, contains only two conditions:

(1) Phenomena are essences.

(2) Phenomena are intuited.

The reason for identifying phenomena with essences is instructive. As we saw, it was claimed that there are some entities by virtue of which statements in phenomenology are said to be true or false. These entities (or phenomena) are not particular observable objects by reference to which empirical statements are confirmed or disconfirmed. Instead, the phenomenologists say, they are the necessary and invariant features of objects. Phenomenology explicates those features of any given object without which it could not truly be said to be the object that it is. These most general, necessary, and invariant features of objects have been called "essences" by other philosophers, and, following that terminological tradition, the phenomenologists also talk about essences.

Many philosophers in the past have held that statements about essences are empirical statements, arrived at by comparison of many examples of a type of object and extracting from the descriptions of all these examples the common features by means of some kind of generalization. Such a process has often been called abstraction. Abstract statements, since they are logically dependent on empirical descriptions of particular cases, are themselves empirical statements. Phenomenological statements, on the other hand, are, for the reasons given, not empirical statements. Hence, phenomenological statements are not reached by abstraction. They are, phenomenologists say, derived from a scrutiny of particular cases by seeing, intuition, or intuition of essences (*Wesensschau*).

The identification of phenomena as essences brings us one step closer to the goal of clarifying the particular way of looking at objects that reveals objects as phenomena. It turns out to be a species of intuition. Phenomenology is a form of intuitionism and has, accordingly, acquired the ill repute of all intuitionisms of being no more than a veiled refusal to provide evidence for one's philosophical statements. But sometimes such a refusal can be justified. Intuitionism is objectionable only if the philosopher is not willing to argue either about the nature of his intuition or about the justification for appealing to it in this case—if his appeal to intuition is merely intended to terminate philosophical debate. The phenomenologists' appeal to intuition is not of this kind. Hence more can, and must, be said about intuition.

Intuition seems to be a psychological term. Its German counterpart, *Anschauung*, often means no more than "seeing." The objects of seeing, in its ordinary sense, are empirical objects. Essences are not empirical objects, so they cannot be seen in any ordinary sense of that term. Hence, intuition must be seeing of some extraordinary kind. One might suggest that the phenomenologists claim to have discovered one more human cognitive faculty than had been known before, but such a discovery of an actual human faculty would have to be couched in empirical statements. Phenomenologists do not make empirical statements, so they cannot claim—nor do they—to discover previously unknown cognitive faculties.

The point of introducing intuition is not psychological but epistemological. To appeal to intuition is not to make a psychological statement about the causal origins of certain statements but an epistemological one about the sort of evidence that will be relevant to them. To say that we know essences by intuition is to say, negatively, that the truth or falsity of statements about essences is not dependent on the truth about empirical statements.

The appeal to intuition makes another positive, epistemological point: Our acquaintance with essences possesses an epistemological feature also possessed by our sensory acquaintance with empirical objects. This logical feature is sometimes described by saying that what we see is described in self-validating statements. A statement, "*P*," about particular objects is self-validating if the strongest evidence that we can adduce for it is a statement like "I have seen that *P*" or "I have observed that *P*." We cannot, therefore, claim that "*P*" is true because there is some other true statement, "*Q*," from which "*P*" can be inferred and that is not equivalent to "*P*." Statements about essences are self-validating in the same sense. Given any statement, "*E*," of the form "________ is the essence of ________," we cannot claim that "*E*" is true because

there is some other true statement, "*F*," which is not equivalent to "*E*" and from which we can infer "*E*." Of course, some statements about the existence of particular objects may be deducible from other statements, and it is similarly true that some statements about essences may be deducible from other statements. But such a deduction does not provide stronger evidence for statements about empirical existence or about essences than do self-validating statements.

Phenomenological statements are not derived by means of abstraction from particular statements, since, if they were so derived, they would not be self-validating. But they are not the only self-validating statements; empirical statements are also self-validating. An adequate account of phenomena must state more than that phenomena are revealed in the intuition of essences; it must specify this intuition to clarify in what respects it differs from the simple seeing of objects of sensory observation.

BRACKETING EXISTENCE: FREE IMAGINATIVE VARIATION OF EXAMPLES

In the light of the problem about the meaning of intuition, the reason for introducing a further condition defining "phenomenon" becomes clearer. This condition is not accepted by all phenomenologists but was regarded as necessary by Husserl, Pfänder, Reinach, and Scheler. We are in a position, they said, to describe objects as phenomena only after we have "bracketed existence" or "suspended our belief in the existence of objects." Husserl calls this the "phenomenological epoche" or the "phenomenological reduction." *Epoche* was borrowed from the Skeptics, but Husserl's use of it differed from theirs.

These references to "bracketing" or "suspending belief in existence," together with the talk about essences, led to the view that phenomenology is a kind of essentialism and, as such, is diametrically opposed to existentialism. There is no room here to bring out all the confusions that produced this fairly common interpretation; suffice it to say that the phenomenological epoche is not achieved by resolving to make no more statements about existence or what exists. To bracket existence is not to eliminate existence in general or existing entities in particular from the list of possible objects for phenomenological study.

In the light of Husserl's repeated insistence on the close similarities between his phenomenology and René Descartes's methodical doubt, the phrase "suspending belief in the existence of objects" is often taken as a description of Cartesian doubt. But this is a misunderstanding, for Husserl insisted on distinguishing suspending belief in existence from doubting existence. This distinction cannot, therefore, be simply ignored.

Suppose a young woman states that she has direct evidence that she is terribly attractive to red-haired men. Her statement is not derived from a psychological law about the preferences of red-haired men or from a physiological one about their exceptional susceptibility to her figure and coloring. Her statement, a direct inductive generalization, is the result of her own experiences with red-haired men and tells us something about many or all of the members of the class of red-haired men. Besides all being red-haired and male, they have one further property: They cannot resist the charms of this young woman. In order to substantiate such a statement, she would have to cite cases of a number of red-haired men who at various times, under various circumstances, have given indubitable proof of their devotion. Two things are important here: that the red-haired men really exist and that their devotion to her is real. The truth of the inductive generalization depends at least on those two conditions. On the other hand, if the generalizations are correct, it follows that there exist (or existed) several red-haired men in this particular condition. If, however, the red-haired men do not exist or if their attachment is a figment of this young woman's imagination, then the general statement is false (unless evidence of a different kind can be found).

The story of this young woman was told in order to exemplify the relation of empirical generalizations to particular empirical statements—of "I am irresistible to red-haired men" to, for instance, "A red-haired matinee idol in New York committed suicide over me," and of both of these to the facts of the case. These relations were exemplified with an imaginary example, for it is quite unimportant that I do not know any young woman of this description. Where a description serves as an example in this sense (*example* is an ambiguous word), it is quite irrelevant whether the object described exists or not. If, on the contrary, I am interested in making a general statement about objects observed, it makes all the difference in the world whether the particular objects covered in my generalization exist and exist as described.

This is one sense of "bracketing existence." When existence has been bracketed in this sense, the descriptions of objects or situations do not serve as premises for an inductive generalization (or an abstraction), but as examples. But "example" is used in several senses. Sometimes it is used to designate one instance of an empirical generalization, but this is not the sense used here. At

other times, examples serve a merely pedagogical function. I might have told my story about the young woman merely to provide a concrete illustration of abstract truth about empirical generalizations, in order to make the abstract statement easier to understand. In a third sense—"example" is used in phenomenology in this sense—the example both serves as an illustration and has evidential functions. In that case, the truth of the statement about empirical generalizations depends on the accuracy of the description of the example. I claim that my general statement is true because the description of the particular example is accurate, but how do I know whether a description is accurate so that it can have evidential force as an example? Since we have bracketed existence, I cannot say that the description is accurate because the case described has actually been observed to exist in a particular place and at a particular time, for examples need not be actual existents.

In order to understand this sense of bracketing existence, we must be able to answer two questions: (1) When can the description of an example rightly be said to be accurate? (2) How is a phenomenological statement to be derived from an example?

In this context Husserl talked about a procedure that he called "free imaginative variation," comparable to what Anglo American philosophers call the method of "counter-examples." Here we describe an example and then transform the description by adding or deleting one of the predicates contained in the description. With each addition or deletion, we ask whether the amended description can still be said to describe an example of the same kind of object as that which the example originally described was said to exemplify. Sometimes we shall have to say that if we add this predicate to the description or take that one away, what is then described is an example of a different kind of object from that exemplified by the original example. At other times the additions or deletions will not affect the essential features of the kind of object exemplified by the different examples.

In this way we discover the necessary and invariant features of a given kind of thing that the example must possess in order to be an example of that kind of thing. We also discover which features are accidental and hence irrelevant to the question whether this object, as described, is or is not an example of a certain kind of thing. What we discover is what phenomenologists call the "essence" of objects.

For example, let us suppose that we meet someone who does not have the usual five senses but only three: sight, touch, and hearing. We might be perplexed, but we should still call him a person. The same would hold if he had three more senses than normal persons. But suppose we met someone who looked like a person but seemed to be deaf and blind, and without any tactile, olfactory, or gustatory sensations. He would still be regarded as a person, although as a seriously defective one. But suppose further that we find that this creature looks like a human being except that it has no sense organs at all. Would he nonetheless be called a person? No. An animal? No. A plant? Not really. We have no word in our language for such a being. We would not know what to say about it.

Here we have varied in imagination an example of a person with reference to one predicate, "possessing sense organs." We find that in order for anything to be a person, it must have sense organs of some kind; there is an essential (necessary and invariant) relation between "person" and "possessing sense organs." The results of free imaginative variation are statements of such essential connections. Since statements about phenomena are one kind of statement about essences (and vice versa), the statements resulting from this procedure are phenomenological statements.

"EPISTEMOLOGICAL CIRCLE." Phenomenological statements are made while existence is and remains bracketed. If true, they are so not because they describe something that we have directly observed. Nor are they true because they are warranted by a series of observations of particular objects or events. Hence, they do not imply the past or present existence of particular objects in just the way in which empirical generalizations imply it. All that is asserted in the phenomenological statement is that if any being is an example of a person, then it must have sense organs. We are, therefore, making an assertion about the necessary relations of properties: Whatever has the property of being a person must also possess the property of having sense organs.

This is the method of free imaginative variation. It would seem to provide an answer to the second question raised earlier—how a phenomenological statement is to be derived from an example. But the same procedure can also be said to provide an answer to the first question, how we decide whether an example is described accurately—whether the description contains all the essential predicates so that the thing described may rightly be said to be an example of a certain kind of object. For, once we have made clear the invariant features of the sort of thing exemplified, we are in a position to say whether the example contains all those necessary features. But to use free imaginative variation to answer both questions is, of

course, circular; we derive the phenomenological statement from any given example by means of free imaginative variation and then confirm that the original example was accurately described because it possesses the invariant features expressed in the phenomenological statement. It would seem that we need an independent criterion for deciding the accuracy of the description of any given example, but there is no discussion of such an independent test in the writings of the phenomenologists. The phenomenological method appears, therefore, to be circular in a second sense that might be called the "epistemological circle."

Phenomenology, as we saw, is circular because it clarifies its own method while using it (the methodological circle); it is also circular, we see now, because it confirms its statements by reference to examples and then attests the accuracy of the descriptions of these examples by reference to the statements derived from them (the epistemological circle). We must now show that what we claimed earlier for the methodological circle—that its presence cannot be construed as an argument against phenomenology—is true for the epistemological circle as well. This will be argued for by an examination of a second sense of "bracketing existence." In this second sense, "bracketing existence" refers to the transition from nonreflective to reflective thinking.

BRACKETING EXISTENCE: PHENOMENOLOGY AND REFLECTION

In free imaginative variation we ask ourselves about any given property of an example, "Is this a necessary feature for being a such and such? Is that?" For our answer we do not appeal to empirical observation. Neither do we give an answer simply by deciding to regard some particular feature as essential. We do not define our terms arbitrarily; instead, with each variation, we ask ourselves whether the example described could still be recognizable as an example of the same sort of thing as that exemplified before. We ask ourselves what features an object must have in order to be recognized as an example of a certain kind of object. What we discover are necessary conditions for recognizing a certain kind of thing.

But recognition presupposes previous acquaintance. I cannot recognize someone whom I meet for the first time, unless I have seen pictures of him or have been given his description or perhaps dreamed of him before. But if we can recognize only what we know already, then we must already know the necessary features of the objects that we are able to recognize. In that case, there would seem to be no need to bracket existence and to vary the examples freely in imagination in order to discover their essential features, since the entire procedure presupposes that we know these essential features all along.

The resolution of this difficulty comes when we consider that the word *know* has two radically different senses, which some English philosophers have called, respectively, "knowing how" and "knowing that." The latter refers to knowledge expressed in statements. To "know that" something is the case is to be able to put what is known into words. I can show that I know a person by describing his looks; however, it is of course also possible that I should know a person and yet be quite unable to give any sort of adequate description of his looks. It is often very difficult to give a good description of those persons whom we know very well. I know them, not in the sense that I can describe them but that I could recognize them anywhere. I can pick them out of a crowd without hesitation. I can identify them by their voice or their walk, although I might be hard put either to describe in words or to imitate them. This second kind of knowledge is "knowing how"; in the example, I know how to recognize a person.

These two kinds of knowledge are independent of each other. It is not a necessary condition for being able to do something, such as recognize someone, that I should be able to *say* that he is a person of a certain description. Conversely, it is not necessary that I should be able to do a certain kind of action, such as ride a bicycle, in order to be able to give a detailed and accurate description of riding a bicycle. It is, furthermore, possible that for certain kinds of knowing how there is no corresponding knowledge that.

Of some performances I can say: This time I did it right; last time I did it badly. Therefore, I possess criteria for proper performance. If asked what these criteria are, I may not be able to put them into words, but I know them in the sense that I use them and, in many cases, I can, upon reflection, state what they are. I have then, by means of reflection, produced knowledge that ________ corresponding to the knowledge how ________ which I possessed all along. This is what happens when I vary an example freely in imagination: I am always able to discriminate between the thing that I would recognize as a certain object and the thing that I would either take as a different kind of object or about which I would not know what to say. But only upon reflection can I verbalize the criteria implicit in such a recognition by stating the essential features of any given kind of object.

REFLECTIVE THINKING. When I vary examples freely in imagination, I reflect about the criteria implicit in my ability to recognize examples of the given sort of object; I now put into words the criteria that previously were merely implicit in my performances. This description of the two sides of the process called "bracketing existence" accords perfectly with Husserl's explanations of it. Phenomenology, he stated, is a reflective enterprise. In its reflection it brings to light what was previously "anonymous" or "latent" in our "performances" (*Leistungen*). But phenomenological reflection is a very special kind of reflection. In phenomenology we do not reflect about facts ("Did I see right? Was that really Jones lying in the gutter?") or about specific actions ("Should I have lectured Jones on the evils of drink?"). Phenomenological reflection does not produce any factual statements, nor does it employ factual statements as premises or as the starting points of reflection. In phenomenology we reflect about examples, in the sense explained; the result of such a reflection is not a factual statement or an empirical generalization but a statement about the necessary conditions for any object's being an example of the sort of thing considered in our reflection.

"Bracketing existence" and the other phrases applied in this context are used ambiguously. Why did Husserl fail to distinguish these two senses? We have already uncovered one source of this ambiguity by showing that we can employ the method of free imaginative variation of an arbitrarily chosen example in order to clarify the essential feature of any object only if we reflect about the example. Hence, treating a given case merely as an example (bracketing in the first sense) presupposes that we have made the transition from nonreflective to reflective thinking (bracketing in the second sense). Although the two kinds of bracketing are distinct, they must occur together.

But there is a second source of the ambiguous use of all these phrases. "Bracketing existence" and "suspending our belief in the existence" of an object seem to be particularly apt in describing important features of the transition from nonreflective to reflective thinking. Reflection involves questioning—more specifically, questioning something that I believed before or regarded as properly done. When I reflect, I ask, "Was that really Jones in the gutter?" or "Should I have helped him up?" Such questioning requires awareness that there are questions to be asked in this situation and that they are not pointless. Before I can reflectively question my earlier belief that it was Jones whom I saw lying in the gutter, I must be open to the possibility that it was not Jones. Hence, as I begin to reflect, I suspend my belief in the existence of Jones in that condition in that place, or I put his existence in brackets. "Bracketing" in this sense means that I become aware of the possibility that something which I believed to exist does not exist as I thought it did, that a statement which I considered true is not, or that some act which I considered right when I did it might have been wrong. Once I have become aware of that possibility, I am ready to reflect.

The insight that phenomenological statements are the product of reflection resolves the methodological and the epistemological circles. The methodological circle arises because the method must be used to clarify what the method itself consists of. It seems, therefore, that we can use the method only if we know what it consists of, but we can know what it consists of only if we have already used it. Therefore it would seem that we can never get started. But since phenomenology is reflective, it does not presuppose knowledge that the phenomenological method consists of certain procedures; it only presupposes that we know how to use it (to reflect about the essential features of arbitrarily chosen examples), even if we cannot describe it. Such a description is not a necessary condition for using the method, so there is no problem here.

The epistemological circle is resolved in a similar manner. In the method of free imaginative variation, it seemed that we could know that a given phenomenological statement, "*P*," is true only if we know that the description, "*E*" of the corresponding example is accurate. But we can know that "*E*" is accurate only if we know that "*P*" is true. Hence, it would seem that we cannot know either that "*P*" is true or that "*E*" is accurate. But phenomenological reflection begins with my being able to recognize the example described in "*E*." I know that I describe the example accurately to the extent that I recognize the object in my description of it. Both the accuracy of "*E*" and the truth of "*P*" are tested by the criteria implicit in my ability to recognize the object. Hence, there is no difficulty in this case either.

NONEMPIRICAL STATUS OF PHENOMENA. In the search for a complete definition of phenomenon we have now discovered three conditions defining phenomena: (1) phenomena are essences, (2) phenomena are intuited, (3) phenomena are revealed by bracketing existence. The third requirement is twofold: Phenomena are known only upon reflection of a specific sort, namely, reflection about the essential features of arbitrarily chosen examples. Once again the question must be raised whether this definition of phenomena is complete. The criterion of

completeness used earlier was that a definition of "phenomenon" is complete only if it is consistent with the first of the three requirements for phenomenological statements—that they are nonempirical. We must ask, therefore, whether phenomena as defined can be described in empirical statements or whether our definition has ruled out that possibility.

It may seem obvious that the definition of phenomenon is complete by this criterion because it seems impossible that phenomena as defined—as being revealed only by bracketing existence—could be described in empirical statements, for statements about phenomena are not statements about single, observed particulars or about series of such single, observed particulars. They are, rather, statements about the necessary relations between the properties of some example of a certain kind of thing in which we do not consider whether the description of our example refers to an actually existing object.

But can we really conclude from this fact, namely, that no observation of actually existent objects is consulted in phenomenological reflection, that the truth of phenomenological statements is independent of the truth of empirical observation statements? We must distinguish between the description of the process by which we arrive at phenomenological statements and the logical conditions that these statements must fulfill in order to be true. The former merely describes how I discover certain statements, but it reveals nothing about the truth conditions of my statements. It is said, for instance, that some Greek geometers discovered certain statements about plane figures by measuring and weighing actual plane figures of tin. They arrived at their statements by means of observations; they were able to make certain statements in geometry after observing actual physical objects, but their statements are no more empirically true (or false) than are the same statements when they appear as theorems in Euclid's *Elements*.

This example presents a case in which statements whose truth or falsity is independent of empirical observation are discovered through empirical observations. It is possible that statements about phenomena constitute a converse case where empirical statements are discovered without explicitly consulting observation of sensory particulars. For instance, it was stated in the preceding section that the phenomenologist does not necessarily consult actual observations when he describes phenomena; his example may be purely imaginary. But it is possible that the statements that he is thus able to make are nevertheless empirical statements. All that was said was that the making of a phenomenological statement is not immediately preceded by observations of existent objects.

Perhaps, however, this is not necessary, since we know the necessary conditions for anything to be an example of a certain kind of thing because we have observed examples of this kind of thing many times and have, as it were, performed an unconscious induction all along. If this is true, then phenomenological statements may still be empirical statements. That they are not empirical statements has not been proved by stating that they are not discovered by means of explicit and deliberate observation of existing objects. The description of "bracketing existence" and of the subsequent reflection has revealed something about the method of discovering statements in phenomenology, but it has not shown that the statement so discovered may not nevertheless be empirical in the sense of being either verifiable by reference to observations of particulars or confirmable or at least refutable by reference to such observations.

There is reason to suspect that the phenomenologists who required that existence be bracketed in phenomenology thought that this requirement assured them that the statements so discovered would not be empirical in any of the senses mentioned. But, as has been shown, they have no such assurance. Hence, they can have no assurance that what is discovered once we have bracketed existence is a phenomenon, in the sense of being the referent of a nonempirical statement. We need further argument to show that bracketing existence does reveal phenomena in the required sense, in all or at least in some cases. Some of the phenomenologists, notably Husserl, have brought forward a number of considerations that provide the arguments needed here. These considerations can best be approached by considering intentionality.

INTENTIONALITY

It was said earlier that reflection undertaken after we have bracketed existence yields, if successful, descriptions of activities that we perform with ease in everyday life but are not able at the same time to describe. Concerning such activities we also know when they miscarry, when they are performed incorrectly or in an improper context, or when someone mistakes such an activity for a different one. We possess criteria for correct and appropriate performance and identification of such activities but are, ordinarily, unable to formulate them. Reflection subject to bracketing of existence yields formulations of these criteria. The phenomenologists regarded all statements resulting from such reflection as nonempirical, but there is no ground for thinking that this is true. These phe-

nomenologists also believed that all the activities that are reflectively described and clarified after bracketing existence are intentional activities. This view can also be shown to be open to objections, but from these two doubtful assertions we can extract a more defensible characterization of phenomena than the one reached so far. So far three necessary conditions for phenomena have been listed: (1) They must be essences that are (2) intuited (3) as the result of the exemplary reflection that requires bracketing existence. We now add a fourth condition, namely, that statements about phenomena must be limited to statements about intentional acts.

The noun *intentionality* does not refer to a thing (as does, for instance, *sodality*) but to the state of an entity—the state of being intentional. Although Husserl used *intentional* in all kinds of contexts, in its primary sense it is an adjective modifying "act"; being intentional is a characteristic of acts. In this employment, "intentional" has an ordinary meaning as a synonym for "deliberate" or "done on purpose," and a philosophical meaning different from, although related to, its ordinary, nonphilosophical meaning. The philosophical use of the term dates back to scholastic philosophy. Later, it completely disappeared from the philosophical vocabulary until it was reintroduced in 1874 by the Austrian philosopher Franz Brentano. Husserl, a student of Brentano's, gives credit to Brentano for reintroducing intentionality into philosophical discussion but adds that intentionality became a fruitful philosophical concept only in phenomenology.

Intentional acts have four aspects, and there are four distinct questions we can ask about them. The sentence "Luther thought that the devil was in his cell" is the complete description of an intentional act. We can ask who is performing an intentional act, and the answer consists of a proper name ("Luther"). It could also be a personal pronoun ("I" or "we") or a definite description ("the father of the Reformation"). We can, in the second place, ask what this person is doing, and the answer will consist of the inflected form of a verb ("thinks," "thought"). The third question concerns the intentional object of the act, what the act is about. In the example, Luther is thinking about the devil. Finally, we can ask in what manner or under what description the intentional object is object of the act; in the example, what is Luther thinking about the devil? "The devil is in my cell."

The intentional act, having four elements, is a tetradic relation. So, for instance, is the relation described in the sentence "I place the book on the table." Here also there are four elements: the subject or agent (myself), my action (placing), what I place (the book), and the table on which I place it. There is, however, an important difference between the two cases. The second statement is false unless there is a table on which I place the book. If the statement as a whole is true, the final of the four terms in the tetradic relation must also exist. It would be self-contradictory to say "I place the book on the table … but there is no table."

We can therefore infer the existence of the table from the truth of the statement "I place the book on the table." This is not so in the case of intentional acts. If it is true that Luther thought that the devil was in his cell, it is not therefore true that the devil exists, let alone that he was in Luther's cell. Luther might have had hallucinations; he might have been the victim of religious madness; or he might have been drunk. All three of these are situations in which we are inclined to see things that are not there or to believe that things exist which in fact do not. Nor can we conclude from the truth of the original sentence that the devil does not exist or was not in Luther's cell. The same holds of whatever is thought or believed to be the case. A belief that my wallet was stolen or that there are leprechauns does not allow the inference that there was a thief who stole my wallet or that there are leprechauns. The same is true of perceiving, of hoping, expecting, doubting, fearing, and all similar activities. The truth of a statement describing someone's intentional act does not allow the inference of either the existence or the nonexistence of what the act is about. This distinguishes intentional acts and their four elements from genuine tetradic relations, where the existence of all four elements can be inferred from the truth of a description of the relation.

THE NONINFERENCE CRITERION. The usual discussion of Husserl's doctrine of intentionality presents intentionality as (1) the defining characteristic of consciousness in the ordinary sense of that term, which (2) consists in the fact that all consciousness is consciousness of something. The first point is false; the latter is true but trivial. It merely asserts that to be conscious is to be related to something. But I am also related to something if, for instance, I own property. In that case I am the owner of something. But being the owner of something is not an intentional act because the existence of the object owned can be inferred from the fact that I own it. The existence or nonexistence of the object of the intentional act, however, cannot be inferred from the true description of that act. (We shall call this the "noninference criterion"). This, rather than merely being related to an object, is the property of intentional acts that distinguishes them from all other kinds of tetradic relations. Hence, it is a defining characteristic.

Two examples will show that intentionality is not the exclusive property of consciousness. Consider the sentences "Luther threw an inkwell in order to injure the devil" and "The rat pushes the lever in order to obtain food." Both sentences express tetradic relations: the agent (Luther, the rat), what he does (throwing, pushing), what he does it with (the inkwell, the lever), and the object of the activity (injuring the devil, obtaining food). It may be said that these are not intentional acts because the object in each case is not what the act is about but is, rather, an aim or a purpose. The acts described in these two sentences are intentional in the ordinary sense of being purposive, but according to the noninference criterion, they are also intentional in the philosophical sense because we cannot infer from the first sentence that the devil was injured and hence we cannot infer that the devil exists or does not exist, nor can we infer from the second that food was obtained by pushing the lever.

The acts described in the two sentences are not acts of consciousness or mental acts in the traditional sense. Throwing and pushing have traditionally been regarded as physical acts, but they differ according to the purpose served. When throwing something at a person in order to injure, one throws differently (much harder, for instance) than when one throws someone a cigarette in order to be helpful. Although physical, both of these acts are intentional in the philosophical sense. Hence intentionality is not, as Brentano thought and Husserl thought at certain times, the defining characteristic of consciousness in the ordinary sense. Husserl became aware of this and redefined "consciousness," in his later writings, by extending the term beyond its ordinary meaning to apply not only to mental acts but also to all kinds of activities, even to those usually regarded as physical, as long as they are intentional. Here intentionality became the defining characteristic of consciousness because this was how consciousness was defined. Husserl would perhaps not have wanted to apply "consciousness" to the behavior of animals, but his views on this point are not well known.

Inference. The verb "to infer" is used in a variety of senses in English, so it must be made clear in what sense it is used in the formulation of the noninference criterion. Suppose I see my foot as it sticks out unshod from my trouser leg and I say, "There's my foot." If someone asks me why I think that my foot is there (exists), I answer, "Because I see it" (or "Because I see something that looks like my foot"). In a loose sense of *infer*, I may be said to infer the existence of my foot from the fact that I see it. In this sense of *infer*, therefore, the correct description of an intentional act ("I see what looks like my foot") allows me to infer the existence of what I see (my foot). But this is inference in a loose sense. The conclusion does not follow necessarily from the premises. It is possible that the premise should be true and the conclusion false, as happens, for instance, when I am having hallucinations. There I see what looks like my foot, but the foot is not there. Common examples of this are the so-called phantom feelings—an amputee feels his foot long after it has been amputated. It is true that he feels his foot, but it is false that his foot is there. But if I say that I know my foot is under the table because I feel it, the inference (in the loose sense) is correct.

The sense of "to infer" used in the noninference criterion is stricter. In this sense we say that something is inferred from a premise or set of premises if the falsity of the conclusion is incompatible with the truth of the premise(s). In this sense it was said earlier that we can infer from the truth of "I place the book on the table" that there is a table. It would be self-contradictory to say "I place the book on the table … but there is no table" and to claim that both parts of this compound statement are true. It is in this stringent sense of "to infer" that the noninference criterion denies that we can infer the existence of the object of the intentional act from a true description of the act itself. The noninference criterion does not deny that feeling my foot, for instance, is often sufficient ground for saying that my foot is there. But it does deny that my foot must exist necessarily if I feel it. Intentional acts differ from other tetradic relations in that it is not inconsistent in the case of intentional acts to deny the existence of the final term of the four-term relation and to assert that the relation is described truly, but it is inconsistent to do this in the case of all nonintentional four-term relations.

Criterion is nonempirical. It is now easy to show that a statement of the noninference criterion is a nonempirical statement in the sense that no empirical statement can show it to be false. In this sense mathematical statements are nonempirical—no measurement of angles or lines in a triangle can show that geometrical statements about triangles are false. If there does appear to be a conflict between actual measurements and measurements predicted on the basis of certain geometrical propositions, we do not reject the geometrical proposition underlying our prediction; rather, we conclude that the measurements are false. The reason for this is, of course, that the procedures used in measuring presuppose the truth of the pertinent statements in geometry. In order to show that the statement of the noninference criterion is false, there must be at least one intentional act in which the

existence of what the act is about or aims at follows with necessity from a true description of the act. But philosophers agree that no necessary relations are observed, or can be inferred from observations, so no statement about a necessary relation can be an empirical statement. Hence, the case needed to refute the noninference criterion cannot be described in empirical statements. It follows that the statement of the noninference criterion, not being refutable by means of an empirical statement, is not itself an empirical statement.

INTENTIONALITY AS A PHENOMENON. The statement of the noninference criterion satisfies the fourth condition laid down for phenomena: It is a statement about intentional acts. It is easy to show that it also satisfies the other three conditions for phenomena: (1) The preceding analysis consisted of reflection subject to bracketing of existence. (2) It brought to light certain essential features of intentional acts. (3) The truth of the statements rests on intuition, in the sense discussed earlier. Intentionality is, therefore, not only one mark of phenomena but is also itself a phenomenon. It has also been shown that the description of this phenomenon contains at least one nonempirical statement, namely, the noninference criterion. There is, then, at least one statement about phenomena, as now defined, that is nonempirical. This suggests that the four conditions for phenomena constitute a complete definition. However, the four conditions for phenomena are not sufficient for a complete definition, so a fifth condition must be added—that, with respect to intentional acts, phenomena serve as criteria of coherence.

CRITERIA OF COHERENCE

Intentional acts are of two kinds; they are either purposive or about something. Purposive acts may be said to be adequate to their intentional object if the means chosen accomplish their purpose. Acts that are about some intentional object may be said to be adequate if what is believed or asserted about an object is really true, if what is questioned is questionable, if what is doubted is doubtful. Whether a given purpose is pursued correctly by using certain means depends on the nature of the purpose and of the means chosen, and on the way the means are used. Whether Luther throws the inkwell correctly at the devil depends on the weight of the inkwell, the distance between him and the devil, and how he throws. There are correct and incorrect ways of throwing inkwells or anything else. Which ways are correct and which are not is a matter of empirical fact, to be discovered by empirical study. Hence, rules about correct performance of this kind of intentional act are empirical rules. Similarly, it is in many cases an empirical question whether my beliefs are true, whether what I question is questionable, or whether what I doubt is doubtful. It can be shown that at least some of these rules satisfy all four defining conditions for phenomena; hence, they can be regarded as statements about phenomena, as defined so far. This, in turn, shows that the four conditions laid down do not constitute a complete definition of “phenomenon,” for phenomena, under this definition, are capable of being described in empirical statements. We need a fifth condition.

The following consideration will yield the required fifth condition for a complete definition of “phenomenon.” Before we can ask whether any given intentional act is correctly performed—whether it is adequate to its intentional object—we must be certain that what we are asking about is a genuine intentional act. Since intentional acts have four elements—the subject (or agent), the action, the intentional object, and either the means used or what is asserted about the intentional object—we need certain rules to determine which subject can be combined with what actions, which intentional objects, and which means or assertions to form coherent intentional acts. Not just any member of each of these four classes of elements can be combined with any other to form a coherent and intelligible intentional act.

COHERENCE AND INTELLIGIBILITY. The meanings of “coherent” and “intelligible” are best indicated by examples of their opposites, intentional acts that are incoherent or unintelligible. Purposive acts are not coherent and not intelligible (they “make no sense”), for instance, where the action and the means used are inappropriate to the intentional object. Someone might have said to Luther that it made no sense to throw anything at the devil because the devil is not a person but merely a symbol of evil. Not being a person, the devil has no body—and hence no location—and therefore cannot be made the target of any physical missile. A different case of an incoherent purposive act is that in which the means are inappropriate to the action. “Killing a person with kindness” is a metaphorical expression precisely because it literally makes no sense; the means chosen for killing a person are utterly inappropriate. They are not inappropriate merely in the sense that someone might try to use kindness as a murder weapon and discover that it does not do the job. It is not at all clear how one would proceed literally to try to kill someone with kindness. “Killing a person with kindness” is therefore not an intelligible or coherent intentional act. Similar incoherences can be

found in the other relations among the four elements of intentional acts.

Corresponding incoherences appear in intentional acts that are about something. If what I believe about something is utterly inappropriate to its intentional object, such as "The Pythagorean theorem is mellifluous and sweet-smelling," there is no way of telling or even of finding out whether the statement is true. Asserting this sentence is not an intelligible intentional act, and hence the assertion is neither true nor false. Similar incoherences can occur between the action (for instance, "I predict") and its intentional object (for instance, some past event) or what is being predicted (that something happened yesterday).

So far the notions of coherence and incoherence, of intelligibility and unintelligibility, have been exhibited within single intentional acts. Husserl pointed out that there is also coherence and intelligibility of series of acts.

Suppose that Luther, rage suffusing his face, threw an inkwell at the devil with all his might and the very next moment rushed up to him, saying, "My dear fellow, I am so sorry. How very clumsy of me. Here, let me help you." This would be very surprising because the first action seemed clearly intended to injure, the second to placate. The change between the two is unmistakable and can be described by saying that the second act has a different intentional object from the first. As juxtaposed, the two acts make no sense because they seem to be members of two incompatible series of acts. The first act seems part of a series intended to enrage or injure the devil, and the second seems part of a different series aimed at mollifying the devil. The first action clearly leads to the expectation of another angry action. The second one disappoints that expectation, so the two actions make no sense, although each by itself makes sense. As single acts they are intelligible or coherent, but they do not make sense when they come in the above order. No one can understand what Luther is up to. We know what a man is up to if we understand a sequence of his actions and have correct expectations about what he is going to do next. If our expectations are disappointed, we may conclude that the agent has changed his mind or that we did not understand him to begin with. We understand or do not understand what someone is up to if his purposive actions form a coherent or incoherent series, respectively.

All this is true irrespective of whether the series of acts is performed well or badly. Hence, there are two sets of rules governing series of acts that correspond to the two sorts of rules governing individual acts: those which govern the coherence of act series and those which govern the adequacy of the act series to its collective purpose. What a man is up to in a series of acts can be inferred only from the sequence of acts performed. But not all sequences of acts are coherent. There are, therefore, rules about intentional acts determining the conditions for coherence of any series of intentional acts. Only if a series is coherent corresponding to the rules governing coherence can the question whether the actions and the means chosen are adequate to the aim pursued in the whole series be answered in the light of the relevant facts. Empirical statements about the adequacy of actions and means to their collective end are to be distinguished from statements about the coherence of such collections of acts.

It is not necessary to cite more examples to show that a series of acts which are about something are coherent or incoherent, intelligible or unintelligible, in analogous ways. A single act of belief, assertion, or questioning may be perfectly coherent and intelligible by itself but may be entirely out of context with what precedes or follows, and it is not understood what this person, in this act, is talking about, what he is trying to say.

HORIZON. Husserl used the term *horizon* to refer to the relations of coherence and incoherence of intentional acts. *Horizon* was not intended to refer to the place where sky and earth meet but to the edge of the perceptual field, which moves and changes with movements of the head or of the entire body. The horizon metaphor suggests that as the edge of the perceptual field (the horizon) leads us to expect a continuation of what lies before us, so any given intentional act suggests further acts that would be continuous or coherent with it. What is said in one act or done in one purposive action leads one to expect a second assertion or a second action continuous with the first. The second statement is continuous with the first if it is about the same object as the first; if in the second action one is up to the same thing as in the first, the two are continuous. I know what you are talking about or what you are up to when I know what sort of thing you will say or do next.

The horizon metaphor also implies that these relations between intentional acts are necessary conditions for any act being intentional, just as it is a necessary condition for the existence of a perceptual field that it have a horizon. Something is an act of asserting, for instance, if and only if I can repeat what I said in another way; if I can amplify, clarify, explain what I said; or if I can confuse, muddle up, and utterly obfuscate what my assertions are about. It is impossible that an intentional act should be

without horizons, that is, unrelated to any other intentional act.

Criteria of coherence. As the horizons of the perceptual field are to some degree indefinite, so are the horizons of intentional acts. I cannot infer from any given assertion or activity of yours that you will next assert one particular statement or do some particular action and no other. When I see a church steeple on the horizon, I know that, when I come closer, I will not see a hippopotamus at its base. But there is definitely a point in coming closer to discover what the church or the building that resembled a church from a distance looks like.

Similarly, there is a point in listening to you to find out what your next statement is going to be or in watching what you are going to do. If I understand what you are talking about or what you are up to, I have some idea of what you are going to say or do next. I know the minimum conditions for your next statement and action; I know the limits beyond which your next action will not be continuous with the last or your next statement will not be about the same object as the last. Horizons are the necessary conditions for any series of assertions or activities to be intelligible. Different kinds of intentional acts have different kinds of horizons. Linguistic acts are related in terms of their meaning; purposive activities, by reference to the purpose. It is the task of phenomenology to clarify the different sorts of horizons (conditions for intelligibility) and to put into words what the horizons of individual examples of each kind of act are. Husserl called the clarification and formulation of horizons "intentional analysis." The results of such intentional analyses are statements of the criteria for the coherence of intentional act series.

Having understood what Husserl meant by "horizon" and that there are criteria for the coherence of single acts corresponding to the horizons in act series, we have found the fifth condition defining "phenomenon." Statements about phenomena must, besides satisfying the first four conditions, be about the criteria of coherence of single intentional acts or of sequences of intentional acts. When we look at any object as a phenomenon, we are trying to discover the criteria for coherence of those intentional acts in which the object (or its name or description) can figure.

ARE PHENOMENOLOGICAL STATEMENTS A PRIORI?

Traditionally philosophers have called statements "a priori" if they are (1) nonempirical and (2) necessarily true. Phenomenologists have always held that their statements are a priori. The two parts of this claim must be examined separately.

It has been shown that phenomenologists agree that their statements are nonempirical, although they disagree about the description of phenomena. Some phenomenologists were content to describe them as essences intuited, but others regarded this as insufficient and added that phenomenological descriptions must be preceded by bracketing existence. But bracketing existence also turned out to be an inadequate guarantee that phenomenological statements are nonempirical. Therefore some members of the phenomenological movement, notably Husserl, added further requirements for statements about phenomena. The preceding discussion can be summarized by stating the five conditions that any statement must satisfy if it is to be a statement about phenomena:

(1) It must be about essences.

(2) It must be self-validating (intuitive).

(3) It must be the result of bracketing existence.

(4) It must be about intentional acts.

(5) It must lay down the criteria of coherence (or intelligibility) of intentional acts.

We must now, once again, ask: Are statements of this kind nonempirical?

THE SENSES OF *EMPIRICAL*. The above question is not easy to answer because the term *empirical* has several meanings. We must examine some of them.

Statements asserting particular matters of fact, such as "There is a fire burning in the fireplace," are true if observation shows them to be true and false if observation shows them to be false (for instance, that the fire has gone out). They are empirical because one observation will show them to be true or false.

General statements, such as "Continuous nervous tension produces high blood pressure," are neither confirmed nor refuted by one observation or even by a few observations but only by a series of carefully controlled observations. This case concerns generalizations about observable connections.

There is a further sense of "empirical" that applies to statements about objects which are in principle nonobservable, such as "ideal gases" or "perfectly elastic bodies." Such entities cannot be observed because they do not exist, and hence we cannot frame empirical statements about them in either the particular or the general sense of "empirical." These entities, which cannot be described in observation statements, are instead defined in a series of

statements constituting a scientific theory. From such a theory statements can be deduced that can be tested by reference to direct experience. If observation shows the deduced statements to be false, we must reject the theory, and hence our theoretical statements about the unobservable entities are indirectly refuted by observation. These statements are therefore, in this indirect way, empirical because observations can serve to show them to be false.

PHENOMENOLOGICAL STATEMENTS. Phenomenological statements, as described in the preceding sections, are not empirical in the first two senses of the term. They are not empirical in the first sense because they are never statements about individual existing intentional acts but only about the criteria governing types of acts; only particular statements are empirical in the first sense. Empirical in the second sense are generalizations derived by induction from a series of observations of particulars. Such inductive generalizations presuppose that we know what particulars belong to the class of objects to be observed. If we want to make a generalization about the relation between nervous tension and high blood pressure, we must have a very precise idea of what must count as examples of nervous tension and what blood pressure counts as "high" blood pressure. Similarly, we cannot inductively arrive at statements about intentional acts unless we are already able to differentiate a coherent intentional act from an incoherent collection of each of the four kinds of elements of intentional acts.

The same applies to generalizations about coherent series of intentional acts. Nothing said by the phenomenologists should exclude the possibility of framing empirical (in the general sense) statements about intentional acts. All that is argued is that the criteria of coherence of individual acts as well as of series of acts are presupposed and therefore are not established by such inductive generalizations. Therefore, statements formulating these criteria cannot themselves be empirical generalizations.

It is undoubtedly a task for phenomenology to differentiate the different senses of "empirical," that is, to describe the different kinds of intentional acts involved in what we call experience and the criteria of coherence belonging to each kind of act. Oddly enough, the phenomenologists so far have barely begun to undertake such an examination, and hence their conviction that statements about phenomena, as now defined, are nonempirical is not supported by adequate phenomenological analyses. This important shortcoming in the theory of the phenomenological method is all the more serious because there are good reasons for thinking that there is one perfectly good sense of the words *experience* and *empirical* in which statements about phenomena, as defined, are empirical.

EMPIRICAL PHENOMENA STATEMENTS. In a scientific theory, the terms are defined in relation to one another in such a way that if we alter the definition of one term, the definitions of some of the other terms are also changed. The effect of such a set of interrelated definitions is to limit the contexts in which these terms may be applied. A set of phenomenological statements has a similar function; it limits the contexts in which given intentional acts may be performed. The limits imposed on these intentional acts in the phenomenological statements are interrelated as the definitions in a theory are. If we alter the limits of one intentional act, those of other acts are also changed. History and ethnology provide many examples of such changes.

Among the Trobriand Islanders, for instance, successful gardening requires the use of magic. Before seedlings are planted, a spell must be spoken over them. It is very important that the magician's mouth be as close as possible to the seedlings, for otherwise some of the power of the spell will be dissipated. The power of the spell resides not in the sound waves produced by the magician but in the meaning of the terms used, something that we would not regard as a physical phenomenon. Yet the power of the magical words is here treated as if it were a physical force that varies with the distance from the object it affects. It is clear that the Trobriander does not draw a distinction between the physical and the mental, so it makes perfect sense for him to say something that makes no sense to us—that the spell must be spoken as close as possible to the seedlings in order to be effective. He imposes different limits on his intentional acts—what makes literal sense to him is to us at best symptomatic of the confusions of the "primitive" mind—and these various limits are interrelated. We can formulate them in a set of phenomenological statements that we regard as false and he regards as true. This example shows the analogy between the limitations imposed on theoretical terms by their implicit definitions in a scientific theory and the mutual limitations imposed on intentional acts and expressed in phenomenological statements.

Statements in a scientific theory limit the application of the terms. If the limits imposed allow the use of the terms in false factual statements, these limits must be

altered; the theory is invalid. In analogous ways phenomenological statements may be invalidated by experience. Phenomenological statements express the limits imposed on intentional acts, and if these limits are such that we cannot distinguish true factual statements from false ones, the limits must be altered; the phenomenological statements are invalid.

In order to make true generalizations about gardening and distinguish them from false ones, we need a clear notion of causation. Causal relations as discussed in science exist only between spatiotemporally contiguous events, and this implies that only spatiotemporally located events can be either causes or effects. A clear notion of causal relations, therefore, presupposes a clear distinction between events that are and those that are not spatiotemporally located, or between physical and mental events. Where such a distinction is not drawn, no clear understanding of causal relations is possible. The Trobriander does not differentiate physical events from mental events (and forces); hence he cannot clearly differentiate causal relations from noncausal relations. As a result, he cannot make general statements about gardening that are always true or always false as tested by the information available to us. They may, of course, be always true (or false) tested by what he knows. His generalizations are about classes containing very heterogeneous types of relations, both causal and noncausal. Statements about the causal are true under very different conditions from statements about the noncausal, so his generalizations are sometimes true and sometimes false, and he does not have the vocabulary necessary to reformulate them in such a way that they are always true or always false. This shows that the Trobriander's lack of scientific information about biology is not accidental. It is impossible for him to do natural science because his language lacks the requisite distinctions. Scientific statements cannot be made in his language, which is clear proof that it is inadequate and that the phenomenological statements describing his linguistic acts as well as the nonlinguistic ones, such as those associated with garden magic, are therefore invalid.

This argument as stated is not conclusive, but it can be strengthened to make a rather formidable case for holding that the phenomenologists are mistaken in their claim that their statements about phenomena are nonempirical in all senses of that term. This conclusion shows that the question asked at the very outset—what are the truth conditions of phenomenological statements—remains unanswered. In the preceding a good deal has been said about these truth conditions, but it has been shown that that answer is incomplete. The phenomenologists' account of their method not only lacks a complete theory of experience in its different forms but also a complete theory of truth, at least as that term applies to the statements in phenomenology.

THE SENSES OF "NECESSARY." The second aspect of a priori statements is their necessity. A priori statements are necessary because they are nonempirical; if they are true at all, they are true independently of facts about the world. Even if all the statements about this world that are now true were false, and if, therefore, our world were very different from what it is now the a priori statements would still be true. They are true whatever happens to be the case in the world. Hence we may say that, if true at all, they must be true regardless of any facts. For this reason the term *necessary* has often been explicated as "true for all possible worlds." A different world from ours is one whose description requires factual statements to be true that are false of our world. Since necessary statements are true whatever factual statements may or may not be true, they are true for all possible worlds. A statement is necessary, therefore, to the extent that its truth is logically independent of the truth or falsity of empirical statements. It follows that there are different senses of "necessity" to correspond to the different senses of "empirical." There are, therefore, also different senses of "a priori." Hence, phenomenological statements are clearly a priori insofar as they are not empirical in the first two senses of that term. But phenomenological statements are empirical in a third sense and are therefore not a priori in that sense of "a priori" that contrasts with this third sense of "empirical."

NECESSARY PHENOMENA STATEMENTS. In the sense explained, statements are necessary if they are *true* necessarily. But if statements about phenomena are a priori—necessarily true *and* nonempirical—they are necessary in a second sense: Their truth is a necessary condition for any empirical statement to be capable of being either true or false. An empirical statement can be either true or false only if it is meaningful, and that depends on the coherence of the intentional act and of the intentional act series in which it is asserted. But as was seen, the coherence of such acts and act series is presupposed by any question about the adequacy of intentional acts to their intentional objects. Hence the statements that lay down the criteria for coherence of all kinds of intentional acts, including acts of asserting, must be true if we are to be able to decide whether any given intentional act is adequate to its intentional object—for instance, whether an assertion is

true or a purposive action is successful. Insofar as phenomenological statements are a priori, they are, therefore, necessary in this second sense; they are presuppositions for the adequacy or inadequacy of any intentional act to its intentional object. The truth of phenomenological statements is logically prior to the truth or falsity of all empirical statements and to the correctness of all purposive actions.

CONTEMPORARY PHENOMENOLOGY

Political events in Europe and the shifting winds of doctrine caused the phenomenological movement to lose much of its original momentum after Husserl's death in 1938. The best-known twentieth-century philosophers who used the term *phenomenology* in descriptions of their own work were Martin Heidegger in Germany and Jean-Paul Sartre and Maurice Merleau-Ponty in France. All three used the term *phenomenology* in appreciably different senses from the phenomenologists previously discussed.

HEIDEGGER. Heidegger was a student of Husserl's and at one time was a coeditor of the *Jahrbuch*. In that journal (Vol. 8, 1927) appeared his first major work, *Sein und Zeit* (translated by J. Macquarrie and E. Robinson as *Being and Time*, New York, 1962). The phenomenologists so far discussed all agreed that it is the task of phenomenology reflectively to bring to light the criteria implicit in the intentional acts we perform in everyday life, in which we act in, get to know about, and learn to master that everyday world which Husserl christened the *Lebenswelt* ("world in which we live"). The emphasis here is on putting into words what is commonly and familiarly done without one's knowing how to describe accurately what he is doing. Heidegger also regarded phenomenology as a sort of reflection but not a reflection designed to put into words what is familiar in performance.

On the contrary, Heidegger's brand of phenomenology tried to open the way back to what had, he thought, become completely unfamiliar, what he calls *Sein* (being). He recognized that "being" had become a philosophically empty word. Hence we cannot gain a better understanding of being by reflecting only about the world insofar as it is familiar to us, for in that world "being" has become almost meaningless; there are very few contexts in which it makes sense to talk about "being." Thus, reflection about the criteria of intelligibility, which we use now, will not reveal much about being. Rather than reflect on these criteria, Heidegger proposes to ask why "being" has become almost meaningless to us. But since a question is intelligible only to the extent that we can specify the sort of answer we expect, and since an answer to Heidegger's question would require a language in which "being" is meaningful, even an intelligible formulation of his question involves him in the attempt to re-create a very different language, in which "being," far from being an empty word, is the richest and most important concept. This language, he believed, is the language used by the pre-Socratic philosophers. Heidegger's phenomenology thus led him into an enterprise utterly unfamiliar to the other phenomenologists, the attempt to develop a new philosophical language by re-creating that of the pre-Socratic philosophers.

SARTRE. Sartre's major work, *L'être et le néant* (Paris, 1943; translated by H. E. Barnes as *Being and Nothingness*, New York, 1957), bears the subtitle *An Essay in Phenomenological Ontology*. The work does not, however, contain any explicit discussion of phenomenology, nor did Sartre explain his conception of phenomenology at length in any other work. More than once he differentiated phenomenology from science by saying that phenomenology makes statements about essences; science, about facts. In one long essay, "La transcendence de l'égo" (*Recherches Philosophiques*, Vol. 6, 1936–1937; translated by F. Williams and R. Kirkpatrick and published in book form as *The Transcendence of the Ego*, New York, 1958), he takes sharp issue with Husserl's transcendental phenomenology, particularly with the claim that in phenomenology we discover that there is a transcendental ego.

It would seem, then, that Sartre was a phenomenologist who, like many others, adopts the descriptive approach to essences but refuses to follow Husserl in his later developments of the theory of the phenomenological method. But Sartre differs radically insofar as he was not averse to constructing philosophical theories. His major work is an example of constructive philosophy in precisely that sense in which phenomenologists attacked it in their polemic against reductionism. Sartre's conception of phenomenology is no clearer if we look at his actual practice of the method than if we consider his sparse statements about it. If Sartre practiced phenomenology at all, the term as used by him and as applicable to his procedures has a different meaning from the one explicated in this discussion.

MERLEAU-PONTY. Merleau-Ponty's major work bears the title *Phénoménologie de la perception* (Paris, 1945; translated by Colin Smith as *Phenomenology of Perception*, London, 1962). Unlike Sartre, he includes an introduction devoted to a clarification of "phenomenology."

The clear and explicit result of this discussion is that Merleau-Ponty has interpreted the notion of phenomenology in a sense rather different from that subscribed to wholly or partly by members of the phenomenological movement, as well as from that used by either Heidegger or Sartre.

These three philosophers used "phenomenology" in appreciably different ways from those in which it has been used by the phenomenologists discussed. To be sure, there were also radical and profound disagreements among the latter about the nature and presuppositions of the phenomenological method, but they regarded these differences as different results arrived at by applying the same method. In this sense these philosophers—Husserl, Pfänder, Geiger, Becker, and Reinach, among others—can be regarded as belonging to one school of philosophy. All of them shared certain common views at the outset, and they believed that they were using the same method. But Heidegger, Sartre, and Merleau-Ponty began doing their respective brands of phenomenology by explaining what they considered phenomenology to be and how their conception differed from that of Husserl. They did not begin with the same common views, as did the earlier phenomenologists; and they did not regard their method as identical with that of Husserl and the other phenomenologists. For this reason they do not belong to the same school of philosophy.

See also Binswanger, Ludwig; Brentano, Franz; Existentialism; Existential Psychoanalysis; Heidegger, Martin; Intentionality; Life, Meaning and Value of; Psychologism; Sartre, Jean-Paul; Scheler, Max; Time, Consciousness of.

Bibliography

ORIGINAL WORKS IN GERMAN

Husserl, Edmund. *Husserliana.* Vol. I: *Cartesianische Meditationen*. The Hague: Nijhoff, 1950. The shortest, though not always easy, introduction to Husserl's mature conception of phenomenology.

Husserl, Edmund. *Husserliana.* Vol. VI: *Die Krisis der europaischen Wissenschaften und die transcendentale Phänomenologie*. The Hague: Nijhoff, 1954. A very late work of Husserl's; introduces the important notion of the *Lebenswelt*, the world of everyday life.

Husserl, Edmund. *Husserliana.* Vols. VII and VIII: *Erste Philosophie*. The Hague: Nijhoff, 1956–1959. Husserl's lectures on the history of philosophy and phenomenology; more accessible than many of his other writings.

Husserl, Edmund et al., eds. *Jahrbuch für Philosophie und phänomenologische Forschung.* 11 vols. Halle: Niemeyer, 1913–1930. Contains representative writings of all the major phenomenologists.

Reinach, Adolf. "Über Phänomenologie." In his *Gesammelte Schriften*, edited by Hedwig Conrad-Martius. Munich, 1921. A brief and very lucid statement of an early conception of phenomenology.

ORIGINAL WORKS IN TRANSLATION

Husserl, Edmund. *Cartesian Meditations*. Translated by Dorion Cairns. The Hague: Nijhoff, 1960.

Husserl, Edmund. *The Idea of Phenomenology*. Translated by William P. Alston and George Nakhnikian. The Hague: Nijhoff, 1964. Translates a series of lectures given in 1907 in which the transcendental-phenomenological reduction is introduced.

Husserl, Edmund. *Ideas*. Translated by W. R. Boyce Gibson. New York: Macmillan, 1931; New York: Collier, 1963.

Husserl, Edmund. "Phenomenology." Translated by C. V. Solomon. In *Encyclopaedia Britannica*, 14th ed., Vol. XVII, 699–702. 1927.

Husserl, Edmund. *Phenomenology and the Crisis of Philosophy: Philosophy as Science and Philosophy and the Crisis of European Man*. Translated with notes and introduction by Quentin Lauer. New York: Harper, 1965. Contains a long essay, published in 1910, in which Husserl provides an expanded version of his earlier polemics against psychologism. This book also translates portions of Husserl's late work, *Krisis*.

BOOKS ON PHENOMENOLOGY

Bachelard, Suzanne. *La logique de Husserl*. Paris, 1951. A lucid and detailed discussion of Husserl's *Formale und Transcendentale Logik* (1929).

Berger, Gaston. *Le cogito dans la philosophie de Husserl*. Paris: Aubier, 1941. One of the earliest and still one of the best monographs on one aspect of Husserl's thought.

Brand, Gerd. *Welt, Ich und Zeit*. The Hague: Nijhoff, 1955. Attempts a summary of Husserlian phenomenology on the basis of unpublished manuscripts. Rather general but often illuminating.

Brentano, F. *Psychology from an Empirical Standpoint*. Translated by A. C. Rancurello, D. B. Terrel, and L. L. McAlister; edited by L. L. McAlister. London: Routledge, 1969.

Cooper, D. E. *Heidegger*. London: Claridge Press, 1996.

Derrida, Jacques. *Speech and Phenomena*. Translated by Gayatri Chakravorty Spivak. Baltimore: Johns Hopkins Press, 1976.

Dreyfus, H. L. *Being-in-the-World: A Commentary on Heidegger's* Being and Time, *Division I*. Cambridge, MA: MIT Press, 1991.

Farber, Marvin, ed. *Philosophical Essays in Memory of Edmund Husserl*. Cambridge, MA: Harvard University Press for University of Buffalo, 1940. A collection of stimulating and sometimes informative essays on different aspects of Husserl's work.

Heidegger, M. *Being and Time*. Translated by J. Macquarrie and E. Robinson. London: SCM Press, 1962.

Merleau-Ponty, M. *Phenomenology of Perception*. Translated by C. Smith. London: Routledge and Kegan Paul, 1981.

Mohanty, J. N. *Edmund Husserl's Theory of Meaning*. The Hague: Nijhoff, 1964. An interesting discussion of Husserl's early work in the light of the treatment given the same questions by Bertrand Russell, G. E. Moore, and others.

Mohanty, J. N. *Transcendental Phenomenology: An Analytic Account*. Oxford: Blackwell, 1989.

Moran, D. *Introduction to Phenomenology*. London: Routledge, 2000.

Schrag, C. O. *Experience and Being*. Evanston, IL: Northwestern University Press, 1969.

Smith, B., and D. W. Smith, eds. *The Cambridge Companion to Husserl*. Cambridge, U.K.: Cambridge University Press, 1995.

Smith, D. W. *Husserl and Intentionality: A Study of Mind, Meaning, and Language*. Dordrecht: Reidel, 1982.

Sokolowski, R. *Introduction to Phenomenology*. Cambridge, U.K.: Cambridge University Press, 2000.

Spiegelberg, Herbert. *The Phenomenological Movement.* 2 vols. The Hague: Nijhoff, 1960; 3rd ed., The Hague: Kluwer Academic, 1982. Discusses phenomenology in general as well as individual contributors to the movement. Extremely informative.

ARTICLES ON PHENOMENOLOGY

Ayer, A. J., and Charles Taylor. "Phenomenology and Linguistic Analysis." *PAS*, Supp. 33 (1959).

Boehm, Rudolf. "Basic Reflection on Husserl's Phenomenological Reduction." *International Philosophical Quarterly* 5 (1965): 183–202.

Downes, Chauncey. "On Husserl's Approach to Necessary Truth." *Monist* 19 (1965): 87–106.

Findlay, J. N. "Phenomenology." In *Encyclopaedia Britannica*, 1965 ed.

Ryle, Gilbert. "Phenomenology." *PAS*, Supp. 11 (1932).

Smith, D. W. "Phenomenology." Stanford Encyclopedia of Philosophy. Available from http://www.plato.stanford.com.

Spiegelberg, Herbert. "Toward a Phenomenology of Experience." *American Philosophical Quarterly* 1 (1964): 1–8.

Richard Schmitt (1967)

Bibliography updated by Benjamin Fiedor (2005)

PHENOMENOLOGY [ADDENDUM]

The development of "phenomenology" is a consequence of the interpretation of the texts of the major figures, especially Edmund Husserl, and of independent phenomenological research. Quite often, the two projects have gone hand in hand. One major factor in the development of phenomenology during the period under review has been the ongoing publication of the Nachlass of the major figures (*Husserliana,* Martin Heidegger's *Gesamtausgabe,* as well as Maurice Merleau-Ponty's lectures). Another is the continuing conversation with analytic philosophy in the English-speaking countries, with structuralism and deconstructionism in France, and with hermeneutics, critical theory, and the tradition of German idealism in Germany.

One major starting point in the conversation with analytic philosophy has been Dagfinn Føllesdal's (1969) paper, which argues that Husserl's concept of Noema is a generalization of the Fregean notion of *Sinn.* Both the Sinn and the Noema are abstract entities, to be distinguished from the object toward which an intentional act may be directed. While the historical claim underlying this thesis—namely, that Gottlob Frege's was a major influence on the development of Husserl's thinking around the turn of the twentieth century—has been challenged (e.g., by Jitendra N. Mohanty), the systematic thesis of Føllesdal (as opposed to Aron Gurwitsch's thesis, that the Noema is the perceived object qua perceived and the object intended is but a system of noemata), has been influential.

Jaakko Hintikka developed another aspect of Husserl's theory of intentionality by construing the Noema as a function from possible worlds to individuals in those worlds. The resources of the semantics of Frege and of possible worlds have been pulled together to interpret Husserl in the work of David Smith and Ronald McIntyre. Mohanty and Frederick Seebohm have cautioned against reducing the intentional thesis of Husserl to an extensional thesis of possible worlds and have emphasized the need for a theory of constitution of possible worlds, if the latter are not to be posited in a naively ontological thesis. Still others, notably R. Sokolowski and Daniel Bell, have questioned the validity of ascribing to Husserl a Fregean-type theory. Sokolowski takes the Husserlian Noema to be identical with the object (with the proviso "as intended"), and Bell reads Husserlian Gegenstand to be a component of the intentional act and so quite unlike the Fregean reference.

From another perspective, John Searle has found the Husserlian intentionality thesis useful for his own work but goes beyond Husserl by appropriating, from Heidegger via Hubert Dreyfus, the idea of Background of skills and practices, and more recently by developing a theory of we-intentionality that is irreducible to I-intentionalities (reminiscent of the Hegelian *Geist* as well as of a thesis advanced by David Carr). This last discussion connects with the way phenomenology has related itself to cognitive science. Jerry Fodor's methodological solipsism has been related by Dreyfus to Husserl's, while Searle's emphasis on Background clearly falls on the Heideggerean side of the divide.

The tension between Husserlian phenomenology and hermeneutics lies in that the former is concerned with consciousness, its contents and structures, the latter with the individual's ontological relatedness to his world

and to others. This issue becomes, Is interpretation to be construed as the gift of a transcendental ego, or is it to be construed as an ontological feature of the mode of being of *Dasein*? Hans-Georg Gadamer's theory of interpretation develops the latter alternative, while Paul Ricoeur comes closest to mediating between Husserlian thinking, especially of the *Logical Investigations,* and an ontologically construed hermeneutics. We must also recall Ricoeur's work on metaphor, in which, going beyond the traditional rhetorical and semantic theories of analytic philosophy, Ricoeur integrates them in such a manner as permits the poetic and disclosive dimension of language to emerge. Ricoeur's researches have also sought to mediate between time (the most radical subjectivity) and narrative (by which reality is redescribed, as by metaphors) and reestablish a certain reciprocity between them.

The most influential critique of classical phenomenology is offered by Jacques Derrida. While it is more common to look upon Derrida's work as refuting Husserl's transcendental phenomenology, it is also possible to maintain that Derrida's work is a further radicalization of Husserl's genetic phenomenology, an alleged result of which is the demonstration that constitution involves a perpetual deferral and difference, also that a radicalization of Husserl's concept of horizonal character of intentionality would call into question all fixity and univocity of meanings, and that possibilities of nonfulfillment of intention are necessarily inherent in all intentionality. But those who ascribe to Husserl a metaphysics of presence fall into the opposite trap of reifying "absence." As Sokolowski has shown, Husserl's thinking rather exhibits a mutual involvement of presence and absence.

Of those from analytic philosophy who have pursued some kind of phenomenology, mention must be made of Castañeda's rich phenomenology of indexical reference and of "I" thought. In the latter context, he distinguishes between the ground floor of empirical I-guise and successive phases of transcendental I-guises, among all of which there is a sameness that is yet not strict identity.

In the United States there is a continuing tendency, inaugurated by Dreyfus and Richard Rorty, to see in Heidegger a pragmatist philosopher, whereby clearly Heidegger's ontological concern with the meaning of being and the historical concern with the historicity of understanding of Being are either underplayed or sought to be altogether set aside. While it was at first usual to look upon Heidegger as an antiscience thinker, now—largely owing to the work of C. F. von Weizsäcker, J. Kockelmans, and Patrick Heelan—one has come to realize that Heidegger's thinking could form the basis for an understanding and appreciation of science and technology. In general, phenomenological thinking about science has exhibited three distinct features: First, following Husserl in the *Crisis,* some have attempted to reestablish the proper connection between science and lifeworld. The most important work on this front is due to J. Mittelstrass. Second, following also Husserl's work in the *Crisis,* but more inspired by Heideggerean thought about historicity of *Dasein* as also by Thomas Kuhn's work on history of science, some have looked upon science as a historical accomplishment marked by epochal changes, epistemological breaks, shifts of paradigm—thereby rejecting the prevailing obsession with the logical structure of scientific theories and also the reigning prejudice in favor of a naively realistic and positivist theory of science. But within phenomenology itself, this time following Husserl's original concern, there is also a continuing concern with the nature and structure of logic and mathematics as theories and with the origin of such theories, their relation to practice and also to the lifeworld, on the one hand, and the transcendental, thinking ego on the other.

Heelan has developed the view, using the conceptual resources of Husserl, Heidegger, and Merleau-Ponty, that scientific observation, like all perception, is hermeneutical. Hermeneutical phenomenology of science focuses, in his view, not so much on theory as on experimental phenomena. Heelan defends a sort of realism called by him hermeneutic or horizonal realism as opposed to the instrumentalism of some phenomenologists. Thus, according to Heelan, in particle physics many phenomena have actual existence only within the context of the measurement processes. Kockelmans emphasizes what he regards as the ontological aspect of science: he draws attention to the role of "objectifying thematization," which lies at the root of every scientific activity. In this latter concept he brings together Husserl's idea of "thematization" and Heidegger's idea that a certain fundamental understanding of being makes possible science, philosophy, and technology. Although Kockelmans accepts the Kuhnian thesis of epochs in the history of science, he nevertheless holds that history of science is guided by an ideal of reason and that each new paradigm is necessarily a historical synthesis.

From its inception phenomenology had a special relation of love and hate toward psychology; at a later phase, it developed a special interest in history. With regard to psychology, there has been a long tradition of original work in what is known as phenomenological psychology. To the period under review belong some

works of Medard Boss, Aron Gurwitsch, Hermann Minkowski, and Ricoeur. Boss has applied his Heideggerean conception of *Daseinsanalytik* to such contexts as sexual perversion, dream, and psychosomatic illness. Drawing upon his work on lived space and lived time, Minkowski studies how these can undergo modifications in psychoses, schizophrenia, manic-depression and hallucinations. Gurwitsch's *Marginal Consciousness,* posthumously published, continues the work done in *The Field of Consciousness.* However, for research in descriptive psychology, possibly the most important results are to be found in Edward Casey's two books on imagining and remembering. This research has opened out new fields of investigation. For example, in his work on remembering, Casey explores a number of neglected, nonrepresentational forms of remembering, including body memory and place memory, reminiscing and commemorating.

In the phenomenology of history, a brief reference may be made to the important work done by Ricoeur, who seeks to mediate between lived time and cosmic time. The past is irrevocably gone, and our access to it across the historical distance is made possible by creative imagination. Here fiction, by its quasi-historical character, comes to our help. History is not a totality, an absolute mediation. Nevertheless, there is a search for meaning, which is open-ended without a Hegelian *Aufhebung.* The idea of one history is a Kantian-type regulative idea.

See also Being; Cognitive Science; Critical Theory; Deconstruction; Derrida, Jacques; Frege, Gottlob; Gadamer, Hans-Georg; Heidegger, Martin; Hermeneutics; Husserl, Edmund; Indexicals; Intentionality; Kuhn, Thomas; Merleau-Ponty, Maurice; Modality, Philosophy and Metaphysics of; Phenomenological Psychology; Phenomenology; Ricoeur, Paul; Solipsism.

Bibliography

Bell, D. *Husserl.* London: Routledge, 1990.

Casey, E. S. *Imagining: A Phenomenological Study.* Bloomington: Indiana University Press, 1976.

Casey, E. S. *Remembering: A Phenomenological Study.* Bloomington: Indiana University Press, 1987.

Castañeda, H.-N. *Thinking, Language, and Experience.* Minneapolis: University of Minnesota Press, 1989.

Derrida, J. *Speech and Phenomena.* Translated by D. Allison. Evanston, IL: Northwestern University Press, 1973.

Dreyfus, H., ed. *Husserl, Intentionality, and Cognitive Science.* Cambridge, MA: MIT Press, 1982.

Føllesdal, D. "Husserl's Notion of Noema." *Journal of Philosophy* 66 (1969): 680–687.

Gurwitsch, A. *Marginal Consciousness,* edited by L. Embree. Athens: Ohio University Press, 1985.

Gurwitsch, A. *Studies in Phenomenology and Psychology.* Evanston, IL: Northwestern University Press, 1966.

Heelan, P. *Space Perception and the Philosophy of Science.* Berkeley: University of California Press, 1983.

Hintikka, J. *The Intentions of Intentionality and Other New Models for Modalities.* Dordrecht: Reidel, 1975.

Kockelmans, J. *Ideas for a Hermeneutic Phenomenology of the Natural Sciences.* Dordrecht: Kluwer Academic, 1993.

Kockelmans, J., ed. *Phenomenological Psychology: The Dutch School.* Dordrecht: Nijhoff, 1987.

Mittelstrass, J. *Die Möglichkeit von Wissenschaft.* Frankfurt am Main: Suhrkamp, 1974.

Mohanty, J. N. *Husserl and Frege.* Bloomington: Indiana University Press, 1982.

Mohanty, J. N. *The Possibility of Transcendental Philosophy.* Dordrecht: Nijhoff, 1985.

Ricoeur, P. *Time and Narrative.* Translated by K. Blarney and D. Pellauer. Chicago, 1988.

Searle, J. *Intentionality.* New York: Cambridge University Press, 1983.

Smith, D., and R. McIntyre. *Husserl and Intentionality.* Dordrecht: Reidel, 1982.

Sokolowski, R. "Husserl and Frege." *Journal of Philosophy* 84 (1987): 523–528.

von Weizsäcker, C. F. *Zeit und Wissen.* Munich: Hanser, 1992.

Jitendra N. Mohanty (1996)

PHILODEMUS

(c. 110–c. 40 BCE)

Philodemus of Gadara was an Epicurean philosopher and epigrammatic poet of the first century BCE. Born in Gadara in Palestine, he was taught philosophy in Athens by the head of the Epicurean school Zeno of Sidon (c.150–70s BCE) and by Demetrius Lacon, Zeno's younger contemporary. In the 80s or 70s he moved to Italy and earned the patronage of Lucius Calpurnius Piso Caesoninus, father-in-law of Julius Caesar. He seems to have spent part of his life at Herculaneum in Campania, probably in Piso's villa, and to have formed around him an Epicurean community of pupils and friends. His writings constitute the largest surviving portion of the library of the villa, which was buried beneath the mud and ashes when Vesuvius erupted in 79 CE and was partly excavated in the mid-eighteenth century. Thirty-seven distinct works are known or conjectured to be his and are contained in carbonized papyrus rolls in various states of fragmentation and corruption. However, it has been possible to gain considerable knowledge of Philodemus's methods and views. He emerges as a prolific writer with a wide range of interests who advances a conception of Epicurean orthodoxy first

defined by Zeno and exhibits the intellectualism characteristic of the school of Athens. He was respected in educated Roman circles: Cicero speaks well of him, and he seems acquainted with both Vergil and Horace.

About thirty of his poems are preserved in the *Palatine Anthology*, and additional evidence suggests that he may have written hundreds. It is controversial whether there are relations between Philodemus's poetic output, his poetic theory, and his philosophical commitments. Except for a poetic invitation to Piso to participate in a festival in Epicurus's honor, his elegant epigrams make no mention of Epicureanism. Most of them concern love, and several contain autobiographical elements. They can be read as illustrating Philodemus's thesis that poetry as such does not benefit but only entertains.

In *On Poems*, he develops and defends his views, arguing both against literary theorists who held that a good poem be morally useful (Heraclides of Pontus, Neoptolemus of Parium, and an unnamed Stoic philosopher) and against formalists (notably, Crates of Mallos) who judged a poem only by reference to its form and aesthetic quality. He considers poetry an imitative art appreciated by reason, which requires careful composition in order to present clearly certain thoughts and move the listener. What makes a poem good is appropriate thoughts expressed in appropriate diction; changing the arrangement of words can destroy the poetic goodness of a verse. However, *On the Good King According to Homer* shows how to derive benefit from the poetry of Homer, especially how to extract both warning and advice from Homer's portrayal of different rulers. *On Music*, too, dissociates moral profit from artistic form. Music as such has no mimetic character. It is sound, an irrational element that causes pleasure to the ear. It affects the soul only via poetry, texts, or thought, which, however, are external to the musical art. *On Rhetoric* suggests a comparable approach to sophistic or epideictic rhetoric. Refuting Epicurean rivals who deny that rhetoric is an art, Philodemus holds that while forensic and political rhetoric are not arts, epideictic or sophistic rhetoric is. It consists mainly in the transmissible method of using the one naturally correct language to write clear and persuasive compositions, and the criteria pertaining to it are independent of its utility.

Philodemus gained credit for his historical work as well. The *Arrangement of the Philosophers*, especially the two books on the Academics and the Stoics, contains biographical and doxographical material and, occasionally, summaries of philosophical doctrines. The *Works on the Records of Epicurus and Some Others* relates the early history of the Epicurean school whereas the treatise *On Epicurus* eulogizes the founder and alludes to rituals of the Epicurean communities. The polemical treatises *On the Stoics* and *Against the …* should also be mentioned. Historical information about the theological doctrines of philosophers from the Presocratics to the Stoic Diogenes of Babylon is found in Philodemus's theological work *On Piety*, which offers a powerful defense of Epicurus's piety and reflects Zeno of Sidon's interpretation of Epicurus's views about the nature of the gods and our concepts and knowledge of them. *On the Gods* discusses our fear of the gods whereas *On the Way of Life of the Gods* treats aspects and attributes of divine existence. Both in theology and in other areas, Philodemus endorses the epistemological positions of his school, some of which may have been mentioned in a work on perception. *On Signs* confirms that he is also committed to the Epicurean methodology developed and defended against Stoic criticisms by Philodemus's teachers in Athens—in particular, the similarity method (a method of sign-inference based on analogy and induction) and the related procedure of comparative assessment (*epilogismos*). Two other works, one of which is subtitled *From the Lectures of Zeno*, make remarks about scientific methodology.

Philodemus engages in both theoretical and practical ethics, often in connection with moral psychology. *On Choices and Avoidances* rehearses canonical theses such as the cardinal principles of Epicurus's doctrine, the criteria of moral choice, the so-called fourfold medicine (*tetrapharmakos*), and the relation between the virtues and pleasure. *On Frank Speech* is the central piece of the ensemble *On Characters and Ways of Life*, to which *On Gratitude* and *On Conversation* also belong. It discusses frank speech (*parrhesia*), the principal educational method of late Epicurean schools and a major tool of moral and psychological therapy, and it reflects the views of Zeno on whose lectures the treatise is based.

Another major work is *On Vices and the Opposite Virtues and the People in whom they occur and the Situations in which they are found*. There survive the extant remains of three books that analyze and treat, respectively, the vices of flattery, arrogance, and greed, as well as other vices of professional administrators and money makers. The fragmentary contents of *On Wealth* are thematically related to this last topic.

The books *On Folly*, *On Lack of Proper Measure*, *On Erotic Love*, and, possibly, *On Envy* belong to the multivolume project *On the Passions*. We know very little about them whereas a good deal survives of *On Anger*, which describes the nature and consequences of anger and

draws a distinction between violent rage and natural anger. Philodemus condemns the former but allows room for the latter, steering what might seem a middle course between the Peripatetic approval of rightful anger and the Stoic aim of eradicating the emotion altogether. *On Death* is conceptually related to the group *On the Passions* and may have belonged there. The surviving text addresses the question of whether the moment of death is always *physically* painful, and also examines cases in which death may cause great *emotional* pain, such as dying prematurely, ingloriously, or unjustly and leaving behind grieving friends. Philodemus's analyses and arguments, and his concession that it is sometimes natural to feel *bites* of sorrow, constitute significant contributions to moral psychology. Moreover, his methods of treating the emotions occupy an important place in the therapeutics of the Hellenistic era.

See also Cicero, Marcus Tullius; Epicureanism and the Epicurean School; Epistemology; Ethics; Hellenistic Thought; Peripatetics; Stoicism.

Bibliography

Armstrong, David, Jeffrey Fish, Patricia A. Johnston, and Marilyn. B. Skinner. *Vergil, Philodemus, and the Augustans*. Austin: University of Texas Press, 2004.

Blank, David. "Philodemus on the Technicity of Rhetoric." In *Philodemus and Poetry*, edited by Dirk Obbink, 178–188. Oxford: Oxford University Press, 1995.

De Lacy, Philip H., and Estella A. De Lacy. *Philodemus: On Methods of Inference*. Naples: Bibliopolis, 1978.

Gigante, Marcello. *Philodemus in Italy. The Books from Herculaneum*. Translated by Dirk Obbink. Ann Arbor: University of Michigan Press, 2002.

Indelli, Giovanni. *Filodemo, L'ira*. Naples: Bibliopolis, 1988.

Indelli, Giovanni, and Voula Tsouna-McKirahan. *Philodemus, On Choices and Avoidances*. Naples: Bibliopolis, 1995.

Janko, Richard. *Philodemus. On Poems Book I*. Oxford: Oxford University Press, 2000.

Konstan, David, Diskin Clay, Clarence E. Glad, et al. *Philodemus. On Frank Criticism. Introduction, Translation, and Notes*. Society of Biblical Literature. Atlanta, GA: Scholars Press, 1998.

Obbink, Dirk, ed. *Philodemus and Poetry*. Oxford: Oxford University Press, 1995.

Obbink, Dirk. *Philodemus, On Piety Part I*. Oxford: Oxford University Press, 1996.

Schofield, Malcolm. "*Epilogismos*: An Appraisal." In *Rationality in Greek Thought*, edited by Michael Frede and Gisela Striker, 221–237. Oxford: Oxford University Press, 1996.

Sider, David. *The Epigrams of Philodemos*. New York: Oxford University Press, 1997.

Sedley, David N. "On Signs." In *Science and Speculation*, edited by Jonathan Barnes. Jacques Brunschwig, Myles Burnyeat, and Malcolm Schofield, 239–272. Cambridge, U.K.: Cambridge University Press, 1982.

Sedley, David N. "Philodemus and the Decentralisation of Philosophy." *Cronache Ercolanesi* 33 (2003): 31–41.

Tsouna, Voula. "Philodemus the Therapy of Vice." *Oxford Studies in Ancient Philosophy* 21 (2001): 233–258.

Voula Tsouna (2005)

PHILO JUDAEUS

(fl. 20 BCE–40 CE)

Philo Judaeus, the Jewish Hellenistic philosopher, was the son of a wealthy and prominent Alexandrian family. Philo was well educated in both Judaism and Greek philosophy. Little is known about the actual events of his life except that in 40 CE the Jewish community of Alexandria sent him as the head of a delegation to Emperor Caligula to seek redress from the wrongs which the Gentile population inflicted upon the Jews. His *Legacy to Gaius* tells the story of this mission. Although he also wrote moral and philosophic treatises on problems then current, the main bulk of his writings are philosophic discourses on certain topics of the Hebrew Scripture. In content they are, on the one hand, an attempt to interpret the scriptural teachings in terms of Greek philosophy and, on the other, an attempt to revise Greek philosophy in the light of those scriptural traditions.

The scriptural teachings with which Philo set out to revise Greek philosophy contained certain definite conceptions of the nature of God and his relation to the world but only vague allusions to the structure and composition of the world. In dealing with the latter, therefore, he felt free to select from the various views of Greek philosophers whichever seemed to him the most reasonable, although occasionally he supported the selection by a scriptural citation. In dealing with the conception of God, however, he approached Greek philosophic views critically, rejecting those that were diametrically opposed to his scriptural traditions and interpreting or modifying those which were plastic enough to lend themselves to remolding.

GOD, PLATONIC IDEAS, CREATION

Of the various conceptions of God in Greek philosophy, Philo found that the most compatible with scriptural teaching was Plato's conception, in the *Timaeus,* of a God who had existed from eternity without a world and then, after he had brought the world into existence, continued to exist as an incorporeal being over and above the corporeal world. But to Plato, in the *Timaeus,* besides the eternal God, there were also eternal ideas. Philo had no

objection to the existence of ideas as such, for he held that there was a scriptural tradition for the existence of ideas. But he could not accept the eternity of the ideas, for, according to his scriptural belief, God alone is eternal. By a method of harmonization that had been used in Judaism in reconciling inconsistencies in Scripture, Philo reconciled the *Timaeus* with the scriptural tradition by endowing the ideas with a twofold stage of existence: First, from eternity they existed as thoughts of God; then, prior to the creation of the world they were created by God as real beings. He may have found support for the need of such a harmonization in the many conflicting statements about the ideas in Plato's dialogues.

The ideas, which in Plato are always spoken of as a mere aggregation, are integrated by Philo into what he terms "an intelligible world," an expression that does not occur in extant Greek philosophic writings before him. Then, following a statement by Aristotle that the "thinking soul" (that is, nous), "is the place of forms" (that is, ideas), Philo places the intelligible world of ideas in a nous, which, under the influence of scriptural vocabulary, he surnamed Logos. Accordingly, he speaks also of the Logos as having the aforementioned two stages of existence.

For the same reason that he could not accept the view that the ideas are eternal, Philo also could not accept the view commonly held by contemporary students of Plato that the preexistent matter out of which, in the *Timaeus,* the world was created was eternal. But as a philosopher he did not like to reject altogether the reputable Platonic conception of a preexistent matter. And so here, too, he solved the difficulty by the method of harmonization. There was indeed a preexistent matter, but that preexistent matter was created. There were thus to him two creations, the creation of the preexistent matter out of nothing and the creation of the world out of that preexistent matter. For this too, it can be shown, he may have found support in certain texts of Plato.

In the *Timaeus,* Plato describes the creation of the world as an act that God "willed" (*ἐβουλήθη*), and similarly the indestructibility of the world is described by him as being due to the "will" (*βούλσις*) of God. Presumably, by will in its application to God, Plato here means the necessary expression of God's nature, so that the creation of the world, and of this particular world of ours, was an act that could not be otherwise; and similarly the indestructibility of the world is something that cannot be otherwise. Philo, however, following the scriptural conception of God as an all-powerful free agent, takes the will by which God created the world to mean that had God willed, he could have either not created the world or created another kind of world. And similarly, if it be his will, he can destroy the world, although, on the basis of a scriptural verse, Philo believed that God would not destroy it.

LAWS OF NATURE, MIRACLES, PROVIDENCE

The scriptural conception of God as an all-powerful free agent is extended by Philo to the governance of the world. Finding scriptural support for the belief in causality and in the existence of certain laws of nature current among Greek philosophers, except the Epicureans, Philo conceived of God's governance of the world as being effected by intermediary causes and by laws of nature which God had implanted in the world at the time of its creation. He even tried his hand at classifying the laws of nature that happen to be mentioned by various Greek philosophers. But in opposition to the Greek philosophers, to whom these laws of nature were inexorable, he maintained that God has the power to infringe upon the laws of his own making and create what are known as miracles. These miracles, however, are not created arbitrarily. They are always created with design and wisdom for the good of deserving individuals or deserving groups of individuals or humankind as a whole, for, to Philo, God governs by direct supervision not only the world as a whole but also the individual human beings within the world.

To express this particular departure of his from the generality of Greek philosophers, Philo gave a new meaning to the Greek term *πρόνοια,* "providence." To those Greek philosophers who made use of this term it meant universal providence, that is, the unalterable operation of the inexorable laws of nature whereby the continuity and uniformity of the various natural processes in the world are preserved. To Philo it means individual providence, that is, the suspension of the laws of nature by the will and wisdom and goodness of God for the sake of human beings whose life or welfare is threatened by the ordinary operation of those laws of nature. With this conception of individual providence, Philo takes up the discussion of the human soul.

SOUL AND WILL

On the whole, Philo's conception of the soul is made up of statements derived from various dialogues of Plato. He distinguishes between irrational souls, which are created together with the bodies of both men and animals, and rational souls, which were created at the creation of the world, prior to the creation of bodies. Of these preexis-

tent rational souls, some remain bodiless but others become invested with bodies. The former are identified by Philo with the angels of Scripture. Having in mind certain passages in Plato where such unbodied souls are identified with the popular Greek religious notions of demons and heroes, but knowing that Plato himself and also Aristotle and the Stoics dismissed these popular notions as mere myths, Philo says that the angels of Moses are what philosophers call demons and heroes, but he warns the reader not to take the existence of angels as mere myths. With regard to the preexistent rational souls that become embodied, he says, following Plato, that they are equal in number to the stars and are to be placed in newly born human beings whose bodies are already endowed with irrational souls. Again following Plato, Philo says that the irrational souls die with the bodies, whereas the rational souls are immortal. But he differs from Plato in his conception of the immortality of the soul. To Plato, the soul is immortal by nature and is also indestructible by nature. To Philo, immortality is a grace with which the soul was endowed by the will and power of God, and consequently it can be destroyed by the will and power of God if it has proved itself unworthy of the grace bestowed upon it.

A similar revision was also introduced by Philo into the Greek philosophic conception of the human will. In Greek philosophy, a distinction is made between voluntary and involuntary acts. But since all the Greek philosophers, except the Epicureans, believed in causality and in the inexorability of the laws of nature, for them the human will, to which they ascribed the so-called voluntary acts, is itself determined by causes and is subject to those inexorable laws of nature which govern the universe, including man, who is part of it. To all of them, except the Epicureans, no human act was free in the sense that it could be otherwise. The term *voluntary* was used by them only as a description of an act which is performed with knowledge and without external compulsion. To all of them, therefore, there was no free will except in the sense of what may be called relative free will. To Philo, however, just as God in his exercise of individual providence may see fit to infringe upon the laws of nature and create miracles, so has he also seen fit to endow man with the miraculous power to infringe upon the laws of his own nature, so that by the mere exercise of his will man may choose to act contrary to all the forces in his nature. This conception of free will is what may be called absolute free will.

KNOWLEDGE

Philo also revised the philosophic conception of human knowledge, including the philosophic conceptions of man's knowledge of God. Human knowledge, like all other events in the world, including human actions, is, according to Philo, under the direct supervision of God. Like all other events in the world, which are to Philo either natural, in the sense that they are operated by God through the laws of nature which he has implanted in the world, or supernatural, in the sense that they are miraculously created by God in infringement upon those laws of nature, so also human knowledge is either natural or supernatural, called by Philo "prophetic," that is, divinely revealed.

Under natural knowledge, Philo deals with all those various types of knowledge from sensation to ratiocination that are dealt with by Greek philosophers, especially Plato and the Stoics. He presents prophetic knowledge as a substitute for that type of knowledge that in Greek philosophy is placed above the various senso-ratiocinative types of knowledge and is described as recollection by Plato, as the primary immediate principles by Aristotle, and as the primary conceptions by the Stoics. Like all miracles, prophetic knowledge is part of God's exercise of his providence over individuals, groups of individuals, or humankind in general. An example of prophetic knowledge due to God's exercise of his providence over individuals is Philo's account of his own experience: Often, in the course of his investigation of certain philosophic problems, after all the ordinary processes of reasoning had failed him, he attained the desired knowledge miraculously by divine inspiration. An example of prophetic knowledge due to God's exercise of his providence over a group of individuals, as well as over humankind in general, is Philo's recounting of the revelation of the law of Moses.

HUMAN KNOWLEDGE OF GOD

Corresponding to the two kinds of human knowledge are two ways by which, according to Philo, man may arrive at a knowledge of God—an indirect ratiocinative way and a direct divinely revealed way. Philo describes the indirect way as the knowledge of the existence of God which the "world teaches" us, and he deals with the various proofs for the existence of God advanced by Greek philosophers. Most acceptable to him is the Platonic form of the cosmological proof in the *Timaeus,* inasmuch as it is based on the premise of a created world. He modifies the Aristotelian form of the cosmological proof so as to establish the existence of a prime mover, not of the motion of the

world but of its existence. He similarly modifies the Stoic proof from the human mind to establish the existence not of a corporeal God immanent in the world but of an incorporeal God above the world.

In his discussion of the direct way of knowing God, however, Philo makes no mention of the Stoic proof of the innateness of the idea of God. His own direct way of knowing God he describes as a "clear vision of the Uncreated One." But as he goes on to explain it, this direct way of knowing God is only another version of the various indirect ways of knowing him and is similarly based upon the contemplation of the world. The difference between the indirect and direct ways is this: In the case of the various indirect ways, both the knowledge of the world and of the existence of God derived therefrom are attained laboriously by the slow process of observation and logical reasoning; in the case of the direct way, both the knowledge of the world and of the existence of God derived therefrom are flashed upon the mind suddenly and simultaneously by divine inspiration.

But the knowledge of God that may be gained by either of these two ways is, according to Philo, only a knowledge of his existence, not a knowledge of his essence; for as Philo maintains, "it is wholly impossible that God according to his essence should be known to any creature." God is thus said by him to be "unnamable" *ἀκατόνομαστος*), "ineffable" *ἄρρητος*), and "incomprehensible" *ἀκατάληπτος*). This distinction between the knowability of God's existence and the unknowability of his essence does not occur in Greek philosophy prior to Philo. In fact, in none of the extant Greek philosophic literature prior to Philo do the terms *unnamable, ineffable,* and *incomprehensible,* in the sense of incomprehensible by the mind, occur as predications of God. Moreover, it can be shown that both Plato and Aristotle held that God was knowable and describable according to his essence. Philo was thus the first to introduce this view into the history of philosophy, and he had arrived at it neither by Scripture alone nor by philosophy alone. He had arrived at it by a combination of the scriptural teaching of the unlikeness of God to anything else and the philosophic teaching that the essence of a thing is known through the definition of the thing in terms of genus and specific difference, which means that the essence of a thing is known only through its likeness to other things in genus and species. Since God is unlike anything else, he is, as Philo says, "the most generic being" (*τὸ γενικώτατον*), that is, the *summum genus,* and hence he cannot be defined and cannot be known.

As a corollary of this conception of the unknowability and ineffability of God, it would have to follow that one could not properly speak of God except in negative terms, that is, in terms which describe his unlikeness to other things. But still Scripture repeatedly uses positive terms as descriptions of God. All such terms, explains Philo, whatever their external grammatical form, whether adjectives or verbs, are to be taken as having the meaning of what Aristotle calls property, and the various terms by which God is described are to be taken as mere verbal variations of the property of God to act, in which he is unlike all other beings. For to act is the unique property of God; the property of all created beings is to suffer action.

THEOCRATIC GOVERNMENT

Philo widened the meaning of the conception of natural law in its application to laws governing human society. To Greek philosophers, with the exception of the Sophists, this application of the conception of natural law (or, as they would say, law in accordance with nature) meant that certain laws enacted by philosophers in accordance with what they described as reason or virtue were also in a limited sense in accordance with nature, that is to say, in the mere sense that they were in accordance with certain impulses, capacities, rational desires which exist in people by nature. The Greek philosophers assumed, however, that no law enacted for the government of humans, even when enacted by philosophers in accordance with reason and virtue, can be regarded as natural law in the sense of its being fully in harmony with the eternal and all-embracing laws of nature by which the world is governed. Philo agrees with the philosophers as to the limited sense in which enacted human law may be regarded as natural law but argues that a law revealed by God, who is the creator of the world (as, to Philo, the law of Moses was), is fully in harmony with the laws of nature, which God himself has implanted in the world for its governance. To Philo, therefore, natural law came to mean a divinely revealed law.

This widened conception of natural law led Philo to answer the question raised by Greek philosophers as to what was the best form of government. To both Plato and Aristotle no form of government based upon fixed law can be the best form of government, and Plato explicitly maintains that the best form of government is that of wise rulers who are truly possessed of science, whether they rule according to law or without law and whether they rule with or without the consent of the governed.

Against this, Philo argues that the best form of government is that based upon fixed law, not indeed upon manmade fixed law, but upon a divinely revealed fixed law. In a state governed by such a divinely revealed law, every individual has his primary allegiance to God and to the law revealed by God. Whatever human authority exists, whether secular, governing the relation of person to person, or religious, governing the relation of humanity to God, that authority is derived from the law and functions only as an instrument of the application of the law and its interpretation. Such a state, whatever its external form of government, is really ruled by God, and Philo came near coining the term *theocracy* as a description of it; the term was actually so coined and used later, by Flavius Josephus. But Philo preferred to describe it by the term *democracy,* which he uses not in its ordinary sense, as a description of a special form of government in contradistinction to that of monarchy and aristocracy, but rather as a description of a special principle of government, namely, the principle of equality before the law, which to him may be adopted and practiced by any form of government.

VIRTUE

In the course of his attempt to analyze the laws of Moses in terms of Greek philosophy, Philo injects himself into the controversy between the Peripatetics and the Stoics over the definition of virtue. Guided by scriptural tradition, he sides with Aristotle in defining virtue as a mean between two vices; hence, in opposition to the Stoics, he maintains that virtue is not the extirpation of all the emotions, that some emotions are good, that there is a difference of degree of importance between various virtues and various vices, and that the generality of human beings are neither completely virtuous nor completely wicked but are in a state which is intermediate between these two extremes and are always subject to improvement. He maintains, however, that by the grace of God some exceptional persons may be born with a thoroughly sinless nature.

Following Plato and Aristotle, both of whom include under the virtue of justice certain other virtues which they consider akin to justice, but guided also by scriptural tradition, Philo includes under justice two virtues that are entirely new and are never mentioned in any of the lists of virtues recorded under the names of Greek philosophers. Thus, on the basis of the scriptural verse (Genesis 15:6) that "Abraham had faith (*ἐπίστευσεν*) in God and it was counted to him for justice (*δικαιοσύνην*)," Philo includes "faith" (*πίστις*), which he takes to mean faith in the revealed teachings of Scripture, as a virtue under what the philosophers call the virtue of justice. Similarly, because the Hebrew term *ṣedakah* in Scripture is translated in the Septuagint both by *δικαιοσύνη,* "justice" (Genesis 18:19) and by *ἐλέημοσύνη,* "mercy," "alms" (Deuteronomy 6:25, 24:13), Philo includes "humanity" (*φιλανθροπία*), in the sense of giving help to those who are in need of it, as a virtue under the philosophic virtue of justice. But on the basis of Scripture only, without any support from philosophy, he describes also "repentance" (*μετάνοια*) as a virtue. In Greek philosophy, repentance is regarded as a weakness rather than as a virtue.

His scripturally based conception of free will as absolute led Philo to give a new meaning to the voluntariness of virtue and the voluntariness of the emotion of desire as used in Greek philosophy. Both Aristotle and the Stoics, using the term *voluntary* in the relative sense of free will, agree that virtue is voluntary, but they disagree as to the voluntariness of the emotions. To Aristotle, all emotions are involuntary, except the emotions of desire and anger, the latter of which by the time of Philo was subsumed under desire; to the Stoics, all emotions are voluntary. Philo, however, using the term *voluntary* in its revised sense of absolute free will, maintains that in this revised sense the term *voluntary* is to be applied, as in Aristotle, to virtue and to the emotion of desire.

Philo similarly gave a new meaning to the philosophic advice that virtue is to be practiced for its own sake. To Plato, Aristotle, and the Stoics, this advice was meant to serve as a principle of guidance to those who, like themselves, did not believe in individual providence and were not impressed by the explanations offered in the popular Greek religious theodicies as to why virtue is not always rewarded and vice not always punished. The reason underlying this advice was that since there is no certainty as to what external goods or evils would follow the practice of either virtue or vice, it is preferable for man to take his chance on the practice of virtue. This reasoning was presumably based on the common human experience that it is easier for one to induce in himself a feeling of happiness in the misery that may follow a life of virtue than it is to induce in himself a feeling of happiness in the misery, and sometimes even in the joy, that may follow a life of vice.

To Philo, however, the advice to practice virtue for its own sake is based upon his belief that providence is individual; that, despite common observation to the contrary, no virtue goes unrewarded; that acts of virtue are of graded merits; and that the reward is always in accordance with the merit of the act. With all this in the back

of his mind, Philo's advice to practice virtue for its own sake (which he expresses in a different context by the statement that man is to serve God out of love and not out of expectation of a reward) means that such a practice of virtue is of the highest degree of merit, and the reward for it, which ultimately is of a spiritual nature in the hereafter, will be in accordance with its merit.

PHILOSOPHY OF HISTORY

Finally, Philo's belief in God as a free agent who acts by will and design in the world as a whole, as well as in the life of individual human beings, has led him to a theo-teleological philosophy of history. Alluding to passages in Polybius's *Histories,* in which the rise and fall of cities, nations, and countries are explained by analogy to the Stoic conception of cosmic history as a cyclical process which goes on infinitely, by necessity and for no purpose, Philo describes the cyclical changes in human history as being guided by "the divine Logos" according to a preconceived plan and toward a goal which is to be reached in the course of time. The preconceived plan and goal is that ultimately "the whole world may become, as it were, one city and enjoy the best of polities, a democracy." His description of the ultimate best of polities is an elaboration of the Messianic prophecies of Isaiah and Micah as to what will come to pass in the end of days.

This is a brief synopsis of Philo's revision of Greek philosophic conceptions of the nature of God and his relation to the world and man. The historical significance of Philo is that his revision became the foundation of the common philosophy of the three religions with cognate Scriptures—Judaism, Christianity, and Islam. This triple religious philosophy, which originated with Philo, reigned supreme as a homogeneous, if not a completely unified, system of thought until the seventeenth century, when it was overthrown by Benedict de Spinoza, for the philosophy of Spinoza, properly understood, is primarily a criticism of the common elements in this triple religious philosophy.

See also Aristotelianism; Aristotle; Emotion; Epicureanism and the Epicurean School; Hellenistic Thought; Jewish Philosophy; Logos; Love; Plato; Platonism and the Platonic Tradition; Spinoza, Benedict (Baruch) de; Stoicism; Virtue and Vice.

Bibliography

The present article is based upon H. A. Wolfson, *Religious Philosophy: A Group of Essays* (Cambridge, MA: Belknap Press of Harvard University Press, 1961), and *Philo: Foundations of Religious Philosophy in Judaism, Christianity and Islam,* 3rd ed., rev., 2 vols. (Cambridge, MA: Harvard University Press, 1962).

Philo's writings have been translated in the Loeb Classical Library: *Philo,* translated by F. H. Colson and G. H. Whitaker, 10 vols., and 2 supp. vols. translated by Ralph Marcus (Cambridge, MA, and London, 1929–1962). *Selections from Philo,* edited with an introduction by Hans Lewy, are available in paperback in *Three Jewish Philosophers,* edited by Hans Lewy, Alexander Altmann, and Isaak Heinemann (New York: Meridian, 1960).

Works on Philo written from various points of view and of interest to students of philosophy are N. Bentwich, *Philo-Judaeus of Alexandria* (Philadelphia: Jewish Publication Society of America, 1910); Émile Bréhier, *Les idées philosophiques et religieuses de Philon d'Alexandrie,* 2nd ed., rev. (Paris, 1925); J. Daniélou, *Philon d'Alexandrie* (Paris, 1958); J. Drummond, *Philo Judaeus,* 2 vols. (London, 1888); E. R. Goodenough, *By Light, Light; the Mystic Gospel of Hellenistic Judaism* (New Haven, CT: Yale University Press, 1935); Isaak Heinemann, *Philons griechische und jüdische Bildung* (Breslau, 1932); E. Herriot, *Philon le juif* (Paris, 1898); C. Siegfried, *Philon von Alexandria* (Jena, Germany, 1875); M. Stein, *Pilon ha-Alexandroni* (Warsaw, 1937); and W. Völker, *Fortschritt und Vollendung bei Philon von Alexandrien* (Leipzig, 1938). See also H. L. Goodhart and E. R. Goodenough, "A General Bibliography of Philo," in E. R. Goodenough, *The Politics of Philo Judaeus* (New Haven, CT: Yale University Press, 1938), pp. 125–348.

Harry A. Wolfson (1967)

PHILO JUDAEUS [ADDENDUM]

The original entry on Philo Judaeus was written by Harry Wolfson, one of the preeminent scholars of medieval religious philosophy. A major premise of his general work is that Philo's philosophical project stands as the foundation for the religious philosophizing common to the three monotheistic cultures: Judaism, Christianity, and Islam. Though Philo, a Hellenized Alexandrian Jew of the first century CE, had little impact upon his own people, he had a manifest impact upon the church fathers, and according to Wolfson his "attempt to interpret the scriptural teachings in terms of Greek philosophy" was common philosophical coin until Spinoza, another Jew, in the seventeenth century tore down Philo's harmonizing project.

Philo scholarship was abundant throughout the last few decades of the twentieth century. There originated an annual conference, *The Studia Philonica Annual*. Much recent work has emphasized the Greek (Alexandrian) milieu that incubated Philo and his philosophy. Philo almost certainly knew no Hebrew and was familiar (only)

with the Septuagint version of scripture. Further, his project of teasing out the inherent philosophicality of scripture took the form of allegory—a method adopted from the Stoic method of allegorical exegesis of Homer—that reached final form in the work of Crates of Mallos in the second century BCE. "Armed with Greek allegorical exegesis," writes David Winston, "which seeks out the hidden meanings that lie beneath the surface of any particular text, and given the Middle Platonist and Neo-Pythagorean penchant to read back new doctrines into the works of a venerable figure of the past, Philo was fully prepared to do battle for his ancestral tradition" (1981, p. 6). This passage by Winston describes the tool, and the philosophical prejudices, that motivated Philo to reveal the deepest truths of Scripture. As Maimonides adapted Aristotelian categories for purposes all his own, so Philo is to be understood "as essentially adapting contemporary Alexandrian Platonism, which was itself heavily influenced by Stoicism and Pythagoreanism, to his own exegetical purposes" (Dillon 1977, p. 182). Caught between two cultures, Philo stands as the first monotheistic thinker to find a manifest use for Greek philosophy for explicating his own religious tradition.

See also Aristotelianism; Homer; Maimonides; Platonism and the Platonic Tradition; Pythagoras and Pythagoreanism; Spinoza, Benedict (Baruch) de; Stoicism.

Bibliography

WORKS BY PHILO JUDAEUS

Philo of Alexandria On the Creation of the Cosmos according to Moses: Translation and Commentary. Translated by David T. Runia. Philo of Alexandria Commentary Series 1. Leiden, Netherlands: Brill, 2001.

Philo of Alexandria: The Contemplative Life, the Giants, and Selections. Translated by David Winston. New York: Paulist Press, 1981.

WORKS ON PHILO JUDAEUS AND HIS PHILOSOPHICAL BACKGROUND

Chadwick, Henry. "Philo and the Beginnings of Christian Thought." In *The Cambridge History of Later Greek and Early Medieval Philosophy*," edited by A. H. Armstrong, 135–157. Cambridge, U.K.: Cambridge University Press, 1967.

Dillon, John. *The Middle Platonists*, 139–183. London: Duckworth, 1977.

Hay, David M., ed. *Both Literal and Allegorical: Studies in Philo of Alexandria's Questions and Answers on Genesis and Exodus*. Brown Judaic Series 232. Atlanta: 1991.

Hay, David M. "Philo's References to Other Allegorists." *Studia Philonica* 6 (1979–80): 41–75.

Hengel, Martin. *Judaism and Hellenism*. 2 vols. Philadelphia: Fortress, 1974.

Mendelson, Alan. *Philo's Jewish Identity*. Brown Judaic Series 161. Atlanta: 1988.

Nadler, Steven. "Spinoza and Philo: The Alleged Mysticism in the *Ethics*." In *Hellenistic and Early Modern Philosophy*, edited by J. Miller and B. Inwood, 232–250. Cambridge, U.K.: Cambridge University Press, 2003.

Runia, David. *Exegesis and Scripture: Studies on Philo of Alexandria*. Variorum Collected Studies Series. London: 1990.

Runia, David. *Philo and the Church Fathers: A Collection of Papers*. Vigiliae Christianae Supplements 32. Leiden, Netherlands: Brill, 1995.

Runia, David. *Philo in Early Christian Literature: A Survey*. Compendia Rerum ad Novum Testamentum III Vol. 3. Assen-Minneapolis: 1993.

Runia, David. *Philo of Alexandria: An Annotated Bibliography 1937–86*. With R. Radice. Supplements to Vigiliae Christianae 8. Leiden, Netherlands: Brill, 1988. 2nd edition, Leiden, Netherlands: Brill, 1992.

Runia, David. *Philo of Alexandria: An Annotated Bibliography 1987–96*. Supplements to Vigiliae Christianae. Leiden, Netherlands: Brill, 2000.

Runia, David T. *Philo of Alexandria and the* Timaeus *of Plato*. PhD diss. Free University of Amsterdam. 2 vols. Amsterdam, 1983. Rev. ed., *Philo of Alexandria and the* Timaeus *of Plato*, Philosophia Antiqua 44. Leiden, Netherlands: Brill, 1986.

Sandmel, Samuel. *Philo of Alexandria: An Introduction*. New York: Oxford University Press, 1979.

Sterling, Gregory E., ed. *The Ancestral Tradition: Hellenistic Philosophy in Second Temple Judaism. Essays of David Winston*. Studia Philonica Monographs 4, Brown Judaic Series 331. Providence, RI: 2001.

Winston, David. "Aspects of Philo's Linguistic Theory." *The Studia Philonica Annual* 3 (1991): 109–125.

Winston, David. "Judaism and Hellenism: Hidden Tensions in Philo's Thought." *The Studia Philonica Annual* 2 (1990): 1–19.

Winston, David. *Logos and Mystical Theology in Philo of Alexandria*. Cincinnati, OH: Hebrew Union College Press, 1985.

Winston, David. "Philo and the Hellenistic Jewish Encounter." *The Studia Philonica Annual* 7 (1995): 124–142.

Winston, David. "Philo's Mysticism." *The Studia Philonica Annual* 8 (1996): 74–82.

Winston, David. "Philo's *Nachleben* in Judaism." *The Studia Philonica Annual* 6 (1994): 103–110.

Winston, David. "Plato's Conception of the Divine Nature." In *Neoplatonism and Jewish Thought*, edited by L. E. Goodman, 21–42. Albany: State University of New York Press, 1992.

Winston, David, and John Dillon. *Two Treatises of Philo of Alexandria: A Commentary on De Gigantibus and Quod Deus Sit Immutabilis*. Brown Judaic Series 25. Chico, CA: Scholars Press, 1983.

Daniel H. Frank (2005)

PHILOLAUS OF CROTON
(c. 470–385 BCE)

Philolaus of Croton (a Greek city in southern Italy) was a philosopher/scientist in the Pythagorean tradition. He was a contemporary of Socrates, being born c. 470 BCE, twenty years after Pythagoras died, and living until c. 385. On his first trip to Italy, Plato may have met an aged Philolaus; he mentions him as a teacher of the Thebans Simmias and Cebes in the *Phaedo*. A large body of pseudo-Pythagorean writings appeared in the first century BCE, and a number of these were forged in Philolaus's name, because he was one of the three most famous early Pythagoreans (along with Pythagoras himself and Archytas). Some fifteen fragments and a number of testimonia survive from these forged works. Philolaus, in fact, wrote one book, *On Nature*, which was probably the first book in the Pythagorean tradition (Pythagoras wrote nothing). Approximately eleven genuine fragments of that book have survived along with a number of testimonia. Aristotle discusses Pythagorean philosophy extensively but does not assign this philosophy to Pythagoras himself but rather to the "people called Pythagoreans," whom he treats as slightly older contemporaries of the atomists. This dating fits Philolaus exactly, and the agreement between the philosophy described by Aristotle and the fragments of Philolaus's book shows that Philolaus was the primary source for Aristotle's account.

Philolaus argued that the nature of the cosmos as a whole and of all things in it was to be explained in terms of two types of elements, unlimiteds and limiters. The unlimiteds include the material elements favored by his predecessors in the pre-Socratic tradition, such as earth, air, fire, and water but also continua such as space and time. Philolaus is emphatic, however, that such principles are not adequate to explain the cosmos because (1) limits, such as shapes, are also part of the cosmos humans can observe, and (2) such limiting features cannot have arisen from what is unlimited. Philolaus's cosmogony illustrates the role of these principles; the first thing to emerge was the central fire, which is a combination of the unlimited, fire, and the limiter, center. This central fire then draws in other unlimiteds such as time, void, and breath, which will be combined with limits to produce the cosmos known to humans.

Philolaus introduces harmony as an essential third principle, which specifies the way in which limiters and unlimiteds are combined. The central example is the musical scale in which the unlimited continuum of sound is limited by specific notes; harmony insures that these notes do not have a haphazard order, however, but are "fitted together" in accordance with whole number ratios. This idea depends on the earlier Greek discovery that, if a person plucks two strings, one of which is twice the length of the other, we will hear the interval of the octave between the two sounds, so that the octave corresponds to the ratio 2: 1. Similarly, the fifth will correspond to the ratio 3: 2 and the fourth to the ratio 4: 3. Philolaus appears to regard the cosmos as a whole as structured according to the ratios that determine diatonic scale. Plato may be influenced by Philolaus in using this same scale to construct the world soul in the *Timaeus*.

By specifying the "formula" according to which limiters and unlimiteds combine, numbers also define the essence of a given thing and thus play an important epistemological role for Philolaus: "And indeed all things that are known have number. For it is not possible that anything whatsoever be understood or known without this" (Fr. 4). On the one hand, Aristotle is clearly right that numbers are not separate from things in this system, as they were later in Plato; Philolaus and his successor Archytas were interested in the numbers of things, not in numbers separated from things. On the other hand, Aristotle's suggestion that the Pythagoreans thought that things just were numbers or that they were made of numbers is not supported by the fragments of Philolaus, where it is clear that things are made of limiters and unlimiteds. According to Philolaus, our senses reveal a world composed of unlimiteds and limiters (e.g., stuffs and shapes), but on further examination the phenomena point to the numerical ratios that govern them. It is doubtful that Philolaus had explicitly addressed the metaphysical status of these ratios. Aristotle may have thought that if numbers reveal the essence of things, then things are, in an important sense, numbers; but this is Aristotelian interpretation. Philolaus prefers to say that things are composed of limiters and unlimiteds and known through the numerical ratios in accordance with which the limiters and unlimiteds are combined.

Philolaus is the first person to move the earth from the center of the universe and make it a planet, and Copernicus saw Philolaus as an important predecessor. The earth does not orbit around the sun in Philolaus's system, however; the fixed stars, five planets, sun, moon, earth, and an enigmatic counter-earth all orbit around the central fire. The system may have some origins in a religious cosmology in which the central fire is identified with Tartarus, a region under the earth where the guilty are punished in Greek mythology. Aristotle suggests that the counter-earth was introduced to satisfy the *a priori*

requirement that there be ten heavenly bodies around the central fire, because the Pythagoreans regarded ten as the perfect number. On the other hand, there is clear evidence that Philolaus intended the system to explain astronomical phenomena as well as satisfying *a priori* or religious requirements. The system can explain basic phenomena and is the first to include the five known planets in correct order, although it cannot account for such things as the apparent retrograde motion of planets. Philolaus clearly responded to objections to his system, which were based on the phenomena, arguing that the motion of the earth around the central fire did not produce a parallax effect, because the distance from the earth to the central fire was small in comparison to the distance between the earth and the planets. Similarly human beings never see the central fire or counter-earth, because the side of the earth on which they live is always turned away from the center, the earth rotating once on its axis during each orbit of the central-fire.

Philolaus argued that in each area of inquiry it was necessary to begin by identifying the minimum number of principles required to explain the phenomena. Limiters, unlimiteds, and harmony are the basic metaphysical principles; bile, blood, and phlegm explain disease; intellect, sensation, nutrition/growth, and generation are the basic psychic faculties. Philolaus drew an analogy between the birth of the cosmos and the birth of a human being, arguing that the embryo is initially hot and draws in cooling breath immediately upon birth, just as the cosmos begins with the central fire drawing in breath from the unlimited. It may be that he regarded the soul as a harmony of physical opposites, a view that Plato, perhaps in criticism of Philolaus, shows in the *Phaedo* to be inconsistent with a belief in an immortal soul.

In the *Philebus*, "the method of the men before our time," which Plato adapts to address problems in his own metaphysics, is clearly the metaphysical system of Philolaus, which thus had a significant impact on Plato's later metaphysics. Some have argued that Philolaus's metaphysics must go back to Pythagoras, but Aristotle clearly dates it to the time of Philolaus, and the system itself—with its emphasis on the necessity of limiters in addition to unlimiteds makes most sense, if it arose after Pythagoras—at a time when Parmenides had championed the role of limit in explaining reality.

See also Archytas of Tarentum; Pythagoras and Pythagoreanism.

Bibliography

TEXTS AND COMMENTARY

Huffman, C. A. *Philolaus of Croton: Pythagorean and Presocratic*. Cambridge, U.K.: Cambridge University Press, 1993.

DISCUSSIONS

Barnes, J. *The Presocratic Philosophers*. London: Routledge, 1982.

Burkert, W. *Lore and Science in Ancient Pythagoreanism*. Translated by E. Minar. Cambridge, MA: Harvard University Press, 1972.

Kahn, Charles H. *Pythagoras and the Pythagoreans*. Indianapolis: Hackett, 2001.

Kingsley, Peter. *Ancient Philosophy, Mystery and Magic*. Oxford: Clarendon, 1995.

Meinwald, Constance Chu. "Plato's Pythagoreanism." *Ancient Philosophy* 22 (1) (2002): 87–101.

Schibli, H. S. "On 'The One' in Philolaus, Fragment 7." *The Classical Quarterly*, n.s., 46 (1) (1996): 114–130.

Sedley, David. "The Dramatis Personae of Plato's *Phaedo*." In *Philosophical Dialogues: Plato, Hume, Wittgenstein*, edited by T. J. Smiley, 3–26. Oxford: Oxford University Press, 1995.

Carl A. Huffman (2005)

PHILO OF LARISSA
(159/8–84/3 BCE)

Philo of Larissa was a student of Clitomachus (187/6–110/9 BCE), whom he succeeded as head of the Academy in 110/09 BCE. In 88 BCE Philo transferred his activities from Athens to Rome, where Marcus Tullius Cicero, among others, studied under him. Present-day evidence does not allow one to say for certain whether Philo was the last head of the Academy or was succeeded by his student, Antiochus of Ascalon.

Philo taught rhetoric as well as philosophy, and an extended analogy of his between the way in which philosophy cares for the soul and the way in which medicine cares for the body has been preserved. But he seems to have been chiefly interested in epistemology, then the dominant concern of the Academy, and scholars are best informed about his views in this area.

It is likely that Philo first upheld Clitomachus's version of Academic skepticism, which endorsed the two theses for which Academics had argued in the their controversy with the Stoa since the time of Arcesilaus, who was head of the Academy in the mid-third century BCE. These are that nothing can be known—or a conclusion that amounts to this in the context of the debate with the Stoa—and that, in consequence, one should suspend judgment about all matters. As head of the school, how-

ever, he defended a mitigated form of skepticism that continued to embrace the thesis that nothing can be known, but now permitted assent to probable impressions, among them the impression that nothing can be known. The account of probability on which this view depends (*probabile* is Cicero's Latin for the Greek *pithanon*, meaning "persuasive") had been developed by Carneades (214/3–129/8 B.C.E.), Clitomachus's teacher, as an alternative to cognitive impressions that the Stoics had made the foundation of their epistemology and that supposedly afforded an absolutely secure guarantee of truth. This position had been anticipated by Metrodorus of Stratonicea, another pupil of Carneades, and seems to have been the position to which Aenesidemus, the one-time Academic who revived the Pyrrhonian school of skepticism in the first century BCE, objected.

In Rome, however, Philo came to hold that knowledge is possible. He did this not by renouncing the Academy's arguments against the Stoa, but by reinterpreting them. He now took them to show, not that knowledge is impossible, but that knowledge is impossible on the Stoic conception of knowledge, which is therefore mistaken. The fault lay with their insistence on a foundation of impressions that could not be false, a condition that the Academy had long argued could not be met and that Philo now held need not be met. And he maintained that his Academic predecessors had never intended to show anything else by their arguments. These new views were opposed by Academics who remained attached to skepticism and Antiochus, who had become convinced that knowledge is possible precisely because the Stoic conditions could be satisfied.

None of Philo's writings have survived. Though he probably wrote other works on epistemology and ethics, the only books we know of are the so-called "Roman Books," in which Philo set out his late views on knowledge and the history of the Academy.

See also Ancient Skepticism; Antiochus of Ascalon; Arcesilaus; Carneades; Cicero, Marcus Tullius; Stoicism.

Bibliography

Barnes, Jonathan. "Antiochus of Ascalon." In *Philosophia Togata: Essays on Philosophy and Roman Society*, edited by Miriam Griffin and Jonathan Barnes. New York: Oxford University Press, 1989.

Brittain, Charles. *Philo of Larissa: The Last of the Academic Sceptics*. New York: Oxford University Press, 2001.

Striker, Gisela. "Academics Fighting Academics." In *Assent and Argument: Studies in Cicero's Academic Books*, edited by Brad Inwood and Jaap Mansfeld. Leiden, Netherlands: Brill, 1997.

James Allen (2005)

PHILO OF MEGARA

(c. 400s BCE)

Very little is known about the life of Philo of Megara, or Philo Dialecticus. Since he was a pupil of Diodorus Cronus at the same time as Zeno of Citium, the founder of the Stoa (cf. Diogenes Laertius, DL 7.16), he was very probably active in Athens in the last decade of the 4th century BCE. He was not, as is assumed in the older literature, a member of the Megarian school of philosophy, but belonged to a separate sect, the Dialecticians. Hence there is no reason to make Megara his birthplace. From the titles of two lost treatises by the Stoic Chrysippus that were directed against Philo, we learn that Philo wrote *On Signs* (DL 7.191) and *On Moods (of Argument)* (DL 7.194). He also wrote a dialogue called the *Menexenus*, in which the five daughters of the Dialectician Diodorus Cronus, all of them also Dialecticians, were made to appear. It is possible that the theory of signs referred to in Pseudo-Galen's *Historia philosopha* c. 9 as belonging to the "dialecticians" goes back to Philo's treatise. The logical terminology in this report is in accordance with that used by the Dialecticians, and the epistemological terminology does not yet show Stoic influence. Signs are here defined as a special class of conditionals, namely sound conditionals with a true antecedent revealing the consequent. We are on safer ground with two other claims attributed to Philo, one concerning implication, the other the definition of modal concepts.

Philo argued that a conditional is true if and only if it is not the case that its antecedent is true and its consequent false (cf. Sextus Empiricus, SE *Adv. Math.* 8.113–114). Hence Philo seems to have given for the first time a truth-functional definition of the conditional. Against this claim, Diodorus Cronus held that a conditional is true if and only if it was not possible and is not possible that its antecedent is true and its consequent false (cf. SE, *Adv. Math.* 8.115–117). Thus the conditional "If it is day, I am talking," which proves to be true, according to Philo, provided that I am talking while it is day, will be false according to Diodorus. Although Sextus Empiricus in his report on this dispute has the consequent "follow" from the antecedent, it is not clear whether Philo and/or Diodorus want to make their criteria for the truth of the conditional a sufficient condition for the validity of an argument. It would have rather bizarre consequences in both cases: For Philo, any true propositions would entail each other, and for Diodorus, any true propositions about the past would entail each other.

Philo defines the possible as that "which, by the intrinsic nature of the proposition, is receptive of truth"; he defines the necessary as that "which, since it is true, by its own nature, is never receptive of falsehood". Similarly, the non-necessary is defined as that "which by its own nature is receptive of falsehood" and the impossible as that "which according to its own nature could never receive truth" (cf. Boethius, *De interpretatione* ii, 234). Here again he disagrees with Diodorus, who defines the possible as that "which either is or will be (true)." For Diodorus there can thus be no unrealized possibilities, whereas this is possible with Philo. Philo's modal logic, like that of Aristotle, seems to be based on an essentialist epistemology.

See also Chrysippus; Diodorus Cronus; Zeno of Citum.

Bibliography

WORKS BY PHILO OF MEGARA

The testimonia on Philo are to be found together with the material on Diodorus Cronus in vol. 1 of G. Giannantoni, *Socratis et Socraticorum reliquiae*. Naples: Bibliopolis, 1990, pp. 414–435. For a possibly Philonian theory of signs see: H. Diels. *Doxographi Graeci*. Berlin: Reimer, 1879, p. 605 (Pseudo-Galen, *Historia philosopha* c. 9).

WORKS ABOUT PHILO OF MEGARA

Kneale, W., and M. Kneale. *The Development of Logic*. Oxford: Clarendon, 1962 (pp. 122, 125f., 128–134).

Ebert, Theodor. "The Origin of the Stoic Theory of Signs in Sextus Empiricus." *Oxford Studies in Ancient Philosophy* 5 (1987): 83–126.

Theodor Ebert (2005)

PHILOPONUS, JOHN

(490–570)

John Philoponus of Alexandria, a sixth-century philosopher and theologian, is best known for his radical attempts to refute fundamental tenets of contemporary Aristotelian–Neoplatonic school philosophy. His main historical significance lies in the fact that he anticipated by centuries the early modern emancipation of natural philosophy from Aristotelian dogmatism. Philoponus (literally *Lover of Work*), or John the Grammarian, as he called himself, is commonly labeled a Christian Neoplatonist, but this epithet is misleading. Philoponus was a Christian, most likely by birth, and he received the standard philosophical training available at Alexandria in his day. Thus, his philosophical orientation was not a matter of choice, and his fierce rationalism, which he employed also as a tool to resolve controversial questions that divided Christianity, bears no resemblance to the genuine Christian Neoplatonism of Pseudo-Dionysius, the Areopagite (c. 500). Roughly 100 years after his death, the Third Council of Constantinople (680–681) condemned his theological doctrines as heresy and thereby curtailed the overall philosophical influence he could have had in later centuries.

Almost everything about Philoponus's life remains a matter of hypothesis. He was born presumably around 490 CE, but it is not known where (Kaster 1988); in the early sixth century, he studied in Alexandria, reading philosophy under Ammonius, Son of Hermias (c. 440–520), who had been a pupil of Proclus at Athens. In the early 520s, Philoponus taught both grammar and philosophy at Alexandria; some of his early commentaries on Aristotle are based on Ammonius's lectures, but in the process of multiple revisions, Philoponus added explanations, observations, and criticisms of his own. In the late 520s, early 530s, around the time of Justinian's eviction of the pagan philosophers in Athens (c. 529), Philoponus turns to writing polemical commentaries (on Proclus and Aristotle), which no longer aim at elucidation but at refutation, especially of the pagan doctrine of eternalism. These works provoked immediate condemnation of by Simplicius of Cilicia, a contemporary member of the Athenian School, the last great pagan mind of antiquity and expert commentator on Aristotle.

Although Philoponus was one of the most powerful and independent thinkers of his time, he never succeeded Ammonius as professor of philosophy. The reasons for this are unclear; although unknown external, personal or political circumstances may have played a role, a likely explanation is that Philoponus had reached a point where his fundamental disagreement with the philosophical establishment compromised his ability to continue the pedagogical tradition of the school. Leadership of the philosophical school remained in pagan hands well into the second half of the sixth century. Philoponus's later writings, from the 540s onward, deal with contemporary issues of theological controversy. He expounded his theological views with philosophical rigor, whether rejecting the orthodox belief in the divine–human duality of the nature of Christ (miaphysitism) or defending the substantial distinctness of the hypostases of the Trinity (tritheism). Philoponus must have died around 570.

Philoponus' *œuvre*, which bears witness to his interests in grammar, philosophy, psychology, medicine, mathematics, astronomy, and theology, may be divided into three related yet distinct parts (Scholten 1996): (1)

The commentaries on Aristotle (*Categories, Prior and Posterior Analytics, Physics, On Generation and Corruption, Meteorology*, and *On the Soul*); (2) the treatises of the critical period, notably, the two monumental polemical treatises *On the Eternity of World against Proclus* (shortly after 529) and the influential *On the Eternity of World Against Aristotle* (early 530s, extant only in fragments); and (3) a number of works on theological doctrine, some of which are only extant in Syriac translation; most important of the last group is a still-extant commentary on the biblical creation myth (*On the Making of the World*, written between 546 and 560) which also targets the naïve Christian cosmography of Cosmas Indicopleustes.

In his philosophical works, one can roughly distinguish between two kinds of criticism: on the one hand, the grappling with implausible Aristotelian theories, mostly physical, and on the other hand, outright repudiation of fundamental cosmological doctrines. Aristotle's definition of light as an incorporeal and instantaneous transition from the potentiality (*dunamis*) of a medium to be transparent to the actuality (*energeia*) of transparency fails to account for the laws of optics and for the calefactory property of the sun. Philoponus proposes to interpret light as an *incorporeal activity* rather than a state, capable of warming bodies and comparable to the soul in animals. Later, in the *Meteorology* commentary, which may be the transcript of his last lecture series on Aristotle, he argues materialistically that light and heat are consequences of the fiery nature of the sun, and that heat is generated when the rays emanating from the sun are refracted and warm the air through friction.

The *Physics* commentary contains one of his most celebrated achievements, the theory of the impetus, which is commonly regarded as a decisive step from an Aristotelian dynamics toward a modern theory based on the notion of inertia. To what extent Philoponus was influenced by previous philosophical or theological authors is a matter of controversy (Fladerer 2003). His own discussion, at any rate, commences with the expression of dissatisfaction with Aristotle's explanation that a projectile continues to move on account of the air's turbulence generated by the projectile itself. Philoponus proposes instead that a projectile moves on account of a kinetic force that is impressed on it by the mover and that exhausts itself in the course of the movement. In short, the medium contributes nothing to a projectile's motion; rather, it impedes it. Moreover, Philoponus holds that there is nothing to prevent motion from taking place in the void.

Occasionally, Philoponus resorts not only to thought experiments but also to pertinent observations that resemble physical experimentation. Aristotle's verdict that the speed of a falling body is proportional to its weight and indirectly proportional to the density of the medium is challenged by the same kind of empirical evidence that Galileo mustered centuries later.

Philoponus is critical of Aristotle's conception of space. He substitutes Aristotle's definition of the *place* of a body (the inner surface of that which contains it) with a conception of three-dimensional extension, its volume. Likewise, the most fundamental level of physical reality is not some mysterious *prime matter* but three-dimensional, indeterminate, and unqualified *corporeal extension*, a concept reminiscent of Descartes's *res extensa*.

The issue at stake in the two polemics against Proclus and Aristotle is the question of the contingency of the world. The earlier work obliterates a pamphlet of eighteen arguments for the eternity of the world written in the previous century by the powerful Neoplatonist Proclus. The lost *Against Aristotle* tackled influential arguments for eternity in *On the Heavens I* and *Physics VIII*. In both cases, Philoponus succeeds in pointing out numerous contradictions, inconsistencies, fallacies, and improbable assumptions. One clear casualty is Aristotle's peculiar postulate of an incorruptible celestial element (ether). The observable irregularities in the heavens, their complexity and changes in color, undermine the thesis of the radical ontological difference between the celestial and sublunary regions. Dissecting the text in unprecedented ways, Philoponus even paves the way for influential demonstrative arguments for noneternity. Although Philoponus concedes that in nature nothing comes to be from nothing, he offers the first *philosophical* defense of the Christian belief that God created the world *ex nihilo*. In the late theological treatise *On the Making of the World*, Philoponus suggests in passing that the celestial bodies were set to spin by a powerful impetus at the time of their creation and that they now continue to move not on account of their own nature but by the will of God.

It is impossible to gauge how Philoponus's ideas resonated with Christians during his lifetime. He was read and admired by Syrian and, to some extent, Islamic philosophers, but the anathema of 681 severely hampered the further propagation of his theological and philosophical work. As Simplicius before them, later thinkers like Thomas Aquinas (1224–1274) and Zabarella (1533–1589) roundly rejected Philoponus. Eventually, the arguments against eternity persuaded Bonaventure

(1217–1274) and Gersonides (1288–1344), and the theory of the impetus was reaffirmed by Buridan (1295–1356) and Oresme (1325–1382). In the sixteenth century, the first editions as well as numerous translations (into Latin) of the commentaries and the treatise against Proclus began to appear in print. In particular, Philoponus's criticism of Aristotle in the *Physics* commentary was widely discussed and persuaded such diverse thinkers as Gianfrancesco Pico della Mirandola (1469–1533) and Galileo Galilei (1564–1642).

See also Aristotelianism; Bonaventure, St.; Buridan, Jean; Galileo Galilei; Gersonides; Impetus; Neoplatonism; Oresme, Nicholas; Pico della Mirandola, Gianfrancesco; Proclus; Pseudo-Dionysius; Simplicius; Thomas Aquinas, St.; Zabarella, Jacopo.

Bibliography

MAJOR WORKS

On the Use and Construction of the Astrolabe, edited by H. Hase. Bonn: Weber, 1839 (ibid., Rheinisches Museum für Philologie 6, 1839, 127–171). Reprinted and translated into French by Alain P. Segonds, *Jean Philopon, traité de l'astrolabe*. Paris: Librairie Alain Brieux, 1981. Translated into English by H. W. Green in *The Astrolabes of the World*. Vols. 1 & 2. Robert T. Gunther. Oxford: University Press, 1932; Reprinted, London: Holland Press, 1976, pp. 61–81.

Commentary on Aristotle's Physics. In *Commentaria in Aristotelem Graeca*, XVI–XVII, edited by H. Vitelli. Berlin: Reimer, 1887–1888. Philoponus's most important commentary in which he challenges Aristotle's tenets on time, space, void, matter, and dynamics; there are signs of revision. Partial translations: *Philoponus, On Aristotle's Physics 2*, by A. R. Lacey. London: Duckworth, 1993; *Philoponus, On Aristotle's Physics 3*, by M. Edwards. London: Duckworth, 1994; *Philoponus, On Aristotle's Physics 5–8*, by Paul Lettinck. London: Duckworth 1993–1994; *Philoponus, Corollaries on Place and Void*, by David Furley. London: Duckworth, 1991.

On the Making of the World (*De opificio mundi*), edited by Wilhelm Reichardt. Leipzig: Teubner, 1897. Greek text and German translation by Clemens Scholten. *De opificio mundi = Über die Erschaffung der Welt*. Freiburg, New York: Herder, 1997. A theological–philosophical commentary on the Creation story in the book of Genesis, written probably 557–560.

On the Eternity of the World against Proclus (*De aeternitate mundi contra Proclum*), edited by Hugo Rabe. Leipzig: Teubner, 1899. Reprinted, Hildesheim: Olms, 1984. Translated by Helen S. Lang and Anthony D. Macro. *De aeternitate mundi. English and Greek*. Berkeley: University of California Press, 2001. A detailed criticism of Proclus's eighteen arguments in favor of the eternity of the world.

On the Eternity of the World against Aristotle (*De aeternitate mundi contra Aristotelem*). Not extant; fragments reconstructed and translated by Christian Wildberg. *Philoponus, Against Aristotle on the Eternity of the World*. London: Duckworth, 1987. A refutation of Aristotle's doctrines of the fifth element and the eternity of motion and time consisting of at least eight books.

FURTHER READING

Fladerer, Ludwig. "Johannes Philoponos, Gregor von Nyssa und die Genese der Impetustheorie." In *Hommages à Carl Deroux V*, edited by Pol Defosse. Collection Latomus 279 (2003): 138–151.

Haas, Frans A. J. de. *John Philoponus' New Definition of Prime Matter: Aspects of its Background in Neoplatonism and the Ancient Commentary Tradition*. Leiden, New York: E. J. Brill, 1997.

Kaster, Robert A. *Guardians of Language. The Grammarian and Society in Late Antiquity*, 334–338. Berkeley: University of California Press, 1988.

Sambursky, Samuel. *The Physical World of Late Antiquity*, 154–175. London: Routledge and Kegan Paul, 1962.

Scholten, Clemens. *Antike Naturphilosophie und christliche Kosmologie in der Schrift "De opificio mundi" des Johannes Philoponos*. Berlin: De Gruyter, 1996.

Sorabji, Richard R. K. *Matter, Space, and Motion*, 227–248. London: Duckworth, 1988.

Sorabji, Richard R. K., ed. *Philoponus and the Rejection of Aristotelian Science*. London: Duckworth, 1987.

Sorabji, Richard R. K. *Time, Creation and the Continuum*, 193–231. London: Duckworth, 1983.

Verrycken, Koenraad. "The Development of Philoponus' Thought and Its Chronology." In *Aristotle Transformed*, edited by Richard R. K. Sorabji, 233–274. London: Duckworth, 1990.

Wildberg, Christian. "Impetus Theory and the Hermeneutics of Science in Simplicius and Philoponus." *Hyperboreus* 5 (1999): 107–124.

Wildberg, Christian. *John Philoponus' Criticism of Aristotle's Theory of Aether*. Berlin: De Gruyter, 1988.

Christian Wildberg (2005)

PHILOSOPHICAL ANTHROPOLOGY

Modern philosophical anthropology originated in the 1920s. During the 1940s it became the representative branch of German philosophy. It arose with, and has absorbed, *Lebensphilosophie,* existentialism, and phenomenology, although it is not identical with them. It has affinities with pragmatism and the sociology of knowledge. Although it is historically based on certain German traditions, it is also indebted to, and largely anticipated by, the eighteenth-century "science of human nature." It combines the critical traditions of the Enlightenment with an emphasis on dogmatic certitude.

HISTORICAL BACKGROUND

Following Bernhard Groethuysen, philosophical anthropology is often conceived as embracing all previous philosophy, insofar as previous philosophy dealt with man's place in the world. But this wide conception blurs the distinctive features of philosophical anthropology. Its history is best restricted to those authors and ideas whose impact is either admitted or can be traced in the literature of modern philosophical anthropology.

The impact of Søren Kierkegaard, Karl Marx, and Friedrich Nietzsche is pervasive. Other generally acknowledged forerunners are Blaise Pascal, Johann Gottfried Herder, Johann Wolfgang von Goethe, Immanuel Kant, G. W. F. Hegel, and Ludwig Feuerbach. Pascal's influence is discernible in philosophical anthropology's conception of man as self-contradictory and mysterious, capable of surpassing his natural limits in quest of authenticity. Pascal's distinction between the organic *esprit de finesse* and the abstract and lifeless *esprit géométrique* was accentuated by Kant's distinction between the phenomenal world of the senses, with its quest for happiness (in the sense of egotistic pleasure), and the noumenal world of the thing-in-itself, between a world of determinate law and a world of transcendental choice. These concepts reveal themselves in the philosophical anthropologists' assumption of an unbridgeable gap between value and reason, between the ideal and the practical. Kant's basic questions—"What can I know? What ought I to do? What may I hope?"—are universally accepted in philosophical anthropology.

Herder was the first German author to correlate biology and the philosophy of man. From him stems the conception of man as a deficient being who must compensate for his lack of natural tools and weapons by the creative use of weapons and technology. Hegel's theory of alienation and its Marxist version have become a vital element in philosophical anthropology's comprehension and critique of society. Feuerbach formulated the claim that man can be used as the common denominator of philosophy, the true *ens realissimum,* embracing reason, will, and emotion. He held that philosophical anthropology was to take the place of theology; and indeed, contemporary philosophical anthropology may be regarded as secularized theology. Feuerbach conceived of God as a projection and objectification of the human spirit, reflecting the categorial structure of the human mind and its conceptual tools. This, as well as the corresponding Hegelian view of the divine spirit as being reflected in human history, is one of the recurring themes of cultural philosophical anthropology.

In a specifically German version and modified by the methodology of the practitioners of the *Geisteswissenschaften,* the "science of human nature," which stemmed from Thomas Hobbes, John Locke, and the Earl of Shaftesbury and reached its culmination in the eighteenth century, is the principal root of philosophical anthropology. David Hume's *Treatise of Human Nature* provided a program for philosophical anthropology. "There is no question of importance whose decision is not comprised in the science of man.… In pretending to explain the principles of human nature we in effect propose a complete system of the sciences" (Everyman ed., Vol. I, p. 5). Philosophical anthropology took up Hume's empiricism with regard to the moral sciences, as well as his conception of religion.

Adam Smith's spectator theory of the moral sentiments was an early statement of the excentric position of man. The "Newtonian-Baconian" school of Scottish and French social thought of the eighteenth century (Francis Hutcheson, Adam Ferguson, John Millar, Dugald Stewart, Pierre-Louis Moreau de Maupertuis, Denis Diderot, and Jean Le Rond d'Alembert), which culminated in John Stuart Mill's sociology, was a direct precursor of philosophical anthropology in its aim of putting the study of man on an empirical biological basis. This school sought to elucidate and bridge the gap between man's distinctive nature and the sociocultural order in "the belief that it was natural for man to make an order of life different from that in which the race was nurtured earlier, that it was in the nature of his equipment that he should react intelligently and creatively to the situations in which he found himself" (G. Bryson, *Man and Society,* Princeton, NJ, 1945, p. 173).

The more widely recognized forerunners of philosophical anthropology—Herder, Christian Garve, and Wilhelm von Humboldt—were directly influenced by the Scottish and French anthropologists and Encyclopedists, who had undermined Cartesian dualism. Thus, at the end of the eighteenth century, there was a wide acceptance of certain propositions concerning man's creative powers, his individuality, and his sociability. The Scottish and French precursors, however, had intended to develop more rigorous methods of investigation than those used by contemporary philosophical anthropologists.

SUBJECT MATTER, ATTITUDE, AND GOAL

Like existentialism and *Lebensphilosophie,* philosophical anthropology studies man's existence, his experiences, and his anxieties, combining the subjectivism of existen-

tialism with the cultural objectivism of *Lebensphilosophie.* It uses the phenomenological methods of *Verstehen* and reduction. Philosophical anthropology shares with existentialism, phenomenology, and *Lebensphilosophie* a critique of society. Yet these currents are not identical; Heidegger and Karl Jaspers, for example, refuse to be identified with philosophical anthropology, despite their great impact on it.

Philosophical anthropology seeks to interpret philosophically the facts that the sciences have discovered concerning the nature of man and of the human condition. It presupposes a developed body of scientific thought, and accordingly, in its program it aspires to a new, scientifically grounded metaphysics. It seeks to elucidate the basic qualities that make man what he is and distinguish him from other beings. It combines, and mediates between, what Kant designated as physiological and pragmatic anthropology.

Physiological anthropology studies man's natural limitations; pragmatic anthropology deals with man's potentialities, with what he, as a free agent, makes of himself, or is able and ought to make of himself. Thus, philosophical anthropology studies both man as a creature and man as the creator of cultural values—man as seen by a scientific observer and man as interpreted by himself (*Aussen*- and *Innenansicht*). Accordingly, most philosophical anthropologists wish to combine scientific methods with an imaginative philosophical approach.

Philosophical anthropology seeks to correlate the various anthropologies that have developed with the specialization of the sciences. Max Scheler distinguished between scientific, philosophical, and theological anthropologies, or interpretations of the fundamental structure of human activities, which know nothing of one another. In order to stem what its followers describe as anarchy of thought and the "loss of the center," philosophical anthropology offers itself as a coordinating discipline. With the dissolution of traditional beliefs in guidance by gods, by kings and feudal leaders, by God, or by nature, there is today a general lack of direction. Man is now, as he was for Protagoras, the only possible measure. By coordinating and interpreting fragmented knowledge, philosophical anthropology aims at a new understanding of man's essential qualities and potentialities. It aims to accomplish this by the development of suitable methods, by a factual elucidation of the perplexities inherent in human institutions, and by borderline research (coordinating different branches of the sciences) used as a basis for a new "map of knowledge."

Since philosophical anthropology arose as an interpretation of various scientific disciplines, it has practitioners in many fields. Although there are only a few academic chairs of philosophical anthropology (Göttingen, Nijmegen), the number of professed philosophical anthropologists is large, chiefly in the German-speaking countries, but also in the Netherlands, Spanish-speaking countries, the United States, and France. Modern French humanism, whether existentialist, religious, or Marxist, is both historically and analytically allied with philosophical anthropology. Many philosophical anthropologists stress that they are theological, historical, political, juristic, biological, phenomenological, or cultural philosophical anthropologists. Much so-called philosophical anthropology is best treated under metaphysics, ontology, theory of value, epistemology, theology, philosophy of science or of history, or under the related contemporary philosophies. This entry will discuss only the distinctive features.

Philosophical anthropology embraces most of the social sciences. Some leading practitioners, such as Arnold Gehlen, emphasize the concept of action, rather than man, as the distinguishing feature of philosophical anthropology, and define it as a new empirical discipline, *Handlungswissenschaft* (similar to "behavioral science" and the "theory of action"), as distinct from the natural sciences and the *Geisteswissenschaften.*

Philosophical anthropology is an attempt to construct a scientific discipline out of man's traditional effort to understand and liberate himself. At the same time, however, it is pervaded by the same antiscientific currents that mark existentialism, *Lebensphilosophie,* and phenomenology. But it is its dialogue with science that gives philosophical anthropology its peculiar character.

THE CRISIS OF SCIENCE

Philosophical anthropologists see a "crisis of science," a crisis first brought into view by three "humiliations of man." First, the humiliation of Copernican astronomy removed man's habitat, the earth, from the center of the universe; second, Charles Darwin's biological evolutionism "shamed and degraded" man; and third, the historical schools revealed the relativity of religious and national cultural values. The crisis in science has been brought to a head by modern developments in depth psychology, post-Euclidean mathematics, and the indeterminacy principle in nuclear physics. From the scientific point of view, these developments represent advances rather than a crisis. However, German philosophers since Kant have conceived of science as being fixed in a rigid mathemati-

comechanical determinism. According to philosophical anthropologists, this basic concept has broken down. There is a wide consensus among Continental thinkers that nineteenth-century materialism has been overcome and that the methods of the *Geisteswissenschaften* and phenomenology have been vindicated.

These methods seek the meaning immanent in events and in the works of man rather than the causal nexus between events. They aim to interpret other minds (both individual and collective), their peculiar intentions and tendencies, and the institutions through which their ideas have found expression. They investigate the conscious and unconscious actions of human beings and the structure of interpersonal (social and cultural) relationships. These methods are descriptive, interpretative, organic, and concrete, rather than explanatory, mechanical, and abstract, as in the natural sciences. This distinction of two methodologies—causal explanation on the one hand and *Verstehen* and phenomenological reduction on the other—takes up the emphasis of what is known in English as the Germano-Coleridgean school on, in the words of J. S. Mill, a philosophy of society in the form of a philosophy of history seeking a philosophy of human culture.

THEORY OF KNOWLEDGE

The crisis of science, according to philosophical anthropologists, evinces a deep crisis in the theory of knowledge—a crisis that makes imperative the adoption of pragmatic theories of truth. Traditional epistemology, they claim, was occupied with only one of the functions of consciousness. It failed to take into account what Pascal called the *logique du coeur* or *esprit de finesse,* which was akin to Samuel Taylor Coleridge's "imagination" and John Henry Newman's "illative reason." And consciousness itself is only a part of the forces that shape human reasons. For philosophical anthropologists, as for sociologists of knowledge, knowledge is determined by dispositions and by outside factors. Erich Rothacker claims that all knowledge is based on the particular ways of thought (*dogmatische Denkformen*) of national and sectional cultures, which determine both the questions asked and the answers given. Questions and answers have no validity apart from their appropriateness to the cultural environment (*Umwelt*). On the other hand, Scheler sought to establish an objective scale of values that would take into account nonrational elements. He distinguished in an ascending order the strata of vitality, intellectuality, and holiness (*Herrschaftswissen, Leistungswissen,* and *Heilswissen*). Despite his epistemological relativism, Rothacker has applied a similar scheme of "lower" and "higher" values in his psychological theory. Although most philosophical anthropologists profess value relativism, implicit value scales may be discerned underlying their methodological views and cultural criticism.

METHODOLOGY

Philosophical anthropology rejects the Cartesian dualism of body and soul: Man is not part animal and part spirit but a being *sui generis,* distinct from animals in physical condition and in aspirations. This attitude, together with philosophical anthropology's theological roots, may account for a nearly universal (although currently weakening) rejection of Charles Darwin and Sigmund Freud for allegedly appealing to the forces of primitivism and animality in man. At the same time, many philosophical anthropologists reject modern intellectualism; their rejection of rationality, like that of many existentialists and *Lebensphilosophs,* has its roots in the romantic reaction to the Enlightenment and the French Revolution. In its suspicion of *Verwissenschaftlichung* ("scientism"), philosophical anthropology perpetuates the traditional German attacks on *Reflexionsphilosophie,* in which the nonrational aspects of reality are alleged to be ignored.

Philosophical anthropology's conception of method was formulated by Wilhelm Dilthey and Edmund Husserl. Husserl's nonempirical phenomenological approach to philosophical questions was claimed to be presuppositionless, wholly scientific, and logically prior to the natural sciences. It is concerned with meanings, an intuitive comprehension of directly experienced essences, and it involves a distinct method for "analyzing" (or rather, interpreting) facts, qualities, relationships, and the basic categories of human nature and culture—a method of analysis different from that which results in an explanatory theory. However, such thinkers as the biologist Adolf Portmann and the psychologist Karl Jaspers attempt to combine the scientific and interpretative approaches.

Ludwig Binswanger, for example, does not exclude the methods of natural science, but raises two objections to reveal their inherent limitations. One is that all abstractions are transpositions and simplifications of reality. The other is that the registration of stimuli in experimental psychology restricts the field of investigation so as to make the perception of meaningful wholes impossible; it precludes the essential selective and synthesizing activities.

Helmuth Plessner sees philosophical anthropology as the paradigm of borderline research. Although there is

still a methodological gap between the physical and the social sciences, there has been spectacular progress toward methodological and substantive unification of physics, chemistry, and mineralogy, and of physiology and biochemistry. This progress supplies a model for philosophical anthropology. In its physical concerns, philosophical anthropology should correlate the work of medicine, zoology, chemistry, and physics, and in its nonphysical concerns, it should correlate the work of psychology, psychoanalysis, psychiatry, and the cultural sciences.

The physical and nonphysical concerns correspond to the traditional divisions of body and soul and of empiricism and subjective idealism. The division between body and soul emphasizes the ineluctable natural limitations of man and the determined aspects of his nature, and thus ignores his freedom and historicity, while the division between empiricism and subjective idealism has traditionally lost itself in metaphysical speculation. Philosophical anthropology tries to avoid both extremes; it sees man as essentially *homo absconditus,* inscrutable, an open question. Man must formulate his destiny so that he is not held rigidly in one role but safeguards his creative freedom. The direction in which this freedom permits man to fulfill himself is not amenable to scientific discovery, and thus science is devalued. Man's choices depend on his philosophical understanding of his own position in the world.

An infinite variety of choices is open to man. What distinguishes man's nature is not how he chooses, but that he does choose—that he is not determined by his biological and physiological constitution but is formed in the light of cultural values he himself has created and internalized. Philosophical anthropology's contribution to the study of cultures is its emphasis on the creative element in the unfolding of the various conceptions of man's position in the world. Therefore, man's self-understanding, or self-image, is a central theme of philosophical anthropology.

THE SELF-IMAGE OF MAN

Formerly, man was threatened not primarily by man, but by nature. Through science, nearly all natural phenomena have been or can be brought under man's control. Man is threatened neither by nature nor by the God who made nature, but by his own use of nature. Man's enemy is man, manmade structures, or the God who made man.

Again, even in coming to know nature, man (or his scientific representatives) meets himself rather than nature. Man no longer seeks nature as such, but nature as we question it for specific scientific purposes and in the specific contexts of axiomatic frameworks that we ourselves have determined.

Thus, man is inescapably confronted by man. We have reason to ask, What is this man? But what causes us to ask questions about the form in which man's subjective image of himself appears in his consciousness?

Man's subjective image determines what he makes of himself. Animals are as nature has created them, but man must complete his character; nature has supplied only the rudiments of it. Man must form his own personality, and he does so according to his image of what he can and should be. Scheler has delineated a historical typology of Western man's self-images, or "reality-worlds."

Man first saw himself as *homo religiosus,* a view based on the Judeo-Christian legacy of supernaturalism and its ensuing feelings of awe and of inherited guilt. The next stage was *homo sapiens,* rational man in harmony with the divine plan. Since the Enlightenment, this image has been largely superseded by the naturalistic, pragmatic image of *homo faber*—man as the most highly developed animal, the maker of tools (including language), who uses a particularly high proportion of his animal energy in cerebral activities. Body and soul are regarded as a functional unity. Human being and development are explained by the primary urges of animal nature—the desire for progeny and the desire for food, possessions, and wealth. Machiavellianism, Marxism, racism, Darwinism, and Freudianism, it is claimed, are based on this interpretation of man.

These three self-images of man have in common a belief in the unity of human history and in a meaningful evolution toward higher organization. The images of *homo dionysiacus* and *homo creator* break with this tradition and herald a new orientation of anthropological thought. In the image of man as *homo dionysiacus,* man sees decadence as immanent in human nature and history. Typical exponents of this view are Arthur Schopenhauer, Nietzsche, and neoromantics like Ludwig Klages, Oswald Spengler, and Leo Frobenius. Man is seen as the "deserter" or the faux pas of life; as a megalomaniac species of rapacious ape; as an infantile ape with a disorganized system of inner secretions; or as essentially deficient in vital powers and dependent for survival on technical means. Man's power of thought is an artificial surrogate for missing or weak instincts, and his "freedom to choose" is a euphemism for his lack of direction. Human social institutions are pitiful crutches for assuring the survival of a biologically doomed race. Reason is regarded as separate from the soul, which belongs to the

vital sphere of the body. Reason is the destructive, "demoniac" struggle with, and submergence of, the healthy activity of the soul.

The image of man as *homo creator* is likewise derived from Nietzsche, and also from Feuerbach. But the Nietzschean superman has been transformed into a stricter philosophical conception by Nicolai Hartmann, Max Scheler, and the Sartrean existentialists. Scheler called this view a "postulatory atheism of high responsibility." Man has no ontological knowledge of an ultimate being. Contrary to Kant's postulate of the ethical need for a God, in the new view there must be no God—for the sake of human responsibility and liberty. Only in a mechanical, nonteleological world is there the possibility of a free moral being. Where there is a planning, all-powerful God, there is no freedom for man responsibly to work out his destiny. Nietzsche's phrase "God is dead" expresses the ultimate moral responsibility of man; the predicates of God (predestination and Providence) are to be related to individual man.

Man's awareness of his own self-images illuminates the whole range of his genuine potentialities so that his choice of an authentic form of life is not restricted by narrowness of view.

THE MAJOR BRANCHES

Philosophical anthropology shares with French humanism a particular critical analysis of society, but before this analysis can be presented, it is necessary to make a survey of the important branches of philosophical anthropology and of their results.

BIOLOGICAL PHILOSOPHICAL ANTHROPOLOGY.

The reaction to determinism in the physical sciences has given rise to biological philosophical anthropology, or bioanthropology. Bioanthropology scrutinizes biological theories philosophically, primarily to correlate man's creative achievements and attitudes with his physiological organization. Man's cultural role—his character as a symbol-making being capable of abstraction, forethought, language, and intersubjective communication—is depicted as an irreducible function of his physiological constitution.

Among many important practitioners of bioanthropology are the biologists F. J. J. Buytendijk and Adolf Portmann and the philosopher Arnold Gehlen. Important starting points of bioanthropological thought have been Walter Garstang's concept of paidomorphosis and Jakob von Uexküll's concept of milieu (*Umwelt*), which was developed earlier, in philosophical terms, by Edmund Husserl. Paidomorphosis emphasizes the embryonic qualities that are preserved in man but lost in adult animals, as well as man's retarded extrauterine development. Gehlen has used the concept of man as a fetal ape to account for man's cultural achievements which, he claims, are conditioned by man's helpless status in the world. Devoid of instincts and of natural weapons and tools, man has been compelled to compensate for his shortcomings by active responses to the challenges of his environment and of his physiological urges. Man defends himself by his actions, whose scope, direction, and intensity, in contrast to instinctive reactions, are within his discretion. He transforms the natural environment into a system of action (*Handlungskreis*), the responses to which are perpetuated in institutions and language. Man's cultural environment is thus both a physiological condition of his survival and a distinctive criterion of his nature.

Uexküll. From his investigations into animal physiology, Uexküll derived a theory of the specific environmental determination of human life. Each species of animal lives in its own *Umwelt*; its consciousness of sense data is strictly limited by its innate capacities of perception. The range of these capacities corresponds to the teleology immanent in the "life plan" of different animals and is strictly limited to the life plan's specific tasks. Uexküll started from Kant's theory that the categories of the understanding determine the perception and conception of the data of the senses. It was Uexküll's teleological interpretation that distinguished his work from that of Western contemporaries who independently developed the sociology of animals. In the German romantic tradition, Uexküll was concerned with fighting the "mechanistic," positivistic conception of science that he saw represented in biochemistry and behaviorism.

Buytendijk. Buytendijk's physiological and psychological investigations have been undertaken in close contact with such phenomenologically oriented thinkers as Scheler, Plessner, Viktor von Weizsäcker, and V. E. von Gebsattel. Like Uexküll, Buytendijk rejects Cartesian dualism and its mechanical interpretation of bodily processes; unlike Uexküll, he rejects the hypothesis that man is determined by his *Umwelt.* Through his detailed comparisons of animal and human physiology and psychology, Buytendijk has sought to work out man's unique condition as expressed in his capacity for abstraction and symbolization (the ability to create signs representing what is bodily absent), and in his capacity for the logical correlation of signs. For Buytendijk, biology is a historical science that must be understood in motivational, teleological terms. He conceives of motives and processes as

value-related and spontaneous, derived from the built-in planning capacity of a self-structuring organism. Parallels with *verstehende* sociology are obvious, but Buytendijk's impressive ability to rest his philosophical views on a biological basis cannot conceal the fact that he held his views prior to his scientific illustration and testing of them.

Portmann. Adolf Portmann's work represents the culmination of bioanthropology. It aims at an integration of biological with psychological, sociological, and anthropological thought. According to Portmann, human biology has turned into anthropology, because the life of man, despite superficial similarities to animal life, is something sui generis. Portmann emphasizes the uniqueness of human action, language, foresight, and upright carriage, and of the human growth rhythm—duration of pregnancy, bodily proportions, extrauterine babyhood, and late formation of the female pelvis. These qualities, he claims, arise from a characteristic interpenetration of the hereditary process and teleological, sociocultural processes. Man's individuality (which continues to grow while the body decays) and man's sociability combine to establish his undetermined "openness," in contrast to determination of the animal by his *Umwelt.*

Portmann's central concept is "internality," the fact that individuals are centers of purposeful activity who use the external shell of the body as a means of self-expression and of communication with other individuals. Portmann does not claim that the affirmation of man's individuality and sociability provides the "meaning of life." Although specific mysteries of man's biological structure have been solved, he claims, the "basic fact" for philosophical anthropology continues to be man's "mysteriousness." Man has no built-in evolutionary mechanism leading to an equilibrium; there is only a creative variability (*Disponibilität*) of the human situation. Man's spontaneous individuality creates new self-images; his sociability spreads and maintains them.

Portmann has sought, however, to advance beyond the limits of functional morphology to a vantage point that will illuminate the hierarchy of values—a vantage point whose need has increased in view of the tremendous potential power of biotechnical advances to influence and change the human condition, and perhaps human nature. However, as in the biophilosophies of Henri Bergson, Pierre Teilhard de Chardin, and Julian Huxley, it is easier to discern the philosophical basis of Portmann's biological hypotheses than it is to discover any positive contribution that biology has made to his philosophical thought. He first developed his conception of man as functional unit as a philosophical hypothesis. "Openness" has been a theme of philosophical anthropology since the time of Herder and Kant.

In general, it must be said that no substantive lesson is to be drawn from either functional or analytical biology, except that it is of man's essence to create structured and meaningful systems of action. The biological foundation of man's creativity entails no concrete guide to what man ought to do. Nothing would appear to follow from the fact that creativity has biological roots except that man cannot permit himself to be altogether determined by any given environment. He must transcend it creatively, and he must be guided by ideas and leitmotifs rather than by instincts, by decisions rather than by reactions to stimuli. But the questions of what decisions man will take and what ideas he will adopt are not answered by bioanthropology, which emphasizes the malleability of human nature as a basic fact. Any insight into the potential content of human achievement must therefore be based on the plurality of the cultures that have unfolded in history. Bioanthropology thus leads into cultural philosophical anthropology.

CULTURAL PHILOSOPHICAL ANTHROPOLOGY. Like American cultural anthropology, cultural philosophical anthropology is concerned with man and his works, with culture history and culture sociology, and with historical morphology and the philosophy of history. It is interested primarily in developed societies—"high cultures" that have created a style of their own beyond the biological and trivial uniformities of the tribal state. Like German sociology, it emphasizes the multiformity rather than the uniformity of human nature, and the history rather than the theory of cultures. Like Portmann's bioanthropology, it finds an ultimate mystery in man—the mystery of archetypes and racial dispositions.

Cultural anthropology combines Dilthey's historicism with the phenomenological method. Man comes to know and liberate himself through history. A comparative study of societies elucidates the human situation and the human predicament. But this study results in the same merely formal characteristics elaborated by bioanthropology—the adaptability of the human mind, the need for a "sane" worldview, sociability with its ensuing problems, a common growth rhythm, and common basic physiological urges.

Arnold Gehlen and Erich Rothacker are the most representative cultural philosophical anthropologists, while Werner Sombart is the most opinionated. Gehlen and Rothacker present integrated theoretical systems that have an ultimately psychological basis. Their psycholo-

gies, like that of Dilthey, are essentially descriptive and interpretative, and their psychological interpretations mirror their cultural philosophies.

Rothacker has classified cultural factors in a scale by "laws of polarity." He seeks to understand individual cultures by a process of "reduction" to "national souls" (attitudes and dispositions that generate *Weltanschauungen*) and to myths. These ur-experiences are not further reducible; they are embodied in the racial inheritance. Therefore, although people do create and develop the *Umwelt* of their national cultures, the possibilities that are thereby realized are ultimately determined. Rothacker's historicist relativism is less free from ethnocentrism than one might be led to expect by the emphasis of philosophical anthropology on the openness of man.

Gehlen's psychology is rooted in the archaic stage of cultural development. The values of this stage serve as criteria for the evaluation of late cultures, which accordingly appear as falls from grace.

In Sombart's anthropology ethnocentric traits are also emphasized. Thus, man's irreconcilable diversity rather than his potential openness is seen as distinguishing the human situation.

Ernst Cassirer, on the other hand, sought to discover the basic function of human cultural achievements (language, myth, religion, art, science, history) behind their innumerable forms and to trace them to a common origin in man's symbol-making power—the power to build up an "ideal world" of his own.

PSYCHOLOGICAL PHILOSOPHICAL ANTHROPOLOGY. Bioanthropology and cultural philosophical anthropology are the most important branches of philosophical anthropology. Among other branches, only psychological philosophical anthropology and theological philosophical anthropology require separate mention.

Psychological philosophical anthropology is the most successful post-Freudian development in psychiatry on the Continent and, through existential psychoanalysis, is exerting considerable influence in the English-speaking world. The outstanding figures in this movement are Ludwig Binswanger, Erwin Straus, and Medard Boss. Erich Fromm seeks to incorporate his psychology within philosophical anthropology, and Rollo May in the United States and R. D. Laing in Britain follow similar lines. Their common belief is that traditional experimental psychology requires the assistance of philosophical thought to arrive at satisfactory results. Some psychological philosophical anthropologists oppose the empirical hypotheses and inductive statistical methods of experimental psychology; most of them combine experimental methods with a specific philosophical or phenomenological approach.

Since psychological philosophical anthropology deals with individual cases, it lends itself to concrete and descriptive investigations. Analyses have been made of laughter and weeping, fantasy, shame, resentment, pleasure, love, and fear. These analyses do not consist in mere registration of stimuli but in selective and synthesizing acts of interpretation by phenomenological "reduction" to an intuition of essential qualities. Plessner has traced the capacity for laughter and weeping to man's "excentricity," his ability to transcend his innate nature and to observe, judge, and respond to situations. Human moods (*Stimmungen*) are typically described as obstacles to the achievement of authenticity. The irrational elements in moods undermine the continuity of character, which is man's potential ability to give meaning and direction to his life. Accidental attitudes that arise from the challenge of situations thus deprive man of his right to make responsible choices; they tie him to an impoverished, one-sided anthropology.

Binswanger developed existential analysis from Freud's psychoanalysis. He describes Freud's positivist, "utilitarian" anthropology as one-sided and negative. Its culture concept, he claims, concentrates negatively on the taming of natural urges rather than positively on a teleological image of man's potentiality. Freud's "somatographic" or "somatomorphic" conception of existence stresses the scientific analysis of sleep, dream, passion, and sensuality while, according to Binswanger, it neglects the historical and cultural aspects of existence, such as religion, art, ethics, and myth, all of which are as important as science. In Binswanger's view psychological investigation should be directed toward the self-transcending, exercise of man's liberty to make authentic choices. The psychologist's task is to illuminate the "inner life history" of the patient, his self-structuring in the light of his inner motivation. Self-structuring is equivalent to character or to the response that the individual makes to the challenge of the world around him. St. Augustine, to whom we owe the beginnings of autobiography, is a case in point. Illness prevented him from carrying out his ambition to become an orator. He transcended his natural disability by turning toward the spiritual world and thus arrived at his essential "real being." He could have reacted otherwise—by resentment or frustration, by neurosis or suicide. These and other potentialities held out to Augustine the temptation to restrict his character by the impoverishment inherent in giving in to an irresponsible choice—a

choice suggested by the logic of the situation. Augustine chose an autonomous life that preserved his access to a full range of human values. Psychosis is explained as an "abortive encounter" with existence, or a form of existential misdirection. Diagnosis of a psychosis therefore depends on a valid interpretation of what constitutes an authentic existence. An authentic existence, according to Binswanger, consists in a life in keeping with a legitimate cultural (religious or national) tradition; in a dialogue with other beings (the "Thou"); or in the ability to act in character in the face of situational challenges.

However, the first of these criteria depends on values that are subject to unresolved doubt; the other two are so devoid of specific content that they hardly invite contradiction. Existential analysis, even more than psychoanalysis, obliterates the line between the normal and the abnormal and reduces psychological problems to questions of *Weltanschauung.*

Viktor von Weizsäcker, V. E. von Gebsattel, Erwin Straus, and Harald Schultz-Hencke have carried out structural analyses of inhibited character types and, in particular, of sexual perversions. Health is defined as openness to all potentialities of life, and obsessional urges are therefore interpreted as disturbed worldviews that enslave the individual in rigid, one-sided, compulsive attitudes and interfere with his social "I-Thou" relationships. Sexual perversions, in particular, have been construed by Gebsattel as obsessional urges that preclude a lasting I-Thou relationship based on mutual freedom, and as thus being incapable of providing ultimate satisfaction. Medard Boss, however, arguing from an equally existential basis, stoutly rejects this view. Gebsattel's apotheosis of the procreative element in love, however, points to the close affinity of philosophical anthropology with "secularized theology."

THEOLOGICAL PHILOSOPHICAL ANTHROPOLOGY. Theological anthropology emphasizes the Biblical conception of man in a dialogue with God. Martin Buber, Emil Brunner, and Dietrich Bonhoeffer are remarkable representatives of this movement, although their work is best studied in its theological context. However, the openness of man and his individuality and sociability are dominant themes of their work. The human difficulty of making the right choices is paralleled by the theological conception of man as simultaneously just and sinning.

A merely intellectual and logical exposition of God's message, in their view, is not enough for an understanding of God's revelation. An emphatic existential I-Thou relationship between man and God, based on the *logique du coeur,* is required. What matters is not that something is true, but how it can be made to come true. Belief in God has been explained by theological philosophical anthropologists, following Feuerbach, in terms of the self-understanding and the creative self-image of man.

The need for a postscientific interpretation of the Creed that is appropriate to a "mature" humanity and avoids theological sophistry has become a leading motif of theological anthropology, and this makes it difficult to distinguish between its tenets and those of secular philosophical anthropologies.

CRITIQUE OF SOCIETY

Philosophical anthropology shares with contemporary French humanism the conception that there is a crisis of the sciences that reflects a radical crisis of European society. It rejects contemporary bourgeois society, from either a romantic or a Marxist viewpoint, for the alleged dehumanizing tendencies it has developed in the process of rationalization following the breakup of feudal and religious institutions.

The rise of scientific rationalism is not regarded as a process of liberation from the shackles of superstition, conventions, and fallacies, but as a process that has deprived Western man of his "center of gravity" and has alienated him from his authentic nature through the replacement of value by "means-end" relationships, by neutral experiment, and by mechanico-mathematical abstraction. In the view of philosophical anthropologists, the "age of transition," or "age of crisis," which heralded the acceptance of utilitarianism in the English-speaking world, is still unresolved. Man's salvation from alienation is not seen as a continuous process of improvement or of piecemeal social engineering but as a radical challenge that is less concerned with practical reform than with a utopian rejection of the modern world.

The central theoretical insights of philosophical anthropology consist in an affirmation of the individuality and sociability of man as ultimate values. This theory would seem to suggest a social organization that combines an optimum of free choice with the minimum encroachment on individual liberty that is compatible with a viable social coexistence. This is in fact the utilitarian image of man that has prevailed since the early nineteenth century in the English-speaking world, where this image of man has been internalized to such an extent that the discussion of ultimate metaphysical questions has predominantly given way to the discussion of means to assure the accepted end of mutual accommodation and individual discretion. By contrast, on the Continent, and

especially in Germany, the romantic reaction to the French Revolution precluded the acceptance of the philosophy of the Enlightenment. No commonly accepted concept of society was developed to counterbalance an unbridled individualism except the radical panaceas of nationalism and totalitarianism. By emphasizing the importance of both individuality and sociability, philosophical anthropology is returning to the type of position that gave birth to utilitarianism, and it may therefore be a step toward a utilitarian view of the world. Although most of its representatives present ethnocentric or nihilistic conclusions, these are not inevitable consequences of philosophical anthropology's affirmation of the creativity and sociability of man.

(See "Philosophical Anthropology" in the index for articles on philosophers who have especially concerned themselves with the topic.)

See also Alembert, Jean Le Rond d'; Augustine, St.; Bergson, Henri; Binswanger, Ludwig; Bonhoeffer, Dietrich; Brunner, Emil; Buber, Martin; Cartesianism; Cassirer, Ernst; Coleridge, Samuel Taylor; Darwinism; Determinism, A Historical Survey; Diderot, Denis; Dilthey, Wilhelm; Encyclopédie; Enlightenment; Existentialism; Ferguson, Adam; Feuerbach, Ludwig Andreas; Freud, Sigmund; Garve, Christian; Gehlen, Arnold; Geisteswissenschaften; Goethe, Johann Wolfgang von; Hartmann, Nicolai; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Herder, Johann Gottfried; Historicism; Hobbes, Thomas; Humanism; Humboldt, Wilhelm von; Hume, David; Husserl, Edmund; Hutcheson, Francis; Idealism; Jaspers, Karl; Kant, Immanuel; Kierkegaard, Søren Aabye; Klages, Ludwig; Locke, John; Marxist Philosophy; Marx, Karl; Maupertuis, Pierre-Louis Moreau de; Mill, John Stuart; Newman, John Henry; Nietzsche, Friedrich; Pascal, Blaise; Plessner, Helmut; Racism; Scheler, Max; Schopenhauer, Arthur; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Smith, Adam; Sombart, Werner; Spengler, Oswald; Stewart, Dugald; Teilhard de Chardin, Pierre; Utopias and Utopianism.

Bibliography

GENERAL AND HISTORICAL

Bryson, Gladys. *Man and Society.* Princeton, NJ: Princeton University Press, 1945.

Groethuysen, Bernhard. *Philosophische Anthropologie: Handbuch der Philosophie,* Part III, 1–207. Munich and Berlin, 1928.

Häberlin, Paul. *Der Mensch: Eine philosophische Anthropologie.* Zürich: Schweizer Spiegel, 1941.

Landmann, Michael. *De Homine,* 543–614. Freiburg and Munich, 1962. A comprehensive bibliography.

Landmann, Michael. *Philosophische Anthropologie.* Berlin, 1955.

Landsberg, P. L. *Einleitung in die philosophische Anthropologie.* Frankfurt, 1949.

Lipps, Hans. *Die Wirklichkeit des Menschen.* Frankfurt, 1954.

Litt, Theodor. *Die Selbsterkenntnis des Menschen.* Leipzig, 1938.

Löwith, Karl. *Das Individuum in der Rolle des Mitmenschen.* Munich, 1928.

Pappé, H. O. "On Philosophical Anthropology." *Australasian Journal of Philosophy* 39 (1961): 47–64.

Pfänder, Alexander. *Die Seele des Menschen.* Halle, 1933.

BIOLOGICAL PHILOSOPHICAL ANTHROPOLOGY

Buytendijk, F. J. J. *Allgemeine Theorie der menschlichen Haltung und Bewegung.* Berlin, 1956.

Buytendijk, F. J. J. *Mensch und Tier.* Hamburg, 1958.

Portmann, Adolf. *Die Biologie und das neue Menschenbild.* Bern, 1942.

Portmann, Adolf. *Zoologie und das neue Bild des Menschen.* Hamburg, 1956.

Pringle, J. W. S. *The Two Biologies.* Oxford: Clarendon Press, 1963.

PSYCHOLOGICAL PHILOSOPHICAL ANTHROPOLOGY

Bally, Gustav. *Der normale Mensch.* Zürich, 1952.

Boss, Medard. *Psychoanalyse und Daseinsanalyse.* Bern, 1957. Translated by Ludwig B. Lefebre as *Psychoanalysis and Daseins-analysis.* New York and London, 1963.

Gebsattel, V. E. von. *Prolegomena einer medizinischen Anthropologie.* Berlin: Springer, 1954.

Laing, R. D. *The Divided Self: A Study of Sanity and Madness.* London: Tavistock, 1960.

May, Rollo, ed. *Existence.* New York, 1958.

Schultz-Hencke, Harald. *Der gehemmte Mensch.* Leipzig: G. Thieme, 1940.

Strasser, Stefan. "Phenomenological Trends in European Psychology." *Philosophy and Phenomenological Research* 18 (1957): 18–34.

Weizsäcker, Viktor von. *Der kranke Mensch: Eine Einführung in die medizinische Anthropologie.* Stuttgart, 1951.

THEOLOGICAL PHILOSOPHICAL ANTHROPOLOGY

Kuhlmann, Gerhardt. *Theologische Anthropologie im Abriss.* Tübingen, 1935.

Loewenich, Walther von. *Menschsein und Christsein bei Augustin, Luther und J. Burckhardt.* Gütersloh, Germany, 1948.

Michel, Ernst. *Ehe: Eine Anthropologie der Geschlechtsgemeinschaft.* Stuttgart, 1948.

Rosenstock-Huessy, Eugen. *The Christian Future.* New York: Scribners, 1946.

H. O. Pappé (1967)

PHILOSOPHY

Defining *philosophy* is itself a philosophical problem. Perhaps a great many philosophers would agree that whatever else philosophy is, it is the critical, normally systematic study of an unlimited range of ideas and issues. But this characterization says nothing about what sorts of ideas or issues are important in philosophy or about its distinctive methods of studying them. Doing this will require some account of the special fields of the subject, its methods, its connections with other disciplines, its place in the academy, and its role in human culture. The task is large. Philosophy pursues questions in every dimension of human life, and its techniques apply to problems in any field of study or endeavor. It may be described in many ways. It is a reasoned pursuit of fundamental truths, a quest for understanding, a study of principles of conduct. It seeks to establish standards of evidence, to provide rational methods of resolving conflicts, and to create techniques for evaluating ideas and arguments. Philosophy may examine concepts and views drawn from science, art, religion, politics, or any other realm.

The best way to clarify these broad characterizations of philosophy is to describe its principal subfields (all of which are addressed in more detail in entries in this Encyclopedia devoted to them alone). It is appropriate to start with what might be called *traditional* subfields of philosophy, most commonly taken to be epistemology, ethics, logic, metaphysics, and the history of philosophy. These remain central in philosophical research; and although they are by no means its exclusive focus, they are intimately connected with virtually every other field of philosophical research and are widely treated as core areas in the teaching of the subject.

FIVE TRADITIONALLY CENTRAL SUBFIELDS OF PHILOSOPHY

EPISTEMOLOGY. Epistemology concerns the nature and scope of knowledge and justification. What does it mean to know (the truth), and what is the nature of truth? What sorts of things can be known, and can we be justified in our beliefs about what goes beyond the evidence of our senses, such as the inner lives of others or events of the distant past? Is there knowledge beyond the reach of science? What are the limits of self-knowledge? Can there be genuine moral knowledge? Quite apart from the depth, modality, or subject matter of knowledge, we may also ask: What are its basic sources? They have been widely thought to be perception, memory, introspection, and reason (understood as a kind of reflection). But what of testimony? And can any substantive knowledge, say in mathematics, be utterly independent of experience in the way a priori (reason-based) knowledge is sometimes held to be?

A major epistemological problem connected with all of these sources is the status of skepticism. Skepticism has many forms, depending on the kind of knowledge or justification it represents as unattainable. What is commonly called Humean skepticism (deriving from David Hume's writings on causation and inductive inference) challenges the belief that any inductive arguments (*probable arguments*, in Hume's terminology) can ground knowledge. Cartesian skepticism, powerfully stated in Descartes's *Meditations*, challenges the belief that we have knowledge at all. Quite apart from whether there can be knowledge or justified belief, there is the question of the structure that a body of knowledge or of justified beliefs must have. Must it, for instance, contain beliefs possessing a kind of axiomatic status, or can it consist of elements that all lack that status or, indeed, are in no way privileged relative to other elements? Traditional *foundationalists*, such as Descartes, have held a view of the first kind; moderate foundationalists (represented by a large proportion of epistemologists since the middle of the twentieth century) hold that foundational cognitions are necessary in a body of knowledge or justified belief but need only be in a certain way noninferentially justified as opposed to indefeasibly justified; and coherentists and other nonfoundationalists have posited various ways aimed at accounting for knowledge and justification without appeal to foundational elements.

ETHICS. Ethics is the philosophical study of morality, particularly conceived as a set of standards of right and wrong conduct. Its most theoretical branch (commonly called *metaethics*) concerns the meanings or, more broadly, the logic, of our moral concepts—such as right action, obligation, and justice—the kinds of evidence we have for propositions about the corresponding subject matter, and the sorts of properties that apparently underlie the application of the concepts. On some major ethical views, such as J. S. Mill's utilitarianism, our obligations derive from our potential contributions to enhancing what is good. For this reason, among others, the concept of the good and the distinction between intrinsic and instrumental goodness are also major concerns of ethical inquiry. On other major ethical views, such as Immanuel Kant's, moral obligatoriness is a property possessed by acts themselves by virtue of their falling under nonconsequentialist principles, for instance, a

principle that, quite apart from the consequences of lying, prohibits it.

Normative ethics is commonly contrasted with metaethics and is concerned to formulate and assess principles meant to guide moral decisions, whether in private or public life. A major question it raises is what moral specific obligations we have. Another is what moral rights persons as such have and, related to this, what legal rights a just society must accord its citizens. Still another is what constitutes a valid excuse for wrongdoing. Any moral philosopher may be concerned with the broad question of how moral disagreements may be rationally settled, and here we have a question that has both metaethical and normative aspects.

LOGIC. Logic is concerned to provide sound methods for distinguishing valid from invalid arguments or, on a wider conception, good from bad arguments in terms of criteria for determining how much support the conclusion receives from the premise(s). Arguments may be considered ordered sequences of propositions in which some—the premise(s)—are conceived as supporting another—the conclusion. A standard example is the following syllogism, which has a very common form: its premises are that *all human beings are mortal* and that *Socrates is mortal*; its conclusion is that *Socrates is a human being*. Deductive logic is concerned with appraising arguments in relation to the question whether the premises *entail* (or logically imply) the conclusion, as with the syllogism just presented. Inductive logic is concerned with appraising arguments in relation to probabilistic support. From premises about the factors that cause influenza, medical experts may conclude that millions of people will be infected during the next flu season. Inductive logic addresses the problem of how we may tell what probability this conclusion has given those premises. More generally, logic helps us to assess how well our premises support our conclusions, to see what we are committed to accepting when we hold a view, and to avoid adopting positions for which we lack supporting reasons. As applied to everyday thinking, the use of logic also helps us to find arguments where we might otherwise simply see a set of loosely related statements, to discover assumptions we did not know we were making, and to formulate the minimum claims we must establish if we are to prove (or inductively support) our point.

METAPHYSICS. Metaphysics seeks basic criteria for determining what sorts of things are real. Criteria of this kind are the special concern of ontology, which is central in metaphysics. Among major ontological questions are these: Are there mental, physical, and abstract things (such as numbers)? Or is there just the physical and the spiritual? Might there be merely matter and energy? Are persons highly complex physical systems, or do they have properties not reducible to anything physical? How much can a person—or other kind of thing—change and remain the very person or thing it is? In the case of persons, this question is central for the problem of personal identity, which, in turn, is crucial for understanding the possibility of nonembodied life. Another question about persons is whether they can be free in a sense not possible for lower animals and whether their freedom is possible if the world should be a deterministic system, that is, one in which every event is entailed by a universal law of nature and some simultaneous or antecedent event. What constitutes a law of nature, and, in particular, what constitutes a causal law, are themselves major questions in metaphysics. Metaphysics has also been traditionally taken to include cosmology, which is concerned with the nature of the universe as a whole and pursues such questions as whether it must have a beginning in time, whether it can be infinite, and whether it must have been created and, if so, by what kind of being or in what way. The nature of time is itself an important metaphysical question.

HISTORY OF PHILOSOPHY. The history of philosophy might be thought to be a branch of the discipline of history rather than of philosophy, much in the way the history of science is a branch of history and not itself a branch of science. This conception would be quite inadequate to the standard conception of the history of philosophy in the field of philosophy. On that conception the history of philosophy is a genuine subfield of philosophy: It is the historical and *philosophical* study of the history of the subject. It commonly includes more in the way of philosophical interpretation and—sometimes—philosophical appraisal of major texts than historiographic studies of either a single philosopher or whole periods in the history of the subject. This is in part because the interpretation—and certainly the proper appraisal—of a philosopher is itself a philosophical problem, often involving epistemological or metaphysical theorizing. A study of a single philosophical work important in the history of philosophy may thus count as a contribution to the history of philosophy and not just to the study of its author.

The history of philosophy, then, examines major philosophers, the influence of one philosopher on another (say, Aristotle on Aquinas, Husserl on Heidegger, or Frege on Russell) or entire periods in the development

of philosophy, such as the Ancient, Medieval, Modern, Nineteenth Century, and Twentieth Century periods. It seeks to understand great figures, their influence on others, and their importance for perennial and contemporary issues. The history of philosophy in a single nation is often separately studied, as in the case of American Philosophy. So are major movements within a nation, such as German Idealism, as well as international movements with a substantial history, such as Existentialism, Logical Positivism, and Phenomenology.

From the wide scope of many of the questions pursued in these philosophical fields, it should be clear that philosophy has a kind of generality possessed by no other field. Metaphysics, for instance, concerns the basic categories encompassing *everything* that exists, and epistemology concerns standards of evidence that apply in *any* kind of thinking. It will also be evident that every other discipline presupposes answers to certain philosophical questions. All of the sciences, for example, presuppose that facts about the past can yield knowledge or justified beliefs about the future. Finally, it should be apparent that, although there are distinctively philosophical questions, no subject matter is (in all its aspects) beyond the reach of philosophical inquiry. Any subject matter can raise philosophical questions: about (for instance) the kinds of entities it concerns, its epistemological presuppositions, and its connection with other subjects.

OTHER MAJOR SUBFIELDS OF PHILOSOPHY

Many branches of philosophy have grown from the traditional core areas just described. What follows is a sketch of a number of the major ones. Comprehensiveness is not possible here, but a wider conception can be formed by reading the entries devoted to the subfields that will be described.

PHILOSOPHY OF MIND. This subfield has emerged largely from metaphysical concerns with mental phenomena. The philosophy of mind addresses not only the possible relations of the mental to the physical (for instance, to brain processes) but to the many concepts having an essential mental element: belief, desire, intention, emotion, feeling, sensation, passion, will, personality, and others. To what extent are any of these concepts explicable in terms of behavioral tendencies? Quite apart from that, what is the relation between mental properties and physical ones? Are the former dependent on the latter, and if so, what kind of dependence is in question? Could two biological beings, for instance, be alike in all their physical properties and still differ in their mental ones? A number of major questions in the philosophy of mind cluster in the area of *action theory*: What differentiates actions, such as raising an arm, from mere body movements, such as the rising of an arm? A common answer has been that actions but not bodily movements must be caused by such mental events as volitions. But must mental elements, such as intentions, beliefs, and emotions enter into adequate explanations of our actions, or can actions be explained by appeal to ordinary physical events? And is a kind of mental causation, or at least the absence of a certain kind of deterministic causation, required for our actions to be *free*?

PHILOSOPHY OF RELIGION. Another traditional concern of metaphysics is to understand the concept of God, including special attributes such as being all-knowing (omniscient), all-powerful (omnipotent), and wholly good (omnibenevolent). Does omnipotence, for instance, entail the ability to alter the laws of logic? Both metaphysics and epistemology have been concerned to assess the various grounds offered to justify one or another form of theism (these include the famous cosmological and ontological arguments, among others treated in this encyclopedia). The philosophy of religion—also called *philosophical theology*—systematically examines these topics and many related subjects, such as the relation between faith and reason, the nature of religious language, the relation of religion and morality, and the question of how a God who is wholly good could allow the kind and amount of evil the world apparently contains. Here the philosophy of religion overlaps the theory of value, a branch of ethics. It is common for a major question to cross philosophical fields in this way, and the same holds for the relation between theology and ethics, for instance in relation to the question whether the rightness of actions could be equivalent to divine commandedness.

PHILOSOPHY OF SCIENCE. This is probably the largest subfield, generated in substantial part by epistemology and in part by metaphysics. Philosophy of science has been commonly divided into philosophy of the natural sciences and philosophy of the social sciences. It has recently been divided further, into philosophy of physics, of biology, of psychology, of economics, and of other sciences. Philosophy of science clarifies both the quest for scientific knowledge and the results yielded by that quest. It does this by exploring the logic of scientific evidence; the nature of scientific laws, explanations, and theories; the nature of the theoretical entities posited in explaining observable phenomena; and the possible connections

among the various branches of science. How, for instance, is psychology related to brain biology, and biology to chemistry? And how are the social sciences related to the natural sciences? Are they methodologically like the latter but incapable of discovering universal as opposed to statistical laws? Must they work with mentalistic concepts such as belief and desire? Does explanation have the same form across the several sciences?

SUBFIELDS OF ETHICS. From ethics, too, have come major subfields. *Political philosophy* concerns the justification—and limits—of governmental control of individuals; the meaning of equality before the law; the basis of economic freedom; and many other problems concerning government. It also examines the nature and possible arguments for various competing forms of political organization, such as laissez-faire capitalism, welfare democracy (capitalistic and socialistic), anarchism, communism, and fascism. *Social philosophy*, often taught in combination with political philosophy (which it overlaps), treats moral problems with large-scale social dimensions. Among these are the ethics of journalism and the media, the basis of compulsory education, the possible grounds for preferential treatment of minorities, the justice of taxation, and the appropriate limits, if any, on free expression in the arts. The *philosophy of law* explores such topics as what law is, what kinds of laws there are—for instance, only positive (enacted) law or also, as Thomas Aquinas held, natural law—and how law is or should be related to morality. It also examines the sorts of principles that should govern punishment and criminal justice in general (ethical questions about law do not exhaust the philosophical questions about it but have been among those central in the philosophy of law). *Medical ethics* addresses many problems arising in medical practice and medical science. Among these are standards applying to physician–patient relationships; moral questions raised by special procedures, such as abortion and ceasing of life-support for terminal patients; and ethical standards for medical research, for instance, genetic engineering and experimentation using human subjects. *Business ethics* addresses such questions as the place of business in society, how moral obligations may conflict with the profit motive, and how these conflicts may be resolved. Other topics often pursued are the nature and scope of the social responsibilities of corporations, their rights in a free society, and their relations to other kinds of organizations.

PHILOSOPHY OF ART (AESTHETICS). This is one of the oldest subfields. It concerns the nature of art, including both the performing arts and literature, painting, and sculpture. Major questions in aesthetics include how artistic creations are to be interpreted and evaluated and how the arts are related to one another, to natural beauty, and to morality, religion, science, and other important elements of human life. Aesthetics also deals with epistemological questions concerning the kinds of evidence we can have about an artwork and—sometimes—the kinds it can give us about the world, particularly about human beings. There is also a metaphysics of the aesthetic: What kind of property is beauty in a painting, power in a symphony, or unity in a poem, and is a poem a physical entity existing where it is written or remembered, or is it something more abstract of which these mental and physical entities are in some sense vehicles?

PHILOSOPHY OF LANGUAGE. This field has close ties to both epistemology and metaphysics and, in the latter connection, to the philosophy of mind. It treats a broad spectrum of questions about language: the nature of meaning, the relations between words and things, the various theories of language learning, and the distinction between literal and figurative uses of language. A major concern in the field is the theory of reference: What, for instance, is required for us to succeed in referring to Socrates by using that name when we have never met him nor even read anything written by him? And if our thoughts are mental and *in* the mind, how can their content be about external objects? A question connected with all of these problems is the relation between the linguistic and the conceptual. To what extent, for instance, is it possible to have concepts at all without linguistic terms to express them, and is thought itself possible apart from language? Since language is crucial in nearly all human activity, the philosophy of language bears on our understanding both of other academic fields and of much of what we ordinarily do.

OTHER IMPORTANT SUBFIELDS. There are many other subfields of philosophy, and it is in the nature of philosophy as critical inquiry to develop new subfields when new directions in the quest for knowledge, or in any other area of human activity, raises new intellectual problems. There is no limit to the number of variety of possible subfields of philosophy. Among the subfields not yet mentioned, but often a focus or research or teaching (at least as a part of other courses), are Philosophy of Logic, Philosophy of History, Philosophy of Mathematics, Philosophy of Medicine, Philosophy of Education, Philosophy of Feminism, Philosophy of Linguistics, Philosophy

of Criticism, Philosophy of Culture, Philosophy of Film, and Philosophy of Sport.

PHILOSOPHICAL METHODS

The *Dialogues of Plato* made famous what might be called the *Socratic method* in philosophy. It is the dialectical method, pursued by Socrates as represented by Plato in the *Dialogues*, in which ideas are set out, explored in relation to their meaning and implications, and assessed by such criteria as consistency and plausibility in relation to various standards, sometimes including common sense. In both Plato and Aristotle, we find early examples of what may plausibly be called *conceptual analysis*. Aristotle provides a particularly good example of how this may be conceived. In his *Nicomachean Ethics*, for instance, he seeks to give an account of the concept (or anyway of *a* concept) of virtue. He saw himself as clarifying the essence of the phenomenon of virtue; but if this essentialist view is understood in terms of his philosophical practice, it seems consistent with construing some of what he did as a kind of conceptual analysis. He is guided by the use of the relevant Greek terms in what we may suppose was educated parlance; yet he is not talking merely about linguistic usage. This is not to assimilate his kind of conceptual analysis to a Platonic kind on which concepts are to be understood by intellectual apprehension of them as abstract entities accessible to reflection. Indeed, if there are times when his analytic technique recalls Plato, there are others when his attention to usage and to what is said brings to mind some moments in the later work of Ludwig Wittgenstein.

A major question here, on which there is persisting difference of judgment among philosophers, is the extent to which these intellective procedures (whether Aristotelian or Platonic) are genuinely different from *linguistic analysis*. A related question is the degree of the authority of linguistic usage in determining the content of a concept. As important as dialectical method and conceptual analysis are in philosophy, however, neither can be described as *the* method of philosophy. It may be that every major philosopher has used at least one of them at some point; but even supposing (what is certainly controversial) that philosophy cannot be competently pursued on a large scale without some measure of at least the latter, there are other methods of inquiry that should be considered philosophical.

An important route to understanding philosophy and, especially, philosophical method, is a comparison of philosophical method with scientific method. From at least the middle of the twentieth century, and in at least much of the Western philosophical tradition, there has been a (sometimes tacit) belief in scientific method as the paradigm of an objective, rational method of seeking truth. There has been an associated belief, or presupposition, that philosophy must, in methodology as well as doctrine, take account of the progress of science. This is not to say that the (or a) method of science, or some interpretation of scientific method, has become the dominant *philosophical* method. But there is a widely held assumption—which we might call the *assumption of the philosophical primacy of scientific method*—that scientific method is the primary model of the rational pursuit of truth, in a sense implying both that our philosophical method, if not itself scientific, should bear an appropriate resemblance to scientific method and that our philosophical results are probably mistaken if they are at odds with, or even unable to account for the possibility of, well-established scientific findings. It will help to describe this primacy assumption in the three major areas of concern in this entry: epistemology, metaphysics, and methodology.

EPISTEMOLOGY. Where scientific method currently has the primacy that has been mentioned, then, first of all, we might expect the assumption of its primacy to have an antirationalist thrust. For despite the rationalist point that a priori truths do not compete with scientific statements in explanation or theorizing, such truths are also traditionally conceived as beyond refutation by scientific procedures and as knowable by a nonscientific method (a kind of reflection). The second point is positive: The influence of scientific method as a model of rational belief formation has given impetus to the view that much of what we know is discovered by inference to the best explanation (a kind of inductive inference), and much of what we understand is understood in terms of underlying theoretical states or entities. Thus, even self-knowledge can be taken to be not only constituted by corrigible belief (roughly, belief whose justification can be defeated) but, often at least, to comprise beliefs arrived at by unconscious (or at least unnoticed) inference from appropriate data. The fallibilism that comes with a deep appreciation of scientific method has similar implications in other areas of apparent human knowledge.

METAPHYSICS. In metaphysics, the assumption of the philosophical primacy of scientific method implies a tendency to take science as the arbiter of the real. The obvious point here is that we should tend to countenance as real whatever our best confirmed scientific theories posit as such, or at least posit as explanatorily basic. (Granted,

it is not always clear what this is even if we can decide what our best-confirmed theories are). But there is a further implication. We must also countenance as real whatever must be posited to understand science itself, for instance properties, numbers, or sets. And, in part on the basis of assuming Occam's Razor (roughly, the principle that in providing explanations we should not posit more entities or types of entities than necessary), many philosophers think we need countenance nothing else.

One good illustration of the point here is the effort to support realism in ethics by arguing (against both noncognitivist and epiphenomenalist views in metaethics) that moral properties have causal and explanatory power and hence can play an explanatory role substantially similar to the role of theoretical entities in the sciences. Moral realists need not be causalists, however; they all agree in holding the cognitivist metaethical view that moral claims have truth value (hence are true or false), but rationalists among them may deny that moral properties—even if in some way grounded in nonmoral properties, such as lying, beating, and killing, that have causal power—are themselves causal properties. Most philosophers would grant, however, that whether or not genuine properties *must* have causal power, whatever does have that power is real.

METHODOLOGY. If what has been said about the metaphysical implications of the assumption of the primacy of scientific method is correct, it should be easy to understand some of the methodological implications for philosophy. For in a way, the second metaphysical implication is methodological: Its basis is largely a commitment to scientific method as so well established, and so near to being self-evidently essential in the search for truth, that we should countenance whatever realities must be posited to account for its success and need *not* countenance any others. A further methodological implication is a tendency to solve philosophical problems, so far as possible, by construing them in a way that lends itself to scientific treatment. The mind–body problem is a good case in point, and eliminative materialism (which claims that explanations of behavior do not ultimately depend on appeals to the mental) illustrates how what seems unnecessary for scientific treatment of a problem may be ontologically discountenanced. Where the assumption of the philosophical primacy of scientific method is at its most influential, philosophical method is conceived as only locally autonomous: Scientific method and the results of its application are the basic determinants of both our standards of rationality and our inventory of reality.

Quite apart from the role in their thinking of scientific method as a model for philosophical inquiry, it may be that philosophers naturally tend to take one or the other of two central philosophical domains, epistemology or metaphysics, or some account developed therein, as primary, as *first philosophy*, in a suggestive but now uncommon terminology. If we give priority to epistemology, we tend to produce an ontology that posits the sorts of objects about which our epistemology says we can have knowledge or justified belief. If we give metaphysics priority, we tend to produce an account of justified belief which allows knowledge or justified belief about the sorts of things our ontology countenances as real. One's philosophical method affects both one's epistemology and metaphysics and one's sense of the relation between them. If our method is dominated by a priori reflection, we are likely to be rationalists in epistemology and realists in metaphysics, at least to the extent of countenancing whatever abstract objects must be posited to ground a priori knowledge. If our method is dominated by observation and experiment, or even by the idea that philosophical claims are ultimately responsible to observation and experiment, we are likely to tend toward empiricism in epistemology and, in metaphysics, to seek an ontology that countenances as real only what is either experienceable or necessary to account for our knowledge of what we experience.

Like epistemology or metaphysics, philosophical method can be primary in shaping a philosophical outlook. It is doubtful that it can wholly determine such an outlook; for apart from certain epistemological and metaphysical commitments, one cannot develop or even use a method. Similarly, one cannot develop an epistemology without making at least tentative metaphysical commitments or construct a metaphysics without making at least tentative epistemological commitments. Philosophers seem to accept as apparently axiomatic that what is knowable is in some sense real; and though, as many philosophers would regard as a lesson of skepticism, it is not self-evident that what is real is knowable, many philosophers cannot easily give up the conviction, or the quest to establish, that this is so. If this apparent asymmetry concerning the knowable and the real is genuine, then taken together with the primacy of our experience in our relations to others and the world, it may explain why epistemology tends, in at least many philosophers, to contribute even more than metaphysics to determining their overall views.

If philosophical method is to be clarified by the comparison with scientific method and not obscured by

assimilation to the latter, it is essential that we distinguish scientific method from something of which it is an immensely impressive special case: *theoretical method*. The former is empirical and, broadly speaking, experimental. The latter is the more general method of building and rebuilding theories in relation to data: raising questions, hypothesizing, comparing and evaluating hypotheses in relation to data, revising theories in the light of the comparisons and evaluations, and adopting theories through assessing competing accounts of the same or similar problems. This distinction has not always been recognized or fully appreciated. For one thing, given the influence of empiricism (an influence to which few in modern philosophy are entirely immune), some thinkers tend to see scientific method as the only kind of theoretical method, at least outside logic and mathematics. But theoretical method is not the property of empiricism; rationalists can also use it, and so can both nonphilosophers and philosophers who are uncommitted with respect to, say, empiricism, rationalism, and pragmatism.

What is here called the theoretical method is very old—as ancient as systematic philosophy itself. It is illustrated in the Socratic attempt to refine definitions by revising them in response to examples and counterexamples; and it, or some major element in it, figures in all of the general philosophical methods considered here. However, the assumption of the primacy of scientific method and with it the often tacit view that scientific method is the only rational theoretical method outside logic and mathematics, is far from obvious.

Consider metaphysics: Properties and propositions, for example, far from being banished, are indispensable for many philosophers, including many who are scientifically oriented. Quite properly, this is in part because of what is required to understand science. But it may be in metaphysics, philosophy of language, and philosophy of mind that we find the greatest impetus toward preserving these common targets of Occam's Razor. Consider epistemology: There is to date no consensus that the traditional domain of the a priori has been accounted for on scientific or, especially, empiricist lines. If only a limited number of philosophers are willing to defend the view that there are synthetic a priori propositions (roughly, substantive propositions, such as basic moral principles, knowable on the basis of reflection on their content), increasingly, many philosophers are alive to the possibility that there may be. This is not to say that the analytic–synthetic distinction has been adequately clarified or is even important in many of the ways it has been thought to be. The suggestion is only that the categories of the analytic and the a priori are less and less widely thought to have been shown unintelligible or empty or even equivalent.

THE AUTONOMY OF PHILOSOPHY

Given what has been said in this entry, it should be plain that philosophy is a distinctive area of inquiry. Even if its concerns overlap those of various other disciplines, it has its own problems and at least some of its own methods. But distinctiveness is not the same as autonomy, which, as applied to a field of inquiry, implies a kind of independence of other such fields. Is philosophy autonomous in this sense? Positively, a rationalistic perspective can provide a stronger basis for the autonomy of philosophy than can an account of philosophy based on assuming the philosophical primacy of scientific method. The reference here is to *hard autonomy*—the kind grounded in a distinctive conceptual and methodological status. This is quite different from *soft autonomy*—the sociological and institutional independence of the discipline manifest chiefly in its generally having its own academic departments.

Soft autonomy is sustainable even if one's philosophical perspective is that of *naturalism*, which, in a strong form, might be described in rough terms as the view that nature is the whole of reality, and the only basic truths are truths of nature. On a form of this view associated with W. V. Quine, philosophy is continuous with natural science. This implies that there is no radical difference in the kinds of claims they can justify or in their standards of evidence: Indeed, epistemology itself is taken to be a kind of psychological inquiry into our cognitive standards and practices. The recently developed field of cognitive science, moreover, may from this perspective be viewed as a kind of naturalized philosophy of mind though its range may include more than problems addressed in that subfield of philosophy. This naturalistic approach to philosophy does not imply that there are no philosophical questions appropriately answered by reflection rather than through scientific inquiry, but the status of the answers is empirical rather than a priori; they are ultimately responsible to observation, as are scientific hypotheses, if in a less direct way. By contrast, on the traditional view that at least some major philosophical theses are a priori, it is clear why they are accountable to distinctively philosophical standards and need not be judged by the evidence drawn from sensory observation or scientific experiments.

To be sure, on the view that philosophy is simply more general than science or asks questions different in

subject matter from those of the special sciences, a de facto autonomy may be sustained, an autonomy that is more than sociological and less than conceptual. But on that view, philosophy does not stand apart from science in the same way nor does it possess autonomous standards of assessment, particularly in normative matters. If, as has been common in the history of philosophy, it is seen as an autonomous cultural resource, as a normative critical enterprise responsible to its own standards, it would seem desirable that philosophy stand apart from science in the suggested way. But distinctness is not opposition nor does distinctness entail competition. Moreover, supposing the hard autonomy thesis is mistaken, soft autonomy may be retained with renewed emphasis. If (in ways to be sketched below) philosophy is, or at least should be, a cultural resource, then whatever philosophers think about hard autonomy, they have reasons to preserve the soft, sociological autonomy of the discipline.

PHILOSOPHY IN RELATION TO OTHER DISCIPLINES

There are many other disciplines, and here it is possible only to indicate how philosophy is related to some of the major ones. The place to begin is with the idea that philosophy is in a sense the *metadiscipline*, the one whose proper business includes accounting for the structure, methodology, and, indeed, the implicit metaphysics and epistemology, of the other disciplines.

For understanding other disciplines, philosophy is indispensable. Many important questions about a field, such as the nature of its concepts and its relation to other disciplines, do not belong to that discipline, are not usually pursued in it, and are philosophical in nature. Philosophy of science, for instance, is needed to supplement the understanding of the natural and social sciences, which may be derived from scientific work itself. Philosophy of literature and philosophy of history are of similar value in understanding the humanities, and philosophy of art is important in understanding the arts. Philosophy is, moreover, essential in assessing the various standards of evidence used by other disciplines. Since all fields of knowledge employ reasoning and must set standards of evidence, logic and epistemology have a general bearing on all of these fields.

Normative disciplines and their subfields—those subfields that overlap normative ethics or properly propose broadly ethical standards—deserve special comment. These include (among others) law, theology, and aesthetics.

LAW. The field of law generates many philosophical questions. One concerns the very nature of law, which some have held to imply a connection with morality and others have taken to be entirely a matter of institutional realities, such as a structure of promulgations and enforcements. On either view, philosophy bears directly on important questions of what relation the law *should* have to morality. It also bears on the relevant standards of evidence. What, for instance, constitutes proof of guilt, and what should determine who counts as a reasonable person in relation to standards of negligence and due care? The topics of moral and legal responsibility, including the problem of diminished capacity and partial blameworthiness, are also areas in which philosophical and legal concerns overlap.

THEOLOGY. Theology is another field that overlaps philosophy. Philosophy of religion concerns not only the problem of adequately characterizing the divine nature but the related question of the rationality conditions for religious faith. Another major question pursued in both philosophy and theology is the relation between ethics and religion. Both areas of inquiry are connected with understanding the nature of evil—whether moral, as with wrongdoing, or natural, as in the case of death from floods—and how evil is possible (in various kinds and degrees) in a world under a god who is all-knowing, all-powerful, and wholly good. Historically, philosophy has influenced theology, just as theology has influenced philosophy. Although it is widely thought that either can be pursued in abstraction from the other, philosophical assumptions are both inevitably presupposed and commonly discussed in the field of religion.

AESTHETICS. Philosophy of art has been mentioned; aesthetics also includes the theory of natural beauty and related questions concerning aesthetics value. Although it should be granted that practitioners of the arts need not know even the rudiments of the philosophy of their art, this is rarely, if ever, so for professional critics and interpreters of the arts. Even if it is possible for critics, philosophy provides a way of conceiving the work and products of the artist that helps critics to appreciate it and to see its place in the culture to which it belongs. Literature in particular may either raise philosophical questions in its own creative works or invite their philosophical interpretation. Philosophy itself constructs mininarratives as central examples, uses dialogue—implicitly or explicitly—and not infrequently relies on metaphors and other literary devices. It is a literary medium from the vantage point of which other kinds of literature can be viewed in

relation to kindred standards of coherence, plausibility, clarity, and profundity.

The relation of philosophy to the professions should also be considered here. Its bearing on law has been noted. Not all of the professions can be mentioned, but it is appropriate to say something briefly about medicine, journalism and communication, and the broad field of business and economics.

MEDICINE AND OTHER HEALTH PROFESSIONS. The very notion of health is normative, particularly in the case of mental health. In this connection, ethics is clearly pertinent; so is philosophy of mind, with its emphasis on understanding the human person. Philosophy of science may yield a better understanding of—and even a greater capacity for—the integration of medical research with medical practice. Philosophy of religion can lead to a better understanding of many patients and of various other people with whom physicians work closely. Aesthetics and the history of philosophy may enhance the common ground practitioners can find with patients or colleagues who are from other cultures or have unusual orientations or views. Philosophy of medicine and medical ethics are obviously of direct relevance.

JOURNALISM AND COMMUNICATION. Journalists face a number of challenges on which philosophy bears. One is determining what is important enough to need coverage. Another is what constitutes objectivity in reporting on events and balance in editorializing. A third is ascertaining the quality of evidence on a given issue; this may be crucial in deciding whether to trust a source or to rely on an anonymous one. A comparative and, in some cases, a historical perspective is highly desirable (and arguably obligatory) in journalism; in achieving perspectives of these kinds, philosophical reflection is useful and sometimes indispensable. There are also more specific ways in which philosophy bears on journalism and communication: Philosophy of language, for example, should enhance understanding of communication, and philosophy of science should cast light on some of the technical subjects with which many people in journalism and communication must deal. Beyond this, political and social philosophy can deepen understanding of society and social institutions. For journalists with special interests, aesthetics, philosophy of law, and philosophy of religion are highly pertinent to the questions they face.

BUSINESS. For many people in business and (applied) economics, the bearing of philosophy on the world of commerce seems at best tenuous. But what we have seen about business ethics alone should belie that impression. A sound ethical perspective is essential for producing a sound code of ethics; philosophical training is valuable in providing a clear, adequately comprehensive, and defensible code. Economic justice, as with employment policy and fair competition, is a major concern that is clarified by work in ethics. So are the nature and responsibilities of corporations, unions, and political parties. Moreover, if cost-benefit analysis is to be mastered, the understanding and assessment of probabilities is essential. These topics are treated by inductive logic and epistemology.

THE PLACE OF PHILOSOPHY IN THE ACADEMY

Some of what should be brought out here is implicit in what has been said: That philosophy is a basic and comprehensive field of knowledge and, as such, has a place in higher education should now be evident. Philosophy also contributes to the capacity for problem solving in any field. In this respect its value is interdisciplinary and subject matter neutral.

CRITICAL THINKING. The first thing to note in this connection is that the study of philosophy helps to develop both the capacity and the inclination to do critical thinking. Logic is the most general philosophical field that develops this ability. Ethics alone is quite general. Studies in the subject should show how philosophical reflection is applicable to moral problems of many kinds. Courses in ethics commonly aim both at giving students a better understanding of moral problems and at helping them develop a reasonable moral outlook from which to approach the moral problems that confront them in their own lives. No other discipline treats these problems in the same comprehensive and systematic ways. Indeed, scientists and others often explicitly hold that such problems are outside their professional domain. Epistemology may be cited as the only discipline that examines standards of evidence and criteria of rational belief systematically and in ways applicable to any subject matter whatsoever. A similar point holds for many other topics that are treated in depth by philosophy and are important for critical thinking; they include definition, knowledge, explanation, causation, justification, communication, meaning, and truth.

NORMATIVE ISSUES. Philosophy provides a unique and systematic approach to normative issues—those concerning what ought or ought not to be, what is right or wrong, what is intrinsically desirable or undesirable, and so on—as opposed to what is as a matter of fact simply the case.

What are the basic moral rights of persons? What moral obligations do people in a society have to one another? What constitutes justice in the distribution of goods and in the determination of punishment? Inquiries in such areas as ethics, political philosophy, philosophy of law, and aesthetics treat normative questions in depth. Courses in these fields usually examine several theories proposed by philosophers in answering these questions, and typically, students in them are encouraged to formulate and defend their own answers to the questions using the methods and concepts introduced in the courses. Given the importance that moral, social, aesthetic, and other value questions have in human life, the contribution philosophy can make in a balanced curriculum is incalculable. It might be thought that these questions do or can receive adequate treatment in the social sciences or perhaps in literature and history. These other disciplines, however, do not, and do not claim to, deal with normative questions in the way philosophy does; and many of the important normative problems philosophers study are not raised in other fields.

INTERDISCIPLINARY PERSPECTIVE. An important function of philosophy is to foster interdisciplinary perspective. For instance, although scientific explanation is, in one form or another, common to all the sciences, conceptual questions about its nature and comparative questions about its logic in the different sciences belong to the philosophy of science. Some of these questions have been treated by scientists but rarely with the comprehensiveness and generality required for a synoptic understanding of the topic. Every discipline generates some essentially philosophical questions about itself, and many questions about relations among different disciplines are also philosophical. Both kinds of questions are examined in such areas as philosophy of science, philosophy of art, philosophy of law, philosophy of history, and philosophy of language. Philosophy also critically examines methods of inquiry, both in science and in everyday life. Its approach in this is usually conceptual, evaluative, and comparative; and typically the philosophical study of these topics differs from other approaches in the techniques used, in the questions pursued, and in the scope of the theories produced in answering these questions. Both in exploring the interrelations among other disciplines and in examining their methods of inquiry, philosophy fulfills a unique and important role as a metadiscipline. It provides a kind of understanding of the other disciplines—particularly of their presuppositions, standards of evidence, and modes of explanation—which other fields of study neither attempt nor are able to provide.

WRITING AND EFFECTIVE COMMUNICATION. A major aim of higher education is to contribute to the quality of discourse in and beyond its institutions of learning. The study of philosophy generally requires analytical writing, critical reading, and formulating intellectual problems and proposed solutions to them. For these reasons, work in philosophy can greatly improve writing and communication skills. Even if writing is taught virtually throughout the curriculum, philosophy can play a major and distinctive part in the task. No other discipline emphasizes, in the same ways, either verbal argumentation or conceptual analysis. Few other disciplines emphasize, to the same degree, students' producing their own theories or critical assessments as opposed to exposition of existing material. In addition, clarity, accurate interpretation, due consideration for others' positions, and the importance of using concrete examples are also stressed in competent teaching of the writing that philosophy requires. These qualities of philosophical training in writing and speaking make the study of philosophy especially valuable in preprofessional pursuits as well as for those seeking a more general education.

THE CULTURAL SIGNIFICANCE OF PHILOSOPHY

INTELLECTUAL HISTORY AND CROSS-CULTURAL VISION. In its historical and cross-cultural investigations, philosophy provides a sense of intellectual history and contributes to one's understanding of one's own culture in relation to other cultures. Most philosophy departments and institutes have programs of research and teaching that address at least ancient, modern, and contemporary philosophy. Many departments offer courses in philosophies produced by cultures other than their own. Studies in these areas help people to locate themselves historically and culturally, to work out a reasonable system of values, and to achieve an understanding of alternatives among values, cultural patterns, and intellectual traditions.

EXAMINATION OF WORLD VIEWS. A presupposition of higher education is that most reflective people seek a coherent view of the world that makes sense of their experience, guides them in certain major decisions, and gives them at least tentative answers to some of the perennial problems concerning human life and its place in the universe. The study of philosophy helps one to formulate and assess such views, whether they are drawn from the history of thought in a particular part of the world from comparative cross-cultural studies, from popular inter-

pretations of current science, or from the one's own—perhaps quite unarticulated—reactions to one's experience. Among the (partial) world views commonly examined in philosophy are materialism, which construes everything there is, including persons, as essentially physical; dualism, which takes minds and hence persons to be radically different from purely physical entities; and, of course, theism in many of its forms. Often, sociopolitical orientations, such as liberal democracy and Marxian socialism, are associated with world views. In examining these positions and world views, the approach of philosophy is holistic, conceptual, and evaluative. Moreover, whatever world view philosophers may hold, in teaching philosophy, they normally make it their business to present forcefully arguments for *and* against their own positions. Their most characteristic concern in this kind of endeavor is to develop a framework for making rational decisions on world views and sociopolitical orientations, not to inculcate any particular one.

ARTICULATION AND CRITIQUE OF PUBLIC POLICY. A huge number of public policy issues are *mainly* moral, and most of them have significant parts that are moral. Normative ethics thus has special bearing on their proper resolution. Abortion and prostitution are mainly moral issues; this is because the chief disagreements are generally over moral rights and principles rather than over nonmoral facts. Distribution of wealth and the structure of the health care system are largely moral issues; but nonmoral factual questions, such as what effects one or another system has, are relatively more important for these issues than for the former two. Moral philosophy speaks directly to problems of public policy. For one thing, they involve questions of justice and of human rights. It is a major task of moral philosophy to develop an adequate theory of justice and a related theory of moral rights. These theories attempt to answer such questions as whether justice requires an equal distribution of wealth; whether everyone has a right to material well-being; whether punishment, as distinct from rehabilitation, is morally justified; and what moral obligations rich nations have to help poor nations. The abortion issue is of particular concern here. This is because a major aspect of it concerns the metaphysical question (also debated in theological contexts) of what constitutes a human person. The issue cannot be adequately understood, then, without a degree of both ethical and metaphysical sophistication.

Philosophers, like others, are divided on these questions, but on one important point they are largely agreed: that there *are* ways of distinguishing good from bad reasoning on moral questions and that some answers to these questions are better than others. In any case, it should be clear that philosophical reflection may help in clarifying issues, evaluating or constructing arguments on each side, determining the full range of policy options, framing definitions (particularly in drafting legislation), deducing consequences from a position so that we can see what it commits us to, eliciting and criticizing basic assumptions, and evaluating a moral issue in the light of the best theories and principles available in moral philosophy.

THE PHILOSOPHER. Philosophy is so broad and complex that no one is an expert in all of its fields. This does not entail that there is nothing of a general kind that can be said about what constitutes a philosopher. The simplest thing to say is that any philosopher will have a high level of competence in at least one of the subfields described here. That will imply using at least one method sketched above or a substantially similar method; it will also imply having a sense of some of the other subfields of philosophy. It does not imply taking any particular view or reflecting on any particular problem. Philosophical training and dialectic are, however, sources of intellectual versatility. In this and other ways, philosophy can add to the depth, scope, and acuity of the wise, much as wisdom can add to the powers of discernment and judgment of the philosopher.

It is widely known that, etymologically, philosophy is the love of wisdom. There is also a strong association—perhaps partly derived from the emphasis on practical wisdom in both Plato and Aristotle—of philosophical reflection with wisdom. In part for these reasons, some people have assumed that a philosopher must be wise, particularly in practical matters. If wisdom in a domain (such as human relations) is taken to be knowledge and soundness of judgment in that domain, it is true that philosophical reflection has high potential for leading to a degree of wisdom, at least in some important domains. It is certainly true that wisdom is a characteristic of many philosophers and inclines many who have it to appreciate one or another philosophical problem. But philosophical competence is no guarantee of wisdom, and wisdom of many kinds is possible for nonphilosophers.

Perhaps the most positive point to be made here is that philosophical competence in a subject-matter area will reveal at least a substantial proportion of the truths and some of the conceptual resources that are needed by a person who has wisdom in that domain. Much depends on the area in question: The more *conceptual* or norma-

tive it is, the greater the bearing of philosophy. Philosophical competence brought to the field of law, for example, can go a long way: Major questions in the law concern evidence, conceptual distinctions, and such normative notions as justice and blameworthiness. These are areas in which epistemology and ethics have much to contribute. The connection of philosophy to computer science may be less close; but even apart from the importance of logic in this field, there are ethical questions of, for instance, privacy and intellectual property rights, for which competence in ethics is of great value.

Quite apart from whether philosophers are characteristically wise, their cultural role includes criticism of major elements in their culture, particularly those that are intellectual, ethical, aesthetic, religious, or political. Certain important kinds of philosophical criticism are in a certain way neutral: The charge of inconsistency or incoherence is morally neutral; the point that an argument is invalid is logical and leaves open whether the argument's constituent propositions are true. A not uncommon view among philosophers has been that, qua philosophers, they should remain neutral in this way, abstaining from moral and political positions. On this view, taking these positions is appropriate for philosophers in their role as citizens but not in their role as professional philosophers.

A less restrictive view is that philosophers as a group, as represented by, for instance, the American Philosophical Association, should not take moral or political positions in official resolutions; and a still less restrictive position would apply this restriction to political but not moral issues. Nonetheless, just as there are philosophical works that systematically defend normative ethical views, there are some defending normative political positions. Why, it may be asked, should philosophers who have well-developed normative political positions not put them forward for the general public as philosophically well grounded? Publication itself may be regarded as a step in this direction, particularly if the style of the work and the medium of publication lend themselves to wide reading by the general public. Moreover, as electronic publication becomes more widespread and more readily accessible to the general public, the distinction between what is published for a professional audience and what is addressed to a wide public audience may become harder to draw.

Disagreement among philosophers about the proper cultural role of philosophy is likely to continue, and they can quite reasonably hold different views on the kinds of public moral or political positions appropriate for wide dissemination by philosophers as individuals as opposed to philosophers acting institutionally or as a corporate body. But we may safely say that, particularly with the declining influence of positivism from the middle of the twentieth century to the present time, few philosophers now believe that taking normative positions in ethics, politics, and elsewhere is not properly philosophical. One way to put a major part of this point is to say that philosophers as such may be prescriptive as well as descriptive. Indeed, even counseling people to avoid slipshod reasoning is prescriptive. Moreover, quite apart from any explicit prescriptions, criticisms of reasoning or counterexamples to proposed ideas are implicitly prescriptive: Plainly, one should not rely on bad reasoning or maintain an idea to which there are clear counterexamples. As a critical enterprise, philosophy is implicitly normative. As appraising major guiding ideas in human life, it is implicitly prescriptive.

CONCLUSION

Philosophy is the systematic and critical study of ideas and issues, a reasoned pursuit of fundamental truths, a quest for a comprehensive understanding of the world, a study of principles of conduct, and much more. Every domain of human existence raises questions to which its techniques and theories apply, and its methods are applicable in the study of any subject or the pursuit of any vocation. Its inquiries encompass the critical study of knowledge and reality, of value and obligation, of religion and science, of language and literature, of art and the professions. In the academy, philosophical studies enhance the capacity for problem solving, the ability to understand and express ideas, and the power to frame cogent arguments. In the culture in which it is practiced, philosophy can be a critical voice, a defender of ideals, a creator of visions.

Philosophy also develops understanding and enjoyment of things whose absence impoverishes many lives: aesthetic experience, communication with many different kinds of people, discussion of current issues, the discerning observation of human behavior, and intellectual zest in the pursuit of knowledge. For individuals in or outside the academy, the study of philosophy provides a major route to developing a well-reasoned vision of the good life and an ability to communicate this vision, defend it, and where necessary modify it. A well-reasoned vision of what human life ought to be yields an ordered set of long-term goals and a sense of the significance of life; it provides, often, the steady intellectual stimulation of comparing a theory of human experience with the constantly changing, ever-surprising panorama that our

experience is; and it anchors our relations with others in a framework that enables us to conceive human conduct with some measure of clarity and understanding.

Bibliography

This entry contains many proper names, as well as many terms common in philosophy, that have entries devoted to them in this encyclopedia. Readers seeking an overall perspective on the nature of philosophy are urged to consider entries on these philosophers or philosophically important terms. One may also find much of relevance to understanding what philosophy is by consulting the entries on special fields, say epistemology, ethics, metaphysics, and the philosophy *of* subfields, such as philosophy of mind, of religion, or of science. Also recommended are the philosophy entries in the first edition of this work (1967) and its supplement (1996). Some of the material in this entry is drawn (with permission) from parts of two documents (of which the author was principal writer) published by the American Philosophical Association with the idea of clarifying the nature of the field and its academic study: "The Role of Philosophy Programs in Higher Education." *Proceedings and Addresses of the American Philosophical Association* 53 (3) (1980): 363–370; and "Philosophy: A Brief Guide for Undergraduates." *Proceedings and Addresses of the American Philosophical Association* 56 (2) (1982): i–xviii. Some material is also based on the author's "Realism, Rationality, and Philosophical Method," written with a similar purpose and appearing in *Proceedings and Addresses of the American Philosophical Association* 61 (1987): 65–74.

Robert Audi (2005)

PHILOSOPHY, HISTORIOGRAPHY OF

See *History and Historiography of Philosophy*

PHILOSOPHY BIBLIOGRAPHIES

See *"Philosophy Bibliographies" in Vol. 10*

PHILOSOPHY DICTIONARIES AND ENCYCLOPEDIAS

See *"Philosophy Dictionaries and Encyclopedias" in Vol. 10*

PHILOSOPHY JOURNALS

See *"Philosophy Journals" in Vol. 10*

PHILOSOPHY OF BIOLOGY

Biology refers both to the systematic investigation of living things, and to the body of knowledge that is the product of that investigation. Throughout biology's history, however, some important questions debated by biologists have not been so much about the organisms being studied, but about the nature of life, the proper way to investigate it and the form biological knowledge should take. When inquiry shifts from questions about living things to questions about proper and improper ways of asking, or answering, or adjudicating, such questions, it shifts to a philosophical level. One need not, of course, be trained in a department of philosophy to contribute to such an inquiry. Indeed many of the most significant contributions to the subject have been made by people trained in the sciences. Nevertheless such contributions are to the subject designated as philosophy of biology.

One assumption implicit in the very name is that the biological sciences are distinctive enough from other sciences that a general inquiry into the nature of science will not suffice. It was common among logical empiricists to suppose it would—in texts written in that tradition one often finds the biological and social sciences dealt with in chapters late in general books in philosophy of science (Braithwaite 1953, Hempel 1966, Nagel 1961). Two early challenges to this assumption were *The Ascent of Life* by Thomas Goudge (1961) and *The Biological Way of Thought* by Morton Beckner (1959).

These two early contributors were followed by a number of introductions to philosophy of biology written in the 1970s and 1980s, most of which were focused narrowly on evolution and genetics, and a standard set of associated philosophical questions (Hull 1974, Ruse 1973, Rosenberg 1985, Sober 2000). But in 1982 an NEH Summer Institute in Philosophy of Biology organized by Richard Burian and Marjorie Grene attracted a group of philosophers ready to focus more or less exclusively on the biological sciences. In 1984 *Philosophy of Science* devoted a "special issue" to the philosophy of biology; shortly thereafter Michael Ruse played a pivotal role in organizing both a journal (*Biology and Philosophy*) and a society (International Society for History, Philosophy, and Social Studies of Biology) devoted exclusively to the biological sciences. Since that time, the scope of research has broadened dramatically, with important contribu-

tions focusing on the biomedical sciences, physiology, cell biology, neurobiology, and developmental biology (Amundson 2005; Bechtel and Richardson 1993; Fox Keller 2000; Oyama, Griffiths, and Gray 2001; Robert 2004; Schaffner 1993; Sterelny and Griffiths 1999). And it is now common in general introductions to philosophy of science to have two chapters on philosophy of biology (for example, Salmon et al. 1992; Machamer and Silberstein 2002). In addition, there has been a tendency to integrate advances in the history of biology into these philosophical discussions.

This entry focuses on issues associated with three related biological domains: genetics, evolution, and development. Some of the most interesting recent philosophical work is focused on developmental biology and its relationship to the other two domains just mentioned. But important work is also being done on areas such as ecology, ethology, and neurobiology: each raises its own special philosophical questions.

DARWIN, MENDEL AND A PARTIAL SYNTHESIS

Two publications in the mid-nineteenth century—Charles Darwin's *On the Origin of Species* and Gregor Mendel's "Experiments in Plant Hybridization"—were to have a lasting impact on the structure of the scientific study of life, an impact still evident in the way philosophers think about biology as a science. Darwin self-consciously characterized the theory of evolution by natural selection presented in his book as one that would provide a theoretical unity to the study of life. As he put it in a letter to the philosopher Sir John Herschel:

> … I find so many young and middle-aged truly good workers in different branches, either partially or wholly accepting my views, *because they find that they can thus group and understand many scattered facts*. This has occurred with those who have chiefly or almost exclusively studied morphology, geographical Distribution, systematic Botany, simple geology & palaeontology.
>
> (BURKHARDT 1994, PP. 135–136)

Darwin argued that central to explanations in all these domains were a set of "laws," which modern scholars identify as the principles at the core of the theory of evolution by natural selection:

Variation. The characteristics of the individual members of a species vary to a greater or lesser degree.

Inheritance. Some of that variation is heritable, transmitted from parents to off-spring.

Geometric increase. Populations tend to increase at a geometric rate.

Struggle for existence. Given limited resources, predation, disease, and so on, the tendency to geometric increase is checked, leading to a struggle for survival.

Differential survival. Individuals with advantageous variations tend on average to survive longer and leave more off-spring.

Differential reproduction. The offspring of parents with advantageous variations tend to have the same advantageous variations.

Darwin used the term "Natural Selection" to refer to the last two principles: "I have called this principle, by which each slight variation, if useful, is preserved, by the term Natural Selection…" (Darwin 1859/1964, pp. 117, 127) And, since the above theory neither provides for the introduction of new variation nor for the divergence within species that actually leads to speciation, he followed the presentation of his theory with lengthy discussions of divergence of character and laws of variation. Crucially, he saw divergence leading to new species and higher taxa as simply long run extrapolations of the same processes that lead to the production of varieties within a species; and he decoupled the causes of new variation from the adaptive needs of the organism.

Modern presentations of the theory often reduce it to a combination of the production of heritable variation and the differential perpetuation of variation. Many philosophical problems emerge from this reduction, and a number of philosophers have been urging a formulation of the theory more in tune with Darwin's.

In 1866, just seven years after Darwin's *Origin*, a scientifically trained monk published the results of nine years of careful experimentation. Mendel's work was revolutionary both in its methods and its conclusions. Trained at the University of Vienna in experimental physics and statistics as well as botany (where he learned about recent developments in agricultural plant hybridization), Mendel realized that the combination of experimental controls and statistical analysis could be used to solve the puzzles of plant hybridization. In the varieties of pea plants with which he experimented, he established that a number of factors were inherited independently, that if one crossed plants with alternate forms of a factor (for example, green and yellow peas) all their offspring would appear like one or the other (the dominant form), and that the next generation of plants would

reveal a ratio of the forms of approximately three dominant to one recessive. Further experimental analysis would reveal that the "dominant" plants were a predictable mixture of pure dominants and plants with a dominant and a recessive factor.

But he went beyond just outlining his experiments and providing plausible inductive inferences from them. In the last section of his paper he suggested an underlying causal mechanism that could account for the observed regularities that he had reduced to a mathematical law of the development of hybrids.

Largely ignored for the remainder of the nineteenth century, the basic idea was developed in quite different ways by many researchers from different disciplinary backgrounds in the first decade of the twentieth century. Scholars have from then until now referred to this theory of inheritance as Mendelian genetics, and the regularities he uncovered experimentally as "Mendel's laws." And because this theory had essentially two distinct components—one related to inheritance, and one related to development—it in theory provided a way of unifying branches of the study of life that Darwin admitted he had not. A number of key steps taken between 1905 and 1920 isolated Mendel's "factors" (genes) to chromosomes and paved the way for generalizing Mendel's principles of inheritance in a mathematical form that would allow their investigation in large populations that do not breed under strictly controlled experimental conditions. This permitted the integration of Mendelian genetics with Darwinism, and it is no surprise that the leading figures in creating this synthesis—J. B. S. Haldane, Ronald Fisher, and Sewall Wright—all had a passion for both mathematics and natural history (Provine 1971, Plutynski 2004, Sarkar 1992). And all had, by quite different routes, fallen under the spell of Darwin's theory of evolutionary descent driven primarily by natural selection.

The basic idea behind their synthesis was remarkably simple: think of populations of organisms in terms of the frequencies of the genes associated with the various traits found in those populations, and think of evolution in terms of gradual changes in those frequencies, under such influences as the migration of organisms in and out of the population, randomly occurring genetic mutations, genetic recombinations of various kinds and, above all, natural selection. Mathematical models were developed which permitted one to predict changes in the genetic make up of future populations given information (or, more often, assumptions) about these variables, the number of alleles of genes for given traits and assumptions about the relative fitnesses of different combinations of these alleles, know as "genotypes." The crucial step in developing these models was, of course, that each of these potential influences on gene frequencies was treated as a quantitative variable—including fitness.

It should be noted that all of these people—and the founder of experimental population genetics, Thomas Hunt Morgan, should be added—treat "genetics" as the study of the transmission of genes in reproduction. Indeed, even in the twenty-first century a synonym for Mendelian genetics is transmission genetics. The Evolutionary Synthesis did not include the study of development or developmental genetics, except for the use of embryological evidence in constructing evolutionary phylogenies. We will come back to this omission later.

Concluding the preface of *Evolution: The Modern Synthesis*, Julian Huxley declared: "The need today is for concerted attack and synthesis" (Huxley 1942, p. 8). That Synthesis was in the making when he wrote (the key publications by Theodosius Dobzhansky, Ernst Mayr, and George Simpson were published between 1937 and 1944), and by the 1959 centenary of the publication of Darwin's *Origin* it was declaring itself triumphant. Most of the philosophical issues related to evolutionary biology under discussion in the 2000s are a direct consequence of the form that the "Synthetic" theory takes This entry mentions five that are critical and discusses four of them in detail.

1) The concept of chance in evolutionary theory and the theory's probabilistic nature.

2) Fitness and selection.

3) Units and levels of selection.

4) The nature selection/adaptation explanations.

5) The ontological status of species and the epistemological status of species concepts.

THE ROLE OF CHANCE. Chance is a contrastive concept; to say that some outcome is chance is typically to deny that it resulted from some cause or other. In evolutionary theory "chance" plays a key role both in discussing the generation of variation and the perpetuation of variation (a distinction owed to John Beatty; see also Sober 1984, ch. 4). Consider the following variation grid, created by considering whether the contribution to fitness of a variation does or does not play a role in either the generation or the perpetuation of that variation:

	Variations	
	Generation	Perpetuation
Fitness Biased	Lamarck	Darwin
Not Fitness Biased	Darwin Neutralism	Lamarck Neutralism

The uniquely Darwinian position is that the generation of variation is chance in that it is not biased by fitness differences (as it is for Lamarckian theories), but the perpetuation of variation typically is biased by fitness differences. Neutralism, to be discussed shortly, will claim that a significant amount of evolutionary change is due to randomly generated variation that is perpetuated by chance as well.

But now consider the following discussion of chance and selection:

> In Darwin's scheme of things, recall, chance events and natural selection were consecutive rather than alternative stages of the evolutionary process. There was no question as to which was more important at a particular stage. But now that we have the concept of random drift taking over where random variation leaves off, we are faced with just such a question. That is, given chance variations, are further changes in the frequencies of those variations more a matter of chance or more a matter of natural selection?
>
> (BEATTY 1984, P. 196)

In the first two sentences, as often, the generation of variation is characterized as a "chance" process because selection plays no role at that stage—the generation of variations is not biased by the adaptive requirements of the organism. The concept of "random variation" is often used by neo-Darwinians as a synonym for "chance variation" in precisely this sense, as in the following from a product of Morgan's "fruit fly lab" and one of the architects of the evolutionary synthesis:

> … mutation is a random process with respect to the adaptive needs of the species. Therefore, mutation alone, uncontrolled by natural selection, would result in the breakdown and eventual extinction of life, not in the adaptive or progressive evolution.
>
> (DOBZHANSKY 1970, P. 65)

The generation of variations is a "chance" process in the sense that the probability assignments are not biased by "adaptive needs" or "fitness."

The remainder of the quotation from Beatty concerns the perpetuation of variations, and in particular how to distinguish variations perpetuated by selection from those perpetuated by another process known as "random drift," in which traits that are selectively neutral may become fixed in a population simply as a result of what statisticians call "errors of sampling." Suppose, for example, that a pair of bats get blown to an island far away from their colony. They mate, and their offspring mate, and a number of genes become fixed in the growing populations simply because they were present in the founding pair, not because they are favored by selection. In the above quoted paper Beatty argues that "it is conceptually difficult to distinguish natural selection from random drift" (Beatty 1984, p. 196). As the entry discusses, this problem arises from a standard way of characterizing "fitness."

Genetic drift plays a critical role in one primary challenge to the neo-Darwinian synthesis, "neutralism," and the concept of chance is often used to draw the contrast. In the following quote, one prominent champion of the neutral theory of molecular evolution characterizes his position:

> … the great majority of evolutionary changes at the molecular (DNA) level do not result from Darwinian natural selection acting on advantageous mutants but, rather, from random fixation of selectively neutral or very nearly neutral mutants through random genetic drift, which is caused by random sampling of gametes in finite populations.
>
> (KIMURA 1992, P. 225)

Here genetic drift refers to a process whereby a selectively neutral allele becomes fixed in a population as a result of a "random (chance) sampling of gametes." This is a rival to neo-Darwinism only because of Kimura's claim that this produces a majority of the evolutionary changes at the molecular level. The contrast between "chance" and "fitness biased" processes is used by Kimura to distinguish means of perpetuating certain variations. We are contrasting two sampling processes.

There is currently a lively debate about whether to characterize this contrast by reference to differences in the sampling processes (Millstein 2002, 2005) or by reference to the expected outcomes of sampling (Brandon and Carson 1996, Brandon 2005). On Millstein's view it is realistically possible for the outcomes to be identical, and thus she seeks to defend a view according to which selection is defined as discriminate sampling (based on selectively relevant differences) and drift as indiscriminate sampling. Both samplings are "probabilistic," of course, but that in no way obviates the above contrast.

FITNESS AND SELECTION. All parties to this dispute are now realizing that wider issues about the nature of probability, explanation, mathematical abstraction, and causation are likely at stake in such disagreements. As one case in point, at least part of the dispute over differentiating drift from selection derives from the tendency to characterize natural selection so that it is indistinguishable from random drift (Brandon 1990, Lennox 1992, Lennox and Wilson 1994).

If we think of selection as a discriminative or biased sampling process, that natural raises the question of the basis of the biasing. Typically, the answer is that it is differences in fitness, the values assigned to different genotypes in the models of population genetics, which some readers will think of as different degrees of adaptation to the relevantly characterized environment.

But as noted above, it is not uncommon to find characterizations of the fitness of a genotype in terms of its relative contribution to the gene pool of future generations—the genotype contributing the larger percentage being the fitter. The expression "survival of the fittest" has essentially been eliminated from any serious presentation of Darwinian selection theory but the concept of "fitness" plays a prominent, and problematic, role. In the mathematical models used in population genetics "fitness" is represented by the variable *W*. Here is a rather standard textbook presentation of the relevant concepts:

> In the neo-Darwinian approach to natural selection that incorporates consideration of genetics, fitness is attributed to particular genotypes. The genotype that leaves the most descendants is ascribed the fitness value $W = 1$, and all other genotypes have fitnesses, relative to this, that are less than 1. … Fitness measures the relative evolutionary advantage of one genotype over another, but it is often important also to measure the relative penalties incurred by different genotypes subject to natural selection. This relative penalty is the corollary of fitness and is referred to by the term selection coefficient. It is given the symbol s and is simply calculated by subtracting the fitness from 1, so that: $s = 1 - W$.
>
> (SKELTON 1993, P. 164)

In this passage evolutionary advantage is equated with reproductive success and fitness is treated indifferently as a quantitative measure of both. But since, as we have seen, natural populations can evolve (via drift) in the absence of natural selection, and since balancing selection may prevent a population from evolving, it is clear that establishing, by measuring different reproductive rates among its members, that the genetic makeup of a population has changed does not establish that natural selection was the source of that change; nor does the fact that no change has been measured establish that natural selection is not operative.

The most widely accepted solution to this problem is to argue that fitness measures a reproductive propensity of organisms (Brandon 1978, Mills and Beatty 1979, Richardson and Burian 1992). Brandon tends to equate fitness in this sense with "adaptedness," and to contrast it with "realized fitness"—differences in realized fitness are explained by differential adaptation to a common selective environment. This suggests that fitness is in some sense relational, enhancing chances of reproducing relative to an environment (Lennox 1992, Lennox and Wilson 1994). In any case, as Millstein has insisted, characterizing fitness as a reproductive propensity raises the question of how to understand this propensity and its organic basis (Millstein 2003).

UNITS AND LEVELS OF SELECTION. A number of challenges to Darwinian selection theory have emerged since the mid-twentieth century. Those challenges can be placed into two broad categories: (1) proposed limitations on natural selection as the primary cause of evolutionary change; and (2) expansions of the scope of natural selection to include new "targets" and "levels." It will be noted that in neither case is it obvious that the theory itself requires modification in the face of such challenges—in principle these might be nothing more than challenges to the theory's range of application. However, if it turned out that most evolutionary change could be explained without recourse to natural selection, this would be grounds for arguing that evolutionary biology was no longer Darwinian (see Godfrey-Smith in Orzack and Sober 2001.)

Darwin conceived of natural selection as almost exclusively an interaction between individual organisms and their organic and inorganic environments. Taking that as our starting point, we can see two challenges to Darwinism today with respect to the units of selection. One comes from those defending a strong form of genic selectionism, such as G. C. Williams (1966, 1992) and Richard Dawkins (1976, 1982), who argue that selection is always and only targeting genes. Here is a clear statement:

> These complications [those introduced by organism/environment interactions] are best handled by regarding individual [organismic] selection, not as a level of selection in addition

> to that of the gene, but as the primary mechanism of selection at the genic level.
>
> (WILLIAMS 1992, P. 16)

Dawkins' preferred mode for making the same point is to refer to organisms—or interactors, to use language introduced by David Hull—as the vehicles of their genes (the replicators), in fact vehicles constructed by the genome for its own perpetuation. Neither Williams nor Dawkins deny that there is interaction between phenotype and environment that plays a role in the differential perpetuation of genes. Their argument is that those interactors are part of the "genic selection mechanism," as Williams worded it above.

This view has been extensively challenged by philosophers of biology on both methodological and conceptual grounds (Brandon 1996, ch. 8; Mitchell 2003, ch. 4; Moss 2003, ch. 1; Sober 1984, chs. 3, 7; Sterelny and Griffiths 1999, chs. 4–5), though there are, among philosophers, also enthusiastic supporters (Dennett 1995). In all the give and take, it is seldom noticed how odd it is that defenders of this view claim to be carrying the Darwinian flag (Gayon 1998 and Gould 2003 are exceptions). Yet it is certainly not a position that Darwin would recognize—and not merely because he lacked a coherent theory of the units of inheritance. It is not a Darwinian view because for Darwin it was differences in the abilities of organisms at various stages of development to respond to the challenges of life that had causal primacy in the explanation of evolutionary change. Gene selectionism was explicitly challenged on these grounds by key figures in the Synthesis (for example, Ernst Mayr).

The Darwinian view of the units of selection also has challenges from the opposite direction. In the 1970s a number of biologists working in the fields of paleontology and systematics challenged the Neo-Darwinian dogma that you could account for "macro-evolution" by simple, long-term extrapolation from the processes modeled by population genetics. (The case was enhanced by parallel and contemporaneous developments in embryology and functional morphology that are discussed in the last section of this entry.) Stephen Jay Gould (2003), in a chapter titled "Species as Individuals in the Hierarchical Theory of Selection" combines two conceptually distinct theses: first, the thesis defended by Michael Ghiselin (1997) and championed and refined by David Hull (2001), that species are in a robust sense of the term "individuals"; and second, that there may well be selection among groups of organisms, *qua* groups. This approach brings us to the brink of problem (5) on the list, how to understand the species category and species as taxa, questions discussed only briefly.

Gould exemplifies one approach to group selection—the unit of selection is always the individual, but there are individuals other than individual organisms that are subject to selection. A very different result emerges if one assumes that groups of organisms such as demes, kin-groups, or species, though not individuals, are nevertheless possible units of selection. Adding to the conceptual complexity, some researchers propose that "group selection" be restricted to the process whereby group-level traits provide advantages to one group over another, in which case there are strict conditions delimiting cases of group selection, while others focus solely on group level effects. Thus a debate analogous to that earlier discussed regarding the definitions of "fitness" emerges here—by group selection do we mean a distinct level of causal interaction, or merely a tendency within certain populations for some well defined groups to displace others over time? (For further discussion, see Sterelny and Griffiths 1999, 151–179; Hull 2001, 49–90.) It is now common to characterize "selection," "interactor," and "replicator" abstractly and to specify the conditions under which an entity is properly identified as a unit of selection. This allows one to leave it an open and essentially empirical question whether, under the right conditions a particular "unit" could be subject to selection. With the modular picture of development that is emerging, the "developmental module" will likely be added to the list.

SELECTION, ADAPTATION, AND TELEOLOGY. Perhaps the central promise of Darwinism, and the reason it was rightly seen as a challenge to the Argument from Design for a benevolent creator, was that it provides a scientific explanation for both phylogenic continuity and adaptive differentiation by means of the same principles. The nature of "selection explanations" is a topic to which much philosophical attention has been devoted in recent years (Allen, Bekoff, and Lauder 1998; Sober 1984). How does one account for the apparently teleological character of explanation by natural selection?

The appearance of teleology is certainly present in Darwinian explanations, and has been since Darwin spoke of natural selection working solely for the good of each being. The appearance of teleology stems from the ease with which both evolutionary biology and common sense take it for granted that animals and plants have the adaptations they do because of some benefit or advantage to the organism provided by those adaptations. But in

what sense can the adaptive advantage be the cause of the presence of the adaptation?

Some insist it cannot (Ghiselin 1997). Others argue that such explanations are actually masked appeals to the past effects of selection (see the papers in Allen, Bekoff, and Lauder, section 3). This entry sketches a case that shows selective explanations of adaptations are robustly teleological (see Lennox 1992, 1993; and the papers in Allen, Bekoff, and Lauder 1998, section 1).

Are the functions performed by confirmed adaptations a central and irreducible feature of explanations of the presence of those adaptations? If the answer is yes, the explanations are teleological. Take the following example. In research combining painstaking field work and laboratory experimentation, John Endler demonstrated that the color patterns of males in certain Caribbean guppy populations resulted from a balance of mate selection and predator selection. For example, he demonstrated that a group of males with a color pattern that matched that of the bottoms of the streams and ponds they populated except for bright red spots have that pattern because a common predator in those populations, a prawn, is color blind for red. Thus red spots did not put their possessors at a selective disadvantage, and were attractors for mates\ (Endler 1983).

Their pattern of coloration was a complex adaptation serving the functions of predator avoidance and mate attraction—and it is an adaptation, as that term is used in Darwinism, only if it is a product of natural selection (Williams 1966, Brandon 1985, Burian 1983). In order for it to be a product of natural selection, there must be an array of color variation available in the genetic/developmental resources of the species wider that this particular pattern but including this pattern. Which factors are critical, then, in producing differential survival and reproduction of guppies with this particular pattern in a shared homogeneous environment? The answer would seem to be the *value*-consequences this pattern has compared to others available in promoting viability and reproduction. In popular parlance (and the parlance favored by Darwin), this color pattern is good for the male guppies that have it, and for their male offspring (Binswanger 1990, Brandon 1985, Lennox 1992). This is a robust version of "consequence etiology" accounts of selection explanations (Bekoff, Allen, and Lauder 1998, section 1), which stresses that selection ranges over value differences which are causally relevant to one among a number of color patterns having a higher fitness value. Selection explanations are, then, a particular kind of teleological explanation, an explanation in which that for the sake of which a trait is possessed, its valuable consequences (avoiding predation, attracting mates), account for the trait's differential perpetuation and maintenance in the population.

SPECIES AND TAXONOMY. Darwin at one point in the *Origin* says that he considers the term "species" one that is given arbitrarily, for convenience. He based that comment on a review of the taxonomic work of his day, and a similar review today would have the same result. Equally competent taxonomists will disagree about whether to rank a group of similar organisms as members of the same species or as members of two distinct species. This issue takes on philosophical import because speciation—the "origin of species," to use Darwin's language—is taken to be the key step in the evolution of life. One would hope to have a clear way of deciding, at least in principle, when that step has been completed! But every attempt to give a clear account of what makes a taxonomic unit a member of the species category runs up against rather compelling problems. Surrounding this topic, which has generated an enormous literature, are both epistemological issues regarding the basis for our species concepts and ontological issues about the nature of species. Interested readers should consult the work of Marc Ereshefsky (1992, a collection of essays defending various views of the species category) or Kim Sterelny and Paul Griffiths (1999, chapter 9; a readable and current overview of the issues).

GENES

In standard texts in the philosophy of biology in the 1970s and 1980s (as well as in most of the more technical journal articles) genetics played a key role in the discussion of two philosophical topics: reductionism, and the structure of evolutionary theory (Hull 1974, ch. 1; Schaffner 1969; Ruse (1973), ch. 10; Rosenberg 1985, ch. 4). The discussion began by importing a theory of reduction that had been developed with physical theories in mind, and asking whether, on such models, there had been a reduction of Mendelian or transmission genetics to the molecular level. This model, developed most clearly by Ernst Nagel (1961), imagined two theories formalized with axioms and laws. Reduction would require that the laws of one theory be, in some clear sense, deducible from the fundamental laws of the other, as, with appropriate corrections, Kepler's planetary laws could be from those of Newtonian celestial mechanics. Typically, this would also require that the key concepts in the two theories be interdefinable. This model was developed into a "general reduction/replacement model" by Kenneth Schaffner (1969) in a paper in which he argued

for the potentially application of such a model to the case of genetics.

David Hull (1974) pointed out that a critical problem in the way of achieving this goal was that the two theories had essentially different goals and domains—hereditary transmission of differences versus genetic input into the biochemistry of development. This suggested to him not only the impossibility of a reduction, but its irrelevance to biology. All parties to this discussion concluded that one needed a much more elaborate account of both biological theory structure and explanation to even try to answer the question.

A number of recent discussions have stressed that understanding both biological investigation and explanation in terms of mechanisms and their operations provides a more realistic picture of fields such as neurobiology and molecular biology (see Machamer, Darden, and Craver 2000; Waters 1994, 2000). It may also provide a more tractable notion of "reduction" in terms of "underlying mechanisms." Detailed histories of the development of genetics played a very important role in this discussion, and thus it was one important area driving the integration of history and philosophy of biology. A fine review of that topic, as well as a carefully hedged defense of genic reductionism which takes into account the complex, interlevel nature of typical biological theories, can be found in Schaffner 1993, chapter 9 (and see Waters 1994 for a somewhat different defense).

One of the puzzles that emerges from reviewing the literature on genetics and reduction is that Hull's point, mentioned above, about the fundamentally different aims of molecular genetics and transmission genetics only really gets serious consideration once developmental biology comes to the fore in the 1990s (see Waters 1994 and the papers in Beurton, Falk, and Rheinberger 2000). This is still very much a discussion in process, so the entry touches on some of the philosophical and historical questions being raised about different uses of the gene concept (or, alternatively, different gene concepts) in evolutionary and developmental contexts.

The traditional gene concept associated with Mendelian genetics that formed the basis of evolutionary biology was important because it was the basis of heritable differences in populations. Genes, or more precisely alleles, were the sources of heritable variation in populations, and thus provide "the material basis for evolution." In the context of developmental biology, however, the focus of research has always been on the genes as sources of deep relationships among species within and even across phyla. Here is a succinct expression of the difference:

> In the Modern Synthesis of population genetics and evolution, genes become manifest by differences in alleles that are active in conferring differential reproductive success in adult individuals. The gene is though to act as a particulate, atomic unit. In current syntheses of evolution and developmental genetics, important genes are manifest by their similarities across distantly related phyla, and they are active in the construction of embryos. These developmental genes are thought to act in a context-dependent network.
>
> (SCOTT GILBERT, IN BEURTON, FALK, AND RHEINBERGER 2000, P. 178)

In a defense of genic selectionism George Williams argues that genes should be understood as units of information.

> Only DNA provides the durable archive for most of the earth's organisms. This constraint should not blind us to the fact that it is information we are concerned with, and that DNA is the medium, not the message. A gene is not a DNA molecule; it is the transcribable information coded by the molecule.
>
> (WILLIAMS 1992, P. 11)

Williams praises philosophers for adopting the distinction between replicators and interactors discussed earlier, but he is critical of them for regarding "replicators as material objects and miss[ing] the codex concept" (Williams 1992, p. 12). Notice that what this approach does is allow nominal acceptance of advances in our understanding of cellular mechanisms at the molecular level while continuing to treat the gene as an "atomic" unit differentiated by reference to phenotypic differences. That is, development can continue to be "black boxed" by taking the gene to be any selectively relevant bit of the "codex" for the organism. In principle it should allow Williams to take on board a suggestion made by Sterelny and Griffiths; on grounds that lots of things get replicated in reproductive cycles "gene selectionism should be generalized to 'replicator selectionism'"(Sterelny and Griffiths 1999, p. 69).

However, taking this approach also raises a new set of concerns, namely those involved with the application of concepts from information theory in the characterization of genes and gene action. This way of talking became extremely popular after the "breaking of the genetic code" in the 1960s. Complementary strands of this "double

helix" consist of only four bases, two purines (Adenine [A] and Thymine [T]) and two pyramidines (Guanine [G] and Cytosine [C]); and since proteins consist of polypeptide chains made from only twenty different amino acids, if DNA is to contain the "instructions" for synthesizing all the possible proteins, the simplest possible "code" would be one in which three bases combined to specify each amino acid. The bases came to be represented as the "letters" of the genetic "alphabet"; they combine into syllables, words and "reading frames"—the book of life! The coded script is "transcribed" and "translated"; there is an "encoding" and "decoding" process; and with the discovery of the complexities of DNA transcription, it is not surprising that terms like "editing" and "proofreading" got added. It is not inevitable that these metaphors should lead researchers to present the genome as both the architect and the blueprint for building an organism—but it was natural. (For a compelling story of the history, see Keller 2000.)

Are there problems with it? A number of philosophers of biology think so, and they are discussed in this entry in two parts. The entry first discusses those problems that are not specific to developmental biology, and then discusses the philosophical debates around Developmental Systems Theory and "evo-devo."

The aforementioned quotation from George Williams (1992) establishes that some evolutionary biologists want to take the information metaphor one step further, and allow it to float free from its source in the discovery of the relationship between DNA sequence and amino acid differences. Genes are units of information, pure and simple. The value of doing this is that it allows one to avoid the troubling fact that the causal complexity of the processes involved in biological development make it quite meaningless to talk about some relatively short and self-contained DNA sequence as a gene for anything other than an amino acid, perhaps.

What could "information" be, in this case? It seems not to be information in the sense of mathematical information theory. One suggestion is to see it from a "teleosemantic" point of view; that is, genes are something akin to units of "meaning," their meaning being what they are present for, the phenotypic trait whose selection insures the replication of that gene (Sterelny and Griffiths 1999, pp. 82–92). Critics have argued that this simply severs completely the causal connection between DNA sequences and phenotypic traits. And it looks as if the original impulse for gene selectionism gets lost. It looks like the interactor—whether it is a colony, an animal, or a gamete—is the only serious causal determinant of differential reproduction. "A purely functional notion of a gene, untied to anything constant at the molecular level, is not a definition suitable for gene selection theory, whatever its other uses might be" (Sterelny and Griffiths 1999, p. 90).

A very different defense of the language of information in biology would in fact tie it very tightly to the detailed machinery of molecular biology, and therefore takes seriously the role of the analogies based on this language in the development of molecular genetics, such as treating DNA as a "reading frame" made up of triplet units, which then allows one to see certain mutations as analogous to "frame shifts" that create nonsense (Maynard Smith 2000, p. 184). But even here, the sense of "information" that seems relevant to the analogy is again a semantic notion tied to meaning and intentionality, not that of the 'signals', 'channels', and 'sources' of "information theory." And thus all the problems associated with that notion are still present. (See the replies to Maynard Smith in Godfrey-Smith 2000; Sarkar 2000.) There is consensus here, however, that the language associated with codes and information storage and retrieval was extremely important in the development of molecular biology. Insofar as there is disagreement, it is over whether these metaphors have outlived their usefulness and are now in fact the source of significant misunderstanding.

THE CHALLENGE OF DEVELOPMENT

Much recent philosophy of biology has focused on the process of development. There are at least two reasons for this. First, the model of the gene described in the previous section is deeply problematic, and one response was the philosophical defense of a developmentalist alternative based on the work of Susan Oyama, known as Developmental Systems Theory or DST. In an important and productive exchange on "the developmentalist challenge" to this sort of "genetic primacy," Ken Schaffner (1998) identified eleven theses of DST, and focused on four with which he thinks serious researchers in biomedical molecular genetics would, in one form or another, agree. The basic idea is that development is a product of a complex time series of interactions among many cellular and extra-cellular factors, among which "genes" (and the quotation marks are important) are just one. As a consequence, DST denies what I have called "information theoretic determinism." In so far as the information metaphor has value (this is currently much disputed), it is applicable only to the developing system—genes carry

information, if at all, only as aspects of developmental systems.

Another obvious consequence of DST is the rejection of various common themes of behavioral genetics (Lewontin 1995). Schaffner's approach to these DST theses is to compare them with work done on a "simple model system," one of the organisms at the center of human genome research, the nematode worm *C. elegans*. This worm became a model organism due to Sidney Brenner adopting it to investigate the development, from zygote to mature adult, of an entire nervous system (and behavior). It was ideal for many reasons, not the least of which was its simplicity—a nervous system containing only 302 neurons forming roughly 5,000 synapses. Schaffner compares the DST theses about genes, development and behavior with the results of the massive, worldwide research assault on *C. elegans* to see how they hold up. This entry cannot follow the details, but one can see from his eloquent conclusion that at least some of the DST argument is acceptable to him:

> Characterizing simple "genes for" behaviors is, accordingly, a drastic oversimplification of the connection between genes and behavior, *even when we have the (virtually) complete molecular story*. The melody of behavior represents no solo performance—it is the outcome of an extraordinarily complex orchestra—and one with no conductor.
>
> (SCHAFFNER 1998, P. 247)

This paper was the target for responses from philosophers of biology and biologists more or less sympathetic to DST (Gilbert and Jorgensen 1998, Wimsatt 1998) and Schaffner was given the last word in reply. This selection of papers constitutes the best introduction to the DST reply to gene-centered research (see Waters 2005).

Independently of philosophical discussion of DST, there are compelling reasons for philosophers to be interested in evolutionary developmental biology, or "Evo-devo." Given the long history of both developmental biology and evolutionary biology, and the long history of their interactions, one might wonder why the goal of integration has appeared on the horizon only in the twenty-first century. The answer is a complex of historical, philosophical, and biological components.

According to one historical narrative (Beurton, Falk, and Rheinberger 2000; Burian 2005), the rapid development of new investigative techniques in molecular biology, driven in part by the medical and agricultural potential of the methods of genetic modification, and the field of "genomics" that evolved along with the Human Genome Project, provided the means for investigating development at the molecular level. This gave rise to a number of quite revolutionary discoveries; this entry notes only two: (1)The "Hox" regulatory genes encode a special sort of protein with a stretch of amino acids known as a "homeodomain." These proteins attach to quite specific segments of DNA, regulating the expression of a series of genes. These proteins act in concert and have "modular" effects on such things as organ formation, body segmentation and bilateral duplication of body parts (a clear introduction can be found in Burian 2005, chs. 11 and 12). (2) Molecular genetics is providing a highly complex, "interactive" picture of gene regulation. It will be noted that in the description of the Hox genes it became clear that certain proteins were responsible for their regulation. In fact all sorts of signals, some coming from within the cell and some from the extracellular environment, play a role in gene expression. This is now so widely accepted that philosophers and historians refer to it as "the interactionist consensus." According to this picture, genes are one of many interacting factors all of which must play their roles in order to give rise to an organism—the study of this interactive process is termed "epigenetics," though it is unclear to what extent its practitioners understand development as a truly epigenetic process (Robert 2004, ch. 1).

On this view, the integration of evolutionary and developmental biology will be—is being—effected by the long overdue integration of molecular genetics, and the molecular understanding of development, into evolutionary studies. At least one advocate of this view (Burian 2005) has stressed the modularity of this view of development, and the implication of the semi-autonomy of these developmental modules (body segmentation and bilateralism, organ systems, limb structure, and so on) for the way evolution can possibly work.

There is another way of viewing the history, being developed in different ways by Alan Love and Jason Robert (Love 2003, Robert 2004). Love focuses on what the proponents of Evo-devo claim their investigations can do that the current evolutionary "synthesis" cannot. Many proponents of this field put the explanation of evolutionary novelties such as feathers, tetrapod limbs, or jaws as the central contribution of development to evolution. While they are happy to concede to population genetics and ecological genetics the explanation of gradual evolutionary changes in traits associated with one or a few Mendelian genes, they argue that the explanation of the appearance of novelties at particular phylogenetic junctures requires an understanding of the network of

changes in the organization of developmental resources needed to produce the novel structure, and an understanding of its functional morphology. The history of work on evolutionary novelties focuses attention on a number of research programs in developmental biology, functional morphology and paleontology, all focused on understanding the first appearance of novel structures and behaviors—and all more or less ignored by the evolutionary synthesis.

Jason Robert argues for the primacy of the organism as it develops from zygote to maturity, and thus for a seriously "top down" or "whole to part" view of developmental causation. This allows us to see the analytic tools that allow us to understand the details of the developmental mechanisms as a first step, with true understanding of development coming when we have an integrated understanding of how those mechanisms interact. Pretty much everyone looking at this rapidly developing area of biology agrees with the following sentiment from Burian:

> During the next few decades, I believe, biologists will highlight the roles played in constructing organisms by dynamic regulatory systems above the level of the genome. The result will be a non-vitalist but much more holistic, vision of the organism, one that places the integration of the organism at the focus of attention. In short, our new understanding of the apparatus regulating gene expression has undermined classical genetic determinism.
>
> (BURIAN 2005, P. 243)

As Robert as pointed out, this prediction for the future sounds remarkably like a return to the "organismic" biologists, such as E. S. Russell, writing in the 1920s and 1930s, against the then rising tide of a population genetic centered evolutionary synthesis (Robert 2004). There are, of course, critics of this viewpoint. While the aforementioned text indicated that Schaffner's review of *C. elegans* research encouraged him to accept, at least in a modified form, some of the theses of "the developmentalist challenge" to genetic determinism, the modifications were significant. And some would likely say he has gone too far, arguing that what we have in this new molecular understanding of development is a vindication of reductionism (Waters 1994, 2004, 2005; Rosenberg 1985, discussed in Robert 2004, pp. 12–15).

Evo-devo once again brings into focus the question of the unity of biology as a science. As stressed earlier, one thing that the evolutionary synthesis provided for philosophers of biology was an image of how the biological sciences could be unified that was decidedly unlike the standard models based on the physical sciences. The attempt to unify evolutionary biology and developmental biology may complicate that image considerably. The fields omitted from the synthesis share key concepts (for example, gene, homology) with evolutionary biology, but appear to deploy them in very different ways. Moreover, the methods of investigation in functional morphology, developmental biology and population genetics or ecology are extremely different. The central problems and questions to be answered are very different, because the basic research agendas of the fields are very different. A field that focuses on "the production of the tetrapod limb" and a field that thinks of populations as gene pools of heritable variation being sampled by selection do not appear to look at organisms in the same way (Amundson, in Orzack and Sober 2001; Love 2003). As this proposed "synthesis" or "integration" takes place, philosophers of biology can both test their models of theoretical unification against the accomplishment of evo-devo, and can provide its advocates with ideas about adequacy conditions for a successful integration. One thing appears certain at this point: evo-devo specialists who have explicitly written on this topic see a special set of problems that will require an integration of concepts and techniques from evolutionary biology and developmental biology; they do not imagine one field being gradually "reduced" to another.

What, then, are the logical and conceptual prerequisites for such an integrated investigation? If we look back to where we started in this entry, it will be recalled that the "integration" of Darwinism with the Mendelian genetics of populations, required the concepts of "fitness" and "selection" to be reshaped into a mathematical form; and what began as a cytologically and developmentally based genetics eventually "black-boxed" development in the interests of focusing on the transmission of genetic "information" from one generation to the next. These changes, in the interests of integration or synthesis, gave rise to a host of philosophical problems. Perhaps, with philosophers and historians inextricably involved with this new synthesis, at least some problems can be avoided.

See also Darwin, Charles Robert; Evolutionary Theory (Natural Selection); Special Sciences.

Bibliography

Allen, Collin, Mark Bekoff, and George Lauder, eds. *Nature's Purposes: Analyses of Function and Design in Biology*. Cambridge, MA: MIT Press, 1998.

Amundson, Ronald. *The Changing Role of the Embryo in Evolutionary Thought*. Cambridge, U.K.: Cambridge University Press, 2005.

Ayala, Francisco. *Population Genetics: A Primer*. San Francisco: Benjamin Cummings, 1982.

Ayala, Francisco. "Teleological Explanations in Evolutionary Biology." In *Nature's Purposes: Analyses of Function and Design in Biology*, edited by Collin Allen, Mark Bekoff, and George Lauder. Cambridge, MA: MIT Press, 1998.

Beatty, John. "Chance and Natural Selection." *Philosophy of Science* (Special Issue: Philosophy of Biology) 51 (2) (1984): 183–211.

Bechtel, William, and Robert Richardson. *Discovering Complexity*. Princeton, NJ: Princeton University Press, 1993.

Beckner, Morton. *The Biological Way of Thought*. New York: Columbia University Press, 1959.

Beurton, Peter, Raphael Falk, and Hans-Jörg Rheinberger, eds. *The Concept of the Gene in Development and Evolution*. Cambridge, U.K.: Cambridge University Press, 2000.

Binswanger, Harry. *The Biological Basis of Teleological Concepts*. Los Angeles: ARI Press, 1990.

Braithwaite, Richard B. *Scientific Explanation*. Cambridge, U.K.: Cambridge University Press, 1953.

Brandon, Robert. *Adaptation and Environment*. Princeton, NJ: Princeton University Press, 1990.

Brandon, Robert. "Adaptation Explanations: Are Adaptations for the Good of Replicators or Interactors?" In *Evolution at a Crossroads*, edited by Bruce Weber and David Depew. Cambridge, MA: MIT Press, 1985.

Brandon, Robert. *Concepts and Methods in Evolutionary Biology*. Cambridge, U.K.: Cambridge University Press, 1996.

Brandon, Robert. "The Difference between Drift and Selection: A Reply to Millstein." *Biology and Philosophy* 20 (1) (2005): 153–170.

Brandon, Robert, and Richard Burian, eds. *Genes, Organism, Populations: Controversies over the Units of Selection*. Cambridge, MA: MIT Press, 1984.

Brandon, Robert, and Scott Carson. "The Indeterministic Character of Evolutionary Theory: No 'No Hidden Variables' Proof but No Room for Determinism Either." *Philosophy of Science* 63 (3) (1996): 315–337.

Burkhardt, Frederick, et al. *The Correspondence of Charles Darwin*. Vol. 9. Cambridge, U.K.: Cambridge University Press, 1994.

Burian, Richard. "Adaptation." In *Dimensions of Darwinism*, edited by Marjorie Grene. Cambridge, U.K.: Cambridge University Press,1983.

Burian, Richard. *The Epistemology of Development, Evolution, and Genetics: Selected Essays*. Cambridge, U.K.: Cambridge University Press, 2005.

Darden, Lindley, and J. Cain. "Selection Type Theories." *Philosophy of Science* 56 (1989) 106–129.

Darwin, Charles. *On the Origin of Species: A Facsimile of the First Edition* (1859). Cambridge, MA: Harvard University Press, 1964.

Dawkins, Richard. *The Extended Genotype*. Oxford: Oxford University Press, 1982.

Dawkins, Richard. *The Selfish Gene*. Oxford: Oxford University Press, 1976.

Dennett, Daniel. *Darwin's Dangerous Idea*. New York: Simon & Shuster, 1995.

Depew, David, and Bruce Weber. *Darwinism Evolving*. Cambridge, MA: MIT Press, 1995.

Dobzhansky, Theodosius. *Genetics of the Evolutionary Process*. New York: Columbia University Press, 1970.

Endler, John. "Natural and Sexual Selection on Color Patterns in Poeciliid Fishes." *Environmental Biology of Fishes* 9 (2) (1983): 173–190.

Ereshefsky, Marc. *The Units of Evolution: Essays on the Nature of Species*. Cambridge, MA: MIT Press, 1992.

Gayon, Jean. *Darwinism's Struggle for Survival*. Cambridge, U.K.: Cambridge University Press, 1998.

Ghiselin, Michael. *Metaphysics and the Origin of Species*. Albany, NY: Student University of New York Press, 1997.

Gilbert, Scott, and Erik Jorgensen. "Wormwholes: A Commentary on K. F. Schaffner's 'Genes, Behavior, and Developmental Emergentism.'" *Philosophy of Science* 65 (2) (1998): 259–266.

Godfrey-Smith, Peter. "Information, Arbitrariness, and Selection: Comments on Maynard Smith." *Philosophy of Science* 67 (2) (2000): 202–207.

Godfrey-Smith, Peter. "Three Kinds of Adaptationism." In *Adaptationism and Optimality*, edited by Steven Orzack and Elliott Sober. Cambridge, U.K.: Cambridge University Press, 2001.

Goudge, Thomas. *The Ascent of Life*. Toronto: University of Toronto Press, 1961.

Gould, Stephen Jay. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press, 2003.

Griffiths, Paul, and Robin Knight. "What Is the Developmentalist Challenge?" *Philosophy of Science* 65 (2) (1998): 253–258.

Hempel, Carl. *The Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall, 1966.

Hull, David. *The Philosophy of Biological Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

Hull, David. *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. Cambridge, U.K.: Cambridge University Press, 2001.

Hull, David, and Michael Ruse, eds. *The Philosophy of Biology*. Oxford: Oxford University Press, 1998.

Huxley, Julian. *Evolution: The Modern Synthesis*. London: George, Allen and Unwin, 1942.

Keller, Evelyn Fox. *The Century of the Gene*. Cambridge, MA: Harvard University Press, 2000.

Keller, Evelyn Fox, and Elisabeth Lloyd, eds. *Keywords in Evolutionary Biology*. Cambridge, MA: Harvard University Press, 1992.

Kellert, Stephen, Helen Longino, and Kenneth Waters, eds. *Scientific Pluralism*. Minneapolis: University of Minnesota Press, 2005.

Kimura, Motoo. "Neutralism." In *Keywords in Evolutionary Biology*, edited by Evelyn Fox Keller and Elisabeth Lloyd. Cambridge, MA: Harvard University Press, 1992.

Lennox, James G. "Darwin *Was* a Teleologist." *Biology and Philosophy* 8 (4) (1993): 408–421.

Lennox, James G. "Philosophy of Biology." In *Introduction to the Philosophy of Science*, edited by Merrilee Salmon et al. Englewood Cliffs, NJ: Prentice-Hall, 1992.

Lennox, James G. "Teleology." In *Keywords in Evolutionary Biology*, edited by Evelyn Fox Keller and Elisabeth Lloyd. Cambridge, MA: Harvard University Press, 1992.

Lennox, James G. "Teleology by Another Name: A Reply to Ghiselin." *Biology and Philosophy* 9 (4) (1994): 493–495.

Lennox, James G., and Bradley Wilson. "Natural Selection and the Struggle for Existence." *Studies in History and Philosophy of Science* 25 (1) (1994): 65–80.

Lewontin, Richard. *Human Diversity*. New York: Scientific American Library, 1995.

Love, Alan. "Evolutionary Morphology, Innovation, and the Synthesis of Evolutionary and Developmental Biology." *Biology and Philosophy* 18 (3) (2003): 309–345.

Machamer, Peter, Lindley Darden, and Carl Craver. "Thinking about Mechanisms." *Philosophy of Science* 67 (1) 2000: 1–25.

Machamer, Peter, and Michael Silberstein. *The Blackwell Guide to the Philosophy of Science*. Oxford: Blackwell Publishers, 2002.

Maynard Smith, John. "The Concept of Information in Biology." *Philosophy of Science* 67 (2) (2000): 177–194.

Millikan, Ruth Garrett. "In Defense of Proper Functions." In *Nature's Purposes: Analyses of Function and Design in Biology*, edited by Collin Allen, Mark Bekoff, and George Lauder. Cambridge, MA: MIT Press, 1998.

Millstein, Roberta. "Are Random Drift and Natural Selection Conceptually Distinct?" *Biology and Philosophy* 17 (2002): 33–53.

Millstein, Roberta. "Selection vs. Drift: A Response to Brandon's Reply." *Biology and Philosophy* 20 (1) (2005): 171–175.

Mitchell, Sandra. *Biological Complexity and Integrative Pluralism*. Cambridge, U.K.: Cambridge University Press, 2003.

Mitchell, Sandra. "Functions, Fitness, and Disposition." In *Nature's Purposes: Analyses of Function and Design in Biology*, edited by Collin Allen, Mark Bekoff, and George Lauder. Cambridge, MA: MIT Press, 1998.

Moss, Lenny. *What Genes Can't Do*. Cambridge, MA: MIT Press, 2003.

Nagel, Ernst. *The Structure of Science*. New York: Harcourt, Brace and World, 1961.

Orzack, Steven, and Elliott Sober, eds. *Adaptationism and Optimality*. Cambridge, U.K.: Cambridge University Press, 2001.

Oyama, Susan, Paul Griffiths, and Russell Gray. *Cycles of Contingency: Developmental Systems and Evolution*. Cambridge, MA: MIT Press, 2001.

Plutynski, Anya. "Explanation in Classical Population Genetics." *Philosophy of Science (Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association)* 71 (5) (2004): 1,201–1,214.

Provine, William. *The Origins of Theoretical Population Genetics*. Chicago: University of Chicago Press, 1971.

Robert, Jason. *Embryology, Epigenesis, and Evolution: Taking Development Seriously*. Cambridge, U.K.: Cambridge University Press, 2004.

Rosenberg, Alexander. *The Structure of Biological Science*. Cambridge, U.K.: Cambridge University Press. 1985.

Ruse, Michael. *The Philosophy of Biology*. London: Hutchinson and Company, 1973.

Salmon, Merrilee, et al. *Introduction to the Philosophy of Science*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

Sarkar, Sahotra, ed. *The Founders of Evolutionary Genetics: A Centenary Reappraisal*. Dordrecht: Kluwer Academic Publishers, 1992.

Sarkar, Sahotra. *Genetics and Reductionism*. New York: Cambridge University Press, 1998.

Sarkar, Sahotra. "Information in Genetics and Developmental Biology: Comments on Maynard Smith." *Philosophy of Science* 67 (2) (2000): 208–213.

Schaffner, Kenneth. *Discovery and Explanation in Biology and Medicine*. Chicago: University of Chicago Press. 1993.

Schaffner, Kenneth. "Genes, Behavior, and Developmental Emergentism: One Process, Indivisible?" *Philosophy of Science* 65 (2) (1998): 209–252.

Skelton, Peter. *Evolution: A Biological and Palaeontological Approach*. Harlow, U.K.: Pearson Education Ltd., 1993.

Sober, Elliott. *The Nature of Selection*. Cambridge, MA: MIT Press, 1984.

Sober, Elliott. *The Philosophy of Biology*. 2nd ed. Boulder, CO: Westview Press, 2000.

Sterelny, Kim, and Paul Griffiths. *Sex and Death: An Introduction to the Philosophy of Biology*. Chicago: University of Chicago Press, 1999.

Waters, Kenneth. "Genes Made Molecular." *Philosophy of Science* 61 (2) (1994): 163–185.

Waters, Kenneth. "Molecules Made Biological." *Revue Internationale de Philosophie* 4 (214) (2000): 539–564.

Waters, Kenneth. "A Pluralist Interpretation of Gene-centered Biology." In *Scientific Pluralism*, edited by Stephen Kellert, Helen Longino, and Kenneth Waters. Minneapolis: University of Minnesota Press, 2005.

Williams, George C. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton, NJ: Princeton University Press, 1966.

Williams, George C. *Natural Selection: Domains, Levels, and Challenges*. Oxford: Oxford University Press, 1992.

Wilson, David Sloan. "Individual Selection and the Concept of Structured Demes." In *Genes, Organism, Populations: Controversies over the Units of Selection*, edited by Robert Brandon and Richard Burian. Cambridge, MA: MIT Press, 1984.

Wimsatt, William. "Simple Systems and Phylogenetic Diversity." *Philosophy of Science* 65 (2) (1998): 267–275.

Wright, Larry "Functions." In *Nature's Purposes: Analyses of Function and Design in Biology*, edited by Collin Allen, Mark Bekoff, and George Lauder. Cambridge, MA: MIT Press, 1998.

Wright, Larry. *Teleological Explanation*. Berkeley: University of California Press, 1976.

James Lennox (2005)

PHILOSOPHY OF ECONOMICS

Why would philosophers be interested in economics? There are at least two answers. First, lessons from eco-

nomics bear directly on moral and political philosophy, as well as on theorizing about rationality. Second, economics provides a case study of some of the most challenging problems in the philosophy of science.

ECONOMICS AS MORAL PHILOSOPHY

What is the ethical basis of economics? If economics is grounded in a theory of the right, what kind of theory is it? Is it a theory of the right grounded in a utilitarian conception of the greatest good for the greatest number, or a Kantian conception of the sovereignty of individual economic agents? Or, if economics is grounded in a theory of value, is the value to be understood in utilitarian or contractarian terms (as an aggregate, or as a matter of mutual advantage)?

PLATO. Alfred North Whitehead described philosophy as a series of footnotes to Plato. What about economics? Plato's *Republic* describes the emergence of a society not by social contract or by conquest but spontaneously, through the workings of the market. "The barest notion of a state must include four or five men" (Book II, 369D). People need food, shelter, and clothing, but "all things are produced more plentifully and easily and of a better quality when one man does one thing" (Book II, 370B). People thus start to specialize in farming, carpentry, and weaving. It quickly becomes obvious, though, that "more than four citizens will be required, for the husbandman will not make his own plough … Neither will the builder make his tools—and he too needs many; and in like manner the weaver and shoemaker" (Book II, 370C). Commercial society thus emerges as an unplanned consequence of the transparent advantages of the division of labor.

ADAM SMITH. In more substantial ways, economics is a footnote to yet another philosopher, born some twenty centuries later. It was Adam Smith, professor of logic and of moral sciences at the University of Glasgow, whose work led more or less directly to the rise of economics as a separate academic discipline. The first three chapters of Smith's *Wealth of Nations* (1981 [1776]) explain the role that division of labor plays in a prosperous society, culminating in a brilliant critique of protectionist trade policy. Using the manufacture of pins as an example, Smith notes that a solitary worker could scarcely make one pin per day, but in a pin-making factory employing ten workers, "one man draws out the wire, another straights it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head … ; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations … Those ten persons, therefore, could make among them upwards of forty-eight thousand pins a day" (p. 15).

Smith then explains how the division of labor is facilitated by the propensity to truck and barter. We do not build factories for our own personal consumption. We specialize to that degree only when we have opportunities to serve large communities. Smith's next insight is that the extent of specialization is limited by the size of the market. A rural carpenter specializes in anything made of wood; a carpenter in a large city specializes in residential house construction; a carpenter serving national and international markets can specialize in making childproof doorknobs. The wealth of nations depends on economic agents being able to reach far beyond their small circle of friends. The farther they can reach, the larger the markets they can reach both as producers and as consumers, the greater will be the division of labor, and the richer they and everyone with whom they trade will be. Because economic agents work with suppliers, distributors, and customers on a global scale, they can produce thousands of pins a day, rather than a small handful at best.

The homage Smith pays, then, is not so much to the self-interest of butchers and bakers as to the division of labor that enables artisans to continuously be renewing, reinventing, and extending the limits of their craft. These opening chapters of *Wealth of Nations* are perhaps the most insightful part of the most insightful work of economics ever written.

Smith's most pointed argument on behalf of a lightly regulated economy, though, is probably to be found in his less famous work, *The Theory of Moral Sentiments* (1976 [1759]). There, Smith argues that a "man of public spirit" will not be a fanatical reformer but instead "will respect the established powers and privileges even of individuals, and still more those of the great orders and societies … When he cannot conquer the rooted prejudices of the people by reason and persuasion, he will not attempt to subdue them by force" (p. 223). By contrast, a "man of system," "is apt to be very wise in his own conceit, and is often so enamored with the supposed beauty of his own ideal plan of government that he cannot suffer the smallest deviation from any part of it. He goes on to establish it completely and in all its parts, without any regard either to the great interests, or to the strong prejudices which may oppose it. He seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess board. He does not consider that the pieces upon the chessboard have no other principle of motion besides

that which the hand impresses upon them; but that in the great "chess board" of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it" (1976 [1759], p. 234).

SMITH'S LEGACY. Smith anticipates Marx in expressing reservations about the alienating aspects of repetitive labor in a factory setting. Smith also comes close to anticipating James Buchanan and Gordon Tullock's insight (1962) that legislators respond to incentives as much as anyone else does—that legislators are not philosopher-kings, above the fray, but instead make their moves on a chess board of society, like everyone else. None of the pieces gets to decide where all the other pieces will be at a given moment. Political order, like economic order, biological order, or any other complex order, will take its shape not because any particular designer intended it to take that shape, but simply because that is what happened when the pieces came together, all with plans of their own. The insight of Smith, Buchanan and Tullock, and others, is that our world is strategic—all the way up. Even a country's most powerful politicians can do no more than hope to exert some influence. It is hard to incorporate this insight into moral and political philosophizing. What would it be like to develop a theory of how to live, and how to pursue social change, in a world where no one is in charge?

Alexander Rosenberg (1988) observes that the products of natural selection are exquisitely functional and almost unimaginably complex, despite no one being in charge. Unplanned economies likewise are functional, indeed typically more functional than centrally planned ones. How is this possible? Friedrich Hayek (1994) argues that a free economy economizes on rationality, morality, and knowledge in a way that a central plan cannot. Central planning models assume central planners will know what they need to know, and will use such knowledge wisely, and for purposes other than their own. Starting with such assumptions, advocates of central planning aim to invest planners with enough power (to implement the "right" decisions) that other agents with less benign plans will be unable to interfere. Unfortunately, giving central planners that much power to do the right thing also gives them that much power to do the wrong thing: to repay debts to their most powerful supporters, to cover up mistakes, to eliminate enemies (anyone who criticizes them), and so on. According to Hayek, there is a fatal conceit involved in thinking that economies would work better if, per impossible, central planners were in charge.

To Ludwig von Mises (the other main protagonist in the "Socialist Calculation Debate," along with Hayek), economics is a value-free, a priori science, more or less like mathematics. But Daniel Hausman and Michael McPherson (1996) plausibly conclude that, "economics remains partly a moral science. It can't be done without moral presuppositions, and it's hard to do it well without addressing moral issues intelligently. Similarly, moral philosophy can't be done without beliefs about human interactions, and it's hard to do it well without knowledge of the kind that economists seek" (p. 8). For example, Hausman and McPherson ask whether market competition results in firms with moral scruples being driven into bankruptcy. They give several reasons to think the answer is no, but their main point is simply that the question matters, and matters in economics, not only in moral philosophy. It bears whether there is any point in being in favor of the market competition that economists study.

Extending Hausman and McPherson's point, and relating it back to Adam Smith, if a firm would need to dominate a small town market in order to do a profitable volume of business, it may find itself needing to cater to the interests of a "lowest common denominator." Or at least, the firm that survives to serve that small market in the long run will be the one that best serves the majority of clients in that small market. By contrast, if a firm can operate on a global scale (advertising on the Internet, perhaps), then capturing even 1 percent of the market can be hugely profitable. In this way, globalization makes possible a proliferation of specialized firms catering to especially discerning clientele, raising free-range poultry, growing organic broccoli, auctioning nineteenth-century German marbles, manufacturing parachutes out of recycled newspaper, or whatever entrepreneurs think of next. (Israel Kirzner, a student of Mises, criticizes how standard equilibrium models treat entrepreneurial innovation. As Kirzner sees it, standard models treat innovations as exogenous shocks, when in fact entrepreneurial innovation is a central driving force in all but the most repressive states, which is why real economies are *always* in disequilibrium.)

ECONOMICS AS SCIENCE

Lionel Robbins (1935) defined economics as "the science which studies human behavior as a relationship between ends and scarce means which have alternative uses" (p. 16). It is amazing how much can be derived from a premise that economic agents put scarce resources to their most efficient use. But is the premise true? Milton Friedman (1984) seems to say it makes little difference; the

unrealism of a theory's assumptions are unimportant, so long as the theory's predictions are correct. Hausman describes Friedman's essay as the most influential work on economic methodology of the twentieth century. Analogously, Hausman says, Ptolemy's astronomy is still used for navigational purposes. Is Hausman right? The idea is theory-laden, and more technical than it appears. It is true that the Copernican revolution did not require us to make any radical changes in our ways of navigating, but is that like saying our navigational methods are premised on the earth being at the center of the universe? Probably not, but Hausman's main point still stands: We do not need to know the rock-bottom truths of astronomy, astrophysics, or anything else in order to have theories that track relevant facts well enough to enable us to navigate. Likewise, in economics, the statement that economic agents are pursuing their own self-interest is close enough to the mark for many purposes, and accordingly has, for many purposes, a lot of explanatory and predictive power. We better understand much of what we see around us when we grasp that self-interest is a more or less ubiquitous motive. Yet, we also see every day that people are motivated by things other than self-interest: by benevolence, vengefulness, and also (as Hobbes observed) by pigheaded, self-destructive vainglory.

THE SCIENTIFIC ATTITUDE. One of the biggest methodological blunders we could make would be to retreat from this messy empirical reality to the empty platitude that people do whatever they do, and this is all we really mean when we say all action is self-interested. When we give up the willingness to let our generalization be tested (and sometimes disconfirmed) by reality, we also give up the generalization's relevance as a tool for understanding reality. As a sometimes disconfirmed generalization, the postulate of self-interest lets us know when to regard a behavior as surprising, worthy of scientific curiosity, and so on. We may find that seemingly altruistic behavior turns out to be, in some previously unnoticed way, self-interested after all. But so long as we avoid the trap of assuming this *must* be the case, no matter what, we leave ourselves open to learning something new. (The new direction of progress may not be economics per se. New directions tend to evolve into new sciences. Just as moral philosophy helped spawn economics, economics can help spawn new disciplines or subdisciplines.) Meanwhile, so long as we understand the postulate of self-interest as a simplification of reality, one that abstracts from messy empirical details, the postulate will be useful.

Karl Popper sought to distinguish between science and nonscience. The real issue is about scientific attitude—whether a theory's proponents treat the theory as something to scrutinize rather than to zealously defend. In any case, it is hard to confirm an economic theory, or any other kind of theory. We give theory a chance to fail, and are impressed by and more confident in it as it survives repeated testing. But as scientists we acknowledge that surviving a test does not put us in a position to be supremely confident. Real science does not work that way; its fruits are not indubitable certainties.

What are the limits of a general theory's ability to help us understand? Daniel Little suggests, "The abstract analysis of the firm based on rational agents arriving at efficient outcomes must be supplemented with more detailed analysis of the specific circumstances and arrangements within which the firm took shape" (1995, p. 6). This is not a throwaway line but is in fact rather disturbing. It suggests there are severe albeit vague limits on the prospects for general explanation.

In the same way, one might see the history of philosophy as pointing to a similar conclusion. Namely, the search for general explanations, general theories, even general definitions, has a history of butting up against recalcitrant limits. There is a point to analyzing knowledge as justified true belief; yet, we now know of cases where this analysis is not good enough. It seems that no matter how much we tweak a theory or a definition, perfection is not an option. When cartographers try to map a three-dimensional terrain by projecting it onto two dimensions, there is no such thing as a representation of the terrain without distortion. A Mercator projection makes Greenland look as large as Africa, and anything we do to correct this distortion of relative size will distort something else in the process. This is an example of a problem for which a perfect solution simply does not exist, and theorists in all sorts of philosophical disciplines, confident though they may be that there is an objective truth about the three dimensional terrain out there, and that their job as theorists is to provide an accurate map of that objective reality, are finding themselves facing the reality that, as a rule rather than as an exception, there are no perfectly accurate theories. Our theorizing needs to be supplemented by knowledge of the local terrain. There is no denying our need for practical wisdom, or as Little puts it, for "detailed analysis of specific circumstances and arrangements" (1995, p. 6).

ECONOMICS AS AN EXPERIMENTAL SCIENCE. Experimental economics starts with the idea that economic

hypotheses are testable in replicable ways in laboratory settings. What is an experiment? What is a theory? What would count as testing a theory? When we test a theory, are we trying to prove it, or disprove it? Are the meanings of economic concepts exhausted by their verification conditions, or are economic theories important and meaningful apart from any efforts we make or could make empirically to test them?

For example, Hausman says, “one might argue that preferences and beliefs are in some sense unobservable” (1984, p. 15). But inferential bases for the ascription of preferences can rather unambiguously be observed in laboratory settings. Experimental subjects can be given opportunities to buy and sell “widgets” that are stripped of all properties other than resale value—the widgets are nothing more than entries on a computer-kept ledger. Experimenters specify those resale values (that is, how much money subjects will be paid for any widgets they possess at the end of the experiment). Thus, much information that is hidden outside the laboratory can be known and controlled in a laboratory setting, enabling researchers to draw reasonably well-grounded inferences about subjects’ motives and strategies.

For example, if we interpret subjects as being in a prisoner’s dilemma situation, such that declining to cooperate is a dominant strategy, and then we see subjects cooperating instead, we are free to hypothesize that the situation is not really a prisoner’s dilemma, retreating to a view that by definition subjects will act to maximize their payoff, and therefore by definition subjects will decline to cooperate in a genuine prisoner’s dilemma. In a laboratory setting, we can do better than that. We can specify all the payoffs and communicate them unambiguously to experimental subjects. We can train them over a sequence of trial runs to make sure they understand their situation. Then we can observe and learn. If subjects do not behave as our theories predict, or if 60 percent behave as predicted and 40 percent do not, then so be it. There is no such thing as being in a situation where there is exactly one theory that fits the observed facts. It is a truism in philosophy of science that any given set of observations will be compatible with an infinite number of theoretical explanations.

Nevertheless, what we learn to accept in the laboratory is that subjects do not consistently act to maximize their monetary payoff. They show inclinations to cooperate, to trust, to be “fair,” and so on, that go beyond anything it is reasonable for us to try to explain in terms of the hypothesis that subjects are acting to maximize their monetary payoff. We can even design the experiment so as to yield fine-grained information about why subjects decline to cooperate, when they do. For example, we can suppose that the two main reasons not to cooperate in a prisoner’s dilemma are greed (the preference to get the good for free when one expects others to cooperate in producing the good) and fear (the preference not to cooperate when one expects others not to). In the field, it may be impossible to tell the difference between greed and fear, since all we observe is whether subjects are cooperating. Laboratory experiments, though, can be designed to tell the difference. That is, we can go beyond the hypothesis that everyone will defect when defection is a dominant strategy to test the hypothesis that when people defect, it is because they are afraid their partners will defect, not because they hope to exploit partners who cooperate. In the laboratory there is much defection, but also much cooperation, and much more cooperation when fear is eliminated as a motive for defection, even when the motive to free ride is left untouched, indeed, even when defecting remains a dominant strategy (see Mark, Schmidtz, and Walker 1989). Perhaps this takes us from economics proper into fields such a psychology, sociology, and so on. But economists probably should find move encouraging, inasmuch as it indicates that their simplest behavioral postulates, in virtue of being disconfirmable (and sometimes disconfirmed), are at the same time fruitful and interesting.

ENVIRONMENTAL ECONOMICS AND ENVIRONMENTAL PHILOSOPHY. Environmental economists are presumed to be advocates of conservationist “wise use” policies, where environmental philosophers are presumed to be advocates of preservationist “no use at all” policies regarding scarce environmental assets. Perhaps the picture never was this simple, but in any case it is changing. Philosophers like Bryan Norton (1991) and Mark Sagoff (2004) are, in various ways, going beyond simple dichotomies in search of new policy paradigms that make sense from both long-term environmental and medium-term economic perspectives.

THE PSYCHOLOGICAL AND INSTITUTIONAL PREREQUISITES OF MARKET ECONOMIES. Since the fall of the Soviet Union, and the subsequently mixed results of formerly communist countries in establishing market economies, wiser and humbler economists have been exploring the idea that market economies cannot be invented, manufactured, or decreed but must instead be treated as organically evolving systems that grow over time. Citizens of the former Soviet Union, it seems, do not understand instinctively how to behave as market

agents. If they grow up in a world where the only examples of entrepreneurship involve bribery and theft, then they will think of entrepreneurs as a species of predator and will not grasp the concept of mutual advantage in the way that owners of small businesses in free countries do.

ECONOMICS OF CULTURE. Economists likewise have begun to turn their attention to the intertwined evolution of economy and culture. Explosions of cultural innovation seem to occur in cities that are at the same time undergoing explosive growth as worldwide commercial centers.

BEYOND HOMO ECONOMICUS. As noted earlier, the postulate of self-interest is most illuminating when treated as a testable empirical hypothesis, so that when behavior fails to conform, it will not simply be ignored but will instead be seen as of scientific interest. One of many cases in point is the "ultimatum" game. Two subjects are assigned the task of dividing a fixed amount of money. The first subject, Proposer, makes a proposal about how to divide the money. The second subject, Responder, has two options: reject the proposal, in which case neither subject gets anything; or accept the proposal, in which case the subjects split the money as proposed. The game is not repeated, so a Responder who is rational as per the Homo economicus model ought to accept any proposal that offers Responder a positive payoff. A bit more tenuously, Proposer, expecting Responder to be rational as per the Homo economicus model, ought to offer responder the smallest possible positive payoff. In fact, neither of these predictions is born out in the laboratory. Proposers most commonly offer to split the money fifty-fifty. Cristina Bicchieri (2005) reports that in a variety of trials and conditions, including in different cultures, responders tend to reject offers below 20 percent of the total, even when the stakes are substantial relative to prevailing wage rates in the subjects' community.

NEUROECONOMICS. Kevin McCabe tested a variation of the ultimatum game while recording subjects' brain activity with functional MRI (Kevin McCabe, et al, 2001). In some trials, subjects were informed that their partner was a computer program playing a fixed probabilistic strategy; these were paired with trials where subjects were informed that their partner was another human subject. Roughly half the subjects chose not to cooperate with human partners. Their brain activity was similar in the computer partner and human partner trials. Subjects who did cooperate, roughly half the total, showed markedly greater brain activity in the prefrontal cortex. The implication: subjects who cooperate are not treating trials with human partners as situations calling simply for payoff calculation. The prefrontal cortex is thought to be the part of the brain dealing with social situations, not with arithmetic calculation. Cooperators evidently are treating the transaction not only as an economic exchange but also as a social exchange, calling for empathetic understanding of the motivations of another agent. It is too early to say where this line of research is leading, but it suggests we may hope some day for a unified explanation of departures from the postulate of self-interest, including the above-reported departures from dominant strategy in the prisoner's dilemma. That is, subjects who do not conform to the predictions of Homo economicus models may be departing from the models in virtue of perceiving the situation as calling not for calculation of their possible payoffs, but instead for something else, such as an exchange of tokens of mutual respect. In any case, our sensitivity to economic motives is a variable. What gets us to focus on the economic bottom line—the numbers—rather than on friendships, grudges, self-esteem, status, and so on, is interestingly complex.

See also Decision Theory; Game Theory; Philosophy of Social Sciences.

Bibliography

ECONOMICS AS MORAL PHILOSOPHY

Buchanan, James, and Gordon Tullock. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press, 1962.

Coase, Ronald. *Essays on Economics and Economists*. Chicago: University of Chicago Press, 1994.

Gauthier, David. *Morals By Agreement*. New York: Oxford University Press, 1986.

Griswold, Charles. *Adam Smith and the Virtues of Enlightenment*. New York: Cambridge University Press, 1999.

Hausman, Daniel M., and Michael S. McPherson. *Economic Analysis and Moral Philosophy*. Cambridge, U.K.: Cambridge University Press, 1996.

Hayek, F. A. *Road to Serfdom*. Chicago: University of Chicago Press, 1994.

Kirzner, Israel. *Competition and Entrepreneurship*. Chicago: University of Chicago Press, 1973.

Mises, Ludwig. "Economic Calculation in the Socialist Commonwealth." In *Collectivist Economic Planning: Critical Studies of the Possibilities of Socialism*, edited by F. A. Hayek, 87–130. London: Routledge & Kegan Paul, 1935.

Plato. *Republic and Other Works*. Translated by B. Jowett. New York: Anchor Books, 1973.

Rosenberg, Alexander. *Philosophy of Social Science*. Boulder, CO: Westview Press, 1988.

Sen, Amartya Kumar. *On Ethics and Economics*. Oxford, U.K.: Blackwell, 1987.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations* [1776], edited by R. H. Campbell and A. S. Skinner. 2 vols. Indianapolis: Liberty Fund, 1981.

Smith, Adam. *The Theory of Moral Sentiments* [1759], edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Fund, 1976.

ECONOMICS AS SCIENCE

Friedman, Milton. "The Methodology of Positive Economics." In *Philosophy of Economics*, edited by Daniel M. Hausman, 210–244. New York: Cambridge University Press, 1984.

Hausman, Daniel M., ed. *The Philosophy of Economics*. New York: Cambridge University Press, 1984.

Little, Daniel, ed. *On the Reliability of Economic Models: Essays in the Philosophy of Economics*. Boston: Kluwer, 1995.

Little, Daniel. *Varieties of Social Explanation*. Boulder, CO: Westview, 1991.

McCloskey, Deirdre. *The Rhetoric of Economics*. 2nd ed. Madison: University of Wisconsin Press, 1998.

Robbins, Lionel. *An Essay on the Nature and Significance of Economic Science*. 2nd ed. London: Macmillan, 1935.

EXPERIMENTAL ECONOMICS

Mark, Isaac R., D. Schmidtz, and James M. Walker. "The Assurance Problem in a Laboratory Market." *Public Choice* 62 (1989): 217–236.

Smith, Vernon. "Constructivist and Ecological Rationality in Economics." Nobel Prize lecture, December 8, 2002.

ENVIRONMENTAL ECONOMICS AND PHILOSOPHY

Norton, Bryan. *Toward Unity among Environmentalists*. New York: Oxford University Press, 1991.

Sagoff, Mark. *Price, Principle, and the Environment*. New York: Cambridge University Press, 2004.

Schmidtz, David. "When Preservationism Doesn't Preserve." *Environmental Values* 6 (1997) 327–339.

PSYCHOLOGICAL AND INSTITUTIONAL PREREQUISITES OF MARKET ECONOMIES

Brennan, Geoffrey, and Philip Pettit. *The Economy of Esteem: An Essay on Civil and Political Society*. Oxford: Oxford University Press, 2004.

Fukuyama, Francis. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press, 1996.

North, Douglass. *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press, 1990.

ECONOMICS OF CULTURE

Appiah, Kwame Anthony. *In My Father's House: Africa in the Philosophy of Culture*. New York: Oxford University Press, 1992.

Cowen, Tyler. *Creative Destruction: How Globalization Is Changing the World's Cultures*. Princeton, NJ: Princeton University Press, 2002.

Cowen, Tyler. *In Praise of Commercial Culture*. Cambridge, MA: Harvard University Press, 1998.

BEYOND HOMO ECONOMICUS

Bicchieri, Cristina, *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press, 2005.

Guth, W., R. Schmittberger, and B. Schwarze. "An Experimental Analysis of Ultimatum Games." *Journal of Economic Behavior and Organization* 3 (1982): 367–388.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richar McElreath. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 91 (2001): 73–79.

Kahneman, Daniel, and Amos Tversky. "Choices, Values, and Frames." *American Psychologist* 39 (1984): 341–350.

Schmidtz, David. *Rational Choice and Moral Agency*. Princeton, NJ: Princeton University Press, 1995.

Simon, Herbert. *Models of Bounded Rationality and Other Topics in Economics*. Cambridge, MA: MIT Press, 1982.

NEUROECONOMICS

McCabe, Kevin, Daniel Houser, Lee Ryan, Vernon Smith, and Theodore Trouard. "A Functional Imaging Study of Cooperation in Two-Person Reciprocal Exchange." *Proceedings Of the National Academy of Science* 98 (2001): 11832–11835.

David Schmidtz (2005)

PHILOSOPHY OF EDUCATION, EPISTEMOLOGICAL ISSUES IN

Epistemological issues have always enjoyed a central place (along with metaphysical, moral, and social/political issues) in philosophical thinking about education. In the entry "Philosophy of Education, History of" in this encyclopedia, Kingsley Price skillfully treats the entire history of the subject, from the Presocratics to John Dewey. This entry covers the intervening decades, focusing on epistemological issues.

By the time of Dewey's death in 1952, philosophy in the English-speaking world was becoming increasing dominated by the analytic movement, which emphasized as methodological matters the importance of clarity, careful analysis, rigorous argumentation, and detailed attention to language, and philosophy of education was no exception to this general trend. The key figures in the development of analytic philosophy of education were Israel Scheffler in the United States, and Richard Peters and Paul Hirst in the United Kingdom. While their work exemplified two different strands of analytic philosophy—Peters and Hirst worked in the 'ordinary language' tradition of analytic philosophy, which emphasized the explication of meanings as manifested in ordinary language, while Scheffler's brand of analysis took more seri-

ous account of logic and its associated formal techniques, and was more inclined to overrule ordinary language when theoretical improvement could be so gained—both sought to bring a level of clarity and sophistication to an area of philosophy that did not always enjoy these, and to integrate philosophy of education with general philosophy. The following discusses some central epistemological issues in philosophy of education.

EPISTEMIC AIMS OF EDUCATION

What is the fundamental epistemic aim of education? For educators, is the highest aim that of *truth* and the bringing about of true belief in students? Or is it, rather, *rationality* and the fostering of rational (or perhaps *justified*) belief? Perhaps that aim is the more encompassing one of *knowledge*, which includes and integrates both of the previous possibilities? Or could the aim be that of enhancing student *understanding*? Each of these has its advocates and deserves brief explication.

TRUTH. The most important contemporary advocate of truth as the fundamental epistemic aim of education is Alvin Goldman (1999). On his '*veritistic*' view, the fundamental epistemic aim of education is the production of true belief in students, along with the development of student ability to discover new (to them) truths by way of inquiry.

Goldman's view has much to recommend it, although it faces some difficulties as well. First, not all modes of transmitting truths to students—brainwashing, indoctrination, deception, and the like—are educationally acceptable, despite their efficacy in producing true belief. Second, from the educational point of view it matters not only *that* students believe truths, but also *on what basis* they believe them: Mindless or otherwise unjustified true belief is not typically the intended aim of educational activities, despite the truth of the relevant student belief. Third, the general failure to enjoy 'direct access' to truth suggests that the relevant educational aim is not true belief, but rather student ability to estimate or judge the truth competently (Scheffler 1965, p. 54). These difficulties suggest that the fundamental epistemic aim of education is not true belief but rather *rational* belief.

RATIONALITY/CRITICAL THINKING. The great majority of historically significant philosophers of education have endorsed the fostering of student rationality, or its educational cognate *critical thinking*, as the (or at least a) basic epistemic aim of education. On this view, educational efforts ought to strive to foster the abilities and dispositions conducive to rational student belief, the latter conceived as belief properly based on *good reasons*. Accordingly, educational activities are epistemically successful just to the extent that they result in enhanced student ability to evaluate candidate reasons for belief fairly and competently, and strengthened student disposition both to so evaluate and to believe accordingly. The dispositional or 'critical spirit' element of the view connects epistemic matters with matters of *character*, and the view as a whole is justified in terms of an appeal to the moral duty to treat students with *respect as persons*: Treating students with respect requires educating them in ways intended to foster critical thinking and thereby their autonomy, independence of judgment, and ability to shape—as far as possible—their own minds and lives (Siegel 1988, 1997; Bailin and Siegel 2003).

Although versions of this view enjoy considerable support from both philosophers (historical and contemporary) and educators, it faces the important objection that rationality and critical thinking are arguably best thought of not as ends in themselves, but rather instrumentally, as *means* to the end of true belief: Why think that the former are epistemically valuable, other than as an effective route to truth (Goldman 1999)? This raises two questions: Can rationality/critical thinking be thought to be valuable other than instrumentally, as a means to truth? Can the virtues of both these putative epistemic ends of education be suitably combined (Siegel 2005b)?

KNOWLEDGE. Taking the fundamental epistemic aim of education to be knowledge has the advantage that, suitably understood in its 'strong' sense, that aim includes both truth and rationality/justification. This better captures the sense in which educators are concerned with the fundamental epistemic aim *of education*, since, from the educational point of view, mere true student belief is less adequate than true belief that is justified, rational, or otherwise based on good reasons; and justified or rational but nevertheless false belief is less adequate than such belief that is also true. This view, that knowledge (in the 'strong' sense that includes both truth and rational justification as conditions of knowledge) is the fundamental epistemic aim of education, is defended by several contemporary authors (Scheffler 1960, 1965, 1989; Adler 2003; Siegel 2005b). It appears to capture the strengths of both the previous views and to meet the objections to them outlined above.

UNDERSTANDING. The way in which all these putative epistemological aims of education involve student *under-*

standing is less than crystal clear, and a plausible case has been made by Catherine Z. Elgin, furthering a philosophical approach pioneered by Nelson Goodman, that it is the latter—rather than truth, rationality or knowledge—that deserves to be regarded as the fundamental epistemic aim of education (Elgin 1999a, 1999b). Whether or not understanding can be integrated successfully with the other main proposed epistemic aims of education, can be shown to be less fundamental than those others, or deserves pride of place as the fundamental such aim, remains the subject of ongoing debate.

TESTIMONY, TRUST, AND TEACHING

Should students believe what their teachers tell them? Arguably, they should, and recent work on the epistemology of testimony suggests as much (Goldman 1999). But student belief in the otherwise unsupported testimonial pronouncements of their teachers conflicts with the view that critical thinking is an important aim of education, since such belief seems clearly enough not to be belief based on reasons subjected to critical scrutiny by the believer/student. Live issues concerning the epistemology of testimony are helpfully illuminated by the educational case. This is obviously not the place to tackle the broad question of the epistemology of testimony. But the educational case concerning testimony in the classroom setting deserves brief comment.

First, it is important to be clear about the sort of student under consideration. Very young children/students cannot evaluate the testimonial pronouncements of their teachers; they lack the cognitive capacity to do so. Such capacity develops gradually; before it is substantially achieved, trust in their teachers' pronouncements seems unproblematic. But how long is the period during which students enjoy such a holiday from the ordinary demands of responsible oversight of their cognitive lives? This is, at least in part, an empirical matter concerning the facts of psychological/cognitive development. Once such development has taken place and students are able to monitor and evaluate the epistemic standing of their beliefs, do those testimony-based beliefs enjoy positive justificatory status if the only thing the student can say in their defense is "my teacher said so"? Here the *reductionist* (who, like David Hume, holds that testimony-based beliefs are justified only if that justification can be reduced to testimony-independent good reasons to trust the speaker's testimony on a given occasion) and the *antireductionist* (who, like Thomas Reid, holds that testimony is itself a basic source of justification) will divide in the predictable way. But the latter will have to explain why the aim of fostering critical thinking (discussed above) can or should be abandoned in the case of teacher testimonial pronouncements, and how so abandoning it can be reconciled with the duty to treat students with respect as persons. It is not meant here to suggest that the antireductionist is doomed to failure. But the educational case does provide a sharp test case of epistemological views concerning testimony.

It should also be noted that the case in which students have nothing to justify their belief in the testimonial pronouncements of their teachers other than the pronouncements themselves is arguably relatively rare and certainly not typical. Just as believers typically have considerable evidence for the general reliability of testimony, so that their trust in testimonial pronouncements is accompanied by testimony-independent evidence that sanctions such trust (Adler 2002), so, too, do students typically have such evidence concerning their teachers' pronouncements. For even when students begin a class with no testimony-independent reason for believing what their teacher tells them, as the class proceeds and students observe their teacher lecture, explain, answer questions, and extemporize, such observation itself provides testimony-independent reason for trusting the teacher's testimonial pronouncements concerning the subject matter at hand (Siegel 2005b).

INDOCTRINATION, TEACHING, AND BELIEF

Questions concerning the places of testimony and trust in the classroom lead naturally to questions concerning teaching and indoctrination. During the decades in which the analytic approach dominated the field, philosophers of education devoted considerable effort to the analysis of the concept of indoctrination (Snook 1972, Spiecker and Straughan 1991, Siegel 1988). The theories of indoctrination developed then divided into three broad types, which located indoctrination in either the *aim* or *intention* of the teacher/indoctrinator (namely, to get students to believe matters independently of the evidence for them), the *method* employed in transmitting the relevant beliefs (that is, in a way that precludes student questioning or demand for reasons), or the character or *content* of the doctrines transmitted (that is, content that does not admit of rational support or that is believed independently of such support). These three ways of understanding indoctrination have in common that (successful) indoctrination results in beliefs that students do not, will not, and/or cannot subject to critical scrutiny. That is, indoctrination, when successful, results

in student acquisition of both specific beliefs and of habits or dispositions to believe independently of the evidential status of the indoctrinated beliefs. In this way indoctrination appears to be incompatible with most of the epistemic aims of education canvassed above, most obviously that concerning the fostering of rationality/critical thinking.

However, the seemingly obvious view that educators should eschew indoctrination in favor of more respectable epistemic educational practices is not so quickly established. First, can education be nonindoctrinating, either in principle or in practice, or is indoctrination inevitable? One might think it unavoidable since, as was suggested above, at least at early stages of development, students do not in fact have the cognitive capacity to challenge, evaluate, or critically consider that which they are taught. If it is for this reason unavoidable, is indoctrination as a consequence not necessarily or always a bad thing, something to be avoided by responsible educators? After all, if students are incapable of subjecting teacher testimonial pronouncements to critical scrutiny until after a certain cognitive-developmental stage is reached, language and concepts acquired, and an appropriate level of reasoning ability attained, it is hard to see how teachers can help bring students to the point at which they can exercise their critical abilities except by indoctrinating them. The alternative view, namely, that indoctrination is avoidable, requires a distinction between indoctrination and nonindoctrinating belief inculcation, but such a distinction is often thought to be controversial (Siegel 1988, 2005b).

Second, (why) should we value educational processes that result in student ability to subject candidate beliefs to critical scrutiny? Philosophers of education who differ in their answers to the question of the fundamental epistemic aim of education will differ in their answers to this one. Veritists will answer that we should value such processes because that ability will increase student acquisition of true belief. Advocates of critical thinking will answer, rather, that we should value them because student acquisition of rational/justified belief will be enhanced, and, moreover, that desirable dispositions will be fostered. Advocates of knowledge (in the strong sense) will embrace both these answers. Those who think indoctrination inevitable may well deny that we should value such processes at all (and may deny that there are, in fact, any such processes).

OPEN-MINDEDNESS, BELIEF, AND COMMITMENT

A further epistemic good related to critical thinking, often regarded as a basic educational aim, is that of *open-mindedness*: Roughly, the ability to regard one's beliefs as fallible and subject to rational rejection or revision in light of evidence and critical reflection (Hare 1979, 1985). But how can open-mindedness be reconciled with the aim of fostering student knowledge or rationality, given that the latter involve student *belief*? That is, how can students be expected both to believe those belief-candidates that reasons and evidence indicate are worthy of belief, and at the same time to remain open-minded about those very beliefs? This tension is insightfully addressed by Jonathan Adler (2004), who urges that open-mindedness be conceived as a meta-attitude toward one's beliefs rather than as a weakening of one's degree of belief or a weakened commitment toward the beliefs themselves, and that it be understood in terms of our general interest in attaining knowledge; he relates these matters to other fundamental issues concerning tolerance, autonomy, and authority that have long animated philosophers of education.

FURTHER TOPICS

There is a range of further issues concerning epistemological dimensions of education that should be mentioned, even though they cannot be addressed in detail here. They include the following issues.

FURTHER ISSUES CONCERNING CRITICAL THINKING. Partly because of its enduring status as a favored educational ideal, considerable philosophical energy has been expended on issues concerning critical thinking other than those already addressed. A particularly animated discussion involves the question of its *generalizability*: Is critical thinking generalizable—that is, applicable to a broad range of topics, domains, and issues—or is it rather *subject-specific*, such that critical thinking in one domain or discipline is importantly different from critical thinking in other areas? A range of views on the question can be found in *The Generalizability of Critical Thinking* (Norris 1992). A further issue is the place of *domain-specific knowledge* in critical thinking; here William Hare (1995) is particularly helpful. The relation between critical and creative thinking has also attracted considerable attention, with some arguing that these are fundamentally distinct and others arguing against such a sharp distinction. The topic has been insightfully treated in a series of works by Sharon Bailin,

who challenges the distinction; see Bailin and Siegel 2003 and references therein.

CURRICULUM. It seems obvious enough that the curriculum should contain that knowledge/information thought to be most important for students to have, but the value and epistemological status of particular sorts of curricular content is controversial. Should a given subject, say mathematics, enjoy pride of place in the curriculum because it is in some sense intellectually central, or is its place secured by virtue of its practical importance or in some other way? More broadly, do particular content areas—science, language and literature, history, and the like—deserve their place in the curriculum because they constitute distinct "forms of knowledge" that are in some sense epistemologically fundamental, intrinsically important, and therefore the stuff of which all "liberally educated" students should be familiar (Hirst 1974)? Can this "forms of knowledge" view of traditional school subjects be sustained (Phillips 1987, pp. 120–136)? Moreover, does this idea of "liberal education" overemphasize the traditional and theoretical to the detriment of the practical, and/or does it reflect a culturally biased "Eurocentric" view of reason, knowledge, and education's character and priorities (Siegel 1997, Bailin and Siegel 2003)?

TEACHING AND LEARNING. How should teaching and learning be conceived and the former conducted? The issues here are many and complex and depend for their resolution on psychological matters as well as on philosophy of mind and other areas of philosophy, yet they are rightly thought to be epistemological (in part) in so far as teaching is thought to involve knowledge transmission and the development of the ability to acquire knowledge, and learning is thought to involve such acquisition. (Passmore 1980, Pearson 1989, Hare 1993).

"GROUP EPISTEMOLOGIES" AND FEMINIST, MULTICULTURALIST, AND POSTMODERNIST CHALLENGES TO IDEALS OF REASON IN EDUCATION. By the 1970s analytic philosophy began to lose its dominant position in the field and, again, philosophy of education followed the trend established in the parent discipline. The rise of Feminism, Multiculturalism, and Postmodernism brought with them important challenges to traditional views concerning the universality and neutrality of 'reason' and rationality and, indeed, to the nature of knowledge itself. While space precludes serious attention to these challenges here, or even a clear articulation of the issues, they are an important part of the contemporary scene in the philosophy of education. (For further discussion and references, see Bailin and Siegel 2003; Siegel 1997, 2004, 2005).

See also Dewey, John; Feminist Epistemology; Multiculturalism; Philosophy of Education, Ethical and Political Issues In; Philosophy of Education, History of; Postmodernism.

Bibliography

Adler, Jonathan E. *Belief's Own Ethics*. Cambridge, MA: MIT Press, 2002.

Adler, Jonathan E. "Knowledge, Truth, and Learning." In *A Companion to the Philosophy of Education*, edited by Randall Curren. Oxford: Blackwell, 2003, p. 285–304.

Adler, Jonathan E. "Reconciling Open-mindedness and Belief." *Theory and Research in Education* 2 (2) (2004): 127–142.

Bailin, Sharon, and Harvey Siegel. "Critical Thinking." In *The Blackwell Guide to the Philosophy of Education*, edited by Nigel Blake, Paul Smeyers, Richard Smith, and Paul Standish. Oxford: Blackwell, 2003, p. 181–193.

Elgin, Catherine Z. "Education and the Advancement of Understanding." In *Proceedings of the 20th World Congress of Philosophy*. Vol. 3, edited by David M. Steiner. Philosophy Documentation Center, 1999a, p. 131–140.

Elgin, Catherine Z. "Epistemology's Ends, Pedagogy's Prospects." *Facta Philosophica* 1 (1999b): 39–54.

Goldman, Alvin I. *Knowledge in a Social World*. Oxford: Oxford University Press, 1999.

Hare, William. "Content and Criticism: The Aims of Schooling" *Journal of Philosophy of Education* 29 (1) (1995): 47–60.

Hare, William. *In Defence of Open-mindedness*. Kingston: McGill-Queen's University Press, 1985.

Hare, William. *Open-mindedness and Education*. Kingston: McGill-Queen's University Press, 1979.

Hare, William. *What Makes a Good Teacher? Reflections on Some Characteristics Central to the Educational Enterprise*. London, ON: Althouse, 1993.

Hirst, Paul H. *Knowledge and the Curriculum: A Collection of Philosophical Papers*. London: Routledge & Kegan Paul, 1974.

Hirst, Paul H., and R. S. Peters. *The Logic of Education*. London: Routledge & Kegan Paul, 1970.

Norris, Stephen P., ed. *The Generalizability of Critical Thinking: Multiple Perspectives on an Educational Ideal*. New York: Teachers College Press, 1992.

Passmore, John. *The Philosophy of Teaching*. Cambridge, MA: Harvard University Press, 1980.

Pearson, Allen T. *The Teacher: Theory and Practice in Teacher Education*. New York: Routledge, 1989.

Phillips, D. C. *Philosophy, Science, and Social Inquiry: Contemporary Methodological Controversies in Social Science and Related Applied Fields of Research*. Oxford: Pergamon Press, 1987.

Scheffler, Israel. *Conditions of Knowledge: An Introduction to Epistemology and Education*. Glenview, IL: Scott, Foresman, 1965.

Scheffler, Israel. *The Language of Education*. Springfield, IL: Charles C. Thomas, 1960.

Scheffler, Israel. *Reason and Teaching*. 2nd ed. Indianapolis, IN: Hackett, 1989.

Siegel, Harvey. *Educating Reason: Rationality, Critical Thinking, and Education*. London: Routledge, 1988.

Siegel, Harvey. "Epistemology and Education: An Incomplete Guide to the Social-Epistemological Issues." *Epistéme* 1 (2) (2004): 129–137.

Siegel, Harvey. "Israel Scheffler." In *Fifty Modern Thinkers in Education: From Piaget to the Present*, edited by Joy A. Palmer. London: Routledge, 2001, 142–148.

Siegel, Harvey. *Rationality Redeemed? Further Dialogues on an Educational Ideal*. New York: Routledge, 1997.

Siegel, Harvey. "Truth, Thinking, Testimony, and Trust: Alvin Goldman on Epistemology and Education." *Philosophy and Phenomenological Research* (2005).

Snook, I. A., ed. *Concepts of Indoctrination: Philosophical Essays*. London: Routledge & Kegan Paul, 1972.

Spiecker, Ben, and Roger Straughan. *Freedom and Indoctrination in Education: International Perspectives*. London: Cassell, 1991.

Harvey Siegel (2005)

PHILOSOPHY OF EDUCATION, ETHICAL AND POLITICAL ISSUES IN

Education is the promotion of learning and development. Educational activities include attending to explanations, lectures, or demonstrations, but it does not follow that teaching or direct instruction is the whole or the essence of education. Education also involves the communication of care and the transmission of elements of at least one culture, but these, too, are only part of what constitutes education. Additionally, education is a form of governance; *to educate* has always meant "to rear, bring up, instruct, train, discipline, develop," but its Latin root, *ēducāre*, is related to *ēducĕre* (from *ē*, "out," and *dūcĕre*, "to lead"), a term of governance. The terms *pedagogy* and *pedagogue* (schoolteacher) derive similarly from the Greek *paidagôgos*, a term of governance (from *paidion*, "child," and *agô*, "to lead") referring to the household slave who supervised the children and led them out into the city from one teacher and place of learning to another. The modern term *governess*, signifying a woman employed to educate the children of a household, is similarly and conspicuously a term of governance.

As a form of governance, education requires justification, and it entails responsibilities, aims, a manner of going about its business, and substance or a communicated content. These are the fundamental aspects of governance, and the philosophy of education can be organized by categories corresponding to them: the authority to educate (justification), the adequate and equitable provision of education (responsibilities), the aims of education (aims), pedagogy and educational ethics (manner), and curriculum (substance or content).

On this account of the divisions of philosophy of education, it becomes evident that an ethic of governance would provide a unifying normative structure. The most obvious and durable illustration of this is an ethic of respsect for persons as self-determining agents. Ethics of this kind have dominated philosophy from the time of Socrates, and they have implications for each of the five named aspects of governance and for each of the five corresponding divisions of philosophy of education. The primary aim and responsibility of educators is to promote autonomy or effective self-determination, and to do so equitably, displaying equal respect within their sphere of educational authority. The scope of the educational authority they possess, the manner in which they exercise that authority, and the content of the education they provide will in turn be limited and shaped by the character of this responsibility. They endeavor to cultivate the intellectual and moral virtues essential to good judgment, to nurture capabilities that will provide the basis of lives worth living, and to enable each student to understand the circumstances of his or her own life and the possibilities that lie before him or her. While promoting autonomy or effective self-determination in such ways, educators teach in a manner respectful of their students and the values inherent in the subjects they teach.

An influential alternative to such an ethic of respect is the ethic of care championed by Nel Noddings (1992) and others. Considered as an ethic of education, it assigns great importance to caring for students. It proposes the development of caring in students as the central purpose or aim of education and sets forth a conception of curriculum based not on the diverse forms of disciplinary knowledge but on the diverse forms of human developmental potentials and diverse "centers of care" or objects of potential interest and devoted attachment. Advocates of this view are less clear about its implications for matters of educational justice and authority, but in addressing the latter, they begin from the presumption that care and control are incompatible. They concede that an ethic of care does not constitute a comprehensive moral point of view, but the debate, which originated not in moral theory but in the psychology of moral development, has

been framed as an opposition and subsequent reconciliation between justice and care.

An alternative would be to hold that the literature of care offers not a competing ethic or ethical theory, but a cluster of important empirical observations about the fundamental place in human development and well-being of being cared for, coming to care about and for oneself, and forming attachments. These tenets have been acknowledged by liberal theorists who regard a deontological ethic of respect as morally fundamental. Examples include the attention to continuity and quality of relationships in schools in the work of Randall Curren (2000, 2003) and conceptions of teaching and the curriculum as providing potential objects of attachment and fulfillment, as discussed in the work of Kenneth Strike (2003) and Harry Brighouse (2005).

Within the educational framework established by an ethical-political orientation, there are roles to be played by guiding norms of other sorts, such as epistemic rationality, craftsmanship, and artistry. If self-determination is enhanced by knowledge and understanding, then curricula must communicate, and teachers display respect for, the epistemic norms pertaining to knowledge and understanding. If the promotion of autonomy or meaningful choice among satisfying lives requires that students have opportunities to experience and develop competence in pursuits that are fulfilling and allow them to make their way in the world, then curricula must communicate, and teachers display respect for, the norms of craft and artistry proper to such pursuits.

Attempts have been made to undermine the distinction between epistemic and moral-political norms that is assumed here. Postmodernists and some varieties of feminists and neo-Marxists hold that the norms of epistemic rationality, at least in their familiar forms, are aspects of systems of oppression and have no objective standing. Such views have had many defenders within the philosophy of education in recent years, but the moral principles they appear to rely on are no more radical than those of the dominant liberal-democratic tradition, which has itself always been at least latently egalitarian. What distinguishes these contemporary critical stances is the assumptions of fact they employ, their salutary attention to previously neglected forms of inequality and disrespect, and—more problematically—their epistemic and metaphysical doctrines.

Although many of the ethical and political issues in philosophy of education were addressed by R. S. Peters and others in, and opposed to, the analytical philosophy of education movement of the 1960s and 1970s, philosophical exploration of them has become more common since the 1980s. This growth of interest in such issues includes debates about parental choice in schooling, public support for religious schools, moral education, inclusion of students with disabilities in regular classrooms, accountability and high-stakes testing, affirmative action in university admissions, and the limits of academic freedom.

EDUCATIONAL AUTHORITY

The question of how to apportion authority over education between parents and public authorities has become important since the early 1980s, as parents in the United States have challenged public school curricula and have increasingly chosen home schooling, usually on religious grounds. What role should parental wishes and rights play in determining the content of public education? When it comes to regulating private, religious, and home schooling, how are parents' interests in the faith and character of their children to be balanced against the protection of children's interests and the need to prepare them for citizenship in a multicultural society? Is it acceptable to exempt religious schools from laws that protect girls and women from discrimination on the basis of sex?

William Galston and other defenders of wide parental discretion argue that parents can be trusted more than government authorities to know and protect their children's interests, that parents have a strong and legitimate interest in transmitting their values to their children, and that it is in the interest of children to be educated in the "thick" cultural traditions that faith communities can provide but that public institutions constrained by requirements of neutrality cannot. James Dwyer (1998) and others have argued in response that it is incoherent to attribute to parents an individual liberty that entails a right to control or predetermine the life course of another person, even a child. Amy Gutmann (1987), Eamonn Callan (1997), Stephen Macedo (2000, 2002), and others have argued that respect for reasonable pluralism cannot be secured by unlimited accommodation of the wishes of parents whose own cultural communities are intolerant. Civic virtues of respectful and reasoned engagement with the views and values of others must be educationally nurtured if a political culture of tolerance and mutual respect is to survive, and it follows from this that educational policy must favor, if not absolutely insist upon, universal standards of civic education. Dwyer, Brighouse, Meira Levinson (1999) and others argue that liberal respect for children as persons in their own right requires policies that ensure that all chil-

dren enjoy an education that introduces them to a variety of cultural and ethical traditions and enables them to think critically about the circumstances and conduct of their own lives.

A related debate over school choice and privatization has taken on significance as schemes to promote parental choice among schools (for example, providing government vouchers redeemable for all or part of tuition) have spread to many parts of the world. Defenders of such schemes have argued that they are necessary to eliminate the differential impact of ability to pay on the freedom of parents to practice their religions, but also that a free market in educational services would promote efficiency and superior educational results. The debate is fraught with empirical speculation on all sides, but Colin Crouch (2003) has made a strong case for the view that privatization would abandon the idea that education is a right of citizenship, and others have addressed the ethical and political principles involved in ways that set the empirical issues aside. Curren has examined the grounds on which a public system of schools might be considered necessary, and he and Brighouse have arrived at similar requirements of justice for any system of education to be deemed acceptable (Curren 2000, Brighouse 2000). Both argue that some choice schemes might satisfy those requirements, that responsibility lies with the state to ensure that those requirements are met, and that public authority over education must be retained at least to the extent necessary to fulfill that responsibility.

A third debate concerns the professional authority of educators themselves. The authority to teach is typically granted through processes of certification and selective employment. But once teachers are employed, by what means are they, schools, and those who supervise them to be held accountable for their performance? Debate has focused on the promise and perils of high-stakes testing as a mechanism of accountability, and there is clearly much of ethical significance at stake. To what extent do extensive testing regimes undermine student motivation to learn? To what extent do they limit the exercise of sound professional judgment and thereby undermine good teaching?

EDUCATIONAL RESPONSIBILITIES

How are educational adequacy and equity to be understood? One debate concerns the kind of educational equality to be achieved and the degree to which equality is a requirement of justice. The major divide has been between those who argue that schooling is to be distributed so as to promote equality of opportunity to live well and those who defend one or another threshold of educational adequacy. Best known among the latter views is Gutmann's argument that in order for the rights of citizenship to be meaningful, every citizen must be provided an education sufficient to make possible *effective* participation in democratic processes (Gutmann 1987).

Another area of lively debate concerns the diversity of students served by schools. The main topics have been religious diversity and the free exercise of religion, gender equity, racial justice and antiracist education, the rights of linguistic minorities, and justice for students with disabilities.

As regards higher education, the focus has been on access or who gains admission. The issue of whether the use of standardized admissions tests such as the SAT (Scholastic Assessment Test) is racially discriminatory has been examined in detail by Robert Fullinwider and Judith Lichtenberg (2004), and countless philosophers have contributed to the debate over the merits of affirmative action in admissions as a way to promote racial and gender equity.

All such views are dismissed as insufficiently transformative, socially and politically, by Paulo Freire and other advocates of revolutionary pedagogies. Because they view the content of conventional schooling as inherently exclusionary and oppressive in ways that sustain unjust regimes, they hold that justice demands forms of teaching that liberate oppressed populations by promoting critical consciousness and action.

EDUCATIONAL AIMS

Does the aim of educating children for their own good conflict with the aim of educating them for the common good? Is the point of transmitting culture to sustain the culture, to benefit the child, or both? Is the point of civic education to stabilize governments that may be corrupt, to prepare citizens to be vigilant in discouraging government corruption, or both? Is the point of education to promote a thriving economy, to enable the child to earn a living, or both? For example, if the economy needs more engineers, how far can schools go in developing the required science curriculum in a preprofessional direction without violating the spirit of a "general" education? What makes the potential for conflict more than conjectural is the existence of other models of the science curriculum. Instruction in science might aim for a broad humanistic and historical understanding of science or an appreciation of the relationships between science, technology, and society; and such aims would not require the

emphasis on mastery of equations and their application that is characteristic of preprofessional instruction.

The hope of reconciling education's worthy aims has rested largely with the enterprise of identifying a highest aim. The dominant choice through much of the Western tradition has been fostering good judgment in matters both public and private; but the dominant choice in recent decades has been autonomy. Although its meaning is often not well defined, *autonomy* seems to signify much the same thing as practically applied good judgment. The coherence and adequacy of the concept of autonomy have been questioned, usually on the grounds that it ignores the social context of personal identity, choice, and efficacy. Defenders of autonomy argue that the metaphysical assumptions of autonomy are not what critics suppose.

PEDAGOGY, DISCIPLINE, AND THE ETHICS OF TEACHING

The landscape of pedagogy has been dominated by different versions of the contrast between pedagogies of content delivery and pedagogies of critical thinking, some more politically charged than others. Friere frames this as a contrast between the "banking" and "problem solving" models, others as a contrast between transmission and construction(ism), and still others as a contrast between teaching that does or does not promote active learning and critical thinking. Defenders of problem solving, constructionist, and critical-thinking pedagogies all offer ethical and emancipatory rationales.

The matter of how coercive classroom management should be has been discussed in connection with pedagogy, classroom dialogue, and theories of motivation and basic psychological needs. A key issue is whether the organization of work and social life in the classroom creates the opportunities for all students to satisfy their basic psychological needs in acceptable ways. If it does, then problems of classroom management will be small, and if it does not, then it will be both more necessary and less just to penalize unwanted conduct.

While most work on the ethics of teaching addresses specific issues, Strike (2003) offers a general account that incorporates ideals of promoting growth, exemplifying civic virtues, and teaching one's subject with integrity or in a way that is true to its inherent virtues. Work on the ethics of higher education has addressed issues of academic freedom, tenure, institutional neutrality, university-business partnerships, sexual harassment, diversity, research ethics, ethical issues in student-life policies, athletics, and the professional responsibilities of faculty and administrators.

THE SUBSTANCE OF SCHOOLING

Discussion of the content of education has often taken the idea of an education in the liberal arts as its point of departure, and multicultural calls to broaden the "canon" or textual basis of liberal education have proliferated. The purpose of a multicultural curriculum is variously described as providing a more accurate view of the world, promoting the self-esteem of those not born into the culturally dominant class or race, correcting the self-perceptions of those who do belong to the dominant class or race, or promoting intercultural or interracial understanding, harmony, mutual respect, or global citizenship. A more radical strand of critique, advanced by Walter Feinberg (1983) and others, holds that the function of schooling is to reproduce social and economic inequality and that school curricula are systems of exclusionary knowledge codes, which mediate that function.

In recent years the major debates about moral education have revolved around three kinds of models and how to move beyond them. Cultural-transmission models call for initiating children into the prevailing moral order by immersing them in a school culture that replicates and teaches it through rituals, moralistic literature, and the like. These models are faulted primarily for their lack of progressivism. Romantic or child liberationist models trust children to spontaneously develop moral sensibilities and commitments but are faulted for their empirical shortcomings. Intellectualist or neo-Kantian models have attempted to sidestep debates over the content of morality and moral education by focusing on the form of morality and moral reasoning. Lawrence Kohlberg's cognitive-developmental variant of this model has been widely influential and widely criticized for ignoring the motivational aspect of moral development and for promoting an ethic of justice that is at odds with the patterns of female moral development, which are said to pertain more to care and inclusion. Alternative models include an ethics of care that emphasizes the nurturing of natural sympathy, neo-Aristotelian approaches that defend roles for both habituation and critical reason, and mixed developmental approaches that consider the moral sentiments, social and community factors, and identity formation together with the cognitive aspects of moral development.

See also Affirmative Action; Authority; Ethics, History of; Feminist Epistemology; Multiculturalism; Philosophy of Education, History of; Rationality; Respect; Socrates.

Bibliography

Brighouse, Harry. *On Education*. London: Routledge, 2005.

Brighouse, Harry. *School Choice and Social Justice*. Oxford: Oxford University Press, 2000.

Callan, Eamonn. *Creating Citizens: Political Education and Liberal Democracy*. Oxford: Oxford University Press, 1997.

Carr, David, and Jan Steutel, eds. *Virtue Theory and Moral Education*. London: Routledge, 1999.

Curren, Randall. *Aristotle on the Necessity of Public Education*. Lanham, MD: Rowman & Littlefield, 2000.

Curren, Randall, ed. *A Companion to the Philosophy of Education*. Oxford: Blackwell, 2003.

Crouch, Colin. *Commercialization or Citizenship: Education Policy and the future of Public Services*. London: Fabian Society, 2003.

Dwyer, James. *Religious Schools v. Children's Rights*. New York: Cornell University Press, 1998.

Feinberg, Walter. *Understanding Education*. Cambridge, U.K.: Cambridge University Press, 1983.

Fenner, David, ed. *Ethics in Education*. New York: Garland, 1999.

Friere, Paulo. *The Pedagogy of the Oppressed*. New York: Seabury, 1970.

French, Peter. *Ethics and College Sports*. Lanham, MD: Rowman & Littlefield, 2004.

Fullinwider, Robert, ed. *Public Education in a Multicultural Society*. Cambridge, U.K.: Cambridge University Press, 1996.

Fullinwider, Robert, and Judith Lichtenberg. *Leveling the Playing Field: Justice, Politics, and College Admissions*. Lanham, MD: Rowman & Littlefield, 2004.

Gutmann, Amy. *Democratic Education*. Princeton, NJ: Princeton University Press, 1987.

Hirst, Paul, and Patricia White, eds. *Philosophy of Education: Major Themes in the Analytic Tradition*. Vol. 3, *Society and Education*. London: Routledge, 1998.

Katz, Michael S., Nel Noddings, and Kenneth Strike, eds. *Justice and Caring: The Search for Common Ground in Education*. New York: Teachers College Press, 1999.

Levinson, Meira. *The Demands of Liberal Education*. Oxford: Oxford University Press, 1999.

Macedo, Stephen. *Diversity and Distrust: Civic Education in a Multicultural Society.* Cambridge, MA: Harvard University Press, 2000.

Macedo, Stephen, and Yael Tamir, eds. *Moral and Political Education*. New York: New York University Press, 2002.

McDonough, Kevin, and Walter Feinberg, eds. *Citizenship and Education in Liberal Democratic Societies*. Oxford: Oxford University Press, 2003.

Noddings, Nel. *The Challenge to Care in Schools*. New York: Teachers College Press, 1992.

Nussbaum, Martha. *Cultivating Humanity: A Classical Defense of Reform in Liberal Education*. Cambridge, MA: Harvard University Press, 1997.

Peters, R. S. *Ethics and Education*. London: George Allen & Unwin Ltd., 1966.

Pritchard, Michael. *Reasonable Children: Moral Education and Moral Learning*. Lawrence: Kansas University Press, 1996.

Sellars, Mortimer. *An Ethical Education*. Oxford: Berg Publishers, 1994.

Simon, Robert. *Neutrality and the Academic Ethic*. Lanham, MD: Rowman & Littlefield, 1994.

Strike, Kenneth. "The Ethics of Teaching." In *A Companion to the Philosophy of Education*, edited by Randall Curren. Oxford: Blackwell, 2003.

Wolfe, Alan, ed. *School Choice: The Moral Debate*. Princeton, NJ: Princeton University Press, 2003.

Randall Curren (2005)

PHILOSOPHY OF EDUCATION, HISTORY OF

There was probably a time when human culture was transmitted spontaneously from one generation to another. The young of the species cannot survive to maturity unless they assimilate some beliefs about the world, some attitudes toward it, and some skill in solving the practical problems it presents; and the only source from which they can derive this minimal wisdom is the culture of their elders. The tendency to imitate offers a ready-made mechanism for inheritance, and in primitive communities, where benign surroundings allowed a leisurely and spontaneous association with children or where a harsh environment spared no time from the effort to keep soul and body together, the education of the young must have proceeded without much thought or care. In societies that were a little more advanced, the need for instruction in tribal ceremonies and the apprenticeship of sons to fathers and of daughters to mothers may have covered spontaneous education with a thin veil of deliberateness. Still, in uncivilized communities generally, culture must have been passed on without the agency of persons especially devoted to that purpose.

Through time, beliefs accumulate, attitudes grow more diversified, skills become more numerous and more complex. This increase in the volume of culture must have rendered obsolete the deliberate spontaneity of its transmission. Mastering what there was to know required special and enduring effort; teaching others to master it demanded more than a casual supervision of their lives. A culture thus enhanced could find lodgment only in a special class of persons—those who were able to encompass it. And this class—seers, priests, and scholars—must have become its chief dispenser to succeeding generations.

BEGINNINGS IN GREECE

There are two important consequences of the concentration of culture in the hands of a specialized class. Conscious of their possession, scholars naturally came to ask

how it might be improved and purified; and this question led to the beginning of research. Second, because they were held responsible for instruction, both scholars and laymen came to expect that some good purpose should be served by their teaching—that it not only should preserve and extend culture but that teaching should serve some other purpose as well.

The earliest records show that the first of these effects, the beginning of research, began to appear in Europe near the beginning of the sixth century BCE. For a long time, no doubt, the learned had looked upon the things of sensory experience as irreducible constituents of the world and, relying upon ancient religious belief, had explained the origin and changes of those things by reference to the gods who presided over them. Now, however, a torrent of speculation deprived sensory things of their irreducible reality and the gods of their explanatory force. Water, pure matter, air, fire—each was advanced as the ultimate stuff of things by some. Other thinkers preferred a substance which possessed all the qualities of sensory things and that was broken into many small bits. Some regarded sensory things as nothing but atoms moving in the void; others resolved their hitherto independent reality into numbers or mathematical structures. And others, still, saw their independence disappear into the absolute unity that was the only reality. Almost all saw the things of ordinary sensory experience as resulting from natural forces working upon the elements or somehow breaking up the unity. The more ancient wisdom was improved by pointing out that the world was really something different from what it seemed to the senses and by disallowing any explanatory value to myth.

SOPHISTS. The second effect of the concentration of culture, the desire to serve a higher purpose, began to appear about the middle of the fifth century BCE. The diversity of opinions concerning the nature of things, their origin and change, and related topics, led in some minds to a profound skepticism. Gorgias (c. 480–380 BCE) argued that nothing exists; that if something did, no one could know it; and that if one could know it, he could communicate his knowledge to no one else. Protagoras (c. 490–c. 421 BCE) held that man is the measure of all things. Each concluded that belief is properly an individual concern and that what is good and right is similarly dependent upon individual interests. They did not draw the conclusion that one might do as he pleased, however; they urged, rather, that conformity to custom and convention furthers the interest of the individual person more than flouting does. They and their fellow Sophists moved through the cities of Hellas, giving instruction in the practical arts, in the humane and literary subjects, in rhetoric, in law and politics, and in the more theoretical considerations out of which their natural and egoistic principles grew. They asked a fee for their instruction, and that procedure was an innovation. But an even greater novelty was their view of their own function as teachers—a view of the transmission of culture not for its own sake merely, or for *ad hoc* purposes, but in order to help their pupils achieve the comprehensive goal of a practically successful life at home, in the court, or in the legislative assembly.

SOCRATES. Socrates (c. 470–399 BCE), to judge from Plato's presentation of him, was even more conscious of his mission as a teacher than were the Sophists. He shared their skepticism toward physical and cosmological theories, but unlike them he refused to leave unchallenged any dogmatic trust in conventional morality. In his hands rhetoric became dialectic; and in his teaching the purpose to which the pupils of the Sophists put the former—the persuasion of others to whatever view the speaker finds most useful—became the discovery of truth, in the dialectical search for which all barriers of personal prejudice and social dogma must give way. He was convinced that the human mind could discover the truth about the physical world and about the life of man in it, although he was equally certain that no one had yet achieved this knowledge. His mission as a teacher, he thought, was to free his pupil's mind from confusion and dogma in order that it should be able to find and recognize the truth—especially the truth about the good or virtue. Confusion and dogma would disappear upon examination of the unclear and unfounded ideas that constituted them. Thus, although Socrates' purpose was positive, his teaching often shows a primarily negative aspect. The skepticism and conventionality of the Sophists brought an objective of prudence to their education; but the skepticism and rationality of Socrates gave to his instruction the purpose of a life of virtue whose discovery required a clarification of the ideas involved in ordinary discourse.

PLATO. Plato (427–347 BCE), influenced by the Sophists as well as by the speculative scientists and metaphysicians and inspired by the instruction of Socrates, gave us the first fully developed philosophy of education—that is, the first explicit, philosophical justification of a theory of education. In his *Republic,* on the basis of observation, he ascribed to all human beings, but in varying degrees, three distinct abilities: the ability to reason, which seeks the good life, the ability for appetition, which is connected with the body and is somewhat wayward, and the

ability to enforce the decisions of reason about what is good against the inclination of appetites. He ascribed to all states, on a similar basis, three functions: that of legislation, that of economic production and distribution, and that of armed enforcement of law and foreign policy.

Plato recommended that education be employed as the chief method of reforming both the individual's character and the state. In a just character each of the three abilities is exercised to the height of its power: Reason recognizes what is good, the appetites freely conform, and the ability to enforce the decisions of reason assures that conformity. In a just state each adult citizen performs that function for which he is best fitted: The highly rational engage in legislation, the predominantly spirited (Plato's name for the ability to enforce reason's decisions about the good) enforce it, and the chiefly appetitive operate the economy. Justice consists in a harmony that results when each part of a thing performs the function proper to it and refrains from interfering with the function of any other part. Reform in individual character and in the state is movement toward personal and social justice.

A system of universal, compulsory, public education from birth to maturity ought to be instituted to bring about this individual and social improvement. All should be taught to read, to write, to count, to appreciate the traditional poetry and drama (highly censored for the young), and to engage in gymnastic exercise. Some should learn the military art, and others should study the sciences and dialectic—the search for the fundamental principle that explains all reality and value. Each student should be tested to discover which ability dominates his soul and should be sent into the state to perform the function appropriate to it when he reaches the limit of his development, which the testing reveals. Thus, each class in the state would be recruited from those best fitted to perform its function. Such a system of education would produce individuals whose souls are as just as their abilities allow and a state whose parts or classes are similarly harmonious.

Plato's philosophical justification of his theory of education consists of three parts. First, he shows that the just state or republic and the just individual are good. For every class of things, there is a Form, or Idea, existing in a supernatural realm, resemblance to which determines the class. The resemblance between a member of the class and its Form is its goodness. The Form for the class of states is that pattern into which the three constituent classes fall when each performs its proper function. The Form for the class of human beings is that pattern into which the parts of the soul fall when each is properly developed. Thus, insofar as a person is just, he is also good, for he resembles the Form of humanity. And insofar as the state is just, it is also good, for it resembles the Form of states. The goodness of a just character and of a just state warrants Plato's recommending them to our efforts.

Besides this ethical support for his recommendations Plato provides a metaphysical explanation for the facts upon which he rests them—the facts of human nature and of society. Every particular falls into some class, and the class is made what it is by virtue of the Form copied by all the members of that class. If we ask, then, why every human being should possess the three abilities (reason, appetite, and spirit) and why every state should perform the three functions (legislation, economic production and distribution, and law enforcement), the answer is that they cannot fail to possess and perform them since exactly that is required by their Forms.

Plato's epistemology gives a third support to his theory of education. First, his contention that we can know only the Forms in their logical connections, coupled with the view that the entire realm of becoming is a copy of that of the Forms, leads to the conclusion that even though knowledge is not an infallible guide to the course of nature it is more useful than mere opinion. In this way he argues that knowledge is useful in the pursuit of justice. He holds, second, that the only method appropriate to acquiring knowledge is that of purely rational inference. Assuming that the method of learning is identical with that of discovering truth, he argues that instruction should follow the path of deduction wherever that is possible.

Plato's philosophy of education resembles in some respects the thought of the metaphysicians and physicists of the fifth and sixth centuries; with them it shares the faith that the human mind can achieve knowledge of what exists. It resembles the thought of the Sophists in its insistence that the world of ordinary sensory experience cannot be known. But of their reliance on conventional morality, it shows no trace at all. Rather, Plato shares with Socrates the conviction that virtue can be known and that it is the business of education to reform conventional morality in its direction.

DEFINITION OF "PHILOSOPHY OF EDUCATION." Plato's work, especially in the *Republic,* serves as a paradigm of a definition of the phrase "philosophy of education." He sets forth an educational theory—that is, a view about the facts of human nature and society on which are based recommendations about the curriculum, the meth-

ods, and the administration of education, regarded as means to the ultimate goal of just and good citizens living in a just and good society. His ethical theory justifies this goal; his metaphysical theory supports the recommendations ancillary to the goal; and his epistemology explains the effectiveness of some of the teaching methods he advocates as well as our capacity to perceive truth generally. "Philosophy of education" means any body of thought like this one—any body of thought that includes a theory of education, an ethics that justifies the goal that the theory adopts, a metaphysics that explains the psychological and sociological parts of the theory of education, and an epistemology that explains why certain methods of teaching and learning are effective and demonstrates our ability to know the truth of any thought whatsoever.

Many philosophies of education do not contain reference to all the subjects with which Plato was concerned. Nonetheless, his reflections on education fix the meaning of the phrase by constituting a model, resemblance to which (at least to some degree and in some respect) allows any body of thought to be called philosophy of education.

HELLENISTIC THOUGHT. After Plato's work, nothing very novel was added to philosophy of education for some seven centuries. There is extant some work of Aristotle's (384–322 BCE), but it is fragmentary and a part of a theory of education rather than a philosophical treatment of such a theory. Epicurus (341–270 BCE) and his followers Zeno of Citium (336?–265? BCE) and the Stoics advocated a tranquility in life—the Epicureans through cultivation of quiet pleasures easily obtained, the Stoics through willing acceptance of the lot for which one is necessarily determined and (among the later members of the school) through a love for all humankind viewed as a brotherhood. But Epicureans and Stoics, as far as we know, themselves developed neither a theory nor a philosophy of education. In the first century CE, Quintilian (c. 35–c. 95) published his *Institutio Oratoria* (*The Training of an Orator*). Quintilian recommends that in his training an orator be given appropriate objectives toward which he can direct his native but unformed impulses. The life of the orator, he dimly suggests, is good because it meets the Stoic requirements of indifference to external circumstance and utility to fellow citizens. His book harks back to the humanistic curriculum of the educator and orator Isocrates (436–338 BCE) and to the Sophists. It was of much influence in later antiquity and again, after its rediscovery, on humanistic education in the Renaissance, but it embodies a theory of education rather than philosophical reflection upon education. Other authors, for example, Plutarch (c. 46–120 CE) and Quintus Septimius Florens Tertullian (c. 160–c. 220), comment on education, but not in a philosophical way.

Although the literature of the Hellenistic age shows little that is new in philosophy of education, two ideas of great importance for change in that philosophy were, nonetheless, gradually coming to dominate men's minds. One is the idea that a chief factor in the good life is obedience to law; the other, that a necessary ingredient in that same life is the happiness of a love that unites all those who obey the law as well as each of them to the lawgiver himself. The Christian ideal of the brotherhood of men under God, their creator, is the expression these ideas assumed, and the movement of Christianity, although influenced by Plato, not to mention Plotinus (205–270), produced a new philosophy of education.

MIDDLE AGES

AUGUSTINE. The new philosophy is the work of St. Augustine (354–430). Human nature, according to his view, must be described in terms of substance and faculties influenced by historical forces. Every human being is a combination of body and soul; the soul possesses the faculties of knowing, feeling, and willing. The first enables us to know whatever we sense and remember to have sensed, certain abstract principles which the mind carries within itself, and the world of sensible things as they are ordered by those principles. The faculty of feeling enables us to desire and to feel emotions which center on desires. The faculty of willing enables us to choose from among differing desires those we want to realize—an ability which exercises itself freely and which, when exercised correctly, employs rules of choice that flow from divine commands.

Human nature cannot be accounted for in terms of substance and faculties alone, however. A historical force always determines how these faculties operate. Before the Fall, Adam and Eve used their faculties in the right way—especially their faculty of desire, directing its operation upon what they ought to desire, centering their love on God and on one another in communion with him, and choosing freely to obey his commands whenever the clamor of bodily appetite opposed itself to the right. But from their original sin, of which the Fall was a natural consequence, flows the force which determines their descendants to act as sinfully as they—to choose freely to disobey God's command by selecting egoistic and carnal desires for realization. Human nature must be painted in terms of substance and faculties corrupted by early events in human history.

Human society is constituted by the direction of the activities of its members toward a single goal, but, like the human soul, it cannot be understood merely in terms of this abstract function. The unity of purpose that in principle constitutes the family, the city, the empire, and the community of humans and angels is disrupted by inherited self-seeking. Another historical force determines two other communities—the city of earth and the city of God—each of which is reflected in the four just mentioned. The advent of Christ signifies God's wish to enable men, despite their sinfulness, to merit salvation. The city of earth is made up of those who refuse to believe in Christ's mission and to repent; its members will not be saved. The city of God is composed of those who believe in that mission and feel genuine repentance; its members will enter upon eternal communion with God after the day of judgment.

The ultimate objective of education grows out of the corruption of human nature and God's concern over it. Like the ultimate objective of the church, that of education is conversion and repentance. On the elementary level the curriculum should be the seven liberal arts—a program of studies prefigured by Plato's curriculum; on the advanced level it should consist in philosophy and theology. The method appropriate to the lower level involves censorship and the prevention of idleness in order to stifle sinful desires. The liberal arts should be taught in an authoritative manner because not all who seek elementary instruction are sufficiently rational to know the truth and since no more than belief is required for salvation. On the higher level, authority gives way to proof since those who advance thus far are able to achieve knowledge. The liberal arts, coupled with religious worship and instruction, ensure correct belief about the nature and order of the universe and about God's relation to man; philosophy and theology show the more able—those destined for the hierarchy of the church—why those beliefs are true.

Augustine's philosophical reflections upon his theory of education stem from his conception of God. He advances, first, a theory of language according to which every word means what it names, and every sentence, the combination of things named by its component words. He concludes that since on this theory no one can tell someone else what he does not already know, each man must learn for himself by consulting things as they are illuminated in a light of divine origin. Teaching is not informing; it is reminding others or ourselves of the knowledge supplied by God.

Second, from the concept of God flows the justification of the objective of education. The goodness of each created thing consists in its resemblance to the idea held before God's mind as the pattern for its creation; this idea is its exemplar. The exemplar for men is the obedience to God's commands and love for him and for one another in him that gave perfection to life before the Fall. To be happy is to possess what one wants at the time of wanting it; since God is the only eternal thing, he is the only dependable object of desire. To be happy is to illustrate the exemplar for man, and conversion, the objective for education, consists in achieving that condition.

Augustine finds in God, also, a metaphysical explanation of human nature and society. In the first moment God created everything either in actuality or in potentiality. All history—each person's repentance or failure to repent, each society's deeds, both good and bad—is the unfolding of what was first merely potential; what happens is what must happen because of the initial creation and God's all-comprehending providence. Human nature and society must be corrupt; hence, conversion must be the ultimate purpose of education.

Later medieval thought. During the centuries that followed the death of Augustine the interest in another world became so dominant that education diminished in importance, and reflection upon it very nearly ceased. Attention was centered on the otherworldly results of repentance or its failure at the expense of training for terrestrial existence; and so dogmatic was the assurance of the need for conversion that any effort to justify this objective appeared useless if not impious. The clergy, then Europe's teacher, offered a meager training to those working toward holy orders and some understanding of religious ritual to the laity. But the transmission of culture diminished greatly. The widespread acceptance of the otherworldly objective of education stifled philosophical reflection upon it. Comment on education is found in the writings of the Venerable Bede (673?–735), of Alcuin (735–804), and of Hrabanus Maurus in the early ninth century; but they are at most casual and at least unphilosophical. Thomas Aquinas (1224?–1274) devoted some systematic attention to the philosophy of education, but his chief contribution to it concerns not the objective of training but the nature of teaching—a discussion which continues the thought of Augustine on that subject.

RENAISSANCE. With the Renaissance came a revival of interest in ancient learning and a recognition of value in terrestrial life. In accord with this change of outlook some

writers assigned to education an egoistic and prudential purpose like that of the Sophists. Reformationist thought—at least in Martin Luther's case—demanded universal, compulsory, state-controlled education in order that religion should be national and God's word available directly to all. Ignatius Loyola (1491–1556), through the Society of Jesus, established a widespread system of schools and universities; and in 1599 the society established a plan of education for them (*ratio studiorum*) that exercised much influence on Catholic education. But Reformationist and Counter-Reformationist literature reveals much more polemic and dogma than philosophical reflection upon education.

MODERN PERIOD

COMENIUS. In the seventeenth century, philosophical reflection upon education began anew, and its history from that time to the present is that of the gradual secularization and naturalization of the Christian objective assigned to education by Augustine. The work of John Amos Comenius (1592–1670) begins this process. (In particular, see his *The Great Didactic* and *The Way of Light.*)

Like Augustine, Comenius holds that human nature is corrupted by inherited sin, but he also asserts that it is capable of absolute perfection. The soul contains the possibilities of erudition (perfect knowledge), of virtue (adherence to the rules of right conduct), and of piety (love of God, the author of humankind). Like Augustine, Comenius viewed history as a decline from innocence, but he held, nonetheless, that there is a zigzag pattern in history, leading to an age of perfect terrestrial existence before the last judgment devoid of international strife and ruled over by Christ. In this last age the possibilities in the human soul realize themselves in perfect knowledge, virtue, and piety, and all societies unite in a single international brotherhood. The reward for striving after this perfection is immortal blessedness. Comenius held that a system of public, universal, state-supported schools, from childhood to maturity, should further the full actualization of the soul's possibilities and assist history toward its goal. The curriculum should constitute a cyclical development from the simple and abstract elements of science, art, language, literature, and religion to their complex and concrete forms. The methods of instruction should consist in the uniform application to the young of the human species of principles observed in the development of the young of other species, both plant and animal.

Comenius's philosophical reflection on his theory of education centers, like Augustine's, around the notion of God. God made humankind in his own image, and, because God is perfect, humans may become so as well. To achieve perfect knowledge is to make perfectly clear to ourselves the things our sensations reveal and to order them according to innate principles which reason brings to light. To perfect conduct is to identify the rule of one's will with a command of God, and to perfect piety is to love God in one's obedience to him. Human nature and human history find a metaphysical explanation in divine providence, which manifests itself through the opposed forces of light and darkness. The business of education is to perfect individuals in the three ways mentioned. It also makes the personal life of each human being perfectly Christian and aids history in its progress toward final social perfection.

LOCKE. Late in the seventeenth century, not long after Comenius, John Locke (1632–1704) published *Some Thoughts concerning Education.* In this book, in *An Essay concerning Human Understanding,* and in *Second Treatise of Civil Government,* he carried further the secularization of the objective of education started by Comenius. With Augustine and Comenius, Locke held that man is free, but in opposition to them he denied that man is inherently sinful by virtue of his racial history. Each person is a mental substance joined to a bodily substance, as Augustine asserted; mental activity, however, can be described wholly without reference to substance, in terms of two faculties, understanding and will. The faculty of understanding enables man both to know and to desire, but what man knows is determined by the ideas his environment allows to enter his mind, and what he desires is determined by the objectives his environment supplies to a few native instincts. The second faculty is the will, and its exercise consists in choosing desires for realization where they conflict.

Society in the state of nature is based on a natural division of labor and on the need to care for offspring. In that state the original "common" of the world was largely transformed into private property, and the function of primitive society was to enforce natural law, or the law of God according to which private property ought to be respected. Disputes inevitably arose, and, since everyone possessed the power to enforce the law of nature, they often could not be settled amicably. Political society came into existence as a guarantee against such disputes. It is based upon a contract or agreement between the community and others according to which each member of the community agrees not to exercise his power to

enforce natural law provided that the others who constitute the government will exercise it for him. It follows that the exercise of governmental power is legitimate only where it protects private-property rights. A government of the kind instituted after the Glorious Revolution, having popular representation, Parliamentary determination of the sovereign, majority rule, and separation of legislative from executive power, Locke held, is best suited for achieving this objective because it can most efficiently check unnecessary governmental activity.

The purpose of education is to produce people who will advance the happiness of the community. They must be of good character and properly disposed toward learning. Good character consists in the habits of acting virtuously, prudently, and with good breeding. The proper disposition toward learning is not possession of it but an esteem for it and the habit of acquiring it when the need arises. These habits and dispositions can best be acquired by a tutorial education at home, by a method of pitting one instinctual desire against another in order to establish them, and by presentation of clear and distinct ideas to the pupil in the order and connection possessed by their objects. In both moral and intellectual training one should appeal to the interests of the child, bring him to learn for himself, and give public approbation to his success. The child who will benefit from such instruction and who will contribute to the happiness of the community is the son of landed gentry, who can look forward to a place in government. The poor should be given sufficient education to make them religious and self-supporting.

The production and maintenance of a good society is the chief objective of Locke's theory of education. Such a society is one in which men find pleasure or happiness in the performance of duty, and Locke's ethical reflection endeavors to justify this conception of the good life. Duty is obedience to natural law as embodied in civil law concerning the protection of private property. Like all moral principles, it can be known with certainty to be valid; it can be demonstrated from the ideas of God, of his creature man, and of the relation between them. The moral and intellectual training of the gentleman will cause him to find his pleasure in doing his duty; the exercise of this duty through government as well as through more informal social controls will spread a similar happiness throughout all levels of society.

Locke's theory of knowledge led him to conclude that we can be perfectly certain of any proposition whose truth we can intuit, demonstrate, or perceive through our senses or through our memory of such perception. Since the validity of duty can be demonstrated, we can know that it is right to perform it; and in this way, his emphasis on moral education is justified. Since the theory holds that we can know very little of the sensible world—only what we remember having perceived through our senses or are now perceiving through them—the de-emphasis of intellectual pursuits is also justified. We must accept many propositions about nature on faith or as merely probable; hence, we do not need to busy the heads of the young with any detailed consideration of them.

ROUSSEAU. Jean-Jacques Rousseau (1712–1778) advanced three distinct philosophies of education; in the most influential of the three he varied the social theme found in Locke's thought. In his discussion of a new constitution for Poland he advocated a highly nationalistic program on the ground that where a nation's institutions are in good health, education should support and renew them. In *Émile,* he set forth a program appropriate to women, holding that their education should give them charm, ability for household management, and thorough-going dependence on their husbands in matters not pertaining to the home. But the major part of *Émile* deals with the education of gentlemen, embodies a theory of education that has exerted much influence upon educational practice, and assigns to education a social ideal quite as secular and political as Locke's but applied in an altogether different way.

Rousseau described human nature, as did Locke, as independent of historical influences and as initially perfectly innocent. A human being is a substance with faculties—those of pleasure and pain, of sense, of reason, of desire and emotion, and of will. These faculties emerge clearly at different stages in the life of the individual according to a general pattern, and the personality is more or less stable according as the newly emerged faculty is made to harmonize with the exercise of others already established. Despite the general pattern, each individual differs from others and must achieve stability through a procedure adapted to his own case.

In the early history of humankind there was no society. Men were independent and therefore equal. With improvement in techniques of hunting, fishing, and farming, they acquired property; with property, they acquired families, differentiation of economic function, interdependence, and inequality. As society became more complex, greed, ambition, and deliberate selfishness entered the soul; in time, men developed government and law in order to protect the property of the wealthy against one another's greed and against the greed of the poor.

Inequality is fixed in the structure of eighteenth-century society and is due for removal by revolutionary action.

Rousseau presented detailed recommendations for educating gentlemen to live happily in these circumstances. They differ for each stage of development, but he urged that in all the child must learn for himself through personal observation of and active participation in the world of nature and society. A tutor who devotes his entire career to one pupil should attend to the pupil's individual interests and instruct him by rousing those interests into activity. The young man who completed this education would have enjoyed to the full each of the stages in his development and would be possessed of a strong body and stable mind. This stability would consist in his possessing no desire for whose realization he did not also possess the requisite power. It would make him neither learned nor urbane, but it would lead him to adopt a rural life in which he could survive the social storm Rousseau anticipated.

Rousseau advanced three criteria for knowledge: sensory experience of the consequences of action, the dictates of the heart, and practical utility. The first he transformed into a method of instruction—the method of letting the child experience for himself the consequences of acting upon his ideas in order to learn what is true about nature and society. The second he employed to warrant his inclusion in education of a considerable amount of simple religious doctrine. The third he relied upon to exclude from education a great deal of philosophy and other literature that he found devoid of practical consequence.

Rousseau's metaphysical reflection led him to hold that all of nature, including men's bodies and their actions, is governed by law but that since duty often requires one to act in ways other than those determined by this law, there is a supernatural realm in which duty presides. To act according to duty is to use the right rule for selecting one desire from among many as a basis for action, and since this selection and realization runs counter to nature, we must be exercising free will when we act rightly.

Rousseau's thought about morality concluded with the view that the good life is one in which there is neither the shallowness of desires that have been multiplied to match excess in power nor the discontent of an excess of power over desires but the happiness which occurs when power to fulfill desires equals the desires one harbors and is exercised to realize only those which are in accord with duty—a view not unlike Locke's. Duty Rousseau understood in terms of the general will. This is the welfare of the nation as opposed to the corporate will, or the welfare of a smaller group, and to the particular will, or the welfare of the individual.

It is our duty to act for the general will where that is possible. But in the major nations of Europe all institutions have been subverted to the service of corporate and particular wills. The social contract (which is, whatever the historical account of it may be, the agreement to act in accord with duty rather than for some lesser goal) has been betrayed by those in authority. Consequently, the ideal of duty cannot serve as the purpose of education generally. The realization or preservation of one's own will must be put in its place. In this way Rousseau justified the individualistic effort at internal peace that informs the theory of education with which he was most concerned.

PESTALOZZI. The educational proposals of Johann Heinrich Pestalozzi (1746–1827), unlike those of Rousseau, whom he greatly admired, bear no trace of direct revolutionary inclinations. But he had a warm sympathy for the downtrodden, and he advocated education for all as a condition of social reform. By his example and his books he contributed greatly to the common-school movement in Europe and America. The influence of Rousseau on his thought is evident chiefly in Pestalozzi's insistence on treating children in ways appropriate to the process of development through which they all must pass.

This process exhibits three stages. The contents of the child's mind are at first blurred and indistinct. Next, objects stand out in consciousness characterized by explicit forms and qualities. Last, these objects are understood as examples of general concepts; they are, to use Pestalozzi's word, defined. Throughout the process the person is himself active in securing and clarifying images and in transforming them into ideas that contain knowledge. Each child should be dealt with in accord with the place he occupies in this threefold process, and a major part of teaching consists in enabling him to work out for himself his own knowledge or definition of things.

Knowledge always contains three elements: the number of things known, the form they exhibit, and the language that embodies them. Pestalozzi concluded that learning must start with the elements into which each of these may be analyzed. The elements of number are units, and arithmetic (operations with units) must be mastered in order to understand number. Form Pestalozzi seems to have thought of as visual and tactual; its elements, consequently, are lines, angles, curves, etc. The student must understand these elements before he can understand

form. The elements of language are ultimately letters, and the mastery of language depends on mastering their spoken and written forms.

Pestalozzi set forth detailed methods for teaching the elements of number, form, and language. They grew out of those he thought natural to a mother's dealings with her children. In the family situation a mother can know in what stage of development each of her children finds himself; she can teach him to count, to draw, and so on, through use and observation of ordinary materials in the context of the economic employment, such as spinning and weaving, in which the family engages; and she can assure herself that he comes to perceive objects clearly and to define them for himself according as his stage of growth permits. These methods, directed toward enabling each child to acquire knowledge based on his own perception (*Anschauung*) of things, Pestalozzi thought could be employed in a school situation. The schools he operated in Switzerland, taking the Swiss village family as their model, attracted imitators from many parts of Europe and America.

Besides knowledge of things, teaching should bring children to a knowledge of skills which exhibit their physical or motor capacities as knowing does their intellectual abilities; and Pestalozzi thought that the performance of deeds could be analyzed into elements just as knowledge could. He was convinced that learning how to do things required the mastery of elementary motions, just as coming to know required the mastery of the elements of number, form, and language. The teaching of morality and religion—more important than that of knowledge and skill—involved transferring the child's feelings of dependence on the mother to other persons in society and to God. But Pestalozzi's treatment of the development of the motor and moral capacities is not so detailed and clear as his discussion of the education of the intellect although he insisted upon the inseparable unity of the three capacities.

The direct influence of Pestalozzi on philosophy of education is negligible. He was not interested in it. Still, his schools and his writings on the theory of education strongly influenced some who were.

FROEBEL. Pestalozzi's younger contemporary Friedrich Froebel (1782–1852) spent several years working in one of Pestalozzi's schools. Froebel was also much given to philosophical reflection, upon which, he thought, the theory and practice of schools depended—especially that of the kindergarten, which he invented almost single-handed.

Froebel's speculations found the goal of education in the full and integrated development of all the powers of the individual and in the internal harmony, as well as the harmonious relations with society, nature, and God, that this development assures. This goal cannot be imposed upon the student; he must achieve it for himself through activities expressive of the powers he harbors. One who has accomplished the goal exhibits a steadiness and solidity of character that gives him integrity in all situations and the intellectual habits (not a store of remembered facts) that enable him to acquire knowledge when necessary.

The process by which this goal may be reached, the process of education, consists in the unfolding of what is present in infancy. Each person is like a plant, and as a plant develops toward a given stage of maturity, so the life of each human being consists in the filling out, through increase of varied detail, of a pattern present from the start. This process is also one of increasing clarity of self-expression and culminates in a clear consciousness of the self. The development of the individual is altogether continuous, and the stages of infancy, childhood, boyhood, youth, and maturity into which it is divided are characterized not by the emergence of novelties, as Rousseau had suggested, but by an increase of clarity in consciousness of the tendencies present in all.

Froebel worked out methods of education in accord with this view of individual development. The methods applicable in the earlier stages should merely enable spontaneous expression of the pupil's self; methods applicable to the later stages should supervise and direct that development. His treatment of the stage of childhood amounts to the nearly single-handed invention of the kindergarten—an institution that spread quickly, especially throughout the United States. His treatment of boyhood involved considerable innovation in the methods, materials, and curriculum of elementary schools.

In the first stage, the infant should be nurtured and cared for. In the second, the senses and language develop, and the child's tendencies toward this development should be permitted free expression. Play is the most important method for this expression. Froebel invented various apparatus (called "gifts") to serve as educative toys; introduced activities (called "occupations"), such as drawing and clay modeling, which, along with the gifts, develop sense perception; emphasized song and spontaneous conversation to develop language and prescribed games, often played in a circle (to which figure he attached cosmic if obscure significance), to develop the sociality inherent in the child.

The stage of boyhood should be developed by instruction. The boy is becoming self-conscious; in order to develop steadiness of character, he should participate in the administration of the school through school government. The study of nature, stories, learning in groups, family work, making things—all these further steadiness of character and habits of intellectual readiness. Froebel insisted that instruction, the direction of development, should not aim at the practically useful but at that self-consciousness of integrated and developed powers which is the proper objective of individual and social evolution. About the stages of youth and maturity Froebel had little comment.

Froebel saw education—the early, spontaneous, and the later, but directed, unfolding of the essential powers of each individual—in a metaphysical setting, tinged with mysticism, obscurantism, and incoherence and indebted heavily to the absolute idealism of his day. The Absolute embraces everything and is continually evolving as force in nature and as mind in man. This cosmic evolution proceeds from action to reaction to equilibrium, from simple to complex, from unconsciousness to self-consciousness. Froebel identifies the Absolute with God and its evolution with his creation. Everything has a purpose that unifies it and that binds it into larger organic wholes, by virtue of evolution or creation. The evolution of the Absolute is reflected in miniature in that of humanity. The human race has developed through five stages, and the life of each individual reflects this racial and cosmic evolution. Education, Froebel thought, ought to enable this process to fulfill itself in each person without hindrance. It ought to be the minister to individuals of a cosmic and racial evolution.

The best life for man is the fullest realization of a consistent will—the consciousness of the best self that he can develop. This self-consciousness is awareness of purposes inherent in him; in becoming aware of them, man becomes free. Evil is the distortion by some external factor of a tendency native to the self; all tendencies are naturally good if allowed to develop into self-conscious, harmonious freedom. Although some education should direct, the fundamental early education is chiefly negative; that is, preventive of external obstruction to the development of natural tendencies.

Froebel's metaphysical and ethical doctrines inspired him to activity that had enormous practical effect upon the schools directly, and while the chief influence on his thought lies in the practical work of Pestalozzi, Immanuel Kant (1724–1804) influenced it indirectly, at least through the pervasive effect of his theories on German thought in general.

KANT. The impact on educational theory of the work of Pestalozzi and Froebel was an emphasis on developing individuality in the student, and this impact may be traced to the thought of Rousseau. In the work of Kant a greater optimism than Rousseau's gave a less individualistic objective to education.

Kant conceived of human nature in terms of three faculties: cognition, which organizes sensory elements into the orderly world of experience; desire, which exercises itself in an instinctive effort at lawless, egoistic domination over others; and will, which selects desires for realization according to a rule. Human society grows out of the exercise of these faculties. The instinctual desire for domination leads to conflict between individuals; the faculty of cognition yields knowledge about how this conflict can be avoided—by association in republics; and the will leads to actual societies of this kind for mutual protection. But between republics conflict breaks out anew; and in order to avoid it, these states tend to unite in a peaceful international community. This community is the natural result of the unimpeded development of human faculties; and since we must believe that all things develop their capacities fully, we must believe that it stands at the end of historical progress.

It is the ultimate objective of education not to advance the welfare of individual students, but to promote the realization of the peaceful international state as the embodiment of human perfection. Accordingly, teachers should not regard the economic or other success of their charges but should center attention upon the fullest possible development of their faculties. This development can be assured by supplying to the cognitive faculty the general truths it should use to organize sensory elements into nature as we experience it, by rigorously disciplining the faculty of desire in order to eliminate the instinct for lawless behavior, and by enabling the will freely to use the right rules in organizing the remaining desires. The result of such instruction will be a perfected character and intellect, which, through the progress of generations, will assist history to realize the educational ideal.

Kant's ethical theory supplies a criterion for the kind of conduct which makes the international state possible. It is conduct which embodies rules that can be generalized without absurdity—rules which fit into the famous "categorical imperative." "Break your promise when you wish to" is not such a rule; for if instead of applying it to

your own desires alone, you try to imagine all persons using it in selecting some for realization, the notion of a promise completely disappears. The rule degenerates into the nonsensical "Break a promise which no one ever takes to be a promise when you wish to." "Always keep your promise" is a necessary moral rule, and like all rules which fit the categorical imperative it is so because we cannot imagine the generalization of its opposite without imagining something rationally absurd. In the international state the character of each person will be so perfected that each will act upon such a rule when it is necessary to make a moral choice. Thus, the state will be both realized and preserved. Kant's philosophy of morals, in this way, clarifies part of the notion of an ideal social order which education should subserve.

Kant's metaphysics makes a great deal of the distinction between two realms—the realm of things we can experience, or phenomena, and the realm of things which transcend experience, or noumena. Following Rousseau, Kant held that human beings dwell in both realms and that in the former their desires and actions are determined by natural laws, whereas in the latter they are governed by right rules or duties. To act rightly requires that a person freely employ a right rule and that he not act in a way determined by a law of nature. Hence, whenever one acts rightly he acts as a free citizen of the noumenal world—he freely applies a rule to his desires to decide which one to act upon. This proposition of Kant's ethical theory illumines his method of training the will; that is, his method of preventing the growth of habit and of requiring that children freely adopt a rule in some hypothetical situation of choice.

Kant's views about history provide a goal for his theory of education, and his ethical and metaphysical theories explain part of that ideal and the method proposed for arriving at it.

FICHTE. Rousseau's despair of achieving the national welfare led him to advocate the cultivation of individual self-sufficiency; and while it was no part of their theories, the effect of the work of Pestalozzi and of Froebel was to further attention to the individual student in the practice of education. Kant's enthusiasm for international well-being led him to advocate a future achievement for the entire race through the fostering of universal faculties rather than through the development of individuality. Enthusiasm for national existence as opposed both to individuality and to internationalism brought Johann Gottlieb Fichte (1762–1814) to advocate an objective more like the one Rousseau would have recommended if he had been more hopeful about the national institutions of his day.

Addressing the German people during the subjection of Prussia to Napoleon Bonaparte, Fichte urged that education be used to unite all Germans in a state that, through purity of race and character, would lead the world. Education was the only independent action allowed by the French; if all German children were separated from their parents, reared in a partially self-governing community in which each individual might learn directly the responsibilities of citizenship, taught through the energizing force of interest rather than by reward and punishment, and thus prepared for an adult life of wholehearted and unswerving duty, this possibility of independent action could be turned to the advantage of all Germany. It would lead to the creation of a reformed and unified German state, devoted to the right, and worthy (unlike others) of world dominance. This nationalistic objective of his somewhat fanciful proposals Fichte might have supported by his view that the best state is highly authoritarian—one in which the fulfillment of each man's duty to work is made possible by the state's provision of the opportunity and compensation for work and the complete control of the economy required by that guarantee. This socialistic ideal, in turn, he might have supported by his view that the physical world must be understood as the means and medium by use of which and in which duty becomes embodied in fact. This view is consonant with his metaphysical idealism, according to which the ego posits itself and its objects for the purpose of doing what it ought—a position Fichte developed out of his criticism of Kant's doctrine concerning noumena.

HERBART. Like Fichte, Johann Friedrich Herbart (1776–1841) gave much thought to the doctrine of noumena; but unlike him, he arrived at a kind of realism, to be described later, opposed to the metaphysical idealism Fichte, G. W. F. Hegel, and others made current in the Germany of his day. He relied upon it to advance an objective of education which assigns importance both to individuality and to sociality—both to being a person of the best possible sort and to being a citizen of the best possible society.

There are five criteria, the "moral ideas," all of which must be exhibited by a person with the best possible character and a society of the best possible sort. Applied to a person, the first two of these ideas are relations between his will and other aspects of his character, while the last three are relations between his will and other persons.

When one knows what he wills and approves of it, he is "inwardly free," and inward freedom is the only freedom men enjoy. When one's will is strong, directed toward many things, or "many-sided," and constituted by inclinations toward objectives systematically ordered by the teleological relations they bear to one another, he possesses "perfection." When one directs his will toward enabling the wills of others to be realized, for the sake of that realization rather than for his own benefit, he is "benevolent." The remaining two ideas apply not to wills alone but to the embodiment in action of one person's will with respect to others. When several persons deliberately live according to a principle or law, thus preventing conflict, each individual acts "rightly"; and when a person willfully benefits or harms another, the idea of "equity" or "requital" requires that a corresponding benefit or injury be visited upon the doer of the deed.

A society—a political state or group of any kind—to which the five moral ideas apply is one in which law prevails because of the general relinquishment of rights whose exercise leads to conflict; one in which there is a system of rewards which makes requital to each citizen for that relinquishment; one in which an administrative system exhibits benevolence by assuring to all the greatest satisfaction of will; one in which many interests or wills, both individual and collective, find coherent realization or perfection in a cultural system; and one in which the society, being "inwardly free," knows its own will and approves of it—a trait that requires a soul for society not unlike that of the individual person.

Assuming that if the individuals in a group acquire the moral ideas the group will also, Herbart holds that the immediate objective of education is to produce individuals who exhibit them; and the production of such persons consists in the appropriate use of truths of psychology. These truths describe the relations of ideas or representations, and Herbart is distinguished in the history of psychology as having been among the first to have endeavored to state those relations in a rigorous, mathematical way. He regarded the propositions of his psychology as based on introspection and as justified by metaphysical reflection. Released from its technical form, his psychology may be stated, in part, as follows.

Each idea, Herbart held, endeavors to preserve itself and succeeds in that endeavor to some degree, that is, *is* itself, more or less. The degree of its success depends upon its relations to other ideas, and these are of three kinds: of opposition, of mere dissimilarity, and of similarity. Red and blue (not-red) are opposed to each other, and short of some third idea that combines them, such as the idea of a substance red on one surface and blue on another, they cannot both be present in the same consciousness. Red and circular are merely dissimilar; consequently, they may both present themselves either in combination in a red circle or in simple juxtaposition or may be present separately. A red rose and a red apple are similar ideas; consequently, one may come to be attached to the other. The effort of each idea to preserve itself—an effort which cannot be completely canceled—succeeds insofar as we are conscious of the idea. The greater the success, the greater is the clarity of our awareness of it; the less the success, the dimmer our consciousness of it. But the degree of the success of any idea depends upon the aid and attack it sustains from others; so that the clarity or obscurity of any idea—its place with respect to the threshold that separates conscious from unconscious ideas (a piece of psychological apparatus made current by Herbart)—depends upon the context of other ideas in which it occurs. Where they oppose it and are stronger than it is, it disappears into unconsciousness and becomes an unconscious impulse, striving to emerge into consciousness the moment it is not prevented by the occurrence there of its stronger opposites. Where the context includes merely dissimilar ideas, it may remain in consciousness, but not for long. The flux of experience will soon bring ideas into consciousness that will drive it down into the dark through opposition or keep it in the light by uniting with it through similarity. Where other ideas are similar, they come to its aid, forming a strong union that, so long as it remains, draws to itself its similars, inward from new sensory perceptions and upward from the storehouse of unconscious old sensations. Such a union of ideas is an "apperceptive mass" or "circle of thought"—another piece of psychological apparatus Herbart helped to make current. The psychology upon which Herbart based his educational procedures informs us that new pieces of information can be mastered only insofar as they become united with some apperceptive mass of ideas and that insofar as they are not so united, they are transformed into unconscious strivings, able to present themselves to consciousness only when a lack of their opposites there allows it or the presence of their similars there draws them up into it.

A person consists of ideas that dwell on two levels. On the level of consciousness he is a succession of ideas, each of which originates either in physiological activity or in sensation and quickly unites with some apperceptive mass or is pressed down into unconsciousness by the success of others striving to occupy consciousness. On the level of unconsciousness are all the ideas whose weakness or whose lack of similarity to those in consciousness

chains them in that dark domain. On the level of consciousness the succession of ideas is punctuated by acts of attention. These are simply ideas in which we are, more or less, completely "absorbed." Some, like loud sounds, are involuntary; others, like highly discriminated shades of color or purposefully held thoughts, are voluntary. As objects of attention these ideas are isolated, but they either quickly become unconscious or acquire "meaning" and connection by drawing up into consciousness those "circles of thought," or apperceptive masses of similar ideas, in whose context they acquire significance. An idea attended to much or clearly, together with its circle of thought, is an interest—a desire to bring into existence that which it represents in some future time. The apperceptive mass to which the idea belongs, together with the relations of that mass to others, presents a framework for its suppression, its mere entertainment, or its realization and makes it a desire rather than a free-floating fancy—a part of the person rather than a casual caprice. An act of will is a desire together with the intention that what it refers to should occur. The ego is the central point of the person—the present idea from which memories radiate into the past, interests (desires, acts of will, etc.) into the future, and to which entire apperceptive masses are drawn from the domain of the unconscious or forced down into it.

Ideas, thus arranged and centered, exhaust the person as an introspectible entity. They result from the exercise of no faculties (Herbart seems both to have used this concept and to have declared it nonsignificant), for the soul possesses none. To think of something, to desire it, to will it, to have a feeling toward it—all this is nothing but, in different ways, to be conscious of an idea as connected with others.

Herbart's view of the nature of a person provided him with a method of education which became widespread both in theory and in practice. Education, he held, is instruction, and instruction should consist in four steps. (His followers made them five, prefixing "preparation" for it to "presentation" of an idea.) First, the idea or information to be learned must be "presented" to the student's clear attention; second, the idea thus presented must be allowed to draw up from the student's unconsciousness all ideas whatsoever whose similarity attracts them to it; third, through comparison most of these associations should be eliminated in favor of those which give the idea its proper meaningfulness in a circle of thought; last, to strengthen the idea's bonds in that circle, the student should be brought to "apply" the idea to new situations. This procedure, based upon the flux of ideas from the center of attention into the apperceptive mass to which they belong, gives the student mastery over new information; and mastery, or the ability to reproduce ideas, is the purpose of instruction.

To instruct a person is to construct him; since feelings, desires, etc., are all ideas, providing the student with ideas is providing him with all the materials of personality. But the instructor, by arranging the conditions in which the student acquires new ideas, determines not merely the materials out of which he is formed but also the organization or form those materials assume. And a person, as we have seen, is simply ideas organized in a certain way.

But education is not merely the construction of a person; it is also the effort to construct one who exhibits the five moral ideas. Herbart refers to this aim as the production of "character," and he deals chiefly with the production of "perfection," or "many-sidedness." If the child's attention is called to many things in his own experience, and if the store of this experience is supplemented vicariously through communication with other persons—a communication based on sympathy with them—his interests will naturally become numerous, and by control of the natural mechanism of apperception, well organized and strong. Perfection of will or character, tinged with an inevitable individuality, is a necessary ingredient in the objective of education, but it is also essential to sensible choices in adult life.

Herbart advanced a metaphysical view as a ground for his psychology. Reality is neither mental, as the prevalent idealism held, nor physical in the sense of being extended in space and time. Its characteristics are quite unknowable except for those of being independent of our minds and composed of perfectly simple entities (*Realen*), not unlike the monads of Gottfried Wilhelm Leibniz. These simple reals conflict with one another from time to time, and on such occasions, there occurs an effort on the part of each to preserve itself from destruction. In a body, this act of self-preservation is its state; in a soul, such an act of self-preservation is an idea that represents, so far as that is possible, the attacking entity. Being simple, the soul cannot engage in more than one act at once; hence the struggle of ideas against one another and the inevitable fall into unconsciousness or into the unity of some apperceptive mass.

The ethical theory by which Herbart justified the five moral ideas as the standard for personal and social existence is one which holds that moral judgments are a species of aesthetic judgments. As such, they neither need nor can be given justification. The human taste prefers

persons and societies that live up well to the five ideas, but the validity of the standard by which they are measured is still nothing different in kind from the taste we enjoy for music, painting, and the natural landscape.

J. S. MILL. In determining the objective for education, John Stuart Mill (1806–1873) disregarded several distinctions emphasized by his predecessors. He ignored the distinction between national and international well-being, speaking of society without qualification, and he argued that individual and social interests might be identified. But his work resembles that of Herbart in some ways: He endeavored to make use of psychology to achieve his educational objective, and the psychology he employed, although it regarded the elements of the mind in a different way, attributed relations to them—those of association—not altogether unlike those Herbart thought he had found.

Mill conceived of a human being in terms of a body and mind, but although they occur in his thought he scarcely makes use of the ideas of substance and of faculty in understanding human nature. The body, with the help of external things, determines what our sensations are like, and it harbors physiological structures which cause us to find activities and things of certain kinds instinctually pleasant or painful. The mind is a series of sensations and ideas with attendant feelings and emotions, held together by connections of an associative kind. Conscious elements are connected in these ways when they have been associated in past experience in certain circumstances. Under these conditions, when one element recurs in consciousness it brings its associates with it. The conditions of association are never repeated from one person to another; hence, every human being is unique.

In his *Utilitarianism* Mill holds that the best society is one in which there is the greatest amount of happiness for the greatest number of people. He understands happiness as constituted by pleasure properly proportioned between higher and lower activities, individual self-realization, and fulfillment of duty.

The chief purpose of education is to bring individuals closer to this social ideal. Careful attention to the content of the curriculum can develop the proper proportion between higher and lower desires and consequently between higher and lower pleasures. The method of instruction can ensure individual self-realization by making room for free discussion and personal discovery of truth. The most difficult task is so to associate egoistic pleasures with fulfillment of duties as to connect them in all subsequent experience. The success of this effort will be a person who finds pleasure in doing what he ought even though doing so involves personal sacrifice. Compulsory elementary education for all and higher education for those who can benefit from it will go a long way toward a society in which happiness is at its maximum.

Mill supported his theory of education by providing a justification of the utilitarian ideal by a theory of meaning according to which free discussion of the consequences of our ideas is the best way to make their meaning clear and by a theory of knowledge according to which we can, by using his famous canons of empirical inquiry, come to be perfectly certain about the sequences of things in nature whose use enables the development of that type of character which will advance the good society.

SPENCER. Herbert Spencer (1820–1903) advanced as the objective for education a life for the individual suffused with pleasure and as full as possible. Its fullness consists in the satisfaction of five kinds of interests, listed here in order of decreasing importance: those pertaining to one's own preservation directly, to it indirectly as does making one's living to begetting and rearing a family, to political and social affairs, and to aesthetic enjoyments. The only knowledge that enables the adequate satisfaction of these interests is scientific, and education of the intellect should be concerned to propagate it rather than knowledge of the classics. Moral education should consist in allowing the natural consequences of mistakes to strengthen knowledge of how to satisfy these interests, and physical education should provide a body that would further their satisfaction. Each individual is charged with finding his own happiness, and the function of government should be merely that of preventing others from infringing upon his pursuit of it. Consequently, education itself should be privately sought and conducted rather than socially compelled and supported.

Spencer held the metaphysical view that reality is unknowable, that it manifests itself in the individual life as phenomena—some vivid and some faint—and that it is expressed in the cosmic dimension as evolution—as change from homogeneous to heterogeneous conditions through differentiation and integration. In evolution survival goes to the fittest; and the fittest are those who find the phenomenon of vivid pleasure associated with the useful and utility in those actions that bring about or constitute "complete living." Education should assist in realizing this end that, in any case, evolution marks out for man.

DEWEY. In the work of John Dewey (1859–1952), the most influential of the twentieth-century philosophers of education, Mill's ideal for education is somewhat simplified and his doctrine of the meaning of ideas, together with Spencer's emphasis on the utility of knowledge, transformed into a criterion for distinguishing knowledge from belief. As we have seen, Mill thought that happiness consists of three distinct factors—pleasure, duty, and self-realization—and he held that education should promote the greatest amount of happiness for the largest number of people. In the place of pleasure Dewey put activity that is satisfactory to the person acting; in the place of duty, the most satisfactory activity; and in the place of self-realization, the fact that the most satisfactory activity is that which the individual most genuinely prefers. The best life, Dewey held, is one in which the most genuinely satisfactory activity is most widespread throughout society. This view depends on his view of human nature.

Human nature cannot be understood in terms of substance and faculty, for there are no such things. Consequently, there can be no single set of activities that characterize all human life, as traditional philosophers and psychologists have supposed. All human beings begin life as biological organisms, filled with unformed energy or impulse, ready to assume whatever direction experience assigns; and since each environment generates a different experience of the world—a different set of patterns of response to it—human beings vary as much as do their environments. The habits that impulse takes on sometimes cease to provide a satisfactory release for it, and in these situations intelligence enters into life to solve the problems thus created. We form hypotheses as to how impulses can be reorganized, look forward to the consequences in action, select those whose anticipation makes us prefer them, act to secure them, and thus test the hypothesis from which they were inferred. Intelligence is the master habit of readjusting others when they break down, and while it characterizes human beings, it does so in no specific way since its possession brings with it no special knowledge but only the ability to acquire any knowledge whatever by finding it in the consequences of action.

Dewey thought of society in terms of group habits. A nation is composed of political parties, religious institutions, courts, etc., and each of these is a complex habit of acting in which many people take part. A society is a set of group habits or institutions that fit together. A good society is one which, by virtue of the ways in which its subordinate institutions fit together, enables growth in satisfaction for its citizens.

Education, according to Dewey, is the process of imposing on the impulse of infants the society or the set of group habits into which the infants are born; it is the perpetuation of society. But it is also a good deal more. For since one of the habits to be imposed upon impulse is that of acting intelligently, education must also foster the reform of society toward an ever better condition. To perpetuate intelligence is to begin its use, and the schools are thus the basis for social progress.

Since there is no single set of abilities running throughout human nature, there is no single curriculum which all should undergo. Rather, the schools should teach everything that anyone is interested in learning. Since a child can learn nothing without using his intelligence, and since this comes into play only when some habit breaks down, he should be inspired with interest in the subject matter he should learn and then made to feel some problem in not actuating that interest or habit. This method requires individual attention to discover particular interests and capabilities. Since the child learns best when he is working with others, he should be given a certain measure of participation in school affairs. In the light of these strictures on curriculum, method, and administration Dewey hoped to produce a child highly endowed with intelligence and disposed to reform society in the direction of the ideal of continually growing satisfactions.

Dewey's ethical ideal was advanced as a justification for this pedagogical objective. To be morally good is to be a set of consequences, deliberately intended and capable of satisfying impulse better than would any other set to which it is preferred; it is a preferred activity. To say that such activity satisfies impulse better than does some other which is rejected is to say that it makes possible more satisfactions in oneself and others than does the other—that it contains the possibility of greater growth. Democracy is a better society than any other because it permits more satisfaction of impulse on the part of more people than does any other. And the intelligent person leads a better individual life than does one who acts from some other habit, such as superstition, because his life contains the opportunity for more satisfactions than does that of one who is hemmed in by dogma. The criterion of growth shows that the objective of education ought to be the democratic society and the intelligent man.

Dewey's theory of knowledge lends support to the reformist tendency in education. The truth of a proposition is its utility, and to know something is to be aware of how to use the known proposition to secure some desir-

able consequence. Consequently, any genuine teaching will result, if successful, in someone's knowing how to bring about a better condition of things than existed earlier. Knowledge is knowing how to do what is useful—a view that may have resulted from Dewey's consideration of Spencer and Mill. This theory of knowledge helps to give the pragmatic flavor to Dewey's philosophy of education.

Dewey's metaphysical reflection helps in the same direction. Traditional metaphysics, such as Plato's, has erred in supposing that truth is a passive apprehension of the real and that its object is eternally separated from the vicissitudes of experience. Traditional metaphysical reflection has forgotten that to mean something is to act to secure certain consequences, and it has therefore overlooked the truth that knowing what is real consists in meaning it or in acting in a certain way to bring about certain consequences. What is real is a set of experiences, each of which is meant by some agent and all of which are connected together in one thing or event by his activity. Dewey used this notion of what is real to justify his method of learning by doing, his view of the curriculum as whatever interests of each student enable him to organize into a unity on his own, and of method as the procedure for arousing interest in organizing or reorganizing the elements of a subject matter.

In Plato's philosophy of education the supernatural realm of the Forms, by lending validity to the just person and the just state, supported the program of education. In St. Augustine's work the educational ideal was organized wholly around God and the theological view of his relation to things; a similar description applies to Thomas Aquinas's thought about education. Comenius also centered his philosophy of education around religious and theological doctrines, but his insistence on the future perfection of human life on earth and on the observation of nature in the search for effective teaching methods marks a beginning in the process of naturalizing the wholehearted supernatural Christian ideal of his predecessors.

Locke found a basis for the goal of education in God's will, but the national welfare, which God's law or the law of nature promotes, and the analysis of it partly in terms of pleasure are additional worldly conditions whose emphasis constitutes a different facet of the disintegration of the supernatural ideal. Rousseau held that God exists, but the chief justification of his objective for education—an internally peaceful life apart from society—lies not in God's having ordained it but in the notion of the general will and its absence from national institutions. Froebel, a follower of Pestalozzi and of Rousseau, made much use of religious language, but by identifying God and the Absolute he removed philosophy of education still further than did Rousseau from a religious center.

Kant held that we cannot avoid belief in God, although he also held that the belief can have no experiential content; but this position effects his educational goal in no way. The chief moral component of that goal is the categorical imperative—a notion Kant wished to conceive wholly in logical terms. The peaceful international state is not justified by being God's will but by being the result of a social life which embodies duty and which constitutes the perfect realization of our intellectual and moral powers. Fichte found the ideal for education in a national existence that would assure Germany of a position of world importance, and Herbart held that individuals and societies that are morally worthwhile are those that satisfy the aesthetic demands of human beings. Spencer made no use of religious propositions in his philosophy of education; nor did Mill, although he regarded great religions as great works of the imagination. Dewey's ideal of a society, containing the possibility of most growth in satisfaction, is completely devoid of religious affiliation. He would probably have said that interest in achieving it can become religious—that, indeed, it should—but by "religious" he would have meant little more than enthusiastic.

The history of philosophy of education reflects a movement evident in other phases of thought—a successive contribution on the part of antiquity to the Christian ideal for transmitting culture from one generation to another and then a gradual elimination from that ideal of supernatural and Christian elements. Of course, at no time has there been a wholehearted and single-minded devotion to any ideal, and there are many who do not accept naturalism today. Nonetheless, one way of understanding the history of philosophy of education is to regard the attitude of philosophers toward the justification and explanation of educational theory as having been expressed first in Plato's classic supernaturalism, next in Augustine's Christian supernaturalism, and then as undergoing a gradual alteration into the wholly non-Christian and naturalistic view represented by John Dewey.

See also Philosophy of Education, Contemporary Issues; Philosophy of Education: Epistemological Issues in; Philosophy of Education: Ethical and Political Issues in.

Bibliography

INDIVIDUAL WORKS

Augustine. "De Ordine." In *Ancient Christian Writers; The Works of the Fathers in Translation*, edited by Johannes Quasten and Joseph C. Plumpe. Vol. V, *Lord's Sermon on the Mount*. Westminster, MD, 1950.

Augustine. "The Teacher." In *Ancient Christian Writers; The Works of the Fathers in Translation*, edited by Johannes Quasten and Joseph C. Plumpe. Vol. XI. *Greatness of the Soul [and] The Teacher*. Westminster, MD, 1950.

Comenius, John Amos. *The Great Didactic*. Translated by M. W. Keatinge. London, 1896.

Comenius, John Amos. *The Way of Light*. Translated by E. T. Campagnac. London, 1938.

Dewey, John. *Democracy and Education*. New York: Macmillan, 1916.

Fichte, Johann Gottlieb. *Addresses to the German Nation*. Translated by R. H. Jones and G. H. Turnbull. Liverpool, 1922.

Froebel, Friedrich. *The Education of Man*. Translated by W. N. Hailmann. New York, 1887.

Herbart, Johann Friedrich. *Outlines of Educational Doctrine*. Translated by Alexis F. Lange. London, 1901.

Herbart, Johann Friedrich. *Science of Education, Its General Principles Deduced from Its Aim*. Translated by Henry M. Felkin and Emmie Felkin. Boston, 1893.

Herbart, Johann Friedrich. *A Text-book in Psychology*. Translated by Margaret K. Smith. New York, 1894.

Kant, Immanuel. "Lecture Notes on Pedagogy." In *The Educational Theory of Immanuel Kant*, edited and translated by Edward Franklin Buchner. Philadelphia, 1908. Also translated by Annette Churton as *Education*. Ann Arbor, MI, 1960.

Kant, Immanuel. "The Natural Principle of the Political Order Considered in Connection with the Idea of a Universal Cosmopolitan History." In *Kant's Principles of Politics*, edited and translated by W. Hastie. Edinburgh, 1891.

Kant, Immanuel. "Perpetual Peace." In *Kant's Principles of Politics*, edited by W. Hastie. Edinburgh, 1891.

Locke, John. *Second Treatise of Civil Government*.

Locke, John, "Some Thoughts concerning Education." In *Locke on Education*, edited by R. H. Quick. London, 1913.

Mill, John Stuart. "On Genius." *Monthly Repository* 6 (1832): 649–659. Partly reprinted in Kingsley Price, *Education and Philosophical Thought*, 449–459. Boston, 1962.

Mill, John Stuart. *On Liberty*.

Mill, John Stuart. *Utilitarianism*.

Pestalozzi, Heinrich. *How Gertrude Teaches Her Children*. Translated by L. E. Holland and F. C. Turner. London, 1894.

Plato. *Republic*. In *Dialogues of Plato*. Translated by Benjamin Jowett, 4th ed. Oxford, 1953.

Quintilian. *Institutio Oratoria*. Translated by H. E. Butler as *The Training of an Orator*. 4 vols. New York, 1920–1922. Loeb Classical Library.

Rousseau, Jean-Jacques. "Considerations concerning the Government of Poland and Its Projected Reform." Edited and translated by Kingsley Price in Ch. IV of *Education and Philosophical Thought*. Boston, 1962.

Rousseau, Jean-Jacques. *A Discourse upon the Origin and Foundation of the Inequality among Mankind*. London, 1761.

Spencer, Herbert. *Education, Intellectual, Moral and Physical*. London, 1861; rev. ed., 1883.

HISTORICAL SURVEYS

Boyd, William. *The History of Western Education*. London, 1952.

Eby, Frederick. *The Development of Modern Education*. New York: Prentice-Hall, 1952.

Laurie, S. S. *Studies in the History of Educational Opinion from the Renaissance*. Cambridge, U.K.: Cambridge University Press, 1903.

Windelband, Wilhelm. *A History of Philosophy*. Translated by James H. Tufts. New York, 1950.

CONTEMPORARY LITERATURE

Adler, Mortimer J., and Milton Mayer. *The Revolution in Education*. Chicago: University of Chicago Press, 1958.

Bode, Boyd Henry. *Progressive Education at the Crossroads*. New York and Chicago: Newson, 1938.

Brameld, Theodore Burghard Hurt. *Patterns of Educational Philosophy; A Democratic Interpretation*. Yonkers-on-Hudson, NY: World Book, 1950.

Broudy, Harry S. *Building a Philosophy of Education*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1961.

Brubacher, John S. *Modern Philosophies of Education*. New York: McGraw-Hill, 1939; 3rd ed., 1963.

Dewey, John. *Experience and Education*. New York: Macmillan, 1938.

Henry, Nelson B., ed. *National Society for the Study of Education Forty-first Yearbook*. Chicago, 1941. Pt. I, *Philosophies of Education*.

Henry, Nelson B., ed. *National Society for the Study of Education Fifty-fourth Yearbook*. Chicago, 1955. Pt. I, *Modern Philosophies and Education*.

Hutchins, Robert M. *Education for Freedom*. Baton Rouge: Louisiana State University Press, 1943.

Kilpatrick, W. H. *The Philosophy of Education*. New York, 1951.

Maritain, Jacques. *Education at the Crossroads*. New Haven, CT: Yale University Press, 1943.

Morris, Van Cleve. *Existentialism in Education; What It Means*. New York: Harper and Row, 1966.

Nunn, Thomas Percy. *Education: Its Data and First Principles*, 3rd ed., rev. New York and London, 1962.

O'Connor, D. J. *Introduction to the Philosophy of Education*. New York: Philosophical Library, 1957.

Peters, R. S. *Authority, Responsibility and Education*. London, 1959.

Phenix, Philip H. *Philosophy of Education*. New York, 1958.

Price, Kingsley. *Education and Philosophical Thought*. Boston: Allyn and Bacon, 1962.

Reid, L. A. *Philosophy and Education*. London, 1962.

Scheffler, Israel. *The Language of Education*. Springfield, IL, 1960.

Smith, B. Othanel, and Robert H. Ennis, eds. *Language and Concepts in Education*. Chicago, 1961.

Ulich, Robert. *Philosophy of Education*. New York, 1961.

Walton, John, and James L. Kuethe, eds. *The Discipline of Education*. Madison: University of Wisconsin Press, 1963.

Kingsley Price (1967)

PHILOSOPHY OF FILM

In one way, the philosophy of film is almost as old as the technology of film; in another way, it is a phenomenon that only emerges fully a century after the earliest screenings of films in 1895. Philosophizing about film has been with us for around a century in the form of the lively debates about the nature of the new medium that sprang up in the wake of its invention. As early as 1907 Henri Bergson had adopted the *cinematographic illusion* as a key metaphor of the scientific, and classical philosophical, conception of time and movement. And in 1916 we see the publication of the first extended philosophical treatise on film, as medium and art form, with the publication of Hugo Münsterberg's (1863–1916) *The Photoplay: A Psychological Study* (2002). So the two-way traffic between film and philosophy—the new medium as a source of philosophical insight and the application of philosophy to the problems thrown up by it—begins.

The publication of Münsterberg's study inaugurated the tradition of *film theory*—reflection on the nature of the medium of film, philosophical in all but name, but typically written by filmmakers, writers, art historians, and cultural critics rather than philosophers per se. Film theorists were preoccupied with the ontology of cinema and with the nature of representation and expression in film, and discussion of these matters typically revolved around the concept of *medium specificity*—the notion, in Münsterberg's words, that the new technology constituted a "specific form of artistic endeavor" (Münsterberg 2002, p. 65) with specific properties and potentials, which demarcated it from the established arts. Münsterberg contrasted film—the *photoplay*—with the stage play, and argued that the key to the power of film was its ability to express human intentional states, such as attention, memory, imagination, and emotions. "The close-up," wrote Münsterberg of one of the key techniques of film, "has objectified in our world of perception our mental act of attention and by it has furnished art with a means which far transcends the power of any theater stage" (Münsterberg 2002, p. 87). This principle of contrast with established art forms, which was ubiquitous in discussions of film through at least its first half century, arose from the desire to demonstrate that film was not merely a technological curiosity, a fairground novelty, or a means of recording and reproduction that might serve to disseminate paintings or plays, but, precisely, a legitimate art form on a par with any of the established arts.

This emphasis on the specificity of film was a legacy of philosophical aesthetics and especially the attempts from the eighteenth century onward to establish a *system of the arts*. Two of the most significant theorists, for example, filmmaker Sergei Eisenstein (1898–1948); and Rudolf Arnheim (1904–, who would become best known as a psychologist of art, made significant allusions to G. E. Lessing's (1729–1781) *Laocoön: An Essay on the Limits of Painting and Poetry* (1766). Lessing had argued that poetry and painting each have their characteristic domains of representation—the temporal and the spatial—and corresponding limits to what they can effectively represent. In his 1938 essay "A New *Laocoön*," Arnheim follows Lessing's example by arguing that theater and cinema, similarly, need to be understood as distinct media with distinct essential features and thus different aesthetic advantages and deficits. Without denying the existence of 'composite' artistic forms—such as opera, in which drama and music are combined—Arnheim argues that, ultimately, theatre is the art of dialogue, while cinema is the art of the moving visual image. Formulating his ideas in the wake of the introduction of the *talkie* in the late 1920s, Arnheim argued that the addition of speech to the movies was a kind of contamination or corruption of the medium proper.

Eisenstein developed the notion of *montage*, which he regarded as the definitive feature of the art of film, through both his filmmaking practice and his theoretical writings, as well as in dialogue with other major filmmakers and theorists of the period—including Lev Kuleshov (1899–1970), V. I. Pudovkin (1893–1953), Dziga Vertov (1896–1954), and Béla Balázs (1884–1949). Initially referring narrowly to the editing of shots, Eisenstein widened the reference of montage to include any technique that involved the interaction of more basic elements: In this sense, one can speak of montage within a shot or between whole sections of a film as much as the montage between two shots literally cut together.

Eisenstein's essay on the *Laocoön* makes reference to Lessing's work in a manner quite different from Arnheim. Where Arnheim draws an analogy between painting and poetry, on the one hand, and cinema and theater, on the other, for Eisenstein, it is the substance of Lessing's claims about painting and poetry and the relationship of cinema to these two forms that is at stake. Cinema—or more particularly, montage—synthesizes the temporality of poetry with the spatiality of painting. Eisenstein ranges widely and generously across literature, painting, theater, and music, and where Arnheim and Münsterberg are concerned to distinguish the characteristics of theater and cinema, Eisenstein more often than not discerns protocinematic techniques in these other art forms. The

specificity of cinema thus emerges for Eisenstein more in terms of the realization and culmination of techniques evident in older media and art forms rather than in the addition of a new medium of art, which stands alongside the traditional forms.

Münsterberg, Eisenstein, and Arnheim were all rooted in the era of *silent* cinema. André Bazin (1918–1958) is widely regarded as the first major theorist of the sound era, and while he, like his precursors, was concerned with the specificity of cinema and often explored the nature of cinema by comparative examination of other media, his perspective on film marks a departure from those theories emerging from the silent era. In "The Ontology of the Photographic Image" (1945), Bazin argues that what is distinctive of and crucial to film is its ability to *capture* the phenomenal world, in the most literal sense; a film is like a fingerprint of reality. Bazin does not wish to deny that films are, like all works of art, the products of those who design them. Greta Garbo (1905–1990) may have been carefully groomed and lit for the camera, but it is still, in a strong sense, the real Garbo that we see in the film. Thus, in contrast to Münsterberg's focus on the rendering of intentional states in films, the distinctive capacity of film as an art for Bazin lies in the way in which human intentionality is bypassed at a certain vital moment in the production of a film, allowing reality to impress itself upon the film unmediated by human intentions or interests.

In *Image and Mind*, Gregory Currie terms this dimension of film and photography "natural counterfactual dependence," which contrasts with the "intentional counterfactual dependence" of painting (Currie 1995, p. 55): The properties of a photograph or a film depend directly on visible properties of the scene before the camera whereas the properties of a painting of the same scene are "mediated by the beliefs of the artist" (Currie 1995, p. 54). And this facet of film is something that filmmakers can facilitate, as in the practice of location shooting where the artifice of studio set construction and the control that such artifice brings with it is foregone in favor of the relative unruliness of real spaces. Such techniques bring out the special kind of realism that (on this account) is inherent in the medium as such.

CONTEMPORARY PHILOSOPHY OF FILM AND THE PROBLEM OF SPECIFICITY

Given the existence of a rich tradition of film theory—the surface of which is only scratched here—in what sense is it true to say, that the philosophy of film only coalesced as a field of debate a century after the invention of the medium? Following Bergson, there have been other important contributions by professional philosophers, including Maurice Merleau-Ponty (1908–1961), Bazin's contemporary and an influence upon him. The American philosopher Stanley Cavell has made a distinctive contribution (to which we will return) by developing and elaborating an ontology of cinema incorporating Bazin's key insights. And there have been other isolated philosophical essays on film. But it is not until the 1990s that a continuous debate about film emerges among professional philosophers, eventually establishing itself as a subdomain within aesthetics and the philosophy of art—a field of debate sufficiently developed to warrant a separate entry in this encyclopedia. Two rather divergent areas of debate have emerged that, for good or ill, generally fall in line with the division between modern analytic and Continental philosophy. In relation to the latter, there is a substantial literature around the work of Gilles Deleuze. Alongside the literature on and by Deleuze stands work by other contemporary Continental philosophers, such as Jean-François Lyotard, Paul Virilio (1932–) and Slavoj Žižek (1949–). Through much of this work, the influence of psychoanalysis is evident.

Deleuze's approach to cinema, as advanced in his two-volume *Cinema* (1992), is based on a fundamental revaluation of Bergson's remarks on the relationship between cinema, movement, and time. In *Creative Evolution* (1907), Bergson argued that both classical philosophical and modern scientific conceptions of movement in fact eliminated movement as an authentic phenomenon by representing motion as a series of immobile instants strung together. The mechanism of cinema realized this conception in literal terms: a succession of still frames which, when projected in sufficiently quick succession, generate an impression or illusion of movement. Deleuze argues, however, that the cinema also enables, and is a part of, the recognition of movement as an irreducible phenomenon. And as cinema evolves over the course of the twentieth century, it provides us not only with an image of movement but one of time—in the Bergsonian sense of *duration* of time as a continuous, experiential whole.

The conception of the *philosophy of film*, and of philosophy more broadly embodied by Deleuze's approach to cinema, is—at least on its own understanding—in marked contrast to more widely accepted notions of philosophy. Rejecting the idea that the philosophy of film *reflects on* the phenomenon of film, Deleuze argues instead that the philosophy of film—like philosophy

more generally—is a creative activity and in this sense is parallel to the activity of filmmaking rather than standing above or outside it. Where filmmakers create through the medium of cinema, in the form of sequences of movement and duration, philosophers create concepts. Deleuze is thus eager not only to play up the creative character of philosophy as he understands it but to emphasize the conceptual value of filmmaking.

The growing literature on cinema within Anglo-American aesthetics comprises the second major branch of contemporary philosophy of film. The main intellectual reference points here are analytic philosophy of mind and language, cognitive psychology, and Wittgenstein. The two contemporary conversations on the philosophy of film are largely separate even if the participants in each can hear the other conversation and occasionally might even talk to each other. There are certainly points of connection: Deleuze's claim that the cinema provides us with an image of movement resonates with the debate in analytic philosophy of film concerning the sense in which the motion we see in a film is real (rather than merely illusory) while his claim that the postwar era witnesses the flourishing of a cinema that privileges time rather than movement echoes the claim within Anglo-American film theory that much art cinema liberates time and space from their traditional subordination to the demands of narrative.

Among the philosophers who have helped to establish the analytic strand of the philosophy of film, none have contributed more than Noël Carroll (1947–); and among the many orthodoxies that Carroll has challenged is the very idea of medium specificity.

In "Forget the Medium!" (2003), Carroll questions both the coherence of the concept *the* medium of film as well as the prescriptivism that typically follows on from the positing of specific qualities that are thought to be distinctive of the medium. He points out that if we think of the medium in terms of the tools and materials of an artistic practice, few, if any, art forms will be defined by a single, fixed medium of expression (and, more radically, he suggests that some art forms may not have *a medium* at all). *Painting* really encompasses a whole range of possible means of marking a surface in order to create a visual design, just as the creation of music encompasses a vast array of instruments for shaping sound. We can, however, understand why earlier film theorists may have focused on the idea of a new medium since the technology of film ushered in a type of depiction that was different in kind, and not merely in degree, from anything that preceded it: moving, photographic depiction. The developments in the basic technology of film were, for the first thirty years of cinema, all refinements of this technology, and so it could appear to have an underlying stability and unity that made it apt to think of in terms of a single medium.

Later technological developments, however, begin to strain the concept of a single and stable medium—Arnheim's alarm at the coming of synchronous sound was shared by many filmmakers and theorists of the time. The advent of television and video raises equally difficult questions—if film is a unique and distinctive medium, should we posit still another new medium of the electronic moving image? And still another one for the digital moving image? Many have answered these questions in the affirmative, erecting boundaries between the various types of moving image. The emergence of new moving image technologies has often led to attempts at distinguishing the specificity of each of these media—such specificity usually taking account not only of the material nature of the technology but of its institutional and social deployment: Thus, television is said to have its own specificity, distinct from that of film, not only because of the electronic basis of broadcasting but the corporate nature of most television output; its continuous *flow*; and the small-screen, domestic context of television viewing. Video, in turn, has been defined dialectically against television, focussing on the portability, immediacy, capacity for instant replay and live feedback, and nonnarrative experimentation characterizing video art and activist video.

Carroll, however, contends that the positing of a succession of media *specificities* only compounds the error of thinking of film as a medium and proposes, instead, that we engage in some conceptual pruning and relandscaping. In place of the medium of film, we should think in terms of the *art form of the moving image*. This superordinate category captures what was new when cinema first emerged and what continues to mark works of this type off from paintings, photographs, operas, novels, and so forth, but it does so without tying it to any particular technology.

From another angle, the emergence of computer-generated imagery as a pervasive feature of mainstream narrative filmmaking has led some theorists to argue that there really was something importantly distinctive about the prototypical live action, photographic film characteristic of the first century of cinema but that that distinctiveness is now disappearing. As the computerized rendering of moving picture settings and characters becomes commonplace—whether through the modifica-

tion of a live action source or through digital creation from the ground up—the Bazinian idea of film as an imprint of reality is weakened. As we watch *The Lord of the Rings*, we really cannot be sure which parts of the image were created through the act of photographic recording (of a set, a real location, a performer) and which were generated digitally; all we can be confident of is that the film as a whole represents a blending of these methods. As a consequence, according to Lev Manovich (1960–) in *The Language of New Media*: "cinema can no longer be clearly distinguished from animation." Far from being distinct from painting, by virtue of the direct causal relationship between image and referent, cinema in the digital age has become, instead, "a subgenre of painting" (Manovich, p. 295). Manovich's view of digital media forms the mirror image of and complement to the realist ontology of film favored by Currie, for whom both animation and abstract film are, at best, marginal instances of film.

OTHER DEBATES

So we find in contemporary philosophy of film a continued debate about the very idea of *film*, as a unified phenomenon and coherent field of study. However, there are a multitude of other debates underway, intersecting at various angles with arguments about the ontology of the medium. The themes and questions being addressed include the following:

(1) *The perception of moving images*. What do we see when we look at a moving photographic image? Do we see a representation? Or is such a moving image *transparent*, in the sense that we see the objects depicted through the moving image, as Kendall Walton (1939–) has argued? Do we *imagine seeing* that which is depicted, or do we engage in *perceptual imagining*, as Currie contends, in which we imagine that certain things are true, based on the moving images we see, but we do not imagine seeing those things? To what extent is our ability to comprehend moving pictures dependent on certain natural perceptual capacities and to what extent on learned conventions?

(2) *Identification, emotional response and ethics*. In what sense and to what extent might we be said to *identify* with the characters, or the camera, when we watch a film? Do we typically empathize or sympathize with characters? Are we subject, in any sense, to an *illusion*? Are our emotional responses to film largely irrational and paradoxical, or is there a kind of rationality to them? Do these emotions have a significant relationship to the ethical value of cinema—its ability, in small or large ways, either to corrupt or to educate? Does the medium of film, or particular forms of filmmaking, embody ideological values and beliefs, such as those bearing on gender or ethnic identity?

(3) *Authorship, intention, and expression*. Given the collective basis of almost all film production, can a film be *authored* in just the same way as a poem or a painting? Does the fact of multiple authorship affect the expressive capacities of film, relative to other art forms, or the way in which we interpret and appreciate films?.

(4) *Fiction and nonfiction*. How does the psychology of watching fiction differ from the psychology of watching documentaries? Does a filmic fiction share more with a novel than a documentary film; does a documentary share more with written history or reportage than it does with a fiction film? Is there a sense in which all films have a documentary dimension?.

FILM AS PHILOSOPHY

One important question that has become a focus of debate asks: To what extent might film be a vehicle of philosophy as opposed to its subject? Can film serve as a distinct medium through which the act of philosophy might be undertaken as opposed to a phenomenon to which philosophy is applied? Can film philosophize? Eisenstein was one of the most forthright and ambitious defenders of the idea that film might act *as* philosophy, with plans for a film version of Karl Marx's *Das Kapital* and a host of arguments in support of *intellectual cinema*. But Eisenstein was not alone. According to Deleuze, the cinema creates new concepts by its own distinctive means. And Cavell has argued that certain key cinematic genres, such as the "Comedy of Remarriage" discussed in *Pursuits of Happiness* (1981), give expression to the philosophical problem of skepticism, insofar as they dramatize, within the intimate arena of romance, the difficulty of knowing the thoughts and feelings of others. Moreover, in the hands of some writers, the *film as philosophy* thesis is very much akin to the treatment of literature as a kind of philosophy, a proposal advanced most explicitly by Martha Nussbaum.

To a considerable degree, the plausibility of the proposal depends on the conception of philosophy that is assumed within it; so the debate is ultimately driven onto the terrain of metaphilosophy. On the one hand, to the

extent that one conceives of philosophy as a professional discipline whose central goals are the posing of questions and the making of arguments in a reasonably robust and formalized sense, then the idea that film might act as an effective medium for such goals looks strained. On the other hand, to the extent that one thinks of philosophy more broadly as a form of self-conscious reflection on any aspect of life that we usually accept unthinkingly, then film—along with art in general—looks much more promising as a means of engaging in such reflection and thus as a form that philosophy might take.

Among the proponents of the film as philosophy thesis are Stephen Mulhall (1962–) and Thomas Wartenberg. Where Cavell focusses on *classical Hollywood* films (especially screwball comedies and melodramas), Mulhall has developed and extended Cavell's approach to encompass contemporary Hollywood filmmaking through studies of the *Alien* tetralogy and the *Mission: Impossible* films. For Mulhall the series of *Alien* films embody philosophical reflection not only on the overt themes of the films, such as human embodiment and the process of reproduction, but on various aspects of the nature of commercial filmmaking itself, including stardom, authorship, and sequeldom. Mulhall is emphatic about the strength of his claims, stating that certain films should be seen "as thinking seriously and systematically about [philosophical views and arguments] in just the ways that philosophers do" (Mulhall 2002, p. 2).

Wartenberg has emphasized the various ways in which films might make genuine contributions to philosophy even if they cannot be construed as making arguments in any conventional sense, including the creation of thought experiments that challenge habitual assumptions and the provision of illustrations that are integral to a philosophical claim, and thus cannot be discarded without damage to the claim in question. Wartenberg has argued that the first *Matrix* film engages us philosophically by creating a thought experiment resembling René Descartes's image of the evil demon, challenging our confidence in the knowledge we gain from sense experience. In other work Wartenberg has emphasized the insights that films may proffer on the terrain of social, political, and moral philosophy. Other authors have made parallel claims about the philosophical significance of various art and avant-garde films, but what unites Cavell, Mulhall, and Wartenberg and makes them distinctive is their emphasis on popular filmmaking, the type of filmmaking that might seem the least congenial—and thus offering the greatest challenge—to the film as philosophy hypothesis.

Counterarguments to these proposals stress the special nature of philosophical knowledge (in normative, if not descriptive terms); the central role of explicit reasoning and argument within it; and the distinctness of philosophy from cognition, self-reflection, and knowledge considered more generally. Paisley Livingston (1951–) has argued that proponents of the *bold* version of the thesis, for whom films can make original philosophical contributions exploiting the specific properties of the medium, are faced with a disabling dilemma: If the contribution can be paraphrased, then any uniqueness premised on medium specificity disappears; if it cannot be paraphrased, then it is difficult to see how a contribution is being made to philosophy proper, when conceived as a discursive discipline. Murray Smith (1962–) has made the complementary point that the *nonparaphrasability* of art is one of its most significant values, and one that brings it into tension with the widely accepted philosophical goals of clarity and explicitness.

Wherever one stands on this issue, and on the question of specificity, the emergence of a debate on the idea of film as philosophy, alongside the diversity of other questions and debates described here, testifies to the seriousness with which the moving image is now taken by philosophers and the consolidation of the philosophy of film at the outset of cinema's second century.

See also Aesthetic Qualities; Art, Expression in; Art, Formalism in; Art, Interpretation of; Art, Ontology of; Art, Performance in; Art, Representation in; Cavell, Stanley; Continental Philosophy; Deleuze, Gilles; Descartes, René; Lyotard, Jean François; Marx, Karl; Nussbaum, Martha; Visual Arts, Theory of the.

Bibliography

Allen, Richard. "Looking at Motion Pictures." In *Film Theory and Philosophy*, edited by Richard Allen and Murray Smith, 76–94. Oxford: Clarendon Press, 1997.

Allen, Richard, and Murray Smith. *Film Theory and Philosophy*. Oxford: Clarendon Press, 1997.

Arnheim, Rudolf. *Film as Art* [1933]. London: Faber, 1983.

Arnheim, Rudolf. "A New *Laocoön*: Artistic Composites and the Talking Film." In *Film as Art*, 164–189. London: Faber, 1983.

Bazin, André. "The Ontology of the Photographic Image." In *What Is Cinema*? Translated by H. Gray, 9–16. Berkeley: University of California, 1967.

Bergson, Henri. *Creative Evolution*. Translated by Arthur Mitchell. New York: Modern Library, 1944.

Carroll, Noël. *The Philosophy of Horror; or, Paradoxes of the Heart*. New York: Routledge, 1990.

Carroll, Noel. "Forget the Medium !" In *Engaging the Moving Image*, 1–9. New Haven, CT: Yale University Press, 2003.

Cavell, Stanley. *Pursuits of Happiness: The Hollywood Comedy of Remarriage*. Cambridge, MA: Harvard University Press, 1981.

Cavell, Stanley. *The World Viewed: Reflections on the Ontology of Film*, 2nd ed. Cambridge, MA: Harvard University Press, 1979.

Cubitt, Sean. *Videography: Video Media as Art and Culture*. New York, 1993.

Currie, Gregory. *Image and Mind: Film, Philosophy, and Cognitive Science*, Cambridge, U.K.: Cambridge University Press, 1995.

Deleuze, Gilles. *Cinema 1–2*. Translated by Hugh Tomlinson and Barbara Habberjam. London: Athlone Press, 1992.

Eisenstein, Sergei. *Selected Works 1: Writings 1922–1934*. Translated and edited by Richard Taylor, London: British Film Institute, 1988.

Eisenstein, Sergei. *Selected Works 2: Towards a Theory of Montage. 1937–40*. Translated by Michael Glenny, edited by Michael Glenny and Richard Taylor. London: British Film Institute, 1991.

Freeland, Cynthia. "Feminist Frameworks for Horror Films." In *Post-Theory: Reconstructing Film Studies*, edited by David Bordwell and Noel Carroll, 195–218. Madison: University of Wisconsin Press, 1996.

Gaut, Berys. "Film Authorship and Collaboration." In *Film Theory and Philosophy*, edited by in Richard Allen and Michael Smith, 149–172. Oxford: Clarendon Press, 1997.

Livingston, Paisley. "Cinematic Authorship." In *Film Theory and Philosophy*, edited by in Richard Allen and Michael Smith, 132–148. Oxford: Clarendon Press, 1997.

Livingston, Paisley. "Theses on Cinema as Philosophy." *Journal of Aesthetics and Art Criticism*. Spec. issue on film as philosophy (January 2006): 1–8.

Manovich, Lev. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

Mulhall, Stephen. *On Film*. London: Routledge, 2001.

Münsterberg, Hugo. *The Photoplay: A Psychological Study, and Other Writings*, edited by Allan Langdale. New York: Routledge, 2002.

Nussbaum, Martha. *Love's Knowledge*. New York: Oxford University Press, 1990.

Plantinga, Carl. *Rhetoric and Representation in Nonfiction Film*. New York: Cambridge University Press, 1997.

Smith, Murray. *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press, 1995.

Smith, Murray. "Film Art, Argument, and Ambiguity." *Journal of Aesthetics and Art Criticism*. Spec. issue on film as philosophy (January 2006): 26–35.

Walton, Kendall. "Transparent Pictures: On the Nature of Photographic Realism." *Critical Inquiry* 11 (1984): 246–277.

Wartenberg, Thomas E. "Philosophy Screened: Experiencing *The Matrix*." In *The Philosophy of Film: Introductory Text and Readings*, edited by Thomas E. Wartenberg and Angela Curran, 270–83. Malden, U.K.: Blackwell, 2005.

Wartenberg, Thomas E. *Unlikely Couples: Movie Romance as Social Criticism*. Boulder, CO: Westview Press, 1999.

Wilson, George M. *Narration in Light: Studies in Cinematic Point of View*. Baltimore, MD: Johns Hopkins University Press, 1986.

Murray Smith (2005)

PHILOSOPHY OF HISTORY

The term "philosophy of history" probably covers a larger variety of endeavors than similar terms such as "philosophy of law" or "philosophy of science." It is hard to bring under one definition the many philosophical questions and responses that are concerned with history. One reason for this, which has long been acknowledged, is that the English term "history," like its cognates in many Western languages (*histoire*, *Geschichte*), is normally used to refer to two distinct, though related, things. On the one hand it refers to the temporal progression of large-scale human events, primarily but not exclusively in the past; on the other hand, "history" refers to the discipline or inquiry in which knowledge of the human past is acquired or sought. Thus "philosophy of history" can mean philosophical reflection on the historical process itself, or it can mean philosophical reflection on the knowledge we have of the historical process. Philosophers have done both sorts of things, and this has led to a distinction between "substantive" (or sometimes "speculative") and "critical" (or "analytical") philosophy of history. The first is usually considered part of metaphysics, perhaps analogous to the "philosophy of nature," whereas the second is seen as epistemology, as in the "philosophy of science." While this distinction has been useful, it becomes blurred when we find some philosophers doing a mixture of both, and others, while certainly reflecting philosophically on history, doing neither. This entry begins with the standard distinction, only to see it lose some of its usefulness in the course of the exposition.

1. "SUBSTANTIVE" PHILOSOPHY OF HISTORY: PHILOSOPHICAL REFLECTION ON THE HISTORICAL PROCESS

The term "philosophy of history" originates with Voltaire in the 1760s, but it is most closely associated with German philosophers of the Enlightenment and post-Enlightenment periods: Kant, Herder, Hegel, and Marx. Hegel's "Lectures on the Philosophy of History," delivered in the 1820s and published shortly after his death, have dominated the discussion. The lectures represent Hegel at the height of his influence, and their relatively brief (less than a hundred pages) introduction is as clear and straightforward as it is comprehensive. Soon translated into other languages (e.g., English in 1857), it is probably the most widely read of Hegel's works. So great was Hegel's impact that his approach to history became paradigmatic not only for many who followed his lead, but

also for those who later attacked the very project of the philosophy of history. What is more, philosophers who reflected on history before Hegel are often thought to have been engaged in the same kind of inquiry he was. But this is anachronistic, and misleading. The substantive philosophy of history is often described, in keeping with Hegel, as the search for the meaning and purpose of world history, and for the force that drives history toward its goal. While this describes many instances of reflection on the historical process, it is a simplification and is not necessarily an apt description of philosophical thought about history prior to Hegel. The most general description of the substantive philosophy of history is that the philosopher tries to "make sense" of the historical process, usually in the face of evidence to the contrary. But the "sense" that the philosopher seeks varies considerably: sometimes it is rational sense, sometimes moral sense, sometimes religious sense.

Philosophical reflection on the historical process seems to originate in early Christian philosophy, which is in turn indebted to the Jewish conception of time. The Hebrew scriptures introduce historical time into a world dominated by cyclical and ahistorical conceptions of time. Indian, Persian, and Greek thought are based on unchanging patterns and eternal recurrence, in which individual events, both natural and human, get whatever significance they have from reflecting, imitating, or instantiating these timeless forms. The sequence of individual events is not "going anywhere." Their essence, what gives them their being, lies outside of time altogether. In spite of the compelling historical accounts left by Herodotus and Thucydides, for Greek philosophers even political arrangements—constitutions such as aristocracy, monarchy, democracy—are portrayed, in the classical texts of Plato and Aristotle, for example, as following cyclical patterns of rise, fall and repetition.

By contrast, for the ancient Jews, human events—both political and religious—get their significance not from a "vertical" and imitative relation to eternal patterns, but from a "horizontal" relation backward and forward to other events in real time: backwards to creation, Adam's fall, God's covenant with his people, its captivity, exile, rulers, and heroes, and so on; forward to the redemption of God's people with the coming of the Messiah. Time is the story of a people's progress from creation through perils, dangers, and risks to final salvation. Christianity takes up this historical conception of time and intensifies it, first by affirming the coming of the Messiah as a central, real historical event, in the *middle* of history, as it were, pointing ahead to a final salvation in the second coming; and second, by extending the promise of salvation to all mankind through a progressive spread and universal triumph of Christianity. Creation, the fall, incarnation, and last judgment are unique, unrepeatable occurrences, and individual events and deeds, both human and divine, are arrayed along a line of time that extends from beginning to end. Given this conception, events are coming from somewhere and are going somewhere in time. Origin and destiny give meaning to human events and actions.

This conception of historical time is not itself a philosophy of history but a cultural and religious worldview. Philosophical reflection begins when this conception generates problems, as it did in the age of Augustine. This philosopher struggled with problems of good and evil, freedom and divine justice, responsibility and punishment. History entered the picture when these concepts were projected onto the stage of the large-scale social events of his own time. The conversion of the Roman Empire under Constantine (323 CE) was seen by early Christian theologians as the vindication of their religion and the harbinger of its eventual triumph throughout the world. During Augustine's time (354–430) the empire was under attack by barbarians, Rome itself had been invaded, and the empire seemed in danger of destruction. Pagans took this as a sign that Christianity was responsible for the demise of the empire, and Christians wondered why God seemed to be punishing Rome rather than rewarding it for its conversion and crowning it with glory. Here it was historical developments, rather than just evil deeds and events, that seemed at odds with religious doctrine, and this constituted the problem Augustine felt the need to solve, addressing both pagan and Christian audiences.

In response, Augustine denied that salvation and divine justice were to be sought in human secular history or its political or even religious institutions. Instead, they were to be found in the City of God, whose citizens have their real life outside secular time. Augustine had already considered the notion of time as limited by eternity in trying to reconcile free will and God's foreknowledge. Augustine's response to the problem of history was to seek the meaning and purpose of history not in history itself, but rather outside of time altogether. In Augustine's thought, the Platonic conception of the timeless realm triumphs over the religious view of history handed down from Judaism and Christianity. As often occurs in the history of Christian thought, Greek philosophy comes to the rescue of the religious worldview. At the same time Augustine inaugurates the tradition of Christian apolo-

getics, later called theodicy: justifying God's ways to humans. Because of the presuppositions that frame Augustine's whole discussion, his project might best be called a theology of history.

Two things should be noted about history as Augustine conceives it: First, as we have noted, its purpose and goal lie not in historical time but outside and beyond it; second, in spite of Augustine's emphasis on human freedom, the driving force behind historical change, what links human events to their ultimate purpose, is the divine will. These two features of history remained more or less constant in the Christian tradition until the time of the enlightenment. Jacques Bénigne Bossuet's *Discourse on Universal History* (1681) still shares in this conception. He sees the world in apparent moral disorder, with the authority of the church being challenged, but assures his readers of the guidance of divine providence and the ultimate salvation of the faithful.

Giambattista Vico, in the *New Science* (1725–1730), also appeals to the idea of providence, but his approach to history is more novel and more modern, because he thinks of providence as embodied in rational, developmental laws rather than acts of divine intervention. He also believes that providence uses narrow human self-interest and self-love to further its own higher ends, a concept usually seen as foreshadowing Hegel's idea of the cunning of reason. Vico is also known for dignifying historical knowledge, in the face of both ancient and modern disdain for it when compared to our knowledge of nature. Because human beings make history through their own acts, Vico believes, they are capable of knowing it. Because God creates nature, only he can truly know it. In this Vico challenges his contemporaries, the Cartesian defenders of the new mathematical science of nature as the paradigm for all knowledge.

In the French Enlightenment, humans take center stage and their reason makes them capable of shaping their own destinies. Human events come under calculation and control. The future is no longer something to be prophesied or predicted, but something to be produced. The legitimacy of rulers can be questioned, and the people can overthrow them. History begins to look like a progress from a past of darkness and superstition into the light of reason and human self-determination. The purpose and goal of history now lies not outside and beyond it, but within it at some attainable point in the future. It is the result of human rather than divine agency, and it is now conceived not as salvation but as emancipation.

Even though Voltaire introduces the term "philosophy of history" it is possible to argue that his view of history, shared by the enlightenment *philosophes* and the revolutionaries of the eighteenth century, was not so much a philosophical reflection on history but again, like the religion of the Jews and early Christians, an emerging political and cultural worldview. The philosophy of history begins, as before, when this worldview generates problems. The late enlightenment period produced a vast new literature of discovery and travel, which led among other things to the beginnings of history as something like an academic discipline with critical methods and justifiable assertions. While this trend was not completed until the nineteenth century, even its beginnings allowed for a new distinction between our warranted knowledge of the past and our beliefs about history's overall direction and goal.

Thus Kant's forays into the philosophy of history tend to raise critical questions about what the enlightenment philosophers never doubted. A late text (1798) bears the title "An old question raised again: is the human race constantly progressing?" But even his earlier essay, "Idea for a Universal History from a Cosmopolitan Point of View" (1784), his major contribution to the philosophy of history, argues only for the limited thesis that the course of past history "permits us to hope" for "a steady and progressive though slow evolution" toward a better state for mankind (1963, p. 11). Kant wants to share the enlightenment point of view, just as he wants to endorse the claims of natural theology, but his critical reason forces him to limit its pretensions. As should be expected when reading Kant, of course, in no way is the idea of divine providence taken for granted. Progress in history, should it be found, would be toward "the achievement of a universal civic society which administers laws among men" (p. 16), which is "the most difficult and the last [problem] to be solved by mankind" (p. 17). He discusses at some lengths the difficulties of such an achievement, asserting as he does elsewhere that it would require solving "the problem of a lawful external relation among states" (p. 18). This is the greatest difficulty of all, because we can see the same antagonism among states as among individuals, which has led again and again to war. But after "devastations, revolutions, and even complete exhaustion," nature brings states to the realization that they must move "from the lawless condition of savages into a league of nations" (p. 19).

By the time he reaches this point the status of Kant's discourse on history should be clear to the reader. He is not making claims about the actual course of history; rather, he is outlining the ideal conditions under which alone, he thinks, history could exhibit any progress.

Because these conditions are in his day far from having been realized, Kant's claims are clearly prescriptive and moral in character. Thus he can assure practicing historians that he is making no attempt to displace their work, because he is propounding an Idea of world history based upon an *a priori* principle (p. 25), an "[I]dea of how the course of the world must be if it is to lead to certain rational ends" (p. 24).

By using the term "Idea," a *terminus technicus* from the *Critique of Pure Reason*, which the translators signal by means of capitalization, Kant indicates a rational concept whose empirical reality not only is not, but, according to the *Critique*, *cannot* be exhibited in experience. But, like human freedom itself, neither can its possibility be empirically denied. Thus the course of history does not provide evidence that the "civic union of the human race" will ever be achieved, but neither does it prove that it never will be. Its realization must at least be regarded as possible, and the Idea that we have of it may help bring it to pass (p. 24). Kant is telling us not where history is going but where it *ought* to be going. Only in this minimal sense can philosophy help "make sense" of history, namely by articulating the "cosmopolitan standpoint" from which alone it can be freed from its apparent moral chaos. And by showing that its moral realization is at least possible, it "permits us to hope" for a better future. Kant's concept of hope is usually associated with his philosophy of religion and refers to the individual's hope for salvation in the world to come. But here he rationally justifies hope for a better future for mankind on earth.

In *Idea for a Universal History*, the concepts of a universal civic society, or league of nations, and of history as progressing toward it, legitimize certain political choices. They are Ideas capable of guiding our action in the social sphere. Kant is anticipating the project of expanding his ethical principles, with such notions as a kingdom of ends, into a political theory. Ethics and politics alike belong to Kant's practical philosophy, not his theoretical philosophy. Their central concern is not with what is the case but with what we ought to do. And the same is true of his philosophy of history.

Johann Gottfried von Herder, a younger contemporary of Kant's, is another German philosopher who reacts critically to the enlightenment's views of history. In his *Ideen zur Philosophie der Geschichte der Menschheit* (1784–1791), he undertakes a universal history, and for him, as for Voltaire, this means expanding the traditional scope of history to include non-European peoples. But Herder takes this insight in a different direction. While the thinkers of the French Enlightenment sought proof of the universality of human reason, Herder by contrast is struck by the diversity and particularity of human nature, embodied in distinct peoples and cultures. Rejecting the Enlightenment's emphasis on reason, legislation and science, Herder sees human nature in the expressions of feeling, such as art, music, poetry, and custom.

The Enlightenment philosophers saw the growth of scientific rationality expanding to the political realm and imagined a future in which reason triumphed over the dark forces of superstition and emotion. Herder, with his emphasis on diversity and culture, was less convinced that history was moving in any unified direction, much less a progressive one. True, his devout Protestantism kept him from embracing the complete cultural relativism that many would later draw from his work. But in contrast to Kant, whose sympathies still lie with the Enlightenment, Herder becomes one of the first great figures of the Romantic movement that grew up in opposition to it.

It is against this background of the Enlightenment and its German critics that Hegel's classic text must be understood. He begins by distinguishing a "philosophische Weltgeschichte" from history proper; philosophy, he says, has "thoughts of its own," *a priori* thoughts, to bring to the study of history (1988, p. 10). But the "only" thought that philosophy brings to the study of history is that of reason—"that reason rules the world," and thus that world history like everything else can be seen as a rational or reasonable (*vernünftig*) affair (p. 12). Reason not only sets the goal for history but also governs the realization of that goal. Hegel did not invent this idea, he reminds us; the idea that reason rules the world is that of Anaxagoras, and it has also been expressed in the idea of divine providence. This too suggests a rational plan, God's plan, but providence is usually portrayed as being hidden from us. Unwilling to settle for pious ignorance, Hegel believes that the rationality of providence can be known and explained. If we take seriously the idea of providence, the demonstration of its rationality would amount to a theodicy or "justification of God" (p. 18).

The embodiment of reason is spirit (*Geist*), both in individuals and in peoples, whose nature is to be conscious and self-conscious, and whose actualization is to be autonomous and self-sufficient, that is, to be free. But this actualization is a temporal process, and that process is history. Spirit actualizes itself and achieves freedom through history, drawing its energy from human passions and intentions; but the result of this process is often at odds with the actual intentions of the individuals and peoples involved. It is here that Hegel's speaks of the "cunning of reason" (p. 35), because reason achieves ends

of its own by using the ends of others. In history, it is only when individuals and peoples organize themselves into states that freedom can finally be truly actualized. It is here, in law, the ethical life of the community and political order, not in the mere absence of constraint, that the "positive reality and satisfaction of freedom" are to be found (p. 41).

The actual course of history can be seen as the display of human perfectibility leading toward the realization of freedom. This pathway is not a smooth one, however, but consists in the spirit's "hard and endless struggle against itself." Spirit hides its own nature from itself, and is even "proud and full of enjoyment in this self-estrangement" (p. 59). Individuals and peoples struggle against each other, and many morally good and virtuous people suffer unjustly. But history moves on a different plane, and here the acts of individuals, especially those of the great figures of history, are not to be judged by moral standards. It is the spirit of peoples, not individuals, that are the agents of history, but these, "progressing in a necessary series of stages, are themselves only phases of the one universal Spirit: through them, that World Spirit elevates and completes itself in history, into a self-comprehending *totality*" (p. 82). The self-comprehension of world spirit is philosophy itself.

In several places Hegel presents in the broadest outlines the necessary stages through which the world spirit has passed on its path toward the realization of freedom. In the ancient "oriental" world only one—the emperor or tyrant—is free. In the Greek and Roman worlds only *some* persons are free. It was first the "Germanic peoples, through Christianity, who came to the awareness that every human is free by virtue of being human" (p. 21). The realization of freedom is the goal that gives meaning to what happens in history, and this realization takes place within history itself, not beyond it. Moreover, it has occurred or is occurring in "our world," "our time" (Hegel 1956, p.524).

Karl Marx is usually seen as a continuation of the classical period of the philosophy of history. Marx admitted some indebtedness to Hegel, but thought of himself as the anti-Hegel, whose idealism "stands on its head" and "must be turned right side up again." More important, Marx rejected not only Hegel, and Hegel's philosophy of history, but academic philosophy as a whole, wanting to be read and understood strictly as a social theorist and reformer. Yet it seems beyond doubt that Marx expounds a philosophy of history in the "classical" sense. Even understood as a blueprint for reform or revolution, his work is founded on and cannot be understood apart from an account of history. This account is summarized neatly by his collaborator, Friedrich Engels, in his preface to the 1888 English edition of the *Communist Manifesto*, in which he states what he calls the "fundamental proposition of Marxism." "In every historical epoch," Engels writes, "the prevailing mode of economic production and exchange and the social organization necessarily following from it" form the basis of that epoch. "Consequently the whole history of mankind … has been a history of class struggles, contests between exploiting and exploited, ruling and oppressed classes." The outcome of this history is that "nowadays, a stage has been reached" where the emancipation of the exploited and oppressed class—the proletariat—from the exploiting and ruling class—the bourgoisie—would entail "at the same time, and once and for all, emancipating society at large from all exploitation, oppression, class distinctions, and class struggles" (Marx and Engels 1998, p. 48). The notion of history as class struggle recalls Hegel's description of the spirit's "hard and endless struggle against itself," its "self-estrangement" in which it "must overcome itself as its own truly hostile hindrance" (Hegel 1988, p. 59). In the background of these descriptions is Hegel's famous account in his *Phenomenology of Spirit* of the struggle between master and servant, an account that can be interpreted in economic and material terms, and which is certainly an account of exploitation and oppression. As Marx admits, this is the origin of a "dialectic" account of the movement of history, which Marx appropriates for his own purposes.

Different as they are from each other to their adherents, Hegel and Marx both reveal their indebtedness to the Enlightenment. For both, it is human affairs and strivings, not divine actions, that drive history, and its purpose or culmination, conceived not as salvation but emancipation, lies within history, not outside or beyond it. Yet unlike the Enlightenment idea of progress, their conception seems to require an end of history. Hegel often speaks as if it has already arrived, and Marx projects it into the near future. Both are unclear what happens after that.

This was but one of many conceptual problems that led to widespread criticism of Hegel's and Marx's philosophies of history and to a general mistrust of the whole project. The idea of attributing a purpose or goal to history as a whole became suspect. Hegel's speculative idealism fell on hard times, and his philosophy of history was seen as the worst manifestation of its extravagant pretensions. It was also read by many, rightly or wrongly, as a glorification of the Prussian monarchy as the culmi-

nation of history. Marx's apparent belief in an inevitable outcome of history was not widely accepted by philosophers, even those sympathetic to his proposed political and social reforms; only the official orthodoxy of the Soviet Union and other communist states took it seriously. Sweeping treatments of history as a whole and the rise and fall of civilizations, such as Oswald Spengler's *Decline of the West* (1918–1922) and Arnold Toynbee's *A Study of History* (1934–1954), were reviewed in the popular press, but not taken seriously by academic philosophers.

The criticism of the philosophy of history reached a high point in the years following World War II and came from different directions. Karl Loewith (*Meaning in History*, 1949) argued that the classical philosophy of history was a secularized version of the Christian story of salvation, that is, religion in disguise. Karl Popper (*The Poverty of Historicism*, 1957) denounced it as pseudoscience. Both studies linked it to the development of twentieth-century totalitarianism. Positivists and analytic philosophers rejected it as an incoherent and unrealizable philosophical project.

Something resembling the classical philosophy of history stayed alive, in milder form, in European and North American Marxism. With the discovery and publication of Marx's early writings in the early 1930s and after, a fuller picture emerged of Marx the thinker, different from the Marx of Soviet propaganda. In particular, the full sense of Marx's indebtedness to Hegel, and his connection to the young, "left" Hegelians became clearer, something that had already been argued by Georg Lukacs in his *History and Class Consciousness* (1923). Marx also influenced the work of many historians, especially in Britain and France. Thus in Western eyes Marx took his place belatedly as a "respectable" philosopher in the Hegelian and post-Hegelian tradition, a development Marx himself would probably not have welcomed. This in turn led to a new assessment of Hegel himself in light of his influence on Marx.

Thus a tendency developed in the 1930s and after to read Hegel through the eyes of Marx and vice versa. This happened in France under the influence of Alexandre Kojeve and Jean Hyppolite, and in Germany through the "Frankfurt School" of Herbert Marcuse, Max Horkheimer, and Theodore Adorno. In this tradition Hegel and Marx were read not so much as making metaphysical or quasi-scientific claims about the direction or outcome of history as offering blueprints for political action and social analysis. Like Kant, they were outlining the conditions under which history *could* make sense, rather than asserting that it does.

Western Marxism remained strong in Europe and later in America through the Cold War period, but by the 1980s French philosophers began to turn away. The "grand narratives" of both Marxism and the capitalist idea of progress were seen by such thinkers as Jean-Francois Lyotard and Michel Foucault as belonging to a period of "modernity" that was coming to an end and giving way to a "postmodern" age. These and other philosophers, who came to be identified with the "postmoderns" label, thought of themselves as continuing the attack on the substantive philosophy of history that had begun a century before, but broadening it to include the Enlightenment idea of human progress, linked to science and technology, still held by many in the West. Defenders of the Enlightenment project, such as Jürgen Habermas, feared that this wholesale rejection of the Enlightenment was a new kind of antirationalism and a rejection of important human values. The postmoderns tend to see in any overarching or "totalizing" set of values the specter of oppression.

These debates have generally not been interpreted as continuations of the classical philosophy of history, but both sides can be seen as thinking about history and its direction in broad terms. And both sides share the ultimate value of emancipation as the key to progress in history. Though the explicit pursuit of questions in the style of the classical philosophy of history is rare, there have been recent examples. The collapse of the Soviet Union, and the trend away from dictatorships and toward democracies in Latin America and elsewhere in the 1990s, inspired Francis Fukuyama (*The End of History and the Last Man*, 1992), to revive Hegel's idea of the End of History. The march toward freedom announced by Hegel, he argued, long discredited by the atrocities of the twentieth century, could now be seen to be back on track. Fukuyama's thesis did not attract many adherents; was soon thought, like Hegel's, to be refuted by events; and was treated by many as an artifact of its time. The same, of course, could be said of the grandiose claims of Hegel and Marx—or indeed of any other philosopher.

The persistence and recurrence of philosophical reflections on the course of history as a whole, as in the case of the debates about modernity and of Fukuyama's book, indicate that the substantive philosophy of history may not have completely disappeared. Perhaps the need to make sense of history, and the continued existence of cultural worldviews about history, such as the idea of progress, will always push philosophers to look at history

as a whole in search of its meaning and purpose—or to deny that it has any.

2. "CRITICAL" PHILOSOPHY OF HISTORY: PHILOSOPHICAL REFLECTION ON HISTORICAL KNOWLEDGE

Serious discussion of questions about historical knowledge began in the nineteenth century, when the substantive philosophy of history had passed its peak in Hegel and history had established itself as a serious discipline in the academy. Prior to the late Enlightenment period, history was generally conceived as a literary genre more valued for the moral and practical lessons it could derive from past events than for its accuracy in portraying them. In some ways the substantive philosophy of history, looking for purpose and meaning in the whole of history, was simply a more sweeping and more pretentious version of ordinary historical discourse. By the middle of the nineteenth century, important new historical studies of antiquity and the middle ages had appeared. Beginning in Germany, history had acquired the dignity and trappings of a *Wissenschaft*, complete with critical methods for evaluating sources and justifying its assertions. The great historian Leopold von Ranke, one of the leading figures of the "historical school" in Germany, was explicitly repudiating the idea of history as edifying moral discourse when he famously claimed that the purpose of his historical work was simply to show the past "as it really was" (*zeigen, wie es eigentlich gewesen*).

For philosophers from Descartes through Kant, mathematics and mathematical natural science had served as the paradigm case of knowledge of the real world. How did the newly flourishing knowledge of the historical past fit in? Some philosophers, such as John Stuart Mill and those in the "positivist" tradition inaugurated by Auguste Comte, argued for the unity of all knowledge and tried to assimilate history to science. Just as physics formulated the laws of nature, and explained events by their means, the science of society would seek out social laws; history was just a case of applying these laws to the past.

Led by the neo-Kantians (e.g., Wilhelm Windelband, Heinrich Rickert,) and by Wilhelm Dilthey, German philosophers questioned this understanding of historical knowledge, focusing on the fact that its object is not natural occurrences but human actions. With history in mind, they began to work out the idea of *Geisteswissenschaften* or sciences of the human spirit, in contrast to the sciences of nature. Not only is the object of history different from that of the natural sciences, they maintained, its aim is also different: it is concerned with individual events and courses of events for their own sake, not in order to derive general laws from them (it is "idiographic" rather than "nomothetic"). Moreover, because human actions are at the center of historical concern, to give an account is often to understand the subjective thoughts, feelings, and intentions of the persons involved rather than to relate external events to their external causes ("understanding" rather than "explanation"). For some philosophers, this made it inevitable that the historian's value judgments would enter into the account of events and actions, and that the "objectivity" so prized in natural science was neither attainable nor desirable.

This opposition between "positivists" and what we might call the "humanists" on the status of historical knowledge, begun in the nineteenth century, continued to shape the epistemology of history well into the twentieth century. Those positivists who accepted the humanists' description of historical knowledge could not consider history to be a genuine science. Those humanists who wanted to defend history as offering genuine knowledge of the past had to contend that the natural sciences did not offer the only model for what qualifies as knowledge. Among the latter, two notable attempts to characterize historical knowledge are those of Benedetto Croce and R. G. Collingwood (1999). Both argued that historical understanding of the past requires moving from action as an external event (e.g., Caesar leading his army across the Rubicon) to the reconstruction of the "inside" of the event: the experience or thought of the agent that motivated it.

Some of the issues that concerned philosophers of history were reflected in the work of historians as well. With the rise of the social sciences in the twentieth century (sociology, anthropology, political science), many historians coveted a place among them, arguing that history had to be "objective" and "value-free." If that meant ignoring the subjective motivations of historical agents, so be it. They borrowed quantitative methods from the social sciences and applied them to the study of the past. Leading the way were the historians of the Annales school in France, beginning in the 1930s. Its best-known theoretician, Fernand Braudel, argued that history should shift its focus from the "surface" ripples of political history to the deeper-lying and slower-moving currents of social, economic, and geographical change. The move toward social history had a large impact on the discipline, and it was partly motivated by the desire to make history more "objective"—but only partly. Braudel's view

reflected something closer to the substantive than to the critical philosophy of history, namely a belief about what the historical process really is.

Among philosophers, the positivist conception of historical knowledge was revived in the 1940s, under the aegis of the unity-of-science movement in analytical philosophy, by Carl G. Hempel. The focus was on the idea of *historical explanation*: Does history merely describe events, or does it try to explain them? And if it explains them, how does its mode of explanation compare with explanation in natural science?

Hempel argued that history does attempt to explain events, not merely describe them, and it does so according to a pattern no different from that found in the natural sciences: it brings events under general laws that allow us to show how they follow from their antecedents. Given such a law, the event to be explained should be logically deducible from its antecedents. Critics such as William Dray (1989) objected to Hempel's "covering law theory" (as Dray called it) on several grounds. Dray did not dispute the claim that history often tries to explain events, but, following Collingwood, he argued that a satisfying historical explanation often consists of reconstructing the reasons behind an action rather than finding its external causes. Further, it is hard to find general laws, of the kind that would be comparable to physical laws, being articulated in historical work.

Hempel conceded that historical accounts bear little surface resemblance to scientific explanations, that they seem to offer merely probabilistic rather than deductive explanations, and that their accounts are often just "sketches" of more complete explanations. But in doing so, he revealed the strongly prescriptive character of his account—a character it shared with much of the epistemology of his day. The implication was that if history could not live up to the standard of natural science, it could not qualify as genuine knowledge. Dray's larger objection to Hempel's approach was that philosophers should pay attention to what historians actually do, and to the wide variety of conceptual strategies in their work, rather than prescribing standards derived from abstract logical analysis or reducing their work to an imitation of a different, and equally idealized, endeavor. In this he was a harbinger of a trend in analytic epistemology that eventually extended even to the philosophy of natural science itself.

Nevertheless, the discussion of history among analytic philosophers in the 1950s was dominated by the theme of causal explanation, and above all by the contrast with the natural sciences. Hempel's proposal set the tone. Even those such as Dray, who argued for the autonomy of historical knowledge, shared this preoccupation. Thus the confrontation of "positivists" with "humanists" continued. At the same time, the discussion extended to other, related topics.

One distinction that was much discussed in this literature was that between history and chronicle. It was agreed that history had to do more than just list facts. As Morton White put it schematically in his *Foundations of Historical Knowledge* (1965):

> The chronicler is likely to tell us: "The king of England died, and then the queen of England died, and then the prince of England died, and then the princess of England died"… But a corresponding history may read: "The king of England died, so the queen of England grieved. Her grief led to her death. Her death led the prince to worry, and he worried to the point of suicide. His death made the princess lonely, and she died of that loneliness.…" (1965, p. 223)

A chronicle simply lists a series of events in the order in which they happened, but according to White, "a history contains causal statements" (p. 223). But what kind of causation do emotions have? Even they seem to have the teleological character of reasons. The distinction between chronicle and history raises further problems. The chronicle involves more than a simple statement of facts. The historian has *selected*, from all the possible facts there are, some that are relevant to the story that is to be told. The problem of selection relates to the problem of historical objectivity, because even if facts are established by careful critical methods, the decision of which ones to look for, and which to include in a historical account, may derive from the interests and values of the historian.

Another problem, related to explanation, had to do with the nature of the *explanandum* in historical accounts. What do historians explain? The distinction between explanation and understanding, or between explanation by causes vs. explanation by reasons, may be relevant to the discussion of individual persons and their actions. But in history the focus is more often on large-scale entities such as nations, peoples, and classes, and on events such as wars, revolutions, and economic crises. We often impute actions or mental states to states or groups, as when we say that "Congress decided," "Japan was offended," "organized labor was fed up," and the like. To what extent are these expressions just shorthand for references to the actions or feelings of individuals? If these large-scale entities do not themselves act and feel, are they subject to causal explanation, and if so what kind? Are

there social laws governing the behavior of such entities and the occurrence of such events, which can be discovered independently of reference to the individuals that make them up, as methodological holists believe? Or must everything be traced, at least implicitly, to individuals? These are questions, of course, that arise in the social sciences generally and are not peculiar to history.

Positivism, reductionism, and the unity-of-science movement gradually lost their hold on analytic philosophy, largely under the influence of the later Wittgenstein, and philosophy of science was itself transformed. Arthur Danto, whose *Analytical Philosophy of History* appeared in 1965, later wrote an essay called "The Decline and Fall of the Analytical Philosophy of History" (1995). Danto claimed that Hempel's project was one of the many casualties of Thomas Kuhn's *The Structure of Scientific Revolutions* (1962). In an ironic reversal of fortune, worthy of a good novel, the attempt to absorb the philosophy of history into the philosophy of science was upended when science was reconceived as an essentially historical phenomenon and the philosophy of science became a branch of the philosophy of history—or at least of history proper. Epistemology was now devoted to describing what scientists actually did, rather than producing idealized and prescriptive accounts, and this meant following their work historically.

Danto was too hard on himself, however, when he described himself retrospectively as pursuing a Hempelian program. His *Analytical Philosophy of History* was actually itself part of a revolution going on the philosophy of history in the 1960s. The model for the philosophical understanding of history was shifting from *science* to *literature*. The old idea of history as a literary genre was revived. While Danto continued to think of history as explaining events causally, his account of how it does this drew heavily on the concept of *storytelling* or *narrative*. The concept of narrative had been used before in analytic philosophy, to distinguish between chronicle and history, but Danto's sophisticated treatment of it was explicitly modeled on literary narratives such as novels. At the heart of Danto's account is the idea that in a historical narrative, as in a good story, events are selected and described retrospectively with reference to later events. Thus the temporal character of events, and the temporal position of the narrator in relation to them, determines the structure of a historical account.

But Danto was not alone in looking to the literary model. W. B. Gallie had published a book called *Philosophy and Historical Understanding* (1964) whose premise was that "history belongs to the genus 'story.'" With the work of Louis Mink in the early 1970s (later collected in *Historical Understanding*, 1987), the trend was well under way to look at narrative as a "cognitive instrument" and history as "mode of comprehension" (these are Mink's terms) based on narrative. Some analytic philosophers (e.g., Maurice Mandelbaum and Leon Goldstein) objected to the emphasis on narrative for favoring the literary presentation of history over the hard work of discovery, evaluation of sources and critical hypothesis that lies behind it. History, they said, is a disciplined inquiry whose goal is knowledge. Narrative is merely the way—indeed only one way—its results are "written up" for public consumption. But Mink's idea is that narrative is more than just literary presentation. It constitutes a conceptual framework for dealing with human events, utterly distinct from scientific explanation, which is entirely appropriate to history. Danto later calls narrative the "metaphysics of everyday life" (Danto 1985, p. xiv).

In literary theory, of course, the study of narrative had a long tradition and had produced a number of classic studies in the English-speaking world. The rise of French structuralist literary theory in the 1960s had also involved considerable focus on narrative, drawing on the earlier work of theorists from Eastern Europe such as Roman Jakobson and Vladimir Propp. But literary theory and the philosophy of history had little contact until the appearance of Hayden White's *Metahistory* in 1973. Drawing on the literary theories of Northrup Frye, Roland Barthes, and others, White produced a theory of narrative in general that he then applied to history by examining the work of both classical historians (Ranke, Michelet) and philosophers of history (Hegel, Marx). White (1973) argues that their work is guided by the same plot structures—romance, comedy, tragedy, and satire—that govern the production of literary texts. White's book was widely influential but also highly controversial, especially among historians, because White seemed to be portraying their work as guided by literary motives, or motifs, rather than by the project of telling the truth about the past.

By this time the study of narrative was burgeoning on all sides, with a lot of emphasis on the fact that narrative or storytelling is a cross-cultural and cross-disciplinary phenomenon sui generis, turning up not only in history and fiction, but also in films, folktales, medical case histories, psychotherapy, medieval altar paintings and tapestries, comic strips, court testimony, and so on. Some theorists proposed a new discipline, to be called "narratology," which would seek out the common features of narrative in all its manifestations. Under the broadening

influence of both Hayden White (1973) and structuralist and poststructuralist theories of literature, the works of historians were studied as examples of narrative form.

At a time when many historians, as noted earlier, were trying to escape traditional approaches by shifting the focus of history away from human actions, there was much difference of opinion on whether narrative was essential to history at all. *Annales* historians in France, and quantitative historians ("cliometricians") elsewhere, disdained traditional historical language and thought narrative dispensable. Those who followed the trend toward the history of "*mentalites*," or social attitudes and thought patterns, implicitly agreed. The point was made that histories have not always told stories. White, by contrast, argued that even such standard examples of nonnarrative history as Burkhard's *Civilization of the Renaissance in Italy* and Huizinga's *Waning of the Middle Ages*, were implicit or truncated literary narratives. Paul Ricoeur in *Time and Narrative* (1983) made a similar claim about Braudel's *The Mediterranean and the Mediterranean World*, the example *par excellence* of the *Annales* school's nonnarrative approach, arguing that large-scale "quasi-persons" turned up in "quasi-plots" in Braudel's work, a kind of narrative in disguise.

3. POSTMODERN SKEPTICISM AND ITS CRITICS

To the outside observer it might seem that with this shift to the discussion of narrative, the epistemological questions that originally motivated the "critical" philosophy of history were gradually fading from view. In the work of Danto, Mink, and Gallie, the concept of narrative had evolved, partly in reaction to the positivist program of Hempel, within the world of analytical philosophy, and it was undoubtedly part of the critical or epistemological reflection on historical knowledge. Even though these thinkers increasingly took literature as their model for understanding history, they were still interested in history's cognitive role. But when this tradition collided with structuralism in Hayden White's work, and with the larger, more literary world of narratology, the problem of knowledge seemed to lose its interest. The focus had shifted from history as knowledge to the historical text as literary artifact (as White called it). While this development is sometimes called the "linguistic turn" in the philosophy of history, it is more properly called the turn to the text. Literary analysis had apparently replaced epistemology.

This is only partly true, however, as there was more to the structuralist and poststructuralist treatment of history than just literary interest. Their analysis contained a profoundly skeptical view of history as a claim to knowledge. They were inclined to see narrative structure as an *a priori* cultural form *imposed* on the real world, an alien structure that by its very nature distorted or misrepresented the messy and chaotic character of human life and action. Their model was fiction, and they saw narrative originating in the literary imagination or the archetypical plot structures embedded in culture. As for history, which pretends to represent the past as it really was, here narrative inevitably achieves the opposite effect, according to them. At best it dresses up reality, reflecting our need for satisfying coherence, and, if we really believe it, derives from wishful thinking. Far from reflecting reality, it escapes from it. At worst, narrative in its role as the "voice of authority" seeks to put across a moral view of the world in the interests of power and manipulation. This skeptical view was increasingly expressed in the writings of Hayden White, after *Metahistory*, and to some extent in those of Mink as well.

There is some irony in this development. The turn to narrative had begun as an attempt to defend the autonomy of history against the claim that it had to be transformed into science in order to be genuine knowledge. It was another chapter in the ongoing battle of the humanists against the positivists. For the humanists, narrative, like "understanding," as opposed to "explanation," was supposed to be capable of telling us about the past as it really was—human actions and intentions—whereas scientific reduction was the alien framework imposed from outside. Now the narrativists seemed to join the positivists in believing that the literary form of traditional history stands in the way of its epistemic pretensions. As we have seen, the antinarrative historians of the *Annales* school, and many other social and economic historians, agreed with them. The only difference was that the poststructuralists, unlike the positivists and the working historians, held no brief for the epistemic pretensions of the sciences and social sciences either. All was linguistic construction, all was imposed on reality—if indeed it makes any sense to speak of a "reality" outside our constructions.

Thus epistemology had not completely disappeared from the narrative treatment of history; there was still a concern for its epistemic status. But the consensus among the most influential poststructuralist or postmodern theorists (the latter term came to prevail) was that it had none. Many of the issues associated with the critical philosophy of history—objectivity, the role of evidence, the nature of explanation—were simply not treated at all. To that extent the project of the critical philosophy of history had been transformed, if not eclipsed.

One theorist who had a lot to say about historical knowledge was Michel Foucault, whose work gradually took on enormous importance from the late 1960s on, first in France and then elsewhere. Foucault's early work was in the history of medicine and psychiatry, but it engaged fundamental social and philosophical issues such as the normal vs. the abnormal and reason vs. insanity. His middle works (*The Order of Things* [1970], *The Archaeology of Knowledge* [1972]) dealt more broadly with knowledge in the human sciences. In keeping with the "linguistic turn," his focus was on forms of discourse, and his treatment took the form of contrasting widely divergent historical examples of scientific theory. His thoughts on history came through primarily in his defense of his own approach against more traditional treatments. He contrasted his own method, which he called "archaeological," with what he called the "history of ideas." He opposed the latter not only because he wanted to look beyond the surface level of ideas to the "discursive practices" that lay behind them; but also because the traditional historical approach tended to view the science of the past as a deficient form of knowledge striving toward the present. Rather than being a teleological continuum, according to Foucault, history manifests discrete breaks between radically different periods, which cannot properly be compared at all as if their sciences were all trying to do the same thing. Foucault was clearly criticizing traditional historians for imposing a teleological structure on the past; but he was doing so by arguing for an alternative conception of historical reality. Thus his work perhaps belongs as much to the substantive as to the critical philosophy of history. And while it differs in some ways from the more literary approach to history of other contemporary trends, it is like them in treating historical knowledge as conceptual construction. The question of its truth does not arise.

This did not sit well with many historians, who were still toiling away, reading documents, sifting and evaluating evidence, attempting to tell the truth, and to distinguish it from falsity, about the past. Historians on the whole had never had a great deal of patience with the philosophy of history; now many were further alienated, if not openly hostile. It is true that White, Barthes, and others had opened the hostilities by portraying professional history, in effect, as a powerful establishment managing the past for political purposes. Now many historians argued that, on the contrary, by questioning the idea of historical truth, the postmoderns were fostering an "anything goes" attitude that opened the doors to Stalinist-style rewriting of history, Holocaust denial, and other falsifications. Postmodern theory provided no way of distinguishing between history and fiction, in the view of its critics. Some historians, it is true, were intrigued by skeptical doubts about history's capacity to know the past. Robert Novick noted (*That Noble Dream*, 1988) that even the respected American historian Charles Beard, in the 1930s, had called historical objectivity a "noble dream" that could never be fulfilled; and Novick went on to argue, with the help of postmodern theories, for an even stronger skepticism about the past. As could be expected, his 1988 book stirred much controversy among professional historians.

But historians were not the only ones who were unhappy with the postmodern turn. Philosophers in the analytic tradition (McCullagh, Bunzl) were prompted by the controversy over Novick's book to mount arguments against the skeptical relativism it represented. While generally admitting the role of culture and language in shaping our approach to the past, these authors adduce some of the standard arguments about the self-refuting character of skepticism and defend the place of evidence and critical judgment in distinguishing better from worse historical accounts. Paul Ricoeur (1984–1988), a continental philosopher who also drew heavily on the analytical philosophy of history, attempted to soften the excesses of postmodernism by reconnecting narrative texts with their roots in human experience. Ricoeur believed that narrative, in both fictional and historical form, "humanizes" the experience of time, bringing order and measure to human existence. He argued that history and fiction draw on each other and often intersect in important ways. But he did not agree with the tendency of his French contemporaries to reduce history to fiction, or to blur the distinction between them. In writing about history, he devoted careful attention to the restraining and guiding role of document and evidence in historical discourse. He also believed that narrative texts build on structures that are already present in ordinary experience, transforming them, and then affecting and enriching ordinary experience in their turn.

Other philosophers of history countered the views of White and the postmoderns by arguing against the idea that narrative is an alien framework imposed on a nonnarrative reality. What reality is meant? Human reality, which history is about, is the temporal flow of experiences and actions that engage persons in their social context. While it may not always have the crafted contours of a novel's plot, neither is it a chaotic absence of order or a meaningless one thing after another. According to this argument, human experience, and especially human action, are ordered in a manner that foreshadows the

structures of narrative itself. Events are experienced as temporal configurations with beginnings, middles, and ends; actions project an end and organize the means for achieving it. The agent grasps a sequence of events together in a temporal order much as a narrator organizes the events of a story; it is as if the agent is constructing and telling himself a story and then acting it out. On this view the narrative we find in historical writings—and in fictional writings too—is not a merely literary device at odds with the human world, it is something more like an extension of human existence by other means.

According to this "continuity theory" (as some have called it), narrative structures constitute "the metaphysics of everyday life," as Danto called it, and offer the key to understanding not only experience and action, but also the self who acts (1985, p. xiv). The self can be seen as constructing itself by implicitly or explicitly telling, and of course also revising, its life story. This theory can be extended from individual to social life, where it becomes relevant to history. Communities, large and small, may be said to constitute themselves in the stories they tell themselves about themselves. Here historical consciousness and historical writing have their place. Written history can be seen as the collective memory that permits a society to hold itself together and plan its future.

Critics of the continuity theory have argued that it does not succeed in answering the skepticism of the postmoderns, which was seemingly its intention. It counters the theory that historical narrative is in principle incapable of portraying the past by arguing against the radical discontinuity between narrative and the real world. But even if it succeeds in demonstrating the protonarrative character of everyday action and experience, and in extending this to the social level, it does not account for the differences between these protonarrative structures and fully formed narratives we find in novels and histories. As regards historical knowledge, this theory, according to its critics, fails to provide a positive account of how narrative can succeed in arriving at historical truth and distinguishing it from falsehood.

4. HISTORICITY, HISTORICISM AND THE HISTORICIZATION OF PHILOSOPHY

These criticisms inadvertently reveal something about the discussion of narrative and history, especially when it draws on continental philosophy for its inspiration, that once again raises questions about how to classify it as philosophy of history. We already found that the focus on historical narratives as literary texts, under the influence of White and the structuralists, moved away from traditional epistemological questions without completely abandoning them. Historical knowledge took a back seat to the literary properties of historical writing. Some of the attempts we have been discussing, designed to counter the influence of poststructuralism on the philosophy of history, similarly defy the standard classification. This is because they draw heavily on the phenomenological and hermeneutical tradition going back to Husserl and Heidegger. These philosophers reflect on history in a way that is indeed related to traditional epistemological and even metaphysical concerns, but not in the way associated with the standard distinction between the substantive and the critical. In this tradition, the key concept is "historicity."

"*Geschichtlichkeit*," sometimes translated as "historicality," is a term used by Husserl and Heidegger in the 1920s and 1930s in their phenomenological descriptions of consciousness and human existence. The importance of this notion attests to the influence on both philosophers of Dilthey, who had died in 1911 but whose posthumously published work was still studied intensely. We have encountered Dilthey as the philosopher of the *Geisteswissenschaften*, whose project of working out a "critique of historical reason" made him an important contributor to the epistemological debates about history. But he also believed that historical knowledge is rooted in certain features of human existence. "We are historical beings before we become observers of history," he wrote, "and only because we are the former do we become the latter." (Dilthey 2002, p. 297)

Husserl and Heidegger, following Dilthey's lead, expand in slightly different ways on what it means to be a "historical being." The phenomenological concept of "world" is central for both: The human world is not merely a container for human beings but a complex of meanings. Past and future are part of that world, and both philosophers devote extensive analysis to temporality. Human experience is not confined to the present but consists of a temporal grasp, holding on to the past and anticipating or projecting its future. The self is not simply a substance that persists through time, but a self-constituting unity of temporal interrelations. These are all essential, ontological features of human existence: it is not as if the human being existed first and then just happened to come up against the world, the past, the future. An existence without these would not be a human existence at all.

The same can be said of the social dimension of existence—Husserl speaks of intersubjectivity and Heidegger

of being-with-others. Taking this dimension into account, we can see that past and future take on broader meanings. The social past—history—has meaning for us and figures in our lives prior to and independently of explicit historical representation and disciplined inquiry. Husserl asserts in his late works that all human activity, even that of a science such as mathematics, has to be understood historically. According to Heidegger, we appropriate our history in an act of self-interpretation, and it becomes part of the future we project for ourselves. Our history is part of our self-understanding and in that sense part of our being. Like the world and others, history is an essential feature of our existence, not something added on or something we could be without.

Though the term "narrative" is not used in these early treatments of the concept of historicity, the idea is implicit in it. Dilthey did compare self-understanding to the composition of an implicit autobiography. The German term *Geschichte*, like the French *histoire*, can mean both "history" and "story," and both senses of the term are often implied." Husserl writes that "the ego constitutes itself for itself, so to speak, in the unity of a *Geschichte*," suggesting that the temporal synthesis of past, present and future, in which the self takes shape, is like telling the story of one's life (Husserl 1999, p. 75). It is easy to see in these concepts the prefiguration of the narrative conception of human time that later theorists apply to history in the larger, social sense.

How does the discussion of historicity fit into the philosophy of history? Clearly it qualifies as philosophical reflection on history, but it does not correspond to the standard categories with which we began. It does have some bearing on the understanding of history as a discipline, in the sense that it seeks the roots of historical knowledge in human existence. It addresses the question of why we seek to know about the past at all. It suggests that the past is more than just an object of curiosity for us, because it corresponds to a dimension of our being. Knowing about the past is knowing where we have come from and thus who we are. History as a disciplined, critical inquiry, as it has developed in the academy, is thus just an extension and intensification of the project of self-knowledge. But while this addresses the nature of historical inquiry, it is not raising the traditional epistemological questions about whether genuine knowledge of the past is possible, how or whether objectivity can be achieved, etc. It is interested in historical inquiry as a human activity, and seeks to understand its significance within human existence as a whole.

If these questions are not epistemological, it may be argued that they are metaphysical. Understanding human nature, after all, has always been a central metaphysical endeavor. This does not mean, however, that these questions are part of the substantive philosophy of history. The latter has traditionally set out to understand the whole process of human history, and this is different from the focus on what is essential to individual human existence. We find few pronouncements in the phenomenological, hermeneutical or narrativist literature about the meaning and purpose of history as a whole.

The concept of historicity became an issue in the French structuralist attack on the phenomenological tradition in the 1960s. The anthropologist Claude Levi-Strauss argued that many of the non-Western societies he studied were "peoples without history" in the sense that they devalue temporal change. The primary purpose of social organization in these societies is to prevent change or contain it as much as possible within an interpretive framework in which its significance can be denied. Their sense of themselves as individuals and as societies is not derived from a consciousness of the difference between past, present and future. Unlike Western societies, they have no interest in their past origins, nor do they ponder their future destiny; in this sense they are not characterized by historicity at all. Levi-Strauss famously attacked Jean-Paul Sartre for making historicity essential to humanity and by implication excluding "peoples without history" from the human race. Either they are somehow less than human, or they are relegated as "primitive peoples" to some remote prehistory, even though they still exist in the present. Levi-Strauss's attack foreshadows the postmodern view that the emphasis on history is a "Eurocentric," and thus provincial and limited, conception.

A related trend in twentieth-century philosophy might be seen as an extension of the notion of historicity, though it does not necessarily follow from it. If human existence is through-and-through historical, then all human endeavor is dependent on and limited to its historical position, including the search for truth. Truths thought to be timeless turn out to be nothing more than reflections of their historical age. Historical relativism of this sort is sometimes called "historicism" (though that term has also been used in a different sense—notably by Karl Popper, who used it to mean "historical determinism"). We have already encountered skeptical relativism about historical knowledge itself, and we have noted that some philosophers are skeptics about scientific knowledge as well. But to attribute the relativity of all knowledge to history in particular is a special form

of skepticism. Like all skepticism, this form has self-referential problems, because the alleged relativity would extend to the relativist thesis itself.

But some philosophers have not flinched at this prospect, propounding the radical historicization even of philosophy. Thus the later Heidegger, and more recently Richard Rorty, view philosophy itself as a large-scale episode in Western history that is nearing or has reached its end. Perhaps this is the ultimate inversion of Hegel's grand design for the philosophy of history: He thought history had come to an end by being fully comprehended in thought. Philosophy ultimately triumphs over history. For Heidegger and Rorty, it is philosophy that has come to an end, and the triumph belongs to history.

See also Adorno, Theodor Wiesengrund; Anaxagoras of Clazomenae; Aristotle; Augustine, St.; Barthes, Roland; Bossuet, Jacques Bénigne; Collingwood, Robin George; Comte, Auguste; Continental Philosophy; Croce, Benedetto; Danto, Arthur; Determinism in History; Dilthey, Wilhelm; Engels, Friedrich; Enlightenment; Foucault, Michel; Geisteswissenschaften; Habermas, Jürgen; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Hempel, Carl Gustav; Herder, Johann Gottfried; Historicism; Horkheimer, Max; Husserl, Edmund; Hyppolite, Jean; Kant, Immanuel; Kuhn, Thomas; Lukács, Georg; Lyotard, Jean François; Marx, Karl; Marxist Philosophy; Mill, John Stuart; Plato; Popper, Karl Raimund; Positivism; Progress, The Idea of; Rickert, Heinrich; Ricoeur, Paul; Romanticism; Rorty, Richard; Sartre, Jean-Paul; Spengler, Oswald; Structuralism and Post-structuralism; Thucydides; Toynbee, Arnold Joseph; Vico, Giambattista; Voltaire, François-Marie Arouet de; Windelband, Wilhelm.

Bibliography

Augustine, Saint, Bishop of Hippo. *The City of God against the Pagans*, edited and translated by R. W. Dyson. New York: Cambridge University Press, 1998.

Carr, David. *Time, Narrative, and History*. Bloomington: Indiana University Press, 1986.

Collingwood, R. G. *The Principle of History: And Other Writings in Philosophy of History*, edited by William H. Dray and W. J. van der Dussen. New York: Oxford University Press, 1999.

Danto, Arthur. *Narration and Knowledge: Including the Integral Text of Analytical Philosophy of History*. New York: Columbia University Press, 1985.

Dilthey, Wilhelm. *The Formation of the Historical World in the Human Sciences*. Vol. 3: *Selected Works*, edited by Rudolf A. Makkreel and Frithjof Rodi. Princeton and Oxford: Princeton University Press, 2002.

Dray, William H. *On History and Philosophers of History*. Leiden, Netherlands: Brill, 1989.

Fukuyama, Francis. *The End of History and the Last Man*. New York: Harper Collins, 1992.

Hegel, Georg Wilhelm Friedrich. *Introduction to the Philosophy of History*. Translated by Leo Rauch. Indianapolis and Cambridge: Hackett, 1988.

Hegel, Georg Wilhelm Friedrich. *The Philosophy of History*. Translated by J. Sibree. New York: Dover, 1956.

Husserl, Edmund. *Cartesian Meditations*. Translated by Dorion Cairns. Dordrecht/Boston/London: Kluwer Academic Publishers, 1999.

Kant, Immanuel. *On History*, edited and translated by Lewis White Beck. New York: Macmillan, 1963.

Lemon, M. C. *Philosophy of History: A Guide for Students*. London: Routledge, 2003.

Löwith, Karl. *Meaning in History*. Chicago: University of Chicago Press, 1949.

Lyotard, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Translated by Geoff Bennington and Brian Massumi. Minneapolis: University of Minnesota Press, 1984.

Marx, Karl, and Friedrich Engels. *The Communist Manifesto*. Edited by David McLellan. Oxford: Oxford University Press, 1998.

Mink, Louis. *Historical Understanding*, edited by Brian Fay, Eugene O. Golob, and Richard T. Vann. Ithaca, NY: Cornell University Press, 1987.

Novick, Peter. *That Noble Dream: The "Objectivity Question" and the American Historical Profession*. Cambridge, U.K.: Cambridge University Press, 1988.

Popper, Karl. *The Poverty of Historicism*. London: Routledge, 1957.

Ricoeur, Paul. *Time and Narrative*. 3 vols. Translated by Kathleen McLaughlin and David Pellauer. Chicago: University of Chicago Press, 1984–1988.

White, Hayden. *Metahistory*. Baltimore, MD: Johns Hopkins University Press, 1973.

White, Morton. *Foundations of Historical Knowledge*. New York: Harper and Row, 1965.

ANTHOLOGIES

Ankersmit, Frank, and Hans Kellner, eds. *A New Philosophy of History*. Chicago: University of Chicago Press, 1995.

Burns, Robert, and Hugh Rayment-Packard, eds. *Philosophies of History: From Enlightenment to Postmodernity*. Oxford, U.K.: Blackwell Publishers, 2000.

Dray, William H., ed. *Philosophical Analysis and History*. New York: Harper and Row, 1966.

Fay, Brian, Philip Pomper, and Richard T. Vann, eds. *History and Theory: Contemporary Readings*. Oxford, U.K.: Blackwell Publishers, 1998.

Roberts, Geoffrey, ed. *The History and Narrative Reader*. London: Routledge, 2001.

David Carr (2005)

PHILOSOPHY OF LANGUAGE

What, if anything, can philosophy teach us about language? It is a feature of English that its adjectives come before its nouns, as in *green table*. This syntactic fact distinguishes English from French. In English there is a difference in sound between words that begin with a b and ones that begin with a p. This phonological fact distinguishes English from other languages. Some Arabic languages, for example, have trill sounds. This phonetic feature distinguishes these Arabic languages from English. Are any of these linguistic features philosophically interesting?

It is doubtful whether any philosopher seriously believes that, qua philosopher, they have anything interesting to say about the syntactic, phonetic, and phonological features of languages in general or of English in particular. Why, then, should it be any different for all of the other features of language? For example, that in English a relative pronoun proceeds the noun phrase it modifies or that English declarative sentences are of the subject-verb-object variety, are interesting facts about English syntax, but why should any of this be of philosophical interest?

Many theorists claim that philosophers of language are interested in answering questions of the sort: What need someone know in order to understand his or her language? Do they need to know the sorts of facts just mentioned? In some sense of *know*, they must. Someone who speaks English, normally, can recognize another as a non-English speaker, as a nonnative English speaker, or not a perfectly fluent English speaker simply by virtue of the fact that this speaker employs syntactic structures or phonemes that are not a part of English, or fails to recognize differences between distinct phonemes of English. For example, if someone failed to recognize a difference between an articulation of the words *bit* and *bet*, this would constitute partial evidence that the individual in question does not (fully) grasp English. But why is this philosophical? It is not! Still, philosophy does matter to language. Why anyone should think so is a complicated matter; one that an answer to will be sketched in the sections that follow.

COMMUNICATIVE ABILITIES

It is uncontroversial that linguistic expressions carry meaning. Right now, you are looking at ink marks on a piece of paper. These marks are in English, they have meaning, and should you know these meanings, you can figure out what they say. We spend a lot of our lives exercising our communicative abilities; abilities to produce utterances (spoken, written, felt, etc.) that others can interpret; and, abilities to interpret utterances that others have produced. These abilities in assigning meanings to expressions—simple and complex—are required in order to ask for help, read traffic signs, interest others, surf the net, read newspapers, write e-mails, watch movies, comfort others, listen to lectures, order food, read a bus schedule, buy wine, quarrel, and make jokes.

One of the central topics in philosophy of language today is to provide an explicit and systematic account of whatever knowledge we have of the meanings of the expressions of our language that enables us to communicate with it. Surrounding these projects are a number of subtle philosophical issues.

WHAT IS MEANING?

What is the meaning of an expression? Traditional scholarly books and articles all weigh in with one analysis or another about the nature of meaning. Some posit that the meaning of an expression is what it applies to (*apple* means the set of apples), the idea that we associate with it (*God* means, say, the idea of a benevolent omnipotent omniscient being), or the characteristic behavior that its uses evince (*Fire*! means run for safety), and so on.

Criticisms run that this or that analysis cannot be right, because if meaning were this, then two expressions that differ in meaning would turn out to be synonymous, or that a meaningful expression would turn out to be, on the proposed analyses, meaningless. For example, a critic of the view that the meaning of an expression is what it applies to might argue that even though the two sentences "Cicero was Roman" and "Tully was Roman" are not synonymous as the referents of *Cicero* and *Tully* are the same. A critic of the view that the meaning of an expression is the idea(s) we associated with it might argue that even though someone can associate the idea of warm weather with the word *grass*, the idea of warm weather is still not part of its meaning. Anyone who denies this should visit Ireland in January.

Though neither argument is definitive (after all, paraphrasing Ludwig Wittgenstein, theories do not get refuted; they just become no longer interesting to defend), they still illustrate how theories of meaning can be, and often are, evaluated. In traditional criticisms, intuitions about what we believe expressions to mean are dominate. The question of what the relationship is between theories (the sole aim of which is to provide an analysis of an important concept) and theories (the aim

of which is to explain various phenomena) is left open by this to and fro (for more on the analysis of the concept of meaning, see William P. Alston's *Philosophy of Language*).

A major shift in the philosophical study of meaning took place about fifty years ago with the abandonment of efforts to analyze the concept of meaning (Quine 1953). But, if it is not an analysis of the concept of meaning that philosophers are after, what, then, warrants evaluations of various claims about what meaning is?

Whatever meaning is, it is relatively uncontroversial what knowledge of it enables us to do: It enables us to understand language. Because we know what the expressions of our language mean, we understand English. In rejecting an account we are saying in effect that this cannot be what we know that enables to understand English, because if it were we need not understand English. Thus, if you were taught the referent of every English word, you would not understand an English interlocutor. On this account, being asked, "Was Cicero the same man as Tully?," should produce bewilderment. On the referential theory, it is analogous to being asked whether bachelors are unmarried men. But if it is not knowledge of the referent of an expression that enables one to understand it, what does enable one to understand it?

The picture that understanding a word is learning to associate an idea with it goes back at least to the early empiricist Thomas Hobbes. It is a bad theory, for suppose you were told, "Though grass covers Ireland in January, it is not warm there then." Were your understanding of the word *grass* to include the idea of warmth, you should find this comment linguistically confused, much like being told "Though John is a bachelor, he has a wife!" But if understanding consists neither in knowing the referents of your words, nor the ideas you associate with them, what then might you know that would enable you to understand English?

The picture that dominated theories of meaning throughout most of the last century is (various versions of) linguistic behaviorism (Skinner 1957). Linguistic competence with an expression is knowing how to behave appropriately when confronted with its uses. For example, suppose you are told "Go get a coke!" In virtue of understanding English, what should you do? Should you automatically get a coke? Presumably not, for that would render linguistically competent English speakers all very active. Perhaps you need only know what you are supposed to do. But what are you supposed to do when someone asks you for a coke? Good manners might require that you should do something when asked, but understanding English requires nothing of you. These various critical points are intended to establish that no particular behavior is associated with language understanding, and so they scream out for clarification from anyone who wants to be a behaviorist about linguistic competence, clarification that was never forthcoming (Chomsky 1959).

MEANING IS RELATIONAL, EXTRINSIC, VAGUE, AND CONVENTIONAL

Beginning with a banality such as understanding a language requires knowing the meanings of its expressions, as philosophers well know, is a necessary precaution against a rampant background of skepticism in some philosophical quarters about the notion of meaning. Some of this skepticism generates from the consideration that whatever is alleged to carry meaning does not do so inherently. For instance, there is nothing about English words that requires "Snow is white" to mean that snow is white. In another language, they might mean grass is green, and so it follows that whatever words mean depends partly on the language from which these words originate. But this sort of relativity should not compromise the reality of what words mean. After all, no one is inherently a father. The relational property of fatherhood depends on a relationship to someone else—a child. Likewise, whether or not a string of words means that snow is white depends on this string's relationship to a specific language.

This issue concerning the meaning of words should not be confused with reservations about the reality or truth of conventions. Being married, like fatherhood, is a relational property. But unlike fatherhood, marriage is not grounded in biology. It is, so to speak, a matter of convention or social arrangement who is married to whom. But, extant conventions might easily have been different. Everyone who is currently married might just as easily not have been without suffering any substantial change to their being—rather, only a change in convention. It is a mistake, however, to infer from this possibility that there really is no such thing as marriage. Likewise, if it is a matter of convention that *dog* means dog and not cat, then it does not follow that there should be a dispute over what *dog* means.

The reality of meaning is equally left uncompromised by considerations about vagueness or borderline cases. Two words translate or paraphrase each other just in case they share the same meaning. In many instances, we are simply unsure whether two words translate or paraphrase each other; and there is no higher source to which we can appeal to settle our doubts. In short, that

meaning is relational, extrinsic, vague, or conventional does not compromise its reality.

LANGUAGE AND USE OF LANGUAGE: SEMANTICS AND PRAGMATICS

Of course, linguistic meaning is not our only employment of the concept of meaning. We sometimes speak of another's action as meaningful, as when identifying purpose as our aim. In seeking the meaning for which Bill burned down his house, however, it need not be assumed that Bill's act of burning down his house is meaningful in the same way as the English sentence "Bill burned down his house" is. For one, it is not conventional meaning we seek in another's act, but rather the underlying intentions. For what reason did Bill carry out his sorry deed?

Similarly, people use words with intent. John might assert "Snow is white" because he wants to alert his listeners to the fact that English is his native tongue. No one would conclude on this basis that the words "Snow is white" mean that English is John's language. We can see clearly that with these speech acts, the notion of meaning enters twice. First, in choosing a vehicle to express our message, words whose conventional meaning best conveys that message are employed. And, secondly, in interpreting a linguistic act, an attributed meaning can and often does exceed this conventional meaning.

An audience can exploit context and individual histories in order to discern an agent's purpose or message. Why did he tell me, "I love you," when he knows that I am fully aware of it? Does he mean to reassure me? Or, does he dread losing me, and so, means by his words for me to feel guilt about our imminent separation? Such exegetical issues concern us all whenever we try to size up what others mean by their particular use of words. With conventional linguistic meaning, speakers rely on a prior comprehension in order to convey successfully a message; with these other sorts of meaning, speakers hope—wittingly or not—to exploit presumed shared beliefs and expectations in discerning nonconventional meaningful aspects of linguistic acts.

In summary: When theorizing about meaning, it is crucial to distinguish between language and the use of language. Languages, such as English, exist independently (in a sense that requires clarification) of what anyone happens to do with them. If these sentences together in this order had never been assembled, it would have made no difference to the existence of English. English words and sentences would have meant whatever they do. Speakers simply exploit the meanings of these words in their writings, and a reader exploits those same meanings in order to understand what is written. For an example, consider sentence (1): Some American musicians are scared of a small Norwegian troll.

Most likely, (1) has never before been written. That, of course, does not prevent it from meaning whatever it does in English. It has its meaning independently of ever having been uttered or thought about. So far our discussion has been primarily concerned with the meaning that sentences have in English (by virtue of being English sentences)—that is, their conventional or literal meaning. The study of the literal meaning of words and sentences is often called semantics.

Conventional meaning, however, is as we have seen not the be all and end all of communication. We often (maybe always) use sentences to communicate contents quite different from their conventional meaning, as observed in the following conversation. Sam asks Chris in sentence (2): Can you help Alex with his paper tonight? Chris in sentence (3) responds: I'm driving into New York to see Jill. By uttering (3), Chris can succeed in telling Sam that she cannot help Alex with her paper that night. Of course, that's not the literal meaning of sentence (3). The literal meaning of that sentence is that Chris is driving into New York to see Jill. But by uttering (3), Chris can succeed in communicating to Sam more than the literal meaning of the sentence she uttered. The study of how words and sentences can be used to communicate contents that go beyond their literal meaning is often called pragmatics. The goal of pragmatics is to study the various mechanisms that speakers exploit to communicate content that goes beyond literal meaning (for more on the distinction between pragmatics and semantics, see H.P. Grice's *Studies in the Ways of Words*). But in ascribing conventional meaning, one can incur theoretical costs.

REPRESENTATIONAL AND COMPOSITIONAL MEANING (SEMANTIC) THEORIES

Philosophers of language and linguists talk about the vehicles that carry meaning as both representational and compositional. Representations represent—so the sentence "Bill Clinton is tall" represents Bill Clinton as tall; however, the sentence "The president of the United States in 1999 is tall" also is true of Bill Clinton, and also represents him as tall, but it does so in a different manner. But it differs not only inasmuch as it uses a different vehicle. The Italian sentence "*Il presidente degli Stati Unitii in 1999 e' alto*" represents Bill Clinton as tall in exactly the same way that "The president of the United States in 1999 is tall" does, even though these two vehicles of representa-

tion are distinct. With these two sentences, however, the vehicles are synonymous—they carry the same meaning, whereas the first two are not synonymous, though both vehicles happen to be true in the same circumstances.

Suppose, for instance, that someone else had been president in 1999; then, the latter two sentences with definite descriptions might be false, but the first sentence with a proper name would still be true. So, whatever meaning is, it would appear to be more fine-grained than a mere symbol-object relationship. If words were merely tags for objects, no two co-tags would differ in meaning. It would seem that vehicles denote objects under representational guises, and these guises are part of what that expression means.

There has been much written about the nature of this guise, yet little of it has been clear. Whatever guises are, we have seen that they must be more fine-grained than the objects to which expressions apply because expressions with the same referent can differ in meaning, but guises must also be more coarsely-grained than the ideas speakers associate with expressions. Two people might use the same expression but associate different ideas with it; for you, *snow* might connote misery but for a skier it might connote joy.

Synonymous sentences in the same or distinct languages are supposed to share guises; those that are nonsynonymous do not, even if the sentences happen to be about the same objects, events, or state of affairs. Like the shadows in Plato's cave allegory, guises suggest existing somewhere in between linguistic items and idiosyncratic ideas associated with expressions by individual speakers, on the one hand, and the objects to which they are conventionally attached, on the other.

Guises are what determine whether a linguistic item is about one thing and not another; they are the concepts that enable us to understand the linguistic items we use. The definite descriptions *the forty-second president of the United States* and *the husband of Hilary Clinton* pick out the same person, Bill Clinton, but they do so in different ways. The ways in which they pick him out are another way to think about the guises associated with expressions. The former expression picks out Bill Clinton partly by virtue of his having the property of being the forty-second president of the United States; and the latter expression picks him out partly by virtue of his being Hilary Clinton's husband. Thus, these two expressions each represent the same individual, but they do in different ways—under different guises.

But there is more to the concept of a guise than is evidenced by representational powers. Natural languages are essentially productive and systematic. They exhibit productivity in that there are no obvious upper bounds on the number of creative linguistic acts that can be performed through speech. Novel sentences can be formed by conjoining any two meaningful indicative sentences—as in, "John left, but Mary stayed"—or by prefacing any meaningful indicative sentence with a psychological verb—as in, "Carl believes that Martha is ill" or "Carl fears that Martha is ill."

Because humans lack magical abilities, this capacity to produce and comprehend novel linguistic acts requires explanation. The standard explanation is that speakers of a natural language must have learned rules that enable them to determine the meaning of a complex expression strictly on the basis of its significant parts. The existence of such compositional rules explains our capacity with productive representational systems—by assuming that any unbounded representational system is compositional, we have an explanation for mastery over productive representational systems (for further discussion of compositionality, see Jerry Fodor's and Ernie Lepore's *Compositionality Papers*).

The property of compositionality can also be invoked in order to explain the following feature: It is a distinctive feature of English that when a grammatical sentence of the form "A R's B" is meaningful, then if "B R's A" is grammatical, not only is it also meaningful, but its parts are presumed to make the exact same meaningful contribution that they do in the original configuration. This aspect of a representational system is referred to as its systematicity.

The existence of a set of compositional rules accounts for systematicity as well as productivity. Compositionality requires that meaningful expressions compose in systematic ways to produce meaningful complexes. The expressions *the red shoe*, *the table*, and *fell on* mean what they mean regardless of whether they are configured to read "The red shoe fell on the table," or "The table fell on the red shoe." To be more specific: reconsider (1). Its literal meaning and, indeed, the literal meaning of any English sentence, depends on two factors: A) the meaning of the words (i.e., *some*, *American*, *musicians*, *are* and *troll*; and B) the way in which these words are assembled. Put together as in (4), what results is a sentence entirely different in meaning from (1): (4) Some Norwegian musicians are scared of a small American troll.

From these apparently obvious facts we can derive the idea that languages have compositional meaning theories. The idea is that the literal meaning of a sentence (its literal or conventional content) is the result of the (literal/conventional) meaning of its parts (the words in it) and the manner in which these parts are put together (their mode of composition).

Furthermore, as we have already noted, in addition to the systematicity of our sentences, speakers are also able to understand and produce indefinitely many sentences—sentences neither they nor anyone else in their community has ever uttered before. This shows that their knowledge of language must be productive; it must extend beyond a fixed lexicon of predefined static elements, and must include a generative system that actively composes linguistic knowledge so as to describe arbitrarily complex structures. The hallmark of productivity in language is recursion. Recursive patterns of complementation, as in (5), and recursive patterns of modification, as in (6) and (7), allow phrases to be nested indefinitely many times within a single sentence: (5) Chris thinks that Kim thought that Robin wanted Sandy to leave; (6) Chris bought a gorgeous new French three-quart covered copper saucepan; (7) Chris is writing a book that describes inventors that have built machines that changed the world that we live in.

Speakers' capacity to formulate and recognize an open-ended array of possible sentences shows how acute a problem it is to coordinate meaning across speakers. When we learn the meanings of expressions of our native language, we must generalize from the finite record of our previous experience to an infinity of other expressions and situations. If we thereby arrive at a common understanding of the meanings of these expressions, it must be because language is structured by substantive and inherent constraints that we are able to exploit. More generally, if our discoveries in the theory of meaning are to help explain how speakers can use language meaningfully, we should expect that the generative mechanisms we postulate as theorists will be compatible with the psychological mechanisms that underlie speakers' abilities.

There are many ways to implement this idea of a compositional meaning theory. One that has been prominent in the philosophical literature is that a theory of meaning for a natural language, L, should consist of a finite set of axioms specifying the meaning of the words and the rules for how they can be composed. These axioms would then permit the derivation of theorems that specify the meaning of complex expressions (such as *some American musicians*) and sentences, such as (1)–(7). So understood, a semantic theory is a formal theory from which we can derive the meaning of an infinity of English sentences. The reason why (1)–(7) mean what they mean in English is that their meanings are encoded, so to speak, in the basic axioms of a correct meaning theory for English.

A straightforward way, then, for a philosopher of language to explain productivity and systematicity is to assume that the meanings of particular sentences can be calculated by inference from general facts about meaning in the language. For example, consider the compositional meaning theory presented in (8)–(10): (8) *Snow* is a noun phrase and refers to the stuff snow; (9) *White* is an adjective phrase and refers to the property whiteness; and (10) If N is a noun phrase and refers to the stuff S and A is an adjective phrase and refers to the property P, then N is A is a sentence and is true if, and only if, S is P.

From this theory, we can derive (11) as a logical consequence: "Snow is white" is true if, and only if, snow is white. Why should we think of (11) as a characterization of the meaning of the English sentence "Snow is white?" We can because it links up this sentence with a condition in the world stated in objective terms—in this case, the condition that snow is white. As theorists of meaning, we can utilize this kind of theory, which Donald Davidson calls an interpretive truth-theory, to provide a general account of how sentences link up with conditions in the world (Davidson 1967, 2001; Lepore and Ludwig 2005).

We use atomic formulas to axiomatize the meanings for elementary structures in the language and use conditional formulas to describe the meaning of complex structures in the language as a function of the semantics of their constituents. We then reason logically from the axioms to associate particular sentences with particular conditions in the world. As in (8)–(10), this inference will be compositional, in that the conclusions we derive will be inferred through a logical derivation that mirrors the syntactic/grammatical derivation of the sentence.

There are two ways to view interpretive truth-theories such as (8)–(10). We can exploit an interpretive truth-theory to formulate a theory of meaning for a new language. For example, we could be pursuing translation. In this case, we are interested in systematically articulating translations of sentences in the object language in terms of sentences in our own; we understand these translations to be derived by inference from the axioms of the theory. Another way to view interpretive truth-theories (and other sorts of compositional theories of meaning), such as (8)–(10), is as ingredients of the speakers' psychology. On this view, we regard the axioms of a

theory of meaning as generalizations that native speakers know tacitly about their own language. When speakers formulate or recognize particular utterances, they reason tacitly from this implicit theory to derive conclusions about specific new sentences. On this understanding, interpretive truth-theories offer an explanation of how speaker knowledge of meaning and inference underlie linguistic competence.

FORMALISM IN PHILOSOPHY OF LANGUAGE

The view we just described invites an analogy between the semantics of natural languages and the semantics of the artificial languages of formal logic. The analogy goes back to Gottlob Frege (1879), who took logic to clarify the features of natural language essential for correct mathematical thought and communication. The work of Richard Montague (1974) took the analogy further. Montague explicitly advocated an exact parallel between the semantic analysis of English—what ordinary speakers actually know about their language—and the semantics of intensional higher-order logic. In fact, many techniques originally developed for giving semantics to logical languages turn out to be extremely useful in carrying out semantic analysis.

INDIRECT SPEECH ACTS

Interpreting a dream partly involves assigning it meaning, but does this imply that dreams are representational in the way that language is? In one sense, they are obviously so. This is the sense in which we might say of any image that it is representational. An image of a horse is of a horse, and not of sheep. But this is a notion of representation irrelevant to our current concerns in the philosophy of language, because it appeals to a natural (and not a conventional) relation between an image and its corresponding object. If dreams are supposed to be representational in the same sense in which photographs or other sorts of images are, then talk of a compositional theory of interpretation or meaning of dreams is not anything like the sort of theory that one invokes for systems of representations such as natural language. For one, photographic images are neither productive nor systematic, nor are they even fine-grained in the way in which linguistic representational systems are. An image of Bill Clinton is an image of the president of the United States, and nothing short of an election can pull them apart. More famously, an image of John giving Bill a toy is indistinguishable from an image of Bill receiving a toy from John, though these inseparable acts are distinct. It is clear that the sort of systematicity that occurs so naturally within bona fide linguistic representational systems cannot be applied to images with the same ease.

We return now to our earlier contrast between literal/conventional meaning and meaning in purpose or what we might call agent meaning. When the subject is employing so-called indirect speech acts, then what one means by one's words must take into consideration background factors. So, for example, suppose Janet says, "It's raining outside." Her words mean that it is raining outside, but she herself might mean for her audience to bring their umbrellas. When Janet spoke she intended her audience to come to believe what she was trying to get across. In order for her words to have meant that her audience is to take their umbrellas, she must have intended her audience to recognize her ulterior motive.

Speaker meaning in contrast to literal/conventional meaning, then, requires (at least) two sorts of intentions, one about what a speaker is trying to get their audience to believe by their utterance and another about getting them to recognize what he or she is trying to do. More specifically, what a speaker means by their words depends on what they intend their audience to come to believe, and what he or she intends them to recognize him or her as intending them to come to believe. Both component intentions, tacitly or not, must accompany an utterance in order for the speaker to mean something by what they say. By Janet's utterance of "It's raining," she means for her listener to bring their umbrella just in case she intends them to come to believe this and she intends them to recognize that she intends them to come to believe this. She intends for them to come to believe they are to bring their umbrella, and she intends them to recognize that she intends them to come to believe they are to bring their umbrella.

Implicit in our discussion is, of course, the assumption that speaker meaning can exceed word meaning. For you to bring your umbrella is not what Janet's words "It's raining" literally/conventionally mean, nor is it implied by anything that these words literally/conventionally mean. Speaker meaning is determined by word meaning alone just in case it is either expressed or implied by what the words used mean; conversely, it is not determined by literal meaning alone if it is neither expressed nor implied by what the speaker's words literally mean. A simple test separates the former distinction from the latter. If we try to deny speaker meaning determined by word meaning, then we end up making inconsistent claims. Because Janet can consistently assert that it is raining outside without intending for you to bring any umbrella, what

she means is neither expressed nor implied by what her words mean (Grice 1989).

Inquiries about speaker meaning not determined by word meaning are about nonlinguistic motives, beliefs, desires, wishes, fears, hopes, and other psychological states that provoke verbal expression. Speaking is an action; it is what we do with meaningful words. This requires reasons, and reasons not entirely about what our words mean. Linguistic and nonlinguistic psychological states both come into play.

SENTENCES MEANING AND UNDERSTANDING

To sum up: One chief goal of philosophy of language is to show how speaker knowledge of a natural language allows speakers to use utterances of sentences from their language meaningfully. As we have seen, one rough and tentative answer has been: If speakers know a recursive compositional meaning theory for their language, then they can use its rules and axioms to calculate interpretive truth conditions for arbitrarily complex novel sentences. But we have also seen that even if speakers can infer the truth conditions of sentences from their language on the basis of (tacitly) employing a compositional meaning theory for their language, such knowledge alone cannot account for all of what goes on in communication. Communication invariably takes us further than the literal/conventional meaning of our words. How do we go further in a communicative exchange than what our words literally mean?

A preliminary, approximate answer is this: We begin by idealizing the information mutually available to us in a conversation as our common ground (Stalnaker 1973). The common ground settles questions about whose answers are uncontroversial, in that interlocutors know the answers, know that they know the answers, and so forth. Meanwhile, the common ground leaves open a set of possibilities about which there is not yet agreement: Maybe there is a matter of fact that could turn out (for all that the interlocutors know) to be one of various ways, or maybe the interlocutors actually do know how it turns out but do not realize that the knowledge is shared—so it could be that the others know, and it could be that they do not—and so forth. We might represent these possibilities in the common ground as a set of possible worlds (situations).

Let the set of possible worlds in which a given sentence is true represent the proposition associated with the sentence. If we adopt this picture, then we can formalize the effect that asserting a formula has on the common ground. When interlocutor A asserts a formula f, he or she introduces into the conversation the information that f is true. Suppose that f expresses the proposition that p. Before A asserts f, the common ground is some set of worlds C. After, the common ground must also take into account f. This formula f restricts the live possibilities by requiring the worlds that are in the common ground to make true the further proposition that p. So, the change that occurs when A asserts f is that the common ground goes from C to C together with the proposition that p. This concise model forms the basis of a range of research characterizing the relationship between truth-conditional semantics (literal/conventional meaning) and conversational pragmatics in formal terms (van Benthem and ter Meulen 1997).

This idealization obviously has its limits. And it is easy to come up with strange puzzles when one moves (perhaps inadvertently) beyond the limits of these idealizations. Before considering one such puzzle, we digress to discuss perhaps one of the most important results from one of the most important research programs in the philosophy of language in the last half-century.

SAUL KRIPKE AND HILARY PUTNAM ON TWIN EARTH

Imagine a planet exactly like Earth, except that where Earth has water, this other planet, Twin Earth, has another mysterious substance, say, XYZ. To human senses, this substance seems exactly the same as water; nevertheless, it has a fundamentally different chemical structure. Imagine further that it is still the year 1700, and chemical structure has yet to be discovered. Still, we judge that the English word *water*, on Earth, means water, whereas the Twin English word *water*, on Twin Earth, means XYZ. Moreover, if an earthling were suddenly teleported to Twin Earth, they would still speak English, and their word *water* would still mean water—this despite the fact that they might have exactly the same dispositions as Twin Earthers have to accept or reject statements about their new surroundings. In short, the unfortunate earthling would think they were surrounded by lots of water, and would be completely wrong.

What moral should we draw from Putnam's (1975) Twin Earth thought experiments? Should we conclude that when you look at how a speaker is disposed to respond to English sentences, *water* can be interpreted equally well as water, XYZ, or even the disjunction of the two? These interpretations are different, and they assign distinct truth values to English sentences in meaningful (but ultimately inaccessible) situations. In fact, though,

when we say *water* in English means water, according to Kripke, we are applying a standard based on our recognition that English speakers intend to pick out a particular kind of stuff in their own environment.

As a community, English speakers have encountered this stuff and named it *water*. And as a community, English speakers work together to ensure first that the community maintains the referential connection between the word *water* and that stuff, and only secondarily, that the individuals in the community can themselves identify examples of the stuff in particular situations. When as observers we recognize that *water* means water, we are not summarizing the epistemic abilities of particular speakers. Rather, we are summarizing social commitments and causal connections in the community that have worked across speakers to hook the word *water* up with the stuff, and keep it that way. What philosophers of language do, ultimately, is to explain how speakers can use language to refer in shared ways to shared aspects of the world.

Kripke (1972) motivates his account with an analogy between words for kinds, such as *water*, and proper names, such as *Richard Feynman*. In the case of proper names, we can point to the social practices that initially fix the reference of a name and transmit that reference within the community. A baby boy is born. His parents call him by a certain name. They talk about him to their friends. Others meet him. The name spreads from link to link much like a chain. To use another example: Let us say that a speaker on the far end of a similar type of chain, who hears about Richard Feynman, may be referring to him even though they cannot remember from whom they first heard his name. They know Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. The speaker is then referring to Feynman even though he or she cannot identify him uniquely. He or she does not know what a Feynman diagram is and does not know what the Feynman theory of pair production and annihilation is. Not only that, the speaker would have trouble distinguishing between Gell-Mann and Feynman (Kripke 1980).

The result is that we can judge a speaker's reference with a proper name independently of sentences that the speaker would accept or reject. In the case of common nouns such as *water*, the word has had its reference since time immemorial. Nevertheless, new speakers still link themselves into chains of reference that participate in and preserve the connection between *water* and water. So analogously, we take an English speaker's word *water* to refer to water, independently of sentences the speaker would accept or reject.

Most philosophers of language find the Kripke/Putnam views about the meanings of names and so-called natural kind terms satisfying; it offers a close fit to an intuitive understanding of ourselves. It seems that we really do commit to use our words with the same reference as our community. And when others make claims about the world, it seems that we really do assess and dispute those claims with respect to the common standard in the community.

For example, on the Kripke/Putnam view, we inevitably focus on certain aspects of an agent's verbal behavior and not others when we assign meanings to their utterances. We do so because we locate the theory of meaning as part of a broader science of the mind, which combines a theory of language with a theory of action (including an account of our intentions and social relationships) and a theory of perception (including an account of the limits and failings of our observation). The theory of meaning in itself explains only so much—and, not surprisingly, just because we understand the meaning of someone's sentences, we do not ipso facto understand them.

Crucially, this new view predicts that some statements are necessarily true solely in virtue of the meanings of the words involved. We have already seen that it is a fact about meaning that *Richard Feynman* names Richard Feynman, or that *water* names water. We can go further. *Hesperus* names the planet Venus, *Phosphorus* names the planet Venus, *is* names the identity relation. So sentence (12) follows, just as a matter of meaning alone: (12) Phosphorus is identical to Hesperus. Given that *Hesperus* and *Phosphorus* are both names for the planet Venus, (12) must be true. There is no way that that planet could have failed to be that planet. Like sentence (12), the other facts that follow from the meanings of our language are necessarily true.

However, on the Kripke/Putnam account, facts about meaning turn out not to be knowable a priori. We discover them. To illustrate, imagine that, early on, the ancient Greeks were in an epistemic situation that left it open whether the bright object that sometimes appeared in the morning sky was the same as the bright object that sometimes appeared in the evening sky. They could not distinguish themselves from their doubles on a Twin Earth where the morning star and the evening star actually were distinct objects (alien satellites, we might suppose). These Twin Earthers would speak a language in which (12) translates into a false sentence—indeed, a

necessarily false sentence. For the ancient Greeks, however, the translation of (12) was necessarily true. Eventually, the ancient Greeks advanced their science, and improved their epistemic situation. They realized that, in our case, there is only one celestial object. At the same time, then, they discovered that (12) is necessarily true.

When we reflect on the generality of Twin Earth thought experiments, it is clear that facts about meaning are knowable a posteriori. We can imagine being quite wrong about what our world is like. In these imaginary situations, our empirical errors extend to errors we make about what our words mean. And, of course, we can also imagine disagreeing with others about what the world is like. Though we are committed to use our words with the shared reference of our community, we must be prepared to resolve our dispute by giving up facts that we think are necessarily true—facts that we think characterize the meanings of our words and the contents of our thoughts. With this model of how proper names and common nouns attach to the world before us, we are now ready to return to the puzzles alluded to above in connection with assertion.

ASSERTION

Why would a speaker ever assert an identity statement like (12)? The trigger for a puzzle comes from arguments that sentence (12) must be true. If this is so, then consider what happens when A asserts C. We update the common ground C by intersecting it with the set of all possible worlds (situations)—the proposition expressed by Hesperus is Phosphorus—leaving exactly the same set C. A, therefore, on this model, has done nothing; the interlocutors' information has not changed at all! But obviously this result is absurd. What has gone wrong?

In fact, in assuming that assertions update the context with the proposition they express, we have implicitly assumed that the participants in the conversation have certain and complete knowledge of their language. For example, interlocutors can calculate that Hesperus is Phosphorus expresses a necessarily true proposition only if they can calculate that Hesperus names Venus and Phosphorus names Venus. Of course, under such circumstances, they do not learn anything from the sentence. It is easy to see how this assumption could go unnoticed.

In discussion, we typically assume the reference of our terms—precisely what matters in the "Hesperus is Phosphorus" case—is not at issue. However, consider how to formalize uses of sentences in more realistic situations (as we do so, we must be careful to respect the intuitions of Kripke's and Putnam's thought experiments [Stalnaker 1978]). Suppose an interlocutor B does not know that Hesperus is Phosphorus. What that really means is that B cannot distinguish between two possible situations. In the first, there is only one heavenly body out there, and B's community speaks a language $English_1$ where both Hesperus and Phosphorus are names for that body. In the second, there are two distinct heavenly bodies, and B speaks a language $English_2$ where Hesperus is a name for one of them and Phosphorus is a name for the other. Because these possibilities are open for B, they must both also be represented in the common ground.

Now, we need a correspondingly expressive notion of assertion. When interlocutor A says something, A is committed that it is true according to the standards for reference that prevail in the community. Any assertion that A makes should turn out to be true in the language A speaks. What we have just seen is that any point of evaluation w in the common ground could potentially have its own language $English_w$ with relevant differences from English as spoken in the real world. Adapting Stalnaker's (1978) terminology, we can associate any utterance u with a diagonal proposition; this proposition is true at a point w if the proposition that u expresses in $English_w$ is true in w.

In the case of Hesperus being Phosphorus, the effect of A's assertion is to intersect the common ground with this diagonal proposition. Concretely, we retain in the common ground worlds of the first kind, where $English_1$ is spoken, Hesperus and Phosphorus are necessarily the same and A's assertion is necessarily true. However, we discard from the common ground worlds of the second kind, where $English_2$ is spoken, Hesperus and Phosphorus are necessarily different and A's assertion is necessarily false (there is substantially more to be said about the relationship between utterance meaning and the information that interlocutors convey).

METHODOLOGICAL ISSUES

We have seen that many important philosophical issues have to be settled in advance before a theorist can construct a compositional meaning theory for, say, English in order to account for linguistic competence with English. For example, the theorist is required to be guided by some idea of what counts as getting it right. If the goal is to get a set of axioms from which the theorist can infer the literal meaning of all possible English sentences, he or she needs to have some idea of how to determine that a particular theory implies the correct literal meanings. Here are four interrelated philosophical topics devoted to such methodological issues:

- Semantics-Pragmatics Distinction: Within the totality of communicated content (all the information communicated by an utterance) it is difficult to distinguish between the literal content and that which is generated through various pragmatic mechanisms (it has proved exceedingly difficult to distinguish between semantic content and pragmatic content). Any theory of meaning must incorporate criteria that distinguish different kinds of content and tells us how to classify content. Many debates in philosophy of language are based, in part, on different ways of drawing the semantics—pragmatics distinction.
- Role of Appeals to Intuitions: Most arguments in philosophy of language appeal to intuitions. We appeal to intuitions about what was said about grammaticality, about inferential connections, and sometimes about what would be true in other possible worlds. No position in the philosophy of language can be defended without various appeals to intuitions. That raises two questions: Why should we think intuitions provide us with reliable evidence? What kinds of intuitions should we rely on?
- The Nature of Meaning: How a philosopher of language goes about constructing a theory of meaning will depend on what he thinks meaning is. Are meanings entities? Is meaning reducible to something else? Do we even need to appeal to meaning or can we leave it out of theory of communication? The meanings of sentences are often referred to as propositions. What are propositions? These foundational issues have dominated discussion in philosophy of language for centuries.
- The Nature of Languages: There is an ongoing philosophical debate about what languages are, what kind of objects they are. Some think they are abstract objects, some think they are social/public objects, some think they are psychological structures, some think natural languages such as English should play an important theoretical role, some think they are superfluous in a serious meaning theory.

WIDER PHILOSOPHICAL IMPLICATIONS

To the noninitiated, research in the philosophy of language can seem technical and without deep philosophical implications. However, any such perception is simply the result of ignorance. Debates in the philosophy of language have wide-reaching implications for all branches of philosophy and research in those other branches inevitably make assumptions about issues that belong under the rubric *philosophy of language*. Indeed, it is not possible to do serious work in any branch of philosophy today without a solid training in the philosophy of language.

The list of such important connections between the philosophy of language and the rest of philosophy could be made very, very long indeed. Limitations of space require we restrict attention to a few topics—epistemology will be one of them. Some of the most discussed contemporary positions in contemporary epistemology draw in a very direct way on views from the philosophy of language.

David Lewis (1996) claims that the epistemological skeptic (i.e., someone who argues that knowledge is impossible) can be refuted once the correct theory of meaning for *know* is adopted. According to Lewis, the correct theory for *know* is one that assigns it a context sensitive meaning, much as with the expressions *I*, *you*, and *here*. Obviously, once someone claims that the meaning of an expression is context sensitive, they become accountable to the philosophy of language. The theory of meaning for context sensitive expressions such as *I* is well-evidenced, and so, if *know* is like them it will have to stand up to certain qualifying tests all such expressions satisfy.

Putnam (1982) argues that his theory of meaning and reference implies that the skeptic's central argument is incoherent. His argument is based on a philosophical position on the nature of meaning. To the extent that his theory of meaning stands up to the scrutiny of the philosopher of language, skepticism may be refuted.

Kripke (1972) argues, as we saw above, that his theory of proper names refutes the traditional view (going back at least to Immanuel Kant) that necessary truths can only be knowable a priori and contingent truths only a posteriori. According to Kripke, it follows from the theories of meaning for proper names and natural kind terms such as *gold* and *tiger* that we can discover necessary truths empirically (many scientific discoveries turn out to be discoveries of necessary truths), and it turns out that we can gain knowledge of contingent facts a priori.

Some of the most discussed contemporary positions in contemporary metaphysics also draw in a very direct way on views from the philosophy of language. Kripke (1972) argues that his theory of reference implies that mental states cannot be physical states (i.e., that materialism is false).

Some of the most discussed contemporary positions in contemporary value theory draw in a very direct way on views from the philosophy of language also. One of the central strands in contemporary ethics is called expressivism. This is the view that sentences containing moral terms (e.g., *good*, *bad*, *should* and so on) cannot be true or false. They serve simply to express attitudes. Expressivism is a view about the meaning of words (Ayer 1946).

See also Artificial and Natural Languages; Conditionals; Content, Mental; Contextualism; Davidson, Donald; Frege, Gottlob; Hobbes, Thomas; Intuition; Kant, Immanuel; Kripke, Saul; Language; Lewis, David; Meaning; Montague, Richard; Phonology; Plato; Pragmatics; Propositions; Putnam, Hilary; Reference; Rule Following; Semantics; Semantics, History of; Sense; Syntactical and Semantical Categories; Syntax; Vagueness; Wittgenstein, Ludwig Josef Johann.

Bibliography

Alston, W. *Philosophy of Language*. Englewood, NJ: Prentice Hall, 1964.

Ayer, A.J. *Language, Truth, and Logic* (1936), 2nd ed. London: Gollancz, 1946.

Benthem, Johan van and Alice ter Meulen, eds. *The Handbook of Logic and Language*. Cambridge, MA: MIT Press, 1997.

Chomsky, Noam. "A Review of B. F. Skinner's Verbal Behavior." *Language* 35 (1959): 26–58.

Davidson, Donald. *Inquiries into Truth and Interpretation*. 2nd ed. Oxford: Clarendon Press, 2001.

Davidson, Donald. "Truth and meaning." *Synthese* 17 (1967): 304–323.

Fodor, Jerry A. and Ernie Lepore. *The Compositionality Papers*. Oxford: Oxford University Press, 2002.

Frege, Gottlob. *Begriffsschrift, eine der arithmetischen nachgebildete formelsprache des reinen denkens*. Halle: L. Nebert, 1879.

Grice, H.P. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

Kripke, Saul A. *Naming and Necessity*. Cambridge MA: Harvard University Press, 1980.

Kripke, Saul A. *Naming and Necessity*. In *Semantics of Natural Language*, edited by D. Davidson and G. Harman. Reidel, 1972.

Lepore, Ernest and Kirk Ludwig. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press, 2005.

Lewis, David. "Elusive Knowledge." *Papers in Metaphysics and Epistemology* (1996): 549–567.

Ludlow, Peter. *Readings in the Philosophy of Language*. Cambridge, MA: MIT Press, 1997.

Montague, Richard. "English as a Formal Language." In *Linguaggi nella società e nella tecnica*. Milan: Edizioni di Comunità, 1970.

Montague, Richard. *Formal Philosophy: Selected Papers of Richard Montague*, edited and with an introduction by Richmond H. Thomason. New Haven, CT; London: Yale University Press, 1974.

Montague, Richard. "The Proper Treatment of Quantification in Ordinary English." In *Approaches to Natural Language*, edited by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes. Dordrecht: Reidel, 1973.

Putnam, Hilary. "The Meaning of 'Meaning.'" In *Philosophical Papers II: Mind, Language, and Reality*. Cambridge, U.K.: Cambridge University Press, 1975.

Putnam, Hilary. *Reason, Truth, and History*. Cambridge, U.K.: Cambridge University Press, 1982.

Quine, W.V.O. "Two Dogmas of Empiricism." In *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953.

Skinner, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, 1957.

Stalnaker, Robert. "Assertion." In *Syntax and Semantics 9*, edited by Peter Cole. New York: Academic Press, 1978.

Stalnaker, Robert. *Context and Content*. Oxford: Oxford University Press, 1999.

Stalnaker, Robert. "Presuppositions." *Journal of Philosophical Logic* 2 (4) (1973): 447–457.

Ernest Lepore (2005)

PHILOSOPHY OF LANGUAGE IN CONTINENTAL PHILOSOPHY

The task of the philosophy of language within the tradition of continental European philosophy has been to overcome the idea of language as an instrument or as a means at the disposal of human beings. Although it has proved possible retrospectively to see Georg Wilhelm Friedrich Hegel (1770–1831) as a resource for this task, Johann Georg Hamann (1730–1788) and Wilhelm von Humboldt (1767–1835) both contributed more. Hamman was the first to give centrality to language and Humboldt, with his formulation that language is an *energeia* not an *ergon*, an activity not a work, opened the door to a more dynamic approach to it. However, it was not until the second half of the twentieth century that these insights were fully explored and decisively surpassed.

MARTIN HEIDEGGER

Martin Heidegger's attempt to go beyond the instrumentalist and expression theories of language is most pronounced in his later thought, especially in *On the Way to Language*. His formulation, *Die Sprache spricht* (language

speaks) is an effort to displace the centrality of the human subject in accounts of language: It is not primarily the human being, but language, that speaks. The human being speaks only in response to language. This insight arose when he shifted his focus from everyday speech, which is explored at length in *Being and Time* as part of his account of everyday existence, to the poetic word. Already in 1936, in "The Origin of the Work of Art" (2002 [1950]), Heidegger claimed that it was not the human being, as in *Being and Time*, but art and most specifically poetry, that brings beings into the open and gives to human beings their outlook on themselves. This led directly, some ten years later, to the famous formulation of his *Letter on "Humanism"* (1998 [1947]) that "language is the house of Being" (p. 239). It announces not only the sense in which humans inhabit language, but also the sense in which the human being belongs to the historical destiny of Being and is called to respond to it. The implications of this account emerge not only in his readings of the poetry of, for example, Friedrich Hölderlin, Stefan George, Rainer Maria Rilke, and Georg Trakl—readings that are directed to undergoing an experience with language such that language transforms people—but also in his reading of the history of philosophy, where thinkers are understood to be saying the word of Being for their time. The words of Being function, somewhat like the work of art, to found a world.

Heidegger's approach to language is directed against the tendency to understand language in terms of something else, such as activity, spirit, or world view. That is why the focus falls on experiencing language. It is Heidegger's view that language shows itself as language only when language comes to be infused with silence. Language comes to be infused with silence mundanely when language fails people so that they are lost for words. For Heidegger, the thinker experiences something similar at a more profound level at the end of European and North American metaphysics. At that time the thinker lacks a word for Being and so can no longer accomplish the philosophical task of naming Being, for example, as *idea*, *energeia*, *subjectum*, or will. Indeed, for Heidegger it is only the lack of a word for Being in our epoch that gives rise to the insight that naming Being was the philosopher's task. However, this is not a negative experience. It is in the experience of language that Heidegger positions his thought as no longer metaphysical, albeit it is not yet beyond European and North American metaphysics. That Heidegger's clearest accounts of this experience arise in the course of his readings of Hölderlin and George show the extent to which his own self-understanding was moulded by the dialogue between poetry and thinking.

MAURICE MERLEAU-PONTY

Like Heidegger, Maurice Merleau-Ponty in *Phenomenology of Perception* (1962 [1945]) distinguished a creative or speaking speech that formulates for the first time, which he called *parole parlante*, from ordinary or spoken speech, *parole parlée*. What unites all of Merleau-Ponty's texts on language is a concern for the creative aspect of language, its capacity to say what has never been said, which he explored as an antidote to the dream of some philosophers to develop a transparent, algorithmic, language. However, unlike Heidegger, Merleau-Ponty's approach to language was from the outset already informed by psychology, and by the late 1940s he had begun to incorporate developments in linguistics. This tendency culminates in "Indirect Language and the Voices of Silence," which begins with Ferdinard de Saussure's insight that meaning is a function of the differences between words, their divergence from each other. Words do not directly signify anything; they are not tied to a preestablished signification. There is thus an "instructive spontaneity" of speech that leads Merleau-Ponty to the insight that people do not speak of Being so much as Being speaks in them, a formulation with clear Heideggerian echoes. The vitality of speech is also apparent when in a conversation one can no longer tell, as Merleau-Ponty famously puts it, what comes from one's dialogue partner and what is one's own contribution.

HANS-GEORG GADAMER

Dialogue is also at the core of Hans-Georg Gadamer's account of language. In his philosophical hermeneutics, which he developed in most detail in *Truth and Method* (1989 [1960]), he highlighted how in dialogue one seeks to reach an understanding with a living person or a text about some topic. However, underlying the effort to reach agreement was an already existing agreement because every dialogue presupposes a community of language as the element in which the dialogue takes place. Hence he conceived the task of a hermeneutical reflection on language not as that of investigating how each language in spite of its differences from other languages could say everything it wants to say, which he characterized as a concern of the philosophy of language and linguistics. His question was rather how to make sense of the intimacy of thought and language because language is not a prison, which is evident because one can readily come to understand a foreign language. Gadamer's answer was to reject accounts of language that relied on conventionalism and preschematization in favor of an account that emphasized its generative and creative power. This led

Gadamer to formulate the idea of the virtuality of language, by which he meant its inexhaustibility, its capacity to exceed what has already been said. Gadamer's account of the infinite resources of language can be seen as an attempt to resist Heidegger's account of the breakdown of the function of language within European and North American metaphysics, but he shared with Heidegger the conviction that language has people in its grip, that it speaks people more than people speak it. As evidence for this view he cited that the time when a text was written can be more precisely determined by its linguistic usage than from its author.

JACQUES DERRIDA

At the heart of Jacques Derrida's understanding of language is his identification of European and North American metaphysics with logocentrism, such that the alleged primacy of presence within European and North American metaphysics is reflected in the alleged transparency of speech and the speaker's mastery over it. By contrast, writing, even before it reaches its destination, is organized around the absence, and possible death of the sender or the addressee, or both. Derrida's deconstruction of logocentrism is sometimes mistakenly understood as a championing of writing to compensate for its previous reduction to the status of a mere supplement to speech, for example, as when Plato presented it as an aid to memory. Nevertheless, Derrida's interest is not so much in what is normally understood by writing as in what he calls arche-writing or protowriting, which is the condition of all forms of language, indeed of all organized systems. Derrida's use of the word "writing" in this contest is strategic: It is intended to reverse the priority of speech over writing, but only as a prelude to passing beyond the opposition between them both.

As Derrida explained in *Of Grammatology* (1976), the inflation of the sign *language* is the inflation of the sign itself. He presented this as a symptom of the historical epoch in which what had finally been gathered under the name *language* came to be summarized as *writing*. Derrida thus does not advocate grammatology in the sense of a science of writing, but, engaging in his own form of grammatology in the sense of a provisional science of textuality, he finds that both linguistics and psychoanalysis fail to recognize the resistance of language to pure ideality, and thus fail to escape logocentrism. However, once made thematic, this tension need not be regarded negatively. By shifting the focus to textuality Derrida draws attention to the way that one can find, for example within the language of Saussure and Sigmund Freud, both the symptoms of the logocentrism of European and North American metaphysics and the trace of what it represses. Derrida performs a similar operation on Heidegger, from whom he had initially drawn the basic outline of his account of European and North American metaphysics. In this way, Derrida continued Heidegger's project of overcoming the conception of language as instrument or medium, but without relying on poetic language to accomplish the task, as had been the case with Heidegger, Merleau-Ponty, and Gadamer.

See also Derrida, Jacques; Gadamer, Hans-Georg; Heidegger, Martin; Hermeneutics; Merleau-Ponty, Maurice.

Bibliography

Derrida, Jacques. *De la grammatologie*. Paris: Minuit, 1976. Translated by G. Spivak as *Of Grammatology* (Baltimore, MD: Johns Hopkins University Press, 1976).

Gadamer, Hans-Georg. *Wahrheit und Methode*. Tübingen, Germany: J.C.B. Mohr, 1960. Translated by Joel Weinsheimer and Donald G. Marshall as *Truth and Method* (New York: Continuum, 1989).

Heidegger, Martin. *Sein und Zeit*. Halle: Niemeyer, 1927. Translated by Joan Stambaugh as *Being and Time* (Albany: SUNY Press, 1996).

Heidegger, Martin. *Brief über den Humanismus*. Bern, Switzerland: A. Franke, 1947. Translated as "Letter on 'Humanism'," in *Pathmarks* (Cambridge, U.K.: Cambridge University Press, 1998).

Heidegger, Martin. "Der Ursprung des Kunstwerkes." In *Holzwege*. Frankfurt, 1950. Translated by Julian Young and Kenneth Haynes as "The Origin of the Work of Art," in *Off the Beaten Track* (Cambridge, U.K.: Cambridge University Press, 2002).

Heidegger, Martin. *Unterwegs zur Sprache*. Pfullingen, Germany: Neske, 1959. Translated by Peter Hertz as *On the Way to Language* (New York: Harper and Row, 1971).

Merleau-Ponty, Maurice. *Phénoménologie de la perception*. Paris: Gallimard, 1945. Translated by C. Smith as *Phenomenology of Perception* (London: Routledge, 1962).

Merleau-Ponty, Maurice. "Le langage indirect et les voix du silence." In *Signes*. Paris: Northwestern University Press, 1960. Translated by R. C. McCleary as "Indirect Language and the Voices of Silence." In *Signs* (Evanston, IL: Northwestern University Press, 1964).

Robert Bernasconi (2005)

PHILOSOPHY OF LANGUAGE IN INDIA

The earliest Indian thinking about language, found in Vedas (Arapura and Raja 1990), is speculative, but later

discussions involve sophisticated arguments among various schools of thought. These discussions, which concern speech units (Sanskrit *śabda*, "sound, speech element, word") and associated meanings (*artha*), share certain themes. One is epistemological. Sounds are evanescent; an instant after they are pronounced they disappear. Consequently, the question arises: How can one rightly speak of complex units like words (*pada*) and sentences (*vākya*) as perceptible entities? Similarly, though one speaks of actions and things involved in them, it is also arguable that acts and things which are thought to be perceived as wholes actually are not so; there is a stream of instants, none of which lasts long enough to enable a qualified cognition of complex external entities. How, then, can one maintain that speech units signify actual actions and things? The second point concerns theory and procedure. Indian scholars operate with constructs in order to account for facts and behavior. This approach was evident already at an early period (ca. 7th c. BCE), when Vedic scholars posited constructed analyzed texts (*padapāṭha*) from which the Vedic texts as continuously recited (*saṁhitāṭha*) were derived by rules.

Indian thinkers accept certain means of acquiring knowledge, referred to as *pramāṇa* (a derivate of *pra-mā* [3rd sg. pres. *pramimīte*], apprehend"). At least two *pramāṇa*s are generally accepted: direct perception (*pratyakṣ*) and inference (*anumāna*). A third, verbal transmission (*śabda āgana*), is accepted by others, including Patañjali's yoga system. A means of knowing through similarity of one thing to another (*upamāna*) makes up a set of four *pramāṇa*s adopted by a major school of logicians, Nyāya. Not all thinkers, however, accept *śabda/āgama* as a separate *pramāṇa*; some account for knowledge acquired verbally through inference.

MEANINGFUL UNITS AND SYMBOLS

Systematic speech sounds—vowels (*a*, *ā*, *i*, *ī*, etc.) and consonants (*k*, *kh*, *g*, *gh*, *ṅ*, etc.) are distinguished from mere sounds (*dhvani*) such as the noise made by a drum. Classes of larger units are also recognized, the major ones being nominal forms (*nāman*), verbs (*ākhyāta*), preverbs (*upasarga*), and particles (*nipāta*); for example, *gauḥ* (nom. sg.), "cow, ox," *gacchati* (3rd sing. pres.), "goes, is going," *upa* in *upa gacchati*, "approaches," and *vā*, "or," respectively.

At an early stage, represented in pre-Pāṇinian texts and alluded to in later works like Patañjali's great mid-second-century BCE commentary (*Mahābhāṣya*) on Pāṇini's c. fifth-century BCE grammar, verbs and nouns were defined semantically. In one view, verbs signify varieties of being (*bhāva*): something comes into being (*jāyate*, "is born"), continues to be (*asti*, "is"), undergoes change while remaining the same entity (*viparaṇamate*, "changes"), increases (*vardhate*, "grows"), decreases (*apakṣīyate*, "diminishes"), and ceases to be (*vinaśyati*, "perishes"). Some scholars reduce these to three stages, with the second encompassing the third, fourth, and fifth. Alternatively, verbs are considered to signify particular actions (*kriyā*, *karman*), the most general action being signified by *kṛ*, "do." This definition is supported by usage: (1) *devadattaḥ pacati*, "Devadatta is cooking," is an appropriate answer to (2) *kiṁ karoti devadattaḥ*, "What is Devadatta doing?" These two views are superseded by considering that whatever a verbal base (*dhātu*) signifies—now spoken of as *kriyā* or *bhāva*—is conceived of as involving continuity in time, always associated with some time. As a consequence, not only terms such as *kṛ* (*karoti*), "do," *pac* (*pacati*), "cook," and *vraj* (*vrajati*), "go," but also ones like *as* (*asti*), "be," *ās* (*āste*), "be seated," and *sthā* (*tiṣṭhati*), "come to a stand, be in place," are now part of a single class of units signifying *kriyā/bhāva*. The canonical statement of this position, which can be seen already in the *Mahābhāṣya*, appears in Bhartṛhari's mid-fifth-century *Vākyapadīya* (3.8.1): whatever is always spoken of as something to be brought to accomplishment, whether it is already accomplished or not, is referred to as *kriyā* ("action") by virtue of its taking on a sequential status.

Contrasting with such semantic definitions, there is a formal approach, epitomized by the grammarian Pāṇini, who assigns to his class of units called *dhātu* verb bases listed in an appendix to the main corpus of rules and to items derived from both verbs and nominal forms (Cardona 1997).

There is a conception of units under which words are groups of sounds and larger units are groups of words. This view is represented in the section of Kauṭilya's Arthaśāstra (disputed date but not later than the third century CE) that deals with writing edicts (2.1.13–14) and in other works. It is already reflected in an argument Kātyāyana (third century BCE) mentions when he speaks of a group of sounds (*varṇasaṅghāta*) as being meaningful (*arthavat*). The same view is presupposed in the *Ṛgvedaprātiśākhya* (2.2), which describes the continuous recitation of a text (*saṁhitā*) as consisting in one's continuously putting together the last sounds of words (*padāntān*) with the initial sounds (*padādibhiḥ*) of following words, without any temporal separation (*kālāvyavadhānena*). Pāṇini himself (1.4.109: *paraḥ sannikarṣaḥ saṁhitā*) states that the maximum drawing together (*paraḥ sannikarśaḥ*) of sounds is called *saṁhitā*. One may

consider that such a procedure accepts that sounds do actually come together to form larger units, but it is also possible that this is an artifact necessary for the proper description by rules of what is found in a language.

From an argument presented by Kātyāyana and Patañjali, it is clear they were aware that one cannot speak of physical sounds truly co-occurring in immediate succession, for example, *gauḥ*, "cow" = *g-au-ḥ*. For speech does not produce two sounds simultaneously, since sounds (*varṇānām* [gen. pl.]) have the property of disappearing immediately upon being pronounced (*uccaritapradhvaṁsitvāt* [abl. sg.], Kātyāyana's vārttika 10 on 1.4.109: … *uccaritapradhvaṁsitvāca varṇānām*). As Patañjali explains: when *g* is pronounced, there is no *au* or *ḥ*; when *au* is pronounced, there is no *g* or *ḥ*; and when *ḥ* is pronounced, there is *g* or *au*.

If there is no physical composite unit such as *gauḥ*, the question arises: What is it that it is understood to signify? Two approaches were taken on this issue. One involves memory. It is assumed that when a sound is perceived, this experience leaves in one a lasting trace (*saṁskāra*, *vāsanā*); the last sound uttered in a given stretch produces a cognition accompanied by the traces left from preceding cognitions of sounds, and this final cognition is what produces an understanding of meanings of words and the sentences they make up. Alternatively, sounds are considered merely to manifest (*vyañj*) actual meaning bearers. These signifiers are posited elements called *sphoṭa*, distinct from physical sounds but manifested (*vyaṅgya*, "to be manifested") by them. Three major sphoṭa types are assumed: sentence (*vākyasphoṭa*), word (*padasphoṭa*), and subword meaningful elements (*varṇasphoṭa*) such as bases and affixes.

The first of these views was proposed at least by the time of the Mīmāṃsaka commentator Śabara (second century) and was accepted by adherents of different schools. The sphoṭa theory was first expounded fully by the grammarian-philosopher Bhartṛhari and remained basically the position of grammarians. Each of these positions was subjected to criticisms. Arguments against the first revolve about the nature of memory, what is recalled, and in what manner; the main argument against the sphoṭa position is that it requires positing units which one can do without.

WORD-MEANING RELATIONS

Speakers and hearers communicate and understand messages by means of words and sentences of a language they share. It is therefore universally accepted that a relation (*sambandha*) holds between words (*śabda*, *pada*) and meanings (*artha*) and that this relation can be a direct or indirect one, respectively called *śakti* ("capacity") or *abhidhā* ("signifying") and *lakṣaṇā* ("secondary meaning relation, metaphor"). A term that directly signifies (*vācaka*) a meaning is qualified as *śakta* ("capable") and its meaning as *śakya*, the object of this capacity. For example, *gaṅgā* directly refers to a flow of water, the river Gaṅgā. By *lakṣaṇā*, the same term can refer to the banks of the river. Thinkers of different schools engaged in arguments concerning both the nature of what is signified and the relations that link words and their meanings.

Concerning what words signify, at one extreme there is the view that terms like *ghaṭa*, "clay pot," *aśva*, "horse," *pac*, "cook, bake," refer to actual external entities, including actions one can witness. Other positions start from the observation that what one can actually perceive is not such an external thing (*vastu*) or action (*kriyā*): The latter is a stream of moments (*kṣaṇa*) that are beyond direct perception, and the former also can be broken down into such moments. The putative wholes treated as having identity are mental constructs (*vikalpa*).

One view consequent on this observation, adopted by certain Buddhistic thinkers, is that signification applies negatively, being a removal or differentiation (*apoha*) of all that is not the momentary entity in question, which is thus differentiated (*apoḍha*, "removed") from all others. The relation between a word—itself a construct—and its significand is then one of cause (*kāraṇa*) and effect (*kārya*): Words have mental constructs as their sources and bring about a comprehension of mental constructs. Although they accept that words and their meanings are related as signifier and significand (*vācyavācakabhāvasambandha*), Pāṇinian grammarians such as Bhartṛhari—with earlier precedents—also consider the cause and effect relation acceptable and conceive of the significands as word-meanings (*śabdārtha*) that are mental (*bauddha*) and not necessarily external objects (*vastu*).

In this connection, grammarians speak of a *vivakṣā*, a desire to speak about things in a particular manner. For example, it is in the nature of things that a sword (*asi*) serves as a means of cutting; one says, for example, (3) *devadattḥ asinā chinatti*, "Devadatta is cutting with a sword," using the instrumental *asinā* to refer to a sword as a means. If a sword is quite sharp, one may also appropriately say (4) *asiḥ sādhu chinatti*," the sword (*asiḥ*, nom. sg.) cuts well (*sādhu*)," speaking of a sword as an agent of cutting in the same way that (3) refers to Devadatta as an agent. In order to account for such usages, Pāṇini orders a group of rules that assign direct participants in the

accomplishment of actions (*kāraka*) to particular categories in such a manner that the participant in question is eligible to be assigned both to the category of participants called *karaṇa*, "instrument," by virtue of being the means (*sādhakatama*, "which is means more than any other") of accomplishing an act and, by a later rule, to the category of agents (*kartṛ*) by virtue of being an independent (*svatantra*) participant. Since a sword cannot be spoken of as an independent participant in the act of cutting without one's simultaneously considering it a means used by someone, this involves a conflict (*vipratiṣedha*), and Pāṇini provides explicitly that in case of such conflicts, what is provided for by a subsequently stated rule takes precedence over the provision of a preceding rule (see Cardona 1974). In connection with such situations, Patañjali notes that this involves what he calls *laukikī vivakṣā*, "communal desire to speak"; that is, it is not a matter of individual preference but of the way a community of speakers (*loka*, "world") expresses itself.

There is also the point of view that words have a natural relation of fitness (*yogyatā*) with their meanings, comparable to the fitness of different sense faculties with respect to what is perceived. Moreover, words and their meanings are commonly identified with each other.

However one conceives of the relation, each generation acquires a knowledge of words related to their meanings by observing how people interact. For example, a child witnesses an interaction between his father (F) and his grandfather (G): G says (5) *gām ānaya*, "bring the cow," to F, who then brings a cow, but F brings a horse when G says to him (6) *aśvam ānaya*. The child learns therefrom that *gām* and *aśvam* respectively designate a cow and a horse. This is an instance of reasoning from concurrent presence (*anvaya*) and absence (*vyatireka*): (a) $x \rightarrow y$, (b) $\bar{x} \rightarrow \bar{y}$. If both hold, then *x* which precedes *y* is its cause. Thus, if a given meaning is understood when a given term or member of a set of terms is used and not understood when this is not used, then the comprehension of the meaning in question is said to be caused by the term, to which this meaning is attributed.

Assuming that words designate positive significands, in ordinary usage one thinks of the term *go*, "cow," as referring to something that one can see and speak of repeatedly, using the same term. Moreover, in order to account for the repeated cognition of a cow each time one is seen, which can be verbalized saying (7) *iyaṁ gauḥ*, "this is a cow," it is assumed that each cow belongs to a class characterized by a class property (*jāti*, "generic property") that inheres in every member: *gotva* ("the property of being a cow").

If one assumes that a word-meaning relation is learned between an instance of the term *go* and a particular cow and also assumes that when another instance of *go* is used it too can refer to this particular cow, then the reasoning procedure shown above is violated, since one now has *y* in the absence of *x*. To assume that a separate relation is grasped between each instance of *go* and each individual (*vyakti*) cow has the consequence that no speaker can acquire the knowledge of such an infinite number of relations. Various solutions are proposed to remedy the situation (see Deshpande 1992 and Scharf 1996). One view, espoused by Mīmāṁsakas, is that the primary word-meaning relation is between a term and the class property (*jāti*). A sentence such as (5) is used, however, with the intention (*tātparya*) that someone bring a cow, not a class property. This is accounted for by assuming that in such an utterance *gām* signifies not only a class property, through a primary relation (*śakti*), but also a particular cow, through the secondary relation called *lakṣaṇā*. An alternative to this position is adopted by grammarians and logicians of the Nyāya school: A term like *go* signifies an individual (*vyakti*) qualified (*viśiṣṭa*) by its class property.

There are other instances where *lakṣaṇā* is said to operate. Consider (8) *kuntān praveśaya*, "have the javelins (*kuntān* [acc. pl.]) come in (*praveśaya*)." *Praveśaya* is a form (2nd sg. imper.) of a causative verb whose non-causal is *praviś* (3rd sg. pres. *praviśati*) "enter." Javelins cannot enter a room of themselves, so they cannot be caused to perform this act in the same way that one can cause people to enter a room. In order to make sense of the intent (*tātparya*) of a speaker who uses (8), it is accepted that *kunta* here bears a secondary relation with the men who bear javelins. In the same vein, consider (9) *gaṅgāyāṁ matsyāḥ*, "there are fish in the Gaṅga," and (10) *gaṅgāyāṁ ghoṣaḥ*, "there is a dairy colony on the Gaṅga." Assuming that *gaṅgāyām* (loc. sg. fem.) is used to refer to a locus in or on which something is located, (9) makes immediate sense, but (10) is hard to understand: fish can live in a river but a village of dairymen cannot be located physically in or on a body of flowing water. It is assumed, then, that in (10) *gaṅā*, which bears a primary word-meaning relation with a river, now bears a secondary relation with its bank (*tīra*).

(10) involves an assumed semantic incompatibility such that it is not possible for the primary meanings of *gaṅgāyām* and *ghoṣaḥ* to be related. However, it is not sufficient to say that what prompts one to understand a sec-

ondary meaning here is solely the impossibility of the referents being connected (*anvayānupapatti*). For this could be resolved also under the assumption that *ghoṣa* has a secondary relation with fish, so that (10) is understood to say what (9) says. Yet this is not the case: A person who hears (10) understands it to say that a dairy colony is located on the edge of the Gaṅga. Accordingly, the major reason prompting a secondary word-meaning relation is considered to be the impossibility of reconciling the primary meaning of a term with the intention (*tātparya*) of a speaker.

Understanding (8) and (10) in the way shown involves setting aside the primary meanings of *kunta* and *gaṅgā*. Consider now (11) *arko' staṁ gataḥ*, "the sun (*arkaḥ*) has set (*gataḥ*, "gone," *astam*, "home")." This can have its literal meaning. Without rejecting this meaning, moreover, there are several possible meanings that can be suggested (*vyaṅgya*, "to be made manifest"), depending on contexts and the persons uttering (11). For example, a go-between saying this to a woman who is to meet a lover suggests it is time to set out, but a servant saying this to a Brahmin means to imply that it is time for his master to perform the evening prayer. Another function of words is therefore considered, called *vyañjanā* (usually translated "suggestion"). This is principally accepted by theoreticians of poetics, though later Pāṇinian grammarians accept it, mainly because under the theory that a meaning bearer is a *sphoṭa*, which is manifested (*vyaṅgya*) by physical sounds.

SENTENCE AND SENTENCE MEANING

Adherents of various schools of thought in ancient and medieval India adopted different views concerning sentences and their meanings. One position—most systematically elaborated and defended first by Bhartṛhari—is that the true unit of communication is an atomic (*akhaṇḍa*) sentence (*vākya*), associated with an equally atomic sentential meaning, considered to be the object of a single flash of knowledge, hence referred to as *pratibhā* ("flash"). This thesis can be justified in so far as actual communication involves whole utterances, but it encounters the problems mentioned earlier in connection with words and their meanings: it is not possible for one to acquire a knowledge of all relations between all possible atomic sentences and their meanings. Moreover, a grammarian's aim is to give a generalized description of all possible sentences in terms of their structures, both formal and semantic, which is impossible if this thesis is taken strictly. Hence, Pāṇinians agree that at least one lower level—of words—must be accepted in terms of both language learning and description. They maintain, however, that words and their constituent bases, affixes, and so forth are constructs posited in order to account for whole utterances.

Under another view, held by some Mīmāṁsakas, there are no sentences qua distinct meaningful units. The sentential meaning of any stretch one calls a sentence is now accounted for indirectly, through the meanings of individual words. A parallel is drawn with the effect produced by utterances such as (12) *putras te jātaḥ*, "You've had a son" (*putraḥ* [nom. sg.], "son," *jātaḥ* [pptcple. nom. sg. m.], "born," *te* [dat. sg.], "to you") or (13) *garbhiṇī te duhitā*, "Your (*te* [gen. sg.]) daughter (*duhitā*) is pregnant (*garbhiṇī*)." Each of the words of these sentences signifies its particular meaning. These word meanings are then related to each other in accordance with the speaker's intention (*tātparya*) to convey a message and the hearer's semantic expectation (*ākāṅkṣā*) that each meaning has to be linked to other meanings of words in the utterances. The effects are happiness on the part of a man who learns he has had a son and sadness on the part of a man who learns his unmarried daughter is pregnant. Similarly, the words of all utterances denote only their individual meanings, which are then related to each other. An intermediate position is taken by logicians of the Nyāya system, who consider that the meaning of an utterance is apprehended through the intermediary of related words: The first word is first cognized as shown earlier, with the consequent memory of the word-meaning relation and a memory trace of the word and its meaning, then this process is repeated until, with the perception of the last word, a cumulative memory trace results of all the words and their meanings related to each other.

Whatever position one takes, two requirements apply to sentences. First, constituents must be in proximity (*āsatti*): each word following the first word of a sentence is uttered immediately after the preceding word, without the intervention of any term that is not syntactically related to the others. Secondly, there must be semantic expectancy (*ākāṅkṣā*), so that a hearer expects that the meaning signified by a word such as *gām* in (5) is connected with an action, since it contains an object-signifying suffix, and *ānaya* requires an object. As shown, the intention of a speaker (*tātparya*) also comes into play. Another requirement must be met if one is mainly interested in an utterance's serving as a means of conveying true knowledge: semantic compatibility (*yogyatā*, "the property of being connectible"). For example, each word of (14) *agninā puṣpāṇi siñcati*, "… is irrigating (*siñcati*) flowers (*puṣpāṇi* [acc. pl.]) with fire (*agninā* [instr. sg.]),"

conveys a meaning that is immediately understood. (14) cannot, however, convey a meaning acceptable in our world, where the act of irrigating requires a liquid (*dravadravya*) as a means. Accordingly, Naiyāyikas would deny that (14) has the status of *pramāṇa*. One might go so far as to deny that (14) produces a verbal cognition (*śābdabodha*). Against this, the following is pointed out. Upon hearing (14), a person would respond by asking how one can speak of irrigating with something that is not a liquid? The hearer has indeed related the meanings of the words in the well-formed utterance (14) according to their syntax, but the resulting sentence meaning is not acceptable in the world as we experience it.

Adherents of different schools differ also concerning the ways in which verbal cognitions (*śābdabodha*) are portrayed. Pāṇinian grammarians, logicians of the Nyāya school, and Mīmāṁsakas of the Bhāṭṭa school can agree that (15) *devadattaḥ kaṭaṁ karoti*, "Devadatta is making (*karoti* [3rd sg. pres.]), a mat (*kaṭam* [acc. sg.])," speaks of a given man making a mat. On the other hand, they give different paraphrases reflecting what they consider to be the *śābdabodha* prompted by this sentence, reflecting the preoccupations and theoretical premises of different schools of thought (see Cardona 1975 and Matilal 1985). Pāṇini accounts for the structure of Sanskrit through a set of derivational rules starting from semantics, and this is most efficently done under the assumption that the principal meaning of (15) is the action. Naiyāyikas are principally interested in the values of utterances as conveyors of valid knowledge, and within this system they operate with subjects and predications, so that the main qualificand in (15) is the person referred to by *devadattaḥ*. Mīmāṁsakas deal chiefly with the exegesis of Vedic utterances related to ritual performance, and in this context the principal meaning of an utterance is the act of bringing about a result.

CONCLUSION

These different interests and the fact that adherents of these systems and others either accepted the authority of Pāṇinian grammar or reacted to it led to ongoing arguments and counterarguments, with successive refinements over millennia, making India a center for the intense study of language and the philosophy of language.

See also Brahman; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Logic, History of: Logic and Inference in Indian Philosophy; Mind and Mental States in Indian Philosophy; Truth and Falsity in Indian Philosophy; Universal Properties in Indian Philosophy.

Bibliography

Aklujkar, Ashok. 1990. "Bhartṛhari: Trikāṇḍī or Vākyapadīya, with Vṛtti on Books 1 and 2." In *The Philosophy of the Grammarians*, edited by Harold G. Coward and K. Kunjunni Raja, 172–174. Princeton, NJ: Princeton University Press, 1990.

Arapura, John G., and K. Kunjunni Raja. "Philosophical Elements in Vedic Literature." In *The Philosophy of the Grammarians*, edited by Harold G. Coward and K. Kunjunni Raja, 101–106. Princeton, NJ: Princeton University Press, 1990.

Biardeau, Madeleine. *Théorie de la connaisance et philosophie de la parole dans le brahmanisme classique*. Paris: Mouton, 1964.

Cardona, George. "Pāṇini's *kārākas*: Agency, Animation and Identity." *Journal of Indian Philosophy* 2 (1974): 231–306.

Cardona, George. *Pāṇini: His Work and Its Traditions*. Vol. 1, *Background and Introduction*. 2nd ed. Delhi: Motilal Banarsidass, 1997.

Cardona, George. "Paraphrase and Sentence Analysis: Some Indian Views." *Journal of Indian Philosophy 3* (1975): 259–281.

Coward, Harold G., and K. Kunjunni Raja, eds. *The Philosophy of the Grammarians*. Vol. 5 of *Encyclopedia of Indian Philosophies*. Princeton, NJ: Princeton University Press, 1990.

Coward, Harold G., and K. Kunjunni Raja. "Introduction to the Philosophy of the Grammarians," In *The Philosophy of the Grammarians*, edited by Harold G. Coward and K. Kunjunni Raja, 1–97. Princeton, NJ: Princeton University Press, 1990.

Deshpande, Madhav M., trans. *The Meaning of Nouns: Semantic Theory in Classical and Medieval India = Nāmārthanirṇaya of Kauṇḍabhaṭa*. Translated and annotated. Dordrecht and Boston: Kluwer, 1992.

Ganguli, Hemanta Kumar. *Philosophy of Logical Construction: An Examination of Logical Atomism and Logical Positivism in the Light of the Philosophies of Bhartṛhari, Dharmakīrti and Prajñākaragupta*. Calcutta: Sanskrit Pustak Bhandar, 1963.

Matilal, Bimal K. *Logic, Language and Reality: An Introduction to Indian Philosophical Studies*. Delhi: Motilal Banarsidass, 1985.

Raja, K. Kunjunni. *Indian Theories of Meaning*. 2nd ed. Adyar, Madras: Adyar Library and Research Centre, 1969.

Ruegg, David Seyfort. *Contributions à l'histoire de la philosophie linguistique indienne*. Paris: Boccard, 1959.

Scharf, Peter M. *The Denotation of Generic Terms in Ancient Indian Philosophy: Grammar, Nyāya and Mīmāṁ*. Philadelphia: American Philosophical Society, 1996.

George Cardona (2005)

PHILOSOPHY OF LAW, HISTORY OF

The problems of authority, law and order, obligation, and self-interest first became central topics of speculation in the thought of the Sophists (late fifth and early fourth centuries BCE). The most famous Sophists all stressed the distinction between nature (*physis*) and convention (*nomos*), and they put laws in the latter category. They generally attributed law to human invention and justified obedience to law only to the extent that it promoted one's own advantage. Laws were artificial, arrived at by consent; the majority of acts that were just according to the law were contrary to nature; the advantages laid down by the law were chains upon nature, but those laid down by nature were free. In the time of the Sophists notions of law, justice, religion, custom, and morality were largely undifferentiated; yet in this same period some of the crucial problems of legal philosophy were first formulated, and attempts were made at a formal definition of law. Thus, Xenophon (*Memorabilia* I, 2) reported that Alcibiades, who associated with both Critias and Socrates, remarked to Pericles that no one can really deserve praise unless he knows what a law is. Pericles replied that laws are what is approved and enacted by the majority in assembly, whereby they declare what ought and what ought not to be done. He admitted that if obedience is obtained by mere compulsion, it is force and not law, even though the law was enacted by the sovereign power in the state. Xenophon also reported an alleged conversation between Socrates and the Sophist Hippias in which both maintained an identity between law, or what is lawful, and justice, or what is right, while admitting that laws may be changed or annulled (ibid. IV, 4). Socrates claimed that there are "unwritten laws," uniformly observed in every country, which cannot conceivably be products of human invention. They are made by the gods for all men, and when men transgress them, nature penalizes the breach.

Socrates and the Sophists, as presented in Plato's dialogues, disagreed concerning human nature. The Sophists conceived of man as egoistically motivated and antisocial, whereas for Socrates, as for Plato and Aristotle, man was a social being with other-regarding as well as self-regarding motives, who finds fulfillment in social life. By contrast, the Sophist Callicles, in Plato's *Gorgias*, holds that man is no exception to the law of nature, according to which the stronger rules; manmade laws and social institutions violate human nature. The less radical Sophists, although they could not identify law with some feature of reality, still accepted its practical usefulness.

PLATO AND ARISTOTLE

PLATO. There is hardly any problem of legal philosophy not touched upon by Plato. He wrote during the decline of the Greek polis, when law and morality could appear as mere conventions imposed by shifting majorities in their own interest and the harmony between the legal order and the order of the universe could not easily be maintained. Plato sought to restore, as far as possible, the traditional analogy between justice and the ordered cosmos. Justice, or right action, cannot be identified with mere obedience to laws, nor can a truly moral life be reduced to conformity with a conventional catalog of duties. Duties involve a knowledge of what is good for man, and this bears an intimate relation to human nature. The question "What is justice?" dominates Plato's *Republic*. Plato conceived of justice as that trait of human character which coordinates and limits to their proper spheres the various elements of the human psyche, in order to permit the whole man to function well. In order to understand the operation of justice in the human soul, Plato examined human nature writ large, the city-state. The state functions well when it is governed by those who know the art of government, and the practice of this art requires a positive insight into the Good. In a just society every citizen performs the role of which he is best capable for the good of the whole. Similarly, in the moral economy of the individual's life, justice prevails when reason rules and the appetites and lower passions are relegated to their proper spheres. A just social order is achieved to the extent to which reason and rational principles govern the lives of its members.

Plato's emphasis on reason found its way into his definition of law. Law is reasoned thought (*logismos*) embodied in the decrees of the state (*Laws* 644D). Plato rejected the view that the authority of law rests on the mere will of the governing power. The *Laws* contains a detailed discussion of many branches of law and is an attempt at a formulation of a systematic code to govern the whole of social life. In contrast with the ideal polis of the *Republic*, in which there would be little need for legislation, in the *Laws* Plato accepted "law and order, which are second best" (*Laws* 875D).

ARISTOTLE. Aristotle, who discussed law in numerous contexts, nowhere gave a formal definition of it. He wrote variously that law is "a sort of order, and good law is good order" (*Politics* 1326a), "reason unaffected by desire"

(ibid. 1287a), and "the mean" (ibid. 1287b). However, these must be taken not as definitions but as characterizations of law motivated by the point Aristotle was making in the given context.

Following Plato, Aristotle rejected the Sophistic view that law is mere convention. In a genuine community—as distinguished from an alliance, in which law is only a covenant—the law concerns itself with the moral virtue of the citizenry (*Politics* 1280b). Aristotle sharply distinguished between the constitution (*politeia*) and laws (*nomoi*); the constitution concerns the organization of offices within the state, whereas the laws are "those according to which the officers should administer the state, and proceed against offenders" (ibid. 1289a). The constitution of a state may tend to democracy, although the laws are administered in an oligarchical spirit and vice versa (ibid. 1292b). Legislation should aim at the common good of the citizens, and justice—what is equal—should be determined by the standard of the common good (ibid. 1283a). Yet Aristotle recognized that the law is often the expression of the will of a particular class, and he stressed the role of the middle class as a stabilizing factor.

In his discussion of the forms of government in Book III of the *Politics*, Aristotle took up the Platonic problem of rule by the best man versus rule according to laws. A society of equals by its very nature excludes the arbitrary rule of one man. In any case, even the best man cannot dispense with the general principles contained in laws; and legal training helps to make better officers of government. Furthermore, administrators, like all men, are subject to passion, and it is thus preferable to be judged by the impersonal yardstick of the laws. This in no way conflicts with the need to change the law through legislation when it has been found by experience to be socially inadequate. But not all law is the product of legislation; customary law is in fact more important than the written law.

Aristotle's discussion of the judicial process foreshadows many modern notions. Although it is better to have written laws than to rely completely on discretion, "some matters can be covered by the laws and others cannot" (ibid. 1287b20). General rules are insufficient to decide particular cases (ibid. 1286a26), although "well-drawn laws should themselves define all the points they possibly can and leave as few as may be to the decision of the judges" (*Rhetoric* 1354a32). Aristotle seems to have had two considerations in mind. First, judicial decision making is practical—it involves deliberation—and as such cannot be completely determined in advance. Second, the resolution of disputed issues of fact in a particular case, on which the decision depends, cannot be settled in advance by legislation. This stress on the insufficiency of general rules connects with Aristotle's influential discussion of equity (*epieikeia*). Equity is just, "but not legally just but a correction of legal justice" (*Nicomachean Ethics* 1137b10). Aristotle sometimes seems to suggest that equity comes into play when there are gaps in the law, so that it consists in the judge's acting as the lawgiver would act if he were present. Yet he also seems to suggest that equity corrects the harshness of the law when adherence to the written law would work an injustice. Principles of equity are thus closely related to the unwritten universal laws "based on nature," a "natural justice" binding on all men, even those who have no association or covenant with each other. Nevertheless, what is naturally just may vary from society to society.

The locus classicus of Aristotle's discussion of justice is Book V of the *Nicomachean Ethics*. Generically, justice has to do with one's relations to others, and there is a sense of "justice" that refers to the complete moral virtue of the member of the community in such dealings. There is also a sense in which "justice" refers to a particular virtue involving the fair dealings of individuals in matters handled by private law. Two kinds of rights fall under this special virtue: rights in division (where each individual claims his fair share of goods, honors, and so on) and rights in redress (for wrongs done by one individual to another, such as failure to fulfill a contract).

ROME

STOICS. The Stoics, who conceived of the universe as a single, organic substance, exercised a lasting influence on legal thought. Nature, which exhibits structure and order, and man both partake of intelligence, or reason (*logos*). An animal is directed by a primary impulse toward self-preservation that adapts it to its environment. In man, reason is the "engineer of impulse," and man's actions may be evaluated only within the framework of the whole of nature. The criterion of moral action is consistency with the all-determining law of nature (*koinos logos*). This conception of a law of nature that is the ultimate standard of human laws and institutions was combined with Aristotelian and Christian notions to form the long-standing natural-law tradition of medieval legal philosophy. Another important Stoic contribution was the belief in the equality of all men in a universal commonwealth and a rejection of Aristotle's doctrine of slavery.

CICERO AND SENECA. The writings of Marcus Tullius Cicero (106–43 BCE) were important in transmitting classical legal thought to the medieval world. Although he

was a professional arguer of legal cases, Cicero's philosophical treatment of law in his *De Legibus* disclaims any interest in "clients' questions" or the "law of eaves and house-walls." His legal philosophy was essentially Stoic; he denied that the positive law of a community (written or customary), even when universally accepted, is the standard of what is just. Nor is mere utility the standard: "Justice is one; it binds all human society, and is based on one law, which is right reason applied to command and prohibition" (*De Legibus* I, 15). An unjust statute is not a true law. Law and morality are logically connected, and only that which conforms to the law of nature is genuine law. This view exercised a lasting influence on natural-law thinking and reappeared in the thought of Thomas Aquinas.

Like Cicero, Lucius Annaeus Seneca (c. 4 BCE–65 CE) aided in transmitting Stoic notions to later thinkers. He reiterated the conception of the equality of all men under natural law, but perhaps more important was his conception of a golden age of human innocence, a prepolitical state of nature. Legal institutions became necessary as human nature became corrupted.

ROMAN LAW. The influence of Stoicism may be traced in pronouncements of the Roman jurists. It is disputed whether these were any more than remarks designed to ornament legal texts, but they nevertheless influenced the thought of later ages. The jurists distinguished three kinds of law: *jus naturale, jus gentium*, and *jus civile*. In practice, the last originally referred to the law of the city of Rome, but ultimately it was applied to any body of laws of a given community. The *jus gentium* first meant the law applied to strangers, to whom the *jus civile* was not applicable, and was later extended to those legal practices common to all societies. Gaius (mid-second century), who systematized the Roman law in his *Institutes*, identified the *jus naturale* and *jus gentium* as universal principles of law agreeable to natural reason and equity. Thus, law was not a mere expression of human will or institution but that which is rationally apprehended and obeyed. The *jus gentium* was not an ideal law by which the positive law was judged but the rational core of existing legal institutions.

Ulpian (c. 170–228) distinguished *jus naturale* from *jus gentium* by stating that *jus naturale* is not peculiar to human beings but is taught by nature to all animals. Thus, among animals there is an institution similar to human marriage. Slavery and its attendant rules are products of the *jus gentium*, for by the *jus naturale* all men were born free. It is not clear, however, that Ulpian regarded slavery as bad. To him we owe the oft-repeated definition of justice: "the constant wish to give each his due" (*Digest* I, 1, 10). Following Celsus (c. 67–c. 130), he defined law (*jus*) as "the art of the good and the equitable" (ibid. I, 1, 1). Again, it does not seem that Ulpian thought of the *jus naturale* as an ideal law opposed to the *jus civile* or to the *jus gentium*. It has been suggested that behind Ulpian's thought was a conception of a natural state antecedent to the conditions of organized society.

The doctrines of the Roman jurists owe their lasting influence to their incorporation into the *Corpus Juris Civilis* of Justinian (sixth century), principally in the section called the *Digest*. The compilers of Justinian's *Institutes* (a section of the *Corpus Juris*) seem to have distinguished the *jus naturale* from the *jus gentium* and seem to have regarded the former as a set of immutable divine laws by which the positive law may be morally evaluated (*Institutes* I, 2, 11; III, 1, 11). The *Corpus Juris* also preserved statements of the Roman jurists concerning the source of the authority to make and unmake the laws constituting the civil law. According to a number of these statements, this authority resides in the consent of the people; however, the statement that "what pleases the prince has the force of law" (*Digest* I, 4, 1) was probably a more accurate view of the facts. Justinian seems to have combined these views theoretically in his reference to a (nonexistent) "ancient law" by which the Roman people transferred all their powers to the emperor (*Codex* I, 17, 1, 7).

EARLY MIDDLE AGES

To the legal thought of the Stoics and the Roman philosophers and jurists the Church Fathers added a distinctively Christian element. The law of nature was no longer the impersonal rationality of the universe but was integrated into a theology of a personal, creative deity. The relationship among the Mosaic law, the Gospels, and natural law emerged as a specific problem; the notion of *jus divinum* (divine law) as a distinct type of law, along with the three recognized by the jurists, was crystallized. The notion of the fall of man from a state of perfection (which may be compared with the view of Seneca) played an important role. Thus, according to St. Ambrose (340–397) the Mosaic law—a law of sin and death (see Romans 8:2)—was given because man failed to obey the law of nature. The fact that many legal institutions, such as slavery and private property, deviate from this ideal law does not necessarily imply that they are unjust or illegitimate; for the natural law is adapted to man only in a condition of innocence.

Of the Church Fathers, St. Augustine (354–430) was perhaps the most original and complex: Only one point

in his thought will be noted here. Cicero maintained that nothing can be nobler than the law of a state (*De Legibus* I, 14) and that if a state has no law, it cannot truly be considered a state (ibid. II, 12). The law of the state must therefore embody justice, for without *justitia* there is no *jus*. Augustine considered this position in *The City of God*, Book XIX. According to Augustine, since Rome had no justice, Cicero's position has the inconvenient consequence that Rome was no state at all. We must therefore seek another definition of "state" (*populus*) in which justice is not an essential element. Augustine stressed the notion of order—"a harmonious multitude"—with the suggestion that legal order need not be moral or just. There are passages in Augustine, however, which seem to uphold a more orthodox natural-law position. In any event the terms of his discussions are somewhat different; his main points of contrast are divine and human law, rather than *jus naturale* and *jus civile*.

The sources of the natural-law theories that were to dominate Western legal philosophy for many centuries were the writings of the Greek and Roman philosophers and poets, Justinian's *Corpus Juris Civilis*, and the Church Fathers. Isidore of Seville (c. 560–636), an encyclopedist and an important transmitter of Roman thought to later writers, concisely expressed the natural-lawyer's ideal regarding positive law: "Law shall be virtuous, just, possible to nature, according to the custom of the country, suitable to place and time, necessary, useful; clearly expressed, lest by its obscurity it lead to misunderstanding; framed for no private benefit, but for the common good" (*Etymologies* V, 21).

MIDDLE AGES AND RENAISSANCE

CIVILIANS AND CANONISTS. In the revived study of Roman law in the twelfth century, associated with the glossators, legal philosophy received a fresh stimulus. Of special interest are the attempts at reconciling differences among the Roman jurists on the definition of law and the classification of its branches. In the main, the civilians were in the broad tradition of natural-law thinking; *jus* flows from *justitia*, although it must always fall short of perfect justice, which is God's alone. Irnerius (c. 1050–c. 1130) thus claimed that statutes ought to be interpreted in the light of equity. Strict law requires that all agreements be kept, but equity allows exceptions to the rule. This equity, according to Azo (c. 1150–c. 1230), must be written, rather than a principle found in the judge's heart.

The middle of the twelfth century also saw the systematization of the canon law. In the *Decretum* of Gratian a high degree of jurisprudential competence was brought to this task. The tripartite division of law of the Roman lawyers was verbally accepted, but the leading conceptions were Augustine's *jus divinum* and *jus humana*. Natural law was identified with the former, while the distinctive feature of the latter (covering both *jus gentium* and *jus civile*) was custom. Natural law is contained in the Mosaic law and the Gospels; the command to do unto others what we would have them do unto us is its fundamental principle. Natural law relates to man's rational nature and is immutable; the *mistica*, the cultic regulations found in Scripture, are part of the natural law only in their moral aspect. The commentators on Gratian further divided natural law so as to include not only commands and prohibitions but also *demonstrationes*, which point to what is good for humankind, such as possession of all things in common. In man's fallen condition custom has legitimately modified the *demonstrationes* in permitting private property and slavery. The other branches of natural law may not be abrogated and are the standards by which even the ecclesiastical law must be judged. Gratian (if not all his commentators) seems to have generally maintained a clear distinction between natural (divine) law and canon law.

AQUINAS. The rediscovery of Aristotle in the thirteenth century greatly influenced the further development of legal philosophy. The culmination of the natural-law tradition is the theory of Thomas Aquinas (c. 1224–1274), who integrated Stoic, Christian, and Aristotelian elements within a comprehensive philosophic system. Laws are standards of conduct that have a binding, or obligatory, character. This can be understood only if laws have some kind of rational origin. Combining this view with a teleological conception of nature and social order, Aquinas regarded legal control as purposive. Laws, he concluded, are ordinances of reason promulgated for the common good by the legitimate sovereign. Four types of law may be distinguished: eternal law, an expression of God's rational ordering of the universe; divine law, which guides man toward his supernatural end; natural law, which guides man toward his natural end; and human law, which regulates through the prospect of punishment the affairs of men in a given community in the light of that community's special requirements. Crucial to the concept of natural law are the notions of natural inclinations and right reason. "All those things to which man has a natural inclination are naturally apprehended by reason as being good and consequently as objects of pursuit, and their contraries as evil, and objects of avoidance" (*Summa Theologiae* I–II, 94). The relationship between inclination and reason, accounting for the apprehension of the natu-

ral law, has been variously interpreted. The precepts of natural law have as their common foundation the principle "Do good and avoid evil." Natural law is a standard to which human law must conform, and Aquinas employed Aristotle's conception of practical reasoning in explaining the derivation of human law from natural law by the legislator, thus accounting for differences between legal systems and for the possibility that rational men should disagree as to what human laws ought to be. He affirmed the long-standing view that an unjust law is no law; but although an unjust law is not binding in conscience, considerations of utility may require one to obey it. Aquinas allowed that such "laws" may be said to possess a "legal" character insofar as they are promulgated under the color of law by the legitimate prince.

Aquinas discussed in detail and with great acuity all of the problems treated by his predecessors. His influence may be traced in the English writers John Fortescue (c. 1394–c. 1476), Thomas Hooker (c. 1586–1647), and Christopher St. Germain (1460–1540). According to St. Germain, natural law is nothing other than the common-lawyer's notion of "reasonableness." More recent Thomist thinkers, such as François Gény (1861–1959) and Jean Dabin, have advanced novel ideals within the Thomistic tradition.

OCKHAM. Some medieval writers seem to have espoused a protopositivism in their emphasis on the primacy of the will; this is characteristic of the Augustinian-Franciscan tradition. Thus, William of Ockham (c. 1285–1349) regarded the divine will as the norm of morality. "By the very fact that God wills something it is right for it to be done." Nevertheless, it is doubtful that Ockham would have affirmed that what the sovereign commands is just. His position is somewhat unclear, however, for he—like all medieval writers—continued to use the rhetoric of natural law in his *Dialogus*: In one of its senses *jus naturale* is composed of universal rules of conduct dictated by natural reason. A right, such as the immutable right of private property, is a dictate of right reason.

RISE OF ABSOLUTISM. A tendency to combine natural-law doctrines with a theory of royal absolutism began in the fourteenth century. A group of civilians, known as the postglossators, undertook to forge a workable system of law out of the older Roman law, which they regarded as the *jus commune* of Europe. The technically trained administrators in the rising nation-states, they were naturally concerned with fundamental problems of legal theory. Bartolus of Sassoferrato (1314–1357) maintained that the ruler is not bound by the laws, although it is "equitable" that he should voluntarily submit to them. The *jus gentium*, however, is immutable. Lucas de Penna (1320–1390) discussed jurisprudential questions in detail. Law is the articulation of the ethical virtue of justice, and reason is the foundation of law. At the same time he maintained, as did many civilians, that the prince's lordship rests on divine authority. The ruler is responsible to God alone and not to the people; law is not the expression of the will of the community. Nonetheless, although the prince is unfettered by the laws, bad laws (those that contradict divine law) have no binding force. It is not clear, in Lucas's view, whether the obligation to obey law derives primarily from the rationality of law or from the divine grant of authority to the ruler.

LATER RENAISSANCE

BODIN. Jean Bodin (1530–1596), the great exponent of unlimited sovereignty under natural law whose views were apparently influenced by the fourteenth-century civilians, like them appears to have had difficulty in adapting Christian legal thought to the conditions of the secular nation-state. In his *Six Books of the Commonwealth* Bodin was emphatic that "law is nothing else than the command of the sovereign in his exercise of sovereign power." But although the prince "has no power to exceed the law of nature," which is decreed by God, it seems plain that Bodin no longer thought of right reason as linking natural and positive law. Bodin's endorsement of the command theory also appears in his treatment of custom. The relative weights of positive law and custom had long been debated by the medieval lawyers, but Bodin was one of the first to hold that custom owes its legal authority to the sufferance of the ruler. In this he anticipated the idea of tacit command expressed by Thomas Hobbes and John Austin.

INTERNATIONAL LAW. The emergence of nation-states also brought the problem of the rational foundation of international law to the forefront of legal thinking. This development may be seen in the writings of the Spanish Thomists Francisco de Vitoria (1492/1493–1546) and Francisco Suárez (1548–1617) and of Hugo Grotius (1583–1645), a Dutch Protestant jurist with broad humanistic leanings. According to Vitoria, the *jus gentium* either belongs to or is derivable from the natural law and consists in prescriptions for the common good in the widest sense, namely, for the international community. Rights and obligations are thus conferred upon nations acting through their rulers.

The conception of a law of nations was developed in great detail by Suárez. Although his *De Legibus* is Thomistic in many respects, Suárez explicitly stated that Aquinas's account of law is inadequate. Suárez began by distinguishing laws in the prescriptive sense from laws of nature in the descriptive sense, which are laws only metaphorically. (Many positivists trace the origin of natural-law thinking to the tendency to confuse these two types of law.) With regard to prescriptive laws, Suárez defined a law (*lex*) as "the act of a just and right will by which the superior wills to oblige the inferior to this or that" or as "a common, just and stable precept, which has been sufficiently promulgated" (*De Legibus* I, 12). The reference to stability is notable: Laws generally survive both the lawgiver and the populace living when they are enacted, and they are valid until abrogated. Such considerations have led recent writers to reject the identification of laws with mere acts of will; but although Suárez rejected the voluntaristic notion of natural law associated with the Ockhamists, he held that the civil law is enacted "more by the will than by reason." It is not derived from natural law by logical inference but by "determination," and hence is, in a sense, arbitrary (ibid. II, 20). Most medieval writers tended to use *lex* and *jus* interchangeably; Suárez, however, defined the latter as "a certain moral power which every man has, either over his own property or with respect to what is due to him" (ibid. I, 2). Although Aquinas briefly discussed *jus naturale* as contrasted with *jus positivum (Summa Theologiae* II–II, 57), the concept of a "natural right" was almost entirely absent from his thought. It is clearly present in Suárez, who, in the style of John Locke (1632–1704) and the Enlightenment philosophers, formulated a list of natural rights. Nevertheless, the individualism of these writers is not present in Suárez. His attitude was quite remote from eighteenth-century natural-law and natural-right theorists, who thought that a perfect system of law could be deduced from the natural law.

Despite Grotius's tendency to underestimate his predecessors, his *De Jure Belli ac Pacis* (1625) clearly showed the influence of such writers as Vitoria and Suárez. He developed their notion of a "just war," a topic that was still discussed by Hans Kelsen (1881–1973) and other twentieth-century theorists concerned with the problem of sanctions in international law. Just wars presuppose the existence of laws governing relations between sovereign states; such laws have their origin in natural law and in treaties, which in turn presuppose precepts of the law of nature. The denial of the existence of natural law supposes that men are egoistically motivated, accepting law as a "second best." However, following Aristotle and the Scholastics, Grotius held that man is social, altruistic, and rational. Therein lies the origin of law, which would be binding whether or not God exists. This statement has been regarded by historians as epoch making; they claim that Grotius separated jurisprudence from theology. More important, perhaps, is the tendency in Grotius and others who followed him to identify natural law with certain rational principles of social organization, and thus to loosen its tie with the Stoic metaphysical conception of the law of nature.

SEVENTEENTH TO LATE NINETEENTH CENTURIES

HOBBES AND MONTESQUIEU. Thomas Hobbes (1588–1679) was perhaps the most important of the seventeenth-century legal philosophers. His break with the tradition of natural law provoked much controversy. Hobbes employed the terminology of "natural right," "laws of nature," and "right reason." But the first was for him simply "the liberty each man hath to use his own power as he will himself, for the preservation of his own nature; that is to say, of his own life" (*Leviathan* 14); the second are principles of self-interest, which are often identified with the third. There is no right reason in nature (*Elements of Law* II, 10, 8). The natural condition of humankind is one of perpetual war, in which common standards of conduct are absent. There is no right or wrong, justice or injustice, mine or thine in this situation. The crucial steps in Hobbes's theory are the identifications of society with politically organized society and of justice with positive law. Laws are the commands of the sovereign; it is in reference to such commands that the members of a society evaluate the rightness or justness of their behavior. An "unjust law" is an absurdity; nor can there be legal limitations on the exercise of sovereign power. No writer has put forward a positivistic conception of law with greater style and forcefulness than Hobbes. Difficulties in his position emerge from his concession that although the sovereign cannot commit an injustice, he may commit iniquity; the idea of injury to God in the state of nature; and the treatment of conscience in *De Cive*. Hobbes solved the problem of the source of the obligation to obey the sovereign's command by his "social contract" doctrine, the interpretation of which is still discussed by scholars. His unfinished *Dialogue between a Philosopher and a Student of the Common Laws of England* examines various doctrines of the English law as put forward by Sir Edward Coke, and it is notable for its critical examination of Coke's statement that reason is the life of the law.

The *Second Treatise of Civil Government* by Locke, primarily an attack on Robert Filmer's "divine right" theory, contains certain implied criticisms of Hobbes. Its interest for legal philosophy lies in its use of a version of the social contract to treat the question of the obligation to obey the law, its conception of limitations on sovereign power, and its individualistic view of natural inalienable rights, particularly rights in property. Locke's influence was enormous, and his view of natural rights had a profound effect on the development of law in the United States.

A new approach to the understanding of law and its institutions was put forward by Baron de Montesquieu (1689–1755). He, too, spoke the language of natural law and defined laws as "necessary relations arising from the nature of things" (*The Spirit of the Laws* I, 1). But his special importance lies in his attempt to study legal institutions by a comparative historical method, stressing the environmental factors that affect the development of law. This suggestion had been anticipated by Bodin, and Giambattista Vico (1668–1744) had also applied a historical method to the study of Roman law, but Vico's work had little immediate influence. Montesquieu's doctrine of the separation of powers had an extraordinary influence. His sharp separation of judicial from legislative and executive power reinforced the conception that the judge is a mere mouthpiece of the law and that judges merely declare the existing law but never make it. In 1790, in his *Reflections on the Revolution in France*, Edmund Burke turned the historical approach to a practical political use when he protested against proceeding a priori in the "science of constructing a commonwealth."

KANTIANISM. Immanuel Kant (1724–1804) contributed to legal philosophy as he did to other branches of philosophy. The keynote of his legal philosophy was inspired by Jean-Jacques Rousseau (1712–1778), who set as the problem of his *Social Contract* the reconciliation of social coercion and individual freedom. Kant's legal philosophy may be called a philosophy of justice in which the concept of freedom plays a central role. Kant sought a systematic understanding of the principles underlying all positive laws that would enable us to decide whether these laws are in accordance with moral principles. Positive law "proceeds from the will of a legislator," and any viable legal system will take into account the particular conditions of the given society. With these conditions the theory of law has no concern. The theory is an application of the results of moral philosophy to the conditions of "men considered merely as men." This endeavor covers both the domain of law (*Recht*) and the domain of ethics; the principle that right action is action in conformity with universalizable maxims holds for both juridical and moral laws. A law (*Gesetz*) is a formula expressing "the necessity" of an action. Juridical and moral laws are distinguished in that the former regulate external conduct irrespective of its motives. (But this does not mean that a judge should necessarily ignore the lawbreaker's motives when passing sentence upon him.) Any man, as a morally free agent, is entitled to express his freedom in activity so long as it does not interfere with the similar freedom that others possess. This is the principle underlying all legislation and "right." Juridical law also involves the authority to compel conformity and to punish violations. The necessary and sufficient condition for legal punishment is that the juridical law has been broken. It must be recognized, however, that the domain of such law is restricted by the limits of compulsion. While it is morally wrong to save one's own life by killing another, even where this is the only expedient, it can never be made legally wrong to kill in such a case. The principle of law receives content in Kant's application of it to particular private rights in external things and in his analysis of the methods for acquiring such rights.

Kant's influence on jurisprudence, after being somewhat eclipsed by Hegelianism, reemerged at the end of the nineteenth century. One of the most important neo-Kantians was Rudolf Stammler (1856–1938), who invented, but eventually discarded, the phrase "natural law with variable content." Accepting the Kantian distinction between "form" and "matter," he attempted to discern the form of all laws. He defined law as "exceptionless binding volition." Just law is an ideal involving principles of respect and cooperation.

UTILITARIANISM AND POSITIVISM. While Kant and his followers may be said to have fostered a variety of natural-law thinking (although different from the Stoic and Thomistic types), Jeremy Bentham (1748–1832) and his followers (notably John Stuart Mill) claim to have rejected such thinking entirely. Of the influences on Bentham, two may be briefly noted. David Hume (1711–1776) argued that moral distinctions are not derived from reason; passion, or sentiment, is the ultimate foundation of moral judgment. Justice is grounded in utility. Second, the Italian criminologist Cesare Beccaria (1738–1794), in his *Of Crimes and Punishments* (1764), subjected the existing institutions of criminal law and methods of punishment to relentless criticism. His standard of judgment was whether "the greatest happiness of the greatest number" was maximized. Bentham acknowledged his debt to Beccaria, and this "principle of

utility" was the base of Bentham's voluminous projected "codes." He did not, however, define the nature of law by reference to utility. In his *The Limits of Jurisprudence Defined* (published in 1945) he defined a law as the expression of "the will of a sovereign in a state." Bentham's views, which were well suited to deal with the problems engendered by the industrial revolution in England, were of immense importance in effecting legal reform. In 1832, the year of his death, the Reform Act was passed, largely as a result of the work of his followers. Mill's *On Liberty* (1859) is an attempt to treat the limits of legal coercion by the state along modified utilitarian lines.

In legal philosophy Bentham's influence affected the English-speaking world especially through the thought of John Austin (1790–1859), the seminal figure in English and American legal positivism and analytic jurisprudence. Austin tried to find a clear demarcation of the boundaries of positive law, which would be antecedent to a "general jurisprudence" comprising the analyses of such "principles, notions, and distinctions" as duty, right, and punishment, which are found in every legal system; these analyses in turn were to be employed in "particular jurisprudence," the systematic exposition of some given body of law. Austin began by distinguishing "law properly so called" and "law improperly so called." The former is always "a species of command," an expression of a wish or desire, analytically connected with the ideas of duty, liability to punishment (or sanction), and superiority. The last notion led Austin to his famous and influential analysis of "sovereignty"; "laws strictly so called" (positive laws) are the commands of political superiors to political inferiors. From this it follows that international law is merely "positive international morality" rather than law in a strict sense. (Some writers, viewing this as an unfortunate and perhaps dangerous consequence, were led to various revisions of Austinianism.) Austin's "separation" of law and morality is often taken as the hallmark of legal positivism. "The existence of law is one thing; its merit or demerit is another," he wrote in *The Province of Jurisprudence Determined* (V, note). Yet Austin was a utilitarian; in distinguishing between the law that is and the law that ought to be, he did not mean that law is not subject to rational moral criticism grounded in utility, which he took to be the index to the law of God. At this point Austin was influenced by such "theological utilitarians" as William Paley.

Austin's views were subjected to vigorous discussion both without and within the traditions of positivism and analytical jurisprudence. And as the disciplines of history, anthropology, and ethnology assumed an increasing importance during the nineteenth century, rival approaches to the understanding of law developed. Thus, Sir Henry Maine (1822–1888), who formulated the historical law that legal development is a movement from status to contract, argued in his *Early History of Institutions* (London, 1875) that the command-sovereignty theory of law has no application in a primitive community, where law is largely customary and the political "sovereign," who has the power of life or death over his subjects, never makes law. The Austinian view can be saved only by maintaining the fiction that what the "sovereign" permits, he commands. Nonetheless, Austin had many followers at the turn of the twentieth century, such as T. E. Holland (1835–1926) and J. W. Salmond (1862–1924), who attempted to preserve the imperative and coercion aspects of his theory while introducing revisions.

The role of the courts was increasingly emphasized. In the United States, John Chipman Gray (1839–1915) wrote *The Nature and Sources of the Law* (New York, 1909; 2nd ed., New York, 1921), one of the most important American contributions to the subject. Acknowledging his debt to Austin, Gray defined law as "the rules which the courts [of the State] lay down for the determination of legal rights and duties." This required him to construe statutes, judicial precedents, custom, expert opinion, and morality as sources of law rather than as law. All law is judge-made. The machinery of the state stands in the background and provides the coercive element, which does not enter into the definition of "law." Gray's influence may be traced in the realist movement in the United States.

HEGELIANISM AND THE HISTORICAL SCHOOL. While England was largely under the sway of the utilitarians, Kantianism, Hegelianism, the historical school, and legal positivism flourished in Germany, both singly and in various combinations. In his *Philosophy of Right*, G. W. F. Hegel (1770–1831) developed some Kantian themes in his own characteristic way. Law and social-political institutions belong to the realm of "objective spirit," in which interpersonal relationships, reflecting an underlying freedom, receive their concrete manifestations. In attempting to show the rightness and the rationality of various legal relationships and institutions in given moments of the development of "spirit," and in seeing them as natural growths, Hegel formulated a theory of law and the state that was easily combined with various historical, functional, and institutional approaches to legal phenomena.

Friedrich Karl von Savigny (1779–1861) is often regarded as the founder of the historical school. His *Of the Vocation of Our Age for Legislation and Jurisprudence* (1814) was published before Hegel's work and was probably influenced by Johann Gottlieb Fichte (but not by Fichte's *Grundlage des Naturrechts*, 1796), whose notion of the "folk-spirit" was widely known. Law, like language, originates spontaneously in the common consciousness of a people, who constitute an organic being. Both the legislator and the jurist may articulate this law, but they no more invent or make it than does the grammarian who codifies a natural language. Savigny believed that to accept his conception of law was to reject the older notions of natural law; nevertheless, it is often claimed that Savigny's conception was merely a new kind of natural law standing above, and judging, the positive law.

Otto von Gierke (1844–1921), the author of *Das deutsche Genossenschaftsrecht*, clearly fits into the tradition of the historical school. Gray, in *The Nature and Sources of the Law*, subjected the theories of Savigny and his American follower, James C. Carter (1827–1905), to severe criticism. It should be noted that Maine's views have nothing in common with those of Savigny; in Maine's work the metaphysics of the *Volksgeist* is entirely absent.

LATE NINETEENTH CENTURY TO MID-TWENTIETH CENTURY

JHERING AND GERMAN POSITIVISM. Rudolf von Jhering (1818–1892), eminent both as a historian of law and as a legal theorist, rejected both Hegel and Savigny: Hegel, for holding the law to be an expression of the general will and for failing to see how utilitarian factors and interests determine the existence of law; Savigny, for regarding law as a spontaneous expression of subconscious forces and for failing to see the role of the conscious struggle for protection of interests. However, Jhering shared the broad cultural orientation of many of the Hegelians, and he was grateful to Savigny for having overthrown the doctrine of "immutable" natural law. Jhering's contribution was to insist that legal phenomena cannot be comprehended without a systematic understanding of the purposes that give rise to them, the study of the ends grounded in social life without which there would be no legal rules. Without purpose there is no will.

At the same time there are strong strains of positivism in Jhering: Law is defined as "the sum of the rules of constraint which obtain in a state" (*Der Zweck im Recht*, p. 320). In this respect he was close to the German positivists, who emphasized the imperative character of law. Karl Binding (1841–1920), an influential positivist, defined law as "only the clarified legal volition [*Rechtswille*] of a source of law [*Rechtsquelle*]" (*Die Normen und ihre Uebertretung*, p. 68). In this period the slogan of German positivism, "All law is positive law," emerged. Yet Jhering opposed many of the claims of the analytical positivists; his essay "Scherz und Ernst in der Jurisprudenz" (Leipzig, 1885) ridiculed their "heaven of jurisprudential concepts."

SOCIOLOGICAL AND ALLIED THEORIES. Jhering's work foreshadowed many of the dominant tendencies of twentieth-century legal philosophy. Hermann Kantorowicz regarded Jhering as the fountainhead of both the "sociological" and "free-law" schools. The former term covers too wide a group of writers to be surveyed here, some of whom were concerned solely with empirical work, while others combined empirical work with a philosophical outlook. Proponents of the jurisprudence of interests (*Interessenjurisprudenz*) eschewed Jhering's inquiries into the metaphysical and moral bases of purposes, claiming that he did not sufficiently attend to the conflict of interest behind laws; law reflects dominant interest. (Similar analyses were made in the United States; for example, the "pressure-group" theory of politics advanced by A. F. Bentley [1870–1957] in *The Process of Government*, Chicago, 1908.) Much attention was devoted to the analysis of the judicial process and the role that the "balancing" of interests plays in it. As Philipp Heck, one of its leading exponents, remarked: "The new movement of 'Interessenjurisprudenz' is based on the realization that the judge cannot satisfactorily deal with the needs of life by mere logical construction" (*Begriffsbildung und Interessenjurisprudenz*, p. 4).

This sentiment was endorsed by the closely allied "free-law" movement. According to this group, "legal logic" and the "jurisprudence of conceptions" are inadequate for achieving practicable and just decisions. The judge not only perforce frequently goes beyond the statute law, but he also often ought to go beyond it. The "free-law" writers undertook the normative task of supplying guidelines for the exercise of judicial discretion, and the judicial function was assimilated to the legislative function. The focus on such problems reflected the enormous change, occasioned by the industrialization of Western society, in the functions of the state. No longer did the nation-state exist merely to keep the peace or protect preexisting rights; rather, it played a positive role in promoting social and individual welfare. The philosophy of law thus became increasingly concerned with the detailed working out of the foundations of legal policy.

The "free-law" theorist Eugen Ehrlich (1862–1922), who influenced such American theorists as Karl N. Llewellyn (1893–1962) and other representatives of legal realist tendencies, summarized his *Grundlegung der Soziologie des Rechts* as follows: "At the present as well as at any other time, the center of gravity of legal development lies not in legislation, not in juristic science, nor in judicial decision, but in society itself." He rejected the positivistic tenet that only norms posited by the state are legal norms, for in any society there is always more law than is expressed in legal propositions. The "inner order" of an association is the basic form of law. Ehrlich also engaged in empirical study of the "legal facts" (*Rechtstatsachen*) and "living law" of various communities in the Austro-Hungarian Empire. Ehrlich may thus be said to have considered custom as law in its own right. However, many positivists would argue that he was not able to account for the normative character of custom.

MARXISM. The Marxist stress on economic interests was often combined with the sociological and free-law views. Central to the Marxist position are the notions of "class" (usually defined in terms of legal relationship to property and the means of production) and "class interest," which leads to the analysis of the role of law in different societies with differing class structures. Addressing their critics, Karl Marx and Friedrich Engels wrote: "Your law [*Recht*] is but the will of your class exalted into statutes [*Gesetz*], a will which acquires its content from the material conditions of existence of your class" (*Communist Manifesto*, 1848). This suggests that law is merely part of the ideological superstructure and has no effect on the material organization of society. It raises the question of whether law exists in all societies—for instance, in primitive society or in the "classless" society arising after the triumph of socialism—and the further question of the nature and function of law in the transitional period from capitalism to socialism. The issue of "revolutionary legality" or "socialist legality" was treated by V. I. Lenin, E. Pashukanis, and Andrei Vishinsky. An important Marxist study of the relationship between law and the economy is that of the Austrian socialist Karl Renner (*Die Rechtsinstitute des Privatrechts und ihre soziale Funktion*, 1929).

PURE THEORY AND RELATIVISM. Although the sociological approaches to law have many practitioners, the most controversial and perhaps the most influential twentieth-century view was that of Hans Kelsen, a leading exponent of legal positivism. Influenced by the epistemology of the neo-Kantians, Kelsen distinguished sharply between the "is" and the "ought," and consequently between the natural sciences and disciplines, such as legal science, which study "normative" phenomena. Legal science is a descriptive science—prescriptive and valuational questions cannot be scientific—and Kelsen's "pure theory" aimed at providing the conceptual tools for studying any given legal system irrespective of its content. The theory is "pure" in that it is divorced from any ideological or sociological elements; it attempts to treat a legal system simply as a system of norms. Kelsen's view was thus similar to the analytical jurisprudence of Austin, but Kelsen regarded legal norms as "de-psychologized commands." In order to understand an act of will as a norm-creating act, we must already employ a norm that serves as a "schema of interpretation." The jurist who seeks to understand legal phenomena must ultimately presuppose a basic norm (*Grundnorm*), which is not itself a positive legal norm. Legal systems are sets of coercive norms arranged in hierarchical fashion; lower norms are the "concretizations" of higher norms. In Kelsen's analysis the "dualisms" of state and law and public and private law disappear, and the relationship between international law and national legal systems is seen in a fresh light.

Unlike Kelsen, Gustav Radbruch (1878–1949) did not found a school. His position, which he called relativism, has many affinities with that of Kelsen; but Radbruch maintained that law, which is a cultural phenomenon, can be understood only in relation to the values that men strive to realize through it. He attempted to analyze these values in relation to legal institutions, showing the "antinomies" among these values that led to his relativism. World War II raised the question in the minds of many legal philosophers whether the separation of law and morals of legal positivism, which was popular in Germany, contributed to the rise of Nazism. Concern over this problem seems to have caused Radbruch to move away from his earlier relativism toward a kind of natural-law position.

REALISM AND OTHER RECENT TRENDS. In the United States, until the mid-twentieth century, legal philosophy had largely been the province of lawyers rather than of professional philosophers. This may account for its sociological and realistic tone. The erudite Roscoe Pound (1870–1964) was its most prolific writer. Pound recognized the influence of Josef Kohler (1849–1919) and his notion of jural postulates and, especially, of Jhering. The pragmatism of William James also contributed to the development of his views. In an early article, "Mechanical Jurisprudence" (*Columbia Law Review* 8 [1908]: 605–610), Pound argued for an understanding of the interests that the law seeks to protect. Introducing a dis-

tinction between "law in books" and "law in action," he maintained the need for a close study of the actual operation of legal institutions. On both scores his influence in the United States has been momentous, but it is difficult to summarize his position; he is often associated with a "social engineering" approach to law. Law contains both precepts and ideal elements. Among precepts Pound distinguished rules, principles, conceptions, doctrines, and standards. It is pointless to isolate some canonical form to which all laws are reducible. The ideal element consists of received ideals "of the end of law, and hence of what legal precepts should be and how they should be applied." Pound offered an elaborate, although tentative, survey of the individual, public, and social interests secured by law. This list was criticized and amended by Pound's Australian disciple Julius Stone (*The Province and Function of Law*, 1946). In his later years Pound moved toward a kind of natural-law thinking, arguing for a more intimate connection between law and morality; he abjured the realist tendencies, which had been influenced by his earlier thought, as "give it up" philosophies.

It is exceedingly difficult to characterize the legal realists; they disclaim a common doctrine but recognize an interest in a common set of problems. With J. C. Gray, the spiritual godfather of American legal realism was Justice Oliver Wendell Holmes Jr. (1841–1935). In his seminal essay "The Path of the Law" (*Harvard Law Review* 10 [1896]: 457–478), he advocated viewing law as the "bad man" would, in terms of the practicable remedies afforded individuals through the medium of the courts. Holmes presented in that article his famous definition of law as "the prophecies of what the courts will do in fact." It may be argued, however, that this definition, while perhaps adequate from the advocate's viewpoint, can hardly apply to the judge. When the judge asks what the law is on some matter, he is not trying to predict what he will decide.

Joseph W. Bingham was one of the first realists. In "What Is the Law?" (*Michigan Law Review* 11 [1912]: 1–25 and 109–121), Bingham argued that legal rules, like scientific laws, have no independent existence, being simply mental constructs that conveniently summarize particular facts. Laws are really judicial decisions, and the so-called rules or principles are among the (mentally) causative factors behind the decision. This nominalism and behaviorism, which characterized much of early realist writing, was criticized by Morris R. Cohen (1880–1947), until recently one of the few academic philosophers in the United States concerned with legal philosophy. "Behavior analysis" was advocated by Karl N. Llewellyn, who extended it beyond judicial behavior to "official" behavior (*Jurisprudence*, Chicago, 1962; collected papers).

The so-called myth of legal certainty was attacked by Jerome Frank (1889–1957) in his *Law and the Modern Mind* (New York, 1930), which explained the genesis of the myth in Freudian terms. In the sixth edition (New York, 1949) Frank was somewhat friendlier toward natural-law thinking, characterizing his change of attitude as going from an earlier "rule-skepticism" to "fact-skepticism" (*Courts on Trial*, Princeton, NJ, 1949). Other important realists include Thurman Arnold, Leon Green, Felix Cohen, Walter Nelles, Herman Oliphant, and Fred Rodell. Both positivism and realism were attacked by Lon L. Fuller (*Law in Quest of Itself*, Chicago, 1940), a leading American exponent of non-Thomistic natural-law thinking (*The Morality of Law*, New Haven, CT, 1964). The revival of natural-law doctrines is one of the most interesting features of current legal thought. Recent contributions and criticisms may be found in the journal *Natural Law Forum*.

The Scandinavian countries are a center of legal philosophy, and many of their leading writers are realists. They are more consciously philosophical than their American counterparts. The leading spirit was Axel Hägerström (1868–1939), who rejected metaphysical presuppositions in legal philosophy and insisted on an understanding of legal phenomena in empirical terms. Many legal concepts can be understood only as survivals of "mythical" or "magical" thought patterns, which should ideally be eliminated. Vilhelm Lunstedt (*Legal Thinking Revised*, Stockholm, 1956) was most radical in his rejection of metaphysics. Values are expressions of emotion and should be excluded from legal science. The "method of social welfare" should be substituted for the "method of justice." Alf Ross (*On Law and Justice*, London, 1958) argued that the first method is as "chimerical" as the second and presents an analysis of legal policy-making as a kind of rational technology. Laws, Ross argued, are directives to courts. The concept "valid law" as used by jurists and legal philosophers cannot be explicated in purely behavioristic terms; inner psychological attitudes must also be included. A similar view is presented by Karl Olivecrona (*Law as Fact*, London, 1939), who wrote important realist analyses of legal language and severely criticized command theories of law, such as Austin's. In *Inquiries into the Nature of Law and Morals* (translated by C. D. Broad, Cambridge, U.K., 1953), Hägerström argued that Kelsen's "pure theory" never escapes the "will" element either, and hence falls subject

to all the criticisms that may be leveled against the command theories.

In the mid-twentieth century, the most influential legal philosopher in the English-speaking world was H. L. A. Hart. In his *Concept of Law* (Oxford, 1961) he developed a view of law as consisting of a "union of primary and secondary rules." The former are rules imposing duties; the latter are rules of recognition, change, and adjudication. The first of the secondary rules (those for recognizing the rules of a system) seems to be crucial to his account of all three. His position was in many respects similar to that of Kelsen. He gave an interesting analysis, allied to Ross's account, of what it means to say that a rule exists. Hart saw the relationship between law and morals as contingent, in contrast with the Thomistic view of a logical connection between the two; this led him to an interpretation of natural law not unlike that presented by some Renaissance writers. In a number of important articles Hart focused on the nature of definition in jurisprudence, the analysis of psychological concepts in the law, legal responsibility, and the principles of punishment.

See also Aristotelianism; Aristotle; Augustine, St.; Austin, John; Beccaria, Cesare Bonesana; Bentham, Jeremy; Bodin, Jean; Burke, Edmund; Celsus; Cicero, Marcus Tullius; Cohen, Morris Raphael; Engels, Friedrich; Enlightenment; Fichte, Johann Gottlieb; Filmer, Robert; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Hägerström, Axel; Hart, Herbert Lionel Adolphus; Hegelianism; Hippias of Elis; Historical School of Jurisprudence; Hobbes, Thomas; Hume, David; James, William; Justice; Kant, Immanuel; Kelsen, Hans; Legal Positivism; Lenin, Vladimir Il'ich; Locke, John; Marx, Karl; Marxist Philosophy; Medieval Philosophy; Mill, John Stuart; Montesquieu, Baron de; Natural Law; Neo-Kantianism; Patristic Philosophy; Plato; Positivism; Pragmatism; Radbruch, Gustav; Realism; Renaissance; Rousseau, Jean-Jacques; Savigny, Friedrich Karl von; Seneca, Lucius Annaeus; Socrates; Sophists; Stammler, Rudolf; Stoicism; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Utilitarianism; Vico, Giambattista; Vitoria, Francisco de; William of Ockham; Xenophon.

Bibliography

Ago, Roberto. "Positive Law and International Law." *American Journal of International Law* 51 (1957): 691–733.

Allen, C. K. *Law in the Making*, 7th ed. Oxford: Clarendon Press, 1964.

Berolzheimer, Fritz. *System der Rechts- und Wirtschaftsphilosophie*, 5 vols. Munich: Beck, 1904–1907. Vol. II translated by R. S. Jastrow as *The World's Legal Philosophies*. Boston: Boston Book, 1912.

Bodenheimer, Edgar. *Jurisprudence*. Cambridge, MA: Harvard University Press, 1962.

Cairns, Huntington. *Legal Philosophy from Plato to Hegel*. Baltimore: Johns Hopkins Press, 1949.

Carlyle, R. W., and A. J. Carlyle. *A History of Mediaeval Political Theory in the West*, 6 vols. Edinburgh and London: Blackwood, 1903–1936.

Cohen, M. R. *Law and the Social Order*. New York: Harcourt Brace, 1933.

Cohen, M. R. *Reason and Law*. Glencoe, IL: Free Press, 1950.

Davitt, T. E. *The Nature of Law*. St. Louis: Herder, 1951.

Del Vecchio, Giorgio. *Justice*. Translated by Lady Guthrie. Edinburgh: Edinburgh University Press, 1952.

Dias, R. W. M. *Bibliography of Jurisprudence*. London: Butterworths, 1964.

Ebenstein, William. *The Pure Theory of Law*. Madison, WI, 1945.

Ehrlich, Eugen. *Grundlegung der Soziologie des Rechts*. Munich, 1913. Translated by Walter L. Moll as *Fundamental Principles of the Sociology of Law*. Cambridge, MA: Harvard University Press, 1936.

Entrèves, A. P. d'. *Natural Law*. London: Hutchinson, 1951.

Friedman, Wolfgang. *Legal Theory*, 4th ed. London: Stevens, 1960.

Friedrich, C. J. *The Philosophy of Law in Historical Perspective*, 2nd ed. Chicago: University of Chicago Press, 1963.

Gierke, Otto von. *Das deutsche Genossenschaftsrecht*, 4 vols. Berlin: Weidmann, 1868–1913. Partially translated by F. W. Maitland as *Political Theories of the Middle Age*. Cambridge, U.K.: Cambridge University Press, 1900. Also partially translated by Ernest Barker as *Natural Law and the Theory of Society*. Cambridge, U.K.: Cambridge University Press, 1934.

Gurvitch, Georges. *Sociology of Law*. New York: Philosophical Library and Alliance Book, 1942.

Hall, Jerome. *Studies in Jurisprudence and Criminal Theory*. New York: Oceana, 1958.

Hamburger, Max. *The Awakening of Western Legal Thought*. Translated by B. Miall. London, 1942.

Hart, H. L. A. "Philosophy of Law and Jurisprudence in Britain (1945–1952)." *American Journal of Comparative Law* 2 (1953): 355–364.

Heck, Philipp. *Begriffsbildung und Interessenjurisprudenz*. Tübingen, 1952.

Jaeger, Werner. "In Praise of Law." In *Interpretations of Modern Legal Philosophies*, edited by P. L. Sayre. New York: Oxford University Press, 1947.

Jennings, W. I., ed. *Modern Theories of Law*. London: Oxford University Press, H. Milford, 1933.

Jhering, Rudolf von. *Die Normen und ihre Uebertretung*. Leipzig, 1872.

Jhering, Rudolf von. *Der Zweck im Recht*, Vol. I. Leipzig: Breitkopf and Härtel, 1877. Translated by I. Husik as *Law as a Means to an End*. Boston: Boston Book, 1913.

Jolowicz, H. F. *Lectures on Jurisprudence*. London: University of London, Athlone Press, 1963.

Jones, J. W. *Historical Introduction to the Theory of Law*. Oxford: Clarendon Press, 1940.

Kantorowicz, Hermann. *The Definition of Law*. Cambridge, U.K.: Cambridge University Press, 1958.
Kelsen, Hans. *What Is Justice?* Berkeley: University of California Press, 1957.
Kocourek, Albert. "The Century of Analytical Jurisprudence since John Austin." In *Law: A Century of Progress*, edited by A. Reppy. New York: New York University Press, 1937.
Lloyd, Dennis. *Introduction to Jurisprudence*. London: Stevens, 1959.
Lottin, O. *Le droit naturel chez saint Thomas d'Aquin et ses prédécesseurs*. Bruges, 1931.
Macdonnell, J., and E. W. D. Manson, eds. *Great Jurists of the World*. London: Murray, 1913. Biographical.
Morris, C. "Four Eighteenth Century Theories of Jurisprudence." *Vanderbilt Law Review* 14 (1960): 101–116.
Paton, G. W. *A Text-Book of Jurisprudence*, 2nd ed. Oxford: Clarendon Press, 1951.
Patterson, E. W. *Jurisprudence*. New York: Foundation Press, 1953.
Pollock, Frederick. *Jurisprudence and Legal Essays*. London: Macmillan, 1961.
Pound, Roscoe. *Jurisprudence*, 5 vols. St. Paul, MN: West, 1959. A collection of much of his work.
Renner, Karl. *Die Rechtsinstitute des Privatrechts und ihre soziale Funktion*. Tübingen: Mohr, 1929. Translated by A. Schwarzschild as *The Institutions of Private Law and Their Social Function*. London, 1949.
Rommen, Heinrich. *Die ewige Wiederkehr des Naturrechts*. Leipzig, 1936. Translated by T. R. Hanley as *The Natural Law*. St. Louis, 1947.
Spencer, A. W., ed. *Modern French Legal Philosophy*. New York, 1921.
Stone, Julius. *The Province and Function of Law*. Sydney: Associated General Publications, 1946.
Ullmann, Walter. *The Medieval Idea of Law*. London, 1946.
Verdross, A. *Abendländische Rechtsphilosophie*. Vienna: Springer, 1958.
Villey, M. *Leçons d'histoire de la philosophie du droit*. Paris: Dalloz, 1957.
Vinogradoff, Paul. *Outlines of Historical Jurisprudence*, 2 vols. London: Oxford University Press, 1910–1922.

M. P. Golding (1967)

The article below reproduces some portions and revises other portions of the previous article.

PHILOSOPHY OF LAW, HISTORY OF [ADDENDUM]

The problems of authority, law and order, obligation, and self-interest first became central topics of speculation in the thought of the Sophists, in the late fifth and early fourth centuries BCE. The most famous Sophists stressed the distinction between nature (*physis*) and convention (*nomos*), and they put laws in the latter category. They generally attributed law to human invention and justified obedience to law only to the extent that it promoted one's own advantage. Laws were artificial, arrived at by consent; most acts that were just according to the law were contrary to nature; the advantages laid down by the law were chains upon nature, but those laid down by nature were free.

In the time of the Sophists notions of law, justice, religion, custom, and morality were largely undifferentiated; yet in this same period some of the crucial problems of legal philosophy were first formulated, and attempts were made at a formal definition of law. Thus, Xenophon (*Memorabilia* I, 2) reported that Alcibiades, who associated with both Critias and Socrates, remarked to Pericles that no one can really deserve praise unless he knows what a law is. Pericles replied that laws are what is approved and enacted by the majority in an assembly, whereby they declare what ought and what ought not to be done. He admitted that if obedience is obtained by mere compulsion, it is force and not law, even though the law be enacted by the sovereign power in the state. Xenophon also recounted an purported conversation between Socrates and the Sophist Hippias in which both maintained an identity between law—or what is lawful—and justice—or what is right—while admitting that laws may be changed or annulled (*Memorabilia* IV, 4).

Socrates claimed that there are "unwritten laws," uniformly observed in every country, that cannot conceivably be products of human invention. They are made by the gods for all men, and when men transgress them, nature penalizes the breach.

Socrates and the Sophists, as presented in Plato's dialogues, disagreed concerning human nature. The Sophists conceived of humans as egoistically motivated and antisocial, whereas for Socrates, as for Plato and Aristotle, people are a social beings, other-regarding as well as self-regarding, who find fulfillment in social life. By contrast, the Sophist Callicles, in Plato's *Gorgias*, holds that we are no exception to the law of nature, according to which the stronger rules; man-made laws and social institutions violate human nature. The less radical Sophists, although they could not identify law with some feature of reality, still accepted its practical usefulness.

PLATO AND ARISTOTLE

PLATO. There is hardly any problem of legal philosophy not touched upon by Plato. He wrote during the decline of the Greek polis, when law and morality could appear as mere conventions imposed by shifting majorities in their own interest and the harmony between the legal

order and the order of the universe could not easily be maintained. Plato sought to restore, as far as possible, the traditional analogy between justice and the ordered cosmos. Justice, or right action, cannot be identified with mere obedience to laws, nor can a truly moral life be reduced to conformity with a conventional catalogue of duties. Duties involve a knowledge of what is good for human beings, and such knowledge bears an intimate relation to human nature.

The question "What is justice?" dominates Plato's *Republic*. Plato conceived of justice as that trait of human character that coordinates and limits to their proper spheres the various elements of the human psyche. In order to understand the operation of justice in the human soul, Plato examined human nature writ large: the city-state. The state functions well when it is governed by those who know the art of government, and the practice of this art requires a positive insight into the Good. In a just society every citizen performs the role to which he or she is best suited for the good of the whole. Similarly, in the moral economy of the individual's life, justice prevails when reason rules and the appetites and lower passions are relegated to their proper spheres. A just social order is achieved to the extent to which reason and rational principles govern the lives of its members.

Plato's emphasis on reason found its way into his definition of law. Law is reasoned thought (*logismos*) embodied in the decrees of the state (*Laws* 644D). Plato rejected the view that the authority of law rests on the mere will of the governing power. The *Laws* contains a detailed discussion of many branches of law and is an attempt at a formulation of a systematic code to govern the whole of social life. In contrast with the ideal polis of the *Republic*, in which there would be little need for legislation, in the *Laws* Plato accepted "law and order, which are second best" (*Laws* 875D).

ARISTOTLE. Aristotle, who discussed law in numerous contexts, nowhere gave a formal definition of it. He wrote variously that law is "a sort of order, and good law is good order" (*Politics* 1326a), "reason unaffected by desire" (*Politics* 1287a), and "the mean" (*Politics* 1287b). However, these must be taken not as definitions but as characterizations of law motivated by the point Aristotle was making in the given context.

Following Plato, Aristotle rejected the Sophistic view that law is mere convention. In a genuine community—as distinguished from an alliance, in which law is only a covenant—the law concerns itself with the moral virtue of the citizenry (*Politics* 1280b). Aristotle sharply distinguished between the constitution (*politeia*) and laws (*nomoi*); the constitution concerns the organization of offices within the state, whereas the laws are "those according to which the officers should administer the state, and proceed against offenders" (*Politics* 1289a). The constitution of a state may tend to democracy, although the laws are administered in an oligarchical spirit and vice versa (*Politics*1292b). Legislation should aim at the common good of the citizens, and justice—what is equal—should be determined by the standard of the common good (*Politics*1283a). Yet Aristotle recognized that the law is often the expression of the will of a particular class, and he stressed the role of the middle class as a stabilizing factor.

In his discussion of the forms of government in Book III of the *Politics*, Aristotle took up the Platonic problem of rule by the best man versus rule according to laws. A society of equals by its very nature excludes the arbitrary rule of one individual. In any case, even the best person cannot dispense with the general principles contained in laws; and legal training helps to make better government officials. Furthermore, administrators, like all people, are subject to passion, and it is thus preferable to be judged by the impersonal yardstick of the laws. The importance of the rule of law in no way conflicts with the need to change the law through legislation when it has been found by experience to be socially inadequate. But not all law is the product of legislation; customary law is in fact more important than the written law.

Aristotle's discussion of the judicial process foreshadows many modern notions. Although it is better to have written laws than to rely completely on discretion, "some matters can be covered by the laws and others cannot" (*Politics*1287b20). General rules are insufficient to decide particular cases (*Politics*1286a26), although "well-drawn laws should themselves define all the points they possibly can and leave as few as may be to the decision of the judges" (*Rhetoric* 1354a32). Aristotle seems to have had two considerations in mind. First, judicial decision-making is practical—it involves deliberation—and as such cannot be completely determined in advance. Second, the resolution of disputed issues of fact that determine the outcome of a particular case cannot be settled in advance by legislation. This stress on the insufficiency of general rules connects with Aristotle's influential discussion of equity (*epieikeia*). Equity is just, "but not legally just but a correction of legal justice" (*Nicomachean Ethics* 1137b10).

Aristotle sometimes seems to suggest that equity comes into play when there are gaps in the law, so that it

consists in the judge's acting as the lawgiver would act if he were present. Yet he also seems to suggest that equity corrects the harshness of the law when adherence to the written law would work an injustice. Principles of equity are thus closely related to the unwritten universal laws "based on nature," a "natural justice" binding on all persons, even those who have no association or covenant with one another. Nevertheless, what is naturally just may vary from society to society.

The *locus classicus* of Aristotle's discussion of justice is Book V of the *Nicomachean Ethics*. Generically, justice has to do with one's relations to others, and there is a sense of justice that refers to the complete moral virtue of the member of the community in such dealings. There is also a sense in which "justice" refers to a particular virtue involving the fair dealings of individuals in matters handled by private law. Two kinds of rights fall under this special virtue: rights in division (where each individual claims his fair share of goods, honors, and so on) and rights in redress (for wrongs done by one individual to another, such as failure to fulfill a contract).

ROME

STOICS. The Stoics, who conceived of the universe as a single, organic substance, exercised a lasting influence on legal thought. In their view, nature, which exhibits structure and order, and man both partake of intelligence, or reason (*logos*). An animal is directed by a primary impulse toward self-preservation, which adapts it to its environment. In humans, reason is the "engineer of impulse," and our actions may be evaluated only within the framework of the whole of nature. The criterion of moral action is consistency with the all-determining law of nature (*koinos logos*). This conception of a law of nature that is the ultimate standard of human laws and institutions was combined with Aristotelian and Christian notions to form the long-standing natural-law tradition of medieval legal philosophy. Another important Stoic contribution was the belief in the equality of all people in a universal commonwealth and a rejection of Aristotle's doctrine of slavery.

CICERO AND SENECA. The writings of Cicero (106–43 BCE) were important in transmitting classical legal thought to the medieval world. Although Cicero was a professional arguer of legal cases, his philosophical treatment of law in his *De Legibus* disclaims any interest in "clients' questions" or the "law of eaves and house-walls." His legal philosophy was essentially Stoic; he denied that the positive law of a community (written or customary), even when universally accepted, is the standard of what is just. Nor is mere utility the standard: "Justice is one; it binds all human society, and is based on one law, which is right reason applied to command and prohibition" (*De Legibus* I, 15). An unjust statute is not a true law. Law and morality are logically connected, and only that which conforms to the law of nature is genuine law. This view exercised a lasting influence on natural-law thinking and reappeared in the thought of Thomas Aquinas.

Like Cicero, Seneca (c. 4 BCE–65 CE) aided in transmitting Stoic notions to later thinkers. He reiterated the conception of the equality of all persons under natural law, but perhaps more important was his conception of a golden age of human innocence, a prepolitical state of nature. Legal institutions became necessary as human nature became corrupted.

ROMAN LAW. The influence of Stoicism may be traced in pronouncements of the Roman jurists. It is disputed whether these were any more than remarks designed to ornament legal texts, but they nevertheless influenced the thought of later ages. The jurists distinguished three kinds of law: *jus naturale, jus gentium*, and *jus civile*. In practice, the last originally referred to the law of the city of Rome, but ultimately it was applied to any body of laws of a given community. The *jus gentium* first meant the law applied to strangers, to whom the *jus civile* was not applicable, and was later extended to those legal practices common to all societies. Gaius (mid-second century), who systematized the Roman law in his *Institutes*, identified the *jus naturale* and *jus gentium* as universal principles of law agreeable to natural reason and equity. Thus, law was not a mere expression of human will or institution but that which is rationally apprehended and obeyed. The *jus gentium* was not an ideal law by which the positive law was judged but the rational core of existing legal institutions.

Ulpian (c. 170–228) distinguished *jus naturale* from *jus gentium* by stating that *jus naturale* is not peculiar to human beings but is taught by nature to all animals. Thus, among animals there is an institution similar to human marriage. Slavery and its attendant rules are products of the *jus gentium*, for by the *jus naturale* all people were born free. It is not clear, however, that Ulpian regarded slavery as bad. To him we owe the oft-repeated definition of justice: "the constant wish to give each his due" (*Digest* I, 1, 10). Following Celsus (c. 67–c. 130), he defined law (*jus*) as "the art of the good and the equitable" (*ibid*. I, 1, 1). Again, it does not seem that Ulpian thought of the *jus naturale* as an ideal law opposed to the

jus civile or to the *jus gentium*. It has been suggested that behind Ulpian's thought was a conception of a natural state antecedent to the conditions of organized society.

The doctrines of the Roman jurists owe their lasting influence to their incorporation into the *Corpus Juris Civilis* of Justinian (sixth century), principally in the section called the *Digest*. The compilers of Justinian's *Institutes* (a section of the *Corpus Juris*) seem to have distinguished the *jus naturale* from the *jus gentium* and seem to have regarded the former as a set of immutable divine laws by which the positive law may be morally evaluated (*Institutes* I, 2, 11; III, 1, 11). The *Corpus Juris* also preserved statements of the Roman jurists concerning the source of the authority to make and unmake the laws constituting the civil law. According to a number of these statements, this authority resides in the consent of the people; however, the statement that "what pleases the prince has the force of law" (*Digest* I, 4, 1) was probably a more accurate view of the facts. Justinian seems to have combined these views theoretically in his reference to a (nonexistent) "ancient law" by which the Roman people transferred all their powers to the emperor (*Codex* I, 17, 1, 7).

EARLY MIDDLE AGES

To the legal thought of the Stoics and the Roman philosophers and jurists the Church Fathers added a distinctively Christian element. The law of nature was no longer the impersonal rationality of the universe but was integrated into a theology of a personal, creative deity. The relationship among the Mosaic law, the Gospels, and natural law emerged as a specific problem; the notion of *jus divinum* (divine law) as a distinct type of law, along with the three recognized by the jurists, was crystallized. The notion of the fall of man from a state of perfection (which may be compared with the view of Seneca) played an important role. Thus, according to St. Ambrose (340–397) the Mosaic law—a law of sin and death (see Romans 8.2)—was given because humans failed to obey the law of nature. The fact that many legal institutions, such as slavery and private property, deviate from this ideal law does not necessarily imply that they are unjust or illegitimate; for the natural law is adapted to us only in a condition of innocence.

Of the Church Fathers, St. Augustine (354–430) was perhaps the most original and complex: Only one point in his thought will be noted here. Cicero maintained that nothing can be nobler than the law of a state (*De Legibus* I, 14) and that if a state has no law, it cannot truly be considered a state (*ibid.* II, 12). The law of the state must therefore embody justice, for without *justitia* there is no *jus*. Augustine considered this position in *The City of God*, Book XIX. According to Augustine, since Rome had no justice, Cicero's position has the inconvenient consequence that Rome was no state at all. We must therefore seek another definition of "state" (*populus*) in which justice is not an essential element. Augustine stressed the notion of order—"a harmonious multitude"—with the suggestion that legal order need not be moral or just. There are passages in Augustine, however, which seem to uphold a more orthodox natural-law position. In any event, the terms of his discussions are somewhat different; his main points of contrast are divine and human law rather than *jus naturale* and *jus civile*.

The sources of the natural-law theories which were to dominate Western legal philosophy for many centuries were the writings of the Greek and Roman philosophers and poets, Justinian's *Corpus Juris Civilis*, and the Church Fathers. Isidore of Seville (c. 560–636), an encyclopedist and an important transmitter of Roman thought to later writers, concisely expressed the natural-lawyer's ideal regarding positive law: "Law shall be virtuous, just, possible to nature, according to the custom of the country, suitable to place and time, necessary, useful; clearly expressed, lest by its obscurity it lead to misunderstanding; framed for no private benefit, but for the common good" (*Etymologies* V, 21).

MIDDLE AGES AND RENAISSANCE

CIVILIANS AND CANONISTS. The revived study of Roman law in the twelfth century, associated with the glossators, gave a fresh stimulus to legal philosophy. Of special interest are the attempts at reconciling differences among the Roman jurists on the definition of law and the classification of its branches. In the main, the civilians were in the broad tradition of natural-law thinking; *jus* flows from *justitia*, although it must always fall short of perfect justice, which is God's alone. Irnerius (c. 1050–c. 1130) thus claimed that statutes ought to be interpreted in the light of equity. Strict law requires that all agreements be kept, but equity allows exceptions to the rule. This equity, according to Azo (c. 1150–c. 1230), is a principle that must be written, not merely lodged in the judge's heart.

The middle of the twelfth century also saw the systematization of the canon law. In the *Decretum* of Gratian a high degree of jurisprudential competence was brought to this task. The tripartite division of law of the Roman lawyers was accepted, but the leading conceptions were Augustine's *jus divinum* and *jus humana*. Natural law was identified with the former, whereas the distinctive feature

of the latter (covering both *jus gentium* and *jus civile*) was custom. Natural law is contained in the Mosaic law and the Gospels; the command to do unto others what we would have them do unto us is its fundamental principle. Natural law relates to our rational nature and is immutable; the *mistica*, the cultic regulations found in scripture, are part of the natural law only in their moral aspect.

The commentators on Gratian further divided natural law so as to include not only commands and prohibitions but also *demonstrationes*, which point to what is good for all humans, such as possession of all things in common. In our fallen condition custom has legitimately modified the *demonstrationes* in permitting private property and slavery. The other branches of natural law may not be abrogated and are the standards by which even the ecclesiastical law must be judged. Gratian (if not all his commentators) seems to have generally maintained a clear distinction between natural (divine) law and canon law.

AQUINAS. The rediscovery of Aristotle in the thirteenth century greatly influenced the further development of legal philosophy. The culmination of the natural-law tradition is the theory of Thomas Aquinas (1224–1274), who integrated Stoic, Christian, and Aristotelian elements within a comprehensive philosophic system. Laws are standards of conduct that have a binding, or obligatory, character. This idea can be understood only if laws have some kind of rational origin. Combining this view with a teleological conception of nature and social order, Aquinas regarded legal control as purposive. Laws, he concluded, are ordinances of reason promulgated for the common good by the legitimate sovereign.

According to Aquinas, four types of law may be distinguished: eternal law, an expression of God's rational ordering of the universe; divine law, which guides us toward our supernatural end; natural law, which guides us toward our natural end; and human law, which regulates through the prospect of punishment the affairs of people in a given community in the light of that community's special requirements. Crucial to the concept of natural law are the notions of natural inclinations and right reason. "All those things to which man has a natural inclination are naturally apprehended by reason as being good and consequently as objects of pursuit, and their contraries as evil, and objects of avoidance" (*Summa Theologiae* I–II, 94). The relationship between inclination and reason, accounting for the apprehension of the natural law, has been variously interpreted. The precepts of natural law have as their common foundation the principle "Do good and avoid evil." Natural law is a standard to which human law must conform, and Aquinas employed Aristotle's conception of practical reasoning in explaining the derivation of human law from natural law by the legislator, thus accounting for differences between legal systems and for the possibility that rational men should disagree as to what human laws ought to be. He affirmed the long-standing view that an unjust law is no law; but although an unjust law is not binding in conscience, considerations of utility may require one to obey it. Aquinas allowed that such "laws" may be said to possess a "legal" character insofar as they are promulgated under the color of law by the legitimate prince.

Aquinas discussed in detail and with great acuity all of the problems treated by his predecessors. His influence may be traced in the English writers John Fortescue (c. 1394–1476), Thomas Hooker (c. 1586–1647), and Christopher St. Germain (1460–1540). According to St. Germain, natural law is nothing other than the common-lawyer's notion of "reasonableness." Among late-twentieth century Thomist scholars, the works of John Finnis have been especially influential.

OCKHAM. Some medieval writers seem to have espoused a protopositivism in their emphasis on the primacy of the will; this is characteristic of the Augustinian-Franciscan tradition. Thus, William of Ockham (c. 1285–1349) regarded the divine will as the norm of morality. "By the very fact that God wills something it is right for it to be done." Nevertheless, it is doubtful that Ockham would have affirmed that what the sovereign commands is just. His position is somewhat unclear, however, for he—like all medieval writers—continued to use the rhetoric of natural law in his *Dialogus*: in one of its senses, *jus naturale* is composed of universal rules of conduct dictated by natural reason. A right, such as the immutable right of private property, is a dictate of right reason.

THE RISE OF ABSOLUTISM. A tendency to combine natural-law doctrines with a theory of royal absolutism began in the fourteenth century. A group of civilians known as the postglossators undertook to forge a workable system of law out of the older Roman law, which they regarded as the *jus commune* of Europe. The technically trained administrators in the rising nation-states, they were naturally concerned with fundamental problems of legal theory. Bartolus of Sassoferrato (1314–1357) maintained that the ruler is not bound by the laws, although it is "equitable" that he should voluntarily submit to them. The *jus gentium*, however, is immutable. Lucas de Penna (1320–1390) discussed jurisprudential questions in

detail. Law is the articulation of the ethical virtue of justice, and reason is the foundation of law. At the same time, he maintained, as did many civilians, that the prince's lordship rests on divine authority. The ruler is responsible to God alone and not to the people; law is not the expression of the will of the community. Nonetheless, although the prince is unfettered by the laws, bad laws (those that contradict divine law) have no binding force. It is not clear, in Lucas's view, whether the obligation to obey law derives primarily from the rationality of law or from the divine grant of authority to the ruler.

LATER RENAISSANCE

BODIN. Jean Bodin (1530–1596) was a great exponent of unlimited sovereignty under natural law whose views were apparently influenced by the fourteenth-century civilians. Like them, he appears to have had difficulty in adapting Christian legal thought to the conditions of the secular nation-state. In his *Six Books of the Commonwealth*, Bodin was emphatic that "law is nothing else than the command of the sovereign in his exercise of sovereign power." But although the prince "has no power to exceed the law of nature," which is decreed by God, it seems plain that Bodin no longer thought of right reason as linking natural and positive law. Bodin's endorsement of the command theory also appears in his treatment of custom. The relative weights of positive law and custom had long been debated by the medieval lawyers, but Bodin was one of the first to hold that custom owes its legal authority to the sufferance of the ruler. In this he anticipated the idea of tacit command expressed by Thomas Hobbes and John Austin.

INTERNATIONAL LAW. The emergence of nation-states also brought the problem of the rational foundation of international law to the forefront of legal thinking. This development may be seen in the writings of the Spanish Thomists Francisco de Vitoria (c. 1492–1546) and Francisco Suárez (1548–1617) and of Hugo Grotius (1583–1645), a Dutch Protestant jurist with broad humanistic leanings. According to Vitoria, the *jus gentium* either belongs to or is derivable from the natural law and consists in prescriptions for the common good in the widest sense—namely, for the international community. Rights and obligations are thus conferred upon nations acting through their rulers.

The conception of a law of nations was developed in great detail by Suárez. Although his *De Legibus* is Thomistic in many respects, Suárez explicitly stated that Aquinas's account of law is inadequate. Suárez began by distinguishing laws in the prescriptive sense from laws of nature in the descriptive sense, which are laws only metaphorically. (Many positivists trace the origin of natural-law thinking to the tendency to confuse these two types of law.) With regard to prescriptive laws, Suárez defined a law (*lex*) as "the act of a just and right will by which the superior wills to oblige the inferior to this or that" or as "a common, just and stable precept, which has been sufficiently promulgated" (*De Legibus* I, 12). The reference to stability is notable: Laws generally survive both the lawgiver and the populace living when they are enacted, and they are valid until abrogated. Such considerations have led recent writers to reject the identification of laws with mere acts of will; but although Suárez rejected the voluntaristic notion of natural law associated with the Ockhamists, he held that the civil law is enacted "more by the will than by reason." It is not derived from natural law by logical inference but by "determination," and hence is, in a sense, arbitrary (*De Legibus* II, 20).

Most medieval writers tended to use *lex* and *jus* interchangeably; Suárez, however, defined the latter as "a certain moral power which every man has, either over his own property or with respect to what is due to him" (*De Legibus* I, 2). Although Aquinas briefly discussed *jus naturale* as contrasted with *jus positivum (Summa Theologiae* II–II, 57), the concept of a "natural right" was almost entirely absent from his thought. It is clearly present in Suárez, who, in the style of Locke and the Enlightenment philosophers, formulated a list of natural rights. Nevertheless, the individualism of these writers is not present in Suárez. His attitude was quite remote from eighteenth-century natural-law and natural-right theorists, who thought that a perfect system of law could be deduced from the natural law.

Despite Grotius's tendency to underestimate his predecessors, his *De Jure Belli ac Pacis* (1625) clearly showed the influence of such writers as Vitoria and Suárez. He developed their notion of a "just war," a topic still discussed by theorists concerned with the problem of sanctions in international law. Just wars presuppose the existence of laws governing relations between sovereign states; such laws have their origin in natural law and in treaties, which in turn presuppose precepts of the law of nature. The denial of the existence of natural law supposes that people are egoistically motivated, accepting law as a "second best." However, following Aristotle and the Scholastics, Grotius held that humans are social, altruistic, and rational. Therein lies the origin of law, which would be binding whether or not God exists. This statement has been regarded by historians as epoch-making;

they claim that Grotius separated jurisprudence from theology. More important, perhaps, is the tendency in Grotius and others who followed him to identify natural law with certain rational principles of social organization and thus to loosen its tie with the Stoic metaphysical conception of the law of nature.

THE SEVENTEENTH TO LATE-NINETEENTH CENTURIES

HOBBES AND MONTESQUIEU. Thomas Hobbes (1588–1679) was perhaps the most important of the seventeenth-century legal philosophers. His break with the tradition of natural law provoked much controversy. Hobbes employed the terminology of "natural right," "laws of nature," and "right reason." But the first was for him simply "the liberty each man hath to use his own power as he will himself, for the preservation of his own nature; that is to say, of his own life" (*Leviathan* 14); the second are principles of self-interest, which are often identified with the third. There is no right reason in nature (*Elements of Law* II, 10, 8). The natural condition of mankind is one of perpetual war, in which common standards of conduct are absent. There is no right or wrong, justice or injustice, mine or thine in this situation. The crucial steps in Hobbes's theory are the identifications of society with politically organized society and of justice with positive law. Laws are the commands of the sovereign; it is in reference to such commands that the members of a society evaluate the rightness or justness of their behavior. An "unjust law" is an absurdity; nor can there be legal limitations on the exercise of sovereign power.

No writer has put forward a positivistic conception of law with greater style and forcefulness than Hobbes. Difficulties in his position emerge from three areas: his concession that, although the sovereign cannot commit an injustice, he may commit iniquity; the idea of injury to God in the state of nature; and the treatment of conscience in *De Cive*. Hobbes solved the problem of the source of the obligation to obey the sovereign's command by his "social contract" doctrine, the interpretation of which is still discussed by scholars. His unfinished *Dialogue Between a Philosopher and a Student of the Common Laws of England* examines various doctrines of the English law as put forward by Sir Edward Coke, and it is notable for its critical examination of Coke's statement that reason is the life of the law.

The *Second Treatise of Civil Government* by John Locke (1632–1704), primarily an attack on Robert Filmer's "divine right" theory, contains certain implied criticisms of Hobbes. Its interest for legal philosophy lies in its use of a version of the social contract to treat the question of the obligation to obey the law, its conception of limitations on sovereign power, and its individualistic view of natural inalienable rights, particularly rights in property. Locke's influence was enormous, and his view of natural rights had a profound effect on the development of law in the United States.

A new approach to the understanding of law and its institutions was put forward by Montesquieu (1689–1755). He, too, spoke the language of natural law and defined laws as "necessary relations arising from the nature of things" (*The Spirit of the Laws* I, 1). But his special importance lies in his attempt to study legal institutions by a comparative historical method, stressing the environmental factors that affect the development of law. This suggestion had been anticipated by Bodin, and Giambattista Vico (1668–1744) had also applied a historical method to the study of Roman law, but Vico's work had little immediate influence. Montesquieu's doctrine of the separation of powers had an extraordinary influence. His sharp separation of judicial from legislative and executive power reinforced the conception that the judge is a mere mouthpiece of the law and that judges merely declare the existing law but never make it. In 1790, in his *Reflections on the Revolution in France*, Edmund Burke turned the historical approach to a practical political use when he protested against proceeding a priori in the "science of constructing a commonwealth."

KANTIANISM. Immanuel Kant (1724–1804) contributed to legal philosophy and other branches of philosophy. The keynote of his legal philosophy was inspired by Jean-Jacques Rousseau (1712–1778), who set as the problem of his *Social Contract* the reconciliation of social coercion and individual freedom. Kant's legal philosophy was a philosophy of justice in which the concept of freedom plays a central role. Kant sought a systematic understanding of the principles underlying all positive laws, one that would enable us to decide whether these laws are in accordance with moral principles. Kant held that positive law "proceeds from the will of a legislator," and any viable legal system will take into account the particular conditions of the given society: With these conditions the theory of law has no concern. The theory is an application of the results of moral philosophy to the conditions of "men considered merely as men." This endeavor covers both the domain of law (*Recht*) and the domain of ethics; the principle that right action is action in conformity with universalizable maxims holds for both juridical and moral laws.

A law (*Gesetz*) is a formula expressing "the necessity" of an action. Juridical and moral laws are distinguished in that the former regulate external conduct irrespective of its motives. (But this does not mean that a judge should necessarily ignore the lawbreaker's motives when passing sentence upon him.) Any person, as a morally free agent, is entitled to express his freedom in activity so long as it does not interfere with the similar freedom that others possess. This is the principle underlying all legislation and "right." Juridical law also involves the authority to compel conformity and to punish violations. The necessary and sufficient condition for legal punishment is that the juridical law has been broken. It must be recognized, however, that the domain of such law is restricted by the limits of compulsion. While it is morally wrong to save one's own life by killing another, even where this is the only expedient, it can never be made legally wrong to kill in such a case. The principle of law receives content in Kant's application of it to particular private rights in external things and in his analysis of the methods for acquiring such rights.

Kant's influence on jurisprudence, after being somewhat eclipsed by Hegelianism, reemerged at the end of the nineteenth century. One of the most important Neo-Kantians was Rudolf Stammler (1856–1938), who invented but eventually discarded the phrase "natural law with variable content." Accepting the Kantian distinction between "form" and "matter," he attempted to discern the form of all laws. He defined law as "exceptionless binding volition." Just law is an ideal involving principles of respect and cooperation.

UTILITARIANISM AND POSITIVISM. Although Kant and his followers may be said to have inspired a variety of natural-law philosophies (although different from the Stoic and Thomistic types), Jeremy Bentham (1748–1832) and his followers (notably John Stuart Mill) claim to have rejected such thinking entirely. Of the influences on Bentham, two may be briefly noted. David Hume (1711–1776) argued that moral distinctions are not derived from reason; passion, or sentiment, is the ultimate foundation of moral judgment, and justice is grounded in utility. Second, the Italian criminologist Cesare Beccaria (1738–1794), in his *Of Crimes and Punishments* (1764), subjected the existing institutions of criminal law and methods of punishment to relentless criticism. His standard of judgment was whether "the greatest happiness of the greatest number" was maximized. Bentham acknowledged his debt to Beccaria, and this "principle of utility" was the base of Bentham's voluminous projected "codes." He did not, however, define the nature of law by reference to utility. In his *The Limits of Jurisprudence Defined* (published in 1945) he defined a law as the expression of "the will of a sovereign in a state." Bentham's views, which were well suited to deal with the problems engendered by the Industrial Revolution in England, were of immense importance in effecting legal reform. In 1832, the year of his death, the Reform Act was passed, largely as a result of the work of his followers. Mill's *On Liberty* (1859) is an attempt to treat the limits of legal coercion by the state along modified utilitarian lines.

In legal philosophy Bentham's influence affected the English-speaking world especially through the thought of John Austin (1790–1859), a seminal figure in English and American legal positivism and analytic jurisprudence. Austin tried to find a clear demarcation of the boundaries of positive law, which would be antecedent to a "general jurisprudence" comprising the analyses of such "principles, notions, and distinctions" as duty, right, and punishment, which are found in every legal system; these analyses in turn were to be employed in "particular jurisprudence," the systematic exposition of some given body of law.

Austin began by distinguishing "law properly so called" and "law improperly so called." The former is always "a species of command," an expression of a wish or desire, analytically connected with the ideas of duty, liability to punishment (or sanction), and superiority. The last notion led Austin to his famous and influential analysis of "sovereignty": "laws strictly so called" (positive laws) are the commands of political superiors to political inferiors. From this it follows that international law is merely "positive international morality" rather than law in a strict sense. (Some writers, viewing this as an unfortunate and perhaps dangerous consequence, were led to various revisions of Austinianism.) Austin's "separation" of law and morality is often taken as the hallmark of legal positivism. "The existence of law is one thing; its merit or demerit is another," he wrote in *The Province of Jurisprudence Determined* (V, note). Yet Austin was a utilitarian; in distinguishing between the law that is and the law that ought to be, he did not mean that law is not subject to rational moral criticism grounded in utility, which he took to be the index to the law of God. At this point Austin was influenced by such "theological utilitarians" as William Paley.

Austin's views were subjected to vigorous discussion both without and within the traditions of positivism and analytical jurisprudence. And as the disciplines of history, anthropology, and ethnology assumed an increasing

importance during the nineteenth century, rival approaches to the understanding of law developed. Thus, Sir Henry Maine (1822–1888), who formulated the historical law that legal development is a movement from status to contract, argued in his *Early History of Institutions* that the command-sovereignty theory of law has no application in a primitive community, where law is largely customary and the political "sovereign," who has the power of life or death over his subjects, never makes law. The Austinian view can be saved only by maintaining the fiction that what the "sovereign" permits, he commands. Nonetheless, Austin had many followers at the turn of the twentieth century, such as T. E. Holland (1835–1926) and J. W. Salmond (1862–1924), who attempted to preserve the imperative and coercion aspects of his theory while introducing revisions.

The role of the courts was increasingly emphasized. In the United States, John Chipman Gray (1839–1915) wrote *The Nature and Sources of the Law*, one of the most important American contributions to the subject. Acknowledging his debt to Austin, Gray defined law as "the rules which the courts [of the State] lay down for the determination of legal rights and duties." This required him to construe statutes, judicial precedents, custom, expert opinion, and morality as sources of law rather than as law itself. All law, on this view, is judge-made. The machinery of the state stands in the background and provides the coercive element, which does not enter into the definition of "law." Gray influenced the realist movement in the United States.

HEGELIANISM AND THE HISTORICAL SCHOOL. While England was largely under the sway of the utilitarians, other trends flourished in Germany: Kantianism, Hegelianism, the historical school, and legal positivism. In his *Philosophy of Right*, G. W. F. Hegel (1770–1831) developed some Kantian themes in his own characteristic way. In his view, law and social-political institutions belong to the realm of "objective spirit," in which interpersonal relationships, reflecting an underlying freedom, receive their concrete manifestations. In attempting to show the rightness and the rationality of various legal relationships and institutions in given moments of the development of "spirit," and in seeing them as natural growths, Hegel formulated a theory of law and the state that was easily combined with various historical, functional, and institutional approaches to legal phenomena.

Friedrich Karl von Savigny (1779–1861) is often regarded as the founder of the historical school. His *Of the Vocation of Our Age for Legislation and Jurisprudence* (1814) was published before Hegel's work and was probably influenced by Fichte (but not by Fichte's *Grundlage des Naturrechts*, 1796), whose notion of the "folk-spirit" was widely known. Law, like language, originates spontaneously in the common consciousness of a people, who constitute an organic being. Both the legislator and the jurist may articulate this law, but they no more invent or make it than does the grammarian who codifies a natural language. Savigny believed that to accept his conception of law was to reject the older notions of natural law; nevertheless, it is often claimed that Savigny's conception was merely a new kind of natural law standing above, and judging, the positive law.

THE LATE-NINETEENTH AND TWENTIETH CENTURIES

JHERING AND GERMAN POSITIVISM. Rudolf von Jhering (1818–1892), eminent both as a historian of law and as a legal theorist, rejected both Hegel and Savigny: Hegel, for holding the law to be an expression of the general will and for failing to see how utilitarian factors and interests determine the existence of law; Savigny, for regarding law as a spontaneous expression of subconscious forces and for failing to see the role of the conscious struggle for protection of interests. However, Jhering shared the broad cultural orientation of many of the Hegelians, and he was grateful to Savigny for having overthrown the doctrine of "immutable" natural law. Jhering's contribution was to insist that legal phenomena cannot be comprehended without a systematic understanding of the purposes that give rise to them, the ends grounded in social life without which there would be no legal rules. Without purpose there is no will.

At the same time there are strong strains of positivism in Jhering: Law is defined as "the sum of the rules of constraint which obtain in a state" (*Der Zweck im Recht*, p. 320). In this respect he was close to the German positivists, who emphasized the imperative character of law. Karl Binding (1841–1920), an influential positivist, defined law as "only the clarified legal volition [*Rechtswille*] of a source of law [*Rechtsquelle*]" (*Die Normen und ihre Ueber-tretung*, p. 68). This period saw the emergence of the slogan of German positivism: "All law is positive law." Yet Jhering opposed many of the claims of the analytical positivists; his essay "Scherz und Ernst in der Jurisprudenz" ridiculed their "heaven of jurisprudential concepts."

SOCIOLOGICAL AND ALLIED THEORIES. Jhering's work foreshadowed many of the dominant tendencies of

twentieth-century legal philosophy. Hermann Kantorowicz regarded Jhering as the fountainhead of both the "sociological" and "free-law" schools. The former term covers too wide a group of writers to be surveyed here, some of whom were concerned solely with empirical work, whereas others combined empirical work with a philosophical outlook. Proponents of the jurisprudence of interests (*Interessenjurisprudenz*) eschewed Jhering's inquiries into the metaphysical and moral bases of purposes, claiming that he did not sufficiently attend to the conflict of interest behind laws; law reflects dominant interest. (Similar analyses were made in the United States; for example, the "pressure-group" theory of politics advanced by A. F. Bentley [1870–1957] in *The Process of Government*). Much attention was devoted to the analysis of the judicial process and the role that the "balancing" of interests plays in it. As Philipp Heck, one of its leading exponents, remarked: "The new movement of 'Interessenjurisprudenz' is based on the realization that the judge cannot satisfactorily deal with the needs of life by mere logical construction" (*Begriffsbildung und Interessenjurisprudenz*, p. 4).

This sentiment was endorsed by the closely allied "free-law" movement. According to this group, "legal logic" and the "jurisprudence of conceptions" are inadequate for achieving practicable and just decisions. The judge not only perforce frequently goes beyond the statute law, but he also often ought to go beyond it. The "free-law" writers undertook the normative task of supplying guidelines for the exercise of judicial discretion, and the judicial function was assimilated to the legislative function. The focus on such problems reflected the enormous change, occasioned by the industrialization of Western society, in the functions of the state. No longer did the nation-state exist merely to keep the peace or protect preexisting rights; rather, it played a positive role in promoting social and individual welfare.

The philosophy of law thus became increasingly concerned with the detailed working out of the foundations of legal policy. The "free-law" theorist Eugen Ehrlich (1862–1922), who influenced such American theorists as Karl N. Llewellyn (1893–1962) and other representatives of legal realist tendencies, rejected the positivistic tenet that only norms posited by the state are legal norms, for in any society there is always more law than is expressed in legal propositions. The "inner order" of an association is the basic form of law. Ehrlich also engaged in empirical study of the "legal facts" (*Rechtstatsachen*) and "living law" of various communities in the Austro-Hungarian empire. Ehrlich may thus be said to have considered custom as law in its own right. However, many positivists would argue that he was not able to account for the normative character of custom.

MARXISM. The Marxist stress on economic interests was often combined with the sociological and free-law views. Central to the Marxist position are the notions of "class" (usually defined in terms of legal relationship to property and the means of production) and "class interest," which lead to the analysis of the role of law in different societies with differing class structures. Addressing their critics in *The Communist Manifesto* (1848), Marx and Engels wrote: "Your law is but the will of your class exalted into statutes, a will which acquires its content from the material conditions of existence of your class" (p. 24). This suggests that law is merely part of the ideological superstructure and has no effect on the material organization of society. It raises the question of whether law exists in all societies—for instance, in primitive society or in the "classless" society arising after the triumph of socialism—and the further question of the nature and function of law in the transitional period from capitalism to socialism. The issue of "revolutionary legality" or "socialist legality" was treated by Lenin, E. Pashukanis, and Andrei Vishinsky. An important Marxist study of the relationship between law and the economy is that of the Austrian socialist Karl Renner (*Die Rechtsinstitute des Privatrechts und ihre soziale Funktion*, 1929).

PURE THEORY AND RELATIVISM. Although the sociological approaches to law had many practitioners, a more controversial and perhaps more influential twentieth-century view was that of Hans Kelsen, a leading exponent of legal positivism. Influenced by the epistemology of the Neo-Kantians, Kelsen distinguished sharply between the "is" and the "ought," and consequently between the natural sciences and disciplines, such as legal science, which study "normative" phenomena. Legal science is a descriptive science—prescriptive and valuational questions cannot be scientific—and Kelsen's "pure theory" aimed at providing the conceptual tools for studying any given legal system irrespective of its content. The theory is "pure" because it is divorced from any ideological or sociological elements; it attempts to treat a legal system simply as a system of norms. Kelsen's view was thus similar to the analytical jurisprudence of Austin, but Kelsen regarded legal norms as "de-psychologized commands." In order to understand an act of will as a norm-creating act, we must already employ a norm which serves as a "schema of interpretation." The jurist who seeks to understand legal phenomena must ultimately presuppose

a basic norm (*Grundnorm*), which is not itself a positive legal norm. Legal systems are sets of coercive norms arranged in hierarchical fashion; lower norms are the "concretizations" of higher norms. In Kelsen's analysis the "dualisms" of state and law and public and private law disappear, and the relationship between international law and national legal systems is seen in a fresh light.

Unlike Kelsen, Gustav Radbruch (1878–1949) did not found a school. His position, which he called relativism, has many affinities with that of Kelsen; but Radbruch maintained that law, which is a cultural phenomenon, can be understood only in relation to the values that men strive to realize through it. He attempted to analyze these values in relation to legal institutions, showing the "antinomies" among these values that led to his relativism. World War II raised the question in the minds of many legal philosophers whether the legal positivism that was popular in Germany, with its separation of law and morals, contributed to the rise of Nazism. Concern over this problem seems to have caused Radbruch to move away from his earlier relativism toward a kind of natural-law position.

REALISM. In the United States, legal philosophy had largely been the province of lawyers rather than of professional philosophers. This may account for its sociological and realistic tone. The erudite Roscoe Pound (1870–1964) was the most prolific writer on this subject. Pound recognized the influence of Josef Kohler (1849–1919) and his notion of jural postulates and, especially, of Jhering. The pragmatism of William James also contributed to the development of his views. In an early article, "Mechanical Jurisprudence," Pound argued for an understanding of the interests that the law seeks to protect. Introducing a distinction between "law in books" and "law in action," he maintained the need for a close study of the actual operation of legal institutions.

On both scores his influence in the United States has been momentous, but it is difficult to summarize his position; he is often associated with a "social engineering" approach to law. Law contains both precepts and ideal elements. Among precepts Pound distinguished rules, principles, conceptions, doctrines, and standards. It is pointless to isolate some canonical form to which all laws are reducible. The ideal element consists of received ideals "of the end of law, and hence of what legal precepts should be and how they should be applied." Pound offered an elaborate, although tentative, survey of the individual, public, and social interests secured by law. This list was criticized and amended by Pound's Australian disciple Julius Stone (*The Province and Function of Law*, 1946). In his later years Pound moved toward a kind of natural-law thinking, arguing for a more intimate connection between law and morality; he abjured the realist tendencies, which had been influenced by his earlier thought, as "give it up" philosophies.

It is difficult to characterize the legal realists; they disclaimed a common doctrine but recognized an interest in a common set of problems. Along with J. C. Gray, the spiritual godfather of American legal realism was Justice Oliver Wendell Holmes, Jr. (1841–1935). In his seminal essay "The Path of the Law," he advocated viewing law as the "bad man" would, in terms of the practicable remedies afforded individuals through the medium of the courts. Holmes presented in that article his famous definition of law as "the prophecies of what the courts will do in fact." It may be argued, however, that this definition, while perhaps adequate from the advocate's viewpoint, can hardly apply to the judge. When the judge asks what the law is on some matter, he is not trying to predict what he will decide.

Joseph W. Bingham was one of the first realists. In "What Is the Law?" Bingham argued that legal rules, like scientific laws, have no independent existence, being simply mental constructs that conveniently summarize particular facts. Laws are really judicial decisions, and the so-called rules or principles are among the (mentally) causative factors behind the decision. This nominalism and behaviorism, which characterized much of early realist writing, was criticized by Morris R. Cohen (1880–1947). "Behavior analysis" was advocated by Karl N. Llewellyn, who extended it beyond judicial behavior to "official" behavior (*Jurisprudence*, Chicago, 1962; collected papers).

The so-called myth of legal certainty was attacked by Jerome Frank (1889–1957) in his *Law and the Modern Mind*, which explained the genesis of the myth in Freudian terms. In the sixth edition Frank was somewhat friendlier toward natural-law thinking, characterizing his change of attitude as going from an earlier "rule-skepticism" to "fact-skepticism." Other important realists are Thurman Arnold, Leon Green, Felix Cohen, Walter Nelles, Herman Oliphant, and Fred Rodell. Both positivism and realism were attacked by the Harvard legal philosopher Lon L. Fuller (1902–1978), a leading American exponent of non-Thomistic natural-law thinking.

The Scandinavian countries were a center of legal philosophy, and many of their leading writers have been realists. They have been more consciously philosophical than their American counterparts. The leading spirit was

Axel Hägerström (1868–1939), who rejected metaphysical presuppositions in legal philosophy and insisted on an understanding of legal phenomena in empirical terms. Many legal concepts can be understood only as survivals of "mythical" or "magical" thought patterns, which should ideally be eliminated. Vilhelm Lunstedt (*Legal Thinking Revised*, Stockholm, 1956) was most radical in his rejection of metaphysics. Values are expressions of emotion and should be excluded from legal science. The "method of social welfare" should be substituted for the "method of justice." Alf Ross (*On Law and Justice*, London, 1958) argued that the first method is as "chimerical" as the second and presents an analysis of legal policymaking as a kind of rational technology. Laws, Ross argued, are directives to courts. The concept "valid law" as used by jurists and legal philosophers cannot be explicated in purely behavioristic terms; inner psychological attitudes must also be included. A similar view was presented by Karl Olivecrona (*Law as Fact*, London, 1939), who wrote important realist analyses of legal language and severely criticized command theories of law, such as Austin's. In *Inquiries Into the Nature of Law and Morals* (translated by C. D. Broad, Cambridge, 1953), Hägerström argued that Kelsen's "pure theory" never escapes the "will" element either, and hence it falls subject to all the criticisms that may be leveled against the command theories.

CRITICAL LEGAL STUDIES. The critical legal studies (CLS) movement, associated with the work of Duncan Kennedy, among many others, borrowed much from legal realism. CLS scholars shared the rule-skepticism of the realists and their rejection of legal formalism. Both groups emphasized the role played by extra-legal factors in shaping the law. For CLS scholars, however, the realists did not go far enough in developing a "critique" of the ideological bias concealed within legal doctrines. A central preoccupation of the critical scholars was the indeterminacy and vagueness of the law. CLS writers attempted to "deconstruct" the law by exposing its inconsistencies and tracing them to the conflicting social and economic forces responsible for shaping it.

H. L. A. HART AND POSITIVISM. One of the most influential legal theorists of the last half of the twentieth century, H. L. A. Hart (1907–1992), in his *Concept of Law*, developed the view that the law is a union of "primary" and "secondary" rules. Primary rules impose duties; secondary rules specify how primary rules may be changed, interpreted, and recognized as valid. A rule of recognition specifies what is to count as law in a given system. In a series of works beginning with *Taking Rights Seriously*, Ronald Dworkin attacked Hart's theory, maintaining that when courts reason about "hard" cases, they invoke standards or principles that cannot be captured by a Hartian rule of recognition. Principles (such as "no man should profit from his own wrongdoing") are part of the law, Dworkin argued, and so Hart's positivism is descriptively inaccurate. In *Law's Empire*, Dworkin argued that law is an "interpretive" concept, so that a judge facing a difficult case must seek to identify the best "constructive interpretation" of the legal doctrine of his community, viewing the legal materials normatively, in their "best light." Law is the product of an interpretation that best sums up the legal texts and principles of a given community into a coherent and attractive whole.

Hart's work spurred debate among legal positivists regarding the proper understanding of Hart's rule of recognition. Hart maintained that moral norms are not necessarily a part of the criteria for the validity of law. But could there be legal systems that *do* incorporate moral criteria of legal validity? "Exclusive" legal positivists, such as Joseph Raz, responded in the negative; "inclusive" positivists, such as Jules Coleman, answered affirmatively.

LATE-TWENTIETH-CENTURY DEVELOPMENTS. A resurgence of interest in natural law characterized the end of the twentieth century, with works by Robert George (*In Defense of Natural Law*) and especially John Finnis, beginning with his *Natural Law and Natural Rights*. Several other prominent jurisprudential "schools" also emerged in the last decades of the twentieth century. Among these were feminist jurisprudence and the law and economics movement. Work by Catherine MacKinnon and other feminist lawyers sought to expose the patriarchal assumptions underlying purportedly neutral legal doctrine; and scholars led by scholar and judge Richard Posner argued that an economic analysis of the formation and function of legal rules and doctrines provides the best explanation for existing law.

See also Aristotelianism; Aristotle; Augustine, St.; Austin, John; Beccaria, Cesare Bonesana; Bentham, Jeremy; Bodin, Jean; Burke, Edmund; Celsus; Cicero, Marcus Tullius; Cohen, Morris Raphael; Dworkin, Ronald; Engels, Friedrich; Enlightenment; Fichte, Johann Gottlieb; Filmer, Robert; Grotius, Hugo; Hägerström, Axel; Hart, Herbert Lionel Adolphus; Hegel, Georg Wilhelm Friedrich; Hegelianism; Hobbes, Thomas; Hume, David; James, William; Just War Theory; Kant, Immanuel; Kelsen, Hans; Legal Positivism; Lenin, Vladimir Il'ich; Locke, John; Marx, Karl; Marxist Phi-

losophy; Mill, John Stuart; Montesquieu, Baron de; Natural Law; Neo-Kantianism; Ockhamism; Paley, William; Patristic Philosophy; Plato; Positivism; Posner, Richard; Radbruch, Gustav; Realism; Rousseau, Jean-Jacques; Savigny, Friedrich Karl von; Seneca, Lucius Annaeus; Social Contract; Socrates; Sophists; Stammler, Rudolf; Stoicism; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Utilitarianism; Vico, Giambattista; Vitoria, Francisco de; William of Ockham; Xenophon.

Bibliography

Christie, George C., and Patrick H. Martin. *Jurisprudence: Text and Readings on the Philosophy of Law*. 2nd ed. St. Paul: West, 1995.

Coase, Ronald. "The Problem of Social Cost." *Journal of Law and Economics* 3 (1960): 1–44.

Cohen, M. R. *Law and the Social Order*. Hamden: Archon Books, 1967.

Coleman, Jules. "Negative and Positive Positivism," *Journal of Legal Studies* 11 (1982): 139–164.

Coleman, Jules. and Scott Shapiro, eds. *The Oxford Handbook of Jurisprudence and Legal Philosophy*. Oxford: Oxford University Press, 2002.

Del Vecchio, Giorgio. *Justice*. Translated by Lady Guthrie. Edinburgh: University Press, 1956.

D'Entreves, A. P. *Natural Law*. London: Hutchinson, 1977.

Duxbury, Neil. *Patterns of American Jurisprudence*. Oxford: Clarendon Press, 1995.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

Dworkin, Ronald. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press, 1977.

Ehrlich, Eugen. *Fundamental Principles of the Sociology of Law*. Translated by Walter L. Moll. New York: Arno Press, 1975.

Finnis, John. *Aquinas: Moral, Political, and Legal Theory*. New York: Oxford University Press, 1998.

Finnis, John. *Natural Law and Natural Rights*. Oxford: Clarendon Press, 1980.

Friedrich, C. J. *The Philosophy of Law in Historical Perspective*. 2nd ed. Chicago: University of Chicago Press, 1965.

Gierke, Otto von. *Das deutsche Genossenschaftsrecht*. Vols. 1–4. Berlin: Weidmannsche Verlagsbuchhandlung, 1868–1913. Partially translated by Ernest Barker as *Natural Law and the Theory of Society*. Cambridge, U.K.: Cambridge University Press, 1958.

George, Robert P. *Natural Law Theory: Contemporary Essays*. Oxford: Clarendon Press, 1992.

Golding, Martin, and William Edmundson, eds. *Philosophy of Law and Legal Theory*. Oxford: Blackwell, 2005.

Hall, Jerome. *Studies in Jurisprudence and Criminal Theory*. New York: Oceana, 1958.

Hamburger, Max. *The Awakening of Western Legal Thought*. Translated by B. Miall. Westport: Greenwood Press, 1970.

Hart, H. L. A. *The Concept of Law*. 2nd ed. New York: Oxford University Press, 1994.

Hart, H. L. A. *Essays in Jurisprudence and Philosophy*. Oxford: Clarendon Press, 1983.

Hart, H. L. A. "Philosophy of Law and Jurisprudence in Britain (1945–1952)." *American Journal of Comparative Law* 2 (1953): 355–364.

Hayman, Robert L., Jr., et al. *Jurisprudence Classical and Contemporary: From Natural Law to Postmodernism*. 2nd ed. St. Paul: West, 2002.

Jhering, Rudolf von. *Law as a Means to an End*. Translated by I. Husik. New York: A. M. Kelley, 1968.

Jones, J. W. *Historical Introduction to the Theory of Law*. Westport: Greenwood Press, 1970.

Kairys, David. *The Politics of Law*. New York: Pantheon Books, 1982.

Kelly, J. M. *A Short History of Western Legal Theory*. Oxford: Oxford University Press, 1992.

Kennedy, Duncan. "Form and Substance in Private Law Adjudication." *Harvard Law Review* 89 (1976): 1685–1778.

King, Peter J., *Utilitarian Jurisprudence in America: The Influence of Bentham and Austin on American Legal Thought in the Nineteenth Century*. New York: Garland, 1986.

Kocourek, Albert. "The Century of Analytical Jurisprudence Since John Austin." In *Law: A Century of Progress*, edited by A Reppy. New York: New York University Press, 1937.

Lobban, Michael. *The Common Law and English Jurisprudence, 1760–1850*. Oxford: Clarendon press, 1991.

Mackinnon, Catharine A. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press, 1989.

Maine, Henry S. *Ancient Law*. Boston: Beacon Press, 1963.

Minow, Martha. *Making All the Difference*. Ithaca: Cornell University Press, 1991.

Morison, W. L. *John Austin*. Stanford: Stanford University Press, 1982.

Morris, C. "Four Eighteenth-Century Theories of Jurisprudence." *Vanderbilt Law Review* 14 (1960): 101–116.

Patterson, Dennis, ed. *A Companion to Philosophy of Law and Legal Theory*. Oxford: Blackwell, 1996.

Pollock, Frederick. *Jurisprudence and Legal Essays*. Westport: Greenwood Press, 1978.

Posner, Richard A. *Economic Analysis of the Law*. 6th ed. New York: Aspen, 2002.

Postema, Gerald. *Bentham and the Common Law Tradition*. Oxford: Clarendon, 1986.

Pound, Roscoe. *Jurisprudence*. 5 vols. St. Paul: West, 1959.

Raz, Joseph. *The Authority of Law*. Oxford: Oxford University Press, 1979.

Raz, Joseph. *The Concept of a Legal System*. Oxford: Clarendon, 1970.

Rhode, Deborah. *Justice and Gender*. Cambridge, MA: Harvard University Press, 1991.

Rommen, Heinrich. *The Natural Law*. Translated by T. R. Hanley. Indianapolis: Liberty Fund, 1998.

Shuchman, Philip. *Cohen and Cohen's Readings in Jurisprudence and Legal Philosophy*. 2nd ed. Boston: Little, Brown, 1979.

Smith, Patricia. *Feminist Jurisprudence*. New York: Oxford University Press, 1993.

Somek, Alexander. "German Legal Philosophy and Theory in the Nineteenth and Twentieth Centuries." In *A Companion*

to Philosophy of Law and Legal Theory, edited by Dennis Patterson. Oxford: Blackwell, 1996.

Stone, Julius. *The Province and Function of Law*. London: Stevens Press, 1961.

Tushnett, Mark. "Critical Legal Studies: A Political History." *Yale Law Journal* 100 (1991): 1515–1544.

Ullmann, Walter, *The Mediaeval Idea of Law*. New York: Barnes and Noble, 1969.

Unger, Roberto Magabeira, *The Critical Legal Studies Movement*. Cambridge, MA: Harvard University Press, 1983.

Villey, M. *Leçons d'histoire de la philosophie du droit*. Paris: Dalloz, 1957.

Vinogradoff, Paul. *Outlines of Historical Jurisprudence*. 2 vols. New York: AMS Press, 1971.

M. P. Golding (1967)
Updated by David M. Adams (2005)

PHILOSOPHY OF LAW, PROBLEMS OF

The existence of legal systems, even the most rudimentary, has afforded the opportunity for a variety of academic disciplines. Of these some are, or purport to be, empirical: They include the historical study of particular legal systems or specific legal doctrines and rules, and sociological studies of the ways in which the content and the efficacy of law and the forms and procedures of law-making and law-applying both influence and are influenced by their economic and social setting, and serve social needs or specific social functions. But since law in most societies soon reaches a very high degree of complexity, its administration requires the special training of judges and professional lawyers. This in turn has created the need for a specific form of legal science concerned with the systematic or dogmatic exposition of the law and its specific methods and procedures. For this purpose the law is divided into distinct branches (such as crime, tort, and contract), and general classifications and organizing concepts are introduced to collect common elements in the situations and relationships created by the law (such as rights, duties, obligations, legal personality, ownership, and possession) or elements common to many separate legal rules (such as act and intention).

No very firm boundaries divide the problems confronting these various disciplines from the problems of the philosophy of law. This is especially true of the conceptual schemes of classification, definition, and division introduced by the academic study of the law for the purpose of exposition and teaching; but even some historical and sociological statements about law are sufficiently general and abstract to need the attention of the philosophical critic. Little, however, is to be gained from elaborating the traditional distinctions between the philosophy of law, jurisprudence (general and particular), and legal theory, although importance has often been attributed to them. Instead, as with other branches of philosophy, it is more important to distinguish as belonging to the philosophy of law certain groups of questions which remain to be answered even when a high degree of competence or mastery of particular legal systems and of the empirical and dogmatic studies mentioned above has been gained. Three such groups may be distinguished: problems of definition and analysis, problems of legal reasoning, and problems of the criticism of law. This division is, however, not uncontroversial; and objections to it are considered in the last section of the article.

PROBLEMS OF DEFINITION AND ANALYSIS

THE DEFINITION OF LAW. All the obscurities and prejudices that in other areas of philosophy surround the notions of definition and of meaning have contributed to the endlessly debated problems of the definition of law. In early arguments the search for the definition of law was assumed to be the task of identifying and describing the "essence" or "nature" of law, and thus the uniquely correct definition of law by reference to which the propriety of the use, however well established, of the expressions "law" and "legal system" could be tested. It is frequently difficult to distinguish from this search for the essence of law a more modest conception of definition that, while treating the task as one of identifying and describing the standards actually accepted for the use of these expressions, assumes that there is only one "true," "strict," or "proper" use of them and that this use can be described in terms of a single set of necessary and sufficient conditions. A wide range of different considerations has shown how unrealistic or how sterile this assumption is in the case of law and has compelled its surrender. Among these considerations is the realization that although there are central clear instances to which the expressions "law" and "legal system" have undisputed application, there are also cases, such as international law and primitive law, which have certain features of the central case but lack others. Also, there is the realization that the justification for applying general expressions to a range of different cases often lies not in their conformity to a set of necessary and sufficient conditions but in the analogies that link them or their varying relationships to some single element.

Lexical definitions and deviant cases. The foregoing are difficulties of definition commonly met in many areas of philosophy, but the definition of law has peculiar diffi-

culties of its own. Thus, the assumption that the definition of law either has been or should be lexical, that is, concerned with the characterization or elucidation of *any* actual usage, has been challenged on several grounds. Thus it is often asserted that in the case of law, the area of indeterminacy of actual usage is too great and relates to too many important and disputed issues, and that what is needed is not a characterization or elucidation of usage but a reasoned case for the inclusion in or exclusion from the scope of the expressions "law" and "legal system" of various deviations from routine and undisputed examples. These deviant cases include not only international law and primitive law but also certain elements found in developed municipal legal systems, such as rules to which the usual sanctions are not attached and rules that run counter to fundamental principles of morality and justice.

Pragmatic definitions. In the above circumstances some theorists disclaim as necessarily deceptive any aim to provide an analysis or definition of law which is a neutral description or elucidation of usage; instead, they speak of the task of definition as "stipulative," "pragmatic," or "constructive," that is, as designed to provide a scheme or model for the demarcation and classification of an area of study. The criterion of adequacy of such pragmatic definitions is not conformity to or the capacity to explain any actual usage but the capacity to advance the theorists' specific aims, which may differ widely. Thus, a definition of law to be used for the instruction or assistance of lawyers concerned primarily with the outcome of litigation or court proceedings will differ from the definition used to demarcate and unify the fruitful area of historical study and will also differ from the definition to be used by the social critic concerned with identifying the extent to which human interests are advanced or frustrated by modes of social organization and control.

Structural problems. Neither the legitimacy of pragmatic definitions nor their utility for deliberately chosen objectives need be disputed. But it is clear that they avoid rather than resolve many of the long-standing perplexities that have motivated requests for the definition of law and have made it a philosophical problem. The factors that have generated these perplexities may be summarized as follows: Notwithstanding the considerable area of indeterminacy in their use, the expressions "law," "a law," "legal system," and a wide range of derivative and interrelated expressions ("legislation," "courts of law," "the application of law," "legal adjudication") are sufficiently determinate to make possible general agreement in judgments about their application to particular instances. But reflection on what is thus identified by the common usage of such terms shows that the area they cover is one of great internal complexity; laws differ radically both in content and in the ways in which they are created, yet despite this heterogeneity they are interrelated in various complex ways so as to constitute a characteristic structure or system. Many requests for the definition of law have been stimulated by the desire to obtain a coherent view of this structure and an understanding of the ways in which elements apparently so diverse and unified. These are problems, therefore, of the structure of law.

Coercion and morality. Reflection on the operations of a legal system discloses problems of another sort, for it is clear that law as a mode of influence on human behavior is intimately related to and in many ways dependent upon the use or threat of force on the one hand and on morality and justice on the other. Yet law is also, at points, distinct from both, so no obvious account of these connections appears acceptable: They appear to be not merely contingent, and since they sometimes fail, the statement of these connections does not appear to be any easily comprehensible species of necessary truth. Such tensions create demands for some stable and coherent definition of the relationships between law, coercion, and morality; but definitions of law have only in part been designed to make these important areas of human experience more intelligible. Practical and indeed political issues have long been intertwined with theoretical ones; and as is evident from the long history of the doctrines of natural law and legal positivism, the advocacy of a submissive or a critical attitude to law, or even of obedience or disobedience, has often been presented in the form of a persuasive definition of the relationship between law and morality on the one hand and between law and mere force on the other.

THE ANALYSIS OF LEGAL CONCEPTS. Although legal rules are of many different types and may be classified from many different points of view, they have many common constituents; and although the law creates for both individuals and groups a great variety of different situations and relationships, some of these are constantly recurrent and of obvious importance for the conduct of social life. Both lawyers and laypeople have frequent occasion to refer to these common elements and situations, and for this purpose they use classifications and organizing concepts expressed in a vocabulary which has bred many problems of analysis. These problems arise in part because this vocabulary has a more or less established use apart from law, and the points of convergence and divergence between legal and nonlegal usage is not always

immediately obvious or easily explicable. It is also the case that the ways in which common elements in law or legal situations are classified by different theorists in part reflect and derive from divergent conceptions of law in general. Therefore, although different writers use such expressions as "rights" and "duty" in referring to the same legal situations, they select different elements or aspects from these situations. A third factor calling for clarification is the fact that many of the commonest notions used in referring to legal phenomena can be explicated only when certain distinctive ways in which language functions in conjunction with practical rules have been understood. These problems of analysis are illustrated in the case of the concepts of (1) legal obligation or duty, (2) a legal transaction, and (3) intention. (Certain distinctions once made between the notions of a legal obligation and a legal duty are no longer of importance and will be disregarded.)

Legal obligations or duties. The situation in which an individual has a legal duty to do or to abstain from some action is the commonest and most fundamental of all legal phenomena; the reference to duty or its absence is involved in the definition of such other legal concepts as those of a right, a power, a legal transaction, or a legal personality. Whenever the law of an effective legal system provides for the punishment of those who act or fail to act in certain ways, the word *duty* applies. Thus, to take a simple example, if the law requires under penalty that persons of a certain age shall report for military service, then such persons have, or are "under," a legal duty to do so. This much is undisputed, however much theorists may dispute over the analysis of "duty" or its application to situations created not by the criminal law but by the law relating to torts or to contract.

However, even the above simple situation can be viewed from two very different standpoints that give rise to apparently conflicting analyses of duty. From one of these (the predictive standpoint), reporting for military service is classified as a duty simply because failure to report renders likely certain forms of suffering at the hands of officials. From the other standpoint (the normative standpoint), reporting for military service is classified as a duty because, owing to the existence of the law, it is an action that may be rightly or justifiably demanded of those concerned; and failure to report is significant not merely because it renders future suffering likely but also because punishment is legally justified even if it does not always follow disobedience.

From Jeremy Bentham onward the predictive analysis of duty as a chance or likelihood of suffering in the event of disobedience to the law has been advocated by important writers for a variety of theoretical and practical reasons. On the one hand it has seemed to free the idea of legal duty from metaphysical obscurities and irrelevant associations with morals, and on the other to provide a realistic guide to life under law. It isolates what for some people is the only important fact about the operation of a legal system and what for all people is at least one important fact: the occasions and ways in which the law works adversely to their interests. This is of paramount importance not only to the malefactor but also to the critic and reformer of the law concerned to balance against the benefits which law brings its costs in terms of human suffering.

By contrast, the normative point of view, without identifying moral and legal duty or insisting on any common content, stresses certain common formal features that both moral and legal duty possess in virtue of their both being aspects of rule-guided conduct. This is the point of view of those who, although they may not regard the law as the final arbiter of conduct, nevertheless generally accept the existence of legal rule as a guide to conduct and as legally justifying demands for conformity, punishment, enforced compensation, or other forms of coercion. Attention to these features of the idea of duty is essential for understanding the ways in which law is conceived of and operative in social life.

Although theorists have often attributed exclusive correctness to these different standpoints, there are various ways in which they may be illuminatingly combined. Thus, the normative account might be said to give correctly the meaning of such statements as that a person has a legal duty to do a certain action, while the predictive account emphasizes that very frequently the point or purpose of making such statements is to warn that suffering is likely to follow disobedience. Such a distinction between the meaning of a statement and what is implied or intended by its assertion in different contexts is of considerable importance in many areas of legal philosophy.

Legal transactions. The enactment of a law, the making of a contract, and the transfer by words, written or spoken, of ownership or other rights are examples of legal transactions which are made possible by the existence of certain types of legal rules and are definable in terms of such rules. To some thinkers, such transactions (acts in the law, or juristic acts) have appeared mysterious—some have even called them magical—because their effect is to change the legal position of individuals or to make or eliminate laws. Since, in most modern systems of law, such changes are usually effected by the use of words, written or spoken, there seems to be a species of legal

alchemy. It is not obvious how the mere use of expressions like "it is hereby enacted …," "I hereby bequeath …," or "the parties hereby agree …" can produce changes. In fact, the general form of this phenomenon is not exclusively legal, although it has only comparatively recently been clearly isolated and analyzed. The words of an ordinary promise or those used in a christening ceremony in giving a name to a child are obvious analogues to the legal cases. Lawyers have sometimes marked off this distinctive function of language as the use of "operative words," and under this category have distinguished, for example, the words used in a lease to create a tenancy from the merely descriptive language of the preliminary recital of the facts concerning the parties and their agreement.

For words (or in certain cases gestures, as in voting or other forms of behavior) to have such operative effect, there must exist legal rules providing that if the words (or gestures) are used in appropriate circumstances by appropriately qualified persons, the general law or the legal position of individuals is to be taken as changed. Such rules may be conceived from one point of view as giving to the language used a certain kind of force or effect which is in a broad sense their meaning; from another point of view they may be conceived as conferring on individuals the legal power to make such legal changes. In Continental jurisprudence such rules are usually referred to as "norms of competence" to distinguish them from simpler legal rules that merely impose duties with or without correlative rights.

As the expressions "acts-in-the-law" and "operative words" suggest, there are important resemblances between the execution of legal transactions and more obvious cases of human actions. These points of resemblance are of especial importance in understanding what has often seemed problematic—the relevance of the mental or psychological states of the parties concerned to the constitution or validity of such transactions. In many cases the relevant rules provide that a transaction shall be invalid or at least liable to be set aside at the option of various persons if the person purporting to effect it was insane, mistaken in regard to certain matters, or subjected to duress or undue influence. There is here an important analogy with the ways in which similar psychological facts (*mens rea*) may, in accordance with the principles of the criminal law, excuse a person from criminal responsibility for his action. In both spheres there are exceptions: In the criminal law there are certain cases of "strict" liability where no element of knowledge or intention need be proved; and in certain types of legal transaction, proof that a person attached a special meaning to the words he used or was mistaken in some respect in using them would not invalidate the transaction, at least as against those who have relied upon it in good faith.

Attention to these analogies between valid legal transactions and responsible action and the mental conditions that in the one case invalidate and in the other excuse from responsibility illuminates many obscure theoretical disputes concerning the nature of legal transactions such as contract. Thus, according to one principal theory (the "will" theory) a contract is essentially a complex psychological fact—something that comes into being when there is a meeting of minds (*consensus ad idem*) that jointly "will" or "intend" a certain set of mutual rights and duties to come into existence. The words used are, according to this theory, merely evidence of this consensus. The rival theory (the "objective" theory) insists that what makes a contract is not a psychological phenomenon but the actual use of words of offer and acceptance, and that except in special cases the law simply gives effect to the ordinary meaning of the language used by the parties and is not concerned with their actual states of mind. Plainly, each side to this dispute fastens on something important but exaggerates it. It is indeed true that, like an ordinary promise, a legal contract is not made by psychological facts. A contract, like a promise, is "made" not by the existence of mental states but by words (or in some cases deeds). If it is verbally made, it is made by the operative use of language, and there are many legal rules inconsistent with the idea that a *consensus ad idem* is required. On the other hand, just because the operative use of language is a kind of action, the law may—and in most civilized legal systems does—extend to it a doctrine of responsibility or validity under which certain mental elements are made relevant. Thus a contract, although made by words, may be vitiated or made void or "voidable" if a party is insane, mistaken in certain ways, or under duress. The truths latent among the errors of the "will" theory and the "objective" theory can therefore be brought together in an analysis that makes explicit the analogy between valid transactions made by the operative use of language and responsible actions.

Intention. The fact that the law often treats certain mental states or psychological conditions as essential elements both in the validity of legal transactions and in criminal responsibility has thrust upon lawyers the task of distinguishing between and analyzing such notions as "will," "intention," and "motive." These are concepts that have long puzzled philosophers not primarily concerned with the law, and their application in the law creates fur-

ther specific problems. These arise in various ways: There are divergencies between the legal and nonlegal use of these notions which are not always obvious or easily understood; the law, because of difficulties of proof or as a matter of social policy, may often adopt what are called external or objective standards, which treat certain forms of outward behavior as conclusive evidence of the existence of mental states or impute to an individual the mental state that the average man behaving in a given way would have had. Although statutes occasionally use such expressions as "maliciously," "knowingly," or "with intent," for the most part the expressions "intentionally" and "voluntarily" are not the language of legal rules but are used in the exposition of such rules in summarizing the various ways in which either criminal charges or civil claims may fail if something is done—for instance, accidentally, by mistake, or under duress.

The problems that arise in these ways may be illustrated in the case of intention. Legal theorists have recognized intention as the mental element of central importance to the law. Thus, an intention to do the act forbidden by law is in Anglo American law normally the sufficient mental element for criminal responsibility and also is normally, although not always, necessary for responsibility. So if a man intends to do the act forbidden by law, other factors having to do with his powers of self-control are usually irrelevant, although sometimes duress and sometimes provocation or deficient ability to control conduct, caused by mental disorder, may become relevant. In fact, three distinct applications of the notion of intention are important in the law, and it is necessary to distinguish in any analysis of this concept (1) the idea of intentionally doing something forbidden by law; (2) doing something with a further intention; and (3) the intention to do a future act. The first of these is in issue when, if a man is found to have wounded or killed another, the question is asked whether he did it intentionally or unintentionally. The second is raised when the law, as in the case of burglary defined as "breaking into premises at night with the intention of committing a felony," attaches special importance or more severe penalties to an action if it is done for some further purpose, even though the latter is not executed. The third application of intention can be seen in those cases where an act is criminal if it is accompanied by a certain intention—for instance, incurring a debt with the intention never to pay.

Of these three applications the first is of chief importance in the law, but even here the law only approximates to the nonlegal concept and disregards certain elements in its ordinary usage. For in the law the question whether a man did something intentionally or not is almost wholly a question concerning his knowledge or belief at the time of his action. Hence, in most cases when an action falling under a certain description (such as wounding a police officer) is made a crime, the law is satisfied, insofar as any matter of intention is concerned, if the accused knew or believed that his action would cause injury to his victim and that his victim was in fact a police officer. This almost exclusively cognitive approach is one distinctive way in which the law diverges from the ordinary idea of intentionally doing something, for in ordinary thought not all the foreseen consequences of conduct are regarded as intended.

A rationale of this divergence can be provided, however. Although apart from the law a man will be held to have done something intentionally only if the outcome is something aimed at or for the sake of which he acted, this element which the law generally disregards is not relevant to the main question with which the law is concerned in determining a man's legal responsibility for bringing about a certain state of affairs. The crucial question at this stage in a criminal proceeding is whether a man whose outward conduct and its consequences fall within the definition of a crime had at the time he acted a choice whether these consequences were or were not to occur. If he did, and if he chose that insofar as he had influence over events they would occur, then for the law it is irrelevant that he merely foresaw that they would occur and that it was not his purpose to bring them about. The law at the stage of assessing a man's responsibility is interested only in his conscious control over the outcome, and discards those elements in the ordinary concept of intention which are irrelevant to the conception of control. But when the stage of conviction in a criminal proceeding is past, and the question becomes how severely the criminal is to be punished, the matter previously neglected often becomes relevant. Distinctions may be drawn at this stage between the individual who acted for a certain purpose and one who acted merely foreseeing that certain consequences would come about.

The second and third applications of the notion of intention (doing something with a further intent and the intention to do a future action) are closer to nonlegal usage, and in the law, as elsewhere, certain problems of distinguishing motive and intention arise in such cases.

PROBLEMS OF LEGAL REASONING

Since the early twentieth century, the critical study of the forms of reasoning by which courts decide cases has been a principal concern of writers on jurisprudence, espe-

cially in America. From this study there has emerged a great variety of theories regarding the actual or proper place in the process of adjudication of what has been termed, often ambiguously, "logic." Most of these theories are skeptical and are designed to show that despite appearances, deductive and inductive reasoning play only a subordinate role. Contrasts are drawn between "logic" and "experience" (as in O. W. Holmes Jr.'s famous dictum that "the life of the law has not been logic; it has been experience") or between "deductivism" or "formalism" on the one hand and "creative choice" or "intuitions of fitness" on the other. In general, such theories tend to insist that the latter members of these contrasted sets of expressions more adequately characterize the process of legal adjudication, despite its appearance of logical method and form. According to some variants of these theories, although logic in the sense of deductive and inductive reasoning plays little part, there are other processes of legal reasoning or rational criteria which courts do and should follow in deciding cases. According to more extreme variants, the decisions of courts are essentially arbitrary.

LEGISLATION AND PRECEDENT. In Anglo American jurisprudence the character of legal reasoning has been discussed chiefly with reference to the use by the courts of two "sources" of law: (1) the general rules made by legislative bodies (or by other rule-making agencies to which legislative powers have been delegated) and (2) particular precedents or past decisions of courts which are treated as material from which legal rules may be extracted although, unlike legislative rules, there is no authoritative or uniquely correct formulation of the rules so extracted. Conventional accounts of the reasoning involved in the application of legislative rules to particular cases have often pictured it as exclusively a matter of deductive inference. The court's decision is represented as the conclusion of a syllogism in which the major premise consists of the rule and the minor premise consists of the statement of the facts which are agreed or established in the case. Similarly, conventional accounts of the use of precedents by courts speak of the courts' extraction of a rule from past cases as inductive reasoning and the application of that rule to the case in hand as deductive reasoning.

In their attack on these conventional accounts of judicial reasoning, skeptical writers have revealed much that is of great importance both to the understanding and to the criticism of methods of legal adjudication. There are undoubtedly crucially important phases in the use of legal rules and precedents to decide cases which do not consist merely of logical operations and which have long been obscured by the traditional terminology adopted both by the courts themselves in deciding cases and by jurists in describing the activities of courts. Unfortunately, the general claim that logic has little or no part to play in the judicial process is, in spite of its simple and monolithic appearance, both obscure and ambiguous; it embraces a number of different and sometimes conflicting contentions that must be separately investigated. The most important of these issues are identified and discussed below. There are, however, two preliminary issues of peculiar concern to philosophers and logicians which demand attention in any serious attempt to characterize the forms of legal reasonings.

Deductive reasoning. It has been contended that the application of legal rules to particular cases cannot be regarded as a syllogism or any other kind of deductive inference, on the grounds that neither general legal rules nor particular statements of law (such as those ascribing rights or duties to individuals) can be characterized as either true or false and thus cannot be logically related either among themselves or to statements of fact; hence, they cannot figure as premises or conclusions of a deductive argument. This view depends on a restrictive definition, in terms of truth and falsehood, of the notion of a valid deductive inference and of logical relations such as consistency and contradiction. This would exclude from the scope of deductive inference not only legal rules or statements of law but also commands and many other sentential forms which are commonly regarded as susceptible of logical relations and as constituents of valid deductive arguments. Although considerable technical complexities are involved, several more general definitions of the idea of valid deductive inference that render the notion applicable to inferences the constituents of which are not characterized as either true or false have now been worked out by logicians. In what follows, as in most of contemporary jurisprudential literature, the general acceptability of this more generalized definition of valid inference is assumed.

Inductive reasoning. Considerable obscurity surrounds the claim made by more conventional jurisprudential writers that inductive reasoning is involved in the judicial use of precedents. Reference to induction is usually made in this connection to point a contrast with the allegedly deductive reasoning involved in the application of legislative rules to particular cases. "Instead of starting with a general rule the judge must turn to the relevant cases, discover the general rule implicit in them …. The outstanding difference between the two methods is the

source of the major premise—the deductive method assumes it whereas the inductive sets out to discover it from particular instances" (G. W. Paton, *A Textbook of Jurisprudence,* 2nd ed., Oxford, 1951, pp. 171–172).

It is of course true that courts constantly refer to past cases both to discover rules and to justify their acceptance of them as valid. The past cases are said to be "authority" for the rules "extracted" from them. Plainly, one necessary condition must be satisfied if past cases are in this way to justify logically the acceptance of a rule: The past case must be an instance of the rule in the sense that the decision in the case could be deduced from a statement of the rule together with a statement of the facts of the case. The reasoning insofar as the satisfaction of this necessary condition is concerned is in fact an inverse application of deductive reasoning. But this condition is, of course, only one necessary condition and not a sufficient condition of the court's acceptance of a rule on the basis of past cases, since for any given precedent there are logically an indefinite number of alternative general rules which can satisfy the condition. The selection, therefore, of one rule from among these alternatives as the rule for which the precedent is taken to be authority must depend on the use of other criteria limiting the choice, and these other criteria are not matters of logic but substantive matters which may vary from system to system or from time to time in the same system. Thus, some theories of the judicial use of precedent insist that the rule for which a precedent is authority must be indicated either explicitly or implicitly by the court through its choice of facts to be treated as "material" to a case. Other theories insist that the rule for which a precedent is authority is the rule which a later court considering the precedent would select from the logically possible alternatives after weighing the usual moral and social factors.

Although many legal writers still speak of the extraction of general rules from precedents, some would claim that the reasoning involved in their use of precedents is essentially reasoning from case to case "by example": A court decides the present case in the same way as a past case if the latter "sufficiently" resembles the former in "relevant" respects, and thus makes use of the past case as a precedent without first extracting from it and formulating any general rule. Nevertheless, the more conventional accounts, according to which courts use past cases to discover and justify their acceptance of general rules, are sufficiently widespread and plausible to make the use of the term *induction* in this connection worth discussing.

The use of *induction* to refer to the inverse application of deduction involved in finding that a past case is the instance of a general rule may be misleading: It suggests stronger analogies than exist with the modes of probabilistic inference used in the sciences when general propositions of fact or statements about unobserved particulars are inferred from or regarded as confirmed by observed particulars. *Induction* may also invite confusion with the form of deductive inference known as perfect induction, or with real or alleged methods of discovering generalizations sometimes referred to as intuitive induction.

It is, however, true that the inverse application of deduction involved in the use of precedents is also an important part of scientific procedure, where it is known as hypothetic inference or hypotheticodeductive reasoning. Hence, there are certain interesting analogies between the interplay of observation and theory involved in the progressive refining of a scientific hypothesis to avoid its falsification by contrary instances and the way in which a court may refine a general rule both to make it consistent with a wide range of different cases and to avoid a formulation which would have unjust or undesirable consequences.

Notwithstanding these analogies, the crucial difference remains between the search for general propositions of fact rendered probable by confirming instances but still falsifiable by future experience, and rules to be used in the decision of cases. An empirical science of the judicial process is of course possible: It would consist of factual generalization about the decisions of courts and might be an important predictive tool. However, it is important to distinguish the general propositions of such an empirical science from the rules formulated and used by courts.

DESCRIPTIVE AND PRESCRIPTIVE THEORIES. The claim that logic plays only a subordinate part in the decision of cases is sometimes intended as a corrective to misleading descriptions of the judicial process, but sometimes it is intended as a criticism of the methods used by courts, which are stigmatized as "excessively logical," "formal," "mechanical," or "automatic." Descriptions of the methods actually used by courts must be distinguished from prescriptions of alternative methods and must be separately assessed. It is, however, notable that in many discussions of legal reasoning these two are often confused, perhaps because the effort to correct conventional misdescriptions of the judicial process and the effort to correct the process itself have been inspired by the realization of the same important but often neglected

fact: the relative indeterminacy of legal rules and precedents.

This indeterminacy springs from the fact that it is impossible in framing general rules to anticipate and provide for every possible combination of circumstances which the future may bring. For any rule, however precisely formulated, there will always be some factual situations in which the question whether the situations fall within the scope of the general classificatory terms of the rule cannot be settled by appeal to linguistic rules or conventions or to canons of statutory interpretation, or even by reference to the manifest or assumed purposes of the legislature. In such cases the rules may be found either vague or ambiguous. A similar indeterminacy may arise when two rules apply to a given factual situation and also where rules are expressly framed in such unspecific terms as "reasonable" or "material." Such cases can be resolved only by methods whose rationality cannot lie in the logical relations of conclusions to premises. Similarly, because precedents can logically be subsumed under an indefinite number of general rules, the identification of *the* rule for which a precedent is an authority cannot be settled by an appeal to logic.

These criticisms of traditional descriptions of the judicial process are in general well taken. It is true that both jurists and judges, particularly in jurisdictions in which the separation of powers is respected, have frequently suppressed or minimized the indeterminacy of legal rules or precedents when giving an account of the use of them in the process of decision. On the other hand, another complaint often made by the same writers, that there is an excess of logic or formalism in the judicial process, is less easy to understand and to substantiate. What the critics intend to stigmatize by these terms is the failure of courts, when applying legal rules or precedents, to take advantage of the relative indeterminacy of the rules or precedents to give effect to social aims, policies, and values. Courts, according to these critics, instead of exploiting the fact that the meaning of a statutory rule is indeterminate at certain points, have taken the meaning to be determinate simply because in some different legal context similar wording has been interpreted in a certain way or because a given interpretation is the "ordinary" meaning of the words used.

This failure to recognize the indeterminacy of legal rule (often wrongly ascribed to analytical jurisprudence and stigmatized as conceptualism) has sometimes been defended on the ground that it maximizes certainty and the predictability of decisions. It has also sometimes been welcomed as furthering an ideal of a legal system in which there are a minimum number of independent rules and categories of classification.

The vice of such methods of applying rules is that their adoption prejudges what is to be done in ranges of different cases whose composition cannot be exhaustively known beforehand: Rigid classification and divisions are set up which ignore differences and similarities of social and moral importance. This is the burden of the complaint that there is an excessive use of logic in the judicial process. But the expression "an excessive use of logic" is unhappy, for when social values and distinctions of importance are ignored in the interpretation of legal rules and the classification of particulars, the decision reached is not more logical than decisions which give due recognition to these factors: Logic does not determine the interpretation of words or the scope of classifications. What is true is that in a system in which such rigid modes of interpretation are common, there will be more occasions when a judge can treat himself as confronted with a rule whose meaning has been predetermined.

METHODS OF DISCOVERY AND STANDARDS OF APPRAISAL. In considering both descriptive and prescriptive theories of judicial reasoning, it is important to distinguish (1) assertions made concerning the usual processes or habits of thought by which judges actually reach their decisions, (2) recommendations concerning the processes to be followed, and (3) the standards by which judicial decisions are to be appraised. The first of these concerns matters of descriptive psychology, and to the extent that assertions in this field go beyond the descriptions of examined instances, they are empirical generalizations or laws of psychology; the second concerns the art or craft of legal judgment, and generalizations in this field are principles of judicial technology; the third relates to the assessment or justification of decisions.

These distinctions are important because it has sometimes been argued that since judges frequently arrive at decisions without going through any process of calculation or inference in which legal rules or precedents figure, the claim that deduction from legal rules plays any part in decision is mistaken. This argument is confused, for in general the issue is not one regarding the manner in which judges do, or should, come to their decisions; rather, it concerns the standards they respect in justifying decisions, however reached. The presence or absence of logic in the appraisal of decisions may be a reality whether the decisions are reached by calculation or by an intuitive leap.

CLEAR CASES AND INDETERMINATE RULES. When the various issues identified above are distinguished, two sets of questions emerge. The first of these concerns the decisions of courts in "clear" cases where no doubts are felt about the meaning and applicability of a single legal rule, and the second concerns decisions where the indeterminacy of the relevant legal rules and precedents is acknowledged.

Clear cases. Even where courts acknowledge that an antecedent legal rule uniquely determines a particular result, some theorists have claimed that this cannot be the case, that courts always "have a choice," and that assertions to the contrary can only be ex post facto rationalizations. Often this skepticism springs from the confusion of the questions of methods of discovery with standards of appraisal noted above. Sometimes, however, it is supported by references to the facts that even if courts fail to apply a clearly applicable rule using a determinate result, this is not a punishable offense, and that the decision given is still authoritative and, if made by a supreme tribunal, final. Hence, it is argued that although courts may show a certain degree of regularity in decision, they are never bound to do so: They always are free to decide otherwise than they do. These last arguments rest on a confusion of finality with infallibility in decisions and on a disputable interpretation of the notion of "being bound" to respect legal rules.

Yet skepticism of this character, however unacceptable, does serve to emphasize that it is a matter of some difficulty to give any exhaustive account of what makes a "clear case" clear or makes a general rule obviously and uniquely applicable to a particular case. Rules cannot claim their own instances, and fact situations do not await the judge neatly labeled with the rule applicable to them. Rules cannot provide for their own application, and even in the clearest case a human being must apply them. The clear cases are those in which there is general agreement that they fall within the scope of a rule, and it is tempting to ascribe such agreements simply to the fact that there are necessarily such agreements in the use of the shared conventions of language. But this would be an oversimplification because it does not allow for the special conventions of the legal use of words, which may diverge from their common use, or for the way in which the meanings of words may be clearly controlled by reference to the purpose of a statutory enactment which itself may be either explicitly stated or generally agreed. A full exploration of these questions is the subject matter of the study of the interpretation of statute.

Indeterminate rules. The decisions of cases that cannot be exhibited as deductions from determinate legal rules have often been described as arbitrary. Although much empirical study of the judicial process remains to be done, it is obvious that this description and the dichotomy of logical deduction and arbitrary decision, if taken as exhaustive, is misleading. Judges do not generally, when legal rules fail to determine a unique result, intrude their personal preferences or blindly choose among alternatives; and when words such as *choice* and *discretion,* or phrases such as "creative activity" and "interstitial legislation" are used to describe decisions, these do not mean that courts do decide arbitrarily without elaborating reasons for their decisions—and still less that any legal system authorizes decisions of this kind.

It is of crucial importance that cases for decision do not arise in a vacuum but in the course of the operation of a working body of rules, an operation in which a multiplicity of diverse considerations are continuously recognized as good reasons for a decision. These include a wide variety of individual and social interests, social and political aims, and standards of morality and justice; and they may be formulated in general terms as principles, policies, and standards. In some cases only one such consideration may be relevant, and it may determine decision as unambiguously as a determinate legal rule. But in many cases this is not so, and judges marshal in support of their decisions a plurality of such considerations which they regard as jointly sufficient to support their decision, although each separately would not be. Frequently these considerations conflict, and courts are forced to balance or weigh them and to determine priorities among them. The same considerations (and the same need for weighing them when they conflict) enter into the use of precedents when courts must choose between alternative rules which can be extracted from them, or when courts consider whether a present case sufficiently resembles a past case in relevant respects.

Perhaps most modern writers would agree up to this point with this account of judicial decision where legal rules are indeterminate, but beyond this point there is a divergence. Some theorists claim that notwithstanding the heterogeneous and often conflicting character of the factors which are relevant to decision, it is still meaningful to speak of a decision as *the* uniquely correct decision in any case and of the duty of the judge to discover it. They would claim that a judicial choice or preference does not become rational because it is deferred until after the judge has considered the factors that weigh for and against it.

Other theorists would repudiate the idea that in such cases there is always a decision that is uniquely correct, although they of course agree that many decisions can be clearly ruled out as incorrect. They would claim that all that courts do and can do at the end of the process of coolly and impartially considering the relevant considerations is to choose one alternative that they find the most strongly supported, and that it is perfectly proper for them to concede that another equally skilled and impartial judge might choose the other alternative. The theoretical issues are not different from those that arise at many points in the philosophical discussions of moral argument. It may well be that such terms as *choice, discretion,* and *judicial legislation* fail to do justice to the phenomenology of considered decision: It is the law felt involuntary or even inevitable character that often marks the termination of deliberation on conflicting considerations. Very often the decision to include a new case in the scope of a rule or to exclude it is guided by the sense that this is the "natural" continuation of a line of decisions or carries out the "spirit" of a rule. It is also true that if there were not also considerable agreement in judgment among lawyers who approach decisions in these ways, we should not attach significance and value to them or think of such decisions as reached through a rational process. Yet however it may be in moral argument, in the law it seems difficult to substantiate the claim that a judge confronted with a set of conflicting considerations must always assume that there is a single uniquely correct resolution of the conflict and attempt to demonstrate that he has discovered it.

RULES OF EVIDENCE. Courts receive and evaluate testimony of witnesses, infer statements of fact from other statements, and accept some statements as probable or more probable than others or as "beyond reasonable doubt." When it is said that in these activities special modes of legal reasoning are exhibited and that legal proof is different from ordinary proof, reference is usually intended to the exclusionary rules of the law of evidence (which frequently require courts, in determining questions of fact, to disregard matters which are logically relevant), or to various presumptions that assign greater or lesser weight to logically relevant considerations than ordinary standards of reasoning do.

The most famous examples of exclusionary rules are those against "hearsay," which (subject to certain exceptions) make inadmissible, as evidence of the facts stated, reports tendered by a witness, however credible, of statements made by another person. Another example is the rule that when a person is charged with a crime, evidence of his past convictions and disposition to commit similar crimes is not admissible as evidence to show that he committed the crime charged. An example of a rule that may give certain facts greater or less probative weight than ordinary standards do is the presumption that unless the contrary is proved beyond reasonable doubt, a child born to a woman during wedlock is the child of both parties to the marriage.

The application of such rules and their exceptions gives rise to results which may seem paradoxical, even though they are justifiable in terms of the many different social needs which the courts must satisfy in adjudicating cases. Thus, one consequence of the well-known exception to the hearsay rule that a report of a statement is admissible as evidence of a fact stated if it is made against the interest of the person who stated it, is that a court may find that a man committed adultery with a particular woman but be unable to draw the conclusion that she committed adultery with him. A logician might express the resolution of the paradox by saying that from the fact that p entails q it does not follow that "it is legally proved that p" entails "it is legally proved that q."

Apart from such paradoxes, the application of the rules of evidence involves the drawing of distinctions of considerable philosophical importance. Thus, although in general the law excludes reports of statements as evidence of the facts stated, it may admit such reports for other purposes, and in fact draws a distinction between statements of fact and what J. L. Austin called performatory utterances. Hence, if the issue is whether a given person made a promise or placed a bet, reports that he uttered words which in the context amounted to a promise or a bet are admissible. So, too, reports of a person's statement of his contemporary mental states or sensations are admissible, and some theorists justify this on the ground that such first-person statements are to be assimilated to behavior manifesting the mental state or sensation in question.

PROBLEMS OF THE CRITICISM OF LAW

ANALYSIS AND EVALUATION. A division between inquiries concerned with the analysis of law and legal concepts and those concerned with the criticism or evaluation of law prima facie seems not only possible but necessary, yet the conception of an evaluatively neutral or autonomous analytical study of the law has not only been contested but also has been taken by some modern critics to be the hallmark of a shallow and useless legal posi-

tivism allegedly unconcerned with the values or ends which men pursue through law.

Objections to pure analysis. Many different objections to a purely analytical jurisprudence have been made. By some it has been identified with, or thought to entail commitment to, the view that a legal system is a closed logical structure in which decisions in particular cases are "mechanically" deduced from clear antecedent rules whose identification or interpretation presents no problem of choice and involves no judgment of value. Other critics have contended that any serious demand for the definition of a legal concept must at least include a request for guidance as to the manner in which, when the relevant legal rules are unclear or indeterminate, particular cases involving the concept in question should best be determined. It is assumed by these critics that any question concerning the meaning of expressions such as "a right" or "a duty," as distinct from the question of what rights or duties should be legally recognized, are trivial questions to be settled by reference to a dictionary. Still others have urged that since the maintenance of a legal system and the typical operations of the law (legislation, adjudication, and the making of legal transactions) are purposive activities, any study that isolates law or legal phenomena for study without considering their adequacy or inadequacy for human purposes makes a vicious abstraction that is bound to lead to misunderstanding.

Replies to objections. None of the above seem to constitute serious objections. The difficulties of decision in particular cases arising from the relative indeterminacy of legal rules are of great importance, but they are distinct from analytical questions such as those illustrated earlier, which remain to be answered even when legal rules are clear. Thus the isolation and characterization of the normative and predictive standpoints from which law may be viewed and the precise manner of interplay between subjective and objective factors in legal transactions are not things that can be discovered from dictionaries. But attention to them is indispensable in the analysis of the notion of a legal obligation, a legal right, or a contract. There is of course much justice in the claim that in order to understand certain features of legal institutions or legal rules, the aims and purposes they are designed to fulfill must be understood. Thus, a tax cannot be distinguished from a fine except by reference to the purpose for which it is imposed; but to recognize this is not to abandon an analytical study of the law for an evaluative one. The identification of something as an instrument for certain purposes leaves open the question whether it is good or bad, although such identification may indicate the standards by reference to which this question is to be answered. In any case, there are many features of legal rules that may profitably be studied in abstraction from the purposes which such rules may be designed to achieve.

CRITERIA OF EVALUATION. Nonetheless, protests against the severance of analytical from critical or evaluative inquiries, even if misdirected in their ostensible aim, often serve to emphasize something important. These protests are usually accompanied by and sometimes confused with a general thesis concerning the standards and principles of criticism specifically appropriate to law. This is the thesis (which has appeared in many different forms in the history of the philosophy of law) that, whatever may be the case with value judgments in other fields or with moral judgments concerning the activities of individuals, the criteria which distinguish good law from bad do not merely reflect human preferences, tastes, or conventions, which may vary from society to society or from time to time; rather, they are determined by certain constant features of human nature and the natural environment with which men must contend.

The doctrine of natural law in its various traditional forms embodies this thesis. There are, however, obscurities and metaphysical assumptions involved in the use by natural-law theorists of the notions of nature and reason that make their formulations unacceptable to most modern secular thought; and they often confuse their important arguments concerning the principles by which law and social institutions should be judged with arguments designed to show that a reference to morality or justice must be introduced into the definition of law or legal validity. Nonetheless, it is possible to segregate these tangled issues, and some important modern philosophical arguments concern the possibility of restating in an acceptable form the claim that there are certain objective and rationally determined criteria for the evaluation and criticism of law. These arguments will be sketched here in relation to substantive law, procedural law, and the ideas of justice and utility.

SUBSTANTIVE LAW. The purposes that human beings pursue in society and for the realization of which they employ law as an instrument are infinitely various, and individuals may differ in the importance they attach to them and in their moral judgments about them. But the simplest form of the argument that there are certain constant criteria for the evaluation of a legal system consists in the elaboration of the truth that if law is to be of any value as an instrument for the realization of human pur-

poses, it must contain certain rules concerning the basic conditions of social life. Thus it is not only true that the legal system of any modern state and any legal system which has succeeded in enduring have contained rules restricting the use of violence, protecting certain forms of property, and enforcing certain forms of contract; it is also clear that without the protections and advantages that such rules supply, people would be grossly hampered in the pursuit of any aims. Legal rules providing for these things are therefore basic in the sense that without them other legal rules would be pointless or at least would operate only fitfully or inefficiently. Criticism of a legal system on the grounds that it omitted such rules could be rebutted only by the demonstration that in the particular case they were unnecessary because the human beings to which the system applied or their natural surroundings were in some way quite extraordinary, that is, that they lacked certain of the salient characteristics that persons and things normally have. This is so because the need for such rules derives from such familiar natural facts as that people are both vulnerable to violence and tempted to use it against each other; that the food, clothes, and shelter necessary to existence do not exist naturally in limitless abundance but must be grown or manufactured by human effort and need legal protection from interference during growth and manufacture and safe custody pending consumption; and that to secure the mutual cooperation required for the profitable development of natural resources, people need legal rules enabling them to bind themselves to future courses of conduct.

Argument along these lines may be viewed as a modest empirical counterpart to the more ambitious teleological doctrine of natural law, according to which there are certain rules for the government of human conduct that can be seen by men endowed with reason as necessary to enable people to attain the specifically human optimum state or end (*finis, telos*) appointed for human beings by Nature or (in Christian doctrine) by God. The empirical version of this theory assumes only that, whatever other purposes laws may serve, they must, to be acceptable to any rational person, enable men to live and organize their lives for the more efficient pursuit of their aims. It is, of course, possible to challenge this assumption and to deny that the fact that there are certain rules necessary if fundamental human needs are to be satisfied has any relevance to the criticism of law. But this denial seems intelligible only as a specifically religious doctrine that regards law as the expression of a divine will. It may then be argued that people's lives should be regulated by the law not in order to further any secular human purposes but because conformity to God's will is in itself meritorious or obligatory.

A more serious objection to the empirical argument conducted in terms of human needs for protection from violence to the person and property and for cooperation is the contention that although these are fundamental human needs, the coercive rules of a legal system need not provide for them. It may be said that the accepted morality of all societies provides a system of restraint which provides adequately for these needs, and that the vast majority of people abstain from murder, theft, and dishonesty not from fear of legal sanctions but for other, usually moral, reasons. In these circumstances it may be no defect in a legal system that it confines itself to other matters in relation to which the accepted morality is silent.

It seems clear, however, that social morality left to itself could not provide adequately for the fundamental needs of social life, save in the simplest forms of society. It may well be that most individuals, when they believe themselves to be protected from malefactors by the punishments, threats of punishment, and physical restraints of the law, will themselves voluntarily submit to the restraints necessary for peaceful and profitable coexistence. But it does not follow that without the law's protections, voluntary submission to these restraints would be either reasonable or likely. In any case, the rules and principles of social morality leave open to dispute too many questions concerning the precise scope and form of its restraints. Legal rules are needed to supply the detail required to distinguish murder and assault from excusable homicide and injury, to define the forms of property to be protected, and to specify the forms of contract to be enforced. Hence, the omission of such things from the legal system could not be excused on the ground that the existence of a social morality made them unnecessary.

PROCEDURAL LAW. Laws, however impeccable their content, may be of little service to human beings and may cause both injustice and misery unless they generally conform to certain requirements which may be broadly termed procedural (in contrast with the substantive requirements discussed above). These procedural requirements relate to such matters as the generality of rules of law, the clarity with which they are phrased, the publicity given to them, the time of their enactment, and the manner in which they are judicially applied to particular cases. The requirements that the law, except in special circumstances, should be general (should refer to classes of persons, things, and circumstances, not to indi-

viduals or to particular actions); should be free from contradictions, ambiguities, and obscurities; should be publicly promulgated and easily accessible; and should not be retrospective in operation are usually referred to as the principles of legality. The principles that require courts, in applying general rules to particular cases, to be without personal interest in the outcome or other bias and to hear arguments on matters of law and proofs of matters of fact from both sides of a dispute are often referred to as rules of natural justice. These two sets of principles together define the concept of the rule of law to which most modern states pay at least lip service.

These requirements and the specific value that conformity with them imparts to laws may be regarded from two different points of view. On the one hand, they maximize the probability that the conduct required by the law will be forthcoming, and on the other hand, they provide individuals whose freedom is limited by the law with certain information and assurances that assist them in planning their lives within the coercive framework of the law. This combination of values may be easily seen in the case of the requirements of generality, clarity, publicity, and prospective operation. For the alternative to control by general rules of law is orders addressed by officials to particular individuals to do or to abstain from particular actions; and although in all legal systems there are occasions for such particular official orders, no society could efficiently provide the number of officials required to make them a main form of social control.

Thus, general rules clearly framed and publicly promulgated are the most efficient form of social control. But from the point of view of the individual citizen, they are more than that: They are required if he is to have the advantage of knowing in advance the ways in which his liberty will be restricted in the various situations in which he may find himself, and he needs this knowledge if he is to plan his life. This is an argument for laws that are general in the sense of requiring courses of action and not particular actions. The argument for generality in the sense of applicability to classes of persons is different: It is that such rule confer upon the individual the advantage of knowing the restrictions to which the conduct of others besides himself will be subject. Such knowledge in the case of legal restrictions that protect or benefit the individual increases the confidence with which he can predict and plan his future.

The value of the principles of natural justice which concern the process of adjudication are closely linked to the principles of legality. The requirement that a court should be impartial and hear arguments and proofs from both sides of a dispute are guarantees of objectivity which increase the probability that the enacted law will be applied according to its tenor. It is necessary to ensure by such means that there will be this congruence between judicial decisions and the enacted law if the commitment to general rules as a method of government is taken seriously.

Care must be taken not to ascribe to these arguments more than they actually prove. Together they amount to the demonstration that all who have aims to pursue need the various protections and benefits which only laws conforming to the above requirements of substance and procedure can effectively confer. For any rational person, laws conferring these protections and benefits must be valuable, and the price to be paid for them in the form of limitations imposed by the law on one's own freedom will usually be worth paying. But these arguments do not show, and are not intended to show, that it will always be reasonable or morally obligatory for people to obey the law when the legal system provides them with these benefits, for in other ways the system may be iniquitous: It may deny even the essential protections of the law to a minority or slave class or in other ways cause misery or injustice.

JUSTICE AND UTILITY. The equal extension to all of the fundamental legal protections of person and property is now generally regarded as an elementary requirement of the morality of political institutions, and the denial of these protections to innocent persons, as a flagrant injustice. Even when these protections are denied, lip service is often paid to the principle of equal distribution by the pretense that the persons discriminated against are either criminal in intention, if not in deed, or are like children who are incapable of benefiting from the freedom which laws confer and are in need of some more paternalistic regime.

Inadequacy of utilitarianism. Different moral philosophies offer different vindications of the principle of equality. The matter is considered here in order to illustrate the philosophical problems that arose in the criticism of law concerning the relative place of the notions of utility and justice. The central principle of utilitarianism, insofar as it supplies a moral critique of law, may be stated as the doctrine that there is only one vice in legal arrangements, namely, that they fail to produce the greatest possible total of happiness in the population within their scope. The concept of a total of happiness or pleasure or satisfaction is of course open to well-known objections. But on any interpretation, utilitarian princi-

ples, if unrestricted, must endorse legal or social arrangements if the advantages they give to some persons outweigh the disadvantages imposed on others. For a consistent utilitarian there can be no necessary commitment to any principles requiring an equal distribution.

However, in some cases, if allowance is made for principles of diminishing marginal utility, it may be shown that an equal distribution is the most efficient, in the sense of producing the greatest total of happiness. But for the utilitarian this is a contingent matter to be established in each case, not a matter of moral principle or justice; and where the question concerns the distribution of the fundamental legal protections of person and property, there seems no compelling utilitarian argument in favor of an equal distribution. Thus, a slave-owning class might derive from the system of slavery benefits outweighing the misery of the slaves. Bentham urged that this was not the case, owing to the inefficiency of slave labor, and therefore he rejected slavery; but he rejected it as inefficient rather than as unjust. Plainly, this form of argument is a very insecure foundation for the principle that all people are morally entitled to the equal protection of the laws, and it seems clear that utilitarian principles alone cannot give any account of the moral importance attached to equality and in general to the notion of the just, as distinguished from an efficient, distribution as a means of happiness.

Moral argument for equality. The simplest moral argument in support of the equal distribution of the law's fundamental protections is one that combines the idea that no rational person could wish himself to be denied these fundamental legal protections with the principle of the universalizability of moral judgment: Moral judgments concerning social and legal arrangements must conform to the requirement that no man could regard as morally acceptable the withholding from others with needs and in circumstances similar to his own of those benefits which he would not wish to be withheld from himself. If this principle is admitted, it follows that it cannot be a sufficient moral ground for accepting legal arrangements that the advantages they give to some outweigh the disadvantages for others. The equal extension to all of the law's protections satisfies both the principle of utility, which requires that the law should advance human happiness, and the independent principle of justice, that the gain in happiness should be distributed fairly. According to this qualified form of utilitarianism, the best legal and social arrangements realize the most efficient of just distributions.

More ambitious arguments have been advanced to show that in spheres other than the distribution of the fundamental protections of the law, utilitarianism is acceptable only if qualified by independent principles of just distribution, and also to demonstrate that the distribution required by justice is in all spheres prima facie that of equality, unless inequalities can be shown to work ultimately for the equal benefit of all. Whatever the strength of these more general arguments may be, it is true that in relation to many legal institutions, utilitarianism unrestricted by other principles of justice yields results which would not be regarded as morally tolerable. This is particularly true of punishment. In all civilized legal systems it is recognized that no man should be punished except for his own conduct, and (with certain exceptions in the case of minor offenses) only then for such of his actions as were voluntary or within his power to control. Such limitations on the scope of punishment seem obvious requirements of justice to the individuals punished, but it is at least doubtful whether they can be adequately supported on purely utilitarian grounds.

THE OBLIGATION TO OBEY THE LAW. The philosophical investigation of the obligation to obey the law requires a distinction between the utilitarian and other moral aspects of this subject similar to that outlined in the case of justice. It seems clear that the mere existence of a legal system, irrespective of the character of its laws, is not sufficient in any intelligible theory of morality to establish that a person ought morally to do what its laws require him to do. Yet there are also powerful arguments against a purely utilitarian theory of the obligation to obey law which would regard this obligation as simply a special case of the obligation to promote happiness, with the corollary that disobedience to bad laws is justified if the consequences of disobedience (including any harm done to others through the weakening of the authority of the legal system) are better in utilitarian terms than the consequences of obedience. Among features of the moral situation for which this utilitarian theory fails to account there are two of peculiar importance. The first is that the obligation to obey law is one which is considered as owed by the citizen specifically to the members of his own society in virtue of their relationship as fellow members, and is not conceived merely as an instance of an obligation to men in general not to cause harm, injury, or suffering. Second, men are often held to be subject to an obligation to obey the law even though it is clear that little or no harm will be done to the authority of the legal system by their disobedience, as in cases (like that of the conscien-

tious objector) where those who disobey the law willingly submit to punishment.

The theory of a social contract focused on these two aspects of the obligation of obedience to law, and it is possible to detach from what is mythical or otherwise objectionable in contract theory certain considerations which show that the obligation to obey the law may be regarded as the obligation of fairness to others, which is independent of and may conflict with utility. The principle involved, stated in its simplest form, is that when a number of persons restrict their liberty by certain rules in order to obtain benefits that could not otherwise be obtained, those who have gained by the submission of others to the rules are under an obligation to submit in their turn. Conflicts between this principle and the principle of utility are possible because often the benefits secured by such restrictions would arise even if considerable numbers failed to cooperate and submit to the rules in their turn. For the utilitarian, there could be no reason for anyone to submit to rules if his cooperation was not necessary to secure the benefits of the system. Indeed, if a person did cooperate, he would be guilty of failing to maximize the total happiness, for this would be greatest if he took the benefits of the system without submitting to its restraints. The consideration that the system would fail to produce the desired benefits or would collapse if all were to refuse their cooperation is irrelevant in a utilitarian calculation if, as is often the case, it is known that there will be no such general refusal.

See also Analytic Jurisprudence; Bentham, Jeremy; Equality, Moral and Social; Ethics and Morality; Guilt; Historical School of Jurisprudence; Justice; Legal Positivism; Natural Law; Persons; Philosophy of Law, History of; Property; Punishment; Religion and Morality; Responsibility, Moral and Legal; Rights; Utilitarianism.

Bibliography

DEFINITION OF LAW

The major classics of jurisprudence in which definitions of law are elaborated include Thomas Aquinas, *Summa Theologiae* I, 2, 90–97; Jeremy Bentham, *The Limits of Jurisprudence Defined* (New York, 1945) and *The Comment on the Commentaries* (Oxford: Clarendon Press, 1928); John Austin, *The Province of Jurisprudence Determined* (London: J. Murray, 1832); J. C. Gray, *The Nature and Sources of the Law* (New York, 1909); and Hans Kelsen, *General Theory of Law and State,* translated by Anders Wedberg (Cambridge, MA: Harvard University Press, 1945).

More recent contributions include Hermann Kantorowicz, *The Definition of Law* (Cambridge, U.K.: Cambridge University Press, 1958); H. L. A. Hart, *The Concept of Law* (Oxford: Clarendon Press, 1961); and Richard Wollheim, "The Nature of Law," in *Political Studies* 2 (1954): 128–142.

ANALYSIS OF LEGAL CONCEPTS

On legal obligations and duties, see Jeremy Bentham, "Essay on Logic," in *The Works of Jeremy Bentham,* edited by J. Bowring, 11 vols. (Edinburgh, 1838–1843), Vol. VIII, pp. 213–293 and *Fragment on Government* (London, 1776), Ch. 6; and O. W. Holmes Jr., "The Path of the Law," in *Collected Legal Papers* (New York: Harcourt Brace, 1920), pp. 167–202.

On legal transactions, see Axel Hägerström, *Inquiries into Law and Morals,* translated by C. D. Broad (Uppsala, 1953), Ch. 5; Morris R. Cohen, *Law and the Social Order* (New York, 1933), Ch. 2; John L. Austin, *How to Do Things with Words* (Oxford, 1962); and Wesley N. Hohfeld, *Fundamental Legal Concepts* (New Haven, CT: Yale University Press, 1964).

On intention, see Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (London, 1789), Ch. 8; John Austin, *Lectures on Jurisprudence* (London, 1869), Chs. 18–19; G. L. Williams, *The Criminal Law: The General Part,* 2nd ed. (London: Stevens, 1961), Ch. 2; and G. E. M. Anscombe, *Intention* (Oxford, 1957).

PROBLEMS OF LEGAL REASONING

On logic and law, see Julius Stone, *The Province and Function of Law* (London: Stevens, 1947), Chs. 6–7; J. C. Jensen, *The Nature of Legal Argument* (Oxford: Blackwell, 1957); Norberto Bobbio, "Diritto e logica," in *Il problema della giustizia* (Milan, 1962), pp. 11–44; and Lars Bergstrom, "Imperatives and Ethics," in *Filosofiska Studier* 7 (1962): 94ff.

On rule and discretion in judicial decision, see John Dickinson, "Legal Rules: Their Function in the Process of Decision," in *University of Pennsylvania Law Review* 79 (1931): 833ff., and "The Problem of the Unprovided Case," in *Recueil d'études sur les sources de droit en honneur de F. Gény* (Paris, 1934), Vol. 11; Jerome Frank, *Law and the Modern Mind,* 6th printing (New York, 1949); Edward Levi, *Introduction to Legal Reasoning* (Chicago, 1949); John Wisdom, "Philosophy, Metaphysics, and Psycho-analysis," in his *Philosophy and Psycho-analysis* (Oxford, 1952); and Richard Wasserstrom, *The Judicial Decision* (Stanford, CA: Stanford University Press, 1961) and "The Obligation to Obey the Law," in *University of California Law Review* 10 (1963): 780–807. Also see *Jurimetrics: Law and Contemporary Problems* 28 (1963).

PROBLEMS OF THE CRITICISM OF LAW

On problems of the criticism of law, see Alf Ross, *On Law and Justice,* translated by Margaret Dutton (London: Stevens, 1958); John Rawls, "Justice as Fairness," in *Philosophical Review* 67 (1958): 164–194; H. L. A. Hart, *The Concept of Law* (Oxford: Clarendon Press, 1961), Chs. 8–9; Torstein Eckhoff, "Justice and Social Utility," in *Legal Essays: A Tribute to Frede Castberg* (Oslo, 1963), pp. 74–93; and Lon Fuller, *The Morality of Law* (New Haven, CT: Yale University Press, 1964).

H. L. A. Hart (1967)

PHILOSOPHY OF LAW, PROBLEMS OF [ADDENDUM]

One of the dominant issues in philosophy of law since Hart's main entry was published has been the dispute between Hart and Ronald Dworkin about the best way to characterize a legal system and the modes of legal reasoning (especially by judges) most appropriate to it.

RULES AND SOCIAL PRACTICES

THE RULE OF RECOGNITION. Hart identified two main kinds of rules in a complex and mature legal system. There are rules that tell people what to do or not do (tax laws, criminal laws, traffic laws), and there are rules that tell people how to do certain kinds of things (in order to accomplish such legal transactions as making valid wills or binding contracts and conveying property). Among the latter kind of rules he identified a small set that he regarded as fundamental to all but the most primitive legal systems: These rules tell how to identify a particular legal system and, within it, how to make laws and adjudicate claims arising under law. Hart's main entry does not address these fundamental rules.

The first kind of fundamental rule Hart famously styled the "rule of recognition." It identifies the primary sources of law (e.g., the Queen-in-Parliament) and it prioritizes these sources (e.g., statute law > common law > "customary law"). Because this and the other fundamental rules determine what is to count as valid law within that system, they have normative legal force there but are not themselves properly called *valid* laws.

SOCIAL PRACTICES AND LEGAL SYSTEMS. Hart's rule of recognition is more like a social practice (or, better, the presuppositions of such a practice) than it is like a black-letter rule of any sort. To follow or engage in a social practice is to conform reflectively to an existing, ongoing pattern or template as a matter of appropriate conduct. The practice functions as a standard and serves as a basis for criticizing deviations. Officials (almost all of them most of the time, in the standard case) simply follow the social practice: They presuppose it internally in what they actually do when they make and enforce given laws. They do so not out of fear of sanctions, but rather because so acting is the regular and expected thing to do. Ordinary citizens need not be aware of the authoritative sources of law (or the other fundamental rules) in their country; but they do need to know what the laws are, for it is these they follow or conform to. In the standard case, a substantial number of them do so in the same way as the officials do—by taking an internal point of view. This concordance between officials and ordinary citizens constitutes law as a social practice. One of Hart's main objects in invoking the idea of a social practice (or rule) is to say that a system of laws, as an exemplar of such a practice, is distinguishable from a large-scale scheme of coercion.

VALID LAWS AND JUDICIAL DISCRETION. Any legal system, insofar as it is a social practice, is an effective legal system, one where laws are conformed to most of the time by most of the people. When laws and court decisions in an effective system are made (or almost always made) in strict conformity to the fundamental rules, such a system would be a full and proper legal system. Here all the laws and court decisions that are made in accordance with the fundamental rules would be valid ones.

Hart does not think that a given law or decision (simply as valid) can cover and determine the correct outcome for all the instances that come within its proper range. For reasons that he spells out in the main entry, there will always be some such cases where the "law runs out." In those cases judges and executive officials will have to use "discretion"; they will have to supplement the law with what he calls (in the main entry) "interstitial legislation."

PRINCIPLES AND INTEGRITY

ONE RIGHT ANSWER: HERCULEAN JURISPRUDENCE. Ronald Dworkin was Hart's main critic in the last three decades of the twentieth century. One of his main criticisms is that legal systems have inbuilt features such that judges, taking the law as it is, can be said to have a duty to make the best decision. In simplest terms, then, Dworkin closes Hart's alleged gaps in law (which allow for judicial discretion) by turning to the character of legal reasoning itself, within a determinate legal system.

Dworkin's theory, using a model judge (named Hercules) for purposes of illustration, is called "law as integrity." Dworkin's main argument may be put this way: If two different judges, both committed to law as integrity agreed literally on everything—agreed "preinterpretively" on what counts as law in a given system (an agreement one can expect from all lawyers, judges, and jurisprudents in the determinate legal system within which they work, say, the United States or the United Kingdom); agreed on the relevant facts of the case; agreed on the law (the relevant propositions of law) and on the history of politics/law and on an interpretation of the political insti-

tutions in their country; agreed about the relevance of the same governing principles in the case at hand; and agreed about the main substantive principles embedded in their own legal system (especially justice [e.g., rights] and fairness [e.g., democratic decision making]) and about the interpretation and the preferred ordering of each of these—they'd reach the same decision. Thus, there is one (and only one) determinate decision a given Herculean judge would reach in a given case within the existing resources of the law, in a determinate legal system.

The theory that there is always one and only one right answer (though it may not always be reached) runs into real problems, however, when one considers a panel of such judges who must reach a single decision through discussion and voting. Here different judges, deploying somewhat different interpretive choices than Judge Hercules, may come up with answers that are significantly different from Hercules's own answers. Such judges would reach their decisions in the right way, in accordance with the ideals and procedures of law as integrity; and each judge's decision, based on convictions grounded in the law's resources, would be a wholly sound one. There could in principle, then, be more than one right answer. Given the way the world is and given Dworkin's own statement in the matter at the end of *Law's Empire* (1986, pp. 412–413), there probably would in fact be more than one right answer.

This reading does not supplant the orthodox reading for a single judge; it continues to be the case here that there is, for that judge, one and only one determinate best answer in a given case. But it does force an amendment on the "one and only one right answer" thesis for a panel of judges, or for a whole judicial system. Here, though there continues to be no need for Herculean judges ever to go outside the law's resources to reach a judicial decision, and no need for them to use discretion (or "interstitial legislation") to fill in gaps in the law, more than one right answer is possible—indeed, is to be expected.

CONVERGENCES. Hart conceded, in the "Postscript" to his *Concept of Law* (1994), that he had not given sufficient attention to principles in the law or found an appropriate role for them in his theory. He also allowed that given legal systems could have a set of embedded substantive principles (a public morality, as Dworkin called it); such principles are, for Hart, typically enshrined in a written constitution and in judicial reasoning about that constitution.

On the other hand, Dworkin's acknowledgment of the important place of near unanimous "preinterpretive" agreement on what counts as law in a given system marks an almost wholesale acceptance of Hart's idea of the nature and importance of a rule (or norm) of recognition in a mature and complex legal system. And there's much merit to Hart's observation that "Dworkin's later introduction of interpretive ideas into his legal theory [in *Law's Empire*] … brought the substance of [h]is position very close to my own" in recognizing that the courts have to deal interpretively with underdetermination in the written law. Hart continues, "Arguably [though] before the introduction of interpretive ideas into his theory there seemed to be a great difference between our respective accounts of adjudication.…" (Hart, "Postscript" to *Concept of Law* [1994], note to p. 272 on p. 307.)

UTILITARIANISM AND BASIC RIGHTS

Hart alluded in his main entry to difficulties utilitarianism had in accommodating within its normative frame the central issues of justice (that is, distribution of basic benefits and protections equally to all). But since the time at which Hart's main entry was written, significant attempts have been made within utilitarianism (under the name "indirect" utilitarianism) to address and perhaps resolve this problem.

Many people in the 1970s and 1980s—including John Rawls, Ronald Dworkin, and even thinkers broadly sympathetic to utilitarianism, such as David Lyons—concluded that utilitarianism was somehow incompatible in particular with basic rights (human or constitutional), or at least with the priority habitually given to such rights.

The problem they see is that no one can think that acting in accordance with any given right (especially if the social rules that formulate such things are kept fairly simple and easy to follow) will on every occasion yield up a result that is compatible with the general happiness principle. Sometimes deviating from that policy will have the greater welfare value. And, given the general happiness principle itself, the principle that the greater benefit should be preferred to the lesser and that normative requirements on action can always be set to achieve the greater benefit, that deviation should be taken. Sometimes a right ought to yield to these considerations: It should do so when so doing holds the prospect of greater well-being.

INDIRECT UTILITARIANISM. In an effort to deal with the problem the critics had identified, this new version of utilitarianism shifts the focus of attention from Jeremy Bentham, who did not countenance the idea of basic moral rights, to J. S. Mill, who did. Roughly, the theorists

of indirect utilitarianism assert that direct appeals to general welfare are self-defeating, all things considered, and that putting standing constraints on the principle—such as a system of moral rules (typically relatively simple and easily followable rules) or a coherent set of civil or constitutional rights justifiable by the standard of general happiness—in fact produces the greater well-being.

Indirect utilitarians do not, however, assert that moral rules should never be overridden nor individual rights ever broached. Rather, on their view, where rules or rights conflict (as they inevitably will, many have argued), some sort of appeal to the general happiness is in order.

Here is where the notion of an indirect utilitarianism comes crucially into play. Its advocates argue that the principle of general happiness should not directly determine what is to be done even here. Rather, the principle operates only indirectly in all such cases. It bears down, not on individual actions per se but on the rules themselves. Here the general welfare principle is used merely to help determine which rule is weightier, a determination that occurs gradually (over time and with experience) and cumulatively, or used to help determine a policy (a second-order rule of conduct), all things considered, for conduct when these particular moral rules (or these particular rights) conflict.

Thus, on their account it is possible to have policies for action (to have both moral rules and rights) that are justifiable by the standard of general happiness and at the same time to shield these policies from direct confrontation with (and possible overthrow by) the happiness principle on individual occasions. Thus, indirect utilitarianism (if all its arguments and presumptions are allowed) seemingly establishes that utilitarianism is compatible with basic constitutional rights and their priority—at least in the case of those rights that are themselves justifiable in accordance with the general happiness principle.

CRITICISM. But considerations of corporate good and of aggregate welfare (including those that amount to nothing more than the increased well-being of some individuals at the expense of others) can and do in fact override constitutional rights on given occasions. Indirect utilitarians cannot really deny this. If they do, then the jumping-off point of indirect utilitarianism would disappear along with the problem it was designed to solve. There would simply be no point to a strategy of shielding moral rules and constitutional rights from being overridden by corporate or aggregate political policies on those occasions when such policies were arguably supported as preferable by direct reference to the standard of general happiness.

Thus, indirect utilitarians are in effect forced to admit that social policies could override constitutional rights, within the utilitarian frame they have devised. After all, social policies in their view merely reflect, cumulatively, the results of applying general welfare considerations to occasions of acting in accordance with those policies. And they have admitted, necessarily, that sometimes corporate or aggregate political policies would in fact be supported as preferable over moral rules and constitutional rights by direct reference to the standard of general happiness.

If this is so, the general happiness principle could not support the assignment of constitutionally guaranteed benefits and protections to each and every individual person in advance, so to speak, and across the board. It could not do so if, in effect, such rights tied the utilitarian politician's hands against allowing corporate or aggregate interests to override or supersede constitutional rights when, cumulatively and all things considered, those aggregate interests could be seen to conduce to greater benefit. Indirect utilitarians cannot allow for politically fundamental constitutional rights that have a built-in, standing, and overriding priority over corporate or aggregate considerations. To this degree, then, philosophical utilitarianism is incompatible with the notion of basic rights (human or constitutional rights) as that idea is commonly understood.

RECENT CRITICAL PHILOSOPHY OF LAW: MODERN AND POSTMODERN

Recent decades have witnessed the birth of several noteworthy developments or movements within the philosophy of law. Broadly these divide into two camps. Those belonging to the first remain more or less faithful to a generally modernist and liberal orientation to legal philosophy. They include law and economics and the liberal humanist strand of feminist jurisprudence. Those belonging to the second camp take up a generally postmodernist and postliberal orientation to legal philosophy. They include critical legal studies in its various manifestations along with the more radical strands of feminist jurisprudence and critical race theory.

Characteristic of modernist liberal legal philosophy are the following assumptions:

(i) human reason is univocal and universal;

(ii) language represents reality and truth is correspondence to reality;

(iii) knowledge requires justification from foundations;

(iv) the methodological path to foundations is analysis (often drawing on methodological individualism or social atomism in the social sciences);

(v) all persons share some morally significant basic freedom and equality;

(vi) to be legitimate government must be constitutional and limited;

(vii) law serves legitimate government through its institutional subordination of power to reason; and

(viii) the true path to historical and moral progress is that marked by the rule of law.

Characteristic of postmodern and postliberal legal philosophy is the rejection of several if not all of these assumptions. Thus: human reason is multivocal and relativistic; language shapes or determines reality; truth is largely coherence; knowledge does not require justification from foundations; and so on. The most significant and general feature of postmodernist and postliberal legal philosophy, however, is its unwillingness to affirm the rule of law as either an empirical possibility or normative goal. On the postmodernist and postliberal view, it is not reason, but power, will, desire, the subconscious, the chance of history, or the forces of nature to which law is always in the end subordinate and through which any historical or moral progress must ultimately be won.

LAW AND ECONOMICS

As a development or movement within legal philosophy, law and economics took flight in the 1970s with Richard Posner's *The Economic Analysis of Law* (published originally in 1973). But its roots reach back to work in the early 1960s by Guido Calabresi, Ronald Coase, and others, as well as to legal realism's instrumentalist stance toward law and associated efforts to bring economic analysis to bear on legal issues in the early twentieth century. What unifies the law and economics movement is a commitment to putting the concepts, methods, and principles of microeconomics to work center stage in the study of law. Several law and economics theses have been advanced.

One thesis was straightforwardly descriptive. Some or all of the law was said to be best described exclusively or primarily in terms of economic efficiency. The law of tort, for example, was best understood as an institutional attempt to minimize the costs of accidents overall for society, including the cost of preventing accidents. A second thesis was straightforwardly normative. Some or all of the law was said to be properly criticized or evaluated exclusively or primarily in terms of economic efficiency. Wherever the law failed to promote or realize economic efficiency, it was to be criticized and reformed. Subsequent theses claimed that considerations of economic efficiency were the key to making accurate predictions of future legal developments, or to explaining legal history, or to giving the best interpretation of various legal systems (e.g., the United States or the United Kingdom). The normative thesis remains today the most widely affirmed and discussed. But taken as a thesis about the primary or overriding aim of law it is not compelling.

ECONOMIC EFFICIENCY. Economic efficiency is a property of transactions or relations between persons and was developed as a proxy for aggregate utility, which was thought unmeasurable given the impossibility of interpersonal utility comparisons. If a transaction or relation makes all those it affects better off or at least no worse off by their own lights, then there is good reason to believe that it increases aggregate utility (though it is not possible to know by how much). Such a transaction or relation is Pareto superior to its status quo ante. Any state of affairs from which no Pareto superior transactions or relations is possible is Pareto optimal. The set of Pareto optimal states of affairs marks the limit of our ability rationally to act so as to improve aggregate utility.

Of course, some non-Pareto optimal states of affairs may actually represent gains in aggregate utility over any or all Pareto optimal states of affairs. But without being able to do interpersonal utility comparisons, there is no way of reliably picking them out. From a utilitarian perspective, then, using the law to facilitate or produce Pareto superior transactions and relations up to but not beyond a point of Pareto optimality is a normatively sound ambition. The law may do this in at least three ways: (i) distributing legal rights and entitlements to those who value them most; (ii) redistributing the costs and benefits of some transaction or relation so as to render it efficient on the Pareto criteria; or (iii) sustaining an open and transparent market with low transaction costs and few incentives for strategic holdout behavior so that persons can voluntarily exchange until they arrive at a Pareto optimal state of affairs (which, according to the Coase theorem, they will do).

The Pareto criteria of economic efficiency have limited application because most transactions or relations between persons generate transaction costs or adverse third-party effects. The Kaldor-Hicks criterion of effi-

ciency accounts for this by picking out as efficient any transaction or relation in which those who gain enough that they could in principle (but need not actually) compensate from their gain those who lose, such that no person impacted by the transaction or relation would be made worse off by it relative to its status quo ante. The Kaldor-Hicks criterion, however, is problematic at the level of application because two different states of affairs may be reciprocally Kaldor-Hicks efficient (the Scitovsky paradox). As more refined criteria of efficiency continue to be introduced, the underlying idea remains the same: Economic efficiency is a proxy for aggregate utility.

DESCRIPTIVE THESIS. Whereas it is possible superficially to describe many areas of the law in terms of economic efficiency, the extent to which the law is well described in such terms is difficult to determine. It may be more efficient (reduce costs overall) for the law to deal with accidental harms through liability rather than property rules because the latter would require those who cause accidents to undertake the costly project of reaching agreements with their victims ex ante. But, because of the complexity and general unavailability of the information required, it is nearly impossible to defend any particular liability rule as privileged from the point of view of economic efficiency. The expected costs associated with any particular rule will be a function not just of the degrees and probabilities of harm from accidents covered by the rule, but also of such things as the costs of the care required to avoid liability and of administering and enforcing the rule. The descriptive thesis advanced by law and economics becomes less compelling as the picture of law to which it is applied is made more realistic and fine-grained.

NORMATIVE THESIS. Attention has shifted over recent years to the normative claim that regardless of how the law as it stands is best described, surely it ought primarily to aim at economic efficiency. This claim is problematic. First, prescriptions that make use of highly simplified economic models inattentive to the kinds of information alluded to above are of marginal use. But the costs (e.g., of information gathering) of building more useful models are likely prohibitive. Second, it is not clear why efficiency should be taken as normatively primary for the law. Whereas there may be good utilitarian reasons to insist that legal reforms always be efficient relative to their status quo ante, there are no good utilitarian reasons to insist that legal reforms either be Pareto optimal regardless of the path to them or be Kaldor-Hicks efficient, because neither guarantees a gain in aggregate utility over the status quo ante. It is unlikely that there are any other good moral reasons (of fairness, or consent, or respect for autonomy) to privilege Pareto optimality or Kaldor-Hicks efficiency as the overriding aim of the law. Thus, the case for grounding legal criticism and reform exclusively in considerations of economic efficiency is weak. Still, economic efficiency may (and probably should) play a subordinate role in legal criticism and reform.

FEMINIST JURISPRUDENCE

Characteristic of feminist jurisprudence are two claims, one descriptive and explanatory, the other normative. The former is that the patriarchical oppression of women is fundamentally realized through law. The latter is that the ending of patriarchical oppression must rank at or near the top of the list of aims in terms of which the law is properly criticized and reformed. Apart from these claims, however, there is little general consensus within feminist jurisprudence. Positions vary with respect to whether women and men share the same fundamental interests, whether those interests are rooted in a biologically or psychologically given human nature, the extent to which those interests are malleable regardless of their genesis, and the proper relationship of the law to those interests.

Liberal humanist feminists generally regard the abolition of patriarchy as a substantially completed task, the completion of which is possible without radical change to the basic structure of modern liberal legal institutions and theory. They aspire to an egalitarian humanism realized under the rule of law. They endeavor to reveal and reform those remaining areas of the law—for example, rape law, employment law, and marriage law—through which patriarchical oppression continues to operate. Progressive feminists also generally regard the abolition of patriarchy as a substantially completed project, the completion of which is possible under the rule of law. But they argue for more radical substantive changes to modern liberal legal theory—for example, the recognition of special rights for women as distinct from men, or the redrawing of the lines marking a private domain presumptively immune to state intervention. These more radical changes to substantive law may be argued for on the grounds that women possess at least some fundamental interests distinct from men, or that under current conditions privacy merely secures a social space for the unchecked reproduction of patriarchical self-understandings.

So, for example, whereas liberal humanist feminists insist that the free speech and privacy rights common to

men and women properly protect the private consumption of pornography, progressive feminists typically endorse legal restrictions on the private consumption of pornography, or at least that pornography that depicts women as mere sexual objects or as subordinate to men. The call for a "battered woman's syndrome" defense to homicide is also a progressive feminist initiative; it is a carefully limited but substantively radical revision to a particular legal doctrine (concerning intent) necessary if the fundamental interests of women are to be secured under the rule of law.

Whereas liberal humanists and progressive feminists divide over the means necessary and appropriate to a final victory over patriarchy, they both seek that victory within and through the rule of law and thus share a modernist orientation toward the law. Radical feminists are different. They argue that patriarchy depends on and is at least partially constituted through the rule of law. They reject the modernist aspiration to historical and moral progress through law and seek a more radical revision to the legal status quo ante. Radical feminists argue that the categories most basic to modern liberal legal theory and practice—such as due process, equal rights, fairness, state neutrality, consent, individual responsibility, privacy, justice, objectivity, impartiality, and rules—underwrite and obscure patriarchical oppression. They seek both to illuminate this fact and to suggest alternative, typically nonlegal or extralegal, frameworks for thinking about and realizing social order.

Feminist jurisprudence has been and remains theoretically diverse and rich. This is in part because it remains politically and methodologically open. Feminist legal theorists have often allied themselves with and drawn on the work of those pursuing other emancipatory political agendas. In its various strands, feminist jurisprudence draws on neo-Marxist and poststructuralist critical theory, queer theory, race theory, neopragmatism, Lacanian psychoanalytic theory, and rational choice theory.

CRITICAL LEGAL STUDIES

Critical Legal Studies (CLS) grew out of a conference in the 1970s that sought to bring together the New Left politics of the 1960s, American Legal Realism's instrumentalist stance toward law, and European social theory (structuralism and poststructuralism). At its inception, then, CLS was divided between modernist and postmodernist orientations toward the law, drawing from Nietzsche, Marx, Weber, Habermas, Foucault, and Derrida. In time, this division was settled in favor of a postmodernist orientation. What began as a radical critique of law under conditions of modern capitalism became a more radical critique of the idea of law itself. At its most provocative, at least in the United States, CLS called into question the possibility of realizing justice under or through law.

Though CLS had some presence in England and Germany, it was and remains (to the extent that it remains at all) primarily an American development. Throughout its history, CLS organized itself generally around two theses. The first was that legal systems, both in their content and operations, were best understood as ideological systems of legitimation. The second was that legal systems were indeterminate and thus incapable of subordinating the exercise of coercive political power to reason. Together these theses underwrite the proposition that law is always and everywhere only the politics of power by another name.

CLS, like American Legal Realism, understood the content and structure of the law to derive in the end from nonlegal normative commitments. And, again like Legal Realism, it sought honesty about that fact. Just as legal realists had undertaken to show that much of American law was determined by a laissez-faire political ideology rather than any science of legal reasoning, so too did CLS scholars. What was not so determined was determined, on the CLS view, by patriarchical or racist or other morally suspect political commitments. Of course, legal realists sought to expose the ideological bases of law so as to place law in the service of morally more reputable nonlegal or extralegal political commitments (generally utilitarian and progressive). CLS scholars generally rejected this instrumentalist approach to law. They tended to argue that the law was always an effect, and could never be the genuine cause, of underlying political, social, and economic change. The point of demystifying the law and exposing it as ideological in nature was not to put it in the service of a more noble cause, but rather to encourage non- or extralegal means to social reform.

That the content and methods of mature legal systems almost always underdetermine the answer to at least some legal questions is neither a radical nor particularly controversial claim. By the 1960s, few legal philosophers thought mature legal systems were or could be fully autonomous and possessed of sufficient internal resources to generate, mechanically as it were, a single determinate answer to every legal question. The existence of so-called hard cases was taken for granted. That all cases were hard cases, however, was not. It is this thesis that CLS, in its most ambitious moments, advanced: that mature legal systems (or at least particular legal systems, e.g., the United States and the United Kingdom) are rad-

ically indeterminate and, accordingly, that the rule of law is impossible.

Three lines of argument were advanced for this thesis. The first and least ambitious rooted the indeterminacy of the law in the formal structures of law. Legal rules competed not only with one another, but with more flexible standards and principles. Precedents, often diverse themselves, could always be read narrowly or broadly. Principles of statutory construction pointed in multiple directions. And so on. While sufficient to debunk any vision of legal reasoning as scientific or mechanical, this argument is not sufficient to establish the radical indeterminacy of law. The sheer number of easy cases never litigated suggests that legal reasoning is not inherently radically indeterminate.

A second and more ambitious argument rooted the indeterminacy of the law in the inconsistency or incoherence of liberal political morality (presumably foundational at least in the United States and the United Kingdom and other contemporary liberal democracies). Liberal political morality valued both individual self-interest and the collective or common good, saw the individual as ultimately free and responsible and socially constituted, and committed itself to state neutrality while privileging secular modernist humanistic conceptions of the good. It was, in short, inconsistent and incoherent. But competing principles and commitments are not necessarily inconsistent or contradictory. Liberal political morality may indeed express and undertake to mediate rationally and reasonably the tension between several competing principles and commitments. It need not, for all that, be reducible to an irrational self-contradiction or to incoherence.

The third and most ambitious argument for the indeterminacy of the law appealed to the structure of language and thought itself. The argument here, drawn from poststructuralist linguistic and social theory, was that the possibility of language and thought, the possibility of meaning itself, presupposed for any particular utterance or expression the existence of a multiplicity of meanings. If legal language could mean even one thing, then, it must necessarily mean or potentially mean many things. For several years many CLS scholars made the case for this proposition by using "deconstructive" strategies of critical reading to "trash" legal propositions privileged within the conventional order of legal reasoning. But this argument ultimately proved to be its own undoing. It dissolved the purposeful human subject in an endless proliferation of meanings and reduced progressive politics to the obscure mysticism of such slogans as "deconstruction is justice."

The future of CLS as a movement in legal philosophy remains unclear. Its most provocative claims have been largely abandoned, whereas its more modest but also more plausible claims (about the relationship of law to politics and underdetermination within the law) have been largely assimilated into more mainstream jurisprudential thinking.

See also Bentham, Jeremy; Derrida, Jacques; Dworkin, Ronald; Feminist Legal Theory; Foucault, Michel; Habermas, Jürgen; Hart, Herbert Lionel Adolphus; Humanism; Justice; Legal Positivism; Legal Realism; Marx, Karl; Mill, John Stuart; Nietzsche, Friedrich; Philosophy of Law, History of; Rawls, John; Responsibility, Moral and Legal; Rights; Utilitarianism; Weber, Max.

Bibliography

HART AND DWORKIN ON SOCIAL PRACTICES AND PRINCIPLES

Coleman, Jules, ed. *Hart's Postscript: Essays on the "Postscript" of "The Concept of Law."* New York: Oxford University Press, 2001.

Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

Dworkin, Ronald. *A Matter of Principle*. Cambridge, MA: Harvard University Press, 1985.

Dworkin, Ronald. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press, 1978. Originally published 1977; important appendix added in 1978.

Hart, H. L. A. *The Concept of Law*. 2nd ed., with a postscript, edited by Penelope Bulloch and Joseph Raz. Oxford: Oxford University Press, 1994.

Rawls, John. *Political Liberalism*. 2nd ed. Lectures 6 and 8. New York: Columbia University Press, 1996.

UTILITARIANISM AND BASIC RIGHTS

Gray, John. *Mill on Liberty: A Defence*. London: Routledge, 1983.

Hart, H. L. A. *Essays in Jurisprudence and Philosophy*. Chs. 8 and 9. Oxford: Clarendon Press, 1983. Both of these essays were originally published in 1979.

Lyons, David. *Rights, Welfare, and Mill's Moral Theory*. New York: Oxford University Press, 1994. Collects together Lyons's "Utility and Rights," referred to below, with other of his relevant essays on Mill.

Lyons, David. "Utility and Rights." In *Ethics, Economics, and the Law*, *NOMOS* 24, edited by J. R. Pennock and John W. Chapman, 107–138. New York: New York University Press, 1982.

Lyons, David. "Utility as a Possible Ground of Rights." *Nous* 14 (1980): 17–28. This essay and "Utility and Rights" criticize the thesis that indirect utilitarianism is compatible with basic rights, but many of his earlier essays support that compatibility.

Martin, Rex. *A System of Rights*. Oxford: Clarendon Press, 1993. See especially chapter 13, section 1.

Rawls, John. *A Theory of Justice*. 2nd ed. Harvard University Press, 1999.

RECENT CRITICAL PHILOSOPHY OF LAW AND SUBSEQUENT SECTIONS

Altman, Andrew. *Critical Legal Studies: A Liberal Critique*. Princeton, NJ: Princeton University Press, 1990.

Calabresi, Guido, and A. D. Melamed. "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral." *Harvard Law Review* 85 (1972): 1089–1128.

Coleman, Jules, and Jeffrey Murphy. "Law and Economics." In *Philosophy of Law*. 2nd ed. Boulder, CO: Westview Press, 1990.

Frug, Mary Jo. *Postmodern Legal Feminism*. New York: Routledge, 1992.

Gilligan, Carol. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press, 1982.

Kairys, David, ed. *The Politics of Law: A Progressive Critique*. New York: Pantheon Books, 1982.

Kelman, Mark. *A Guide to Critical Legal Studies*. Cambridge, MA: Harvard University Press, 1987.

MacKinnon, Catherine. *Feminism Unmodified: Discourses on Life and Law*. Cambridge, MA: Harvard University Press, 1987.

Minow, Martha. *Making All the Difference*. Cambridge, MA: Harvard University Press, 1990.

Patterson, Dennis. *Law and Truth*. New York: Oxford University Press, 1996.

Polinsky, A. *An Introduction to Law and Economics*. 2nd ed. Boston, MA: Little Brown, 1989.

Posner, Richard. *The Economic Analysis of Law*. 5th ed. Boston, MA: Little Brown, 1998.

Smith, Patricia. *Feminist Jurisprudence*. New York: Oxford University Press, 1993.

Symposium: "Critical Legal Studies." *Stanford Law Review* 36 (1984): 1–674.

Symposium: "Deconstruction and the Possibility of Justice." *Cardozo Law Review* 11 (1990): 919–1734.

Symposium: "Efficiency in the Law." *Hofstra Law Review* 8 (1980): 485–770.

Unger, Roberto. *The Critical Legal Studies Movement*. Cambridge, MA: Harvard University Press, 1986.

Rex Martin (2005)
David Reidy (2005)

PHILOSOPHY OF MEDICINE

The subject matter unique to philosophy of medicine—as opposed to those issues that are best seen under the heading of philosophy of biology—is clinical medicine and its underlying methodology and assumptions. Crucial to philosophy of medicine is the family of terms *disease*, *malady*, *health*, *normal*, *abnormal*, *condition*, and *syndrome*, all of which have evaluative aspects to their definitions. For all its scientific base, medicine must be a value-laden practice guided by the values of its practitioners and its public. It is in this regard—but not only in this regard—that the claim "Medicine is an art and a science" should be understood.

DISEASE, HEALTH, AND NORMALITY

A stable departure from physiological normality that causes death, disability, pain, loss of pleasure, or inability to achieve pleasure is the sort of entity that is called disease (Clouser, Culver, and Gert 1981). The departure has to be stable enough so that it causes similar problems in similar people and so that it is recognizable by different medical practitioners as the same disease entity. When the departure is less clearly individuatable than a disease, the entity is referred to as a syndrome.

Normality and health are relative terms. They are relative to species, age, gender, (perhaps) social status, race, and ultimately to one's own physiology. A healthy (normal) eighty-five-year-old is different from a healthy (normal) twenty-year-old; and a healthy (normal) professional athlete is different from a healthy (normal) philosophy professor. Normal health is also relative to one's values. Unless a person feels comfortable doing what she wants to do, she can claim to be unhealthy by saying things like: "I just don't feel up to par." In this sense health is a theoretical state of a person.

The concept of biological variability derives its useful sense from the relativity of *normal*. Biological variability makes generalization problematic in a way that generalizing from one billiard ball to any such object is not. Biological variability—meaning that no two organisms are exactly alike—is trivially true. It is unhelpful, except as a reminder that generalization is problematic.

Diseases are real to the extent that they are stable departures from normality (sometimes called "baseline") as defined above. Obviously, diseases are not like traditional physical objects. They can overlap and be in two places at the same time. (Mental diseases present their sorts of problems, which parallel issues in philosophy of mind and philosophy of psychology.) Diseases are real in that they cause real pain, disability, or both; they are real in the sense that they can be reduced to physiological occurrences. Diseases are theoretical in the sense that they are not traditional physical objects, and they are identified only relative to a value structure that then becomes part of the medical theory. For example, given the current medical theory of European and North American scien-

tific medicine, chronic fatigue syndrome is a disease. But against the backdrop of eighteenth-century medicine, it would have been seen primarily as a characteristic of some women and lazy men. Chiropractic medicine sees disease only in terms of misalignment of vertebrae. The reality of disease, a sense for reduction, and the theory-ladenness of disease exemplify traditional questions in philosophy of science.

What is classifiable as a disease is also a function of what physicians are willing to do, what they are interested in, and what will be reimbursed. Thus, infertility is treated as a disease in large part because it is a terrible burden to some, it is interesting to deal with medically, and people are willing to pay for treatment. Being short is also treated as a condition worth reversing (in children) for the same sorts of reasons. This makes disease relative to culture and economic conditions.

Treating a condition as if it were a disease makes it a disease in a stipulative sense but not in the physiological sense. Baldness and bad breath would be conditions that might be troublesome, most effectively treated medically, and yet still not classified as diseases. However, if they are caused by a disease, they may be considered signs of an underlying medical condition. Psychiatry periodically re-decides whether certain psychological conditions should be considered diseases.

Genetics adds an interesting twist to defining disease and thinking about health. Consider a disease such as sickle-cell anemia, where homozygous recessive is a serious disease but the heterozygous condition can be beneficial in malarial areas (heterozygotes have a better survival rate from malaria than do either homozygotic forms) but still can, in rare instances, cause serious medical problems. Thus, in a nonmalaria infested area, the heterozygous condition might be called a disease. Huntington's disease is caused by a dominant gene whose effects do not manifest themselves until (usually) middle age. Should one consider a teenager with the dominant gene diseased? One could say that the person is healthy now even though the gene is one that will cause a disease later in life. But this is not really correct, because there are subtle changes in body chemistry caused by the dominant gene even when there are no Huntington's symptoms. One normally would say something such as, "a mutation in the normal gene is what causes Huntington's disease." This locution is odd for two reasons: (1) the person (almost assuredly) never had a normal gene to mutate; and (2) using "normal" when speaking of the gene would seem to imply that "abnormal" and "diseased" might be usable for genes as well. But, of course, the gene for Huntington's disease is not a normal gene with a disease.

THE LOGIC OF DIAGNOSIS

Diagnosis and scientific explanation present similar philosophical problems, especially with respect to explanation, causality, and laws. Diagnosis begins with history taking and moves on to the physical examination. The standard history questions assume that disease entities have a typical natural history to them.

Signs are objective characteristics, such as blood pressure and broken bones. Symptoms are the subjective characteristics reported by the patient—for example, pain and lightheadedness. The signs and symptoms of disease vary with the stage of the disease. Thus, an early stage of any disease may be confused for the later stage of another. Physicians look for the best overall explanation for the condition, given the patient's individuating factors such as age, gender, occupation, stress factors, and so forth. The best explanation is assumed to be the most probable explanation, where the disease is considered to be the cause of the condition being investigated.

A standard procedure in diagnosis is the rule-out test. A physician limits the diagnosis to a few conditions and then does a test, which, if negative, will rule out one of the possible causes. This procedure is repeated until only one likely answer is left. This is in keeping with a simplistic version of falsification.

Doctors also use a simple confirmation strategy in diagnosis. Usually, more than one confirmatory test result is required before the diagnosis is accepted. Other predictions will have to be borne out by test results as well as physical findings and consistent history. Laboratory tests are crucial to modern-day diagnoses, although they present problems. Results are subject to false positives (disease reported when absent) and false negatives (disease not reported when present). The best test has a high true positive ratio and a low false positive ratio. Bayes's theorem can be used to calculate the probability that a person with a positive test actually does have the disease in question.

Because test results are continuous, cutoff points must be chosen. The cutoff points are chosen based on how serious an error would be. If a disease is fatal and can be treated safely, then a high false positive rate would be acceptable. For less worrisome conditions, compromise between the two figures is possible. Again, values are part of what looks like a objective aspect of medicine. In this sense, medical diagnosis may be different from the usual

picture of the scientific method. There are other differences as well.

Some of the crucial aspects of physical diagnosis—for example, interpreting heart sounds and kinds of rashes—are subjective and cannot be taught so much as they must be learned by practice. The apprenticeship of medical students and physicians (residents) is, in this sense, different from the time graduate students in science spend learning bench laboratory skills. Also, anecdotes play a role in diagnosis in a way that they would not in physics or most other sciences. Related to the reliance on anecdotes is that the best physicians just seem to sense that, no matter where the facts are pointing, something else is going on. Subjectivity, anecdotes, and intuition seem not, in general, to be good scientific methodology, and yet it seems to be precisely what separates the great clinicians from the ordinary ones. The key to understanding these great diagnosticians is probably pattern recognition.

Physicians often wait in order to let a disease show itself more clearly, sometimes confirming their diagnoses by follow-up: Did the condition follow its predicted course? Did the treatment have the expected effect and in the expected manner? If not, the diagnosis may well have been incorrect. Even if the follow-up is consistent with the diagnosis, the actual condition may have been different and may have remitted on its own or have been similar enough to the disease suspected so that it responded to the treatment. In these sorts of cases, physicians do not know that they were wrong; they will count these cases as successes and so use them to support a similar diagnosis the next time. There is no practical defense against this failing.

HOLISM AND REDUCTIONISM

Holistic medicine assumes that diseases are primarily a function of lifestyle and life events of the patient. A holistic approach to diagnosis will focus as much on psychosocial history as it will on traditional signs and symptoms. Stress as a factor in disease is important in holistic accounts. Reductionistic medicine focuses more on physiology as the key to diagnosis, treatment, and taxonomy of disease. The reductionistic approach is the legacy of scientific medicine begun in the mid-nineteenth century.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Causation: Philosophy of Science; Explanation; Laws, Scientific; Laws of Nature; Philosophy of Biology; Philosophy of Mind; Reduction; Reductionism in the Philosophy of Mind.

Bibliography

Bursztajn, H., R. Feinbloom, R. Hamm, and A. Brodsky. *Medical Choices, Medical Chances*. pts. 1 and 2. New York: Routledge, Chapman and Hall, 1990.

Clouser, K. D., C. M. Culver, and B. Gert. "Malady: A New Treatment of Disease." *Hastings Center Report* 11 (June 1981): 29–37.

Collins, D. "Genetics of Huntington's Disease." University of Kansas Medical Center. Available from http://www.kumc.edu/hospital/huntingtons/genetics.html.

Kark, J. "Sickle Cell Trait." Information Center for Sickle Cell and Thalassemic Disorders. Available from http://sickle.bwh.harvard.edu/sickle_trait.html.

Kelley, W., ed. *Textbook of Internal Medicine*. Chps. 6–8. Philadelphia: Lippincott, 1992.

Margolis, J. "Thoughts on Definitions of Disease." *Journal of Medicine and Philosophy* 11 (3) (1986): 233–236.

Maull, N. "The Practical Science of Medicine." *Journal of Medicine and Philosophy* 6 (2) (1981): 165–182.

McNeil, B., et al. "Primer on Certain Elements of Medical Decision Making." *New England Journal of Medicine* 293 (5) (1975): 211–215.

Merskey, H. "Variable Meanings for the Definition of Disease." *Journal of Medicine and Philosophy* 11 (3) (1986): 215–232.

Munson, R. "Why Medicine Cannot Be a Science." *Journal of Medicine and Philosophy* 6 (2) (1981): 183–208.

Murphy, A. E. *The Logic of Medicine*. Baltimore, 1976.

Passmore, R., ed. *A Companion to Medical Studies*. Vol. 3. Oxford: Blackwell Scientific Publications, 1974.

Schaffner, K. "Philosophy of Medicine." In *Philosophy of Science*, edited by M. Salmon, et al. Englewood Cliffs, NJ: 1992.

Wulff, H. *Rational Diagnosis and Treatment*. Oxford: Blackwell Scientific Publications, 1976.

Arthur Zucker (2005)

PHILOSOPHY OF MIND

The mind seems to occupy a special place in the world. It is the seat of thought and feeling, of rationality and moral concern. Is it fundamentally different from the other things we find in the natural world? Is it possible for the mind to be investigated scientifically? Can one ever really know what is going on in the mind of someone else?

Such questions delineate the subject matter of the philosophy of mind. The central problem in this area is the *mind-body problem*: the project of finding an account of the mind that locates it in the broader physical world. While this problem does not exhaust the philosophy of mind, one's response to it imposes substantial constraints on what one may say about other questions in this area.

One of these questions concerns mental causation. It seems obvious that what happens in the mind can bring about physical events in one's body and vice versa. If, however, the mental is radically different from the physical, such causal commerce may seem problematic. A related question concerns the prospects for psychology. If there is difficulty in supposing that causal laws govern the mental, what sorts of results can we expect from the scientific investigation of the mind?

Another question concerns epistemology. The *problem of other minds* is the project of explaining how one can know about the minds of others. The problem arises as a result of an asymmetry between how one knows about one's own mental states and how one knows about others' mental states. I know what I am thinking or feeling by a peculiar means devoid of inference from more basic bits of knowledge. I do not have to make observations on the basis of which I find out what is going on in my mind. Indeed, I may be unable to err about my own mental states. By contrast, I cannot know what someone else is thinking or feeling without observing their behavior, and the inferences I make are plainly subject to error. One may wonder if such inferences are ever justified. Even if they are justified, one may wonder how it is possible that the very same kind of phenomena may be known in such radically different ways.

These questions presume, of course, that we know how to sort the mental from the nonmental in the first place. Two features seem especially characteristic of the mental. First, many mental states exhibit what is known as *intentionality*: They have content directed at the world; they are about things. The belief that the earth is flat, for example, has as its content the proposition that the earth is flat; the fear of flying, for another example, is about flying. Second, any mental state involving an experience displays the striking feature that it makes sense to speak of *what it is like* to be a particular kind of creature having that kind of experience. Mental states having this what-it-is-like character may be called *phenomenal* states, and if someone is in such a state, we may say that that person is *phenomenally conscious.*

Some remaining questions in the philosophy of mind aim at more specific kinds of mental phenomena. What is the difference between an emotion and a mood? How is an intention to act related to a desire to act? Such questions often branch into other areas of philosophy: the philosophy of action, of responsibility, and so on.

THE MIND-BODY PROBLEM

Leaving the notion of reduction at an intuitive level, we may distinguish the two dominant positions on the mind-body problem as follows. Materialism (or physicalism) is the thesis that that the mental reduces to the physical. More cautiously, every mental entity is ultimately nothing above and beyond the physical entities that exist. Once certain physical entities are in place, nothing extra is needed for the mental entities to exist as well. By contrast, dualism is the thesis that the mental and physical are ultimately distinct, so that neither reduces to the other. What makes the mind-body problem a problem is that we seem to have powerful evidence for both of these incompatible positions.

On the one hand, the physical workings of the brain apparently suffice to account for our behavior. If the mind is not in some way reducible to the brain, it is hard to see what room there could be for the mind to play any role in our behavior. Yet it surely does play such a role. Further, it is plain that events affecting the brain have systematic effects on the mind as well. So simplicity favors eliminating the mind as an extra entity.

On the other hand, the mind resists such a reduction. It is hard to see how the physical aspects of anything could add up to its having thoughts or feelings. Any putative creature with a mind may well be a mindless automaton. The point is made most vivid if we consider a physical organism built out of the very same physical ingredients as you, the reader. It seems possible that such an organism may yet be mindless, despite its physical and behavioral similarity to a creature that has a mind. If this is right, then clearly what makes you a creature with a mind does not automatically accompany your physical characteristics; it is something over and above the physical. Hence, materialism is false.

Materialism and dualism are not the only options. One other option is idealism (or phenomenalism), according to which it is the physical that ultimately reduces to the mental. Further, there is the view that neither the mental nor the physical reduces to the other. Rather, both reduce to some third, neutral entity. This position is sometimes known as neutral monism. Neither alternative has been widely endorsed. Idealism is apt to seem simply incredible, and neutral monism may seem frustratingly mysterious.

DUALISM

René Descartes developed the most famous form of dualism, a form known as Cartesian substance dualism. On

this view, the mind is a substance (and hence capable of existing on its own without any distinct supporting entities, such as a body). It is distinct from any physical object in that it is essentially without spatial extension. A dualist need not adopt all of these tenets, however. A dualist need hold only that mental properties—such as being in pain or intending to leave the room—are in principle independent of any combination of physical properties.

As noted above, there is considerable pressure to opt for materialism in order to make sense of the causal role that the mental has in affecting our behavior. One route that the dualist might take in the face of this pressure is to endorse epiphenomenalism—the view that, contrary to appearances, mental events never cause physical events. There would then be no need to accommodate the causal role of the mind. This advantage is offset, however, by a grave difficulty in accounting for one's knowledge of other minds. When I know what someone else is thinking, my primary evidence is that person's behavior. In treating this as evidence, I presume that part of what brings about such behavior is the person's mental states. Epiphenomenalism undercuts this presumption and throws into doubt the value of such evidence.

There is no requirement that a dualist be an epiphenomenalist. Perhaps the most attractive form of dualism is one that maintains causal interaction between mind and body while rejecting the supposition that the mind is a substance. The view known as *emergentism* does exactly this. (For a classic defense of the view, see C. D. Broad's *The Mind and Its Place in Nature.*) Emergentism may be characterized by three theses. First, it rejects the view that the mind is a substance, maintaining only that there are two types of properties, mental properties and physical properties (property dualism). Second, it claims that once an organism reaches a certain level of complexity, the laws of nature dictate that it will then have various irreducible mental properties. Third, it holds that these mental properties subsequently make a difference to the organism's behavior. More precisely, in the presence of mental properties, the physical elements of the organism behave differently from what one would expect on the basis of just the general laws governing those physical elements when not assembled in this special fashion. The emergentist thus makes the bold empirical claim that we have in effect only an incomplete view of the laws of physics, that if one were to examine the physical events occurring in creatures with minds, one would find that the usual laws do not apply.

LOGICAL BEHAVIORISM

In the first half of the twentieth century, logical behaviorism held sway as the main alternative to dualism. On this view, any statement about the mental can be translated into a statement about behavioral dispositions. A statement such as "Amy is in pain" is synonymous with some such statement as "Amy is disposed to wince, cry out, etc." Since wincing and crying out are themselves physical events, it seems that something purely physical can be disposed to undergo such events. If being in pain is merely being thus disposed, then a purely physical thing can be in pain. (A seminal statement of this view may be found in Gilbert Ryle's *The Concept of Mind*.)

Logical behaviorism has several attractive aspects. First, it makes good sense of our knowledge of other minds. If pain is simply the disposition to wince, cry out, or the like, then I can know that someone is in pain simply by observing those behaviors. Second, it fits happily with the familiar picture of how we come to learn psychological terms, specifically, that one learns what others mean by the word "pain" by observing that pain is attributed to people on the basis of their overt behavior. Third, the view explains why we have apparent a priori knowledge of the links between certain mental states and certain behaviors. We do not have to gather empirical evidence to support the claim that wincing is typically a sign of pain.

Logical behaviorism nonetheless faces a very basic problem, namely, that no proposed translation is in fact plausible unless it makes use of further mental terms. Consider again the statement "Amy is in pain." It seems possible for this to be true even when she is not disposed to wince, cry out, or the like. Suppose, for instance, that she wishes not to let anyone discover her pain and is thus determined to suppress any overt indications of it. She will then not be inclined to behave in those ways.

Of course, we could try to understand the behavioral disposition in a more complex fashion. We might unpack the disposition claim as follows: "Amy is in such a state that *if she were to feel uninhibited*, she would wince, cry out, etc." The situation in which she never displays such behaviors because she is determined to suppress such signs is no counterexample to this translation. But if this is the translation on offer, then we have not succeeded in showing how a purely physical entity could be in pain, since the complex characteristic assigned to Amy is already mental in part in that it refers to feeling uninhibited—a mental characteristic in its own right. In general, any attempt to characterize a behavioral disposition seems bound to include such a reference.

THE IDENTITY THEORY

The identity theory rose to prominence in the middle of the twentieth century, succeeding logical behaviorism as the leading materialist theory. The view is simple: Every type of mental state is identical with some type of physical state, probably a neurophysiological state. (An extensive overview of the sorts of considerations that helped lead many philosophers to the identity theory can be found in Herbert Feigl's extensive essay "The 'Mental' and the 'Physical'.")

What is novel in the identity theory is not so much its simple positive claim as its disavowals: The identity claim is not accompanied by any claim about translation. This feature of the identity theory enables it to avoid many of the usual objections to materialism. This virtue can be illustrated by working with a well-known example. Consider the claim that being in pain is identical with some type of brain process, say, having one's C-fibers firing. Identity theorists suppose that the relation between "being in pain" and "having one's C-fibers firing" is analogous to the relation between "the morning star" and "the evening star": the terms have different senses, but the same referent. Other favored examples include the identity of lightning with a kind of electrical discharge, or the identity of heat with molecular motion. In each case, the identity can only be discovered empirically; it cannot be discovered by a priori analysis of the meanings of the terms.

While the a posteriori character of these identity claims is a key appeal of the identity theory, it is also the source of important objections. One such objection was made famous as "objection 3" in J. J. C. Smart's classic paper "Sensations and Brain Processes" (1959). If an identity statement of the form "$M = P$" is a posteriori, different concepts must be associated with "M" and "P." Those different concepts involve different properties that pin down the referent of "M" and "P." For example, with the identity "the morning star = the evening star," the first name is associated with certain properties, such as being visible in the morning, that are not associated with the second name. Now turn to the alleged identity of pain with C-fiber firing. By analogy, we should conclude that despite the truth of this identity, there are nonetheless two distinct sets of properties, those associated with "pain" and those associated with "C-fiber firing." The objection, finally, is that the property associated with "pain" is a mental property that has yet to be identified with anything physical. Any a posteriori identity between the mental and physical will leave an unreduced residue of mental properties, and these mental properties undermine materialism.

Smart's response to this objection was to acknowledge that there must be different senses associated with the mental and physical terms of the identity while insisting that the sense of the mental term can be explained without appeal to any further mental properties. For instance, he claims that the sentence "I see a yellowish-orange afterimage" is equivalent in meaning to "There is something going on in me which is like what goes on in me when I see an actual orange in good light." The vocabulary in this second statement is topic-neutral in the sense that it is silent on the nature of what is going on; it may or may not be a physical process. When we identify the experience of a yellowish-orange afterimage with a type of brain process, that identity is justified by the empirical evidence that shows that the named type of brain process is in fact what is going on when one sees an actual orange in good light.

It is worth stressing here that, while the identity theorists advertised a lack of commitment to translations of psychological sentences, this sort of objection seems to force them to providing translations nonetheless. Their translations might prove to be just as dubious as the behaviorist's translations.

ANOMALOUS MONISM

One important challenge to the identity theory is posed by anomalous monism, the view championed by Donald Davidson and made famous in his essay "Mental Events." The view may be defined as a combination of one positive thesis and one negative thesis. Positively, it holds that each particular mental event is also a particular physical event, though categories of mental events cannot be equated with categories of physical events. Anomalous monism thus endorses a thesis of token identity, but not type identity. The negative thesis is that the mental is anomalous: there are no strict laws involving mental events as such. This anomalism allegedly blocks the discovery of laws relating the mental and the physical, laws apparently needed to justify a claim of identity between mental and physical properties.

The negative claim is aimed directly at the identity theory; it seeks to undercut potential sources of empirical support for that view. It is worth noting that even if anomalism is consistent with the identity theory, it is certainly significant for psychology, since it rules out the ambition of psychology to uncover strict laws governing the mental.

The positive thesis also challenges the identity theory, albeit indirectly, in that it suggests that one can be a materialist without being an identity theorist. If it suffices for materialism to say that each particular mental event is identical with some physical event, then a materialist may rest content with such instead of holding out for the more ambitious theory of type identity. Yet few philosophers are convinced that a thesis of token identity is sufficient for materialism. Intuitively, a materialist must hold, at a minimum, that how someone is mentally depends on how that person is physically. The thesis of token-event identity does not secure this result.

SUPERVENIENCE

The idea that how things are physically must determine how things are mentally may be captured by the notion of *supervenience*, also introduced into the philosophy of mind by Davidson. To say that mental properties supervene on physical properties is (roughly) to say that any two creatures that are exactly alike physically must also be exactly alike psychologically. There may be no neat match-up of mental and physical properties, but supervenience implies that how things are mentally is fixed by how things are physically.

The notion of supervenience is in this way useful for formalizing a kind of dependence of the mental on the physical, although there have been many subtly different ways of making the notion precise. There is an important limitation to any supervenience thesis, however, in that the thesis itself leaves unanswered questions as to why and how the mental is determined in this fashion. To answer these questions, it seems that a more committed theory of the nature of the mental is needed.

FUNCTIONALISM

A distinct challenge to the identity theory came in the form of functionalism. This is the view that mental properties are functional properties, that is, properties defined by the causal or functional roles they play. Consider the property of being a laundry detergent. Something is a laundry detergent if and only if it can combine with water in a washing machine to clean clothes. Various different chemicals can play this role equally well. When a particular chemical plays this role, it is said to *realize* the property of being a detergent. Since many different chemicals can play this role, the property of being a detergent is multiply realized and cannot be identified with any one of its realizers. (A seminal paper advocating functionalism is Hilary Putnam's "The Nature of Mental States.")

One motivation for functionalism is the conviction that it is in fact very unlikely that there is a single physical property to be found in all creatures sharing a given mental state. Functionalism accommodates this conviction by allowing mental properties to be multiply realized, and it does so without giving up on materialism, as an individual can have a functional property solely in virtue of his physical characteristics.

Functionalism has also been found attractive because of the apparent similarity between minds and computers. Consider what it is for a computer to run a program. The same program can be run by many different sorts of machines, so long as they have distinguishable states that play the right roles relating inputs, outputs, and each other. If the mind is akin to a computer, mental states may plausibly be classified as functional, relating sensory inputs, behavioral outputs, and different internal states. (For an important challenge to this analogy, see John Searle's "Minds, Brains, and Programs.")

A further appeal of functionalism is that it promises a degree of autonomy for psychology. If mental properties are multiply realized, then one can investigate what mental properties do without worrying about the specific physical characteristics of the underlying realizers. It is, of course, controversial how much autonomy this provides.

Even if we opt for functionalism, there remains much work to be done by way of locating the right sorts of functional properties to identify with various mental properties. The two distinctive features of the mind mentioned earlier—intentionality and phenomenal consciousness—provide targets for such work.

INTENTIONALITY

Theories of intentionality have generally taken either of two forms. They differ primarily in whether they determine the content of a mental state by appeal to the overall functioning of the mind in question or by appeal to individual mental states in isolation. On the former (interpretational) approach, a subject *S* has the belief that *P* just in case the belief that *P* appears in the overall assignment of intentional states providing the best interpretation of *S*. The details of the theory depend on what it takes to amount to a good interpretation. Typically, the idea is that the theory must predict the behavior of *S* and make *S*'s thoughts and actions by and large *rational* for someone in that environment.

The other (causal/informational) approach, which focuses on specific connections between particular brain states and states in the world, is encouraged by the idea

that we may be able to distinguish within states such as believing that *P* and hoping that *P* a common element—a *representation* that *P*—for which an independent physicalist theory can be given. The physical state *N* may represent that *P* by virtue of a causal link, in that someone in state *N* has the information that *P*. On a very simple version of this view, *N* represents that *P* if and only if the only thing that can cause someone to be in *N* is the fact that *P*. This simple version fails to make room for false representations, however; some way is needed to distinguish those causes of the representing state that fix its content and those that do not. (Seminal works in this area include Fred Dretske's *Knowledge and the Flow of Information* and Jerry Fodor's *Psychosemantics*. A useful survey may be found in the anthology *Mental Representation: A Reader*, edited by Stich and Warfield.)

CONTENT EXTERNALISM

Whatever theory of content one develops, an important constraint is imposed by content externalism. This is the view that the content of someone's mental states is determined not solely by that person's intrinsic features; the larger social and historical environment in which that person is embedded makes a difference. An easy route to seeing the point is to consider beliefs about particular individuals. Suppose that Amy and Basil are friends, that Amy believes that Basil is intelligent, and further, that Basil has a twin about whom Amy knows nothing. Amy's belief is plainly about Basil, not his twin. Yet if the situation were reversed, so that Amy was acquainted with Basil's twin instead of Basil, her belief would have had a different content, even though she would have been intrinsically the same in both cases. Hence, the contents of one's mental states may vary while one's intrinsic features remain unchanged. (Two fundamental papers about content externalism are Hilary Putnam's "The Meaning of 'Meaning' " and Tyler Burge's "Individualism and the Mental.")

This observation has raised two concerns. First, some worry that externalism is problematic for the view that intentional mental states can play a causal role in determining behavior. The worry, crudely put, is that since content is determined by wider environmental factors, content can play a causal role in behavior only if those wider environmental factors themselves play a causal role, which seems mysterious. A second concern is that externalism may be incompatible with the privileged access to our own minds that we seem to have. We need not investigate our environment to know what we think; yet if the contents of our thoughts depend on that environment, it may seem mysterious how we manage such a feat. These two problems have motivated some philosophers to introduce a notion of narrow content—mental content determined solely by the intrinsic features of the agent. If there is such a thing as narrow content, any theory of intentionality needs to accommodate it as well as content individuated in a more ordinary fashion.

PHENOMENAL CONSCIOUSNESS

The second distinctive aspect of the mind with which materialists must contend is phenomenal consciousness. What sort of physical and/or functional property can ensure that its bearer is undergoing an experience?

Many positive approaches to phenomenal consciousness take their cue from the fact that phenomenal states seem bound up with intentionality. Consider, for instance, what it is like to look at a bright red tomato. That experience plausibly represents the world as being a certain way: as containing a bright red tomato. One may even argue that all phenomenal states include such content. The state of pain, for instance, may represent one's body as being damaged.

What makes a state phenomenal, however, is not simply its having a certain content. Something else must be added to distinguish the mere belief that there is a bright red tomato in front of one from the visual experience of a bright red tomato in front of one. A variety of proposals have been offered as to what might make the difference. On one option, the content of a phenomenal state plays a rather different functional role in the overall system than the content that attaches to a mere belief. On another, a phenomenal state is a representational state that itself is represented by some other, higher-order representational state.

Whatever the merits of these theories, few would hold that they can be seen to be true simply as a matter of conceptual analysis. It is simply too easy to imagine situations in which the proposed physical and functional conditions are met even while nothing is experienced at all. Indeed, it seems quite conceivable that a being could have all the various physical and functional properties that we ourselves have and yet be devoid of phenomenal consciousness. Such creatures are known as philosophical zombies—physical duplicates of ourselves for whom all is "dark inside." (For influential discussions, see David Chalmers, *The Conscious Mind*; Peter Ludlow et al., eds., *There's Something about Mary*; and Ned Block et al., eds., *The Nature of Consciousness*.)

The fact that we can easily conceive of such zombies does not, of course, settle the issue in favor of dualism. As is familiar from the work of the identity theorists, the identity in question may be a posteriori. Consider the case of heat and molecular motion again. There is, in fact, no possible situation in which heat is present without molecular motion; nonetheless, we can apparently conceive of such a situation. We may explain away that apparent conceivability, however, by pointing out that we could then be imagining a world in which something other than heat *appears* to be heat, because, we imagine, this other thing produces heat sensations. We have misdescribed the genuine possibility we imagined.

The materialist appears to be obligated to offer a similar sort of story explaining away our apparent ability to conceive of zombies. There is, however, an important difference between the psychophysical case and the case of phenomenal states. In the heat example, we could distinguish between the appearance of heat and the heat itself, but in the case of phenomenal states, it is unclear that a comparable distinction can be drawn. (This well-known argument is found in Saul Kripke's *Naming and Necessity*.)

The difficulty here is related to one discussed earlier—namely, that made famous as "objection 3" in Smart's classic defense of the identity theory. There the worry turned on the implications of saying that mental and physical terms are associated with quite distinct concepts. The materialist needs to offer some story about those concepts that allows us to explain the a posteriori character of the identity claim, and the apparent possibility of zombies, in a way consistent with the claim that *all* properties are ultimately nothing over and above physical properties. Whether any such story is available remains an extremely controversial question.

See also Behaviorism; Dualism in the Philosophy of Mind; Functionalism; Mind-Body Problem; Physicalism; Reductionism in the Philosophy of Mind.

Bibliography

Armstrong, David. *A Materialist Theory of the Mind*. Rev. ed. New York: Routledge, 1993.

Block, Ned, ed. *Readings in Philosophy of Psychology*. Cambridge, MA: Harvard University Press, 1980.

Block, Ned, Owen Flanagan, and Güven Güzeldere, eds. *The Nature of Consciousness: Philosophical Debates*. Cambridge, MA: MIT Press, 1997.

Broad, C. D. *The Mind and Its Place in Nature*. Paterson, NJ: Littlefield Adams, 1960.

Burge, Tyler. "Individualism and the Mental." *Midwest Studies in Philosophy* 4 (1979): 73–121.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford, U.K.: Oxford University Press, 1996.

Chalmers, David J., ed. *Philosophy of Mind: Classical and Contemporary Readings*. Oxford, U.K.: Oxford University Press, 2002.

Davidson, Donald. "Mental Events." In *Essays on Actions and Events*. Oxford, U.K.: Oxford University Press, 1980.

Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

Feigl, Herbert. *The "Mental" and the "Physical": The Essay and a Postscript*. Minneapolis: University of Minnesota Press, 1967.

Fodor, Jerry. *Psychosemantics*. Cambridge, MA: MIT Press, 1987.

Fodor, Jerry. *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 1981.

Foster, John. *The Immaterial Self*. London: Routledge, 1991.

Kim, Jaegwon. *Supervenience and Mind*. Cambridge, U.K.: Cambridge University Press, 1993.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

LePore, Ernest, and Brian McLaughlin, eds. *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford, U.K.: Blackwell, 1985.

Levine, Joseph. *Purple Haze: The Puzzle of Consciousness*. Oxford, U.K.: Oxford University Press, 2001.

Lewis, David. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (1972): 249–258.

Ludlow, Peter, Yujin Nagasawa, and Daniel Stoljar, eds. *There's Something about Mary: Essays on Frank Jackson's Knowledge Argument*. Cambridge, MA: MIT Press, 2004.

Putnam, Hilary. "The Meaning of 'Meaning'." In his *Mind, Language, and Reality*. Vol. 2 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1975.

Putnam, Hilary. "The Nature of Mental States." In his *Mind, Language, and Reality*, Vol. 2 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1975.

Ryle, Gilbert. *The Concept of Mind*. Chicago: University of Chicago Press, 1984.

Searle, John. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (1980): 417–424.

Smart, J. J. C. "Sensations and Brain Processes." *Philosophical Review* 68 (1959): 141–156.

Stich, Stephen P., and Ted A. Warfield, eds. *Mental Representation: A Reader*. Oxford, U.K.: Blackwell, 1994.

D. Gene Witmer (2005)

PHILOSOPHY OF PHYSICS

The philosophy of physics investigates the logical, conceptual, metaphysical, and epistemological foundations of the physical sciences, especially fundamental physics. It is concerned with general issues such as the subject matters and aims of physics, the nature of physical laws, the

direction of time, and issues specific to particular theories such as the measurement problem in quantum mechanics and the status of the second law in thermodynamics. The philosophy of physics is enormously relevant to traditional metaphysics because it addresses the implications of physical theories for fundamental ontology, the natures of time and space, laws, causal relations, counterfactuals, and natural kinds. This encyclopedia includes entries on general, specific, and metaphysical issues in the philosophy of physics, so this entry will serve mainly as a guide to the main problems in philosophy of physics and to direct the reader to more specific articles.

The best short characterization of physics derives from Aristotle's view that physics is the science of motion and the causes of motion of material bodies; paradigmatically the motions of planets, projectiles, and pointers. The primary aim of fundamental physics has been to find a true theory (or theories) that specifies a fundamental ontology, spatiotemporal structure and laws, and that provides a complete (or as complete as possible) account of the motions of such material bodies. Many natural phenomena (e.g., the tides, the weather, rainbows, the growth of plants, the movements of animals, light, and even mental phenomena) either involve the motions of material bodies or are the causes of motions of material bodies. It follows that the scope of physics includes most everything. A true theory that accounted for the motions of all material bodies would be a theory of everything or at least of everything capable of making a difference to the positions and motions of material bodies.

The possibility of their being a complete physical theory was given a tremendous boost by the development of Newtonian or classical mechanics (see the "Classical Mechanics, Philosophy of" entry). The ontology of classical mechanics consists of dimensionless particles that possess inertial mass and certain other intrinsic properties and that move in a three-dimensional space in accordance with certain laws. Macroscopic material bodies are identified with more or less stable configurations of dimensionless particles and the motions of a particle are described in terms of the change of spatial position (or relative spatial position) over time.

The motion of a particle (and so the motions of the material bodies) is determined by the forces acting on it via the single dynamical law $F=m(p)a$ where F is the total force (the vector sum of all forces acting on particle p) on p, m(p) is the inertial mass of p and a is p's acceleration. A free particle (one on which the total force is 0) moves at a constant velocity. Newtonian forces are determined by the intrinsic natures of particles (their masses, charges, and so on) and their relative positions. For example, the attractive gravitational force particles exert on one another is given by $F = Gm1m1/r^2$; where m1, m2 are the gravitational masses of the two particles and r is the distance between the two particles. Classical mechanics was enormously successful in accounting for the motions of material bodies in circumstances where the total force on a body could be (approximately) determined as in the motions of the planets, comets, projectiles, and so on.

Classical mechanics is usually understood as supporting determinism. This means, roughly, that the state of the universe at time t together with the dynamical laws determines the state of the universe at any other time. Pierre Simon de Laplace made this vivid by imagining a supreme intelligence that ascertains the state of the universe at one time and then, knowing the laws of mechanics, is able to predict the state of the universe at any other time. There are subtle issues concerning the relations between determinism and prediction and also issues concerning whether classical mechanics is genuinely deterministic. It is only given certain qualifications concerning the nature of the force laws and the assumption that the system is isolated (see the "Determinism and Indeterminism" entry). Many philosophers think that the issue of the truth of determinism has significant implications for issues concerning free will (see the "Determinism and Freedom" entry) and more generally the place of mind in nature.

During the latter half of the nineteenth century classical mechanics was extended to include light and other electromagnetic phenomena. This involved introducing electromagnetic fields and dynamical equations (Maxwell's equations) that described the dynamics of electromagnetic fields and interactions between the motions of charged particles and fields. Light was understood as a kind of wave disturbance in the aether—a posited substance that was supposed to fill all space and provide the ground for electromagnetic fields. Also, toward the end of the nineteenth century it became increasingly plausible that matter is composed of atoms of various kinds and that these can be identified with Newtonian particles. By the last decade of the nineteenth century the package of Newtonian mechanics, Maxwellian electromagnetic theory, and the atomic theory of matter looked like good candidates for the sought after complete theory of the motions of material bodies. Of course this turned out not to be so.

One of the main philosophical discussions inspired by classical mechanics concern the natures of space and time. Isaac Newton thought of space as a kind of arena in

which particles move. Newtonian space is absolute in that its existence and nature is independent of the particles it contains. It is three-dimesional, infinite in each dimension, homogeneous, and Euclidian. The positions of and so distances between particles are defined in terms of their locations in space. It follows that for Newton there are distinct possible universes in which all the distances among particles are identical but the positions are translated. Famously, G.W. Leibniz argued that it is difficult to make sense of absolute space. He observed that God would have no reason to place the material contents of space in one region of absolute space rather than another and concluded that absolute space offends the principle of sufficient reason.

The dispute between Newton and Leibniz blossomed into a debate between those (absolutists) who think of space as an independent entity that provides spatial structure and those (relationists) who think of spatial relations between particles as primary. There are famous arguments on both sides. Relationists observe that by the lights of Newtonian physics, absolute position and motion are empirically inaccessible. Empiricist considerations suggest to them that we should not believe that absolute space exists. Absolutists respond that although we cannot determine absolute motion, absolute space is required to provide an adequate explanatory theory including the explanation of possible distance relations and of rotations. Relationists reply by arguing that the empirical content of Newtonian theory is that trajectories are physically possible only if they can be embedded into absolute space and satisfy the Newtonian laws, but that reference to absolute space is merely a convenient fiction. Only spatial relations are real. This debate has survived the demise of Newtonian mechanics and continues in discussions of the interpretations of relativity theories (see the entries "Space in Physical Theories" and "Relativity Theory").

Newton also thought of time as absolute. He suggests that time flows throughout the universe at a constant rate. It is assumed that a free particle traverses equal absolute distances in equal intervals of absolute time and so free motion measures absolute time. There is also an absolutist/relativist issue concerning the nature of time. Relativists observe that Newtonian theory provides no empirical access to absolute temporal locations. Again empiricist considerations suggest that physics can do without Newtonian time.

Some relativists claim that the empirical content of classical mechanics involve only facts about temporal sequences of interparticle distances. On a sophisticated relativist account, the laws of classical mechanics specify which sequences of interparticle distances are physically possible (see Julian Barbour's *The End of Time* and the "Time in Physics" entry for further detail). Exactly how far one can go in dispensing with apparent spatial and temporal structure in favor of spatial and temporal relations while maintaining the empirical core of classical mechanics—or relativity and quantum mechanics—remains a lively topic of discussion.

Another issue concerning time in classical mechanics involves the apparent direction or arrows of time. Many apparently lawful processes, in particular those associated with thermodynamics, are temporally directed. For examples, gasses diffuse, ice in warm water melts, and electromagnetic waves emanate from moving charged particles and the entropy of isolated systems never decreases. In addition, causation, counterfactuals, memory, decision, and so forth are temporally directed. However, the dynamical laws of classical mechanics are temporally symmetric in that for any sequence of particle positions that are in accord with those laws (i.e., is physically possible), the reversed sequence of positions is also physically possible. Where then does the arrow of time come from? Newton seems to have thought of time as possessing an intrinsic direction of flow. But it is hard to see how this flow—whatever "flow" amounts to—can account for the temporal asymmetries. It seems that the solution must lie in physical laws or conditions rather than the metaphysical nature of time.

There has been much work within physics on the problem of reconciling temporally asymmetric processes, in particular those of thermodynamics, with temporally symmetric fundamental laws. (see the "Philosophy of Statistical Mechanics" entry). Ludwig Boltzmann observed that most of the micro states compatible with, say, a block of ice floating in warm water are ones that evolve toward the future in accordance with their dynamical laws to ones in which the ice block melts. *Most* is determined relative to a natural measure on the set of micro states and Boltzmann understood this to mean that it is very likely that the ice block will melt. However, it turns out that, relative to the same probability measure, it is very likely that the ice cube evolved from one that was more melted in the past! This follows from the temporal symmetry of the laws. One response to this problem is that the explanation of temporal asymmetries lies in the macro state of the very early universe. It is posited that this state was one of enormously low entropy (and satisfies certain further conditions) and it is also posited that there is a probabil-

ity distribution over micro states that realize this state. This is called "the past hypothesis" (Albert 2000).

It has been argued that it follows from the past hypothesis and the dynamical laws that macroscopic systems that become approximately energetically isolated (e.g., an ice cube in warm water) will satisfy (the appropriate statistical versions of) the laws of thermodynamics. Some philosophers have pursued his idea further and claimed that all of the temporal arrows are ultimately derivable from the past hypothesis and the dynamics. The foundations of statistical mechanics and the relations between fundamental laws of physics and special science laws (see the "Special Sciences" entry) remain controversial philosophical issues.

The idea that the package of classical mechanics, electromagnetic theory, the atomic theory of matter, and statistical mechanics constitute the complete theory of motion was undermined during the first decades of the twentieth century as it became clear that these theories are incompatible with one another and inadequate as a theory of the very small—atomic structure—and the very big—cosmology. One big problem is that in Maxwell's equations the speed of light appears as a constant of nature. It was thought that this speed is relative to the aether. This suggests that it ought to be possible to measure the absolute velocity of the Earth relative to the aether by sending light rays in various directions. However, experiments designed to measure the velocity of the Earth relative to the aether yielded null results. It appeared that measurements of the speed of light yield the same result no matter the velocity of the source or receiver. Obviously, some modification of classical mechanics/electromagnetic theory was required.

H. A. Lorentz proposed modification of the Newtonian laws so that clocks and measuring rods, which are in motion with respect to absolute space, systematically slow down and shorten. As a consequence, although there are facts about the velocities of bodies with respect to absolute space, it also turns out that those velocities cannot be detected. Albert Einstein's special theory of relativity (STR) makes a quite different and revolutionary proposal. It rejects absolute Newtonian space-time as the framework for the motions of matter in favor of Minkowski space-time.

In Minkowski space-time the fundamental notion is that of the space-time interval between events. Einstein posited that the interval between any two events connected by a light ray is 0. This has the consequence that there are no absolute (frame independent) facts about the elapsed time or spatial distance between two events. It also follows that there are pairs of events (events that cannot be connected by a light ray) for which there are no absolute facts about their temporal order. Einstein's proposal entails the same phenomena as Lorentz's as a result of changing the underlying spatiotemporal structure.

The change from Newtonian space and time to Minkowski space-time suggested to Einstein the possibility of accounting for gravitation not as a force between bodies, but rather as a feature of space-time itself. He succeeded in doing this in the general theory of relativity (GTR). The main idea of the general theory is that the geometry of space-time itself has a geometrical structure that is not Euclidian (i.e., flat) but, rather, depends locally on the distribution of matter and energy. According to GTR, bodies freely move on geodesics (the shortest paths between points in space-time) and what counts as a geodesic is given by the geometry. Because gravitation is an effect of space-time in the GTR, not a force as in classical mechanics, it follows that it acts on all bodies in the same way. This is quite different from Newtonian mechanics in which gravitation is a force that acts the same on all bodies only because inertial mass and gravitational mass are equal. Where in Newtonian mechanics space is an inert arena, in the GTR space-time is a dynamical entity that changes over time through interactions with matter. Both the STR and the GTR are spectacularly successful in their empirical predictions (see "Relativity Theory" entry).

The STR and the GTR have been the objects of much discussion in the philosophy of physics. Among the main issue are: paradoxical scenarios; for example, the twin paradox and the possibility of closed causal loops and apparent time travel, the extent to which the metric of space-time is a real fact or is, to some extent, conventional (see the entry on "Conventionalism"), descendents of the absolutist/relationist dispute within relativistic frameworks, the formulation and viability of determinism within relativity theory (see the "Hole Argument" entry), the compatibility relativity, and quantum mechanics.

The other major failure of classical mechanics/electromagnetic theory concerned its inadequacy as accounts of atomic structure and interaction between atoms and light. According to these theories, atoms should be unstable. Further, it was found, contrary to these theories, that matter emits radiation only with certain specific frequencies, that light behaves in particle-like as well as wave-like ways, and that electrons behave in wave-like as well as particle-like ways. Over the first third of the twentieth century a novel theory—quantum mechanics—developed to account for these and many other phenomena. In

quantum mechanics the state of a system at t is characterized by a wave function Ψ(t).

Ψ(t) specifies the values of certain "observables" (position, momentum, spin, and so on) and the probabilities of obtaining various measurement results. A novel feature of quantum mechanics is that Ψ(t) specifies the values of only some observables; for example, if it specifies the value of x-spin, say spin up (in which case it is said to be an eigenstate of x-spin with value spin up), it specifies no value for other spin observables (e.g., y-spin). This is an instance of Werner Heisenberg's uncertainty principle. Ψ also specifies the probabilities of the results of measurements of other spin observables. If Ψ1 is an eigenstate of observable O with value v1 and Ψ2 is an eigenstate of O with value v2, then the superposition c1Ψ1+ c2Ψ2 is a well-defined state that specifies no value for O but says that the probability of a measurement of O yielding value v is c^2. Ψ(t) evolves deterministically by Schrödinger's law except when measured. When measured, Ψ collapses probabilistically to an eigenstate of the measured observable.

Quantum mechanics is beset with puzzles. The dominant way of thinking about quantum mechanics—the Copenhagen interpretation—holds that an observable possesses a determinate value only when the state is an eigenstate of that observable. What does it mean for an electron to possess a position but no determinate momentum (or the other way round) and yet for there to be a probability of a measurement yielding a particular value? It turns out that, in typical (nonmeasurement) interactions, the macroscopic system will evolve into a state that is not an eigenstate of ordinary properties. This is the situation of Erwin Schrödinger's cat that ends up in a state that is not an eigenstate specifying whether it is alive or dead (see "Quantum Mechanics" entry). What can that mean?

Further, that measurement appears in the fundamental laws is immensely implausible and completely unsatisfactory without a precise characterization of measurement. There is also the novel feature that typical quantum states are nonlocal. As Einstein observed and John Bell demonstrated (see entries on "Einstein, Albert" "Bell, John, and Bell's Theorem," and "Non-locality"), there are quantum states involving pairs of particles for which a measurement on one of the pair instantaneously changes the probabilities of certain measurement results for the other particle. This appears to be a kind of influence at a distance that seems incompatible with special relativities apparent prohibition on superluminal causal influences. Whether or not the conflict is genuine is a subtle issue

For most of its history and up until the present, these problems encouraged an instrumentalistic construal of quantum mechanics (see entries on "Scientific Realism" and "Copenhagen Interpretation"). Instrumentalism amounts to giving up the ambition of a complete true theory of motion. However, in the last few decades a number of realist interpretations of quantum mechanics have been proposed. These include Bohmian mechanics, many world/minds theories, and spontaneous collapse theories. Each of these interpretations specify an explicit ontology (that interprets the wave function realistically and may include other items) and laws governing that ontology that yield results matching (or approximately matching) the predictions of orthodox quantum theory. In some, such as Bohmian mechanics, the dynamical laws are completely deterministic, whereas in others, such as the GRW collapse theory, are probabilistic. Because these interpretations are empirically equivalent (or approximately empirically equivalent), they provide an interesting real example of theory underdetermination (see the entry on "Underdetermination Thesis, Duhem-Quine thesis").

Among the notable features of realist interpretations of quantum mechanics are the difficulty squaring it with relativity theories. Currently, there is no satisfactory quantum version (realist or not) of general relativity. Producing such an account is one of the urgent problems of contemporary physics. Less often appreciated is the difficulty in reconciling quantum mechanics and Einstein's Minkowski formulation of special relativity. A realist understanding of the wave function seems to require (because of nonlocal states) more space-time structure than Minkowski space-time provides. Interpretations of quantum theory and connections with relativity will be of central concern in the philosophy of physics in the twenty-first century.

See also Bell, John, and Bell's Theorem; Classical Mechanics, Philosophy of; Conventionalism; Copenhagen Interpretation; Determinism and Freedom; Determinism and Indeterminism; Einstein, Albert; Hole Argument; Non-locality; Philosophy of Statistical Mechanics; Quantum Mechanics; Relativity Theory; Scientific Realism; Space in Physical Theories; Special Sciences; Time in Physics; Underdetermination Thesis, Duhem-Quine Thesis.

Bibliography

Albert, David Z. *Quantum Mechanics and Experience*. Cambridge, MA: Harvard University Press, 1992.

Albert, David Z. *Time and Chance*. Cambridge, MA: Harvard University Press, 2000.

Barbour, Julian. *The End of Time: The Next Revolution in Physics*. Oxford; New York: Oxford University Press, 2000.

Maudlin, Tim. *Quantum Non-locality and Relativity*. Oxford, U.K.; Cambridge, MA: Blackwell, 1994.

Barry Loewer (2005)

PHILOSOPHY OF RELIGION

Analytical "philosophy of religion," still in its infancy in the 1960s, has developed markedly since then. Other approaches have certainly continued to play a part in philosophy of religion written in English, even more so in other languages. Process philosophy, for example, inspired by the thought of Alfred North Whitehead and exemplified in the ongoing work of Charles Hartshorne and others, has retained influence in philosophy of religion and in theology, probably more than in other areas of philosophy. Phenomenology, postmodernism, and other approaches characteristic of the European continent inspire important contributions to the subject. Indeed, there is often not a sharp line between different approaches. Continental writers such as Søren Kierkegaard figure extensively in undoubtedly analytical writing about religion, and analytical philosophy of religion makes such extensive use of medieval material as to be more or less continuous with neoscholastic treatments of the subject.

Although there had been a few earlier analytical essays about various religious issues, the main development of analytical philosophy of religion may be said to have begun in the 1950s with discussion of the "logical positivist" challenge to the cognitive significance of religious language. Most analytical philosophers then held, or were strongly tempted to hold, as an empiricist principle, that every (logically contingent) assertion, in order to have any cognitive meaning, must be verifiable or, more broadly, testable, in principle, by experience. It was charged, by Alfred Jules Ayer, Antony Flew, and others, that the affirmations of religious belief typically do not satisfy this criterion of meaning (A. Flew, R. M. Hare, and B. Mitchell in Brody 1974).

How, then, were the apparent truth claims of religions to be understood? Some were prepared, with Ayer, to treat major religious assertions as mere expressions of emotion, without any cognitive significance. Others sought ways of understanding such assertions as empirically verifiable in principle. John Hick (in Brody 1974) argued, for instance, that "eschatological verifiability," in a life after death, provides at least a partial solution to the problem. Still others, while granting that empirical testability is decisive for the meaning of typical factual assertions, sought to establish a different, and not merely emotive, type of meaning that could be ascribed to religious assertions. The most influential attempts of this type were inspired by the later writings of Ludwig Wittgenstein, particularly by his account of "language games" and their relation to forms of life.

The Wittgensteinian approach, as developed, for example, by Norman Malcolm (in Brody 1992) and D. Z. Phillips (1970), has generated very interesting studies of the relation of religious language to religious life. It is widely criticized, by some as giving inadequate weight to the apparent straightforwardly realistic intent of typical religious assertions, and by some as improperly shielding religious claims from rational criticism by relativizing them to religious language games. It remains, nevertheless, an important strand in contemporary discussion. Of all that has been done in analytical philosophy of religion, it is probably the discussion of religious language in general, and Wittgensteinian themes in particular, that have most interested professional theologians, perhaps because these themes have seemed more relevant than more metaphysical discussions to the work of interpretation and reinterpretation of traditions in which theologians are so much engaged.

Within analytical philosophy during the 1950s the verifiability criterion of meaning was already undergoing severe criticism and has since been virtually abandoned in anything like its original form. Many analytical philosophers continue to consider themselves empiricists and seek alternative ways of excluding claims that they regard as objectionably metaphysical. Many others, however, see the permanent contribution of analytical philosophy, not in a form of empiricism, or in any set of doctrines, but in a method, style, or discipline that can be applied to virtually all the historic issues of metaphysics and ethics and can be used in developing and espousing almost any of the classic philosophical doctrines.

The majority of work done in analytical philosophy of religion since the 1960s has been inspired by the later conception of analytical philosophy and has not focused on issues about religious language. It is characterized by metaphysical realism, taking the religious claims under

discussion to be straightforwardly true or false. (For defense of this stance, see, e.g., Swinburne 1977, chaps. 2–6.) Some have suggested calling it philosophical theology rather than philosophy of religion, because the principal subject of most of it is God rather than human religious phenomena, though atheists as well as theists have certainly been important participants in the discussion. On this basis, mainly since 1960, a very substantial body of literature, dealing with most of the traditional issues of philosophical theology and some new ones too, has been created.

Among the traditional topics the attributes of God received rather early analytical attention. (For general treatments see Swinburne 1977, chaps. 7–15; Kenny 1979; Wierenga 1989.) Analysis of the concept of God was easily seen as an appropriate subject for analytical philosophy, and issues about the attributes had been connected, since the Middle Ages, with problems about predication, an appealing point of entry into philosophical theology for those interested in the philosophy of language. According to some of the most influential medieval theologians, God is so different from creatures that positive attributes of creatures cannot in general be predicated of God univocally, that is, in the same sense in which they are predicated of creatures. How then can we predicate anything of God? Various Scholastic theologians developed various solutions, the best known being the theory of analogical predication of Thomas Aquinas. Analytical philosophers of religion have taken up the problem and some of the medieval views, along with more contemporary concerns—for instance, about the ascription of psychological predicates to a being who is supposed not to have a body. (Cf. Maimonides, Thomas Aquinas, and Alston in Brody 1992).

The two divine attributes that have received the most extensive analytical discussion are omniscience and eternity. The central issue about eternity is whether to understand it (as medieval and early modern theology generally did) as involving existence outside of time or rather as involving existence without beginning or end in time, as many contemporary thinkers have proposed. Critics of divine timelessness, such as Nelson Pike (1970) and Nicholas Wolterstorff (in Brody 1992), have questioned the compatibility of timelessness with God's consciousness or action or interaction with creatures. Eleanor Stump and Norman Kretzmann, however, have presented an influential defense of the traditional timeless conception (in Brody 1992), and the issue remains vigorously debated.

Omniscience and eternity are related topics, for one of the most discussed issues about God's knowledge concerns God's relation to time: Does God have complete knowledge of the future? In particular, does God know, infallibly and in every detail, how free creatures will use their freedom? Traditional theologies generally gave an emphatically affirmative answer to this question; but some modern philosophers and theologians have disagreed, arguing that the doctrine of total, infallible foreknowledge compromises the freedom of the creatures. The extensive analytical literature on this issue (e.g., in Fischer 1989) is continuous with older discussions, and opinion remains divided.

A related old debate, recently revived, concerns what has been called "middle knowledge": Does God know, completely and infallibly, what every actual and even merely possible free creature would freely do (or would have freely done) in every possible situation in which that creature could act freely? In the late sixteenth century, Luis de Molina, a Jesuit, proposed an ingenious theory of divine providence according to which God uses such subjunctive (and largely counterfactual) conditional knowledge to control the course of history without having to interfere metaphysically with the freedom of creatures. This theory of middle knowledge was widely embraced by Jesuits, but opposed by Dominicans, who argued that there cannot be such determinate conditional facts about everything that would be freely done by particular creatures in all possible circumstances. This historic controversy was introduced into current analytical discussion by Anthony Kenny (1979) and Robert Adams (1987), who have both defended the Dominican objection to middle knowledge; but the opposite position has been argued by a vigorous school of contemporary Molinists, including Alvin Plantinga (1974) and Alfred Freddoso (1988).

Regarding the relation of God to ethics, it was almost universally held in the 1960s that fundamental ethical principles must be independent of theology and that an acceptable theological account of the nature of ethical facts is impossible. Since then, however, it has come to be widely held by theists, and granted by many nontheists, that facts about God, if God exists, could play a central role in explaining the nature of ethics and that theistic philosophers should be expected to avail themselves of this possibility. The most discussed type of theological theory in this area is the divine-command theory of the nature of ethical obligation, or of right and wrong (Helm 1981). Several thinkers, such as Philip Quinn (1978), have tried to reformulate and explain the theory in such a way as to defend it against the traditional objections to it.

Adams (1987) has proposed a form of the theory that rests on semantical assumptions very similar to those of some of the most influential contemporary exponents of metaethical naturalism but employs different (theistic) metaphysical assumptions.

The grounds proposed for belief or disbelief in the existence of God have naturally claimed at least as much analytical attention as the attributes of God. This is a subject so intensively discussed for centuries that one might have expected little novelty in the treatment of it. But in fact investigations have been rather innovative, and the state of debate has changed significantly since 1960. One striking change is that the traditional arguments for the existence of God, then widely dismissed, even by theologians, as hopelessly discredited, have many defenders at the turn of the twenty-first century.

This is connected with a more general phenomenon, which is that analytical philosophers, especially those inclined to construct and defend constructive metaphysical theories, demand less of arguments than has commonly been demanded in the past. Virtually no one thinks any one "theistic proof" conclusive; but if arguments must be either conclusive or worthless, there would be little useful reasoning about any of the most important philosophical issues. Theistic apologists are accordingly less apt to seek a single "knockdown" proof than to try to show that several traditional (and perhaps also novel) arguments have something of value to contribute to a "cumulative case" for theism, an approach exemplified by Richard Swinburne (1979). Extensive work has been done interpreting, developing, and criticizing all the main types of theistic arguments. Those that have probably received the most attention and development are the "ontological" and the "teleological" (to give them their Kantian names).

The fallaciousness of any ontological argument and the contingency of all real existence had become such commonplaces, especially among empiricists, that it had a certain "shock value" when Norman Malcolm in 1960 published a defense of an ontological argument (reprinted in Brody 1992). Malcolm claimed to find in Anselm's *Proslogion,* besides the famous argument of its second chapter, a second ontological argument in which it is not existence but necessary existence that figures as a perfection. Malcolm also held that necessary existence cannot be excluded from theology on general philosophical grounds. Whether a statement expresses a necessary truth, he argued, depends on the language game in which it figures; and a religious language game can treat the existence of God as a necessary truth. These two features of Malcolm's article foreshadow the main tendencies in the development of ontological arguments since then: (1) attention to more modal versions of the argument and (2) the attempt to rehabilitate the idea of necessary existence.

Ontological argument studies have been greatly influenced by the dramatic development of modal logic, which was gathering momentum in the 1960s and burst into the center of American philosophical consciousness in the 1970s. In 1962 Hartshorne published a modal proof of the existence of God relying only on the premises that God's existence must be necessary if it is actual and that God's existence is at least possible. Subsequent discussion has established that this proof, and related proofs from slightly slenderer assumptions, are valid in the system of modal logic (S5) most widely thought to be appropriate for the context. David Lewis (in Brody 1974) and Plantinga (1974) have given the argument a form that takes account of developments in modal predicate logic as well as modal propositional logic (or in *de re* as well as *de dicto* modality). The argument is still of limited value for proving the existence of God, because those who would otherwise doubt the conclusion are likely to doubt the possibility premise, given the rest of the argument. But the modal development of the argument is helpful in structuring discussion of questions about necessary existence.

In the 1950s it was the opinion of almost all analytical philosophers that the existence of a real being, such as God (as distinct from merely abstract objects, such as numbers), cannot be necessary in the strongest, "logical" sense. This opinion has come to be widely doubted, however, and the traditional view that God should be conceived as an absolutely necessary being has regained a following. (For contrasting views see Adams 1987, chaps. 13–14, and Swinburne 1977, chaps. 13–14.) Several factors have contributed to this change. The identification of necessity with analyticity, on which the rejection of necessary existence was commonly based, is under attack. W. V. O. Quine's influential doubts about the adequacy of the notion of analyticity led Quine himself to skepticism about necessity. But others, influenced in some cases by an interest in necessity *de re,* have been inspired to seek a more robustly metaphysical conception of necessity. Since a conception of the latter sort was generally held by the great philosophers of the Middle Ages and the seventeenth century, a growing and more sympathetic understanding of those periods of the history of philosophy has also tended to undermine the most dismissive attitudes toward the idea of necessary existence.

The most popular argument for the existence of God in the eighteenth century was the teleological or design argument, usually in a pre-Darwinian form drawing its evidence largely from biological adaptations. This type of argument was discredited both by the devastating critique it received in David Hume's *Dialogues concerning Natural Religion* and by the development of an alternative explanation of the biological phenomena in terms of natural selection. A major rehabilitation of the design argument has been undertaken by Swinburne (1979). Instead of the biological evidence, he takes as his principal evidence the most pervasive, highest-level regularities in the universe. Since they constitute the most fundamental laws of nature, to which all scientific explanations appeal, he argues, there cannot be any scientific explanation of them. There may therefore be no viable alternative to a theological explanation for them, if they are to be explained at all. Deploying the apparatus of Bayesian probability theory, and responding to Hume's objections, Swinburne tries to establish that a theological explanation is indeed more plausible than no explanation at all. Swinburne's argument depends at some points on controversial metaphysical theses and has inspired an extended atheistic response by J. L. Mackie (1982); but the teleological argument has at least been shown to have much more philosophical life in it than had been thought.

The leading argument for atheism, aside from the various critiques of theistic arguments, has long been the argument from evil. The evils that occur in the world are incompatible, it is argued, with the existence of an omnipotent, omniscient, perfectly good God. In the earlier years of analytical philosophy of religion this was usually a charge of demonstrable, logical incompatibility; and attempts to provide theists with a "solution" to the "problem of evil" concentrated accordingly on trying to show the possibility of a perfect deity having permitted the evils. Borrowing a Leibnizian idea, for instance, Pike argued that for all we know, this might be the best of all possible worlds (in Adams and Adams 1990). Plantinga (1974) developed a much-discussed version of the traditional "free will defense," arguing that even if there are possible worlds containing less evil, and as much moral good, as the actual world, an omnipotent God may have been unable to create them because it may be that creatures (whether humans or angels) would not have freely done what they would have to do freely in order for one of those worlds to be actual. The adequacy of such theistic responses to the "logical" form of the argument from evil has been keenly debated, but it has probably become the predominant view that the argument does not afford much hope of a tight, demonstrative proof of atheism.

There has therefore been increasing interest in probabilistic arguments from evil, as presented, for example, by William Rowe (in Adams and Adams 1990), whose thesis is that evils show theism to be implausible, or at least constitute evidence against theism, which might contribute to a cumulative case for atheism. Theistic responses to this type of argument must address issues of plausibility and not merely of possibility. Some have been methodological, attempting to show that the relevant probabilities cannot be determined, or that the explanatory structure of the situation keeps the evils from being even relevant evidence (e.g., Stephen Wykstra in Adams and Adams 1990). Others have tried to give plausible accounts of why evils might have been necessary for greater goods. One widely debated hypothesis, developed in different ways by Hick (in Adams and Adams 1990) and Swinburne (1979), for instance, is that evils, and possibilities of evil, play an essential part in making the world a context for the moral and spiritual development of free creatures.

All such explanations of why God would permit great evils have seemed to some morally or religiously objectionable. Among theists who take this view, Marilyn Adams has argued that we should accept that we simply do not know why God has permitted horrendous evils but that within a religion that affirms, as Christianity does, God's love for individuals who suffer them, it is important to have a coherent account of how God may be seen as redeeming them (Adams and Adams 1990). She points to traditional religious ideas of suffering shared with God or with Christ as suggesting how horrendous evils might be "defeated" by forming an organic whole with incommensurably great religious goods.

One of the more dramatic developments of the period under review is the development of a defense of the rationality of theism that professes not to be based on arguments or evidence. Plantinga maintains that belief in the existence of God can be "properly basic," a basic belief being one that is not inferentially based on any other belief (Plantinga and Wolterstorff 1983). It has been held by many that some beliefs (formed, perhaps, in sensation or memory) do not need inferential support from other beliefs for their justification. Plantinga argues that more beliefs than some have supposed are reasonably held without being based on the evidence of other beliefs and that there is no compelling reason to deny that some religious beliefs have this basic status. He suggests that religious beliefs not based on "evidence" constituted by other

beliefs may nonetheless be based on other sorts of "grounds," which might be found, for example, in religious experience. Plantinga's view (which he has dubbed "Reformed epistemology") has been keenly debated. One of the most discussed issues is whether it allows an adequate basis for distinguishing between rational and irrational religious beliefs. (For a moderately critical view see R. Audi in Audi and Wainwright 1986).

A related but importantly different view has been developed by William Alston (1991). Religious experience has been a major subject of discussion in philosophy of religion (e.g., W. James, W. T. Stace, and C. B. Martin in Brody 1974; Wainwright, 1981), as it has been in modern theology. Not all of the discussion has been epistemological or focused on the justification of belief. Pike (1992), for instance, has written about the phenomenology of mysticism, arguing, against the older theory of Stace, that there are mystical experiences of theistic as well as nontheistic content. Alston's approach is thoroughly epistemological, however, and he focuses on the experience of more ordinary religious believers rather than of those adepts typically singled out as "mystics."

Relying on carefully discussed analogies with sense perception, Alston argues that in some circumstances experiences as of God addressing, or being present to, a person can reasonably be regarded as perceptions of God. His argument is placed in the context of a "doxastic practice" conception of the justification of beliefs. He argues that we are able to form and justify beliefs only in socially established practices in which we have learned to be responsive to such factors as experiential cues and communal traditions as well as to beliefs that we hold. In Alston's view we have no choice but to rely on socially established doxastic practices, and it is presumptively rational to do so, even though we typically have little or no independent evidence of the reliability of the practice. He argues that this presumption of rationality applies also to religious doxastic practices that are socially established, and in particular to practices in which participants have learned to form beliefs of having perceived God in various ways. Alston offers vigorous rebuttals of several major objections to basing religious beliefs on religious experience. In his opinion the most serious problem for his view, which he treats at some length, is that posed by the existence of diverse religious traditions whose well-established doxastic practices lead them to form apparently conflicting beliefs on the basis of their religious experience.

For philosophy of religion as for contemporary theology, the problem of conflicting truth claims of different religions is, if not a new issue, one that is coming into increasing prominence. Hick (1989) has done much to draw attention to it. He argues that it is not plausible to suppose that one traditional form of religious experience is veridical while others are not, and he tries to articulate a way in which many apparently conflicting forms could all be at bottom veridical, proposing to regard them as apprehending different "phenomenal" manifestations of a single "noumenal" transcendent "reality." Not that Hick thinks all religious beliefs equally acceptable; the main criterion he proposes for the value of religious traditions and belief systems is their fruitfulness in producing morally and spiritually recognizable saints, people notably advanced in a transformation from self-centeredness to Reality-centeredness. Among the issues in the vigorous debate about Hick's view are the adequacy of the conceptual apparatus he borrows from Immanuel Kant and whether it is compatible (as he means it to be) with a fundamentally realist and cognitivist conception of religious belief.

See also Atheism; Ayer, Alfred Jules; Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Empiricism; Epistemology, Religious; Evil, The Problem of; God, Concepts of; Hare, Richard M.; Hume, David; James, William; Kant, Immanuel; Kierkegaard, Søren Aabye; Mackie, John Leslie; Malcolm, Norman; Modal Logic; Mysticism, Nature and Assessment of; Ontological Argument for the Existence of God; Phenomenology; Postmodernism; Quine, Willard Van Orman; Religious Experience; Religious Experience, Argument for the Existence of God; Religious Pluralism; Stace, Walter Terence; Teleological Argument for the Existence of God; Theism, Arguments For and Against; Thomas Aquinas, St.; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

Adams, M. M., and R. M. Adams, eds. *The Problem of Evil.* New York: Oxford University Press, 1990.

Adams, R. M. *The Virtue of Faith and Other Essays in Philosophical Theology.* New York: Oxford University Press, 1987.

Alston, W. P. *Perceiving God.* Ithaca, NY: Cornell University Press, 1991.

Audi, R., and W. J. Wainwright, eds. *Rationality, Religious Belief, and Moral Commitment: New Essays in the Philosophy of Religion.* Ithaca, NY: Cornell University Press, 1986.

Brody, B. A., ed. *Readings in the Philosophy of Religion: An Analytic Approach.* Englewood Cliffs, NJ: Prentice-Hall, 1974; 2nd ed., 1992. Somewhat different selections, both comprehensive and both excellent, in the two editions.

Fischer, J. M., ed. *God, Foreknowledge, and Freedom.* Stanford, CA: Stanford University Press 1989.
Hartshorne, C. *The Logic of Perfection and Other Essays in Neoclassical Metaphysics.* La Salle, IL: Open Court, 1962.
Helm, P., ed. *Divine Commands and Morality.* Oxford: Oxford University Press, 1981.
Hick, J. *An Interpretation of Religion.* New Haven, CT: Yale University Press, 1989.
Kenny, A. *The God of the Philosophers.* Oxford: Clarendon Press, 1979.
Mackie, J. L. *The Miracle of Theism.* Oxford: Clarendon Press, 1982.
Molina, L. de. *On Divine Foreknowledge* (Part IV of the *Concordia*). Translated and with an introduction by A. J. Freddoso. Ithaca, NY: Cornell University Press, 1988.
Phillips, D. Z. *Faith and Philosophical Enquiry.* New York: Schocken, 1970.
Pike, N. *God and Timelessness.* New York: Schocken, 1970.
Pike, N. *Mystic Union.* Ithaca, NY: Cornell University Press, 1992.
Plantinga, A. *The Nature of Necessity.* Oxford: Clarendon Press, 1974.
Plantinga, A., and N. Wolterstorff, eds. *Faith and Rationality.* Notre Dame, IN: University of Notre Dame Press, 1983.
Quinn, P. L. *Divine Commands and Moral Requirements.* Oxford: Clarendon Press, 1978.
Swinburne, R. *The Coherence of Theism.* Oxford: Clarendon Press, 1977.
Swinburne, R. *The Existence of God.* Oxford: Clarendon Press, 1979; 2nd ed., 1991.
Wainwright, W. J. *Mysticism.* Madison: University of Wisconsin Press, 1981.
Wierenga, E. R. *The Nature of God.* Ithaca, NY: Cornell University Press, 1989.

Robert M. Adams (1996)

PHILOSOPHY OF RELIGION [ADDENDUM]

Philosophy of religion has recently focused on a range of issues regarding God: semantic (concerning the meaning of the term *God*), metaphysical (concerning the reality and attributes of God), epistemological (concerning justified belief and knowledge regarding God's reality), and ethical (concerning the bearing of God on personal and social morality and the meaning of life). The entry by Robert M. Adams illustrates some of these issues; longer representative discussions can be found in William E. Mann (2005) and William J. Wainwright (2005). Another area of recent philosophy of religion concerns whether, and if so how, claims regarding God fit with the natural sciences. With growing recognition that the natural sciences are not deterministic, many philosophers have found room for a God who freely acts in history and in human lives (see Draper [2005] and the exchange between Worrall and Ratzsch [2004]).

The term *God* joins *religion* as among the most elusive in English. Its uses are remarkably diverse, and this contributes significantly to the difficulty in settling many apparent disagreements regarding God. If I use the term *God* in one way, and you use it in a different way, then we may find ourselves appearing to disagree about God but actually talking at cross-purposes. For example, if I use the term in such a way that God is capable of suffering, rejection, and even incarnation, but you do not, we will diverge significantly in our questions and answers regarding God. The underlying semantic divergence regarding *God* will yield, sooner or later, divergence in claims deemed acceptable or true regarding God. As a result, philosophical illumination of one's concept of God continues to serve a valuable purpose.

One important semantic lesson is that the term *God* is typically used as a title rather than as a personal name (on which see Pike 1970). Such use can easily avoid begging the question whether God exists. The title *God* can have an intelligible use even if no one satisfies the title. The title *God*, however, does not enjoy just one understanding among its users, even its philosophical users. For instance, some philosophical writers use the title to connote a timeless transcendent agent, whereas others allow for a God in time.

Philosophers of religion have recently pursued the following longstanding question: What cognitive support, if any, is there for the claim that God exists? The question attracts a wide variety of interpretations of *cognitive support*. The most familiar understanding of cognitive support is in terms of evidence, that is, what indicates, even if fallibly and nondeductively, that a proposition is true. Evidence for the claim that God exists indicates, perhaps fallibly and only probabilistically, that it is true that God exists. Evidence can come in differing strengths and can enable a claim to be beyond reasonable doubt.

DIVINE HIDDENNESS

We can now approach the problem of *divine hiddenness* that is beginning to occupy many philosophers of religion: If God exists, why do not all competent people have evidence that makes it beyond reasonable doubt for them that God exists? Many competent people claim not to have adequate evidence (for reasonable belief) that God exists. Some philosophers, however, deny that an all-loving God would be hidden in a way that permits rea-

sonable doubt about God's existence (see Schellenberg 2004).

FREEDOM RESPONSE

Proponents of the Freedom Response to divine hiddenness maintain that God hides to enable people freely to love, trust, and obey God (see Hick 1985, Murray 2002). Seeking to form truly loving relationships with people, God does not coerce people to respond in a particular way. As an all-loving being, God hides to avoid coercing people to respond, and some philosophers hold that this allows for inculpable nonbelief regarding God's existence. The Freedom Response, however, prompts this question: Could not God supply less obscure self-revelation without abolishing our freedom in responding to that revelation? God could, evidently, be significantly less hidden while preserving our freedom to deny that God exists. Some revelations of God's power would overwhelm us in a way that removes our freedom, but the removal of divine hiddenness seems not to require any such overwhelming revelation.

PROPER-MOTIVATION RESPONSE

A second response to divine hiddenness, the Proper-Motivation Response, implies that God hides to discourage a human response based on improper motives (see Pascal 1995, Swinburne 1992). According to this response, God's self-revelation without hiding would prompt us to selfish fear or arrogance. Aiming to discourage such fear and arrogance, God hides and, according to some philosophers, thereby allows for inculpable nonbelief regarding God's existence. However, the Proper-Motivation Response must face this issue: Could not God supply a less obscure self-revelation without eliciting improper motives in our response to that revelation? Must a world where God is less obscure be less susceptible to human pursuit of God that is humble and passionate? The mere fact of less obscurity in God's self-revelation seems not to undermine humble and passionate seeking after God. God could readily promote such seeking in a setting of less obscure divine revelation.

HIDDENNESS AND SIN

A third response to divine hiddenness is that human sinfulness accounts for typical failure to appreciate the evidence of God's reality through creation, history, and conscience (see the discussions in Moroney 2000, Plantinga 2000). Some proponents hold that every competent adult who does not believe that God exists culpably fails to believe and thus that there is no need to explain how an all-loving God could allow for inculpable nonbelief that God exists, at least among competent adults. The main problem with this response is that it offers no straightforward way for itself to be justified. We seem to lack the needed avenue to evidence to infer that, with regard to every person who does not believe that God exists, that person is culpable, owing to sin, for nonbelief. Some people cannot plausibly be diagnosed so readily.

MULTIPURPOSE RESPONSE

A fourth response to divine hiddenness, the Multipurpose Response, acknowledges that God has various purposes in hiding and that we are not in a position to identify all of God's specific purposes in hiding (see Moser 2002). Divine hiding is sometimes a constructive effort on God's part to encourage (deeper) human focus, longing, and gratitude toward God. God thus aims to take us to our own deepest resources and their ultimate inadequacy, where we acknowledge our needing God at all times. In apprehending God's absence, we can achieve a deeper appreciation of God's presence. According to the Multipurpose Response, occasional divine hiding occurs in the context of God's main desire to have people lovingly know God and thereby to become loving as God is sacrificially loving. According to this response, God's primary aim is to include all people in God's kingdom family as beloved children under God's lovingly righteous guidance. So, God wants humans to love God and thus to treasure God, not just to believe, however reasonably, that God exists. Mere reasonable belief that God exists will not meet God's primary aim for humans. For our own benefit, according to this response, God is after something more profound and more transforming than simple reasonable belief that God exists.

COGNITIVE IDOLATRY

If we reject or neglect transformation toward God's character of sacrificial love, we may be blinded by our own counterfeit "intelligence" and "wisdom." We will then lack the kind of filial obedience and humility appropriate to relating, cognitively and otherwise, to God. We will then have assigned the authority of God to ourselves or to some other part of creation. We would then be guilty of idolatry, the mistake of exchanging God's rightful authority for a false authority. We commit cognitive idolatry when we demand a certain sort of knowledge or evidence of God inappropriate to a filial relationship with God (on which see Moser 2002). We thereby run afoul of God's rightful authority in the cognitive domain. The Multipur-

pose Response implies that we are in no position to demand that God be revealed in a particular way.

The problem of divine hiddenness has affinities with the traditional problem of evil. One might think of inculpable nonbelief as a certain sort of evil that would not exist if there were a loving God. In any case the problem of divine hiddenness occupies many philosophers of religion in ways that bear on epistemology, semantics, and metaphysics.

See also Hiddenness of God.

Bibliography

Draper, Paul. "God, Science, and Naturalism." In *The Oxford Handbook of Philosophy of Religion*, edited by William J. Wainwright, 272–303. New York: Oxford University Press, 2005.

Hick, John. *Evil and the God of Love*. 2nd ed. London: Macmillan, 1985.

Howard-Snyder, Daniel, and Paul K. Moser, eds. *Divine Hiddenness: New Essays*. New York: Cambridge University Press, 2002.

Mann, William E., ed. *The Blackwell Guide to the Philosophy of Religion*. Malden, MA: Blackwell, 2005.

Moroney, Stephen K. *The Noetic Effects of Sin: A Historical and Contemporary Exploration of How Sin Affects Our Thinking*. Lanham, MD: Lexington Books, 2000.

Moser, Paul K. "Cognitive Idolatry and Divine Hiding." In *Divine Hiddenness: New Essays*, edited by Daniel Howard-Snyder and Paul K. Moser, 120–148. New York: Cambridge University Press, 2002.

Murray, Michael J. "Deus Absconditus." In *Divine Hiddenness: New Essays*, edited by Daniel Howard-Snyder and Paul K. Moser, 62–82. New York: Cambridge University Press, 2002.

Pascal, Blaise. *Pensées and Other Writings*. Translated by Honor Levi. New York: Oxford University Press, 1995.

Peterson, Michael L., and Raymond J. VanArragon, eds. *Contemporary Debates in Philosophy of Religion*. Malden, MA: Blackwell, 2004.

Pike, Nelson. *God and Timelessness*. New York: Schocken Books, 1970.

Plantinga, Alvin. *Warranted Christian Belief*. New York: Oxford University Press, 2000.

Schellenberg, J. L. "Divine Hiddenness Justifies Atheism." In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. VanArragon, 30–41. Malden, MA: Blackwell, 2004.

Swinburne, Richard. *Revelation: From Metaphor to Analogy*. Oxford, U.K.: Clarendon Press, 1992.

Wainwright, William J., ed. *The Oxford Handbook of Philosophy of Religion*. New York: Oxford University Press, 2005.

Worrall, John, and Del Ratzsch. "Does Science Discredit Religion?" In *Contemporary Debates in Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. VanArragon, 59–94. Malden, MA: Blackwell, 2004.

Paul K. Moser (2005)

PHILOSOPHY OF RELIGION, HISTORY OF

It is not easy to say when strictly philosophical thought about religion began, for religion has always involved thought or belief of some kind. Even in other fields much of our thought is incipiently philosophical, but this is much more so in an interest that tends to be all-embracing. Religion has always had a cognitive factor, observances of various kinds had a meaning and these would often be of a far-reaching kind, involving beliefs about an afterlife or the influence upon us of beings other than those who inhabit this world. At what stage such beliefs come to be questioned, and not just accepted as a matter of course or tradition, is difficult to determine. But there is evidence of early questioning of this kind, and of the consequent defense and speculation, in some cultures, for example in India. It is a moot point how much of this we would consider strictly philosophical. But it is certain that the period, from the eighth to the fourth century BCE, which saw such an upsurge of intellectual interest and culture simultaneously (and seemingly without much mingling of cultures) in different parts of the world, produced philosophical thought of a very explicit kind, including philosophical reflection about religion.

EASTERN TRADITIONS

HINDUISM. Perhaps the earliest example of philosophical reflection about religion is found in the *Upaniṣads.* These were committed to writing about the eighth century BCE but they reflect much that had been going on before. They are part of the corpus of Indian sacred writings known as the *Vedānta.* Even the earliest and simplest of these contain distinctive and shrewd anticipations of the views about life and the universe that came to be explicitly formulated in the *Upaniṣads,* and it would thus be misleading to say that religious thought began in India with the composition of the *Upaniṣads.* But it is in the body of writings known by that name that we have the first sustained and deliberate thought about religion in a form that has affinity with what we know as philosophy.

The *Upaniṣads* vary much in quality and purpose. There is also much variety within their more strictly philosophical content, but the dominant theme is that of the unity of the universe. This is sometimes thought of in a sense that eliminates all plurality, anticipating much that some mystics have held at later times. For others "the One" is involved in all things in a way which is transcendent and absolute but which leaves it vague what status is

to be accorded to finite things. This comes closer to the way God's transcendence has been understood generally in Western thought. But on occasion the *Upaniṣads* venture to be more explicit; some of their themes come close to those of G. W. F. Hegel and of post-Hegelian idealists in the nineteenth century; there is a clear insistence on the interdependence of whole and part in an all-inclusive system of reality, and this led also to speculations about the nature of the system and the function of the parts within it which suggest much that we read in idealist writings in our own times. There are also parts of the *Upaniṣads* that come closer to the Western notion of God as creator of a world of beings distinct from Himself and from one another. This is not unlike Christian theism and, in this respect, some passages of the *Upaniṣads* anticipate much which has since been central in Christian thought.

"The One breathed breathless" is a typically cryptic summing up of much of the teaching of the *Upaniṣads.* What it expresses is the profound and persistent sense of some ultimate nature of reality which escapes our understanding. The world does not wholly explain itself, it is rooted in mystery, and this means more than that there are things which are beyond our particular understanding at a certain time. All things point beyond themselves to a mystery that is in principle beyond our grasp or to some unity of things in the universe which is in some way more complete and final than the interrelations of things as we trace them in our normal understanding of the world. This is the significance of the terms that occur so often in Indian thought—"not this, not that" and "I am that." In this context these reflect a sense of some ultimate transcendent reality which is very vigorously presented in the *Upaniṣads* and whose implications are sometimes very explicitly set forth. It is indeed a very significant fact that there should be so shrewd a philosophical grasp of this notion at such an early date, and this makes the *Upaniṣads* a work of considerable significance for our understanding of religion in general. They contain also much explicit philosophical argument that is highly relevant to philosophical controversies about religion today. This covers many aspects of religion besides those that directly concern the dominant theme of the unity of all reality.

The *Upaniṣads* contain also much reflection upon our practical attitudes. This tends to be of the "world-denying" type and severely ascetic; that is not surprising where the dominant theme is the ultimate oneness of all things. But we find also in the *Upaniṣads* much emphasis on social service, on compassion, virtue, and welfare. Even if the views adopted on such matters seem to Western eyes too strictly determined by the sense of ultimate union with the whole, and even if it is true, as even some leading Hindus have stressed, that the otherworldly feature of Indian religion has led to apathy and indifference to present concern, there is also much to be learned from the insights we find in the *Upaniṣads,* as in later Indian thought, about the true nature of compassion and selflessness.

DAOISM AND CONFUCIANISM. Not much later than the time the *Upaniṣads* were committed to writing, there appeared in China philosophical teaching and writing about religion which had also at the center of it a sense of some ultimate unity of all reality. This is the essential significance of the doctrines of Dao (expounded in the *Daode Jing* traditionally ascribed to Laozi—born 604 BCE—and in later writings like those ascribed to Liezi and Zhuangzi); and this in turn reflects a generally more basic notion that lies behind most early Chinese thought about religion, the idea of a "heaven and earth relationship." What this implies is that there is some character of reality beyond what we find in the world around us but which cannot be explicitly defined or grasped. We can only know it in its requirements and in the sense of some kind of justice operative in the universe at large. The "beyondness" of the power which works for righteousness in this way is deliberately softened; it is almost as if it could only be known from within. But this is itself a very significant fact, and the elusiveness of the influence to which our lives are subject in this way in Chinese thought is no mean indication of the subtleness of their philosophical and religious insights. It has in fact led sometimes to the view that Chinese religion, and especially Confucianism, is entirely a moral or religious system. That impression could easily be derived from *The Analects* of Confucius (551–478 BCE), since they are concerned mainly with ethical and social matters, especially those which concern the appropriate "orders" in society. But Confucianism is in fact extensively determined and overlaid by notions like that of a heaven and earth relationship mentioned above. The distinctive thing, for philosophy, about early Chinese religion and thought about religion is the shrewd sense that the nature of what lies beyond present existence and gives it meaning is best discerned by following a Way or path. The goal is, as it were, best reflected for us in the way it is to be attained. If this is not the whole truth, it is a significant pointer to it.

BUDDHISM. At a slightly later date we have the founding of the Buddhist religion in India. This led to the compo-

sition of the Pali Canon, containing, it is alleged, the substance of the teaching of Buddha. The canon was closed in the reign of King Aśoka (273–231 BCE) but not committed to writing until the first century BCE. It is not implausible to conclude that it does reflect fairly closely the actual teaching of the historical Buddha. The Pali Canon is of exceptional interest to philosophers today. It contains acute philosophical thinking, and some incline even to think of Buddhism as being more a philosophical system than a religion. That is certainly a mistaken impression, but we have in Buddhism a very shrewd grasp of the nature of religion as philosophy illuminates it. The purport of this has often been grievously misunderstood, not least in the assumption that Buddhism is a religion without God. The mystery of transcendent being is at the center of Buddhism and has remained so through most of its history and in its many varieties. This may not take quite the same form as in the West or find closely parallel expressions elsewhere, but it is unmistakable to anyone who knows his way about the subject.

A peculiarly distinctive feature of the doctrines of the Pali Canon is the subtle understanding of the difficulty of characterizing a reality that is "beyond" in the sense in which the infinite must be. It is in this context that we are told that we must not say that God exists or that He does not exist. At one point we have a list of sixty-two typical metaphysical questions that must not be asked. This is closely in line with much that has been maintained today in various forms of antimetaphysical philosophy, and it is strange how little appreciation there has been, on the part of recent positivists and agnostics, of how much grist of a sort there is to their mill in the doctrines of the Pali Canon. But it might all the same not be grist they could altogether accept, least of all if they fully grasped its implications in its contexts. For here we have skepticism and positivism with a difference. It springs less from a radically empiricist outlook than from a profound sense of the elusiveness of transcendent reality, and this makes much of the teaching of the Pali Canon uniquely relevant to philosophical controversies about religion today. The account of such matters as Buddha's enlightenment reinforces this, for while this can plausibly in fact be given an atheistic interpretation, it does point suggestively to a subtle grasp of the transformation of present reality through the invasion of it by a reality of an entirely different order which beggars all description. In these and kindred ways the Pali Canon, like related further aspects of Buddhist and of Hindu thought, has close and instructive points of affinity with the cruxes of religious thought today; and this is being increasingly understood by some experts in this field.

PHILOSOPHERS. There has been a long line of impressive Asian thinkers who have attempted variations and refinements on the themes just outlined. Among the most important are Śankara (c. 788) and Rāmānuja (c. 1017). In recent times the more traditionalist type of Hindu thought is well represented in the works of Radhakrishnan, while we have in the very liberal writings of Śri Aurobindo an attempt at reform that is sharply opposed to the objectionably otherworldly aspect of Hinduism and that tries to come to grips with the notion of some divine disclosure which leaves the individual a free and responsible creature.

GREECE

In the Western tradition philosophy begins with the Greeks, and to give a full indication of the course of religious philosophy in the West would be to outline the main continuous progress of philosophy from the Greeks to the present day. For almost all the main philosophical notions and the main divisions of opinion in philosophy (realist, nominalist, idealist, and so forth) have entered into religious controversy in one way or another. The matters that can be noted in the remainder of this entry must thus be highly selective.

PARMENIDES AND HERACLITUS. In Greek thought, as in that of the Orient, there has been a central preoccupation with the problem of the one and the many. In the work of Parmenides this took a very distinctive and influential form. He proceeded by way of analysis of the nature of thought. This he found to involve predication, the affirmation of one thing about something else. To think is to say of an identifiable *A* that it is *B*; it is some relating of terms in a system that makes the relations possible. But there is an element of exclusion in such predication. If I say that this book is blue, that precludes its being black, although of course it says nothing about its being round or square, etc. All determination, as it is put, is negation. But does not this raise peculiar problems? For negation seems to be some odd sort of affirmation of what is not the case. It appears thus to deal with what is not. But what is not, Parmenides thought, is just altogether unreal—and no one can think or affirm this. But if negation becomes impossible in these ways, affirmation appears also to stand condemned, and there seems thus to be something radically unsatisfactory about thought itself and about the world as thinking apprehends it. Parmenides concluded that it was a mistake to suppose that the universe was a system of terms in relation, of the many which change and come into being and go, and that we must therefore think of all reality as one undifferenti-

ated whole—conceived by him also as a sphere extending in the same way in all directions. There was given in this way a logical form to a profound religious sense of some ultimate all-embracing unity.

By contrast we find, in the work of other Greek philosophers, an emphatic insistence on the reality of the here and now and the world of variety and change. Protagoras took this to the length of insisting, in anticipation of much later empiricism and relativism, that nothing is real except as it appears. Neither the external world nor our moral ideas have any independent or objective reality; and this view of things received distinctive expression also in the thought of Heraclitus, who insisted that all things were in flux and that "we cannot step twice into the same river." But this was supplemented by Heraclitus by the notion of a pattern of change in which some principle or "logos" was expressed. For him, as for Parmenides, this carried with it a poetically mystical religious undertone. The idea of fire, as a central element, functioned as a symbol of that.

PLATO. In due course Plato was to take up the problems presented in the way described above. He carefully restated and developed the difficulties that troubled Parmenides and Heraclitus and started a program of reconstruction by dealing firmly with the problem of negation. He observed that this does not involve reference to a wholly unreal, to mere nothing. It could be amply provided for within the notion of terms in relation, for to say that something is not is just to say that it is other than something else, to indicate precise location within a system of interrelations. But if thought, as involving determination of this kind, is to function accurately, the system within which it operates must be a strict and tight one. Where is this to be found? Plato thought he found it preeminently in mathematics, and he thus came to regard mathematics as the true propaedeutic to philosophy and a paradigm of its method. The realities which could be properly thought and known had thus to be quasi-mathematical ones, and they consisted of general forms or principles which were real in their own right and bestowed on all other things whatever reality those could properly claim. This left Plato with the hard problem of accounting for the particulars and the changing course of things in the world, and it is not certain that he arrived at a view of this question which contented him. He sometimes spoke of particulars imitating the forms and sometimes of their participating in the reality of the forms, but the individual and unique existent had never more than a problematic place in Plato's philosophy.

Difficulties also arose in yet another way, for even in its more rarefied instances, as in mathematics, there appears to be something essentially inadequate about the process of relating terms in a system. Every relation, including the relation of whole to part, seems to require yet another, or another system, to make it possible. All explanations of one thing in terms of others leave us with further questions and matters unexplained—there is no natural limit to the process of thought—and for the Greeks in particular that which is without proper limit is unsatisfactory—evil, they said, is of the infinite. Plato was led in this way to the notion of some yet more perfect reality, some quite different mode of unified existence in which present imperfect relatedness disappeared, and he held that everything had its reality exhaustively determined by this ultimate nature of the universe. To this he gave the name "the Good," and he declared that, in the sense indicated, this Good was "beyond being and knowledge." He did not mean that it was not real, or a mere notion—far from it. But it could not be given the sort of determinate existence and intelligibility which we ascribe to the sort of entities our minds can understand and encompass.

This is the first explicit formulation in Western thought of the idea of transcendence as it came to dominate much subsequent thinking. It is evident that it owes much, not only to Parmenides' puzzles about predication and nonbeing, but also more directly to Parmenides' insistence on some ultimate all-encompassing unity of being. But it does not involve the elimination of all plurality. When his system seems to involve that, Plato turns back on himself in vigorous protest—as in the famous passage in the *Sophist* where he insists that there must be "place in that which is perfectly real" for "change, life, soul, understanding." The specific forms, metaphysical as well as mathematical, had their place in the one universe in which everything derived its significance from the central all-encompassing reality of the good, and these forms lent some sort of reality to the particulars and to individual lives in the normal sense. The relation of particular to universal and of this to the Good, the ultimate supreme reality, may not have been worked out in a satisfactory way. But at least we have the notion that all we find in the world derives eventually from some one transcendent source in which all imperfection is resolved.

The formulation of these ideas owed much to the influence upon Plato of the Eleusinian mysteries and Orphic cults with which he came into contact—and also to the religiously orientated teaching of Pythagoras. In turn, it affected his teaching on what may appear to be

more specifically and recognizably religious conceptions, like his doctrine (in the *Timaeus* and the *Laws*) about the Demiurge who fashions the world according to the eternal patterns and his belief in preexistence and immortality. But it is not primarily in what he says about these more conventionally religious notions that Plato shows his main penetration or had his more abiding influence on religious thought. His notion of a system of forms held together in the transcendent unity of the Good was a more radically instructive and formative notion—although the teaching of the *Laws* and the *Timaeus* prescribed much of the form of later natural theology. It accorded best also with the element of mysticism which tempered the rationalism of his precursor to whom he was deeply indebted, namely Socrates. It is thus in the notion of the Form of the Good that Plato comes nearest to the idea of God in subsequent theism, but his approach to the subject left him no way in which his supreme and central principle of the Good could acquire the character of a person. That was precluded by the severely rationalist nature of Plato's main approach to his task and the consequent exclusion of any kind of revelation of an active concern, which could only be mediated through the actual particulars of life and history that figured in such an ambiguous and unimpressive way in Plato's philosophic outlook.

ARISTOTLE. Our next main landmark is the philosophy of Aristotle. He did not separate the universal as completely as Plato did from the particular, although it is a moot point, still much debated, how ultimate is the difference between Plato and Aristotle here. But the difference did lead in due course to notions of the union of form and matter and of mind as the informing principle of the body by which much subsequent thinking on questions of this kind was directed. For Plato the properly mental side of human life was sharply separated from the body, and along with this went a low estimate of the body—although the body was not thought to be evil, as in much subsequent teaching. The mind is apt to be thought of by Plato as imprisoned in the body and awaiting its release. On the slant given to the subject by Aristotle there is a much closer integration of mind and body and this has been the model for a great deal of later thinking about human personality and the belief in resurrection. The mind is thought to require at least some kind of body, and there are philosophers who regard mind and matter as coextensive in the universe in general. Others have taken the Platonic lead in propounding a very sharp dualism of mind and body.

In strictly religious matters the difference between Plato and Aristotle here seems to become narrow; for although we have no strict equivalent to the Form of the Good in Aristotle or the same insight into the transcendent character of the ultimate religious reality, we do have an "Unmoved Mover" whose relation to the course of events He affects is a somewhat remote and detached one. The God of Aristotle is little involved in the world; it would have been a sign of inferiority and imperfection for Him to be so. This reflected a typically Greek attitude. To be affected by something external to your self is an indication of weakness, and in Aristotle's ideal of the "Great-Minded Man" this is very marked—he will not be cruel to his inferiors just because they are beneath such notice.

The Stoics came later to pride themselves on their independence and self-sufficiency. Likewise the God of Aristotle is absorbed in contemplation of His own perfection; He takes no overt interest in other things, but He moves all other things by attraction. This is in sharp contrast with subsequent Christian teaching and represents the main way in which Christianity is "foolishness to the Greek." But the idea of an Unmoved Mover did nonetheless have a very extensive influence on later religious thought: It provided the model for the famous causal arguments for the existence of God. We have somehow to account for the world, and since we cannot account for it in terms of the way events determine one another within the world, we must have recourse to some altogether different mode of determination and explanation; and in due course this consideration became one of the main ways in which religious thinkers presented the idea that the world as we find it is dependent on some reality which is altogether "beyond" or transcendent. Here, as elsewhere, Aristotle determined very closely the style, if not always the substance, of later religious arguments.

This is evidenced specially in the way some of the further leading notions of Aristotle's philosophy, such as his distinction of potential and actual and his analysis of four types of cause and his notion of substance, became formative ideas in the religious thinking of later Christian times. It is in these ways, more than by very distinctively religious insight, that Aristotle made his main contribution to the philosophy of religion.

There is one further notion of great importance which had its place in Aristotle's system and became subsequently very influential. It is the idea of a law of nature. At times this was understood in a very relativistic way. To "follow nature" was taken to mean abiding by your own whims or impulses. It was sharply contrasted with con-

vention, and the latter came to be much derided in some quarters in the period after Aristotle and Plato in Greece—indeed earlier to some extent among the Sophists. Here we see again, in an extreme form, the ideal of being self-sufficient. This was carried by some of the Cynics and Epicureans and the early Stoics to the extent of trying to "return to nature" and doing without society and its irksome restrictions altogether—a cry that was sounded vigorously again in the seventeenth century. But it came to be realized that this policy led to absurdity and chaos, in personal life and in society; and thus the idea of "Nature" underwent complete transformation—it came to be taught that there was a nature to the universe at large ("Nature" with a capital N, as it were) and that this disclosed itself to men's reason. This led, in the fusion of the idea of law of nature with the Roman idea of a "law of nations," to the conception of a number of basic moral principles which were bound up with our rational nature and which, for many, further owed their firmness and objectivity to their foundation in the ultimate nature of the universe. This notion had a long and varied history and played a very important part in Christian accounts of morality and its relation to religion. It has a close affinity with the teaching of early Chinese religions and the notion of some power from beyond the world working for righteousness within it and prescribing our basic moral principles. Reflection upon this affinity can be very fruitful in seeking the way forward with such problems in the way they present themselves today.

EARLY AND MEDIEVAL CHRISTIANITY

The thought of early Christian times was extensively affected by Greek philosophy. This is evident even in the New Testament itself, not only in the way its authors write about matters like soul and body, but also in the central theme of "the Word" or Logos which became flesh. The Greek notion of Logos provided the basic concept in terms of which the doctrine of the Incarnation was to be understood. Directly, the concept of Logos came into philosophical thought in Christian times from the Stoics, for whom it meant originally an immanent World-Soul. But it was later combined with the Platonic idea of *nous* and so was conceived as acting in accordance with archetypal patterns. The basic problem was how is it possible to have knowledge of a strictly transcendent being, and for this a solution was sought in terms of an intermediary, in this case a logos, which was also induced in due course to fill other roles and help in the solution of further problems. These procedures came into Christian thought in the first place through the work of a gifted Jewish philosopher of the first century, namely Philo, and it had a prominent place in the subsequent Christology of formative thinkers like the Alexandrians, of whom Origen has most interest for philosophers. But what we have in the main during early Christian centuries is not so much philosophy of religion in the strict sense as theological writings that make extensive use of philosophical concepts. There were also some theologians of this period, as there have been of later times, who resented the intrusion of philosophy into the domain of faith. Of these the most outstanding was Tertullian.

The main exception to the normal course of thought in the early Christian period was Neoplatonism. Here we revert again to a profound sense of the Oneness of the Universe in a way that puts particulars and plurality in jeopardy, as they had been to some extent in the philosophy of Plato. But some account must be given of particulars, and there was developed in this way the difficult notion of emanation. God is the ultimate unity and He transcends all the categories of thought, but finite beings exist in the form of some falling away from the original perfection. This comes to terms in some fashion with the facts of finite existence and the reality of evil that occupied the minds of thinkers of this period a great deal. But it is very hard to make sense of the notion of emanation without calling in question the all-embracing nature of the one ultimate reality. The insistence on the latter notion did, however, influence the course of mystical thought and practice extensively. It also led, as in the case of Oriental mysticism, to attempts to draw away altogether from our present existence, with its limitations and evil, and to pass beyond the world of intellect as well as sense into total union with ineffable Being.

In sharp contrast to this teaching we have the position of thinkers who reflected anew on the significance of the Hebrew-Christian doctrine of creation. The Hebrews had come early to understand the elusive and transcendent character of God, and this had found very remarkable expression in parts of the Old Testament, the most famous passage here being the story of Moses at the burning bush. But this carried with it in Hebrew thought a subtle appreciation of the way a true discernment of God's transcendence required the recognition of our own distinctness as beings dependent on God. This sharpened, however, the question how such beings could in any way come to know God. The Hebrew answer was in terms of God's disclosure of Himself in history and experience, and this was deepened and extended in specifically Christian claims about the work and person of Christ. In this context the problem of revelation becomes a crucial one,

and it has remained at the center of Christian philosophy at all times except when insistence on the distinctness of faith precluded all rational consideration of it.

AUGUSTINE. Preoccupation with the way human beings, being finite, can come to know an Infinite Being lies at the center of the more specifically philosophical parts of the writings of Augustine. In his attack on the problem Augustine gives prominence to our reflection on what we find our own souls to be like as a clue to our understanding of the relation of God to the world. He set the pattern for much subsequent reflection on our own nature and started a concern for the inward aspect of personality which persisted through formative later thinkers, such as René Descartes and George Berkeley, to such nineteenth-century theologians as F. R. Tennant and the phenomenologists and existentialists of the present day. This side of Augustine's achievement is, however, often obscured by another. For although he emphasized the distinctness and freedom of finite beings, he came in another way to put these ideas in considerable jeopardy. In seeking to account for the redemptive work of Christ he posited the notion of an initial abuse of man's freedom leading to subsequent enslavement to sin. This gave considerable impetus to a doctrine of the Fall which, although not prominent in this form in earlier Christian times, became a central theme of much later theology and Christian profession of faith. The personal experiences of Augustine and his African background are thought to have greatly influenced his view in these respects, and there have certainly been voices, like those of Pelagius in his own time and Abelard later, raised in sharp protest against the rigors of the Augustinian doctrine of humanity's sin. The doctrine of the Fall has also been invoked to simplify the problem of our knowledge of God by blunting the strictly epistemological character of the problem; this came about through emphasis on the way our own allegedly corrupted nature made us spiritually blind and stood in the way of a vision of God. In the same context the idea of a law of nature became the idea of what is practicable in the present sinful state of humankind and society by contrast with the ideal law of God. This distinction was given much prominence by St. Augustine and has been reaffirmed, in the sense in which he understood it, by his most notable followers to the present day.

ANSELM. The question of particulars and universals became prominent again in the controversy of realism and nominalism in the early Middle Ages. It had many implications for religious thought. For example, the view that individuals do not exist in themselves was thought to culminate in pantheism in the sense that "all visible things pass into intellectual, and intellectual into God." This period also saw further attempts to provide a rational defense of the faith, although without denying that faith had a firm foundation of its own. An outstanding feature of this activity in philosophical thinking is the formulation of the Ontological Argument by St. Anselm. This was intended to show that sound understanding of the idea of God yields us the necessity of His existence. The idea of God, it was urged, is the idea of a being than whom nothing greater can be conceived. But a being that does not exist is inferior to one who has the additional attribute of existence. Many changes have since been rung on this argument and it is being much canvassed at the present day.

THOMAS AQUINAS. The most impressive achievements of the Middle Ages in religious thought came about initially through the work of Muslim scholars (Mohammad al-Ghazali and Averroes in particular) who were much concerned about the question of reason and revelation in their own faith. Among these there had also been preserved important works of Greek philosophy, especially those of Aristotle, which were not properly known by Christian scholars. There came about in this way a revival of the study of Aristotle and a new concern about the way a transcendent being could be known by limited finite ones. This culminated in the very comprehensive work of St. Thomas Aquinas, which ranged over most religious questions, seeking a synthesis of religious claims and established philosophical principles. It set up firmly one of the main forms of natural theology. For Thomas this covered two things. First we have the attempt to establish the existence of God by argument. This took the form of the famous "Five Ways." The first three of these are variations on the Cosmological Argument, as the term came to be used in due course. They seek to pass from the limited or contingent nature of finite things to an ultimate First Cause or Ground. The least elaborate, and also the most plausible, is the third way, which proceeds directly from the contingency of the world to its absolute Source without presupposing any particular view of cause and effect as we understand it. This argument, in one form or another, has been central to a great deal of subsequent philosophy of religion. Many hold today that it gets us at least very near the truth about the initial relation of God to the world and the way we know this. The other two "Ways" depend on notions of a scale of being and value and on the adaptation of things to their purposes, which are at least alien to the way we normally think about the world today—though they have their defenders.

The second prong of natural theology was that which sought, through an extremely subtle and cautious doctrine of analogy, to determine the attributes of God more precisely. It was urged that we cannot know God as He is in Himself, we can only know that He must be; and because God is a transcendent Ground of all things, He cannot be mirrored in the world He has made in the way an effect normally tells us something about its cause. Thomas and his followers were therefore well aware of the need to move very circumspectly here, and what they maintained was that God must be thought to have certain attributes, like goodness or power, in whatever way is necessary for Him to be the Author of those in the form in which they appear in the created world. In presenting this doctrine some very careful distinctions were drawn between various types of analogies. The main difficulty which this approach involves is that of determining whether anything of substance is added in this way to what is originally claimed in regarding God as a transcendent Being. There is in any case needed in addition extensive recourse to revealed truth to supply the particular affirmations of a faith like the Christian one. These truths of faith could not, according to Thomas, conflict with the truths of reason, but they go beyond them.

WILLIAM OF OCKHAM. The most formidable opponent of natural theology was William of Ockham, who questioned the ability of natural reason to discover in any measure the inscrutable will of God or reduce the mystery of transcendent being. His methods of procedure, involving the reduction of our postulates to the minimum that the facts require, anticipates many features of modern thought where skepticism about affirmations and alleged entities which pass beyond the facts of sensible experience and science is sometimes combined with a dogmatic affirmation of faith in which reason plays no part.

MODERN PHILOSOPHY

Outstanding formative philosophers of the modern period, roughly the last five hundred years, were of two main sorts, rationalists and empiricists. The former, including Descartes, Benedict de Spinoza, and Gottfried Wilhelm Leibniz, had great confidence in the power of reason alone to establish ultimate metaphysical truths.

RATIONALISM. Descartes claimed to prove his own existence by the power of reason alone and drew a sharp distinction between mind and body. He then sought, by severely rational arguments, to prove the existence of God. Two of these arguments invoke the causal principle, although they require also our having the idea of God; the third is a special form of the Ontological Argument; it contends that if we think of a being who does not exist we are withholding from our conception of it a "perfection," namely existence, which is essential to our conception of a perfect being. These arguments are not usually thought to succeed as they stand, but they can nonetheless be thought to be significant as indications of the insight into there having to be an ultimate reality in which essence and existence are one. They also illustrate the futility of seeking to establish the existence of such a being by arguments involving consideration of what limited finite things are like. Descartes's causal arguments are particularly illuminating in this way, as he imports into his premises, at every step in an elaborate argument, certain considerations derived from the notion of an infinite being which it is the aim of the argument to defend.

A further feature of Descartes's work is the insistence on the freedom of the individual—"liberty of indifference." This is bound up with the insistence on the distinctness of persons as nonmaterial entities. The same theme is taken up in Leibniz's monadology, in which every being is a distinct mental monad. But the genuineness of our freedom is jeopardized by Leibniz in his doctrine of preestablished harmony and the way each monad consistently unfolds in its history some destiny which its own nature prescribes for it from the start. In the ingenious monistic system of Spinoza freedom comes to be thought of in terms of accepting our place and destiny in the universe with adequate understanding and forbearance rather than in the form of genuine "liberty of indifference." Descartes's doctrine of the self as a distinct mental substance has been subjected to considerable criticism from time to time, not least at the present day. But there are many also who consider it an essential ingredient in a sound understanding of the relations of God to man and who stress, as did Descartes, the "interiority" and unextended character of the mind.

EMPIRICISM. Empiricism inclines to skepticism and is severely skeptical in its stricter forms. The great British empiricists did not all hold to their principle with the same consistency. We find John Locke departing from his avowed aim of showing that knowledge derives from sense impressions, not only in his theory of knowledge and his account of material and mental substances, but also in his expressly religious thought where he claimed, for example, that the existence of One Infinite Mind can be proved with the same certainty as we find in mathematics. There is much in fact in Locke's presentation of the causal argument in Chapter X, Book IV, of his *Essay*

concerning Human Understanding that has close relevance to controversies about the subject today. Likewise Berkeley, while dispensing with the notion of independent material substances, found in his account of the world of nature as dependent on its being perceived a firm foundation for the belief in a Divine Being on whose Mind the whole world of nature depends. To Berkeley we owe also a subtle appreciation of the distinctiveness of the way minds are known and the essential inwardness of personality which is so central a feature of religious philosophy today.

David Hume, however, was little attracted to these compromises and, although he confessed to some admiration for the argument which seeks to prove God's existence from the evidence of design in the universe, he adhered generally to a ruthlessly empiricist position. This involved total skepticism about God, immortality, and all properly religious notions. Hume contended that religion had started in a thoroughly naturalistic way with the personification of natural objects and so forth and that only at a late and sophisticated stage of culture did people arrive at some unification of religious notions and the belief in one God. His presentation of this view is delightfully lucid and it set the pattern for much of the anthropological treatment of religion later in the nineteenth century. In Hume's *Dialogues* there are also canvassed some of the main arguments that are used to support or reject religious beliefs, ranging from the general belief in God to belief in miracle.

KANT'S CRITICISMS. The "critical" philosophy of Immanuel Kant sought to arrest the skepticism of Hume without retreating to the strict rationalism of Descartes and his followers. Kant's main contention was that the sort of experience of the world which we undoubtedly have presupposes a unified world of objects presented to an abiding subject. The modes of unification thereby involved, the necessary conditions of experience, provided a new basis for confident belief in causality and substance, though not in the same sense as that of Descartes; but it was also implied that knowledge is confined to the world of our experience and the principles involved in this, sometimes thought to be imposed by the mind itself. This did certainly yield us the belief in an unobservable subject of experience, but nothing could be known of this beyond its being required to account for the sort of knowledge we have of the external world. There was also a tendency to isolate this inner self so completely from the external world of known reality that the functioning of the "pure self," especially as will or active agent, became very hard to conceive and set for Kant some of his main difficulties, especially in his ethics.

The limitations involved in the alleged "critical" account of knowledge were, however, extensively corrected by Kant in his insistence that we have certain grounds for "faith," which supplements what we can strictly know. These grounds of faith are found in the operation of our practical reason or moral awareness which sets before us certain moral obligations, largely in the form of strictly universal rules, which have in turn far-reaching implications. It was urged, for example, that there is a moral requirement that justice be rewarded, but that, since the ethical motive would be impaired if we set our own happiness as the aim of moral actions, God must be postulated to guarantee the eventual relation between happiness and virtue in the universe. Freedom and immortality were similar postulates of practical reason. These contentions have been subjected to much criticism, and doubt has been cast on the success of even the limited undertaking of postulating certain principles of a unified world of experience. Religious thinkers have urged that "faith" in its Kantian form has little in common with properly religious faith and that the severely rationalist character of the appeal to postulates of practical reason neglects the distinctively religious element in religious belief. On the other hand the prominence given to moral considerations in religious thought has been widely welcomed, and many writers have sought to provide versions of the moral and teleological arguments which are not open to the difficulties of those provided by Kant.

IDEALIST RESPONSES TO KANT. A great deal of post-Kantian philosophy was concerned with the gap in the Kantian system between the world as we apprehend it and the ultimate or "noumenal" reality of the world as it really is. For Kant these tended to be two separate worlds, but many thought this unsatisfactory and sought in various ways to understand the ultimate reality or "thing-in-itself" as some completion of the world as we find it—a notion that is in many ways anticipated in some of Kant's own reflections. There were thus initiated various metaphysical enterprises concerned especially with finding within the world of our own experience some reliable clue to the nature of the universe as a whole. The most influential of these was that of Hegel, who found the ultimate principle of reality in reason. We cannot exhaustively understand the universe but the universe is in principle capable of being understood through and through as a system where everything has its place and nature determined by rational necessity. Others (like

Arthur Schopenhauer) gave to will the preeminent place as a metaphysical clue.

There were many variations on these themes in the nineteenth century, including the work of British idealists such as Thomas Hill Green, F. H. Bradley, and Bernard Bosanquet and of American thinkers such as Josiah Royce. Idealism became the dominant philosophical view, and within the perspective of it many views were advanced about the relation of God to the world, taking distinctive features of our own experience as the clues to what lies beyond it. This tended to leave nothing essentially or irreducibly mysterious about religion. But the leading post-Hegelian idealist, namely Bradley, argued that there were radically contradictory features of present experience which implied that the ultimate nature of the universe was suprarational. And with this emphasis we come back again to the idea of some transcendent reality on which everything depends in some way that in principle we cannot understand. It was argued also, in criticism of the more rationalist type of idealism, that it left little room for the distinctness and freedom of the individual, since all beings came to be regarded as elements or "phases" or "appearances" of an ultimate all-inclusive system—and in the same way the problem of evil became a very acute one for idealist defenders of religion.

NATURAL THEOLOGY. In correction of the rationalist temper of idealist philosophy many voices were raised from time to time during the nineteenth century, stressing the mystery and elusiveness of religion. The most impressive and influential of these were those of Friedrich Schleiermacher and Rudolf Otto, the former giving prominence to the "feeling of absolute dependence" in religion and the latter stressing our sense of the holy or the numinous, the *mysterium tremendum et fascinans.* Otto claimed, in sharp contrast to the earlier naturalistic theories of Hume and his nineteenth-century followers, that there was ample evidence of this sense of the holy in the rawest beginnings of religion and he sought to describe the way it became schematized and moralized to give riper and more distinctive forms of religion. Other writers sought to correct the somewhat a priori approach of idealist philosophers by resorting to what they described rather incorrectly as an empiricist defense of religion that consisted in drawing out the implications of various features of our experience. This was the form that much natural theology took in the late nineteenth century, exemplified especially in the work of F. R. Tennant. Even if this approach fails to do justice to the factor of transcendence in religion, it could nonetheless be thought to have provided many of the ingredients of a sound understanding of religious experience.

Toward the close of the nineteenth century and early in the twentieth century there appeared, however, a strong reaction against what was thought to be the facile and too liberal rationalization of religious philosophy at that time. This found expression most of all in the insistence, by Karl Barth and other eminent theologians such as Reinhold Niebuhr and Emil Brunner, on the "wholly other" character of God and the need, as they understood it, to fall back on a dogmatically orthodox theological position in which the central place was accorded to the idea of an exclusive revelation. This presented considerable difficulties, not least on the ethical side where elementary ethical principles seemed to be put in serious jeopardy. But it did give prominence again to the idea of God's transcendence, which is a focus for controversies about religion among philosophers of the present day.

PHILOSOPHY OF RELIGION. The philosophy of our time has become extensively empiricist again. This trend had been preparing for some time in America in aspects of the work of William James and Charles Sanders Peirce. But it gathered its momentum in the work of the Vienna circle and those, such as Ludwig Wittgenstein, who extended its influence in England, notably at Cambridge and Oxford. Recent empiricism represents a sharp reaction against the ambitious and occasionally turgid speculations of nineteenth-century metaphysical philosophers. It set a premium on clarity and claimed to be tough-minded and down to earth. Its policy was extensively that of Hume, and it reflected much of the skepticism of the period subsequent to World War I. To Hume's empiricism was added, however, an alleged linguistic technique which was intended, in its main early forms at least, to account for the persistence of seemingly bold nonempirical notions, like the idea of the soul or of God, by ascribing them to confusions engendered by misleading forms of speech. This set off a spate of philosophical criticism of religion aimed at showing that its basic conceptions were logically improper. This is sometimes known as the linguistic veto. A desperate attempt to save religion was undertaken by several other empiricist philosophers who seemed willing to sacrifice the strictly nonempirical elements in religion and reinterpret the main features of religious belief in terms of present experience—for example, by regarding religion as a matter of satisfying certain distinctive emotions or by identifying it, in essentials, with ethics.

There have been considerable recent variations on this theme of the attenuation of religious faith. The same method of apologetics has appealed also to many theological writers, some equating religion with morality and others finding the essence of religion in a certain depth and earnestness of our own activities. The most outstanding of these theologians have relied heavily on the work of existentialist philosophers who have brought into prominence the importance of certain searching present experiences and of deep inner aspects of them. Neither they nor their existentialist mentors are very systematic or lucid thinkers, and it is thus not very clear how far they mean to go in interpreting religion in terms of our human experience in the here and now.

RELIGIOUS EXISTENTIALISM. A typically elusive representative of this kind of philosophical theology was Paul Tillich. It is never quite clear whether he meant by his central conceptions of "the Ground of Being" and the "New Being" a transcendent reality (or some impact of this upon us) or some profound depth of our own experience and natures. Nor is it clear how far this skepticism about traditional beliefs, reinforced by much skepticism in the field of biblical scholarship, is meant to go; for the writers in question often give expression to seemingly skeptical views in the language of orthodoxy. The position is not made easier by considerable borrowings from phenomenological thinkers like Martin Heidegger who combine unusual perceptiveness with a veritable genius for elaborate and obscure modes of utterance.

LINGUISTIC APOLOGETIC. Equally uncertain and difficult is the work of certain more strictly philosophical thinkers who take their start from a new emphasis in linguistic philosophy derived largely from the later and much modified form of Wittgenstein's work. They stress the open texture and varieties of language and, on this basis, press the claims of religious language to a status not impaired by its not complying with the conditions of ordinary language or scientific language. This leaves the door open for a cautious but less skeptical approach to religion. But the question remains how much is accomplished unless we indicate how the distinctive language of religion is to be understood and what criteria may be applied to it. There is a tendency for some linguistic apologists of religion to be content with stressing the alleged oddity of religious language and thereby also to conflate major notions, like freedom and immortality, and to leave it very unclear in what sense the various affirmations made in religion are to be understood. These writers also tend to draw much support from existentialist insistence on the importance of formative and challenging present experiences. The details of their work, as in the case of I. T. Ramsey, is illuminating and imaginative, but it is not clear how much it can accomplish until their kind of sensitivity to religious language is accompanied by rigorous heed to the centrality and discipline of the more strictly epistemological considerations.

RESPONSE TO EMPIRICIST CRITICISMS. Epistemological considerations have again been uppermost in the work of a further body of recent philosophers who have taken up the challenge of empiricist and linguistic critics more boldly. They have welcomed the challenge in particular as a way of sharpening the question of the place of evidence in religious belief. They maintain that evidence is not strictly relevant to the question of the existence of God; we apprehend the necessity of God's existence in the contingent character of everything else. This, they maintain, is the element of truth misleadingly presented in the traditional arguments. Pioneers of this position in recent philosophy are Austin Farrer and E. L. Mascall, while another severe critic of linguistic empiricism, C. A. Campbell, has arrived, by way of some modifications of Bradley's thought, at a not dissimilar renewal of the emphasis on the suprarational character of the object of religious worship.

This takes the sting out of the challenge, given sharpness by John Wisdom and later by Antony Flew, to indicate what would count for or against the existence of God. The answer, it is said, is "nothing," for we are not here accounting for the way the world goes or some particular feature of it, but for there being anything at all. The question "Why is there something rather than nothing?" is regarded even by some skeptical philosophers as a significant one. This new appreciation of the uniqueness of the idea of God and of God's relation to the world has opened the way also for subtler understanding of religions other than Christianity, especially Buddhism, and with this has come a renewed philosophical interest in world religions. This is a more discerning interest than the one motivated by superficial notions of syncretism at the turn of the century.

But there has been accentuated in turn the problem of particular religious affirmations. Some have attacked this afresh through new presentations of the traditional doctrine of analogy; some, like A. C. Ewing, persist in a cautious restatement of idealism; others turn to fresh examination of the nature and sanction of religious imagery. There has also been much recourse to the analogy with our knowledge of one another, and in this con-

text it has been thought, by the present writer among others, that a fresh examination of the nature of religious experience and of features of it that could afford justification of the claim to revelation in Scriptures and history, holds the best promise of a solution of the epistemological problems of religious faith. Some who follow this course are apt to lapse from a steady epistemological study, which their initial problem requires, into a psychological or phenomenological one; but when they do so, in the case of Gabriel Marcel for example, they may nonetheless provide highly relevant material for those who manage to keep the epistemological task steadily in mind. That may also be supplemented by the perceptive analysis of those whose concern is not mainly religious or who may be strictly atheistic like Jean-Paul Sartre. The work of Maurice Merleau-Ponty is thought by many to be especially suggestive and illuminating in this way.

Consideration of religious experience may thus prove the point of convergence of many of the approaches to religion which hold most promise today of deepening our understanding of its perennial problems. Advances in fields other than strictly religious studies, most of all perhaps the study of paranormal phenomena, will have much relevance to the present tasks of the philosophy of religion; and some writers, such as H. H. Price, C. D. Broad and C. J. Ducasse, have considered closely the implications of matters like paranormal phenomena for our general view of the world and for relevance to specific questions like immortality. Psychological studies, notably those that investigate the unconscious and the unconscious matrix of conscious imagery, have considerable relevance to the philosophers' problems. A further major preoccupation of those who study the philosophy of religion today is the relation of ethics to religion, not only in the form of fresh examination of the problems of freedom and grace or of variations on the traditional "moral argument," but also in reflections on the role of moral experience within the totality of religious experience. There have likewise been fresh examinations of the claims made for mystical experience, and one writer at least, namely W. T. Stace, is prepared to defend a very extreme form of monism as the ultimate truth about the universe to which mystical experience points. Other philosophers, including some such as J. N. Findlay who took their orientation at one time from Wittgensteinian philosophy, are beginning to embark on bold—too bold?—speculative ventures in the field of religious thought.

In these ways the philosophy of religion, of which fashionable philosophers fought very shy about twenty years ago, has become again one of the liveliest interests of philosophers. It is of considerable significance also that some of the major themes of contemporary fiction, including those that seem to have little overtly to do with religion, are found to bear closely on aspects of religion that have most importance for the philosophy of religion. In the blend of new philosophical investigations of religion, sharpened in the challenge and discipline of tough-minded philosophy, and a perceptive understanding of contemporary cultures (in their limitations as well as in their achievements) in other regards may be found a means of genuine advance in the life of religion itself which will enable it to have its place effectively in the sophistications of a developing culture and rapidly changing state of society.

See also Islamic Philosophy; Jewish Philosophy.

Bibliography

Allen, E. L. *Existentialism from Within.* London: Routledge and Paul, 1953.

Blackham, H. J. *Six Existentialist Thinkers.* London: Routledge and Paul, 1952.

Campbell, C. A. *Selfhood and Godhood.* London and New York, 1957.

Ewing, A. C. *Idealism.* London: Methuen, 1934.

Farrer, Austin. *Finite and Infinite.* 2nd ed. Naperville, IL, 1959.

Ferré, Frederick. *Language, Logic, and God.* New York: Harper, 1961.

Flew, A. G. N., and Alasdair MacIntyre, eds. *New Essays in Philosophical Theology.* London: SCM Press, 1955.

Hartshorne, Charles. *The Logic of Perfection.* La Salle, IL: Open Court, 1962.

Henson, H. H. *Christian Morality.* Oxford: Clarendon Press, 1936.

Hepburn, Ronald. *Christianity and Paradox.* London: Watts, 1958.

Hick, John. *Philosophy of Religion.* Englewood Cliffs, NJ: Prentice-Hall, 1963.

Hughes, E. R., and K. Hughes. *Religions in China.* London, 1950.

James, E. O. *The Worship of the Sky God.* London: Universtiy of London, Athlone Press, 1963.

Jayatilleke, K. N. *Early Buddhist Theory of Knowledge.* London, 1963.

Knowles, David. *Evolution of Medieval Thought.* London: Longmans, 1962.

Lewis, H. D. *Our Experience of God.* London: Allen and Unwin, 1959.

Lewis, H. D. "Survey of the History of Religion." *Philosophical Quarterly* (15) (April 1954) and (16) (July 1954).

Lewis, H. D. *Teach Yourself the Philosophy of Religion.* London, 1965.

MacGregor, Geddes, and J. W. Robb. *Readings in Religious Philosophy.* Boston: Houghton Mifflin, 1962.

Mascall, E. L. *Existence and Analogy.* London, 1949.

Mascall, E. L. *He Who Is.* London: Longmans, Green, 1943.

Mascall, E. L. *The Secularization of Christianity.* London: Darton, Longman and Todd, 1965.
Mascall, E. L. *Words and Images.* New York: Ronald Press, 1957.
Mitchell, Basil, ed. *Faith and Logic.* London: Allen and Unwin, 1957.
Moore, Edward C. *American Pragmatism.* New York: Columbia University Press, 1961.
Murti, T. R. V. *The Central Philosophy of Buddhism.* London: Allen and Unwin, 1955.
Paton, H. J. *The Modern Predicament.* London: Allen and Unwin, 1955.
Perry, R. B. *Present Philosophical Tendencies.* London: Longmans, Green, 1929.
Philosophy (122) (July 1957). On empiricism and religion.
Pringle-Pattison, Andrew Seth. *The Idea of God.* Oxford: Clarendon Press, 1917.
Ramsey, Ian. *Prospect for Metaphysics.* New York: Philosophical Library, 1961.
Ramsey, Ian. *Religious Language.* London: SCM Press, 1957.
Rogers, A. K. *A Student's History of Philosophy.* New York, 1940.
Rose, H. J. *Ancient Greek Religion.* London: Hutchinson's University Library, 1948.
Slater, R. H. L., and H. D. Lewis. *World Religions and World Community.* New York: Columbia University Press, 1963.
Smart, Ninian. *Historical Selections in the Philosophy of Religion.* New York: Harper and Row, 1962.
Smart, Ninian. *Reasons and Faiths.* London: Routledge and Paul, 1958.
Smith, John E. *Philosophy of Religion.* New York: Macmillan, 1965.
Taylor, A. E. "Theism," in *Encyclopedia of Religion and Ethics,* edited by James Hastings. New York, 1959.
Wolfson, H. A. *Religious Philosophy.* Cambridge, MA: Belknap Press of Harvard University, 1961.

H. D. Lewis (1967)

PHILOSOPHY OF RELIGION, HISTORY OF [ADDENDUM]

A remarkable revival of interest in philosophy of religion occurred during the final third of the twentieth century. The demise of logical positivism had freed philosophers of religion from their preoccupation with responding to its verificationist challenge to the meaningfulness of religious language. A narrow empiricism in epistemology ceased to play a foundational role in the philosophy of the Anglophone world. The philosophical community became more pluralistic in its views about the methods and assumptions that may fruitfully be brought to bear on the study of religion. At the same time, the philosophical community was growing in size, largely as a result of the rapid expansion of systems of higher education in countries such as the United States. As a result of these developments, philosophical reflection on religion came to be conducted by more philosophers with a greater variety of points of view than at any time in the past. Many philosophers approached religion from within well-established traditions of thought such as pragmatism, process philosophy, phenomenology, and Thomism. Remarks on religion in the later works of Ludwig Wittgenstein inspired discussions of religious forms of life. Some philosophers cast fresh light on classical topics in philosophical theology, whereas others focused on issues that contemporary culture has made salient. This entry provides a brief survey of some of the highlights of this flowering of philosophy of religion.

Natural theology, the enterprise of giving arguments for God's existence, was among the classical topics that attracted attention. Alvin Plantinga (2000) constructed a modal ontological argument that resembles earlier arguments discussed by Charles Hartshorne and Norman Malcolm, and William Rowe (1975) set forth a cosmological argument that resembles the version proposed by Samuel Clarke. Both of these arguments are clearly valid. However, as Plantinga and Rowe point out, each of them depends on a premise that one may rationally reject, and so neither is a successful proof of God's existence. Richard Swinburne (1979) produced a probabilistic cumulative case argument for God's existence. Making use of Bayesian reasoning, he tried to show that each one of several factors such as cosmic order, the existence of consciousness, and religious experience increases the probability of God's existence. According to Swinburne, the cumulative effect of all these factors is to render theism slightly more probable than not.

A controversial challenge to the view that belief in God is irrational or in some other way improper unless it is supported by arguments or other propositional evidence was mounted by Plantinga. Many theists do not, in fact, base their belief in God on such propositional evidence. Plantinga argued that such basic belief in God can be epistemically proper under certain conditions, typically conditions pertaining to how the belief is directly grounded in experience. In later work, he has gone on to contend that basic belief in God can have a good deal of warrant, which is the epistemic characteristic enough of which converts true belief into knowledge, despite its lack of support by arguments or other propositional evidence. Plantinga describes the position for which he has argued as Reformed epistemology because he finds it suggested in the writings of John Calvin. Influenced by Wittgenstein and Thomas Reid, William Alston (1991) argued for the practical

rationality of engaging in a nonsensory perceptual practice whose outputs are beliefs about how God is manifested in the experience of the practitioner. Both Plantinga and Alston espouse views in religious epistemology according to which belief in God can have positive epistemic status even in the absence of a successful natural theology.

Of course evil constitutes a potential defeater for the positive epistemic status that theistic belief can acquire from experience. According to the logical problem of evil, the existence of an omnipotent, omniscient, and perfectly good God is logically inconsistent with the existence of evil. In his celebrated free will defense, Plantinga (2000) argued convincingly that the existence of God is consistent with the existence of evil The main focus of debate subsequently shifted to the evidential problem of evil, according to which evils of a certain sort count as evidence that renders the existence of God improbable. An influential version of this problem formulated by Rowe (1975) started from the notion of pointless suffering, which is defined as suffering an omnipotent and omniscient being could have prevented without losing some greater good or permitting some evil equally bad or worse. Rowe maintained that instances of suffering known to us are apparently pointless and hence count as compelling evidence against the existence of God. In an attempt to rebut Rowe, Stephen Wykstra argued against concluding that such instances of suffering are apparently pointless. On his view, even if such suffering has a point because without it God cannot secure some greater good, it is very likely that its point is completely beyond our ken.

As a result of the increasing religious pluralism of modern societies, religious diversity has become a salient threat to the positive epistemic status of conflicting systems of religious belief. Major world religions disagree about even such fundamental issues as whether the ultimate religious reality is a personal deity or an impersonal absolute. John Hick proposed that our response to this situation should be to adopt the hypothesis that all the world religions are somehow in touch with a single noumenal reality to which no substantive human concepts apply. On this hypothesis, the ultimates of different world religions are equally real but merely phenomenal realities, all of which are in part products of human cultures and traditions. Opponents of Hick's proposal, such as Alston and Plantinga, have argued for the rationality of remaining within the belief systems of particular religions despite the negative impact religious diversity has on the epistemic status of such beliefs, at least for those who are sufficiently aware of the conflict to which this diversity gives rise.

Important work has also been done on several other topics. One example is the metaphysics of theism. Philosophical reflection on God's nature has produced new accounts of such divine attributes as omniscience, omnipotence, eternity, and simplicity. Another example is religious ethics. John Finnis has developed a natural law theory influenced by the thought of Thomas Aquinas; Alasdair MacIntyre has argued in favor of moral inquiry within the tradition of Aquinas; Robert Adams and Philip Quinn have formulated and defended divine command theories of moral obligation. John Caputo and Merold Westphal have played a significant role in drawing to the attention of Anglophone philosophers the religious implications of the writings of major French and German thinkers. And promising first steps have been taken toward feminist and comparative philosophies of religion.

See also Islamic Philosophy; Jewish Philosophy.

Bibliography

Adams, Marilyn M., and Robert M. Adams, eds. *The Problem of Evil*. Oxford and New York: Oxford University Press, 1990. Reprints important contributions to the discussion of logical and evidential problems of evil.

Alston, William P. *Perceiving God: The Epistemology of Religious Experience*. Ithaca and London: Cornell University Press, 1991. Defends the rationality of engagement in nonsensory perceptual religious practices of belief formation.

Long, Eugene, T. *Twentieth Century Western Philosophy of Religion: 1900–2000*. Dordrecht: Kluwer, 2000. Offers a comprehensive historical account of Western philosophy of religion in the twentieth century.

Plantinga, Alvin. *Warranted Christian Belief*. Oxford and New York: Oxford University Press, 2000. Develops and argues for the author's mature position on topics in religious epistemology.

Quinn, Philip L., and Kevin Meeker, eds. *The Philosophical Challenge of Religious Diversity*. Oxford and New York: Oxford University Press, 2000. Reprints major contributions to the debate about the epistemological consequences of religious diversity.

Quinn, Philip L., and Charles Taliaferro, eds. *A Companion to Philosophy of Religion*. Oxford: Blackwell, 1997. Provides a comprehensive survey of philosophy of religion at the end of the twentieth century in seventy-eight original articles by leading contributors to the field.

Rowe, William L. *The Cosmological Argument*. Princeton, NJ: Princeton University Press, 1975. Contains a thorough analysis of Samuel Clarke's version of the cosmological argument for the existence of God.

Swinburne, Richard. *The Existence of God*. Oxford: Clarendon Press, 1979. Sets forth in detail the author's Bayesian cumulative case argument for the existence of God.

Philip L. Quinn (2005)

PHILOSOPHY OF RELIGION, PROBLEMS OF

The term *philosophy of religion* is a relative newcomer to the philosophical lexicon, but what is now so designated is as old as philosophy itself. One of the earliest spurs to philosophical reflection, in ancient Greece and elsewhere, was the emergence of doubts concerning the religious tradition; and religious beliefs and conceptions have always formed much of the staple of philosophical discussion.

If one surveys the various things philosophers have done in thinking about religion, it is difficult to find any unifying thread other than the fact that they all spring from reflection on religion. Philosophy of religion is occupied to a large extent with the consideration of reasons for and against various fundamental religious beliefs, particularly the various arguments for the existence of God. But we find many other matters treated in books that are regarded as being within the philosophy of religion. These include the nature and significance of religious experience, the nature of religion, the relation between religion and science, the nature of religious faith as a mode of belief and/or awareness, the nature of revelation and its relation to the results of human experience and reflection, the place of religion in human culture as a whole, the logical analysis of religious language, the nature and significance of religious symbolism, and possibilities for reconstructing religion along relatively nontraditional lines.

CENTRAL AIM

Some justification can be found for grouping all these topics under the heading "philosophy of religion" if we view them all as growing out of a single enterprise, the rational scrutiny of the claims of religion—the critical examination of these claims in the light of whatever considerations are relevant—with a view to making a reasonable response to them. A highly developed religion presents us with a number of important claims on our belief, our conduct, our attitudes and feelings. It gives answers to questions concerning the ultimate source of things, the governing forces in the cosmos, the ultimate purpose(s) of the universe, and the place of man in this scheme. It tells us what a supreme being is like, what demands he makes on men, and how one can get in touch with him. It offers a diagnosis of human ills, and it lays down a "way of salvation" that, if followed, will provide a way to remedy these ills and satisfy man's deepest needs. All this is very important. If the claims of a given religion on these points are justified, discovering this is a matter of the greatest moment. At bottom the philosophy of religion is the enterprise of subjecting such claims to rational criticism.

It is worth noting that such claims are not made by religion in general but by particular religions exclusively and that although generally we can find claims of all these sorts in any given religion, the specific content will differ widely from one religion to another. This will have important consequences for the direction taken by the philosophizing that arises in response to each religion. This article is largely concerned with the Western tradition, and thus the philosophy of religion represented has grown out of concern with some aspect of the Judeo-Christian tradition, either through support or opposition. Philosophical reflection on a very different religious tradition will give rise to different preoccupations. Thus, Western philosophers, unlike their Indian counterparts, are much concerned with arguments for and against the existence of a supreme personal deity and with whether or not the occurrence of miracles is compatible with the reign of natural law. However, in a religious tradition like the Hindu or the Buddhist, which does not feature the notion of a supreme personal deity who has active personal dealings with his creatures, these problems do not arise. Philosophers in such a tradition, by contrast, will be concerned with trying to clarify the relation of a supreme ineffable One to the various things in the world that constitute its manifestations and with considering arguments for the ultimate unreality of the empirical world. There is, however, enough in common among different religions to ensure that all philosophy of religion will be directed to recognizably identical problems, though in very different forms.

Philosophers have raised critical questions about the justifiability and value of religious beliefs, rites, moral attitudes, and modes of experience. However, philosophers have largely focused their critical powers on the doctrinal (belief) side of religion. This selectivity might be attributed to an occupational bias for the intellectual, but there is a real justification for it. If our basic interest is in questions of justifiability, then it is natural that we should concentrate on the belief side of religion, for the justification of any other element ultimately rests on the justification of some belief or beliefs. If one asks a Roman Catholic why he goes to Mass, or what the value is of so doing, he would, if he knew what he was about, appeal to certain basic beliefs of his religion: that the universe, and all its constituents, owes its existence to and depends for its ultimate fate on a supreme personal being, God; that

man inevitably fails to live up to the moral requirements God lays down for him; that God became a man in the person of Jesus of Nazareth and suffered death in order to save man from the fatal consequences of his sinfulness; that as a part of a program designed to enable men to benefit from this, God has ordained that they should participate in the rite of the Mass, in which, in some mysterious way, they actually incorporate the body and blood of Jesus and so partake of the salvation effected through him. The ritual, as conceived by the participants, is a reasonable thing to do if and only if these beliefs are justified.

However, the attention of philosophers is generally more narrowly concentrated than this. Not all the beliefs of a given religion, not even all the beliefs considered crucial by that religion, receive equal attention. In works on the philosophy of religion, one finds little discussion of relatively special doctrines that are peculiar to a given religion, such as the virgin birth of Jesus, the divine mission of the church, or the special status of the priesthood, however important these doctrines may be for the religion in question. Instead, attention is focused primarily on what might be called the metaphysical background of the doctrinal system, the worldview of the religion—the view of the ultimate source and nature of the universe; the nature of man; man's place in the universe; the end to which man is, or should be, tending; and so on. This preferential treatment is partly due to a desire to make philosophical discussions relevant to more than one religion; for example, roughly the same worldview underlies Judaism, Christianity, and Islam. It is also partly due to a conviction that philosophical reflection will yield definite results only with respect to the more general aspects of a religious outlook. Very few philosophers have supposed that one can establish the virgin birth by philosophical argument.

It might also be argued that if we abstract from commitment to any particular religion, the worldview aspect of religion is the most undeniably significant one. Without presupposing some particular religious beliefs, it would be difficult to show that the acceptance of elaborate theological dogmas like that of the Trinity, or participation in rites, or singling out certain objects as sacred is an essential part of a fully human life. However, it can be argued on the basis of facts concerning the nature of man and the conditions of human life that human beings have a deep-seated need to form some general picture of the total universe in which they live, in order to be able to relate their own fragmentary activities to the universe as a whole in a way meaningful to them; and that a life in which this is not carried through is a life impoverished in a most significant respect. This would seem to be an aspect of religion that is important on any religious position; and so it seems fitting that it should be at the center of the picture in a general philosophical treatment of religion.

OTHER INVESTIGATIONS AND THE CENTRAL AIM

In presenting, defending, and criticizing arguments for and against such fundamental beliefs as the existence of a supreme personal deity, the immortality of the human personality, and the direction of the universe toward the realization of a certain purpose, philosophers are directly engaged in critical evaluation. The other major topics listed at the beginning of this article do not have exactly this status, but they are all directly relevant to rational criticism of fundamental religious beliefs. In order to conduct a systematic scrutiny of such beliefs, one must start with an adequate conception of the nature and range of religion, so that he can be sure that he is dealing with genuine religious beliefs and with those which are most fundamental for religion, and so that he will not be unduly limited by the particular interests with which he starts.

Moreover, one needs an adequate understanding of the nature of religious belief in order to filter out irrelevant considerations and arguments. The charge of irrelevancy has been most trenchantly leveled against the traditional enterprise of presenting metaphysical arguments for the existence of God by Søren Kierkegaard, who maintained that anyone who tries to give an argument for the existence of God thereby shows that he has misunderstood the special character of religious belief. Whether or not such charges are justified, the mere fact that they can be made with any plausibility shows that it is incumbent on the philosopher of religion to look into the character of religious faith and to try to determine its similarities to and differences from other modes of belief; for example, those in everyday life and in science. With an increasing realization of the way in which thought and belief are shaped by language, this kind of investigation has increasingly taken the form of an inquiry into the type of utterances that express religious belief, an attempt to make explicit the logic of religious discourse—the special ways in which terms are used in religious utterances, the logical relations between religious statements themselves and between religious statements and statements in other areas of discourse, the extent to which religious statements are to be construed as expressive of feelings or

attitudes or as directions to action, rather than as factual claims. Also, an appreciation of the extent to which language is used symbolically in religion can easily lead to a general concern with the nature and function of religious symbolism.

All the concerns listed thus far involve investigation of the relation of religion to other segments of human culture, such as science, art, and literature. The question of the relation of science and religion has a special importance for one who is critically examining religious beliefs in our society. For the last few hundred years the main challenges to religious doctrine in Western society have been made in the name of science. With respect to many segments of science, from Copernican astronomy through Darwinian biology to Freudian psychology, it has been claimed that certain scientific discoveries disprove, or at least seriously weaken, certain basic religious doctrines. Discussions of whether this ever does, or can, happen—and if so, what is to be done about it—have bulked large in works on philosophy of religion.

Philosophers of religion also investigate the nature of religious experiences because it is often claimed that such experiences provide direct warrant for the existence of God, or of other objects of religious worship. One is naturally led into a survey of the types of religious experience and into questions of their psychological bases. Finally, if a philosopher has decided that the basic beliefs of the traditional religion(s) of his society are unacceptable, he is naturally faced with the question of what to do about it. If he feels that religion is a crucially important aspect of human life, he will want to find some way of preserving religious functions in a new form. Hence, naturalistic philosophers, who reject the supernaturalistic beliefs of our religious tradition, sometimes attempt to sketch the outlines of a religion constructed on naturalistic lines. This will usually involve the substitution of some component(s) or aspect(s) of the natural world for the supernatural deity of the Judeo-Christian tradition. This may be Humanity (Auguste Comte), human ideals (John Dewey), those natural processes which make a contribution to the realization of the greatest good (H. N. Wieman), or some combination of these.

RELATIONS TO OTHER DISCIPLINES

The philosophy of religion is distinguished from theology and from sciences dealing with religion (such as psychology of religion and sociology of religion) in opposite ways. It is distinguished from theology by the fact that it takes nothing for granted, at least nothing religious; in the course of its examination it takes the liberty of calling anything into question. Theology, in a narrow sense of that term, sets out to articulate the beliefs of a given religion and to put them into systematic order, without ever raising the ultimate question of their truth. The philosophy of religion is distinguished from sciences of religion by the fact that it is addressed to questions of value and justification and tries to arrive at some sort of judgment on religious claims. The psychology of religion—for instance, when pursuing strictly psychological questions—studies religious beliefs, attitudes, and experiences as so many facts, which it tries to describe and explain, without attempting to pass judgment on their objective truth, rationality, or importance.

The philosophy of religion, conceived of as an attempt to carry out a rational scrutiny of the claims made by a given religion, will always start from concern with some particular religion or type of religion and will basically aim at a judgment of that religion. It certainly is historically accurate to think of philosophy of religion as arising in this way and, furthermore, it may be taken as its common and most basic form. However, it is also possible for a philosopher to concern himself directly with the fundamental issues involved in the religious claims in question—the ultimate source of things, the destiny of man, and cosmic purpose, for example—without approaching them through the consideration of answers given to these questions by some organized religion. Benedict de Spinoza's *Ethics* is an outstanding example of this kind of investigation. Other examples are Samuel Alexander's *Space, Time and Deity* (2 vols., London, 1920) and Henri Bergson's *L'évolution créatrice* (*Creative Evolution,* New York, 1911). Whether we call philosophizing of this kind philosophy of religion is not important, but it is important to realize that these questions can be considered outside the context in which we are explicitly concerned with religion as such.

VARIOUS APPROACHES

One should not suppose that every philosopher of religion concerns himself with the whole range of problems. On the contrary, a given philosopher will usually restrict his attention because of his special interests, his conception of religion, and/or his general philosophical position. The second and third of these factors deserve further notice. Concerning the second, the types of problems that a given philosopher emphasizes will sometimes be influenced by the particular aspect of religion he regards as essential. Thus, the concentration on problems connected with religious belief in traditional philosophy of religion is partly due to the fact that most philosophers

of religion have thought of religion primarily as a kind of belief (although this may, in fact, be less important than other factors). W. T. Stace in *Time and Eternity,* for example, considers mystical experience to be the essence of religion. Stace concentrated his main efforts on interpreting and justifying religious doctrine conceived as basically an expression of mystical experience. On the other hand, Kierkegaard thought of religion as basically a matter of an individual maintaining a certain general stance in life, and he devoted himself to an elaborate description of a variety of such stances, combined with indirect recommendations of one of these; he rarely mentioned any of the problems customarily discussed by philosophers of religion.

The operation of the third factor, the individual's philosophical position, is more apparent and, perhaps, more powerful. A few examples, selected more or less at random, will be helpful. Philosophers who are primarily speculative metaphysicians—Plato, Thomas Aquinas, Gottfried Wilhelm Leibniz, G. W. F. Hegel, and A. N. Whitehead—naturally take very seriously the enterprise of constructing metaphysical arguments for or against the existence of God, whereas predominantly antimetaphysical philosophers—David Hume, Immanuel Kant, and Dewey—will either criticize such arguments or, as is more common in recent times, ignore them altogether. Those who subscribe to the thesis that the only proper job of philosophy is the analysis (clarification) of concepts will observe the appropriate restrictions when and if they turn their attention to religion. There is a great deal of work of this kind to be done with the concepts of God, creation, revelation, faith, and miracle, to name a few. Traditionally this has been done in connection with attempts to reach substantive conclusions on the existence of God, immortality, and other major issues, but if one thinks that conclusions on such matters cannot be attained by philosophical reflection, as analytic philosophers do, he may still seek to make explicit the concepts involved in religious belief. Such philosophizing will regard itself as a humble servant of theology or of more ordinary religious belief and will pretend to no judicial functions, except where it locates internal confusions or inconsistencies.

The influence of philosophical orientation is clearly exemplified in naturalistic philosophers, who generally rule out all supernaturalism on the basis of their general philosophical position, without giving particular supernaturalistic beliefs any detailed examination. Naturalists devote their energies to revising religious belief and practice so that they will be acceptable within a naturalistic framework.

Finally, one may consider Hegel, who devoted his lectures on the philosophy of religion to demonstrating a dialectical progression in the history of religion. This reflected Hegel's basic philosophical conviction that reality consists of the process of the Absolute coming to full self-consciousness, that this process exhibits a dialectical pattern, and that it is manifested in the history of every cultural form.

In the task of classifying the positions that have been taken in the philosophy of religion, one confronts the difficulty that not all philosophers of religion, even in a single religious tradition, are dealing with the same problems. However, there is a common task underlying all the different approaches. All philosophy of religion is ultimately concerned with arriving at a rational judgment of the religion under discussion and, if the judgment is negative, to present some sort of alternative. The initial principle of division can then be taken as the affirmative or negative character of this judgment. (This cannot be absolutely clear-cut, partly because often some part of the religion is affirmed and some is rejected, partly because it is not absolutely clear what is to be included in the religion in question.) It can then be asked of those whose judgment is affirmative what the basis of their judgment is.

One major group, which includes the great majority of philosophers of religion, presents various arguments in support of such beliefs as the existence of God and the immortality of the soul, arguments that take their start from premises that are not themselves religious doctrines and that, it is assumed, any reasonable man would accept. In other words, they attempt to support religious belief by resting it on nonreligious premises. A smaller but still considerable group regards religious belief as not needing any such support from the outside; they regard it as somehow self-justifying or at least as justified by something from within religion. Some of them (Bergson and James) suppose that the belief in the existence of God, for example, is justified by religious experience. One can directly experience the presence of God, and therefore one does not need to prove his existence by showing that he must be postulated to explain certain facts. Others regard religious faith as different from other modes of belief in such a way that it does not need support of any kind, either from argument from effect to cause or from direct experience. Kierkegaard, Emil Brunner, and Paul Tillich, for example, all take this position, though there are great differences between them. (The case of Tillich

illustrates the point that in some cases it is difficult to distinguish between those who accept the religious tradition and those who reject it. Tillich considered himself a Christian theologian, but his interpretation of Christian doctrine is so unorthodox that many feel he reconstrued it out of recognition and therefore should be classed with those who substitute a symbolic reinterpretation for traditional beliefs.)

In the other major group we can distinguish between those who simply reject traditional religion (Baron d'Holbach and Bertrand Russell) and those who in addition try to put something in its place. In the latter group we can distinguish between those who try to retain the trappings, perhaps even the doctrinal trappings, of traditional religion but give it a nonsupernaturalistic reinterpretation, usually as symbolic of something or other in the natural world (George Santayana), and those who attempt to depict a quite different sort of religion constructed along nonsupernaturalistic lines (Comte, Dewey, and Wieman).

Outside this classification are those analytical philosophers who restrict themselves to the analysis of concepts and types of utterances. We may regard them as not having a major position in the philosophy of religion, but rather as making contributions that may be useful in the construction of such a position.

See also Religion.

Bibliography

Edwin A. Burtt, *Types of Religious Philosophy,* rev. ed. (New York: Harper, 1951), and Robert Leet Patterson, *An Introduction to the Philosophy of Religion* (New York: Holt, 1958), are useful introductory textbooks. A wide variety of readings in the field can be found in *Philosophy of Religion,* edited by George L. Abernethy and Thomas A. Langford (New York: Macmillan, 1962), and in *Religious Belief and Philosophical Reflection,* edited by William P. Alston (New York, 1963).

The following works are important treatments of a wide variety of topics in this area: John Baillie, *The Interpretation of Religion* (New York: Scribners, 1928); H. J. Paton, *The Modern Predicament* (London: Allen and Unwin, 1955); A. E. Taylor, *The Faith of a Moralist* (New York: Macmillan, 1930); and F. R. Tennant, *Philosophical Theology* (Cambridge, U.K.: Cambridge University Press, 1928). These works are written from a standpoint more or less sympathetic to traditional theism. For fairly comprehensive discussions from a more critical standpoint, see J. M. E. McTaggart, *Some Dogmas of Religion* (London: Arnold, 1906), and Bertrand Russell, *Religion and Science* (New York: Holt, 1935).

Works dealing with the nature and significance of religious experience include William James, *The Varieties of Religious Experience* (New York: Longman, 1902), and Rudolf Otto, *The Idea of the Holy,* translated by J. W. Harvey (New York: Oxford University Press, 1958). The nature of religion is discussed in Josiah Royce, *The Sources of Religious Insight* (New York: Scribners, 1912), and in Julian Huxley, *Religion without Revelation* (New York: Harper, 1957). For the relation of religion and science, see Bertrand Russell, op. cit., and Michael Pupin, ed., *Science and Religion; a Symposium* (New York, 1931). J. H. Newman, *A Grammar of Assent* (London: Burns, Oates, 1870), and Paul Tillich, *Dynamics of Faith* (New York: Harper, 1957), cover the nature of religious faith as a mode of belief and/or awareness.

In Emil Brunner, *The Philosophy of Religion from the Standpoint of Protestant Theology,* translated by A. J. D. Farrer and B. L. Woolf (New York: Scribners, 1937), the nature of revelation and its relation to the results of human experience and reflection are considered. The place of religion in human culture as a whole is dealt with in G. W. F. Hegel, *Lectures on the Philosophy of Religion,* translated by E. B. Speirs and J. B. Sanderson, 3 vols. (London: K. Paul, Trench, Trubner, 1895), and in George Santayana, *Reason in Religion* (New York: Scribners, 1905). For the logical analysis of religious language, see A. G. N. Flew and Alasdair MacIntyre, eds., *New Essays in Philosophical Theology* (London: SCM Press, 1955), and C. B. Martin, *Religious Belief* (Ithaca, NY: Cornell University Press, 1959). Edwyn Bevan, *Symbolism and Belief* (Boston: Beacon Press, 1957), and W. T. Stace, *Time and Eternity* (Princeton, NJ: Princeton University Press, 1952), discuss the nature and significance of religious symbolism. Possibilities for reconstructing religion along relatively nontraditional lines appear in Immanuel Kant, *Religion within the Limits of Reason Alone,* translated by T. M. Greene and H. H. Hudson (La Salle, IL: Open Court, 1934); John Dewey, *A Common Faith* (New Haven, CT: Yale University Press, 1934); and Julian Huxley, op. cit.

William P. Alston (1967)

PHILOSOPHY OF SCIENCE, HISTORY OF

Philosophy of science emerged as a distinctive part of philosophy in the twentieth century. Its defining moment was the meeting (and clash) of two courses of events: the breakdown of the Kantian philosophical tradition and the crisis in the sciences and mathematics in the beginning of the century. But what we now call philosophy of science has a rich intellectual history that goes back to the ancient Greeks. It is intimately connected with the efforts made by many thinkers to come to terms with the distinctive kind of knowledge (*epistēmē*, *scientia*) that science offers. Though science proper was distinguished from natural philosophy only in the nineteenth century, the *philosophy* of natural philosophy had almost the very same agenda that current philosophy of science has.

ARISTOTLE

Aristotle (384–322 BCE) thought that there was a sharp distinction between our understanding of facts and our understanding of the reasons for those facts. Though both types of understanding proceed via deductive syllogism, only the latter is characteristic of science, because only the latter is tied to the knowledge of causes. In *Posterior Analytics*, Aristotle illustrates this difference by contrasting the following two instances of deductive syllogism:

> **Syllogism A**
> Planets do not twinkle.
> What does not twinkle is near.
> Therefore, planets are near.
>
> **Syllogism B**
> Planets are near.
> What is near does not twinkle.
> Therefore, planets do not twinkle.

Syllogism A, Aristotle said, demonstrates the fact that planets are near, but does *not* explain this fact, because the syllogism does not state its causes. However, syllogism B is explanatory because the syllogism gives the *reason why* planets do not twinkle: *because* they are near. Aristotle's point was that, besides being demonstrative, explanatory arguments should also be asymmetric: The asymmetric relation between causes and effects should be reflected in an asymmetric relation between the premises and the conclusion of the explanatory arguments: The premises should explain the conclusion, and not the other way around.

For Aristotle, scientific knowledge forms a tight deductive-axiomatic system whose axioms are *first principles*, which are "true and primary and immediate, and more known than and prior to and causes of the conclusion" (71b19–25). Being an empiricist, he thought that knowledge of causes has experience as its source. But experience on its own cannot lead, through induction, to universal and necessary first principles that state ultimate causes. Nor can first principles be demonstrated, on pain of either circularity or infinite regress. So something besides experience and demonstration is necessary for knowledge of first principles. This is the process of abstraction based on intuition, a process that reveals the essences of things, that is, the properties by virtue of which a thing is what it is. Though Aristotle called first principles "definitions," they are not verbal, but rather state the essences of things. In Aristotle's rich ontology, causes are essential properties of their effects and necessarily give rise to their effects. He thought that the logical necessity by which the conclusion follows from the premises of an explanatory argument mirrors the physical necessity by which causes produce their effects.

ARISTOTELIANISM

By the 1250s, Aristotle's works had been translated into Latin, either from the original Greek or through Arabic translations, and a whole tradition of writing commentaries on these works flourished. Aristotle's *Organon* was the main source on issues related to logic and knowledge. At about the same time, the first universities were founded in Paris and Oxford, and natural philosophy found in them its chief institutional home. Aristotelianism was the dominant philosophy throughout the Middle Ages, though it was enriched by insights deriving from religious beliefs and many philosophical commentaries. The new Aristotelianism put secular learning on almost equal footing with revealed truth, especially at the University of Paris.

Thomas Aquinas (c. 1225–1274) argued that science and faith cannot have the same object, since the object of science is something seen, whereas the object of faith is the unseen. He found in Aristotle's views the mean between two extremes, one being Plato's view, which demeaned experience and saw in it just an occasion in the process of understanding the realm of pure and immutable forms, the other being the Democretian atomist view, which reduced all knowledge to experience. Aristotelianism, Aquinas thought, was the golden mean. Experience is necessary for knowledge, since nothing can be in the mind if it is not first in the senses. But thought is active in that it extends beyond the bounds of sense and states the necessary, universal, and certain principles on which knowledge is based.

Aquinas inherited (and suitably modified) much of Aristotle's rich metaphysics. Aristotle, drawing a distinction between matter and form, argued that when a change takes place, the matter perdures (persists), while the form changes. He conceived of change as the successive presence of different (even opposing) forms in the substratum. Scholastic philosophers differentiated this substratum from the ordinary matter of experience and called it "prime matter" (*materia prima*). The form that gives prime matter its particular identity (making it a substance of a particular kind) they called "substantial form." Substantial forms were individuating principles that accounted for the specific properties of bodies (which all shared the same prime matter). Aquinas added that prime matter is pure potentiality, incapable of existing by itself. He adopted the view that change (as well as

motion) was the passage from potentiality to actuality. Since a thing cannot be both actual and potential at the same time, he took it to be obvious that nothing can be the active source of its own motion, and hence that motion always requires a mover. Aquinas found solace in the Aristotelian doctrine of the first unmoved mover (the source of all motion), which immediately lent itself to being identified with God.

THE PROBLEM OF MOTION

The status of motion was heavily debated among the Scholastics. One central Aristotelian axiom was that everything that moves requires a mover. Another central axiom was that the mover is in contact with the thing moved. This might be borne out in ordinary experience, but some cases created problems. One of them was projectile motion, and another concerned natural motion, that is, motion toward the natural place of a thing. In both cases, it is not obvious that something does the moving, let alone by being in contact with the thing moved. There was no easy way out of these problems. Underlying them was the very issue of what motion is. Is motion merely the final form momentarily attained by the moving object at any instant? Or is it something in addition, a flux or transformation of forms (in medieval terminology, *forma fluens* or *fluxus formae*)?

The radical answer to this question was sharpened by William of Ockham (c. 1280–1349), who argued that motion is nothing over and above the moving body and its successive and continuous termini. He was a nominalist who thought that only particulars exist. He denied that universals exist and claimed that general terms, or predicates, refer to concepts that apply to many particulars. He argued that the key to the problem of motion was thus held by the abstract noun "motion." It is wrong, he claimed, to think that this and other abstract nouns refer to distinct and separately existing things. Only individual bodies, places, and forms are needed to explain what motion is. Another view came from Jean Buridan (c. 1295–1358). He argued that local motion involves *impetus*, a motive force transmitted from the mover to the moving body, which acts as an internal cause of its continued motion.

ARGUMENT ACCORDING TO IMAGINATION

On March 7, 1277, Etienne Tempier, Bishop of Paris, issued an act condemning 219 propositions drawn from the works of Aristotle and his commentators (including Aquinas). These propositions were supposed to be in conflict with Christian faith and in particular with the omnipotence of God. They included such claims as that the world is eternal, that God could not make several worlds, that God could not make an accident exist without a subject, that God could not move the entire cosmos in straight line. Ironically, this act opened up new conceptual possibilities that were hitherto regarded as closed. If Aristotle could err in matters theological, could he not err in matters philosophical too?

On the premise that only the law of noncontradiction constrains God's actions, it was argued that anything that can be conceived without contradiction is possible. This led to a new type of argumentation: arguing according to the imagination (*secundum imaginationem*). If something could be consistently imagined, then it was possible. New ideas were pursued on this basis, unconstrained by claims concerning the actual course of nature (*secundum cursus naturae*). Central elements of Aristotelian doctrine were given close logical scrutiny. For instance, in the Aristotelian scheme of things, where there is no void and the entire cosmos occupies no place, it made no sense to say that the entire cosmos could move. But what if, Buridan asked, God made the whole cosmos rotate as one solid body? Freed to inquire into the logical possibility of this rotation, Buridan argued that since we can imagine it, there must be something more to motion than the moving body, its forms, and the places it acquires. For if these were all there were to motion, then, contrary to our assumption, the entire cosmos could not move, simply because there would be no places successively acquired.

Ockham pushed argument according to imagination to its limits by arguing that there is no a priori necessity in nature's workings. God could have made things other than they are. Hence, all existing things are contingent. There are no necessary connections between distinct existences, and there is justification for inferring one distinct existence from another, Ockham forcefully argued. Accordingly, all knowledge of things comes from experience. Ockham claimed that there could never be certain causal knowledge based on experience, since God might intervene to produce the effect directly, thereby dispensing with the secondary (material) cause. Ockham thus gave a radical twist to empiricism, putting it in direct conflict with the dominant Aristotelian view.

FIRST PRINCIPLES

The status of scientific knowledge was heavily debated in the thirteen and fourteenth centuries. John Duns Scotus (c. 1265–1308) defended the view that first principles are

knowable with certainty, as they are based only on the natural power of the understanding to see that they are self-evident, ultimately by virtue of the meanings of the terms involved in them. For him, the understanding is not caused by the senses, but only occasioned by them. Once it has received its material from the senses, the understanding exercises its own power in conceiving first principles. Interestingly enough, Scotus thought that there could be certain causal knowledge coming from experience. He asserted as self-evident a principle of induction. He held that this principle is known a priori by the intellect, since a free cause (that is, an act of a free agent) leads by its *form* to the effect that it is ordained to produce. It was then an easy step for him to extend this principle from free causes to natural causes: "Whatever happens frequently through something that is not free, has this something as its natural *per se* cause."

Ockham disagreed with Scotus's account of the first principles, but his central disagreement with his predecessors was about the *content* of first principles. Since he thought there was nothing in the world that corresponded to general concepts (such as universals), he claimed that first principles are, in the first instance, about mental contents. They are about concrete individuals only *indirectly* and insofar as the general terms and concepts can be predicated of concrete things. Ockham is famous for the principle known as Ockham's razor: Entities must not be multiplied without necessity. In fact, this principle of parsimony was well-known in his time. Robert Grosseteste (c. 1168–1253) had put it forward as the law of parsimony (*lex parsimoniae*).

Ockham's most radical follower, Nicolas of Autrecourt (c. 1300–after 1350), rejected the demand for certainty altogether and claimed that only probable knowledge is possible. He endorsed atomism, claiming that it is at least as probable as its rival, Aristotelianism. In reaction, the fourteenth-century Parisian masters—Buridan, Albert of Saxony (c. 1316–1390), and others—claimed that empirical knowledge can be practically certain and wholly adequate for natural science. For Buridan, if we fail to discover an instance of *A* that is not *B*, then it is warranted to claim that all *A*s are *B*. On the basis of this principle, he defended on empirical grounds the Aristotelian claim that there is no vacuum in nature, since, he said, we always experience material bodies.

THE PREROGATIVES OF EXPERIMENTAL SCIENCE

Despite their engagement with philosophical issues in natural science, thinkers such as Ockham and Scotus were little concerned with natural science itself. They saw little role for mathematics, the science of quantity, in physics. They neglected experiment altogether. This was a drawback of their thought in relation to some earlier medieval thinkers. Grosseteste was one of the first to emphasize the role of mathematics in natural science. Roger Bacon (1214–1292) went further by arguing that all sciences rest ultimately on mathematics, that facts should be subsumed under mathematical principles, and that empirical knowledge requires active experimentation. Bacon put forward three virtues of experimental science. First, it criticizes by experiment the conclusions of all the other sciences. Second, it can discover new truths (not of the same kind as already known truths) in the fields of science. Third, it investigates the secrets of nature and delivers knowledge of future and present events.

The emphasis on the mathematical representation of nature exerted important influence on the work of the masters of Merton College in Oxford, who, in the fourteenth century, by and large put aside the philosophical issues of the nature of motion and focused instead on its mathematical representation. Walter Burley (c. 1275–c. 1345), Thomas Bradwardine (c. 1295–1349), William of Heytesbury (before 1313–1372/1373), Richard Swineshead (d. c. 1355), known as the Mertonians, most of whom where nominalists, engaged in a project to investigate motion and its relation to velocity and resistance in an abstract mathematical way. Similar research, though more concerned with the physical nature of motion, was undertaken in Paris by Buridan, Albert of Saxony, and Nicole Oresme (c. 1320–1382), known as the Paris terminists. The mathematical ingenuity of the Mertonians and the Parisians led to many important mathematical results that spread throughout Western Europe and germinated in the thought of many modern thinkers, including Galileo Galilei (1564–1642). By the end of the fourteenth century, a protopositivist movement, concerned not with the ontology of motion, but with its measurement, started to spread.

THE COPERNICAN TURN

In *De revolutionibus orbium coelestium* (On the revolutions of the celestial spheres), Nicolaus Copernicus (1473–1543) developed his famous heliocentric model of the universe. The unsigned preface of the book, which was published posthumously in 1543, firmly placed it within the saving-of-appearances astronomical tradition favored by Plato and endorsed by many medieval thinkers. As it turned out, the preface was written not by Copernicus himself but by Andreas Osiander, a Lutheran

theologian. Copernicus emphatically refused to subscribe to this tradition. He had a *realist* conception of his theory, according to which, as Pierre Duhem put it, "a fully satisfactory astronomy can only be constructed on the basis of hypotheses that are true, that conform to the nature of things" (1908, p. 62).

Before Copernicus, the dominant astronomical theory was that of Claudius Ptolemy (c. 85–c. 165). Pretty much like Aristotle and Plato, Ptolemy had assumed a geocentric model of the universe. To save the appearances of planetary motion, he devised a system of deferents (large circles centered on the earth) and epicycles. There were alternative mathematical models of the motion of the planets (e.g., one based on a moving eccentric circle), but Ptolemy thought that since all these models saved the appearances, they were good enough. The issue of their physical reality was not raised (though at least some medieval philosophers understood these models realistically). Geometry was then the key to studying the celestial motions, but there was no pretense that the world itself was geometrical (though Plato, in the *Timaeus*, did advocate a kind of geometrical atomism). The Copernican heliocentric model, though it made the earth move around the sun, continued to use epicycles. But Copernicus argued that his theory was true. He based this thought mostly on considerations of harmony and simplicity: His own theory placed astronomical facts into a simpler and more harmonious mathematical system.

THE BOOK OF NATURE

Galileo Galilei (1564–1642) famously argued that the book of nature is written in the language of mathematics. He distinguished between logic and mathematics. Logic teaches us how to derive conclusions from premises, but does not tell us whether the premises are true. Mathematics is in the business of demonstrating truth. Though Galileo emphasized the role of experiment in science, he also drew a distinction between appearances and reality, which set the stage for his own, and subsequent, explanatory theories of phenomena, which posited unobservable entities. He accepted and defended the Copernican system and further supported it with his own telescopic observations, which spoke against the dominant Aristotelian view that the heavens are immutable. But the possible truth of Copernicus's theory suggested that the world might not be as it is revealed to us by the senses. Indeed, Galileo understood that the senses can be deceptive, and hence that proper science must go beyond merely relying on the senses. The mathematical theories of motion that he advanced were based on idealizations and abstractions. Experience provides the raw material for these idealizations (frictionless inclined planes, ideal pendula), but the key method of science was extracting, via abstraction and idealization, the basic structure of a phenomenon so that it could be translated into mathematical form. Then mathematical demonstration takes over and further consequences are deduced, which are tested empirically. So Galileo saw that understanding nature requires the use of creative imagination.

Galileo also distinguished between primary qualities and secondary qualities. Primary qualities—such as shape, size, and motion—are possessed by objects in themselves and are immutable, objective, and amenable to mathematical exploration. Secondary qualities, such as color and taste, are relative, subjective, and fleeting. They are caused on the senses by the primary qualities of objects. The world that science studies is the world of primary qualities. Subjective qualities can be left out of science without any loss. Galileo set for modern science the task of discovering the objective and real mathematical structure of the world. This structure, though mathematical, was also mechanical: All there is in the world is matter in motion.

THE INTERPRETATION OF NATURE

The emerging new science was leaving Aristotelianism behind. But it needed a new method. Better, it needed to have its method spelled out so that the break with Aristotelianism, as a philosophical theory of science, could be complete. Aristotelianism offered two criteria of adequacy for scientific method: epistemological adequacy and metaphysical adequacy. For epistemological adequacy, the scientific method had to meet some philosophical requirements as to what counts as knowledge. For metaphysical adequacy, the metaphysical presuppositions of scientific theories should coincide with the metaphysical presuppositions of philosophical theories. To different extents, the theories of scientific method developed in the seventeenth century were attempts to challenge these criteria, for they were considered more as fetters to science than enablers of its development.

In *Novum organum* (*The New Organon*; 1620/1960), Francis Bacon (1561–1626) placed method at center stage and argued that the world is knowable but only after a long process of trying to understand it—a process that begins with experience and is guided by a new method of induction by elimination. This new method differed from Aristotle's on two counts: on the nature of first principles and on the process of attaining them. According to Bacon, the Aristotelian method (which Bacon called

"anticipation of nature") starts with the senses and particular objects but then flies to first principles and derives from them further consequences. He contrasted this method to his own, which aims at an interpretation of nature, and which gradually and carefully ascends from the senses and particular objects to the most general principles. He rejected induction by enumeration as childish (since it takes account only of positive instances).

Bacon's alternative proceeds in three stages. Stage 1 involves compiling a natural and experimental history to derive a complete inventory of all instances of natural phenomena and their effects. Here observation rules. Then at stage 2, one constructs tables of presences, absences, and degrees of variation. Take, for example, the case of heat, which Bacon discussed in some detail. The table of presences records all phenomena with which the nature under examination (heat) is correlated (e.g., heat is present in light, etc.). The table of absences is a more detailed examination of the list of correlations of the table of presences that seeks to find absences (e.g., heat is not present in the light of the moon). The table of degrees of variation consists of recordings of what happens to correlated phenomena if the nature under investigation (heat) is decreased or increased in its qualities. Stage 3 is induction. Whatever is present when the nature under investigation is present or increases, and whatever is absent when this nature is absent or decreases, is the *form* of this nature. The crucial element in this three-stage process is the elimination or exclusion of all accidental characteristics of the nature under investigation. On the basis of this method, Bacon claimed that heat is motion and nothing else.

Bacon's forms are reminiscent of Aristotelian substantial forms. Yet he also claimed that the form of a nature is the law(s) it obeys. Indeed, Bacon's view was transitional between the Aristotelian view and a more modern conception of laws of nature. Bacon, in his view of science, found almost no place for mathematics, however, though he did favor active experimentation and showed great respect for alchemists because they had laboratories. In an instance of a fingerpost, he claimed that an essential part of interpreting nature by the new method of induction consists in devising a crucial experiment that judges between two competing hypotheses for the causes of an effect. Accordingly, Bacon distinguished between two types of experiments: those that gather data for a natural and experimental history and those that test hypotheses.

THE METAPHYSICAL FOUNDATIONS OF SCIENCE

René Descartes (1596–1650) too sought to provide an adequate philosophical foundation of science. But unlike Bacon, he felt more strongly the force of the skeptical challenge to the very possibility of knowledge of the world. So he took it upon himself to show how there could be certain (indubitable) knowledge and, in particular, how science can be based on certain first principles. Knowledge, he thought, must have the certainty of mathematics. Though Bacon was fine with some notion of virtual certainty, Descartes was after metaphysical certainty, that is, knowledge beyond any doubt. But in the end, Descartes accepted that in science a lot of things (other than the basic laws of nature) can be known only with virtual certainty. He distinguished all substances into two sorts: thinking things (*res cogitans*) and extended things (*res extensa*). He took the essence of mind to be thought and of matter extension. The vehicles of knowledge he took to be intuition and demonstration. We can be certain only of things that we can form clear and distinct ideas of or truths that we can demonstrate. Descartes tried to base his whole foundation for knowledge on a single indubitable truth, namely, "Cogito, ergo sum" ("I think; therefore I exist"). But having demonstrated the existence of God, he took God as guaranteeing the existence of the external world and, ultimately, of our knowledge of it.

Descartes was not a pure rationalist who thought that *all* science could be done a priori. Nor was he an empiricist either, obviously. He did not think that all knowledge stemmed from experience. In *Principia philosophiae* (*Principles of Philosophy*; 1644/1985), he argued that the human mind, by the light of reason alone, can arrive at substantive truths concerning the fundamental laws of nature. These laws (for instance, that the total quantity of motion in the world is conserved) are discovered and justified a priori, as they supposedly stem directly from God's immutability. Accordingly, the basic structure of the world is discovered independently of experience, is metaphysically necessary, and is known with metaphysical certainty. But once this basic structure has been laid down, science can use hypotheses and experiments to fill in the details. This is partly because the basic principles of nature place constraints on whatever else there is and happens in the world, without determining it uniquely. The less fundamental laws of physics are grounded in the fundamental principles, but are not directly deducible from them. Hypotheses are needed to flesh out these principles. Hypotheses are also needed to

determine particular causes and matters of fact in the world, such as the shape, size, and speeds of corpuscles. It is only through experience that the values of such magnitudes can be determined. Accordingly, Descartes thought that the less fundamental laws could be known only with virtual certainty. Descartes's view of nature was mechanical: Everything can be explained in terms of matter in motion.

NEWTON

The real break with the Aristotelian philosophical and scientific outlook occurred with the consolidation of empiricism in the seventeenth century. Empiricists repudiated the metaphysics of essences and the epistemology of rational intuition, innate ideas, and infallible knowledge. Modern philosophical empiricism was shaped by the work of three important figures: Pierre Gassendi (1592–1655), Robert Boyle (1627–1691), and Isaac Newton (1642–1727). Gassendi revived Epicurean atomism and stressed that all knowledge stems from experience. Boyle articulated the mechanical philosophy and engaged in active experimentation to show that the mechanical conception of nature is true.

Newton's scientific achievements, presented in his monumental *Philosophiae naturalis principia mathematica* (*Mathematical Principles of Natural Philosophy*) of 1687, created a new scientific paradigm. The previous paradigm, Cartesianism, was overcome. Newton's methodological reflections became the point of reference for all subsequent discussion concerning the nature and method of science. Newton demanded certain knowledge but rejected the Cartesian route to it. By placing restrictions on what can be known and on what method should be followed, he thought he secured certainty in knowledge. His famous dictum "Hypotheses non fingo" ("I do not feign hypotheses") was supposed to act as a constraint on what can be known. It rules out metaphysical, speculative, and nonmathematical hypotheses that aim to provide the ultimate ground of phenomena. Newton took Descartes to be the chief advocate of hypotheses of the sort he was keen to deny.

His official conception of the method of science was deduction from the phenomena. He contrasted his method with the broad hypothetico-deductive method endorsed by Descartes. Newton's approach was fundamentally mathematical and quantitative. He did not subscribe to the idea that knowledge begins with a painstaking natural and experimental history of the sort suggested by Francis Bacon. The basic laws of motion, in a sense, stem from experience. They are neither true a priori nor metaphysically necessary. Newton strongly disagreed with Gottfried Leibniz (1646–1716), who thought that laws of nature are contingent but knowable a priori through considerations of fitness and perfection. The empirically given phenomena that Newton started with are laws (e.g., Kepler's laws). Then, by means of mathematical reasoning and the basic axioms or laws of motion, he drew further conclusions, for example, that the inverse-square law of gravity applies to all the planets. This kind of deduction from the phenomena has been described as demonstrative induction. It is induction, since it ultimately rests on experience and cannot deliver absolutely certain knowledge. But it is demonstrative, since it proceeds in a mathematically rigorous fashion.

THE REVIVAL OF EMPIRICISM: LOCKE AND HUME

In his preface to *An Essay concerning Human Understanding* (1689), John Locke (1632–1704) praised "the incomparable Mr. Newton" and took his own aim to be "an Under-Labourer in clearing some Ground a little, and removing some of the Rubbish, that lies in the way of Knowledge." Locke was an empiricist and a nominalist. He thought that all ideas come from impressions and claimed that whatever exists is particular. He adopted as fundamental the distinction between primary and secondary qualities. He also drew a distinction between real essences and nominal essences. The real essence of a thing is its underlying internal constitution, based on its primary qualities. The nominal essence concerns the observable characteristics of a thing and amounts to the construction of a genus or a species. The nominal essence of gold, for instance, is a body yellow, malleable, soft, and fusible. Its real essence is its microstructure. Being a nominalist, he thought that real essences are individuals, whereas nominal essences are mere concepts or ideas that define a species or a kind. Though Locke argued that proper knowledge amounts to knowing the real essences of things, he was pessimistic about the prospects of knowing real essences. As he said, he suspected "that natural philosophy is not capable of being made a Science" (1689/1975, IV.12.10). To be sure, knowledge of nominal essences can be had, but Locke thought that this knowledge is trivial and uninteresting, since it is ultimately analytic. Even though Locke's famous book appeared after Newton's *Principia*, it is a pre-Newtonian work. It does not share Newton's optimism that the secrets of nature can be unlocked.

All empiricists of the seventeenth century accepted nominalism and denied the existence of universals. This

led them to face squarely the problem of induction. Realists about universals, including Aristotle, who thought that universals can exist only *in* things, could accommodate induction. They claimed that after a survey of a relatively limited number of instances, thought ascended to the universals shared by these instances and thus arrived at truths that are certain and unrevisable. This route was closed for nominalists. They had to rely on experience through and through, and inductive generalizations based on experience could not yield certain knowledge. This problem came in sharp focus in the work of David Hume (1711–1776).

The subtitle of Hume's *A Treatise of Human Nature* (1739/1978) was *Being an Attempt to Introduce the Experimental Mode of Reasoning into Moral Subjects*. This was an allusion to Newton's achievement and method. Hume thought that the moral sciences had yet to undergo their own Newtonian revolution. He took it upon himself to show how Newton's rules for philosophizing were applicable to the moral sciences. All ideas should come from impressions. Experience must be the arbiter of everything. Hypotheses should be looked upon with contempt. His own principles of association by which the mind works (resemblance, contiguity, and causation) were the psychological analogue of Newton's laws.

Being an empiricist, Hume argued that all factual (and causal) knowledge stems from experience. He revolted against the traditional view that the necessity that links cause and effect is the same as the logical necessity of a demonstrative argument. He argued that there can be *no* a priori demonstration of any causal connection, since the cause can be conceived without its effect and visa versa. Taking a cue from Nicolas Malebranche (1638–1715), he argued that there is no perception of a supposed necessary connection between cause and effect. Hume also went one step further. He found worthless his predecessors' appeals to the power of God to cause things to happen. Hume completely secularized the notion of causation. He also found inadequate, because circular, his predecessors' attempts to explain the link between causes and effects in terms of powers, active forces, and the like.

But his far-reaching point was that the alleged necessity of the causal connection cannot be empirically proved either. As he famously argued, any attempt to show, on the basis of experience, that a regularity that has held in the past *will* or *must* continue to hold in the future is circular and begs the question. It presupposes a principle of uniformity of nature. But this principle is *not* a priori true. Nor can it be proved empirically without circularity. For any attempt to prove it empirically will have to assume what needs to be proved, namely, that since nature has been uniform in the past, it will or must continue to be uniform in the future. Hume's challenge to any attempt to establish the necessity of causal connections on empirical grounds has become known as his *skepticism* about induction. But Hume never doubted that people think and reason inductively. He just took this to be a fundamental psychological fact about human beings that cannot be accommodated within the confines of the traditional conception of Reason. Indeed, Hume went on to describe in detail some basic "rules by which to judge of causes and effects" (1739/1978, p. 173).

KANT'S AWAKENING

Hume's critique of necessity in nature awoke Immanuel Kant (1724–1804) from his "dogmatic slumber," as he famously stated. Kant thought that Hume questioned the very possibility of science, and Kant took it upon himself to show how science was possible. He claimed that although all knowledge starts with experience, it does not arise from it. It is actively shaped by the categories of the understanding and the forms of pure intuition (space and time). The mind, as it were, imposes conceptual structure on the world, without which no experience could be possible. His central thought was that some synthetic a priori principles must be in place for experience to be possible.

Unlike Newton, Kant thought that proper science is not possible without metaphysics. Yet his understanding of metaphysics contrasted sharply with that of his predecessors. Metaphysics, Kant thought, was a science, in particular, the science of synthetic a priori judgments. Mathematics is a key element in the construction of natural science proper; without mathematics no doctrine concerning determinate natural things is possible. On these grounds, Kant argued that the chemistry of his age was more of an art than a science. The irony, Kant thought, was that though many past great thinkers (Newton in particular) repudiated metaphysics and relied on mathematics to understand nature, they failed to see that such reliance on mathematics made them unable to dispense with metaphysics. For, in the end, they had to treat matter in abstraction from any particular experiences. They postulated universal laws without inquiring into their a priori sources.

As Kant argued in his *Critique of Pure Reason* (1781/1965), the a priori source of the universal laws of nature is the transcendental principles of pure understanding. These constitute the object of knowledge in general. Thought (that is, the understanding) imposes on objects in general certain characteristics in virtue of

which objects become knowable. Phenomenal objects are constituted as objects of experience by the schematized categories of quantity, quality, substance, causation, and community. If an object is to be an object of experience, it must have certain necessary characteristics: It must be extended; its qualities must admit of degrees; it must be a substance in causal interaction with other substances. In his three Analogies of Experience, Kant tried to prove that three general principles hold for all objects of experience: that substance is permanent, that all changes conform to the law of cause and effect, and that all substances are in thoroughgoing interaction. These synthetic a priori principles make experience possible. In particular, there is the universal law of causation, namely, that "everything that happens, that is, begins to be, presupposes something upon which it follows by rule." This is nothing like an empirical generalization. Rather, it is imposed by the mind on objects.

Yet these transcendental principles make no reference to any objects of experience in particular. In his *Metaphysical Foundations of Natural Science* (1786/1970), Kant sought to show how these principles could be concretized in the form of laws of matter in motion. Kant thus enunciated the law of conservation of the quantity of matter, the law of inertia, and the law of equality of action and reaction, and he thought that these laws were the concrete mechanical analogues of his general transcendental principles. These laws were metaphysical laws in that they determined the possible behavior of matter in accordance with mathematical rules. They determine the pure and formal structure of motion, where motion is treated *in abstracto* purely mathematically. It is no accident, of course, that the last two of these laws (the law of inertia and the law of equality of action and reaction) are akin to Newton's laws and that the first law (the law of conservation of the quantity of matter) was presupposed by Newton too. Kant intended his metaphysical foundations of (the possibility of) matter in motion to show how Newtonian mechanics was possible. But Kant also thought that there are physical laws that are discovered empirically. Though he held as true a priori that matter and motion arise out of repulsive and attractive forces, he claimed that the laws of particular forces, even the law of universal attraction as the cause of gravity, can only be discovered empirically.

His predecessors, Kant thought, had failed to see the hierarchy of laws that make natural science possible: transcendental laws that determine the object of possible experience in general, metaphysical laws that determine matter in general, and physical laws that fill in the actual concrete details of motion. Unlike the third kind, laws of the first two kinds require a priori justification and are necessarily true. Though philosophically impeccable, Kant's architectonic suffered severe blows in the nineteenth and early twentieth centuries. The blows came, by and large, from science itself. Creating an explosive mixture that led to the collapse of Kant's synthetic a priori principles were the crisis of Newtonian mechanics, the emergence of Albert Einstein's special and general theories of relativity, the advent of quantum theory, the emergence of non-Euclidean geometries and their application to physics, Gottlob Frege's claim that arithmetic, far from being synthetic a priori, was a body of analytic truths, and David Hilbert's arithmetization of geometry, which proved that no intuition was necessary. It is no exaggeration to claim that much of philosophy of science in the first half of the twentieth century was an attempt to come to terms with the collapse of the Kantian synthetic a priori and to re-cast (or even cast to the wind) the concepts of the a priori and the analytic so as to do justice to developments in the sciences.

WHEWELL VERSUS MILL

The nineteenth century saw the culmination of Newtonian mechanics, mostly in the able hands of Pierre-Simon Laplace (1749–1827) and his followers. The Newtonian framework was extended to capture other phenomena, from optics, to heat, to electricity and magnetism. But Kant's philosophy was very much the doctrine that almost every serious thinker about science had to reckon with. William Whewell (1794–1866) took from Kant the view that ideas (or concepts) are necessary for experience in that only through them can facts be bound together. He noted, for instance, that induction gives rise to a "new mental element." The concept of elliptical orbit, he thought, was not already there in the astronomical data employed by Johannes Kepler, but was a new mental element added by Kepler. But, unlike Kant, he thought that history (and the history of science in particular) had a key role to play in understanding science and its philosophy. He analyzed this role in *The Philosophy of the Inductive Sciences, Founded upon Their History* (1840). Each science grows through three stages, Whewell thought. It begins with a "prelude," in which a mass of unconnected facts is collected. It then enters an "inductive epoch," in which the useful theories of creative scientists bring order to these facts—an act of "colligation." Finally, a "sequel" follows, where the successful theory is extended, refined, and applied. Whewell strongly emphasized the role of hypotheses in science. Hypotheses can be proven true, he

thought, by a "consilience of inductions," by which he meant the theoretical unification that occurs when a theory explains data of a kind different from those it was initially introduced to explain, and when a theory unifies hitherto unrelated domains. Indeed, Whewell found in the consilience of inductions a criterion of truth.

His contemporary John Stuart Mill (1806–1873) took an empiricist turn. Mill was a thoroughgoing inductivist who took all knowledge to arise from experience through induction. He even held that the law of universal causation, namely, that for every event there is a set of circumstances upon which it follows as an invariable and unconditional consequent, is inductively established. Hence, Mill denied that there could be any certain and necessary knowledge. But Mill also tried to delineate the scientific method so that it leads to secure causal knowledge of the world. In *A System of Logic, Ratiocinative and Inductive* (1843/1911) he put forward the method of agreement and the method of difference. According to the first, the cause is the common factor in a number of otherwise different cases in which the effect occurs. According to the second, the cause is the factor that is different in two cases that are similar except that the effect occurs in one, but not the other. In effect, Mill's methods encapsulate what is going on in controlled experiments. Mill was adamant, however, that his methods work *only if* certain substantive metaphysical assumptions are in place: that events have causes, that events have a *limited* number of possible causes, and that the same causes have the same effects, and conversely.

Mill was involved in a debate with Whewell concerning the role of novel predictions. Unlike Whewell, Mill thought that no predictions could *prove* the truth of a theory. He suggested that a hypothesis could not be proved true on the basis that it accounts for known phenomena, since other hypotheses may fair equally well in this respect. He added that novel predictions cannot provide proof either, since they carry no extra weight over predictions of known facts. Mill's target was not just the crude version of the method of hypothesis. He wanted to attack the legitimacy of the rival substantive assumption featured in Whewell's more sophisticated view, namely, that elimination of rival hypotheses can and should be based on explanatory considerations. The difference between Mill and Whewell was over the role of substantive explanatory considerations in scientific method. The debate continues.

CONVENTIONALISM

The inductivist tradition that flourished in England in the nineteenth century was challenged by the rise of French conventionalism. The work of Henri Poincaré (1854–1912) on the foundations of geometry raised the question of whether physical space is Euclidean. In *La science et l'hypothèse* (*Science and Hypothesis*; 1902/1952), Poincaré took this question to be meaningless, because, he suggested, one can make physical space possess *any* geometry one likes, provided that one makes suitable adjustments to one's physical theories. Consequently, he called the axioms of Euclidean geometry "conventions" (definitions in disguise). He extended his geometric conventionalism further by arguing that the principles of mechanics are also conventions. Conventions, for Poincaré, are general principles that are held to be true but whose truth can neither be the product of a priori reasoning nor be established on a posteriori grounds. But calling general principles "conventions" did not imply, for Poincaré, that their adoption (or choice) was arbitrary. He stressed that some principles were more convenient than others. He thought that considerations of simplicity and unity, as well as certain experiential facts, could and should guide the relevant choice. Indeed, he envisaged a hierarchy of the sciences in which the axioms of Euclidean geometry and the principles of Newtonian mechanics are in place (as ultimately freely chosen conventions) so as to make possible empirical and testable physical science.

Though Poincaré took scientific theories to be mixtures of conventions and facts, he favored a structuralist account of scientific knowledge that was Kantian in origin. The basic axioms of geometry and mechanics are (ultimately freely chosen) conventions, and yet, he thought, scientific hypotheses proper, even high-level ones such as Maxwell's laws, are empirical. Faced with discontinuity in theory change (the fact that some basic scientific hypotheses and laws are abandoned in the transition from one theory to another), he argued that there is, nonetheless, substantial continuity at the level of the mathematical equations that represent empirical and theoretical relations. From this, he concluded that the theoretical content of scientific theories is structural, by which he meant that a theory, if successful, correctly represents the *structure* of the world. In the end, the structure of the world is revealed by structurally convergent scientific theories.

THE RISE OF ATOMISM

The beginning of the twentieth century was marked by a heated debate over atomism, an emergent scientific theory that posited unobservable entities, atoms, to account

for a host of observable phenomena (from chemical bonding to Brownian motion). Though many scientists adopted atomism right away, there was strong resistance to it by other eminent scientists. Ernst Mach (1838–1916) resisted atomism on the basis of the empiricist claim that the concept of atoms was radically different from ordinary empirical concepts, and hence problematic. Resistance to atomism was best exemplified in the writings of Pierre Duhem (1861–1916). In *La théorie physique, son objet, sa structure* (*The Aim and Structure of Physical Theory*; 1906/1954), he put forward an antiexplanationist form of instrumentalism that sharply distinguished science and metaphysics, and claimed that explanation belongs to metaphysics and not to science.

But Duhem's theory of science rested on a restricted understanding of scientific method that can be captured by the equation "scientific method = experience + logic." On this view, whatever cannot be proved from experience with the help of logic is irredeemably suspect. To be sure, theories, as hypothetico-deductive systems, help scientists classify and organize the observable phenomena. But, for Duhem, the theoretical hypotheses of theories can never be confirmed or accepted as true. At best, they can be appraised as convenient or inconvenient, empirically adequate or empirically inadequate, classifications of the phenomena. Ironically, Duhem himself offered some of the best arguments against his own instrumentalist conception of theories. The most central one comes from the possibility of *novel* predictions. If a theory were just a "rack filled with tools," it would be hard to understand how it can be "a prophet for us" (Duhem 1906/1954, p. 27).

Duhem was a strong critic of inductivism. He argued that observation in science is not just the act of reporting phenomena. It is the interpretation of phenomena in the light of some theory and other background knowledge. This thesis, known as the view that observation is theory-laden, resurfaced in the 1960s, at that time drawing on a mass of empirical evidence coming from psychology to the effect that perceptual experience is theoretically interpreted. Duhem also stressed that there can be no crucial experiments in science, since no theory can be tested in isolation from other theories (and auxiliary assumptions), and consequently, that any theory can be saved from refutation by making suitable adjustments to collateral theories or auxiliary assumptions.

THE A PRIORI SET IN MOTION

Though battered by developments in physics and mathematics, the Kantian conception of a priori principles did find a place of sorts in the work of the neo-Kantian school of Marburg, Germany. In *Substance and Function* (1910/1923), Ernst Cassirer (1874–1945) argued that, though mathematical structures are necessary for experience, in that phenomena can be identified, organized, and structured *only* if they are embedded in such structures, these structures need not be fixed and immutable for all time. He thought that mathematical structures, though a priori (since they are required for objective experience), are revisable yet convergent: Newer structures accommodate old ones within themselves.

But it was Hans Reichenbach (1891–1953), in *The Theory of Relativity and A Priori Knowledge* (1921/1965), who unpacked the two aspects of Kant's conception of the a priori: that a priori truths are necessarily true, and that they structure objects of knowledge. Reichenbach rejected the first aspect of a priori knowledge, but insisted that the second aspect was inescapable. Knowledge of the physical world, he thought, requires principles of coordination, that is, principles that connect the basic concepts of the theory with reality. These principles he took to structure experience. Mathematics, he thought, was indispensable precisely because it provided a framework of general rules for coordinating scientific concepts and reality. Once this framework is in place, a theory can be presented as an axiomatic system, whose basic axioms (what Reichenbach called "axioms of connection") are empirical. Against Kant, Reichenbach argued that a priori principles of coordination, though they structure objects of knowledge, can be rationally revised in response to experience. He was naturally led to conclude that the only workable notion of the a priori is one that is *relativized*.

LOGICAL POSITIVISM

The influence of Moritz Schlick (1882–1936) on the philosophical course of events can hardly be exaggerated. Armed with the notion of convention, he and his followers, the logical positivists, tried to show that there can be no synthetic a priori at all. They extended conventionalism to logic and mathematics, arguing that the only distinction possible is between empirical (synthetic a posteriori) principles and conventional (analytic a priori) ones. In particular, though they thought that empirical science requires a logico-mathematical framework to be in place before theories can get any grip on reality, this conventional and analytic framework is purely *formal* and is *empty* of factual content. Accordingly, all a priori knowledge is analytic. Moreover, the logical positivists' conventionalist account of analyticity implies that grasping a priori (or analytic) truths requires no special faculty of intuition and that having epistemic access to a priori

(or analytic) truths presents no deep philosophical problem. Accompanying the doctrine that analytic truths are definitions or stipulations was the so-called linguistic doctrine of necessity: that all and only analytic truths are necessary. In the spirit of Hume, this doctrine excised all necessity from nature, and had already played a key role in Ludwig Wittgenstein's *Tractatus Logico-Philosophicus*.

The logical positivists adopted an empiricist criterion of meaning known as the verification principle. Nonanalytic statements, that is, synthetic empirical statements, are meaningful (cognitively significant) if and only if their truth can be verified in experience. In slogan form, the meaning is the method of verification. The logical positivists used this criterion to show that statements of traditional metaphysics were meaningless, since their truth (or falsity) made no difference in experience.

Soon after the foregoing criterion of meaning was adopted, a fierce intellectual debate started among members of the Vienna Circle, a debate that spanned a good deal of the 1930s and came to be known as the "protocol-statements debate." Protocol statements were supposed to capture the content of scientists' observations in such a basic form that they can be immediately verified. One issue was whether protocol statements are (should be) expressed in physical-object language ("The needle points to 2 on the dial") or in phenomenal language ("A black line overlies a "2" shape on a white background"). Though the balance soon turned in favor of the former, Rudolf Carnap (1891–1970), following Schlick, did toy with the idea that protocol statements need no justification, for they constitute the simplest states in which knowledge can be had. But he was soon convinced by the arguments of Otto Neurath (1882–1945) that there are neither self-justified protocol statements nor statements not subject to revision, if only because the processes that yield them are fallible. Instead of abandoning the claim that science provides knowledge, on the grounds that this knowledge cannot be certain, Carnap opted for the view that scientific knowledge falls short of certainty. Armed with Alfred Tarski's account of truth, he claimed that the truth of a scientific statement is no less knowable than the statement itself.

In the course of the 1930s, the concept of verifiability moved from a strict sense of being provable on the basis of experience to the much more liberal sense of being confirmable. The chief problem was that the strong criterion of cognitive significance failed to deliver the goods. In addition to metaphysical statements, many ordinary scientific assertions, those that express universal laws of nature, turn out meaningless on this criterion, precisely because they are not, strictly speaking, verifiable.

According to the logical positivists, Hilbert's approach to geometry and the Duhem and Poincaré hypothetico-deductive account of scientific theories, if combined, offer a powerful and systematic way to present scientific theories. The basic principles of the theory are taken to be the axioms. But the terms and predicates of the theory are stripped of their interpretation, or meaning. Hence, the axiomatic system itself is entirely formal.

The advantage of the axiomatic approach is that it lays bare the logical structure of the theory, which can then be investigated independently of the meaning, if any, one may assign to its terms and predicates. However, as a formal system, the theory lacks any empirical content. For the theory to acquire such content, its terms and predicates have to be suitably interpreted. It was a central thought of the logical positivists that a scientific theory need not be completely interpreted to be meaningful and applicable. They claimed that it is enough that only *some* terms and predicates, the so-called observational ones, be interpreted. The other terms and predicates of the theory, in particular, those that, taken at face value, purport to refer to unobservable entities, were deemed theoretical and were taken to be only partially interpreted by means of correspondence rules. It was soon realized, however, that the correspondence rules muddle the distinction between the analytic (meaning-related) part and the synthetic (fact-stating) part of a scientific theory—a distinction that was central in the thought of the logical positivists. For, on the one hand, the correspondence rules specify (even if only partly) the meaning of theoretical terms, and on the other hand, they contribute to the factual content of the theory.

A GHOSTLY DISTINCTION

A key idea developed in Carnap's *Logical Syntax of Language* (1934/1937) was that the development of a general theory of the logical syntax of the logico-mathematical language of science would provide a neutral framework in which scientific theories are cast and studied, scientific concepts (e.g., explanation, confirmation, laws, etc.) are explicated, and traditional metaphysical disputes are overcome. The project required a sharp analytic-synthetic distinction. Philosophical statements would be analytic (about the language of science), and scientific statements would be synthetic (about the world). A central (and stable) tenet of Carnap's was the principle of tolerance. Since the choice of a language is a conventional matter (to be evaluated only in terms of its practical fruitfulness), the

aim of philosophy of science, Carnap held, is to make clear the different language *forms* adopted by rival parties in philosophical and scientific disputes (e.g., the dispute between logicists and intuitionists in mathematics, or between realists and idealists, Platonists and nominalists, scientific realists and instrumentalists in philosophy of science). Far from being genuinely factual, these disputes, Carnap thought, center on suitable choices of a language. The principle of tolerance is thus part of Carnap's attempt to eliminate metaphysical "pseudoproblems" from the sciences. It formulates a metatheoretical standpoint in which issues of ontology are replaced by issues concerning logical syntax.

Carnap's project in *The Logical Syntax of Language* came to grief. This was the result of many factors, but prominent among them were Tarski's work on truth (which suggested that truth is an irreducibly semantic notion) and Kurt Gödel's incompleteness theorem. Though Carnap was fully aware of Gödel's limitative results, his own attempt to provide a neutral, minimal metatheoretical framework (the framework of "General Syntax" [1934/1937, pt. IV]) in which the concept of analyticity was defined fell prey to Gödel's proof that some mathematical truths are not provable within such a system.

The notion of analytic a priori truths came under heavy attack from W. V. O. Quine (1908–2000). In "Two Dogmas of Empiricism" (1951), Quine argued that the notion of analyticity is deeply problematic, since it requires a notion of cognitive synonymy (sameness of meaning) and there is no independent criterion of cognitive synonymy. Quine's chief argument against the analytic/synthetic distinction rested on the view that "analytic" was taken to mean unrevisable. If analytic statements have no empirical content, experience cannot possibly have any bearing on their truth-values. So analytic statements cannot undergo truth-value revision. But, Quine argued, nothing (not even logical truths) is unrevisable. Hence, there cannot be any analytic truths. Here Quine took a leaf from Duhem's book (and also from Carnap's book). Confirmation and refutation are holistic; they accrue to systems (theories) as a whole and not to their constituent statements, taken individually. If a theory is confirmed, then everything it says is confirmed. Conversely, if a theory is refuted, then *any* part of it can be revised (abandoned) to restore accord with experience. The image of science that emerged had no place for truths with a special status: all truths are on a par. This leads to a blurring of the distinction between the factual and the conventional. What matters for Quine is that a theory acquires its empirical content as a whole, by issuing in observational statements and by being confronted with experience.

The cogency of Quine's attack on the a priori rests on the cogency of equating the notion of a priori with the notion of unrevisable. We have already seen a strand in post-Kantian thinking that denied this equation, while holding onto the view that some principles structure experience. It might not be surprising, then, that Carnap was not particularly moved by Quine's criticism. For he *too* denied this equation. Quine, however, did have a point. For Carnap, (a) it is rational to accept analytic statements within a linguistic framework; (b) it is rational to reject them when the framework changes; and (c) all and only analytic statements share some characteristic that distinguishes them from synthetic statements. Even if Quine's criticisms are impotent against (a) and (b), they are quite powerful against (c). The point was simply that the dual role of correspondence rules (and the concomitant Hilbert-style implicit definition of theoretical terms) made drawing this distinction impossible, even *within* a theory. Carnap spent a great deal of effort to develop the characteristic specified in (c). In the end, he had to reinvent Ramsey sentences to find a plausible way to draw the line between the analytic and the synthetic (Psillos 1999, chap. 3).

The challenge to the very possibility of a priori knowledge was a key factor in the *naturalist turn* in the philosophy of science in the 1960s. The emergence of naturalism was a real turning point in the philosophy of science, because it amounted to an ultimate break with neo-Kantianism in all its forms. By the 1960s, philosophy of science had seen the advent of psychologism, naturalism, and history of science.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Constructivism and Conventionalism; Laws of Nature; Laws, Scientific; Philosophy of Science, Problems of; Scientific Realism.

Bibliography

Aristotle. *Posterior Analytics*. 2nd ed. Oxford, U.K.: Clarendon Press, 1993.

Bacon, Francis. *The New Organon* (1620), edited by Fulton H. Anderson. New York: Macmillan, 1960.

Burtt, Edwin A. *The Metaphysical Foundations of Modern Physical Science*. 2nd ed. London: Routledge and Kegan Paul, 1932.

Carnap, Rudolf. *The Logical Syntax of Language* (1934). Translated by Amethe Smeaton. London: Kegan Paul, 1937.

Cassirer, Ernst. *Substance and Function* (1910). Translated by William Curtis Swabey and Marie Collins Swabey. Chicago: Open Court, 1923.

Cohen, I. Bernard. *The Birth of a New Physics*. London: Penguin, 1985.

Descartes, René. *Principles of Philosophy* (1644). In *The Philosophical Writings of Descartes*, Vol. 1. Translated by John Cottingham, Robert Stootfoff, and Dugald Murdoch. Cambridge, U.K.: Cambridge University Press, 1985.

Duhem, Pierre. *The Aim and Structure of Physical Theory* (1906). Translated by Philip Wiener. Princeton, NJ: Princeton University Press, 1954.

Duhem, Pierre. *To Save the Phenomena* (1908). Translated by Edmund Doland and Chaninah Mascher. Chicago: University of Chicago Press, 1969.

Duns Scotus, John. *Philosophical Writings: A Selection*. Translated by Allan Wolter. Indianapolis, IN: Hackett, 1987.

Gower, Barry. *Scientific Method: An Historical and Philosophical Introduction*. London: Routledge, 1998.

Grant, Edward. *Physical Science in the Middle Ages*. Cambridge, U.K.: Cambridge University Press, 1977.

Hume, David. *A Treatise of Human Nature* (1739), edited by L. A. Selby-Bigge. 2nd ed. edited by P. H. Nidditch. Oxford, U.K.: Clarendon Press, 1978.

Kant, Immanuel. *Critique of Pure Reason* (1787). Translated by Norman Kemp Smith. New York: St. Martin's Press, 1965.

Kant, Immanuel. *Metaphysical Foundations of Natural Science* (1786). Translated by James Ellington. Indianapolis, IN: Bobbs-Merrill, 1970.

Lindberg, David C., ed. *Science in the Middle Ages*. Chicago: University of Chicago Press, 1978.

Locke, John. *An Essay concerning Human Understanding* (1689). Oxford, U.K.: Clarendon Press, 1975.

Losee, John. *A Historical Introduction to the Philosophy of Science*. 4th ed. Oxford, U.K.: Oxford University Press, 2001.

Mill, John Stuart. *A System of Logic, Ratiocinative and Inductive* (1843). 8th ed. London: Longmans, Green, 1911.

Ockham, William. *Philosophical Writings: A Selection*. Translated by Philotheus Boehner. Indianapolis, IN: Hackett, 1990.

Poincaré, Henri. *Science and Hypothesis* (1902). New York: Dover, 1952.

Psillos, Stathis. *Scientific Realism: How Science Tracks Truth* (1999). London: Routledge.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43.

Reichenbach, Hans. *The Theory of Relativity and A Priori Knowledge* (1921). Translated by Maria Reichenbach. Berkeley: University of California Press, 1965.

Thayer, H. S., ed. *Newton's Philosophy of Nature: Selections from His Writings*. New York: Hafner, 1953.

Whewell, William. *The Philosophy of the Inductive Sciences, Founded upon Their History*. London: J. W. Parker, 1840.

Whewell, William. *Theory of Scientific Method*. Indianapolis, IN: Hackett, 1989.

Stathis Psillos (2005)

PHILOSOPHY OF SCIENCE, PROBLEMS OF

The scope of the philosophy of science is sufficiently broad to encompass, at one extreme, conceptual problems so intimately connected with science itself that their solution may as readily be regarded a contribution to science as to philosophy and, at the other extreme, problems of so general a philosophical bearing that their solution would as much be a contribution to metaphysics or epistemology as to philosophy of science proper. Similarly, the range of issues investigated by philosophers of science may be so narrow as to concern the explication of a single concept, considered of importance in a single branch of science, and so general as to be concerned with structural features invariant to all the branches of science, taken as a class. Accordingly, it is difficult to draw boundaries that neatly separate philosophy of science from philosophy, from science, or even from the history of science, broadly interpreted. But we can give some characterization of the main groups of problems if we think of science as concerned with providing descriptions of phenomena under which significant regularities emerge and with explaining these regularities. Problems thus arise in connection with terms, with laws, and with theories where a theory is understood as explaining a law and a law is understood as stating the regularities that appear in connection with descriptions of phenomena.

TERMS

Ordinary language provides us the wherewithal to offer indefinitely rich descriptions of individual objects, and, as a matter of logical fact, no description, however rich, will exhaustively describe a given object, however simple. Science chooses a deliberately circumscribed vocabulary for describing objects, and scientists may be said to be concerned only with those objects described with the vocabulary of their science and with these only insofar as they are so describable. Historically, the terms first applied by scientists were continuous with their cognates in ordinary speech, just as science itself was continuous with common experience. But special usages quickly developed, and an important class of philosophical problems concerns the relation between scientific and ordinary language, as well as that between those terms selected for purposes of scientific description and other terms that, though applicable to all the same objects as the former, have no obvious scientific use. Scientists from Galileo Galilei to Arthur Eddington have sometimes tended to impugn as unreal those properties of things not

covered by scientific description or at least have thought that the question of which are the real properties is an important one. Certainly, it would destroy the very concept of science to suppose it possible to account for all the distinctions between things under all the descriptions of them that are feasible, but there is no recipe for selecting the scientifically relevant predicates.

In practice, terms have been chosen when there seem to be interesting and systematic patterns of change in the properties picked out by these terms—for instance, between the distance a body travels and the time it takes to do so, between the temperature and the pressure of a gas, between the density of a fluid and the deviation from a norm of a light ray passing into it, and so forth. It has often been immensely difficult to set aside manifest and cherished differences among objects and the subtle language for expressing these in favor of the spare vocabulary of science under which such seemingly crucial distinctions are obliterated, as, for example, between celestial and terrestrial objects or between "noble" and base metals.

Not only do scientific terms cut across the distinctions of common sense, but they also permit distinctions not ordinarily made and allow comparisons more precise than ordinarily demanded—for example, between differential amounts and precisely determinable degrees. For the class of terms discussed here are those that may be said to apply or not to apply to a given object by means of an act of observation rendered precise through some device of mensuration—for example, that the distance traveled is *n* units along a scale, that the temperature of a gas is *n* degrees along another scale, that the density of a fluid is *m* grams per cubic centimeter. The last measurement, which involves reference to different scales—namely, measures of mass and volume—is sometimes called a "derived" in contrast with a "fundamental" measurement, where only single scales are involved. But even when we speak of derived measurements, as with pressure (in terms of foot-pounds), velocity (in terms of feet per second), or stress (in terms of force per unit area), we remain within the domain of observation; the coincidence of a needle with a mark on a gauge, the angle of a balance, the appearance of a color, a bubble between lines, or a certain buzz, inform us that a given term is true or false with respect to whatever we are studying.

Philosophers may press for a further reduction of the observational language of a science to a favored idiom—for example, to a sense-datum language—but within science observational vocabulary enjoys a certain ultimacy. There are many questions as to whether observational language, thus construed, is sufficient for the entire conduct of science, whether the whole language of science can be expressed in purely observational terms so that recourse need never be made to covert entities, hidden processes, or occult structures unamenable to direct observation and measurement. This issue cannot be fruitfully discussed until we come to the topic of theories, but it has been recognized that while observation has an essential role to play as the occasion for framing and the basis for testing scientific hypotheses, the no less important feature of measurement sets a limit on the program of thoroughgoing observationalism. For the algorithms, in connection with which it first makes scientific sense to assign numerical values and to apply scales, require use of the real number system, the class of whose values has the power of the continuum.

Hence, as Carl G. Hempel remarked, "A full definition of metrical terms by means of observables is not possible." Nevertheless, it has been through the efforts of reductionists to assimilate the entirety of scientific language to observation terms that other sorts of terms, having logically distinct roles within science, have been discovered, and a main task in philosophy of science has been to identify and determine the relation between terms occurring at different levels, and variously related to observation, within the idiom of developed scientific theories.

LAWS

One cannot very readily treat the syntactical features of laws in isolation from their semantic properties or, for that matter, from pragmatic considerations. Syntax here concerns the formal conditions of "lawlikeness" for sentences, and semantics concerns the truth conditions for lawlike sentences, it being customary to define a law as a true lawlike sentence. But some philosophers will reject this definition since it might rule out any sentence as having the status of a law, inasmuch as laws are not, they feel, the sorts of sentences that it makes sense to regard as admitting truth-values in the normal way or even at all; for these a law would be a lawlike sentence which has a certain use.

It is commonly supposed that a universally quantified conditional sentence—$(x)(Fx \supset Gx)$—is the simplest form with which a lawlike sentence may be expressed. The chief syntactical problems arise, however in connection with the nonlogical terms *F* and *G*. For an important class of cases these will be observational, so that it is in principle possible to determine whether a given instance is both *F* and *G*, and the law is generally based upon some

known favorable instances, Yet there are cases in which the terms satisfy observational criteria, in which there are a large class of favorable instances and no known counterinstances, and still the appearance of these terms in a lawlike sentence *L* disqualifies *L* as a law even if it is true. Such terms are unduly restricted in scope, whereas it is thought that the terms suitable for laws should be unrestricted in scope. "All the hairs on my head are black" employs the restrictive term "the hairs on my head" and thus is disqualified as a law.

A criterion sometimes advanced for identifying restrictive terms as antecedents in possible laws is that if the requisite universal conditional supports a true counterfactual, it is a law, but if the counterfactual is false, as (with reference to a certain white hair) "If that hair were on my head, it would be black" is false, then the corresponding sentence is not a law, and the term is restricted. However, this criterion begs the question insofar as it seems that counterfactuals must be analyzed in terms of general laws; at any rate, the analysis of counterfactuals, as well as the basis for distinguishing true from false counterfactuals, remains to be given by philosophers. In what sense "the hairs on my head" is restrictive, whereas *ravens* in "All ravens are black" is not, is difficult to specify, though the former does refer to a specific object (my head) and it is believed that the terms in a law must not make such references. This restriction, however, makes Johannes Kepler's laws laws in name only and forestalls the possibility of any laws for the universe as a whole. And though Kepler's laws may be retained since they are derivable from laws that employ unrestricted and generally referential terms, the laws of the universe hardly could be thus derived; moreover, it could be argued that "All the hairs on my head are black" might be derivable from some general laws of hirsuteness, making use only of purely qualitative predicates. Thus, precise and rigorous criteria for lawlikeness are difficult to specify.

If the terms of a lawlike sentence *L* must be unrestricted, *L* cannot be known as true through induction by finite enumeration; since there must in principle always be uninspected instances under *F*, the law $(x)(Fx \supset Gx)$ cannot be known true no matter how many known favorable instances there are. Of course, laws are not always (and perhaps not even often) inductive generalizations from large samples—Galileo's laws, for instance, were based upon few observations indeed—and it has been maintained by anti-inductivists (chiefly Karl Popper and his followers) that observations function as tests rather than inductive bases for laws; in this view laws need not be generalizations from observation but only be in principle falsifiable on the basis of observation. Some lawlike sentences may be known false, at least to the extent that they admit of observational consequences, but often the antecedent of a lawlike sentence is sufficiently hedged with *ceteris paribus* riders, to which we may add indefinitely, that one need not surrender a law save as an act of will.

This suggests that the criteria for accepting a lawlike sentence as a law are more complex than either inductivists or their opponents have recognized, and an instrumentalist position may be taken, in accord with which laws are neither true nor false but serve as instruments in the facilitation of inference—"inference-tickets," as Gilbert Ryle put it. In this view, as Stephen Toulmin pointed out, the question is not "'Is it true?' but 'When does it hold?'" Here laws are regarded not as sentences about the world but as rules for conducting ourselves in it, and semantic considerations thus yield to pragmatic ones in that there is surely some agreement that a criterion for accepting *L* as a law is that it should, in conjunction with information, furnish successful predictions. Whether, in addition, a successful law is true and, if so, in what sense it is true other than that it successfully enables predictions cannot be discussed independently of larger philosophical considerations.

Many laws in science are statistical in form, but the suggestion that a law may be truly scientific and yet affirm a merely probable connection among phenomena has been offensive to scientists and philosophers with antecedent commitments to determinism as a metaphysical fact or a scientific ideal. For these nothing less than deterministic (nonstatistical) laws are ultimately tolerable, so that statistical laws, while countenanced as interim makeshifts, are, ideally, to be replaced in every instance with deterministic ones. As a program, however, the projected reconstruction of statistical laws and the theories that contain them has encountered an impressive obstacle in the quantum theory of matter, upon which the whole of atomic physics is based, for the laws here are demonstrably irreducible to deterministic form.

To be sure, there is a logical possibility that quantum theory could be replaced in toto. But there is no way—for instance through the discovery of hidden variables—in which its laws may be rendered deterministic, and since there is scant evidence for any alternative and the evidence for quantum theory is overwhelming, most members of the scientific community are reconciled to an obdurate indeterminism at the core of one of its most fundamental theories. If the quantum theory should be true, certain events are objectively probable, or indeter-

ministic; that is, they are probable independently of the state of our knowledge or ignorance.

An epistemological sense of probability, connected with our concepts of induction and confirmation, is not incompatible with determinism; we may even speak of the probability of a deterministic law, meaning that relative to our evidence its degree of confirmation is equal to a number between 0 and 1. It is nonepistemological probability, according to which we could conceivably be certain that a given event were objectively probable, which is allegedly repugnant to determinism. It should be pointed out, however, that indeterministic laws may be deterministic in at least the sense that the values of certain probability variables are precisely determined by the values of other variables. At any rate, the extent of incompatibility between determinism and indeterministic laws and the precise explication of the two kinds of probability are topics of continuing philosophical investigation and controversy.

Laws are believed to play an important role in explanation as well as in prediction. It has been maintained that a necessary condition for explaining an event *E* consists in bringing *E* under the same general law with which it could have been predicted. Hempel regards the temporal position of the scientist vis-à-vis the event as the sole difference between explaining and predicting that event. This symmetry has been challenged (notably by Israel Scheffler), but we might still maintain Hempel's thesis by distinguishing among laws. Not every law used in prediction has explanatory force if we think of explanations as causal explanations, for causal laws do not exhaust the class of scientific laws, which also includes functional expressions of covariation among magnitudes, statistical laws, and so on, all of which are used in predicting. Even so, it has been questioned whether even causal explanation requires the use of causal laws, either in science or in history or the social sciences, where this controversy has been chiefly focused.

Be this as it may, the explanation of particular events has less importance in science proper than the explanation of regularities, and it is therefore the explanation of laws that characterizes scientific achievement in its most creative aspect. This brings us to theories, for it is commonly held that to explain a law *L* is to derive *L* from a theory *T* when *T* satisfies certain conditions.

THEORIES

Let us characterize a law all of whose nonlogical terms are observational as an empirical law. A theory may be regarded as a system of laws, some of which are empirical. Not every empirical law is part of a theory, nor are all the laws of a theory empirical, for some of a theory's laws employ theoretical terms, which are nonobservational. Theoretical terms, if they denote at all, refer to unobservable entities or processes, and it is with respect to changes at this covert level that one explains the observed regularities as covered by empirical laws. Thus one explains the regularities covered by the Boyle–Charles law (all the terms of which are observational) in terms of the (unobservable) behavior of the gas molecules of which the gas is theoretically composed. The status of theoretical terms (and the theoretical entities they would designate if they designated anything) has been the subject of intense philosophical investigation. It is not mere unobservability—Julius Caesar is at this point in time unobservable though his name is not a theoretical term—but unobservability in principle that characterizes these entities; it is unclear whether there would be any sense in speaking of observing, say, Psi-functions, electrons, fields, superegos, and the like. Moreover, the behavior of theoretical entities, supposing the theory to be true, is (as with certain fundamental particles) often so grossly disanalogous to the behavior of the entities they are invoked to explain that our ordinary framework of concepts fails to apply to them.

Yet theoretical terms seem deeply embedded in scientific language. Empiricist strategies of eliminating them by explicit definition in observational language or of tying them to observation by reduction sentences have failed, although there exist techniques by which they may be formally replaced with striking ease. William Craig demonstrated that any theory containing both theoretical and observational predicates may be replaced with another employing only observational ones but yielding, nevertheless, all the observational theorems (or empirical laws) of the original. Craig's result, however, has not been a victory for empiricism; the reasons for this are somewhat obscure, but it is due in part at least to the realization that theoretical terms play a role and have a meaning in terms of the total structure of the theory and therefore cannot be neatly extricated to leave anything to be called a "theory." Indeed, it often happens that rather than theoretical terms being defined in observational terms, observational terms are defined with reference to the theoretical vocabulary, so that one must, in effect, master the theory in order to make the relevant observations.

With the elaboration of a theory, however, the inferential route from observation to (predicted) observation becomes complex (there may be many intervening steps and intermediate computations) and far removed from

the simple universal conditional used to represent a law. A theory, in Hempel's words, "may be likened to a complex spatial network [which] floats, as it were, above the plane of observation and is anchored to it by rules of interpretation." Theories, that is, impinge upon experience as wholes but not in all their parts, and the rules of interpretation, or correspondence, which permit them to be applied, are not part of the theory; indeed, the same formal theoretical network might, through different interpretations, have application to different domains of experience.

We may think of a theory as a formal system distinguishable, in principle, from its interpretation, regarding the former (in R. B. Braithwaite's terms) as a calculus and the latter as its model. In point of scientific history and practice, however, model and calculus emerge together. The distinction first began to be clear through the advent of non-Euclidean geometries and the consequent agitated question of which was physically descriptive, and geometry, perhaps because it has been almost paradigmatic of axiomatic systems, has served as a pattern, at least for analytical purposes, for the calculi of theories generally. Thus, philosophers think of theories as employing primitive and derived terms, primitive and derived sentences, satisfying explicit formation and transformation rules, and the like. But whether, apart from the purposes of philosophical representation, actual scientific theories exhibit axiomatized form and whether axiomatization is even a desideratum for scientific theory-formation are moot points.

At any rate, the framing of theories in the course of history has almost always involved some intuitive model on the scientist's part, the pattern of thought being (whether this is or is not the "logic of discovery" that N. R. Hanson suggested) this, that the regularities for which explanation is sought would hold as a matter of course if certain states of affairs (those postulated by the theory) held in fact. Whether the theoretical states do hold in fact is, of course, the immediate question, and it is through the obligation to provide an answer that the scientific imagination is disciplined. Without the formal means of deriving testable consequences from a theory, the theory would merely be ad hoc, and one wants more than the mere deduction of the laws that the theory was intended to explain. Indeed, it is by and large the ability of a theory to permit derivations far afield from its original domain that serves as a criterion for accepting a theory, for in addition to the obvious fruitfulness such a criterion emphasizes, such derivations permit an increasingly broad and diversified basis for testing the theory. The great theories in the development of science—Isaac Newton's, Albert Einstein's, Paul Dirac's—have brought into a single comprehensive system great numbers of phenomena not previously known to have been connected.

It is impossible to say, of course, whether the whole of scientific knowledge might someday be embraced in a single unified theory, but piecemeal assimilation of one theory to another is constantly taking place, and the conceptual issues that arise through such reductions are of immense philosophical interest. The careful elucidation of the logic of scientific reduction—of thermodynamics to mechanics, of wave and matrix mechanics—draws attention to features that lie, far more obscurely, within the oldest philosophical problems and controversies: problems of emergence, of natural kinds, of free will and determinism, of body and mind, and so on. The treatment of these questions is often not so much philosophy of science proper as the philosophical interpretation of science, in which the philosophy of science serves as a technique of philosophical clarification, illuminating topics remote from the conceptual issues of science as such.

See also Braithwaite, Richard Bevan; Eddington, Arthur Stanley; Empiricism; Explanation; Force; Galileo Galilei; Hempel, Carl Gustav; Laws, Scientific; Matter; Popper, Karl Raimund; Quantum Mechanics; Ryle, Gilbert; Thought Experiments in Science.

Bibliography

The literature on the philosophy of science is immense and often technical. Many influential papers are anthologized in *Philosophy of Science*, edited by Arthur C. Danto and Sidney Morgenbesser (New York: Meridian, 1960), and in *Readings in Philosophy of Science*, edited by Herbert Feigl and May Brodbeck (New York: Appleton-Century-Crofts, 1953).

The following may be consulted as representative and excellent discussions. On preliminary definition of the field see R. B. Braithwaite, *Scientific Explanation* (Cambridge, U.K.: Cambridge University Press, 1953), pp. 1–9, and Israel Scheffler, *The Anatomy of Inquiry* (New York: Knopf, 1963), pp. 3–15. On the observational bases of science see Carl G. Hempel, *Fundamentals of Concept Formation in Empirical Science* (Chicago: University of Chicago Press, 1952), especially pp. 20–50, and Scheffler, op. cit., pp. 127–222. On measurement see Ernest Nagel "Measurement," in Danto and Morgenbesser, op. cit., pp. 121–140. On laws see C. F. Presley, "Laws and Theories in the Physical Sciences," in Danto and Morgenbesser, op. cit., pp. 205–215; Ernest Nagel, *The Structure of Science* (New York: Harcourt Brace, 1960), especially pp. 29–78; and Stephen Toulmin, *Philosophy of Science* (London: Hutchinson, 1953), pp. 57–104. On the nonlogical terms of laws see Carl G. Hempel and Paul Oppenheim, "The Logic of Explanation," in Feigl and Brodbeck, op. cit., pp. 319–352. On laws and

falsifiability see Karl Popper, *The Logic of Scientific Discovery* (New York: Hutchinson, 1959), pp. 27–48. On explanation see Hempel and Oppenheim, op. cit. On the parity between explanation and prediction see Scheffler, op. cit., pp. 43–88. On theories see Braithwaite, op. cit., pp. 50–114; Nagel, *The Structure of Science*, op. cit., pp. 79–152; Presley, op. cit., pp. 215–225; Toulmin, op. cit., pp. 105–139. On Craig's theorem see Carl G. Hempel, "The Theoretician's Dilemma," in *Minnesota Studies in the Philosophy of Science*, edited by Herbert Feigl et al., Vol. II (Minneapolis: University of Minnesota Press, 1958), and Scheffler, op. cit., pp. 193–203. On the logic of discovery see N. R. Hanson, *Patterns of Discovery* (Cambridge, U.K.: Cambridge University Press, 1958), passim. On reduction of theories and related issues, see Nagel, *The Structure of Science*, pp. 336–397.

Arthur C. Danto (1967)

PHILOSOPHY OF SEX

In the last quarter of the twentieth century a distinct, new subarea of philosophy came to life, the philosophy of sex. Many philosophical books and professional journal articles on various aspects of sex appeared in print during this period; university-level courses devoted substantially or entirely to the philosophy of sex proliferated, as did textbooks for these courses (the first, the anthology *Philosophy and Sex*, was published in 1975, edited by Robert Baker and Frederick Elliston); and in 1977 a professional organization, The Society for the Philosophy of Sex and Love, was founded.

The new philosophical investigation of sexuality emerged partially in concert with second-wave feminism's critique of both the politics of sexual difference, including gender discrimination, and the politics of sexual desire and behavior, including widespread social and legal contempt for the sexual preferences and lifestyles of gays, lesbians, transsexuals, and the transgendered. But the philosophy of sex was (and has been) historically and thematically separate from any particular ethical, political, metaphysical, or religious perspective. Indeed, the discipline encompasses a host of viewpoints, schools, approaches, and methods, as shown by its eclectic teaching and research materials, for example, Igor Primoratz's collection *Human Sexuality* (1997) and Alan Soble's encyclopedia *Sex from Plato to Paglia* (2005).

By the early twenty-first century, scholars working in the philosophy of sex had exhumed much of its history, although many figures and movements remained to be explored. They had also written about numerous conceptual, ontological, ethical, and political matters. In addition to "sexual activity" and "sexual desire," perhaps the two fundamental concepts (or phenomena) of the area, subjects investigated included marriage (same- and other-sex), fidelity and adultery, consent and coercion, seduction, exploitation, sexual objectification, sexual harassment, rape, date and acquaintance rape, pornography, prostitution (and other sex work), sexual perversion, incest, pedophilia, group sex, masturbation, sexual orientation, sadomasochism, and sex with and without love, commitment, or psychological intimacy (casual sex, promiscuity). Analytic, existentialist, phenomenological, poststructuralist, postmodernist, evolutionary, conservative, liberal, feminist, Marxist, and diverse religious philosophers have all had their say.

A HISTORY OF THE PHILOSOPHY OF SEX

The philosophical discussion of sex in the West began with the ancient Greek philosopher Plato (427–347 BCE). His dialogues *Symposium* and *Phaedrus*, which are about *eros* (identified in the former work as a powerful passion to possess the good and beautiful), are provocative, astute, and an indispensable foundation for anyone interested in pursuing the philosophy of sex. Although Plato's student Aristotle (384–322 BCE) had little to say about *eros*, he meditates at length in his *Nicomachean Ethics* (books 8, 9) about *philia* (friendship-love), arguing that genuine friends improve each other's virtue and want the good for each other for each other's sake. Those who engage in research in the philosophy of sex commonly also study the related phenomena of love and friendship. Furthermore, the philosophy of sex generates its most instructive results when approached interdisciplinarily, that is, when it pays attention not only (and most obviously) to the psychology of sex and love but also to the sociology and history of mating practices and marriage forms, the anthropology of sexual and fertility rites and rituals, and the anatomical, physiological, and genetic findings of biomedical science.

Between antiquity and the twenty-first century, many philosophers, theologians, and others in the humanities made significant contributions to the richness of the philosophy of sex. Among the figures who made a lasting impact is St. Augustine (354–430), the Bishop of Hippo (in North Africa). Augustine was a profound thinker about sex and the human condition, as can be seen in his *The City of God* (for example, book 14), in which he expresses apprehension (as Plato did) about the threat to self-mastery and individual contentment by the forcefulness of the sexual impulse. Also noteworthy are the people with whom Augustine had theological dis-

putes over the nature of the prelapsarian sexuality of Adam and Eve and the effects on sexuality of the Fall: on the one side, the radically more sexually ascetic St. Jerome (the translator of a Latin Vulgate bible, in 380) and, on the other, the much more sexually relaxed Pelagians, including Julian (c. 386–454), Bishop of Eclanum—battles recounted well by Princeton University historian of religion Elaine Pagels (1988). Innumerable later medieval theologians were also important (see Brundage 1987), from Peter Abelard and his student, lover, and wife Heloise, whose tragic lives and impassioned letters are lessons in ardent sexual desire and an equally ardent Christianity, to St. Thomas Aquinas, tutored by Albertus Magnus (who also set about to merge Catholicism with Aristotle). In his stupendous *Summa theologiae* (1265–1273), Aquinas formulated a natural law theory that eventually (1879) became the authoritative foundation of Catholic teaching about sexuality.

After the medievalists, there came, from 1500 to 1900, a stream of colorful scholars: the skeptic Michel De Montaigne (1533–1592), author of the famous essay "On Friendship" and the lesser known "Of the Power of the Imagination," on sexuality; the French mathematician and rationalist philosopher René Descartes, whose last book (1649) was *The Passions of the Soul*; the Scottish empiricist philosopher David Hume, who proposed in his monumental *A Treatise of Human Nature* (1739–1740) that the amorous passion "betwixt the sexes" was composed of three discordant elements: kindness, lust, and a response to beauty (2.2.11); the Englishman Thomas Hobbes (life in the state of nature, he wrote in 1651, is "solitary, poor, nasty, brutish, and short" [*Leviathan*, sec. 1.13]), who contended in his earlier "Human Nature" (sec. 9.15) that sexual desire is actually composed of two distinct desires, a desire to be sexually pleased by the other person and (as anomalous as it sounds) a desire to please the other; his adversary, a defender of the state of nature, Jean-Jacques Rousseau (1712–1778), who promulgated terrifying warnings about the evils of self-abuse (the solitary vice) in his autobiography, *Confessions*, and in a treatise devoted to educational techniques, *Emile*; the philosopher and physician Bernard Mandeville, who, in *A Modest Defence of Publick Stews* (1724), praised prostitution in part because it prevented self-abuse, or so he was convinced; the bachelor Immanuel Kant (1724–1804), who alleged that sexual love not combined with "human love" is merely an appetite that, when satisfied, discards the other person like a lemon sucked dry (*Lectures on Ethics* 1997, Ak 27:384); the Marquis de Sade, whose inventory of acrobatic and monstrous sexual feats in *120 Days of Sodom* (c. 1785) proclaims that "anything goes," and who died in the Charenton insane asylum; G. W. F. Hegel, who, wielding dialectical logic in "On Love" (1797–1998), claimed that during sex (only during good sex?) "consciousness of a separate self disappears, and all distinction between the lovers is annulled" (p. 307); the Danish Christian-existentialist philosopher Søren Kierkegaard, whose brilliant "Diary of a Seducer" and portrayal of the aesthetic/sensual and ethical stages of life in *Either/Or* (1843) began the decade-long analysis of his broken engagement with his beloved Regine Olsen; a German fan of Kant, Arthur Schopenhauer (1788–1860), whose nineteenth-century metaphysics, philosophy of mind, and deification of the reproductive function of sexuality in *World as Will and Representation* uncannily anticipated both Charles Darwin and Sigmund Freud; Karl Marx and Friedrich Engels, whose *Communist Manifesto* (1848) equated being a prostitute and being a bourgeois wife, an idea far from dead among contemporary feminist scholars; John Stuart Mill, the author of the definitive feminist treatise *Subjection of Women*, who employed, in *On Liberty* (1859), his liberal utilitarianism to exonerate Mormon polygyny and pimps or brothels; and, closer to the fin-de-siècle, a German fan of Schopenhauer, Friedrich Nietzsche, who, with myriad scattered, sharp aphorisms about the sexes to his credit, still failed to negotiate benignly his crush on the vamp Lou Salomé and ended up dying in an insane asylum.

After Plato and Augustine, philosophical deliberation about sex became less urgent. With the exception of the thorough Thomas and the obsessed Sade, those mentioned above did most of their philosophy in epistemology, ontology, ethics, economics, and political theory, writing only sporadically on sexuality. The twentieth century, however, witnessed an outpouring of candid, sometimes shocking, inquiries into human sexuality. First was Sigmund Freud's *Three Essays on the Theory of Sexuality* (1905), which audaciously challenged myths about childhood sexual innocence and postulated that human sexual nature was polymorphously perverse. Freud's legacy includes the maverick psychoanalyst Jacques Lacan (1901–1981), who explored sex and language, and Lacan's Slovenian student Slavoy Žižek, who has explored nearly everything, from the role of power in human sexuality to cultural variations in the technology of toilets. Later came Bertrand Russell's *Marriage and Morals* (1929), which combined a prescient and formidable feminism with a well-reasoned critique of marital sexual fidelity. *Marriage and Morals*, called a "lecherous" book by some, cost Russell an appointment at the City University of New York. Then, during the thick of World War II (1943), Jean-Paul Sartre's *L'être et le néant* was published. Sartre

unabashedly exposed the "bad faith" of the woman who allows an unwelcome male hand to remain on her knee without so much as a mild squawk. In sexual interactions, for Sartre, we always desire to capture the freedom of the other. That endeavor, however, is doomed to failure; consequently, he argued, sexual relations reduce to masochism or sadism.

Soon afterwards appeared *Le deuxième sexe* (*The Second Sex*) by Simone de Beauvoir, Sartre's long-standing companion, with its primordial yet fertile feminist accounts of love, sex, and gender: "One is not born, but rather becomes, a woman" (p. 267). Beauvoir's "Must We Burn Sade?" helped garner for the Divine Marquis a persisting scholarly interest. Coming before and after Sartre and Beauvoir were some social philosophers—Wilhelm Reich (1897–1957), Herbert Marcuse (1898–1979), Erich Fromm (1900–1980), and Norman O. Brown (1913–2002)—who tried to solder an alliance between Freud's psychology and Marx's humanist economics in the name of liberating sexuality from oppressive Victorian morality and twentieth-century political tyranny. (Marcuse's *Eros and Civilization* [1955] is a worthy successor to Freud's 1930 *Civilization and Its Discontents*.) Outside philosophy, Alfred Kinsey and his associates at Indiana University stirred up a hornet's nest by investigating in the late 1940s the extent of homosexual and other atypical sexual behaviors in America.

More recently, New York University philosopher Thomas Nagel domesticated Sartrean insights and fashioned from them, in "Sexual Perversion" (1969), an H. P. Gricean theory of psychologically natural human sexuality. It is routinely acknowledged that this essay inaugurated contemporary philosophy of sex. It was followed almost immediately by a swarm of sophisticated discussions and rebuttals that also boosted the field, including essays by Sara Ruddick, Robert C. Solomon, Janice Moulton, Jerome Shaffer, Robert Gray, and Alan Goldman. In his wide-ranging and erudite *Sexual Desire: A Moral Philosophy of the Erotic* (1986), politically conservative British philosopher Roger Scruton rehabilitated nearly everything traditional, from sexual fidelity in marriage to Rousseau's condemnation of the solitary vice and, in an already sexual-orientation sensitive climate, Scruton fearlessly raised doubts about homosexuality. In *Sex and Reason* (1992), law professor and Judge Richard Posner expounded a no-nonsense, pragmatic/utilitarian ethical and legal philosophy of sex, and articulated what we should expect sexually from *homo economicus* (e.g., male pederasty tends to increase in locales in which there is a relative scarcity of women).

Another law professor and political philosopher, Catharine MacKinnon, after her early innovative writings on sexual harassment, dramatically escalated (along with Andrea Dworkin) the feminist battle against sexism. In *Feminism Unmodified: Discourses on Life and Law* (1987) and *Toward a Feminist Theory of the State* (1989), she argued that women's consent to sex in patriarchy is chimerical, implying that all heterosexual intercourse is rape. A third philosopher with legal training, John Finnis, joined by other New Natural Lawyers and the Catholic theologian Germain Grisez, overhauled Thomistic philosophy of sex. Finnis defended, in the *Notre Dame Law Review* (1994), the crucial but, for many critics, dubious moral distinction between the permitted coital acts of a sterile heterosexual couple and the prohibited sexual acts of a lesbian or gay couple.

This distinction in Catholic ethics has affinities with another one, well worth contemplating, between (illicit) heterosexual coitus in which procreative potential is deliberately impeded by contraceptive devices and (licit) intercourse that is unlikely to be procreative because the couple has deliberately restricted engaging in the act to the infertile period in the wife's cycle (see Anscombe 1976, Wojtyła 1981, and Noonan 1986). The unconventional feminist Camille Paglia frankly told university women, in *Sex, Art, and American Culture* (1992), that if they go to fraternity parties and willingly drink excessively, it is partially their own foolish fault if their panties come down on a billiard table—thereby adding the cool voice of a humanist public intellectual to the often tempestuous debate in philosophical and legal circles about date and acquaintance rape.

Of special significance is the French Renaissance man Michel Foucault, who caused a thunderstorm among philosophers, historians, and social theorists of sex with the three volumes of his *Histoire de la sexualité* (1976–1984). Foucault sparked "genealogical" studies informed by the heuristic idea that not only are patterns of sexual desire and behavior socially engineered but also that the very concepts of our sexual discourse are "socially constructed." (He was in part reacting against the discourse of "natural" sexuality found in Reich and Marcuse.) Foucault influenced feminism, gender studies, queer theory, and the debate about the resemblance and continuity, or lack of them, between ancient same-sex relationships and their contemporary counterparts. (These questions are pursued in the collections edited by Edward Stein, by Nussbaum and Sihvola, and by David Halperin and his colleagues. This venture is sharply criticized by Paglia in "Junk Bonds and Corporate Raiders," in

her *Sex, Art, and American Culture.*) One contested issue is whether homosexuality as a sexual orientation was first recognized in 1869 when the Magyar sexologist Károly Mária Benkert coined a word for it ("homosexuality"), a word unknown to the ancients, who could very well have invented it had they deemed that doing so was philosophically, socially, or medically meaningful. It was late nineteenth-century European sexology that detected value in picking out and labeling a class of persons as homosexual.

CONCEPTUAL ANALYSIS

Related to the question of the "birth" of the modern homosexual, there is the analytic task of defining "sexual orientation" and each of the various sexual orientations. It seems that neither sexual orientation in general nor any specific sexual orientation can be adequately understood in terms solely of behavior. Because there are many reasons and motives to engage in sex, and many intentions and desires are involved, outward behavior might not reveal anything interesting about a person's core sexual psychology (orientation). A closeted gay male who engages in coitus with his wife to impregnate her does not thereby make or declare himself heterosexual; the frustrated straight male in prison who reluctantly succumbs to mutual masturbation does not thereby become gay; the prostitute who participates in sexual acts with both the male and the female of a couple who has hired her for an evening is not thereby bisexual; an abstinent person who engages in no sexual activity, not even self-abuse, does not necessarily have an "asexual" orientation but may be heterosexual, homosexual, bisexual, or polysexual. What the examples suggest is that preferred sexual activity, or activity that one would engage in purely out of desire and for no other reason, is a better indicator of sexual nature than behavior, which might be induced by nonsexual motives. Counterfactual questions such as "What would you prefer to do, given your druthers and all real-life obstacles eliminated?" as well as straightforward questions about sexual fantasies, perhaps those entertained during the solitary vice, and about what a person finds arousing in anticipation (even if not during the anticipated act itself) are more revealing of sexual psychology than an accounting of acts performed. Orientation, then, is largely understood in terms of what sexual desire attaches to and the sources of sexual pleasure. But what are sexual desire and sexual pleasure?

Among the central concepts in the philosophy of sex are sexual desire, sexual activity, sexual pleasure, sexual perversion, sexual arousal, and sexual satisfaction. Philosophers have worked on these concepts, striving to provide clear analyses of them as well as illumination about the role and significance of sexual desire, and the others, in human life. Analytic philosophy of sex attempts to indicate, for example, how sexual desire is different from other kinds of desires; to explain how acts can be specifically sexual instead of some other kind of act; to discover what it is that makes a feeling or sensation one of sexual pleasure; and to determine what meaning, if any, can be given to the idea that some sexual acts (but not others) are unnatural or perverted. In the process of analyzing these central concepts, philosophers of sex have discerned or proposed that understanding any one of them might require understanding some other central concept. A chief case is sexual activity, which might be defined as activity that aims to satisfy sexual desire, or is motivated by sexual desire, or is intended to produce (or does produce) sexual pleasure. These candidate analyses seem to be on the right track, yet they all suffer from the same apparent defect.

The principal problem is that if sexual activity is defined as activity that is motivated by sexual desire or is intended to yield sexual pleasure (which works well for many paradigmatic instances), there are activities that are presumably sexual, are not uncommon, and yet are not captured by these or similarly fashioned definitions. Acts performed by a prostitute may produce pleasure for the paying client or are done by him to satisfy his sexual desires, but these definitions cannot explain why the acts of the prostitute (e.g., fellatio or coitus) are still sexual for her, assuming, which is plausible, that she participates for payment and not out of sexual desire for her client and that she derives no sexual pleasure from what she does or has done to her. The problem is not only that, given this type of analysis, the single act that the client and the prostitute perform together might be a sexual act for the client but not for the prostitute. The conundrum, more specifically, is that the feature (if any) in virtue of which her contribution to the act is sexual is not clear. It might be proposed that sexual activity be analyzed, instead, in terms of the involvement of salient sexual body parts—say, the genitals. If so, acts performed by a prostitute are sexual when and because her genitals are involved. But "involves the genitals" (or any other body part) seems neither necessary nor sufficient for an act to be sexual: some sexual acts are not genital (rubbing the breasts) and some acts that involve the genitals are not sexual (a gynecological exam). Perhaps "sexual body part" should be analyzed in terms of "sexual activity" (a body part is sexual exactly on those occasions when it is employed in a sexual act) rather than the other way around.

Analytic philosophy also tackles "derivative" sexual concepts, a large group of concepts (or phenomena) that include reference to sexuality. Derivative concepts that philosophers have attended to include adultery, jealousy, sexual harassment, casual sex, promiscuity, seduction, flirting, cybersex, and sexual fantasy.

Intriguing questions can be asked about adultery, in addition to standard moral questions, which are also explored by philosophers of sex. Does a nonmarried person who engages in sexual activity with a married person commit adultery? (In the law, the answer varies by jurisdiction.) Does a person commit adultery if she believes falsely that her spouse is deceased? Is adultery altogether a physical act or could desires and fantasies be not only adulterous in spirit but adultery itself? (See Matthew 5:28.) Some claim that in vitro fertilization, if carried out with donor (nonspousal) sperm, constitutes adultery. Can such a judgment be sustained? Casual sex and promiscuity, too, suggest questions beyond the ethical: For how many partners over what period of time is the judgment "promiscuous" accurate? Can one engage in casual sex with one's spouse? (Theologians argue that marital sex can be unchaste. Perhaps in this way it can be casual.) What distinguishes promiscuity from casual sex? Are there moral or perfectionist criticisms that can be made about casual sex and promiscuity other than condemning them for the absence of love, marriage, or commitment? There are difficulties in defining "sexual harassment"—what counts as a sexual advance, an improper sexual comment, or hostile work environment?—and explaining what is wrong with it, when (if) it is wrong—as sexual discrimination, immoral sexual conduct, or misuse of power, authority, or institutional position?

Seduction poses the analytic problem of carving out distinct logical space between rape, on the one side, and completely consensual sexual activity, on the other, and hence may pose novel ethical questions beyond those that apply to the other cases. But the moral issues concern not only the perpetrator of seduction. What about the person who welcomes and encourages being seduced, perhaps to be reassured of attractiveness or power? Sexual fantasy is a ubiquitous human phenomenon that suggests provocative questions: Does sexually fantasizing about a person "use" that person in any robust sense? Is it possible to criticize morally a person who fantasizes sexually about a third party during sexual activity with a partner, while not objecting to sexual fantasy *tout court*? What is the relationship between fantasy and sexual desire: Do we fantasize about something (or someone) because we desire it or do we desire it because we have fantasies about it? Jealousy, because of its intentional structure (its dependence upon beliefs), might arise in response to a fantasy. Is the fault with sexual jealousy (if it is faulty) exhausted by its being caused by a false belief or one arrived at negligently? Or can sexual jealousy be deplored because it frequently betrays a wrongful attitude of owning another person?

Cybersex highlights the intentionality of sexuality, because cybersexual arousal depends exquisitely on beliefs about unseen persons; it forces us to ask why another person's body is apparently so important—or not so important, after all—in sexual experiences, which also raises questions about masturbation; and cybersex makes us ponder whether some sexual activity—and therefore, for example, some adultery—may involve no physical touching in the ordinary sense (as does telephone sex). Similarly, flirting might be a sexual activity that falls somewhere between faithfulness and infidelity. To which is it closer? Does this depend on with whom one flirts, why, or the extent to which one is tempted or willing to turn flirting into physical contact? Flirting is interesting also because it is occasionally misread, conveying to some optimistic or deluded recipients an explicit invitation to engage in sex instead of registering merely as playful or teasing. As a result, flirting might sometimes precipitate date or acquaintance rape.

The derivative concept "rape" has long presented special problems. One controversial matter is whether rape should be defined in terms of the absence of consent or the presence of force. This has implications for how the occurrence of rape is established in a court of law. The choice is difficult: A force definition of rape might place too much emphasis on whether or to what extent a woman resists, which many see as irrelevant. A consent criterion implies that tough issues about mens rea become important: Did the accused believe that the woman had consented, even if she didn't; is the accused liable for something he might or should have believed but did not believe (that is, that consent was absent)? The difference between a force and a nonconsent criterion may be illustrated with acquaintance rape. A force criterion tends not to classify such acts as rape, whereas proponents of a nonconsent criterion argue that rape includes all nonforcible yet nonconsensual sex (see McGregor 2005). Further, like prostitution, rape seems to provide a counterexample to the analytic proposal that sexual activity be understood in terms of sexual desire or sexual pleasure.

Perhaps because both prostitution and rape are activities that involve coercion or are not engaged in (fully) voluntarily, they resist being characterized as sexual acts in terms of desire or pleasure. Indeed, it has been argued, on various grounds, that rape is not a sexual activity at all. (Maybe this point applies to prostitution as well.) If a woman, a virgin, is raped and does not thereby, automatically, lose her virginal sexual status, she has not taken part in a sexual activity, at least not one that was sexual for her. But the derivative concept "virginity" and similar notions—abstinence, chastity, celibacy—require careful analysis in their own right. Are they merely a matter of behavior or anatomical characteristics or does state of mind play a role, and how? Another issue concerns the extension of "rape," which accentuates problems in spelling out the meaning of coercion or consent and in deciding why and when coerced or nonconsensual sex is wrong. Suppose a man badgers his wife for sex until she acquiesces, and they engage in sexual activity even though she much prefers not to. Has she been coerced and therefore raped, and is this the reason the act is morally stained? Perhaps badgering does not amount to coercion, but it is still morally suspect. By contrast, some would say that even if the badgering coerces her into sex, it is not especially morally objectionable. Or suppose a woman hints to her husband, "No sex until you buy me that fur coat."

HUMANS AND OTHER ANIMALS

One debate in the philosophy of sex concerns the relevance of animal sexuality for understanding and judging human sexuality. Some philosophers, for example Thomas Aquinas in *Summa contra gentiles* (chap. 122, sec. 6), argue from observations of animal sexual behavior to the nature of human sexuality and draw ethically conservative conclusions. These philosophers emphasize (a subset of) that which is common between humans and animals. For example, many animals engage in sexual relations only to reproduce and that, too, is what is significant about human sexuality. Then there are philosophers—those who are sympathetic to sociobiology or evolutionary psychology are among them—who similarly stress what is common to animals and humans, yet draw ethically liberal conclusions. We are fundamentally animals and that fact should not be ignored or minimized; the robust sexuality that is due to our animal nature is suppressed at our peril.

What may distinguish the first group of philosophers from the second is the animal species invoked in drawing conclusions about humans. If one selects as the argument's observational basis monogamous birds (swans) and mammals (wolves), different conclusions will emerge than if one selects more sexually adventuresome species (dogs, the bonobo). The question—Which is the right animal model?—is murky, although similarity of DNA, testicle size, and other traits are potentially useful links. (Why even assume that the same animal model will be the right one for both human males and females?) Regardless, we must avoid the circularity of arguing that a species is the right model because these creatures are remarkably like humans—unless our methodology is a sophisticated "reflective equilibrium." Further, once we select some animal species from which to argue, we must take the "bad" with the "good": The aggression, dominance, promiscuity, and oddness (e.g., urolagnia in some llamas) of animal sexuality, along with its attractive features, have to be extrapolated to humans as well. Against both the conservative and the liberal who argue from animal sexuality to ethics, it can be protested that doing so commits the naturalistic fallacy. What cannot be excluded is that comprehensively studying animals can tell us something about human nature. It is a dangerous leap from there to ethics.

Some philosophers, by contrast, even though acknowledging that humans, as embodied, are undeniably in part animals, perceive sharp discontinuities or differences of kind, not degree, between animals and humans. There are physiological differences such as concealed ovulation and the absence of oestrus in human females that have extensive implications for sexual psychology and behavior. But more striking is the human cerebral cortex and hence cognitive differences between humans and animals. This view can also be taken in an ethically conservative or liberal direction. Conservatives—Scruton, for one, and many theologians—say that humans have mind or soul, something that lifts us above animals, so that even if we have animal urges, we can and should transcend them. Behaving in a humanly civilized fashion is to be accomplished by virtue of our spirit and for the sake of our spirit. But the discontinuity is also compatible with liberal sexual ethics. Nagel, in formulating his theory of psychologically natural human sexuality, emphasizes the differences between animal and human sexuality that result from the nearly unique faculties of the human mind, primarily intentionality and self-consciousness (which also figure prominently in Scruton's philosophy of sex). Yet Nagel comfortably embraces Millian liberal sexual ethics.

Further, for social constructionists animal and human sexuality are of course different, and nothing

much is to be gained by comparing them. Human sexuality and sexual discourse vary as much as human culture varies, whereas animals have (by and large) no culture or language that might construct their sexuality or their (nonexistent) conceptions of it. The sociobiologists and their philosophical sympathizers retort: Yes, society constructs much of human sexuality, but human sexuality (to use E. O. Wilson's metaphors) is a twig bent at birth; it is on a leash, tied ultimately to a biological post, a substrate upon which society can work—and which it requires in order to work—its constructionist miracles. As suggested by the mixed results of the medical management of intersex conditions (neonates of ambiguous sex), the social cannot make everyone male, female, straight, or gay. How much of human sexual nature is due to animal biology, and how much to culture, is as difficult to resolve as analogous nature-nurture quandaries about the contribution of race or biological sex on various skills and personality traits. Often these disputes are replaced by (prematurely, perhaps, but not altogether baselessly) brute political machinations, à la Plato's Thrasymachus in the *Republic*.

One reason for looking at animal sexuality is that this knowledge may serve as a guide to what human sexuality would be like were it not for social interference, that is, in the absence of all cultural influence (although, unlike Freud, social constructionists do not speak of the cultural as an "interference" but as necessarily constitutive). It does not strain the imagination to conceive of cultureless animals as expressing pure state-of-nature sexuality. If humans arrange their sexuality consistently with what is seen among animals (by peeling back various social influences), we can have some faith that we are not too far away from humanly natural, healthy, satisfying sexuality. Such thinking builds on an absorbing and plausible thought, that animal sexuality cannot in any way be unnatural or "perverted." If nothing about morality can be learned from animal sexuality, at least we can get glimpse of normality. There is probably too much Rousseauvian utopianism in this thinking, and of course such a view remains vulnerable to the hitch of which animal model confers the best insight into "normal" humanity. Alternatively, well-founded speculation about the sexuality of prelapsarian Adam and Eve might, for some theologians, supply that information. The Garden of Eden is their Hobbesian state of nature.

SEXUAL PERVERSION

As far as popular culture and ordinary folk are concerned, the terms "[sexual] perversion," "[sexually] perverted," and "[sexual] pervert" are not problematic, even if they might not always be in good taste and cause distress to those singled out. "Sexual" is bracketed because "pervert" in ordinary talk implies that the domain of discourse is the sexual. (At least, that is the default position.) By contrast, some philosophers, psychologists, and other academics have argued that "sexual perversion" is outmoded, ontologically groundless, confused, offensive, unscientific, not applicable to anything in human sexual behavior, and hence happily dispensable. Despite the counsel of philosophers and other experts that "perversion" be extirpated from the language, ordinary people use it unflinchingly, as does the Religious Right. The American Psychiatric Association (APA) no longer officially uses "perversion" to refer to sexual disorders but has, since 1980, opted for the clinical "paraphilia," even if an ordinary person's list of perversions is nearly identical to the paraphilias listed in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*). ("Paraphilia" is not a total improvement. It is an unlikely group that includes "Philadelphia," "philanthropy," "philosophy"—and "paraphilia.")

The fundamental problem about sexual perversion is distinguishing natural from unnatural sexuality. In carrying out this task, it is necessary to explain not only how certain sexual behaviors (desires, preferences) are perverted but also how they are sexual to begin with. For example, if being potentially procreative is the feature that defines sexual activity, then being nonprocreative cannot be a mark of the sexually perverted, because whatever is not procreative is not sexual. (The acts might still be "nonsexually [or fill in the blank] perverted.") Or if sexual activity is defined as activity that tries to satisfy sexual desire and sexual desire is defined, in turn, as desire for physical contact with another human being, the perversions cannot be sexual, because they typically do not involve desire for that contact: consider the wide variety of fetish objects that excite men. Some would call it special pleading or adhockery, whereas others would see it as a stroke of genius, to say that the fetishist does desire physical contact with a person, unconsciously, and achieves that in a psychologically safe way by substituting the fetish object.

Philosophers and psychologists have tried, with unclear success, to formulate theories about sexual perversion. An obvious contender, that only potentially reproductive sexual acts—acts that are reproductive in their anatomical and physiological forms—are natural, and all others perverted, has seemed plausible to many thinkers (Catholics and some evolutionists, mostly) but

implausible to others. Certainly, being nonprocreative is a property that many (preanalytically) perverted sexual acts share: zoophilia, cross-dressing, exhibitionism, voyeurism, klismaphilia, necrophilia, urolagnia, sadomasochism.

But analyzing perversion as nonprocreative sexuality is not straightforward. Some nonprocreative sexual acts are not especially, or at all, perverted: masturbation (solitary or mutual) and oral sex to orgasm. And some purportedly perverted acts (cross-dressing, light sadomasochism) for some people often or regularly culminate in heterosexual intercourse, as if functioning as foreplay. Also note that both vertical (parent-child) and horizontal (sibling-sibling) incest can be procreative, yet many have thought them considerably unnatural (or maybe only repulsive). The sexual practices that are supposed to be subsumed under the label "perversion" or "paraphilia" are extraordinarily diverse, other than being nonprocreative, so finding common, essential features may be doomed—a reason to dispatch the concept. We could still investigate, without using "perversion," behaviors that are unusual, bizarre, harmful, or are done compulsively or exclusively, in preference to every other sexual activity (which category may well include a narrow interest in heterosexual coitus). That "unusual," "bizarre," and "harmful" are to a greater or lesser extent evaluative or culturally bound is why these features of sexual acts cannot be used to develop an objective, scientific, universally sound theory of sexual perversion. Social constructionists applaud this result.

Another question about sexual perversion has to do with its morality. The Roman Catholic position, that what is perverted is for that reason sinful, has not won over many secular adherents. "Premodern" philosophy of sex, which derives from the older Plato, Augustine, and Aquinas, understands sexual perversion teleologically as behavior that is incompatible with the (perhaps divinely ordained) species design. Premoderns frequently add that in virtue of this deviation, deliberately performed sexually perverted acts are immoral. But perhaps not every deviation is wrong. Mutual masturbation, cunnilingus, and fellatio, which in themselves are nonprocreative and hence unnatural, might be permissible when they function as preparation for heterosexual marital coitus. "Modern" philosophy of sex dates from the late nineteenth century and the rise of scientific sexology (e.g., Iwan Bloch, Magnus Hirschfeld, Richard Krafft-Ebing, Havelock Ellis, Freud). Some modern philosophers of sex retain the biological, teleological account of perversion, whereas others (Freud, Nagel) replace that with a more sophisticated psychological account.

What the two branches of modern philosophy of sex share is a refusal to judge perverted sex immoral merely because it is perverted. Many modern philosophers of sex have reached, instead, for the evaluation "psychologically unhealthy." It is worthwhile to think of premodern judgments of sinfulness as superseded by modern judgments of sickness, as social authority residing over sexual perversion passed from the clergy and organized religion to the physician and biomedical science. The fate of homosexuality illustrates this progression, from being condemned as sin by all Western religions to being deprecated as sickness (although excused, in keeping with the medical model) by most Western psychology and psychiatry through the mid-twentieth century. But in 1973, the APA removed homosexuality from its list of mental disorders in *DSM*, thereby helping to usher in "postmodern" philosophy of sex, according to which no nonharmful, consensual sexual behaviors are perversions, sinful, or sick, but alternative sexual choices. The APA has not gone completely postmodern. It still classifies some innocuous sexual practices (fetishism, transvestism) as sexual mental disorders.

The American Psychiatric Association distinguishes between sexual dysfunctions and the paraphilias, which, even though they involve unusual or bizarre sexual desires or acts, do not necessarily involve inadequate functioning of the sexual organs. When homosexuality was still a mental sexual disorder, there was no doubt that gay men could sport firm erections and did not suffer from ejaculatory problems merely in virtue of their orientation. In addition to premature ejaculation, an inability to achieve or maintain an erection, insufficient lubrication, and pain during coitus, the APA includes as a dysfunction "Hypoactive sexual desire disorder," a deficit or absence of sexual desire that causes psychic distress or interpersonal (e.g., marital) problems (*DSM-IV*, sec. 302.71). Critics have pointed out that the clinical judgments that a person has too little sexual interest and is bothered too much by a perceived lack of desire are routinely influenced by all manner of social factors that seem irrelevant to a diagnosis of mental disorder. The *DSM* also lists a more extreme variant, "Sexual Aversion Disorder" (sec. 302.79), but (asymmetrically) contains no "hyperactive sexual desire disorder." The APA did, however, briefly flirt with Patrick Carnes's innovation, "sexual addiction" (a type of obsessive-compulsive promiscuity), as a sexual mental disorder, which was included only in the revised version of *DSM-III* (1987). Speaking of naturally pleasurable sexual activity as "addic-

tive" is highly disputable, as is whether promiscuity (such as homosexuality) is sinful, sick, or a mere variation in human sexuality.

SEXUAL USE

Being unnatural is of course not the only way sexual activity might go astray morally. In the Kantian tradition, the central way that sexual activity is morally wrong is when one person uses another person sexually, treating the other as a means or object, thereby violating the second formulation of the categorical imperative. Coercing another person, as in rape or quid pro quo sexual harassment (boss to employee: "Have sex with me or you're fired"), or deceiving someone in order to obtain sexual relations (an identical twin sliding into the bed of his brother's wife) are frequently cited cases of treating another person as a means. On a Kantian view, and on some utilitarian views (such as Mill's), it is necessary for the moral permissibility of a sexual event that all parties furnish free and informed consent. Other instances of possible use are difficult to settle; even among confirmed Kantians, exactly what treating another person as a means or an object amounts to has long been disputed. One disagreement between conservative and liberal Kantians is over whether an adult's consent is sufficient (ceteris paribus) for the morality of sexual activity. Kant answered "no," arguing that sexual activity avoided mere use in, and only in, marriage, or that marriage made mutual sexual use permissible. (How to interpret Kant is an issue for Kant scholarship. See *Lectures on Ethics*, Ak 27:388.)

In this respect many conservative Kantians, such as Karol Wojtyła (Pope John Paul II, 1920–2005), have followed Kant, insisting that mutual consent alone neither eliminates nor blesses the mutual use in sexual relations that must occur if the persons are not married (although some conservatives would be satisfied were sexual relations confined to a genuinely committed even if nonmarital relationship). In any event, Kant and the conservative Kantians need to explain—a challenging task—how the additional ingredient, marriage or commitment, changes sexual activity from mere mutual use to something morally permissible, and why only commitment or marriage and nothing else (say, consent) has the ability to do this. For liberal Kantians, mutual consent is powerful enough by itself to make sexual acts permissible in the absence of marriage. The presence of consent, they argue, satisfies the demand of the second formulation of the categorical imperative for the reciprocal acknowledgment by each person of the rational autonomy (the humanity) of the other. In virtue of consent, much sex is permissible that is condemned morally by Kantian and other conservatives: same-sex sexual acts, group sex, casual sex (say, between strangers), even adultery if all parties consent. Consent is sufficient only ceteris paribus for the liberal Kantian and the Millian utilitarian because third parties might be harmed or have their legitimate interests disregarded by the consensual sex of others (as often happens in adultery). For some conservative Kantians, mutual consent to use each other not only is not sufficient, but makes for an especially morally corrupt situation, for they take, as did Kant, the often slighted part of the second formulation seriously: one may not treat the humanity in one's own person merely as a means. This is what one does to oneself—willingly makes an object of oneself—when consenting to be sexually used by another person, even if that use is mutual. It is an interesting question how it might be decided whether mutual consent cancels or compounds the moral faults of mere use.

The opposite of sexual objectification is sexual personification, which occurs when, to mention the key instance, a person or a couple gives a name to an erotic body part. (Christening the genitals is an important theme in D. H. Lawrence's *Lady Chatterley's Lover*.) An example of nonsexual personification might be worshiping an idol, a golden lamb, treating a mere material symbol of the Almighty as if it were the Almighty. Personification can be understood as raising something's ontological status or treating it as if had a higher status. This is what happens when a couple gives proper names to their genitals, treating them as persons. In objectification, by contrast, one person reduces (or attempts to reduce) the ontological status of another. If a person manipulates another so that a goal of the first person is thereby attained, the first has used the second, has treated him or her as a mere material object, in that the second's personhood-defining feature, rational autonomy, has been minimized or ignored. One person is acting toward another as if the latter were no more ontologically elegant than an inanimate thing or a subhuman animal. In sexual objectification, even if there is no coercion or deception, a person is treated as a usable object fundamentally capable of (only) satisfying another's sexual desire. It is often claimed, by both Kantian conservatives and many feminist philosophers, that this is exactly what is morally wrong with prostitution and pornography: women are not respected fully as the persons they are but are seen and treated only or primarily as consumable and fungible providers of sexual pleasure, even when they consent to participate. Some theorists go further, claiming that these considerations apply as well to the institution of heterosexual marriage.

PORNOGRAPHY AND PROSTITUTION

Arguments about consent occur when philosophers, legal theorists, political activists, and women and men sex-industry workers discuss pornography and prostitution. If consent is present in a given instance of prostitution (which can be defined, but not unproblematically, as exchanging sexual activity for compensation), if neither the client nor the provider of sexual services is subject to coercion or is deceived, or if those hired to perform sexual acts in front of a camera (a type of prostitution) in the production of pornography (variously definable, notoriously with difficulty) have, similarly, freely and with reasonably full and relevant information agreed to do so, the issue still arises whether their consent is sufficient. In this debate, one side (the liberal, the libertarian, perhaps the Milton Friedman capitalist) points out that if consent is sufficient for other kinds of paid labor, from slinging slop in a fast food pub and collecting garbage to executing proctological examinations and fighting in a volunteer army, there can be no objection to a person's engaging in sex for payment. Anything is fodder for the market or, at least, nothing differentiates selling sexual services and performing other tasks that some people, but not all, find too repugnant or risky to undertake even for substantial financial compensation. The other side (some conservative theologians, Marxists, and feminists) insists, however, that sexuality is "different," that it does or should involve a quality of intimacy that is undermined by its being bought and sold, or that it is demeaning when sexuality is the means of making a living, or that sexuality is metaphysically or anthropologically too crucial an aspect of human personality or identity to be commodified. Doing so entails an immeasurable cheapening of humanity. Whether these claims about how sexuality differs from other aspects of human life are culturally bound (hence not so compelling?) or are deep, sustainable philosophical truths about the human person is unclear. Note that if they are overblown, exaggerating the significance of sexuality in an overall picture of the human person, it might be more difficult to explain why rape is an especially grievous harm (see Murphy 1994).

However, that women sex-industry workers participate consensually is debatable. There are various reasons, often advanced by feminists and Marxists, for doubting that the consent of the women who make pornography or sell sexual services is genuine (see, e.g., MacKinnon's *Only Words*). They might have been indoctrinated to devalue themselves and their sexuality or have been as children victims of sexual abuse, and in either case, they may be exceptionally vulnerable to being manipulated into prostitution and the production of pornography. Further, to the extent that women who participate in these activities come from the lower economic levels of society, the lure of making decent money despite lacking education or vocational training can be coercive, if their alternatives are even more dismal. The possibility of compulsion may be greater when the women, in addition to being relatively impoverished, are members of a disparaged ethnic minority or have dependent children. Their dire need creates a situation in which being offered money for sexual activity is coercive, even if engaging in those sexual events seems to them, at the moment, a small sacrifice of their sexual integrity.

It might also be argued that because women are willing to sell sexual services in either prostitution or pornography, this is by itself evidence that something is amiss in their rational autonomy; doing such things is not what someone "in her right mind" would choose to do. Several responses to this account of the plight of women sex-industry workers have been advanced. One rebuttal is that it overstates the victimization of women and underestimates their strength and resourcefulness. Another is that citing financial need as coercive may imply too much. Most people who sell their labor have financial needs, are in no position to refuse to work, and they, too, would have to be described as coerced. Finally, there are women who relish the opportunity to make good money in the sex industry and would not describe their situation as one in which they are pressured into doing something they prefer not to do.

CONCLUSION

Our personal understandings of the nature of sexuality and its significance in our lives, public discussion of ethical, religious, and social issues, and technical matters about sex that arise in medicine, social science, and the law—all these can profit from philosophical study. Students who take courses in the philosophy of sex are exposed to material they are unlikely to encounter elsewhere, material that gives them an opportunity to scrutinize their beliefs about sexuality and habitual behaviors. The law benefits from the philosophical analysis of concepts such as rape, harassment, and consent; theology is in a position to learn from the elaboration of theories of natural human sexuality and the examination of the conceptual connection between the goodness of the natural and the goodness of human actions; social scientific surveys of the frequency of sexual activity (by age, education, ethnicity, and other parameters) and the extent of non-heterosexual sexual orientations depend on analyses of

"sexual activity," "sexual desire," and "sexual preference" and effective ways of identifying and counting or measuring them; the pronouncements of psychiatry and medicine on sexual health, both physical and mental, can be (and have been) improved by the deliberations of philosophers who investigate the concepts of sexual perversion and mental illness. The philosophy of sex has proven that it is no idle enterprise.

See also Abelard, Peter; Albert the Great; Affirmative Action; Aristotle; Augustine, St.; Beauvoir, Simone de; Darwin, Charles Robert; Descartes, René; Engels, Friedrich; Feminist Social and Political Philosophy; Foucault, Michel; Freud, Sigmund; Grice, Herbert Paul; Hegel, Georg Wilhelm Friedrich; Heterosexism; Hobbes, Thomas; Hume, David; Kant, Immanuel; Kierkegaard, Søren Aabye; Lacan, Jacques; Mandeville, Bernard; Marx, Karl; Mill, John Stuart; Montaigne, Michel Eyquem de; Nagel, Thomas; Nietzsche, Friedrich; Nussbaum, Martha; Pelagius and Pelagianism; Plato; Reich, Wilhelm; Renaissance; Rousseau, Jean-Jacques; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Schopenhauer, Arthur; Sexism; Thomas Aquinas, St.; Thomism; Utilitarianism; Wilson, Edward O.

Bibliography

Anscombe, G. E. M. "Contraception and Chastity." *The Human World*, no. 7 (1972): 9–30. Reprinted and revised in *Ethics and Population*, edited by Michael D. Bayles. Cambridge, MA: Schenkman, 1976.

Baker, Robert B., and Frederick A. Elliston, eds. *Philosophy and Sex*. 3rd ed., edited by Robert B. Baker, Kathleen J. Wininger, and Frederick A. Elliston. Amherst, NY: Prometheus, 1998.

Beauvoir, Simone de. "Must We Burn Sade?" Translated by Annette Michelson. In *The Marquis de Sade: The 120 Days of Sodom and Other Writings*, edited by Austryn Wainhouse and Richard Seaver. New York: Grove Press, 1966.

Beauvoir, Simone de. *The Second Sex*. Translated by Howard M. Parshley. New York: Knopf, 1953.

Ben-Ze'ev, Aaron. *Love Online: Emotions on the Internet*. Cambridge, U.K.: Cambridge University Press, 2004.

Blackburn, Simon. *Lust: The Seven Deadly Sins*. New York: Oxford University Press/New York Public Library, 2004.

Brundage, James A. *Law, Sex, and Christian Society in Medieval Europe*. Chicago: University of Chicago Press, 1987.

Burgess-Jackson, Keith, ed. *A Most Detestable Crime: New Philosophical Essays on Rape*. New York: Oxford University Press, 1999.

Cornell, Drucilla, ed. *Feminism and Pornography*. Oxford, U.K.: Oxford University Press, 2000.

Corvino, John, ed. *Same Sex: Debating the Ethics, Science, and Culture of Homosexuality*. Lanham, MD: Rowman and Littlefield, 1997.

Devine, Philip E., and Wolf-Devine, Celia, eds. *Sex and Gender: A Spectrum of Views*. Belmont, CA.: Wadsworth, 2003.

Dworkin, Andrea. *Intercourse*. New York: Free Press, 1987.

Dworkin, Andrea. *Pornography: Men Possessing Women*. New York: Penguin, 1989.

Epstein, Louis M. *Sex Laws and Customs in Judaism*. New York: Bloch, 1948. Reprint, New York: Ktav, 1967.

Finnis, John M. "Law, Morality, and 'Sexual Orientation.'" *Notre Dame Law Review* 69 (5) (1994): 1049–1076.

Foucault, Michel. *The History of Sexuality*. Translated by Robert Hurley. 3 vols. New York: Vintage, 1988–1990.

Freud, Sigmund. *Three Essays on the Theory of Sexuality*. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud* Vol. 7, translated and edited by James Strachey. London: Hogarth Press, 1953–1974.

Gert, Bernard. "A Sex Caused Inconsistency in DSM-III-R: The Definition of Mental Disorder and the Definition of Paraphilias." *Journal of Medicine and Philosophy* 17 (2) (1992): 155–171.

Gilbert, Paul. *Human Relationships: A Philosophical Introduction*. Oxford, U.K.: Blackwell, 1991.

Giles, James. *The Nature of Sexual Desire*. Westport, CT: Praeger, 2004.

Gruen, Lori, and George F. Panichas, eds. *Sex, Morality, and the Law*. New York: Routledge, 1997.

Gudorf, Christine E. *Body, Sex, and Pleasure: Reconstructing Christian Sexual Ethics*. Cleveland: Pilgrim Press, 1994.

Halperin, David M., John J. Winkler, and Froma I. Zetlin, eds. *Before Sexuality: The Construction of Erotic Experience in the Ancient Greek World*. Princeton, NJ: Princeton University Press, 1990.

Halwani, Raja. *Virtuous Liaisons: Care, Love, Sex, and Virtue Ethics*. Chicago: Open Court, 2003.

Hegel, G. W. F. "Love." In *On Christianity: Early Theological Writings*, translated by Thomas M. Knox and Richard Kroner, 302–308. New York: Harper, 1961.

Hobbes, Thomas. "Human Nature, or the Fundamental Elements of Policy." In *The English Works of Thomas Hobbes of Malmesbury*. Vol. 4, edited by Sir William Molesworth. Aalen, Germany: Scientia Verlag, 1966.

Horowitz, Gad. *Repression: Basic and Surplus Repression in Psychoanalytic Theory; Freud, Reich, and Marcuse*. Toronto: University of Toronto Press, 1977.

Jackson, Stevi, and Sue Scott, eds. *Feminism and Sexuality: A Reader*. New York: Columbia University Press, 1996.

Kant, Immanuel. *Lectures on Ethics* [*Vorlesung*]. Translated by Peter Heath, edited by Peter Heath and J. B. Schneewind. Cambridge, U.K.: Cambridge University Press, 1997.

LeMoncheck, Linda, and Mane Hajdin. *Sexual Harassment: A Debate*. Lanham, MD: Rowman and Littlefield, 1997.

MacKinnon, Catharine A. *Feminism Unmodified: Discourses on Life and Law*. Cambridge, MA: Harvard University Press, 1987.

MacKinnon, Catharine A. *Toward a Feminist Theory of the State*. Cambridge, MA: Harvard University Press, 1989.

MacKinnon, Catharine A. *Only Words*. Cambridge, MA: Harvard University Press, 1993.

Mandeville, Bernard. *A Modest Defence of Publick Stews*. Edited by Richard I. Cook. Los Angeles: Augustan Reprint Society, 1973.

McGregor, Joan. *Is It Rape? On Acquaintance Rape and Taking Women's Consent Seriously*. Aldershot, U.K.: Ashgate, 2005.

Montaigne, Michel De. "Of the Power of the Imagination." In *The Complete Essays of Montaigne*. Translated by Donald M. Frame. Stanford, CA: Stanford University Press, 1958.

Murphy, Jeffrie G. "Some Ruminations on Women, Violence, and the Criminal Law." In *In Harm's Way: Essays in Honor of Joel Feinberg*, edited by Jules Coleman and Allen Buchanan. Cambridge, U.K.: Cambridge University Press, 1994.

Nagel, Thomas. "Sexual Perversion." *Journal of Philosophy* 66 (1) (1969): 5–17. Reprinted and revised in his *Mortal Questions*. Cambridge, U.K.: Cambridge University Press, 1979.

Noonan, John T. *Contraception: A History of Its Treatment by the Catholic Theologians and Canonists*, enlarged ed. Cambridge, MA: Harvard University Press, 1986.

Nussbaum, Martha C. *Sex and Social Justice*. New York: Oxford University Press, 1999.

Nussbaum, Martha C., and Juha Sihvola, eds. *The Sleep of Reason: Erotic Experience and Sexual Ethics in Ancient Greece and Rome*. Chicago: University of Chicago Press, 2002.

Pagels, Elaine. *Adam, Eve, and the Serpent*. New York: Vintage, 1988.

Paglia, Camille. *Sex, Art, and American Culture: Essays*. New York: Vintage, 1992.

Plato. *The Collected Dialogues of Plato*. Edited by Edith Hamilton and Huntington Cairns. Princeton: Princeton University Press, 1961.

Posner, Richard A. *Sex and Reason*. Cambridge, MA: Harvard University Press, 1992.

Priest, Graham, ed. *The Monist* ["Perversion"] 86 (1) (2003).

Primoratz, Igor. *Ethics and Sex*. London: Routledge, 1999.

Primoratz, Igor, ed. *Human Sexuality*. Aldershot, U.K.: Ashgate, 1997.

Rich, Adrienne. "Compulsory Heterosexuality and Lesbian Existence." *Signs* 5 (4) (1980): 631–660. Reprinted in *Blood, Bread, and Poetry: Selected Prose 1979–1985*. New York: Norton, 1986.

Robinson, Paul. *The Freudian Left: Wilhelm Reich, Geza Roheim, Herbert Marcuse*. New York: Harper and Row, 1969.

Rubin, Gayle S. "Thinking Sex: Notes for a Radical Theory of the Politics of Sexuality." In *Pleasure and Danger: Exploring Female Sexuality*, edited by Carole S. Vance. London: Routledge, 1984.

Russell, Bertrand. *Marriage and Morals*. London: George Allen and Unwin, 1929.

Sartre, Jean-Paul. (1943) *Being and Nothingness: An Essay on Phenomenological Ontology*. Translated by Hazel E. Barnes. New York: Philosophical Library, 1956.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne. New York: Dover, 1966.

Schulhofer, Stephen. *Unwanted Sex: The Culture of Intimidation and the Failure of Law*. Cambridge, MA: Harvard University Press, 1998.

Scruton, Roger. *Sexual Desire: A Moral Philosophy of the Erotic*. New York: Free Press, 1986.

Shelp, Earl E., ed. *Sexuality and Medicine*. 2 vols. Dordrecht, Netherlands: Reidel, 1987.

Shrage, Laurie. *Moral Dilemmas of Feminism: Prostitution, Adultery, and Abortion*. New York: Routledge, 1994.

Snitow, Ann, Christine Stansell, and Sharon Thompson, eds. *Powers of Desire: The Politics of Sexuality*. New York: Monthly Review Press, 1983.

Soble, Alan. *Pornography, Sex, and Feminism*. Amherst, NY: Prometheus, 2002.

Soble, Alan. *Sexual Investigations*. New York: New York University Press, 1996.

Soble, Alan. *The Philosophy of Sex and Love: An Introduction*. St. Paul, MN: Paragon House, 1998.

Soble, Alan, ed. *The Philosophy of Sex: Contemporary Readings*. 4th ed. Lanham, MD, 2002.

Soble, Alan, ed. *Sex from Plato to Paglia: A Philosophical Encyclopedia*. Westport, CT: Greenwood Press, 2005.

Soble, Alan, ed. *Sex, Love, and Friendship: Studies of the Society for the Philosophy of Sex and Love, 1977–1992*. Amsterdam: Rodopi, 1997.

Solomon, Robert C., and Kathleen M. Higgins, eds. *The Philosophy of (Erotic) Love*. Lawrence: University Press of Kansas, 1991.

Stein, Edward, ed. *Forms of Desire: Sexual Orientation and the Social Constructionist Controversy*. New York: Routledge, 1992.

Stewart, Robert, ed. *Philosophical Perspectives on Sex and Love*. New York: Oxford University Press, 1995.

Strossen, Nadine. *Defending Pornography: Free Speech, Sex, and the Fight for Women's Rights*. New York: Scribner, 1995.

Taylor, Richard. *Having Love Affairs*. Buffalo, NY: Prometheus, 1982. Reissued, *Love Affairs: Marriage and Infidelity*. Amherst, NY: Prometheus, 1997.

Thomas Aquinas, St. *On the Truth of the Catholic Faith: Summa Contra Gentiles*. Book 3, pt. 2, *Providence*, translated by Vernon J. Bourke. Garden City, NY: Image Books, 1956.

Thomas Aquinas, St. (1265–1273) *Summa Theologiae*. Cambridge, U.K.: Blackfriars, 1964–1976.

Thurber, James, and White, E. B. *Is Sex Necessary? Or Why You Feel the Way You Do*. New York: Harper and Row, 1975.

Trevas, Robert, Arthur Zucker, and Donald Borchert, eds. *Philosophy of Sex and Love: A Reader*. Upper Saddle River, NJ: Prentice Hall, 1997.

Tuana, Nancy, and Laurie Shrage. "Sexuality." In *The Oxford Handbook of Practical Ethics*, edited by Hugh LaFollette. Oxford, U.K.: Oxford University Press, 2003.

Vannoy, Russell. *Sex Without Love: A Philosophical Exploration*. Buffalo, NY: Prometheus, 1980.

Verene, Donald, ed. *Sexual Love and Western Morality: A Philosophical Anthology*. 2nd ed.. Boston: Jones and Bartlett, 1995.

Watson, Francis. *Agape, Eros, Gender: Towards a Pauline Sexual Ethic*. Cambridge, U.K.: Cambridge University Press, 2000.

Wertheimer, Alan. *Consent to Sexual Relations*. Cambridge, U.K.: Cambridge University Press, 2003.

Wilson, Edward O. *On Human Nature*. New York: Bantam, 1979.

Wojtyła, Karol [Pope John Paul II]. *Love and Responsibility*. New York: Farrar, Straus and Giroux, 1981.

Alan Soble (2005)

PHILOSOPHY OF SOCIAL SCIENCES

The "philosophy of social sciences" comes in three varieties, as the metaideology, the metaphysics, and the methodology of the disciplines involved. The metaideology looks at how far different, traditional legitimations of social sciences succeed. The metaphysics looks at questions having to do with what social science posits—what things it says there are—and at how far those posits are consistent with more or less commonplace beliefs. And the methodology looks at questions regarding the nature of observations, laws, and theories in social science, the logic of induction and confirmation, the requirements of understanding and explanation, and so on.

METAIDEOLOGY

The social sciences were conceived and pursued, from the very beginning, under the influence of ideals (particularly of scientific objectivity and progress) deriving from the eighteenth-century enlightenment (Hawthorn 1976). The first social scientists were economists and sociologists, as we would call them today, and they were self-consciously concerned about producing something that would count, not as philosophy, not as literature, not as common sense, but as science: as a project faithful to the image forged by natural science.

The scientific intention—the intention to make science—has remained characteristic of work in the social sciences. It puts social scientists, paradoxically, under an obligation of an ideological kind: the obligation to show that the sort of analysis they pursue is of a properly scientific kind. The metaideology of social science interrogates and assesses the ideologies whereby the social sciences try to legitimate what they do, to show that what they do is genuinely scientific in character.

Broadly speaking, there are three main ideologies that have been invoked—individually or in various combinations—by social scientists in the scientific legitimation of their enterprise. Each of these marks a feature that putatively distinguishes social science from mere common sense, mere social lore. The first ideology hails social science as an explanatory enterprise of culturally universal validity; the second as an enterprise that is interpretatively neutral, not being warped by people's self-understanding; and the third as an enterprise that enjoys evaluative independence: value-freedom. The universality, neutrality, and independence claimed are each meant to establish social science as objective, and therefore scientifically respectable, in a way in which common sense is not; each notion offers an explication of what scientific objectivity involves. Some approaches in the metaideology of social science, particularly those of a postmodern cast (Rosenau 1992), reject all three ideologies out of hand: They reject any notion of objectivity in the area (others consider them one by one, under the assumption that they may come apart.

Social lore is always lore about a particular social milieu and culture, and an aspiration to cultural universality, if it can be vindicated, would certainly give social science a distinctive status. Such an aspiration is supported in a variety of traditions: among anthropologists and sociologists of a Durkheimian cast, among many Marxist scholars, and among those economists who think that all human behavior, and the patterns to which it gives rise, can be explained by reference to *homo economicus.*

But the metaideologists of social science have claimed many reasons to question the possibility of any universalist, or at least any straightforwardly universalist, theory. Hermeneutic philosophy, which has long been dominant in Germany, and the analytical tradition sponsored by the work of the later Ludwig Wittgenstein both suggest that any explanation of human behavior has to start with the culturally specific concepts in which people understand their environment and cannot aspire, therefore, to a substantive universality (McCarthy 1978, Winch 1958). The debate on these questions ranges widely, encompassing issues of cultural and other forms of relativism (Hollis and Lukes 1982).

Social lore is not only particularistic, it is also designed to represent people as subjectively understandable or interpretable. We, the local consumers of such lore, know what it is like to be creatures of the kind represented and know how we would go about communicating with them. The second, and perhaps least persuasive, ideology of social science suggests that this disposition to represent people as subjectively understandable comes of a limited perspective that social science transcends. It suggests that social science can aspire to an objective explanation of people's behavior without worrying about whether the explanation fits with their self-understanding: without being anxious to ensure that it makes native sense of them and facilitates interpersonal communication. The ideology suggests that social science, in the received phrases, can aspire to a form of *Erklären,* or explanation, that need not service the needs of interpersonal *Verstehen,* or understanding.

Metaideologists of social science have claimed many reasons to question this aspiration to *Verstehen*-free

explanation. Hermeneutic and Wittgensteinian thinkers both reject the idea that people can be properly understood without facilitating communication (Winch, 1958). And the many philosophers who follow the lead of Donald Davidson on interpretation argue that there is no interpreting human subjects without representing them as more or less rational and more or less interpersonally scrutable (Macdonald and Pettit 1981).

Social lore is often evaluatively committed as well as particularistic and oriented to subjective understanding. It takes a form premised on an evaluative characterization of the status quo. Thus, it may characterize the beliefs and explain the behavior of rulers on the assumption that the regime they sustain is unjust. The third and most common legitimating ideology of social science, one associated in particular with the German sociologist Max Weber, holds that in this respect—and perhaps in this respect only—social science can do scientifically better than social lore. It can acknowledge that the agents in the society have evaluative beliefs, and it can take account of these in its explanation of what they do, without itself endorsing any such beliefs; it can be objective, in the familiar sense of remaining uncommitted on evaluative questions.

Metaideologists of social science have also sought reasons to doubt this claim, but the debate has been confused by differences over what sorts of evaluative commitments would really be damaging to the pretensions of social science. The critique of social science on the grounds of not escaping a commitment to value has been nurtured by the appearance, in the later part of the century, of a variety of realist positions on the nature of value. If values are taken to be objective features of the world, then a social scientist's beliefs as to what those features are may well affect their interpretation of how certain subjects think and act; interpretation, after all, is bound to be influenced by the interpreter's view of the subject's environment (Hurley 1989, chap. 5; Macdonald and Pettit 1981, chap. 4; Taylor 1981).

The metaideology of social science may concern itself with other issues: for example, whether the models used in social science, in particular within economics, are really empirical, scientific models and not just pieces of mathematics or exercises in a conversational rhetoric (Hausman 1991, McCloskey 1985, Rosenberg 1992). These issues are not discussed here.

METAPHYSICS

The metaphysics of social science usually takes it as granted that there is no society without individual intentional agents: without subjects who apparently act, other things being equal, on the basis of their beliefs and desires (Pettit 1993, pt. 1). The question that metaphysics raises bears on what more we should include in our metaphysical stock-taking of society; and on how the more we should include, if there is any, relates to individual intentional subjects.

There are two aspects of social life that are particularly relevant to this question. There is the social interaction between individuals in virtue of which various relationships get formed: relationships involving communication, affection, collaboration, exchange, recognition, esteem, or whatever. And there is the social aggregation of individual attitudes and actions in virtue of which various institutions get established: These institutions will include common instrumentalities such as languages, cultures, and markets; groups such as the club, union, or party, whose essence it is to have a mode of collective behavior; groups that may have only a nonbehavioral collective identity such as genders, races, and classes; and shared resources of the kind illustrated by museums, libraries, and states.

The metaphysics of social science concerns itself both with issues raised by interaction and with questions associated with aggregation, specifically with social interaction and aggregation. (On the definition of "social," see Ruben 1985.)

On the side of interaction the main issue in social philosophy is that which divides so-called atomists from nonatomists (Taylor 1985). The atomist holds that individual human beings do not depend—that is, noncausally or constitutively depend—on social relationships for the appearance of any distinctive, human capacities. The nonatomist holds that they do. The atomist defends an image of human beings under which they come to society with all the characteristic properties that they will ever display; social life does not transform them in any essential manner. The nonatomist denies this, believing that it is only in the experience of social relationships that human beings come properly into their own.

The debate between atomists and nonatomists has centered on the connection between thought and language. Atomists have taken their lead from Thomas Hobbes, who argues that, however useful language is for mnemonic, taxonomic, and communicative purposes, thinking is possible without speech, even without any inchoate form of speech. Nonatomists have tended to follow Jean-Jacques Rousseau and the Romantic tradition with which he is associated—a tradition also encompassing Johann Gottfried Herder and G. W. F. Hegel—in

arguing, first, that language is social and, second, that thought requires language.

The atomist tradition has been dominant in English-speaking philosophy, while the nonatomist has had a considerable presence in France and Germany. One source of nonatomism in the English-speaking world has been the work of the later Wittgenstein, in which it is suggested that following a rule—and, therefore, thinking—is possible only in the context of social practices and relationships (Wittgenstein 1968). This very strong nonatomist thesis may also be weakened, so that the claim is that following a rule of a characteristic kind—say, a suitably scrutable kind—requires such a social context (Pettit 1993, chap. 4). Another source of nonatomism in recent English-speaking philosophy has been the argument that the content of a person's thoughts is fixed, not just by what goes on in his head, but by the linguistic community to which he belongs and to which he aspires to remain faithful (Burge 1979, Hurley 1989).

What now of the issues generated by the aggregative aspect of society? There are a number of interesting questions raised by the aggregative structure of society, some having to do with the reducibility of aggregative theory to theory of a more psychological cast, others having to do with the status of aggregative individuals and the standing of the causal relevance we ascribe to such entities (Gilbert 1992, James 1984, Ruben 1985, Tuomela 1996). Perhaps the most pressing question, however, is whether the entities that appear with the social aggregation of individual attitudes and actions give the lie to our ordinary sense of intentional agency: whether it means that, contrary to appearances, we are in some way the dupes of higher-level patterns or forces (Pettit 1993, chap. 3). The individualist, to use a name that also bears further connotations—see under "Methodology"—denies that aggregate entities have this effect; the nonindividualist insists that they do.

One extreme sort of individualism would say that intentional agency is not compromised by any aggregate, social entities, because in strict truth no such entities exist. A more plausible form of the doctrine would say that while there are indeed a variety of aggregate entities, there is nothing about those entities that suggests that our received, commonplace psychology is mistaken. No doubt, there are aggregate regularities associated with such entities: For example, a rise in unemployment tends to be followed by a rise in crime; the fact that something is in an organization's interest generally means that agents of the organization will pursue it; and so on. But the individualist will argue that those regularities do not signal the presence of forces unrecognized in commonplace psychology or the operation of any mechanism—say, any selection mechanism—that belies the assumptions of that psychology. That the regularities obtain can be explained within that psychology, given the context in which the relevant agents find themselves and given their understanding—perhaps involving relevant aggregate-level concepts—of that context.

METHODOLOGY

There are two sorts of methodological questions raised in the philosophy of social science: first, questions imported from the methodology of natural science having to do with such matters as observations and laws and theories, realism and nonrealism in theory interpretation, statistical inference, confirmation, and explanation; second, questions that arise only, or arise distinctively, within the social sciences. Perhaps the two major questions of the latter kind bear on whether it is good explanatory practice to follow the individualistic and economistic assumptions, respectively, that characterize much social science. Here the emphasis will be on the issues of individualism and economism.

The methodological individualist, as characterized in the literature, is associated with a number of more or less outlandish doctrines: for example, that individuals each play indispensable roles, so that things would always have been significantly different if the actual individuals had not been around or if they had not done the things they actually did; that individuals are unaffected by their circumstances, or their relationships with one another, in the things they come to think and want; or that all social facts can be expressed in terms of a nonsocial psychology and that all social laws can be derived from the laws of such a psychology.

Methodological individualism is better understood, however, as a doctrine that has more clearly had respectable defenders as well as opponents: specifically, as the doctrine that it is always good explanatory practice to try to explain social events in terms of finer-grain, individualistic factors rather than by reference to aggregative antecedents. Such an explanatory individualism has been defended by Jon Elster (1985). He argues that aggregative antecedents are causally relevant in virtue of the causal relevance of individual factors and that staying at the aggregative level means leaving the productive mechanism in a black box; it amounts to a willful neglect of relevant facts.

Suppose that we have found a good aggregative explanation of some social phenomenon: say we find that

secularization is explained adequately by urbanization or a rise in crime by a rise in unemployment. We gain further information about the causal history of such a phenomenon as we are informed about the individual-level factors at work in producing secularization or crime. But it may still be that the aggregative story gives us equally important causal information. It may be, for example, that while we learn more about the detail of the actual causal process in going individualistic we learn more about what would be enough to ensure an increase in secularization or crime—that there should be urbanization or unemployment—in spotting the aggregative connections. After all, we might have known the individual-level explanations without having come to recognize the aggregative connections. Perhaps the right line is neither explanatory individualism nor explanatory nonindividualism but explanatory ecumenism (Jackson and Pettit 1992).

The second question bears on whether it is a good explanatory strategy in social science to make economistic assumptions about individual agents: to assume, as economists tend to do, that agents are rational in the way they form and reform their preferences and that their preferences are generally egoistic in character. There are lots of persuasive arguments for following an economistic strategy: arguments that point to the precision in model building and prediction that economistic assumptions allow (Becker 1976). But it seems manifest, on the other hand, that the economistic story is not the whole truth about human beings (Hollis 1977). For example, it is surely obvious that most of us do not make our decisions on the self-concerned, calculative basis that that story would seem to suggest.

But this consideration may not be decisive against economism. For what is possible is that while agents often do not calculate economistically, they tend sooner or later to give up on patterns of behavior that are not at least comparatively satisfactory in economistic terms (Pettit 1993, chap. 5). Perhaps the fact that a pattern of behavior satisfies such economistic constraints is necessary to explain the resilience, if not the actual production and reproduction, of the behavior.

See also Confirmation Theory; Davidson, Donald; Durkheim, Émile; Enlightenment; Explanation; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hermeneutics; Hobbes, Thomas; Induction; Marxist Philosophy; Philosophy of Science, History of; Philosophy of Science, Problems of; Postmodernism; Realism; Rousseau, Jean-Jacques; Weber, Max; Wittgenstein, Ludwig Josef Johann.

Bibliography

Becker, G. *The Economic Approach to Human Behavior.* Chicago: University of Chicago Press, 1976.

Burge, T. "Individualism and the Mental." *Midwest Studies in Philosophy* 4 (1979): 73–121.

Elster, J. *Making Sense of Marx.* Cambridge, U.K.: Cambridge University Press, 1985.

Gilbert, M. *On Social Facts.* Princeton, NJ, 1992.

Hausman, D. *The Inexact and Separate Science of Economics.* Cambridge, U.K.: Cambridge University Press, 1991.

Hawthorn, G. *Enlightenment and Despair: A History of Sociology.* Cambridge, U.K.: Cambridge University Press, 1976.

Hollis, M. *Models of Man: Philosophical Thoughts on Social Action.* Cambridge, U.K.: Cambridge University Press, 1977.

Hollis, M., and S. Lukes, eds. *Rationality and Relativism.* Cambridge, MA: MIT Press, 1982.

Hurley, S. *Natural Reasons: Personality and Polity.* New York: Oxford University Press, 1989.

Jackson, F., and P. Pettit. "In Defence of Explanatory Ecumenism." *Economics and Philosophy* 8 (1992): 1–21.

James, S. *The Content of Social Explanation.* Cambridge, U.K.: Cambridge University Press, 1984.

Macdonald, G., and P. Pettit. *Semantics and Social Science.* London: Routledge and Kegan Paul, 1981.

McCarthy, T. *The Critical Theory of Jürgen Habermas.* Cambridge, MA: MIT Press, 1978.

McCloskey, D. *The Rhetoric of Economics.* Madison: University of Wisconsin Press, 1985.

Papineau, D. *For Science in the Social Sciences.* London: Macmillan, 1978.

Pettit, P. *The Common Mind: An Essay on Psychology, Society, and Politics.* New York: Oxford University Press, 1993.

Rosenau, P. M. *Post-Modernism and the Social Sciences: Insights, Inroads, and Intrusions.* Princeton, NJ: Princeton University Press, 1992.

Rosenberg, A. *Economics: Mathematical Politics or Science of Diminishing Returns?* Chicago: University of Chicago Press, 1992.

Ruben, D.-H. *The Metaphysics of the Social World.* London: Routledge and Kegan Paul, 1985.

Ryan, A. *The Philosophy of the Social Sciences.* London: Macmillan, 1970.

Taylor, C. *Philosophy and the Human Sciences.* Cambridge, U.K.: Cambridge University Press, 1985.

Taylor, C. "Understanding and Explanation in the *Geisteswissenschaften.*" In *Wittgenstein: To Follow a Rule,* edited by S. H. Holtzman and C. M. Leich. London: Routledge and Kegan Paul, 1981.

Tuomela, R. *The Importance of Us.* Stanford, CA: Stanford University Press, 1995.

Winch, P. *The Idea of a Social Science and Its Relation to Philosophy.* London, 1958.

Wittgenstein, L. *Philosophical Investigations,* 2nd ed. Translated by G. E. M. Anscombe. Oxford: Blackwell, 1968.

Philip Pettit (1996)

PHILOSOPHY OF STATISTICAL MECHANICS

Probabilistic modes of description and explanation first entered into physics in the theory of statistical mechanics. Some aspects of the theory that are of interest to the general philosopher of science are the nature of probability and probabilistic explanations within the theory, the kind of intertheoretical relation displayed between this theory and the nonprobabilistic theory it supplants, and the role to be played in scientific explanations by the invocation of cosmological special initial conditions. In addition, this theory provides the framework for attempts to account for the intuitive sense that time is asymmetric by reference to asymmetric physical processes in time.

HISTORY OF THE THEORY

It was in the seventeenth century that thinkers first realized that many material systems were describable by a small number of physical quantities related to one another by simple laws—for example, the ideal gas law, relating the volume, temperature, and pressure of a gas.

It was soon understood that a fundamental notion was that of equilibrium. Left alone, systems might spontaneously change the value of their parameters, as when a gas expands to fill a box. But they would soon reach an unchanging final state, that of equilibrium. And it was realized that this process was asymmetrical in time, in that systems went from earlier non-equilibrium states to later equilibrium states, but not from earlier equilibrium states to later states of non-equilibrium.

Studies of steam engines initiated by S. Carnot showed that stored heat could be converted to mechanical work, but only by a process that converted stored heat at a higher temperature to residual heat at a lower temperature. This result was made mathematically elegant by R. Clausius, who introduced the notion of entropy as a measure of heat's ability to be converted into external work into physics. That heat was a form of stored energy and that the total amount of energy in heat and work was conserved became a fundamental principle of physics, as did the idea that energy could spontaneously only go from a more ordered to a more a less orderly state. These results were formalized in the First and Second Laws of Thermodynamics. But why were these laws true?

The latter half of the nineteenth century and the beginning of the twentieth saw the development of an intensive debate about the place of thermodynamics within the more general sciences that dealt with dynamics and with the constitution of matter. P. Duhem, E. Mach, and others argued that the laws should be understood as autonomous principles. But others sought an account of heat as the hidden energy of motion of the microscopic constituents of matter. This was later understood for gases in terms of a simple model of molecules in free motion except for collisions among them. The early work on kinetic theory of W. Herepath and J. Waterston, followed by work of A. Kronig, made this a rich area for theoretical exploration. J. C. Maxwell and L. Boltzmann discovered laws governing the distribution of velocity of the molecules in the equilibrium state, and they developed a law governing how such distributions changes as a system in nonequilibrium approached equilibrium, at least for the simple system of a nondense gas.

The theory of approach to equilibrium soon met with profound objections. J. Loschmidt pointed out that the apparently demonstrated time-asymmetrical approach to equilibrium was hard to understand in light of the fact that the laws governing the underlying dynamics of the molecules allowed for the time reverse of each possible process to be possible as well. Later H. Poincaré showed that the kind of systems being dealt with would, except possibly for exceptional initial conditions in a class of probability zero, return over infinite time infinitely often to states arbitrarily close to their initial states. Once again this seemed incompatible with the monotonic increase of entropy described by thermodynamics and apparently deduced from the dynamics in kinetic theory.

Both Maxwell and Boltzmann introduced probabilistic elements into their theory. The equilibrium distribution might be thought of as the most probable distribution of the molecules in space and in velocity. Alternatively, in an approach later systematically developed by J. W. Gibbs, equilibrium values might be calculated by computing the average of macroscopic features over all possible distributions of the molecules. Both Maxwell and Boltzmann also argued that approach to equilibrium should also be thought of probabilistically. Maxwell discussed the possibility of a "demon" who could, by inspecting molecules one by one, change an equilibrium state of a system to a nonequilibrium state without doing external work on the system. Critics such as S. Burbury and E. Culverwell noted that the introduction of probabilistic notions was not sufficient by itself to overcome the puzzles of reversibility and recurrence.

In his last view of the theory, Boltzmann, following his assistant Dr. Scheutz, offered a time-symmetrical version of the theory. On this view, isolated systems spend most of their life near equilibrium over very long periods

of time. There would be occasional fluctuations away from equilibrium. A system found in a nonequilibrium state would probably be closer to equilibrium both in the past and future. Our local region of the universe, a universe that as a whole was itself in equilibrium, was one such fluctuation. Scientists could only exist in such a nonequilibrium regions because only such a region could support sentient creatures. Why do we find our local world approaching equilibrium in the future and not in the past? Because the time direction of increase in entropy determined the future just as the local direction of gravitational force determined the down spatial direction.

In an important study of the foundations of the theory in 1910, P. and T. Ehrenfest (1959) surveyed the basis of the theory as understood in different ways by Maxwell, Boltzmann and Gibbs. They also offered an important interpretation of Boltzmann's equation describing approach to equilibrium in which the solution of the equation described not the inevitable or even probable behavior of an individual system but rather the sequence of states that would be found dominant at each time in a collection of systems all of whose members started in the same macroscopically nonequilibrium condition.

PROBABILITY AND STATISTICAL EXPLANATION

Probability is characterized formally by simple mathematical postulates, the additivity of probabilities over disjoint sets of events being the most important of these. Philosophers have long debated the interpretation of probability. Some interpretations are subjectivist, taking probabilities to be measures of partial belief. Others are logical, holding probabilities to represent partial entailments. Other interpretations are objectivist. Some varieties of this last are frequency, limits of frequency, or dispositional interpretations.

At least one proposal (by E. Jaynes) has held that the probabilities in statistical mechanics are subjective, or rather of a kind of logical sort resting upon a principle of indifference. Most interpreters of statistical mechanics hold to objectivist interpretations of probability, but even among them there is much debate. Are the probabilities somehow dependent on the underlying dynamical laws, as ergodic approaches suggest? Or are they reflective of a deeper lawlike structure of tychistic chance, as Albert suggests, referring to Ghirardi-Rimini-Weber (GRW) stochastic theories introduced in the interpretation of quantum mechanics? Or is it the case, rather, that the probabilities have an autonomous place within the theories requiring their independent postulation?

Philosophers analyzing statistical explanations have usually focused on uses of probabilistic explanation in everyday circumstances or in the application of statistics to such fields as biology. Here some suggestions have been that high probability is explanatory, that increased probability is what matters, or that explanations are only genuinely probabilistic when pure tychistic chance is relevant.

In statistical mechanics explanation in the nonequilibrium theory has many aspects that fit familiar patterns of statistical explanation as analyzed by philosophers. Within the theory the main areas of controversy are over the nature and rationale for the particular kind of probabilistic explanation that does justice to the empirical facts. In the equilibrium theory a kind of transcendental use of probability in the statistical explanations offered by ergodic theory is quite unlike the usual kind of causal-probabilistic explanations familiar in other contexts.

THE THEORY OF EQUILIBRIUM

Boltzmann and Maxwell developed a standard method for calculating the equilibrium values of the macroscopic parameters of a system. This became formalized by Gibbs as the method of the microcanonical ensemble. Here a probability distribution is placed over the microstates possible for the system, given its constraints. For each microstate the values of the macroscopic parameter are calculable. One takes as the observed equilibrium values the average value of these parameters calculated over all the possible microstates, using the stipulated probability distribution. But why does the method work? What rationalizes the choice of probability distribution and the identification of average values with equilibrium quantities?

Boltzmann argued that the method could be partly justified if one thought of equilibrium values as average values over an infinite time as the system changes its microstates under dynamic evolution. Another component of this way of thinking is a claim that, given the large numbers of molecules in a system, average values would coincide with overwhelmingly most probable values for a macroscopic parameter. Boltzmann and Maxwell argued that one could identify such time averages with so-called phase averages, calculated using the posited probability distribution over the microscopic conditions possible for the system, if one thought of any one system as going through all possible microstates compatible with the macroscopic constraints on the system as time went on.

This became formalized by the Ehrenfests in the form of the Ergodic Hypothesis.

Early versions of the Ergodic Hypothesis were provably false. Weaker versions, such as the claim that the microstate of the system would come arbitrarily close to every possible microstate over infinite time, were impossible to demonstrate and could not support the equality of time and phase averages even if true.

These early ideas gave rise to the mathematical discipline of ergodic theory. The results of J. von Neumann, and, in stronger form, those of G. Birkhoff, showed that for certain idealized dynamical systems, except for a set of initial conditions of zero probability in the standard probability distribution, the time average of quantities calculated from the microstate of the system over infinite time would, indeed, equal the phase average of that quantity calculated using the standard-probability distribution over all possible microstates of the system.

But did any realistic models of a system meet the conditions needed for these theorems to hold? Many decades of work, culminating in that of Sinai, showed that a familiar model of a dilute gas, hard spheres in a box, was a model of an ergodic system. On the other hand, important work in theoretical dynamics showed that more realistic models of the gas would necessarily fail to be strictly ergodic (the KAM theorem). So any hope of applying ergodicity to rationalize the standard theory would require subtle reasoning involving the fact that the system was composed of vast numbers of molecules and might be, therefore, "ergodiclike."

From ergodicity many consequences follow. Except for a set of initial points of probability zero, infinite time averages of a phase quantity will equal the phase average of that quantity. For any measurable region of the phase space, the proportion of time spent by the system in that region over infinite time will equal the probabilistic size of that region. Most important is the following: Boltzmann realized that the standard probability distribution was invariant over time under the dynamics of the system. But could there be other such time invariant distributions? If the system is ergodic, one can show that the standard distribution is the unique time-invariant distribution, which assigns zero probability to regions assigned zero probability by the standard distribution.

These results provide us with a kind of transcendental rationale for the standard equilibrium theory. Equilibrium is an unchanging state. So if we are to identify macroscopic features of it with quantities calculated by using a probability distribution over the microstates of the system, this probability distribution should be unchanging under the dynamics of the system. Ergodicity shows us, with a qualification, that only one such probability distribution, the standard one, will do the trick.

But as a full rationale for the theory, ergodicity must be looked at cautiously. Real systems are not genuinely ergodic. We need to simply swallow the claim that we may ignore sets of conditions of probability zero in the standard measure. And the kind of rationale we get seems to ignore totally the place of equilibrium as the end point of a dynamic evolution from nonequilibrium conditions.

THE THEORY OF NONEQUILIBRIUM

Maxwell and Boltzmann found equations describing the approach to equilibrium of a dilute gas. Later a number of other such kinetic equations were found, although attempts at generalizations to such situations as dense gases have proved intractable.

But how can such equations, whose solutions are time asymmetric, possibly be correct if the underlying dynamics of the molecules are symmetrical in time? Careful analysis showed that the Boltzmann equation depended upon a time-asymmetrical assumption, the Stosszahlansatz. This posited that molecules had their motions uncorrelated with one another before, but not after, collisions. Other forms of the kinetic equations made similar assumptions in their derivation. Two general approaches to deriving such equations are that of the master equation and the approach that works by imposing a coarse graining of cells over the phase space available to the system and postulating fixed transition probabilities from cell to cell. But the time-asymmetrical assumption must be imposed at all times and might even be inconsistent with the underlying deterministic dynamics of the molecules.

Many attempts have been made to understand the kinetic equations and to resolve the paradoxes. Some of these explore how an initial probability distribution over a collection of systems can, in a "coarse-grained" sense, distribute itself over the increased phase volume available to a system. This way of looking at things was first described by Gibbs. The coarse-grained spreading of the probability distribution is taken to represent the approach to equilibrium of the system. This interpretation fits with the understanding of the solution curve of the Boltzmann equation outlined by the Ehrenfests.

To show that such spreading of the initial probability distribution occurs, one relies upon the underlying dynamics and generalizations of the results of ergodic

theory. Systems can be characterized as randomizing in a variety of senses of increasing strength such as being a mixing system, a K-system, or a Bernoulli system. Then one can rely upon the model of the system—hard spheres in a box, for example—and the dynamics to show the system randomizing in the specified sense. This approach often relies upon many idealizations, such as calculating what happens in the infinite time limit. And the results often depend upon the use of unrealistic models of systems. For these reasons the applicability of the results to real systems and their real finite time behavior requires care.

Crucially these results, following as they do from the time-symmetrical dynamics, cannot by themselves introduce time asymmetry into the account. To do that one must make a time-asymmetrical assumption about how the initial probability distribution over the microstates of the system is constrained. This problem was studied by N. Krylov and others. Krylov's solution was a kind of nonquantum uncertainty principle applicable to the preparation of systems. Others look for the solution in cosmological facts, as we shall later note. Still others seek to modify the underlying dynamics by postulating some time-asymmetrical fundamental physical principle in play, such as the time-asymmetrical GRW stochastic field proposed in some interpretations of quantum mechanics.

There are ways of trying to understand an approach to equilibrium quite at odds with the mixing approach just described. O. Lanford, for example, has produced a "rigorous derivation of the Boltzmann equation." Going to an idealized limit, the Boltzmann-Grad limit, Lanford imposes an initial probability distribution, and then shows that with probability one systems will evolve for a short time as described by the Boltzmann equation. Because the results can be proved only for very short times—less than the mean free time to the first collision—their applicability to the real world is again in question. As usual, interesting issues about time asymmetry arise, here in the form of the choice of the initial probability distribution.

IRREVERSIBILITY

Why is it that, although the underlying dynamic principles are symmetrical in time, the thermodynamic laws describe a world asymmetrical in time, a world in which entropy spontaneously increases in one time direction but not the other? Merely introducing probabilities into the account by itself will not provide the grounds for understanding the physical origins of irreversibility.

Throughout the history of thermodynamics and statistical mechanics, the suggestion has been repeatedly made that the source of thermodynamic time asymmetry lies in the existence of some time-asymmetrical law governing the underlying dynamics. The recent invocation of time asymmetric GRW stochastic influences is the latest such proposal.

Sometimes it has been suggested that the entropic increase experienced by an "isolated" system is to be accounted for in terms of the fact that systems can never really be fully causally isolated from their external environment. Even the most carefully insulated system, for example, has its molecules' motion influenced by gravitational forces exerted by matter outside the system. Whether the fact that isolation is an idealization is really relevant to thermodynamic time asymmetry has been much debated. Of great importance to this debate is the existence of systems that seem to show the usual macroscopic entropic increase familiar from thermodynamics, but which are systems sufficiently isolated from their surrounding environments such that a simple external trigger can have their microstates follow a reverse course, with the system recurring to its original nonequilibrium state—spin-echo experiments, for example. For these systems seem to show that a kind of entropic increase cannot be accounted for in terms of external interference with the system.

As noted above, it was Boltzmann's assistant, Dr. Scheutz, who first suggested a cosmological solution to the problem. Scheutz suggested that the universe as a whole is in a time-symmetrical equilibrium state, with our local portion of the cosmos in a rare fluctuation away from equilibrium. Such a region would be very likely, from a time symmetrical probabilistic perspective, to evince higher entropy in one time direction but lower entropy in the other direction of time, since it is unlikely to be at the turning point of maximal deviation from equilibrium. Boltzmann then supplemented this with his assertion that the very meaning of the future is that is the time direction in which entropy is increasing.

Current cosmological theories describe a very different sort of universe, one that, to the best of our knowledge, is in an overall nonequilibrium state and that has entropic increase in the same time direction in all its regions. In current Big Bang cosmology the universe is said to be spatially expanding from a singularity some tens of billions of years ago. Some theorists take the thermodynamic time asymmetry to have its roots in the cosmic expansion. The more general opinion is that this cannot be correct, since, according to the prevailing but

not universal opinion, even if the universe began to contract, entropy would continue to increase.

In the dominant opinion, rather, the source of entropic increase is found in a special physical condition of the universe just after the Big Bang. In these accounts the matter of the universe is taken to be, at that early date, in thermal equilibrium. But matter is thought to be smoothly distributed in space. This is a very low entropy state because of the fact that gravity, unlike intermolecular forces in a gas, is a purely attractive force. The theory goes on to propose a clumping of matter into dense galactic clusters, galaxies, and stars, leaving most of space almost devoid of matter. This results in an enormous increase in spatial-gravitational entropy. Matter so clumped goes into a lower entropy state than its original equilibrium, since it now consists of hot stars in cold interstellar space. The general increase of entropy from the Big Bang onward is then accounted for by positing both the usual time-symmetrical probability assumptions and initial low entropy for the universe as a whole.

One question that then arises is why the initial state should be one of such low entropy. Here one is up against the usual perplexities that arise if we ask for an answer to a why question about "the initial state of everything." Why is such a low-probability state the one we find? Should one posit many universes, of which our low-probability case is a rare example? Here one is reminded of the speculation of Scheutz about our region of the universe just being an improbable sample from the whole. Can one explain why we find ourselves in such a universe by some version of the anthropic principle, first used by Boltzmann to explain why we find ourselves in a low-entropy region of his speculated high-entropy universe? Can one attribute probabilities to initial singular states or to universes at all? Here one thinks of the criticism offered by D. Hume of the teleological argument for the existence of God.

The second law of thermodynamics is not concerned, of course, with the entropy change of the entire cosmos, but rather with the parallel in time-entropic increases of small systems temporarily causally isolated from their external environments. The study of the connection between cosmic entropy increase and that of the "branch systems" was initiated by H. Reichenbach. Many of the arguments in the literature claiming to derive changes of entropy of branch systems that are parallel in time to the entropy increase of the cosmic whole are badly flawed, but a reasonable inference can likely be constructed using probabilistic posits that themselves do not smuggle time asymmetry into the derivation.

THERMODYNAMICS AND STATISTICAL MECHANICS

We often speak of an older theory being reduced to a newer theory, and it is often said that thermodynamics has been reduced to statistical mechanics. But, as we have learned in general, the relation of older theory to newer theory may be of some complexity and some subtlety.

Thermodynamics, traditionally, was not a theory framed in probabilistic terms. Its laws, especially the second law, could not be exactly true, as Maxwell noted, in the light of the new probabilistic account. Alternative ways of dealing with this problem are available. One way is to stick with traditional thermodynamics and offer an account of the relation between newer and older theory that is far from a simple derivation of the latter from the former. Another possibility is to use the new knowledge of the probabilistic aspects of thermal phenomena to construct a novel statistical thermodynamics that imports probabilistic elements directly into the older theory.

There must be a high degree of complexity in the relations between the concepts of the older theory—such as volume, pressure, temperature and entropy—and those of the newer theory—such as concepts dealing with molecular constitution, the dynamics governing the molecules, and probabilistically framed concepts dealing either with the distribution of states of constituents of the individual system or with the distribution of microstates of systems in a collection of systems characterized by some macroscopic parameters.

Consider, for example, thermodynamic entropy. Associated with it are many distinct entropy concepts in statistical mechanics. Boltzmann entropy, for example, is defined as the fluctuating property of an individual system, defined in terms of the actual spatial and momentum distribution of the molecules of the system at a time. Gibbs's entropies, on the other hand, are defined in terms of some probability distribution imposed over some imagined ensemble of systems characterized by some specified constraints. To make matters even more complicated, there is Gibbs's fine-grained entropy, defined by the probability distribution alone and useful for describing the equilibrium states of systems, and Gibbs's coarse-grained entropy, whose definition requires a specification of some coarse-grained partition of the phase space as well as the probability distribution, and whose place is in characterizing the approach to equilibrium of nonequilibrium systems. Other notion of entropy, such as those defined in terms of topology rather than measure theory, exist as well.

None of this complexity shows that one is wrong in thinking that in some appropriate sense, statistical mechanics explains the success of thermodynamics or that it might be plausible to speak of a reduction of thermodynamics to statistical mechanics. The complexity and subtlety of the relations between the two theories informs the philosopher of science of just how varied and complicated such reductive relations might be.

Philosophers outside the field of philosophy of physics might take some interest in the relationship that thermodynamics bears to the underlying physical description of the systems to which thermodynamic concepts are applied. A material object composed of atoms or molecules, for example, can exist in equilibrium with a system of electromagnetic radiation, leading physicists to speak of both such systems as having a common temperature. What this shows is that concepts such as entropy and temperature have a kind of functional role, with their meanings fixed by the place they play in a theory that is applicable to physical systems of many different kinds. This bears some analogy with the claim, so familiar in the philosophy of mind, that mental terms are functional and that mental states are multiply realizible in physical systems of varied natures.

THE DIRECTION OF TIME

The claim that our very notion of the asymmetry of time is rooted in entropic asymmetries of physical systems in time was first made by Boltzmann, as we have noted. The claim has often been repeated but remains controversial. Much needs to be done to provide a completely convincing case that our deepest intuitions about the difference between past and future are somehow grounded in entropic asymmetries.

A first question relates to what an entropic theory of the direction of time is claiming. It certainly cannot be that we find out which direction of time is the future by somehow checking up directly on the entropic behavior of systems around us, for that claim has little plausibility. So what does the claim come down to?

What intuitively distinguishes future from past? We think we have a direct insight into which of a pair of events is later than the other. We take it that we have asymmetric epistemological access into past and future, there being memories and records of the past and not of the future. We usually take it that causation goes from an earlier event as cause to a later event as effect. We are anxious about future events but not about past events, although we may regret the latter. We often think of the past as being over and done with and hence not subject to change, whereas the future is open to many possibilities. Some philosophers have argued that past events have determinate reality, whereas there is no such thing as a determinate being to the future.

The most plausible version of the entropic theory of the direction of time is best understood by looking at the analogy introduced by Boltzmann. What lies behind our intuitions that space is distinguished by an asymmetry because one direction is down and its opposite up? Surely it is the existence of gravitational force that fully accounts for the down-up distinction. It is gravity that explains why rocks fall down and, in our atmosphere, flames and helium balloons go up. Even the fact that we can tell, directly and without using our sensory awareness of the external world, which direction is down is explained in terms of the local direction of gravitational force. For it is the behavior of fluids in our semicircular canals that tells us which way is up, and the behavior of that fluid is entirely explained in terms of its gravitationally induced weight. In regions of the universe with no gravitational field, there is no distinction between the up and the down direction to be drawn.

The entropic theorist of the direction of time argues that the situation is exactly analogous to the case of down directionality and gravity. The claim is that we can account for all the intuitive differences by which we distinguish past from future by a scientific account at whose core are entropic asymmetries in the behavior of systems in time. If there were regions of the cosmos in which entropic changes were antiparallel to one another in time, the entropic theorist claims, the inhabitants of such regions would take opposite directions of time to be the future direction of time. And in regions of the cosmos in equilibrium, there would be no past-future distinction, although, of course, there would still be opposite directions in time.

There have been numerous proposals, starting with the seminal work of H. Reichenbach, to try to justify the claim that is it is, indeed, entropic change that lies at the heart of any explanation of why we have memories and records of the past and not of the future, of why we think of causation as going from past to future, of why we have differential concerns about past and future, and of why we think of the past as determinate but think of the future as an open realm of mere possibilities. Despite much important work on this problem, however, the very possibility of constructing such entropic accounts remains controversial.

See also Causal Approaches to the Direction of Time; Counterfactuals; Physics and the Direction of Time.

Bibliography

Albert, D. *Time and Chance*. Cambridge MA: Harvard University Press, 2000.

Brush, S. *The Kind of Motion That We Call Heat*. Amsterdam: North-Holland, 1976.

Brush, S., ed. *Kinetic Theory*. Oxford: Pergamon Press, 1965.

Ehrenfest, P. and T. *The Conceptual Foundations of the Statistical Approach in Mechanics*. Ithaca, NY: Cornell University Press, 1959.

Guttman, Y. *The Concept of Probability in Statistical Physics*. Cambridge, U.K.: Cambridge University Press, 1999.

Price, H. *Time's Arrow and the Archimedean Point*. Oxford: Oxford University Press, 1996.

Reichenbach, H. *The Direction of Time*. Berkeley: University of California Press, 1956.

Sklar, L. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge, U.K.: Cambridge University Press, 1993.

Lawrence Sklar (2005)

PHILOSOPHY OF TECHNOLOGY

The philosophy of technology brings logical, metaphysical, epistemological, ethical, and political philosophical questions to bear on the making and using of artifacts. The particular balance among these questions will differ within related regionalizations of philosophy, such as the philosophy of science or the philosophy of art. In the philosophy of technology, for instance, epistemology typically plays a lesser role than in the philosophy of science but a greater role than in the philosophy of art. Any philosophical assessment of technology is thus partially defined by its own inner balance in relation to philosophy as a whole.

HISTORICAL EMERGENCE

Although limited discussions of *techne* and associated or derivative phenomena can be found in ancient, medieval, and early modern philosophy, it was not until the late nineteenth and early twentieth centuries that technology, as something distinct from technics or technique, became a subject for theoretical examination. Among the earliest contributing texts, the mechanical engineer Franz Reuleaux's *Theoretische Kinematik* (1875) developed an extended conceptual analysis of different types of tools and machines. More generally, Ernst Kapp's *Grundlinien einer Philosophie der Technik* (1877), in the first book to use "philosophy of technology" in its title, outlined a theory of culture grounded in technics understood as the extension and differentiation of human anatomy and physiology. The hammer, for instance, functions as an extension of the fist, the camera as an extension of the eye, and the railroad as an extension of the circulatory system; and vice versa, the fist can be said to be like a hammer, the eye like a camera, and rail lines like blood vessels. Elaborations of this view of technology as organ projection are representative of a school of what Carl Mitcham (1994) calls engineering philosophy of technology, an approach that was further developed in the work of thinkers as diverse as the Russian Peter Englemeier, the German Friedrich Dessauer, the Frenchman Gilbert Simondon, and the Spaniard Juan David García Bacca (all of whom have been largely ignored in Anglo American philosophy).

The research engineer Dessauer, for instance, developed a neo-Kantian critique of the transcendental possibility of technological invention that sees technology as bringing noumenal power into the world. Dessauer was also instrumental in promoting philosophical discussion within the Verein Deutscher Ingenieure (VDI; Society of German Engineers). The psychologist Simondon explored relations among parts, artifacts, and technical systems and the evolutionary manifestation of what he called technicity. The engineer Englemeier and the philosopher García Bacca both saw technological change engendering world-historical transformations that were at once humanizing and transcending of the merely organically human. Additional contributions to this school can be found in theoretical discussions about cybernetics and artificial intelligence. Also illustrative of achievements in engineering-oriented philosophy of technology are the scientific philosopher Mario Bunge's (1985) systematic metaphysics, epistemology, and ethics of technology and the engineer Billy Vaughn Koen's (2003) brief for engineering as the one right method for problem solving.

In its emergence, however, philosophy of technology was more commonly associated with what might be called a counterphilosophy that interprets technology not as extending but as encroaching on or narrowing the dimensions of human experience. Following Immanuel Kant's attempt "to deny [scientific] *knowledge*, in order to make room for *faith*," this humanities philosophy of technology has sought to limit technological thought and practice to make room for human culture in all its rich diversity. A case in point is the public intellectual Lewis

Mumford's (1967) criticism of what he calls monotechnics, the technics of power, in contrast to poly- or biotechnics. The problem with monotechnics is that it promotes the pursuit of physical power and control at the expense of other aspects of human flourishing such as friendship and art. For Mumford the "myth of the machine" is to think that power is the source of all human benefit. In fact, it constitutes an unrealistic narrowing of human activity. Some version of this argument has been promoted especially by the continental European philosophical tradition in the works of José Ortega y Gasset (1939), Martin Heidegger (1954), and Jacques Ellul (1954). Indeed, even more broadly, the relation between technology and life—whether in the sense of *zoe* (organic existence) or *bios* (human flourishing)—has become one of the most crucial issues in both the metaphysics and ethics of technology.

Until the latter half of the twentieth century, the argument for delimitation had the unintended side effect of relegating technology to marginal status in professional philosophy. Only as technology became more than an engineering interest or a social problem has it begun to be a mainstream topic in philosophy. One of the challenges in the twenty-first century will be to pursue the professional development of philosophical reflection on technology in ways that bridge the oppositions inherent in its bimodal historical origins without compromising their basic if divergent concerns.

ETHICAL AND POLITICAL ISSUES

Because of their prominence in public affairs, the philosophy of technology properly highlights ethical and political issues. Indeed, contemporary work in practical or applied ethics—as in nuclear, environmental, biomedical, and computer ethics—emphasizes the moral challenges of technology, although in ways that sometimes reduce the field to an aggregate of different ethics for different technologies. Such subspeciation can deprive ethics of possible synergistic strengths. Access equity issues, for instance, occur in both biomedicine and computers, and the concepts and principles for dealing with one might well inform or enhance the other. Speaking generally, then, one can identify at least six competing and overlapping interpretations of technology as an ethical or political problem. Three of these arose initially before World War II, although they have continued to cast a shadow of concern, often in new and distinctive forms.

First, there is a problem of the just distribution of technological products and powers—that is, technology as a political issue. Since the Industrial Revolution the social-justice question has found numerous expressions in authoritarian and democratic regimes, in developing and developed countries. Authoritarian regimes have often justified themselves as acting to promote access to technological benefits against entrenched special scientific, technical, or corporate interests or against those whose commitment to equality undermines the invention and production of goods and services. Democratic regimes have placed more emphasis on promoting equality by means of due process and regulatory agencies. One aspect of due process that has been given special philosophical attention concerns the legal protocols to promote free and informed consent, extending the concept from human experimentation to engineering at large (Martin and Schinzinger 2005).

With the engineered design of new products and processes social justice issues have often taken special form in association with some otherwise morally neutral concepts. The advent of electronic computer and Internet communications, for instance, has helped impart ethical significance to questions of privacy and the so-called "digital divide." Additionally, according to Ulrich Beck (1992), concerns for the fair distribution of goods and services were, during the late twentieth century, superseded by those dealing with the fair distribution of dangers and risks, thus giving social justice debates a special twist. One of the strongest criticisms of some of the resulting twists and turns has been Kristin S. Shrader-Frechette's (1991) careful dissecting of the antidemocratic assumptions of much risk-cost-benefit analysis.

Second is the problem of the alienation of workers from their labor in the industrial means of production, which has been presented especially by Marxists as an economic and by some non-Marxist social scientists as a psychological issue. Langdon Winner's (1977) analysis of the theory of autonomous technology or the idea that technology as resistant to human control is a more general statement of the issue. Critical theory work by Herbert Marcuse (1964) and Andrew Feenberg (1991, 1999) extended the classic Marxist discussion into situations reconfigured by consumerist culture and globalization. Opposing Marcuse's pessimism about transformation, Feenberg (especially 1995) has been more optimistic about alternative possibilities. Environmentalists, however, have further argued that technology in general alienates human beings from nature.

Don Ihde's (1990) phenomenology of the technolifeworld offers another take on this issue through an analysis of human—technology—world relations. Two fundamental types of such engagements are instrumental

relations, in which the technology is integrated into the human sensorium as its extension (the blind man's cane), and hermeneutic relations, in which the technology becomes part of the world to be interpreted (a thermometer). Both engagements manifest an invariant structure that amplifies some aspect of the world (exact metric of temperature) while simultaneously reducing others (general sense of climate). The former tends to bring humans closer to the world, the latter to distance (or alienate) them from it.

Third is the problem of the destruction or transformation of culture by modern science and technology—either directly through new weapons and forms of military conflict or indirectly through the impact of new means of transportation, communication, and media. The destruction of World War I, the most violent in human history, was a manifestation of technology that only became worse during World War II with the development of nuclear weapons. The long cold war practice of nuclear deterrence and the early twenty-first-century challenges of terrorism present special problems for learning to manage the destructive potential in technology.

Between the two world wars concern for the more indirect technological transformation of culture took on special salience, as variously illustrated by the cultural lag theory of the American sociologist William Fielding Ogburn, the elegiac ruminations of the Catholic theologian Romano Guardini, or the active nihilistic enthusiasms of Ernst Jünger. In the latter half of the twentieth century the issue found small-scale manifestation in personal efforts to come to terms with new choices (e.g., in diet, drugs, and consumer lifestyle options) and large-scale manifestation in debates about the dynamics of sociotechnical change (e.g., the role of technology in economic development and technological determinism versus social constructionism). Questions can also arise about the transformed character of cultural life under the influence of information and image technologies, from television to the Internet and virtual reality machines.

Since World War II three more issues have emerged to ethical and political prominence. One is that of democratic participation. An anticipatory version of this issue emerged in interwar proposals for technocracy. For some theorists (such as Thorstein Veblen) rule by technical elites offered a better alternative than rule by economic or political elites. However, in the postwar revival of democratic theory, and with recognition that technology (like law) is a creation that also influences the creators, it was argued that the principle of "no taxation without representation" should be extended to "no innovation without representation" (Goldman 1992). Winner, for instance, describes "technologies as forms of life" and calls for the abandonment of "technological somnambulism" (1986, p. 10) in favor of public debate about the design of technological projects as diverse as highway bridges, tomato harvesters, and nuclear power plants. Efforts to determine how such democratic participation should be structured both within communities of technical expertise and in the negotiations between technical experts and the nontechnical public have been the subject of ongoing debates (see Sclove 1995).

Fifth is the industrial pollution of the natural environment, which has contributed to attempts to develop an appropriate environmental or ecological ethics. What is the difference between artifice and nature—and the moral status of wilderness or the nonhuman environment? As nature is humanly transformed, to what extent should contemporary technological action take into account the welfare of future generations, whether human or nonhuman? What is the relation between values that are divided between the anthropocentric and ecocentric, extrinsic or instrumental and intrinsic?

Another morally relevant concept, closely related to issues of both participation and environmentalism, is that of unintended consequences. To what extent are scientists and engineers responsible for the unexpected and perhaps even unforeseeable results of their technological actions? Two attempts to deal with the plethora of environmental issues, especially in relation to the challenge of unintended consequences, are those associated with sustainable development and the precautionary principle—with competing interpretations of both becoming major themes of moral and political deliberations.

Finally, there is the issue of responsibility: How are humans to respond ethically to the power placed in their hands by modern technology? Such a question has personal, professional, and policy dimensions. At the personal level, quantitatively and qualitatively enhanced choices, with expanding knowledge production relevant to such choices (scientific research and consumer reports), place existential pressures on individuals to increase conscious reflection. The principle of free and informed consent appears to require not only that medical professionals inform the subjects of human experimentation about the risks and benefits of their participation but also that medical patients of all sorts become reflective participants in their own treatment—and that consumers of any technological goods or services weigh multiple costs and benefits as if they were

engineers designing their lives. Are such demands both reasonable and possible?

At the professional level, scientists and engineers, falling under similar existential pressures to expand the conscious exercise of responsibility, have formulated codes of conduct for technical practices related to both research and design. In engineering ethics, for instance, the primacy of protecting public safety, health, and welfare is now a well-established general principle. In what sense, however, are engineers qualified to make such judgments? Does technical expertise provide any basis for determining appropriate levels of public safety, health, or welfare?

Finally, at the level of public policy, responsibility takes two closely related forms. Policy for science and technology seeks out the best ways to fund or regulate developments in science and technology. Science and technology for policy searches for the best ways to bring scientific knowledge to bear on political decision making while making technological power most effectively available for political action. Responding to and exemplifying these dual drives scientific and technological research agencies such as the U.S. National Science Foundation, the Human Genome Project, and the National Nanotechnology Initiative have created specific programs to promote ethical reflection on the creation and use of new scientific knowledge and technological products, processes, and systems.

Again speaking broadly, it is possible to identify two fundamental attitudes toward this spectrum of ethical and political issues. One attempts to explain modern technology as rooted in human nature and culture (engineering philosophy of technology), the other interprets modern technical methods and effects as deformations of human action, however preferable in particular instances to those of nature (humanities philosophy of technology). The engineering approach in its expansive confidence calls in one way or another for more and better technology, the humanities approach in its restrictive questioning for some relinquishment or delimitation of technology. The tensions between such alternative attitudes repeatedly come to the fore in analysis of such key concepts as privacy, risk, participation, and the environment, and in assessments of new opportunities in virtual reality construction, biotechnological design, and nanotechnological research and development.

There is also a tendency for the engineering school to make alliances with the Anglo American analytic tradition in philosophy, and for the humanities school to find a convenient partner in the European phenomenological tradition. The former, viewing technology as a complex amalgam of artifacts, knowledge, activities, and volitions, each with diverse structural features scattered across historical epochs and societal contexts, prefers to deal on a case-by-case basis with one technology after another. The latter strives for bolder generalizations about technology as a whole, at least across each historical or societal context. From the phenomenological perspective, too great an emphasis on individual technological rocks can obscure the extent to which such geological specimens are constituents of mountains extended in both space and time.

METAPHYSICAL ISSUES

The attempt to speak of *technology* rather than *technologies* rests on an attempt to identify some inner or essential feature of diverse technologies. This hypothetical essential feature may be termed *technicity*. One can then immediately note that, before the modern period, technicity was at a minimum scattered throughout and heavily embedded within a diversity of human engagements, and indeed that philosophy took a stand against any separating of technicity from its embedding context. Plato's argument in the *Gorgias* is precisely an argument against disembedding *techne* from social or cultural contexts and traditions, not to mention ideas of the good. For Aristotle, *techne* is an intellectual virtue, and thus properly subordinate to the flourishing of human nature. What is distinctive about modern philosophy, by contrast, is the attempt, beginning with Galileo Galilei, Francis Bacon, and René Descartes to disembed technics from particular human activities, to study them in systematic ways, and thus to create technology.

John Stuart Mill in his *Logic* (1843) already assumes the success of this disembedding project when he explains the practical value of science. For Mill the rationality of any art is grounded in a corresponding science.

> The art proposes to itself an end to be attained, defines the end, and hands it over to the science. The science receives it, considers it as a phenomenon or effect to be studied, and, having investigated its causes and conditions, sends it back to art with a theorem of the combinations of circumstances by which it could be produced. Art then examines these combinations or circumstances, and according as any of them are or are not in human power, pronounces the end attainable or not.

(*LOGIC*, BOOK 6, CH. 12, SECTION 2)

Remarkably, Mill's analysis does not recognize art (or traditional technics) as including any knowledge of means. Art is concerned solely with determining an end, to achieve which it deploys appropriate means as determined by science. It is the scientific study of means that constitutes what even during Mill's lifetime was coming to be called technology. Modern technicity may thus be defined as a systematic or scientific study of means that suspends examination of ends. Does such an approach have distinctive social and cultural implications, independent of any particular technologies and contexts?

Among the first philosophers to analyze such a disembedding of means from ends was Ortega. In the English translation of his *La rebelión de las masas* (1929), Ortega writes that "[t]hree principles have made possible [the] new world: liberal democracy, scientific experiment, and industrialism. The two latter may be summed up in one word: technicism" (1939, p. 56). Ortega himself actually uses the word *técnica*, but the term *technicism* is significant, and this in fact constitutes one of its first English occurrences with this sense. (Before the 1930s, *technicism* simply meant excessive reliance on technical terminology. The previous decade Max Scheler used the cognate *Technizismus* to name the industrial ethos.)

As part of a further "Meditación de la técnica" (1939), Ortega outlined a historical movement from the chance inventions that characterize archaic societies, through the trial-and-error techniques of the artisan, to the scientific technologies of the engineer. According to Ortega, the difference between these three forms of making lies in the way they create the means to realize a human project—that is, in the kind of technicity involved. In the first epoch, technical discoveries are accidental; in the second, techniques emerge from intuitive skill. In both instances they are preserved and elaborated within the confines of myth and craft traditions. In the third, however, the engineer undertakes scientific studies of technics and, as a result, "prior to the possession of any [particular] technics, already possesses technics [itself]" (*Obras*, 5:369). It is this third type of technicity that constitutes modern technicism (and here Ortega himself uses the term *tecnicismo*).

But technicism, understood here as the science of how to generate all possible technical means, disembedded from any lived making and using, creates a unique challenge. Before the modern period human beings were commonly limited by circumstances, within which they inherited a way of life and the technical means to achieve it. Now, however, they are given in advance many possible ways to live and a plethora of technical means but little in the way of a substantive vision of human flourishing. "To be an engineer and only an engineer is to be everything possibly and nothing actually," all form and no content (*Obras*, 5:366). There is in the midst of modern technicism what Ortega describes as a hidden ethical challenge to imagination and choice. Insofar as people can be anything they want, why should they take the trouble to be any one thing at all? Will not some extranatural motivation (not to say fanaticism) not be needed to help Buridan's cyborgs select among (rejecting some) the equally liberal options that surround them?

According to Heidegger modern technology is a challenge not just to ethics but to ontology. For Heidegger (1954) scientific technics constitutes a new kind of truth: truth not as correspondence, not as coherence, and not as functional knowledge, but as disclosure or revelation. Technology discloses Being in a historically unique way: as *Bestand* or resource. A castle constructed with traditional technics on a cliff overlooking the Rhine makes more fully present than before the stone that invests the landscape with its particular contours, while it sets off the curve of the river against the backdrop of its walls and towers. It invites people to settle near and experience the particularities of this place. By contrast, a poured concrete, hydroelectric power station compels the river to become an energy resource and converts the landscape into, not a place of human habitation, but a machine for the generation of electricity. It encourages people to draw on its energy for multitasking business in production and travel. The distinctly modern technicity that manifests itself in the disclosure of nature as resource Heidegger names *Gestell* (enframing).

Gestell at first sight appears to be a human work, something human beings in the course of history have chosen to practice for their own benefit. It gives them power over nature. However, as it digitalizes nature physically (dimensioned vectors), geographically (longitude and latitude), chemically (molecules, atoms, and subatomic particles), and biologically (genetic mapping), it also transforms language (computer signal processing) and art (pixel imaging) so that impact outstrips original intentions. Hidden in the midst of *Gestell* is Being as event, that which lets this dominating transformation come to pass. *Gestell* is at once destiny and, precisely because it appears so clearly to be the result of a human activity, an obscuring of the transhuman imparting of a destiny that is its ground.

In the same year that Heidegger's *Die Frage nach der Technik* appeared, Jacques Ellul published *La Technique*,

later translated into English as *The Technological Society* (1954). For Ellul, too, what is happening is something transhuman, or at least transindividual, the emergence of a new social order in which people give themselves up to the systematic analysis of actions into constituent means that are then evaluated in terms of output/input metrics. The scientific analysis of techniques extends technoscientific methods into economics, politics, education, leisure, and elsewhere creating what he calls the technical milieu. After the milicux of nature and of society, technology is the third great epoch of human history. Ellul's characterology of this new reality—describing its rationality, artificiality, self-directedness, self-augmentation, indivisibility, universality, and autonomy—reveals the technical milieu as something more than simply human. Although more hospitable to human biological existence, it nevertheless also manifests certain inexorable laws of artifice (such as those of economics). Just as the natural milieu once provided a framework for human life, a differentiated but overriding order to which human beings adapted in a variety of ways, so now a much more homogeneous technical milieu presents itself, not simply as a realm of freedom that human beings have constructed, but as that which also constructs and constrains them even when they fail to recognize it.

FROM METAPHYSICS TO ETHICS

Efforts to make phenomenological metaphysics fruitful for ethics can be found in the work of two German American philosophers, Hans Jonas and Albert Borgmann. Jonas's (1966) work begins with a fundamental inquiry into the phenomenon of life, arguing that in the organic world there emerges a new kind of being. For Jonas the key features of human inner life (introspection and subjectivity) are present in embryo in the most primitive organisms, and in metabolism there emerges the primordial form of freedom. In metabolism a detachment enters the world insofar as being becomes distinguished from physical identity. However, in the materialism of modern science this unique reality is easily overlooked. Adopting a teleological approach to ontology, Jonas argues that only from the perspective of the more fully realized freedom manifest in humans can the reality of the organic as a whole be recognized for what it is. On this ontological basis Jonas (1984) undertakes an extended philosophical scrutiny of the technological projects of nuclear weapons and biomedical health care. In the presence of technical powers to end or alter human life Jonas reformulates the Kantian categorical imperative as: "Act so that the effects of your action are compatible with the permanence of genuine human life" (p. 11). Such a reformulation of the fundamental deontological principle constitutes an attempt at the re-embedding of technology in moral philosophy.

More broadly and in sustained dialogue with a range of discussions about the place of technology in human affairs, Borgmann's (1984) work draws a fundamental distinction between two kinds of artifice and action. On the one side are technological devices that obscure their inner functions to deliver without engagement commodities for easy and effortless consumption. This constitutes what Borgmann calls the device paradigm, an ideal type at which the products and processes of modern technology aim. On the other are focal things and practices whose workings are more transparent and that demand of their users some reordering of interests if they are to be used. The model for the first is the central heating system that only needs its thermostat set, for the second the wood-fired hearth.

In a series of studies arguing the nondeterminist importance of material culture to ethics and politics, Borgmann (1992, 1999) calls on citizens in the high-tech world to reconsider their ways of life to develop a deeper sense for the possibilities of human flourishing in the midst of liberal options for self-determined self-fulfillment. For Borgmann the ideal is not a forced return to the past but a voluntary recovery of the commanding presence of things in the technological present. As he concludes in a volume devoted to the critical assessment of his thought:

> Science makes reality ever more transparent, and technology makes it more and more controllable. But at the end of our inquiries and manipulations there is always something that reflects rather than yields to our searchlight and presents itself as given to us rather than constructed by us. It is intelligible not because we have seen through it or designed it but because it speaks to us [in the form of] an unforethinkable and uncontrollable reality. (Higgs, Lights, and Strong 2000, pp. 368–369)

It is such a reality to which human flourishing is ultimately in thrall even in the midst of its highest exercises of insight and mastery.

EPISTEMOLOGICAL ISSUES

Epistemology has often been treated as a stepchild in the philosophy of technology family of philosophical interests. Technological forms of knowledge are commonly

thought to be derivative of scientific knowledge, so that any attempt to bring the theory of knowledge to bear in the examination of technology has regularly been part of a discussion of the relation between technology and science. At the same time this common privileging of science has been philosophically criticized, although the criticism has taken different forms in the European phenomenological and in the Anglo American analytic philosophical traditions.

From a phenomenological perspective the argument has been that technology is not so much applied science as science is theoretical technology. In his historico-philosophical studies of the scientific and technological revolutions of the seventeenth century and after, for instance, Jonas (1974) argues that from its origins modern science was animated by a technological interest that gives it an inherently applicable or technological character. Related studies of the dependency of science on technological instrumentation, from Galileo's telescopes to particle accelerators and PCR (polymerase chain reaction) machines (e.g., see Ihde 1991), suggest that science might even be described as applied technology. This approach to the epistemology of technology has parallels with the pragmatic tradition of conceiving scientific knowledge in fundamentally instrumentist terms (see Hickman 2001). The Venezuelan phenomenologist Ernesto Mayz Vallenilla (2004) likewise offers a more Husserlian-based but complementary effort to describe the unique epistemological features of what he calls meta-technical instruments.

From the analytic perspective there has been more of an effort to identify distinctive types of knowledge operative in technology. Summarizing the results from such an approach, Mitcham (1994) draws attention to at least four types of distinctly technological knowledge: sensorimotor skills, technical maxims (including rules of thumb and recipes), descriptive laws or technological rules (which take an "if A then B" form), and technological theories (either grounded in scientific theory or bringing scientific method to bear on human-technology interactions). German philosophers of technology such as Hans Lenk, Gunter Ropohl, and Bernhard Irrgang, all associated with the VDI promotion of philosophical reflection on technology, are pursuing efforts to develop epistemological analyses of the engineering sciences. And Joseph C. Pitt (2000) makes a determined effort to identify the distinctive forms of technological and engineering knowledge, drawing especially on the careful analyses of aeronautical engineering history by Walter G. Vincenti (1990) to argue that engineering design possesses its own cognitive features.

Important issues for any theory of technological knowledge remain the characterization of whatever basic epistemic criteria might be analogous to those operative in science such as truth, simplicity, coherence, and explanation. There may be distinctive technological forms of such criteria. But two major candidates for uniquely technological criteria are effectiveness and efficiency. Certainly, many propositions of engineering knowledge are assessed in terms of effectiveness and efficiency more than truth or explanation. A further epistemological challenge is to explicate the distinctive character of models and modeling in the technological and engineering contexts. The relevance of such epistemological analyses nevertheless remains of problematic relevance to ethics and politics.

EMPIRICAL, ANTHROPOLOGICAL, AND POLICY TURNS

Concern for the adequacy of metaphysical definitions of technology—and perhaps exhaustion with endless ethical and political difficulties (with hopes that new approaches might prove more fruitful)—has given rise to what has been called an empirical turn in the philosophy of technology. As advocated by the Dutch philosophers Peter Kroes and Anthonie Meijers, this program argues that "philosophical reflection should be based on empirically adequate descriptions reflecting the richness and complexity of modern technology" (2000, p. xix) and promotes a greater analysis of what technologists and engineers actually do over any extended exegesis of texts, whether those of other philosophers of technology or even engineers and technicians. As such, a natural alliance has developed with social constructivist approaches to science, technology, and society studies in the pursuit of richer metaphysical or ontological understandings of artifacts, epistemological analyses of technical practice, and even ethical decision making among professional engineers. From the perspective of Jozef Keulartz et al. (2002), this also provides a solid opportunity for advancing a pragmatist ethics for technological culture.

Two topics of prominence in the empirical turn from the interpretation of texts to the interpretation of technical artifacts have been those of design and function. Design is often identified as the essence of engineering, and there have been numerous technical studies of design methodology. At the same time engineering design must be distinguished from aesthetic design as well as design by means of evolutionary processes in nature. Even

within the realm of engineering design, studies such as those by Vincenti (1990), Louis Bucciarelli (1994), and Richard Buchanan and Victor Margolin (1995) have very different implications for assessing proposals for consumer, green, sustainable, or participatory design. With regard to technical functions, analyses have focused on the relation between functions in organisms, social institutions, and artifacts; on the relation between functional and physical descriptions of artifacts; and on the extent to which functions are determined by design or use.

A different sense for new beginnings has emerged in relation to prospects in the development of the new fields of bioengineering and biotechnology—especially when applied to humans. The leader in this case is the medical scientist and philosopher Leon Kass, the chair of the Bush administration's President's Council on Bioethics. In his turn Kass has tried to go outside the boundaries of standard bioethics in at least four ways: to promote thinking that enrolls more than professional bioethicists, that does more than piecemeal or specialized analyses, that references human nature as a norm, and that builds toward policy results. As in *Beyond Therapy: Biotechnology and the Pursuit of Happiness* (2003), Kass et al. at the council seek to raise broad issues about what it means to be human in the presence of possibilities for the reengineering not just of the external world but of the inner world of human birth, growth, and experience. He has been especially concerned about the possibilities for the deformation of humanity not from above by totalitarian governmental use of technology but from below by positive consumer endorsement of behaviors that would from a traditional perspective be assessed as temptations.

Beyond the policy-oriented work of Kass and colleagues, policy questions have become increasingly central not just as aspects of ethical responsibility but as issues in their own right. What precisely is technological policy, as opposed to technological politics? Does policy decision making take different forms in relation to science and to engineering? How are policies to be formulated and assessed?

The extent to which these turns in the philosophy of technology will define its future are questions that the professional community must examine. Any such examination will also need to include a self-criticism that considers the special responsibilities of a regionalization in philosophy that, more than the philosophy of science or of art, has as part of its heritage public responsibilities and a large measure of ethical concerns.

See also Applied Ethics; Aristotle; Artificial Intelligence; Bacon, Francis; Bioethics; Categorical Imperative; Computationalism; Computer Ethics; Descartes, René; Engineering Ethics; Environmental Ethics; Epistemology, History of; Ethics, History of; Galileo Galilei; Genetics and Reproductive Technologies; Heidegger, Martin; Human Genome Project; Kant, Immanuel; Machine Intelligence; Marxist Philosophy; Metaphysics, History of; Mill, John Stuart; Neo-Kantianism; Ortega y Gasset, José; Philosophy of Biology; Philosophy of Science, History of; Philosophy of Science, Problems of; Veblen, Thorstein Bunde.

Bibliography

Beck, Ulrich. *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main, Germany: Suhrkamp, 1986. Translated by Mark Ritter as *Risk Society: Toward a New Modernity*. London: Sage, 1992.

Borgmann, Albert. *Crossing the Postmodern Divide*. Chicago: University of Chicago Press, 1992.

Borgmann, Albert. *Holding on to Reality: The Nature of Information at the Turn of the Millennium*. Chicago: University of Chicago Press, 1999.

Borgmann, Albert. *Technology and the Character of Contemporary Life: A Philosophical Inquiry*. Chicago: University of Chicago Press, 1984.

Buccarelli, Louis. *Designing Engineers*. Cambridge, MA: MIT Press, 1994.

Buchanan, Richard, and Victor Margolin, eds. *Discovering Design: Explorations in Design Studies*. Chicago: University of Chicago Press, 1995.

Bunge, Mario. "Technology: From Engineering to Decision Theory." In *Treatise on Basic Philosophy, Vol. 7, Philosophy of Science and Technology*, Part 2, *Life Science, Social Science, and Technology*. Boston: D. Reidel, 1985.

Ellul, Jacques. *La technique ou l'enjeu du siècle*. Rev. ed. Paris: Economica, 1990. Translated by John Wilkinson as *The Technological Society* (New York: Vintage, 1965). See also two supplements: *Le système technicien* (Paris: Calmann-Lévy, 1977), translated by Joachim Neugroschel as *The Technological System* (New York: Continuum, 1980); and *Le bluff technologique* (Paris: Hachette, 1988), translated by Geoffrey W. Bromiley as *The Technological Bluff* (Grand Rapids, MI: W. B. Eerdmans, 1990).

Feenberg, Andrew. *Alternative Modernity: The Technical Turn in Philosophy and Social Theory*. Berkeley: University of California Press, 1995.

Feenberg, Andrew. *Critical Theory of Technology*. Rev. ed. New York: Oxford University Press, 1991.

Feenberg, Andrew. *Questioning Technology*. New York: Routledge, 1999.

Feenberg, Andrew. *Transforming Technology: A Critical Theory Revisited*. New York: Oxford University Press, 2002.

Goldman, Steven. "No Innovation without Representation: Technological Action in a Democratic Society." In *New Worlds, New Technologies, New Issues*, edited by Stephen H.

Cutcliffe et al., 148–160. Bethlehem, PA: Lehigh University Press, 1992.

Heidegger, Martin. "Die Frage nach der Technik." In *Vorträge und Aufsätze*. Pfullingen, Germany: Neske, 1954. Translated by William Lovittas "The Question concerning Technology," in *The Question concerning Technology and Other Essays* (San Francisco: Harper and Row, 1954).

Hickman, Larry. *Philosophical Tools for Technological Culture: Putting Pragmatism to Work*. Bloomington: Indiana University Press, 2001.

Higgs, Eric, Andrew Lights, and David Strong, eds. *Technology and the Good Life?* Chicago: University of Chicago Press, 2000.

Ihde, Don. *Instrumental Realism: The Interface between Philosophy of Science and Philosophy of Technology*. Bloomington: Indiana University Press, 1991.

Ihde, Don. *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press, 1990.

Jonas, Hans. *The Phenomenon of Life: Toward a Philosophical Biology*. New York: Harper and Row, 1966.

Jonas, Hans. *Philosophical Essays: From Ancient Creed to Technological Man*. Englewood Cliffs, NJ: Prentice Hall, 1974.

Jonas, Hans. *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Frankfurt: Germany, 1979. Translated by Hans Jonas, with the collaboration of David Herr, as *The Imperative of Responsibility: In Search of an Ethics for the Technological Age* (Chicago: University of Chicago Press, 1984).

Keulartz, Jozef, Michiel Korthals, Maartje Schermer, and Tsjalling Swierstract, eds. *Pragmatist Ethics for a Technological Culture*. Dordrecht, Netherlands: Kluwer Academic, 2002.

Koen, Billy Vaughn. *Discussion of the Method: Conducting the Engineer's Approach to Problem Solving*. New York: Oxford University Press, 2003.

Kroes, Peter, and Anthonie Meijers, eds. *The Empirical Turn in the Philosophy of Technology, Vol. 20, Research in Philosophy and Technology*. Amsterdam, Netherlands: Elsevier, 2000.

Marcuse, Herbert. *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Boston: Beacon Press, 1964.

Martin, Mike W., and Roland Schinzinger. *Ethics in Engineering*. New York: McGraw-Hill, 1983. 4th ed., 2005.

Mayz Vallenilla, Ernesto. *Fundamentos de la meta-técnica*. Caracas, Venezuela: Monte Avila, 1990. Translated by Carl Mitcham as *The Foundations of Meta-Technics*. Lanham, MD: University Press of America, 2004.

Mitcham, Carl, ed. *The Encyclopedia of Science, Technology, and Ethics*. 4 vols. Detroit, MI: Macmillan Reference, 2005.

Mitcham, Carl. *Thinking through Technology: The Path between Engineering and Philosophy*. Chicago: University of Chicago Press, 1994.

Mumford, Lewis. *The Myth of the Machine, Vol. 1, Technics and Human Development*. New York: Harcourt, Brace, 1967.

Ortega y Gasset, José. "Meditación de la técnica." In *Ensimismamiento y alteración*. Buenos Aires, Argentina: Espasa-Calp, 1939.

Ortega y Gasset, José. *Revolt of the Masses*. New York: W. W. Norton, 1932.

Pitt, Joseph C. *Thinking about Technology: Foundations of the Philosophy of Technology*. New York: Seven Bridges Press, 2000.

President's Council on Bioethics. *Beyond Therapy: Biotechnology and the Pursuit of Happiness*. New York: Regan, 2003.

Sclove, Richard. *Democracy and Technology*. New York: Guilford Press, 1995.

Shrader-Frechette, Kristin S. *Risk and Rationality: Philosophical Foundations for Populist Reforms*. Berkeley: University of California Press, 1991.

Vincenti, Walter G. *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*. Baltimore, MD: Johns Hopkins University Press, 1990.

Winner, Langdon. *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. Cambridge, MA: MIT Press, 1977.

Winner, Langdon. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: University of Chicago Press, 1986.

Carl Mitcham (1996, 2005)

PHONOLOGY

"Phonology" is the branch of linguistics concerned with the articulatory and auditory domain of grammar—that is, with the theory of what John Langshaw Austin (1962) called phonetic acts. Its subject matter links with but is distinct from that of syntax, semantics, and pragmatics. It covers the forms in which the sounds of words are kept in memory and the manner in which the motions of speech organs are shaped by grammar.

Unlike syntax, semantics, and pragmatics (but like closely related morphology), phonology has been largely ignored by philosophers. On the whole, philosophers consider the fact that natural languages are primarily spoken rather than written as of little interest for what Michael Dummett (1986) calls a "philosophical explanation" of language. This attitude stems largely from the mistaken but widely held view that spoken signs are arbitrary sounds whose individuating traits are those of noises. On that view, utterances contemplated apart from their semantic and syntactic features are merely tokens of acoustical types, bereft of grammatical properties, fully described by the physics of noises, and available for human communication simply because humans can perceive and produce them; there is nothing intrinsically linguistic about them. Nor is this attitude an accident. Historically, philosophers have had little incentive to reflect on the sound of language. Most belong to traditions that admit no crucial differences (except perhaps those that pertain to pragmatics) between natural lan-

guages and notational systems developed by scientists, mathematicians, or philosophers for the elaboration of their theories. Such notational systems have a syntax and a semantics of sorts, but they have no phonology. Their constituent elements are typically spatial ideographs that share little with the phonological structures of natural languages. Studying language with such a bias offers few reasons, if any, to focus on what is spoken rather than written. It can, however, entrap one in a false conception of linguistic signs, so false, in fact, as seriously to weaken philosophic doctrines built on it.

Phonology rests on a series of presumptions—each supported by a vast body of observations—that together entail that the sounds of natural languages are not arbitrary human noises, on a par with grunts or snorts, whose individuating attributes lie entirely outside the domain of grammar.

The first such presumption is that when people acquire a word they memorize the underlying phonological representation of that word, a representation that defines—but often only partially—how the word is pronounced. These representations have the structure of linearly arrayed discrete timing positions that are assigned pointers to articulatory organs (lips, blade of tongue, dorsum of tongue, root of tongue, velum, vocal cords) implicated in the pronunciation of the word, and pointers to actions these organs execute during speech. The first timing position for the English *pin,* for instance, points to the lips, the vocal cords, the velum, full closure of the first, stiffening of the second, and nonlowering of the third.

A second presumption is that these pointers (called phonological features) on timing positions are drawn from a finite repertoire, common to all languages, and that they are combined within and across timing positions in rule-governed ways. Some rules are common to all languages and reflect innate linguistic endowments, others are language specific and reflect the influence of linguistic exposure. No language, for instance, avails itself of nasal snorts. French admits rounding of the lips in combinations of features that English excludes (thus the sound *ü* in French but not in English). Korean, unlike English (except for *h*), admits aspiration in underlying phonological representations. German, unlike English, admits initial sequences corresponding to sounded *k* followed by sounded *n.* All languages assemble features in similar (three-dimensional-like) structures.

A third presumption is that underlying phonological representations, in isolation or when compounded in complex words, are subject to rule-governed processes that add, subtract, or modify phonological features, which group them into syllables, feet, and prosodic words, which assign stresses and (in some languages) tones, and which ultimately yield final articulatory instructions, so-called surface phonological representations related to, but often very different from, the underlying representations in memory. Processes of this sort account for the fact that, for example, *leaf* occurs as *leavz* (with *v* instead of *f*) in the plural, or that *serene* is pronounced differently when alone than when a constituent of *serenity,* or that *p* gets aspirated in *pin* though not in *spin.* The details of these rules, the manner of their application, the universality of their formats, and the options fixed by different languages are all objects of intense research and controversies. But the evidence in behalf of their reality seems irrefutable.

Phonology is of philosophic interest, not only because it brings into question analogies between contrived notational systems and natural languages, but also because it raises conceptual issues of its own. Two can be mentioned here.

First, individual spoken utterances are analyzable in both acoustical and phonological terms. No generalizable exact correspondences between these two analyses are known. None may be forthcoming. For instance, nothing acoustical corresponds to word division. How can this dualism be reconciled? Is there a cogent sense in which the objects of speech production are the same (or belong to the same types) as those of speech perception? Offhand, the problem resembles that raised by other events amenable to multiple descriptions. But in this case solutions must be attuned to much that is already understood about both phonology and acoustics. It is not a simple task.

Second, phonological theory associates multiple representations with each utterance—including an underlying representation and a surface one—and it describes them all in the same notation. Surface representations can be conceptualized as instructions (or intentions) to move articulators in certain ways; their ontological status, though unclear, is at least comparable to that of other familiar cases. Not so the other phonological representations. They do not have familiar analogues. The semantic domain of phonological notation therefore cannot be ontologically homogeneous. Furthermore, part of that domain is deeply perplexing.

See also Austin, John Langshaw; Dummett, Michael Anthony Eardley; Philosophy of Language; Pragmatics; Semantics; Syntax.

Bibliography

Anderson, S. R. *Phonology in the Twentieth Century.* Chicago: University of Chicago Press, 1985.

Austin, J. L. *How to Do Things with Words.* Cambridge, MA: Harvard University Press, 1962.

Bromberger, S., and M. Halle. "The Ontology of Phonology." In *On What We Know We Don't Know,* edited by S. Bromberger. Chicago: University of Chicago Press, 1992.

Dummett, M. In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson,* edited by E. LePore. Oxford and New York: Blackwell, 1986.

Kenstowicz, M. *Phonology in Generative Grammar.* Cambridge, MA: Blackwell, 1994. An introduction to the field and a complete bibliography.

Quine, W. V. O. *Word and Object,* chap. 3. Cambridge, MA: MIT Press, 1960.

Sylvain Bromberger (1996)
Morris Halle (1996)

PHRONÊSIS

See Appendix, Vol. 10

PHUSIS

See *Nomos and Phusus*

PHYSICALISM

Physicalism, of which materialism is a historical antecedent, is primarily an ontological doctrine concerning the nature of reality and, specifically, mental reality. It is the view that reality is ultimately constituted or determined by entities—objects, events, properties, and so on—that are physical. This thesis is often combined with a claim about the explanatory supremacy of physical theory (physics).

Any formulation of physicalism raises the question, What is meant by "physical"? It is difficult to formulate a conception of the physical that is neither too strong, making physicalism obviously false, nor too weak, making physicalism trivially true. For example, what is physical may be simply identified through the language of physics. However, a problem arises over the conception of physics appealed to. Current physics seems too narrow because future extensions of physics would not count as physical; but the idea of a completed physics is too indeterminate because there is no clear idea of what that physics might include. One could attempt to characterize the physical in more general terms such as having spatial location or being spatiotemporal. However, this threatens to make physicalism trivially true because mental phenomena seem clearly to have spatial location in virtue of having subjects—persons—who have bodies. It may be preferable to appeal to the idea of a completed physics. Although at any particular time people may not know exactly what is physical and what is not (because they may not know whether they have completed physics), nevertheless what is physical is all and only what a completed physics countenances.

There are two main types of physicalist theses. First, there is eliminative materialism, or physicalism. According to this there are not, and never have been, any mental entities, events, properties, and so forth. Strictly speaking, this is not a view about the nature of mental reality. Second, there is a group of doctrines that fall under the general heading of identity theories, some of which are stronger than others. These can be divided into two main categories. The stronger doctrines may be called type-type identity theories, or type physicalist theories (Armstrong 1968, Lewis 1966, Place 1956, Smart 1959), and the weaker doctrines may be called token identity theories, or token physicalist theories (Davidson 1970, Macdonald 1989, Macdonald and Macdonald 1995).

Physicalist theories need to account for at least two different kinds of mental phenomena. First, there are the sensations, such as color experiences, pains, tingles, itches, and the like, which are typically, and perhaps essentially, identified in terms of how they feel to their subjects. Then there are the intentional states or events, such as beliefs, hopes, desires, and thoughts, which are typically, and perhaps essentially, identified in terms of their intentional contents, or their "aboutness." For example, a person's belief that water is transparent has the intentional content, *water is transparent*; a content that represents the world around that person in a certain way, irrespective of whether the world happens to be that way. One of the biggest difficulties for physicalism is accounting for both of these kinds of mental phenomena. In the late twentieth and early twenty-first centuries, philosophers have expressed skepticism as to whether a thoroughgoing physicalist position is possible, and have maintained that physicalism (either token, or type) is true of at most one of these two kinds of mental phenomena (Chalmers 1996, Kim 1998).

TYPE PHYSICALISM

Consider any mental phenomenon, such as being in pain now, or thinking right now that water is transparent. It is possible to talk about this phenomenon as an individual

occurrence of a certain kind in the mental life of a person and discuss its properties. It is also possible to talk about the kind of phenomenon—pain, or the thought that water is transparent—of which this event is an individual instance. Physical phenomena too can be discussed in both of these ways. Type physicalism is the view that the mental types, properties, or kinds under which mental phenomena fall are identical with physical types, properties, or kinds. For example, pain—that type of phenomenon, occurrences of which are individual pains—is identical with some single type of physical phenomenon such as C-fiber stimulation.

Type physicalism has its origins in the doctrines espoused by the logical positivists and central-state materialists (Place 1956, Smart 1959). It is a strong form of physicalism because it is reductionist. Many who endorse it believe that nothing short of it counts as a proper physicalism. They argue that even if it is in practice impossible for sentences containing mental terminology to be translated into or replaced by sentences containing physical and topic-neutral terminology, any view that holds that all mental phenomena are physical phenomena, but mental properties or kinds are not physical properties or kinds, is not worthy of the name "physicalism."

THE FIRST OBJECTION TO TYPE PHYSICALISM. Type physicalism suffers from two serious objections. The first, from phenomenal properties, specifically concerns sensations such as color experiences, pain, afterimages, and the like. It is that phenomena of these kinds or types have "felt" properties, such as being reddish, stabbing, or vivid, whereas phenomena of physical types do not. Given this, and given Leibniz's principle of the indiscernibility of identicals, it follows that sensation types are not identical with physical types because the phenomena that fall under them do not share all of the same properties. A variant of this objection focuses on the distinctive point of view a subject has on its own experiences: A subject knows what it is like to have experiences in a way that others do not, and this subjective mode of access reveals the phenomenal aspect of the experience, whereas an "other"-oriented point of view does not (Nagel 1974).

One response is to argue that the problem is purely conceptual and does not threaten physicalism, which is an ontological view about what sorts of things there are in the world, not a view about concepts (Levine 2001, Loar 1997, Tye 1999). Consider the type-type identity expressed by "Brain State B is the red-feeling sensation." To the objection that such identities are false because first-person access to experiences reveals them to have properties that physical states do not have, the response is that the apparent difference in properties arises from the distinctive nature of human experiential (or phenomenal) concepts alone. Certain concepts, such as the concept *red-feeling sensation* (or *reddish sensation*), are ones that can only be possessed by being put into direct contact with experiences that fall under them, without the mediation of other information or concepts that one might have of those states. Because the phenomenal concept *red-feeling sensation* enables subjects to be put in direct contact with their own red-feeling experiences in a way in which no concept of Brain State B could do, it puts them in a position to recognize directly and in an immediate way their own phenomenal red-feeling experiences. Possession of the concept *Brain State B* could not put any subject in a position to recognize directly and in an immediate way its own red-feeling experiences. So, even having met the experiential requirement on the possession of the concept *red-feeling experience*, a subject might be under the illusion of thinking that the red-feeling sensation has a property that Brain State B lacks. Whether or not this response succeeds depends on whether, in acquiring a new concept, such as the concept *red-feeling sensation*, one learns a new fact about the world that one did not know before, despite being in possession of the concept *Brain State B*.

THE SECOND OBJECTION TO TYPE PHYSICALISM. The second objection to type physicalism is that from multiple realizability. This claims that mental kinds or properties may be realized in physically diverse types of ways, hence there is no single physical property with which a given mental property may be identified. The point is that even if each mental property were in fact to be realized by a single physical one, it is possible for it to be realized by physically diverse ones. The reason is that the introspective and behavioral basis upon which attributions of mental properties are typically made is silent on the potential internal physical realizers of them. Given the claim that identical things are necessarily identical, the mere possibility that a given mental property should be realized by a physical property other than that which in fact realizes it is sufficient to refute the claim that that mental property is identical with any physical property that may realize it. This objection is not independent of a modal argument that trades on the thesis that identical things are necessarily identical (Kripke 1980). This begins with the conceivability of a mental state type's existence in the absence of any physical type of state, and argues that, because what is conceivable is possible, it is possible that mental state types could exist in the absence of any

physical state type. The argument concludes that, because it is possible that mental types should exist in the absence of any type of physical phenomenon, mental state types are not identical with any type of physical phenomenon. A version of this argument is held to be particularly decisive against type physicalism with respect to sensation states.

One response is to argue that mental types are identical with disjunctions of physical types. For example, pain may not be identical with C-fiber stimulation, but it may be identical with the disjunctive property, C-fiber stimulation, or A-fiber stimulation, or … , and so on (properties picked out by predicates formed by disjoining predicates that pick out all the possible physical realizers of mental properties). However, it is unclear whether these are bona fide properties. They do not have a unity of their own, viewed from a physical perspective; and it is arguable that a reason is needed, apart from the fact that they all realize a given mental property, to think that they are properties in their own right (Macdonald 1989).

Against this, it might be claimed that because any given mental *predicate* may correlate with an indefinite number of physical *predicates*, this may pose problems for formulating laws connecting mental with physical properties; but it does not follow that there is not a single physical *property* that is the extension of a given mental predicate. Mental properties are identical with the physical properties picked out by disjunctive physical predicates, but their autonomy is secured by their participation in real regularities, and so they do have a unity of their own, despite being identical with disjunctive physical properties (Antony 2003).

In a similar but more radical vein, it might be claimed that although there are mental and physical predicates, there really are only physical properties, so there are no type-type identities of any kind that might be problematic for physicalism (Kim 1998). This reductionist response avoids the problem of multiple realizability altogether, but only by taking an eliminativist stand on mental properties. Alternatively, it might be claimed that the only type-type identities licensed by physicalism are species-specific (as in, for example, that expressed by "pain in humans is identical with C-fiber stimulation"). None of these claims is unproblematic: the first, because it threatens to make mental properties non-nomic, which seems to undermine the commitments of type physicalism; the second, because it is eliminativist; and the third, because it leaves questions such as "What makes pain in humans and pain in dogs both pain?" unanswered.

TOKEN PHYSICALISM

Many consider one or the other of the above objections to be decisive against type physicalism and have opted instead for a weaker view: token physicalism. According to this, each individual mental event or phenomenon is identical with some physical event. One influential version of this is the view known as anomalous monism (Davidson 1970). Token physicalism is compatible with the multiple realizability of mental properties by physical ones because it is not committed to the view that each individual occurrence of a given mental kind is identical with an occurrence of the same type of physical phenomenon. It also appears to avoid the objection from phenomenal properties in its original form because it can concede that mental kinds have associated with them felt aspects with which no physical kinds are associated. To the objection that mental events are not identical with physical events because it is no part of the nature of any physical event that it have a felt aspect, the following reply can be made. If token physicalism is true, no physical event is essentially of a mental type; but given that it is of a given mental type, it has what is essential to being of that type. Thus, if this pain is identical with this C-fiber stimulation, then it is not essentially a pain. However, given that it is, as it happens, a pain, it has (though not essentially) what is essential to being of that type, namely being felt.

Without an explanation of how mental types relate to physical ones, token physicalism threatens to succumb to the charge that it is dualist because it countenances the existence of nonphysical properties or types. A common strategy is to advance a supervenience doctrine concerning the relation between mental and physical properties, according to which physical properties, although distinct from mental ones, in some sense determine them (Hellman and Thompson 1975). There are many varieties of supervenience theses. One difficulty is in finding a thesis strong enough to do justice to the claim that physical properties determine mental ones without being so strong as to entail identities between mental and physical properties or types, and with these, reducibility. Another, related problem, is explaining how it could be that mental types or properties supervene on physical ones in a way that dispels the worry that mental properties have no causal powers of their own.

See also Causal Closure of the Physical Domain; Dualism in the Philosophy of Mind; Functionalism; Mind-Body Problem; Nonreductive Physicalism; Philosophy of Mind; Reduction; Reductionism in the Philosophy of Mind.

Bibliography

Antony, L. "Who's Afraid of Disjunctive Properties?" *Philosophical Issues* 13 (2003): 1–21. An argument for the identity of mental properties with disjunctive physical ones.

Armstrong, D. M. *A Materialist Theory of Mind*. London: Routledge & Kegan Paul, 1968. Defense of a type-type identity theory of the mental and physical.

Block, N., ed. *Readings in the Philosophy of Psychology*. Vol. 1. Cambridge, MA: Harvard University Press, 1980. Articles on type-type identity theories, token identity theories, reductionism, and functionalism.

Chalmers, D. *The Conscious Mind*. New York: Oxford University Press, 1996. A conceivability argument against type physicalism, specifically with regard to sensation states.

Davidson, D. "Mental Events." In *Experience and Theory*, edited by Lawrence Foster and J. W. Swanson. Amherst, MA: University of Massachusetts Press, 1970: 79–101. Highly influential argument for a token identity of the mental and physical.

Hellman, G., and F. W. Thompson. "Physicalism: Ontology, Determination, Reduction." *Journal of Philosophy* 72 (1975): 551–564. Defense of a supervenience doctrine.

Jackson, F., R. Pargetter, and E. Prior. "Functionalism and Type-Type Identity Theories." *Philosophical Studies* 42 (1982): 209–225. Discussion of the relation between functionalism and type-type identity theories.

Kim, J. "Supervenience as a Philosophical Concept." *Metaphilosophy* 12 (1990): 1–27. Discusses supervenience as a covariance relation and its relation to reduction.

Kim, J. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998. A reductionist physicalist position regarding intentional states only that suggests handling the multiple realizability objection either by denying that mental properties are genuine or by maintaining that mental properties are identical with disjunctive physical ones.

Kripke, S. *Naming and Necessity*. Oxford: Basil Blackwell, 1980. A modal argument against type-type and token identity theories.

Levine, J. *Purple Haze*. Oxford: Oxford University Press, 2001. A defense of the view that phenomenal property problems with type physicalism are purely conceptual.

Lewis, D. "An Argument for the Identity Theory." *Journal of Philosophy* 63 (1966): 17–25. An argument for a type-type identity theory.

Loar, B. "Phenomenal States." In *The Nature of Consciousness*, edited by N. Block, O. Flanagan, and G. Guzeldere, 597–616. Cambridge, MA: MIT Press, 1997. A defense of the view that phenomenal property problems with type physicalism are conceptual.

Macdonald, C. *Mind-Body Identity Theories*. London: Routledge, 1989. A survey of type-type and token identity theories, and a defense of a token identity theory.

Macdonald, C., and G. Macdonald. "How to Be Psychologically Relevant." In *Philosophy of Psychology: Debates on Psychological Explanation*, edited by C. Macdonald and G. Macdonald. Oxford Basil Blackwell, 1995. Advances an argument for a token identity theory that attempts to explain how mental properties could supervene on physical ones in a way that dispels the worry that mental properties have no causal powers of their own.

McGinn, C. "Can We Solve the Mind-Body Problem?" *Mind* 98 (1989): 349–366. An argument for the view that the mind-body problem cannot be solved because of the conceptual disparities between the mental and the physical.

Nagel, T. "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435–450. Argues against identity theories of the mental and the physical on the basis of a disparity between the mental and physical in points of view.

Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47 (1956): 44–50. A defense of central-state materialism for sensations.

Smart, J. J. C. "Sensations and Brain Processes." *Philosophical Review* 68 (1959): 141–156. A defense of central state materialism for sensations.

Tye, M. "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion." *Mind* 108 (1999): 705–725. A defense of the view that phenomenal property objections to type physicalism are due to a confusion about phenomenal concepts.

Cynthia Macdonald (1996, 2005)

PHYSICAL REALITY

See *Philosophy of Science, History of; Philosophy of Science, Problems of*

PHYSICOTHEOLOGY

"Physicotheology" is the aspect of natural theology that seeks to prove the existence and attributes of God from the evidence of purpose and design in the physical universe. The argument is very ancient, but it is from the Greeks that its medieval and modern forms principally spring. Socrates revolted against the materialist tendencies of earlier philosophers, and his pupil Plato sought to show that the order and harmony exhibited in the world sprang from the action of mind. Plato argued that since matter cannot move itself, motion is evidence of the presence of mind in nature. All the activity and change in the world have their origin in a supreme mind that moves itself and creates subordinate souls or gods, the heavenly bodies. The outer sphere of the universe is set in motion by the direct action of the changeless, transcendent God. Aristotle expounded more emphatically a teleological or purposive view of nature in which the members of the hierarchy of natural classes in the universe seek to realize their beings according to their stations. This perspective presupposes a rational design, a universal aspiration to fulfillment, and in one passage Aristotle describes God as the perfect being whom all things desire.

The theological aspects of Greek views of nature passed into later science and were readily translated into Christian thought. The animistic view of natural knowledge may be seen in the work of Galen (second century), for whom the processes of the human body are divinely planned. During the earlier medieval period the natural world appeared to the eye of faith to be a scene of symbols and ciphers veiling moral and spiritual doctrines. Later medieval philosophers were fond of discerning marks of providential direction in the operations of nature, and Thomas Aquinas rests one of his proofs of the existence of God upon the cooperation of all types of natural objects to make the order of the world and the pointing of that order to an intelligent author who devised it. There was abundant recourse to this argument during the later Middle Ages.

SEVENTEENTH CENTURY

The golden age of the Argument from Design was the two centuries following the rise of science in the seventeenth century, and it took place principally in England. The new philosophy of nature abandoned belief in the intrinsic teleology of physical objects. In place of the analogy with a creator of living organisms or an artist creating works of beauty it substituted the analogy of an inventor and manufacturer of elaborate machines. The new scientists combined faith in the sovereignty of God in nature and belief in the mechanistic bases of phenomena by conceiving the deity as the skillful contriver of instruments, a consummate engineer.

In England the doctrine was promoted by two trends of thought, the Baconian gospel of controlled observation and the revival of Greek atomism. The Baconian method inspired groups of inquirers in London and Oxford to collect a mass of detailed information in which they saw the confirmation of their religious faith; and it was the descriptions of the zoologists and botanists, such as Nehemiah Grew and Francis Willoughby, that strikingly illustrated the marvelous skill of the Creator. The second doctrine, the atomic, or corpuscular, theory of matter, incurred charges of materialism and atheism from moralists because of its association with Epicurean atomism, and in order to divide themselves from these imputations the virtuosi were intent on attaching theological conceptions to the elements of the material world. They were also acutely sensitive to the materialist dangers in the dualist philosophy of René Descartes. Neither their religion, which formed the frame of all their thought, nor their reason, which saw the marks of purpose and planning in nature, allowed them to accept the idea that the world originated in the chance combination of material atoms. Ralph Cudworth, in his *True Intellectual System of the Universe* (1678), spoke for all the experimental philosophers when he argued at length that greater perfections and higher degrees of being cannot possibly arise out of senseless matter. The ancient metaphysics of cause, securely rooted in Christian theology, precluded any doctrine of natural evolution, and it is interesting to observe that when writers on biology mentioned the hypothesis that creatures have been produced by "millions of trials," as did John Ray, the hypothesis was dismissed with scorn. Species had been finally and completely created. There was no conceivable alternative to the Argument from Design.

ROBERT BOYLE. The Argument from Design was expounded with eloquence by Robert Boyle (1627–1691). In his multifarious researches he was concerned with the evidence of benevolent and ingenious contrivance in nature and found on all sides "curious and excellent tokens and effects of divine artifice." But first we may notice the way in which he associated the atomic view of matter with supernatural power. In embracing the corpuscular or mechanical philosophy, he writes, he is far from supposing with the Epicureans that atoms accidentally meeting in an infinite vacuum were able by themselves to produce a world and all its phenomena. The philosophy he pleads for teaches that in the beginning God gave motion to matter and so guided the motions of its parts as to "contrive them into the world he designed they should compose," establishing those rules of motion that we call the laws of nature (*The Excellence and Grounds of the Mechanical Philosophy,* 1674). In *The Origin of Forms and Qualities* (1666) he explains that the diversity of bodies must arise from motion and that motion in the beginning was from God, for it is not inherent in matter.

In the realm of animate nature Boyle points to numerous instances of ingenious design, such as the human eye, and he constantly speaks of organisms as engines or machines. For him an animal as a whole is an engine, and each part of it is a subordinate engine excellently fitted for some subordinate use. Here he reverts to a famous analogy that in a simpler context goes back to Cicero and even to Xenophon, the analogy of organisms and the world with clocks and watches. In Boyle's day, clocks were the most complex examples of machines available for comparison, and he takes a celebrated clock as a model of the machine of the world, the cathedral clock at Strasbourg, in which "the several pieces making up that curious Engine are so fram'd and adapted, and are

put into such motion, as though the numerous wheels and other parts of it knew and were concerned to do its Duty" (*The Usefulness of Experimental Natural Philosophy,* 1663). The popularity of the analogy between a watchmaker and the author of nature in the following age issued largely from the writings of Boyle.

ROBERT HOOKE. During the early years of the Royal Society proofs of design multiplied. Robert Hooke's *Micrographia* (London, 1665) disclosed the astonishing beauty and ingenuity of the minute creatures revealed by the microscope, and in his Cutlerian lectures he spoke of the divine providence that in the eye "has so disposed, ordered, adapted, and empowered each part so to operate as to produce the wonderful effects which we see."

JOHN RAY. Before the end of the century there appeared treatises by the greatest zoologist of the age that were wholly devoted to the evidences in nature of the existence of God. John Ray's *The Wisdom of God Manifested in the Works of Creation* was first published in London in 1691, enlarged in three later editions before Ray's death in 1705, and reprinted more than twenty times by 1846. In the preface he declares that his discourse will serve to demonstrate the existence of the Deity and illustrate his principal attributes, his infinite power and wisdom. He proceeds to show the futility of attributing the world to the operation of chance events; it manifests all the marks of deliberate creation. Inanimate bodies are reviewed in order, the system of the stars and their planets, and the services performed for animals and man by water, air, fire, meteors, rain, and winds. Passing to regions of life, he ascends through the vegetable and animal kingdoms, discovering everywhere a complex arrangement of parts that contribute to the welfare of the plant or animal and to the uses of man.

Ray was too close an observer of nature to accept the crude doctrine that organisms are complex machines constructed by a divine watchmaker. His physicotheology borrowed from Cudworth the theory of plastic nature or vital force by which the growth, adaptation, and instinctive activities of living creatures are directed. This plastic virtue acts sympathetically, without reason, informing the movements of material bodies. Ray therefore diluted his physicotheology with an immaterial energy, a form of animism. But the plastic nature is nonetheless a subordinate instrument of divine providence, although it transcends the operations of local motion. Its relative independence of the immediate direction of God allowed Ray to meet a cardinal difficulty in the Argument from Design; he could accept the aberrations of nature without making the Deity responsible for them. Faced with this problem, Boyle has preserved his mechanistic view of creation by asserting that the irregularities we find in nature may serve ends that lie concealed in God's unsearchable wisdom.

Ray presided over the subsequent course of the Argument from Design, and theologians drew freely on his *Wisdom of God in Creation.* They studied also his *Three Physico-theological Discourses* (London, 1692), which supports the biblical narratives of the creation, the deluge, and the final dissolution of the world by arguments from natural philosophy.

ISAAC NEWTON. The appearance of Isaac Newton's *Principia* in 1687 had provided the argument with a great deal of new material. Natural theology became absorbed by the cosmology of the *Principia,* and preachers and poets acclaimed the almighty hand that "poised, impels and rules the steady whole." Newton's great treatises offered at many points notable arguments for the belief that the universe is the work of an intelligent being; indeed, Newton told Richard Bentley that in writing the *Principia* he had had an eye upon arguments for a belief in a deity, and in the *Opticks* he declared that the main business of natural philosophy was to deduce causes from effects until we arrive at the First Cause, which cannot be mechanical. In the *General Scholium* added to the second edition of the *Principia* and in the *Queries* of the Latin translation of the *Opticks* (1706), he set forth the religious conceptions that underlay his mathematical physics of the universe. Why is it, he asks, that all planets move the same way in concentric orbits? What prevents the stars from falling on one another? And, with a glance at the evidence of Boyle and Ray, how, he asks, did the bodies of animals come to be contrived with so much art? Whence, in short, arose all that order and beauty that we see in the world? Does it not appear that there is a Being incorporeal, living, intelligent, and omnipresent, who created the world?

For Newton, however, the admirable system of nature was not imposed by the deity upon an infinitely complex material mechanism; immaterial forces were introduced into the heart of the mechanism of nature. Newton asserted the atomic theory of matter in the manner of Boyle: It seemed probable that God in the beginning formed matter in solid, massy, hard, impenetrable, movable particles, but the forces that cause the particles to cohere and to form larger bodies are immaterial. It is not the business of experimental philosophy to discuss the nature of these forces, but it is clear that they provide

the world with its structure and order. They could not have arisen from chaos by the mere laws of nature; the wonderful uniformity of the planetary system, for example, must be the effect of choice and must proceed from the counsel and dominion of an intelligent and powerful Being.

Other fundamental principles of Newton's system of physics are associated with theology. Absolute space is immovable, homogeneous, indivisible, and distinct from matter; like other thinkers of the time, Newton accorded space some of the attributes of God. He described infinite space as the boundless sensorium of the omnipresent God, whereby he perceives all things. Motion also presupposes a metaphysical agent, for if the motion of moving bodies is derived from the impact of bodies already in motion, some other principle was necessary for putting bodies in motion in the first instance and for conserving the motion of those in movement. The agent must be an all-powerful immaterial being, for pressure is constantly brought to move bodies throughout the universe. Furthermore, the variety of motion is always decreasing because at every impact between bodies, some motion is lost. It must be renewed by an immaterial power.

EIGHTEENTH CENTURY

The natural theology of Newton crowned the Argument from Design, and by the beginning of the eighteenth century the main stock of theory and of evidence on which the argument relied had been provided. Numerous writers repeated and enforced the case pronounced by John Locke that the works of nature everywhere sufficiently evidence a Deity. Prominent among those who vindicated the conclusions of the great men of the seventeenth century were the Boyle lecturers in the series instituted in Boyle's will with the purpose of confuting atheism. The lectures were inaugurated in 1692 by Richard Bentley, a renowned scholar who corresponded with Newton while preparing the lectures. In his letters to Bentley, Newton maintained that there are many features of the universe that cannot be explained in terms of mechanical principles, and he went on to assert that the cause that constructed the planetary system cannot be blind and fortuitous but must be one very skilled in mechanics and geometry. Bentley faithfully reported these opinions in the lectures.

CLARKE AND LEIBNIZ. The second Boyle lecturer was the celebrated Samuel Clarke, who delivered the course called "A Demonstration of the Being and Attributes of God" in 1704, an excellent survey of the accepted picture, with some fresh touches. His famous correspondence with Gottfried Wilhelm Leibniz on natural theology was published in 1717; he probably received advice from Newton in composing his replies, and the letters further reveal Newton's position on such important topics as the divinity of space. But the vital interest of this correspondence is the conflict between Leibniz's conception of nature as mechanical, determined, self-sufficient, and self-perpetuating and the doctrine, defended by Clarke, of God's providential guidance of the world. Leibniz rejected the Newtonian contention that God corrects aberrations of the cosmic order, such as certain inequalities of planetary motions, as a watchmaker cleans and mends a watch—a view that implies that the creation of the system was imperfect and that God is lacking in foresight. Clarke, on his part, accused Leibniz of restricting the liberty of God to act as he will, independently of the laws of nature; indeed, but for his constant intervention, the world would lapse into chaos. The doctrine of supernatural intervention began to recede from the physics of astronomy and found its home before the end of the century in the realms of geology and biology.

JOSEPH BUTLER. The deists, in their war against revelation, caught at the notion that God, having created the world in the distant past, had left it to the action of the laws of nature. Deism provoked a stream of hostile pamphlets and treatises, but orthodox churchmen who opposed deism continued to harp on law, order, and design and the divine artificer. The greatest of these apologists was Bishop Butler. *The Analogy of Religion* (1736) shows that he had closely studied Newton, but his natural theology rises above that of other writers of the age in its candid recognition of the defects of nature, which he ascribes to our ignorance of God's purposes.

Another Boyle lecturer was William Derham, whose *Physico-theology* (London, 1713) and *Astro-theology* (London, 1715) rehearsed the testimony of Ray and of Newton at prodigious length, with some superficial reflections of his own. Many other utterances must be passed over. It is interesting to observe the large number of writers who discussed Clarke's (and Newton's) theology of space.

DAVID HUME. In the later years of the eighteenth century, natural theology encountered the penetrating criticism of David Hume, although few scientific theologians were shaken by it. In the *Dialogues concerning Natural Religion,* published posthumously in 1779, Hume exploded the logic of the Argument from Design, especially in the form in which it was presented by the disci-

ples of Newton, such as the Scottish mathematician Colin Maclaurin. Hume confronted the analogy between the maker of a machine and the maker of the world with the point that while scientists like Nicolas Copernicus and Galileo Galilei made fruitful use of reasoning by analogy, the associations between cause and effect that provided the material of their arguments were derived from observation. The inference from machines and their makers to a world and its maker is not parallel. Order, arrangement, or the adjustment of final causes is not by itself any proof of design, but only insofar as it has been seen to be produced by design; since we have no experience of the invention and production of a world or of nature, we cannot maintain that an orderly universe must arise from thought or art. For all that we can know a priori, matter may contain the source of order within itself.

Hume attacked this argument by a reductio ad absurdum. If we are confined to speculative, a priori explanations of the origins of the world, they can lead to disturbing conclusions. Some natural philosophers have found nature to resemble an organism, a vegetable or an animal, and its origin ought to be ascribed to generation and vegetation rather than to reason or design. When the analogy with the manufacturers of machines is pressed, we might infer that several deities combine in contriving and framing the world. Hume now introduced fatal evidence against the belief in a benevolent Creator. The curious artifices of nature embitter the life of every living being. "The whole presents nothing but the idea of a blind nature, impregnated by a great vivifying principle, and pouring forth from her lap, without discernment or parental care, her maimed and abortive children." Faced with these difficulties the defender of traditional doctrine in the *Dialogues* is compelled to admit that belief in a beneficent Creator of the world cannot be rationally sustained. The sources of such a belief are "temper and education," and the defender of the Argument from Design falls back on utilitarian supports; belief in divine design promotes morality.

NINETEENTH CENTURY

Hume's *Dialogues concerning Natural Religion* failed to confound the deep-seated prepossessions of the natural theologians, nor were they discomposed by the refutation of the Argument from Design by Immanuel Kant in the *Critique of Pure Reason* (1781).

WILLIAM PALEY. At the turn of the century the argument was revived in William Paley's *Natural Theology* (1802). It marks the apotheosis of the analogy between a watch and a natural object, opening, in fact, with the discovery of a watch lying on a heath. The instrument must have been made by a being who comprehended its construction and designed its use. If we suppose that the watch contains a mechanism by which it can produce another watch (a supposition that exhibits the deficiency of the mechanical analogy), our admiration of the maker's skill will increase. Paley proceeds to describe numerous examples of natural contrivances, drawn from anatomy, physiology, botany, and entomology: the eyes of fish, animals, and men, the construction of the ear, the webbed feet of water birds, the elongated tongue of the woodpecker, and a catalog of other instances. These marvels of adaptation prove the existence of a superhuman designer, God. As for the suffering that nature displays, Paley attempts to minimize the spectacle; the pain of animals, he thinks, is exaggerated, and their happiness outweighs their pain. Even venomous bites and the preying of one species on another are shown to be necessary features of benevolent design.

BRIDGEWATER TREATISES. Leading men of science in this period duly acknowledged the action of divine providence in natural phenomena. In geology John Playfair and Sir Charles Lyell discovered in the adjustment of the strata of the earth to the accommodation of living creatures clear proofs of divine foresight, and James Prescott Joule saw in the interconvertibility of natural forces evidence of the sovereign will of God. The most sustained defense of the Argument from Design was advanced in the Bridgewater Treatises of the 1830s. Eight men of science, four of whom were clergymen, were chosen to discharge the intentions of the earl of Bridgewater to explore "the Power, Wisdom and Goodness of God, as manifested in the Creation." These writers added a wealth of new information from astronomy, physics, chemistry, and anatomy to the old theses of Ray and Derham, and they outstripped Paley in showing how all aspects of nature have been thoughtfully arranged for the comfort of the world's inhabitants and especially for man. John Kidd, Regius professor of medicine in the University of Oxford, in *On the Adaptation of External Nature to the Physical Condition of Man* (London, 1833); Peter Roget, secretary to the Royal Society, in *On Animal and Vegetable Physiology Considered in Relation to Natural Theology* (London, 1834); and William Buckland, professor of geology at Oxford, in *On Geology and Mineralogy* (London, 1836), showed how climates have been fitted to the character of the various races of humankind, horses invented for man's transport, minerals for his adornment, and water for his ablutions. In short, much of the reasoning of these

writers recalls that of the lady who praised the goodness of the Creator in causing a great river to flow through the main cities of Europe.

Sir Charles Bell, the most distinguished physiologist of the time, in his *The Hand, Its Mechanism and Vital Endowments as Evincing Design* (London, 1833), argues that species were successively created to fit the conditions of geological epochs, changes in their anatomy being deliberately shaped to meet the circumstances of the creatures' life. Man is the center of a magnificent system, which has been prepared for his reception by a succession of revolutions affecting the whole globe, and the strictest relation is established between his intellectual capacities and the material world. The celebrated William Whewell, in his *Astronomy and General Physics considered with Reference to Natural Theology* (London, 1833), makes play with the ambiguous sense of the word *law,* a common procedure among scientific theologians of the period, confusing the idea of uniform sequence with the idea of legal and moral law; the confusion arose from Whewell's demonstration that the laws of nature, terrestrial and celestial, provide evidence of selection, design, and goodness. The tenacity and ingenuity with which the scientists vindicated the sovereignty of God over nature are illustrated in Charles Babbage's *Ninth Bridgewater Treatise* (London, 1837), where by means of his calculating machine he proves mathematically that miraculous interruptions of scientific laws can be predicted, and that the Being who called the laws into existence must have chosen them with the breaches of continuity in view.

The Bridgewater Treatises marked the final stage of the general confidence of men of science in the old natural theology, although religious thinkers long continued, and still continue, to appeal to it. However, when the treatises appeared the classical form of the Argument from Design was weakening. Whewell had difficulty in understanding the bearing of cosmology upon the support and comfort of sentient creatures, and geologists, led by James Hutton and Lyell, were abandoning the view that there had been sudden changes in the crust of Earth, occasioned by the mediation of God. The catastrophic picture of geological change was yielding to the uniformitarian view in which the laws operating at present could in the slow process of ages have caused all the changes of the past. The range of natural law in time and space was being extended, but the scientists failed to account for the processes by which fresh species had originated, and faith in the periodic agency of the Creator was encouraged.

CHARLES DARWIN. Charles Darwin opened a notebook on the transmutation of species in 1837, and in the unpublished "Essay on Species" of 1844 he proposed the machinery by which new species might result from the natural selection of fortuitous variations. The notion of special creations, he recorded in his private notebook, explains nothing, and the Essay concluded with a forceful reductio ad absurdum of the Argument from Design. *The Origin of Species* (1859) brought a wealth of material to substantiate the theory of natural selection in the evolution of species and in adaptations of the organs of living creatures to their circumstances, and it is interesting to see Darwin using the same examples that Paley did to show evidence of contrivances resulting not from purpose but from chance. By abolishing both transcendent and immanent teleology, Darwin undermined the ground on which physicotheology had stood since the seventeenth century. Yet in the last chapter of the *Origin* Darwin himself assumed a First Cause, though not a beneficent one, and he declared in 1873 that the impossibility of conceiving that this great and wondrous universe arose through chance seemed to him the chief argument for the existence of God. In the end, however, Darwin became a complete agnostic, as is shown most clearly in the unexpurgated edition of his *Autobiography* (first published in 1958).

J. S. MILL. In his *Three Essays on Religion,* published posthumously in 1874, J. S. Mill allowed some value to the Argument from Design, for the world contains marks of deliberate contrivance, and our experience of such devices is associated with an intelligent mind. Mill here seems to have exposed himself to Hume's objections against arguing from cases within the world to the world as a whole. But Mill recognized many features of the world that are incompatible with beneficent design, and he thought that God may be a limited Being circumscribed by matter and force. Mill maintained that if Darwin's doctrine of evolution were shown to be valid it would greatly weaken the evidence for the work of a divine intelligence in nature.

SUPPORT FROM SCIENTISTS. Other scientists contrived to fit the theory of natural selection into the frame of divine purpose. Samuel Houghton, a fellow of the Royal Society, described expressions of supernatural intentions in his book *Principles of Animal Mechanics* (London, 1873). Another book that exercised great influence was Professors P. G. Tait and Balfour Stewart's *The Unseen Universe* (1875), in which it was contended that science upheld the ideas of religion on the transcendental

world and its connection with the physical world. A succession of eminent scientists proclaimed that nature is the sacred book of God. The most popular and, it must be added, most muddle-headed work that applied evolution to theistic principles was Henry Drummond's *Natural Law in the Spiritual World* (1883). The tendency of these scientific writers was to assert the view that Darwin's theory had deepened and widened the belief in the operation of purpose in nature, a view that was characterized as misplaced zeal by those who stood more closely to Darwin's findings.

A number of physicists of the period also employed classical versions of the design argument. The Celestial Engineer was reinstated by O. M. Mitchell in his widely read *The Orbs of Heaven* (4th ed., London, 1853) at the middle of the century, in which, after the manner of Newton, the deity is invoked to secure the stability of the solar system. It was a notion of the earlier apologists that the identical character of the fundamental materials of the physical world in all parts of the natural order indicated the action of an intelligent maker. The idea had been adopted by Sir John Herschel in his *Study of Natural Philosophy* (1830), and it was now revived by the greatest mathematical physicist of the age, James Clerk Maxwell. At the meeting of the British Association in 1873, he pointed out that every type of molecule in the universe is identical with every other type; a molecule of hydrogen, whether it occurs in Sirius or in Arcturus, executes its vibrations in precisely the same time. No theory of evolution accounts for this identity, for the molecule is not subject to change. Its similarity to other molecules proves that it is the product not of chance but of design. It is a manufactured article, and because they are the work of a Creator, the foundation stones of the material universe remain, whatever catastrophes may occur in the heavens. Even the argument from miracles reappeared in the *Natural Theology* (London, 1891) of a later mathematical physicist, Sir George Stokes: "If the laws of nature are in accordance with God's will, he who willed them may will their suspension." Stokes assumed that God's action in nature cannot be detected within the laws of physics but by interventions from beyond. *Natural Theology* embraces the arguments of physicotheology in the period.

A monumental exposition in a modern setting of the Argument from Design appeared in *Philosophical Theology* (London, 1928–1930) by F. R. Tennant. Recent discussions of the argument have abandoned the old mechanical analogies and have dwelled on the evidence for various types of vitalism in biology. On these views evolution is guided no longer from outside but by directive activities within organisms. In the human psychosocial phase of evolution these self-directed activities point toward moral ends; history becomes the education of humankind in the fulfillment of God's design. Teleological doctrines of this kind have drawn support from philosophers such as Samuel Alexander and A. N. Whitehead, who contend that the universe is informed by an immanent nisus to divinity. Present theological discussions, however, ignore natural theology, and for contemporary linguistic philosophers the Argument from Design possesses no validity whatsoever and is logically and morally indefensible, although it may serve to heighten religious emotions.

See also Alexander, Samuel; Atheism; Atomism; Boyle, Robert; Butler, Joseph; Cicero, Marcus Tullius; Clarke, Samuel; Copernicus, Nicolas; Cudworth, Ralph; Darwin, Charles Robert; Deism; Descartes, René; Epicureanism and the Epicurean School; Galileo Galilei; God, Concepts of; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Materialism; Matter; Maxwell, James Clerk; Mill, John Stuart; Motion; Motion, A Historical Survey; Newton, Isaac; Paley, William; Plato; Socrates; Teleological Argument for the Existence of God; Tennant, Frederick Robert; Thomas Aquinas, St.; Whewell, William; Whitehead, Alfred North; Xenophon.

Bibliography

In addition to original works mentioned in the text, the reader may consult the following works as guides to the immense literature on the subject.

Adams, Robert. *Leibniz: Determinist, Theist, Idealist.* New York: Oxford University Press, 1994.

The Bridgewater Treatises on the Power, Wisdom and Goodness of God as Manifested in the Creation. 12 vols. London, 1833–1836.

Burtt, E. A. *The Metaphysical Foundations of Modern Physical Science.* London: K. Paul, Trench, Trubner, 1925.

Carré, M. H. "The Divine Watchmaker." *Rationalist Annual* (1965): 83–91.

Clark, Ronald. *The Survival of Charles Darwin: A Biography of a Man and an Idea.* New York: Random House, 1984.

Cohen, Bernard, and George Smith, eds. *The Cambridge Companion to Newton.* Cambridge, U.K.: Cambridge University Press, 2002.

Cohen, I. B., and H. M. Jones, eds. *Science before Darwin.* London: Deutsch, 1963.

De Beer, Gavin. *Charles Darwin.* London: Nelson, 1963. Bibliography.

Flew, Antony, and Alasdair MacIntyre, eds. *New Essays in Philosophical Theology.* New York, 1955.

Gillispie, Charles C. *Genesis and Geology.* Cambridge, MA: Harvard University Press, 1951.

Hodge, Jonathan, and Gregory Radick, eds. *The Cambridge Companion to Darwin.* Cambridge, U.K.: Cambridge University Press, 2003.

Hooykaas, R. *Natural Law and Divine Miracle.* Leiden: Brill, 1959.

Hurlbutt, R. M., III. *Hume, Newton and the Design Argument.* Lincoln: University of Nebraska Press, 1965.

Jolley, Nicholas, ed. *The Cambridge Companion to Leibniz.* Cambridge, U.K.: Cambridge University Press, 1994.

Koyré, Alexandre. *From the Closed World to the Infinite Universe.* Baltimore: Johns Hopkins Press, 1957.

LeMahieu, Dan. *The Mind of William Paley.* Lincoln: University of Nebraska Press, 1976.

Metzger, Henri. *Attraction universelle et religion naturelle chez quelques commentateurs anglais de Newton.* Paris, 1938.

Mossner, E. C. *Bishop Butler and the Age of Reason.* New York: Macmillan, 1936.

Norton, David, ed. *The Cambridge Companion to Hume.* Cambridge, U.K.: Cambridge University Press, 1993.

O'Connor, David. *Hume on Religion.* New York: Routledge, 2001.

Pendelhum, Terence, and Ted Honderich, eds. *Butler: The Arguments of the Philosophers.* London: Routledge, 1985.

Pilkington, Roger. *Robert Boyle; Father of Chemistry.* London: Murray, 1959.

Raven, C. E. *John Ray.* Cambridge, U.K., 1950.

Raven, C. E. *Natural Religion and Christian Theology.* Cambridge, U.K.: Cambridge University Press, 1953.

Redwood, John. *Reason, Ridicule, and Religion: The Age of Enlightenment in England, 1660–1750.* Cambridge, MA: Harvard University Press, 1976.

Ruse, Michael. *The Darwinian Paradigm: Essays on Its History, Philosophy and Religious Implications.* New York: Routledge, 1989.

Rutherford, Donald. *Leibniz and the Rational Order of Nature.* New York: Cambridge University Press, 1995.

Sargent, Rose. *The Diffident Naturalist: Robert Boyle and the Philosophy of Experiment.* Chicago: University of Chicago Press, 1995.

Strong, E. W. "Newton and God." *Journal of the History of Ideas* 13 (2) (April 1951): 147–167.

Westfall, R. S. *Science and Religion in Seventeenth Century England.* New Haven, CT: Yale University Press, 1958.

Westfall, Richard. *The Life of Isaac Newton.* New York: Cambridge University Press, 1993.

Young, Robert. *Darwin's Metaphor: Nature's Place in Victorian Culture.* Cambridge, U.K.: Cambridge University Press, 1985.

Meyrick H. Carré (1967)
Bibliography updated by Christian B. Miller (2005)

PHYSICS, PHILOSOPHY OF

See *Philosophy of Physics*

PHYSICS AND THE DIRECTION OF TIME

Our experience of the temporality of things seems to be an experience of a radically asymmetric feature of the world. Although we do know some things about what the future will be like, we have an access to past events that is not given to us of events in the future. We take ourselves as having memories of the past but not of the future and as having records of the past but not of the future. In our explanatory accounts of what happens in the world, we explain present and future by reference to what happened in the past, but we typically do not explain the past by referring to the future. We take it that there is causation in the world—that events determine one another to occur. But, intuitively, we think of causation and determination as directed from past to future. We have distinctive *attitudes* to past and future. Of the past we may have regrets, for example, but our concern for the future will be rather things such as anxiety or anticipation. So profound are these apparent differences between past and future that they are often promoted into the realm of deep metaphysics. Sometimes it is argued that the past is *fixed*, subject to some version or another of immutability, whereas the future remains merely a domain of open possibilities. In an even more extreme view it is argued that what is past has a *determinate reality* whereas the future remains a realm to which we cannot even attribute any kind of determinate being.

One might take these asymmetric features of time as irreducible, primitive properties of the world. And one might take our awareness of these features as somehow direct and not further explicable. Alternatively, one might argue for some basic, asymmetrical, metaphysical aspect of time as grounding all the asymmetries discussed above. For example, there are proposed *branching* models of the world in which a tree of possibilities is constantly pruned into a single actuality as time goes on and the present moves inexorably into the future. One problem with any such model is the need to respect the results of modern physics, especially special and general relativity, so as to reconcile the usual assumption in the metaphysical models of a unique global present with the denial of any such objective feature of the world in the relativistic accounts of spacetime. Another alternative would be to take a temporally asymmetric notion of causation as primitive and argue that all the other intuitive asymmetries follow from the fundamental asymmetry of causation.

NATURALISTIC THEORIES OF TIME ASYMMETRY

On the other hand, one could seek for some *naturalistic* account of the temporal asymmetries. Here, one looks at what our best available scientific theories tell us about the actual physical structure of the world in the hopes of finding some physical process characterized by fundamental physics that could serve to *ground* or *explain* the existence and nature of the fundamental temporal asymmetries. Much work has been done in this direction, but more needs to be done to make such a naturalistic account fully convincing. It is to this approach that this entry is directed.

Physics presents us with a *paradox*. Although most of its fundamental laws are often alleged to be time-reversal invariant and unable, it is therefore claimed, to ground any fundamental asymmetry in time of processes in the world, physics also describes a number of alleged time-asymmetric features of the world at a very general level. Measurement processes in quantum mechanics are often alleged to be asymmetric in time. We see radiation outbound in spherical waves from accelerated charged particles but not spontaneous collapsing spheres of radiation converging on a particle and accelerating it. Subtle experiments seem to show that some of the interactions of the elementary particles show asymmetries in time that may, indeed, require positing a fundamental law governing them that itself describes a lawlike asymmetry in time for the world. Most importantly, thermodynamics seems to reveal to us a world that is time asymmetric. A metal bar hot at one end and cold at the other when kept in isolation becomes uniformly warm all over. But an isolated uniformly warm bar does not spontaneously become hot at one end and cold at another.

A naturalistic account of the *direction of time* requires more than finding some physical process that is time asymmetric. It also requires even more than finding a fundamental process that has such time asymmetry. Suppose the weak interactions of the elementary particles obey a time-asymmetric law. How would such a fact be of any use in accounting for our intuitive sense that there are records and memories of the past and not of the future, or that causation proceeds from past to future, or that the past is determinate and fixed and the future a realm of mere possibilities? What is needed from a naturalistic theory of the direction of time is some appropriate connection between the physical facts introduced in the account as grounding the direction of time and those features that characterize our intuitive, deeply rooted sense of the asymmetry of time.

THERMODYNAMIC ASYMMETRIES

The thermodynamic asymmetries, being pervasive elements of our everyday experience, provide the most promising physical basis for a naturalistic account of the direction of time. Here, two fundamental questions must be explored: (1) Why does the world show these deep asymmetries in time of physical processes? (2) How could the existence of these asymmetric processes account for the intuitive asymmetries we attribute to the world in time? Neither question is trivially answered.

The contemporary explanation of the thermodynamic laws starts with the realization that macroscopic objects are composed of a vast number of microscopic constituents. Macroscopic thermodynamic properties, then, are thought of as grounded in such microscopic features of a system as the total energy of its microscopic constituents or the average energy of some one of these. The microscopic constituents are assumed to obey the standard dynamical laws, originally classical dynamics and now quantum mechanics. How the system behaves, then, will depend upon these laws and upon whatever initial conditions are possessed by the microscopic constituents, with the system also subject to such macroscopic constraints as exist (such as a gas being confined to a box).

Probabilistic methods were soon invoked to deal with the behavior of the vast number of microscopic constituents. These led to such theories as the kinetic theory of gases and the more abstract statistical mechanics. One consequence of this new viewpoint was a rethinking of the thermodynamic laws to allow for such possibilities as fluctuations away from the equilibrium state, even for an isolated system. A deep conceptual problem for this theory is the understanding of why the probabilistic posits that are made, and that work so well for prediction and explanation, hold in the world. Are they brute posits to be otherwise unexplained? To what degree can they be extracted from the dynamical laws governing the behavior of the microscopic constituents? Need the fundamental dynamical laws be modified to find an appropriate explanation for the fundamental probabilistic posits (and, perhaps, to solve other outstanding problems as well, such as the nature of the measurement process in quantum mechanics)? Another crucial question is the degree to which the probabilistic posits can be shown consistent with the underlying dynamics and the degree to which they can be weakened with the empirical results still forthcoming.

Furthermore, arguments that have existed from the early days of the theory indicate that introducing proba-

bility into the theory is not, by itself, enough to ground a theory of the direction of time. Probabilistic considerations would suggest that the world we live in should be a world where all systems are at equilibrium, not at all like the world we actually live in with its vast pool of nonequilibrium systems and its parallel movement from nonequilibrium to equilibrium of temporarily isolated systems. In addition, to obtain the desired nonequilibrium results in the theory, the theory's probabilistic posits must be applied in a temporally asymmetric way, being taken as correctly applicable to temporally initial, but not to temporally final, states of isolated systems.

COSMOLOGICAL CONSIDERATIONS

From early days of the theory, cosmology has been invoked as providing the needed supplementary posits. Ludwig Boltzmann's assistant Scheutz suggested the possibility that the cosmos was, in general, in equilibrium, with the part of it with which we were familiar in a local (if large from our perspective) fluctuation away from equilibrium. Our local cosmos, then, was in equilibrium in the past and will be again in the future. Boltzmann added the *anthropic* observation that we could not find ourselves in one of the pervasive equilibrium regions of the cosmos since such a region could not support the flows of energy necessary for life. To this Boltzmann added the additional proposal that the reason we found our region heading toward equilibrium in the future time direction and not in the past time direction is that our very meaning of the *future* direction of time was that the future time direction was determined by that temporal direction in which our local region of the cosmos had a succession of states closer to equilibrium (of higher entropy).

Current cosmology, insofar as it is a discipline open to observation and empirical test, is doubtful of this early cosmological speculation. The current model, rather, is of a universe (at least as far as we know) that is distinctly unsymmetrical in time. In particular, it is posited that there is a singularity in which the cosmos is all *at a single spatial point* in the past time direction some tens of billions of years ago, the so-called Big Bang cosmology.

Even accepting this model of the universe is not enough to get the thermodynamic asymmetries out of the cosmology. Instead, it is generally agreed, one must make a specific assumption about the Big Bang state of the cosmos, that it is a low-entropy, that is, a highly nonequilibrium, state. The usual posit is that the space of the world at the Big Bang is highly smooth, this being for gravitational force the low-entropy condition. The idea is that as matter clumps from uniformly distributed into stars (and galaxies, etc.), the matter, initially in thermal equilibrium, becomes highly nonuniform and in a grossly nonequilibrium state, with hot stars radiating out to cold space. The decrease in the matter's entropy is continually being paid for by the great increase in the entropy of the spatial distribution that has gone from uniform to clumped.

The idea, then, is that the universe as a whole must be posited to have an initial highly nonequilibrium starting point. It is this posit that must be added to the standard probabilistic assumptions to get us the result of a predicted nonequilibrium condition for the world as we find it, and a predicted, temporally asymmetric, approach to equilibrium in the future and not into the past, for system temporarily isolated from their environments. Here the grand cosmic initial condition is being invoked to generate the temporal asymmetry unobtainable from the allegedly time-symmetric dynamical laws alone. Getting the result about the temporarily isolated subsystems of the universe requires a little more, in the form of a demonstration that from the temporally asymmetric behavior of the universe as a whole one can derive, with either no additional temporally asymmetric assumptions at all or with some posited additional asymmetry of dynamics, parallel increase of entropy of so-called *branch systems* in the same time direction in which the entropy of the cosmos as a whole is increasing.

FROM ASYMMETRIES IN TIME TO THE DIRECTION OF TIME

But then there is the second question noted above as well. Why should we think that this pervasive asymmetry of physical systems in time has anything to do with our intuitive sense of the asymmetry of time itself? Once again, the mere fact that there is some asymmetry in time of systems of the world, even a lawlike temporal asymmetry, is not enough to establish the naturalist's claim. What else is required?

Boltzmann hinted at an answer in his famous paper "On Zermelo's Paper 'On the Mechanical Explanation of Irreversible Processes'" where he claimed that what we took to be the future direction of time was just the direction of time in which the entropy of our local portion of the universe was on the increase. In that paper he drew a trenchant analogy between the temporal case and an intuitive spatial asymmetry accounted for by gravity. Originally we might think of space as being asymmetric, with one direction being *down* and its opposite *up*. Eventually, though, we realize that what we call the *down* direction is just the direction of the local gravitational

force. On antipodal points of the earth, the local *downward* directions are directed oppositely to one another. And in a region of space in which there was no gravitational force, there would be no downward direction. Just so in a region of the universe not in equilibrium, Boltzmann maintained, the direction of time in which entropy was increasing would be the future time direction, and in a region of the universe at equilibrium, there would be no future direction of time and no past direction (although there would still be two oppositely directed directions of time).

What makes Boltzmann's remarks about gravity and *down* so plausible? It is the fact that we have in gravity and its consequences a complete explanation of all the facts that we take as distinguishing the downward direction of space from all the other directions. Down is the direction in which, for example, rocks fall. We even have an explanation, invoking the fluid in our semicircular canals, of how it is that we can tell without external observation which direction is the downward one.

The analogous argument in the case of the direction of time would require a sustained argument to the effect that it is the existence of the asymmetric processes of systems in the world described by thermodynamics, and explained by statistical mechanics combined with cosmology, that provides a full explanatory account of all those features of the world that we take as marking out an asymmetric nature to time. What would need explaining is why we have memories and records of the past and not of the future, why we take causation as going from past to future and not the other way around, why we think of the past as fixed and determinate and of the future as a realm of possibility, and, also, of how it is that we think we can tell, without inference, of any pair of events which is the earlier and which the later.

Sometimes it is claimed that the entropic theory of time direction is supported by the fact that we cannot tell of a film of events whether it is being run in the right or the wrong direction except by reference to entropic facts portrayed by the film. Sometimes it is argued against the entropic account that we can tell of events in the world what their order is in time without noting any entropic features of them. Both arguments are misguided. What would be required of an entropic account would be some kind of explanation of all the intuitive asymmetries that constitute our sense of the direction of time, not a demonstration that our judgments of time order are always inferences from directly observed entropic facts.

A number of tentative suggestions have been made in this direction. Hans Reichenbach suggested that records might be analyzed as causal interventions that induced a macro low-entropy change into what would otherwise be a macro high-entropy state. But many records do not fit his paradigm. There is no fully developed extant argument to the effect that the very existence of records, and presumably those mental records we call memories, can have their time asymmetry directly accounted for by the entropic asymmetry of processes in time.

One might argue that such intuitive asymmetries as the direction of causation and the difference in *fixedness* of past and future are derivative from the asymmetry of records, our taking as the fixed and the determining that which we can know to be the case from records. Or, alternatively, one might try to directly account for the asymmetry of causation out of entropic-like facts. An example of that can again be found in a tradition stemming from Reichenbach where it is noted that spatially separated correlated events can often have their correlation explained by some common past event casually connected to both correlated events but not by any such correlating event in the future of the correlated pair (the so-called *fork* asymmetry). Another approach stems from David Lewis. Here, causation is analyzed in terms of counterfactual conditionals. It is suggested that the fact that an even has a numerous extended range of effects in its future, but not in its past, grounds our intuition that there are forward looking but not *back tracking* counterfactual assertions that we accept, and that this underlies our intuitions about the time asymmetry of the causal relation.

Even though no fully worked-out account of these sorts exists, the naturalistic approach to the direction of time remains the only plausible alternative to *metaphysical* accounts. The latter remain hard to explicate, and it is hard to understand how they might provide new insights into the intuitive asymmetries of time. The former, in its usual thermodynamic version, is at least clear in its intentions and of an intrinsic plausibility. The further pursuit of this naturalistic program is well worthwhile.

See also Boltzmann, Ludwig; Lewis, David; Reichenbach, Hans; Time; Time, Being, and Becoming.

Bibliography

Albert, D. *Time and Chance*. Cambridge, MA: Harvard University Press, 2000.

Boltzmann, L. "On Zermelo's Paper 'On the Mechanical Explanation of Irreversible Processes.'" In *Kinetic Theory*. Vol. 2, edited by S. Brush, 239–245. Oxford: Pergamon Press, 1965–1972.

Davies, P. *Physics of Time Asymmetry*. Berkeley: University of California Press, 1974.

Earman, J. "An Attempt to Add a Little Direction to 'The Problem of the Direction Of Time.'" *Philosophy of Science* 41 (1974): 15–47.

Horwich, P. *Asymmetries in Time*. Cambridge, MA: MIT Press, 1987.

Lewis, D. "Counterfactual Dependence and Time's Arrow." In *Philosophical Papers*. Vol. 2. edited by D. Lewis, 32–66. Oxford: Oxford University Press, 1983–1986.

McCall, S. *A Model of the Universe*. Oxford: Oxford University Press, 1994.

Mehlberg, H. *Time, Causality and the Quantum Theory*. Dordrecht, Netherlands: Reidel, 1980.

Mellor, H. *Real Time, II*. London: Routledge, 1998.

Price, H. *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*. New York: Oxford University Press, 1996.

Reichenbach, H. *The Direction of Time*. Berkeley: University of California Press, 1956.

Sklar, L. *Physics and Chance*. Cambridge, U.K.: Cambridge University Press, 1993.

Sklar, L. "Up and Down, Left and Right, Past and Future." *Nous* 15 (1981): 111–129.

Strevens, M. *Bigger Than Chaos*. Cambridge MA: Harvard University Press, 2003

Tooley, M. *Time, Tense, and Causation*. Oxford: Clarendon Press, 1997.

Lawrence Sklar (1996, 2005)

PHYSIS AND NOMOS

See *Nomos and Phusus*

PIAGET, JEAN

(1896–1980)

Jean Piaget, the psychologist and philosopher, was born in Neuchâtel, Switzerland. He studied zoology at the university there and in 1918 received his doctorate for a thesis on the subject of land mollusks in the Valais Alps. He then studied psychology for a year at Zürich and, from 1919 to 1921, abnormal psychology, logic, and the philosophy of science at the Sorbonne. From 1921 to 1925, he was director of studies at the Institut J.-J. Rousseau (now the Institut des Sciences de l'Éducation) in Geneva; he was its assistant director from 1929 to 1932 and became codirector in 1932. In 1925 he was appointed professor of philosophy at the University of Neuchâtel; in 1929, professor of the history of scientific thought at the University of Geneva; and in 1940, professor of experimental psychology and director of the psychological laboratory at Geneva. He served as professor of child psychology at the Sorbonne from 1952 to 1963. From 1955 to 1980 he was director of the Centre International de l'Épistémologie Génétique at Geneva. Piaget also took an active interest in international educational projects. He was director of the Bureau International de l'Éducation from 1929 to 1967 and was associated with UNESCO as its assistant director general.

THOUGHT

Although Piaget is usually considered a psychologist working in the field of child thought, his interests were always, broadly speaking, philosophical. As a young man he read widely in philosophy, and while in Paris he studied with André Lalande and Léon Brunschvicg. Even his earliest work, which appeared between 1925 and 1932, dealt with such topics as thought, causality, moral judgment, and the development of language. His logical and epistemological interests show themselves particularly in his later studies, starting about 1937. By means of simple, although highly ingenious experiments, Piaget set out to make a detailed investigation of the way in which logical, mathematical, and physical concepts develop in the individual. He thus studied experimentally many of the concepts and principles that philosophers had discussed in the past on a purely a priori level. Piaget would say that what he was really doing in this work was reexamining the whole question of the Kantian categories. This reexamination formed for him the basis of a new discipline that he called genetic epistemology.

In a series of studies Piaget examined in some detail the development not only of abstract concepts such as classes, relations, and numbers but also of physical concepts like space, time, speed, atomism, conservation, and chance, all of which he has regarded as constructed from our behavioral activities. In starting from the facts of observable child behavior rather than from adult introspections (or sensations), Piaget differed from such thinkers as Ernst Mach, Karl Pearson, and Bertrand Russell by the importance he attached to the part played by overt activities in building up the conceptual machinery of thought. Throughout his work Piaget placed considerable emphasis on the pragmatic aspect of logical and mathematical operations, as, for example, the way we actually handle symbols and formulas. From this point of view Piaget's account bears a marked resemblance to the views of Jules Henri Poincaré and the intuitionists; the construction of number, for example, had for Piaget a definite psychological aspect.

ABSTRACT CONCEPTS. Piaget believed that logical and mathematical notions first show themselves as overt activities on the part of the child and only at a later stage take on a conceptual character. They are to be conceived as internalized actions in which things are replaced by signs, and concrete actions by operations on these signs. Rational activity occurs in the child when his trial-and-error gropings attain a definite pattern of order that may be inverted in thought. At this rational stage, if the child makes a mistake in performing a task, he is able to return to his starting point. This characteristic of thought that enables us to reverse a train of ideas or actions Piaget calls "reversibility." It is the basis of our ability to perform mental experiments, as well as the psychological foundation of the deductive process.

Piaget contended that the more elementary forms of logical behavior in which the child compares, distinguishes, and orders the objects around him are largely concerned with the creation of concrete classificatory and relational systems. It is from these systems that we develop our later, more abstract, logical and mathematical modes of thinking. Piaget would rather not speak of the intuition of number before the child has developed logical concepts of invariance and has thereby grasped the operation of reversibility. The transition to number occurs in the child just when his activities of classifying and ordering objects take on the form of simple logical systems. What emerges from Piaget's experimental researches is that numerical concepts in their psychological development are ultimately based on simple logical notions. There is thus some resemblance between the way number comes to be constructed in a child's thought and the attempt on a purely normative plane by Russell and others to define number in logical terms.

PHYSICAL CONCEPTS. Among the other concepts studied by Piaget, those of time and space are of particular philosophical relevance. Immanuel Kant, for example, believed that these concepts were objects of an a priori intuition. Piaget, however, found that the abstract notion of time arises at a relatively late stage; at first time is connected with space. For example, the child first confuses the notion of age with that of height or other visible signs of age. As far as space is concerned, his ability to make spatial judgments is initially fairly rudimentary. He can differentiate between open and closed figures but has difficulty in distinguishing one shape from another. He is also incapable of imagining a perspective other than his own. Only at a later stage is he able to take account of several relations at once (before and behind, right and left) and to coordinate them into a general system of perspectives.

PERCEPTION. For Piaget learning played an important part not only in the elaboration of intellectual structures but also in the field of perception. It is this that distinguishes his view from that of the Gestalt psychologists. For the latter, the perceptual constancies of shape and size belong directly to the perceived objects and are independent of age and ability. For Piaget, however, perception of figures is built up as a result of a series of random eye and other muscular movements, which are gradually corrected. The young child does not attribute a constant size or even identity to the objects around him. Piaget believed that the logical forms of activity that emerge in child behavior, namely classifying, relating, and so forth, arise as a result of his trial-and-error activities.

Piaget's views on perception have certain philosophical implications. In the past, he points out, philosophers have assumed a definite psychology of perception in their epistemologies. A good example of this is John Locke's sensationalism, in which it is assumed (1) that empirical facts are passively given in perception and (2) that they correspond to a certain range of linguistic expressions that designate them. For Piaget, however, even the notion of an object, one of the simplest forms of perceptual invariants, requires a definite learning process. Before the child is able to use linguistic expressions to refer unequivocally to definite objects, he must first have developed concrete classificatory and relational activities. Even the simple statement, "This is green," implies the acquisition of such skills and hence cannot be regarded as a reference to a simple perceptual datum. When we talk intelligently of green, this presupposes that we have learned to classify objects according to their color and to differentiate one color from another.

BEHAVIOR AND LOGIC. Piaget's work might be dismissed as philosophically irrelevant by philosophers of a Platonic turn of mind. It might be said that philosophical discussions of conceptual thinking are largely concerned with questions of validity and not with questions of origin. Piaget does not deny that logical notions as they appear in purely formal discussions differ from those occurring in ordinary thought. However, he asserts (1) that even our simpler kinds of intellectual performance have a logical character about them, which we can study formally, and (2) that when the logician constructs logical systems, performs deductions, tests for validity, and so on, his logical behavior can be studied in the same direct way as that of the child or unsophisticated adult. Piaget

also believed that it may be illuminating to compare the simpler logical structures inherent in our behavior with the purely formal systems constructed by the logician, as we may find some continuity between them.

See also Brunschvicg, Léon; Intuitionism and Intuitionistic Logic; Kant, Immanuel; Mach, Ernst; Number; Pearson, Karl; Perception; Poincaré, Jules Henri; Psychology; Russell, Bertrand Arthur William.

Bibliography

WORKS BY PIAGET

Piaget was a prolific writer, and among the numerous volumes he wrote, the following have a specifically logical or philosophical character:

Le jugement et le raisonnement chez l'enfant. Paris, 1924. Translated by M. Warden as *Judgment and Reasoning in the Child*. New York: Harcourt Brace, 1928.

Le langage et la pensée chez l'enfant. Paris, 1924. Translated by M. Warden as *The Language and Thought of the Child*. New York: Harcourt Brace, 1926. 2nd ed. translated by M. Gabain. London, 1932.

La représentation du monde chez l'enfant. Paris: Alcan, 1926. Translated by J. Tomlinson and A. Tomlinson as *The Child's Conception of the World*. New York: Harcourt Brace, 1929.

La causalité physique chez l'enfant. Paris, 1927. Translated by M. Gabain as *The Child's Conception of Physical Causality*. New York: Harcourt Brace, 1930.

Le jugement moral chez l'enfant. Paris: Alcan, 1932. Translated by M. Gabain as *The Moral Judgment of the Child*. New York: Harcourt Brace, 1932.

La genèse du nombre chez l'enfant. Paris, 1941. Written with A. Szeminska and translated by C. Gattegno and F. M. Hodgson as *The Child's Conception of Number*. New York: Humanities Press, 1952.

Le développement des quantités chez l'enfant. Paris, 1941. Written with B. Inhelder.

Le développement de la notion du temps chez l'enfant. Paris: Presses Universitaires de France, 1946.

Les notions de mouvement et de vitesse chez l'enfant: Paris: Presses Universitaires de France, 1946.

La représentation de l'espace chez l'enfant. Paris: Presses Universitaires de France, 1948. Written with B. Inhelder and translated by F. J. Langdon and J. L. Lunzer as *The Child's Conception of Space*. New York: Humanities Press, 1956.

La géometrie spontanée chez l'enfant. Paris: Presses Universitaires de France, 1948. Written with B. Inhelder and A. Szeminska and translated by E. A. Lunzer as *The Child's Conception of Geometry*. New York: Basic, 1960.

La genèse de l'idée de hazard chez l'enfant. Paris: Presses Universitaires de France, 1951.

De la logique de l'enfant à la logique de l'adolescent. Paris: Presses Universitaires de France, 1955. Translated by Anne Parsons and Stanley Milgram as *The Growth of Logical Thinking*. New York: Basic, 1958.

La genèse des structures logiques elementaires. Paris, 1959. Written with B. Inhelder.

WORKS ON PIAGET

There have been few philosophical discussions of Piaget's work, but W. Mays, "The Epistemology of Professor Piaget," in *PAS* (London, 1953–1954) compares his epistemology with the views of contemporary philosophers. C. Parsons in "Inhelder and Piaget's 'The Growth of Logical Thinking: II.' A Logician's Viewpoint," in the *British Journal of Psychology* (1960), criticizes Piaget's logic from a theoretical standpoint.

John H. Flavell in *The Developmental Psychology of Jean Piaget* (Princeton, NJ: Van Nostrand, 1963) gives a good summary of Piaget's work from a psychological point of view. The book contains an excellent bibliography of primary and secondary sources. W. Mays, "How We Form Concepts," in *Science News* (1954) gives a simple introduction from a more philosophical viewpoint.

K. Lovell in *The Growth of Basic Mathematical Concepts in Children* (London, 1961) provides an introduction to Piaget's ideas from an educational standpoint. Z. P. Dienes in *Building up Mathematics* (London, 1960) shows how Piaget's work has influenced the introduction of new methods in the teaching of school mathematics.

OTHER RECOMMENDED WORKS

Battro, Antonio M. *Piaget: Dictionary of Terms*. New York: Pergamon Press, 1973.

Boden, Margaret A. *Jean Piaget*. New York: Viking Press, 1980.

Bringuier, Jean Claude, and Jean Piaget. *Conversations with Jean Piaget*. Chicago: University of Chicago Press, 1980.

Elkind, David. *Child Development and Education: A Piagetian Perspective*. New York: Oxford University Press, 1976.

Elkind, David, and John H. Flavell, eds. *Studies in Cognitive Development: Essays in Honor of Jean Piaget*. New York: Oxford University Press, 1969.

Furth, Hans G. *Piaget and Knowledge: Theoretical Foundations*. Englewood Cliffs, NJ: Prentice-Hall, 1969.

Ginsburg, Herbert, and Sylvia Opper. *Piaget's Theory of Intellectual Development*. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1988.

Kegan, Robert. *The Evolving Self: Problem and Process in Human Development*. Cambridge, MA: Harvard University Press, 1982.

Murray, Frank B., and Millie Corinne Almy. *The Impact of Piagetian Theory: On Education, Philosophy, Psychiatry, and Psychology*. Baltimore: University Park Press, 1979.

Piaget, Jean. *The Essential Piaget*, edited by Howard E. Gruber and J. Jacques Vonèche. Northvale, NJ: J. Aronson, 1977, 1995.

Piaget, Jean, and Bäbel Inhelder. *The Child's Conception of Space*. New York: W. W. Norton, 1967.

Piattelli-Palmarini, Massimo, ed. *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press, 1980.

Wolfe Mays (1967)

Bibliography updated by Michael J. Farmer (2005)

PICO DELLA MIRANDOLA, COUNT GIOVANNI
(1463–1494)

Count Giovanni Pico della Mirandola, the Renaissance philosopher, was born in Mirandola, near Modena. He was a younger son in a family of feudal lords who ruled the small territory of Mirandola and Concordia in northern Italy. He seems to have received at an early age his first humanistic training in Latin and, perhaps, in Greek. Being destined by his mother for a career in the church, he was named papal protonotary at the age of ten and began to study canon law at Bologna in 1477. Two years later he began the study of philosophy at the University of Ferrara, which he continued at the University of Padua from 1480 to 1482.

After a number of journeys that took him to Paris and repeatedly to Florence, Pico studied Hebrew and Arabic under the guidance of several Jewish teachers and in 1486 composed 900 theses, offering to defend them in Rome the following year in a public disputation to which he invited scholars from all parts of Europe. When some of these theses met with objections from various theologians, Pope Innocent VIII appointed a committee to have them examined. As a result of the investigation 7 theses were condemned as unorthodox, and 6 more were declared to be dubious. When Pico published a defense of these 13 theses, the pope condemned all 900, although Pico had signed an act of submission. Pico fled to France, where he was arrested in 1488 on the request of papal envoys.

Upon the intervention of several Italian princes Pico was released from prison by King Charles VIII. He returned to Italy and was allowed by the pope to settle in Florence, under parole, as it were, and under the personal protection of Lorenzo de' Medici. Except for a few short visits to Ferrara, Pico spent the remainder of his life in Florence and there wrote, or began to write, his most important works, remaining in close touch with the circle of the Medici, with the Platonic Academy of Marsilio Ficino, and with Girolamo Savonarola. In 1493 he was acquitted of all ecclesiastical censures and restrictions by Alexander VI. He died in 1494 on the very day (November 17) on which Charles VIII of France made his entry into Florence after the expulsion of Piero de' Medici.

Pico's numerous writings reflect the wide range of his interests. He composed Italian and Latin poems of which only some have survived. A number of his humanistic letters were published posthumously, as was his famous *Oration*, originally composed for the projected disputation. To the scholastic aspect of his work we may assign the 900 theses (1486) and especially the *Apologia* (1487), his defense of the condemned theses. Another early work is his lengthy commentary on the Platonic love poem of his friend Girolamo Benivieni (1486). His mature philosophical works include the *Heptaplus* (1489), a sevenfold interpretation of the first verses (1:1–27) of Genesis, and his *De Ente et Uno* (On Being and Unity), written in 1491 but published posthumously. His most extensive work is his *Disputationes Adversus Astrologiam Divinatricem* (Disputations against Astrology), in twelve books, published posthumously. To this we may add a few short religious and theological writings and several fragments of a commentary on the Psalms that have been preserved in a number of scattered manuscripts and are still for the most part unpublished.

A characteristic document of Pico's attitude on history and philosophy from his earlier years is his correspondence with Ermolao Barbaro (1485). Barbaro, a distinguished Venetian humanist and student of the Greek texts of Aristotle, had stated in a letter to Pico that the medieval philosophers were uncultured and barbarous and did not deserve to be read or studied. Pico replied in a long letter in which he praises and defends the medieval philosophers and insists with great eloquence that what counts in the writings of philosophers is not their words but their thoughts. Unlike Barbaro and many other humanists who despised the scholastic philosophers for their lack of elegance and classical learning, Pico is willing to recognize the solidity of their thought and to learn from them whatever truth they may have to offer. The line between humanism and Scholasticism, rhetoric and philosophy, is thus clearly drawn, and Pico, although deeply imbued with humanist learning, throws his weight on the side of Scholasticism or, at least, of a synthesis of both sides. Many years after Barbaro and Pico died, Philipp Melanchthon wrote a reply to Pico's letter in defense of Barbaro's position.

SYNCRETISM

Pico's defense of the scholastic philosophers was merely a special instance of a much broader historical and philosophical attitude that has been rightly emphasized as his syncretism. Pico was convinced that all known philosophical and theological schools and thinkers contained certain valid insights that were compatible and hence deserved to be restated and defended. This was the underlying idea of his projected disputation, for the 900 theses relied on the most diverse sources—Hermes Trismegistus, Zoroaster, Orpheus and Pythagoras, Plato and Aris-

totle and all their Greek followers and commentators, Avicenna and Averroes and other Arabic philosophers, Thomas Aquinas and John Duns Scotus and several other medieval Latin thinkers, and the Jewish kabbalists.

In using all these sources, Pico wished to emphasize his basic conviction that all of these thinkers had a genuine share in philosophical truth. His notion of a universal truth in which each of the schools and thinkers participates to some extent constitutes an attempt to deal with the apparent contrasts and contradictions in the history of philosophy. It may be compared with the positions of the ancient eclectics and of G. W. F. Hegel, yet it differs from both of them. For Pico truth consists in a large number of true statements, and the various philosophers participate in truth insofar as their writings contain, besides numerous errors, a number of specific statements that are true. That this was Pico's intent we may gather from the second part of his *Oration* and from a passage in the *Apologia* that repeats it almost verbatim. He insists that he is not bound by the doctrines of any master or school but has investigated all of them. Instead of confining himself to one school, he has chosen from all of them what suits his thought, for each has something distinctive to contribute.

Pico's syncretism presupposes that of Ficino, who had proposed a theory of natural religion; had traced the Platonic tradition back to Hermes, Zoroaster, and other early theologians; and had insisted on the basic harmony between Platonism and Christianity. Yet Pico made these notions part of a much wider and more comprehensive synthesis. He explicitly includes Aristotle and all his Greek, Arabic, and Latin followers, and he adds to these previously known sources the Jewish kabbalists, with whom he became acquainted through his Hebrew studies, thus being probably the first Christian scholar to use kabbalistic literature. This attitude toward Aristotelianism and kabbalism clearly distinguished Pico from Ficino and other predecessors; it was to find further development in Pico's own later thought and to exert a strong influence on the philosophy of the sixteenth century. Pico's broad syncretism has been rightly praised by several historians as a steppingstone to later theories of religious and philosophical tolerance.

Pico's use of kabbalism consisted not so much in accepting specific kabbalist theories as in gaining recognition for kabbalism in general. Some of the theories that he seems to have borrowed from kabbalist authors were not necessarily of kabbalistic origin, such as the scheme of the three worlds—elementary, celestial, and angelic—which he uses for the first three sections of his *Heptaplus*. Pico accepted the claim made by the followers of kabbalism that their writings were based on a secret tradition that went back, at least in oral form, to biblical times. Kabbalism thus acquires a kind of authority parallel to that of the Bible and similar to that held by Hermes and Zoroaster in the eyes of Ficino and Pico. Moreover, Pico applied to kabbalism a principle that had been used for the Old Testament by all Christian writers since St. Paul: He tried to show that the kabbalistic tradition, no less than the Hebrew Scripture, was in basic agreement with Christian theology and hence could be taken as a prophecy and confirmation of Christian doctrine. With this argument he laid the foundation for a whole tradition of Christian kabbalism that found its defenders in Johannes Reuchlin, Giles of Viterbo, and many other thinkers in the sixteenth and later centuries.

In Pico's own work the kabbalistic influence is most noticeable, after the time of the 900 theses, in his *Heptaplus* and in his fragmentary commentary on the Psalms. In a manner that goes far beyond the usual medieval scheme of the four levels of meaning Pico assigns to the text of Scripture a multiple meaning that corresponds to the various parts of the universe. He also uses the kabbalistic method of scriptural interpretation, which assigns numerical values to the Hebrew letters and extracts secret meanings from the text by substituting for its words other words with comparable numerical values.

The other distinctive aspect of Pico's syncretism, his tendency to assume a basic agreement between Plato and Aristotle, also remained one of his major preoccupations during his later life. We know that he planned to write an extensive treatise on the agreement between Plato and Aristotle. The idea that Plato and Aristotle were in basic agreement, although differing in their words and apparent meaning, was not new with Pico. We find it in Cicero, who probably took it from his teacher Antiochus of Ascalon. It is also attributed as a program to Ammonius Saccas, the teacher of Plotinus, and endorsed by Boethius. We may also compare certain trends in recent scholarship that have attempted to bridge the gap between Plato's dialogues and Aristotle's extant later writings by interpolating the oral teaching of Plato and the lost early writings of Aristotle.

Pico's approach is known to us through his *De Ente et Uno*, a small treatise composed toward the end of his life and the only surviving fragment of his projected larger work on the harmony of Plato and Aristotle. The question he discusses is whether being and unity are coextensive, as Aristotle maintains in the tenth book of the *Metaphysics*, or whether unity has a broader diffusion

and higher status than being, according to the view of Plotinus and other Neoplatonists. Following the scholastic doctrine of the transcendentals, Pico sets out to defend the position of Aristotle. He then tries to prove that Plato did not hold the opposite view, as claimed by the Neoplatonists. In support of his claim Pico cites a passage from Plato's *Sophist* and dismisses the testimony of the *Parmenides*, arguing that this dialogue is merely a dialectical exercise.

In the course of his discussion Pico sharply distinguishes between being itself and participated being, and it is thus possible for him to maintain that God is identical with being in the first sense but above being in the second. The harmony between Plato and Aristotle that Pico tries to establish turns out to be Aristotelian, at least in its wording, but in another sense it is neither Platonic nor Aristotelian, and the distinction between being itself and participated being is evidently indebted to the same Neoplatonists whom Pico tries to refute on the major issue of the treatise. As a result Pico's position was criticized, on the one side, by Ficino, who, in his commentary on the *Parmenides*, defended Plotinus and, on the other, by the Aristotelian Antonio Cittadini, who formulated a series of objections that were answered first by Pico himself and then by his nephew and editor Gianfrancesco Pico.

Another aspect of Pico's syncretism is his treatment of classical mythology. An allegorical interpretation of the myths of the Greek poets had been developed by the ancient Stoics and Neoplatonists, and for them it had been a device for reconciling pagan religion with philosophical truth. When the medieval grammarians continued to interpret the classical poets in this manner, they minimized the pagan religious element and emphasized the implied universal, or even Christian, truth that would justify the study of these authors. The method was taken over and further developed by the humanists and Ficino. Pico tends to be even more elaborate in his discussion and interpretation of ancient myths, especially in his commentary on Benivieni's love poem. Here he repeatedly mentions his plan to write a treatise on poetic theology, a work that probably remained unwritten. Pico apparently intended to construct a detailed system of the theology implicit in the myths of the ancient poets and thus to include them in his universal syncretism.

DIGNITY OF MAN

Much more famous than the ideas thus far discussed is Pico's doctrine of the dignity of man and his place in the universe. The *Oration*, in which this doctrine is developed, is probably the most widely known document of early Renaissance thought. In many editions the work is titled "Oration on the Dignity of Man," but this title properly belongs only to the first part of the oration; the original title was simply *Oration*. Man and his dignity are often praised by the early humanists, and some of them dedicated entire treatises to the subject. The topic was taken up by Ficino, who assigned to the human soul a privileged place in the center of the universal hierarchy and made it, both through its intermediary attributes and through its universal thought and aspirations, the bond of the universe and the link between the intelligible and the corporeal world. In his *Oration* Pico went beyond Ficino in several ways. He did not discuss the question merely in passing or in the context of a large work dedicated to other subjects but displayed it prominently in the opening section of a short and elegant speech. Moreover, he lays the accent not so much on man's universality as on his freedom; instead of assigning to him a fixed though privileged place in the universal hierarchy, he puts him entirely apart from this hierarchy and claims that he is capable of occupying, according to his choice, any degree of life from the lowest to the highest. He has God tell Adam:

> Neither a fixed abode nor a form that is thine alone, nor any function peculiar to thyself have We given thee, Adam, to the end that according to thy judgment thou mayest have and possess what abode, what form, and what functions thou thyself shalt desire. Constrained by no limits, in accordance with thine own free will, in whose hand We have placed thee, thou shalt ordain for thyself the limits of thy nature.… Thou shalt have the power to degenerate into the lower forms of life, which are brutish. Thou shalt have the power, out of thy soul's judgment, to be reborn into the higher forms, which are divine.

These words have a modern ring, and they are among the few passages in the philosophical literature of the Renaissance that have pleased, almost without reservation, modern and even existentialist ears. It is not absolutely certain that they were meant to be as modern as they sound, and it is hard to believe what has often been said—that when Pico wrote them, he had denied or forgotten the doctrine of grace. After all, the words are attributed to God and are addressed by him to Adam before the Fall. Yet they do contain an eloquent praise of human excellence and of man's potentialities, and they receive added vigor when we think of what the reformers,

and even great humanists like Michel Eyquem de Montaigne, were to say about man's vanity and weakness.

Some scholars have tried to minimize Pico's praise of human dignity and regard it as a piece of mere oratory. This view is refuted by the testimony of the *Heptaplus*, a work written several years later and for an entirely different purpose. Here again, Pico places man outside the hierarchy of the three worlds—the angelic, celestial, and elementary—treats him as a fourth world by himself, and praises him and his faculties, although within a more obvious theological context.

Pico's insistence on man's dignity and liberty also accounts, at least in part, for his attack on astrology, to which he dedicates his largest extant work, probably composed during the last few years of his life. The *Disputationes Adversus Astrologiam Divinatricem* is full of detailed astronomical discussions and displays an amazing mastery of the astrological and antiastrological literature of previous centuries. It has often been hailed by historians as a landmark in the struggle of science against superstition. In fact, Pico does state that the stars act upon sublunar things only through their light and heat, not through any other occult qualities that may be attributed to them, and this statement sounds very sober, if not necessarily modern. Moreover, we learn that even a scientist such as Johannes Kepler at least modified his initial belief in astrology under the influence of Pico's treatise.

In Pico's time, however, the belief in astrology was more than a superstition, and the rejection of it was not necessarily scientific. As a general system astrology was closely linked with the scientific cosmology of the age and hence widely accepted not only by quacks but also by serious thinkers. There is no evidence that Pico was especially guided by scientific considerations in his polemics against astrology, and we must face the fact that he accepted natural magic while rejecting astrology. We happen to know that his work against astrology was composed as a part of a larger work he planned to write against the enemies of the church. The basic impulse of his attack was religious and not scientific, and he indicates more than once what his chief objection to astrology was—the stars are bodies, and our selves are spirits; it cannot be admitted that a corporeal and, hence, lower being should act upon a higher being and restrict its freedom.

Pico's conception of the relation between philosophy and religion is also significant. He became increasingly concerned with religious problems during his later years, a development in which his shock at the papal condemnation of his theses and the influence of Savonarola must have played a part. The fact appears in the religious and theological content of several of his later writings and in the religious motivation of his treatise against astrology. It also finds an unexpected expression in certain passages of the *De Ente et Uno*, a work that deals fundamentally with a very different problem. Here Pico tells us that God is darkness and that philosophical knowledge can lead us toward God only up to a certain point, beyond which religion must guide us. Unlike Ficino, Pico seems to regard religion as a fulfillment of philosophy; religion helps us to attain that ultimate end for which philosophy can merely prepare us.

Pico did not live long enough to develop his ideas into a coherent system. Fragmentary as his work was, it had wide repercussions for a long time. His universal syncretism came closer to subsequent efforts at formulating a universal religion than that of any of his predecessors, including Ficino. His study of Hebrew and Arabic, although not entirely without precedents, served as a widely known example and gave a powerful impulse to these studies in Christian Europe, leading to a study of the Hebrew Scripture and to many new translations of Jewish and Arabic texts. His study of the kabbalah started a broad and powerful current of Christian kabbalism, which flourished throughout the sixteenth century and included many distinguished scholars and thinkers. In his attempt to harmonize the traditions of Platonic and Aristotelian philosophy, of Hermetic and kabbalistic theology, and of the various strands of Arabic and scholastic thought with one another and with Christian doctrine, Pico pointed the way toward intellectual freedom and a universal truth that stands above the narrow limits of particular schools and traditions.

See also Antiochus of Ascalon; Aristotelianism; Aristotle; Averroes; Avicenna; Boethius, Anicius Manlius Severinus; Cicero, Marcus Tullius; Duns Scotus, John; Ficino, Marsilio; Hegel, Georg Wilhelm Friedrich; Humanism; Italian Philosophy; Kabbalah; Kepler, Johannes; Melanchthon, Philipp; Montaigne, Michel Eyquem de; Neoplatonism; Pico della Mirandola, Gianfrancesco; Plato; Platonism and the Platonic Tradition; Plotinus; Pythagoras and Pythagoreanism; Renaissance; Stoicism; Thomas Aquinas, St.

Bibliography

WORKS BY PICO

Three volumes of a critical edition of Pico's works by Eugenio Garin have appeared: *De Hominis Dignitate, Heptaplus, De Ente et Uno, e scritti vari* (Florence: Vallecchi, 1942) and *Disputationes Adversus Astrologiam Divinatricem*, 2 vols. (Florence: Vallecchi, 1946–1952). For the other works

(especially the *Conclusiones, Apologia*, and *Letters*) one of the numerous editions of Pico's works must be used. The earliest and best was published in Bologna (1496); the most accessible is the Basel edition of 1572. For additional letters and texts see Léon Dorez, "Lettres inédites de Jean Pic de la Mirandole," *Giornale storico della letteratura italiana* 25 (1895): 352–361, and Eugenio Garin, *La cultura filosofica del Rinascimento italiano* (Florence: Sansoni, 1961). A few of Pico's letters and short religious works, along with the biography of Pico by his nephew, were translated by Sir Thomas More as *Pico, His Life by His Nephew*, edited by J. M. Rigg (London, 1890). The commentary on Benivieni was translated by Thomas Stanley in 1651 and later appeared as *A Platonick Discourse upon Love*, edited by Edmund G. Gardner (Boston: Merrymount Press, 1914).

There is a modern English version of the *De Ente et Uno*, translated by Victor M. Hamm as *Of Being and Unity* (Milwaukee, WI: Marquette University Press, 1943), and no less than three versions of the *Oration—The Very Elegant Speech on the Dignity of Man*, translated by Charles G. Wallis (Annapolis, MD: St. John's Book Store, 1940); *Oration on the Dignity of Man*, translated by Elizabeth L. Forbes in *The Renaissance Philosophy of Man*, edited by Ernst Cassirer, Paul Oskar Kristeller, and John H. Randall Jr. (Chicago: University of Chicago Press, 1948), and published separately with the Latin text (Lexington, KY, 1953); and *Oration on the Dignity of Man*, translated by A. Robert Caponigri (Chicago: Gateway, 1956). The correspondence with Ermolao Barbaro was translated by Quirinus Breen as "Giovanni Pico della Mirandola on the Conflict of Philosophy and Rhetoric," *Journal of the History of Ideas* 13 (1952): 384–426.

WORKS ON PICO

For Pico's thought the chief monograph is Eugenio Garin, *Giovanni Pico della Mirandola* (Florence: Le Monnier, 1937). Important is Ernst Cassirer's "Giovanni Pico della Mirandola," *Journal of the History of Ideas* 3 (1942): 123–144, 319–346. See also Giovanni Semprini, *La filosofia di Pico della Mirandola* (Milan, 1936); Eugenio Anagnine, *G. Pico della Mirandola* (Bari, Italy: Laterza, 1937); Pierre Marie Cordier, *Jean Pic de la Mirandole* (Paris: Nouvelles Éditions Debresse, 1957); E. Monnerjahn, *Giovanni Pico della Mirandola* (Wiesbaden, Germany: Steiner, 1960); and Eugenio Garin, *Giovanni Pico della Mirandola* (Mirandola, Italy, 1963).

For Pico's sources see Pearl Kibre, *The Library of Pico della Mirandola* (New York: Columbia University Press, 1936). For the condemnation of his theses see Léon Dorez and Louis Thuasne, *Pic de la Mirandole en France* (Paris: Leroux, 1897). For his scholastic background see Avery Dulles, *Princeps Concordiae: Pico della Mirandola and the Scholastic Tradition* (Cambridge, MA: Harvard University Press, 1941). For the *De Ente et Uno* and its background see Raymond Klibansky, "Plato's *Parmenides* in the Middle Ages and the Renaissance," *Mediaeval and Renaissance Studies* 1 (1941–1943): 281–330. For his kabbalism see Joseph L. Blau, *The Christian Interpretation of the Cabala in the Renaissance* (New York: Columbia University Press, 1944); F. Secret, *Le zôhar chez les kabbalistes chrétiens de la Renaissance* (Paris, 1958) and "Pico della Mirandola e gli inizi della Cabala cristiana," *Convivium*, n.s. 25 (1957): 31–47.

For Pico's influence on the iconography of Renaissance art see Edgar Wind, *Pagan Mysteries in the Renaissance* (New Haven, CT: Yale University Press, 1958).

Several long papers by Eugenio Garin, Robert Weiss, Paul Oskar Kristeller, and Frances A. Yates are included in *L'opera e il pensiero di Giovanni Pico della Mirandola nella storia dell'umanesimo. Convegno internazionale, Mirandola, 15–18 settembre 1963*, 2 vols. (Florence: Nella Sede dell'Istituto, 1965).

See also Brian P. Copenhaver, "The Secret of Pico's Oration: Cabala and Renaissance Philosophy," *Midwest Studies in Philosophy* (26 [2002]: 56–81).

Paul Oskar Kristeller (1967)
Bibliography updated by Tamra Frei (2005)

PICO DELLA MIRANDOLA, GIANFRANCESCO
(1469–1533)

Gianfrancesco Pico della Mirandola was the nephew of the great Florentine humanist Giovanni Pico della Mirandola. He, like his uncle, became interested in the reform movement of Girolamo Savonarola (1452–1498) that was centered in the Convent of San Marco. The younger Pico della Mirandola moved into the convent and joined the group of scholars who took part in the daily discussions of philosophy and religion. His uncle moved into the convent in 1492 and placed his library there. Among the manuscripts brought to the convent by Pico della Mirandola and other scholars were five manuscript copies of Sextus Empiricus. Savonarola became interested in making these texts in Greek available to modern readers and asked two of his monks to begin preparing an edition of the writings of Sextus. This project never came to fruition, but some of it seems to be incorporated in the younger Pico della Mirandola's own publications.

He edited his uncle's work on astrology that was left in 1494, after Pico della Mirandola had died. He himself authored another work criticizing astrology, as did Savonarola. Pico della Mirandola was writing in praise of Savonarola up to the moment when the latter was arrested, tried, and burned at the stake. Thereafter, his disciples had to flee for their lives. Pico della Mirandola went back to his ancestral home of Mirandola, Italy, and struggled for about ten years to secure control of his family's property.

He wrote on a variety of philosophical and theological subjects, supporting the views of his teacher, Savonarola. In 1520 he published the first presentation of Greek skepticism in modern times, *Examen vanitatis doc-*

trinae gentium et veritatis Christianae disciplinae: Distinctum in libros sex, quorum tres omnem philosophorum sectam universim: reliqui Aristoteleam et Aristoteleis armis particulatim impugnant: Ubicunque autem Christiana et asseritur, et celebratur disciplina (Examination of the vain doctrine of the gentiles and the true Christian teaching). The work was apparently written over at least fifteen to twenty years. Besides presenting arguments and analyses out of Sextus, it also contains a text by John Philoponus and Hasdai Crescas. It is curious that Pico della Mirandola includes the material from Crescas, which had not yet been published and only circulated in Hebrew manuscript. He may have gotten a text and its translation from Judah León Abrabanel (c. 1460–c. 1521), with whom he was in contact. Pico della Mirandola's skeptical work did not have wide circulation. It is cited by several people writing on philosophical topics, but it does not seem to have encouraged people to look further into skeptical thought. He was read by Gentian Hervetius (1499–1584), the translator of Sextus, and probably by Francisco Sanches, Pierre Gassendi, and Gottfried Wilhelm Leibniz. The more serious impact of Sextus on modern thought had to await the presentation of his doctrines in Michel Eyquem de Montaigne's writings.

See also Pico della Mirandola, Count Giovanni; Sextus Empiricus; Skepticism, History of.

Bibliography

WORKS BY PICO DELLA MIRANDOLA

Examen vanitatis doctrinae gentium, et veritatis Christianae disciplinae. Mirandulae, Italy: N.p., 1520.

WORKS ABOUT PICO DELLA MIRANDOLA

Cao, Gian Mario. "Gianfrancesco Pico, Reader of Sceptics." Paper presented at the Conference on Renaissance Scepticism, New York, April 2004.

Floridi, Luciano. *Sextus Empiricus: The Transmission and Recovery of Pyrrhonism*. New York: Oxford University Press, 2002.

Popkin, Richard H. *The History of Scepticism: From Savonarola to Bayle*. Rev. ed. New York: Oxford University Press, 2003.

Schmitt, Charles B. *Gianfrancesco Pico della Mirandola (1469–1533) and His Critique of Aristotle*. The Hague: Nijhoff, 1967.

Schmitt, Charles B. "The Rediscovery of Ancient Skepticism in Modern Times." In *The Skeptical Tradition*, edited by Myles Burnyeat. Berkeley: University of California Press, 1983.

Viti, Paulo, ed. *Pico, Poliziano e l'umanesimo di fine quottrocento: Beblioteca medicea laurenziana*, November 4–December 31, 1994. Florence, Italy: Leo S. Olschki editore, 1994.

Richard Popkin (2005)

PIETISM

Since the seventeenth century "Pietism" has been an important movement within German Protestantism, and it is still influential in some parts of Germany. It began as a reaction against the formal and conventional character that appeared in Protestantism in the aftermath of the Reformation. Pietism opposed on the one hand the intellectualism implicit in the orthodox tendency to equate faith with the giving of assent to correct doctrine, and on the other, the tendency to identify Christianity with conformity to the ecclesiastical establishments that had been set up in various parts of Germany. By stressing experience, feeling, and personal participation as essential to a true Christian faith, Pietists hoped to bring new life into the Lutheran Church. One can point to similar movements in other parts of Christendom, in the English-speaking world the movement most akin to Pietism was Methodism.

The founder of German Pietism was Philipp Jakob Spener (1635–1705). Influenced by the extreme Protestant sect of Jean de Labadie, he undertook the task of raising the devotional level of his congregation in Frankfurt am Main and eventually, he hoped, of German Protestantism as a whole. Devotional meetings in his home were the beginnings of the famous *collegia pietatis*. At its meetings his sermons were considered, the New Testament was expounded, and there was conversation on religious topics. Spener gave clear expression to the aims of his movement in *Pia Desideria* (Frankfurt am Main, 1675), in which he laid down six goals to be realized: (1) greater study of the Bible but with the aim of personal devotion rather than academic competence; (2) a serious commitment to Martin Luther's belief in the priesthood of all Christian believers, so that the laity might really participate in the life of the church instead of merely conforming outwardly; (3) a realization that Christianity is a practical faith rather than an intellectual belief and that this faith expresses itself in love; (4) corresponding to this, a new style in apologetics and controversy that must aim not so much at intellectual conviction as at winning the allegiance of the whole man; (5) following from the last two points, the reorganization of theological education in order to lay stress on standards of life and conduct rather than on academic achievement; (6) the renewal and revitalizing of preaching as an instrument for building up a genuine piety among the people.

Spener continued to advocate his views in many other writings, including *Das geistliche Priesterthum* (1677), *Des thätigen Christenthums Nothwendigkeit*

(1679), *Die allgemeine Gottesgelehrtheit aller gläubigen Christen und Rechtschaffenen Theologen* (1680), *Klagen über das verdorbene Christenthum* (1684), *Natur und Gnade* (1687), and *Evangelische Glaubenslehre* (1688), which were all published at Frankfurt. He became engaged in stormy controversies, both attracting supporters and arousing opposition. Through the support of the elector of Brandenburg, the University of Halle became a center for Pietist views. Spener himself seems to have been a reasonable man who avoided the extravagances of some of his followers and performed a genuine service for the Lutheran Church.

Also important in the history of Pietism is August Hermann Francke (1663–1727). He taught at the University of Halle and is noteworthy for his development of the practical emphasis of Pietism. He founded a school for the poor and an orphanage and also took an interest in the cause of foreign missions. Like Spener, he encountered opposition, especially among some of the theologians, because of his indiscriminate attacks on intellectualism and his depreciation of the academic disciplines in the interests of devotion and philanthropy. Francke, however, had his supporters and was favored by King Frederick William I of Prussia. Mention should also be made of Count Nikolaus Ludwig von Zinzendorf (1700–1760), a pupil of Francke, who spread the spirit of Pietism to Holland, England, and North America by founding communities there. He maintained close relations with John Wesley and the Methodists. Like the other Pietists, he stressed feeling and personal devotion in what seems to have been a mixture of mysticism and emotionalism.

The chief characteristics of Pietism can be seen from this sketch of its origins and early history. It made claims for the affective and sometimes also the conative aspects of religion, in devotion and in practical service, at the expense of the cognitive element. While this may have been a healthy corrective to a sterile dogmatic orthodoxy, it tended to lead to dangerous excesses. Its insistence on intense inward experience could easily lead to the emotionalism that is common in evangelical religion and to the contempt for intelligence and common sense that sometimes accompanies it. The moralistic tone encourages utopianism. Some of those who have been caught up in the enthusiasm of Pietism have underrated the complexities of the moral life and the limitations of what is possible for man; as a result they have shared with the Methodists a belief in perfectionism. Apart from these dangerous excesses, Pietism has contended for the breadth of the human spirit and guarded against too narrow a rationalism. That the tenets of Pietism can receive a sober formulation worthy of respectful consideration is shown above all by the work of F. D. E. Schleiermacher, whose analysis of religion in terms of the feeling of absolute dependence is a direct reflection of the Pietist tradition in Germany.

The influence of Pietism on philosophy is largely indirect. The Pietists themselves tended to be antiphilosophical, but their spirit and teaching became part of the German heritage and eventually influenced even philosophy. This influence showed itself above all in the rise of *Lebensphilosophie* of which the religious variety, as expressed in the work of Rudolf Christoph Eucken, comes nearest to being a philosophical version of Pietism. Yet even the nonreligious varieties of this philosophy probably owe something to the anti-intellectualism that Pietism has encouraged.

See also Eucken, Rudolf Christoph; Luther, Martin; Mysticism, Nature and Assessment of; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

Crowner, David, and Gerald Christianson, eds. and trans. *The Spirituality of the German Awakening*. New York: Paulist Press, 2003.

Grünberg, Paul. *P. J. Spener*, 3 vols. Göttingen: Vandenhoeck and Ruprecht, 1893–1896.

Mahrholz, A. *Der deutsche Pietismus*. Berlin: Furche-verlag, 1921.

Nagler, Arthur W. *Pietism and Methodism*. Nashville, TN, 1918.

Petig, William E. *Literary Antipietism in Germany During the First Half of the Eighteenth Century*. New York: P. Lang, 1984.

Pinson, Koppel S. *Pietism as a Factor in the Rise of German Nationalism*. New York: Columbia University Press, 1934.

Ritschl, A. *Geschichte des Pietismus*, 3 vols. Bonn: Marcus, 1880–1886.

Sachsse, E. *Ursprung und Wesen des Pietismus*. Wiesbaden: Niedner, 1884.

Schmid, H. *Die Geschichte des Pietismus*. Nördlingen: Beck, 1863.

Stoeffler, Fred E. *German Pietism During the Eighteenth Century*. Leiden: Brill, 1973.

John Macquarrie (1967)
Bibliography updated by Tamra Frei (2005)

PISAREV, DMITRI IVANOVICH
(1840–1868)

Dmitri Ivanovich Pisarev, the Russian literary critic and social philosopher, was educated at St. Petersburg Uni-

versity (1856–1861). His studies were interrupted by a nervous breakdown requiring four months of institutionalization. At this time he twice attempted suicide. Pisarev was imprisoned from 1862 to 1866 for his outspoken criticism of the tsarist regime. He drowned while swimming in the Baltic Sea, under circumstances that suggest suicide, at the age of twenty-eight.

Pisarev called himself a "realist" and praised "fresh and healthy materialism," but his own philosophical position was a sense-datum empiricism. In his early writings on ethics and social philosophy, in the years 1859 to 1861, he advocated the "emancipation of the individual person" from social, intellectual, and moral constraints but particularly stressed the preservation of the wholeness of human personality in the face of the fragmenting pressures of functional specialization and the division of labor.

Among the constraints that the free individual must discard are those resulting from "the timidity of his thought, caste prejudices, the authority of tradition, the aspiration toward a common ideal" (*Polnoye Sobraniye Sochineniy,* Vol. I, Col. 339). Pisarev declared that common ideals have "just as little *raison d'être* as common eyeglasses or common boots made on the same last and to the same measure" (Col. 267). Eyes differ, feet differ, individuals differ; hence eyeglasses, boots, and ideals (for "every ideal has its author") should be individually fitted. Pisarev's moral relativism anticipated contemporary emotivist or noncognitivist doctrines in ethics—the claim that moral judgments are expressions of individual taste or preference. "When it is a matter of judging port or sherry," Pisarev wrote, "we remain calm and cool, we reason simply and soundly …, but when it is a question of lofty matters, we immediately … get up on our stilts.… We let our neighbor indulge his taste in hors d'oeuvres and desserts, but woe unto him if he expresses an independent opinion about morals" (Col. 266).

In his later writings Pisarev adopted a utilitarian ethics modified by the principle of "economy of intellectual energies." In the situation of cultural and intellectual deprivation of Russia at mid-century, he argued, the greatest-happiness principle precludes such luxuries as esoteric art that can be enjoyed "only by a few specialists" and abstruse science that is "in its very essence inaccessible to the masses" (Col. 366).

See also Empiricism; Ethics, History of; Noncognitivism; Russian Philosophy; Social and Political Philosophy; Utilitarianism.

Bibliography

WORKS BY PISAREV

Polnoye Sobraniye Sochinenii (Complete works). St. Petersburg, 1894. Six volumes.

Selected Philosophical, Social and Political Essays. Moscow: Foreign Languages Publishing House, 1958.

WORKS ON PISAREV

Coquart, Armand. *Dmitri Pisarev (1840–1868) et l'idéologie du nihilisme russe.* Paris: Institut d'études Slaves de l'Université de Paris, 1946.

Masaryk, Tomáš G. *Die geistigen Strömungen in Russland.* 2 vols. Jena, Germany, 1913. Translated by E. and C. Paul as *The Spirit of Russia: Studies in History, Literature, and Philosophy.* 2 vols., Vol. II, 53–81. New York: Macmillan, 1955.

George L. Kline (1967)

PLANCK, MAX

(1858–1947)

The German physicist Max Planck was the discoverer of the quantum of action, also called Planck's constant. Born in Kiel, he studied physics and mathematics at the University of Munich under Philipp von Jolly and at the University of Berlin under Hermann von Helmholtz and Gustav Kirchhoff. After receiving his Ph.D. at Munich (1879), he taught theoretical physics, first in Kiel, then (starting in 1889) in Berlin, as Kirchhoff's successor. "In those days," he wrote later, "I was the only theoretician, a physicist *sui generis,* as it were, and this circumstance did not make my *début* so easy." At this time Planck made important, and indeed quite fundamental, contributions to the understanding of the phenomena of heat, but he received hardly any attention from the scientific community: "Helmholtz probably did not read my paper at all. Kirchhoff expressly disapproved of its contents." The spotlight was then on the controversy between Ludwig Boltzmann and the Wilhelm Ostwald–Georg Helm–Ernst Mach camp, which supported a purely phenomenological theory of heat. It was via this controversy, and not because of the force of his arguments, that Planck's ideas were finally accepted. "This experience," he wrote, "gave me an opportunity to learn a remarkable fact: a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die."

Nevertheless, the discovery of the quantum of action in 1900, for which Planck received the Nobel Prize in physics (1918), was a direct result of these earlier studies. In 1912 Planck became permanent secretary of the (then)

Prussian Academy of Sciences, a post that he retained with only minor interruptions for the rest of his life. He used this position with excellent judgment for furthering the international collaboration of all scientists. From 1930 to 1935 he was president of the Kaiser-Wilhelm-Institut, which later became the Max-Planck-Institut.

Politically Planck was conservative, loyal to the Prussian ideas of the state and of honor, and loyal to Wilhelm II. During World War I he more than once expressed his devotion to the cause of the German people united in battle, and he received the order of "pour le mérite," one of the highest orders of Wilhelm's Germany. However, he opposed the Nazi regime. He defended Albert Einstein, first against his scientific opponents, then against his political enemies. Despite severe criticism by Johannes Stark, Phillip Lenard, and Ernst Müller, he continued to defend Einstein and other Jewish scientists (such as Walther Nernst) even after 1933. He later personally demanded of Adolf Hitler that those scientists who had been imprisoned be freed; as a consequence he was removed as president of the Physical Society, was refused the Goethe Prize of the city of Frankfurt (he was awarded it after the war, in 1946), and finally was forced to witness the execution of his only son, who had been connected with the German resistance. Antiquated as some of his political ideas may have been, he nevertheless put individual justice above all and defended it even at the risk of his own life. At the end of the war he was rescued by the Allied forces. He spent the last years of his life in Göttingen.

APPROACH TO SCIENCE

Planck's research was guided by his belief "of the existence in nature of something real, and independent of human measurement." He considered "the search for the absolute" to be the highest goal of science. "Our everyday starting point," he explained, "must necessarily be something relative. The material that goes into our instruments varies according to our geographical source; their construction depends on the skill of the designers and toolmakers; their manipulation is contingent on the special purposes pursued by the experimenter. Our task is to find in all these factors and data, the absolute, the universally valid, the invariant that is hidden in them."

This point of view was not allowed to remain a philosophical luxury, without influence upon the procedures of physics. One of the main objections that Planck raised against the positivistic creed was its sterility in the promotion of theory. "Positivism lacks the driving force for serving as a leader on the road of research. True, it is able to eliminate obstacles, but it cannot turn them into productive factors. For … its glance is directed backwards. But progress, advancement requires new associations of ideas and new queries, not based on the results of measurement alone."

SCIENTIFIC DISCOVERIES

Of new ideas Planck himself produced essentially two. He recognized and clearly formulated those properties of heat that separate it from purely mechanical processes, and he introduced and applied to concrete problems the idea of an atomistic structure not only of matter but of radiation also. In his doctoral dissertation he had already separated thermodynamic irreversibility from mechanical processes and had interpreted Rudolf Clausius's entropy as its measure. Later he showed (independently of Willard Gibbs) that "all the laws of physical and chemical equilibrium follow from a knowledge of entropy."

His conviction that the principle of the increase of entropy was a genuine and independent physical law and his belief in the universal (or, to use his term, *absolute*) validity of all physical laws led him to apply thermodynamic reasoning in domains that until then had been regarded as inaccessible to it. For example, he determined that the lowering of the freezing point of dilute solutions could be explained only by a dissociation of the substances dissolved, thus extending the science of thermodynamics to electrically charged particles. This tendency to strain laws to the limit rather than to restrict them to the domain of their strongest evidence caused a temporary clash with Boltzmann, who was quite unperturbed by the fact that in his approach the entropy of a system could both increase and decrease. But it also led to Planck's greatest triumph—his discovery of the quantum of action.

Planck was the only one to correlate the relevant features of radiation with the entropy, rather than the temperature, of the radiant body. "While a host of outstanding physicists worked on the problem of spectral energy distribution, both from the experimental and theoretical aspect, every one of them directed his efforts solely towards exhibiting the dependence of the intensity of radiation on the temperature. On the other hand I suspected that the fundamental connexion lies in the dependence of entropy upon energy. As the significance of the concept of entropy had not yet come to be fully appreciated, nobody paid attention to the method adopted by me, and I could work out my calculations completely at my leisure." These calculations furnished a formula that agreed with experiment and contained the existing theoretical results (Wien's formula and the

Raleigh-Jeans law) as limiting cases. In the attempt to find a rationale for this result, Planck used Boltzmann's statistical interpretation of entropy and was thus led to the discovery of the "atomic," or discontinuous, structure of action (energy).

REALISM, DETERMINISM, AND RELIGION

The discovery of the quantum of action was brought about not only by the specific physical arguments used but also by the philosophical belief in the existence of a real world behaving in accordance with immutable laws. The intellectual climate of the late nineteenth century was opposed to such a belief (Boltzmann was almost the only other figure to uphold it). This climate not only found expression in the philosophical superstructure but influenced physical practice itself. Laws were regarded as summaries of experimental results and were applied only where such results were available. However, it was the "metaphysics" of Planck, Boltzmann, and, later on, Einstein (whom Planck interpreted as a realist from the very beginning) that made possible many of the theories that are now frequently used to attack realism and other "metaphysical" principles.

Planck never accepted the positivistic interpretation of the quantum theory. He distinguished between what he called the "world picture" of physics and the "sensory world," identifying the former with the formalism of the ψ waves, the latter with experimental results. The fact that the ψ-function obeys the Schrödinger equation enabled him to say that while the sensory world might show indeterministic features, the world picture, even of the new physics, did not. His belief in the existence of objective laws also provided him with an important steppingstone to religious belief. Planck argued that the laws of nature are not invented in the minds of men; on the contrary, external factors force us to recognize them. Some of these laws, such as the principle of least action, "exhibit a rational world order" and thereby reveal "an omnipotent reason which rules over nature." He concluded that there is no contradiction between religion and natural science; rather, they supplement and condition each other.

See also Quantum Mechanics.

Bibliography

WORKS BY PLANCK

Theory of Heat Radiation. Translated by Morton Masius. Philadelphia: P. Blakiston's Son, 1914; 2nd ed., New York: Dover, 1959.

Eight Lectures on Theoretical Physics. Translated by A. P. Wills. New York: Columbia University Press, 1915. Lectures given at Columbia University in 1909.

The Origin and Development of the Quantum Theory. Translated by H. T. Clarke and L. Silberstein. Oxford: Clarendon Press, 1922. Nobel Prize address.

A Survey of Physics; A Collection of Lectures and Essays. Translated by R. Jones and D. H. Williams. London: Methuen, 1925. Reissued as *A Survey of Physical Theory.* New York: Dover, 1960.

Treatise on Thermodynamics. Translated by Alexander Ogg. London: Longmans, Green, 1927; 3rd rev. ed., New York, 1945.

Introduction to Theoretical Physics. Translated by Henry L. Brose, 5 vols. London, 1932–1933; New York, 1949. Includes *General Mechanics, The Mechanics of Deformable Bodies, Theory of Electricity and Magnetism, Theory of Light,* and *Theory of Heat.*

Scientific Autobiography and Other Papers. Translated by Frank Gaynor. New York: Philosophical Library, 1949.

The New Science. Translated by James Murphy and W. H. Johnston. New York: Meridian, 1959. Includes *Where Is Science Going?* (a defense of determinism with a preface by Albert Einstein), *The Universe in the Light of Modern Physics,* and *The Philosophy of Physics.*

WORKS ON PLANCK

Schlick, Moritz. "Positivism and Realism." In *Logical Positivism,* edited by A. J. Ayer. Glencoe, IL: Free Press, 1959. This essay was a direct reply to the criticisms of positivism that Planck expressed in *Positivismus und reale Aussenwelt.* Leipzig, 1931.

Vogel, H. *Zum philosophischen Wirken Max Plancks.* Berlin: Akadamie, 1961. Excellent biography with detailed bibliography.

Paul K. Feyerabend (1967)

PLANTINGA, ALVIN
(1932–)

Born in Ann Arbor, Michigan, Alvin Plantinga is one of the most important and influential philosophers of the twentieth and early twenty-first centuries. His publications range over a wide variety of fields, but his most enduring contributions have been in metaphysics, epistemology, and, especially, the philosophy of religion. He is best known for his work on the metaphysics of necessity and possibility, for his defense of the view that knowledge is to be analyzed partly in terms of proper function, for his development of the "free will defense" against the so-called "logical problem of evil," for his many and vigorous defenses of the rationality of religious belief, and for his much-discussed "evolutionary argument against naturalism."

Plantinga earned his BA in philosophy and psychology from Calvin College in 1953, and he cites his experi-

ence at Calvin as perhaps the single most significant intellectual influence in his life. There he studied with Henry Stob and Harry Jellema, the latter of whom played an especially formative role in his intellectual development. Plantinga received his MA in philosophy from the University of Michigan in 1955, and his PhD from Yale in 1958 under the supervision of Paul Weiss. He was elected to the American Academy of Arts and Sciences in 1975, co-founded the Society of Christian Philosophers in 1978, served as President of the Central Division of the American Philosophical Association in 1981 and 1982, and served as President of the Society of Christian Philosophers from 1983 to 1986. He has given the prestigious Gifford Lectures twice, and in a 1980 article he was heralded by *Time* magazine as "America's leading orthodox Protestant philosopher of God."

METAPHYSICS

Plantinga's most influential work in metaphysics has focused primarily on the metaphysics of modality. In *The Nature of Necessity* (1974), as well as in various papers, a central theme is the exposition and defense of a *realist* and *actualist* construal of possible worlds and modal properties. On his view, the standard possible worlds semantics for modal logic is to be taken with metaphysical seriousness: it is not a mere heuristic device; possible worlds are not merely useful fictions. Rather, possible worlds exist. They are abstract states of affairs of a certain kind—something like total or complete ways that a world history might have gone. Moreover, individual things have *modal properties*—properties such as *being possibly seven feet tall*, or *being necessarily even*—and, Plantinga thinks, realism about such properties requires one to believe that individual things exist in other worlds. On his view, a thing exists in another possible world only if, had that world been actual, the thing itself, not a mere stand-in or counterpart, would have existed. Thus, if Fred has the property *being possibly seven feet tall*, then there is a possible world such that, had that world been actual, Fred himself would have existed and would have been seven feet tall. Ultimately, this understanding of modal properties, together with his commitment to strong form of actualism, leads Plantinga to endorse the controversial view that objects have *individual essences*—properties that are both essential and essentially unique to them.

EPISTEMOLOGY

Plantinga's major works in epistemology are the volumes that comprise his *Warrant* trilogy (1993a, 1993b, 2000). Warrant, according to Plantinga, is that property or quantity that distinguishes knowledge from mere true belief. The main goals of the *Warrant* books are to identify the necessary and sufficient conditions for warrant, and to defend an affirmative answer to the question whether distinctively Christian belief can be warranted.

In the first volume of the trilogy, Plantinga surveys a broad range of post-Gettier analyses of knowledge, arguing that all of them founder on counterexamples involving cognitive malfunction. The basic problem is that, for each candidate analysis of "S knows that *p*," the conditions listed as necessary and sufficient for knowledge could be satisfied by a person whose cognitive faculties are malfunctioning in such a way that, intuitively, beliefs produced by faculties behaving in that way fail to count as knowledge. He also argues for the striking and controversial conclusion that *justification*, construed at least in part as a matter of epistemic duty-fulfillment, is not necessary for knowledge at all.

In the second volume, Plantinga articulates and defends a new analysis of knowledge, according to which (roughly) S knows that *p* if and only if S believes that *p*, *p* is true, and S's belief that *p* is the product of faculties that are properly functioning, successfully aimed at truth, and operating in an appropriate environment. The notions of *proper function* and *appropriate environment* are normative notions; but, Plantinga says, the normativity involved is of a sort commonly invoked in the natural sciences. Thus, Plantinga regards his analysis as, strictly speaking, an instance of "epistemology naturalized." But he also argues that his brand of epistemology naturalized flourishes best in the context of a supernaturalistic metaphysics. Toward establishing this conclusion, he begins by arguing that proper function is an irreducibly normative notion that does not admit of a purely naturalistic analysis. He then attacks naturalism directly, arguing that belief in naturalism and evolutionary theory together is epistemically self-defeating and therefore irrational. This latter argument is the so-called "evolutionary argument against naturalism."

The third volume of the *Warrant* trilogy applies the account of knowledge defended in the second volume in the service of an argument for the possibility of warranted Christian belief. The central and striking thesis of the book is that if Christian belief is true, then it is warranted. This conclusion is important because it implies, contrary to widespread opinion, that objections against the *rationality* of Christian belief are not independent of objections against the *truth* of Christian belief. In order to defend the conclusion that Christian belief is unwarranted, one must first defend the conclusion that it is

false. In support of his central thesis, Plantinga begins by arguing that the only really philosophically interesting question about the rationality of Christian belief is the question that asks whether Christian belief is or can be warranted. He then notes that, in light of the analysis of knowledge proposed in the second volume, Christian belief can be warranted so long as it is produced by properly functioning faculties that are successfully aimed at truth and functioning in a suitable environment. Much of the rest of the volume, then, is devoted to establishing the conclusion that if Christian belief is true, then these conditions are satisfied.

PHILOSOPHY OF RELIGION

Plantinga's work in the philosophy of religion has focused on what is sometimes referred to as "negative apologetics": the task of showing that objections to religious belief are unsuccessful. Thus, to take just a few examples, Plantinga has argued that the proposition that God exists is logically consistent with the proposition that evil exists; that the existence of evil does not constitute a defeater for the rationality of Christian belief; that widespread religious pluralism and intractable disagreement on religious matters do not provide reason to doubt that one knows that one's own religious beliefs are true; and that what he takes to be the correct views about human freedom and moral responsibility are not in tension with the traditional belief that God has perfect knowledge of the future.

Plantinga's focus on negative apologetics stems in part from his view that the warrant for Christian belief need not, and, in the ordinary case, does not come from philosophical argument but rather from something like religious experience. This view is a central theme in his work on religious epistemology, especially in the third volume of his *Warrant* trilogy (discussed above), but also in two earlier works: *God and Other Minds* (1967), and "Reason and Belief in God." Nevertheless, he does make occasional forays into the territory of positive apologetics. For example, in *The Nature of Necessity* and *God, Freedom and Evil* (1977), Plantinga argues that a modal version of the ontological argument for God's existence is both valid and plausibly sound. Likewise, Plantinga has devoted considerable energy to rebutting naturalism and its common companion, materialism.

Besides introducing important arguments into the literature on the philosophy of religion, however, Plantinga has also played an important role in shaping the way in which many religious philosophers now approach topics in their own fields of specialization. In his highly influential paper, "Advice to Christian Philosophers," (1984) Plantinga urges philosophers who share his Christian worldview to allow the presuppositions of that worldview to inform their work not only on topics in the philosophy of religion but elsewhere as well. He advises them not to become swept up in projects that arise out of and embody presuppositions of rival worldviews (such as naturalism or creative anti-realism), but to pursue a philosophical agenda in which one explores how a person with a Christian perspective ought to think about the various topics central to her discipline. This advice, itself an apt expression of one of the distinctive features of Plantinga's own work, has had a significant impact on the sorts of philosophical projects that have been undertaken in the late twentieth and early twenty-first centuries.

See also Evil, the Problem of; Ontological Argument for the Existence of God.

Bibliography

PRIMARY WORKS

God and Other Minds. Ithaca, NY: Cornell University Press, 1967.

The Nature of Necessity. New York: Clarendon Press, 1974.

God, Freedom, and Evil. Grand Rapids, MI: Eerdmans, 1977.

"Reason and Belief in God." In *Faith and Rationality*, edited by Alvin Plantinga and Nicholas Wolterstorff, 16–93. Notre Dame: University of Notre Dame Press. 1983.

Warrant: The Current Debate. New York: Oxford University Press, 1993a.

Warrant and Proper Function. New York: Oxford University Press, 1993b.

"Advice to Christian Philosophers," *Faith and Philosophy* 1: 253 – 71. Reprinted in Sennett, 1998.

The Analytic Theist: An Alvin Plantinga Reader. Edited by James F. Sennett. Grand Rapids, MI: Eerdmans, 1998.

Warranted Christian Belief. New York: Oxford University Press, 2000.

Essays in the Metaphysics of Modality, edited by Matthew Davidson. New York: Oxford University Press, 2003.

SECONDARY WORKS

Tomberlin, James, and Peter van Inwagen, eds. *Alvin Plantinga*. Dordrecht, Netherlands: D. Reidel, 1985.

Michael C. Rea (2005)

PLATO

(428/427 BCE–337/336 BCE)

The philosopher Plato was born to an aristocratic Athenian family. His father Ariston was said to be descended from the legendary King Codrus; the family of his mother Perictione was prominent in more historical times.

Dropides, an ancestor of Perictione, was a relative and friend of Solon (as Plato himself reports in the *Timaeus*, 20e). After Plato's father's death, Perictione was remarried to Pyrilampes, a political associate of Pericles and Athenian ambassador to the Persian king. Perictione's brother Charmides and her cousin Critias had a more sinister career, as members (and in Critias's case, ringleader) of the Thirty Tyrants who ruled Athens in a bloody junta, after the defeat by Sparta in 404 BCE.

Plato's family is well represented in the dialogues, perhaps to compensate for his own absence. In the *Charmides*, situated thirty years before the rule of the Thirty, Plato introduces his uncle Charmides as a promising young nobleman, under the influence of his older cousin Critias. The reference here to Charmides' family allows Plato to sing the praises of his own household, as the union of two outstanding families "than which no more noble union can be found in Athens" (*Charmides* 157e). The two families in question are those of Perictione and Pyrilampes, Plato's mother and stepfather. It is Plato's cousin Critias the tyrant (and not, as some scholars have supposed, the tyrant's grandfather) who appears again as introductory speaker in the *Timaeus* and as narrator in the unfinished *Critias*. Plato's older brothers Glaucon and Adeimantus are the chief interlocutors in the *Republic*, his stepbrother Demos is mentioned as a reigning young beauty in the *Gorgias* (481d), and his half brother Antiphon appears in the *Parmenides* as the one who preserves the memory of the philosophical conversation between Socrates, Zeno, and Parmenides. Plato had no occasion to mention his sister Potone, the mother of his nephew Speusippus who succeeded him as head of the Academy.

We are largely dependent on the autobiographical sections of the *Seventh Letter* for information about Plato's life. (The authenticity of this *Letter* is disputed, but even scholars who doubt its authenticity generally assume that the author was well informed.) The author of the *Letter* reports that Plato's relatives in the antidemocratic coup of 404 BCE invited him to join them, and that he, as an upper-class young man of twenty-three with political ambitions, was initially sympathetic; he expected these men to lead the city "from a life of injustice to a just government." But Plato observed that in a short time "they made the previous (democratic) regime look like a Golden Age" (*Epist.* VII, 324d). Thus Plato was repulsed by the behavior of Critias and the oligarchs; on the contrary, he admired the courage of Socrates in refusing to obey the tyrants' command, when they ordered him to lead a death squad against a prominent democrat. Plato's political ambitions revived in the restored democracy after 403, but after watching the politics of Athens for ten or fifteen years he concluded that the situation was hopeless, and that "the races of mankind would not cease from evils until the class of true philosophers come to political power or the rulers of the cities practice true philosophy" (*Epist.* VII 326b). At the age of about forty, Plato then departed for the Greek cities of southern Italy and Sicily.

Sometime after his return to Athens from Syracuse (c. 387 BCE), Plato began to gather together the group of students and researchers in science, mathematics and philosophy that became known as the school of the Academy. The early fourth century saw the creation in Athens of the first fixed schools of higher education, replacing the wandering Sophists of the fifth century. Antisthenes, the follower of Socrates, and the famous orator Isocrates had both recently established their schools. Unlike these institutions, Plato's community of scholars seems to have had no formally enrolled students and no tuition fees.

We know very little about the functioning of the Academy. The physical basis was a small estate with a garden owned by Plato, in the suburban neighborhood named after a park and gymnasium dedicated to the hero Academos. Formal instruction was probably offered in the gymnasium; the communal meals, or *syssitia*, presumably took place in Plato's villa. We happen to know of one public lecture given by Plato "On the Good." There is no evidence for a curriculum in mathematics and dialectic modeled on the studies of the guardians in *Republic* VII, as some scholars have supposed. There is in fact no evidence for any fixed curriculum. The only contemporary report (other than veiled attacks from Isocrates as head of the rival school) consists of quotations from Attic comedy, which make fun of the haughty manners and elegant dress of intellectuals from the Academy, and of their elaborate pedantry in the botanical classification of a pumpkin.

The intellectual caliber of the school is attested by the quality of its associates: on the one hand, Aristotle, who worked in the Academy for twenty years before Plato's death, and on the other hand Eudoxus, a great mathematician and astronomer, who seems to have maintained close contact with the school over many years, despite philosophical disagreement with Plato on central issues. Clearly, the members of the Academy were as much concerned with ethics and politics as with science and theoretical philosophy; the school is sometimes represented as a training program for statesmen. Plato's personal prestige is reflected in Aristotle's elegy to Friend-

ship, where Plato is called "a man whom the bad do not even have the right to praise, who alone or who first among mortals clearly showed, in his own life as in his teachings, that to become good is also to become happy (*eudaimôn*)."

Plato's quiet life in the Academy was interrupted in 367 and 361 by two invited voyages to the court of Dionysius II in Syracuse. Plato was persuaded to accept the invitation by his close friend Dion, the uncle of the tyrant, whom Plato had converted to philosophy on his first visit to Syracuse some twenty years earlier. Since the young Dionysius displayed a passion for philosophy, Plato was unwilling to reject this opportunity to influence the politics of the most powerful Greek city of the time. He proved quite ineffective in the intrigues of the Syracusan court, and was barely able to escape safely from his final visit to Dionysius in 360 BCE at the age of sixty-eight. The *Seventh Epistle* presents a detailed account of the Syracusan adventure from Plato's point of view. It ended in disaster both for Plato and for Sicily. After driving the tyrant out of Syracuse, Dion himself was murdered in 353 BCE. Plato, at seventy-five, responded with an elegy on the death of Dion, ending with the verse "Dion, you who once drove my heart mad with *erôs*."

WRITINGS

Of the thrirty-six dialogues preserved in the traditional canon (presumably as edited by Thrasyllus in the first century CE), some twenty-six or twenty-seven are generally recognized as the work of Plato. (The authenticity of the *Hippias Major* is contested; some scholars would also defend the *First Alcibiades* and perhaps a few others usually regarded as spurious.) The traditional corpus includes thirteen *Epistles*, most of them now recognized as spurious. Two or three of the *Epistles* have some claim to be authentic; the most important of these, for both philosophical and biographical reasons, is *Epistle* VII.

The only reliable guide to the chronology of the dialogues is the division into three stylistic groups, established by Campbell and Ritter in the late nineteenth century.

Group I: *Apology*, *Charmides*, *Crito*, *Cratylus*, *Euthydemus*, *Euthyphro*, *Gorgias*, [*Hippias Major*], *Hippias Minor*, *Ion*, *Laches*, *Lysis*, *Menexenus*, *Meno*, *Phaedo*, *Protagoras*, *Symposium*

Group II: *Republic*, *Phaedrus*, *Parmenides*, *Theaetetus*

Group III: *Sophist-Statesman*, *Philebus*, *Timaeus-Critias*, *Laws*

Group III was identified first (as the "late group") on the basis of several independent studies. These six dialogues are marked by very strong stylistic peculiarities typical of the *Laws*, which we know to have been written towards the end of Plato's life. Group II includes dialogues stylistically akin to the *Republic*, which show relatively few distinctive features of Plato's late style. Group I is the default class, the remaining sixteen or seventeen dialogues, from the *Apology* to the *Symposium* and *Phaedo*, in which Plato's brilliant conversational style bears none of the distinctive marks of the late period.

This chronological division into three groups is only partially in agreement with a conventional division of Plato's dialogues into early, middle, and late. The dialogues of Group III are all truly "late." (There was a brief attempt to date the *Timaeus* earlier for philosophical reasons, but that attempt has generally been recognized as a failure. The style of the *Timaeus* was from the beginning recognized as belonging to the latest period.) But the usual classification of "middle" dialogues ignores chronology altogether. It combines two dialogues of Group II (*Republic*, *Phaedrus*) with two from Group I (*Symposium*, *Phaedo*) solely on grounds of philosophical content. Despite their stylistic differences, all four works present the classical version of Plato's doctrine of Forms. A popular view of Plato's development locates these dialogues in a "middle period," divided on the one hand from the more "Socratic" dialogues of an earlier period, and, on the other hand, from the attack on the theory of Forms in the *Parmenides* and hence from the more critical philosophy of the *Theaetetus* and *Sophist*. This tripartite division is not supported by the Campbell-Ritter chronology of the three groups, since stylistically the *Parmenides* and *Theaetetus* are not later than the *Republic*. The notion of a "Socratic" period depends upon a particular interpretation of the role of Socrates in the earlier dialogues.

SOCRATES

The figure of Socrates appears in every Platonic dialogue except the *Laws*, and he is the chief speaker in all but five. This raises two difficult problems for interpreting Plato's work. How far does Socrates speak for Plato? And what is the relation between the Socrates of Plato's dialogues and the historical figure? We deal here with the historical question.

Since Socrates wrote nothing, we are entirely dependent on other writers for knowledge of his thought. The traditional attitude of historians has been to rely on the picture of Socrates presented in Plato's earlier dia-

logues, supplemented or confirmed by information from Xenophon and Aristotle. The result has been to take dialogues such as the *Laches*, *Euthyphro*, and *Protagoras* as providing a historical account of Socrates, as a moral philosopher who identifies virtue with knowledge, denies the reality of *akrasia* (weakness of will), and systematically pursues definitions of the moral virtues. In these "Socratic" dialogues Plato is thought to be closely following the thought and methodology of his master. On this view, Plato will only gradually develop his own philosophy, first with the doctrine of recollection in the *Meno* and then with the theory of Forms in the *Symposium* and *Phaedo*. This view can be supported by evidence from Aristotle, whose references to Socrates match the picture given in these "early" dialogues.

This account of Socrates has been treated with skepticism in much recent scholarship, because of a realization (pioneered by Gigon but developed by others) that the Socratic literature is a form of fiction rather than of historical biography. This fictional status is particularly clear in the remains of Socratic dialogues by other authors, such as Aeschines or Phaedo, but also in Xenophon's *Symposium* and in Platonic dialogues such as the *Menexenus*. Plato's portrait of Socrates is no doubt generally faithful to the moral character of the man as he saw him. But in regard to details of Socratic philosophy and argumentation, Plato would be at least as free as a modern novelist would be in dealing with historical figures. Furthermore, in the view of skeptical critics, Aristotle cannot serve as a reliable witness. He arrived in Athens as a youth more than thirty years after Socrates' death, and his picture of Socrates can be explained as his own inference from the Platonic dialogues. Judged as a historian of philosophy, Aristotle has serious faults. He generally sees his predecessors through the prism of his own scheme, and his account of the development of Plato's thought is particularly suspect. Aristotle's report of Cratylus's early influence on Plato is scarcely compatible with Plato's own portrait of Cratylus in the dialogue of that name; and Aristotle's claim that Plato's Theory of Forms was derived from the Pythagoreans is not supported by his own account of Pythagorean doctrine. Aristotle never mentions Plato's much more profound debt to Parmenides' concept of Being.

For all these reasons, a critical reader may well doubt that we have any reliable information about the philosophy of Socrates. It is perhaps in Plato's *Apology* that we can best catch a glimpse of the historical Socrates. The *Apology* is a special case among Plato's writings, since it is not a fictitious dialogue but a courtroom speech, the representation of a public event at which Plato claims to have been present. From this and other sources we can form a vivid picture of Socrates' powerful personality, his strong moral character, and his remarkable skill in elenchus, that is, in arguing his interlocutors into contradiction. But beyond the firm refusal to act unjustly and the conception of virtue (*aretê*) as care of one's self, or care of one's soul, our historical knowledge of Socrates' philosophical views seems to be limited to a handful of moral paradoxes: that no one does wrong voluntarily, that it is better to suffer than to do wrong, that virtue is knowledge, and that no evil can happen to a good man. In order to put philosophical flesh on this skeleton of doctrine, we must turn to the dialogues. But then we can no longer distinguish what derives from Plato's memory of the historical Socrates from what has its source in Plato's own artistic and philosophical imagination.

THE FIRST DIALOGUES: *APOLOGY*, *CRITO*, *ION*, *HIPPIAS MINOR*, *GORGIAS*

Although we do not know the chronological order of the dialogues in Group I, it is natural to begin with the two dialogues directly concerned with Socrates' trial and death, *Apology* and *Crito*, and with two other very short dialogues, *Ion* and *Hippias Minor*. We connect with this group a much more substantial work, *Gorgias*, which many scholars would put later. These five dialogues serve to illustrate the wide range of Plato's philosophical concerns, while at the same time revealing no trace of the metaphysics and epistemology that we recognize as distinctly Platonic.

Although it may have been written ten or twelve years after Socrates' death, the *Gorgias* presents a systematic exposition of Socratic moral doctrine and a strong defense of this view against anti-moralist attack. The *Gorgias* repeatedly recalls Socrates' trial and matches it with a judgment myth, in which the souls of those who are truly guilty of injustice will be punished. In the *Crito*, Socrates had formulated as his fundamental moral principle that one should never act unjustly, never return a wrong for a wrong. Socrates is prepared to die for this principle, and is unwilling to save his life by an unjust escape from prison. "It is not living that is of chief importance, but living well, and that is living honorably and justly" (*Crito* 48b). Crito agrees. Socrates recognizes that, between those who accept and those who reject these principles, "there is no common basis for discussion, no *koinê boulê*, but they must despise one another's views" (49d). In the *Gorgias* there is no such agreement, and the principle of justice is itself at stake. The Greek conception of justice

(*dikaiosunê*) is broad enough to cover morality generally, understood as respect for the rights of others. (Thus Aristotle defines "justice" as virtue in regard to others.) Socrates in the *Gorgias* has much the same task as later in the *Republic*: to defend the principle of morality against opponents who endorse the ruthless pursuit of wealth and power. The question, "Why be moral?" is posed here in dramatic form, against the background of Socrates' own fate as a martyr for moral principle.

The Socratic elenchus as practiced in the *Gorgias* is able to show that the anti-moral positions of Polus and Callicles are basically incoherent, but Socrates is less successful in his positive defense of the principle of morality. He relies here on an analogy between virtue or moral integrity, as an internal order of the soul, and the role of order and harmony in other domains: in the health of the body, in the order of the cosmos, and in the successful products of the arts. But an argument from analogy has its limitations. What is lacking here is a positive psychological theory (like the tripartite theory of the *Republic*) as the basis for a constructive argument in support of the conception of virtue as the harmony and health of the soul.

The most important positive doctrine of the *Gorgias* is the claim that all actions are done "for the sake of the good," that is, for a goal or *telos* that the agent perceives as good (467c–468b). This remains the fundamental axiom in action theory for both Plato and Aristotle; it reappears in the *Republic* as the supreme Form of the Good, "which every soul pursues, and for the sake of which it performs all its actions" (*Rep*. 505d 11). In the argument for this principle in the *Gorgias*, Plato deliberately blurs the distinction between good-for-the-agent and intrinsically or absolutely good. Polus will acknowledge that what people really want is something good, namely good for them or in their interest, but he denies that this is necessarily the honorable or moral thing to do (*to kalon*). Plato's point will be that moral knowledge consists precisely in the recognition that what is good absolutely (i.e., virtue) is also good for you, so that it is in your interest to be virtuous. This is Plato's reading of the Socratic paradox that no one is voluntarily unjust.

The *Gorgias* thus expounds, both by paradox and by systematic argument, the principles of Socratic moral philosophy as exemplified in the *Apology* and *Crito*. By contrast, in the *Hippias Minor* we find Socrates arguing for a more perverse paradox, namely, for the blatantly false proposition that anyone who commits unjust and dishonorable actions *voluntarily* is a better person than the one who does such actions unintentionally. The interlocutor is unconvinced, and we can only wonder what point Socrates is supposed to be making. This is probably an indirect way of calling attention to the more authentic Socratic claim that in fact no one does such actions voluntarily. But why not? Why does the analogy fail with arithmetic, for example, where the good mathematician makes mistakes on purpose, whereas the bad mathematician does so unintentionally? If moral virtue is a form of knowledge, why is it not to be understood on the model of the arts and sciences? The implicit Platonic answer seems to point to the role of intentions (the verb *boulesthai*, "to want," is systematically repeated at 366b–367a), and thus to the universal desire (*boulêsis*) for the good recognized in the *Gorgias*. Whatever the implied answer to this paradox may be, the *Hippias Minor* demonstrates Plato's early preoccupation with the problem of moral knowledge.

Finally, in the *Ion* Plato develops a different Socratic theme concerning knowledge: the refutation of knowledge claims on the part of the poets (*Apology* 22b). Instead of attacking the poets directly, Plato begins with their representative, the *rhapsode* or performer. Socrates' argument in the *Ion* is a direct refutation only of the claim to knowledge or art (*technê*) on the part of the rhapsode, but the positive theory of poetic inspiration applies to the poet as well. According to this theory, the power of poetry comes from the Muse and is transmitted via the poet to the rhapsode, like the attractive power that is transmitted from the magnet stone via iron rings to other pieces of iron. Hence neither the poet nor the rhapsode needs to understand what is going on. Their divine inspiration is non-cognitive: being possessed by a god, they are out of their mind.

The *Ion* thus presents Plato's first move in the ancient quarrel between philosophy and poetry, a quarrel that will be dramatically represented in the confrontation between Socrates and the poets in the *Symposium* and will assume canonical form in the criticism of poetry in the *Republic*. The *Ion* is indirectly invoked in the last scene of the *Symposium*, since it provides us with the argument that the narrator has forgotten. In this final episode Socrates is proving to Agathon and Aristophanes, a tragic and a comic poet, that anyone who knows how to compose tragedy by art (*technê*) will know how to compose comedy as well (223d). The needed premise is given by a proposition of the Ion, namely, that anyone who possesses the relevant knowledge (*technê*) will be able to deal with poetry as a whole, since it is a single art (532c). With a slight revision, Socrates' argument against Ion will serve as well against Agathon and Aristophanes. In contrast

with the *Republic*, where Plato criticizes poetry first on moral grounds (in Book 3), and then on principles of ontology (in Book 10), in the *Ion* and *Apology* Plato's criticism is more Socratic, rejecting the claims of poetry to be recognized as a kind of knowledge. It is thus aimed directly against the traditional conception of poets as *sophoi* or sages, sources of wisdom. For Plato the quarrel between philosophy and poetry is ultimately a culture war, a competition for the moral leadership of Greek society.

In the course of his attack on poetry in the *Ion*, Plato introduces the important epistemological principle of a one-to-one mapping between a *technê* and its object: "necessarily, the same art will know the same subject matter, and if the art is different, it will know a different subject" (538a). This correlation between a form of cognition and a definite content or subject matter appears frequently in other dialogues (for example *Gorgias* 464b, *Charmides* 171a). In the *Republic* this principle is invoked to show that knowledge and opinion must have different ontological objects (V, 478a); in the *Timaeus* a similar principle is implied as premise in an argument for positing Forms (51d). Problematic in its particular applications, this principle reflects Plato's fundamental realism in epistemology. Truth in cognition reflects reality in the object known: "What *is* completely is completely knowable; what is not in any respect is unknowable in every respect" (*Rep*. 477a 3).

DEFINITION AND APORIA: *LACHES*, *CHARMIDES*, *EUTHYPHRO*, *PROTAGORAS*

On a traditional view, these four dialogues provide something like a philosophical portrait of the historical Socrates: pursuing the topic of moral virtue, seeking definitions of the virtues (courage in *Laches*, temperance in *Charmides*, piety in *Euthyphro*), identifying virtue as a kind of knowledge, and denying the reality of *akrasia*. Most descriptions of the philosophy of Socrates are based upon the evidence of these dialogues, as supported by Aristotle's account. But if Aristotle's account of Socrates is derived from his own reading of these dialogues, his testimony is of no independent historical value. In at least one case Aristotle's report can be seen to be directly dependent on a Platonic dialogue, since for the Socratic denial of *akrasia* he quotes the *Protagoras* verbatim (*N.E.* VII.2, 1145b 24, citing *Prot*. 352c 1).

On the fictional view of the dialogues proposed above, what we have in the *Protagoras* and the dialogues of definition is not documentary evidence for the historical Socrates but rather Plato pursuing Socratic themes in his own way, and with his own philosophical goals in view. Thus the *Laches* and the *Euthyphro* offer a subtle lesson in the logic of definition, which will be completed in the *Meno*. And in the *Protagoras* we find something entirely new and problematic: a hedonistic anticipation of rational choice theory that is unparalleled in other dialogues.

Whatever Socrates' own concern with definition may have been (and there is no trace of this either in the *Apology* or the *Crito*, nor in the *Ion* and *Hippias Minor*), the treatment of definition in the *Laches-Charmides-Euthyphro-Meno* has a systematic quality and an epistemic orientation that is distinctly Platonic. Unlike the more straightforward search for a definition of rhetoric in the *Gorgias*, which does not raise epistemological issues but leads instead to a formula acceptable to all parties, the attempt to define virtues in these four dialogues of definition is formally aporetic and regularly unsuccessful. Although the search for a definition always fails, in two cases it points incidentally to an account of virtue as the knowledge of good and bad (*Laches* 198c–199e, *Charmides* 174b–e). In the *Protagoras* (as also in *Meno* and *Euthydemus*) virtue is again identified with some kind of knowledge.

The teachability of virtue is a topic debated at length in the *Protagoras* and *Meno*, and raised also in the *Laches* for the special case of courage. The claim of teachability seems to stand or fall with the conception of virtue as knowledge. The *Meno* makes explicit the principle implied at the end of the *Protagoras*: Virtue is teachable if and only if virtue is a kind of knowledge (*Meno* 87b). This assumption reflects the Greek sense that *technê* and teaching go together. But this principle raises the problem posed in the *Hippias Minor*: If virtue is a kind of knowledge, how is it different from other, more professional forms of *technê*? This question is briefly discussed at the beginning of the *Protagoras*: The young Hippocrates wants to study with Protagoras not for professional reasons, in order to become a sophist, but for liberal education, the training appropriate for a free man and citizen (312ab). This leaves open the question of what such training should consist in. We must wait for the *Republic* to get a definite answer to the question of the teachability of virtue. The *Protagoras* and *Meno* present arguments for both sides of the question (see below).

The dialogues of definition direct us to the theory of knowledge by two routes: first, by the suggestion that virtue, the target of definition, is itself a kind of knowledge. And second, by the claim that knowledge as such

depends on knowledge of essences. Thus in the *Laches*, where the two generals Laches and Nicias are being consulted as experts on training in virtue, the request for a definition is proposed as a test of their expertise: "if we know *what virtue is*, we should be able to say *what it is*" (190c). For if we did not know at all what virtue is, how could we advise anyone how to acquire it? (190b). Similarly, if Charmides is temperate, he should have some notion of what temperance is (*Charmides* 159a). In the *Meno* this type of question is justified by the general principle of priority of definition: One cannot know anything whatsoever about *X* unless one knows what *X* is (*Meno* 71b). We will return to this principle below, in discussing the *Meno*.

It is in the *Euthyphro* that the notion of essence or whatness, what *X* is, is most fully articulated as the object to be captured in a definition. To define piety one must specify something quite general, for the pious is "the same as itself in every action… similar to itself and having some one character (*idea*)" (*Euthyphro* 5d). The definiens must be not only coextensive with the definiendum but explanatory of it; necessary and sufficient conditions are not enough for a Platonic definition. Socrates wants to find "the very feature (*auto to eidos*) by which all pious things are pious." Only then will he be able to "look to this character (*idea*) and use it as a model (*paradeigma*), so that when any action is of this sort I will say that it is pious, and when it is not of this kind I will say that it is not pious" (6e). The definition offered by Euthyphro ("piety is what is loved by the gods") turns out to fail this test; it is a proprium, an attribute uniquely true of piety, but not an explanatory essence. Socrates complains to Euthyphro: "When you were asked *what the pious is*, you were not willing to reveal to me its essence (*ousia*, literally its being or is-ness), but you gave me instead an attribute (*pathos*), saying that it belongs to the pious to be loved by all the gods" (11a).

The distinction between an essence and an accidental attribute, so fundamental for Aristotle's philosophy, is here sharply delineated for the first time, but without clear metaphysical implications. In the dialogues of definition, including the *Meno*, essences are presented as logical or epistemological concepts, as items corresponding to a definition, an item true of all the cases, and hence able to serve as a criterion for the use of a term, but without any definite ontological interpretation. Despite the terminology of *eidos* and *idea*, which in later dialogues will serve to designate the Forms of classical Platonic theory, the essences of the *Euthyphro* and *Meno* are not articulated as structures in the nature of things, neither as immanent nor as transcendent forms. In this situation the reader is free to assume either that the author of these dialogues has not yet decided on an ontological interpretation for his definienda, or that he has chosen to reserve this task for other dialogues, such as the *Symposium* and the *Phaedo*.

TRANSITIONAL DIALOGUES? *LYSIS*, *EUTHYDEMUS*, AND *MENO*

These three dialogues present or allude to typical Platonic themes in epistemology and metaphysics, but without any definite formulation of what will be the standard theory of the *Phaedo* and *Republic*. Hence they are sometimes described as "transitional." It is again left to the reader to regard these statements either as deliberately incomplete or as reflecting Plato's own indecision.

Lysis and *Euthydemus* form with *Charmides* a literary group of dialogues with similar introductory episodes, presenting a charming school scene in which Socrates converses with handsome boys or adolescents. (The setting of the *Laches* is comparable, but in that dialogue Socrates converses only with the fathers and not with the boys.) The question of education is implicitly raised by the setting in each case, and discussed at length in the *Euthydemus* and *Meno*. Aside from the literary setting and the general theme of education, in other respects these three "transitional" dialogues are very different from one another.

The *Lysis* is concerned with the topic of friendship and love, a topic discussed below in connection with the *Symposium* and *Phaedrus*. There are a number of parallels between the *Lysis* and *Symposium*), the most striking of which is the concept of a final object of love for the sake of which everything else is loved. In seeking to explain in the *Lysis* why anything is dear or desirable (*philon*), Socrates suggests that one thing is dear for the sake of another, as a doctor is desirable for the sake of health, but that such a regress cannot go on indefinitely: "we must either give up or come to some starting-point (*archê*), which will no longer refer to some other dear, but we will come to "that which is primarily dear" (*prôton philon*), for the sake of which we say that all other things are dear … This is what is truly dear; the other dear things are like its images" (*Lysis* 219c 5–d5). Since the form of the argument resembles Aristotle's thesis (in N.E. I.7) that happiness is the supreme good, for the sake of which everything else is good, some scholars have used this parallel to interpret the Lysis passage as a reference to happiness. But there is nothing in the text to justify this interpretation. On the contrary, the formula "for the sake

of which" refers to the good in passages cited above from the *Gorgias* and *Republic* (section IV). Furthermore, the context in the *Lysis* identifies the "dear" (*philon*) as the good and the beautiful (216c 6–d2). Above all, the formula "that which is primarily dear" (*ekeino ho esti prôton philon*) is a close approximation to the standard terminology for the Forms in other dialogues, and specifically for the Form of Beauty in the *Symposium* (*auto ho esti kalon* 211d 1). This anticipation of the technical language for Forms, together with the generally quite abstract form of the arguments about friendship, sets the *Lysis* apart from more typical "early Socratic" dialogues such as the *Laches* or the *Euthyphro*.

The *Euthydemus* is equally non-standard for other reasons. Plato presents an entertaining satire on two elderly sophists, the brothers Euthydemus and Dionysodorus, who claim to teach virtue by a shortcut method, and who display their art by confounding the student with a rapid series of fallacious arguments. Their art of unscrupulous refutation, or eristic, is designed to provide the sharpest possible contrast with the genuine Socratic elenchus, represented here not by the usual refutation but by a constructive protreptic in which Socrates argues that wisdom is the only good, ignorance the only evil, and hence that in order to enjoy happiness and a good life (*eu prattein*) one must pursue wisdom and knowledge.

Both Socrates' protreptic and several of the Sophists' refutations contain enigmatic allusions to Platonic doctrines presented in other dialogues. In the most surprising of these allusive passages, the young Clinias compares mathematicians to hunters because they must turn over their findings to someone else. Just as hunters turn over their catch to cooks, who know how to make good use of it, mathematicians, if they are wise, will turn over their discoveries about reality (*ta onta*) to dialecticians (*hoi dialektikoi*) to make use of (*Euthydemus* 209c). This subordination of mathematics to dialect is scarcely intelligible without the epistemology of Books VI and VII of the *Republic*. But this is not the only case where the *Euthydemus* anticipates doctrines to be developed in later dialogues, including an allusion to recollection (296d 1) and a hint that the relativism of Protagoras may be self-refuting. (Compare *Euthydemus* 286c 2–4 with the *peritropê* argument of *Theaetetus* 170a–171c.) There is also a rough version of the principle of non-contradiction (293b 8–d 1), and a kind of caricature of the problem of the presence of "the beautiful itself" in the many beautiful things (300e–301a). The *Euthydemus* is thus one of the most comical and also one of the most puzzling of all the dialogues.

MENO AND RECOLLECTION

The *Meno* introduces the doctrine of recollection, which plays an important role in two later dialogues, the *Phaedo* and *Phaedrus*. Like the sixteenth-century theory of innate ideas which it inspired, Plato's doctrine of recollection is an antecedent both for the Kantian notion of a priori knowledge and for contemporary theories of innatism in psychology. The fundamental thesis of the Platonic doctrine is that there is something in the nature of the human mind that predisposes it to grasp the nature of reality: "the truth of beings (*ta onta*) is forever in our psyche" (*Meno* 86b 1). The supernatural form this doctrine takes in Plato is determined by its association with the Pythagorean doctrine of transmigration, which implies a previous existence for the human soul. The *Phaedrus* give a mythical account of prenatal experience, in which human souls travel with the gods outside the heavens, to a vision of ultimate reality described in terms of the Platonic theory of Forms. It is our recollection of this prenatal vision of transcendent Beauty that explains the phenomenon of falling in love.

In the *Phaedo* as well recollection takes as its object the eternal Forms, illustrated in this dialogue by the Equal itself, as distinct from sensible equals. This choice of the Form of Equality in the *Phaedo* connects recollection with mathematics, as in the *Meno*, where recollection is illustrated by the geometry lesson to an uneducated slave boy. (The boy is led by a series of questions to see, first, that he is unable to double a square by numerical additions to the side, and then to recognize the solution when Socrates draws the diagonal.) But it is not only mathematical concepts but conceptual thought generally that is involved in recollection. As the *Phaedrus* insists, a human soul must be able "to understand what is said according to a form (*eidos*), passing from many sense perceptions to a unity gathered together by rational thought. And this is recollection of what our soul once saw when it traveled together with a god and looked beyond what we now call reality and was able to rise up into the really Real" (*Phaedrus* 249bc). The myth of the *Phaedrus* thus represents Plato's most brilliant expression of the classical Greek view that reason (*nous*), the cognitive capacity to understand the world, constitutes the immortal, godlike element in the human psyche.

The *Meno* presents a simpler version of the doctrine, without any explicit reference to the theory of Forms. Recollection is introduced in response to Meno's paradox about learning something new, or seeking for something you do not know. Meno in turn is responding to the principle of "Priority of Definition," which claims that you

cannot know anything about *X* unless you know "what *X* is." How then do we ever get started? Recollection answers that what we learn is not new; we only need to be reminded. In the fuller doctrine formulated in the *Phaedo* and *Phaedrus*, it is not Socratic questioning but sense perception that serves to trigger a conceptual understanding (of equality, beauty, and the like) that is provided by the mind from its innate resources.

The "transitional" status of the *Meno* is indicated not only by the fact that it presents the simplest version of recollection, but also by tentative statements of other themes that are more fully developed in the *Phaedo* and *Republic*: the distinction between knowledge and opinion, the method of hypothesis, and two levels of virtue, one dependent on right opinion and the other on knowledge (*Meno* 99a–100a).

PLATO'S THEORY OF *ERÔS*

Love is a central topic in three Platonic dialogues (*Lysis*, *Symposium*, *Phaedrus*); it also plays an important role in the moral psychology of the *Phaedo* and *Republic*. The fundamental idea is expressed symbolically in Plato's etymological reading of *philo-sophia* as love of wisdom or passion for knowledge (*Phaedo* 66e2, 68a). In the psychological theory of the *Republic*, all three parts of the soul are characterized as distinct forms of love: desire for learning (*to philomathes*), desire for honor, desire for pleasure and wealth. Thus for the rational part the object of desire is "to know the truth" (581b). Like the religious mystics, Plato makes use of the language of sexuality to express philosophical passion: the true lover of knowledge will not be relieved of his pangs of *erôs* "until he grasps the nature of each Form with the appropriate part of the soul, and clinging to and mingling with the truly real, he begets truth and understanding (*nous*)" (490b). Plato anticipates the Freudian notion of sublimation in his account of the channeling of desire (485d); the notion of unconscious Oedipal desires is recognized in his description of criminal dreams (571c–d). There is also a superficial analogy between Plato's tripartite psychology and the Freudian trio of ego, superego, and id, but the second principle is in fact quite different in each case. Plato's *thumos* or "spirit" is a principle of anger, pride, and self-assertion, in contrast to the guilt-producing and self-punishing aspects of the Freudian superego. What the two psychological theories have in common is the understanding of psychic conflict in terms of deep divisions within the soul.

Plato's theory of *erôs* has been criticized for devaluing the love for an individual person in favor of love for an abstract principle like the Forms. Thus in the ladder-of-love passage in the *Symposium*, the lover who follows Diotima's instructions will leave behind his initial passion for an individual beauty in order to rise to more spiritual beauties and finally to the Beautiful itself. Even in the *Phaedrus*, where the philosophical lovers assist one another in growing the wings of their souls and escape together from the cycle of rebirth, their real love is for the Form of Beauty. But it is misleading to evaluate the Platonic conception of *erôs* as if it were a contribution to the modern theory of love. Plato's concern with interpersonal love is better illustrated by his treatment of friendship (*philia*), as depicted in the case of the two boys Lysis and Menexenus in the *Lysis*. So it is the *Lysis* that provides Aristotle with the starting point for his own theory of friendship. The philosophical importance of *erôs* for Plato lies not in its role as a relation between persons but rather in its function as the energy driving us to pursue what we take to be good (or good-and-beautiful) and hence, when properly enlightened, to pursue the Good itself. Rightly directed, *erôs* is *philo-sophia*, the passion for wisdom. Only wisdom can recognize the true nature of the Good, "which every soul pursues and for the sake of which it performs all its actions" (*Rep*. VI, 505d11). It is in this sense, as knowledge of the good, that wisdom is equivalent to virtue, since it guarantees that the *erôs* of the wise will be directed to what is objectively good. The emotional drive in question is, however, intrinsically ambivalent; in the absence of wisdom, *erôs* can also become the criminal passion that impels the tyrant to psychological destruction in *Republic* Book IX.

VIRTUE AND KNOWLEDGE: PLATO'S MORAL PSYCHOLOGY

Many scholars have followed Aristotle in holding that Socrates identified virtue with wisdom and thus ignored the power of irrational emotion to influence action. The conception of virtue as a form of knowledge is represented in a number of dialogues. The neglect or denial of irrational emotion is most extreme in the *Protagoras*, where Socrates interprets *akrasia* as an error in measuring future pleasures and pains. What is generally understood as being overcome by passion is there explained as an intellectual mistake. No other Platonic text explicitly denies the reality of *akrasia*. But several passages in the *Gorgias* and *Meno* have been taken to imply this, by suggesting that everyone desires good things, and hence that virtue consists only in the knowledge of good and bad, that is, in the ability to choose the goal of action correctly.

Nowhere, however, does either Plato or Socrates maintain that *all* desires are desires for the good. On the contrary, the *Gorgias* implicitly distinguishes between *boulesthai* as desire for good things and *epithumia* as desire for pleasure (so explicitly at *Charmides* 167e; this distinction between rational desire or *boulêsis* and non-rational desire or *epithumia* becomes fixed in Aristotelian terminology). The doctrine that virtue is a kind of knowledge can be understood as a paradoxical exaggeration, designed to focus attention on the practical importance of a correct conception of the good, and hence on the value of the Socratic elenchus in leading interlocutors to recognize their own ignorance. But in the face of this exclusive focus on moral knowledge, the existence of *akrasia* (that is, of people acting against their better judgment) is a challenge. The last section of the *Protagoras* was written in response to this challenge. But some readers will doubt that either Plato or Socrates ever held the extreme view presented in this dialogue, namely, that the intellect is all-powerful in the control of human action, so that *akrasia* is simply an error of judgment and vice is always due to ignorance.

What is clear, in any case, is that if Plato ever held such an intellectualist view, he abandoned it in the *Republic*. The exposition of the tripartite psychology includes an unmistakable description of *akrasia* in the story of *Leontius* (who is disgusted at his own weakness in "being overcome by the desire" to gaze at corpses, *Republic* 440a 1). In this tripartite theory, two out of three psychic principles represent emotional drives that can conflict with, and sometimes overcome, the rational judgment of the *logistikon* (the calculating part) as to what is best to do. These two principles are the *thumos*, or "spirit" of anger and pride, and the *epithumêtikon* of animal appetite—hunger, thirst, and sexual desire. The division into three parts rests upon a careful distinction between sheer desire, for example thirst as desire to drink, and the rational desire for something good, as desire for a good drink. The aim of Plato's tripartite division is precisely to account for the phenomena of psychic conflict, in this case between the desire of a thirsty man to drink and his rational judgment that the water is not good to drink.

On the basis of this division into three parts of the soul, each with its characteristic desire, Plato provides a psychological definition of the virtues in terms of the harmonious working together of all three parts. It is the function of the rational part (*logistikon*) to rule over the others in deciding what is the best thing to do; and wisdom is the excellence of this part in judging well. Courage is the excellence of the spirited part, maintaining its loyalty to the commands of reason and law in the face of danger and temptation. The other virtues consist in cooperation, that is, in willing obedience to the commands of the rational part. Hence virtue can be defined as psychic harmony, and vice defined as psychic disorder or *stasis*, civil war between the parts of the soul.

By this assimilation of virtue to psychological health, vice to psychological disorder, Plato formulates his first answer to the challenge to morality (formulated by Thrasymachus in Book I, reformulated by Glaucon and Adeimantus in Book II). But the *Republic* actually represents two different views of psychic disorder. In Book IV the vices are described in terms of disobedience or revolt on the part of the irrational emotions; in this context, there is no distinction to be drawn between vice and *akrasia*, conceived as unruly behavior by the lower parts. (This is also the picture of vice presented by the behavior of the disobedient horse in the *Phaedrus* myth.) In Books VII and IX, on the other hand, the irrational desires are presented not as disobedient subjects but as successful rebels, who have driven reason from the throne and taken its place as rulers in the acropolis of the soul (*Rep*. 553d, 560b–d). The *logistikon* now appears as their subject, carrying out their commands. Thus we have in Books VII–IX a conception of vice represented not as *akrasia*, not as a failure of reason to control the emotions, but rather as moral ignorance, that is, a mistaken conception of the good (as in Aristotle's distinction between vice and *akrasia*).

This Platonic distinction between two conceptions of vice, only one of which corresponds to *akrasia*, is developed in different ways in several later dialogues. Thus the *Sophist* (228a–229a) distinguishes moral ignorance from *ponêria*, vice as a kind of disease; the former is to be treated by instruction, the latter by punishment. The *Timaeus* 86b–e proposes a similar distinction between moral ignorance and madness due to excessive pleasures and pains; the latter is caused by a disordered condition of the body. The Socratic paradox will be maintained for both kinds of vice, since the loss of self-control from bodily causes can be seen as involuntary (*Tim*. 86e 3). The connection of the non-rational desires with the body rather than with the soul proper, hinted at in the *Phaedo* and in *Republic*, is most systematically developed in the *Timaeus* (42a–e), where the non-rational soul is created by the lesser gods in connection with their creation of the body.

POLITICAL CONSTRUCTION: FROM THE *REPUBLIC* TO THE *LAWS*

The tripartite psychology of the *Republic* has an exact parallel in the tripartite social structure of the envisaged

polis. Corresponding to reason, spirit and appetite are the three classes of rulers, soldiers, and producers (the latter class consisting of farmers and craftsmen). Scholars have suggested that the psychic tripartition is an artifact of this parallelism, and that Plato's moral psychology would more properly take the form of a bipartition into reason and emotion, as in Aristotle's *Nicomachean Ethics* (I.13) and in modern theories of action based on belief and desire. (Plato actually flirts with such a bipartite psychology in Socrates' first speech in the *Phaedrus* 237d–e). However, Plato remains loyal to the tripartite psychology in non-political settings as well (in Socrates' second speech in the *Phaedrus*, and in the *Timaeus*). There is a theoretical advantage to recognizing more than one type of non-rational emotions, some of which are more amenable than others to rational control.

It is essential to the scheme of the *Republic* that the city is conceived as a great organism, just as the psyche is conceived as a micro-community. Unity and cohesion are fundamental principles of excellence for the city as much as for the individual. Plato's political aim, the greatest good for the city, is for the citizens to share one another's joys and sorrows with a unanimity like that of the parts of a human body, where the whole person suffers if a single part is in pain (V, 462a–e). But this organic unity can be achieved only on the basis of a functional division of labor between the three social classes. Thus the political definition of justice in terms of each group doing its proper job (in Book IV) is prefigured by the initial division of labor through which the city comes into being (in Book II). The first society arises from the mutual need of individuals for one another: one to grow food, one to build houses, one to make clothes. Hence the fundamental principle of specialization: one person, one work.

Instead of a social contract theory, in which civil society is conceived as an artifact designed to bring people out of the state of nature, Plato claims to find a natural basis for social life in reciprocal need and the advantages of cooperation (II, 369–370). He thus sees human beings as by nature friendly and cooperative, in deliberate contrast to the Hobbesian view of human nature presented by Glaucon in the ring of Gyges story earlier in the same book (II, 358e–362c). Since the division of labor is to the advantage of all in the political as well as in the economic sphere, the city of the *Republic* will have a natural cohesion that is absent from the historical cities of Greece, which are (as Socrates observes in a moment of Marxian insight) really two cities, the city of the rich and the city of the poor (IV, 422e–423a). This pathological split will be avoided in Plato's city, because there the ruling classes will have no private property, no money, and no nuclear family to generate selfish preferences. The needs of the rulers will be provided for by the farmers and craftsmen, who alone will have private belongings and wealth. Thus the ideal city will radically separate economic power from political power; the rulers, who alone possess the latter, will be systematically excluded from the former.

The political structure of the *Republic* is built up in successive stages, beginning with cooperation and division of labor, then the division into three classes, followed by three culminating waves of paradox in Book V. The first wave is the principle of equal education and access to political power for gifted women; the second wave is the community of wives and children, in other words, the abolition of the nuclear family. (This innovation brings with it some extraordinary marriage arrangements requiring a great deal of systematic deception on the part of the rulers. The principle of benevolent deception was established earlier, in presenting the myth of metals as a noble lie in Book III, 414b–415c.) The third wave, and the condition of possibility for the entire scheme, is rule by philosopher-kings. Only philosophers are competent to rule the city, because only philosophers have access to the Form of the Good and the Form of Justice, the knowledge of which is strictly necessary if the rulers are to make the city just and good. The system of education designed to produce these rulers will be discussed below.

Did Plato abandon these ideals in his later work? An answer to this question is provided in two documents, the *Statesman* and the *Laws*. The *Statesman* is a puzzling work. It purports to define the statesman, or *politikos*, and to show how he is different from the philosopher. It then defines *politikê*, the art of statesmanship as a kind of knowledge or understanding that is competent in giving orders, that is, in ruling. But the dialogue never specifies the content of this expertise. It says only that the possession of such knowledge by the ruler (or rulers) is the one indispensable condition of a genuine constitution; all other constitutions can be no better than imitations of this model. Constitutions are ranked by two criteria; the old classification according to rule by one, few or many is now crossed with the new criterion of lawful or lawless. As lawless one-man rule, tyranny is still the worst form of government, but democracy is now the least bad; the best imitation of the model is a constitutional monarchy (302b–303).

How is the ideal model of the *Statesman* essentially different from the constitution of the *Republic*? More precisely, how is this ruler with authoritative knowledge dif-

ferent from a philosopher-king? If we assume that the *Republic* is in the background, we can see Plato as returning here to familiar territory but from a very different point of view. The Forms are not mentioned as objects of the statesman's expertise (although the dialogue does refer to incorporeal and non-sensible realities); nothing whatever is said about the content of the statesman's knowledge or the nature of his training. We are told only that he will act with justice, and so as to make the city and the citizens better (293d8–9, 297b 2). So presumably the perfect statesman must know what is justice and what is good. Whether or not he knows them as transcendent Forms is left for the reader to surmise.

The one point of general theory that is carefully discussed in the *Statesman* is whether or not the true ruler's knowledge should in principle be supreme over and above the law, and the answer of this dialogue is a resounding "yes." The regime of legality is an imitation, a second-best, in the absence of the scientific ruler. But nothing in the human world can be superior to genuine knowledge.

At first sight, the position of the *Laws* is diametrically opposed, for here Plato provides the first philosophical argument in favor of the principle that a city should be ruled by laws rather than by men, and that human rulers should be servants of the law (*Laws* IV, 715c–d). Law, indeed, is said to be "the dispensation of reason (*nous*)" (714a). But on a closer look the two texts are not so far apart, since the omniscient ruler of the *Statesman* is not to be found among us, and according to that dialogue also the best human constitutions must be law-abiding. In the *Laws*, despite the shift in favor of the rule of law, Plato still yearns nostalgically for the unfettered authority of the truly wise ruler. He is now convinced that human nature is too weak to bear such unlimited power and still remain uncorrupted (IV, 713c with IX, 875b, the source of Acton's principle that absolute power corrupts absolutely). But if such a man could be found, he would not need to be controlled by laws. "For neither law nor any order is superior to knowledge; and it is not right for reason (*nous*) to be subordinate to anything" (IX, 875c). This is precisely the thesis of the *Statesman*.

But the author of the *Laws* has given up hope of the messianic politics sketched in the *Republic*. The detailed constitution of the twelve books of the *Laws* presents a complex political system tightly controlled by an extremely precise legal code, with many invasions of individual liberty, and a social structure very different both from that of the *Republic* and also from that of fourth-century Athens. The society of Plato's last city prefigures that of Aristotle's *Politics*, Book VII. In both constructions one social class possesses all the property and is the only group to bear arms and to have political rights, while the mass of the population—the producer class of the *Republic*—is disenfranchised and reduced to slavery or limited to foreign residents. In the *Laws* the city has become a club of the leisured class, whose members can devote all of their time to the practice of political virtue, to the study of the law code. and to ritual celebrations in song and dance.

The city of the *Laws* is an entirely new project, based upon a different political philosophy in which the rule of law is supreme. The constitution includes several realistic political institutions, representing a compromise with Athenian democracy, which introduce a career of public service into the utopian life style of this privileged class of citizens. But despite all these innovations, one fundamental principle of the *Republic* has been preserved. Although there is no place for a supreme philosopher-king in this law-bound aristocracy, a kind of counterpart is nevertheless preserved in the institution of the Nocturnal Council, introduced at the end of the work. This Council is a group of high officials meeting daily to study the philosophical foundations of legislation, and to revise the laws if need be. To this extent the author of the Republic remains loyal to himself. The construction of a good constitution will still require the presence of philosophy in a position of the highest influence.

RHETORIC AND DIALECTIC: *GORGIAS* AND *PHAEDRUS*

Rhetoric, the art of public speaking, was developed by the Sophists into a powerful instrument of political leadership; and Plato's chief rival as an educator was the orator Isocrates. Corresponding to its important role in Greek society, rhetoric is a frequent topic of the dialogues, notably in the *Gorgias* and the *Phaedrus*, but implicitly also in the *Protagoras*. In the latter dialogue Socrates presents his own art of question-and-answer as an alternative to, and ultimately a victor over, the art of long speeches represented by Protagoras.

The contest between Socrates and Protagoras is thus a contest between two forms of *logos*, two methods for winning an argument. In the *Gorgias* this contrast of methods reflects the deeper contrast between values. The goal of Socrates' rhetorical opponents is wealth and power, and their speeches aim to persuade the majority. Socrates' goal is virtue and knowledge, and his methodology is designed to get only the agreement of his interlocutor (472b). Socrates' characteristic device is the

elenchus: deriving a denial of the interlocutor's thesis from premises that the interlocutor will accept. This is the method that Plato describes retrospectively in the *Sophist*: if someone claims to have knowledge who is in fact ignorant, "since his opinions are confused, it is easy to examine him and to bring these opinions together in discussion and, setting them side by side, to show that they contradict one another" (230b). In the *Gorgias*, Socrates refers to this as the art of conversation (*dialegesthai*) in contrast to the art of speech-making (*rhetorikê*) (448d 10). But Socratic dialectic must also be distinguished from eristic, the pursuit of contradiction for its own sake (illustrated by the notorious behavior of the two sophists in the *Euthydemus*, above in section VI). Unlike this frivolous form of refutation, the Socratic elenchus is designed to free the interlocutor from the false conceit of knowledge, so that the way is opened for him to begin to learn.

In the *Republic* Plato will transform dialectic, as the art of question and answer, into a much more ambitious and constructive method. We look first at his treatment of rhetoric, which is quite different in the *Gorgias* and the *Phaedrus*. In the *Gorgias*, rhetoric is represented by Socrates' opponents, and in particular by Gorgias, the most famous orator of the late fifth century, and teacher of Plato's rival Isocrates. Gorgias stands for the political power of unscrupulous persuasion, and thus for power without moral responsibility or even, in the case of his followers Polus and Callicles, for power without moral restraint. In the *Gorgias*, Socrates argues that the rhetorical practice of public persuasion, without principles of justice and without knowledge, is not an art at all, not a *technê* but a mere empirical knack. To qualify as a *technê* rhetoric would need the theoretical clarity and contact with truth that are characteristic of knowledge. As seen in the *Gorgias*, rhetoric clearly lacks both.

In the *Phaedrus*, by contrast, Plato is concerned with rhetoric not as an instrument of political power but as the form of prose literature, and his sample is not a political speech but a series of epideictic displays on the topic of love. Socrates surprises his interlocutor by not limiting the rhetorical art to speeches in law courts and in public assemblies but generalizing it to cover "the bewitchment of the soul through discourse" (*psychagogia dia logôn*, 261a8). Rhetoric is here conceived as the art of speaking and writing well. Plato makes one of his notable contributions to literary criticism in the discussion of what he calls "literary necessity" (*logographikê anankê*) linking the parts of a composition to one another. Socrates observes that a discourse (*logos*) should have an organic form, like a living creature, "so as to be neither headless nor footless, with middle parts and extremities that are fitting both to one another and to the whole" (262bc). It turns out that to produce discourse with this quality, the author must be able to gather similar things into unity, and also divide them by kinds. The art of these collections and divisions is called "dialectic" (266c), and it seems that a true art of writing or speaking must include or presuppose dialectic. If rhetoric is to be a *technê*, it will not follow the path of the professional orators (269d). True rhetoric would, for instance, require a philosophic understanding of the psyche and of its natural varieties (271d). Like the *Gorgias*, the *Phaedrus* ends by rejecting the claims of ordinary rhetoric to be regarded as a technê. But if Plato's judgment of rhetoric in this dialogue tends to be much more positive than in the *Gorgias*, that is because the art of *logoi* is here conceived constructively as the art of writing, including philosophical writing, and hence as an application of dialectic rather than an alternative to it.

While dialectic was introduced in the *Gorgias* and elsewhere as the Socratic art of conversation (*dialegesthai*) conducted in question-and-answer form, in the *Republic* it becomes the highest method of philosophy, the method by which the intellect ascends to the cognition of transcendent Forms. More specifically, it is the method of passing beyond the assumptions (hypotheses) that function as premises of reasoning in the deductive sciences of mathematics. Dialectic thus presupposes the method of hypothesis developed in the *Meno* and *Phaedo*, a method derived from mathematics, according to which a problem can be solved conditionally on the basis of an explicit assumption. By subjecting these assumptions to critical scrutiny, dialectic is somehow able to rise above them and thus reach the *anhypotheton*, the object of unconditional knowledge, in other words the Forms (VI, 511b). The actual practice of dialectic is not described, but its study follows ten years of training in mathematics. Its connection with the conversational method of question-and-answer is preserved in the requirement that the dialectician must be able to "give an account (*logos*) of the being (*ousia*) of each thing" (VII, 534), that is, to give a systematic answer to the question "What is it?" Giving such an account will necessarily involve a reference to permanent essences or Forms.

Dialectic is described quite differently in the later dialogues, but it remains the highest form of knowledge, the essential method of philosophy. It continues to proceed by question and answer, and to seek the definition of essences in answers to the question "What is X?" According to the *Philebus*, dialectic still takes as its object "true

being which is forever unchanging," the reality "which neither comes to be nor passes away" (58a2, 61e2)," precisely the kind of Being represented by the Forms in Plato's classical theory. In the *Sophist* and *Statesman*, however, as in the *Phaedrus*, dialectic is described in more formal terms, as the method of gathering pluralities into unities and dividing them into kinds (*genê*), where the term *eidos*, which designates a transcendent entity in the classical theory of Forms, seems to be used in the more strictly logical sense of "species" or sub-kind. Instead of the relation to mathematics and the method of hypothesis, which is fundamental for the conception of dialectic in the *Republic*, it is the method of Division that is central for dialectic in the later dialogues, from the *Phaedrus* to the *Philebus*. This shift in the description of dialectic corresponds to a different, less metaphysical way of referring to the objects of knowledge (see further below).

ESTHETICS AND EDUCATION: PLATO AGAINST THE POETS

Can virtue be taught? That is the question raised dramatically in the school scenes of several early dialogues, and discussed at length in the *Protagoras* and *Meno*. The conclusion of the *Meno* is problematic. Socrates insists that we must first define virtue before we can answer this question. Since we have no definition, we must answer it conditionally. If (and only if) virtue is a kind of knowledge, it is clearly teachable. But such virtue is hard to find. What about virtue based on correct opinion (*doxa*)? It might give the same results as virtue based on knowledge, but would it be teachable? The *Meno* ends without any clear statement on the question of teachability.

If there is a Platonic answer to this question, it must be found in the educational scheme of the *Republic*. There is a different but parallel answer in the scheme of education in the *Laws*. For Plato (as later for Aristotle), an essential function of the city is to make its citizens good, that is, virtuous. Hence education is a central concern in both dialogues. The *Republic* describes two stages of education, one for the wider guardian class (in Books II and III) and one for the select group of future rulers (in Book VII). Corresponding to these stages we have two accounts of virtue, one based on right opinion (in Book IV) and one on philosophic knowledge (Books VI–VII). The limitations of the initial account of the virtues in Book IV are visible only retrospectively, after the distinction between knowledge and opinion is drawn in Book V. Only after this introduction of philosophy can we appreciate the ambiguous status of wisdom, and hence of virtue generally, as defined in Book IV.

In order to become virtuous, the entire guardian class must have the basic system of education described in terms of music and gymnastics. Only a smaller group will enjoy the training in philosophy, consisting of ten years of mathematical science followed by dialectic and culminating in the vision of the Form of the Good. The first stage of education will produce "citizen excellence" (*politikê aretê*); the higher education, accompanied by years of public service, will yield the unqualified virtue of the philosopher-kings. If we take this as Plato's answer to Meno's question "Is virtue teachable?" the answer is: yes, but not by the available means of education. Only a fundamental change in the conditions of social life would make it possible to produce in a regular way the kind of excellence that occurs sporadically today, by good luck or (as the *Meno* says) by divine dispensation.

Under the more favorable conditions of Plato's city, the character of the guardians will be shaped by a carefully controlled cultural environment, that will include a radical change in the literary and musical content of their education. Plato here defines his position in the culture war he describes as the ancient quarrel between philosophy and poetry (*Rep*. X, 607b; see above). All of the arts will play an essential role in the moral education of the young guardians, but it is poetry that is the center of Plato's attention, because of the fundamental influence of Homer and the tragedians on Greek moral thought. Since Plato regards their influence as essentially malignant, he would eliminate from his educational scheme major themes of Greek poetry (Books II–III). Following Xenophanes and others, he attacks as immoral the Homeric depiction of the gods. His basic theological principle is that the gods are good, and are therefore (by the law of transitive causation) cause only of the good, and they must be represented accordingly (III, 379a–380c). Plato thus avoids the thickets of theodicy; there is no problem of justifying the action of the gods, since they are never responsible for evil. The actions of glorified heroes must also be represented in such a way as to provide a moral paradigm for the young guardians.

Finally, when Plato returns in Book X to the restrictions on poetry, he attacks the imitative arts generally on epistemic grounds, as being at third remove from truth. He also blames the emotional impact of epic and tragedy for relaxing the moral discipline of the soul. Hence the poets are to be banned from Plato's city, and readmitted only if their influence can be morally justified (697b–e). This is a famous challenge to future aesthetic theory. Aristotle's *Poetics* and Sir Philip Sydney's *Defence of Poesie*

count among the more noteworthy responses to Plato's challenge.

At the same time, a properly controlled aesthetic environment is recognized as decisive for the development of virtue in the young. This includes the visual arts, but poetry and music are of particular importance, since rhythm and harmony, more than anything else, penetrate deep into the soul (III, 401d). Because of the close connection between the beautiful and the good, the young should be surrounded by beauty in all its forms, so that later, when moral principles are presented to them in rational teaching (*logos*), they will recognize these as familiar and congenial (402a). The positive use of the arts in education is developed further in the *Laws*, with special attention to dance, since there will be choruses for the citizens at all ages (Books I–II). Literature and lyre-playing will be essential in education, and the Athenian Stranger who speaks for Plato in the *Laws* holds up the Platonic dialogue, and specifically the text of the *Laws*, as a model for the literature to be used in schools (*Laws* VII, 802 ff; 811c-e). In both the *Republic* and the *Laws*, the content of literature and music is interpreted in moral terms: "rhythms and the performing arts as a whole (*pasê mousikê*) are the imitations of the characters of better and worse human beings" (*Laws* VII, 798e).

Plato's positive evaluation of poetry, implicit in his use of literature in these proposed schemes of education, receives a theoretical development in the account of poetic inspiration as divine madness in the *Phaedrus* (245a). In the *Ion* (as in the *Apology*) the notion of divine possession for the poet was employed ironically, in order to emphasize the poet's lack of cognitive competence. In the *Phaedrus*, on the other hand, the madness of artistic inspiration is presented as a positive force, in parallel with the divine madness of love which carries us back in recollection to a prenatal vision of the Forms. Plato never says that artistic experience, like erotic experience, can trigger recollection of the Forms. But it is easy to see how a later Platonist such as Plotinus (and his followers, such as Proclus), less fearful than Plato of the moral and intellectual dangers from poetry, could make use of the *Phaedrus* parallel between poetry and love to develop a powerful conception of art as a privileged mode of access to a higher level of metaphysical reality. This was a theory much in vogue among the Romantics of the nineteenth century, who took over Plato's theory of poetry as divine possession, deprived it of its ironic sting, and transformed it into a theory of creative genius.

THE CLASSICAL DOCTRINE OF FORMS

The centerpiece of Platonic philosophy is the metaphysical theory of Forms or Ideas, presented in three dialogues (*Symposium*, *Phaedo*, and *Republic*), utilized in two others (*Cratylus* and *Phaedrus*) and criticized in a sixth (*Parmenides*). Whether some version of this theory reappears in dialogues later than the *Parmenides* is a question to be discussed below. The term "idea" is a transliteration of *idea*, one of Plato's terms for the Forms. Since the English word suggests something mental or psychological, "idea" seems misleading as a designation for Platonic Forms, which are clearly intended to be mind-independent realities.

As we have seen above, the dialogues of definition present the object of definition as the being or essence (*ousia*) of the subject under discussion and distinguish it from an ordinary property or attribute (*pathos*). The essence is not only true of all and only instances of the subject, but it is also explanatory of *being the thing in question*. The answer to a question "What is *X*?" should say what *X* is, in the sense of explaining what makes something an *X*. (*Meno* 72c 8.; *Euthyphro* 6d 11). Thus being dear to the gods, although true of all and only pious actions, does not say *what pious is*, because it does not tell us why the gods favor some actions rather than others (*Euthyphro* 11a). These logical properties of essences prepare for, but do not imply, the metaphysical doctrine of the *Phaedo* and *Republic*.

Similarly, the terminology for *definienda* in the dialogues of definition prefigures the later terminology for the Forms, but in a pre-theoretical way: *eidos* and *idea* are ordinary terms for features, structures, or kinds of things. Aristotle says that Socrates was pursuing universal definitions, but that he did not separate the universals as Plato did (*Met.* 1078b 30). Hence some scholars have interpreted the essences of *Meno* and *Euthyphro* as immanent (Aristotelian) rather than transcendent (Platonic) forms. But the texts do not support such a distinction. For example, the idea of piety is described as a model (*paradeigma*) for deciding whether a given action is pious (*Euthyphro* 6e); but the *Euthyphro* does not tell us whether this model would be located in the mind or in the nature of things. The ontology of the *definienda* in these dialogues is left strictly indeterminate.

Plato supplies a metaphysical framework for the objects of definition in the *Symposium* and *Phaedo*, with a further development in the Republic. These dialogues introduce the conception of eternal, unchanging Being as location for the objects pursued in dialectic. Plato has taken over from Parmenides this notion of Being or

What-is (*to on*) as an unchanging reality accessible only to thought or rational understanding (*nous*), defined by contrast with the changing realm of Becoming that is accessible to the senses. Plato's conception differs from that of Parmenides in two respects: Platonic Being exists in the plural (*ta onta* or The-things-that-are), corresponding to the plurality of Forms, while for Parmenides, Being is a unique One; and Becoming is allotted a certain measure of reality, whereas its ontological status for Parmenides seems to be that of appearance only.

This metaphysical conception of the Forms, which is assumed throughout the argument of the *Phaedo*, is most succinctly formulated in the final description of the Beautiful in Diotima's lesson on love, as reported by Socrates in the *Symposium* (210e–211b). The Form of Beauty (literally "the Beautiful itself") is there distinguished from the many beautiful things by (1) being one (unique) rather than many; (2) being a Being rather than a Becoming, that is, being eternally and unchangeably beautiful, rather than becoming beautiful at one time and not beautiful at another; (3) being only and always beautiful, rather than beautiful in one respect or for one observer, but not beautiful in another respect or for another observer. Hence (4) the being of the Form, which is accessible only to thought or understanding, is distinct from its appearance in becoming, which is accessible to opinion (*doxa*) and sense-perception. (5) Anything else that is beautiful is such only because of its dependence on the Beautiful itself. This ontological dependence is described in terms of participating or sharing in the Form, or imitating the Form by being an image of the Form. (6) Reflecting this dependence is the notion of eponymy: everything called an *F* is named after the *F* itself. (7) The converse of the eponymy relation is the principle of one over many: For every plurality of things called *F*, there is the Form *F* itself.

The relation between Forms and their sensible eponyms is the most obscure feature in Plato's theory. In the *Phaedo*, Socrates insists on the derivation of sensible beauty from the Form, but expresses uncertainty as to how this derivation is to be understood: "Nothing else makes it [the sensible thing] beautiful except the presence or communion or whatever connection there may be with the Beautiful itself—I am not sure about this, but [I am sure] that it is by the Beautiful that all the many beautiful become beautiful" (100d). The terminology of participation occurs once in the *Symposium*, repeatedly in the *Phaedo*, and once again in the *Republic*. But this notion of participation as a Form-sensible relation is subjected to a withering critique in the *Parmenides* (131a–e). In the *Republic* participation is generally replaced by the language of imitation and imaging or copying; and it is this terminology that reappears later in the *Timaeus*.

Difficulties with the classical theory will be discussed in the next section. We consider here the intended scope and motivation of Plato's theory. It is often presented as a solution to the problem of universals. This, however, is not only anachronistic but inaccurate, since the concept of universals (which are not properly *ousiai*, not substances in a strict sense) was introduced by Aristotle precisely as an alternative to Plato's conception of Forms. In the *Republic* the Being of the Forms is introduced on epistemic grounds as the object of knowledge, in contrast to the imperfect reality of the sensible manifold as object of *doxa*. (The deficient reality of the many beautiful things is reflected in the fact that they are beautiful in some respects, not beautiful in other respects. Hence they *are* in some respects, but they *are not* in other respects. The *is* of predication is thus taken to express a reality claim for the subject.) The underlying assumption, often reasserted in later dialogues, is that an object of knowledge must be eternally invariant; otherwise the cognition of it at one time would become false at another time (so explicitly at *Cratylus* 440a). But knowledge must be always true; hence an object of knowledge must be eternally unchangeable.

This is the argument underlying the presentation of Forms as invariant objects of knowledge in *Republic* V. To knowledge strictly understood corresponds Being strictly understood: "what is completely real (*to pantelôs on*) is completely knowable" (*Rep*. 477a 3). Anything less real can be the object only of imperfect cognition and partial truth. Plato hesitates to present this as an argument, however, since it might seem to imply the priority of epistemic considerations. That would be misleading. Epistemology and ontology go hand in hand for Plato, but it is the real that determines what is knowable and not conversely. It is the stability of Being that makes reliable cognition possible.

This Parmenidean insight constitutes the permanent basis for Plato's metaphysical speculation. It is worked out for the first time in the classical theory of Forms, but it persists as well in later dialogues such as the *Philebus* and *Timaeus*. It is in the *Sophist* that we have the most explicit statement that without stability and invariance there can be no knowledge or understanding (*nous*) whatsoever (249b–d). What is distinctive of the classical theory is not the invariance of Being but the one-many and eponymy relations between Forms and sensibles, as expressed at *Rep*. X, 596a: "We are accustomed to posit some one Form concerning each plurality to which we

assign the same name." Thus there will be one Form of Beauty corresponding to all beautiful things, one Form of Good corresponding to all specific goods, and so on. But the passage just quoted from *Republic* X is destined to cause trouble, for many reasons. For example, it suggests that Forms will be as plentiful as common nouns and adjectives. Plato will have to speak more cautiously about cutting nature at the joints (*Phaedrus* 265e) and thus make clear that not every distinction between words will mark a distinction between Forms or Kinds of things (*Statesman* 262b–263b). The scope of the classical theory is originally undefined, but it does seem to be committed to Forms for artifacts as well as for natural kinds. Thus there is a Form for shuttle at *Cratylus* 389b and a Form for bed at *Rep*. 596b, 597a.

Less obvious than the Parmenidean-epistemic motivation for the doctrine of Forms, but equally important, is the distinctively Platonic conception of philosophy as a form of love or *erôs*, the conception expressed in Plato's etymological reading of *philo-sophia* as "the love of wisdom" (see above). This notion of the philosopher as lover with the Forms as the beloved object provides the original context for the introduction of the theory in the *Symposium*, where the Beautiful itself appears as the ultimate object of philosophic passion. In the *Phaedo* the philosopher is said to be ready for death because it is only when liberated from the body that he can hope to obtain the object of his desire, namely, full knowledge of the truth (67e–68b). The Form of Good is the ultimate Form, not only because it is the source of being and knowability for the other Forms, but also because it is "what every soul pursues and for the sake of which it performs all of its actions" (*Rep*. 505d 11). The doctrine of Forms is thus designed, from the beginning, to provide not only an epistemology and an ontology but also a philosophy of life, that is to say, a theoretical basis for ethics and politics. It is in virtue of his or her access to the Forms, and above all to the Form of the Good, that a philosopher-king is uniquely qualified to govern, since only such access enables them to know what is a good life for individuals and for the city.

These powerful practical implications of the theory of Forms reflect its origin in the Socratic conception of philosophy as a form of life and in the Socratic concern with defining the virtues as the mark of a good life. No interpretation of Plato's theory can be adequate unless it takes into account this profoundly practical bent of Plato's conception of philosophy. The unique character of Plato's metaphysics lies in this convergence between the Parmenidean demand for eternally unchanging reality and the Socratic pursuit of what makes a human life worth living. Thus the original focus of the theory is not on the problem of meaning for general words or concepts but specifically on what we may identify as value terms: the noble and beautiful first of all (*to kalon*), the good (*agathon*) and the just (*dikaion*). The first generalization of the theory is to mathematical concepts (the equal, the greater, and the smaller) and then to health and strength and to every term defined in dialectic, that is, to every essence (*ousia*) "on which we put the stamp of what-it-itself-is" (*Phaedo* 65d 12, 75d 1: *auto to ho esti*, the most technical expression for the Forms). How far this generalization of the theory is meant to extend is a question to be raised and partially answered in later dialogues.

PARMENIDES AND THE CHALLENGE TO THE CLASSICAL THEORY

In the dialogue *Parmenides* Plato brings the two Eleatic philosophers, Parmenides and Zeno, to Athens for an imaginary confrontation with Socrates. This is the first of a series of dialogues in which Socrates is no longer the chief speaker, being replaced here by Parmenides. Since Parmenides was almost certainly dead by 450 BCE (the alleged time of the conversation, when Socrates was about twenty), Plato has ignored chronology in order to introduce Parmenides as a masterful critic of the doctrine of Forms. It is no accident that, in the dialogues generally, Parmenides is the only philosopher who is allowed to win an argument with Socrates. Furthermore, in view of the Eleatic inspiration of Plato's own conception of Being, Parmenides can be trusted as a sympathetic critic of the theory. He is the first to recognize that to give up the theory completely would mean to abandon philosophy (135b–c).

The dialogue divides into two parts. Part I present a series of objections to the Forms, objections that are never explicitly answered by Plato either in this dialogue or elsewhere. Part II contains eight rigorous deductions from the hypotheses *That the One Is* and *That the One Is Not*. The conclusions come in contradictory pairs. According to Deduction 1, the One has no properties; according to Deduction 2 the One has all properties, including contraries (e.g., it is both at rest and in motion, both greater than itself and smaller than itself). How the deductions of Part II are related to one another and to the objections in Part I are matters of extreme obscurity. Parmenides introduces these arguments simply as an example of how a philosopher should be trained before attempting to formulate a theory of Forms. These deductions are thus presented as a logical "exercise" (*gymnasia*)

preliminary to philosophy proper. They nevertheless represent the only fully developed examples of formal dialectic in the dialogues.

Interpretation of the eight (or on some counts nine) baffling arguments of Part II has been a subject of controversy since antiquity. Skeptics saw these apparently contradictory deductions as purely destructive, whereas Plotinus identified the first three hypotheses with his three principal hypostases: the One, Nous and Soul (*Ennead* V.1.8). Modern views have emphasized the overlap with the mingling of Kinds in the *Sophist* and other topics discussed in the late dialogues, such as whole and part, rest and motion. Several interpreters have found in Part II Plato's answers to the difficulties raised in Part I. Some commentators assume that all the arguments of Part II are intended as valid; others regard some of the deductions as so obviously fallacious that the detection of fallacy must be intended as an essential part of the training.

Part I begins with a brief statement by Socrates of the classical theory of Forms, presented as a response to Zeno's paradoxes about plurality. Zeno is quoted as showing that, if things are many, they must have incompatible properties, for example they must be both similar and dissimilar. Socrates agrees that such contraries will be true of the sensible many but not of the corresponding Form: Similarity itself will never be dissimilar, and the One itself will never be plural. Socrates' brief statement here of the classical theory is peculiar in two respects: the relation between the many and the corresponding Form is consistently described as participation (*metechein*, *metalambanein*), and the Forms are said at one point to be "separate" (*chôris*) from their participants (129d 7). In responding, Parmenides will seize upon this last point: "And do you divide as separate certain Forms themselves, on the one hand, and as separate on the other hand the things which participate in them? And is there in your view some Similarity itself separate from the similarity that we have?" (130b 1–5). Socrates agrees, and thus accepts a fatal replication of the Forms as immanent properties.

Both features—the reliance on the concept of participation and the distinction between Magnitude itself and the magnitude in us—accurately reflect the formulation of the doctrine in the *Phaedo* (e.g. 102d 7). And both will be exploited by Parmenides in his criticism, where the notion of participation is shown to be incoherent, while the separation between Similarity itself and "the similarity that we have" (or the similarity "in us") leads to a two-world ontology in which our world is structured by immanent forms. In that case the transcendent Forms of Plato's theory become irrelevant and unknowable. This is the conclusion of the last difficulty, which Parmenides describes as the greatest (133b–134e).

As a consequence of Parmenides' criticism, two features of the classical theory as formulated here must be abandoned: namely, participation taken literally as the "sharing" of Forms by sensibles, and the existence of "forms that we have" or "forms in us" separate from the Forms themselves. Among Parmenides' other objections the best known is the so-called Third Man argument, according to which the one-over-many principle of *Republic* X (that for every group of *F*s we posit a Form, the *F*-itself) leads to an infinite regress. The nerve of this argument is the implicit premise that the *F*-itself is *F*; hence if we add the *F*-itself to the first group of *F*s, we get a larger group of *F*s calling for another *F*-itself; and so on indefinitely. Some scholars have claimed that this premise (the so-called self-predication principle, that *F*-itself is *F*) reflects a logical confusion on Plato's part between being a property and having a property. However, the *Sophist* makes clear that Plato remained committed to this principle, and recent interpretations have shown that no fallacy need be entailed. At the same time, the second implicit premiss required for the regress, the so-called Non-identity principle (that for any larger group of *F*s, a new and different *F*-itself is needed), has no deep Platonic motivation, and its role in generating the regress can be blocked in several different ways. More problematic than the Third Man argument is the parallel objection against the conception of Forms as models (*paradeigmata*), where the dependent relation of participation is understood in terms of similarity or being a likeness of the Form (132d–133a). This objection seems to attack the central concept of imaging or imitation, which replaces participation in the doctrine as reformulated in the *Republic* and *Timaeus*. How much of the classical theory of Forms can be thought to survive the critique of the *Parmenides* will depend in part on the interpretation of this model-copy relation as developed in the *Timaeus*.

THEAETETUS AND *SOPHIST*: SURVIVAL OF THE FORMS? THE LATER DIALECTIC

The *Theaetetus* and *Sophist* stand in the shadow of the *Parmenides*: both dialogues refer to the conversation between Socrates and Parmenides as if it were a historical event (*Theaet*. 183e 7; *Soph*. 217c 5). As a consequence, both dialogues distance themselves from the classical theory of Forms. Neither dialogue denies the existence of

Forms, and both refer to concepts or entities that recall Forms. But neither dialogue asserts the metaphysical dualism of the classical theory. The *Sophist* even subjects this theory to a new round of criticism. It is as if Plato in the *Parmenides* had wiped the slate clean, and was prepared to make a fresh start in the later dialogues in addressing the basic issues of epistemology and metaphysics.

The *Theaetetus* is almost the last dialogue in which Socrates appears as the chief speaker (only the *Philebus* is later), and the last one in which his elenchtic function is dramatically displayed. In fact the negative character of the elenchus is uniquely underscored here in the comparison of Socrates to a midwife. The official role of Socrates in this dialogue is not to produce theories on his own (as he did in the *Phaedo*, *Republic* and *Phaedrus*, and will again in the *Philebus*) but solely to extract definitions of knowledge from Theaetetus.

Theaetetus's attempts to define knowledge fall into two categories, dividing the dialogue into two unequal parts. The first and longer section corresponds to the initial definition of knowledge as sense perception (*aisthêsis*). This definition is ultimately rejected on the grounds that truth, and therefore knowledge, is not accessible to sense perception as such but only to the rational psychic activity that Theaetetus calls *doxazein*, "having an opinion" (187a). The remainder of the dialogue is then devoted to various accounts of knowledge and error based on this notion of *doxa*, that is, opinion, belief, or judgment. The results of this second section are equally negative, so that the *Theaetetus* has the external form of an aporetic dialogue like the *Laches*or *Meno*—an unsuccessful attempt to define knowledge. The philosophical content of the *Theaetetus* is, however, extremely productive in arguments and insights for epistemology and philosophy of mind. Why then is the outcome so negative?

If we relate this discussion to Plato's theory of knowledge as formulated in the *Republic*, we can see why the enterprise of the *Theaetetus* was doomed to fail. According to the view of *Republic* V–VI (reasserted in the *Timaeus*), sense perception and opinion (*doxa*) take as their object the realm of sensory Becoming, whereas knowledge proper takes as its object only invariant Being. Thus in the Divided Line of *Republic* VI, both sense perception and *doxa* belong to the lower sections of the line, devoted to the visible realm, but knowledge belongs at the top with the Forms as its object. In the *Republic* and *Timaeus* this view of knowledge as metaphysically grounded is presented as a basic assumption, without detailed supporting argument. In the *Theaetetus*, in contrast, all attempts to define knowledge avoid any recourse to Parmenidean ontology or to the classical doctrine of Forms. This systematic departure from Plato's classical epistemology can be seen as an application of the method proposed and exemplified by Parmenides in the dialogue named after him: See what follows not only from your own assumption but also from its denial (136a 1). Accordingly, in the *Theaetetus* we pursue an acount of knowledge from the opposing, non-Platonic point of view. Let us assume that knowledge can be defined either on the basis of sense perception, or on the basis of *doxa*, and see what follows from either assumption. The *Theaetetus* thus has the form of a double *reductio*. Since neither alternative gives a satisfactory result, we are justified in returning to our original point of view. There is still no explicit argument for the Parmenidean postulate (that knowledge in the full sense takes as its object Being in the full sense). But this assumption is supported indirectly, by the failure of the alternative attempt in the *Theaetetus* to give an account of knowledge that avoids this postulate.

Although the general form of the *Theaetetus* is thus negative, the positive content is extremely rich. The first section develops a subtle theory of subjective perceptual qualities within the framework of Protagorean relativism, on the basis of a neo-Heraclitean doctrine of flux. Commentators disagree on whether this theory of perception should be read as merely hypothetical or whether it in fact represents Plato's own view of the subject. A decision must depend upon whether or not the *Theaetetus* account of perception is compatible both with Plato's own version of cosmic flux in the Timaeus and also with his mechanistic account of sense qualities in that dialogue. Of great interest also is the argument known as the *peritropê*, or "overturning," according to which Protagorean relativism is shown to be self-refuting; since it could be true at most for those who believe it, but false for everyone else, therefore even those who believe it must admit its falsity for the others, that is, for most people (*Theaet*. 170a–171c).

The final rejection of sense perception as a candidate for knowledge relies upon a new distinction between sense-perception proper, that is, information derived through the sense organs of the body, and "common thoughts" (*koina*) like "same" and "different," "one" and "many," that apply to more than one sense modality. The argument concludes that the *being* of predication and existence, and hence of truth, is not available to sense perception as such. "But if one fails to grasp the truth of

something, one cannot have knowledge of that thing" (186c 9). Hence sense perception cannot be knowledge.

The "common thoughts" or concepts (*koina*) introduced by this argument include "beautiful" and "ugly," "good" and "bad," as well as "same" and "different," "similar" and "dissimilar" (185a–186a). As non-sensible notions, these *koina* are clearly suggestive of Forms, but nothing whatsoever is said about their ontological status. There is a closer hint of the classical theory in the famous moral digression of the *Theaetetus* (where virtue is defined as *homoiôsis theôi*, "becoming like god" 176b 1): There resemblance at the human level is said to connect us with transcendent models (*paradeigmata*) of justice and injustice "established in reality" (*en tôi onti hestôta* 176e 3). These two paradigms represent two lives, one of which, as a model of injustice, is "godless and most wretched." The context of the digression clearly invokes both the judgment myths of *Phaedo* and *Republic* and the moral spirit of the *Gorgias*; but there is no unambiguous reference here to Forms as defined in the classical theory.

In the *Sophist*, Plato returns to questions of ontology with a vengeance. The central theme of the dialogue is the problem of Not-Being, and it is argued that the concept of Being is equally problematic, so that the two concepts must be clarified together. Accordingly, the dialogue surveys a series of metaphysical positions, including both Parmenidean monism and a materialist view that reduces Being to bodily existence. A clearly recognizable version of Plato's classical theory is discussed as the doctrine of "the friends of the Forms." As in the *Parmenides*, a sympathetic critique is guaranteed here by the presence of a metaphysically oriented philosopher as protagonist. As a pupil of Parmenides, this "visitor from Elea" can subject both Parmenides' account of Being and Plato's own theory to constructive criticism. In particular, the Stranger's critique of the Friends of Forms shows that the classical theory must expand its ontology to make room for motion and change as a kind of Being. How this is to be done is left for discussion elsewhere, presumably in the *Timaeus*. The *Timaeus* also pursues the most puzzling suggestion of the Stranger's critique, namely that there must be a place among the Forms for Intelligence (*nous*) and hence for life and soul (*Sophist* 248e–249d).

The doctrine of Forms reappears in the constructive argument of the *Sophist* as a theory of Kinds (*genê*) that are capable of combination or participation with one another; dialectic is accordingly redefined as the science of "dividing according to Kinds," knowing "which Kinds harmonize with which, and which do not admit one another" (253b–d). Although in this dialogue we set out to define the Sophist, we seem to have found the philosopher instead, since this dialectical art belongs only "to one who purely and rightly philosophizes" (253e). The description of the philosopher appeals here to the visual imagery of the classical theory: The philosopher is said to be so hard to see because of the brightness of the region "where he is attached always in reasoning to the form (*idea*) of Being; for the eyes of the soul of most people cannot bear for long the sight of what is divine" (254a). The metaphysical discussion is, however, left incomplete. The Eleatic Stranger speaks of participation only between Forms or Kinds; nothing is said of the relation between Forms and their sensible eponyms.

Instead of metaphysics the new theory of participation between Kinds offers something like transcendental logic. "It is through the weaving-together of Forms (*eidê*) with one another that rational discourse (*logos*) has been given to us" (259e). The most elementary weaving-together (*symplokê*) is between noun and verb to form the basic *logos* of a sentence or statement (262c 6). Plato thus introduces the subject-predicate analysis of sentence structure that served as the basis for Aristotle's own theory of predication. Exactly how this analysis is applied in the detailed account of Not-Being is a matter of dispute, but it is clear that the Form of Not-Being is explained by reference to two other Forms, Being and Otherness. (In effect, negation is analyzed in terms of non-identity.) The *Sophist* thus opens up an entirely new dimension in the theory of Forms: a network of logical and semantic relations between concepts or Kinds, such as whole-part or logical inclusion, combination or extensional overlap, and mutual exclusion.

This conception of dialectic as "dividing according to Kinds" is reflected in the method of Collection and Division that was described in the *Phaedrus* (265d–266c) and is systematically applied here in both dialogues, in successive definitions of the *Sophist* and the *Statesman*. As was noted above, in these definitions the terms *genos* (kind) and *eidos* (form) seem to be used in their logical sense simply as "genus" and "species," and the ontology of the Forms is apparently left indeterminate. At the same time, the Eleatic Stranger speaks more definitely of "incorporeal beings, the greatest and finest," which have no images adapted to sense perception but can be clearly indicated only by rational discourse (*logos*); it is for the sake of these beings that the dialectical definitions are pursued (*Statesman* 285e–286a). In such a passage, as in the reference to the divine *idea* of Being at *Sophist* 254a, there is a clear reminder of the classical theory. But nothing is said in either the *Sophist* or *Statesman* to indicate how the dual-

ism of the *Phaedo* and *Republic* is to be altered or preserved.

PHILEBUS AND THE RETURN OF SOCRATES

In the *Philebus* the problems of ontological dualism and participation are directly confronted for the first time since the *Parmenides*. These issues are presented here within the broader context of relations between the One and the Many. As in the *Parmenides*, the problem of participation is distinguished from superficial or eristic ways of being at the same time one and many (as one subject with many properties, or one whole with many parts). The serious problem arises only when we distinguish unities that do not belong to "what comes to be and perishes" but are truly beings and truly unities, like the one Beautiful and the one Good. (Among the examples of ungenerated and imperishable unities listed at *Philebus* 15a are One Human Being and One Ox, thus providing a partial answer to the Population Problem of *Parmenides* 130c 1. Another partial answer is given in the discussion of Forms of Fire and other elements at *Timaeus* 51b–52a.) The question then is how such unities, "admitting neither generation nor corruption, can remain one and the same while coming to be in many and infinite cases of becoming, either one unity being scattered and becoming many, or (most impossible of all) being separate from itself as a whole" (15a–b, recalling the critique of participation at *Parmenides* 131a–c).

A full discussion of these metaphysical issues is avoided in the *Philebus*, however, because of pressure from the prior question whether pleasure or knowledge is the good and the cause of a good human life. The relation of eternal Forms or Monads to sensible becoming is reformulated here in the light of "an immortal and unaging attribute (*pathos*) of discourse (*logoi*)," an attribute rather cryptically identified as the claim that "the identity of one and many generated by discourse (*logoi*) circulates in every way among everything that is ever said" (15d). As the best way out of this confusion, the dialectical method of collecting unities and distinguishing pluralities is presented as a gift from the gods and the basis for all art or science (*technê*, 16c 2). The discussion thus shifts from the problems of ontological dualism to the dialectical project of discerning unity and plurality in the various kinds of pleasure and knowledge. Instead of metaphysics we are given the method of Division, based on the principle (tossed down from heaven by some Prometheus) that "things that are said to be in every case (or "things said to be forever," *ta aei legomena einai*) are derived from one and many, and hence have Limit and Unlimited in their nature" (16c).

These principles of Limit and the Unlimited, introduced here by Plato for the first time, are apparently borrowed from the Pythagorean philosopher Philolaus, who claimed that "Nature in the world-order has been fitted together from unlimited [constituents] and from limiting ones, both the world-order as a whole and everything within it" (fragment 1). In the *Philebus* these two principles provide the basis for a fourfold cosmic scheme that includes several ideas figuring also in the cosmology of the *Timaeus*. "All the beings that are now present in the universe" are analyzed as a blended Mixture of Limit and Unlimited, under the causal influence of Intelligence (*nous*). In this scheme, as in the *Timaeus*, causality is interpreted as the purposeful act of a maker, or *dêmiourgos*. Also common to the *Timaeus* is the introduction of a world soul (*Philebus* 30a–d). But the *Philebus* principles of Limit and Unlimited do not correspond exactly to anything in the *Timaeus*; they figure here as immanent components of Becoming, entering as ingredients into a Mixture that represents both cosmic order and a good human life (23b–27c).

This fourfold scheme of Unlimited, Limit, Mixture, and rational Cause is said to be required in order to decide the contest between pleasure and knowledge for recognition as the good. It has already been settled that neither candidate deserves first place; pleasure and knowledge are each shown to be less choiceworthy alone than the Mixed Life that contains both (20d–22c). The issue for the rest of the dialogue is to assign second place in the competition for the good or, more precisely, to determine the relative position of knowledge and pleasure in accounting for the goodness of the good life. It will turn out that, in the ranking of ingredients in the final Mixture, forms of knowledge occupy third and fourth place, while a selected group of pleasures comes in only fifth. The first two constituents of the good life are principles first of measure (*metron*, *metrion*) and next of beauty and proportion (*kalon*, *symmetron*). The fourfold scheme permits Socrates to identify pleasure as a part or species of the Unlimited, while knowledge and intelligence (*nous*) belong to the genus of the Cause of successful mixtures.

The central section of the *Philebus* is a classification of different kinds of pleasure and knowledge. Socrates proceeds to give a subtle analysis of a number of kinds of pleasure, both mental and physical, in order to distinguish pleasures that are true and pure from various mixed and false pleasures. Only pure pleasures of sense and

intellect will be admitted into the final construction of the good life. Although all forms of knowledge will be admitted, a ranking is nevertheless carried out between different forms of expertise, in a new version of the Knowledge Line of *Republic* V. The lowest division is between various manual crafts, including music; for such arts the level of cognition depends upon the extent and precision of the mathematical component. Mathematics in turn is divided into two, with philosophical mathematics representing a higher standard of precision. (Pure mathematics here recalls, but does not exactly correspond to, the higher form of measurement based upon "due measure" in the *Statesman*. The concept of due measure, to *metrion*, does, however, return to define the first constituent of the Good Life in the final ranking of the *Philebus*.)

Finally, the highest form of knowledge is identified as dialectic, which ranks above natural philosophy and cosmology on ontological grounds familiar from the classical theory. For only dialectic is concerned with what is "really real," with Being that is eternal and unchanging; whereas the science of the natural world is a study of what has comes to be and perishes (*Philebus* 59a, 61e; this is the same ontological contrast that will serve as foundation for the cosmology of the *Timaeus* 28a–b). Dialectic is here described in terms of classical dualism, including the epistemic contrast with *doxai* (at 59a 1). But the reader is inevitably reminded of the quite different account of dialectic given earlier in the *Philebus*, where there is reference not to Being and Becoming but rather to the recognition of unities and pluralities (16b–17a). The old and the new conceptions of dialectic are thus both presented but left uncombined. A similar ambivalence can be seen in the *Philebus* regarding the problems of metaphysical dualism, which are recognized but not resolved. And in another respect we are left with expectations unfulfilled. Much of the dialogue raises the question of the Good as such and the good-making properties of any mixture, but we reach at the end only "the threshold of the good," in a list of the essential ingredients of a good human life. Any hopes for an account of the Form of the Good are left unsatisfied. It is no wonder that the dialogue ends (67b 11) with the interlocutor reminding Socrates that something has been left out!

We may wonder why Plato brings Socrates back as protagonist in the *Philebus*, after replacing him with an Eleatic Visitor in the *Sophist-Statesman*, and again replacing him with Timaeus and an Athenian Stranger in the other late dialogues. No doubt the role of pleasure in the good life was familiar Socratic terrain. But the presence of Socrates might equally serve as a reminder of the dualism expounded by the same figure in the *Phaedo* and *Republic*, and also of the unresolved problems raised against Socrates' presentation of this doctrine in the *Parmenides*. Although the *Philebus* is not formally aporetic like the *Theaetetus*, it certainly concludes on a note of incompleteness. If there is a Platonic response to the metaphysical problem recalled here, one must look for it elsewhere, perhaps in the *Timaeus*.

TIMAEUS AND THE PLATONIC COSMOS

The *Timaeus* was for many centuries the most influential of all of Plato's works. After the rise of Christianity, it could be regarded as a philosophical exegesis of the creation story in the Book of Genesis. But the profound influence of the *Timaeus*derives from its mathematical conception of nature, which has also attracted modern admirers from Kepler and Galileo to Whitehead and Heisenberg. For students of Plato the *Timaeus* has the special interest of offering Plato's only radical reformulation of the classical theory of Forms. The introduction of a spatial Receptacle, on the one hand, and an intelligent Maker of cosmic order, on the other hand, permits Plato for the first time to give a systematic account of the natural world, while deploying new resources to counter the challenges to the classical theory that were formulated in the *Parmenides*.

In addition to the Receptacle and the Demiurge, Plato's new theory makes use of two other notions developed in the late dialogues: (1) The idea presented in the *Sophist* that the realm of Being must be enlarged to include motion and change is reflected in the theory of mixture in the *Philebus*, where the analysis of phenomenal unities gives rise to the new, paradoxical expressions *genesis eis ousian*, "becoming into being" (26d 8) and *gegenêmenê ousia*, "being that has come to be" (27b 8). Although the *Timaeus* reverts to the classical antithesis between Being and Becoming, the cosmological theory deals in fact almost exclusively with Becoming. (2) Without using the terms "Limit" and "Unlimited" from the *Philebus*, the *Timaeus* presents a comparable analysis of Becoming as the mixed result of an interaction between two principles, represented here allegorically as Reason and Necessity. The victory of the former over the latter is spelled out in the creation narrative as the shaping of the chaotic motions of the Receptacle by the purposeful action of the Demiurge, "structuring [the pre-cosmic elements] with figures (*eidê*) and numbers" (53b 4).

The *Timaeus* thus interprets the cosmic act of the divine Maker in terms of the normative notion of mathematical measure (*to metrion*, *to symmetron*) expounded in the *Statesman* and *Philebus*. Whereas in the epistemology of the *Republic* mathematics points only upward, to raise the mind towards the Forms, in the cosmology of the *Timaeus* (and, by anticipation, in the *Statesman* and *Philebus* as well) the function of mathematics is also directed downward, to impose order on the mixed products of Becoming, on the good human life as on the order of nature.

By the formal device of Timaeus's monologue, Plato has inserted into this dialogue a prose treatise *peri physeôs* in the Pre-Socratic tradition, applying a revised theory of Forms to produce his own account of the nature of things, that is to say, of the world of perceptible order and natural change. One goal of this account must be to avoid the "greatest aporia" of the *Parmenides* by giving an account of the visible cosmos, including human beings, that does not "separate" the phenomenal world from the Forms. Hence, instead of a sensible realm of immanent forms, Timaeus posits as an entity independent from the Forms only the Receptacle, the place where the Forms are imaged. As joint offspring of Forms and Receptacle, the sensible images are like the Mixtures of the *Philebus*, with no existence independent of their two principles. On the one hand, as modifications of the Receptacle their existence is adjectival rather than substantival. On the other hand, they are no more independent or separable from the Forms than the images in a mirror are independent from the objects mirrored. The *Timaeus* is insistent on this fact of double dependence. "Since in the case of an image even that on which it depends does not belong to it, but it is always carried about as an appearance (*phantasma*) of something else [namely, the Form], it is fitting that it come to be in something else [namely, the Receptacle], on pain of being nothing at all" (52c 2–5). This is Plato's strategy for avoiding the fatal separation between Forms and immanent features of the sensible world, conceived as the "forms that we have" or "forms in us." Properly conceived, images exist only as fleeting determinations of the Receptacle under the influence of one or more Forms.

Of course many questions are left open, including the problem of how Plato can avoid the reciprocity of Similarity which in the *Parmenides* (in a version of the Third Man argument at 132d–133a) threatens to undermine the explanatory role of images and likeness. Images are said to be impressed on the Receptacle " in an amazing way, hard to describe" (50c 6). The promise to return to this topic is not fulfilled, unless we take the theory of elementary triangles, introduced at 53b, to be the promised account of how images of the Forms are produced in the Receptacle. It is in any case the geometry of these invisibly small triangles that replaces the atomism of Democritus with a more strictly mathematical theory. And it is the same geometric account that provides the mechanism by which the mathematics of Limit and "due measure" imposes order and goodness on the realm of sensory flux. It would seem that the theory of elementary triangles in the *Timaeus* is the physical expression of the notion of normative mathematics developed in the *Statesman* and *Philebus*.

A famous problem, debated already in Plato's school, is whether the creation story is to be taken literally, as positing a chaotic condition of the Receptacle before the Demiurge goes to work, or whether the myth of creation is to be interpreted as an expository device to distinguish different explanatory factors. A non-literal reading of creation would avoid the apparent incompatibility between the *Timaeus* account of pre-cosmic motions before the creation of the world-soul and the account given in the *Phaedrus* and the *Laws*, where the soul as self-mover is the source of all motion and change. A non-literal reading would also dispense with some vexing problems about the ontological status of the Demiurge and his relation to the Forms. (He is described as *noêtos*, "intelligible" like the Forms at 37a 1.) If we do not have to take creation literally, the Maker simply represents the principle of reason as a causal agency among the Forms.

Some problems will nevertheless remain. Why is the eternal and unchanging model for creation presented as a *panteles zôon*, a "complete living thing," containing within itself as parts or species all the intelligible living beings (30c–31b)? On the one hand, this eternal model is described in terms that clearly identify it with the Forms of the classical theory. (Thus at 39e 8 the model is referred to as to *ho estin zôon*, "the what-living-thing-is." This technical expression for the Forms occurs in no other dialogue later than the *Parmenides*.) On the other hand, nothing in the classical theory prepares us for this conception of the Forms as alive. It is as if Plato in the *Timaeus* chose to respond to the criticism of the Eleatic Stranger in the *Sophist*, who complained that the Friends of Forms conceive the highest Being as neither living nor thinking, "but standing immobile like a pious statue, without intelligence" (249a 1). Since it is a fixed doctrine that intelligence (*nous*) requires a mind or *psychê*, and *psychê* entails life (*Sophist* 249a, *Philebus* 30c 9, *Timaeus* 30b), by bringing the Forms to life the *Timaeus* evades

this criticism. But the reader is left without a clue as to how the life and thought of the Forms are to be understood. Does the divine Intelligence of the Forms possess a divine Psyche of its own, before the creation of the World-Soul? And how is this Intelligence among the Forms related to the divine Psyche established as first source of motion by the argument of *Laws* X? These are some of the many questions that the myth of the Demiurge allows Timaeus to avoid.

See also Aristotle; Platonism and the Platonic Tradition; Socrates.

Bibliography

WORKS BY PLATO

Editions and Translations

The standard edition of the Greek text is by John Burnet, *Platonis Opera*. 5 vols. Oxford: Clarendon Press, 1899–1907. Volume I has appeared in a second edition, edited by E.A. Duke, W. F. Hicken, D. B. Robinson, and J. C. G Strachen, 1995, and a new edition of the *Republic* has appeared as *Platonis Respublica*, edited by S.R. Slings, New York: Oxford University Press, 2003. There are also editions of the Greek text with French translations in the "Les Belles Lettres" series (13 vols. in 25 parts, Paris: Les Belles Lettres, 1920–1956), and with English translations in the Loeb series (12 vols., London and New York: Harvard University Press, 1921–1935). The best English translations are collected in J. M. Cooper, ed., *Plato: Complete Works* (Indianapolis, IN: Hackett, 1997).

Modern Editions and Translations of Individual Dialogues

Adam, James. *The Republic of Plato*. 2 vols. Cambridge, U.K.: Cambridge University Press, 1902.

Burnyeat, Myles. *The Theaetetus of Plato*. Translated by M. J. Levett. Indianapolis, IN: Hackett, 1990.

Cornford, F. M. *Plato and Parmenides*. Translation of the *Parmenides* with commentary. London: Routledge & Kegan Paul, 1939.

Cornford, F. M. *Plato's Cosmology*. Translation of the Timaeus with Commentary (1937). Indianapolis, IN: Hackett,1997.

Cornford, F. M. *Plato's Theory of Knowledge*. Translation of the *Theaetetus* and *Sophist* with commentary. London: Routledge & Kegan Paul, 1935.

Morrow, G. R. *Plato's Epistles*. Indianapolis, IN: Bobbs-Merill, 1962.

Plato. *Gorgias*. Translation and commentary by E. R. Dodds. Oxford: Clarendon Press, 1959.

Plato. *Meno*, edited with translation and commentary by R. W. Sharples. Warminster, U.K.: Aris & Phillips, 1985.

Plato. *Parmenides*. Translated by M. L. Gill and P. Ryan with an introduction by M. L. Gill.Indianapolis, IN: Hackett, 1993.

Plato. *Phaedo*. Translated with commentary by D. Gallop. Oxford: Clarendon Press, 1975.

Plato. *Phaedo*, edited by C. J. Rowe. Cambridge, U.K.: Cambridge University Press, 1993.

Plato. *Phaedrus*, edited with translation and commentary by C. J. Rowe. Warminster, U.K.: Aris & Phillips, 1986.

Plato. *Philebus*. Translated with Introduction by D. Frede. Indianapolis, IN: Hackett, 1993.

Plato. *Statesman*, edited with translation and commentary by C. J. Rowe. Warminster, U.K.: Aris & Phillips, 1995.

Plato. *Symposium*, edited by K. J. Dover. Cambridge, U.K.: Cambridge University Press, 1980.

Plato. *Symposium*, edited with translation and commentary by C. J. Rowe. Warminster, U.K.: Aris & Phillips, 1998.

Plato. *Timaeus*. Translated with introduction by D. J. Zeyl. Indianapolis, IN: Hackett, 2000.

WORKS ABOUT PLATO

General Accounts

Crombie, I. M. *An Examination of Plato's Doctrines*. 2 vols. London: Routledge & Kegan Paul, 1962–1963.

Friedländer, P. *Plato*. Translated by H. Meyerhoff. 3 vols. London: Routledge & Kegan Paul, 1958–1969.

Grote, G. *Plato and the Other Companions of Socrates*. 2nd ed., 3 vols. London: Murray, 1867.

Grube, G. M. A. *Plato's Thought* (1935). Indianapolis, IN: Hackett, 1980.

Guthrie, W. K. C. *A History of Greek Philosophy*. Vols. 3–5. Cambridge, U.K.: Cambrdige University Press, 1969–1978.

Kraut, R., ed. *Cambridge Companion to Plato*. Cambridge, U.K.: Cambridge University Press, 1992.

Robin, Léon. *Platon* Paris: F. Alcan, 1935.

Shorey, Paul. *What Plato Said*. Chicago: University of Chicago Press, 1933.

Taylor, A. E. *Plato: The Man and His Work*. London: Methuen, 1926.

Willamowitz-Moellendorff, Ulrich von. *Platon*. 2 vols. (1919). 3rd ed., vol. I. Berlin: Weidmann, 1948.

Zeller, Eduard. *Die Philosophie der Griechen*. 6th ed. Hildesheim, 1963.

Particular Aspects of Plato's Work

Allen, R. E., ed. *Studies in Plato's Metaphysics*. London: Routledge & Kegan Paul, 1965.

Bambrough, Renford, ed. *New Essays on Plato and Aristotle*. London: Routledge & Kegan Paul, 1965.

Bobonich, C. *Plato's Utopia Recast: His Later Ethics and Politics*. Oxford: Oxford University Press, 2002.

Burnyeat, M. "Plato on Why Mathematics is Good for the Soul." *Proceedings of the British Academy* 103 (2000): 1–81.

Cherniss, H. F. *Aristotle's Criticism of Plato and the Academy*. Baltimore, MD: Johns Hopkins Press, 1944.

Cherniss, H. F. *The Riddle of the Early Academy*. Berkeley: University of California Press, 1945.

Frede, M. "Prädikation und Existenzaussage." *Hypomnemata* 18. Göttingen: Vandenhoeck and Ruprecht, 1967.

Gosling, J. C. B., and C. C. W. Taylor. *The Greeks on Pleasure*. Oxford: Oxford University Press, 1982.

Irwin, T. *Plato's Moral Theory* (1977). Revised as *Plato's Ethics*. Oxford: Oxford University Press, 1995.

Kahn, C. H. *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*. Cambridge, U.K.: Cambrdige University Press, 1996.

Klosko, G. *The Development of Plato's Political Theory*. New York: Methuen, 1986.

Meinwald, C. C. *Plato's Parmenides*. Oxford: Oxford University Press, 1991.

Morrow, G. R. *Plato's Cretan City: A Historical Interpretation of the Laws*. Princeton, NJ: Princeton University Press, 1960.

Patterson, R. *Image and Reality in Plato's Metaphysics*. Indianapolis, IN: Hackett, 1985.

Robinson, Richard. *Plato's Earlier Dialectic*. 2nd ed. Oxford: Clarendon Press, 1953.

Ross, W. D. *Plato's Theory of Ideas*. Oxford: Clarendon Press, 1953.

Santas, G. X. *Plato and Freud: Two Theories of Love*. Oxford: Blackwell, 1988.

Skemp, J. B. *The Theory of Motion in Plato's Later Dialogues*. Cambridge, U.K.: Cambridge University Press, 1942.

Solmsen, Freidrich. *Plato's Theology*. Ithaca, NY: Cornell University Press, 1942.

Stenzel, Julius. *Plato's Method of Dialectic*. Translated from the 2nd German edition by D. J. Allan. Oxford: Clarendon Press, 1940.

Vlastos, G. *Introduction to Plato's Protagoras*. New York: Liberal Arts Press, 1956.

Vlastos, G. *Platonic Studies*. 2nd ed. Princeton, NJ: Princeton University Press, 1981.

Vlastos, G. *Studies in Greek Philosophy*. Vol. 2. Princeton, NJ: Princeton University Press, 1995.

Wedberg, A. E. C. *Plato's Philosophy of Mathematics*. Stockholm: Almqvist & Wiksell, 1955.

On Socrates

Gigon, O. *Sokrates, sein Bild in Dichtung und Geschichte*. Bern: A. Francke, 1947.

Momigliano, A. *The Development of Greek Biography* Cambridge, MA: Harvard University Press, 1971.

Patzer, A., ed. *Der Historische Sokrates*. Darmstadt: Wissenschaftliche Buchgellschaft, 1987.

Vlastos, G. *Socrates, Ironist, and Moral Philosopher*. Cambridge, U.K.: Cambridge University Press, 1991.

On the Early Academy

Dillon, J. *The Heirs of Plato: A Study of the Old Academy*. Oxford: Clarendon Press, 2003.

Ostwald, M., and J. P. Lynch. "Plato's Academy." Cambridge Ancient History VI, 602–616. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1994.

On the Chronology of the Dialogues

Brandwood, L. *The Chronology of Plato's Dialogues*. Cambridge, U.K.: Cambridge University Press, 1990.

Campbell, L. *The Sophistes and Politicus of Plato*. Oxford: Clarendon Press, 1867.

Kahn, C. H. "On Platonic Chronology." *New Perspectives on Plato, Modern and Ancient*, edited by J. Annas and C. Rowe, 93–127. Cambridge, MA: Harvard University Press, 2002.

Ritter, C. *Untersuchungen über Plato*. Stuttgart: n.p., 1888.

Bibliographies

Brisson, Luc, and Benoît Castelnérac. *Platon, 1995–2000: Bibliographie*. Paris: Vrin, 2004.

Brisson, Luc, and Frédéric Plin. *Platon, 1990–1995: Bibliographie*. Paris: Vrin, 1999.

Cherniss, H. F. "Plato Studies, 1950–1957." *Lustrum* 4 (1959): 5–308; 5 (1960): 323–648.

Rosenmeyer, T. G. "Platonic Scholarship, 1945–1955." *Classical Weekly* 50 (1957): 172–182, 185–196, 197–201, 209–211.

Saunders, Trevor, and Luc Brisson. *Bibliography on Plato's Laws*, revised and completed with an additional *Bibliography on the Epinomis* by Luc Brisson. Sankt Augustin, Germany: Academia Verlag, 2000.

Charles H. Kahn (2005)

PLATONISM, MATHEMATICAL

See *Realism and Naturalism, Mathematical*

PLATONISM AND THE PLATONIC TRADITION

The term "Platonism" is so widely used in modern scholarship that it is difficult to determine its meaning precisely as applicable either to a particular group of thinkers or to a specific collection of doctrines. Ancient sources frequently describe "Platonists" as those philosophers who further developed the known or presumed teaching of Plato himself and "Academics" as those who pursued the skeptical methodology believed to have been initiated by the Socrates of Plato's earlier dialogues. However, the substantive "Platonism" seems first to occur in scholarly literature only around the beginning of the eighteenth century when it was used to characterize doctrines that were not only derived from but also combined with Plato's own teaching by later exegetes.

In order to apply this relatively modern usage of the term "Platonism" legitimately to the history of Western philosophy in general, it is useful to distinguish between: (1) Platonism in the sense of a Platonic *tradition*, or a set of ideas that is viewed in a strongly historical sense in connection with Plato or his early exegetes and is sufficiently extensive and coherent to overwhelm any influences from other traditions; and (2) Platonism in the sense of a Platonic *influence*, or a set of ideas that is viewed in a weakly historical sense in connection with Plato or his early exegetes and is not sufficiently extensive or coherent to overwhelm any influences from other traditions. Within the former category, it is useful to distinguish further (a) the *direct* Platonic tradition, that is, various philosophical ideas which we know to form part of the Platonic legacy and which their proponents characterized similarly, and (b) the *indirect* Platonic tradition, that is, those philosophical ideas which we know to form

part of the Platonic legacy but which their advocates characterized differently.

Throughout the ancient period of Western thought, there was a Platonic tradition when Platonic philosophers were either members of Plato's Academy or claimed to revive and continue the "Academy." For this discussion, the medieval period is considered in terms of distinct Byzantine, Arabic, Jewish, and Latin cultural components, and here the distinction between direct and indirect traditions of Platonism becomes important, especially with respect to the Arabic tradition in which a type of indirect Platonism was viewed as "Aristotelianism." During the modern period of Western thought there has been initially a Platonic tradition, when Platonic philosophers again claimed to revive and continue the "Academy," but subsequently only Platonic influence.

Although such a procedure risks oversimplification, it may be useful to introduce the detailed historical analysis with a statement of the "essence of Platonism," that is, the set of philosophical assumptions underlying Plato's own written works or oral teachings in the view of his immediate successors in the Academy. Scholars may perhaps be guided by the ancient summary of Platonism in Apuleius's *On Plato and His Doctrine* (2nd century CE), which can be shown to depend on the early Peripatetics and on the early Academy—both with respect to the individual doctrines attributed to Plato and the pedagogical framework presenting them. According to Apuleius, Plato developed his own philosophical viewpoint after being introduced to the teachings of Heraclitus, studying with Socrates, encountering the Pythagoreans, and absorbing the dialectics of Parmenides and Zeno—the philosophical notions influencing Plato here being obviously those of the world as a continuous flux (Heraclitus), of the pursuit of universal definitions and of the primacy of the moral sphere (Socrates), of number as the underlying reality and of the immortality of the soul (the Pythagoreans), and of the contrast between real being and mere appearance (Parmenides).

Also according to Apuleius, Plato brought philosophy to perfection by combining the physics, ethics, and logic that had been pursued independently by the Pythagoreans, Socrates, and the Eleatics respectively into a single curriculum organized into three parts. On the basis of these historical data, one might therefore summarize the "essence of Platonism" as follows: Platonism is specifically characterized by the establishment of a contrast between the realm of being that is the object of knowledge or reasoning and is not subject to change and the realm of becoming that is the object of opinion or sensation and is liable to change. The two realms are linked by the soul, which exists indestructibly before, during, and after the temporal period of its combination with the body and for which assimilation either to the realm of being or to the realm of becoming represents the primary ethical choice.

ANCIENT PLATONISM

Modern scholars customarily divide Platonism in the ancient world into four main periods by using a mixture of ancient and modern terminology.

The "Old Academy" (347–267 BCE) is what Cicero called the original succession of philosophers within the Academy itself. The first of these philosophers was Speusippus (the scholarch, or "head of the school," 347–339 BCE), whose written works do not survive but whose doctrines can be reconstructed somewhat from later reports. Apparently Speusippus was influenced by the Pythagoreans into advocating as the first principles of reality, the One and the Dyad, the former transcending being, goodness, and intellect and the latter coinciding with matter. Speusippus abandoned Plato's own doctrine that the Forms were Ideal Numbers, yet emphasized Plato's teaching regarding the mathematicals intermediate between intelligibles and sensibles. He also explained the various levels of being as resulting from the relation between the One and different levels of matter.

Whereas Speusippus's theories were not influential until the time of the Neoplatonists, what became the standard type of Old Academic doctrine seems to have originated with his successor Xenocrates (scholarch, 339–314 BCE). Although the latter's works do not survive, it is possible on the basis of later reports to conclude that he produced the official edition of Plato's works and that he began a process of systematizing Platonic thought. For example, he established the formal tripartite division of philosophy into physics, ethics, and logic and he continued to develop the Pythagorean side of Plato's oral teaching. As first principles of reality, Xenocrates opposed the monad conceived as good to the dyad conceived as evil—the former corresponding to a self-thinking intellect containing the Forms or Ideal Numbers—and derived the entire cosmos from their interaction. The higher and lower worlds were mediated by a soul that was defined as a "self-moving number": in other words, it was self-moving like the soul of the *Phaedrus* and mathematically structured like that of the *Timaeus*.

Xenocrates' successor was Polemo (scholarch, 314–267 BCE), who seems to have differed from his two

predecessors in that he placed somewhat greater emphasis on ethics. According to later testimonies, Polemo advocated the view that the goal of human existence was "life according to nature," this principle however required neither the rejection of external goods nor the extirpation of passions. Besides the three scholarchs, the Old Academy included other significant thinkers, including Crantor of Soli, the first known author of formal commentaries on the dialogues of Plato.

The New Academy (267–80 BCE) is distinguished by Cicero from the Old Academy on the basis of its shift from a dogmatic to a skeptical mode of philosophizing. Although this radical change of direction seems to have occurred in reaction to the extreme dogmatism of the current Stoic school, it appealed to the aporetic method illustrated by Socrates in the early dialogues of Plato for its historical justification. Arcesilaus (scholarch, 267–241 BCE), who followed the Socratic practice of writing nothing, argued that the degree of cognitive certitude claimed by the Stoic notions of perspicuity and assent was unattainable and that the correct epistemological attitude to the physical world was one of "withholding assent" (*epochē*). In fact, Arcesilaus did not hold to the position that nothing could be known, but more radically to the viewpoint that one cannot be certain whether anything can be known or not.

Later thinkers in the New Academic tradition slightly modified Arcesilaus's teaching. Carneades (scholarch, c. 160–129 BCE) agreed that it would be possible to reject the Stoic notions of perspicuity and assent while being guided in practical matters by observing three levels of probability. The end of the New Academy seems to have been occasioned by a dispute, the precise details of which are somewhat obscure, between Philo of Larissa (c. 130–69 BCE) and Antiochus of Ascalon (160–80 BCE). According to one reading of the evidence, Philo attempted to reconcile the New Academy and the Old Academy, whereas Antiochus, who was particularly enraged by the interpretation gaining currency that Arcesilaus and Carneades had endorsed the skeptical position publicly while indulging in dogmatic activities in private, preferred to reestablish the Old Academy entirely.

Modern historians call the next phase of ancient Platonism (80 BCE–c. 250 CE) "Middle Platonism." This terminology has been established in order to characterize Platonism in the period between the revival of dogmatism in the Academy by Antiochus of Ascalon and the innovations of doctrine introduced by Plotinus. Although it is applied to a number of philosophers working at different times and in different places, it is perhaps possible to identify certain methods and doctrines as typical of this phase of the tradition. From the viewpoint of methods, the Middle Platonists concentrated on the dogmatic aspects of the tradition—although aporetic and dogmatic elements co-exist in the work of Plutarch of Chaeronea (c. 45–125 CE)—and within the dogmatic approach there is a strong tendency toward systematization. The practice becomes fully established of writing commentaries on Plato's work: Eudorus of Alexandria (fl. c. 25 BCE) is reported to have followed Crantor in commenting on the *Timaeus*—and also of producing handbooks of Platonic teachings—examples of this genre are extant in the form of the *Didaskalikos* of "Alcinous" (fl. c. 130 CE) and *On Plato and his Doctrine* by Apuleius of Madaura (b. c. 125 CE). The tendency toward systematization is accompanied by a tendency toward syncretism. From Aristotelianism, Plutarch can adopt the ethical doctrine of the mean and Alcinous the logical doctrine of the categories. The combination of Pythagoreanism and Platonism implicit in the assumption of monad and dyad as first principles continues with figures like Eudorus, this development being associated with the rise of Platonizing pseudo-Pythagorica around this time (for example, the treatises *On the Soul of the Universe and On Nature* by "Timaeus of Locri" and *On the Nature of the Universe* by Ocellus Lucanus).

From Stoicism, Antiochus of Ascalon can adopt the physical doctrine of active and passive principles and Atticus (fl. c. 170 CE) the ethical doctrine of extirpating passions. From the viewpoint of doctrines, the following physical ideas may be considered as particularly characteristic of Middle Platonism: (1) controversy over the corporeality or the incorporeality of the first principle—here the position of Antiochus should be contrasted very clearly with that of Eudorus and the rest of the tradition; and (2) postulation of a triadic group of first principles consisting of a first God that is One as in Pythagoreanism and Good as in the *Republic* and corresponds to a self-thinking Intellect containing the Forms; a second God having affinities with the Demiurge of the *Timaeus* and the Logos of Stoicism; and a World Soul sharing features with the principle of the same name in the *Timaeus* and the Indefinite Dyad of the Pythagoreans; and (3) tentative emergence of a first principle above Being itself in the work of Numenius of Apamea (fl. c. 150 CE). Among the ethical ideas characteristic of Middle Platonism might be mentioned the debate over the goal of human life. Here, the Antiochean notion of assimilation to nature should be contrasted with the Eudoran ideal of assimilation to God.

The phase in the history of Platonism initiated by the philosophy of Plotinus and in the twenty-first century called "Neoplatonism" may be divided into several "schools," in the sense of being associated with certain leading thinkers: namely, that of Plotinus and his students Porphyry and Amelius, that of Iamblichus and his followers, and the Athenian school from Plutarch of Athens, through Syrianus and Proclus, to Damascius. This last school claimed to be the successor of the ancient Academy and was closed by the Emperor Justinian in 529 CE.

Plotinus (204–269 CE) studied with Ammonius Saccas in Alexandria and later established his own school in Rome. He set out a metaphysical system, which, with various additions and modifications, became foundational for Platonic philosophy and for the reading of Plato until modern times. Thanks to the complete corpus of Plotinian writings called the *Enneads* and the biography attached by Porphyry to his edition of the latter, historians can understand the methods and doctrines of Plotinus more than they can those of any previous Platonist. The *Enneads* reveal precisely how Plato's works yielded systematic metaphysical tenets: The *Republic* provided the notion of the Good above Being; the *Parmenides* provided the postulation of the One, the One-Many, and the One-and-Many as the three first principles; the *Symposium* provided the identification of Beauty and Intellect; the *Sophist* provided the five Kinds constituting Intellect; the *Phaedrus* provided the relation between universal and individual Soul; the *Phaedo* provided the individual soul's attachment and detachment from the body and the notion of virtue as purification; the *Theaetetus* provided the notion of assimilation to the divine; and the *Timaeus* provided the distinction between being and becoming, the notion that the divine has no envy, the treatment of the intelligible living creature as a phase of Intellect, the treatment of the Demiurge as an intellective phase of Soul, the indivisible and divisible components of Soul, the cosmological reading of the lower gods, and the identification of the Receptacle and Matter.

Plotinus's philosophical approach was sometimes based on the interpretation of a specific passage, often quite brief, in Plato's dialogues, sometimes based on the discussion of a particular problem (e.g., that of the relation between Intellect and intelligible objects raised by Porphyry and recorded in *Enneads* V. 5), sometimes based on the critique of some false interpretation of Plato (e.g., that of the evil nature of the visible world maintained by the Gnostics and reported in *Enneads* II. 9) but usually based on a combination of the above. Porphyry's *Life of Plotinus* describes the role of sources other than Plato in these discussions, Aristotle's *Metaphysics* being particularly influential (a statement corroborated by Plotinus's use of the doctrines of potency and act and of the self-thinking intellect), both Platonic and Peripatetic commentators (e.g., Gaius and Alexander respectively) being sources of inspiration, and Stoic doctrines also being utilized (a statement corroborated by Plotinus's dematerialized reading of the *pneuma* as "procession and reversion").

The system emerging from this analysis might perhaps be summarized as follows. According to Plotinus, reality—understood dynamically as a descending hierarchy of "procession" (ontological founding and at certain points ethical fall) and as an ascending hierarchy of "reversion" (ontological completing and at certain points ethical perfecting)—consists of three principles or "hypostases": the One or Good (described less determinately as the Beyond, the Supreme, the First), which is cause or power; Intellect—a macrocosmic unity and microcosmic plurality that timelessly thinks itself is logically distinguishable into the five Platonic Kinds of Being, Sameness, Otherness, Motion, and Rest, and metaphysically contains the Platonic Forms; and Soul—a macrocosmic unity and microcosmic plurality that generates time and receives the Platonic Forms into itself as reason-principles. This hypostasis also contains a higher and a lower aspect: Soul proper and Nature. Below these three principles is the nonprinciple of Matter (in some but not all contexts called Evil), which receives the unfolding of Soul's lower aspect by projecting the Forms into three-dimensional space. The reversion is the more complex of the two dynamic aspects of reality given that it also comprises the epistemological transition from the discursive and propositional reasoning of Soul to the intuitive and nonpropositional thinking of Intellect to that which is approached in an entirely noncognitive manner.

Iamblichus (c. 245–325 CE) presided over an influential philosophical school at Apamea in Syria. He devoted himself to formal commentary on both Plato and Aristotle, a practice in which he followed his teacher Porphyry, and wrote an extensive study of Pythagorean mathematics. His approach to philosophy initiated certain tendencies especially characteristic of later Neoplatonism: namely, increasing emphasis on systematic and religious elements. In the former case, Iamblichus reinforced both the continuity and the discontinuity between the Plotinian hypostases by introducing numerous mediating terms; in the latter, he postulated a more radical fall of the human soul that could only be reversed by ritual

observances. For Iamblichus, the systematic and religious aims came together since the discernment of more levels of reality provided a metaphysical foundation for traditional polytheism.

Proclus (412–485 CE) was the most influential representative of the Athenian school of Neoplatonism. In a number of extant works that include commentaries on Plato's *Alcibiades*, *Cratylus*, *Parmenides*, *Republic*, and *Timaeus*, a commentary on Euclid's *Elements*, and such independent treatises on Platonic philosophy as *The Platonic Theology* and the *Elements of Theology*, Proclus extended the emphasis on systematic and religious aspects of philosophy already detectable in Iamblichus. The systematization was particularly influential. This can be seen in his *Commentary on the Parmenides*, where he interpreted the famous dialectical discussion starting from the hypothetical proposition "If it is (there is a) one, the one will not be many" by applying the first five hypotheses to the One, the "ones" or gods together with the beings participating in them, nondivinized souls, Forms in Matter, and Matter; by associating three senses of "One" (above Being, with Being, and below Being) with the first three hypotheses; and by showing that all the attributes denied of the One in the first hypothesis are affirmed of the gods in the second.

Systematization can also be seen in the *Elements of Theology* where Proclus applied a method reminiscent of Euclidean geometry in order to "demonstrate" through a series of propositions, proofs, and corollaries and starting from certain initial propositions such as "All that is unified is other than the One itself" what philosophers must believe regarding the One itself (propositions 1–6), regarding the relation between the One and the other hypostases of the expanded post-Iamblichean order of being (propositions 7–112), and regarding the other hypostases themselves (propositions 113–211).

MEDIEVAL PLATONISM

The medieval Platonic tradition can be divided into the non-Latin and Latin traditions, the former in its turn being divisible into the Byzantine, Arabic, and Jewish traditions. But before turning to these, a few comments are necessary regarding certain transformations of ancient philosophy by patristic writers that formed a basis for later developments.

The most important intermediary between ancient and medieval Platonism in the West was Augustine of Hippo (354–430 CE). In the autobiographical *Confessions*, Augustine reported his encounter with "certain Platonic books translated from Greek into Latin" (Confessions VII. 9)—assumed to be writings of Plotinus and Porphyry by most scholars—and his consequent liberation from the dualistic and materialistic tenets of Manichaeism that had formerly impeded his progress toward Christian truth. What is being described here in narrative terms is the discovery of that synthesis of Platonism (specifically Neoplatonism) and Christianity that becomes a standard feature of Augustine's writing. This synthesis included two versions of a Platonic theory of first principles: (1) the identification of the Neoplatonic One and Being/Intellect with the Trinitarian Father and Word respectively (as in *Confessions* VII. 9); and (2) the identification of the One and Being/Intellect with God and the angels respectively (as in *On the Literal Interpretation of Genesis* II. 15ff.).

These two versions of Platonism are moving in opposite directions, since in the former case the universal aspect of the second principle is intensified while the hierarchical relation between the first and second principles is weakened; in the latter, the universal aspect of the second principle is weakened while the hierarchical relation between the first and second principles is intensified. The most important intermediary between ancient and medieval Platonism in the East was "Dionysius the Areopagite." This otherwise unknown fifth-century Christian achieved a posthumous authority by writing an important group of theological treatises, including *On the Celestial Hierarchy, On Divine Names* and *On Mystical Theology*, under the pseudonym of the first-century Dionysius famously converted by St. Paul.

On Divine Names in particular provides a skillful Christian adaptation of late pagan Neoplatonism in which the negative and affirmative predicates of hypotheses I and II of Plato's *Parmenides* are applied not to the One and the gods respectively—as in Proclus's commentary—but to a God or "Thearchy"—who is simultaneously transcendent of and immanent in created things. This important transformation in the direction of monotheism has as further philosophical consequences that the distinction between the transcendence and immanence of the deity by being partially mind-dependent introduces an element of idealism into the realist ontology characteristic of traditional Platonism. The Augustinian and Pseudo-Dionysian versions of the Neoplatonic theory of first principles should especially be compared with regard to their handling of the theory of Forms and the doctrine of Soul. With respect to the Forms, both writers understood Forms in the sense of physical paradigms as contained in the divine Intellect but Forms in the sense of moral absolutes as equivalent to

divine attributes. With respect to the Soul, both authors removed the universal Soul from their system but, with suppression of the idea of transmigration between bodies, retained the function of individual souls.

The most important thinker within the Byzantine tradition of medieval Platonism is Michael Psellus (1018–1078). This author's claim to have revived the discipline of philosophy single-handedly is justified to the extent that, in an environment dominated by orthodox Christianity and methodological Aristotelianism, he reestablished the patristic notion of Platonism as a forerunner of Christianity and the later Neoplatonic notion of a relation between Aristotle and Plato in which the former's physics serves as an introduction to the latter's theology. Although Psellus is hardly responsible for metaphysical innovations in works like *On Plato's Psychogony* and *On the Ideas Which Plato Mentions*, the fact that he discussed philosophy by explicitly combining pagan Platonic sources such as Plato, Proclus, and Plotinus with Christian Platonic sources such as Gregory of Nyssa, Pseudo-Dionysius, and Maximus the Confessor represents an innovation in textual practice.

More specifically, this practice might be characterized as selective in that it isolates only certain aspects of traditional Platonism as compatible with Christianity—for example, by removing all theurgic elements (in *On the Activity of Demons*)—as allegorical in interpreting metaphysical principles in pagan texts as symbols of metaphysical principles in Christian scripture, and as combinatory in that it juxtaposes groups of notions drawn from traditional Platonism and Christianity without reducing the conflicting elements—for example, by combining Proclus's metaphysical interpretation of Jupiter's relation to the lower gods with the pseudo-Dionysius's of the Thearchy's relation to the angelic ranks (in *On Homer's Golden Chain*). This highly original textual manipulation of Platonism established an intellectual tradition that endured until the fall of Constantinople. Later representatives include John Italos (c. 1025–after 1082), Eustratios of Nicaea (fl. at the end of the eleventh and early twelfth centuries)—author of a commentary on Aristotle's *Nicomachean Ethics*, which includes material from Proclus and was translated into Latin by Grosseteste, and Nicholas of Methone (mid-twelfth century), author of a "refutation" of Proclus's *Elements of Theology*.

That Arabic writers were able to make a major contribution to the development of medieval Platonism not only in the Islamic but also subsequently in the Christian world resulted from a fortunate circumstance: the availability of some reasonable translations as sources. Under the Umayyad and ʿAbbāsid caliphs, a vast enterprise of translating scientific and philosophical works from Greek into Arabic (sometimes through the intermediary of Syriac) was undertaken by such figures as Hunain b. Ishāq (808–873) and Qustā b. Lūqā (tenth century) with the result that all of Aristotle except the *Politics*, a certain amount of Plato, and many Greek philosophical commentaries, became available. It was in such a milieu that an important group of philosophical *apocrypha* arose.

This group consists of: (1) an Arabic "Plotinian" corpus (possibly the remains of a translation and commentary on *Enneads* IV–VI produced in the circle of al-Kindī [b. late eighth century and d. after 866]) comprising the *Theology of Aristotle*, the *Letter on Divine Science*, and the *Sayings of a Greek Sage*; (2) the adaptation of Proclus's *Elements of Theology*, later known to the Latins as the *Book of Causes* (the Arabic original of which was produced before 992); and (3) an Arabic translation of approximately thirty-five propositions from Proclus's *Elements of Theology* and *Elements of Physics*. These works are connected through their expression of metaphysical teachings that depart from their Plotinian or Proclean originals in identical ways: namely, in describing the first principle as Pure Being—meaning Being without Form—rather than as the One above Being; and as creating, without any preexistent term or materiate substratum—rather than as causing—all subsequent principles. Moreover, that the first principle or Creative Being does not relate indirectly—through the mediation of an order of gods or "ones"—but directly to the second principle or Created Being is the common doctrine of the *apocrypha*.

During the next few generations, various writers developed a uniquely Arabic approach to the reading of the philosophical tradition in which a Neoplatonic doctrinal component drawn from sources of the type mentioned was inserted into an overtly Aristotelian context. Within this tradition al-Fārābī (Latin: *Alfarabius* [d. 950]) outlined a program of harmonizing Aristotelian and Platonic doctrine in his *Reconciliation of the Two Sages*, *Philosophy of Plato*, *Philosophy of Aristotle*, *Attainment of Happiness*, and *Opinions of the Inhabitants of the Virtuous City*. According to his metaphysical system, the Supreme Being or One produces a series of intellects, each of which can think its cause (thereby giving rise to a further intellect) and itself (thereby giving rise to a celestial sphere), this theory being understandable as the transfer of the emanative causal mechanism from the Neoplatonic hypostases to the Aristotelian unmoved movers.

Also within this tradition Ibn Sinā (Latin: *Avicenna* [980–1037]) organized knowledge into logic, physics, and metaphysics along Aristotelian lines in his encyclopedic *Book of Healing* and its abridgement the *Book of Salvation*. He further developed al-Fārābī's metaphysical system in proposing that, when the Supreme Being produces the subsequent terms, the first intellect in a threefold process first thinks the Being necessary in itself (thereby producing by emanation the second intellect), then itself as necessarily existing through its cause (thereby producing the soul of the first heaven), and finally itself as contingently existing through itself (thereby producing the body of the first heaven), this process being repeated until all the intellects, souls, and heavens have been generated. The inevitable reaction came when Ibn Rushd (Latin: *Averroes* [1126–1198]) attempted to liberate the authentic Aristotle from such Neoplatonizing tendencies. Of his two most famous interpretative innovations, the doctrines that the intellects are not connected by emanation and that there is a single agent and materiate intellect for all humanity, the first but not the second obviously runs counter to Neoplatonism. Ethical and political thought was not neglected by the Arabs and, since both Plato's *Laws* and *Republic* were available in Arabic, in this area they tended to be more Platonic than Aristotelian. Among examples of their work are al-Fārābī's compendium of the former dialogue and Ibn Rushd's commentary on the latter.

The most important thinker within the Jewish tradition of medieval Platonism is Solomon ben Judah ibn Gabirol (Latin: Avicebron [c. 1021–1058]). As the author of some excellent poetry in Hebrew, including the famous *Kingly Crown* and one philosophical treatise in Arabic, Ibn Gabirol stands within two cultural traditions. The philosophical work, which survives only in the Latin translation by Iohannes Hispanus and Domenicus Gundissalinus under the title of *Fountain of Life*, continues the speculative approach of the Arabic *apocrypha* but also develops the latter in an original style. Ibn Gabirol argued that the duality of form and matter underlies both the spiritual and the corporeal levels of reality, this combination of the formal and the material being used in subtle ways to explain the relation between unity and plurality. Although form and matter are also two closed doors between the human intellect and its Creator that are difficult to pass through, one can describe the Creator as Wisdom, Unity, and Will inasmuch as he is the cause of form and as Being inasmuch as he is the cause of matter.

Ibn Gabirol's duality of form and matter in created things represents the moments of determination and undetermined within an emanation, as indicated by his references elsewhere to the dynamic process whereby the inferior comes forth from and strives for union with the superior. Although a Hebrew translation of certain extracts was subsequently made by Ibn Falqera and there may have been some influence on the Jewish mystics of the Gerona circle, the philosophical afterlife of the *Fountain of Life* was mainly in the world of Latin scholasticism.

Because only the *Timaeus* was available in Latin (translated up to 53c with commentary by Calcidius [fourth century CE]) throughout the Middle Ages, the translations of the *Meno* and *Phaedo* (by Aristippus of Catania [d. 1166]) and of the *Parmenides* (included in Proclus's commentary up to 142a translated by William of Moerbeke [c. 1215–1286]) achieving only limited circulation towards the end of the period, one refers to a predominantly indirect transmission of doctrine in speaking of a "Platonic" tradition in the medieval Latin world. However, even this restricted definition of the latter is problematic given that the doctrines concerned are usually combined with Christianity and, during the thirteenth to fifteenth centuries especially, with Aristotelianism. One way of approaching the medieval Latin tradition of Platonism is perhaps to distinguish certain doctrinal clusters; that is, groups of philosophical teachings that exhibit sufficient coherence among themselves and predominate sufficiently in the context where they occur, and then to track the evolution of these clusters through medieval thought. The most important clusters are the following:

(1) A "Timaean" cluster. This group of doctrines, which is presented in passages of Augustine's *Against the Academics*, *On the City of God*, and *On Eighty-Three Different Questions* (qu. 46), based on Cicero's works (including his partial translation of the *Timaeus*) and also in Calcidius's *Commentary on the Timaeus* and Macrobius's *Commentary on the Dream of Scipio*, represents a systematic and cosmological Platonism. It emphasizes the metaphysical principles of Soul and Nature, interprets the transcendent Forms as thoughts in the divine mind, and in general has affinities with the Middle Platonic doctrine of Antiochus of Ascalon.

(2) A psychological and Augustinian cluster. Based on Augustine's *Soliloquies* and *On the Trinity*, this group identifies the relations between the macrocosmic and microcosmic aspects of the Neoplatonic hypostases of the One and Intellect respectively along Porphyrian lines in order to ground human cognition or rather supracognition of the First Principle.

(3) A mathematical cluster. Based on the ancient tradition dating back to Xenocrates, it is transmitted to the Middle Ages by Boethius's *On Arithmetic*, *On Music*, and *On the Consolation of Philosophy*. This group emphasizes the relations between the monadic and dyadic aspects of the hypostases and between monad, dyad, and number series.

(4) A "Proclean" cluster. This group of doctrines, which is presented in different ways by Latin translations of the Pseudo-Dionysian corpus (by John Scottus Eriugena [d. c. 877–879] and several later writers), of the Arabic *Book of Causes* (by Gerard of Cremona between 1160 and 1187), and of Proclus's *Elements of Physics* (by an unknown translator c. 1160), *Elements of Theology*, *Commentary on the Timaeus*, *Commentary on the Parmenides*, and *Minor Theological Tractates* (all by William of Moerbeke between 1268 and 1286), represents a systematic and theological Platonism. It emphasizes the metaphysical principles of the One and Intellect, interprets the transcendent Forms as divine attributes or names, and in general is aligned with the Neoplatonic doctrines of the Athenian School.

(5) A psychological and Avicennian cluster. Based on the Latin translation of the psychological portions of Ibn Sīnā's *Book of Healing* (probably by Ibn Daud [d. c. 1180] and Dominicus Gundissalinus [fl. 1126–1150]), this group equates the relation between the macrocosmic and microcosmic aspects of the Neoplatonic hypostasis of Intellect with the conjunction between the separate agent intellect and the human intellect used by Arabic Aristotelianism to combine the abstraction of universals with the emanation of Forms.

In the medieval Latin world, these clusters occur in the following combinations and sequence. In John Scottus Eriugena's *On Natures*, cluster 4 as it occurs in the pseudo-Dionysian corpus is developed into a comprehensive metaphysical doctrine in which everything that is and is not can be divided on the one hand into the four quasi-species of "creating and not created," "creating and created," "not creating and created," and "not creating and not created" and on the other into a procession and a reversion of the First Cause with respect to its effects and of the effects with respect to their First Cause.

In philosophers of the twelfth century there was a tendency to combine clusters 1 and 3. For example, Adelard of Bath (fl. c. 1110–1125), who also translated the writings of Euclid from Arabic, elaborated within the context of cluster 1 a view of nature and of reason as theologically quasi-independent and also a theory of universals designed to harmonize the opinions of Plato and Aristotle; see his *On the Same and the Different and Natural Questions*. William of Conches (d. c. 1154), in his *Glosses on Plato's Timaeus*, *Glosses on Macrobius*, and *Philosophy of the World*, wrote extensively on an issue central to a naturalistic cosmology but problematic for Christian theology: namely the status and function of the Platonic world soul. Thierry of Chartres (fl. 1121–1148), who was described by contemporaries as the greatest Platonist of his era, elaborated within the contexts of cluster 1 and cluster 3 a metaphysics in which the interaction between God's unity and Matter produces the multiplicity of Forms equivalent to numbers, the Trinitarian nature of God also being expressible arithmetically as $1 \times 1 = 1$; see his *Commentary on Boethius's On the Trinity* and *On the Works of the Six Days*.

With the appearance of translations from Arabic into Latin and the rise of the medieval university after circa 1200, Platonism had to compete with Aristotelianism: a task that it accomplished most successfully within the sphere of what modern scholars term "Latin Avicennism." In the anonymous *Book of Avicenna on the First and Second Causes and the Emanation of Being*, cluster 2, cluster 4 as it occurs in the Pseudo-Dionysius and in the *Book of Causes*, and cluster 5 are combined to produce a metaphysical system in which the procession and reversion of effects with respect to the First Cause begins with the production of the first created intellect by Pure Being, and in which cognition takes place when the human soul ascends from the looking of reason to the vision of intellect and the tenth created intellect or agent intellect combines with the human intellect. This text represents a kind of standard late medieval Platonism to which all serious thinkers will have to react whether they are predominantly Aristotelian (e.g., Albert the Great [c. 1200–1280]) or predominantly Platonic (e.g., Dietrich of Freiberg [c. 1240–1318/1320], Meister Eckhart [c. 1260–1327]) in tendency.

MODERN PLATONISM

The modern Platonic tradition can be divided into a phase beginning with the impact of the early-fifteenth-century humanist movement and a phase beginning with Friedrich Daniel Ernst Schleiermacher's German translation of Plato (published 1804–1809). The former phase might also be termed the "early modern" or "Renaissance" phase of Platonism.

Humanism can be defined as an ideal of liberal education based on the study of grammar, rhetoric, poetry, history, and moral philosophy especially through the recovery of authoritative texts in Greek and Latin, the term "humanism" itself corresponding to the *studia humanitatis* advocated by the Roman rhetorician and philosopher Cicero. Although the beginnings of the humanistic movement can be detected in Northern France during the early twelfth century, the main development is usually traced from Francesco Petrarca (in English, "Petrarch," 1304–1374).

Taking their cue from the latter's pointed praising of Plato in preference to Aristotle, Italian humanists together with their Byzantine associates produced during the next century and a half a series of Plato translations based on newly imported manuscripts. These included Latin versions of the *Republic* by Manuel Chrysoloras and Uberto Decembrio, by Pier Candido Decembrio, and by Antonio Cassarino, versions of the *Phaedo*, *Gorgias*, several *Letters*, *Phaedrus* (partial), *Crito*, *Apology*, *Symposium* (partial) by Leonardo Bruni, of the *Axiochus* by Cencio de' Rustici, of the *Ion* by Lorenzo Lippi, of the *Crito*, *Axiochus*, and *Euthyphro* by Rinuccio Aretino, of several *Letters* and the *Euthyphro* by Francesco Filelfo, and of the *Charmides* (partial) by Angelo Poliziano. From this list of titles, one may conclude that the "humanists" interest in Plato was primarily focused on the literary, ethical, and political aspects of Plato's work.

The first Platonic philosopher affected by humanism was Nicholas of Cusa (originally Niklaus Krebs, 1401–1464), a fact indicated by his commissioning of a Latin translation of the *Parmenides* by the Byzantine émigré George of Trebizond, the manuscript of which exists, together with his own marginal notes (Volaterranus 6201, f. 61r–86v), in the twenty-first century. Although Nicholas was not familiar with the complete *Parmenides* when he wrote his most celebrated philosophical work *On Learned Ignorance* (1440)—and probably not even with the part reproduced in Moerbeke's Latin translation of Proclus's commentary—the teaching of the dialogue fitted naturally into the philosophical system already developed for that work on the basis of medieval sources.

In summary, that system involves the threefold distinction of an "absolute maximum" (God), a "contracted maximum" (the Universe), and a simultaneously absolute and contracted maximum (Christ). With respect to the absolute maximum (and also to the relation between the three maxima), one can discern a Pythagorean and Trinitarian metaphysical structure comprising unity, equality, and connection, which is applied to a Dionysian structure based on the contrast of negative theology (indicating divine transcendence) and affirmative theology (indicating divine immanence) in order to produce an original Cusan metaphysical structure consisting of what surpasses opposites, opposites as such, and the "coincidence of opposites." Although this system has many affinities with doctrines advocated during the Middle Ages, it is innovative in emphasizing the subjectivity of the negative-affirmative theological antithesis (by transferring the teaching of pseudo-Dionysius's *On Mystical Theology* to the cosmological sphere), in its frequent recourse to mathematical images: for example, the maximum, infinity, the circles and triangles of geometry, and the concord of music, and in emphasizing the coincidence inherent in opposition (again adapting the teaching of *On Mystical Theology* to a cosmological use), the combination of the first and last points epitomizing the "learned ignorance," which provided Nicholas with his title.

Marsilio Ficino (1433–1499) is a truly seminal figure who established a pattern of interpreting Platonic philosophy that remained fundamental for the next three centuries. By the late 1450s Ficino had acquired a sufficient reputation as a Platonic thinker and as a Greek scholar to be requested by the Florentine ruler Cosimo de' Medici in 1462 to translate Plato's complete works into Latin from a newly acquired manuscript, this translation appearing in a first edition in 1484 and a second edition in 1491. In addition to this commission, Ficino translated the Hermetic corpus (under the title *Pimander*), the *Enneads* of Plotinus (published in 1492), and various treatises by Porphyry, Iamblichus, Proclus, Synesius, and Michael Psellus for the first time, and also made a fresh translation of the pseudo-Dionysian corpus.

Historians rate him highly not only as an exegete of Plato and Platonism (on the basis of his translations and the commentaries published with the latter) but also as a constructive Platonic thinker (on the basis of his substantial independent treatise titled *The Platonic Theology or On the Immortality of the Soul*). Ficino is important as an exegete because he considered for the first time since antiquity the complete writings of Plato and was therefore able to draw material from dialogues unavailable during the Middle Ages and to engage more fully with the argumentative context of Plato's teaching. Moreover, he proposed a special interpretation of the history of philosophy under the influence of late ancient and Byzantine writers and of Diogenes Laertius's *Lives of the Philosophers* according to which the Christian revelation beginning from Moses is confirmed by a unified and

harmonious system of pagan theology emerging as a six-fold transmission linking Hermes Trismegistus with Orpheus, Aglaophemus, Pythagoras, Philolaus, and Plato.

Dionysius the Areopagite plays a pivotal role in this theory, which basically unifies disparate ideas through allegorical reading. As a thinker of unique inspiration and apostolic authority, Dionysius disclosed the truths concealed in the ancient system to pagan Platonists like Plotinus, the latter in turn transmitting those truths back to Christian writers such as Augustine. Ficino is important as a constructive thinker because he developed the Neoplatonism not only of Proclus (which had become known toward the end of the medieval period) but also of Plotinus (which was almost totally unknown during the Middle Ages) in directions more consistent with the Christian sense of human dignity and individuality. For example, the hypostatic system is sometimes recast so that Soul, instead of being simply the lowest of the three principles of the One, Intellect, and Soul, becomes the third member of a series of five terms God, Angel, Soul, Quality, and Body. This arrangement not only gives Soul a mediating and therefore sustaining function but supplies a novel argument for Soul's immortality in that if Soul were dissoluble then the entire order could likewise suffer dissolution. The hypostatic system is also sometimes modified in that Soul, instead of ascending or descending by identifying with the adjacent term of the series conceived dynamically in the upward direction, ascends or descends by passing through static regions formed by the adjacent terms on both sides.

One work by Ficino was particularly influential both inside and outside philosophy: namely, *On Love* written in 1469 and published in 1484. In fact, it is largely owing to this free commentary on Plato's *Symposium* that Platonism was to become among all doctrines in the history of philosophy the most influential on literature, the visual arts, and music.

Another Platonic philosopher affected by humanism was Giovanni Pico della Mirandola (1463–1494). In his *Conclusiones*—a set of 900 philosophical theses that he would have defended in a public disputation had the Pope not intervened by declaring some of them heretical—Pico attempted to extend the notion of a universal system underlying philosophy by adding the Jewish Kabbalah to the Egypto-Hellenic tradition described by Ficino. On the basis of the number of theses drawn from different schools of philosophy in the more historical first part of the work, it would seem that Proclus (supplying fifty-five theses) and the Kabbalists (supplying forty-seven theses) are the most important influences, the organization of the project itself into a set of propositions recalling Proclus's methodology (as in the *Elements of Theology*) and the ascription of numbers to the propositions reflecting the Kabbalistic approach (900 being the numerical value of the cruciform Hebrew letter *tsade*).

Other writings by Pico also respond to Ficinian ideas. The *Oration* (called *Oration on the Dignity of Man* after Pico's death) and the *Heptaplus* elaborate the notion of Soul as central in the order of reality, while *On Being and Unity* (part of a projected work *On the Concord of Plato and Aristotle*) discusses the question whether among first principles Unity is prior to Being or not. In the latter essay, Pico's conclusion that Unity is not prior to Being according to either Plato or Aristotle required him to argue that hypothesis I of the *Parmenides* forms part of a purely dialectical exercise and thereby to sustain the Porphyrian, Arabic, and Latin rather than the Plotinian version of the Neoplatonic theory of first principles. In both these cases his theories deviate from Ficino's normal view.

The question of the impact of Platonism on the generations after Ficino is an extremely complex one. Despite the reading of Plato's dialogues in Greek courses at the traditional universities, the attempt of Francesco Patrizi (1529–1597) to establish courses on Platonic philosophy at the Universities of Ferrara and Rome, and the rise of numerous Platonically inclined literary "academies" in Italy and France, Platonism never displaced Aristotelianism institutionally. In fact, with respect to the sixteenth century it is necessary to speak of Platonic influence rather than of a Platonic tradition. Platonism during this period is partially a continuation of earlier tendencies. This description would apply to various discussions of Soul, for example when the Lateran Council of 1512–1517 proclaimed the immortality of the human soul as official dogma perhaps under the influence of Ficino, and when Giordano Bruno (1548–1600) incorporated the Plotinian theory of the World Soul into his cosmological speculation. Closely connected with the theory of Soul and disseminated by the various "academies" were the Platonic doctrines of spiritual love (derived from Ficino's reading of the *Symposium*) and of divine madness (derived from Ficino's reading of the *Phaedrus* and *Ion*) whose influence can be detected in Bruno's *Eroici Furori* and Patrizi's *Della Poetica* respectively.

Sixteenth-century Platonism is also partially an adaptation to newer ideas. Here, Platonism was rightly seen as having more in common with the rising mathematical sciences and quantitative thought than did Aristotelianism. Of the two main concepts of traditional

mathematical Platonism, the notion of the *a priori* validity of numbers and of the symbolic power of numbers, Johann Kepler (1571–1630) applied both the first and the second and Galileo Galilei (1564–1642) the first but not the second to the astronomical-physical sphere.

With respect to the seventeenth and eighteenth centuries it is even more necessary to speak of Platonic influence rather than of a Platonic tradition. Platonism during this period is partially a continuation of earlier tendencies. This description applies to the philosophy of inner spirituality advocated by the "Cambridge Platonists" Henry More (1614–1687) and Ralph Cudworth (1617–1688), the last European thinkers to explicitly place themselves within the Platonic tradition. Seventeenth- and eighteenth-century Platonism is also partially an adaptation to newer ideas. Here, the notion of the intellectual love of God in Benedictus de Spinoza (1632–1677), the notion of reality as a system of spiritual monads each of which reflects the entire universe from its own viewpoint in Gottfried Wilhelm Leibniz (1646–1716), and the notion of thought reaching the sphere of things-in-themselves in the precritical thought of Immanuel Kant (1724–1804) are particularly important.

Between the sixteenth and eighteenth centuries there is also a remarkable example of a Platonism that might be considered as standing on the borderline between Platonic tradition and Platonic influence. The treatise *Siris* by George Berkeley (1685–1753) is a recommendation of tar-water as a panacea taking the form of a chain of reflections linking the properties of this liquid first with the chemical and physical phenomena of air and fire respectively and secondly with the spiritual world ascending to God. The main philosophical aims of Berkeley's study, which obviously blends the chain of reflections with the chain of being itself, are to oppose the mechanistic, materialist, and pluralistic view of the universe—well established by his own day—with a spiritual, immaterialist, and unified one, and also to supplement the sensuous immaterialism of his own earlier works—notably the *Treatise concerning the Principles of Human Knowledge*—with a theological idealism. The substantial final section of the treatise achieves its aims by mustering an impressive array of explicitly cited Platonic sources, including Plato's dialogues *Alcibiades I*, *Phaedo*, *Republic*, *Theaetetus*, *Timaeus*, *Parmenides*, and *Letters*, Plotinus's *Enneads*, Proclus's *Platonic Theology*, Bessarion, Ficino (especially his commentary on the *Enneads*), Patrizi, and Cudworth.

On the basis of these authorities, it then argues that the three hypostases of Plotinian Neoplatonism are a reflection of the Christian Trinity. Here, the most important points to emerge are that Unity and Being are mutually convertible, that the placing of the hypostasis of the One before the hypostasis of Intellect or Being does not imply any atheism because there is nevertheless no time at which the One was without Intellect—an argument seemingly unprecedented within the Platonic tradition—and that the purely notional distinction between divine attributes allows the first point to be compatible with the second. The Platonic teachings quoted in *Siris* are clearly not to be taken too literally: Rather, the philosophical maxims of ancient times are proposed, as Berkeley put it, not as principles of logical demonstration but as hints to awaken and exercise the inquiring mind.

Schleiermacher's German translation of the writings of Plato (1804–1809), in which the necessity of distinguishing Plato's own doctrine from the teachings of later "Platonists" and the suggestion that the authentic teaching of the dialogues is superior to the pedantic systematizations of their later admirers was stressed, is rightly seen as the watershed in the history of Platonic interpretation. This new approach had already been gaining ground during the seventeenth and eighteenth centuries, as evidenced by various comments in Leibniz and reference books like J. J. Brucker's *Historia Critica Philosophiae* (1742–1744) and Denis Diderot's *Encyclopédie* (1751–1765). This approach underlines the change from the perception of a unified Platonic tradition to that of more fragmentary Platonic influences. But these changed circumstances present a new set of problems for any interpreter wishing to apply the term "Platonism" henceforth. In short, to what extent is it reasonable to speak of "Platonism" after 1800? A few comments on the "afterlife" of Platonism are perhaps in order.

One should begin by considering what might be termed *modern historical studies* on the question of "Platonism." Of relevance to the historical question are the distinctions intended by modern interpreters when employing the terms "Middle Platonism" (occasionally "Pre-Neoplatonism") and "Neoplatonism" with respect to the ancient tradition (see especially the works of Willy Theiler and Heinrich Dörrie). Although the application of such terminology assumes the principle of distinguishing between Plato's own doctrine and that of later exegetes, it does not exclude the possibility that a particular teaching originates with Plato himself, something that must be ascertained on a case by case basis. Also of

relevance to the historical question is the notion that certain doctrines central to Plato's thought that were taught orally by the master but not included in his dialogues can be identified using the tools of modern criticism (see especially the works of Hans-Joachim Krämer and Konrad Gaiser). The study of such doctrines can yield clarification regarding both the meaning of certain teachings expressed obscurely in the dialogues and the origination of various doctrines associated with Middle or Neoplatonism.

These points represent historiography rather than philosophy in the wake of "Platonism." In order to identify a trajectory of modern philosophical Platonism, one might consider the following three cases in which influential doctrines have been or could be associated with Platonism:

> (1) Georg Wilhelm Friedrich Hegel (1770–1831) in his *Lectures on the History of Philosophy* interpreted the Proclean triad of Being, Life, and Intellect within the intelligible world as corresponding to thought-determinations within the Hegelian Idea. One could tentatively propose this as a case of Platonism in that Hegel was explicitly reading a Platonic text and because his doctrine combines similarities with Platonism (the triadic structure occurs in Neoplatonism) with differences (the Platonic structure is abstractly universal whereas the Hegelian is concretely universal). But Hegel is obviously less a Platonist in either of the aforementioned senses than a creative reader of Platonism.

> (2) Gottlob Frege (1848–1925) in his *Foundations of Arithmetic* postulated purely logical objects, which inhabit a logical realm of the objectively nonreal in contrast with the physical realm of the objectively real and the psychical realm of the subjectively real, and which especially include numbers. Scholars frequently describe his thinking as "logical Platonism" in that, although Frege was not explicitly reading a Platonic text, his doctrine combines similarities with Platonism (the establishment of an *a priori* element) with differences (the Platonic element is an essence whereas the Fregean is a proposition). But scholars label Frege a Platonist in an extremely loose sense, given that what is common to Platonism and Frege does not enter into any recognizably systematic structure of Platonism.

> (3) Martin Heidegger (1889–1976) in his *The Essence of Truth: On Plato's Cave Allegory and Theaetetus* interpreted the Platonic Forms not as what is real as opposed to what is apparent but as the interplay of appearance and concealment. One could again tentatively propose this as a case of Platonism in that Heidegger was explicitly reading a Platonic text and because his doctrine combines similarities with Platonism (the dual structure occurs in Neoplatonism) with differences (the Platonic duality is metaphysical in character whereas the Heideggerian is phenomenological in character). But Heidegger is again less a Platonist in either of the senses distinguished previously than a creative reader of Platonism.

These ideas in Hegel, Frege, and Heidegger are undoubtedly among the more powerful philosophical thoughts since the 1800s. They reveal clearly that, although Platonism declined in significance as a tradition between 1600 and 1800, it has continued to provide a stimulus to philosophical activities of all kinds. There is no reason to think that this will not continue to be the case in the twenty-first century and beyond.

See also Agent Intellect; Alcinous; al-Fārābī; Ancient Scepticism; Antiochus of Ascalon; Arcesilaus; Augustine, St.; Averroes; Avicenna; Carneades; Cudworth, Ralph; Eckhart, Meister; Ficino, Marsilio; Frege, Gottlob; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Ibn Gabirol, Solomon ben Judah; More, Henry; Neoplatonism; Nicholas of Cusa; Numenius of Apamea; Petrarch; Philo of Larissa; Pico Della Mirandola, Count Giovanni; Plato; Plotinus; Plutarch of Chaeronea; Porphyry; Proclus; Pseudo-Dionysius; Pythagoras and Pythagoreanism; Schleiermacher, Friedrich Daniel Ernst; Socrates.

Bibliography

Copenhaver, Brian P., and Charles B. Schmitt. *Renaissance Philosophy*. New York: Oxford University Press, 1992.

Dillon, John M. *The Middle Platonists: A Study of Platonism 80 BC to AD 220*. London: Duckworth, 1977.

Fakhry, Majid. *A History of Islamic Philosophy*. 2nd ed. New York: Columbia University Press, 1983.

Gersh, Stephen. *From Iamblichus to Eriugena: An Investigation of the Prehistory and Evolution of the Pseudo-Dionysian Tradition*. Leiden, Netherlands: E. J. Brill, 1978.

Gersh, Stephen. *Middle Platonism and Neoplatonism: The Latin Tradition*. Notre Dame, IN: University of Notre Dame Press, 1986.

Hankins, James. *Plato in the Italian Renaissance*. Leiden, Netherlands: E. J. Brill, 1990.

Klibansky, Raymond. *The Continuity of the Platonic Tradition during the Middle Ages*. Millwood, NY: Kraus International Publications, 1982.

Kristeller, Paul Oskar. *Eight Philosophers of the Italian Renaissance*. Stanford, CA: Stanford University Press, 1964.

Kristeller, Paul Oskar. *Renaissance Thought and Its Sources*. Edited by Michael Mooney. New York: Columbia University Press, 1979.

Merlan, Philip. *From Platonism to Neoplatonism*. 3rd ed. The Hague: Nijhoff, 1968.

Rosenthal, Franz. *The Classical Heritage in Islam*. Translated by Emile and Jenny Marmorstein. Berkeley: University of California Press, 1975.

Tigerstedt, Eugène N. *The Decline and Fall of the Neoplatonic Interpretation of Plato. An Outline and Some Observations*. Helsinki: Societas Scientiarum Fennica, 1974.

Tigerstedt, Eugène N. *Interpreting Plato*. Stockholm: Almqvist and Wiksell International, 1977.

Wallis, Richard T. *Neoplatonism*. London: Duckworth, 1972.

Zervos, Christian. *Un philosophe néoplatonicien du XIe siècle: Michel Psellos: Sa vie, son œuvre, ses luttes philosophiques, son influence*. Paris: Laroux, 1920. An English-language translation was published by Burt Franklin (New York) in 1973.

Stephen Gersh (2005)

PLEASURE

The concept of "pleasure" has always bulked large in thought about human motivation and human values and standards. It seems clear to most people that pleasure and enjoyment are preeminent among the things worth having and that when someone gets pleasure out of something, he develops a desire for it. Moreover, from the time of Plato much of the discussion of the topics of motivation and value has consisted in arguments for and against the doctrines of psychological hedonism (only pleasure is desired for its own sake) and ethical hedonism (only pleasure is desirable for its own sake). One can make an intelligent judgment on these doctrines only to the extent that he has a well-worked-out view as to the nature of pleasure. Otherwise he will be unable to settle such questions as whether a putative counterexample, for instance, a desire for the welfare of one's children, is or is not a genuine example of desiring something other than pleasure for its own sake.

DEMARCATION OF THE TOPIC

Pleasure and pain have usually been regarded as opposite parts of a single continuum. As pain diminishes, it tends toward a neutral point; by continuing in the same "direction" we move toward increasing intensities of pleasure. Thus Jeremy Bentham regarded amounts of pain as negative quantities to be algebraically summated with amounts of pleasure in computing the total hedonic consequences of an action or a piece of legislation. This was in accordance with the utilitarian principle that an action is justified to the extent that it tends to produce pleasure and the diminution of pain. Since *pain* is most commonly used as a term for a kind of bodily sensation, it is natural to think of pleasure as having the same status. And indeed there are uses of the term *pleasure* in which it seems to stand for a kind of bodily sensation. Thus we speak of "pleasures of the stomach" and thrills of pleasure. But as hedonists have often insisted, in any sense of the term in which psychological or ethical hedonism is at all plausible, the term *pleasure* must be used so as to embrace more than certain kinds of localized bodily sensations. When someone maintains that pleasure is the only thing which is desirable for its own sake, he certainly means to include states of the following sort:

(1) Enjoying (taking pleasure in) doing something, such as playing tennis.

(2) Getting satisfaction out of something, such as seeing an enemy humiliated.

(3) Having a pleasant evening; hearing pleasant sounds.

(4) Feeling good, having a sense of well-being.

(5) Feeling contented being.

It seems clear that phenomena of these sorts do not consist in localized bodily sensations of the same type as headaches, except for being of an opposite quality. When someone has enjoyed playing tennis, it makes no sense to ask where (in his body) he enjoyed it. Nor does it make sense to wonder whether the pleasure he got from the tennis came and went in brief flashes, or whether it was steady and continuous; but these would be sensible questions if getting pleasure from playing tennis were a localized bodily sensation like a headache. This is not to deny that various localized sensations might be involved in his enjoyment of the game, such as a swelling in his chest after making a good shot, or a sinking sensation in his stomach after muffing a shot. The point is that his enjoyment of the game cannot be identified with such sensations, for he could be enjoying the game throughout its duration, even though such sensations cropped up only from time to time.

In fact we are confronted with two distinguishable positive-negative dimensions. There is the pleasure-pain dimension, a dimension of bodily sensations ranging from intense pains to intense localizable pleasures of the sort experienced in sexual orgasm. To specify the other dimension we need a terminological convention. We shall use the term *getting pleasure* as a general designation for an experience like those specified in the above list. Thus, enjoying listening to music and feeling good on arising in

the morning are special forms of "getting pleasure." Getting pleasure can, then, be thought of as the positive segment of a dimension, the negative segment of which will be termed *getting displeasure* and will include such things as feeling bad, feeling discontented, having a miserable time, being uncomfortable, being displeased by someone's action, being "pained" or distressed at the sight of something, and so on. We have variations of degree in this "pleasure-displeasure" dimension, as well as in the "pleasure-pain" dimension. One can enjoy oneself more or less and be displeased at something more or less. Moreover, it would seem that there is an intermediate neutral point at which one is neither pleased nor displeased at what is happening, neither enjoying oneself nor feeling miserable, and so on. It is the pleasure-displeasure dimension that philosophers are really trying to understand when they discuss "pleasure and pain." Hence we shall take the problem of the nature of pleasure to be the problem of understanding what it is to "get pleasure." For simplicity of exposition we shall largely confine the discussion to the positive segment of the pleasure-displeasure dimension; when dealing with the entire dimension we shall use the term *hedonic tone.*

It is important to realize that in posing the problem in this way philosophers (and psychologists) have assumed that there is something fundamental which is common to enjoying something, getting satisfaction out of something, being pleased at something, feeling good, and so on. It is conceivable that this assumption is mistaken, in which case virtually all the discussions of the problem have been misguided. In this article we shall follow tradition in supposing that there is an important common element to be found.

PLEASURE AS A NONLOCALIZED SENSATION

Admitting all the above, it still might be supposed that pleasure is a nonlocalized bodily sensation on the order of fatigue or "feeling energetic." (If pleasure is a sensation, it must be a bodily sensation rather than visual, auditory, tactile, olfactory, or gustatory; for it is evident that pleasure is not simply a function of the stimulation of external sense receptors.) If so, to get pleasure out of playing tennis would be to have the pleasure sensation while playing tennis. This view has been made the target of some acute critical attacks, most notably by the Oxford philosopher Gilbert Ryle. The main criticisms are as follows:

(1) Any sensation can be either pleasant or unpleasant, depending on further features of the context. A thrill can be either a thrill of pleasure or a thrill of horror. A masochist even gets pleasure out of painful sensations. Some sensations are generally pleasant (moderate warmth), others generally unpleasant (strong electric shock); but the fact that what one enjoys in a particular case depends on factors other than the kind of sensation involved, shows that we cannot identify taking pleasure in something with having a certain kind of sensation.

(2) It would seem that any sensation, if it becomes sufficiently acute, will tend to monopolize consciousness and interfere with concentration on anything else. On the view under consideration, the more pleasure we get out of, say, playing the piano, the more intense the sensation of pleasure would become, the more our attention would be taken up with the sensation of pleasure, and the harder it would become to concentrate on the playing. But the reverse is the case. The more pleasure we get out of doing something, the easier it is to concentrate on *it.*

(3) Any kind of sensation could conceivably occur without its usual conscious accompaniments and could, indeed, occupy the whole of consciousness. Even if sinking sensations in the stomach normally coincide with a perception or thought of something as dangerous, it is quite possible for one to have such sensations without being aware of anything else at the moment. Thus, on the sensation theory one could conceivably have the pleasure of playing tennis all by itself, without having it in conjunction with one's awareness that one is playing tennis. Pleasures do not seem to be detachable in the way this theory requires them to be. However, to this argument the sensation theorist could reply that we do have cases in which the pleasure sensation occurs all alone, such as feeling good or having a sense of well-being without consciously feeling good *about* anything in particular. Of course we cannot get the enjoyment of playing tennis without playing tennis, but that is just because of the way the complex phrase "enjoying playing tennis" is defined. We would not label the pleasure we get "the pleasure of playing tennis" unless the pleasure sensation occurred in conjunction with the awareness that one is playing tennis. But this verbal point does not disprove the contention that what makes enjoying playing tennis a case of getting pleasure is the presence of the

same sensation which occurs alone in feeling good (about nothing in particular).

(4) A more serious difficulty is posed by another respect in which the sensation theory represents enjoyment as loosely connected with what is enjoyed. According to the theory, to enjoy something is to have the pleasure sensation in conjunction with that something. But if "in conjunction with" means merely "in consciousness at the same time as," we are faced with the following difficulty. Let us suppose that while enjoying playing tennis at a given moment I am aware not only of playing tennis but also of oppressive humidity in the atmosphere and of a plane flying overhead. The pleasure sensation occurs in consciousness at the same time as all these cognitions. Therefore the sensation theory implies that I must be enjoying the oppressive humidity and the plane just as much as I am enjoying playing tennis. But this is contrary to the facts. A person knows immediately which of the various things he is aware of at the moment he is taking pleasure in; and the sensation theory can give no account of this discrimination. We must posit some more intimate connection between the pleasure and its object than simply being together in consciousness at the same time. But it seems that so long as we interpret getting pleasure as having a certain kind of sensation, no more intimate bond can be specified.

VARIANTS OF THE "CONSCIOUS-QUALITY" THEORY

The heavy emphasis on the bodily sensation theory in recent philosophical discussion has tended to obscure the fact that there are a number of other theories that belong to the same family, some of which have been much more important historically than the sensation theory. The general sort of view, of which the sensation theory is a variant, can be described as the view that pleasure is one of the ultimate immediate qualities (or data) of consciousness (experience). To say that it is a quality of consciousness is to say that it constitutes one of the ways in which one state of consciousness differs from another with respect to its own intrinsic nature rather than its relations to other things. (To say that a state of consciousness is a visual sensation of redness is to say something about its intrinsic nature, while to say that it belongs to Jones is not.) It is an immediate quality of consciousness because one is aware of it immediately, just by virtue of its presence; nothing further is required to get at it. Analogously, in a visual sensation one is aware of the color just by virtue of having the sensation; the color is not something that could be there without being the object of awareness. It is an ultimate quality of consciousness, because it cannot be analyzed in any way with respect to its intrinsic nature. Again we may use the less problematic sensory qualities to illustrate the point. A felt pressure differs from a felt warmth, or a seen color from a heard sound, in a way that cannot be further analyzed. To know what the difference is, one must have experienced both. Henceforth, we shall use the terms *pleasantness* and *unpleasantness* for the supposed ultimate qualities, the awareness of which is, on this kind of theory, essential for getting pleasure or displeasure.

The thesis that

(*A*) Pleasure is a kind of bodily sensation (more exactly stated, a quality that defines a kind of bodily sensation)

is one variant of this view; for qualities that do define kinds of bodily sensation are ultimate immediate qualities of experience—tingling, nausea, dizziness, and so on. However, there are other variants that are deserving of more respect.

(*B*) Pleasure is a kind of feeling, or a quality that defines a kind of feeling, where feelings are taken to be elements of consciousness distinguishable from sensations, including bodily sensations.

(*C*) Pleasure is a quality that can occur only as one aspect or attribute of some larger conscious complex, as a certain pitch or timbre occurs only as an aspect of a sound that has other aspects. Theories of this sort differ according to the sort of conscious element pleasure is thought to qualify: sensations, complexes of sensations, feelings, and so on. However, once we abandon the project of identifying pleasure with a certain kind of mental element, there is no reason not to take the most liberal alternative and consider the quality of pleasantness attachable to any sort of conscious state. This would have the advantage of not forcing us to explain away the fact that thoughts, realizations, memories, and mental images all seem to be accompanied by pleasure in the same way as sensations. For purposes of further discussion we shall take as our formulation of (*C*): *Pleasure is a quality that can attach to any state of consciousness.*

Let us consider whether the arguments against (*A*) cited above have any force against (*B*) and (*C*). Both the

first argument (that any sensation can be pleasant or unpleasant) and the second (that any sensation is capable of monopolizing consciousness) depend on specific features of bodily sensations; one could hardly expect them to have any bearing on theories that do not identify getting pleasure with having a certain kind of bodily sensation. With respect to thesis (*C*), it is not clear that every quality of conscious states is inherently neutral between being pleasant and unpleasant, nor is it clear that every quality of conscious states will monopolize attention in proportion to its degree. With respect to thesis (*B*), there are, of course, feelings that are, or essentially involve, bodily sensations (feeling nauseated, feeling tired), and the arguments do apply to these. But thesis (*B*) identifies pleasure with feelings that are distinct from bodily sensations. Apart from this qualification there are feelings, ordinarily so called, which, no matter how "strongly" one has them, do not tend to monopolize attention (feeling calm), and there are feelings that are not, by their nature, neutral between pleasantness and unpleasantness (feeling contented, feeling distressed). Such examples show that the consideration adduced in the first two arguments cannot be used to rule out the possibility that pleasure is some kind of feeling.

The third argument (that any sensation should be capable of occurring without its usual conscious accompaniments), on the other hand, does rule out the possibility of pleasure being a feeling, if a feeling is conceived as a mental element that could occur alone. However, we must remember that thesis (*B*) is distinguishable from thesis (*A*) only to the extent that it is restricted to feelings that are not identifiable, in whole or in part, with bodily sensations. And insofar as such feelings exist, it is doubtful that they are capable of occupying the whole of consciousness.

To make this point more concrete, let us look at the way position (*B*) developed. Its historical roots are to be found in the tripartite division of the mind into faculties of cognition, will, and feeling, a scheme developed in Germany in the eighteenth century by such men as Moses Mendelssohn and Immanuel Kant. Roughly speaking, the faculty of feeling is the faculty of being consciously affected, positively or negatively, by things of which one becomes aware through the faculty of cognition. Already the suggestion appears that a feeling is something that arises only in reaction to one or more cognitions and hence does not have the essential autonomy of a sensation. The introspective psychologists of the nineteenth and early twentieth centuries who tried to work out a doctrine of feeling as a distinctive kind of element of consciousness, most notably Wilhelm Wundt and E. B. Titchener, wound up with a notion of feelings as, in effect, simply hypostatized bearers of the supposed ultimate qualities of pleasantness and unpleasantness. Wundt, indeed, tried to incorporate other qualities into feelings, namely, the dimensions of strain-relaxation, and excitement-quiescence; but other workers in the field tended to regard these as features of associated bodily sensations.

More generally, it seems likely that insofar as two feelings, ordinarily so called, differ in their immediate "feel," other than with respect to pleasantness and unpleasantness, this difference can be attributed to the bodily sensations involved. Thus, if we contrast feeling homesick and feeling relieved, or feeling distressed and feeling contented, the difference in "feel," apart from different degrees of pleasantness and unpleasantness, will come down to differences in the kinds of bodily sensations involved. Hence, we are left with pleasantness and unpleasantness as the only qualitative dimension of feelings, construed as elements distinguishable from bodily sensations. Since it was generally held that such feelings could occur only in reaction to "cognitive" mental elements, including sensations, the third argument has no force against the thesis that getting pleasure out of something consists in having a pleasant feeling in conjunction with that something. But immunity from those criticisms is purchased at the price of any significant distinction between theses (*B*) and (*C*). Instead of saying that pleasantness and unpleasantness are qualities of special mental elements termed *feelings,* which can occur only in conjunction with other mental elements, we might just as well say that pleasantness and unpleasantness are qualities that can attach to any mental element. For since on the feeling theory nothing can be said about the intrinsic nature of feelings, other than that they "bear" the qualities of pleasantness and unpleasantness, it would be in principle impossible to determine by introspection whether, when I am relieved at discovering that my child is out of danger, the pleasantness I experience attaches to my awareness of the situation or to a feeling that occurs in response to my awareness. There would be a point in adopting the more complex categorization of the experience in terms of special feeling-elements if the postulation of such elements were needed for the construction of a theory as to the causes and/or effects of getting pleasure and displeasure. But the notion of feeling-elements has not so far demonstrated any theoretical fertility. Thus, when probed, thesis (*B*) reduces to thesis (*C*).

Thesis (*C*)—that pleasure is a quality that can attach to any state of consciousness—escapes the third and fourth arguments, as well as the first two. The third argument obviously has no application since, according to this thesis, pleasure can exist only as a quality of some more concrete entity. It escapes the fourth argument (that according to the sensation theory, pleasure would attach to any awareness present in consciousness at the same time) because it is possible that the quality of pleasantness would attach to one apprehension and not another, even if both are in the same consciousness at the same time. Thus, in the example given, pleasantness could attach to my awareness of playing tennis but not to my awareness of the humid atmosphere, even though I am aware of both simultaneously.

Thus thesis (*C*) emerges as the only serious contender from the ranks of quality-of-consciousness theories, and historically most such theories can be regarded as approximations to it. John Locke treated pleasure and pain as "simple ideas obtained both from sensation and reflection," and for David Hume pleasure and pain were "impressions of sensation." Neither Locke nor Hume distinguished in any systematic way between kinds of sensations, qualities of sensations, feelings, and qualities of feelings. If we look at the way they actually used the notions of an "idea of pleasure" or "impression of pleasure," we can see that in effect they took pleasure to be a qualitative feature that can attach to any state of consciousness. The "sensationist" psychologists, such as David Hartley and James Mill (whose psychology, in the hands of Jeremy Bentham and John Stuart Mill, became the basis of the utilitarian ethics and social philosophy), took pleasure and pain to be ultimate, unanalyzable properties of sensations, copies of sensations (ideas), and combinations of sensations and ideas; pleasure and pain were thought to be transferred, via association, to any mental content. None of these thinkers distinguished between the pleasure-pain and the pleasant-unpleasant dimensions, but once we clear up that confusion their view, as applied to the latter, can be seen to be a form of thesis (*C*).

CONSIDERATION OF CONSCIOUS-QUALITY THEORY

The main support for the conscious-quality theory comes from the fact, already noted, that a person knows immediately when he is getting pleasure from something. He knows it in a way no one else could conceivably know it—just by virtue of being the one who is getting the pleasure. He has an epistemologically "privileged access" to the fact. Since it is natural to take the awareness of sensory qualities, especially visual ones, as a paradigm of immediate knowledge of one's psychological states, it is natural to construe what one knows when he knows that he is enjoying something as some ultimate quality of consciousness.

Nevertheless, on further probing, the thesis that pleasure is a quality that can attach to any state of consciousness is not very plausible phenomenologically. When we reflect on a wide variety of cases of getting pleasure, as indicated by the list at the beginning of this article, we are unable to isolate a felt quality that they all share, in the way in which we can easily isolate a quality of redness which a number of different visual sensations share, or a quality of painfulness that a number of different bodily sensations share. On the contrary, enjoying playing tennis feels very different from getting satisfaction out of seeing an enemy in distress, and both feel very different from the sense of well-being one has when, in good health, one arises carefree from a good night's sleep. Nor does it seem possible to find *in* these experiences some respect in which they are qualitatively the same, as two sounds, otherwise very different, can be the same in pitch. Even if we stick to one term in the "pleasure family," such as *getting satisfaction,* it seems equally implausible to suppose that there is some felt quality common to getting satisfaction out of seeing an enemy in distress and getting satisfaction out of the realization of a job well done. The enjoyment or satisfaction seems to take whatever felt quality it has from what one is enjoying or getting satisfaction from. Thus John Stuart Mill was on sound ground in insisting, against Bentham, that there are qualitative differences between "pleasures."

These doubts are reinforced by the fact that here we are without external support for the postulation of basic conscious qualities. In the case of sensory qualities, at least those of the external senses, we can tie down the quality to a certain kind of stimulation; people ordinarily get red visual sensations when and only when their optic nerves are stimulated by stimuli of a certain physical description. Moreover, certain kinds of variations in the physical properties of the stimulus can be correlated with judgments of degrees of properties of the sensation, such as hue, saturation, and shade. These correlations support our confidence in purely introspective discriminations between visual qualities. Nothing of the sort is possible with pleasantness. This quality, if such there be, does not vary with variations in physical stimuli in any discernible fashion. Nor can anything much better be found on the response side. It is true that there are gross typical differ-

ences in bearing and manner between a person enjoying himself and a person having a miserable time, between a person satisfied with the way things are going and a person who feels terribly frustrated. On the positive side of these contrasts we are more likely to get relaxation, expansiveness, and smooth coordination; on the negative side tenseness, constriction, and disruption of ongoing activities. But these manifestations differ so much from case to case because of other factors—general personality characteristics and state of health, for example—that they cannot be taken as reliable indications of how much pleasure or displeasure a given person is getting at the moment.

MOTIVATIONAL THEORIES

No doubt there is something that all the experiences we have classified under "getting pleasure" have in common. If it is not an immediately felt quality, what is it? In searching for an alternative we might well take note of a different tradition in which the notion of pleasure was analyzed motivationally, in terms of the realization of the good, of the object of striving. In many systematic schemes of the "passions of the soul," the basic notion is appetite, inclination, striving, or tendency of the person toward some object he apprehends as good or desirable. Pleasure, delight, or joy is then defined as the state in which this object is actually present, in which the appetite has reached fruition. Versions of this view are to be found in Thomas Aquinas, Thomas Hobbes, Benedict de Spinoza, and many other philosophers, as well as in some more-recent psychologists, notably William McDougall. The basic presuppositions of this approach to the subject are quite different from those of Locke and Hume. For Locke and Hume, and British empiricists generally, the way to understand any psychological concept is either to find it among the immediate data of introspection or to show how it is to be analyzed into such data. This approach ultimately stems from the Cartesian insistence that one knows one's own states of consciousness better than anything else, in particular, better than physical objects and events, since it is possible to doubt the existence of all the latter but not of all the former. Hence it is natural for one in this tradition freely to posit immediate qualities of consciousness whenever there is any plausibility to doing so. Thinkers in the other tradition have a more objectively oriented epistemology, according to which conscious experience has no priority over, for instance, overt behavior as an object of investigation and an object of knowledge. This leaves them free to explore the possibility of analyzing the notion of pleasure in terms of notions like appetite, or tendency, which could not be regarded as immediate objects of introspection.

Their view of the nature of pleasure might be formulated as follows:

(*D*) To get pleasure is to be in a state of consciousness which includes the awareness that one has obtained something one wants.

There are serious difficulties with this version of a motivational theory of pleasure. No doubt there are many pleasures that do presuppose a want in the absence of which no such pleasure would be forthcoming. I would not take pleasure in the discomfiture or prosperity of a certain person unless I wanted him to be discomfited or to prosper, as the case may be. But it seems that there are many pleasures which do not presuppose any such preexisting want. Simple sensory pleasures, such as the pleasure of eating a good steak, are the most obvious cases. Having found steak pleasant, we may then develop a desire for a steak; but here the want presupposes the prior experience of pleasure, not vice versa. The view under consideration does not deny that wants can be reinforced or strengthened by the experience of pleasure in their satisfaction. But it does deny that one can get pleasure from anything except by way of that thing satisfying some previously existing want. And this seems contrary to experience. Surely infants take pleasure in many things, such as throwing a ball, when they encounter them for the first time. Prior to this encounter they could not have had a desire for it, for they did not yet know what throwing a ball is. It is noteworthy that proponents of this position maintain it in the face of these difficulties only by generously positing instincts and other nonconscious "tendencies" and "strivings."

However, there are other versions of a motivational theory that do not presuppose a preexisting desire for each pleasure. The most promising is a view put forward by Henry Sidgwick, among others:

(*E*) To get pleasure is to have an experience that, as of the moment, one would rather have than not have, on the basis of its felt quality, apart from any further considerations regarding consequences.

This account makes pleasure a function not of a preexisting desire but of a preference one has at the moment of the experience. To say that one has the preference at the moment is not to say that one expresses the preference even to oneself; it is not to say anything about what is before one's consciousness at the moment. It is, rather, to say something dispositional—for example, that one would choose to have an experience just like this rather

than not if one were faced with such a choice at this moment and if no considerations other than the quality of the experience were relevant. This, unlike thesis (*D*), allows for the possibility of taking pleasure in something one did not previously have a tendency to seek. On the other hand, it is also clearly distinct from the conscious quality theory. According to thesis (*E*), when one says that he is enjoying something, he is saying something about the quality of his experience; he is saying that the quality of his experience is such that on that basis alone he would prefer to have it rather than not to have it. But he is not saying what the quality of his experience is; he is saying, rather, how it is related to his preferences, likes, or desires. More particularly, he is not saying that there is some particular quality, "pleasantness," present in the experience. On this view, the felt qualities on the basis of which the experience is valued can be as diverse as the range of human likes. They can involve calm, excitement, warmth, cold, thrills, and sinking feelings.

It might seem that the strongest reason for the conscious-quality view, the fact that pleasure is something to which the subject has privileged access, would pose a difficulty for thesis (*E*), but this is not necessarily so. It is natural to think that the only things an individual can know about immediately, in a way no one else can, are the qualities of his experience; and indeed sensory qualities have this status. But there are many things to which an individual has privileged access that cannot be regarded as immediately felt qualities, such as intentions, attitudes, and beliefs. If I intend to quit my job tomorrow, I know that I have this intention without having to do any investigation to find out; I know just by virtue of having the intention; I know this as immediately as I know that I am now aware of a reddish patch. And it is in principle impossible for anyone else to know in this way that I have that intention. Yet an intention is neither a felt quality nor a complex of felt qualities. Hence the epistemological status of pleasure is not a conclusive reason for construing it as a quality of experience. The epistemological status of pleasure does place a constraint on the range of possible theories; we cannot identify pleasure with something to which the subject does not have privileged access, such as a certain pattern of neuron firings in the brain. However, among the nonsensory quality items to which a person has privileged access are his likes, preferences, and wants. It seems reasonable to suppose that a person's knowledge that he would choose to have an experience just like his present one on the basis of its felt quality can be just as immediate as his knowledge that he is aware of a red patch.

Motivational theories have the following superiority over conscious-quality theories. It does not seem to be merely a contingent fact that pleasure is desirable, or that the fact that an activity is enjoyable is a reason for doing it. "I get a lot of satisfaction out of teaching, but I see absolutely no reason to do it" sounds like a self-contradiction. This is not to say that the fact that one will get pleasure out of something is a conclusive reason for doing it; there may well be other considerations that outweigh this. I would enjoy playing tennis now, but if an urgent job has to be completed, that is a good reason for not playing tennis. What we are suggesting to be necessarily true is (*P*) the fact that one gets pleasure out of *x* is a reason for doing or seeking *x*. This reason must be put into the balance along with other relevant reasons in making a decision in any particular case. The conscious-quality theory can throw no light on this necessity. If pleasure is an unanalyzable quality of experience, there is nothing about the meanings of the terms involved in (*P*) that would make it necessarily true. Why should it be necessarily true that a certain unanalyzable quality of experience is something to be sought? It would seem that any such quality is something that would or would not be taken as desirable by a given person, or people in general, depending on further factors. A motivational theory, on the other hand, analyzes the concept of pleasure in such a way as to make principles like (*P*) necessary. If to enjoy an experience *is* just to be disposed to choose an experience exactly like it if nothing other than the felt quality is relevant, then it follows trivially that the fact that something involves enjoyment is a reason for choosing it.

Superficially it might appear that opting for a motivational theory would involve a commitment to psychological hedonism, but this would be a mistake. The motivational theory commits us to holding that pleasure is (always) intrinsically desirable, but it carries no implication that pleasure is the only thing intrinsically desirable. One could adopt thesis (*E*) as his theory of the nature of pleasure and still regard other things as intrinsically desirable, such as fulfillment of one's potentialities and intellectual consistency, independent of any pleasure they might bring. It is an analysis of desire in terms of pleasure that would stack the cards in favor of psychological hedonism. If we hold that to desire something is to think of it as pleasant, it does follow that we do not desire anything except pleasure or what is believed to lead to pleasure.

THE MEASUREMENT OF PLEASURE

The problem of measuring hedonic tone has occupied both psychologists and philosophers. Psychologists have addressed themselves to such problems as the physiological basis of pleasure, the dependence of pleasantness on various aspects of sensory stimulation (such as contrast), and the effect of pleasure and displeasure on the speed and efficiency of learning. To deal with these problems they have to study the effect of variation of sensory stimulus conditions, for instance, on degree of hedonic tone, or the effect of variations in hedonic tone on something else, such as ease of recall of learned material. To do this, one must be able to specify the degree or amount of hedonic tone present at a given moment. Philosophical concern with the measurement of pleasure has grown out of utilitarianism and other hedonistic ethical theories. According to utilitarianism, an action is justified if and only if it will probably lead to a greater balance of pleasure over displeasure for everyone affected than any possible alternative action. Applying this principle to a particular case would involve estimating the total quantity of pleasure and displeasure that would be produced by each of the possible choices. To do this we would first have to list the ways in which one choice or another would make the situation, patterns of activities, and so on of a given person different from what they would be if that choice had not been made. Second, we would have to obtain information concerning how much pleasure or displeasure that person has derived from the situations and activities in question. Third, we would have to project how much pleasure and displeasure the person would derive from each of these in the future, taking into account any changes in circumstances, age, and so on that could be expected to make a difference. Fourth, we would have to sum up the hedonic consequences for that person. Fifth, having done this for each person likely to be affected, we would have to sum these results, arriving at a figure representing the probable total hedonic consequences of that choice.

Some of the problems relevant to these procedures fall outside the scope of this article. These include the problem of determining just what the objective consequences of a choice are likely to be, the problem of determining what features of a situation are responsible for the pleasure or displeasure felt, and the problem of projecting probable future pleasure from past pleasures. These are all essentially general problems of inductive reasoning. The problems having to do specifically with the measurement of pleasure are (1) How can one determine the degree of pleasure or displeasure experienced by a given person at a given moment? (2) How can one compare the amount of pleasure felt by one person at a given time with the amount of pleasure felt by another person at a given time?

In everyday discourse we compare pleasures and displeasures. We say things like "I didn't enjoy that party as much as the last one," "I get more pleasure out of gardening now than I used to," and "That interview was not as unpleasant as I had expected it to be." Even granting the reliability of such comparative judgments, the utilitarian needs something more. He needs to be able to specify the hedonic value of particular experiences in numbers that he can meaningfully subject to arithmetical operations, so that if a person gets four positive units (pleasure) from one minute of playing tennis and one negative unit (displeasure) from the next minute of playing tennis, the total hedonic value of the two minutes is greater than that of two minutes spent lying in the sun, from which he derived one positive unit per minute.

An obvious move is to try to refine everyday comparative judgments in such a way as to yield these kinds of results. (In fact, all the methods that have actually been used have been of this sort.) We might ask the subject to consider a large number of his past experiences and to make a comparative judgment on each pair. Possibly after ironing out a few inconsistencies, we would arrange a series such that each experience is more pleasant, or less unpleasant, than any experience lower in the series. We could then have the subject locate a point of hedonic indifference, after which we could assign positive and negative integers to the ranks diverging in either direction from the point of indifference. This would constitute a hedonic scale for that individual. Any other experience would be assigned a number by matching it with an experience on the scale from which it is hedonically indistinguishable. (If it fell between two experiences on the scale, the scale would have to be revised.)

Even assuming that subjects make responses that would enable us to set up an unambiguous scale, one might still doubt that it provides an adequate measuring procedure. First, it relies heavily on the subject's memory of how much pleasure or displeasure he got out of something in the past, and such memories are notoriously fallible. Second, even if we have constructed a scale such that, given two adjoining experiences, the subject is unable to think of an experience which would lie between them, it is still an open question whether the intervals between the items are equal. We have as successive items (*a*) taking a shower after a game of tennis, (*b*) being complimented on a performance, and (*c*) seeing one of one's

children receive a prize. What reason is there to think that (*c*) is just exactly as much more pleasant than (*b*) as (*b*) is than (*a*)? And yet we have to make that assumption if we are going to use the numerical assignments to compare one "sum" of pleasures with another.

A different procedure would be to have the subject rate an experience, when it happens, by an absolute scale, for instance, a nine-point scale ranging from +4 to –4. This would avoid the problem about memory, but it brings fresh difficulties in its stead. Why suppose that the subject is in fact using the same standards every time we get him to make a rating? For that matter, why suppose that ratings which people are forced to make on an artificially constructed scale correspond to any real differences in experience at all? Moreover, there is still the question of whether the intervals on our "absolute scale," as used by the subject, reflect equal differences in actual degree of hedonic tone. If one of these procedures yielded measurements that entered into well-confirmed hypotheses relating hedonic tone to, for example, various properties of learning, this would bolster our confidence in the procedure. At least it would show that we were measuring something important. But such results have not been obtained to any considerable extent.

Even if all the above problems were surmounted, it would still be very difficult to compare the amount of pleasure or displeasure experienced by two different people. Suppose that I am trying to determine whether the total balance of pleasure over displeasure (or the reverse) is greater for my wife or for myself with respect to a given party. Even if the foregoing problems could be surmounted and we could find a valid way of assigning a hedonic number for each of us, relative to a scale for each, how are we to calibrate the two scales? How are we to determine whether a rating of +3 on my scale represents the same amount of pleasure as a rating of +3 on her scale?

So long as we restrict ourselves to refinements of the method of introspective judgment, the problem of intersubjective comparison seems insoluble. On the other hand, if there were some intersubjectively measurable variable, or complex of variables, which we had reason to think is intimately related to hedonic tone and which correlated well enough with rough introspective judgments to be taken as a measure of hedonic tone, all problems would be solved. Such a development is still in the future. Attempts to correlate introspective hedonic judgments with gross physiological variables on the order of pulse rate or patterns of respiration have not been fruitful. There has been no end of speculation concerning the neurological basis of hedonic tone. Pleasantness has been thought to depend on the degree to which assimilation counteracts dissimilation in the activity of any group of central neurones (A. Lehmann), the degree of the capacity of a neural element to react to stimulation (H. R. Marshall), the average rate of change of conductance in the synapses (L. T. Troland), and so on. Thus far, none of these theories has yielded effective physiological measures.

See also Bentham, Jeremy; Cartesianism; Empiricism; Good, The; Hartley, David; Hedonism; Hobbes, Thomas; Hume, David; Kant, Immanuel; Locke, John; McDougall, William; Mendelssohn, Moses; Mill, James; Mill, John Stuart; Pain; Plato; Ryle, Gilbert; Sensa; Sidgwick, Henry; Spinoza, Benedict (Baruch) de; Thomas Aquinas, St.; Utilitarianism; Wundt, Wilhelm.

Bibliography

Plato discusses the nature of pleasure in his dialogues *Philebus* and *Timaeus.* Aristotle's main discussion is in Book X of the *Nicomachean Ethics.* A very illuminating history of the topic is H. M. Gardiner, R. C. Metcalf, and J. G. Beebe-Center, *Feeling and Emotion, A History of Theories* (New York, 1937).

Important discussions by contemporary analytical philosophers include Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), Ch. 4, and *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954), Ch. 4; a symposium, "Pleasure," between Ryle and W. B. Gallie in *PAS,* Supp. 28 (1954): 135–164; Terence Penelhum, "The Logic of Pleasure," in *Philosophy and Phenomenological Research* 17 (1957): 488–503; Anthony Kenny, *Action, Emotion, and Will* (London: Routledge and Kegan Paul, 1963), Ch. 6; P. H. Nowell-Smith, *Ethics* (London: Penguin, 1954), Chs. 8–10; and R. B. Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1959), Ch. 12.

Karl Duncker's article "Pleasure, Emotion and Striving," in *Philosophy and Phenomenological Research* 1 (1940): 391–430, is a brilliant phenomenological analysis of the varieties of pleasure and their relation to desire. The psychological research on hedonic tone is well summarized in J. G. Beebe-Center, *The Psychology of Pleasantness and Unpleasantness* (New York, 1932). More recent developments are surveyed in Beebe-Center's article "Feeling and Emotion," in *Theoretical Foundations of Psychology,* edited by Harry Helson (New York: Van Nostrand, 1951).

The measurement of pleasure is discussed in Jeremy Bentham, *Introduction to the Principles of Morals and Legislation* (2nd ed., Oxford, 1907), Ch. 4; Henry Sidgwick, *The Methods of Ethics* (London: Macmillan, 1874), Book II, Chs. 2 and 3 (also good on the nature of pleasure); R. B. Perry, *General Theory of Value* (New York: Longmans Green, 1926), Ch. 21; and Robert McNaughton, "A Metrical Conception of Happiness," in *Philosophy and Phenomenological Research* 14 (1954): 172–183.

William P. Alston (1967)

PLEKHANOV, GEORGII VALENTINOVICH
(1856–1918)

Georgii Valentinovich Plekhanov, the Russian Marxist, revolutionary, philosopher, sociologist, and historian of social thought, was the son of a poor nobleman. After graduating from a military academy in Voronezh, he studied at the Mining Institute in St. Petersburg. As a student he joined the revolutionary movement and became one of the leaders of the revolutionary organization of the Narodniki (Populists), called Zemlia i volia (Soil and freedom). After Zemlia i volia split into the terroristic Narodnaia volia (People's freedom) and the Bakuninist-anarchist Chernyi peredel (Redistribution of soil) groups, Plekhanov became the leading theoretician of the Chernyi peredel group.

In the beginning of 1880, Plekhanov emigrated to France and then settled in Switzerland. Between 1880 and 1882 he turned from Populism to Marxism, and in 1883 he founded in Geneva the first Russian Marxist group, Osvobozhdenie truda (The emancipation of labor). In the summer of 1889 he took part in the founding congress of the Second International. In the late 1890s Plekhanov was one of the first to criticize both the international revisionism of Eduard Bernstein and its Russian variant, "economism."

In 1900, Plekhanov's group joined forces with a new group headed by V. I. Lenin. The two groups organized the second congress of the Russian Social-Democratic Labor Party in London in 1902. The congress accepted a party program written mainly by Plekhanov. Disagreements over the nature of the party led to the split of the party into Bolsheviks and Mensheviks. Plekhanov supported Lenin at the congress, but he became neutral soon afterward and even leaned to the Menshevik side.

During the first Russian revolution (1905), Plekhanov severely criticized the tactics of the Bolsheviks, but after the defeat of the revolution he again came closer to Lenin. The onset of World War I led to the final parting of Plekhanov and Lenin. Plekhanov urged socialists to support the Allied governments, but Lenin declared war on the imperialist war.

After the February revolution of 1917 Plekhanov returned to Russia. Believing that Russia was not yet sufficiently mature for socialism, he regarded the October revolution as a fateful mistake. Nevertheless, he refused to engage in active struggle against Soviet authority.

As the founder of the first Russian Marxist group, Plekhanov is rightly called the father of Russian Marxism and of Russian social democracy. He was also an outstanding leader of the Second International. But the workers' movement is indebted to Plekhanov for his theoretical work, especially in philosophy, even more than for his practical organizational activity.

GENERAL PHILOSOPHICAL VIEWS

Plekhanov regarded himself as an orthodox follower of Karl Marx and Friedrich Engels and severely criticized those who tried to "revise" the basic teachings of Marx or to "supplement" them with the ideas of Immanuel Kant, Ernst Mach, or some other philosopher. But he insisted that the views of Marx and Engels should be developed further.

In his early writings Plekhanov exhibited the tendencies to reduce philosophy to the philosophy of history and to regard philosophy as a preliminary to science. He later stressed the independent tasks and problems of philosophy and defined philosophy in a broader way, as a study of the basic principles of being and knowledge and of their mutual relationships. Whereas Marx and Engels often insisted on the methodological character of their philosophy, Plekhanov stressed its systematic character. Marxist philosophy, according to Plekhanov, is a system, which Plekhanov named dialectical materialism.

Following Engels, Plekhanov maintained that the basic question of every philosophy was "the question about the relationship of subject to object, of consciousness to being," and he regarded materialism and idealism as two basic answers to the question. Dualism was a possible, but weaker, answer. A consistent thinker must choose between an idealistic and a materialistic monism, but vulgar materialism is not the only alternative to idealism. The real solution is dialectical materialism.

As the concept of matter was not clearly defined by Engels, Plekhanov made several attempts to do so. His formulations were more or less modifications of the traditional materialist view that matter is what exists independently of man's consciousness, affects his sense organs, and produces sensations. Plekhanov tried to show that opposing philosophies that maintain the world exists only in the consciousness of one man (solipsism), only in the consciousness of humankind (solohumanism), or only in that of some superindividual objective spirit (objective idealism) all lead to contradictions. The belief in the existence of the external world is, according to Plekhanov, an unavoidable leap in philosophy. Lenin reproached Plekhanov for such Humean terminology,

and Soviet philosophers later exploited this criticism to accuse Plekhanov of Humeanism.

In criticizing idealistic views that *mind, spirit, consciousness*, or *psyche* (he used these terms more or less interchangeably) is the only reality, Plekhanov at the same time rejected the view of those materialists who regard mind as a part of matter or (as Engels did) as a form of the movement of matter. Nevertheless, he held that mind is one of the properties of substance, or matter. In some earlier writings Plekhanov affirmed that mind is merely a mode of matter, a property characteristic of matter organized in a certain way. Later he modified his view, maintaining that mind is an attribute of matter, a property that, at least to a minimal, nonobservable degree, belongs to all matter. This theory led to his being accused of hylozoism. Plekhanov first thought that mind could be regarded as a consequence of another, more fundamental property of matter, movement. Later he changed this view and asserted that consciousness is an "inner state" of matter in motion, a subjective side of the same process whose objective side is motion.

Accepting the traditional correspondence theory of truth, Plekhanov tried to explain in a more specific way the character of correspondence or agreement holding between thought and reality. Against naive realism he stressed that "correspondence" does not mean "similarity." He maintained that sensations are "hieroglyphs" because although they can adequately represent things and their properties, they are not "similar" to them. To avoid misinterpretation of his views, Plekhanov later renounced this terminology; nevertheless, he was severely criticized for it by some Soviet philosophers, who held that it was a concession to Kantianism.

Plekhanov often stressed that Marxist philosophy is *dialectical* materialism and that dialectics is the soul of Marxist philosophy. But in explaining his conception of dialectics, he added little to what had already been said by Marx and Engels. He was more original in his view of the relationship between formal logic and dialectics. Starting from Engels, who likened the relationship between the two to that between lower and higher mathematics, Plekhanov maintained that thinking according to the laws of formal logic is a special case of dialectical thinking. By the help of a number of distinctions, like those between motions and things, between changing and relatively stable things, and between simple and compound things, he tried to determine more precisely the limits of fields in which the two logics could be applied. These explanations, although they gave no final clarification of the problem, nevertheless were the most explicit treatment of the problem in classical Marxist literature and served as the starting point for many later discussions.

PHILOSOPHY OF HISTORY

Plekhanov's views on the philosophy of history have sometimes been misinterpreted. The fault is partly his own. Trying to present Marx and Engels's view on the relations between the economic foundation and the superstructure in a simple schematic way, he produced a formula involving:

> 1. The state of the forces of production; 2. Economic relations conditioned by these forces; 3. The socio-political regime erected upon a given economic foundation; 4. The psychology of man in society, determined in part directly by economic conditions and in part by the whole socio-political regime erected upon the economic foundation; 5. Various ideologies reflecting this psychology. (*Fundamental Problems of Marxism*, edited by D. Ryazanov, p. 72)

This formula may be regarded as an adequate schematization of economic materialism, the theory according to which the economic factor (the forces of production) is ultimately predominant in history. However, in other places Plekhanov maintained that neither man as man nor society as society can be characterized by a constant relationship between economic and other factors because such relationships are always changing. He even explicitly criticized the view that the economic factor must always be decisive and called it a "libel against mankind." Plekhanov admitted that so far men have been the "slaves of their own social economy," but he insisted that "the triumph of human reason over the blind forces of economic necessity is possible" (*Izbrannye filosofskie proizvedeniia* [Selected philosophical works], Vol. II, p. 233).

In his best writings Plekhanov criticized not only the theory of the predominant role of the economic factor but also the theory of factors as such. In polemics against those who attributed the theory to Marx, he maintained that genuine materialists are averse to dragging in the economic factor everywhere and that "even to ask which factor predominates in social life seems to them pointless" (*The Materialist Conception of History*, p. 13). The question is unjustified because, "strictly speaking, there exists only one factor of historical development, namely—social man" (*Izbrannye Filosofskie Proizvedeniya*, Vol. V, p. 363); different branches of the social sciences—ethics, politics, jurisprudence, political economy—investigate one and the same thing, the activity of social man.

AESTHETICS

Plekhanov was one of the few Marxist thinkers interested in aesthetics and the sociology of art. Criticizing the view that art expresses only feelings, he insisted that it expresses both feeling and thoughts, not abstractly, however, but in lively pictures. He added that the pictorial expression of feelings and thoughts about the world is not an end in itself but is done in order to communicate one's own thoughts and feelings to others. Art is a social phenomenon.

The first task of an art critic, according to Plekhanov, is to translate the idea of a work of art from the language of art into the language of sociology in order to find what could be called the sociological equivalent of a literary phenomenon. After the first act of materialistic criticism, the second act—the appreciation of the aesthetic values of the work in question—must follow.

Investigating the social roots of the theory of art for art's sake and of the utilitarian view of art, Plekhanov came to the conclusion that the inclination toward art for its own sake emerges from a hopeless separation of the artist from the surrounding social milieu, whereas the utilitarian view of art emerges when a mutual understanding between the larger part of society and the artist exists. The utilitarian view of art can thus be combined with both conservative and revolutionary attitudes.

The value of a work of art is primarily dependent on the value of the ideas it conveys, but correct ideas are not enough for a valuable work. A work of art is great only when its form corresponds to its ideas.

IMPORTANCE AND INFLUENCE

Although Plekhanov is not one of those greatest of philosophers who have opened up new vistas to humankind, he was not a mere popularizer of Marxist philosophy. Starting from Engels's interpretation of Marxist philosophy, he improved it and developed it in many directions. He greatly influenced Lenin's conception of Marxist philosophy, and through both his own works and Lenin's he decisively influenced Soviet philosophy between the two world wars. The leaders of the Soviet "philosophical front" in the 1920s, A. M. Deborin and Deborin's most outstanding opponent, L. I. Aksel'rod, were Plekhanov's immediate disciples.

In 1930 a new period in Soviet philosophy began, a period that included severe criticism of Plekhanov. All kinds of accusations were made against Plekhanov, but the Stalinist criticism abated in the 1940s and 1950s, and Plekhanov's philosophical views survived. Nevertheless, the publication of previously unpublished writings of Marx in the 1930s and 1940s and new discussions of Marx's philosophy in the 1950s and the 1960s seem to have produced an interpretation of Marxist philosophy that is more profound than that offered by Engels and developed by Plekhanov and Lenin.

See also Aesthetics, History of; Art, Expression in; Art, Value in; Correspondence Theory of Truth; Deborin, Abram Moiseevich; Dialectical Materialism; Engels, Friedrich; Kant, Immanuel; Lenin, Vladimir Il'ich; Mach, Ernst; Marx, Karl; Marxist Philosophy; Panpsychism; Russian Philosophy.

Bibliography

WORKS BY PLEKHANOV

Collections

Sochineniya (Works). 24 vols, edited by D. Ryazanov. Moscow and Leningrad, 1922–1927.

Literaturnoe nasledie G. V. Plekhanova (The literary heritage of Plekhanov). 8 vols. Moscow, 1934–1940.

Selected Philosophical Works. 5 vols. Moscow: Progress, 1974–76.

Other Works

Anarchism and Socialism. Westport, CT: Hyperion Press, 1981.

The Development of the Monist View of History. Moscow: Progress Publishers, 1974.

Essays in the History of Materialism. Translated by Ralph Fox. New York: H. Fertig, 1967.

The Materialist Conception of History. New York: International Publishers, 1964.

K Voprosu o Razvitii Monisticheskogo Vzglyada na Istoriyu. St. Petersburg, 1895. Translated by Andrew Rothstein as *In Defense of Materialism: The Development of the Monist View of History*. London, 1947; 2nd ed., published as *The Development of the Monist View of History*, Moscow: Foreign Languages Publishing House, 1956.

"K Voprosu o Roli Lichnosti v Istorii." *Nauchnoe Obozrenie*, Nos. 3–4 (1898). Translated as *The Role of the Individual in History*. New York, 1940.

Fundamental Problems of Marxism. Moscow: Progress Publishers, 1977.

Art and Society & Other Papers in Historical Materialism. New York: Oriole Editions, 1974.

Materialismus militans: Reply to Mr. Bogdanov, edited by Richard Dixon, Moscow: Progress, 1973.

WORKS ON PLEKHANOV

Baron, Samuel H. *Plekhanov in Russian History and Soviet Historiography*. Pittsburgh: University of Pittsburgh Press, 1995.

Baron, Samuel H. *Plekhanov: The Father of Russian Marxism*. Stanford, CA: Stanford University Press, 1963.

Fomina, V. A. *Filosofskie vzgliady Plekhanova* (Philosophical views of Plekhanov). Moscow, 1955.

Hook, Sidney. *The Hero in History*. New York: John Day, 1943.

Howard, M. C., and J. E. King. "The Political Economy of Plekhanov and the Development of Backward Capitalism." *History of Political Thought* 10 (1989): 329–344.
Kolakowski, Leszek. "Plekhanov and the Codification of Marxism." In *Main Currents of Marxism*. Vol. 2. Oxford: Oxford University Press, 1978.
Petrović, Gajo. *Filozofski pogledi G. V. Plehanova* (Philosophical views of G. V. Plekhanov). Zagreb, Yugoslavia, 1957.
Steila, Daniela. *Genesis and Development of Plekhanov's Theory of Knowledge: A Marxist between Anthropological Materialism and Physiology*. Dordrecht: Kluwer Academic, 1991.
Vaganyan, V. *G. V. Plekhanov: Opyt kharakteristiki sotsial'no-politicheskikh Vozzrenii* (G. V. Plekhanov: An essay on the characteristics of his sociopolitical views). Moscow, 1924.
Vaganyan, V. *Opyt bibliografii G. V. Plekhanova* (Bibliographical essay on G. V. Plekhanov). Moscow and Leningrad, 1923.
Volfson, S. *G. V. Plekhanov*. Minsk, U.S.S.R., 1924.
Walicki, A. *A History of Russian Thought from the Enlightenment to Marxism*. Oxford: Clarendon Press, 1980.
Yovchuk, M. T. *G. V. Plekhanov i ego trudy po istorii filosofii* (G. V. Plekhanov and his works in the history of philosophy). Moscow, 1960.
Zinoviev, Grigori. *G. V. Plekhanov*. Petrograd, 1918.

Gajo Petrović (1967)
Bibliography updated by Vladimir Marchenkov (2005)

PLESSNER, HELMUT

(1892–1985)

Helmut Plessner, was, with Max Scheler, the founder of modern philosophical anthropology. Born in Wiesbaden, Germany, he studied medicine, and then zoology and philosophy, at the universities of Freiburg, Heidelberg, and Berlin. He received a doctorate in philosophy from Erlangen in 1916 and his *Habilitation* in philosophy with Scheler and Hans Driesch at Cologne in 1920. His academic career in Germany was terminated by the National Socialist regime, and in 1934 he went to Groningen, the Netherlands, first as a guest of the Physiological Institute (where he was associated with F. J. J. Buytendijk), then as Rockefeller fellow, and from 1929 to 1942 as professor of sociology. Again dismissed by the Nazis, he was reinstated at Groningen by the Dutch in 1945 and occupied the chair of philosophy from 1946 to 1951. In 1951 he accepted the chair of sociology at the University of Göttingen in Germany. He became professor emeritus in 1962 and lectured as a visiting professor at the New School for Social Research in New York in 1962–1963. He received an honorary doctorate from Groningen in 1964.

Plessner's work—he published twelve books and approximately ninety monographs, essays, and papers—ranges over an extraordinarily wide area, including animal physiology, aesthetics, phenomenology, the history of ideas, the history of philosophy, sociological theory, sociology of knowledge, sociology of education, and political sociology. Most of these studies are linked to the problems of philosophical anthropology, the discipline to which he devoted his most important publications. His background in zoology and physiology, his phenomenological training under Edmund Husserl, and his sociological orientation led him to redefine the problems and findings of the modern sciences of man.

Plessner agrees with the view that man artificially creates his nature, or more precisely, that what man makes of himself is contingent on history. However, man is bound by the structural principle of his position in the world; in contrast to the centricity of animals, who are, simply, what they are as organisms, in their *Umwelt*, man is "eccentric." Plessner rejects the dualism of spirit and matter present in Scheler's anthropology. He sees man as being a body (with such organically determined traits as upright posture, impoverishment of instincts, and drive surplus) and consequently exposed to his environment, and also as having a body and acting by means of it, as being open, within certain limits, to the world. Man is both "inside" and "outside" himself. Social and historical order is based on the precarious balance of these two dialectical moments. This order enables man to maintain a distance from things, from situations, and from himself, making it possible for him to use language and to plan actions. Man's eccentricity leads him to enter history, "to make himself" in history. However, when man faces ambivalent or insuperable situations, the balance on which order is founded is disrupted; planned action, speech, and all historically determined "orderly" ways of coming to terms with the world are blocked. His indirect, socially mediated relationship with the world momentarily breaks apart. In such marginal situations man responds in a prehistorical, presocial, and yet peculiarly human manner: by laughter or by tears.

See also Driesch, Hans Adolf Eduard; Husserl, Edmund; Philosophical Anthropology; Scheler, Max.

Bibliography

Plessner's most important works are *Die Einheit der Sinne* (Bonn: Cohen, 1923), *Die Stufen des Organischen und der Mensch* (Berlin and Leipzig: de Gruyter, 1928; 2nd ed., Stuttgart, 1964), *Das Schicksal deutschen Geistes im Ausgang seiner bürgerlichen Epoche* (Zürich and Leipzig: Niehans, 1935; 2nd ed. published as *Die verspätete Nation*, Stuttgart: Kohlhammer, 1959), *Lachen und Weinen* (Munich, 1941), and *Zwischen Philosophie und Gesellschaft* (Berlin and Munich, 1953).

For literature on Plessner see Jürgen Habermas, "Anthropologie," in *Fischer Lexikon*, Vol. II, *Philosophie*, edited by Alwin Diemer and Ivo Frenzel (Frankfurt, 1958).

OTHER RECOMMENDED WORKS

Essbach, Wolfgant, et al. *Plessners "Grenzen der Gemeinschaft": eine Debatte*. Frankfurt am Main: Suhrkamp, 2002.

Marx, Werner, and Helmut Plessner. *Das Spiel, Wirklichkeit und Methode*. Freiburg i. Br., Schulz, 1967.

Schultz, Walter, and Helmut Fahrenbach, eds. *Wirklichkeit und Reflexion: Walter Schulz zum 60. Geburtstag*. Pfullingen: Neske, 1973.

Thomas Luckmann (1967)
Bibliography updated by Michael Farmer(2005)

PLETHO, GIORGIUS GEMISTUS

(c. 1355–1452)

Giorgius Gemistus Pletho, the leading Byzantine scholar and philosopher of the fifteenth century, was born in Constantinople, the son of a cleric. Pletho is noted primarily for advocating a restoration of ancient Greek polytheism and, above all, for inspiring the interest of the Italian humanists of the *quattrocento* in the study of Plato. His studies followed the usual pattern of Byzantine education, emphasizing the classical Greek heritage. Influenced by certain of his teachers, Pletho became interested primarily in the philosophy of Plato, whose writings had again been brought into vogue in Byzantium during the eleventh-century renaissance under the influence of the Neoplatonic philosopher-statesman Michael Psellus. In 1380, Pletho went to the Turkish court at Brusa, or Adrianople, where he is reputed to have studied under the Jewish scholar Elisaeus. There Pletho presumably received training in the Muslim commentators on Aristotle, in Zoroastrianism, and in Chaldean astronomy and astrology and was encouraged by Elisaeus to further his study of Greek philosophy. Indeed, Gennadius Scholarius, who later condemned Pletho for his belief in polytheism, credits Elisaeus with leading Pletho to apostasy. About 1390, Elisaeus was burned at the stake by the Turks, probably for heterodoxy, and Pletho returned to Constantinople, from which he moved in 1393 to Mistra in the Peloponnese, near the ancient site of Sparta. It was at this administrative and cultural center of Mistra, which ranked second only to Constantinople and Thessalonica, that he spent the most important years of his life.

In 1438 Pletho appeared as adviser to the Greek delegation at the Council of Ferrara-Florence, convoked in order to effect a union between the Eastern and Western churches. An antiunionist and in some respects even anti-Christian, he took little interest in the council's proceedings. He preferred to consort with the Italian humanists, themselves fascinated by his knowledge of the works of Plato, which had for centuries been virtually unknown to the West. He left the council before the final ceremony of union to return to Mistra, where he remained until his death.

Pletho's works reveal a deep insight into Platonic philosophy and, remarkably, a devotion to Greece rather than to the crumbling Byzantine Empire. Many of his treatises aim at the revivification and restoration of Greece's ancient glory. In his famous tract "On the Differences between Plato and Aristotle," he asserts the superiority of Platonism to Aristotelianism, and his *Laws,* inspired by Plato's *Laws* and *Republic,* advocates a return to the polytheism of ancient Greece. Two memoirs based on a Platonic reconstruction of the state present a systematic plan of social and economic reform for Greece. Pletho felt that the collapse of the Byzantine Empire was due primarily to Christianity, the adoption of which had caused the alteration of the institutions of ancient Greece. In order to restore Greece to its former greatness it was necessary to foster a return to the ancient religion and to adopt a philosophy based on Platonic principles, which could serve as a guide in the process of governing. Pletho's numerous works include treatises on Zoroastrianism, Chaldean astronomy, music, history, rhetoric, the "philosophic virtues," geography, and various theological subjects. Among his theological writings is a treatise on the procession of the Holy Spirit composed in response to the Latin view presented at the Council of Ferrara-Florence.

Despite some modern opinion to the contrary, Pletho's apostasy from Christianity seems certain. Scholarius, his Aristotelian opponent, condemns him for advocating paganism in his *Laws,* and George of Trebizond quotes Pletho as asserting that a new religion, neither Christian nor Islamic but similar to that of the ancient Greeks, would sweep the world. Why then did Pletho attend the Council of Ferrara-Florence and evidently acquiesce in the act of union? Pletho was taken to the council by the Byzantine emperor John VIII, probably as a learned layman philosopher who could buttress the arguments of the theologians. Pletho's opposition to union was more on nationalistic grounds than dogmatic. As a patriot he feared that the consummation of union would precipitate a fresh Turkish attack on Constantinople. Moreover, he seemed to fear the Latinization of the Greeks, as for example in the possible suppression of

Greek in favor of Latin in the ritual of the church. Finally, as a propagandist for the formation of a "Greek" nation and a restored Hellenism (in contrast to a "Byzantine" or, more correctly, "Roman" state), he was opposed to the international papal control implicit in the union of the two churches. His acceptance of the union can then be explained only as an act of political expediency with the aim of aiding Greece, not as the result of conviction that any particular doctrinal position was correct.

Almost every Greek humanist scholar of the fifteenth century was in some way influenced by Pletho, the most notable being his pupil, Cardinal Bessarion. A great many Italian humanists were also influenced by his writings and presence at the council. Through Pletho, ancient doctrines of the Chaldeans and Pythagoreans were transmitted to the West. More important, he set in motion at Florence the passionate interest in Platonism that was soon to permeate much of western Europe. Marsilio Ficino credits Pletho with inspiring Cosimo de' Medici to found the famous Platonic Academy. By introducing into Italy (especially through Paolo dal Pozzo Toscanelli) the geographical concepts of Strabo, Pletho may have prepared the ground for the correction of Ptolemy's geographical errors. Pletho consequently helped to alter the Renaissance conception of the configuration of Earth, thus indirectly influencing Christopher Columbus, for whom Strabo was an important authority. The high esteem in which Pletho was held by the Italian humanists is attested by the transfer of his remains from Mistra to Rimini, where they were interred in the Church of St. Francis.

See also Aristotelianism; Aristotle; Byzantine Philosophy; Ficino, Marsilio; Greek Academy; Humanism; Neoplatonism; Plato; Platonism and the Platonic Tradition; Pythagoras and Pythagoreanism; Zoroastrianism.

Bibliography

TEXTS

Traité des vertus. Edited by Brigitte Tambrun-Krasker. Athens: Academy of Athens, 1987.

Contra Scholarii pro Aristotele objections. Edited by E. V. Maltese. Leipzig: Teubner, 1988.

Oracles Chaldaïques. Recension de Georges Gémiste Pléthon. Edited by Brigitte Tambrun-Krasker. Athens: Academy of Athens, 1995.

STUDIES

Athanassiadi, Polymnia. "Byzantine Commentators on the Chaldaean Oracles: Psellos and Plethon." In *Byzantine Philosophy and Its Ancient Sources,* edited by Katerina Ierodiakonou, 237–252. Oxford: Oxford University Press, 2002.

Karamanolis, George. "Plethon and Scholarios on Aristotle." In *Byzantine Philosophy and Its Ancient Sources,* edited by Katerina Ierodiakonou, 253–282. Oxford: Oxford University Press, 2002.

Lagarde, B. "Le *De Differentiis* de Pléthon d' après l'autographe de la Marcienne." *Byzantion* 43 (1973): 312–343.

Tardieu, M. "Pléthon lecteur des Oracles." *Mêtis* 2 (1987): 141–164.

Woodhouse, C. M. *Gemistos Plethon: The Last of the Hellenes.* Oxford: Clarendon Press, 1986.

Deno J. Geanakoplos (1967)

Bibliography updated by Katerina Ierodiakonou (2005)

PLOTINUS

(c. 205–270)

Plotinus, usually considered the founder of Neoplatonism, was probably born in Lykopolis, Upper Egypt, and he may have been a Hellenized Egyptian rather than a Greek. He turned to the study of philosophy when he was twenty-eight. Disappointed by several teachers in Alexandria, he was directed by a friend to Ammonius Saccas, who made a profound impression on him. Of Ammonius's teachings we know extremely little, but a promising line of investigation has been opened up in a comparison of Plotinus's doctrines with those of Origen the Christian, also a student of Ammonius. Of other students of Ammonius, Origen the Pagan and Longinus deserve special mention.

Plotinus was Ammonius's pupil for eleven years. He left Ammonius to join the expeditionary army of Emperor Gordianus III that was to march against Persia, hoping to acquire firsthand knowledge of Persian and Indian wisdom, in which he had become interested through Ammonius. When Gordianus was slain in Persia in 244, probably at the instigation of his successor, Philip the Arabian, Plotinus had to flee from the army camp—which could mean that he was politically involved in some way. Plotinus reached Antioch in his flight and from there proceeded to Rome, where he arrived in the same year. In Rome he conducted a school of philosophy and after ten years started writing. At about this time he gained influence over, or the confidence of, the new emperor, Gallienus, and it is possible that his philosophy was meant to aid the emperor in some way in his attempted rejuvenation of paganism. In any case, Plotinus asked the emperor to grant him land in order to found some kind of community, the members of which would live according to the laws (or *Laws*) of Plato.

Despite the emperor's favorable attitude, a cabal of courtiers brought the plan to nothing, indicating that they may have seen in it some political implications. However, because the contents of Plotinus's writings and some facts of his life seem to point to a complete absence of political interests, the problem of Plotinus's involvement in affairs of state is controversial. Nevertheless it is strangely coincidental that his literary activity began in the first year of Gallienus's rule. Moreover, when Plotinus died (probably from leprosy, about two years after the assassination of Gallienus), he was not in Rome but on the estate of one of his friends (of Arabic origin), and only one of his pupils, a physician, was present. These circumstances make it difficult to rule out the possibility that Plotinus had left Rome and that his pupils had all dispersed at the death of Gallienus (between March and August of 268) because he and they were afraid they would be affected by the anti-Gallienus reaction; this would again contradict a completely apolitical interpretation of Plotinus.

Plotinus's works, which were all written in the sixteen years after 253, have come down to us only in the edition by his pupil Porphyry. Porphyry arranged the works according to content into six sections called enneads because each contains nine treatises; he arbitrarily created some treatises by dissecting or combining the originals. Independent of this arrangement, he indicated when each treatise was written by assigning it to one of three periods in the life of Plotinus: before Porphyry became Plotinus's student, 253–263; while Porphyry was his student, 263–268; after Porphyry left him, 268–270. Whether Porphyry numbered the treatises within each period in strictly chronological order is open to some doubt. The presentation of Plotinus given here follows the three periods of Porphyry with only a few forward or backward references. The standard citation to Plotinus's work designates the number of the ennead first, by Roman numeral; the treatise second, by Arabic numeral; and the place of the treatise in Porphyry's chronological enumeration third, in brackets. The chapter number and, where relevant, the line number are also given in addition to the standard citation.

Contrary to the frequent attempts to present Plotinus's philosophy as a consistent whole, this presentation will stress all tensions by which the philosophy is permeated and leave it an open question whether Plotinus succeeded in reconciling them.

INFLUENCES

To understand the philosophy of Plotinus, a knowledge of some of the doctrines of Plato, Aristotle, the Neo-Pythagoreans, and the Stoics is very important.

In his dialogues Plato divided all reality into the realm of ideas (intelligibles) and the realm of sensibles, treating intelligibles alone as that which truly is (*ousia*), which implied that they are eternal and changeless (but see below). One of these ideas, the idea of the Good, he elevated above others, calling it beyond being (*epekeina ousias*). Comparable to the sun, it is the source of being and cognizability of all existents. In a lecture (or a lecture course) he seems to have identified the Good with the One.

Plato discussed the concept of the One in his dialogue *Parmenides*, ostensibly without any conclusion. In one passage he asserts hypothetically that if the One existed, it would be ineffable and unknowable. Whether this assertion was supposed to reveal the self-contradictory and, therefore, unacceptable character of the One, or on the contrary to express Plato's positive assertion as to the character of the One, is controversial. In another dialogue, *The Sophist*, Plato seems to contradict his standard doctrine concerning the unchangeable character of the ideas by ascribing life, change, and knowledge to the realm of ideas.

As to the realm of the sensible, Plato in his *Timaeus* explains the origin of the cosmos in the form of a myth—as the work of a divine artisan (demiurge) who uses an ideal cosmos as model and fashions it out of something Plato calls "receptacle" and describes as void of any qualities, after ideas have in some way "entered" this void and by so doing created rudiments of the four elements. In addition to the physical universe the demiurge also fashions a cosmic soul and the immortal part of individual souls. The cosmic soul and the individual souls consist of a mixture of the same ingredients, on which mixture the demiurge imposes a numerical and a geometrical structure.

The immaterial and substantial character of the individual souls (or at least part of them) guarantees their preexistence and postexistence (immortality). They are all subject to the law of reincarnation.

In the *Second Letter* (the authenticity of which was never doubted in antiquity, though today it finds virtually no defender), Plato, in a brief, and entirely obscure passage, seems to compress his whole philosophy into a formula reading: There are three realms, the first related to

"the king," the second to the second, the third to the third. Plotinus was convinced that Plato is here describing the three realms of the One, Intelligence, and the Soul (whereas many Christian writers were convinced that Plato must have darkly anticipated the doctrine of the Trinity).

From Aristotle, Plotinus drew an important presentation of Plato's philosophy, ostensibly different from the one professed by Plato in his dialogues. According to Aristotle, Plato had assumed a realm of mathematicals mediating between ideas and sensibles (other sources identified this realm with that of the soul). Aristotle also attributed to Plato the view that two opposite principles, the One and the Indeterminate Dyad, are the supreme principles constitutive of everything, particularly of ideas and mathematicals—a doctrine Aristotle related to a similar, equally dualistic doctrine of the Pythagoreans. Aristotle represented Plato as having identified the Indeterminate Dyad with the receptacle and as having seen in it the principle of evil.

Plotinus also adopted Aristotle's doctrine of Intelligence (nous) as superior to the rest of the soul. Aristotle implied that it alone is immortal, the rest being merely the "form" of the body, hence incapable of separate existence. Aristotle designated the supreme deity as Intelligence contemplating (that is, intelligizing) itself; the cognitive activity of the Intelligence differed from sensation in that its objects (immaterial intelligibles) are identical with the acts by which Intelligence grasps them.

Plotinus was also aware of Academic and Neo-Pythagorean attempts to take over and modify the two-opposite-principles doctrine by elevating the One above the Indeterminate Dyad (sometimes above another One, coordinated with the Dyad), which thus changed Plato's dualism into monism culminating in a transcendent One. Plotinus also knew of the syntheses of Plato's and Aristotle's philosophy attempted by some Platonists, especially of the second century CE, most prominently Albinus and Apuleius. Another influence was the strictly materialistic and immanentistic Stoic doctrine of the omnipresence of the divine in the cosmos. Finally, two Neo-Pythagorean teachers are particularly relevant as sources for Plotinus: Moderatus, who seems to have taken his cue from Plato's *Parmenides*, distinguishing a first One above being from a second and a third; and Numenius, who distinguished the supreme god from the divine artisan, creator of the cosmos.

PLOTINUS'S PHILOSOPHY FIRST PERIOD, 253–263

Plotinus subdivided Plato's realm of intelligibles into three: the One, Intelligence, and the Soul (presupposed in IV 8 [6], Ch. 6; V 4 [7], Ch. 1; VI 9 [9], Chs. 1f.; V 1 [10], Ch. 10; V 2 [11]).

THE ONE. Following what are at best hints in Plato, Plotinus developed a full-fledged theory of the One as the highest principle, or cause. Precisely because it is the principle of everything that is—and is therefore omnipresent—it is itself above being (absolutely transcendental: VI 9 [9], Ch. 4, ll. 24f., Ch. 7, ll. 28f.; V 4 [7], Ch. 1, ll. 4–8; V 2 [11], Ch. 1). Since it is above being, it is fully indetermined (qualityless), although it may be called the Good as the object of universal desire. Because it is one, it is entirely undifferentiated (without multiplicity: V 4 [7]; VI 9 [9], Ch. 3, ll. 39–45). As every act of cognition, even of self-cognition, presupposes the duality of object and subject, Plotinus repeatedly and strongly states that the One is void of any cognition and is ignorant even of itself (VI 9 [9], Ch. 6, l. 42; III 9 [13], Chs. 7, 9). He tries to mitigate this statement in some places, hesitatingly attributing to the One some kind of self-awareness (V 4 [7], Ch. 2, l. 16) or quasi awareness of its "power" to engender being (V 1 [10], Ch. 7, l. 13). In other places he distinguishes the ordinary kind of ignorance from the ignorance of the One and says that there is nothing of which the One is cognizant but that there is also nothing of which it is ignorant (VI 9 [9], Ch. 6, ll. 46–50).

INTELLIGENCE. The realm of the One is "followed" by that of Intelligence (intellect, spirit, mind—all somewhat inadequate translations of the Greek word *nous*). Here, for the first time, multiplicity appears. Roughly, this realm (hypostasis) corresponds to Plato's realm of ideas and, therefore, to that of true being. But whereas Plato's ideas are self-sufficient entities outside the Intelligence that contemplates them, Plotinus develops a doctrine of the later Platonists (perhaps originating with Antiochus of Ascalon) that interpreted ideas as thoughts of God and insists that intelligibles do not exist outside the Intelligence (V 9 [5], Chs. 7f.; III 9 [13], Ch. 1). The structure of the second hypostasis also differs from that of Plato's ideal realm in that Plotinus assumes the existence of ideas of individuals; the resulting difficulty that the infinity of individuals would demand an infinity of ideas Plotinus meets by assuming that the sensible world is, as the Stoa had it, subject to cyclical destruction and regeneration and that in each of these worlds the same indistinguish-

able individuals, for which one idea would suffice, would exist (V 7 [18], Ch. 1).

Another difference between Plato's and Plotinus's realm of ideas is that Plotinus assumed the existence of souls in this realm (IV 8 [6], Ch. 3). This doctrine creates a special problem. The ideal Socrates, unlike the soul of Socrates, must be composed of soul and body. It should follow that the soul of the empirical Socrates should be only a copy of that of the ideal Socrates, a consequence that, however, Plotinus rejects in places (V 9 [5], Ch. 13; VI 4 [22] Ch. 14) and approaches in others (III 9 [13], Ch. 3; V 2 [11], Ch. 1, l.19). Finally, Plotinus's realm of Intelligence contains even archetypal matter.

Despite all this multiplicity Intelligence remains one. In it everything is contained in everything without losing its identity, just as in mathematics every theorem contains all the others and, thus, the totality of mathematics (V 9 [5], Chs. 6, 9; IV 3 [27], Ch. 2).

Plotinus found it necessary to relate his doctrine of the One and Intelligence to the doctrine of the two opposite principles that figures in Aristotle's obscure presentation of Plato's philosophy in the *Metaphysics* (A6, 987a29ff.). In that difficult passage (the text of which may be faulty), Plato is said to have identified ideas with numbers. Plotinus also found it necessary to relate his philosophy to the doctrine identifying the soul with number, the best-known example of which was Xenocrates' definition of the soul as self-changing number. Thus, Plotinus calls the realm of Intelligence the realm of number and calls the soul number (V 1 [10], Ch. 5). But as he conceives number to be derived from the interaction of One with plurality and yet elevates the One above the realm of Intelligence (being), he seems to assign to his One a double role, a doctrine very close to the Neo-Pythagorean assumption of a double One, one superior and transcendental and another inferior, present in the realm of Intelligence, or number (V 1 [10], Ch. 5).

SOUL. Below the hypostasis of Intelligence Plotinus locates that of the Soul. Some souls remain unembodied; others "descend" into bodies. These bodies are either celestial or terrestrial. Celestial bodies offer no resistance to the soul's dwelling in them and thus these souls do not suffer from their incarnation (IV 8 [6], Ch. 2); terrestrial bodies, however, do offer resistance, and governing them may involve the soul to such an extent that it becomes alienated from Intelligence, its true home, and thus "sinks." In addition to these souls of individual bodies, Plotinus also assumes the existence of a cosmic soul (IV 8 [6], Ch. 7; III 9 [13], Ch. 3; II 2 [14], Ch. 2; I 2 [19], Ch. 1); thus, the world at large is one living organism. Probably the realm of the Soul does not consist of these individual souls alone; rather, they are all only individualizations of something we could call Soul in general (compare IV 3 [27], Ch. 4). In any case, all souls form only one Soul, and this unity implies that all souls intercommunicate by extrasensory means (IV 9 [8]).

Plotinus sometimes proves, sometimes merely assumes, not only the incorporeality, substantiality, and immortality of all the individual souls of humans, animals, and even plants (IV 7 [2], Chs. 2–8[iii], 14), but also proves or assumes reincarnation, in the course of which the same soul may pass from the body of a human into that of a beast or a plant (III 4 [15], Ch. 2). Plato's best-known proof of immortality is based on the absolute simplicity and, therefore, indissolubility of the human soul. But Plato also taught that the soul is tripartite, and perhaps in an effort to reconcile these two doctrines, Plotinus assumes that the simple and, therefore, immortal soul on its "way" to the body receives additional, lower parts as accretions. This seems to be similar to a doctrine usually associated with Gnosticism—a downward journey of the soul, during which it passes the several planetary spheres, each of which adds something to it.

EMANATION. The explanation of the relation of the three hypostases to one another leads to one of the most characteristic doctrines of Plotinus, but it is a strangely ambiguous one. This relation is described as "emanation," or "effulguration," of Intelligence from the One and of Soul from Intelligence—an emanation that, however, leaves the emanating entity undiminished (VI 9 [9], Ch. 9; V 1 [10], Chs. 3, 5–7; compare III 8 [30], Ch. 8, l. 11). The emanating entity thus remains outside of its product and yet is also present in it (VI 4 [7], Ch. 3; VI 9 [9], Ch. 7), a position sometimes described as dynamic pantheism to distinguish it from immanentist pantheism. This emanation Plotinus describes as entirely involuntary: What is full must overflow, what is mature must beget (V 4 [7], Ch. 1, ll. 26–41; V 1 [10], Ch. 6, l. 37; V 2 [11], Ch. 1, l. 8; compare IV 3 [27], Ch. 13).

Seen in this way, there is no fault, no guilt involved in emanation, nor is any justification of why the One had to become multiple necessary. On the contrary, the process deserves praise; without it the One would have remained mere potentiality, and its hidden riches would not have appeared (IV 8 [6], Ch. 5f.). But sometimes, particularly when discussing the Soul's descent, Plotinus speaks of emanation in an entirely different manner. Even the emanation of Intelligence from the One, let alone that of Soul

from Intelligence, he describes in such terms as *apostasy* and *falling away*. It is recklessness and the desire to belong to nobody but oneself that cause Intelligence to break away from the One (VI 9 [9], Ch. 5, l. 29). The Soul is motivated to break away from Intelligence by the desire to govern, which causes the Soul to become too immersed in bodies; by a craving for that which is worse; by a will to isolation (V 2 [11], Ch. 1; IV 8 [6], Ch. 4, I. 10; V 1 [10], Ch. 1). Matter emanates from Soul as the result of the Soul's wish to belong to itself (III 9 [13], Ch. 3). The "lowest" kind of Soul (the vegetative) is called the most foolhardy (V 2 [11], Ch. 2, l. 6). Thus, instead of an outflow, we should speak, rather, of a fall—with all its implications of will, guilt, necessity of punishment, and so on. These two interpretations—we shall call the former optimistic and the latter pessimistic—are difficult to reconcile.

INTELLIGENCE AND SOUL. Let us now consider the constitution of the second and third hypostases in additional detail. On the whole, Plotinus teaches that the One is in no way engaged in producing Intelligence. But sometimes he speaks as if Intelligence were the result of some kind of self-reflection of the One: The One turns to itself; this turning is vision; and this vision is Intelligence (V 1 [10], 7, l. 6—but the text is uncertain). Once more, we see that it is not easy for Plotinus to deprive the One of all self-awareness (consciousness). In any case, Intelligence is already multiple and, thus, less perfect than the One. However, the outflow from the One would not be sufficient to produce Intelligence. Rather, this flow must come to a stop—congeal, as it were. Incipient Intelligence must turn back to its source to contemplate it, and only by this act does Intelligence become fully constituted (V 2 [11], Ch. 1, l. 10). The emanation continues, and Soul emerges, again constituted by its turning toward the source, which is Intelligence (V 1 [10], Ch. 6, l. 47; V 2 [11], Ch. 1, l. 18; III 9 [13], Ch. 5). In Soul, multiplicity prevails over unity, and perfection has therefore decreased.

From Soul emanates matter, the totally indetermined (III 9 [13], Ch. 3; III 4 [15], Ch. 1). Because Plotinus tends to split the Soul into a higher, lower, and lowest kind, it is only the lowest that is the source of matter. Matter, when illuminated by the Soul, becomes the physical world, the model of which is in the realm of Intelligence (Soul thus corresponds to Plato's divine artisan, the demiurge). Thus, Plotinus's system would seem to be entirely monistic. But sometimes Plotinus speaks as if matter existed by and in itself, "waiting" to be ensouled (IV 8 [6], Ch. 6, ll. 18–20; V 2 [11], Ch. 1).

Emanation must be described in temporal terms. But, of course, it is in fact an entirely timeless event (VI [10], Ch. 6, l. 19). Once the sensible world, particularly the human body, has been constituted, the Soul in the acts of incarnation becomes submerged in the realm of the temporal. The clash between a pessimistic and an optimistic evaluation of the emanative process can now be repeated in Plotinus's evaluation of incarnation.

INCARNATION. The Platonist cannot easily ignore either the myth of the *Phaedrus*, implying that souls "fall" by some kind of failing, or the otherworldly mood of the *Phaedo*, implying that the soul should try to flee the body and be polluted as little as possible by it. But just as it is difficult for a Platonist to forget that according to the *Timaeus*, the first incarnation of the soul is the work of the divine artisan himself and, thus, a blameless event, so it is equally difficult for him to forget the myth of the *Republic*, according to which embodiment seems to be the result of some universal necessity. As a result, Plotinus had to resolve a contradiction. Sometimes he did so by trying to prove that there is no true contradiction (IV 8 [6], Ch. 5). But recognizing that such an assertion is in the last resort unsatisfactory, even when it is assumed that only part of the Soul descends (IV 7 [2], Ch. 13, l. 12; IV 8 [6], Ch. 7, l. 7), he adopted a theory that he explicitly claims as his innovation (he otherwise presents himself as an orthodox Platonist).

According to this theory, a true fall has never taken place. Actually, even when in a body, the soul still lives its original "celestial" life and remains unseparated from Intelligence. Only we are not aware of this "hidden" life of the soul; in other words, we are partly unconscious of what happens in our minds (IV 8 [6], Ch. 8). What is true of the Soul in relation to Intelligence is even truer of the relation between our embodied selves and Intelligence. Not even when present in us does Intelligence discontinue its activity (V 1 [10], Ch. 12).

Plotinus also makes an optimistic and a pessimistic evaluation of the deterioration that has taken place in the soul as a result of its incarnation. On the whole, he tries to prove that no real deterioration has taken place, but he often feels that he must find reasons why the soul should try to escape the body and return home. One of these reasons is that the body prevents the soul from exercising the activity peculiar to it (IV 8 [6], Ch. 2, l. 43), which means, of course, that some deterioration does take place.

DUALISM. There are some dualistic traits in the philosophy of Plotinus, particularly the recognition of the Inde-

terminate Dyad (as opposed to the One), to which he also refers simply as the Indeterminate (II 4 [12], Ch. 11, l. 37). Aristotle presented Plato's philosophy as a dualistic system, identifying the Indeterminate Dyad with Plato's receptacle and also with matter, in his own sense of the word; in other words according to him, Plato's ideas, being the product of the interaction of the two opposite principles, contain matter. Aristotle furthermore asserted that the Indeterminate Dyad is also the principle of evil. Plotinus is willing to recognize the Indeterminate as a second principle and to see in matter the principle of evil, but he refuses to recognize the existence of evil in the realm of Intelligence (ideas). He is thus forced to recognize the existence of two kinds of matter, one in the realm of the sensible and the result of the last emanative step, the other in the realm of Intelligence ("intelligible matter"), which does not have some of the properties usually associated with matter—specifically, it is not evil. He justifies this by the assumption that everything, including matter in the physical world, must have its archetype in the realm of Intelligence (II 4 [12], Chs. 2f., 11, 14). Whether the assumption of intelligible matter can be reconciled with monism appears dubious; its "origin" is never made clear by Plotinus.

As to matter in the realm of the sensible, it is sheer indeterminacy, incorporeal, and, thus, different from the Stoic conception of matter (II 4 [12], Chs. 1, 4, 9, 10). It remains as unaffected by the ideas (or "ratios," *logoi*, by which Stoic term Plotinus often designates ideas as present in the soul qua formative powers) as the mirror is unaffected by what it reflects. Precisely because this matter is indeterminate, it is evil (II 4 [12], Ch. 16, l. 19), which means that evil is not something positive, but sheer privation.

There is a strange parallelism between matter and the One, because both are entirely indeterminate. Therefore, they both elude ordinary concepts, and Plotinus faces the question of what it means to know them. As far as matter is concerned, Plotinus likens it sometimes to darkness, and the mental act by which we grasp it to "unthinking thinking," or the soul's reduction to indefiniteness (II 4 [12], Chs. 6, 11)—concepts reminding us of Plato's pseudo thinking (*nothos logismos*), declared by him to be the appropriate way to think the receptacle.

KNOWLEDGE OF THE ONE. But much more important for Plotinus is the problem how the One, in spite of its being ineffable, can be known. In the pseudo(?)-Platonic *Epinomis* (992B), the author insists that in order to know the One (whatever "knowledge" means here), the soul must itself become one; the Platonic *Letters* also seem to teach some kind of suprarational insight. Perhaps starting from passages such as these and also from passages in Aristotle and Theophrastus in which some kind of infallible knowledge of certain objects is described as a kind of touching (*thinganein*), Plotinus asserts that to "know" the One means to become one with it, which the soul can accomplish only by becoming as simple or as "alone" as the One. In the moment of such a union the soul has become God or, rather, is God; the soul has reascended to its original source (VI 9 [9], Ch. 9f.). Among the terms Plotinus uses to describe this condition are *ecstasy, simplicity, self-surrender, touching*, and *flight of the alone to the alone* (VI 9 [9], Chs. 3, 11). This ecstasy—repeatedly experienced by Plotinus himself—is undoubtedly the climactic moment of man's life. It is not expressible in words (compare Plato, *Epistle VII*, 341D); only he who has experienced it knows what it means to be ravished away and full of God.

For this reascent man prepares himself by the acquisition of all the perfections (virtues, *aretai*). However, each of these perfections acquires different meanings according to the level on which man's spiritual life takes place—thus, there is a social fairness, above it another kind of fairness, and so on. Man also prepares himself by the exercise of dialectics (I 2 [19]; I 3 [20]). The preliminary stages of achievement Plotinus calls "becoming Godlike" (I 6 [1], Ch. 8), a condition often described by Platonists preceding Plotinus as the ultimate goal of Plato's philosophy.

FREE WILL AND DEMONOLOGY. Among the other topics treated in this period, Plotinus's defense of the freedom of the will—only "reasonable" souls are free; others are subject to fate, εἱμαρμένη (III 1 [3])—and his demonology deserve special mention. In regard to demonology Plotinus tries to steer a middle course between two theories, one identifying demons with the supreme parts of our soul, and the other assuming the existence of demons as extrapsychical beings (III 4 [15]).

SECOND PERIOD, 263–268

POLEMICS. More than two-fifths of Plotinus's total literary output was produced during the brief period between 263 and 268, when Porphyry was studying with Plotinus. Perhaps Porphyry's presence worked as a powerful stimulus. A considerable part of the output of this period is devoted to polemics with other schools, notably on the doctrine of categories and against Gnosticism.

Categories. Plotinus rejects both the Aristotelian and the Stoic versions of this doctrine, adhering to the principle that there can be no categories common to the realms of the sensible and the intelligible. In application to the realm of the sensible he corrects and modifies Aristotle's categories; to the realm of Intelligence he tries to apply Plato's five genera—being, identity, diversity, rest, and change (VI 1–3 [42–44]).

Ideal numbers. Aristotle presented Plato as professing the existence of ideal numbers (twoness, threeness, and so on, as distinguished from ordinary numbers—two, three, and so on). And he devoted much effort to the criticism of the theory of ideal numbers. Plotinus defends the theory of ideal numbers—which differ from nonideal numbers in that they do not consist of addible unities and are therefore not addible themselves (V 5 [32], Ch. 4)—and, objecting to any nominalist or abstractionist theory of numbers, attributes to them subsistence. Specifically, after having divided the realm of Intelligence into three layers—Being, Intelligence (in a restricted sense of the word), and the original Living Being—he assigns ideal numbers to the uppermost layer and explains that only because of their existence can Being divide itself into beings (VI 6 [34]), Chs. 8, 16). In this context he also introduces a peculiar concept of infinity: The truly infinite is a thing that has no limits imposed on it from without but only from within (VI 6 [34], Chs. 17f., but compare V 5 [32], Ch. 4).

Polemic against Gnosticism. Of all the polemics of Plotinus, the most significant is the one against Gnosticism. One could say that when facing Gnostic pessimism point-blank, Plotinus overcompensates for the pessimistic and Gnostic strand present in himself and responds with an almost unlimited optimism. The fundamental mood underlying Gnosticism is alienation from a hostile world, and Gnosticism undertakes to explain this mood and to open the road to escape from the world. The explanation is in the form of a history of the origin of the visible cosmos; according to Gnosticism, this cosmos is the result of the activity of an evil god sometimes identified with the Creator-God of the Old Testament or with Plato's divine artisan. This evil god is only the last in a succession of beings. The manner in which this succession takes place consists in a number of voluntary acts by which divinities of an ever lower order originate. The relation between these deities is often personal, based on such traits as curiosity, oblivion, daring, ambition. Man, as he exists in this evil world, contains in himself a spark of what was his original, divine substance, now imprisoned in his body owing to the scheming of the evil god. At a certain moment a messenger-savior in some way breaks the power of the evil god and makes it possible for those who hear the whole story (acquire gnosis) to regain their original standing and free themselves from the tyranny of the evil god.

Plotinus treats Gnosticism as a strictly philosophic system. He simply compares its doctrines with his own and with those of Plato; its salvationary aspects are of little interest to him (compare III 2 [47], Ch. 9). In the succession of divine beings he sees only a superfluous multiplication of the three hypostases of his own system (compare V 5 [32], Chs. 1f.). To the cosmic drama that results in the creation of the visible cosmos he opposes his view of a totally undramatic, unconscious emanation, a product of necessity without arbitrariness and, contradicting even Plato's *Timaeus* (40B–45A), without planning (V 8 [31], Ch. 7) and, therefore, entirely blameless. The cosmos, product of the activities of the Soul (or Intelligence or both), he considers to be beautiful. Whereas Gnosticism sees the visible universe filled with spirits inimical to man, most outstanding among them being the rulers of the celestial bodies (planets), Plotinus sees in these spirits powers related to man in brotherly fashion. What is true in Gnosticism can, according to him, be found in Plato. The Gnostic objection that Plato did not penetrate the mysteries of the intelligible world Plotinus considers ridiculously presumptuous (II 9 [33]; compare V 8 [31], Ch. 8).

PROBLEMS. In the second period Plotinus was also concerned with the problems inherent in his own system, especially with the relation between the intelligible world and the sensible world and with the structure of the intelligible world.

The One. First, Plotinus tries to elucidate the nature of the One still further. He does this particularly in the context of a discussion concerning the nature of human freedom, in which he also asks whether the One should be considered as a necessary being or as a free one (*ens necessarium* or *ens liberum*)—in theistic terms, whether God must exist or has freely chosen to exist. In what is perhaps his most profound theological discussion, Plotinus tries to establish the concept of the One as Lord of itself and thus not having to serve even itself, so that in the One freedom and necessity coincide (VI 8 [39], Chs. 7–21). And without any vacillation he excludes any kind of consciousness from the One (V 6 [24], Chs. 2, 4f.).

Intelligence and Soul. As far as Intelligence is concerned, Plotinus reiterates his doctrine that it contains ideas within itself (V 5 [32], Chs. 1f.), and he again tries

to explain how, in spite of being one, it still contains multiplicity (VI 4 [22], Ch. 4; VI 5 [23], Ch. 6). With regard to souls Plotinus tries to explain how they can remain distinct from one another although they all are only one soul (VI 4 [22], Ch. 6; IV 3 [27], Chs. 1–8; compare IV 9 [8], Ch. 5).

Both Intelligence and Soul are supposed to be present in the sensible world and, therefore, present in what is extended, although they themselves are not extended. Starting from the famous discussion in Plato's *Parmenides* (131B), in which the attempt is made to explain how one idea can be present in many particulars, Plotinus tries to show that just because Intelligence and Soul are not extended, they can be omnipresent and ubiquitous in what is extended (VI 4 [22], especially VI 5 [23], Ch. 11). And also in this context he tries to establish the concept of differentiated unity (VI 4 [22], Ch. 4), that is, the noncontradictory character of "one" and "many."

Intelligence, Soul, change. Probably the most formidable difficulty facing Plotinus is the result of his theory treating Intelligence and Soul as metaphysical principles on the one hand and as present in man on the other (that is, as both transcendent and immanent) and, therefore, in some way engaged in mental life, particularly in sensing and remembering. As metaphysical principles—that is, members of the realm of the intelligible—Intelligence and Soul should be unchangeable, whereas in man they seem to be involved in change. From this difficulty Plotinus tries to extricate himself in many ways, of which two will be presented.

On the one hand he keeps even the human soul away as much as possible from the processes of sensing, remembering, desiring, experiencing pleasure and pain, and so on (III 6 [26], Ch. 1–5). Sometimes he insists that the soul simply notices all these processes without being affected by what it perceives (IV 6 [41]; IV 4 [28], Ch. 19). Sometimes he insists that it is not the soul itself but only some trace of it which is engaged in these activities (IV 4 [28], Chs. 18f.; compare VI 4 [22], Ch. 15, l. 15), and this ties in with the theory that the soul did not really—or not in its entirety—descend (VI 4 [22], Ch. 16). Sometimes he introduces the concept of a double soul, a higher and a lower, with only the lower being changeable. This doubling of the soul Plotinus carries to such extremes that he assumes two imaginative faculties and two faculties of memory, each belonging to its respective soul and each remembering in a different manner and different events. This is particularly the case after man's death; the higher soul no longer remembers anything it experienced while in the body, whereas the lower soul still remembers (IV 3 [27], Chs. 25–32; IV 4 [28], Ch. 1, l. 5). Sometimes he suggests that all the mental activities involving change happen not to the soul but to the composite of soul and body (IV 4 [28], Ch. 17), leaving undecided how anything can affect a whole without affecting the part that belongs to it.

On the other hand, when it comes to Intelligence and Soul as metaphysical principles (and even to the world soul and astral souls), Plotinus disallows them memory entirely (IV 4 [28], Chs. 6–17). As to sensing, he distinguishes two kinds, one serving such practical purposes as self-preservation, the other purely theoretical; it is only the theoretical kind that he ascribes to metaphysical entities, the implication obviously being that this kind of sensation does not cause any change in the perceiver (IV 4 [28], Ch. 24). Why they should still be called Intelligence and Soul remains somewhat unclear. Perhaps the most striking example of the real effects of the Soul's falling away from Intelligence (despite everything said by Plotinus to minimize these effects) is that the cosmic soul, as it falls away, engenders time because of an inability to contemplate the totality of Intelligence simultaneously (III 7 [45], Ch. 11).

Ethics. The difficulties created for the explanation of the cognitive aspects of man's mental life without the assumption of a real change (passibility) of the soul return with even greater significance in the field of ethics. If there is no actual fall of the soul and if no deterioration of its nature has taken place as the result of incarnation (III 6 [26], Ch. 5), why is purifying the soul necessary? Yet the concept of purification plays a central role in the ethics of Plotinus (compare I 6 [1]; I 2 [19]); he even describes the perfections—wisdom, self-control, justice, courage—as purifications. Plotinus tries to help himself by a metaphor: The soul is merely covered with mud, which, however, has never penetrated it. According to another explanation, what the soul has acquired because of its fall is nothingness, and all it has to do, therefore, is to get rid of nothing (VI 5 [23], Ch. 12, ll. 16–23).

Cosmic sympathy. The insistence that memory and sensation, in their ordinary senses, are absent from the realm of Intelligence and even from that of the celestial sphere Plotinus explains with his theory that the universe is one animated organism. The sympathy existing among parts of one organism make memory and sensation superfluous, since the mutual affection need not be perceived. This leads to characteristic explanations of the efficacy of magic, prayers, and astrology. All these activities (and prophecies) are made possible by the fact that each part of the universe affects the others and is affected by them, not by mechanical causation nor by influencing

the will of deities—particularly stars—but exclusively by mutual sympathy (IV 3 [27], Ch. 11; IV 4 [28], Chs. 40f.). In this doctrine of sympathy many scholars see the influence of the Stoa, particularly Posidonius, on Plotinus.

Matter. As to matter, Plotinus in the writings of this period—with less ambiguity than in other periods—characterizes it as the result of the last step of the emanative process, thus fully preserving the monistic character of his system (II 5 [25], Ch. 5; compare I 8 [51], Ch. 7). Some other problems discussed by Plotinus are distinctly occasional pieces and somewhat peripheral with regard to the system. Thus, we find a theory of vision, explained by sympathy (IV 5 [29]; II 8 [35]); a discussion of the Stoic concept of the complete interpenetration of bodies (II 7 [37]); a cosmology without the assumption of ether (II 1 [40]).

THIRD PERIOD, 268–270

As is to be expected, some earlier themes recur in the third period. In fact, one of the essays of the third period (V 3 [49]) contains what is perhaps the most comprehensive presentation of the basic tenets of Plotinus's philosophy. Plotinus proves that there must be a One preceding all multiplicity and that this One must be ineffable (V 3 [49], Chs. 12f., 17). To explain its presence in us and the fact that we know about it although we do not know it, he says that those full of and possessed by the divine also feel that something greater than themselves is present in them, although they cannot say what it is (V 3 [49], Ch. 14). Once more facing the problem of how the One, which is absolutely simple, can be the source of multiplicity, Plotinus is on the verge of admitting that the One is at least potentially (though it is a potentiality sui generis) many (V 3 [49], Chs. 15f.; compare VI 5 [23], Ch. 9). The same essay contains what is probably the most detailed and the most impressive description of the upward journey of the soul to reach the goal of ecstatic union, described by the formula "through light light" (V 3 [49], Ch. 17, ll. 28–37; compare V 5 [32], Chs. 4–9). As advice on how to achieve this union, Plotinus says "strip yourself of everything" (V 3 [49], 17, l. 38). Furthermore, Plotinus still feels he must prove that ideas are not external to Intelligence (V 3 [49], Chs. 5–13).

On the whole, the writings of Plotinus's last period are dominated by two themes. The first concerns theodicy, the origin and justification of evil, and the second asks what man's true self is.

THEODICY. To explain the origin of evil, Plotinus tries to reconcile the view that matter, though void of any quality and actually only deficiency, is still evil in some sense of the word and is the source of all evil (I 8 [51], Chs. 8, 10). In so doing, he sometimes comes dangerously close to the Gnostic theory that matter imprisons the soul (I 8 [51], Ch. 14, ll. 48–50) and to a completely dualistic system (I 8 [51], Ch. 6, l. 33). Nevertheless, his optimism is particularly strong in this period; he has high praise for the beauty of the visible cosmos (III 2 [47], Ch. 12, l. 4), and rejects the idea of an evil creator of the cosmos (III 2 [47], Ch. 1). His theodicy is a blend of Platonic arguments, drawn especially from Book X of the *Laws*, and Stoic arguments. Perfection of the whole demands imperfection of the parts (III 2 [47], Chs. 11, 17; III 3 [48]) and the existence of evil (I 8 [51], Chs. 8–15). At the same time he minimizes the importance of evil by insisting that it exists only for the wicked one (III 2 [47], Ch. 6). Furthermore, he points out that the cosmic order rewards and punishes everybody according to his merits and assigns each one an appropriate place, thus making for a completely harmonious whole (III 2 [47], Ch. 4). Ultimately, his theodicy is based on convictions characteristic of most theodicies—that to designate a particular as evil is to lose sight of the whole, that everything participates in the good as far as it can, and that evil is only absence of the good (III 2 [47], Chs. 3, 5; I 8 [51], Chs. 1–5).

Providence. Closely connected with the problem of theodicy is the problem of providence. Plotinus insists on the all-pervasive character of providence, thus rejecting Aristotle's dichotomy of the universe into a sublunar sphere dominated by necessity and a supralunar world to which providence is restricted. He replaces Aristotle's distinction by the dichotomy of good and wicked men; only the wicked are subject to necessity (III 2 [47], Ch. 9; compare III 1 [3], Chs. 8–10). But this providence is entirely impersonal (compare VI 7 [38], Ch. 1) and actually coincides with the order of the universe.

TRUE SELF AND HAPPINESS. The second major theme of Plotinus's last period is that of ascertaining what man's true self is—that is, of ultimately obeying the divine command "Know thyself." Attendant subproblems are the explanations of wherein man's true happiness consists and of the concept of self-knowledge. It is extremely difficult for Plotinus to give a consistent account of what constitutes man's true self. He cannot simply identify it with Intelligence or Soul (as he did in IV 7 [1], Ch. 1, l. 24 or in I 4 [46], Chs. 8–16, where it is identified with the "higher" soul), precisely because both, in their character of metaphysical entities, remain transcendent; however, he rejects the idea that man is truly the composite of soul and body (I 4 [46], Ch. 14, l. 1) because this would grant

the body too much importance. One of the solutions favored by Plotinus is that Intelligence is man's true self, but only if and when he succeeds in identifying himself with it. On the other hand, no such identification is actually necessary, because Intelligence is always in and with us even though we are not aware of it. (*Mutatis mutandis* this can also be applied to the relation of man and whatever is to be conceived the highest divinity: compare VI 5 [23], Ch. 12). Once more the concept of the unconscious plays a decisive role in the system of Plotinus (I 4 [46], Chs. 9f.; V 3 [49], Chs. 3f.). All this ties in with the idea that self-knowledge occurs only when the subject, the act, and the object of knowledge coincide—which takes place only on the level of Intelligence—whereas neither man as a whole nor Soul can possess full self-knowledge (V 3 [49], Chs. 3, 6). The One is, of course, above any kind of self-knowledge (V 3 [49], Chs. 10–13).

The thesis that only Intelligence is man's true self (if and when he makes full use of it) serves also as a basis for a discussion of the problem of man's happiness. If by "man" we mean the composite of body and soul, man cannot experience happiness, nor can he if he is body alone. However, if by "man" we mean the true self, it is obvious that happiness consists in the exercise of Intelligence—that is, in contemplation. But as the activity of Intelligence is uninterrupted (here in the argument Plotinus switches from Intelligence as immanent to transcendent Intelligence; see I 1 [53], Ch. 13, l. 7) man is actually always happy, although he may remain unconscious of it (I 4 [46], Chs. 4, 9, 13–16). Why this should apply only to the sage remains unclear.

The formidable problem of how the soul, the essence of which is unchangeability, can ever become evil also vexed Plotinus to the end (compare I 8 [51], Ch. 4, 12, 15). In the work of his last period he explains that as the soul at its descent acquires additional parts, evil resides only in them. Thus, the ethical task of man is not so much to separate the soul from the body as it is to separate it from these adventitious parts (I 1 [53], Ch. 12, l. 18). In this context the problem of who is the subject of punishments in afterlife also emerges; Plotinus answers that it is that "composite" soul (I 1 [53], Ch. 12). Why we should call soul an entity that is or can become evil, "suffer" punishment, and so on, after Soul has been presented as belonging to the realm of the unchangeable, remains unanswered; so do virtually all questions resulting from the dual character of Intelligence and Soul as metaphysical (transcendental) entities on the one hand and human (immanent) entities on the other.

There is almost something providential in the fact that the very last of Plotinus's essays, written at a time when death was approaching him, reasserts that all things participate in the One (the Good) and discusses the question of how to reconcile the two theses that life is good and yet death no evil, though it deprives us of something good (I 7 [54], Ch. 3). The battle between the pessimistic and the optimistic strands in Plotinus continued to the very end of his activity. Optimism ultimately won: Life is good—though not for the wicked one; death is good, because it will permit the soul to live an unhampered life.

See also Alcinous; Antiochus of Ascalon; Aristotle; Categories; Cosmos; Emanationism; Evil, The Problem of; Gnosticism; Good, The; Neoplatonism; Nous; Numenius of Apamea; Origen; Plato; Platonism and the Platonic Tradition; Porphyry; Posidonius; Pythagoras and Pythagoreanism; Socrates; Stoicism.

Bibliography

WORKS BY PLOTINUS

The standard edition of the Greek text of the *Enneads* and of Porphyry's *Life of Plotinus* is by Paul Henry and Hans-Rudolf Schwyzer, *Plotini Opera*, in 3 vols. (Oxford: Clarendon Press, 1964–1982) (*editio minor* with concise apparatus). For full critical apparatus and an English translation of the important Arabic tradition see their *editio maior*, in 3 vols. (Paris: Desclée de Brouwer, 1951–1973).

Modern translations (some including ancient text and/or valuable notes) are available in English by A. H. Armstrong, in 7 vols. (Cambridge, MA, and London: Harvard University Press and Heinemann, 1966–1988); in French by É Bréhier, in 6 vols. (Paris: Belles Lettres, 1924–1938); in German by R. Harder, R. Beutler, and W. Theiler, in 5 vols. (Hamburg: F. Meiner, 1956–1962); in Spanish by J. Igal, in 3 vols. (Madrid: Gredos, 1982–1998); and in Italian by M. Casaglia, C. Guidelli, A. Linguiti, and F. Moriani, in 2 vols. (Torino: Unione Tipografico-Editrice, 1997). A new French translation of individual treatises with extensive commentary, under the direction of Pierre Hadot (Paris: Cerf, 1988–), comprises, so far, 8 vols.

WORKS ON PLOTINUS

A complete lexicon for Plotinus' writings is J. H. Sleeman and Gilbert Pollet, *Lexicon Plotinianum* (Leiden: Brill, 1980). For a comprehensive bibliography, see Richard Dufour, *Plotinus: A Bibliography 1950–2000* (Leiden: Brill [*Phronesis*, vol. 46, No. 3], 2001). This is updated in: http://rdufour.free.fr/BibPlotin/Anglais/Biblio.html.

GENERAL WORKS

Gerson, Lloyd P. *Plotinus*. London: Routledge, 1994.

Hadot, Pierre. *Plotinus, or the Simplicity of Vision*. Translated by M. Chase, with an introduction by A. I. Davidson. Chicago, IL: University of Chicago Press, 1993.

O'Meara, Dominic J. *Plotinus: An Introduction to the* Enneads. Oxford: Clarendon Press, 1993.

Rist, John M. *Plotinus: The Road to Reality*. Cambridge, NY: Cambridge University Press, 1967.

Schwyzer, Hans-Rudolf. "Plotinos." In *Realencyklopädie der classischen Altertumswissenschaft*, edited by A. Pauly and G. Wissowa. Vol. 21 (1951): 471–592, with supplement in *SupplementBand*, 15 (1978): 310–328.

COMMENTARIES ON INDIVIDUAL TREATISES

Atkinson, Michael. Ennead *V, 1, On the Three Principal Hypostases: A Commentary with Translation*. Oxford: Oxford University Press, 1983.

Fleet, Barrie. Ennead *III, 6, On the Impassivity of the Bodyless*. Oxford: Clarendon Press, 1995.

Meijer, P. A. *Plotinus On the Good or the One (*Enneads *VI, 9): An Analytical Commentary*. Amsterdam: Gieben, 1992.

Tornau, Christian. *Plotin:* Enneaden *VI, 4-5 (22–23): Ein Kommentar*. Stuttgart: Teubner, 1998.

STUDIES OF DETAIL

Adamson, Peter. *The Arabic Plotinus: A Philosophical Study of the* Theology of Aristotle. London: Duckworth, 2002.

Armstrong, A. H. *Plotinian and Christian Studies*. London: Variorum Reprints, 1979.

Blumenthal, Henry J. *Plotinus' Psychology: His Doctrines of the Embodied Soul*. The Hague: Nijhoff, 1971.

Bussanich, John R. *The One and its Relation to Intellect in Plotinus: A Commentary on Selected Texts*. Leiden: Brill, 1988.

Corrigan, Kevin. "Body's Approach to Soul. An Examination of a Recurrent Theme in the *Enneads*." *Dionysius* 9 (1985): 37–52.

Corrigan, Kevin. *Reading Plotinus: A Practical Introduction to Neoplatonism*. West Lafayette IN: Purdue University Press, 2005.

Dillon, John M. "Plotinus and the Transcendental Imagination." In his *Religious Imagination*. Edinburgh: Edinburgh University Press, 1986.

Dillon, John M. "Plotinus, the first Cartesian?" *Hermathena* 149 (1990): 19–31.

Emilsson, Eyjólfur Kjalar. *Plotinus on Sense-Perception: A Philosophical Study*. Cambridge, U.K.: Cambridge University Press, 1988.

Emilsson, Eyjólfur Kjalar. "Plotinus on the Objects of Thought." *Archiv für Geschichte der Philosophie* 77 (1995): 21–41.

Ferwerda, R. *La signification des images et des métaphores dans la pensée de Plotin*. Groningen: J. B. Wolters, 1965.

Gerson, Lloyd P., ed. *The Cambridge Companion to Plotinus*. Cambridge, NY: Cambridge University Press, 1996.

Hadot, Pierre. "Les niveaux de la conscience dans les états mystiques selon Plotin." *Journal de Psychologie Normale et Pathologique* 77 (1980): 243–266.

Igal, Jesus. "The Gnostics and 'The Ancient Philosophy' in Porphyry and Plotinus." In *Neoplatonism and Early Christian Thought: Essays in Honour of A. H. Armstrong*, edited by H. J. Blumenthal and R. A. Markus. London: Variorum Publications, 1981.

Jackson, B. Darrell. "Plotinus and the *Parmenides.*" *Journal of the History of Philosophy* 5 (1967): 315–327.

Kalligas, Paul. "Forms of Individuals in Plotinus: a Re-examination." *Phronesis* 42 (1997): 206–227.

McCumber, J. "*Anamnesis* as Memory of Intelligibles in Plotinus." *Archiv für Geschichte der Philosophie* 60 (1978): 160–167.

O'Brien, Denis. *Plotinus on the Origin of Matter: An Exercise in the Interpretation of the* Enneads. Napoli: Bibliopolis, 1991.

O'Daly, G. J. P. *Plotinus' Philosophy of the Self*. Dublin: Irish University Press, 1973.

Rist, John M. "Monism: Plotinus and Some Predecessors." *Harvard Studies in Classical Philology* 69 (1965): 329–344.

Schibli, Hermann S. "Apprehending our Happiness. *Antilepsis* and the Middle Soul in Plotinus, *Ennead* I, 4, 10." *Phronesis* 34 (1989): 205–219.

Schniewind, Alexandrine. *L'éthique du sage chez Plotin: Le paradigme du* spoudaios. Paris: J. Vrin, 2003.

Schroeder, Frederic M. *Form and Transformation: A Study in the Philosophy of Plotinus*. Montréal, London: McGill-Queen's University Press, 1992.

Sorabji, Richard. "Myths about Non-propositional Thought." In *Language and Logos: Studies in Ancient Greek Philosophy Presented to G. E. L. Owen*, edited by M. Schofield and M. C. Nussbaum. Cambridge, U.K.: Cambridge University Press, 1982.

Strange, Steven Keith. "Plotinus, Porphyry and the Neoplatonic Interpretation of the *Categories*." In *Aufstieg und Niedergang der römischen Welt: Geschichte und Kultur Roms im Spiegel der neueren Forschung*, edited by Wolfgang Haase. Vol. II, 36.2. Berlin: De Gruyter, 1987.

Strange, Steven Keith. "Plotinus on the Nature of Eternity and Time." In *Aristotle in Late Antiquity*, edited by L. P. Schrenk. Washington DC: The Catholic University of America Press, 1994.

Philip Merlan (1967)
Bibliography updated by Paul Kalligas (2005)

PLOTINUS [ADDENDUM]

What is it, Plotinus asks (Plotinus 1956), that lures the eye toward a beautiful sight and that draws the ear to a beautiful sound? It is the thrill that the soul feels in sensing its affinity with the noble being that manifests itself in those beautiful sights and sounds. Material things become beautiful by sharing in Form and thus in Unity. This applies to the productions of artists as much as to the beauties of nature. Thus the true objects of artistic imitation are Form and Unity; and so the artist is always entitled to "add where nature is lacking" (Plotinus 1956). Beauty is also found in noble conduct, in excellent laws, and in human virtue. The virtuous soul acquires Beauty and becomes godlike by purifying itself from evil. Thus Beauty in general has a metaphysical significance through its relation to Form, the One, and to the divine. In Plotinus's eyes, Beauty's significance is not only metaphysical but quasi-religious, not only because of its relationship to the divine, but also because Beauty is what draws the soul

onwards in its ascent to the suprasensible world whence it came (Plotinus 1956).

See also Beauty.

Bibliography

Beardsley, Monroe C. *Aesthetics from Classical Greece to the Present: A Short History*. New York: Macmillan, 1966.

Plotinus. *The Enneads*. Translated by Stephen McKenna, revised by B. S. Page, with a foreword by E R Dodds and an introduction by P Henry. 2nd ed. London: Faber and Faber, 1956.

Paul Thom (2005)

PLOUCQUET, GOTTFRIED
(1716–1790)

Gottfried Ploucquet, the German philosopher and logician, studied philosophy and theology at Tübingen, experiencing both Wolffian and Pietist influences. After serving as a pastor, he was professor of logic and metaphysics at Tübingen from 1750 to 1782. He was elected to the Berlin Academy in 1748. Ploucquet was one of the few logicians between Gottfried Wilhelm Leibniz and George Boole to study a symbolic calculus. In metaphysics, despite his Wolffian training, he developed a quite personal position inspired by René Descartes and Nicolas Malebranche and aimed at revising Leibnizianism on a theological basis.

Ploucquet regarded the problems of theology, cosmology, and psychology as inextricably intertwined, with theology as the predominant discipline. There were some variations in Ploucquet's doctrines, but typically he held that a monad is a spiritual substance, and that even being is spiritual. Spiritual substances and material things can interact because God represents both and connects them. Human perceptions are an effect of God's "real vision." Spiritual and material things are both real because God represents them; material things are real in a further sense, as *phaenomena substantiata,* insofar as God represents them as real. This divine representation is the cause of the real existence of things; but we perceive only an appearance of this real existence. Ploucquet showed, by an examination of the logical difficulties of the concept of infinity, that space and time cannot exist outside of human representation.

Ploucquet's philosophy was basically a pronounced metaphysical subjectivism and phenomenalism. But in order to escape the consequent idealism of this position, Ploucquet reintroduced a variety of realism based on God. Ploucquet's was one of the most significant attempts before Immanuel Kant to develop a phenomenalism that asserted the real existence of things but denied (contrary to Leibnizian and Wolffian phenomenalism) that we can know such things on the basis of their appearances.

See also Descartes, René; Leibniz, Gottfried Wilhelm; Logic, History of: Precursors of Modern Logic; Malebranche, Nicolas; Phenomenalism; Wolff, Christian.

Bibliography

PRINCIPAL WORKS BY PLOUCQUET

Primaria monadalogiae capita. Berlin, 1747.

Principia de substantiis et phaenomenis. Frankfurt and Leipzig, 1752.

Fundamenta philosophiae speculativae. Tübingen, 1759.

Institutiones philosophiae theoreticae. Stuttgart, 1772.

Sammlung der schriften, welche den logischen Calkul des Herrn Prof. Ploucquets betreffen, mit neuen Zusätzen. Edited by F. A. Böck. Tübingen, 1773.

Elementa philosophiae contemplativae. Stuttgart, 1778.

Commentationes philosophicae selectiores. Utrecht, 1781.

WORKS ON PLOUCQUET

Aner, Karl. *Gottfried Ploucquets Leben und Lehre.* Bonn, 1909.

Bornstein, Paul. *Gottfried Ploucquets Erkenntnistheorie und Metaphysik.* Potsdam: Buchdr. von A.W. Hayn's Erben, 1898.

Giorgio Tonelli (1967)

PLURALISM

Pragmatism and Continental hermeneutics combined to produce a decided turn toward forms of "pluralism" in twentieth-century philosophy (Geyer 1993, B. Singer 1990). This has led to the rejection of any one favored epistemological method (e.g., the scientific method, scriptural exegesis, introspection) and any one favored basis for the reconstruction of reality (e.g., mind, matter). Neopragmatists propose to replace the notion of truth with notions such as "fitting," "useful," and "warranted." Given that what is "fitting" is relative to the problem being faced and the means at one's disposal, we are left with the possibility of a plurality of ways of conceiving the world and of achieving our aims within it.

Moral pluralism opposes the monistic view that there is any one method of determining what is morally right (e.g., the utilitarian calculus or Kantian universalizability), and it also opposes the relativistic view that all things have value only with respect to a particular cul-

tural context. Pluralists insist that a good life typically involves the desire, not for one, but for many kinds of "goods," often of incommensurable value; moreover, the realization of certain "goods" may conflict with and even preclude the realization of others. As such, pluralists believe that moral conflicts are inevitable and that there are not one but many alternative ways of resolving such conflicts (Kekes 1993). The trend toward pluralism has also been influenced by our growing awareness of different cultures with nonequivalent conceptions of reality and "the good life."

The modern nation-state has evolved beyond the belief that it manifests the cultural orientation of a single "race," usually its majority. The reality is that every nation is composed of numerous groups with different cultural orientations. And the state is considered the primary guarantor that minority views will be presented, respected, and given a voice in determining policy (Guttman 1993). The rejection of the view that a Eurocentric male-dominated culture is the norm to be achieved universally has led to the demand that the cultures of non-Europeans, women, and minorities be recognized and granted equal voice (Taylor 1992). In this way pluralism is considered by many to be an essential part of the liberal democratic state, and this has manifested itself in terms of educational policy as the rejection of monoculturalism and the demand for a multicultural orientation.

One form of multiculturalism has focused on the need of suppressed groups to have their cultures recognized. Such a demand for recognition may motivate certain proposals—for example, to replace a Eurocentric focus with an Afrocentric focus or a male-centered orientation with a feminist-centered orientation. Some argue that because of the past harms inflicted upon such groups, ostensibly because they were different, they are justified in embracing those differences in order to cleanse them of the negative valuations imposed by the hegemonic culture. It is right for such groups to adopt a separatist posture if this is the best means of achieving a redefinition of themselves that is positive and self-affirming (Young 1990). Where members of the hegemonic culture have inflicted unjust harms on members of an oppressed group, some argue that the oppressed group has the right to cultural restitution. The domination of culture A by culture B may not be the result of culture A's not offering viable options; rather, it may be the result of unjust injuries and harms visited on culture A by culture B. In such cases groups sharing culture A have a right to "moral deference," affirmative action, and the preservation of their culture (Mosley 1990, Nickel 1994, Thomas 1992–1993).

Many have been concerned that multiculturalism might degenerate into a bedlam of different groups, each espousing its own brand of cultural authenticity. Critics argue that this would amount to merely replacing one culture's hegemony with another culture's hegemony. Multiculturalism in this sense would fail to reflect the pluralist maxim that no orientation is "fitting" for every situation and that for a given end there may be several equally "fitting" means (West 1993, Yates 1992).

An alternative form of multiculturalism, closer to pluralism, emphasizes the importance of diversity and cross-cultural communication. On this view the more cultural orientations there are for consideration, the better the likelihood of finding or constructing a "fitting" adaptation to some current problem (Rorty 1992). For this reason every culture should be allowed the opportunity of articulating itself to the public at large and of thereby influencing the manner in which individuals construct their character.

Pluralism does not end with the insistence on an equal voice for every culture but extends itself to the view that different biological species often have interests that may conflict with the interests of human beings. Some have argued that, just as racism and sexism accord special preference to white males and victimize women and non-Europeans, so speciesism accords special preference to the interests of human beings and unjustly victimizes nonhuman species (P. Singer 1990). The insistence on a plurality of interests and capacities has been extended to include the interests of other animal species, as well as trees, rivers, and ecological systems (Wenz 1990).

See also Affirmative Action; Animal Rights and Welfare; Pragmatism; Racism; Sexism; Social and Political Philosophy; Speciesism.

Bibliography

Baghramian, Maria, and Attracta Ingram, eds. *Pluralism: The Philosophy and Politics of Diversity.* London: Routledge, 2000.

Barry, Brian M. *Culture and Equality: An Egalitarian Critique of Multiculturalism.* Cambridge, MA: Harvard University Press, 2001.

Geyer, M. "Multiculturalism and the Politics of General Education." *Critical Inquiry* 19 (1993): 499–533.

Guttman, A. "The Challenge of Multiculturalism in Political Ethics." *Philosophy and Public Affairs* 22 (3) (Summer 1993): 171–206.

Kekes, J. *The Morality of Pluralism.* Princeton, NJ: Princeton University Press, 1993.

Mosley, A. "Preferential Treatment and Social Justice." In *Terrorism, Justice, and Social Values,* edited by C. Peden and Y. Hudson. Lewiston, NY: Mellen Press, 1990.

Nickel, J. W. "Ethnocide and Indigenous Peoples." *Journal of Social Philosophy* 25 (1994): 84–98.

Okin, Susan Moller. *Is Multiculturalism Bad for Women?* Edited by Joshua Cohen, Matthew Howard, and Martha Nussbaum. Princeton, NJ: Princeton University Press, 1999.

Rorty, A. "The Advantages of Moral Diversity." *Social Philosophy and Policy* 9 (2) (Summer 1992): 38–62.

Singer, B. "Pragmatism and Pluralism." *Monist* 75 (4) (October 1992): 477–491.

Singer, P. *Animal Liberation.* New York: New York Review of Books, 1990.

Taylor, C. *Multiculturalism and the Politics of Recognition.* Princeton, NJ: Princeton University Press, 1992.

Thomas, L. "Moral Deference." *Philosophical Forum* 24 (1–3) (Spring 1992–1993): 233–250.

Walzer, Michael. *On Toleration.* New Haven, CT: Yale University Press, 1997.

Wenz, P. "Minimal, Moderate, and Extreme Moral Pluralism." *Environmental Ethics* 15 (1993): 61–74.

West, C. *Beyond Eurocentrism and Multiculturalism.* Monroe, ME: Common Courage Press, 1993.

Yates, S. A. "Multiculturalism and Epistemology." *Public Affairs Quarterly* 6 (1992): 435–456.

Young, M. Y. *Justice and the Politics of Difference.* Princeton, NJ: Princeton University Press, 1990.

Albert Mosley (1996)
Bibliography updated by Philip Reed (2005)

PLURALS AND PLURALITY

Plurality falls under the concept of grammatical number. So, one prefaces the discussion of plurality with a brief overview of grammatical number. Since this entry is written in English, one can consider grammatical number in English.

English nouns are either plural or singular, which is usually signaled by the presence or absence of the inflectional ending *s*. Thus, *book* (singular) contrasts with *books* (plural). However, some nouns have peculiar forms for singular and plural. For example, the plural of *louse* (singular) is *lice* (plural). Some nouns, like *deer*, do not take the suffix *s*, yet behave as both singular and plural. This is shown by the form of its preceding determiner and, should the noun be in the subject position, by the form of the main verb. Thus, in the sentence *That deer is crossing the road*, *deer* behaves like a singular noun, while in the sentence *Those deer are crossing the road*, it behaves like a plural noun. Still other nouns, such as *police*, behave only as grammatically plural.

While every English noun must appear in either a singular or plural form, not every English noun may appear in both forms. On the contrary, English pronouns have both a singular (e.g., *he*, *she*, or *it*) and a plural form (*they*), which, for the most part, share no stem. In addition, as illustrated earlier, many English common nouns, known as count nouns, occur as both singular and plural nouns. By contrast, English proper nouns appear in the singular or plural form, but not both. The singular proper noun *Aristotle* does not occur in the plural (in the same relevant sense), nor does the plural proper noun *the Andes* occur in the singular. Moreover, English common nouns, such as *dust* and *advice*, called mass nouns, occur typically only in the singular. This division between nouns that can occur both in the singular and in the plural and those that do not occur is crosscut by words that cannot be preceded by the full range of English determiners and those that can be. Thus, English nouns can be partitioned into four classes. On the one hand, proper nouns and pronouns cannot be preceded by determiners, while common nouns can be; and on the other, count nouns and pronouns occur in both singular and plural forms, while mass nouns and proper nouns do not.

	Occurs with a determiner	Admits the contrast of singular and plural
Proper name	–	–
Pronoun	–	+
Mass noun	+	–
Count noun	+	+

The contrast between singular and plural forms is signaled by the inflection of the noun, but the distinction applies to the noun phrase containing the noun. This is manifested by the fact that conjoined proper nouns behave as though they are plural. For example, while the sentence *Russell and Whitehead was coauthors* is unacceptable in English, the sentence *Russell and Whitehead were coauthors* is not.

Bearing in mind these facts about English grammatical number, one may ask what contribution grammatical number makes to a noun phrase. The commonsensical view, the one of traditional grammar, maintains that a plural noun phrase, such as *these books*, denotes more than one thing, whereas a singular noun phrase, such as *this book*, denotes precisely one thing. Matters, however, are not so simple. Some singular noun phrases, such as *Pegasus*, and some plural noun phrases, such as *the Furies*, denote nothing at all. Some singular nouns denote more than one thing. The proper noun *Benelux* denotes Belgium, The Netherlands, and Luxemburg; the collective

count noun phrase *the team* denotes the people making up the team; and *this furniture* may denote a roomful of furniture comprising, say, two tables and a sofa, each of which is, of course, a piece of furniture. At the same time, plural count nouns such as *these pants* (compare *this pair of pants*) may denote only a single thing. Finally, what single thing, if any, does the singular noun phrase *the average Roman legionnaire* denote?

Common nouns contrast with pronouns and proper names in that they tolerate being preceded either by almost any determiner or by no determiner. When a common count noun is not preceded by any determiner, it must appear in the plural form. Such noun phrases are known as bare plurals. As Greg Carlson (1977) notes, such noun phrases are liable to different construals. The noun *dogs* in the sentence *dogs are barking* can be paraphrased as *Some dogs are barking*; however, when it occurs in the sentence *dogs bark* it is not paraphrased as *some dogs bark*. Rather, it seems to express a quasi-universal statement, something like *almost all dogs bark*, often known as the generic construal. Carlson notices that a similar contrast applies to mass nouns in the bare usage. *Water is liquid* as opposed to *water is dripping* (see Carlson and Pelletier 1995).

Further questions arise with quantified noun phrases. The singular noun phrase *some boy* might be thought to contrast with the plural noun phrase *some boys* because the former pertains to a single boy, while the latter pertains to more than one boy. This contrast does not appear to obtain for the singular noun phrase *each boy* and the plural noun phrase *all boys*, nor for the singular noun phrase *no woman* and the plural noun phrase *no women*.

An important source of data for the investigation of plural noun phrases is their susceptibility to so-called collective and distributive construals. One useful way to determine what these construals consist in is to use an equivalence between plural noun phrases and conjoined noun phrases, where the conjoined noun phrases contain proper nouns. If *the men* denotes Bertrand Russell and Alfred Whitehead, then (1.0) is paraphrasable by (1.1):

(1.0) The men wrote a book.

(1.1) Whitehead and Russell wrote a book.

It has long been recognized that sentences such as (1.0) and (1.1) have different construals, distinguishable with the help of adverbs:

(2.0) The men wrote a book.

(2.1) The men wrote a book *together*.

(2.2) The men *each* wrote a book.

These are paraphrasable by the following sentences, respectively:

(3.0) Whitehead and Russell wrote a book.

(3.1) Whitehead and Russell wrote a book *together*.

(3.2) Whitehead and Russell *each* wrote a book.

The sentences in (1) are true on the collective construal, since *Principia Mathematica* was written as a collaborative effort of Whitehead and Russell. This construal can be forced by the use of the adverb *together*, as in (2.1) and (3.1). The sentences in (1) are also true on the distributive construal, since Russell wrote at least one book on his own, for example, *An Inquiry into Meaning and Truth*, and Whitehead also wrote a book on his own, for example, *A Treatise on Universal Algebra*. This construal can be enforced by the use of the adverb *each*, as in (2.2) and (3.2).

As shown by the next example, the susceptibility of plural noun phrases to collective and distributive construals is not confined to collaboration:

(4.0) These two suitcases weigh fifty kilograms.

(4.1) These two suitcases *each* weigh fifty kilograms.

(4.2) These two suitcases weigh fifty kilograms *together*.

Moreover, collective and distributive construals seem to be the extremes of a range of construals. If *the men* denotes Richard Rodgers, Oscar Hammerstein, and Lorenz Hart, it is true to say that

(5) The men wrote musicals.

even though none of them wrote a musical on his own and the three never wrote a musical together. What is true is that Rodgers and Hammerstein wrote musicals and Rodgers and Hart wrote musicals (see Gillon 1987).

Next, it should be noted that susceptibility of collective and distributive construals is not confined to plural noun phrases in the subject position. Every argument position containing a plural noun phrase—be it the subject, object, indirect object, or object of a preposition—is liable to these construals, regardless of whether the noun phrase is an argument of a verb or of a noun (see Gillon 1996).

(6.1) Isabelle gave the girls a cookie.

(6.2) Rick drove through the Redwood trees.(Compare *Rick drove through the Redwood tree.*)

(6.3) The two suitcases' weight is fifty kilograms.

(6.4) The writing of *Principia Mathematica* by Russell and Whitehead.

Finally, even singular count nouns give rise to collective and distributive construals. Suppose that someone has two suitcases and says:

(7) This luggage weighs fifty kilograms.

The sentence could be taken to mean that altogether the luggage weighs fifty kilograms or that each piece of luggage weighs fifty kilograms.

Two crucial questions arise for the semantics of plural noun phrases: First, what do plural noun phrases denote? Second, how does one account for the various construals to which they are liable?

One can begin with the first question. According to the earliest researchers to address the question, such as Michael Bennett (1974) and Roland Hausser (1974), plural noun phrases denote sets. This view was roundly criticized by Godehard Link (1983) and Peter Simons (1983), who argued, independently of each other, that plural noun phrases do not denote sets, but what Simons called pluralities. Whereas a set of concrete individuals is an abstract mathematical entity, without spatial or temporal location, a plurality of concrete individuals is a concrete entity, with the spatial and temporal location of its membership. However, for both a set and a plurality, identity is determined by membership.

A plurality, then, is nothing more than the sum of its members. At the same time, a plurality is different from a collective, which may be more than the sum of its members. Thus, while a plurality is identified by its membership, so that if it acquires or loses a member, it becomes a different plurality, a collective is not identified simply by its members, for it can remain the same, even if its membership changes. Thus, an orchestra can remain the same, even though its members change. Inversely, the exact same individuals might constitute two collectives. Indeed, Simons (1982) reports that once the same musicians made up the Chapel Orchestra, the Court Opera Orchestra, and the Vienna Philharmonic. Nonetheless, a plurality can also be seen as the limiting case of a collective: that is, a plurality is a collective without conditions governing its constitution (Simons 1987, chapter 4.4).

The set of pluralities on a finite domain has the structure of a join semilattice. For example, consider three people: Dan, Paul, and Rick. They can form seven pluralities: three improper—Dan, Paul, and Rick; and four improper—Dan + Paul, Dan + Rick, Paul + Rick, and Dan + Paul + Rick. The algebraic operation symbolized here by + , is a join operation. It is idempotent ($x + x = x$), since there is no difference between Dan and Dan + Dan; it is commutative, since there is no difference between Dan + Paul and Paul + Dan; and it is associative, since Dan + (Paul + Rick) is the same plurality as (Dan + Paul) + Rick. The seven pluralities are all concrete individuals.

The join semilattice just described is isomorphic to the join semilattice obtained by assigning each plurality, proper and improper, a set: An improper plurality is assigned a singleton set. Thus, Dan is assigned {*Dan*}, a plurality comprised of two people is assigned a doubleton set. Thus, Dan + Rick is assigned {*Dan, Rick*}. And the plurality comprising three people is assigned a set of three people. The operation on these sets corresponding to + is that of union. Since every join semilattice of pluralities is isomorphic to a join semilattice of sets, a number of semanticists, including Fred Landman (1989a, 1989b), Roger Schwarzschild (1996), and Yoad Winter (2001), are content to treat pluralities as sets.

Link (1983) develops a semantics for a formal notation, designed to simulate singularity and plurality. Like Simons (1983, 1987), Link views the denotations of plural count noun phrases as distinct from the denotations of singular mass noun phrases, the former having their denotation based on individuals, the latter on so-called masses (see the mass noun entry). This distinction in denotation seems implausible, in view of the near synonymy of mass nouns such as *footwear*, *luggage*, *traffic*, and *advice*, with count nouns such as *shoes*, *suitcases*, *vehicles*, and *suggestions*. In light of such facts, Gillon (1992) provides a semantics of common nouns whereby a plural noun phrase such as *shoes* and a singular noun phrase such as *footwear* may have the same denotation; after all, all shoes are footwear, even if some footwear are not shoes. Another semanticist to provide a uniform domain for the interpretation of mass nouns and count nouns is Almerino Ojeda (1993), who takes all nouns to denote, in the first instance, kinds.

One can now turn to the second question: How are the various construals of plural noun phrases to be explained? A few authors such as Gillon (1987, 1992, 1996) and Schwarzschild (1996) think that the collective and distributive construals are extremes of a variety of construals, which, in their view, is pragmatically determined. However, the preponderance of authors recognize only two construals—the collective and distributive construals of traditional grammar—and take them to be the result of an ambiguity arising from the presence or

absence of an unpronounced adverb. For some, like Link (1983), the adverb is essentially a phonetically null version of the English adverb *each*. For others, like Landman (1989a, 1989b), it is a phonetically null collectivizing operator applying to noun phrases. In fact, each of these views require no less than three kinds of phonetically null operators. Since virtually every plural noun phrase, no matter where it occurs in a sentence, is liable to collective and distributive construals, no fewer than three such phonetically null words are required (see Gillon 1996). Finally, several authors (Schein 1993, Lasersohn 1995, Landman 2000) have tried to develop a theory of events and their parts and participants to account for collective and distributive construals.

See also Generics; Nouns, Mass and Count.

Bibliography

Bäuerle, Rainer, Christoph Schwarze, and Arnim von Stechow, eds. *Meaning, Use, and Interpretation of Language*. Berlin, Germany: Gruyter, 1983.

Bennett, Michael R. "Some Extensions of a Montague Fragment of English." PhD diss., University of California, Los Angeles, 1974.

Carlson, Greg. "A Unified Analysis of the English Bare Plural." *Linguistics and Philosophy* 1 (1977): 413–457.

Carlson, Greg, and Jeff Pelletier, eds. *The Generic Book*. Chicago: University of Chicago Press, 1995.

Gillon, Brendan S. "Collectivity and Distributivity Internal to English Noun Phrases." *Language Sciences* 18 (1–2) (1996): 443–468.

Gillon, Brendan S. "A Common Semantics for English Count and Mass Nouns." *Linguistics and Philosophy* 15 (6) (1992): 597–639.

Gillon, Brendan S. "The Readings of Plural Noun Phrases in English." *Linguistics and Philosophy* 10 (1987): 199–220.

Hausser, Roland. *Quantification in an Extended Montague Grammar*. PhD diss., University of Texas, Austin, 1974.

Landman, Fred. *Events and Plurality: The Jerusalem Lectures*. Dordrecht, Netherlands: Kluwer Academic, 2000.

Landman, Fred. "Groups I." *Linguistics and Philosophy* 12 (5) (1989a): 559–606.

Landman, Fred. "Groups II." *Linguistics and Philosophy* 12 (6) (1989b): 723–744.

Lasersohn, Peter. *Plurality, Conjunction, and Events*. Dordrecht, Netherlands: Kluwer Academic, 1995.

Link, Godehard. "The Logical Analysis of Plurals and Mass Terms." In *Meaning, Use, and Interpretation of Language*, edited by Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 302–323. Berlin, Germany: Gruyter, 1983.

Ojeda, Almerindo. *Linguistic Individuals*. Stanford, CA: Center for the Study of Language and Information, 1993.

Schein, Barry. *Plurals and Events*. Cambridge, MA: MIT Press, 1993.

Schwarzschild, Roger. *Pluralities*. Dordrecht, Netherlands: Kluwer Academic, 1996.

Simons, Peter M. "Class, Mass, and Mereology." *History and Philosophy of Logic* 4 (1983): 157–180.

Simons, Peter. *Parts: A Study in Ontology*. New York: Oxford University Press, 1987.

Simons, Peter M. "Plural Reference and Set Theory." In *Parts and Moments: Studies in Logic and Formal Ontology*, edited by Barry Smith, 199–260. Munich, Germany: Philosophia (Analytica), 1982.

Smith, Barry, ed. *Parts and Moments: Studies in Logic and Formal Ontology*. Munich, Germany: Philosophia (Analytica), 1982.

Winter, Yoad. *Flexibility Principles in Boolean Semantics: The Interpretation of Coordination, Plurality, and Scope in Natural Language*. Cambridge, MA: MIT Press, 2001.

Brendan S. Gillon (2005)

PLUTARCH OF CHAERONEA

(c. 46–after 119, before 127 CE)

Plutarch, a Greek biographer and Platonic philosopher, was born in Chaeronea, Boeotia. His teacher was Ammonius, an Egyptian Platonist who resided in Athens and was head of a school that he called the Academy. After his studies (c. 90?) Plutarch established a philosophical school in Chaeronea. Plutarch held important public offices and was a priest at Delphi for twenty years or more. His extant writings include forty-eight biographies and various other works (*Moralia*): dialogues; diatribes; theoretical treatises; essays; collections of anecdotes; moralistic lectures; and polemical, antiquarian, and exegetical works. Several dialogues have Delphi as their setting and are concerned with the oracle and other religious problems. *Socrates' Daemonic Sign* has a historical setting. It portrays Plutarch's circle of friends and students. *Table-Talks* is a long collection of conversations on a wide range of questions.

INFLUENCES

Plato's dialogues, especially the *Timaeus*, but also Platonic school philosophy, as it could be found in manuals and introductory works, provide the basis of Plutarch's philosophy. In Plutarch's day, Platonism was dominated by Pythagorean tendencies, most importantly the tendency to construct a hierarchy of metaphysical principals based on an ontological derivation from the principals "one" and "dyad." Plutarch himself, however, was just as much influenced by the skepticism of the Hellenistic Academy, though in the mitigated form it took under Philo of Larissa. This influence shows in the limited epistemic sta-

tus he granted to empirical science, his cautious attitude regarding the epistemic claims of popular religion, and his reflections on the unreliability of the senses. This epistemology can be traced back to Plato's *Timaeus*, and Plutarch explicitly did so. He developed a kind of fallibilism that allowed him provisionally to accept various physical doctrines, for example, about the nature of the moon, or the function of specific organs of the body. The Hellenistic Academy provided Plutarch with numerous arguments against Stoics (*Common Notions*, *Stoic Contradictions*) and Epicureans (*Reply to Colotes* and *That Epicurus Actually Makes a Pleasant Life Impossible*).

COSMOLOGY AND METAPHYSICS

Plutarch devoted a separate treatise to Plato's description of the composition of the world soul in the *Timaeus* and discussed this issue in several other places. Contrary to the large majority of his fellow Platonists, Plutarch understood Plato's story of the creation of the cosmos by a divine craftsman literally, in that he believed that the cosmos had existed only for a finite time. It came into being when the craftsman, or demiurge, imposed order on a preexisting chaos. Previous to his intervention, there was matter and a precosmic soul, as the principle of motion, both in a disordered state. The Platonic forms too existed, as their existence is eternal, but the world did not yet participate in them. When the demiurge imparted something of himself—namely intelligibility, or mathematically expressible rationality—to this preexisting soul, it became the world soul. The world soul then started to organize matter and create a structured, beautiful world (or cosmos).

Time, in the Platonic sense of succession characterized by cyclic regularities, was born together with the world. Plutarch leaves unspecified the relation between the craftsman and the forms. The forms and the craftsman belong to the same realm, and when the craftsman imparts something of himself to the preexisting soul, the latter, and through it the world, partake of the forms. The world is not perfect, as the original irrational soul, now integrated into the world soul, at times makes its influence felt. Soul itself, that is, soul in abstraction from the order it has received, is thus Plutarch's principle of evil. Plutarch espouses a mitigated metaphysical dualism: The rational and the irrational, order and disorder, good and evil are engaged in an unending struggle, but the good always dominates. The good he attributed to the gods, whereas higher forces responsible for evil can be mere demons, not gods. Plutarch linked his dualist views to an antagonism, at the level of metaphysical principles, between the One and the indeterminate Dyad. This doctrine was attributed to Plato from as early as Aristotle and was cherished by Pythagorean Platonists. Plutarch equates the demiurge with the highest deity. In his dialogue *The Delphic E*, Plutarch has his master Ammonius define the supreme god as true being, eternity, and absolute unity, and call this god the One. In his treatise on Egyptian religion, *Isis and Osiris*, Plutarch interprets Egyptian myths allegorically and explains how they conform with Plato's cosmology and metaphysics, as he understands them.

MORAL PSYCHOLOGY AND ETHICS

The human soul, being an image of the world soul, is analogously constituted. It too consists of rational and irrational parts, the latter being more prominent than it is in the world soul, however. The irrational is part of the human soul itself, is the cause of disorder and the passions, but is also the dynamic force of our mental life. Rationality is intellect and the truly divine coming from outside.

In the eschatological myth at the end of *The Face in the Moon*, Plutarch develops a theory of a double death: In "ordinary death," the human soul frees itself from the body and ascends to the moon; after purification a second death ensues wherein the intellect sheds the irrational part. In *Moral Virtue*, Plutarch transposes his cosmological views onto the human soul and on this basis erects a theory of virtue as the mean and the moderation of the passions (*metriopatheia*). Plutarch's virtue ethics stands in a Peripatetic tradition, yet has its theoretical foundations in Platonic traditions as well. Our souls have a rational and an irrational part or force—the passions. The passions have to be made obedient to reason. Reason imposes limit and structure, or even in a sense *is* the limit, establishing the right mean between extremes, moral virtue between opposite vices. When the passions obey reason, the human soul achieves psychic harmony, which is a necessary and perhaps even sufficient condition for happiness in this life (though not necessarily for success in one's undertakings) and leads to felicity in the next. This is also the fundamental lesson of Plutarch's texts on practical ethics. Plutarch was a keen observer of human behavior, virtues, and vices. His *Lives* essentially consists of character studies, and some two dozen of his *Moralia* are on moral themes. Titles include *Advice to Bride and Groom*, *How to Tell a Flatterer from a Friend*, *Inoffensive Self-Praise*, *Exile*, *Compliancy*, *Superstition*, *Control of Anger*, *Tranquillity of Mind*, *Brotherly Love*, *Talkativeness*. Moral considerations dominate his

approach to literature in *How to Study Poetry*. He even wrote on the behavior of animals: *The Cleverness of Animals* and *Beasts Are Rational*.

PLATONISM, STOICISM, AND EPICUREANISM

Plutarch incorporated ideas, examples, and terminology from other schools into his texts, but he subordinated them to his overall Platonism. This is especially obvious in his dialogues: He presents and examines various views; this typically leads to a Platonic position in which he combines what is sound in the views of other schools and adds an additional, transcendental, perspective. Plutarch construed his Platonism as occupying a middle position between Stoicism and Epicureanism. Whereas the Epicureans denied providence and the Stoics made the gods responsible for everything, the Platonic god is causally responsible for good things only. Plutarch combated the Stoic monolithic view of the mind and the Stoic ideal of being passionless: The passions constitute an intrinsic, indelible part of our psychic make-up; hence we have to learn to manage and control them.

See also Ancient Skepticism; Aristotle; Epicureanism and the Epicurean School; Neoplatonism; Plato; Pythagoras and Pythagoreanism; Stoicism; Virtue Ethics.

Bibliography

All the extant works of Plutarch have been edited and translated in *The Loeb Classical Library*. Among the more important spurious works included in the Plutarchan corpus are *Philosophical Opinions* (a doxography), *Music* (drawn from Peripatetic sources), *Fate* (a product of syncretistic Platonism of the second century CE), and, on the level of popular philosophy, *The Education of Children* and *Consolation to Apollonius*.

GENERAL INTRODUCTIONS TO PLUTARCH

Russell, D. A. *Plutarch*. London: Duckworth, 1973.

Sirinelli, Jean. *Plutarque de Chéronée: Un philosophe dans le siècle*. Paris: Fayard, 2000.

Ziegler, Konrat. *Plutarchos von Chaironea*. In Pauly's Realenzyklopädie, 41. Halbband, 1951: 636.18-962.14 (also published as a monograph: Stuttgart, 1949)

IMPORTANT STUDIES ON PLUTARCH'S PHILOSOPHY

Babut, Daniel. *Plutarque et le stoïcisme*. Paris: Presses Universitaires de France, 1969.

Brenk, Frederick E. "An Imperial Heritage: The Religious Spirit of Plutarch of Chaironeia." *Aufstieg und Niedergang der Römischen Welt* 2.36.1 (1987): 248–349.

Dillon, John. *The Middle Platonists: A Study of Platonism, 80 B.C. to A.D. 220*. Rev. ed. London: Duckworth, 1996.

ON THE *LIVES*

Duff, Tim. *Plutarch's Lives: Exploring Virtue and Vice*. Oxford, U.K.: Clarendon Press, 1999.

Jan Opsomer (2005)

PNEUMA

Ancient Greek thought early posited a connection between breath and life. The notion that wind or breath—*pneuma*—accounted for the functions of living things persisted in philosophical and medical accounts of organisms, sometimes alongside the notion of an immaterial soul or *psychē*. The idea that a distinct kind of *pneuma* played a role in the functioning of organisms seems to have developed in early medical theory. Some texts refer to *pneuma* as a kind of nutriment. The idea that there is a specifically "psychic" *pneuma* is found in the doctor Diocles of Carystus (fourth century BCE), who had connections to Aristotle's school.

In Aristotle's biology an innate *pneuma* is mentioned in connection with a number of functions of the organism and is even compared to the *ether*, the fifth element from which the heavenly bodies are composed. In the case of sexual generation *pneuma* is used to explain the ability of the male seed to convey its movements to the female matter without contributing matter to the resulting embryo; in animal movement it helps explain the movement of the limbs. There is room for doubt about how systematically Aristotle used the concept, however, or its relationship to the elements. His second successor, Strato of Lampsacus, seems to have considered *pneuma* to be the material substance of the soul, perhaps in recognition of the discoveries of Hellenistic medicine; a treatise on *pneuma* survives in the Aristotelian corpus.

Praxagoras (fourth century BCE), who distinguished veins from arteries, theorized that the latter contain only *pneuma*; this was eventually rejected by Galen. The Hellenistic doctors Herophilus (c. 335–c. 280 BCE) and Erasistratus (flourished c. 250 BCE) recognized a system of *neura* or nerves originating from the brain, responsible for motor and perceptual functions. Because some nerves were seen to be hollow, they were thought to contain a special kind of *pneuma* suited to their functions. In Galen's physiology the "vital *pneuma*" is distributed through the arteries; the brain refines this into "psychic *pneuma*," which, through the nerves, is the instrument by which the soul performs its functions.

Unlike these medical theories associating *pneuma* with the vascular systems, Epicurus describes the material

soul as like, or partly composed of, *pneuma*. In Stoic philosophy it played a broader role. The Stoics hypothesized that *pneuma*—for them, a kind of hot air—is distributed throughout all other matter in the cosmos. Supposing that all action happens by bodies in contact, yet needing to account for apparent cases of action at a distance, the Stoics held that the pervasiveness of this single material accounted for the "sympathy" between distant bodies, as well as the cohesiveness of the cosmos as a whole and the qualities of individual things. Associated with the divine intelligence pervading the cosmos, the part of the cosmic *pneuma* pervading living things is the soul.

The Greek term *pneuma* was later used in religious contexts and associated with spirit and the divine. The physiological use of *pneuma* to account for functions of living things is echoed in the early modern notion of "animal spirits."

See also Aristotle; Epicurus; Stoicism; Strato and Stratonism.

Bibliography

Freudenthal, Gad. *Aristotle's Theory of Material Substance: Heat and Pneuma, Form and Soul*. Oxford, U.K.: Clarendon Press, 1995.

Solmsen, Friedrich. "Greek Philosophy and the Discovery of the Nerves." In *Kleine Schriften*. Vol. 1, 536–582. Hildesheim, Germany: Georg Olms Verlagsbuchhandlung, 1968.

Staden, Heinrich von. "Body, Soul, and Nerves: Epicurus, Herophilus, Erasistratus, the Stoics, and Galen." In *Psyche and Soma: Physicians and Metaphysicians on the Mind-Body Problem from Antiquity to Enlightenment*, edited by John P. Wright and Paul Potter, 79–116. Oxford, U.K.: Clarendon Press, 2000.

Sylvia Berryman (2005)

POINCARÉ, JULES HENRI
(1854–1912)

Jules Henri Poincaré, the French mathematician and philosopher, was born into a distinguished family at Nancy. His cousin Raymond was both prime minister and president of the Third French Republic. At an early age Poincaré showed an interest in natural history and the classics, and at the age of fifteen he developed an interest in mathematics. However, he trained first as a mining engineer, studying mathematics on his own during this training. In 1879 he was appointed to teach courses in mathematical analysis in the Faculty of Science at Caen. In 1881 he moved to the University of Paris, where he was soon given charge of the courses in mathematics and experimental physics. He lectured on mechanics, mathematical physics, and astronomy. Poincaré wrote an enormous number of papers on mathematics and physics and several important books on the philosophy of science and mathematics, as well as popular essays on science. His most important mathematical contributions were in differential equations, number theory, and algebra. In 1887 he was elected a member of the Académie des Sciences, and in 1899 he was made a knight of the Légion d'Honneur for his work on the three-body problem. In 1906 he became president of the Académie des Sciences, and in 1908 he was elected to the Académie Française.

Poincaré's work in the philosophy of science was in the tradition of Ernst Mach and Heinrich Hertz, and he admitted a debt to Immanuel Kant. His work was clearly influenced by his mathematical approach, and his interest was largely in the formal and systematic character of theories in the physical sciences. He showed less concern with epistemological problems connected with their support and establishment although he did write on the psychology of discovery. Albert Einstein had a profound respect for his work in both mathematics and the philosophy of science. He is often claimed as an ancestor of logical positivism, although the justification is not always easy to see.

AIMS AND GENERAL CHARACTER OF SCIENCE

Underlying scientific procedures, Poincaré held, is a belief in a general order in the universe that is independent of us and our knowledge. This is what mainly distinguishes the sciences from mathematics, which presupposes, if anything, merely the ability of the human mind to perform certain operations. The aim of the scientist is to discover as much as possible of the order of the universe, a point which must be borne in mind when Poincaré's view is called "conventionalism."

The method of discovery is basically inductive, proceeding by generalizing from observed facts; its lack of finality is due to its basis in a belief in a general order, since we can never be sure that the discovered order is absolutely general. Modifications in scientific conclusions spring from the constant pursuit of this generality. The discovery of facts depends upon observation and experiment, but these, in turn, depend upon selection because scientists cannot observe and absorb everything at once. There must be some principle of selection, but this principle must not be one of morality or practical utility. The search for an acceptable principle of selection led Poincaré to the idea of simplicity and a somewhat unusual

defense of this idea. The best scientists are motivated by disinterested curiosity about how the world is, and their interest in general truths leads them to select those facts that "have the greatest chance of recurring." These are simple facts—that is, facts with few constituents. On grounds of probability there is more chance of the recurrence of a few constituents together than of the recurrence of many constituents together. However, familiar facts are more likely to appear simple to us than are unfamiliar facts. This seems to involve an unresolved conflict between two conceptions of simplicity.

What did Poincaré mean by "facts"? This is a question to which he gave less attention than it deserves. He held that science is to some extent objective. He toyed with sensationalism, but as a means of obtaining the necessary objectivity, he asserted that many sensations have external causes. Thus, he cannot strictly be regarded as a sensationalist. Objects are groups of sensation but not merely this; the sensations are "cemented by a constant bond," and science investigates this bond, or relation. Our sensations reflect whatever it is in the external world between which relations hold; science teaches us not the true nature of things but only the true relations between things. Scientific conclusions may thus be true of the world since they can give us a picture of its structure, though not of its content. We should expect theories of light, for example, to tell us not what light is but only what relations hold between the various occurrences of whatever light is.

The two main aims of scientific investigations are to relate what previously appeared unrelated and to enable us, by using these relations, to predict new phenomena.

CONVENTIONS

Poincaré constantly compared the physical sciences with pure mathematics and said that their methods of discovery are similar even though their methods of supporting conclusions are different. His view of science emerges most clearly from his comparison of it with geometry, in *Science and Hypothesis.* The space of geometry is not the space of sense experience; we can arrange conditions so that two things that look equal to a third thing do not look equal to each other. The mathematical continuum is invented to remove this disagreement with the law of contradiction; then, in mathematics things equal to the same thing are equal to one another *whatever* our senses tell us. This is one of those axioms of analysis, not geometry, which Poincaré called "analytical a priori intuitions."

Some geometrical axioms look superficially like this—for example, the Euclidean axiom that through one point only one line parallel to a given line can be drawn. The development of non-Euclidean geometries has shown that such axioms do not, as was formerly supposed, state fundamental properties of observable space. Coherent systems of geometry can be constructed based on the denial of Euclid's axioms, and these new geometries, when suitably interpreted, are translatable into Euclidean geometry. Moreover, they have physical applications. The applicability of the various systems is a function of context, or scale. The representation is purely structural.

Poincaré concluded that geometrical axioms are not synthetic a priori truths, for they are not of necessity true, and not experimental truths, for geometry is exact. They are conventions, or disguised definitions. It does not follow, as some critics have supposed, that they are arbitrary, for our choice is controlled by observation, experiment, and the need to avoid contradictions; nevertheless, such axioms cannot be either true or false. They are adopted because in certain contexts they are useful for saying how the world is. For most purposes Euclidean geometry is the most convenient. The application of geometry to the world involves an idealization. "Thus we do not *represent* to ourselves external bodies in geometrical space, but we *reason* about these bodies as if they were situated in geometrical space." No experimental support for Euclidean or any other geometry is possible, since experiments tell us only about the relations between bodies and nothing about the relations of bodies to space or of one part of space to another.

The physical sciences contain a conventional element as well as experimental, mathematical, and hypothetical elements, a fact which has been missed by most scientists. For example, the principle of inertia, according to which a body under the action of no force can move only at a constant speed in a straight line, is neither a priori nor experimental. It was originally conceived as experimental but has become a definition and so cannot now be falsified by experiment. Scientific conclusions are always conventional to some extent since alternatives to any hypothesis are always possible and, other things being equal, we choose those that are most economical. Because we have no means of knowing that the qualitative features of our hypothesis correspond to the reality, it does not make sense to regard the chosen hypothesis as the one true hypothesis.

In the physical sciences there are two kinds of statement—laws, which are summaries of experimental results and are approximately verified for relatively isolated systems, and principles, which are conventional

postulates, completely general, rigorously true, and beyond the reach of experimental testing because for reasons of convenience we have made them so. Science is not entirely conventional because it does not consist wholly of principles. We begin with a primitive law, or experimental conclusion, but this is broken up into an absolute principle (definition) and a revisable law. Poincaré's example is the empirical statement "The stars obey Newton's law," which is broken up into the definition "Gravitation obeys Newton's law" and the provisional law "Gravitation is the only force acting upon the stars." Gravitation is an invented, ideal concept, but the provisional law is empirical and nonconventional because it predicts verifiable facts. The law of the conservation of energy is an outstanding example of a convention; it defines the concept of energy.

Prediction involves generalization, and generalization involves idealization. We connect a number of points on a graph by a smooth curve which does not pass through every one of them, and so we presuppose that the law we seek is best represented by a smooth curve even if this does not exactly fit the experimental results.

Points chosen midway between the existing points have a much better chance of showing which curve we should draw by eliminating one of them. A hypothesis is most strongly supported when it passes the tests that it was most likely to fail.

UNITY AND SIMPLICITY

We can obtain new knowledge only through experiment, and the role of mathematics in the physical sciences is to direct our generalizations from experiment. But experiment and generalization depend on presuppositions, most of which we make unconsciously. Among our presuppositions the most important are beliefs in the unity and simplicity of nature. Unity involves the possibility that various parts of the universe act upon one another as do the various parts of the human body, in the limited sense that to understand and describe one phenomenon, we may have to investigate other, superficially unrelated phenomena. The presupposition of simplicity is weaker: We can generalize any fact in an infinite number of ways, and we actually generalize in the simplest way until we have evidence against this way.

Two opposing trends can be discerned in the history of science. There is a movement toward simplicity and unity when we discover new relations between apparently unconnected objects and a movement toward complexity and diversity when, with the help of improved techniques, we discover new phenomena. The progress of science depends upon the first tendency, for "the true and only aim is unity." The second tendency is important, but it must ultimately give way to the first. Poincaré argued, referring to the growing unification of the studies of light, magnetism, and electricity, that there are signs of a continued victory for the tendency toward unity. But there are also signs that this does not always go along with simplicity since unity can sometimes be achieved only by revealing the increased complexity in things when shown to be related. However, unity is essential and simplicity merely desirable.

Poincaré's account, like many others, suffers from a lack of clarity concerning precisely what is meant by "simplicity."

HYPOTHESES

Poincaré distinguished three kinds of hypotheses. The first kind he called "natural and necessary," and they are the very general hypotheses that we use in making judgments of relevance—for instance, when in physics we judge that the effect of very distant bodies is negligible. These form the common basis of theories in mathematical physics and should be the last to be abandoned.

The second kind he called "indifferent," and these are useful artifices for calculation or pictorial aids to understanding. Hypotheses are of this kind when they are alternatives that cannot be distinguished by experiment. Thus, he said, the two hypotheses that matter is continuous and that matter has an atomic structure are indifferent because experiment cannot establish the real existence of atoms. Such hypotheses may be useful, but they may also be seriously misleading if they are not seen for what they are.

The third kind of hypotheses he called "real generalizations." They are direct generalizations from observations and are indefinitely open to further testing. Whether or not they are finally accepted, they are always valuable, if only for their suggestiveness.

THEORIES AND THE ROLE OF MATHEMATICS

The aim of experiments in physical science is to break up complex phenomena into simple ones with respect to time and space, to connect each moment in the development of phenomena with immediately contiguous moments and each point in space with immediately contiguous points. We also aim to break up complex bodies and events into elementary bodies and events. Because observable phenomena may be analyzed in this way and

be regarded as the result of large numbers of elementary phenomena similar to one another, they are conveniently described by differential equations. This accounts for the ease with which scientific generalization takes a mathematical form. Mathematical physics depends upon the approximate homogeneity of the matter studied, since this enables us to extrapolate.

A physical theory may be superseded by another that uses qualitatively different concepts but the same differential equations; the equations are merely given different interpretations in the two theories. The superseded theory will be just as valuable for prediction because it contains the same relations as the new one, and as long as these stand up to testing, we can say that these are the real relations between things in the world. Both theories are true in the only way in which it makes sense to talk of the truth of a theory. Any advantage that the new theory has over the old will be merely psychological and will lie in its suggestions rather than in its implications. It is relatively unimportant that one theory of light refers to the movement of an ether and another refers to electric currents; what is important is the extent to which their equations agree, and it is on this that their truth must be judged.

Theories do not set out to explain, although they may provide possible explanations. They are devices enabling us to connect and predict phenomena but not to describe reality in all its details. The assertion that, for example, atomic theories explain the behavior of matter implies that we are able to establish the actual existence of atoms as delineated by the theories. But this is a metaphysical and not a scientific assertion because such existence can never be established by scientific means.

MATHEMATICS AND LOGIC

In mathematics Poincaré was, on the whole, an intuitionist, holding that the integers are indefinable and that underlying all mathematics is the principle of mathematical induction whose validity is intuitively recognized—that is, synthetic a priori.

In his last years Poincaré made a lively attack on the logic of Giuseppe Peano, Bertrand Russell, and others, especially on the logistic attempt to reduce mathematics to logic (*Mathematics and Science: Last Essays,* Chs. 4–5). He thought it important to study not only the consequences of adopting given conventions but also the reasons for adopting these conventions rather than others. He argued that it is impossible to derive all mathematical truths from the accepted logical principles without further appeals to intuition. He pointed, for example, to the difficulty of defining numbers without begging the question, and he saw even in the foundations of Russell's logic a reliance, inescapable on any satisfactory account, on synthetic a priori principles. He objected to the idea of an actual infinity, which he claimed was essential to Russell's system, and held that the logical paradoxes could be avoided by excluding nonpredicative definitions—that is, definitions of particular members of a class which refer to all the members of that class (*Science and Method,* Book II, Chs. 4–5). He expressed a general dissatisfaction with the extensional interpretation of logical constants.

See also Mathematics, Foundations of.

Bibliography

WORKS BY POINCARÉ

Oeuvres de Jules Henri Poincaré. 11 vols. Paris: Gauthier-Villars, 1928–1956. Contains all Poincaré's important scientific papers together with part of his own account of his work, a biography by G. Darboux (Vol. II), and centenary lectures on his life and work (Vol. XI).

La science et l'hypothèse. Paris: Flammarion, 1902. Translated by W. J. Greenstreet as *Science and Hypothesis.* London, 1905. Poincaré's most important book on the philosophy of science and mathematics.

La valeur de science. Paris: Flammarion, 1905. Translated by G. B. Halsted as *The Value of Science.* New York: Science Press, 1907.

Science et méthode. Paris: Flammarion, 1908. Translated by Francis Maitland, with a preface by Bertrand Russell, as *Science and Method.* London: Nelson, 1914.

Dernières Pensées. Paris: Flammarion, 1912. Translated by John W. Bolduc as *Mathematics and Science: Last Essays.* New York: Dover, 1963.

"Analyse des travaux scientifiques de Henri Poincaré, faite par lui-même." *Acta Mathematica* 38 (1921): 1–385. Partly reprinted in *Oeuvres.* This is a Poincaré issue of the journal and contains an extensive bibliography and appreciations by J. Hadamard, H. A. Lorenz, P. Painlevé, M. Planck, L. Fuchs, and others.

WORKS ON POINCARÉ

Carnap, Rudolf. *Logical Syntax of Language.* London: Kegan Paul, 1937.

Einstein, Albert. "Geometry and Experience." In *Sidelights on Relativity,* translated by G. B. Jeffery and W. Perrett. London, 1922.

Frank, Philipp. "Einstein, Mach and Positivism." In *Albert Einstein, Philosopher-Scientist,* edited by P. A. Schilpp. Evanston, IL: Library of Living Philosophers, 1949. Other articles also mention Poincaré.

Frank, Philipp. *Modern Science and Its Philosophy.* Cambridge, MA, 1949. Logical positivism's debt to Poincaré.

Grünbaum, Adolf. "Carnap's Views on the Foundations of Geometry." In *Philosophy of Rudolf Carnap,* edited by P. A. Schilpp. La Salle, IL: Open Court, 1964.

Hadamard, J. S. *The Early Scientific Work of Henri Poincaré.* Rice Institute Pamphlet (Houston, TX) 9 (3) (1922).

Hadamard, J. S. *The Later Scientific Work of Henri Poincaré.* Rice Institute Pamphlet (Houston, TX) 20 (1) (1933).

Le Roy, É;douard. "Science et philosophie. "*Revue de métaphysique et de morale* 7 (1899): 375, 503, 706; 8 (1900): 37.

Popper, K. R. *The Logic of Scientific Discovery.* London: Hutchinson, 1959. For a somewhat misleading criticism of Poincaré's conventionalism.

Revue de métaphysique et de morale 21 (1913): 585–718. Obituary number of journal.

Rey, A. *La théorie de la physique chez les physiciens contemporains.* Paris, 1907.

Schlick, Moritz. *Raum und Zeit in der gegenwärtigen Physik* Berlin, 1917. Translated by H. L. Brose as *Space and Time in Contemporary Physics.* Oxford, 1920.

Volterra, V. "H. Poincaré." *Book of the Opening of the Rice Institute* 3 (Houston, TX, 1912): 899–928.

Peter Alexander (1967)

POLITICAL PHILOSOPHY, HISTORY OF

The history of political philosophy is the succession of notions about the actual and proper organization of people into collectivities and the discussion of those notions. It is philosophical in character, because it is concerned with obedience and justice as well as with description; the persistent preoccupation of political philosophers has been the definition of justice and of the attitude and arrangements that should create and perpetuate justice.

A distinctive characteristic of political philosophizing is that it has usually been undertaken in response to some particular political event, or possibility, or threat, or challenge. This has led to a raggedness, even an incoherence, in works devoted to it and to an emphasis on intuitive argument which compares unfavorably with the content of other philosophical literature. Political philosophy has sometimes been supposed to confine itself to a particular entity called "the state," but in fact political philosophers have always concerned themselves with the collectivity as a whole, even when they have drawn a distinction between "state" and "society."

Problems of definition and description might appear to be prior to problems of analysis and prescription in political philosophy. In fact, however, ethical doctrine has always had a powerful effect on the view that a political thinker takes of the collectivity; he has tended to see it in terms of what he thinks it ought to be. Nevertheless, it has become usual to separate the empirical element from the normative. Empirical study has been further divided into sociology and political science. These definitions and divisions are no more satisfactory than others devised for similar purposes, and although we talk with some confidence of "sociologists," "political scientists" have only very recently emerged as an independent class of thinkers.

It is often useful to look upon political philosophy as in some sense systematic, proceeding from a view of reality and knowledge (ontology and epistemology) to a view of the individual (psychology) and a view of the social bond (sociology), and so to a general ethic, a political ethic, and finally to a set of recommendations about the form of the state and about political conduct. The expression "political philosophy" will be used in this sense here, and it will be considered solely in terms of the Mediterranean-European tradition.

CRITIQUE OF THE SUBJECT

There are several ways in which the history of political philosophy has been found important. Every thinker who engages in speculating about state and society and in formulating principles concerning them is anxious to know of the performance of his predecessors, to learn from them and to share their minds. Every thinking citizen is in this position too, to some extent, at least in the democracies: The questions raised in political life are frequently philosophical questions. Both thinkers and citizens, moreover, have good reason to believe that the intellectual and cultural life which they share with their contemporaries, together with the institutions which make political and social life possible for them, in some sense embody notions inherited from past political philosophy and philosophies. Certainly neither political attitudes nor political behavior nor political machinery can be understood without knowledge of this kind.

These various requirements have led to differing standards for the study. Insofar as it is the record of thought about state and society, its level of accuracy has to be as high as possible. For academic historical purposes, every word of the text of Aristotle, or Marsilius of Padua, or Jefferson must be correctly registered, his intentions known, the circumstances of the writing and publication of his work discovered and recorded. But neither the conscientious citizen nor the inquiring political theorist need be much affected by the particular version of a given work which he reads, even if it is an indifferent version, clumsily translated and abbreviated perhaps, or a brief and tendentious summary in a general history. The complete book need not be known, nor the attitude of its author. It may even help if little fables are allowed to grow up around such works. The misunderstanding of one political philosopher by another, or the misreading of

authoritative books by citizens and constitution makers, has often been fruitful.

Moreover, historians of thought and of society have not been content with the role of annalists or of mere recorders of what was once written. They have sought to discover why the works were composed at all, to trace interconnections and influences covering whole generations, whole centuries of intellectual development. More recently they have been concerned to study literature in the light of ideology and to see in the writings of political philosophers especially the "reflection" or "expression" of the social structure at the time of writing, with its discontinuities, inconsistencies, and ambivalencies. Classics have come to be regarded not only as determined in this way but also as instruments in the social process, intellectual weapons in the hands of interested men and groups of men.

Although these differing motives can be distinguished in the historiography of political philosophy, individual commentators are seldom moved by one alone and often fail to see them as distinct. To this confusion must be added the unfortunate consequence of confining attention to a particular selection of authorities, a selection perhaps made originally for good philosophical reasons but which persists for reasons of convenience, curriculum, or plain conservatism. This, which is itself an example of a confusion between the interests and outlook of the historian and of the philosopher, has led to the creation of a canon of "classics" which alone go to make up "the history of political philosophy." Taken together, these circumstances are responsible for a number of persistent weaknesses in the study of this subject, some of which are listed below:

(1) The scripturalist tendency to criticize works as if their authors should have written out the final truth with complete coherence and as if, therefore, their failure to do so, their incoherencies and inconsequences must conceal some inner truth to be unraveled.

(2) The philosophizing tendency to relate the select thinkers to each other and to no others, as if contrasts between them and them alone are significant and as if they can be thought of as addressing each other. The reader's task becomes that of welding the various works into some philosophic whole.

(3) The tendency to mistake the theoretical interest of a work for its significance in other directions. This tendency is the general form of the failure to distinguish the separable interests and objectives of historians (as annalists and explainers), of philosophers, and of citizens.

(4) The tendency toward what might be called "naive sociologism": The particular circumstances of a thinker are seen as expressed in his thinking in a literal and unconvincing way, and the dominant social conditions of the present are read almost unchanged into apparently analogous conditions of the past.

Each of these tendencies can be disabling enough in itself; when they are present in combination, the results can be strange indeed. The search for Hobbist elements in John Locke, for example (tendency 2), can become an attempt to prove that he was really a Hobbist altogether and that his work on government must be examined for cryptic signs of those elements. More familiar are the exaggerations that come from stressing the relations of influence between the canonical works (tendency 2) and seeing all other intellectual elements as "anticipations" and "derivations" of these to such an extent that the relationships between bodies of thought and past societies are entirely distorted (tendency 3). Worst of all, perhaps, is a commentator who allows his thought to be so dominated by his experiences as a citizen in his own day that he betrays himself into an extreme form of the fourth tendency. When this happens, not only do Plato's or Rousseau's politics appear "totalitarian," but they are also made distantly responsible for the totalitarian proclivities of the twentieth century.

Weaknesses of this kind, however, do not necessarily deprive the commentaries concerned of their interest. In the historiography of political philosophy, as in many other inquiries, the intrusion of obvious but stimulating fallacies helps to maintain the enterprise.

GREEK POLITICAL PHILOSOPHY

The Greek city-state, or polis, gave us the word *political* and is usually supposed to have been the social organization which provided the necessary conditions for men to take for the first time a rational-critical view of the relation of the individual to the collectivity. The claim might be made that only in completely autonomous, small-scale, urban societies, like those of the Mediterranean area from the tenth century before Christ on, could an attitude of this kind develop. Because of the small size of these political entities, deliberations could take place, and decisions be made, in face-to-face discussion among all citizens, who could also see their collectivity as parallel with numerous other collectivities of the same character. It is certainly the case that the mold in which political philosophy has been set ever since is patently recognizable as Greek, and the assumption of face-to-face discus-

sion and decision persists to this day, with not entirely fortunate results.

SOCRATES AND PLATO. The issues of freedom versus tyranny, of the various forms of the state (monarchy, aristocracy, or democracy), and of the nature and operation of law are not certainly known to have been debated until very close to the time of Socrates, who was born about 470 BCE, well into the famous fifth century. The Sophists, or teachers of the art of rhetoric and persuasion for use in the law courts and in Greek public life generally, are usually credited with initiating political discussion properly defined. Although he was unsparing in his criticism of these professionals in the techniques of influence, of *sophistry* in fact, it is hard not to classify Socrates himself as a Sophist.

A determined effort has been made, by Karl Popper and others, to separate the political doctrine of Socrates, the champion of the critical discussion of dogmas and of institutions, from that of Plato, "the enemy of the open society," and their thinking has been related to the political events of late fifth-century Athens in a way which betrays many of the weaknesses described above. It seems best, however, to take Socrates and Plato as the dual spokesmen in the first known critical inquiry into the nature of the collectivity, with the peculiarity that one of them, Plato, did all the recording. The point at issue was the perennial point of how justice can be secured between men, organized as they have to be for the purposes of making a livelihood, propagating their kind, and cultivating the humane arts and accomplishments.

The answer given in Plato's *Republic,* probably composed about 365 BCE and the most powerful of his dialogues, is straightforward enough in principle, perhaps even a little banal, but it is argued on the very loftiest plane. Justice is secured only when every member of the polis is doing what he is best suited to do, and those who are best suited to do the ruling are the philosophers themselves—lovers of wisdom, those who really know. "Unless," says Socrates at the end of Book V, "either philosophers become kings in our states or those whom we now call kings and rulers take to the pursuit of philosophy seriously and adequately, and there is a conjunction of these two things, political power and philosophic intelligence, … there can be no cessation of troubles for our states, dear Glaucon, nor I fancy for the human race either."

The steps of the argument before and after this passage are by no means a matter of formal political-theoretical demonstration, and the *Republic* is at one and the same time many different treatises, a characteristic which it shares with most of its successors as classics of political philosophy. What has probably sunk deepest into the European political imagination is its utopian element, the description of an ideal condition of the collectivity when it is ruled by a select society of guardians.

The famous Platonic guardians were to be brought into the world in accordance with premeditated principles of eugenics and were not to know who their parents were. They were to live in conditions of complete communism and poverty, without privacy and outside the family; both men and women were to spend their whole lives in the service of the polis and to undergo thirty years of education—gymnastics and military training to prepare the body, music and philosophical instruction to prepare the mind. Although it is implied that the guardians would be a small minority of the whole population, and that their undisturbed rulership would ensure justice, their actual relationship with the other two elements in the polis, the soldiery and the consumers (by which term Plato presumably meant the mass of handicraftsmen and peasants, producing and consuming), is never specified. These divisions of the polis are presented as analogues of the divisions of the soul; indeed, the polis is the soul writ large. Insofar as there is a positive political doctrine in this most famous of all works of political philosophy, it seems to be hypothetical—if the polis-soul could be constructed in this way, then all problems would be solved.

Several other Platonic dialogues are concerned with political issues, and the last of them, the *Laws,* can be looked upon as the complete recasting of the Socratic-Platonic political philosophy in the light of a lifetime's reflection and experience, some of it Plato's own practical experience in advising a pupil of the Platonic Academy in the administration of the polis at Syracuse, in Sicily. But although Plato's *Politicus* (otherwise called the *Statesman*) presents an account of political life and political ideals rather different from that of the *Republic,* and although his *Laws* clashes at certain points with the *Politicus,* the ideal state of the *Republic* is that element of the political thought of Socrates and Plato which has interested posterity and influenced its thinking, almost to the exclusion of their other views.

ARISTOTLE. Aristotle, Plato's pupil, was the first of many later philosophers and thinkers who addressed themselves to the Platonic utopia, and he rejected a great deal of it. Aristotle was even more of a synoptic thinker than Plato and was much more interested in the amassing and

classification of knowledge. The gathering of information about politics and political organization was, therefore, only one of the many tasks on which Aristotle spent his extraordinarily industrious life (384–322 BCE), along with his Herculean studies of logic, psychology, biology, literature, economics, physics and other subjects. But there is evidence to show that, like Plato and other Greek thinkers, Aristotle considered politics the most important subject of all.

The Aristotelian treatises on political philosophy, the *Eudemian Ethics* and *Nicomachean Ethics* and the *Politics* itself, appear to have been based on a monumental assemblage of material of a political-scientific character, including a record of no fewer than 158 constitutions of Greek poleis. These writings had even more impressive experience behind them, because Aristotle, a Macedonian by birth, had actually been tutor to Alexander the Great, who in Aristotle's lifetime subjugated Greece and Athens. Nevertheless, Aristotle's political theory was properly philosophical, that is, it proceeded from a general view of the world and of knowledge.

He was no more disposed than any other citizen of the polis to see the individual as a reality apart from the collectivity, but he did provide a critique of the reasons why human life implied compulsory association. Man, he claimed, is a species of animal that possesses intelligence and is found in intelligently collaborative groups; therefore "man is a political animal." The natural unit of the human family forms part of the natural unit of the village, which in turn forms part of the natural unit of the polis; but the polis is not merely the family enlarged, it is an association for leading the good life, which is otherwise incapable of realization—and this means a difference in classificatory, in logical, order. States (poleis—Aristotle significantly dismisses all larger organizations as capable of ordered living only by religious means) must be judged by the extent to which they enable citizens to become virtuous and to live the good life, a life of moderation, the mean. This line of argument led Aristotle to sketch his own ideal state, but it also led him, in the *Politics,* to raise a series of crucial issues which have endured almost unchanged as decisive questions for political science as well as for political philosophy.

Probably the most conspicuous are the claims of fundamental inequality between humans: Slaves and barbarians are by nature inferior to Greeks and to citizens, although Aristotle conceded that inequality in some respects does not mean inequality in all respects. Within every collectivity, however, quite apart from the division between citizens and those incapable of citizenship, there are three classes: an upper class of aristocrats; a middle class of substantial men, mainly merchants, craftsmen, and farmers; and a lower class of laborers and peasants. The interests of these classes conflict: in sharp contrast with Plato and his anxiety for a harmony, a unity, in the polis-soul, Aristotle recognized politics as a conflict-defining, conflict-resolving activity. The actual distribution of political power among these classes—Aristotle himself insisted on the political virtue of the middle class—together with the web of manmade laws, goes to make up the particular constitution (*politeia,* the same word as the Greek title of Plato's *Republic*) of that polis. In spite of his fundamental inegalitarianism and his Greek inability to conceive of consent or representation as relevant to politics, Aristotle has often been hailed as the initiator of constitutionalism, as "the first Whig."

JUDAIC AND CHRISTIAN POLITICAL PHILOSOPHY

It is conventional to reckon the death of the polis at the death of Aristotle in 322 BCE and to believe that nothing new of importance to political philosophy appeared until the Roman Stoics evolved the universalistic dogmas of natural law. It is undoubtedly true that no systematic philosophical discussion of political principles can be traced in Judaic thought or in early Christian thought. But it is important to recognize that the symbols and the symbol system of subsequent political thinking derives from Judaic as well as from Greek sources and that its psychological assumptions are deeply tinged with Christian revelationism.

The three social institutions of the ancient Hebrews, whose significance for the history of political thinking has only recently come to be recognized, are patriarchalism, the sense of the people, and kingship. The text of the Old Testament that proclaimed the duty of obedience as the basis not only of political discipline but of all social order, including economic order, was the commandment "Honor thy father and thy mother." Throughout the Christian centuries, therefore, all questions of obedience were seen in a patriarchal context, and the political power of the Hebraic patriarch (Judah, who condemned his daughter to death for playing the harlot, or Abraham, with his fighting army of servants) was the model for the power exercised by kings and ministers. Quite as significant was the Judaic sense of the chosen people, the people led by the hand of God through the wilderness because they had an enduring purpose and being. Whenever Christian political theorists thought of the people as having a voice in the appointment of a king or a regime,

or of the king as having a duty to his people, their model was the peculiar people of Israel. European kingship was also conceived in biblical terms, and the tribal hero-king whose actions committed the people before God and whose power came from God can be seen behind the western European dynastic regimes.

Even more authoritative, of course, were the words of Jesus himself on political matters, and the few texts which could be made to bear at all upon them have been perpetually cited throughout the Christian era. Christ's submission to the Roman authority, his use of an inscription on a Roman penny ("Render unto Caesar the things that are Caesar's"), and his repeated insistence that his kingdom was not of this world made it difficult to find authority in the New Testament for any doctrine of resistance. Saint Paul's sayings pointed in the same quietist direction ("The powers that be are ordained of God"). But more interesting to the twenty-first century are those fragments of evidence from the apostolic era that make it possible to believe that Christ's immediate followers lived a communistic existence.

ROMAN STOICISM AND NATURAL LAW

The belief that there is a universal and eternal moral ordering which is common to all men and which therefore carries weight on certain issues in every collectivity is a widespread ethical and religious notion, and it need have very little specific content. Its origins have been sought in Plato's immutable Ideas and, further back, in Greek poetry. The source most often favored, however, is the religious-philosophical sect of the Stoics, who took their name from the *stoa,* or porch, before which Zeno, their reputed founder, preached and taught in Athens soon after the time of Aristotle, about 390 BCE. Stoicism was brought to Rome during the classical generations of Roman republicanism, and it continued to be a system widely accepted, although changing in content, from the time of the Scipios (about 100 BCE) until about 200 CE, when even the great Roman political families began to feel the attraction of Christianity.

The orator-statesman Cicero, although eclectic in his intellectual outlook and not usually thought of as a philosopher, wrote probably the most widely read of all works in political philosophy until recent times, *On the Laws* (*De Legibus,* c. 46 BCE) and *On the Duties of the Citizen* (*De Officiis,* a year or two later). The *Laws* was composed in deliberate imitation of Plato and was intended to complement Cicero's *De Re Publica* (his *Republic* of a year or two before), a work that was lost until 1820. *De Re Publica* contains, however, the classic text for the universalistic theory of natural law as it entered into political philosophy:

> True law is right reason in agreement with Nature; it is of universal application, unchanging and everlasting … there will not be different laws at Rome and at Athens, or different laws now and in the future, but one eternal and unchangeable law will be valid for all nations and all times, and there will be one master and one ruler, that is, God, over us all, for He is the author of this law, its promulgator and its enforcing judge. (Book III, Ch. 22, Sec. 33)

The cosmopolitan character of this doctrine—a society of all humanity ruled by one God—is in sharp contrast with the earlier Greek outlook, which assumed that only the small-scale polis could embody political good. The individual is recognizably the unit of this universal society and is the subject of the rights conferred on all citizens, all Roman citizens, by the Roman law. The identification of law with reason must be noticed in this process; reason carries its own claims to the individual's obedience. The final sanction of law and authority is placed here outside the collectivity altogether, in the Deity. Nevertheless, nothing in Stoicism could be taken as an argument against the deification of the later emperors, and one of them, Marcus Aurelius, was himself a Stoic thinker. So also was Epictetus, who began life as a slave. A rough doctrine of original freedom and equality, even the use of the contractarian model for the collectivity, has been read into Stoic texts—"All seats," so the Stoic proverb went, "are free in the theatre, but a man has a right to the one he sits down in"—but it was religious rather than specifically social equality. Much of the intellectual groundwork, in fact, of subsequent political philosophy can be sighted in the intellectual-religious tradition of Stoicism, and it is only the philosophizing tendency of historians which has prevented its attracting more attention than it has done.

ST. AUGUSTINE

The City of God (*De Civitate Dei*), written between 410 and 423 by St. Augustine, bishop of Hippo in north Africa (354–430), traditionally occupies an important place in the canon of great works on political philosophy. This extraordinary treatise raises in an acute form the problem of the historical reputation and effect of a body of thought in contrast with its actual content and the intention of its writer. *The City of God* was undoubtedly read in medieval times and afterward as the authoritative

statement of the superiority of ecclesiastical power over the secular, because it was believed to identify the visible Christian church with the mystical city of God, thought of as the bride of Christ or, even more mystically, as the body itself of the Christian Saviour. But it is very doubtful whether this was St. Augustine's intention or is even implied by his text. What is more, the conscientious political scientist finds it very difficult to decide whether *The City of God* contains any positive political doctrine at all, theoretical or otherwise.

Very recent political philosophy might, therefore, justifiably claim this work as an antipolitical classic, stating in very different terms the position sketched out by Karl Marx and V. I. Lenin as "the withering away of the state." There is the same tendency to identify all arrangements in the collectivity with evil, with the unjustifiable exercise of naked power, and the same confidence that in the fullness of time this monstrous regimentation will disappear. Moreover, Augustine was a historicist: He sought to show how God's plan to fill up the places left in Heaven when Satan and his angels revolted was being fulfilled. The creation of man and the world was intended to reveal candidates for the heavenly choir, and some few men on earth at any one time, the pilgrims (*peregrinati*), were destined at the last trumpet to be among them. They and they only were the living members of the City of God, but no one would know who constituted this select few until the judgment. It seems to have been a matter of almost complete indifference to St. Augustine how those who were to be saved behaved toward society, secular or spiritual, or what was the nature of political arrangements.

The occasion of Augustine's beginning *The City of God* was the sack of Rome by Alaric the Goth in 410, and the fall of the Roman Empire, which this event presaged, could not possibly affect the Christian who held such views about history, state, and society. The complement of the City of God was the city of the devil (*civitas diaboli*), and although it seems unjustifiable to identify the one city with the church, it seems that Augustine did quite often refer to the Roman Empire as the other. Since the heathen Romans could not possibly do justice to God and since kingdoms without justice are but great robberies (*Remota itaque justitia quid sunt regna nisi magna latrocinia?*—Ch. 4, Bk. 4), what could the Roman Empire be but thievery on a colossal scale? If by the Roman Empire Augustine implied all possible forms of the collectivity—and there are passages to confirm this assumption—then he must indeed be supposed to have had a completely negative political philosophy. Justice could never be found in any of them. In this final work of ancient political theory, then, the overriding concern is with justice, just as it had been with Socrates at the very beginning, but in it justice is viewed from an anarchist, antipolitical outlook.

MEDIEVAL POLITICAL PHILOSOPHY: POPE AND EMPEROR

Apart from the development of natural law in Christian form, the Middle Ages did not give rise to much speculation about the nature of the collectivity that has affected subsequent attitudes, nor to any great body of specifically political philosophy. Before the time of St. Thomas Aquinas in the thirteenth century, what little critical analysis there was seems to have been dominated by the Church Fathers and especially by Augustine. Although these early medieval thinkers knew of the great Greek philosophers, the actual treatises of Plato, Aristotle, and others had been lost in the West. There seems to have been a certain amount of political awareness among the subjects of the Germanic kingdoms which had come to spread over Europe, and during the nineteenth century a great deal was made of the primitive Germanic sense of community (*Gemeinschaft*), people (*Volk,* folk), and corporation (*Gesellschaft*). But unless jurisprudence is counted a part of political philosophy, neither these arrangements nor the universal social institutions associated with feudalism seem to have been the subjects of much corresponding theorization. It is remarkable how little headway the analysis of political theories in ideological terms has made with the Middle Ages.

ST. THOMAS. John of Salisbury's *Policraticus* (1159) was still Ciceronian and Augustinian in content, in spite of the fact that by his time the text of Aristotle had already reached the Latin West from the Arabs. It was left to St. Thomas to arrange the enormous access of Aristotelian information and principle in a form acceptable to a Christian Europe, which he did in his great *Summa Theologiae.* The frank acceptance of natural man—man as revealed by Aristotelian science; man not incurably maimed by sin and therefore indifferent to social-political arrangement; man whose nature is perfected, not taken away by the grace of God (*gratia non tollit naturam, sed perfecit*)—distinguished the sociology of Thomas from that of his predecessors. But although of enduring importance for politics, indeed still the final authority for the Thomist thinkers of our own day, the *Summa* and its Christian doctrine of natural law contains no developed political philosophy. For this we must turn

to the *De Regimine Principum* (*Of the Rulership of Princes*) and other works, including Thomas's commentaries on Aristotle's *Politics* and *Ethics.*

In these works St. Thomas presented his theory of the relationship between pope and emperor, which had already preoccupied Christian Europe for centuries and would continue to do so until the end of the medieval period. He developed the traditional distinction of *regnum* and *sacerdotum* (secular and spiritual jurisdiction) in Aristotelian terms, in terms of ends, the ends of humanity. "We are confronted," as A. P. d'Entrèves says, "with the doctrine of the distinction and interrelation of two great spheres of human life within one single society—the Christian society, *respublica christiana.*" But although Thomas is moderate in his claims for the pope against the emperor, although he never talks of the direct sovereignty of the pope, he is firmly convinced that all kings in Christendom should be subject to the Vicar of Christ as to Christ himself. Yet willing as he was to temper Aristotelian inegalitarianism with Christian grace, anxious as he was to give every Christian his share in the affairs of the collectivity, Thomas was absolutely intolerant of the Jew and the infidel: They remain outcasts in the Christian community.

Authority in St. Thomas's system must be legitimate, otherwise it may be resisted. An evil ruler exceeding his powers and burdening his subjects must be resisted—resisted not by the individual citizen in virtue of his individual rights (Thomas had no room for such rights) but presumably by the church. This is the sense in which Thomas's thinking has been hailed, like that of Aristotle, as the forerunner of constitutionalism.

DANTE AND MARSILIUS. The other two medieval thinkers usually accorded a place in the history of political philosophy are Dante Alighieri, the supreme poet of the city of Florence, whose political essay *Monarchia* was composed between 1310 and 1313, and Marsilius of Padua, whose *Defensor Pacis* (Defender of peace) was completed in 1324. Both were imperialists, on the opposite side of the pope-emperor controversy from St. Thomas, but both were Aristotelians. Dante's work was an idealization of the position of the medieval European emperor, who was in fact a ruler of Germany to whom the traditional trappings of the Western Roman emperor still attached as the secular ruler of all humanity, whose powers were derived directly from God and not indirectly through the pope. Marsilius approached somewhat closer to realism and had a recognizably empirical sociology: He insisted on the Aristotelian class analysis of political society and regarded the clergy as one among the classes, and therefore not in the privileged position which papal theory claimed.

The twenty-first-century observer is far more at home in the Greek polis or in a Roman province than at the papal *curia* or the court of a feudal king. So much was the medieval collectivity a religious whole, embracing not only all the territory occupied by Christians but also the whole of intellectual and cultural life, that it may be doubted whether there existed anything which corresponds to the term *state* as political philosophers ordinarily use it. Apart from the metaphysics of the papal-imperial argument, most "political thought" of the European Middle Ages is recognizable as advice to a ruler, wise reflections on commonplace situations that are entirely traditional in context and object and show no trace of the analytic attitude. Nevertheless, the medieval collectivity and the reflections of medieval theologians upon it can be appreciated under more headings than that of record.

Apart from the paradigm for the metaphysical approach to the final problem of ethics and politics provided by Thomas, the medieval situation provides the extreme example of territorial political relationships, in which the psychological mechanism usually called religious can be seen most clearly at work in providing the consensus on which such collective action as went forward had to rely. Any properly empirical account of how a collectivity in fact works, at any time, has to recognize that this mechanism is still very much in operation and that the mistake of supposing it to be replaced by rational-technical cooperation has still to be properly appreciated.

MACHIAVELLI AND REALPOLITIK

Although the polis began to lose its independence of policy as early as the lifetime of Aristotle, the towns of the Roman Empire continued to maintain a collective life that differed very little from the life of the classical polis. The decline of the cities was the outstanding feature of the fall of the empire, but they never entirely disappeared, at least in Italy. By the time of Dante and Marsilius such cities as Florence, Venice, and Milan were again in the formal position which Athens had occupied: They were independent urban communities having diplomatic relationships with each other and with the territorial monarchies. The cities possessed their own hinterlands, too, and colonies. It is not surprising, therefore, that the rational-critical attitude reappeared and that a consuming interest in ancient culture, in Plato and Aristotle, in Rome and

Greece, led to an appreciation of classical political philosophy on something like its own terms.

Nevertheless, Niccolò Machiavelli's *The Prince* (written 1513, first printed 1532), in some ways the most effective and interesting of all works of political philosophy, was in form merely one more piece of advice to a ruler. It was not presented as a philosophical work, and it contained neither abstract argument about politics nor any systematic discussion of the nature of state and society. Its analysis is confined to situations between a prince and his people and between princes (or cities) themselves. Its method is historical, the citing of significant instances. The outcome of discussion is advice, with occasional reflective aphorisms. Some of these aphorisms have become famous, and all of them show an astonishing realism and insight: "Above all a prince should abstain from the property of others; because men sooner forget the death of their father than the loss of their patrimony." "Whoever is responsible for another becoming powerful ruins himself." "Fortune is a woman and if she is to be submissive it is necessary to beat and coerce her." The headings of the twenty-six brief chapters of *The Prince* are even more significant than the sayings; Chapter 17 is titled "Cruelty and Compassion, Whether It Is Better to Be Loved Than to Be Feared."

Machiavelli's well-known answer is that it is far better to be feared than to be loved, if you cannot be both. His cool discussion of the effects of cruelty and unscrupulousness, his detached attitude toward Christianity and the traditional virtues, and his professed admiration for men of his time who are known to have been villainous and contemptible, especially the political gangster Cesare Borgia, have given Machiavelli the reputation of being the theorist of power politics, deliberate immoralism, and irresponsible, tyrannical government. But the contents of his major work on politics, the *Discourses on Livy,* have been cited to show that he was a believer in republican, not monarchical, government, and they have been used with the famous last chapter of *The Prince* itself to demonstrate that he was in fact a virtuous, patriotic Italian, worthy of the reputation he enjoyed among the English Whigs, for example, for political probity and insight. It has even been suggested, not for the first time in our generation, that *The Prince* was a satirical work. But there can be no doubt that from the time of its appearance this book was regarded as a textbook for tyrants and an exposition of the principles of power politics.

THE REFORMATION AND SECULAR NATURAL LAW

If Machiavelli's writing is looked upon as philosophical in intent, its most remarkable feature is its failure even to mention the doctrine of Christian natural law, which since the time of Thomas had dominated discussion of the nature of the collectivity and of the duties of citizens. The arrival of Protestantism raised the question of political obligation in an acute form for the first time in the history of political philosophy. It challenged a believing Lutheran or Calvinist to decide whether he should go on obeying a Catholic prince, and a Catholic subject to make the same decision about a Protestant prince. This had the effect of emphasizing, crystallizing, and codifying natural-law doctrine, since it was only under a legal or quasi-legal system of natural law that most citizens felt that they could claim a right to disobey and ultimately to resist political authority which commanded actions against their faith. Once this codification was made, systematic reflection on the philosophical problems raised by political allegiance began in earnest, and in the process natural law began to lose its exclusively religious sanction and become secularized.

It took a long time for the breakdown of universal religious consensus to have effects of this kind, even though many other influences going far back into the Middle Ages tended toward the secularization of political life. Martin Luther himself offered no systematic political teaching, certainly no doctrine of the right to resist princes for conscience sake'. In fact, in his treatise *Of Good Works* (1520) Luther wrote out traditional patriarchal rules for submission in a particularly emphatic form. John Calvin preached nonresistance too, but the religious wars in France in the later sixteenth century gave rise to a multitude of theories of the social contract that provided justification for disobedience and even for revolution on the basis of natural law. In England the Calvinists went even further, or so it seemed to the great doctor of the English Reformed church, Richard Hooker, when he sat down to write *The Laws of Ecclesiastical Polity* (written in the 1590s, first four books published in 1594 but not in print complete until 1662). Hooker believed that the claims to inspiration made by the extreme Puritans amounted to a denial of the efficacy of reason itself and to a complete rejection of natural-law principles. His response was a majestic reformulation of Thomas's natural-law philosophy that took account of the changes brought about by the Reformation, particularly of the doctrine of the final sovereignty of each individual state and its ruler, which had come to replace the ultimate

authority of emperor or pope in Christendom. The absolute sovereignty of the secular ruler, from whose decree there was no appeal, a doctrine which might be called that of ethical self-sufficiency of every political system, was given its classical expression in the *Six Books of the Republic,* published by the eminent French lawyer Jean Bodin in 1576.

Along with these developments went another that can be seen very clearly, as early as Machiavelli. This was the recognition that the body politic—the people and their political instruments, such as their parliament or their local institutions—might itself be an object of governmental action, worked on and molded by an enlightened ruler, just as the body politic might in its turn take action against government, rebel against it, replace and change its constitution. Meanwhile, secular natural law was providing a framework within which such processes could go forward and within which—as a code of international law—the various sovereign states could negotiate with one another. By the time that Hugo Grotius came to write that source book of all subsequent international law, *De Jure Belli ac Pacis* (The law of war and of peace; 1625), these relationships had come to include Islamic and Buddhist societies and societies entirely alien to the Christian point of view, even societies with no apparent belief in a deity. Natural law therefore had to become independent of Christian revelation, and Grotius stated that his principles would endure even if God did not exist. The stage was set for the first great classic of modern European, as opposed to classical ancient, political philosophy, the *Leviathan* of Thomas Hobbes (1651).

HOBBES

Although Hobbes is rightly regarded as above all a philosopher, with his own view of knowledge and of the nature of the physical world, his point of departure was political, as much as Plato's or Aristotle's was. Hobbes's declared object was "to set before men's eyes the mutual relation between protection and obedience, of which the condition of human nature, and the laws divine require an inviolable observation." This relation required the absolute submission of each individual to the dictates of an arbitrary sovereign, of "That great LEVIATHAN, or rather (to speak more reverently) of that *Mortal God,* to which we owe under the Immortal God, our peace and defence" (*Leviathan,* Ch. 17). Political science—though Hobbes did not use the phrase itself, he insisted that the proper name for the knowledge he was examining was in fact "science," on the geometrical model then beginning to take hold on men's minds—implied absolutism, despotism.

But Hobbesian political doctrine was no doctrine of the divine right of kings, nor even of one-man rule, for in this system democracies, aristocracies, and monarchies should all equally be absolute sovereigns, whose every dictate is law. Monarchy was to be preferred, as might be expected, and democracy, "the government of a few orators," was least desirable. The power of government is a part of the divine providence, but its sanctions are much more tangible. They rest on the unqualified alienation of all the rights of every individual into the hands of the sovereign at the time of the making of the social contract—of compact, as Hobbes called it—and thereafter every attribute of every citizen, even his property, depended on the sovereign's will. So anxious was Hobbes to remove any possible grounds that might be used to justify resistance to authority that he advanced two positions entirely unacceptable to most of his contemporaries. One was the reformulation of natural law in a form that gave no rights to the citizen and the other was to confer on the sovereign the function of pronouncing on the interpretation of Scripture itself.

Perhaps the most famous element in the Hobbesian system was the account of the state of nature, and the best-remembered passage reads:

> during the time men live without a common power to keep them all in awe, they are in that condition which is called war; and such a war, as is of every man against every man.… In such condition, there is no place for industry, because the fruit thereof is uncertain: and consequently no culture of the earth, no navigation, nor use of the commodities that may be imported by sea; no commodious building; no instrument of moving, and removing such things as require much force; no knowledge of the face of the earth; no account of time; no arts; no letters; no society; and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish and short. (*Leviathan,* Ch. 13)

If this fighting anarchy is in fact the natural state of man, then it does seem to follow that the only possibility of cooperation in the collectivity is by absolute submission, and every human value must depend on the existence and efficacy of "the great Leviathan." The law, or rather the laws, of nature did exist at that repulsive stage of human development but only as rules of prudence, for "Reason suggesteth convenient articles of peace, which otherwise

are called the laws of nature." Whatever the status of these principles, they could not possibly be used to justify resistance to the sovereign, although Hobbes did provide for the transfer of allegiance to another sovereign when the one established can no longer provide protection. He also allowed to the individual the right to refuse to confess to a crime or to take his own life. The appeal to revelation and to conscience, which Hobbes believed was responsible for the political instability of his own time, and especially for the Puritan rebellion in England, was completely precluded by his interpretation of the claims of his sovereign.

In spite of Hobbes's confident belief that his elucidation of the true principles of political science would resolve conflict, his work aroused immediate opposition and has given rise to unending controversy. There is first the question of whether his state of nature, succeeded by a covenant, or social contract, was intended to be taken literally as a historical and anthropological claim, or whether it was simply hypothetical. A recent ideological interpretation has claimed that the state of nature was hypothetical but that the aggressive, competitive emphasis arose from Hobbes's observing the possessive individualism informing the increasingly capitalist society in which he lived. The second question concerns the continuity between his state of nature and his state of society. How could men with the characteristics Hobbes gives them ever form themselves into a collectivity? A third question is whether he ever intended men to be morally obliged to obey the sovereign, or, if this was his intention, whether he succeeded in tying them down ethically. A further question is how far he was indeed abandoning the whole natural-law position and advancing an entirely utilitarian political ethic; men obey always and only because they see it is to their advantage.

WHIG CONSTITUTIONALISM AND LOCKE

Hobbes was not the first writer to invoke what came to be called the "pleasure-pain principle" in political discussion, and his radical contemporaries, the Levellers of the English Civil War, also made claims which seemed to rest on strictly utilitarian grounds, although in an unphilosophical and unsystematic form. The appearance of writings of this character, which have claims to be the first emanating from the common man, raises an important issue about the career of political philosophy from the seventeenth century on. The Levellers were democrats, and the political rights they claimed were meant to be exercised by a far greater proportion of the population than ever had been previously contemplated, even by the English Parliamentarians locked in their struggle with the house of Stuart. It has been recently and justifiably questioned whether all individuals were intended to be covered by Leveller declarations, or even all male householders, but from that time on, there is a recognizable class content in the doctrines of the political philosophers. Until the late eighteenth century most thinkers continued to share the universal assumption that "citizen" must be confined to the fully literate, propertied, elite minority, but they showed an increasing awareness that this was a tiny minority and that the right of this minority to stand for the whole might need justification.

Paradoxically enough, this crucial question was raised in an awkward form by one of Hobbes's exact contemporaries, Sir Robert Filmer, a traditionalist rather than a progressive. Sovereignty is a patriarchal matter, Filmer claimed, a matter of natural subordination, and unless this is recognized, the inequality of distribution of property and the subjection of poor men, men without the vote, servants, and women could never be justified. Much of Filmer's thinking, and that of the commonsensical Englishmen who came to accept his authority, is present in the writing of Hobbes. Nevertheless, for historical reasons it was against Filmer rather than against Hobbes that in the years 1679 and 1680 Locke wrote out the classic statement of Whig constitutionalism and government by consent, *Two Treatises of Government* (revised and published in 1689).

This modification of the accepted account of the relation of Locke to Hobbes is due to very recent scholarship, and the same evidence goes to show that the work of Benedict de Spinoza, the only immediate follower Hobbes had among philosophers, was more of an intellectual preoccupation for Locke than Hobbes ever was. Spinoza (*Tractatus Theologicopoliticus,* 1670; *Tractatus Politicus,* 1677), if easily the least influential, was in some ways the most engaging of all the political thinkers of the early modern age in Europe. Unfortunately, we cannot dwell here on his modification of the Hobbesian system; his overt insistence that the contract was hypothetical; his specific insistence that all obligations had to be utilitarian, based on self-interest; or his attempt to ensure that the enlightened sovereign must seek the welfare of his people.

Locke's *Second Treatise,* with its subtitle *Of Civil Government,* seems to have been the first composed of the two, and it begins with the following assertion against Filmer's claim that all men are born unfree, unequal, and in patriarchal subjection:

> To understand political power right, and derive it from its original, we must consider what state all men are naturally in, and that is, a state of perfect freedom to order their actions and dispose of their possessions, and persons, as they think fit, within the bounds of the law of nature, without asking leave, or depending on the will of any other man. A state also of equality, wherein all the power and jurisdiction is reciprocal, no one having more than another. (Sec. 4)

The law of nature, then, was real, and it governed all men in the peaceable condition which preceded the foundation of the collectivity, when order was maintained by what Locke called "the executive power of the law of nature" in the hands of every man. This law of nature gave men tangible rights, even before the contract. It ensured them the right to their religious opinions (not argued for, or even mentioned, in the work on government but in a succession of *Letters on Toleration,* the first published in 1689); it guaranteed them the right to property, whose acquisition was brought about by men "mixing their labour" with the goods of nature; it made it legitimate for every person to take some political responsibility and in due course to act as sovereign himself or as part of the sovereign power, for the vital political right was that of insisting that government rested on the consent of the governed, the consent of the majority expressed constitutionally through representation. The stage of contract came about because the predominantly peaceful state of nature was liable to war and because property was insecure under it. When it arrived, political power was "a right of making laws for the regulating and preserving of property, and of employing the force of the community, in the execution of such laws, and in defence of the commonwealth from foreign injury, and this only for the common good" (*Second Treatise,* Sec. 3).

Contract, to Locke, was an agreement to pool the natural political virtue of individuals and to establish a sovereign power thereby which was in a perpetual trust relationship with the people. If the trust was broken, the people had a right to cashier their governors and put others in their place or, if necessary, to alter the constitution, and all this without the return of the state of nature. In this sense, and in allowing a final appeal to God if the compact itself was dissolved, Locke can be said to have held to a doctrine of the sovereignty of the people and to a perpetual reserved right of revolution. He believed in a form of the separation of powers and in the rule of majorities, but he shows little sympathy with representative democracy.

Recent studies have shown that Locke's political philosophy, as contrasted with his general philosophy, was much less influential in the eighteenth century than had been supposed. Nevertheless, the Lockean outlook, along with that of his friend and contemporary Sir Isaac Newton, must be counted as the point of departure of the intellectual movement known as the Enlightenment.

THE ENLIGHTENMENT AND MONTESQUIEU

Locke could not deal adequately with Newtonian mathematics, but in spite of the intellectual barrier between them, the two men shared one passionate curiosity: to know all that could be known about societies, customs, and religions outside Europe. Confidence in the efficacy of mathematico-physical methods to solve all problems, including those of social and political organization, and cultural relativism leading to doubt about religious revelation and the necessary value of any familiar institution underlie much Enlightenment thought. Meanwhile, the steady spread of literacy and the consequent growth of the size of the politically conscious, curious, and ambitious community, especially in France and England, was changing the conditions of political and social speculation.

The result was a proliferation of works of political philosophy which from now on defeats the summary historian. Sir Isaiah Berlin has said that "the conflict of the rival explanations (or models) of social and individual life had by the late eighteenth century become a scandal." Except as a critical movement, compelling all established dogma to give an account of itself, the Enlightenment cannot be called a uniform current of thought at all. Of the multiple works of Voltaire, Baron de Montesquieu, David Hume, Claude-Adrien Helvétius, Adam Ferguson, Jean-Jacques Rousseau, Gabriel Bonnot de Mably, D'Argenson, Richard Price, Thomas Paine, Thomas Jefferson, Edmund Burke, and their successors, we can comment here on only one or two that find a place in the traditional canon.

MONTESQUIEU. Charles Louis de Secondat, Baron de Montesquieu, may serve as the example of the early sociological attitude, presented with great literary skill and at considerable length in his *Esprit des lois* (in preparation from 1734, published 1748). To Montesquieu, who sought to examine and record social uniformities, natural laws describe necessary human behavior, and because they are necessary, they also oblige men ethically, or, rather, they are the basis of legal systems which men are

morally obliged to obey. At this point it is usual to say that Montesquieu's attitude touches that of Hume in his *Treatise of Human Nature* (1739), containing his famous aphorism about all systems of morality imperceptibly changing from propositions containing "is" and "is not" to propositions containing "ought" and "ought not." But the French author's interest was not in obligation as such; rather, it was in the structure of the collectivities which men find themselves obeying and in the ways in which these structures or their "spirits" (*esprits*) express environment.

ROUSSEAU AND THE GENERAL WILL. Montesquieu is scarcely representative of the most characteristic feature of the political philosophy of his age, at least when viewed from the somber century we now inhabit, because he was neither an optimist nor a believer in the perfectibility of man. Rousseau was skeptical of progress too, for in some moods he seems to have believed that human nature had once been perfect but had been corrupted by society. This was the position which he defended in his first *Discourse* (1751). In his second *Discourse,* the *Discourse on Inequality* (1755), not society but property was the evil attacked.

Neither of these works contained Rousseau's specific contribution to political philosophy. In the *Social Contract* (*Du contrat social,* 1762) Rousseau elaborated a doctrine that was both original and potentially revolutionary; the relation of the individual to the collectivity was seen as a matter of will, not of agreement, and the solution of the problem of obligation was the discovery of a general will directed to universal moral ends, which the individual had only to obey in order to secure justice. Rousseau presented the general-will model in individualistic, contractarian terms:

> Man was born free, and everywhere he is in chains. What is it that can make this legitimate? … The moment men leave the state of nature and set up society, that act of association brings into being a moral, collective body in the place of the particular persons of each contracting party, composed of as many members as there are voices in the assembly, which from this same act receives its unity, its common personality (*moi commun*), its life and its will. This passage from the state of nature to the state of society produces a very remarkable change in man, in substituting justice for instinct in his conduct, and giving to his actions the morality which before they lacked. (*Du contrat social,* Book I, Chs. 1 and 6)

In spite of the care that Rousseau took to effect a moral reconciliation of the will of the individual and that of society, the collectivist possibilities of his approach to political obligation are evident. Since he insisted that a collectivity which has no general will is unworthy of the obedience of its citizens, its revolutionary potentialities are also obvious. The most conspicuous element supporting the interpretation that the *Social Contract* is a statement of tyrannical revolutionary nationalism is its final chapter, "The Civil Religion," which can be interpreted as justifying the condemnation to death of anyone who flouts Rousseau's own dogmatic statement about the relation of the individual to the state.

***THE FEDERALIST,* BURKE, AND PAINE**. The supposed direct relationship of Rousseau's thinking with the revolutionary movements of the late eighteenth century, particularly with the American and French revolutions—even with the Reign of Terror and the despotism of Napoleon Bonaparte—is a conspicuous example of that interplay between intellectual speculation and political movement in which both citizens and historians seem to want to believe. It is of course doubtful whether any element from the multifarious theorization about politics which went on during the Enlightenment could ever be shown to be causally related to what happened in France after 1789, and it is certain that the rebelling American colonists took little trouble to justify their actions in philosophical terms. Nevertheless, the foundation of the American political attitude is of importance to political philosophy, and *The Federalist* (written jointly by James Madison, Alexander Hamilton, and John Jay in the form of a collection of papers published in the New York press in 1787 and 1788) is an outstanding instance of a book's being taken as a compendium of the theoretical content of a nation's political outlook. Max Beloff has said that the sociology of this work was static; in their day there had been founded in America a society, a prefabricated, premeditated structure that would endure unchanged forever. It had the characteristic common to all ethically justified institutions: "Justice is the end of government. It is the end of civil society. It ever has been and ever will be pursued, until it be obtained, or until liberty be lost in the pursuit." But justice is not the imposition of equality—it is the protection of the weak against the stronger. Government will otherwise be content to hold the ring, and liberty will be ensured by the separation of the powers and by the balance between the state and federal governments.

Edmund Burke was a champion of the Americans against the arbitrary powers of the British crown, and he

must have approved of much of the argument of the *Federalist,* especially that concerning the benefits of unequal distribution of property. The exercise of political power was the greatest challenge to the wisdom and responsibility of an individual and to his capacity to decide weighty issues on behalf of others. Where were such men to be found but among those experienced in the proper administration of great possessions and of the people who went with them?

Each of Burke's voluminous writings on politics, which occupied his whole life, contains a remark or two of importance to the philosophy of politics. But the work that has caught the eye of posterity is the one he wrote in horrified protest against the actions of the French revolutionaries, *Reflections on the Revolution in France* (published in 1790). The famous passage remembered from this book goes as follows:

> Society is indeed a contract. Subordinate contracts for objects of mere occasional interest may be dissolved at pleasure—but the state ought not to be considered as nothing better than a partnership in a trade of pepper and coffee, calico or tobacco, or some other such low concern, to be taken up for a little temporary interest, and to be dissolved by the fancy of the parties. It is to be looked upon with other reverence; because it is not a partnership in things subservient only to the gross animal existence of a temporary and perishable nature. It is a partnership in all science; a partnership in all art; a partnership in every virtue, and in all perfection. As the ends of such a partnership cannot be obtained in many generations, it becomes a partnership not only between those who are living, but between those who are living, those who are dead, and those who are to be born. Each contract of each particular state is but a clause in the great primeval contract of eternal society, linking the lower with the higher natures, connecting the visible and invisible world, according to a fixed compact sanctioned by the inviolable oath which holds all physical and moral natures each in their appointed place. (pp. 163–164)

The extravagance of the language and the lamentable vagueness of the statements are typical of Burke, and typical also of the uncritical acceptance of the contractarian model long after it had become unnecessary. Indeed, Burke's account of obligation, insofar as he presented one at all, was far closer to Rousseau's general-will argument than he would have admitted.

But the phrases that have interested posterity are those that limit the freedom of each generation to act against the expectations of the past and the interests of the future, and those in which he condemns as immoral the action of any society which allows fundamental revolution. It was an offense against all humanity to act as the French revolutionaries were doing. The very language of abstract natural right was excoriated by Burke, and he challenged all subsequent political thinkers with the problem of the status of political principles in relation to political action and practice.

Burke's effusive, skeptical conservatism was too much for Thomas Paine, his acute Anglo American contemporary, whose *The Rights of Man* (Part I, 1791, a direct answer to Burke) is often acclaimed a minor classic of political philosophy. There has been no writer more optimistic about the effects of violent political action, or more indifferent to the existence of established government. "The instant formal government is abolished, society begins to act. A general association takes place and common interest produces common security." But in the second part of *The Rights of Man* (1792) Paine identified himself with the nascent working class, and added to the responsibilities of government policies that were hitherto scarcely contemplated and are hailed in our day as the first discernible sign of welfare legislation, even down to family allowances and maternity benefits. The talk of property, representation, and the will and wants of all, which had increased steadily since the time of Hobbes, had issued at last into something like universalistic claims for participation in political activity, into that "numerical democracy" which has characterized the industrialized world ever since.

THE UTILITARIAN TRADITION

BENTHAM. "It is the greatest happiness of the greatest number that is the measure of right and wrong." This famous tag appears in the second paragraph of Jeremy Bentham's *Fragment on Government* (1776) and may be looked upon as the original formulation of the utilitarian principle for specifically political purposes, although Bentham had the law in mind. (Utilitarian ethics of course goes back as far as Hobbes, and Bentham's use of it may be directly referred to Hume.) Bentham went on to offer a definition of the collectivity which was followed more or less faithfully by all his successors in the utilitarian tradition: "When a number of persons (whom we may style subjects) are supposed to be in the habit of pay-

ing obedience to a person, or an assemblage of persons, of a known and certain description (whom we may call governors) such persons altogether (subjects and governors) are said to be in a state of political society."

The unsatisfactory character of crude utilitarian ethics is plain in Bentham's best-known book, the *Introduction to the Principles of Morals and Legislation* (1789). "It seems to me," John Plamenatz has said of this work, "that Bentham, without quite knowing what he is doing, is trying to reconcile two couples of irreconcilable doctrines; egoistic hedonism with utilitarianism on the one hand, and a psychological with an objective theory of morals on the other." But in clarifying legal principles and in giving directions to lawyers and politicians, Bentham was much more effective, perhaps the most effective writer of principle for the purpose of advice. So anxious was he to make it crystal clear what men should do tomorrow that he went so far as to proclaim that the motives from which men act are morally irrelevant; only the consequences matter. Carrying out this advice made Bentham into an advocate of the doctrine that government is a necessary evil, since all that government can do is to coerce, and coercion must be kept to that minimum (Bentham's coinage) which will prevent even greater pain. In this way, with Paine as well as with Bentham, utilitarianism was used to justify equality between citizens and representative democracy.

J. S. MILL. The logical difficulties of utilitarian ethics and the possible dangers of numerical democracy—leaving every man to make up his mind about his own and the general happiness and giving him an equal right to a part in decisions about them—are also evident in the classic statement of liberalism, John Stuart Mill's *On Liberty* (written 1854, published 1859). It was followed in 1861 by *Utilitarianism* and *Representative Government.*

Mill's *On Liberty* shares some of the social unreality that is so evident in Bentham's definition of the collectivity, but to a very much smaller degree. "Wherever," says Mill, "there is an ascendant class, a large portion of the morality of the country emanates from its class interests, and its feelings of class superiority." In his later life Mill might well have described himself as socialist. But the doctrinal legacy of his text is very different:

> The object of this Essay is to assert one very simple principle, … that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection, … to prevent harm to others.… The only part of the conduct of any one, for which he is amenable to society, is that which concerns others.… The only freedom which deserves the name, is that of pursuing our own good in our own way, so long as we do not attempt to deprive others of theirs. (*On Liberty,* Ch. 1)

This principle of other-regarding actions being distinguished from self-regarding actions, and being alone amenable to control from outside, is one of extreme difficulty in practice but of great convenience in argument. With it goes a deep suspicion of the "tyranny of the majority," not simply as expressed in governmental action but even more in the form of intolerant conformism of opinion. Mill is at his most persuasive when he argues that "all silencing of discussion is an assumption of infallibility" and when he insists that it is to the universal advantage that the truth should be known. His book may be regarded as the most forceful of all pleas for freedom of thought and expression. He ends it by insisting on three very general reasons against "government interference." States should not do things better done by individuals, things which it is better for the individuals to do themselves, and things which might unnecessarily add to governmental power.

SIDGWICK. Mill was by no means the last of the utilitarian thinkers, although the positive grounds for freedom and justice put forward by the idealists were already beginning to replace the negative arguments summarized above. Henry Sidgwick's *Elements of Politics* (1891) may be taken as the final statement of political utilitarianism, although in its later editions it is marked by repeated concessions to socialism, always referred to in quotes. Sidgwick's definition of the collectivity is still Bentham's, although he admits that the principles of politics are not absolutely true but are based on psychological propositions approximately true of civilized man. He adopts from the great utilitarian jurist, John Austin, the claim that in every state the legislature must be legally unlimited, but he also qualifies this. He comes down emphatically on the side of individualism, "which takes freedom—the absence of physical and moral coercion—as the ultimate and sole end of governmental interference."

GERMAN IDEALISM

KANT. The general-will model associated with Rousseau underwent some development at the hands of the great German philosopher Immanuel Kant in various works written in the 1780s and 1790s. His idea of a "general and

public will" is not a particularly lucid concept, but it does express for political purposes the supreme ethical principles of the Kantian philosophy that each man should treat each other man as if he were an end, never as a means, and that each act should be such that it might become a universal law. V. F. Carritt has also praised him highly for the recognition that obligation is a condition of political societies, not a product of them. More influential for subsequent political philosophy, however, was Kant's theory of history. In the course of this complex argument he proposes that the attainment of political society which shall enforce justice requires that man have a master to force him to be free and that this master be the will of the community.

HEGEL. Most philosophers have tried to bring to bear on the problems of political philosophy an overall view of the world and of knowledge. No philosopher has been so devoted to system and the whole as Georg Wilhelm Friedrich Hegel. Political philosophy has its appointed and necessary place within the dialectic exposition of reality. Reality is spiritual, the Absolute, and collectivities have their part to play in the teleological "unfolding" of the Absolute. Collectivities—the family, "civil society," and the state—are manifestations of objective spirit, and the state is the culmination of objective spirit. Collectivities arise when the manifestation of objective spirit in the individual reveals itself as inadequate. The individual can be truly himself only in some society. Formal ethics is bare and empty, and it must be made concrete. Concrete ethics can only be social. Thus the family is a dialectical necessity.

But the family is not a permanent institution; although the members of the family are united in the family and hence are one, the children grow up and leave the family. This "negation" of the family is negated in a new collectivity, civil society. Civil society embraces the economic order and the economic organizations and institutions through which it is expressed, as well as the legal system and the enforcement facilities necessary to it. But the legal system implies something over and above civil society, namely, the state, without which a legal system is impossible. Family and civil society are both embraced within the state; they are at the same time fulfilled by it and manifestations of it. The same is true of the individual. In the state the individual rises above his mere particularity to become a person and truly free.

What the concept of a state fully embraces can be known only through the historical development of actual states. Among the many possible forms of the actual state, the most rational is a monarchy. A corporative state, in which individuals participate in governmental affairs by virtue of their standing in the corporative bodies of civil society rather than as individuals, is more rational than representative democracy, in which individuals are represented merely as individuals. Nevertheless, the constitution which is best for any particular state is that one which has developed slowly in that state over the course of centuries. A constitution imposed artificially is bound to fail.

It might seem that Hegel's conceptual scheme would require that the state be embraced in some other form of collectivity, but this is not the case. The state is the highest form of objective spirit, and, at this point of the dialectic, objective spirit is negated by absolute spirit—the realm of art, religion, and philosophy. Thus Hegel rejected the Kantian notion of a federation of states and regarded war as not only natural but the motive force of history.

GREEN AND BOSANQUET

The meaning and implications of Hegel's political philosophy provoked immediate and lasting controversy. The central points of discussion have been the relation of the individual to the collectivity, whether state, society, race, or nation; the meaning of the notion of state; and the application of dialectic to the discovery of a necessary pattern in political history. The first point was the dominant problem of the social thought of the British idealist philosophers of the later nineteenth and early twentieth centuries. In political philosophy the two chief figures, with rather opposed views, were T. H. Green and Bernard Bosanquet. Green undertook the task of updating British liberalism to meet the changing circumstances of a rapidly industrialized society. To do so, he sought to divorce liberalism from the ethical egoism of utilitarianism and the laissez-faire economic doctrines of David Ricardo and to replace them with an idealist theory of society based broadly on Kant and Hegel.

For Green, as for earlier liberals, the effect upon freedom was the criterion by which a piece of legislation was to be judged. Did it tend to enlarge or to restrict freedom? Green held that Benthamite liberals had arbitrarily identified freedom with absence of legal restraint, implying that any piece of legislation must necessarily restrict freedom. Green pointed out that it had become evident that a person could be legally free and still not have the power to act for his own benefit. Where one party to a contract has all the powers of coercion on his side and the other party cannot help but agree to the terms proposed by the

first party, then the state has the right and the obligation to interfere to restore the original freedom. There are other restraints on freedom than those imposed by the state.

Nevertheless, freedom was not, for Green, a natural right, for he held that there are no natural rights in the eighteenth-century sense. No one possesses abstract rights independent of his membership in a society in which the members recognize some common good as their own ideal good. Thus Green, more a Kantian than a Hegelian, held that the basis of all political obligation is the moral obligation to treat the other members of one's own society as ends in themselves, as having wills whose realization should not be interfered with. The state, on Green's view, has the duty to foster the conditions that permit each member so to act, and to lead him to regard and treat the other members as ends. The members in turn obey the state because they recognize it as the embodiment of their common right.

Green's liberalism stressed the positive function of the state in supporting the moral well-being of all its citizens, and it was not far from the Fabian conception of a national minimum of physical well-being below which the state should not allow any of its citizens to fall—for otherwise they could not participate fully as moral and political beings in society. The liberal side of Green's thought has greatly influenced British political philosophy, which has tended to remain idealist or partially idealist long after idealism passed out of fashion in other areas of British philosophy. But it has been certain antiliberal tendencies which have come to be generally thought of as most typical of idealist political thought, especially since the publication of L. T. Hobhouse's *The Metaphysical Theory of the State* (London and New York, 1918). This work was a direct attack on Hegel and on Bosanquet's *Philosophical Theory of the State* (London, 1899).

Bosanquet developed the notion of the relation between individual and society beyond Green's claim that individuals are individuals only insofar as they are social. He claimed that society itself is more real and more of an individual than any of its members can ever be. And within each member of society it is the social self, rather than any purely individual desires or aims, that is most real. The social self is somehow identical with society, and thus social coercion is coercion by the higher, social self of the lower, individual self. In short, social coercion is self-mastery and true freedom.

Hobhouse charged that this revival of Rousseau's (and Locke's) notion that a man can be forced to be free is in itself dangerous and illiberal. He further charged that this notion, combined with Bosanquet's failure to distinguish properly between society and the state, or indeed to give any clear or unambiguous definition of the state, leads to the doctrine that the state can do no wrong, and hence to the justification of almost any action on the part of the government in power. There is no doubt that idealist claims have in fact so been used; however, Bosanquet held not that individual governments can do no wrong but that they can do wrongs of a kind totally different from those which individuals can commit—a government can confiscate property, but it cannot commit theft. And individual states can be judged by how well or poorly they fulfill the functions of a state.

MARX AND MARXISM

The Marxian development of Hegelianism is of an entirely different order from the academic philosophies of Green and Bosanquet. The difference is epitomized in Marx's famous eleventh thesis on Ludwig Feuerbach: "The philosophers have only *interpreted* the world, in various ways; the point is to *change* it."

Karl Marx, the great theoretician of socialism, applied the Hegelian dialectic of history to the Hegelian analysis of collectivities. Hegel's family, civil society, and state are not three eternal ideas partially or imperfectly manifested at all periods of history. Rather, they are abstractions from the particular socioeconomic arrangements of the period in which Hegel and Marx lived. Hegel was right in stressing the central role of the economic function in civil society and in holding that, as now constituted, civil society requires a police power and hence a state. But he failed to see that civil society is not necessarily the same as capitalist, bourgeois society (civil society and bourgeois society are designated by the same phrase in German), and he did not see that those who determine the economic arrangements of society are not abstract individuals but are those who exercise control over the economic resources and forces available at the time. Since all others are excluded from having a voice in these economic arrangements, the result is class divisions and the need for the dominant class to defend its economic and political position against the other classes. Thus, as Hegel said, the state is necessary, but it is necessary as an instrument of the oppression of one class by another and not as something inherent in the very notion of social life. If class divisions were done away with, then there would be no one to oppress and the state would disappear. Civil society would be all that there was.

Marx, of course, believed that although in all previous periods (except for an initial period of primitive communism) the state had been necessary, the economic forces of capitalism had so developed that it was not only possible but also necessary for the state to disappear. The complexity of previous class divisions was becoming polarized into two antagonistic classes: the bourgeoisie, who controlled the instruments of production, and the proletariat, who had no choice but to work for the bourgeoisie at subsistence wages. Once the proletariat rises up and takes over the means of production from the bourgeoisie, there will be no more classes to oppress. In the classless society the state, the government of persons, will be "replaced by the administration of things and by the conduct of processes of production" (Friedrich Engels, *Anti-Dühring,* Moscow, 1962, p. 364).

Three intellectual tasks emerge from this view of the historical situation: a study of the laws according to which one era passes into another; a study of the present bourgeois era to discover in it those forces and movements tending toward its breakup and the emergence of the inevitable next era of the classless society; and some sort of preparation and anticipation, however blind, of the period of transition and its aftermath. Thus, economic history and political sociology become pressing practical subjects, and the central problem of politics becomes that of revolution.

The problem of justifying revolution had often been raised before. For Marxists, justification is no longer in question; revolution is inevitable, and only its date is unknown. Marxists must know how to bring about a revolution, whether it must be violent, and whether the revolution can be hastened if the productive forces are not yet ripe. Marx was sure that the bourgeoisie would not yield power without a struggle and that the revolution must be violent. He also held that it could not be hastened: "No social order ever disappears before all the productive forces for which there is room in it have been developed" (*A Contribution to the Critique of Political Economy,* translated by N. I. Stone, Chicago, 1904, preface).

Those later developments of Marxist thought that have been serious and not merely propagandistic justification of a position have generally been attempts at adjusting or revising the theory of revolution to changing historical situations—the growth of mass socialist parties with the apparent possibility of their coming into power by peaceful means; abortive revolutionary governments like those of the Paris Commune of 1870 and the soviets of workers and peasants of the Russian revolution of 1905 (both as interpreted somewhat mythically by Marxist writers); the rapid succession in 1917 of a bourgeois revolution in Russia by a proletarian one before all the possibilities of the bourgeois era could come to flower; the conspiratorial character ascribed to that proletarian revolution; the imposition of socialist regimes in Eastern Europe by Soviet intervention; and the greater or lesser success of Marxist-inspired revolutions in countries, notably China, where modern bourgeois capitalism had only the most tenuous foothold. These revolutions in countries with precapitalist economies were totally inexplicable on classical Marxist grounds, and interpretations of them generally rely on some variant of Lenin's doctrine that in the latter part of the nineteenth century capitalism developed into a higher, final phase of international imperialism, with a corresponding internationalized proletariat and an interaction between the proletariat of the imperialist states and of the populations of the colonies.

ANARCHISM

Socialism, both Marxist and non-Marxist, has since the time of Marx generally favored some sort of centralized control at least of economic life, despite the Leninist prominence given to Friedrich Engels's phrase "the withering away of the state." Although in general it has been held impossible to predict the exact character of a communist society, it has not been claimed that there would be no central authority. In opposition to this collectivist view were most of those early socialists whom Marx classified as utopian, as well as the anarchists and the later guild socialists, such as G. D. H. Cole.

The anarchists differed enormously in their attitudes toward social and economic arrangements, especially in their attitudes toward the institution of private property, but they were united in their opposition to the state, and hence to any centralized authority and to any participation in governmental functions. Engels expressed the Marxist's difference with the anarchist ideal succinctly:

> In this society there will, above all, be no *authority,* for authority = state = absolute evil. (How these people propose to run a factory, operate a railway, or steer a ship without a will that decides in the last resort, without single management, they of course do not tell us.) The authority of the majority over the minority also ceases. Every individual and every community is autonomous, but as to how a society of even only two people is possible unless each gives up some of his autonomy Michael Bakunin again

maintains silence. (Letter to Theodor Cuno, January 24, 1872)

The anarchists see the primary fault of the present economic order not in the economic arrangements, as do socialists, but in the existence of the state. The state is to be overthrown (although many anarchists, despite the popular identification of anarchism with terrorism, would stop short of violence), and then society will take care of itself. The actual order that will emerge is variously pictured as anything from an extreme individualism to voluntarily cooperating groups of various sizes. Marxists deny this primacy to the state, which, they hold, will collapse when the economic order of which it is the instrument collapses.

Ideas resembling the doctrines of the anarchist thinkers can be found in writings of various periods from Greek times onward, but the first fully articulated anarchist theory is to be found in William Godwin's *Enquiry concerning Political Justice* (1793). Like later anarchists, Godwin was as much an ethical writer as a political theorist. All social organization, and especially all governments, are necessarily corrupting. Society creates prejudices—preconceived ideas. We see people in terms of their social function and status rather than as individuals, and we judge in terms of false ideals—honor in a monarchy and public-spiritedness, a concern for the good of the state rather than of the individual, in a republic. Neither is a substitute for the ideal of benevolence. Godwin's solution is a small, classless community without rules in which individuals cooperate without compulsion, out of friendship, understanding, and benevolence.

Pierre-Joseph Proudhon, a self-educated Besançon printer, was the first theorist to describe himself as an anarchist. Despite his famous definition, "Property is theft," Proudhon was not against property as such but only against its unequal distribution. His ideals were equality and independence. As political science discovers the natural laws according to which society functions, then the arbitrary laws of governments become unnecessary. Proudhon favored individual ownership of the means of production by peasants and artisans. As political science revealed their mutual interests to them, they would freely join together in an ever-widening system of interlocking economic contracts that would make government unnecessary. Only in the case of some large-scale industries and public utilities would workers' syndicates be necessary.

With Bakunin anarchism became associated with the nineteenth-century revolutionary tradition. The son of a Russian nobleman, Bakunin was involved in a number of revolutionary movements from the 1840s on, took part in abortive revolutions in France, Prague, Dresden, and Bologna, and was imprisoned in Saxony, Austria, and Russia. Bakunin was influenced by Proudhon but also by Hegel, Comte, Arnold Ruge, Charles Darwin, and Marx. Like Proudhon, he held that what is produced should be distributed according to the amount of labor the recipient has provided, but he differed in advocating public ownership of the means of production. He differed from Marx in advocating the early destruction of the state rather than its seizure by the workers.

Another Russian writer, Prince Peter Kropotkin, was also influenced by Proudhon. His chief differences from Proudhon and Bakunin were that he favored the small local community as the unit of social organization and argued that goods should be distributed on the basis of need rather than on the basis of what the recipient had produced. Thus he envisaged warehouses where goods would be distributed freely rather than earlier schemes of distribution based on some measure of the recipient's production. Kropotkin also tended to stress the notion that man is naturally social, which was a factor in earlier anarchist theories, even going so far as to find that cooperation, and not merely competition, is a factor in animal evolution.

Far too complex in his views to be classed merely as an anarchist is the French philosopher Georges Sorel. Sorel is important less for his programmatic views than for his analysis of social systems into consumers' and producers' societies, each with its own system of morality, and of the roles of violence and of political myths in revolutionary movements. In a consumers' society the good is things to be obtained—welfare, prosperity, distributive justice, or the classless society. The consumers' society is based on envy. A producers' society sees the good in the cooperative creative endeavor of self-reliant individuals. But this creative endeavor tends in the end to decay into a consumers' society. Violence is a sign of moral health in a revolutionary movement. It ranges from a violence of principles to, occasionally, physical violence. It is intended as much to discourage the "reasonable" sympathizer who feels the time is not ripe for revolution and the man of good will seeking reconciliation as it is to intimidate the enemy. A myth is the revolutionary morality stated in terms of a hoped-for future. Thus, the notion of the general strike may be self-contradictory, but this is beside the point. It is not scientific prophesy but the expression of the aspirations of the revolutionary masses.

FASCISM AND NATIONAL SOCIALISM

Marxism and anarchism are representative of a modern tendency to see political arrangements in terms of a program and often of one dominant idea. There have been others, notably racism and the various forms of nationalism, but only two can be mentioned here. Like Marxism, fascism and national socialism were official philosophies, justifications of particular revolutions and of the regimes that ensued from them. Unlike Marxism, they were not coherent doctrines, and their proponents never made more than a pretense of reconciling theory and practice. New situations called out new theoretical pronouncements in diametrical opposition to earlier ones—but the earlier pronouncements were deliberately allowed to remain as part of the doctrine, with no attempt at harmonizing them with the new claims. Complicating any systematic interpretation is their deliberate irrationalism. Benito Mussolini tended to glorify action—any action; Adolf Hitler relied on his own intuition.

Of these two ideologies, fascism had the twin advantages for clarity and consistency, if not for ideological use, of being largely confined to a conception of the right arrangement of politico-economic life and of having an official formulation compiled by a philosopher, Giovanni Gentile (although Gentile's formulation was worked over by Mussolini himself). Both fascism and national socialism pretended to be nationalist and socialist. In Italy this meant the corporative state and the denial of class antagonisms. Political power was supposed to pass upward through organizations embracing all those who worked in an industry, workers and owners alike, but these organizations would naturally merge their own interests in the national interest. In practice, although not as efficiently as in Germany, this meant totalitarian political control. The fascist glorification of the leader and the attempted revival of the glories of the Roman Empire seem peripheral to fascism when compared with the role played by similar claims in national socialist doctrine.

The tenets of national socialism, unlike those of fascism, were purposely left vague and were allowed to shift as circumstances warranted. The actual doctrines could only be what Hitler said they were, yet he deliberately tolerated or encouraged conflicting outlines of national socialism by Alfred Rosenberg and others. Even statements by Hitler himself were authoritative for the doctrine only at the time they were made. What can be said is that national socialism, like anarchism, was an antipolitical doctrine, but at the same time it was paradoxically a doctrine that aimed at total control. It was antipolitical in that this control was centered outside the state even though it might work through the state. The authority of the governmental workers and even of national socialist party leaders was diffused, indistinct, and broken on the lower levels so that it could be centered at the top. Hitler's own authority was held to derive not so much from his political position as chancellor of the Reich as from his being the *Führer,* or leader, of the people. He somehow embodied, and knew nonrationally, their strivings and desires; his will was theirs.

Of the various doctrines of national socialism, the central one was undoubtedly that of the racial war between Aryans and Jews. In this war the Jews were seen as the aggressors. They were guilty of constant and unceasing conspiratorial attacks on the superior Aryan race, which in self-defense was forced to undertake their extermination. All other violence instigated by Hitler, both against other nations and against the Germans themselves, was an incidental means to the strengthening of the Aryan race in its main battle. Nevertheless, even the race doctrine could have been dropped unceremoniously, or aimed at some other target, if circumstances had seemed to warrant, just as, for expediency, Hitler dropped first the anticapitalist claims of national socialism and then its anti-Bolshevist ones.

TWENTIETH-CENTURY POLITICAL THOUGHT

With the growing professionalization of political thought into political science and its various branches, and the development of related sociological disciplines, there has been a decline in the Anglo-Saxon countries of political philosophy in the tradition with which Hobbes, Locke, Burke, Mill, and Green are identified. Books of traditional political philosophy have continued to be written, but not generally by those who are writing the most vital works in the more central areas of philosophy, and the new works have not generally been regarded as major contributions to philosophy by those working in the newer analytic modes of philosophy. Perhaps only the subtle and persuasive Burkean traditionalism of Michael Oakeshott has attracted the continuing interest, if not the agreement, of contemporary analytic philosophers.

The dearth of major systematic treatises of the nineteenth-century kind written by contemporary philosophers does not mean that they have completely neglected political philosophy. Despite the recent claim that political philosophy is dead, contemporary philosophers have applied new techniques developed in other fields to the study of the political realm. The apparent death of one tradition of political philosophizing has per-

haps been confused with the death of political philosophy. Two main contemporary trends, which overlap to some extent, can be distinguished.

METHODOLOGY. The first trend consists in the application of the insights gained by the logical positivists and other philosophers of science into the logical status of laws, theories, and concepts in the physical sciences to the problems of political philosophy and to the methodology of political science. The most eminent representative of this trend was Karl Popper. Popper's conception of politics depended on his conception of scientific research, and its exposition is closely intertwined with his critique of earlier political philosophies. It is thus difficult to do justice to his view on how politics should be practiced without explaining his scientific methodology and his reasons for holding that the notions of historical development held by Hegel, Marx, Comte, and Mill are mistaken, and that therefore their notions of what the aims and methodology of the social and political sciences should be are fallacious. But in general he took a cautious attitude toward social change. He used the analogy of scientific investigations to advocate what he terms "piecemeal engineering"; small-scale social changes are to be preferred, because our predictions are always fallible, and mistakes on a small scale are more easily rectifiable than large-scale ones. A total change of society, or the prophecy of the results of a total change, is logically impossible; but the broader the change, the more factors which we must predict and which may go wrong or be overlooked. Connected with this viewpoint is his limited utilitarianism: It is better to attempt to alleviate pain by rectifying an existing evil than to try to increase pleasure by initiating some apparently beneficial change.

The writings of Popper and others on the logic and methodology of the social and political sciences has pioneered in a field that was little more than discovered in the nineteenth century by Mill, Comte, and Spencer—a field in which there is much important work to be done. For example, philosophers have begun to study the logic of political decision making, a subject that has heretofore been left largely to the political scientists themselves.

ANALYTIC POLITICAL PHILOSOPHY. The other main trend in contemporary political philosophy consists in the manipulation of the methods of philosophical analysis developed in the English-speaking countries in the middle decades of the twentieth century. However, neither the variety of philosophical tasks undertaken nor the results achieved present a unified picture, since the approach analytic philosophers take to political philosophy is no more unified than their approach to other groups of philosophical problems.

The first full-scale analytic treatment of the problems of political philosophy, T. D. Weldon's *The Vocabulary of Politics* (Harmondsworth, U.K., 1953), is popularly supposed to have proclaimed the death knell of political philosophizing. Weldon claimed that the various philosophical theories put forth as foundations for liberal democracy, communism, and authoritarianism cannot do what they are held to do. Either they are logically empty and thus have no consequences, or they are mistaken and harmful empirical generalizations open to refutation. Thus Weldon made short work of the social contract theory. Assume, he said, that the Mayflower Compact was shown to be a forgery and that the laws of Massachusetts are held to be based on it. If the citizens of Massachusetts then claimed that because the compact was a forgery, they had lost faith in their democratic institutions, we would feel that this reason was a cover for some other reason.

But despite his denial of the usefulness or the possibility of providing foundations for a political viewpoint, Weldon's alternative description of the political process is a good example of philosophizing about politics, and he himself claimed that "a great deal needs to be done about the language in which discussions of political institutions are conducted" (p. 172).

Other contemporary philosophers have not taken as negative an attitude toward traditional philosophizing about politics as Weldon's. Rather than rejecting out of hand notions like the social contract or general will, they have sought to give new interpretations of such notions, regarding them, for example, as models of the political process. When so interpreted, new sorts of questions arise, questions appropriate to the relation between a model and reality rather than to the analysis of an empirical description. Many other new analyses of traditional political problems and of earlier answers to them are being given, particularly of such problems as sovereignty and natural law, on the borderline between philosophy of law and political philosophy. But the variety of work being done precludes any overall description.

See also General Will; Natural Law; Social Contract.

Bibliography

GENERAL HISTORIES

The standard work is still G. H. Sabine, *A History of Political Theory* (New York, 1938, and revisions). More recent inclusive works, such as John Plamenatz, *Man and Society,* 2

vols. (London: Longman, 1963), tend to be much more restricted in range.

CRITIQUE OF THE SUBJECT

No reasoned survey has yet appeared, but in such works as Karl Popper, *The Open Society and Its Enemies* (Vol. I, *The Spell of Plato,* Vol. II, *The High Tide of Prophecy, Hegel, Marx and the Aftermath*; London: Routledge, 1945; Princeton, NJ: Princeton University Press, 1950), and C. B. Macpherson, *The Political Theory of Possessive Individualism* (Oxford: Clarendon Press, 1962), examples of contemporary critical attitudes will be found. They are themselves instances of an approach criticized in the first part of the article; for works sharing the view taken there, see the continuing collections titled *Philosophy, Politics and Society,* edited by Peter Laslett, W. G. Runciman, et al. (Oxford: Blackwell, 1957–).

ANCIENT POLITICAL PHILOSOPHY

Ernest Barker's books are the most useful for the ancient period: *Greek Political Theory, Plato and His Predecessors,* 3rd ed. (London: Metheun, 1947), *The Politics of Aristotle* (Oxford: Clarendon Press, 1946), and *From Alexander to Constantine* (Oxford: Clarendon Press, 1956). The editions of the ancient classics are innumerable, but the student is recommended to use the Loeb editions if he possibly can, with the original and its English translation on facing pages. All the relevant works (Plato, Aristotle, Cicero, Augustine, etc.) are now in print, although the edition of *The City of God* has yet to be completed.

MEDIEVAL POLITICAL PHILOSOPHY

The great works on medieval political philosophy are those of R. W. Carlyle and A. J. Carlyle (*A History of Mediaeval Political Theory in the West,* 6 vols., London: Blackwood, 1903–1936) and of Ernst Troeltsch (*The Social Teaching of the Christian Churches,* translated by Olive Wyon, 2 vols., New York: Macmillan, 1931). The books of Walter Ullmann, beginning with *The Medieval Idea of Law* (London: Methuen, 1946), contain a stimulating if highly individual interpretation. The important texts are available in Thomas Aquinas, *Selected Political Writings,* edited by A. P. d'Entrèves (Oxford: Blackwell, 1948); Marsilius of Padua, *Defensor Pacis,* translated with an introduction by Alan Gewirth as Vol. II of his *Marsilius of Padua, Defender of Peace* (New York: Columbia University Press, 1956); and Dante, *De Monarchia,* translated and annotated by P. H. Wicksteed (1896).

MACHIAVELLI AND THE SIXTEENTH CENTURY

The Prince was edited in Italian by L. H. Burd (Oxford: Clarendon Press, 1891), but there is a more recent critical edition in English of the *Discourses* (by L. J. Walker, London: Routledge and K. Paul, 1950) which is valuable for Machiavelli generally. A useful if uninspired book is J. W. Allen, *A History of Political Thought in the 16th Century* (reprinted, London, 1957). Jean Bodin's *Republic* has been edited in English by K. D. Macrae (Cambridge, MA: Harvard University Press, 1962). Hooker is still best read in John Keble's Victorian edition of his *Works,* 3 vols. (London, 1836).

THE SEVENTEENTH AND EIGHTEENTH CENTURIES

The general authority on the seventeenth and eighteenth centuries is Otto von Gierke, *Natural Law and the Theory of Society,* translated by Ernest Barker (Cambridge, U.K.: Cambridge University Press, 1934). Hobbes's *Leviathan* has been edited by Michael Oakeshott (Oxford, 1947); Spinoza's *Political Works* by A. G. Wernham (Oxford: Blackwell, 1958); Locke's *Two Treatises* by Peter Laslett (Cambridge, U.K.: Cambridge University Press, 1960); and Robert Shackleton has written a standard work on Montesquieu: *Montesquieu, a Critical Biography* (Oxford: Oxford University Press, 1961); Montesquieu's *De l'esprit des lois* is available in English, edited by F. Neumann (New York, 1949). Rousseau studies are still dominated by C. E. Vaughan, *The Political Writings* (1915; reprinted, Oxford: Blackwell, 1962); there is also a translation of the *Social Contract* by F. M. Watkins (London, 1953). There are many reprints, but so far no critical editions, of the books of Burke, Paine, Bentham, Mill, and Green quoted in the text.

THE ENLIGHTENMENT AND UTILITARIANISM

Ernst Cassirer wrote a definitive work, *The Philosophy of the Enlightenment,* translated by F. C. A. Koelln and J. P. Pettegrove (1932, English ed., Princeton, NJ: Princeton University Press, 1951), and J. L. Talmon one with a more tendentious if stimulating thesis, *The Origins of Totalitarian Democracy* (London: Secker and Warburg, 1952). John Plamenatz prefixed a brilliant essay, "The English Utilitarians," to his reprint of Mill's *Utilitarianism* (Oxford: Blackwell, 1949). Élie Halévy, *The Growth of Philosophic Radicalism,* translated by Mary Morris (London: Faber and Gwyer, 1928), is still important.

HEGEL AND GERMAN IDEALISM

Hegel's main work on political philosophy is *Naturrecht und Staatswissenschaft im Grundrisse* (Berlin, 1821), 2nd ed. edited by E. Gans as *Grundlinien der Philosophie des Rechts* (Berlin, 1833), translated by T. M. Knox as *The Philosophy of Right* (Oxford: Clarendon Press, 1942). His *Phänomenologie des Geistes* (Würzburg and Bamberg, 1807), translated by J. B. Baillie as *Phenomenology of Mind* (London: S. Sonnenschein, 1910), and *Vorlesungen über die Philosophie der Geschichte,* edited by E. Gans (Berlin, 1837) and translated by J. Sibree as *Lectures on the Philosophy of History* (London: Bohn, 1857), should also be consulted. See also *Hegel's Political Writings,* translated by T. M. Knox with an introductory essay by Z. A. Pelczynski (Oxford: Clarendon Press, 1964). On Hegel's political thought, see M. B. Foster, *The Political Philosophies of Plato and Hegel* (Oxford: Clarendon Press, 1935); Franz Rosenzweig, *Hegel und die Staat,* 2 vols. (Oldenburg, 1920); Eric Weil, *Hegel et l'état* (Paris: J. Vrin, 1950), and the works by Popper and Plamenatz cited above.

Hermann Lübbe, ed., *Die Hegelsche Rechte* (Stuttgart and Bad Canstatt, 1962), and Karl Löwith, ed., *Die Hegelsche Linke* (Stuttgart and Bad Canstatt, 1962), contain selections from right-wing and left-wing German successors of Hegel, respectively. The second is more directly relevant to political philosophy. From the voluminous writing on this period, see Sidney Hook, *From Hegel to Marx* (New York: Humanities Press, 1950); Georg Lukács, *Die Zerstörung der Vernunft* (Berlin: Aufbau, 1954); and Herbert Marcuse, *Reason and Revolution: Hegel and the Rise of Social Theory,* 2nd ed. (New York: Humanities Press, 1954). Johann Gottlieb Fichte was an idealist contemporary of Hegel whose writings are of considerable political interest. See

especially his *Der geschlossene Handelsstaat* (The closed commercial state; Tübingen, 1800).

BRITISH IDEALISM

Green's *Lectures on the Principles of Political Obligation* were first published in *The Works of Thomas Hill Green,* edited by R. L. Nettleship, 3 vols. (London, 1885–1888). See Melvin Richter, *The Politics of Conscience: T. H. Green and His Times* (London: Weidenfeld and Nicolson, 1964). On Bosanquet, besides Hobhouse, see F. Houang, *Le néo-Hegelianisme en Angleterre: La philosophie de Bernard Bosanquet* (Paris, 1954). For a general account of British Neo-Hegelian political and social thought, see A. J. M. Milne, *The Social Philosophy of English Idealism* (London: Allen and Unwin, 1962). *The Philosophy of Loyalty* (New York: Macmillan, 1908) by the American idealist Josiah Royce shows a related development. On Royce, see J. E. Smith, *Royce's Social Infinite* (New York: Liberal Arts Press, 1950). Other works by Hobhouse are *Liberalism* (London: Williams and Nirgate, 1911) and *The Elements of Social Justice* (London: Allen and Unwin, 1922). See J. A. Hobson and Morris Ginsberg, *L. T. Hobhouse, His Life and Work* (London: Allen and Unwin, 1931). Of the many British political writings broadly following in the tradition of Green, the following may be mentioned: Ernest Barker, *Reflections on Government* (London: Oxford University Press, 1942); A. D. Lindsay, *The Modern Democratic State* (London: Oxford University Press, 1943); and J. D. Mabbott, *The State and The Citizen* (London: Hutchinson's University Library, 1948). A curious wartime idealist work with an intent similar to that of Hobhouse's *Metaphysical Theory of the State* is R. G. Collingwood's *The New Leviathan* (Oxford: Clarendon Press, 1942).

MARXISM

Almost any writing of Marx or Engels is relevant to their political philosophy. See especially Marx's *Das Kapital,* 3 vols. (Hamburg, 1867–1894), translated by Samuel Moore, Edward Aveling, and Ernest Untermann as *Capital,* 3 vols. (Chicago, 1915); Marx and Engels's *Die deutsche Ideologie,* edited by V. Adoratsky (Vienna, 1932), translated as *The German Ideology,* edited by S. Ryazanskaya (Moscow: Progress, 1964); Marx and Engels's *Manifest der kommunistischen Partei* (London, 1848), translated as *The Communist Manifesto,* edited with an introduction by Harold Laski (London: Allen and Unwin, 1948); and Engels's *Der Ursprung der Familie, des Privateigentums und des Staat* (Zürich, 1884), translated by Ernest Untermann as *The Origin of the Family, Private Property, and the State* (Chicago: Kerr, 1902). Two convenient anthologies are Lewis S. Feuer, ed., *Basic Writings on Politics and Philosophy* (Garden City, NY: Doubleday, 1949), and T. B. Bottomore and Maximilien Rubel, eds., *Selected Writings in Sociology and Social Philosophy* (London: Watts, 1956). Of the writings of Lenin, see especially *Chto Delat?* (*What Is to Be Done?*; Stuttgart: Dietz, 1902), *Shag Vperyod, Dva Shaga Nazad* (*One Step Forward, Two Steps Back*; Geneva: Partii, 1904), *Imperializm, kak Vysshara Stadiya Kapitalizma* (*Imperialism, the Highest Stage of Capitalism*; Petrograd: Zhizn' i znznie, 1917), and *Gosudarstvo i Revolutsiya* (*State and Revolution*; Petrograd: Zhizn' i znznie, 1918). There are various English editions of all of these. Of the many other Marxist writers on political philosophy, one of the most interesting is Antonio Gramsci. See his *Opere,* 6 vols. (Turin, 1947–1954), and *The Modern Prince and Other Writings,* translated by Louis Marks (London: Lawrence and Wishart, 1957). For other Marxist writings and for writings on Marxism, consult the bibliographies to the entries Dialectical Materialism, Historical Materialism, and Marxist Philosophy.

ANARCHISM

Among the chief anarchist works are William Godwin, *An Enquiry concerning Political Justice and Its Influence on General Virtue and Happiness* (London: GGJ and J. Robinson, 1793); the writings of Michael Bakunin, translations of which appear in *The Political Philosophy of Bakunin: Scientific Anarchism* (Glencoe, IL: Free Press, 1953); Prince Peter Kropotkin's *The State, Its Part in History* (London: Freedom office, 1898), *Mutual Aid, a Factor of Evolution* (London, 1902), and *Modern Science and Anarchism* (Philadelphia, 1903); Pierre-Joseph Proudhon's *Qu'est-ce que la Propriété?* (Paris, 1840), translated by Benjamin R. Tucker as *What Is Property; An Inquiry into the Principle of Right and of Government* (New York: Humboldt, 1890); Henry David Thoreau, "Civil Disobedience," in *The Writings of Henry David Thoreau,* Vol. X (Boston and New York, 1863); Benjamin R. Tucker, *Instead of a Book: A Fragmentary Exposition of Philosophical Anarchism* (New York, 1897); and Georges Sorel, *Réflexions sur la violence* (Paris: Librairie de "Pages Libres," 1908), translated by T. E. Hulme and J. Roth as *Reflections on Violence* (New York: Huebsch, 1914). On anarchism, see George Woodcock, *Anarchism: A History of Libertarian Ideas and Movements* (Cleveland, OH: World Publishing, 1962); James Joll, *The Anarchists* (London: Eyre and Spottiswoode, 1964); and Alexander Gray, *The Socialist Tradition* (London: Longmans, 1946).

NATIONAL SOCIALISM AND FASCISM

For further pronouncements by national socialists, see Josef Goebbels, *Goebbels Tagebücher,* edited by Louis Lochner (Zürich, 1948); Adolf Hitler, *Mein Kampf,* 2 vols. (Munich, 1925–1927), and *Hitler's Secret Conversations 1941–1944* (New York: Farrar, Straus and Young, 1953). German Philosophy and National Socialism contains an extensive bibliography of relevant works, which may be supplemented by bibliographies in many of the works cited there. On fascism, consult Benito Mussolini, *Scritti i discorsi,* 12 vols. (Milan, 1934–1939) and *The Doctrine of Fascism,* translated in *Social and Economic Doctrines of Contemporary Europe,* edited by Michael Oakeshott, 2nd ed. (New York, 1942); and Giovanni Gentile, *Che cosa è il fascismo* (Florence: Vallecchi, 1925) and *Origini e dottrine del fascismo* (Rome: Libreria del Littorio, 1929).

CONTEMPORARY POLITICAL THOUGHT

The best picture of contemporary analytic political philosophy can be gathered from the series of collections titled *Philosophy, Politics and Society,* edited by Peter Laslett, W. G. Runciman et al. (Oxford, 1957–). Popper's main works on political philosophy are *The Open Society and Its Enemies* and *The Poverty of Historicism* (London: Routledge, 1957). Weldon also published *States and Morals* (London: J. Murray, 1946). On Oakeshott, consult his inaugural address in the first volume of *Philosophy, Politics and Society* and his

Rationalism in Politics (New York: Basic, 1962). Other examples are H. L. A. Hart, "The Ascription of Responsibility and Rights," *PAS* 49 (1948–1949): 179–194, reprinted in *Essays on Logic and Language,* edited by A. G. N. Flew (Oxford: Blackwell, 1951); Margaret Macdonald, "The Language of Political Theory," in *PAS* 41 (1940–1941), reprinted in Flew, op. cit.; J. W. N. Watkins, "Epistemology and Politics," in *PAS* 58 (1957–1958): 79–102; and S. I. Benn and R. S. Peters, *Social Principles and the Democratic State* (London: Allen and Unwin, 1959).

Peter Laslett (1967)
(Introduction through Kant)
Philip W. Cummings (1967)
(Hegel through recent political thought)

POLITICAL PHILOSOPHY, HISTORY OF [ADDENDUM]

Political philosophy, theory, and thought all focus on the arguments that have been advanced—by prominent thinkers from around the world and throughout human history—for various conceptions of a just human community. Different schools of such political thinking have developed over time and here some of these schools will be sketched and the major contributors will be listed. The bibliography lists sources for further study of these ideas.

LIBERTARIANISM AND CAPITALISM

Libertarianism is the political system wherein the highest political good is the protection of the individual citizen's right to life, liberty, and property. Capitalism is the economic system of libertarianism because in libertarian societies the institution of the right to private property, that is, to own anything of value (not, of course, other human beings, who are themselves owners), is fully respected and protected.

Libertarian law rests on the idea that the individual is the most important member of society, with all groups to be formed by the consent of individual members, including the military, corporations, universities, clubs, and the government itself. What is primarily prohibited in a libertarian society is involuntary servitude. What is primarily promoted via the political administration is the liberty of all persons to advance their own objectives provided they do not in this process violate anyone's equal rights. The major contributors to libertarian political thought have been Murray N. Rothbard, Ayn Rand (although she eschewed that term, preferring *radical capitalism* instead), Robert Nozick, Loren Lomasky, Jan Narveson, Douglas B. Rasmussen, Douglas J. Den Uyl, and Tibor R. Machan.

There is dispute about the label *capitalism* as the proper way to call the economic order under libertarianism, mostly because its definition is often a precondition of having either a favorable or unfavorable view of the system. Some have insisted on the use of laissez-faire, in memory of the French entrepreneurs who responded to the king's question as to what the government can do to help the economy by exclaiming: "Laissez-faire, lassize passe," or "Leave us to do, leave us to act." Some use F.A. Hayek's term *the spontaneous order* to stress such a system's support of uncoerced behavior. There is also the more popular term *free enterprise.*

Yet capitalism is most widely used, by both critics and supporters of an economic order in which individuals have the right to own property and to use of it on their own terms. By itself capitalism is an economic arrangement of an organized human community or polity. Often, however, entire societies are called capitalist, mainly to stress their thriving commerce and industry. More rigorously understood, however, capitalism presupposes a libertarian legal order governed by the rule of law in which the principle of private property rights plays a central role. Such a system of laws was historically grounded on various classical liberal ideals in political thinking. These ideals can be defended by means of positivism, utilitarianism, natural rights theory and/or individualism, as well as notions about the merits of laissez-faire (no government interference in commerce), the invisible hand (as a principle of spontaneous social organization), prudence and industriousness (as significant virtues), the price system as distinct from central planning (for registering supply and demand), and so on.

Put a bit differently, *capitalism* or *economic libertarianism* are the terms used to describe that feature of a human community whereby citizens are understood to have the basic right to make their own (more or less wise or prudent) decisions concerning what they will do with their labor and property, or whether they will engage in trade with one another involving nearly anything they may value. Thus capitalism includes freedom of trade and contract, the free movement of labor, and the protection of property rights against both criminal and official intrusiveness.

The concept of freedom plays a central role in the understanding of both libertarianism and capitalism. There are two prominent ways of understanding the nature of freedom as it pertains to human relationships. The one that fits with capitalism is negative freedom: the

condition of everyone in society not being ruled by others with respect to the use and disposal of themselves and what belongs to them. Citizens are free, in this sense, when no other adult person has authority over them that they have not granted of their own volition. In short, in capitalism one enjoys negative freedom, which amounts to be free from others' intrusiveness. The other meaning of freedom is that citizens have their goals and purposes supported by others or the government so as to prosper. Under this conception of freedom one is free to progress, advance, develop, or flourish only when one is enabled to do so by the efforts of capable others.

In international political discussions the concept of capitalism is used very loosely, so that such very diverse types of societies as Italy, New Zealand, the United States of America, Sweden, and France are all considered capitalist. Clearly, no country today is completely capitalist. None enjoys a condition of economic laissez-faire in which governments stay out of one's commercial transactions except when conflicting claims over various valued items are advanced and the dispute needs to be resolved in line with due process of law. But many Western type societies protect a good deal of free trade, even if they also regulate most of it as well. Still, just as those countries are called democratic if there is substantial suffrage—even though many citizens may be prevented from voting—or if there exists substantial free trade and private ownership of the major means of production (labor, capital, intellectual creations, and so on), the country is usually designated as capitalist.

The most common reason among political economists for supporting capitalism is this system's support of wealth creation. This is not to say that such theorists do not also credit capitalism with other worthwhile traits, such as encouragement of progress, political liberty, innovation, and so on. Those who defend the system for its utilitarian virtues—its propensity to encourage the production of wealth—are distinct from others who champion the system—or the broader framework within which it exists—because they consider it morally just.

The first group of supporters argue that a free market or capitalist economic system is of great public benefit, even though this depends on private or even social vice, such as greed, ambition, and exploitation. As Bernard Mandeville, the author of *The Fable of the Bees*, put it, this system produces "private vice, public benefit." Many moral theorists see nothing virtuous in efforts to improve one's own life. They believe, however, that enhancing the overall wealth of a human community is a worthwhile goal. Those who follow along lines of Mandeville in the twentieth century, including Ludwig von Mises, Milton Friedman, F. A. Hayek, Gary Becker, and James Buchanan, stress the practical merits of this economic system rather than its moral justification.

Those who stress the moral or normative merits of capitalism, mostly libertarians, say the system rewards prudence, hard work, ingenuity, industry, entrepreneurship, and personal or individual responsibility in all spheres of human life, and this is all to the good. This alone makes the system morally preferable to alternatives. Yet, another reason given why libertarianism or capitalism is not only useful but morally preferable is that it makes possible the exercise of genuine moral choice and agency, something that would be obliterated in noncapitalist, collectivist systems or economic organization. Most of the libertarians (see previous paragraph) advance this type of normative argument for capitalism.

Capitalist theorists note that most critics of capitalism demean wealth. Indeed, they virtually attack the pursuit of human individual well-being itself and, especially, luxury, anytime there are needy people left anywhere on earth, as well as, more recently, if any portion of nature is overrun by human beings (as if they were not natural creatures). But, the champions of capitalism argue, this stems from utopian thinking and has the consequence of begrudging anyone a measure of welfare because some people will always be poor some of the time and nature will continue to be transformed by people.

Yet the capitalist advocate need not be seen as reckless toward the environment. Indeed, arguably the strict and consistent institution of the principle of private property rights—through, for example, privatization and prohibition of dumping waste into other private as well as public realms—may solve the environmental problems we face better than any central planning champions of the environment tend to propose. Libertarians and capitalists think that the environment suffers worst when the "tragedy of the commons" is permitted, whereby commonly owned values are overused because everyone is deemed to have a right to such use, while no one in particular is left with the responsibility to care for it.

Capitalism rests in large part on the belief that human beings are essentially individuals and a society's laws must value individuals above all else. Most historians of ideas admit that whether the importance of human individuality should have been recognized in earlier times, it certainly was not much heeded until the modern age. Even in our time it is more often that groups—ethnic, religious, racial, sexual, national, and cultural—are taken to have greater significance than individuals. The

latter are constantly asked to make sacrifices for the former. In capitalism, however, the individual (e.g., as the sovereign citizen or the consumer) is king. Undoubtedly a capitalist system does not give prime place to economic equality among people, something that group thinking seems to favor because, in groups, all are deemed to be entitled to a fair share.

WELFARE STATISM

The welfare state or, from the economic viewpoint, the mixed economy, may be understood as a combination of the principles of capitalism and socialism. Sometimes the emphasis in this system is placed not so much on economic dilemmas as on certain moral considerations. Basically the welfare state consists of a legal system that aims at securing for everyone the negative right to liberty and the positive right to well-being. The main defenders of this system in the later twentieth century are John Rawls, Amartya Sen, Martha Nussbaum, and Jurgen Habermas.

The welfare state, which is to say most Western countries, balances the two values that together seem to its advocates to be the bedrock of a civilized society. No one ought to have his or her sovereignty seriously compromised, nor should anyone be permitted to fall below a certain standard of living. This is difficult to maintain because at different times one or another of these objectives will probably take priority and, in mostly democratic systems, political leaders will vacillate between giving more support to one or the other. The right to strike, for example, which is the negative liberty to quit one's job in an effort to gain respect for one's terms of employment, may conflict with the positive right to be provided with various services (e.g., health care, mail delivery, or education).

It is indeed a prominent feature of the welfare state that both negative and positive rights receive their legal protection. Negative rights involve respect for a person's life, liberty, and property—that is, everyone is by law supposed to abstain from interfering with these. Positive rights, in turn, involve respect for a person's basic needs—that is, everyone who is unable to secure the requirements of survival, and even flourishing, is supposed to have those provided by way of the appropriate public policy (e.g., taxation, mandated services, public education, and national health care).

The moral underpinnings of the welfare state can be utilitarianism, altruism, or certain intuitively held moral precepts. Utilitarianism requires that the general welfare be pursued by all and whatever public policies to facilitate this were needed would be justified. Although many utilitarians believe that the general welfare is best achieved when government operates in a largely laissez-faire fashion, there is no objection to government intervention in social affairs if without those many in the society may fail to achieve a decent and prosperous form of life. Altruists, in turn, often hold that to make certain that people fulfill their primary obligation to help others, it is necessary to introduce public measures that will secure such help, given that many might wish to breach their duty to do the right thing. Finally, there is the claim that, by our common intuitions, it is evident that both a measure of personal liberty and social welfare must be guaranteed to all, lest the quality of life in society fall below what it should be.

Whereas the welfare state is objected to by people from several other perspectives, it is thought by its supporters to be the most stable modern political order. Although it is characterized by much dispute and controversy, in the long run, its supporters maintain, the system seems to be overall satisfactory and just.

COMMUNITARIANISM

Communitarianism could be viewed as a sort of halfway house between the collectivist system of socialism and the individualist one of capitalism. The idea is less capable of being sharply defined than these others. Roughly it comes to the view that human beings are necessarily or essentially parts of distinct human groups, communities, with their diverse values, histories, priorities, practices, laws, and cultures. The organizing principles of these different groups will themselves vary. There is no overriding true social and political order, not even any universal ethics. Rather it is the particular character of the communities that establish for its parts or members what is the proper way to live, what laws should be enacted, and what aesthetic and religious values need to be embraced.

Some communities can be Spartan, others Stoic, yet others bohemian and so forth. Each can have its peculiar way of life without implying any objective condemnation of some alternative form. Yet participation in the community's form of life is not a matter of individual consent. Such an idea derives from a mistake: There is a transcendent or general human nature that requires every community to adhere to certain minimal standards of justice. No such transcendent human nature exists, as far as many communitarians see things, so those that, say, grant individuals certain rights are not superior to those that do not—they are simply different. Among those who are prominent communitarians, Charles Taylor, Amitai

Etzioni, Thomas Spragens, Michael Sandel, and Richard Rorty stand out.

Actually there is not much more that can be said about communitarianism because there are simply too many types of community, each with its own framework and priorities. The main point is that the rules, laws, ideals, and so forth are all the result of the often slowly evolving consensus or collective practices of the community's membership. Just as socialism sees humanity as the whole to which individuals belong, communitarianism sees different ethnic, national, racial, gender, cultural, professional, or similar distinguishable groups as the whole to which the individual member belongs. One may imagine, for example, that languages have developed, in part, to meet the requirements, imagination, and circumstances of different linguistic communities, with no language superior or completely translatable to any other.

Communitarians often unite in their criticism of bourgeois society or liberal capitalism because of their emphasis on individuality, privacy, personal freedom, consent, and competition. Communitarians believe that the view of human nature underlying such liberal capitalist views is seriously flawed. They are convinced, also, that the central idea of liberal capitalism is what has come to be known as *homo economicus* or *economic man*. That idea figures heavily in economic analysis and views individuals as autonomous entities who enter the world fully formed, ready to make choices in the market, and self-sufficient. While there are other conceptions of the human individual that might support liberal capitalism, it is this that has occupied the attention of communitarians and it is in contrast to this view that they have advanced their position.

ISLAMIC POLITICAL THEORY

Muslims are divided into two communities, the Sunni majority and Shii minority, and they adhere to different ideas as to political rule. They are known as the Sunni caliphate and the Shii Imamate.

When Muhammad died, most Muslims, since they thought that Muhammad did not name a successor, relied upon the decision of a group of his cohorts. The caliphate, chosen by way of consultation (called *shura*) and agreement (called *ijma*), an oath of loyalty (called *baya*) that is sworn by those who elect him, and the compact (called *ahd*) with the people to govern by Islamic law (Sharia) developed into what is widely regarded as legitimate government for Sunni Islam.

But the Shii rejected the Sunni caliphs and regarded them as subverting Islamic law. They adhered to the idea that Muhammad had selected Ali, who was reported to be his cousin and son-in-law, to be the ruler (*Imam*) of Muslims. They held that the oldest (male) descendant (*Ahl al-Bait*) must be the divinely anointed, religious, and political chief. Abbasid rule (750–1250) formed Islamic political theory as theocratic, with theologians as the legal authorities who had royal privilege and professed to uphold the divine goal for the Muslim community under Abbasid edicts. In the last analysis, as matters now stand, there is no unified Muslim political theory that enjoys widespread acceptance.

In geopolitical affairs a very influential version of Muslim politics comes from the clerics and adherents of the Wahhabi branch of radical Islam, based mainly in Saudi Arabia and considered to be the most virulently anti-Western in light of the belief that any accommodation of Western values is an intolerable compromise with the words of the Prophet. The main point of contention is that the West legally tolerates freedom of religion and even nonbelief, which undermines the virtuous life demanded of the faithful, leading to their corruption.

JEWISH POLITICAL THEORY

Jews, as such, do not adhere to a firm political creed, unlike many Muslims, but tend to embrace varieties of democratic, even liberal, institutions, while also encouraging some socialist economic practices and certain mild forms of theocracies, depending on the version of Judaism they embrace. Jewish political ideas derive mainly from the belief that Jews are a separate, unique—chosen—people, not merely adherents to a different religion or a system of moral principles that emerge from such a religion (of course this idea is shared by nearly all traditional and organized religious groups). Jewish political ideas pertain to how the Jews as a unified people have held on to a political community throughout the centuries, without becoming amalgamated into communities wherein they lived as exiles and how they shaped these by giving clear expressions of their own culture and forms of political conduct.

Jews often choose to demonstrate a Jewishness via political means and this for many of them consists of loyalty to modern Israel as well as various Jewish missions, including various communal groups (for example, the kibbutz) constituted almost exclusively by Jews. As is common in politics everywhere, Jews will often stress the need for power as they advance the causes of their various

groupings, although this also includes extensive education and proselytizing.

CONCLUSION

None of the systems we have sketched here are fully exemplified anywhere, although some—for example, Islamic theocracy—are approximated in some parts of the world (e.g., Iran). There are, however, no purely capitalist, socialist, or communist societies and the welfare states are also quite different, with various ways of balancing the values of personal autonomy and social security. Instead, most societies—countries—exhibit mixed systems and often where democratic decision-making takes place, the main topic of debate is which of these values should be stressed more, as well as how much state support should be given to various special interests.

See also Civil Disobedience; Cosmopolitanism; Postcolonialism; Republicanism

Bibliography

Mitchell, Joshua. *Not by Reason Alone: Religion, History, and Identity in Early Modern Political Thought*. Chicago: University of Chicago Press, 1996.

Novak, David. *Covenantal Rights: A Study in Jewish Political Theory*. Princeton, NJ: Princeton University Press, 2000.

Skoble, Aeon J., and Tibor R. Machan, eds. *Political Philosophy: Essential Selections*. Upper Saddle River, NJ: Prentice Hall, 1999.

Strauss, Leo, and Joseph Cropsey, eds. *History of Political Philosophy*. 3rd ed. Chicago: University of Chicago Press, 1987.

Watt, William. *Islamic Political Thought*. Edinburgh: Edinburgh University Press, 1998.

Tibor Michan (2005)

POLLA

See *Hen/Polla*

POMPONAZZI, PIETRO

(1462–1525)

Pietro Pomponazzi, the Italian Renaissance Aristotelian, was born in Mantua. He studied philosophy at the University of Padua, where, after obtaining his degree, he became extraordinary professor of philosophy in 1488 and ordinary professor in 1495. When war caused the university to close in 1509, he left Padua. After a short period at Ferrara he became a professor of philosophy at the University of Bologna, where he taught from 1512 until his death. He married three times and had two children.

Of Pomponazzi's writings only a few were published during his lifetime. Best known is the treatise *De Immortalitate Animae* (On the immortality of the soul, 1516), which immediately provoked a large controversy. It was publicly attacked by several philosophers and theologians and was followed by the author's two treatises in defense—the *Apologia* (1518) and the *Defensorium* (1519)—which were longer than the original work. Probably as a result of this experience Pomponazzi did not publish anything else except for a few short philosophical questions that he added to the 1525 reprint (*Tractatus Acutissimi*) of his three writings on immortality. Equally important are his treatises *De Incantationibus* (On incantations) and *De Fato* (On fate), both written about 1520, which were published posthumously in Basel by a Protestant exile in 1556 and 1567, respectively. A sizable body of other writings has been preserved in manuscript, and the study and publication of this material have barely begun. The most important among these unpublished writings are questions on Aristotelian and other problems, which Pomponazzi probably worded himself and that therefore directly reflect his thought. A much larger group consists of his class lectures on various works of Aristotle. Since they were taken down by students and show a certain amount of oscillation from year to year and from copy to copy, they must be used with caution in any attempt to reconstruct Pomponazzi's thought and philosophical development.

Pomponazzi was a product and in many ways a typical representative of the tradition of scholastic Aristotelianism that flourished at Bologna, Padua, and other Italian universities from the thirteenth to the seventeenth century. This school, often referred to as Paduan Averroism, had no institutional or doctrinal connections with theology, as did its northern counterparts, but rather with medicine, and this accounts for its secular orientation. In the study of Aristotle, whose writings served as the prescribed texts for the teaching of the philosophical disciplines, the emphasis was, as in Paris and elsewhere, on logic and natural philosophy rather than on ethics and metaphysics.

Pomponazzi's main sources were the writings of Aristotle and of his commentators, and his style, far removed from classical or humanistic elegance, is a rather harsh example of scholastic terminology and argument, although he was at times capable of concise formulation and caustic wit. His reasoning shows great subtlety and

acumen, but he is repetitious and sometimes inconsistent. He obviously enjoyed spinning out an argument and following reason wherever it led, and out of intellectual honesty he was prepared to admit his puzzlement before certain dilemmas and to modify his views whenever he felt compelled to do so by some strong argument. Thus, we may well understand the outburst in *De Fato* (III, 7) in which he compares the philosopher with Prometheus. In his efforts to understand the secrets of God the philosopher is eaten up by his continual worries and thoughts; stops eating, drinking, and sleeping; is held up to ridicule by all; is taken as a fool and a faithless person; is persecuted by the Inquisition; and is laughed at by the multitude.

In spite of his general scholastic orientation Pomponazzi was by no means unaffected by other currents. He knew and respected Plato and was clearly influenced by Marsilio Ficino (and Giovanni Pico della Mirandola) in his remarks about the place of man in the universe and perhaps in his preoccupation with the immortality of the soul. Like the humanists he cultivated the monographic treatise in addition to the question and the commentary, occasionally injected personal remarks about himself, and cited such sources as Cicero and Plutarch. His doctrine that virtue is its own reward has Stoic rather than Aristotelian antecedents, and his insistence that the end of man consists in practical virtue rather than in contemplation is at variance with Aristotle and may owe something to Cicero and to such humanists as Leonardi Bruni and Leon Alberti.

One may even link with humanism Pomponazzi's interest in Alexander of Aphrodisias. Alexander was not entirely unknown during the Middle Ages, but his writings acquired a much wider diffusion through new translations around the turn of the sixteenth century. The label of Alexandrism often attached to Pomponazzi is dubious and misleading. We know from a question composed by Pomponazzi in 1504 that his view on the problem of immortality, as adopted in his treatise of 1516, was derived from that of Alexander. We also learn that the writing of his treatise *De Fato* was occasioned by his reading a new Latin translation of Alexander's treatise on the subject (Pomponazzi knew no Greek). However, *De Fato* is actually a defense of the Stoic position against Alexander.

Pomponazzi's *De Incantationibus* is an attempt to offer natural explanations for a number of occurrences popularly ascribed to the agency of demons and spirits. The effects ascribed to the stars by the astrologers form for Pomponazzi a part of the system of natural causes. This work is the only one by Pomponazzi that was once on the Index of Prohibited Books (it no longer is) because of its implied criticism of miracles. It contains an interesting passage on prayer that shows a certain affinity to some ideas expressed in the treatise on immortality. The value of prayer, he said, consists not in the external effects it may have but in the pious attitude it produces in the person who prays.

The *De Fato*, which is divided into five books, is by far the longest of Pomponazzi's works. He discusses in great detail and with a great number of intricate arguments the problems of fate, free will, and predestination. His conclusions are by no means simple or clear-cut, but it appears from his final remarks that he regarded the Stoic doctrine of fate, on purely natural grounds, as relatively free from contradictions. Yet, because human wisdom is subject to error, Pomponazzi was willing to submit to the teaching of the church and to accept the doctrine that God's providence and predestination are compatible with man's free will. However, he was not satisfied with the way in which this compatibility is customarily explained and tried to propose an explanation that he considered more satisfactory.

De Fato has been unduly neglected by students of Pomponazzi, perhaps because of its length and difficulty. It is now available in a critical edition and may be studied within the twofold historical context in which it belongs: first, the philosophical controversy between determinism and indeterminism as it appeared in antiquity in the works of the Stoics and Alexander and again in more modern discussions and, second, the specifically theological problem of reconciling providence and predestination with free will. The second question has occupied Christian theologians of all centuries; it had been discussed before Pomponazzi by Lorenzo Valla in his treatise on free will, and it was to be debated by Martin Luther, Desiderius Erasmus, and many other theologians during and after the Reformation.

DE IMMORTALITATE ANIMAE

Pomponazzi's treatise *De Immortalitate Animae* is much better known, and it had far wider repercussions during the sixteenth century and even later. Pomponazzi explains the origin of the treatise as follows: He had stated in a class lecture that Thomas Aquinas's view on immortality, though perhaps true, did not agree with Aristotle's, and he was subsequently asked by a Dominican friar who was his student to express his own opinion on the question, staying strictly within the limits of natural reason. In complying with this request, Pomponazzi begins with the

statement that man is of a manifold and ambiguous nature and occupies an intermediary position between mortal and immortal things (Ch. 1). The question is in what sense such opposite attributes as mortal and immortal may be attributed to the human soul (Ch. 2). Pomponazzi first lists six possible answers, and after having discarded two of them because they had never been defended by anybody, he promises to discuss the remaining four (Chs. 2–3).

The first of the four answers is the view attributed to Averroes and others, according to which there is only one immortal soul common to all human beings and also an individual soul for each person, which, however, is mortal. Pomponazzi rejects this opinion at great length (Ch. 4). The Averroist position maintains that the intellect is capable of acting without a body and can therefore be considered as separable and immortal. Yet in our experience, Pomponazzi argues, the intellect has no action that is entirely independent of the body, and therefore we have no evidence that the intellect is separable. If we wish to understand the relationship of the intellect and the body, we must distinguish between being in the body as having the body for its organ or subject or substratum and depending on the body as having the body, its perceptions, and imaginations for its object. Pomponazzi insists that the intellect does not have the body as its subject as do the souls of animals and the lower faculties of the human soul. Yet the human intellect cannot know anything without the perceptions or imaginations offered to it by the body, and this fact alone proves that the intellect is not separable from the body.

Second, Pomponazzi discusses an opinion he attributes to Plato, according to which each person has two souls, one immortal and the other mortal (Ch. 5). This position is rejected on the ground that the subject of perception and that of intellectual knowledge must be the same and that it is therefore impossible to distinguish two separate natures within the human soul (Ch. 6).

Third, he examines the view, attributed to Thomas Aquinas, which holds that the human soul has but a single nature and that it is absolutely (*simpliciter*) immortal and only in some respects (*secundum quid*) mortal (Ch. 7). Elaborating on some of the arguments he had already advanced against Averroes, Pomponazzi insists that he finds no evidence to prove the absolute immortality of the soul. He has no doubt, he adds, that the doctrine of the absolute immortality of the soul is true, since it is in accordance with Scripture, but he wonders whether it is in agreement with Aristotle and whether it can be established within the limits of natural reason without recourse to the evidence of faith and revelation (Ch. 8).

Fourth, Pomponazzi discusses a position according to which the human soul, having only one nature, is absolutely mortal and only in certain respects immortal (Ch. 9). He then proceeds to defend this position, which he had identified elsewhere as that of Alexander of Aphrodisias. Insisting once more on the middle position of humankind, he argues that the human intellect, unlike that of the pure intelligences, always needs the body for its object and has no way of acting without the help of the images of sense or imagination. It must therefore be considered absolutely mortal and only relatively, or improperly speaking, immortal. However, unlike the souls of the animals, the human intellect does not have the body as its subject because it does not use a bodily organ in knowing. If it resided in an organ, the intellect could not reflect on itself or understand universals. The fact that the human intellect is capable of some knowledge of itself and of universals shows that it participates somewhat in immortality and, hence, that it is in some respect immortal. This interpretation of immortality is claimed to be more probable than the others and to be more in accordance with the teachings of Aristotle (Chs. 9–10).

Having reached this conclusion, Pomponazzi continues in good scholastic fashion to formulate several sets of objections to his view (Chs. 11 and 13) and to answer these objections in great detail (Chs. 12 and 14). In addition to repeating and elaborating some of the same arguments presented in the preceding chapters, he introduces, especially in Chapter 14, several new arguments and conclusions that are of great intrinsic interest.

Along with other objections to his view Pomponazzi cites (Ch. 13) the argument that, according to Aristotle's *Ethics*, the ultimate end of man is contemplation and that the satisfactory fulfillment of this end requires immortality. In his reply he states that man has a threefold intellect—speculative, practical, and technical. Only a few persons have a share in the speculative intellect, whereas the technical intellect is shared by some animals. We may thus conclude that the practical intellect, in which all human beings and only all human beings share, is the faculty peculiar to human beings. Every normal person can attain the practical intellect in a perfect way, and a person is called absolutely good or bad with reference to this practical intellect but merely in some respect good or bad with reference to the other two intellects. For a man is called a good man or a bad man with reference to his virtues and vices, yet a good metaphysician with reference to his speculative intellect and a good architect with ref-

erence to his technical intellect. However, a good metaphysician or a good architect is not always a good man. Hence, a man does not mind so much if he is not called a good metaphysician or a good architect, but he minds very much if he is called unjust or intemperate, for it seems to be in our power to be good or wicked, but to be a philosopher or an architect does not depend on us and is not necessary for a man. The ultimate end must thus be defined in terms of the practical intellect, and every man is called upon to be as virtuous as possible.

By contrast, it is neither necessary nor even desirable that all men should be philosophers or architects but only that some of them should be. Moreover, since the perfection of the practical intellect is accessible to almost everybody, a farmer or a craftsman, a poor man or a rich man, may be called happy and is actually called happy and is satisfied with his lot whenever he is virtuous. In other words, Pomponazzi departs in this important respect from Aristotle and identifies the end of human life with moral virtue rather than with contemplation, because this end is attainable by all human beings.

There had been another objection—that God would not be a good governor of all things unless all good deeds found their reward and all bad deeds their punishment in a future life. To this Pomponazzi replies that the essential reward of virtue is virtue itself, and the essential punishment of vice is vice itself. Hence, it makes no difference whether the external or accidental reward or punishment of an action is sometimes omitted, since its essential reward and punishment are always present. Moreover, if one man acts virtuously without the expectation of a reward and another with such an expectation, the act of the latter is not considered to be as virtuous as that of the former. Thus, he who receives no external reward is more fully rewarded in an essential way than he who receives one. In the same way the wicked person who receives no external punishment is punished more than he who does, for the punishment inherent in guilt itself is much worse than any punishment in the form of some harm inflicted upon the guilty person.

Pomponazzi further develops this idea in reply to another objection. It is true that religious teachers have supported the doctrine of immortality, but they have done so in order to induce ordinary people to lead virtuous lives. Yet persons of a higher moral disposition are attracted toward the virtues by the mere excellence of these virtues and are repelled from the vices by the mere ugliness of these vices; hence, they do not need the expectation of rewards or punishments as an incentive. Rejecting the view that without a belief in immortality no moral standards could be maintained, Pomponazzi repeats that a virtuous action without the expectation of a reward is superior to one that aims at a reward and concludes that those who assert that the soul is mortal seem to preserve the notion of virtue much better than those who assert that it is immortal. In thus stating that moral standards, as defined by the philosopher, do not depend on religious sanctions, he does not deny the validity of religious beliefs but asserts the autonomy of reason and philosophy, drawing upon certain passages in Plato and above all on Stoic doctrine and anticipating to some extent the views of Benedict de Spinoza and Immanuel Kant.

Having presented all arguments against the immortality of the soul, Pomponazzi states in the last chapter that the question is a neutral one, as is that of the eternity of the world. That is, he does not believe there are any natural reasons strong enough to demonstrate the immortality of the soul or to refute its mortality, although he knows that many theologians, notably Thomas Aquinas, have argued otherwise. Since the question is thus doubtful on purely human grounds, it must be resolved by God himself, who clearly proved the immortality of the soul in the Holy Scriptures. This means that the arguments to the contrary must be false and merely apparent. The immortality of the soul is an article of faith, for it is based on faith and revelation. It must thus be asserted on this ground alone and not on the basis of inconclusive or unconvincing rational arguments.

This conclusion and a similar one found in the *De Fato* have given rise to a variety of interpretations on the part of Pomponazzi's contemporaries and of modern historians. The statement made by some that Pomponazzi simply denied the immortality of the soul is patently false. He merely said that the immortality of the soul cannot be demonstrated on purely natural grounds or in accordance with Aristotle but must be accepted as an article of faith. This position is widely and somewhat crudely referred to as the theory of the double truth. The term is inadequate, for neither Pomponazzi nor anybody else ever said that something is true in theology and its opposite true in philosophy. What Pomponazzi did say, and what many respectable thinkers before and after him said, is that one theory—for example, that of the immortality of the soul—is true according to faith but that it cannot be demonstrated on the basis of mere reason and that its opposite would seem to be supported by equally strong or even stronger probable arguments.

This view has been called absurd by many modern historians and, ironically, by some who actually take a similar position themselves, though perhaps on other issues and with different words. Yet the persistent charge made against Pomponazzi and against many other medieval and Renaissance thinkers who took a similar position has been that the so-called theory of the double truth is merely a hypocritical device to disguise their secret disbelief and to avoid trouble with the church authorities. Thus, in saying that immortality cannot be demonstrated and that mortality may be defended by strong rational arguments whereas immortality is to be held as an article of faith, Pomponazzi, according to these historians, merely concealed his opinion that the soul was really mortal and substituted for it a formula that would protect him against ecclesiastic censure or punishment.

This is a serious and delicate problem. We cannot deny that a thinker of the past may have entertained opinions that we do not find expressed in his writings or that he may have put into writing views which he did not hold in his innermost heart. On the other hand, unless we have some text or document in support of this assertion, we are not entitled to claim that a thinker held some specific views that he failed to express in his writings or that are even in contrast with his expressed views. As a theologian of the eighteenth century said on this matter, we must leave it to God to look into Pomponazzi's heart and to see what his real opinion was. The human historian has no basis other than the written document, and the burden of proof, in history as in law, rests with those who want to prove something that is contrary to the overt evidence. Neither innuendo nor the assertions made by unfriendly critics or extremist followers can be accepted as valid evidence in lieu of some original statement or testimony concerning the author's view.

According to this standard, we have no real grounds for maintaining that Pomponazzi was hypocritical. The position he takes in the treatise on the immortality of the soul is fundamentally retained in two lengthy works composed afterward in defense of the first and, with a few dubious exceptions, also in his questions and class lectures. He was attacked by some theologians but defended by others, and his treatise was not condemned by the church authorities. The general position that immortality could not be rationally demonstrated, if not all the specific opinions that Pomponazzi associated with it, was held also by John Duns Scotus and even by the leading Thomist of Pomponazzi's time, Cardinal Cajetan. After the first excitement had passed, Pomponazzi continued to teach at a university located in the papal states, had among his students many clergymen who apparently found nothing offensive in what he said, and died peacefully as a widely respected scholar. The pupil who took his remains to his hometown and erected a monument for him was Ercole Gonzaga, later a cardinal and president of the Council of Trent. If there is any presumptive evidence, it hardly favors the opinion that Pomponazzi was a secret disbeliever or atheist.

INFLUENCE

Pomponazzi's influence, although not easily traceable, was considerable. The school of Italian Aristotelianism to which he belonged flourished for a hundred years or more after his death, and within this tradition his name remained famous and his views on such questions as the immortality of the soul and the unity of the intellect continued to be cited and discussed, if not adopted. The posthumous publication of several of his writings later in the century also gives testimony to his continued fame. His lectures and questions were copied in a large number of manuscripts, an indication of his popularity among his students; moreover, a considerable number of manuscripts containing the *De Incantationibus* and the *De Fato* prove that these works circulated widely, although, or perhaps because, they were not published during the author's lifetime. A few anecdotes associated with his name that we find in biographies, short stories, and dialogues of the period suggest that he made some personal impression even on the larger public outside university circles. He obviously was read by students and writers who did not belong to the Aristotelian tradition, and we may cite as an example Giulio Cesare Vanini, who seems to have used him as one of his favorite sources.

During the seventeenth century the Aristotelian school that had dominated the teaching of philosophy for such a long time finally lost its hold, especially in the field of natural philosophy, which was gradually replaced by the new mathematical physics of Galileo Galilei and his successors. Aristotelianism persisted much longer in the fields of logic, biology, and metaphysics. Yet because physics was the center and stronghold of medieval and Renaissance Aristotelianism, especially in Italy, most of Pomponazzi's specific teachings lost their immediate validity when the Aristotelian system within which he had developed his ideas came to be abandoned. Nevertheless, we may say that his view of the relation between natural reason and faith was capable of being reformulated in terms of the new physics and that in certain instances this did happen.

Even more important is another development. The seventeenth century, and still more the eighteenth, witnessed the rise and diffusion of free thought and overt atheism, especially in France. Some of the freethinkers who set out to discard faith and established religion came to consider the Aristotelian rationalists such as Pomponazzi as their forerunners and allies. Pomponazzi's treatise on the immortality of the soul was praised by the free thinkers and condemned by Catholic apologists, although moderate thinkers like Pierre Bayle tried to preserve a proper perspective. Pomponazzi's treatise was even reprinted in a clandestine edition with a false early date.

The use to which the French Enlightenment put Pomponazzi and the other Italian Aristotelians has had a strong influence on modern historians of the school, beginning with Ernest Renan. Again, a distinction is needed. It is one thing to say that Pomponazzi and the Aristotelians held the same views as later freethinkers, and it is another to state that they represent an earlier stage in a development that was to produce the views held by the freethinkers. In the first sense Pomponazzi was a forerunner of the freethinkers; in the second sense the evidence says he was not. Hence, we should not praise or blame him, depending on our own preferences and values, for being a freethinker, since we lack the factual basis for judgment. Yet in a different sense we may praise him. He belongs to the long line of thinkers who have attempted to draw a clear line of distinction between reason and faith, philosophy and theology, and to establish the autonomy of reason and philosophy within their own domains.

See also Alexander of Aphrodisias; Aristotelianism; Aristotle; Averroes; Averroism; Bayle, Pierre; Cajetan, Cardinal; Cicero, Marcus Tullius; Duns Scotus, John; Erasmus, Desiderius; Ficino, Marsilio; Galileo Galilei; Humanism; Kant, Immanuel; Luther, Martin; Pico della Mirandola, Count Giovanni; Plato; Plutarch of Chaeronea; Reformation; Renan, Joseph Ernest; Spinoza, Benedict (Baruch) de; Stoicism; Thomas Aquinas, St.; Valla, Lorenzo.

Bibliography

WORKS BY POMPONAZZI

Pomponazzi's *De Immortalitate Animae* is available in several modern editions—edited by G. Gentile (Messina, 1925), by G. Morra (Bologna, 1954), and by W. H. Hay II in a facsimile of the original edition of 1516 (Haverford, PA, 1938). It is also included, along with the *Apologia*, the *Defensorium*, and several shorter treatises, in the collection titled *Tractatus Acutissimi* (Venice, 1525).

The *De Naturalium Effectuum Causis Sive de Incantationibus* was printed separately in Basel (1556) and with the *De Fato* in the *Opera* (Basel, 1567); the volume contains only these two works. There is also a French translation—*Les causes des merveilles de la nature*, translated by Henri Busson (Paris, 1930).

There is a modern critical edition of the *De Fato* edited by Richard Lemay (Lugano, 1957). Of the *De Immortalitate* there is an English translation by W. H. Hay II, first published with the facsimile edition of the text and then in a revised and annotated version in *The Renaissance Philosophy of Man*, edited by Ernst Cassirer, Paul Oskar Kristeller, and John H. Randall Jr. (Chicago: University of Chicago Press, 1948).

For the unpublished works of Pomponazzi see L. Ferri, "Intorno alle dottrine psicologiche di Pietro Pomponazzi …," in *Atti della Reale Accademia dei Lincei, Memorie della classe di scienze morali, storiche e filologiche*, Series 2, Vol. 3 (1875–1876), Part III, 333–548; C. Oliva, "Note sull'insegnamento di Pietro Pomponazzi," in *Giornale critico della filosofia italiana* 7 (1926): 83–103, 179–190, 254–275; Bruno Nardi, "Le opere inedite del Pomponazzi," in *Giornale critico della filosofia italiana*, 29–35 (1950–1956); and Paul Oskar Kristeller, "A New Manuscript Source for Pomponazzi's Theory of the Soul …," in *Revue internationale de philosophie* 2 (1951): 144–157, and "Two Unpublished Questions on the Soul of Pietro Pomponazzi," in *Medievalia et Humanistica* 9 (1955): 76–101, and 10 (1956): 151.

WORKS ON POMPONAZZI

For the general background see Ernest Renan, *Averroès et l'averroïsme* (Paris: Durand, 1852); Bruno Nardi, *Saggi sull'Aristotelismo Padovano dal secolo XIV al XVI* (Florence, 1958), and John H. Randall Jr., *The School of Padua and the Emergence of Modern Science* (Padua, 1961).

For Pomponazzi's doctrine see Francesco Fiorentino, *Pietro Pomponazzi* (Florence: Successori Le Monnier, 1868); Andrew H. Douglas, *The Philosophy and Psychology of Pietro Pomponazzi* (Cambridge, U.K.: Cambridge University Press, 1910); E. Weil, "Die Philosophie des Pietro Pomponazzi," in *Archiv für Geschichte der Philosophie* 41 (1932): 127–176; Ernst Cassirer, *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*, Vol. I (Berlin: Cassirer, 1922), 105–117; and Bruno Nardi, *Studi su Pietro Pomponazzi* (Florence, 1965). For the immortality controversy see Giovanni Di Napoli, *L'immortalità dell'anima nel Rinascimento* (Turin: Società editrice internazionale, 1963).

See also Martin Pine, "Pomponazzi and Double Truth," *Journal of the History of Ideas* (29 [1968]: 163–176; John L. Treloar, John L., "Pomponazzi: Moral Virtue in a Deterministic Universe," *Midwest Studies in Philosophy* (26 [2002]: 44–55; John L. Treloar, "Pomponazzi's Critique of Aquinas' Arguments for the Immortality of the Soul," *Thomist* (54 [3] [1990]: 453–470; Margaret M. Van De Pitte, Margaret M., "Pietro Pomponazzi and the Debate Over Immortality," in *Philosophy and Culture, Vol. 3*, edited by Venant Cauchy (Montreal: Montmorency, 1988).

Paul Oskar Kristeller (1967)
Bibliography updated by Tamra Frei (2005)

POPE, ALEXANDER
(1688–1744)

Alexander Pope, England's leading poet of the Age of Reason, was born in London, the son of a prosperous Roman Catholic linen draper. His Catholicism barred him from public school and university; and he was educated by private tutors and by extensive reading and study on his own, largely at Binfield in Windsor Forest, where his father had retired. About the age of twelve, a severe illness stunted Pope's growth and deformed his spine, and for the rest of his life he was infirm. His devotion to poetry came early, and his genius was immediately recognized by William Wycherley and William Walsh. Early publications of note include the *Pastorals* (1709), *An Essay on Criticism* (1711), *The Rape of the Lock* (1712, enlarged 1714), and *Windsor Forest* (1714). During frequent visits to London, he became the friend of many prominent literary figures: Jonathan Swift, Joseph Addison, Richard Steele, John Arbuthnot, John Gay, and Lord Bolingbroke. Although not an ardent party man, Pope inclined more to the Tory than to the Whig. In 1718, after the death of his father, he removed to Twickenham, on the Thames near London. Pope's translations of the *Iliad* (1715–1720) and the *Odyssey* (1725–1726) were well received and financially successful. The edition of William Shakespeare appeared in 1725.

Author of the *Essay on Man* (1733–1734), *Moral Essays* (1731–1735), and *Imitations of Horace* (1733–1737), and of the *Dunciad* (1728–1743) and various other satires, Pope was a philosopher-moralist-poet. He was generally so regarded throughout the eighteenth century, both at home and abroad. There is little of the original in Pope's thought, nor did he pretend to any, the very notion of originality being distasteful to the rationalistic mind. In the *Essay on Criticism*, he stated that his aim was to present "What oft was thought, but ne'er so well expressed." His writing in general admirably fulfills this precept, and his memorable formulations of traditional and familiar ideas bear the stamp of literary genius.

Despite frequent allegations to the contrary, Pope was not a deist. Indeed, in the *Dunciad* he specifically attacks Anthony Collins, Bernard Mandeville, Thomas Morgan, Matthew Tindal, John Toland, and Thomas Woolston, the leading deists of the day. He eschewed the role of Christian (Catholic) poet, however, preferring to represent what he considered the best in Western thought, both pagan and Christian. His universality is best seen in the *Essay on Man*, where in Epistle I a rationalistic metaphysics is presented, centering on the "Great Chain of Being," a concept as old as Plato's *Timaeus* that was a part of the heritage of Western man and was influential until well into the eighteenth century. The rationalistic myth of a "chain of being" extending from the Godhead at the one extreme to the lowliest atom at the other, with man as the middle link between the pure reason of angelic spirits and the pure instinct of lower animals, is presented by Pope as a means of chastising presumptuous man for attempting to be too rational, for attempting to deny the earthbound aspect of his nature. Such generic "pride" on the part of man would necessarily push him into a higher link and thus destroy the entire chain. The moral is clear: "The bliss of Man (could Pride that blessing find)/Is not to act or think beyond mankind." Man must submit to his ordained place in the universe because "Whatever is, is Right."

Pope has been frequently ridiculed for ending Epistle I on this seeming note of "easy optimism," as it has been erroneously labeled. A moment's recollection, however, of the fact that Pope devoted much of his career to satirizing contemporary mores and morality will make it evident that his "optimism" was not ordinary or glandular optimism but strictly metaphysical optimism, which is not necessarily of any comfort to humankind. Granted the "chain of being" as ordained by Deity, that plan and that chain must be right, even though, according to the "principle of plenitude," evil is just as necessary as good. Thus, apart from the totality of cosmic rightness, many circumstances of life may not be good for man himself. Small comfort, therefore, to man to be assured that what seems evil to him personally is actually good from the cosmological point of view: God, but not man, can afford to be optimistic. In fact, the theme of the entire *Essay* is the problem of reconciling the contrary, apparently irreconcilable elements of man's nature with the infinite wisdom of a God of order and harmony. Thus it is that in the opening lines of Epistle II, Pope makes an effort to dismiss the prior metaphysical optimism with the homely precept: "Know then thyself, presume not God to scan;/The proper study of Mankind is Man." The remainder of the *Essay* is concerned with the world of real existence, insofar as this is possible given the background of rationalistic formalism. Epistle II treats of man as an individual; Epistle III treats of man and society; and Epistle IV treats of man and happiness. Here there is little "easy optimism."

Pope teaches that self-love is superior to reason and that the passions are requisite for action. The "dominant passion" (which varies from man to man) rules life in different ways, and virtue and vice are joined in man's mixed

nature. In the second epistle reason is "The God within the mind" that distinguishes between virtue and vice, to which in the third epistle are added instinct and social love. The fourth epistle, after much deliberation, declares that only in virtue is happiness to be found. Pope then ends the *Essay* with the affirmation that he has

> Shew'd erring Pride, *Whatever is, is Right*;
> That *Reason, Passion*, answer one great aim;
> That true *Self-Love* and *Social* are the same;
> That *Virtue* only makes our Bliss below;
> And all our Knowledge is, *Ourselves to Know*.

The major sources of Pope's philosophy have been much disputed, with Gottfried Wilhelm Leibniz, the earl of Shaftesbury, Bolingbroke, and William King the most frequently mentioned modern authors. There is no direct evidence that Pope knew Leibniz, and he specifically denied any influence by him. Pope had certainly read parts of Shaftesbury's *Characteristics* and undoubtedly acquired something from the reading. The case for Bolingbroke's *Fragments or Minutes of Essays* was widely accepted until recent investigations adduced evidence that the *Fragments* were composed later than Pope's *Essay;* what Pope may have received from Bolingbroke in the course of conversation, however, remains unknown. Archbishop King's *De Origine Mali* (1702), probably in Edmund Law's translation of 1731, contains much of the metaphysical thinking of the first epistle of the *Essay on Man*; and there is little doubt that Pope found much useful information and many references in Law's elaborate notes. Gleanings from the ancient Platonists, Neoplatonists, and Stoics are to be assumed, as are, of course, some from the Christian tradition.

The *Essay on Man* first appeared anonymously, and Pope did not claim it until 1735. On the Continent it was translated (poorly) into French prose in 1736 and the following year into French verse (even more poorly). It ran through several editions with considerable praise until attacked in 1737 by J. P. de Crousaz in his *Examen de l'essai de M. Pope sur l'homme*. The Swiss theologian, ignorant of English, deliberately used the poem as a means of assailing the Spinozistic and the Leibnizian philosophies, of which Pope was innocent. The attack was taken up by several English pamphleteers until William Warburton (later bishop of Gloucester and editor of Pope's *Works*), that colossus of controversy, came to the defense with a series of articles in the *History of the Works of the Learned*, published as a book in 1739 and revised in 1742. Warburton vindicated Pope against allegations of unorthodoxy, including that of deism.

Another Continental attack came in 1742 from Louis Racine in a poem titled *La religion*. In 1755 Gotthold Lessing and Moses Mendelssohn, in *Pope ein Metaphysiker!*, ridiculed both the Prussian Royal Academy for using a poet as the subject of a prize essay in philosophy and Pope for attempting to be a metaphysician in poetry. To Immanuel Kant, on the contrary, Pope was a favorite poet from whom he quoted frequently and whose thought he took seriously. Arthur O. Lovejoy has ventured the statement that "it would be hardly excessive to say that much of Kant's cosmology is a prose amplification and extension of the 'philosophy' of the First Epistle of the *Essay on Man*." Scorned or admired, at any rate, Pope's venture into verse philosophy was exceedingly popular, as is indicated by its translation into at least fifteen European languages and by scores of editions in English during the eighteenth century. And his century was the last that would have approved of such a venture.

Pope's original plan as poetical philosopher and moralist was ambitious, although somewhat vague. His magnum opus, to be titled "Ethic Epistles," was to consist of four books: the *Essay on Man*, as we now have it in four epistles; four more epistles dealing with "the extent and limits of human Reason," arts and sciences both "useful" and "unuseful," "the different Capacities of Men," and the "Use of Learning," science and wit; the "Science of Politics," to treat "of Civil and Religious Society in their full extent"; and "Private Ethics or Practical Morality." The plan—but not the philosophy—is curiously reminiscent of that of David Hume as stated in the "Advertisement" to the *Treatise of Human Nature* (1739). (Incidentally, Hume probably took from Pope such terms as "the science of man," "the science of human nature," "the soul's calm sunshine," and "the Feast of Reason.") In 1741 Hume was to devote an entire essay, "That Politics may be reduced to a Science," to the refutation of Pope's lines (*Essay on Man*, III, 303–304): "For Forms of Government let fools contest;/Whate'er is best admister'd is best."

The *Essay on Man* was the only part of the magnum opus completed as planned. However, the *Epistles to Several Persons*, commonly known as the *Moral Essays*, constitute part of the original design and would have been portions of the fourth book, "Private Ethics or Practical Morality." These four epistles or essays are "To Cobham" ("Of the Knowledge and Character of Men"); "To a Lady" ("Of the Characters of Women"); "To Bathurst" ("Of the Use of Riches"); and "To Burlington" (also "Of the Use of Riches"). Pope was always the philosopher-moralist-poet whose description of his own career (*Epistle to Dr. Arbuthnot*, ll. 340–341) is essentially accurate: "not in

Fancy's Maze he wander'd long,/But stoop'd to Truth, and moraliz'd his song."

See also Addison, Joseph; Bolingbroke, Henry St. John; Collins, Anthony; Deism; Gay, John; Hume, David; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Lessing, Gotthold Ephraim; Lovejoy, Arthur Oncken; Mandeville, Bernard; Mendelssohn, Moses; Morgan, Thomas; Neoplatonism; Plato; Platonism and the Platonic Tradition; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Stoicism; Swift, Jonathan; Tindal, Matthew; Toland, John; Woolston, Thomas.

Bibliography

Primary sources include *The Twickenham Edition of the Poems of Alexander Pope*, John Butt, general editor, 6 vols. in 7 (London, 1939–1961); and *Pope: The Correspondence*, edited by George Sherburn, 5 vols. (Oxford, 1956).

Secondary sources include Arthur Friedman, "Pope and Deism," in *Pope and His Contemporaries: Essays Presented to George Sherburn*, edited by J. L. Clifford and Louis A. Landa (Oxford: Clarendon Press, 1949); Arthur O. Lovejoy, *The Great Chain of Being* (Cambridge, MA: Harvard University Press, 1936); Geoffrey Tillotson, *The Moral Poetry of Pope* (Newcastle upon Tyne, U.K.: Literary and Philosophical Society of Newcastle upon Tyne, 1946), *On the Poetry of Pope*, 2nd ed. (Oxford, 1950), and *Pope and Human Nature* (Oxford: Clarendon Press, 1958); Jonathan Barnes, "Partial Wholes," *Social Philosophy and Policy* (8 [1] [1990]: 1–23; David J. Leigh, "Alexander Pope and Eighteenth Century Conflicts about Ultimacy," *Ultimate Reality and Meaning* (20 [1] [1997] 23–40.

Ernest Campbell Mossner (1967)
Bibliography updated by Tamra Frei (2005)

POPPER, KARL RAIMUND

(1902–1994)

Karl Raimund Popper, the Austrian philosopher of natural and social science, was born in Vienna and was a student of mathematics, physics, and philosophy at the university there. Although he was not a member of the Vienna circle of logical positivists and was in sharp disagreement with many of its doctrines, he shared most of the group's philosophical interests and was in close touch with several of its members, having a considerable influence on Rudolf Carnap. His first book, *Logik der Forschung,* was published in 1935 in the circle's series Schriften zur wissenschaftlichen Weltauffassung. In 1937 Popper went as senior lecturer to Canterbury University College in Christchurch, New Zealand, and remained there until his move in 1945 to a readership at the London School of Economics in the University of London. From 1949 to 1969 he was professor of logic and scientific method at the London School of Economics, and then became professor emeritus. He was knighted in 1964.

REJECTION OF VERIFIABILITY THEORY

The foundation of Popper's wide-ranging but closely integrated philosophical reflections is the bold and original form he first gave in 1933 to the problem of demarcating science from pseudo science in general and from metaphysics in particular. The logical positivists had taken this problem to be one of distinguishing meaningful from meaningless discourse and had proposed to solve it by making empirical verifiability the necessary condition of a sentence's meaningfulness or scientific status—in their eyes one and the same thing. Popper dissented both from their formulation of the problem and from their solution. His view had always been that the important task is to distinguish empirical science from other bodies of assertions that might be confused with it: metaphysics, such traditional pseudo sciences as astrology and phrenology, and the more imposing pseudo sciences of the present age, such as the Marxist theory of history and Freudian psychoanalysis. To identify this distinction with that between sense and nonsense is, he held, to make an arbitrary verbal stipulation. It is also an unreasonable stipulation because the line between science and pseudo science is neither precise nor impermeable. Pseudo science, or "myth," as he sometimes called it, can both inspire and develop into science proper: Indeed, the general progress of human knowledge can be considered as a conversion of myth into science by its subjection to critical examination.

FALSIFIABILITY CRITERION

A crucial difficulty for the verifiability theory of meaning was David Hume's thesis that inductive generalization was logically invalid. Being unrestrictedly general, scientific theories cannot be verified by any possible accumulation of observational evidence. Moritz Schlick sought to interpret scientific theories as rules for the derivation of predictive statements from observational ones and not as statements themselves at all, but this attempt came to grief on the fact that theories can be empirically falsified by negative instances. This logical asymmetry in the relation of general statements to observations underlies Popper's view that falsifiability by observation is the criterion of the empirical and scientific character of a theory. He maintained, first, that scientific theories are not, in fact,

arrived at by any sort of inductive process. The formation of a hypothesis is a creative exercise of the imagination; it is not a passive reaction to observed regularities. There is no such thing as pure observation, for observation is always selective and takes place under the guidance of some anticipatory theory. Second, even if induction were the way in which hypotheses were arrived at, it would still be wholly incapable of justifying them. As Hume showed, no collection of particular observations will verify a general statement; nor, Popper added, is such a statement partially justified or rendered probable by particular confirming instances, since many theories that are known to be false have an indefinitely large number of confirming instances.

For Popper the growth of knowledge begins with the imaginative proposal of hypotheses, a matter of individual and unpredictable insight that cannot be reduced to rule. Such a hypothesis is science rather than myth if it excludes some observable possibilities. To test a hypothesis, we apply ordinary deductive logic in order to derive singular observation statements whose falsehood would refute it. A serious and scientific test consists in a persevering search for negative, falsifying instances. Some hypotheses are more falsifiable than others; they exclude more and thus have a greater chance of being refuted. "All heavenly bodies move in ellipses" is more falsifiable than "All planets move in ellipses," since everything that refutes the second statement refutes the first but much that refutes the first does not refute the second. The more falsifiable a hypothesis, therefore, the less probable it is, and by excluding more, it says more about the world, has more empirical content. Popper goes on to show that the obscure but important concept of simplicity comes to the same thing as falsifiability and empirical content. The proper method of science is to formulate the most falsifiable hypotheses and, consequently, those that are simplest, have the greatest empirical content, and are logically the least probable. The next step is to search energetically for negative instances, to see if any of the potential falsifiers are actually true.

CORROBORATION

If a hypothesis survives continuing and serious attempts to falsify it, then it has "proved its mettle" and can be provisionally accepted. But it can never be established conclusively. The survival of attempted refutations corroborates a theory; the corroboration being greater to the degree that the theory is falsifiable. Popper's critics have fastened on this theory of corroboration as the point at which the inductive procedure he ostensibly rejects makes an implicit reappearance. Is there any real difference, they ask, between the view that a theory depends for justification on the occurrence of confirming instances and the view that it depends on the failure of falsifying ones to occur?

Furthermore, his critics claim, there is apparently an inductive inference embedded in Popper's doctrine—the inference from the fact that a theory has thus far escaped refutation to the conclusion that it will continue to do so. Popper could reasonably reply that the formal likeness between confirming and falsifying instances conceals an important difference in approach—that between those who glory in confirmations and those who ardently pursue falsifications. However, a certain disquiet about the inductivist flavor of the positive support that his theory allows a hypothesis to derive from the failure of attempted refutations is expressed in Popper's leanings toward a rather skeptical view of the status of unrefuted hypotheses: "Science is not a system of certain, or well-established, statements.… Our science is not knowledge (epistēmē): it can never claim to have attained truth, or even a substitute for it, such as probability.… *We do not know: we can only guess.*" *(The Logic of Scientific Discovery,* Ch. 10, Sec. 85, p. 278).

EMPIRICAL BASIS

To complete his account of the growth of scientific knowledge, Popper had to explain the empirical basis of the falsificatory operation, that is, he had to make clear the formal character of the observation statements that are logically deduced from theories. It follows from the falsifiability criterion that unrestricted existential statements of the form "There is (somewhere at some time) an *X*" are unempirical because however many spatiotemporal positions have been examined for the presence of an *X,* an infinity of further positions remains to be examined. This is not true, however, of circumscribed existential statements reporting the existence of something at a specified place and time. Popper takes the basic observation statements to be of this form, to refer to publicly observable material objects, and to be capable of being straightforwardly affirmed or denied as true or false. Such basic statements are motivated by perceptual experiences, but they do not, as they are held to in the usual empiricist tradition, describe them. They can themselves be empirically tested in the light of the further basic statements that follow from them, together with accepted scientific theories. The infinite regress that this conception involves is not a vicious one: It can be halted by a conventional assignment of truth to basic statements at any point. But

this convention is not dogmatic, since it is only provisional; if the basic statements in question are challenged, they can always be exposed to empirical tests.

EPISTEMOLOGY

In his later writings Popper drew many further inferences from his initial body of ideas. One is that knowledge has no foundations or infallible sources, either in reason or the senses. He sees the rationalist and empiricist epistemologies of the modern age as united in a determination to replace one sort of authority—a sacred text or an institution—with another—a human mental capacity. Both kinds of intellectual authoritarianism hold the mistaken opinion that truth is manifest and consequently that error is a sin and its propagation the outcome of some kind of conspiracy to deceive. There is no more comprehensive critique of the quest for certainty in the work of any other modern philosopher.

A second conclusion Popper drew is that the traditional empiricist account of concept formation—essentially Hume's idea that concepts are acquired by perceiving the similarity of sets of particular impressions—is mistaken because it embodies the same inductivist error as Francis Bacon's and J. S. Mill's accounts of scientific knowledge. Resemblance is not passively stumbled upon; rather, we classify things together in the light of antecedent preconceptions and expectations. Popper rejects innate ideas strictly so called but believes that we approach the world of experience with innate propensities—in particular, with a general expectation of regularity that is biologically explicable even if not logically justifiable. The influence of Immanuel Kant is especially evident in this side of Popper's thought. In a sense the proposition that nature contains regularities is for him synthetic a priori: It is neither a logical truth nor an empirical truth (since it is unfalsifiable), but it has a kind of psychological necessity as a general feature of the active human intellect.

THEORETICAL ENTITIES

Popper's dissent from the usual empiricist and positivist view that private, experiential propositions constitute the empirical foundation of knowledge and his insistence on the provisional and incompletable nature of scientific theorizing together determine his attitude to the subject matter or ontological significance of scientific theory. He rejects the essentialism of the rationalist philosophy of science, which conceives the goal of inquiry to be a complete and final knowledge of the essences of things, on the grounds that no scientific theory can be completely justified and that the acceptance of a new theory creates as many problems as it solves. He is equally opposed to the instrumentalist or conventionalist doctrine of those who, like Ernst Mach, Henri Poincaré, and Pierre Duhem, take the theoretical entities of science to be logical constructions, mere symbolic conveniences to assist us in the prediction of experience. The entities of scientific theory (such as molecules and genes) are not distinguishable in nature from the medium-sized public observables (such as chairs and trees) referred to in basic statements: Both are possible objects of genuine knowledge.

PROBABILITY

A difficulty arises for Popper's falsifiability criterion from the presence in normal scientific discourse of statements about probability in the sense of frequency. No finite sequence of *A*'s of which none are *B* decisively refutes the proposition that most *A*'s are *B.* In his first book Popper put forward a modified version of Richard von Mises's view that the probability of the occurrence of a property in an unrestrictedly open class is the limit of the frequencies of its occurrence in finite segments of the open sequence, a version that made probability statements accessible to decisive empirical refutation. Since then he had argued that probability statements, although they may rest on statistical evidence, should not themselves be interpreted statistically but rather as ascribing objective propensities to natural objects.

DETERMINISM AND VALUE

Popper's conviction that the mind is essentially active in the acquisition of knowledge and that its progress in discovery cannot be subsumed under a law and made the subject of prediction led him far beyond the philosophy of natural science, with which his central doctrines were concerned. Scientific knowledge is a free creation; it follows that the mind is not a causal mechanism. He contended that no causal model of the most elementary acts of the mind in empirical recognition and description can be constructed, since such a model would leave out the intention to name that is essential to any real act of description. Although the pursuit of knowledge is guided by an innate propensity to expect deterministic regularity in the world, the existence of knowledge as developed by a series of unanticipatable novelties is the strongest reason for rejecting general, metaphysical determinism.

Popper's theory of mind and knowledge also has ethical implications. Judgments of value are not empirical statements but decisions or proposals. Our valuations are not determined by our natural preferences but are the

outcome of autonomous acts of mind—a further link with Kant. Popper's own basic moral proposal was, however, not very Kantian. Popper was a negative utilitarian for whom the primary moral imperative is "diminish suffering."

HISTORY AND SOCIETY

In *The Open Society and Its Enemies* (1945) and in *The Poverty of Historicism* (1957), Popper applies his theory of knowledge to humankind and society in the form of an attack on historicism, the doctrine that there are general laws of historical development that render the course of history inevitable and predictable. In *The Open Society* historicism is examined in three influential versions, those of Plato, G. W. F. Hegel, and Karl Marx. In *The Poverty of Historicism,* historicism is formally refuted and attributed to two oppositely mistaken views about the nature of social science. The formal objection is that since the growth of knowledge exercises a powerful influence on the course of history and itself depends on the anomalous initiatives of original scientific genius, neither the growth of knowledge nor its general historical effects can be predicted. Some historicists have been motivated by the mistaken idea that a science of society would have a general evolutionary law as its goal. This is a naturalistic error. The evolutionary process is not a lawlike regularity at all; rather, it is a loosely characterized trend whose phases exemplify the laws of genetics, for example. The historicists who have made this error are right in believing that scientific method applies to society, but they have a false idea of what scientific method is. However, among historicists there are antinaturalists who hold that ordinary scientific method does not apply to society, for which laws of a special historicist form must be found. Popper asserts that scientific method applies both to nature and to society, and in the same way—to particular isolable aspects of the whole. Social science can discover laws that make clear the unintended consequences of human action, but there can be no laws of the whole system. It follows that social reform must proceed by piecemeal social engineering, not by total revolutionary reconstructions of the social order. Popper presents the central problem of politics in a characteristically falsificationist way: The question "Who should rule?," he says, should be replaced by the question "How can institutions be devised that will minimize the risks of bad rulers?"

PHILOSOPHY AND KNOWLEDGE

Popper did not believe, as do most analytic philosophers, that philosophy is sharply distinguishable from science, either in its methods—which, like science's, must be those of trial and error, conjecture and attempted refutation—or in its subject matter—which is not only language but also the world to which language refers. Furthermore, there is no uniquely correct philosophical method. Both the examination of actual language and the construction of ideal languages can contribute to the philosophical understanding of particular problems, but they are not universal keys to truth. Popper believed that if philosophy is to be of any general importance, it must stand in a close relation to the work of other disciplines. When it is isolated, as a special autonomous craft, from the general pursuit of knowledge, it degenerates into scholasticism and triviality.

See also Basic Statements; Carnap, Rudolf; Confirmation Theory; Conventionalism; Determinism in History; Duhem, Pierre Maurice Marie; Hegel, Georg Wilhelm Friedrich; Historicism; Hume, David; Induction; Kant, Immanuel; Laws, Scientific; Logic, History of: Modern Logic; Logical Positivism; Mach, Ernst; Marx, Karl; Philosophy of Science, History of; Plato; Poincaré, Jules Henri; Political Philosophy, History of; Probability and Chance; Progress, The Idea of; Schlick, Moritz; Scientific Method; Verifiability Principle.

Bibliography

WORKS BY POPPER

Books

Logik der Forschung. Berlin: Springer, 1935. Translated by Popper, with the assistance of Julius Freed and Lan Freed, as *The Logic of Scientific Discovery.* New York: Basic, 1959.

The Open Society and Its Enemies. 2 vols. London: Routledge, 1945; 4th, rev. ed., with addenda, London, 1961.

The Poverty of Historicism. London: Routledge, 1957; 2nd ed., with some corrections, 1961.

Conjectures and Refutations; The Growth of Scientific Knowledge. London: Routledge, 1963. Collected essays.

Essays

"Logic without Assumptions." *PAS,* n.s., 47 (1946–1947): 251–292.

"New Foundations for Logic." *Mind,* n.s., 56 (1947): 193–235; corrections and additions, n.s., 57 (1948): 69–70.

"Indeterminism in Quantum Physics and in Classical Physics," I and II. *British Journal for the Philosophy of Science* l (1950–1951): 117–133; 173–195.

"On the Theory of Deduction." Proceedings of the *Koninklijke Nederlandse Akademie van Wetenschappen* 51 (1 and 2) (1948).

"Probability, Magic or Knowledge out of Ignorance?" *Dialectica* 2 (1957): 354–374.

"The Propensity Interpretation of Probability." *British Journal for the Philosophy of Science* 10 (1959): 25–42.

WORKS ON POPPER

Bunge, Mario, ed. *The Critical Approach to Science and Philosophy.* Glencoe, IL: Free Press, 1964. Contains 29 articles largely concerned with the whole range of Popper's views. Includes a bibliography of Popper's publications, complete up to the beginning of 1964.

Kaufmann, Walter. *From Shakespeare to Existentialism.* New York, 1959. Chapter 7, "The Hegel Myth and Its Method," is a sympathetic but powerful criticism of Popper's account of Hegel in *The Open Society.*

Levinson, Ronald B. *In Defense of Plato.* Cambridge, MA, 1953. This substantial critique of Plato's modern opponents gives Popper pride of place.

Neurath, Otto. "Pseudorationalismus der Falsifikation." *Erkenntnis* 5 (1935): 290–294.

Reichenbach, Hans. "Über Induktion und Wahrscheinlichkeit. Bemerkungen zu Karl Poppers *Logik der Forschung.*" *Erkenntnis* 5 (1935).

Schilpp, P. A., ed. *The Philosophy of Karl Popper.* 2 vols. La Salle, IL: Open Court, 1974.

Warnock, G. J. Review of *The Logic of Scientific Discovery. Mind,* n.s., 69 (1960): 99–101.

Anthony Quinton (1967)

POPPER-LYNKEUS, JOSEF

(1838–1921)

Josef Popper-Lynkeus was an Austrian inventor, social reformer, and philosopher. Now almost completely forgotten, Popper enjoyed great fame in the early years of the twentieth century and on several topics his writings are far from dated.

LIFE AND WORKS

Popper grew up in the ghetto of the small Bohemian town of Kolin. At the age of sixteen he began his studies in mathematics and physics at the German Polytechnikum in Prague. Four years later he moved to Vienna, where he attended lectures first at the Imperial Polytechnikum and later at the University of Vienna. In spite of his acknowledged brilliance, Popper was not able to secure a teaching position, partly because he was Jewish and partly because of his radical opinions on religious and social questions. For some time he had a minor clerical job with the National Railways in southern Hungary. Returning to Vienna, he earned his living as a private tutor and as the owner of a scientific-technical literary agency. He attended scientific conferences and lectures, taking notes in longhand. These he wrote up, making ten to twelve carbon copies which he sold to the city's newspapers. In his autobiography, Popper recalls that during those years his income barely equaled that of the lowest-paid unskilled laborer. Popper's extreme poverty came to an end at the age of thirty with his invention of the so-called *Kesseleinlagen*—a device that significantly improved the working capacity of engine boilers. Although this, as well as several other of Popper's inventions, became generally used, he did not acquire wealth and it was not until he was almost sixty that he could retire from active participation in the production and selling of his various appliances in order to devote himself to literary pursuits.

During the last twenty years of his life, when Popper's books on social and philosophical questions had a very wide circulation, he became the center of what amounted almost to a cult. Popper's books give the impression of a man of transparent honesty and uncompromising hostility to every kind of humbug, especially of the kind that infested German public life in the late nineteenth and early twentieth centuries, but they do not, according to those who knew him, convey an adequate idea of his character and personal impact. His friends and admirers included Ernst Mach, Wilhelm Ostwald, Albert Einstein, Sigmund Freud, Arthur Schnitzler, Hermann Bahr, Stefan Zweig, Philipp Frank, and Richard von Mises. Mach referred to him as a "genius of freethinking"; Einstein, who visited Popper when a young man, spoke of him as a "saintly and prophetic person"; and all who met Popper were impressed by his deep serenity, warmth, and unusual and genuine kindness.

Popper was not a scientist of the first rank, but several of his publications dealing with problems in physics are favorably mentioned in standard histories of the subject. He was the first person to suggest the possibility of transmitting electric power, he was a pioneer in aerodynamics, and he was one of the first to see the full implications of the work of Robert Mayer. Popper's treatise "Über die Quelle und den Betrag der durch Luftballons geleisteten Arbeit" (On the sources and the amount of the work done by balloons; *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften,* 1875) led to correspondence with Robert Mayer, who requested Popper to review the second edition of his *Die Mechanik der Wärme* (*Mechanics of Heat,* 1874). Popper's article, published under the title "Über J. R. Mayer's Mechanik der Wärme" in the periodical *Das Ausland* (1876), did not confine itself to a discussion of Mayer's conservation principle but also contained a statement of a phenomenalistic philosophy of physics. In its "sharpness and fresh originality," according to Philipp Frank, "it equals the best that is found in Mach's works." In this essay there are also some

remarkably perceptive criticisms of the common view that the law of entropy implies the "heat-death" of the universe. In his later work, *Physikalische Grundsätze der elektrischen Kraftübertragung* (Physical principles of the transmission of electricity; Vienna, 1884), Popper emphasized the analogies between different forms of energy and suggested that every type of energy be regarded as a product of two factors, one of which can be regarded as a kind of quantity and the other as a "difference of level." This idea was subsequently employed in the "energetics" of Georg Ferdinand Helm and Ostwald, both of whom made due acknowledgment to Popper.

Popper's first work dealing with religious and social questions was published in Leipzig on May 30, 1878, the hundredth anniversary of Voltaire's death. It was titled *Das Recht zu Leben und die Pflicht zu Sterben, sozialphilosophische Betrachtungen, anknüpfend an die Bedeutung Voltaires für die neuere Zeit* (The right to live and the duty to die, social-philosophical reflections in connection with Voltaire's significance for our times). This work contains most of the ideas that Popper was to develop in later writings—a defense of the value of the individual in opposition to the national policies of all existing states, proposals for various social welfare measures totally at variance with the prevailing laissez-faire philosophy, recommendations for drastic reforms of the criminal law and judicial procedures, and reflections about the baleful influence of religion and metaphysics, accompanied by suggested methods for their elimination from the human scene. Both here and in a later more detailed study, *Voltaire, eine Charakteranalyse* (Voltaire—a character analysis; Vienna, 1905), Popper went out of his way to rebut the charges of German nationalists and romantics about Voltaire's disruptive (*zersetzende*) influence on morals and society, praising Voltaire for his great honesty, humanity, and courage, which, in Popper's opinion, were not matched by any of his German detractors.

In 1899 Popper published, under the pseudonym of Lynkeus (Lynkeus was the helmsman of the Argonauts, famous for his keen sight), a two-volume book titled *Phantasien eines Realisten* (Fantasies of a realist), which consisted of eighty sketches in the form of short stories or dialogues, most of them centering on some controversial philosophical or social topic. One story, "Gährende Kraft eines Geheimnisses" (The fermenting power of a secret), is set in fifteenth-century Florence and deals with the incestuous relations between a mother and her adolescent son, both of whom were burned at the stake. The *Phantasien* was banned in Vienna, and clerical members of the Austrian parliament demanded a criminal prosecution of the author. Since the book was published in Dresden and the German authorities took no action, it remained in circulation and went into no fewer than twenty-one editions. Philosophically of more interest than "Gährende Kraft eines Geheimnisses" are various sketches illustrating the influence of religion on human life, including an imaginary conversation between David Hume, Denis Diderot, Baron d'Holbach, and other outstanding figures of the French Enlightenment. One of the stories, "Träumen wie Wachen" (Dreaming like waking), independently arrived at several of the key doctrines of Freud's theory about dreams. Like Freud, Popper insisted that there is a continuity between waking thought and dream content and that dreams cannot be dismissed as "nonsense." Freud did not read Popper's story until after the first edition of *The Interpretation of Dreams* had been published, but later he repeatedly complimented Popper on his insights.

Of Popper's other books, three deserve special mention. *Über Religion* (Vienna, 1924), which was written in 1905 but could not be published before the overthrow of the monarchy with its clerical censorship, contains the fullest statement of Popper's criticism of religion and metaphysics. *Das Individuum und die Bewertung menschlicher Existenzen* (The individual and the evaluation of human lives; Dresden, 1910) is the most complete statement of Popper's individualistic ethics and his objections to the many theorists from G. W. F. Hegel to Friedrich Nietzsche whose writings bristle with contempt for the common man.

Popper himself regarded *Die allgemeine Nährpflicht* (Vienna, 1912) as his most important work. It develops in detail the system which, in Popper's words, should replace "our dreadful economic conditions" by such as are "good and moral." Society, according to Popper, has the duty to secure every individual against want, irrespective of his talents and qualifications. He classifies goods and services into "necessities" and "luxuries," the former including food, clothing, shelter, medical attention, and basic education. To ensure for every human being a "guaranteed subsistence-minimum," Popper proposes a term of labor service in the *Nährarmee* (Nourishment army). Utilizing an elaborate analysis of agricultural and industrial conditions in Germany at the beginning of the century, he calculates that twelve years of service by men and seven by women, working a thirty-five-hour week, would be sufficient for this purpose. There is to be a double economy: The provision of necessities is to be regulated by the state, while private enterprise is to handle the production and distribution of luxuries. After a person has completed his

term of service, he is free to work in any occupation he chooses, or not to work at all. In the latter event, he is still fully entitled to receive all "necessities." As technology advances, the period of service in the Nourishment Army will become progressively shorter.

Popper deliberately used the term *Nährpflicht* (literally "the duty to furnish nourishment") to express the key concept of his program, since it rhymes with *Wehrpflicht,* the German for compulsory military service, which Popper resolutely opposed. Popper's idea of a "compulsory civil service" is similar to one proposed by William James in his essay "The Moral Equivalent of War," but Popper anticipated James by several decades. If Popper's ideas about the duty of society to secure the individual against economic uncertainty do not sound exciting to the contemporary reader now that the concept of the welfare state is accepted by the majority of the populations of western Europe and the United States, and even the notion of a guaranteed income is advocated by leading economists, it should be remembered that at the time of their first publication, these ideas were extremely radical and were in fact received with violent hostility. In 1878 the great majority of political theorists, economists, and statesmen still adhered to the view that people are poor because of their laziness and ineptitude and that any state intervention in economic matters is a highly dangerous tampering with natural laws.

In spite of his courage and independent spirit, Popper failed to emancipate himself in some important areas of thought from the prejudices of his times. For example, he accepted without any question the view that masturbation "shatters" (*zerrütet*) the nervous system. He also had no doubt about the soundness of the prevailing hereditarian theories, according to which mental disturbances are largely the result of an innately weakened nervous system, and Popper frequently indulged in generalizations about the basically weak or strong nervous system of this or that national group. Although he knew of Freud's high esteem of his own work, Popper had no appreciation whatsoever of any of the ideas of psychoanalysis. Fritz Wittels, a psychoanalyst who was one of Popper's most devoted and trusted followers, called his attention to Freud's books and there was some polite correspondence between Popper and Freud. However, according to Wittels, Popper scarcely did more than look at Freud's books. In one case, when the subject was society (Freud's *Group-Psychology and the Analysis of the Ego*), Popper went to the trouble of reading the book. "I enjoyed what he quoted from the Frenchman [Le Bon]," he later told Wittels, but as for Freud's own theories, Popper added, "I must tell you that I did not understand one word."

THE SANCTITY OF HUMAN LIFE

None of Popper's theories is philosophically more interesting than the ethical individualism on which he bases his program of social reform. On the opening page of *Das Individuum und die Bewertung menschlicher Existenzen* (from now on referred to as *Das Individuum*) Popper announces what he calls his "motto," and the rest of the book consists of its elaboration and defense as well as of detailed criticism of the anti-individualist positions of various influential writers, including Hegel, Nietzsche, Thomas Carlyle, Herbert Spencer, Heinrich von Treitschke, and Popper's own friend Wilhelm Ostwald. Popper formulates the motto as follows:

> **Basic Principle of a Moral Social Order**
>
> When any individual, of however little account, but one who does not deliberately imperil another's existence, disappears from the world without or even against his will, this is a far more important happening than any political or religious or national occurrence, or the sum total of the scientific and artistic and technical advances made throughout the ages by all the peoples of the world.
>
> Should anybody be inclined to regard this statement as an exaggeration, let him imagine the individual concerned to be himself or his best beloved. Then he will understand and accept it.

To make clear what he means, Popper lists a number of propositions that he terms "the value-arithmetic" of human lives. The valuation of a person's life by the person himself, he writes, is something indefinite, varying, according to the mental state of the individual, from nothing to infinity. His life means nothing to him in moments of extreme unhappiness or when he is willing to sacrifice it for a cause in which he believes; but in other circumstances he regards it as possessing infinite value. "From an ethical point of view," Popper writes, "the existence of a stupid peasant-boy is just as infinitely valuable as the existence of a Shakespeare or a Newton" (*Das Individuum,* p. 193). "There is not the remotest equivalence," he remarks, "between the existence of a human being who wants to go on living and who is not trying to destroy another one, and any other value; the former exceeds the latter infinitely" (p. 189). Let us suppose that the angel of death were to allow William Shakespeare and Isaac Newton, in the most creative periods of their lives, to go on

living only on condition that we surrender to him "two stupid day-laborers or even two incorrigible thieves." As moral beings we must not so much as consider an exchange of this kind. It would be far better if Shakespeare and Newton were to die. One may call attention, as much as one wishes, to the pleasure produced in countless future ages by Shakespeare's plays; one may point to the immense progress of science which would be the consequence of the prolongation of Newton's life—by comparison with the sacrifice of a human being, these are mere "luxury-values."

However, all of these considerations, Popper repeatedly insists, apply solely to "non-aggressive individuals." A person whose life is threatened by another may, in self-defense, kill the aggressor without having to feel the slightest remorse or misgivings. In such a case, the person's own life rightly counts as something infinite, while the life of the aggressor, be he one or many, counts as nothing. It is in fact a person's *duty,* and not merely his right, to defend himself in such a case with all means at his disposal. In addition to helping himself, he also "exerts a beneficial influence on millions of others if he demonstrates to them by his example what importance and value a non-aggressive human being attaches to his life" (p. 218). In one place Popper goes so far as to assert that it would be better if all the aggressors in the world, even if they numbered millions, were to be destroyed than if a single human being succumbed to them without resistance.

On occasions Popper concedes that his own principles cannot be proved and that the principles of his assorted opponents cannot be disproved, but for the most part he maintains that they can be shown to be "true" by means of an "evident deduction" from premises granted by most civilized men (p. 64). He employs two types of arguments, the first of which consists in calling attention to the way in which civilized persons actually judge and behave in a great many situations, when their vision is not clouded by special bias or prejudice. Suppose, for example, a fire were to break out in the Louvre; in such a situation. Popper maintains, it would not occur to any of the firemen or any of the voluntary helpers to save the paintings in preference to the human beings present. If somebody were to save a painting and let a human being die, his behavior would be generally condemned and he might in fact be subjected to punishment. It is true, Popper admits, that sometimes when people *hear* that in a fire in some distant location a number of human beings perished but that certain valuable manuscripts or collections were saved, they respond with greater satisfaction than if it had been the other way around; but this only proves that distance from the place of a disaster produces indifference and makes people forget the enormous value of somebody else's life. "It becomes altogether different," Popper observes, "if one stands in front of the burning house." To take another illustration, in all civilized nations a person may not be subjected to vivisection or become the involuntary subject of a medical experiment, regardless of the benefits that might accrue to medical science and, indirectly, to future generations.

Popper also considers at great length another type of case that, in his opinion, shows particularly clearly that civilized people do in fact adhere to his principles. In fortresses or on ships, where the shortage of food may become so acute as to necessitate the sacrifice of some individuals, civilized men would always decide the issue by the casting of lots; in such a situation it would not occur to anybody to refer to the special literary or scientific talents of some member of the group. Shakespeare and Newton would here count no more than anybody else, and nobody would dare to propose that a less talented person be killed so that the great dramatist or the great physicist be kept alive instead. This is very evident in a case of this kind because "once the terror of death is so close, everybody perceives that the naked existence of a human being is something so elevated and infinite that compared with it everything else—be it genius, scholarship, or physical beauty—becomes quite inferior in value and a mere luxury" (ibid., p. 208).

The analysis of these and many other cases makes it clear, Popper contends, that his principles, which seem so strange and unrealistic when first stated in general terms, are quite commonly invoked. It is true that they are widely ignored when it comes to certain questions, such as compulsory military service, the death penalty, and the duty of society to guarantee the basic subsistence of every human being. However, in these cases it can be shown that people are simply inconsistent and have not perceived the implications of their own principles.

Popper's second type of argument, which is already indicated in his "motto," is much more interesting and original. It may not unfairly be labeled an ad hominem technique. Arguments of this type consist of two steps: (*a*) If a person, *X,* recommends a policy that involves the killing of one or more nonaggressive human beings, we extract from him the admission that the policy would not be justified if he, *X,* were the individual to be killed; (*b*) we then extract from him the admission that other human beings have the same right to live and not to be sacrificed to some biological, cultural, or aesthetic goal.

Popper observes that, except in special "periods of hate," most human beings are ready to make the latter of these admissions, at the very least for other members of their own nation or class. It does not, of course, mean, Popper explains, that a human being should mourn the death of any given person the way he mourns the death of somebody close to him; but human beings should realize that the mourning of somebody else in a similar situation is as justified as one's own and that to this other person his life or the life of somebody dear to him is more important than anything else in the world.

Popper employed his ad hominem strategy with relish in dealing with assorted philosophers and aesthetes who flaunted their readiness to approve the killing or enslavement of millions of ordinary human beings if this were necessary to achieve a biologically superior race or to produce great works of art. Thus, Popper devoted a good deal of attention to Spencer's conclusions that in giving artificial aid to the weakest members of a society, its physical and moral qualities are undermined and that, furthermore, all acts by the state to protect the weak and the sick are a "sin against the natural laws of life." After pointing out the dubious analogies on which such conclusions are based and the arbitrary preference for the value of future lives to those now in existence, Popper turns to his "frequently employed method." Suppose, he writes, Spencer or those taking such a "biological viewpoint" were themselves to become sick or unable to look after themselves. Would they approve of a society that turned to them and said: "Perish miserably! To help you is to make future generations less perfect." Will Spencer and his followers then be prepared to be treated as damaged goods, as refuse in a human breeding institution? Will they then still hold to the theories which they so calmly advocated while they were in good health and *others* were sick and in need of assistance?

Apparently nobody, not even the "monstrous" Nietzsche, irritated Popper more than the anti-Semitic historian and aesthete Heinrich von Treitschke, who in his essay "Der Sozialismus und seine Gönner" (Socialism and its patrons) had claimed that "the one statue of Phidias more than makes up for all the misery of the millions of slaves in Antiquity." One may well believe, Popper comments, that Treitschke can look at the statue of Phidias with great delight when *others* were compelled to labor as slaves. "A person holding such a view," Popper proceeds, "ought to have his own principles applied to himself to determine whether he will adhere to them after he has come to feel in his own person what they mean" (ibid., p. 166). It would have been a good idea to condemn Treitschke to five years of service as a slave and then offer him an apartment in the Berlin Museum, where he could spend all his days admiring antique statues. That would be the time to ask Treitschke how he feels about Phidias and the slaves. Perhaps this is the only method, Popper concludes, to make people like Treitschke have some respect for human life.

It would lead too far afield to attempt a detailed assessment of Popper's principles here, particularly of his rather curious "value-arithmetic" of human lives. A few words, however, are perhaps in order about his *ad hominem* technique, both because arguments of this kind are in fact very common (although few employ them with Popper's deliberateness and persistence) and because there may be a tendency to dismiss them too readily. Anybody with a training in logic is apt to regard all such arguments as flagrant instances of the fallacy of *ignoratio elenchi.* If a person makes a moral judgment but violates it in his own behavior, this is surely no argument against the soundness of the moral judgment. We all tend to smile at the familiar stage figure of the preacher of temperance who takes out his whiskey flask as soon as the congregation has departed, but his failure to practice what he preaches does not by itself invalidate his preaching—it does not even prove that he is insincere. A doctor, unable to break his own smoking habit, is not necessarily giving bad advice and also may be perfectly sincere when he advises his patients to stop smoking. Turning to one of Popper's examples, if Spencer, after becoming ill and helpless, were to abandon his views concerning the social or biological undesirability of aiding the weak, this would not disprove his views; nor, conversely, would it be evidence for Spencer's position if, upon falling ill, he refused all aid and cheerfully disintegrated in the belief that he was thereby promoting biological progress.

Yet surely this is not the end of the matter. In reading Popper, one cannot help feeling that he is doing a great deal more than expressing his indignation at the defenses of callousness and inhumanity by writers like Spencer, Nietzsche, and Treitschke. Granting that Popper's ad hominem arguments do not disprove the positions he attacks and that they do not prove his own ethical individualism, it might nevertheless be held that his strategy helps to bring out at least two points of some interest. In the first place, Popper may be said to call attention to a double use of "understand" and related expressions which seems of special importance in ethical controversy.

Bernard Shaw once remarked that nobody should be allowed to be a judge unless he had spent at least six months in prison. The average judge, he explained, does

not really know what he is doing when he sends a man to prison. In a sense this is no doubt false, but in another and deeper sense it may well be true. A judge can of course understand the statement "You are hereby sentenced to imprisonment for a period of five years" without having been a prisoner and even without having visited a prison—he obviously knows the difference in meaning between "two years" and "five years," and he also knows when to apply and when not to apply the word *prison.* At the same time, however, he might not know what he is doing in the sense that he has no clear conception of what it is like to languish for years in prison—what conditions really prevail in most prisons and what such a term of imprisonment frequently does to a man's character.

It may very plausibly be held that when intellectuals like Nietzsche, Spencer, and Treitschke advocate or condone the destruction or enslavement of millions of men, they do *not,* in this latter sense of the word, understand what they are recommending and that they could properly understand their own recommendations only if they became slaves or if they themselves experienced the prospect of being forcibly done away with. If we are satisfied that a person who recommends a certain policy does not himself understand, in this deeper sense, what he is recommending, this does not indeed show his policy to be mistaken, but it does undermine his standing in the discussion. For it means that he is ignorant of relevant, perhaps crucially relevant, facts, and hence, on almost any normative theory, his recommendation would not be adequately supported.

Second, Popper's strategy may help to determine the true *status* of the recommendations under discussion. Most people would want to make a distinction between a genuine moral or evaluative judgment and the mere expression of a desire or feeling; and it is the mark of the former but not of the latter—so, at least, a defender of Popper would argue—that it is universalizable: In passing a moral judgment on somebody, one is, in virtue of its being a moral judgment, committed to passing the same judgment about *anybody else* in similar circumstances, including oneself and those one cares for. Now, the writers whom Popper was opposing presumably wished their pronouncements to be treated as genuine evaluative judgments, as the advocacy of certain ideals and not merely as expressions of their desires. However, unless they were willing to maintain that they, too, ought to be enslaved or killed or left without assistance in order to further the goals in question, their original assertions will not qualify as genuine evaluations.

It will be instructive to see how Popper's challenge, thus interpreted, helps to determine the status of Treitschke's recommendation. Treitschke, we will assume, has just declared that certain "inferior" human beings ought to be enslaved for the purpose of producing a sublime work of art. Let us also assume that, in the sense under discussion, Treitschke admits that he, as well as his children (whom he loves), is "inferior." Now, if Treitschke, in this hypothetical situation in which he imagines himself and his children to be inferior, is ready to maintain that he and his children, no less than other inferior human beings, ought to be enslaved, his original declaration has the status of a genuine evaluative judgment. If, however, Treitschke wishes to exempt himself and his children, not merely in the sense that he would resist any attempt to be sold into slavery but in the sense of declaring that he and his children, although inferior beings, ought not to be enslaved, it would follow that his initial statement was not a genuine evaluation—that "ought" was not used there in its moral or evaluative sense. (More accurately: It would follow either that Treitschke was not offering a genuine evaluation or that he was inconsistent in denying a proposition entailed by one asserted previously.) Popper would probably add to this that in actual fact the great majority of those who talk like Treitschke, and very likely Treitschke himself, would insist that they and those they love ought not to be enslaved or otherwise mistreated. While it may be disappointing to realize that the callous positions against which Popper wrote have not been refuted, it is not a mean achievement to have shown that certain pronouncements masquerading as value judgments are in fact nothing more than the expressions of certain desires.

ELIMINATION OF RELIGION AND METAPHYSICS

Popper's positivism, like that of Mach, may be regarded as a midway stage between the philosophy of Auguste Comte and the logical positivism of the Vienna circle. Although he knew a great deal about mathematics, Popper did not advance beyond J. S. Mill's position that mathematical statements are extremely well supported empirical propositions. Metaphysics he dismissed as futile, but he wavered between dismissing metaphysical questions as meaningless and treating them as meaningful but unanswerable.

He never wavered, however, in regarding metaphysics, and more especially the theological varieties associated with Western religions, as exceedingly harmful. No change in economic arrangements, however rational

and beneficial it may be, can bring about a happy world unless all forms of supernaturalism are banished. There can be no peace in the world, Popper insists, as long as there is the slightest vitality in organized religious superstition, which is something "necessarily aggressive." Some of Popper's more conservative followers have done their best to play down his antireligious sentiments. It is therefore necessary to insist that he himself regarded the *Ausrottung* (extermination) of religion and metaphysics—and of all "enthusiasm for transcendent ideals"—as an essential part of his philosophical and social program, one that was necessarily implied by his humanitarian individualism. Margit Ornstein, his literary executor, relates how Popper, very shortly before his death, when be was revising the manuscript of *Über Religion,* remarked to her with a smile, "This is my Parthian arrow," adding, "When the Parthians left the battle scene they turned around once more to aim a final arrow at the enemy" (*Über Religion,* p. 3).

Purely ceremonial or "civil" religions, such as those practiced by the ancient Greeks and Romans or most of the people of China and Japan, are relatively harmless: Unlike the religions that we know in the West, they lack any kind of metaphysical foundation, anything that can be called a theological system, and above all, they do not possess a powerful priestly caste. Religion begins to have an evil influence only when it is given a systematic formulation and when it becomes "an affair of the heart." Popper's condemnation is sweeping and is meant to apply to the kind of belief fostered by rationalistic theologians no less than to the pietistic enthusiasm found in many religious groups all over the world. "At first it [religious zeal] is just nonsense, then it becomes obstinacy and spite, and in the end it is wildness and insanity beyond all limits" (*Über Religion,* p. 2). The harmfulness of religion is exactly proportional to the degree of religious fervor. Popper approvingly quotes Pierre Bayle's saying that "the person who is convinced that he is promoting the Kingdom of God by the extermination of heretics will step on all moral laws," and he offers numerous examples from the history of the "genuine positive," as opposed to the merely ceremonial and civil religions, to support his indictment that the former increase bad feeling in the world, that they encourage malicious tendencies which are then covered up and justified in high-sounding language, that they place love of man below the love of religious conceptions, that they multiply situations of strife and conflict by promoting the intervention of priests in even the most intimate details of everyday living, that they weaken and indeed destroy respect for truth and justice, and, finally, that they use, wherever they can, the power of the state for their purposes, especially in matters of education.

Popper disliked Christianity most of all, and in a section of *Dos Individuum* (A digression on the valuation of human lives in the Christian religion) he undertakes to correct the long-standing and, he claims, erroneous notion that Christianity encourages respect for the individual. Christianity does indeed speak of the value of the individual *soul,* but both in doctrine and in practice this notion has coexisted with contempt for the individual's body and life here on Earth. Popper does not deny that now and then religious belief has given people hope and consolation and that some of the expressions of religious devotion have been touchingly beautiful. However, such considerations must not be allowed to affect our overall judgment—"the burning of one heretic more than cancels ten thousand beautiful and deep feelings" (*Dos Individuum,* p. 72).

Popper had no doubt that the ideal of a "superstition-free culture," which, for him, meant a world without religion, was entirely attainable. He repeatedly takes issue with the widespread view that religious belief or religious needs are innate. This, he argues, is clearly disproved by the existence of entire nations without religion and of numerous persons in our own culture who are entirely devoid of religious belief and whose lives are no less happy or responsibly conducted than those of most believers. Moreover, the existing statistics on the prevalence of religious faith are suspect in the sense that, as far as religious issues are concerned, most people are not allowed to develop freely but live under the constant pressure of proreligious propaganda and the threat of social disapproval and economic loss if they avow their disbelief. "The masses of Europe," he writes, live in effect "in a religious penitentiary" (*Dos Individuum,* p. 59). Once the social and political power of the churches is shattered and education, uninfected by proreligious bias, becomes universal, religious belief is bound to vanish. "A person who has learned about the history and origin of religions, including Christianity, who has absorbed the main results of the sciences and the relations of these to the claims of religion, will not for a moment be afraid of or express gratitude to imaginary entities or persons" (p. 223).

Prior to the elimination of religious influences from the public schools, freethinkers must band together into a powerful "International League for the Liberation from Superstition." Such a league would publish and obtain the vast circulation of what Popper calls "counter-books"—works written in simple and clear prose, which would refute point by point the fallacies, the lies, and the distor-

tions in the religious and proreligious textbooks used in the schools. This league would also open "counter-schools" and train "wandering counter-preachers," whose function it would be to bring enlightenment to the peasant population. The counter-preachers would conduct meetings in the villages immediately after the Sunday services. In the beginning the peasants, incited by the priests, would try to chase away the "godless intruders," but with some courage and persistence it would be possible to receive a hearing, to catch the interest of the peasants, and in the end to make them see the soundness and good sense of the unbeliever's position. In his first formulation of this program in 1878, Popper estimated that such a "gigantic cleansing operation" would take several hundred years, but writing thirty years later, apparently encouraged by the constant decline in religious belief, he thought that a "few generations" would be quite sufficient.

In some places Popper admits that the teaching of science and of the history of religions and the exhibition of the conflict between scientific conclusions and religious assertions is not enough to banish supernaturalism. We also have to take into account the "metaphysical need" which is commonly found in Europeans, though it is for the most part lacking in the peoples of east Asia. This metaphysical need can be eliminated by "improved epistemological instruction." The metaphysical need is "nothing other than the longing to find a resting place in the exploration of the universe, to reach a stage at which there will be no urge to ask new questions" (p. 62). It is however, a senseless drive and must be recognized as such if we are to have a healthy mental constitution. Our knowledge of the world consists in the establishment of functional relations between experienced data (Mach's "elements"). Knowing the world means discovering correlations and subsuming these under ever wider correlations. "We cannot do anything further," writes Popper, "than to determine ever richer relations between elements already known or to insert new ones as connecting links between them." The world may be likened to a carpet spread out in front of us, between whose webs we go on weaving ever-new webs without limit. It is a vain effort "to try to see behind the carpet," as the metaphysicians and mystics do, in the hope of finding there all kinds of wonderful happenings. In discovering causal relations, "we do not descend step by step into the Ground of the World … rather we crawl like an insect on that colorful carpet which we call the world and which, as a consequence of our explorations, becomes ever more dense" (p. 63). This carpet has no "other side" transcending the one we explore.

See also Bayle, Pierre; Carlyle, Thomas; Comte, Auguste; Diderot, Denis; Dreams; Einstein, Albert; Ethics, History of; Freud, Sigmund; Hegel, Georg Wilhelm Friedrich; Holbach, Paul-Henri Thiry, Baron d'; Holism and Individualism in History and Social Science; Hume, David; James, William; Logical Positivism; Mach, Ernst; Mill, John Stuart; Newton, Isaac; Nietzsche, Friedrich; Ostwald, Wilhelm; Positivism; Voltaire, François-Marie Arouet de.

Bibliography

In addition to the works by Popper mentioned in the body of the article, the following deserve to be mentioned. *Fürst Bismarck und der Antisemitismus* (Vienna, 1886) is an examination of the violent anti-Semitic fulminations of Eugen Karl Dühring and Richard Wagner, as well as of the milder anti-Semitic arguments of Eduard von Hartmann. Popper's *Selbstbiographie* (Leipzig, 1917) reprints the complete text of "Über J. R. Mayer's Mechanik der Wärme," as well as the correspondence between Mayer and Popper. *Die Philosophie des Strafrechts* (Vienna: R. Löwit, 1924) presents the details of Popper's objections to existing penal systems and his own alternative, based on his ethical individualism. Parts of a major epistemological treatise that Popper had planned to write were posthumously published under the title "Über die Grundbegriffe der Philosophie und die Gewissheit unserer Erkenntnisse" in *Erkenntnis* 3 (1932–1933): 301–324.

Very little by Popper is available in English. "Dreaming and Waking," translated by A. A. Brill, can be found in *Psychoanalytic Review* 34 (1947): 188–197. The story about incest is translated by S. Rosenzweig as Appendix I of his article "The Idiocultural Dimension of Psychotherapy—Pre- and Post-History of the Relations between Sigmund Freud and Popper-Lynkeus," in *Psychoanalysis and the Social Sciences* 5 (1958): 9–50. Extracts from various of Popper's writings are translated in H. 1. Wachtel, *Security for All and Free Enterprise; A Summary of the Social Philosophy of Josef Popper-Lynkeus* (New York: Philosophical Library, 1955), which has an introduction by Einstein.

A. Gelber, *Josef Popper-Lynkeus, sein Leben und sein Wirken* (Vienna, 1922), and F. Wittels, *Die Vernichtung der Not* (Vienna, 1922), are full-length studies of Popper's life and work. The latter is available in English, translated by Eden and Cedar Paul as *An End to Poverty* (London: Allen and Unwin, 1925). There is a shorter but very informative study by Richard von Mises in Vol. VII of the series Neue Österreichische Biographic (Vienna, 1931), pp. 206–217. Popper's scientific work is discussed in P. Frank, "Josef Popper-Lynkeus zu seinem achtzigsten Geburtstag," in *Physikalische Zeitschrift* 19 (1918): 57–59; and in T. von Karman, "Lynkeus als Ingenieur und Naturwissenschaftler," in *Die Naturwissenschaften* 6 (1918): 457–463. Popper's contributions to "energetics" are discussed in G. Helm, *Die Energetik nach ihrer geschichtlichen Entwickelung* (Leipzig: Veit, 1898), Part VII, Ch. 2. A most interesting excerpt from the correspondence between Mach and Popper, containing a remarkable anticipation of the quantum theory, is reprinted

in H. Löwy, "Historisches zur Quantentheorie," in *Die Naturwissenschaften* 21 (1933): 302–303.

Freud's estimate of Popper is found in his article "My Contact with Josef Popper-Lynkeus," which is reprinted in Vol. V of Freud's *Collected Papers* (New York, 1959) and also in his *Character and Culture* (New York, 1963). Popper's remark about Freud quoted in this article will be found in F. Wittels, "Freud's Correlation with Popper-Lynkeus," in *Psychoanalytic Review* 34 (1947): 492–497.

In recent years there has been a good deal of discussion of ad hominem arguments of the kind employed by Popper against writers like Spencer, Nietzsche, and Treitschke. This discussion is in large measure due to the work of the influential British philosopher R. M. Hare, who in his *Freedom and Reason* (Oxford: Clarendon Press, 1963) employed a strategy strikingly similar to that used by Popper. Among discussions of how much (or how little) can be established by means of such arguments, the following are especially noteworthy: A. C. Ewing, "Hare and the Universalization Principle," in *Philosophy* 39 (1964): 71–74; D. H. Munro, "R. M. Hare's *Freedom and Reason.*" in *Australasian Journal of Philosophy* 42 (1964): 119–134; G. Madell, "Hare's Prescriptivism," in *Analysis* 26 (1965): 37–41; and G. Ezorsky, "Ad Hominem Morality," in *Journal of Philosophy* 63 (1966): 120–125.

There is a complete bibliography of writings by Popper on philosophical, political, and scientific topics in H. I. Wachtel, *Security for All and Free Enterprise* (see above).

Paul Edwards (1967)

POPULAR ARGUMENTS FOR THE EXISTENCE OF GOD

Argument about the existence of God is rare, for religious beliefs are effectively supported in our society by means that are not principally rational. It is common to answer the question "Why are you a believer?" with "Because I was taught to be," uttered in the tone of voice, or in the context, of one presenting reasons, not mere causes, of belief. It is even more common to speak of faith in God as if this were a specially compelling reason for belief and, moreover, one beyond logical criticism. Faith, however, is merely determination to believe and no kind of reason. Literature giving such justifications is not considered in this entry. Despite this omission of the greater part of the popular writing and what one might call the traditional verbal folklore of religion, a vast quantity of material remains that can be considered argumentative. After omitting further the grossest absurdities among these arguments, it has still been necessary to choose in a rather arbitrary way what should be dealt with, and no claim to completeness is made.

GENERAL REMARKS

Most of the arguments in popular literature may be seen as variants of the more strictly philosophical arguments, such as the Cosmological and Teleological arguments, or those from morals and common consent. The variants are popular largely because they are posed as probable rather than as valid arguments; that is, they are not offered as arguments whose premises entail their conclusions. Almost all of them fall into a common class of arguments of the form "The universe contains some puzzling feature, *F* (design, an objective morality). God's existence explains *F*, and no other known hypothesis does. Therefore, God exists." That they have this form is a matter of no small importance; it affects the whole question of what kind of objection is likely to succeed against a given popular argument.

It is beside the point to demonstrate the formal invalidity of such arguments, although their invalidity is very easy to show in almost every case. However, it is entirely relevant to require of such an argument that it should make clear just how God's existence explains *F.* (Similarly, the real force of the well-known infinite regress counter to the Cosmological, or First Cause, Argument, is that it demonstrates the failure of this argument to provide the promised explanation. The argument merely postpones the explanation. That God's nature is mysterious does not, of course, fill any explanatory bill.) On this score, popular arguments are universally unsatisfactory, appealing tacitly (for the most part) to the claim on the one hand that all things are possible to God and on the other that, God being a transcendental mystery, it is presumptuous to expect any account of his efficacy to be actually intelligible. As the substance of an explanation, this is thin. Further, it is an entirely relevant question to ask whether any explanation is required of some singled-out feature, and whether alternative explanations are simply not known or whether there appears to be a reason to suppose there are none.

ARGUMENT FROM COMMON CONSENT

The argument from common consent is an old and constantly recurring popular argument (see J. A. O'Brien, *God: Can We Find Him?*). The argument has a large measure of plausibility, despite the fact that it is formally invalid; for it is very often overwhelming evidence for some view that the majority holds it. For example, if a huge majority of spectators at a football game believes that a certain team won the game, that is exceedingly good evidence that this team indeed won it; and any

minority dissent can be written off in some way, such as irrational partisanship for the beaten team. However, the proportion of majority to minority views is not the only, or by any means the most important, factor in such situations. It is also crucial whether the majority has any competence to judge the issue. On the outcome of football games the majority of spectators is well placed to judge, but on the significance of some scientific experiment the majority is not at all well placed. Obviously, the general run of humankind has always been and still is poorly placed to pronounce on such a question as the existence of a Deity. This requires a competence in logical reasoning on highly abstract matters and an ability to assess complex evidence that the majority does not possess. Their vote carries no weight on this issue.

ARGUMENT FROM MORALS

An argument widely used, especially by evangelists who aim at the most general audience, is the argument from the intelligibility of morals. (On a more sophisticated level it has been argued by A. E. Taylor in *The Faith of a Moralist.*) Many who urge it seem to have dimly in mind an essentially rather sophisticated argument, encapsulated in naive remarks like "But if God doesn't exist, why do you not murder or plunder?" and "If God doesn't exist, then a morality could amount only to doing what you please." The rather sophisticated argument thus hinted at is as follows: To call an action moral (immoral) is, first, to provide a motive for doing (avoiding) it. Second, the claim that an action is moral can be a subject of rational dispute, which requires that the claim be not simply a disguised subjective remark about the speaker's tastes. The existence of God explains these two features of normal discourse. Therefore, God exists.

As was pointed out earlier, the first question must be "Does the existence of God explain these features of moral discourse?" If the question whether an action is moral is equivalent to the question whether the action is consistent with God's commands, then moral questions are not purely subjective. On the other hand, it is doubtful whether the theory accounts for the sort of discussion that actually goes on when moral issues are argued. God's commands must, according to the hypothesis, be arbitrary. It cannot be that he consults something beyond his own will, since that external thing or principle would then be the source of morality and God its mere interpreter and announcer, not its creator. However, moral reasoning surely requires empirical knowledge of other persons and the world generally—and a very great deal of intelligence if the reasoning is to be satisfactory. It is far from clear that the hypothesis allows for the relevant play of intelligence and knowledge in arriving at moral conclusions.

Again, it is rarely stated just which motive for behaving morally is provided under the hypothesis of God's existence. It cannot be suggested that we have a moral duty to obey God's commands because the whole point of the proposed explanation is that his commands are the source of all moral duties. It could be claimed that terror of punishment and desire for reward are perfectly adequate motives for obeying the commands. However, despite the undoubted efficacy of these motives, they are seldom urged because they do not adequately account for what we feel our motives really are in moral behavior. The most satisfactory suggestion as to the motive provided under the hypothesis seems to be that one obeys the commands out of love of God.

In sum, it is uncertain how the hypothesis clearly explains the required features of moral discourse. Further, it seems quite possible to account for them at least as well without being committed to the theistic view. For if love of God is an adequate motive for moral behavior, why should not love of one's fellows also be adequate? And if it is, then it further seems an objective empirical question that courses of action promote those almost universally desired ends of continuance of life, adequate food and shelter, and freedom from violence, as well as less fundamental and more subtle ends that promote smooth social intercourse.

TELEOLOGICAL ARGUMENTS

Versions of the classical Teleological Argument are by far the most popular of all popular arguments. The variety of changes rung upon this old theme in respect of its premises is astonishingly wide, as may be gathered from the following brief examples: The smallness of the human gene has been cited by A. C. Morrison, for no very clear reason, as an instance of God's designing hand, and so has the immensity of the orbital velocity of an electron. More markedly odd are such suggestions as "This old world has three times as much water as land but with all of its twisting and turning not a drop sloshes off into space" (*Ebony* symposium, November 1962, p. 96) and that the annual progress of Earth round the sun, although it is much more rapid, is also much smoother than the most sophisticated jet airliner yet designed. Although it is difficult to see what relevance these considerations may be thought to have, they perhaps involve a confusion between a good argument to the conditional conclusion that if these things are designed, then the technology of their produc-

tion is well beyond our present reach, and a bad argument to the conclusion that these things have, in fact, been designed.

An ingenious variant, heard in conversation but apparently never published, neatly turns the tables on a standard polemic against belief in a God that stems from Freudian psychology—that such belief is caused by a psychological mechanism arising from various sexual stresses in an infant's relationship with its father. This mechanism, it is claimed, far from showing that belief in God is pathological and irrational, really demonstrates his loving care for his creatures in providing a psychological mechanism that promotes belief, thus preventing the damnation of his creatures as heretics and infidels. This does not at all answer the point that insofar as belief depends upon the psychological stresses, it is irrational and pathological. (Irrational and pathological beliefs may, of course, be true.)

ARGUMENTS FROM THE SCIENCES

Only more recent arguments taken from the biological and the physical sciences will be discussed. First, however, there is a general argument from the very existence of science, or as it is more likely to be put, from the intelligibility of nature (see D. Elton Trueblood, *Philosophy of Religion,* pp. 94–98). It is felt that the universe must be rational if science, using logic and mathematics, is able to comprehend it. But logic and mathematics are concerned with deriving some propositions or formulas from others. It is not the conclusions or the premises of arguments that may properly be called rational, but only the procedure of deriving conclusion from premises. This procedure reflects no rational process in nature. It would be more accurate (although still not very accurate) to call this a linguistic procedure. We can move from "If there is lightning, then there is thunder" and "There is lightning" to the conclusion "There is thunder" by the rational procedure known as *modus ponens,* but it is not even intelligible to suppose that *modus ponens* is a natural physical process by means of which lightning produces thunder. Scientists may discover the important equation that relates the speed of a falling body to the square of the time of its fall. They may differentiate this equation, $v = t^2$, to show that the body's acceleration is constant. Differentiation is a mathematical procedure of derivation, but it is not intelligible to say that the body or the gravitational field in which it falls undergoes any such process of differentiation, or that it undergoes some nonmathematical counterpart of it.

ARGUMENTS FROM BIOLOGY. It has been argued—by Pierre André Lecomte du Noüy and Pierre Teilhard de Chardin, for example—that the pattern of evolution as displayed by modern biology shows clear marks of a designing hand. The direction of evolution, it is claimed, is toward progressively more intelligent life forms, thus showing the desire of the Creator (Omega, as Teilhard de Chardin called him) to bring about beings like himself. The claim is highly dubious. It induces "a certain shuffling of the feet" (to quote P. B. Medawar's review) in Teilhard, when he discusses the fact that insects and plants do not seem to evolve in this way at all. Lecomte du Noüy solved the difficulty by defining the problematic cases not as evolutions but as adaptations. The direction of adaptation is toward usefulness; that of evolution, toward liberty. Thus he made the claim perfectly, if trivially, safe. Even so, there is a difficulty, for if it is all a plan, why does God not bring about immediately and at a stroke the desired state of affairs now being so laboriously approached with such a plethora of wasteful products? Lecomte du Noüy's apparent answer is merely that since God is an eternal Being, what seems to us simple mortals as a drear immensity of wasted time is to him but the twinkling of an eye. The irrelevance of this to the original objection is obvious enough. The waste is still waste, and the existence of so many pointless dinosaurs (whose lives played no part in future evolution) can scarcely have escaped the attention of him who takes note of the fall of a sparrow.

One prevalent argument, put forward by Morrison, among others, is based on the allegedly remarkable hospitality of our planet to complex forms of life. Temperatures are neither too high nor too low, and there is an abundance of water and oxygen and an atmospheric blanket against lethal doses of cosmic radiation. But the argument inverts the situation. We now have good reasons (of a Darwinian kind) to believe that the surviving life forms are those that adapt to the environment rather than those for whom the environment has been adapted by a beneficent Overseer. So far as is known, only one of the nine major planets of our particular star is hospitable to complex life forms. It might be surprising if every planet of every star fulfilled the quite detailed set of conditions that favor life as we know it and that prevail over most (not all) of our planet. But that there is one such planet is not so surprising that we need recourse to metaphysical entities to explain it.

Similar arguments from alleged improbabilities also spring from biology. Lecomte du Noüy and others have claimed that life is inconsistent with the Second Law of

Thermodynamics. This law states that entropy increases, which means, roughly, that in any isolated system energy breaks down from various differentiated forms that are usable in doing work to an undifferentiated state of uniform heat. In statistical thermodynamics, increase of entropy is defined roughly as increase of the randomness of systems, that is, their movement toward more probable forms. But, it is said, living organisms decrease in entropy as they grow; they build up differentiated forms of energy and hence are improbable structures.

However, the phenomena of life are quite consistent with the law, for living organisms are not thermodynamically isolated systems. In whatever way life may be improbable, it is certainly not improbable in any sense that makes it inconsistent with statistical thermodynamics.

A second, more plausible, claim of this kind is that even a simple protein molecule is a highly improbable structure, so improbable that it is simply incredible that it should ever have come into existence by pure chance. A calculation cited by V. H. Mottram puts the odds against a chance "manufacture" of a simple protein molecule as 10^{160} to 1, a small chance by any standards. Mottram also claimed that 10^{243} years would be needed for such an event to occur on this planet (a much longer period than that accepted for the cool Earth) and that it would require sextillion sextillion sextillion times more material than is believed to be in the entire universe. Another calculation shows that the probability of such a molecule's arising by chance manipulation of amino acids (already quite complex structures) is still as low as $1{:}10^{48}$ and hence very improbable indeed.

The ways of statistical arguments are notoriously complex. We must always ask "Relative to what assumptions are these probability figures reached?" This was not made clear by Mottram. Presumably we are to assume at least that the atoms are rearranged in various positions by a process of mechanical shuffling of some sort in which all the rearrangements so envisaged are equally probable.

The possibility of such a rearrangement is very dubious. Even elementary chemistry informs us that certain combinations are not possible—for example, five hydrogen atoms may not be linked to one carbon atom. There is no evidence that such groups were excluded from the class of equiprobable arrangements considered in constructing this figure. If one considers the various linkages of more complex groups in which, say, a group of fifty atoms hooks on to another group of fifty, the number of chemically possible combinations is, presumably, very small. But this cannot have been taken into consideration in constructing the figures, because we do not have sufficient knowledge of the chemical possibilities at this level. The theists appear to have committed at this point the fallacy of assuming equal probabilities in cases where we have no positive knowledge of what the probabilities are.

Consider a liter of hydrogen containing, say, 10^{22} atoms. If we attempt to assign a number to all the conceivable arrangements of those atoms, the number is enormous. Yet we invariably find them divided into hydrogen molecules, 0.5×10^{22} pairs of atoms extremely close together. The improbability of this always coming about as a random arrangement of atoms is immense, and certainly far greater than any of the figures quoted by Mottram, yet this is presumably not evidence of design. Without more information about and justification of the assumption of equiprobability on which Mottram's calculation is based, plainly no reliance can be placed upon it.

ARGUMENTS FROM PHYSICS. Perhaps even more than biology, modern physics has given rise to a group of widely circulated arguments purporting to show that, despite the fact that God nowhere appears in the calculations of physicists, modern physics demands, suggests, or allows for the existence of God.

Although most apologists agree that the views of a scientist have no special authority outside the field of his expertise, this does not prevent their citing a vast mass of material produced by those physicists who spend their less strenuous hours philosophizing on their findings. The view almost universally favored among such writers, and perhaps most forcefully expressed by Sir Arthur Eddington and Sir James Jeans, is that modern physics establishes the subjectivity of all knowledge and that reality is mental, not material. It is often further concluded that physics has shown the world to be a nonrational place about which clear logical argument is out of place.

Relativity theories are alleged to have shown the subjectivity of all knowledge and to have confirmed Protagoras's doctrine that man is the measure of all things. But the special theory of relativity is concerned with relations between inertial systems (a notion definable wholly within objective dynamics). It is not at all concerned with any observers who may be reading clocks or using measuring rods within these systems. The general theory only extends the results of the special theory to cover relations between systems of a wider class. Neither theory is subjectivist or mentalistic.

A similar example of needless obscurantism concerns the primary place given the concept of energy by the relativistic notion that mass (matter) may be con-

verted into energy, and vice versa. Few of us are sure just what energy is; and, when a scientist such as E. J. Bing informs us that everything is energy, that it may exist in the form of electromagnetic vibration, and that it is a vehicle of universal thought (a gratuitous addition), we are apt to think that, while we do not know what this really means, perhaps everything is, in some obscure way, thought and hence in the mind of God.

Trueblood (op. cit., pp. 102–105) has invoked the science of thermodynamics to yield a theistic conclusion. The Second Law of Thermodynamics shows that the universe is steadily increasing its thermodynamic randomness—it is dissipating its stores of differentiated energy usable in doing work. It also shows that, as we trace the history of the universe in time according to the law, we come to a state of minimum energy, a sort of beginning in time of the universe. But this is far from lending support to the theistic hypothesis. It simply means that the law leads us to a point beyond which it will not take us. It gives no warrant for the conclusion that the minimum entropy state has a supernatural cause.

The greatest number of arguments are derived from the difficult and puzzling field of quantum mechanics. It is possible to give some indication of the relevant state of affairs in physics in terms of two features: (1) The Schrödinger wave equation, which is fundamental to quantum physics, contains the ψ function. This gives as its square the probability that an electron, for example, is in a certain spatiotemporal region. This feature leads to the result that the exact later states of electrons are unpredictable even from the fullest statement of their earlier states. (2) Beams of radiation or of electrons show some features characteristic of beams of particles but others characteristic of beams of waves, although their being particles is inconsistent with their being waves.

Feature (2) leads directly to such distortions as "If an electron can be two wholly inconsistent things, it is a little narrow to expect so much less of God." The electron, of course, is not, nor can it be, two inconsistent things—and (2) does not entail this. But the claim, together with the breakdown of the Laplacean view that given the complete mechanical state of the universe at any one time, any future or past state could be rigorously deduced in every detail, is generally hailed by religious apologists. Very few apologists claim that quantum physics actually provides evidence for God's existence. It is simply that in quantum theory mechanical determinism breaks down and there is no mechanical picture of quantum processes that is an adequate interpretation of the mathematical formalism of the theory. To religious apologists it appears that these facts allow for occult nonphysical causes and forbid rational understanding. They appear to feel that in the overthrow of reason itself lies their best defense.

More specific in their trend toward the admission of occult or physically transcendent causes are the following characteristic arguments. Arguing from the bad habit some physicists have of speaking about unpredictable electron jumps as the electron "choosing" one rather than another energy state, E. J. Bing wrote, "Let's call a spade a spade. To say that an electron 'chooses' to do anything *is to attribute free will to the electron.*" The theory gives no warrant for taking this obvious metaphor literally. It is quite unclear what real meaning there could be for such terms as *choice* and *free will* if their use is extended from describing living things to describing those that are nonliving. Such extension can result only in confusion.

Some physicists (Jeans, for example) have an equally deplorable habit of speaking of the Schrödinger wave equation as "waves of knowledge" in discussing the behavior of subatomic particles. This is presumably because the Schrödinger equation, which describes the behavior, is a wave equation and contains a function whose square is a probability. Apparently they regard probability as purely a matter of knowledge and thus suppose that some occult mental principle is at work in the quantum world. These suggestions won no assent from such authoritative quantum physicists as Niels Bohr and Werner Heisenberg, who most strongly insisted on the indeterminacy of quantum physics. Their notion is not that quantum phenomena have occult causes (acts of free will on the part of electrons) or unknown causes, but that they have no causes at all. Although there have been many distinguished scientists, including Albert Einstein, who believe it is possible that in the future we shall have a fully deterministic theory of the subatomic world, they have all taken for granted that the theory would postulate only physical causes.

See also Common Consent Arguments for the Existence of God; Cosmological Argument for the Existence of God; Degrees of Perfection, Argument for the Existence of God; Moral Arguments for the Existence of God; Ontological Argument for the Existence of God; Religious Experience, Argument for the Existence of God; Teleological Argument for the Existence of God.

Bibliography

FAVORABLE VIEWS

Bing, E. J. "Modern Science Discovers God." *American Mercury* (June 1941).

Corbishley, Thomas. *Religion Is Reasonable.* London: Burns & Oates, 1960.

Ebony, symposium. "Why I Believe in God" (December 1961 ff.).

Eddington, A. S. *Nature of the Physical World.* London: Cambridge University Press, 1928.

Gittelsohn, R. B. "Have We Outgrown God?" *Saturday Review* (September 16, 1961).

Griffith, A. L. *Barriers to Christian Belief.* London: Hodder and Stoughton, 1962.

Jeans, Sir James. *The Mysterious Universe.* London: Cambridge University Press, 1930.

Lecomte du Noüy, Pierre. *Human Destiny.* London: Longmans Green, 1947.

Lewis, C. S. *Miracles.* London: Bles, 1948.

Morrison, A. C. "Seven Reasons Why a Scientist Believes in God." *Reader's Digest* (October 1960).

Mottram, V. H. "Scientific Basis for Belief in God." *Listener* (April 22, 1948).

O'Brien, J. A. *God: Can We Find Him?* New York: Paulist Press, 1942.

Robinson, J. A. T. *Honest to God.* London: SCM Press, 1963.

Rosten, Leo, ed. *Guide to the Religions of America.* New York: Simon and Schuster, 1955.

Taylor, A. E. *The Faith of a Moralist.* London: Macmillan, 1930.

Taylor, F. Sherwood. *Man and Matter.* London: Chapman and Hall, 1951.

Teilhard de Chardin, Pierre. *The Phenomenon of Man.* Translated by Bernard Wall. New York: Harper, 1959.

Trueblood, D. Elton. *Philosophy of Religion.* New York: Harper, 1957.

Weaver, Warren. "A Scientist Ponders Faith." *Saturday Review of Literature* (January 3, 1959).

Whittaker, Sir Edmund. "Religion and the Nature of the Universe." *Listener* (June 1, 1950).

SKEPTICAL VIEWS

Cohen, Chapman. *God and Me.* London, 1946.

Cohen, Chapman. *God and the Universe.* London: Pioneer Press, 1946.

Feyerabend, Paul. "Niels Bohr's Interpretation of the Quantum Theory." In *Current Issues in the Philosophy of Science,* edited by Herbert Feigl and Grover Maxwell. New York: Holt, Rinehart and Winston, 1961. The relevant features of quantum physics are well discussed in this difficult but not highly mathematical paper.

Jack, H. "A Recent Attempt to Prove God's Existence." *Philosophy and Phenomenological Research* 25 (1965): 575–579. Critically discusses one of the arguments by A. C. Morrison and a very similar argument by Lecomte du Noüy in *Human Destiny.*

Medawar, P. B. "Critical Notice of 'Phenomenology of Man.'" *Mind* 70 (1961): 99–106.

Russell, Bertrand. *Religion and Science.* London: Butterworth-Nelson, 1936.

Russell, Bertrand. *The Scientific Outlook.* London: Allen and Unwin, 1931.

Russell, Bertrand. *Why I Am Not a Christian.* New York: Simon and Schuster, 1957.

Stebbing, L. Susan. *Philosophy and the Physicists.* London: Methuen, 1937.

OTHER RECOMMENDED TITLES

Behe, Michael. *Darwin's Black Box: The Biochemical Challenge to Evolution.* New York: Free Press, 1998.

Behe, Michael et al., eds. *Science and Evidence for Design in the Universe.* San Franciso: Ignatius Press, 2000.

Buckman, Robert. *Can We Be Good without God? Biology, Behavior, and the Need to Believe.* Amherst, MA: Prometheus, 2002.

Capra, Fritjof. *The Tao of Physics,* 4th ed. Boston: Shambhala, 2000.

Davies, Paul. *God and the New Physics.* New York: Simon and Schuster, 1984.

Davis, Jimmy, and Harry Poe. *Designer Universe: Intelligent Design and the Existence of God.* Broadman and Holman, 2002.

Dawkins, Richard. *The Blind Watchmaker.* New York: Norton, 1996.

Dembski, William. *Intelligent Design: The Bridge between Science and Theology.* Downers Grove, IL: InterVarsity Press, 2002.

Dembski, William, ed. *Mere Creation: Science, Faith and Intelligent Design.* Downers Grove, IL: InterVarsity Press, 1998.

Dennett, Daniel. *Darwin's Dangerous Idea: Evolution and the Meanings of Life.* New York: Simon and Schuster, 1995.

Hare, John. *Why Bother Being Good? The Place of God in the Moral Life.* Downers Grove, IL: InterVarsity Press, 2002.

Johnson, Philip. *Darwin on Trial.* Downers Grove, IL: InterVarsity Press, 1993.

Ross, Hugh. *The Fingerprint of God.* New Kensington, PA: Whitaker House, 2000.

Schroeder, Gerald. *The Hidden Face of God: Science Reveals the Ultimate Truth.* New York: Free Press, 2002.

G. C. Nerlich (1967)
Bibliography updated by Christian B. Miller (2005)

PORPHYRY

(c. 232–c. 304)

Porphyry, one of the principal founders of Neoplatonism, was born of Syrian parents at Tyre. He studied philosophy at Athens. In 263 he went to Rome, joined the group that regarded Plotinus as its master, and, apparently some years after Plotinus's death, took over his school. He died some time in the first six years of the fourth century.

Porphyry can be called a founder of Neoplatonism because, while the philosophy he upheld was in the main that of Plotinus, he made it possible for this philosophy to become, as it did, an institution throughout the Roman Empire. He arranged Plotinus's lectures for publication in their present form; he defended and developed their content in independent works of his own; third, he enabled

some of the much more systematic, not to say more teachable, philosophy of Aristotle to be included even by Platonic professors in a university curriculum.

In the so-called *Sententiae ad Intelligibilia Ducentes* (Aids to the Study of the Intelligibles; a short, difficult summary, incomplete as we have it, of Neoplatonism) he presents methodical proofs of two Plotinian theses which were unacceptable to the more conservative Platonists and to Porphyry himself when he first came to Rome: the independence and priority of the One to Being or Intellect, and the identity of Intellect or Thought with its objects. Plotinus, however, had been ambiguous over the extent to which the lower hypostases, Intellect (embracing the Platonic forms) and Soul (embracing nature and the Aristotelian forms), each existed in its own right. It is the monistic strand that seems to dominate in Porphyry: Everything that is not the One is an appearance of the One and is the result of the inadequacy of our thought about the One. The serious consequence of this doctrine is for the ordinary notion of personality. The individual, embodied soul and intellect, themselves appearances (he also calls them parts) of some universal soul and intellect, will be unreal; Porphyry calls the individual soul "the soul in a relation"—for it is related to a body—which implies its nonsubstantiality according to Aristotle's doctrine of categories. This consequence was vigorously challenged by Iamblichus. Union with the One can be achieved, according to Porphyry, by the unaided effort of intellect, but we do not have enough evidence to know how he met the philosophical problems of this thesis even if he had a consistent doctrine about it.

Porphyry's ethics followed Plotinus in stressing the universal equation between pursuit of the good, becoming what one "essentially" is, the self-awareness that accompanies thought, and "reversion" to the "cause" of one's being. Evil, together with matter, was the result of a "deviation from reality." In schematizing *Ennead* I 2 [19], Porphyry gave Plotinus's scale of virtues a nomenclature which became conventional for later Neoplatonists. *A*, the virtues of the soul, are (1) civic, (2) purificatory; *B*, the virtues of the intellect, are (3) contemplative, (4) paradigmatic. Less abstractly and on less philosophical grounds he was attracted like many Neoplatonists by the asceticism and taboos of Pythagoreanism.

Nothing has survived of a book that Porphyry wrote comparing Platonism and Aristotelianism. It undoubtedly maintained that there was no substantial conflict between the two, which was commonplace for Platonists of the empire. His commentaries on Plato have perished too; so have those on Aristotle, except for the introduction to the *Categories* known as the *Isagoge* and an elementary commentary on the same work. But his views were often quoted; and it is clear that what is distinctive about his treatment of Aristotle is twofold—a facility in expounding him without trying to Platonize him or to score against him, and a remarkable gift of clear exposition that does not depend (as it does in some later commentators) on ignoring the difficult issues. Most of the formulas that aimed at accommodating the metaphysical presuppositions of Aristotle's logic to Platonism had probably been worked out already. But since it was only the metaphysics that was objectionable, the way was open to the full acceptance of a purely formal logic. This meant not the Aristotelian logic of terms from which the nonexistent, the negative, and the particular were excluded, but something roughly equivalent to the Boolean algebra of classes.

This logic without metaphysics is roughly, too, what we find in Porphyry; and it is what has sometimes been inaptly called Porphyry's nominalism. With some debt to the Stoics, it enabled logic to develop as an autonomous science. For his *Isagoge* was translated into Arabic and Syriac as well as Latin, and his more advanced work was incorporated in Boethius's logic. The *Isagoge* is traditionally said to have made species a fifth predicable in place of definition. If it had it would have misrepresented Aristotle by implying that the subject was not a universal term, like those of the other predicables, but a particular. The implication might not have disturbed Porphyry, but in fact the *Isagoge*, or *Quinque Voces*, is not about predicables but what it says it is about, the five words that are essential to the understanding of the *Categories*. It does, however, introduce "inseparable accidents" which are an uneasy intermediate between essential attributes and pure or separable accidents.

Porphyry was a man of wide learning and wide interests. He studied many of the religious beliefs and practices with which he came into contact, and though generally sympathetic to them as various if inferior ways to salvation, he was renowned for centuries as the author of a detailed work against the Christians. But this and ventures of a more or less occultist nature—allegorical interpretations of poetry, descriptions of the soul's "vehicles," and the like—have mostly survived only in statements from controversial sources; and while respectable as philosophy in their day they are of small philosophical interest in the modern sense.

See also Logic, History of; Neoplatonism; Plotinus.

Bibliography

Porphyry's "Life of Plotinus" is included in editions of Plotinus' works.

Sententiae ad Intelligibilia Ducentes (Aids to the study of the intelligibles), edited by Erich Lamberz. Leipzig, 1975.

Opuscula Selecta, edited by A. Nauck. Leipzig, 1886, rpr. 1963.

On Abstinence from Killing Animals. Translation and commentary by Gillian Clark. London, 2000.

Isagoge, with Boethius' translation. In *Commentaria in Aristotelem Graeca*. Vol. IV, Part I, edited by A. Busse. Berlin, 1887.

Porphyry Introduction (Isagoge). Translation and commentary by Jonathan Barnes. Oxford, 2003.

Porphyrii Fragmenta, edited by Andrew Smith. Leipzig, 1993.

On Porphyry's life and works see:

Beutler, R. "Porphyrios." No. 21 in Pauly and Wissowa, *Realencyklopädie der classischen Altertumswissenschaft*. Vol. XXII, Sec. 1. Stuttgart, 1953.

Bidez, J. *Vie de Porphyre, Le philosophe néo-platonicien*. Ghent, 1913, rpr. 1964.

Dörrie, H., et al. *Porphyre*. Vandoeuvres-Genève, 1966.

Smith, Andrew. *Porphyry's Place in the Neoplatonic Tradition*. The Hague, 1974.

Smith, Andrew. "Porphyrian Studies since 1913." *Aufstieg und Niedergang der Römischen Welt* II 32 (1987): 717–773.

Zambon, Marco. *Porphyre et le Moyen-Platonisme*. Paris, 2002.

A. C. Lloyd (1967)
Bibliography updated by Andrew Smith (2005)

PORTER, NOAH

(1811–1892)

Noah Porter was an American Congregationalist clergyman, philosopher, and psychologist, and president of Yale College from 1871 to 1886. As a student in the Yale Divinity School, Porter had become a disciple of Nathaniel W. Taylor's modified version of New England Calvinism. For ten years he preached Taylorism at churches in New Milford, Connecticut (1836–1843), and Springfield, Massachusetts (1843–1846). He was then appointed Clark professor of moral philosophy and metaphysics at Yale, holding this chair throughout his tenure of the presidency of the college. On retiring from the office of president, he resumed a small teaching load until his death.

Porter's thought until 1853 was dominated by the conventional Scottish commonsense realism that pervaded American colleges. Then two years spent in Europe, largely in study at the University of Berlin, increased his familiarity with more recent and more daring philosophical systems. He became particularly interested, through the German philosopher Friedrich Adolf Trendelenburg, in the central epistemological problems of modern philosophy. Porter was convinced that these problems had to be solved before any advance in ontology could be expected. Moreover, he believed that the epistemological questions themselves required a foundation in scientific psychology.

This conviction and a much keener appreciation of the value of the history of thought than was usual among American philosophers of his time, led Porter to the preparation and publication of his important treatise *The Human Intellect,* the best work on psychology in English before William James. Porter presented and critically examined the leading ideas of both English and European (chiefly German) schools of psychology, as well as summarizing earlier work in the field. Because he regarded psychology as a necessary prelude to epistemology which, in turn, he considered prior to metaphysics, he insisted that psychology had to be an inductive science and roundly criticized G. W. F. Hegel for attempting to ground psychology in his metaphysical system. Although inductive, however, psychology cannot be a material or experimental science. Its subjects are the data of consciousness, which must be discovered introspectively; physiological experiments and investigations must be kept in mind by the psychologist, but these studies are ancillary to the direct study of the data of consciousness.

The influence of this major work and of Porter's many lesser writings was one of the chief forces in liberating academic philosophy in America from domination by naive realism and in introducing the study of German philosophy and psychology.

Among the nonphilosophical activities of Porter, special note should be taken of his editorship, with Chauncey A. Goodrich, of a revised edition of Noah Webster's *An American Dictionary of the English Language* (Springfield, MA, 1864). This work was revised under Porter's sole supervision as *Webster's International Dictionary of the English Language* (1890).

See also Common Sense; Consciousness; Hegel, Georg Wilhelm Friedrich; James, William; Psychology; Realism.

Bibliography

WORKS BY PORTER

The Human Intellect. New York: Scribners, 1868.

The American Colleges and the American Public. New Haven, CT: Chatfield, 1870.

Science and Sentiment. New York: Scribners, 1882.

The Elements of Moral Science. New York: Scribners, 1885.

Kant's Ethics. A Critical Exposition. Chicago: S.C. Griggs, 1886.

WORKS ON PORTER

Blau, Joseph L. *Men and Movements in American Philosophy.* New York: Prentice-Hall, 1952.

James, Walter T. *The Philosophy of Noah Porter.* 1951. An unpublished doctoral dissertation, available in typescript in the Columbia University library.

Merriam, George S., ed. *Noah Porter: A Memorial by Friends.* New York: Scribners, 1893. Biographical and expository discussion.

Schneider, H. W. *A History of American Philosophy.* New York: Columbia University Press, 1946. Contains additional comment.

J. L. Blau (1967)

PORT-ROYALISTS

See *Arnauld, Antoine; Logic, History of; Nicole, Pierre*

POSIDONIUS
(135–51? BCE)

Posidonius of Apameia, the Stoic philosopher, was famous in his own time and continued to influence writers into the first and second centuries CE. Soon after, his writings seem to have been lost, and even his name is rarely mentioned. Known to modern historiography mainly from the mention of his views in Cicero, Strabo, Seneca, and Galen, he was considered from the Renaissance to the beginning of the nineteenth century as a minor figure in the development of Stoicism. Then his thought began to be discovered in an ever-increasing number of writers, who were believed to follow him although they do not quote him, and he was established as the mediator between the Orient and the Occident, the reconciler of philosophy with religion and mysticism, the foremost representative of dualism. In the early twentieth century the reconstruction of Posidonius's work through *Quellenforschung* ("source criticism") was replaced by a reconstruction based on the inner form of his thought, and Posidonius was represented as a visual thinker, the defender of monism, the proponent of the doctrines of cosmic sympathy and vitalism, and the last Hellenistic philosopher. Both interpretations pay little attention to the fragments preserved under the name of Posidonius and therefore remain largely conjectural. What will be said here is based exclusively on the attested material.

This material leaves no doubt about the fundamentally dualistic character of Posidonius's system. His ethics, which is the best-known part of his thought, teaches, contrary to the general Stoic dogma, that passions are not simply false judgments but an irreducible force in human nature. This distinction is also echoed in Posidonian physics in the again unorthodox definition of matter as endowed with its own form and quality, which is merely reshaped and remodeled by divine reason. His logic establishes reason as a criterion of truth independent of sense perception. On the other hand, the duality of matter and reason is bridged by the realm of mathematical forms; among the Stoics only Posidonius was a mathematical realist. The macrocosm and the microcosm are in the end viewed as gradated, as hierarchies as it were, in which reason governs the subordinate irrational forces. God pervades the world; the passions follow the leadership of rational insight; man is here to contemplate and to act.

The Platonic and Aristotelian elements in this Stoicism were noted even by ancient critics. In Posidonius's opinion the founders of the Stoa, Zeno and Cleanthes themselves, had been Platonizing and Aristotelianizing. The strict monism of the school was due to Chrysippus, whose work Posidonius thought had to be undone. Yet although Posidonius harked back to the older teaching and in this sense remained in the Greek tradition—he was innocent of the later Orientalizing—he undoubtedly made an original contribution to philosophy. His ethics is a greatly refined analysis of the emotions that refutes the rationalistic position by pointing to its inner inconsistency and its inconsistency with observed facts. He stressed the importance of the will. Although only a few details of his physics can be rediscovered, it is clear that he was intent on explaining things; he was famous for his etiologies, and he carefully distinguished the various causes, assigning first place to teleology. Cosmic sympathy is but one of the factors he invoked in his exegesis of nature. His logical investigations furthered the understanding of syllogistic thinking, which seemed to him validated not by linguistic connections but by implied axioms. In short, his system marks a step forward in the history of Greek rationalism, and this is in accord with Posidonius's belief in the gradual development of knowledge and in the idea of progress, which he, like so many earlier Greek rationalists, upheld.

Posidonius's contributions were, however, not restricted to the field of philosophy proper. He wrote a history of his own time and in it, if not separately, dealt copiously with the rise of civilization, which he claimed began with practical inventions made by philosophers. In the historical process itself he detected the dominance of freedom over circumstance. Several of his books were devoted to natural sciences, such as astronomy and mete-

orology; he also investigated problems of mathematics and of military tactics. Perhaps the greatest significance of these works lies in the fact that they do not isolate scholarly and scientific research but put it in a philosophical framework. Events are seen as part of the history of the cosmos. Scientific explanations are hypotheses, the correctness and adequacy of which must be judged through philosophical reflection. It was as a philosopher that Posidonius felt impelled to reject the heliocentric theory in favor of the geocentric theory. Although he erred in this respect, he did enforce the idea of the hypothetical character of all scientific knowledge and did restore the unity of the sciences which Hellenistic thought had destroyed.

The stoa of the empire, initially influenced by Posidonius, tended more and more to follow Chrysippus. Thus, the philosopher Posidonius soon lost importance. His scientific writings kept the Greek heritage alive much longer and carried it, through Seneca's *Naturales Quaestiones,* into the Middle Ages. If one judges his achievement and his influence, one cannot compare him with Plato, Aristotle, or Democritus or with Zeno, Epicurus, or Plotinus. It is fair to say, however, that his personality, which he allowed to intrude into his work, makes him one of the most attractive figures among ancient philosophers. He was a man of dignity and not without a sense of irony and humor. He lived the dogma he preached, studying and teaching as well as participating in the political affairs of Rhodes, his adopted city. The variety of his gifts is amazing—his dialectical skill, traced by Galen to his mathematical erudition; the keenness of his powers of observation of men and things, which is especially marked in his reports on the travels that took him throughout almost the whole of the then-known world; and the strength of his analytical ability, along with his love of literature and art. It was perhaps the universalism of his nature that made it possible for him not only to attempt a new explanation of the universe in all its aspects, doing justice to both man's cognitive and his practical concerns, but also to root human existence—for the last time in antiquity, it seems—in the world of reality without depriving this world of the reign of human reason, which he considered of the same nature as the divine spirit ruling the cosmos.

See also Aristotelianism; Aristotle; Chrysippus; Cicero, Marcus Tullius; Epicurus; Galen; Hellenistic Thought; Leucippus and Democritus; Mysticism, History of; Plato; Platonism and the Platonic Tradition; Plotinus; Rationalism; Renaissance; Seneca, Lucius Annaeus; Stoicism; Vitalism; Zeno of Citium.

Bibliography

The approach of *Quellenforschung* has been criticized, and criticized fairly, by J. F. Dobson, "The Posidonius Myth," in *Classical Quarterly* 12 (1918): 179ff. For Posidonius as a monist see K. Reinhardt, *Poseidonios* (Munich: C.H. Beck, 1921); see also his *Kosmos und Sympathie* (Munich: C.H. Beck, 1926) and A. Pauly and G. Wissowa, *Realencyclopädie der classischen Altertumswissenschaft,* Vol. XXII (Stuttgart, 1953), Part I, Cols. 558–826. I. Heinemann, *Poseidonios' metaphysiche Schriften,* 2 vols. (Breslau, 1921–1928), considers Posidonius especially in relation to his predecessor Panaerius.

For a reconstruction of Posidonius's philosophy according to the attested fragments see Ludwig Edelstein, "The Philosophical System of Posidonius," in *American Journal of Philology* 57 (1936): 286ff. For the historical fragments see F. Jacoby, *Die Fragmente der griechischen Historiker,* Vol. II, No. 87 (Berlin: Weidmann, 1926). The collection of fragments by I. Bake, *Rhodii Reliquiae Doctrinae* (Leiden, 1810), is antiquated.

Ludwig Edelstein (1967)

POSIDONIUS [ADDENDUM]

Modern study of Posidonius has been transformed since the mid-1960s by the collection of ancient evidence compiled by Ludwig Edelstein and Ian G. Kidd (1972), and minutely analyzed by Kidd (1988, 1999), which contains only texts that name Posidonius explicitly. The picture presented is undoubtedly too narrow, and an accurate assessment of Posidonius's achievement and influence must await further study of other texts in which his influence may be reliably detected. But even the newly circumscribed picture has made it increasingly clear that Posidonius largely adhered to basic Stoic doctrines and principles and that his main innovations lie in his breathtakingly comprehensive effort to integrate both natural and human sciences into Stoic cosmology, epistemology, and ethics. His range was encyclopedic, and while the bulk of his massive output was in physics (embracing also metaphysics, theology, and the special sciences), there is little he neglected.

In metaphysics Posidonius sought to reconcile Stoic materialism with its quasi-dualist principles of matter and God (which are thoroughly blended together throughout the universe), and to explicate the incorporeal status of time, void, and bodily limits (points, lines, and surfaces). In logic relatively little is securely attested: work on the logic of relations and on axiomatic method in mathematics. He also analyzed the structure of scientific explanation (etiology): Subordinating mathematical

sciences to philosophy, he emphasized material and teleological factors in ways that suggest Aristotelian influence (Rhodes, where he worked most of his life, had a tradition of Aristotelian studies, and Andronicus of Rhodes [first century BCE], a younger contemporary, had a prominent role in reviving study of Aristotle's treatises).

Posidonius's scientific work had substantial impact on many later Stoics (notably Lucius Annaeus Seneca, Cleomedes [fl. c. 100 CE], and Geminus [10 BCE–60 CE]) and on ancient science and philosophy more widely (including Strabo [c. 64 BCE–after 23 CE] and Galen). Spanning astronomy, meteorology, geophysics, and geography, his work shows a concerted effort to extend the scope and empirical basis of Stoic theories. Problems he tackled include the size and distance of the sun and moon, the size and climatic zones of the earth, eclipses, comets, rainbows, clouds, thunder, winds, earthquakes, volcanoes, hydrodynamics, and mineralogy. Especially impressive is his theory of oceanic tides, which he correlated with the daily, weekly, and annual periodic motions of the moon; detailed observation here revealed systematic links between celestial and terrestrial phenomena that exemplify the principle of cosmic interaction (sympathy) underlying Stoic determinism and its providential design.

In ethics Posidonius upheld the central doctrines of Stoic Eudaemonism: virtue is a form of knowledge, only it (and anything possessing it) is genuinely good, and it is entirely sufficient for happiness (eudaimonia). He also brought new rigor to Stoic psychology by subjecting previous accounts of emotion and emotional behavior to precise critical analysis. Tendentious evidence in Galen has convinced many scholars that Posidonius rejected the monistic psychology of Chrysippus in favor of a Platonizing dualism, but recent studies (Cooper, Tieleman) argue that he sought rather to defend Stoic intellectualism by analyzing the structure of human motivation more closely. Similar concerns are evident in his massive *History* (fifty-two books covering 146 to 80s BCE—from a Roman defeat of federated Greece to an invasion of Athens), where ethics and ethnography combine with climatology and geography to explain both customs and historical events.

See also Aristotle; Chrysippus; Eudaimonia; Galen; Plato; Seneca, Lucius Annaeus; Stoicism.

Bibliography

TEXTS AND TRANSLATIONS

Edelstein, Ludwig, and Ian G. Kidd, eds. *Posidonius. Vol. 1, The Fragments*. New York: Cambridge University Press, 1972. The standard collection of ancient testimony.

Kidd, Ian G. *Posidonius. Vol. 2, The Commentary*. New York: Cambridge University Press, 1988. Meticulous analysis, with full reference to earlier scholarship.

Kidd, Ian G. *Posidonius. Vol. 3, The Translation*. New York: Cambridge University Press, 1999.

Theiler, Willy. *Die Fragmente/Poseidonios*. 2 vols. Berlin: de Gruyter, 1982. An insufficiently cautious collection of texts but often helpful for its commentary.

STUDIES

Algra, Keimpe. "Posidonius' Conception of the Extra-cosmic Void: The Evidence and the Arguments." *Mnemosyne* 46 (4) (1993): 473–505. Establishes his Stoic orthodoxy on some basic questions of cosmology.

Cooper, John M. "Posidonius on Emotions." In *The Emotions in Hellenistic Philosophy*, edited by Juha Sihvola and Troels Engberg-Pedersen. Dordrecht, Netherlands: Kluwer Academic, 1998. Argues that Posidonius's account has affinities with Plato but remains fundamentally Stoic.

Tieleman, Teun. *Chrysippus' On Affections: Reconstruction and Interpretation*. Leiden, Netherlands: Brill, 2003. Argues, contrary to prevailing views, that Posidonius sought to refine rather than undermine Stoic moral psychology.

Stephen A. White (2005)

POSITIVISM

The term *positivism* was used first by Henri, Comte de Saint-Simon to designate scientific method and its extension to philosophy. Adopted by Auguste Comte, it came to designate a great philosophical movement which, in the second half of the nineteenth century and the first decades of the twentieth, was powerful in all the countries of the Western world.

The characteristic theses of positivism are that science is the only valid knowledge and facts the only possible objects of knowledge; that philosophy does not possess a method different from science; and that the task of philosophy is to find the general principles common to all the sciences and to use these principles as guides to human conduct and as the basis of social organization. Positivism, consequently, denies the existence or intelligibility of forces or substances that go beyond facts and the laws ascertained by science. It opposes any kind of metaphysics and, in general, any procedure of investigation that is not reducible to scientific method.

The principal philosophical sources of positivism are the works of Francis Bacon, the English empiricists, and

the philosophers of the Enlightenment; but the cultural climate that made it possible was that of the eighteenth-century Industrial Revolution and the grand wave of optimism to which the first successes of industrial technology gave rise. Positivism made this climate into a philosophical program—that is, a universal project for human life. It exalted science without concerning itself (as does contemporary positivism) with the conditions and the limits of the validity of science, and it claimed that not only ethics and politics but also religion would become scientific disciplines. In one direction, this led to an attempt to establish a "positive" religion in place of traditional theological religions.

Through its acceptance of the concept of the infinity of nature and of history and, therefore, of necessary and universal progress, positivism had affinities with the other important nineteenth-century philosophical movement, absolute idealism, and belongs with it in the general range of romanticism.

There are two fundamental kinds of positivism: social positivism, with a professedly practicopolitical character, and evolutionary positivism, with a professedly theoretical character. Both share the general idea of progress, but whereas social positivism deduces progress from a consideration of society and history, evolutionary positivism deduces it from the fields of physics and biology. Comte and John Stuart Mill are the principal representatives of social positivism, and Herbert Spencer of evolutionary positivism. A materialistic or spiritualistic metaphysics is often associated with evolutionary positivism. A third, critical type of positivism, also known as empiriocriticism, should be distinguished from both social and evolutionary positivism. Contemporary forms of positivism—logical positivism and neopositivism—are directly connected with critical positivism.

SOCIAL POSITIVISM

Social positivism arose in France through the work of Saint-Simon and other socialistic writers (Charles Fourier, Pierre Joseph Proudhon) and in England through that of the utilitarians (Jeremy Bentham and James Mill), who, in turn, considered their work closely associated with that of the great economists Thomas Malthus and David Ricardo. Social positivism sought to promote, through the use of the methods and results of science, a more just social organization. According to Saint-Simon, men now lived in a critical epoch because scientific progress, by destroying theological and metaphysical doctrines, had eliminated the foundation of the social organization of the Middle Ages. A new organic epoch, in which positive philosophy would be the basis of a new system of religion, politics, ethics, and public education, was required. Through this system society would regain its unity and its organization by basing itself on a new spiritual power—that of the scientists—and a new temporal power—that of the industrialists. In his last writing, *The New Christianity* (1825), Saint-Simon considered the new organic epoch to be a return to primitive Christianity.

COMTE. Saint-Simon's ideas inspired the work of Auguste Comte. The point of departure of Comte's philosophy is his law of the three stages. According to this law, both the general history of humanity and the development of the individual man, as well as that of every branch of human knowledge, passes through three stages: the *theological,* or fictitious, stage in which man represents natural phenomena as products of the direct action of supernatural agents; the *metaphysical* stage, in which the supernatural agents are replaced by abstract forces believed to be capable of generating the observable phenomena; and, finally, the *positive* stage, in which man, refusing to seek the ultimate causes of phenomena, turns exclusively toward discovering the laws of phenomena by observation and reasoning. The positive stage is that of science, whose fundamental task is to predict phenomena in order to use them.

"Science whence comes prediction; prediction whence comes action" is the formula in which Comte epitomized his theory of science. The formula, as Comte himself recognized, expresses exactly Francis Bacon's point of view. The law of the three stages permits the classification of the sciences according to the order in which they entered into the positive phases—an order determined by the degree of simplicity and generality of the phenomena which are the objects of each science as it reaches the positive phase. Thus, according to Comte the following hierarchy constitutes "a necessary and invariable subordination": astronomy, physics, chemistry, biology, and sociology. Mathematics remains outside this order because it is at the basis of all the sciences; psychology, because it is not a science, also remains outside. Psychology should be based on introspective observation. But introspective observation is impossible, because the observed and observing organ would have to be identical. The apex of the hierarchy of sciences is sociology, or social physics, which Comte divided into social statics, or theory of order, and social dynamics, or theory of progress.

Progress is a necessary law of human history: The realization of progress is entrusted not to individuals, who are only the instruments of progress, but to the true subject of history—humanity, conceived as the Great Being in which past, present, and future beings partake. "We always work for our descendants, but under the impulse of our ancestors, from whom derive the elements and procedures of all our operations" (*Politique positive,* Vol. IV, pp. 34–35). Humanity is the continuous and uninterrupted tradition of the human race, and it is the divinity that must replace the God of traditional religions. The wisdom and providence of humanity preside infallibly over the realization of progress. At the end of progress there is sociocracy, a new absolutist social regime based on science and the religion of humanity and directed by a corporation of positivist philosophers. Sociocracy, by limiting liberties, will make impossible any deviation from the fundamental beliefs of the positivistic cult.

In his last work, *Philosophy of Mathematics* (1856), Comte proposed a new kind of religious trinity, the Great Being (humanity), the Great Fetish (Earth), and the Great Way (space). The religious aspect of Comte's philosophy drew a great number of followers and generated the greatest wave of enthusiasm. Pierre Lafitte and Émile Littré in France, Richard Congreve and G. H. Lewes in England were the most philosophical of Comte's first disciples. The influence of Comte's religious thought, however, rapidly exhausted itself, except among small groups of devotees, while his philosophical ideas (the law of the three stages; the conception of science as description and prediction; the theory of progress; and sociology as a positive science) have exercised a lasting influence on science and philosophy.

BENTHAM AND THE MILLS. Comte's English contemporaries, the utilitarians Jeremy Bentham and James Mill, presented with equal force, although more modestly, the fundamental requirement of positivism: that every kind of valid knowledge be included within science. They sought to establish a science of mind based on facts, as is the science of nature, and tried to make ethics itself, as Bentham used to say, an "exact science." They considered the mind to be an associative mechanism, ruled by precise laws whose constitutive elements are sensations, which were regarded as the ultimate facts of mind. Traditional ethics was substantially a theory of the end of human conduct: It established by a priori means what that end was and deduced from it the rules of conduct. Bentham and Mill intended to substitute for traditional ethics a theory of the motives of conduct—that is, of the specific causes of conduct. If it were ascertained what are the motives and the rules that human beings obey, Bentham and Mill believed, it would be possible to direct human conduct in the same way that nature can be controlled by knowing its causal laws.

These principles remained fundamental in later developments of positivism, first in the work of John Stuart Mill, who was influenced by both Saint-Simon and Comte. Mill, like Saint-Simon and Comte, spoke of reorganizing society on new foundations. He rejected, however, the doctrinaire political and religious absolutism of Comte and defended instead the freedom and development of the individual, to whom he considered the social organization subordinate. Mill's classic *Principles of Political Economy* (1848) concluded by determining the limits of governmental intervention in economic affairs—limits required so that there would be in human existence "a sacred fortress safe from the intrusion of any authority."

Mill's *System of Logic* (1843), which is perhaps the most important work of nineteenth-century positivism, contains a fundamental correction of Comte's view of science. Comte had stressed the rational aspect of science and considered its experimental basis, the verification of facts, as merely preparatory to the formulation of laws. He had excluded the notion that once they were formulated, laws could again be subjected to the test of facts and eventually placed in question by "a too detailed investigation," and he had prescribed for scientific investigation a series of limitations to keep it from being transformed into "a vain and at times a seriously disturbing curiosity." Mill's logic, instead, appealed to a radical empiricism and avoided any dogmatizing of scientific results. The very principles of logic, according to Mill, are generalizations of empirical data, and induction is the only method that science has at its disposal. The basis of induction itself, the principle of the uniformity of the laws of nature, is, in turn, an inductive truth, the fruit of many partial generalizations. Prediction is possible in science only on the basis of past experience, which alone furnishes the evidence both for the major premise and for the conclusion of the traditional syllogism. "'All men are mortal' is not the proof that Lord Palmerston is mortal; but our past experience of mortality authorizes us to infer *both* the general truth and particular fact with the same degree of certainty for one and the other" (*System of Logic,* Bk. II, Ch. 3).

Like the other utilitarians, John Stuart Mill held that the human mind has the same structure as natural phenomena and is knowable in the same ways. "If we knew the person thoroughly, and knew all the inducements

which are acting upon him, we could foretell his conduct with as much certainty as we can predict any physical event" (*System of Logic,* Bk. VI, Ch. 2, 2). To make such predictions possible, he held that a new science, ethology, was needed to study the laws of the formation of character. Mill placed this science alongside Comtian sociology, to which he attributed the task of discovering the laws of progress that make it possible to predict social events infallibly (ibid., Ch. 10, 3).

Mill held that even religion should be based on experience. Experience, by suggesting that there is a limited and imperfect ideological order in nature, permits belief in a divinity of limited power, a kind of demiurge. Such belief encourages a religion of humanity based upon an altruistic ethics and the "supernatural hopes" of humankind.

SOCIAL POSITIVISM IN ITALY AND GERMANY. In Italy social positivism had two defenders, Carlo Cattaneo and Giuseppe Ferrari. Both were influenced by the work of Saint-Simon, and both saw him as a continuer of the work of Giambattista Vico, whom they credited with having founded "a science of man in the very heart of humanity."

The German social positivists Ernst Laas, Friedrich Jodl, and Eugen Dühring appealed to Ludwig Feuerbach rather than to Saint-Simon and Comte. But faith in science, in progress based on science, and in a perfect social form to which this progress must lead was the inspiration of all social positivists.

EVOLUTIONARY POSITIVISM

Evolutionary positivism shared the faith in progress of social positivism but justified it in a different way. Evolutionary positivism is based not on society or history but on nature, the sphere of physics and biology. Its immediate forerunners were the work of the geologist Charles Lyell and the doctrine of biological evolution. Lyell, in *The Principles of Geology* (1833), demonstrated that the actual state of Earth is the result not of a series of cataclysms (as Georges Cuvier had argued) but rather of the slow, gradual, and imperceptible action of the same causes that are acting before our eyes. The doctrine of evolution triumphed in 1859 with the publication of Charles Darwin's *Origin of Species,* which first presented adequate proofs of biological evolution and formulated the doctrine in a rigorous way. Lyell's and Darwin's doctrines made possible the formulation of the idea of a natural and necessary progress of the whole universe, beginning with a cosmic nebula and, through the uninterrupted development of the inorganic and organic world, continuing into the "superorganic" development of the human and historical world. It is superfluous to note that the scientific theories that furnish the occasion for the rise of the idea of evolutionary positivism do not constitute the elements of a sufficient proof of it, since it is so highly generalized a hypothesis that it seems to be of a metaphysical nature. Darwin himself remained "agnostic" (to use the term created by another biological evolutionist, T. H. Huxley) with respect to all problems that concern the universe in its totality.

SPENCER. The importance of Herbert Spencer, however, and the lasting influence of his work, depends on his defense of universal progress as a continuous and unilinear evolution from a primitive nebula to the more refined products of human civilization. Spencer used the term *evolution* in preference to progress in an early programmatic article of 1857, and even then he saw universal progress as modeled on biological evolution. His definition of evolution as "the passage from the homogeneous to the heterogeneous" or from the simple to the complex was suggested by the development of vegetable and animal organisms, whose parts are chemically and biologically indistinct at first but which then differentiate to form diverse tissues and organs. Spencer held that this process can be discovered in all fields of reality and that each of these fields has a specific science whose task is to recognize and clarify its characteristics. Philosophy is (as Comte conceived of it) the most generalized knowledge of the process of evolution. The role of philosophy begins with the widest generalizations of the individual sciences; from these generalizations it seeks to realize a "completely unified" knowledge. However, neither philosophy nor science, according to Spencer, can take the place of religion.

The truth of religion is that "the existence of the world with all that it contains and all that it encompasses is a *mystery* that always needs to be interpreted" (*First Principles,* London, 1862, Par. 14). All religions, however, fail in giving this interpretation; therefore, the sole task of authentic religion is to serve as a reminder of the mystery of the ultimate cause. The task of science, on the other hand, is to extend indefinitely the knowledge of phenomena. Like William Hamilton and Henry Mansel, Spencer held that human knowledge is enclosed within the limits of the relative and the conditioned, that is, within the limits of phenomena. Beyond these limits there is the unlimited and unknown force on which all phenomena depend. The unknowability of this force is revealed in the insolubility of certain problems at the limits of philosophy and science, such problems as those concerning the

essence of space, of time, of matter, and of energy, the duration of consciousness (whether finite or infinite), and the subject of thought (whether it is the soul or not).

If Comte's religion of humanity had little success among philosophers and scientists, Spencer's agnosticism found many adherents among them, and for a few decades it was a required attitude for intellectuals generally. Other positivists, however, such as Roberto Ardigò, rejected agnosticism and denied that one could speak of an "unknowable" in an absolute sense. Ardigò;, moreover, wanted to redefine the process of evolution by considering it as "a passage from the indistinct to the distinct," referring to psychological experience rather than to biology.

Spencer wrote on many fields of knowledge—biology, sociology, ethics, politics, and education. When he turned his attention to sociology, he attempted to rescue it from the practical and political task that Comte had assigned to it and to consider it as a theoretical discipline whose task is to describe the development of human society to its present state. This change was accepted by such positivist sociologists as John Lubbock, Edward Tylor, Émile Durkheim, and William Graham Sumner, who were strongly influenced by Spencer.

Evolutionary positivism is, in its more rigorous form, as far from materialism as it is from spiritualism. Spencer affirmed (*First Principles,* Par. 194) that the process of evolution can be interpreted both in terms of matter and movement and in terms of spirituality and consciousness: The Absolute that it manifests can be defined neither as matter nor as mind. Positivism embraces both trends that interpret the concept of evolution materialistically and trends which interpret it spiritualistically. The laws of the conservation of matter discovered by Antoine Lavoisier (1789) and the laws of the conservation of energy implicit in Robert Mayer's discovery of the equivalence of heat and work (1842) were taken as proofs of the hypothesis that a single substance, of which matter and energy are inseparable attributes, is the eternal subject of cosmic evolution and necessarily determines all its characteristics.

HAECKEL AND MONISM. The German philosopher Ernst Haeckel termed the view that matter and energy are inseparable attributes of one basic substance "monism" and utilized it to combat the dualism that he held was proper to all religious conceptions based on the duality of spirit and matter, of God and the world. Haeckel also found a decisive confirmation of biological evolution and of its necessity in what he termed the "fundamental biogenetic law" of a parallelism between ontogeny, the development of an individual, and phylogeny, the development of the species to which that individual belongs. Monism was accepted by many chemists, biologists, and psychologists and became popular through the diffusion of Haeckel's writings and of such other works as Ludwig Büchner's *Force and Matter* (1855).

Monism also inspired literary and historical criticism. A passage from the introduction to Hippolyte Taine's *History of English Literature* (1863) has remained famous as an expression of this tendency: "Vice and virtue are products just as vitriol and sugar are, and every complex datum is born from the encounter of other simpler data on which it depends."

LOMBROSO. The positive school of penal law, founded by Cesare Lombroso, drew its inspiration from materialistic and especially from deterministic positivism. This school taught that criminal behavior depends on inevitable tendencies which are determined by the organic constitution of the delinquent. The structures of this constitution would be analyzed by a corresponding science—criminal anthropology.

WUNDT. Evolutionary positivism was also interpreted spiritualistically, notably by Wilhelm Wundt, who sought to substitute "psychophysical parallelism" for materialistic monism. Wundt's doctrine was that mental events do not depend on organic events but constitute a causal series by themselves and correspond point for point to the series of organic events. He made this doctrine the basis of his psychological investigations (Wundt founded the first laboratory of experimental psychology), and for many decades it remained the working hypothesis of experimental psychology. Wundt cultivated, moreover, a "psychology of peoples" that is descriptive sociology, in Spencer's sense. Like Spencer, Wundt intended it to be the study of the evolutionary process that produces institutions, customs, languages, and all the expressions of human society.

INFLUENCE OF EVOLUTIONARY POSITIVISM. Evolutionary positivism has left as a legacy to contemporary philosophy the idea of a universal, unilinear, continuous, necessary, and necessarily progressive evolution—an idea that forms the background and the explicit or implicit presupposition even of many philosophies which do not recognize their debt to positivism and which, in fact, argue against it. The idea of evolution is fundamental to the philosophies of C. S. Peirce, William James, and John Dewey, as well as to those of George Santayana, Samuel

Alexander, and A. N. Whitehead. Some of these philosophers have sought to remove the necessitarian character from the idea of evolution and to include within it an element of chance or freedom (Peirce, James, Dewey) or of novelty and creativity (Henri Bergson, C. Lloyd Morgan). Bergson, who interpreted evolution in terms of consciousness and insisted upon its creative character, explicitly acknowledged his debt to Spencer (*La pensée et le mouvant,* 3rd ed., Paris, 1934, p. 8). It is not without reason that his disciple Édouard Le Roy termed Bergson's doctrine a "new positivism," which means a new spiritualistic interpretation of cosmic evolution.

The vitality and the broad diffusion of the legacy of positivism is no sign of its validity. No scientific discipline is as yet able to adduce any sufficient proof in favor of a unilinear, continuous, and progressive cosmic evolution. In fact, in the very field where the phenomena of evolution have been most closely considered—biology—evolution seems to lack precisely those characteristics that positivism attributes to it.

CRITICAL POSITIVISM

EMPIRIOCRITICISM. In the last decade of the nineteenth century, positivism took on a more critical form through the work of Ernst Mach and Richard Avenarius. In Germany and Austria this critical positivism was known as empiriocriticism. Mach and Avenarius both held that facts (which for them, as for the other positivists, constituted the only reality) were relatively stable sets or groups of sensations connected to and dependent on each other. Sensations are the simple elements that figure in the constitution both of physical bodies and of perceptions or consciousness or the self. These elements are neutral, neither physical nor psychical, and every substantial difference between the physical and the psychical disappears. From this point of view, a "thing" is a set of sensations and the thought of the thing is the same set considered as "perceived" or "represented." For Avenarius, however, the process of interiorization, which he called introjection, and by which the thing is considered as a modification of the subject or as a part of consciousness, is a falsification of "pure" (that is, authentic or genuine) experience. For Avenarius and Mach, science, and knowledge in general, is only an instrument that the human organism uses to confront the infinite mass of sensations and to act in the light of those sensations in such a way as to conserve itself. The function of science is, therefore, economic, not contemplative or theoretical. It conforms to the principle of least action, and its end is the progressive adaptation of the organism to the environment.

Theories concerning concepts, scientific laws, and causality very different from those of classical positivism are the chief results of empiriocriticism. According to Mach a concept is the result of a selective abstraction that groups a large number of facts and considers those elements of these facts that are biologically important—that is, those adapted to excite the appropriate reaction in the organism. Since the variety of the biologically important reactions is much smaller than the variety of facts, the first task is to classify and simplify the facts by means of concepts, each of which constitutes the project of an appropriate reaction. And since the interests with which people confront facts are different, there are different concepts which refer to the same order of facts. The laborer, the doctor, the judge, the engineer, and the scientist all have their own concepts, and they define them in those restricted ways which are appropriate for stimulating the reaction or set of reactions in which each is interested.

The concept of law, which classical positivism conceived of as a constant relationship among facts (a relationship which in turn was considered as a fact) underwent a radical transformation in critical positivism. The Englishman Karl Pearson, in *The Grammar of Science* (1892), gave a kind of *summa* of the fundamental principles of the science of the time. Although Pearson's work utilized Machian concepts, it supplied Mach himself with many inspirations. Pearson affirmed that scientific law is a description, not a prescription: It "never explains the routine of our perception, the sense-impressions we project into an 'outside world.'" Instead of description, Mach preferred to speak of a restriction that the law prescribes on our expectation of phenomena. In any case, he added, "Whether we consider it a restriction of action, an invariable guide to what happens in nature, or an indication for our representations and our thought which bring events to completion in advance, a law is always a limitation of possibilities" (*Erkenntnis und Irrtum,* Leipzig, 1905, Ch. 23).

Mach and Pearson sought to free the notion of causality from the notion of force, which they regarded as an anthropomorphic interpolation. Mach held that the mathematical notion of function should be substituted for that of cause. When science succeeds in gathering various elements into one equation, each element becomes a function of the others. The dependence among the elements becomes reciprocal and simultaneous, and the relation between cause and effect becomes reversible (*Die Mechanik in ihrer Entwicklung,* 4th ed., Leipzig, 1901, p. 513). From this point of view, time, with its irreversible order, is real at the level of sensations and as a sensation.

The time of science is, on the other hand, an economic notion which serves for the ordering and prediction of facts.

Along the same lines, a disciple of Mach, Joseph Petzoldt, proposed to substitute for the principle of causality the "law of univocal determination," which would also be applicable to cases of reciprocal action. According to this law, one can find for every phenomenon means that permit determination of the phenomenon in a way which excludes the concurrent possibility of different determinations. According to Petzoldt this law permits the choosing, from among the infinite conditions that either determine a phenomenon or are interposed between it and its cause, of those conditions which effectively contribute to the determination of the phenomenon itself.

Pearson drew from his descriptive concept of law the consequence that scientific laws have only logical, not physical, necessity: "The theory of planetary motion is in itself as logically necessary as the theory of the circle; but in both cases the logic and necessity arise from the definition and axioms with which we mentally start, and do not exist in the sequence of sense-impressions which we hope that they will, at any rate, approximately describe. The necessity lies in the world of conceptions, and is only unconsciously and illogically transferred to the world of perceptions" (*The Grammar of Science,* 2nd ed., London, 1900, p. 134).

The empiriocritical branch of positivism is the immediate historical antecedent of the Vienna circle and of neopositivism in general. The sense impressions spoken of by Pearson and the sensations spoken of by Mach, Avenarius, and Petzoldt as neutral elements that constitute all the facts of the world, both physical and psychical, correspond exactly to the objects (*Gegenstände*) spoken of by Ludwig Wittgenstein in his *Tractatus Logico-philosophicus* as the constituents of atomic facts and to the elementary experiences (*Elementarerlebnisse*) spoken of by Rudolf Carnap in *Der logische Aufbau der Welt.* The restriction of necessity to the domain of logic, and the consequent reduction of natural laws to empirical propositions, is also a characteristic of the neopositivism of Wittgenstein, Carnap, and Hans Reichenbach. The critique of the principle of causality frequently recurs in neoempiricism reinforced by consideration of quantum mechanics (Philipp Frank, Reichenbach). The emphasis on prediction, important at all levels of science, is also a result of both empiriocriticism and logical positivism, as is the principle of the empirical verifiability of scientific propositions and the need to test and correct them constantly.

What empiriocriticism lacks is the stress on logic and language that is central to contemporary neopositivism. This stress developed out of work done in mathematical logic, especially by Bertrand Russell. Empiriocriticism lacks the concern with logic and the preoccupation with the nature of mathematics and of logical principles that is characteristic of contemporary neopositivism. The view that the proper business of philosophy is the clarification of concepts or the analysis of meanings derives largely from Russell, as does the preoccupation with problems about the status of logical and mathematical principles. The so-called linguistic theory about the nature of logical and mathematical principles, although subsequently endorsed by Russell, was developed by Wittgenstein. The use of the verifiability principle to demarcate meaningful from meaningless sentences and questions derives ultimately from David Hume's theory of impressions and ideas, but it is not to be found in any systematic form prior to the publications of the Vienna circle.

See also Logical Positivism.

Bibliography

There are no complete studies on positivism. For the individual philosophers, see J. Watson, *Comte, Mill and Spencer: An Outline of Philosophy* (New York: Macmillan, 1895); Leslie Stephen, *The English Utilitarians,* 3 vols. (London: Duckworth, 1900); D. G. Charlton, *Positivist Thought in France during the Second Empire, 1852–1870* (Oxford: Clarendon Press, 1959); and W. M. Simon, *European Positivism in the Nineteenth Century* (Ithaca, NY: Cornell University Press, 1963), which is limited to Comte's positivism and reactions to it.

The best comprehensive exposition of positivism as a philosophy and general world view is Richard von Mises, *Kleines Lehrbuch des Positivismus: Einführung in die empiristische Wissenschaftsauffassung* (Den Haag: van Stockum, 1939), translated under the author's supervision as *Positivism: A Study in Human Understanding* (Cambridge, MA: Harvard University Press, 1951). A briefer and more historical account, by another member of the logical positivist movement, is Hans Reichenbach, *The Rise of Scientific Philosophy* (Berkeley, CA: University of California Press, 1951). A less partisan overview (by a non-positivist) of positivist thought as a whole, emphasizing its unity while acknowledging its diverse ramifications and placing each episode in historical context, is Leszek Kolakowski, *The Alienation of Reason: A History of Positivist Thought*, translated by Norbert Guterman (New York, NY: Doubleday, 1968).

On the Enlightenment forerunners to nineteenth-century positivism, Charles Coulston Gillispie, *Science and Polity in France: The Revolutionary and Napoleonic Years* (Princeton, NJ: Princeton University Press, 2004) provides the political and social background, while the ideas are the focus in Keith Michael Baker's *Condorcet: From Natural Philosophy to*

Social Mathematics (Chicago: University of Chicago Press, 1975).

There are a number of valuable studies of the major nineteenth-century figures. On Comte, Robert C. Scharff, *Comte after Positivism* (Cambridge, U.K.: Cambridge University Press, 1995) and Juliette Grange, *La philosophie d'Auguste Comte: Science, politique, religion* (Paris: Presses Universersitaires de France, 1996) both focus primarily on philosophical ideas. On Mill, John Skorupski, *John Stuart Mill* (London: Routledge, 1989) also puts philosophical content in the foreground. On the evolutionary positivists, however, most studies have focused on social, political, and cultural aspects. David Weinstein, *Equal Freedom and Utility: Herbert Spencer's Liberal Utilitarianism* (Cambridge, U.K.: Cambridge University Press, 1998), for instance, focuses entirely on political ideas, and the Monist movement is situated in its social context by Gangolf Hübinger. "Die monistische Bewegung: Sozialingenieure und Kulturprediger," in *Kultur und Kulturwissenschaften um 1900 II: Idealismus und Positivismus*, G. Hübinger, R. von Bruch, and F.W. Graf, eds. (Stuttgart: Steiner, 1997, 246–259). Two of the three major figures of critical positivism have been the subjects of informative life-and-works studies: John T. Blackmore, *Ernst Mach: His Work, Life, and Influence* (Berkeley: University of California Press, 1972), and Theodore M. Porter, *Karl Pearson: The Scientific Life in a Statistical Age* (Princeton, NJ: Princeton University Press, 2004). The importance of Mach in particular for later positivist thought is brought out by Richard von Mises in "Ernst Mach and the Scientific Conception of the World," in *Unified Science: The Vienna Circle Monograph Series Originally Edited by Otto Neurath, Now in an English Edition*, edited by Brian McGuinness, translated by Hans Kaal (Dordrecht: Reidel, 1987: 166–190), and by Philipp Frank in "The Importance for our Times of Ernst Mach's Philosophy of Science," in his *Modern Science and its Philosophy* (Cambridge, MA: Harvard University Press, 1949: 61–78).

A great deal of scholarly effort has been devoted since the 1980s to the excavation and philosophical reconstruction of logical positivism, particularly the Vienna Circle. One important strand in this literature has regarded the neo-Kantian roots of logical positivism as more important than the positivist influence going back to Comte, Mill, and the western Enlightenment; exemplary for this trend is Michael Friedman, *Reconsidering Logical Positivism* (Cambridge, U.K.: Cambridge University Press, 1999). The continuity between the Enlightenment and logical positivism, in contrast, has been stressed by Thomas Uebel, e.g. "Enlightenment and the Vienna Circle's Scientific World-Conception," in *Philosophers on Education; Historical Perspectives*, edited by A. O. Rorty (London: Routledge, 1998, pp. 418–438), and *Vernunftkritik und Wissenschaft: Otto Neurath und der erste Wiener Kreis* (Vienna: Springer, 2000). The occlusion of the political, social, and educational dimensions in logical positivism after its main figures emigrated to North America is discussed by George Reisch, *How the Cold War Transformed Philosophy of Science: To the Icy Slopes of Logic* (Cambridge, U.K.: Cambridge University Press, 2005). A useful handbook with comprehensive bibliographies of the major figures and many peripheral ones is Friedrich Stadler, *The Vienna Circle: Studies in the Origins, Development, and Influence of Logical Empiricism*, translated by Camilla Nielsen et al (Vienna: Springer, 2001).

Nicola Abbagnano (1967)
Translated by Nino Langiulli
Bibliography updated by A. W. Carus (2005)

POSITIVISM, LEGAL

See *Legal Positivism*

POSITIVISM, LOGICAL

See *Logical Positivism*

POSNER, RICHARD

(1939–)

Richard Allen Posner, legal academic and federal court judge, was born in 1939 in New York. He was educated at Yale and Harvard Law School and has taught at the University of Chicago for many years. He was appointed to the federal appellate bench in 1981 and served as the chief judge of his court from 1993 to 2000. He is a leading advocate of the economic analysis of law and, by his own description, a legal pragmatist.

ECONOMICS OF LAW

Posner has argued that the various doctrines of the common law can best be explained as wealth maximizing. To say that a transaction or institution is wealth maximizing is to say that it creates more wealth than alternative possible transactions or institutions. Wealth, in this usage, is the value that goods have in the hands of their owners, and the value that a thing will have in the hands of a particular person is, qualifications aside, the amount that that person is willing to pay for it. Thus, the goal of wealth maximization is reached when goods are placed in the hands of those who would be willing to pay the most for them. An example of a wealth-maximizing rule, according to Posner, is the negligence rule in tort law: Under the rule of negligence, properly understood, injurers are liable for the losses they cause only when they could have taken precautions that would have prevented the accidents for less than the expected cost (that is, the cost discounted by the likelihood) of the accidents themselves. If prospective injurers take precautions when and only when it would be cost effective to do so—which the rule of negligence gives them an incentive to do—then

the cost of accidents overall will tend to be minimized and the wealth of society will tend to be maximized.

Similarly, but more controversially, Posner has offered an economic explanation of the criminal law. Its major function, according to Posner, is to prevent people from bypassing the market system of voluntary exchange. When goods are exchanged voluntarily, as in a sale, wealth is increased since parties necessarily value what they have received in an exchange more highly than what they traded for it. When the market is bypassed, as in theft, there is no guarantee that the stolen good is valued more highly by the thief than by its owner. Similarly, Posner has argued (thereby creating a great deal of controversy) that one of the things wrong with rape is that it bypasses the marriage and sex market so that wealth tends to be decreased. For Posner this is one of the virtues of wealth maximization over utilitarianism: Wealth maximization can explain why rape is always a crime whereas he believes that utilitarianism would have to condone rape if the enjoyment of the rapist were greater than the pain and unhappiness caused to the victim.

Even if the common law does promote the maximization of wealth, the question remains whether it *should*. Posner believes that wealth maximization is an ethically attractive guide not only for the common law but for social institutions generally. A system that maximized wealth overall would maximize everyone's chance for a higher income and thus would elicit *nearly* universal consent *ex ante*—though Posner's consenting parties would not have to do so in ignorance of their personal attributes. All persons would know of their own productive capacity—the extent to which they can benefit others—so they would know approximately how they would do under wealth-maximizing laws. It is only the unproductive who would not consent: They would be less well off under a wealth-maximizing system.

LEGAL PRAGMATISM

Posner believes that philosophical pragmatism is largely irrelevant to the law. By contrast, he believes that what he calls *everyday pragmatism* has a great deal to say. The everyday pragmatist—for example, the pragmatic judge—is an instrumentalist in law as in other things. Pragmatic judges are not bound by some conception of the law as an immutable body of rules but rather use their office to try to achieve reasonable resolutions to legal disputes. They reject moral, legal, and political theory (including constitutional theory) as guides to decision making. They are not bound by precedent, but neither are they bound to ignore it. Wise judges realize the virtues of following precedent—the value of certainty in law, the importance of the reliance interest, the wisdom that inheres in some of the common law—but they are free to ignore it when they can do more good by ignoring it. When pragmatic judges must look beyond the law to settle legal disputes, as they often must, they will find no help in academic moral theory. They must rely on common sense and economics and other sciences, as well as on values that are widely shared.

Although Posner's pragmatic judges are free to follow precedent or not, as they see fit, Posner counsels restraint in constitutional adjudication, placing himself among those judges and theorists that belong to what he calls the *outrage* school: The problem is that most judges are lacking the factual knowledge and expertise in social science that would justify them in striking down legislation. Hence, judges should only declare legislation unconstitutional when it stirs a strongly negative reaction in them. When in the future judges do in fact have a better grasp of social science and the factual underpinnings of the various areas of law, the need for law itself as we understand it will begin to disappear—the *supersession* thesis. Antitrust law and administrative law are two areas of American law that illustrate the thesis: "It is fair to say that at the beginning of its second century antitrust law has become a branch of applied economics" (Posner 1999, p. 229).

Posner calls himself a moral relativist. He believes that there is no rational road to agreement with those of fundamentally different moral beliefs and—what is now largely uncontroversial—that there is no way to reach certainty in moral matters. It follows, he believes, that we cannot call the actions of someone in another culture immoral unless we add *by our lights*, though he does not explain what the difference is between saying that something is immoral and saying that it is immoral by our lights.

See also Ethics and Economics; Philosophy of Law.

Bibliography

WORKS BY POSNER

"Utilitarianism, Economics, and Legal Theory." *Journal of Legal Studies* 8 (1979): 103.

"The Ethical and Political Basis of the Efficiency Norm in Common Law Adjudication." *Hofstra Law Review* 8 (1980): 487.

Problems of Jurisprudence. Cambridge, MA: Harvard University Press, 1990.

Overcoming Law. Cambridge, MA: Harvard University Press, 1995.

Problematics of Moral and Legal Theory. Cambridge, MA: Harvard University Press, 1999.

Frontiers of Legal Theory. Cambridge, MA: Harvard University Press, 2001.

Michael Louis Corrado (2005)

POSSIBILITY

The subject of possibility is a central topic in philosophy. It was frequently discussed in the history of philosophy, and it is actively debated by contemporary philosophers.

HISTORICAL DEVELOPMENTS

ARISTOTLE. The first comprehensive treatment of possibility occurs in the work of Aristotle. Aristotle's writing on this subject is difficult and confusing, but he seems to have held that the idea of possibility is derivative from that of necessity and negation, "It is possible that P" meaning "It is not necessary that not-P" (see *On Interpretation* 13.22b). Necessity of this basic kind is absolute necessity, and like absolute possibility it is applicable to sentences or propositions (*logoi*). According to his *Posterior Analytics* (4.21), a necessary proposition truly predicates something of a thing's essence; an example would be "A man is a rational being." A possible proposition, one that may be asserted to be such by a proposition containing the words "It is possible that … ," attributes an accident to a thing, an accident being a character that, because it is not excluded by a thing's essence, may or may not belong to it, as being seated may or may not belong to a man or woman. Because Aristotle held that what belongs to a thing's essence is given by a "real" definition, necessary propositions for him are either real definitions or logical consequences of such definitions.

Formal possibility. Although Aristotle's explicit remarks on absolute necessity relate to his theory of essences, he also uses a formal notion of necessity and, thus, of possibility, as when he argues that "Necessarily, every S is L" follows from "Necessarily, every M is L" and "Necessarily, every S is M." That the necessity and, correlatively, the possibility involved here is not the same as the real necessity and possibility just discussed is evident from the fact that the necessity of the conclusion "Every S is L" (and the impossibility of "Some S is not L") is justified wholly by the logical connection signified by "Every … is …" and by the sub-occurrences of "necessarily" in the modal syllogism. Important as this type of necessity and possibility obviously is to his theory of modal syllogisms, Aristotle does not seem to have reached the point of formulating its meaning explicitly. (See the discussion of Aristotle's modal syllogisms in *The Development of Logic* [1963] by William Kneale and Martha Kneale.) There can be little doubt, however, that this formal notion of necessity is rooted in the necessity of the first principles of all reasoning, such as the principle of contradiction. These principles cannot be demonstrated, Aristotle said, because all demonstration presupposes them (see *Posterior Analytics* 1.3.72b). They are known immediately and intuitively, and they cannot be consistently questioned.

Relative possibility. In the *Prior Analytics* (1.19.23a) Aristotle distinguishes absolute from relative necessity, and he implicitly makes a similar distinction for possibility in various passages of the *Organon* (for instance, in *De Sophisticis Elenchis* 4.166a22–166a30). Just as a proposition that does not state an absolute necessity may be considered necessary relative to certain other propositions (as a contingent statement constituting the conclusion of a valid deductive argument may be considered necessary relative to the truth of the premises), so a proposition like "Jones is walking" may be considered impossible relative to the proposition "Jones is sitting," and "Jones is sitting" may be considered possible relative to "Jones is not running." Although this distinction is intuitively clear, Aristotle does not explicitly say whether relative necessity and relative possibility are to be understood by reference to the sort of real absolute necessity and possibility discussed earlier or whether, as is likely, they are to be understood in relation to the formal notions that he sometimes uses but does not explicitly define.

Potentiality. Another sort of possibility discussed by Aristotle is potentiality, for certain possibilities can be said to exist as potentialities of concrete things. The possibility of a person's reading this or that may be understood in relation to a potentiality (we would say an ability) that the person has. For Aristotle a person who can read is a potential reader. Although the notion of potentiality is basic to Aristotle's metaphysics, he thought it could be understood only by analogy: "As a man who is building is to one who knows how to build, as waking is to sleeping, that which sees to that which has sight but has eyes shut, that which is shaped out of matter to its matter, the finished product to the raw material, so in general is actuality to potentiality" (*Metaphysics* 1048b).

MEGRIANS AND STOICS. A definition of possibility widely accepted in the Hellenistic period was that of Diodorus Cronus of Megara, who said, "The possible is that which either is or will be true" (Kneale and Kneale

1963, p. 117f). This identification of possibility with, in effect, present and future actuality was challenged by the Stoics (for example, by Chrysippus), who defined real possibility as "that which is not prevented by any thing from happening even if it does not happen" (Kneale and Kneale 1963, p. 123). Because the Stoics tended to be strict determinists, holding that whatever happens is necessitated by something else, they typically argued that our assessment of nonactuals as possibles could be based only on ignorance, for any conceivable occurrence that does not take place at some time or other is presumably prevented from taking place by the course of nature. Thus, their conception of real possibility developed into a conception of what is now known as epistemic possibility, or possibility as consistency with our knowledge.

Because the Stoics were especially interested in formal logic, they had another conception of possibility, however. According to this conception, necessary propositions (that is, necessary sentences) are those that are always true, such as the propositions of logic and mathematics. Possible propositions are those that are sometimes true. Since today's utterance of "A sea battle will occur tomorrow" is sometimes true according to the Stoics, then even though the course of nature may determine its truth with respect to tomorrow, the fact it states still belongs to the category of the possible (in the sense of sometimes true). It is perhaps worth adding that some commentators—for instance, Jaakko Hintikka (1959)—find this conception of possibility in Aristotle as well.

NEOPLATONISTS. The next distinctive conception of possibility, which turned out to be of great importance in medieval and modern philosophy, was worked out by the Neoplatonists—although it can be said to have its roots in Plato. According to this tradition, possibilities are not facts or states of affairs (that is, items properly expressed by sentences or propositions) but beings or essences that belong to Nous or Intelligence, the "first emanation of the One." Aristotle had spoken of potential beings inherent in various matters—for instance, a statue of Hermes existing potentially in a chunk of marble—but the idea of a possible being, which cannot be understood in relation to what substances or matter will become under certain conditions or when operated on in a certain way, is evidently new.

Admittedly, the idea may in a sense be traced back to Plato, for a possible being thus conceived is essentially something thinkable or intelligible, and Plato identified the intelligible with the world of Ideas or Forms. But Plato's Ideas were always general rather than specific, of humanity rather than of Socrates, and this means that the only possibilities, in this sense, that Plato could accommodate were kinds or species. Such Neoplatonists as Plotinus admitted Ideas of individual souls, and these, being nongeneral, may be regarded as the prototypes of the possible beings that occur in the theories of later philosophers such as Leibniz.

An extremely important aspect of the Neoplatonist treatment of possibles is that all possible beings were held to be actualized; possibility and actuality were regarded, that is, as precisely coextensive. The basic reason for this was that the infinite perfection or "goodness" of the One, which "overflows" into the emanation constituting the world of diverse actuality, requires that every possible being be brought into existence or actualized. This principle of plenitude among actualities was thought to be necessary according to the nature of things, because it is an essential feature of the One's perfection "to produce otherness" and "necessarily to do this in the maximum degree" (Lovejoy 1936, p. 66).

The Neoplatonic conception of possibles as Ideas in a divine mind that, owing to the perfection of that mind, are necessarily actualized was a recurrent and problematic theme in medieval philosophy. As A. O. Lovejoy pointed out in *The Great Chain of Being* (1936), medieval writers tended to conceive the love or goodness of the Christian God (in whose mind the Ideas were now said to exist) as an "immeasurable and inexhaustible energy," a love of which "the only beneficiaries … were not actual sentient creatures or already existing moral agents, but Platonic ideas, conceived figuratively as aspirants for the grace of actual existence" (p. 68).

ABELARD. Abelard, writing in the early twelfth century, was led to maintain that what can be is the same as what can be produced by God and that "it is intrinsically impossible for God to do (or make) or to leave undone (or unmade) anything other than the things that he actually does at some time do or omit to do; or to do anything in any other manner at any other time than that in which it actually is done" (Lovejoy 1936, p. 71).

AQUINAS. Because Abelard's view of possibility and actuality seemed not only to deny God's divine freedom but also, in implying that the created world was so good that it could not be better, to "make the creation equal to the Creator," it was regarded as heretical (Lovejoy 1936, p. 73). Accordingly, other Schoolmen, who like Aquinas agreed that "all things preexist in God by their types (*rationes*)," had to maintain that the creation involved a

selection among the ideas. In this view not all possibles are actual, and what is actual is not necessary: There are, that is, possible beings that God could have created but did not create, and he did not have to create the things that he did create. To square this claim with God's goodness, Aquinas found it necessary to invoke the Aristotelian distinction between absolute and relative necessity and possibility. Although it is absolutely possible for God, good as he is, to have created things other than what he did create, it may nevertheless be admitted that, relative to his choice, which was "becoming to" rather than necessary to his goodness, the existence of what is actual is necessary and could not be otherwise. That is, relative to this premise, it is impossible for anything to exist that does not sooner or later actually exist.

Even granting the distinction between absolute and relative possibility, it might be objected that Aquinas is still imposing a limit on God's freedom. If what actually exists is determined by God's selection from a class of possibilities, it would appear that God could not, in an absolute sense, have created anything not belonging to this class. In reply to this Aquinas maintained that what is absolutely impossible is self-contradictory and that what is self-contradictory is contrary to God's nature, repugnant to being, and therefore not an object at all. ("So it is better to say that what involves a contradiction cannot be done rather than God cannot do it," *Summa Theologica* 1.25.3–4.) In making this reply, Aquinas may seem to be introducing a formal notion of absolute possibility of the sort defended in more recent times. Yet, as with Aristotle, the category of possibility in question is grounded not in linguistic or purely logical considerations but wholly in intelligible essences ("intelligible forms"). In other words, the definitions relevant to ascertaining the consistency or intelligibility of a term or idea are "real" rather than nominal or analytical, which means that the possibility in question is the absolute kind espoused by Aristotle, not the formal or conceptual sort allowed by most modern philosophers.

HOBBES. In the modern period we find in Hobbes a view that not only contrasts vividly with the typical medieval one but which, confused as it is, is occasionally defended by philosophers of the twenty-first century. Hobbes's view contrasts with the medieval one because he held that conceivable beings are not necessarily possible beings. If a being is conceivable, the only conclusion Hobbes would draw is that words standing for it are not gibberish. To be possible, the necessary conditions for a thing's existence must be satisfied. Hobbes therefore contended that every possible being, event, or state of affairs is actual at some time or another: "If it shall never be produced, then those things will never concur which are requisite for the production of it" (*Elements of Philosophy* 10.4). Because for Hobbes whatever exists does so by virtue of necessary causes, we can call something possible (or contingent), as opposed to necessary, in his opinion only when we do not know the cause that will produce it. This view plainly goes back to that of the Stoics, for it implies that the only legitimate possibilities that are not also necessities are epistemic possibilities—that is, things or states of affairs whose existence is consistent with our knowledge at a given time.

DESCARTES. Descartes's approach to possibility is important mainly because it is essentially psychologistic: what is possible is what is clearly and distinctly conceivable. Descartes admitted that if the idea of a thing involves a contradiction, the thing is impossible, but he held that its impossibility is owing to the fact that contradictory ideas cannot be clearly and distinctly conceived. This latter criterion is basic for Descartes because some impossibilities do not, in his view, involve contradictions. As he saw it, there are *a priori* truths that are necessary and guaranteed to be true by the goodness of God but whose denials, which state impossibilities, are consistent. To know firsthand whether a given idea—for instance, the idea of a circular polygon—does represent a possibility, one must therefore be able to form a clear and distinct idea of it. If one is able to form such an idea, one has God's assurance that it represents a real possibility, the sort of thing that God could actualize if he chose to do so.

SPINOZA. According to Spinoza, "A thing is said to be impossible either because the essence of the thing itself or its definition involves a contradiction, or because no external cause exists determinate to the production of such a thing" (Spinoza, *Ethics*, 1, prop. 33, note I). Because Spinoza in effect adopted the Neoplatonic principle of plenitude, he held that if the idea of a thing does not involve a contradiction, it must be actual, for all self-consistent beings are determined to exist, and necessarily exist, by the very nature of reality, which he calls "God":

> [Accordingly, a] thing cannot be called contingent unless with reference to a deficiency of our knowledge. For if [and here Spinoza introduces the notion of epistemic possibility] we do not know that the essence of a thing involves a contradiction, or if we actually know it involves no contradiction, and nevertheless can affirm nothing with certainty about its existence because the order of causes is concealed from us, that thing

> can never *appear to us* as necessary or impossible, and therefore we call it either contingent or possible (Spinoza, *Ethics*, 1, prop. 33, note I).

LEIBNIZ. To general readers, Leibniz is best known for his metaphysical optimism, the doctrine that this is the best of all possible worlds. He conceived of a possible world as a maximal collection of absolutely possible beings each of which is "compossible" with the other beings contained in that world; the totality is maximal in the sense that it contains everything compossible with its contents. Two things are compossible, Leibniz said, when it is absolutely possible for them to exist together; and something is absolutely possible, for him, when God's conception of it is free from contradiction. Because Leibniz held that God's concept of a thing includes all facts about it, including such apparently accidental facts as that it once crossed a certain river in Peru or that it once was bitten by a dog called "Rover," he concluded that if a thing is absolutely possible, it is so only relative to its place in a possible world, one including certain possible rivers, perhaps, and certain possible dogs. A possible being is strictly a being, therefore, whose existence is compossible with the members of a possible world. This conception comprehends the less restrictive idea, common in recent metaphysics, that a possible thing or state of affairs is one that "exists at," or belongs to, some possible world.

Like Aristotle, Leibniz drew a distinction between absolute and relative possibility. (Leibniz used the term "hypothetical" here instead of "relative," but his distinction was the same as Aristotle's.) Because God created the best of all possible worlds, any existing thing that is not, like God, an absolutely necessary being depends on God's creative choice. A thing that is absolutely possible but dependent this way on God's creative choice is hypothetically necessary: its nonexistence is hypothetically impossible, ruled out by the choice God actually made. Everything that has occurred, will occur, or is now occurring is necessary in this hypothetical sense, according to Leibniz. But hypothetical necessity is not the same as absolute necessity, he insisted; Diodorus Cronus (see above) erred in not recognizing this important fact. All human behavior is hypothetically necessary, but it is not thereby inevitable in an absolute sense. This is why one can rightly maintain that free choice remains possible for human beings. A free action, for Leibniz, is one that results from a "rationally spontaneous" choice; its originating principles lie within the agent. Free actions spring from motives and other causes, but these "incline without necessitating," he said; absolutely necessity is not imposed upon them (see Mates 1986, p. 119)

HUME. The British empiricists, typically rejecting the claims of conceptualism as defended by most epistemic rationalists, seemed to embrace more fully the idea that possibility is a matter of logical consistency. In remarking that, "The contrary of every matter of fact is still possible, because it can never imply a contradiction," Hume appears firmly committed to a view of logical possibility. But in adding to the quoted sentence, "And is conceived by the mind with the same facility and distinctness," Hume discloses his tacit commitment to a psychologistic conception of possibility (what is possible is what is conceivable), which was held by Descartes and is often assumed even today. (See Hume's *Enquiry concerning Human Understanding*, 4.1.)

KANT. In Kant there is not only a clear identification of *a priori* possibility but an explicit distinction between logical and physical (or nomological) possibility. For philosophers like Spinoza, who identified the logical with the real order, there was plainly no sense in this distinction, and there was little place for it in the philosophies of the Greek and medieval thinkers. It is, however, essential to the contemporary outlook. Kant expresses the distinction a bit clumsily thus:

> A concept is always possible [he means "represents a possibility"] if it is not self-contradictory. This is the logical criterion of possibility, and through it objects are distinguished from the *nihil negativum*. But it may nonetheless be an empty concept, unless the objective reality of the synthesis through which the concept is generated has been specifically proved; and such proof … rests on principles of possible experience, and not on the principle of analysis (the law of contradiction). This is a warning against arguing directly from the logical possibility of concepts to the real possibility of things. (*Critique of Pure Reason*, A597/B625, note)
>
> [Thus, the possibility of such things as] a special fundamental power of our mind to *intuit* the future (not merely, say, to deduce it), or, finally, a faculty of our mind to stand in a community of thoughts with other men (no matter how distant they may be)—these are concepts the possibility of which is entirely groundless, because it cannot be grounded in experience and its known laws, and without this it is an arbitrary combination of thoughts that, although it contains no contradiction, still can make no claim to objective reality, thus to the possibility of the

> sort of object that one would here think. (*Critique of Pure Reason* A223/B270)

To ascertain that such things are empirically (as opposed to merely logically) possible, we must ascertain whether the nature of things so described agree with the formal conditions of actual experience.

CONTEMPORARY DEVELOPMENTS

Not all the conceptions of possibility discussed in the previous section on the history of philosophy are equally acceptable to contemporary philosophers, and new conceptions are topics of current debate. Generally speaking, possibility is now discussed in relation to two principal subjects: basic metaphysics, which takes some kind of absolute possibility as fundamental, and the compatibility of freedom and determinism, which introduces possibilities of other kinds. The conceptions of possibility now considered tenable by most philosophers (there is disagreement on this) can be identified by reference to these two subjects.

BASIC METAPHYSICS. Until the 1970s, most analytic philosophers described absolute *a priori* possibilities as "logical possibilities" and identified them, as Leibniz did, by reference to logical consistency: An absolute possibility is something that can be exhaustively described without contradiction. In logic a contradiction has the form of "p and not-p" however; and this syntactical structure is not explicit in many statements that fail to express genuine possibilities: it is not present, for instance, in "Some bachelors are married" or "Mary is both taller and shorter than Sally." To expose the contradictions implicit in these statements one must make use of definitions and conceptual truths such as "For any x and y, if x is taller than y then x is not shorter than y." Conceptual truths and statements true by definition were called "analytic" truths, and the full range of absolute possibilities was generally conceded to be identifiable only by reference to them. An absolute possibility was then said to be expressed by a statement that is consistent with all relevant analytic truths. According to this conception, a statement that is not so consistent would fail to express a genuine possibility.

This way of identifying absolute possibilities was undermined by Saul Kripke in lectures given in 1970 and subsequently published under the title *Naming and Necessity* (1980). Kripke's criticism featured two striking examples. The first involved what most philosophers would call an analytic truth pertinent to the standard meter located in Paris. The truth is that the rod is one meter long. Although this truth is a consequence of an arbitrarily chosen standard specifying what is to count as a meter in length, and thus would be acknowledged to be analytic by most philosophers, it is not necessary because the rod in question does not of necessity possess its current length: it could have a different one. This latter possibility is genuine, but it is identified by reflection on how the rod might change, what might happen if, say, it were heated—not by the consistency of "The rod is not a standard meter long" with the truth that the length it now has equals one meter. The analytic consistency conception of absolute possibility does not give the right result in this kind of case.

Kripke's second example concerned the identity of Hesperus and Phosphorus, the morning star and the evening star. The statement that Hesperus = Phosphorus is not an analytic truth; it was discovered to be true by empirical investigation. The two "stars" turned out to be a single planet, Venus, seen in the sky at different times and presumed to be different. The fact that the statement is not an analytic truth does not prove that it is not necessary, however. It is in fact necessary because it concerns a single planet, and that planet, like everything else, is necessarily self-identical. Because the identity of Hesperus and Phosphorous had to be discovered empirically, the necessity of their identity had to be inferred from the fact of their identity. If "a" and "b" are used "rigidly," as Kripke said, to pick out the same objects in actual as well as counterfactual situations, then the following principle provides a basis for the inference: If $a = b$ then it is necessary that $a = b$. Because the necessity of "a" being "b" is equivalent to the impossibility of *a* not being *b*, a certain possibility is ruled out by our empirical investigation: We learn that it is not possible for *a* to differ from *b*. This impossibility is not known *a priori* by the discovery that some statement (or proposition) is self-contradictory or analytically inconsistent.

In developing his metaphysical views, Kripke drew a distinction between *de dicto* and *de re* necessity and possibility. A *de dicto* possibility is in effect the possible truth of some proposition; it is expressed in words by a sentence beginning "It is possible that …" A *de re* possibility, by contrast, is attached to a particular thing, such as a person or chair. We are concerned with such possibilities, Kripke said, when we wonder whether a certain person might have done this or that in some counterfactual situation. Kripke spoke of "contingent properties" in describing such possibilities. A property is contingent for a thing when the thing may or may not possess it in some situation or other. Such a property contrasts with a necessary or "essential" one, this being a property that a thing pos-

sesses in every situation, actual and counterfactual, in which it may exist. *De re* possibilities correspond to Aristotle's potentialities; *de re* necessities correspond to his "actualities," or the components, as he conceived them, of a thing's "form" or essence.

Kripke emphasized that the notions of necessity and possibility he discussed belong to metaphysics, not epistemology, and he sometimes spoke of them as metaphysical necessity and metaphysical possibility (see Kripke 1980, p.19). In commenting on the formal semantics he invented for the logic of statements affirming such necessities and possibilities, Kripke used Leibniz's notion of a possible world. A statement, S, is necessary with respect to the actual world, Kripke said, just when S is true with respect to all possible worlds—more exactly, all worlds that are possible relative to the actual world. S is possible with respect to the actual world (it is, for members of this world, possible that S) just when S is true with respect to some possible world—with some world that is possible relative to the actual world. Kripke spoke of worlds possible "relative to" the actual world because different assumptions may be made about this relativity, and these different assumptions are associated with modal principles that are characteristic of different systems of modal logic (see Kripke 1971).

Although Kripke informally used the notion of a possible world in describing the truth-condition for statements affirming metaphysical possibilities and necessities, he did not believe that such statements were understandable only in relation to possible worlds or that the framework of possible worlds provides a reductive analysis of modal discourse. In fact, to avoid philosophical confusions and anxieties regarding possible worlds, he recommended that "possible state (or history) of the world" or "counterfactual situation" might provide a preferable terminology (see Kripke 1980, pp. 18f). As far as modal knowledge is concerned, he seems to believe that intuitiveness (or perhaps intuitive obviousness) is basic. As he put it in *Naming and Necessity* (1980), "Some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking" (p. 42). In speaking of intuitive content this way Kripke appears to favor an epistemically rationalist (or Cartesian) view of modal knowledge, but he did not discuss the matter in greater detail, and it remains uncertain what the details of his view actually are.

An influential writer about possibility who appears to regard possible worlds and the possible individuals that compose them as basic realities is David Lewis (1986). Lewis believes that all possible worlds actually exist but that only one world, at least from our perspective, is actual: our world. Like Leibniz, Lewis holds that the possible individuals of other possible worlds do not include the individuals of our world; in fact, he thinks the individuals of different worlds cannot be shared. When we consider a counterfactual possibility involving a person belonging to our world—George W. Bush, say—the possibility is grounded in (or actually involves) a counterpart to that person, a being relevantly similar to him, belonging to another possible world. Lewis accepts this counterpart theory because he thinks a given thing cannot have incompatible features. If a thing belonged to two different worlds, the worlds would overlap in it, and this could happen only if the thing's nonrelational features were exactly the same in both worlds: A thing cannot possibly differ from itself. Lewis ably defends his position against a multitude of objections in *On the Plurality of Worlds* (1986), and he also provides a non-Cartesian account of how he thinks we can have genuine knowledge of worlds that, although existing, are possible rather than actual.

FREEDOM AND DETERMINISM. The conceptions of possibility relevant to this topic are brought to mind by the question, "If the world is a deterministic system, is it possible for human beings to do anything that they do not actually do?" Not every responsible philosopher agrees that this question requires an affirmative answer if human beings can reasonably be considered capable of acting of their own free will, but the question is commonly asked and different kinds of possibility are mentioned in answering it (see Austin 1961).

Possibility as ability. This kind of possibility corresponds to Aristotle's potentiality. We often have this sort of possibility in mind when we wonder what a person is capable of doing, and what he or she could do in specific circumstances. Can Tom do fifty push-ups? Can he do that many after a big meal? What is relevantly possible here? The basic idea pertinent in answering these questions is that of an ability or capacity. To have an ability or capacity a person must be capable of doing something; and to be thus capable is to be such that if conditions are of the right kind, appropriate behavior will occur. In discussions involving human freedom the abilities under consideration are voluntary: they are abilities that a person can manifest "at will." If Sally has the ability to swim, then she will normally succeed in swimming if she is

immersed in water and attempts to swim. The qualification "normal" is important here because a failure to swim would not be evidence of an inability to swim if one's legs were encased in concrete. Success is required only in "favorable" conditions.

Sometimes we are concerned with what a person can do in special conditions, which may be far from what are considered favorable. Can Tom swim in a rough sea? Can Betty solve an algebra problem when her roommate's stereo is pounding in her ears, when she is seething with irritable frustration? The relevant test here is success under the specified conditions. In a particular case the test to be satisfied is specified by a conditional statement in the subjunctive mood: If conditions C were to obtain and the subject attempted to exercise the relevant capacity, the subject would succeed in the attempt.

The most important recent work on the logic of subjunctive conditionals is contained in Davis Lewis's book *Counterfactuals* (1973). Lewis gives the truth-conditions for these statements by reference to possible worlds. A statement of the form "If it were the case that p, it would be the case that q" is true, according to Lewis, just when q is true at the possible worlds that satisfy p and are otherwise most similar to the actual world. (There may or may not be a single most similar p-world.) Thus, to decide whether Tom could do fifty push-ups after a certain meal, one in effect has to decide, Lewis says, whether a possible world in which he (or his "counterpart") does fifty push-ups after such a meal would be minimally different from the actual world, or whether it would require him to have undergone a course of training, say, that he did not experience in the actual world. Because the negation, according to Lewis, of the conditional "If A were to happen, B would happen" is "If A were to happen, B might not happen," one can use his theory to identify another kind of possibility, which might be called a "contingent" possibility. Suppose it is false both that if A were to happen, B would happen and that A does happen. Under these circumstances it could be said that B's not happening is a contingent possibility.

Relative or hypothetical possibility. A conception of possibility ultimately vital to the subject of human freedom is that of what is possible given the laws of nature and the occurrence of remotely prior causal factors. Aristotle and Leibniz both acknowledged this conception, but the idea that it represents a genuine kind of possibility is often questioned by contemporary philosophers. Benson Mates (1986), in his commentary on Leibniz, says that the distinction between absolute and hypothetical necessity (and therefore between absolute and hypothetical possibility) seems to originate in a confusion of "Necessarily, if P then Q" and "If P, then necessarily Q." There is no doubt that this confusion is often made, but it was certainly not made by Leibniz, who explicitly distinguished statements of these kinds and accused Diodorus Cronos, who denied that any possibility could fail to be a necessity, of confusing hypothetical necessity with absolute necessity (see Mates 1986, pp. 117ff).

Peter van Inwagen (1983, 2000), wishing to avoid the confusion Mates mentioned, introduced a new modal operator in formulating an argument against the compatibility of freedom and determinism. The formula "Np" containing his special operator "N" is to be understood as meaning "p [is true] and no one has or ever had any choice as to whether p." If "O" is a modal operator representing a kind of necessity, there is no doubt that an argument having "Op" and "O(if p then q)" as premises and "Oq" as a conclusion is valid. Accordingly, van Inwagen formulates a corresponding argument featuring his operator "N" and argues that it is valid. The remotely prior causes C occurred and no one now has or ever had any choice about their occurrence; hence "N(C)." Similarly, the laws of nature hold true and no one has or ever had any choice about this fact. The laws also imply that if C then B, where B is a representative item of behavior in a deterministic world. Because this implication is necessary and something no one has or ever had any choice about, van Inwagen concludes that N(B)—that B occurs and no one has or ever had any choice about it: an alternative to B is out of the question.

Van Inwagen's argument has been seriously criticized since his book was published in 1983, and he has gone on to sketch a new argument to express his sense of the "sheer inescapablity" of determined behavior (see van Inwagen 2000). But it is obvious that the sheer inescapability of B is tantamount to the fact that it is relatively (or hypothetically) necessary in Leibniz's sense, and van Inwagen's conclusion "N(B)" amounts to nothing more than an assertion that B is a logical consequence of natural laws and the occurrence of initial conditions (or previous causes) that cannot be altered when B occurs. Van Inwagen's worry about human freedom depends, in effect, on the relative impossibility of behavior that does not occur. So this sense of possibility is vital to the freedom-determinism issue, at least as philosophers such as van Inwagen understand it.

See also Analytic and Synthetic Statements; Modal Logic.

Bibliography

POSSIBILITY IN THE HISTORY OF PHILOSOPHY

Aquinas, Thomas. *Basic Writings of Saint Thomas Aquinas*, edited by Anton C. Pegis. 2 vols. New York: Random House, 1945.

Aristotle. *The Complete Works of Aristotle: The Revised Oxford Translation*, edited by Jonathan Barnes. 2 vols. Princeton, NJ: Princeton University Press, 1984.

Copleston, Frederick. *A History of Philosophy*. 8 vols. London. 1959–1965.

Descartes, René. *The Philosophical Writings of Descartes*. Translated and edited by John Cottingham, Robert Stoothoff, and Dugald Murdoch. 2 vols. Cambridge, U.K.: Cambridge University Press, 1984.

Hintikka, Jaakko. "Necessity, Universality, and Time in Aristotle." *Adjatus* 20 (1959): 65–90.

Hobbes, Thomas. *Hobbes: Selections*, edited by Frederick J. E. Woodbridge. New York: Scribners, 1930.

Hume, David. *Enquiry concerning Human Understanding*. 3rd ed., edited by L. A. Selby-Bigge. Revised by P. H. Nidditch. Oxford: Clarendon Press, 1975.

Kant, Immanuel. *Critique of Pure Reason*. Translated and edited by Paul Guyer and Allen Wood. Cambridge, U.K.: Cambridge University Press, 1998.

Kneale, William, and Martha Kneale. *The Development of Logic*. Oxford: Clarendon Press, 1963.

Leibniz, Gottfried von. *Leibniz: Selections*, edited by Philip P. Wiener. New York: Scribners, 1951.

Lovejoy, A. O. *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, MA: Harvard University Press, 1936.

Mates, Benson. *The Philosophy of Leibniz: Metaphysics and Language*. New York: Oxford University Press, 1986.

Spinoza, Baruch. *Philosophy of Benedict de Spinoza*. Translated by R. H. M. Elwes. New York, 1955.

POSSIBILITY IN CONTEMPORARY PHILOSOPHY

Austin, J. L. "Ifs and Cans." In his *Philosophical Papers*. Oxford: Clarendon Press, 1961.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Kripke, Saul. "Semantical Considerations on Modal Logic." In *Reference and Modality*, edited by Leonard Linsky. London: Oxford University Press, 1971.

Lewis, David. *Counterfactuals*. Oxford: Blackwell, 1973.

Lewis, David. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.

van Inwagen, Peter. *An Essay on Free Will*. Oxford: Clarendon Press, 1983.

van Inwagen, Peter. "Free Will Remains a Mystery." In *Philosophical Perspectives*. Vol. 14: *Action and Freedom, 2000*, edited by James E. Tomberlin. Boston: Blackwell, 2000.

Bruce A. Aune (1967, 2005)

POSSIBLE WORLDS

See *Modality, Philosophy and Metaphysics of*

POSTCOLONIALISM

Not unlike the Renaissance, the Enlightenment, and postmodernism, postcolonialism refers not only to a temporal marker that signals a shift in mentalities and metaphilosophical questioning but also to a decolonizing movement enabled by new material conditions and to a theoretical and philosophical methodology. As a temporal marker postcolonialism is caught in a series of paradoxes. On the one hand, postcolonialism signals the alleged *end* of colonialism and the beginning of a new historical period. On the other hand, at the center of postcolonialism is the exploration of what postcolonial theorists have called the *postcolonial present*, namely, the enduring legacy of colonialism in contemporary times. Still, one of the most basic goals of postcolonialism is to foreground the movements of decolonization that began as early as the end of World War II, peaked during the 1950s and 1960s, and have lasted into the twenty-first century. For this reason many postcolonial theorists argue that postcolonialism is less an "ism" that describes an already past movement, but is more a series of philosophical issues that emerge from the ongoing process of decolonization in the midst of the global hegemony of Europe and the United States.

Undoubtedly, postcolonialism also refers to all the movements of decolonization that emerged during the 1950s, movements that predominantly took the form of so-called Third World nationalism. These movements of national liberation and anti-European imperialism and decolonization spread throughout the so-called Third World, a noun that conceals the specific Cold War context of many of these anticolonial struggles. *Third World* makes reference to all those recently created nations that were part neither of the developed, capitalist, industrialized, democratic *First World* nor the developing, industrializing, and socialist *Second World*. Critical theorist Robert J. C. Young (1950–) has for this reason argued that instead of referring to Third World postcolonialism, we should make reference to *Tricontinentalism*, by which he means, the deliberate and explicit joining of former colonial societies in Latin America, Africa, and Asia in anticolonial struggles. What postcolonialism as the collective name for a series of movements seeks to foreground is precisely the engagement with what is called by some postcolonial theorists the postcolonial condition, or postcoloniality. Latin American sociologist and critical theorist Anibal Quijano has called this condition the *postcoloniality of power*, a felicitous expression that expresses what Homi K. Bhabha (1949–) has called the *ongoing colonial present*.

As a theoretical and philosophical methodology, postcolonial theory is no less heterogeneous and at times internally contradictory than the Renaissance and the Enlightenment were. Postcolonial theory finds many of its philosophical sources in the discourse of, to use Paul Ricoeur's apropos phrasing, *the hermeneutics of suspicion*: Marxism, psychoanalysis, deconstruction, semiotics, structuralism, and postmodernism. More concretely, most of postcolonial thinking takes place through demystifying readings of canonical figures in Western philosophy. Such demystification is exemplified in the works of Louis Althusser (1918–1990), Jacques Derrida, Michel Foucault, Sigmund Freud, Jacques Lacan, Karl Marx, Friedrich Nietzsche, Jean-Paul Sartre. As a methodology, postcolonialism submits both the production and effects of all cultural artifacts, whether they be novels, philosophical texts, or sociological treatises, to a type of X-ray that shows the ways in which these texts and their effects are caught in the dialectical tension between colonialism, imposed and internalized, and anticolonialism, both internal to the West, and from without, from the colony, the liberated postcolonial nations, and emergent social movements.

This type of double reading that traces the effects of colonialism on colonial consciousness and culture, and that unearths and names the voice and gaze of the colonial other, has been amply developed by what has been called *postcolonial criticism*. With this term some critics seek to differentiate between the kinds of work that literary criticism performs from that which theory or philosophy produces. Yet, the attempt to differentiate between postcolonial criticism and theory reproduces one of the most contested *disciplinary* divisions that postcolonialism, as a methodology of analysis, continuously aims to challenge. As the works of Bhabha and Gayatri Chakravorty Spivak (1942–) illustrate and explore, literary criticism cannot be separated from and made to dispense with philosophical analysis, and the latter cannot dispense with literature nor be made to speak in a language purified of rhetoric, simile, metaphor, and the thick historicity of its diction. Furthermore, postcolonial theorists can neither negate nor neglect the ways in which disciplinary divides have been utilized to silence and deauthorize other forms of questioning—in what postcolonial theorists call the production of knowledge—precisely because of postcolonial theory's own hybrid and interdisciplinary sources.

Postcolonialism can be said to be a phenomenology of the social world that analyses in tandem the mutually conditioning effects of the objective on the subjective and vice versa. Social existence conditions the ways in which subjects are able to live and experience their subjectivities, and such subjectivities in turn, whether subjugated or insurrected, transform the social world. Postcolonialism is therefore also simultaneously a type of critical epistemology and historical ontology that studies the sources and effects of modes of representation and the ways in which social being is historically conditioned. As Spivak has put it, appropriating and displacing the phenomenological hermeneutics of Martin Heidegger, colonialism has *worlded*—that is to say, woven a thick web of material relationships that made possible meanings and the subjects that are mediated by them—the worlds of both the colonizer and the colonized. What makes postcolonialism different from other forms of phenomenology, ontology, and epistemology is that it has deliberately sought to disclose the world worlded by colonialism from the standpoint of the subaltern. By the term *subaltern* postcolonial critics mean those agents who have been expropriated, exploited, marginalized, racialized, bestialized, and rendered part of the fauna of continents empty of people and subjects. Every social agent and epistemic subject occupies a location, whether this location be literally geographical or figuratively political, epistemological, racial, or gendered. Edward W. Said (1935–2003) has called the analysis of this localization of all agents the *geographical inquiry into historical existence.*

Postcolonial theorists argue that to analyze the world from the perspective of the colonizer—the sovereign European political subject ensconced on the pedestal of racial privileged—would distort at best and conceal at worst the ways in which the colonized, the subaltern of colonial cultures, have been disempowered, rendered invisible and silent, reduced to a mere tabula rasa for the evangelizing, civilizing, and commercializing mission of Europe. The postcolonial critique of Western domination is simultaneously a critique of the imposition of a global economic system of structural inequality, or what is also called the globalization of capitalism through colonialism and imperialism. For this reason postcolonial theory shares many important insights and methods with standpoint feminist epistemological critique. All social location, as both of these positions argue, is mediated by representations: cultural, gendered, racial, religious. Postcolonial critique, as a form of Marxism, thus also aims to unmask the fetishizing and alienating effects of the systems of cultural representation imposed by European colonialism.

Postcolonialism, therefore, maintains that since no cultural or personal identity exists outside representation,

and all representation is mediated by the history of its production, imposition, or rejection, all identities are thus contaminated by instability, hybridity, or creolization. A postcolonial corollary to the hybridity of all identity is that there is no subjectivity and agency that is not affected by power. All subjectivity and agency, argue postcolonial theorists, are forms of power. The postcolonial analysis, however, maintains that some forms of power are genocidal, subjugating, and narcissistic while others are enabling, benign, and indispensable. Power, in this analysis, is neither a stable substance nor a force that emanates from a center but a configuration of relationships that condition modes of social interaction. For postcolonial theorists, however, the uses and abuses of power are discerned from the standpoint of its effects on the subaltern in history and society.

At the center of postcolonial theory, notwithstanding its variegated sources and heterogeneous forms of articulation, is a series of epistemological innovations. Whether one studies the work of Bhabha, Frantz Fanon (1925–1961), Said, or Spivak, to mention just some of the canonical figures in postcolonial theory, we encounter an in-depth and unmitigated analysis of what has been called variously the space of enunciation, the discursive fields, or the structure of attitude and reference. Postmodern theorists mean by these terms that all epistemic locations, statements, and responses of affect are either allowed or disallowed by certain rules, syntax, or injunctions. To claim epistemic authority, make statements, and submit to feelings is to be interpellated by the syntax of a discursive matrix that already also anticipates their assent, response, or evocation.

Some postcolonial theorists have focused their attention on the structures of attitude and reference that condition how subjects and agents are made to know, speak, and feel *from* a location of privilege and plenipotentiary sovereignty *about* other subjects and agents who are located somewhere else in history and space. Said's classic work *Orientalism: Western Representations of the Orient* (1979) documented and analyzed the ways in which orientalism, the collective name for a group of disciplines that studied the Orient, operated as a power-knowledge apparatus that interpellated European agents to adopt imperial affective, epistemological, and enunciative spaces and comportments. Other postcolonial critics have focused on the knowing, speaking, and feeling *to* of all colonial discourse and the ways in which their reifying, objectivifying, and alienating effects are both unsustainable and contested by the other of their addressee. In Fanon's work, for instance, we discover one of the most elaborate phenomenologies of oppression and liberation as well as a psychoanalytical analysis of the devastating effects of the powers of torture on both colonizer/torturer, and colonized/tortured. Yet other postcolonial theorists have focused on the *how* and by *what* means the mater-slave relationship between colonizer and *subaltern* have been mediated in such a way that neither the master nor the slave are entirely inured to each other's power of conquest or resistance. Spivak's work is without a doubt the most sophisticated, extensive, and sustained engagement with this dialectic of complicit and resisted knowledge production and insurrected agency.

Not unlike how Immanuel Kant illustrated his transcendental method by way of antinomies, postcolonial critics have developed a type of critical philosophy that proceeds also by way of the disclosure of a series of antinomies at the heart of contemporary Western thinking: universalism versus European exceptionalism; rationalism versus racial supremacy; humanism versus racial genocide; technophilia versus Romantic idolatry of the primitive; historicism versus teleological theodicy. As a critical methodology that inherits the discourse of what has been called a second Enlightenment, namely, the discourse of suspicion (Marx, Freud, Nietzsche), postcolonialism can be said to constitute a third Enlightenment, one that awakens the postcolonial world to the enduring legacies of five centuries of colonialism, imperialism, and now, globalization.

Postcolonialism is neither anti-Western, obdurately rejecting all European thinking, nor Third-Worldist, naively celebrating all that is produced and thought by the subaltern. Postcolonialism is a type of thinking that aims to situate us beyond the epistemological, ontological, and phenomenological Manichaeisms that have informed colonialism and postcolonial nationalism. Postcolonialism urges us to think beyond the either/or, for/against and in the proper space of the hybrid of the neither/nor, and/but, not with/but not without. For postcolonial thinkers, the philosophical inheritance of the West, of Europe, is at stake, not solely because it bears the traces of its complicity with colonialism, but because it is also the archive of resistance to that colonialism.

See also Deconstruction; Derrida, Jacques; Enlightenment; Epistemology; Foucault, Michel; Freud, Sigmund; Heidegger, Martin; Humanism; Kant, Immanuel; Lacan, Jacques; Mani and Manichaeism; Marxist Philosophy; Nietzsche, Friedrich; Ontology; Phenomenology; Postmodernism; Renaissance; Ricoeur, Paul; Romanticism; Psychoanalysis; Sartre,

Jean-Paul; Structuralism and Post-structuralism; Teleology.

Bibliography

Appiah, Anthony Kwame. *In My Father's House: Africa in the Philosophy of Culture*. New York: Oxford University Press. 1992.

Bhabha, Homi K. *The Location of Culture*. London: Routledge, 1994.

Brydon, Diana, ed. *Postcolonialism: Critical Concepts in Literary and Cultural Studies*. 5 vols. London: Routledge, 2000.

Cabral, Amilcar. *Return to the Source: Selected Speeches*. New York: Monthly Review Press, 1973.

Del Sarto, Ana, Alicia Ríos, and Abril Trigo, eds. *The Latin American Cultural Studies Reader*. Durham, NC: Duke University Press, 2004.

Fanon, Frantz. *Black Skins, White Masks*. Translated by Charles Lam Markmann. New York: Grove Press, 1967.

Fanon, Frantz. *A Dying Colonialism*. Translated by Haakon Chevalier. New York: Grove Press, 1967.

Fanon, Frantz. *The Wretched of the Earth*. Translated by Richard Philcox. New York: Grove Press, 2004.

Moore-Gilbert, Bart. *Postcolonial Theory: Contexts, Practices, Politics*. London: Verso, 1997.

Said, Edward W. *Culture and Imperialism*. New York: Alfred A. Knopf. 1993.

Said, Edward W. *Orientalism: Western Representations of the Orient*. New York: Vintage, 1979.

Said, Edward W. *Reflections on Exile and Other Essays*. Cambridge, MA: Harvard University Press, 2000.

Spivak, Gayatri Chakravorty. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Cambridge, MA: Harvard University Press, 1999.

Spivak, Gayatri Chakravorty. *In Other Worlds: Essays in Cultural Politics*. New York: Routledge, 1988.

Spivak, Gayatri Chakravorty. *Outside in the Teaching Machine*. New York: Routledge, 1993.

Young, Robert J. C. *Postcolonialism: An Historical Introduction*. Malden, MA: Blackwell, 2001.

Eduardo Mendieta (2005)

POSTMODERNISM

The term "postmodernism" first emerged in the 1950s to describe new architectural and literary movements that opposed commonly accepted canons regarding the unity and coherence of narratives and artistic styles. Sociologists, meanwhile, have used "postmodernism" to indicate discordant trends such as the parallel growth in cosmopolitan globalization and parochial traditionalism. The term has also been appropriated by mainly French and German philosophers to designate a criticism of reason, regarded as a universal and certain foundation for knowledge and morality, and of modern culture, understood as a progressive unfolding of knowledge and morality. An examination of the works of these philosophers shows that many of the postmodern themes regarding the fragmentation (or deconstruction) of the rational subject and its object can be explained from the standpoint of conceptual tensions implicit within post-Kantian philosophy, which remains the main target of postmodern criticism.

A true son of the Enlightenment, Immanuel Kant defended reason as a universal faculty whose untrammeled exercise irresistibly leads to questioning all dogma and all authority, and from there leads to the complete emancipation of all individuals from the fetters of tradition. Faith in reason as it is deployed in science and morality fuels faith in the interminable progress of humanity. However, as postmodernists like Michel Foucault point out, the very reason that develops modern culture disintegrates under its own self-critical gaze, thereby issuing in oscillating and often discordant trends between absolutism and nihilism, totalitarianism and anarchism, humanism and multiculturalism, and universalism and parochialism. The end of rational idealism in turn spells the end of the subject as an autonomous, self-identifying, and self-determining locus of agency.

Ironically, it was Kant himself who initiated the critique of reason that later inspired postmodern philosophy. Kant observed that reason recognizes no limits in questioning the ultimate metaphysical grounds underlying reality, but that any answer it gives in response to its own questions entails contradiction. Rational inquiry must therefore be limited to phenomena within everyday experience. Kant's critique of pure, experience-transcending reason already anticipated postmodern skepticism regarding the completeness of our knowledge of things in their totality, while rejecting such skepticism with regard to our knowledge of things in their experiential finitude. Kant's rejection of this latter form of skepticism, whose main exponent is David Hume, requires that reason be seen as a synthetic power that infuses experience of objects with causal necessity as it imposes rational identity on the experiencing subject. However, to reconcile the causal necessity of the world with the uncaused freedom of the moral subject, Kant had to divide reason into two *opposed* deployments—theoretical and practical—only one of which was a source of knowledge (he later added a third, aesthetic deployment to mediate between the moral and the theoretical). Subsequent postmodernists continued to divide reason into an indefinite number of context-specific applications, thereby undermining any certain belief in a common reason, a common world, and a common humanity.

Also postmodern is Kant's view that reason questions even its own authority as a certain foundation of knowledge. As G. F. W. Hegel astutely noted, this self-referential (or reflexive) use of reason is paradoxical. By limiting the valid deployment of cognitive reason to natural science, critical philosophy undermines its own claim to validity as a *nonscientific* form of reflective knowledge. Conversely, by grounding natural science in a nonnatural form of transcendental subjectivity, it unwittingly shows natural science and its object to be partial and superficial forms of cognition and reality, respectively.

According to postmodernists, Hegel's system marks the last great attempt to resolve the crisis of reason bequeathed by Kantian philosophy. It does so by affirming what Kant had denied: reason's infinite demand to know the infinite totality. As noted above, this demand issues in contradiction. However, Hegel thought that this was true only if philosophical reflection did not completely grasp all possible metaphysical categories in a manner that showed how each implied all the others. Hegel's circular reasoning would show that the apparent contradictions implicit in metaphysical reasoning ultimately establish a closed system of resolved complementarities. In contrast, any attempt to found one kind of belief deductively on another in a noncircular way, as Kant had proposed, must issue in unresolved contradiction.

Postmodernists question whether reason can establish a complete and coherent system of thought. From Hegel's thought they retain his dialectical view that the fundamental reasons that define, categorize, and ground our beliefs about things effectively refer to properties that are thought to be external or opposed to these things. Thus, while logic (analytic reason) seeks to establish categorical distinctions between self and other, nature and society, reason and unreason, philosophical reasoning about logic undermines these distinctions. Postmodernists therefore conclude that nothing is certain and definite, not even our certainty that we as rational subjects exist.

The undermining of categorical distinctions has an important bearing on the meaningfulness of language. Postmodernists point to the futility of trying to ground the meaning of concepts in empirically verifiable objects or in what is immediately given in experience. As Ludwig Wittgenstein noted in his *Tractatus Logico-philosophicus*, the logical and philosophical metalanguage that is supposed to ground the meaningfulness of the object language in immediate experience is not itself an object language referring to immediate experience. Citing similar self-referential paradoxes made famous by Bertrand Russell, Kurt Gödel, and Werner Heisenberg, Jean François Lyotard has argued that epistemic and logical indeterminacy, incompleteness, and uncertainty necessarily infect any scientific or philosophical metanarrative that claims to be all-encompassing. At the beginning of the twenty-first century, the common acceptance by philosophers of language that meaning is relative to context and usage and yet is underdetermined by them has led philosophers as diverse as Donald Davidson and Jacques Derrida to suggest that meaning is at the very least an indefinite project of textual interpretation, if not, as Lyotard and Foucault argue, an anarchic war of contesting and inventing.

For postmodernists, acknowledging the uncertainty, ambiguity, and loss of identity that comes with the demise of rationalism, humanism, and idealism need not commit us to nihilism. On the contrary, as Friedrich Nietzsche observed, by insisting on impossible norms of certainty, clarity, and identity, we end up devaluing those common unfathomable and uncanny modes of moral and religious experience that open us up to novelty, fantasy, and vulnerability. Worse, by insisting on these impossible norms, we become arrogant and drunk on our own "will to power." It was in the name of pure reason, after all, that "enlightened" Europeans sought to eliminate or assimilate to themselves the "unenlightened" peoples of Africa, Asia, and the Americas. Genuine postmodern responsibility, by contrast, endeavors to promote an active, nondomineering receptivity to the other, no matter how different it may appear.

See also Art, Interpretation of; Art, Value in; Foucault, Michel; Language; Lyotard, Jean François; Rationality.

Bibliography

Best, Steven, and Douglas Kellner. *The Postmodern Turn*. New York: Guilford Press, 1997.

Derrida, Jacques. *Margins of Philosophy*. Translated by Alan Bass. Chicago: University of Chicago Press, 1982.

Derrida, Jacques. *Writing and Difference*. Translated by Alan Bass. Chicago: University of Chicago Press, 1978.

Foucault, Michel. *The Order of Things*. New York: Pantheon, 1970.

Habermas, Jürgen. *The Philosophical Discourse of Modernity*. Translated by Frederick G. Lawrence. Cambridge, MA: MIT Press, 1987.

Horkheimer, Max, and Theodor W. Adorno. *The Dialectic of Enlightenment*. Translated by John Cumming. New York: Herder and Herder, 1972.

Kristeva, Julia. *Revolution in Poetic Language*. Translated by Margaret Waller. New York: Columbia University Press, 1984.

Lyotard, Jean François. *The Postmodern Condition*. Translated by Geoff Bennington and Brian Massumi. Minneapolis: University of Minnesota Press, 1984.

Lyotard, Jean François. *The Postmodern Explained: Correspondence, 1982–1985*. Translated by Julian Pefanis and Morgan Thomas. Minneapolis: University of Minnesota Press, 1993.

Lyotard, Jean François. *Toward the Postmodern*. Translated by Robert Harvey and Mark S. Roberts. Atlantic Highlands, NJ: Humanities Press, 1993.

Natoli, Joseph, and Linda Hutcheon, eds. *A Postmodern Reader*. Albany: State University of New York Press, 1993.

Silverman, Hugh J., ed. *Postmodernism: Philosophy and the Arts*. New York: Routledge, 1990.

Vattimo, Gianni. *The End of Modernity*. Translated by Jon R. Snyder. Baltimore, MD: Johns Hopkins University Press, 1988.

David Ingram (2005)

POTENTIALITY

See *Possibility*

POWER

The meanings of *power, influence, control,* and *domination* are uncertain, shifting, and overlapping. Although two of these words may be interchangeable in one context, in another context one of the words may refer to a genus and another to a species, or one may refer to a cause and another to an effect. To substitute *power* for *influence* would not matter much in the sentence "The United States has very great influence in South American politics," but to interchange them would radically change the meaning of the sentence "Colonel House's power derived not from any constitutional authority but from his influence over President Wilson."

Shifts like this account for much of the intractability of problems associated with power. For instance, power is often said to be a relation (Lasswell 1950, Friedrich 1950, Partridge 1963), yet we talk about the distribution of power, about the power of speech, about seeking power as a means to future enjoyment (Hobbes 1946), or about power as "the production of intended effects" (Russell 1938). If power is a relation, between what kind of terms or things does it hold? Does power over men require a minimum of acquiescence, consent, or cooperation (Hume; Friedrich 1956–1957), or can it be analogous to a physical force acting on an otherwise inert object? Is to exercise power always to succeed in what one intends (Russell 1938, Lasswell 1950), or can a man exercise power in ignorance of what he is doing (Dahl 1961, Partridge 1963, Oppenheim 1961), like a ruling elite that neither knows nor cares about the effects of its actions on other classes?

Instead of seeking a single analysis of *power,* it is more helpful to think of diverse uses of *power* and of associated words like *influence* as instances of different members of a family of concepts that do not all share any one particular characteristic but have various relations and resemblances by which they are recognizably kin. One might construct a power paradigm combining as many of these family features as possible. Thus, "*A,* by his power over *B,* successfully achieved an intended result *r*; he did so by making *B* do *b,* which *B* would not have done but for *A*'s wishing him to do so; moreover, although *B* was reluctant, *A* had a way of overcoming this."

There are five main features of this paradigm: (1) an intention manifest in the exercise of power; (2) the successful achievement of this intention; (3) a relationship between at least two people; (4) the intentional initiation by one of actions by the other; and (5) a conflict of interest or wishes engendering a resistance that the initiator overcomes. Not every feature would be present, of course, in every instance in which we properly speak of power; but we can examine how different instances are related to the paradigm and to one another, and thus throw some light on a few of the questions listed above.

POWER AND CONFLICT

Some instances of power do not involve overcoming resistance to an initiative. A charismatic leader's power over his followers consists in being able merely by suggestion to move them to do willingly what he wants, even though their interests might have led them to act differently. The family of power concepts might be arranged along a conflict scale (Partridge 1963): At the end at which conflict is least would lie instances of influence, while at the other end would lie instances of domination, and in between, instances of authority. In the extreme case, exercising influence would not involve overcoming resistance, for to manipulate a man's actions by shaping what he considers to be his interests is not to impose action upon him *in the face of* his interests. Yet this would still be an instance of power satisfying the first four features of the paradigm.

The limiting case at the end of the scale at which conflict is least would be rational persuasion, for to offer a man good reasons for doing something is not to exercise power over him, although it may influence his decision. One possible difference between influence and

power, then, seems to be that power generally implies a difference of standing between the two parties: The one stimulates, the other reacts. Rational persuasion, on the contrary, to the extent that it criticizes and invites criticism, presupposes at least the possibility of a dialogue between equals. To the extent that persuasion is really rational, the influence is not so much that of the persuader as of his arguments; the same arguments from anyone else would do as well. (By contrast, a threat of violence is more effective coming from a strong man than from a weak man.) Of course, if *A* rationally persuades *B* to help him, *A* may get power—not over *B,* however, but over *C* or *D,* or even simply the power to do something he could not otherwise do.

POWER, INJURY, AND INTEREST

In the case of the man who punishes another for disobedience, conditions (1), (3), and (5) of the paradigm would be satisfied, but not (2) and (4), for the initiative has been refused. Instead, it suffices for an instance of power if the power-holder successfully and intentionally makes the subject suffer for refusing the initiative. And by yet a further extension of meaning, one can exercise power over someone by deliberately making him suffer, whether or not he has refused an initiative. Just as in the limiting case of rational persuasion one could speak of influence but hardly of power, so at the other end of the scale one can talk of power but not of influence, for influence is manifest in what a man is, does, or believes, not in what is simply made to happen to him by another man.

A stoic would probably resist the extension of the concept of power to cover the mere infliction of suffering. By not caring about physical pain or external conditions, he might say, one can remove oneself from the power of another man. So too Martin Luther believed that a true Christian is free because no outer things can touch him at any significant point. It would seem that what characterizes a power situation of this kind is not just the ability to make someone suffer, which after all a dentist possesses, but rather to do him harm—that is, to attack his interests. Thus, by revising the notion of a man's interests, and therefore the notion of harm, the stoic or the Christian can deny the reality of one man's power over another, since nothing that another man can do to me can affect my real interests; I am always free, if once I see what those interests are. This argument is a little odd, because the concept of power generally implies a restriction on choice; but according to the stoic or Christian view, one can always choose to make the restriction insignificant, and therefore one can choose whether to be in the power of another. In that case, there could not be a real restriction, and all power would be illusory. But then, what would power be like if it were real?

The stoic argument demonstrates, however, that whether one man has power over another depends not merely on what he can do to the other but also on the importance to be attached to his action and on whether the subject can reasonably be expected to disregard it. One would not say that *X* was in *Y*'s power if one thought that what *Y* could do to *X* was trivial—something that *X* could or should readily ignore.

Again, although threats of real harm are an exercise of power, bribes or promises of reward are not, unless some special feature of character or situation makes them irresistible—that is, unless no one so placed could reasonably be expected to resist them (although some in fact might). This is not to say, of course, that a man cannot exercise power by bribery. However, it need not be power *over* his hirelings but power over others *through* them; or it may be power only in the still more general sense of an ability to bring about an intended result. Thus, we speak of power in situations in which a man could either successfully determine another's actions or do him harm. An ability to do him some good is not in itself power over him, although the threat of withholding a good that he has come to count on may well be.

PROBLEMS OF POWER AS A RELATION

Power may not be a relation between people but between a person and a thing. There is a nonsocial kind of power that is simply an ability to produce an intended result, like a tenor's power to smash a tumbler with a high C. And even in a social context, the financier's power to destroy a government comes very close to this, for in this instance too power is manifest merely in the active achievement of an intended result. Although the financier no doubt works by initiating actions on the part of others, the relation between him and his object (the government) is that which exists between agent and patient. This case can be distinguished both from that in which power is exercised by punishing a subject for noncompliance and from that in which power is used to inflict deliberate injury. For in the present case the object of the exercise may be only to remove an obstacle. The manifestation of power does not consist in the government's being made to suffer, for it would be just as much a manifestation of power if the financier had chosen instead to prop it up or if the government welcomed its downfall as a blessed release from responsibility. Power is manifest simply in that what happens is the result of the financier's intentional action, just

as the tenor's power is manifest in his being able to break a glass whenever he likes.

Power is of course relational in a logical sense in that it requires more than one term for a complete statement; and if more than one of the terms is a person, and the relation presupposes institutions, rules, and so forth, power will certainly be a social relation. But writers who stress that power is a relation usually mean that it is an initiative-response relationship of the kind that C. J. Friedrich had in mind when he wrote, "The power seeker must find human beings who value the things [he controls] sufficiently to obey his orders in return" (*Constitutional Government and Democracy,* p. 12).

Now, Friedrich's point is substantially true in those instances in which power implies a successful initiative and even perhaps in those instances in which power tends to injure its subject. To set about hurting someone, one must know how to get the right kind of response: There is no point in depriving nonsmokers of tobacco. It is not so clear, however, that the financier's power is of this type, for he does not secure a response from the government; he merely makes something happen to it. Although his agents respond to his initiatives, one must distinguish the power he has over them from the power he has over the government. These powers would be of the same kind only if he were able not just to destroy the government, but to use it as he wished. But it is presumably because he cannot do this that he uses his power to destroy it.

This analysis further elucidates the relation between power and consent. We have seen that at one extreme a man may exercise power over another by influencing his desires, or a man may do whatever he is told by another because he believes that he ought to do so, which is an instance of authority. Both cases imply some measure of consent or acquiescence, if not to the particular initiative, then to the right of the initiator to issue it. But in cases in which power depends on threats or on physical coercion, the subject's acquiescence amounts to no more than that he continues to value whatever is being used as a lever against him—an acquiescence that only the stoic, perhaps, would seriously regard as a matter of choice. However, political power cannot be entirely coercive. The few can rule the many because the many believe either that the few are entitled to do so or that they could harm them if they disobeyed. But they would not think that coercion were possible if they did not also believe that most of the people were prepared to obey without coercion. A political power situation, therefore, must almost always contain some elements of acquiescence as well as coercion—*almost* always because it is at least theoretically possible that a reign of terror might enslave a whole people simply by sowing such mistrust that its opponents could never know their own potential strength.

POWER AND INTENTION

Still further from the paradigm is the case in which one says quite generally that a person is powerful, or that he seeks power, without specifying the range of possible intended action or the persons subject to the power. Usually it would not be difficult to supply terms to complete either one or both of these blanks. Political theorists commonly insist that comparisons of power, without reference to its "domain" and "scope," are meaningless (Lasswell and Kaplan 1950, Oppenheim 1961). However, some have tried to generalize the concept by disregarding intentionality. R. A. Dahl defines power as "the difference in probability of an event, given certain actions by *A,* and the probability of the event given no such action by *A*" ("The Concept of Power," p. 214). At this level of abstraction, *power* is freed not only from intentionality but also from achievement and conflict; what remains is a relation between a stimulus and a reaction. Elsewhere (*Modern Political Analysis,* p. 40), Dahl defines *influence* as a relation among actors in which one induces others to act in some way in which they would not otherwise act. Dahl would want to purge, if he could, the hint of intentionality in the word *induce.* Like a field of force in mechanics, power is a potential for creating disturbance, like the potential of a stone cast in a pond for creating ripples. But this has some odd results. Instead of suffering a *loss* of power, the crashing financier who brings down thousands with him in his fall would be exercising a power that is perhaps greater than ever before. Admittedly, it is a mark of power if a man's actions cause disturbances, even if he is careless or even ignorant of them. Nevertheless, if powerful men cause incidental and unintended disturbances, they do so in the course of *getting what they want.* (C. Wright Mills's conception of a "power elite" seems to be of this kind.) One would not call someone powerful who, like a careless smoker constantly causing fires, was forever causing disturbances but never achieving anything he intended; nor is it clear that any useful methodological purpose in political science would be served by a definition of power that permitted the production of unintended effects alone to serve as a criterion.

To possess power or to be powerful is, then, to have a generalized potentiality for getting one's own way or for bringing about changes (at least some of which are intended) in other people's actions or conditions. *Influence,* it is true, is used in a more general sense. If a parent

has the unintended influence of stiffening his child's determination to be as different from him as possible, this would not be described as an instance of power: It is more like "the influence of climate on national character." The use of the term *influence* suggests that there is a causal relationship between the behavior of the parent and that of the child (cf. P. H. Partridge, "Some Notes on the Concept of Power," p. 114). "A writer's influence on succeeding generations" stands somewhere between this case and that of influence by rational persuasion. For a writer may have influence only to the extent that other writers recognize his merits and choose to imitate him. Although such influence may not be intended, still it is not a cause, at least in the sense that climate is a possible cause of national character. In any case, none of these is an instance of an influence in the sense that House *had* influence with Wilson. "To *use* one's influence" usually implies actively and intentionally working through or on other people, and one who can do this recurrently "*has* influence." Of course, people who have power (that is, who can do many things they want and induce many other people to accept their initiative) are likely on that account to influence (that is, to have effects on) other aspects of society in ways that neither they nor their social inferiors necessarily understand. Other classes, envying and admiring them, may imitate their tastes and practices, and in this sense they may be influenced by them. But this influence is not a manifestation of power; it is only one of its effects.

See also Authority; Feminist Legal Theory; Freedom; Luther, Martin; Rights; Sovereignty; Stoicism; Violence.

Bibliography

CLASSIC THEORIES OF POLITICAL POWER

Aiken, H. D., ed. *Hume's Moral and Political Philosophy.* New York: Hafner, 1948.

Hobbes, Thomas. *Leviathan,* edited by M. Oakeshott. Oxford: Blackwell, 1946.

Hume, David. "Of the First Principles of Government." In *Essays Literary, Moral and Political* (1741). London, 1963.

Hume, David. *Treatise of Human Nature* (1740), edited by L. A. Selby-Bigge. Oxford, 1951. Book III.

Plamenatz, John. *Man and Society.* 2 vols. London: Longmans, 1963. A critical discussion of classic theories.

Spinoza, Benedict de. *The Political Works.* Translated and edited by A. G. Wernham. Oxford: Clarendon Press, 1958.

Spinoza, Benedict de. *Tractatus Politicus* (1677).

Spinoza, Benedict de. *Tractatus Theologico-Politicus.* Hamburg, 1670.

GENERAL PHILOSOPHICAL DISCUSSIONS

Acton, H. B. "Logique et casuistique du pouvoir." In *Annales de philosophie politique: Tomes I et II: Le Pouvoir,* Vol. II, 69–86. Paris, 1956–1957.

Emmet, Dorothy. "The Concept of Power." *PAS* 54 (1953–1954): 1–26.

Friedrich, C. J. "Le probleme de pouvoir dans la theorie constitutionnaliste." In *Annales de Philosophie Politique: Tomes I et II: Le Pouvoir,* Vol. I, 33–51. Paris, 1956–1957.

Jouvenel, Bertrand de. *Le pouvoir.* Geneva, 1945. Translated by J. F. Huntington as *On Power; Its Nature and the History of Its Growth.* Boston, 1962.

Jouvenel, Bertrand de. *Pure Theory of Politics.* Cambridge, U.K.: Cambridge University Press, 1963.

Partridge, P. H. "Politics and Power." *Philosophy* 38 (1963): 117–136.

Partridge, P. H. "Some Notes on the Concept of Power." *Political Studies* 2 (1963): 107–125.

Polin, R. "Sens et fondement du pouvoir chez John Locke." In *Annales de Philosophie Politique: Tomes I et II: Le Pouvoir,* Vol. I, 53–90. Paris, 1956–1957.

Russell, Bertrand. *Power.* London: Allen and Unwin, 1938.

SOCIOLOGY AND METHODOLOGY OF POLITICAL SCIENCE

Bierstedt, Robert. "An Analysis of Social Power." *American Sociological Review* 15 (1950): 730–738.

Dahl, R. A. "The Concept of Power." In *Introductory Readings in Political Behavior,* edited by Sidney S. Ulmer. Chicago: Rand McNally, 1961.

Friedrich, Carl J. *Constitutional Government and Democracy.* Boston: Ginn, 1950.

Friedrich, Carl J. *Man and His Government.* New York: McGraw-Hill, 1963.

Lasswell, Harold D., and Abraham Kaplan. *Power and Society: A Framework for Political Inquiry.* New Haven, CT: Yale University Press, 1950.

Merriam, Charles E. "Political Power." In *A Study of Power,* by H. D. Lasswell, C. E. Merriam, and T. V. Smith. Glencoe, IL: Free Press, 1950.

Oppenheim, Felix E. *Dimensions of Freedom.* New York: St. Martin's Press, 1961.

Walter, E. V. "Power and Violence." *American Political Science Review* 58 (1964): 350–360.

Weber, Max. *From Max Weber; Essays in Sociology.* Translated and edited by H. H. Gerth and C. Wright Mills. London: Routledge and Kegan Paul, 1948.

ADDITIONAL SOURCES

Arendt, Hannah. *On Violence.* New York: Harcourt Brace, 1970.

Connolly, William E. *The Terms of Political Discourse,* 2nd ed. Princeton, NJ: Princeton University Press, 1983.

Dahl, R. A. *Modern Political Analysis,* 5th ed. Englewood Cliffs, NJ: Prentice Hall, 1984.

Foucault, Michel. *Power,* edited by James D. Faubion; translated by Robert Hurley. New York: New Press, 2000.

Lukes, Steven. *Power: A Radical View.* London: Macmillan, 1974.

Morriss, Peter. *Power: A Philosophical Analysis.* Manchester, U.K.: Manchester University Press, 1987.

Parsons, Talcott. *Politics and Social Structure.* New York: Free Press, 1969.

Scott, John. *Power.* Cambridge, U.K.: Blackwell, 2001.

Wrong, Dennis H. *Power: Its Forms, Bases, and Uses.* New York: Harper and Row, 1979.

Stanley I. Benn (1967)
Bibliography updated by Philip Reed (2005)

PRACTICAL REASON

Reason can and should guide one in deciding what to believe, at least in large part. But can reason also guide one's actions and the goals that one aims to achieve through them? This question is at the heart of philosophical interest in practical reason. One's thoughts and discourse about practical matters are full of references to reason, and each day brings with it a fresh round of deliberation over such things as the costs and benefits of alternative lines of conduct. Disagreement over how best to understand these phenomena has focused on two distinct questions: First, is reason itself ever a genuine source of considerations for or against conduct, or is our everyday thought and discourse simply a *façon de parler*? Second, to what extent, if at all, can such considerations make a difference to what one does? Under the first question, which is address in the first three sections of this entry, the central issues concern whether and the extent to which the deliberative process that culminates in a decision or intention can be dubbed *reasoning*. The second question, with which the article ends, concerns the nature of motivation and action and, in particular, what role (if any) reason plays in the explanation of one's behavior.

INSTRUMENTAL PRACTICAL REASON

Most agree that if any conduct is contrary to reason, then not acting to achieve one's goals with some level of efficiency and effectiveness is. Once one decides to lose weight, for instance, overeating seems unreasonable. But what precisely is reason's role here? Many, such as those who follow the eighteenth-century Scottish philosopher David Hume, argue that its role is limited to delivering and evaluating beliefs about connections between behaving in certain ways and achieving goals. If one's goal is to lose weight, reason's work is done once it delivers the news that eating less will bring that about. This implies that reason concerns itself only with delivering causal information about how to realize one's goals and hence does work in the realm of action that is no different from the work it does in the realm of belief.

To see precisely how little reason does on such a view, consider the following: Suppose my goal is to have the doorbell to ring and I am told to push the button. However, I perversely insist that it is a trick door and that standing motionless will make it ring. If I stand motionless, I will as a consequence frustrate my goal. Reason seems against my conduct. For the minimalist, however, that means only that the belief on which my conduct was based was false. Suppose, alternatively, that I have no idea how to make a doorbell ring. I stand in puzzled silence, and again reason fails to support my behavior. This time, the problem is not that I have incorrect beliefs about how to achieve my goal; it is that I have no beliefs about this at all. Nevertheless, it is again really just in lacking a belief that I've fallen foul of reason. Conforming to practical reason, on this minimalist view, means simply ensuring that I have the right stock of beliefs about how to achieve my ends. Reason does not pass judgment on what I do per se. It is thus not practical in this more interesting sense. Indeed, when I do what I falsely believe will bring results, my action displays a kind of fit with my belief, even if my belief is itself defective.

Suppose, however, that my actions did not display this kind of fit with my beliefs, even while my beliefs were flawless. Imagine, for instance, I failed to push the button when my goal was that the doorbell ring and I correctly believed that pushing the button would achieve this. Was my conduct *then* contrary to reason insofar as it did not fit with my goal and my true means–ends belief? A minimalist such as Hume would deny that it was. An action itself cannot be contrary to reason because an action cannot be evaluated for its truth or falsity. *A fortiori* reason cannot justify an action either, for justification of something is just support for its truth. Thus, no action seems contrary to reason in the way that a false or unjustified belief is.

Reason will be practical in an interesting sense, it seems, only if one of two things is true: Either there exists some special realm of facts about the *to-be-doneness* of certain actions themselves, information about which reason can deliver, or else reason is more than an information-delivering faculty. Philosophers have tended to avoid views requiring special facts although in the case of ethical reasoning, some have thought the idea worth developing. This case will be returned to below. For many, a more attractive strategy is to argue that reason issues distinctive rules of conduct. The most likely candidate for such a rule would be a rule of instrumental reason, for instance: Do what is necessary to achieve your goals. For an action to be contrary to reason would then be for it to fail to con-

form to such a rule. The key issue, then, is whether practical reason is indeed *normative* in this sense, that is, whether there are any genuine rules of reason.

Arguments that a given rule is a norm of reason can be grouped into two kinds: those appealing to the concept of reasonableness, and those appealing to substantive considerations beyond that concept. The eighteenth-century Prussian philosopher Immanuel Kant, for instance, employed the first style of argument regarding conformity to a *hypothetical imperative*. The concept of *reasonable behavior*, he argued, contains the idea of conformity to the rule *take the means necessary to achieve your goals*. The twentieth-century political philosopher John Rawls is an example of a philosopher who also employed the second style of argument. Rawls argued that reasonableness includes a willingness to propose and abide by fair terms of cooperation if assured that others will likewise do so, on the grounds that, although it is not a conceptual truth, the contention enjoys much intuitive support.

REASONING ABOUT GOALS

Goals can share many of the above features of actions. Suppose, again, my goal is that the doorbell ring. Typically, I don't just want that. Perhaps I believe that the ring will bring my friend to the door. My goal is really an instrumental goal, a goal that is desirable because its achievement is instrumental to achieving a further distinct goal. Suppose, however, that I am standing in front of the wrong house. Even though I am right to believe that pushing the button will achieve the ring, reason is against my pushing the button because it is against achieving the ring. To this instrumental extent at least, our goals can be contrary to reason.

Minimalists will be led say about goals *mutatis mutandis* what they say about action: The defect, as in the case of action, is in the belief that the ring will bring my friend to the door. It is only because of this false belief that my goal falls foul of reason. Goals are just like actions in the sense that they cannot be evaluated as true or false, and *a fortiori* cannot be justified or unjustified either. So, if reason were practical in any interesting sense, there would either have to be a distinctive realm of facts about the *to-be-pursuedness* of certain goals or else reason would have to issue distinctive norms concerning goals such as *pursue intermediate goals necessary to reach your primary goals*.

This sort of reasoning need not exhaust practical reasoning about ends. For instance, suppose I have not one, but two goals: that the doorbell ring and that those behind the door not be disturbed. Do I conform to reason if I push, or rather fail to push, the button? Given the bell cannot ring and leave the inhabitants undisturbed, the answer must wait until I resolve this conflict. Having goals that are not *jointly realizable* seems contrary to reason. However, goals are jointly realizable only if some can be dropped in favor of others in cases of conflict. We could do this willy-nilly, of course. But ranking seems more reasonable. We should decide whether having the doorbell ring is more or less important than disturbing those behind the door. Given reason counsels joint realizability, it thus also counsels ranking. Moreover, rankings conform to requirements of consistency. For instance, they are transitive: If ringing is ranked above not disturbing the inhabitants, and not disturbing them above not wearing out the button very slightly, then ringing should be ranked above not wearing out the button. This would explain why we would think it unreasonable for me to worry about wearing out the button given that I'm not worried about the more important fact that it will disturb them.

Presumably, one does not pursue all of one's goals for the sake of other goals, however. Some things one cares about for their own sakes; they are *final* goals. Can reason evaluate such final goals? One way that it might is this (Schmidtz 1995): Suppose I am a philistine, but then decide to become the sort of person for whom art is a final end. Suppose further that I decide this because I believe that becoming that sort of person will enhance my standing in the eyes of others. I aim, in other words, to come to pursue something for its own sake, but my reasoning is clearly instrumental. If I find out that learning to love art for arts sake will not lead others to think better of me, then reason will counsel me not to learn that.

REASON IN ETHICAL DELIBERATION

When one deliberates about what to do, one often considers whether what one proposes is morally permissible, right, virtuous, and so on. One seems to care about such things for their own sakes, so this seems to constitute a final end. But does reason ever really guide one to moral conclusions?

Those who think that it does can be divided into two camps: those who think that moral reasoning can be explained in terms of reasoning from individual goals, and those who think it involves a special kind of reasoning. The former think that moral reasoning is, in fact, not fundamentally different from the above forms of practical reasoning but in some way facilitates the achievement of one's goals—typically, by being based on principles of social conduct that reasonable individuals would accept and act on. Such, for instance, is found in game-theoretic

explanations of morality. On a standard version, game theorists argue that people seeking to achieve their goals will want ground rules for their interactions with each other. They will thus freely engage in a series of bargains with others in which each person tries to secure practices most favorable to their goals. Bargaining would continue until no viable alternative agreement can be struck under which someone would be better off. Moral practices represent these agreements, and because they do, they are justified in light of their being the upshot of these reasonable goal-oriented bargains. Along these lines, David Gauthier (1986) argues that reasonable agents will be disposed to cooperate with others who likewise cooperate, even when doing so will not be the best way to achieve their own goals (as is often the case in moral matters).

A more controversial idea is that moral reasoning is fundamentally different from nonmoral reasoning. There are two main lines of thought here: The first is that there is a distinct realm of moral facts, as real as any scientific fact but accessible only through the exercise of a special faculty of reason. On this view, practical reason operates quasi-perceptually to deliver putative moral facts such as that lying is wrong. Some (McDowell 1979) have held that this is analogous to sense perception, such as is exercised by informed palates when they perceive differences between wines. Others (Ross 1939) think of it as more akin to intellectual perception, such as is exercised in the perception of mathematical truths. Many, however, find this postulation of a *sui generis* faculty of reason too mysterious to accept.

The second line of thought does not appeal to the exercise of a special faculty and access to special facts but to a special rule distinct from those connected to advancing individual goals. The most famous attempt to defend this line of thinking comes from Kant. Moral reasoning is based, he argued, on a rule he referred to as the *Categorical Imperative*. This rule requires one not to act in ways that one could not want everyone else also to act. Every rational agent is committed to this rule, Kant argued, simply by engaging in practical reasoning. Committing oneself to this rule is a presupposition of taking up the point of view of practical deliberation. Therefore, he concluded, it is a rule of practical reason. Few have found Kant's arguments convincing. Nevertheless, some contemporary philosophers have tried to develop and defend some version of Kant's ideas. Rawls's idea of reasonableness is one attempt. Another is Thomas Scanlon (1998) who argues that reasonableness requires being responsive to the appropriateness of principles of conduct to serve as foundations for mutual recognition and accommodation.

REASON AND MOTIVATION

Suppose deliberating to conclusions about what to do is genuinely a form of reasoning. These conclusions may still make no difference to what one does. Reason, that is, may not be practical in another sense—in the sense that it cannot motivate one to comply with its conclusions. When one acts contrary to conclusions of practical deliberation, is one unreasonable in the sense of being insufficiently motivated by this deliberation?

Internalists about practical reason hold that one can be: The conclusions of practical reasoning must motivate reasonable agents. This is especially the case, they argue, in moral reasoning: It is not possible to believe it to be wrong to lie, for instance, yet remain unmotivated to tell the truth. One reason internalism is attractive is that it explains the magnetism conclusions of practical reasoning exhibit. To be sure, the conclusions of practical reasoning do not always motivate everyone. If one is depressed or weak-willed, for instance, practical conclusions may have no motivational effect on one. So, internalists must stipulate which psychological condition a reasonable agent is in such that that agent must be motivated. This has not proven to be an easy task.

Internalism, however, appears inconsistent with an attractive conception of motivation, often referred to as the Humean view. On this view, motivation requires, in addition to belief, a desire. Michael Smith (1995) has offered an influential defense of this view. Briefly, the leading idea is that the best functional account of belief and desire gives them different directions of fit with the world. A belief is an attitude toward a given proposition *p*, such that the perception of not-*p* disposes the believer to change attitude to not-*p*. Desire has the reverse direction of fit: an attitude toward *p* such that the perception of not-*p* disposes the desirer to change the world to *p*. If these accounts are basically right, then three things seem clear: Motivation requires a desire, beliefs and desires are only contingently related, and no state could have both directions of fit. Smith himself argues that, nonetheless, one's beliefs about what one has reason to do must motivate agents in the right psychological condition. His position is controversial, however, and the prospects for internalism remain unclear.

See also Decision Theory; Game Theory; Hume, David; Kant, Immanuel; Normativity; Rationalism in Ethics (Practical Reason Approaches); Rawls, John; Reason.

Bibliography

Blackburn, Simon. *Ruling Passions*. Oxford: Clarendon, 1998.

Bratman, Michael. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.
Broome, John. "Normative Practical Reasoning." 175–193. *Aristotelian Society: Supplementary Volume*. 2001.
Dancy, Jonathan. *Moral Reasons*. Oxford: Basil Blackwell, 1993.
Darwall, Stephen. *Impartial Reason*. Ithaca, NY: Cornell University Press, 1983.
Gauthier, David. *Morals by Agreement*. Oxford: Oxford University Press, 1986.
Gibbard, Alan. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press, 1992.
Hill, Thomas E., Jr. *Dignity and Practical Reason*. Ithaca, NY: Cornell University Press, 1992.
Hume, David. *A Treatise of Human Nature*, edited by D. F. Norton and M. Norton. New Ed. Oxford: Oxford University Press, 2000.
Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, edited by A. Zweig and T. E. Hill, Jr. Translated by A. Zweig. Oxford: Oxford University Press, 2002.
Korsgaard, Christine. "The Normativity of Instrumental Reason." In *Ethics and Practical Reason*, edited by Garrett Cullity and Berys Gaut. Oxford: Oxford University Press, 1997.
Korsgaard, Christine. "Skepticism about Practical Reason." *Journal of Philosophy* (1986): 5–25.
McDowell, John. "Are Moral Requirements Hypothetical Imperatives?" *Aristotelian Society Supplementary Volume* (1978): 13–29.
McDowell, John. "Virtue and Reason." *Monist* 62 (1979): 331–50.
Nagel, Thomas. *The Possibility of Altruism*. Princeton, NJ: Princeton University Press, 1978.
Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
Rawls, John. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77 (1980): 515–572.
Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1972.
Raz, Joseph. *Practical Reason and Norms*. Oxford: Oxford University Press, 1975.
Ross, W. D. *Foundations of Ethics*. Oxford: Clarendon Press, 1939.
Scanlon, Thomas. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
Schmidtz, D. *Rational Choice and Moral Agency*. Princeton, NJ: Princeton University Press, 1995.
Smith, Michael. *The Moral Problem*. New York: Routledge, 1995.
Stocker, Michael. "Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy* 76 (1979): 738–53.
Velleman, David. "The Possibility of Practical Reason." *Ethics* 106, No. 4 (1996): 694–726.
Wallace, R. Jay. "How to Argue About Practical Reason." *Mind* 99 (395) (1990): 355–385.
Wallace, R. Jay. "Normativity, Commitment, and Instrumental Reason." *Philosopher's Imprint*. 1 (3) (2001): 1–26, http://www.philosophersimprint.org/001003.
Wiggins, David. "Deliberation and Practical Reason." In *Essays on Aristotle's Ethics*, edited by A. O. Rorty. Berkeley: University of California Press, 1980.
Williams, Bernard. "Internal and External Reasons." In *Moral Luck: Philosophical Papers, 1973–1980*. Cambridge, U.K.: Cambridge University Press, 1981.

Robert N. Johnson (2005)

PRACTICAL REASON APPROACHES

See *Rationalism in Ethics (Practical Reason Approaches)*

PRAGMATICISM

See *Peirce, Charles Sanders; Pragmatism*

PRAGMATICS

"Pragmatics" was defined by Charles W. Morris (1938) as the branch of semiotics that studies the relation of signs to interpreters, in contrast with semantics, which studies the relation of signs to designata. In practice, it has often been treated as a repository for any aspect of utterance meaning beyond the scope of existing semantic machinery, as in the slogan "Pragmatics = meaning minus truth conditions" (Gazdar 1979). There has been some doubt about whether it is a homogeneous domain (Searle, Kiefer, and Bierwisch 1980).

A more positive view emerges from the work of Herbert Paul Grice, whose *William James Lectures* (1967) are fundamental. Grice showed that many aspects of utterance meaning traditionally regarded as conventional, or semantic, could be more explanatorily treated as conversational, or pragmatic. For Gricean pragmatists, the crucial feature of pragmatic interpretation is its inferential nature: the hearer is seen as constructing and evaluating a hypothesis about the communicator's intentions, based, on the one hand, on the meaning of the sentence uttered, and on the other, on contextual information and general communicative principles that speakers are normally expected to observe. (For definition and surveys see Levinson 1983.)

THE SEMANTICS-PRAGMATICS DISTINCTION

In early work, the semantics-pragmatics distinction was often seen as coextensive with the distinction between truth-conditional and non-truth-conditional meaning

(Gazdar 1979). On this approach, pragmatics would deal with a range of disparate phenomena, including (a) Gricean conversational inference, (b) the inferential recognition of illocutionary-force, and (c) the conventional meanings of illocutionary-force indicators and other non-truth-conditional expressions such as *but, please, unfortunately* (Recanati 1987). From the cognitive point of view, these phenomena have little in common.

Within the cognitive science literature in particular, the semantics-pragmatics distinction is now more generally seen as coextensive with the distinction between decoding and inference (or conventional and conversational meaning). On this approach, all conventional meaning, both truth-conditional and non-truth-conditional, is left to linguistic semantics, and the aim of pragmatic theory is to explain how the gap between sentence meaning and utterance interpretation is inferentially bridged. A pragmatic theory of this type is developed in D. Sperber and D. Wilson (1986).

IMPLICATURE

Grice's distinction between saying and implicating crosscuts the semantics-pragmatics distinction as defined above. For Grice, "what is said" corresponds to the truth-conditional content of an utterance, and "what is implicated" is everything communicated that is not part of what is said. Grice saw the truth-conditional content of an utterance as determined partly by the conventional (semantic) meaning of the sentence uttered, and partly by contextual (pragmatic) factors governing disambiguation and reference assignment. He saw conventional (semantic) implicatures as determined by the meaning of discourse connectives such as *but, moreover* and *so,* and analyzed them as signaling the performance of higher-order speech acts such as contrasting, adding and explaining (Grice 1989). An alternative analysis is developed in D. Blakemore (1987).

Among nonconventional (pragmatic) implicatures, the best known are the conversational ones: These are beliefs that have to be attributed to the speaker in order to preserve the assumption that she was obeying the "cooperative principle" (with associated maxims of truthfulness, informativeness, relevance, and clarity), in saying what she said. In Grice's framework, generalized conversational implicatures are "normally" carried by use of a certain expression, and are easily confused with conventional lexical meaning (Grice 1989). In Grice's view, many earlier philosophical analyses were guilty of such confusion.

Grice's account of conversational implicatures has been questioned on several grounds:

(1) The status and content of the cooperative principle and maxims have been debated, and attempts to reduce the maxims or provide alternative sources for implicatures have been undertaken (Davis 1991, Horn 1984, Levinson 1987, Sperber and Wilson 1986).

(2) Grice claimed that deliberate, blatant maxim-violation could result in implicatures, in the case of metaphor and irony in particular. This claim has been challenged, and alternative accounts of metaphor and irony developed, in which no maxim-violation takes place (Blakemore 1992, Hugly and Sayward 1979, Sperber and Wilson 1986).

(3) Pragmatic principles have been found to make a substantial contribution to explicit communication, not only in disambiguation and reference assignment, but in enriching the linguistically encoded meaning in various ways. This raises the question of where the borderline between explicit and implicit communication should be drawn (Sperber and Wilson 1986, 1995). It has even been argued that many of Grice's best-known cases of generalized conversational implicature might be better analyzed as pragmatically determined aspects of what is said (Carston 1988, Recanati 1989).

(4) The idea that the context for utterance interpretation is determined in advance of the utterance has been questioned, and the identification of an appropriate set of contextual assumptions is now seen as an integral part of the utterance-interpretation process (Blakemore 1992, Sperber and Wilson 1986).

PROSPECTS

Within the cognitive science literature, several approaches to pragmatics are currently being pursued. There are computational attempts to implement the Gricean program via rules for the recognition of coherence relations among discourse segments (Asher and Lascarides 1995, Hobbs 1985). Relations between the Gricean program and speech-act theory are being reassessed (Tsohatzidis 1994). The cognitive foundations of pragmatics and the relations of pragmatics to neighboring disciplines are still being explored (Sperber and Wilson 1995, Sperber 1994). Despite this diversity of approaches, pragmatics now seems to be established as a relatively homogenous domain.

See also Cognitive Science; Grice, Herbert Paul; Metaphor; Philosophy of Language; Reference; Semantics.

Bibliography

Asher, N., and A. Lascarides. "Lexical Disambiguation in a Discourse Context." *Journal of Semantics* 12 (1995): 69–108.

Blakemore, D. *Semantic Constraints on Relevance.* Oxford: Blackwell, 1987.

Blakemore, D. *Understanding Utterances.* Oxford: Blackwell, 1992.

Carston, R. "Explicature, Implicature and Truth-Theoretic Semantics." In *Mental Representations: The Interface between Language and Reality,* edited by R. Kempson. Cambridge, U.K.: Cambridge University Press, 1988.

Davis, S., ed. *Pragmatics: A Reader.* New York: Oxford University Press, 1991.

Gazdar, G. *Pragmatics: Implicature, Presupposition and Logical Form.* New York: Academic Press, 1979.

Grice, H. P. "Logic and Conversation." *William James Lectures.* Cambridge, MA, 1967.

Grice, H. P. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press, 1989.

Hobbs, J. "On the Coherence and Structure of Discourse." Center for the Study of Language and Information (October 1985).

Horn, L. "A New Taxonomy for Pragmatic Inference: Q-Based and R-Based Implicature." In *Meaning, Form and Use in Context,* edited by D. Schiffrin. Washington, DC, 1984.

Hugly, P., and C. Sayward. "A Problem about Conversational Implicature." *Linguistics and Philosophy* 3 (1979): 19–25.

Levinson, S. "Minimization and Conversational Inference." In *The Pragmatic Perspective,* edited by J. Verschueren and M. Bertuccelli-Papis. Amsterdam: Benjamins, 1987.

Levinson, S. *Pragmatics.* Cambridge, U.K.: Cambridge University Press, 1983.

Morris, C. "Foundations of the Theory of Signs." In *International Encyclopedia of Unified Science,* edited by O. Neurath, R. Carnap, and C. Morriss. Chicago: University of Chicago Press, 1938.

Recanati, F. *Meaning and Force.* Cambridge, U.K.: Cambridge University Press, 1987.

Recanati, F. "The Pragmatics of What Is Said." *Mind and Language* 4 (1989): 295–329.

Searle, J., F. Kiefer, and M. Bierwisch, eds. *Speech-Act Theory and Pragmatics.* Dordrecht: Reidel, 1980.

Sperber, D. "Understanding Verbal Understanding." In *What Is Intelligence?,* edited by J. Khalfa. Cambridge, U.K.: Cambridge University Press, 1994.

Sperber, D., and D. Wilson. *Relevance: Communication and Cognition.* Oxford, 1986.

Sperber, D., and D. Wilson. "Postface" to the second edition of *Relevance.* Oxford, 1995.

Tsohatzidis, S., ed. *Foundations of Speech-Act Theory: Philosophical and Linguistic Perspectives.* London: Routledge, 1994.

Deirdre Wilson (1996)

PRAGMATICS [ADDENDUM]

A major focus of post-Gricean pragmatics is the role that pragmatic inference plays in determining the explicit content of utterances (as opposed to their conversational implicatures). As well as disambiguation and reference fixing, there are pragmatic processes of propositional completion, as in the examples in (1), and, more controversially, processes of "free" enrichment, as in (2):

(1)	a.	It's too late.	*[for what?]*
	b.	Cotton is better.	*[than what?]*
(2)	a.	I've had breakfast.	*[today]*
	b.	John's car hit Tom's and Tom stopped illegally.	*[causal relation]*

The pragmatic completions in (1) are mandated by aspects of the linguistic semantics of the sentences, specifically by the lexical items *too* and *better*. However, this does not seem to be the case for the examples in (2), which express complete, truth-evaluable propositions without the bracketed addition. These pragmatic inferences seem to be entirely pragmatically motivated (i.e., "free" from linguistic indication); they are undertaken in order to satisfy standing communicative presumptions concerning the informativeness and relevance of utterances. For instance, (2a) is strictly speaking true provided the speaker has had breakfast sometime in her life, but in most contexts a speaker intends a more specific proposition and relevant implications hinge on the enriched content (e.g., "that she is not hungry at this moment"). Another kind of free pragmatic process is "lexical modulation": the encoded meaning of a word may be narrowed down in context (e.g., *drink* used to mean "alcoholic drink"), broadened (e.g., *square* used to mean "squarish") or metaphorically extended (e.g., *nightmare* used to mean "unpleasant experience").

The view that "free" pragmatic inferences can affect explicit content in these ways is labeled "truth-conditional pragmatics" and is held by pragmatists across different theoretical persuasions. Various accounts of the phenomenon and its relation to conversational implicature are being developed. Stephen Levinson (2000) argues for a system of "default" pragmatic inferences triggered by particular linguistic forms (e.g., *and*, *some*, *drink*), which are distinct from the kind of inferences responsible for more context-specific implicatures. François Recanati (2003) makes a different distinction between two kinds of pragmatic processes: "primary" processes, such as free enrichment, which contribute to truth-conditional con-

tent, are a matter of local associative processing, whereas "secondary" ones, which account for implicatures, are cases of global propositional inference, constrained by Gricean maxims. Relevance theorists, led by Dan Sperber and Deirdre Wilson, argue that all pragmatic inference can be accounted for by a single principle geared to the recovery of an "optimally relevant" interpretation and that pragmatic enrichment of explicit content often occurs in order to ensure an inferentially sound basis for an antecedently derived conversational implicature.

An alternative, more semantically oriented position, represented by Jason Stanley (2000), denies the existence of processes of "free" pragmatic enrichment and claims that all aspects of an utterance's truth-conditional content are indicated in its linguistic form. So the examples in (2) are to be explained in the same way as the examples in (1): There is a covert indexical element in their linguistic form and it is this that triggers the pragmatic process of finding the relevant contextual value.

Which of these views is correct (if either) remains to be seen.

See also Grice, Herbert Paul; Metaphor; Non-Truth-Conditional Meaning; Reference; Semantics; Semantics, History of.

Bibliography

Bach, Kent. "Speaking Loosely: Sentence Nonliterality." In *Midwest Studies in Philosophy*. Vol. 25, *Figurative Language*, edited by Peter French and Howard Wettstein. Oxford, U.K.: Blackwell, 2001.

Carston, Robyn. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford, U.K.: Blackwell, 2002.

Levinson, Stephen. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press, 2000.

Recanati, François. *Literal Meaning*. Cambridge, U.K.: Cambridge University Press, 2003.

Stanley, Jason. "Context and Logical Form." *Linguistics and Philosophy* 23 (2000): 391–434.

Wilson, Deirdre, and Dan Sperber. "Relevance Theory." In *Handbook of Pragmatics*, edited by Laurence Horn and Gregory Ward. Oxford, U.K.: Blackwell, 2004.

Robyn Carston (2005)

PRAGMATISM

"Pragmatism" was the most influential philosophy in America in the first quarter of the twentieth century. Viewed against the widely diversified intellectual currents that have characterized American life, pragmatism stands out as an energetically evolved philosophical movement. As a movement it is best understood as, in part, a critical rejection of much of traditional academic philosophy and, in part, a concern to establish certain positive aims. It is in these respects, rather than because of any one idea or exclusive doctrine, that pragmatism has been the most distinctive and the major contribution of America to the world of philosophy. Among the Continental thinkers it has influenced and with whose philosophy it has been in harmony are Georg Simmel, Wilhelm Ostwald, Edmund Husserl, Hans Vaihinger, Richard Müiller-Freienfels, Hans Hahn, Giovanni Papini (leader of the Pragmatist Club in Florence), Giovanni Vailati, Henri Bergson, and Édouard Le Roy.

BACKGROUND

The origins of pragmatism are clear in outline, if not in detail. The familiar capsule description is as follows: Pragmatism is a method of philosophizing—often said to be a theory of meaning—first developed by Charles Sanders Peirce in the 1870s; revived and reformulated in 1898 by William James, primarily as a theory of truth; further developed, expanded, and disseminated by John Dewey and Ferdinand Canning Scott Schiller.

This glossing of the facts is useful as a summary or for directing us where to look if we want to find out more about pragmatism. But it can be misleading. A reexamination or rewriting of the history is not to be embarked upon here; but the following cautionary points deserve mention. The specific formative conditions of the early evolution of pragmatism are not entirely clear for several reasons. The historical occasion of the birth of pragmatism is complicated because it was to some extent the product of cooperative deliberation and mutual influences within the "Metaphysical Club," founded by Peirce, James, and others in the 1870s in Cambridge. This may be one of the very few cases in which a philosophy club produced something notable philosophically (compare John Locke's account of the "club" in the 1670s that stimulated the writing of his great *Essay*). But the paper (now lost) that Peirce drew up as a memento lest the club dissolve without leaving behind anything substantial, the paper in which pragmatism was first expressed, was not the free creation of one mind, even though the major credit surely goes to Peirce. Years later, undertaking to write on pragmatism, Peirce queried James: "Who originated the term *pragmatism*, I or you? Where did it first appear in print? What do you understand by it?" And James replied with the reminder: "You invented 'pragmatism' for which I

gave you full credit in a lecture entitled 'Philosophical Conceptions and Practical Results.'"

In addition to some uncertainty as to the facts in the evolution of pragmatism, there are—as we shall see—several problems of interpretation. Peirce and James often gave very different accounts of what they understood by "pragmatism." Usually this is explained by holding James responsible for distorting or even misunderstanding Peirce's ideas. That there were differences between Peirce and James on this score is clear. Peirce, despairing of what James (and his followers) were making of the idea, rebaptized his own view as "pragmaticism," a word ugly enough, he commented, to keep it safe from kidnapers. Historians usually side with Peirce, tending to discredit James's overzealous pronouncements upon pragmatism and applications of it to issues of the moral value and truth of religious belief. But with equal justice it can be maintained that James was developing a substantially different approach to a different type of philosophical problem, related in some ways to Peirce's thought, but mostly superficially; only his habitual overgenerosity led him to call what he was doing "pragmatism" and to cite Peirce as the "inventor."

There is, however, a more serious and persistent problem of interpretation entrenched in the history of pragmatism. This is the problem of determining with some precision what "pragmatism" means or stands for as a philosophical doctrine. As already suggested, pragmatism, by virtue of being an evolving philosophical movement, is to be viewed as a group of associated theoretical ideas and attitudes developed over a period of time and exhibiting—under the differing influences of Peirce, James, and Dewey—rather significant shifts in direction and in formulation. We have the advantage of historical perspective and can make use of it to survey and select distinctive themes and phases in the formation of pragmatism, but a single definitive statement of a single thesis is not to be hoped for.

In the heyday of pragmatism its rapidly changing character proved to be a source of embarrassment and confusion to pragmatists and critics alike. Arthur O. Lovejoy, in a welcome effort at clarification, in 1908 distinguished thirteen possible forms of pragmatism. And Schiller, in an almost intoxicating pluralistic spirit, commented that there were as many pragmatisms as there were pragmatists (at the time a considerable company). Additional confusion over pragmatism was caused by the tendency of its spokesmen to find the philosophical past well populated with pragmatists. Thus Socrates, Protagoras, Aristotle, Francis Bacon, Benedict de Spinoza, Locke, George Berkeley, David Hume, Immanuel Kant, J. S. Mill, and an assorted variety of scientists were included in the fold.

These perplexities, once hotly debated in the journals, are now only of historical interest. They need not concern us in surveying and assessing what are undoubtedly the leading ideas of pragmatism. It suffices to note the irony in the fact that while pragmatism was supposed to have made its appearance in the paper by Peirce titled "How to Make Our Ideas Clear" (1878), pragmatists continued to have so much trouble in doing so.

CHARLES SANDERS PEIRCE

What has come to be known as Peirce's pragmatism grew out of his study of the phenomenology of human thought and the uses of language. For Peirce, the investigation of thought and language—and, therefore, the way into specific studies of all kinds of claims, assertions, beliefs, and ideas—depended upon the understanding of "signs." One of Peirce's lasting ideals, resolutely pursued but never completely achieved, was to work out a general theory of signs—that is, a classification and analysis of the types of signs and sign relations and significations that, in the broadest sense, make communication possible. A sign is anything that stands for something else. While this ancient way of putting it admits of a trivial construction (signs are signs), for Peirce, the main thing was that signs are socially standardized ways in which something (a thought, word, gesture, object) refers us (a community) to something else (the interpretant—the significant effect or translation of the sign, being itself another sign). Thus, signs presuppose minds in communication with other minds, which in turn presupposes a community (of interpreters) and a system of communication.

PRAGMATIC METHOD. Put roughly, Peirce's pragmatism is a rule of procedure for promoting linguistic and conceptual clarity—successful communication—when men are faced with intellectual problems. Because the emphasis is upon method, Peirce often remarked that pragmatism is not a philosophy, a metaphysic, or a theory of truth; it is not a solution or answer to anything but a technique to help us find solutions to problems of a philosophical or scientific nature.

One of Peirce's best-known statements of the technique was in "How to Make Our Ideas Clear" (1878): "Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then our conception of these effects is the whole of

our conception of the objects." In a somewhat clearer account he said that "in order to ascertain the meaning of an intellectual conception one should consider what practical consequences might conceivably result by necessity from the truth of that conception; and the sum of these consequences will constitute the entire meaning of the conception" (*Collected Papers*, Vol. V, paragraph 9).

While Peirce often spoke of pragmatism as a method of clarifying the meaning variously of words, ideas, concepts (sometimes of objects), we can take his intended purpose to be as follows:

(1) Pragmatism is a method of clarifying and determining the meaning of signs. We must note the comprehensive status Peirce gives to signs in this connection, for example: "All thought whatsoever is a sign, and is mostly of the nature of language." The pragmatic method, however, does not apply to all the various kinds of signs and modes and purposes of communication. Peirce considered pragmatism "a method of ascertaining the meaning of hard words and abstract concepts" or, again, "a method of ascertaining the meanings, not of all ideas, but … 'intellectual concepts,' that is to say, of those upon the structure of which, arguments concerning objective fact may hinge."

(2) The aim of the method is to facilitate communication, and in particular cases, the degree to which this is accomplished determines the relevance and justification of the method. This aim takes two main forms illustrated in Peirce's writings. The first is of a critical nature: Where disputes or philosophical problems seem to have no discoverable or agreed-upon solution, pragmatism advises that words are being used in different ways or without definite meaning at all. For example, says Peirce, pragmatism will "show that almost every proposition of ontological metaphysics is either meaningless … or else … absurd." And it is in this critical capacity that Peirce remarked: "Pragmatism solves no real problem. It only shows that supposed problems are not real problems."

But the second role the method performs is much less negative: Where signs (that is, ideas, concepts, language) are unclear, the method supplies a procedure for reconstructing or explicating meanings. Here the method is directed to translating (or systematically replacing) unclear concepts with clearer ones. It is in this spirit that Peirce offered his explications of the concepts of "hardness," "weight," "force," "reality." His procedure consisted in translating and explicating a sign (a term, such as *hard*, or sentences of signs, such as "*x* is hard") by providing a conditional statement of a given situation (or class of situations) in which a definite operation will produce a definite result. Thus, to say of some object *O* that it is "hard" is to mean that "if in certain situations the operation of scratch-testing is performed on *O*, then the general result is: *O* will not be scratched by most substances." The sign (or concept) "hard" in statements asserting that some object is hard is replaceable and clarified pragmatically with a conditional statement of the sort just given. Peirce refers to this method of conditional explication of signs as a "prescription" or "precept." The conditionals are recipes informing us what we must do if we wish to find out the kind of conditions determining the meaningful use of a sign.

MEANING. For Peirce, two points are of considerable importance in the pragmatic procedure for determining meaning, (*a*) Where one cannot provide any conditional translation for a sign, its (pragmatic) meaning is empty. This is what Peirce intended by such characteristic statements as that our conception of an object is our conception of its "practical effects" or "sensible effects." He did not mean (as James sometimes did) that the meaning of a concept is the practical effect it has in particular cases when you use it. All Peirce argued was that a concept must have some conceivable consequences, or "practical bearings," and that these must be specifiable in the manner just discussed if the concept is to play a significant role in communication, (*b*) Peirce's pragmatism thus is offered as a schema for getting at the meaning, or empirical significance, of language. As a schema it is not a theory of meaning in the sense of some general definition of meaning; it is a theoretical device for getting at the empirically significant content of concepts by determining the roles they play in classes of empirically verifiable statements. This procedure, or schema, clearly foreshadowed the later programs of operationalism and the verifiability theory of meaning.

Despite some serious difficulties that jeopardize portions of Peirce's method, the general aspects of his approach appear to be sound canons of scientific practice. Peirce's recondite statements of pragmatism have created considerable confusion. But Peirce seemed less concerned with the problem of providing an accurate and complete statement of the "maxim" of pragmatism than with its use and justification. This he attempted to show in much of his later philosophical inquiries of a scientific and metaphysical sort.

Peirce's schema, or prescriptive method, for "determining the meaning of intellectual concepts" has several sources in addition to his familiarity with scientific technique. Suggestions of it are to be found in Berkeley and in Kant. Peirce's view that meanings take a general form expressed in schema or formulas that prescribe kinds of operations and results and conceivable consequences and rules of action was directly linked to Kant. Peirce says he was led to the method of pragmatism by reflecting on Kant's *Critique of Pure Reason* and on the Kantian use of *pragmatisch* for empirical, or experimentally conditioned, laws, "based on and applying to experience."

INQUIRY AND TRUTH. It should be noted, finally, that Peirce's pragmatism is part of a more general account of "inquiry," aspects of which he elaborated with some care and most of which was taken up into Dewey's extensive construction of a theory of inquiry. Peirce described the function of thought as a form of behavior initiated by the irritation of doubt and proceeding to some resolution in a state of belief. Belief is a condition of organic stability and intellectual satisfaction, but these latter do not determine the truth of beliefs. Peirce outlined a scientific and pragmatic method of clarifying and justifying belief. It was this aspect of Peirce's analysis of inquiry and belief that suggested a pragmatic theory of truth. On this matter he was unclear and wavering. Sometimes truth and pragmatic meaning overlapped or coalesced in his discussions of them. But Peirce also argued that truth theory and pragmatism are entirely separate considerations. Generally, the idea of truth, for Peirce, is drawn from Kant and is to be understood as a regulative idea, one that functions solely to order, integrate, and promote inquiry. Taken as a "correspondence" or "coherence" theory—or criticized from the point of view of such theories—Peirce's account of truth looks strange, cumbersome, and naive.

WILLIAM JAMES

It was James who launched pragmatism as a new philosophy in a lecture "Philosophical Conceptions" in 1898; it was under his leadership that pragmatism came to be famous; and it was primarily his exposition that was received and read by the world at large.

Although Peirce and James were lifelong friends and exerted much intellectual influence upon each other, they differed in ways that had important effects upon their respective versions of pragmatism. Peirce was a realist (calling himself a scholastic realist); James was far more of a nominalist. Where Peirce sought meaning in general concepts and formulas of action, James sought meaning in experienced facts and plans of action. James looked to the concrete, immediate, practical level of experience as the testing ground of our intellectual efforts; for Peirce, the immediate sensory experience is all but destitute of "intellectual purport." Furthermore, while Peirce's pragmatism took a logical and scientific character, James, despite being an eminent man of science, was first and foremost a moralist in his pragmatism.

VALUE. Moral interests and moral language appear in almost every important passage of James's writing on pragmatism. In *Pragmatism* James made his moral conception of philosophy unmistakably evident in saying that "the whole function of philosophy ought to be to find out what definite difference it will make to you and me, at definite instants of our life, if this world-formula or that world-formula be the true one." The phrase "what definite difference … at definite instants of our life" is by and large James's way of critically judging the meaning and truth of ideas. For James, meaning and truth are included in a more fundamental category of value; to determine the meaning or truth of ideas one must evaluate their "practical consequences," "usefulness," "workability." In several famous pronouncements, James spoke of truth as what is good or expedient in our beliefs. In a phrase that permanently shocked some of his readers, James described the meaning and truth of ideas as their "cash value."

Generally, for James, the function of thought is that of assisting us to achieve and sustain "satisfactory relations with our surroundings." The value of ideas, beliefs, and conceptual dealings is to be determined accordingly, on each of numerous occasions, by their effectiveness and efficiency as the means of carrying us propitiously "from any one part of our experience to any other part, linking things satisfactorily, working securely, simplifying, saving labor."

James was thus primarily concerned with issues of belief and conceptual renditions of experience in their role of enabling men to deal with environments and to enrich the fare of daily experience. It is the level of life experience that interested James. Hence, his own statements of pragmatism resemble those of Peirce but emphasize the importance of immediate experience and practical consequences and clues to action. For James, our thoughts of an object pragmatically considered lead us to "what conceivable effects of a practical kind the object may involve—what sensations we are to expect from it, and what reactions we must prepare. Our con-

ception of these effects, whether immediate or remote, is then for us the whole of our conception of the object." If we compare this statement from *Pragmatism* with those cited earlier from Peirce, it is not difficult to see that in James's pragmatism the emphasis is upon the way individuals interpret environing conditions for purposes of successful action. The passage also reflects how James's view differed from Peirce's Kantian conception; James explained "pragmatism" as coming from the Greek *πράγμα*, meaning "practice," "action." Indeed, so fundamental are action, exploration, and life experience in James's philosophy that some of his critics have taken great pains to demonstrate the value of inaction and the general uselessness of philosophy. In this endeavor, it may be said, they have been on the whole successful.

BELIEF. It was James's conception of truth that became a *cause célèbre* for pragmatism and its critics, until eventually James, tiring of the matter, turned his attention to other philosophical pursuits, leaving to Dewey the defense and development of pragmatism. Aside from truth, the other major critical issue in pragmatism was James's argument for the justification of moral and religious belief. James's interest in the meaning and function of belief was that of a skilled and perceptive psychologist and moralist. His general view was this: When, for a given person *P*, a belief *B* answers or satisfies a compelling need (of *P* to see or interpret the world in a certain way), the "vital good" supplied by *B* in the life of *P* (the difference it makes as a beneficial causal condition in the psychological and physiological behavior of *P*) justifies *B*. It must be noted that James argued for this justification procedure only when (*a*) the choice of *B* or not-*B* is, for a given individual at a given time, "live," "forced," and "momentous"; (*b*) the evidence for or against *B* is equal, or admits of no rational adjudication of one over the other; (*c*) the effect or consequences of *B* are a "vital benefit." These three qualifications work against ascribing to James some popular defense or universal apologia for religious belief. He thought he was correct in pointing to a psychological and moral right to belief analogous to the justification of postulates or posits (in Kantian and Fichtean transcendental philosophizing) or of certain theoretical hypotheses in science.

Peirce and Dewey, among others, were highly critical of this defense of the will to believe. James the psychologist and literary artist brilliantly described the working consequences of types of religious belief for characteristic types of persons. But James the philosopher tended to confuse a descriptive analysis of how belief functions and why men believe with questions of the evaluation or verification of specific cases of belief. (Thus, for example, the fact that *B* answers a need of *P* is not of itself evidence that the content of belief *B* is warranted or that *P* has correctly understood his "need.")

However, it was this side of James that was enthusiastically received as the moral core of his pragmatism by Schiller in England and Giovanni Papini in Italy. Here also James's views have affinities with those of Bergson, Vaihinger, and Simmel. James seemed to be a democratic, energetic, and lovable Johann Gottlieb Fichte, an artist and scientist exhorting men to trust their beliefs and, above all, to leave the classroom and cloister and start living and acting in the world.

JOHN DEWEY

In the article "The Development of American Pragmatism," Dewey described Peirce's views as stemming from an "experimental, not a priori, explanation of Kant" and James's pragmatism as inspired by British empiricism. But he also noted this difference: "Peirce wrote as a logician and James as a humanist." There was, in fact, a cross-fertilization of these strains; but the characterization is apt and traceable enough in the history of pragmatism and in Dewey, too, to be of expository aid. Dewey began to appreciate James while still under the influence of Hegelian and Kantian idealism; later he recognized the importance of Peirce, whose insights and ideas were in many respects anticipations of those Dewey had started to work out on his own. The Hegelian synthesis of the logical and humanistic sides of pragmatism was achieved by the disenchanted Hegelian Dewey.

INSTRUMENTALISM. Through Dewey's patient, critical, and indefatigable efforts, pragmatism was carefully and thoroughly reformulated into what Dewey called Instrumentalism, "a theory of the general forms of conception and reasoning." Instrumentalism was a single philosophical theory within which the two evolving aspects of pragmatism found coherent expression. Instrumentalism was both theory of logic and a guiding principle of ethical analysis and criticism. For Dewey, this theory bridged the most persistent and noxious of "dualisms" in modern thought—the separation of science and values, knowledge and morals.

Instrumentalism was Dewey's theory of the conditions under which reasoning occurs and of the forms, or controlling operations, that are characteristic of thought in establishing future consequences. In the paper cited above, Dewey wrote:

> Instrumentalism is an attempt to constitute a precise logical theory of concepts, of judgments and inferences in their various forms, by considering primarily how thought functions in the experimental determinations of future consequences … it attempts to establish universally recognized distinctions and rules of logic by deriving them from the reconstructive or mediative function ascribed to reason. It aims to constitute a theory of the general forms of conception and reasoning.

A suggestive and vital feature of this theory for Dewey was that while the subject matters of scientific inquiry and moral and social experience differ, the method and forms of thought functioning "in the experimental determinations of future consequences" do not differ in kind. The method of thought and the forms of reflective behavior exhibit a common functional pattern whenever problematic situations become resolved through inquiry yielding "warranted assertion."

INQUIRY AND TRUTH. "Warranted assertion" is the term for Dewey's version of truth. Inquiry is initiated in conditions of doubt; it terminates in the establishment of conditions in which doubt is no longer needed or felt. It is this settling of conditions of doubt, a settlement produced and warranted by inquiry, which distinguishes the warranted assertion. Whereas Dewey once defined "truth" as the "working" or "satisfactory" or "verified" idea or hypothesis, he was led, later—partly as a result of several critical controversies over truth with Bertrand Russell during the 1930s and 1940s—to restate his view of truth as warranted assertion.

In his *Logic* Dewey gave his general definition of inquiry as "the controlled or directed transformation of an indeterminate situation into one that is so determinate in its constituent distinctions and relations as to convert the elements of the original situation into a unified whole." The theory of inquiry was developed over many years and in many writings; into it went the products of Dewey's reflections on the nature of thought, his contributions to psychology and education, the influence of the biological and functional aspects of James's *Principles of Psychology*, and the influence of Peirce on the nature of scientific inquiry. In his analysis of the biological and cultural conditions of inquiry and in his account of intelligence as a function of these interacting conditions in a particular situation with respect to a problem and its outcome, Dewey was also guided by some of the basic ideas in the philosophical social psychology of G. H. Mead, once Dewey's colleague at Michigan and Chicago and one of his closest friends. The definitive statement of the theory is in Dewey's *Logic: The Theory of Inquiry* (1938).

For Dewey, the theory of inquiry is a generalized description of the organic, cultural, and formal conditions of intelligent action. Such action is provoked by problems of diverse kinds—political, ethical, scientific, and aesthetic. But irrespective of the specific content of human problems or the nature of problem situations, inquiry is a reflective evaluation of existing conditions—of shortcomings and possibilities—with respect to operations intended to actualize certain potentialities of the situation so as to resolve what was doubtful. The purpose of inquiry is to create goods, satisfactions, solutions, and integration in what was initially a wanting, discordant, troubled, and problematic situation. In this respect all intelligence is evaluative, and no separation of moral, scientific, practical, or theoretical experience is to be made. So commanding an achievement was Dewey's last-mentioned work that "pragmatism" is often identified with the position he expounded there as a naturalistic logic for evaluating and reconstructing human experience.

MORE RECENT TENDENCIES

A somewhat different articulation of pragmatism, deriving less from James and Dewey than from Peirce, was set forth by C. I. Lewis in the 1920s as "conceptualistic pragmatism." Lewis emphasized the role of mind in supplying the a priori principles and categories by which we proceed to organize and interpret sense experience. But he also stressed the plurality of categories and conceptual schemes by which experience can be interpreted and the evolutionary character of our systems. Because a priori principles impose no necessary order on the world or upon sense experience (determining only our ways of organizing experience), Lewis argued for a "pragmatic a priori." Decisions to accept or reject conceptual principles, indeed the very function of those principles, rest upon socially shared needs and purposes and upon our interest in increased understanding and control over experience. According to Lewis (in *Mind and the World Order*), "The interpretation of experience must always be in terms of categories … and concepts which the mind itself determines. There may be alternative conceptual systems giving rise to alternative descriptions of experience, which are equally objective and equally valid.… When this is so, choice will be determined, consciously or unconsciously, on pragmatic grounds."

Lewis's pragmatism resulted in a theory of conceptual and empirical meaning and in an analysis of empiri-

cal judgments as probable and evaluative modes of acting upon passing and future experience.

In more recent literature, under the influence of Dewey and Lewis as well as Rudolf Carnap, Charles Morris, Ernest Nagel, Willard Van Orman Quine, and others, "pragmatism" connotes one broad philosophical attitude toward our conceptualization of experience: Theorizing over experience is, as a whole and in detail, fundamentally motivated and justified by conditions of efficacy and utility in serving our various aims and needs. The ways in which experience is apprehended, systematized, and anticipated may be many. Here pragmatism counsels tolerance and pluralism. But, aside from aesthetic and intrinsic interests, all theorizing is subject to the critical objective of maximum usefulness in serving our needs: Our critical decisions, in general, will be pragmatic, granted that in particular cases decisions over what is most useful or needed in our rational endeavors are relative to some given point of view and purposes.

An expression of this attitude that is of current interest was advanced by Peirce, James, and Dewey, as well as by F. P. Ramsey, the brilliant English philosopher influenced by Peirce and James. This is an interpretation of the laws and theories of science as "leading principles," or instrumental procedures, for inferring stated conditions from others. Construed as leading principles, theories function as guides for logical inference, indicating how certain formulations are to be derived from other formulations of events, rather than as descriptively true statements of reality serving as premises from which conclusions are deduced. Pragmatically, theories are inference policies, neither true nor false (except pragmatically) but nonetheless critically assessable as to their utility and clarity and the fruitfulness of the consequences that result from adopting them.

While there continues to be an interest in the philosophies of Peirce, James, Dewey, and Schiller, pragmatism as a movement, in the form outlined in these pages, cannot be said to be alive today. But pragmatism has succeeded in its critical reaction to the nineteenth-century philosophical background from which it emerged; it has helped shape the modern conception of philosophy as a way of investigating problems and clarifying communication rather than as a fixed system of ultimate answers and great truths. And in this alteration of the philosophical scene, some of the positive suggestions of pragmatism have been disseminated into current intellectual life as practices freely adopted and taken for granted to an extent that no longer calls for special notice.

The measure of success pragmatism has achieved in encouraging more successful philosophizing in our time is, by its own standards, its chief justification. To have disappeared as a special thesis by becoming infused in the normal and habitual practices of intelligent inquiry and conduct is surely the pragmatic value of pragmatism.

See also Pragmatics; Pragmatist Epistemology.

Bibliography

GENERAL WORKS ON PRAGMATISM

Ayer, A. J. *The Origins of Pragmatism: Studies in the Philosophy of Charles Sanders Peirce and William James*. San Francisco: Freeman, Cooper, 1968.

Dewey, John. "The Development of American Pragmatism." In his *Philosophy and Civilization*. New York: Minton Balch, 1931. Ch. 2.

Murphy, J. P. *Pragmatism: From Peirce to Davidson*. Boulder, CO: Westview Press, 1990.

Quine, W. V. O. "Two Dogmas of Empiricism" In his *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1951.

Rorty, Richard. *Objectivity, Relativism, and Truth*. Cambridge, U.K.: Cambridge University Press, 1991.

Rorty, Richard. "A Pragmatist View of Contemporary Analytic Philosophy." *Utopia y Praxis Latinoamericana* 7 (2002): 29–40.

Scheffler, I. *Four Pragmatists: A Critical Introduction to Peirce, James, Mead, and Dewey*. London: Routledge and Kegan Paul, 1974.

Schneider, H. W. *History of American Philosophy*. New York: Columbia University Press, 1946. Chs. 39–41.

Thayer, H. S. *Meaning and Action: A Critical History of Pragmatism*. Indianapolis: Bobbs-Merrill, 1968.

Thayer, H. S. "Pragmatism." In *A Critical History of Western Philosophy*, edited by D. J. O'Connor, 437–462. New York: Free Press of Glencoe, 1964.

Wiener, Philip P. *Evolution and the Founders of Pragmatism*. Cambridge, MA: Harvard University Press, 1949.

WORKS BY PEIRCE

Collected Papers. 8 vols, edited by C. Hartshorne, P. Weiss, and A. W. Burks. Cambridge, MA: Harvard University Press, 1931–1958. Vols. V and VIII especially.

Charles S. Peirce: Selected Writings, edited by P. P. Weiner. New York: Dover, 1966.

WORKS ON PEIRCE

Buchler, Justus. *Charles Peirce's Empiricism*. New York: Harcourt Brace, 1939.

Gallie, W. B. *Peirce and Pragmatism*. Harmondsworth, U.K.: Penguin, 1952.

Murphey, Murray G. *The Development of Peirce's Philosophy*. Cambridge, MA: Harvard University Press, 1961.

Rosenberg, Jay F. "How not to Misunderstand Peirce—A Pragmatist's Account of Truth." In *What Is Truth?*, edited by Richard Schantz. Berlin: de Gruyter, 2002.

WORKS BY JAMES

"Philosophical Conceptions and Practical Results." University of California *Chronicle* (1898). Reprinted in his *Collected Essays and Reviews*, 406–437. New York: Longman, 1920.

Pragmatism and the Meaning of Truth. Cambridge, MA: Harvard University Press, 1979.

The Will to Believe and Other Essays in Popular Philosophy. Cambridge, MA: Harvard University Press, 1979.

WORKS ON JAMES

Myers, G. E. *William James: His Life and Thought*. New Haven, CT: Yale University Press, 1986.

Perry, R. B. *The Thought and Character of William James.* 2 vols. Boston: Little Brown, 1935.

WORKS BY DEWEY

Democracy and Education. New York: Macmillan, 1916.

Essays in Experimental Logic. Chicago: University of Chicago Press, 1916.

Reconstruction in Philosophy. New York: Holt, 1920. Reprinted in paperback edition with a new introduction. New York, 1950.

Human Nature and Conduct. New York: Holt, 1922.

Logic: The Theory of Inquiry. New York: Holt, 1938.

The Early Works, the Middle Works, the Later Works. Carbondale: Southern Illinois University Press, 1990.

WORKS ON DEWEY

Geiger, G. R. *John Dewey in Perspective*. New York: Oxford University Press, 1958.

Hook, Sidney. *John Dewey*. New York: Day, 1939.

Thayer, H. S. *The Logic of Pragmatism*. New York: Humanities Press, 1952.

White, Morton. *The Origin of Dewey's Instrumentalism*. New York: Columbia University Press, 1943.

OTHER PRAGMATISTS

Abel, Reuben. *The Pragmatic Humanism of F. C. S. Schiller*. New York: King's Crown Press, Columbia University, 1955.

Lewis, C. I. *An Analysis of Knowledge and Valuation*. La Salle, IL: Open Court, 1946.

Lewis, C. I. *Mind and the World Order*. New York: Scribners, 1929.

Lewis, C. I. "A Pragmatist Conception of the A Priori." *Journal of Philosophy* 20 (1923): 169–177.

Mead, George H. *Mind, Self, and Society*. Edited by C. W. Morris. Chicago: University of Chicago Press, 1934.

Mead, George H. *The Philosophy of the Act*. Chicago: University of Chicago Press, 1938.

Mead, George H. *Selected Writings*. Edited by Andrew J. Reck. Indianapolis: Bobbs-Merrill, 1964.

Schiller, F. C. S. "Axioms as Postulates." In *Personal Idealism*, edited by Henry Sturt, 47–133. London: Macmillan, 1902.

Schiller, F. C. S. *Logic for Use*. London: G. Bell, 1929.

Schiller, F. C. S. *Studies in Humanism*. New York: Macmillan, 1907.

H. S. Thayer (1967)

Bibliography updated by Benjamin Fiedor (2005)

PRAGMATISM [ADDENDUM]

Not unexpectedly, given that "pragmatism" is not a doctrine but a method (as Charles Sanders Peirce put it), the tradition of classical pragmatism is formidably diverse. Even the method—the pragmatic maxim—is differently interpreted by different pragmatists; and this diversity is compounded by the different doctrines and interests of the various pragmatists. But there is a pattern discernible within the diversity: a shift from Peirce's reformist, scientific philosophy, anchored by his realism about natural kinds and laws and about the objects of perception, through William James's more nominalist pragmatism, his insistence that "the trail of the human serpent is over everything" (1907, p. 37), through John Dewey's proposal that the concept of warranted assertibility replace the concept of truth, to the radicalism of Ferdinand Canning Scott Schiller's avowedly Protagorean relativization of truth to human interests.

Contemporary pragmatisms are no less diverse, but the spectrum has shifted to the left. The more conservative neopragmatists are as akin to James as to Peirce, and the most radical go beyond Schiller's relativism to an antiphilosophical, sometimes antiscientific, even anti-intellectual, stance—a stance so much at odds with the aspirations of the founders of pragmatism as to put one in mind of Peirce's complaints about writers who persisted in "twisting [the pragmatists'] purpose and purport all awry" (*Collected Papers*, 5.464).

Nicholas Rescher describes his philosophy as pragmatic idealism: idealism, because it holds that "reality … as humans deal with it is *our* reality—our thought-world as we conceive and model it" (1994, p. 377); pragmatic, because it holds that, though our picture of reality is a mental construction, it is not a free construction but is objectively constrained by success or failure in practice, in prediction and attainment of purpose.

In some ways—not least in philosophizing unapologetically in the grand systematic manner—Rescher is much like Peirce; indeed, his conception of the interlocking cognitive, evaluative, and practical aspects of rationality takes him further than Peirce into some of the territory of value theory. In other ways Rescher's pragmatism is more reminiscent of James: inter alia, for its stress on practical consequences and on a pluralism of perspectival truth-claims. So, too, is his idealism. Qua pragmatist Peirce denies the intelligibility of the in-principle-incognizable: Qua "objective idealist" he maintains that "matter is just effete mind" (*Collected Papers*, 6.25).

Rescher's idealism sounds more like the Jamesian serpent—or Deweyan interactionism.

In repudiating metaphysical realism and endorsing internal realism, Hilary Putnam evinced some sympathy with Peircean conceptions of truth and reality. But his conceptual relativism—"Our language cannot be divided into two parts, a part that describes the world 'as it is anyway' and a part that describes our conceptual contribution" (1992, p. 123)—sounded more like James. However, his argument against the irrealism of Nelson Goodman (himself classifiable as a left-wing Jamesian of the boldest nominalist stripe) stressed the distinction between wholly conventional names such as "Sirius" and only partially conventional general terms such as *star*. Putnam thus recalled Peirce's realism of natural kinds, and perhaps divided our language after all. It is not surprising, then, to find that most recently, in his Dewey lectures, he tends to a more realist stance.

Sympathetic in the 1950s and 1960s to the positivists' aspiration to a scientific single theory that explains everything" (Putnam 1992, p. 2), Putnam is since then inclined to a pluralistic, problem-centered approach to philosophy. Here, as in his defense of democracy as a precondition for the application of intelligence to the solution of social problems, he acknowledges Dewey.

A year before the publication of W. V. O. Quine's "Two Dogmas," Morton G. White had invoked Dewey in describing the analytic-synthetic distinction as "an untenable dualism." Rejecting that distinction, adopting a holism of verification, insisting on the underdetermination of theory by data, Quine describes himself as going beyond C. I. Lewis's pragmatic a priori to a "more thorough pragmatism" that emphasizes pragmatic considerations in theory-choice generally. "Pragmatic" here suggests the relatively unconstrained rather than, as in Rescher, a kind of constraint. Quine refers approvingly to Schiller's view of truth as manmade as one of pragmatism's main contributions to empiricism. But he hopes to avoid Schiller's relativism by means of a naturalism that views philosophy as internal to science. This differs significantly from Peirce's and Dewey's aspiration to make philosophy scientific by applying the method of science to philosophical questions.

As another of pragmatism's main contributions Quine mentions Peirce's and Dewey's connecting belief and meaning to behavior. But Quine's behaviorism is more stringent, in part because of the influence of B. F. Skinner, and in part because Quine's extensionalism leaves him uneasy, as Peirce was not, with any irreducibly dispositional talk.

As Putnam's allusions to the existentialist character of James's ethics indicate, some hope a neopragmatism might heal the analytic-Continental rift. One example is Karl-Otto Apel's grafting of pragmatic elements from Peirce and Jürgen Habermas onto Alfred Tarski's semantic conception of truth. Another is Joseph Margolis's attempt, emphasizing both the biological roots and the "deep historicity" of human injury, and proposing a reconciliation of a modest realism with a weak relativism, to marry themes from Peirce with themes from Martin Heidegger.

Richard Rorty describes himself as accommodating themes from Dewey with themes from Heidegger. Maintaining that "revolutionary movements within an intellectual discipline require a revisionist history of that discipline" (1983, p. xvii), Rorty dismisses Peirce as having merely given pragmatism its name. And he urges in the name of pragmatism that the project of a philosophical theory of knowledge should be abandoned; that science is exemplary only as a model of human solidarity; that philosophy is more akin to literature than to science: that it should be in the service of democratic politics; that truth is "not the kind of thing one should expect to have a philosophically interesting theory about" (1983, p. xiv) and that to call a statement true is just to give it "a rhetorical pat on the back" (1983, p. xvii); that pragmatism is antirepresentationalism.

There is some affinity between Rorty and Schiller. But Peirce, who was a pioneer of the theory of signs, of representation, and who desired "to rescue the good ship Philosophy for the service of Science from the hands of the lawless rovers of the sea of literature" (*Collected Papers*, 5.449), would disagree with Rorty's pragmatism in every particular. So too, except perhaps for his description of the best ethical writing as akin to "novels and dramas of the deeper sort" (1891, p. 316), would James. And so, most to the point, would Dewey, who hoped to renew the philosophical theory of knowledge by making it more scientific, and whose political philosophy is infused by the hope that the application of scientific methods would enable intelligent social reform, and by the conviction that a free society is a prerequisite of a flourishing science.

See also Behaviorism; Democracy; Dewey, John; Empiricism; Goodman, Nelson; Habermas, Jürgen; Heidegger, Martin; Idealism; James, William; Lewis, Clarence Irving; Naturalism; Peirce, Charles Sanders; Putnam, Hilary; Quine, Willard Van Orman; Rationality; Realism; Schiller, Ferdinand Canning Scott; Tarski, Alfred.

Bibliography

Apel, K.-O. "C. S. Peirce and the Post-Tarskian Problem of an Adequate Explication of the Meaning of Truth: Towards a Transcendental-Pragmatic Theory of Truth." In *The Relevance of Charles Peirce*, edited by E. Freeman. La Salle, IL: Open Court, 1983.

Bernstein, R. "The Resurgence of Pragmatism." *Social Research* 59 (1992): 813–840.

Brandom, Robert. "Pragmatics and Pragmatisms." In *Pragmatism and Realism*, edited by Hilary Putnam. New York: Routledge, 2002.

Dewey, John. *The Early Works, the Middle Works, the Later Works*. Carbondale: Southern Illinois University Press, 1990.

Dewey, John, and J. Tufts. *Ethics*. New York: Holt, 1932.

Goodman, N. *Ways of Worldmaking*. Hassocks, U.K.: Harvester Press, 1978.

Haack, Susan. "The Relevance of Psychology to Epistemology." *Metaphilosophy* 6 (1975): 161–176.

James, William. *Essays in Radical Empiricism*. Cambridge, MA: Harvard University Press, 1976.

James, William. "The Moral Philosopher and the Moral Life." *International Journal of Ethics* 1 (1891). Reprinted in *William James: Selected Writings*, edited by G. Bird. 1994.

James, William. *Pragmatism,* edited by F. Burkhardt and F. Bowers. Cambridge, MA: Harvard University Press, 1975.

Joas, H. *G. H. Mead: A Contemporary Re-examination of His Thought*. Translated by R. Meyer. Cambridge, MA: MIT Press, 1985.

Lewis, C. I. "A Pragmatist Conception of the A Priori." *Journal of Philosophy* 20 (1923): 169–177.

Margolis, J. *Pragmatism without Foundations*. Oxford: Blackwell, 1986.

Mead, G. H. *Selected Writings: George Herbert Mead,* edited by A. J. Reck. Chicago: University of Chicago Press, 1964.

Murphy, J. P. *Pragmatism: From Peirce to Davidson*. Boulder, CO: Westview Press, 1990.

Okrent, M. *Heidegger's Pragmatism*. Ithaca, NY: Cornell University Press, 1988.

Peirce, C. S. *Collected Papers,* edited by C. Hartshorne, P. Weiss, and A. Burks. Cambridge, MA: Harvard University Press, 1931–1958.

Putnam, Hilary. *Renewing Philosophy*. Cambridge, MA: Harvard University Press, 1992.

Putnam, Hilary. "Sense. Nonsense and the Senses: An Inquiry into the Powers of the Human Mind." *Journal of Philosophy* 91 (9) (1994): 447–517.

Putnam, Hilary. *Words and Life*. Cambridge, MA: Harvard University Press, 1994. Sec. 3, "The Inheritance of Pragmatism."

Putnam, Hilary, James Conant, and Urszula Zeglen, eds. *Pragmatism and Realism*. New York: Routledge, 2002.

Quine, W. V. O. "The Pragmatists' Place in Empiricism." In *Pragmatism: Its Sources and Prospects*, edited by R. J. Mulvaney and P. M. Zeltner. Columbia: University of South Carolina Press, 1981.

Quine, W. V. O. "Two Dogmas in Retrospect." *Canadian Journal of Philosophy* 21 (1991): 265–274.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43. Reprinted in his *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1951.

Rescher, Nicholas. *Pluralism: Against the Demand for Consensus*. New York: Oxford University Press, 1993.

Rescher, Nicholas. *A System of Pragmatic Idealism.* 3 vols. Princeton, NJ: Princeton University Press, 1992–1994.

Rescher, Nicholas et al. *Philosophy and Phenomenological Research* 54 (1994): 377–457.

Rorty, Richard. *The Consequences of Pragmatism*. Hassocks, U.K., 1983.

Rorty, Richard. *Objectivity, Relativism, and Truth*. Cambridge, U.K.: Cambridge University Press, 1991.

Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." In his *Science, Perception, and Reality*. London: Routledge and Kegan Paul, 1956.

Stich, Stephen. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press, 1990.

Thayer, H. S. *Meaning and Action: A Critical History of Pragmatism*. Indianapolis: Bobbs-Merrill, 1968.

White, M. G. "The Analytic and the Synthetic: An Untenable Dualism." In *John Dewey, Philosopher of Science and Freedom*, edited by Sidney Hook. New York: Dial Press, 1950.

Susan Haack (1996)
Bibliography updated by Benjamin Fiedor (2005)

PRAGMATIST EPISTEMOLOGY

William James's observation that "when … we give up the doctrine of objective certitude, we do not thereby give up the quest or hope of truth itself" (1956, p. 17) succinctly expresses one important epistemological theme of traditional pragmatism: accommodation of a thoroughgoing fallibilism with a modest optimism about the possibility of successful truth seeking. Also characteristic of that tradition is its naturalism, its acknowledgment of the biological, and the social as well as the logical elements in the theory of knowledge, and its respect for science as, in Charles Peirce's words, "the epitome of man's intellectual development" (*Collected Papers*, 7.49). Since 1968 these ideas have been variously worked out by some who are fully aware of their roots in pragmatism and have also entered the thinking of many who are not. More surprising, some self-styled neopragmatists defend epistemological positions (or antiepistemological positions) quite unlike these classically pragmatist themes.

Both fallibilism and naturalism are prominent themes in W. V. O. Quine's epistemology, themes of which he acknowledges the pragmatist ancestry; his fallibilism, furthermore, like Peirce's, extends to mathematics and

logic, and his naturalism, like Peirce's, has an evolutionary character. And he shares the pragmatists' regard for science. However, he seems drawn beyond a view of epistemology as resting in part on empirical assumptions about human cognitive capacities to conceiving of it as internal to the sciences of cognition; and thence, under pressure of the implausibility of supposing that psychology or biology could answer the questions about evidence, justification, and so forth, with which epistemology has traditionally been concerned, he seems drawn to a revolutionary scientism that would abandon the traditional questions in favor of questions the sciences can be expected to answer. Unlike his fallibilism and his modest, reformist naturalism, neither his scientism nor his revolutionary displacement of epistemology falls within the tradition of pragmatism.

Nicholas Rescher's approach, from its insistence that we humans "cannot function, let alone thrive, without knowledge of what goes on around us" (1994, p. 380) to its stress on the provisional, tentative character of all our estimates of truth, is unambivalently within the pragmatist tradition. But Rescher takes issue with Peirce's definition of truth, and therefore conceives of progress in terms of improvement over earlier stages rather than closeness to a supposed final stage.

Focusing on criteria of evidence and justification rather than on guidelines for the conduct of inquiry, Susan Haack adapts from the pragmatist tradition: Her fallibilism, expressed in the thesis that justification comes in degrees; her weak, reformist naturalism, expressed in the thesis that our criteria of evidence have built into them empirical presuppositions about human cognitive capacities; her account of perception; and her strategy for the metajustification of criteria of justification.

In stark contrast to Rescher or Haack, Richard Rorty urges in the name of pragmatism that the philosophical theory of knowledge is misconceived; and, in contrast to Quine, that epistemology should be, not replaced by the psychology of cognition, but simply abandoned. Rorty likens his repudiation of epistemology to John Dewey's critique of the "spectator theory." What Dewey intended, however, was to reform epistemology, to replace the quest for certain knowledge of eternal, unchanging objects with a realistic account of fallible, experimental, empirical inquiry. Rorty's revolutionary attitude derives from his conception of justification as a matter exclusively of our practices of defending and criticizing beliefs, not grounded in any connection of evidence and truth. This "conversationalist" conception of justification is motivated by his rejection of any conception of truth as meaning more than "what you can defend against all comers."

Often accused of relativism, Rorty denies the charge. He escapes it, however, only by shifting from contextualism ("*A* is justified in believing that *p* iff (if and only if) he can defend *p* by the standards of *his* community") to tribalism (" … iff he can defend *p* by the standards of *our* community" [1979, p. 308]). But tribalism is arbitrary if our practices of criticizing and defending beliefs are, as Rorty holds, not grounded in any connection of evidence and truth.

In not-so-stark contrast to Rorty, Stephen Stich (1990) urges in the name of pragmatism that it is mere epistemic chauvinism to care whether one's beliefs are true, and that justified beliefs are those that conduce to whatever the subject values. True, Stich cheerfully embraces relativism (and rejects tribalism since he thinks our epistemic practices too preoccupied with truth); and he looks to the sciences of cognition to help us "improve" our cognitive processing so as better to achieve what we really value. But, as more overtly in Rorty, the effect is profoundly antiepistemological and "pragmatist" in quite another sense than the traditional one.

See also Dewey, John; Epistemology; James, William; Naturalism; Peirce, Charles Sanders; Quine, Willard Van Orman.

Bibliography

Haack, S. *Evidence and Inquiry*. Oxford: Blackwell, 1993.

James, W. *The Will to Believe* (1897). New York, 1956.

Peirce, C. S. *Collected Papers,* edited by C. Hawthorne, P. Weiss, and A. Burks. Cambridge, MA: Harvard University Press, 1931–1958.

Quine, W. V. O. "Epistemology Naturalized." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1967.

Quine, W. V. O. "Natural Kinds." In *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1967.

Quine, W. V. O. "The Pragmatists' Place in Empiricism." In *Pragmatism: Its Sources and Prospects*, edited by R. J. Mulvaney and P. M. Zeltner, 21–40. Columbia: University of South Carolina Press, 1981.

Quine, W. V. O. *Pursuit of Truth*. Oxford: Oxford University Press, 1990.

Rescher, N. "Précis of *A System of Pragmatic Idealism*," *Philosophy and Phenomenological Research* 54 (1994): 377–390.

Rescher, N. *A System of Pragmatic Idealism.* Vol. 1: *Human Knowledge in Idealistic Perspective*. Princeton, NJ: Princeton University Press, 1992.

Rorty, Richard. *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press, 1982.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Stich, S. P. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Susan Haack (1996)
Bibliography updated by Benjamin Fiedor (2005)

PRECOGNITION

Etymologically, *precognition* is simply the Latin equivalent of *foreknowledge*. But it has come to have a more specialized meaning as a semitechnical term for one of the phenomena or putative phenomena of parapsychology (psychical research). This entry touches on the wider issues of foreknowledge only insofar as they appear in a rather special form in the narrower context of parapsychology. Again, since the philosophical problems centering on some of the other concepts of parapsychology are examined at length elsewhere, telepathy, clairvoyance, and psychokinesis are mentioned here only when necessary to the main goal of becoming clearer about the logical geography of parapsychological precognition. Nor will there be any discussion of what the facts actually are. We shall be concerned only with theoretical questions of implication and explanation.

Precognition is one of a group of terms that also includes *telepathy*, *clairvoyance*, and—more peripherally—*psychokinesis* (PK). Telepathy is thought of, initially at any rate, as consisting in the acquisition of information by one person from another without the use of any of the senses normally indispensable to communication. Clairvoyance, at the same initial stage, is conceived of as being generically identical with telepathy; the specific difference is that in the case of clairvoyance the information is supposed to be obtained not from another person but from an object. Telepathy would be termed "precognitive" if the information so acquired was not going to become available to the other person until later. Clairvoyance would be termed precognitive if the information so acquired was not, until later, even going to become available in things, as opposed to minds.

It is thus possible to consider precognitive telepathy and precognitive clairvoyance as being two species of the genus precognition. Straight telepathy, straight clairvoyance, and both sorts of precognition are all supposed to be both nonsensory and noninferential. It is partly for this reason that all these alleged phenomena are frequently classed together as varieties of extrasensory perception (ESP). It is important to recognize that both these negative characteristics are in all four cases defining. To show that the information was acquired by the use, whether conscious or unconscious, of sensory cues, clues, or signs is a sufficient reason for disqualifying as genuine telepathy, or what have you, any ostensible case of telepathy or other such phenomenon. Similarly, to show that this acquisition was the result of a feat of inference, however heroic and remarkable in itself, again constitutes a completely sufficient reason for insisting that we are not confronted with a genuine case of precognition. At most we must describe it as a pseudo precognition, "precognition" only in quotation marks.

Suppose someone has an intuition or a dream or a waking vision that is found to correspond to some actual later happening. Suppose that it seems out of the question either (1) to account for the correspondence as the result of successful inference, conscious or unconscious, from materials available to the subject at the time, or (2) to trace it back to some causal ancestor common to both the "anticipation" and the "fulfillment," or (3) to say that the "fulfillment" was somehow a result of the "anticipation," or (4) even to refuse to account for the correspondence in any way on the grounds that it was just a coincidence. (The counterargument in this last case would be that some intuitions, dreams, visions, and so forth, are bound to prove veridical and that presumably this was just one of those striking cases that is—as the catch phrase has it—"by the law of averages" bound to occur occasionally.) If such an intuition, or what have you, were to occur we would—provided that all four conditions seemed to be met—have at least a prima facie case of precognition. Three theoretical questions must then be considered.

OPERATIONAL DISTINCTIONS

The first question is whether there are real operational distinctions to be made between all the supposed varieties of ESP or whether any of them can be regarded as alternative descriptions of the same logically possible phenomena. For instance, some ingenuity is required to work out an experimental design that would enable us to distinguish decisively between straight clairvoyance and precognitive telepathy.

To make this clear, consider a stylized ESP experiment. The experimenter equips himself with a pack of cards, perhaps the special Zener type, which consists of five suits of five identical cards. He devises a procedure for randomizing the order in which the cards are to be offered as targets. He recruits a subject whose function is to guess the values of the cards chosen as targets. The

experimenter takes drastic and thorough precautions to ensure that it is quite impossible for the subject to tell by any combinations of inference and sensory perception what is or is going to be the value of any target card. (This is, of course, very much more easily said than done. But here our concern is with theory only.) The subject in due course makes his guesses, and these guesses are recorded. If enough guesses are made—provided always that the experiment has been properly designed and properly executed—we should expect "by the law of averages" that when the guesses are scored against their targets about one-fifth of the total will turn out to have been right and the remaining four-fifths wrong. But if significantly more hits have been scored than this mean-chance expectation, then it seems that *some* ESP factor must have been involved.

Suppose now that the experimenter has taken care to ensure that no one at all, himself included, should know, at the time when the subject makes his guesses, what is the value of each target card. It might seem that his experimental results can be interpreted as evidence only for clairvoyance and not for telepathy. But once we have allowed the possibility of precognition, then these same results can be described equally well in terms of precognitive telepathy. The subject is perhaps precognitively "picking" the brains of whoever later does the scoring.

The problem is further complicated if one is also prepared to allow the possibility of PK. Literally, "psychokinesis" means movement by the mind. The idea is that perhaps some people sometimes may be able, whether consciously or unconsciously, to move or otherwise affect things without pushing or pulling them and, indeed, without in any way touching either the things in question or any other things involved in the process. Perhaps, it is suggested, these people or, indeed, all of us really can in some conditions bring about changes in things by simply "willing," as a gambler might wish that by simply "willing" and without any detectable cheating he could get dice to fall in the ways he desires.

Once this suggestion is allowed there seems to be room for an alternative description of many experiments that might otherwise have appeared to be unambiguous evidence of the reality of precognition. Such a description will be in terms of psychokinesis, guided perhaps by a measure of straight telepathy or straight clairvoyance. The subject may not, after all, really be precognizing the target. Perhaps he or somebody else is consciously or unconsciously influencing psychokinetically the target-determining mechanism in order to increase the degree of correspondence between the guess series and the target series. With appropriate alterations the same suggestion can be applied to spontaneous, as opposed to experimental, cases of ostensible precognition. The "fulfillment" or "fulfillments" become partly or wholly the results of the "anticipations," and, by specification, any such cases are disqualified from being classed as genuinely precognitive. Confronted by this kaleidoscopically changing confusion of alternative descriptions, we need not wonder that PK was once described as the parapsychological equivalent of a universal solvent.

IMPLICATIONS

The second sort of theoretical question concerns the implications of precognition. Suppose it were to be established that there really is such a phenomenon, which actually does satisfy all the conditions stipulated; what would follow?

THE FUTURE AS PRESENT. One consequence that has often been thought to follow from the existence of precognition is that, sensationally, the future must somehow be already here—or at any rate there. This is usually derived from a conception of precognition as a mode of perception, of ESP. Thus, J. W. Dunne, in *An Experiment with Time* (3rd ed., London, 1939, p. 7), claims that in precognition "we habitually observe events before they occur." By valid inference from this misdescription he concludes that the future must therefore really be present. Upon this absurdity he proceeds to erect his logical extravaganza "the serial theory of time." Or again, in a useful survey of the field, D. J. West remarks: "precognition—foreseeing arbitrary events in the future that could not by any stretch of the imagination be inferred from the present—that is something which is almost impossible for our minds to grasp. How can anyone see things which do not yet exist?" (*Psychical Research Today*, London, 1954, p. 104).

Now it is necessarily true that if anything is to be seen or otherwise perceived—and not just "seen" or "perceived" (in discrediting quotation marks)—that thing must be presently available. (We ignore for present purposes the peripheral problems presented by very distant stars.) West is therefore more right than perhaps he realizes in suggesting that it is inconceivable that anyone should be able to see things that do not yet exist. Nevertheless, the correct conclusion to draw is not, as some have been inclined to think, that precognition is logically impossible. The correct conclusion is, rather, that if the phenomenon specified was to occur, it could not be conceived of as any sort of perception. The argument reduces

to absurdity not the notion of precognition as such but the assumption that such precognition can be assimilated to perception. (There are indeed further reasons, applying equally to all varieties of ESP, which tend to destroy this analogy and therefore make unfortunate the use of the expression "extrasensory perception." But the present reason, applying only to precognition, is in this case by itself entirely decisive.)

PRECOGNITION AS FOREKNOWING. Suppose one begins by thinking of precognition not as foreseeing but as foreknowing. Suppose then that one happens to be one of those who conceives of cognition on the model of perception. This is, of course, a misconception, but one with a most ancient and distinguished pedigree. One relevant reason for insisting that this model is inapplicable is that whereas it is logically possible for me to know now that certain things happened in the past and that other things will happen in the future, it is not logically possible for me now to perceive anything but what is now available to be perceived. Thus, anyone who thinks of precognition as a form of knowing and of knowing as a sort of perceiving will arrive by a rather longer route at exactly the same conclusions—that the future is present—as the person who begins by thinking of precognition as a type of perception. In either case the treatment indicated is essentially the same.

C. D. Broad comments:

> The fact is that most people who have tried to theorize about non-inferential precognition have made needless difficulties for themselves by making two mistakes. In the first place, they have tried to assimilate it to sense-perception, when they ought to have assimilated it to memory. And, secondly, they have tacitly assumed an extremely naive prehensive analysis … [which] is simply nonsensical when applied to ostensible remembering or ostensible foreseeing. ("The Philosophical Implications of Foreknowledge")

By "prehensive analysis" Broad means believing, mistakenly, that for an occurrence to be remembered it must somehow be present.

FATALISM. The model of memory is, as Broad urged, much less inapt than that of perception. But it, too, has its dangers. It has beguiled some into thinking that precognition must necessarily involve fatalism. The suggestion is that precognition would be an exact analogue and complement of memory, but where memory operates backward, precognition would be remembering forward. (See, for instance, Lewis Carroll, *Through the Looking Glass*, Ch. 5.) Now, if someone remembers that he himself killed Cock Robin, and provided that he really does remember and that he is not merely claiming, mistakenly or even dishonestly, to remember that he committed this crime, then it follows necessarily that he did kill Cock Robin. But if he has done it, then he has done it, and it must now be too late for anyone to intervene to save the victim. It is, notoriously, a tautology that what is done is done and cannot be undone. The past is unalterable. The temptation is to argue that the same must, in exactly the same sense, apply to the future. If I can truly precognize that I will kill Cock Robin—provided that it really is a precognition and that I am not merely claiming mistakenly, or even dishonestly, to be precognizing—then it follows necessarily that I will kill Cock Robin.

The false step is to go on to urge that by parity of reasoning, since he will do it, then he will do it, and therefore it must now be too late for anyone to save Cock Robin. For the conclusion does not follow. From the proposition that he will kill Cock Robin we are entitled to infer that he will kill Cock Robin and hence that no one will in fact save the bird. But what we are not entitled to infer is that it must now be too late to take any steps to save Cock Robin, that no one could possibly do anything to help. It is one thing to know that some catastrophe will in fact occur; it is quite another to know that there is now nothing that anyone could do to prevent it, even if he so wished. To know that he will in fact do it, it is sufficient to know that he in fact will: tautology. It is not necessary also to know, what may very well not be the case, either that he would not have been able to do otherwise had he been going to want to or that no one else would have been able to stop him had they been going to be so inclined.

This point is, of course, involved in the much wider question of whether foreknowledge in the general sense must carry any such fatalist implications. The wider question is beyond the scope of this article, but the argument offered here is as applicable to the wider context as to this narrower one. The problem remains why it should be thought, as obviously it often is, that to establish the reality of noninferential precognition, even as an extremely weak and rare faculty, ought to raise fatalist anxieties in a much more acute form than does, for instance, the present possibility of inferring the outcome of some not too distantly future election—on the basis of a knowledge of the present preferences, psychological traits, beliefs, and expressed voting intentions of the electors concerned.

The threat to autonomy. One possible suggestion is that it may be thought that whereas predictions on the

basis of knowledge of human beings do not constitute any threat to the autonomy and dignity of the persons concerned, a precognitive forecast about someone's future actions, made without reference to his peculiar characteristics, plans, and desires, would tend to show that his decisions to act in those ways will not be as causally necessary as he might like to believe. To show that human wishes, plans, and decisions do not affect what happens would indeed be to demonstrate a fatalist conclusion; for this is precisely what "fatalism" means. But to show that someone can know, without reference to that other person's wishes and plans, what another person is going to do is, surely, not sufficient to show that those wishes and plans will not determine his course of action.

It might be argued that knowledge presupposes grounds and that, insofar as the grounds contain no reference to the wishes and plans of the agent, this shows that he cannot properly be held responsible for what he is going to do. This argument would have more force if knowledge of what is going to occur always had to be grounded on knowledge of the presence of particular causes sufficient to bring about the occurrence. But quite apart from any question of whether it is true that all knowledge must be grounded on something else, the argument must be ineffective as long as we have to allow that some knowledge is quite sufficiently grounded simply on a recognition of reliable signs. Suppose precognition does actually occur, and suppose that it is properly to be classed as a form of knowledge; then it can be only either a variety that is not grounded at all or one which is based upon just such a recognition of signs—the recognition, namely, that some particular class of guesses, intuitions, visions, or whatnot are in fact reliable pointers to the future. For any inference, whether conscious or unconscious, from any knowledge, however acquired, of the causes of what is going to happen to the true conclusion that just that is indeed going to happen must by definition disqualify that conclusion as a genuine noninferential precognition.

Perceptual model and fatalism. A second suggestion is that the special anxiety felt in this case of precognition is just one more consequence of thinking in terms of a perceptual model. If in having a precognitive experience you were, as it were, seeing the future, then indeed it would be absurd to insist, once that experience has taken place, that there are any steps that anyone could take that could prevent the fulfillment of the precognition. It would be absurd so to insist because on this assumption of a literal foreseeing, the event precognized would by now have been seen happening. But once an event has happened there cannot be anything that anyone could possibly do to prevent it from happening.

Precognitive infallibility. A third suggestion is adapted to a rather different conception of the problem. It is common enough to find people who (at any rate, in their most self-consciously philosophical moments) would be reluctant to concede that there is any such thing as real knowledge of future events, or at least of future human actions. To such a person precognition might appear to present a special problem precisely because of the analogy to memory. This might, of course, be because he naively assimilated memory to perception. But he might in a rather more complicated way be arguing that since from the occurrence of a genuine memory one is entitled to deduce that the past was as that memory represents it to have been, therefore the occurrence of an authentic precognition would, insofar as precognition is to be conceived on the model of memory, provide a similarly inexpugnable guarantee that the future must necessarily be as it is precognized to be going to be. The idea would be, presumably, that whereas inferences can be invalid and their conclusions false, memory is necessarily infallible. Thus, if precognition is a reality, and if it is a faculty exactly analogous to memory, then it, too, must be similarly infallible. In that case there can be nothing which anyone could do to prevent the fulfillment of any such precognitive anticipations.

Insofar as this claim really represents a different contention from any so far considered, and it is not altogether clear that it does, the crucial error seems to lie in a confusion between remembering and mistakenly or dishonestly claiming to remember. True memory is, if you like, infallible, but only in the weak sense that "I remember doing it" entails "I did it," not in the strong sense that "I claim to remember doing it" entails "I did it." This is because it is always possible that in making such a memory claim I may either be mistaken or be acting dishonestly. Thus, to be exactly analogous to memory, precognition would have to be infallible in this and only this sense. But this sort of infallibility pertains equally to knowledge: for "He knows that the dogmas of his Roman Catholic faith are true" entails "The dogmas of his Roman Catholic faith are true"; whereas "He claims with absolute conviction that he knows that the dogmas of his faith are true" is by itself not even evidence for "The dogmas of his faith are true." And we have already devoted enough space to urging that from the possibility of knowledge as such of future human actions no fatalist conclusions follow necessarily.

"Forward memory" and fatalism. Another, and perhaps the most important, consideration encouraging the

idea that parapsychological precognition must constitute a fatalist threat more serious than any arising from ordinary possibilities of foreknowledge is that what we remember is always and necessarily something in which somehow we ourselves were previously involved: We remember, that is, only what we have learned or what happened to us or what we did. Therefore, insofar as precognition is to be thought of as "remembering forward," its contents must be similarly restricted to what we shall later come to know by other means, to what will happen to us, or to what we will do. But now, as long as I remain the sort of creature that I am, it will clearly not be possible for me to precognize something very unpleasant as going to happen to me without my casting about for ways in which the unpleasantness may be avoided.

Hence, if there is to be precognition, at least one of three further conditions must be satisfied: Either (1) the contents of my precognitions must be restricted to terms that even in an unchanged universe would not provoke me to effective avoiding action, or (2) I as the precognizer must be so changed that I no longer attempt any avoiding action, or (3) the universe around me must be so changed that my attempts are all in fact now ineffective. Obviously both the second and the third of these options would constitute major steps towards a fatalistic universe. Yet neither of these represents a necessary corollary of precognition as such. On the other hand, to take the first option is to accept a limitation that drastically reduces the analogy between precognition and memory. The conclusion is that any fatalist consequences belong to precognition as a faculty fully analogous to memory, not simply to precognition as such.

CAUSE AND EFFECT. It has sometimes been suggested that to establish the reality of precognition would be to show that in some cases effects can precede their causes. Surprising and disturbing though the effects reported certainly are, this at least is something that neither these nor any other phenomena could ever establish. The reason is, quite simply, that "a cause must either precede or be simultaneous with its effect" is a necessary truth. It is no more possible to discover an effect preceding its cause than to light upon a bachelor husband—and the impossibility is of the same sort in both cases.

Someone who had appreciated this point might well be inclined to dismiss it as merely verbal and trifling. He might claim that nevertheless we have here some radically new and theoretically highly recalcitrant facts and that to take account of them we must revise some of our old ideas.

Not every verbal point is trifling, however, and not all matters of definition are mere matters of definition. What looks like a piece of obstructive lexicography can be justified at a deeper level. The implicit definitions to which appeal was originally made are grounded on a more fundamental necessity. We cannot simply brush off the objection by prescribing a small revision in usage whereby causes may in future be spoken of as succeeding their effects, and then proceed exactly as before. The crux is that causes are—and in principle can always be used by us as—levers for bringing about their effects. But a cause that succeeded its effect could not be, or be used as, a lever for producing it. Once the "effect" has happened it must be too late for any "cause" to bring it about—and too late also for it to be prevented by preventing the occurrence of this "cause." To make this suggested change in the usage of the terms *cause* and *effect* would be not to modify but to disrupt the concept of cause. The refusal to accept the claim that in precognition we would be confronted with causes operating backward in time may therefore spring from something less discreditable than complacency. It might even be one manifestation of a conviction that to accommodate such a phenomenon we should need something much more radical and much more ratiocinative than a paradoxical but really not particularly significant set of adjustments in the usage of one or two common terms.

POSSIBLE EXPLANATIONS

The third kind of theoretical question about precognition is "What sort of explanation or account could we hope to find, supposing it were to be definitely established that precognition does indeed occur?" Presumably this would have to cover whatever other parapsychological phenomena were also found to be genuine. To provide such a theory would be enormously difficult, if not impossible. In any case, in the present confusing and apparently contradictory state of the evidence in this field, a state that should no doubt be attributed (at least in part) to the lack of any theory adequate to serve as even the most tentative of working hypotheses, it is impossible to say with confidence and precision just what are the phenomena of which we need to take account. Nevertheless there are three suggestions that it may perhaps be useful to consider.

CAUSAL EXPLANATION. The first suggestion concerns the possibility of interpreting precognitive correlations in causal terms. To give a causal account of the subsistence of a statistically significant correlation between two series of events *A* and *B* involves showing either (1) that *A*

results from *B*, or (2) that *B* results from *A*, or (3) that both *A* and *B* result from some third cause or set of causes, or (4) that both *A* and *B* are causally independent results of separate chains of causation. Suppose *A* is a series of precognitive guesses or anticipations and *B* a series of fulfillments or verifications. Series *A* cannot result from series *B*, for that would involve the logical impossibility of future occurrences bringing about events in the past. Series *B* cannot result from series *A*, for if it does, then the case is ipso facto disqualified by definition. And *A* and *B* cannot both result from some third cause or set of causes, for if they do, then again the case is by definition disqualified from rating as genuinely precognitive. The only remaining possibility is to say that *A* and *B* are both the causally independent results of separate chains of causation.

But to say this is precisely not to display a causal connection between *A* and *B*; it is, rather, to imply that the statistically significant correlation between the two series is a coincidence. This conclusion may be disturbing, but at least it has the merit of not involving any actual self-contradiction. For to establish a statistically significant correlation between two series of events is not thereby and necessarily to establish that these series are in any way connected causally. In the face of any correlation, however perfect and however extended, it is always significant, although often foolishly misguided, to insist that there is nevertheless no causal connection. Statements of constant conjunction do not entail statements of causal connection. Anyone who insists on a stronger sense of statistical significance, which would entail the subsistence of a causal connection, and who then proceeds to stipulate that a precognitive correlation would have to be statistically significant in this stronger sense, will succeed only in making his concept of precognition self-contradictory from the start.

COINCIDENCE. It seems that any explanation or, if that now becomes too strong a word, any account of precognition as such will have to center on the notion of coincidence or of something very like it. The laws, if there are any laws to be discovered, will describe the conditions under which we may expect to find precognitive correlations. One is reminded of C. G. Jung's talk about "synchronicity phenomena." For "synchronicity phenomenon" is in fact only a pretentious neologism for "coincidence," with perhaps a built-in suggestion that such phenomena are both more common and also somehow more significant than might be thought. It is a similarity that might easily be overlooked because of Jung's terminological peculiarities, because he associates the idea with many of his own more bizarre inventions, and because he exploits it for his own, it seems, often willfully antiscientific and antirational ends. A law of the kind suggested might paradoxically but pointedly be characterized as a law about the regularities in the conditions for the occurrence of a certain sort of coincidence.

STATISTICAL EXPLANATION. Theorists seem to have taken far too little notice of the surely remarkable fact that it seems to be impossible either for the subjects or for anyone else to achieve any significant success in identifying, without reference to the targets, the particular guesses that are going to prove to be hits. Another similar and similarly neglected fact is that even after the guesses have been scored against the targets we have no criterion for distinguishing any particular hit as precognitive. In each case the reason for talking of precognition is not that any particular guess can, at some stage, be identified as precognitive but that, after the guesses have been checked against the targets, the proportion of hits in a series of guesses is found to be significantly above mean-chance expectation.

With appropriate alterations the same thing seems to be true of all ostensible parapsychological phenomena. It is usually argued that whereas this perhaps has to be allowed in the case of quantitative experiments in card guessing, dice throwing, and so forth, it does not apply at all to what appear to be spontaneous cases of telepathy and clairvoyance, precognitive or straight. But this is surely wrong. For suppose we find that someone who had no means of inferring that the *Titanic* might meet disaster nevertheless had a dream that is later found to have corresponded in amazing detail with what actually happened on the night when that great ship went down. Still, our only warranty for describing his dream as precognitive lies precisely in that extraordinary degree of correspondence: Any single item of correspondence might be dismissed as something that was bound to happen "by the law of averages," and so no single item can be picked out as unequivocally precognitive.

Of course this situation may conceivably at any time be transformed by the progress of the research. But at the time of writing it remains true that all the putative varieties of ESP, precognition in particular, are and must be defined in essentially statistical terms. This is no reason to ignore or to dismiss the evidence. But it may very well prove to be a significant theoretical pointer.

See also Parapsychology.

Bibliography

For general discussion consult C. J. Ducasse, "Broad on the Relevance of Psychical Research to Philosophy," pp. 375–410, A. G. N. Flew, "Broad on Supernormal Precognition," pp. 411–436, and C. D. Broad, "A Reply to My Critics," pp. 709–830, in *The Philosophy of C. D. Broad*, edited by P. A. Schilpp (New York: Tudor, 1959); and W. G. Roll, "The Problem of Precognition," in *Journal of the Society for Psychical Research* 41 (1961): 2ff. Roll's article is valuable especially for its bibliography.

There have been many ingenious discussions of ways of making operational distinctions between the various forms of ESP phenomena in the parapsychological journals since about 1930. For an excellent example, see C. W. K. Mundle, "The Experimental Evidence for PK and Precognition," in *Proceedings of the Society for Psychical Research* 49 (1949–1952): 61–78.

For a criticism of Dunne's "Serial Theory of Time," see A. G. N. Flew, *A New Approach to Psychical Research* (London: Watts, 1953), Appendix II: "An Experiment with 'Time.'"

C. D. Broad's "The Philosophical Implications of Foreknowledge" was published in *PAS*, Supp. 16 (1937): 177–209. Broad referred the empirically curious to H. F. Saltmarsh's "Report on Cases of Apparent Precognition," in *Proceedings of the Society for Psychical Research* 42 (1934): 49–103. With Saltmarsh's paper one may compare D. J. West's considerably more skeptical "The Investigation of Spontaneous Cases," in *Proceedings of the Society for Psychical Research* 48 (1948): 264–300. The weight of both evidence and research has now shifted away from ostensible spontaneous cases of ESP toward quantitative experiments in card guessing. The classic series is that reported by S. G. Soal and K. M. Goldney in "Experiments on Precognitive Telepathy," in *Proceedings of the Society for Psychical Research* 47 (1942–1945): 21–150. This work was hailed by Broad in 1944 in "The Experimental Establishment of Telepathic Communication." Soal and Frederick Bateman have since produced a general survey, *Modern Experiments in Telepathy* (New Haven, CT: Yale University Press, 1954).

On the infallibility of precognitive experiences and on cause and effect, see M. A. E. Dummett, "Can an Effect Precede Its Cause?" in *PAS*, Supp. 28 (1954): 27–44, and the reply with the same title by A. G. N. Flew in that volume, on pp. 45–62. See also Flew's *Hume's Philosophy of Belief* (London: Routledge and Paul, 1961), Ch. 6.

For a fuller criticism of Jung's theory of synchronicity phenomena, see A. G. N. Flew's "Coincidence and Synchronicity," in *Journal of the Society for Psychical Research* 37 (1953–1954): 198–201.

OTHER RECOMMENDED TITLES

Braude, Stephen E., ed. *The Limits of Influence: Psychokinesis and the Philosophy of Science*. New York: Routledge, 1991.

Braude, Stephen E. "Psi and Our Picture of the World." *Inquiry* 30 (1987): 277–294.

Brier, Robert. "Mundle, Broad, Ducasse, and the Precognition Problem." *Philosophy Forum* 14 (1974): 161–169.

Brier, Robert. *Precognition and the Philosophy of Science: An Essay on Backwards Causation*. New York: Humanities Press, 1974.

Craig, William Lane. "Divine Foreknowledge and Newcomb's Paradox." *Philosophia* 17 (1987): 331–350.

Dummett, Michael. "Causal Loops." In *The Nature of Time*, edited by Raymond Flood and Michael Lockwood. Oxford: Blackwell, 1986.

Lucas, J. R. "Foreknowledge and the Vulnerability of God." In *The Philosophy of Christianity*, edited by Godfrey Vesey. Cambridge, U.K.: Cambridge University Press, 1989.

Meehl, Paul. "Precognitive Telepathy II." *Nous* 12 (1978): 371–395.

Werth, Lee F. "Normalizing the Paranormal." *American Philosophical Quarterly* 15 (1978): 47–56.

Zagzebski, Linda. *The Dilemma of Freedom and Foreknowledge*. Oxford: Oxford University Press, 1991.

Zagzebski, Linda. "Recent Work on Divine Foreknowledge and Free Will." In *The Oxford Handbook of Free Will*, edited by Robert Kane. Oxford: Oxford University Press, 2002.

Antony Flew (1967)
Bibliography updated by Benjamin Fiedor (2005)

PREDESTINATION

See *Determinism, A Historical Survey*

PREDICATE

See *Subject and Predicate*

PREFERENTIAL TREATMENT

See *Affirmative Action*

PRESCRIPTIVISM

See *Metaethics; Noncognitivism*

PRE-SOCRATIC PHILOSOPHY

"Pre-Socratic" is the term commonly used (and the one that will be used here) to cover those Greek thinkers from approximately 600 to 400 BCE who attempted to find universal principles that would explain the whole of nature, from the origin and ultimate constituents of the universe to the place of man within it. Yet 400 was the last year of Socrates' life, and among the Sophists, who are also excluded, Protagoras and Gorgias were older than he and others were his contemporaries. "Pre-Socratic" there-

fore indicates not so much a chronological limit as an outlook and a range of interests. This outlook Protagoras and Socrates deliberately attacked, condemning natural philosophy as worthless compared with the search for a good life, the discussion of social and political questions, and individual morality. Socrates also dismissed its explanations as inadequate because expressed predominantly in terms of origins and internal mechanisms. In his view explanation should be functional, looking to the end rather than the beginning. Thus, for the last sixty or so years of the fifth century, both points of view existed, and a lively controversy went on between them. It was not that the natural philosophers excluded human nature from their investigations but that they saw man and society in a larger framework, as a particular late stage in cosmic development, whereas the others deliberately turned their backs on the external world. The universal and speculative character of pre-Socratic thought was also combated by some of the fifth-century medical writers, and it was in the fields of physiology and hygiene that observational science reached its highest point in this period.

NATURE OF THE EVIDENCE

Before attempting to describe the pre-Socratic doctrines, it is necessary to emphasize the peculiar nature of our sources of knowledge. None of the pre-Socratics' works has survived independently. We have a few references in Plato, some more systematic discussion in Aristotle, and information from later compilers and commentators of which the greater part goes back to a history by Aristotle's pupil Theophrastus. Actual quotations occur and are in some cases extensive, as with the prose fragments of Heraclitus and the 450 surviving lines of Empedocles. Yet, from Aristotle onward, the men who passed on this information were not historians in the modern sense but wrote from a particular philosophical viewpoint (most often Peripatetic), searching the past for anticipations of their own ideas and selecting and arranging their material accordingly. The task of reconstruction and interpretation is thus very different from and more precarious than that of interpreting a philosopher whose original writings are still available for study.

THE MILESIAN SCHOOL

Pre-Socratic philosophy differs from all other philosophy in that it had no predecessors. Philosophy has been a continuous debate, and even highly original thinkers can be seen developing from or reacting against the thought of a predecessor. Aristotle is unimaginable without Plato; Isaac Newton, without René Descartes, Johannes Kepler, Galileo Galilei, and many others. But with the Greeks of the sixth century the debate begins. Before them no European had set out to satisfy his curiosity about the world in the faith that its apparent chaos concealed a permanent and intelligible order, and that this natural order could be accounted for by universal causes operating within nature itself and discoverable by human reason. They had predecessors of a sort, of course. It was not accidental that the first pre-Socratics were citizens of Miletus, a prosperous trading center of Ionian Greeks on the Asiatic coast, where Greek and Oriental cultures met and mingled. The Milesian heritage included the myths and religious beliefs of their own peoples and their Eastern neighbors and also the store of Egyptian and Babylonian knowledge—astronomical, mathematical, technological. The influence of this heritage was considerable. Yet the Milesians consciously rejected the mythical and religious tradition of their ancestors, in particular its belief in the agency of anthropomorphic gods, and their debt to the knowledge of the East was not a philosophic one. That knowledge was limited because its aim was practical. Astronomy served religion; mathematics settled questions of land measurement and taxation. For these purposes the careful recording of data and the making of certain limited generalizations sufficed, and the realm of ultimate causes was left to dogmatism. For the Greeks knowledge became an end in itself, and in the uninhibited atmosphere of Miletus they gave free play to the typically Greek talent for generalization, abstraction, and the erection of bold and all-embracing explanatory hypotheses.

Consciously, the revolt of the Milesian philosophers against both the content and the method of mythology was complete. No longer were natural processes to be at the mercy of gods with human passions and unpredictable intentions. In their place was to come a reign of universal and discoverable law. Yet a whole conceptual framework is not so easily changed. Poetic and religious cosmogonies had preceded the schemes of the Milesians, and the basic assumptions of these can be detected beneath the hypotheses of their philosophic successors. Nevertheless, the achievement of abandoning divine agencies for physical causes working from within the world itself can hardly be overestimated.

It was common to the mythologies of Greece and neighboring civilizations (and, indeed, to others) that the world arose from a primitive state of unity and that the cosmogonic process was one of separation or division. This was the first act of the Hebrew Creator. In the Babylonian *Enuma Elish* the original state of the universe was an undefined mass of watery cloud. The Greek theogony

of Hesiod speaks of Heaven and Earth, conceived as anthropomorphic figures, lying locked in an embrace until their son forced them apart as Marduk formed heaven and earth by splitting apart the body of the monster Tiamat. Euripides relates an old tale according to which earth and heaven were once "one form" and after their separation brought to birth the whole variety of living things. In Egypt (like Babylonia, a river culture) everything arose out of the primeval waters.

THALES. It is not surprising, therefore, that the first people to seek a universal explanation of the world along rational lines assumed that it was in substance a unity from which its variety had been produced by some process of segregation. The key, they thought, lay in identifying the single substance that must satisfy the condition of being able to produce variety out of itself. Thales (active in 585 BCE), who chose water or moisture, may still have had the myths at the back of his mind. For him the earth floated on water as it did for the Egyptians. Little else certain is known of him, and we can only guess at his reasons. Water can be seen as solid, liquid, and vaporous. Aristotle thought it more probable that Thales was influenced by the essential connection of moisture with life, as seen in such substances as semen, blood, and sap. With the removal of external personal agents, the world must initiate its own changes, and at this early stage of speculation the only possibility seemed to be that life of some kind is everywhere and that the universe is a growing, organic structure. This may be the explanation of the saying attributed to Thales: "Everything is full of gods."

ANAXIMANDER. With Anaximander, Thales' younger contemporary, there emerges the notion of the four primary opposites that later, when the concepts of substance and attribute had been distinguished, gave rise to the four elements adopted by Aristotle and destined for a long and influential history. Anaximander spoke of only the hot and the dry, which were inevitably in conflict with the cold and the wet. This led him to a momentous idea. The original substance of the universe could not be anything definitely qualified like water, for how could the cold and wet produce their opposites, the hot and dry? Water quenches fire; it cannot engender it. Prior to all perceptible body there must be an indefinite something with none of the incompatible qualities implied by perceptibility. Although still regarding all that exists as corporeal, Anaximander is the first to find ultimate reality in the nonperceptible.

This primary substance he called the apeiron, a word of many meanings all related to the absence of limits—everlasting, infinite, indefinite. Because it was imperishable, the origin of all things, and the author of their changes, he called it (says Aristotle) divine. From it all things have been "separated out," though in what sense they were previously "in" it while the apeiron itself remained a unity is a question that probably did not present itself to him. Somewhere in the apeiron, Theophrastus asserts, a "germ" or "seed" of hot and cold was separated off, and from the interaction of these two flowed the whole cosmic process. A sphere of flame enclosed a moist mass, more solid at the center where the earth formed, vaporous between. The sphere burst into rings around which the dark vapor closed, leaving holes through which we see what appear as sun, moon, and stars. Wet and dry continue to separate, forming land and sea, and finally life itself is produced by the same action of heat (sun) on the cold and moist portions of the earth. The first animals were born in water and crawled onto dry land. Human infants were originally born and nurtured within the bodies of fishlike creatures, for under primitive conditions unprotected babies could not have survived.

Earth, a flat cylinder, hangs freely in space because of its equal distance from all parts of the spherical universe. The sun is the same size as Earth. Eclipses are caused by the closing of the holes in the vapor tubes of the sun and moon. In this first of all attempts at a rational cosmogony and zoogony, the sudden freedom from mythical modes of thought is almost incredible.

ANAXIMENES. Further reflection led Anaximenes, the youngest member of the Milesian school, to a different conclusion about the primary substance: It was air. In its elusiveness and invisibility as atmospheric air, it could almost match the apeiron, and, whereas apeiron, once differentiated into a universe, could no longer be so called, air could become hotter and colder, rarer and denser, and still remain the same substance. Moreover, this theory allowed Anaximenes to break with the notion of separation, which was, at bottom, mythical, and account for the universe by the extension of a known natural process. This was condensation and rarefaction, the former of which he associated with cold and the latter with heat. Air as it rarefies becomes fire; condensed, it turns first to wind, then to cloud, water, earth, and stones. In other words, it is all a question of how much of it there is in a given space, and for the first time the idea enters science that qualitative differences are reducible to differences of quantity. This is Anaximenes' main achievement,

although there is no evidence that he applied the principle with any mathematical exactness.

With air as his basic, self-changing substance, Anaximenes could find room for the ancient belief that life was identical with breath. Macrocosm and microcosm were animated by the same principle: "Just as our soul, which is air, integrates us, so breath and air surround the whole cosmos."

The few details that we have of his cosmology suggest that compared with Anaximander's, it was reactionary and timid. His contribution lies elsewhere.

THE PYTHAGOREANS

Pythagoras (c. 570–490 BCE) was also an eastern Greek but migrated from his native Samos to Croton in southern Italy. As a result the western or Italian Greek philosophers, even when not actual members of his school, became known for a characteristic outlook very different from that of the materialistic and purely rational Milesians and stamped with the impress of his remarkable genius. He founded a brotherhood dedicated to *philosophia* (the word was believed to be his invention) as a way of life, with a strong religious, and also a political, element. Philosophically, his importance lies in the shift of interest from matter to form. Inspired, it is said, by the discovery that the musical intervals known to the Greeks as consonant (and marked by four fixed strings on the seven-stringed lyre) were explicable in terms of ratios of the numbers 1 through 4, Pythagoras saw the universe as one glorious *harmonia*, or mathematico-musical structure. Number was the key to nature. This idea had incalculable consequences for science even if it led at the time to some rather fanciful equations of natural objects and moral qualities with particular numbers. In spite of that, by the time of Socrates the school had made real progress in mathematics. Since the cosmic harmony included everything, all life was akin. The soul was immortal and underwent a series of incarnations, both human and animal. Philosophy was the effort to understand the structure of the cosmic harmony, with the ultimate aim of integrating the philosophic soul more closely into that harmony on the principle that knowledge assimilates the knower to its object. This aim also demanded the observance of certain religious precepts of which the most important was abstention from animal food.

HERACLITUS

Heraclitus (active c. 500 BCE) objected to the Pythagorean emphasis on harmony, maintaining that, on the contrary, strife and opposition were the life of the world. Life was maintained by a tension of opposites fighting a continuous battle in which neither side could win final victory. Thus, movement and the flux of change were unceasing for individuals, but the structure of the cosmos remained constant. This law of individual flux within a permanent universal framework was guaranteed by the Logos, an intelligent governing principle materially embodied as fire, the most subtle element and identified with soul or life.

Philosophy had thus far meant the search for an essentially simpler reality underlying the bewildering confusion of appearances. The answers fell into two broad categories, matter and form: Reality was a single material substance (the Milesians) or an integral principle of structure that could be expressed in terms of numbers (the Pythagoreans). Heraclitus, with a statement like "You cannot step twice into the same river," reaches the logical conclusion of the materialistic answer. The water will be different water the second time, and, if we call the river the same, it is because we see its reality in its form. The logical conclusion of form-philosophy is the opposite of flux—namely, a belief in an absolute, unchanging reality of which the world of change and movement is only a quasi-existing phantom, phenomenal, not real. (This conclusion was reached in the idealism of Plato, which was largely of Pythagorean inspiration.)

ELEATIC SCHOOL: UNITY OF REALITY

At this time the direction of philosophy was changed by the precocious and uncompromising logic of Parmenides of Elea, who was perhaps twenty-five years younger than Heraclitus. For the first time abstract, deductive reasoning is deliberately preferred to the evidence of the senses: "Ply not eye and ear and tongue, but judge by thought." He concluded that if there is any reality at all (in the language of his time, if "it is"), it must be (1) one only (for if more than one, its units could be separated only by "what is not"); (2) eternal and unchanging (for to speak of change or perishing is to say that reality at some time "is not" what it was, but to say of "what is" "it is not" is contradictory and impossible); (3) immovable (this follows from his statement that "all is full of what is"; since it cannot admit discontinuity or lack of homogeneity and since "what is not is not," the spatial requirements of locomotion cannot be provided).

In this way he "proved" that, on the premise of his predecessors that reality is one, differentiation of the real can never occur. It remains one—a timeless, changeless, motionless, homogeneous mass, which he compared to a sphere. The multiple, changing world of appearances is

an illusion of our senses. Only as a concession to human weakness, and in recognition of our practical need to come to terms with the show of a natural world, did he append a cosmology of the conventional type, beginning with two principles, heat-light and cold-darkness. Cosmogony from a single origin was no longer possible, yet he explicitly warns his hearers that reality is in truth a unity and that the cosmos is only a deceitful appearance to mortals.

It is disputed whether the One Reality of Parmenides is material. The question can hardly be answered, since we are still in a period before the distinction between material and nonmaterial could be drawn. The important thing is that it was nonsensible and could be reached only by thought. Parmenides was the first philosopher to distinguish explicitly between the sensible and the intelligible and to condemn the former as unreal. Plato himself, though fully aware of the distinction between material and spiritual, usually preferred to call them sensible and intelligible, and it is very doubtful whether the philosophy of Platonic idealism would ever have been possible without Parmenides.

ZENO AND MELISSUS. Parmenides had two followers, who, with him, are known as the Eleatic school. Zeno of Elea (born c. 490 BCE) concentrated on a defense of the proposition that reality is one and immovable by the dialectical method of showing up absurdities in the contrary view. His famous paradoxes are aimed at demonstrating the impossibility of plurality and movement. Melissus of Samos (active in 440 BCE) modified Parmenides' ideas to the extent of saying that reality is infinite. He explicitly denied the possibility of empty space (which Parmenides had only hinted at) and said that if there were many things, each would have to have the characteristics of the Parmenidean One. It is therefore probable that the atomists had him especially in mind when they boldly explained the world in terms of space plus tiny entities, each of which had many of the Eleatic qualities—indivisibility, homogeneity, unalterability.

The naïveté of Parmenides' logic and the purely linguistic nature of some of his difficulties seem obvious now, but at the time his questions appeared unanswerable. There were only two ways out: either to abandon monism and admit the ultimate plurality of the real or to admit the unreality of the natural world. The latter solution was Plato's, with his contrast between "what always is and never becomes" and "what is continually becoming (like the flux of Heraclitus) but never truly is." The remainder of pre-Socratic thought is occupied with attempts to save the phenomena by adopting some form of pluralism.

THE PLURALISTS: EMPEDOCLES

The first of the pluralistic systems was that of Empedocles (c. 490–430 BCE), a Sicilian poet-philosopher steeped in the Western tradition, with its combination of rationalism and mystical religion so different from the purely scientific outlook of the Ionians. His proposal was the first clear enunciation of the four-element theory. Fire, air, water, and earth are the ultimate roots of all things, themselves ungenerated and indestructible. Everything in nature comes into being and perishes by the mixture and separation of these substances. The first premise is no longer "It is" but "They are." Thus, trees and animals, clouds and rocks, are not mere illusion. However, since they are only temporary combinations of the four "realities" in varying proportions, we can admit that they themselves are not "real." Nor need the forbidden concepts of "becoming" and "perishing" be invoked; mixture and separation will account for all. Locomotion is, of course, necessary, and, although he accepts the Eleatic denial of empty space, Empedocles seems to have thought that this could occur by some reciprocal and simultaneous exchange of place, the whole remaining full.

The four elements are not self-moving (another concept that Parmenides had rendered difficult), and the blend of mystic and rationalist in Empedocles appears especially in his motive causes. These were two, Love and Strife, the former bringing disparate elements together and the latter drawing them apart. They are in endless opposition and prevail in turn, bringing about a double evolutionary cycle. Under Love all four elements are indistinguishably fused in a sphere; under Strife the same sphere contains them in separate layers. During the contest, when neither Love nor Strife is in complete control and when the elements are partly joined and partly separated, a world like our own is formed. Nothing existent is as yet incorporeal, though Love and Strife are of finer and more tenuous substance than the elements. Their names are no metaphors, nor is their action purely mechanical. Under Love the elements are dear to and desired by one another; Strife makes them grim and hostile. Nothing is purely inanimate, and everything has its share of consciousness.

Besides his poem on nature, Empedocles also wrote a religious one, in which the moral character of Love and Strife is emphasized—Love is good, Strife evil. In the present world Strife is gaining, and men have fallen from a previous blessed state by giving themselves to Strife and

sin, above all the sin of killing and eating animals. All life is akin, as it was to the Pythagoreans, and our souls are fallen spirits that must undergo a series of incarnations before they can win back their former state by abjuring Strife and cultivating Love. What the substance of the spirits was is not clearly stated, but most probably in their pure state they were portions of Love that are now contaminated with Strife.

ANAXAGORAS

Anaxagoras of Clazomenae (c. 500–428 BCE) brings us back to Ionia both geographically and in spirit. His motive is rational curiosity entirely uncomplicated by religious preoccupations. Even Parmenides, a Westerner like Empedocles, had written in verse and represented his deductive arguments as a revelation from a goddess. In his return to prose, as in his purely scientific aims, Anaxagoras is the heir of the Milesians. At Athens, where he lived until exiled for atheism, he was a member of the brilliant and freethinking circle of Pericles. His prosecution seems to have had a political flavor, but the charge is nevertheless significant: He declared the sun to be not a living divinity but a lump of incandescent rock larger than the Peloponnese.

To save the phenomena without admitting the coming into being or destruction of what exists, he adopted an extreme form of pluralism plus a first cause of motion, which he called Mind. It is described as knowing all things and having the greatest power, and, in order to control the material world, it is entirely outside the mixture of which the material world is formed. It is not easy to be sure whether Anaxagoras is at last trying to express the notion of incorporeal being without an adequate vocabulary or whether he still thinks of Mind as an extremely subtle and tenuous form of matter. At any rate, its separateness from the constituents of the cosmos is emphasized at every turn. In spite of the references to its knowledge and power, it action seems to be confined to the earliest stages of cosmogony, except in the case of living creatures. They are an exception to the rule that Mind is in nothing else, and them it still controls.

In the beginning "all things were together," a stationary mass in which nothing could be distinguished. Mind is the agent that has produced from this an ordered cosmos. It did so by starting a rotatory movement or vortex, which by its own increasing speed brought about the gradual separation of different forms of matter. Anaxagoras's highly subtle and ingenious theory of matter seems to have been especially prompted by the need to explain nourishment and organic growth: How can flesh and hair come out of the not-flesh and not-hair of the food we eat? After Parmenides the coming into being of new substances is disallowed. Anaxagoras answered that there is a portion of everything in everything—that is, every distinguishable substance, in however small a quantity, contains minute particles of every other but is characterized by that which predominates. He boldly asserted the existence of the infinitesimal (which Zeno had denied) in the words: "Of the small there is no smallest."

THE ATOMISTS

Perhaps around 430 BCE Leucippus promulgated the much simpler theory of atomism, which was further developed by his famous pupil Democritus of Abdera (born c. 460 BCE). Like the other theories, this one arose in direct response to the Eleatic challenge. Its most striking innovation for its time was the assertion of the existence of genuine empty space. Thus far, everyone had believed that "what is" must be some form of body, and, when Parmenides brought into consciousness the implicit consequence that space, not being "what is," must be "what is not" (that is, nonexistent), his conclusion seemed logically inescapable. Hence, even the atomists had to use the paradoxical expression that it is no more correct to say of "what is" than of "what is not" that it *is*.

At this particular point in the philosophic debate, this was the only way of expressing the conviction that, though not any kind of stuff, space must be assumed if the plain facts are to be explained. Democritus, said Aristotle, is to be commended for refusing to be dazzled by the abstract logic of Parmenides and for relying on the kind of argument more proper to a natural scientist. Reality consists of innumerable microscopic and indivisible (*a-tomos* = uncuttable) bodies in motion in infinite space. They are solid and homogeneous but infinitely variable in size and shape. At different places in the infinite, they have collided and become entangled. Projections hook together, convex fits into concave, and so on. Their continued motion sets up a vortex in which the larger and heavier fall into the center and the smaller and lighter are extruded to the circumference; in this way a cosmos is formed. There are many worlds, and not all are similar to our own. The first atomists appear to have provided no separate cause of motion, perhaps because they deemed it sufficient to free the atoms by setting them loose in infinite space. After all, the chief Eleatic arguments against motion had been the continuity of being and the nonexistence of a void.

Only atoms and the void exist. Sensible qualities other than size and shape are subjective, caused by inter-

action between the atoms of external objects and those in our own bodies. This was worked out in considerable detail. For instance, hard objects have their atoms more closely packed than do soft. Sweet flavors are caused by smooth atoms, bitter and astringent by sharp or hooked. Colors vary according to the positions of surface atoms, which cause them to reflect in different ways the light that falls upon them. Objects are continually throwing off films of atoms, and sight is the reception of these films by the eye. The soul, or life principle, is composed of smooth, round atoms that are even more mobile than the rest and impart to the body the power of motion and cognition, for "soul and mind are the same"—that is, composed of the same kind of atoms. Soul is dispersed throughout the body, alternating with body atoms, but the mind appears to have been a collection of these finest particles that is located probably in the breast. Although the direct objects of sight and hearing, taste and smell, are unreal, they lead the mind to the truth about reality, and Democritus quoted with approval a saying of Anaxagoras: "Phenomena are a glimpse of the unseen."

Ancient atomism (including its revival by Epicurus a century or more after Democritus) has acquired a partly adventitious reputation through its resemblances to nineteenth-century physical theories, but its hard, solid, unbreakable particles have little in common with the ultimate entities of modern science. Its most striking features are the distinction between primary and secondary qualities (upheld by Descartes, Galileo, and John Locke), the explanation of directly observable objects by hypothetical constituents below the level of perception, and the outspoken championship of discrete quanta as opposed to a continuum. Its inadequacy in allowing no mode of action other than direct contact, collision, and interlocking was evident in some physical problems—for example, in its attempted explanation of magnetism and, most of all, in the effort to include within its purview the phenomena of life and thought. The atomic structure of matter has indeed been a fruitful hypothesis, but the intention of its authors is best understood in the context of their time and as an attempt to escape the Eleatic dilemma, rather than as an anticipation of postmedieval science.

DIOGENES OF APOLLONIA

The teleological explanation, which one would naturally associate with Anaxagoras's adoption of Mind as first cause, appears more strongly in the second half of the fifth century in a less gifted thinker, Diogenes of Apollonia. He put Mind back into the mixture by returning to Anaximenes's idea that the primary substance is air or breath and by identifying this air in its purest (dry and warm) state with intelligence. The regularity of cosmic events he regarded as evidence of intelligent control, going so far as to say that anyone who reflects will agree that all is arranged in the best possible way. Breath is also the life of humans and animals, so that all owe their soul and mind to the same material principle—"a small portion of the god"—which they share in varying degrees of purity. He probably thought he avoided the Eleatic arguments against a materialistic monism by the admission of void, which, by the time he wrote (after Melissus and Leucippus), would in any case be recognized as necessary for the process of condensation and rarefaction by which air produced the variety of nature.

When we consider the grotesqueness of some of the mythological background from which the pre-Socratic thinkers started, we must be amazed by the intellectual insight and firm grasp of universal principles that at their best they were capable of displaying. But a dispassionate assessment of their contribution to the history of philosophy would probably show that, to use a metaphor, although they manufactured many of the pieces and set them on the board, Plato and Aristotle were the first players who learned the rules and started the game. The pieces are those opposed concepts by means of which philosophical discussion is maintained: being and becoming, sensible and intelligible, analytic and synthetic, appearance and reality, time and eternity, materialism and idealism, mechanism and teleology, and so forth. Once these stand out clearly, a philosopher may champion one or the other, but the pre-Socratics could not yet do this. One cannot speak realistically of a controversy among them between, say, materialists and idealists. The achievement of their intellectual effort and controversy was that by the end of this period a clear notion of what was meant by matter and mind, sensible and intelligible, phenomenal and real, and the rest was at last emerging, so that succeeding generations had the set in their hands and could begin the game in earnest. For the first of all philosophers, this was no mean achievement.

Their interests were, of course, in modern terms, as much scientific as philosophical, and in this sphere also they could claim some remarkable results. For instance, before the end of the period the true cause of both lunar and solar eclipses had been discovered (probably by Anaxagoras), and certain Pythagoreans had abandoned the geocentric cosmology, asserting that Earth, the sun, and the planets all circled round a central fire. But it is probably fair to say that their scientific discoveries

appeared only as by-products of the main controversies and of the few universal principles from which they confidently deduced even the details of the physical world. The true and lasting discoveries were not picked up and developed as they would have been by post-Renaissance scientists because, owing to the different preoccupations of philosophy at their time, they had no firm basis in established fact and did not in any way stand out from other and, to us, more fanciful assumptions.

See also Alcmaeon of Croton; Anaxagoras of Clazomenae; Anaximander; Anaximenes; Apeiron/Peras; Appearance and Reality; Archē; Chaos Theory; Cosmology; Cosmos; Diogenes of Apollonia; Empedocles; Hen/Polla; Heraclitus of Ephesus; Infinity in Theology and Metaphysics; Leucippus and Democritus; Logos; Materialism; Melissus of Samos; Monism and Pluralism; Orphism; Parmenides of Elea; Philolaus of Croton; Pre-Socratic Philosophy; Pythagoras and Pythagoreanism; Thales of Miletus; Xenophanes of Colophon; Zeno of Elea.

Bibliography

Fränkel, Hermann. *Early Greek Poetry and Philosophy: A History of Greek Epic, Lyric, and Prose to the Middle of the Fifth Century*. Translated by Moses Hadas and James Willis. New York: Harcourt Brace Jovanovich, 1973.

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 1: *The Earlier Presocratics and the Pythagoreans*. Cambridge, U.K.: Cambridge University Press, 1962. Vol. 2: *The Presocratic Tradition from Parmenides to Democritus*. Cambridge, U.K.: Cambridge University Press, 1965.

COLLECTIONS OF ESSAYS

Allen, R. E., and David J. Furley, eds. *Studies in Presocratic Philosophy*. Vol. 2: *The Eleatics and the Pluralists*. London: Routledge & Kegan Paul, 1975.

Curd, Patricia, and Daniel W. Graham, eds. *The Oxford Handbook of Presocratic Philosophy*. Oxford: Oxford University Press, 2006.

Furley, D. J., and R. E. Allen, eds. *Studies in Presocratic Philosophy*. Vol. 1: *The Beginnings of Philosophy*. London: Routledge & Kegan Paul, 1970.

Long, A. A., ed. *The Cambridge Companion to Early Greek Philosophy*. Cambridge, U.K.: Cambridge University Press, 1999.

Mourelatos, A. P. D., ed. *The Pre-Socratics: A Collection of Critical Essays*. Garden City, NY: Doubleday, 1974; second printing, with revised introduction and bibliography, Princeton, NJ: Princeton University Press, 1993.

Taylor, C. C. W., ed. *From the Beginning to Plato*. Routledge History of Philosophy, 1. London: Routledge, 1997.

PARTICULAR THINKERS OR SCHOOLS

Burkert, Walter. *Lore and Science in Ancient Pythagoreanism*. Translated by E. L. Minar, Jr. Cambridge, MA: Harvard University Press, 1972.

Huffman, Carl A. *Philolaus of Croton: Pythagorean and Presocratic, Commentary on the Fragments and Testimonia with Interpretive Essays*. Cambridge, U.K.: Cambridge University Press, 1993.

Kahn, Charles H. *Anaximander and the Origins of Greek Cosmology*. New York: Columbia University Press, 1960. Repr. Philadelphia: Centrum, 1982. Repr. Indianapolis: Hackett, 1994.

Kahn, Charles H. *The Art and Thought of Heraclitus: An Edition of the Fragments with Translation and Commentary*. Cambridge, U.K.: Cambridge University Press, 1979.

Mourelatos, A. P. D. *The Route of Parmenides*. New Haven. CT: Yale University Press, 1970.

Sider, David. *The Fragments of Anaxagoras*. Beiträge zur klassischen Philologie, 118. Meisenheim am Glan: Hain, 1981.

"Phoenix Pre-Socratics," a series published by the University of Toronto Press, comprises volumes that provide ancient text, English translation, and explanatory comments: (1) David Gallop, *Parmenides of Elea*, 1984, rev. ed. 1991; (2) Robinson, T. M., *Heraclitus*, 1987; (3) Inwood, Brad, *The Poem of Empedocles*, 1992, rev. ed. 2001; (4) Lesher, J. H., *Xenophanes of Colophon*, 1992; (5) Taylor, C. C. W., *The Atomists: Leucippus and Democritus*, 1999.

W. K. C. Guthrie (1967)
Bibliography updated by Alexander Mourelatos (2005)

PRESUPPOSING

The notions of "presupposing" and of contextual implication, which we shall compare and contrast in what follows, have come to play increasingly prominent roles in the philosophical literature of the English-speaking world since the 1940s. This development is not accidental but arises from the stress the twentieth century put upon analysis as a fundamental mode of philosophical inquiry. The notions of presupposing and of contextual implication play both negative and positive roles within this general orientation. Negatively, they are devices that contemporary thinkers employ in order to minimize the tendency of philosophers and other reflective persons to view the world in terms of oversimplified conceptual models. Positively, they function as instruments in the dissection and ultimate understanding of certain human activities, especially those that involve the efforts of human beings to communicate with one another, as in promising, stating, saying, implying, a task that, some philosophers feel, is hindered or obstructed by the natural disposition of reflective individuals to subsume such activities under excessively simple descriptions. The appeal to the notions of presupposing and of contextual implication has thus served to widen—and at the same time to make more accurate—our conceptions of the cir-

cumstances in which human communication takes place. This entry will describe the history (all of it relatively recent, of course) of the major developments that have taken place with regard to these subjects, and will in this way attempt to bring out their essential features.

SIMILARITIES AND DIFFERENCES

It is no simple matter to show why presupposing and contextual implication are two separate concepts, since the differences between them are subtle. Most writers have, in fact, not discriminated between them, in part because both notions are slippery but also because they have similar functions. Their similarities may be elucidated as follows. If we distinguish between what a person explicitly states, or asserts, when he utters certain words in certain circumstances and what he (or perhaps his statement) implies, then the concepts of presupposing and of contextual implication belong to the latter category rather than to the former. This crude distinction must be refined further, however, for the sense of "implies" that is being marked out here is not that of logical implication in any of the various senses of that term—for example, the sense involved in saying that "*X* is a husband" implies "*X* is married." Indeed, both presupposing and contextual implication are to be contrasted with logical implication.

The kinds of implications that fall into this category may be indicated by simple examples. In saying "alas!" in certain circumstances, I am normally taken as implying that I am unhappy. But I am not taken to be asserting that I am unhappy, as I would be if I were to utter the words "I am unhappy." Or, to vary the example, when a person says, "All my children are now in college," he is normally taken to be implying that he has children (although not to be asserting that he has), and his auditors are justified in making this assumption. Or again, when one says in such sorts of contexts, "Smith has just gone out," he implies, or his words imply, that he believes or knows that Smith has gone out, and those to whom he is speaking are justified in assuming that he does. That the sense of "implication" expressed by these examples is not that of logical implication may be illustrated by the observation that there is no formal contradiction in asserting "All my children are in college, but I have no children" or in asserting "Smith has gone out, but I don't believe he has." Indeed, in standard systems of mathematical logic, the first statement is true whenever the speaker has no children, and the second is true whenever Smith has gone out but the speaker does not believe he has.

Sentences like "All my children are in college, but I have no children" and "Smith has gone out, but I don't believe he has" thus satisfy the rules of logical syntax and, indeed, the rules for correct English. Yet they fall upon the ear as decidedly odd. If employed at all in everyday speech, they would occur only in unusual circumstances—"I don't believe he has" might be whispered as an aside to a confederate, for example. But except for situations like this, they would be perplexing things to say. What, then, is the source of their oddity, given that they do not involve any formal mistake?

It is now generally agreed that the oddity we feel upon hearing such sentences stems from a disparity between the conditions we assume will have been satisfied whenever someone is trying to communicate with another and the utterances we expect will be employed in those circumstances. In effect, this is to say that certain assumptions, or presuppositions, that communicating human beings make in the everyday give-and-take of verbal intercourse, assumptions that thus form the ground of such intercourse, fail to hold or are violated in such circumstances.

Talk about presuppositions and talk about what is contextually implied by a speaker's words thus have in common a reference to the background conditions normally expected to obtain when an utterance is made. If stating and asserting are conceived of as elements constituting part of the foreground of the situation in which communication takes place (that is, as activities that bring an item of information into the immediate focus of attention), then presupposing and contextual implication may be thought of as elements constituting part of the background of the situation (that is, as factors that remain implicit unless they are otherwise articulated but that nonetheless are essential factors in communication). Part of the task that faces the student of informal logic is to specify what these conditions are, how they contribute to the background that makes communication possible, and what sorts of relations exist between them and the utterances that occupy the foreground during the transmission of information.

Let us then call the concepts referring to such conditions background concepts. Because such concepts play covert roles in daily discourse and because their functions are remarkably similar, it is not surprising that many writers have failed to discriminate between them. But not all writers have blurred the distinction. Isabel C. Hungerland is one notable exception. In her important paper "Contextual Implication" (*Inquiry*, Vol. 4, 1960, pp. 211–258), she writes, "The relation (presupposing)

defined by Strawson is *not* that of contextual implication…. The relation between the two may be indicated as follows: When *S* presupposes *S*', a speaker in making the statement *S*, contextually implies that he believes that *S*'" (p. 239). Following Mrs. Hungerland's suggestion and overlooking the many subtleties a full treatment of the subject would demand, we may say that the key distinctions that mark off the one notion from the other are those of scope: Neither the conditions subsumed under the two notions nor the range of entities to which the notions apply are in all cases the same.

Presupposing is a concept referring to those conditions that must be satisfied before an utterance can count as a statement, or if "statement" is so defined that statements need be neither true nor false (see P. F. Strawson, "Identifying Reference and Truth-Values," in *Theoria* 30 [2] [1964]), then presupposing applies to those conditions that must be satisfied before statements can be either true or false. Contextual implication, on the other hand, is a concept that applies to those conditions that must be satisfied before an utterance can count as "normal" in the circumstances in which it is made—that is, it applies to those beliefs a speaker has when he makes the utterance he does in certain circumstances and which rule out that he is lying or deliberately deceiving someone. The range of entities thus referred to by the concept of presupposing is either the class of statements as such or the class of those statements that are either true or false, whereas the range of entities referred to by the notion of contextual implication is the class of beliefs held by the speaker (and, derivatively, by his auditors).

Examples may be invoked at this point to illuminate the above remarks. Suppose during the course of a conversation I say, "The store on the corner sells such goods," not realizing that there is no longer a store on the corner. My remark in this circumstance is neither true nor false; as R. G. Collingwood puts it, the question of its truth or falsity "does not arise." For it is a presupposition of my using that utterance to make a statement (that is, an utterance that can be either true or false) that there be such a store. We may say in such a case that it is a condition of the truth or falsity of the remark that the store exist. But I may well believe that there is such a store, and in making the remark, I imply that I have this belief at the time of my utterance. One of the conditions for the normality of the remark (that is, that I was not lying) is that I had this belief at the time of saying what I did. We may say therefore that the conditions determining the normality of the background from which my remark issued and the conditions determining the background from which a statement would have issued are different conditions. It is this sort of difference in the background conditions that determines the difference between the concepts of presupposing and of contextual implication.

HISTORY OF CONTEXTUAL IMPLICATION

The genesis of the notions of contextual implication and of presupposing differs considerably. As a philosophical subject, contextual implication, under another name, has a longer traceable history in the modern period than does presupposing. The history of contextual implication is mainly connected with developments in moral philosophy, especially with efforts to give a correct analysis of the use of moral language. In G. E. Moore's *Ethics* (London, 1912), for example, we find the following comments:

> There is an important distinction, which is not always observed, between what a man *means* by a given assertion and what he *expresses* by it. Whenever we make any assertion whatever (unless we do not mean what we say) we are always *expressing* one or other of two things—namely, either that we *think* the thing in question to be so, or that we *know* it to be so." (p. 125)

In the subsequent history of moral philosophy the distinction referred to by Moore became the key distinction invoked by those authors who espoused the emotive theory of ethics. According to advocates of this doctrine, the sorts of utterances used in moral contexts ("That's good," "Stealing is wrong") are not being used to make assertions and hence are neither true nor false, as both naturalists and nonnaturalists had assumed. The primary use of such utterances is to express the attitude or the feelings of the speaker toward whatever he is talking about and to arouse comparable attitudes in the auditor. The later history of contextual implication is deeply concerned with the import of this distinction, and the main works in which it is discussed, sometimes critically, are *Language, Truth and Logic* by A. J. Ayer (London, 1936); *The Philosophy of G. E. Moore*, edited by P. A. Schilpp (Evanston, IL, 1942), pp. 540–554; *Ethics and Language* by C. L. Stevenson (New Haven, CT, 1944); *Ethics* by P. H. Nowell-Smith (Harmondsworth, U.K., 1954); *The Emotive Theory of Ethics* by Avrum Stroll (Berkeley, CA, 1954); *The Logic of Moral Discourse* by Paul Edwards (Glencoe, IL, 1955); and "Contextual Implication" by Isabel Hungerland (see above). Various formulas are proposed by some of these writers.

Nowell-Smith says, for example, "A statement *p* contextually implies *q* if anyone who knew the normal con-

ventions of the language would be entitled to infer *q* from *p in the context in which they occur*" (*Ethics*, p. 80). According to Hungerland all such early attempts to characterize the relation that obtains between what a speaker expressly asserts and what he implies suffer either from vagueness or from mistakenly thinking that the relation is a special case of inductive inference. Her own contention is that it is neither vague nor a case of inductive inference, but is, rather, the presumption that in a situation of communication, acts of stating are normal. She thus likens contextual implication to the juridical principle that a man is presumed to be innocent until proved guilty, a principle that is not arrived at inductively, by surveying the evidence, but which serves to place the onus of proof in a legal contest upon the prosecution. As she puts it, "Contextual inference (if we wish to use the word) is a matter, rather, of a communal assumption in the absence of evidence to the contrary, that, in a situation of communication, acts of stating are normal" (p. 233). Her view is that contextual implication depends upon three factors: (1) The presence of a stating context (since the question of a man's believing what he says does not arise in a nonstating context); (2) the presumptions of normality (that is, that within a stating context the implication holds only if the presumptions are principles of communication); and (3) rules for the correct use of an expression (that is, whether belief is implied when a man says *p* will be in part determined by rules for the correct use of *p*).

HISTORY OF PRESUPPOSING

Unlike contextual implication, the notion of presupposing has its genesis in logical theory, especially in those developments involving alternative accounts of Bertrand Russell's theory of descriptions and of the so-called square of opposition. The writer most closely identified with both of these matters is P. F. Strawson of Oxford University. He has dealt with the theory of descriptions in his papers "On Referring" (*Mind*, 1950), "Presupposing" (*Philosophical Review*, 1954) and "Identifying Reference and Truth-Values" (see above) and in his book *Individuals* (London, 1959; Ch. 8 especially). In *Introduction to Logical Theory* (London, 1952) Strawson considers both the theory of descriptions and the square of opposition.

In the works that deal only with the theory of descriptions, Strawson rejects Russell's analysis of sentences containing definite descriptive phrases (that is, phrases of the form "the so and so" used in the singular in English). According to Russell, the analysis of a sentence like "The queen of England is beautiful" contains in part an assertion to the effect that the queen of England exists. Strawson argues, cogently, that this statement is not an explicit part of what is asserted by "The queen of England is beautiful" but is presupposed by a speaker who would use such a sentence in normal circumstances to make a statement. In *Introduction to Logical Theory*, Strawson goes on to define the statement "*S* presupposes *S*'" as follows: "The truth of *S*' is a necessary condition of the truth or falsity of the statement that *S*" (p. 175).

This characterization has been objected to by various writers, including David Rynin, who points out that when "necessary condition" and "truth or falsity of the statement that" are interpreted in the ordinary, truth-functional way, the definition has the paradoxical consequence that all presupposed statements are true. Rynin's demonstration is that $(S \supset S')$ and $(-S \supset S')$, but $(S \vee -S)$; therefore *S*'. Avrum Stroll has also suggested that Strawson's account suffers from the difficulty that if "The king of France no longer exists" is used to make a true statement, then by Strawson's criterion one who employs it thereby presupposes the existence of the king of France. It is now generally agreed that neither Russell's nor Strawson's analysis does full justice to all uses of sentences in everyday English containing "the" phrases in the singular. But regarded as proposals for the development of explanatory models for subparts of everyday discourse, each has considerable merit. In this interpretation Strawson's doctrine belongs to the logical tradition of analyzing descriptive phrases initiated by Gottlob Frege in "Über Sinn und Bedeutung" (1892) and supported by David Hilbert and Paul Bernays in their *Grundlagen der Arithmetik* (Berlin, 1934; Vol. I, p. 384) and by Rudolf Carnap in *Meaning and Necessity* (Chicago, 1947; pp. 33–42).

Strawson has also argued that if universal statements ("All my children are in college") are interpreted as presupposing the existence of the items mentioned by the subject term, paradoxes stemming from modern symbolic interpretations of the square of opposition can be eliminated without affecting the logical relations that one intuitively feels ought to hold between the elements of the square. This matter is persuasively discussed by S. Peterson in "All John's Children" (in *Philosophical Quarterly*, 1960).

PRESUPPOSING IN METAPHYSICS

The notion of presupposition plays an important role in various metaphysical constructions, including Collingwood's *An Essay on Metaphysics* (Oxford, 1940) and Michael Polanyi's *Personal Knowledge* (Chicago, 1958). Collingwood distinguishes (Chs. 3–4) between absolute

and relative presuppositions, arguing that the former are neither true nor false and that metaphysics is the science that ascertains what these absolute presuppositions are. His view is that absolute presuppositions form the basis of the civilizations developed at various times in history and the ground of the science developed in such civilizations. When a civilization changes, its presuppositions change and are succeeded by others. According to this view, metaphysics is therefore a branch of the historical sciences.

See also Ayer, Alfred Jules; Carnap, Rudolf; Collingwood, Robin George; Entailment, Presupposition, and Implicature; Frege, Gottlob; Hilbert, David; Moore, George Edward; Questions; Russell, Bertrand Arthur William; Strawson, Peter Frederick.

Bibliography

Anscombe, G. E. M. *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson, 1959. Ch. II.

Asher, Nicholas, and Alex Lascarides. "The Semantics and Pragmatics of Presupposition." *Journal of Semantics* 15 (1998): 239–300.

Baker, A. J. "Presupposition and Types of Clause." *Mind* 65 (1956): 368–378.

Bar-Hillel, Yehoshua. "Analysis of 'Correct' Language." *Mind* 55 (1946): 328–340.

Black, Max. *Models and Metaphors*. Ithaca, NY: Cornell University Press, 1962. Ch. 4.

Black, Max. *Problems of Analysis*. Ithaca, NY: Cornell University Press, 1954. Chs. 2–3.

Cavell, Stanley. "Must We Mean What We Say?" *Inquiry* 1 (1958).

Collingwood, R. G. *An Autobiography*. London: Oxford University Press, 1939. See especially pp. 66–76.

Donagan, Alan. *The Later Philosophy of R. G. Collingwood*. Oxford: Clarendon Press, 1962.

Frege, Gottlob. "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik* 100 (1892): 25–50. Translated by Max Black as "On Sense and Reference." In *Translations from the Philosophical Writings of Gottlob Frege*, edited by P. T. Geach and Max Black. New York: Philosophical Library, 1952.

Geach, P. T. "Russell on Meaning and Denoting." *Analysis* 19 (1959): 69–72.

Grant, C. K. "Pragmatic Implication." *Philosophy* 33 (1958): 303–324.

Grice, Paul. "Presupposition and Conversational Implicature." In *Radical Pragmatics*, edited by Peter Cole. New York: Academic Press, 1981.

Griffiths, A. P. "Presuppositions." *Analysis* 15 (1955).

Hall, Roland. "Assuming: One Set of Positing Words." *Philosophical Review* 67 (1958): 52–75.

Hall, Roland. "Presuming." *Philosophical Quarterly* 11 (1961): 10–21.

Hampshire, Stuart. "On Referring and Intending." *Philosophical Review* 65 (1956): 1–13.

Hancock, Roger. "Presuppositions." *Philosophical Quarterly* 10 (1960): 73–78.

Hart, H. L. A. "A Logician's Fairy Tale." *Philosophical Review* 60 (1951): 198–212.

Kemeny, John G. *A Philosopher Looks at Science*. Princeton, NJ: Van Nostrand, 1959. Ch. 3.

Langford, C. H. "The Notion of Analysis in Moore's Philosophy." In *The Philosophy of G. E. Moore*, edited by P. A. Schilpp. Evanston, IL: Northwestern University Press, 1942.

Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.

Llewelyn, John E. "Collingwood's Doctrine of Absolute Presuppositions." *Philosophical Quarterly* 11 (1961): 49–60.

Llewelyn, John E. "Presuppositions, Assumptions, and Presumptions." *Theoria* 28 (1962): 158–172.

Lycan, William. "The Myth of the 'Projection Problem for Presupposition.'" *Philosophical Topics* 15 (1987): 169–175.

MacIver, A. M. "Some Questions about 'Know' and 'Think.'" *Analysis* 5 (1937–1938).

Nelson, Everett J. "Contradiction and the Presupposition of Existence." *Mind* 55 (1946): 319–327.

Nowell-Smith, P. H. "Contextual Implication and Ethical Theory." *PAS*, Supp. 36 (1962).

O'Connor, D. J. "Pragmatic Paradoxes." *Mind* 57 (1948): 358–359.

Quine, W. V. "Meaning and Inference." In *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953. See especially pp. 164–166.

Rynin, David. "Donagan on Collingwood on Metaphysics." *Review of Metaphysics* 18 (1964).

Sellars, Wilfrid. "Presupposing." *Philosophical Review* 63 (1954): 197–215.

Shwayder, D. S. "Self-Defeating Pronouncements." *Analysis* 16 (1956).

Shwayder, D. S. "Uses of Language and Uses of Words." *Theoria* 26 (1960).

Soames, Scott. "Presupposition." In *Handbook of Philosophical Logic*, edited by D. Gabbay. Dordrecht: Kluwer, 1989.

Stalnaker, R. C. "Pragmatic Presupposition." In *Pragmatics: A Reader*, edited by S. Davies. New York: Oxford University Press, 1991.

Strawson, P. F. *Introduction to Logical Theory*. London: Methuen, 1952.

Strawson, P. F. "On Referring." *Mind* 59 (1950): 320–344. Reprinted in *Logico-Linguistic Papers*. London: Methuen, 1971.

Strawson, P. F. "A Reply to Mr. Sellars." *Philosophical Review* 63 (1954): 216–231.

Stroll, Avrum. "The Paradox of the First Person Singular Pronoun." *Inquiry* 6 (1963): 217–233.

Urmson, J. O. "Parenthetical Verbs." In *Essays in Conceptual Analysis*, edited by A. Flew. London: Macmillan, 1956.

Wolterstorff, Nicholas P. "Referring and Existing." *Philosophical Quarterly* 11 (1961): 335–349.

Avrum Stroll (1967)

Bibliographic assistance by J. Ornstein (1967)

Bibliography updated by Benjamin Fiedor (2005)

PRESUPPOSITION

Consider the following famous example from Bertrand Russell.

(1) The present king of France is bald.

According to Russell, (1) is false because it asserts the existence of the present king of France. However, following P. F. Strawson (1952), a number of philosophers and linguists have maintained that, if there is no present king of France, an utterance of (1) fails to have a determinate truth-value—in Strawson's words, the question of whether (1) is true or false "does not arise." On this view, (1) therefore does not assert or even entail the existence of the present king of France but rather "presupposes" his existence.

THE RANGE OF PHENOMENA

Sentences like (1) are argued to presuppose the existence of a particular individual, but there are many other "presupposition" effects. It has been argued, for example, that factive verbs such as *know* and *regret* presuppose the truth of their complement clauses and that "certain aspectuals"—a class of verbs such as *quit* and *continue*—also presuppose certain actions having taken place (this class covers the example "Have you stopped beating your dog?"). It also appears that a number of modifiers introduce presupposition effects, for example *again, too, even,* and so forth. L. Karttunen (1973) argued that in propositional-attitude environments such as "Fred wants to sell his unicorn" it is presupposed that Fred believes he has a unicorn. A number of additional constructions that invoke presupposition effects have been explored, including those triggered by phonological stress. So, for example, if I say "I didn't go to the baseball game," it arguably presupposes that I went to some other kind of game.

PRESUPPOSITION VERSUS ENTAILMENT

The philosophical controversy surrounding presupposition comes in at the very beginning—determining whether these are genuine cases of presupposition or are merely cases of entailment. To illustrate, consider (2)–(4):

(2) Fred stopped washing the dishes.

(3) Fred didn't stop washing the dishes.

(4) Fred had been washing the dishes.

According to the presupposition thesis, both (2) and (3) presuppose (4). Hence, if (4) is false, then (2) and (3) must lack determinate truth-values. Alternatively, according to the entailment analysis, (2) entails (4). Should (4) be false, then according to the entailment analysis (2) will be false and (3) will be true. This dispute has all the makings of a stalemate, since it turns on speakers' intuitions about whether sentences lack genuine truth-values under the relevant conditions or are merely false. Indeed, Strawson (1964) came to doubt whether the matter could in fact be settled by "brisk little formal argument[s]" and offered that each view could be reasonable, depending on one's interests. Others have put more stock in brisk little formal arguments, notably D. Wilson (1975), who offered an extensive critique of the presuppositional analysis.

THE PROJECTION PROBLEM

One of the most interesting questions to surface is the so-called projection problem for presupposition, first observed by D. T. Langendoen and H. Savin (1971). This problem involves the question of what happens when a construction with a presupposition is embedded in more complex constructions (e.g., in propositional-attitude constructions or in the scope of negation). To illustrate, when (2) is negated, yielding (3), it continues to presuppose (4)—the presupposition is said to be projected. Other constructions, such as "doubts that," do not always project presuppositions, and still others (such as the "wants" case from Karttunen, discussed above) project something weaker than the original presupposition. The question is therefore whether projection presupposition is arbitrary or whether it obeys certain specific rules. Much subsequent work has attempted to articulate those "projection rules" (see Gazdar, 1979, Heim 1991, Karttunen 1973, and Soames 1979, 1982, for important examples).

SEMANTIC VERSUS PRAGMATIC PRESUPPOSITION

If one accepts that there are genuine instances of presupposition, there remains the question of whether presupposition is a reflex of semantics or pragmatics—that is, whether the presupposition follows from the meaning of the sentence or is merely part of the conversational background. R. Stalnaker (1974) gave several arguments in favor of the pragmatic alternative, including the interesting observation that, in a case like (5),

(5) If Eagleton hadn't been dropped from the Democratic ticket, Nixon would have won the election

there seems to be a presupposition that Nixon lost, although the effect is weak, and, in the right context or given appropriate information, that presupposition can

be overruled. This graded effect suggests that pragmatic phenomena are in play. Stalnaker also observed that the pragmatic alternative is useful in separating the question of entailment relations from the question of presupposition and in working out solutions to the projection problem. (But see Wilson 1975 for criticism of pragmatic accounts of presupposition.)

APPLICATIONS

The doctrine of presupposition remains somewhat controversial, but at the same time it has found interesting applications. For example, B. van Fraassen (1968, 1970) argued that presupposition might be employed in the treatment of the "liar paradox" and proposed that liar sentences are neither true nor false owing to a presupposition failure. Presupposition has also played an important role in work on the semantics of propositional attitudes, much of it extending from the work of Karttunen (1973). I. Heim (1992), for example, has updated the initial Karttunen analysis with features of Stalnaker's presuppositional analysis. Still other research (including unpublished work by Saul Kripke) has investigated the interplay of presupposition and the analysis of discourse anaphora.

See also Anaphora; Kripke, Saul; Liar Paradox, The; Philosophy of Language; Russell, Bertrand Arthur William; Strawson, Peter Frederick.

Bibliography

Gazdar, G. *Pragmatics.* New York: Academic Press, 1979.

Grice, P. "Presupposition and Conversational Implicature." In *Radical Pragmatics,* edited by P. Cole. New York: Academic Press, 1981.

Heim, I. "On the Projection Problem for Presuppositions." In *Pragmatics,* edited by S. Davis. New York: Oxford University Press, 1991.

Heim, I. "Presupposition Projection and the Semantics of Attitude Verbs." *Journal of Semantics* 9 (1992): 183–221.

Karttunen, L. "Presuppositions of Compound Sentences." *Linguistic Inquiry* 4 (1973): 169–193.

Langendoen, D. T., and H. Savin. "The Projection Problem for Presupposition." In *Studies in Linguistic Semantics,* edited by C. Filmore and D. T. Langendoen. New York: Holt, Rinehart and Winston, 1971.

Soames, S. "How Presuppositions Are Inherited: A Solution to the Projection Problem." *Linguistic Inquiry* 13 (1982): 483–545.

Soames, S. "A Projection Problem for Speaker Presuppositions." *Linguistic Inquiry* 10 (1979): 623–666.

Stalnaker, R. "Pragmatic Presuppositions." In *Semantics and Philosophy,* edited by M. Munitz and D. Unger. New York, 1974.

Strawson, P. "Identifying Reference and Truth-Values." *Theoria* 3 (1964): 96–118.

Strawson, P. *Introduction to Logical Theory.* New York: Wiley, 1952.

van Fraassen, B. "Presupposition, Implication, and Self-Reference." *Journal of Philosophy* 65 (1968): 136–152.

van Fraassen, B. "Truth and Paradoxical Consequences." In *The Paradox of the Liar,* edited by R. L. Martin. New Haven, CT: Yale University Press, 1970.

Wilson, D. *Presuppositions and Non-Truth-Conditional Semantics.* London: Academic Press, 1975.

Peter Ludlow (1996)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

PRICE – SEXTUS EMPIRICUS

8
volume

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON GALE™

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

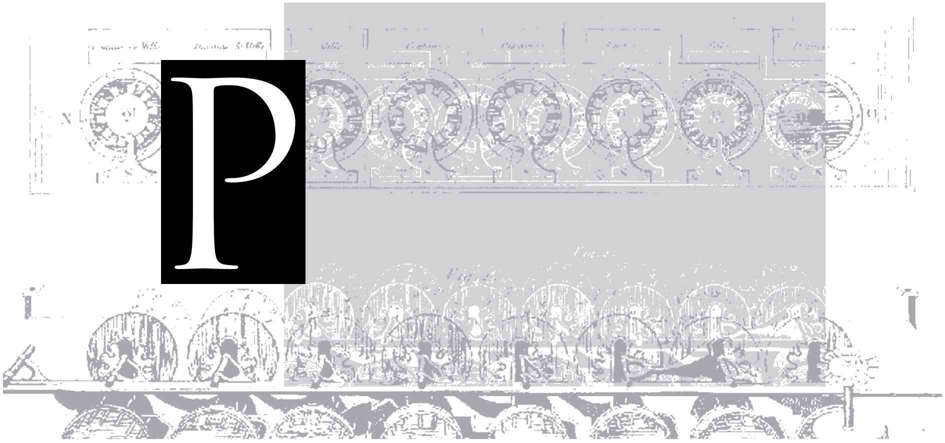

P

PRICE, RICHARD

(1723–1791)

Richard Price, a Welsh dissenting preacher, moral philosopher, and actuary, was born at Tynton, Llangeinor, Glamorganshire. His father, Rees, was a dissenting minister with extreme Calvinist opinions. Richard Price was educated at a number of different academies, finally entering Coward's Academy in London, where he remained for the years 1740–1744. He was ordained at the age of twenty-one and began his ministerial career as a domestic chaplain. He later served a number of London congregations, notably those at Stoke Newington, where he lived, and at the Gravel-Pit Meeting House in Hackney. Price was buried in the cemetery at Bunhill Fields; his friend Joseph Priestley preached the funeral oration.

In addition to his writings on moral philosophy, Price wrote with considerable influence on financial and political questions. His papers on life expectancy and on calculating the values of reversionary payments were instrumental in reforming the actuarial basis of the insurance and benefit societies of the time. His paper on the public debt is said to have led William Pitt, the prime minister, to reestablish the sinking fund to extinguish England's national debt. In his pamphlet *Observations on the Nature of Civil Liberty, the Principles of Government, and the Justice and Policy of the War with America* (London, February 8, 1776), Price defended the American cause. The widespread circulation and generally favorable acceptance of this work is said to have encouraged the American decision for a declaration of independence. Price had become friendly with Benjamin Franklin during the latter's stay in London, and in 1778 the Continental Congress moved to grant Price American citizenship if he would come to America and serve as an adviser on the management of American finances. He was grateful for the invitation but did not accept it. Price also regarded the French Revolution with approval, which he expressed, along with an appeal for reform in England, in his *Discourse on the Love of Our Country* (1789). Edmund Burke's *Reflections on the Revolution in France* (1790) was written in reply.

Price is also the author of *Four Dissertations:* I. "On Providence"; II. "On Prayer"; III. "On the Reasons for expecting that virtuous Men shall meet after death in a State of Happiness"; IV. "On the Importance of Christianity, the Nature of Historical Evidence, and Miracles" (London, 1767). In the fourth of these dissertations Price criticized David Hume's "Of Miracles." Hume was grateful for the civility with which Price argued, and he wrote

to Price that the light in which he put this controversy was "new and plausible and ingenious, and perhaps solid. But I must have some more time to weigh it, before I can pronounce this judgment with satisfaction to myself."

MORAL PHILOSOPHY

Price's contribution to moral philosophy is *A Review of the Principal Questions in Morals* (London, 1758; corrected editions in 1769 and 1787). Price criticized the moral-sense doctrines of Francis Hutcheson in order to clear them away and make room for an account of immutable right and wrong, derived from Samuel Clarke.

Price says that we may have three different perceptions concerning the actions of moral agents. We may notice whether they are right or wrong, whether they are beautiful or ugly, and whether they are of good or ill desert. By talking of perceptions here, he shows that he has accepted the premise, of Lockean origin, that all knowledge is to be accounted for as some kind of perception by one of our faculties. Thus, Price's first question, "How do we know right?," is treated as a search for the faculty by means of which we obtain our ideas of right and wrong. He considers Hutcheson's answer that our moral ideas come to us by the way of a moral sense, and he understands Hutcheson to be claiming that this sense is "a power within us, different from reason; which renders certain actions pleasing and others displeasing to us."

Price objects to this doctrine because of certain consequences that he believes are implied by it. Our approval and disapproval of actions appear to depend on the way our minds work or, to carry the matter back a step, on the way God has made them to work. Thus, our judgments of right and wrong depend on the mere good pleasure of our Maker, who created us in a certain way. But if he had pleased, he might have made us to be pleased or displeased by quite different actions, even actions contrary to those that now please and displease us. Thus, right and wrong would be only matters of taste, only a certain effect in us, and nothing in actions themselves.

For his part, Price is convinced that morality is equally unchangeable with all truth and that right and wrong are real characteristics of actions and not mere sensations derived from the particular way in which our minds are framed. To show the immutability of right and wrong, Price argues that these ideas are derived not from a special sense but from the understanding. As Price sees it, the only debatable issue in morals is not what actions are right and wrong but what is the faculty by which we discern right and wrong.

Price prefaces his argument for regarding the understanding as our moral faculty with the preliminary claim that the understanding is a source of new ideas. He objects to interpreting John Locke as saying that sensation and reflection are the sources of all our ideas. Price argues that Locke may have meant only that all our ideas are ultimately grounded on ideas derived from sensation and reflection. Thus, Price makes room for certain new ideas that may arise as the understanding compares the objects of thought and judges them. Some of these new ideas are solidity, inertia, substance, accident, duration, space, cause or power, entity, possibility, and actual existence.

Price locates these new ideas in a revised classification of simple ideas. He divides simple ideas into those implying nothing real outside the mind and those that denote real and independent existence distinct from sensation. The first class of simple ideas consists, on the one hand, of tastes, smells, and colors and, on the other, of such notions as order, happiness, and beauty. The second class of simple ideas has three subclasses: the real properties of external objects, such as figure, extension, and motion; the actions and passions of the mind, such as volition, memory, and so on; and those new ideas noted above which arise as the understanding considers the ideas it has been supplied with. It is important to note that Price does not regard the second class of simple ideas as constructions of the mind. The real properties of external objects are in the objects, and such new ideas as cause, duration, and space are of properties in a real world.

Armed with his reclassification of simple ideas, Price is now prepared to locate our ideas of moral right and wrong in the scheme and thus establish that they are perceptions of the understanding. Price first considers the question of whether moral right and wrong are simple ideas. He declares that they must be, for we cannot give definitions of them that are more than synonymous expressions. It is Price's recognition of this point which has led contemporary students to declare him one of the first to recognize the naturalistic fallacy, although he does not use that term. Having established that our ideas of right and wrong are simple ideas, Price then locates them in his scheme as two of those new ideas which arise in the understanding.

Hutcheson had simply assumed that if right and wrong are immediately perceived, they must be perceptions of an implanted sense. But the question of how we perceive these ideas may be settled by simply considering the nature of our own perceptions.

> Let anyone compare the ideas arising from our *powers of sensation,* with those arising from our *intuition of the natures of things,* and enquire which of them his ideas of right and wrong most resemble.… It is scarcely conceivable that anyone can impartially attend to the nature of his perceptions, and determine that when he thinks gratitude or beneficence to be *right,* he perceives nothing *true* of them, and *understands* nothing, but only receives an impression from a sense.

Price notes that some impressions of pleasure or pain, satisfaction or disgust, generally attend our perceptions of moral right and wrong; the proponents of a moral sense may have confused these impressions with our actual perceptions of right and wrong.

But there is an assumption in Price's own system on which much depends and for which he offers insufficient argument. He tells us that "all actions undoubtedly have a *nature.* That is, *some character* certainly belongs to them, and somewhat there is to be *truly* affirmed of them." It is the task of the understanding to perceive these truths. Price regards actions in this way because it enables him to say that their rightness or wrongness is in them, not in the mind of the person judging the actions, but apart from noting the advantage to his own moral philosophy, Price offers no justification for the claim that actions have natures. It is unfortunate that he does not, for he rests his contention that morality is eternal and immutable on this claim.

When Price turns to our ideas of the beauty and deformity of actions, the second kind of perception of actions which he promised to account for, he finds that these perceptions are feelings of delight or detestation which may accompany our perceptions of the rightness or wrongness of actions. These feelings of delight and detestation are the effects on us of the actions we consider, and it is very likely that they arise from an arbitrary structure of our minds, which may be called a sense. Price allows that there is a distinction between noting that an action is right and approving it. We are made, however, in such a way that we cannot perceive an action to be right without approving it, for in humans it is necessary that the rational principle, or the intellectual discernment of right and wrong, should be aided by instinctive determinations. When these feelings of the heart support the perceptions of the understanding, we are provided with the motivation for moral behavior. Here Price agrees with Hutcheson, pointing out that he has never disputed that we owe much to an implanted sense and its determinations. He means to resist only the claim that we owe our knowledge of right and wrong to such a sense.

Our ideas of the good and ill desert, the third sort of perception concerning actions which Price notes, carry the mind to the agent. He finds that we cannot but love a virtuous agent and desire his happiness above that of others. Quite apart from any advantage which we may gain from someone else's virtuous behavior, we have an immediate approbation of making the virtuous happy and of discouraging the vicious.

Price distinguishes between abstract and practical virtue. Abstract virtue denotes "what an action is independently of the sense of an agent; or what, in itself and absolutely, it is right such an agent, in such circumstances, should do." But Price recognizes that the actual practice of virtue depends on the opinion of the agent concerning his actions. Thus, an agent may be mistaken about his circumstances but sincere about what he believes he ought to do. In this respect practical virtue may diverge from abstract virtue but be no less obligatory insofar as the agent acts from a consciousness of rectitude. The ideal state of affairs is a correspondence of practical virtue with abstract virtue. Its achievement depends on the liberty and intelligence of the agent. These constitute the agent's capacity for virtue, and intention gives virtue actual being in a character. Price takes a short way with the question, "Why be moral?" "The knowledge of what is right, without any approbation of it, or concern to practise it, is not conceivable or possible. And this knowledge will certainly be attended with *correspondent, actual practice,* whenever there is nothing to oppose it." Why a person chooses to do what he knows he should do is a question "which need not and should not be answered."

Benevolence is not the sole virtue. We also have duties to God and to ourselves, and there is room for many other sorts of good behavior, such as veracity, sincerity, and gratitude. As a measure of virtue Price offers the rule that "the virtue of an agent is always less in proportion to the degree in which natural temper and propensities fall in with his actions, instinctive principles operate, and rational reflexion on what is right to be done, is wanting."

Price discusses at length the relation of morality to the divine nature. Just as moral right and wrong are independent of man's mind, they are also absolutes for God. Were this not so, there would be no sense in which God's will could be good.

FREEDOM OF THE WILL

Price and Priestley published a set of letters as *A Free Discussion of the Doctrines of Materialism and Philosophical Necessity* (London, 1778). The correspondence had its origin in Price's criticism of Priestley's *Disquisitions Relating to Matter and Spirit.* The letters cover the nature of matter, the human mind, the mortality of the soul, the essence of the deity, and the doctrine of necessity. The last topic is the one that is treated in the most interesting way. Priestley contended that there can be no human liberty because "liberty" must mean someone's willing without a motive, which he regards as impossible. Price enlarges on the account of liberty that he offered in *A Review of the Principal Questions in Morals.* He argues that human agents are not physical objects but unique entities capable of self-determination. Consider the difference between a man who is dragged by a superior force and a man who follows a guide for a reward. Both of these examples may be certainties, but having different foundations, they are of totally different natures. "In both cases the man might in common speech be said to *follow*; but his following in the one case, however certain in event, would be *his own* agency: In the other case, it would be the agency of another. … In the one case, superior power moves him: In the other he moves himself."

See also Burke, Edmund; Clarke, Samuel; Hume, David; Hutcheson, Francis; Liberty; Locke, John; Moral Principles: Their Justification; Moral Sense; Priestley, Joseph; Properties; Responsibility, Moral and Legal.

Bibliography

Price's *Works* were published in 10 volumes (London, 1816), with a memoir of his life by W. Morgan. *A Review of the Principal Questions in Morals* has been published with a critical introduction by D. D. Raphael (Oxford, 1948). This is a reprint of the third edition (1787) with an appendix and "A Dissertation on the Being and Attributes of the Deity."

For biography, see Carl B. Cone, *Torchbearer of Freedom, the Influence of Richard Price on Eighteenth Century Thought* (Lexington: University of Kentucky Press, 1952). Other works on Price include Joseph Priestley, *A Discourse on the Occasion of the Death of Dr. Price* (London, 1791); Leslie Stephen, *History of English Thought in the Eighteenth Century* (London, 1876; 2nd ed., London, 1902); and Roland Thomas, *Richard Price* (London: Oxford University Press, 1924).

Elmer Sprague (1967)

PRIESTLEY, JOSEPH
(1733–1804)

Joseph Priestley, the English scientist, nonconformist minister, educator, and philosopher, was born at Birstall, Yorkshire, the son of a cloth dresser. His mother died in 1740, and in 1742 Priestley was adopted by a childless well-to-do aunt, Mrs. Keighley, a convinced but unbigoted Calvinist. A sensitive child, Priestley suffered greatly because he could not convince himself that he had experienced the "new birth" essential, on the Calvinist scheme, for his salvation. As a result of these childhood miseries Priestley was left, he tells us, with "a peculiar sense of the value of rational principles of religion" as opposed to the "ignorance and darkness" of Calvinism.

Until the age of sixteen Priestley was educated at a conventional grammar school. For the next three years, his health being too poor for regular studies, he in large part educated himself, reading his way into mathematics, physics, and philosophy and undertaking the study of European and Middle Eastern languages. In 1752 his health improved and he entered Daventry Academy, a university-type institution set up by nonconformists because Oxford and Cambridge would not admit nonconformists to a degree.

At Daventry the emphasis was on free discussion, and the curriculum was considerably broader than at Oxford or Cambridge. Priestley was introduced to David Hartley's *Observations on Man* (1749) and was at once—and permanently—converted to Hartley's general outlook. The simplicity and generality of Hartley's associationist psychology appealed to Priestley's maturing scientific instincts; it provided a theoretical foundation for his belief in perfectibility through education; and it offered a psychological alternative to the doctrine of free will, which Priestley's reading of Anthony Collins's *Philosophical Inquiry concerning Human Liberty and Necessity* (1714) had already caused him to reject.

In 1755 Priestley entered the ministry, taking over a decaying congregation at Needham Market, Suffolk. Stammering and unorthodox, he was not a success as a minister. He moved in 1758 to a more sympathetic but equally impoverished congregation at Nantwich in Cheshire. In an attempt to increase his income he set up a school where, perhaps the first to do so, he taught experimental science with the help of an "electrical machine" and an air pump.

Appointed in 1761 as "tutor of the languages" at Warrington Academy in Lancashire, Priestley taught oratory, literary criticism, grammar, history, and law, as well

as languages. Characteristically, on all these latter topics Priestley developed ideas that he sooner or later published. *The Rudiments of English Grammar* (1761), many times reprinted, is typical of his innovating boldness, insofar as he tried to simplify English syntax by removing from it the complications introduced by classically trained grammarians. His *A Chart of Biography* (1765) and *A New Chart of History* (1769) were even more enthusiastically received; they won for him not only his sole academic distinction, the doctorate of laws of the University of Edinburgh, but also his fellowship of the Royal Society.

Priestley's days of relative isolation were now over. In 1762 he married an ironmaster's daughter, Mary Wilkinson, an intelligent woman with a sense of humor and considerable force of character—qualities she was to need in the years to come. His duties at Warrington left him free to visit London for a month each year, where he came into contact with an active group of scientists, philosophers, and political thinkers, including Benjamin Franklin and Richard Price. Franklin encouraged Priestley's project of writing a history of electrical experiments. The work that resulted, *The History and Present State of Electricity, with original Experiments* (1767), is a notable contribution to the history of science. Describing a number of important original experiments, it is also in some respects the most theoretically adventurous of Priestley's scientific works. It contains as well Priestley's reflections on the use of hypotheses in scientific procedures as a guide to experimentation.

EDUCATION AND GOVERNMENT

Like many of his fellow dissenters, Priestley was greatly interested in educational reform. Education had, he thought, thus far concentrated unduly on the needs of the clergy. His *An Essay on a Course of Liberal Education for Civil and Active Life* (1765) is a plea for a curriculum that should be suitable for men of affairs, emphasizing history and public administration rather than the classical languages. Priestley did much to encourage the teaching of history in the nonconformist academies. A set of lectures that he delivered at Warrington (published in 1788 as *Lectures on History and General Policy*) provided not only the academies but also the new American colleges with a text suitable for their needs; it was, indeed, recommended even at Cambridge. It is a summary account of the main historical sources, with an emphasis on commerce, law, and administration, rather than a historical textbook of the ordinary kind.

Priestley's political theory was closely related to his interest in education and his experience as a member of a minority group. In an appendix to his *Essay on a Course of Liberal Education* he developed an argument against the introduction of a state system of education, which would inevitably, he thought, favor the *status quo* and produce a quite undesirable uniformity of conduct and opinion. Like John Stuart Mill after him, Priestley gloried in diversity; uniformity, he said, is "the characteristic of the brute creation."

These reflections were more fully worked out in *An Essay on the First Principles of Government* (1768), which bears the subtitle *On the Nature of Political, Civil and Religious Liberty.* For Priestley, the preservation of civil liberty was the crucial political issue. Deciding who should participate in government—who, that is, should possess political, as distinct from civil, liberty—was, he thought, a practical matter, to be settled by considering what groups in the community are most likely, if they possess political power, to act for the greatest happiness of the greatest number. Such groups remain entitled to power only as long as they continue so to act. Legislation, on Priestley's view, should be kept to the minimum. What that minimum is cannot be determined a priori but only as a result of political experiment. But we can see at once, Priestley thought, that legislation that restricts civil and religious liberty is bound to be against the interests of the community. Unlike most nonconformist upholders of toleration and unlike his master John Locke, Priestley was uncompromising on this point; he upheld unbounded liberty of expression even to atheists and Roman Catholics.

In Priestley's eyes, the noblest of occupations was that of the clergyman, not the lecturer, and in 1767 he accepted a call to Mill Hill, Leeds, a congregation to whom his religious views were exceptionally congenial. The years Priestley spent at Mill Hill were extremely important in his development; his salary, although small, sufficed for his needs, and his duties left him considerable leisure.

UNITARIANISM

Priestley had long before abandoned both the doctrine of the atonement, on which he wrote critically in *The Scripture Doctrine of Remission* (1761), and orthodox Trinitarianism. Now he took what was to be the final step in his transition from Calvinism to Unitarianism. Christ, he argued, although the Messiah, was a man, and not even a perfect man. Priestley's subsequent theological writings were in large part an attempt to prove—most maturely in

his *History of Early Opinions concerning Jesus Christ* (1786)—that Unitarianism was the doctrine of the early church. He defended his unorthodoxies both against clerical attack, as in his *Letters to Dr. Horsley* (1783–1786), and, as in his *Letters to a Philosophical Unbeliever* (Pt. I, 1780; Pt. II, 1787), against those who, like Edward Gibbon, could not understand why Priestley did not make a complete break with Christianity. Priestley valued his theological writings above all his other work. A firm belief in Providence is everywhere evidenced in his writings. Few men have committed themselves so often and so absolutely to the doctrine that "all is for the best in the best of all possible worlds," although he also believed that the future world could—and therefore would—be better.

SCIENTIFIC ACHIEVEMENT

It was as a scientist that Priestley won his international reputation. He published in 1772 what was intended to be the second section of a general history of science, *The History of the Present State of the Discoveries relating to Vision, Light and Colours*; but this work, invaluable though it still is to historians of science, did not arouse a great deal of interest. Priestley therefore abandoned his large-scale historical project and concentrated instead on chemistry. His first chemical publication, in 1772, was of an unusually practical character: It described a method of producing "mephitic julep," or soda water. But it was the paper "On Different Kinds of Air," which he read in that same year to the Royal Society," that at once established his reputation as a chemist. In 1774 he prepared the first edition of *Experiments and Observations on Different Kinds of Air*; this he republished in a series of editions, with important changes in contents, in method of organization, and even in title, until 1790.

By the end of that period Priestley, following up the work of Joseph Black and Henry Cavendish, had considerably enlarged our knowledge of the chemical properties of gases. He differentiated between nine gases, of which only three had previously been known to science, and described a method of collecting them. Of particular importance was his preparation of "dephlogisticated air" (oxygen), which he produced on August 1, 1774, by heating red mercuric oxide. It then became clear that air was not an element. Priestley went on to examine the properties of oxygen; in a series of chemicobiological experiments he brought out its importance for animal life.

As a resourceful experimenter, using simple and economical methods, Priestley has had few equals. But it was left to others, to Cavendish and Antoine Lavoisier, to appreciate the theoretical significance of his work. Priestley had isolated oxygen and had observed its importance in combustion; he had passed a spark through a mixture of hydrogen and oxygen and had noticed that dew was formed. Yet his last scientific work (1800) bore the title *The Doctrine of Phlogiston established and that of the Composition of Water refuted.* Although he had himself carried out important quantitative experiments, he did not appreciate the significance of the quantitative considerations by which Lavoisier overthrew the phlogiston theory.

PHILOSOPHY

Much of Priestley's most important scientific work was carried out at Shelburne, where from 1772 until 1780 he acted as "librarian and literary companion" to the Earl of Shelburne. During these same years Priestley embarked upon his most substantial metaphysical works. He began in 1774 with *An Examination of Dr. Reid's Inquiry into the Human Mind on the Principles of Commonsense, Dr. Beattie's Essay on the Nature and Immutability of Truth, and Dr. Oswald's Appeal to Commonsense on Behalf of Religion,* commonly referred to as *An Examination of the Scotch Philosophers.* This is a vigorous polemic, which sets out to demonstrate the superiority of Hartley's psychology to the philosophy of the Scottish commonsense school, a philosophy that Priestley thought obviously reactionary insofar as it substituted for the simple Locke-Hartley theory of mind "such a number of independent, arbitrary, instinctive principles that the very enumeration of them is really tiresome." All the so-called instinctive beliefs of common sense can, Priestley set out to show, be derived from the operations of associative principles working on the materials provided by sensation. He came to regret in later life the tone of this publication but never its doctrines.

MATERIALISM. Hoping to make Hartley's views better known, Priestley published an abridged version of Hartley's *Observations on Man* in 1775 as *Hartley's Theory of the Human Mind on the Principle of the Association of Ideas.* In his preface, Priestley somewhat tentatively suggested that all the powers of the mind might derive from the structure of the brain. Even as a suggestion this created a considerable uproar, but Priestley was not to be intimidated by clerical clamor. Convinced that materialism was the natural metaphysical concomitant of Hartley's associative psychology, he set out, therefore, in his *Disquisitions Relating to Matter and Spirit* (1777) to demonstrate that materialism was theologically, scientifically, and metaphysically superior to orthodox dualism.

On the theological side, materialism had commonly been objected to on the ground that it is incompatible with immortality. Man, Priestley replied, is not "naturally" immortal; he is immortal only because, as we know from revelation, God chooses to resurrect him; this resurrection is of the body and therefore also of the body's mental powers. As for the commonplace metaphysical objections to materialism, these are based, according to Priestley, upon an untenable conception of matter as being by nature inert and therefore incapable of exerting mental activity. To such a concept of matter Priestley opposed the physical theories of his friend and fellow scientist John Michell and the Jesuit mathematician Roger Boscovich. Material objects, on their view, are centers of force; if this is the nature of matter, Priestley argued, there is no good reason for denying that mental operations are part of the activity of a material object. On the other hand, there are very good reasons for objecting to the traditional dualism, which is quite incapable of explaining how mind and body can enter into any sort of relationship.

DETERMINISM. Priestley had been a determinist long before he became a materialist, but not until 1777, in *The Doctrine of Philosophical Necessity Illustrated,* did he fully present his case against free will; indeed, even then he thought of himself as supplementing Thomas Hobbes, Collins, David Hume, and Hartley with illustrations rather than as working out an entirely independent position. The doctrine of free will, he argued, is theologically objectionable because it cannot be reconciled with the existence of an all-seeing Providence; from a metaphysical standpoint, it makes human actions quite unintelligible, and ethics has no need of it. As a basis for our everyday moral judgments, the distinction between acting voluntarily and acting under compulsion is certainly important, but this distinction does not, according to Priestley, rest upon a metaphysical conception of free will.

Priestley's metaphysical unorthodoxies considerably disturbed his old friends, provoking a candid but good-tempered correspondence with Richard Price, published in 1778 as *A Free Discussion of the Doctrines of Materialism and Philosophical Necessity Illustrated.* Developing his views on the relation between moral judgments and determinism, Priestley admitted that the determinist will prefer to avoid describing people as blameworthy or praiseworthy. He will say of them, rather, that they have acted, or have not acted, from good principles—from principles, that is, that are conducive to the general happiness. But the determinist's different method of describing moral conduct has, Priestley thought, no practical consequences, and if determinism is in some respects inconsistent with everyday usage, this is even more true of libertarianism.

LATER YEARS

There was a real risk, however, that Priestley's reputation for materialism might endanger the earl of Shelburne's political ambitions. Perhaps for this reason Priestley and Shelburne parted amicably in 1780, when Priestley, refusing Shelburne's offer of a post in Ireland, took up residence in Birmingham. There he had a circle of congenial friends who were prepared to offer him financial as well as intellectual support. He became a member of the Lunar Society, with which were associated men of the caliber of Erasmus Darwin and James Watt, and he enjoyed the friendship and help of the scientifically minded potter Josiah Wedgwood, who supplied him with apparatus specifically designed for his chemical experiments. Much of Priestley's scientific work in this period, under Alessandro Volta's influence, conjoined his two main scientific interests: electricity and gases. He examined the effect of passing electrical sparks through a variety of gases and studied their thermal conductivity.

He was by no means unsympathetic to the laissez-faire sociopolitical attitude of Birmingham industrialists. In *Some Considerations on the State of the Poor in General* (1787) he strongly criticized the poor laws and elsewhere opposed apprenticeship laws and laws for regulating interest rates. On his view, any sort of social welfare legislation "debased the very nature of man" by treating him as someone who had to be provided for. Although Priestley warmly supported schemes for cooperative insurance against hardship, he was opposed to any legislation that might diminish independence or increase the power of the state over individuals.

POLITICAL RADICALISM. In general terms, Priestley's life at Birmingham was a continuation and development of his earlier activities; theological controversy continued to be his main interest. But one event transformed his life and modified his political attitudes: the French Revolution. Reacting to that revolution, the British government became steadily more intolerant and conservative, and Priestley came to think that extensive political innovations were a necessary condition for the preservation of civil liberty. He moved toward political radicalism of the nineteenth-century kind in his *Letters to Edmund Burke occasioned by his Reflections on the Revolution in France* (1791) and in the anonymously published *A Political Dia-*

logue on the General Principles of Government (1791). He had formerly been accustomed to describe himself as "a Unitarian in religion but a Trinitarian in politics" because he had accepted the view that liberty rested on the balance between king, Commons, and House of Lords. He now came to feel that there should be but one source of political power, the will of the people as it would be represented in a reformed House of Commons.

On July 14, 1791, the Friends of the Revolution organized a dinner at Birmingham (Priestley was not present) in order to commemorate the fall of the Bastille. This was the last straw. With the encouragement, it would seem, of the authorities, an angry mob attacked the nonconformist chapels, then turned their attention to Priestley's house, destroying his books and furniture. Priestley was persuaded by his friends to leave Birmingham for London where he was, however, shunned by his scientific colleagues.

LIFE IN AMERICA. For some years, Priestley had been contemplating migration to the United States, where his three sons had already gone. In 1794 he left for New York and finally settled in Northumberland, Pennsylvania. There, still supported by his old friends, he continued to experiment and to write, mainly on theological questions.

He was disappointed, however, by the orthodoxy of the American clergy and alarmed by the growth of intolerance in the United States. Although he took no part in politics, he wrote an uncompromising exposition of his political and religious views in *Letters to the Inhabitants of Northumberland* (1799). There was talk of his being deported under the Aliens Act, but John Adams would not permit the application of the act to "poor Priestley." With the election of Thomas Jefferson to the presidency, Priestley was not only secure but also at last on good terms with authority. Jefferson consulted him on educational questions, and Priestley's *Socrates and Jesus Compared* (1803) precipitated Jefferson's "Syllabus" of his religious beliefs. Another of Priestley's works, *The Doctrines of Heathen Religion Compared with those of Revelation* (1804), awoke in Adams an enthusiasm for comparative religion. Priestley's last years, from 1801 until his death, were marred by ill health and bereavements, but his diversified intellectual interests remained with him until the end.

See also Boscovich, Roger Joseph; Collins, Anthony; Darwin, Erasmus; Determinism, A Historical Survey; Determinism and Freedom; Franklin, Benjamin; Hartley, David; Hobbes, Thomas; Hume, David; Jefferson, Thomas; Libertarianism; Locke, John; Materialism; Mill, John Stuart; Price, Richard.

Bibliography

Priestley's scientific writings have never been collected. The standard edition of his philosophical, theological, and miscellaneous writings is *The Theological and Miscellaneous Works of Joseph Priestley,* edited by John Towill Rutt, 25 vols. in 26 (London: G. Smallfield, 1817–1832). John Arthur Passmore, *Joseph Priestley* (New York, 1965) includes a short list of Priestley's writings, an introductory essay, and selections from his major works.

See also Thomas Henry Huxley, "Joseph Priestley," in *Science and Culture* (London, 1881); Thomas Edward Thorpe, *Joseph Priestley* (New York: Dutton, 1906); Edgar Fahs Smith, *Priestley in America, 1794–1804* (Philadelphia: P. Blakiston's Son, 1920); Wallace Ruddell Aykroyd, *Three Philosophers: Lavoisier, Priestley and Cavendish* (London: Heinemann, 1935); Arthur Handley Lincoln, *Some Political and Social Ideas of English Dissent, 1763–1800* (Cambridge, U.K., 1938); Francis Edward Mineka, *The Dissidence of Dissent* (Chapel Hill: University of North Carolina Press, 1944); Stephen Edelston Toulmin, "Crucial Experiments: Priestley and Lavoisier," in *Journal of the History of Ideas* 18 (2) (1957): 205–220; and the detailed bibliographies in the *Dictionary of National Biography* and the *Dictionary of American Biography.*

John Passmore (1967)

PRIMARY AND SECONDARY QUALITIES

The distinction between "primary and secondary qualities," first stated and thus named by Robert Boyle, received its classical formulation in John Locke's *Essay*. There Locke states that apart from ordinary causal properties or "powers," material objects possess five primary qualities—extension (size), figure (shape), motion or rest, number, and solidity (impenetrability)—and many secondary qualities, such as color, taste, smell, sound, and warmth or cold. This distinction was made in the context of representative realism; that is, it was presupposed that the qualities of objects are quite distinct from, and are in fact causes of, "ideas" (representations or sensa), which are the only immediate objects of sensory awareness. The basis of the distinction was twofold. First, perceived size, shape, motion, number, and solidity are ideas caused by and exactly resembling the corresponding primary qualities of objects; perceived color, taste, smell, sound, and so on are caused by, but do not resemble, the corresponding secondary qualities. Second, the primary qualities are inseparable from matter and are found in every part of it;

the secondary qualities are not true qualities of matter but are merely powers in the objects to produce sensory effects in us by means of the primary qualities in their minute parts. Thus, red as experienced (idea or sensum) is the effect of the secondary quality red, which is merely the power possessed by a special texture or surface structure of the object to reflect certain light frequencies and to absorb others.

This formulation is rather clumsy, and since George Berkeley the custom has been to apply the first part of the distinction to the qualities of the ideas or sensa. The primary or spatiotemporal qualities of these data may then be said to characterize the object as well, for instance, the sensum is square and so is the object; but the secondary qualities are said not to characterize the object at all except in a derivative way, for instance, the sensa may be red and fragrant, but the object itself is intrinsically neither colored nor scented; it is red and fragrant only in the secondary sense that it causes the appropriate data of color and smell in the percipient. The doctrine is thus essentially the same as Locke's, but the language is slightly different. This second formulation will be used here.

Though Boyle and Locke invented and popularized the distinction and the terminology of primary and secondary qualities, the distinction dates back in principle to Democritus, who said that sweet and bitter, warm and cold, and color exist only by convention (*νόμῳ*), and in truth there exist only the atoms and the void (Fr. 9, Diels and Kranz). The distinction was revived by Galileo Galilei and accepted by René Descartes, Isaac Newton, and others.

ARGUMENTS FOR THE DISTINCTION

RELATIVITY AND MEASUREMENT. The relativity argument is the most important one: Secondary qualities are affected by the condition of our sense organs and nervous system, by our distance from the object or its motion relative to us, by the lighting or by such intervening media as fog. Since secondary qualities thus vary according to, and depend for their nature on, factors quite external to the physical object, they cannot be intrinsic properties of it. This point was elaborated by Locke in various examples, two of which follow: (*a*) If one takes three bowls of water, one judged hot, one judged cold, and one judged medium, and places one hand in the hot water and the other hand in the cold, and then transfers both hands to the middle bowl, the water in that bowl will feel hot to the hand that has been in the cold water and cold to the hand that has been in the hot water. But since it cannot be both hot and cold, hot and cold are therefore not intrinsic properties of the water. (*b*) Marble is not colored in the dark; its color appears only in the light. But presence or absence of light cannot alter its real properties, so that the perceived color cannot be included among them.

If we grant the position of representative realism that hot, cold, and color, as experienced, are qualities of ideas or representations, then it is plausible to suppose on these grounds that they do not also characterize objects or resemble properties of objects. (Locke does not always make it clear that representative realism is to be presupposed). But this claim is apparently open to the insuperable objection, stated by Berkeley, that the primary qualities also vary: The object's apparent shape or size varies just as much as its color or sound. This would mean that shape and size as perceived do not characterize objects or resemble the actual properties of the object, thus subverting the whole basis of the distinction. That Locke did not see this may have been partly because he felt that he had to argue against the commonsense assumption that all sensible qualities characterize objects, and partly because the belief that primary qualities characterize all matter was apparently guaranteed by the physics of his day.

Although this objection is valid against Locke's position, it does not destroy the distinction between the primary and secondary qualities, which it is natural to recast and support by a revised relativity argument. This new point is that, in contrast with the secondary qualities, the main primary qualities—shape, size, and motion—can all be measured (solidity cannot, but it is dubious anyhow, in that most physical objects, even atoms, are far from solid or impenetrable; number, whether there is one object or two, seems scarcely a quality at all; strictly also in the case of shape, what is measured are various dimensions—diameters, angles, and so on—of the object, and supporters of the distinction must maintain that these are the differentiae of the shape). A plate may look elliptical, but by measuring its diameters and seeing that they are equal, we can establish that it is round; one man may look taller than another, but their relative heights may be settled by measurement, as can the speed of objects relative to the earth. The measured size and shape of a plate may thus be held to characterize it, and the sensible size and shape may agree with and resemble them, so that one can say that size and shape (and motion) are primary. Nevertheless, only in favorable circumstances does a given primary sensible quality also characterize the object (for instance, both object and sensum are round); otherwise, there is only a projective relationship, as between elliptical sensum and round object.

Measurement is objective and does not vary significantly because it is an operation that depends on the coordination of a number of separate perceptions and that may be performed by a number of different persons. Consequently, variations due to the measurer on any particular occasion are compensated for and do not affect the final result, and the various actions confirm that one is not simply establishing the qualities of representations. Measurement also leads to conclusions regarding the dimensions and positions of objects in physical space that can be verified by further activities or operations, such as fitting the objects together, moving one's hand between them, rolling an object to confirm that it is round, and so on. By contrast, the variation found in the sensory qualities seems to be caused by their being simply the content of one single act of perception limited to one person at one time.

If all this is so, the list of primary qualities must be somewhat amended. Shape, size, and motion remain, but one should substitute mass for solidity. Temperature is more difficult: Since it can be measured, it seems at first primary. But what is measured is the property of causing expansion in fluid or metal; this property in no way resembles felt warmth, and in physical theory it is a form of energy. Hence, temperature should not be regarded as a separate primary quality. Material objects do, of course, possess many other properties—causal and dispositional ones, for example—as Locke realized by his doctrine of "powers," but part of the distinction is that only the primary ones are intrinsic (that is, possessed without reference to other objects) and that all such powers are ultimately due to patterns of primary qualities. Even so, the distinction would have difficulty in coping with some intrinsic "scientific" properties, such as energy or electric charge.

Apart from this, various objections have been made to the distinction in terms of measurement. First, measured motion and size must be stated in terms of some standard, such as a yard or meter; hence, they are purely relational and are not intrinsic properties of the object. But one can reply that it is only the description or labeling of the measurement that is thus relational; the motion or extension labeled, which is actually measured, seems intrinsic to the object.

Second, since colors and sounds may be measured, are they not also primary? But this objection seems based on a misunderstanding of the processes of measurement, for one way of "measuring" color might be to compare a given shade with a standard on a shade card; but that would be like comparing the sensible size of two objects, not measuring them. Proper measurement goes beyond this kind of sensory experience, and even if one gives the shade a number, one cannot calculate with the results as one can with the dimensions of objects. Normally, however, measurement of colors or sounds is either the measurement of the amplitudes or lengths of light waves or sound waves, or a mixture of wave measurement and the comparison of experiences. If one brings up a decibel meter and says that the sound to be measured is 80 decibels, it is the amplitude of the sound waves that is ultimately responsible for the movement of the pointer to 80. It should be noted, however, that the logarithmic scale is used because of a characteristic of human ears—that experienced loudness is related logarithmically to wave amplitude.

Third, measurement is a perceptual process—at least it relies on and largely uses perception—so it may be only producing various correlations of sensa and never getting through to the supposed properties of material objects at all. This objection is made from the point of view of phenomenalism, however, while the whole primary-secondary quality distinction presupposes representative realism. Supporters of the latter would say that the best explanation of the correlation is that the sense experiences arise in the measurement of actual physical objects.

ARGUMENTS FROM SCIENCE. Science can adequately explain and describe the nature of the physical world solely in terms of primary qualities; hence, while primary qualities must characterize objects, there is no need to suppose that secondary qualities must also. The latter would be otiose, and on the principle of economy, or Ockham's razor (that entities should not be multiplied more than is necessary), it would be unscientific to suppose that they exist as intrinsic properties of objects. The objection to this argument is partly that the science of one's day is not final (thus, Locke was persuaded by seventeenth-century science to include solidity in the list of primary qualities), and mainly that scientific theory and description are not the whole truth—they describe only one aspect of the world, being limited by their quantitative approach and their instruments. Secondary qualities may thus be real properties of matter with biological or aesthetic functions; Ockham's razor oversimplifies the facts pertaining to living things.

Investigation of the causal processes on which perception depends shows that the only variables capable of transmitting information about the properties of external objects are spatiotemporal ones, which are associated with primary qualities. Thus, light waves (energy distributed in space and time) pass from the object to the per-

cipient, but nothing resembling experienced color and sound is transmitted. But the main force of this argument, since it applies to all the senses, is neurological. The nerves from the different sense organs to the brain are all similar, and therefore the only variables are the frequencies of the impulses (which convey the intensity of the stimulus), their different neural pathways, and their different destinations in the brain. Indeed, it seems to be the different destinations that primarily govern the type and quality of the sensation. And although one can conceive of primary qualities being transmitted by spatiotemporal variables, it is difficult to conceive of color, warmth, taste, or smell being so transmitted. (It may be objected that radio and television can transmit color and sound by converting them into electrical impulses for transmission and then reconverting them. But, strictly speaking, what is converted is not color or sound but light waves or sound waves; moreover, the radio or television station must use microphones and cameras to effect the conversion, and there is no evidence of such conversion devices at the objects we see or hear.)

BERKELEY'S CRITICISMS

Berkeley's formidable criticisms of the distinction between primary and secondary qualities have convinced many people. We have mentioned his objection concerning relativity, which, though valid against Locke, can be avoided by restating the distinction on the basis of measurement. He also has nothing to say on the scientific considerations, which were not explicit in Locke. But he did have some further well-known criticisms. First, he stated, "An idea can be like nothing but an idea." In other words, our sensa, being private, mental, and directly perceivable, cannot resemble properties of material objects that are public, physical, and not objects of direct awareness. But resemblance is claimed only for primary qualities; and though sensa cannot be extended in physical space, it seems reasonable to claim a structural resemblance, a similarity in form, between the spatial relations that they sensibly possess and those attributed to objects by measurement; thus, it can be confirmed by measurement that various relations between the sides of a square sensum hold in the object. A similar resemblance seems plausible in the case of motion. There are, however, some underlying difficulties here. In the older representative realism, sensa were mental; and since the mind was held to be unextended, they could hardly have spatial relations. But newer versions would allow some sensible or subjective space different from physical space; certainly sensa seem spatial, and there seems to be no reason why what is directly perceivable and what is not should be unable to have a similar form or character.

Second, matter consisting only of primary qualities—for instance, possessing extension but no color, taste, sound or smell—is inconceivable. This objection is beside the point: Admittedly one cannot conceive, in the sense of "imagine" or "picture to oneself" (Berkeley's sense of the word), any such thing, for what we can imagine is limited by past experience and perception. But the range of possible existents need not be confined to this, and there is much in science, particularly in modern physics, that cannot be imagined or pictured.

See also Berkeley, George; Boyle, Robert; Colors; Descartes, René; Galileo Galilei; Leucippus and Democritus; Locke, John; Newton, Isaac; Pain; Perception; Realism; Sensa; Sound.

Bibliography

CLASSICAL WORKS

Anticipations of the distinction are to be found in Democritus, Fragments 9 and 125, in H. Diels and W. Kranz, *Fragmente der Vorsokratiker;* Galileo, *Il saggiatore* (The Assayer; 1623), in his *Opere*, edited by F. Flora (Milan, 1952), pp. 311–314; René Descartes, *Principles of Philosophy*, Part IV, Nos. 188–203, and *The World or Essay on Light*, Chs. 1–2. Robert Boyle adumbrated the distinction in his *The Origin of Forms and Qualities* (London, 1666), especially the section titled "An Excursion about the Relative Nature of Physical Qualities"; John Locke stated it in his *Essay concerning Human Understanding* (London, 1690), Book II, Ch. 8; Newton's view is best stated in his *Opticks*, Book I, Part II (paperback, New York, 1952, p. 124), but that was not published until 1704. Berkeley's criticisms are in his *Principles of Human Knowledge*, Secs. 9–20. Edwin A. Burtt, *The Metaphysical Foundations of Modern Physical Science* (London: K. Paul, Trench, Trubner, 1925), contains useful discussion of the early versions.

MODERN WORKS

No book is devoted to the distinction, and substantial treatments are rare, but helpful contributions may be found in the following:

A more recent scientific comment on the distinction is Russell Brain, *Mind, Perception and Science* (Oxford: Blackwell, 1951). L. Susan Stebbing, *Philosophy and the Physicists* (London: Methuen, 1937), attacks scientific versions of it, especially those of Newton and of Arthur S. Eddington (see his *The Nature of the Physical World*, Cambridge, U.K.: Cambridge University Press, 1928).

Among the general works available are Charlie Dunbar Broad, *Scientific Thought* (London: Kegan Paul, 1923) and *The Mind and Its Place in Nature* (London: Kegan Paul, 1925); William Pepperell Montague, *The Ways of Knowing* (London: Allen and Unwin, 1925); Roderick Chisholm, *Perceiving: A Philosophical Study* (Ithaca, NY: Cornell University Press, 1957); and more fully, R. J. Hirst, *The*

Problems of Perception (London: Allen and Unwin, 1959). D. M. Armstrong, *Perception and the Physical World* (New York: Humanities Press, 1961), Chs. 14 and 15, discusses the subject from a direct realist position, and Gilbert Ryle's criticisms of the distinction in his *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954) are well known.

Among articles on the distinction are Reginald Jackson, "Locke's Distinction between Primary and Secondary Qualities," in *Mind* 38 (149) (1929): 56–76; J. J. C. Smart, "Colours," in *Philosophy* 36 (137) (1961): 128–142 (reprinted with additions in his *Philosophy and Scientific Realism*, New York: Humanities Press, 1963); Colin Strang, "The Perception of Heat," in *PAS* 61 (1960–1961): 239–252; and for a more general treatment, W. C. Kneale, "Sensation and the Physical World," in *Philosophical Quarterly* 1 (2) (1951): 109–126.

OTHER RECOMMENDED TITLES

Averill, Edward Wilson. "The Primary Secondary Quality Distinction." *Philosophical Review* 91 (1982): 343–362.

Ayers, Michael. *Locke: Epistemology and Ontology*. London: Routledge, 1991.

Bennett, Jonathan. *Locke, Berkeley, Hume: Central Themes*. Oxford: Oxford University Press, 1973.

Cummins, Robert. "Two Troublesome Claims about Qualities in Locke's *Essay*." *Philosophical Review* 84 (1975): 401–418.

Dicker, Georges. *Perceptual Knowledge*. Dordrecht: Reidel, 1980.

Fales, Evan. *A Defense of the Given*. Lanham, MD: Rowman & Littlefield, 1996.

Fumerton, Richard. *Metaphysical and Epistemological Problems of Perception*. Lincoln: University of Nebraska Press, 1985.

Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.

Levin, Janet. "Physicalism and the Subjectivity of Secondary Qualities." *Australasian Journal of Philosophy* 65 (1987): 400–411.

Mackie, John. *Problems from Locke*. Oxford: Clarendon Press, 1976.

McGinn, Colin. *The Subjective View*. Oxford: Clarendon Press, 1983.

Sellars, Wilfrid. *The Metaphysics of Epistemology*. Atascadero, CA: Ridgeview Press, 1989.

Smith, A. D. "Of Primary and Secondary Qualities." *Philosophical Review* 99 (2) (1990): 221–254.

Wright, Crispin. "Moral Values, Projection, and Secondary Qualities." *PAS*, Supp. 62 (1988): 1–26.

R. J. Hirst (1967)
Bibliography updated by Benjamin Fiedor (2005)

PRINGLE-PATTISON, ANDREW SETH
(1856–1931)

Andrew Seth Pringle-Pattison, the Scottish personal idealist, was born Andrew Seth, in Edinburgh. (He adopted the surname Pringle-Pattison at the age of forty-two as a condition of inheriting a family estate in Scotland.) He studied philosophy at Edinburgh University under Campbell Fraser. Two years of study in Germany convinced him that it was the worst place for the study of German idealism but resulted in his completing, at twenty-four, his Hibbert essay, *The Development from Kant to Hegel.* From 1880 to 1883 he served as Fraser's assistant at Edinburgh and then took the foundation chair of philosophy in the University College of South Wales at Cardiff. He left Cardiff in 1889 for the chair of logic and metaphysics at the University of St. Andrews. This he relinquished in 1891, when he succeeded Fraser at Edinburgh. In 1919 he resigned, after thirty-nine influential years as a university teacher.

Philosophy for Pringle-Pattison was a serious enterprise of the human spirit, which he did not distinguish strictly from a statement of his own findings in religion and morality. His writing is clear and eloquent but not very original. He sought to advance his subject through critical interpretation of the great philosophers, especially Immanuel Kant and G. W. F. Hegel. He was skeptical about the value of philosophical systems, holding that we cannot know the universe as we can know its individual parts; only God can do this. Rather, "the ultimate harmony may justifiably be spoken of as an object of faith—something which I am constrained to believe, even though I do not fully see it."

Pringle-Pattison was a Scottish Hegelian with a difference. Rebelling against the absolutism of Hegel and of such Hegelians as Francis Herbert Bradley and Bernard Bosanquet, for whom the individual is merged in the universal, he insisted on the uniqueness of the individual person. It is only as knower that the self is a unifying principle. As a real being it is separate and distinct, impervious to other selves, even to God. "I have a centre of my own—a will of my own—a centre which I maintain even in my dealings with God Himself." We feel this to be so; it neither needs to nor can be established by argument. But God too is a Person; we cannot deny him self-consciousness, because this is the highest source of worth in ourselves. Hegel and the Hegelians were at fault here also.

Philosophy, Pringle-Pattison held, cannot do justice to "the individual within the individual—those memories, thoughts, and feelings which make each of us a separate soul" (*Hegelianism and Personality,* p. 217). Religion and poetry go further and deeper than philosophy, and this, as he said, is why he drew so frequently on the poets.

Our knowledge of the Absolute starts from experience—our experience "of the concrete worlds of morality, of beauty, of love or of the passion of the intellectual life."

It is, however, a postulate of reason that the world is a cosmos, not a chaos, which we can gradually explore but never grasp in its entirety. Pringle-Pattison described his philosophy as "a larger idealism" that reconciles the dictates of morality and religion with the findings of science, purpose being the supreme category.

He was cautious in his claims about immortality. The nature of the soul is such that it is reasonable to entertain the hypothesis of its survival, and since human spirits must be "values for God" they were surely not made to be constantly destroyed and replaced by others. Yet if there is personal immortality, it is not the inherent possession of every human soul but must be won by the continuous effort needed to develop a coherent self. Morality does not depend on personal immortality, nor need immortality be the central article of philosophy or religion. In the apprehension of Truth, Beauty, and Goodness—eternal realities—man has already tasted eternal life and so should not be much concerned about personal survival.

See also Absolute, The; Bosanquet, Bernard; Bradley, Francis Herbert; Hegel, Georg Wilhelm Friedrich; Idealism; Kant, Immanuel.

Bibliography

WORKS BY PRINGLE-PATTISON

Scottish Philosophy: A Comparison of the Scottish and German Answers to Hume. Edinburgh: Blackwood, 1885.

Hegelianism and Personality. Edinburgh: Blackwood, 1887.

Man's Place in the Cosmos. Edinburgh: Blackwood, 1897.

The Idea of God in the Light of Recent Philosophy. Oxford: Clarendon Press, 1917.

The Idea of Immortality. Oxford: Clarendon Press, 1922.

Studies in the Philosophy of Religion. Oxford: Clarendon Press, 1930.

The Balfour Lectures on Realism (delivered 1891), with a memoir by editor G. F. Barbour. Edinburgh: Blackwood, 1933.

A. K. Stout (1967)

PRIOR, ARTHUR NORMAN
(1914–1969)

Arthur Norman Prior was born on December 4 at Masterton, near Wellington, New Zealand. He acknowledged an early philosophical debt to John Findlay. But his first academic post was at Canterbury University College, where he succeeded Karl Popper. He was the visiting John Locke Lecturer at Oxford in 1956, and in 1958 he was appointed a professor of philosophy at the University of Manchester. After short periods as a visiting professor at the University of Chicago and at the University of California at Los Angeles, he moved in 1966 to a tutorial fellowship at Balliol College, Oxford, and Oxford University appointed him to a concurrent readership.

Prior's early intellectual interests were very much religious in character. He was influenced for several years by the theologian Arthur Miller, who combined a strict adherence to Presbyterian doctrine with an equally strong support for socialism and opposition to nationalism. But Prior's pacifism weakened, and he served from 1942 to 1945 in the New Zealand air force. And the central focus of his interests gradually shifted—helped by an occasional bout of atheism—from theology to ethics and logic. He exchanged ideas with a wide circle of friends and acquaintances, and his hospitality to students was legendary.

Prior's first book, *Logic and the Basis of Ethics* (1949) traced seventeenth-, eighteenth-, and nineteenth-century anticipations of G. E. Moore's criticism of the so-called naturalistic fallacy. But his main claim to fame lies in his pioneering work on the formal logic of temporal relationships. His most important investigations in this field were published in *Time and Modality* (1957), *Past, Present, and Future* (1967), and *Papers on Time and Tense* (1968). But he also wrote on several logical topics in this encyclopedia; he published a substantial survey of the current state of logical inquiry under the title of *Formal Logic* (1955; 2nd ed., 1962); and a posthumous volume of papers, *Objects of Thought* (1971), was edited by P. T. Geach and A. J. P. Kenny.

Prior almost always used the Polish style of notation in the discussion of logical proofs and principles and was a convinced, though largely unsuccessful, champion of its virtues. The major inadequacy in his tense logic, however, was a failure to discuss or accommodate aspectual differences—roughly, differences between the meanings expressed by verbs in a perfect tense and those expressed by verbs in an imperfect tense (see Galton, 1984). Other criticisms may be found in L. J. Cohen's (1958) review of *Time and Modality* and in his subsequent controversy with Prior (*Philosophy* 34 [1959]). In his *Formal Logic* Prior displayed an impressively wide acquaintance with logical systems outside the field of tense logic, and this book remains a useful text for anyone interested in comparisons between different axiomatizations of the propositional calculus, between different kinds of logical quantification, between different modal logics, or between different three-valued or institutionist logics. But the treatment of metalogical issues in the book is

occasionally rather selective: for example, in its discussion of completeness proofs for the predicate calculus as against its treatment of completeness proofs for the propositional calculus.

Outside the brilliant originality of his work on tense logic, perhaps Prior's most striking idea was expressed in "The Runabout Inference-Ticket" (1960), where he argued that, if the meaning of a logical connective consisted just in the logical uses to which it can be put (as many seemed to hold), then it would be easy to invent a connective with a meaning that would enable one to infer any conclusion from any premises.

See also Atheism; Modal Logic; Moore, George Edward; Nationalism; Pacifism; Popper, Karl Raimund; Socialism.

Bibliography

WORKS BY PRIOR

Logic and the Basis of Ethics. Oxford: Clarendon Press, 1949.

Formal Logic. Oxford: Clarendon Press, 1955. 2nd ed., 1962.

Time and Modality. Oxford: Clarendon Press, 1957.

"The Runabout Inference-Ticket." *Analysis* 21 (1960): 38–39.

Past, Present, and Future. Oxford: Clarendon Press, 1967.

Papers on Time and Tense. Oxford: Clarendon Press, 1968.

Objects of Thought, edited by P. T. Geach and A. J. P. Kenny. Oxford: Clarendon Press, 1971.

WORKS ON PRIOR

Cohen, L. J. Review of *Time and Modality. Philosophical Quarterly* 8 (1958): 266–271.

Gallon, A. *The Logic of Aspect.* Oxford: Clarendon Press, 1984.

L. J. Cohen (1996)

PRIORITY OF KNOWLEDGE, THE

See *Knowledge, The Priority of*

PRISONER'S DILEMMA

See *Decision Theory*

PRIVATE LANGUAGE PROBLEM

The private language problem is essentially the question of whether or not a language as a system of symbols that are means of thinking is, of necessity, a language as a system of symbols that are means of communication. Defining "private language" as language (in the sense of means of thinking) which in principle the speaker alone can understand (so that it cannot serve as a means of communication), our question is roughly equivalent to: "Is a private language possible?" Many philosophers, following Ludwig Wittgenstein, have made the claim (here called the private language thesis, abbreviated PLT) that private languages are impossible. Armed with it, they have argued against solipsism, phenomenalism, the analogical or empirical view of one's knowledge of other minds, and against mind-body dualism. Some of them have gone on to argue for certain versions of philosophical behaviorism as well as for the view that the meaning of a word consists of its use or employment in a social practice and not in its referring to something or its designating a kind of entity.

Thus, the PLT has been a central principle in the cluster of Wittgensteinian doctrines. It is not clear, however, that exactly the same thesis figures in all the arguments in question, since the idea of a private language varies in different contexts. There is, therefore, a multiple problem: First, to differentiate the several propositions which pass as the PLT by clarifying the sense of "private language" being used; second, to determine which ones are true; and third, to explain why they are supposed to be intimately related. These problems differ from the question, debated around 1930, of whether or not it is possible to start with a private language about one's sensations or "raw" feelings and arrive at the intersubjective and communicable language of science. (On this question, see Rudolf Carnap, "Psychology in Physical Language," and J. R. Weinberg, *An Examination of Logical Positivism.*)

THE SENSE OF "IMPOSSIBLE"

In all the interpretations of the PLT, the word *impossible* is understood in a strong sense that is not easy to characterize precisely. Some philosophers speak of "logical impossibility," but they do not necessarily mean that private languages are impossible in the sense that unbounded triangular figures are impossible. The expression "unbounded triangular figure" reduces to the formal self-contradiction "*unbounded* figures *bounded* by three lines" by means of a substitution allowed by the definition of "triangle." But few philosophers would suggest that there is a similarly ready definition of "language" by means of which we can produce a formal self-contradiction "private so-and-so which is not private." The impossibility at issue is like (1) the impossibility of unextended red things (that is, the impossibility that something be red and yet lack

width or length), or (2) the impossibility of a cube with fewer than eight edges. These do not lead straightforwardly to formal contradictions, since there are no definitions for all the terms involved; they depend on implication relations that constitute the concepts involved in their statement. In the last analysis, the persistent rejection of (1) and (2) evidences the failure to understand the meanings of all the words involved, that is, the lack of some of the relevant concepts. But (1) is unprovable and obvious, and (2) only needs a trivial argument, while the PLT (if true) requires careful reasoning. We shall speak of *conceptual impossibility* to refer to any formal self-contradiction, to any impossibility which entails a formal self-contradiction, and to any a priori impossibility such as that found in the above examples (1) and (2).

THE PRIVATE LANGUAGE THESIS

The most important propositions often discussed as the PLT, each embodying a different idea of private language, are the following:

PLT*: It is impossible for a man to use a word with a meaning that nobody else could, even in principle, understand.

PLT-1: It is impossible for a man to use words that refer to private objects, that is, objects that nobody else could—even in principle—know. (For subtheses arising out of the ambiguities of "know," see H.-N. Castañeda, "The Private-Language Argument.")

PLT-2: It is impossible for a man who has always lived in isolation to possess a language, even if his sounds are understandable by another person.

Here the expressions "could not in principle" and "impossible" are meant to express conceptual impossibility. PLT* allows that a man may use words with meanings that nobody else in fact understands, provided that they are understandable to other people in the appropriate circumstances. PLT-1 allows that a man may refer to objects that, in fact, he alone knows, but again others must be capable of knowing them in the appropriate circumstances. PLT-2 allows that a man, like Robinson Crusoe, keeps possession of a language he learned previously while living in a community of speakers.

Many philosophers assume that it is conceptually impossible for two persons to share one and the same immediate sensation. Many also hold that, in a strict sense of "know," others do not really know whether one has a certain immediate sensation or not, precisely because they cannot share it. On these assumptions, a language about one's own immediate sensations would be a language of the sort that PLT-1 claims to be impossible. Indeed, such a language is customarily regarded as the would-be prototype of private language.

In general, on the assumption that (direct) knowledge of the referent of a word is required for understanding the meaning of the word in question, PLT* entails PLT-1. On this assumption, a language about one's own immediate sensations is also private in the way that PLT* claims to be impossible.

PLT-1 does not entail PLT*. A word might have a meaning understandable to only one person because the word itself is a private object in the sense of PLT-1, even though everybody may be acquainted with the physical objects it refers to. For example, the words of a person's language might all be mental images of German written words, so that all his thinking would be a sort of mental reading of German. In this case, the referents of the words would be public, but the words themselves would be private and hence unintelligible to others.

PLT-2 neither entails nor is entailed by PLT-1. If PLT-2 is true, then if on the previous assumptions about sensations, one's language about one's own sensations is private in the sense of PLT-1, then one could still, in principle, invent such a language. Conversely, the truth of PLT-1 does not by itself make it impossible for an isolated person to invent a language about physical objects. Similarly, PLT-2 neither entails nor is entailed by PLT*.

APPLICATIONS OF PRIVATE LANGUAGE THESES. The important claims made with the help of the PLT do require other assumptions, which in their turn play roles, as we shall see, in the defense of the PLT itself. The most natural and pervasive of these assumptions is the following:

(*A*) In the sense of "thinking" in which one can both have a false (or true) thought and draw inferences from what one thinks, it is conceptually impossible to think without possessing a language that is a means of thinking.

From this assumption and PLT-2, one can conclude that the fact that one thinks, guarantees the existence of other persons, namely, one's fellow speakers of the same language. Thus, the solipsist who merely asserted that it is possible that he alone exists at the time he is thinking would be contradicting himself (an argument of this sort can be constructed with premises suggested by Rush Rhees in "Can There Be a Private Language?"). Of course, many philosophers have serious objections to (*A*).

The existence of hallucinations, illusions, and visual perspective leads many philosophers to characterize every

case of perception in terms of our apprehension of sense data or immediate impressions. Some have proceeded to espouse a phenomenalistic program of "logical reconstruction" of physical objects and minds as systems of sense data; others, however, have subscribed to some form of realism, that is, the complete irreducibility of physical objects and minds to sense data. But all of them have recently been criticized on the ground that the language of sense data is private in either the sense of PLT* or the sense of PLT-1. Here, in addition to (*A*), the critics need the following assumption:

(*B*) If it is conceptually impossible that there be a language about entities of a sort *T*, then there are no entities of sort *T*.

Again, some philosophers would claim against (*B*) that if PLT* or PLT-1 is true, then sense data or the given in experience are simply ineffable.

Many philosophers have subscribed to some form or other of a principle of verification, for example:

(*C*) It is conceptually impossible to understand a sentence without knowing what state of affairs would verify the statement made with it.

Assumption (*C*) leads to the view that language about states of consciousness is private, if we add to it and (*A*) and (*B*) the following principle:

(*P*) Only the person himself can verify conclusively and directly that he has certain experiences.

On this view, for instance, when someone else speaking about me says, "He is in pain," he cannot understand or mean exactly the same thing that I understand and mean when, of myself, I say, "I am in pain." But if PLT* is accepted, one is involved in a contradiction. Here many philosophers have given up (*P*), and in order to guarantee that everybody else can know what somebody is feeling or thinking, some philosophers have espoused some form of behaviorism, that is, a view according to which every description of a person's experiences or mental states is really shorthand for (synonymous with) a description of his bodily movements, his relations to other bodies, and his abilities to perform further movement. This is often supplemented with the supposition that first-person utterances like "I have a headache" do not make statements of direct knowledge but are, rather, learned responses, analogous to the natural responses of moaning, crying, and so on, which are said to constitute the person's ache. As is to be expected, other philosophers have preferred to keep (*P*) and reject one or more of the other premises, in particular (*C*) or PLT*. (See Castañeda, op. cit., Part B, for a discussion of the privacy of experiences.)

THE MAIN ARGUMENTS FOR THE PLT. There are many arguments seeking to prove that being private makes it impossible for a language to have a property required for the existence of a language. Most of the arguments depend on the following assumption:

(*D*) A language is a system of rules, and to speak or write a language is to follow rules.

On this assumption, it suffices to establish the PLT to show that a man (say, Privatus) cannot be following rules when he is using a private language (to be called Privatish). This is, in fact, what a series of arguments suggested by Wittgenstein purports to do. The gist of the argument is as follows: A rule is, by its very nature, the sort of thing that can be misapplied (or disobeyed), but Privatus cannot misapply the rules of Privatish; hence, when speaking Privatish, Privatus is not following rules. The specific arguments are meant to support the crucial premise:

(1) Privatus cannot misapply the rules of Privatish.

A fair objection to (1) is that Privatus can certainly make slips; he may call something of kind *A* "*B*," whatever "*A*" and "*B*" may mean in Privatish. Slips of the tongue are precisely ways in which one violates the rules (if there are such) of natural languages. For instance, if there are rules of English governing the application of color words to physical objects; whenever one commits a slip of the tongue and calls a red object "blue," then one misapplies either a rule governing the use of "red" or one governing the use of "blue."

This reply to (1) is often met by several rejoinders. The first claims both that a slip counts as a misapplication of a linguistic rule only if there is a way in which the speaker can in principle detect and correct his slip and that Privatus cannot detect or correct his slips. This rejoinder, however, changes the issue, since premise (1) says nothing about verifying the existence of a misapplication of a rule. Nevertheless, the rejoinder has a point, for if to use words is to apply rules, then one must at least sometimes be able both to know of one's misapplications of the rules for the use of one's words and to know how to make the appropriate corrections. The question of whether or not Privatish allows this is discussed below under premise (2).

The second rejoinder is that to obey a rule is a *custom* (use, institution), but Privatus's actions cannot constitute a custom (see Ludwig Wittgenstein, *Philosophical Investi-*

gations, Sec. 199). This rejoinder would establish PLT-2 but not PLT* or PLT-1. For it may be a custom in a tribe that people use words which they alone understand in the ways required by PLT* or PLT-1. But as an argument for PLT-2 the rejoinder is by itself question-begging. It must be supported by an argument which shows that obeying a rule is indeed a custom.

The third rejoinder is that Privatus's slips do not count as violations of the rules of Privatish because we cannot be corrected or taught by others what is the correct thing to say (see Wittgenstein, op. cit., Sec. 378, and Norman Malcolm, "Discussion of Wittgenstein's *Philosophical Investigations,*" pp. 536f.). If the "cannot" here is taken to mean conceptual impossibility, the rejoinder does not apply to PLT-2. If it is taken in a weaker sense, that is, a sense in which a person may be in the position of being in fact corrected by other persons, then the rejoinder supports PLT-2, but it would not allow that there be just one language-user in the universe. Besides, it is not clear that it would allow that Antonia Udina, for example, used language when, as we normally say, he spoke Dalmatian as the last speaker of Dalmatian. Although a person who uses words must be capable of self-correction, it is not immediately obvious that a person's sounds cannot count as utterances of words if nobody else can (in some sense) correct him. The need for others' possible corrections has to be established by an argument. Thus, we are again thrown back to the other lines of reasoning.

The fourth rejoinder is that Privatus's slips do not count because another person, by noting Privatus's behavior and circumstances, cannot discover that his use of the word is correct or incorrect (adopting Malcolm, op. cit., p. 537). This rejoinder also leaves PLT-2 unsupported if "cannot" is understood as expressing conceptual impossibility. While it must be conceptually possible for Privatus to know whether his uses of language are correct or incorrect, it is not at all clear that it must be possible for others to know this fact. The principle that it must be possible for others to know whether his uses of language are correct or incorrect requires an independent argument to support it. However, the present rejoinder has a point. It reminds us that if there is no way at all of telling, for any word of Privatish, whether or not Privatus used it correctly (however coherent the concept of a private language is), it would be a completely gratuitous hypothesis that Privatus spoke a private language. Although our topic here is only the conceptual possibility of private language, we should note that the claim that somebody's entire language is of the type described in PLT* is certainly gratuitous. Yet the claim that someone has a mixed language, part of which is private in the sense of PLT*, does not seem gratuitous.

The fifth rejoinder dismisses mere slips on the ground that they show at most a breakdown of a linguistic habit. The rejoinder asks us to consider the case of Privatus trying deliberately to apply a rule of Privatish and failing to comply with it. The rejoinder claims that, for Privatish, "thinking one was obeying a rule would be the same thing as obeying it," but "to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately'" (Wittgenstein, op. cit., Sec. 202). This rejoinder does not require that every utterance of a word be a case of deliberately attempting to obey the corresponding linguistic rule(s). Conjoined with assumption (*A*), this view would lead to a vicious infinite regress. For then, in order to say something, one would have to be aware of the rules governing the words one intends to utter, and these rules in their turn would be formulated in some words the rules governing which one would have to be aware of through some other words, and so on ad infinitum. Therefore, to use language is, of necessity, to use most of the words from habit, not in intended obedience of the linguistic rules. The rejoinder cannot even demand that Privatus sometimes be aware of the rules of Privatish: A being might speak a language without ever rising to the level of formulating any of his rules. But if, by assumption (*D*), languages are made up of rules, then if it were conceptually impossible for Privatus to be at least sometimes aware of the rules of Privatish, Privatish would be a very defective language indeed, incapable of discharging the philosophical duties that private languages are alleged to discharge. Thus, the rejoinder is right in urging that

(*a*) For every rule *R* of a language *L* and every speaker *S* of *L,* it is conceptually possible that sometimes *R* applies to *S*'s situation while *S* thinks that he is obeying *R* without *S*'s actually obeying *R.*

Presumably, a rule of language is here of the form "If *x* is ϕ, you may (must) call it '…,'" but the meaning of "call" is difficult. In one normal sense of "call," slips of the tongue are, again, ways in which (*a*) is true. Clearly, a person may think that he is calling a thing "red" in deliberate compliance with the English rule for "red," without realizing that he actually called it "blue" because he is deaf or because he simply did not hear what he said. In the same sense of "call," (*a*) can be true because the speaker deliberately calls a red thing "blue," if he thinks that the rule in question allows (or prescribes) his calling it "blue." In particular, suppose that the rule *R* allowing (or prescrib-

ing) that one call a thing "red" is the rule Gaskon typed yesterday and that today, confusedly, Gaskon thinks that the rule he typed yesterday allows (or prescribes) that a certain thing be called "blue," and he calls the thing in question "blue," thinking that he is complying with the rule. Here, in spite of his deliberately calling a certain thing "blue," Gaskon's use of "blue" and the rule he thinks he is complying with both satisfy (*a*). Both ways of satisfying (*a*) are open to Privatus. It might be argued that Privatus's deliberately calling one of his private objects "*A*" instead of "*B*" has no point or "function" (see Wittgenstein, op. cit., Sec. 260), since he is not talking to others. This is, however, false. Privatus may very well play word games involving miscallings of things. But more importantly, whether or not there is a point in Privatus's flouting of the rules of Privatish has nothing to do with the issue about the possibility of private language.

The rejoinder often uses a stronger sense of "call." In this sense, by a natural development of assumption (*A*), to *think* that something is, for example, red is to call it "red." This stronger sense appears in an argument given in support of PLT-1. As said above, language about one's own immediate sensations is often regarded as the paradigm of private language in the sense of PLT-1. Now, one knows incorrigibly that one's sensations have immediately sensible qualities. That is to say, if one believes that one has a pain (itch, tickling, feeling of discomfort), then one knows that one has a pain (itch, tickling, feeling of discomfort). So it is impossible to have no pain while one thinks that one has a pain. Thus, if one thinks that one is obeying the rule of the form "If *x* is a pain, you may (must) call it 'pain,'" one surely thinks that one is in pain and the rule cannot fail to apply. Similarly, since one also has incorrigible knowledge of the absence of one's immediate sensations, if the objects that Privatus can think about in Privatish are only his immediate sensations, then when he thinks that a rule of Privatish does not apply, the rule does not, in fact, apply. But if "call" is taken in its normal sense, neither of these two features of the rules of Privatish implies that Privatus cannot think that he is obeying a rule (which then applies) without actually obeying it, since slips and deliberate miscallings are still available as violations of the rule. However, if "call" is taken in the strong sense (in which thinking can be calling), then if Privatus thinks that he is obeying a rule of the form "If *x* is *A,* you may (must) call it '*A*,'" he surely thinks that the rule applies, that is, he thinks that the object *x* is *A*; if *A* is a sensible property of Privatus's immediate sensation *x,* then *x* is *A,* and Privatus is both calling *x* "*A*" and unavoidably obeying the rule. Thus, if Privatish is a private language about Privatus's immediate sensations and their sensible properties, then (*a*) above and (*b*) below are both false:

(*b*) For every rule *R* of a language *L* and every speaker *S* of *L,* it is possible that sometimes *S* thinks that he is obeying *R* while he is not.

Since (*a*) is true, Privatish is not a private language.

This argument does not by itself support PLT-2; it may or may not support PLT*, depending on how one interprets the phrase "knowing the meaning of a word."

There is, however, a difficulty with the above argument. Consider the rule of English: "If *x* is a cat, you may (must) call *x* 'cat'; that is, you may (must) think that *x* is a cat." This rule differs from the above rule for the Privatish word "*A*" in that thinking that one is obeying the rule for "cat" does not imply that the rule for "cat" applies to the situation in question. For to think that one is obeying the latter rule implies that one thinks that it applies, and this implies that one thinks that some object *x* is a cat. But surely one can be mistaken about *x*'s being a cat. Yet the rule for "cat" also fails to satisfy condition (*a*). Suppose that the rule applies; then the object *x* in question is a cat. And suppose that one thinks that one is obeying the rule; then it is true that one thinks that if *x* is a cat one may (must) think that *x* is a cat, and that one thinks that *x* is cat. Thus, one is in fact obeying the rule! Therefore, the strong sense of "call" included in the concept of language rule *R* makes (*a*) an impossible condition.

Now, in the case in which a rule *R* does not apply to a man's situation, we are often reluctant to say that when such a man thinks that he is obeying *R,* he is not obeying *R.* But we could say this with no great distortion, and if we did, we could say that the above rule for the English word "cat" satisfies condition (*b*). For in a situation in which an object *x* is not a cat and the rule does not apply, we may very well both misperceive or otherwise think that *x* is a cat and think that, in accordance with the rule, we may (must) think that *x* is a cat. Thus, if we raise (*b*) as the crucial condition that linguistic rules must satisfy, then we can claim that PLT-1 is established in the sense that a pure language of sensations is impossible. But this answer is inconclusive. Besides the small amount of distortion involved, there is the fact that (*b*) is not a general condition of rules. This is shown by the following rule which a man might give to his son: "If you think that you need to delay your action, think that $1 + 2 + 3 + \cdots + 24 = 300$." Since to think that one thinks that *p* entails that one thinks that *p,* if the boy thinks that the rule applies, he thinks he needs to delay his action, and the rule applies. If he thinks that he is obeying the rule, he thinks

both that it applies and that $1 + 2 + 3 + \cdots + 24 = 300$; hence he thinks that $1 + 2 + \cdots + 24 = 300$; hence, the rule applies and he obeys it. Thus, to defend PLT-1 by means of (*b*) requires an independent argument showing that rules of language must, in any case, comply with (*b*), distorted as suggested.

Let us turn now to a subtler line of argument. Some defenders of the PLT do not argue for (1) but for

(2) Privatus cannot distinguish his correct uses of Privatish words from his incorrect uses.

Suppose, then, that Privatus is debating whether something is *A* or not. Suppose that Privatish is private in the sense of PLT-1. Here the defenders of the PLT adduce (*a*) that Privatus lacks a *criterion* of correctness, that is, "something *independent* of his impression" that he is correctly using the Privatish rule governing the use of "*A*" by means of which he can "prove his impression correct" (Malcolm, op. cit., p. 532), and (*b*) that his impression that he remembers what objects of kind *A* appeared like before is of no help, since memory "is not the highest court of appeal" (Wittgenstein, Sec. 56) and the "process [of checking memories] has got to produce a memory which is actually *correct*" (Sec. 265). Now, these points exaggerate Privatus's predicament. Privatus's private objects may be related among themselves by entailment, by coexistence, by similarities, by causal relationships, and so on. Privatus can resort to any of these to test whether he is, on the present occasion, using the term "*A*" correctly. For instance, in Privatish, "being *A*" may be logically equivalent to "being *B* and becoming *C* in the presence of another *C*." Indeed, Privatus may even employ paradigms. The very first object he calls "*A*" may very well be enduring, so that he can compare the next objects of kind *A* with it. The same applies to languages of the type mentioned in PLT-2. Furthermore, memory *is* the highest court of appeal when it comes to our knowledge of the past. True, we have records and other historical evidence, but all of this only provides inductive evidence, not a proof, and our inductions involve the acceptance of unchallenged memories.

Nevertheless, Privatus is not only in no position to question the correctness of all of his uses of words, but he also cannot *prove* that the uses he questions are correct unless he is allowed the ability to identify certain properties of objects without criteria and without challenging his memory. But exactly the same happens with the speakers of *any* language. In the case of terms like "red" and "straight," for instance, there is nothing at all to which an English speaker *E* can resort in order to "prove" that he has correctly called an object red or straight. His fellow speakers may all utter in unison, "Not red but blue." Yet this choral utterance is not a proof; the speakers may be lying, may all be victims of a hallucination, or may just be rehearsing a new song—or the whole proceedings may be just *E*'s hallucination. In any case, for *E* to accept the correction, he must correctly identify the words expressing it without the use of criteria and remember correctly the meanings of these words. A vicious infinite regress would ensue if *E* were required to have a proof that he both remembers this correctly and identifies the objects the words apply to.

Moreover, there is nothing to prove each corrector's use of words correct. Suppose, for example, that one corrector learned the meaning of "blue" with the help of object *O* and that he continuously stares at *O* during the preceding two minutes before correcting Privatus. He still must *remember* correctly that *O* has the same color it had two minutes before, that the color of *O* is called "blue," that the name of the color sounds "b–l–u–e," that the noise "red" uttered by *E* has the same meaning that makes red and blue incompatible, and so on. Thus, either somebody *just* identifies some words or objects correctly and remembers some qualities of objects and the meanings of some words correctly, or else nobody can be corrected by another speaker. In sum, demands (*a*) and (*b*) cannot be adduced against the possibility of a private language.

LOGICAL WORDS. Often it is claimed that a private language cannot have logical words or syntactical rules, both of which are necessary for the existence of logical relationships. Clearly, if a private language is allowed no implications or entailments, it would certainly be no language. But if "private language" is meant in the sense of PLT* or PLT-1 or PLT-2, this contention appears to be false. Often this contention is defended on the ground that a really private language does not have words with meanings in common with the words of another language (Wittgenstein Sec. 261; Malcolm p. 537). Now, private language in this sense is impossible. A language is a system of words of which some refer to objects, some signify properties or relations, and some express logical connections; the words expressing logical connections must be capable of being understood by anybody else and must, therefore, be common to all languages. This is an important result. But it is not the same as PLT*, which requires that every single word of a language must be understood by persons other than the speaker. Likewise, the impossibility of languages without logical words does not imply that a language cannot have some nonlogical words which refer to private objects, that is, it does not imply that PLT-1 is true. Again, that a language must have

logical words implies nothing about the possibility of a single man developing a language for and by himself, that is, does not imply that PLT-2 is true.

"**THE SAME**." Apparently Wittgenstein knew that there are no criteria (in the sense of something independent) which prove that words have been used incorrectly. He also knew that the correctness of an application of a word is not determined by a rule whose formulation serves as a recipe or canon. His fundamental opposition to private language derives from his profound investigations into the nature of concepts and his strong inclination toward an extreme nominalism. This opposition is never crystallized in a definite argument, but its gist is, in crude form, as follows. Postulate:

(*E*) The similarities and samenesses we find in things do not exist *in rerum natura,* that is, do not exist in things as we find them, independently of our finding them or of our referring to them in the way we do; they "come from the language" (Rhees p. 80) and at bottom consist of the fact that we "call" the things in question the same (Wittgenstein Secs. 146, 149, 185–190, 208–223, 348–352).

On a rigorous interpretation of (*E*), we find a rationale for assumptions (*A*), (*B*), and (*C*), as well as for the fact that the PLT has a chameleon-like and pervasive character. If we take (*E*) literally, then to find a property in several things is to find that we "call" the things in question "the same" or refer to them with the same word. Thus, it is impossible to think that something is such-and-such without a language in which there is an expression (even if a very long phrase) which "constitutes" the such-and-such in question. This is assumption (*A*). Also, (*B*), without an expression "constituting" a type *T,* there is no type *T* for things to belong to. Similarly, to understand an expression is not to apprehend an independently existing (or subsisting) property but simply to know how and to what to apply it, and this includes knowing how to call certain utterances "true" in which the expression is correctly applied. This is, in fact, a generalization of assumption (*C*).

We cannot say that a man in doubt about whether or not he used a word correctly must simply identify certain features of things without criteria and, armed with these identifications, test his uses of words. For on the extreme interpretation of assumptions (*A*) through (*E*) to identify a feature is to "call" a thing something. So, when the use of a word is at issue, the identification and nature of the thing is precisely what is at issue. The referents of one's previous uses of the word, as well as the uses themselves, are irrelevant. If one "calls" something "*A*," then it is *A* and *a fortiori* similar to the previous *A*'s; if one withholds the name "*A*" from it, then it is not an *A,* and a fortiori it is dissimilar to all *A*'s with respect to being an *A.* Clearly, it does not matter whether one's language is about private or about public objects; one's uses of words simply fail to be capable of being incorrect. They would seize reality so well that each "would have to be at once a statement and a definition" (Rhees p. 82).

Thus, the following question arises. If, on assumption (*D*), language is a matter of rules and rules are the sort of thing that can be misapplied or not, how, then, is language possible after all? At this stage, obviously, we are not interested in proving anything but are anxious to find an explanation. Wittgenstein seems to suggest one: A man's uses of words can be incorrect only if they are compared with those of his fellow speakers. His "calling" something "*A*" is correct if his cospeakers now also call it "*A*." Then it is *A* and a fortiori similar to the things he and his cospeakers previously called "*A*." That is why obeying a rule of language is a practice (Wittgenstein, op. cit., Sec. 202). It is not necessary that the speakers of the language should call the thing in question "*A*" or that they call it "*A*" afterward. Nor is it necessary that they call it "*A*" or anything at all, or that they call it the same thing. It is just a contingent fact that they coincide in calling it "*A*." But this coincidence (or agreement) is an empirical fact that is necessary for the existence of language.

Such is the underlying argument of Wittgenstein's remarks (Secs. 146, 149, 185–190, 208–223; for a discussion of the role of Wittgenstein's extreme nominalism in his views about necessary truth, see Michael Dummett's "Wittgenstein's Philosophy of Mathematics"). He builds a Heraclitean picture of language as something living only in our actual use of it and changing according to our needs. But is this a true picture of the connection between language and reality?

Here we cannot discuss the whole issue of nominalism, but to this writer it seems indefensible. We could doubtless have classified objects in entirely different ways from the ways we in fact do. For instance, we might have had no color words, no terms for species of plants or animals, and instead have used, say, "sha" for some elephants and white roses and reddish sand, and "sho" for female elephants, eggs, and rivers. But even so, we should have had to *find* features of similarity in the things so classified, and these features would have provided tests for the correct application of our words. At any rate, the view that things are the same because we "call" them "the same" or because we refer to them with the same words can get off the ground only by postulating our recogni-

tion of the samenesses of words, that is, the similarities of noises whose application to things constitutes the similarities of the latter. A serious infinite regress would ensue if we also hold that our words are similar only because we "call" them so.

The several propositions that are often debated as the claim that private languages are impossible can be linked to each other only under the assumption of extreme nominalism. None of the arguments given for the claim appear to be successful. There may be no conclusive way of either proving or refuting this claim. Perhaps the only course is to build detailed and rigorous philosophical views on each alternative and assess the adequacy of such views by their consequences. This topic continues to be widely discussed in the literature, and many philosophers adopt a position different from that advocated in the present article.

See also Carnap, Rudolf; Malcolm, Norman; Mind-Body Problem; Rule Following; Solipsism; Wittgenstein, Ludwig Josef Johann.

Bibliography

THE PROBLEM BEFORE WITTGENSTEIN

Carnap, Rudolf. "Psychology in Physical Language." In *Logical Positivism,* edited by A. J. Ayer. Glencoe, IL: Free Press, 1959. Ch. 8. Translated by Frederic (there called George) Schick from the German version, which appeared in *Erkenntnis* 3 (1932–1933).

Weinberg, J. R. *An Examination of Logical Positivism.* New York and London, 1936. Ch. 11.

WITTGENSTEIN'S VIEWS

Albritton, Rogers. "On Wittgenstein's Use of the Term 'Criterion.'" *Journal of Philosophy* 56 (1959): 845–857. An excellent collation and exegesis of the passages in which Wittgenstein discusses criteria; brings out certain internal tensions in Wittgenstein's conception.

Dummett, Michael. "Wittgenstein's Philosophy of Mathematics." *Philosophical Review* 68 (1959): 324–348.

Malcolm, Norman. "Wittgenstein's *Philosophical Investigations.*" *Philosophical Review* 63 (1954): 530–559. An extended review and discussion that presents Wittgenstein's arguments against the PLT, with emphasis on the private language speaker's lack of a criterion of correctness.

Rhees, Rush. "Can There Be a Private Language?" *PAS,* Supp., 28 (1954): 77–94. A symposium with A. J. Ayer containing an exposition of Wittgenstein's views on private language understood mainly as PLT-2; conveys Wittgenstein's nominalistic rationale very well, without discussing it as such.

Wittgenstein, Ludwig. *Philosophical Investigations.* German with facing translation by G. E. M. Anscombe. Oxford: Blackwell, 1953.

THE PROBLEM AFTER WITTGENSTEIN

Ayer, A. J. "Can There Be a Private Language?" *PAS,* Supp., 28 (1954): 63–76. A symposium with Rush Rhees in which Ayer criticizes the PLT and emphasizes the need for criterionless identification at some point.

Carney, James. "Private Language." *Mind* 64 (1960): 560–565. Favors the PLT.

Castañeda, Héctor-Neri. "Criteria, Analogy, and Knowledge of Other Minds." *Journal of Philosophy* 59 (1962): 533–546. Discusses some problems of Malcolm's views on criteria.

Castañeda, Héctor-Neri. "The Private-Language Argument." In *Knowledge and Experience,* edited by C. D. Rollins. 1962 Oberlin Philosophy Colloquium. Pittsburgh: University of Pittsburgh Press, 1963. A symposium with V. C. Chappell and J. F. Thomson; surveys critically the arguments for the PLT and also replies to the other symposiasts.

Chappell, V. C. "Comments." In *Knowledge and Experience,* edited by C. D. Rollins. Vigorously defends the PLT from criticisms raised by Castañeda's discussion.

Garver, Newton. "Wittgenstein on Private Language." *Philosophy and Phenomenological Research* 20 (1960): 389–396. Favors the PLT.

Hardin, Clyde L. "Wittgenstein on Private Language." *Journal of Philosophy* 56 (1959): 517–528. Attacks the PLT.

Hervey, Helen. "The Private Language Problem." *Philosophical Quarterly* 7 (1957): 63–79. Accepts PLT* but claims that the real problem is whether or not one can recognize and classify one's sensations apart from their outward expression.

Tanburn, N. P. "Private Languages Again." *Mind* 72 (1963): 88–102. Attacks PLT-2 and the thesis that all entities referred to in conversation must be publicly available.

Thomson, J. F. "Comments." In *Knowledge and Experience,* edited by C. D. Rollins. In his part of the symposium with Castañeda and Chappell, Thomson contends that the PLT is too unclear for decision.

Todd, W. "Private Languages." *Philosophical Quarterly* 12 (1962): 206–217. Argues that "a private language, in the sense of a personal sensation language, is logically possible," refutes certain arguments of Malcolm, and attempts to show that a private language need not lack criteria of correctness.

Wellman, Carl. "Wittgenstein and the Egocentric Predicament." *Mind* 68 (1959): 223–233. Discusses critically some of Wittgenstein's arguments for the PLT in connection with the view that all knowledge is based on private experiences.

Héctor-Neri Castañeda (1967)

PRIVATE LANGUAGE PROBLEM [ADDENDUM]

Although the proper formulation and assessment of Ludwig Wittgenstein's argument (or arguments) against the possibility of a private language continues to be disputed, the issue has lost none of its urgency. At stake is a broadly

Cartesian conception of experiences that is found today in much philosophy of mind.

WHAT IS A PRIVATE LANGUAGE?

In §243 of *Philosophical Investigations* (1967; see also §256) Wittgenstein introduces the idea of a language in which "a person could write down or give vocal expression to his inner experiences—his feelings, moods, and the rest—for his private use. … The individual words of this language are to refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language." In subsequent sections (according to some commentators, up to as far as §315) Wittgenstein criticizes the possibility of such a "private language," and this is where "the private language argument" is usually supposed to be located.

Following the main essay, suppose that Privatus speaks Privatish, a private language. §243 suggests that Privatish has two features:

(1) Privatish contains a referring expression n that refers to one of Privatus's "immediate private sensations" s.

(2) s "can only be known to" Privatus.

According to Wittgenstein Privatish has a third feature, which he apparently thinks follows from (1) and (2):

(3) n can only be understood by Privatus (and, hence, Privatish can only be completely understood by Privatus).

At this point three questions arise. First, what does (2) mean? Second, what sort of referring expression is n? Third, why is (3) supposed to follow from (1) and (2)?

By (2) Wittgenstein seems to mean that only Privatus can know whether he is having s. "The essential thing about private experience is … that nobody knows whether other people also have *this* or something else" (§272, see also §246). Of course, this conception of sensations is held by Wittgenstein's opponent (a defender of the possibility of a private language), not Wittgenstein himself.

As to the referring expression n, it is a name, not a description (e.g., "the private sensation caused by pinpricks") (see §§256–257). Not even Wittgenstein's opponent would accept that to understand a description that in fact refers to sensation s one has to know that it refers to s.

Is n a proper name of a token sensation, or is it a common noun referring to a type of sensation? If n refers to a token sensation, something occurring only in the mind of Privatus, then Wittgenstein's opponent looks exactly like Bertrand Russell in "The Philosophy of Logical Atomism" (1918/1956; see also Candlish 2004). According to Russell, "[i]n order to understand a proposition in which the name of a particular occurs, you must already be acquainted with that particular" (1918/1956, p. 204; see also Russell 1912, chapter 5). Since, on Russell's view, the only particular things with which one is acquainted are private items he calls *sense-data* (and, in addition, perhaps one's self), no two people can be acquainted with the same particular, and so no two people can understand a genuine name (as Russell puts it, a name in the "logical sense"). Hence, Russell thinks, if Privatus's name n refers to a token sensation, no one else can understand it.

However, it is clear that Wittgenstein takes the sensations in question to be types, not tokens (see, in particular, §258); accordingly, the name n is a common noun. But then Russell's views about acquaintance and understanding play no role in the justification of (3), for Russell holds that two people can be acquainted with the same property (or type), including properties of private objects. Thus, if Privatus is acquainted with a certain type of sensation s, that is no barrier, on Russell's view, to others also being acquainted with s.

So why does Wittgenstein think that (3) follows from (1) and (2)? His argument may be this: one cannot know that Privatus's name n refers to s, so one cannot know what n means, and hence one cannot understand it. But it is not obvious that knowledge that n refers to s is necessary for understanding n, or for successful communication using it: perhaps all that is required is that one believes that Privatus's name n refers to s. The upshot is that Wittgenstein's double characterization of a private language as one "which describes my inner experiences and which only I myself can understand" (§256) is contentious. (This point is due to Edward Craig [1982]; for further discussion see Craig [1997].)

Given that the two characterizations of a private language should be separated, it is probably better to use the first, leaving the second as a disputed consequence. Thus, a private language may be explained as one containing names for types of inner experiences, with the further stipulation that, if there are any inner experiences, no one knows whether others have the same types of inner experiences as him- or herself.

§258

§258 contains the famous example of keeping "a diary about the recurrence of a certain sensation," and most commentators identify it as the core of the private language argument. The key move in this section is to cast doubt on whether the diary keeper can "impress on [himself] the connexion between the sign ["S"] and the sensation," and so "remember the connexion *right* in the future." Anthony Kenny points out, against some commentators, that "remembering the connexion right" does not mean that one correctly applies "S" to one's sensation, but that one remembers the meaning of S (1973, pp. 191–193).

Supposedly, there is no fact about the meaning of "S" for the diarist to remember because there is "no criterion of correctness." Here, there is little consensus on what the missing criterion amounts to, or whether its absence does indeed show that the diarist fails to attach a meaning to "S." For some representative examples of exegesis, see Malcolm Budd (1989, chapter 3), Stewart Candlish (1980, 2003), John V. Canfield (1991, 2001), Robert J. Fogelin (1987, chapter 12), P. M. S. Hacker (1986, chapter 9; 1990, 61–67); Colin McGinn (1997, chapter 4), David Pears (1988, chapters 13, 14, 15), Scott Soames (2003, chapter 2), and Crispin Wright (1986).

KRIPKE'S WITTGENSTEIN

In *Wittgenstein on Rules and Private Language* (1982), Saul A. Kripke suggests (without unreservedly endorsing) a novel and exciting interpretation of the private language argument (see also Fogelin 1987, p. 241, n. 10). On this interpretation the main argument appears in the earlier long discussion of following a rule starting around §138. As Kripke observes, the conclusion of the private language argument is stated in §202, well before the argument's traditional location, "Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it." Kripke takes the sections following §243 to be a discussion of a purported counterexample—namely, sensation language—to the conclusion argued for earlier.

The argument Kripke extracts from Wittgenstein is in two parts. The first part purports to establish that there are no facts that make it the case that an individual (Jones, say) means something by a word (addition by "+," say). This conclusion is reached by canvassing all the plausible candidates for such meaning-constituting facts and finding them all wanting. A *skeptical paradox* ("our paradox" of §201) looms, "no facts, no truth conditions, correspond to such statements such as "Jones means addition by '+'" (1982, p. 77). The second part offers a skeptical solution: "skeptical" because the paradoxical conclusion is embraced; a "solution" because sentences such as "Jones means addition by '+'" remain assertible, despite the lack of any "corresponding fact." And the account of why such sentences are assertible essentially involves a linguistic community, so that if Jones is "considered in isolation," he cannot be said to mean anything by his words. This is the most general sense in which a "private language" is impossible: An individual considered in isolation from other speakers cannot be said to speak a language (see Kripke 1982, pp. 109–110).

Most commentators have not endorsed Kripke's interpretation (in particular, see Baker and Hacker 1984, chapter 1; McGinn 1984, chapter 2). However, Kripke's Wittgenstein has become a philosopher of considerable interest in his own right.

THE COMMUNITY VIEW

Kripke's book revived interest in the issue of whether the private language argument and related material on rule following is supposed to exclude a Robinson Crusoe isolated from birth from speaking a language (discussion of this topic goes back to Alfred J. Ayer [1954] and Rush Rhees [1954]; see also Kripke [1982, p. 110, n. 84]). While the characterization of a private language in §243 seems to leave room for such a Crusoe, other sections, notably §198, suggest the opposite. Norman Malcolm (1986, 1989) offers a defense of the "community view," and is countered by G. P. Baker and Hacker (1990). The community view is rejected by most commentators; for further discussion and references, see Canfield (1996).

See also Rule Following.

Bibliography

Ayer, Alfred J. "Can There Be a Private Language?" In *The Philosophy of Wittgenstein*. Vol. 9, edited by John V. Canfield. New York: Garland, 1986. Ayer's essay originally published in the journal *Proceedings of the Aristotelian Society* in 1954.

Baker, G. P., and P. Hacker. "Malcolm on Language and Rules." *Philosophy* 65 (1990): 167–179.

Baker, G. P., and P. M. S. Hacker. *Scepticism, Rules, and Language*. Oxford, U.K.: Basil Blackwell, 1984.

Budd, Malcolm. *Wittgenstein's Philosophy of Psychology*. London: Routledge, 1989.

Candlish, Stewart. "Private Language." *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 2003. http://plato.stanford.edu/archives/spr2004/entries/private-language/

Candlish, Stewart. "The Real Private Language Argument." *Philosophy* 55 (1980): 85–94.
Canfield, John V. "The Community View." *Philosophical Review* 105 (1996): 469–488.
Canfield, John V., ed. *The Philosophy of Wittgenstein, Vol. 9: The Private Language Argument*. New York: Garland, 1986.
Canfield, John V. "Private Language." In *Wittgenstein's Philosophical Investigations: Text and Context*, edited by Robert L. Arrington and Hans-Johann Glock. London: Routledge, 1991.
Canfield, John V. "Private Language: The Diary Case." *Australasian Journal of Philosophy* 79 (2001): 377–394.
Craig, Edward. "Meaning and Privacy." *A Companion to the Philosophy of Language*, edited by Robert Hale and Crispin Wright, 127–145. Oxford, U.K.: Blackwell, 1997.
Craig, Edward. "Meaning, Use, and Privacy." *Mind* 91 (1982): 541–564.
Fogelin, Robert J. *Wittgenstein*. 2nd ed. London: Routledge and Kegan Paul, 1987.
Hacker, P. M. S. *Insight and Illusion: Themes in the Philosophy of Wittgenstein*. Rev. ed. New York: Oxford University Press, 1986.
Hacker, P. M. S. *Wittgenstein: Meaning and Mind*. Oxford, U.K.: Basil Blackwell, 1990.
Kenny, Anthony. *Wittgenstein*. London: Penguin, 1973.
Kripke, Saul A. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Oxford, U.K.: Blackwell, 1982.
Malcolm, Norman. *Nothing Is Hidden*. Oxford, U.K.: Blackwell, 1986.
Malcolm, Norman. "Wittgenstein on Language and Rules." *Philosophy* 64 (1989): 5–28.
McGinn, Colin. *Wittgenstein on Meaning*. Oxford, U.K.: Basil Blackwell, 1984.
McGinn, Marie. *Routledge Philosophy Guidebook to Wittgenstein and the Philosophical Investigations*. London: Routledge, 1997.
Pears, David. *The False Prison*. Vol. 2. Oxford, U.K.: Clarendon Press, 1988.
Rhees, Rush. "Can There Be a Private Language?" In *The Philosophy of Wittgenstein*, edited by John V. Canfield. New York: Garland, 1986. Rhees's essay originally published in the journal *Proceedings of the Aristotelian Society* in 1954.
Russell, Bertrand. "The Philosophy of Logical Atomism" (1918). In *Logic and Knowledge: Essays, 1901–1950*, edited by Robert Charles Marsh. London: George Allen and Unwin, 1956.
Russell, Bertrand. *The Problems of Philosophy*. New York: Henry Holt, 1912.
Soames, Scott. *Philosophical Analysis in the Twentieth Century*. Vol. 2. Princeton, NJ: Princeton University Press, 2003.
Wittgenstein, Ludwig. *Philosophical Investigations*. 2nd ed. Translated by G. E. M. Anscombe. Oxford, U.K.: Blackwell, 1967.
Wright, Crispin. "Does *Philosophical Investigations I. 258–60* Suggest a Cogent Argument against Private Language?" In *Subject, Thought, and Context*, edited by Philip Pettit and John McDowell, 209–266. New York: Oxford University Press, 1986.

Alex Byrne (2005)

PROBABILITY AND CHANCE

The weather report says that the chance of a hurricane arriving later today is 90 percent. Forewarned is forearmed: Expecting a hurricane, before leaving home I pack my hurricane lantern.

Probability enters into this scenario twice, first in the form of a physical probability, sometimes called a chance, quantifying certain aspects of the local weather that make a hurricane very likely, and second in the form of an epistemic probability capturing a certain attitude to the proposition that a hurricane will strike, in this case one of considerable confidence.

It is not immediately obvious that these two probabilities are two different kinds of thing, but a prima facie case can be made for their distinctness by observing that they can vary independently of one another: For example, if the meteorologists are mistaken, the chance of a hurricane may be very low though both they and I am confident that one is on its way.

Most philosophers now believe that the apparent distinctness is real. They are therefore also inclined to say that my belief that the physical probability of a hurricane is very high is distinct from my high epistemic probability for a hurricane. There must be some principle of inference that takes me from one to the other, a principle that dictates the epistemic impact of the physical probabilities—or at least, of my beliefs about the physical probabilities—telling me, in the usual cases, to expect what is physically probable and not what is physically improbable. One can call such a principle, mediating as it does between two different kinds of probability, a probability coordination principle.

The three principal topics of this entry will be, in the order considered, epistemic probability, physical probability, and probability coordination. Two preliminary sections will discuss the common mathematical basis of epistemic and physical probability and the classical notion of probability.

THE MATHEMATICAL BASIS

What all probabilities, epistemic and physical, have in common is a certain mathematical structure. The most important elements of this structure are contained in the axioms of probability, which may be paraphrased as follows:

(1) All probabilities are real numbers between zero and one inclusive (for any proposition *a*, $0 \leq P(a) \geq 1$).

(2) The probability of an inconsistent proposition is zero; the probability of a logical truth, or tautology, is one.

(3) The probability that either one or the other of two mutually exclusive propositions is true is equal to the sum of the probabilities of the individual propositions. (Two propositions are mutually exclusive if they cannot both be true; the *cannot* is interpreted as a matter of logical consistency, so that the axiom says that for any two propositions *a* and *b* such that $a \vdash \neg b$, $P(a \vee b) = P(a) + P(b)$.)

The axioms as stated here assume that probabilities are attached to propositions, such as the proposition that "A hurricane will strike New York at some time on the afternoon of January 20, 2005." The axioms may also be stated in a way that assumes that probabilities attach to events. It is more natural to attach epistemic probabilities to propositions and physical probabilities to events, but when the two kinds of probability are discussed side by side it is less confusing, and quite tolerable, to take propositions as the primary bearers of both kinds of probability. Nothing important is thought to turn on the choice.

The three axioms of probability, though simple, may be used to prove a wide range of interesting and strong mathematical theorems. Because all probabilities conform to the axioms, all probabilities conform to the theorems. It is possible, then, to do significant work on probability without presupposing either epistemic or physical probability as the subject matter, let alone some particular construal of either variety. Such work is for the most part the province of mathematicians.

Philosophical work on probability may also be mathematical, but is most often directed to one or the other variety of probability, usually attempting a philosophical analysis of probability statements made in a certain vein, for example, of probability claims made in quantum mechanics or evolutionary biology (both apparently claims about physical probability) or of probability claims made in statistical testing or decision theory (both apparently claims about epistemic probability).

Two important notions encountered in statements of the mathematical behavior of probability are conditional probability and probabilistic independence. Both are introduced into the mathematics of probability by way of definitions, not additional axioms, so neither adds anything to the content of the mathematics.

CONDITIONAL PROBABILITY. The probability of a proposition *a* conditional on another proposition *b*, written $P(a|b)$, is defined to be $P(ab)/P(b)$, where *ab* is the conjunction of *a* and *b*. (The conditional probability is undefined when the probability of *b* is zero.) For example, the probability of obtaining three heads on three successive tosses of a coin, conditional on the first toss yielding heads, is the probability of obtaining three heads in a row, namely one-eighth, divided by the probability of obtaining heads on the first coin, namely one-half—in other words, one-quarter.

Some writers suggest taking conditional probability as the basis for all of probability mathematics, a move that allows, among other things, the possibility of conditional probabilities that are well defined even when the probabilities of the propositions conditionalized on are zero (Hájek 2003). On this view, the mathematical posit stated above linking conditional and unconditional probabilities is reinterpreted as an additional axiom.

The act of conditionalization may be used to create an entirely new probability distribution. Given an old probability distribution $P(\dot{s})$ and a proposition *b*, the function $P(\dot{s}|b)$ is provably also, mathematically speaking, a probability distribution. If *k* is a proposition stating all of one's background knowledge, for example, then a new probability distribution $P(\dot{s}|k)$ can be formed by conditionalizing on this background knowledge, a distribution that gives, intuitively, the probabilities for everything once one's background knowledge is taken into account. This fact is especially important in the context of epistemic probability.

PROBABILISTIC INDEPENDENCE. Two propositions *a* and *b* are probabilistically independent just in case $P(ab) = P(a)P(b)$. When the probability of *b* is nonzero, this is equivalent to $P(a|b) = P(a)$, or in intuitive terms, the claim that the truth or otherwise of *b* has no impact on the probability of *a*.

Several of the most important and powerful theorems in probability mathematics make independence assumptions. The theorem of most use to philosophers is the law of large numbers. The theorem says, very roughly, when a large, finite set of propositions are independent, but have the same probability *p*, then the proportion of propositions that turn out to be true will, with high probability, be approximately equal to *p*. (The generalization to countably infinite sets of propositions is easy if the propositions are ordered; substitute *limiting frequency* for *proportion*.)

For example, the propositions might all be of the form "Coin toss *x* will produce heads," where the *x* stands for any one of a number of different tosses of the same coin. If the probability of each of the propositions is one-half, then the law of large numbers says, in effect, that provided the tosses are independent, it is very likely that about one-half will yield heads.

It is natural to interpret the probabilities in this example as physical probabilities, but the law of large numbers applies equally to any kind of probability, provided that independence holds. There are, in fact, many variants of the law of large numbers, but the details are beyond the scope of this entry.

CLASSICAL PROBABILITY

The development of the mathematics, and then the philosophy, of probability was spurred to a perhaps surprising degree by an interest, both practical and theoretical, in the properties of simple gambling devices such as rolled dice, tossed coins, and shuffled cards. Though there was from the beginning a great enthusiasm for extending the dominion of the "empire of chance" to the ends of the earth, gambling devices were—and to some extent are still—the paradigmatic chance setups.

A striking feature of gambling devices is their probabilistic transparency: The discerning eye can "read off" their outcomes' physical probabilities from various physical symmetries of the device itself, seeing in the bilateral symmetry of the tossed coin a probability of one-half each for heads and tails, or in the six-way symmetry of the die a probability of one-sixth that any particular face is uppermost at the end of a roll (Strevens 1998).

The classical definition of probability, paramount from the time of Gottfried Wilhelm Leibniz to the time of Pierre Simon de Laplace (the late seventeenth century to the early nineteenth century) takes its inspiration from the alignment of probability with symmetry. The best-known formulation of the classical account is due to Laplace:

> The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible. (1902, pp. 6–7)

As many commentators note, this formulation, typical of the classical probabilists, appears to involve two parallel definitions, the first based on the notion of equal possibility and the second on the notion of equal undecidedness. Laplace's relation of equal possibility between two cases probably ought to be understood as picking out a certain physical symmetry in virtue of which the cases have equal physical probabilities. All classical probabilities, on the equal possibility definition, have their basis in such physical symmetries, and so would seem to be physical probabilities. The relation of equal undecidedness between two cases refers to some sort of epistemic symmetry, though perhaps one founded in the physical facts. A probability with its basis in undecidedness would seem to be, by its very nature, an epistemic probability. Classical probability, then, is at the same time a kind of physical probability and a kind of epistemic probability.

This dual nature, historians argue, is intentional (Hacking 1975, Daston 1988). In its epistemic guise classical probability can be called on to do work not normally thought to lie within the province of an objective notion of probability, such as measuring the reliability of testimony, the strength of evidence for a scientific hypothesis, or participating in decision-theoretic arguments such as Blaise Pascal's famous wager on the existence of God. In its physical guise classical probability is able to cloak itself in the aura of unrevisability and reality that attaches to the gambling probabilities such as the one-half probability of heads.

The classical definition could not last. Gradually, it came to be acknowledged that although the epistemic probabilities may, or at least ought to, shadow the physical probabilities wherever the latter are found, they play a number of roles in which there is no physical probability, nor anything with the same objective status as a physical probability to mimic. The classical definition was split into its two natural parts, and distinct notions of physical and epistemic probability were allowed to find their separate ways in the world.

At first, in the middle and late nineteenth century, physical probability commanded attention almost to the exclusion of its epistemic counterpart. Developments in social science, due to Adolphe Quetelet (1796–1874), in statistical physics, due to James Clerk Maxwell and Ludwig Boltzmann, and eventually (around 1930) in the synthesis of evolutionary biology and genetics, due to Ronald Aylmer Fisher and many others, turned on the successful deployment of physical probability distributions. Begin-

ning in the early twentieth century, however, epistemic probability came into its own, freeing itself over the decades from what came to be seen as the classical probabilists' futile attempt to provide strict guidelines dictating unique rational epistemic probabilities in every situation.

Modern philosophy remade itself in the twentieth century, imposing a historical horizon at around 1900. The story of the interpretation of probability is often told beginning near that year, with the result that the development of epistemic probability, and logical probability in particular, comes first—a convention that will be followed here.

EPISTEMIC PROBABILITY

Epistemic probability takes two forms. In its first form, it is a measure of a person's degree of confidence in a proposition, increasing from zero to one as his or her attitude goes from almost total disbelief to near certainty. This kind of epistemic probability is called credence, degree of belief, or subjective probability. The propositional attitude one gets when one attaches a subjective probability to a proposition is sometimes called a partial belief.

In its second form, associated most often with the term *logical probability*, epistemic probability measures the impact of a piece or pieces of evidence on a proposition. Its elemental form may not be that of a probability distribution, in the usual sense, but it is related to a probability distribution in some straightforward way, and as will be seen shortly, is quite capable of providing a basis for a complete system of epistemic probability.

There is a foundational dispute between the proponents of the two forms of epistemic probability. It is not a fight for existence but for primacy: The question is which of the two kinds of epistemic probability is the more epistemologically basic.

LOGICAL PROBABILITY. The second form of epistemic probability has, since 1900, most often taken the guise of logical probability. A logical probability is attached not to a proposition but to a complete inductive inference. It is a measure of the degree to which the evidence contained in the premises of an inductive inference, considered in isolation, probabilifies the conclusion. The idea of probabilistic inference was an important part of classical probability theory, but from the post-1900 perspective it is associated first with John Maynard Keynes (1921)—who was more famous, of course, as an economist.

In explaining the nature of logical probability, and in particular the tag *logical* itself, Keynes draws a close analogy with deductive inference: Whereas in a deductive inference the premises entail the conclusion, in an inductive inference they partially entail the conclusion, the degree of entailment being represented by a number between zero and one, namely, a logical probability. (Note that a degree zero entailment of a proposition is equivalent to full entailment of the proposition's negation.) Just as the first form of epistemic probability generalizes from belief to partial belief, then, the second form generalizes, in Keynes's hands, from entailment to partial entailment.

For example: Take as a conclusion the proposition that the next observed raven will be black. A proposition stating that a single raven has been observed to be black might entail this conclusion only to a relatively small degree, this logical probability representing the slightness of a single raven's color as evidence for the color of any other raven. A proposition stating that many hundreds of ravens have been observed to be black will entail the conclusion to some much greater degree.

It is an objective matter of fact whether one proposition deductively entails another; so, Keynes conjectured, it is in many cases a matter of objective fact to what degree one proposition partially entails another. These facts themselves comprise inductive logic; the logical probabilities are at base, then, logical entities, just as the name suggests.

Although exact logical probabilities are for Keynes the ideal, he allows that in many cases logic will fix only an approximate degree of entailment for an inductive inference. The presentation in this entry will for simplicity's sake focus on the ideal case.

Keynes's logical probability is not only compatible with subjective probability, the other form of epistemic probability; it also mandates certain values for a person's subjective probabilities. If the premises in an inductive inference are known for certain, and they exhaust the available evidence, then their inductive impact on the conclusion—the degree of entailment, or logical probability attached to the inference, from the premises to the conclusion—is itself the degree of belief, that is, the subjective probability, that a rational person ought to attach to the conclusion, reflecting as it does all and only the evidence for the conclusion.

Keynes uses this argument as a basis for taking as a formal representation of logical probabilities the probability calculus itself: The degree to which proposition *b* entails proposition *a* is written as a conditional probabil-

ity $P(a|b)$. Note that these probabilities do not change as the evidence comes in, any more than facts about deductive entailment can change as the evidence comes in. The logical probability $P(a|b)$ must be interpreted as a quantification of the inductive bearing of *b* alone on *a*, not of *b* together with some body of accepted knowledge.

The unconditional probability $P(a)$, then, is the inductive bearing on *a* of an empty set of evidence—the degree to which *a* is entailed, in Keynes's sense, by the set of logical truths, or tautologies, alone. One might think that the degree of entailment is zero. But this cannot be right: If one has no evidence at all, one must set one's subjective probabilities for both *a* and its negation equal to their respective degrees of entailment by the tautologies. But one cannot set both subjective probabilities to zero—it cannot be that one is certain that neither *a* nor its negation is true, since one of the two must be true. One's complete lack of evidence would be better represented by setting both subjective probabilities to intermediate values, say one-half. The logical probabilist, in endorsing this assignment, implicitly asserts that the empty set of evidence, or the set of tautologies, entails both *a* and its negation to degree one-half.

Although its subject matter is the bearing of evidence on hypotheses, then, logical probability theory finds itself having to take a position on what one should believe when one has no evidence (under the guise of the question of the tautologies' partial entailments). To answer this question, it has turned to the principle of indifference, which recommends—when there is no evidence favoring one of several mutually exclusive possibilities over the others—that the available probability be equally distributed among them. This is, of course, the same principle that comprises one strand of the classical definition of probability: Laplace suggested assigning equal probabilities to cases "such as we may be equally undecided about in regard to their existence" (Laplace 1902, p. 6). It has also played an important role in the development of the theory of subjective probability, and so is discussed in a separate section later in this entry.

As the role of indifference shows, logical probability is close in spirit to the epistemic strand of classical probability. It posits, at least as an ideal, a single system of right reasoning, allowing no inductive latitude whatsoever, to which all rational beings ought to conform. Insofar as rational beings ever disagree on questions of evidential impact, it must be because they differ on the nature of the evidence itself.

Many philosophers find this ideal of inductive logic hard to swallow; even those sympathetic to the idea of strong objective constraints on inductive reasoning are often skeptical that the constraints take the form of logical truths, or something analogous to logical truths. This skepticism has two sources.

First is the perception that inductive practices vary widely. Whereas there exists a widespread consensus as to which propositions deductively entail which other propositions, there is no such consensus on degrees of evidential support. That is not to say, of course, that there is disagreement about every aspect of inductive reasoning, but there is far less agreement than would be necessary to build, in the same way that deductive logic was constructed, a useful inductive logic.

Second, there are compelling (though not irresistible) reasons to believe that it is impossible to formulate a principle of indifference that is both consistent and strong enough to do the work asked of it by logical probabilists. These reasons are sketched in the discussion of the principle later on.

Rudolf Carnap (1950) attempted to revive the idea of a system of induction founded on logic alone in the mid-century. His innovation—drawing on his general philosophy of logic—was to allow that there are many systems of inductive logic that are, from a purely logical viewpoint, on a par. One may freely choose from these a logic, that is, a set of logical probabilities, that suits one's particular nonlogical ends.

Carnap relativized induction in two ways. First, his version of the principle of indifference was indexed to a choice of language; how one distributes probability among rival possibilities concerning which one knows nothing depends on one's canonical system for representing the possibilities. Second, even when a canonical language is chosen, Carnap's rule for determining inductive support—that this, degrees of entailment or logical probabilities—contains a parameter whose value may be chosen freely. The parameter determines, roughly, how quickly one learns from the evidence. Choose one extreme, and from the observation of a single black raven one will infer with certainty that the next raven will also be black (straight induction). Choose the other extreme, and no number of black ravens is great enough to count as any evidence at all for the blackness of the next raven. A sensible choice would seem to lie somewhere in the middle, but on Carnap's view, logic alone determined no preference ranking whatsoever among the different choices, rating all values apart from the extremes as equally good.

Carnap did give extralogical arguments for preferring a particular value for the parameter, arriving at an inductive rule equivalent to Laplace's rule of succession. Given that, say, i out of n observed ravens have been black, both Carnap and Laplace assign a probability of $(i + 1)/(n + 2)$ to the proposition that the next raven will be black.

One awkward feature of Carnap's system is that, no matter what value is chosen for the inductive parameter, universal generalizations cannot be learned: The inductive bearing of any number of black ravens on the hypothesis "All ravens are black" is zero.

Carnap's system is of great intrinsic interest, but from the time of its presentation, its principal constituency—philosophers of science—was beginning to move in an entirely different direction. Such considerations as Nelson Goodman's new riddle of induction and arguments by Bayesians and others that background knowledge played a part in determining degrees of inductive support, though not beyond the reach of Carnap's approach, strongly suggested that the nature of inductive support could not be purely logical.

Today, the logical approach to inductive inference has been supplanted to a great extent by (though not only by) the Bayesian approach. Still, in Bayesianism itself some have seen the seeds of a new inductive logic.

SUBJECTIVE PROBABILITY. Whereas logical probability is a logical entity—a quantification of the supposed logical facts about partial entailment—the other kind of epistemic probability, subjective probability, is a psychological entity, reflecting an actual cognitive fact about a particular person or (if they are sufficiently agreed) a group of people. The rationality of a person's subjective probabilities may be a matter of logic, then, but the probabilities themselves are a matter of psychology.

That for a number of propositions one tends to have a degree of confidence intermediate between the extremes associated with total disbelief and total belief, no one will deny. The advocates of subjective probability as a key epistemological notion—who call themselves Bayesians or simply subjectivists—go much further than this. They characteristically hold that humans have, or ought to have, well-defined subjective probabilities for every proposition and that these subjective probabilities play a central role in epistemology, both in inductive inference, by way of Thomas Bayes's (1702–1761) conditionalization rule, and in practical deliberation, by way of the usual mechanisms of decision theory.

The subjectivist's first challenge is to give a substantial characterization of subjective probability and to argue that subjective probabilities are instrumental in human cognition, while at the same time finding a foothold in the descriptive, psychological scheme for the normative concerns of epistemology. Much of this groundwork was laid in Frank Plumpton Ramsey's influential paper "Truth and Probability" (1931).

Ramsey does not define subjective probability as such, and even goes so far as to acknowledge that the ideal of a definite subjective probability for every proposition is just that—an ideal that goes a long way toward capturing actual human epistemology without being accurate in every respect. What he posits instead is a connection—whether conceptual or empirical he does not say—between the value of a person's subjective probability for a proposition and his or her betting behavior.

If one has a subjective probability p for a proposition a, Ramsey claims, one will be prepared to accept odds of up to p: $(1 - p)$ on the truth of a. That is, given a game in which one stands to win \$$n$ if a is true, one will pay up to \$$pn$ to play the game; equivalently, if one will pay up to \$$m$ to play a game in which one stands to win \$$n$ if a is true, one's subjective probability for a must be m/n. (Decision theorists, note, talk about utility, not dollars.)

Importantly, all human choice under uncertainty is interpreted as a kind of betting. For example, suppose I have to decide whether to wear a seat belt on a long drive. I am in effect betting on whether I will be involved in an auto accident along the way. If the cost of wearing a belt, in discomfort, inconvenience, and forsaken cool, is equivalent to losing \$$m$, and the cost of being beltless in an accident, in pain, suffering, and higher insurance premiums, is \$$n$, then I will accept the risk of going beltless just in case my subjective probability for there being an accident is less than or equal to m/n. (Here, the "prize" is negative. The cost of playing is also negative, so just by agreeing to play the game, I gain something: the increase in comfort, cool, and so on. My aim is to play while avoiding a win.) The central doctrine of decision theory is, then, built into the characterization of subjective probability.

Ramsey (1931) uses this fact to argue that, provided a person's behavior is coherent enough to be described, at least approximately, by the machinery of decision theory, his or her subjective probabilities for any proposition may be inferred from his or her choices. In effect, the person's subjective probabilities are inferred from the nature of the bets, in the broadest sense, he or she is prepared to accept. Because one's overt behavior can be systematized,

approximately, using a decision-theoretic framework, one must have subjective probabilities for every proposition, and these probabilities must play a central role in one's decision theory.

What is the force of the *must* in the preceding sentence? That depends on the nature of the posit that one having a certain subjective probability for a proposition means that one is prepared to accept certain odds on the proposition's being true. Some writers, especially in the midcentury heyday of conceptual analysis and psychological behaviorism, interpret the posit as a definition of subjective probability; on this view, one having certain subjective probabilities just is one having a certain betting behavior. Others, like Ramsey (1931), opt for a looser connection. On any approach, there is a certain amount of latitude in the phrase "prepared to accept." If I am prepared to accept certain odds, *must* I play a game in which I am offered those odds? Or only if I am in a betting mood? The former answer vastly simplifies the subjectivist enterprise, but at a cost in psychological plausibility: It is surely true that people frequently gamble in the broad sense that they take measured risks, but it is not nearly so obvious that they are compulsive gamblers intent on taking on every favorable risk they can find. Work on the psychology of decision making also suggests that it is a mistake to found the subjectivist enterprise on too strong a conception of the connection between subjective probability and betting behavior.

Subjective probabilities are supposed to conform, as the name suggests, to the axioms of probability theory. In a theory such as Ramsey's (1931), a certain amount of probability mathematics is built into the technique for extracting the subjective probabilities; that humans not only have subjective probabilities, but arrange them in accord with the axioms, is a condition for the success of Ramsey's (1931) project.

Insofar as subjective probability is not simply defined as whatever comes out of the Ramsey project, however, there is a question whether subjective probabilities obey the axioms. If they do not, there is little that they are good for, so the question is an important one for subjectivists, who tend to follow Ramsey in giving a normative rather than a descriptive answer: It is rational to arrange one's subjective probabilities in accordance with the axioms. (It is not unreasonable, of course, to see this normative claim, if true, as evidence for the corresponding descriptive claim, since humans are in certain respects reliably rational.)

The vehicle of Ramsey's argument is what is called the Dutch book theorem: It can be shown that, if one's subjective probabilities violate the axioms, then one will be prepared to accept certain sets of bets (which bets depends on the nature of the violation) that will cause one a sure loss, in the sense that one will lose whether the propositions that are the subjects of the bets turn out to be true or false.

The details of the argument are beyond the scope of this entry (for a more advanced introduction, see Howson and Urbach 1993), but an example will illustrate the strategy. The axioms of the probability calculus require that the probability of a proposition and that of its negation sum to one. Suppose one violates this axiom by assigning a probability of 0.8 both to a certain proposition a and to its negation. Then one is prepared to accept odds of 4:1 on both a and $\neg a$, which means a commitment to playing, at the same time, two games, in one of which one pays $8 and wins $10 (i.e., one's original $8 plus a $2 profit) if a is true, and in one of which one pays $8 and wins $10 if a is false. Whether a is true or false, one pays $16 but wins only $10—a certain loss. To play such a game is irrational; thus, one should conform one's subjective probabilities to the probability calculus. Needless to say, the Dutch book argument works best on the dubious interpretation of "prepared to accept" as equivalent to "compelled to accept"; there have been many attempts to reform or replace the argument with something that makes weaker, or even no, assumptions about betting behavior.

Subjectivism has been developed in several important directions. First are various weakenings or generalizations of the subjectivist machinery. The question of the connection between subjective probability and betting behavior is, as noted, one locus of activity. Another attempts to generalize the notion of a subjective probability to a subjective probability interval, the idea being that where one does not have an exact subjective probability for a proposition, one may have an approximate level of confidence that can be captured by a mathematical interval, the equivalent of saying that one's subjective probability is indeterminately somewhere between two determinate values.

Second, and closely related, is all the work that has been put into developing decision theory over the last 100 years (e.g., see Jeffrey 1983). Finally, subjectivism provides the foundation for the Bayesian theory of inference. At the root of the Bayesian system is a thought much like the logical probabilist's doctrine that, if k is one's background knowledge, then one's subjective probability for a hypothesis a ought to be $P(a|k)$. Whereas for a logical probabilist a conditional probability $P(a|b)$ is a timeless

logical constant, for a subjectivist it is something that constantly changes as further evidence comes in (even holding *a* and *b* fixed). For this reason, the subjectivist theory of inference must be an inherently dynamic theory; what is perhaps its best-known weakness, the "problem of old evidence," arises from this fact.

Subjectivism had almost entirely eclipsed logical probabilism by the late twentieth century; as the celestial metaphor unwittingly implies, however, there is a cyclic aspect to philosophical history: An interest in the central notion of logical probability theory, evidential weight, is on the rise.

There are three strands to this new movement. First is the perception among philosophers of science that scientific discourse about evidence is almost never about the subjective probability scientists should have for a hypothesis, and almost always about the degree of support that the evidence lends to the hypothesis. Second is the development of new and safer (though limited) versions of the principle of indifference. Third is technical progress on the project of extracting from the principles of Bayesian inductive inference a measure of weight. Note that this third project conceives of inductive weight as something derived from the more basic Bayesian principles governing the dynamics of subjective probability, a view opposed to the logical probabilists' derivation of rational subjective probabilities from the (by their lights) more basic logical principles governing the nature of inductive support.

INDIFFERENCE. The principle of indifference distributes probability among various alternatives—in the usual case, mutually exclusive and exhaustive propositions—concerning which little or nothing is known. The principle's rationale is that certain probability distributions reflect ignorance better than others. If I know nothing that distinguishes two mutually exclusive possibilities, picked out by propositions *a* and *b*, then I have no reason to expect one more than the other: I should assign the propositions equal probabilities. Any asymmetric assignment, say assigning twice the probability to *a* that I assign to *b*, would reflect some access on my part to facts supporting *a* at the expense of *b*. Thus, ignorance and probabilistic symmetry ought to go hand in hand—or so the principle of indifference would have it.

The principle is an essential part of logical probability theory, for the reasons given earlier, but there have always been subjectivists who appeal to the principle as well. It is most useful within the Bayesian approach to inductive inference.

The epistemic strand of classical probability theory also invokes the principle, of course, blending it with the discernment of "equally possible cases" in the paradigmatic gambling setups. This conflation has confused the discussion of the principle ever since, with proponents of the principle continuing to take aid and comfort in the principle's apparent virtuoso handling of cases such as the one-half probability of heads. One's reasoning about the gambling probabilities, however, as the classical probabilists for the most part themselves dimly saw, is a matter of inferring physical probabilities from physical symmetries, not of setting epistemic probabilities to reflect symmetric degrees of ignorance (Strevens 1998).

The most famous arguments against the principle of indifference were developed in the nineteenth century, which was a time of hegemony for physical over epistemic probability. They take their name from Joseph Bertrand (1822–1900), who pointed to the difficulty of finding a unique symmetry in certain indifference-style problems.

Consider, for example, two leading theories of dark matter in the universe: the MACHO and the WIMP theories. Each posits a certain generic form for dark matter objects, respectively large and small. If one has no evidence to distinguish them, it seems that the principle of indifference directs one to assign each a probability of one-half (assuming for the sake of the argument that there are no other possibilities). But suppose that there are four distinct schools of thought among the MACHO theorists, corresponding to four distinct ways that MACHOs might be physically realized, and eight such schools of thought among WIMP theorists. Now there are twelve possibilities, and once probability is distributed equally among them, the generic MACHO theory will have a probability of one-third and the WIMP theory a probability of two-thirds. Cases such as this make the principle seem capricious, if not simply inconsistent (as it would be if it failed to pick out a privileged symmetry).

Matters become far worse, as Bertrand noted, when there are uncountably many alternatives to choose among, as is the case in science when the value of a physical parameter, such as the cosmological constant, is unknown. Even in the simplest of such cases, the principle equivocates (Van Fraassen 1989, chapter 12). As noted earlier, some progress has been made in solving these problems, with Edwin T. Jaynes (1983) being a ringleader. Most philosophers, though, doubt that there will ever be a workable principle of indifference suited to the needs of general inductive inference.

PHYSICAL PROBABILITY

The paradigms of physical probability are the probabilities attached to gambling setups; there are, however, many more interesting examples: the probabilities of quantum mechanics and kinetic theory in physics, the probabilities of population genetics in evolutionary theory, actuarial probabilities such as the chance of dying before reaching a certain age, and the probabilities in many social science models. It is by no means clear that there is a single phenomenon to be explained here; the physical probabilities ascribed to phenomena by the best scientific theories may differ in their makeup from theory to theory. There is a commonality in the phenomena themselves, however: Whenever the notion of physical probability is put to scientific work, it is to predict or explain what might be called probabilistic patterns of outcomes. These patterns are characterized by a certain kind of long-run order, discernible only over a number of different outcomes, and a certain kind of short-term disorder, the details of the order and disorder depending on the variety of probability distribution.

The simplest and best-known of the patterns is the Bernoulli pattern, which takes its name from the corresponding probability distribution. This is the pattern typical of the outcomes produced by gambling devices, such as the pattern of heads and tails obtained by tossing a coin. The long-term order takes the form of a stable frequency equal to the corresponding probability. In the case of the tossed coin, this is of course the one-half frequency with which heads and tails occur (almost always) in the long run. The short-term disorder, though an objective property of the pattern itself, is perhaps best gotten at epistemically: Once one knows that the long-run frequency of heads is one-half, the outcome of one toss provides no useful information about the outcome of the next. The law of large numbers implies that a chance setup will produce its characteristic probabilistic patterns in the long run with very a high (physical) probability. When discussing physical probability, it is more natural to talk of probabilities attaching to events than to propositions; what follows will be formulated accordingly.

THE FREQUENCY THEORY The frequentist theory of physical probability has its roots in the empiricist interpretation of law statements according to which they assert only the existence of certain regularities in nature (on the regularity theory, see Armstrong 1983). What is usually called the actual frequency theory of probability understands physical probability statements, such as the claim that the probability of a coin toss's yielding heads is one-half, as asserting in a like spirit the existence of the appropriate probabilistic patterns—in the case of the coin toss, for example, a pattern of heads and tails in the actual outcomes of coin tosses exemplifying both the order and the disorder characteristic of the Bernoulli patterns.

The characteristic order in a Bernoulli pattern is a long-run frequency approximately equal to the relevant probability; in the case of the coin, then, it is a long-run frequency for heads of one-half. It is from this aspect of the pattern that frequentism takes its name. (One complication: A distinction must be made between the case in which the set of events exemplifying the pattern is finite and the case in which it is countably infinite. In a finite case, what matters is the proportion or relative frequency, whereas in the infinite case, it is instead the limiting frequency, that is, the value of the relative frequency in the limit, if it exists, as it must for the Bernoulli pattern to exist.)

Although their account is named for frequencies, most frequentists insist also on the presence of appropriate short-term disorder in the patterns. It is less easy to characterize this disorder in the purely extensional terms implicit in a commitment to regularity metaphysics. Suffice it to say that there is a broad range of characterizations, some strict, some rather lax. Among frequentists, Richard von Mises (1957) tends to a strict and Hans Reichenbach (1949) to a lax requirement (though Reichenbach holds, characteristically, that there is no uniquely correct level of strictness; for a discussion of the technical problems in constructing such a requirement, see Fine [1973]).

The probability that a particular coin toss lands heads is one-half, according to frequentism, because the outcome of the toss belongs to a series that exemplifies the Bernoulli pattern with a frequency of one-half. The truth-maker for the probability claim is a fact, then, about a class of outcomes, not just about the particular outcome to which the probability is nominally attached. But which class? If one is tossing an American quarter, does the class include all American quarters? All American and Canadian quarters? All fair coins? Or—ominously—all coin tosses producing heads? To give an answer to this question is to solve what has become known as the problem of the reference class.

The standard frequentist solution to the problem is to understand probability claims as including a (perhaps implicit) specification of the class. All physical probability claims are, in other words, made relative to a reference class. This doctrine reveals that the frequency theory is

best seen as an account, in the first instance, of statements of statistical laws. A claim about the one-half probability of heads, for example, is on the frequency interpretation in essence a statement of a probabilistic law concerning a class of coin tosses, not a claim about a property of a particular toss.

The kinship between the regularity account of deterministic laws and the frequency account of probability is, then, even closer than it first appears. Note that the regularity account has its own analog of singular probability claims, namely, singular claims about deterministic tendencies, such as a particular brick's tendency to fall to earth when released. Regularity theorists interpret a tendency claim not as picking out an intrinsic property of the object possessing the tendency, but as a veiled law statement.

The case of probability introduces a complication, however, that is not present in the case of exceptionless regularities: A particular coin toss will belong to many reference classes, some with different frequencies for heads. There may be, then, no determinate fact of the matter about an individual coin toss's probabilistic tendency to produce heads, or equivalently, about what are often called single case probabilities. Frequentists have made their peace with this consequence of their view.

Opponents of the frequency view argue that single-case probabilities are metaphysically, inductively, and explanatorily indispensable. Are they right? Here is the case for metaphysical indispensability: Some writers, especially propensity theorists, hold that there is clearly a fact of the matter about the value of the probability that some particular coin toss lands heads, independent of any choice of reference class. Frequentists may simply deny the intuition or may try explain away the appearance of a single-case fact (for related versions of the explanation, see Reichenbach 1949, §68; Strevens 2003, pp. 61–62).

And here is the case for predictive indispensability: To settle, for predictive and decision-theoretic purposes, on a rational subjective probability for an event using the probability coordination principle, a corresponding physical probability must be found (see the discussion of probability coordination later on). The corresponding probability is often understood to be the physical probability of that very event, hence, a single-case probability. Frequentists must find an alternative understanding. Reichenbach proposes using the frequentist probability relative to the narrowest reference class "for which reliable statistics can be compiled" (1949, p. 374).

The case for explanatory indispensability rests principally on the intuition that the probabilistic explanation of a single outcome requires a single-case probability. The philosophy of scientific explanation, much of it developed by regularity theorists and other metaphysical empiricists, offers a number of alternative ways of thinking about explanation, for example, as a matter of showing that the outcome to be explained was to be expected, or as a matter of subsuming the outcome to be explained under a general pattern of outcomes (both ideas proposed by Carl Gustav Hempel). The fate of frequentism, and more generally of the regularity approach to laws of nature, depends to some extent, then, on the adequacy of these conceptions of explanation.

Why be a frequentist? The view has two principal advantages. First is its light metaphysical touch, shared with the regularity account of laws. Second is the basis it gives for the mathematics of probability: Frequencies, as mathematical objects, conform to almost all the axioms of probability. Only almost all because they violate the axiom of countable additivity, an extension to the countably infinite case of the third axiom described earlier. Countable additivity plays an important role in the derivation of some of probability mathematics' more striking results, but whether it is necessary to provide a foundation for the scientific role of physical probability claims is unclear.

There is more than one way to be a frequentist. A naive actual frequentist holds that there is a probability wherever there is a frequency, so that, in a universe where only three coin tosses have ever occurred, two coming up heads, there is a probability for heads of two-thirds. This view has been widely criticized, though never held. Compare with the naive regularity theory of laws (Armstrong 1983, §2.1).

What might be called ideal actual frequentism is the theory developed by Reichenbach (1949) and von Mises (1957). On this view, probability statements are construed as ideally concerning only infinite classes of events. In practice, however, they may be applied to large finite classes that in some sense come close to having the properties of infinite classes. Thus, Reichenbach distinguishes the logical meaning of a probability statement, which asserts the probabilistic patterning of an infinite class of outcomes, and the finitist meaning that is given to probability claims in physical applications, that is, in the scientific attribution of a physical probability (Reichenbach 1949). On the finitist interpretation, then, a physical probability claim concerns the probabilistic patterning of some actual, finite class of events—albeit a class large

enough to have what Reichenbach calls a practical limiting frequency. (Reichenbach's wariness about logical meaning owes as much, incidentally, to his desire to have his theory of probability conform to the verifiability theory of meaning as to a concern with, say, the validity of probability claims in a finite universe.)

David Lewis (1994), reviving Ramsey's account of laws of nature, proposes that the fundamental laws are nothing but the axioms of the theory that best systematizes, or unifies, the phenomena. A systematization is good to the degree that it is simple, that it makes claims about a large proportion of the phenomena (ideally all the phenomena, of course), and that its claims are accurate. Lewis (1994) extends the definition of accuracy, or as he calls it, fit, to accommodate axioms attributing physical probabilities: A set of phenomena are a good fit to a physical probability statement if the phenomena exemplify the probabilistic patterns appropriate to the probability ascribed. A system of probabilistic axioms will be a good systematization, then, only if the physical probabilities it assigns to the phenomena are reflected, for the most part, in corresponding probabilistic patterns.

In this respect, Lewis's view is a form of frequentism. Although there is not some particular set of outcomes whose probabilistic patterning is necessary and sufficient for the truth of a given probabilistic law statement, it is nevertheless the world's probabilistic patterns, taken as a whole, that provide the basis for all true statements of probabilistic law.

Some writers suggest that a claim such as "The probability of obtaining heads on a toss of this coin is one-half" is equivalent to the claim that, if the coin were tossed infinitely many times, it would yield heads with a limiting frequency of one-half. The truth-makers for physical probability claims, then, are modal facts (except in the case where there actually are an infinite number of tosses). This view is known as hypothetical frequentism.

Though much discussed in the literature, hypothetical frequentism is seldom advocated. Reichenbach (1949) and von Mises (1957) are sometimes labeled hypothetical frequentists, but the textual evidence is thin, perhaps even nonexistent. Colin Howson and Peter Urbach (1993) advocate a hypothetical frequency view. Bas C. van Fraassen's (1980) frequencies are also hypothetical, but because he holds that the literal meaning of theoretical claims is irrelevant to the scientific enterprise, the spirit of his account of probability is, in its empiricism, closer to Reichenbach's ideal actual frequentism.

The weaknesses of frequentism are in large part the weaknesses of the regularity theory of laws. An interesting objection with no parallel in the regularity account is as follows: In the case of reference classes containing countably infinite numbers of events, the value (indeed, the existence) of the limiting frequency will vary depending on how the outcomes are ordered. There appear to be no objective facts, then, about limiting frequencies. Or rather, if there are to be objective facts, there must be some canonical ordering of outcomes, either specified along with the reference class or fixed as a part of the scientific background. How serious an impediment this is to the frequentist is unclear.

THE PROPENSITY THEORY. If frequentism is the regularity theorist's natural interpretation of physical probability claims, then the propensity account is the interpretation for realists about laws, that is, for philosophers who believe that law statements assert the existence of relations of nomic necessity and causal tendencies (Armstrong 1983). For the propensity theorist, probabilities are propensities, and propensities are a certain kind of distinctly probabilistic causal tendency or disposition.

The propensity theorist's home territory is single-case probability, the kind of probability attached to a particular physical process or outcome independently of the specification of a reference class or ordering of outcomes. Because propensities are supposed to be intrinsic properties of token processes, on the propensity view every probability is a single-case probability. Given some particular outcome that one wishes to predict or explain, then, there is an absolute fact of the matter as to the physical probability of the outcome that one may—and presumably, must—use in one's prediction or explanation.

Of course, knowledge of this fact, if it is to be obtained by observing the statistics of repeated experiments, will require the choice of a reference class, the aim being to find a class containing processes that are sufficiently similar that their statistics reveal the nature of each of the underlying propensities in the class. Furthermore, by analogy with the case of deterministic causal tendencies, propensities may owe their existence to probabilistic laws governing classes of processes. Thus, something not unlike the frequentist's reference classes may turn up in both the epistemology and the metaphysics of propensities, but this does not detract from the fact that on the propensity view, there are real, observer-independent single-case probabilities.

To identify probabilities with propensities is revealing because one thinks that one has a good intuitive sense

of the nature of propensities in the deterministic case; one is reasonably clear on what it is to be fragile, aggressive, or paramagnetic. Though the metaphysics of dispositions is still a matter of dispute, it seems that one comes to deterministic propensities, at least at first, by grasping what they are propensities for: for example, breaking, violent behavior, and magnetic attraction. To adopt a propensity theory of probability, then, with the sense of familiarity the word *propensity* brings, is to make an implicit commitment to elucidating what probabilistic propensities are propensities for.

A straightforward answer to this question was given by Karl R. Popper (1959) in one of the earliest modern presentations of the propensity theory: A probabilistic propensity is a disposition to produce probabilistically patterned outcomes. A particular coin's probability for heads of one-half, then, is a disposition to produce a sequence of heads and tails that is disordered in the short term, but in the long term contains heads with a frequency of one-half. (Popper in fact omits the disorder requirement and allows that the sequence may be long and finite or infinite.) On Popper's view, then, a probabilistic propensity differs from a deterministic propensity not in the means of production, but only in what is produced: a probabilistic pattern over a long series of trials, rather than a single discrete episode of, say, shattering or magnetic attraction.

Popperian propensity theory is committed to the claim that, if the probability of a tossed coin's landing heads is one-half (and remains so), then continued tossing of the coin will eventually yield a set of outcomes of which about one-half are heads. But this sits badly with the intuitive conception of the workings of probability: If the probability of heads is one-half, then it is possible, though unlikely, that it will produce all heads for as long as one likes, even forever.

This intuition has an analog in probability mathematics. The law of large numbers prescribes a very high probability that the long-run frequency with which an outcome occurs will match its probability; by the same token, however, there is a nonzero probability that any (finite) long run will fail to produce a probability-matching frequency. There is some physical probability, then, that a probabilistic propensity will fail to produce what, according to the Popperian propensity view, it must produce. If this physical probability is itself a Popperian propensity—and surely it is just another manifestation of the original one-half propensity for heads—then it must produce, by Popper's definition, a matching frequency, which is to say that it must occasionally produce the supposedly impossible series of heads. If it is to be consistent, Popper's definition must be carefully circumscribed. (There is a lesson here for frequentists, too.)

Most propensity theorists accept that probabilistic setups will occasionally fail to produce probability-matching frequencies. Thus, they repudiate Popper's version of the propensity theory. What, then, can they say about the nature of the propensity? Typically, they hold that the probability of, say, heads is a propensity to produce the appropriate probabilistic patterns with a high physical probability (Fetzer 1971, Giere 1973)—thus, such a probabilistic propensity is probabilistic not only in its characteristic effect, which is, as on Popper's definition, a probabilistic pattern, but also in its relation to the effect. (D. H. Mellor [1971] offers an interesting variant on this view.)

Whereas the Popperian definition comes close to inconsistency, this new definition is manifestly circular. Its proponents accept the circularity, so committing themselves to the ineffability of probabilistic propensities.

The ineffability of propensities, it is asserted, is not a problem provided that their values can be inferred; the usual apparatus of statistical inference is tendered for this purpose. Critics of the post-Popperian propensity interpretation naturally fasten on the question of whether it succeeds in saying anything substantive about probability at all—anything, for example, that illuminates the question of why physical probabilities conform to the axioms of the probability calculus or explain the outcomes that they produce. It does seem that modern propensity theorists are not so far from what is sometimes called the semantic interpretation of probability, on which probabilities are considered to be model-theoretic constructs that ought not to be interpreted at all, but simply accepted as formal waypoints between evidence and prediction in probabilistic reasoning (Braithwaite 1953). Compare Carnap's (1950) notion of partial interpretation and Patrick Suppes (1973).

A characteristic doctrine of the propensity theory is that probabilistic propensities, hence probabilities, are metaphysically irreducible: They are in some sense fundamental building blocks of the universe. The corollary to this doctrine is that the physical probabilities science assigns to outcomes that are deterministically produced—including, according to many philosophers, the probabilities of statistical mechanics, evolutionary biology, and so on—are, because they are not irreducible, they are not propensities, and because they are not

propensities, they are irreducible. Ronald N. Giere (1973) writes that they must be given an "as if" interpretation, but propensity theorists offer no account of "as if" probability's scientific role.

On a broader understanding of the nature of a propensity, however, at least some of the physical probabilities assigned by science to the outcomes of deterministic processes might count as probabilistic propensities. As explained in the entry on chaos, certain subclasses of chaotic systems have dynamic properties in virtue of which they tend to generate probabilistic patterns of outcomes (Strevens 2003). These dynamic properties may be understood, then, as endowing the systems with a propensity to produce probabilistic patterns, and the propensity itself may be identified with the physical probabilities that science ascribes to the outcomes.

There is one, not inconsiderable, complication: The systems in question will generate the probabilistic patterns only given appropriate initial conditions. Almost all, but not all, initial conditions will do. This raises two important questions that need to be answered if chaos is to provide a part of the foundation for the metaphysics of physical probability. First, ought the necessary properties of the initial conditions to be considered a part of the propensity? If so, the propensity seems not to be an intrinsic causal property of the process. Second, the initial conditions are, in this context, most naturally described using a probability distribution. Thus, the basis of the probabilistic propensity is a further probabilistic element itself in need of analysis.

THE SUBJECTIVIST THEORY. It is something of a mystery why the mathematics of the probability calculus should be useful both for capturing elements of belief and inductive inference and for describing the processes that give rise to probabilistic patterns, or in other words, why two such different things as epistemic and physical probability should share the same formal structure.

According to the subjectivist theory of physical probability, there is no mystery at all: Physical probabilities are nothing but a certain kind of subjective probability. The intuition that, say, the probability of heads is a quantification of some physical property of the tossed coin is, on the subjectivist approach, an illusion: There are frequencies and mechanical properties out in the world, but physical probabilities exist entirely in the descriptive apparatus of people's theories, or in their minds.

For the principal architect of subjectivism, Bruno de Finetti, the appeal of the theory is not only its neoclassical reunification of epistemic and physical probability but also its empiricism: Subjectivism is most at home in what is now called a Humean world. Of course, frequentism is also a theory of physical probability that the metaphysical empiricist can embrace; the main advantage of subjectivism over frequentism is its provision—if such is truly necessary—of single-case probabilities (de Finetti 1964).

Subjectivism asserts the identity of the subjective probability for heads and the physical probability for heads. But it does not claim that, say, one's subjective probability for the MACHO theory of dark matter is also a physical probability for the theory. Rightly so, because one does not acknowledge the existence of physical probabilities wherever there are subjective probabilities. A plausible subjectivism must have the consequence that one projects only a small subset of one's subjective probabilities onto the world as physical probabilities.

At the heart of the subjectivist theory, then, must be a criterion that picks out just those subjective probabilities that are experienced as physical and that accounts for their particular, peculiar phenomenology. The key notion in the criterion is one of resilience: Unlike most subjective probabilities, which change as more evidence comes in, the subjective probabilities one calls physical have attained a certain kind of stability under the impact of additional information. This stability gives them the appearance of objectivity, hence of reality, hence of physicality, or so the subjectivist story goes. Brian Skyrms (1980) employs this same notion of resilience to give a projectivist account of causal tendencies and lawhood in the deterministic as well as the probabilistic case; subjectivism, then, like frequentism and the propensity theory, can be seen as a part of a larger project embracing all causal and nomological metaphysics.

There is an obvious difficulty with the subjectivist position as elaborated so far: My subjective probability for an outcome such as a coin's landing heads may very well change as the evidence comes in. I may begin by believing that a certain coin is fair, and so that the physical probability of its yielding heads when tossed is one-half. As I continue to toss it, however, I may come to the realization that it is biased, settling eventually on the hypothesis that the physical probability of heads is three-quarters. Throughout the process of experimentation, I project (according to the subjectivist) a physical probability distribution onto the coin, yet throughout the process, because the projected physical probability for heads is changing, increasing from one-half to three-quarters, my subjective probability for heads is also changing. Where is the resilience?

De Finetti's (1964) achievement is to find a kind of resilience, or constancy, in my subjective probabilities even as my subjective probability for heads is changing. This resilience is captured by the property de Finetti calls exchangeability. Consider my subjective probability distribution over, say, the outcomes of the next four tosses of my coin. Every possible sequence of four outcomes will be assigned some subjective probability. The probability assignment—the subjective probability distribution—is said to be exchangeable if any two sequences having the same number of heads and tails are assigned equal probabilities. For example, exchangeability implies that HTHT and HHTT, each having two heads and two tails, are assigned the same probability, but allows this probability to differ from that assigned to, say, HHHT. In an exchangeable distribution, then, the probability assigned to a sequence of heads and tails depends only on the relative frequency with which heads and tails occur in the sequence (in the case of infinite sequences, which de Finetti uses in his mathematical construction, substitute *limiting frequency*).

If my subjective probability distribution over heads and tails is exchangeable, then the order in which the heads and tails come in as I experiment with my coin will not in itself affect my subjective probability for heads. The frequency with which heads and tails come in will, by contrast, most definitely affect my subjective probability. Thus, exchangeability is a kind of partial resilience; it is resilience to information about order, but not frequency.

De Finetti (1964) claims, uncontroversially, that one's subjective probability distributions over future sequences of heads and tails (and the outcomes of other Bernoulli setups) are exchangeable. He goes on to prove a theorem—his celebrated representation theorem—that shows that the following two reasoners will be outwardly indistinguishable: First, a reasoner who has various hypotheses about the physical probability of heads and updates the subjective probabilities for these hypotheses in the usual way as evidence comes in, and second, a reasoner who has no beliefs about physical probabilities, but simply has an exchangeable subjective probability distribution over future sequences of outcomes. The only difference between the two reasoners, then, will be that the first will claim, presumably as a result of introspection, to be learning about the values of physical probabilities in the world.

The subjectivist's sly suggestion is that people are all in fact reasoners of the second kind, falsely believing that they are reasoners of the first kind. Or, in a more revisionist mood the subjectivist may argue that, though they are reasoners of the first kind, they will give up nothing but dubious metaphysical commitments by becoming reasoners of the second kind.

Critics of subjectivism question the aptness of exchangeability as a psychological foundation for probabilistic reasoning. The sole reason that people assign exchangeable subjective probability distributions to certain classes of sequences, according to these writers, is that they believe the sequences to be produced by physical probabilities (Bernoulli distributions, to be exact) and they know that an exchangeable subjective probability distribution is appropriate for outcomes so produced. Note that this argument has both a descriptive and normative dimension: Against a descriptive subjectivist, who holds that beliefs about physical probability play no role in people's probabilistic reasoning, the critic proposes that such beliefs cause them to assign exchangeable distributions. Against a normative subjectivist, who holds that beliefs about physical probability should not play a role in people's probabilistic reasoning, the critic proposes that such beliefs are required to justify their assigning exchangeable distributions.

A different line of criticism targets subjectivism's metaphysics: Why not identify physical probability with whatever produces the probabilistic patterns? Why not say that the probability of heads is a quantification of, at least in part, the physical symmetry of the coin? Such a position has its problems, of course, but they are not obviously insurmountable. More generally, given the rich array of options available for understanding the nature of physical probability, the subjectivist's flight from any attempt to give a metaphysics seems to many, as yet, insufficiently motivated.

PROBABILITY COORDINATION

It is generally accepted that it is rational, in normal circumstances, to set one's subjective probability for an event equal to the physical probability ascribed by science to that event or to that type of event. Returning to the first paragraph of this entry, if the physical probability of a hurricane is high, I should expect—I should assign a high subjective probability to—a hurricane strike. This is the principle of probability coordination.

Because the equation of physical and epistemic probability is made explicit in the classical definition of probability, classicists are probability coordinators par excellence. Leibniz, for example, articulates what appears to be an early formulation of the probability coordination principle when he writes "quod facile est in re, id probabile est in mente" (Hacking, 1975, p. 128); Ian Hacking glosses this as "our judgment of probability 'in the mind'

is proportional to (what we believe to be) the facility or propensity of things" (the parenthesized phrase is not in the Latin; 1975, p. 128). But strictly speaking, of course, classicists cannot conceive of this as a coordination of different kinds of probability, since they allow only one kind of probability.

In the twentieth century, probability coordination was introduced as a topic in its own right by David Miller, who argued, as a part of a Popperian case against inductive inference, that a probability coordination principle would have to be inconsistent. Commentators soon pointed out that there are consistent versions of the principle, and some years later David Lewis wrote what is still the most influential paper about the proper form of a principle of coordination and its role in scientific inference, conjecturing that such a principle "capture[s] all we know about [physical probability]" (1980, p.266).

Modern attempts at a formulation of a probability coordination principle contain two elements not present in Leibniz's maxim. First is the modification interpolated by Hacking: The principle commands that one sets one's subjective probabilities equal not to the corresponding physical probabilities, but to what one believes the values of those probabilities to be, or more generally, to the mean of the different possible values, weighted by one's subjective probability that each value is the correct one. Such a principle might be loosely interpreted as saying that one should do one's best to set one's subjective probabilities equal to the physical probabilities.

Second is a restriction of the range of the principle: When one possesses certain kinds of information, probability coordination is not necessarily rational. Suppose, for example, that I know for some science-fictional reason that the coin I am about to toss will land heads. Then I should set my subjective probability for heads equal to one, not equal to the physical probability of one-half. The information that the coin will land heads is what Lewis (1980) calls inadmissible information; in the presence of inadmissible information, the principle of probability coordination does not apply. Note that what is admissible is relative to the outcome in question; knowing how the coin lands is admissible when I am setting my subjective probability for the outcome of a different toss.

An attempt at a probability coordination principle might, then, have the following form: one's subjective probability for an event e, conditional both on the proposition that the physical probability of e is$p$ and on any admissible information k, should be set equal to p. (One's unconditional subjective probability for e, then, will be the weighted sum of the physical probabilities, as mentioned earlier.) In symbols: If one's background knowledge is admissible, then set

$$C(e|tk) = P_t(e),$$

where $C(\cdot)$ is one's subjective probability distribution, t is the proposition that the correct physical probability distribution for e is $P_t(\cdot)$, and k is any other admissible information.

Note that propositions such as t are normally consequences of two kinds of fact: probabilistic laws of nature and some properties of e in virtue of which it falls under the laws. For example, if e is the event of a particular coin toss's landing heads, then the law might be "All tosses of a fair coin land heads with physical probability one-half" and the additional fact the fairness of the coin in question. In what follows it is assumed that the latter facts are part of the background knowledge, and that t simply asserts some probabilistic law of nature, as suggested by the previous notation.

The most puzzling aspect of the probability coordination principle is the nature of admissibility. Lewis proposes a working definition of admissibility (he says that it is a "sufficient or almost sufficient" condition for admissibility) on which information is admissible either if it is historical—if it concerns only facts about the past up to the point where the principle is invoked—or if it is purely probabilistic, that is, if it is information about physical probabilities themselves.

The definition is problematic for two reasons. One difficulty is explicitly identified by Lewis (1980) and for many years prevented him from advancing the frequency-based theory of physical probability that he wished to give. As noted earlier, when coordinating probabilities for a given outcome, information about the future occurrence or otherwise of that outcome ought to be counted inadmissible. It turns out that frequency-based probabilities provide information of this sort. Lewis, then, has three choices. The first is to revise the working definition of admissibility so as to rule out such information, in which case information about physical probabilities will be inadmissible and the resulting probability coordination principle will be useless. The second is to stay with the working definition of admissibility, allowing the information provided by frequency-based probabilities to count as admissible by fiat. It can be shown, however, that the resulting principle—that is, Lewis's original principle—clearly sets the wrong subjective probabilities in certain circumstances: There are certain complex facts about the future that a frequency-based probability distribution entails cannot

obtain, yet assigns a nonzero probability. If such a probability distribution is known to be the correct one, then the right subjective probability for the facts is zero, but probability coordination results in a nonzero subjective probability. The third option is to abandon probability coordination as such. Lewis takes the third way out, proposing a new kind of probability coordination principle that has the form (using the notation from earlier) $C(e|tk) = P_t(e|t)$. Michael Strevens (1995) points out that both Lewis's new principle and his original principle are consequences of a more general probability coordination principle according to which conditional subjective probabilities should be set equal to conditional physical probabilities. This principle yields Lewis's original principle when information about physical probability distributions is admissible and Lewis's new principle when it is not.

A different problem with Lewis's working definition of admissibility is that it makes no sense of probability coordination in deterministic systems. If one conditionalizes on the exact initial conditions of a coin toss, one ought not to set one's subjective probability for heads to the physical probability of heads, one-half, but either to zero or to one depending on whether those particular initial conditions cause the coin to land heads or tails. If a probability coordination principle is to be applied to the probability of heads, exact information about initial conditions must therefore be ruled inadmissible. Lewis's (1980) working definition of admissibility counts initial conditions, like all historical facts, as admissible.

Lewis (1980) does not regard this as a problem, since he agrees with the propensity theorists that in deterministic systems there could be only ersatz physical probabilities. Even if this is correct as a metaphysical doctrine, however, it remains a matter of fact that one coordinates one's subjective probabilities with such ersatz probabilities all the time, as when one forms expectations about the outcomes of a tossed coin. Whatever one calls it, then, there is a coordination principle for systems such as gambling devices that apparently has the same form as the genuine probability coordination principle (for a reconciliation of Lewis's account of physical probability and probability coordination in deterministic systems, see Loewer 2001).

There is clearly more work to be done elucidating the form of the probability coordination process, and in understanding admissibility in particular. A different project attempts to justify the practice of probability coordination, by giving an a priori argument that subjective probabilities should track physical probabilities, or beliefs about such. Lewis himself says no more than that he can "see dimly" why probability coordination is rational. Howson and Urbach (1993) attempt a full-blown justification. Strevens (1999) argues that Howson and Urbach's argument appeals implicitly to a principle of indifference and goes on to make a case that there is a strong parallel between providing an a priori justification for probability coordination and providing an a priori justification for inductive inference, that is, solving the problem of induction.

A final question about the relation between epistemic and physical probability was adumbrated earlier: Why should the same formal structure be central to one's understanding of two such different things as the production of the probabilistic patterns and the nature of inductive reasoning?

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Confirmation Theory; Decision Theory; Determinism and Indeterminism; Explanation; Statistics, Foundations of.

Bibliography

Armstrong, David Malet. *What Is a Law of Nature?* New York: Cambridge University Press, 1983.

Braithwaite, Richard Bevan. *Scientific Explanation: A Study of the Function of Theory, Probability, and Law in Science*. Cambridge, U.K.: Cambridge University Press, 1953.

Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: Chicago University Press, 1950.

Daston, Lorraine. *Classical Probability in the Enlightenment*. Princeton, NJ: Princeton University Press, 1988.

De Finetti, Bruno. "Foresight: Its Logical Laws, Its Subjective Sources." In *Studies in Subjective Probability*. 2nd ed., edited by Henry E. Kyburg Jr. and Howard E. Smokler. New York: Wiley, 1964.

Fetzer, James H. "Dispositional Probabilities." *Boston Studies in the Philosophy of Science* 8 (1971): 473–482.

Fine, Terrence L. *Theories of Probability: An Examination of Foundations*. New York: Academic Press, 1973.

Giere, Ronald N. "Objective Single-Case Probabilities and the Foundation of Statistics." In *Logic, Methodology, and Philosophy of Science: Proceedings*, edited by Patrick Suppes, et al. Amsterdam, Netherlands: North Holland, 1973.

Gillies, Donald. *Philosophical Theories of Probability*. London: Routledge, 2000.

Hacking, Ian. *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction, and Statistical Inference*. New York: Cambridge University Press, 1975.

Hájek, Alan. "What Conditional Probability Could Not Be." *Synthese* 137 (2003): 273–323.

Howson, Colin, and Peter Urbach. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court, 1993.

Jaynes, Edwin T. *E. T. Jaynes: Papers on Probability, Statistics, and Statistical Physics*, edited by R. D. Rosenkrantz. Dordrecht, Netherlands: D. Reidel, 1983.

Jeffrey, Richard C. *The Logic of Decision*. 2nd ed. Chicago: Chicago University Press, 1983.

Keynes, John Maynard. *A Treatise on Probability*. London: Macmillan, 1921.

Laplace, Pierre Simon de. *A Philosophical Essay on Probabilities*. Translated by F. W. Truscott and F. L. Emory. New York: Wiley, 1902.

Lewis, David. "Humean Supervenience Debugged." *Mind* 103 (1994): 473–490.

Lewis, David. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*. Vol. 2, edited by Rudolf Carnap and Richard C. Jeffrey, 83–132. Berkeley: University of California Press, 1980.

Loewer, Barry. "Determinism and Chance." *Studies in History and Philosophy of Modern Physics* 32 (2001): 609–620.

Mellor, D. H. *The Matter of Chance*. Cambridge, U.K.: Cambridge University Press, 1971.

Popper, Karl R. "The Propensity Interpretation of Probability." *British Journal for the Philosophy of Science* 10 (1959): 25–42.

Ramsey, Frank Plumpton. "Truth and Probability." In *Philosophical Papers*, edited by D. H. Mellor. Cambridge, U.K.: Cambridge University Press, 1931.

Reichenbach, Hans. *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability*. 2nd ed. Translated by Ernest H. Hutten and Maria Reichenbach. Berkeley: University of California Press, 1949.

Skyrms, Brian. *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven, CT: Yale University Press, 1980.

Strevens, Michael. *Bigger than Chaos: Understanding Complexity through Probability*. Cambridge, MA: Harvard University Press, 2003.

Strevens, Michael. "A Closer Look at the 'New' Principle." *British Journal for the Philosophy of Science* 46 (1995): 545–561.

Strevens, Michael. "Inferring Probabilities from Symmetries." *Noûs* 32 (2) (1998): 231–246.

Strevens, Michael. "Objective Probability as a Guide to the World." *Philosophical Studies* 95 (1999): 243–275.

Suppes, Patrick, et al., eds. *Logic, Methodology, and Philosophy of Science: Proceedings*. Amsterdam, Netherlands: North Holland, 1973.

Suppes, Patrick. "New Foundations of Objective Probability: Axioms for Propensities." In *Logic, Methodology, and Philosophy of Science: Proceedings*, edited by Patrick Suppes et al. Amsterdam, Netherlands: North Holland, 1973.

Van Fraassen, Bas C. *Laws and Symmetry*. New York: Oxford University Press, 1989.

Van Fraassen, Bas C. *The Scientific Image*. New York: Oxford University Press, 1980.

Von Mises, Richard. *Probability, Statistics, and Truth*. 2nd ed. New York: Macmillan, 1957.

Michael Strevens (2005)

PROCLUS

(412–485)

Proclus was born in Constantinople into a Lycean family that was still faithful to the old Hellenic religion in a society already dominated by Christianity. The talented young man forsook a career as a lawyer and decided to devote his life entirely to philosophy. After studies in Alexandria, he arrived in 431 in Athens where he joined the Platonic school of Syrianus. After the death of his venerated master, he became the leader of this school and remained in that position for almost fifty years until his death in 485. As we know from his biographer Marinus, his whole life was devoted to teaching and writing. Proclus was also a deeply devout person. In the community of the school he continued to practice with his disciples the rituals of the old Hellenic religion as well as the theurgical rituals of the Chaldeans. For Proclus, Plato was more than a philosopher intent upon the search for the truth; he was also a divinely inspired prophet showing the soul a way of salvation. Reading Plato had become more than just a scholarly exercise—it was a religious activity of paramount importance.

Proclus was convinced that the truth had been revealed by the gods in many different ways, in obscure oracles, myths and symbols. He saw himself as the interpreter whose task it was to explain the hidden significance of those religious traditions in a civilization where they were doomed to disappear. It was his ambition to prove the harmony between Plato and the other sources of divinely inspired wisdom, in particular the Chaldean Oracles and the Orphic poems. In his view, only a philosophical approach could offer the framework and rational arguments needed for this interpretation. For that reason Plato remained for him the ultimate authority in all matters, divine and human. Aristotle, on the contrary, was given only a subsidiary role, as he never developed a proper theology. His significance was restricted to matters of logic and physics.

Proclus wrote commentaries on the dialogues of Plato that were part of the curriculum of the Neoplatonic school. The course started with the reading of the *Alcibiades I* about self-knowledge, which was regarded as an introduction to philosophy, and culminated in the explanation of the two major dialogues of the Platonic corpus: the *Timaeus* about the generation of the physical world and the *Parmenides*, which was thought to offer Plato's doctrine on the first principles. Proclus also wrote a series of interpretative essays on the *Republic*. The commentaries of Proclus are masterpieces of their genre: They not only

offer a systematic interpretation of the text, but also introduce his own philosophical views and provide us with a wealth of information about the discussions within the Platonic tradition. That is particularly the case with the commentary on the *Timaeus*: Proclus defends Plato's explanation of the physical world as superior to that in Aristotle's *Physics* because only Plato discovered the ultimate (i.e., divine) causes of the physical phenomena.

Besides his work on Plato, Proclus composed a remarkable commentary on Euclid's *Elements*, the prologues of which offer a philosophical introduction to the study of the mathematical sciences. The *Hypotyposis*, or *Outline of the Astronomical Hypotheses*, is another indication of his interest in science and cosmology. Proclus also wrote short treatises on diverse subjects, such as the treatise *On the Existence of Evils*, in which he attempts to explain the existence of evil in a world proceeding from an absolutely good principle. If all agents act for some good and yet, evil occurs, it is unintended and uncaused. Evil cannot exist in its own right and no proper cause explains it. Its existence is parasitic (*para-hypostasis*), supervening upon substances and acts. This doctrine, which was adopted by a Christian author writing under the pseudonym of Pseudo-Dionysus in his celebrated treatise *On the Divine Names*, became for centuries the dominant view in philosophical and theological debates on evil.

Besides his commentaries, Proclus owes his reputation to his two great syntheses of Platonic philosophy, the *Elements of Theology* and the *Platonic Theology*. Theology is for Proclus a rational investigation into the first causes and principles of all things. The first philosophers only admitted corporeal entities, such as fire or water, as first causes. Later philosophers recognized souls as principles of life and movement and thus discovered noncorporeal being. Aristotle posited unmoved intellects above the self-moving souls and considered the first intellect to be the ultimate divine cause. Only Plato, however, recognized a cause beyond intellect, beyond being, beyond knowledge and discourse, namely the One, from which all things including matter derive their existence (*Platonic Theology* I,1). Therefore, Plato's theology is for Proclus (and for the entire Neoplatonic school since Plotinus) the accomplishment of all preceding theological speculation, since it reveals the "three principal hypostases": the One, the Intellect, the Soul.

ELEMENTS OF THEOLOGY

In this work, the metaphysical counterpart of Euclid's *Elements*, Proclus demonstrates "in a geometrical manner" the most fundamental theorems of the theological or metaphysical science as he understands it. The work contains 211 propositions, each of them followed by a demonstration. The first part (props. 1–112) examines the fundamental principles that govern the structure of all reality, such as the relation between the One and the many, the cause and the effect, the whole and the parts, transcendence and participation, procession and reversion, continuity and discontinuity, Being, Life and Intellect, limit and limitedness, self-movement and self-constitution, act and potency, eternity and time.

In the second part (props. 113–211) Proclus gives a survey of all degrees of reality, applying to them the general metaphysical principles he had demonstrated before. He discusses successively the gods (or "henads"), the intellects and the souls. The physical realm falls outside the scope of this theological metaphysics. The *Elements of Theology* is without doubt the most original work of Proclus, not so much because of its content (which offers the standard doctrine of the Athenian school), but because of its extraordinary attempt to develop the entire Neoplatonic metaphysics from a set of axioms. It also had a tremendous influence, in particular through the Arabic adaptation that was made in the ninth century in the circle of Al-Kindī. In the middle of the twelfth century this Arabic treatise was translated into Latin. The *Liber de Causis*, as it was named, circulated as the work of Aristotle and thus obtained a great authority in medieval Scholasticism. The systematic character of the *Elements* and its rigorous method make it for the student the best introduction to the complicated thought of Proclus.

ONE AND MULTIPLICITY. The *Elements* begin with the proposition that "every manifold in some way participates in unity." Without some form of unity a multitude would fall apart into an infinity of infinite things. A multitude cannot, however, be itself the unity it participates in. It is not the One, but a unified manifold, having unity as an attribute, and is therefore posterior to the One itself upon which it depends. All things, then, derive their being ultimately from the One from which they proceed.

This One must be identified with the Good, since it is the proper function of the One to hold together all things and maintain them in existence, which is also the function of the Good. For to hold a thing together and make it one is to give it its perfection and well-being, whereas dispersion is the cause of its destruction and evil. Since the Good is what unifies things and the One is what gives them perfection, the One and the Good are names

designating the absolute principle from which and toward which all things eternally proceed and return.

Having demonstrated the necessity of the One as the first principle, Proclus explains how all things in all their specificity proceed from this One. It is impossible to admit that the utmost multiplicity of the material world with its particularized bodies would proceed immediately from the first principle. Plotinus had already argued that from the One comes first the Intellect and from the Intellect proceeds the Soul, which stands itself at the beginning of time, division, and movement. For Proclus, this Plotinian understanding of the procession is unsatisfactory, in particular with regard to the second level, the Intellect, which is identical with true Being and Life. If we respect the "law of continuity," which governs the procession of all things along the "chain of being," we cannot admit that the Intellect (which contains already the Forms of all things) comes into existence immediately after the absolute One. There must be "mean terms" connecting the extremities. From the One comes forth Being, from Being Life, from Life the Intellect. Whereas Being is the ultimate intelligible object (*noēton*), the Intellect, which contains in its thought the paradigmatic Forms, is the properly intellectual level (*noeron*). The intermediate realm of Life is the intelligible-intellectual. In this triad of hypostases, the superior level has the most comprehensive and farthest reaching causality: for all things participate in being, but not all are living or capable of thinking. The causality of the One reaches even further than that of Being, since matter, the indeterminate substrate of the physical realm, does depend on the One, though it does not really "exist."

THE TRIADIC DYNAMIC STRUCTURE OF REALITY. Many propositions in the *Elements* concern causality (hence the Latin-Arabic adaptation is appropriately called *De causis*): they are not about the physical causes, which are for Proclus only auxiliary causes, but about the "true causes," which always transcend their effect. Whatever produces something must be superior to the effect, which owes its existence to it. If this effect has itself the power to produce, it will produce again something inferior to it, until the procession comes down to what is altogether unproductive, that is, matter. Although the effect is inferior to its cause, it is also somehow similar to that which has produced it. The effect is in a secondary manner what its cause is primarily. Insofar as the effect is similar to its cause and shares its character, it is said to "remain" in its cause without yet having its proper existence. On the superior level, it exits "causally" or "potentially" (if "potency" is understood as a productive power).

A being only acquires its proper existence (*hyparxis*) when it proceeds from its cause and becomes distinguished from it. Through the procession it becomes somehow dissimilar to the cause. Yet the procession from the cause cannot go on infinitely: the effect must also revert upon the cause from which it proceeds. Through this "return" (*epistrophē*) the effect strives to be connected again with its cause and becomes similar to it. If things have their being through procession, they attain their well-being or perfection through reversion. For the cause of their well-being can only come from the origin of their being. The final cause is thus identical with the efficient, since all things desire as ultimate end that which is the principle of their procession. As Proclus formulates it: "All that is produced by a cause both remains in it and proceeds from it" (*Elements of Theology*, § 30). "All that proceeds from something reverts upon that from which it proceeds" (§ 31). Therefore, "all that proceeds from a principle and reverts upon it has a cyclical activity"(§ 33). All beings remain in their causes, proceed from them, and return to them, in an eternal circularity, since the end is identical with the origin. Proclus finds this triadic dynamic structure on all levels of reality.

PARTICIPATION AND NONPARTICIPATION. When attempting to understand the relation between the Forms and the many things that are similar to them, Plato introduced the metaphor of "participation." Participation, however, raises as many problems as it solves, as Plato shows in the aporetic discussion of the *Parmenides* (which offered ammunition for Aristotle's subsequent criticism). The term seems to suggest that the many things sharing in the same Form take "parts" of it. How can one reconcile the transcendence of the Forms with their presence in the many things? If participation is real, the Forms must be immanent in the things sharing them and hence will be divided. But how, then, can the transcendence of the Forms be preserved? If, on the other hand, we stress the unity and the indivisibility of the Forms, we end up making participation impossible.

Proclus's solution to this problem is the distinction between the participated and unparticipated mode of a *hypostasis*. What is participated in by the particular things cannot be the ideal Form itself, but must be a form that comes forth from it and is present in them. These immanent forms are somehow comparable to the Aristotelian Forms in matter. However, whereas Aristotle rejects the transcendent Forms as an unnecessary duplication of reality, Proclus argues that the unparticipated Forms are necessary to guarantee the universal character of the forms in matter. The participated form belongs entirely

to the particular thing sharing it. Since what inheres in one thing cannot be present in another, there is no explanation of the fact that the many things, though obtaining a proper form, have this form in common. By postulating an unparticipated Form, which exists prior to all participated forms proceeding from it, the Platonists can explain how the *eidos* is common to all that can share in it and nevertheless the same in all. As is said in proposition 23: "all that is unparticipated brings forth from itself the participated; and all the participated hypostases extend back to the unparticipated."

The distinction between the participated and the unparticipated not only applies to the Forms, but to all levels of reality: Soul, Intellect, and even the One. Within each realm a distinction must be made between the first unparticipated term (the "monad") and the "series" or multiplicity of beings of a similar nature coordinated with it. Thus, besides the many souls that are participated in on various levels by different bodies—the particular souls by which each human being exists as a particular animal, the souls of demons, the planetary divine souls—we must postulate the existence of the unparticipated Soul, from which a multiplicity of souls proceeds according to diverse modes of participation.

Similarly, besides the many particular intellects participated in by different divine and human souls, there must also exist an absolute unparticipated Intellect, which comprises the totality of all Forms. The many intellects proceed from this absolute Intellect and form together with it a coordinate series of a similar intellectual nature. Following the same line of reasoning, we must also posit after the One, which is absolutely transcendent and can in no way be participated by the inferior levels, a manifold series of "ones," "units," or "henads" consequent upon the primal One, which are participated in by the different classes of being. Those henads are not the modalities of unity acquired by beings, but self-subsisting units which remain transcendent above the beings that depend upon them. Though they are in themselves beyond being and beyond knowledge, as is the primal One, in which they remain co-united, their distinctive properties can be inferred indirectly from the different classes of beings dependent upon them. "For differences within an order of participants are determined by the distinctive properties of the principles participated in" (*Elements of Theology* § 123).

In view of the different classes of beings depending upon them, we can distinguish, for example, intelligible, intellectual, hypercosmic, or encosmic henads. Yet, insofar as they are all self-subsisting units, they remain unified in the One itself. If the One stands for the first divine cause, the different henads constitute the different classes of the gods. "For every god except the One is participable" (§ 116). With this doctrine of the henads, Proclus can defend—against Christian monotheism—both the unity and multiplicity of the divine. In his view, it is the main task of a Platonic philosopher to explain in a rational system the procession and the distinctive properties of all the classes of the gods we know through the diverse religious traditions. That is the purpose of Proclus's last magnum opus.

THE PLATONIC THEOLOGY

Proclus distinguishes four types of theological discourse. Divinely inspired poets use dramatic stories (talking about sexual relations, births, fights, cuttings of organs) to symbolically indicate the processions of the divine principles and their mutual relations. This mythological discourse is characteristic of the ancient Hellenic theology, as known through the Orphic poems and the works of Homer and Hesiod. In oracular discourse (in particular the *Chaldean Oracles*) prophets reveal the names and properties of the gods without resorting to the dramatic scenery of mythology. The Pythagoreans resort to mathematical analogies and similitudes (numbers, circles, spheres) to disclose the divine orders. Finally, there is scientific or dialectical theology, which investigates the divine classes and their properties using strictly rational arguments and an abstract philosophical vocabulary: one and many, being, whole and part, identity and otherness, similarity and dissimilarity.

This scientific theology has been brought to perfection by Plato in his dialogue *Parmenides*. In Proclus's interpretation, this dialogue displays the fundamental axioms and basic concepts needed for the development of a scientific theology. In the second part of this dialogue, Parmenides examines in a dialectical exercise the hypothesis of Unity, considering the consequences following from the position of the One and from its denial, both for the One and for what is other than the One. If we start from the hypothesis of the One, only negative conclusions seem to follow: the One has no parts and is not a whole; it is not in something nor in itself; it is neither similar nor dissimilar. One cannot even say that it "is" or "is one." In short, no names, no discourse, no knowledge of it is possible. Parmenides therefore has to restate his original hypothesis, now emphasizing the existence of the One. All attributes that were denied in the first hypothesis can be predicated of this One-that-is.

The interpretation of the different hypotheses of the *Parmenides* (of which we mention only the first two) led to a lively debate in the Neoplatonic school, as we know from Proclus's commentary. Proclus defends a theological interpretation of this dialectical discussion about the One and the Many. If the "One" stands for the first principle, the successive hypotheses of the *Parmenides* must refer to the different principles of the whole of reality. The One of the first hypothesis, of which no discourse is possible, is the absolute, unparticipated One or primal god. In the second hypothesis, Parmenides deduces, through the subsequent conclusions following from the position of the One-that-is, the different modes of unity ("henads") that are participated in by the different degrees of being. Whereas the first hypothesis leads to a negative theology, the deductions from the second hypothesis give the articulations of a positive theology. "In this dialogue proceed all the divine classes in good order from the first cause and demonstrate their mutual connection" (*Platonic Theology*, I, ch. 7).

When interpreted in this way, the *Parmenides* provides a framework in which the other discourses about the gods can be integrated and decoded: the mythological stories about Zeus and Kronos from the Hellenic and Orphic traditions, the strange divine names revealed in the *Chaldean Oracles*, the mathematical *theologumena* of the Pythagoreans, the various scattered remarks about the gods in the other dialogues of Plato. In the Renaissance, Marsilio Ficino adapted the model of Proclus's theology in an original way to integrate the revealed truth of Christianity.

CONCLUSION

It is difficult to evaluate the originality of a thinker who, in most of his works, claims to be nothing but a faithful follower of his master Syrianus. It is Proclus, however, who put his mark on the development of the later tradition of Neoplatonism in Byzantine, Arabic, and Latin medieval thought. His huge influence—much greater than that of Plotinus—is to be explained mainly by two important indirect channels: the Christian reception of his theology by Pseudo-Dionysus and the Arabic adaptation of the *Elements* in the *Liber de Causis*. And yet it is no historical accident that Proclus gained this fame. The *diadochos* (or successor) of Plato, as he was named, has been the authoritative commentator of Plato and the great systematizer of Neoplatonic metaphysics.

See also Liber de Causis; Neoplatonism; Plato; Plotinus; Pseudo-Dionysus.

Bibliography

THE LIFE OF PROCLUS

Marinus, Vie de Proclus. Edited and translated by Henri Saffrey and Alain Segonds. Paris: Les Belles Lettres, 2003.

ANNOTATED EDITIONS AND TRANSLATIONS

The Elements of Theology. 2nd ed. Edited and translated by E. R. Dodds. Oxford: Clarendon Press, 1992.

A Commentary on the First Book of Euclid's Elements. Translated by Glenn Morrow. Princeton, NJ: Princeton University Press, 1970.

Proclus' Commentary on Plato's Parmenides. Translated by Glenn Morrow and John M. Dillon. Princeton, NJ: Princeton University Press, 1987.

Commentaire sur le Timée. 5 vols. Translated by A. J. Festugière. Paris : Vrin, 1966–1968.

Commentaire sur la République. 3 vols. Translated by A. J. Festugière. Paris: Vrin, 1970.

Théologie platonicienne. 6 vols. Edited and translated and edited by Henri Saffrey and Leendert Westerink. Paris: Les Belles Lettres, 1968–1997.

Commentaire sur le Premier Alcibiade. 2 vols. Edited and translated by Alain Segonds. Paris: Les Belles Lettres, 1986. English translation by W. O'Neill (2nd ed., The Hague: Nijhoff, 1971).

Commentaire sur le Parménide. Vol. 1. Edited and translated by Alain Segonds and Carlos Steel. Paris: Les Belles Lettres, 2005.

On the Existence of Evils. Translated by Jan Opsomer and Carlos Steel. London: Duckworth, 2003.

Proclus' Hymns. Translated and annotated by R. M. Van den Berg. Leiden, Boston, Köln: Brill, 2001.

BIBLIOGRAPHIES

Steel, Carlos, ed. *Proclus: Fifteen Years of Research: A Critical Bibliography (1990–2004)*. *Lustrum* 44 (2005).

For a complete survey of editions and translations, see *Plato Transformed*. Leuven: De Wulf-Mansion Centre, Katholieke Universiteit Leuven. http://www.hiw.kuleuven.ac.be/dwmc/plato/proclus/index.htm.

SECONDARY LITERATURE

Siorvanes, Lucas. *Proclus. Neo-Platonic Philosophy and Science*. New Haven and London: Yale University Press, 1996.

Segonds, Alain, and Carlos Steel, eds. *Proclus et la théologie platonicienne*. Leuven-Paris: Leuven University Press—Les Belles Lettres, 2000.

Carlos Steel (2005)

PRODICUS OF CEOS

Prodicus of Ceos, the Greek Sophist, was probably born before 460 BCE and was still alive at the time of the death of Socrates in 399 BCE. He traveled widely as an ambassador for Ceos and also earned a great deal of money lecturing in various Greek cities, especially in Athens. His writings are known to have dealt with physical doctrines,

with religious and moral themes, and above all with distinctions between the meanings of words usually treated as synonyms. Socrates attended a lecture by him on the last of these topics and regularly claimed to be a pupil of Prodicus in the art of synonymy (*Protagoras* 341A, *Meno* 96D).

In physics he appears to have treated the four elements of Empedocles as divine, and no doubt they formed the basis of the cosmology of Prodicus, to which Aristophanes refers in the *Birds* (1.692), although the fanciful cosmology that follows is probably not based on that of Prodicus. Prodicus further held that those natural objects and powers that are useful to human life were made the objects of cult and treated as gods by men. Inevitably, he was later classed as an atheist, but it is more likely that he offered an account of the origin of the gods that was not intended to deny their existence.

In a work titled the *Horae* (Hours) he included the since famous story "Heracles Where the Road Divides," of which we have a fairly full summary in Book II of Xenophon's *Memorabilia*. Vice and Virtue appear to Heracles personified as women and invite him to choose between them. Each describes what she has to offer, and Heracles chooses the arduous tasks of Virtue rather than the pleasures of Vice.

Of greater philosophic interest is the ethical relativism attributed to Prodicus in the pseudo-Platonic dialogue the *Eryxias*. There he is apparently quoted as arguing that what is good for one man is not good for another man, so that we cannot speak of anything as good *simpliciter*. On the other hand, the goodness of a thing does not depend on the goodness of the user (although some scholars have interpreted him this way). Rather, the value of a thing inheres in the thing itself in such a way that it will be good in relation to one person and not good in relation to another, according to the person and the way in which it is used.

The discussion of synonyms and the right use of words clearly involved fine distinctions of meaning between words. Many examples quoted are ethical, and a term of narrower application is commonly distinguished from one of wider application that includes in its range of meaning the meaning of the first term. The value of such distinctions is clear in rhetorical argument. But Prodicus was also eager to reject the kind of view found in Democritus, according to which there can be different names for the same thing since names are attached to things by convention only. Prodicus maintained, it would seem, that no two words have the same meaning, and in this he at least prepared the way for the search for precisely stated meanings that later fascinated Socrates and Plato.

See also Ethical Relativism; Sophists.

Bibliography

TEXTS AND TRANSLATIONS

Diels, H., and W. Kranz, eds. *Fragmente der Vorsokratiker*. 10th ed. Vol. II. Berline, 1961.

Sprague, R. K., ed. *The Older Sophists: A Complete Translation by Several Hands of the Fragments in die Fragmente der Vorsokratiker, ed. by Diels-Kranz*. Columbia, SC: University of South Carolina Press, 1972.

STUDIES

Guthrie, W. K. C. *The Sophists*. Cambridge, U.K.: Cambridge University Press, 1971.

Kerferd, G. B. *The Sophistic Movement*. Cambridge, U.K.: Cambridge University Press, 1981.

Kerferd, G. B. "The 'Relativism' of Prodicus." *Bulletin of the John Rylands Library* 37 (1954): 249–256.

G. B. Kerferd (1967)
Bibliography updated by Paul Woodruff (2005)

PROGRESS, THE IDEA OF

In broad terms a popular belief in "progress" means the rejection of an attitude that has characterized most human communities throughout history. Normally, people have believed that the future would repeat the past. When they have expected that human life was going to change, they have usually supposed that this change was going to take place suddenly and radically, by supernatural intervention. And if they have permitted themselves to hope for the improvement of the human condition, the hope has commonly been directed toward salvation from the world rather than reform of the world. By and large, historical change, when people have been aware of it at all, has been viewed as a sign of mortality and the proof of a lapse from ideal standards. Indeed, in many societies there has been a popular conviction that humankind's condition has changed in the course of history but for the worse. Characteristically, when people have believed in a golden age, they have put that age in the past rather than the future.

In contrast, in modern Western societies change and innovation have a different place in the popular imagination. Not everyone assumes that all change is necessarily for the better, but it is widely assumed, even by conservatives, that only a society that has a general capacity to change is capable of surviving. And despite wars and depressions a large proportion of the members of Western societies have tended to expect that, short of a cataclysm, their children would live happier and better lives

than they. They have supposed that this improvement would be cumulative and continuing and that although temporary setbacks, accidents, and disasters might take place, human knowledge, power, and happiness would increase over the long run.

The emergence of this idea is the product of a variety of circumstances, such as the accumulation of an economic surplus, the increase of social mobility, and the occurrence of major inventions that have dramatically increased human power over nature. Over and above these, however, the idea of progress is peculiarly a response to the emergence of the unique social institution of organized scientific inquiry.

HISTORY OF THE IDEA

Seeds of the faith in progress can be found in the works of the two great spokesmen for the new science, Francis Bacon and René Descartes. The fundamental elements of the idea itself were developed in the course of the so-called quarrel of the ancients and the moderns, which occupied writers and critics in the last part of the seventeenth century. At the heart of this controversy was a dispute over the authority that should be attributed to the opinions and examples left by the ancient writers. Was it the task of scholars to stand as sentinels at the gate, guarding against innovation and protecting established styles and beliefs? The controversy implicitly raised not only literary questions but the larger question of what attitude toward the past should govern the intellectual life.

In developing their position, the moderns argued that the partisans of the ancients were misled by a false analogy. They looked upon the ancients as their forefathers and therefore thought of the ancients as older and, in consequence, wiser than themselves. But just as the individual grows older and presumably wiser as time goes by, so does humanity. The so-called ancients were really the young men of humanity, and those alive today were the true ancients. They stood on the shoulders of their predecessors and could see farther; their wisdom and authority was greater than the wisdom and authority of their predecessors. This argument was developed with particular force by Bernard de Fontenelle in his *Digression sur les anciens et les modernes* (published in 1683).

The analogy between the history of humankind and the life of an individual had already been developed, however, by a number of writers. Blaise Pascal, for example, used it in drawing a belief in intellectual progress from an examination of the nature of scientific inquiry. In 1647, Pascal had published a study, *Nouvelles expériences touchant le vide*, which encountered immediate objections from many scientists and philosophers, including Descartes, on the ground that it denied the time-honored truth that nature abhorred a vacuum. Pascal replied to one of his critics, Father Noel, that an appeal to inherited authority had no force where the study of physics was concerned. And in a longer essay, *Fragment d'un traité du vide*, he went on to give general reasons for moderating the respect for received authority. "The experiments which give us an understanding of nature multiply continually," he pointed out, "from whence it follows … that not only each man advances in the sciences day by day, but that all men together make continual progress in them as the universe grows older." Pascal believed, however, that such progress took place only where the experimental methods of the sciences were relevant. In theology received authority set the final limits to inquiry, for there the object was not to add to the knowledge provided by ancient authority but only to understand as fully as possible what that authority revealed.

During the eighteenth century, however, and particularly in France an increasing number of intellectuals came to believe that the methods and spirit of science should be applied to all fields. In consequence, the idea of progress came to include a concept of social and moral progress. The cumulative improvement in human knowledge and power that had been brought about in the physical sciences could also be brought about in the organization of human society and the character of human conduct, it was asserted, if only the barriers that existed against the employment of rational methods in morals, religion, and politics could be removed. The Encyclopedists, chief among whom were Denis Diderot and Jean d'Alembert, led in the dissemination of this point of view. The most complete and moving expression of this faith in progress was the Marquis de Condorcet's *Esquisse d'un tableau historique des progrès de l'esprit humain*, written in 1793.

In the nineteenth century a new kind of historicist philosophy emerged that rejected the eighteenth-century conception of reason and the sharp dichotomy between the present and the past that had been made by believers in progress. This philosophy, best represented by G. W. F. Hegel, substituted the view that history followed its own inherent course of development and that this course of development embodied rational principles higher than those of merely human reason. Since this form of historicist philosophy identifies all conceivable changes as elements in an unfolding rational purpose, it deprives the idea of progress of definite meaning.

The more definite and combative eighteenth-century conception of progress, however, also continued to be a central theme in the thought of the nineteenth century. In one form or another, major figures of the century, such as Karl Marx, Auguste Comte, and John Stuart Mill, all propounded the doctrine. Although Marx, Comte, and Mill were influenced, each in his own way, by historicist ideas, each retained the characteristic eighteenth-century emphasis on the struggle between reason and superstition, on the movement of humankind away from theological and metaphysical modes of thought to positive or empirical habits of mind, and on the importance of extending the standards and methods of the sciences to all domains.

In the twentieth century the idea of progress continued to have adherents, particularly among American pragmatists, Marxists, and logical empiricists. For obvious historical reasons, however, advocates of the belief in progress have become steadily more modest in their claims since World War I, and since the turn of the twentieth century the idea of progress has been seized on by an increasing number of philosophers, theologians, and social critics as the prime fallacy of the tradition of liberalism and rationalism.

ANALYSIS OF THE IDEA

In tracing the history of the idea of progress, it is useful to distinguish between two motifs. Generally speaking, the belief in progress has been supported by an appeal to the progress of the sciences. In many cases, however, this appeal has consisted in showing that the sciences—usually some particular science—had uncovered fundamental truths that had been previously unknown and that progress would now take place if only these truths were accepted as guides to practice. Thus, progress has been said to be guaranteed if people lived by the fundamental principles disclosed by the science of economics, if they accepted the laws of historical development revealed by a scientific approach to history, or if they extended to the government of human society the Darwinian doctrine of evolution by natural selection. Progress has also been thought to be guaranteed if people could only come to recognize certain rational moral principles, such as universal natural rights. Such universal principles, though antecedent to any particular science, were nevertheless closely identified with science, for it was assumed that their validity would be apparent to anyone who could disencumber himself from the superstitions and prejudices of the past and that this process of disengagement was immensely accelerated by the advent of science. This conception of the nature and conditions of progress lends itself to Utopian and Messianic interpretations of progress when understood as an ideal but to the reduction of the idea, in G. M. Young's phrase, "from an aspiration to a schedule" when associated with rigid, a priori approaches to the problem of improving the human condition.

A second motif in the theory of progress, however, has associated progress not with any particular discoveries of science or reason but with the unique, self-corrective methods of science. From this point of view the essential conditions for progress are the rejection of absolutes and fidelity to the principles of free, fallibilistic, experimental inquiry in all domains of thought and action. Even if we assume that it is valid to assert that the methods of science are universally applicable, this approach obviously imposes practical conditions for progress that are immensely difficult and perhaps impossible to realize. Accordingly, those who adopt this approach to the idea of progress can be taken to be saying only that there is a possibility of progress or, at best, a slow and uneven historical tendency that is characteristic only of societies possessing an appropriate ethic and social order and whose continuation is by no means ensured. In the past many proponents of the idea of progress undoubtedly underestimated the difficulties of domesticating within society at large the attitudes and habits of mind exemplified in scientific investigation. Nevertheless, insofar as their concept of progress depended simply on an appeal to the character of scientific procedure, they cannot be said merely to have offered a secularized version of older religious beliefs in a heavenly city, and criticisms of them for having done so, which are standard in much of the literature related to the history of the idea of progress, are a source of considerable confusion.

To be sure, the theories of progress that were developed in the eighteenth and nineteenth centuries are often based on a combination of these two motifs. In Condorcet's thought, for example, there can be found Utopian as well as realistic formulations of the idea of progress. Nevertheless, it is a mistake, on the whole, to associate the idea—particularly as it arose in eighteenth-century France—with the naive hope that human beings and human society could be made perfect. If we study the specific predictions that Condorcet made with regard to the future of humanity, for example, we find that he pointed ahead, with extraordinary prescience, to what are now such commonplace facts as the lengthening of life expectancy, social insurance, and the guarantee of equal

legal rights to all citizens. Although none of these has brought the happiness and general reasonableness that Condorcet assumed they would, it was historical realism on his part, not juvenile innocence, to make such predictions. An inability to imagine the wretchedness of the past, not a cold, unillusioned understanding of the present, lies behind the failure to appreciate why reasonable men in the eighteenth and nineteenth centuries should have been rhapsodic about the possibility of changes in the human condition that, in the light of contemporary heightened expectations, may tend to appear fairly modest.

SCIENTIFIC PROGRESS. What can be said with regard to the validity of the idea of progress? We must first ask what meaning can be assigned to the notion of scientific progress.

One frequent argument against the validity of the belief in scientific progress is that it contains a self-contradiction. The belief that there is scientific progress is usually attached to the argument that science is continually self-corrective. But if science never does anything but correct itself, is there any sense in speaking of scientific progress? Does not the concept of progress presuppose a fixed end or standard, and does not science, at any rate as interpreted by those who emphasize its fallibilism, deny that there can be fixed ends or standards? *Progress*, in short, appears to be a term without meaning, according to this view, unless it can be attached to metaphysical standards, such as absolute truth, whose status is antecedent to science.

This view fails, however, once it is recognized that progress can also refer to the solution of particular problems, not only to the movement toward a general and abstract goal. For example, meaning can obviously be assigned to the statement that science has made progress in determining the causes of malaria or in describing the characteristics of the other side of the moon. Such statements mean that there are now answers to questions to which there were no answers before and that these answers are in accord with the procedures of inquiry in force among competent scientific investigators. Once scientific progress is defined in terms of the solutions to particular problems, sense can also be given to the notion of cumulative scientific progress, for the general scientific capacity to solve problems has also tended to grow.

Some doubt has been thrown on these conclusions, however, by recent philosophers of science. Karl Popper, for example, argued that scientific theories and hypotheses are never genuinely confirmed but at best succeed only in resisting successive efforts to falsify them. Since the capacity of a scientific conclusion to survive a series of such efforts does not prove that it will always be able to do so, it would seem to make no sense to speak of successful or true solutions of scientific problems. Popper's view, however, seems to involve an unnecessarily paradoxical way of stating the truism that all scientific conclusions are subject to correction in the future. The survival of a scientific conclusion despite successive efforts to overthrow it adds to the degree of reliability that may reasonably be ascribed to it. It is just as possible to describe the critical position of scientists toward accepted conclusions as efforts to extend the range and reliability of these conclusions as it is to describe it as the expression of a compulsion to destroy what has been inherited. The accumulation of increasingly well-tested and continuously powerful ideas by the sciences is an obvious fact of their history, but as seen by Popper, it seems almost an accidental by-product.

Doubt has also been thrown on the belief in scientific progress by the view that the history of science is the record of revolutions in scientific theory so radical in character that it is impossible to establish the continuity between the ideas of one generation and the ideas of a later one. If this were true, it would be impossible a fortiori to establish a concept of progress, since such a concept presupposes a measure of continuity in the sequence of events under examination. Underlying this view is the thesis that the confirmation by experiment of particular hypotheses always entails the use of a specific theoretical framework. When this theoretical framework changes, observations are simply run through a different set of conceptual categories. Accordingly, it makes little sense, it is argued, to say that the sciences have improved or extended their knowledge, for all that has happened is that one body of beliefs has been substituted for another. This point of view raises epistemological and methodological questions of great complexity, and there is no room to discuss them sufficiently here. It appears to leave out of account, however, the consideration that, for example, fundamental principles of Newtonian physics can, with appropriate modifications, be absorbed into modern physical theories. It also appears to underestimate the implications of the fact that these principles, without substantial modification, continue to provide reliable instruments for the explanation and prediction of events in large sectors of macrophysics.

SOCIAL AND MORAL PROGRESS. Assuming that both meaning and truth can be assigned to the idea of progress in science, what is the status of the belief in social and

moral progress? Obviously, the answer to this question depends in part on the standards employed as the touchstones of progress. However, some of the difficulties involved in stating and defending such standards can be circumvented if in this sphere we also define progress in terms of the successful solution of specific problems. Thus, there has been striking progress in the control of disease, in methods of farming, in material productivity, in the reduction of backbreaking labor, in the techniques of rapid mass communication, in the spread of literacy, and probably in the reduction of the amount of violence in everyday life.

Of course, it is theoretically possible to hold a moral code from whose standpoint one or more of these historical trends would be regarded as retrogressive rather than progressive. In fact, however, even though members of different contemporary cultures (and members of the same culture) hold widely disparate moral outlooks, there are few informed and disinterested observers, whatever their moral outlooks, who regard any of these trends, considered in themselves, as movements in the wrong direction. And most would also look upon many other historical trends that have characterized the modern world—for example, the development of more humane attitudes in penology, the abolition of slavery and serfdom, the spread of the doctrine of basic human rights—in a similarly favorable light. To this extent it is possible to speak with a measure of precision and truth of social and moral progress.

But this answer, of course, goes only part of the way. On at least two scores it is incomplete. First, it is reasonable to ask whether the gains that have been mentioned have not been bought at a cost that more than cancels them out; second, it is possible to ask how we are to vindicate the moral principles in terms of which we assess these gains as gains.

The cost of progress. It is not possible, of course, to give a wholly unequivocal answer to the question of the cost of progress. The notion that large-scale historical trends can be neatly categorized as good or bad belongs to eschatology, not to mature historical analysis. If the reduction of civil violence, considered in itself, is a progressive trend, contemporary mass warfare and genocide must be considered retrogressive; if rapid mass communication is a benefit to humankind, the use of the facilities of communication for totalitarian thought control is a calamity. Moreover, the successful solution of many problems often creates new and more difficult ones. The control of disease, for example, has created a serious threat of overpopulation. And by what calculus can one measure the gains brought about, for example, by industrial innovations against the losses brought about by mass warfare or cyclical unemployment? A moral accounting system for judging even much simpler matters than these does not exist.

Nevertheless, if the span of time we measure is sufficiently long, it remains true that on the whole the physical lot of most ordinary people has considerably improved in modern societies and that this has largely been due to the application of rational techniques to the economy. The cost has been grievous, and many of the sacrifices this progress has entailed could probably have been avoided if people had employed reasonable forethought and had shown reasonable respect for the equities. Admittedly, too, it is difficult to say whether this physical progress has made individuals "happier"; indeed, it is doubly difficult to say this, for "happiness" is in part a function of what people expect, and physical progress has meant an enormous expansion of their expectations. Nevertheless, it is doubtful that most of those who put forward the view that the costs of material progress outweigh the benefits would willingly exchange places with any but the most privileged members of past societies if they actually had the chance.

Nor must we confine ourselves to a belief in purely physical or material progress. The role of fantasy, ignorance, superstition, and fanaticism in determining the world's affairs continues to be enormous. It is doubtful, however, whether so many members of human societies, from housewives to statesmen, have ever before thought it reasonable to make decisions on the basis of carefully acquired and sifted information, and never before have societies possessed as much knowledge about themselves and their workings as they do now, shaky and scattered though that knowledge is. Only if one thinks it morally dangerous to seek reliable information before making decisions or thinks it mistaken to try to employ rational methods in the study of human affairs can he declare such long-range social trends to be anything but progressive. Indeed, the very reason that the members of an educated modern society bear a particularly heavy burden of responsibility for the emergence of doctrines such as Nazism is that they have opportunities to be informed and judicious which members of other societies did not have. In sum, although it is not possible to say in wholesale terms that there has been moral progress, it is possible to assert that the context of human behavior has changed and that the collective capacity to achieve human purposes, whether good or ill, has enormously increased. The expectations that it is reasonable to

impose on modern social arrangements are therefore justifiably higher than those that may have been reasonable in the past. In this modified but important sense it is fair to speak of moral progress.

Justification of moral standards. All the preceding reflections, however, obviously presuppose the validity of a secular, liberal, and rationalistic moral code. In the end, as must be obvious, objections to the idea of progress usually turn on fundamental differences in values. Whether the validity of one fundamental moral outlook as against another can be demonstratively proved is an issue that falls beyond the scope of the present article. If we assume, however, that we cannot resolve these differences in a way that will satisfy traditional standards of demonstrative certainty, there is no so-called ultimate answer to the question of whether modern society has been the scene of genuine progress.

It is possible, however, to show that a relativistic moral philosophy is perfectly compatible with a belief in progress, for it is not true that a relativistic philosophy cannot make any meaningful statements about progress because it has to grant that there are different moral standards and that all are equally valid. First, even if there is no way of proving the absolute validity of a moral outlook, there is still a way of intelligently and objectively assessing its credentials. The moral ideals that underlie the indictment of modern civilization for its excessive individualism and egalitarianism made by T. S. Eliot, for example, would require, if they were to be seriously employed as positive programs for action, the dismantling of large segments of industrial society. Since we may assume that those who put forward such criticisms would wish medical science to continue its work, for example, and would accept a world population at something like its present size, we must conclude that their announced preferences are both unrealistic and incoherent because they are incompatible with other values that they also hold. An examination of available resources, of the costs of maintaining or instituting alternative systems of values, and of the utility of these systems as guides to the resolution of definite historical problems provides a way of choosing among competing moral outlooks and makes the choice something more than a matter of personal whim or social convention.

Second, although the philosophical relativist may believe that apart from the specification of definite problems in determinate historical contexts, there is no way of showing that a moral code is valid, this does not mean that he does not himself hold any moral standards or that he is any less attached to them than an absolutist would be. A twenty-first-century American looking at slavery in ancient Rome, for example, will regard it as a change for the better that slavery is now illegal in Western society, and he will do so whether or not he is a relativist. And to say that he might feel different if he were a Roman is irrelevant, for he is a twenty-first-century American, not a Roman, and it would be a different person with a different identity, not he, who felt different in the hypothetical circumstances. Similarly, if the standards of people in the future change, they may well disagree with us in regard to what has been progressive in history. But if these future judgments reverse present judgments, that does not bind a relativist living here and now to accept them. Nothing in his position requires him to say that progress is any historical trend that comes to be thought desirable.

Progress as a moral standard. As a final consideration, it is important to recognize that the idea of progress in its most important aspect is itself a regulative moral ideal, not simply a belief about history. It represents a directing principle of intellectual and social action, instructing human beings to regard all social arrangements with a critical eye and to reject any claim that any human problem has been finally solved or must be left finally unsolved. To the extent that this idea of progress is embodied in moral codes and social systems, these codes and systems will contain deliberate provision for self-reform. The idea of progress thus represents the social application of the principle that inquiry should be kept open and that no bounds can legitimately be set to the authority of such free inquiry. As such, it would appear to be an indispensable belief for a fully liberal civilization.

See also Alembert, Jean Le Rond d'; Bacon, Francis; Comte, Auguste; Condorcet, Marquis de; Conservatism; Descartes, René; Diderot, Denis; Eliot, Thomas Stearns; Encyclopédie; Fontenelle, Bernard Le Bovier de; Hegel, Georg Wilhelm Friedrich; Liberalism; Logical Positivism; Marx, Karl; Marxist Philosophy; Mill, John Stuart; Pascal, Blaise; Popper, Karl Raimund; Pragmatism; Rationalism; Scientific Revolutions; Utopias and Utopianism.

Bibliography

Alembert, Jean d'. *Discours préliminaire de l'Encyclopédie*. Edited by F. Picavet. Paris, 1894.

Becker, Carl L. *The Heavenly City of the Eighteenth Century Philosophers*. New Haven, CT: Yale University Press, 1932.

Bishop, Michael, and Stich, Stephen. "The Flight to Reference, or How 'Not' to Make Progress in the Philosophy of Science." *Philosophy of Science* 65 (1998): 33–49.

Bury, J. B. *The Idea of Progress*. London: Macmillan, 1920.

Butterfield, Herbert. *The Whig Interpretation of History*. New York: Scribners, 1951.

Campbell, James. "Arthur Lovejoy and the Progress of Philosophy." *Transactions of the Charles S. Peirce Society* 39 (2003): 617–643.

Carrier, Martin. "How to Pile up Fundamental Truths Incessantly: On the Prospect of Reconciling Scientific Realism with Unending Progress." In *Science at Century's End: Philosophical Questions on the Progress and Limits of Science*, edited by Martin Carrier, Gerald Massey, and Laura Ruetsche. Pittsburgh, PA: University of Pittsburgh Press, 2000.

Comte, Auguste. *Cours de philosophic positive*. Paris, 1835–1852.

Condorcet, Marie Jean de. *Esquisse d'un tableau historique des progrès de l'esprit humain*, edited by O. H. Prior. Paris, 1933.

Fontenelle, Bernard de. *Digression sur les anciens et les modernes*. In *Oeuvres completes*. 5 vols. Paris, 1825. Vol. IV.

Frankel, Charles. *The Case for Modern Man*. New York: Harper, 1956.

Inge, W. R. *The Idea of Progress*. Oxford: Clarendon Press, 1920.

Kitcher, Philip. "Mathematical Progress." *Revue Internationale de Philosophie* 42 (1988): 518–540.

Maritain, Jacques. *True Humanism*. Translated by M. R. Adamson. New York, 1954.

Mill, John Stuart. *A System of Logic*. London, 1961. Book VI, Chs. 10–11.

Moulines, Ulises. "Is There Genuinely Scientific Progress?" In *On Comparing and Evaluating Scientific Theories*, edited by Adam Joniksz. Atlanta, GA: Rodopi, 2000.

Niebuhr, Reinhold. *The Nature and Destiny of Man*. New York, 1951.

Nisbet, Robert. *History of the Idea of Progress*. Revised edition. New Brunswick: Transaction Books, 1994.

Pascal, Blaise. *Fragment d'un traité du vide*. In *Oeuvres*, edited by Léon Brunschvicg and Émile Boutroux. Paris, 1904–1914. Vol. II.

Rousseau, Jean-Jacques. "Discourse on the Arts and Sciences." In *The Social Contract and Discourses*. Translated by G. D. H. Cole. London, 1913.

Spencer, Herbert. *Illustrations of Universal Progress*. New York, 1881.

Teggart, F. J., ed. *The Idea of Progress*. Rev. ed. Berkeley: University of California Press, 1949. Introduction by George H. Hildebrand.

Todd, A. J. *Theories of Social Progress*. New York: Macmillan, 1918.

Turgot, A. R. J. *On the Progress of the Human Mind*. Translated by McQ. de Grange. Hanover, NH, 1929.

Charles Frankel (1967)
Bibliography updated by Benjamin Fiedor (2005)

PROJECTIVISM

"Projectivism" has its roots in David Hume's remark in the *Treatise* about the mind's "propensity to spread itself over external objects." We sometimes speak of properties of objects where in fact the features we notice are "projections" of our internal sentiments (or other qualities of our experience). The family of metaethical views claiming that value is a projection of our conative and affective physiological states is called projectivism by Simon Blackburn (1984), and the name has stuck. Blackburn proposes that "we say that [we] *project* an attitude or habit or other commitment which is not descriptive onto the world, when we speak and think as though there were a property of things which our sayings describe, which we can reason about, know about, be wrong about, and so on" (1984, pp. 170–171). In ethics projectivism is popular because it provides an explanation of how it is that moral judgment can have the logical role that it seems to have in deciding what to do. Believing that something has some property typically provides me with a reason to act only in conjunction with the desire to promote (or oppose) the realization of that property. But believing that something is good is (or has been taken historically to be) sufficient by itself to provide a person with a reason to act. Nor is this a coincidence; it is not that we humans happen to like good things, as we happen to like to eat sugary things. Rather, it is part of the logic of judgments of goodness that they provide reasons. How can this be? Projectivists explain: the judgment that something is good is the projection of our affinity toward it, our "appetite," as Thomas Hobbes puts it.

There are three varieties of projectivism to distinguish. The most straightforward is the error theory, advanced by J. L. Mackie (1977, see also Robinson 1948), according to which our projection of value into the world is an illusion. Ordinary moral judgments presuppose an objectivity or independence of moral properties that is simply not to be had, and so they are in error. Mackie sees moral thought and language much as an atheist sees religious talk and language. The believers are not conceptually confused, but they are ontologically mistaken. The second sort of projectivism regards moral properties as Lockean "secondary qualities," not illusions, but real properties that consist in dispositions to affect human perceivers in certain ways. According to John MacDowell (1987), a leading exponent, just as we do not understand what the blueness of an object is except as the disposition to look blue to us, so we do not understand what goodness is except as the disposition to seem good to us. The projection involved in attribution of secondary qualities, including values, involves no error at all.

A third sort of projectivism is noncognitivism, or as it is more commonly called in discussions of projec-

tivism, expressivism. The expressivist holds that moral judgments do not state propositions at all but rather serve to express some noncognitive mental state of the judge. Like secondary-quality theorists, expressivists deny that there is any mistake involved in moral judgment; true, there are no moral properties, and we speak as though there are, but this "speaking as though" is just a misleading feature of the surface grammar. In fact, according to expressivists, moral judgments do not serve the same semantic function as most declarative sentences, even though they look the same.

Blackburn's projectivist position (the most influential one of the 1980s) develops an expressivist analysis of moral language with enough logical richness and complexity to model real moral deliberation and argument. His idea is easier to make out against the background of common criticisms of expressivism. Richard Brandt (1959), among others, noted that people's ordinary thinking about moral judgments runs contrary to expressivism. We have generally believed that normative judgments are used to state facts, that they are true or false, and when we change our moral views we come to regard our earlier views as mistaken, not merely as different. (By contrast, when one's taste in dessert changes, one generally regards the old preference as merely different or, at worst, childish.) Brandt complained that expressivists had given no explanation of why we are so confused. Blackburn's theory is designed to meet such objections. While maintaining an underlying expressivist semantics, he tries to show why we speak and think as though moral judgments state facts, can be true or false, and so on.

Imagine that people initially spoke about ethics in a language like English but having a quite explicitly expressivist structure. Rather than saying, "Voting for this health-care bill is morally wrong," they said, "Boo, voting for this health-care bill!" Now imagine that these speakers valued a kind of consistency of sentiment, so that it was regarded as a confusion if someone said, "Boo, eating mammals, and hooray, eating cows!" And suppose they also believed that some moral sensibilities could never survive reflection by a rational person, so that expressing one of those sensibilities would be conclusive evidence that the speaker simply had not thought carefully about the subject. The expressivist community might "invent a predicate answering to that attitude, and treat commitments as if they were judgments, and then use all the natural devices for debating truth" (Blackburn 1984, p. 195). Since Blackburn's theory seeks to defend realist-style reasoning without realist metaphysics, he calls it "quasi-realism."

An important objection to Blackburn's quasi-realism is made by Crispin Wright (1988) and Bob Hale (1990). Our moral language has a realist surface structure, and quasi-realism seeks to vindicate this structure without giving in to realist metaphysics. But if quasi-realism is successful—if every realist-sounding thing we say can be endorsed in good faith by the quasi-realist—then how will a quasi-realist be distinguishable from a full-blooded realist? As Wright puts it, Blackburn's program confronts a dilemma: Either it does not account for all the realist logical features of moral language, in which case it fails, or it succeeds in accounting for all of them, "in which case it makes good all the things which the projectivist started out wanting to deny: that the discourse in question is genuinely assertoric, aimed at truth, and so on" (1988, p. 35).

Despite these difficulties, projectivism deserves to be taken seriously, not just in the metaphysics of value, but in other metaphysical domains as well. For example, there have been projectivists about mental states (Dennett 1987—judging that someone has intentional states is taking "the intentional stance" toward the person), causes (saying that one event caused another is projecting one's psychological propensity to associate events of the first kind with events of the second in temporal sequence), probability (Finetti 1972—judgments of probability project one's degree of credence into the world), and logical impossibility (Blackburn 1984—projecting a certain kind of inconceivability). With the exception of the first, all of these sorts of projectivism are plausibly attributed to Hume, who should be regarded as the prototype projectivist.

See also Error Theory of Ethics; Hobbes, Thomas; Hume, David; Mackie, John Leslie; Metaethics; Noncognitivism; Realism.

Bibliography

Ayer, A. J. *Language, Truth, and Logic.* Harmondsworth, U.K., 1971.

Blackburn, S. *Essays in Quasi-Realism.* New York Oxford University Press, 1993.

Blackburn, S. *Spreading the Word.* Oxford: Clarendon Press, 1984.

Brandt, R. *Ethical Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1959.

de Finetti, B. *Probability, Induction, and Statistics.* London: Wiley, 1972.

Dennett, D. *The Intentional Stance.* Cambridge, MA: MIT Press, 1987.

Geach, P. T. "Ascriptivism." *Philosophical Review* 69 (1960): 221–225.

Gibbard, A. *Wise Choices, Apt Feelings.* Cambridge, MA: Harvard University Press, 1990.

Hale, B. "Can There Be a Logic of Attitudes?" In *Reality, Representation, and Projection,* edited by J. Haldane. New York: Oxford University Press, 1990.

Harman, G. *The Nature of Morality.* New York: Oxford University Press, 1977.

MacDowell, J. *Projection and Truth in Ethics.* Lawrence, KS, 1987.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* Harmondsworth, U.K.: Penguin, 1977.

Robinson, R. "The Emotivist Theory of Ethics." *Proceedings of the Aristotelian Society* 22 (1948): supplement, 79–106.

Smart, J. J. C. *Ethics, Persuasion, and Truth.* London: Routledge and Kegan Paul, 1984.

Stevenson, C. L. *Facts and Values.* New Haven, CT: Yale University Press, 1963.

Wright, C. "Realism, Anti-Realism, Irrealism, Quasi-Realism." In *Realism and Antirealism,* edited by P. French, T. Uehling, and H. Wettstein. Midwest Studies in Philosophy, Vol. 12. Minneapolis: University of Minnesota Press, 1988.

James Dreier (1996)

PROMISES

Promising is a device for obligating oneself. In a culture in which promising is available, people have a normative power that they would lack in the absence of this institution. By exercising this power—standardly through the utterance of a linguistic formula—one can bring about changes in the expectations of others in ways that enhance one's ability to pursue their goals and that foster relations of familiarity and trust.

The existence of a normative power of this kind seems philosophically puzzling: How can the utterance of a linguistic formula cause a change in the normative relations that obtain in the world? In response to this question, it might help to situate promising within the general theory of speech acts, noting that it is one of a range of illocutionary acts that may be performed with words (to be set aside such acts as asserting or commanding). In addition, one might point to the conventional aspect of promising, observing that it is a contingent social practice, sustained by interlocking sets of human dispositions, expectations, and sanctions, which enables people to coordinate their behavior in ways that promote the common good. An analogy might be to the institution of contract in the law.

But there is a normative complexity to promises that these remarks fail to capture. Agents who make a promise incur a distinctively moral obligation, opening themselves to corresponding moral complaint if they should fail to do what they have promised. Moreover this moral dimension seems crucial to the ordinary operation of promises. Promises serve to assure the promisee that something the promisee values will in fact take place—this is one way in which they differ from threats. But they achieve this effect through the promiser's implicit acknowledgement of the moral obligation that is brought into existence by the act of promising itself. Promisers give promisees to understand that they have a distinctively moral reason to do what has been promised, a reason that is strong enough to lead to performance even in the absence of independent reasons for so acting. This in turn grounds the promisee's assurance that the promised performance will take place. What accounts for the moral obligation that thus figures at the center of promissory interactions?

One answer to this question stresses the value to the agent of the normative power involved in promising. The ability to obligate oneself is a great advantage when it comes to pursuing one's projects and developing interpersonal relationships of depth and commitment. One would potentially deprive oneself of this advantageous capacity if one failed to do what one had promised, insofar as people would be less inclined to take one's promissory acts seriously. It may be doubted, however, whether this approach provides a complete account of the moral obligations brought into existence through promising. One issue is the directionality of promissory obligations. The normative powers approach focuses on the moral importance to the promiser of the ability to obligate oneself in this way. But when one fails to do what they have promised, the moral objection to their conduct turns primarily on the effects of their behavior on others.

This dimension of promissory obligations is central to a second approach, the practice view. On this view, the moral wrong involved in promise-breaking derives from the nature of promising as a valuable convention. This basic idea might be developed in a variety of ways, depending on the more general moral theory one favors. Thus, utilitarians invoke the duty to promote the impartial good, arguing that it is a violation of that duty to act in ways that undermine a highly beneficial social practice such as promising. Other theorists appeal to the idea of fairness, contending that it would be unfair to fail to do one's part to sustain a beneficial practice that one has profited from oneself. The moral duty that promising brings into existence is thus traced to fundamental social duties, in accordance with moral principles of utility or fairness.

There are two potential problems with this approach, however. First, it still does not capture the specific directionality of the moral duty involved in promissory acts. On the practice view, everyone who potentially benefits from the useful convention of promising could equally be said to be wronged when a person breaks a promise. Intuitively, however, it appears that the promisee, in particular, has a privileged ground for moral complaint. Second, it would seem possible to wrong another person in precisely the same way without exploiting a social practice such as promising. Thus, even in the absence of a promise A might deliberately lead B to believe that A will do X, where X is something A knows B wants A to do. Under these circumstances, A's failure to do X would appear to wrong B in just the same way a broken promise would have done. Yet this wrong cannot be explained by appeal to more general duties to sustain beneficial practices.

A third approach, the fidelity view, holds that promissory obligations derive from more general duties not to disappoint the expectations one has deliberately raised in others. This approach accounts well for the specific directionality of promissory obligations, and it does so in a way that explains the similarities between breaking a promise and other cases of dashed expectations. But the fidelity view encounters a different problem. It holds that the moral duty to keep one's promise is in place only when the promisee has come to expect that the promiser will perform. But as was seen above, in the promising case this kind of expectation is supposed to derive from the promiser's acknowledgement of the moral obligation to perform. There is thus a potential circularity in the interpretation of promissory interactions that is suggested by the fidelity view.

Much of the philosophical interest of promises derives from their normative complexity. An account that is adequate to this complexity might need to draw on several of the strategies sketched above, in a kind of hybrid approach. For instance, the practice view might explain how the act of promising brings into existence an initial moral obligation that is independent of the promisee's expectations. Perhaps it is the promiser's acknowledgement of this practice-based obligation that generates a corresponding expectation in the promisee. Once such an expectation is in place, the fidelity view could explain why promisers incur a further and specifically directional obligation to perform. Finally, the normative powers approach illuminates the value of the social practice of promising, highlightin the advantages gained from having the ability to obligate oneself through promissory acts.

See also Dentological Ethics; Moral Rules and Principles.

Bibliography

Anscombe, G. E. M. "Rules, Rights, and Promises." In *Ethics, Religion, Politics: Collected Philosophical Papers*. Minneapolis: University of Minnesota Press, 1981.

Fried, Charles. *Contract as Promise*. Cambridge, MA: Harvard University Press, 1981.

Kolodny, Niko and R. Jay Wallace. "Promises and Practices Revisited." *Philosophy and Public Affairs* 31 (2003): 119–154.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Rawls, John. "Two Concepts of Rules." *The Philosophical Review* 64 (1955): 3–32.

Raz, Joseph. "Promises and Obligations." In *Law, Morality, and Society. Essays in Honour of H. L. A. Hart*, edited by P. M. S. Hacker and Joseph Raz. Oxford: Clarendon Press, 1977.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1999.

Searle, John R. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge, U.K.: Cambridge University Press, 1969.

Thomson, Judith. *The Realm of Rights*. Cambridge, MA: Harvard University Press, 1990.

Warnock, G. M. *The Object of Morality*. London: Methuen, 1971.

R. Jay Wallace (2005)

PROOF THEORY

The background to the development of "proof theory" since 1960 is contained in the entry "Mathematics, Foundations of." Briefly, Hilbert's program (HP), inaugurated in the 1920s, aimed to secure the foundations of mathematics by giving finitary consistency proofs of formal systems such as for number theory, analysis, and set theory, in which informal mathematics can be represented directly. These systems are based on classical logic and implicitly or explicitly depend on the assumption of "completed infinite" totalities. Consistency of a system S (containing a modicum of elementary number theory) is sufficient to ensure that any finitarily meaningful statement about the natural numbers that is provable in S is correct under the intended interpretation. Thus, in David Hilbert's view, consistency of S would serve to eliminate the "completed infinite" in favor of the "potential infinite" and thus secure the body of mathematics represented in S. Hilbert established the subject of proof theory as a technical part of mathematical logic by means of which his program was to be carried out; its methods are described below.

In 1931 Kurt Gödel's second incompleteness theorem raised a prima facie obstacle to HP for the system Z of elementary number theory (also called Peano arithmetic—PA) since all previously recognized forms of finitary reasoning could be formalized within it. In any case Hilbert's program could not possibly succeed for any system such as set theory in which *all* finitary notions and reasoning could unquestionably be formalized. These obstacles led workers in proof theory to modify HP in two ways. The first was to seek reductions of various formal systems S to more constructive systems S'. The second was to shift the aims from foundational ones to more mathematical ones. Examples of the first modification are the reductions of PA to intuitionistic arithmetic HA and Gentzen's consistency proof of PA by finitary reasoning coupled with quantifier-free transfinite induction up to the ordinal ϵ_0, TI(ϵ_0), both obtained in the 1930s. The second modification of proof theory was promoted especially by Georg Kreisel starting in the early 1950s; he showed how constructive mathematical information could be extracted from nonconstructive proofs in number theory. The pursuit of proof theory along the first of these lines has come to be called relativized Hilbert program or reductive proof theory, while that along the second line is sometimes called the program of unwinding proofs or, perhaps better, extractive proof theory. In recent years there have been a number of applications of the latter both in mathematics and in theoretical computer science. Keeping the philosophical relevance and limitations of space in mind, the following account is devoted entirely to developments in reductive proof theory, though the two sides of the subject often go hand in hand.

METHODS OF FINITARY PROOF THEORY

Hilbert introduced a special formalism called the epsilon calculus to carry out his program (the nomenclature is related neither to the ordinal ϵ_0 nor to the membership symbol in set theory), and he proposed a particular substitution method for that calculus. Following Hilbert's suggestions, Wilhelm Ackermann and John von Neumann obtained the first significant results in finitary proof theory in the 1920s. Then, in 1930, another result of the same character for more usual logical formalisms was obtained by Jacques Herbrand, but there were troublesome aspects of his work. In 1934 Gerhard Gentzen introduced new systems, the so-called sequent calculi, to provide a very clear and technically manageable vehicle for proof theory, and reobtained Herbrand's fundamental theorem via his cut-elimination theorem. Roughly speaking, the latter tells us that every proof of a statement in quantificational logic can be normalized to a direct proof in which there are no detours ("cuts") at any stage via formulas of a complexity higher than what appears at later stages. Sequents have the form $\Gamma\to\Delta$, where Γ and Δ are finite sequences of formulas (possibly empty). $\Gamma\to\Delta$ is derivable in Gentzen's calculus LK just in case the formula $A \supset B$ is derivable in one of the usual calculi for classical predicate logic, where A is the conjunction of formulas in Γ and B is the disjunction of those in Δ.

INTRODUCTION OF INFINITARY METHODS TO PROOF THEORY

Gentzen's theorem as it stood could not be used to establish the consistency of PA, where the scheme of induction resists a purely logical treatment, and for this reason he was forced to employ a partial cut-elimination argument whose termination was guaranteed by the principle TI(ϵ_0). Beginning in the 1950s, Paul Lorenzen and then, much more extensively, Kurt Schütte began to employ certain infinitary extensions of Gentzen's calculi (cf. Schütte, 1960, 1977). This was done first of all for elementary number theory by replacing the usual rule of universal generalization by the so-called ω-rule, in the form: from $\Gamma\to\Delta,A(\mathbf{n})$ for each n = 0,1,2, …, infer $\Gamma\to\Delta,(x)A(x)$. Now derivations are well-founded trees (whose tips are the axioms $A\to A$), and each such is assigned an ordinal as length in a natural way. For this calculus LK_ω, one has a full cut-elimination theorem, and every derivation of a statement in PA can be transformed into a cut-free derivation of the same in LK_ω whose length is less than ϵ_0. Though infinite, the derivation trees involved are recursive and can be described finitarily, to yield another consistency proof of PA by TI(ϵ_0). Schütte extended these methods to systems RA_α of ramified analysis (α an ordinal) in which existence of sets is posited at finite and transfinite levels up to α, referring at each stage only to sets introduced at lower levels. Using a suitable extension of LK_ω to RA_α, Schütte obtained cut-elimination theorems giving natural ordinal bounds for cut-free derivations in terms of the so-called Veblen hierarchy of ordinal functions. In 1963 he and Solomon Feferman independently used this to characterize (in that hierarchy) the ordinal of predicative analysis, defined as the first α for which TI(α) cannot be justified in a system RA_β for $\beta<\alpha$. William Tait (1968) obtained a uniform treatment of arithmetic, ramified analysis, and related unramified systems by means of the cut-elimination theorem for LK extended to a language with formulas built

by countably infinite conjunctions (with the other connectives as usual). Here the appropriate new rule of inference is: from $\Gamma\rightarrow\Delta,A_n$, for each n = 0,1,2, …, infer $\Gamma\rightarrow\Delta,A$, where A is the conjunction of all the A_n's.

Brief mention should also be made of the extensions of the other methods of proof theory mentioned above, concentrating on elimination of quantifiers rather than cut elimination. In the 1960s Burton Dreben and his students corrected and extended the Herbrand approach (cf. Dreben and Denton, 1970). Tait (1965) made useful conceptual reformulations of Hilbert's substitution method; a number of applications of this method to subsystems of analysis have been obtained in the 1990s by Grigori Mints (1994). Another approach stems from Gödel's functional interpretation, first presented in a lecture in 1941 but not published until 1958 in the journal *Dialectica*; besides the advances with this made by Clifford Spector in 1962, more recently there have been a number of further applications both to subsystems of arithmetic and to subsystems of analysis (cf. Feferman 1993). Finally, mention should be made of the work of Dag Prawitz (1965) on systems of natural deduction, which had also been introduced by Gentzen in 1934 but not further pursued by him; for these a process of normalization takes the place of cut elimination. While each of these other methods has its distinctive merits and advantages, it is the methods of sequent calculi in various finitary and infinitary forms that have received the most widespread use.

PROOF THEORY OF IMPREDICATIVE SYSTEMS

The proof theory of impredicative systems of analysis was initiated by Gaisi Takeuti in the 1960s. He used partial cut-elimination results and established termination by reference to certain well-founded systems of ordinal diagrams (cf. Takeuti 1987). In 1972 William Howard determined the ordinal of a system ID_1 of one arithmetical inductive definition, in the so-called Bachmann hierarchy of ordinal functions; the novel aspect of this was that it makes use of a name for the first uncountable ordinal in order to produce the countable (and in fact recursive) ordinal of ID_1. In a series of contributions by Harvey Friedman, Tait, Feferman, Wolfram Pohlers, Wilfried Buchholz, and Wilfried Sieg stretching from 1967 into the 1980s, the proof theory of systems of iterated inductive definitions ID_α and related impredicative subsystems of analysis was advanced substantially. The proof-theoretic ordinals of the ID_α were established by Pohlers in terms of higher Bachmann ordinal function systems (cf. Buchholz et al. 1981). The methods here use cut-elimination arguments for extensions of LK involving formulas built by countably and uncountably long conjunctions. In addition, novel "collapsing" arguments are employed to show how to collapse suitable uncountably long derivations to countable ones in order to obtain the countable (again recursive) ordinal bounds for these systems. An alternative functorial approach to the treatment of iterated inductive definitions was pioneered by Jean-Yves Girard (1985).

In 1982 Gerhard Jäger initiated the use of the so-called admissible fragments of Zermelo-Fraenkel set theory as an illuminating tool in the proof theory of predicatively reducible systems (cf. Jäger 1986). This was extended by Jäger and Pohlers (1982) to yield the proof-theoretical ordinal of a strong impredicative system of analysis; that makes prima facie use of the name of the first (recursively) inaccessible ordinal. Michael Rathjen (1994) has gone beyond this to measure the ordinals of much stronger systems of analysis and set theory in terms of systems of recursive ordinal notations involving the names of very large (recursively) inaccessible ordinals, analogous to the so-called large cardinals in set theory.

SIGNIFICANCE OF THE WORK FOR HP AND REDUCTIVE PROOF THEORY

Ironically for the starting point with Hilbert's aims to eliminate the "completed infinite" from the foundations of mathematics, these developments have required the use of highly infinitary concepts and objects to explain the proof-theoretical transformations involved in an understandable way. It is true that in the end these can be explained away in terms of transfinite induction applied to suitable recursive ordinal notation systems. Even so, one finds few who believe that one's confidence in the consistency of the systems of analysis and set theory that have been dealt with so far has been increased as a result of this body of work. However, while the intrinsic significance of the determination of the proof-theoretic ordinals of such systems has not been established, that work can still serve behind the scenes as a tool in reductive proof theory. It is argued in Feferman (1988) that one has obtained thereby foundationally significant reductions, for example of various (prima facie) infinitary systems to finitary ones, impredicative to predicative ones, and nonconstructive to constructive ones. With a field that is still evolving at the time of writing, it is premature to try to arrive at more lasting judgments of its permanent value.

See also Gödel, Kurt; Gödel's Incompleteness Theorems; Hilbert, David; Logic, History of; Mathematics, Foundations of; Neumann, John von; Set Theory.

Bibliography

Buchholz, W., S. Feferman, W. Pohlers, and W. Sieg. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies.* Lecture Notes in Mathematics 897. New York: Springer-Verlag, 1981.

Buchholz, W., and K. Schütte. *Proof Theory of Impredicative Systems.* Naples, 1988.

Dreben, B., and J. Denton. "Herbrand-Style Consistency Proofs." In *Intuitionism and Proof Theory,* edited by J. Myhill, A. Kino, and R. E. Vesley, 419–433. Amsterdam: North-Holland, 1970.

Feferman, S. "Gödel's *Dialectica* Interpretation and Its Two-Way Stretch." In *Computational Logic and Proof Theory,* edited by G. Gottlob et al. Lecture Notes in Computer Science 713. New York: Springer-Verlag, 1993.

Feferman, S. "Hilbert's Program Relativized: Proof-Theoretical and Foundational Reductions." *Journal of Symbolic Logic* 53 (1988): 364–384.

Gentzen, G. *The Collected Papers of Gerhard Gentzen,* edited by M. Szabo. Amsterdam: North-Holland, 1969.

Girard, J.-Y. "Introduction to Π^1_2 Logic." *Synthese* 62 (1985).

Girard, J.-Y. *Proof Theory and Logical Complexity.* Naples: Bibliopolis, 1987.

Hilbert, D., and P. Bernays. *Grundlagen der Mathematik.* Vols. 1–2, 2nd ed. Berlin: Springer, 1968–1970.

Howard. W. "A System of Abstract Constructive Ordinals." *Journal of Symbolic Logic* 37 (1972): 355–374.

Jäger, G. *Theories for Admissible Sets: A Unifying Approach to Proof Theory.* Naples: Bibliopolis, 1986.

Jäger, G., and W. Pohlers. "Eine beweistheoretische Untersuchung von (Δ^1_2-CA)+(BI) und verwandter Systeme." *Sitzungsber. Bayerische Akad. Wissenschaft Math. Nat. Klasse* (1982): 1–28.

Kreisel, G. "A Survey of Proof Theory." *Journal of Symbolic Logic* 33 (1965): 321–388.

Mints, G. E. "Gentzen-Type Systems and Hilbert's Epsilon Substitution Method. I." In *Logic, Methodology, and Philosophy of Science IX,* edited by D. Prawitz et al. Amsterdam and New York: Elsevier, 1994.

Mints, G. E. *Selected Papers in Proof Theory.* Naples: Bibliopolis, 1992.

Myhill, J., A. Kino, and R. E. Vesley, eds. *Intuitionism and Proof Theory.* Amsterdam: North-Holland, 1970.

Pohlers, W. "Contributions of the Schütte School in Munich to Proof Theory." In *Proof Theory,* edited by G. Takeuti. Amsterdam: North-Holland, 1987.

Pohlers, W. *Proof Theory: An Introduction.* Lecture Notes in Mathematics 1407. New York: Springer-Verlag, 1989.

Prawitz, D. *Natural Deduction.* Stockholm: Almqvist and Wiksell, 1965.

Rathjen, M. "Admissible Proof Theory and Beyond." In *Logic, Methodology, and Philosophy of Science IX,* edited by D. Prawitz et al. Amsterdam and New York: Elsevier, 1994.

Schütte, K. *Beweistheorie.* Berlin: Springer, 1960.

Schütte, K. *Proof Theory.* Berlin: Springer-Verlag, 1977.

Sieg, W. "Hilbert's Program Sixty Years Later." *Journal of Symbolic Logic* 53 (1988): 338–348.

Simpson, S. G. "Partial Realizations of Hilbert's Program." *Journal of Symbolic Logic* 53 (1988): 349–363.

Tait, W. "Normal Derivability in Classical Logic." In *The Syntax and Semantics of Infinitary Languages,* edited by J. Barwise. Lecture Notes in Mathematics 72. New York: Springer-Verlag, 1968.

Tait, W. "The Substitution Method." *Journal of Symbolic Logic* 30 (1965): 175–192.

Takeuti, G. *Proof Theory.* 2nd ed. Amsterdam: North-Holland, 1987.

Solomon Feferman (1996)

PROPER NAMES AND DESCRIPTIONS

A *singular term* is an expression whose semantic function, when used in a particular context, is to *refer to* (denote, designate)—that is, to stand for—a single thing. A *definite description* is a singular noun phrase beginning with the definite article "the" or with a possessive noun or pronoun, as "the author of *Waverley*" and "my brilliant career." Proper names, such as "Shakespeare" and "London," are generally classified along with definite descriptions, individual variables, pronouns, and some other indexicals as singular terms. A French speaker who utters the words "*Londres est jolie*" asserts the same thing as an English speaker uttering "London is pretty." The thing asserted is a *proposition*, that London is pretty. The fundamental semantic role of a declarative sentence is to *express* (or to *contain*) a proposition (*q.v.*), which is the *semantic content* of the sentence. The proposition that Sir Walter Scott is ingenious has some component in common with the proposition that Scott is ingenuous, because both of these are directly about Scott, and some other component again in common with the proposition that Shakespeare is ingenious. These two proposition components are separately correlated with the proper name "Scott" and the predicate "is ingenious." The proposition component semantically correlated with an expression is the expression's *semantic content*. The principal philosophical controversy regarding proper names (and other singular terms) concerns the question: What are their semantic contents? The theories of John Stuart Mill (1806–1873), Bertrand Russell (1872–1970), and Gottlob Frege (1848–1925) provide rival answers.

1. THE NAIVE THEORY AND THE MILLIAN THEORY

One natural theory of semantic content is the naive theory, whose main theses are: (i) the semantic content of any singular term, as used in particular context, is its referent (bearer; the individual referred to); (ii) any semantically contentful expression refers to its semantic content; and (iii) the proposition semantically contained in a sentence is a complex, structured entity whose constituents are the semantic contents of expressions making up the sentence, typically the simple (noncompound) component expressions. (The theory may allow particular sorts of exceptions, as for example those generated by the use of quotation marks.) On the naive theory the proposition contained in "Shakespeare is ingenious" is a *singular proposition*—composed partly of things such as properties, relations, and concepts, and partly of the very individual(s) the proposition is about. By contrast, a (*purely*) *general proposition* is made up exclusively in a certain way of the former sorts of entities. On the naive theory, semantic content and reference collapse into one.

Definite descriptions pose a difficulty for the naive theory because they contain proper parts with semantic content. In *A System of Logic* (1893), Mill proffered a variant of the naive theory on which the proposition contained in "The author of *Waverley* is ingenious" is composed of something involving the attribute of authorship of *Waverley* in place of Scott himself. Mill distinguished between *denotation* (referent) and *connotation*. A general term ("concrete general name") was said by Mill to "denote" the class of individuals to which the term applies. Mill used the term "connotation" for a semantic content consisting of attributes or properties. General terms were held to have both denotation and connotation. According to Mill, definite descriptions also have both connotation and (typically) denotation, whereas proper names have only denotation. Mill's theory strongly suggests a systematic modification of the naive theory. The central theses of the Millian theory are: (i) the semantic content of any *simple* (noncompound) singular term is its referent; (ii) any expression refers to its extension; and (iii) the semantic content of a typical contentful compound expression (e.g., a definite description) is a composite entity whose constituents are the semantic contents of expressions making up the compound expression, typically the simple component expressions. (Mill's actual theory was somewhat more complex, but also somewhat less plausible.)

2. THE PUZZLES

The naive and the Millian theories give rise to philosophical puzzles concerning substitution and nonreferring names. Frege's puzzle arises from certain sentences, especially identity sentences. The sentence "Hesperus is Phosphorus" (or "The Evening Star is The Morning Star"), by contrast with "Hesperus is Hesperus," is informative. Its semantic content apparently extends knowledge. It is also a posteriori and synthetic. Yet according to both the naive theory and the Millian theory, the semantic contents of both sentences are composed of the same components, evidently in precisely the same way. Those theories thus ascribe the same semantic content to both sentences. In his early work, *Begriffsschrift* (1972 [1879], §8), Frege proposed solving this puzzle by reading the predicate for numerical identity as covertly metalinguistic: It was held that "Hesperus is Phosphorus" contains a substantive proposition concerning the names "Hesperus" and "Phosphorus," to the effect that they are co-referential. There are serious difficulties with this account, however, and Frege came to reject it. Most significantly, the account fails to solve the general problem of which "Hesperus is Phosphorus" is a special case. Unless the theory is part of a more sweeping proposal concerning all expressions and not just that of identity predicates, there is no explanation for the analogous difference in epistemic and semantic status between "Hesperus is a planet if Phosphorus is" (synthetic *a posteriori*) and "Hesperus is a planet if Hesperus is" (analytic *a priori*).

A second puzzle is the apparent failure of substitution in special contexts, especially those of propositional attitude. Jones may sincerely and reflectively assent to "Hesperus appears in the evening sky" and sincerely and reflectively dissent from "Phosphorus appears in the evening sky," even while fully grasping their semantic content. This appears to violate the classical logical rule of Leibniz's law, or the substitutivity of equality. Both the naive theory and the Millian theory treat "Jones believes that Hesperus appears in the evening" and its substitution instance "Jones believes that Phosphorus appears in the evening" as having the same content, and therefore also the same truth-value.

A further nest of problems concerns sentences involving nonreferring proper names. The sentence "Sherlock Holmes is addicted to cocaine" clearly has content. Yet on both the naive theory and the Millian theory, the semantic content of any sentence will lack a necessary component if any contained name lacks a referent. It is evident, moreover, that this sentence (taken as a statement of real fact, rather than as a statement made from

within the fiction) cannot be counted literally true. But, it seems, neither can its negation—"Sherlock Holmes is not addicted to cocaine"—be truly uttered. This seems to violate the classical law of excluded middle. These puzzles are especially pressing with regard to negative existentials, such as "Sherlock Holmes does not exist." This sentence is true if and only if "Sherlock Holmes exists" is false, and therefore, it would seem, if and only if the referent of "Sherlock Holmes" lacks existence. Yet the negative existential itself implies that the name does not so much as have a referent. How, then, can it be true? Indeed, how can it have any content at all?

3. RUSSELL'S THEORY OF DESCRIPTIONS

Russell's semantic theory (post-1904) is a supplement to the naive theory. Russell employed *propositional functions* in lieu of attributes. A propositional function assigns to any objects in its domain a singular proposition concerning those objects. Russell's general theory of descriptions, or of what he called "denoting phrases," consisting of a noun phrase preceded by a determiner such as "all" or "some," assigns a content to sentences in which they figure while denying that the determiner phrases themselves are meaningful units. The theory analyzes sentences of both the Aristotelian *A* form, "Π(all *S*)" (e.g., "All millionaires are wealthy"), and the *I* form, "Π(some *S*)," where Π is a monadic predicate. (More generally, Π may be the result of filling all but one of the argument positions of an n-adic predicate, $n \geq 1$.) The *A* form is analyzed as "For everything x, if x is a *S*, then Π(x)"—more colloquially as, "Everything is such that: if it is a *S*, then Π(it)" ("Everything is, if a millionaire, then wealthy"). The complex predicate "is such that: if it is a *S*, then Π(it)" stands for a certain propositional function, whereas the quantifier "everything" stands for a higher-level propositional function, which assigns to any first-level propositional function, *F*, the proposition that *F* is "always true"—that is, the proposition that *F* yields a true proposition for each and every argument.

Russell analyzed "Π(some *S*)" as "Something is such that: it is a *S* and Π(it)"—wherein the complex predicate "is such that: it is a *S* and Π(it)" stands for a certain propositional function said to be "sometimes true"—that is, to yield a true proposition for at least one argument. An English phrase of the form "all *S*" thus corresponds to the incomplete string, "everything is such that: if it is a *S*, then it … ," and a phrase of the form "some *S*" corresponds to the incomplete string, "something is such that: it is a *S* and it.…" Russell called phrases of either form *incomplete symbols*. The sentences in which such phrases figure have content, though the phrase, in and of itself, does not contribute a proposition-component to the proposition expressed. As Russell put it in "On Denoting," "denoting phrases have no meaning in isolation."

The introduction of a quantifier ("everything," "something") into the analysis gives rise to ambiguities analogous to that of "every boy kissed a girl" when the simple Aristotelian sentential form occurs within the scope of a governing operator, such as "not," "necessarily," or "Jones believes." Thus, on Russell's general theory of descriptions, a sentence of the form "not Π(all *S*)" (e.g., "All millionaires are not wealthy") may be analyzed by giving the indefinite description "all *S*" *primary occurrence* (over "not"), yielding: "Everything is such that: if it is a *S*, then not-Π(it)." This reading is equivalent to the Aristotelian *E* form, "Π(no *S*)." Alternatively, and nonequivalently, "not Π(all *S*)" may be analyzed by giving the phrase "all *S*" *secondary occurrence*, yielding the reading, "Not everything is such that: if it is a *S*, then Π(it)." (The latter analysis—equivalent to the Aristotelian *E* form—is obtained by letting the negation in "not Π(all *S*)" govern the entire *A* form, not just its predicate Π.) Similarly, "Jones believes Π(some *S*)" may be analyzed as "Something is such that: it is S and Jones believes that Π(it)" (primary occurrence), or alternatively, and nonequivalently, as "Jones believes: that Π(some *S*)" (secondary).

In most cases, only one of the two readings is plausibly intended (as with "Jones believes some husbands are bachelors"). If the simple Aristotelian *A* or *I* form occurs with two or more governing operators, the number of readings is compounded. For example, "Jones believes some millionaires are not wealthy" may be analyzed alternatively, and nonequivalently, as: (*i*) "Someone is a millionaire and Jones believes he/she is not wealthy" (*wide scope*); (*ii*) "Jones believes: that someone is both a millionaire and not wealthy" (*intermediate scope*); or (*iii*) "Jones believes: that no one is both a millionaire and wealthy" (*narrow scope*).

The central tenet of Russell's theory of definite descriptions is that a description such as "the author of *Waverley*" (used in the sense of "the sole author of *Waverley*") is semantically equivalent to the corresponding uniqueness-restricted existential quantifier "some unique author of *Waverley*," in the sense of "something such that it, and nothing else, wrote *Waverley*." The restricted quantifier falls under the purview of Russell's general theory of descriptions. On Russell's theory, then, "the author of *Waverley*" corresponds to the string "Someone is such that: he or she uniquely wrote *Waverley* and he or

she … ," making definite descriptions also "incomplete symbols" which have "no meaning in isolation." The words "The author of *Waverley* is ingenious" are not directly about Walter Scott, but about the complex propositional function, *being a unique author of*Waverley *who is also ingenious*, expressing that this function yields a true proposition for at least one individual. There is nothing that the phrase "the author of *Waverley*" contributes on its own to this proposition.

As with "some *S*," sentences that position a definite description within governing operators yield multiple readings. For example, "Jones believes the author of *Waverley* is not ingenious" may be analyzed alternatively, and nonequivalently, as: (i) "Someone uniquely wrote *Waverley* and Jones believes he is not ingenious"—that is, Jones believes *of Waverley*'s sole author that *he* is not ingenious (*wide scope*); (ii) "Jones believes: that someone both uniquely wrote *Waverley* and is not ingenious"—that is, Jones believes that whoever wrote *Waverley* single-handedly is not ingenious (*intermediate scope*); or (iii) "Jones believes: that no one both uniquely wrote *Waverley* and is ingenious" (*narrow scope*). The wide-scope reading is consistent with Jones's belief not involving a conception of Scott as sole author of *Waverley*. The narrow-scope reading attributes a belief that is consistent with *Waverley* not having a sole author.

A definite description is said to be *proper* when there is someone or something that uniquely answers to the description, and is *improper* otherwise. Russell artificially, and misleadingly, extended Mill's term "denotation" to the semantic relation that obtains between a proper definite description and the individual uniquely described, even though a definite description is supposed not to be a singular term. He might instead have called this relation "simulated denotation." Russell retained the term "meaning" for semantic content.

Both the Millian theory and Russell's theory deny that the individual that uniquely answers to a definite description is itself a component of the content of sentences involving the description. Those theories are able to solve the puzzles in the special case where the terms involved are definite descriptions rather than proper names, by reading sentences involving definite descriptions as containing propositions involving corresponding attributes or propositional functions. In particular, Russell's claim that definite descriptions are not singular terms, but quantificational constructions, blocks substitutivity of equality, which is applicable only to singular terms, from licensing the substitution of "the first Postmaster General" for "the inventor of bifocals" in the secondary-occurrence reading of "Jones believes that the inventor of bifocals was clever." (By contrast, the envisioned substitution is indeed licensed by logical principles, including substitutivity as applied to variables, when the sentence takes on its primary-occurrence reading.)

Russell handled the same difficulties in the case of proper names (and such devices as demonstratives) through his thesis that names are ordinarily not used as "genuine names" (singular terms). Instead they were held to be "disguised" or "abbreviated" definite descriptions. The proposition expressed by a sentence involving a typical name is to be analyzed in accordance with Russell's theory of descriptions. This blocks substitution in sentences such as "Jones believes that Hesperus appears in the evening." Russell acknowledged the possibility of "names in the strict, logical sense" (logically proper names), which function in accordance with the naive theory. The class of admissible semantic contents for usable genuine names was severely limited by Russell's principle of acquaintance, that every proposition one can grasp must be composed entirely of constituents with respect to which one has a special sort of intimate and direct epistemic access, (*direct*) *acquaintance*. This restriction seems sufficient to prevent the puzzles from arising with logically proper names. (Russell did not countenance genuine names lacking a referent. Curiously, he claimed that singular existential and negative existential statements involving genuine names are without meaning. It would have been better to say that such sentences are always trivially true and trivially false, respectively.)

4. FREGE'S THEORY OF *SINN* AND *BEDEUTUNG*

In his classic paper, "*Über Sinn und Bedeutung*" (1892), Frege abandoned the naive theory in favor of a richly elegant philosophy of semantics, which extends the Millian theory's two-tiered semantics for definite descriptions and predicates to include all meaningful expressions. (Like Mill, and unlike Russell, Frege counted definite descriptions as singular terms.) Frege distinguished between the referent (*Bedeutung*) of an expression and its sense (*Sinn*). The sense of an expression contains a purely conceptual manner of presenting the name's referent. Individuals that are not themselves senses—such as persons and even their sensations—cannot be constituents of a genuine Fregean sense. Furthermore, the sense of a singular term secures the term's referent. An expression's sense is a conception of something, and the expression's referent, if there is one, is whatever uniquely fits the concept. The reference relation is thus the relative product of

a purely semantic relation (that between an expression and its sense) and a nonlinguistic relation (that between a sense and the object that fits it). Third, the sense of an expression is the semantic content. Expressions having the same sense must have the same referent, but importantly, expressions having the same referent may differ in sense. Frege illustrated his notion of sense by means of three lines that intersect in a single point. Then the phrases "the point of intersection of *a* and *b*," "the point of intersection of *a* and *c*," and "the point of intersection of *b* and *c*" converge in reference but diverge in sense.

The observation that proper names have a sense, as distinct from the referent, is tailor-made to solve both Frege's Puzzle and the problem of how sentences involving nonreferring names can have content. Frege's solution to the substitution problem is more complex. Crucial to Frege's theory are the principles of extensionality and compositionality. They hold that the referent or sense, respectively, of a complex expression is a function of the referents or senses, respectively, of the component expressions. In the latter case Frege spoke metaphorically of the sense of a constituent expression as a *part* of the sense of the complex expression, so that the sense of the whole is composed of the senses of the parts.

Thus, if a constituent expression is replaced by one having the same sense, the sense of the whole is preserved, whereas if a constituent expression is replaced by one having the same referent but a different sense, the referent of the whole is preserved even though the sense is not. In particular, Frege held as a special case of extensionality that a compound expression having a nonreferring part must be nonreferring ("Sherlock Holmes's older brother"). Frege argued, using extensionality, that the *cognitive value* (*Erkenntniswerte*) of a sentence is not the referent of the sentence, but is fixed by its sense, and that the referent of a sentence is one of two truth values, truth and falsity ("the true" and "the false"). Because a sentence refers to its truth-value, and a sentence involving a nonreferring name itself refers to nothing, such a sentence as "Sherlock Holmes is addicted to cocaine" is neither true nor false. (Frege held that the sentence *presupposes*, without asserting, that Sherlock Holmes exists.)

Frege argued that certain expressions create a special context in which subordinate expressions do not refer to their customary referent. When occurring within quotation marks (for example, in "direct discourse" reporting the words used by a speaker) an expression refers to itself. Analogously, expressions occurring subordinate to operators such as "Jones believes that" and "Jones said that" (the latter occurring in "indirect discourse" reporting the content of a speaker's utterance) refer to their *ungerade* (indirect, oblique) referent, which is the customary sense. Extensionality is to be understood as requiring the validity of substituting for a name in a sentence any expression having the same referent *in that same position*. (Scattered remarks suggest that Frege might have applied his doctrine of semantic shifting also to the problem of negative existentials.)

5. THE THEORY OF DIRECT REFERENCE

Despite a fundamental disagreement over the matter of singular propositions, there is common ground between Russell and Frege in regard to ordinary proper names. Both held a strong version of the theory that names are *descriptional*. On their view, if "St. Anne" is analyzable as "the mother of Mary," it must be analyzable even further, because "Mary" is also supposed to be descriptional. But even "the mother of the mother of Jesus" must be in this sense further analyzable. If "α" is a nondescriptional singular term referring to Mary, then it may be said that the description "the mother of α" is *descriptional relative to* Mary. A *thoroughly descriptional* term is one that is descriptional but not descriptional relative to anything. The orthodox theory, shared by Russell and Frege, is the theory that proper names and similar devices are either thoroughly descriptional or descriptional relative only to items of direct acquaintance. Frege held the stronger thesis (which is retained by contemporary variants of Frege's theory, such as that of John Searle) that proper names are thoroughly descriptional. Any departure from the stronger thesis would constitute a rejection of fundamental Fregean theory.

In recent philosophy the orthodox theory has been forcefully challenged, most notably by Keith Donnellan (1972), David Kaplan, Saul Kripke (1972, 1979), Ruth Barcan Marcus, and Hilary Putnam. These philosophers favor the theory of direct reference, which holds that proper names (and similar devices) are nondescriptional. Importantly, this theory does not deny that particular names may exhibit any or all of the three aspects of a Fregean sense mentioned in the previous section. What is denied is that the conceptual representation carried by a name secures the referent. But the direct-reference theory is significantly stronger than a simple denial of Russell's doctrine that ordinary names are abbreviated definite descriptions. The theory holds that names are not even similar to definite descriptions. An immediate consequence is that a great many definite descriptions fail to be thoroughly descriptional or descriptional relative only to

items of direct acquaintance, because many contain names of ordinary individuals.

Three main kinds of arguments have been advanced in favor of the direct-reference theory. The modal and epistemological arguments are due chiefly to Kripke. Suppose for simplicity that the name "Shakespeare" simply means "the English playwright who wrote *Hamlet*, *Macbeth*, and *Romeo and Juliet*." If the orthodox theory of names is correct, then the sentence, "Someone is Shakespeare iff he is an English playwright who is sole author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*," should express a necessary, a priori truth. On the contrary, however, it might have come to pass that Shakespeare elected to enter a profession in law instead of becoming a writer. Furthermore, it is possible, and is not ruled out solely by semantic reflection, that Francis Bacon should go on to write these plays. These intuitions are supported by a complementary intuition: that "Shakespeare" continues to refer to the same person even with respect to nonactual possible worlds in which Shakespeare lacks the distinguishing characteristics that people actually use to identify him—that is, even in discourse about such a counterfactual scenario. One important consequence of the direct-reference theory is that any proper name is a *rigid designator* (Kripke)—that is, it designates the same thing with respect to every possible world in which that thing exists and does not designate anything else with respect to other possible worlds.

One example of the semantic arguments for the direct-reference theory comes from Donnellan: According to the orthodox theory, the semantic content of the name "Thales" is determined by a description such as "the Greek philosopher who held that all is water." But suppose that the man referred to by writers from whom the use of the name "Thales" derives never genuinely believed that all is water but was thought to, owing to some error or hoax, and that, by coincidence, there was a Greek hermit who did hold this bizarre view, though he bears no historical connection to anyone. Contrary to the orthodox theory, the name "Thales" would nevertheless refer to the first of the two. This argument seems to reveal also that the surrounding settings in which speakers find themselves, and not merely the concepts evoked in them, are crucial to determining the referents of the names they use. In a word, the securing of a referent for a name is a *contextual* phenomenon. Donnellan and Kripke have provided partial accounts of the securing of a referent for a name by means of historical chains of communication. Putnam has given a similar account of certain terms designating something by means of a "division of linguistic labor." Because of these accounts the direct-reference theory is sometimes called the causal theory of reference.

6. THE MILLIAN THEORY RECONSIDERED

What, then, is the semantic content of a name? It is tempting to answer that it is, or at least includes, a descriptive or conceptual "mode of presentation." Although this proposal does not require that the associated mode of presentation secure the referent, it faces some of the same difficulties as the orthodox theory. A more general difficulty arises because the variations of the argument from Frege's Puzzle against the naive theory and the Millian theory can be mounted against a wide variety of theories of semantic content, including Frege's. The general strategy involved in that argument, however, seems to involve an error. This might be demonstrated through an application to a situation involving expressions for which it is uncontroversial that semantic content is exactly the same.

Suppose that foreign-born Sasha learns the words "ketchup" and "catsup" by actually consuming the condiment and reading the labels on the bottles. Suppose further that, because of his idiosyncratic experience, Sasha comes to believe that the substances so named are different condiments sharing a similar taste, color, and consistency. Whereas "Ketchup is ketchup" is uninformative for Sasha, "Catsup is ketchup" is informative. It would be a mistake, however, to conclude that "catsup" and "ketchup" differ in semantic content for Sasha. The terms are perfectly synonymous in English; indeed, they are arguably the same English word. Most English speakers learned one in a sort of ostensive definition, and the other as a strict synonym (or as an alternative spelling) of the first. If either may be learned by ostensive definition, then both may be—witness Sasha. This discredits the original argument from Frege's puzzle.

One important consideration favoring the Millian theory over the orthodox theory comes by consideration of individual variables. Consider the following propositional-attitude attribution:

(1) The planet Venus is an individual x such that Jones believes that x is a star.

It is characteristic of this *de re* (as opposed to *de dicto*) locution that it does not specify how Jones is supposed to conceive of Venus in believing it to be a star. The Orthodox Theorist contends that this is a result of the allegedly descriptional name "Venus" positioned outside of the scope of the nonextensional operator "Jones believes

that," where it is open to substitution and to existential generalization. What is more significant, however, is that a nondescriptional singular term is positioned within the scope of the nonextensional context: the last occurrence of the variable "*x*" in (1). It follows by the principles of conventional semantics that (1) is true if and only if its component open sentence:

(2) Jones believes that *x* is a star

is true under the assignment of Venus as value for the variable. In turn, (2) is true under this assignment if and only if Jones believes the semantic content of the complement open sentence:

(3) *x* is a star

under the same assignment. But the fundamental characteristic of a variable with an assigned referent is that its semantic content is just its referent. This is precisely the point of using a variable rather than a definite description (such as "the first heavenly body visible at dusk") within the scope of an attitude verb in a *de re* attribution. If a variable with an assigned value had, in addition to its value, a Fregean sense, then (3) would contain a specific general proposition, under the relevant assignment. If (1) is to fail to specify how Jones conceives of Venus, the content of (3) under the assignment of Venus to "*x*" can only be the singular proposition about Venus that it is a star. If the open sentence (3), under the assignment of Venus as the value of "*x*," contains the singular proposition about Venus that it is a star, then so does the closed sentence "*a* is a star," where "*a*" is an individual constant that refers to Venus. It is not the variability of a variable, but its structural simplicity, that gives it the feature that the variable's semantic content, under an assignment of a referent, is just the assigned referent. (An exactly parallel argument proceeds using pronouns in place of variables, using "The planet Venus is such that Jones believes that *it* is a star.")

It is important to note also that at least some aspects of the remaining puzzles would arise even in a language for which it was stipulated that the Millian theory is correct. Suppose, for example, that an authoritative linguistic committee that legislates the grammar and semantics of the language, and to which all speakers of the language give their cooperation and consent, decreed that proper names are to function exactly like the mathematician's variables, "*x*," "*y*," and "*z*," except that they are to remain constant. Ordinary speakers would presumably continue to regard co-referential names as not always interchangeable in propositional-attitude attributions. English speakers who use "ketchup" and "catsup" as exact synonyms may be inclined to assent to "Sasha believes that ketchup is a sandwich condiment, but he does not believe that catsup is." On philosophical reflection, however, it emerges that this expresses a logical impossibility. Similarly, speakers who agree to abide by the legislative committee's decree about proper names might for independent pragmatic reasons be led to utter or to assent to such sentences as "Jones believes that Hesperus appears in the evening, but he does not believe that Phosphorus does." Insofar as the same phenomena that give rise to the puzzles would arise even in the case of a language for which the Millian theory was true by fiat and unanimous consent (and do in fact arise with respect to such straightforward synonyms as "ketchup" and "catsup"), the puzzles cannot be taken as evidence against the Millian theory. A deeper understanding is needed of the puzzles, and a reexamination of the Millian theory in light of this deeper understanding.

See also Demonstratives; Indexicals; Quantifiers in Natural Language.

Bibliography

Church, A. *Introduction to Mathematical Logic*. Vol. 1, secs. 1–6. Princeton, NJ: Princeton University Press, 1956.

Donnellan, K. "Proper Names and Identifying Descriptions." In *Semantics of Natural Language*, edited by D. Davidson and G. Harman, 356–379. Dordrecht, Netherlands: D. Reidel, 1972.

Frege, G. *Begriffsschrift*. 1879. Translated by G. Frege as "Conceptual Notation." In his *Conceptual Notation and Related Articles*. Oxford: Oxford University Press, 1972.

Frege, G. "*Über Sinn und Bedeutung*." 1892. Translated as "On Sense and Meaning" or (preferably) "On Sense and Reference" or "On Sense and Nominatum." Widely reprinted. Available in Harnish, *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Harnish, R. M. *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Kripke, S. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1972. Excerpted in Harnish, *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Kripke, S. "A Puzzle about Belief." 1979. In Salmon and Soames, *Propositions and Attitudes*, 102–148. Also in Harnish, *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Mill, J. S. "Of Names." In *A System of Logic*. 1893. Bk. 1, chap 2. Excerpted in Harnish, *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Russell, B. "Knowledge by Acquaintance and Knowledge by Description." In *Propositions and Attitudes*, edited by N. Salmon and S. Soames, 16–32. Oxford: Oxford University Press, 1988.

Russell, B. "On Denoting." In *Mind*. 1905. Widely reprinted. Available in Harnish, *Basic Topics in the Philosophy of Language*. London: Prentice Hall/Harvester Wheatsheaf, 1993.

Russell, B. *The Philosophy of Logical Atomism*, edited by D. Pears. La Salle, IL: Open Court, 1985.

Salmon, N. "A Millian Heir Rejects the Wages of *Sinn*." In *Propositional Attitudes: The Role of Content in Logic, Language, and Mind*, edited by C. A. Anderson and J. Owens, 215–247. Stanford, CA: Center for the Study of Language and Information, 1990

Salmon, N. "A Problem in the Frege-Church Theory of Sense and Denotation." *Noûs* 27 (2) (June 1993): 158–166.

Salmon, N. "Reference and Information Content: Names and Descriptions." In *Handbook of Philosophical Logic*. Vol. 4, *Topics in the Philosophy of Language*, edited by D. Gabbay and F. Guenthner, 409–461. Dordrecht, Netherlands: D. Reidel, 1989.

Salmon, N., and S. Soames, eds. *Propositions and Attitudes*. Oxford: Oxford University Press, 1988.

Searle, J. "Proper Names." *Mind* 67 (266) (1958): 166–173.

Nathan Salmon (2005)

PROPERTIES

Our every assertion or thought involves "properties" or relations. Most simply, we predicate some property of some thing: Earth is round. Sometimes we refer to properties by name or by description: Red is the color of blood. Sometimes our quantifiers range over properties: Galaxies come in many shapes and sizes.

This familiarity with properties, however, does not reveal what properties are. Indeed, the question is equivocal, both in ordinary and in philosophical discourse. There are different conceptions of properties, equally legitimate, corresponding to the different roles that properties have been called upon to play (Bealer 1982; Lewis 1983, 1986). And for each conception there are different theories as to what sort of entity, if any, is best suited to play the role. The most fundamental division is between abundant and sparse conceptions of properties. On an abundant conception every meaningful predicate expresses some property or relation, including "is blue or round," "is on top of a turtle," "is identical with the planet Mars"; a property's instances need not resemble one another in any intrinsic respect. Abundant properties are needed to serve as "meanings," or components of "meanings," in a compositional semantics for language. On a sparse conception of properties a predicate expresses a property only if the objects satisfying the predicate resemble one another in some specific intrinsic respect; perhaps "has unit positive charge" and "is ten kilograms in mass" are examples. Sparse properties are needed to provide an objective basis for the scientist's project of discovering the fundamental classifications of things and the laws that govern them. Properties, whether abundantly or sparsely conceived, are neither language- nor mind-dependent: They existed before there were beings to talk and think about them; they would have existed even had there never been such beings.

In this entry only conceptions of properties are explicitly distinguished and discussed, although much of what is said applies also to relations and to propositions. Other philosophers' terms for *property* in the abundant sense include *attribute* (Quine 1970), *propositional function* (Russell 1919), and *concept* (Bealer 1982, Frege 1884); *universal* and *quality* have for the most part been interpreted sparsely. Ordinary language allows abundant or sparse readings of *characteristic, feature, trait,* and more.

ABUNDANT CONCEPTIONS OF PROPERTIES

How abundant are the properties on the abundant conception? Whenever there are some things, no matter how scattered or dissimilar from one another, there is the abundant property of being one of those things. Thus, for any class of things, there is at least one abundant property had by all and only the members of that class. It follows that there are at least as many abundant properties as classes of things and that the abundant properties outrun the predicates of any ordinary language. (There are nondenumerably many classes of things—assuming an infinity of things—but at most denumerably many predicates in any ordinary language.) Abundant properties, owing to their very abundance, must be transcendent, rather than immanent: They are not present in their instances as constituents or parts. It is not plausible to suppose that an object has a distinct constituent for each and every class to which it belongs.

If we say that whenever there are some things, there is exactly one property had by all and only those things, then a property may be identified with the class of its instances. For example, the property of being human may be identified with the class of human beings. But there is a well-known objection to this identification (Quine 1970). Consider the property expressed by "is a creature with a heart" and the property expressed by "is a creature with kidneys." If properties are "meanings," or semantic values, of predicates, then the properties expressed by these two predicates are distinct. Yet, these predicates, we may suppose, are coextensive: As a matter of fact, any creature with a heart has kidneys, and vice versa; the class

of creatures with a heart is identical with the class of creatures with kidneys. Thus, distinct properties correspond to the same class and cannot be identified with that class.

Different responses to the objection invoke different criteria of individuation for properties, that is, different criteria for deciding when properties, introduced, say, via predicates that express them, are one and the same. One response simply denies that "is a creature with a heart" and "is a creature with kidneys" express distinct properties. More generally, properties expressed by coextensive predicates are identical. Call this an extensional conception of properties. A property so conceived may be identified with the class of its instances. Extensional conceptions of properties are adequate to the semantic analysis of mathematical language and extensional languages generally (Tarski 1946).

A second response holds that "is a creature with a heart" and "is a creature with kidneys" express distinct properties, because it is logically possible for something to satisfy one predicate without satisfying the other. On this response properties expressed by necessarily coextensive predicates are identical; properties expressed by accidentally coextensive predicates are distinct. Call this an intensional conception of properties. If one accepts the standard analyses of logical possibility and necessity in terms of possibilia, then a property, on the intensional conception, may be identified with the function that assigns to each possible world the set of possible objects that has the property at the world. If one holds that each object exists at, and has properties at, only one world, then a property may more simply be identified with the class of (actual and) possible objects that has the property (Lewis 1986). Properties, on the intensional conception, are appropriate semantic values for predicates of (standard) modal languages and intensional languages generally (Carnap 1947, Kripke 1963).

A third response holds that the properties expressed by "is a creature with a heart" and "is a creature with kidneys" are distinct because they are structured entities with different constituents: The property expressed by "is a creature with a heart" has the property expressed by "is a heart" as a constituent; the property expressed by "is a creature with kidneys" does not. On this response properties have a quasi-syntactic structure that parallels the structure of predicates that express them. Call two predicates isomorphic if they have the same syntactic structure and corresponding syntactic components are assigned the same semantic values. On a structured conception of properties, properties expressed by isomorphic predicates are identical; properties expressed by nonisomorphic predicates are distinct. (Structured conceptions are sometimes called hyperintensional because they allow necessarily coextensive predicates to express distinct properties.) Structured conceptions subdivide according to whether the unstructured semantic values are intensional or extensional and according to whether the relevant structure is surface grammatical structure, or some hypothetical deep structure, or structure after analysis in terms of some chosen primitive vocabulary. Structured properties may be identified with sequences of unstructured properties and other unstructured semantic values. Structured properties, on one version or another, have a role to play in the semantic analysis of propositional attitudes and of hyperintensional languages generally (Carnap 1947, Cresswell 1985).

Thus far, this entry has assumed that predicates of ordinary language are satisfied by objects once and for all. In fact, most ordinary language predicates are tensed; they may be satisfied by objects at some times but not at others. For example, "is sitting" is true of me now, but was false of me ten minutes ago. On a tensed conception of properties, whether or not a property holds of an object may also be relative to times. Most simply, tensed properties may be identified with functions from times to untensed properties. Tensed properties may be taken as semantic values for tensed predicates.

We have, then, a plurality of abundant conceptions of properties. Which is correct? One need not and should not choose. A plurality of conceptions is needed to account for the multiple ambiguity in our ordinary talk of properties. And it seems that both structured and intensional conceptions are needed for compositional semantics: Structured properties are needed to provide distinct semantic values for predicates, such as "is a polygon with three sides" and "is a polygon with three angles," that are necessarily coextensive without being synonymous; intensional properties are needed to provide distinct semantic values for unstructured predicates that are accidentally coextensive. To accept a plurality of conceptions, it suffices to find, for each conception, entities that satisfy that conception's criteria of individuation.

Realists with respect to some conception of properties hold that entities satisfying the individuation criteria for the conception exist. Realists divide into reductionists and antireductionists. Reductionists identify properties, under the various conceptions, with various set-theoretic constructions (in ways already noted): class, functions, or sequences of actual or possible objects (Lewis 1986). Antireductionists reject some or all of these identifications. For some antireductionists, classes are suspect or

esoteric entities; classes are to be explained, if at all, in terms of properties, not vice versa (Bealer 1982, Russell 1919). For other antireductionists the problem is not with classes, but with the possibilia that comprise them (on intensional conceptions). Possible but nonactual entities are to be explained, if at all, in terms of uninstantiated properties, not vice versa (Plantinga 1976). According to the antireductionist, properties are basic or primitive; it is merely posited that there are entities satisfying the appropriate individuation criteria. Some entities, after all, must be taken as basic; according to the antireductionist, properties are an acceptable choice.

Eliminativists hold that, strictly speaking, there are no properties. They take aim, typically, at intensional conceptions, at conceptions with modal criteria of individuation. They claim that modal notions, such as logical possibility and necessity (whether taken as primitive or analyzed in terms of possibilia), incorrigibly lack the clarity and precision required of a rigorous scientific semantics or philosophy (Quine 1970). Eliminativists have the burden of showing how ordinary and philosophical discourse ostensibly referring to properties can be paraphrased so as to avoid such reference; or, failing that, of showing that such discourse is dispensable, merely a *façon de parler.*

SPARSE CONCEPTIONS OF PROPERTIES

On an abundant conception any two objects share infinitely many properties and fail to share infinitely many others, whether the objects are utterly dissimilar or exact duplicates. On a sparse conception the sharing of properties always makes for genuine similarity; exact duplicates have all of their properties in common. Whatever the sparse properties turn out to be, there must be enough of them (together with sparse relations) to provide the basis for a complete qualitative description of the world, including its laws and causal features. The sparse properties correspond one-to-one with a select minority of the abundant properties, on some intensional conception. ("Intensional," because distinct sparse properties may accidentally be instantiated by the same objects.) Those abundant properties that correspond to sparse properties are called natural (or perfectly natural, since naturalness presumably comes in degrees; Lewis 1983, 1986). The naturalness of properties is determined not by our psychological makeup, or our conventions, but by nature itself.

How sparse are the properties, on a sparse conception? First, there is the question of uninstantiated properties. If sparse properties are transcendent, there is no difficulty making room for uninstantiated sparse properties; perhaps uninstantiated sparse properties are needed to ground laws that come into play only if certain contingent conditions are satisfied (Tooley 1987). If, on the other hand, sparse properties are immanent, are present in their instances, then uninstantiated sparse properties must be rejected, because they have nowhere to be (Armstrong 1978, 1989). Of course, uninstantiated sparse properties may nonetheless possibly exist, where this is understood according to one's favored interpretation of modality.

Second, there is the question of the compounding of sparse properties (and relations). Disjunctions and negations of natural properties are not themselves natural: Their instances need not resemble one another in any intrinsic respect. For example, instances of the property having-unit-positive-charge-or-being-ten-kilograms-in-mass need not resemble one another in either their charge or their mass. It follows that there are no disjunctive or negative sparse properties (Armstrong 1978).

The case of conjunctive sparse properties is less clear. There are two views. According to the first, since instances of a conjunction of natural properties, such as having-unit-positive-charge-and-being-ten-kilograms-in-mass, resemble one another in some—indeed, at least two—intrinsic respects, there exists a sparse property corresponding to the conjunction. According to the second view, the sparse properties must be nonredundant; they must be not only sufficient for describing the world but minimally sufficient. On this view conjunctive sparse properties are excluded on grounds of redundancy: A putative conjunctive sparse property would hold of an object just in case both conjuncts hold.

Similarly, structural sparse properties, such as being-a-molecule-of-H_2O, may be admitted on the grounds that they make for similarity among their instances. Or they may be excluded on grounds of redundancy: A putative structural sparse property would hold of an object just in case certain other sparse properties and relations hold among the object and its parts. But the exclusion of structural (and conjunctive) sparse properties faces a problem. It rules out a priori the possibility that some properties are irresolvably infinitely complex: They are structures of structures of structures, and so on, without ever reaching simple, fundamental properties or relations (Armstrong 1978). A sparse conception that allowed for this possibility would have to allow some redundancy; and if some redundancy, why not more? This suggests that conjunctive and structural sparse properties should generally be admitted. (An alternative treatment makes use of degrees

of naturalness and has it that conjunctive and structural properties are natural to some lesser degree than the properties in terms of which they are defined; a world with endless structure has no perfectly natural properties.)

If structural sparse properties are admitted, the sparse properties will not be confined to fundamental physical properties; there will be sparse properties of macroscopic, as well as microscopic, objects. For example, the sparse properties will include specific shape-and-size properties, such as being-a-sphere-ten-meters-in-diameter (which are arguably structural properties definable in terms of sparse distance relations). However, the vast majority of ordinary-language predicates—"is red," "is human," "is a chair," to name a few—fail to express natural properties to which sparse properties correspond; rather, these predicates express properties that, when analyzed in fundamental physical terms, are disjunctive (perhaps infinitely so) and probably extrinsic. (This judgment could be overturned, however, if there are irreducible natural properties applying to macroscopic objects—most notably, irreducible phenomenological properties of color, sound, and such.)

What are the properties on a sparse conception? There are three principal theories (or clusters of theories, since they each subdivide). According to the first, the properties sparsely conceived are just some of the properties abundantly conceived: The properties that are perfectly natural. What makes some properties natural and others unnatural? One version of the theory simply takes naturalness to be a primitive, unanalyzable distinction among abundant properties (Quinton 1957; see also Armstrong 1989, Lewis 1986). But since a property is natural in virtue of the resemblances among its instances, it might seem more appropriate to take instead some relation of partial resemblance as primitive and to define naturalness in terms of resemblance. The resulting version, called resemblance nominalism, can be worked out in different ways with different primitive resemblance relation (Price 1953; see also Armstrong 1989, Goodman 1951; Lewis 1983). The chief objection to the view is that partial resemblance between ordinary objects, no less than naturalness of properties, cries out for analysis. When two objects partially resemble one another, the objection goes, they must have constituents that exactly resemble one another, perhaps constituents that are literally identical. More generally, it is argued, properties must be constituents of objects if properties are to play a role in the explanation of the natures and causal powers of objects; one cannot explain an object's nature or causal powers by invoking a class to which it belongs. Sparse properties, then, must be immanent, not transcendent, entities.

What are these constituents of ordinary things? Not ordinary spatial or temporal constituents—or, at least, not always. For even an object with no spatial or temporal extension might have a complex nature and stand in relations of partial resemblance. If sparse properties are immanent, then they must be nonspatiotemporal constituents of things. There are two prominent theories as to the nature of these constituents. The first theory takes them to be universals (Armstrong 1978, 1989.) They are repeatable: Each of them is, or could be, multiply instantiated. And they are wholly present in their instances: An immanent universal is located—all of it—wherever each of its instances is located. When objects resemble one another by having a sparse property in common, there is something literally identical between the objects. It follows that universals fail to obey commonsense principles of location, such as that nothing can be (wholly) in two places at the same time. But that is no objection. Such principles were framed with particulars in mind; it would beg the question against universals to require them to meet standards set for particulars.

On the other theory of sparse properties as immanent, the nonspatiotemporal constituents of ordinary particulars are themselves particulars, called tropes (Armstrong 1989, Lewis 1986, Williams 1966) or abstract particulars (Campbell 1981). When ordinary particulars partially resemble one another by having some sparse property—say, their mass—in common, then there are distinct, exactly resembling, mass tropes as constituents of each. On a trope theory sparse properties can be identified with maximal classes of exactly resembling tropes (perhaps including merely possible tropes). Exact resemblance between tropes is taken as primitive by trope theory; but it is a simple and natural primitive compared to the partial resemblance relation taken as primitive by an adequate resemblance nominalism.

A possible disadvantage of a universals theory is that it requires two fundamentally distinct kinds of entities: universal and particulars. An ordinary particular cannot simply be identified with a bundle of coinstantiated universals, lest numerically distinct but qualitatively identical particulars be identified with one another. On a universals theory there must be some nonqualitative, nonrepeatable constituent of ordinary particulars to ground their numerical identity. A trope theory, on the other hand, needs only tropes to make a world. Ordinary particulars can be identified with bundles of coinstantiated tropes; numerically distinct but qualitatively identical

particulars are then bundles of numerically distinct but exactly resembling tropes.

The great advantage of a universals theory is that it promises to analyze all resemblance in terms of identity: Exact resemblance is identity of all qualitative constituents; partial resemblance is partial identity, identity of at least one qualitative constituent. But it is unclear whether the promise can be kept. Objects instantiating different determinates of a determinable—such as unit-positive and unit-negative charge—seem to partially resemble one another by both being charged without there being any analysis of this resemblance in terms of the identity of constituent universals or, for that matter, the exact resemblance of constituent tropes. A universals theory and a trope theory would then have to fall back upon primitive partial resemblance between universals, or tropes. Some of the advantages of these theories over resemblance nominalism would be forfeited.

Of the three basic theories of sparse properties—resemblance nominalism, a theory of immanent universals, and a theory of tropes—only one can be true; the theories posit incompatible constituent structure to the world. However, assuming each theory is internally coherent, and adequate to the needs of science, the question arises, What sort of evidence could decide between them? It seems that a choice between the theories will have to be made, if at all, on the basis of pragmatic criteria such as simplicity, economy, and explanatory power. There is as yet no philosophical consensus as to what that choice should be.

See also Armstrong, David M.; Carnap, Rudolf; Eliminative Materialism, Eliminativism; Frege, Gottlob; Goodman, Nelson; Kripke, Saul; Metaphysics; Quine, Willard Van Orman; Realism; Reduction; Reductionism in the Philosophy of Mind; Russell, Bertrand Arthur William; Tarski, Alfred.

Bibliography

Armstrong, D. M. *Universals and Scientific Realism.* 2 vols. Cambridge, U.K.: Cambridge University Press, 1978.

Armstrong, D. M. *Universals: An Opinionated Introduction.* Boulder, CO: Westview Press, 1989.

Bealer, G. *Quality and Concept.* Oxford: Clarendon Press, 1982.

Campbell, K. "The Metaphysic of Abstract Particulars." In *Midwest Studies in Philosophy.* Vol. 6. Minneapolis: University of Minnesota Press, 1981.

Carnap, R. *Meaning and Necessity.* Chicago: University of Chicago Press, 1947.

Cresswell, M. *Structured Meanings.* Cambridge, MA: MIT Press, 1985.

Frege, G. *Die Grundlagen der Arithmetik* (1884). Translated by J. L. Austin as *The Foundations of Arithmetic.* Oxford: Blackwell, 1950.

Goodman, N. *The Structure of Appearance.* Cambridge, MA: Harvard University Press, 1951.

Kripke, S. "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94.

Lewis, D. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61 (1983): 343–377.

Lewis, D. *On the Plurality of Worlds.* Oxford: Blackwell, 1986.

Plantinga, A. "Actualism and Possible Worlds." *Theoria* 42 (1976): 139–160.

Price, H. H. *Thinking and Experience.* London: Hutchinson's University Library, 1953.

Quine, W. V. O. *Philosophy of Logic.* Englewood Cliffs, NJ: Prentice Hall, 1970.

Quinton, A. "Properties and Classes." *Proceedings of the Aristotelian Society* 58 (1957): 33–58.

Russell, B. *Introduction to Mathematical Philosophy.* London: Allen and Unwin, 1919.

Tarski, A. *Introduction to Logic.* 2nd ed. Oxford, 1946.

Tooley, M. *Causation.* Oxford: Clarendon Press, 1987.

Williams, D. C. "The Elements of Being." In *The Principles of Empirical Realism.* Springfield, IL: Thomas, 1966.

Phillip Bricker (1996)

PROPERTY

The institution of property has interested social philosophers in part, at least, because it raises issues of justice. Like government, it is practically universal but varies enough in its particular arrangements to suggest the question What criteria are relevant in assessing the relative merits of various arrangements? Again, because it discriminates between rights and fortune, it invites moral criticism and the demand for justification.

Many of the classical accounts of the origin and function of private property have taken for granted that in nature all things were held "in common." This phrase, however, is ambiguous, for it often meant not a system regulating the use of goods by general agreement but a condition where, there being no rules, everything was *res nullius* (a thing belonging to no one) and the concept "property" was consequently irrelevant. How, then, it was asked, would humans come to appropriate the land and its fruits? How could such appropriation be justified? What would be rational grounds for claiming exclusive possession? And could there be any limit on people's right to do what they would with their own?

THEORIES OF PROPERTY

According to the Church Fathers, property was both the consequence and the social remedy for the sin of covetousness that came with the Fall. But since owners have appropriated what at one time belonged to all, they have a duty to administer it for the benefit of all. "Our property," said Gregory the Great, "is ours to distribute, but not ours to keep." The concept of the owner as steward is the core of the traditional Christian view of property.

NATURAL LAW AND CONVENTIONALISM. By the seventeenth century, property rights came to be grounded in the needs and accomplishments of the individual owner, and ownership implied a natural right to enjoy and dispose of its objects, limited only by the duty to respect the rather narrowly defined interests of others. In John Locke's account, property as an institution is explained by human needs. Although God gave the earth and all its fruits to all people to preserve their lives, still this meant one's making the fruits of the earth exclusively his own, if only by eating them. However, what in nature entitled one to call something one's own was that one made the effort to make it so. To add one's labor to a *res nullius* was to create a title to the whole product. Locke limited this title to whatever one could use before it spoiled; appropriation for waste would be illegitimate.

To appropriate an object implied for Locke not merely a right to enjoy it but also to alienate it at will, so that although the appropriation of *res nullius* could legitimately be effected only by labor, the title, once established, could be freely transferred. It is questionable, however, whether Locke was justified in assuming that because we may appropriate what we need from the common stock, we may therefore transfer what we acquire, but do not need, to whomsoever we choose. Locke needed this right, however, if his theory was not to suggest, as did certain later writers on economic justice, that the laborer was entitled to the entire fruits of his master's fields, if not to the fields themselves. For where all land had long been appropriated, the titles of present owners would depend entirely on the legitimacy of such transfers in the past. So, since the land was no longer *res nullius,* all the laborer could claim was the value of his labor in wages. Moreover, in a market economy and with the introduction of money, wealth might be accumulated and stored indefinitely without spoiling; furthermore, since money had only a conventional value, hoarding it deprived no one of anything of natural value, and its distribution must be taken to be by common consent. Having accounted, then, for the existence of property, and for existing titles, with a theory of natural right, Locke overlaid the theory with a conventionalist theory that neutralized the limitations on appropriation that the original theory prescribed.

Nature and convention are to be found similarly blended, if in varying proportions, in Hugo Grotius, Samuel von Pufendorf, and William Blackstone. In Immanuel Kant, too, there is a blend. Kant deduced the principle of first occupier from the autonomy of the will but conceded that only a universal legislative will—the civil state—could give binding force to the intention to appropriate.

UTILITARIAN POSITIONS. According to David Hume, a man's creation ought to be secured to him in order to encourage "useful habits and accomplishments." Inheritance and the right to alienate were alike valuable as incentives to or conditions for useful industry and commerce. Property rested on convention in the sense of rules upheld by common interests commonly perceived. It was a law of nature, too, but in the sense that men were sufficiently alike the world over for the same general arrangements to be equally to the public advantage. Hume's argument, then, also blends natural law doctrine with conventionalism but reduces both to utilitarianism.

Jeremy Bentham did little more than elaborate Hume's arguments. However, by introducing considerations of utility, Hume and Bentham pointed the way for criticism of the distribution of private property and, indeed, of the institution itself. Already in 1793 William Godwin was arguing that in a consistent application of the principle of utility "every man has a right to that, the exclusive possession of which being awarded to him, a greater sum of benefit or pleasure will result, than could have arisen from its being otherwise appropriated" (*Political Justice,* Book 8). J. S. Mill, though broadly committed to a belief in private property, held that, in the case of land at any rate, private ownership must be conditional on its expediency; the rights associated with it, especially the right to exclusive access and enjoyment, ought to be limited to whatever was required to exploit it efficiently. Mill recognized that the rights of property were not an inseparable bundle, to be justified en bloc; each constituent right had to be independently justified on grounds of utility.

However, Mill's belief that the institution of property would be justified provided that it guaranteed to individuals the fruits of their labor and abstinence is open to question. In a complex industrial society, "the fruits of one's labor" can mean only the value of a given worker's

contribution to the finished product. But value derives from the relations of supply and demand, both for the commodity and for labor of the various kinds needed to produce it. "The fruits of one's labor," understood as one's share in a social dividend, will depend not only on one's efforts but also on the number of other people available to do the same job and on how badly consumers want it done. If for the time being a particular skill is in short supply, is it self-evident that this increases the value of its fruits or that those who have it should be the better off for it?

MARXIST AND HEGELIAN CRITIQUES. Again, the exclusive claims of labor take no account of what men owe to others and to the social interest. Émile Durkheim, for instance, objected that "it is not enough to invoke the rights that man has over himself: these rights are not absolute but limited by the claims of the moral aims, in which a man has to cooperate." Karl Marx was equally critical of the German Social Democrats' Gotha program of 1875, which claimed that labor should receive its produce "unabridged and in equal right." He charged that this formula ignored the need for capital replacement and development, social services, and the support of the incapable. In any case, he said, distribution proportional to contribution would still be only partial justice, bearing in mind differences in natural capacity on the one hand and need on the other. In the truly cooperative society, based on common ownership of the means of production, individual labor would be impossible to separate out, and distribution would be according to need alone. This would be possible, however, only because labor would have ceased to be a burden and would have become "life's principal end."

This last condition suggests why, in a period when hedonistic premises underlay a great deal of psychology, ethics, and economics, the necessary relation between labor and property should have been so generally accepted. On the assumption that work was painful, the only conceivable reason for working was a greater pleasure expected from its fruits. Marx argued that this account of labor was neither an explanation nor a justification, but a consequence, of the system of private property. The worker was alienated from his work, which appeared to him not as a fulfillment but as a burden; he was alienated, too, from the product of his work, which, passing to his employer in surplus value, confronted him as capital—that is, as an instrument of his own bondage.

Despite the stress on labor as the source of value that Marx shared with the English utilitarians and economists, his account of property derives at least as much from G. W. F. Hegel as from the English school. Like Kant, Hegel regarded property as necessary not because it helped to satisfy human needs but because "a person must translate his freedom into an external sphere in order that he may achieve his ideal existence" (*Philosophy of Right,* Sec. 41); because "property is the first embodiment of freedom and so is in itself a substantive end" (Sec. 45). Plato erred, in Hegel's view, in denying private property to the guardians, for he was denying them the conditions necessary for giving concrete realization to their personalities and wills.

Marx and Hegel are alike in seeing the human will objectifying itself in its acquisitions and creations. If for Marx the process is not rationalizing and liberating but alienating and enslaving, it is because the property created is not and cannot be the worker's own. The laborer can transcend this alienation only in the communist society, in which, like Plato's guardians or the members of a monastic community, he gets caught up in a common enterprise where "mine" and "thine" are of no account because life is more than the satisfaction of material needs. In a world in which "sharing in" counted for more than "sharing out," property—like justice—would present no problems.

ECONOMIC AND SOCIAL SIGNIFICANCE. In the course of the past century, legal and social philosophers (Léon Duguit and Karl Renner, for example) have come to think of property increasingly as an institution with social functions and not, like Locke, as simply a guarantee of individual interest. Moreover, because property entails inequalities in power, in claims on the social product, in social status, and in prestige, it must be justified, and not merely in terms of the interests or natural rights of its immediate beneficiaries.

It is difficult, however, to see how any one theory could apply generally to all forms of private property and include all rights of ownership. Individual control of productive resources raises very different issues from the exclusive right to enjoy consumer goods such as clothes and furniture. The right to control the use of mines and factories is not really an instance of the right of a Kantian rational and autonomous being to manipulate mere things for his own needs; it is also an exercise of power over other people. According to A. A. Berle, the United States is gradually extending to such property the limitations traditionally applied to state action in order to protect individual freedom.

Again, could one justify one's title to dividends on the ground that instead of enjoying the fruits of one's labor one had invested them? And would the same justification extend to a corporate title to the yield on investments financed out of undistributed profits? Such claims have certainly flourished under the umbrella of natural rights; but it is difficult to see how any but a utilitarian argument could seriously be proposed in defense of such arrangements.

ANALYSIS OF OWNERSHIP

Talk of property often seems to be talk about things. Things constitute property, however, only inasmuch as they can be assigned to owners; to own something is to have, in respect to it, certain rights and liabilities vis-à-vis other persons or the public at large. Ownership, therefore, is a normative relation or a complex of such relations between owner, object, and third parties, and to refer to something as "property" is to locate it as a term in such a relationship. Some jurists, indeed, insist that *property* refers not to things at all but, rather, to a bundle of rights. And this is obviously true of income titles, such as securities and annuities, and of rights of control over "intellectual property," such as patent rights and copyrights; these are "things" only in a very abstract sense, as characteristic complexes of normative relations.

As the objects of property are diverse, so also are the rights constituting it. Landowners' rights are necessarily different from copyright owners', and the owner of a gun does not have the same unrestricted use and control of it as the owner of a table has of the table. Jurists have nevertheless tried to identify some right necessary to ownership. The rights of exclusive use, possession, or alienation seem to be likely candidates, but each can conceivably be detached (for example, by a lease or an easement, under the terms of a trust, or, in former times, by entail) without the owners' losing property in the object. Accordingly, Sir Frederick Pollock suggested that "we must look for the person having the residue of all such powers when we have accounted for every detached and limited portion of it." But this residue, as held, say, by a ground landlord with a thousand-year tenant, may be very slender indeed, and the owner to whom all the detached rights will revert when the encumbrances reach the end of their term will certainly not be the present owner.

A. M. Honoré suggests a way out of these difficulties by concentrating not on the difficult exceptions but on the standard instance. He defines ownership as "those legal rights, duties and other incidents which apply, in the ordinary case, to the person who has the greatest interest in a thing admitted by a mature legal system." Among the characteristic features are the right to possess and to be secure in possession, to use and to manage the property, to enjoy income arising from it and to alienate, consume, waste, or destroy the capital, and to transmit ownership to one's successors indefinitely; the absence of a fixed date on which the owners' interests terminate; the prohibition of harmful use; the liability of the property to execution for debt or insolvency; and the reversion to the owner on the termination of whatever lesser interests (leases, usufructs) encumber the property. Now, to say that *A* is the owner of *x* is not necessarily to say that he is the present subject of all these incidents; however, provided the kind of property in question can intelligibly be said to be the object of them and in the absence of special conditions or reservations, it is reasonable to infer that he is.

The Scandinavian legal realists—Karl Olivecrona and Alf Ross, for example—have been more radical in their analyses. According to Ross, ownership is "solely … a tool in presentation." Theoretically, one could enunciate a mass of directives to judges, each consisting of a conditioning fact or facts (F) and a legal consequence (C), such as (1) if a person has lawfully purchased a thing (F_1), judgment for recovery of possession should be given in his favor (C_1); (2) if a person by prescription has acquired a thing and raised a loan that is unpaid (F_2), the creditor should be given judgment for satisfaction out of the thing (C_2); and so on. Now, to introduce "ownership" is not, according to Ross, to add something that *accounts* for the connections between the F's and the C's but merely to indicate the systematic connection between them such that $F_1, F_2, F_3, \cdots, F_p$ severally and collectively entail the totality of legal consequences $C_1, C_2, C_3, \cdots, C_n$. The word *ownership* in Ross's view is "without any semantic reference whatever"; it serves only to reduce the complexity of particular rules to a systematic order. There is nothing beyond or in addition to the rules.

Now, it is certainly true that only confusion can result from trying to identify some special kind of a thing, or some special quality of things, which is called "property." Nevertheless, "ownership" does not always imply the same bundle of rights. The possible conditioning facts and the legal consequences are not the same for every case in which one may say that *X* is the owner of *P*. And, therefore, since the relevant rules do not have the rigorous relation to one another that Ross suggests, one can identify them as the rules of property (as distinct from, say, personal rights) only by recognizing some sort of family resemblances between them. Indeed, the terms Ross uses in exemplifying his conditioning facts—"purchase,"

"occupation of *res nullius,*" "acquisition by prescription"—are obviously already impregnated with ownership; to purchase something, for example, is to give money for it—that is, on the understanding that one acquires not merely possession, but also owners' rights, over it.

Deciding who is the owner of a piece of property is, of course, to decide on the basis of certain facts where certain powers and liabilities lie. But to reduce a legal concept like property to a finite set of directives to judges ignores the fact that judges are constantly having to reshape the rules in the very process of applying them. If the rules of ownership are treated as a more or less arbitrary agglomeration, it is difficult to see how judges could make rational decisions at all.

Ross's bundle of conditioning facts and legal consequences is significant, however, because it suggests how one goes about constructing a paradigm case of ownership, or, rather, a family of paradigms related by the fact that different conditioning facts entail broadly similar legal consequences. Deciding ownership in an atypical case would then involve deciding whether it can be assimilated to any of the available paradigms even though some characteristic ownership features are absent or other features that are out of character are present. A judge may have all kinds of reasons for making or refusing such an assimilation; but it is difficult to see how the problem could be presented to him at all without presupposing the standard cases of ownership as agreed starting points for discussion.

See also Bentham, Jeremy; Durkheim, Émile; Godwin, William; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Hume, David; Justice; Kant, Immanuel; Locke, John; Marx, Karl; Mill, John Stuart; Natural Law; Patristic Philosophy; Plato; Pufendorf, Samuel von; Rights; Utilitarianism.

Bibliography

Dias, R. W. M. *A Bibliography of Jurisprudence.* London, 1964. Extensive annotated bibliography, mainly from the juristic standpoint.

GENERAL HISTORIES

Bartlett, J. V. et al., eds. *Property, Its Duties and Rights.* London: Macmillan, 1915.

Grace, Frank. *The Concept of Property in Modern Christian Thought.* Urbana: University of Illinois Press, 1953.

Macpherson, C. B. *The Political Theory of Possessive Individualism: Hobbes to Locke.* Oxford: Clarendon Press, 1962.

Schlatter, Richard. *Private Property: The History of an Idea.* London: Allen and Unwin, 1951.

CLASSICAL THEORIES

Hegel, G. W. F. *Grundlinien der Philosophie des Rechts.* Berlin, 1821. Translated, with notes, by T. M. Knox as *Hegel's Philosophy of Right.* Oxford: Clarendon Press, 1942.

Hume, David. *Treatise of Human Nature* (1640), edited by L. A. Selby-Bigge. 2nd ed. Oxford: Clarendon Press, 1896.

Locke, John. *Two Treatises of Government* (1690). Edited by Peter Laslett. Cambridge, U.K.: Cambridge University Press, 1960.

Marx, Karl. *Critique of the Gotha Programme.* In *Karl Marx and Friedrich Engels: Selected Works.* 2 vols. Moscow, 1949–1951. Vol. II.

Marx, Karl. "Ockonomisch-philosophische Manuskripte aus dem Jahre 1844." In *Marx-Engels Gesamtausgabe,* edited by D. Riazanovsky and A. Adoratski. Berlin, 1932. Part I, Vol. III, edited and translated by T. B. Bottomore as "Economic and Philosophical Manuscripts" in *Karl Marx: Early Writings.* London, 1963.

Mill, J. S. *Principles of Political Economy.* 7th ed. London, 1871.

ECONOMIC AND SOCIAL SIGNIFICANCE

Berle, A. A. "Property, Production and Revolution." *Columbia Law Review* 65 (1) (1965): 1-20.

Berle, A. A., and G. C. Means. *The Modern Corporation and Private Property.* New York: Macmillan, 1933.

Duguit, Léon. *Les transformations générales du droit privé.* Paris, 1912.

Durkheim, Émile. *Leçons de sociologie physique des moeurs et du droit.* Paris: Presses Universitaires de France, 1950. Translated by Cornelia Brookfield as *Professional Ethics and Civic Morals.* London: Routledge, 1957.

Renner, Karl. *Die Rechtsinstitute des Privatrechts und ihre soziale Funktion.* Tübingen, 1929. Translated by Anges Schwarzschild as *The Institutions of Private Law,* edited by Otto Kahn-Freund. London, 1949.

ANALYSIS OF OWNERSHIP

Honoré, A. M. "Ownership." In *Oxford Essays in Jurisprudence,* edited by A. G. Guest. London: Oxford University Press, 1961.

Noyes, C. Reinold. *The Institution of Property.* New York: Longmans, Green, 1936.

Olivecrona, Karl. *Law as Fact.* London: Oxford University Press, 1939.

Pollock, Frederick. *Jurisprudence and Legal Essays.* London: Macmillan, 1961.

Ross, Alf. *On Law and Justice.* London: Stevens, 1958.

Simpson, A. W. B. "The Analysis of Legal Concepts." *Law Quarterly Review* 80 (October 1964): 535–558.

Stanley I. Benn (1967)

PROPERTY [ADDENDUM]

What is property? It is some valued item that belongs to someone. Its existence in society may be collective or individual, although even if collective, it usually emerges from instances of (pooled or expropriated) individual

ownership. And that presupposes the right to *private* property.

PROPERTY IS PRIVATE

The institution of the right to private property is the single most important condition for a society in which freedom in the classical liberal tradition—which means negative liberty, including free trade, freedom of the press, and freedom of religion—is to flourish. Under communism, in contrast, no such right is recognized. Privacy has a negligible role in a system which holds, as Karl Marx (1818–1883) proclaimed, that "the human essence is the true collectivity of man" (1970, p. 126). Even within noncommunist, nonsocialist systems the exact status of property is in dispute—some hold it is a convention established by implicit consensus and maintained by government or law. Some hold it is a natural normative relationship that comes about by means of the creative and productive initiative of persons and the law of property exists to recognize and not to create it.

WHAT IS THE RIGHT TO PRIVATE PROPERTY?

Karl Marx understood the right to private property, although it was John Locke (1632–1704) who tried to justify this right. Marx wrote, in "On the Jewish Question," that "the right of man to property is the right to enjoy his possessions and dispose of the same arbitrarily without regard for other men, independently, from society, the right of selfishness" (1970, p. 53). This, though correct, is not the full story. The right to private property, be it applied to obtaining and holding a toothbrush or, as was Marx's concern (and what Marx found objectionable), an entire factory, does spell out a person's authority to use what he or she owns without regard for other persons. This use may be reckless, prudent, or generous. Its exercise may not, however, violate others' rights. Defenders do not assume that it would be insidious.

The natural right to private property was only discussed in direct terms starting in the eleventh and twelfth centuries. William of Ockham (1285–1347) proposed that "Natural right is nothing other than a power to conform to right reason, without an agreement or pact" (2001, p. 48) or, as Heinrich A. Rommen paraphrased him, "the right to private property is a dictate of right reason" (1954, p. 419), the power to make one's moral choices on one's own, free of others' intrusion. Because such choices are made by persons in the natural world, one of our natural rights is the right to private property.

ONE ROLE OF PRIVATE PROPERTY IN SOCIETY

Property rights weren't explicitly identified in ancient times but the Old Testament ban on stealing implies what was spelled out by Locke and other classical liberals. Moreover, there have been strong philosophical intimations of it in, for example, Aristotle's *Politics* (384–322 BCE). Whereas Plato, his teacher, held that, at least within the ruling class of a political community, there may not be any private property and indeed privacy, at all, Aristotle objected as follows:

> That all persons call the same thing mine in the sense in which each does so may be a fine thing, but it is impracticable; or if the words are taken in the other sense, such a unity in no way conduces to harmony. And there is another objection to the proposal. For that which is common to the greatest number has the least care bestowed upon it. Every one thinks chiefly of his own, hardly at all of the common interest; and only when he is himself concerned as an individual. For besides other considerations, everybody is more inclined to neglect the duty which he expects another to fulfill; as in families many attendants are often less useful than a few. (*Politics*, 1261b34)

Earlier Thucydides (c. 471–c. 400 BCE) said,

> They devote a small fraction of the time to the consideration of any public object, most of it to the prosecution of their own objects. Meanwhile, each fancies that no harm will come to his neglect, that it is the business of somebody else to look after this or that for him; and so, by the same notion being entertained by all separately, the common cause imperceptibly decays. (*The History of the Peloponnesian War*, bk. 1, sec. 141)

So, communal ownership leads to reduction of responsibility and a corresponding lack of attentive involvement with whatever is owned. This does not mean that people are evil. At their homes, this is likely to be different—if one is late and rushes off, the trash will be disposed of upon one's return. At a public place the attitude seems to be, "It will get cleaned up somehow, by someone, at some time." So, it is a systemic problem: people are unable to incorporate the significance of managing the public property within the scale of their values. Each of us knows, directly, how important or not it is for oneself to keep one's backyard clean. So one will take care of it commensurate with that knowledge. It is not possible, however, for an individual to know how important it is for the

community, society, or humanity at large that one keep the air or river or lake clean, and to what degree.

A more recent defense of the right to private property is closer to that which we get from John Locke; namely, that we require this right so as to have a sphere of moral authority—as Robert Nozick (1938–2002) called it, "moral space," or as Ayn Rand (1905–1982) noted,

> Bear in mind that the right to property is a right to an action, like all others: it is not the right to an object, but to the action and the consequences of producing or causing that object. It is not a guarantee that a man will earn any property, but only a guarantee that he will own it if he earns it. It is the right to gain, to keep, to use and to dispose of material values. (1967, p. 322)

Basically, then, the main normative reason given for why one has a right to private property is that it is the means by which one's liberty to act free of others' imposition is secured within a social context. It is also a precondition for individuals to act prudently and productively in human communities without the legal permission for others to take from them what they have earned. Economists tend, in contrast, to defend it as a feature of the infrastructure by which productivity and prosperity is best encouraged in a society. Another support given to the idea is that it makes it possible for individuals to remain sovereign and to distribute resources as they see fit rather than others would demand.

There are innumerable objections to the right to private property, most recently the idea that property is held by the public at large and government merely permits individuals to make use of it to the extent government deems this in the public interest. For why this is a troublesome view the general theory of natural rights would need to be explored and scrutinized.

See also Civil Disobedience; Cosmopolitanism; Postcolonialism; Republicanism.

Bibliography

Aristotle. *Politics*. Chicago: University of Chicago Press, 1984.

Bandman, Elsie and Bertram Bandman, eds. *Bioethics and Human Rights: A Reader for Health Professionals*. Boston: Little, Brown, 1978.

Comte, August. *The Catechism of Positive Religion*. 1852. Clifton, NJ: Kelley, 1973.

Locke, John. *The Second Treatise on Government*, edited by Thomas P. Peardon. New York: Liberal Arts Press, 1952.

Machan, Tibor R. *Individuals and Their Rights*. La Salle, IL: Open Court, 1989.

Machan, Tibor R. *The Right to Private Property*. Stanford, CA: Hoover Institution Press, 2002.

Marx, Karl. *Selected Writings*, edited by David McLellan. London: Oxford University Press, 1970.

Nagel, Thomas, and Liam Murphy. *The Myth of Ownership: Taxes and Justice*. London: Oxford University Press, 2003.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Ockham, William. *Opus Nonaginta Dierum* Lewiston, NY: E. Mellen Press, 2001.

Rand, Ayn. *Capitalism: The Unknown Ideal*. New York: New American Library, 1967.

Rommen, Heinrich A. "The Genealogy of Natural Rights." *Thought* 29 (1954).

Thucydides. *The History of the Peloponnesian War*. Mineola, NY: Dover Publications, 2004.

Tibor R. Machan (2005)

PROPOSITIONAL ATTITUDES: ISSUES IN SEMANTICS

Propositional attitudes like knowledge, belief, and assertion play an important foundational role for semantic theory, the goal of which is to specify the meanings of sentences and their semantic contents relative to contexts of utterance. Meanings are plausibly regarded as functions from such contexts to semantic contents, which in turn are closely related to the assertions made, and the beliefs expressed, by utterances. For example, the semantic content of *I live in New Jersey* in a context C with x as agent and t as time is standardly taken to be the proposition that x lives in New Jersey at t. To understand the meaning of this sentence is, to a first approximation, to know that a competent speaker x who sincerely and assertively utters it in C asserts, and expresses a belief in, this proposition. Roughly put, if p is the semantic content of S in C, then an assertive utterance of S in C is an assertion of p, and is standardly taken as indicating the speaker's belief in p. Whether the semantic content of a sentence is always among the propositions asserted by an utterance of the sentence, and whether, in those cases in which it is, the assertion of any other proposition by the utterance is always parasitic on the assertion of the semantic content, are matters of detail. Though important, they do not affect the foundational point. A semantic theory for a language is part of a larger theory that interprets the assertions and beliefs of its speakers. This, more than any other fact, allows one to subject semantic theories to empirical test. Competent speakers of a language are relatively good at identifying the propositions

asserted and beliefs expressed by utterances. To the extent to which assignments of semantic content issued by a semantic theory lead to verifiably correct characterizations of speakers' assertions and beliefs, the semantic theory is confirmed; to the extent to which these assignments lead to verifiably incorrect characterizations, it is disconfirmed.

ATTITUDE ASCRIPTIONS

This point is closely related to the use of attitude ascriptions

(1a) N asserted that S

(1b) N believed that S

to test different semantic analyses of S. It is convenient to express this in terms of the relational nature of the attitudes. Consider assertion. In each case of assertion there is someone, the agent, who does the asserting, and something, the object of assertion, that is asserted. The term *proposition* is used to designate things that are objects of assertion (and other propositional attitudes) and bearers of truth value. Assertion is a mediated relation holding between agents and propositions. An agent asserts a proposition p by doing something or employing some content-bearing representation associated with p. The most familiar cases are those in which the agent asserts a proposition by assertively uttering a sentence.

Ascriptions like those in (2) report the assertions of agents:

(2a) Edward asserted the proposition that Martha denied.

(2b) Edward asserted the proposition that the Earth is round.

(2c) Edward asserted that the Earth is round.

That in (2a) *asserted* is flanked by two noun phrases suggests that it is a two-place predicate and that a sentence ⌜NP assert NP⌝ is true if and only if the first (subject) noun phrase designates an agent who bears the assertion relation to the entity designated by the second (direct object) noun phrase. This analysis also applies to (2b), which is true if and only if Edward asserted the proposition designated by *the proposition that the Earth is round*. On the assumption that this proposition is also designated by *that the Earth is round*, this analysis can be extended to (2c), which is equivalent to (2b). Similar remarks hold for other propositional attitude verbs, including *believe*, *deny*, *refute*, and *prove*.

With this in mind, one can return to the ascriptions in (1). If, as many theorists believe, (i) ⌜that S⌝ in (1) designates the semantic content of S (in the context), (ii) these ascriptions report relations between agents and those contents, and (iii) sometimes substitution of sentences with necessarily equivalent semantic contents fails to preserve the truth values of such ascriptions, then semantic contents must be more fine-grained than the sets of possible world-states in which they are true. On these assumptions substitution in such ascriptions can be used to discriminate different but intensionally equivalent semantic analyses of S.

PROPOSITIONS, POSSIBLE WORLD-STATES, AND TRUTH SUPPORTING CIRCUMSTANCES

This has significance for possible world semantics. In this framework a semantic theory is a formal specification of truth with respect to a possible context of utterance and circumstance of evaluation. The semantic content of S in C is the set of possible circumstances E such that S is true with respect to C and E. Circumstances of evaluation are traditionally identified with possible world-states—thought of as maximally complete properties that the world genuinely could have had. As a result, the semantic contents of all necessarily equivalent sentences are taken to be identical. This, plus the standard treatment of attitude ascriptions as reporting relations between agents and the semantic contents of their complement clauses, leads to the counterintuitive prediction that substitution of necessarily equivalent sentences in such ascriptions never changes truth value. If one adds the apparently obvious fact that (3a) entails (3b),

(3a) A asserts/believes that P&Q

(3b) A asserts/believes that P&A asserts/believes that Q

one gets the further counterintuitive results (i) that anyone who asserts or believes a proposition p asserts or believes all necessary consequences of p, and (ii) that no one ever asserts or believes anything necessarily false, since to do so would involve simultaneously asserting or believing every proposition.

In 1983 Jon Barwise and John Perry attempted to evade these results by constructing a semantic theory in which metaphysically possible world-states were replaced by abstract situations—thought of as properties that need be neither maximally complete, nor genuinely capable of being instantiated by any parts of the world. This strategy was shown to be unsuccessful by Scott Soames (1987),

where it was demonstrated that variants of the problems posed by attitude ascriptions for standard possible worlds semantics can be re-created for any choice of truth-supporting circumstances used in formal characterizations of truth with respect to a context and a circumstance. Robert C. Stalnaker explored a different approach in 1984. After providing a naturalistic argument that semantic contents must be sets of metaphysically possible world-states, he suggested that counterexamples could be avoided by (i) allowing for exceptional cases in which attitude ascriptions report relations to propositions other than those expressed by their complement clauses, and (ii) resisting the claim that the agent believes the conjunction of p and q in many cases in which the agent believes both conjuncts. These suggestions are rebutted in later work by Mark Richard (1990) and Jeffrey Speaks (forthcoming).

STRUCTURED PROPOSITIONS, MILLIANISM, AND DESCRIPTIVISM

The problems posed by attitude ascriptions for possible worlds semantics have led many theorists to characterize the semantic content of a sentence S as a structured complex the constituents of which are the semantic contents of the semantically significant constituents of S. In essence this was also the classical position of Gottlob Frege (1892/1948) and Bertrand Russell (1905, 1910). A variant of this position, growing out of the possible worlds framework, was championed by Rudolf Carnap (1947). For Carnap, two formulas are intensionally isomorphic if and only if they are constructed in syntactically the same way from constituents with the same intensions (functions from world-states to extensions). In effect, semantic contents of syntactically simple expressions are identified with intensions, while semantic contents of syntactically complex expressions are structured complexes the constituents of which are the semantic contents of their grammatically significant parts. This view was criticized by Alonzo Church (1954), who argued that semantically complex, but syntactically simple, expressions require a stronger notion of synonymy than sameness of intension. Church's modification of Carnap—which relies on rules of sense to induce a notion of synonymous isomorphism—is a variant of the classical Fregean position.

In the late 1980s the assignment of structured semantic contents to sentences was given a neo-Russellian twist by David Kaplan (1986, 1989), Nathan Salmon (1986), and Soames (1987). On the Russellian picture structured propositions are recursively assigned to formulas, relative to contexts and assignments of values to variables. The semantic content of a variable v relative to an assignment f is just f(v), and the semantic content of a closed (directly referential) term relative to a context C is its referent relative to C. Semantic contents of n-place predicates are n-place properties and relations. The contents of truth-functional operators may be taken to be truth functions, while the semantic content of a formula ⌜λx [Fx]⌝ is identified with a propositional function g that assigns to any object o the structured proposition expressed by ⌜Fx⌝ relative to an assignment of o to "x." ⌜∃x [Fx]⌝ expresses the structured proposition in which the property of assigning a true proposition to at least one object is predicated of g. In this framework the attitude ascriptions (1a and 1b) express structured semantic contents in which the relation of asserting or believing is predicated of a pair consisting of an agent and the structured proposition semantically expressed by S. The semantic theory is completed by specifying the intensions determined by structured semantic contents, including the truth conditions of structured propositions in all possible world-states.

The signature commitment of this approach is to the possibility of asserting and believing singular propositions—which include as constituents the very objects they are about. On this approach to believe *de re* of an object that it is F is to believe the singular proposition about that object, which says that it is F. Sentences like (4), involving quantifying-in, are quintessential examples of *de re* belief ascriptions.

> (4) There is a planet x such that when the ancients saw x in the morning they believed that x was visible only in the morning and when they saw x in the evening they believed that x was visible only in the evening.

MILLIANISM. If, as Kaplan (1989) contends, the semantic contents of sentences containing indexicals are also singular propositions, then belief ascriptions containing indexicals in their complement clauses are also *de re* and hence share the basic semantic properties of ascriptions like (4). Salmon (1986) and Soames (2002) take this a step further, arguing for the Millian view that the semantic content of an ordinary proper name is simply its referent. One potentially problematic consequence of this view is that since *Ruth Barcan* and *Ruth Marcus* are coreferential, (5a) is characterized as semantically expressing the same proposition as (5b) and hence as having the same semantically determined truth value, even though it seems evident to many that it is possible to believe that

Ruth Barcan was a modal logician without believing the Ruth Marcus was:

(5a) John believes that Ruth Barcan was a modal logician.

(5b) John believes that Ruth Marcus was a modal logician.

Different Millians respond to this problem in different ways. Salmon (1986) and David Braun (2002) argue that the intuitions that (5a) and (5b) can differ in truth value are mistaken because speakers tend to confuse the identical beliefs ascribed to John by these ascriptions with the different manners of holding these beliefs associated with their different sentential complements. Soames (2002, 2005a) argues that even though the semantic contents of these sentences are the same, assertive utterances of them may indeed result in assertions of propositions with different truth values. In "A Puzzle about Belief" (1979) Saul Kripke takes a different tack. While neither advocating nor denying the Millian view, he argues that substitutivity problems of the sort illustrated here are independent of Millianism and indicate a breakdown of the basic principles underlying our belief-reporting practices.

DESCRIPTIVISM. By contrast, descriptivists, following in Frege's footsteps, have wanted to assign different semantic contents to the two names and hence to the complement clauses in (5a and 5b). The problem has been to find a way of doing this that does not run afoul of Kripke's refutation of descriptivism in *Naming and Necessity* (1972). One of Kripke's arguments holds that since names are rigid designators, their semantic contents cannot be given by any nonrigid descriptions. This argument is not easily avoided by rigidifying candidate descriptions. As shown by Soames (2002), an analysis that takes the semantic content of *Aristotle* to be given by ⌜the actual F⌝ will, all other things being equal, identify the semantic content of *Aristotle was a philosopher* with the singular proposition (about the actual world-state @) that the unique individual who "was F" in @ was also a philosopher. Assuming that the analysis also includes the standard relational treatment of belief ascriptions, one then gets the result that for any possible agent a and world-state w, ⌜x believes that Aristotle was a philosopher⌝ will be true of a with respect to w only if in w a believes that the unique individual who "was F" in @ [not w] was also a philosopher. Since this is obviously incorrect, names can neither be nonrigid descriptions, nor descriptions rigidified using the actuality operator.

What about descriptions rigidified using Kaplan's *dthat* operator? Even if, contra Kripke, a correct reference-fixing description ⌜the x: Dx⌝ could be found for each name, the semantic content of ⌜dthat [the x: Dx]⌝ would simply be its referent, in which case the descriptivist would be saddled with precisely the Millian predictions about attitude ascriptions that the theory was designed to avoid. One possible response, suggested by David Chalmers (2002), is, in effect, to take a belief ascription ⌜a believes that S⌝ to report that the belief relation holds between the agent and pair consisting of the semantic content of S (in the context) and the meaning (function from contexts to such contents) of S. However, now a different problem arises. To avoid Kripke's nonmodal arguments against familiar candidates for reference-fixing descriptions, post-Kripkean descriptivists have had to resort to egocentric, metalinguistic descriptions of the sort *the individual I have heard of under the name "n*." Although this move assigns different objects of belief to the complement clauses of (5a) and (5b), it does not solve the problem. The point, after all, is not simply to assign different belief objects in these cases, but to explain the different information one gathers about John from utterances of (5a) and (5b). As Soames (2005b) argues in *Reference and Description: The Case against Two-Dimensionalism* (2005), it is hard to see how these egocentric, metalinguistic descriptions could, realistically, contribute to this.

DAVIDSON'S LINGUISTIC VIEW

A different approach to problems involving substitutivity is to take ascriptions ⌜x says/asserts/believes that S⌝ as reporting relations either to S itself, or to a complex in which S is paired with its semantic content. Either way, since substitution of one expression for another in S always produces a new complement S', attitude ascriptions that differ in this way always report relations to different objects, whether or not the semantic contents of S and S' are the same. This encourages the thought that such ascriptions can always differ in truth value.

An early and influential version of this approach was developed by Donald Davidson (1968–1969), who argued that (6a) should be understood on the model of (6b), in which *that* is treated as a demonstrative, utterances of which refer to utterances of the independent sentence that follows it:

(6a) Galileo said that the Earth moves.

(6b) Galileo said that. The Earth moves.

On this analysis what is said by an assertive utterance u of (6a) is that one of Galileo's utterances stands in, as Davidson puts it, the samesaying relation to the subutterance u* of *the Earth moves*. Although this analysis promised a simple way of capturing the logic of attitude ascriptions, it foundered on certain recalcitrant facts, including the fact that some ascriptions, like *Every mother said that her son was lovable*, cannot be broken up into separate and independent sentences in the manner of (6b) and the fact that the assertion made by an utterance of (6a) could have been true even if the subutterance u* had never existed, indicating that the Davidsonian truth conditions are incorrect.

LATER LINGUISTIC AND REPRESENTATIONALIST VIEWS

Beginning in the 1990s improvements of Davidson's idea, including, most notably, that of Richard Larson and Peter Ludlow (1993), avoid these difficulties by dispensing with utterances and by treating attitude ascriptions as reporting relations between agents and the interpreted logical forms of their sentential complements. These are abstract, syntax-encoding structures that contain both the expressions occurring in sentences and their referential contents. Abstracting, one has here a version of the structured propositions approach in which linguistic expressions are included in the propositions sentences express. Although this version has potential virtues, it shares a crucial problem with Davidson's original analysis. Just as Davidson's silence about the intension of the samesaying relation prevented his theory from making any predictions about when (if ever) substitution of coreferential names or indexicals in a *says that* ascription changes truth value, so Larson and Ludlow's silence about the intension of the belief relation, alleged to hold between agents and interpreted logical forms, prevents their theory from making any predications about similar substitution in belief ascriptions (see Soames 2002). Since some such substitution clearly does preserve truth value, the problem is a daunting one.

Arguably, the most sophisticated approach of this general type is Richard's (1990), which combines context-sensitivity with linguistically augmented, structured Russellian propositions. For Richard, a belief ascription ⌜x believes that S⌝, used in a context C, is true of an agent a if a accepts some sentence S' with the same Russellian content in a's context as S has in C, while being similar enough to S to satisfy the belief-reporting standards in C. As indicated by Soames (2002), the evaluation of this view crucially depends on identifying similarity standards present in contexts and assessing their impact. Although there are certain evident problems here, opinions of their import vary. Finally, a different sort of context-sensitive view, advocated by Mark Crimmins and John Perry (1989), takes belief ascriptions to report that an agent believes a structured, Russellian proposition by virtue of having ideas of a certain sort—where these are mental particulars in the mind of the agent that are either implicitly demonstrated, or implicitly characterized as being of a certain type, by the one uttering the ascription. This view is usefully criticized by Jennifer Saul (1993).

EXTENSION: INTENSIONAL "TRANSITIVE" VERBS

Example (2a), in which *assert* occurs as an ordinary transitive verb operating on the extensions of its noun-phrase arguments, shows that not all attitude ascriptions contain sentential clauses. The examples in (7) show that there are also verbs, the grammatical objects of which are not overtly clausal, which are intensional in nature:

(7a) John wants a perpetual motion machine.

(7b) John is looking for the fountain of youth.

(7c) John imagined a room full of unicorns.

(7d) John worships many gods.

The relationship between these examples and ordinary propositional attitude ascriptions is a matter of ongoing investigation. How is it that (7a to 7d) can be true even though there are apparently no real entities described by their postverbal arguments? Are some or all these sentences covertly clausal? For example, are (7a) and (7b) to be assimilated to (8a) and (8b)?

(8a) John wants it to be the case that he has a perpetual motion machine.

(8b) John is trying to bring it about that he finds the fountain of youth.

These and related questions have been discussed by philosophical logicians and linguistic semanticists including Richard Montague (1974), Graeme Forbes (2000), Richard (1998), and Marcel den Dikken, Larson, and Ludlow (1997).

See also Intensional Transitive Verbs.

Bibliography

Barwise, John, and John Perry. *Situations and Attitudes.* Cambridge, MA: MIT Press, 1983.

Braun, David. "Cognitive Significance, Attitude Ascriptions, and Ways of Believing Propositions." *Philosophical Studies* 108 (2002): 65–81.

Carnap, Rudolf. *Meaning and Necessity*. Chicago: University of Chicago Press, 1947.

Chalmers, David. "The Components of Content." In *The Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press, 2002.

Church, Alonzo. "Intensional Isomorphism and Identity of Belief." *Philosophical Studies* 5 (1954): 65–73.

Crimmins, Mark, and John Perry. "The Prince and the Phone Booth: Reporting Puzzling Beliefs." *Journal of Philosophy* 86 (1989): 685–711.

Davidson, Donald. "On Saying That." *Synthese* 19 (1968–1969): 130–146.

Dikken, Marcel den, Richard Larson, and Peter Ludlow. "Intensional 'Transitive' Verbs and Concealed Complement Clauses." In *Readings in the Philosophy of Language*, edited by Peter Ludlow, 1041–1053. Cambridge, MA: MIT Press, 1997.

Forbes, Graeme. "Objectual Attitudes." *Linguistics and Philosophy* 23 (2000): 141–183.

Frege, Gottlob. "On Sense and Reference." *Zeitschrift fur Philosophie und philosophische Kritik* 100 (1892): 25–50; translated by Max Black. *Philosophical Review* 57 (1948): 207–230.

Kaplan, David. "Demonstratives." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press, 1989.

Kaplan, David. "Opacity." In *The Philosophy of W. V. Quine*, edited by Lewis Edwin Hahn and Paul Arthur Schilpp. La Salle, IL: Open Court, 1986.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980; originally published in *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman. Dordrecht, Netherlands: Reidel, 1972.

Kripke, Saul. "A Puzzle about Belief." In *Meaning and Use: Papers Presented at the Second Jerusalem Philosophical Encounter, April 1976*, edited by Avishai Margalit. Dordrecht, Netherlands: Reidel, 1979.

Larson, Richard, and Peter Ludlow. "Interpreted Logical Forms." *Synthese* 95 (1993): 305–356.

Montague, Richard. *Formal Philosophy*, edited by Richmond H. Thomason. New Haven, CT: Yale University Press, 1974.

Richard, Mark. "Commitment." *Philosophical Perspectives* 12 (1998): 255–281.

Richard, Mark. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge, U.K.: Cambridge University Press, 1990.

Russell, Bertrand. "Knowledge by Acquaintance and Knowledge by Description." *Proceedings of the Aristotelian Society* 11 (1910): 108–128.

Russell, Bertrand. "On Denoting." *Mind* 14 (1905): 479–493.

Salmon, Nathan. *Frege's Puzzle*. Cambridge, MA: MIT Press, 1986.

Saul, Jennifer. "Still an Attitude Problem." *Linguistics and Philosophy* 16 (1993): 423–435.

Soames, Scott. *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*. New York: Oxford University Press, 2002.

Soames, Scott. "Direct Reference, Propositional Attitudes, and Semantic Content." *Philosophical Topics* 15 (1) (1987): 47–87.

Soames, Scott. "Naming and Asserting." In *Semantics vs. Pragmatics*, edited by Zoltán Gendler Szabó. New York: Oxford University Press, 2005a.

Soames, Scott. *Reference and Description: The Case against Two-Dimensionalism*. Princeton, NJ: Princeton University Press, 2005b.

Speaks, Jeffrey. "The Priority of Belief: Stalnaker on Intentionality." "Is Mental Content Prior to Linguistic Meaning?" *Nous*. Forthcoming.

Stalnaker, Robert C. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Scott Soames (2005)

PROPOSITIONAL ATTITUDES: ISSUES IN THE PHILOSOPHY OF MIND AND PSYCHOLOGY

This entry aims to characterize the philosophical issues surrounding the propositional attitudes. Particular attention is paid to the arguments philosophers have brought to bear when discussing the existence and nature of the attitudes.

SUBJECT MATTER AND PHILOSOPHICAL METHODOLOGY

Discussions of the nature of mind typically distinguish between two fundamental kinds of mental states or properties. One kind of mental state or property involves states that are qualitative in nature: Examples include raw feels, sensations, tickles, and pains. The other kind of mental state or property involves states that are contentful in nature, "pointing to" or "representing" things beyond themselves: Examples include thoughts, desires, fears, and intentions. This distinction is not unproblematic, since it is not clear whether these two categories exhaust the domain of the mental, nor is it clear whether they are mutually exclusive. However, most philosophers of mind accept that there is some important distinction in this region. Propositional attitudes are often cited as the paradigmatic example of the latter kind of mental state.

As their name indicates, the propositional attitudes are *attitudes*—cognitive relations such as belief, desire, fear, hope—that a subject bears to what are typically (though not uncontroversially) taken to be *propositions*. The attitudinal component of a propositional attitude is

a matter of how a particular proposition is being taken: Thus Sally can *believe* that it will rain, *hope* that it will rain, *fear* that it will rain, and so forth. In each case the proposition that is the content of her attitude, *that it will rain*, is the same; what differs is how this proposition is being taken by Sally (believed in, hoped for, or feared). Of course, one and the same attitude can be taken towards different propositions: Thus Sam can believe that the Yankees are a great baseball team, believe that Atlanta is hot and muggy, believe that the office of the U.S. presidency has been demeaned, and so forth. In each case Sam's attitude is the same (belief), what differs is the propositions he believes.

A good deal of the attention philosophers have given to the propositional attitudes is devoted to analyzing the sentences used to ascribe the attitudes. Examples of such *attitude-ascribing sentences* include "Jones believes that it's raining," "Smith worries that State University's soccer team will lose," and "McSorley wants State University's soccer team to lose." Indeed, the very idea that propositional attitudes are cognitive relations that subjects bear to propositions (a variant of which is defended in Jerry Fodor's 1978 article, "Propositional Attitudes") is advanced on the basis of the surface grammar of the sentences used to ascribe the attitudes. Thus "Jones believes that it's raining" is naturally read as saying, of Jones (the reference of "Jones"), that he bears the belief-relation (= the reference of "believes") to the proposition *that it's raining* (= the reference of "that it's raining"). Even sentences such as "McSorley wants State University's soccer team to lose," which do not appear to refer to a proposition at all, can be translated (admittedly with some awkwardness) into equivalent sentences that do, or at least appear to, make such a reference: "McSorley wants it to be the case that State University's soccer team loses." Admittedly, though, such a propositionalist formulation may not be possible in all cases of attitude-ascribing sentences. Consider "Williams fears bats" or "Simon loves ice cream."

Whatever their ultimate nature (more on which below), the propositional attitudes themselves have been thought to be extremely important for the study of human behavior. This is seen when we consider how we go about explaining our own and others' behavior in those cases in which the behavior is taken to be intentional (falling in the domain of human action). In such cases, we explain the behavior as the effect of the subject's propositional attitudes. Thus it seems natural to explain why McSorley walked to the refrigerator in terms of her desire for cold water and her belief that cold water is to be found there; or to explain why Jackson ran away by citing his belief that a dangerous lion was coming his way and his desire not to get attacked. Explanations of this belief-desire sort are used by ordinary folk as we go about trying to predict and explain the actions of our fellows in everyday circumstances.

One philosophical question that arises in this connection concerns the status of such explanations. Suppose, as many philosophers do, that these explanations are sometimes true. What sort of explanation do they offer? Perhaps everyone can agree that they are *rationalizing* explanations, depicting the action in question as rational in light of the subject's corpus of beliefs and desires. But some philosophers hold that, in addition to rationalizing the behavior in question, they also provide a *causal* explanation of it (see Davidson 1963). If so, then the sort of psychology that appeals to the ordinary "folk" explanations of action—what has been termed *folk psychology*—can take its place beside other sciences that seek to characterize the world's causal nexus.

It is noteworthy that the causal-explanatory perspective provides an alternative approach to the nature of the attitudes, one that differs from the approach involving the analysis of attitude-ascribing sentences. Where the sententialist approach (as we might call it) assumes that we can understand the attitudes by making sense of *our talk*about them, the causal-explanatory approach begins by assuming that, whatever their ultimate nature, the propositional attitudes are the causal springs of human action. Taking the latter approach leads one to conceive of the attitudes as whatever plays the relevant causal role in the production of action. Of course, the two approaches might well be complementary: What one learns about the attitudes from analyzing attitude-ascribing sentences might be compatible with (and supplement) what one learns about the attitudes by thinking about them as the causal basis of action. (Indeed, the desire to secure the compatibility of the sententialist approach and the causal-explanatory approach appears to be a core motivation behind Fodor's 1975 hypothesis in *Language of Thought*, according to which propositional attitudes are tokenings of language-like mental symbols in the brain.) But it is also possible that the sententialist and causal-explanatory approaches will turn out to be in tension, with each one yielding some conclusions not sanctioned by, or perhaps even in conflict with, the other. Settling such a matter is perhaps the main burden of philosophical reflection on the nature of the propositional attitudes.

THE NATURE OF THE CONTENTS OF THE ATTITUDES

Common to both the sententialist and the causal-explanatory approaches to the propositional attitudes is the idea that the attitudes are *contentful* mental states. As mental states they are *about* things, typically objects and properties from the nonmental environment. Take Sanchez's belief that his grandmother smothers him with kisses. This belief is about his grandmother, him, and the property of smothering with kisses. This aspect of the propositional attitudes—their being *about* worldly states of affairs—raises a number of interesting and related philosophical issues. How does something (such as a mental state like Sanchez's belief) come to be about another thing (such as Sanchez's grandmother) in the first place? What *determines* what a mental state is about? How does the "aboutness" of mental states relate to other forms of "aboutness"? And finally, what can be said about cases in which a mental state is "about" something that does not really exist—unicorns, for example?

Philosophers have introduced the term "intentionality" to designate the domain of aboutness itself. In speaking of mental states as about the world, we are speaking of their *intentional* properties, just as in speaking of, for example, the sentence "Morty Morris has a big red wart on his nose" as about Morty Morris's big red nose wart, we are speaking of the sentence's intentional properties. Such properties are also called *semantic* properties: Both mental states and sentences—and arguably pictures, maps, models, and perhaps other things as well—have such properties. When something, such as a mental state or a sentence, has intentional or semantic properties, and so is about something, we can speak of what the state is about as *the content* of that state. Talk of the content of a mental state is to be understood in terms of what the mental state *represents* as the case. So Sanchez's belief has a content, which is what that belief represents to be the case: namely, that his grandmother smothers him with kisses.

It is noteworthy that a belief can represent something that is not the case. Suppose that Sanchez's grandmother does *not*, in fact, smother him with kisses (it's all "in his head," so to speak). Then, supposing there is an inventory of all of the facts that make up our world, we would not find in this inventory any fact to the effect that Sanchez's grandmother smothers him with kisses. In short, there is no fact that is represented by his belief. But then what is this shadowy thing we are calling the content of his belief, that which his belief represents to be the case? Above we called this content a "proposition," and we can now see the attraction of the view that the content of an attitude is a proposition. For although it is hard to say exactly what propositions are, we can say at least this much: The existence of a given proposition does not depend on the existence of the corresponding fact that would make the proposition true. That is, there can be *false* propositions. Given that Sanchez's grandmother is not as Sanchez's belief depicts her, the proposition that is the content of Sanchez's belief is itself a false proposition.

The postulation of the proposition as the content of the attitudes raises a bundle of related metaphysical questions. What is the nature of propositions? (Is it essentially a linguistic entity? an abstract one? a mental one?) Do propositions have parts, and if so, what is the nature of those parts? Here we focus on a question bearing more directly on the philosophy of mind: How do propositional attitudes come to have the propositional content they have? More concretely, what makes Sanchez's belief a belief about his grandmother, and not, say, about ice cream sundaes or pink elephants or any of an infinite number of other things? Let us address this by asking which facts *fix* the content of his belief: Which facts are such that, if you fix them, then, no matter what else is going on in the universe, you have fixed the content of his belief that his grandmother smothers him with kisses? A natural first guess would be that the facts in question are facts regarding the mental image(s) in Sanchez's mind at the time that he calls this belief to mind. On such a view, once we fix the mental image(s) "in" his mind, we have fixed what his belief is about.

But this cannot be quite right. First, mental images do not appear to have the right sort of specificity to fix the content of the propositional attitudes. To see this, imagine a scenario in which Sanchez's grandmother has an identical twin, from whom the grandmother herself is indistinguishable, but whom Sanchez has never met or otherwise heard of. Then the image in Sanchez's mind "fits" his grandmother's twin as much as it "fits" his grandmother. But it seems implausible to think that his belief is about the twin, for he has never met or heard of her. Second, in addition to not having the right sort of specificity, mental images are too unstable and subjective to fix the contents of one's attitudes. This is clearest in cases in which the subject matter of the attitude is an abstract one. Precisely what image goes before your mind when you call forth your belief that 1+1=2? And what image is before your mind when you believe that space is (or is not) infinite? Will it be true that any two people who believe e.g. that 1+1=2 will have the same type of image before their minds? Presumably not. But then how

does the image fix the content of their belief? It seems that what they have in common, in virtue of which they both count as believing that 1+1=2, is something other than a particular type of image.

And the same point can be made even in cases in which the subject matter of the attitude is not abstract. Take Roger's belief that Morty Morris has a big red wart on his nose. Since Morty Morris is Roger's best friend, Roger has a vivid mental image of Morty (wart and all). But Mathilde, who (having been told by Roger) also believes that Morty Morris has a big red wart on his nose, has never met Morty, and therefore has no such image. Again it seems that what Roger and Mathilde have in common, in virtue of which they both count as believing that Morty Morris has a big red wart on his nose, is something other than a particular type of mental image in mind.

These arguments (and the examples on which they are based) raise a host of issues regarding how propositional attitudes come to have the propositional content they have. Consider first the relation between such contents and the environment in which one lives and interacts. One plausible account of why Sanchez's belief is about his grandmother, rather than her identical twin, is that his belief was caused and sustained by activities involving one woman and not the other. So it can seem that interaction with one's environment is relevant to the determination of the contents of one's attitudes. Next consider the relationship between language and the content of the attitudes. Recalling that mental images are too unstable and subjective to fix the contents of attitudes, we might ask: Precisely what do Roger and Mathilde have in common, in virtue of which they both count as believing that Morty Morris has a big red wart on his nose? At least part (but only part!) of the answer is that they are both disposed to accept and assert a sentence that *means* that Morty Morris has a big red wart on his nose. Perhaps, then, among the facts that fix the content of one's attitudes we must include facts regarding the meanings of one's words.

These conclusions highlight one of the bigger controversies in the theory of content. In particular, we have seen at least three types of fact that might be regarded as relevant to fixing the content of one's attitudes. We started off with the suggestion that facts regarding the subject's mental images fix the content of her attitudes, but we moved quickly to include facts regarding causal history and then on to facts regarding the meanings of one's words. These correspond roughly to three distinct theoretical options available with respect to the sort of facts needed to fix the content of one's attitudes.

Content internalism is the view that the only facts needed to fix the content of a subject's attitudes are facts that do not presuppose the existence of anything beyond the subject herself. The view with which we started, according to which the facts regarding the subject's mental images fix the contents of her attitudes, is one version of content internalism. But the content internalist can allow other sorts of facts, so long as these do not presuppose the existence of anything beyond the subject herself; and the most plausible versions of content internalism (for which see Searle 1983) include facts about the individual's use of language, where the meanings of her words are not thought to depend on the existence of anything beyond the subject herself. Of the various arguments for content internalism, one of the most influential is what we might call the argument from "intentional inexistence." Consider, to begin, that one can form a belief which is "about" something that does not exist—as with Ponce de Leon's belief that the Fountain of Youth is in Florida, or Roger's belief that the largest natural number is even. What is more, it would seem possible (though of course highly unlikely) that none of our beliefs succeed in being about any existing thing: Perhaps you are suffering an eternal and systematic hallucination in a world containing nothing but your own mind! But in that case, although your beliefs remain the same (or so it might seem), there are no worldly objects for them to be "about." This suggests that the "aboutness" properties of beliefs should be understood in such a way as not to presuppose the existence of anything beyond the thinking subject.

Many philosophers, unconvinced by this sort of argument, have thought that the internalist view is too restrictive in the set of facts it regards as relevant to fixing the content of the attitudes. A second view, *content individualism*, expands the set of content-fixing facts to include not just the facts allowed by the content internalist, but also any facts regarding the thinker's own causal history. (See Davidson 1984 and 2001 for an example of a view that combines *content externalism*, which is the denial of content internalism, with content individualism.) Although the cost of moving from internalism to individualism is that of having to rebut the argument from intentional inexistence—something that forces the individualist to come up with an account of beliefs "about" non-existent "objects"—the payoff of making this move can be made clear in connection with the following development of Sanchez's case. Sanchez

has an identical twin, Twin-Sanchez, separated from Sanchez from birth. Twin-Sanchez has interacted only with twin-granny, the identical twin of Sanchez's grandmother. Further, the course of experience Sanchez has with his grandmother is internally indistinguishable from the course of experience Twin-Sanchez has with twin-granny. So, for example, at the very moment Sanchez sees his grandmother wearing a lovely purple vest and making waving motions as she smiles, Twin-Sanchez sees twin-granny wearing an indistinguishable lovely purple vest making waving motions as *she* smiles; at the very moment Sanchez hears his grandmother singing a lovely melody, Twin-Sanchez hears twin-granny singing an indistinguishable lovely melody; and so forth through time. At one point each of the Sanchez twins, admiring the grandmother in his presence, forms a belief he would express with, "She has a wonderful voice." The natural view is that the contents of their beliefs differ: Sanchez's belief represents *his* grandmother (not twin-granny) as having a wonderful voice, whereas twin-Sanchez's belief represents twin-granny (not granny) as having a wonderful voice.

The content individualist can easily accommodate this natural view, as the difference in content can be fixed by the facts regarding each twins' causal history (with distinct grannies). The content internalist, by contrast, will have trouble accepting the natural view: Since the twins' course of experiences are internalistically indistinguishable, there will be some pressure on the content internalist to treat the twins as having beliefs *with the very same content*. (See Searle 1983 for an attempt by an internalist to avoid this conclusion.)

But if the content individualist has this virtue over the internalist, some philosophers have felt that individualism does not go far enough. A third position, which we might designate as *content anti-individualism*, is still more liberal in the range of facts it regards as relevant to fixing the content of a subject's attitudes. As its name suggests, content anti-individualism is the denial of content individualism. But it is helpful to see why a theorist might deny that "individualistic" facts suffice to fix the content of a subject's attitudes. The controversy has to do with the role of language in fixing the content of the attitudes. In one sense, it is uncontroversial that the meaning of one's words determines the contents of the attitudes one expresses with those words. The controversial matter regards what determines the meaning of one's words. The individualist maintains that no facts beyond those regarding the individual speaker herself—the conditions under which she uses her words, how she herself explicates their meanings—are needed to fix the meaning of her words; whereas the anti-individualist maintains that these "individualistic" facts do not suffice to fix the meanings of her words. The insight (or alleged insight) behind anti-individualism is that individual language users typically defer to, and take themselves to be answerable to, public standards of correct usage. Such standards are not typically fixed by the individual's *own* word usage or meaning-explications, but instead are fixed by the usage of other speakers (Kripke 1972) and the meaning-explicating practices of the relevant experts in her linguistic community (Putnam 1975 and Burge 1979).

Interestingly, the sententialist and causal-explanatory approaches to the attitudes bear on the debate regarding the nature of mental content. For example, among the reasons offered in defense of anti-individualism, Burge notes in "Individualism and the Mental" (1979) that variations in public standards for the correct use of a word lead to differences in *the belief-attributing sentences* that would be used to report a subject's beliefs. And among the reasons offered in defense of content internalism are considerations pertaining to the internal basis of mental causes (for which see Fodor 1980). Although neither argument is decisive, each suggests the core motivations for and potential liabilities of the various positions on mental content.

THE METAPHYSICS OF THE ATTITUDES: VERSIONS OF MATERIALISM

The question regarding the nature of mental content cannot be addressed in isolation from what we might call the metaphysics of the attitudes. What is the nature of the states and properties dubbed "the propositional attitudes"? How do such states and properties relate to the thinker's bodily states and properties? These questions, of course, force us to confront a particular version of the notorious mind-body problem.

The positions that can be taken on the relation between a subject's propositional attitudes and her bodily states and properties correspond to positions familiar from the general mind-body problem. *Attitude dualism* holds that propositional attitudes are *immaterial* states or properties of thinking subjects. But as with dualism generally, attitude dualism runs into trouble in connection with the causal role that the attitudes are thought to play: How do immaterial states or properties affect a subject's body? Most contemporary philosophers take some version of this problem to be decisive against dualism. And of these most go on to endorse *materialism*, according to

which all of the objects and properties of our world are material in nature. So we will restrict our discussion accordingly.

Among materialist views we can begin with the view known as *philosophical behaviorism*, according to which the so-called propositional attitudes really are nothing other than complex behavioral dispositions. Philosophical behaviorism itself (unlike psychological behaviorism) was originally motivated by the verification theory of meaning, according to which the meaning of a sentence consists in the conditions whose obtaining would *verify* the sentence (establish its truth). Since sentences such as "John believes that it's raining" are typically regarded as true or false in virtue of observable behavior (e.g., John's uttering "It's raining!," carrying an umbrella with him, putting on galoshes, and so forth), the result of applying the verification theory of meaning to attitude-ascribing sentences is that each such sentence is to be regarded as equivalent in meaning to a "behavioral translation," a much longer sentence describing all of the behaviors and behavioral dispositions whose presence would verify the original sentence (see e.g. Ryle 1949). However, this view faces two obvious and devastating difficulties.

First, as noted in Putnam "Brains and Behavior" (1965), the view is either false or unacceptably circular. It is false if the translation of the target sentence ("John believes that it's raining") fails to capture all of the conditions whose presence would be taken as evidence for the truth of that sentence. But in order to avoid falsity on this score, the translation will need to make reference to other attitudes the subject has: For example, John's uttering "It's raining!" counts for the truth of "John believes that it is raining" only if he is speaking sincerely and believes that "It's raining" means that it's raining; John's taking an umbrella with him (or putting galoshes on) counts for the truth of "John believes that it's raining" only if he desires not to get wet and believes that the umbrella (galoshes) will prevent him from getting wet; and so forth. In fact, it would appear that the connection between attitudes and behavior invariably involves other attitudes in this way. But in that case, any attempt to translate a target attitude-ascribing sentence will yield a translation which itself contains mention of other attitudes. On pain of circularity, these latter attitude-ascribing components in the translation must also be translated. But then the problem begins again, and the whole approach appears doomed to an unacceptable sort of circularity.

Nor is this philosophical behaviorism's only problem. A second objection is that philosophical behaviorism surrenders the idea of propositional attitudes as *the causes* of behavior. Consider: that a sugar cube dissolves in water is the basis for regarding it as water-soluble; so it would be an empty explanation to regard its solubility in water as *the cause* of its dissolving on a particular occasion. (Compare the doctor spoofed in Molière's play *La malade imaginaire*: He explained the sleep-producing character of a particular drug to its having a "dormative virtue.") Similarly, if beliefs and desires are dispositions to act, then it would be an empty explanation indeed to regard beliefs and desires as *the causes* of action.

Given the failure of philosophical behaviorism, the desire to preserve the causal profile of the propositional attitudes within a materialist framework provided the main motivation behind *identity theory*. Recognizing the role of the attitudes as the causes of intelligent behavior, early identity theorists used the fact (or what they regarded as the fact) that the causes of intelligent behavior are to be found in the states and processes of the central nervous system, to conclude that the propositional attitudes are *identical to* those states and processes of the central nervous system. (For an early formulation of identity theory, albeit in connection with sensory rather than contentful states, see Smart 1962.) The proposed identity was between property-*types*: An "attitudinal" property-type (such as the property of *believing that it is raining*) was held to be identical to a property-type instantiated by the central nervous system (such as the property of *having such-and-such a pattern of neural activation in this-or-that region of the brain*). But this gave rise to an objection from the so-called multiple-realizability of mental states (Putnam 1967): On the assumption that creatures whose underlying neurophysiology is very different from our own might nevertheless be regarded as being the subjects of attitudes, such type-identity claims were much too strong.

Such an objection to type-identity theory acquired additional force in light of the development of sophisticated forms of artificial intelligence. Alan Turing's famous "Turing Test" (1950) taught that a system was to be regarded as "intelligent" so long as it behaved in a way that would lead those with whom it interacted to regard it as intelligent. The implicit idea was that any system with the right sort of *functional complexity*—as seen in its capacity to acquire and process information from its environment and to use this information to guide its subsequent actions—was to be regarded as intelligent (and hence, given some plausible assumptions, as a subject of the propositional attitudes). The result was what is perhaps the most widely accepted view of the metaphysics of

the attitudes: *functionalism*. According to the functionalist, propositional attitudes are best characterized by their functional or causal profile. So just as it would be a mistake to identify the property of *being a carburetor* with the property of *being made of metal and shaped in such-and-such a way*—surely being a carburetor is more a matter of function rather than material—so too it would be a mistake to identify the property of *believing that it is raining* with some particular property of the body. Rather, a subject has this property, and so counts as believing that it is raining, when the subject is in a state with a certain functional or causal profile—one that is caused in certain characteristic sorts of ways (e.g., seeing rain) and interacts with other functionally defined states to bring about certain effects (e.g., producing utterances of "It's raining!," movements to retrieve the umbrella when leaving, and so forth).

According to the functionalist, the first task in connection with the metaphysics of the attitudes is to specify the functional role corresponding to each distinct attitude. Once that task is completed, the functionalist philosopher can then pass on to empirical investigation the task of identifying what particular physical property *realizes* that functional role in a given system. (Compare: Once the functional role of a carburetor has been specified, we can go on and ask which feature of a particular car realizes that role.) Such a view is often advanced as part of a "computational" theory of the attitudes, according to which the functional role of particular mental states is best understood in information-processing terms. So formulated, functionalism, as shown by A. Newell (1980) and David Marr (1982), has been popular not only in the philosophy of mind but also in traditional cognitive science.

Of course, having specified what we take to be the functional role of a particular attitude-type (say, the belief that it is raining), there is no guarantee that there will *be* any state or property of the body or the central nervous system playing that role. Perhaps the very idea that there is a state playing that role is itself part of a mistaken theory of the mind, one whose fundamental postulates (beliefs, desires, and so forth) are as misguided as was the postulation of witches and other spiritual entities by misguided theorists of earlier ages. A number of philosophers, such as P. Churchland (1981), have begun to express such misgivings, arguing that the account of mind which postulates propositional attitudes is part of a "worm-eaten myth" that will be replaced as brain science progresses. Such a view, known as *eliminative materialism*, is perhaps the starkest version of materialism there is, as it combines a general commitment to materialism with the view that there is nothing in the material world that answers to what we take the propositional attitudes to be. Though clearly radical, such a view has challenged mainstream theorists to further clarify what is at issue in the debate over the propositional attitudes.

See also Belief; Belief Attributions; Content, Mental; Intentionality; Language of Thought.

Bibliography

Burge, Tyler. "Individualism and the Mental." In *Studies in Metaphysics*, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. 73–121. Midwest Studies, vol. 4. Minneapolis: University of Minnesota Press, 1979. Reprinted *The Nature of Mind*, edited by David M. Rosenthal. Oxford: Oxford University Press, 1991.

Churchland, P. "Eliminative Materialism and the Propositional Attitudes." *The Journal of Philosophy* 78 (2) (1981): 67–90.

Davidson, Donald. "Actions, Reasons, and Causes." [1963] Reprinted in *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.

Davidson, Donald. *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press, 1984.

Davidson, Donald. *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press, 2001.

Fodor, Jerry. *The Language of Thought*. New York: Thomas Y. Crowell, 1975.

Fodor, Jerry. "Methodological Solipsism considered as a Research Strategy in Cognitive Psychology." *The Behavioral and Brain Sciences* 3 (1) (1980): 63–72. Reprinted in *The Nature of Mind*, edited by David M. Rosenthal. Oxford: Oxford University Press, 1991.

Fodor, Jerry. "Propositional Attitudes." *The Monist* 61 (4) (1978): 501–23. Reprinted in *The Nature of Mind*, edited by David M. Rosenthal. Oxford: Oxford University Press, 1991.

Kripke, Saul A. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Marr, David. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman, 1982.

Newell, A. "Physical Symbol Systems." *Cognitive Science* 4 (2) (1980): 135–183.

Putnam, Hilary. "Brains and Behavior." In *Analytical Philosophy, Second Series*, edited by R. J. Butler. Oxford: Basil Blackwell and Mott, 1965.

Putnam, Hilary. "The Meaning of 'Meaning'." In his *Philosophical Papers*. Vol. 2: *Mind, Language, and Reality*. Cambridge, U.K.: Cambridge University Press, 1975.

Putnam, Hilary. "The Nature of Mental States." [1967] Reprinted *The Nature of Mind*, edited by David M. Rosenthal. Oxford: Oxford University Press, 1991.

Ryle, Gilbert. *The Concept of Mind*. Chicago: University of Chicago Press, 1949.

Searle, John R. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, U.K.: Cambridge University Press, 1983.

Smart, J. J. C. "Sensations and Brain Processes." [1962] Reprinted in *The Nature of Mind*, edited by David M. Rosenthal. Oxford: Oxford University Press, 1991.

Turing, Alan. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433–460.

Sanford Goldberg (2005)

PROPOSITIONAL KNOWLEDGE, DEFINITION OF

The traditional "definition of propositional knowledge," emerging from Plato's *Meno* and *Theaetetus*, proposes that such knowledge—knowledge that something is the case—has three essential components. These components are identified by the view that knowledge is justified true belief. Knowledge, according to the traditional definition, is belief of a special kind, belief that satisfies two necessary conditions: (1) the truth of what is believed and (2) the justification of what is believed. While offering various accounts of the belief condition, the truth condition, and the justification condition for knowledge, many philosophers have held that those three conditions are individually necessary and jointly sufficient for propositional knowledge.

The belief condition requires that one accept, in some manner, any proposition one genuinely knows. This condition thus relates one psychologically to what one knows. It precludes that one knows a proposition while failing to accept that proposition. Some contemporary philosophers reject the belief condition for knowledge, contending that it requires a kind of mentalistic representation absent from many cases of genuine knowledge. Some other contemporary philosophers endorse the belief condition but deny that it requires actual assent to a proposition. They propose that, given the belief condition, a knower need only be disposed to assent to a proposition. Still other philosophers hold that the kind of belief essential to propositional knowledge requires assent to a known proposition, even if the assent need not be current or ongoing. The traditional belief condition is neutral on the exact conditions for belief and for the objects of belief.

The truth condition requires that genuine propositional knowledge be factual, that it represent what is actually the case. This condition precludes, for example, that astronomers before Nicolas Copernicus knew that Earth is flat. Those astronomers may have believed—even justifiably believed—that Earth is flat, as neither belief nor justifiable belief requires truth. Given the truth condition, however, propositional knowledge without truth is impossible. Some contemporary philosophers reject the truth condition for knowledge, but they are a small minority. Proponents of the truth condition fail to agree on the exact conditions for the kind of truth essential to knowledge. Competing approaches to truth include correspondence, coherence, semantic, and redundancy theories, where the latter theories individually admit of variations. The truth condition for knowledge, generally formulated, does not aim to offer an exact account of truth.

The justification condition for propositional knowledge guarantees that such knowledge is not simply true belief. A true belief may stem just from lucky guesswork; in that case it will not qualify as knowledge. Propositional knowledge requires that the satisfaction of its belief condition be suitably related to the satisfaction of its truth condition. In other words, a knower must have adequate indication that a belief qualifying as knowledge is actually true. This adequate indication, on a traditional view of justification suggested by Plato and Immanuel Kant, is suitable evidence indicating that a proposition is true. True beliefs qualifying as knowledge, on this traditional view, must be based on justifying evidence.

Contemporary philosophers acknowledge that justified contingent beliefs can be false; this is fallibilism about epistemic justification, the kind of justification appropriate to propositional knowledge. Given fallibilism, the truth condition for knowledge is not supplied by the justification condition; justification does not entail truth. Similarly, truth does not entail justification; one can lack evidence for a proposition that is true.

Proponents of the justification condition for knowledge do not share an account of the exact conditions for epistemic justification. Competing accounts include epistemic coherentism, which implies that the justification of any belief depends on that belief's coherence relations to other beliefs, and epistemic foundationalism, which implies that some beliefs are justified independently of any other beliefs. Recently, some philosophers have proposed that knowledge requires not evidence but reliable (or truth-conducive) belief formation and belief sustenance. This is reliabilism about the justification condition for knowledge. Whatever the exact conditions for epistemic justification are, proponents of the justification condition maintain that knowledge is not merely true belief.

Although philosophers have not agreed widely on what specifically the defining components of proposi-

tional knowledge are, there has been considerable agreement that knowledge requires, in general, justified true belief. Traditionally, many philosophers have assumed that justified true belief is sufficient as well as necessary for knowledge. This is a minority position now, owing mainly to Gettier counterexamples to this view. In 1963 Edmund Gettier challenged the view that if one has a justified true belief that *p*, then one knows that *p*. Gettier's counterexamples are:

(I) Smith and Jones have applied for the same job. Smith is justified in believing that (i) Jones will get the job, and that (ii) Jones has ten coins in his pocket. On the basis of (i) and (ii), Smith infers, and thus is justified in believing, that (iii) the person who will get the job has ten coins in his pocket. As it turns out, Smith himself will actually get the job, and he also happens to have ten coins in his pocket. So, although Smith is justified in believing the true proposition (iii), Smith does not know (iii).

(II) Smith is justified in believing the false proposition that (i) Jones owns a Ford. On the basis of (i), Smith infers, and thus is justified in believing, that (ii) either Jones owns a Ford or Brown is in Barcelona. As it turns out, Brown is in Barcelona, and so (ii) is true. So although Smith is justified in believing the true proposition (ii), Smith does not know (ii).

Gettier counterexamples are cases where one has a justified true belief that *p* but lacks knowledge that *p*. The Gettier problem is the difficulty of finding a modification of, or an alternative to, the traditional justified-true-belief analysis that avoids difficulties from Gettier counterexamples.

Contemporary philosophers have not reached a widely accepted solution to the Gettier problem. Many philosophers take the main lesson of Gettier counterexamples to be that propositional knowledge requires a fourth condition, beyond the justification, belief, and truth conditions. Some philosophers have claimed, in opposition, that Gettier counterexamples are defective because they rely on the false principle that false evidence can justify one's beliefs. There are, however, examples similar to Gettier's that do not rely on any such principle. Here is one such example inspired by Keith Lehrer and Richard Feldman:

(III) Suppose that Smith knows the following proposition, *m*: Jones, whom Smith has always found to be reliable and whom Smith has no reason to distrust now, has told Smith, his officemate, that *p*: He, Jones, owns a Ford. Suppose also that Jones has told Smith that *p* only because of a state of hypnosis Jones is in and that *p* is true only because, unknown to himself, Jones has won a Ford in a lottery since entering the state of hypnosis. Suppose further that Smith deduces from *m* its existential generalization, *o*: There is someone, whom Smith has always found to be reliable and whom Smith has no reason to distrust now, who has told Smith, his officemate, that he owns a Ford. Smith, then, knows that *o*, since he has correctly deduced *o* from *m*, which he also knows. Suppose, however, that on the basis of his knowledge that *o*, Smith believes that *r*: Someone in the office owns a Ford. Under these conditions, Smith has justified true belief that *r*, knows his evidence for *r*, but does not know that *r*.

Gettier counterexamples of this sort are especially difficult for attempts to analyze the concept of propositional knowledge.

One noteworthy fourth condition consists of a "defeasibility condition" requiring that the justification appropriate to knowledge be "undefeated" in that an appropriate subjunctive conditional concerning defeaters of justification be true of that justification. A simple defeasibility condition requires of our knowing that *p* that there be no true proposition, *o*, such that if *q* became justified for us, *p* would no longer be justified for us. If Smith genuinely knows that Laura removed books from the office, then Smith's coming to believe with justification that Laura's identical twin removed books from the office would not defeat the justification for Smith's belief regarding Laura herself. A different approach claims that propositional knowledge requires justified true belief sustained by the collective totality of actual truths. This approach requires a precise, rather complex account of when justification is defeated and restored.

The importance of the Gettier problem arises from the importance of a precise understanding of the nature, or the essential components, of propositional knowledge. A precise understanding of the nature of propositional knowledge, according to many philosophers, requires a Gettier-resistant account of knowledge.

See also Coherentism; Epistemology; Kant, Immanuel; Plato; Reliabilism; Truth.

Bibliography

Audi, Robert. *Epistemology*. London: Routledge, 1998.

BonJour, L. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.
Carruthers, P. *Human Knowledge and Human Nature*. New York: Oxford University Press, 1992.
Chisholm, Roderick. "Knowledge as Justified True Belief." In *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982.
Chisholm, Roderick. *Theory of Knowledge.* 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1989.
Goldman, A. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.
Greco, John, and Ernest Sosa, eds. *The Blackwell Guide to Epistemology*. Oxford: Blackwell, 1998.
Klein, Peter. "A Proposed Definition of Propositional Knowledge." *Journal of Philosophy* 67 (1971): 471–482.
Lehrer, K. *Theory of Knowledge*. Boulder, CO: Westview Press, 1990.
Lewis, C. I. *An Analysis of Knowledge and Valuation*. La Salle, IL: Open Court, 1946.
Moser, Paul. "Epistemology (1900–Present)." In *Routledge History of Philosophy.* Vol. 10: *Philosophy of the English Speaking World in the 20th Century*, edited by John Canfield. London: Routledge, 1996.
Moser, Paul. *Knowledge and Evidence*. New York: Cambridge University Press, 1989.
Moser, Paul, ed. *The Oxford Handbook of Epistemology*. Oxford: Oxford University Press, 2002.
Moser, Paul, and A. vander Nat, eds. *Human Knowledge: Classical and Contemporary Approaches.* 2nd ed. New York: Oxford University Press, 1995.
Pollock, J. *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman & Littlefield, 1986.
Shope, Robert. *The Analysis of Knowing: A Decade of Research.* Princeton, NJ: Princeton University Press, 1983.
Sosa, E. *Knowledge in Perspective*. New York: Cambridge University Press, 1991.

Paul K. Moser (1996)
Bibliography updated by Benjamin Fiedor (2005)

PROPOSITIONS

On one use of the term, "propositions" are objects of assertion, what successful uses of declarative sentences say. As such, they determine truth-values and truth conditions. On a second, they are the objects of certain psychological states (such as belief and wonder) ascribed with verbs that take sentential complements (such as *believe* and *wonder*). On a third use, they are what are (or could be) named by the complements of such verbs. Many assume that propositions in one sense are propositions in the others.

After some decades of skepticism about the worth of positing propositions, the last quarter of the twentieth century saw renewed interest in and vigorous debate over their nature. This can be traced in good part to three factors: the development in intensional logic of formal models of propositions; (not altogether unrelated) attacks on broadly Fregean accounts of propositions; and a spate of work on the nature of belief and its ascription.

"Possible-worlds semantics" is a collection of methods for describing the semantical and logical properties of expressions such as *necessarily*; these methods developed out of work done by Saul Kripke, Richard Montague, and others in the 1960s. It illuminated the logic and semantics of modal terms such as *necessarily,* of conditionals and tenses, and other constructions as well. In such semantics one assigns a sentence a rule that determines a truth-value relative to various "circumstances of evaluation" (possible worlds, times, whatever); a sentence such as "it is necessary that S" has its truth-value determined by the rule so associated with S. The success of such accounts made it natural to hypothesize that propositions, qua what is named by expressions of the form "that S," could be identified with such rules—equivalently, with sets of circumstances such rules pick out.

Such a conception of proposition provides too crude an account of objects of belief or assertion: It implausibly makes all logically equivalent sentences express the same belief and say the same thing. A partial solution to this problem supposes that propositional identity is partially reflected in sentential structure, taking propositions themselves to be structured. Given the working hypothesis that a proposition's structure is that of sentences expressing it, critical to determining the proposition a sentence (use) expresses are the contributions made by sentence parts (on that use).

Gottlob Frege (1952) suggested that associated with names and other meaningful expressions are "ways of thinking" or *senses* of what the expressions pick out; one might suppose that sense and sentence structure jointly determine proposition expressed. Sense, in the case of names and other singular terms, has standardly been taken to be given by describing how one thinks of the referent. For example, the sense of "Aristotle" for me might be given by "the author of the *Metaphysics*"; if so, my uses of "Aristotle taught Alexander" and "the author of the *Metaphysics* taught Alexander" would, on a Fregean view, express the same proposition.

During the 1970s Kripke, David Kaplan, and others argued convincingly that this view is untenable: It is obvious, on reflection, that the truth conditions of the assertion or belief that Aristotle was *F* depend on Aristotle in a way in which the truth conditions of the assertion or belief that the author of the *Metaphysics* was *F* do not. So either ways of thinking are somehow tied to the objects

they present (so that the way I think of Aristotle could not present anything but Aristotle), or the contributions of expressions to propositions must be something other than senses.

The success of accounts of intensional language that ignored sense in favor of constructions from references, along with the apparent failure of Fregean accounts, led in the 1980s to debate over the merits of what is variously called direct-reference theory, Millianism, and (neo-) Russellianism, espoused at various times by a wide variety of theorists including Kaplan (1989), Mark Richard (1990), Nathan Salmon, and Scott Soames (1988). On such views sense is irrelevant to individuating a proposition; indeed, it is irrelevant to semantics. In particular, what a name contributes to a proposition is its referent: The proposition that Twain is dead is the same singular proposition as the proposition that Clemens is.

Neo-Russellians identify the object of assertion and the referent of a "that" clause with a Russellian proposition. They allow that there is such a thing as a "way of grasping" a proposition and that belief in a singular proposition is mediated by such. Against the intuition that, for example, A: Mo believes that Twain is dead, and B: Mo believes that Clemens is dead, might differ in truth-value, direct-reference accounts typically suggest that a pragmatic explanation is appropriate. Just as an ironic use of a sentence can convey a claim without literally expressing it, so a sentence about Mo's beliefs might convey information about Mo's way of grasping a singular proposition, without that information being part of what the sentence literally says. If this is so, intuitions about A and B are explained pragmatically.

Those unhappy with this account of propositions have looked elsewhere. Many accounts of propositions identify the proposition determined by S with some construction from linguistic items associated with S and the semantic values of S's parts. James Higginbotham has identified the referents of "that" clauses with phrase markers that may be annotated with referents; Richard has suggested that the referent of a "that" clause be identified with something like the singular proposition it determines paired off with the sentence itself. In making linguistic items constitutive of propositions, these views run counter to ones, like Frege's and Bertrand Russell's, that closely tie meaning and synonymity to propositional determination. On linguistic views of propositions the synonymity of *groundhog* and *woodchuck* does not assure the identity of the proposition that groundhogs are pests with the proposition that woodchucks are. Other theorists (Gareth Evans, for example) have attempted to revive a version of Frege's views of propositions.

Many philosophers continue to doubt the utility of positing propositions. Quineans argue that meaning and reference must be determined by behaviorally manifest facts but that such facts woefully underdetermine assignments of meaning and reference; they conclude that there is nothing about language that need or could be explained by positing propositions. Stephen Schiffer has argued that propositions are a sort of "linguistic posit": that we accept nominalizations of the form "that S" as referring to singular terms and have coherent criteria for using sentences in which those terms occur is itself sufficient for its being true that there are propositions. Such a deflationist view implies neither the possibility of a substantive account of propositions (on which, for example, the proposition expressed by a sentence is compositionally determined), nor that propositions play a substantive role in explaining semantic phenomena.

See also Frege, Gottlob; Kripke, Saul; Meaning; Modality, Philosophy and Metaphysics of; Philosophy of Language; Quine, Willard Van Orman; Reference; Russell, Bertrand Arthur William.

Bibliography

Evans, G. *The Varieties of Reference.* Oxford: Clarendon Press, 1982.

Frege, G. "Uber Sinn und Bedeutung." *Zeitschrift fur Philosophie and Philosophische Kritik* 100 (1892): 25–50. Translated by P. Geach and M. Black as "On Sense and Reference." In *Translations from the Philosophical Writings of Gottlob Frege,* edited by P. Geach and M. Black. New York: Philosophical Library, 1952.

Horwich, P. *Truth.* Cambridge, MA: Blackwell, 1990.

Kaplan, D. "Demonstratives." In *Themes from Kaplan,* edited by J. Perry et al. New York: Oxford University Press, 1989.

Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.

Lewis, D. "Attitudes *de dicto* and *de se.*" *Philosophical Review* 88 (1979): 513–543.

Montague, R. *Formal Philosophy,* edited by R. Thomason. New Haven, CT: Yale University Press, 1974.

Quine, W. V. O. *Word and Object.* Cambridge, MA: MIT Press, 1960.

Richard, M. *Propositional Attitudes.* Cambridge, U.K.: Cambridge University Press, 1990.

Salmon, N., and S. Soames, eds. *Propositions and Attitudes.* Oxford: Oxford University Press, 1988.

Schiffer S. "A Paradox of Meaning." *Noûs* 28 (1994): 279–324.

Stalnaker, R. *Inquiry.* Cambridge, MA: MIT Press, 1984.

Mark E. Richard (1996)

PROPOSITIONS [ADDENDUM]

Despite the rearguard efforts of Robert Stalnaker and Max Cresswell, by the late 1990s it became widely acknowledged that sets of possible worlds are too coarse-grained to serve as propositions. It is safe to say that among those philosophers who believe in propositions, most think of them as sententially structured entities, composed out of the contents of the words and phrases in the sentences that express them. Fregeans hold that these contents are Fregean senses; Russellians hold that they are objects, properties, and relations.

Yet the 1990s also saw new challenges and approaches to structured propositions. George Bealer and Michael Jubien have independently argued (i) that it is counterintuitive to hold that we believe and assert structured complexes, and (ii) that theories of structured propositions are subject to the same problem that Paul Benacerraf raised for set-theoretic reductions of arithmetic. On one such reduction, the number 2 is identified with the set {{∅}}; on another, 2 is identified with {∅, {∅}}, where ∅ is the null set. Benacerraf's problem is that there are no principled reasons for preferring one or the other reduction, or any of the infinitely many equally good alternatives, and so none of these reductions can be correct. For similar reasons, the proposition that Jones loves Smith cannot be identified with the ordered set ⟨love, ⟨Jones, Smith⟩⟩, or with ⟨Jones, ⟨love, Smith⟩⟩, or ….

These and other problems led Bealer to reject all reductions of propositions to structured objects and to hold that propositions are unstructured and irreducible. They led Jubien to reject propositions altogether in favor of a Russellian multiple-relation theory of judgment, which dispenses with propositions by analyzing "believes" and other attitude verbs as many-place predicates that relate subjects to objects, properties, and relations instead of to whole propositions. However, (i) is debatable, and (ii) can be avoided if one can provide a rationale for preferring one system of reduction. For example, Jeffrey King holds (roughly) that a structured proposition is obtained by replacing the words of a sentence with their contents while retaining the syntactic relations in the logical form of the sentence. This solves Benacerraf's problem because the structure in propositions is identified with the syntactic structure in the logical form. The connection with syntax provides a principled reason for identifying propositions with the structured objects proposed by King.

Another approach to structured propositions is due to Jon Barwise and John Etchemendy, who use what they call "Austinian propositions," named after the Oxford philosopher J. L. Austin, in their solution to the liar paradox. An Austinian proposition is like a structured Russellian proposition except that it contains a contextually determined situation that the proposition is about. So while the Russellian proposition that Claire is playing cards is true just in case Claire is playing cards, the Austinian proposition that Claire is playing cards is true just in case Claire is playing cards in the contextually determined situation. For every situation *s*, there is a liar proposition *f* about *s* that claims that *f* is false in *s*. In Barwise and Etchemendy's formal development, it turns out that every such *f* is simply false. However, for every *s*, there is an expanded situation *s′*, and there is a true proposition *p* about *s′* that claims that *f* is false in *s′*. The intuition that the liar proposition *f* is both true and false arises out of a failure to keep separate the distinct Austinian propositions *f* and *p*.

A general challenge to propositions has come from Donald Davidson, who has used the so-called slingshot argument to collapse all facts into a single Great Fact, effectively robbing facts of their philosophical utility. Davidson argues that if we give up on facts, we should also give up on entities that represent facts, such as propositions. The slingshot argument can also be used directly against propositions to show that all true propositions collapse into a single Great Proposition. But as Stephen Neale has shown, the slingshot argument can be avoided as long as one holds that sentential operators like "the fact that … is identical to the fact that … " and "the proposition that … is identical to the proposition that …" satisfy certain logical constraints on inference rules involving definite descriptions. This constraint can easily be satisfied if one adopts a Russellian analysis of definite descriptions (which construes "The *f* is *g*" as "There is exactly one *f*, and it is *g*").

See also Meaning; Propositional Attitudes: Issues in Semantics.

Bibliography

Barwise, Jon, and John Etchemendy. *The Liar*. New York: Oxford University Press, 1987.

Bealer, George. "Propositions." *Mind* 107 (425) (1998): 1–32.

Benacerraf, Paul. "What Numbers Could Not Be." *Philosophical Review* 74 (1965): 47–73.

Cresswell, Max. *Structured Meanings*. Cambridge, MA: MIT Press, 1985.

Davidson, Donald. "The Structure and Content of Truth." *Journal of Philosophy* 87 (1990): 279–328.

Jubien, Michael. "Propositions and the Objects of Thought." *Philosophical Studies* 104 (2001): 47–62.

King, Jeffrey. "Structured Propositions and Complex Predicates." *Noûs* 29 (1995): 516–535.

Neale, Stephen. *Facing Facts*. New York: Oxford University Press, 2001.

Stalnaker, Robert. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Peter W. Hanks (2005)

PROTAGORAS OF ABDERA

Protagoras of Abdera in Thrace, most famous of the Sophists, was born not later than 490 BCE and probably died soon after 421 BCE. According to Plato, he was the first to declare himself a professional Sophist. He went from city to city in the Greek world, offering instruction in return for money, and he undertook above all to train young men in the art of politics. He was well known in Athens, where he enjoyed the friendship of Pericles—he produced a theoretical basis for Periclean democracy and was asked by Pericles to draft the constitution for the new colony of Thurii in 443 BCE. He made contributions to grammatical and rhetorical theory, and his views on religion provoked charges of impiety against him in the courts, which led to his exile from Athens at the end of his life and to the public burning of at least one of his books.

His writings were numerous and included "On Truth," "On the Gods," and "Antilogic" (or "Antilogies"). Later writers probably took their information about him mainly from the accounts of Plato, Aristotle, and Sextus Empiricus, but one of his works was read by Porphyry in the third century CE, and in the Hellenistic period he was regarded as sufficiently important for his statue to be set up, together with those of Plato, Aristotle, and other thinkers, in the Serapeum at Memphis in Egypt.

Since the time of Plato, Protagoras's main doctrines have been regarded as possessing considerable philosophical interest, even by those who deny philosophical importance to the Sophists in general; but very divergent interpretations have been propounded. With no surviving works and virtually no fragments, interpretation must depend upon the assessment of the evidence of Plato, Aristotle, and Sextus Empiricus. In what follows, the view is taken that Plato in the *Theaetetus* correctly states the basic position of Protagoras and then proceeds to distinguish certain possible developments of this position not held by Protagoras. The basic position was independently understood in the same way by both Aristotle and Sextus Empiricus, each of whose information was not simply derived from the *Theaetetus*. This would be denied by some scholars.

EPISTEMOLOGY

The starting point must be the famous contention that "man is the measure of all things, of things that are that [or 'how'] they are and of things that are not that [or 'how'] they are not." Theodor Gomperz maintained that "man" is to be understood collectively in the sense of "mankind as a whole" or "the human race." But against this, the evidence of the *Theaetetus* 152A–B seems to show conclusively that it is individual men that Protagoras had in mind in the first instance, although, as will be seen, his theory is capable of easy extension to groups of men, and he probably made this extension himself.

According to Plato's example in the *Theaetetus*, when the same wind appears cold to one person and warm to another person, then the wind is warm to the person to whom it appears warm and is cold to the person to whom it seems cold. It follows that all perceptions are true and the ordinary view is mistaken, according to which, in cases of conflict, one person is right and the other person is wrong about the quality of the wind or of anything else. This clearly was the position held by Protagoras, but it is not clear exactly how he came to this view. It is often held that his position is a kind of subjective idealism similar to that of Bishop Berkeley, according to which qualities in a thing are for the person to whom they seem, so long as they seem to him, but have no existence independent of their seeming.

Against this view, Sextus Empiricus is explicit: All qualities perceived by different persons are actually present in matter. Sextus's introduction of matter may well be anachronistic, but his account suggests an alternative view, accepted by F. M. Cornford among others, according to which opposite qualities are copresent in objects, and in cases of conflict of perceptions between two persons, what happens is that we have a sort of selective perception—one person perceives one quality and the other its opposite, both qualities being present in the situation, waiting to be perceived, as it were, independently of any actual perceiving by a subject. This view seems to have the support of Aristotle, who always treats Protagoras's doctrine as involving the denial of the principle of contradiction, and the view coincides with incidental pointers in Plato's account ("the same wind"—152B; "perception, then, is always of something that is"—152C). It is true that in the "secret doctrine" attributed to Protagoras by Plato (152Cff.) the independent status of sense objects is undermined, but the fact that this is presented as a secret

doctrine is surely conclusive evidence that it was a doctrine not publicly associated with Protagoras.

The "man-measure" doctrine is presented by Plato in the first instance as a doctrine about perception of sensible qualities. But it is clear that Plato supposed that for Protagoras it also applied to moral and aesthetic qualities such as "just" and "beautiful." It is especially in these cases that the extension of the doctrine to groups of people was made by Protagoras—"whatever seems just to a city is just for that city so long as it seems so." Probably Protagoras did not extend his doctrine to apply to all judgments; this was done immediately by his opponents in the famous *peritrope*, or "turning of the tables": Let us suppose that whatever seems true to any person is true for the person to whom it seems so. If this is the doctrine of Protagoras, then Protagoras will hold that those who hold that Protagoras's theory is false are holding the truth (*Theaetetus* 171A). But Plato points out that if Protagoras could pop his head up through the ground, he would surely have an answer to this objection.

At the very least, Protagoras was clear about one point. In the case of conflict about perceived qualities all perceptions are true. But some perceptions are better than others, for example, the perceptions normally found in a healthy man as distinct from those found in a man who is ill. It is the function of a doctor, Protagoras held, to change a man who is ill so that his perceptions become those of a man who is well. Likewise, in moral, political, and aesthetic conflicts it is the function of the Sophist as a teacher to work a change so that better views about what is "just" and "beautiful" will seem true to the "patient"—better, that is, than those that previously seemed true to him. All the "patient's" views are equally true, but some are better than others.

There is nothing to suggest that by "better" Protagoras meant what will seem better. Quite the contrary. Better views are views that have better consequences, and consequences which are better are so as a matter of fact, independently of whether a person thinks them better or not. In other words, Protagoras here made an exception to his man-measure doctrine. There is every reason to suppose that he would have excepted the class of judgments about the consequences of judgments from his principle. Indeed, there is no actual evidence in any ancient author that Protagoras himself ever applied his doctrine to statements other than those about perceived qualities and moral and aesthetic qualities treated on the same plane as visually perceived qualities. What probably happened was that he propounded his doctrine in certain general statements such as "whatever seems to anyone is so for that person," without adding the qualifications that he really intended; thus he gave a handle to his enemies, which enabled them to apply the *peritrope* and similar objections.

The above account rests primarily upon Plato's *Theaetetus*. To it may be added evidence from other sources. According to Diogenes Laërtius, Protagoras was the first to propound the theory that there are two *logoi*, or accounts, to be given about everything. This has sometimes been treated as simply the now familiar rhetorical doctrine that "there are two sides to every question." But this theory was used as a method of argument, and it should probably be related to the man-measure doctrine and to what Plato called "Antilogic," the probable title of one of Protagoras's treatises. In conflicts about perceived qualities, and also moral and aesthetic qualities, there might seem room for an infinite variety of "seemings," but if we take any one as a starting point, for instance, that the wind seems warm, all other seemings may be expressed as the negative of this, namely "not-warm." This was clearly the way in which Plato tended to regard phenomena—as did the antilogicians, too—namely, as always being both "warm" and "not-warm." In this view, Plato was probably following Protagoras. It is possible that Protagoras associated with the two-*logoi* principle the prescription attributed to him by Aristotle "to make the lesser [or 'the weaker'] argument the stronger." This may have been what the Sophist was expected to do when altering a man's opinions for the better.

SOCIAL THEORY

In Plato's dialogue *Protagoras* we are given a coordinated theory of the Sophist in relation to society and of a possible theoretical basis for a Periclean-style democracy. All is completely consistent with the positions attributed to Protagoras in the *Theaetetus*. When Protagoras professes to make men good citizens, Socrates objects that while the Athenians call in experts to advise on technical matters, they regard all citizens as capable of advising them on matters relating to the city. This seems to imply that Athenian democracy leaves no place for expert instruction in citizenship. Protagoras replies with a myth followed by a nonmythical exposition that while all men share in the qualities that make good citizens, they do not do so by nature but acquire these qualities by instruction and by practice. These qualities are beliefs and opinions about what is just and right. In a sense, the whole community teaches its members about these matters, and so all are rightly consulted about political matters. But the expert teacher, such as the Sophist Protagoras, can

improve opinions on such matters, whether it be in the case of an individual or in the case of a whole community.

OTHER VIEWS

Protagoras's doctrines ranged beyond the topics discussed above to cover physical and mathematical problems as well, but it is no longer possible to state his actual teachings on these problems. He seems to have held that a tangent touches a circle not only at one point, but at more than one, clearly arguing from visual experience of drawn lines. Parmenides had rejected the world of seeming in favor of his world of being; Protagoras took the opposite path and attempted to expound a world in which all appearances were true and where there was nothing outside or beyond what appeared. This involved the copresence of opposed and contradictory qualities at many points. Protagoras was prepared to accept and explain this copresence through his "man-measure" principle, either on the basis of a theory of subjective idealism or, more probably, on the basis of a conception of a phenomenal world actually composed of opposites (a conception typical of the pre-Socratics). This conception seemed to Plato to be substantially correct for the phenomenal world, hence his great interest in Protagoras. But Plato felt that this view made it impossible to give any account or explanation of phenomena, and to be able to give an explanation seemed to him essential.

Diogenes Laërtius says that for Protagoras the soul is nothing apart from its perceptions. This suggests a phenomenalistic view of the soul as well as of everything else. Diogenes' account may be correct, although doubts have been cast upon it. If it is correct, however, it probably was not intended to imply any doctrine like the modern theory of neutral monism, but simply to deny the existence of any "submerged," or nonphenomenal, element in the soul.

See also Ethical Relativism; Sophists.

Bibliography

TEXTS AND TRANSLATIONS

Gagarin, Michael, and Paul Woodruff. *Early Greek Political Thought from Homer to the Sophists*. Cambridge, U.K.: Cambridge University Press, 1995.

Sprague, R.K., ed. *The Older Sophists; a Complete Translation by Several Hands of the Fragments in die Fragmente der Vorsokratiker, ed. by Diels-Kranz*. Columbia: University of South Carolina Press, 1972.

Fragments and testimonia of Protagoras' work may be found in H. Diels and W. Kranz, eds., *Fragmente der Vorsokratiker*, 10th ed. (Berlin, 1961), Vol. II. A. Capizzi, *Protagora* (Florence, 1955), is probably the best full discussion, and it includes some material not in Diels and Kranz.

STUDIES

Bett, R. "The Sophists and Relativism." *Phronesis* 34 (1989): 139–169.

Classen, J. "Protagoras' *Aletheia*." In *The Criterion of Truth, Essays Written in Honour of George Kerferd*, edited by P. Huby and G. Neal. Liverpool: Liverpool University Press, 1989.

Guthrie, W. K. C. *The Sophists*. Cambridge, U.K.: Cambridge University Press, 1971.

Kerferd, G. B. "Plato's Account of the Relativism of Protagoras." *Durham University Journal* 42 (1949): 20–26.

Kerferd, G. B. "Protagoras' Doctrine of Justice and Virtue in the 'Protagoras' of Plato." *Journal of Hellenic Studies* 73 (1953): 42–45.

Kerferd, G. B. *The Sophistic Movement*. Cambridge, U.K.: Cambridge University Press, 1981.

Schiappa, E. *Protagoras and Logos: A Study in Greek Philosophy and Rhetoric*. Columbia, SC: University of South Carolina Press, 1991.

Vlastos, G. "Introduction." In *Plato's Protagoras*, translated by M. Ostwald. Indianapolis/New York: Bobbs-Merrill, 1956.

G. B. Kerferd (1967)

Bibliography updated by Paul Woodruff (2005)

PROTOCOL SENTENCES

See *Basic Statements*

PROUDHON, PIERRE-JOSEPH
(1809–1865)

Pierre-Joseph Proudhon has been called the father of anarchism, a title that is accurate insofar as organized anarchist movements throughout the world can be traced to his teachings and to the actions of his disciples. Proudhon was also the first writer deliberately to accept the title of anarchist, which he did in 1840. Before his time the term had been used to denote one who seeks to promote social disorder; Proudhon argued that it could be used with more justice to describe one who seeks social order without authoritarian government. "As man seeks justice in equality, so society seeks order in anarchy," he said. "Anarchy—the absence of a master, of a sovereign—such is the form of government to which we are every day approximating." Such doctrines were not entirely original; the English writer William Godwin had expounded them fifty years earlier without describing them as "anar-

chist," but Proudhon appears to have been uninfluenced by Godwin and to have reached his conclusions independently.

Proudhon prided himself on being a man of the people. He was born in Besançon, capital of Franche-Comté, of Jura peasant stock. His childhood was hard, and after a brief period at the college in Besançon, he received his education largely through his work as a printer; he taught himself Greek and Hebrew and developed a prose style that eventually won the admiration of Charles-Pierre Baudelaire, Gustave Flaubert, and Victor Hugo. The turning point in Proudhon's career came when he was awarded a scholarship by the Besançon Academy in 1838. This took him to Paris and gave him the leisure to formulate his ideas and to write his first important book, *Qu'est-ce que la propriété?* (*What Is Property?,* Paris, 1840). This book, hailed by Karl Marx as "the first decisive, vigorous and scientific examination" of the institution of property, gained notoriety because in one passage Proudhon defined property as "theft." The author's love of telling phrases distorted the nature of his argument, for *Qu'est-ce que la propriété?* was in fact an investigation of abuses that had entered into the institution of property rather than a condemnation of property itself. The arguments that Proudhon put forward in this early book, on the nature of property and the faults of government, are those which he elaborated and gave a deeper philosophical backing in his later works.

Proudhon attacked the existence of private property that allows the exploitation of the labor of others, such as the owning of land by those who do not work it; he had only approval for the "possession" that allows a worker to dispose of what his hands make. "The right to products is exclusive—*jus in re*; the right to means is common—*jus ad rem.*" This is so because the means of production, the heritage of techniques and inventions, have been built up by human cooperation, and no man has a right to use them exclusively for his own benefit. However, for the sake of independence, Proudhon granted the need for each man to control the land or tools he can use. In this early book he still thought in terms of a peasant-and-handcraft society.

Proudhon attacked unreformed property because it negates equality, but he rejected the communist theories of his time (principally those of the French utopian socialists) because they denied independence. Here Proudhon came to the political aspect of his argument—both unreformed property and communism are dependent on forms of authority to maintain themselves. But how far is authority justified? Proudhon contended that it arises from the tendency of social animals and primitive man to seek leaders. As reason develops, criticism, protest, and rebellion arise. Emergent political science finds the laws by which society functions in the nature of things, not in the whims of rulers. At this point anarchy, administration without government, becomes possible. Proudhon, at this stage under the influence of Hegelian ideas imperfectly absorbed from French reviews, created a triad. The thesis is property, which destroys equality; the antithesis is communism, which denies independence; the synthesis is anarchy or liberty, which is embodied in a society of producers bound together by a network of free contracts. In the widening recognition of mutual interests, government becomes unnecessary.

During the 1840s Proudhon served for several years as office manager for a water transport firm in Lyons, work that allowed him to travel frequently to Paris. In these two settings his theory of mutualism—the form of anarchism particularly associated with him—developed. Political radicalism flourished in mid-nineteenth-century Lyons, and Proudhon encountered there the disciples of Étienne Cabet, Charles Fourier, Pierre Leroux, and other socialist prophets. He developed the idea of a worldwide working-class organization on an economic basis rather than a political one. This led him to place faith in various forms of mutual credit systems that might eventually make governmental administration unnecessary; he envisaged such associations as becoming worldwide. In Paris, Proudhon associated with some of the leading European revolutionary theorists, including Marx, Mikhail Bakunin, and Alexander Ivanovich Herzen. However, his personal and theoretical incompatability with Marx soon became evident; the historic conflict between libertarian and authoritarian views of socialism began with the split between Marx and Proudhon, which dates from Marx's attack in *La misère de la philosophie* (Paris, 1847) on Proudhon's *Système des contradictions économiques* (2 vols., Paris, 1846). Bakunin and Herzen, on the other hand, eventually became Proudhon's most important disciples.

During the 1840s Proudhon, an eclectic thinker, took what he found valid from the writings of G. W. F. Hegel, Ludwig Feuerbach, Immanuel Kant, and other German philosophers, as well as from Auguste Comte and the French utopians. He evolved a philosophy that left out the third term of the Hegelian triad, and accepted contradiction as an enduring force tending toward a dynamic equilibrium—the desirable condition of existence. He denied all absolutes, all utopian aspirations to permanent solutions, and, in his *Philosophie du progrès* (Paris, 1853) saw

progress as "the affirmation of universal movement and in consequence the negation of all immutable forms and formulae, of all doctrines of eternity, permanence, or impeccability, and of every subject, or object, spiritual or transcendental, that does not change." He was, deliberately and avowedly, an antisystematic philosopher.

Proudhon assumed the standpoint of a critical independent, and as such he became the most outspoken journalist of the period, giving qualified support to the French revolution of 1848. His *Le représentant du peuple* (1848) was the first anarchist newspaper published with any regularity; harried by suppressions and fines, it survived under various names for more than two years. Proudhon was elected in June 1848 to the Constituent Assembly, where he maintained an intransigent minority position. He also planned a people's bank, based on his mutualist ideas, which never materialized because he was imprisoned for attacks in his paper on Louis Napoleon, then president of the Republic.

Proudhon's three years of imprisonment were light: He was allowed occasional days out on parole, on one of which he married Euphrasie Piégard, and he wrote two of his most important books, *Les confessions d'un révolutionnaire* (Paris, 1850), an analysis of the events of 1848 that states the aim of anarchist revolutionism as "no more government of man by man, by means of the accumulation of capital," and *Idée générale de la révolution au XIX[e] siécle* (*General Idea of the Revolution in the Nineteenth Century,* Paris, 1851). The latter book comes nearer than anything else Proudhon wrote to presenting his view of the ideal libertarian society, based on contract instead of laws, with authority decentralized in communes and industrial associations, with frontiers abolished and flexible federation replacing the centralized national state.

During the early years of the Second Empire, Proudhon was subjected to constant police persecution, and in 1858 he was again sentenced to three years' imprisonment for an offense against the press laws. He fled to Belgium, where, although pardoned in 1860, he lived until 1862. During his final years in Paris, he gained a considerable mutualist following among French workingmen, and before he died early in 1865, he learned that his followers had taken a leading part in the meetings that led to the founding of the International Workingmen's Association.

During his final years Proudhon wrote a number of books that elaborated important aspects of his doctrines. *Du Principe fédératif* (Paris, 1863) summarized his criticism of nationalism and developed his ideas of communal organization leading gradually to world federation. *De la Justice dans la révolution et dans l'église* (3 vols., Paris, 1858) opposed his own theory of an immanent justice to transcendentalist ideas of justice. *De la Capacité politique des classes ouvrières* (published posthumously in Paris, 1865) developed Proudhon's view of the power of the working class to achieve its own liberation by economic means.

Later anarchism and syndicalism were largely influenced by Proudhon's doctrines, as was the populist movement in Russia. As the Russian anarchist Mikhail Bakunin said, "Proudhon was the master of us all."

See also Bakunin, Mikhail Aleksandrovich; Comte, Auguste; Feuerbach, Ludwig Andreas; Fourier, François Marie Charles; Hegel, Georg Wilhelm Friedrich; Herzen, Aleksandr Ivanovich; Kant, Immanuel; Marx, Karl.

Bibliography

WORKS BY PROUDHON

Oeuvres complètes de P. J. Proudhon, 26 vols. Paris, 1867–1870.

Correspondance, 14 vols. Paris, 1874–1875.

What Is Property? Translated by Benjamin Tucker. Princeton, MA: B.R. Tucker, 1876.

System of Economical Contradictions. Translated by Benjamin Tucker. Boston, 1888.

Oeuvres complètes de P. J. Proudhon. Edited by C. C. A. Bouglé and Henri Moysset, 11 vols. Paris, 1920–1939. Never completed.

General Idea of the Revolution in the Nineteenth Century. Translated by John Beverley Robinson. London: Freedom Press, 1923.

WORKS ON PROUDHON

Brogan, D. W. *Proudhon.* London, 1934.

Dolléans, Édouard. *Proudhon.* Paris: Gallimard, 1948.

Hoffman, R. L. *Revolutionary Justice: The Social and Political Thought of P.-J. Proudhon.* Urbana: University of Illinois Press, 1972.

Hyams, Robert S. *Pierre-Joseph Proudhon: His Revolutionary Life, Mind, and Works.* London: Murray, 1979.

Lubac, Henri de. *Proudhon et le christianisme.* Paris, 1945. Translated by R. E. Scantlebury as *The Un-Marxian Socialist: A Study of Proudhon.* London: Sheed and Ward, 1948.

Noland, Aaron. "Proudhon and Rousseau." *Journal of the History of Ideas* 28 (1967): 33–54.

Prion, Gaëtan. *Proudhon et syndicalisme révolutionnaire.* Paris, 1910.

Ritter, Alan Irving. *The Political Thought of Pierre-Joseph Proudhon.* Princeton, NJ: Princeton University Press, 1969.

Sainte-Beuve, Charles A. *P.-J. Proudhon.* Paris, 1872.

Vincent, K. Steven. *Pierre-Joseph Proudhon and the Rise of French Republican Socialism.* New York: Oxford University Press, 1984.

Woodcock, George. *Pierre-Joseph Proudhon.* London: Routledge and Kegan Paul, 1956.

George Woodcock (1967)
Bibliography updated by Philip Reed (2005)

PROUST, MARCEL
(1871–1922)

The French author Marcel Proust was born and educated in Paris. He lived there all his life, leaving only for short holidays or artistic pilgrimages, most of which were to the great cathedral cities of France. His father, a professor of medicine, was Catholic; his mother, whom he adored, was Jewish. Both traditions, as well as his consuming interest in French history and culture, played important roles in his life and art, although he was neither religiously orthodox nor politically chauvinistic. He undertook a considerable and seemingly futile search for a vocation and did some writing, most of which was discarded drafts of his future novel. Suffering terribly from asthma and from certain guilts about his homosexuality, but with economic as well as spiritual means sufficient to indulge and transmute these ills, Proust ensconced himself in his famous cork-lined room to write his masterpiece, *A la recherche du temps perdu.*

PHILOSOPHICAL THEMES

Although Proust compared a work of art in which there are theories to an object on which the price is marked, *A la recherche* is, nonetheless, a philosophical novel. There are two major philosophical themes woven into the novel: that reality is composed of artistic essences and that the search for essences ends in their dissolution. Proust stated only the first theme; the second, however, is implied by much of the action of the novel.

In the last volume of the novel, *Le temps retrouvé,* Proust, as narrator and participant, stated his theory of artistic essences as reality; this theory, because of its role in the context of the whole novel, must be understood as an integral part of it, along with the characterization, dialogue, and plot. According to Proust's theory, we live in a world of people, places, and things, all of which are organized spatially or temporally, in the ordinary sense of space and time, and which impinge on us. Most of us merely react to these phenomena. The true artist, however, like the scientists, attempts to find the laws that govern these phenomena. Whereas the scientist proceeds by his intellect, the artist cannot, for his laws are to be discovered only by intuition. The artist's intellect supplements, but it cannot supplant, intuition. Intuition is that state of mind in which the artist—rooted in past experiences, nourished by suffering, and graced by an involuntary memory of a past sensation joined with a similar present one—extracts the qualitative similarity or essence from these sensations in order to embody that essence in a metaphor which, like the essence, is not subject to the ravages of time. Thus, these essences are the only true reality, and their artistic expression the only true judgment on reality.

Proust, it is important to realize, did not deny the existence of temporal or spatial relations, but he rejected them as unreal. Hence, he must understand by *reality* something quite distinct from *existence*: *reality* for him functioned as an honorific term denoting that which is salvageable from the past and which transcends the present—that, therefore, which is ultimate in the precise sense of being out of time. *Reality,* in effect, denotes the essences extracted by intuition from what exists in relation to what existed.

It has been claimed that Proust's conceptions of time and intuition are Bergsonian. It seems, however, that there are important differences. According to Henri Bergson, time is essentially duration (*durée*). The concepts of the past, present, and future cannot apply to time because they spatialize it. Duration can only be experienced, not thought of or talked about; it is the indivisible, ultimate fact of process in the world, and intuition is the experience of duration, a direct acquaintance with it. For Proust, however, time is not duration; it consists of chronological relations among events. Nor is time ultimate; only the timeless essences are that. Finally, intuition for Proust is an extraction from, not an immersion in, time.

Nor is Proust's theory Platonic, as has sometimes been suggested. Plato's timeless essences are perfect and have their being absolutely independently of the spatial and temporal particulars of this world; the Proustian essences are at most more or less imperfect copies of the truly real forms.

Besides this aesthetic-ontological theme, which Proust integrated magnificently in the novel, there is the nether theme of the dissolution of essences in the very search for them. Although he never stated this theme, much of the novel embodies it. The treatment of love is probably the best single example. Through the narration of many different love relationships, commonly regarded as a major achievement of the novel, Proust dramatized that love has no essence, only an inexhaustible set of properties, none of which is necessary or sufficient. Here

intellect supplants rather than supplements intuition. Proust's observations, analyses, and generalizations harvest a vast multiplicity of criteria that govern our understanding and concept of love. In effect, Proust showed through his characterization, monologue and dialogue, as well as through the plot, that the range of the experience of love renders impossible any traditional essentialist definition of it. To have discovered, explored, and artistically wrought this important truth about our conceptual life and to have shown it a full generation before philosophers stated it is not the least of Proust's accomplishments in his great novel.

See also Appearance and Reality; Bergson, Henri; Intuition; Plato; Platonism and the Platonic Tradition.

Bibliography

WORKS BY PROUST

Portraits de peintres. Paris: Heugel, 1896.

Les plaisirs et les jours. Paris, 1896. Translated by Louise Varese and others as *Pleasures and Days,* with an introduction by F. W. Dupee. Garden City, NY: Doubleday, 1957.

La bible d'Amiens. Paris, 1904. A translation, with preface and notes, of John Ruskin's *Bible of Amiens.*

Sésame et lys. Paris, 1906. A translation, with preface ("Journées de lecture") and notes, of Ruskin's *Sesame and Lilies.*

A la recherche du temps perdu (published by the *Nouvelle Revue Française*). 7 parts, 13 vols. Paris, 1919–1927. One volume had been published in 1913, at Proust's expense. Translated by C. K. Scott Moncrieff (the first 6 parts) and Stephen Hudson (last part only) as *Remembrance of Things Past.* New York: Random House, 1922–1931. Uniform edition, 12 vols. London and New York, 1941. Pléiade edition, 3 vols., Paris, 1954.

Chroniques. Paris: Éditions de la Nouvelle revue francaise, 1927.

Jean Santeuil. With introduction by André Maurois. Paris, 1954. Translated by Gerard Hopkins. London: Weidenfeld and Nicolson, 1955; New York: Simon and Schuster, 1956.

Contre Sainte-Beuve. Paris: Gallimard, 1954. Translated by Sylvia Townsend Warner as *By Way of Sainte-Beuve.* London: Chatto and Windus, 1958.

WORKS ON PROUST

Brée, Germaine. *Marcel Proust and Deliverance from Time.* New Brunswick, NJ: Rutgers University Press, 1955.

Cocking, J. M. *Proust.* London: Bowes and Bowes, 1956.

Delattre, Floris. *Bergson et Proust.* Paris, 1948. Vol. I of *Les études Bergsoniennes.*

Green, F. C. *The Mind of Proust.* Cambridge, U.K.: Cambridge University Press, 1949.

Maurois, André. *The Quest for Proust.* London, 1950.

Weitz, Morris. *Philosophy in Literature: Shakespeare, Voltaire, Tolstoy and Proust.* Detroit: Wayne State University Press, 1963.

Morris Weitz (1967)

PROVABILITY LOGIC

Even though "provability logic" did not come into its own until the early seventies, it has its roots in two older fields: metamathematics and modal logic. In metamathematics, we study what theories can say about themselves. The first—and most outstanding—results are Kurt Gödel's two incompleteness theorems.

If we take a sufficiently strong formal theory *T*—say, Peano arithmetic—we can use Gödel numbering to construct in a natural way a predicate *Prov(x)* in the language of *T* that expresses "*x* is the Gödel number of a sentence which is provable in *T*." About *T* we already know that it satisfies *modus ponens*:

If it is provable that *A* implies *B,* then, if *A* is provable, *B* is provable as well.

Now it turns out that, using Gödel numbering and the predicate *Prov,* we can express *modus ponens* in the language of *T,* and show that in *T* we can actually prove this formalized version of *modus ponens*:

$$Prov(\ulcorner A \to B \urcorner) \to (Prov(\ulcorner A \urcorner) \to Prov(\ulcorner B \urcorner)).$$

When we rephrase both the normal and the formalized version of *modus ponens* using the modal operator $\Box$, reading $\Box A$ as "*A* is provable in *T,*" we get the modal rule

(1) $$\frac{A \to B \qquad A}{B}$$

and the modal axiom

(2) $$\Box(A \to B) \to (\Box A \to \Box B).$$

Indeed both the rule and the axiom are well known from the basic modal logic *K.*

Similarly, we can show that if there is a proof of the sentence *A* in *T,* then *T* itself can check this proof, so *T* proves $Prov(\ulcorner A \urcorner)$—we shall call this principle *Prov*-completeness. Again, though in a less straightforward way than in the case of *modus ponens,* we can formalize the principle itself and see that *T* actually proves:

$$Prov(\ulcorner A \urcorner) \to Prov(\ulcorner Prov(\ulcorner A \urcorner) \urcorner).$$

When we rephrase the principle of *Prov*-completeness and its formalization in modal logical terms, we get the modal rule that is usually called necessitation:

(3) $$\frac{A}{\Box A},$$

and the modal axiom

(4) $$\Box A \to \Box\Box A,$$

which is the transitivity axiom 4 well known from modal systems such as K4 and S4.

Finally, one might wonder whether T proves the intuitively valid principle that "all provable sentences are true," that is, whether T proves $Prov(\ulcorner A \urcorner) \to A$. Unexpectedly, this turns out not to be the case at all. Löb proved in 1953, using Gödel's technique of diagonalization, that T proves $Prov(\ulcorner A \urcorner) \to A$ only in the trivial case that T already proves A itself!

Löb's theorem has a formalization that can also be proved in T. Writing both the theorem and its formalization in modal terms, we get the modal rule

(5) $$\frac{\Box A \to A}{A},$$

and the modal axiom

(6) $$\Box(\Box A \to A) \to \Box A,$$

usually called W (for well-founded) by modal logicians.

Now we can define provability logic, which goes by various names in the literature—*PRL*, *GL* (for Gödel/Löb), *L* (for Löb), and, in modal logic texts, *KW*4. It is generated by all the modal formulas that have the form of a tautology of propositional logic, plus the rules (1),(3),(5) and axioms (2),(4),(6) given above. One can prove that rule (5) and axiom (4) already follow from the rest, so that *PRL* is equivalently given by the well-known system K plus the axiom $\Box(\Box A \to A) \to \Box A$.

The main "modal" theorem about *PRL*—but one with great arithmetical significance—is the "fixed point theorem," which D. de Jongh and G. Sambin independently proved in 1975. The theorem says essentially that "self-reference is not really necessary." Suppose that all occurrences of the propositional variable p in a given formula A are under the scope of $\Box$-es, for example, $A(p) = \neg \Box p$ or $A(p) = \Box(p \to q)$. Then there is a formula B in which p does not appear, such that all propositional variables that occur in B already appear in $A(p)$, and such that $PRL \vdash B \leftrightarrow A(B)$. This B is called a fixed point of $A(p)$. Moreover, the fixed point is unique, or more accurately, if there is another formula B' such that $PRL \vdash B' \leftrightarrow A(B')$, then we must have $PRL \vdash B \leftrightarrow B'$. Most proofs of the fixed point theorem in the literature give an algorithm by which one can compute the fixed point.

For example, suppose that $A(p) = \neg\Box p$. Then the fixed point produced by the algorithm is $\neg\Box\bot$, and indeed we have $PRL \vdash -\Box\bot \leftrightarrow -\Box(-\Box\bot)$. If we read this arithmetically, the direction from left to right is just the formalized version of Gödel's second incompleteness theorem. Thus, if T does not prove a contradiction, then it is *not* provable in T that T does not prove a contradiction.

The landmark result in provability logic is Solovay's "arithmetical completeness theorem" of 1976. This theorem says essentially that the modal logic *PRL* captures *everything* that Peano arithmetic can say in modal terms about its own provability predicate. Before formulating Solovay's theorem more precisely, we turn to the semantics of *PRL*.

Provability logic has a suitable Kripke semantics, just like many other modal logics. Unaware of the arithmetical relevance of *PRL*, Krister Segerberg proved in 1971 that it is sound and complete with respect to finite irreflexive transitive frames, and even with respect to finite trees. This completeness theorem immediately gives a decision procedure to decide for any modal formula A whether A follows from *PRL* or not. Looking at the procedure a bit more precisely, it can be shown that *PRL* is "very decidable": Like the well-known modal logics *K, T,* and *S*4, it is decidable in PSPACE. This means that there is a Turing machine that, given a formula A as input, answers whether A follows from *PRL*; the size of the memory that the Turing machine needs for its computations is only polynomial in the length of A.

The modal completeness theorem was an important first step in Solovay's proof of the arithmetical completeness of *PRL*. Suppose that *PRL* does not prove the modal formula A. Then there is a finite tree such that A is false at the root of that tree. Now Solovay devised an ingenious way to describe the tree in the language of Peano arithmetic. Thus he found a translation f from modal formulas to sentences of arithmetic, such that Peano arithmetic does not prove $f(A)$. Such a *translation* f respects the logical connectives (so, e.g., $f(B \wedge C) = f(B) \wedge f(C)$), and $\Box$ is translated as *Prov* (so $f(\Box B) = Prov(\ulcorner f(B) \urcorner)$). Thus Solovay's arithmetical completeness theorem gives an alternative way to construct many nonprovable sentences. For example, we know that *PRL* does not prove $\Box p \vee \Box\neg p$, so by the theorem, there is an arithmetical sentence $f(p)$ such that Peano arithmetic does not prove $Prov(\ulcorner f(p) \urcorner) \vee$

$Prov(\ulcorner \neg f(p) \urcorner)$. In particular, if we suppose that Peano arithmetic does not prove any false sentences, this implies that neither $f(p)$ nor $\neg f(p)$ is provable in Peano arithmetic.

In recent years, logicians have investigated many other systems of arithmetic that are weaker than Peano arithmetic. They have given a partial answer to the question: "For which theories of arithmetic does Solovay's arithmetical completeness theorem still hold¿" It certainly holds for theories T that satisfy the following two conditions:

1. T proves induction for formulas in which all quantifiers are bounded (like the quantifier $\forall x \leq y + z$) and T proves that for all x, its power 2^x exists. In more technical terms: T extends $I\Delta_0 + EXP$.
2. T does not prove any false Σ_1 sentences.

For such theories, it is also clear that *PRL* is sound if we read $\Box$ as $Prov_T$ (where $Prov_T$ is a natural provability predicate with respect to a sufficiently simple axiomatization of T). To sum up, we have the following theorem: If T satisfies 1 and 2, and A is a modal sentence, then

$$PRL \vdash A \Leftrightarrow \text{for all translations } f, T \vdash f(A).$$

This result shows a strength of provability logic: For many different theories, *PRL* captures exactly what those theories say about their own provability predicates. At the same time this is of course a weakness: For example, provability logic does not point to any differences between those theories that are finitely axiomatizable and those that are not.

In order to be able to speak in a modal language about such distinctions between theories, researchers have extended provability logic in many different ways, only a few of which are mentioned here. One way is to add a binary modality, $\rhd$, where for a given theory T, the modal sentence $A \rhd B$ stands for "$T + B$ is interpretable in $T + A$." It appears that the interpretability logic of $I\Delta_0 + superexp$ is different from the interpretability logic of Peano arithmetic.

Another way to extend the framework of *PRL* is to add propositional quantifiers, so that one can express principles like Goldfarb's:

$$\forall p \forall q \exists r \Box((\Box p \vee \Box q) \leftrightarrow \Box r).$$

Finally, one can of course study predicate provability logic. V. A. Vardanyan proved that the set of always provable sentences of predicate provability logic is not even recursively enumerable, so it has no reasonable axiomatization.

See also Gödel, Kurt; Gödel's Incompleteness Theorems; Kripke, Saul; Logic, History of; Modal Logic; Peano, Giuseppe.

Bibliography

Artemov, S. N., and L. D. Beklemishev. "On Propositional Quantifiers in Provability Logic." *Notre Dame Journal of Formal Logic* 34 (1993): 401–419.

Boolos, G. *The Logic of Provability.* New York: Cambridge University Press, 1993.

de Jongh, D. H. J., M. Jumelet, and F. Montagna. "On the Proof of Solovay's Theorem." *Studia Logica* 50 (1991): 51–70.

Smoryński, C. *Self-Reference and Modal Logic.* New York: Springer-Verlag, 1985.

Solovay, R. M. "Provability Interpretations of Modal Logic." *Israel Journal of Mathematics* 25 (1976): 287–304.

Visser, A. "Interpretability Logic." In *Mathematical Logic (Proceedings, Chaika, Bulgaria, 1988),* edited by P. P. Petkov. New York: Plenum Press, 1990.

Rineke Verbrugge (1996)

PROVIDENCE

The idea of "providence" has three components—foresight, direction, and care. It is normally found in a theistic context. In its fullest sense it means that God foresees and governs (in a word, "provides for") the world that is the object of his care (or love). Divine providence was affirmed by Plato in his *Laws* (887–888), where he condemns the view, later held by the Epicureans, that the gods take no interest in human affairs. The most important later thought upon the subject arose in Stoicism and Christianity.

STOICS

The Stoics held a firm belief in the providence (*pronoia*) of God (or the gods). Thus, Epictetus uses an elementary form of the teleological argument to prove God's supervision of the universe (*Discourse* 1.16). But two factors prevented the Stoics from taking a fully personal view of providence. First, they often conceived God abstractly (as a cosmic logos) and even physically (when they identified him with nature's basic elements, air and fire). Second, and correlatively, they did not stress God's care for persons individually, nor, as a consequence, did they allow that God accomplishes his purpose in and through the free response of human wills to his initiative. On the contrary, they equated providence with destiny or fate

(*heimarmene*). In the words of Cleanthes's Hymn to Zeus, translated by Seneca, *Ducunt volentem fata, nolentem trahunt* ("Fate leads the willing, drags the unwilling on," Epistles 107:11).

CHRISTIANITY

Our primary evidence for Christianity is the teaching of Christ himself. Christ taught that God is a Father who cares for all his children individually. Therefore, they must not be anxious or distressed; rather, they must trust God absolutely (Matthew 6:25–33, 10:29–31). Furthermore, they must approach God freely in prayer in the confidence that he will answer their requests (Matthew 7:7–11). St. Paul made two basic assertions: first, that we know through Christ that God's sovereignty is one of love through which we are "more than conquerors" (Romans 8:35–39) and second, that God accomplishes his purpose by cooperating with our wills, not by demanding our submission to a fait accompli (Romans 8:14–16, Philippians 2:12–13). Hence, St. Paul, like Jesus, affirms the reality of, and the necessity for, petitionary prayer.

Attempts have been made to see providence in nature, history, and individual lives.

NATURE

The theist maintains that God acts in nature both ordinarily, through those laws which science formulates, and extraordinarily, through miracles. Both modes of God's activity signify his wisdom and love to the believing mind. Furthermore, many theists, following Thomas Aquinas in his Fifth Way, believe that it is possible to base an argument for God's existence on the apparent traces of design in nature, but it must be admitted that the fact of evil constitutes prima-facie evidence against the existence of a Designer who is both omnipotent and good.

HISTORY

To what extent can we interpret God's purpose in terms of a "pattern," or "patterns," discernible in historical events? Here one can only summarize a general tendency among modern theologians. Most of them would say that our ability to perceive a pattern or plan is restricted to the main events of the Bible as interpreted by the prophetic and apostolic writers. Perhaps we also have a right to see a *preparatio evangelica* in the achievements of Greece and Rome, but we cannot perceive an analogous plan in either the secular or ecclesiastical history of the postbiblical era. Thus, Josef Pieper writes, "Not that he who philosophizes could reach the point of being able to identify *in concreto* the character of an event in terms of salvation and disaster. We are moving here within the realm of the *mysterious*—in the strictest sense. And even for the believer, the history of salvation 'within' history is not to be apprehended concretely" (*The End of Time,* London, 1954, p. 23).

INDIVIDUAL LIVES

In regard to individual lives we must also distinguish between a general belief in providence and a detailed knowledge of its workings. St. Paul affirmed as a matter of faith that "we know that in everything God works for good with those who love him, who are called according to his purpose" (Romans 8:28). But in 1 Corinthians 13:12 he admits that all our knowledge of God is indirect, partial, and confused. Hence, any claim to see God's purpose in particular events is bound to be provisional and incomplete.

See also Christianity; Cleanthes; Epictetus; Epicureanism and the Epicurean School; Philosophy of Religion, History of; Plato; Seneca, Lucius Annaeus; Stoicism; Teleological Argument for the Existence of God; Thomas Aquinas, St.

Bibliography

Caussade, J. P. de. *L'abandon à la providence divine.* 1867. Translated by A. Thorold as *Self-Abandonment to Divine Providence,* 5th ed. London, 1955. A classic of spiritual theology.

Flint, Thomas. *Divine Providence: The Molinist Account.* Ithaca, NY: Cornell University Press, 1998.

Geach, Peter. *Providence and Evil.* Cambridge, U.K.: Cambridge University Press, 1977.

Hazelton, Roger. *Providence.* London, 1958.

Laird, John. *Mind and Deity,* 173–201. London: Allen and Unwin, 1941.

Pollard, William C. *Chance and Providence.* New York: Scribners, 1958.

Swinburne, Richard. *Providence and the Problem of Evil.* Oxford: Clarendon Press, 1998.

H. P. Owen (1967)
Bibliography updated by Christian B. Miller (2005)

PSEUDO-DIONYSIUS

The writings of Pseudo-Dionysius, first cited at the beginning of the sixth century, have attracted interest partly because the writer has been wrongly identified with Dionysius the Areopagite, who was converted by St. Paul at Athens, and also with St. Denis, the patron saint of

France. Neither of these identifications, however, is possible.

While the thought of Pseudo-Dionysius was a continuation of the Christian Platonism of the early Church Fathers, it is directly influenced by the latest forms of Neoplatonism, as found in Proclus. No other early Christian writer was so clearly influenced by a particular philosopher. The influence of Pseudo-Dionysius on later theologians, philosophers, mystics, and poets was immense. John of Damascus and Thomas Aquinas were both strongly influenced by him. Peter Lombard, Robert Grosseteste, and Albert the Great also acknowledged their debt to him. The poetry of Dante Alighieri and John Milton reflects his heavenly hierarchy.

Four of his treatises—"The Celestial Hierarchy," "The Ecclesiastical Hierarchy," "The Divine Names," and "The Mystical Theology"—and ten of his letters are extant. The problem of the one and the many in the treatises is the problem of the relation of God to the universe, both visible and invisible. The basic propositions of Proclus were that every plurality participates in unity, is both one and not one, and is other than the one itself. The order of the universe is an order that depends on the ultimate unity. It is arranged in different orders of being that descend from and ascend to the first principle. This hierarchical view of the universe goes back to Plato and Aristotle and is found in Philo and the Gnostics, as well as in later Platonism. Proclus and Pseudo-Dionysius represent the final stage of the idea in the ancient world, and Pseudo-Dionysius is the chief transmitter of the idea to later times.

The four treatises exhibit the sequence of Dionysius's thought. Those on hierarchies show the descent and return of the divine goodness, "The Divine Names" shows the nature of God, and "The Mystical Theology" shows the way by which the knowledge of God may be found.

THE HIERARCHIES

"The hierarchy is a holy order, a knowledge and an activity which assimilates to the divine nature as far as possible and which through the light granted from God is raised in due proportion to the imitation of God" ("The Celestial Hierarchy" III, I). The celestial hierarchy contemplates the divine perfection and shares in it, reflecting its light down through its several ranks: Seraphim, Cherubim, Thrones, Dominions, Powers, Authorities, Principalities, Archangels, and Angels. The members of the highest hierarchy are nearest to God and share most fully his vision and his likeness. The other members of the hierarchy become more symbolic and corporeal as they descend. Each member of the hierarchy comes directly from God, in contrast with the emanations of Proclus, which produce one another. The Christian doctrine of creation makes the unity of the hierarchy that of spiritual communion rather than that of progressive generation. On Earth the ecclesiastical hierarchy continues the celestial hierarchy in visible form, with Jesus at the top of this hierarchy as God is at the summit of the celestial hierarchy. The members of the hierarchy in descending triads are chrism, communion, and baptism; bishops, priests, and deacons; monks, laity, and catechumens.

"THE DIVINE NAMES"

The third treatise discusses the names given to God. These names cannot describe God but must be understood in a special sense, since God is above all reason, speech, being, and name. He is above being yet the cause of being, and may be said to be only in a higher sense. His names are not derived from himself but from the manifestation of his providence. He is both nameless and many-named. He is in the world, around the world, above the world, and above the heavens. He is sun, star, fire, water, wind, dew, cloud, stone, and rock—and none of them. Knowledge of God comes through prayer, which draws men to him so that they may know his goodness. How can such a God be the sovereign creator of a world in which evil exists? Only because evil is not real but simply the absence of good. "Evil is then a deprivation, defect, weakness, disproportion, error, and the absence of purpose, beauty, life, understanding, reason and perfection." When night falls, there is nothing positive in its darkness but simply the absence of light. Evil is simply the absence of goodness.

"THE MYSTICAL THEOLOGY"

"The Mystical Theology" describes the way to the knowledge of God by the Neoplatonic method of abstracting visible and invisible qualities until one comes to the knowledge of God by negation or removal. This knowledge of God is mystical and ineffable rather than philosophical and theological, and it involves complete cessation of thought and speech. One penetrates the darkness that is above intelligible things, and in absolute silence one is united to the ineffable. God is absolutely unknowable, and the ecstasy that unites with him is both total ignorance and a knowledge beyond reason.

The distinctive quality of Pseudo-Dionysius is found in the extreme statement of two things—the unity of the world and the unity of God. The unity and order that the divine goodness imposes on the universe is described

most concretely and explicitly. The unity of God is described in negative terms that isolate it completely from all else. The extreme statement of these two opposite things enabled Pseudo-Dionysius to influence succeeding thinkers in their account of an ordered world and of a transcendent God. The opposition is inherent in all Platonism if not in all philosophy. Its explicit exposition is therefore of value.

See also Albert the Great; Aristotle; Dante Alighieri; Grosseteste, Robert; John of Damascus; Milton, John; Neoplatonism; Patristic Philosophy; Peter Lombard; Plato; Platonism and the Platonic Tradition; Proclus; Thomas Aquinas, St.

Bibliography

WORKS BY PSEUDO-DIONYSIUS

Migne, J. P., ed. *Patrologia Graeca*. Vol. III. Paris, 1857.

Suchla, Beate Regina Suchla, et al. *Corpus Dionysiacum*. 2 vols. Patristische Texte und Studien 33, 36, Berlin–New York: Walter de Gruyter, 1990–1991. Critical edition

Luibheid, Colm, tr. *The Complete Works*. Classics of Western Spirituality. Mahwah NY: Paulist Press, 1987. English translation.

WORKS ON PSEUDO-DIONYSIUS

de Andia, Ysabel, ed. *Denys l'Aréopagite et sa postérité en orient et en occident*. Collection des ...tudes Augustiniennes, Série Antiquité 151. Paris: Institut d'Études Augustiniennes, 1997.

de Andia, Ysabel. *Henosis: L'union à Dieu chez Denys l'Aréopagite*. Leiden: E.J. Brill, 1996.

Louth, Andrew. *Denys the Aorepagite*. London: Geoffrey Chapman, 1989.

Roques, René. *L'univers dionysien*. Paris: Aubier, 1954.

Rorem, Paul. *Pseudo-Dionysius. A Commentary on the Texts and an Introduction to Their Influence*. New York: Oxford University Press, 1993.

E. F. Osborn (1967)
Bibliography updated by Andrew Louth (2005)

PSEUDO-GROSSETESTE

Pseudo-Grosseteste was the anonymous author of a *Summa Philosophiae,* written between 1265 and 1275. Because of the reference in the *Summa* to Simon de Montfort's death (1265), it could not have been written by Robert Grosseteste, who died in 1253. Bartholomew of Bologna, Robert Kilwardby, and a disciple of Roger Bacon have all been suggested as the author, but there is no consensus. It does seem probable, however, that he was English and was either a Franciscan or a secular.

The *Summa,* which begins with a history of philosophy similar to that found in Bacon's *Opus Maius,* is a work of considerable subtlety and sophistication, an advanced product of the so-called Augustinian school. It holds that there is a universal wisdom in which both ancients and moderns share, perfected however by Christian revelation. Those concerned with wisdom are theosophists, to whom truth is directly revealed; theologians, who systematize and make more clear what has been revealed to the theosophists; and philosophers. The first two groups are concerned with the infallibly true, and their proper study is of matters relevant to human salvation. Philosophy, on the other hand, while it may often be in error, is completely unrestricted in its scope and may undertake to explain the natures and causes of all things whatsoever.

The *Summa* then treats the whole range of metaphysical questions in separate treatises, beginning with truth and the necessary existence of an uncreated being and ending with psychology, light, the four elements, meteors, and minerals. Its characteristic metaphysical positions are derived largely from the author's explicit hylomorphism. Every created thing is composed of matter and form. Prime matter, the mark of contingency, is not corporeal but is unextended and has three inseparable properties: It is in potency to every form; it has a desire for form; and it is privation of form. Insofar as it is privation of form it is the cause of instability; but its desire for form is a tendency toward stability. It first receives universal form, that is, substance. Substance, or substantial form, is either corporeal or incorporeal and individuates matter. It receives further perfections from other forms, so that there is a plurality of forms in any given body. This leads the author to reject the distinction (except as one of reason) between essence and existence. It also leads him to insist that the Intelligences are compounded of matter and form and differ both according to species and individuality. The human soul, like the Intelligences, is an incorporeal intelligent substance, but unlike them is capable of being joined to a body as well as of existing separately; it too is composed of matter and form. In these points, as in many others throughout the *Summa,* the author seems to be correcting what he considers the errors of Thomas Aquinas.

See also Augustinianism; Bacon, Roger; Essence and Existence; Grosseteste, Robert; Kilwardby, Robert; Revelation; Thomas Aquinas, St.

Bibliography

The text of the *Summa* is printed in L. Baur, *Die philosophischen Werke des Robert Grosseteste, Bischofs von Lincoln* (Münster in Westfalen: Aschendorff, 1912), pp. 275–643. Parts are translated in R. McKeon, *Selections from Medieval Philosophers* (New York: Scribners, 1929), Vol. I, pp. 290–314. An excellent analysis appears in C. K. McKeon, *The Summa Philosophiae of the Pseudo-Grosseteste* (New York, 1948). A good summary is in E. Gilson, *History of Christian Philosophy in the Middle Ages* (New York: Random House, 1955), pp. 265–274.

Richard C. Dales (1967)

PSYCHĒ

"Psychē" in Homer first means life and later means a departed life or ghost. The first identification with soul in the sense of the conscious self is found perhaps in Ionia, and the earliest full identification with the rational as well as with the emotional side of personality has been attributed to Socrates. In all this there was no opposition between soul and body. The doctrine that the soul is a prisoner in the body that Plato took from Orphic doctrine had reached Greece, perhaps from Scythia, before the time of Pythagoras, probably in association with a doctrine of transmigration. Plato, in the *Phaedo,* while recognizing that most people do not believe in survival after death (80D), propounded a view that combines the Socratic and Orphic attitudes. In the tripartite soul of the *Republic,* however, it is the rational part alone that is immortal; this was also Aristotle's view.

The majority of the pre-Socratics regarded the universe as a quasi-living organism, and this view also found expression in Plato's doctrine in the *Timaeus* of a world soul as a source of orderly motion in the universe. Aristotle presented a developed human and animal psychology in his analysis of the soul in the *De Anima* and elsewhere. Whereas Plato regarded the soul as a substance separate from the body, Aristotle's final view treated it as the form of a living body. For the Stoics the soul is an aspect of the all-pervading cosmic logos, while for the Epicureans it is a combination of especially smooth atoms. Within Christian theology Augustinians follow an essentially Platonist view, while Thomists prefer Aristotle's approach.

See also Aristotle; Augustinianism; Epicureanism and the Epicurean School; Homer; Orphism; Plato; Platonism and the Platonic Tradition; Pre-Socratic Philosophy; Pythagoras and Pythagoreanism; Socrates; Stoicism; Thomism.

Bibliography

Burnet, J. "The Socratic Doctrine of the Soul." *Proceedings of the British Academy* 7 (1915–1916): 235–259.

Rohde, E. *Psyche.* Translated by W. B. Hillis. London: Routledge and K. Paul, 1925.

G. B. Kerferd (1967)

PSYCHOANALYSIS

The term *psychoanalysis* pertains to the theory, therapy, and method of inquiry created by Sigmund Freud (1856–1939). The origin of psychoanalysis is often traced to Freud's *Interpretation of Dreams* (1900), but some of its key elements can be found in his earlier writings, especially his *Studies on Hysteria* (1895), cowritten with Josef Breuer.

FREUDIAN THEORY

"Freudian theory" is not a single theory but a set of smaller ones, at least some of which are familiar to most philosophers and educated people. One of the most significant and best known of these is Freud's theory of dreams.

Freud theorized that all dreams are fulfillments of repressed infantile wishes. During sleep, these repressed wishes can enter into the dreamer's consciousness, but only in a disguised form, after the dream censor has altered their appearance. Freud calls what survives the dream censorship the "manifest content"; what exists prior to the censorship is the dream's "latent content." By having a patient free associate to a dream's manifest content, Freud hoped to determine the dream's latent content and ultimately to glean information about a patient's unconscious conflicts.

In *The Interpretation of Dreams*, Freud distinguishes three areas of the human mind: consciousness, the preconscious, and the unconscious. Consciousness contains all that we are immediately aware of; the preconscious contains mental contents that we can easily become aware of; and the unconscious contains mental contents that cannot be brought to consciousness except through the use of psychoanalytic therapy.

The reason that unconscious ideas cannot readily be brought to consciousness, according to Freud, is that they are repressed. Repression and the unconscious are closely linked in his early writings: "Thus we obtain our concept of the unconscious from the theory of repression. The repressed is the prototype of the unconscious for us" (Freud 1923, p. 15). On his early theory of the dynamic

unconscious, what is repressed is unconscious and what is unconscious is repressed.

On September 26, 1922, however, Freud read a short paper at the Seventh International Psycho-Analytical Congress, "Some Remarks on the Unconscious," in which he indicated dissatisfaction with his theory. In an abstract of the paper (the abstract may have been written by Freud himself), it is noted that the speaker (i.e., Freud) had retold the history of the development of the theory of the unconscious and had pointed out that it had been deemed necessary to equate the dynamic unconscious with the repressed. "It has turned out, however, that it is not practicable to regard the repressed as coinciding with the unconscious and the ego with the preconscious and conscious. The speaker discussed the two facts which show that in the ego too there is an unconscious, which behaves dynamically like the repressed unconscious.…" (author unknown, 1923, p. 367). The two facts are resistance proceeding from the ego during analysis and an unconscious sense of guilt.

This short paper and its abstract anticipated the publication of his *The Ego and the Id* (1923), in which Freud makes another important modification of his earlier views. Here he introduces the expression "*das Es*" ("the it"), which he explicitly borrows from Georg Groddeck; it has been translated by Freud's English translators as "the id." On his new theory, the structural theory, the unconscious is not equated with the repressed. All that is repressed is unconscious, but some of what is unconscious is not repressed. Some of what is in the id is repressed, but some of it is not. In addition, Freud now divides the mind into the id, ego, and superego.

The ego is held to be partly conscious and partly unconscious. It negotiates the demands of the outside world and those of the id and the superego. The id is largely unknowable, according to Freud, but we can know that it exists and know some of its properties. The id is entirely unconscious; it seeks satisfaction only of its instinctual needs, and it is the source of much psychic conflict. The superego develops out of the ego and maintains a system of ideals, values, and prohibitions.

At first, Freud tended to equate repression and defense, but in later works he classifies repression as but one type of defense. Other defense mechanisms include projection, reaction formation, sublimation, isolation, and regression. Despite his work on these other types of defenses, Freud still held that repression was the most important type of defense. In fact, he saw repression as the "cornerstone" of the whole structure of psychoanalysis.

Freud appealed to repression as the important causal determinant of parapraxes, which include memory mistakes, slips of the tongue, and neuroses, although not all neuroses. What Freud called "actual neuroses," including anxiety neuroses and neurasthenia, are caused by events in later life and are not explainable by Freudian theory. What Freudian theory does purport to explain are the "psychoneuroses," such as obsessional neurosis, hysteria, and depression. The psychoneuroses are said to arise from the repression of erotic wishes; their symptoms are "compromise formations"—they represent a solution to unconscious conflicts among the id, ego, and superego.

Another significant Freudian theory concerns sexual stages of development. Each of us, it is theorized, goes through four such stages. In the first year, the infant passes through the oral stage, during which its mouth is its primary source of pleasure. The focus then changes in the anal stage, where, during the next three years or so, the infant's interest shifts to its anus. From three to five years, the child passes through the phallic period, and its genitals are of major interest. There is then a latency period lasting until puberty, when an interest in sex reemerges.

How a child reacts to events during the various stages of sexual development can help determine its adult personality. Both Freud and his followers theorized that certain personality clusters, such as obstinacy, parsimony, and orderliness, were causally linked to specific events in one or other of the infantile stages of development.

In addition to the four stages of sexual development, Freud postulated another stage, the oedipal phase, lasting roughly from age three to five years. During this period, the male child unconsciously desires to possess his mother sexually, but because of perceived threats from his father, the child develops what Freud terms the "oedipal complex." The boy begins to fear that his father will cut off his penis and develops castration anxiety. Freud's first published discussion of castration anxiety occurs in his discussion of the case of Little Hans (1909), whose mother told him that if he continued to touch his penis, she would ask the doctor to cut it off.

Some Freudians postulate in little girls a complex analogous to the oedipal complex, the "electra complex." It was Carl Jung, however, not Freud, who introduced this concept. Freud himself doubted that the concept was useful and even that the phenomenon occurred (Freud 1931, p. 229).

PSYCHOANALYTIC THERAPY AND METHOD OF INQUIRY

Standard psychoanalysis, or "analysis," has certain features that distinguish it from other types of psychotherapy. The analysand, the patient, reclines on a couch, while the analyst remains out of sight. The therapy is scheduled for four or five times per week, and, in contrast to short-term psychoanalytically oriented psychotherapy, it typically lasts three years or more.

Some analysts distinguish three phases of an analysis: the beginning, middle, and end. In the beginning phase, the analyst has a preliminary consultation with the patient, sets the fee schedule, determines (in consultation with the patient) the days and hours of analysis, and decides whether the patient is a suitable candidate for analysis. The initial phase of the analysis can last from a week or two to several months.

In the middle phase, which can last years, the heart of the analysis take place. The patient is instructed to report childhood memories and dreams and to free-associate about their contents, saying whatever comes to mind without pause or hesitation. The main methods of inquiry used by the analyst are free association and interpretation. The analyst uses the observed data from the patient's free associations in conjunction with other observations from the therapy sessions to form preliminary interpretations, or hypotheses, about the cause of the patient's problems. These hypotheses are modified as the analysis progresses, with the analyst taking into account some very important factors that tend to emerge later in the analysis.

One of these factors is resistance. Freud explained the resistance that eventually emerges as the attempt to defend against remembering what has been repressed. Resistance can take different forms, including certain verbalizations, expressions of recalcitrant attitudes, and the unwillingness to free-associate. In his *Inhibitions, Symptoms, and Anxieties* (1926 [1925], p. 159), Freud refers to the patient's resistance as "the resistance of the unconscious."

The solution to the resistance problem is to let the patient "work through" the resistances:

> One must allow the patient time to become more conversant with this resistance with which he has now become acquainted, to *work through* it, to overcome it, by continuing, in defiance of it, the analytic work according to the fundamental rule of analysis. Only when the resistance is at its height can the analyst, working in common with his patient, discover the repressed instinctual impulses which are feeding the resistance; and it is this kind of experience which convinces the patient of the existence and power of such impulses. The doctor has nothing else to do than to wait and let things take their course, a course which cannot be avoided nor always hastened" (1914, p.155).

Another significant factor, one of "undreamt-of importance" (Freud 1940, p. 174), is transference. In the course of the analysis, the patient comes to see the analyst as the reincarnation of some important figure in his or her past and "transfers" to the analyst the negative or positive feelings formerly directed to the figure from the past. An important part of the analysis consists of the analyst's attempt to analyze the overt manifestations of the patient's transference in order to reach a final interpretation of the patient's problems.

In the third and last phase of the analysis, the final interpretation is revealed to the patient: The repressed is made conscious. Yet no mere telling of the interpretation is likely to have any lasting therapeutic effect unless the ego has been strengthened enough to enable the patient's acceptance of the interpretation.

In *Analysis Terminable and Interminable* (1937), Freud gives two criteria for terminating the analysis: first, symptom relief, with the patient overcoming his anxieties and inhibitions; and, second, the analyst's judging that so much material has been made conscious and so much resistance conquered that there is no need to fear a repetition of the pathological processes that caused the patient's problems. These criteria are relevant to deciding in *one sense*, Freud says, if there is to be "the end of an analysis," but in another sense, more is required. In asking whether the analysis is at an end in this second sense, we are asking whether the analyst has had such a far-reaching influence on the patient that no further change could be expected to take place in him if his analysis were to continue. "It is as though," Freud writes, "it were possible by means of analysis to attain to a level of absolute psychical normality," as though the analyst had succeeded in resolving every one of the patient's repressions (1937, pp. 219–220).

The above material contains the main outlines of Freud's most important theories, his method of inquiry, and his therapy, but not all of his theories are covered, and important details are necessarily omitted. For brief discussions of additional psychoanalytic concepts, see B. Moore and B. Fine (1990); for more detailed discussions

of nearly all of Freud's theories, the history of the psychoanalytic movement and its development in countries around the world, and the contributions of other major figures to the development of psychoanalysis, see E. Erwin (2002).

FISSURES IN THE MOVEMENT

In the early years of the psychoanalytic movement, two serious schisms occurred: Alfred Adler (1870–1937) broke with Freud in 1911 and, at approximately the same time, Carl Jung (1875–1961) began fighting with Freud and in 1914 resigned from the International Psychoanalytical Association. These figures disagreed with Freud about several matters, but especially about the theoretical importance placed by Freud on infantile sexuality.

After breaking with Freud, Adler went on to develop his own general psychology. One of his key ideas is that the psychologically disturbed individual suffers from extreme feelings of inferiority. One of the main goals of Adlerian therapy is to eliminate this feeling of inferiority and to put in its place a feeling of community and connectedness with others. Carl Jung also developed his own type of psychotherapy and along with it a rich and complex theoretical framework that included the postulation of the collective unconscious, his theory of archetypes, and his distinction between "extroverts" and "introverts."

One could view the theorizing of Adler and Jung as taking psychoanalysis in new directions, but their theories are so radically different from Freud's that it is doubtful that either's theory or therapy is a form of psychoanalysis at all. When Adler left—or rather was pushed out of—the Vienna Psycho-Analytical Society, he started his own group, "The Society for Free Psycho-Analysis," but he quickly changed the name of his theory to "Individual Psychology," a step for which, Freud said, "we are all thankful." ("There is room enough on God's earth, and anyone who can has a perfect right to potter about on it without being prevented; but it is not a desirable thing for people who have ceased to understand one another and have grown incompatible with one another to remain under the same roof." [Freud 1914, p. 52]). Jung, like Adler, also did not characterize his theory or therapy as a form of psychoanalysis; he preferred the name "analytical psychology."

Long after the departure of Adler and Jung, other cracks developed in the psychoanalytic movement, but these were much smaller. One of the first of these resulted from the work of Melanie Klein, a Budapest psychoanalyst who in 1926 moved to London, where she continued her work analyzing children. Klein saw herself as continuing Freud's work, although she did depart from his theories in certain respects, such as postulating the occurrence of oedipal conflicts in little girls and at an earlier time than specified by Freud's theory. Klein claimed to have made a series of important discoveries about infants, including their having a terrifying mental life, populated by beasts and monsters, and having cannibalistic urges causally linked to earlier contact with the mother's breast. Anna Freud, also working in London at the same time, strongly disagreed with some of Klein's theorizing and managed to win the support of the Vienna Psycho-Analytical Society in condemning Klein's views. The result was a bitter dispute between the "Kleinians" and London psychoanalysts who sided with Anna Freud.

A second division occurred because of the development of ego psychology, the groundwork for which was laid first by Freud's *The Ego and the Id* (1923) and developed further by Anna Freud in her work *The Ego and the Mechanisms of Defense* (1946 [1936]). Ego psychology began to flourish within the psychoanalytic tradition with the publication of Heinz Hartmann's *Ego Psychology and the Problem of Adaptation* (1958 [1939]). Hartmann and his colleagues did not see themselves as breaking with the Freudian tradition in any serious way, but they placed far more emphasis than did Freud on the role of the ego, while greatly reducing the theoretical significance of the id and superego.

Two further theoretical sharp turns occurred in the second half of the twentieth century with the development of object-relations theory and self psychology, now two of the most dominant forms of psychoanalysis.

Object-relations theory developed out of the work of British psychoanalysts, among them Melanie Klein, W. D. Fairbairn, and D. W. Winnicott. This theory is also associated with the work of psychoanalysts living in the United States, such as Otto Kernberg. According to traditional Freudian theory, there exists in each individual biological, instinctual urges, the mental representation of which are referred to as "drives." There are two sorts of drives: the sexual drive and the drive for self-preservation. Object-relations theorists reject Freud's biologically oriented drive theory and argue that the infant is motivated not by instinctual urges but by the need to relate to another person, such as the mother. Freud, like the object-relations theorists, also used the term *object* in his discussion of infants, but he was referring not to people or things external to the infant but to the child's mental representation of them.

In contrast, object-relations theorists tend to refer to things or persons in close proximity to the infant as

"objects," but, somewhat confusingly, the theory also talks of "internalized objects," which clearly are not objects in the external world. One leading theorist, W. D. Fairbairn, in his *Psychoanalytic Studies of the Personality* (1952, p. 137) distinguishes between "objects" and "internalized objects" in terms of a contrast between normal and pathological psychology. In the object-relations theory, psychology becomes "the study of the relationships of the individual to his objects, whilst, in similar terms, psychopathology may be said to resolve more specifically into a study of the relationships of the ego to its internalized objects" (Fairbairn, 1952, p. 137).

In self psychology, the key theoretical concept, that of a "self object," also has a double use; it is sometimes applied to persons and at other times to their mental representations. Self psychology was developed by Heinz Kohut and his colleagues. Kohut became known for his theory of the narcissistic personality disorder, said to have a different etiology from the "transference neuroses" talked about by Freudian theory. This disorder, Kohut claimed, can be recognized partly by observing its distinct symptoms and partly by analyzing the different types of transference that develop in the course of the analysis: a mirroring, idealizing, and twinship transference. Each of these transferences reflects the failure of a parent to respond adequately to a different type of need of the infant, such as the child's need to confirm its own sense of greatness (the need for a "mirroring" response) or the need to experience others who resemble it (the need for a "twin" response). The result of these failures to respond is the narcissistic pathology, the subsequent failure of the narcissistic person to develop an intact self.

THE CURRENT STATUS OF PSYCHOANALYSIS

Freud's theorizing has had an enormous influence on psychiatry, clinical psychology, art, cinema, literature, religion, anthropology, history, biography, sociology, and philosophy. The remnants of his theorizing survives through the work of individual psychoanalysts and the work of the psychoanalytic institutes and associations that exist in the United States, Great Britain, Brazil, Sweden, Finland, Mexico, South Africa, France, Austria, and in many other countries.

The work of the breakaway theorists, Alfred Adler and Carl Jung, has been considerably less popular than Freudian theory, but their theories have nonetheless been influential and are still accepted by many. Many Adlerians belong either to the International Association for Individual Psychology or to the North American Society of Adlerian Psychology. There are also Alfred Adler institutes and schools in Chicago, San Francisco, Washington, New York, and other cities.

Many adherents of the theories of Jung belong to the International Association for Analytical Psychology. C. G. Jung institutes and societies are located in this country in New York, Seattle, Portland, Boston, Los Angeles, and other large cities, and in Canada, Australia, Great Britain, and other countries.

The continued influence of various psychoanalytic theories is important, but there is also the question of truth: How much of psychoanalytic theorizing is at least approximately true? Some of the things that Adler and Jung said were rather commonsensical and not controversial or original. If we subtract these propositions, how many of their distinctive and original claims have been shown to be true? Not very many. There are few, if any, formal empirical studies of their theories. The verdict must be that their theories remain little more than interesting but unproven conjectures.

The work of the ego psychologists, the object-relations theorists, and the self psychologists has been the subject of more empirical inquiry, but there is nothing that can be said to constitute a firm body of supporting evidence for any one of these modifications of Freudian theory. This fact has led one prominent psychoanalyst to point out that the developments in ego psychology were not prompted by new data in the psychoanalytic situation but by the recognition of obvious deficiencies in Freudian theory, and that none of these three theories has remedied the epistemological and methodological difficulties associated with Freudian theory (Eagle 1993).

The sheer quantity of the empirical evidence for Freudian theory and therapy is far greater than that of its newer psychoanalytic rivals. It includes not only Freud's case studies but also the published case studies of many of his followers, data from anthropology and the "psychopathology of every day life," and more than 1,500 experimental studies. There are also Freud's arguments to consider: They are designed to show that even without the benefit of controlled studies, his theories receive powerful support from the data obtained from psychoanalytic case studies.

In evaluating the Freudian evidence, one issue concerns its subject. There is a watered-down, commonsensical version of Freud's theories and there are the original, distinctively Freudian versions articulated and modified over the years principally by Freud himself. On the watered-down version, the unconscious exists if a person

has mental states that exist below the threshold of consciousness, whether or not these states can be brought to consciousness without the aid of psychoanalysis. Repression is said to occur whenever one tries to keep something painful out of consciousness, which obviously happens when one tries to forget a sad love affair or a hurtful insult. There are "Freudian slips," it is said, if people make linguistic mistakes with sexual innuendoes, regardless of what causes the errors. Defense mechanisms such as "projection," "reaction formation," and "displacement" are said to be operative so long as certain types of defensive behavior are displayed, such as attributing to others one's own faults or doing just the opposite of what one would like to do, no matter what causal mechanism explains the behaviors.

The evidence for some of the best-known hypotheses of the popularized, watered-down version of Freudian theory is quite strong but not new: the evidence for some sort of unconscious mind, intentional forgetting, slips of the tongue, and defensive behaviors was known to psychologists and philosophers of the nineteenth century, before Freud invented psychoanalysis. Recent historical research has shown that even many of Freud's seemingly distinctive ideas were anticipated, not merely in some vague way but in detail, by the philosophers Arthur Schopenhauer (1788–1860), Friedrich Nietzsche (1844–1900), Eduard von Hartmann (1842–1906), and J. F. Herbart (1776–1844) (Zentner 2002).

If we limit the discussion to what is distinctively Freudian, scholars still disagree about what the evidence shows. Some still claim that the evidence gleaned from clinical case studies strongly supports some parts of Freudian theory, although that view is losing adherents even among Freudians, partly or largely due to the trenchant and systematic criticisms of the Freudian clinical evidence by the philosopher Adolf Grünbaum (2002). If, as Grünbaum argues, the clinical evidence has little probative value with respect to Freudian theory, that leaves mainly the Freudian experimental evidence, said by some to firmly support some central parts of the Freudian corpus (Kline 1981, Fisher and Greenberg 2002). Another review of the very same experimental evidence concludes that it provides almost no support for any distinctively Freudian hypothesis (Erwin 1996).

As regards Freud's therapeutic claims, there are uncontrolled case studies and correlational studies of long term orthodox psychoanalytic therapy, but there has never been a randomized clinical trial studying its effects. Two retrospective studies of long-term psychoanalysis have been published in recent years; some analysts argue that despite their lack of controls, they provide support for the effectiveness of psychoanalysis because of the employment of novel statistical techniques or the presence of other features that obviate the need for experimental controls. These studies and the claims on their behalf are criticized in Erwin (2002).

See also Freud, Sigmund.

Bibliography

WORKS BY FREUD

All references to Freud's works are to *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. 24 vols. Translated by James Strachey. London: Hogarth Press, 1953–1974.

Breuer, J., and S. Freud. *Studies on Hysteria* (1895). *S. E.* 2, 19–305.

The Interpretation of Dreams (1909). *S. E.* 4 and 5: 1–627.

"Analysis of a Phobia in a Five–Year–Old Boy" (1909). *S. E.* 10: 5–149.

"On the History of the Psycho-Analytic Movement" (1914). *S. E.* 14: 3–86.

The Ego and the Id (1923). *S. E.* 19: 12–63.

Inhibitions, Symptoms and Anxiety (1926). *S. E.* 20: 87–172.

"Female Sexuality" (1931). *S. E.* 21: 225–243.

Analysis Terminable and Interminable (1937). *S. E.* 23: 216–253.

An Outline of Psychoanalysis (1940). *S. E.* 23: 174.

WORKS ABOUT FREUD

Freud, A. *The Ego and the Mechanisms of Defense*. New York: International Universities Press, 1946 [1936].

Author Unknown. Abstract. *Int. J. Psycho-Anal.*, 1923, 4: 367.

Eagle, M. "The Dynamics of Theory Change in Psychoanalysis." *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, edited by J. Earman, A. Janis, G. Massey, and N. Rescher. Pittsburgh: Pittsburgh University Press, 1993.

Erwin, Edward. *A Final Accounting: Philosophical and Empirical Issues in Freudian Psychology*. Cambridge, MA: MIT Press, 1996.

Erwin, Edward. *The Freud Encyclopedia: Theory, Therapy, and Culture*. New York: Routledge, 2002.

Fairbairn, W. D. *Psychoanalytic Studies of the Personality*. London: Tavistock Publications and Routledge & Keegan Paul, 2002.

Fisher, S., and R. Greenberg. "Scientific Tests of Freud's Theories and Therapy." In *The Freud Encyclopedia: Theory, Therapy, and Culture*, edited by Edward Erwin. New York: Routledge, 2002.

Grünbaum, A. "Critique of psychoanalysis." In *The Freud Encyclopedia: Theory, Therapy, and Culture*, edited by Edward Erwin. New York: Routledge, 2002.

Grünbaum, A. *The Foundation of Psychoanalysis: A Philosophical Critique.* Berkeley: University of California Press, 2002.

Hartmann, H. *Ego Psychology and the Problem of Adaptation*. New York: International Universities Press, 1958 [1939].

Kline, P. *Fact and Fantasy in Freudian Theory.* London: Methuen, 1981.
Moore, B., and B. Fine, eds. *Psychoanalysis: The Major Concepts*. New Haven: Yale University Press, 1995.

Edward Erwin (2005)

PSYCHOANALYSIS, EXISTENTIAL

See *Existential Psychoanalysis*

PSYCHOANALYTIC THEORIES, LOGICAL STATUS OF

Since psychoanalysis fails to conform to currently accepted methodological models, its prominence on the contemporary scene constitutes a challenge to the methodologist. He must either revise his canons or show the psychoanalyst the error of his ways. Both tacks have been tried, but thus far the second has predominated. This entry will be confined to methodological problems raised by psychoanalytic theory, though as we shall see, such problems cannot be pursued very far without running into questions concerning the clinical interpretation of particular cases.

CONTENT OF PSYCHOANALYTIC THEORY

Within psychoanalytic theory there are diverse strands, and the relations between them are by no means obvious. For one thing, there are theoretical ideas at different levels. Fairly close to actual clinical practice are found the concepts of repression, regression, projection, reaction formation, and transference. At a higher level there is a theoretical model of the mind in terms of psychic energy, which gets attached to various ideals, the transformations of which are governed by quasi-mechanical principles. This is, in fact, designed to be a perfectly general model of the mind, in terms of which, in the last analysis, all psychological processes and states may be conceived. At this level we have also the division of the psyche into the three systems—id, ego, and superego—together with an account of their properties and interrelations.

In addition to the distinction between levels, we have the distinction between developmental and dynamic theories. In the first group is the theory of psychosexual stages—oral, anal, genital—according to which there is a biologically determined order, beginning from infancy, in which first one, then another, area of the body is maximally sensitive to pleasurable stimulation and according to which certain personality traits predominate as one or another stage is prolonged or transcended only with difficulty. For example, passivity and lack of initiative are associated with the oral stage, during which sensuous pleasure comes mostly from taking things into the mouth.

By contrast, the dynamic theories have to do with processes that take place, or can take place, over a short span of time or at least within the same stage of a person's life. Under this heading we have, for example, the theory of defense mechanisms, according to which the person will defend himself against dangerous impulses by various devices—going to the other extreme (reaction formation), attributing the impulses to someone else (projection), and so on. One of the reasons that the distinction between developmental and dynamic theories is important is that many of the philosophical difficulties raised about psychoanalytic theory center on the notion of unconscious psychic processes, and such processes are more central in dynamic than in developmental theories.

In order to have something fairly definite to work with, let us take the following to be an oversimplified formulation of the psychoanalytic theory of psychic conflict, which is basic to all the dynamic theories.

(1) When it is very painful for a person to be aware of the fact that he has a certain desire, he represses it (prevents it from becoming conscious). The pain may stem from a severe conflict between the desire and the person's standards for himself, from fear of the consequences of attempts to satisfy the desire, or from both.

(2) Repressed psychic material exhibits primitive, infantile features. These include the lack of sharp distinctions, which is in turn conducive to the formation of strong associations between a certain desire and many other, often irrelevant, things and a tolerance for lack of realism and for incompatibility of one's desires and thoughts.

(3) A repressed desire (which continues to exist as a desire) can be partially satisfied by happenings, in actual occurrence or in fantasy, which are associated with the object of the desire.

(4) When the substitute satisfactions themselves arouse too much anxiety, the person seeks to ward them off, often in equally derivative ways.

This basic theory is then applied to the explanation of dreams, slips of the tongue, and neurotic symptoms by studying the ways in which such phenomena constitute substitute satisfactions of repressed desires and/or defenses against such satisfactions.

An illustration of these ideas is presented by Sigmund Freud in Lecture 17 of his *General Introduction to Psychoanalysis.* A girl has, for obvious reasons, repressed a strong desire for sexual intercourse with her father. In the unconscious, various things happen to this desire and the ideas involved in it. The dread of carrying out the act generalizes to a dread of sexual activity of any sort. An association is formed between sexual intercourse and breaking a vase. The bolster at the back of the bed is pictured as the girl's father and the back of the bed as her mother. The pressure of this repressed material becomes so great that the girl develops a compulsion to go through an elaborate ritual before going to sleep at night. She arranges the vases in her room so that breakage is impossible, thus symbolically guarding against sexual intercourse, and she takes care lest the bolster touch the back of her bed, thus achieving a substitute satisfaction for her desire to keep her father and mother apart.

METHODOLOGICAL PROBLEMS

Some of the philosophical objections to psychoanalytic theory can easily be shown to have little or no force. For instance, some philosophers object that the theory postulates unobservable entities; others believe that it is self-contradictory to speak of unconscious mental processes, for what is mental is, by definition, conscious.

In answer to the first objection, it can be pointed out that this practice is common in the most respectable parts of science. Electromagnetic fields and energy quanta are as unobservable as unconscious fantasy. They are, nonetheless, scientifically legitimate because of the functions performed by the theories embodying them, a point to which we shall return. In answer to the second objection, it may be admitted that psychoanalytic theory involves some stretching of such terms as "desire" and "thought" (as in the unconscious thoughts believed to underlie the conscious content of a dream). But, again, this is standard practice in scientific theorizing. The submicroscopic particles postulated in the kinetic theory of gases are modeled on familiar physical objects, like baseballs, except that they lack some of the properties of baseballs, like color and texture, and they possess perfect elasticity. One may as well say that it is a contradiction to speak of physical particles that have no color. Difference from familiar concepts is not in itself fatal. Again, the crucial question is what can be done with the concepts thus derived.

The serious difficulties emerge when we try to determine whether psychoanalytic concepts have the kind of status that is required for scientific validity and fruitfulness. This problem has two closely related parts. (1) Do psychoanalytic terms have any empirical significance, and if they do, how can it be exhibited? (2) How can theoretical principles couched in these terms be put to an empirical test? These questions become two sides of the same coin if we make certain assumptions that are widely shared by contemporary philosophers of science. First, a term has the kind of semantic status required for science if and only if statements in which it figures have implications for what would be experienced under certain circumstances. Second, one brings out a term's empirical or scientific significance, as contrasted with its pictorial associations, by tracing out such implications. Third, it is only if statements have such implications that they can be put to an empirical test. Given these assumptions, we can deal with the two questions simultaneously. By showing how statements involving the term *repress* give rise to implications of a sort that make an empirical test possible, we will at the same time be showing what scientific significance the term has over and above any of its pictorial associations—for example, a man firmly clamping a lid down on a pot of molten metal. With this equivalence in mind, the following discussion will be explicitly directed to the second question: How can the theoretical principles of psychoanalysis be empirically tested?

There is a commonly accepted doctrine, largely derived from a consideration of physics, according to which a theory involving unobservables gets empirical significance by virtue of the fact that it, together with subsidiary assumptions, implies various general lawlike hypotheses that can be directly tested empirically. In this way the theory can be assessed in terms of the extent to which it succeeds in explaining and unifying a variety of lower-level laws that have been empirically confirmed and, on the negative side, the extent to which it does not imply lower-level hypotheses that have been empirically disconfirmed. The Bohr theory of atomic structure, which represents an atom as a sort of miniature solar system with electrons revolving in orbits around the nucleus, cannot be tested directly, for an individual atom cannot be observed. However, from the theory we can derive a variety of testable hypotheses—for instance, those concerning the constitution of the spectrum of the light emitted from a given element.

DERIVING TREATABLE HYPOTHESES

One might well expect to have difficulty deriving testable hypotheses from psychoanalytic theory. The theory represents the postulated unconscious processes mediating between events that are accessible to either introspection or observation, just as do unobservable processes within the atom in the Bohr theory.

In a typical sequence we start with conscious Oedipal desires in a child. Tentative attempts at satisfaction of the desires are met with violent opposition, and as a result the child builds up strong fear and/or horror of the realization of the Oedipal desires. Thus far, everything is, in principle, directly accessible to one or more observers. Then, according to the theory, the complex of desires, fears, and guilt is repressed, whereupon it undergoes various transformations, the exact nature of which is influenced by things that happen to the person, these things again being directly observable. In particular, the associations formed in the unconscious are largely determined by conscious experiences of the person. Finally, the unconscious complex is manifested in various ways—dreams, memory failures, slips of the tongue, compulsions, obsessions, psychosomatic illnesses—all of which are again accessible to experience. This being the case, one would suppose that the theory would yield general hypotheses to the effect that whenever strong desires of a certain kind are met with strong internal and/or external opposition, then (perhaps with the further assumption of certain kinds of intervening experiences) abnormal symptoms of certain kinds will be forthcoming. In other words, since unconscious psychic processes are supposed to provide connecting links between observables, a theory about them should imply that certain antecedent observables would lead to certain consequent observables.

In fact, however, we find little of this. Some attempts have been made to derive hypotheses about statistical distributions from parts of the theory. For example, the theory of dreams holds that dreams partially satisfy repressed desires by representing them as satisfied. It would follow from this that if a group of people were prevented from dreaming for several nights, they would then show a higher average level of tension than a control group. This hypothesis has been tested, using eyeball movement as a criterion of the occurrence of dreams. Most efforts of this sort have stemmed from relatively peripheral components of the theory; in particular, virtually nothing has been done to derive testable hypotheses specifying sufficient conditions for the occurrence of abnormal symptoms. It is only if this were done that the theory could be used for the prediction of such phenomena. Perhaps this is because of the psychoanalyst's preoccupation with the treatment of particular cases rather than with controlled testing of general hypotheses.

There are other features of the situation that also make the formulation of testable hypotheses extraordinarily difficult. Psychoanalytic theory has not been developed to the point where one can give sufficient conditions for one outcome rather than another even on the theoretical level of unconscious processes. Repression is said to occur when a desire arouses great anxiety, but just how much anxiety is required? Obviously, the amount is crucial, but the measurement problem has yet to be solved. Again, given a certain level of anxiety aroused by Oedipal desires, repression is not the only possible outcome. There might, instead, be a regression to the oral or anal phase, or the libido might be redirected into homosexual channels. There are some suggestions about what makes the difference—for example, if one never fully outgrew an earlier stage, this makes regression more likely. But at present this is all rather loose.

Moreover, once repression has occurred, the repressed material may develop in a great many different ways. The fear of sexual contact with the mother may or may not generalize, and if it does, it may generalize along various dimensions. Thus, the person may develop a dread of sexual contact with anyone or only with anyone who is like his mother in some respect. A part of the complex may come to be associated with things that have little or no intrinsic connection with it, as the girl in the example cited above formed an association between sexual intercourse and the breaking of a vase. It may well seem impossible to develop principles that would take into account all the determinants of unconscious trains of thought in a way that makes possible, in principle, the prediction of such associations. This impression is reinforced by the fact that these associations are often powerfully influenced by the person's external experiences, which could not be predicted on the basis of psychological facts about him. Thus, in the above example the girl had once broken a vase and cut her finger, which had bled profusely, an incident that then was associated in her mind with the bleeding accompanying defloration.

But even if connections were strong on the level of unconscious processes, there would still remain the job of formulating sufficient conditions for the occurrence of the ultimate facts to be explained. One and the same unconscious complex, given our present powers of discrimination, may issue in a phobia, hysterical paralysis or anesthesia, obsessive concern over bodily symptoms, or a generalized feeling of unworthiness, to mention only a

few possibilities. No doubt the choice of symptom is due to other factors, but the problem has not been investigated sufficiently to yield even promising general hypotheses.

BACKGROUND FOR CLINICAL INTERPRETATION

In view of the extreme difficulty of empirically verifying psychoanalytic theory, one might ask why it should be regarded as anything other than an imaginatively satisfying fantasy. Why does it seem to have an empirical foundation? The answer is that it has significant connections with empirical facts but not connections of the sort insisted on by philosophers of science who take their models from physical theory. Psychoanalytic theory has grown out of the clinical treatment of neurotics, and in that context it has the function of providing suggestions for the interpretation of particular cases. Thus, if we are dealing with a compulsion neurosis, the theory tells us that compulsive behavior simultaneously provides substitute satisfactions for repressed desires (through the realization of states of affairs unconsciously associated with the realization of the desires) and guards against the arousal and/or satisfaction of the desire. (See the clinical case described above.)

Furthermore, the theory tells us what kinds of desires are most often repressed—incestuous, homosexual, aggressive. Also, psychoanalytic theory is associated with certain techniques—the analysis of dreams, of free associations, and of reactions to the analyst—for ferreting out repressed material in particular cases. Thus, the theory provides leads for the analyst. Insofar as it has this function rather than that of explaining and unifying testable hypotheses about the conditions under which, in general, we will get one outcome rather than another, it is no defect that it is largely made up of rather loose statements about what can happen, given certain conditions, and what can be responsible for a given symptom. In explaining an event, *E,* that has already occurred, our needs are simpler than when we are engaged in predicting or establishing general principles. In retrospective explanation we can take advantage of our knowledge that *E* has already occurred; we are reasoning backward to its sources. Therefore, provided we have a list of possible causes and some way of telling which of these are present, we have something to go on, even if each statement of possible cause is only to the effect that *C* can result in *E.* If we were setting out to predict, however, we would need a further specification of the conditions under which *C* will in fact lead to *E.* The knowledge that an unconscious desire for and fear of intercourse with the father, plus an association between intercourse and breaking a vase, can lead to a compulsive tendency to arrange vases so as to minimize chances of breakage is general knowledge of a sort, but not of the sort exemplified by the Newtonian theory of gravitation, in which the general principles enable one to predict one state of the system from any other state of the system.

Thus, one can say that psychoanalytic theory, given the way it has developed up to now, makes contact with empirical reality through being used as a basis for explanations of certain kinds of observable occurrences and that the theory receives empirical support to the extent that such explanations are adequate. To many methodologists this situation is profoundly unsatisfying. If a theory yields predictively confirmed hypotheses, we have a strong indication that contact with something real has been made, for by thinking in these terms, we have succeeded in anticipating the course of nature. But if the theory can provide only suggestions for retrospective explanations, it is not so clear what this shows. More specifically, many have suspected that the success of psychoanalysts in devising explanations of their patients' symptoms is more a function of the analysts' ingenuity than of the soundness of their theory. It is easy to get the impression that a plausible explanation in psychoanalytic terms could be framed for any behavior, no matter what the facts. If it is not a reaction formation from overattachment to mother, then it is a projection of a self-directed death wish, and so on.

ADEQUACY OF CLINICAL INTERPRETATIONS

Clearly, what is needed is a set of objective criteria for the adequacy of an explanation in terms of unconscious psychic factors, criteria that would permit us to assess a proposed explanation on some grounds other than the way it seems to make sense of the phenomena. If and only if such criteria can be formulated can explanations of particular cases provide any empirical basis for the theory.

Within the limits of this article, we can only touch briefly on the problems involved in formulating and defending such criteria. The problems fall into three groups.

STATUS OF THE DATA. Questions have often been raised about the status of the ultimate data to which the psychoanalyst appeals in justifying an interpretation. These consist of the behavior of the patient, verbal and

otherwise, in therapeutic sessions. Criticisms have been of three sorts.

First, the data actually presented are a small sample of all the behavior engaged in by the patient in the presence of the analyst. We are almost never given any reason for supposing that this is a representative sample, that the analyst has not, perhaps unconsciously, selected those items that best support his hypothesis.

Second, a given patient is rarely, if ever, compared with controls who do not have his difficulties. Without this we cannot show that the data cited have bearing on the abnormalities to be explained. For example, if almost anyone would get annoyed when the analyst acts bored with the session, then the fact that patient *A* does so is not likely to reveal anything that is responsible for any idiosyncrasy of his.

Third, the analyst may often be guilty of contaminating the data through, perhaps unconsciously, tipping the patient off about his interpretation, thus implicitly inviting the patient to produce associations that will support that interpretation.

These are serious problems in data collection and assessment, and they will have to be solved if psychoanalysis is to become more respectable scientifically. But since it seems in principle possible to overcome them, they are less crucial for the logical status of the theory than problems in the other groups.

UNCONSCIOUS CAUSES. An explanation of *E* in terms of *C* is not warranted unless *C* actually exists. What objective tests are there for the actual existence of the unconscious psychic factors appealed to by the analyst? Analysts regularly use a number of detection procedures.

Among the things they consider significant are the following: (1) Patterns of behavior that are as they would be if *A* had a desire of which he is not conscious. For example, a seventeen-year-old girl devotes a great deal of time and energy to the small children of a youngish widower friend of the family, though she is not aware of being in love with him. (2) Patterns of feeling that have the same status. In the same example, the girl gets very depressed when the widower does not send her a birthday present. (3) Analysis of dreams and of free associations. Such analysis proceeds in a rather devious fashion and cannot be illustrated briefly. It is based on the principle that unconscious complexes influence conscious thought and fantasy, including dreaming, by producing relatively safe conscious derivatives of these complexes. (4) Final realization by the patient, after treatment, that he had the desire in question all along.

The inferences involved in the use of these procedures are extremely complex, and it is difficult to say just how conclusively anyone has ever demonstrated the existence of certain unconscious material in a given case. It is worth noting that the use of (3) and (4), unlike (1) and (2), requires the assumption of certain parts of the theory. Thus, for example, we cannot take dreams to reveal unconscious desires in the way analysts do unless we assume that dreams are formed in the manner postulated by the theory. This means that insofar as explanations that are supported in part by dream interpretation are adduced in support of the theory, we are going round in a circle.

UNCONSCIOUS COMPLEXES AND SYMPTOMS. The most difficult problem is that of showing that a given unconscious complex is responsible for certain symptoms. Granted that the girl does have a repressed desire for and dread of sexual intercourse with her father, why should we suppose that this is what led her to develop a compulsive tendency to arrange the vases in her room in a certain way before retiring? In order to answer this question, we shall have to decide what kind of explanation this is supposed to be. Freud often gives the impression that it has the ordinary pattern of an "in-order-to" explanation ("I went into the kitchen in order to get a bottle of beer" or "I went into the kitchen because I wanted a bottle of beer"), except that here the want is unconscious. But the ordinary "in-order-to" explanation carries the assumption that the agent believes that the action in question is, or may be, instrumental in the satisfaction of the want in question. Can we say that the girl unconsciously believed that preventing the vases from breaking would be instrumental in preventing intercourse with her father? A strange belief, but Freud did say that the unconscious is quite illogical. Or should we say, rather, that no belief is involved here but only an association between breaking a vase and intercourse? However this issue is resolved, this assimilation will not help us to justify the explanation, for the fundamental method of justifying an ordinary "in-order-to" explanation—getting a sincere report by the agent of why he did what he did—is not available here.

Freud might claim that an analogue is available—the realization by the patient, after treatment, that that was why she had to arrange the vases as she did. However, if one rests the adequacy of the explanation on the patient's posttherapeutic insight, he leaves himself open to the charge of undue influence on the source of data. Moreover, circularity comes up again, for if the patient came to have this conviction as a result of being presented with

this explanation under hypnosis, this would not count in favor of the explanation. Only insight that comes after certain kinds of therapeutic interactions is relevant, and the claim that insight produced in that way is valid depends on the psychoanalytic theory about the effects that can be expected from psychoanalytic therapy. Thus, there are difficulties in construing the explanation on the model of "I went to the kitchen because I wanted a bottle of beer." On the other hand, if we take as our model an everyday explanation in terms of physical causation, like "The window broke because a baseball hit it," we will have to support it by reference to general principles to the effect that factors of the sort cited have results of the kind we are seeking to explain. And the absence of such tested generalizations in psychoanalysis has already been noted.

Thus, it would seem that before psychoanalytic theory can enjoy a firm empirical foundation, its practitioners must either develop explicit and workable objective criteria for the adequacy of interpretations of clinical phenomena in terms of unconscious factors, or do more to derive testable general hypotheses from the theory, or do both.

See also Dreams; Existential Psychoanalysis; Freud, Sigmund; Psychoanalysis; Psychology; Religion, Psychological Explanations of; Unconscious.

Bibliography

The basic expositions of psychoanalytic theory are to be found in the works of Sigmund Freud. See especially his *General Introduction to Psychoanalysis* (New York: Garden City, 1949), *New Introductory Lectures on Psychoanalysis* (New York: Norton, 1933), *The Interpretation of Dreams* (New York: Macmillan, 1937), *The Ego and the Id* (London: Hogarth Press, 1927), *Inhibitions, Symptoms, and Anxiety* (London, 1961), and the theoretical essays in Vol. IV of the *Collected Papers* (London: Hogarth Press, 1948).

Important later additions include Anna Freud, *The Ego and the Mechanisms of Defence* (London: L. and Virginia Woolf at the Hogarth Press, and the Institute of Psycho-analysis, 1937), and the new developments in ego psychology to be found in two articles by Heinz Hartmann, "Comments on the Psychoanalytic Theory of the Ego," in *The Psychoanalytic Study of the Child,* Vol. V (New Haven, CT: Yale University Press, 1950), and "The Mutual Influences in the Development of the Ego and Id," ibid., Vol. VII (New Haven, CT: Yale University Press, 1952).

Useful expositions are to be found in R. L. Munroe, *Schools of Psychoanalytic Thought* (New York: Dryden, 1955); Otto Fenichel, *The Psychoanalytic Theory of Neurosis* (New York: Norton, 1945); C. S. Hall, *A Primer of Freudian Psychology* (Cleveland, OH: World Publishing, 1954); and David Rapaport, "The Structure of Psychoanalytic Theory," in *Psychology: A Study of a Science,* edited by Sigmund Koch, Vol. III (New York: McGraw-Hill, 1959).

Methodological problems are explored in *Psychoanalysis, Scientific Method, and Philosophy,* edited by Sidney Hook (New York: New York University Press, 1959); B. F. Skinner, "Critique of Psychoanalytic Concepts and Theories," in *Minnesota Studies in the Philosophy of Science,* Vol. I (Minneapolis: University of Minnesota Press, 1956); Antony Flew, "Motives and the Unconscious," ibid.; Gustav Bergmann, "Psychoanalysis and Experimental Psychology," in *Psychological Theory,* edited by M. H. Marx (New York: Macmillan, 1951); Ch. 3 of R. S. Peters, *The Concept of Motivation* (London: Routledge and K. Paul, 1958); a series of essays reprinted in *Philosophy and Analysis,* edited by Margaret Macdonald (Oxford, 1954)—Stephen Toulmin's "The Logical Status of Psycho-analysis," Antony Flew's "Psycho-analytic Explanations," and R. S. Peters's "Cause, Cure, and Motive"; and B. A. Farrell's articles "The Criteria for a Psychoanalytic Interpretation," in *PAS,* supp., 36 (1962)—reprinted in *Essays in Philosophical Psychology,* edited by D. F. Gustafson (Garden City, NY: Anchor, 1964), pp. 299–323—and "Can Psychoanalysis Be Refuted?," in *Inquiry* 4 (1961): 16–36.

Attempts to derive testable general hypotheses from psychoanalytic theory are canvassed in R. R. Sears, *Survey of Objective Studies of Psychoanalytic Concepts* (New York: Social Science Research Council, 1943).

William P. Alston (1967)

PSYCHOKINESIS

See *Parapsychology*

PSYCHOLOGISM

"Psychologism" is the term first used in Germany in the first half of the nineteenth century to designate the philosophical trend defended by Jakob Friedrich Fries (1773–1843) and by Friedrich Eduard Beneke (1798–1854) against the dominant Hegelianism. Fries and Beneke advocated a philosophical position based entirely on psychology. They held that the only instrument philosophical inquiry has at its disposal is self-observation (or introspection) and that there is no way to establish any truth other than by reducing it to the subjective elements of self-observation. Psychology becomes, from this point of view, the fundamental philosophical discipline. Logic, ethics, metaphysics, philosophy of law, philosophy of religion, and philosophy of education are all little more than psychology or applied psychology. Beneke wrote, "With all of the concepts of the philosophical disciplines, only what is formed in the human soul according to the laws of its development can be thought; if these laws are understood with certainty and clarity, then a certain and clear knowledge of those disciplines is

likewise achieved" (*Die Philosophie in ihrem Verhältnis zur Erfahrung*, p. xv).

Fries and Beneke, who viewed Immanuel Kant as their predecessor inasmuch as he defended the "rights" of experience, held, nevertheless, that he was mistaken in wanting to institute an inquiry independent of experience which would arrive at knowledge of the a priori forms of intuition and of the categories and in seeking the transcendental ground of truth—the objective validity of human knowledge. This inquiry, Fries claimed, is impossible. The critique of reason can only be a science of experience based on self-observation (*System der Metaphysik*, p. 110). In the same period Vincenzo Gioberti branded as psychologism all of modern philosophy from René Descartes on. He meant by psychologism the philosophical procedure that claimed to go from man (that is, from experience) to God and contrasted it with ontologism, which is the movement from God to man.

The doctrine defended by Fries and Beneke has some connection with certain aspects of English empiricism from John Locke to David Hume in that in both theories experience is not only the instrument of control and the criterion of the truth of knowledge but also the psychological origin of knowledge itself.

Fries and Beneke were correct in accusing Kant of rejecting psychologism, since he had posited the premises for a critique of any psychologism by distinguishing (in a famous passage in the *Critique of Pure Reason*) the *quaestio facti* of the "physiological derivation" of a priori concepts—that is, of their occurrence in the mind or consciousness of man—from the *quaestio juris* of their validity, which demands as a response the transcendental deduction. This distinction, on the basis of which Kant criticized Locke, who would have answered only the first question, is one of the pivotal points of the whole Kantian doctrine—namely, that the truth of empirical knowledge does not depend on the psychological mechanism but on a priori conditions independent of this mechanism; that the validity of the moral norm does not depend on desires or appetites but is a priori as well; and that the validity of aesthetic judgments is in turn based on taste, an a priori faculty.

Toward the middle of the nineteenth century, psychologism was defended in the very field in which it would seem most foreign—logic and mathematics. In John Stuart Mill's *A System of Logic* it is explicitly stated that introspection is the only basis of the axioms of mathematics and the principles of logic; in Mill's *Examination of Sir William Hamilton's Philosophy* logic is classified under psychology and distinguished from it only as the part is distinguished from the whole or art from science. Many logicians in subsequent years accepted this point of view.

The Kantian point of view was developed systematically by Rudolf Hermann Lotze in his *Logik*. The psychological act of thinking is, according to Lotze, completely distinct from the content of thought. The psychological act exists only as a determinate temporal phenomenon, whereas the content has another mode of being—validity. A decade later Gottlob Frege defended the same point of view with regard to mathematics.

> Never take a description of the origin of an idea for a definition, or an account of the mental and physical conditions through which we become conscious of a proposition for a proof of it. A proposition may be thought, and again it may be true; never confuse these two things. We must remind ourselves, it seems, that a proposition no more ceases to be true when I cease to think of it than the sun ceases to exist when I shut my eyes. (*Die Grundlagen der Arithmetik*, introduction)

In the last decades of the nineteenth century, the neo-Kantians argued against the psychologistic presentation of philosophy. The Baden school (Wilhelm Windelband, Heinrich Rickert) defended the independence of values from psychological experience, which could never establish their absoluteness and necessity, and the Marburg school (Hermann Cohen, Paul Natorp) held, similarly, that the validity of science, like that of ethics and aesthetics, does not depend on psychological conditions but on the laws proper to these sciences—that is, on the methodological rules that govern their construction. Cohen and Natorp held, moreover, that "thought" or "consciousness" does not designate a psychic reality subject to introspection but the objectively valid content of knowledge—the totality of the possible objects of knowledge itself and the method used in the development of the sciences.

The systematic critique of psychologism in the fields of logic and mathematics is an important part of Edmund Husserl's *Logische Untersuchungen*. His main objections are that if logical laws were based on psychological laws, then (1) they ought to be, like the latter, vague and approximate, whereas, at least in part, they are so exact that they cannot be guaranteed by an empirical element; (2) they ought to be based, like all empirical laws, on induction, which yields only a probable validity and not the apodictic certainty they manifest; (3) they ought to imply the existence of such psychic events as representation and judgment, whereas they do not con-

cern the reality of psychic life and of other facts (unlike the laws of nature, which are merely probable) but concern necessary relations independently of facts (*Logische Untersuchungen*, Vol. I, Secs. 21–24). Later in his career Husserl wrote, in terms very close to Frege's, "To refer to it [a number] as a mental construct is an absurdity, an offence against the perfectly clear meaning of arithmetic discourse, which can at any time be perceived as valid, and precedes all theories concerning it" (*Ideen*, Sec. 22). He warned against the tendency to "psychologize the eidetic"—that is, to identify essences, which are the authentic objects of knowledge, with the simultaneous consciousness of these essences (ibid., Sec. 61).

The battle between psychologism and antipsychologism is sometimes fought among philosophers with the same point of view. Among the existentialists Martin Heidegger, who adopted as his method Husserl's phenomenology, intended existential analysis as the uncovering of human situations in their essence, not in their psychic occurrence (*Sein und Zeit*, Halle, 1927, Sec. 7), whereas Jean-Paul Sartre, speaking of existential psychoanalysis, seems inclined toward psychologism, although he tried to correct it by affirming that "consciousness is not a mode of particular knowledge but it is the dimension of transphenomenal being in the subject" (*L'être et le néant*, Paris, 1943, p. 17).

Within logical empiricism the argument against psychologism is one of the fundamental points of Rudolf Carnap's first work, *Der logische Aufbau der Welt*. The fundamental theses of *Logische Syntax der Sprache*, especially the principle of tolerance, are incompatible with psychologism, according to which, obviously, there could be only a single language—that determined by psychological laws. Carnap took the same line when he criticized Bertrand Russell's thesis that propositions are mental events in "Empiricism, Semantics, and Ontology." Arguments against psychologism occur frequently in the writings of other logical empiricists, though traces of psychologism can be found in the thesis, deriving from Russell and held by many logical empiricists, of the immediate, private, and incommunicable character of the sense data that are at the basis of empirical propositions.

See also Beneke, Friedrich Eduard; Carnap, Rudolf; Cohen, Hermann; Descartes, René; Empiricism; Existential Psychoanalysis; Frege, Gottlob; Fries, Jakob Friedrich; Gioberti, Vincenzo; Hegelianism; Heidegger, Martin; Hume, David; Husserl, Edmund; Intuition; Kant, Immanuel; Locke, John; Logical Positivism; Lotze, Rudolf Hermann; Mill, John Stuart; Natorp, Paul; Neo-Kantianism; Propositions; Psychology; Rickert, Heinrich; Russell, Bertrand Arthur William; Sartre, Jean-Paul; Windelband, Wilhelm.

Bibliography

WORKS FAVORING PSYCHOLOGISM

Beneke, Friedrich Eduard. *Die Philosophie in ihrem Verhältnis zur Erfahrung, zur Spekulation, und zum Leben*. Berlin, 1833.

Fries, Jakob Friedrich. *System der Metaphysik*. Heidelberg, 1824.

Gioberti, Vincenzo. *Introduzione allo studio della filosofia*. Brussels, 1840.

Lipps, Theodor. *Grundzüge der Logik*. Hamburg–Leipzig, 1893.

Lotze, Hermann. *Logik*. Leipzig, 1874. Translated by Helen Dendy as *Logic*. Oxford, 1884.

Mill, John Stuart. *Examination of Sir William Hamilton's Philosophy*. London, 1865.

Mill, John Stuart. *A System of Logic.* 2 vols. London, 1843.

Sigwart, Christoff. *Logik.* 2 vols. Tübingen, 1873–1878. Translated by Helen Dendy as *Logic.* 2 vols. London, 1890.

WORKS CRITICAL OF PSYCHOLOGISM

Carnap, Rudolf. "Empiricism, Semantics, and Ontology." *Revue internationale de philosophie*, 4th year (11) (1950): 20–40. Reprinted in *Readings in the Philosophy of Science*, edited by Herbert Feigl and May Brodbeck. New York: Appleton-Century-Crofts, 1953.

Carnap, Rudolf. *Der logische Aufbau der Welt*. Berlin: Weltkreis, 1928.

Carnap, Rudolf. *Logische Syntax der Sprache*. Vienna: Springer, 1934. Translated by Amethe Smeaton as *The Logical Syntax of Language*. London: Kegan Paul, 1938.

Frege, Gottlob. *Die Grundlagen der Arithmetik*. Breslau, 1884. Translated by J. L. Austin as *The Foundations of Arithmetic*. Oxford: Blackwell, 1950.

Husserl, Edmund. *Ideen zu einer reinen Phänomenologie und Phänomenologischen Philosophie.* Vol. I. Halle: Niemeyer, 1913. Translated by W. R. Boyce Gibson as *Ideas: General Introduction to Pure Phenomenology*. New York: Macmillan, 1931.

Husserl, Edmund. *Logische Untersuchungen.* 2 vols. Halle: Niemeyer, 1900–1901.

Pap, Arthur. *Elements of Analytic Philosophy*. New York: Macmillan, 1949.

OTHER RECOMMENDED TITLES

Bermudez, Jose Luis. "Psychologism and Psychology." *Inquiry* 42 (1999): 487–504.

Cohen, Jonathan. "Frege and Psychologism." *Philosophical Papers* 27 (1998): 45–67.

Davis, Wayne. "Psychologism and Humeanism." *Philosophy and Phenomenological Research* 67 (2003): 452–459.

George, Rolf. "Psychologism in Logic: Bacon to Bolzano." *Philosophy and Rhetoric* 30 (1997): 213–242.

Kitcher, Patricia. "Revisiting Kant's Epistemology: Skepticism and Psychologism." *Nous* 29 (1995): 285–315.

Nicola Abbagnano (1967)
Translated by Nino Langiulli (1967)
Bibliography updated by Benjamin Fiedor (2005)

PSYCHOLOGY

In the development of "psychology," the study of the mental life and activities of animals and men, three phases can be conveniently distinguished—the presystematic, the systematic but prescientific, and the scientific. The presystematic, by far the longest of the three phases, is that in which men observed and reflected on human ways and embodied their reflections in aphorisms, anecdotes, and fables. Presystematic thinking is important since it has been passed down through the ages and is continually augmented by that amalgam of wisdom, superstition, and dogma that those who claim no professional competence like to describe as the fruits of their experience. The presystematic psychology of contemporary primitive groups has been recorded by anthropologists, but little is known of the corresponding ideas of the precursors of the systematic psychology of the European tradition. The doctrines of the pre-Socratic philosophers are transitional.

SYSTEMATIC PHILOSOPHY OF MIND

MIND, BODY, AND NATURE. Systematic psychology began with Aristotle's *De Anima,* which was of outstanding importance at an early stage because it provided a solid, biologically based conceptual scheme. This involved, first, an elucidation of the concept of soul (*ψύχη*) and such related concepts as mind (*νοῦς*), which were regarded as the differentiating properties of the phenomena to be studied. Aristotle's scheme laid down the lines along which the relationship between various manifestations of soul and mind were conceived until the seventeenth century.

Second, life and mind, being closely connected with the functioning of the body, must be conceived of in a way that does justice to the peculiar intimacy of this relationship. Aristotle paid close attention to this relationship.

Third, there is the problem of how the relationship between psychological phenomena and other phenomena of the natural world is to be conceived. Are psychological concepts and categories of explanation reducible to others? Aristotle, again, was particularly interested in this question because of the attempts of some of his contemporaries and predecessors to show that human behavior fell under the concept of motion, which had a wide applicability in the natural world.

In the exposition of the systematic period of psychology these problems will be employed not simply as a framework for expounding the main lines of Aristotle's system of psychology but also as a framework for picking out the main features of the most important theoretical systems since Aristotle laid the foundation of psychology.

PLATO AND ARISTOTLE. Aristotle (384–322 BCE) insisted on the widest possible definition of soul, thus returning to the pre-Platonic view that soul is virtually the principal of all life. The natural expression for a living thing was *ἔμψυχον σῶμα*—"body with a soul." Aristotle started from the linguistic point that some bodies are so described whereas others are not and asked by what criterion this distinction was made. His answer was that it is life but that there are different levels of life. Intellect, sensation, nutrition, motion, are all forms of being alive. What they have in common, however, is a self-originating tendency to persist toward an end.

This marked both a return to and a great improvement on pre-Platonic views of soul. In early Greek thought soul was thought of simply as that which keeps a man alive and which leaves his body when he dies. It was connected with breathing. Spirit (*θύμος*), on the other hand, was thought of as the generator of movement; it was connected with the movement of the limbs and with emotional states. It was thought of as quite distinct both from soul and from mind, which was regarded as the source of images and ideas. The notion of the soul as a whole of which spirit and mind were attributes emerged only gradually.

Plato (427?–347) tried to combine the concept of the soul as a whole with a stress on the preeminence of mind, which he inherited from Anaxagoras. His account, therefore, of the soul as a whole was constantly confused by the special status that he accorded to mind. In the *Republic* he spoke of the soul as having three parts—reason or mind, spirit, and desire (*ἐπιθυμία*). But he also thought that reason was the defining property of an immaterial substance that survived bodily death whereas spirit and desire passed away with the body. Similarly, in the cognitive sphere he regarded sensation and imagination as inferior to reason and as intimately connected with the body. This represented a fusion of the Orphic belief in the survival of the soul with an exaltation of mathematical reasoning as the only way of obtaining certain knowledge, which Plato took from the Pythagoreans. He thought that in mathematics the soul grasps forms that are eternal and nondeceptive. As like can be known only by like, the soul, in its rational aspect, must also be eternal. Plato's conviction was reinforced by such considerations as those that he adduced in the *Meno,* in which the grasp of mathematical truths was exhibited in an untutored slave. According to

Plato, this indicated that the slave was being made to remember what he had known previous to his embodiment. Thus, Plato's preoccupation with epistemology led him to make a sharp cleavage between the rational and irrational parts of the soul.

Aristotle approached the matter from a biological rather than an epistemological standpoint. Reason, spirit, and desire represented different levels of being alive. To be alive is to possess a self-originating tendency toward an end. This is exhibited at the lowest level in nutrition and reproduction. Thus, plants have a low-grade soul. Animals have sensation, locomotion, and desire superimposed upon nutrition and reproduction. Human beings, in addition, have reason, or mind, by means of which a rule or plan is imposed upon desire. By mind is meant self-direction in accordance with a rational formula.

Aristotle maintained that the lower level of soul is a necessary condition for the higher and that the possession of a higher type of soul also changes the way in which the lower functions. Because humans are rational, they feed, reproduce, perceive, and act in a manner that differs from that of animals.

Soul and body. Plato's view of the special status of reason was plausible at a time when almost nothing was known about the functioning of the brain and nervous system, for abstract thought seems to proceed with little dependence on bodily organs. Furthermore, the identity of a subject of experience through time does not seem to depend entirely on bodily continuity. There is thus a case for Plato's concept of the rational soul as some kind of active agency that inhabits the body for a brief period.

Plato thought that the rational soul inhabits the head because the head is round (the most perfect shape and, hence, an appropriate place for the seat of reason) and the part of the body nearest the heavens. It makes contact with the brain, which was conceived of as a kind of marrow encased in the skull. The irrational soul makes contact with the marrow of the spinal cord in its bony sheath. The better part of the irrational soul, spirit, inhabits the heart and functions in such manifestations of life as energy, courage, and ambition; the worse part, desire, functions below the diaphragm, in appetite, nutrition, and reproduction. The rational and irrational parts affect each other through the liver, which acts as a sort of mirror of thought.

In sleep the soul is shut up, and its motions subside. A few agitations remain, however, and produce dreams. Usually dreams are the expressions of desires that are suppressed—an interesting anticipation of Sigmund Freud's theory of dreams. The good man controls his desires sensibly and so is not unduly disturbed by them in sleep. In the *Republic* Plato also suggested that in sleep the rational soul, if not troubled by irrational desires, can attain truths not otherwise revealed.

Plato thought of sensation as a transmission of motions. The human body receives an impression from without and responds with an inner motion. Some parts of the body—for instance, the hair and the nails—are subject to shock but do not respond with inner movements. Sense organs, however, are good conductors of motion. Thus, hearing, for instance, is the end product of a kind of shock. By means of air in the cavities of the body a blow is transmitted through the ears to the blood and brain and then to the soul. Knowledge does not consist just in sensation but in the activity of the soul in relation to what is thus transmitted. This transmission is complicated by the intervention of memory, imagination, feeling, and association, all of which act as intermediaries between reason and sensation.

Aristotle believed that there was a very intimate connection between soul and body that was a particular case of the more general relationship between form and matter. The soul is "the first actuality of a natural body furnished with organs." He used other examples to illustrate this relationship. If the eye were an animal, he said, eyesight would be its soul, this being the form or capacity of the eye. To speak of soul is to speak of a capacity or propensity to function in a certain way that depends on a certain bodily structure, or it is to speak of the actual exercise of such a capacity or propensity, which is the second kind of actuality. Thus, anger, for instance, can be the appetite of returning pain for pain or the boiling of the blood around the heart, depending on whether the dialectician or the physical scientist is considering it; there is always a biological and a psychological account to be given.

The soul, Aristotle argued, is the cause of the body in three ways. It is its efficient cause in that reference to some concept, such as desire, is required to explain movement. It is the formal cause in that behavior is explained as the exercise of a capacity or tendency. It is the final cause in that reference must be made to "the reason for the sake of which" movements of the body take place. If the behavior is explained by recourse to the rational soul, then plans and rules are imposed on desire. In choice, for instance, means are worked out and adapted to attain an end.

Generally speaking, Aristotle held that soul and body are a particular case of the more general correlatives,

form and matter. When he spoke of theoretical reason rather than practical reason, he suggested that the distinction between matter and form is again exemplified in that reason is both passive and active. But he hinted at another sort of doctrine when he also claimed that active reason comes from without and is divine. It is like a helmsman in a ship. This looks like a concession to the Platonic view of mind.

The details of Aristotle's physiology were carefully related to his idea of the levels of soul. The primary function of the nutritive soul is the absorption of nourishment, but its end is to generate another being like itself. The unity of the species is thus preserved though individual members perish. The stomach was thought of as an oven where animal heat cooks the food and blood in the heart. The heart is the seat of life, sensation, motion, and heat.

Sensation is a discriminative power from which the higher cognitive functions develop. There is the organ, the power to receive sensible forms, and the sense, regarded as constituted of both matter and form. In sensation the sense organ is assimilated to its object—for example, the eye becomes colored. But whereas in nutrition both matter and form of external objects are absorbed, in sensation only form without matter is taken in, like wax taking the imprint of a seal ring. Each sense is sensitive to one or more qualities ranging between extremes. Too little would not register; too much would destroy the organ. This was an application of Aristotle's doctrine of the mean that he developed in relation to moral conduct.

The particular senses are all developments of touch, depending on the intervention of a more refined medium. Taste, for instance, apprehends the savory properties of bodies through the intermediary of moisture; smell, the odorous properties conveyed through the air. In the transmission of sensations to the heart and in the vitality that flows from the heart, the "connatural spirits" play an important role. They were thought of as a kind of inner air quite distinct from the outer air that we breathe. Closely associated with the blood, they acted as a universal internal medium for the transmission of sensation. Besides the specific senses there is *sensus communis,* which is not a sixth sense but a generic power of sensation as such which provides unity for the sensitive soul in its particular manifestations. The ear does not see; however, the man who hears also sees, and some qualities are presented through more than one sense—for example, roundness by sight and touch. By *sensus communis* we also perceive the common sensibles of figure, motion, rest, magnitude, and also what Aristotle called the accidental sensibles, which are the principles of association of ideas—similarity, contiguity, and the like. We also perceive that we perceive through *sensus communis.*

Imagination is a by-product of sensation. Forms provided by sensation are manipulated in the absence of physical objects. Memory is a combination of imagination and *sensus communis.* There is an image of something plus an awareness of its pastness. Recollection is rather different, for it involves the exciting of an image and the release of a whole chain of images joined by habit according to the principles of association. Imagination also provides a link between knowledge and action, for desire presupposes the imagination of an end to be attained. It may be deliberative, if influenced by reason, or merely sensitive. Desire is thus dependent on sensation and thought. In this way Aristotle was able to maintain his three levels of soul by making desire appear at two levels, depending on whether it is rational or irrational.

Psychological and mechanical concepts. Aristotle believed not only that there were certain very general concepts, such as form, matter, and change, which could be applied to everything; he also extended teleological categories of explanation—his ill-fated final causes—to all nature. Nature, he thought, was composed of natural kinds that could be classified by genus and differentia, which all had a natural place, and which all tended toward the realization of their essence. "Nature, like mind, always does whatever it does for the sake of something, which something is its end." Such modes of explanation proved singularly unfruitful when extended to the physical world. But because they were taken from the realm of life, where Aristotle, a marine biologist and the son of a doctor, was particularly acute, they fitted very well, in a general sort of way, that realm of phenomena in which they had their natural home. Aristotle was often accused by later mechanists of being anthropomorphic, but there is not much wrong with being anthropomorphic about men. Indeed, those who later attempted to explain human behavior in mechanical terms applicable to the physical world may well have made the obverse mistake to Aristotle's.

Aristotle himself, in criticizing the mechanists of his day, gave some very interesting arguments to show why the soul, which is the source of movement, cannot itself be moved. Plato had steadfastly claimed that the soul was the source of motion. In a famous passage in the *Phaedo* (98B–99D) he made clear his objection to extending mechanical explanations to cover human conduct. Plato admitted that some kind of physical account could be

given of the movements that led up to Socrates' sitting in his prison cell, awaiting his death. But he scorned the suggestion that this account would be a satisfactory explanation of the situation, for an explanation must include some reference to Socrates' reasons for being there. Plato did not, however, develop elaborate arguments against mechanical theories.

Aristotle, on the other hand, wrote his *De Anima* as part of his systematic attempt to classify the different sciences on the basis of the subject matter with which they were concerned. He was therefore very much concerned both with demarcating the field of application of various families of concepts and with sketching the ways in which they were related to each other. Movement (*κίνησις*) was only a particular type of change. He was most anxious to deny that it was either the only or the fundamental type.

Aristotle argued, first, that a logical mistake is made if the soul as a formal cause is thought of as moved in the physical sense. How can a capacity or tendency be conceived of as moving or being moved? Nor can the actualizations of soul in particular cases be properly conceived of as movements, for in practical thought the processes have unity because they go on for the sake of some end. Their particular type of unity cannot be assimilated to such physical unities as the parts of a spatial magnitude; it is more like the unity of a series of numbers. Reference to an end is a conceptual device for picking out how a series of movements are to be thought of as constituting one action; such an end is not itself an extra movement. In the case, too, of some processes of theoretical thought, such as inferring, "thinking has more resemblance to a coming to rest or arrest than to a movement." The end is, as it were, built into the meaning of the term. "Inferring," "concluding," and even "perceiving" are terms that intimate the attainment of ends or standards that are intrinsic to the processes themselves.

Concept of consciousness. Arguments of the Aristotelian type have been revived in recent times by such philosophers as Gilbert Ryle, who have defended a predominantly Aristotelian concept of mind in opposition to a Platonic or mechanical concept. Such a concept of mind is in keeping with the biological orientation of psychology that followed the impact of Charles Darwin. However, it sprang out of the post-Wittgenstein reaction against privacy as the hallmark of the mental, which had characterized most psychological theories since the time of René Descartes (1596–1650).

It is difficult for modern Western scholars to grasp that the Greeks really had no concept of consciousness in that they did not class together phenomena as varied as problem solving, remembering, imagining, perceiving, feeling pain, dreaming, and acting on the grounds that all these are manifestations of being aware or being conscious. Historically, this emphasis on private experience presupposed the development of individualism as a social movement. The Greeks of the city-states lived in a public world of public feats and public concerns. Their word *ἰδίωτης*, from which we derive the word *idiot,* was a term of disdain for a man who concerned himself only with private matters. Socrates, with his stress on individual self-knowledge and the care of the individual soul, was a moral innovator. With the conquests of Philip and Alexander the Great and the breakup of the small autonomous Greek states, this moral innovation became systematized in the codes of the Stoics and Epicureans. The ideal of individual self-sufficiency developed as a substitute for the much-lauded self-sufficiency of the city-states. Man, it was claimed, was a citizen of the world who should either discipline himself and purify his individual soul (Stoics) or slip through life unobtrusively by cutting down the possible sources of misery (Epicureans). This led to an increase of interest in the will and the emotions and to an emphasis on individual experience.

This turning inward was institutionalized by Christianity, with its stress on personal salvation and the purity of soul. Introspection vied with revelation as a source of knowledge. St. Augustine paved the way for Descartes's first certainty, *cogito ergo sum.* With Descartes the Platonic view of the soul and of knowledge was reinterpreted in the light of the rise of the mathematical sciences, but there was a difference—the stress on the certainty of our knowledge of our own mental states. Mind was no longer simply associated with reason; it was something to which we have private access and whose rational activity it is self-contradictory to doubt. This stress on privacy as a hallmark of the mental was a far cry from Aristotle's view of soul as characterized by a self-originating tendency to pursue an end. A brief mention, however, should be made of some of the intervening systems, though from the point of view of psychological theory, nothing of any great importance happened after the death of Aristotle in 322 BCE until the seventeenth century, when new systems were inspired by the rise of the physical sciences.

STOICS AND EPICUREANS. The Stoics and Epicureans provided an interesting contrast in respect to their views about the relation between soul and the rest of nature. Both attempted a monistic view, but whereas the Stoics reverted to Plato and tried to extend the concept of soul so that it permeated all nature, the Epicureans reverted to

Democritus and extended a mechanical atomistic account of nature to include life and mind.

Stoics. The Stoics thought of everything in the universe as being either active or passive; hence, there was no opposition between dead matter and soul. The ultimate substance is fire, which has different forms at different levels of being, ranging from cohesion at the inorganic level, through growth at the plant level, to life of a rational or irrational type at the animal and human level. Fire is thus the all-pervading principle of activity as well as the reason or regulator of change in the universe. Mental activity as found in men is a concentrated form of the universal reason, creatures being vehicles for the operation of this universal regulation. Hence the Stoic injunction to live according to nature, for in simple instinctive tendencies reason is often manifest in an incorrupted form.

The Stoics believed that the soul of man is a very subtle form of the all-pervasive fire, for the corporeal can be affected only by what is corporeal. The soul is affected by the body; therefore, the soul, too, must be corporeal. It combines heat, mobility, and a high degree of rarefaction. Indeed, it was more or less identified with the "connatural spirits" of Aristotle that course through the body closely associated with the blood, which are transmitted in generation, and which are similar in nature to the warm outer air, which is also essential to life. The breast is the seat of the soul.

Perhaps the most interesting and important contribution of the Stoics to psychology was their application of the Aristotelian categories of activity and passivity, which they thought to be the defining attributes of what is real, to the mind. Mental activity, they held, is characterized by assent (*συγκατάθεσις*), which can be exhibited in perception and memory, as well as in practical and intellectual judgment. This may be justified or erroneous, but truth is natural and error unnatural. When error of a perceptual, intellectual, or practical kind occurs, the explanation is to be sought in the theory of emotions or mental disturbances. Basic to this Stoic account was the notion of impulse, which covered both appetite and aversion and which operates obscurely at the level of sensation as well as at the rational level, when it is transformed into the adoption of ends for action. Emotions are thus unsuccessful attempts at full rational choice. The early Stoics left such failures unexplained; the later Stoics assigned the cause to circumstances and, therefore, to things that are beyond our power. From this came their characteristic emphasis on the assertion of will over adversity, of rational choice over irrational promptings.

Epicureans. The main interest of Epicurean psychology was its anticipation of mechanical theories of the seventeenth and subsequent centuries. Everything, Epicurus (341–270 BCE) believed, was constructed from atoms and, therefore, everything, including minds, could be explained in terms of the mechanical laws governing atoms. The soul differs from other atoms in that it is lighter and more mobile; heat is fundamental to its nature, but it is not identical with fire. It permeates the body like a subtle air and gives it life.

Sensations are effects produced in sense organs by effluxes from objects, differences in sensations being explained in terms of differences in external movements and in the configurations of the underlying atoms. Similarly, ideas are caused by atoms striking the subtle matter of the thinking soul. Incoming impressions set up other motions in the mind, making possible judgment, which is a motion of the mind superimposed upon an impression. Error occurs when impressions are accompanied by irrelevant motions of the mind. The motions of the mind can be linked together to form complex ideas by principles of association. Reason is simply the use of general ideas brought about by the fusion of images into composite pictures.

It is difficult to see how notions such as error and truth could be generated by such descriptions of mere movements of atoms. Indeed, Epicurus did nothing to meet Aristotle's acute criticisms of mechanical descriptions of thought. He did something, however, to meet the charge of fatalism in his notorious doctrine of the swerve of the atom, which was a consequence of the self-motion postulated for all atoms. The power of the mind to incline this way or that constitutes its freedom. People are poised between pain, which is one sort of motion, and pleasure, which is an excessive reaction to pain. Between these two extremes there is an equilibrium, which is more permanently satisfying and which reason can guide men to attain. This he called freedom from disturbance (*ἀταράξια*), which is inseparable from the use of reason.

THEOLOGICAL PSYCHOLOGY. The psychology of the Greeks had always been, in varying degrees, subservient to epistemological and ethical concerns. The account of reason, for instance, or the role ascribed to the passions was a graphic way of presenting solutions to problems about knowledge and conduct. But there was also the Greek passion for speculation about the ultimate nature of things, about the One in the many, and about the status of mind in the universe and its relation to the body. With the coming of Christianity, which brought with it

the biblical account of the creation of the world, this radical metaphysical speculation abated, and the body was seen largely as something that had to be considered as a potent source of temptation. Psychological theory became almost entirely an offshoot of epistemology and ethics, for the supreme purpose of life for thinking men became the knowledge of God and the quest for salvation.

The religious preoccupations of such writers as Plotinus, Clement, and Augustine introduced, of course, a different emphasis into epistemology and ethics. This was manifest before the coming of Christianity in the work of Philo Judaeus (fl. 20 BCE–40 CE), who thought that real knowledge was a possession only of minds that had been so purified that they received divine illumination. Philo was the first systematic thinker to fuse the religious fervor of the Hebrew tradition with a selection from the conceptual schemes of Plato, Aristotle, and the Stoics. Knowledge of God and a divinely sanctioned code of conduct had somehow to be fitted into the speculative schemes of the Greeks. Because neither God nor his purposes are manifest to the senses, increasing importance was attached to inner experience as a way of knowing. Philo even wrote a treatise titled *On Dreams Sent from God.*

This shift of emphasis from the outer world to the inner world is clearly seen in the Neoplatonism of Plotinus (c. 204–270). Plato, like all the Greeks, was supremely interested in action, politics, and the external world. His theory of Forms was, in the main, explanatory—his version of the search of the Greek cosmologists for the One in the many. Even the supreme Form, the Form of the Good, was both the source of the intelligibility of the world and the supreme ideal of action. Plotinus, on the other hand, saw mystical contemplation and absorption in the One as an end in itself. Psychology therefore became harnessed to the exploration and mapping of inner experience. As G. S. Brett remarks in his *History of Psychology*: "In Plotinus, for the first time in its history, psychology becomes the science of the phenomena of consciousness, conceived as self-consciousness" (R. S. Peters, ed., rev. ed., p. 206).

With the adoption of Christianity as the official religion of the Roman Empire a place had to be found for revelation as well as for knowledge found in inner experience. Augustine (354–430) managed to combine these two sources of knowledge. Insofar as there was no revealed doctrine on a matter, he dealt with it within the framework of Platonism penetrated by Christian mysticism. For instance, the growing knowledge of the self and of God was fitted into a Christianized version of Plato's doctrine of reminiscence. Questions about the body, on the other hand, were dealt with by an appeal to the Scriptures. So, too, was the origin of the soul, for it was transmitted into the body when God breathed upon Adam. The lasting influence, however, of Augustine's *Confessions* was the importance attached to introspection and private experience. No man can escape from his own experience; he can obtain knowledge, insofar as he does not rely on revelation, only by working backward to the presuppositions of his experience as a thinking being. In this approach to the mind Augustine anticipated Descartes.

A corrective to this extreme subjectivity was provided by the rediscovery of Aristotle and the meticulous transmission of his texts by Islamic theologians. The adaptation of Aristotle in the service of Christian theology reached its climax in the work of St. Thomas Aquinas (1224?–1274). But using Aristotle as a substructure to support Christian theology was not entirely straightforward. To start with, there was the problem about the status of reason, one of the most debated topics during the Middle Ages. Aristotle's account of the Active Intellect suffered from notorious obscurities, and there was the worry about its relation to revelation as well. Furthermore, the Islamic school, culminating in Averroes, had tended to favor a mildly pantheistic interpretation of Aristotle's doctrine of Active Intellect. Averroes held that the reasons of individuals are but fleeting manifestations of universal reason. Thomas rejected this interpretation, completely following his teacher Albert the Great (c. 1193/1206–1280).

Thomas defined intellect as the faculty of comprehension that each individual possesses as an intelligent being. Nevertheless, reason was still regarded, as by Plato and Aristotle, as the mark of man's difference from animals and as, in some sense, superhuman. It is qualitatively distinct from sensation and any other processes that are intimately connected with the body.

Apart from this query about the status of reason, which was itself a legacy from Aristotle, Thomas tried to stick to the Aristotelian view of the soul as the form of the body. He deliberately rejected the more Platonic theory that a man is a soul using a body. It was not just respect for the authority of Aristotle that influenced Thomas. The fact was that Christianity was committed to the belief in the resurrection of the body. The intimacy of the connection between soul and body postulated by Aristotle was a better foundation for this doctrine than the more Platonic view occasioning that contempt for the body that culminated in the Albigensian heresy that the body had been created by the devil. Thomas followed Aristotle

closely in his account of sensation, *sensus communis,* memory, and imagination. What was lacking was Aristotle's stress on striving toward an end as the defining characteristic of soul. The intuitive certainties of self-consciousness explored by Augustine remained the foundation both of psychology and of epistemology.

Scholasticism has now become a byword for sustained attention to minor questions within a system whose foundations in revelation were not questioned. There is point in such criticisms. Nevertheless, the Schoolmen preserved and spread a tradition of disciplined discussion that is the lifeblood of science and philosophy. Furthermore, in psychology they handed down not only the general outlines of Aristotle's conceptual scheme but also the details of his psychological system.

The great natural philosophers were nurtured in this Aristotelian tradition even though they eventually overthrew it. At Padua, for instance, where Galileo Galilei was trained, there was a flourishing branch of the Averroistic type of Aristotelianism. Descartes was trained by the Schoolmen at La Flèche, and his *Passions of the Soul* bears witness to these early influences. Even Thomas Hobbes, one of the archenemies of Aristotelian essences, relied on Aristotle's *Rhetoric* for the details of his psychology. He merely poured a traditional content into a mechanical mold that he adapted from Galileo, Pierre Gassendi, and the ancient atomists. The Schoolmen provided the thinkers of the seventeenth century with something solid and disciplined to revolt against. And, as with most rebels, these thinkers were really revolting against a mass of assumptions that were deeply embedded in their own consciousness. Indeed, in a certain sense their revolt was only a return to other elements in their intellectual heritage—the precipitates left by the Pythagoreans, Plato, and the atomists.

DESCARTES. Descartes's view of the mind was a return to Plato, enriched by the introspective musings of Augustine and made more precise by developments in the natural sciences.

Nature and mind. The natural sciences had made leaps forward not because of a vast accumulation of new facts, though one of the features of the Renaissance had been man's turning his gaze out toward the natural world; it was, rather, because of the amazing success that had attended the application of geometry to the phenomena of the natural world.

The success of geometric thinking about nature tended to corroborate what Plato had said about the status of reason as contrasted with the senses; it also convinced the new natural philosophers like Johannes Kepler, Galileo, and Descartes that the real qualities of the natural world were those which could be treated geometrically. Matter was homogeneous, as the atomists had said. Qualitative distinctions, which had been exalted by Aristotle into irreducible natural kinds, were appearances of the varying motions and configurations of the underlying bodies. The Aristotelian doctrine of form and matter was banished; so were the final causes that he had postulated in nature.

How, then, was mind to be conceived, once the Aristotelian doctrine of form and matter had been discredited? There were two obvious possibilities. One was to adopt Epicurus's view that soul and mind were configurations of light and mobile atoms. The other was to revert to the Platonic view that mind is an altogether different type of substance that inhabits the body. Descartes adopted the second course, partly because he shared Plato's view about the wonder of reason and its difference from sensation and bodily processes and partly, no doubt, because of his Christian convictions about God, freedom, and immortality.

Mind. Descartes's departure from Aristotle was much more radical in his account of the soul than in his account of the mind. Whereas Aristotle had described the soul, even in its most primitive manifestations, in teleological terms, Descartes attempted to describe all its lower functions, which were connected with the body, mechanically. His account of mind was not dissimilar in its main outlines from Aristotle's account of reason, which was the most Platonic part of his doctrine, for both accounts held that mind comes from without, furnishes the ultimate principles of thought, and may be considered apart from the body. Indeed, Descartes stated emphatically that the mind can think without a body.

For his account of mind Descartes looked into himself in the manner of Augustine, but he rejected that reliance on faith which was epitomized by the protestation *Credo quia absurdum* ("I believe because it is absurd"). Nothing that was not clearly and distinctly present to the mind was to be included in a judgment. Everything must be doubted—even mathematical truths—until a belief can be found that applies to what exists and that it would be self-contradictory to deny. Descartes's *cogito ergo sum*—his more precise rendering of Augustine's intuitive certainty about his existence as a thinking being—was the result.

Descartes explored the rest of what was intimated in this first certainty and tried to spin out of it all sorts of other truths—for example, the existence of God and of

an external world. The details of his attempted demonstration do not concern us here. They effectively established, in Descartes's view, the existence of thinking substances that were innately so constituted that they would come to form clear and distinct ideas of extension, figure, motion, and other simple natures. Ideas are all mental; as images they are presented through bodily processes, images being apparently corporeal.

Minds were thought to be passive in cognition. When a mind is thinking clearly and distinctly, its ideas correspond to the real qualities of objects. But minds are also active in volition. At the intellectual level their activity consists only in assent to the necessary connection between ideas, and volition is one of the most potent sources of error, for there is often assent when ideas are not clear and distinct. Volition is also the cause of action and is operative in attention, recollection, and fantasy.

Body-mind relation. Descartes's account of the body-mind relation was not dictated solely by Platonized Christian piety. It was equally the product of his knowledge of science and his convictions about scientific method. First, Descartes was convinced that the body is a machine and that animals' behavior could be explained mechanically, animals having no souls. He was acquainted with the discoveries of William Harvey that showed the circulation of the blood to be a mechanical process. Furthermore, mechanical models were a feature of the age. Decorative fountains were constructed with model men that were moved hydraulically and even uttered sounds like words. Descartes thought that the body contained tubes like water pipes along which the animal spirits (the up-to-date rendering of Aristotle's "connatural spirits") coursed. Because many movements of the body can be executed without conscious intentions, Descartes assumed that these could be explained in the same way as the movements of the hydraulic men. He has thus been credited with the discovery of reflex actions. He thought that all animal behavior could be explained in this way.

Second, Descartes believed in the principle of conservation of energy. The quantity of motion imparted to and conserved in a system being constant, there could be no extra source of energy deriving from volition. Thus, the relationship between body and mind had to be conceived in a way that was consistent with this principle.

Third, Descartes held that scientific explanation consisted of making deductions from relations grasped between clear and distinct ideas. Clear and distinct ideas were available of the simple natures of body (for example, extension, figure, motion) and of mind (thinking, willing) but not of the relation between them. Descartes held fast to the obvious fact that body and mind interact (for when I will, it is my arm that moves; I feel pain when my body falls and not when a stone falls). But we have only a confused idea of this interaction. His account of the relationship between them was therefore only a likely story with which he was not really satisfied. It only narrowed down the point at which the crucial philosophical difficulties occurred.

Descartes knew that muscles operate in opposing pairs and that nerves are necessary for sensation and movement. He pictured nerves as tubes along which animal spirits flow. Changes in the motion of these animal spirits cause them to open some pores in the brain rather than others. When this happens, the spirits are deflected into muscles that move the body by being distended laterally and, thus, shortened. At the level of instinct and habit this process is purely mechanical. At the level of conscious intention, however, something more had to be postulated, the impact of mind on body at the crucial switching point of the spirits, the pineal gland.

Descartes supposed that in sensation motion was transmitted from the stimulus object through a medium to the sense organ and thence along the spirits in the nerves to the pineal gland in the center of the brain, where an impression was made like that of a seal on wax. This was a material image that stimulated the soul to produce a corresponding idea. Descartes gave a similar account of passions in the narrow sense of emotions and organically initiated disturbances, which have their source in the agitation of the spirits. By passions in a general sense, Descartes meant all things that *happen* to minds, including sensations, lower forms of memory, feelings, emotions, and other disturbances of reason. These he contrasted with the mind's activity. All such incoming stimuli generally give rise to an act of will. Willing again makes contact with the body at the pineal gland, and a chain of events is started in the body terminating with the movement of the muscles, which produces voluntary action.

The soul is like a pilot in a ship in that it can effect the direction but not the amount of bodily movement. Thus, Aristotle's image of active reason could be reconciled with the principle of the conservation of energy. Descartes's hypothesis that interaction between body and mind occurred at the pineal gland did nothing to dispel the philosophical perplexity about how this interaction could be conceived, and then the pineal gland later was shown to be nothing more than an obsolescent eye. Descartes was attached to this idea because the pineal

gland was the only part of the brain that was not duplicated in both halves of the brain. He was convinced that the soul, being unitary, could not affect the body at two points. His hypothesis enabled him to keep his mechanistic account of the body intact.

For a long time it has been fashionable to deride Descartes's rather disastrous form of dualism and even to suggest that he created the body-mind problem. This is a piece of intellectual insularity. Descartes was perhaps the first thinker to formulate the problem at all clearly. It would be possible to deny his basic assumption that body and mind are qualitatively distinct substances and still to claim that apart from this metaphysical extravagance his statement of the problem brought out at least two cardinal points that are involved in it. First, he obviously saw the logical incongruity of explaining mental processes, such as geometric reasoning and deliberating before action, in mechanical terms. There is a logical gap between the types of explanation used, as Aristotle had pointed out in his criticisms of the mechanists who held that the soul was moved. Descartes, in his account of the transactions that were alleged to take place at the pineal gland, must have thought that motion at this point is somehow identical or correlated with the mental activity involved in producing an idea or making an act of will. His hypothesis did much to draw attention to this logical disparity between the two types of description.

Second, Descartes's account did much to establish privacy, rather than Aristotle's criterion of purpose with plans and rules superimposed at the level of the rational soul, as the main hallmark of the mental. As has been indicated, Descartes's theory in this respect marked the culmination of a trend that can be traced back through Augustine and Plotinus to Philo. To attribute mind to something is not just to say that men act in accordance with rules and that their movements persist toward ends. It is to say that they act like this because of their knowledge of rules and because they are conscious of ends. Consciousness is crucial for picking out the obvious respect in which men differ from cunningly contrived machines. Descartes must be credited with the clearheadedness to have stood firm on this cardinal point.

SPINOZA. Benedict de Spinoza's system was a consequence of pushing Descartes's assumptions to their logical conclusions.

Nature and mind. Descartes had accepted the traditional notion of substance as that which is a cause of itself, can be conceived through itself, and needs only itself in order to exist. Spinoza (1632–1677) argued that if this is the definition of substance and if there is such a substance, there can be only one such substance, which can be called either nature or God. Nature, so conceived, must have infinite attributes, but we know only two of them, thought and extension. God is therefore "the place of the world and the whole system of thinking." Everything is a mode or modification of God. Thus, nothing can be adequately explained unless its occurrence can be deduced from principles applying to the system as a whole.

Explanation is deductive in character and accords with mechanical principles. Unlike Descartes, Spinoza envisaged a science of psychology in which mental as well as physical phenomena could be deduced from quantitatively expressed laws. Emotions, he argued, must obey laws just as lines, planes, and bodies do. Human beings, as part of nature, must exhibit the general characteristics of all modifications of God or nature. They must be determined within a system; they must have a mental and a physical aspect; and they must exhibit conatus, or the striving to persist within their own being. These characteristics must now be considered in turn.

In stating that human behavior was determined within a system, Spinoza wished to oppose what he considered to be two basic illusions that human beings had with respect to themselves. The first of these was the illusion of free will. People are convinced that they have free will, he argued, because they are conscious of their actions but ignorant of their causes; thus, they conclude that they are uncaused. If stones were conscious, they, too, would believe in free will. Yet human behavior can be explained just as can the movements of stones. In both cases the explanation will consist in deducing what occurs from the laws of the system of which they both are part, ultimately the system of nature as a whole. The human body is a system of simpler elements maintained in an equilibrium, but this system is part of a broader system, not a self-contained isolable system. Adequate explanation is seeing events as part of the whole system of nature; in this system there are no final causes. Nature just is, like a vast, timeless machine.

Body and mind. How then was the body-mind relation to be conceived? Spinoza was one of the first to point to the difficulties in Descartes's pineal gland hypothesis. Spinoza's solution was to suggest that interaction does not take place for the very good reason that body and mind are correlated attributes of the same underlying substance, not distinct substances. Indeed, Spinoza says that the mind is the idea of the body. This is obvious enough at the level of immediate confused ideas that are

of bodily states. But the changes in a man's body are part of a larger system, which includes the properties of the food absorbed in nutrition. A wider knowledge of the events in a man's stomach is possible for a physiologist who can understand the laws governing them. He would see these events as part of an ever widening network of events which constitute nature. The man's feeling of stomachache, on the other hand, would be confused, fragmentary, and inadequate, an idea of an effect cut loose from its causes.

This illustrates the difference between what Spinoza called the first and second grades of knowledge. The materials of the first grade are the confused ideas of bodily states that we call feelings and sensations. These ideas are connected only by principles of association. This is the level of sense perception and imagery, of uncritical beliefs founded on animal instinct, association, and hearsay. The second grade of knowledge is rational insight. At this level rational connections are grasped as general notions develop that connect an ever widening system of events. The more abstract and general thought becomes, the nearer it approaches the thought of the Cartesian physicist and, ultimately, God's thought. There is also a third grade of knowledge, called *scientia intuitiva* by Spinoza, which is more mystical. It is a return from the abstract laws of the scientist to a grasp of the particular as illuminated by such laws. The role of the body, as that which is correlated with mind and of which mind is an idea, seemed to recede when Spinoza passed to reason, or the second grade of knowledge. Mind as the idea of the body becomes at this point almost as difficult a notion as Descartes's notion of mental activity somehow mirroring movement in the brain, for thinking is not of or about body or brain states any more than it is a form of movement which is similar to or identical with brain states.

Conative aspect of mind. Spinoza's account of mental phenomena was much less intellectualistic than that of Descartes. Indeed, in certain respects he reverted to Aristotle's emphasis on teleology and self-maintenance. Spinoza held that the most important characteristic of every modification of nature was its conatus, its striving to persist in its own essence. In man, as in every other natural modification, there is an inherent tendency to react to all changes in a way that maintains its characteristic unity and equilibrium. A person differs from animals in being self-conscious in this endeavor.

Spinoza employed this homeostatic postulate to rewrite Descartes's account of the passions as presented in *Les passions de l'âme.* Descartes had paid particular attention to the causal influence of animal spirits and had left rather vague the part played by the cognitive grasp of the situation, though he generally put forward an ideomotor theory. Spinoza evinced little interest in the physiology of the matter. Instead, he developed a theory of motivation by harnessing Descartes's passions to his own homeostatic principle. He postulated that whenever a body is acted on by another body, its vitality may be increased, may be diminished, or may remain constant. The awareness of these occurrences is the mental aspect of the psychophysical states which are called emotions. There are thus three primary emotions corresponding to increase, diminution, or maintenance of bodily vitality. These are joy (*laetitia*), grief (*tristitia*), and desire (*cupiditas*). As a result of experience people tend to keep before them what will increase their vitality and remove what will decrease it. "Love" is thus defined as "joy accompanied by the idea of an external cause."

Spinoza drew a sharp distinction between the passive emotions which characterize the first grade of knowledge and the active ones which mark the second and third grades. People are passive when the cause of changes in them lies outside them. In this state of human bondage the emotions that accompany confused, fragmentary ideas are thrust on people; they tend to be sporadic, inordinate, unpredictable, and obsessive. Individuals are subject to panic, jealousy, and overmastering loves and hates. When a man passes to the second grade of knowledge, however, his vitality is increased, and there is a distinctive form of joy that goes with the use of reason. The explanation of human conduct is now to be sought within him, in his clear understanding of the world and of his relation to it. By understanding himself, including his own emotions and history, as part of the system of nature, a man can attain a kind of freedom, which depends upon his acceptance of his own nature. He is then capable of rational self-love and rational benevolence and can attain glimmerings of the greatest good which he can possess—"the knowledge of the union which the mind has with the rest of nature." The attainment of this state brings its own delight.

In making suggestions for attaining this state of blessedness, Spinoza in many respects anticipated later psychoanalytic techniques, as well as the general psychoanalytic aim of replacing subservience to irrational promptings by rational control based on self-knowledge. He thought, for instance, that many irrational reactions could be traced back to an early reaction to an object to which the present object had become associated by irrelevant similarities. Scientific understanding of this might help to dissociate the emotion from the irrelevant stimu-

lus. He was not so naive, however, as to suppose that mere intellectual understanding could free an individual from the obsessiveness of emotion. It takes an emotion to master an emotion. And Spinoza thought that seeing things "under the aspect of eternity" had a specific emotional accompaniment. Hence, the psychological shrewdness as well as the ethical profundity of his remark, "Blessedness is not the reward of right living; it is the right living itself. Nor do we delight in blessedness because we restrain our desires. On the contrary it is because we delight in it that we restrain them."

HOBBES. Hobbes (1588–1679) already subscribed to the deductive model of geometry when he visited Galileo in 1636. He returned replete with concepts and laws that were to form the foundation of his psychology. For the idea had dawned on him, perhaps suggested by Galileo, of applying the new natural philosophy to human behavior. Of course, Epicurus had long ago sketched a mechanical theory of mind, but it was very general. Galileo had worked out the details of a new theory of motion. Could not still further consequences be deduced from the law of inertia? Harvey had deduced the theory of the circulation of the blood from mechanical postulates. Could not Hobbes apply the details of this new theory of motion to psychology and politics?

Body and mind. Hobbes did not really see any particular problem about the relationship between body and mind because for him everything was body. Even God must have a body if he exists, for "substance incorporeal" is a contradiction in terms.

Thus, "conceptions and apparitions are nothing really but motions in some internal substance of the head." Sensation is "some internal motion in the sentient," and pleasure is "nothing really but motion about the heart."

In truth, Hobbes was not much worried by such philosophical niceties as whether, according to his theory, mental phenomena like thinking were being postulated as identical with or merely causally dependent on motions in the head. He was much more interested in working out a mechanical explanation of these phenomena. This is what makes his psychology of absorbing interest. It represents just about the first attempt in the history of psychology to put forward in any detail something that begins to look like a scientific theory.

Mechanical theory of mind. According to Hobbes, in sensation the sense organs were agitated by external motions without which there could be no discrimination and, hence, no sensation. The selectivity of perception was explained by suggesting that while a sense organ retains motion from one object it cannot react to another; similarly, in attention the motion from the root of the nerves persists "contumaciously" and makes the sense organ impervious to the registering of other motions. Imagination was explained by a strict deduction from the law of inertia: "When a body is once in motion, it moveth, unless something else hinder it, eternally; … so also it happeneth in that motion, which is made in the internal parts of man, then, when he sees, dreams, etc.… Imagination therefore is nothing but decaying sense." This decay is not a decay in motion, which would be contrary to the law of inertia. It comes about because the sense organs are moved by other objects. This explains why dreams are so vivid, for in sleep there are no competing motions from the outside world. Thus, the longer the time that elapses after sensing an object, the weaker the imagination. Memory is imagination with a sense of pastness added to it.

This was an exciting and an ingenious theory. The difficulty about it is that the type of distinction implied in the explicanda cannot really be deduced from the mechanical postulates of the theory, for the differences between perceiving, imagining, and remembering are basically epistemological ones implying standards and criteria different from those that might be attributed to mere movements. Hobbes never faced the basic difficulties that Aristotle first formulated in his opposition to the theory that the soul was itself moved. Nevertheless, Hobbes did produce something that looked like a scientific theory. Its conceptual difficulties attend all psychological theories that attempt to translate epistemological distinctions into differences of process.

Mechanical theory of action. In the theory of action Hobbes attempted to get rid of final causes and to substitute efficient causes for them. To do this, he had to introduce the concept of endeavor, which was very different from Spinoza's conatus. He used the term *endeavor* to designate infinitely small motions, which he postulated as occurring in the medium between the object and the sense organ, between the sense organ and the brain, and heart. His theory of motivation was that external objects transmit motions by a medium to the sense organs and from there to the brain and to the heart; this results not only in the production of images but also in some alteration or diversion of vital motions round the heart. When these incoming motions help the circulation of vital motions, it appears to us as pleasure, and the body is guided to preserve the motions by staying in the presence of the stimulating object; and conversely with pain.

Appetite and aversion are thus the first endeavors of animal motion. They are succeeded by the flow of animal spirits into some receptacle near the "original" of the nerves which brings about a swelling and relaxation of the muscles causing contraction and extension of the limbs, which is animal motion.

Hobbes thought this mechanical account of action was quite consistent with ascribing a central role to consciousness, for in Hobbes's view all action was voluntary in the very strong sense that it is preceded by the thought of an end to be attained. He also claimed that the only way to develop a science of human nature was to look into ourselves and analyze what we find there. Hobbes found two basic motions of the mind, "the one arising from the concupiscible part, which desires to appropriate to itself the use of those things in which all others have a joint interest; the other proceeding from the rational that teaches every man to fly a contra-natural dissolution, as the greatest mischief that can arrive to nature." Everything we do is derived from the desire for power or the fear of death. Conflict between manifestations of these basic motions of the mind leads to deliberation. In this "alternate succession of appetite and fear" the one that emerges triumphant is called "will." "Will therefore is the last appetite in deliberation." Free will is an illusion, for the outcome of such conflicts can be explained mechanically.

Theory of passions. On top of this mechanical ground plan Hobbes superimposed an account of the passions taken largely from Aristotle's *Rhetoric.* They are to be distinguished by reference to the objects of appetite and aversion as well as by our opinion of attaining such objects. Ambition, for instance, is desire for office; hope is appetite with an opinion of attaining. Individual differences are due, in the main, to differences in the mobility and agility of the animal spirits. Dullness, for instance, derives from "a grossness and difficulty of the motion of the spirits about the heart." Hobbes even had a theory of laughter, which he thought to be the expression of sudden glory caused by something new and unexpected in which we somehow discover ourselves superior to others.

Hobbes assigned a special place in his theory of the passions to curiosity, which, together with the ability to name things and hence to reason deductively, distinguishes humans from animals.

Hobbes's account of the passions was unusual in that it was so positive. For him passions were not, as for the Stoics, imperfect reasonings; they were a particular case of motion in the natural world on which his account of human nature was erected. Nevertheless, when he dealt with what was distinctive of man, his reason, Hobbes parted company with both naturalism and mechanical theory. The type of reason, called prudence, which enables man to satisfy his desires more efficiently, on the basis of experience, must be sharply distinguished from the reason by means of which men are able to arrive at the universal truths of geometry and philosophy.

Scope of mechanical theory. This is not the place to enter into the tortuous details of Hobbes's nominalist theory of meaning or his conventionist theory of truth. It is important to note, however, that in dealing with these specifically human facets of behavior, just as in his treatment of the foundations of civil society, Hobbes defended a position that stressed above all the role of artifice and convention. He even put forward a kind of contract theory of definition to parallel his social contract theory of government. These accounts were underpinned by a very crude causal theory of signs as well as by a mechanical theory of human nature. But no clear connection was ever made between the conventionist and naturalistic elements. David Hume later tried to make such a connection by suggesting that reason was a wonderful and unintelligible instinct in human nature. Hobbes, however, more or less ignored his own mechanical theory when he dealt with geometry, law, logic, and other such artificial creations of human reason.

Thus, although Hobbes was the first thinker to develop in any detail a mechanical theory of mind, he also, more or less unwittingly, exhibited the glaring difficulties in such an undertaking. Indeed, the things in which he was most interested, apart from politics, were precisely those things which it is very difficult to accommodate within a mechanical theory.

LEIBNIZ. Gottfried Wilhelm Leibniz (1646–1716) understood much better than Hobbes the new natural philosophy; indeed, his discovery of the infinitesimal calculus contributed considerably to it. However, he resisted its mechanistic implications. Descartes had viewed nature, the animal world, and bodies as machines but had stopped short at mind; Hobbes had mechanized mind as well. Leibniz went to the other extreme and mentalized nature. In many respects he reverted to Aristotle.

Nature and mind. The *Monadology* was a brilliant synthesis of Aristotelian logic taken seriously and a variety of trends in the natural sciences. The whole Cartesian philosophy presupposed the subject-predicate view of judgment in which every proposition, when reduced to logical form, has a subject and a predicate. Moreover, the predicate was thought to be contained in the subject. The

Aristotelians thought that this common structure of language mirrored a world of substances composed of various attributes. Leibniz, like Spinoza, took the definition of substance seriously; he thought that it was the cause of itself, could be conceived by itself, and needed only itself in order to exist. But where Spinoza concluded that if this was the definition of substance, there could be only one—namely, God or nature—Leibniz concluded that the world must be composed of countless substances all exhibiting the features picked out in their definition. These monads develop according to an immanent principle that is their force or essence. Everything that will ever happen to them, their predicates, is included in their original notion. The principle of sufficient reason explains the succession of these states in time, the identity of a substance at different times being recognized by "the persistence of the same law of the series." Now I am a substance and know by introspection that I am characterized by appetition and perception. What I know about myself must in general be a paradigm for the basic structure of all substances. But no two substances are alike. In perception they all mirror the universe from a particular point of view. There is no interaction, however. Each monad is windowless and develops because of its own immanent principle, not because of external causal influences. The monads seem to influence one another only because of the preestablished harmony of their immanent development.

This bizarre application of an ancient logical doctrine to the world accorded nicely with various new developments in the sciences. Leibniz naturally regarded it as consistent with his discovery of the infinitesimal calculus, the guiding idea of which was that a succession of states develops according to a law governing the series. The successive states of a monad flow into one another like a series of terms differing infinitesimally, their development being defined by the law of the series. This fitted well with the law of continuity, which held that *natura non facit saltus* ("nature makes no leaps"). Change is a summation of infinitesimal degrees of change. Furthermore, the recent discovery of the microscope revealed that if a piece of cheese or a seemingly empty pool is examined, each will be found to be teeming with life. Could not all nature, therefore, be alive—a vast system of monads at varying levels of development? In embryology, too, the doctrine of preformation was in vogue. The assumption that all the characteristics of an adult animal exist in embryonic form from the moment of generation supported Leibniz's view that from the original notion of the monad all its later states and characteristics could be deduced. His conception of the essence of monads being force or activity was connected, too, with his contribution to the dispute in dynamics about the relationship between force and mass. Leibniz held that his concept of *vis viva* or activity directed toward the future states of the monad was required by his discovery of the conservation of momentum.

The synthesis of Aristotelian logic and these trends in science made Leibniz utterly opposed to the mechanistic picture of nature and of man in which the real world was a world of bodies in motion having only primary qualities whose changes were to be explained only by reference to efficient causes. What is real, he claimed, is not what is mathematically measurable but our experience of activity and perceiving. Nature, as well as man, is characterized by appetition and perception. Final causes are reconciled with the laws of motion by the principle of sufficient reason, which governs the unfolding of the immanent nature of the monads. The difference between substances is only one of degree of clarity in perception and of self-consciousness in appetition. Bare monads have a minimum of perception and appetition. Their perception is confused, and their appetition is blind. Souls, or conscious monads, have memory, feeling, and attention. Animals, or, rather, the dominant monads of animals, are examples. Rational souls, or spirits, are self-conscious; unlike brutes, which are "empirics" and are aware only of particulars, they can reason and understand necessary truths. Extension is only an appearance, the way in which low-grade monads appear to us; the laws of motion are just appearances of the laws of appetition which depend ultimately on God's choice of what is best. Aristotle and Galileo are reconciled, but Galileo's and Isaac Newton's laws are, at best, laws of appearances.

Concept of mind. Leibniz's concept of mind or soul was articulated in what he said about perception and appetition. He regarded perception as marvelous because it cannot be conceived of as an action of the object on the percipient, for the monads are windowless. Perception is better regarded as the expression of a plurality in a unity. One thing may be said to express another when there is a constant and regular relation between what can be said about the one and about the other. It is thus that a projection in perspective expresses its original. The monads are perspectives of the universe from different points of view. Expression is thus the genus of which perception, animal feeling, and intellectual knowledge are species.

Leibniz combined this highly metaphysical account of perception with some shrewd objections to John Locke's tabula rasa theory of the mind. He held that the senses provide us only with instances and by themselves

cannot provide the sort of universal knowledge that we have in science. The mind is active and categorizes experience by means of which it interprets the testimony of the senses. The proper analogy for the mind is not a tabula rasa but a block of veined marble. In this doctrine Leibniz harked back to Aristotle's active reason and laid the foundation for Immanuel Kant's categories. Locke, he argued, had in fact tacitly admitted this in postulating mental operations that are known by reflection.

Leibniz maintained that Locke was wrong in saying that the mind does not always think. We have an infinite number of perceptions of which we are not aware. Habituation and wandering attention, as well as the smallness of the perceptions, explain our failure to notice them. Our attention is often drawn to a sound that has just occurred and that we would not otherwise have consciously noticed, although we registered it. "These insensible perceptions are also the signs of personal identity and its constituents; the individual is characterized by traces of his previous states which these perceptions preserve by connecting them with his present state." They are also the means of recollection. They explain decisions that seem arbitrary to us, like turning to the left rather than to the right; they explain frequent feelings of uneasiness which are not intense enough to be felt as pain. These insensible perceptions, he argued, are "as much use in pneumatics as is the insensible corpuscle in physics." Both are beyond the reach of our senses, and there are as good grounds for believing in one as in the other. Since "nature makes no leaps," these insensible perceptions must accord with the law of continuity. "All this brings us to the conclusion that observable perceptions come by degrees from those which are too small to be observed."

Although Leibniz confused some rather different things in this doctrine—for example, unconscious perceptions, minute perceptions that summate like the noise of waves in the roar of the sea, and confused perceptions—he prepared the ground for the concept of unconscious mental processes which was to prove so important in nineteenth-century thought, and he anticipated later investigations of subliminal perception and "determining tendencies." This shows how a highly speculative theory can lead to the emphasis on facets of experience which may be very important but which have previously been disregarded.

Leibniz's emphasis on appetition as the other main characteristic of monads was a welcome change from the intellectualism of Descartes and Locke. However, Leibniz made no detailed empirical derivations from this notion to match the derivations made from his concept of perception. It had more affinities with Spinoza's "conatus" than with Hobbes's "endeavor," although it was really the Aristotelian conception of the formal and final cause brought up to date and made compatible with dynamic theory. His concept can best be elucidated by quoting him; he calls his concept by the Aristotelian term "entelechy," which is "a power mediating between the simple faculty of acting and the definite or effected act. It contains and includes effort. It is self-determined to action, not requiring to be aided, but only requiring not to be inhibited. The illustration of a weight which stretches the cord it is attached to, or of a bent bow, may elucidate the notion."

Soul and body. Leibniz believed that every living creature is composed of a vast number of special organic structures each developing in its own characteristic way; they are all so coordinated and mutually complementary, however, that together they act as an individual. The unity is the soul or the dominant monad; the multiplicity is the body or assemblage of bare monads. The monads of the body all have their own activity, and they are represented or mirrored in the perceptions of the dominant monad or mind. The mind has no power to interfere with or penetrate the forces that it seems to direct. The activities of the monads of the body subserve the dominant activity of the mind as the players of an orchestra, each playing independent parts, subserve the performance of the symphony, and the symphony is the resultant harmony, which has been preestablished. The manifold activities of the bare monads thus combine to bring about the end of the dominant monad. The body depends on the mind in the sense that the reason of what happens in the body is to be found in the mind (compare to Aristotle's view of soul and body).

Thus, Leibniz reverted to a view of mind and nature which was basically Aristotelian, but he transformed the Aristotelian entelechy by giving it the basic hallmarks of Cartesian mind—thinking and willing as experienced from within. Furthermore, he pressed the emphasis on privacy much further than Descartes by claiming that the monads are windowless and that everything that will ever happen to them is contained in their original notion.

There was, however, another radically different concept of mind which developed out of Descartes's stress on privacy and incorrigibility as the hallmarks of mental states. This was that of British empiricism, which culminated in Hume and the associationists.

HUME. The contribution of Hume (1711–1776) to psychology was not very extensive in its details because his

theorizing about the mind, like that of George Berkeley and Locke, was mainly a way of doing epistemology. And there were special reasons, deriving from his epistemological position, for his eschewing speculation about the relationship between mind and body and the general status of mind in nature. Nevertheless, his general concept of mind was of considerable historical importance. It was the first thoroughgoing attempt to eliminate spiritual substance altogether, and it was the first theory to make reason subservient to the passions and to extol the importance of instinct and habit. It was also the first attempt to develop a Newtonian theory of mind and to erect the principles of the association of ideas into scientific postulates—an undertaking which considerably influenced David Hartley and hence the course of associationist psychology.

Hume's predecessors. John Locke (1632–1704) took from Descartes the assumption that we are confronted with our own ideas, not with things, and that some kind of certainty is both desirable and attainable. He rejected, however, Descartes's doctrine of innate ideas and adopted a Baconian version of empiricism. He postulated simple ideas of sense that made their imprint on the passive tabula rasa of the mind. Once ideas got into the mind, Locke's theory more or less followed Descartes's, for he believed that the active spiritual substance within intuits relations between ideas, the relations which form the foundations of knowledge. Locke, however, did not stick consistently to his "way of ideas." For example, he asserted, like Descartes, that we have intuitive knowledge about our own existence as selves and "sensitive" knowledge of things existing independently of our perceptions of them. They are material substances that support "powers" to produce in us ideas of primary qualities, which are real properties of the things in question, and secondary qualities which are not real.

George Berkeley (1685–1753) stuck more consistently to the way of ideas and eliminated material substance, of which we have and could have no idea because it is a logical absurdity; the representative theory of perception; and the distinction between primary and secondary qualities. He claimed, however, that we have "notions," rather than ideas, of ourselves as active agents and of other minds, including God. We also have a notion of our own causal activity. Berkeley relied on this notion to distinguish ideas of sense from ideas of imagination, for having eliminated the concept of a thing independent of our perceptions, Berkeley had to have a criterion for distinguishing what are commonly called things from the mere coexistence of qualities; imaginary objects, for instance, appear to us as clusters of coexisting qualities. Thus, he claimed that when we see objects, it is God talking the divine sense language and producing ideas in our minds; when we imagine objects, we are doing the producing ourselves and have a notion of our own agency in so doing. Berkeley's stress on the activity of the mind contrasted strongly with Locke's tabula rasa.

Hume simply stuck rigorously to the way of ideas and eliminated Berkeley's "notions." There was no simple idea of material substance, of ourselves and others as spiritual substances, of God, or of causal agency. All that was left, therefore, as genuine components of the mind were ideas themselves and certain links between them. Hume likened the mind to a theater "where several perceptions successively make their appearances, pass, repass, glide away," and to a political organization in which the members come and go but the principles of organization—the principles of the association of ideas—persist.

Hume's contributions. Hume was the first to attempt an explicit distinction between images, which he called impressions, and what we would now call sensations—he called them ideas. He regarded them as two sorts of perceptions. Impressions could not be distinguished from ideas in a Lockian way by their relation to an external object. For Hume, following the way of ideas, disclaimed any possibility of knowledge of a world of objects existing independently of our perceptions. And, because he ruled out notions, Berkeley's appeal to awareness of our causal agency in producing ideas of imagination was not open to him. Of course, like Berkeley, Hume agreed that what we call things exhibit a certain constancy and coherence; they resemble past clusters of qualities. We assume independent existence in order to connect past with present perceptions. But, he argued, we can no more demonstrate the existence of a world independent of us than we can demonstrate that pleasure is preferable to pain.

There are, however, subjective criteria for making the distinction between images and sensations, which is all that remains once belief in a world of independent objects has been ruled out. These are the criteria of vividness and order. Hume suggested that ideas could be picked out because they were faint copies of previous impressions. In other words, impressions are both more vivid than ideas and prior to them. But he gave counterexamples to both these criteria—those of vivid ideas in fever or madness and of forming an idea of a color that had never previously been presented as an impression. In the case of fever or madness Hume suggested that the imagination transfers the vividness of an impression to

an idea. Similarly, our belief in an external world is a work of the imagination.

Hume's recourse to the imagination was of cardinal importance in his account of the mind because it linked his theory of knowledge with his rehabilitation of feeling. It has often been remarked that one of the main features of Hume's philosophy was a reversal of the roles hitherto ascribed to reason and feeling. He brought over into epistemology his ethical theory, which he adapted from Francis Hutcheson's theory of moral sense, that moral judgments are based on feeling. "Reason is, and ought always to be, the slave of the passions." This moral sense was the product of biological properties inherent in the species; it had its counterpart in our judgments of matters of fact and existence. Reasoning is "nothing but a wonderful and unintelligible instinct in our souls." Our belief in the reality of causal connections or in the existence of an external world or that the future will resemble the past are instinctive and indemonstrable. "Nature, by an absolute and uncontrollable necessity, has determined us to judge as well as to breathe and feel." The categories used by scientists in their theories, such as continuity and causality, are largely products of the imagination.

Hume stressed facets of human nature that had been largely neglected since Aristotle. He postulated an original fabric of human nature consisting of various propensities not unlike that of later instinct theorists. He also extolled the place of habit in conduct, not simply in explaining such developed forms of behavior as obedience to government but also in explaining the origin of some indemonstrable beliefs. For instance, he held that the idea of causal connection could be analyzed into the elements of priority in time of event *A* to event *B* and constant conjunction of event *A* with event *B,* together with a conviction of the necessity that *B* must follow *A*. As there was no impression of this necessity given in experience, Hume attributed our belief in it to habit or a "determination of the mind" brought about by experience of such constant conjunction and the force of the imagination.

The passions. Appropriately enough, the details of Hume's psychology consisted mainly of an elaborate and highly complex theory of the passions, stated in Book 2 of his *Treatise of Human Nature.* One of Hume's tasks was to rehabilitate the passions, the natural feelings of decent people, from the Puritans' distrust and the rationalists' disregard. He also had to demolish sophisticated theories, deriving from Hobbes, in which all passions were regarded as forms of self-love. Whereas Bishop Butler attacked psychological hedonism in order to establish the supremacy of conscience, Hume refuted the hypothesis of self-love in order to make way for his rival hypothesis of innate benevolence and sympathy.

He also regarded the sensations of pleasure and pain as part of the original fabric. In a passion one of these sensations is accompanied by an affection. The direct affections include desire and aversion, joy and grief, hope and fear. The difference between these depends on the character of the expectation of good or evil. Desire is for present good, joy for assured good in the future, and hope for probable though remote good in the future. Hume thought that through experience these affections, together with the sensation of pleasure or pain associated with them, can become associated with an object. This generates such indirect passions as pride and humility, when the object is ourselves, or love and hate, when the object is other people. Benevolence and malevolence, however, are not derived from love and hate. Hume classed them as direct and instinctive.

Sympathy occupied a role in Hume's theory of passions somewhat similar to imagination in his theory of belief. The idea of another person's feeling is said to be associated with the idea of oneself, and the required liveliness is thus imparted to the otherwise neutral conception of another person's joy or sorrow.

The idea of the self played an important part in Hume's intricate account of the passions. Like the idea of causality, it presented a serious problem for analysis, for we believe strongly in the reality of both of them. Yet, Hume argued, there was no simple impression of sense from which these ideas derived. Introspection revealed only "some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure." What we call self must therefore be "a bundle of perceptions." Like Locke, Hume then went on to compare the self to an oak, a vegetable, or any type of organism which maintains itself through change by virtue of its relations. Another apt analogy is the self-maintained unity of a political association. But Hume maintained that the unity of this bundle, which makes it a "connected heap," is associative, not real; there are no grounds for ascribing to it the simplicity and permanence which are required for real unity. Perceptions are loose, separate, perishing existences. There can be no real links between them. The problem is to explain how we come to believe that there are.

Hume made the same type of move in relation to the idea of self that he made in the case of causality. He demonstrated that if the way of ideas is followed, there is no ground in experience for believing in the reality of the self; he then embarked upon some speculative psychology

to explain how we come to have this belief. He suggested that members of the bundle are related to one another in a specific way in time, the order being preserved by memory. The members have the relations of resemblance and cause and effect between them. But cause and effect is not a real relation; thus, no real unity characterizes the self. We come to believe in it because of the "felt smoothness" with which we pass from one idea to another once the associative links have been established.

Nature and mind. Although Hume's adherence to the way of ideas ruled out wide speculations about the place of mind in nature, there was a highly imaginative idea behind his positivistic system. Hume regarded himself as the Newton of the sciences of humankind. He made frequent references to his pursuit of the experimental method and thought his rigorous interpretation of the way of ideas to be thoroughly consistent with Newton's methodological canons of economy and simplicity in explanation, testability of hypotheses, and refusal to postulate occult causes. Hume stressed that once we have arrived at the original fabric of human nature, it is futile to attempt to satisfy any further our intemperate desire to search for other causes.

But Hume did not emulate Newton merely in his methodology. He also regarded his concepts in the psychological sphere as parallel to Newton's concepts in the physical. His simple impressions were the equivalent of Newtonian atoms, and his principles of association were likened to the "gentle force" of Newton's principles of gravitational attraction. Indeed, Hume regarded imagination and, perhaps, sympathy as cohesive forces. When imagination works according to the associative principles of resemblance, contiguity, and cause and effect, the result is what Hume called the understanding. When it works capriciously, the result is fancy. Of course, the principles of association were as old as Aristotle, though Aristotle's principles were not the same as Hume's. Hobbes, too, had made use of them, though he believed that thought which was guided by desire or which exhibited a plan was more important. However, in Hume's system for the first time they were looked upon as important *scientific* principles governing the working of the mind. This conception was taken up by Hartley in his theory of vibrations and developed into the associationist school of psychology.

Hume's theory was also important in the history of psychology because it firmly established psychology as the science of the contents of consciousness. Although Descartes's first certainty was rejected in relation to its content, what persisted was the assumption that a man has some incorrigible sort of knowledge about his own mental states. Hume rejected Descartes's search for simple natures, which appear to the mind as clear and distinct ideas, as the foundations of science. Instead, he postulated simple impressions of sense, perishing existences about which we can be certain provided that we make no inferences beyond them. Because Hume, like Locke, consistently confused psychology with epistemology, two parallel traditions developed from his work. On one hand, there was the search in epistemology for sense data which could provide an incorrigible basis for a system of knowledge; on the other hand, there was the development of introspective psychology whose task was envisaged as cataloguing the contents of the mind, analyzing them into simple units, and attempting generalizations about the links between these units which explained the generation of complex ideas and states.

Body and mind. Hume, understandably enough, had little to say about the relationship between mind and body. Body, according to his theory, stood for another bundle of impressions. He did not even connect the idea of self with impressions of bodily states, which might have been an obvious move if he had looked seriously for specific impressions, from which the idea of self is derived. In the Humean tradition William James, for instance, later suggested that the idea of self was intimately connected with impressions of breathing, cephalic movements, and the like. But Hume made no such suggestion. He noted the inexplicability of the fact that "the motion of our body follows upon the command of our will." "Will," he suggested, was another name for the strongest motive (compare to Hobbes's account). But we simply have to accept these de facto connections between events. To speculate further would be to postulate occult causes and thus to sin against both Newtonian methodology and the way of ideas.

KANT. It would be very difficult to sketch the contribution of Kant (1724–1804) to psychology within the framework previously used, partly because he made very little direct and explicit contribution to psychology and partly because his Copernican revolution in philosophy involved a radical reformulation of questions asked under such a framework. Furthermore, though Kant's concept of mind may, in fact, be extremely important insofar as it delimits the sphere of empirical psychology, those who developed empirical psychology in fact paid little heed to the implications of Kant's position. Perhaps that was a pity, for Kant made a sustained effort to separate epistemology from empirical psychology, and until these two are clearly distinguished, there will continue to be confu-

sion in this area, as is demonstrated in the genetic psychology of Jean Piaget. Nevertheless, Kant's influence on psychology was largely negative and indirect; thus, only a short exposition will be given of those parts of his critical philosophy which seem relevant to psychology.

First and foremost, Kant rejected the notion of the empiricists that what is called mind could be explained as the product of ideas arising from experience and systematizing themselves according to laws of association. Kant maintained that the mind must be regarded as a structure regulated by principles of its own activity. These principles could not be arrived at empirically, for they were presupposed by any empirical investigation, including psychology. They could be arrived at only by critical philosophy, which asked the question "What must be presupposed for our experience to be possible?"

Kant was particularly interested in two realms of experience—Newtonian science and the autonomous morality of thinkers of the French Revolution. Kant attempted to reconcile the rationalism of Christian Wolff and Leibniz with the empiricist position of Hume by postulating an active mind whose nature was to impose a structure on experience to make it intelligible. This structure was composed of the categories used by scientists, such as substance, cause and effect, and continuity, which Hume had assigned to the imagination; Kant attributed the structure to reason, which synthesizes the data of sense. The content is provided by the senses, but the form is provided by reason. Thus, what we call nature is in part the work of mind. It is composed ultimately of things-in-themselves, whose real nature must be forever unknowable. We, too, must exist as noumenal selves, as things-in-ourselves. Of course, Hume was right in maintaining that we have no impressions of such selves. At best, we have intimations of such selves behind the appearances in our moral experience as active rational beings.

Human beings have empirical selves insofar as they have bodies and psychic functions—for example, sensation, imagery, feeling, purposes—which depend on embodiment. Such selves can be known by inner sense, and their manifestations can be investigated empirically; Kant called such a study anthropology. Kant made his mark on the history of introspective psychology by imposing on these phenomena the tripartite division—knowing, feeling, and willing—worked out in his *Critique of Judgment.* But he did not note anything particularly novel about the phenomena thus investigated, although he did declare that such investigations could never be properly scientific. He was convinced that science involved quantification and that since the phenomena studied by anthropology could not be subsumed under mathematically expressed laws, psychology could at best be a collection of descriptive material classified under the headings that he suggested. Thus, Kant's extrapolation of Newtonian physics as the paradigm of all sciences had the negative effect of making it incumbent on those who wanted to develop psychology as a science to attempt the quantification of the phenomena to be studied. The result was Gustav Theodor Fechner's psychophysics, Johann Friedrich Herbart's attempt at mathematical laws of consciousness, and countless other premature attempts at quantification.

Another result of Kant's analysis was an increase of interest in the problems connected with the self. The controversy about the existence of a pure self and whether it was a proper object of study occupied most thinkers during the nineteenth century. Of much more importance for psychology, however, was Kant's doctrine that there can be no science of human actions, though its importance has seldom been recognized by those who are committed to empirical psychology. Human actions are the product of human reason, deliberation, and choice, and Kant held that insofar as a man's reason is involved, his behavior is not explicable in terms of the mechanical laws of nature. He acts freely and is determined only by rational laws of his own creation. This was similar to Spinoza's doctrine of freedom and activity. It raises all sorts of problems about the relationship between reason and emotion and between mind and body, problems that Kant did not seriously tackle. His concept of a rational being as a noumenon which was somehow related to a phenomenal embodied self was a metaphysical model that dramatized difficulties connected with the mechanical explanation of thought and rational action which Descartes had used a different model to depict. Kant laid more stress on the concept of will and rational action than did Descartes, but both men picked out a crucial problem for the development of psychology to which no satisfactory answer has yet been given.

TRANSITION FROM PHILOSOPHY TO SCIENCE

The history of psychology as thus far reviewed is in the main a history of the philosophy of mind, and the issues discussed have been mainly philosophical issues. The rest of the history, however, will be concerned with the slow but progressive disentanglement of psychology as an empirical science from philosophical speculation.

Although it is possible to consider Aristotle's *De Anima* as the transition from presystematic to systematic

psychology, the transition from philosophy to empirical science cannot be pinpointed so precisely. This was not so much a transition as a process of differentiation. Indeed, it began with Aristotle, but it becomes unmistakable in the psychologies of Descartes and Hobbes, both of whom were affected by the impact of Galileo's physics. Both framed hypotheses about the physical and physiological mechanisms of consciousness and behavior that were in principle testable by observation and experiment. From Descartes and Hobbes the main line of development in empirical psychology was through the British empiricists Locke, Berkeley, and Hume.

EIGHTEENTH-CENTURY BRITISH PSYCHOLOGY. Locke's new way of ideas laid the foundations for the twin doctrines of sensationism and associationism. The theory was that the mind is composed only of sensations and mental images (mental images being faint copies of sensations), that all complex percepts or ideas are formed through association, and that all trains of thought arise through association. Locke's analysis of mind was not so simple as that. He included ideas of reflection, abstract ideas, and the self, or possessor of sensations and ideas. Berkeley contested the existence of abstract ideas and furthered the development of associationism by giving an associationist explanation of the perception of the third dimension of space—another hypothesis that was to become the subject of experimental study. Hume further refined sensationism by eliminating the self on the basis of the negative result of his attempt to observe it by introspection. The next important step was taken by Hartley, who proposed a neural basis of conscious processes. His hypotheses, too, could in principle be tested by observation and experiment. Further refinements and elaborations of associationism are to be found in the works of James Mill, J. S. Mill, Thomas Brown, and Alexander Bain. The associationist doctrines spread to the Continent and as experimental psychology later returned to England and went to the United States.

A second major influence on the advance of psychology toward the status of an empirical science was provided by the biological sciences, notably in the evolutionary doctrine of Darwin. This influence was later to prove one of the causes of the disruption of associationist psychology.

Hartley. While David Hartley (1705–1757) was practicing medicine, he made many observations of psychological interest and wrote his major opus, *Observations on Man, His Frame, His Duty, and His Expectations* (1749). It was a thoroughgoing attempt to provide a neurophysiological basis for the mental processes of sensation, imagery, and association. Influenced by Newton's *Opticks,* he proposed an explanation of conscious experience and association in terms of vibrations transmitted through nerves, which were conceived of as solid fibers, thus breaking from the earlier conception of nerves as hollow tubes for the conduction of the animal spirits. For every kind of sensation there are different kinds of vibrations or vibrations differently located; corresponding to images or memories, there are vibratiuncles, miniature vibrations that can persist after the larger vibrations have subsided and which form the physical substratum of memory. The associative processes occur by virtue of the fact that if two stimuli occur simultaneously and produce two corresponding vibrations in two regions of the brain—say, vibration *A* arising from a visual stimulus and vibration *B* arising from an auditory stimulus—the repetition of only the visual stimulus producing vibration *A* will arouse vibration *B* in the absence of the original stimulus that produced *B.* This is a simple translation into neurophysiological terms of the traditional principle of association of ideas, explaining, for example, the association of thunder with lightning. Hartley further advanced associationist theory by suggesting ways in which some of the several special laws of association—contiguity in space, contiguity in time, contrast, and similarity—could be reduced to the single law of association by temporal contiguity. He also offered a more detailed account than had yet been given, in terms of association, of the formation of general ideas.

Brown. As professor of moral philosophy in Edinburgh, Thomas Brown (1778–1820) delivered a series of lectures subsequently published under the title *Lectures on the Philosophy of the Human Mind.* Though not himself an associationist, he made very important contributions to the theory of association, which he preferred to describe as suggestion. Two of his ideas were of especial importance. First, he distinguished between simple suggestion, which is association in the commonly accepted sense, and relative suggestion, which is not in any sense an associative process but is a process that was later to be described by Charles Spearman as the "eduction of relations." Second, Brown formulated the secondary laws of association—the principles of recency, frequency, duration, liveliness, and so on. These were later to become the subject of innumerable experimental studies.

NINETEENTH-CENTURY BRITISH PSYCHOLOGY. Brown's philosophy was severely criticized by Sir William Hamilton (1788–1856) in his *Discussions on Philosophy and Literature* (1852) and his *Lectures on Metaphysics and*

Logic (posthumously published in 1859–1860), but Brown was defended with no less force by J. S. Mill in *An Examination of Sir William Hamilton's Philosophy* (1865). Hamilton, who was professor of logic and metaphysics at Edinburgh from 1836 until his death, had been greatly attracted by German philosophy and contributed to the rise of the British idealistic school of philosophy later to be represented by T. H. Green and F. H. Bradley. This school, deriving its inspiration from the intellectualist and idealist thought of G. W. F. Hegel and other Continental philosophers, had no common ground with the mechanistic empiricist and physiological approach of the British psychologists, but in its criticism contributed to the refinement, as well as the demise, of associationism. It was Bradley who, in attacking the atomistic features in associationism, phrased the dictum "Association marries only Universals." This theme was to be developed in an original way in G. F. Stout's doctrines of noetic synthesis and relative suggestion.

James Mill and John Stuart Mill. Associationism reached its zenith in the work of James Mill (1773–1836). An economist and historian rather than a philosopher or psychologist, he learned his philosophy—hedonistic utilitarianism—from Hartley. His psychology, however, was a refinement of Hartley's and his analysis of mind was much more acute. *The Analysis of the Human Mind* appeared in 1829. Mental life was reduced to sensory elements, and the development of complex ideas was explained by the principle of association. Mill gave a clearer account than had Hartley of the way in which the several laws of association could be reduced to the single law of contiguity. He refined previous accounts of emotional experience in terms of sensations. Like Hartley, he attempted to apply the principles of associationism to the explanation of the complex phenomena of conscience and religion.

John Stuart Mill (1806–1873), his son, was a more subtle and acute philosopher than his father. He was certainly more disposed to take seriously any objection to a theory he wished to defend. In his rational and reasonable way he was inclined to make concessions that resulted in his rejecting the original theory. He sacrificed simple hedonism by conceding that pleasures might differ in quality. He gave up associationism by introducing the concept of mental chemistry—the idea that mental compounds, like chemical compounds, might exhibit properties not deducible from the properties of the elements. This breach in the associationist defenses was to be widened later by doctrines of creative synthesis and Gestalt qualities and the biological concept of emergent evolution—ideas all at variance with pure associationist doctrines. J. S. Mill was less concerned with sensationism as a psychological doctrine than with its philosophical counterpart, phenomenalism—the description of material things and the physical world in terms of sense data or "permanent possibilities of sensation."

Bain. Though in the associationist tradition, Alexander Bain (1818–1903) was less interested in the philosophy of mind than in psychology as an empirical science. He was emphatic in his demand that psychology should be cleared of metaphysics. His *Manual of Mental and Moral Science* (1868) was virtually a textbook of empirical psychology. It was a condensation of his two major works, *The Senses and the Intellect* (1855; rev. ed., 1894) and *The Emotions and the Will* (1859). He was thoroughgoing in his insistence on the need for a physiological basis for psychology not merely in general terms but in terms of known physiological facts, about which he made it his business to be well informed. As far as this implied a philosophy of mind, it found expression in his formulation of the principle of psychophysical parallelism. Especially important were his accounts of habit formation and learning. His treatment of habit was in large measure the inspiration behind the eloquent chapter on this topic in William James's *Principles of Psychology.* E. L. Thorndike and other "learning theorists" owe to Bain the first clear formulation of the law of effect, the principle that responses are ingrained by the reward of pleasure. Even his sillier theories contributed to enlightenment. One of the silliest theories in the history of psychology—that maternal love is based on the pleasurable tactile sensations experienced from contact with a baby—foreshadows the subtler theories of Freud concerning erogenous zones in the body and, more remotely, the "releaser mechanisms" of the ethologists. Bain's associationism was not an ideology. It was merely that he had assimilated the dominant features of the current psychological climate of opinion.

Two other developments were to complete the transformation of psychology from a branch of philosophy into an empirical science: (1) the impact of the theory of evolution and (2) the establishment of laboratories for experimental psychology. The theory of evolution had its origin in England in the work of Darwin; the idea of laboratories for experimental psychology came chiefly from the Continent.

EVOLUTIONARY PSYCHOLOGY. Darwin's theory of evolution as set out in his *Origin of Species* (1859) was a very large theory, but it was a scientific, not a philosoph-

ical, theory. It was supported by an enormous body of empirical observations. Theories of evolution date back to antiquity. Charles Darwin's grandfather Erasmus Darwin had adumbrated a Lamarckian theory of evolution. Alfred Russel Wallace anticipated Darwin's theory by a few months. Herbert Spencer (1820–1903), who had propounded philosophical and psychological theories of evolution for some years before the appearance of the *Origin of Species,* was accordingly well placed to capitalize on Darwinism in the development of his own ambitious "synthetic philosophy."

Darwin (1809–1882) himself wrote on distinctively psychological topics. His *Descent of Man* (1871) discusses the similarities between the mental processes of man and of animals. His work *Expression of Emotions in Man and Animals* gives an evolutionary interpretation of changes in features and postures and assigns biological utility to these changes. The evolutionary approach stimulated many studies by amateur and professional naturalists. G. J. Romanes (1848–1894) collected evidence for the continuity of development from the animal to the human mind, and Sir John Lubbock (1834–1913) was among the first to use laboratory techniques in the study of insects. Laboratory studies like these were to be developed later on a grand scale by such American comparative psychologists as E. L. Thorndike and R. M. Yerkes.

Galton. Sir Francis Galton (1822–1911), the versatile cousin of Charles Darwin, contributed to meteorology, anthropology, anthropometry, and psychology and to the development of statistical and other metric methods in psychology. Among his major interests was the inheritance of mental characteristics, for the study of which he devised ingenious methods. He stressed heredity as a determinant of mental life and behavior. His records of the behavior of twins are reminiscent of the Leibnizian concept of a preestablished harmony. According to his records, twins can behave exactly like two clocks each causally insulated from environmental influences and from each other, behaving similarly and thinking in unison almost entirely in consequence of the similarities of their innate constitution. His major psychological works were *Hereditary Genius* (1869) and *Inquiries into Human Faculty* (1883). He set up the first two English psychological laboratories—the first at the International Health Exhibition of 1884 and the second in the South Kensington Museum. He pioneered the application of physical and psychometric tests in schools.

Ward and Stout. Philosophical psychology was to feel the impact of the new biological approach. James Ward's revolutionary article on psychology in the ninth edition of the *Encyclopaedia Britannica* (1886) mounted a devastating attack upon associationism, recasting psychology in terms of a "psychoplasm," or "presentational continuum," which, like bodily tissues, undergoes progressive differentiation and integration. Ward's distinguished pupil Stout wrote *Manual of Psychology,* a standard text for some three decades, in 1898. This was described as being written from a genetic point of view; thereafter, almost every textbook of psychology had a biological orientation.

EMPIRICISM IN EUROPE. The empiricist philosophy was introduced into France by littérateurs and essayists like Voltaire and Denis Diderot, not by philosophers or psychologists. Voltaire had lived in England from 1726 to 1729, and so was in a position to introduce British ways of thought in philosophy into the intellectual life of France. Diderot had a clearer understanding of British empirical psychology. He particularly interested himself in the mental life of persons deprived of one sense—for example, sight.

The first of the French empiricist philosophers to contribute to sensationism was Étienne Bonnot de Condillac (1715–1780). Diderot had been concerned with the mental life of persons deprived of one sense; Condillac started from the imaginary case of a person deprived of all senses except one. He took the case of a statue endowed only with the sense of smell, selecting smell because of its relative simplicity. From this he proceeded to add other senses and to explain in sensationist terms attention, memory, imagination, and reason. He attached no importance to association. He believed that the experience of one sensation after another is *ipso facto* a comparison of the two and that the occurrence of the unpleasant sensation constitutes the will to terminate the sensation. Condillac's sensationism was perhaps the simplest and most elegant form of the doctrine in the history of psychology. His views are set out in the *Traité des sensations* (1754).

Claude-Adrien Helvétius (1715–1771), author of a volume of essays titled *De l'esprit* (1758), was a minor social and political philosopher who seized upon Locke's empiricism and concept of the tabula rasa to defend an extreme doctrine concerning the equality and perfectibility of men. His basic thesis was that all differences between men are due to differences in experience and education. All error was due to passion or ignorance.

The doctrines that Helvétius derived from Locke were to return to England in the works of William Godwin, especially in his *Political Justice* (1793). Like

Helvétius, Godwin taught that all men are equal at birth and that their subsequent differences were due to experience and education. Voluntary actions originate in opinions, which can be changed by rational persuasion. Vice is error, which can be corrected. In Helvétius and in Godwin the association of empirical philosophy with an intellectualist hedonism is displayed in its most extreme form.

Through Condillac the influence of Locke spread to Italy and Switzerland. In Italy this influence is to be seen in the teachings of several all-but-forgotten writers. In Switzerland, Charles Bonnet (1720–1793) of Geneva was the outstanding figure in empirical philosophy. His chief work in psychology was the *Essai analytique sur les facultés de l'âme* (1760). Although he followed Condillac for the most part, Bonnet differed chiefly in the importance he attached to physiological explanations.

GERMAN PSYCHOLOGY AND EXPERIMENTATION. Throughout the seventeenth, eighteenth, and early nineteenth centuries German psychology was dominated by the philosophical doctrines of Leibniz, Kant, and Hegel, each of whom contributed to a rationalist idealism very unfavorable to the development of psychology as a science.

Hegel. Georg Wilhelm Friedrich Hegel (1770–1831) has received scant attention in the histories of psychology, understandably so since his form of rationalism is the most extreme antithesis to the empiricist philosophy that had favored the development of psychology as an empirical science. He is, however, not without importance in the history of psychology.

One of Hegel's theses was that it is a mistake to suppose that complex phenomena are explained only by reference to simpler phenomena, that we can, for example, understand religion in its developed form by the study of cults of primitive people or that we can understand man only through the study of lower forms of animal life. In this he challenged what had long been and still is a basic principle of comparative psychology, but Hegel's thesis survives in the view of psychologists who hold that the proper study of humankind is man and that we should begin with civilized man in advanced societies. It lives on in the contention of Freudian psychologists that the evidence for infantile sexuality can be appreciated only in the light of adult sexual behavior.

Equally important for psychology is the Hegelian dialectical progression—thesis, antithesis, synthesis. When this progression is stated as an empirical observation of movements of thought and action, not as a metaphysical principle or a principle of logic, it illuminates many sequences in the history of politics, philosophy, and science. A dialectical progression is illustrated in the fate of Hegel's own philosophy. Its influence in Germany was short-lived. His rationalistic thesis issued in an empiricist antithesis, Wundtian experimental psychology. The dialectical progression is illustrated by the British vogue for Hegelianism among philosophers who found in it an antithesis with which to confront the prevailing empiricist philosophy and psychology. The progression is illustrated by the sequence from Hegel's idealist thesis to the antithesis of dialectical materialism that was to become a central tenet of communist philosophy. Although it provides no comprehensive philosophy of history, the concept of dialectical progression affords a rather more subtle and articulate account of historical movements than conventional, commonsense accounts in terms of "the swing of the pendulum."

Hegel's doctrines were associated with, and conferred philosophical status upon, a widespread romantic and mystical philosophy of nature according to which everything in nature had some spiritual and symbolical significance. The influence of this philosophy of nature persisted far into the nineteenth century and in the biological sciences favored vitalistic, as opposed to mechanistic, accounts of mind, body, and nature. Psychologists divided progressively into two groups. The first comprised the philosophers—that is, those who primarily taught philosophy and whose philosophy of mind contained much metaphysics. The second group consisted of natural scientists whose approach was from mathematics, physics, and the biological sciences. The distinction is not sharp, since romanticism and metaphysics were in the air that every German student, even students of the natural sciences, breathed.

The first steps in the transition from the philosophy of mind to scientific psychology were taken when Kant challenged psychologists to show that their subject could claim scientific status. This challenge was taken up by Herbart, Ernst Heinrich Weber, and Fechner. That it could be an experimental science was argued by Weber, Fechner, Johannes Müller, Hermann von Helmholtz, and others. Wundt finally established it as a science that required a distinctive kind of laboratory.

Herbart. Johann Friedrich Herbart (1776–1841) set out to establish a basis for psychology other than that of the prevailing "faculty" psychology associated with Christian Wolff (1679–1754), a disciple of Leibniz and precursor of Kant who was much less distinguished than either. Herbart tried to show that the laws describing mental process could be put into precise mathematical form.

Herbart's first achievement was the grafting of associationism onto a rationalist metaphysical root. The soul was retained, serving the traditional function of giving unity to the mind, but the data of empirical psychology were, as in associationism, sensations and ideas. In Herbart's system ideas were not just passively associated. They interacted by attractions and repulsions in accordance with which they were drawn into or forced out of consciousness. The behavior of ideas in Herbart's psychology resembles that of the "reals" in his pluralistic metaphysics. Two "reals"—for instance, *A* and *B*—differing in quality, tend to disturb each other because of their difference, but each also tends to preserve itself by resisting the disturbing effect of the other. This principle of self-preservation is reminiscent of the Spinozistic doctrine that "everything that is in itself endeavors to persist in its own being" and, when applied in Herbart's psychology, foreshadows the concept of homeostasis that was to be current in psychology a century later.

Herbart's account of the way in which ideas enter consciousness and are expelled from it represents a phase in the history of the theory of the unconscious midway between Leibniz and Freud; his concept of the apperceptive mass, a system of ideas bound together by mutual attraction, was still current when psychoanalytic writers were developing the concept of a mental complex. Herbart's metaphysics and mathematics were to be forgotten, and he did not contribute directly to the development of psychology as an experimental science. His most lasting influence was in the field of educational psychology, chiefly in the application of his theory of apperception to the process of learning.

Lotze. Rudolf Hermann Lotze (1817–1881) succeeded Herbart in the chair of philosophy at Göttingen. His most influential work was his *Medizinische Psychologie* (1852), the first systematic work on physiological psychology and one of the very few written by an author qualified in both physiology and philosophy. Against the then prevailing view he defended the thesis that every mental phenomenon has its physiological counterpart and that the laws which apply to inorganic matter also apply to organic matter. Final causes, vital and mental forces, and the soul itself can act only through mechanical causation. He insisted, however, that physiology alone cannot explain mental phenomena. Lotze is best known in psychology for his doctrine of local signs, a contribution to the theory of space perception.

Weber and Fechner. Experimentation and the use of quantitative methods in psychology were greatly advanced by Ernst Heinrich Weber (1795–1878) and Gustav Theodor Fechner (1801–1887), who were colleagues in the University of Leipzig and who both taught Lotze.

Weber taught anatomy and physiology. His early work *De Tactu* (1834) reported studies demonstrating the difference between muscle sense and touch. These studies were extended to pain, pressure, and temperature, through which emerged the concept of thresholds and the famous law that has come to be called Weber's law. This states that the smallest increment in a stimulus required to produce a difference in the sensation experienced is not an absolute amount but is relative to the magnitude of the stimulus in question. Like most German scientists of his time, Weber was to some degree under the spell of the current metaphysics and the romantic philosophy of nature, but neither of these influenced his experimental studies. His metaphysics and his science were kept apart.

With Fechner the case was different. Fechner's intellectual life was a pilgrimage from physics and chemistry, through physiology and medicine, to metaphysics and mysticism. From an early age he had been preoccupied with the problem of the relation between matter and spirit. He was attracted to a form of panpsychism according to which not only man and the lower animals have consciousness but also the earth and the other planets—indeed, all material things. In this view all souls are parts of the soul of the universe.

Fechner concluded, on the obscurest of grounds, that the mystery of the relation between mind and body would be resolved by ascertaining the quantitative relations between stimuli and sensations. He suggested that Weber's law could be put into a quantitative form. Weber's law thus became the Weber-Fechner law, according to which the relation between stimulus and sensation is expressed in the formula $S = k \log R$ where S is the experienced intensity, R is the physical intensity, and k is a constant for the particular sense in question. For the verification of this law Fechner designed what are known as the psychophysical methods. These methods have been used in the most tedious of laboratory exercises to which many generations of students of experimental psychology have since been subjected, and the published results of these exercises are among the most tedious controversies in the history of science. But the possibility of experiment and measurement in psychology was established—paradoxically, by a metaphysical mystic. The metaphysics and the mysticism were soon forgotten, but the exercises live on.

Beneke. Friedrich Eduard Beneke (1798–1854), a contemporary of Herbart, was another philosopher who contributed to the foundation of a science of empirical psychology, which, he claimed, was the basis of all philosophy. Like Herbart, he set out to provide a basis for psychology other than that of a doctrine of faculties, and like Herbart, he stressed the activity of the mind. Among his works on psychology are *Lehrbuch der Psychologie* (1832) and *Die neue Psychologie* (1845). Because of his rejection of the prevailing Hegelian philosophy of the Absolute, Beneke was dismissed from his post in the University of Berlin, but after Hegel's death he was reinstated. His best-known contribution to psychological theory was his doctrine of mental, as contrasted with physiological, traces for the explanation of the facts of memory. This doctrine was later to be developed in Great Britain by Stout.

Müller. Johannes Müller (1801–1858) was a contemporary of Beneke at Berlin. He was the first to hold the title of professor of physiology. (Hitherto, the subject had been taught as a branch of medicine.) He had been under the influence of the prevailing philosophy of nature but contributed to the clarification of the concepts of mind, body, and nature by distinguishing the mental principle, which is restricted in its operation to the nervous system, from the vital principle, which is diffused throughout the organism. He was also preoccupied with the opposition between nativistic and empiricist explanations of space perception as represented, respectively, in the doctrines of Kant and Herbart. Müller reformulated the issue in terms that made it possible to submit the question to experimental tests. He also formulated the theory of specific energies in the nervous system—the hypothesis that the sensory qualities are generated by specific activities of the organs of sense or by specific differentiation in corresponding realities in the brain.

Helmholtz. Hermann von Helmholtz (1821–1894), Müller's distinguished pupil, is acknowledged to be the most outstanding of the physicist-physiologists who have contributed to the development of experimental psychology. In the range of his pioneering studies he has been compared with Francis Galton. His publications were more numerous than Galton's, and his investigations were carried further. He was the first to make a realistic calculation of the speed of nervous impulses, which are important, among other things, in the study of reaction times. He developed Müller's doctrine of specific energies and Thomas Young's three-color theory of vision.

Helmholtz's *Handbuch der physiologischen Optik*, published in three volumes (1856–1866), remained an authoritative text for many decades, although it was not translated into English until 1924–1925. No less outstanding were his contributions to the theory of hearing and the related subjects of phonetics and music. He was essentially a scientist with little interest in philosophy and still less patience with transcendentalism. There is, however, much in his writings of philosophical interest—for example, his puzzling concept of unconscious inference in perceptual judgments. His discussions of the principle of the conservation of energy are important in the history and philosophy of science.

Wundt. The last phase in the transition of psychology from a branch of philosophy to psychology as an independent empirical science is conveniently dated as beginning in 1879, when Wilhelm Wundt (1832–1920) established the first psychological laboratory. Wundt's chief claim to a place in the history of philosophy arises from the conceptual system in terms of which he interpreted the experimental data from his own and other laboratories. His philosophy of mind deviated from the simpler forms of atomistic sensationism in that the ultimate elements of mind were, according to him, of two kinds, sensations and feelings. He and his disciples devoted much energy and skill to defining the differences between sensations and feelings and to elucidating his curious tridimensional theory of feeling, but the general program was to analyze experience into its elements, to define the fundamental attributes of these elements, and to formulate the laws in accordance with which these elements are combined. The account leaned heavily on the principle of association but deviated from traditional associationist doctrines in introducing a concept of creative synthesis. This concept was a variant of the concept of apperception and embodied a theory of attention. It had some points in common with J. S. Mill's conception of mental chemistry and in some degree foreshadowed later theories of emergent properties and the doctrine of the Gestalt psychologists that a complex experience is more than the sum of its parts. His most influential work was *Grundzüge der physiologischen Psychologie* (1873). In later years he published two works that contributed to the incursion of psychology into sociology and anthropology.

Ebbinghaus and Külpe. Among other outstanding experimental psychologists were two of Wundt's pupils, Hermann Ebbinghaus (1850–1909) and Oswald Külpe (1862–1915). Wundt's laboratory research had been chiefly concerned with sensation and perception and with relatively simple processes of reaction and association. Ebbinghaus and Külpe extended the experimental

method into the study of the higher and more complex functions of memory and the processes of thinking.

In a monumental work, *Über das Gedächtnis* (1885), Ebbinghaus published the results of what has been described by J. C. Flügel in his *A Hundred Years of Psychology* as "the most brilliant single investigation that has ever been made in experimental psychology." Ebbinghaus's outstanding achievement was to extend the experimental method to the "higher thought processes." He was the first to establish quantitative laws concerning the process of memorization. In 1894 he succeeded Theodor Lipps, a pupil of Wundt's most widely known for his studies in psychological aesthetics, in the chair of psychology at Breslau. There Ebbinghaus pioneered in the study of intelligence and devised the completion test, which remains an important component of intelligence tests.

Külpe directed the laboratory at Würzburg, which achieved great fame through its investigations of willing and judging. Through the discovery of imageless thoughts these studies contributed both to the breakdown of sensationism and, in consequence of the inconclusive disputes this discovery provoked, to the behaviorist revolt against introspective methods. At Würzburg as at Leipzig confusion arose through the interpretation of experimental data in terms of implicit philosophical concepts and assumptions, and the conclusions drawn have had to wait for review in the light of further clarification of the distinction between empirical psychology and the philosophy of mind.

THE SHIFT TO THE UNITED STATES. In the age of Wundt, psychology was a Germanic science, and Germany was the heart of the empire. Mainly through Wundt's influence upon those who came to Leipzig from the United States, psychology became an American science with the United States as the new seat of dominance. Among those who studied abroad and then returned to America were Stanley Hall, who established the first American psychological laboratory at Johns Hopkins in 1888; J. McKeen Cattell, who after several years as assistant to Wundt founded the laboratory at Pennsylvania; and Hugo Münsterberg, who, having established a laboratory at Freiburg, was invited by William James to Harvard in 1892. In the same year E. W. Scripture took charge of the laboratory at Yale. By 1897 there were fifteen psychological laboratories in the United States, and by the end of the century there were twenty-six, all based, to begin with, on the laboratory in Leipzig. Most of Wundt's American pupils, however, were soon to deviate from the German pattern and to open up approaches characteristically American—allergic to philosophical speculation, distrustful of introspective methods, and much concerned with the practical applications of their science. Hall became famous for his studies of adolescence. Cattell, more influenced by Galton than by Wundt, concentrated on the measurement of individual differences. Münsterberg's interest turned to applications of psychology to industry and criminology. The mantle of Wundt fell upon E. B. Titchener, an Englishman from Oxford who after his studies at Leipzig went to the United States to develop experimental psychology at Cornell.

THE ESTABLISHED ORDER OF 1900. Wundtian psychology was one important form of and ingredient in what has been called the established order of 1900, against which many revolts were to be mounted. There were, in fact, at least two established orders, one in Britain, represented by Ward and Stout, and the other in the United States, represented by Titchener. These were very different establishments, but they had in common a foundation in some form of body-mind dualism and the acceptance of the facts of consciousness, observed by introspection, as defining the subject matter of psychology.

Ward. James Ward (1843–1925) presented his own system as a sort of synthesis of the too objective thesis of Aristotle's psychology and the too subjective antithesis of Descartes's psychology. His basic conceptual framework was doubly tripartite. In his analysis of experience he distinguished the three modes of consciousness—cognition, feeling, and conation; his analysis of each kind of experience referred to a self or ego, an act or mental attitude, and a presentation (a mental object, sensation, or idea). The most interesting features of his system are contained in his detailed analysis of the phases of development from simple sensation to perception and from perception to the construction of a memory thread and an ideational tissue. Though qua psychologist Ward can be treated as a dualist, his background metaphysics was a variant of an idealistic monadology of the Leibnizian type.

Stout. G. F. Stout (1860–1944) developed Ward's psychology in an individual way, creating an original and independent system. As a psychologist Stout, like Ward, developed a dualistic psychology, but as a philosopher he developed an original theory of mind, body, and nature.

Titchener. E. B. Titchener's laboratory at Cornell was the temple of the Wundtian form of the established order, and Titchener (1867–1927) was its high priest. Here as elsewhere, however, empirical psychology continued to

be inextricably entangled with philosophy. Titchener's deviations from sensationist and associationist psychology were less fundamental than he himself believed. He was a dualist, and he confessed to a bias in favor of sensationism. He was reductionist in his treatment of conation. He differed from the classical sensationists in accepting feelings as basic elements; he also differed from them in the treatment of the elements as existences, as contrasted with meanings. He sought to explain complex mental states as arising from the synthesis of elements and thus to display the structure of the mind. Accordingly, he is described as a structural, as opposed to a functional, psychologist. His cardinal tenet, which was to become the major object of attack, was his thoroughgoing proclamation of introspection as *the* distinctive method of psychology. His two most important works were his *Lectures on the Psychology of Feeling and Attention* (1908), a detailed exposition of the thesis that "the system of psychology rests upon a threefold foundation: the doctrine of sensation and image, the elementary doctrine of feeling and the doctrine of attention," and the *Lectures on the Experimental Psychology of the Thought* (1909), an equally thoroughgoing examination of the claims for the discovery of an imageless thought element and a polemic against the doctrine of Franz Brentano and Stout that references to object is the criterion of mind.

REVOLTS AND THE ERA OF THE SCHOOLS. The established order of the United States and the established order of Britain were to become the objects of attack from four directions: (1) The behaviorist attack directed in the main against dualism, the concept of consciousness, and the reliance upon introspection; (2) the attack of the Gestalt psychologists against all forms of psychological atomism; (3) the psychoanalytic attack against the overemphasis on conscious processes and inadequate recognition of the unconscious mind; and (4) the attack of the hormic psychologists, which was directed against the intellectualism of traditional psychology—that is, the overemphasis on cognitive processes and the relative disregard for conation or purposiveness in the explanation of conscious experience and behavior.

In the four revolts the schools were all fighting on more than one front. Each was attacking traditional psychology, and each engaged in polemics with the other revolting schools. Confusion was increased by the fact that within each school there were conflicting factions and by the general failure to distinguish straight empirical issues from issues of philosophy and of linguistic usage.

Behaviorism. The conception of psychology as the study of behavior and as an essentially biological science dates back to Aristotle, but behaviorism as an ideology can be dated precisely. It began in 1914, when J. B. Watson (1878–1958) published *Behavior* while a professor at Johns Hopkins University.

This book was a protest and a revolt against dualism, the concept of consciousness, and *any* use of the introspective method in psychology. Psychology is to be the study of behavior by objective methods. It was a protest in defense of animal psychology, in which statements about the animal mind and the consciousness of animals must be pure guesswork, and it was a protest against the interminable and inconclusive disputes between introspective psychologists about the differentiation of sensations and feelings, the James-Lange theory of emotion, and imageless thought. It was also an attack on the traditional theory of consciousness in which some sort of mental stuff was thought to be the subject matter of psychology.

In *Behavior* and two other important books, *Psychology from the Standpoint of a Behaviorist* (1919) and *Behaviorism* (1924), Watson developed his distinctive account of all the major topics that constitute psychology. Like the structuralists he set out to exhibit complexes in terms of simple elements, complex responses to situations as derived from simple responses, native and acquired. The analysis of behavior was in terms of stimulus and response (an analysis to be elaborated later by E. C. Tolman in terms of intervening variables). Sensation and perception were described as responses to present stimuli and constellations of stimuli, memory and learning as responses to past stimuli and neural traces, feelings and emotions as types of sensorimotor responses, and thinking as subvocal verbal behavior. Introspection itself was redescribed as verbal behavior. In his system Watson included much that was irrelevant to the major principle—for example, a bias toward explanations in terms of environmental influence and a bias against explanations in terms of heredity. He had a special bias against the concept of purpose, though later behaviorists found no difficulty in assimilating purposive behavior as goal-directedness. His laws of conditioning were the old laws of association transformed into generalizations about bonds between simple reflexes instead of between simple ideas.

Watson's writings were naive and often confused, but his behaviorism sailed on the tides of the time. Behaviorism was inevitable. Watson's behaviorism was fortunate in that it was reinforced by the most important philo-

sophical movements of the period, positivism and physicalism. It was also reinforced by the logicians and the methodologists of the inductive school, who maintained that scientific laws state correlations between observables. Watson accordingly assumed that because mind was unobservable, it could not be discussed or referred to in science. When logicians later began to proclaim that scientific systems were hypothetico-deductive, such behaviorists as Tolman and C. L. Hull conceded the importance of unobservables in the form of intervening variables and hypothetical constructs. This return to the methodology of Galileo made any simple form of behaviorism difficult to maintain. Nevertheless, B. F. Skinner stuck to the old inductive concept of scientific method and proclaimed that his findings involved no theory. Behaviorism was further supported by a number of outstanding experimental psychologists—for example, K. S. Lashley and W. S. Hunter—sympathetic to Watson's approach.

Lashley was primarily a neurophysiologist who as a behaviorist was more sympathetic to the views of the Gestalt psychologists than to those of Watson. He contributed in an important way to the advance of knowledge concerning the localization of the higher functions in the cortex.

Hunter, a distinguished experimental psychologist, rallied to the support of behaviorism through an odd philosophical argument, based on a very naive form of realism, that consciousness or experience is merely a name applied to what other people call the environment. This argument is reminiscent of a characteristic doctrine of Ward and Stout that the subject matter of psychology comprises "the whole choir of heaven and earth" as it appears to the observer, a view later to be defended by the Gestalt psychologists in terms of the behavioral, as contrasted with the geographical, environment—another variant of the view that things as they appear are appropriate objects of psychological science.

As professor of psychology at the University of California, E. C. Tolman (1886–1959) developed an original system of purposive behaviorism that had perhaps much more in common with the psychology of McDougall than it had with the psychology of Watson. Watson was preoccupied with responses to stimuli. Tolman described Watson's behaviorism as molecular, for it was concerned mostly with physiological details; his own he described as molar, for it was concerned with external and integrated behavior and with emergent properties.

Clark L. Hull (1884–1952), professor at Yale, is known for his inventiveness and originality. His contribution to behaviorism reflects his own interest in methodological studies and the concept of hypothetico-deductive systems. He constructed a miniature system of this type aimed at a rigorous ordering of some of the basic laws of behavior. His deductive dream and his attempt to develop a Galileo-like resolution of behavior into simple externally initiated movements bore a marked similarity to the mechanistic system of Hobbes.

Behaviorism is not strictly an arguable thesis; it is a pronunciamento, a policy statement. The traditional psychologist declares, "I propose to study consciousness by introspection"; the behaviorist says, "I do not; I propose to study behavior by objective methods." The issue is almost as simple as that. There are, however, many arguable issues in particular systems of behaviorism. Reasons can be given for and against policy decisions. There are larger philosophical issues that cannot be evaded.

Roughly three types of behaviorism have emerged: a metaphysical type that says that consciousness does not exist; a methodological type which says that consciousness is not amenable to scientific procedures of investigation; and a radical analytic type, defended chiefly by philosophers, according to which mental facts can all be analyzed in terms of behavior and dispositions to behavior. In Watson's behaviorism and in many others these issues are confused. The behaviorists, no less than Titchener, confused questions of empirical fact with questions of philosophical analysis. It is not possible to know what an emotion is by the introspective observation of emotional states. A prior decision has to be made concerning what to observe, what is to count as an emotion. In the same way it is not possible to know what behavior is by the objective observation of behavior. A prior decision has to be taken about what to observe and about what is to count and what is not to count as an example of behavior. For example, before describing a movement of the body like raising an arm as signaling to a friend or testing the direction of the wind, a person must know what the agent had in mind. This inadequate attention to the question of what constitutes behavior was one of the major weaknesses of behaviorism. Behaviorism is no less riddled by interminable and inconclusive disputes than is introspective psychology. Nevertheless, it has contributed very effectively to the advance of psychology as a biological and an experimental science.

Gestalt psychology. The term *Gestalt psychology* applies primarily to a school of psychology pioneered by Max Wertheimer (1880–1943), Kurt Koffka (1886–1941), and Wolfgang Köhler (b. 1887). Their polemic was directed chiefly against the atomism of traditional psychology and of the established order. They opposed the

thesis that perceptual experience is to be explained by a bricks and mortar account of the combining of simple sensations. Their positive thesis was that what is experienced is always organized and consists of wholes which are greater than the sum of their parts. Like all revolutionaries, they exaggerated the difference between their own ideology and traditional doctrine. The fact with which they were concerned had preoccupied philosophers and psychologists from the beginnings of systematic thought. Aristotle's formal cause was a Gestalt concept, and Kant had grappled with the problem in his treatment of the categories; Ward and Stout had grappled with it in their accounts of the development of the perception of space, time, thinghood, and causality, and Mill had seen the problem when he wrote about mental chemistry. Christian von Ehrenfels (1859–1932), an Austrian philosophical psychologist, introduced the concept of form qualities. There were also contemporary psychologists—for example, Charles Spearman and Henry J. Watt—who were concerned with the concepts of Gestalt psychology in their own ways.

The outstanding contribution of the Gestalt psychologists was in the number, the variety, and the ingenuity of their experiments. Wertheimer's elegant experiments on the perception of movement were followed by no less elegant experiments by himself, his colleagues, and his disciples on the principles of organization in perceptual experience. In the earlier phases Gestalt psychology was as intellectualist as traditional psychology in its preoccupation with the cognitive experience of the normal adult human mind. Its interest extended, however, to child psychology in the studies by Koffka and to animal psychology in Köhler's studies of insight and learning in apes. Kurt Lewin (1890–1947) used Gestalt concepts in the study of problems of personality and of human motivation. The Gestalt psychologists were distinguished chiefly by their experimental inquiries, but in their writings there are many pronouncements relevant to the philosophy of mind.

The slogan "The whole is more than the sum of its parts" is a near tautology but a useful tautology. The increasing emphasis placed by Köhler and Lewin on field theory (the theory concerning properties of total fields of activity as contrasted to the properties of isolated units) has also contributed to the philosophy of science in its application to psychology.

The concept of isomorphism (the parallelism between phenomenal experience and neural processes) has given a new slant to the discussion of classical theories concerning the relations of body and mind.

The experimental findings of the Gestalt psychologists have been assimilated into empirical psychology. Its evaluation as a philosophy of nature, life, and mind must take into account not only its historical antecedents but also some less well known but important contemporary theories, such as those, for example, of Spearman and Watt.

Alternatives to Gestalt psychology. Charles Spearman (1863–1945) made two significant contributions to the development of psychology in the early decades of the twentieth century. The first was through the development of statistical methods in psychology. Building on the studies of Galton and Karl Pearson, he elaborated his two-factor theory for the analysis of human abilities. His second notable contribution was an attempt to formulate principles of cognition, which he believed to be as basic to psychology as Newton's laws had been basic to physics. It was an ambitious plan in which three noegenetic principles—the apprehension of experience, the eduction of relations, and the eduction of correlates—were set out as necessary and sufficient for the explanation of all the cognitive operations of the human mind. The principles of the eduction of relations had been anticipated by Brown's concept of relative suggestion, but in its detailed elaboration it covered most of the facts of cognitive experience studied by the Gestalt psychologists.

Henry J. Watt (1879–1925) enters the history of psychology through his experimental studies of judgment and the higher thought process at the Würzburg laboratory. After his return to Britain he spent the rest of his life at the University of Glasgow elaborating a comprehensive theory that was finally presented in his *Sensory Basis and Structure of Knowledge* (1925). It is a paradoxical fact that atomism, against which Gestalt psychology was directed, should have received its most precise and systematic formulation by a psychologist preoccupied with precisely the facts that Gestalt psychologists were concerned with.

Watt offered an ingenious alternative to Gestalt theory made possible by the sharp distinction he drew between sensationism and associationism, whereas Titchener had treated them as equivalent doctrines. Watt agreed that traditional psychology rested upon two postulates—(1) that the elements of mind are sensations and (2) that the compounds are produced by association. He not only accepted the first postulate, but he also refined it with great subtlety. He rejected the second postulate, replacing it with the doctrine that complex cognitive experiences arise through a distinct process of integration—a concept to which he gave a new definition and which he illustrated in great detail. Watt produced an

original account of the facts that had previously been interpreted in terms of Mill's mental chemistry, Wundt's creative synthesis, Spearman's noetic principles of eduction, and the principles of Gestalt psychology.

The Gestalt psychologists captured the headlines in the journals of psychology. For a time Spearman had a band of disciples, although Watt's book did not have a second edition. Spearman and Watt had the misfortune of attracting disciples who could neither advance their theory nor excite impassioned critics. Thus, both have been forgotten. Both, however, may be classed among the mute inglorious Miltons of psychology whose works may yet attract the attention of future historians of science.

The philosophy of nature, life, and mind of both Spearman and Watt were, though different from each other, both in the tradition of dualism. That of the Gestalt psychologists was rather different—a dualism of physics and phenomenology. A residual doubt remains. There would appear to be no empirical procedure for deciding between the doctrines of the Gestalt psychologists, of Spearman, and of Watt. The case may again be one in which a choice must be made on grounds of terminological convenience.

Psychoanalysis and derivative schools. The most important revolt against traditional psychology at the turn of the twentieth century was that of Freud and his disciples.

Sigmund Freud created an entirely new psychology—psychoanalysis. This is both a technique of psychotherapy and a body of theory providing a rationale for the technique. The theory developed into an overall account of nature, life, and mind. Freud's philosophy of nature was a conventional nineteenth-century mechanistic materialism predisposing him to an equally conventional preference for physiological explanations of the mind. Thus, it is even more remarkable that his most distinctive and revolutionary doctrines assumed the form of hypotheses to which mechanism and physiology are completely irrelevant.

Central in his system of psychology is the concept of the unconscious. Mind is divided into the conscious, the preconscious, and the unconscious. The conscious is the traditional, familiar, introspectable part of the mind—introspectable thoughts, feelings, and desires. The preconscious consists of all that is out of mind but which can be brought to mind at will or which readily returns to mind in accordance with the accepted laws of association. The unconscious, on the other hand, consists of ideas and wishes, especially wishes, which can be brought into consciousness only by special techniques, of which psychoanalysis is said to be the most fundamental.

Freud's originality did not consist in the discovery of the unconscious, for others before him had hit on this notion, but in postulating that the mind worked in accordance with two different types of laws—those of the primary processes, which included unconscious processes, and those of the secondary processes of thought. The first were ruled by the pleasure principle, the second by the reality principle. The laws of the primary processes were principles of emotive congruence appropriate to wishes. Freud's great contribution to psychological theory lay in postulating these laws of primary processes to explain such phenomena as hysteria, dreams, parapraxes, and so on which were previously unexplained and among which no one had previously seen any connection.

There are some superficial resemblances between Freud's and Herbart's psychology, but these are only superficial. In Herbart's system the contents of the unconscious were ideas; in Freud's system they were mainly wishes. Herbart was concerned with the movement of ideas between consciousness and Freud's preconscious. He had no clear conception of the unconscious mind in Freud's sense. Herbart's explanation of the movements of ideas were formulated in terms of quasi-mechanical forces, efficient causes, whereas Freud's explanatory principles were, in effect, formulated in terms of a truncated type of teleological concept, the Freudian wish. Similarly, Freud's defense mechanisms—sublimation, projection, reaction formation, and the like—were quite unmechanical mechanisms. They were goal-directed procedures for protecting the conscious mind against the unwelcome wishes and ideas that had been repressed.

From first to last Freud was concerned with mental conflict, the conflict between opposing motives. At the beginning he emphasized the conflict between primitive instinctive impulses, mainly sexual, and the need to conform to the rules and norms of society. The emphasis later shifted to the conflict between the life and death wishes. At first the world was astounded and shocked by Freud's theories about sex, especially by his account of infantile sexuality. So prominent was sex in his system that a Freudian explanation of any form of behavior came to be generally thought of as an explanation by reference to unconscious sexual desires. His generalized concept of sex was that all pleasure is essentially the pleasure of sexual experience, including the satisfaction of defecation (anal eroticism) and the satisfaction of sucking and feeding (oral eroticism), as well as the satisfaction derived

from the genital organs (genital eroticism). This general theory of affective experience makes the thesis of infantile sexuality almost tautological. More significant empirically was the thesis of the universality of the Oedipus complex—the thesis that every male child unconsciously wishes to kill his father and have sexual relations with his mother (female children have an Electra complex—the unconscious wish to replace the mother in her relation to the father). These unconscious desires are obvious sources of the conflicts that issue in neuroses and other forms of aberrant behavior.

In Freud's later writings the emphasis was transferred to the conflict between the life-promoting instincts and the desire for death—Eros and Thanatos. When directed outward, the death wish is a source of violence and destruction; when directed inward, it results in suicidal behavior. The concept of the death wish was, however, further generalized. It covered not only the desire to kill and to be destroyed but also the desire to inflict pain and to suffer pain. Thus the odd phenomena of sadism and masochism are explained. As he often did, Freud attempted to reinforce limited hypotheses by highly general theories. The hypothesis of the death wish was based upon the general theory that in all the processes of nature there is a tendency for animated matter to revert to an inorganic state. Slightly less generalized was the theory that all responses to stimuli by an organism were directed to the removal of the stimulus and are thus consummated in unconsciousness, in sleep or death. These speculations were disturbing to his disciples, who felt an obligation to defend them, since these ideas were all but demonstrably mistaken and on the face of it inconsistent with Freud's more basic hedonistic account of human motivation. They were not at all essential to his general theory.

To this phase of Freud's speculations belongs the doctrine that the total personality is organized on three levels—the id, the ego, and the superego. The id consists of the totality of primitive instinctive impulses, and the ego contains the conscious motives. The concept of the superego is the most interesting and original feature of this hierarchy. Although it was often described as the primitive unconscious conscience, Freud explained it as an introjected image of the parent that continued to issue commands and to administer punishment when those commands were disobeyed. Not a few of Freud's disciples have treated the superego as the source of conscience as traditionally conceived and believe it is the explanation of action that accords with moral principles. This, however, was not Freud's view. He was himself a man of great integrity with very definite ethical principles. These principles were not derived from his own superego but are to be explained in terms of the distinction between the pleasure principle and the reality principle. Action in accordance with the pleasure principle is directed to immediate pleasure regardless of consequences; action in accordance with the reality principle is directed to maximizing pleasure in the long run. This may be little more than a terminological variation on traditional hedonism, but as is often the case, terminological innovation can contribute to enlightenment.

By 1950 Freudian theory was the dominating influence in psychology. Neither the technique of psychoanalysis nor the supporting theory has received scientific validation, but no theory of human motivation and no form of psychotherapy can ignore the theories and practice of Freud. Freud himself protested that psychoanalysis does not attempt to explain everything, but in the human and social sciences there is hardly a question to which Freudian theory is quite irrelevant. The theory of the unconscious has been advanced and the techniques of analysis developed by such distinguished disciples as his daughter Anna Freud, Melanie Klein (a specialist in the analysis of children), and many others in Europe and the United States. Theory and techniques have also been developed by many disciples and eclectics. Two of Freud's disciples who deviated from his theories—Alfred Adler and Carl Jung—have had very considerable influence.

Alfred Adler (1870–1937) distinguished his system from psychoanalysis by labeling it individual psychology. Before meeting Freud, he had made a special study of the biological phenomena of compensation for defective bodily organs. After his association with Freud he extended his principles to account for all forms of compensation for inferiority, the "inferiority complex." In deviating from Freud, he assigned less importance to unconscious motivation and to sexuality.

Carl Gustav Jung (1875–1961) labeled his system analytical psychology. He differed from Freud in assigning a less important place to sexual motives and in his account of the unconscious. Jung regarded the libido as an undifferentiated "life force" which became differentiated into a number of instincts or drives. In his long life Jung developed a number of important but highly controversial theories. He elaborated the controversial and obscure concept of the collective unconscious and a theory of archetypal ideas (which has been confused by some with the Platonic concept of archetypes). Less controversial were the results of his experimental studies of word association and his suggestions regarding personality types. His wide-ranging speculations covered alchemy,

mythology, and the psychology of religion. Students of religion have found in Jung much of what they found absent or uncongenial in the writings of Freud. The opposition between Freudian and Jungian psychology has provided a modern parallel to the classical distinction between the Aristotelians and the Platonists.

Hormic psychology. In the Wundtian system as interpreted by Titchener the elements of mind were sensations and feelings. Conative experience had been eliminated by reductive analysis. Similarly, the concept of conative behavior had no place in Watsonian behaviorism. The concept of conation was not prominent in early Gestalt theory. Before 1950, however, the concept of conative or goal-directed behavior had been restored as a key concept in most systems of psychology. Tolman, the most sophisticated of the self-proclaimed behaviorists, established a new purposive behaviorism, and Lewin steered Gestalt psychology into the study of volitional processes. Throughout, Freudian theory is permeated by the facts of goal-directedness. The most thoroughgoing exponent of a conative psychology was William McDougall (1871–1938).

McDougall had a medical education but devoted himself to research in physiology, making several significant discoveries. An important early publication was his brief *Physiological Psychology* (1905), which contains the germs of his later theories. His most important publication was his *Introduction to Social Psychology* (1908). This title was unfortunate since the book contains the essentials of his general theory of motivation. Central to this theory was the thesis that there is a limited number of prime movers by whose conative force every train of thought and every bodily activity is initiated and sustained. These prime movers were first described as instincts, but the objections that were raised to his extreme deviation from the traditional biological conception of an instinct led McDougall to redescribe them as propensities.

In his detailed elaboration of these "propensities" McDougall developed an account of instinctive behavior originally suggested by William James. Prior to James instinct had been regarded as a biological mechanism producing rigid and stereotyped forms of behavior that were neither learned nor modified by experience. James drew attention to the cognitive emotional and impulsive components in instinctive action. McDougall developed this idea within the framework of the tripartite analysis of conscious experience that he had learned from Stout. Stressing the extent to which instinctive dispositions are modified both on the cognitive (receptive) side and on the conative (responsive) side, he suggested that the primary instincts are to be defined by reference to the central or affective components, the "primary emotions." He went on to describe the ways in which instinctive dispositions are modified and the ways in which they are organized into more complex motivating dispositions, the sentiments. A sentiment was conceived of as a system of instinctive disposition organized around an idea. Patriotism, for example, is a complex organization of instincts directed to promoting the welfare of a national group. McDougall's account of the structure of human personality was similar to that first set out in the famous sermons of Bishop Butler on human nature (1726). With McDougall, as with Butler, the motivating forces in man are organized in a three-tiered hierarchy. At the base are the primary instincts or propensities. At the second level in Butler's system were certain regulating and controlling principles, such as benevolence and cool self-love, and at the summit was the ultimate controlling principle, which was identified with conscience.

In McDougall's system the basic instincts are organized into and controlled by the sentiments, which function in a similar way to Butler's principles of benevolence and cool self-love. Thus, the parental sentiment is an organization of the maternal instinct together with other instincts, and in McDougall's view it explains all disinterested altruism. The self-regarding sentiment is an organization of the instincts of self-assertion together with others that exercise a similar control over primitive aggressive instincts. It functions in McDougall's theory in a way similar to Butler's cool self-love and Freud's reality principle. At the head of the hierarchy in McDougall's system as the supreme controlling force is a master sentiment that is an elaborated form of the sentiment of self-regard.

Both Butler's and McDougall's accounts of the structure of human personality, of human motivation, and of the basis of volition or self-control have important similarities with, but also important differences from, Freud's hierarchy of id, ego, and superego. Butler's analysis had greater philosophical subtlety than McDougall's, but McDougall's was developed in much greater detail. The central theses were contained in *Social Psychology.* The details were further elaborated in his later works, such as the *Outline of Psychology* (1923) and the *Outline of Abnormal Psychology* (1926). McDougall was himself surprised, as well as gratified, by the outstanding success of his *Introduction to Social Psychology.* He was to be surprised and disappointed by the reception of what he intended to be his magnum opus, *Body and Mind: A History and*

Defense of Animism (1911). This contained a critical review of the traditional theories of the relations of body and mind in which he eventually decided in favor of interactionism. His general philosophy of nature, life, and mind was that of an orthodox dualist and interactionist. This was later developed into a Leibnizian monadology. The personality of man was conceived as a hierarchy of monads. Every monad is potentially a thinking, striving self, but each differs in degrees of development. At the head of the hierarchy is the supreme monad—the self, which is in command of, and directly or indirectly in communication with, all subordinate monads. The mode of communication was conceived to be telepathic.

McDougall was one of the last of the academic psychologists to attempt a comprehensive system covering all the facts of cognition, feeling, and conation as well as the facts of unconscious motivation. His theories, however, fell out of favor, though not entirely because of specific objections to them. They were outmoded by current trends in both psychology and philosophy. Nevertheless, he exercised a considerable influence on thought and research in motivation theory, not least upon those who differed from him, and he contributed to the reunification of psychology and the biological sciences, which had been separated since Aristotle's day. Indeed, it could be argued that McDougall, like Aristotle, saw that the concept of purpose was both logically irreducible to mechanistic concepts and fundamental for the explanation of human behavior. His mistake was to translate this eminently defensible conceptual doctrine into a genetic doctrine about the origins of behavior. The two do not necessarily go together, for the doctrine that human behavior cannot be explained without recourse to a concept like purpose does not entail the genetic doctrine that men must come into the world equipped with a myriad of built-in purposes.

REACTION AGAINST REACTIONS. The proliferation of schools continued into the 1930s. Carl Murchison's *Psychologies of 1925* was followed by his *Psychologies of 1930,* and at the time no end to such quinquennial volumes could be foreseen. Psychologists, however, began to tire of these battles among the schools, each of which was in revolt against the established order and at war with the others in revolt. There came a revolt against revolt, a reaction against reactions. Robert S. Woodworth (1869–1962), who had written the most influential critical commentary on the schools, *Contemporary Schools of Psychology* (1931), was a leading advocate of a middle-of-the-road psychology. Teaching and practicing psychologists tended to be eclectic; many leaned heavily on one or another of the schools, and only a few remained uncommitted.

Schools were then replaced by "approaches," a term that suggests convergence rather than divergence. Approaches, like viewpoints, are complementary. The new situation favored the emergence of groups of psychologists united in discipleship to a single dominating personality. These groups differed from the schools in that a school was created by several outstanding personalities who, though agreeing on certain basic theses, made individual contributions to the system of psychology defended by the school. There have always been groups of the simpler leader-and-disciples type. Before the age of the schools there were philosophical psychologists with their disciples—for example, Brentano and Alexius Meinong on the Continent, Ward and Stout in Great Britain, James in the United States. In the schools themselves there were subgroups composed of a man and his disciples—the Freudians, the Jungians, the Pavlovians, and so on. After the dissolution of the schools new personalities emerged, each with an individual approach or field of specialization; there were psychologists like Piaget at Geneva, Albert Michotte at Louvain, and Tolman and many others in the United States.

RELATION TO PHILOSOPHY. The history of psychology in the twentieth century is a story of the divorce and remarriage of psychology and philosophy. The trouble began when psychologists claimed the status of empirical scientists. At first the philosophers were the more aggressive, deriding the young science as a bogus discipline. The psychologists hit back and made contemptuous remarks about philosophical logic-chopping and armchair psychology. The arguments were charged with emotion, and neither side emerged with great credit. Slowly, some progress was made toward a diagnosis of the situation, a diagnosis that may well provide the basis for a happy reconciliation.

Psychology has always been, and may well always remain, a parasitic discipline. For twenty centuries it was just a branch of philosophy. To gain emancipation, it entered into willing bondage to the established natural sciences. Increasingly it has claimed to be, and has been increasingly accepted as, a biological science. Aristotle's psychology had a biological orientation, and theories of the temperaments have always had a physiological slant. Since Darwin psychologists have attempted to work down to the biological foundations of mental life, and biologists have extended their field of interest upward to include the more complex functions of organisms tradi-

tionally described as mental—perception, learning, problem solving. In the twentieth century psychologists and biologists found a common approach, frame of reference, and interest in such new special studies as ethology, cybernetics, and information theory and a common lack of interest or only a peripheral interest in problems of the philosophy of mind. There have, however, been other developments that have helped to resolve the conflicts between philosophers and psychologists and to clarify the lines of demarcation between work that can properly be done in an armchair and work that must be done on a laboratory stool, in a birdwatcher's blind, or behind a one-way screen.

The behaviorists, in their revolt against Titchener's introspectionism, had taken over quite uncritically Titchener's greatest error. Hegel had attempted to answer questions of empirical fact by a priori reasoning. Titchener made the opposite mistake, supposing that questions of philosophical analysis could be settled by observations made in a laboratory. His mistake is on record; he recalled that in 1888, when first reading James Mill's *Analysis of the Human Mind,* the conviction flashed upon him, "You can test all this for yourself." He thought he could test it by introspection. The *Analysis* of James Mill was an exercise in philosophical analysis that can be carried out in a soft armchair, perhaps more efficiently there than on a hard laboratory stool. The behaviorists also fell victim to the same error in confusing introspection and philosophical analysis, in failing to see that questions of analysis arise not only in regard to introspective reports but also in regard to behavioral concepts—stimulus, response, and behavior itself.

However, behaviorists and other biologically minded psychologists were little disposed to either philosophical speculation or philosophical analysis. They were content, like most biologists, to think of the world, regardless of consistency, *both* in terms of commonsense realism and in terms of the billiard-ball atomism of nineteenth-century physics, thereby following the physicists whenever they revised their theories. Those who had some interest in philosophy followed the prevailing trend in philosophy to some form of phenomenalism.

Reduction of mental concepts. There had been three centuries of philosophical thinking devoted to the elimination of superfluous psychological concepts. At first a mind was thought of as an immaterial substance that, like a material substance, persists through changing states. As a rod of iron passes through states of being hard and soft or black, red, and white in accordance with changes of temperature, so a mind passes through states of joy, sorrow, and so on in accordance with the success and failure of its endeavors. Descartes had described all modes of consciousness as states of the soul, some of which appear to be states of external bodies, others of which appear to be states of the body in which the soul is embodied, and others that really are, as they appear to be, states of the soul itself. In his new way of ideas Locke redescribed experience in terms of the soul, self, or ego being presented with and attending to objects in the mind that chiefly represent things in the external world. Berkeley pointed out, cogently, that there is no way of comparing these representative ideas with the things they are supposed to represent. There were, he suggested, no reasons for, and there were reasons against, supposing that there are material things to be represented. Exit the material world. Then came Hume, who gave an important negative introspective report. He could not observe this soul, self, or ego to which presentations were said to be presented. Exit the soul.

For a long time attempts were made to defend what Titchener described as an act and content psychology—the doctrine that mind consists in mental contents and acts of willing and attending concerned with these contents (without, however, anyone to perform these acts). Late in the nineteenth century Brentano argued that these acts or attitudes are what is distinctive of mind. G. E. Moore based his refutation of idealism on this thesis by distinguishing in sensation the sensing, which alone is distinctively mental, from the sense datum sensed. But, like Hume, he made another negative introspective report—that the act is diaphanous, unintrospectable. Exit the act, the last claimant to mentality.

This reduction and elimination acquired a temporary finality in Bertrand Russell's neutral monism. Influenced by Moore, Ernst Mach, and William James, he proposed the overall theory that the stuff of which the universe is composed is neutral, not mental or physical. Organized in one way, it issues in the laws of physics; organized in another way, it results in the laws of psychology. Combining these, we have an account of nature. In this long reductive process man first had lost his soul, then his mind, then his consciousness, and finally even his body, which was reduced to a permanent possibility of neutralized sensations.

Linguistic approach. The finality of this form of phenomenalism was short-lived. The conception of philosophy as an inquiry into the ultimate nature of reality was supplanted by the idea that philosophy is the critical analysis of the concepts of science and of common sense. This was in turn replaced by the idea of philosophy as the

study of linguistic usages. Instead of asking what mind is, philosophers set out to disentangle the various uses of the word *mental,* and they became interested in the depth psychologists' uses of new words and of old words in new senses. Philosophers and psychologists began to find a new basis for collaboration. The philosophers clarified concepts; the psychologists attempted to verify by laboratory procedures the hypotheses stated in these concepts.

Not all issues between philosophers and psychologists have been resolved, but there has been notable progress toward a policy of coexistence, and here and there some progress toward cooperation has been made.

See also Animal Mind; Apperception; Behaviorism; Consciousness; Dreams; Emotion; Existential Psychoanalysis; Experience; Gestalt Theory; Guilt; Happiness; Humor; Images; Imagination; Intention; Intuition; Memory; Mind-Body Problem; Pain; Perception; Pleasure; Psychoanalytic Theories, Logical Status of; Religion, Psychological Explanations of; Sound; Thinking; Time, Consciousness of; Touch; Unconscious; Volition.

Bibliography

Boring, E. G. *A History of Experimental Psychology.* 2nd ed. New York: Appleton-Century-Crofts, 1950.

Boring, E. G. *Sensation and Perception in the History of Experimental Psychology.* New York: Appleton-Century, 1942.

Boring, E. G. et al. *History, Psychology, and Science: Selected Papers.* New York: Wiley, 1963.

Flügel, J. C. *A Hundred Years of Psychology, 1833–1963.* London: Macmillan, 1933. Enl. ed., New York: Basic, 1964. The enlarged edition has an additional part, by Donald J. West, on the period from 1933 to 1963.

Hamlyn, D. W. *Sensation and Perception.* London: Routledge and K. Paul, 1961.

Hearnshaw, L. S. *Short History of British Psychology, 1840–1940.* New York: Barnes and Noble, 1964.

Murphy, Gardner. *Historical Introduction to Modern Psychology.* Rev. ed. New York: Harcourt Brace, 1949.

Peters, R. S., ed. *Brett's History of Psychology.* Rev. ed. London and New York, 1962. Consult pp. 769–772 for a bibliography of all major texts in the history of psychology. This work is an abridgment and updating of G. S. Brett's three-volume *History of Psychology* (London: Allen, 1912) and is the only complete history of psychology.

R. S. Peters (1967)
C. A. Mace (1967)

PSYCHOLOGY [ADDENDUM]

In the 1950s and 1960s, scientific psychology underwent a major transformation. Behaviorist, Gestalt, and Freudian views were largely superseded by an approach called cognitive psychology, which treats the mind as a kind of information processor analogous to a computer. Cognitive psychology investigates the mental structures and processes that underlie perception, attention, learning, memory, language, inference, and problem solving. The field retains some behaviorist, Gestalt, and Freudian insights, but provides a coherent alternative that has been highly fruitful both experimentally and theoretically.

THE COGNITIVE REVOLUTION

The roots of cognitive psychology lie partly in the limitations of previous theoretical approaches to psychology, particularly behaviorism. Behaviorism attempted to make psychology scientific by avoiding reference to hypothetical mental entities such as thoughts and concepts. It tried to restrict psychology to the use of observed stimuli to predict observed behavioral responses. Behaviorism was fueled in part by a positivist philosophy of science that failed to recognize that explanation in natural science abounds with hypothetical entities such as atoms and genes. By the 1950s it was becoming apparent that stimulus-response relations were inadequate to account for human verbal behavior and even for learning in rats.

The emergence of an alternative explanatory framework came from several sources. One was information theory, developed by Claude Shannon in the 1940s, which inspired psychologists such as George Miller to try to characterize the capacities of the human mind to process information. Miller's 1956 paper, "The Magical Number Seven Plus or Minus Two," reviewed evidence that minds are inherently limited in their ability to hold only a small amount of information but argued that this limitation is surmounted by representations that chunk pieces of information together. Cognitive psychology has largely abandoned the information-theoretic division of information into discrete bits, but the metaphor of information processing remains pervasive.

A second and ultimately more important source of cognitive psychology was the development in the late 1940s and 1950s of the idea of a computer program. Before the advent of computers, philosophers and psychologists who wanted to give a mechanistic account of mind were limited to relatively simple mechanisms such as clockworks and telephone switchboards. Computer

programs consist of representational structures such as numbers, words, and lists, along with algorithmic procedures that transform the structures to produce new ones. In their 1960 book *Plans and the Structure of Behavior*, George Miller, Eugene Galanter, and Karl Pribram compared the plans that control behavioral operations to computer programs. By 1956, Allan Newell and Herbert Simon had already developed a computer program that could simulate human performance on a high-level task, proving logic theorems. Most theories in cognitive psychology operate with the analogy that human thought applies mechanical processes to mental representations, just as computation applies algorithms to defined structures. In the strongest view, thinking is not just *like* computation, it *is* a kind of computation.

The third conceptual source for cognitive psychology was Noam Chomsky's new approach to linguistics, developed in the 1950s as an alternative to the behaviorist approaches of Zelig Harris that then dominated the field. Chomsky incisively criticized the explanatory adequacy of behaviorist accounts of language learning. He proposed an alternative that postulated mental structures such as an innate universal grammar that makes possible the efficient learning of any human language.

Ideas about information, computation, and mental grammars redirected the experimental research that occupies most psychologists much more than do theoretical matters. Investigation shifted from studies of animal behavior to experiments with human subjects concerning such mental operations as visual pattern recognition, memory, verbal learning, and speech perception. In 1967 Ulric Neisser published the new enterprise's first textbook, *Cognitive Psychology*, and the journal of the same title began three years later. Neisser focused on processes for visual and auditory cognition.

TOPICS IN COGNITIVE PSYCHOLOGY

Later textbooks have addressed a broader range of topics, especially learning, memory, attention, perception, problem solving, language, representation, decision making, and deductive and inductive inference. Many experimental results have accumulated concerning these cognitive processes.

Research on perception has investigated how people recognize objects and other structures such as faces. Perceptual recognition involves both bottom-up processing from physical stimuli registered on sensory receptors such as the retina, and top-down processing influenced by high-level beliefs and concepts. Visual imagery has been a lively area of research, with Stephen Kosslyn and others arguing that evidence supports the view that minds operate with visual as well as verbal representations. The study of attention considers the factors that lead people to focus on and shift their concentration to different aspects of their environment.

Memory researchers distinguish between long-term memory, which permanently stores representations of events and concepts, and working memory, which holds and manipulates information as people perform cognitive tasks. Human memory is very different from computer memory, which stores information exactly as presented to it. Consolidation of events and facts into long-term memory involves reconstruction and blending with previous experience. People are conscious of only part of the contents of working memory and are totally unaware of most of the cognitive processing that constitutes thought. Almost all thinking is unconscious, not because of Freudian repression mechanisms, but because people have little access to most of the operations of their brains.

Cognitive psychologists distinguish between episodic memory for particular events and semantic memory for conceptual relationships. Debate has raged concerning what concepts are, although most psychologists reject the traditional view, still found in philosophy, that concepts can be defined by necessary and sufficient conditions. Alternative theories of concepts maintain variously that they consist of prototypes that specify typical but not universal features, sets of exemplars of objects, general knowledge about things, or patterns of activation in neural networks. These theories have inspired experimental investigations of how concepts are learned.

Cognitive psychologists who study language perform experiments concerning how people comprehend and produce utterances. In the 1960s and 1970s, research in psycholinguistics was heavily influenced by Chomsky's theory of transformational grammar, but later research shifted away from emphasis on the syntactic structure of language to concern with its meaning and communicative functions. There has also been investigation of high-level linguistic processes such as reading and understanding discourse.

Like the Gestalt psychologists, cognitive psychologists have been interested in problem solving and creativity. They have constructed detailed computational models of how people solve different kinds of problems using general rules and/or analogies with previous problem solutions. Theories of expertise have been developed based on different accounts of how different mental

structures constitute the knowledge possessed by people with substantial experience in a particular domain.

Psychological investigations of inference have been at odds with normative models of reasoning popular in philosophy. In contrast to the view that deductive inference involves the application of formal syntactic rules such as those found in propositional logic and predicate calculus, much psychological research has supported the view that the human mind makes inferences in a way that does not sharply distinguish syntax from semantics. People make deductive inferences based on the content of representations, not just their form. Similarly, people do not make decisions using the formalism of expected utility theory, which sharply distinguishes utilities from probabilities. Instead, they use a variety of cognitive and emotional heuristics to evaluate and choose different options. It is controversial whether the deviation of people from normative models of inference shows that they are irrational, or whether there is a need for more psychologically realistic models of rationality.

THEORETICAL AND EXPERIMENTAL DEVELOPMENTS

In the 1970s, cognitive psychology became part of the interdisciplinary field of cognitive science, which also includes philosophy, linguistics, anthropology, neuroscience, and the branch of computer science called artificial intelligence. Computational ideas continued to be at the core of psychological theory. Some theorists such as Allen Newell and Herbert Simon, and later John Anderson, maintained that psychological phenomena are best explained by postulating that the mind uses rules of the form, IF such and such, THEN such and such. These rules are operated by procedures that search through a large number of possible sequences of operations to provide solutions to problems. Rule-based models have been used to provide detailed explanations of problem solving, skill acquisition, and language production and acquisition.

Other psychologists have emphasized the role that structures such as concepts and schemas play in cognition. From this perspective, problem solving is not so much a sequential search through a large range of possible moves, as the application of patterns that enable people to comprehend and respond to situations. Thought is viewed as a kind of pattern matching rather than as a search through a space of operators defined by rules. Computational models of visual processing have inspired new theories and experiments concerning the nature of visual imagery.

During the 1980s, there was an influx of theoretical ideas based on computational models of artificial neural networks. This approach is called *connectionism* because it emphasizes the connections (links) between simple neuron-like processors. Inferences of the sort performed by rules and concepts are supposed to emerge from the interactions of many highly connected units that take in activation from many other units and pass it on to many others. Learning consists in adjusting the strengths of the links between the units. Connectionist models have been applied to many psychological phenomena, including learning from examples and high-level problem solving of the sort discussed by Gestalt psychologists.

From psychology journals such as *Cognitive Psychology*, it is evident that most psychological research is experimental rather than theoretical. Experimenters use a variety of measures such as error rates and reaction times to detect characteristics of human thinking. For example, investigations of eyewitness memory and testimony show that it is sometimes inaccurate, with errors being more likely if the witness has been distracted, if the misinformation is plausible, if there is social pressure, and if eyewitnesses have been given surreptitious positive feedback. Studies of deductive and inductive inference look at the difference between examples where people reason well and examples where they tend to make mistakes. Another experimental measure is reaction time, which compares how fast people are to respond to different stimuli. For example, people are quicker to respond to an item if is preceded by a similar item that primes it.

In the 1990s, cognitive psychology began to draw much closer to neuroscience. Cognitive psychologists have always assumed that mental operations are carried out by the brain, but for decades they lacked experimental techniques to relate human behavior to brain processes. Instruments are now available for imaging what is happening in the brain while people perform cognitive tasks. The most commonly used are PET (positron emission tomography) and fMRI (functional magnetic resonance imaging) scans. Both tools can detect the increase in blood flow to regions of the brain that become active when it is performing a particular task. For example, brain scans can determine what parts of the brain have increased activity when people are asked to rotate mental images. Many cognitive psychologists have turned to performing experiments in which people perform mental tasks while their brains are being scanned.

These experiments have furnished data that are used to suggest and evaluate theories about what brain processes are involved in various cognitive tasks. Increas-

ingly, cognitive theorizing is more brain-oriented, invoking processes that occur within particular brain regions and involve the interactions of multiple brain regions. In contrast to the connectionist models of the 1980s, computational models based on cognitive neuroscience employ much more neurologically realistic ideas about the structure of brain networks and the operation of individual neurons. Whereas cognitive psychology originally ignored the role of motivation and emotion in human thinking, cognitive neuroscience has inspired new models that integrate cognition and emotion in accounts of human decision making. Even the topic of consciousness, assailed by the behaviorists as inherently unscientific, is now being investigated by means of psychological and neurological experiments along with neurocomputational theories.

Cognitive psychologists disagree about the extent to which mental structures and processes are innate. Psycholinguists and evolutionary psychologists such as Stephen Pinker argue that natural selection has furnished the mind with many special purposed inference mechanisms such as a language acquisition device. The alternative view is not that the mind is a blank slate with no innate machinery, but rather that what all humans inherit is a highly flexible learning mechanism that makes possible adaptation to many situations. Although proponents of innateness are more likely to advocate rule-based rather than connectionist theories, there is no inconsistency in maintaining that most rules are learned and that some connections are innate.

OTHER AREAS OF PSYCHOLOGICAL RESEARCH

Cognitive psychology is only one area of psychological research, but it has had a major impact on other areas such as developmental, social, organizational, educational, and clinical psychology. Developmental psychologists study the origins and growth of children's knowledge of language and the world. The ideas of Jean Piaget dominated developmental psychology for decades, but they have been reassessed and revised by means of theories and experiments suggested by cognitive psychology. Developmental psychologists have also constructed new experimental techniques, such as the measurement of infants' attention times used to indicate what kinds of objects and situations are unfamiliar to them.

Social psychology, which concerns how people interact with each other, has become dominated by the field of social cognition, which examines people's cognitive processes. Social cognition investigates how people make sense of each other using concepts, stereotypes, rules, hypotheses, memories, emotions, personality traits, and other mental representations. Social psychology also looks at how cognition can vary across different cultures. Cognitive psychology has had an equally major impact on organizational psychology, which studies such topics as management and industrial development. In particular, work on organizations has been influenced by cognitive theories of decision making and learning. Similarly, cognitive theories of learning such as ones concerned with the acquisition of rules and concepts have had a substantial impact on educational psychology.

Clinical psychology, another area with practical applications, has also been transformed by the cognitive revolution. This area was once dominated by Freudian theories, which generated little success in either experimental or clinical settings. Cognitive therapy is one of the few kinds of psychotherapy that have been shown in controlled experiments to benefit people with emotional disorders such as depression. Unlike psychoanalysis, it does not require detailed discussion of a patient's childhood, but instead concentrates on helping the patient to replace unrealistic beliefs and goals in order to produce more positive appraisals of themselves and their situations. The medical field of psychiatry has also abandoned Freudian theories in favor of neurochemical explanations and treatments of mental illnesses such as depression, mania, dementia, and schizophrenia.

CONTROVERSIES

Although psychology has accumulated an impressive body of experimental findings and theoretical explanations, there remain topics of controversy that indicate directions of future work and generate interesting philosophical issues. These topics include: nature versus nurture, culture versus universality, rules versus connections, images versus propositions, mental logic versus mental models, heuristics and biases versus the adaptive toolbox, and embodiment versus computation.

NATURE VERSUS NURTURE. Psychologists disagree about the extent to which the behavior of humans is the result of innate, genetically transmitted neural structures (nature), compared with the extent to which it is the result of learning from physical and social environments (nurture). The nature side is currently emphasized by evolutionary psychologists such as Stephen Pinker and Leda Cosmides, who argue that many specific mental abilities have developed as the result of evolution by natural selection. They have proposed that the brain con-

tains evolved modules that are specialized in function, such as a face recognition system, a language acquisition device, navigation specializations, and a routine for detecting cheaters in social situations. The alternative view is not that there are no genetically inherited structures in the brain, but rather that the main innate endowment is a flexible learning ability that enables people to adapt to a wide range of environments. Proponents of this view include Jeffrey Elman, Steven Quartz, and Terrence Sejnowski. Psychologists also debate the extent to which general intelligence is inherited, and even whether there is such a thing in contrast to specific kinds of intelligence such as verbal, social, and physical abilities. Resolution of these debates will require further research on the structure and development of the brain.

UNIVERSALITY VERSUS CULTURE. Psychologists have tended to assume that the mental processes they investigate are universal, operating in all human minds. In contrast, anthropologists have tended to emphasize the diversity of different cultures with respect to beliefs and practices. An increasing number of psychologists have been using experimental methods to investigate the impact of culture on cognition, motivation, and emotion. Richard Nisbett and others have explored the impact of cultural differences on aggressive behavior and even on general styles of thinking. Compared to Westerners, East Asians are more likely to notice environments and relations rather than objects, to see change rather than stability, and to explain other people's behavior in terms of situations and relationships rather than personality traits. This issue is not the same as the nature versus nurture issue because there is no evidence of any relevant biological differences between Westerners and East Asians; cognitive differences therefore reflect culture rather than genetics. Further cross-cultural work in cognitive and social psychology will provide more information on the extent to which cognitive processes are universal.

RULES VERSUS CONNECTIONS. Many psychologists follow Chomsky in supposing that the acquisition and use of language depends on the possession of rules. For example, the standard way to form the past tense in English is to add "ed," as in "Sheila argued with Tom." But there are also many exceptions to this rule, such as "threw" and "went." James McClelland and other connectionists have argued that language use does not require rule acquisition and can be understood as the result of learning mechanisms that modify the links in neural networks. In contrast, Stephen Pinker and others argue that rules are necessary to explain patterns of linguistic behavior.

IMAGES VERSUS PROPOSITIONS. Philosophers such as Jerry Fodor have interpreted the computational view of mind to imply the existence of a language of thought that uses a representational scheme akin to verbal propositions. In contrast, most psychologists maintain that human thinking involves more than one kind of code, in particular visual images in addition to verbal propositions. Stephen Kosslyn and other have argued that a combination of psychological and neurological evidence, along with computer simulations, support the hypothesis that minds use visual as well as verbal representations. For example, brain scanning experiments show that when people perform imaging tasks, they use parts of the brain involved in visual processing. However, Zenon Pylyshyn continues to maintain that the evidence does not support the hypothesis that visual imagery is computationally different from verbal inference.

MENTAL LOGIC VERSUS MENTAL MODELS. How do people perform inferences such as the following? All humans think; anything that thinks has a brain; so all humans have brains. Philosophers since Aristotle have used formal logic to identify valid deductive inferences that accord with rules of inference such as *modus ponens*. Some psychologists such as Lance Rips similarly argue that human inference uses a kind of mental logic based on abstract rules. In contrast, Philip Johnson-Laird and others have presented experimental evidence that human inference does not distinguish form and content, but rather works with concrete instantiations that he calls *mental models*. The mental model approach has been applied to syllogistic, propositional, probabilistic, and causal inferences.

HEURISTICS AND BIASES VERSUS THE ADAPTIVE TOOLBOX. Many philosophers since Aristotle have assumed that humans are inherently rational, but many psychologists have investigated common tendencies to make inferential errors in inductive reasoning. Daniel Kahneman and Amos Tversky launched a fertile research program that showed that people often have difficulty making inferences that accord with normative models based on mathematical theories of probability and utility. They proposed instead that people operate with simple mental heuristics that bias them into making inferential errors. For example, people might think there are more words that start with "R" than end with "R," because it is easier to think of examples of the former rather the latter. Gerd Gigerenzer agrees with Kahneman and Tversky that human rationality is bounded by cognitive limitations but argues that biological evolution has provided people

with heuristics that are fast, frugal, and quite effective. For example, he thinks that people have heuristics for reasoning about frequencies that can enable them to avoid many of the reasoning errors identified by Kahneman and Tversky.

EMBODIMENT VERSUS COMPUTATION. Since the cognitive revolution in the 1960s, most psychologists have adopted the view of mind as an information processor analogous to a computer. An alternative view, based on the work of James Gibson and others, is that much of human behavior can be understood in terms of responses by minds to properties of their physical and social environments. On this view, human thought depends heavily on the kinds of bodies that people have, with sensory and kinesthetic abilities that shape our perceptions and thoughts. Extreme versions of this view deny that computation is part of human thinking at all, but more moderate views emphasize special kinds of computations performed by brains that store information using visual and other formats that are tied to the sensory apparatus of human bodies. Then the embodied nature of much of human thinking is a useful supplement to the computational theory of mind, not an alternative to it.

RELATIONS TO PHILOSOPHY

Developments in cognitive psychology have been important to philosophy for two reasons. First, philosophy of mind has responded to changes in the nature of psychological theories and explanations, with implications for metaphysical issues concerning the relation of mind and body. Second, a major strain of epistemology has become naturalistic, viewing theories of knowledge as continuous with cognitive psychology. Philosophical naturalism has also had an impact on ethics.

PHILOSOPHY OF MIND AND PSYCHOLOGY. The common-sense theory of mind is dualism, according to which people consist of two substances, matter and soul. In contrast, materialism maintains that humans, like the rest of the universe, consist only of matter and energy. Materialists from Epicurus to Lucretius to Thomas Hobbes have faced the daunting project of explaining how mental phenomena such as thinking and consciousness arise from brain activity. The development of cognitive psychology, with its view of thinking as computation performed by the brain, has greatly contributed to philosophical theories of mind.

In the 1960s and 1970s, philosophers such as Hilary Putnam, Jerry Fodor, and Daniel Dennett discussed the implications of the new ideas about computation and cognition. In the 1950s, materialist philosophers such as J. J. Smart had advocated an *identity* theory, according to which mental processes are brain processes. But the advent of artificial intelligence and computational theories of mind suggested an alternative view that mental processes are independent of any particular physical realization. This view is called *functionalism* because it understands mental states as functional states that are related to each other computationally. Functionalism is still a version of materialism, because it assumes that mental states have a basis in brains, computers, or some other form of matter and energy. But it differs from the identity theory in not equating mental states with any particular kind of physical state.

Functionalism gained much plausibility from the rise of computational views of mind, but it has been challenged by developments in cognitive neuroscience. Cognitive psychology has moved away from abstract computational theories toward theories embedded in particular accounts of the structure and processes in the human brain. It has therefore become less plausible that mental states are functional states rather than specific brain states. Moreover, progress in the field of artificial intelligence has not been as great as its originators had hoped. There have been some impressive industrial applications, but the prospect of a general-purpose machine intelligence comparable to humans remains distant. In contrast, understanding of how the brain uses the biochemical properties of neurons organized into connected functional areas has expanded rapidly. Hence developments in scientific psychology have lent support to identity over functionalist theories of mind. Some philosophers such as David Chalmers have argued that problems in understanding consciousness require a form of dualism, but many psychologists and neuroscientists remain optimistic that even consciousness will yield to scientific explanation.

What is the nature of the theories, explanations, and experimental results that cognitive psychologists offer? Philosophy of science has often adopted the model, derived from physics, that a theory consists of universal laws that deductively explain universal generalizations from observation. But the results of psychological experiments are usually statistically significant effects or tendencies, not laws. Moreover, psychological theories are rarely stated as universal laws, and the relationship between theories and what they explain is rarely deductive. Instead, a theory in cognitive psychology is a description of a computational or neurological mechanism,

where a mechanism is a system of parts whose properties and relations produce regular changes in other parts. For example, a bicycle is a mechanism consisting of parts such as pedals, wheels, a crank, and a chain that interact with each other. In purely computational theories such as those based on rules, the parts are mental representations, and the changes are brought about by computations on the representations. In neurological theories, the parts are neurons organized into brain areas and the changes are brought about by biochemical processes. Ideally, theories can be both neurological and computational when they show how groups of neurons implement mental representations and how biochemical processes implement computational operations.

PHILOSOPHICAL NATURALISM. Naturalism is the view that philosophical problems are continuous with those in science and amenable to treatment by the same kinds of methods used by scientists, as opposed to a priori theorizing about necessary truths and nonnatural entities such as gods and souls. In the late twentieth century, naturalized epistemology was revived through the work of philosophers such as Willard Van Orman Quine and Alvin Goldman. Quine's naturalism was limited by his adherence to behaviorist psychology, but later work has made full use of the expanding resources of cognitive psychology. Goldman has shown how to link philosophical questions about the origins and justification of knowledge with psychological research on perception, memory, and inference. This kind of naturalism does not use psychology to replace or reduce philosophy, which remains concerned with normative issues about justification that are not studied by psychologists. But naturalism applies epistemic appraisal to psychological processes that operate in human brains. Epistemic naturalism has also influenced the philosophy of science, with the view that scientific theories are mental structures rather than logical entities.

Moral naturalism has also been revived by philosophers who argue that ethics needs to pay close attention to cognitive psychology and neuroscience. According to moral naturalists such as Owen Flanagan, construction of moral theories and projection of moral ideals needs to ensure that the character, decision making, and behavior prescribed are possible for human beings. Understanding and evaluation of moral judgments is improved by appreciating how they arise from cognitive and neural processes such as concept application and empathy, which requires integration of cognitive operations with emotional processes. As with naturalized epistemology, moral naturalism does not purport to reduce ethics to psychology, but rather to develop a richer account of moral justification consistent with rapidly increasing scientific knowledge about the cognitive and neurological sources of human action and judgment. Similarly, some philosophers have become involved in controversies concerning human rationality that arise from the debates about mental logic and heuristics and biases. According to the naturalistic perspective, epistemology, ethics, and metaphysics should continue to evolve hand in hand with further developments in cognitive psychology and neuroscience.

See also Anderson, John; Artificial Intelligence; Aristotle; Behaviorism; Chomsky, Noam; Cognitive Science; Computationalism; Dennett, Daniel Clement; Dualism in the Philosophy of Mind; Epicurus; Fodor, Jerry A.; Functionalism; Gestalt Theory; Goldman, Alvin; Hobbes, Thomas; Lucretius; Materialism; Memory; Naturalized Epistemology; Neuroscience; Piaget, Jean; Putnam, Hilary; Quine, Willard Van Orman; Smart, John Jamieson Carswell.

Bibliography

Anderson, John R. *Cognitive Psychology and Its Implications*. 5th ed. New York: Worth, 2000.

Anderson, John R., with Francis S. Bellezza, et al. *Rules of the Mind*. Hillsdale, NJ: Erlbaum, 1993.

Baars, Bernard. *The Cognitive Revolution in Psychology*. New York: Guilford, 1986.

Bechtel, William, and Adele A. Abrahamsen. *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks*. 2nd ed. Malden, MA: Blackwell, 2002.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press, 1996.

Gardner, Howard. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books, 1985.

Gigerenzer, Gerd. *Adaptive Thinking: Rationality in the Real World*. New York: Oxford University Press, 2000.

Gilovich, Thomas, Dale Griffin, and Daniel Kahneman, eds. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge, U.K.: Cambridge University Press, 2002.

Goldman, Alvin I. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Johnson-Laird, P. N. and Ruth M. Byrne. *Deduction*. Hillsdale, NJ: Erlbaum, 1991.

Kosslyn, Stephen M. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press, 1994.

Kunda, Ziva. *Social Cognition: Making Sense of People*. Cambridge, MA: MIT Press, 1999.

Matlin, Margaret W. *Cognition*. 6th ed. Hoboken, NJ: Wiley, 2005.

May, Larry, Marilyn Friedman, and Andy Clark, eds. *Mind and Morals: Essays on Cognitive Science and Ethics*. Cambridge, MA: MIT Press, 1996.

Miller, G. A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63 (1956): 81–97.
Miller, G. A., E. Galanter, and K. Pribram. *Plans and the Structure of Behavior*. New York: Holt, Rinehart, and Winston, 1960.
Murphy, Gregory L. *The Big Book of Concepts*. Cambridge, MA: MIT Press, 2002.
Neisser, U. *Cognitive Psychology*. New York: Appleton-Century-Crofts, 1967.
Nisbett, Richard E. *The Geography of Thought: How Asians and Westerners Think Differently … and Why*. New York: Free Press, 2003.
O'Reilly, Randall C., and Yuko Munakata. *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press, 2000.
Pinker, Steven. *The Blank Slate: The Modern Denial of Human Nature*. New York: Viking, 2002.
Quartz, Steven R., and Terrence J. Sejnowski. *Liars, Lovers, and Heroes: What the New Brain Science Reveals about How We Become Who We Are*. New York: Morrow, 2002.
Rips, Lance J. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: MIT Press, 1994.
Thagard, Paul. *Mind: Introduction to Cognitive Science*. 2nd ed. Cambridge, MA: MIT Press, 2005.
Wilson, Robert A., and Frank C. Keil, eds. *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press, 1999.

Paul Thagard (2005)

PSYCHOLOGY, GESTALT

See *Gestalt Theory*

PUFENDORF, SAMUEL VON

(1632–1694)

Samuel von Pufendorf, the German political and legal philosopher and historian, was born in Dorfchemnitz, in Meissen, Saxony, the son of a poor Lutheran pastor. A scholarship enabled Pufendorf to attend the famous Prince's School at Grimma. From 1650 to 1656 he attended lectures on Lutheran theology and Aristotelian philosophy at Leipzig. Somewhat later he studied contemporary philosophy at Jena, where he also read newly published books on mathematics and discovered the works of Hugo Grotius and Thomas Hobbes. At Jena he came in contact with Erhard Weigel, a former teacher of Gottfried Wilhelm Leibniz, whose strange but original method of teaching ethics "mathematically" made a lasting impression upon Pufendorf. To Weigel, Pufendorf owed the inspiration for his first work on the general principles of law, *Elementorum Jurisprudentiae Universalis*. In 1658 Pufendorf became a tutor in the house of the Swedish ambassador to Denmark. When war broke out between Sweden and Denmark, Pufendorf was imprisoned for eight months, and it was during this imprisonment that he composed the booklet inspired by Weigel. Upon his release Pufendorf migrated in 1659 to the Netherlands, where the work was published in 1660.

On the recommendation of Grotius's elder son, Pufendorf was offered the chair of natural and international law at Heidelberg, the first such chair at a German university. He was soon appointed also as instructor of the heir to the crown of the Palatinate, and thus he began to mix with the electoral court, where he avidly studied the burning contemporary political problems. Out of this study came a pseudonymous work on the condition of the Holy Roman Empire, *De Statu Imperii Germanici* (1667), a work later famous for its statement that the constitution of the Empire resembles a monster, being neither a monarchy nor an aristocracy nor a democracy.

After his appointment as professor of natural law at the University of Lund in Sweden, Pufendorf wrote his fundamental work on national and international law, *De Jure Naturae et Gentium* (1672). The eight volumes of this compendium, which contains a veritable encyclopedia of the social sciences, are rather difficult reading. Pufendorf therefore produced an abstract of this work, titled *De Officio Hominis et Civis* (1673), which was soon translated into English, French, and German and thus found many readers abroad. By 1684 a Swiss Calvinist theologian was lecturing on the *De Officio Hominis* at Lausanne, but Lutheran theologians in both Sweden and Germany criticized Pufendorf's ideas vehemently. The king of Sweden himself had to protect his professor of law and induce the authorities of the university to defend Pufendorf against the charge of heresy. Pufendorf replied bitterly to the charge, and a long paper war ensued. Finally, Pufendorf published a "sanguinary" (his own description) polemical treatise titled *Eris Scandica* (Frankfurt, 1686), containing all his essays and letters relating to the controversy.

In 1677 Pufendorf was appointed by the king as court historian in Stockholm, where he spent ten years working on his extensive, thirty-three-volume history of Sweden, a work of no importance today except as an example of careful work and precise reporting. His shorter *Einleitung zu der Historie der vornehmster Reiche und Staaten* (2 vols., Frankfurt, 1682–1685) is more highly esteemed.

From 1688 until his death shortly after having been knighted by his former sovereign, the king of Sweden, Pufendorf lived in Berlin, where he had been called as court historian by the elector of Brandenburg.

A noted representative of the Baroque era, Pufendorf was a man of great self-confidence and stolid self-reliance. He had unshakable faith in the power of scientific reason and wished to establish it in the fields of jurisprudence and politics. He believed in the certainty of mathematics and rejoiced in the reunion of philosophy and mathematics then taking place. Although he wished to treat the problematic questions of jurisprudence and politics "mathematically," he was a true empiricist who sought to introduce a "scientific" method into the study of history. He was therefore eager to undertake the thoroughly planned research into public archives that resulted in his history of Sweden.

Pufendorf thus united the two major trends of his age, Baconian empiricism and Cartesian logicism. One of the last polyhistors, he united in his work all the methods of historical, sociological, and juridical thinking. A political figure rather than simply a lawyer, Pufendorf profoundly criticized the constitution of the Holy Roman Empire and its political conception. He argued for the founding of a European federation of sovereign states. He did not defend national or regional absolutism, however popular they were at the time; instead, he tried to unite the Hobbesian doctrine that the state should be governed by the rule of law and based on natural law in the empirical sense of the term (the war of all against all, *status necessitatis*) with the Grotian doctrine that the rule of international law should be based on natural law in an emotional sense (an inclination for society, *ordo amoris*). On this account Pufendorf has often been called a predecessor of eighteenth-century rationalism. Such a view is supported by his letter to his younger friend Christian Thomasius, in which he claimed that he "never had boldness enough to draw the utmost conclusions" from his philosophical rationalism and voluntarism.

Despite Leibniz's opinion that Pufendorf was "a man of no great judgment," his legal thought was of considerable importance and great philosophical interest. He was undoubtedly one of the most outstanding social philosophers on the European continent in the seventeenth century. It may be an exaggeration to call Pufendorf the first "philosopher of culture" (*Kulturphilosoph*) in Germany, but he was the first to grasp the fundamental concept of the sociological theory of law and politics. He saw the social realities of human life as a whole. His structural distinction between physical facts and moral institutions inspired a new way of studying social facts in their independence and uniqueness. Following Weigel, Pufendorf distinguished four elements of social being: personality (*persona*), rank or profession (*status*), quality, and quantity. Every pattern of social order should be examined on the basis of these fundamental structures; for example, a state may be described in terms of its sovereignty, type of government, power, and population.

These elements, the ontological foundations of every community, have simultaneously to be interpreted as fundamental ethical principles of social life. Pufendorf designated three patterns of well-formed communities: humanity, ordered by the law of reason; Christianity, ordered by the law of God; and citizenship, ordered by the law of the state. Natural law, including religious and rational principles, therefore limits both civic and moral duties. Philosophy of law comprises both sociology and political science on the one hand, and jurisprudence and ethics on the other. This new discipline, which Pufendorf called simply natural law, was intended to unite all the tasks of interpreting social order and to combine the scholastic methods of the sixteenth-century Spanish thinkers with the newer ideas of Grotius and Hobbes.

In apparent contradiction to these sources of his thought on social order was Pufendorf's strong belief in reason of state (*ratio status*). Although he often emphasized the self-determination and self-sufficiency of the state, he did not mean by this a totalitarian absolutism. And although he proclaimed the independence of political power against every ecclesiastical claim, he never taught the modern ideology of unlimited government, and his views were therefore not contradictory to the rule of law. What Pufendorf said about the relation of church and state must be interpreted dialectically. He conceded neither decisive authority to reason of state nor the right of moral constraint to the church.

Pufendorf may be called the initiator of the seventeenth-century movement of "scientific" natural law in Germany. By introducing the ideas of Grotius and Hobbes into German thought he made their ideas really effective for the first time. He liberated the natural-law theory from the domination of scholasticism and humanism. In so doing he built up an independent political science that always took into account contemporary history and reason of state. A clever and levelheaded politician, he predicted the decline of the Hapsburg monarchy after the Treaty of Westphalia. In criticizing the "monstrous" constitution of the empire he sought to advance a European commonwealth based on the natural and rational principles of international law. As a histo-

rian, Pufendorf introduced the empirical study of archives and gave an effective example of a new method of historical insight, and he may be regarded as an important predecessor of nineteenth-century historicism.

See also Aristotelianism; Cartesianism; Empiricism; Grotius, Hugo; Historicism; Hobbes, Thomas; Humanism; Leibniz, Gottfried Wilhelm; Natural Law; Rationalism; Scientific Method; Sovereignty; Thomasius, Christian; Voluntarism.

Bibliography

WORKS BY PUFENDORF

Elementorum Jurisprudentiae Universalis Libri Duo. The Hague, 1660.

De Statu Imperii Germanici ad Laelium Fratrem, Dominum Trezolani, Liber Unus. Geneva, 1667.

De Jure Naturae et Gentium Libri Octo. Lund, Sweden, 1672. Translated by Basil Kennett as *Of the Law of Nature and Nations.* London, 1710.

De Officio Hominis et Civis Prout Ipsi Praescribuntur Lege Naturali. Libri Duo. Lund, Sweden, 1673. Reprinted with translation by F. G. Moore. 2 vols. Oxford, 1927. Translated by Andrew Tooke et al. as *The Whole Duty of Man according to the Law of Nature,* edited by Ian Hunter and David Saunders. Indianapolis: Liberty Fund, 2003.

De Habitu Religionis Christianae ad Vitam Civilem. Bremen, 1687. Translated as *Of the Relation between Church and State.* London, 1719.

Of the Nature and Qualification of Religion in Reference to Civil Society. Translated by J. Crull; edited by Simone Zurbuchen. Indianapolis: Liberty Fund, 2002.

WORKS ON PUFENDORF

Carr, Craig L., and Michael J. Seidler. "Pufendorf, Sociality and the Modern State." *History of Political Thought* 17 (3) (1996): 354–378.

Dufour, Alfred. "Pufendorf." In *The Cambridge History of Political Thought, 1450–1700,* edited by J. H. Burns and M. Goldie. Cambridge, U.K.: Cambridge University Press, 1991.

Hunter, Ian. *Rival Enlightenments: Civil and Metaphysical Philosophy in Early Modern Germany.* Cambridge, U.K.: Cambridge University Press, 2001.

Krieger, Leonard. *The Politics of Discretion: Pufendorf and the Acceptance of Natural Law.* Chicago: University of Chicago Press, 1965.

Scheuner, Ulrich. "Samuel Freiherr von Pufendorf." In *Die Grossen Deutschen,* Vol. 5, 126ff. Berlin, 1957.

Schneewind, J. B. "Pufendorf's Place in the History of Ethics." *Synthese* 72 (1987): 123–155.

Welzel, Hans. *Die Naturrechtslehre Samuel Pufendorfs.* Berlin, 1958.

Wolf, Erik. *Grosse Rechtsdenker der deutschen Geistesgeschichte,* 4th ed., 311–366. Tübingen, 1963. Bibliography on pp. 367–370.

Erik Wolf (1967)
Bibliography updated by Philip Reed (2005)

PUNISHMENT

The word *punishment* is used in varying contexts. The punishment meted out by the state to a criminal or by a parent to his children is not the same as the punishment boxers give or receive. The latter, however, is punishment only in a metaphorical sense, for it lacks several of the features necessary to a standard case of punishment. Characteristically, punishment is unpleasant. It is inflicted on an offender because of an offense he has committed; it is deliberately imposed, not just the natural consequence of a person's action (like a hangover), and the unpleasantness is essential to it, not an accidental accompaniment to some other treatment (like the pain of the dentist's drill). It is imposed by an agent authorized by the system of rules against which an offense has been committed; a lynching is not a standard case of punishment. Philosophers who have written on punishment have usually had in mind punishment in the standard sense rather than in any extended or metaphorical sense.

The philosopher's interest in punishment is mainly connected with questions of justification. It is, prima facie, wrong to deliberately inflict suffering or deprivation on another person, yet punishment consists in doing precisely this. What conditions, the philosopher asks, would justify it? Or, more generally, what kind of consideration would count toward a justification? For instance, if a person had already committed a crime, that would clearly be relevant to the question of whether he ought to be punished (although it might not be conclusive). What if he were only expected to commit a crime in the future? Or, again, is it relevant to the question of whether this man should be punished to say that punishing him would deter others? And assuming that criminals ought to be punished, how should we set about deciding appropriate penalties?

It is not, of course, the business of the moral or social philosopher to provide a justification for any particular act or system of punishment or even of the institution of punishment in general. Philosophers are not necessarily apologists for their society and age. They are interested in the procedures and modes of argument that we are committed to by our fundamental conceptions of morality and in criteria of criticism and justification rather than in inquiries into whether actual institutions satisfy them.

Philosophers, it is true, have not always made this distinction; they have often worked on the understanding that a philosophical argument could be seriously shaken by showing that it leads to conclusions inconsistent with some widely approved institution or moral rule. More-

over, for many philosophers, if such a rule or institution seemed to imply a principle inconsistent with other moral principles accepted by the society, there must necessarily be some broader principle, which a philosopher could discover and by which the conflict could be resolved. Applied to the case of punishment, this would mean that a philosopher must reconcile the apparently conflicting principles that wrongdoers should be punished and that it is wrong to deliberately make another person suffer. But this is surely a misconception of the nature of philosophy. There is no point, after all, in asking whether and how punishment can be justified if one assumed in advance that it can. For justification a number of contingent facts are required that the philosopher as such is not qualified to provide. His task is to analyze what is being asked for and so to point out what kinds of facts and arguments are admissible to the discussion.

JUSTIFICATION OF PUNISHMENT

The question of justification arises at two levels. One can take for granted the principle that wrongdoers should be punished and ask whether a particular case of punishment was justified. At this level the philosopher is concerned with the criteria in a general system which any particular act of punishment must satisfy. One can, however, question the very idea of punishment as an institution that involves deliberately inflicting pain or deprivation. This raises the philosophical question of how one justifies a set of rules or an institution like a penal system. Corresponding to these two levels of justification are two broadly opposed approaches to punishment, the retributivist and the utilitarian. Each, in fact, has been taken to offer an answer to the problems at both levels, but the persuasive force of retributivism is mainly in its answers to problems of the first type, and of utilitarianism to questions of the second type. Characteristically, the retributivist stresses guilt and desert, looking back to the crime to justify punishment and denying that the consequences of punishment, beneficial or otherwise, have any relevance to justification. The utilitarian, on the other hand, insists that punishment can be justified only if it has beneficent consequences that outweigh the intrinsic evil of inflicting suffering on human beings.

RETRIBUTIVIST THEORIES. The most thoroughgoing retributivists, exemplified by Immanuel Kant, maintain that the punishment of crime is right in itself, that it is fitting that the guilty should suffer, and that justice, or the moral order, requires the institution of punishment. This, however, is not to justify punishment but, rather, to deny that it needs any justification. To say that something is right or good in itself means that it does not need to be justified in terms of the value or rightness of anything else. Its intrinsic value is appreciated immediately or intuitively. But since at least some people do doubt that punishment is right, an appeal to intuition is necessarily unsatisfactory. Again, to say "it is fitting" or "justice demands" that the guilty should suffer is only to reaffirm that punishment is right, not to give grounds for thinking so.

Some retributivists, while admitting that punishment is, prima facie, evil, maintain that it is nevertheless better that the wicked should be punished than that they should prosper more than the virtuous and, perhaps, at their expense. In this view, the function of criminal law is to punish wickedness or immorality in order to maintain a kind of cosmic distributive justice. However, it is not self-evident that wickedness should be punished any more than it is self-evident that legal guilt should be. Archbishop Temple, himself a retributivist, declared that he had no "intuition that it is good that the wicked should suffer." Nor is it clear that virtue must be rewarded or that universal justice requires the kind of human rectification that this sort of retributivism envisages. Of course, in a universe in which the wicked prospered, there might be no incentive to virtue, but this is essentially a utilitarian mode of argument. Again, evil motives and a bad character are necessary conditions of wickedness but not of legal guilt and criminal liability. The state's function is to punish breaches of those rules which in the public interest ought to be upheld; it is a matter of indifference in law (but not in morals) that some men who observe the rules do so from the unworthy motive of fear and others break them from laudable motives of principle. Conversely, it is at least doubtful whether the criminal law should provide penalties for offenses against morality except where the public interest is at stake—for example, whether it should extend to cases of lying other than, say, false pretenses and perjury.

Though immorality is neither a necessary nor a sufficient condition for punishment, the relation between law and morals is nevertheless a close one, and what punishment is to the one, blame is to the other. Both regulate social intercourse, and in any given society the aims and ideals upheld by the law will usually correspond, more or less, with those upheld by the dominant morality. Moreover, in the family and the school punishment is often used to reinforce moral condemnation as part of the process of moral education. Some writers who regard punishment as moral retribution couple this idea with the argument that the point of punishment is to be found

in what Lord Justice Denning has called "the emphatic denunciation by the community of a crime." In this view, punishment reinforces the community's respect for its legal and moral standards, which criminal acts would tend to undermine if they were not solemnly denounced. There is, however, no intrinsic reason why denunciation should take precisely the form of inflicting suffering on criminals, unless, perhaps, one accepts Ewing's view that punishment has the advantage of impressing both on the criminal and on everyone else that a breach of law and morals is so serious that society must do something to prevent it. That, however, is surely to justify punishment by its utility in maintaining respect for the law. Hastings Rashdall refers to "the enormous importance of the criminal law in promoting the moral education of the public mind," but Rashdall was a utilitarian who justified punishment by reference to "the production of good effects on conscious beings."

For G. W. F. Hegel punishment is necessary to annul the wrong done by the criminal. By this he means something more than restitution or compensation, neither of which is, strictly speaking, punishment. It is, rather, that the criminal has upset the balance of the moral order, which can be restored only by his being made to suffer. Or, in terms of the dialectic, crime is a negation of right and as such a nullity; punishment negates the negation, thus reaffirming the right. But in what sense can punishment be said to restore the balance or annul the wrong, unless it is taken for granted that criminals deserve to be punished? This is precisely the point in question.

UTILITARIAN THEORIES. The utilitarian position is exemplified in Jeremy Bentham's remark that "all punishment is a mischief.… If it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil." By reforming the criminal, by deterring him or others from similar offenses in the future, or by directly preventing further offenses by imprisonment, deportation, or execution, the good that comes out of punishment may outweigh (so the utilitarian argues) the intrinsic evil of suffering deliberately inflicted. Without such effects, or if the suffering inflicted exceeded the suffering avoided, the institution would be unjustified.

The critics of utilitarianism claim that if people generally could be persuaded that an innocent man was guilty, utilitarianism would justify punishing him since as a warning to others he would be just as useful as a genuine offender. Again, offenders might be deterred by threatening to punish their wives and children, particularly, if as is so often the case with political terrorists and resistance fighters, it were difficult to catch the offenders themselves. Or, again, if punishment could be justified as a way of reforming criminals, it would seem better to punish them before, rather than after, they committed their crimes. Retributivists claim that utilitarians are in danger of losing sight of two conditions that are necessary to the very idea of punishment—namely, that an offense should have been committed and that punishment shall be of the offender himself, who alone can be said to deserve it. "Punishment is punishment," wrote F. H. Bradley, "only when it is deserved"; punishment for any other reason is "a crying injustice."

The dilemma of utilitarianism, then, at least in its crude form, is that it justifies punishing innocent people provided that such punishment causes less suffering than might otherwise be caused by the would-be criminals it deters. Some utilitarians argue that in the end the deception would break down, that it could not be used systematically, or that the long-term consequences would be bad for society. But these answers are unsatisfactory because they depend on assumptions of purely contingent consequences. Our revulsion against punishing innocent people seems to go deeper than that. In any case, these answers will not meet the case for punishing hostages, which can certainly be done systematically and requires no deception or secrecy.

PUNISHMENT AND PRINCIPLES OF JUSTICE. To meet the above criticisms, a crude utilitarianism would have to be supplemented by other moral principles—namely, that differences in treatment must be justified by relevant differences in circumstance or condition, where "relevance" is defined in the light of general rules, and that every human being should be treated with at least a minimum of respect as a source of claims and not as a mere instrument for the promotion of the interests of others. It can be argued that punishment of the innocent or of hostages is an abuse not because it necessarily makes for more unhappiness than it prevents but because it treats innocent men in a way that is appropriate only for the guilty and makes an arbitrary difference in treatment between them and other innocent men. Moreover, a legal system is designed to guide conduct by laying down rules and attaching penalties to those who choose to break them. It is acceptable, in the words of J. D. Mabbott, only because "the criminal makes the essential choice; he 'brings it on himself.'" Otherwise, punishment would not be consistent with the principle of respect for persons. The hostage, on the other hand, has no chance to settle his own fate; he is used as a mere lever for manipulating

other people's conduct, and his own interest is subordinate to that of the other members of society. Punishment of the innocent ignores, in short, fundamental procedural rules of justice and morality without which utilitarianism would make little sense, for unless everyone is worthy of equal consideration as a source of claims, whose interest is to count in assessing the utility of a course of action? Whom are we entitled to treat as simply a tool for advancing other men's interests—as Aristotle's "slave by nature"—and what would count as a reason for considering other men before him?

This has bearing, too, on the reasons for accepting as excuses such defenses as duress, unavoidable accident, or ignorance of fact—conditions under which an offender can claim that he could not help doing what he did. Bentham argued that to punish anyone under such conditions would be pointless and, therefore, mischievous, because the threat of penalties could not possibly deter anyone in the future who was similarly placed. Now, it is true that nothing would be lost if such people escaped punishment, provided they could be distinguished from cheats trying to take advantage of such excuses and provided enough offenders without such excuses could be detected to furnish examples for others. The principle of "strict liability," which exists in some legal systems for certain offenses, has been defended on the utilitarian ground that it is impossible to tell a genuine excuse from a pretense. It is questionable, however, whether a person who would otherwise be treated as innocent ought to be treated as guilty because someone else might otherwise escape a merited penalty. Punishing the man who commits an offense through ignorance or accident, because it is too difficult to tell whether he really did it on purpose or because we have to make an example of *someone,* is very like punishing the innocent as a warning to the guilty. The utilitarian case for these excuses is unsatisfactory inasmuch as it makes them subject to such qualifications.

A better ground for such excuses is that punishment is morally acceptable only if it is the consequence of an act freely chosen by the criminal, which it would not be under these conditions. A man acting in ignorance or by accident cannot be said to bring his punishment on himself. Punishment, seen as a way of influencing conduct, cannot be justified if there has been no real possibility of choice. Moreover, the punishment of involuntary offenses introduces into men's lives the possibility of disasters that they can neither foresee nor avert.

Utilitarianism, then, must be supplemented by principles of justice if it is not to clash with other moral principles that are usually considered fundamental. It has, however, the merit, as an approach to the justification of punishment, that it provides a clear procedure for determining whether the institution is acceptable in general terms. This the retributivist approach cannot do because it denies the relevance of weighing advantages and disadvantages, which is what we ultimately must do in moral criticism of rules and institutions. Consequently, a retributivist justification of punishment as an institution usually turns out to be a denial of the necessity for justification, a veiled reference to the beneficial results of punishment (a utilitarianism in disguise), or an appeal to religious authority.

When it is a question of justifying a particular case of punishment, however, the retributivist is in a far stronger position. There would be no point in having a general rule if on every occasion that it had to be applied one had to consider whether the advantages in this particular case warranted acting in accordance with it. Moreover, the point of punishment as deterrent would be quite lost were there no general expectation, based on the general operation of the rule, that the guilty would be punished. Assuming, then, that a penal system can be justified in utilitarian terms, any offense is at least prima facie an occasion for a penalty. Equally, without an offense there is no question of a penalty. The retributivist contention that punishment is justified if, and only if, it is deserved is really applicable, therefore, to the justification of particular instances of punishment, the institution as such being taken for granted.

SEVERITY OF PUNISHMENT

The clash between the utilitarian and retributivist approaches to punishment also arises in considering the criteria by which appropriate punishments are assessed. The retributivist insists that the punishment must fit the crime; the utilitarian relates the penalty to the general aims of the system, to the prevention of further crime, and, perhaps, to the reform of the criminal.

The most extreme form of retributivism is the law of retaliation: "an eye for an eye." This alone, Kant claimed, could provide a just measure of the penalty, since it was the crime itself and nothing else that settled it. However, to try to apply it literally might be monstrously cruel, or, as Kant recognized, it might be absurd. Thieves can be deprived of their property and murderers hanged, but what penalty is appropriate to the dope-peddler, the blackmailer, and the smuggler?

There is not much sense, either, in trying to construct a table of equivalents so that the amount of suffering

inflicted by the criminal could be meted out to him in some other form. How can such a table be drawn up? How many years must a blackmailer spend in jail to experience suffering equal to his victim's? Is it possible, in any case, to make comparisons of suffering between persons? Of course, we do assess the gravity of an offense and try to ensure that the punishment for a trivial offense is less severe than for a serious one. But this is possible only because we take for granted an existing scale of penalties and grade new offenses accordingly. Such grading does not imply an intrinsic relation between the crime and the penalty apart from that established by the scale. Some retributivists admit this but claim nevertheless that the penalties prescribed by the law ought to reflect the moral heinousness of the offense. The most serious offenses against morals deserve the most severe penalties. This, however, only shifts the question a step back, for what makes one moral offense more serious than another?

Utilitarians have tended to concentrate on deterrence, turning away from the actual criminal act except as one of a class of actions that might be prevented by punishing the particular instance severely enough (but only just enough) to make the action unattractive to the offender and to possible future offenders. Unfortunately, there are always people who cannot be deterred or reformed. Beyond a certain point the additional suffering one would have to inflict on all offenders to reduce their number might be so great as to exceed the amount of suffering thereby averted. The aim of the utilitarian, then, would presumably be to select the penalty at which the aggregate of suffering caused by crimes actually committed and punishments actually inflicted would be the smallest possible.

The utilitarian approach has often been criticized as justifying severe penalties for trivial offenses and vice versa. To eliminate parking offenses might need heavier penalties then to eliminate blackmail, which would be monstrous. But this criticism misses the point of the utilitarian case. There would, indeed, be no objection to threatening the severest penalty for any offense providing the threat never had to be carried out. Punishment is only an unfortunate consequence of the fact that the threats, which are the true operative elements in the system, are partially ineffective and would be wholly ineffective if they were not carried out when they failed to deter. In fixing penalties, the utilitarian's problem is not, therefore, to minimize the number of offenses, irrespective of the punishment inflicted, but to minimize the total amount of suffering from both sources. If we call parking offenses trivial, we mean that each one causes relatively little suffering; therefore, we are prepared to put up with a large number of them rather than incur the cost of making offenders suffer heavy penalties. Blackmail, on the other hand, causes so much suffering that if heavier penalties would yield even a small reduction in the number of offenses, there might be a net gain even though offenders would suffer more than they did before. In this way a utilitarian might agree with the retributivist that severe penalties ought to be restricted to serious offenses, but he would argue that we call an offense serious precisely because it causes a great deal of suffering. For the retributivist only serious crimes *deserve* severe penalties; for the utilitarian only serious crimes are worth averting at the cost of severe penalties.

The utilitarian approach to this matter does not supply a procedure for sentencing particular criminals (any more than a justification for punishment as an institution would be a case for any particular application of it). Arguing from expected consequences, one might establish a kind of standard penalty for each class of offense. Officials drafting new rules might consider whether a proposed maximum penalty would keep offenses down to manageable proportions, or people concerned about road accidents might argue that heavier penalties for motoring offenses would make drivers more careful. Deciding the sentence in a particular case, however, is clearly a different matter. The maximum penalty is a limiting factor, but questions like the degree of responsibility, provocation, and the offender's previous record are all relevant. However, one might reasonably ask why, as a matter of principle, they should be relevant.

PUNISHMENT AND RESPONSIBILITY

The problem of responsibility arises in relation to punishment as it does in relation to blame in moral theory. The principle, discussed already, that a man ought not to be punished for doing what he cannot help creates difficulties when extended to actions which a man could not help doing because of his own state of mind instead of external or contingent factors, like duress or ignorance of fact. An insane man, as defined, say, by the M'Naghten rules (that is, one who did not know what he was doing or did not know that what he was doing was wrong), cannot be said to choose his act because he cannot know it for what it is. But sometimes a man may know that what he is doing is wrong yet still be unable to stop himself from doing it. He may be subject, for instance, to an irresistible temptation or provocation. But how is that to be understood? A temptation is not irresistible merely because a particular man has yielded to it or even because

he might have been expected to yield to it. However, a temptation may be so strong that we might expect any ordinary person to yield to it (even though a few people may in fact resist it), or, as one might say, it might be "more than human nature can stand." In that sense it may be "irresistible."

Some people, of course, find it much more difficult than others to resist temptation. Some, like kleptomaniacs, are "impelled" to act in the sense that deliberation neither plays, nor could play, any part in what they do. Such people might be distinguished from plain wrongdoers by the fact that nothing—not blame, punishment, praise, or rational argument—seems to affect their disposition to break the rules. Or, again, their actions may lack any point, or if they can be said to have any point, it is only in relation to a set of aims and standards of achievement so distorted and eccentric that they are intelligible only to a psychiatrist. The kleptomaniac who steals nylon stockings for which he has no possible use (according to ordinary standards of utility) might properly be said to be unable to help stealing them. Far more difficult is the case of the psychopath, who seems to have no wish to resist temptation or, rather, who knows that some of the things he wants to do are wrong in the sense that other people disapprove of them but on whom this knowledge enforces no internal restraint beyond prompting a degree of caution. Criminals of this type would once have been described as "wicked" but are now often described as incapable of self-control. To say, however, that they are not responsible for their acts creates the odd situation that anyone is liable to punishment who usually resists temptation but sometimes fails, whereas the man who never resists is not liable at all.

The determinist has a short way with these difficulties. Since everyone's actions are the response of his character to a given set of circumstances, how can anyone ever be held responsible for his actions? We do what we must, given what we are, and what we are is the end of a causal chain going back to before we were born. If one knew a person well enough, one might predict that under given conditions he would commit a crime. Is this compatible with saying that he can choose whether to do so, or is his belief in his freedom to choose simply an illusion? Can the result of a genuine choice be predicted?

To say that something is predictable is not, however, the same as saying it is unavoidable. We can forecast a man's actions just because we know the kind of choices that he regularly makes. The more we know of his dispositions and his preferences, the more likely we are to be right. But that does not mean that he never acts voluntarily or that he never makes a real choice but only thinks he does. If all choices are illusions, what would a real choice be like? A man's behavior may be predictable because he can be relied upon to do what is reasonable, but to act with good reason is the very reverse of being subject to an inner compulsion. An essential difference between voluntary and involuntary action is that it makes sense to speak of the motives, aims, and reasons for the former but only of the causes of the latter. It is only when a person's behavior seems pointless or when explanations in terms of aims do not seem sufficient that we look for the kind of cause which would justify saying that he could not help himself. Of course, a complete account of voluntary and rational behavior must refer to causes as necessary conditions for action, but such causes would not constitute a sufficient explanation. An account of the electronic activity in the brain would not provide a sufficient explanation of a move in a game of chess unless the move was so completely and absurdly irrelevant that it had to be accounted for simply as the result of a nervous twitch. In that case, however, it would not really be a move in the game at all, not an action, indeed, but something that happens to the player. The weakness of the determinist position, insofar as it purports to undermine the notion of responsibility, is that it treats such abnormalities as the explanatory model for the normal.

It is arguable, in any case, that the concept of responsibility *requires* that human behavior be causally accountable rather than the reverse. As David Hume pointed out in *An Enquiry concerning Human Understanding,*

> [Where actions] proceed not from some *cause* in the character and disposition of the person who performed them, they can neither redound to his honour, if good; nor infamy, if evil. … The person is not answerable for them; and as they proceeded from nothing in him that is durable and constant, and leave nothing of that nature behind them, it is impossible [that] he can, upon their account, become the object of punishment or vengeance.

In Hume's view universal causality is consistent with the concept of choice and is a necessary condition for responsibility and, therefore, for blame and punishment.

Strictly speaking, all that is necessary for a theory of punishment is that human conduct should be capable of being modified by threats. For some people—for instance, compulsive lawbreakers like kleptomaniacs—that is not the case. Others, however, commit crimes believing they can escape punishment; still others, in a spirit of rebellion, indifference, or, more rarely, of mar-

tyrdom, prefer to do what they want and risk the consequences rather than conform. Why they prefer it—what conditions account for their being the men they are—is irrelevant. To say "they prefer it" is to say they might have chosen to do otherwise but did not, and that is all that is necessary for the concept "responsibility." To ask whether they were free to prefer otherwise, being what they were, is to ask whether they could choose to choose, and it is not clear that this really means anything. The experience of punishment may provide a reason for choosing differently next time, but to have a reason for choosing is not to be without a choice and, therefore, without responsibility.

EXTENUATION. Though a criminal may be held responsible for his actions, there may nevertheless be circumstances which, so it is said, diminish responsibility or extenuate guilt. Temptation or provocation, though not irresistible may have been very great. The offender may have had a good character, and there may be no reason to expect any future lapse.

In some cases mitigation of sentence on such grounds can be readily justified in utilitarian terms. Little is to be gained by punishing the obviously exceptional lapse; a very small penalty might be enough to dissuade other respectable people who might otherwise be tempted to imitate it and for whom the shame of being treated as a criminal, whatever the penalty, is usually deterrent enough.

However, it is not easy to show, at least in utilitarian terms, that mitigation is reasonable in all the instances in which it is commonly thought appropriate. Nor does everyone agree on what are extenuating circumstances. It is not self-evident that whoever is sorely (but not irresistibly) tempted should be treated more leniently than people who have done the same thing but under less temptation. A strong temptation might be withstood if there were sufficient counterinducement. Leniency might weaken the resolve of others in the future. Some people treat crimes of passion leniently; others would say that the temptation is so commonly felt that if people were not discouraged from taking the law into their own hands by treating offenses of this kind severely, such offenses would rapidly multiply. Again, some people would accept a plea of drunkenness as an extenuation of an offense, whereas others would consider it an aggravation.

It is doubtful whether our ideas on this aspect of punishment depend on utilitarian considerations. Nor is there any reason to suppose that any system of utilitarian argument could show them to be consistent and rational. It was suggested earlier that though the criteria of morality and law, of blame and punishment, are not identical, they influence one another. If we blame people less for yielding to strong temptation, we also feel they deserve a less severe punishment. But this only shifts the question a step back. Why should temptation mitigate blame?

A possible answer might be that at least some temptations can be pleaded as partial justifications. Thus, a man who pleads that he killed someone to shorten his sufferings or a woman who kills her deformed baby is appealing to another moral principle to excuse the act. Similarly, a man who kills his wife's lover might claim that his victim was violating his rights. These are not complete justifications, as a plea of self-defense would be, but they are excuses that count, as it were, against the initial presumption of guilt and so incline us to look at the offense more sympathetically and more leniently, whatever the advantages of severity in terms of deterrence, prevention, or reform. There is nothing irrational in striking a balance of desert.

But differences of opinion about a criminal's deserts often turn not on the way such a balance is struck but on the extent to which his judges (or their critics) are able to comprehend his action. Anyone who could imagine himself tempted in similar circumstances would probably be more sympathetic than someone who could not and who would therefore see no reason for being indulgent. On the other hand, anyone who suspected that he himself might yield to such a temptation and who flinched from the possibility might react to it with very great severity indeed.

PUNISHMENT AND REFORM

There is no reason to suppose, then, that the sentencing practice of the courts will display rational and consistent principles; furthermore, any attempt to set up criteria of rational judgment on strictly utilitarian principles is likely to cut across deeply rooted moral convictions. Accordingly, some criminologists and psychiatrists, such as Eliot Slater and Bernard Glueck, and some penal reformers, such as Barbara Wootton, have swung away from the general conceptions of punishment and desert. Instead of asking what penalty is warranted by the crime, whether the agent was fully responsible for his action, whether circumstances exonerate him wholly or in part, they prefer to ask what kind of treatment is most likely to rehabilitate him, subject, of course, to the example it might set for others.

This comes very close to repudiating altogether the concept of punishment as a deliberate infliction of suffer-

ing, which the criminal deserves, consequent to a voluntary breach of the law. First, the treatment most likely to rehabilitate him need not be unpleasant (though if it is to instill a measure of discipline, it very well may be). And, second, avoiding the question of moral responsibility, the reformer also avoids the question of what the criminal deserves, because the reformer's prime concern is with the treatment he needs. Criminals would no more deserve punishment than the sick deserve medicine. Indeed, for such writers as Samuel Butler and the American lawyer Clarence Darrow, criminality is a kind of sickness to be treated rather than a wrong to be punished.

Attractive as this approach may seem on humanitarian grounds, it has at least one serious consequence. The concepts of responsibility and desert cannot be discarded without some loss. For it is not a necessary condition of medical treatment that a patient must have shown symptoms of a disease; those exposed to smallpox are vaccinated before they develop a fever. Without the principle that punishment must be deserved, there would be no obstacle to subjecting people likely to become criminals to corresponding forms of penal prophylaxis. Moreover, if we substitute for punishment the idea of rehabilitative treatment, there is nothing against sentencing a person of bad character to a severe course of treatment for the most trivial offense if his character would be better for it in the end. This would clearly be incompatible with the usually accepted principle that trivial offenses should not carry severe penalties.

Reformism of this kind is open to attack from another quarter. The point has been made by Hegel and Bernard Bosanquet, among others, that retributive punishment is a kind of tribute to the moral personality of the criminal. It is precisely as a morally responsible agent, recognized as capable of making reasoned choices and accepting the consequences, that the criminal is punishable. Bosanquet goes so far as to say that punishment is "his right, of which he must not be defrauded." It is to be distinguished, argued Bradley, from the discipline or correction appropriately administered to animals and children. Punishment "is inflicted because of wrongdoing, as desert, the latter is applied as means of improvement." Since rational adults are neither animals nor children, no one has the right to treat them as if they were. It might be similarly argued that lunatics are under tutelage because they are incapable of looking after their own interests and cannot be expected to respect those of other people. The sane criminal, on the contrary, can be made to pay for his antisocial choices in order to demonstrate to him and, through him, to others that crime does not pay, but it diminishes his stature as a rational adult to deny that he is responsible for ordering his own life and to impose upon him ends of another person's choosing.

Nevertheless, retributivists have often been much concerned with moral reformation. They have insisted, however, that this was something the criminal must do for himself. Because it was associated with shame and rejection, punishment could bring the criminal up short and force him to reconsider his life in the light of society's condemnation of his actions. But the remorse that was a necessary condition for self-reformation was entirely dependent on the criminal's recognition that his punishment was deserved. Without that there could be no inward reformation, no reassertion of moral standards, but only a sense of resentment and injustice. Accordingly, punishment can yield the benefits of reform only if it is thought of, above everything else, as retributive—as the appropriate desert of a responsible guilty agent. It is this which distinguishes the retributive approach to moral reformation from the kind of utilitarianism which turns its back on desert and responsibility and is concerned only with the needs of rehabilitation.

It is, of course, an open question whether punishment ever does produce the kind of self-reformation the Hegelians had in mind or whether it does so more often than it produces a moral decay. Indeed, our knowledge of the facts of criminal behavior is probably far too scanty and uncertain for us to know how relevant much of the philosophical discussion of punishment really is. We cannot say for sure that a penal system is justified because it tends to reform criminals. Nor do we know, for that matter, whether the deterrent view of punishment is applicable to all kinds of crime. Many people commit offenses without seeming to take any account of consequences before they act, and they repeat the same offenses again and again in spite of punishment. Perhaps those who do not, would not repeat them even without punishment. Perhaps there would be no more cases of certain classes of crime than there are already; perhaps the only people to commit them are those who also do not take account of consequences before they act. It seems likely that some potential offenders are deterred from evading taxes or from smuggling by the threat of punishment, but is there any certain evidence that the threat of punishment deters anyone who would otherwise commit rape or arson?

Utilitarians tend to assume that punishment as an institution can be justified by its beneficial consequences, but the argument depends on certain a priori assumptions about criminal (or would-be criminal) behavior that may be greatly overintellectualized. However, even though

research should prove the usual utilitarian justifications for punishment groundless, that does not mean that some other, nonutilitarian justification is better. The proper procedure may well be to ask, with the utilitarian, whether the consequences are by and large beneficial; it is equally possible that punishment as an institution might fail that test. A theory of punishment that led to the conclusion that all punishment was wrong need be no more necessarily mistaken than a theory that led to a similar conclusion as regards, say, slavery, which, after all, was accepted as uncritically in Aristotle's day as punishment is today.

See also Aristotle; Bosanquet, Bernard; Bradley, Francis Herbert; Butler, Samuel; Good, The; Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Moral Rules and Principles; Rashdall, Hastings; Responsibility, Moral and Legal; Utilitarianism.

Bibliography

For a broad treatment of the subject, see A. C. Ewing, *The Morality of Punishment* (London: K. Paul, Trench, Trubner, 1929). For classic expositions of retributive positions, see Immanuel Kant, "Metaphysische Anfangsgründe der Rechtslehre," in *Die Metaphysik der Sitten* (1797), in *Werke,* Vol. VII (Berlin, 1922), translated by W. Hastie as *Kant's Philosophy of Law* (Edinburgh, 1887); G. W. F. Hegel, *Philosophie des Rechts* (1821), in *Werke,* Vol. VIII (Berlin, 1833), translated by T. M. Knox as *Hegel's Philosophy of Right* (Oxford: Clarendon, 1942). See also Ch. 8 of Bernard Bosanquet, *Philosophical Theory of the State* (London, 1899; 4th ed., Macmillan, 1923), and Essay 1 of F. H. Bradley, *Ethical Studies* (London, 1876; 2nd ed., rev., Oxford: Oxford University Press, 1927). Extensive bibliographic notes can be found in Giorgio Del Vecchio, *La giustizia,* which was originally published in *Annuario dell'Università di Roma* (1922–1923), 4th ed. (Rome, 1951), translated by Lady Guthrie as *Justice,* edited by A. H. Campbell (Edinburgh, 1952).

More recent analyses, mainly in retributive terms, can be found in J. D. Mabbott, "Punishment," in *Mind* 48 (1939): 152–167; C. W. K. Mundle, "Punishment and Desert," in *Philosophical Quarterly* 4 (16) (1954): 216–228; D. C. Hodges, "Punishment," in *Philosophy and Phenomenological Research* 18 (2) (1957–1958): 209 (a case for viewing punishment as reparation). K. Baier, "Is Punishment Retributive?," in *Analysis* 16 (2) (1955): 25–32, argues that the concept of punishment entails retribution but does not deal with problems of justification. See also A. R. Manser, "It Serves You Right," in *Philosophy* 37 (142) (1962): 293–306. H. J. McCloskey, "The Complexity of the Concepts of Punishment," ibid., 307–326, draws attention to variety of contexts in which "punishment" is used and argues against analysis in terms of a standard case.

For classic statements of utilitarian positions, see C. B. Beccaria, *Dei delitti e delle pene* (1764); *An Essay on Crimes and Punishment, Translated from the Italian,* 4th ed. (London, 1785); Jeremy Bentham, *Introduction to the Principles of Morals and Legislation* (1789), Wilfrid Harrison, ed. (Oxford, 1948), and *Traités de législation, civile et pénale,* E. Dumont, ed. (Paris, 1802), translated by R. Hildreth as *Theory of Legislation,* C. K. Ogden, ed. (London, 1931). See T. H. Green, *Lectures on the Principles of Political Obligation,* in *Works,* edited by R. L. Nettleship, Vol. II (London, 1885), or in a separate edition, with an introduction by A. D. Lindsay (London: Longman, 1941); Green repudiates retributivism but stresses moral reformation as a function of punishment and, therefore, the importance of the concept of desert (cf. Bosanquet, op. cit.). See also Vol. I of Hastings Rashdall, *Theory of Good and Evil,* 2 vols. (Oxford: Clarendon Press, 1907; 2nd ed., 1924).

The following examples of recent analyses of the problems of punishment are mainly utilitarian in emphasis but try to allow some place for retributivism in a general theory. In W. D. Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930), especially Appendix II, Ross argues that the state is entitled to punish offenders in the public interest because by invading the rights of others, they forfeit their own claim to protection. On this subject see also A. G. N. Flew, "The Justification of Punishment," in *Philosophy* 29 (3) (1954): 291–307; H. L. A. Hart, "Prolegomenon to the Principles of Punishment," in *PAS* 60 (1959–1960): 1, and "Murder and the Principles of Punishment: England and the United States," in *Northwestern University Law Review* 52 (4) (1957): 433. Also, see A. Quinton, "On Punishment," and Stanley I. Benn, "An Approach to the Problems of Punishment," in *Freedom and Responsibility,* edited by Herbert Morris (Stanford, CA: Stanford University Press, 1961); J. Rawls, "Two Concepts of Rules," in *Society, Law, and Morality,* edited by F. A. Olafson (Englewood Cliffs, NJ: Prentice-Hall, 1961); Stanley I. Benn and R. S. Peters, *Social Principles and the Democratic State* (London: Allen and Unwin, 1959), especially Chs. 8–9. Patrick J. Fitzgerald, *Criminal Law and Punishment* (Oxford: Clarendon Press, 1962), deals with the philosophical problems in the broad context of English criminal law. For a discussion of the relations of law, morality, and punishment, see Patrick Devlin, *The Enforcement of Morals* (London: Oxford University Press, 1965), and H. L. A. Hart, *Law, Liberty and Morality* (Stanford, CA: Stanford University Press, 1963).

For discussions of free will, determinism, and responsibility in relation to punishment, see H. L. A. Hart, Paul Edwards, and John Hospers, symposium on "Determinism and Freedom in Law and Ethics," in *Determinism and Freedom,* edited by Sidney Hook (New York: New York University Press, 1958); J. D. Mabbott, "Freewill and Punishment," in *Contemporary British Philosophy,* edited by H. D. Lewis, 3r series (London, 1956). In Barbara Wootton, *Social Science and Social Pathology* (London: Allen and Unwin, 1959), there is a survey of recent trends in psychiatric and criminological studies critical of the concept "responsibility"; see especially Ch. 8. For a criticism of these trends, see H. L. A. Hart, *Punishment and the Elimination of Responsibility* (Oxford: Clarendon Press, 1968). *Freedom and Responsibility* (see above) is a useful collection of readings on various aspects of the problem of responsibility, including its relation to punishment, extenuation, and so on; it includes an extensive bibliography.

Stanley I. Benn (1967)

PUNISHMENT [ADDENDUM]

Since 1967, preventive theories of punishment (whether strictly utilitarian or more loosely consequentialist) have entered a long decline, beginning with the virtual disappearance of reform theory in the 1970s. Crowding them out are various alternatives generally categorized as "retributive."

All preventive theories treat punishment as (primarily) a means of controlling objectionable behavior. Insofar as they propose to justify punishment on the assumption that penalties can be tuned to achieve a certain degree of social control, they are empirically vulnerable (as well as morally vulnerable for ignoring justice). What had become clear by the 1970s was that social science could not then, or in the foreseeable future, give preventive theories much empirical content. Social science could not, that is, say what effect, if any, statutory penalties, rehabilitation, exemplary punishments, or even the incapacitation of criminals would have on the crime rate. If even relatively crude tuning of penalties to consequences is in practice impossible, preventive theories cannot justify choosing one system of punishment over another, much less one punishment over another.

In contrast to preventive theories, retributive theories do not seek to justify punishment by pointing to an empirical relation between punishment and "consequences" (such as a certain crime rate). For retributive theories, the relation between punishment and its justification is conceptual ("internal," as Georg Wilhelm Friedrich Hegel would say). No retributivist need deny that punishment has some general tendency to control crime or that that tendency is a reason to have some punishment system rather than none. All retributivists need deny is that punishment's (actual or probable) tendency to control crime matters much for understanding why we should (or should not) have this institution, practice, or act of punishment rather than another.

Retributive theories may be divided into two importantly different kinds: moralistic and legalistic. Moralistic retributivism has three (main) divisions: desert, paternalist, and condemnatory theories. *Desert* theory takes it as (more or less) brute fact that wrongdoing deserves an unpleasant response, that is to say, punishment. Punishment is justified because it is deserved. Nothing more need be said. *Paternalist* theory holds that all justified punishment, or at least all justified punishment of rational agents, must aim at a certain good for those punished. This good may be subjective (R. A. Duff's "penance") or objective (Robert Nozick's "connection with correct values"). *Condemnatory* theory, in contrast, understands punishment as (primarily) an "expressive act" not meant to benefit anyone.

All three varieties of moralistic theory are retributive (in the sense used here) because all seek to achieve a good that is conceptually related to punishment. For desert theories, that good is simply giving wrongdoers what they deserve. Degree of desert determines severity of punishment. For paternalist theories, the justification of punishment lies in the way punishment treats the wrongdoer—for example, as a being capable of learning justice from the punishment appropriate to the crime. The seriousness of the wrong determines what penalty is appropriate to teach the lesson that the crime shows the wrongdoer needs to learn. For condemnatory theories, the justification of punishment lies in what the punishment "expresses." The denunciation should be as emphatic as the crime was bad; the more severe the punishment, the more emphatic the denunciation is.

While desert theories seem to be the direct descendants of traditional retributivism, paternalist theories superficially resemble traditional *reform* theories, and condemnatory theories similarly resemble traditional *deterrence* theories (denunciation resembling a deterrent threat). Both nonetheless differ fundamentally from the corresponding preventive theory. According to the paternalist theory (in its pure form, at least), punishment would be justified even if wrongdoers never repent or learn as a result of punishment. What is important—important because it respects the moral personality of the wrongdoer—is that the right punishment be imposed with the right intention. In much the same way, according to the condemnatory theory (in its pure form), punishment is justified even if emphatic denunciation has no effect on the crime rate or on the individual's later conduct. Reaffirming the wrongness of an act is good in itself, good enough (all else being equal) to justify the punishment.

All moralistic theories share the assumption that punishment belongs to ordinary morality (rather than to the law in particular). Moralistic theories use ordinary moral practices (such as disciplining children) to understand punishment (with legal punishment only a special case). Moralistic theories differ primarily in the part of ordinary morality to which they assign punishment. Desert theory treats punishment as (negative) rewarding; paternalist theory treats punishment as correction or teaching; and condemnatory theory treats punishment as moral statement. Other moralistic theories are possible,

for example, one treating punishment as a form of self-defense or satisfaction of a promise.

Legalistic theories, in contrast, assume that (justified) punishment is a practice (largely) confined to (relatively just) legal systems. The only important form of legalistic retributivism today is "the fairness theory" (also known as "benefits-and-burdens," "reciprocity," or "unfair advantage" theory). It holds that legal punishment (and close analogues) is justified insofar as it supports the (relatively) just distribution of benefits and burdens that a (relatively just) legal system (or similar practice) creates. A relatively just legal system is to be thought of as a cooperative enterprise from which each benefits if others generally do their part and in which doing one's part will sometimes be burdensome. According to the fairness theory (in its pure form at least), the institution of legal punishment is justified if punishment keeps lawbreakers from gaining an unfair advantage over the law-abiding. Punishment, if just, necessarily takes back the unfair advantage the crime as such takes (or, at least, some fair equivalent of that advantage). Though the fairness theory has an obvious affinity to certain theories of distributive justice (especially, Rawlsian social contract), it presupposes no particular theory. All it presupposes is that there can be an equivalence between crime and (just) punishment assuring that (in general at least) legal punishment of certain people in certain ways will (as a conceptual matter) increase (or at least help to maintain) overall distributive justice (however defined). Explaining that presupposition has proved difficult.

See also Hegel, Georg Wilhelm Friedrich; Nozick, Robert; Presupposition; Rawls, John.

Bibliography

For some sense of the history of punishment since 1967, compare Marvin Zalman, "The Rise and Fall of the Indeterminate Sentence," *Wayne Law Review* 24 (1977): 45–94; Nigel Walker, *Punishment, Danger, and Stigma* (Oxford: Blackwell, 1980); and Annika Snare, ed. *Beware of Punishment*, Scandinavian Studies in Criminology, 14 (1995).

For some broad general treatments of punishment theory, see Ted Honderich *Punishment: The Supposed Justifications* (London: Hutchinson, 1969); Hyman Gross, *A Theory of Criminal Justice* (New York: Oxford University Press, 1979); C. L. Ten, *Crime, Guilt, and Punishment: A Philosophical Introduction* (Oxford: Clarendon, 1987); Igor Primoratz, *Justifying Legal Punishment* (Atlantic Highlands, NJ: Humanities Press International, 1989); and Michael Moore, *Placing Blame: A General Theory of the Criminal Law* (Oxford: Clarendon, 1997). Comparison of these works (spaced about a decade apart) suggests the increasing importance of retributivism, even when (as with Ten) the author himself argues for a preventive theory.

For some important criticism of retributivism, see Richard W. Burgh, "Do the Guilty Deserve Punishment?" *Journal of Philosophy* 79 (1982): 193–210; John Cottingham, "Varieties of Retribution," *Philosophical Quarterly* 29 (1979): 236–246; John Braithwaite and Philip Pettit, *Not Just Deserts: A Republican Theory of Criminal Justice* (Oxford: Clarendon, 1990), ch. 8–9; David Dolinko, "Some Thoughts about Retributivism," *Ethics* 101 (1991): 537–559; David Dolinko, "Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment," *Law and Philosophy* 16 (1997): 507–528; and Nigel Walker, "Even More Varieties of Retribution," *Philosophy* 74 (1999): 595–605.

For the best recent statement of the desert version of retributivism, see D. E. Scheid, "Constructing a Theory of Punishment, Desert, and Distribution of Penalties," *Canadian Journal of Law and Jurisprudence* 10 (1997): 441–506. Other important discussions include Lawrence H. Davis, "They Deserve to Suffer," *Analysis* 32 (1972): 136–140; John Kleinig, *Punishment and Desert* (The Hague: Nijhoff, 1973); W. A. Parent, "The Whole Life View of Criminal Desert," *Ethics* 86 (1975): 350–354; and Douglas N. Husak, "Why Punish the Deserving?" *Nous* 26 (1992): 447–464.

For the most influential statements of the paternalistic theory, see Herbert Morris, "A Paternalistic Theory of Punishment," *American Philosophical Quarterly* 18 (October 1981): 263–271; Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981), pp. 363–397; Jean Hampton, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13 (1984): 208–238; and R. A. Duff, *Trials and Punishments* (Cambridge, U.K.: Cambridge University Press, 1986).

Among important discussions of the expressive theory are: A. J. Skillen, "How to Say Things with Walls," *Philosophy* 55 (1980): 509–523; Igor Primoratz, "Punishment as Language," *Philosophy* 64 (1989): 187–205; Michael Davis, "Punishment as Language: Misleading Analogy for Desert Theorists," *Law and Philosophy* 10 (1991): 310–322; Andrew von Hirsch, *Censure and Sanctions* (Oxford: Clarendon, 1993). Though Joel Feinberg's influential "The Expressive Function of Punishment," *Monist* 49 (1965): 397–423, is usually treated as a seminal statement of the expressive theory—as reprinted in Feinberg's *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton, NJ: Princeton University Press, 1970)—it is not. It actually proposes to revise the standard (Hart-Flew) definition of punishment, something necessary because of Feinberg's legal positivism. Since retributivism aims at justice, not mere social control, retributive expressivists do not need to fiddle with the definition of punishment to distinguish between (justified and so, expressive) punishment and mere penalties. Retributivists have therefore generally ignored Feinberg's proposal (while citing him as pointing to the expressive function of punishment).

For some examples of moralistic retributivism relying neither on desert, paternalism, nor condemnation, see: Phillip Montague, *Punishment as Societal Defense* (Lanham, MD: Rowman & Littlefield, 1995); and Jerry Cederblom, "The Retributive Liability Theory of Punishment," *Public Affairs Quarterly* (1995): 305–315.

For important discussions of the fairness theory (and related issues), see: Herbert Morris, "Persons and Punishment," *The Monist* 52 (1968): 475–501; John Finnis, "The Restoration of Retribution," *Analysis* 32 (1971–1972): 131–135; J. Murphy, "Marxism and Retribution," *Philosophy and Public Affairs* 2 (1973): 217–243; Wojciech Sadurski, *Giving Desert its Due: Social Justice and Legal Theory* (Dordrecht: Kluwer, 1985); George Sher, *Desert* (Princeton, NJ: Princeton University Press, 1987); Richard Dagger, "Playing Fair with Punishment," *Ethics* 103 (1993): 473–488; and Anthony Ellis, "Punishment and the Principle of Fair Play," *Utilitas* (1997): 81–97. For an example of a legalist retrubutivism that at least seems not to be an instance of the fairness theory, see Margaret R, Holmgren, "Punishment as Restitution: The Rights of the Community," *Criminal Justice Ethics* (1983): 36–49.

For recent examples of the application of retributivism to various practical problems in the theory of punishment, see Kathleen Dean Moore, *Pardons: Justice, Mercy, and the Public Interest* (New York: Oxford University Press, 1989); Jeffrie G. Murphy and Jean Hampton, *Forgiveness and Mercy* (Cambridge, U.K.: Cambridge University Press, 1990); Michael Davis, *To Make the Punishment Fit the Crime: Essays in the Theory of Criminal Justice* (Boulder, CO: Westview, 1992); R. A. Duff, *Criminal Attempts* (Cambridge, U.K.: Cambridge University Press, 1996); Andrew von Hirsch and Andrew Ashworth, eds. *Principled Sentencing: Readings on Theory and Policy* (Oxford: Hart, 1998); and Stephen Kershnar, "An Argument for the Use of Torture as Punishment," *Hamline Journal of Public Law and Policy* (1998): 497–528.

Recent positive work on deterrence seems to be more than a decade old. See, for example (beside Ten, 1987, above), D. Schmidtz, "Deterrence and Criminal Attempts," *Canadian Journal of Philosophy* 17 (1987): 615–623; Daniel M. Farrell, "On Threats and Deterrence," *Social Theory and Practice* 15 (1989): 367–394; and B. Sharon Bird, "Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution," *Law and Philosophy* 8 (1989): 151–200. Reform theory may be reviving as "restorative justice" within the victim's rights movement (but with a strong retributive overlay). See, for example, Charles K. B. Barton, *Getting Even: Revenge as a Form of Justice* (Chicago: Open Court, 1999).

Michael Davis (2005)

PUTNAM, HILARY

(1926–)

Hilary Putnam, after receiving a BA from Pennsylvania (1948) and a year spent at Harvard (1948–1949), studied at the University of California, Los Angeles, taking his doctorate in 1951 with a dissertation titled "The Concept of Probability: An Application to Finite Sequences." He taught at Northwestern (1952–1953), Princeton (1953–1961) and MIT (1961–1965), becoming Walter Beverly Pearson professor at Harvard in 1965. From 1995 to 2000, he served as Cogan University Professor there, becoming emeritus in 2000. He has been influential in most areas of philosophy, particularly in the philosophy of language, of logic, of mathematics, and of science.

Putnam is sometimes thought of as often changing his mind. (See, for example, the *Dictionary of Philosophers' Names.*) Sometimes he has. But in central respects he has held a single, though developing, position since the mid-1950s, a position that in some aspects resembles the later Ludwig Wittgenstein's. This entry sets out some constant central themes.

Putnam was among those American philosophers to benefit directly from the intellectual exodus from Europe caused by Nazism. He was a student of Rudolf Carnap and of Hans Reichenbach. Though his approach to issues is quite different from theirs, Reichenbach in particular had a lasting and often acknowledged influence on Putnam's thought. Putnam's innovations stand out when it is noted. In *Realism with a Human Face* (1990, p. 289), Putnam remarks,

> In *Theory of Relativity and A Priori Knowledge* (1922) Reichenbach listed a number of statements … each of which Kant would have regarded as synthetic *a priori,* and each of which can be held immune from revision …, but which *collectively imply statements that are empirically testable,* and that Kant would, therefore, have to regard as *a posteriori.*

Certain principles had, in Immanuel Kant's time, as good a claim as any to fix how particular spatial, temporal, and other concepts are to be applied, and thereby which concepts those were; to be intrinsic to the concepts involved, thus "conceptual truths," thus a priori. Relativity theory allows us to see how they are at least jointly testable, so that some may turn out false. Such, it seems, is a fate to which a priori truths are liable.

Putnam reports Reichenbach as making a related point to his classes. Considering questions such as "How can we show that that blackboard is wider than this ashtray?," he argued that any system of measurement, or of observation, treats some propositions that seem empirical (such as "mere translation does not make things grow or shrink") as axiomatic. One cannot sensibly apply the system while doubting these propositions; they are not subject to confirmation or refutation within the system. But it could prove reasonable to replace the system with another in which these propositions are testable, so possibly false. In that sense they are empirical.

There are two contrasting reactions to these points. One is: What this shows is that every concept commits itself to a particular empirical theory. If the theory proves false, then the concept is incoherent, so without application and to be discarded. This was Paul Feyerabend's reaction, and it is also Paul Churchland's.

The other reaction is: If we are confronted with situations that force giving up what seemed conceptual truths, it may appear that the concepts whose applications seemed to be governed by those principles are, in fact, otherwise governed. Perhaps the application of the concept "straight line" to items in the world is not governed by the Euclidean parallel postulate, but rather in such and such other way. That reaction grants face value to Reichenbach's point that the same proposition that is axiomatic in one system may be testable and false in another. This was Putnam's reaction. He developed the position and drew its implications in a powerful series of papers in the 1950s and 1960s (see Putnam, 1962a–d). Part of the idea is that what principles govern the application of a concept depends in part on how the world in fact is. Putnam defined that role for the world in "The Meaning of 'Meaning'" (1975). This last article, though not published until 1975, was completed by 1968.

By the early 1970s, Putnam had begun to emphasize some new themes. For one thing, he became increasingly impressed with what he calls the "interest relativity" of such notions as explanation and cause. The general point is: What a concept counts as applying to—the correct way of applying it—varies with the circumstances in which it is to be applied. A concept may count, on one occasion, as fitting what it does not count as fitting on another. That is continuous with Putnam's earlier reaction to Reichenbach. The point then was: What it is reasonable to judge as to how a concept operates depends on the conditions in which such judgments are made. The point now is: What those conditions are depends not just on how the world is, but may vary from occasion to occasion, given the world as it is. Not coincidentally, this point went along with other developments in Putnam's thought.

The first of these developments is what he calls "internal realism," first presented in 1976, and amplified in his writings of 1981, 1983, 1987a, 1987b, and elsewhere. The position includes four points. First, there are mundane, true things to say about what our words and thoughts are about: "the word 'gold' means (refers to) gold; this is gold"; "This is a chair; this is what 'chair' refers to," and so forth. Second, there are philosophical dicta that sound much like such mundanities, or their denials, or generalizations of these, but that say, or try to say, something quite different. They are bad answers to the following pseudo-problem. On the one hand, there are thoughts and words—items that purport to represent the world as being thus and so; on the other hand, the items the world in fact contains, which are what and how they are independent of what we think, or do not think, about them. How are our words and thoughts related to these items? How, if at all, does their truth depend on how those items are? And how could they be so related? Internal realism holds that the problem rests on a mistake; hence so do any 'solutions', which take it at face value.

Third, the mundane remarks (point one) are correct because they are a feature of how these words are (or are to be) used. But that formulation depersonalizes things misleadingly. The standard for the correctness of a statement cannot be fixed independently of what users of the relevant words and concepts—that is, human beings—are prepared to recognize as correct: What Putnam identifies as our (human) perceptions of rationality and reasonableness. What it is for a statement to be correct depends on the sorts of beings we are, and is not reducible to some set of principles that would have to hold anyway. Fourth, it is part of what we are prepared to recognize as rational that any concept might be applied correctly in different ways in different circumstances. What sometimes counts as the cause of the explosion may not at other times. It is because human rationality is occasion-sensitive that the problem mentioned in point two is a pseudo-problem. We cannot sensibly take a "God's-eye view" of how we relate to the world, trying to say how our concepts would apply without us.

The occasion-sensitivity of rationality does not mean that truth is relative, or that there are no objective facts—given a framework, or setting, in which concepts are to be applied. Nor does giving up on a God's-eye view mean a deflationist account of truth. Putnam insists that we cannot comprehend what truth is without understanding the role of truth in our lives, notably in our activities of asserting, and of treating assertions in the ways we do; and that deflationism does not help us understand the role of truth in human life.

In arguing against the possibility of a God's-eye view, Putnam has produced what are probably his most discussed arguments. In one he identifies the God's-eye view (what John McDowell has called "the view of the cosmic alien") as one from which we may consider our own language as an uninterpreted calculus with a range of possible interpretations, and then ask which interpretation is the right one. In what he first saw as a generalization of the Skolem paradox, he argues that, in that case, *nothing*

could make one interpretation the right one, so we could not ever be talking about anything (or about one thing rather than others). But we cannot pose serious problems without talking about definite things. This is a reductio of the idea of a God's-eye view (see "Models and Reality," Philosophical Papers, Vol. 3).

In another argument Putnam considers the (apparent) question whether we might be brains in a vat: that we are, and always have been, nothing but brains, kept alive by a bath of nutrients, fed computer-generated stimuli through electrodes. He argues that if the God's-eye view is possible, then that we are, and that we are not such brains should both be possibilities. But for the words of the question to mean what a God's-eye view requires them to means we must be using them in ways that entail that we are not brains in vats. For, as argued in "The Meaning of 'Meaning,'" what our words mean depends, inter alia, on how we are in fact connected to the world, and not just on what we may anyway be aware of. For our "brain" to mean *brain,* and our "vat" to mean *vat,* we must be connected to the world as brains in ways vats could not be. So we cannot formulate what, from a God's-eye view, ought to be a possibility, in a way that makes it possible. That is another reductio of the idea of a God's-eye view (see Putnam 1981).

Equally important to internal realism are Putnam's arguments against a causal theory of reference: Arguments, based on the interest-relativity of causation, that our being causally linked to the world as we, in fact, are is not enough in itself to make some one interpretation of our language correct—once it is granted that the language we speak may coherently be viewed by us as less than fully interpreted, so open to interpretation. These arguments appear in many replies to critics, and notably in "Realism with a Human Face" (1990).

At about the time Putnam began to develop internal realism, he also began to change his way of thinking about human psychology, rejecting a picture of it, and with that, a view he once espoused—functionalism. Viewed one way, a human being is an organism constructed in a particular way, a particular battery of mechanisms arranged to interact with each other and the environment in given ways. If, while taking that view, we ask what it is for someone to believe that Mars is a planet, or to have any propositional or other attitude—to be in a mood, experience an emotion, and so on—it is tempting to look for an answer by trying to identify some state(s) of some mechanism(s) such that for someone to believe that is for him to be structured like that. In that frame of mind, for example, one might speak seriously of someone having a "token of a mentalese sentence" in his "belief box." This is the picture Putnam rejects.

Against it Putnam notes that to ascribe belief to someone is to relate that person to the world as we view it, and to ourselves, as on the same side as ours, or a different one, with respect to such and such question as to how the world is, and so on. Given internal realism, this means that there will be different truths to tell on different occasions as to what a given person, as he is at a given time, then believes. So for someone to be as said to be when we say him to believe thus and so, cannot be for him to have some particular mechanism, otherwise identifiable, in some particular state. And so on for other mental states (see Putnam 1989).

Putnam has been refining the ideas discussed above, notably the idea of a distinction between ordinary and philosophic statements, and applying them in new areas, such as philosophy of mathematics. The above indicates a few main themes, omitting Putnam's striking arguments for them.

See also Carnap, Rudolf; Functionalism; Kant, Immanuel; Rationality; Realism; Reference; Reichenbach, Hans; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY PUTNAM

"The Analytic and the Synthetic" (1962a). Reprinted in *Mind, Language and Reality, Philosophical Papers,* Vol. 2. Cambridge, U.K.: Cambridge University Press, 1975.

"It Ain't Necessarily So" (1962b). Reprinted in *Mathematics, Matter and Method, Philosophical Papers,* Vol. 1. Cambridge, U.K.: Cambridge University Press, 1975.

"Dreaming and 'Depth Grammar'" (1962c). In *Philosophical Papers,* Vol. 2. Cambridge, U.K.: Cambridge University Press, 1975.

"What Theories Are Not" (1962d). In *Philosophical Papers,* Vol. 1. Cambridge, U.K.: Cambridge University Press, 1975.

"The Meaning of 'Meaning.'" In *Philosophical Papers,* Vol. 2. Cambridge, U.K.: Cambridge University Press, 1975.

"Realism and Reason" (1976). Reprinted in *Meaning and the Moral Sciences.* London: Routledge and Kegan Paul, 1978.

Reason, Truth and History. Cambridge, U.K.: Cambridge University Press, 1981.

"Models and Reality." Reprinted in *Realism and Reason, Philosophical Papers,* Vol. 3. Cambridge, U.K.: Cambridge University Press, 1983.

The Many Faces of Realism. LaSalle, IL: Open Court, 1987.

"Realism with a Human Face." In *Realism with a Human Face, Philosophical Papers,* Vol. 4. Cambridge, MA: Harvard University Press, 1990.

Renewing Philosophy. Cambridge, MA: Harvard University Press, 1992.

Words and Life. Cambridge, MA: Harvard University Press, 1994.

WORKS ON PUTNAM

Clark, P., and B. Hale, eds. *Reading Putnam.* Oxford: Blackwell, 1994.

Ebbs, G. *Rule-Following and Realism.* Cambridge, MA: Harvard University Press, 1997.

Rorty, R. "Solidarity or Objectivity." In *Post-analytic Philosophy,* edited by J. Rajchman and C. West. New York: Columbia University Press, 1985.

Charles Travis (1996)

PYRRHO

(c. 360–c. 270 BCE)

Pyrrho of Elis is much less well known than the eponymous philosophy he inspired, Pyrrhonism. Diogenes Laertius, in his biography of Pyrrho in *Lives of Eminent Philosophers*, offers his usual mixture of anecdote, scandal, and unreliable doctrinal information (9.61–70). Thus Pyrrho was said to have had no concern for his own safety and to have been rescued from precipices and oncoming traffic by the timely interventions of his (presumably nonskeptical) friends (9.62). Aenesidemus rejected such fables, "saying that while he did philosophize in accordance with suspension of judgment, he did not act in a heedless manner" (9.62). Such stories, as well as ones presumably designed to exalt his image, such as one that claims that he demonstrated his unworldly indifference by washing pigs (9.66), are apocryphal, but they indicate what others apparently took his skeptical detachment to amount to.

He is also said to have traveled with Anaxarchus as far as India, where he consorted with the "gymnosophists," the naked philosophers, which led him "to philosophize in the noblest manner, adopting nonapprehension (*akatalēpsia*) and suspension of judgment (*epochē*); he said that nothing was good or bad, just or unjust, and that in all cases nothing is really true, but that men act by law and custom in all cases; for each thing is no more (*ou mallon*) thus than not" (9.61). This suggests that Pyrrho did indeed institute much of what was to become distinctive about later Pyrrhonism, but it also implies that he was primarily concerned with ethical questions (broadly construed)—an impression confirmed by some later evidence, most noticeably that of Cicero, who treats him exclusively as originating an obsolescent quietist ethics.

However, other testimonies suggest a broader engagement with more general epistemological themes. Pyrrho himself wrote nothing, but a disciple, Timon of Phlius, lauded his master in both prose and poetry, of which some seventy-one fragments survive. Some aredevoted to exalting Pyrrho's imperturbable and noble character at the expense of the "vanity" of other philosophers. But more important for an assessment of Pyrrho's philosophical position is a report of a passage from one of Timon's prose works, embedded in an antiskeptical tract of the Aristotelian Aristocles of Messene, and itself preserved by way of the Christian Eusebius. (Such is the tortuous route of the early skeptical tradition, and attempting to purge such intrinsically hostile reports of later accretions of distortion and selectivity is a serious scholarly challenge.)

Aristocles reports Pyrrho as holding that "we are so constituted as to know nothing," and hence that all inquiry is pointless. Since for Sextus Empiricus and other later skeptics, the first claim would be unacceptably (if negatively) dogmatic, and since Sextus defines the Pyrrhonian way as one of continued (if unrequited) inquiry, this may be a misrepresentation on Aristocles' part. But it is equally possible that Pyrrho did conclude that nothing was knowable and inquiry futile, and that the subsequent rejection of these views was a later development. This latter possibility gains support from recent suggestions (see Bett 2000) that Pyrrho was not in fact the skeptical hero that later skeptics, such as Sextus, wanted to paint him as for ideological reasons of their own. Both Aenesidemus and Sextus treated Pyrrho as an archetypical role model. But Sextus, while remarking that "the skeptic school … is called 'Pyrrhonian' from the fact that Pyrrho seems to have taken up scepticism more thoroughly and conspicuously than any of his predecessors" (1.7), rarely mentions him by name elsewhere in his large *oeuvre*. And Diogenes reports that one Numenius claimed that Pyrrho dogmatized, that is, held positive tenets (9.68). If this is right, then Pyrrho might have become a model because of his legendary imperturbability, rather than because of any practice of skeptical argumentation. Working against this view are some of the citations quoted and Diogenes' claim that Pyrrho was skilled in dialectical argument (9.64), which was to be the hallmark of later skepticisms.

At all events, Timon, as reported by Aristocles, states that according to Pyrrho, "[i] One must consider three questions: First, how are things by nature? Second, what should our attitudes toward them be? Third, what will be the result of adopting such an attitude? … [ii] Pyrrho

declared all things equally indifferent, unmeasurable, and undecidable; [iii] for this reason [or since] neither our sensations nor our judgments are true or false. [iv] Consequently, we should not put our trust in them but should be without opinion, uncommitted, and unswayed, saying of each thing no more [*ou mallon*] [a] that it is than [b] that it is not, or [c] that it both is and is not, or [d] that it neither is nor is not. [v] For those thus disposed, the consequence will be first nonassertion [*aphasia*] then tranquility [*ataraxia*]" (Aristocles, in Eusebius, 14.18.2; = Long and Sedley Fr. 1F–5; my translation).

This is by far our most important philosophical testimony for Pyrrho, and it is based upon (and perhaps reports verbatim) a text of Pyrrho's own pupil. It is, however, multiply difficult to interpret. The following interpretation is one way to try to make sense out of this difficult passage. The "things" (*pragmata*) of [i] may be states of affairs in the world or, more vaguely, subjects of possible cognition. "Indifferent" (*adiaphora*) in [ii] may perhaps be better rendered "undifferentiable" (thus making the claim not about the metaphysical condition of things but rather about our epistemic position with regard to *pragmata*). "Since," the alternative connective of [iii], represents a textual conjecture that has won some scholarly support, and it has the obvious effect of reversing the direction of dependence between [ii] and [iii]. The scope of "no more" in [iv] is unclear (the sentence is syntactically ambiguous); it may govern four disjuncts (including disjuncts [c] and [d]), rather than just the first two. Finally, the precise sense of "aphasia" in [v] is disputed.

With the connective "for this reason" in [iii] (the reading of the manuscripts), the assertion is that the indeterminacy of things renders our sensations and judgments about them neither true nor false. At first sight, this seems to be a strange inference. (Does not indeterminacy simply render them false, insofar as they make positive claims that fail to correspond to the indeterminate facts?) Yet the assertion can be made intelligible if we suppose that for a claim of the form "*x* is *F*" to be true, *x* must be unequivocally *F*, and equally for it to be false, *x* must be wholly not *F*. Thus, because of the indeterminacy in things, any unequivocal statement of the form "*x* is *F*" will be partly true and partly false, and hence neither wholly true nor wholly false. Consequently, withholding strong belief and commitment makes sense. (In favor of the manuscript reading, the sequence of consequence is maintained: from states of affairs, via epistemic consequences, to epistemic attitude, to pragmatic consequence.) In addition, the account of truth just given supports the interpretation that "no more" in [iv] has narrow scope, that is, that [c] and [d] are alternative ways of describing the counterpoise between [a] and [b]. To get this interpretation, we have to understand "it is" to mean "it is unequivocally" in [a] and [b], but not in [c] and [d].

The upshot, then, is that we will make no statements; not that we will literally say nothing, but that we will express no strong commitment to the unequivocal truth of our first-order remarks about the world. And when we attain this state, tranquility will follow like a shadow. As later skeptics put it, once one stops seeking to make (and support) unequivocal claims about the world, all one's initial anxiety (apparently caused by the second-order belief that there *should* be answers to such questions and the consequent frustration of not finding them) vanishes. If all this is right, then Pyrrho really was a recognizable precursor to the later skepticism that took his name. Yet Pyrrho was not a thoroughgoing skeptic, for as the interpretation offered above suggests (but does not demand), Pyrrho did commit himself to speculations at least about the actual (Heraclitean) state of affairs of things in a way that Sextus Empiricus at least (although perhaps not Aenesidemus) would have found anathema.

See also Aenesidemus; Ancient Skepticism; Sextus Empiricus; Timon of Phlius.

Bibliography

Annas, J., and J. Barnes. *Outlines of Pyrrhonism*. Cambridge, U.K: Cambridge University Press, 1994.

Bett, Richard. *Pyrrho, His Antecedents, and His Legacy*. Oxford, U.K.: Oxford University Press, 2000.

Diogenes Laertius. *Lives of Eminent Philosophers*. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1972.

Eusebius of Caesarea. *Preparation for the Gospel*. Translated by Edwin Hamilton Gifford. Oxford, U.K.: Clarendon Press, 1903.

Hankinson, R. J. *The Sceptics*. London: Routledge, 1995.

R. J. Hankinson (2005)

PYRRHONIAN PROBLEMATIC, THE

KNOWLEDGE AND JUSTIFICATION

If a belief is to count as knowledge, then it must be true. But truth is not enough: lucky guesses and, more generally, beliefs that are only accidentally related to the facts

they purport to describe do not amount to knowledge. What else, besides truth, is needed for a belief to count as knowledge, then? There is no agreement regarding how to fully answer that question, but there is a line of thought regarding how to begin such an answer that is widely shared: for a belief to amount to knowledge it has to be justified or supported by reasons, or rationally grounded, or warranted, or have some sort of positive epistemic status. (These, and other, words are sometimes used as synonyms, whereas sometimes they are intended to mark important epistemological distinctions. I use them interchangeably.) The justification in question here is usually qualified as epistemic, to distinguish it from the kind of justification that, for example, an assassin's mother might have in believing that her son is innocent despite mounting evidence against him.

It is possible to adopt many different attitudes with respect to any proposition *p* (say, the proposition that Paris is the capital of France). For instance, it is possible to believe that Paris is the capital of France, to be happy that Paris is the capital of France, to hope that Paris is the capital of France, and so on. Some of these attitudes can be called doxastic attitudes. What distinguishes a doxastic attitude from other attitudes we can adopt toward a proposition is that one can be justified or unjustified (in the epistemic sense) in adopting a doxastic attitude.

There are three basic doxastic attitudes: belief, disbelief, and suspension of judgment. (It might be that there are other attitudes that we might be justified or unjustified, in the epistemic sense, in adopting, but belief, disbelief, and suspension of judgment are basic in the sense that any other doxastic attitude will be such only because it entails one of these three basic attitudes.) To disbelieve that Paris is the capital of France is to believe that it is not true that Paris is the capital of France (and so, depending on how you count, you might think that there are only two basic doxastic attitudes: belief and suspension of judgment). To suspend judgment with respect to the proposition that Paris is the capital of France is to be in a mental state that is opposed both to believing and disbelieving the proposition. Suspension of judgment must therefore be carefully distinguished from having no attitude whatsoever with respect to a certain proposition. There is a difference between never having considered the question whether there is an even number of stars in the Milky Way and, having considered it, suspending judgment with respect to the question.

ACADEMIC AND PYRRHONIAN SKEPTICISM

If the connection between knowledge and justification presented earlier is correct, then we can know a proposition only if we are justified in believing it. Skepticism with respect to a range of propositions is the claim that the only justified attitude with respect to the propositions in that range is to suspend judgment. We are all skeptics, in this sense, with respect to some range of propositions. For instance, it seems obvious that the only correct attitude with respect to the proposition that there is an even number of stars in the Milky Way, once we have considered it, is to suspend judgment. This is ordinary skepticism. But most of us are nonskeptics with respect to many propositions. For instance, it seems obvious that the only justified attitude with respect to the proposition that Paris is the capital of France is to believe it, whereas the only justified attitude with respect to the proposition that Tony Blair is the president of the United States is to disbelieve it. Philosophical skepticism extends well beyond ordinary skepticism, claiming that we should suspend judgment with respect to propositions that we ordinarily think we are justified in believing.

It is customary to distinguish between two different kinds of philosophical skepticism, which can be called, following an ancient tradition, Academic skepticism and Pyrrhonian skepticism. Academic skepticism referred originally to a phase in the history of Plato's Academy that stretched approximately from the third to the early first century BCE. The main figures of Academic skepticism were Arcesilaus (mid-third century BCE), Carneades (mid-second century BCE), and Clitomachus (d. 110/109 BCE). The main sources for Pyrrhonian skepticism are the writings of Sextus Empiricus in the late second century CE.

Academic and Pyrrhonian skepticism differ in the scope of propositions that, according to them, we should suspend judgment about. Let's call those propositions that do not contain any epistemic concepts ordinary propositions and let's call those propositions to the effect that someone knows an ordinary proposition can be called epistemic propositions.

Academic skeptics think that the only justified attitude with respect to most (perhaps all) ordinary propositions is suspension of judgment. However, Academic skeptics do not suspend judgment with respect to epistemic propositions: On the contrary, they think that the only justified attitude with respect to them is to disbelieve them—that is, they think that we are justified in believing that we do not know almost anything of what we take

ourselves to know. (When contemporary authors discuss skepticism, chances are they are referring to this aspect of Academic skepticism: to the claim that we do not know certain propositions that we ordinarily take ourselves to know. However, the tradition is to classify as a skeptic with respect to a certain proposition only someone who thinks we should suspend judgment with respect to that proposition, not someone who thinks that we should dissent from it.)

Pyrrhonian skeptics, meanwhile, extend their skepticism to epistemic propositions as well. Both Academic and Pyrrhonian skeptics leave it open whether Paris is the capital of France or not: maybe it is, maybe it is not, but we are not justified in believing that it is or believing that it is not. According to Pyrrhonian skeptics it is also an open question whether we know that Paris is the capital of France: maybe we do, maybe we do not, but we are not justified in believing that we do or that we do not. Academic skeptics, on the contrary, do not leave this question open: they think we are justified in believing that we do not know that Paris is the capital of France.

THE MODES OF AGRIPPA

From now on, the focus will be on Pyrrhonian skepticism exclusively The Pyrrhonians had a number of ways, or modes, to induce suspension of judgment. The importance of Pyrrhonian skepticism to contemporary epistemology derives primarily from these modes, and in particular from a subset of them referred to collectively as the modes of Agrippa. There are five modes associated with Agrippa, but three of them are the most important: the mode of hypothesis (or unsupported assertion), the mode of circularity (reciprocal), and the mode of regression to infinity.

The three modes of Agrippa function together in the following way. Whenever the dogmatist (Sextus refers to those who are not skeptics as dogmatists) asserts his or her belief in a proposition p_1, the Pyrrhonian will challenge that assertion, asking the dogmatist to justify p_1, to give reasons for thinking that it is true. The dogmatist will then either decline to answer the challenge or adduce another proposition p_2 in support of p_1. If the dogmatist refuses to answer the challenge, the Pyrrhonian will be satisfied that the only justified attitude to take with respect to p_1 is to suspend judgment, because no reason for it has been given (thus appealing to the mode of hypothesis). If the dogmatist adduces another proposition p_2 in support of p_1, then either p_2 will be identical to p_1 or it will be a different proposition. If p_2 is the same proposition as p_1, then the Pyrrhonian will also suspend judgment with respect to p_1, because no proposition can support itself (thus appealing to the mode of circularity). If, however, p_2 is different from p_1, then the Pyrrhonian will ask the dogmatist to justify his or her assertion of p_2. And now the dogmatist offers no reason in support of p_2, offers p_2 itself or p_1 as a reason, or adduces yet another proposition p_3, different from both p_1 and p_2. If the dogmatist offers no reason for p_2, then the Pyrrhonian will invoke the mode of hypothesis again and suspend judgment in accordance with it; if either p_2 itself or p_1 is offered as a reason to believe in p_1, then the Pyrrhonian will invoke the mode of circularity and suspend judgment in accordance with it (because not only can no proposition be a reason for believing in itself but also no genuine chain of reasons can loop); and, finally, if the dogmatist offers yet another proposition p_3, different from both p_1 and p_2, as a reason to believe p_2, then the same three possibilities that arose with respect to p_2 will arise with respect to p_3.

The dogmatist will not be able to continue offering different propositions in response to the Pyrrhonian challenge forever—eventually, either no reason will be offered, or a proposition that has already made an appearance will be mentioned again. The Pyrrhonian refers to this impossibility of actually offering a different proposition each time a reason is needed as the mode of infinite regression. The three Pyrrhonian modes, then, work in tandem to induce suspension of judgment with respect to any proposition whatsoever.

AGRIPPA'S TRILEMMA

The Pyrrhonian use of the three modes of Agrippa to induce suspension of judgment can be presented in the form of an argument, called Agrippa's trilemma. It is at least somewhat misleading to present the Pyrrhonian position in terms of an argument, because in presenting an argument one is usually committed to the truth of its premises and conclusion, whereas Pyrrhonian skeptics would suspend judgment with respect to them. Nevertheless, presenting the Pyrrhonian problematic in the form of an argument does not do much violence to this skeptical position, because what is important is not whether the Pyrrhonian skeptics themselves accept the premises or the validity of the argument, but whether their audience does. Problems still remain regarding the coherence of anyone (be they Pyrrhonian skeptics or not) who accepts the soundness of an argument whose conclusion is that we are not justified in believing anything. It is doubtful, though, whether anyone accepts Agrippa's trilemma: "Dogmatists" certainly do not, and neither do Pyrrhon-

ian skeptics. It is not a coincidence that Wittgenstein's dictum about throwing the ladder after using it to climb echoes Sextus's less-pleasing image of the laxative that purges itself together with the "humours" of the body it is designed to expel. Still, even if we do not think that the argument is sound, we stand to learn something interesting about the structure of an epistemological theory—because each of the premises of the apparently valid argument looks plausible at first sight.

Before presenting a reconstruction of Agrippa's trilemma some definitions need to be introduced. Say that a belief is inferentially justified if and only if it is justified (at least in part) in virtue of its relations to other beliefs. A justified basic belief, by contrast, is a belief that is justified but not in virtue of its relations to other beliefs. An inferential chain is a set of beliefs such that every member of the set is allegedly related to at least one other member by the relation is justified by. Agrippa's trilemma, then, can be presented thus:

(1) If a belief is justified, then it is either a basic justified belief or an inferentially justified belief.

(2) There are no basic justified beliefs.

Therefore,

(3) If a belief is justified, then it is justified in virtue of belonging to an inferential chain.

(4) All inferential chains are such that either (a) they contain an infinite number of beliefs; (b) they contain circles; or (c) they contain beliefs that are not justified.

(5) No belief is justified in virtue of belonging to an infinite inferential chain.

(6) No belief is justified in virtue of belonging to a circular inferential chain.

(7) No belief is justified in virtue of belonging to an inferential chain that contains unjustified beliefs.

Therefore,

(8) There are no justified beliefs.

Premise (1) is beyond reproach, given our previous definitions. Premise (2) is justified by the mode of hypothesis. Step (3) of the argument follows from (1) and (2). Premise (4) is also beyond reproach—the only remaining possible structure for an inferential chain to have is to contain basic justified beliefs, but there are none of those according to premise (2). Premise (5) is justified by appeal to the mode of infinite regression, and (6) is justified by appeal to the mode of circularity. Premise (7) might seem to be a truism, but some authors have argued that denying it is the only plausible way out of Pyrrhonian skepticism.

It is interesting to note that Agrippa's trilemma is perfectly general; in particular, it applies to philosophical positions as well as to ordinary propositions. In fact, when Agrippa's trilemma is applied to epistemological theories themselves, the result is called "the problem of the criterion."

Many contemporary epistemological positions can be stated as a reaction to Agrippa's trilemma. In fact, all of premises (2), (5), (6), and (7) have been rejected by different philosophers at one time or another. In the remainder of this entry, we examine each of these responses.

REJECTING PREMISE (2): FOUNDATIONALISM

Foundationalists claim that there are basic justified beliefs—beliefs that are justified but not in virtue of their relations to other beliefs. In fact, according to foundationalists all justified beliefs are either basic beliefs or are justified in virtue of being inferentially related to a justified belief (or to some justified beliefs). This is where foundationalism gets its name: The edifice of justified beliefs has its foundation in basic beliefs.

But how do foundationalists respond to the mode of hypothesis? If a belief is not justified by another belief, then is it not just a blind assertion? If basic beliefs are justified but not by other beliefs, then how are they justified? What else besides beliefs is there that can justify beliefs?

To this last question, many foundationalists reply: experience (if they are talking about empirical knowledge, of course; *a priori* knowledge raises interesting problems of its own, and it is also subject to Agrippa's trilemma). To a rough first approximation that glosses over many important philosophical issues, experiences are mental states that, like beliefs, aim to represent the world as it is, and, like beliefs, can fail in achieving that aim—that is, experiences can misrepresent. Nevertheless, experiences are not to be identified with beliefs, for it is possible to have an experience as of, for example, facing two lines that differ in length without having the belief that one is facing two lines that differ in length—a combination of mental states that anyone familiar with the Müller-Lyer illusion will recognize.

There are three important questions that any foundationalist has to answer. First, what kinds of beliefs do experiences justify? Second, how must inferentially acquired beliefs be related to basic beliefs for them to be

justified? Third, in virtue of what do experiences justify beliefs?

TRADITIONAL AND MODERATE FOUNDATIONALISM. With respect to the first question, we can distinguish between traditional foundationalism and moderate foundationalism. Traditional foundationalists think that basic beliefs are beliefs about experiences, whereas moderate foundationalists think that experience can justify beliefs about the external world. Take, for example, the experience that you typically have when looking at a tomato under good perceptual conditions—an experience that, remember, can be had even if no tomato is actually there. A moderate foundationalist would say that that experience justifies you in believing that there is a tomato in front of you. The traditional foundationalist, however, would say that the experience justifies you only in believing that you have an experience as of a tomato in front of you. You may well be justified in believing that there is a tomato in front of you, but only inferentially.

A traditional argument in favor of traditional foundationalism relies on the fact that whereas you can be mistaken regarding whether there is a tomato in front of you when you have an experience as of facing a tomato, you cannot, in the same situation, be mistaken regarding whether you are undergoing such an experience. From the point of view of traditional foundationalism, this fact indicates that the moderate foundationalist is taking an unnecessary epistemic risk—the risk of having a foundation composed of false beliefs.

The moderate foundationalist can reply that the traditional foundationalist must undertake a similar risk. For, while it is true that if one is undergoing a certain experience then one cannot be mistaken in thinking that one is undergoing that experience, one can still be mistaken about one's experiences—for instance, perhaps one can believe that one is in pain even if the experience that one is undergoing is actually one of feeling acutely uncomfortable. And if it were just as difficult to distinguish between the true and the false in the realm of beliefs about our own experiences as it is in the realm of beliefs about the external world, then we could be wrong about which of our own beliefs are basically justified and which are not. If this kind of metafallibilism is accepted, then why not accept the further kind according to which basic justified beliefs can be false? Of course, the resolution of this dispute depends on whether, as the moderate believes, we can be mistaken about our own experiences.

DEDUCTIVIST AND NONDEDUCTIVIST FOUNDATIONALISM. What about our second question: How must basic beliefs be related to inferentially justified beliefs? Here, too, there are two different kinds of foundationalism: deductivism and nondeductivism. According to the deductivist the only way in which a (possibly one-membered) set of basic justified beliefs can justify another belief is by logically entailing that other belief. In other words, there has to be a valid argument whose premises are all basic justified beliefs and whose conclusion is the inferentially justified belief in question. Given that the argument is valid, the truth of the premises guarantees the truth of the conclusion—it is impossible for all the premises to be true while the conclusion is false. Nondeductivism allows relations other than logical entailment as possible justificatory relations. For instance, many foundationalists will claim that good inductive inferences from basic justified beliefs provide their conclusions with justification—even though inductive arguments are not valid, that is, even though it is possible for all the premises of a good inductive argument to be true while its conclusion is false. Although these are independent distinctions, traditional foundationalists tend to be deductivists, whereas moderate foundationalists tend to be nondeductivists. Notice that for a traditional, deductivist foundationalist, there cannot be false justified beliefs. Many contemporary epistemologists would shy away from this strong form of infallibilism and take that consequence to be an argument against the conjunction of traditional foundationalism and deductivism.

PRIMITIVIST, INTERNALIST, AND EXTERNALIST FOUNDATIONALISM. The question that is most interesting from the point of view of the Pyrrhonian problematic is our third one: What is it about the relation between an experience and a belief that, according to the foundationalist, allows the former to justify the latter? (Analogous questions apply to nonfoundationalist positions too, and the discussion to follow is not restricted to the specific case of foundationalism.) There are three different proposals about how to answer this question that are the most prominent. The principles that assert that a subject is justified in having a certain belief given that he or she is undergoing a certain experience can be called epistemic principles. Our third question can then be stated as follows: What makes epistemic principles true?

The first proposal, which we shall call primitivism, claims that the question cannot have an intelligible answer. There is no more basic fact in virtue of which epistemic principles obtain. They describe bedrock facts, not to be explained in terms of anything else, but are

instead to be used to explain other facts. Epistemological theorizing, according to the primitivist, ends with the discovery of the correct epistemic principles.

The other two positions are nonprimitivist. Internalist nonprimitivism holds that epistemic principles are true in virtue of facts about ourselves—for instance, one prominent internalist view is that which epistemic principles are true for a given subject is determined by which epistemic principles that subject would accept under deep reflection. Externalist nonprimitivism holds that epistemic principles are true in virtue of facts that are not about ourselves—for instance, one prominent externalist view is that certain experiences provide justification for certain beliefs because the obtaining of those experiences is reliably connected to the truth of those beliefs (reliabilism), or because i.e., it could not easily happen that those experiences obtain without those beliefs being true (an appeal to "sensitivity" or "safety" conditionals).

Both externalists and internalists think that primitivists are overlooking real facts, whereas primitivists think that there are fewer things in heaven and earth than are dreamt of in nonprimitivist philosophy. Within the nonprimitivist camp externalists think that internalists have too subjective a conception of epistemology—to some extent, thinking it so, or being disposed to think it so under conditions of deep reflection, makes it so for at least some traditional internalists. Internalists, for their part, are likely to think that externalists are no longer engaged in the same project that both skeptics and internalist epistemologists are engaged in, the project of determining "from the inside" whether one's beliefs are justified or amount to knowledge, because the obtaining of a relation between a subject's belief and the external world is something that the subject is in no position to ascertain "from the inside."

REJECTING PREMISE (5): INFINITISM

Infinitism, the claim that infinite evidential chains can provide justification to their members, is the answer to Agrippa's trilemma that has received the least attention in the literature. This is due, at least in part, to the fact that infinitism has to deal with what might seem like formidable obstacles. For instance, it seems that no one actually has an infinite number of beliefs. To this objection, the infinitist is likely to reply that actually occurring beliefs are not needed, only implicit beliefs that are available to the subject to continue constructing his or her inferential chain if called on to do so (by others or by him- or herself). The plausibility of this reply depends on whether good sense can be made of the notion of implicit belief and the notion of an implicit belief's being available for a subject.

Even leaving that problem aside, the infinitist, like the coherentist, maintains that justification can arise merely in virtue of relations among beliefs. Infinitists will then have to respond to many of the same objections that are leveled against coherentism—in particular, they would have to respond to the isolation objection mentioned in the next section.

REJECTING PREMISE (6): COHERENTISM

Coherentists reject two related features of the picture of evidential reasons that underlies Agrippa's trilemma. The first feature is the idea that justification is an asymmetrical relation: if a belief p_1 justifies a different belief p_2, then p_2 does not justify p_1. The second feature is the idea that the unit of justification is the individual belief. Putting these two rejections together, the coherentist believes that justification is a symmetrical and holistic matter. It is not individual beliefs that are justified in the primary sense of the word, but only complete systems of beliefs—individual beliefs are justified, when they are, in virtue of belonging to a justified system of beliefs. The central coherentist notion of justification is best taken to be a comparative one: A system of beliefs B1 is better justified than a system of beliefs B2 if and only if B1 has a greater degree of internal coherence than B2. One crucial question that coherentists have to answer, of course, is what it takes for one system of beliefs to have a greater degree of coherence than another. Many coherentists have thought that explanatory relations will be crucial in elucidating the notion of coherence: The more explanatorily integrated a system is, the more coherence it displays.

The main objection that coherentists have to answer is called the isolation objection. The objection centers on the fact that, according to the coherentist, the justification of a system of beliefs is entirely a matter of relations among the beliefs constituting the system. But this runs against the strong intuition that experience has an important role to play in the justification of beliefs. To illustrate the problem, suppose that you and I both have a highly coherent set of beliefs—your system, it is safe to assume, contains the belief that you are reading, whereas mine does not, and it contains instead the belief that I am swimming (because, let us suppose, I am swimming right now). Suppose now that we switch systems of beliefs—somehow, you come to have my set of beliefs and I come to have yours. Given that coherence is entirely a matter of relations among beliefs, your system will be as coherent in

my mind as it was in yours, and vice versa. And yet, our beliefs are now completely unjustified—there you are, reading, believing that you are swimming, and here I am, swimming, believing that I am reading. In other words, certain transformations that preserve coherence in a system of beliefs do not seem to preserve justification.

In reply, coherentists argue that it is possible to give experience a role without sacrificing the idea that justification is entirely a matter of relations among beliefs—one idea is to require that any minimally acceptable system of beliefs contain beliefs about the experiences that the subject is undergoing. It is fair to say that there is no agreement regarding whether this move can solve the problem.

REJECTING PREMISE (7): POSITISM

One position that can be traced back to some ideas in Ludwig Josef Johann Wittgenstein's *On Certainty* (published posthumously in 1969)—and, perhaps, also to José Ortega y Gasset's *Ideas y Creencias* (1940)—is that evidential chains have to terminate in beliefs that are not properly said to be either justified or unjustified. This position, which we shall call positism (not to be confused with positivism), shares many features with foundationalism: for instance, both positists and foundationalists agree that inferential chains have to be finite and noncircular. But, whereas the foundationalist thinks that the starting points of inferential chains are beliefs that are justified by something other than beliefs, the positist thinks that the starting points of inferential chains are beliefs that are not justified by anything—they are posits that we have to believe without justification. Despite this difference between the positist and the foundationalist, the positions are structurally similar enough that analogues of the questions posed to the foundationalist can be asked of the positist.

First, then, which beliefs are such that they are not justified and yet are the starting points of every inferential chain—in other words, how do we identify which are the posits? One answer that can be gleaned from Wittgenstein's *On Certainty*, which we will call relativistic positism, is that this is a matter that is relative both to time and society, because what the posits are is determined by some function of the actual positing practices of the members of one's society at a certain time. Thus, according to Wittgenstein the proposition that no one has been to the moon was a posit for a certain long period of time—it was a proposition that no one felt the need to justify, and that was presupposed in many justificatory practices. For obvious reasons, though, that proposition can no longer appropriately function as a posit. Other epistemologists, nonrelativistic positists, think that which beliefs are properly posited depends on some objective truth about which beliefs have to be presupposed to engage in the practice of justifying beliefs at all. One prime candidate for playing this role is the first-person belief that I am not being deceived by an evil demon into thinking that I am a normally embodied and situated human being.

The second question, regarding how posits must be related to inferred beliefs to justify them, can receive answers that are completely analogous to the foundationalists'. The third question, applied to positism, is the question why certain beliefs are properly posited. Relativistic positists answer that this is so because of a certain societal fact: because they are taken to be so by an appropriate subsector of a certain society at a certain time. Nonrelativistic positists answer that a certain belief is properly taken as a posit just in case every justificatory act that we engage in presupposes that the belief in question is true.

One objection that positists of both sorts have to face is that they are transforming a doxastic necessity into an epistemic virtue—that is, they are concluding that certain beliefs can properly serve as the starting points of inferential chains because that is how in fact they are treated (relativistic positism) or because otherwise it would not be possible to engage in inferential practices at all (nonrelativistic positism). The Pyrrhonian skeptic, of course, will reply that the mere fact that most members of a society accept a certain belief without justification, or even the fact that if we do not do so then we cannot justify anything else, does not mean that it should be accepted without justification.

CONCLUSION

Perhaps one of the most interesting developments in relation to the Pyrrhonian problematic is that more and more epistemologists are arguing that the proper way to reply to Agrippa's trilemma is to combine some of the positions that, for ease of exposition, we have presented as mutually exclusive (this development is explicit in contemporary authors such as Sosa, but, some will argue, it is already present in Descartes). Thus, for example, many contemporary epistemologists put forward theories that contain elements of both internalism and externalism, or foundationalism and coherentism. It is a testament to the endurance of the Pyrrhonian problematic that philosophers continue in this way to grapple with it.

See also Agrippa; Ancient Skepticism; Arcesilaus; Carneades; Classical Foundationalism; Coherentism; Descartes, René; Greek Academy; Internalism versus Externalism; Ortega y Gasset, José; Plato; Pyrrho; Sextus Empiricus; Skepticism, History of; Wittgenstein, Ludwig Josef Johann.

Bibliography

Barnes, Jonathan. *The Toils of Scepticism*. New York: Cambridge University Press, 1990.

BonJour, Lawrence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Burnyeat, Myles. "Can the Skeptic Live His Skepticism?" In *Doubt and Dogmatism: Studies in Hellenistic Epistemology*, edited by Malcolm Schofield, Myles Burnyeat, and Jonathan Barnes. Oxford, U.K.: Clarendon Press, 1980.

DePaul, Michael R., ed. *Resurrecting Old-Fashioned Foundationalism*. Lanham, MD: Rowman and Littlefield, 2001.

Fogelin, Robert J. *Pyrrhonian Reflections on Knowledge and Justification*. New York: Oxford University Press, 1994.

Klein, Peter. "Human Knowledge and the Infinite Regress of Reasons." *Philosophical Perspectives* 13 (1999): 297–325.

Ortega y Gasset, José. *Ideas y Creencias*. Buenos Aires: Espasa Calpe, 1940.

Sextus Empiricus. *Outlines of Scepticism*. 2nd ed., edited by Julia Annas and Jonathan Barnes. New York: Cambridge University Press, 2000.

Sosa, Ernest. "How to Resolve the Pyrrhonian Problematic: A Lesson from Descartes." *Philosophical Studies* 85 (2–3) (1997): 229–249.

Wittgenstein, Ludwig. *On Certainty (Uber Gewissheit)*, edited by G. E. M. Anscombe and G. H. von Wright. Translated by Denis Paul and G. E. M. Anscombe. Oxford: Basil Blackwell, 1969–1975.

Juan Comesaña (2005)

PYTHAGORAS AND PYTHAGOREANISM

Pythagoras was an Ionian Greek born on the island of Samos, probably about 570 BCE. His dislike of the policies of the Samian tyrant Polycrates caused him to immigrate to Crotona in southern Italy. There he founded a society with religious and political, as well as philosophical, aims that gained power in the city and considerably extended its influence over the surrounding area. A certain Cylon, however, stirred up a revolt against the society in which a number of its leading members were killed, and Pythagoras retired to Metapontum. The community recovered its influence until a more serious persecution took place in the middle of the fifth century, from which the survivors scattered to various parts of the Greek world—notably Thebes, Phleius, and Tarentum. In these places "they preserved their original ways and their science, although the sect was dwindling, until, not ignobly, they died out" (in the late fourth century), to quote the epitaph written by a contemporary.

NATURE OF THE EVIDENCE

The obstacles to an appraisal of classical Pythagoreanism are formidable. There exists no Pythagorean literature before Plato, and it was said that little had been written, owing to a rule of secrecy. Information from the Christian era is abundant but highly suspect. Pythagoras himself, though a fully historical figure, underwent a kind of canonization. His life was quickly obscured by legend, and piety attributed all the school's teaching to him personally. Moreover, the dispersion of the school inevitably led to divergences of doctrine in the various groups. Aristotle makes it clear that by the late fifth century some Pythagoreans were teaching one thing and some another. A further reason for division was that the universal genius of Pythagoras, for whom religion and science were two aspects of the same integrated worldview, was beyond the scope of lesser men. Some naturally inclined more to the religious and superstitious; others, to the intellectual and scientific side, as is confirmed by later references to the division between *acusmatici* and *mathematici.*

As early evidence there are several references to Pythagoras in works of his contemporaries or near contemporaries (for instance, Xenophanes satirized his belief in the transmigration of souls), a valuable reference in Plato to the relationship between astronomy and harmonics in the Pythagorean system, a quantity of information from Aristotle (who at least would not confuse the Pythagoreans with Plato, as later writers excusably did), and some quotations from pupils of Aristotle who were personally acquainted with the last generation of the school.

Given the nature of the sources, the following is a fairly conservative summary of Pythagoreanism before Plato.

MAN AND THE COSMOS

In contrast with the Milesians, the Pythagoreans were not motivated by disinterested scientific curiosity. For Pythagoras, philosophy was the basis of a way of life, leading to salvation of the soul. "Their whole life," said a fourth-century writer, "is ordered with a view to following God, and it is the governing principle of their philosophy." At philosophy's center, therefore, were man and his relation to other forms of life and to the cosmos. Purity

was to be sought by silence, self-examination, abstention from flesh and beans, and the observance of other primitive taboos that the Pythagoreans interpreted symbolically. Of the recognized gods they worshiped Apollo, guardian of the typically Greek ideal of moderation ("nothing too much"), of whom Pythagoras was believed to be an incarnation.

Behind both the superstition and the science was the notion of kinship or sympathy. The kinship and essential unity of all life made possible the belief in the transmigration of souls and accounted for the prohibition of meat: a sheep might house the soul of an ancestor. Not only animate nature in our sense but the whole world was akin, for the cosmos itself was a living, breathing creature. The cosmos was one, eternal, and divine; men were divided and mortal. But the essential part of man, his soul, was not mortal; it was a fragment of the divine, universal soul that was cut off and imprisoned in a mortal body. Men should therefore cultivate and purify the soul, preparing it for a return to the universal soul of which it was a part. Until then, since it was still contaminated by the body, it must tread the wheel of reincarnation, entering a new body of man or animal after the death of its previous tenement.

These tenets were also taught by the religious movement known as Orphism, from which the religious side of Pythagoreanism can hardly be separated. (Pythagoras himself was said in the fifth century to have written books under the name of Orpheus.) But whereas the Orphics sought salvation by purely religious means—sacramental ceremonies and the observance of ritual prohibitions—Pythagoras added a new way, the way of philosophy.

Philosophy, for Pythagoras as for others, meant the use of reason and observation to gain understanding of the universe. The link between this procedure and his overriding aim of salvation seems to have been the principle that like is known by like, a widespread tenet of pre-Socratic thought, common to such diverse systems as the philosophicoreligious synthesis of Empedocles and the scientific atomism of Democritus. Hence, an understanding of the divine universe would bring man's nature closer to its own. In this conception we meet the typically Pythagorean conception of *kosmos,* a word that combines in an untranslatable way the notion of orderly arrangement or structural perfection with that of beauty. Closely linked with it is *peras,* meaning limit. An organic whole, particularly one that, like the universe, lives forever, must of necessity exhibit limit and order in the highest degree. What is unlimited has no *telos* (end) and is *a-teles,* which means both "endless" and "incomplete." But the world is a perfect whole, a model of order and regularity, supremely exemplified in Greek eyes by the ceaseless wheeling of the heavenly bodies in (as they believed) perfect circles, bringing about the unvarying succession of day and night and seasons. It was said of Pythagoras that he was the first to call the world *kosmos,* "from its inherent order." By studying this order, we reproduce it in our own souls, and philosophy becomes an assimilation to the divine, as far as that is possible within the limitations imposed by our mortal bodies.

THE DOCTRINE THAT THINGS ARE NUMBERS

The Pythagoreans studied mathematics in a cosmic context, and for them numbers always retained a mystical significance as the key to the divine cosmos. "They supposed the whole heaven to be a *harmonia* and a number," said Aristotle. *Harmonia,* though specially applied to music, could signify any well-organized structure of parts fitted together in due proportion. Its effect in music seems to have burst on Pythagoras as a revelation of the whole cosmic system. We may accept the many later statements that he discovered the numerical ratios underlying the intervals that the Greeks called consonant and used as the basis of their scale. They involve only the numbers 1 to 4–1:2, octave; 3:2, fifth; 4:3, fourth. These numbers add up to 10, a sacred number for the Pythagoreans, which was symbolized by the dotted triangle (*tetractys*), "source and root of everlasting nature." From the discovery that the sounds they recognized as beautiful depended on inherent, objective, mathematical order, they leaped to the conclusion that number was the key to the element of order in nature as a whole.

With this innovation the Pythagoreans would seem to have taken the momentous step from explanation in terms of matter (as the Milesians had sought it) to explanation in terms of form. Yet philosophy was not quite ready for that step, nor could the distinction between matter and form be clearly grasped. They saw simply the ultimate, single nature (*physis*) of things in their mathematical structure. There seems little doubt that probably until well on in the fifth century they thought it possible to speak of things as actually made up of "numbers" that were regarded simultaneously as units, geometrical points, and physical atoms. Lines are made of points; surfaces, of lines; solids, of surfaces; and physical bodies, of solids. In this scheme two points made a line; three, the minimum surface (triangle); four, the minimum solid (tetrahedron). A later theory spoke of the "fluxion" of point into line, line into surface, and so on, which gave a

geometrical progression (1, 2, 4, 8) instead of the arithmetical (1, 2, 3, 4), and the sequence of point, line, square, cube. Based on continuity, it seems designed to avoid the problem of incommensurable magnitudes or irrational numbers.

Whenever they were discovered (probably not much later than 450), incommensurables had dealt a blow to the original "things are numbers" doctrine, the idea that geometrical figures—and thus ultimately the physical world—are based on a series of integers. No ratio between integers can either describe the relation between the diagonal of a square and its side or serve as the basis of construction of a right triangle. If, however, magnitudes are regarded as continuous and hence infinitely divisible, the existence of incommensurable or irrational magnitudes (those which cannot be expressed as a ratio of natural numbers) could be explained and the difficulty overcome.

THE ULTIMATE PRINCIPLES

The analysis went further than that outlined above, for numbers themselves have their elements. The ultimate principles were limit and the unlimited, which were equated with good and bad respectively; moral concepts went side by side with physical concepts in this extraordinary system. Abstractions as well as physical phenomena were equated with numbers; for instance, justice was 4, the first square number, symbolizing equality or requital. After limit and the unlimited came odd and even instances, respectively, of these two. They generated the unit (considered to be outside the number series, and both odd and even), from the unit sprang numbers, and from numbers came the world. There seems no doubt that the scheme goes back to an ultimate duality that corresponds to the moral dualism of Pythagoreanism, but one can also see how monistically minded Neoplatonic commentators could speak of the cosmos as originating from the One. In general terms, *kosmos* was achieved by the imposition of limit on the unlimited in order to make the limited, just as the imposition of definite ratios on the indefinite range of musical pitch produced the *harmonia* of the scale.

COSMOGNY AND COSMOLOGY

Cosmogony starts with the planting of a unit in the infinite. Aristotle called it, among other things, a seed; and since limit was associated with male and unlimited with female, the Pythagoreans probably thought of the generation of the living cosmos as taking place as did that of other animals. It grows by drawing in and assimilating the unlimited outside, that is, by conforming it to limit and giving it numerical structure. Physically the process resembles inspiration, and the unlimited is also called breath.

The unit seed had the nature of fire and in the completed cosmos (which evidently grew from the center outward) became a fire at its center. There are traces of two different cosmological schemes, a geocentric one that spoke of a fire at the center of Earth, and a more remarkable one attributed, in later sources at least, to the fifth-century Pythagorean Philolaus, which made Earth a planet. (Nicolas Copernicus in *De Revolutionibus* says that reading of this Pythagorean doctrine gave him courage to consider explaining the heavenly motions on the basis of a moving Earth.) According to this latter scheme, Earth, planets, sun, and moon—and an extra body called the counterearth—all revolved about the center of the universe, which was occupied by a fire invisible to man because he lived on the opposite side of Earth. It was known that the moon's light is borrowed, and the idea was extended to the sun, whose heat and light were said to be reflected from the central fire. The moon was eclipsed by the interposition of both Earth and the counterearth and, according to some, of further, otherwise unknown, planetary bodies. These caused the comparatively frequent lunar eclipses.

In this system, the mixture of religion and science in Pythagoreanism is well brought out. Fire was given the central position, not for any scientific reason but because it was regarded with religious awe—and the center is the most "honorable" place. It was lauded with such titles as Hearth of the Universe, Tower of Zeus, and Throne of Zeus. Yet the same thinkers were aware that with Earth in orbit "the phenomena would not be the same" as in a geocentric scheme (presumably they were thinking of the lack of stellar parallax and variations in the apparent size of the sun and moon). They pointed out that even with a central Earth, an observer would be separated from the center by the distance of its radius, and they argued that the visible effect would be as negligible in one case as in the other. This assumes that the heavenly bodies are at vast distances from Earth; and it is not known how, if at all, this system was related to the theory later known as the harmony of the spheres.

In any case, there are many divergent versions of this doctrine. In outline, the idea was that large bodies in motion must inevitably produce a sound; that the speeds of the heavenly bodies, judged by their distances, are in the ratios of the musical consonances; and that therefore the sound made by their simultaneous revolution is con-

cordant. We do not hear it because it has been with us from birth, and sound is perceptible only by contrast with silence. It has been plausibly argued that in the original version Pythagoras, like Anaximander, took only three orbits into account (sun, moon, and all the stars); this would relate it to his original musical discovery about the fourth, fifth, and octave. Later versions speak of seven, eight (Plato), and ten orbits. In any form, the doctrine emphasizes the universal importance, in Pythagorean eyes, of mathematical and musical laws and their intimate relation to astronomy.

NEO-PYTHAGOREANISM

The influence of Pythagorean thought on the history of philosophy and religion has been exercised largely through the medium of Plato, who enthusiastically adopted its main doctrines of the immortality of the soul, philosophy as an assimilation to the divine, and the mathematical basis of the cosmos. Later antiquity regarded him as a Pythagorean source, so that post-Platonic writings are of little help in distinguishing Pythagorean from original Platonic material in the dialogues. The Neo-Pythagorean movement, which started in the first century BCE, was an amalgam of early Pythagorean material with the teachings of Plato, the Peripatetics, and the Stoics. All of this material was credited to Pythagoras, who was revered as the revealer of esoteric religious truths. The interests of Neo-Pythagoreanism were religious and, in accordance with the prevailing tendencies of the time, it emphasized the mystical and superstitious sides of the earlier doctrine, its astral theology and number-mysticism, to the detriment of philosophical thinking. It cannot be called a system, but rather is a trend that in different forms continued until the rise of Neoplatonism in the third century CE, when it lost its identity in that broader and more powerful current. Besides contributing to Neoplatonism, it influenced Jewish thought through Philo of Alexandria and Christian thought through Clement of Alexandria. Prominent Neo-Pythagoreans were Cicero's acquaintance, Nigidius Figulus, and Apollonius of Tyana, a wandering mystic and ascetic of the first century CE, credited with miraculous and prophetic powers. Numenius of Apamea in the late second century was called both Pythagorean and Platonist, and was the immediate precursor of Neoplatonism.

See also Apeiron/Peras; Aristotle; Atomism; Cicero, Marcus Tullius; Clement of Alexandria; Continuity; Copernicus, Nicolas; Cosmology; Empedocles; Geometry; Leucippus and Democritus; Neoplatonism; Numenius of Apamea; Philo Judaeus; Philolaus of Croton; Plato; Platonism and the Platonic Tradition; Pre-Socratic Philosophy; Reason; Xenophanes of Colophon.

Bibliography

Delatte, A. *Études sur la littérature pythagoricienne.* Paris: E. Champion, 1915. For specialists.

Fritz, K. von. *Pythagorean Politics in South Italy: An Analysis of the Sources.* New York, 1940.

Guthrie, W. K. C. "Pythagoras and the Pythagoreans." In his *History of Greek Philosophy.* Vol. I, 146–340. Cambridge, U.K.: Cambridge University Press, 1962. References to much of the literature of the subject will be found in this general account.

Minar, E. L., Jr. *Early Pythagorean Politics.* Baltimore: Waverly Press, 1942.

Morrison, J. S. "Pythagoras of Samos." *Classical Quarterly,* n.s., 6 (1956): 135–156.

Thesleff, H. *An Introduction to the Pythagorean Writings of the Hellenistic Period.* Turku, Finland, 1961. For specialists.

Timpanaro-Cardini, Maria. *Pitagorici: testimonianze e frammenti.* 2 vols. Florence, 1958 and 1962. Texts from *Fragmente der Vorsokratiker,* edited by H. Diels and W. Kranz, translated into Italian, with introduction and commentary.

Van der Waerden, B. L. "Die Arithmetik der Pythagoreer." *Mathematische Annalen* 120 (1948): 127–153 and 676–700. For specialists.

Van der Waerden, B. L. *Die Astronomie der Pythagoreer.* Amsterdam: North-Holland, 1951. For specialists. Both this and the other van der Waerden work do not require a knowledge of Greek; texts are given in German.

W. K. C. Guthrie (1967)

PYTHAGORAS AND PYTHAGOREANISM [ADDENDUM 1]

Scholarship on Pythagoras and early Pythagoreanism has undergone a revolutionary change in recent decades. On the one hand, we know much less about Pythagoras and the early school than seemed to be the case a generation ago. On the other hand, it is no longer true that, as W. K. C. Guthrie writes in the original article above, "there is no Pythagorean literature before Plato." Both changes are due to the work of Walter Burkert (1962/1972).

THE NEW SKEPTICISM ABOUT EARLY PYTHAGOREAN PHILOSOPHY

There had always been skeptics who doubted the traditional view of scientific work by Pythagoras and his early followers. Burkert showed decisively how far this tradition derived from a completely unhistorical view of

Pythagoras created in Plato's Academy and popularized by Plato's immediate successors. The striking similarities between Plato's work and the traditional account of Pythagorean philosophy (as given in Guthrie's article) are largely due to this post-Platonic tradition, which in later versions regularly credited Pythagoras with the invention of Platonic philosophy. At the same time, Burkert defended the authenticity of most of the fragments attributed to Philolaus (in the middle or late fifth century BCE), which are now generally recognized as the earliest Pythagorean texts.

Except for the fragments of Philolaus and a single reference in Plato (reporting that the Pythagoreans regarded music and astronomy as "sister sciences" [*Republic* 530d]), there is no account of Pythagorean philosophy before Aristotle. Furthermore, there is no evidence of any Pythagorean writing before Philolaus. So Aristotle, writing a century and a half after Pythagoras's death, was entirely dependent on an oral tradition for information about the teachings of the early pre-Philolaus school. Since Pythagoras became a figure of legend almost in his own lifetime, it is extremely difficult to know how much of Aristotle's account can be traced back to the founder or to his early followers. We do have some early references to Pythagoras, notably by Heraclitus, but these references are hostile and open to diverse interpretations.

PYTHAGOREAN PHILOSOPHY DOWN TO THE TIME OF PLATO

There is no fully reliable account of early Pythagorean doctrine. All we know for certain is that he advocated some version of reincarnation, since Xenophanes, a contemporary, makes fun of Pythagoras for the belief that a human psych ē could be reborn in an animal body. Good sources report that Pythagoras founded a cult society or sect, with special dietary restrictions (ultimately including vegetarianism), whose members played an important political role in the cities of southern Italy for several generations. Burkert regarded Pythagoras as essentially a religious teacher, a charismatic guru who founded a ritual community without scientific or philosophic content. Philolaus in the late fifth century then appears as the first Pythagorean philosopher. This interpretation has been followed by many scholars, including the influential study *Philolaus of Croton* by Carl Huffman (1993). On the other hand, the surviving fragments of Philolaus do not show him to be a profoundly original thinker, and some scholars (including the present author) would regard Philolaus as formulating, and perhaps updating, an older Pythagorean world view.

The references by Heraclitus to Pythagoras as a polymath who "pursued inquiry further than anyone else" suggest that the more archaic features of this world view go back to the founder himself. As a native of Samos, Pythagoras may well have absorbed the new naturalistic cosmology being worked out in neighboring Miletus during his lifetime. The musical elements in the Pythagorean scheme would then be the personal contribution of Pythagoras himself. It is probably from this school that Plato has taken the notion of the music of the spheres, the cosmic harmony produced by the movement of the heavenly bodies. The numerical proportions corresponding to the musical consonances (2:1, 3:2, 4:3) are embedded in the Sacred Tetractys, the number 10 as the sum of 1+2+3+4, said to be the source of natural order and to provide the oath "by which the Pythagoreans swear." (Diels-Kranz, Fragmente der Vorsokratiker 58.B 15) Teachings of this kind may well go back to the first generation in the sixth century. The reports concerning Hippasus of Metapontum (probably dated in the early fifth century) indicate that an interest in the harmonic mean is older than Philolaus. How much work in mathematics or astronomy can be attributed to the early Pythagoreans is another question. But at least one contemporary scholar (Leonid Zhmud 1997) has argued in favor of the traditional view of Pythagoras as founder of a scientific school.

Aristotle reports that the Pythagoreans were pioneers in mathematics and interpreted the whole universe in terms of numbers. However, except for vague reports concerning Hippasus, we know nothing of Pythagorean mathematics in the early period. (The first author to refer to the "Pythagorean theorem" is Plutarch, in the Roman period.) Documentation for Pythagorean teaching begins with Philolaus. He claims that the *kosmos*, or world order, is composed of two opposing principles that are harmoniously fitted together, the principle of Limit or Limiting (*perainonta*) and the Unlimited (*apeiron*). The world is knowable because it is structured by number, and the numbers of special interest are the musical ratios: 2:1, 3:2, 4:3. Aristotle's report of a Pythagorean cosmology in which the earth is a planet like the sun, circulating around a central Fire, seems to be derived from Philolaus. The two Philolaic principles, the Limit and the Unlimited, show up in a famous passage in Plato's *Philebus*, Plato's dialogue, where they are said to have been tossed down from heaven, together with fire, by a certain Prometheus (*Philebus* 16c). Many readers have found it natural to

identify this Prometheus with Pythagoras. This passage may be one of the sources for the ancient story that Plato borrowed his philosophy from Pythagoras.

There is a kernel of truth in the myth that Plato's philosophy is Pythagorean. A genuinely Pythagorean view of the soul as transcending its existence in a human body serves as point of departure for Plato's *Phaedo*. A similar view of the soul is presupposed by the doctrine of recollection in the *Meno* and by the myths of judgment and preexistence in the *Phaedo, Republic*, and *Phaedrus*. Furthermore, the other typically Pythagorean view, the mathematical interpretation of nature and the conception of the cosmos in terms of musical harmony, finds its fullest expression in Plato's *Timaeus*. These two dialogues, the *Phaedo* and the *Timaeus*, form the channel through which Pythagorean ideas have passed into the mainstream of ancient and modern thought.

It is important, however, not to exaggerate Plato's debt to the Pythagorean tradition. In most cases, for instance in the judgment myths, Plato has probably transformed Pythagorean material beyond recognition. The pre-Platonic version would appear crude and primitive by comparison. This is perhaps clearest in the case of recollection. Plato's doctrine takes for granted the Pythagorean view of a cycle of reincarnation for the human psyche. But in the original Pythagorean version, recollection would refer only to Pythagoras' alleged ability to recall his previous incarnations, or perhaps to the soul's need to remember certain ritual instructions for correct behavior in the next world—the need to preserve a memory after death that avoids or survives the drink from the River Lethe (forgetfulness). The notion of recollection as a theory in epistemology, as a priori knowledge preceding sensory experience, is entirely Plato's invention. There is no Pythagorean epistemology, and nothing corresponding to the doctrine of Forms. Plato's theory of recollection represents an allegorical reinterpretation of Pythagorean themes that are originally magical or mythical.

In the case of mathematics, however, Pythagorean influence on Plato may be more substantial. The leading Pythagorean of Plato's day, Archytas of Tarentum, was both Plato's friend and also a great mathematician. Of course, not all the mathematics in Plato's dialogues needs to be derived from Pythagorean sources. (Theaetetus, for example, was not a Pythagorean, nor was his teacher Theodorus.) But Plato does cite the Pythagoreans for their view of the relation between music and astronomy (*Republic* 530d, in what is apparently a quotation from Archytas). Unfortunately, we are very poorly informed on the details of Archytas's work in astronomy and applied mathematics. Hence we cannot tell to what extent his thought inspired Plato's use of geometrical figures and numerical proportions in the cosmology of the *Timaeus*. The role of musical harmonies in the creation of the world soul is at least Pythagorean in spirit. It was the cosmology of the *Timaeus* that became the model for a Pythagorean world view in later centuries, down to the time of Johannes Kepler.

PYTHAGOREANISM AFTER PLATO

Aristotle reports that among Plato's "unwritten doctrines" was a theory of two fundamental principles: the One and the Indeterminate Dyad. These principles are regularly attributed to Pythagoras in the post-Aristotelian doxography, together with a mathematical cosmology based on the *Timaeus*. This grandiose and completely unhistorical picture of Pythagorean philosophy, which was accepted throughout antiquity, seems to have been created by Plato's disciples in the Academy and, in particular, by Speusippus, Plato's successor as head of the school. Speusippus composed a book *On Pythagorean Numbers*, which is largely devoted to the cosmological implications of the number 10. This number generates all things by containing as parts the numbers 1, 2, 3, and 4, corresponding, respectively, to point, line, plane, and solid. Thus, Pythagorean numerology, the doctrine of the symbolical and allegorical significance of the numbers from 1 to 10, seems to have begun with Speusippus.

The personal prestige of Pythagoras remained high in the Hellenistic age, although the Pythagorean school seems to have died out by the end of the fourth century BCE. What we find instead are pseudepigraphic works, treatises claiming as their author was either Pythagoras himself or, more frequently, one of his followers. The most popular author for these forged books is Archytas, but there is also a work ascribed to Timaeus of Locri, the fictitious speaker in Plato's dialogue of that name. This treatise has been preserved intact because it was falsely believed to be the Pythagorean original from which Plato derived his cosmology. In general, these pseudepigraphic works contain doctrines borrowed from Plato and Aristotle, with little or no material that is authentically Pythagorean.

A more genuine Pythagorean revival begins in the first century BCE in Rome with a famous Roman magus named Nigidius Figulus, to whom Cicero dedicated his translation of the *Timaeus*. Combining Oriental lore with Greek wisdom and the gift of second sight, Nigidius seems to have created a kind of Pythagorean society

within the Roman aristocracy. The distinguished scholar Marcus Terentius Varro was attracted by this archaic cult, and Cicero himself showed great respect for the Pythagorean tradition. Pythagoras had always been popular in Rome. At the beginning of the third century BCE, in response to an oracle during the Samnite war, a statue of Pythagoras was erected in the Roman forum. The popularity of Pythagoras in Rome was due in part to the fact that the original Pythagorean community had been located in the Greek cities of southern Italy, and hence Pythagoreanism was known since Aristotle as the *Italian philosophy*. Since most philosophers were from Greece proper, the Romans were pleased to have their own sages from Croton, Metapontum, and Tarentum.

At the same time that Nigidius was reviving Pythagoreanism in Rome in the first century BCE, Eudorus of Alexandria was inaugurating a new line of Platonic philosophers who have come to be called the Neopythagoreans. The name reflects the fact that these philosophers share the view that Plato's philosophy was derived from Pythagoras. Thus, Eudorus reports that "Socrates and Plato agree with Pythagoras that the goal of life (the *telos*) is becoming like god. But Plato articulated it more clearly by adding 'as far as possible.'" (Plato, *Theaetetus* 176b.) In reaction against the tradition of Skepticism that prevailed in the Hellenistic Academy, these philosophers emphasize the metaphysical and theological elements in Plato's philosophy. In this respect, and in the central importance attributed to numbers, the Neopythagoreans return to a position like that of Speusippus and Xenocrates in the Old Academy.

According to Eudorus, the Pythagoreans regard the One as the first principle of all things and the supreme god, but immediately below it are the two opposed principles, the One and the Indefinite Dyad. This notion of a hierarchical system of transcendental principles was developed in a new version of Pythagorean philosophy by Moderatus of Gades in the first century CE. Here there are three levels of nonsensible reality represented by three Ones. (The three Ones are related to the first three hypotheses in Plato's *Parmenides*.) Other Platonists in this tradition refer to a doctrine of three gods as distinctively Pythagorean. The best-preserved view is that of Numenius of Apamea in the second century CE. The first god is pure *nous*, an intellect focussed only on itself, like Aristotle's Prime Mover; the second god is *nous* as the demiurge, responsible for ordering the material universe; the third god is either the visible cosmos or its animating principle, the world soul. There is a significant parallel between this tripartite scheme and the three hypostases of Plotinus (the One, the Intellect, and the Soul), and it is not surprising to learn that Plotinus was accused of borrowing his philosophy from Numenius.

Pythagorean influence continued into late antiquity and the middle ages, both as numerology and as integrated into Neoplatonism, above all in the work of Iamblichus (c. 300 CE), who composed a book *On the Pythagorean Way of Life* as an introduction to his major work *On the Pythagorean School*. Pythagorean harmonics through the influence of Boethius) continued to play a role in music down to the modern age. Finally, in the Renaissance, Pythagorean ideas were revived with the new access to Plato and the Neoplatonists. This leads, on the one hand, to a flowering of occult numerology and theosophy—for example, in the cosmology of Robert Fludd—and on the other hand, to scientific applications of Pythagorean thought in the work of Copernicus and Kepler. It was Kepler who made the last great scientific contribution to the Pythagorean tradition. Taking as his model the *Timaeus* and Ptolemy's *Harmonica*, Kepler published his laws of planetary motion in a work titled *Harmonice Mundi* (The harmonics of the universe), in which he undertook to show how the movements of the planets were designed to illustrate the Pythagorean music of the spheres. Kepler's work brings the story of scientific Pythagoreanism to a happy conclusion. In one sense, the spirit of Pythagoreanism is still alive today in the mathematical interpretation of nature; string theory may be the latest version of the harmony of the spheres.

See also Archytas of Tarentum; Aristotle; Boethius, Anicius Manlius Severinus; Cicero, Marcus Tullius; Copernicus, Nicolas; Epistemology; Fludd, Robert; Hellenistic Thought; Heraclitus of Ephesus; Iamblichus; Kepler, Johannes; Numenius of Apamea; Philolaus of Croton; Philosophy of Science, History of; Plato; Plotinus; Plutarch of Chaeronea; Renaissance; Xenophanes of Colophon.

Bibliography

Burkert, W. *Weisheit und Wissenschaft: Studien zu Pythagoras, Philolaos, und Platon*. Nürnberg: H. Carl, 1962. Revised translation by Edwin L. Minar, *Lore and Science in Ancient Pythagoreanism*. Cambridge, MA: Harvard University Press, 1972.

Dillon, J. *The Middle Platonists*. Ithaca, NY: Cornell University Press, 1977.

Huffman, Carl A. *Archytas of Tarentum: Pythagorean, Philosopher, and Mathematician King*. Cambridge, U.K.: Cambridge University Press, 2005.

Huffman, Carl A. *Philolaus of Croton: Pythagorean and Presocratic*. Cambridge, U.K.: Cambridge University Press, 1993.

Huffman, Carl A. "The Pythagorean Tradition." In *Cambridge Companion to Early Greek Philosophy"*, edited by A. A. Long. Cambridge, U.K.: Cambridge University Press, 1999, pp. 66–87.

Hussey, E. "Pythagoreans and Eleatics." In *Routledge History of Philosophy*, Vol. 1: *From the Beginnings to Plato*, edited by C. C. W. Taylor. London: Routledge, 1997, pp. 128–174.

Kahn, C. H. "Pythagorean Philosophy Before Plato." In *The Presocratics. A Collection of Critical Essays*, edited by Alexander P. D. Mourelatos. Princeton, NJ: Princeton University Press, 1993, pp. 161–185.

Kahn, C. H. *Pythagoras and the Pythagoreans*. Indianapolis, IN: Hackett, 2001.

Riedweg, C. *Pythagoras. Leben, Lehre, Nachwirkung*. Munich: C.H. Beck, 2002.

Zhmud, L. *Wissenschaft, Philosophie und Religion im frühen Pythagoreismus*. Berlin: Akademie Verlag, 1997.

Charles H. Kahn (2005)

PYTHAGORAS AND PYTHAGOREANISM [ADDENDUM 2]

Ideas of Pythagoras and his school (including Philolaos) became known to the Islamic and to a lesser degree to the Jewish world since the end of the ninth century. Doxographical information about them can be found in Arabic translations of Aristotle, Plato, and above all two doxographical sources: the *Placita philosophorum*, which is attributed to Plutarch and is assumed to be compiled by Aetius Arabus (Daiber 1980), and a doxography that is attributed to Ammonius and is available only in an Arabic version (Rudolph 1989), much like the Arabic translation of the *Placita* apparently from the second half of the ninth century.

The impact of these sources, especially of Aetius, on Islamic thought (Rosenthal 1965, Daiber 1980), p. 337f.), on the Islamic philosopher al-Kindī, who died in 866 (Baffioni 1985), on the anonymous encyclopaedia of the Sincere Brethren from the tenth century (Netton 1991), and on the Jewish philosopher Saʿadia ben Joseph in the first half of the tenth century (Efros), was concentrated on the Pythagorean doctrine of numbers, especially of the number four as source of the cosmos and its harmony, and also applied to music. Shahrastani's exposition of the Pythagorean doctrine (Baffioni 1983), pp. 96ff.), which mainly combined the reports of Aetius and in Arabic doxography attributed to Ammonius (Rudolph 1989), shaped the picture of Pythagoras among Islamic thinkers.

Moreover, Neo-Pythagorean texts on ethics contributed to the propagation of Pythagorean thought in Islamic and Jewish circles. Here, an important role was played by the Pythagorean *Carmina aurea* on ethical principles of life such as piety, modesty, justice, and self-examination as ways of the soul's assimilation to God. This text was known to the Arabs in an anonymous Arabic translation from the second half of the ninth century, which was integrated in Ḥunayn ibn Ishāq's *Nawādir al-falāsifa* (Anecdotes of the Philosophers), a collection of wise sayings that was often used by Muslim authors (Baffioni 1994, Miskawayh 1964), and that in the adaptation of Muhammad ibn ʿAlī al-Ansārī was translated into Hebrew (Daiber 1995).

Originally, the *Carmina aurea* were translated into Arabic with the commentary by Iamblichus (250–330 CE), a pupil of the neoplatonic philosopher Porphyry. This commentary, which is lost in its Greek original and preserved in Arabic (Daiber 1995), differs from that attributed to Proclus, which in a similar manner offers Neo-Pythagorean traditions in neoplatonic shape and which is preserved in a redaction by Abū l-Faradj ibn al-Tayyib from the eleventh century (Linley 1984; cf. Daiber, *Islam* 65 1988, 134–137). Iamblichus's commentary continues the discussion of his *De vita pythagorica* and *Protrepticus* and amalgamates Pythagorean, Platonic-neoplatonic, Aristotelian, and Stoic ethics. It found an echo in al-Kindī, who in his "Summary on the Soul According to Aristotle, Plato, and Other Philosophers" describes the ascent and return of the soul to its divine origin through purification and increasing knowledge of God—a doctrine that is developed a century later in the encyclopaedia of the Sincere Brethren (Baffioni 1992) and is aluded to in Ibn Sīnā's (Avicenna) (d. 1037) alleged Pythagoreanism (Chaix-Ruy 1959).

Iamblichus's neoplatonic tradition of the *vita pythagorica* is reflected in a treatise attributed to Plato, "The Exhortation concerning the Education of Young Men," which is preserved only in Arabic (Rosenthal 1941, pp. 383ff.). It can be traced back to his teacher, Porphyry, who in his *History of Philosophy* (lost in Greek and preserved in some Arabic fragments) had included the biography of Pythagoras (Rosenthal 1990). It seems plausible that the same neoplatonic tradition of the *vita pythagorica* also affected the alleged letter by Pythagoras to Hiero, the tyrant of Sicily, which is available in a clumsy ninth-century translation (Rosenthal 1975). Finally, Neo-Pythgorean ethics is mirrored in the numerous sayings

attributed to Pythagoras and transmitted in Syriac and Arabic gnomologia (Gildemeister 1870, Levi della Vida 1910, Gutas 1975).

Bibliography

Baffioni, Carmela. "'Detti aurei' di Pitagora in trasmissione araba." *I moderni ausili all'Ecdotica. Atti del Convegno Internazionale di Studi* (Fisciano–Vietri sul Mare–Napoli, 27–31 ottobre 1990). A cura di Vincenzo Placella e Sebastiano Martelli. Napoli. 1994 (= Pubblizacioni dell'Universita degli stud di Salerno. Sezione atti, convegni, miscellanee. 39): 107–131

Baffioni, Carmela. "Platone, Aristotele e it pitagorismo kindiano." *Annali del'Istituto Universitario Orientale di Napoli. Pubblicazioni* (Napoli) 45 (1985): 135–144.

Baffioni, Carmela. "Traces of 'Secret Sects' in the Rasaʾil of the Ikhwan al-Safaʾ." In *Shiʿa Islam, Sects, and Sufism*: *Historical dimension, religious practice and methodological considerations.* Edited by F. DeJong. Utrecht: M. Th. Houtsma Stichting, 1992, 10–25.

Baffioni, Carmela. "Una 'storia della filosofia greca' nell'islam del XII secolo: III. Pitagora." *Elenchos* (Napoli) 4 (1983): 93–132, 261–345.

Chaix-Ruy, J. "Du pythagorisme d'Avicenne au soufisme d'Al-Ghazali." *Revue de la mediterranneé* (Alger) 19 (1959): 289–327.

Daiber, Hans. *Aetius Arabus: Die Vorsokratiker in arabischer Überlieferung*. Wiesbaden, Germany: Steiner, 1980.

Daiber, Hans. *Neuplatonische Pythagorica in arabischem Gewande: Der Kommentar des Iamblichus zu den Carmina aurea*. New York: North-Holland, 1995.

Efros, Israel Isaac. "Saadyah's Second Theory of Creation in Its Relation to Pythagoreanism and Platonism." *Louis Ginzberg Jubilee Volume* English Section (1945–1946): 133–142.

Gildemeister, J. "Pythagorassprüche in syrischer Überlieferung." *Hermes* 4 (1870): 81–98.

Gutas, Dimitri. *Greek Wisdom Literature in Arabic Translation: A Study of the Graeco-Arabic Gnomologia*. New Haven, CT: American Oriental Society, 1975.

Levi della Vida, Giorgio. "Sentenze pitagoriche in versione siriaca." *Rivista degli studi orientali* 3 (1910): 595–610.

Linley, Neil, ed. and trans. *Ibn at-Tayyib: Proclus' Commentary on the Pythagorean Golden Verses*. Buffalo, NY: Department of Classics, State University of New York, 1984.

Miskawayh. *An Unpublished Treatise of Miskawayh on Justice or Risala fi mahiyyat al-ʿAdl ʿi Miskawayh*, edited and translated by M. S. Khan. Leiden, Netherlands: E. J. Brill, 1964.

Netton, Ian Richard: *Muslim Neoplatonists: An Introduction to the Thought of the Brethren of Purity (Ikhwān al-Safāʾ)*. Edinburgh: Edinburgh University Press, 1991.

Rosenthal, Franz. "Arabischc Nachrichten über Zenon den Eleaten." In *Greek Philosophy in the Arab World: A Collection of Essays*. Aldershot, U.K.: Grower Publishing, 1990. Rosenthal's essay originally published in the journal *Orientalia* N.S. 6 (1937): 27–67.

Rosenthal, Franz. "Fithaghuras." In *Encyclopaedia of Islam 12*, 929–930. Leiden, Netherlands: E. J. Brill, 1965.

Rosenthal, Franz. "From Arabic Books and Manuscripts, XIII–XIV." *Journal of the American Oriental Society* 95 (2) (1975): 209–213.

Rosenthal, Franz. "Some Pythagorean Documents Transmitted in Arabic." *Orientalia* 10 (1941): 104–115, 383–395.

Rudolph, Ulrich. *Die Doxographie des Pseudo-Ammonios: Ein Beitrag zur neuplatonischen Uberlieferung im Islam*. Stuttgart, Germany: Steiner, 1989.

Hans Daiber (2005)

Q

QUALIA

The word *quale* (or *qualia*) derives from the Latin for "quality." As used by C. I. Lewis (1929) and those following him, it refers to the qualities of phenomenal individuals, such as color patches, tastes, and sounds. In this sense the term means what George Berkeley meant by "sensible qualities," or what later philosophers meant by *sensa* or *sense data*. Since the demise of sense data theories, the term *qualia* has come to refer to the qualitative, or phenomenal, character of conscious, sensory states, so that it is mental states, not phenomenal individuals, that are the subjects of predication. Another expression for this aspect of mental life is the "raw feel" of experience, or "what it's like" to have certain sensory experiences. Qualia are part of the phenomenon of the subjectivity of consciousness, and pose one of the most difficult problems for a materialist solution to the mind-body problem.

IDENTITY THEORY

J. J. C. Smart posed the challenge this way in a 1959 article: Consider a sensation like a yellowy-orange after-image. According to the materialist theory known as the "central state identity theory" (or just "identity theory"), the sensation is a brain state. Smart's worry, which he attributed to Max Black, was that even if one accepted that the sensation was itself a brain state, it still seemed as if one had to attribute an "irreducibly psychic" property to the brain state. That is, there is a distinctive qualitative character experienced when having a yellowy-orange after-image, and that property—that yellowy-orange character—does not seem at all like a physical property. So even if all mental states are brain states, we might still be driven to the view that some mental properties—qualia, in particular—are not physical. This would constitute a form of dualism known as "property dualism," a position inconsistent with materialism.

The Max Black objection presented by Smart in 1959 is related closely to the "conceivability argument," a dualist argument going back to René Descartes, and revived in 1980 by Saul Kripke and in 1996 by David Chalmers. Roughly, the idea is this. When one considers simultaneously what it is like to see a yellowy-orange after-image and any description of the firing pattern of an assembly of neurons, it seems perfectly coherent to imagine having the one without the other. That neurons should fire in this or that pattern and that it should be like nothing at all for the subject whose neurons they are seems clearly possible. Yet, if qualia are identical to neural properties, such a situation is not possible. Hence, qualia must be not

neural properties, but nonphysical properties possessed by neural states.

Another closely related dualist argument is Frank Jackson's 1982 "knowledge argument." We are asked to imagine a scientist who knows everything about the physiology of color vision, but who has never seen anything in color. Upon first seeing a red rose, it seems clear that she would learn something new—what it is like to see red. Yet, if qualia are just physical properties of the nervous system, she should have already known what it is like to see red. Hence, Jackson concludes, qualia are not physical properties.

Many materialist philosophers object that these dualist arguments rely on an assimilation of concepts and properties. Concepts are elements of thought, ways of thinking of objects and properties, comparable to words in a language. Just as there can be many distinct words referring to the same object or property, so too there can be distinct concepts that apply to the same property. All the above arguments demonstrate, according to these philosophers, is that we have different ways of conceiving of qualia, and that it isn't obvious that they pick out the same properties. But just as the fact that we had to learn that water is identical to H_2O does not impugn the claim that they are identical, so too the fact that we have to learn that a certain quale is a certain neural property does not refute the claim that they are indeed the very same property.

EXPLANATORY GAP

Proponents of property dualism respond that there are important differences between the water–H_2O case and the case of qualia that undermine the analogy pushed by materialists (Chalmers 1996). However, even if the conceivability and knowledge arguments do not demonstrate that qualia are, as Smart put it, "irreducibly psychical," they do point toward another problem, one that goes under the name of the "explanatory gap" (Levine 1983, 2001). The problem is this. If qualitative sensory experiences are really nothing over and above the interplay of neural firing in the relevant part of the brain, then one would expect that the qualitative character of particular types of sensations could be explained and predicted by reference to their neurophysiological embodiments. Yet, when we consider what it is like to see a red rose or a yellowy-orange after-image, it seems completely arbitrary that it should be the result of this type of neural firing as opposed to some other. In fact, it seems totally arbitrary that it should be like anything at all, merely from a knowledge of the neural properties. In this sense there seems to be an explanatory gap between the underlying level of neurophysiological phenomena and the level of qualitative experience. Thomas Nagel (1974) makes a similar argument about the limits of materialist understanding by noting that as much as we learn about the echolocation sense of bats, we can never learn thereby what it is like to be a bat and to sense the world in this way.

Faced with these strong intuitions that there is something suspect about the connection between physical properties of the nervous system and qualia, materialists have adopted two different strategies. The first is to attempt to straightforwardly dispel these anti-materialist intuitions by coming up with materialist theories of qualia that are intuitively acceptable. The second is to grant the apparent mystery involved in the connection between qualia and neurophysiological properties, but to argue that there are reasons why this connection should appear so mysterious that do not in the end contradict the basic tenets of materialism.

In line with the first strategy, Smart (1959) himself addressed the problem by proposing what he called a "topic-neutral" analysis of qualitative character. His claim was that our notion of qualitative character is neutral with respect to the kind of material in which it is embodied. Rather, to have a sensation of a yellowy-orange after-image, for instance, is to occupy a state that is similar to the state one is in when actually seeing an orange. This idea was then later developed by functionalists such as H. Putnam (1991), who identified mental states of all kinds with causal roles. That is, a particular mental state is defined as a state that is caused in certain characteristic ways (by physical stimuli and other mental states) and has certain characteristic effects (particular forms of behavior as well as other mental states). On this view the connection between a particular qualitative sensory state and a brain state is truly contingent, since it is allowed that any other physical state that filled the same causal role would count as an instance of this qualitative state.

Adopting functionalism for qualia might seem to provide the materialist with a response both to the conceivability argument and to the problem of the explanatory gap. It is conceivable that a creature might experience a certain sensory quality without being in a particular physical state since there are many different physical states that can support the relevant causal role. Also, one might occupy a certain physical state without having the sensory experience because that state is not connected in the right way to other states, and therefore is not playing the appropriate causal role. As for the explanatory gap, the idea is that appeal to the intrinsic

physical properties of a brain state don't explain its mental, or qualitative character, because qualitative character is a function of the relations that physical state maintains with other internal physical states, as well as stimuli and behavior. The proper locus of explanation for the qualitative character of experience is the overall pattern of interactions among the subject's internal states; it is a matter of the structure, not the "stuff" in which the structure is embodied.

However, it turns out that almost the very same problems that attended the identity theory return to haunt functionalism as well. Take the conceivability argument. In the form of the "inverted qualia" and "absent qualia" hypotheses the conceivability argument can be mounted against functionalism as well. The inverted qualia hypothesis is the conjecture that there could be two functionally identical creatures—that is, both creatures, though made of different material, possess a set of internal states that maintain the very same pattern of interactions with each other and the relevant inputs and outputs to the system—that experience very different qualia when occupying the very same functional state.

The standard illustration of this possibility is known as the "inverted spectrum hypothesis." Oversimplifying greatly for now, consider the fact that the color wheel can be inverted in such a way as to maintain all of the similarity relations. That is, if one creature sees blue and green where another sees yellow and red, and vice versa, then all of their judgments about the relative similarities of objects with respect to color would converge. Imagine that this inversion occurred at birth, so they learned to use color terms the same way. Jack and Jill might both call a ripe tomato red, though Jack's experience is qualitatively like what Jill would experience were she looking at a ripe cucumber.

If such an inversion of qualia with respect to functional roles is possible, then the qualitative character of a sensory experience cannot be identified with its functional role. To make matters worse, it seems perfectly coherent to imagine a creature that satisfies the relevant functional description, and yet for whom there is no conscious experience occurring at all. (Often such creatures are known as "zombies" in the literature.) Ned Block (1980) describes a very compelling example. Imagine, he says, the entire nation of China connected by phone lines in such a way that, collectively, they satisfy the same functional description as a human brain. Would we want to say that the entire nation of China, as a single subject, is seeing red, or feeling pain? Certainly it seems at least possible that no genuine experience is going on at all. Hence having a qualitative sensation cannot be merely a matter of possessing internal states that play a certain functional, or causal role.

THE DILEMMA

The objections to both the identity theory and functionalism reveal a deep dilemma for materialists about qualia. The qualitative character of a sensation—the way color looks, the way pain feels—strongly seems to be an intrinsic property of the sensation, a matter of how things are with one at that moment, not a matter of how one is disposed to act or what effects are likely. In this sense the identity theory seems quite appropriate, since it identifies the qualitative character of an experience with a physical property of the brain state one occupies at that moment. The problem is that there seems to be no intelligible connection between the physical properties of brain states and the qualitative properties of sensory experiences.

If, however, we pin qualitative character on the pattern of relations that a sensory state maintains with other states, as well as stimuli and behavior, then we can see how appeal to the physical properties of brain states could play an important explanatory role. We can explain how it is that the neural state one occupies when, say, experiencing a yellowy-orange after-image, interacts with other neural states and stimuli and behavior so as to realize the relevant pattern by appeal to the causal mechanisms of the brain and nervous system. The only problem here is that, as demonstrated by the inverted and absent qualia hypotheses, qualitative character is not convincingly characterizable as a matter of the pattern of interactions among internal states. Thus the materialist is faced with this dilemma: Qualitative character is explicable in physical terms only if it can be characterized as a pattern of causal relations among mental states, but only a theory of qualitative character that treats it as an intrinsic property of mental states will be intuitively acceptable.

Functionalism is a structural theory of qualitative character—a particular quale is identified with a particular niche in the overall system of causal interactions among stimuli, internal states, and behavior. Another structural theory worthy of mention is what we might call the "quality space" theory, proposed by Austen Clark in 1993. On this view we start with the idea that different sensory modes—vision, hearing, etc.—define quality spaces. A quality space is a multidimensional space whose axes are determined by the number of independent parameters along which sensory experiences in a given mode can vary. To take again an admittedly oversimplified example, consider color vision. Colors vary along

three dimensions: hue, brightness, and saturation. A particular color (where this means a determinate shade) can then be identified with a vector representing its values in each of the three dimensions. Assuming colors only vary in these three ways (which is part of the oversimplification), then a person's similarity judgments about colors can be predicted and explained by the distances among the relevant color vectors. A complete map of a sensory system can be drawn once all of the independent parameters of variation have been determined. Qualia, then, are points in quality spaces.

The quality space view differs from functionalism in that the structure by reference to which a quale is defined is not a pattern of causal interactions, but rather a quality space. However, it shares with functionalism the idea that it is structural relations rather than intrinsic features of the experience that determine qualitative character. Also, like functionalism, on the quality space view the appeal to the physical features of neural states comes in to explain how the relevant structure is embodied. This allows for the possibility of many alternative physical embodiments for the same quality space, so long as the overall structure of relations among the elements is preserved. Unfortunately, also like functionalism, the quality space view is subject as well to the problems of the inverted and absent qualia hypotheses. So with respect to the general dilemma facing the materialist it does not improve on functionalism.

REPRESENTATIONALISM

A view that seems to promise a way to overcome the materialist's dilemma is "representationalism," discussed in the work of Fred Dretske (1995), G. Harman (1990), William Lycan (1987), and Michael Tye (1995). One way to motivate the theory is to start from an untenable but nevertheless quite tempting theory of qualia and then see representationalism as a way to capture the spirit of the original view while removing its fatal weakness. The tempting but untenable view is this. Qualia, rather than features of mental states, are properties of external objects. They are the colors, sounds, and textures out there in the world that our senses detect. This view is tempting for two reasons. First, it removes qualia as obstacles to a materialist solution to the mind-body problem, since qualia are no longer features of mental states. Second, it is intuitively plausible. Advocates of the view often defend it by citing the so-called transparency of sensory experience. If asked to describe what it is like to have various sensory experiences, one finds oneself describing the properties of external objects. One says things like "it looks like a lemon," "it tastes like chicken," or "it feels smooth."

The reason the view is untenable is that it cannot handle cases of hallucination or illusion. Suppose one "sees" a pink elephant where there is nothing remotely pink or elephant-like. Clearly one is having a sensory experience with a "pinkish" qualitative character, yet there is no object out there in the world that is pink. Hence the quale cannot be the pink of the elephant, it has to be a feature of one's experience, a property of one's internal mental state. Representationalism comes into play at this point. According to this view, sensory states are mental states that represent the way the world is around us. They differ from belief states in being nonconceptual, more picturelike, but they share with beliefs and thoughts the feature of representing the world. Qualia, then, are the representational contents of sensory experiences. That is, to have a "pinkish" qualitative character is for one's visual state to have the content that something out there in the world is pink. Notice that representationalism shares with the original view the core idea that pinkness is primarily a feature of external objects, but it nevertheless accommodates hallucination and illusion. Just as one can think that there are pink elephants even though there aren't any, so too one can have visual experiences that represent pink elephants even though there aren't any.

While representationalism has many virtues, there are two primary problems. First, the view is less plausible when applied to bodily sensations like pains and itches than when applied to colors and sounds. What does the qualitative character of an itch or a headache represent? Advocates of representationalism maintain that these sensations represent conditions of the body. Whether this view can be sustained is a matter of controversy. The main problem, however, is that representationalism does not overcome the basic challenge facing other materialist theories. Just as functionalism and quality space theory have trouble with inversion and zombie scenarios, so too does representationalism. It seems easy to imagine a creature who normally sees objectively red objects the way others see objectively green objects, and it also seems possible for there to be creatures, or devices, that meet the relevant specifications for representing the qualities of external objects in a "sensory" format but for whom there is nothing it is like to occupy these representational states. Properly programmed computers certainly seem like possible examples. Hence the principal challenge to materialist theories of qualia remains.

PHENOMENAL CONCEPTS

Some materialist philosophers dismiss the inverted and absent qualia hypotheses, along with the conceivability argument, by insisting that the intuitions that underlie these challenges are just that—intuitions—and should not be accorded much significance. Daniel Dennett (1988) and Georges Rey (1997) go so far as to embrace eliminativism, the view that qualia do not really exist. We think we have these features of experience, but in fact they represent a kind of cognitive illusion. However, other materialists, such as Brian Loar (1997), William Lycan (1987), Colin McGinn, David Papineau (2002), Scott Sturgeon (2000), and Michael Tye (1995), insist on the reality of qualia and grant the import of the intuitive resistance to materialism. Their strategy is to attempt to provide a satisfactory materialist theory of the intuitive resistance itself. For many the main tool in this endeavor is the notion of a "phenomenal concept."

Phenomenal concepts are the special concepts of qualitative properties that we employ when thinking of our qualitative states from within the first-person point of view. When one considers what it is like to see a red rose, and then says something like, "How can *that* be merely a matter of neurons firing in a certain pattern?" one is employing a phenomenal concept to think about the experience. The proposal then is to explain the stubborn cognitive resistance to materialist theories (of whatever form) by appeal to peculiar features of phenomenal concepts. It is a feature of our cognitive architecture, on this view, that we cannot come to see how the qualitative character of our experience is just a matter of the way our neurons are firing. We are doomed to suffer from an explanatory gap, but that we are so doomed is itself explicable in perfectly respectable materialist terms.

Whether this appeal to phenomenal concepts can do the work of extricating materialism from the challenges posed by qualia is still a matter of controversy. It appears that the cluster of problems comprising consciousness, qualia, and subjectivity are destined to haunt the philosophy of mind for some time to come.

See also Consciousness; Knowledge Argument.

Bibliography

Block, N., and J. A. Fodor. "What Psychological States Are Not." *The Philosophical Review* 81 (1972): 159–181.

Block, Ned. "Troubles with Functionalism." In *Readings in Philosophy of Psychology*, edited by Ned Block. vol. 1, 268–305. Cambridge, MA: Harvard University Press, 1980–1981.

Chalmers, David John. *The Conscious Mind: In Search of a Theory of Conscious Experience*. Oxford: Oxford University Press, 1996.

Churchland, Paul M. "Reduction, Qualia, and the Direct Introspection of Brain States." *Journal of Philosophy* 82 (1985): 8–28.

Clark, Austen. *Sensory Qualities*. Oxford: Oxford University Press, 1993.

Davies, Martin, and Glyn W. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell, 1993.

Dennett, Daniel C. *Consciousness Explained*. Boston: Little, Brown, 1991.

Dennett, Daniel C. "Quining Qualia." In *Consciousness in Contemporary Science*, edited by A. J. Marcel and E. Bisiach. Oxford: Oxford University Press, 1988.

Dretske, Fred. *Naturalizing the Mind*. Cambridge, MA: MIT Press, 1995.

Hardin, C. L. *Color for Philosophers: Unweaving the Rainbow*. Indianapolis: Hackett, 1988.

Harman, G. "The Intrinsic Quality of Experience." In *Philosophical Perspectives, 4*, edited by J. Tomberlin, 31–52. Atascadero, CA: Ridgeview, 1990.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Kripke, Saul A. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Levine, Joseph. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (1983): 354–361.

Levine, Joseph. *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press, 2001.

Lewis, Clarence Irving. *Mind and the World-Order: Outline of a Theory of Knowledge*. New York: Scribner's Sons, 1929.

Loar, Brian. "Phenomenal States." In *The Nature of Consciousness: Philosophical Debates*, edited by N. Block, O. Flanagan, and G. Güzeldere. Cambridge, MA: MIT Press, 1997.

Lycan, William G. *Consciousness*. Cambridge, MA: MIT Press, 1987.

McGinn, Colin. *The Problem of Consciousness: Essays toward a Resolution*. Oxford: Basil Blackwell, 1991.

Nagel, Thomas. "What Is It Like to Be a Bat?" *The Philosophical Review* 83 (4) (1974): 435–450.

Papineau, David. *Thinking about Consciousness*. Oxford: Oxford University Press, 2002.

Putnam, H. "The Nature of Mental States." In *The Nature of Mind*, edited by David M. Rosenthal. New York: Oxford University Press, 1991.

Rey, Georges. *Contemporary Philosophy of Mind: A Contentiously Classical Approach*. Cambridge, MA: Blackwell, 1997.

Shoemaker, Sydney. *Identity, Cause, and Mind: Philosophical Essays*. Cambridge, U.K.: Cambridge University Press, 1984.

Smart, J. J. C. "Sensations and Brain Processes." *The Philosophical Review* 68 (1959): 141–156.

Sturgeon, Scott. *Matters of Mind: Consciousness, Reason, and Nature*. London: Routledge, 2000.

Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press, 1995.

Joseph Levine (1996, 2005)

QUALITIES, PRIMARY AND SECONDARY

See *Primary and Secondary Qualities*

QUALITY AND QUANTITY

See *Categories*; *Dialectical Materialism*; *Primary and Secondary Qualities*

QUANTIFIERS IN FORMAL LOGIC

Familiarity with classical quantification theory is presupposed here. Some proposed amendments are considered, as are several additions.

ALTERNATIVES TO CLASSICAL QUANTIFICATION THEORY

First-order logic can be reformulated so as to avoid quantifiers and variables. This is only partially done in modal logic, which avoids explicit quantification over possible states of the world in favor of operators $\Box$ and $\Diamond$. However, in principle all quantification is avoidable, if one is willing to admit enough operators and does not worry about their having ordinary-language readings. In practice, however, few have preferred this predicate-functor approach (see Quine 1960, Benthem 1977). Thus, even such dissidents as the intuitionists adopt the classical quantificational language, though the properties they ascribe to the quantifiers are nonclassical. (Thus, while classically $\forall$ and $\neg\neg\forall$ and $\forall\neg\neg$ are equivalent, intuitionistically the first is stronger than the second and the second stronger than the third.)

Classical logic allows terms formed from constants and function symbols, subject to the restriction that each term must denote some element of the domain over which the quantifiers range; but terms are eliminable using Bertrand Russell's theory of descriptions. On the classical Tarskian definition of truth in a model, truth of $\forall x\phi(x)$ (respectively, $\exists x\phi(x)$) is equivalent to the truth of $\phi(t)$ for all (respectively, some) terms t only in special cases, as when each element of the domain is the denotation of some term of the language (which is never so if the domain is uncountable and the language countable). By contrast, the so-called substitutional quantifier $\prod$ (respectively, $\sum$) is defined by the condition that $\prod x\phi(x)$ (respectively, $\sum x\phi(x)$) always counts as true if and only if (iff) $\phi(t)$ is true for all (respectively, some) terms t. There is no technical obstacle to introducing such operators, but whether there is any philosophical advantage to doing so is controversial. In particular, if one has in mind a specific domain, $\prod$ (respectively, $\sum$) will be intuitively equivalent to the ordinary language "for every (respectively, some) element of the domain" only in special cases (see Kripke 1976). Antithetical to substitutional quantification is so-called free logic, which drops the classical restriction that all terms must have denotations and gives up the classical inferences from $\forall x\phi(x)$ to $\phi(t)$ and from $\phi(t)$ to $\forall x\phi(x)$ (see Bencivenga 1983).

EXTENSIONS OF CLASSICAL QUANTIFICATION THEORY

In contrast to the various anticlassical logics just mentioned, by far the largest body of work on quantifiers in formal logic concerns certain extraclassical logics, called model-theoretic logics. These accept classical logic and the Tarskian definition of truth in a model, but introduce additional kinds of quantifiers into the language, indicating their intended meaning by adding clauses for them to the Tarskian definition. There are several kinds (see Barwise and Feferman 1985).

CARDINALITY QUANTIFIERS. Though there are nineteenth-century and even medieval antecedents, the modern theory of such quantifiers as "most" begins with Andrzej Mostowski (1957). Given a formula $\phi(x)$ and a model with domain A, write $\phi[a]$ to indicate that $a \in A$ satisfies $\phi(x)$; also write card B for the cardinality of a set B. Then the truth conditions for the most studied Mostowski-style quantifiers are as shown in Table 1.

All these generalized quantifiers count as logical notions according to the definition of Alfred Tarski (1986) (which requires that any sentence involving a purportedly logical operator that is true in a model remains true if the model is replaced by an isomorphic one). Their theory has been worked out in some detail. For example, for first-order logic plus Q_0 the Löwenheim-Skolem theorem holds but the compactness theorem fails, while for Q_1 the opposite is the case.

PLURAL QUANTIFIERS. So-called second-order and higher-order quantifiers are nowadays generally read as first-order quantifiers, but with a different domain from that of the first-order quantifiers. Thus, one writes "$\exists X(Xy$ & …)" but reads it as something like "There is a class X such that y is a member of X and …" or "There is a concept X such that y falls under X and …" and simi-

TABLE 1

Quantifier	Truth condition
$\mathbf{Most}x\,\phi(x)$	card $\{a: \phi[a]\} >$ card $\{a: \neg\phi[a]\}$
$\mathbf{More}x[\phi(x), \psi(x)]$	card $\{a: \phi[a]\} >$ card $\{a: \psi[a]\}$
$\mathbf{Q}_0 x\phi(x)$	card $\{a: \phi[a]\}$ infinite
$\mathbf{Q}_1 x\phi(x)$	card $\{a: \phi[a]\}$ uncountable
$\mathbf{H}x\,\phi(x)$	card $\{a: \phi[a]\} =$ card A
$\mathbf{R}xy\,\phi(x)$	for some infinite $I \subseteq A$, $\psi[a, b]$ for all distinct $a, b \in I$

larly for the two-place "$\exists X(Xyz \ \& \ \ldots\ )$" and the third-order "$\exists \mathbf{X}(\mathbf{X}Y \ \& \ \ldots\ )$," with *relation* and *class of classes* in place of *class*.

George S. Boolos (1984) suggests a different reading, "There are some things, the *x*s, such that *y* is one of them." Such a reading is available only in the second-order, one-place case, but there it seems to offer a way of avoiding overt quantification over classes or concepts. But it is controversial whether such plural quantification is prior to such notions as that of class, or whether the use of the plural involves a covert "ontological commitment" to something like classes. Boolos argues against the reduction of plural to class quantification, on the grounds that "[t]here are some classes such that any class is one of them iff it is not a member of itself" is true, while "[t]here is a class of classes such that any class" is false.

GAME QUANTIFIERS. Any first-order sentence is equivalent to one in *prenex* form, with all quantifiers out front. Any first-order prenex is equivalent to an existential second-order sentence (quantifying over functions from and to the domain *A* of the first-order variables), called its Skolem form, as with this equivalent pair (where the alternation of quantifiers may go on for any finite number *n* of rounds):

(1) $$\forall x_1\exists y_1\forall x_2\exists y_2 \ldots \phi(x_1, y_1, x_2, y_2, \ldots)$$

(2) $$\exists f_1\exists f_2 \ldots \forall x_1\forall x_2 \ldots \phi(x_1, f(x_1), x_2, f(x_1, x_2), \ldots)$$

Leon Henkin (1961) observes that one can associate to (1) a game for two players: player A chooses some $a_1 \in A$, player E chooses some $b_1 \in A$, A chooses a_2, then E chooses $b_2, \ldots$, and in the end E wins if $\phi[a_1, b_1, a_2, b_2, \ldots]$, and A if not. A strategy for a player is a rule telling that player how to play on each move as a function of the opponent's previous moves. A winning strategy is one such that, if the player plays according to it, then the player will win, regardless of how the opponent plays. A strategy for E can be represented as a pair of functions, one giving E's first move as a function of A's first move, the other giving E's second move as a function of A's first two moves. Then, (2) asserts that there is a winning strategy for E.

The game interpretation is especially useful if one wants to consider infinitely long formulas. A sentence like (1) but with an infinite alternation of quantifiers can be thought of as describing an infinite game—one may imagine each move made twice as fast as the one before—and the assertion that there exists a winning strategy for E is expressible as an infinitely long second-order sentence like (2) with infinite blocks of existential second-order and universal first-order quantifiers. There is this difference, that for a finite game one or the other of the players must have a winning strategy, but not for infinite games except in special cases. One such special case is that where ϕ is a conjunction of formulas $\phi_1, \phi_2, \ldots$, each involving only finitely many of the *x*'s and *y*'s. This game quantifier has a tractable theory in this case (see Moschovakis 1972).

BRANCHING QUANTIFIERS. Henkin (1961) also introduces branching quantifiers and suggests an interpretation in terms of an associated Skolem form, illustrated by the following pair:

(3) $$\begin{matrix}\forall x_1\exists y_1 \\ \forall x_2\exists y_2\end{matrix}\ \phi(x_1, y_1, x_2, y_2)$$

(4) $$\exists f_1\exists f_2\phi(x_1, f(x_1), x_2, f_2(x_2))$$

Note the subtle difference between (4) and (2): In the latter, f_2 is a one-place function. The main result about Henkin quantifiers is the Enderton-Walkoe theorem, asserting that not only is every Henkin quantifier sentence equivalent to an existential second-order sentence but also the converse holds. This means that known results about the logic of existential second-order sentences immediately apply to the logic of Henkin quantifier sentences: the Löwenheim-Skolem theorem, the compactness theorem, the definability of truth for sentences of this class by a sentence of the class, and more.

Jaako Hintikka (1996) introduces a nonbranching notation, in which (3) would be written as follows:

(5) $$\forall x_1\exists y_1\forall x_2\exists y_2/x_1\phi(x_1,y_1,x_2,y_2)$$

The "$/x_1$" is read "independent of x_1." Hintikka, long an advocate of a game interpretation of first-order quantifi-

cation, also suggests a game interpretation of the new quantifiers, in terms of a game of imperfect information, in which at the time of E's second move, E has available only information about A's second move, not about A's first move—which is most easily imagined if one thinks of E as a team, with different members making different moves and having available different information when doing so. Hintikka calls the logic with these quantifiers independence-friendly (or information-friendly) logic and makes strong and controversial claims about the philosophical significance of theorems about existential second-order sentences when restated for "IF" logic (see Hintikka 1996; compare Tennant 1998; see also Hodges 1997; Burgess 2003).

Which quantifiers considered by logicians have natural-natural language counterparts, and how close those counterparts are, is a much discussed question that cannot be addressed in this entry.

See also Artificial and Natural Languages; First-Order Logic; Intuitionism and Intuitionistic Logic; Quantifiers in Natural Language; Types, Theory of.

Bibliography

Barwise, K. Jon, and Solomon Feferman, eds. *Model-Theoretic Logics*. Berlin, Germany: Springer-Verlag, 1985.

Bencivenga, Ermanno. "Free Logics." In *Handbook of Philosophical Logic*. Vol. 3, *Alternatives to Classical Logic*, edited by Dov M. Gabbay and F. Guenther, 73–426. Dordrecht, Netherlands: D. Reidel, 1983.

Benthem, J. F. A. K. van. "Tense Logic and Standard Logic." *Logique et Analyse* 80 (1977): 41–83.

Boolos, George S. "To Be Is to Be the Value of a Variable (or to Be Some Values of Some Variables)." *Journal of Philosophy* 81 (1984): 430–439.

Burgess, John P. "A Remark on Henkin Sentences and Their Contraries." *Notre Dame Journal of Formal Logic* 44 (2003): 185–188.

Enderton, H. B. "Finite Partially-Ordered Quantifiers." *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 16 (1970): 393–397.

Henkin, Leon. "Some Remarks on Infinitely Long Formulas." In *Infinitistic Methods: Proceedings of the Symposium on Foundations of Mathematics, Warsaw, 1959*, 167–183. Oxford, U.K.: Pergamon, 1961.

Hintikka, Jaako. *The Principles of Mathematics Revisited*. New York: Cambridge University Press, 1996.

Hodges, Wilfrid. "Compositional Semantics for a Language of Imperfect Information." *Logic Journal of the IGPL* 5 (1997): 539–563.

Kripke, Saul A. "Is There a Problem about Substitutional Quantification?" In *Essays in Semantics*, edited by Garteth Evans and John McDowell, 324–419. New York: Oxford University Press, 1976.

Moschovakis, Yannis. "The Game Quantifier." *Proceedings of the American Mathematical Society* 31 (1972): 245–250.

Mostowski, Andrzej. "On a Generalization of Quantifiers." *Fundamenta Mathematicæ* 44 (1957): 12–36.

Quine, W. V. O. "Variables Explained Away." *Proceedings of the American Philosophical Society* 104 (1960): 343–347.

Tarski, Alfred. "What Are Logical Notions?" *History and Philosophy of Logic* 7 (1986): 143–154.

Tennant, Neil. Review of *The Principles of Mathematics Revisited*, by Jaako Hintikka. *Philosophia Mathematica* 6 (1998): 90–115.

Walkoe, W. J. "Finite Partially Ordered Quantification." *Journal of Symbolic Logic* 35 (1970): 535–550.

John P. Burgess (2005)

QUANTIFIERS IN NATURAL LANGUAGE

See Appendix, Vol. 10

QUANTITY

See *Measurement and Measurement Theory*

QUANTUM COMPUTING AND TELEPORTATION

In the 1980s and 1990s a series of revolutionary developments in the foundations of quantum mechanics led to what would later become the thriving fields of quantum information, quantum computation, and quantum cryptography. The roots of this revolution lie in the debate between Albert Einstein and Niels Bohr on the interpretation of quantum mechanics, specifically in the notion of "entangled" quantum states at the heart of the Einstein-Podolsky-Rosen argument for the incompleteness of quantum mechanics. What Einstein, Podolsky, and Rosen showed in their 1935 paper "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" was that composite quantum systems, consisting of spatially separated subsystems, could exist in certain states with peculiar nonclassical correlations between the outcomes of measurements on the subsystems. They argued that these correlations are incompatible with the assumption that the quantum state is a complete description of the system.

In a two-part commentary on the paper, Schrödinger referred to these states as being "entangled." Roughly

thirty years later John Bell re-examined the Einstein-Podolsky-Rosen argument and showed that quantum mechanics could not be completed in the way Einstein would have liked, because the correlations of entangled states violate an inequality that an Einsteinian completion of quantum mechanics would have to satisfy. Essentially Bell showed that the correlations are inconsistent with any explanation in terms of a common cause (whether deterministic or stochastic) originating in the preparation of the state.

The salient feature of quantum information-processing tasks is the exploitation of entanglement as a new physical resource. Entanglement can be used to teleport quantum states, to exponentially outperform classical computers, and to implement cryptographic procedures that are impossible classically.

TELEPORTATION AS REMOTE "STEERING"

Schroödinger regarded entangled states as problematic because they allow the possibility of what he called remote "steering," which he regarded as unacceptable in a physical theory. As it turns out the teleportation of quantum states is an experimentally confirmed application of remote "steering" between two separated systems.

Consider Alice and Bob, the traditional protagonists in any two-party communication protocol. Suppose Alice and Bob each holds one of a pair of quantum particles associated with binary-valued physical quantities or "observables." An example would be a pair of spin-½ particles, with two possible values, + and –, for the spin in some direction, say the *z*-direction. Alice's particle might be represented by the pure quantum state $|+\rangle_A$ and Bob's particle by the pure quantum state $|-\rangle_B$. A spin state is represented as a unit vector in a 2-dimensional vector space, a so-called Hilbert space, the representation space for quantum states. The state of the composite two-particle system is a product state:

$$|+\rangle_A|-\rangle_B$$

represented by a vector in the 4-dimensional product Hilbert space for the two particles. An entangled pure state is a linear sum or "superposition" of product states that cannot itself be expressed as a product state. (More generally, for mixed states, representing mixtures or probability distributions of pure states, an entangled state is a state that cannot be represented as a convex combination or probability distribution of product states.)

Suppose Alice and Bob each holds one of a pair of particles in the entangled state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_A|-\rangle_B - |-\rangle_A|+\rangle_B)$$

The coefficients (here $\pm 1/\sqrt{2}$) can be complex numbers in general, and the squares of the absolute values of the coefficients (which are required to sum to 1: here ½ in both cases) represent the probabilities of obtaining the corresponding values of the relevant observables on measurement (+ and –, or – and +, for *A* and *B*). It turns out that Bob's state, which defines the statistics for measurement outcomes on his particle, can be represented as an equal weight mixture of the orthogonal states $|+\rangle_B$, $|-\rangle_B$, but equivalently as an infinity of other mixtures including, to take a specific example, the equal weight mixture of the four nonorthogonal states, represented as superpositions with complex coefficients $\pm\alpha$, $\pm\beta$ in the 2-dimensional Hilbert space of Bob's particle:

$$|\phi_1\rangle_B = \alpha|+\rangle_B + \beta|-\rangle_B$$

$$|\phi_2\rangle_B = \alpha|+\rangle_B - \beta|-\rangle_B$$

$$|\phi_3\rangle_B = \beta|+\rangle_B + \alpha|-\rangle_B$$

$$|\phi_4\rangle_B = \beta|+\rangle_B - \alpha|-\rangle_B$$

If Alice measures the spin observable with outcomes associated with the two possible states $|+\rangle_A$, $|-\rangle_A$ on her particle A, and Bob measures the corresponding spin observable on his particle B, Alice's outcomes will be oppositely correlated with Bob's outcomes (+ with –, and – with +). If instead Alice prepares a spin-½ particle A' in the state $|\phi\rangle_1\rangle_{A'} = \alpha|+\rangle_{A'} + \beta|-\rangle_{A'}$ and measures an observable on the pair of systems $A+A'$ in her possession with possible outcomes corresponding to the four orthogonal states:

$$|1\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{A'}|-\rangle_A - |-\rangle_{A'}|+\rangle_A)$$

$$|2\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{A'}|-\rangle_A + |-\rangle_{A'}|+\rangle_A)$$

$$|3\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{A'}|+\rangle_A - |-\rangle_{A'}|-\rangle_A)$$

$$|4\rangle = \frac{1}{\sqrt{2}}(|+\rangle_{A'}|+\rangle_A + |-\rangle_{A'}|-\rangle_A)$$

(the so-called Bell states), she will obtain the outcomes 1, 2, 3, 4 with equal probability, and these outcomes will be correlated with Bob's states $|\phi_1\rangle_B$, $|\phi_2\rangle_B$, $|\phi_3\rangle_B$, $|\phi_4\rangle_B$ (i.e., if

Bob checks to see whether his particle is in the state $|\phi_i\rangle_B$ when Alice reports that she obtained the outcome i=1, 2, 3, 4, he will find that this is always in fact the case). This follows because:

$$|\phi_1\rangle_{A'}|\psi\rangle = \frac{1}{2}(-|1\rangle|\phi_1\rangle_B - |2\rangle|\phi_2\rangle_B + |3\rangle|\phi_3\rangle_B + |4\rangle|\phi_4\rangle_B)$$

In this sense, Alice can "steer" Bob's particle into any equivalent mixture generating the same statistics by an appropriate local measurement.

Now, remote "steering" in this probabilistic sense is precisely what makes quantum teleportation possible. Suppose Alice and Bob share a pair of spin-½ particles A and B in the entangled state and Alice is given a spin-½ particle A' in an *unknown* state $|\phi_1\rangle$. There is no procedure by which Alice can determine the unknown state, but if Alice measures the composite system A+A' in the Bell basis, she will "steer" Bob's particle into one of the states $|\phi_1\rangle_B$, $|\phi_2\rangle_B$, $|\phi_3\rangle_B$, $|\phi_4\rangle_B$ with equal probability. If Alice tells Bob the outcome of her measurement, Bob can apply a local operation corresponding to a transformation in the Hilbert space of his particle to obtain the state $|\phi_1\rangle_B$.

Note that before Alice sends Bob the outcome of her measurement, the quantum state that Bob assigns to his particle—the information represented by the mixed state—is unchanged by Alice's measurement operation, even though after Alice's measurement the probability is ¼ that the state of Bob's particle is in fact $|\phi_1\rangle$ (in this case the local operation to obtain the state is represented by the identity). The trick that results in the transference of the state $|\phi_1\rangle$ from Alice to Bob, without the particle A' traveling from Alice to Bob, is the ability afforded Alice by the shared entangled state to correlate one of four measurement outcomes (each occurring with probability ¼) with one of four states that together represent a particular decomposition of Bob's mixed state. The transference of the state of A' to Bob's particle is accomplished by Bob's operation, which requires that Alice sends the information about her measurement outcome to Bob. In the teleportation protocol the state of the particle A' is destroyed by Alice's measurement and recreated as the state of Bob's particle by Bob's operation—in fact, the systems A and A' end up in an entangled state as the result of Alice's measurement. (Note that if the state $|\phi_1\rangle$ of A' were not destroyed there would be two copies of the state, which would violate the quantum "no cloning" theorem.)

COMPUTATION VIA ENTANGLEMENT

The field of quantum computation was launched in the 1980s with two seminal papers by David Deutsch in 1985 and Richard Feynman in 1982. The basic idea can be illustrated by the first genuinely quantum algorithm, proposed by Deutsch, later improved by Duetsch and Jozsa in 1992.

Consider a function f that maps an input value x =0 or x =1 onto an output value that is either 0 or 1. The algorithm for computing f might be quite complicated. To take Mermin's example, $f(x)$ might represent the value of the millionth bit in the binary expansion of $\sqrt{2+x}$, so that $f(0)$ is the millionth bit in the expansion of $\sqrt{2}$ while $f(1)$ is the millionth bit in the expansion of $\sqrt{3}$. Suppose we are interested in whether the function $f(x)$ is constant for both values of x or takes different values for both values of x—whether the millionth bit of $\sqrt{2}$ is the same as the millionth bit of $\sqrt{3}$, or not. With a classical computer we would have to run through the algorithm twice to evaluate $f(0)$ and $f(1)$ and then compare these values. With a quantum computer it is possible to answer the question in a single run of the algorithm.

We might represent the computation of f by a classical computer as follows:

$$\langle 0\,\rangle\langle 0\,\rangle \rightarrow \langle 0\,\rangle\langle f(0)\,\rangle$$

$$\langle 1\,\rangle\langle 0\,\rangle \rightarrow \langle 1\,\rangle\langle f(1)\,\rangle$$

where $\langle x\,\rangle\langle f(x)\,\rangle$ represents the input and output registers for the computation, and $\rightarrow$ represents the mapping defined by the algorithm.

In the case of a quantum computer the input and output registers are quantum states, specifically here "qubits," or states represented as orthogonal vectors in a 2-dimensional Hilbert space:

$$|0\rangle|0\rangle \rightarrow |0\rangle|f(0)\rangle$$

$$|1\rangle|0\rangle \rightarrow |1\rangle|f(1)\rangle$$

Here $\rightarrow$ represents the quantum mechanical implementation of the algorithm by quantum transformations of the input state. We could put the input register into a superposition of quantum states $\frac{1}{\sqrt{2}}(|O\rangle + 1\rangle)$, in which case, since the quantum transformations are linear maps, the quantum implementation of the algorithm yields:

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle|f(0)\rangle + |1\rangle|f(1)\rangle)$$

This state is a linear superposition of *both* possible inputs and the associated outputs of f, apparently representing the computation for *all* possible values. Unfortunately however only a single read-out is possible: if we make a measurement on both registers, we obtain just one of the possible results with probability ½, and the original state,

in which all possible inputs and associated outputs are represented, is altered. So there is no advantage over a classical computer.

Now we are interested in a global property of f, whether f is constant or balanced. Remarkably it turns out that we can answer this question in just one run of the algorithm, but at the expense of foregoing any information about the value of the function for either input. The final state is a product state (of the two registers) if $f(0) = f(1)$ and an entangled state if $f(0) \neq f(1)$. By appropriate quantum transformations (see the discussion in Mermin's *Lecture Notes on Quantum Computation*) this state can be transformed to:

$$|1\rangle(|f(0)\rangle - |\overline{f(0)}\rangle)$$

if $f(0) = f(1)$, and to:

$$|0\rangle(|f(0)\rangle - |\overline{f(0)}\rangle)$$

if $f0) \neq f(1)$, where $\overline{0} = 1$ and $\overline{1} = 0$. The outcome of a measurement on the input register, + or –, will distinguish whether $f(0) = f(1)$ or $f(0) \neq f(1)$.

Note that a measurement of the output register will yield the value $f(0)$ or $\overline{f(0)}$ with probability ½, that is, 0 or 1 with probability $\frac{1}{2}$, which provides no information about the value of f for either input. In general a quantum computation involves the evolution of correlations in successive entangled states to a final state in which a measurement can determine the answer to a question about a global property of a function. The global property here is a disjunctive property:

$$(f(0) = f(1) = 0) \vee (f(0) = f(1) = 1)$$

or:

$$((f(0) = 1) \wedge (f(1) = 0)) \vee ((f(0) = 0) \wedge (f(1) = 1))$$

and the computation yields one or other of these disjunctions in the final measurement, which excludes the possibility of recording of the values of the disjuncts. The two alternative disjunctions are represented in the 4-dimensional Hilbert space of the two registers as quantum disjunctions, corresponding to two orthogonal 2-dimensional planes. Depending on whether the function is constant or balanced, the final state of the two registers is represented by a vector lying in one or the other of these two planes, and this can be determined by a measurement of the input register.

The Deutsch-Jozsa algorithm is a simple example of a quantum algorithm. More sophisticated quantum computation algorithms, such as Shor's algorithm for finding the prime factors of a number, demonstrate an exponential speed-up over classical computation. The foundational significance of quantum computation concerns our understanding of computational complexity, that is, the relative efficiency of computational algorithms, rather than our characterization of the class of computable functions as those functions computable by a Turing machine. What quantum computation achieves is the possibility of solving a problem in a run-time (or number of computational steps) that increases as a polynomial function of the size of the input, while the computation using a classical computer would require superpolynomial, typically exponential, time. The difference can be quite dramatic. A classical computer of the sort available as of this writing would take an amount of time longer than the age of the universe to factor a 250-digit number into its two prime factors, using the fastest known algorithm. By contrast a quantum computer using Shor's algorithm could find the factors in minutes.

See also Bell, John, and Bell's Theorem; Bohr, Niels; Einstein, Albert; Hilbert, David; Many Worlds/Many Minds Interpretation of Quantum Mechanics; Quantum Mechanics; Schrödinger, Erwin.

Bibliography

Bell, John S. "On the Einstein-Podolsky-Rosen Paradox." *Physics* 1 (1964): 195–200. Reprinted in *Speakable and Unspeakable in Quantum Mechanics*, John S. Bell. Cambridge, U.K.: Cambridge University Press, 1987.

Bennett, Charles, Gilles Brassard, Claude Crepeau, Richard Jozsa, Asher Peres, and William Wootters. "Teleporting an Unknown Quantum State via Dual Classical and EPR Channels." *Physical Review Letters* 70 (1993): 1895–1899.

Bohr, Niels. "Discussion with Einstein on Epistemological Problems in Modern Physics." In *Albert Einstein: Philosopher-Scientist*, edited by Paul A. Schilpp, 201–241. La Salle, IL: Open Court, 1949.

Brassard, Gilles, Samuel Braunstein, and Richard Cleve. "Teleportation as a Quantum Computation." *Physica D* 120 (1998): 43–47.

Brown, Julian. *Minds, Machines, and the Multiverse*. New York: Simon and Schuster, 2000.

Deutsch, David. "Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer." *Proceedings of the Royal Society of London* A400 (1985): 97.

Deutsch, David, and Richard Jozsa. "Rapid Solution of Problems of Quantum Computation." *Proceedings of the Royal Society of London* A425 (1968): 73–90.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" *Physical Review* 47 (1935): 777–780.

Feynman, Richard P. "Simulating Physics with Computers." *International Journal of Theoretical Physics* 21 (1982): 467–488.

Mermin, David N. *Lecture Notes on Quantum Computation*. Available from http://people.ccmr.cornell.edu/mermin/qcomp/CS483.html.

Nielsen, Michael, and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2000.

Schrödinger, Erwin. "Discussion of Probability Relations Between Separated Systems." *Proceedings of the Cambridge Philosophical Society* 31 (1935): 555–563.

Schrödinger, Erwin. "Probability Relations between Separated Systems." *Proceedings of the Cambridge Philosophical Society* 32 (1936): 446–452.

Jeffrey Bub (2005)

QUANTUM LOGIC AND PROBABILITY

Quantum physics predicts many astonishing physical effects that have been subsequently observed in the laboratory. Perhaps the most significant effect is the violation of Bell's inequality, which implies a failure of classical locality. But the most widely known bit of quantum magic is the experiment of Clinton Davisson and Lester Germer demonstrating interference effects for electrons. Richard Feynman said this is a phenomenon "which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality, it contains the *only* mystery" (1963–1965, Vol. 3, p. 1-1). As we will see, Feynman somewhat overstates the case, but let us first try to get in his frame of mind.

The interference effect is illustrated by the two-slit experiment. If we send a plain water wave toward a barrier with two narrow slits in it, we find that a circular wave is produced on the far side of each slit. As these two circular waves expand, they eventually overlap and interfere. Where the crest of one meets the crest of the other, we get a crest of twice the height; where the trough of one meets the trough of the other, we get a trough of twice the depth; and where the crest of one meets the trough of the other, the waves cancel out. This creates *interference bands*: regions of extreme agitation where the waves meet in phase, crest-to-crest and trough-to-trough, juxtaposed with quiescent regions where the waves meet out of phase, crest-to-trough. The pattern of regions of high and low activity is easy to calculate. Notice, in particular, that there are places where one would observe wave motion if either slit alone were open, but where there are no waves when both slits are open, because of destructive interference.

In quantum theory, a *wave function* represents the physical state of an electron, and for a *single* electron the wave function is mathematically similar to a water wave. (It is not exactly the same, since it is a *complex-valued* function. Moreover, this analogy works only for a *single* electron. The wave function for a pair of electrons is defined on the configuration space of the system, which has more dimensions than physical space.) The dynamics of the wave function is similar enough to the dynamics of water waves to display the same interference effects. That is, in the case of a single electron shot at a screen with a single slit, the wave function that makes it through the slit spreads out on the far side in a sort of circular pattern. And in the case of a single electron shot at a screen with two slits, the wave function that gets through spreads out in two circular patterns, one centered at each slit, and these interfere where they overlap, just like the water waves.

Of course, when we actually *look for* a single electron, we never find it spread out; we always find it at some localized place. We can use the wave function to make predictions about where the electron will appear by squaring the wave function and interpreting this value as the *probability* that the particle will be found at a particular location. If we do many identically prepared experiments, we find that the distribution of the electrons matches the square of the amplitude of the wave function, thereby confirming the predictive accuracy of quantum mechanics. But the mystery is this: To get these interference effects, we do not have to send many electrons *at the same time*. We can send the electrons through the device one at a time, with long gaps between them, and watch the interference bands build up slowly, dot by dot. So it is not that different electrons are somehow interfering *with each other*; it is rather that each electron is somehow interfering *with itself*.

To make the effect even more vivid, consider this fact. We send electrons through the slits one at a time and watch for flashes on a distant screen. There are particular areas on the distant screen where we will sometimes see flashes when only the right slit is open, and sometimes see flashes when only the left slit is open, but *never* see flashes (because of destructive interference) when both slits are open. So each electron sent through when both slits are open must somehow be physically influenced by the fact that both are open, since that region is only forbidden when both are open. But, to put the question fancifully, how can an electron "know" that both slits are open if (being a tiny particle) it only goes through one slit?

This would appear to be a straightforward *physical* question that calls for a *physical* answer. And indeed, two different physical answers are available, corresponding to the two straightforward ways to interpret the quantum formalism. On the one hand, if one thinks that the wave function is *complete*, that is, that it encodes all the physical characteristics of the electron, then one will simply reject that claim that the electron is a tiny particle that can only go through one slit. If the wave function is complete, then when the wave function spreads out enough to go through both slits, the electron itself spreads out enough to go through both slits, and that is how it can interfere with itself. This leaves a mystery, but the mystery is not why there is interference. Rather, it is why the electron makes a small, localized flash on the far screen. This problem is solved, in this approach, by giving an account of wave-function collapse.

The second physical answer maintains that the electron is indeed a tiny particle that always has a well-defined location, and hence goes through one slit or the other. On this approach, the wave function is *not* complete, since it does not indicate what that position is. This account is realized in the "pilot wave" theory of Louis de Broglie and David Bohm. As John Bell has written, "While the founding fathers agonized over the question: 'particle' *or* 'wave', de Broglie in 1925 proposed the obvious answer: 'particle' *and* 'wave'." That is, in the view of de Broglie and Bohm, there is, in addition to the located particle, a wave function that *guides the trajectory* of the particle. The state of the wave function is influenced by the fact that both slits are open, in exactly the way the quantum formalism indicates. So each particle "knows" that both slits are open, even though it goes through only one, because the *wave function* "knows" that both slits are open, and the wave function guides the particle.

These two physical answers to the puzzle of the interference bands are perfectly adequate, and evidently require no adjustments to classical logic or probability theory. The solution of de Broglie and Bohm is even, in certain sense, a classical solution, contrary to Feynman's worry. So there is nothing in the phenomena discovered by modern physics that could *require* us to abandon or modify classical logic or probability theory.

Nonetheless, there have been many attempts, of various sorts, to argue that a change in logic or probability theory is at least *suggested* by the mathematical form of quantum theory, or that a change in logic or probability will produce an interpretation that is both physically adequate and somehow preferable to the two physical solutions outlined above.

At this point one would like a clear account of how classical logic or probability theory might be changed, and how the change might help us understand phenomena like the two-slit experiment without recourse to the sorts of physical hypotheses discussed above (hypotheses that, by the way, are already used to solve the measurement problem in quantum theory). Unfortunately, no such clear account is possible, because despite a long history and many attempts, no such account has ever been produced. So in its place, we must search instead for the reasons that anyone ever thought that classical logic or probability theory is responsible for the "mystery" surrounding these phenomena.

There are several different routes that can lead us to call into question classical logic. One, followed by Feynman in his famous *Lectures on Physics* (1963–1965), proceeds by reasoning about the two-slit experiment. The other, which is the foundation of the technical field of quantum logic, proceeds from an analysis of the mathematical machinery of quantum theory. Let us examine these in turn.

In his analysis of the two-slit experiment, Feynman first introduced proposition A:

> *Proposition A*. Each electron *either* goes through slit 1 *or* goes through slit 2.

Feynman then went on to consider what he calls the *consequences* of this proposition for predictions about the results of the experiment. If proposition A is true, he said, then we ought to be able to calculate the probability that the electron will land at any point of the screen by first determining the probability for electrons that go through slit 1 (by blocking slit 2 and seeing what happens), and then determining the probability for electrons that pass through slit 2 (by blocking slit 1 and seeing what happens). If proposition A it true when both slits are open, Feynman said, then the individual probabilities derived from these experiments should add. With both slits open, there are more ways for any result to come about (since an electron can get to a certain spot either by going through slit 1 or by going through slit 2), and the chance of the result should be just the sum of the chances of each process. This is, of course, not what we see. Because of the interference, there are places on the far screen where electrons appear with either slit open, but where no electrons appear with both slits open. Feynman concluded, "When one does *not* try to tell which way the electron goes, when there is nothing in the experiment to disturb the electrons, then one may *not* say that an electron goes through either hole 1 or hole 2. If one does say that, and starts to make deductions from the statement, he will make errors

in the analysis." That is, Feynman concludes from considerations of how the probabilities ought to add that proposition A is not true.

We can equally well present Feynman's dilemma using only logic rather than probability theory. There are places on the screen where an electron can appear when only slit 1 is open (and the electron goes through slit 1) and also when only slit 2 is open (and the electron goes through slit 2). So if the electron goes through slit 1, it can appear at a certain point, and if it goes through slit 2, it can appear at that same point. From the premise that the electron either goes through slit 1 or goes through slit 2, it then follows by a disjunctive syllogism that it can appear at that point. But with *both* slits open, the electron cannot appear at that point. It seems to follow that when both slits are open, the disjunction is not true. It is not the case that the electron went through slit 1 or went through slit 2.

Something must have gone wrong with Feynman's analysis somewhere. For in the theory of de Broglie and Bohm, electrons always have exact locations, and every electron that gets from the source to the far screen goes either through slit 1 or through slit 2. And the de Broglie and Bohm theory makes all the right predictions: exactly the predictions of quantum theory. Where did Feynman go wrong?

The solution is not hard to seek. Feynman considers first doing an experiment with slit 1 open *and slit 2 closed*, and then an experiment with slit 2 open *and slit 1 closed*. So the experimentally confirmed propositions are that if the electron goes through slit 1 with slit 2 closed, it can appear at a certain spot, and that if it goes through slit 2 with slit 1 closed, it can appear at that spot. The relevant disjunction for using disjunctive syllogism is the following: The particle either goes through slit 1 with slit 2 closed or through slit 2 with slit 1 closed. From this disjunction it does indeed follow that the electron can appear at the spot. But this disjunction tells us nothing at all about what can happen *with both slits open*.

Feynman's thought, evidently, is that if the electron goes through one slit, then it cannot make any difference whether the other slit is open. This is a reasonable conjecture, supported by classical intuitions. But this conjecture is false, and quantum theory shows why it is false: The state of the wave function is influenced by the state of both slits. Indeed, one consequence of quantum mechanics is that the state of the wave function is influenced by the presence or absence of detectors at either slit. Even when both slits are open, a detector at one slit will cause the interference to go away *even when the detector does not fire*. This is a straightforward mathematical consequence of the dynamics of the wave function. The ultimate *physical* moral is that one must take account of the entire experimental arrangement when considering what quantum mechanics predicts. As John Bell put it, "When one forgets the role of the apparatus … , one despairs of ordinary logic… . Hence 'quantum logic.' When one remembers the role of the apparatus, ordinary logic is just fine." And the apparatus in question is the *whole experimental situation*, including elements (such as the presence or absence of detectors that *do not fire*) that would be deemed irrelevant in classical physics.

Feynman's argument is a *physical* argument: It proceeds solely from the observation of experimental results to the (incorrect) conclusion that proposition A cannot be true, since one could deduce a false consequence from it. The field of quantum logic takes the opposite tack. Quantum logicians want to maintain that something like proposition A is *true* when both slits are open. But since false claims can apparently be deduced from proposition A using classical logic, this requires a change in logic itself.

Quantum logicians tend not to start from experiment, as Feynman does, but from observations about the form of the mathematical apparatus used in quantum theory. In particular, they begin with the observation that the space of all wave functions is a complex *vector space*. This means that given any pair of wave functions $|\psi_1\rangle$ and $|\psi_2\rangle$, and any two complex numbers α and β, there exists another wave function of the form $\alpha|\psi_1\rangle + \beta|\psi_2\rangle$. Such a wave function is called a *superposition* of $|\psi_1\rangle$ and $|\psi_2\rangle$.

Suppose that the wave function of an electron that goes through slit 1 with slit 2 closed is $|\psi_1\rangle$, and the wave function of an electron that goes through slit 2 with slit 1 closed is $|\psi_2\rangle$. Then when *both* slits are open, the wave function will be of the form $\alpha|\psi_1\rangle + \beta|\psi_2\rangle$, a superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$. (In particular, in the usual experimental configuration, it will be $(1/\sqrt{2})|\psi_1\rangle + (1/\sqrt{2})|\psi_2\rangle$.) This wave function is evidently neither $|\psi_1\rangle$ nor $|\psi_2\rangle$. It would *not* be correct to say, with classical logical connectives, that the electron is either in state $|\psi_1\rangle$ or in state $|\psi_2\rangle$. So if we allow $|\psi_1\rangle$ now to stand for the *proposition* that the electron is in state $|\psi_1\rangle$, and $|\psi_2\rangle$ to stand for the proposition that the electron is in state $|\psi_2\rangle$, then the classical proposition "$|\psi_1\rangle$ or $|\psi_2\rangle$" is *false* when both slits are open.

What the quantum logician does, though, is to introduce a *new* connective, usually written $\vee$, that is defined so that $|\psi_1\rangle \vee |\psi 2_2\rangle$ *is* true whenever the electron is in a superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$. If one tries to think of $\vee$ as

a sort of disjunction, one can then have a disjunction that is true even though neither disjunct is true—a circumstance that violates classical truth conditions.

More technically, the quantum logician associates propositions with subspaces of Hilbert space, the vector space of the wave function. The "conjunction" of two propositions (written **A** $\wedge$ **B**) is just the *intersection* of the associated subspaces, and the "disjunction" of two propositions (**A** $\vee$ **B**) is the *span* of the subspaces, that is, the subspace consisting of all vectors that can be formed by adding vectors from the two given subspaces. A proposition is *true* just in case the wave function of the system lies in the associated subspace. So if the wave function of the system is $(1/\sqrt{2})|\psi_1\rangle + (1/\sqrt{2})|\psi_2\rangle$, then the proposition $|\psi_1\rangle$ is not true, and the proposition $|\psi_2\rangle$ is not true, but the proposition $|\psi_1\rangle \vee |\psi_2\rangle$ is true.

With this terminology in place, one can easily show that the set of "quantum propositions" form a non-Boolean (nondistributive) lattice under the operations $\vee$ and $\wedge$. This is a straightforward mathematical fact about the structure of subspaces of Hilbert space under these operations. There is no nonclassical logic or probability theory here, just standard mathematics.

Of course, if one starts to pronounce $\vee$ "or" and $\wedge$ "and," then matters can get somewhat confusing. Because the lattice of quantum propositions is nondistributive, (**A** $\vee$ **B**) $\wedge$ **C** can be true while (**A** $\wedge$ **C**) $\vee$ (**B** $\wedge$ **C**) is false. If one presents this fact by saying that "(**A** or **B**) and **C**" is true while "(**A** and **C**) or (**B** and **C**)" is false, then it appears that de Morgan's laws have failed. Hence the supposed need for quantum logic.

If quantum logic is just the study of the structure of subspaces of Hilbert space, then it is a perfectly legitimate, but badly named, enterprise. It is not an alternative to, or replacement for, classical logic, since it studies connectives that are not the classical connectives. Nothing has been shown to be wrong or misleading about classical logic. Rather, the problem lies with our intuitions about experimental conditions, which lead us incorrectly to expect that whether the second slit is open is irrelevant to the behavior of the electron at the other slit. Quantum mechanics shows not that there is anything wrong with classical logic, but rather that the physics of the quantum world is very unlike the physics of Isaac Newton and James Clerk Maxwell. The surprising relevance of experimental conditions is shown by experiments like the two-slit experiment, and the appropriate way to reason about these experiments is, of course, classically.

What about Feynman's proposition A? With both slits open, is it correct or incorrect to say that the electron either went through slit 1 or slit 2? The answer to this question once again depends on physics rather than logic. If the de Broglie and Bohm theory is correct, then the electron always goes through one slit or the other. Retrospectively, one can even tell which slit it went through. Proposition A is therefore true. If one adopts an interpretation according to which the wave function is complete, then the wave function is all there is to the electron, and the wave function "goes through" both slits. Part of it goes through each slit, so it goes neither *entirely* through slit 1 nor *entirely* through slit 2. On a truth-functional reading of "or," proposition A is false. As long one is clear about the exact content of any proposition and about the interpretation of quantum theory at issue, classical logic and probability theory work just fine.

See also Bell, John, and Bell's Theorem; Bohm, David; Hilbert, David; Logic, Non-Classical; Maxwell, James Clerk; Newton, Isaac; Non-locality; Quantum Mechanics.

Bibliography

Birkhoff, Garrett, and John von Neumann. "The Logic of Quantum Mechanics." *Annals of Mathematics* 37 (1936): 823–843.

Feynman, Richard, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*. Reading, MA: Addison-Wesley, 1963–1965.

Kochen, Simon, and Ernst P. Specker. "The Problem of Hidden Variables in Quantum Mechanics." *Journal of Mathematics and Mechanics* 17 (1967): 59–87.

Maudlin, Tim. "The Tale of Quantum Logic." In *The Philosophy of Hilary Putnam*, edited by Yemima Ben Menahem. New York: Cambridge University Press, 2005.

Putnam, Hilary. "Is Logic Empirical?" In *Boston Studies in the Philosophy of Science*. Vol. 5, edited by Robert S. Cohen and Marx Wartofsky. Dordrecht, Netherlands: D. Reidel, 1968. Reprinted as "The Logic of Quantum Mechanics," in his *Mathematics, Matter, and Method.* New York: Cambridge University Press, 1976.

Van Fraassen, Bas. *Quantum Mechanics: An Empiricist View*. New York: Oxford University Press, 1992.

Von Neumann, John. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer-Verlag, 1932. Translated by Robert T. Beyer as *Mathematical Foundations of Quantum Mechanics.* Princeton, NJ: Princeton University Press, 1955.

Tim Maudlin (2005)

QUANTUM MECHANICS

Quantum mechanics has the distinction of being considered both the most empirically successful and the most poorly understood theory in the history of physics.

To take an oft-cited example of the first point: The theoretically calculated value of the anomalous magnetic moment of the electron using quantum electrodynamics matches the observed value to twelve decimal places, arguably the best confirmed empirical prediction ever made. To illustrate the second point, we have the equally oft-cited remarks of Niels Bohr, "Anyone who says that they can contemplate quantum mechanics without becoming dizzy has not understood the concept in the least," and of Richard Feynman, "[We] have always had (secret, secret, close the doors!) we always have had a great deal of difficulty in understanding the world view that quantum mechanics represents." How could both of these circumstances obtain?

For the purposes of making predictions, quantum theory consists in a mathematical apparatus and has clear enough rules of thumb about how to apply the mathematical apparatus in various experimental situations. If one is doing an experiment or observing something, one must first associate a mathematical *quantum state* or *wave function* with the system under observation. For example, if one prepares in the laboratory an electron beam with a fixed momentum, then the quantum state of each electron in the beam will be something like a sine wave. In the case of a single particle it is common to visualize this wave function as one would a water wave: as an object extended in space. Although this visualization works for a single particle, it does not work in general, so care must be taken. But for the moment, this simple visualization works. The wave function for the electron is "spread out" in space.

The second part of the mathematical apparatus is a *dynamical equation* that specifies how the quantum state changes with time so long as no observation or measurement is made on the system. These equations have names like the *Schrödinger equation* (for nonrelativistic quantum mechanics) and the *Dirac equation* (for relativistic quantum field theory). In the case of the electron mentioned earlier the dynamical equation is relevantly similar to the dynamical equation for water waves, so we can visualize the quantum state as a little plane water wave moving in a certain direction. If the electron is shot at a screen with two slits in it, then the quantum state will behave similarly to a water wave that hits such a barrier: circularly expanding waves will emerge from each slit, and there will be constructive and destructive interference where those waves overlap. If beyond the slits there is a fluorescent screen, we can easily calculate what the quantum state "at the screen" will look like: It will have the peaks and troughs characteristic of interfering water waves.

Finally comes the interaction with the screen. Here is where things get tricky. One would naively expect that the correct way to understand what happens when the electron wave function reaches the screen is to build a physical model of the screen and apply quantum mechanics to it. But that is not what is done. Instead, the screen is treated as a measuring device and the interaction with the screen as a measurement, and new rules are brought into play.

The new rules require that one first decide what property the measuring device measures. In the case of a fixed screen it is taken that the screen measures the *position* of a particle. If instead of a fixed screen we had an absorber on springs, whose recoil is recorded, then the device would measure the *momentum* of the particle. These determinations are typically made by relying on classical judgments: There is no algorithm for determining what a generic (physically specified) object "measures," or indeed whether it measures anything at all. But laboratory apparatus for measuring position and momentum have been familiar from before the advent of quantum theory, so this poses no real practical problem.

Next, the property measured gets associated with a mathematical object called a *Hermitian operator*. Again, there is no algorithm for this, but for familiar classical properties like position and momentum the association is established. For each Hermitian operator there is an associated set of wave functions called the *eigenstates* of the operator. It is purely a matter of mathematics to determine the eigenstates. Each eigenstate has associated with it an *eigenvalue*: The eigenvalues are supposed to correspond to the possible outcomes of a measurement of the associated property, such as the possible values of position, momentum, or energy. (Conversely, it is typically assumed that for every Hermitian operator, there corresponds a measurable property and possible laboratory operations that would measure it, although there is no general method for specifying these.)

The last step in the recipe for making predictions can now be taken. When a system is measured, the wave function for the system is first expressed as a sum of terms, each term being an eigenstate of the relevant Hermitian operator. Any wave function can be expressed as a sum of such terms, with each term given a weight, which is a

complex number. For example, if an operator has only two eigenstates, call them $|1>$ and $|2>$, then any wave function can be expressed in the form $\alpha|1> + \beta\,|1>$, with α and β complex numbers such that $|\alpha|^2 + |\beta\,|^2 = 1$. (This is the case, for example, when we measure the so-called spin of an electron in a given direction, and always get one of two results: spin up or spin down.) Recall that each eigenstate is associated with a possible outcome of the measurement: $|1>$, for example, could be associated with getting spin up, and $|2>$ with getting spin down. The quantum mechanical prediction is now typically a probabilistic one: the chance of getting the result associated with $|1>$ is $|\alpha|^2$, and the chance of getting the result associated with $|2>$ is $|\beta\,|^2$. In general, one writes out the wave function of the system in terms of the appropriate eigenstates, and then the chance of getting the result associated with some eigenstate is just the square of the complex number that weights the state.

We can now see how quantum theory makes empirical predictions: So long as one knows the initial quantum state of the system and the right Hermitian operator to associate with the measurement, the theory will allow one to make probabilistic predictions for the outcome. Those predictions turn out to be exquisitely accurate.

If a Hermitian operator has only a finite number of eigenstates, or the eigenvalues of the operator are discrete, then any associated measurement should have only a discrete set of possible outcomes. This has already been in the case of spin; for a spin-1/2 particle such as an electron, there are only two eigenstates for the spin in a given direction. Physically, this means that when we do an experiment to measure spin (which may involve shooting a particle through an inhomogeneous magnetic field) we will get only one of two results: Either the particle will be deflected up a given amount or down a given amount (hence spin up and spin down). In this case the physical quantity is *quantized*; it takes only a discrete set of values. But quantum theory does not require all physical magnitudes to be quantized in this way; the position, momentum, or energy of a free particle is not. So the heart of quantum theory is not a theory of discreteness, it is rather just the mathematical apparatus and the rules of application described earlier.

THE MEASUREMENT PROBLEM

Why, then, is the quantum theory so puzzling, or so much more obscure than, say, classical mechanics? One way that it differs from classical theory is that it provides only probabilistic predictions for experiments, and one might well wonder, as Albert Einstein famously did, whether this is because "God plays dice with the universe" (i.e., the physical world itself is not deterministic) or whether the probabilities merely reflect our incomplete knowledge of physical situation. But even apart from the probabilities, the formulation of the theory is rather peculiar. Rules are given for representing the physical state of a system and for how that physical state evolves and interacts with other systems when no measurement takes place. This evolution is perfectly deterministic. A different set of rules is applied to derive predictions for the outcomes of experiments, and these rules are not deterministic. Still, an experiment in a laboratory is just a species of physical interaction, and ought to be treatable as such. There should be a way to describe the physical situation in the lab, and the interaction of the measured system with the measuring device, that relies only on applying, say, the Schrödinger equation to the physical state of the system plus the lab.

John S. Bell put this point succinctly, "If you make axioms, rather than definitions and theorems, about the 'measurement' of anything else, then you commit redundancy and risk inconsistency" (1987, p. 166). You commit redundancy because while the axioms about measurement specify what should happen in a measurement situation, the measurement situation, considered as a simple physical interaction, ought also to be covered by the general theory of such interactions. You risk inconsistency because the redundancy produces the possibility that the measurement axioms will contradict the results of the second sort of treatment. This is indeed what happens in the standard approaches to quantum mechanics. The result is called the *measurement problem*.

The measurement problem arises from a conflict in the standard approach between treating a laboratory operation as a normal physical interaction and treating it as a measurement. To display this conflict, we need some way to represent the laboratory apparatus as a physical device and the interaction between the device and the system as a physical interaction. Now this might seem to be a daunting task; a piece of laboratory apparatus is typically large and complicated, comprising astronomically large numbers of atoms. By contrast, exact wave functions are hard to come by for anything much more complicated than a single hydrogen atom. How can we hope to treat the laboratory operation at a fundamental level?

Fortunately, there is a way around this problem. Although we cannot write down, in detail, the physical state of a large piece of apparatus, there are conditions that we must assume if we are to regard the apparatus as a good measuring device. There are necessary conditions

for being a good measuring device, and since we do regard certain apparatus as such devices, we must be assuming that they meet these conditions.

Take the case of spin. If we choose a direction in space, call it the *x*–direction, then there is a Hermitian operator that gets associated with the quantity *x*–spin. That operator has two eigenstates, which we can represent as $|x\text{–up}>_S$ and $|x\text{–down}>_S$. The subscript s indicates that these are states of the system to be measured. We have pieces of laboratory equipment that can be regarded as good devices for measuring the *x*–spin of a particle. We can prepare such an apparatus in a state, call it the "ready" state, in which it will function as a good measuring device. Again, we do not know the exact physical details of this ready state, but we must assume such states exist and can be prepared. What physical characteristics must such a ready state have?

Besides the ready state, the apparatus must have two distinct indicator states, one of which corresponds to getting an "up" result of the measurement and the other that corresponds to getting a "down" result. And the key point about the physics of the apparatus is this: It must be that if the device in its ready state interacts with a particle in the state $|x\text{–up}>_S$, it will evolve into the indicator state that is associated with the up result, and if it interacts with a particle in state $|x\text{–down}>_S$, it will evolve into the other indicator state.

This can be put in a formal notation. The ready state of the apparatus can be represented by $|\text{ready}>_A$, the up indicator state by $|\text{“up”}>_A$, and the down indicator state by $|\text{“down”}>_A$. If we feed an *x*–spin up particle into the device, the initial physical state of the system plus apparatus is represented by $|x\text{–up}>_S|\text{ready}>_A$, if we feed in an *x*–spin down particle the initial state is $|x\text{–down}>_S|\text{ready}>_A$. If the apparatus is, in fact, a good *x*–spin measuring device, then the first initial state must evolve into a state in which the apparatus indicates up, that is, it must evolve into $|x\text{–up}>_S|\text{“up”}>_A$, and the second initial state must evolve into a state that indicates down, that is, $|x\text{–down}>_S|\text{“down”}>_A$. Using an arrow to represent the relevant time evolution, then, we have for any good *x*–spin measuring device

$$|x\text{–up}>_S|\text{ready}>_A \rightarrow |x\text{–up}>_S|\text{“up”}>_A \text{ and}$$

$$|x\text{–down}>_S|\text{ready}>_A \rightarrow |x\text{–down}>_S|\text{“down”}>_A.$$

We have not done any real physics yet, we have just indicated how the physics must come out if there are to be items that count as good *x*–spin measuring devices, as we think there are.

The important part of the physics that generates the measurement problem is the arrow in the representations listed earlier, the physical evolution that takes one from the initial state of the system plus apparatus to the final state. Quantum theory provides laws of evolution for quantum states such as the Schrödinger and Dirac equations. These would be the equations one would use to model the evolution of the system plus apparatus as a normal physical evolution. And all these dynamical equations have a common mathematical feature; they are all linear equations. It is this feature of the quantum theory that generates the measurement problem, so we should pause over the notion of linearity.

The set of wave functions used in quantum theory form a *vector space*. This means that one can take a weighted sum of any set of wave functions and get another wave function. (The weights in this case are complex numbers, hence it is a complex vector space.) This property was mentioned earlier when it was noted that any wave function can be expressed as a weighted sum of the eigenvectors of an observable. An operator on a vector space is just an object that maps a vector as input to another vector as output. If the operator *O* maps the vector **A** to the vector **B**, we can write that as

$$O(\mathbf{A}) = \mathbf{B}.$$

A linear operator has the feature that you get the same result whether to operate on a sum of two vectors or you first operate on the vectors and then takes the sum. That is, if *O* is a linear operator, then for all vectors **A** and **B**,

$$O(\mathbf{A} + \mathbf{B}) = O(\mathbf{A}) + O(\mathbf{B}).$$

The dynamical equations evidently correspond to operators; they take as input the initial physical state and give as output the final state, after a specified period has elapsed. But further, the Schrödinger and Dirac equations correspond to linear operators. Why is this important?

We have already seen how the physical state of a good *x*–spin measuring device must evolve when fed a particle in the state $|x\text{–up}>_S$ or the state $|x\text{–down}>_S$. But these are not the only spin states that the incoming particle can occupy. There is an infinitude of spin states, which correspond to all the wave functions that can be expressed as $\alpha|x\text{–up}>_S + \beta|x\text{–down}>_S$, with α and β complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$. Correspondingly, there is an infinitude of possible directions in space in which one can orient a spin measuring device, and each of the directions is associated with a different Hermitian operator. For a direction at right angles to the *x*–direction, call it the *y*–direction, there are eigenstates $|y\text{–up}>_S$ and

$|y\text{–down}\rangle_S$. These states can be expressed as weighted sums of the *x*–spin eigenstates, and in the usual notation

$$|y\text{–up}\rangle_S = 1/\sqrt{2}|x\text{–up}\rangle_S + 1/\sqrt{2}|x\text{–down}\rangle_S \text{ and}$$

$$|y\text{–down}\rangle_S = 1/\sqrt{2}|x\text{–up}\rangle_S - 1/\sqrt{2}|x\text{–down}\rangle_S.$$

So what happens if we feed a particle in the state $|y\text{–up}\rangle_S$ into the good *x*–spin measuring device?

Empirically, we know what happens: About half the time the apparatus ends up indicating "up" and about half the time it ends up indicating "down." There is nothing we are able to do to control the outcome: *y*–up eigenstate particles that are identically prepared nonetheless yield different outcomes in this experiment.

If we use the usual predictive apparatus, we also get this result. The "up" result from the apparatus is associated with the eigenstate $|x\text{–up}\rangle_S$ and the "down" result associated with $|x\text{–down}\rangle_S$. The general recipe tells us to express the incoming particle in terms of these eigenstates as $1/\sqrt{2}|x\text{–up}\rangle_S + 1/\sqrt{2}|x\text{–down}\rangle_S$, and then to take the squares of the weighting factors to get the probabilities of the results. This yields a probabilistic prediction of 50 percent chance "up" and 50 percent chance "down," which corresponds to what we see in the lab.

But if instead of the usual predictive apparatus we use the general account of physical interactions, we get into trouble. In that case, we would represent the initial state of the system plus apparatus as $|y\text{–up}\rangle_S|\text{ready}\rangle_A$. The dynamical equation can now be used to determine the physical state of the system plus apparatus at the end of the experiment.

But the linearity of the dynamical equations already determines what the answer must be. For

$$|y\text{–up}\rangle_S|\text{ready}\rangle_A = (1/\sqrt{2}|x\text{–up}\rangle_S + 1/\sqrt{2}|x\text{–down}\rangle_S)|\text{ready}\rangle_A$$

$$= 1/\sqrt{2}|x\text{–up}\rangle_S|\text{ready}\rangle_A + 1/\sqrt{2}|x\text{–down}\rangle_S|\text{ready}\rangle_A.$$

But we know how each of the two terms of this superposition must evolve, since the apparatus is a good *x*–spin measuring device. By linearity, this initial state must evolve into the final state

$$1/\sqrt{2}|x\text{–up}\rangle_S|\text{“up”}\rangle_A + 1/\sqrt{2}|x\text{–down}\rangle_S|\text{“down”}\rangle_A.$$

That is, the final state of the apparatus plus system must be a superposition of a state in which the apparatus yields the result "up" and a state in which the apparatus yields the result "down." That is what treating the measurement as a normal physical interaction must imply.

So by making axioms about measurements, we have both committed redundancy and achieved inconsistency. The axioms say that the outcome of the experiment is not determined by the initial state; each of two outcomes is possible, with a 50 percent chance of each. But the treatment of the measurement as a normal physical interaction implies that only one final physical state can occur. And furthermore, that final physical state is an extremely difficult one to understand. It appears to be neither a state in which the measuring apparatus is indicating "up" nor a state in which the apparatus is indicating "down," but some sort of symmetric combination of the two. If all the physical facts about the apparatus are somehow represented in its wave function, then it seems that at the end of the experiment the apparatus can neither be indicating up (and not down) nor down (and not up). But we always see one or the other when we do this experiment.

At this point our attention must clearly be turned to the mathematical object we have called the wave function. The wave function is supposed to represent the physical state of a system. The question is whether the wave function represents all of the physical features of a system, or whether systems represented by the same wave function could nevertheless be physically different. If one asserts the former, then one believes that the wave function is complete, if the latter, then the wave function is incomplete. The standard interpretations of the quantum formalism take the wave function to be complete, interpretations that take it to be incomplete are commonly called *hidden variables* theories (although that is a misleading name).

The wave function $1/\sqrt{2}|x\text{–up}\rangle_S|\text{“up”}\rangle_A + 1/\sqrt{2}|x\text{–down}\rangle_S|\text{“down”}\rangle_A$ does not represent the apparatus as indicating up (and not down) or as indicating down (and not up). So if the wave function is complete, the apparatus, at the end of the experiment, must neither be indicating up (and not down) nor down (and not up). But that flatly contradicts our direct experience of such apparatus. This is the measurement problem. As Bell puts it, "Either the wave function, as given by the Schrödinger equation, is not everything, or it is not right" (1987, p. 201).

COLLAPSE INTERPRETATIONS

COLLAPSE TIED TO OBSERVATION. What is one to do? From the beginning of discussions of these matters, Einstein held the argument to show that the wave function is not everything and hence that quantum mechanics is incomplete. The wave function might represent part of the physical state of a system, or the wave function might

represent some features of ensembles, collections, or systems, but the wave function cannot be a complete representation of the physical state an individual system, like the particular *x*–spin measuring device in the laboratory after a particular experiment is done. For after the experiment, the apparatus evidently either indicates "up" or it indicates "down," but the wave function does not represent it as doing so.

By contrast, the founders of the quantum theory, especially Bohr, insisted that the wave function is complete. And they did not want to deny that the measuring device ends up indicating one determinate outcome. So the only option left was to deny that the wave function, as given by the Schrödinger equation, is right. At some times, the wave function must evolve in a way that is not correctly described by the Schrödinger equation. The wave function must "collapse." The standard interpretation of quantum mechanics holds that the wave function evolves, at different times, in either of two different ways. This view was given its canonical formulation in John von Neumann's *Mathematical Foundations of Quantum Mechanics* (1955). Von Neumann believed (incorrectly, as we will see) that he had proven the impossibility of supplementing the wave function with hidden variables, so he thought the wave function must be complete. When he comes to discuss the time evolution of systems, Von Neumann says "[w]e therefore have two fundamentally different types of interventions which can occur in a system **S**. … First, the arbitrary [i.e., nondeterministic] changes by measurement. … Second, the automatic [i.e., deterministic] changes which occur with the passage of time" (p. 351). The second type of change is described by, for example, the Schrödinger equation, and the first by an indeterministic process of collapse.

What the collapse dynamics must be can be read off from the results we want together with the thesis that the wave function is complete. For example, in the *x*–spin measurement of the *y*–spin up electron, we want there to be a 50 percent chance that the apparatus indicates "up" and a 50 percent chance that it indicates "down." But the only wave function that represents an apparatus indicating "up" is $|\text{“up”}>_A$, and the only wave function for an apparatus indicating "down" is $|\text{“down”}>_A$. So instead of a deterministic transition to the final state

$$1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$$

we must postulate an indeterministic transition with a 50 percent chance of yielding $|x\text{–up}>_S|\text{“up”}>_A$ and a 50 percent chance of yielding $|x\text{–down}>_S|\text{“down”}>_A$.

It is clear what the collapse dynamics must do. What is completely unclear, though, is when it must do it. All Von Neumann's rules say is that we get collapses when measurements occur and deterministic evolutions "with the passage of time." But surely measurements also involve the passage of time; so under exactly what conditions do each of the evolutions obtain? Collapse theories, which postulate two distinct and incompatible forms of evolution of the wave function, require some account of when each type of evolution occurs.

Historically, this line of inquiry was influenced by the association of the problem with "measurement" or "observation." If one begins with the thought that the non-linear evolution happens only when a measurement or observation occurs, then the problem becomes one of specifying when a measurement or observation occurs. And this in turn suggests that we need a characterization of an observer who makes the observation. Pushing even further, one can arrive at the notion that observations require a conscious observer of a certain kind, folding the problem of consciousness into the mix. As Bell asks, "What exactly qualifies some physical systems to play the role of 'measurer'? Was the wave function of the world waiting to jump for thousands of millions of years until a single-celled living creature appeared? Or did it have to wait a little longer, for some better qualified system … with a Ph.D.?" (1987, p. 117).

This line of thought was discussed by Eugene Wigner, "This way out of the difficulty amounts to the postulate that the equations of motion of quantum mechanics cease to be linear, in fact that they are grossly non-linear if conscious beings enter the picture" (1967, p. 183). Wigner suggests that the quantum measurement problem indicates "the effect of consciousness on physical phenomena," a possibility of almost incomprehensible implications (not the least of which: How could conscious beings evolve if there were no collapses, since the universe would surely be in a superposition of states with and without conscious beings!). In any case, Wigner's speculations never amounted to a physical theory, nor could they unless a physical characterization of a conscious system was forthcoming.

So if one adopts a collapse theory, and if the collapses are tied to measurements or observations, then one is left with the problem of giving a physical characterization of an observation or a measurement. Such physicists as Einstein and Bell were incredulous of the notion that conscious systems play such a central role in the physics of the universe.

SPONTANEOUS COLLAPSE THEORIES. Nonetheless, precise theories of collapse do exist. The key to resolving the foregoing puzzle is to notice that although collapses must be of the right form to make the physical interactions called “observations” and “measurements” have determinate outcomes, there is no reason that the collapse dynamics itself need mention observation or measurement. The collapse dynamics merely must be of such a kind as to give outcomes in the right situations.

The most widely discussed theory of wave function collapse was developed by Gian Carlo Ghirardi, Alberto Rimini, and Tulio Weber (1986) and is called the spontaneous localization theory or, more commonly, the GRW theory. The theory postulates an account of wave function collapse that makes no mention of observation, measurement, consciousness, or anything of the sort. Rather, it supplies a universal rule for both how and when the collapse occurs. The “how” of the collapse involves localization in space; when the collapse occurs, one takes a single particle and multiplies its wave function, expressed as a function of space, by a narrow Gaussian (bell curve). This has the effect of localizing the particle near the center of the Gaussian, in the sense that most of the wave function will be near the center. If the wave function before the collapse is widely spread out over space, after the collapse it is much more heavily weighted to a particular region. The likelihood that a collapse will occur centered at a particular location depends on the square amplitude of the precollapse wave function for that location. The collapses, unlike Schrödinger evolution, are fundamentally nondeterministic, chancy events.

The GRW collapse does not perfectly locate the wave function at a point. It could not do so for straightforward physical reasons: The localization process will violate the conservation of energy, and the more narrowly the postcollapse wave function is confined, the more new energy is pumped into the system. If there were perfect localizations, the energy increase would be infinite—and immediately evident. (It follows from these same observations that even in the “standard” theory there are never collapses to perfectly precise positions—even after a so-called position measurement.)

Therefore, the GRW theory faces a decision: Exactly how localized should the localized wave function be? This corresponds to choosing a width for the Gaussian: The narrower the width, the more energy that is added to the system on collapse. The choice for this width is bounded in one direction by observation—the energy increase for the universe must be below observed bounds, and particular processes, such as spontaneous ionization, should be rare—and in the other direction by the demand that the localization solve the measurement problem. As it happens, Ghirardi, Rimini, and Weber chose a value of about 10^{-5} centimeters for the width of the Gaussian. This is a new constant of nature.

Beside the “how” of the collapse, the GRW theory must specify the “when.” It was here that we saw issues such as consciousness getting into the discussion: If collapses occur only when measurements or observations occur, then we must know when measurements or observations occur. The GRW theory slices through this problematic neatly; it simply postulates that the collapses take place at random, with a fixed probability per unit time. This introduces another new fundamental constant: the average time between collapses per particle. The value of that constant is also limited in two directions; on the one hand, we know from interference experiments that isolated individual particles almost never suffer collapses on the time scale of laboratory operations. On the other hand, the collapses must be frequent enough to resolve the measurement problem. The GRW theory employs a value of 10^{15} seconds, or about 100 million years, for this constant.

Clearly, the constant has been chosen large enough to solve one problem: Individual isolated particles will almost never suffer collapses in the laboratory. It is less clear, though, how it solves the measurement problem.

The key here is to note that actual experiments record their outcomes in the correlated positions of many, many particles. In our spin experiment we said that our spin measuring device must have two distinct indicator states: $|\text{“up”}>$ and $|\text{“down”}>$. To be a useful measuring device, these indicator states must be macroscopically distinguishable. This is achieved with macroscopic objects—pointers, drops of ink, and so on—to indicate the outcome. And a macroscopic object will have on the order of 10^{23} particles.

So suppose the outcome $|\text{“up”}>$ corresponds to a pointer pointing to the right and the outcome $|\text{“down”}>$ corresponds to the pointer pointing to the left. If there are no collapses, the device will end up with the wave function $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$. Now although it is unlikely that any particular particle in the pointer will suffer a collapse on the time scale of the experiment, because there are so many particles in the pointer, it is overwhelmingly likely that some particle or other in the pointer will suffer a collapse quickly: within about 10^{-8} seconds. And (this is the key), since in the state $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$ all the particle positions are correlated with one another, if the

collapse localizes a single particle in the pointer, it localizes all of them. So, if having the wave functions of all the particles in the pointer highly concentrated on the right (or on the left) suffices to solve the measurement problem, the problem will be solved before 10^{-4} seconds has elapsed.

The original GRW theory has been subject to much discussion. In a technical direction there have been similar theories, by Ghirardi and Rimini and by Philip Perle, that make the collapses to be continuous rather than discrete. More fundamentally, there have been two foundational questions: First, does the only approximate nature of the "localization" vitiate its usefulness in solving the measurement problem, and second, does the theory require a physical ontology distinct from the wave function? Several suggestions for such an additional ontology have been put forward, including a mass density in space-time, and discrete events ("flashes") in space-time.

The addition of such extra ontology, beyond the wave function, reminds us of the second horn of Bell's dilemma: Either the wave function as given by the Schrödinger equation is not right or it is not everything. The versions of the GRW theory that admit a mass density or the flashes postulate that the wave function is not everything, do so in such a way that the exact state of the extra ontology can be recovered from the wave function. The more radical proposal is that there is extra ontology, and its state cannot be read off the wave function. These are the so-called hidden variables theories.

ADDITIONAL VARIABLES THEORIES

According to an additional variables theory, the complete quantum state of the system after a measurement is indeed $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$. The outcome of the measurement cannot be read off of that state because the outcome is realized in the state of the additional variables, not in the wave function. It immediately follows that for any such theory, the additional ontology, the additional variables, had best not be "hidden": since the actual outcome is manifest, the additional variables had best be manifest. Indeed, on this approach the role of the wave function in the theory is to determine the evolution of the additional variables. The wave function, since it is made manifest only through this influence, is really the more "hidden" part of the ontology.

The best known and most intensively developed additional variables theory goes back to Louis de Broglie, but is most intimately associated with David Bohm. In its nonrelativistic particle version, Bohmian mechanics, physical objects are constituted of always-located point particles, just as was conceived in classical mechanics. At any given time, the physical state of a system comprises both the exact positions of the particles and a wave function. The wave function never collapses: it always obeys a linear dynamical equation like the Schrödinger equation. Nonetheless, at the end of the experiment the particles in the pointer will end up either all on the right or all on the left, thus solving the measurement problem. This is a consequence of the dynamics of the particles as determined by the wave function.

It happens that the particle dynamics in Bohmian mechanics is completely deterministic, although that is not fundamentally important to the theory and indeterministic versions of Bohm's approach have been developed. The dynamical equation used in Bohmian mechanics is much more importantly the simplest equation that one can write down if one assumes that the particle trajectories are to be determined by the wave function and that various symmetries are to be respected. If one starts with idea that there are particles and that quantum theory should be a theory of the motion of those particles that reproduces the predictions of the standard mathematical recipe, Bohmian mechanics is the most direct outcome.

Since Bohmian mechanics is a deterministic theory, the outcome of any experiment is fixed by the initial state of the system. The probabilities derived from the standard mathematical recipe must therefore be interpreted purely epistemically: they reflect our lack of knowledge of the initial state. This lack of knowledge turns out to have a physical explanation in Bohmian mechanics: Once one models any interaction designed to acquire information about a system as a physical interaction between a system and an observer, it can be shown to follow that initial uncertainty about the state of the target system cannot be reduced below a certain bound, given by the Heisenberg uncertainty relations.

This illustrates the degree to which the ontological "morals" of quantum theory are held hostage to interpretations. In the standard interpretation, when the wave function of a particle is spread out, there is no further fact about exactly where the particle is. (Because of this, position measurements in the standard theory are not really measurements, i.e., they do not reveal preexisting facts about positions.) In Bohm's interpretation, when the wave function is spread out, there is a fact about exactly where the particle is, but it follows from physical analysis that one cannot find out more exactly where it is without thereby altering the wave function (more properly, with-

out altering the effective wave function that we use to make predictions). Similarly, in the standard interpretation, when we do a position measurement on a spread out particle, there is an indeterministic collapse that localizes the particle—it gives it an approximate location. According to Bohm's theory the same interaction really is a measurement: It reveals the location that the particle already had. So it is a fool's errand to ask after "the ontological implications of quantum theory": the account of the physical world one gets depends critically on the interpretation of the formalism.

Bohm's approach has been adapted to other choices for the additional variables. In particular, interpretations of field theory have been pursued in two different ways: with field variables that evolve indeterministically, and with the addition to Bohmian mechanics the possibility of creating and annihilating particles in an indeterministic way. Each of these provides the wherewithal to treat standard field theory.

There have been extensive examinations of other ways to add additional variables to a noncollapse interpretation, largely under the rubric of *modal interpretations*. Both rules for specifying what the additional variables are and rules for the dynamics of the new variables have been investigated.

A THIRD WAY?

There are also some rather radical attempts to reject each of Bell's two options and to maintain both that the wave function, as given by the Schrödinger equation, is right and that it is everything—that is, it is descriptively complete. Since a wave function such as $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$ does not indicate that one outcome rather than the other occurred, this requires maintaining that it is not the case that one outcome rather than the other occurred.

This denial can come in two flavors. One is to maintain that neither outcome occurred, or even seemed to occur, and one is only somehow under the illusion that one did. David Z. Albert (1992) investigated this option under the rubric the *bare theory*. Ultimately, the bare theory is insupportable, since any coherent account must at least allow that the quantum mechanical predictions appear to be correct.

The more famous attempt in this direction contends that, in some sense, both outcomes occur, albeit in dif-ferent "worlds." Evidently, the wave function $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$ can be written as the mathematical sum of two pieces, one of which corresponds to a situation with the apparatus indicating "up" and the other to a situation with the apparatus indicating "down." The *many worlds theory* attempts to interpret this as a single physical state, which somehow contains or supports two separate "worlds," one with each outcome.

The many worlds interpretation confronts several technical and interpretive hurdles. The first technical hurdle arises because any wave function can be written as the sum of other wave functions in an infinitude of ways. For example, consider the apparatus state $1/\sqrt{2}\ |\text{“up”}>_A + 1/\sqrt{2}\ |\text{“down”}>_A$. Intuitively, this state does not represent the apparatus as having fired one way or another. This state can be called $|D_1>_A$. Similarly, $|D_2>_A$ can represent the state $1/\sqrt{2}\ |\text{“up”}>_A - 1/\sqrt{2}\ |\text{“down”}>_A$, which also does not correspond to an apparatus with a definite outcome. The state $1/\sqrt{2}|x\text{–up}>_S|\text{“up”}>_A + 1/\sqrt{2}|x\text{–down}>_S|\text{“down”}>_A$, which seems to consist in two "worlds," one with each outcome, can be written just as well as $1/\sqrt{2}|y\text{–up}>_S|D_1>_A + 1/\sqrt{2}|y\text{–down}>_S|D_2>_A$. Written in this way, the state seems to comprise two worlds: one in which the electron has *y*–spin up and the apparatus is not in a definite indicator state, the other in which the electron has *y*–spin down, and the apparatus is in a distinct physical state that is equally not a definite indicator state. If these are the "two worlds," then the measurement problem has not been solved, it has been merely traded as a single world without a definite outcome for a pair of worlds neither of which has a definite outcome.

So the many worlds theory would first have to maintain that there is a preferred way to decompose the global wave function into "worlds." This is known as the *preferred basis problem*.

A more fundamental difficulty arises when one tries to understand the status of the probabilities in the many worlds theory. In a collapse theory the probabilities are probabilities for collapses to occur one way rather than another, and there is a physical fact about how the collapses occur, and therefore about frequencies of outcomes. In an additional variables theory the probabilities are about which values the additional variables take, and there is a physical fact about the values they take and therefore about frequencies of outcomes. But in the many worlds theory, whenever one does an experiment like the spin measurement described earlier, the world splits: There is no frequency with which one outcome occurs as opposed to the other. And more critically, that the world "splits" has nothing to do with the amplitude assigned to the two daughter worlds.

Suppose, for example, that instead of feeding a *y*–spin up electron into our *x*–spin measuring device, we feed in an electron whose state is $1/2|x\text{–up}>_S + \sqrt{3}/2\ |x\text{–down}>_S$. By linearity, at the end of the experiment, the state of the system plus apparatus is $1/2|x\text{–up}>_S|\text{“up”}>_A + \sqrt{3}/2\ |x\text{–down}>_S|\text{“down”}>_A$. Even if we have solved the preferred basis problem and can assert that there are now two worlds, one with each outcome, notice that we are evidently in exactly the same situation as in the original experiment: Whenever we do the experiment, the universe “splits.” But the quantum formalism counsels us to have different expectations in the two cases: in the first case, we should expect to get an “up” outcome 50 percent of the time, in the second case only 25 percent of the time. It is unclear, in the many worlds theory, what the expectations are for, and why they should be different.

Another interpretation of the quantum formalism that has been considered is the *many minds theory* of Barry Loewer and Albert. Despite the name, the many minds theory is not allied in spirit with the many worlds theory: It is rather an additional variables theory in which the additional variables are purely mental subjective states. This is somewhat akin to Wigner's appeal to consciousness to solve the measurement problem, but where Wigner's minds affect the development of the wave function, the minds in this theory (as is typical for additional variables theories) do not. The physical measurement apparatus in the problematic case does not end up in a definite indicator state, but a mind is so constituted that it will, in this situation, have the subjective experience of seeing a particular indicator state. Which mental state the mind evolves into is indeterministic. The preferred basis problem is addressed by stipulating that there is an objectively preferred basis of physical states that are associated with distinct mental states.

The difference between the many worlds and the many minds approaches is made most vivid by noting that the latter theory does not need more than one mind to solve the measurement problem, where the problem is now understood as explaining the determinate nature of our experience. A multiplicity of minds are added to Loewer and Albert's theory only to recover a weak form of mind-body supervenience: Although the experiential state of an individual mind does not supervene on the physical state of the body with which it is associated, if one associates every body with an infinitude of minds, the distribution of their mental states can supervene on the physical state of the body.

A final attempt to address the problems of quantum mechanics deserves brief mention. Some maintain that the reason quantum mechanics is so confusing is not because the mathematical apparatus requires emendation (e.g., by explicitly adding a collapse or additional variables) or an interpretation (i.e., an account of exactly which mathematical objects represent physical facts), but because we reason about the quantum world in the wrong way. Classical logic, it is said, is what is leading us astray. We merely need to replace our patterns of inference with quantum logic.

There is a perfectly good mathematical subject that sometimes goes by the name *quantum logic*, which is the study, for example, of relations between subspaces of Hilbert space. These studies, like all mathematics, employ classical logic. There is, however, no sense in which these studies, by themselves, afford a solution to the measurement problem or explain how it is that experiments like those described earlier have unique, determinate outcomes.

THE WAVE FUNCTION, ENTANGLEMENT, EPR, AND NON-LOCALITY

For the purposes of this discussion, the wave function has been treated as if it were something like the electromagnetic field: a field defined on space. Although this is not too misleading when discussing a single particle, it is entirely inadequate when considering collections of particles. The wave function for N particles is a function not on physical space, but on the 3N-dimensional configuration space, each point of which specifies the exact location of all the N particles. This allows for the existence of entangled wave functions, in which the physical characteristics of even widely separated particles cannot be specified independently of one another.

Consider R and L, a pair of widely separated particles. Among the wave functions available for this pair is one that ascribes *x*–spin up to R and *x*–spin down to L, which is written as $|x\text{–up}>_R|x\text{–down}>_L$, and one that attributes *x*–spin down to R and *x*–spin up to L:$|x\text{–down}>_R|x\text{–up}>_L$. These are called *product states*, and all predictions from these states about how R will respond to a measurement are independent of what happens to L, and vice versa.

But besides these product states, there are entangled states like the *singlet state*: $1/\sqrt{2}|x\text{–up}>_R|x\text{–down}>_L - 1/\sqrt{2}|x\text{–down}>_R|x\text{–up}>_L$. In this state the *x*–spins of the two particles are said to be anticorrelated since a measurement of their *x*–spins will yield either up for R and down for L or down for R and up for L (with a 50 percent chance for each outcome). Even so, if the wave function is complete, then neither particle in the singlet state has a

determinate *x*–spin: the state is evidently symmetrical between spin up and spin down for each particle considered individually.

How can the *x*–spins of the particles be anticorrelated if neither particle has an *x*–spin? The standard answer must appeal to dispositions: although in the singlet state neither particle is disposed to display a particular *x*–spin on measurement, the pair is jointly disposed to display opposite *x*–spins if both are measured. Put another way, on the standard interpretation, before either particle is measured neither has a determinate *x*–spin, but after one of them is measured, and, say, displays *x*–spin up, the other acquires a surefire disposition to display *x*–spin down. And this change occurs simultaneously, even if the particles happen to be millions of miles apart.

Einstein found this to be a fundamentally objectionable feature of the standard interpretation of the wave function. In a paper coauthored with Boris Podolsky and Nathan Rosen (EPR 1935), Einstein pointed out this mysterious, instantaneous "spooky action-at-a-distance" built into the standard approach to quantum theory. It is uncontroversial that an *x*–spin measurement carried out on L with, say, an "up" outcome" will result in a change of the wave function assigned to R: It will now be assigned the state $|x\text{–down}>_{R}$. If the wave function is complete, then this must reflect a physical change in the state of R because of the measurement carried out on L, even though there is no physical process that connects the two particles. What EPR pointed out (using particle positions rather than spin, but to the same effect) was that the correlations could easily be explained without postulating any such action-at-a-distance. The natural suggestion is that when we assign a particular pair of particles the state $1/\sqrt{2}|x\text{–up}>_{R}|x\text{–down}>_{L} - 1/\sqrt{2}|x\text{–down}>_{R}|x\text{–up}>_{L}$, it is a consequence of our ignorance of the real physical state of the pair: The pair is either in the product state $|x\text{–up}>_{R}|x\text{–down}>_{L}$ or in the product state $|x\text{–down}>_{R}|x\text{–up}>_{L}$, with a 50 percent chance of each. This simple expedient will predict the same perfect anticorrelations without any need to invoke a real physical change of one particle consequent to the measurement of the other.

So matters stood until 1964, when Bell published his famous theorem. Bell showed that Einstein's approach could not possibly recover the full range of quantum mechanical predictions. That is, no theory can make the same predictions as quantum mechanics if it postulates (1) that distant particles, such as R and L, have each their own physical state definable independently of the other and (2) measurements made on each of the particles have no physical affect on the other. Entanglement of states turns out to be an essential feature—arguably the central feature—of quantum mechanics. And entanglement between widely separated particles implies non-locality: The physics of either particle cannot be specified without reference to the state and career of the other.

The spooky action-at-a-distance that Einstein noted is not just an artifact of an interpretation of the quantum formalism; it is an inherent feature of physical phenomena that can be verified in the laboratory. A fundamental problem is that the physical connection between the particles is not just spooky (unmediated by a continuous space-time process), it is superluminal. It remains unclear to this day how to reconcile this with the theory of relativity.

See also Bohm, David; Bohmian Mechanics; Many Worlds/Many Minds Interpretation of Quantum Mechanics; Modal Interpretation of Quantum Mechanics; Non-locality; Philosophy of Physics; Quantum Logic and Probability.

Bibliography

Albert, David Z. *Quantum Mechanics and Experience*. Cambridge, MA: Harvard University Press, 1992.

Bell, John S. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. Cambridge, U.K.: Cambridge University Press, 1987.

Dürr, Detlef, Sheldon Goldstein, and Nino Zanghi. "Quantum Equilibrium and the Origin of Absolute Uncertainty." *Journal of Statistical Physics* 67 (1992): 843–907.

Ghirardi, GianCarlo, Alberto Rimini, and Tulio Weber. "Unified Dynamics for Microscopic and Macroscopic Systems." *Physical Review* 34 (2) (1986): 470–491.

Maudlin, Tim. *Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics*. Malden, MA: Blackwell, 2002.

Von Neumann, John. *Mathematical Foundations of Quantum Mechanics*. Translated by Robert T. Beyer. Princeton, NJ: Princeton University Press, 1955.

Wheeler, John Archibald, and Wojciech Hubert Zurek, eds. *Quantum Theory and Measurement*. Princeton, NJ: Princeton University Press, 1983.

Wigner, Eugene. *Symmetries and Reflections*. Westport, CT: Greenwood Press, 1967.

Tim Maudlin (2005)

QUESTIONS

See Appendix, Vol. 10

QUINE, WILLARD VAN ORMAN
(1908–2000)

Willard Van Orman Quine, an Edgar Pierce professor of philosophy at Harvard, was born in Akron, Ohio. In 1930 he was graduated from Oberlin, where he majored in mathematics, and he wrote a doctoral dissertation in logic under Alfred North Whitehead at Harvard. He visited Vienna, studied mathematical logic at Warsaw, and at Prague met Rudolf Carnap, whose work was to inspire and influence him.

Some of Quine's publications are in philosophy, some in symbolic logic, and others are concerned with the logical regimentation of ordinary language. It is his philosophy and related aspects of his advocated regimentation of language that concern us here, his contributions to logic being dealt with elsewhere.

ANALYTIC-SYNTHETIC DISTINCTION

Some philosophers have attempted to distinguish between such statements as "A river flows through Brisbane," which, they contend, are true as a matter of fact, and statements like "No bachelor is married," the truth of which is said to be independent of matters of fact. The former have been described as synthetic, the latter as analytic. Quine maintained, first, that the analytic-synthetic distinction has never satisfactorily been made and, second, that there is no good reason for believing that it can be made.

LOGICAL TRUTH. Given a list of logical particles and the notion of truth, with which Quine was comparatively satisfied, we may, he contends, derive the notion of logical truth. "All birds are birds" is logically true because it is both true and such that if we leave its logical parts alone and replace "birds" with some other word, then if we get a statement at all, we get a true one—for example, "All snakes are snakes." But even though this analytic statement is logically true, there are analytic statements like "No bachelor is married" that are not, and thus analyticity remains to be explained. If we replace "bachelor" with the synonymous "unmarried man," we have a logical truth, and it would thus appear that an analytic statement either is a logical truth or is reducible to one by interchange of synonyms.

SYNONYMY. However, according to Quine, an account of analyticity that depends on the notion of synonymy is unsatisfactory. Suppose that all and only Guards officers are very tall soldiers with long hair. Since "Guards officers" and "very tall soldiers with long hair" are coextensive expressions, there are statements whose truth or falsity cannot be affected by interchanging these expressions. But because they are not synonymous expressions, there are also statements like "Necessarily, all and only Guards officers are Guards officers" that can be so affected. In contrast, the truth of the statement "Necessarily, all and only bachelors are bachelors" cannot be affected by interchanging "bachelors" and "unmarried men" because these expressions are synonymous. But to make the last statement is to say that "All and only bachelors are bachelors" is analytic. Thus, we give an account of synonymy in terms of the effects of interchanging expressions in certain contexts. But because these contexts cannot be specified without reference to analyticity or some equivalent notion, we cannot, without circularity, use the notion of synonymy in giving an account of analyticity. Similar difficulties frustrate the derivation of self-contradictoriness from logical falsity.

Quine also discusses the possibility of giving an account of the analyticity of statements in artificial languages, but here, as in natural languages, the difficulty is, he contended, that each of the key notions in the theory of meaning is definable only in terms of the others.

Anyone who produced an account of these notions acceptable to Quine would thereby refute him, but what sort of account this would be remains to be seen. In the meantime the strongest argument against him is ad hominem. "All the illuminated manuscripts are illuminated" is logically true only if "illuminated" has the same meaning in each of its occurrences. Thus, the notion of logical truth, which Quine accepts, is dependent upon the notion of synonymy, which he rejects.

RADICAL TRANSLATIONS. Quine's theory of meaning was further developed in his discussion of the difficulties that would arise if we were to attempt to translate the language of a hitherto isolated tribe. Radical translation, as he calls it, would have to begin not with words but with those sentences that have a comparatively direct relation to stimulus conditions. The stimulus meaning of a sentence for a person is defined in terms of the class that has as its members the kinds of stimulation that would prompt the person's assent to the sentence. Intrasubjective stimulus synonymy is sameness of stimulus meaning for one speaker, and two sentences are socially stimulus-synonymous if they are intrasubjectively stimulus-synonymous for nearly everyone who speaks the language. A sentence is stimulus-analytic for a person if he would

assent to it, if to anything, after every stimulation, and a socially stimulus-analytic sentence is stimulus-analytic for nearly every speaker of the language.

In order to see that these are not our intuitive notions of synonymy and analyticity, we need to distinguish occasion sentences and standing sentences. If every minute or so we are asked to assent to "John has hiccups," we cannot do so without having another look at John on each occasion. In contrast, having once assented to the standing sentence "Salt is soluble in water," we may assent again without observing salt or anything else again. Applied to occasion sentences, intrasubjective stimulus synonymy approximates sameness of meaning; standing sentences, however, are related to experience indirectly, and the kinds of stimulus that would prompt assent to a standing sentence vary from speaker to speaker. Thus, the stimulus meaning of a standing sentence falls short of our intuitive notion of meaning; stimulus synonymy is correspondingly inadequate, and some socially stimulus-analytic sentences would normally be described not as analytic but as conveying information common to the whole community.

Quine demands of those who talk of analyticity and synonymy that they give of their concepts the sort of account in terms of dispositions to verbal behavior that he has given of his.

By observing and testing native speech behavior dispositions, the linguist can come to translate some occasion sentences and to recognize stimulus analyticity and synonymy. But in order to complete the radical translation of a language, he must frame analytical hypotheses. This consists of segmenting what he hears into native words and hypothetically equating these to English expressions. Quine contends that there will be many sets of analytical hypotheses that fit all native dispositions to speech behavior and yet lead to incompatible translations of countless sentences in their language. Suppose that, observing the circumstances in which a native utters "Gavagai," we translate this sentence as "Rabbit!" Whether the word *gavagai* is to be taken to apply to rabbits, temporal stages of rabbits, or something even stranger to us can be settled only when we can ask questions like "Is this the same rabbit as that?" This cannot be done until we have translated the parts of speech that make up the native system of reference, and since this is part of what we do when we adopt a set of analytical hypotheses, there is more than one way of doing it. For example, the sentence translated as "Is this (the same) (rabbit) as that?" might, on another set of empirically satisfactory hypotheses, be translated as "Is this (a rabbit stage) (of the same series) as that?"

In this way Quine arrives at the principle of the indeterminacy of translation, which says that it is possible to compile incompatible manuals for translating one language into another, all of which fit all observable speech dispositions, and that there is no sense in asking which is the right manual. It is only in exceptional cases that we can talk of the meaning of a single sentence, and when our statements about the world conflict with experience, they do so not individually but as a system. Thus, we have what might be called the Quine-Duhem conventionalist thesis that any statement can be held to be true no matter what is observed, provided that adjustments are made elsewhere in the system; it is from this thesis that Quine infers that it is impossible to make the analytic-synthetic distinction.

Quine believed that his discussion of radical translation reveals the possibility of differences between the conceptual schemes of people that are not empirically conditioned. In the case of two compatriot linguists working independently on the radical translation of a language, one linguist might conclude that he and the native see the world in the same way, as consisting of tables, chairs, ducks, and rabbits, while the other finds that the native speaks of rabbit stages, not of rabbits, and concludes that the native's outlook is different from his own. Now, in order to determine what the native's outlook really is, it is necessary to discover which is the correct way of translating the native's language. But according to the principle of the indeterminacy of translation, it does not make sense even to ask this, and consequently it cannot make sense to ask what the native's outlook is. It can be shown that the native is in no better position than the linguist here, and it then becomes hard to see the sense of talking about an outlook when there is no conceivable way of discovering what this outlook is. Quine's position here is not clear. He admitted that these differences of outlook are in principle undetectable and grants that such cultural contrasts are threatened with meaninglessness, but he continued to speak of them.

As radical translation is not known ever to have been undertaken, the absence of incompatible manuals of translation does not count against the principle of indeterminacy. Nevertheless, it might well be contended that until there are more conclusive arguments for it, the principle is to be taken as the incredible consequence of unsound premises. Quine, in discussing meaning, did concentrate on the statement-making function of language, and it has, in fact, been argued that by neglecting

the countless other uses of language, he arrived at a concept of synonymy the inadequacy of which is revealed by the fact that it makes translation indeterminate.

ONTOLOGY

Philosophers have disagreed as to what there is; some have held, for example, that there are only material things, and others have denied this. Quine called such theories "ontic theories" and maintained that they are a part of the sciences distinguished only by extreme generality. Given that there are physical objects, it is the natural scientist who discovers whether there are wombats; and given classes, it is the mathematician who finds out whether there are even prime numbers. Whether there are physical objects and classes, however, is the concern of the philosopher. The integration of established theories, which is one of the aims of scientific work, may lead to any one of many equally satisfactory accounts of the world, each with its ontic theory, and there is no sense in asking which of these accounts is the true one. Thus, Quine took a conventionalist view even of the theses of ontologists.

Today it is commonly maintained that since there is no way of settling an ontic dispute, ontologists have unwittingly concerned themselves with pseudo questions. Quine, in proposing a method of determining the ontic import of a theory, attempted to make such questions decidable and thus real. His method was, in outline, as follows: "$(\exists x)$(x is a cat)" may be read as "There is an x such that x is a cat" or as "There is something such that it is a cat." According to Quine, anyone who makes this statement is thereby committed to the existence of cats. The statement consists of the existential quantifier "$(\exists x)$," the predicate "——is a cat," and an "x" that works like a pronoun and is needed in any but the simplest cases to show under which quantifier a predicate comes. If we add to this equipment such truth-functional words as "*and* and *not,* we can make statements like "$(\exists x)$(x is a book, and x is boring), and $(\exists x)$(x is a book, and x is not boring)." This is a paraphrase of "Some but not all books are boring," which, it is alleged, reveals the ontic import of this statement. Bertrand Russell, Quine, and others have suggested similarly revealing paraphrases of general hypotheticals, of statements containing proper names, and of statements containing such descriptive phrases as "the prime number between 5 and 11." Quine contended that in adopting any theory, we commit ourselves to the existence of certain entities and that by translating the theory into a language in which the only formal devices are predication, quantification, and truth-functional composition, we make these commitments explicit.

ONTIC COMMITMENTS. The commitments revealed in the above manner are incurred when certain words are used in certain ways. We are, according to Quine, committed to the existence of physical objects because of the ways in which physical object terms function in our language. In contrast, we are not committed to such objects as "sakes," because even though we do some things for the sake of others, "sake" functions in only a few of the ways in which a term does. When constructing theories, we are, within limits, free to decide what expressions will function as terms, and by such decisions we might commit ourselves to the existence of atoms, for example, but not to that of meters. We accept the reality of physical objects more readily than we do that of atoms because typical sentences about physical objects are more closely associated with sensory stimulation than are typical sentences about atoms. By this criterion sense data are even more acceptable than physical objects, but this is counteracted by the fact that sense data are a less satisfactory basis for an account of the world. On the grounds of utility for theory, classes are to be preferred to attributes and sentences to propositions.

Many would maintain that it is only when Quine is discussing the considerations that influence ontic decisions that he tackles philosophical problems, and that he does this in a way he himself admitted to be sketchy. He does this sketchily because it has been done in detail by others to whom he refers, and believing that ontologists must take account of scientific theories, he is especially interested in working out how this is to be done. Perhaps the major philosophical problem raised by Quine's proposed criterion of ontic commitment is that of the nature of this commitment: I may know what it is like for a nation to be, or not to be, committed to an isolationist foreign policy, but what is it like to be, or not to be, committed to the existence of physical objects?

REGIMENTATION OF ORDINARY LANGUAGE. The regimentation of language serves purposes other than that of revealing ontic commitments. The logic of ordinary language is difficult to formulate, and consequently it is more economical to theorize in a language that is ordinary except in its logical parts, which are designed to facilitate deduction. And if there are fewer kinds of construction and less obscurity in a regimented language, then in moving into it we simplify and clarify our conceptual scheme.

Because of misgivings about synonymy Quine cannot maintain that for an ordinary-language sentence to be replaced by a regimented one, the two must be synonymous. Indeed, we may be making the replacement just because one sentence is ambiguous and the other is not. Paraphrase into a regimented language consists, he maintains, of replacements that, in certain contexts, forward certain programs. Against this it has been argued that for any two sentences there will be a program that is forwarded by replacing one with the other, and consequently Quine's notion of paraphrase is vacuous unless contexts and programs can be specified. If this can be done, however, the notion of sentence synonymy can be derived. This notion is no less satisfactory, and no more difficult to make adequate sense of, than the notion of paraphrase, without which Quine cannot talk of putting theories into a regimented language.

The bulk of Quine's philosophical work was published after 1947. By 1960 he had combined into a coherent position theses some of which were first put forward ten years earlier. Between 1947 and 1960 certain changes in his views occurred. From declaring, in 1947, that he did not believe in abstract entities, he had come not only to accept such entities but also to claim that he had always done so; from counting phenomenalism, in 1948, as a conceptual scheme suitable for certain purposes, he came to reject it; and from maintaining, in 1951, that in the face of recalcitrant experience we could change our logical laws, he had apparently come to hold that there is nothing that would count as changing our logical laws.

Quine's status as a philosopher never depended upon the number of people who agreed with him. On the contrary, the sign of his achievement is the valuable discussion he provoked by his persistent and penetrating attacks on analyticity and related notions and by his unfashionable conviction that philosophers want to discover what reality is like.

See also Analytic and Synthetic Statements; Analyticity; Artificial and Natural Languages; Carnap, Rudolf; Logic, History of; Ontology; Philosophy of Language; Synonymity; Underdetermination Thesis, Duhem-Quine Thesis; Whitehead, Alfred North.

Bibliography

The topics of the two main sections above were first treated in detail in "Two Dogmas of Empiricism" (1951) and "On What There Is" (1948), reprinted in Quine's *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953). A definitive statement of Quine's position, including an exposition of the indeterminacy thesis, is in his *Word and Object* (New York: MIT Press, 1960), in which a list of Quine's many other philosophical publications will be found. His views on analyticity and logical truth are briefly stated in his "Mr. Strawson on Logical Theory," in *Mind* 62 (1953): 433–451. Important earlier articles are his "Truth by Convention," in *Philosophical Essays for Alfred North Whitehead* (New York: Longmans, Green, 1936), reprinted in Herbert Feigl and Wilfrid Sellars, eds., *Readings in Philosophical Analysis* (New York: Appleton-Century-Crofts, 1949), and "Steps toward a Constructive Nominalism," in *Journal of Symbolic Logic* 12 (1947): 105–122, which he wrote with Nelson Goodman.

Among the many discussions of Quine's views are the following, which are the sources of critical points made in this article. H. P. Grice and P. F. Strawson, "In Defence of a Dogma," in *Philosophical Review* 65 (1956): 141–158, defend the analytic-synthetic distinction. P. F. Strawson, "Propositions, Concepts and Logical Truths," in *Philosophical Quarterly* 7 (1957): 15–25, is an attack on Quine's notion of logical truth. Comments on indeterminacy are made in L. J. Cohen, *The Diversity of Meaning* (London: Methuen, 1962), pp. 67–74. Quine's views on ontology are criticized by G. J. Warnock in "Metaphysics in Logic," in *Essays in Conceptual Analysis,* edited by A. G. N. Flew (London: Macmillan, 1956), pp. 75–93, and from a different standpoint in Rudolf Carnap, *Meaning and Necessity* (Chicago: University of Chicago Press, 1947). The relation between conventionalism and the analytic–synthetic distinction is discussed in G. H. Herburt, "The Analytic and the Synthetic," in *Philosophy of Science* 26 (1959): 104–113. Some points made above are also developed in C. F. Presley, "Quine's *Word and Object,*" in *Australasian Journal of Philosophy* 39 (1961): 175–190.

C. F. Presley (1967)

QUINE, WILLARD VAN ORMAN [ADDENDUM]

Willard Van Orman Quine, the Edgar Pierce Professor of Philosophy Emeritus, at Harvard, author of twenty-one books and scores of journal articles and reviews, made many significant contributions to metaphysics, epistemology, philosophy of language, philosophy of science, philosophy of mind, logic, philosophy of logic, and set theory, and ethics (and ethical theory). These contributions are of a stature that firmly places Quine among the titans of twentieth-century Anglo American philosophy.

In most of his publications following *Word and Object* (1960), Quine sought to sum up, clarify, and expand on various themes found in that book. Quine can occasionally be seen changing his mind regarding some detail of his prior thought, but by and large he remains remarkably consistent.

NATURALISM

The keystone of Quine's systematic philosophy is naturalism. Roughly, naturalism is the view that there is no suprascientific justification for science *and* that it is up to science to determine both what there is (ontology) and how we know what there is (epistemology). Moreover, Quine maintains that the best current science tentatively and fallibly plumps for a physicalist ontology and an empiricist epistemology.

ONTOLOGY: PHYSICALISM. Since he maintains that what a (formalized) theory says there is is determined by the range of values of the bound variables of that theory, and since the bound variables of the best current scientific theory of the world (viz., physics) range over both physical objects and numbers, then, given his naturalism, Quine's physicalism embraces both concrete objects and abstract objects. He is a scientific realist regarding (observable and unobservable) physical objects and a Platonic realist regarding numbers (or sets). However, in *Pursuit of Truth* (1980) Quine downgrades the philosophical importance of ontology, including physicalism. He does so because of ontological relativity (i.e., indeterminacy of reference). The thesis is that a theory's ontology can be supplanted *salva veritate* by any one-to-one mapping of it. Ontological relativity thus engenders an attitude of indifference toward various equally apt ontologies for a given theory, including physical theory so called. At the same time it highlights the importance of a theory's ideology, that is, its lexicon of predicates. The philosophical point of Quine's thesis is, then, that what a theory says there is is less important to our understanding of the world than what a theory says about what there is.

There are two further senses in which Quine may be said to be a physicalist. First, as expected, he rejects Cartesian dualism of mind and body in favor of materialism. In this regard, he endorses Donald Davidson's anomalous monism: token identity, type diversity. Second, he is a physicalist in the sense in which physicalism is opposed to phenomenalism in epistemology (see below).

EPISTEMOLOGY: EMPIRICISM. If the best current scientific theory (tentatively and fallibly) proffers a physicalist answer to the question of what there is, then what does it proffer in response to the question of how we know what there is? The answer is, in a word, empiricism. Quine maintains that it is a finding of science that all that we come to know about the world begins with the activation of our nerve endings.

So, Quine endorses the naturalization of both ontology and epistemology. And although he downgrades the philosophical importance of ontology, he maintains the philosophical importance of epistemology. The central question of epistemology, according to Quine, is How do we acquire our theory of the world and why does it work so well? Any answer to this question must explain the relation between one's empirical data (the "meager input") and one's theory of the world (the "torrential output"). Much of what Quine wrote after *Word and Object* is, ultimately, devoted to answering this question. His own distinctive answer may be called externalized empiricism in order to differentiate it from approaches of other naturalized epistemologists (e.g., Donald Davidson). Quine's empiricism is externalized in the sense that he takes sets of activated nerve endings as his data and sets of sentences as his theory of the world (as opposed, say, to impressions and ideas, respectively).

In Quine's hands, the general relation, R_1, holding between sets of activated nerve endings and sets of sentences gets analyzed into two relations. There is the causal relation, R_2, holding between holophrastically construed observation sentences and their respective patterns of activated nerve endings, and there is the logical relation, R_3, holding between those same observation sentences, now analytically construed, and standing sentences. Quine schematizes how the child or the race, beginning with verbal responses conditioned to their respective patterns of nerve endings (R_2), could have gone on to achieve verbal reference to bodies, substances, unobservables, and abstract objects (R_3). Moreover, his account of R_3 explains how observation sentences are logically related to theoretical sentences in such a way that no bridge principles are needed for linking observation and theoretic sentences. His account also highlights the hypothetico-deductive method of prediction and falsification and the moderately holistic character of theory revision.

RECIPROCAL CONTAINMENT. Externalized empiricism is Quine's contribution to answering the central epistemological question of how we acquire our theory of the world and why it works so well. As such, his epistemology (empiricism) "contains" his ontology (physicalism): *nihil in mente quod non prius in sensu.* However, Quine's epistemologizing always takes place within some accepted theory of the world (the best one he can muster at the time), so his epistemology (empiricism) is itself contained within his ontology (physicalism). This latter containment is the central lesson of naturalism: There is no first philosophy. It is this latter containment that also

makes Quine's epistemology such a radical departure from the tradition.

CHANGES OF MIND

Even though Quine's thought has been remarkably consistent since his first works appeared in the 1930s, he changed his mind on a few important matters. First, he downgraded the importance of ontology, discussed above. Second, in the context of radical translation, Quine dropped the idea that the linguist can translate the native's "Gavagai" as her own "Lo, a rabbit" just in case the native's stimulus meaning for "Gavagai" is approximately the same as the linguist's for "Lo, a rabbit." The problem is with making scientific sense of this "implicit homology assumption" regarding different people's nerve endings. Quine changed to the position that the linguist can translate the native's "Gavagai" as her own "Lo, a rabbit" just in case the linguist can empathize with the native to the extent that she can confidently conjecture that, were she in the native's position when he uttered (or assented to) "Gavagai," then she would have done likewise for "Lo, a rabbit." In this way the linguist is (tentatively) equating the native's "Gavagai" with her own "Lo, a rabbit" without relying on an implicit homology assumption. Third, since, according to Quine's externalized empiricism, the meager input underdetermines the torrential output, then it is conceivable that there could be two (or more) global theories of the world that are empirically equivalent, logically compatible, equally simple, and so forth. Would both be true? Quine's empiricism encourages an ecumenical response: Both would be true. His naturalism encourages a sectarian response: Only one would be true. Quine himself vacillated on the issue but eventually endorsed the sectarian response. This suggests that his commitment to naturalism runs deeper than his commitment to empiricism.

See also Anomalous Monism; Davidson, Donald; Empiricism; Epistemology; Ethics; Logic, History of; Materialism; Metaethics; Metaphysics; Naturalism; Naturalized Epistemology; Ontology; Phenomenalism; Philosophy of Language; Philosophy of Mind; Philosophy of Science, History of; Philosophy of Science, Problems of; Physicalism; Reference; Set Theory; Subject and Predicate.

Bibliography

WORKS BY QUINE

Among Quine's post–*Word and Object* books that bear directly on the topics discussed here are:

The Ways of Paradox and Other Essays. New York: Random House, 1966; enlarged ed., Cambridge, MA: Harvard University Press, 1976.

Ontological Relativity and Other Essays. New York: Columbia University Press, 1969.

Roots of Reference. La Salle, IL: Open Court, 1974.

Theories and Things. Cambridge, MA: Harvard University Press, 1981.

Other of Quine's later books of interest are:

Set Theory and Its Logic. Cambridge, MA: Harvard University Press, 1963; rev. ed., 1969.

Philosophy of Logic. Englewood Cliffs, NJ: Prentice Hall, 1970; Cambridge, MA: Harvard University Press, 1986.

The Web of Belief. With J. S. Ullian. New York: Random House, 1970; rev. ed., 1978.

The Time of My Life. Cambridge, MA: MIT Press, 1985.

Quiddities. Cambridge, MA: Belknap Press of Harvard University Press, 1987.

WORKS ON QUINE

The secondary literature on Quine is immense and still growing. It includes:

Barrett, R., and R. Gibson, eds. *Perspectives on Quine.* Cambridge, MA: Blackwell, 1990.

Davidson, D., and J. Hintikka, eds. *Words and Objections.* Dordrecht: Reidel, 1969.

Gibson, R. F., Jr. *The Philosophy of W. V. Quine: An Expository Essay.* Tampa: University Presses of Florida, 1982.

Hahn, L. E., and P. A. Schilpp, eds. *The Philosophy of W. V. Quine.* La Salle, IL: Open Court, 1986.

Hookway, C. *Quine: Language, Experience, and Reality.* Stanford, CA: Stanford University Press, 1988.

Romanos, G. D. *Quine and Analytic Philosophy.* Cambridge, MA: MIT Press, 1983.

R. F. Gibson (1996)

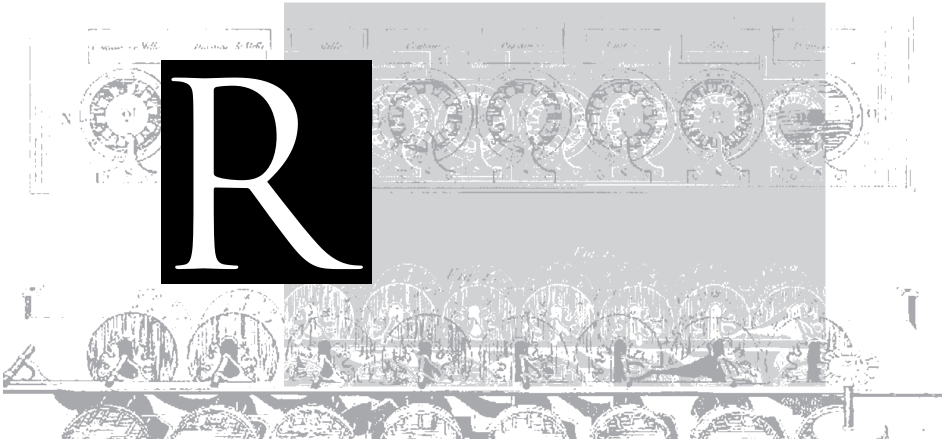

RACISM

"Racism" is the doctrine that one group of men is morally or mentally superior to another and that this superiority arises out of inherited biological differences. Of the modern theories aimed at dividing one portion of humanity from another, it is the most morally reprehensible and the least substantially based. Nationalism has a certain rationale in the existence of nation-states, and it does not, at least not necessarily, imply the inferiority of one nation to another. The various doctrines of the struggle between economic classes can point to a wide assortment of empirical evidence in support of their claims; in the Marxist version the exploiting capitalist is as much a victim of the capitalist system as is the exploited proletarian, and the eventual overcoming of all class distinctions is a moral aim as well as a prophesied event. The tenets of racism, however, lead to moral conclusions that contradict many of the most generally accepted civilized standards and have notoriously led to what on ordinary grounds are inconceivable crimes. It might be claimed that ordinary standards are mistaken and that, for example, it was morally imperative that the Nazis exterminate the Jews—if racist claims had a substantial factual basis. Fortunately for ordinary moral standards, if not for the exterminated Jews and other victims of racial persecution, the tenets of racism are not merely unsubstantiated by the facts but in large measure contradicted by the facts.

Nor have the most important racist theorists been equipped to judge the alleged facts on which they based their claims. The question of race is an enormously complex one, and a judgment on it requires a synthesis of materials from history and prehistory and from a wide variety of biological, anthropological, and psychological disciplines, but primarily from genetics. Many of the necessary facts have only recently become available, and major questions remain unanswered. Yet most racist theories were put forth prior to the accumulation of this evidence, and even most contemporary racist theories are based on outdated biology. Furthermore, most racists—Houston Stewart Chamberlain, with his varied but erratic education, is a possible exception—have lacked the scientific training required to judge whatever evidence was available at the time they wrote. And until a racist theory can be substantiated to a very high degree of probability, the unsavoriness of the conclusion that there are inequalities in the capacities of groups of men requires that the theory be rejected.

OUTLINE OF THE THEORY

Although there are many variations on the racist theme (the number of contradictions among racist claims, notably about which are the privileged races, is enough in itself to cast doubt on the tenability of the whole racist enterprise), a model set of racist tenets, divisible into three groups of claims, can be isolated.

The first group starts with the premise that humankind is now, has been in the past, or ought to be in the future divided up into biologically distinct groups. The different tenses must be distinguished because in some instances the claim is made that the superior race is not now in existence but should be bred from the "best blood" among various existing groups. This claim is the link between racism and eugenics, but although eugenicists often fall into racist language or hold racist beliefs (Sir Francis Galton, the founder of eugenics, rated blacks as about two grades below the Anglo-Saxon "race," and the British pragmatist philosopher F. C. S. Schiller supported both eugenics and the English fascist Sir Oswald Mosley), the connection between the two theories is not inevitable.

The distinction between groups of humankind is held to be based on the common biological heredity of the members of each group. Among the biological distinctions between groups are inherited capacities for certain cultural activities—some races, it is claimed, are more warlike than others, some more musical, others predestined to be dominated. These are factual claims, seemingly open to confirmation or refutation by a scientific examination of the evidence, and the evidence seems overwhelmingly against every one of them. Someone who upheld these views would not necessarily be a racist, but they are essential to the racist position.

In a class by itself is the claim that the mechanism of transmission of group characteristics is the blood. Of all racist claims this is the one most surely refuted, and it would seem to be inessential to the doctrine. Yet the insistent stress on this claim even in the face of overwhelming evidence of its falsity is an index of the nonrational sources of racist thinking. Theories of inheritance through the blood, of blood kinship, of bluebloods, and of good and bad blood are survivals of age-old prescientific thought, on the same order as the view that the soul is the breath.

The final set of doctrines are essential to racism and distinctive of it. Not only are human groups different from one another but some are "better," "stronger," "higher," or "more creative" than others—physically, intellectually, or morally. (The proponent of a particular racist doctrine quite naturally almost always identifies himself with the race he judges superior. Thus, Comte Joseph Arthur de Gobineau, who was born in the south of France and who placed the "Nordic race" at the pinnacle of humanity, devoted considerable research to proving his own descent from the Viking Otto Jarl; the British-born, French and Swiss-educated Chamberlain, who espoused Aryan or Teutonic superiority, included the contemporary English—and the Slavs and Celts—among the Teutons.) The higher race or races, it is claimed, have a moral right to dominate, to enslave, or even to eradicate the lower races. Finally, higher and lower races should not intermarry. Race mixture, or "mongrelization," is against nature. For the superior race it can lead only to the lowering of standards and to racial degeneration. It would seem that race mixture would improve the "lower" race, but this is generally denied either on biological or on historical grounds. Thus, Chamberlain held that the "lower" Jewish race was not improved by an alleged ancient admixture of Aryan blood, which came too little and too late.

CRITICISM

No complete examination of the fallacies of racist doctrines can be presented here. What seems most important is that there are not now and, so far as anthropological evidence shows, have never been any pure races of men and that the very concept of race as applied to groups of human beings is suspect. In the vast number of its traits humankind is one, and there has been constant intermarriage and a consequent diffusion of genetic traits throughout the species. There are obvious dissimilarities among groups of people, but these differences more or less gradually shade off into one another; it is a question of statistical predominance of certain physical or physiological traits in a population rather than of sharply defined group differences.

Estimates of the number of genes in man range from 10,000 to 100,000, whereas the number of genes that control skin color, shape of lips and nose, and hair form are few. Racists want to correlate these obvious differences—which in themselves are purely statistical and thus no certain guide to the ancestry of a particular person—with differences in innate inheritable mental characteristics. Yet the evidence is against any such correlation. Each gene or gene cluster, except for certain linked genes, is inherited individually; on the average, half comes from the father and half from the mother. The number of possible combinations of ancestral genes is astronomical, and the

question whether specific mental characteristics are linked with a particular genetic heritage can almost certainly be answered in the negative for human beings, if not, perhaps, for certain domestic animals.

In any case, humankind has apparently been faced with an environment that puts a premium on intelligence, and there seem to be no detectable group differences in intelligence. It is practically impossible to devise a satisfactory test to determine whether there are biological differences in intelligence. In most cases the available methods of classifying by ancestry those to be tested are quite fallible. It is equally difficult to find two groups genetically distinct and culturally alike, and intelligence tests are quite generally distorted by cultural factors and place a premium on particular cultural achievements that obscures any possible genetic factors in the results.

Finally, if there were any evolutionary reasons for thinking that some race was at one time constitutionally better fitted to one environment than another, the rate of human cultural change is such that this supposed superiority would have been insignificant for many centuries. There is no reason to think that one group of humankind is mentally or physically better fitted than another to cope with the complexities of modern urban civilization and an internationally dispersed technology.

To the above summary and inadequate account of the biological claims that contradict racism should be added the overwhelming historical evidence of constant migrations and intermarriages of human groups and the highly probable inference that movement and mixture was also the rule during the prehistory of the human species. This has been especially true of the two alleged races most notoriously prominent in racist literature, Aryans and Jews. The Aryan is generally presented as a pure and superior race and the Jewish "race" as inferior, contradictorily characterized as both pure and bastardized, often by the same author. However, there neither is nor could be evidence that either race is more or less "pure" than the other. Each group is an amalgam of people of varied ancestry, and mixture has produced no apparent genetic debilitation of the sort that racists inveigh against when they deplore the "mongrelization of the race." Cultural differences exist between Germans and Jews, but there are likewise cultural differences between different groups of Germans and between groups of Jews, as well as cultural similarities between German and Jewish groups. To assign these likenesses and similarities to race rather than to a vast complex of recognized sociocultural factors is to ignore a great bulk of historical evidence.

THE IRRATIONALISM OF RACISTS. Arguing with a proponent of racism is like arguing with someone who would today claim that the earth is flat and at the center of the universe. The evidence that the earth is round is so overwhelming, and so bound up with our very conception of what physical science is, that in the face of someone who claims that the earth is flat we can only point helplessly at the great body of scientific factual claims and scientific laws and ask, "But don't you see?" Similarly, when we are faced with the claims of a racist who persists in his doctrine in the face of our very notions of what constitutes biology and what constitutes historical research, we have no common ground for argument with him. An extreme but typical racist statement can be used as an example:

> It is established for all time: "alien albumen" is the sperm of a man of alien race. The male sperm is partially or completely absorbed by the female and thus enters her bloodstream. One single cohabitation of a Jew with an Aryan woman is sufficient to poison her blood forever. Together with the "alien albumen" she has absorbed the alien soul. Never again will she be able to bear purely Aryan children … they will all be bastards. (Julius Streicher, quoted in Quentin Reynolds, Ephraim Katz, and Zwy Aldouby, *Minister of Death,* New York, 1960, p. 150)

To someone with the most elementary acquaintance with contemporary biology it is unnecessary to point out the false assumptions and false statements in this quotation. But to refute the argument in a way that would satisfy its maker is impossible, because he denies the very grounds on which a scientific refutation as we understand it could be based.

The racist views in Adolf Hitler's *Mein Kampf* likewise seem based on a different biology from the one we know, but in Hitler's thought there is an added historical dimension. The picture Hitler draws of the sociopolitical situation in Germany and Austria during his own lifetime is often shrewd, but it is open to rational criticism: he makes factual claims that can be shown to be historically untrue and historical interpretations that can be challenged by an appeal to evidence and probability. His picture of the Aryans as the only culture-creating people, whose presence in a certain area at a certain time can be demonstrated simply because cultural innovation must have taken place then and there, bears no relation to what we know of the movements of peoples or to our notions of probability. In the chapter "Nation and Race," Hitler

uses few examples, and when examples are given they are used tendentiously to show what they could not prove. Thus, the culture of contemporary Japan, he claims, is the product of European stimulation: It is Western culture and technology with Japanese trimmings. Without continued infusions of Western culture, the culture of contemporary Japan is doomed to decay, and the culture found in Japan by Western explorers must itself have been the ossified remnants of some earlier, but forgotten, Aryan invasion. Hitler's arguments do not generally reach even the level of this one, circular as it is. Yet to show that no such invasion took place in historical times, and probably could not have taken place in prehistoric times, seems no answer to Hitler's claims. The picture he presents of the past is a deliberately mythical one, on a deliberately mythical time scale that bears no apparent relation to the known events and temporal ordering of history. In the absence of such relationships, all appeals to facts become irrelevant, and facts are notably absent from the argument.

RACISM OUTSIDE GERMANY

Although racism as a fully articulated doctrine and the central feature of official policy is notoriously associated with Germany, it has been powerful elsewhere. It was among thinkers of the French Enlightenment—Comte de Boulainvilliers, Comte de Buffon, and Baron de Montesquieu—that the concept of race was first made explicit and the germs of racism were implanted. Gobineau, in the mid-nineteenth century, was the true originator of the doctrine of racism, and throughout the nineteenth century and later, French thinkers vied with one another to show their descent from Gauls, Romans, Gallo-Romans, Celts, or Teutons and the superior Frenchness of one of these purported races over another.

In the United States and England also racism has flourished, and in these countries the complex interconnection of racist doctrines with social and economic factors is most apparent. In English thought racism has been mainly a concomitant of imperialism. The influx of darker-skinned peoples from the Commonwealth has led both to widespread resentment and to the expression of racist sentiments, but not as yet to any new fully developed racist theories. In the United States racism first arose in the South as a defense of slavery, was invoked as a justification of American imperialist expansion into the western Pacific and the Caribbean and for the restriction of the immigration of "undesirable" stock into the United States, and arose again as a defense of segregation.

Twentieth-century arguments that blacks were biologically inferior are not essentially different from earlier ones, of which Samuel Cartwright's "The Prognathous Species of Mankind" (1857) is an example. The argument moves from stressed and exaggerated physiological differences between blacks and whites to the claim of broad mental differences. Features of the "typical negro" are closer to "the simiadiae and the brute creation" than to whites. The standard black color is a shiny, oily black, and lighter colors are the result not of intermixture with whites but of sickness or degeneration. In "the bleaching process of bad health or degeneration" even the pigment of the iris is lost, and the degenerate Negro is clairvoyant at night. The Negro does not have real hair: "the shaft of each hair is surrounded with a scaly covering like sheep's wool, and, like wool, is capable of being felted. True hair does not possess this property.… the negro approximates the lower animals in his sense of smell, and can detect snakes by that sense alone. All the senses are more acute, but less delicate and discriminating than the white man's." Natural history, like the Bible, "proves the existence of at least three distinct species of the genus man, differing in their instincts, form, habit, and color. The white species having qualities denied to the black—one with a free and the other with a servile mind—one a thinking and reflective being, the other a creature of feeling and imitation, almost void of reflective faculties, and consequently unable to provide for and take care of himself."

Several racial theories, notably those of Madison Grant and Lothrop Stoddard, reflected the growing awareness among the descendants of earlier groups of immigrants to the United States of the changing national origins of later groups. The works of these men both promoted the fear of the ultimate extinction of the "white race" (which was often meant to exclude southern and eastern Europeans) by rising birth rates among Asians and Africans and influenced the restrictive immigration laws of the 1920s. But it is doubtful whether these or later writers have added anything substantially new to the racist theses.

See also Affirmative Action; Boulainvilliers, Henri, Comte de; Buffon, Georges-Louis Leclerc, Comte de; Chamberlain, Houston Stewart; Enlightenment; Fascism; Gobineau, Comte Joseph Arthur de; Montesquieu, Baron de; Nationalism; Schiller, Ferdinand Canning Scott.

Bibliography

RACIST WRITINGS

The fountainhead of racist doctrines is Count Joseph Arthur de Gobineau's four-volume *Essai sur l'inégalité des races humaines* (Paris, 1853–1855); Vol. I has been translated by Adrian Collins as *Essay on the Inequality of Races* (London and New York, 1915). The first two volumes were also published at Philadelphia as early as 1856 in a translation by H. Hotz and with an introduction by Josiah C. Nott, both of whom were propagandists for slavery. For an early refutation of Gobineau's views based on moral grounds, see Alexis de Tocqueville's correspondence in his *"The European Revolution" & Correspondence with Gobineau,* edited by John Lukacs (Garden City, NY: Doubleday, 1959).

Richard Wagner published his anti-Semitic essay, "The Jews in Music," in 1850. Wagner later became an enthusiastic supporter of Gobineau, and Gobineau of Wagnerism. Representative writings of Wagner on race are available in *Wagner on Music and Drama,* compiled by Albert Goldman and Evert Sprinchorn (New York, 1964). Also important in disseminating Gobineau's views in Germany was Ludwig Schemann, founder of the Gobineau-Verein, translator, editor, and biographer of Gobineau, and author of such racist works as *Die Rassenfrage im Schrifttum der Neuzeit* (Munich, 1931).

Houston Stewart Chamberlain, Wagner's son-in-law, ranks with Gobineau as a race theorist; the two-volume *Die Grundlagen des neunzehnten Jahrhunderts* (Munich, 1899), his major work, was translated by John Lees as *The Foundations of the Nineteenth Century,* 2 vols. (London and New York: J. Lane, 1911). Chamberlain influenced both Adolf Hitler's *Mein Kampf* (2 vols., Munich, 1925–1927) and Alfred Rosenberg's *Der Mythus des 20. Jahrhunderts* (Munich, 1930).

The term *Aryan* was popularized by Friedrich Max Müller as a label for the speakers of the hypothetical language from which Indo-European languages were allegedly descended. Although Müller later denied that the term had any racial significance, the romantic claim that language expresses the soul of the race made the identification of Aryan speakers with an Aryan race almost inevitable, and Müller's own writings abound in such identifications. See, for example, *Lectures on the Science of Language,* 2 vols. (London: Longman, Green, Longman, and Roberts, 1861–1864).

In defense of slavery on racial grounds, see, in addition to Cartwright's essay, Josiah Nott's *Types of Mankind* (Philadelphia, 1854), parts of which are reprinted with Cartwright's essay and other writings in *Slavery Defended: The Views of the Old South,* edited by Eric L. McKitrick (Englewood Cliffs, NJ: Prentice-Hall, 1963). Other American works are Madison Grant, *The Passing of the Great Race* (New York: Scribners, 1916), and Lothrop Stoddard, *The Rising Tide of Color against White World-Supremacy* (New York: Scribners, 1920).

WORKS ON RACE AND RACISM

Jacques Barzun, *Race: A Study in Superstition,* 2nd ed. (New York: Harper and Row, 1965), is a historical survey. See also Ernst Cassirer, *The Myth of the State* (New Haven, CT: Yale University Press, 1946), Ch. 16, and Hannah Arendt, *The Origins of Totalitarianism,* 2nd ed. (New York: Meridian, 1958), especially Ch. 6.

Ashley Montague, *Man's Most Dangerous Myth,* 4th ed. (New York, 1964), and Ashley Montague, ed., *The Concept of Race* (New York: Free Press of Glencoe, 1964), together survey much of the present relevant biological knowledge and opinion. L. C. Dunn and Theodosius Dobzhansky, *Heredity, Race, and Society,* 2nd ed. (New York: New American Library, 1952), is a clear and useful account of the genetic aspects. Carlton S. Coon, *The Origin of Races* (New York: Knopf, 1962), is a work by a physical anthropologist who believes in the existence of biological differences between human groups that are associated with intellectual differences. Henry E. Garrett, "The Equalitarian Dogma," in *Perspectives in Biology and Medicine* 4 (1961): 480–484, presents a minority view by a former head of the American Psychological Association.

Philip W. Cummings (1967)

RACISM [ADDENDUM]

Racism is the view that (1) the human species is composed of different racial groups, (2) these groups are arranged hierarchically from least to most superior, and (3) superior groups have the right to use inferior groups for the benefit of the superior group. Sexism and speciesism are similarly defined, justifying the right of men to dominate women, and the right of human beings to dominate other species of beings. Before Darwin, racial groups were typically defined in terms of lineage and type: a racial group was the progeny of certain original types, each of which exemplified a distinctive physiognomy and pattern of behavior. The current status of a racial group was then explained by reference to its ancestral sources. Thus, the democratic and enterprising nature of the English was the result of their Anglo-Saxon heritage, whereas the servile position of Africans was the result of their being the progeny of Ham, cursed by Noah to be servants of servants. Some argued that Africans were not the progeny of Adam and Eve, but were a pre-Adamite lower species to be used for human benefit.

Post-Darwinian biology favored the notion of race as a sub-species—a group within a particular species that is isolated genetically from other members of that species, and as a result develops distinctive morphological and/or behavioral attributes. Africans, Asians, and Europeans look different because they have evolved on different continents and have developed different body types and personalities.

Social Darwinism portrayed evolution as a struggle for existence in which superior races survived and inferior ones perished. Eugenicists hoped to enhance natural

selection by using our knowledge of natural phenomena to reproduce superior human beings and avoid reproducing inferior ones. Polygenecists considered the "lower" races to be of a different species than the "higher" races, and therefore subject to the interests of the higher races. They believed that, like the offspring of horses and donkeys, racial "hybrids" were likely to be infertile and dysfunctional. Even when races were acknowledged to be of the same species, "race mixing" was considered dysgenic and debilitating to the "higher" races. Miscegenation laws prohibiting intermarriage and procreation between different races were considered to be in society's best interests and therefore good social policy.

The claim that Africans and African Americans have diminished cognitive and moral capacities compared with Europeans and European Americans was often used to justify slavery and segregation. This claim continues to attract adherents who seek naturalistic explanations of skewed racial achievements such as J. Richard Herrnstein and Charles Murray (1994), Michael Levin (1997), and Stephen Kershnar (2003). According to Allan Chase (1980), Malthusians held that the diminished intelligence of the "lower races" lead them to bad choices and immoral behavior that worsened their plight and made them among the least well off wherever they were.

Some have argued that research on racial differences perpetuates harm by reinforcing racist assumptions and should be curtailed. But while the results of research on racial differences could be used to harm, such knowledge could also be used to help. If, for instance, it were found that people with high melanin content in their skin responded to a particular chemical compound that affected mental functioning, then it might be possible to manipulate that compound to either boost or retard intellectual performance. Research on racial differences degenerates into racism only when racial differences are believed to establish a hierarchical ranking that is biologically fixed and immutable.

Many now consider the very concept of a race to be an artifact of European expansionism, justifying European domination of African, Asian, and Native American people. On this view, Europeans classified "others" as different races to further the ends of domination, and continued use of racial categories merely reinforces that original aim. Rejecting European economic, political, and cultural imperialism requires that we reject both racism and racialism (i.e., classification by races). In a similar fashion, Marxists consider racism to be an ideological ploy that divides the lower classes so that the European and non-European proletariat fight one another instead of fighting capitalists.

The Nazis portrayed Jews as an inherently acquisitive and parasitical race that threatened the evolution of mankind. Their mission was to exterminate the "Jewish race" and establish the unchallenged hegemony of the "Aryan race." In reaction to the atrocities of the holocaust, many scientists and leaders such as Joseph Graves (2001) and Ashley Montagu (1997) and marshaled evidence to show that race was a pseudoconcept with no biological validity. On this view, neither the Jews nor the Aryans were races defined by distinct biological differences. Jews typically were biologically more similar to contiguous non-Jews than they were to Jews in distant locales. Moreover, classifying people into races by skin color, hair texture, and other observable characteristics ignores many other features (such as internal proteins and DNA sequences) that could also be used to classify them. Those who argue that there are no races typically cite evidence that genetic variation is greater within traditional racial groups than between them, thus showing that racial groups aree not reproductively isolated gene pools.

But other biologists and social scientists, such as Phillip Kitchner (1999), Robin Andreasen (2000), and Neil Risch (2003), have insisted that there is compelling biological evidence for the existence of races. These researchers assert that when human populations are classified in terms of their ancestral geographic origins, there is a high correlation between traditional racial groups and genetic clusters.

However, the question of whether or not races exist is independent of whether racism exists. Just as it is possible for witchcraft to exist even though there are no witches, so it is possible for racism to exist even if there are no races. Some, such as Naomi Zack (2002), argue that if we are to move beyond "racial" animosities of the past, we must cease using racial categories, because they have no biological validity. For Lucius Outlaw (1996) and Alain Locke (1999), continued consciousness of racial distinctions may be linked to pride in cultural achievements made under extreme duress, or to demands for restitution for past and present harms. Whether or not races exist, it should be possible to agree that racism is morally wrong, and that racism should be eliminated.

Racist behavior elevates the interests of members of allegedly superior races over the interests of allegedly inferior races. The existence of racist behavior may be independent of individual intent, as is often the case with institutional racism, where certain procedures and practices harm groups historically considered inferior, even if

the implementers do not explicitly intend such harm. Thus, requirements that are unnecessary for successful performance and recruitment limited to traditional networks often serve to perpetuate the effects of overt and egregious racist acts of the past, even if this is not currently intended by those who implement such policies.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism; Social and Political Philosophy.

Bibliography

Andreasen, Robin. "Race: Biological Reality or Social Construct" *Philosophy of Science* 67 (Proceedings) (2000): 653–666.

Appiah, Kwame Anthony. *In My Fathers House: Africa in the Philosophy of Culture*. New York: Oxford University Press, 1992.

Banton, Michael. *Racial Theories*. New York: Cambridge University Press 1987.

Block, N. J., and Gerald Dworkin, eds. *The I.Q. Controversy: Critical Readings*. New York: Pantheon, 1976.

Braun, Lundy. "Race, Ethnicity, and Health: Can Genetics Explain Disparities?" *Perspectives in Biology and Medicine* 45 (2002): 159–174.

Chase, Allan. *The Legacy of Malthus: The Social Costs of the New Scientific Racism*. Urbana: University of Illinois Press, 1980.

Ezorsky, Gertrude. *Racism and Justice: The Case for Affirmative Action*. Ithaca, NY: Cornell University Press, 1991.

Frederickson, George M. *Racism: A Short History*. Princeton, NJ: Princeton University Press, 2002.

Graves, Joseph L. Jr. *The Emperor's New Clothes—Biological Theories of Race at the Millennium*. New Brunswick, NJ: Rutgers University Press, 2001.

Harris, Leonard, ed. *Racism*. Amherst, NY: Prometheus Books, 1999.

Herrnstein, Richard, and Charles Murray. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press, 1994.

Kershnar, Stephen. "Experiential Diversity and *Grutter*." *Public Affairs Quarterly* 17 (2) (April 2003): 163–164.

Kitchner, Phillip. "Race, Ethnicity, Biology, Culture." In Harris, *Racism*, 87–117.

Levin, Michael. *Why Race Matters: Race Differences and What They Mean*. Westport, CT: Praeger, 1997.

Locke, Alain. "The Contribution of Race to Culture." In Harris, *Racism*, 215–218. Amherst, NY: Humanity, 1999.

Montagu, Ashley. *Man's Most Dangerous Myth: The Fallacy of Race*. Walnut Creek, CA: Altamira Press, 1997.

Mosley, Albert. "Are Racial Categories Racist?" In *Research in Afrinca Literatures*, 28.4 (Winter 1997), pp. 101–110.

Outlaw, Lucius T. Jr. *On Race and Philosophy*. New York: Routledge, 1996.

Risch, Neil, Esteban Burchard, Elad Ziv, and Hua Tang. "Categorization of Humans in Biomedical Research: Genes, Race, and Disease." *Genome Biology* 3 (2002): comment 2007.1–2007.12.

Zack, Naomi. *Philosophy of Science and Race*. New York: Routledge, 2002.

Albert G. Mosley (1996, 2005)

RADBRUCH, GUSTAV

(1878–1949)

Gustav Radbruch was a German legal philosopher whose name and work have become widely known outside Germany only since the end of World War II. During his lifetime, the interests and activities of scholar, politician, and reformer of law were closely intermingled. After World War I, Radbruch became active in the Social Democratic Party and twice served as minister of justice of the Weimar Republic. His principal work was the draft of a new criminal code. Later he held a chair of law at the University of Heidelberg, from which he was dismissed by the Nazi regime. After World War II he was recalled and exercised a predominant influence in the reorientation of German legal education and philosophy until his death.

Radbruch's legal philosophy, generally known as "relativism," is closely akin to the position of his friend and teacher Max Weber. Radbruch believed, like Weber, that values could not be scientifically proved and that they were "a matter of conscience (*Gewissen*), not of science (*Wissenschaft*)," This in no way implied indifference to values. Radbruch differed both from Rudolf Stammler, who sought to formulate a theoretically valid concept of justice, and from Hans Kelsen, who detached legal science altogether from a philosophy of values. Radbruch, while starting from the Kantian distinction of "is" (*Sein*) and "ought" (*Sollen*), was guided mainly by the teachings of Heinrich Rickert and Emil Lask in treating law as a *Kulturwissenschaft,* a science directed to the realization of values. He therefore considered that the task of legal philosophy was to relate legal reality to basic ideas. But the truth of specific ideas and values cannot be scientifically proved. Radbruch instead developed—and applied to numerous specific problems of law—a series of antinomies of legal values. Thus, the Aristotelian idea of distributive justice, which directs equals to be treated equally, says nothing about the perspective from which they are to be characterized as equals or unequals. Justice, which cannot yield objective criteria of equality, must be supplemented by a second value, "utility," and a third, "security." Between these three values there is constant tension. In another perspective, law can be directed to individual values, collective values, or work values. Accordingly, a legal system emphasizes either individualism, collec-

tivism, or transpersonalism. For the first, the ultimate idea is liberty; for the second, the nation; and for the third, civilization.

After the war, Radbruch recoiled from the extremes of tolerance—as practiced by the Weimar Republic during the rise of the Nazi movement—having witnessed the unprecedented barbarism of the Third Reich, which was largely covered by a formal notion of law. He tentatively turned to a moderate natural-law philosophy, holding that in certain extreme cases a contradiction between positive law and justice might reach such an intolerable degree that the law as unjust law (unlawful law, *unrechtiges Recht*) must cede to the higher demands of justice. Radbruch died before he could elaborate his thesis beyond the postulate that special courts should be empowered to adjudge the validity of laws.

See also German Philosophy; Justice; Kelsen, Hans; Philosophy of Law, History of; Philosophy of Law, Problems of; Rickert, Heinrich; Stammler, Rudolf; Value and Valuation; Weber, Max.

Bibliography

WORKS BY RADBRUCH

Einführung in die Rechtswissenschaft. Leipzig, 1907. Edited by K. Zweigert. Stuttgart: Koehler, 1961.

Rechtsphilosophie. Leipzig, 1914. 3rd ed. revised and enlarged, Leipzig: Quelle and Meyer, 1932. 5th ed., completed with a preface by Erik Wolf, Stuttgart: Koehler, 1956. Translated from the German edition by Kurt Wilk with introduction by E. W. Patterson in *The Legal Philosophies of Lask, Radbruch and Dabin.* Cambridge, MA: Harvard University Press, 1950.

"Gesetzliches Unrecht und übergesetzliches Recht." *Süddeutsche Juristenzeitung* 5 (1946): 107ff.

Der Geist des englischen Rechts. 2nd ed. Heidelberg, 1947.

Vorschule der Rechtsphilosophie. Willsbach: Scherer, 1947.

WORKS ON RADBRUCH

Campbell, A. H. *Gustav Radbruchs Rechtsphilosophie.* Hanover, 1949.

Cattaneo, M. A. "Il Positivismo giuridico e la separazione tra il Diritto e la Morale." *Rendiconti* (Istituto Lombardo) 94 (1960): 701–742.

Friedmann, Wolfgang. "Gustav Radbruch." *Vanderbilt Law Review* 14 (1960): 191–209.

Wolf, Erik. *Griechisches Rechtsdenken,* 3 vols. Tübingen, 1950–1954.

Wolf, Erik. *Grosse Rechtsdenker der deutschen Geistesgeschichte.* 3rd ed. Tübingen: Mohr, 1951.

Wolf, Erik. *Das Problem der Naturrechtslehre: Versuch einer Orientierung.* Tübingen, 1955. 2nd ed. Tübingen, 1959.

Wolf, Erik. *Rechtsgedanke und biblische Weisung.* Tübingen, 1948.

Wolfgang Friedmann (1967)

RADISHCHEV, ALEKSANDR NIKOLAEVICH

(1749–1802)

Aleksandr Nikolaevich Radishchev was the leading social critic and philosopher of the Russian Enlightenment. He was born in Moscow, the son of a prosperous landowner, and was educated in Moscow, in St. Petersburg, and, from 1766 to 1771, at the University of Leipzig. At Leipzig he studied under the Leibnizian Ernst Platner and read widely in current French philosophy. Upon his return to Russia he pursued a successful career in the civil and military service until 1790, when his radical work *Puteshestvie iz Peterburga v Moskvu* (St. Petersburg, 1790; translated as *A Journey from St. Petersburg to Moscow*) aroused the ire of Catherine the Great and he was exiled to Siberia. Paul I permitted him to return to European Russia in 1796. After the accession of Alexander I, in 1801, Radishchev was appointed to a special legislative commission, but his egalitarian, libertarian proposals went unheeded, and in September 1802 he took his own life in St. Petersburg.

In the *Journey,* Radishchev employed the principles of natural law and the social contract to support a severe critique of Russian social institutions, serfdom in particular. Under the inspiration of Jean-Jacques Rousseau, Voltaire, Guillaume-Thomas-François de Raynal, and other French thinkers, he condemned serfdom as morally wrong and economically inefficient, criticized autocracy, and attacked censorship and other practices that violate men's natural rights to freedom and equality. He advocated immediate reforms to avert revolution and called generally for enlightenment and "naturalness" in social arrangements, manners, and morals.

In Siberia, Radishchev wrote his principal philosophic work, *O cheloveke, o ego smertnosti i bessmertii* (On man, his mortality and immortality; published posthumously, St. Petersburg, 1809), a close examination of the cases for and against personal immortality. In the end he rejected materialistic denials of immortality in favor of various arguments—from personal identity and the conservation of force, among others—that suggest the existence of an incorporeal soul that survives the body and passes into a more perfect state. In epistemology Radishchev adopted a realistic position and accepted experience as the only basis for knowledge but maintained that in addition to sensory experience there is "rational experience" of the relationships of things and that man "feels" the existence of a Supreme Being. He also maintained that things in themselves are unknowable,

asserting that thought, like the verbal expression it employs, is merely symbolic of reality.

Radishchev's treatise *O cheloveke* was one of the first original philosophic works in the Russian language, and the influence his pioneering social criticism had on Alexander Pushkin, the Decembrists, and subsequent generations of Russian reformers and revolutionaries has led to his being regarded as the father of social radicalism in Russia. He was also a poet of considerable talent.

See also Immortality; Natural Law; Rousseau, Jean-Jacques; Russian Philosophy; Social Contract; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY RADISHCHEV

Polnoe sobranie sochinenii (Complete works). 3 vols. Moscow and Leningrad, 1938–1952.

A Journey from St. Petersburg to Moscow. Translated by Leo Wiener. Cambridge, MA: Harvard University Press, 1958.

"On Man, His Mortality and Immortality." Selections translated by Frank Y. Gladney and George L. Kline in *Russian Philosophy,* edited by J. M. Edie, J. P. Scanlan, and M.-B. Zeldin. Chicago: Quadrangle, 1965.

WORKS ON RADISHCHEV

Dynnik, M. A. et al., eds. *Istoriia filosofii.* 6 vols. Moscow, 1957–.

Lapshin, I. I. *Filosofskie vzgliady A. N. Radishcheva* (Philosophical views of A. N. Radishchev). Petrograd, 1922.

Zenkovsky, V. V. *Istoriia russkoi filosofii.* 2 vols. Paris, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy.* 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)

RĂDULESCU-MOTRU, CONSTANTIN

(1868–1954)

Constantin Rădulescu-Motru, the Romanian philosopher of energetic personalism, studied at the universities of Bucharest, Paris, Munich, and Leipzig. He obtained his doctorate from Leipzig in 1893 with a thesis on the development of Immanuel Kant's theory of causality in nature, published in Wilhelm Wundt's *Philosophische Studien.* In 1904 he became professor of psychology and logic at the University of Bucharest. He founded the journals *Noua Revista Română* (1900), *Studii Filosofice* (1905, after 1920 called *Revista de Filosofie*), *Anale de Psihologie* (1935), and *Jurnal de Psihotehnică* (1937), as well as the Romanian Society of Philosophy.

Rădulescu-Motru was the dominant figure in Romanian philosophy from 1905 to 1930. The most articulate expression of his philosophical system is to be found in his *Personalismul Energetic* (1927). Influenced by the work of Wilhelm Ostwald and William Stern, it was an impressive effort to unify the results of natural science, biology, and psychology. Rădulescu-Motru called his system personalistic because the human personality plays the central role within it, and energetic because he considered personality to be the highest form of cosmic energy. The universe is in continuous evolution, and its goal is the creation of energetic personality. Rădulescu-Motru distinguished six stages of the evolutionary process: cosmic energy, adaptation, organic individuality, consciousness, ego, and personality. Personality is both modified and enriched through evolution from the primitive *homo divinans* to *homo faber.* Finally, through Stoicism, Christianity, and science, the energetic personality, the vocational or professional man, emerges. With the achievement of a personality having a total comprehension of the universe, the evolutionary process will come to an end; Nature will have reached its ultimate goal.

See also Kant, Immanuel; Ostwald, Wilhelm; Personalism; Stern, Louis William; Wundt, Wilhelm.

Bibliography

WORKS BY RĂDULESCU-MOTRU

Puterea Sufleteasca. Bucharest, 1907.

Elemente de Metafizică (1912). Definitive edition Bucharest, 1928.

Curs de Psihologie. Bucharest, 1924.

Personalismul Energetic. Bucharest, 1927.

Vocația. Bucharest, 1932.

Românismul. Bucharest, 1936.

WORKS ON RĂDULESCU-MOTRU

Omagiu profesorului C. Rădulescu-Motru. Bucharest, 1932.

Rus, G. "Il personalismo energetico di C. Rădulescu-Motru." *Acta Philosophica et Theologica* 2 (1964): 411–438.

Mircea Eliade (1967)

RAHNER, KARL

(1904–1984)

One of the most significant Roman Catholic theologians of the twentieth century and a formative influence upon

Vatican II, Karl Rahner was born on March 5, 1904, in the city of Freiburg im Breisgau, Germany, the fourth of seven children in the family of Karl and Luise (Trescher) Rahner. Upon graduation from secondary school at the age of eighteen, Rahner followed in the footsteps of his elder brother Hugo and entered the Society of Jesus; he was to remain a Jesuit his entire life. During his novitiate studies from 1924 to 1927, Rahner was introduced to Catholic scholastic philosophy and to the modern German philosophers. He seems especially to have been influenced by the work of Joseph Maréchal (1878–1944), the Belgian philosopher and Jesuit, whose adoption of Kant's transcendental method in his five-volume work, *Le point de départ de la métaphysique*, had led to somewhat of a breakthrough in the appreciation of Kant's philosophy among neo-Scholastics. Maréchal was known as the "father of transcendental Thomism" for his use of St. Thomas Aquinas's epistemology in an attempt to demonstrate that the metaphysical world Kant had secured for practical reason was already inherent in the theoretical.

After teaching Latin at the Feldkirch Novitiate, Rahner studied theology at Valkenburg in the Netherlands (1929–1933), where his Christian spirituality was further nurtured through study of patristic and medieval mysticism, and above all of St. Ignatius of Loyola (1491–1556), founder of the Jesuit order and author of the *Spiritual Exercises*. Following ordination to the priesthood in 1932, Rahner commenced study for his doctoral thesis in philosophy at Freiburg, while at the same time attending lectures by Martin Heidegger, whose philosophy of Dasein, or "being in the world," was to be the other primary philosophical influence upon him. His dissertation, a response to Kant's critique of theoretical metaphysics by means of the transcendental Thomism of Maréchal and the existentialism of Heidegger, was rejected by his doctoral director, Martin Honecker, for its departure from more traditional neo-Scholastic interpretations of Aquinas's epistemology, but was later published as *Geist in Welt* (Spirit in the world).

After failing the doctorate in philosophy, Rahner returned to Austria, where he successfully completed his second dissertation, this time in theology, at Innsbruck in 1936 and was appointed as Privatdozent (lecturer) in the faculty of theology of the University of Innsbruck in 1937. That summer he delivered a series of lectures to the Salzburg summer school on the "Foundations of a Philosophy of Religion," later published as *Hörer des Wortes* (Hearer[s] of the word).

When the Nazis abolished the theology faculty (July 1938) and the Jesuit college (October 1939) at Innsbruck, Rahner left for Vienna, where he did some teaching and served as a consultant at the Pastoral Institute for five years. After a brief stint as a pastor in Bavaria in the final year of the war, he taught dogmatic theology at Berchmanskolleg in Pullach. In 1948 he returned to the theology faculty at Innsbruck, where he was to reside until 1964. There he lectured on a wide variety of topics to later be included in the essays published as *Schriften zur Theologie* (Theological investigations), the first volume of which appeared in 1954. Of particular significance was his scholarly preoccupation with the relationship between nature and grace.

During this prolific period Rahner experienced some difficulties within the Church, beginning as early as 1950 when he was prevented from publishing a book on the Assumption of Mary, and continuing through the following decade until 1962, when he was placed under a censorship regulation from Rome. Suspicions over his orthodoxy subsided, however, when the newly elected Pope John XXIII appointed Rahner as one of the theological experts (*periti*) at the Second Vatican Council, and the censorship upon him was reversed in 1963. Rahner's influence at Vatican II was widespread; particularly noteworthy is his selection as one of the seven theologians who would develop *Lumen Gentium* (Dogmatic constitution on the Church), a document fully explicating the doctrine of the Church, and setting forth explicitly in chapter II the Church's inclusivist stance with regard to salvation.

It was during the Second Vatican Council that Rahner was invited to take the Chair in Christianity and Philosophy of Religion at the University of Munich, where he began teaching in 1964, the same year that a Festschrift, *Gott in Welt* (God in the world) was published in honor of his sixtieth birthday. During his time at Munich, Rahner published a collection of essays in spirituality as the seventh volume of *Schriften zur Theologie*, and together with Edward Schillebeeckx edited the first issue of *Concilium*. In 1967 Rahner accepted the University of Münster's invitation to become Ordinary Professor of Dogmatics and the History of Dogma, where he completed three more volumes of the *Schriften*, before retiring in 1971. Retirement brought him back to Munich, where he prepared *Grundkurs des Glaubens* (Foundations of Christian faith), the most systematic summary of his theology, and to Innsbruck, where in addition to pastoral and moral essays, Rahner worked out the most developed form of his transcendental Christology, and completed the final volumes of the *Schriften*, thus continuing the life

of the diligent scholar until his death in Innsbruck on March 30, 1984.

Rahner has been criticized for a failure to adequately address the problem of evil, especially in light of his experience as a German Catholic living through the Nazi genocide of two-thirds of the Jewish population of Europe. Neither in the above nor in what follows will it be possible to do full justice to the breadth of Rahner's theological output, which covers almost every aspect of religious thought. The focus here is upon Rahner's efforts at *aggiornamento*, or renewal of neo-Scholasticism and the philosophical import of two concepts integral to his theological *weltanschauung*: *Vorgriff auf esse* and *das übernatürliche Existential*.

Geist in Welt focuses upon one of the central problems of philosophy, namely the nature and possibility of metaphysics. In this work, Rahner examines one part of St. Thomas Aquinas's metaphysics of knowledge, specifically that section of the *Summa theologiae* that addresses what appears to sense intuition, *conversio ad phantasmata* (conversion to the phantasm), in light of Kant's critique of speculative metaphysics. Whereas Kant had rejected theoretical knowledge of God in order to secure a place for metaphysics as a practical philosophy, Rahner uses the tools of transcendental and existentialist philosophy, honed through Maréchal and Heidegger respectively, to retrieve the theoretical metaphysics of St. Thomas. *Spirit in the World* is thus Rahner's attempt to demonstrate how, given that human knowledge is wedded to the *a posteriori*, or realm of sensory experience, metaphysics is still possible; and as fraught as the philosophical analysis is, the main arguments are fairly accessible.

Essentially Rahner proffers a teleology of knowledge according to which there is presupposed in every human act of knowing the *Vorgriff auf esse* (the "pre-apprehension of being" in Heideggerian terms), a transcendental awareness of infinite being, or of God, the a priori condition without which no individual act of knowing could occur. In every act of knowing then, the individual, or "spirit in the world," has already reached out beyond the world and known the metaphysical. This awareness of God, which is always indirect and shrouded in mystery (since we cannot know God as if God were an object among other realities that are present to us), presupposes the transcendental orientation of the human knower to God, who is both the source and ultimate goal of the human quest for knowledge. To be human is therefore to be in relation to God, since we implicitly affirm the existence of God in every judgment we make, regardless of whether or not we ever formally acknowledge this. *Ipso facto*, human existence itself implies the transcendental experience of God for Rahner, thus satisfying not only the transcendental Thomism of Maréchal and the existentialism of Heidegger, but also the Ignation impulse to "find God in all things."

Hörer des Wortes is formally an investigation into the relationship between philosophy of religion and theology. Philosophy of religion, according to Rahner, consists in showing human beings to be the infinite spirits who, because of our nature, are turned toward a possible revelation, or self-communication of God, since revelation for Rahner is always personal, not propositional. God, the personal infinite, chooses human history as the place of transcendent self-communication (a divine self-communication that finds concrete historical expression in Jesus Christ); theology begins with the human person who has become attuned to God's self-communication, a hearer of God's word. In order to make this case, Rahner develops his "transcendental arguments" further and grounds them more fully theologically. By means of a "theological anthropology," a metaphysical analysis of human nature, Rahner proposes *Vorgriff auf esse* as a pre-apprehension of infinite being that also elicits the restless yearning of the human spirit (echoing a desire at least as old as Augustine) for fulfillment in and through that absolute being whose self-communication is both the ground and telos of human existence.

This understanding of the human spirit's desire for transcendent meaning, together with God's ineffable self-communication, later has important implications for Rahner's interpretation of the relationship between nature and grace, a relationship examined through his concept of *das übernatürliche Existential*, the "supernatural existential," first coined during his intervention in the *nouvelle théologie* debate in 1950 but still being worked out as late as 1976 in *Grundkurs des Glaubens*. The debate revolved around whether or not the human orientation toward God was natural or supernatural, and Rahner uses the term "supernatural existential" in an attempt to overcome the tendency in neo-Scholastic theology to dichotomize nature and grace, while at the same time safeguarding the gratuity of God's grace. Once again borrowing from Heidegger's vocabulary, Rahner defines an "existential" as a fundamental element in human existence, and claims that the central and abiding existential of human nature is the unconditional desire for grace and for the beatific vision. At the same time, however, he argues that the very fact of this desire already belies God's self-communication, precisely the meaning of grace for Rahner. In other words, since our very existence is per-

meated with God's constant self-giving, human nature is already grace laden. *Das übernatürliche Existential* ultimately entails the universal human experience of grace, and similar to the *Vorgriff auf esse*, this experience, though wedded to the world, is also transcendental and thus can never be directly or concretely realized.

The ubiquitous nature of the "supernatural existential" also undergirds Rahner's Christian inclusivism, itself a corollary of his philosophy of grace, and arguably the theological stance for which he is best known in non-Catholic circles. For Rahner, because God's gracious self-communication has found concrete historical expression in Jesus Christ, all grace is ultimately the grace of Christ; yet, significantly, Christ's universal grace is not narrowly circumscribed by Christianity. If *das übernatürliche Existential* is a universal given, and just as the *Vorgriff auf esse* is never directly or concretely realized, then it is possible that a person may accept this gift of grace without explicit acknowledgment and regardless of whether or not one is formally Christian. It is for this reason that non-Christians living lives of grace are "anonymous Christians" from Rahner's perspective, a title not intended as a subtle form of Christian supersessionism, but rather as a theologically astute commitment to the view that God's grace is active well beyond the confines of Christianity. The religious inclusivism espoused by Rahner had an ecumenical import that has proven vital to Catholic interreligious dialogue in the post–Vatican II era and that presumably will continue to be relevant to the burgeoning interest in religious diversity among philosophers of religion well into the third millennium

See also Heidegger, Martin; Maréchal, Joseph; Thomas Aquinas, St.

Bibliography

WORKS BY RAHNER

Geist in Welt: Zur Metaphysik der endlichen Erkenntnis bei Thomas von Aquin. Innsbruck and Leipzig: Felizian Rauch, 1939; 2nd ed., expanded and reworked by J. B. Metz. Munich: Kösel, 1957. Translated as *Spirit in the World* by W. Dych. New York: Herder and Herder, 1968.

Grundkurs des Glaubens: Einführung in den Begriff des Christentums. Freiburg, Basel, Vienna: Herder, 1976. Translated as *Foundations of Christian Faith: An Introduction to the Idea of Christianity* by W. Dych. London: Darton, Longman and Todd, 1978.

Hörer des Wortes: Zur Grundlegung einer Religionsphilosophie. Munich: Kösel-Pustet, 1941; 2nd ed., reworked by J. B. Metz. Munich: Kösel, 1963. Translated as *Hearer of the Word* by J. Donceel. New York: Continuum, 1994.

Karl Rahner: Sämtliche Werke. 32 vols., edited by K Lehmann, J. B. Metz, K. H. Neufeld, A. Raffelt, and H. Vorgrimler. Freiburg: Herder, 1995ff.

Schriften zur Theologie, 16 vols. Einsiedeln und Zurich, 1954–1984. Translated as *Theological Investigations*, 23 vols. London: Darton, Longman and Todd, 1961–1984.

WORKS ABOUT RAHNER

Bleistein, R., and E. Klinger. *Bibliographie Karl Rahner, 1924–1969*. Freiburg: Herder, 1969.

Bleistein, R., ed. *Bibliographie Karl Rahner, 1969–1974*. Freiburg: Herder, 1969.

Dych, W. *Karl Rahner*. London: Geoffrey Chapman, 1992.

Imhof, P., and H. Treziak. "Bibliographie Karl Rahner, 1974–1979." In *Wagnis Theologie: Erfahrungen mit der Theologie Karl Rahners*, edited by H. Vorgrimler, 579–597. Freiburg: Herder, 1979.

Imhof, P., and E. Meuser, "Bibliographie Karl Rahner, 1979–1984." In *Glaube im Prozess*, edited by E. Klinger and K. Wittstadt, 854–888. Freiburg: Herder, 1984.

Kilby, K. *Karl Rahner: Theology and Philosophy*. London: Routledge, 2004.

Marmion, D., and M. E. Hines, eds. *The Cambridge Companion to Karl Rahner*. Cambridge, U.K.: Cambridge University Press, 2005.

Pedley, C. J. "An English Bibliographical Aid to Karl Rahner." *The Heythrop Journal* 25 (1984): 319–365.

Raffelt, A. "Karl Rahner: Bibliographie der Sekundärliteratur, 1948–1978." In *Wagnis Theologie: Erfahrung mit dem Theologie Karl Rahners*, edited by H. Vorgrimler, 598–622. Freiburg: Herder, 1979.

Raffelt, A. "Karl Rahner: Bibliographie der Sekundärliteratur, 1979–1983 und Nachträge." In *Glaube im Prozess*, edited by E. Klinger and K. Wittstadt, 872–885. Freiburg: Herder, 1984.

Raffelt, A., and R. Siebenrock. "Karl Rahner: Sekundärliteratur, 1984–1993." In *Karl Rahner in Erinnerung*, edited by A. Raffelt, 165–205. Freiburger Akademieschriften, Band 8. Düsseldorf: Patmos, 1994.

Tallon, A. "Rahner Bibliography Supplement." *Theology Digest* 38 (2) (1991): 131–140.

Tallon, A. "Rahner Studies, 1939–1989: Part 1, 1939–1973." *Theology Digest* 36 (4) (1989): 321–346.

Tallon, A. "Rahner Studies, 1939–1989: Part 2, 1974–1989." *Theology Digest* 37 (1) (1990): 17–41.

Elizabeth Galbraith (2005)

RAMSEY, FRANK PLUMPTON

(1903–1930)

Frank Plumpton Ramsey, the Cambridge mathematician and philosopher, was one of the most brilliant men of his generation; his highly original papers on the foundations of mathematics, the nature of scientific theory, probability, and epistemology are still widely studied. He also

wrote two studies in economics, the second of which was described by J. M. Keynes as "one of the most remarkable contributions to mathematical economics ever made." Ramsey's earlier work led to radical criticisms of A. N. Whitehead and Bertrand Russell's *Principia Mathematica,* some of which were incorporated in the second edition of the *Principia.* Ramsey was one of the first to expound the early teachings of Ludwig Wittgenstein, by whom he was greatly influenced. In his last papers he was moving toward a modified and sophisticated pragmatism.

THE FOUNDATIONS OF MATHEMATICS

A stumbling block in the reduction of mathematics to logic attempted in *Principia Mathematica* has long been its appeal to the so-called ramified theory of types, introduced in order to cope with the paradoxes discovered by Russell and others. The excessive restrictions demanded by the theory of types were mitigated by introducing an ad hoc axiom of reducibility, which Ramsey, following Wittgenstein, held to be at best contingently true. Ramsey was one of the first to argue, following Giuseppe Peano, that many of the notorious paradoxes depended on the use of equivocal semantic notions having no place in mathematics. By introducing the notion of "predicative functions"—roughly speaking, truth-functions permitting infinitely many arguments—Ramsey was able to show that the paradoxes could be avoided without appeal to an axiom of reducibility. In order to improve what he regarded as an unsatisfactory conception of identity in *Principia Mathematica,* Ramsey proposed the wider concept of "propositional functions in extension," considered as correlations, not necessarily definable, between individuals and associated propositions. Fully elaborated, this view would seem to lead to a markedly nonconstructivistic set theory, which most contemporaries would find unacceptable. Ramsey's distinction between semantic and logical paradoxes and his rejection of that part of the theory of types that subdivides types into "orders" has been almost universally accepted by his successors.

PHILOSOPHY OF SCIENCE

In a striking paper, "Theories," Ramsey developed a novel method for eliminating overt reference to theoretical entities in the formal statement of scientific theory. The method consists of replacing, in the axioms of the formal system expressing the scientific theory in question, every constant designating a theoretical entity with an appropriate variable and then applying universal quantification over the propositional matrices thus obtained. Ramsey was able to show that the conjunction of the universally quantified statements thus derived from the original axioms would have the same observational consequences as the original axiom system. This technique is of interest to philosophers concerned with the ontological implications or commitments of scientific theory.

PROBABILITY

Ramsey sketched a theory of probability considered as measuring a degree of "partial belief," thereby providing a stimulus to what are sometimes called "subjective" or "personalistic" analyses of probability. His most important idea was an operational test for degree of belief. Suppose somebody, *P,* has no preference between the following options: (1) to receive m_1 for certain, and (2) to receive m_2 if p is true but m_3 if p is false, where p is some definite proposition and m_1, m_2, and m_3 are monetary or other suitable measures of utility for *P.* Then *P*'s degree of belief in p is proposed to be measured by the ratio $(m_1 - m_3)/(m_2 - m_3)$—roughly speaking, therefore, by the betting odds that *P* will accept in favor of *p*'s being true, given the relative values to him of the possible outcomes.

GENERAL PHILOSOPHY

Ramsey's most suggestive idea in general philosophy was that of treating a general proposition, say of the form "all *A*'s are *B,*" as a "variable hypothetical," considered not as a truth-function (as it had been in his earlier papers) but rather as a rule for judging that if something is found to be an *A* it will be judged to be a *B*—that is, as a formula for deriving propositions in certain ways rather than as an authentic proposition having truth-value. This idea is connected with Ramsey's unfortunately fragmentary explorations into the connections between belief, habit, and behavior. Ramsey's papers on facts, propositions, and universals also have not outlived their usefulness.

See also Keynes, John Maynard; Logical Paradoxes; Mathematics, Foundations of; Peano, Giuseppe; Philosophy of Science; Pragmatism; Probability; Russell, Bertrand Arthur William; Scientific Theories; Type Theory; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

A collection of Ramsey's work, including previously unpublished papers, was published posthumously as *The Foundations of Mathematics and Other Logical Essays,* edited by Richard B. Braithwaite (London: K. Paul, Trench, Trubner, 1931). This collection has a preface by G. E. Moore, a useful editor's introduction, and a complete bibliography.

For the definitions of "predicative functions" and "functions in extension," see especially pp. 39–42, 52–53; Ramsey's discussion of theories is mainly on pp. 212–236; the generalized betting definition of degree of belief occurs on p. 179.

For discussions of Ramsey's work, see Israel Scheffler, *The Anatomy of Inquiry* (New York: Knopf, 1963), pp. 203–222, which contains a critical exposition of Ramsey's procedure for eliminating theoretical terms; Herbert Gaylord Bohnert, *The Interpretation of Theory* (PhD diss., University of Pennsylvania, 1961), further elaboration of Ramsey's work on the nature of scientific theory; Leonard J. Savage, *The Foundations of Statistics* (New York: Wiley, 1954), which acknowledges indebtedness to Ramsey's definition of partial belief; and Gilbert Ryle, "'If,' 'So,' and 'Because,'" in *Philosophical Analysis,* edited by Max Black (Ithaca, NY: Cornell University Press, 1950; reprinted, New York, 1963), which is a discussion of hypothetical statements as "inference licenses."

Max Black (1967)

RAMUS, PETER

(1515–1572)

Peter Ramus was a logician, educational reformer, and author of many widely used works on philosophy and letters. He was born Pierre de la Ramée in Cuts (Oise), in northern France, the son of an impoverished descendant of a noble family from Liége. After beginning Latin at Cuts, he went to study at Paris, probably between the ages of eight and twelve, and despite grave financial difficulties received his master of arts degree there at the age of twenty-one. His master's inaugural thesis, according to one still widely circulated but questionable report, was *Quaecumque ab Aristotele Dicta Essent, Commentitia Esse* (Whatever Aristotle has said is a fabrication; the common translation of *commentitia* as "false" is oversimplified).

In 1543, Ramus (he had adopted Petrus Ramus as the Latin form of his name) published two works growing out of his teaching, *Dialecticae Partitiones* (The structure of dialectic," also titled *Institutiones Dialecticae* [Training in dialectic]) and *Aristotelicae Animadversiones* (Remarks on Aristotle), which violently attacked Aristotle and the university curriculum as confused and disorganized. The university faculty, led largely by doctors of medicine, secured from Francis I a decree forbidding the sale of these books and prohibiting their author from teaching publicly and from writing on philosophy (which included all academic subjects other than grammar, rhetoric, medicine, law, and theology). Ramus, however, quietly continued to teach and write and in 1545 moved to the Collège de Presles in Paris, where he was joined by his earlier associate, Omer Talon (Audomarus Talaeus). Ramus soon became principal and dedicated himself, with great success, to promoting more purposeful and effective teaching. In 1547, Henry II lifted the ban against Ramus and in 1551, he appointed him professor of eloquence and philosophy in the body of professors supported by the king, which was later known as the Collège de France; Ramus became its first dean. Earlier an observant Catholic, he embraced the Protestant reform around 1562, withdrawing to Fontainebleau in 1562–1563 during the religious wars and to Rhenish Germany and Switzerland from 1568 to 1570. He returned, however, and was murdered on the third day of the Massacre of St. Bartholomew. Charles Waddington's assignment of his murder to an academic opponent, the physician Jacques Charpentier, is repeated in many encyclopedia articles but is without demonstrable foundation.

WORKS

Ramus's published works run to some sixty-odd titles, supplemented by thirteen additional works of Talon, his frequent collaborator. The works of the two men appeared mostly between 1543 and 1650, in nearly eight hundred (at present) known editions and adaptations (some eleven hundred if works published in collected editions are separately enumerated). Besides the pivotal writings on dialectic, or logic, and on rhetoric, Ramus's works include classical editions and commentaries; lectures on physics, metaphysics, and mathematics; textbooks for grammar, arithmetic, algebra, and geometry; miscellaneous orations and open letters; and the posthumously published *Commentariorum de Religione Christiana Libri Quatuor* (1576), a basically Zwinglian theological work, unoriginal and apparently of little influence. Other works, notably Latin translations from the Greek, remained unpublished at his death. Although most of his writing was in academic Latin, he published a few works in French, including a *Gramere* of the French language (1562) in a reformed spelling that was developed from that of Louis Meigret.

PHILOSOPHY

The striking orderliness of Ramus's philosophy is superficial and is determined by pedagogical serviceability rather than by insight. His *Dialectica* (French, 1555; Latin, 1556, with subsequent revisions), later called also *Logica,* is the key work in the Ramist canon and appeared in nearly 250 extant editions or adaptations, chiefly Latin. The *Dialectica* grew out of his 1543 works and proposed to supplant the highly complex quantified logic of the

Middle Ages, so objectionable to humanists. Actually, it exaggerates—at times grotesquely—the quantifying drives built up in medieval Scholasticism. Following the *De Inventione Dialectica* of Rudolph Agricola, Ramus reduced all argumentation to one "art of discourse" (*ars disserendi,* a Ciceronian definition common during the Middle Ages), which he called indifferently dialectic or logic. He thus did away with dialectic as a separate art that argues from probabilities and is thereby distinct from a scientific logic, which argues from certainties or necessity.

RHETORIC. By the same token, he also dispensed with rhetoric as a separate argumentative art persuading to action. The Ramist *Rhetorica* (1548), published under Talon's name but with Ramus's close collaboration (in some 175 known extant editions or adaptations), reduced rhetoric explicitly to mere "ornamentation," or the application of tropes and figures, conforming to what had been, in fact, a strong trend in medieval thinking about rhetoric. Like Agricola, Ramus treated logic or dialectic as made up of *inventio* (discovery of arguments for any kind of discourse, from mathematics to poetry) and *iudicium* or *dispositio* (the arrangement of arguments, including for Ramus not only syllogism but also method, likewise referable to any and all discourse). Ramus's treatment of syllogism varied somewhat from some previous treatments but in no original or insightful way, and he did nothing to advance formal logic. Still, his influence was vast and symptomatic.

LOGIC. In the wake of Scholasticism, logic had a high prestige value even among humanists. Ramus made it accessible to all by withdrawing it, more than even medieval Scholasticism had done, from the scientifically elusive world of sound and word and by associating it more with the sense of vision through overt or covert resort to spatial constructs or models in his teaching. Most notable among these models were the dichotomized divisions, often arranged in bracketed tabular form, for analysis of everything under the sun. One divided a subject into two parts, subdivided each of these into two, then again dichotomized each subdivision, and so on. The resulting structure somehow corresponded both to extramental actuality and to the contents of the mind. The intensified passion for this far-from-new procedure was associated with the new medium of typography, which reproduced these and other spatial constructs with an ease and conviction unknown in a manuscript-oriented civilization.

METHOD. In this climate Ramists gave the term *logical analysis* its first extensive currency and developed concern with method. Between 1543 and 1547 the treatment of method earlier found largely in rhetoric manuals had been transplanted into logic manuals published separately by Johannes Sturm and Philipp Melanchthon. During this period Ramus effected the same transplantation in a pseudonymous 1546 revision of his *Dialecticae Partitiones,* from which method made its way into the *Dialectica* from 1555 on. For Ramus, method prescribed treating any subject by going from the general to the particular, although for special reasons one could use cryptic method, proceeding from the particular to the general. Dichotomization implemented method.

Metaphysics was absorbed or displaced by logic, which Ramus passionately but unconvincingly identified with Plato's dialectic. Ethics was to be taught by methodized analysis of biography and history, and the physics that had formed so great a part of Scholastic philosophy was replaced, in principle at least, by analytic study of works on natural history such as Vergil's *Georgics.*

INFLUENCE

Ramus's realignments involved him in disputes with Antonio de Gouveia, Joachim de Perion, Pierre Galland, Jacques Charpentier, Adrien Turnèbe, Jean Riolan the elder, and Jakob Schegk, disputes protracted after Ramus's death by hundreds of litigants. Ramist-inspired agitation over method set the stage for René Descartes (who at La Flèche studied a post-Ramist logic textbook with a section on method) and helped make meaningful the application of the nickname "Methodists" to John Wesley's followers. The modern encyclopedia owes a good deal of its organization to the Ramist and semi-Ramist tradition as represented by polymath organizers of knowledge such as Johann Heinrich Alsted. Ramus's followers, numbered by the thousands in the sixteenth and seventeenth centuries, were distributed, in descending abundance, through Germany, the British Isles and their American colonies, France, Switzerland, the Low Countries, and Scandinavia. Anti-Ramists such as Nicolas de Grouchy, Everard Digby, and Francis Bacon and Ramists such as Johann Thomas Freige (Freigius), Gabriel Harvey, and John Milton crossbred to produce various syncretists, such as Bartholomew Keckermann, Andreas Libavius, Alsted, and Robert Sanderson. Ramism and its derivatives were particularly popular in Calvinist "middle" or secondary schools for cultural and psychological rather than directly religious reasons: The Ramist account-book interpretation of knowledge and actuality

appealed strongly to the bourgeois mind. Influence in strictly university circles and on speculative thought was more intermittent or indirect, but extraordinarily pervasive.

See also Aristotle; Bacon, Francis; Descartes, René; Logic, History of; Medieval Philosophy; Melanchthon, Philipp; Milton, John.

Bibliography

Ong, Walter J., S.J. "Ramist Method and the Commercial Mind." *Studies in the Renaissance* 8 (1961): 155–172.

Ong, Walter J., S.J. *Ramus and Talon Inventory.* Cambridge, MA: Harvard University Press, 1958.

Ong, Walter J., S.J. *Ramus, Method, and the Decay of Dialogue.* Cambridge, MA: Harvard University Press, 1958. Includes an exhaustive bibliography.

Risse, Wilhelm. *Die Logik der Neuzeit.* Vol. I. Stuttgart: F. Frommann, 1964.

Waddington, Charles. *Ramus.* Paris: C. Meyrueis, 1855.

Walter J. Ong, S.J. (1967)

RANDOMNESS

See *Chaos Theory; Probability and Chance*

RASHDALL, HASTINGS

(1858–1924)

The English theologian, philosopher, and historian Hastings Rashdall was born in London, the son of an evangelical clergyman. He was educated at Harrow and at New College, Oxford, where he read Classical Moderations and "Greats." He remained at Oxford two years after graduation, reading philosophy and theology and working on an essay on the history of medieval universities, for which he won the chancellor's prize in 1883. Much of his next twelve years was taken up with expanding this essay for publication in 1895 as a work in three volumes.

In 1883 he left Oxford to become a lecturer at St. David's College, a college for the education of the clergy in Lampeter, Wales, and in December of that year he was appointed a tutor in theology at University College, Durham. In 1889 he returned to Oxford as a fellow of Hertford College and in 1894 was appointed for a year as chaplain and divinity tutor at Balliol, without relinquishing his Hertford fellowship. He returned in 1895 to New College as fellow and tutor and dean of divinity. He retained his New College fellowship but not his tutorship on his appointment in 1910 as a canon of Hereford Cathedral. He remained in Hereford until 1917, when he became dean of Carlisle, an office he retained until his death.

Rashdall was primarily a theologian and secondarily a philosopher, although he would have been unwilling to draw a clear distinction between the two. His aim was to keep philosophy religious and religion philosophical. Even his history of medieval universities aimed at establishing the rational foundations of religion and ethics, the close connection between the intellectual and spiritual life, and the place of mind in the constitution of the world.

Rashdall justly described himself as "on the left wing of the Church and the right wing of the philosophers." His liberalism in religion and forthright opposition to bigotry kept getting him into trouble with the defenders of orthodoxy. The last years of his life were clouded by the false charge that he denied the divinity of Christ—a charge based on a newspaper misrepresentation of his observation that Jesus never claimed divinity for himself.

Philosophically Rashdall was a personal idealist. Although he held that there is no matter apart from mind—a personal Mind, "in which and for which everything that is not mind has its being"—he rejected monism. Minds are substantial, and every consciousness is exclusive of every other. Individual minds are produced by the eternal Mind, which is God, but are neither included in it nor adjectives of it. In line both with this metaphysical position and with his general distrust of mysticism, Rashdall held our knowledge of God to be inferential.

Rashdall's most important philosophical work is his two-volume *The Theory of Good and Evil.* Although it made no distinctively original contribution to ethics, it is perhaps the best general introduction to the subject written from an objectivist point of view, before the advent of metaethics and the application of philosophical analysis. Rashdall's treatment is thorough and comprehensive, and the book leaves no doubt about the importance for theory and practice of the issues discussed. Although it is not a history of ethics, it includes illuminating expositions and criticisms of theories of classical moral philosophers where these are relevant to the development of his own theme.

Rashdall's emphasis on the value of human personality found expression in his moral theory. Intuitionism, in the sense of acceptance of impersonal moral laws binding independently of their consequences, was wholly

alien to his thought. He was an uncompromising utilitarian, for whom actions are to be judged by their tendency to produce the greatest good or well-being for human beings. There are, indeed, moral intuitions, but they are about the relative value of ends, not about the rightness of rules of conduct. The good that it is the duty of each to produce for all is a personal good but is not confined to pleasure or happiness. Pleasure is only one element that, in interrelation with other mutually modifying elements, including morality, contributes to form an ideally good pattern of life. It was Rashdall who coined the term *ideal utilitarianism* to distinguish this form of the theory from the traditional hedonistic utilitarianism it has generally replaced, partly through his own influence. One advantage of the abandonment of hedonism claimed by Rashdall is that it enables the utilitarian to include in moral judgment the quality of the act itself as well as of its consequences. Thus, the disposition to promote the general good can be taken as itself part of the good to be promoted.

Much of the second volume of *The Theory of Good and Evil* deals with the metaphysical and theological presuppositions of an absolute objective morality. Rashdall held that only in metaphysics can we find an ultimate defense of the validity of moral judgments and that personal idealism has the best chance of supplying it. One postulate of morality is the existence of individual selves to which actions may be attributed; another is the existence of God, as possessing and willing the absolute moral ideal; and a third is immortality. Although he was a determinist, Rashdall escaped having to hold God responsible for evil in human willing because he regarded God not as strictly omnipotent but as limited by those eternal necessities that are part of his own nature.

See also Analysis, Philosophical; Idealism; Metaethics; Religion and Morality; Utilitarianism.

Bibliography

Rashdall's works include *The Universities of Europe in the Middle Ages,* 3 vols. (Oxford: Clarendon Press, 1895); "Personality Human and Divine," in *Personal Idealism,* edited by Henry Sturt (New York: Macmillan, 1902), *The Theory of Good and Evil,* 2 vols. (London, 1907; 2nd ed., London: Oxford University Press, H. Milford, 1924); *Philosophy and Religion* (London: Duckworth, 1909); *The Problem of Evil* (Manchester, U.K., 1912); *Is Conscience an Emotion?* (Boston: Houghton Mifflin, 1914); and *The Moral Argument for Personal Immortality* (London, 1920).

For a discussion of Rashdall, see P. E. Matheson, *The Life of Hastings Rashdall* (London: Oxford University Press, H. Milford, 1928).

A. K. Stout (1967)

RATIONAL INTUITION

See *Intuition*

RATIONALISM

The term *rationalism* (from the Latin *ratio*, "reason") has been used to refer to several different outlooks and movements of ideas. By far the most important of these is the philosophical outlook or program that stresses the power of a priori reason to grasp substantial truths about the world and correspondingly tends to regard natural science as a basically a priori enterprise. Although philosophies that fall under this general description have appeared at various times, the spirit of rationalism in this sense is particularly associated with certain philosophers of the seventeenth and early eighteenth centuries, the most important being René Descartes, Benedict de Spinoza, and Gottfried Wilhelm Leibniz. It is rationalism of this type that will be the subject of this entry.

Two other applications of the term should, however, be distinguished.

RATIONALISM IN THE ENLIGHTENMENT

The term *rationalism* is often loosely used to describe an outlook allegedly characteristic of some eighteenth-century thinkers of the Enlightenment, particularly in France, who held an optimistic view of the power of scientific inquiry and of education to increase the happiness of humankind and to provide the foundations of a free but harmonious social order. In this connection "rationalistic" is often used as a term of criticism, to suggest a naive or superficial view of human nature that overestimates the influence of benevolence and of utilitarian calculation and underestimates both the force of destructive impulses in motivation and the importance of such nonrational factors as tradition and faith in the human economy. Jean d'Alembert, Voltaire, and the Marquis de Condorcet, among others, are often cited in this connection. Although there is some truth in these criticisms, the naïveté of these and other Enlightenment writers has often been grossly exaggerated. Also, insofar as "reason" is

contrasted with "feeling" or "sentiment," it is somewhat misleading to describe the Enlightenment writers as rationalistic, for many of them (Denis Diderot, for example) characteristically emphasized the role of sentiment. Reason was praised in contrast with faith, traditional authority, fanaticism, and superstition. It chiefly represented, therefore, an opposition to traditional Christianity.

Here there are two contrasts with the seventeenth-century rationalism of Descartes and others. First, this rationalism is not characteristically antireligious or nonreligious; on the contrary, God in some sense, often in a traditional sense, plays a large role in rationalist systems (although Spinoza's notion of God was extremely unorthodox, and it is notable that the opposition of reason and faith is important in his *Tractatus Theologico-Politicus*). Second, the view of science held by such Enlightenment thinkers as Voltaire was different from that of rationalism, being much more empiricist. The central contrast embodied in the term *rationalism* as applied to the earlier systems is that of reason versus experience, a contrast that is certainly not present in the Enlightenment praise of the "rational." Parallel to this difference, there is a difference between the characteristic seventeenth- and eighteenth-century views on the nature and importance of system; the eighteenth century declared itself against the *esprit de système* of the seventeenth century, with its elaborate metaphysical systems, and in favor of an *esprit systématique*, which could be orderly without being speculatively ambitious. (See d'Alembert, *Preliminary Discourse* to the *Encyclopédie* and Condillac, *Traité des systèmes*.)

RATIONALISM IN THEOLOGY

The Enlightenment spirit of rational criticism directed against the supposed revealed truth of the Scriptures also had effects within Christianity itself. In this connection the term *rationalism* is used in a specific theological sense to refer to the doctrines of a school of German theologians that was prominent roughly between 1740 and 1840, and which had great influence on the development of biblical criticism. With their spirit of antisupernaturalism can be associated Immanuel Kant's *Die Religion innerhalb der Grenzen der blossen Vernunft* (1793), in which rational morality is the basis of religious belief.

However, the best-known use of "rationalism" in a religious connection is an entirely negative one, in which it stands for an antireligious and anticlerical movement of generally utilitarian outlook, laying great weight on historical and scientific arguments against theism. This use of the term, a popular rather than a technical one, seems now to be obsolescent, its place being taken by *humanism*.

RATIONALISM VERSUS EMPIRICISM

Rationalism as it will be discussed here is standardly contrasted with empiricism. This contrast (which rests on that contrast between reason and experience which has already been mentioned) is now so basic to the use of the terms that no account can afford to ignore it, and a number of comparisons between views associated with these two outlooks will be made in the course of this entry. It is of course impossible to give a detailed comparison of the two outlooks, and in general comparisons will be introduced incidentally to the account of rationalist ideas. There is, however, one issue, that of innate ideas, which embodies a central disagreement between the two, and regarding which an account merely from the rationalist side would be particularly unilluminating. On this issue the disagreements will be considered in rather greater detail than elsewhere. At the same time, it is hoped that the treatment of this issue will give slightly more insight into the rationalist outlook than can be achieved by what is at other points inevitably a very selective summary of rationalist opinions.

INNATE IDEAS

Descartes distinguished three classes of "ideas" (by which he meant merely whatever it is in a man's mind in virtue of which he can be said to be thinking of a given thing): adventitious, factitious, and innate. The first type came to the mind from experience, the second were constructed by the mind's own activity, and the third were created by God together with the mind or soul itself. The last included what were for Descartes the three fundamental ideas of the basic types of substance: God, mind, and matter (or extension). For the most part Descartes argued negatively for the view that these ideas are innate, trying to show that they could not be derived from experience (where this means, fundamentally, sensation).

His argument had two main points. First, the ideas are pure, containing no sensory material; these ideas are not images, reproductions, or copies of sensory experience. Descartes regarded this as fairly obvious in the cases of God and of mind; and he made a particular effort (as in the argument of the wax in *Meditations* II) to establish the same claim for matter. Second, the fundamental ideas implicitly contain, in different ways, some idea of infinity, and in grasping the idea one thereby grasps the possibility of infinitely many and various modifications to which

mind and matter can be subject. In the case of God this argument goes further, for here we grasp an actual infinity of perfections implicit in the idea. The same point, however, holds for all these ideas: The grasp of infinitely many possibilities must transcend what has been given to us in experience, since experience could have given us at best only a limited set of such conceptions, corresponding to what had actually been experienced.

Even if both points of the argument are granted as showing that these ideas are not totally derived from experience, it might be doubted whether they are enough to show that the ideas are innate. For might they not be grasped in some nonempirical manner at a later stage of life—for example, when (or if) someone comes to think in these very general terms? In Descartes's philosophy there is at least an implicit answer to this objection. Descartes thought that the pure ideas of mind and matter are used in the comprehension of experience even before they become conscious in reflection. It is by reference to these ideas that one forms the ordinary unreflective conceptions of oneself as having a series of thoughts or of a material object as enduring, occupying space, and having various characteristics, even though, before reflection, one's conceptions of these things will be confused. Thus, the operation of the pure idea is implicit in ordinary prereflective experience, and such experience begins to be acquired from the moment of birth; therefore, there is ground for calling the pure ideas innate.

In the case of God the argument is slightly different, since it is less clear that this idea is "put to use" in any prereflective way. Here Descartes may have meant to claim merely that it would be natural to the power and economy of God's operations that he should implant the idea of himself in the soul at its creation, "the mark," as Descartes put it, "of the workman on his work." There is indeed a difficulty in seeing how, for Descartes, there could be an idea in the mind of which the mind is not fully conscious (as this account implies), since for Descartes "mind" and "consciousness" were virtually equivalent. And this difficulty also arises for the ideas of mind and matter, since Descartes explicitly denied (presumably there was no alternative) that the infant or young person is fully conscious of his innate ideas; they are latent and emerge only later—in the process of learning language, for instance. Nevertheless, Descartes's claims for the operation of fundamental ideas in prereflective consciousness, although not quite consistent with his metaphysics of the mind, became an important element in later theories of innate ideas, especially in the debate with empiricism.

INNATE PRINCIPLES. Descartes appealed only to innate ideas, or concepts, the materials of judgments and beliefs. He did not invoke innate principles, or propositions, his view apparently being that granted innate ideas, we have only to grant in addition a certain power of the mind to elicit features implicit in these ideas in order to explain how necessary knowledge could be derived from innate ideas (as he supposed it could).

Leibniz, however, who continued the Cartesian insistence on innate ideas, added a requirement for innate principles. His argument was of the same general type as that ascribed to Descartes with respect to the ideas of mind and matter: If there were no innate and unlearned propositions, we could learn no propositions at all—at least not by way of logical deduction. For, he argued, confronted with any valid inference of the form "*P*, so *Q*," we could not see that *Q* followed from *P* except by having already grasped the necessary truth of the proposition "if *P* then *Q*." Thus, in order to follow any inference and to learn anything by deduction, first premises are required that must themselves be unlearned.

An objection to this argument can be seen from the famous difficulty raised by Charles L. Dodgson (Lewis Carroll) that if there is necessarily a difficulty in seeing the validity of the original inference as it stands, the same difficulty will recur with the inference obtained by the addition of the "innate" major premise; to grasp the validity of this inference, another major premise would seem to be required, and so on, thus starting a vicious regress. Dodgson's point makes it clear that no multiplication of premises can be adequate to extricate the validity of an inference; what is needed is something of a different category, a rule. At this point a characteristic empiricist rejoinder to Leibniz's puzzle is to claim that the rules of inference are not unlearned but are learned in the course of learning a language (they are the rules implicit in the correct use of "if," "then," "not," and so on). This illustrates the natural and perhaps inevitable tendency of empiricism, in contrast with rationalism, to turn to a linguistic account of logical necessity. (Such an account, however, even if adequate in itself, may not dispose of the issues as thoroughly as empiricism has tended to believe; the question remains of what is involved in learning a language.)

Leibniz, in introducing the argument just considered, explicitly stated that he was of the "Platonic" opinion that a priori knowledge (at least) is innate and "recollected" (*New Essays*, Book I). There is a difficulty, however, in knowing how far the doctrine is supposed to range: Leibniz's doctrine that the soul is a monad and

that every monad only develops its own inner potentialities, being unaffected by anything outside, implies that in one sense all thoughts, of whatever kind, are innate. This problem involves major questions in the interpretation of Leibniz—in particular, of his views on sense perception. However, it seems reasonable to say that at least in his remarks on innateness in the *New Essays* Leibniz was distinguishing between kinds of knowledge and ideas, such that some (the pure and a priori) can be said to be innate and others cannot.

Leibniz's remarks were in criticism of the First Book of John Locke's *Essay concerning Human Understanding*, and they constitute a subtle consideration of the issues lying between rationalism and empiricism at this point. Locke's First Book, although called "Of Innate Ideas," is in fact chiefly concerned with innate principles (and in some part with the alleged innate principles of morality that had been advanced by his adversary, Lord Herbert of Cherbury). Locke considered various characteristics supposed to show that a given proposition is innate (that it is universally believed, that it is assented to as soon as understood, and so forth), and had little difficulty in showing that these are inadequate. He then turned to the consideration that tiny children do not display elaborate conceptions of logic and mathematics such as are alleged by rationalists to be innate. His principle in this instance was "There is nothing in the mind of which the mind is not conscious"; if these conceptions were innate, they would be in the infant's mind, hence it would be conscious of them and (presumably) could display this consciousness. Leibniz, in reply, claimed that this so patently follows that Locke, in insisting on the principle about consciousness, was in effect begging the question: This principle is what the issue turns on. But, as has been seen, this was not how Descartes put the matter. Leibniz here made the cardinal point of the discussion his own non-Cartesian doctrine of subconscious perceptions (connected with his general doctrine of continuity).

DEBATE WITH EMPIRICISM. Once the obvious fact is granted that the allegedly innate ideas do not manifest themselves temporally before other experience, it may be wondered whether any point remains to calling them innate. It has sometimes been suggested that the doctrine of innate ideas merely depends on a confusion between a logical and a temporal sense of "prior." However, this is to underestimate the force of the rationalist claims that the allegedly innate material is such that its operation is a precondition of our learning anything else. It is not easy to decide how to evaluate these claims, as against the central empiricist claim that no such preexisting material need be postulated (the so-called tabula rasa theory of the mind). For one thing, empiricism in its first developments tended to make up for the lack of original raw material by crediting the mind with a very elaborate set of operations. This was evidently the case with Locke, who used such notions as "abstraction," "reflection," and "intuition"; who spoke of "ideas" that are not evidently mere copies from sense perception; and who admitted a nonempirical notion of "substance" and its powers. His position retained a number of rationalist elements. The much more economical apparatus of David Hume, which in effect admits nothing but sensations, their copies, and the operations of association, defines a quite distinctive empiricist theory.

If the debate about innate ideas is cast in terms of a Humean empiricism, there remain principally two issues, one logical and one psychological. The logical issue concerns the question whether highly general concepts, such as those used in mathematics and the sciences, are reducible to or analyzable into those sorts of empirical concepts that can plausibly be said to be derived from sense experience. It would be widely agreed that the answer to this question would be "no." The second, psychological issue is whether the acquisition of concepts, such as occurs in language learning—and this would include even the supposedly straightforward empirical concepts—can be adequately explained by a psychological model postulating only the minimum empiricist requirements of sense perception, retention, association, and so forth. There is influential opinion (held by Noam Chomsky and others) that the answer to this, too, must be "no"; any adequate model may well require stringent innate constraints on the direction and nature of generalization from learning situations. How far these restraints might be supposed to approximate to the rationalists' conceptions of innate ideas—or, in other words, whether the model demands an innate analogue to the possession of concepts—remains to be seen.

If this is indeed an open question, then there is an explicitly psychological version of the rationalist view that is still worth serious consideration. This is not, of course, to say that the innate elements in an adequate model would be likely to correspond to the particular sorts of "ideas" that the rationalists selected for this status—such as the metaphysical notions of God, matter, and mind. Also, there was certainly an endemic confusion, in both the rationalist and the empiricist position on this issue, between psychological and logical issues. Nevertheless, there is still some life in the question, in both its logical and its psychological aspects, the occur-

rence of the psychological term *innate* in the original debate not being merely the result of confusion.

KNOWLEDGE

It was remarked above that there would now be wide agreement that many general theoretical concepts of mathematics and the sciences do not admit of total reduction to empirical concepts. In contrast with positivist or operationalist views it would be agreed by many that such concepts as "mass," for instance, are not a mere shorthand for sets of possible observation data. Such agreement, however, although it would constitute a rejection of strict empiricism, would not in itself constitute an acceptance of rationalist views about such concepts. It is possible to think that these concepts "transcend," or "go beyond," the empirical merely in virtue of conventional elements—that they are parts of humanly constructed models of reality which relate the observable by imposing a structure on it.

Essential to rationalism, however, is a realistic view (incompatible with even a modified empiricism) about the relation of these concepts to reality and about the necessary relations obtaining between these concepts themselves. The intellectual grasp of these concepts and the truths involved in them is seen as an insight into an existing and unique structure of the world. It is not easy to express this picture (which in varying degrees dominated the rationalists) in less figurative language, but the picture has at least two consequences: that there is a unique set of concepts and a unique set of propositions employing these concepts that adequately express the nature of the world, and that these propositions form a system and could ideally be recognized as a set of necessary truths. There are, admittedly, difficulties about the last point, particularly with reference to Leibniz (these will be considered in the next section). However, something like this general picture is central to rationalism and leads immediately to the question of how anyone can come to know this uniquely correct representation of the world. This invites two more specific questions: What, in general, is the guarantee that knowledge of the world is possible? how can any individual tell in a particular case whether he has hit on some genuine piece of knowledge?

DESCARTES'S EPISTEMOLOGY. Most rationalists tended to answer the first of the above questions by referring to God; some, but not all, did the same for the second; and they varied in the priority that they assigned to the two questions. Descartes started famously with the second question and found the answer in the "clear and distinct perceptions" of the intellect. Proving, as he supposed, the existence of God via clear and distinct perception, he then employed God's perfection of "being no deceiver" to establish in general terms the reliability of beliefs that went beyond clear and distinct perception. He was, however, so impressed by the thought that it was only in virtue of humanity being created and sustained by God that he could know anything at all, that he was constantly tempted to double back and use the divine perfection to guarantee even the basic clear and distinct perceptions, thus laying himself open to the charge of arguing in a circle.

However this may be, it is notable that in Descartes "clear and distinct perception" is a thoroughly epistemological category. The truths that can be clearly and distinctly perceived do not constitute one homogeneous logical or metaphysical class of truths; the class includes at least statements of contingent existence (his own, in the *cogito*) and of necessary existence (that of God), contingent statements about immediate psychological experience, and necessary truths about the relations of ideas. The status of these last, which Descartes called eternal truths, is somewhat obscure. Descartes held, in the Augustinian-Scotist tradition, that they were the products of God's will; but it is left unclear what it is that God has brought about in creating eternal truths, and hence what it is that one knows in knowing them.

THE CARTESIAN TRADITION. The development of the Cartesian tradition within rationalism tended to emphasize to an even greater extent the theological elements in Descartes's theory of knowledge. Thus Nicolas Malebranche retained for the individual case the test of "clear and distinct perception" in a style that seems to assimilate it to moral perception and the promptings of conscience: "One should never give one's complete assent except to propositions which seem so evidently true that one could not reject them without feeling an interior pain, and secret reproaches of the reason" (*De la recherche de la vérité*, I, Ch. 2; for the moral analogue, see Bossuet, *Traité de la connaissance de Dieu et de soi-même*, Ch. 1, Sec. 7).

Malebranche gives a strongly Augustinian and indeed Neoplatonist turn to the general account of God's guarantee of the possibility of knowledge. His doctrine was that all our knowledge of the external world is mediated by God; the mind of God contains paradigm ideas in whose form he created the world, and it is these same ideas of which we are conscious when thinking about the world. This is the meaning of Malebranche's saying that we see all things in God. This doctrine, apart from serving religious purposes, was also an attempt to get around

the difficulties inherent in Descartes's own causal account of relations between matter and mind (which will be considered more generally later in this entry).

The role of God in the foundations of knowledge takes different and less extreme forms in other areas of the rationalist tradition. The greatest contrast to the Malebranche development of Cartesianism might plausibly be said to be Spinoza's system. It is true that Spinoza did assert that it is the nature of God that guarantees the correspondence of our thoughts to the world, but he so transmuted the notion of God that the doctrine is only verbally similar to Cartesianism. "God" is one name ("Nature" is another) for the one substance, that is, everything that there is. This substance has infinitely many attributes, of which we can comprehend only two, mind and matter. These two attributes are necessarily parallel to one another, and corresponding to any mode of the one attribute there must be a mode of the other. Hence, thought and the material world are inherently adjusted to one another, and the development of knowledge consists in the project of rendering the thought component of this relation as clear (in Spinoza's term, as "active") as possible. It admittedly remains obscure how, within the constraints of Spinoza's determinism, this can be regarded as a "project" at all. Despite this and the other notorious difficulties, Spinoza's system is particularly interesting in the present connection as a thoroughgoing attempt to answer the crucial question that was left very much in the air in Descartes's thought, namely, how any knowledge of a necessary truth, regarded as knowledge of the relations of ideas, could also constitute knowledge of the world.

Leibniz's system, for all its radical differences from Spinoza's, resembles it in one respect having to do with the foundation of knowledge: The general possibility of the correspondence of thought to the world is guaranteed metaphysically by the existence of a correlation between the two. The monads are not affected by anything outside and each develops its own activity from within, but a correspondence between the activities of the monads is given by the "Preestablished Harmony"; and knowledge, the correspondence between "conscious" states of certain monads and other monads, is a special case of this. The Preestablished Harmony, however, depends on God's optimal choice, that is, on God's benevolence. Thus, in a less explicitly epistemological form, Leibniz (in contrast with Spinoza) reverted to the original Cartesian standpoint, in that there is a transcendent and personal God who has a will, and it is a result of his will that there is an ultimate guarantee of the possibility of knowledge.

In general, however, Leibniz was not much concerned with epistemological problems; in particular, he was uninterested in the question that was the starting point for Descartes: How can the individual be certain of the truth of anything? Spinoza was concerned with this question, and tried to develop a theory of knowledge that would avoid the regress latent in Descartes's method, arising from the question of how one knows that one knows. In Spinoza's "degrees of knowledge" it is an essential property of the highest, or intuitive, degree that it is self-guaranteeing. Even so, there is an evident shift in the Spinozistic outlook away from the Cartesian question "What do I know, and how do I know it?" Spinoza, like Leibniz and many other rationalists, gave the metaphysical description of the world from "outside," from a "God's-eye" standpoint rather than from the subjective epistemological standpoint from which Descartes (although unsuccessfully) tried to work. It is, perhaps, a mild irony of the history of philosophy that Descartes's attempt to start with subjective questions of epistemology and to "work out" from there had more influence on the development of empiricism than on later rationalism.

SCIENCE AND SCIENTIFIC METHOD

No attempt will be made here to give an account of the detailed developments of the philosophy of science within rationalist thought, or of the actual scientific conceptions held by or associated with rationalists, although these are of course of great importance, most notably in Leibniz's critique of Cartesian physics and in the development of his concept of force. We shall consider only one or two general points about the rationalists' conception of a completed science and associated notions of scientific method.

Rationalist developments in these matters can usefully be seen in the light of an unresolved conflict within Descartes's system between the method of approaching scientific inquiry and the expected shape of the final product. Descartes favored in principle an approach to inquiry that might be called systematically exploratory. This he called the analytic method; and the straightforward exposition of the results of such an inquiry would be heuristic in style, explaining the resolution of difficulties as they were encountered in the systematic progress. He seems, however, also to have had a picture of a completed science as a complete deductive system, ideally expressed in a unique system of theorems with necessary truths (of a metaphysical character) as its axioms; this he termed the synthetic method of exposition. There is, perhaps, no essential clash between these two ideas of

method and result; but Descartes seems not to have been clear about the relation between the two or how this specific method, fully pursued, would yield this specific result. Ambiguities about this question emerge in Descartes's accounts of the role of experiment, in which he sometimes gives the incoherent impression that he is both engaged in logical deduction of scientific laws from self-evident metaphysical premises and doing experiments to assist him in this deduction. On the whole, it is probably better to regard the idea of a complete formally deductive metaphysico-scientific system as less important in Descartes's thought than is sometimes supposed, and to see him as using certain limiting principles of scientific explanation, within which he constructs models to explain particular phenomena.

The idea of the total deductive system, however, had a powerful effect on rationalism and reached its most extreme expression in the work of Spinoza, where the "synthetic" method of Euclidean demonstration is explicitly regarded as necessary to the highest form of understanding. This was not just an expository preference; it was an expression of the basic Spinozistic outlook, which regarded the relation of cause to effect as that of logical ground to consequence—for Spinoza all explanatory relations were logical and timeless. The parallel orders of thought and matter, remarked on earlier, supposedly guarantee that the logical relations of ideas will constitute a totally adequate expression of the nature of the world. (A singular application of this notion of total parallelism is to be found in Ehrenfried Walter von Tschirnhaus, who in *Medicina Mentis* [1687] argued that an adequate definition of laughter should be able to produce laughter.)

Leibniz, partly under the influence of Erhard Weigel, was also attracted to the "geometrical method." He devoted a good deal of effort to the project of a universal calculus, which would enable arguments on any subject matter to be cast into a rigorous demonstrative form. However, the idea of such a calculus in no way presupposes an ideal of being able to demonstrate scientific truths from metaphysical or other supposedly self-evident axioms, which was the Spinozistic and, on occasion, the Cartesian ideal. Even if Leibniz started with the notion that it should be possible to settle any argument by appeal to the self-evident, he abandoned it in his mature philosophy, in which he made fundamental the distinction between "truths of reason," which can be established by logical insight on the basis of the law of noncontradiction, and "truths of fact," which depend on the principle of sufficient reason and cannot be established on logical grounds alone. There are some notorious difficulties about this distinction, especially concerning the question of the nature of the contingency of "truths of fact," since Leibniz also held the further general principle that in all true propositions the predicate is contained in the subject. It does seem clear, however, that there is an ineliminable contingency about "truths of fact," and hence that the aspiration of reducing all knowledge to a system of deductions from self-evident premises must be impossible in the Leibnizian system.

Francis Bacon said in his *Cogitata et Visa* (1607), "Empiricists are like ants, they collect and put to use; but rationalists, like spiders, spin threads out of themselves." Bacon, of course, preferred the ants. Although there is some element of truth in the image of the spider, as applied to some rationalist thinkers, it does less than justice to the substantial empirical work done under rationalist inspiration. This is all the more so if one counts Galileo Galilei's view of science as fundamentally rationalist. He certainly rejected any kind of Baconian empiricism and shared the rationalist vision of a mathematical structure of reality that intellectual insight could grasp; but he perhaps had a more sophisticated feeling than any of the philosophers for the balance of imagination and experiment in physics. The rationalist tradition certainly embodied fundamental insights (lacking in empiricism) about the nature of science; above all, it saw the importance of mathematical structures in physical explanation and the vital possibility of a theory's making a conceptual jump beyond the observations and not merely (as in empiricism) an advance in generality. Its sense of the activity of the scientific mind, of its restructuring of observations through concepts and models, was very significant. At the same time, empiricism rightly fought for a clearer distinction between pure mathematics and natural science, undermined the aspirations to final certainty that dogged the rationalists, and emphasized the role of laborious observation and experiment in contrast with the rather dreamlike quality of rationalist visions of the universe. No clearer case exists in the history of philosophy of the need for, and eventual occurrence of, a synthesis; one aspect of that synthesis is neatly summed up in a remark of Giorgio de Santillana that "the true scientist has an empiricist conscience and a rationalist imagination."

SUBSTANCE AND CAUSALITY

In the history of classical empiricism the concepts of substance and of active causal power together became progressively weaker and were finally abandoned. Thus Locke employed the full Cartesian array of both material

and mental substances, both possessing causal power; George Berkeley banished material substance, partly on the ground that it could not be conceived of as possessing causal activity, which belongs only to mental substance; Hume maintained that the notions of substance and of causal activity are unintelligible. By contrast, in the rationalist tradition the notion of substance has not declined; developments in the idea of causal activity, although partly parallel to the idea of substance, are very different; in general the fortunes of "substance" and of "causal activity" have not been directly linked, as they have proved to be in empiricism—both have undergone considerable and partly independent variations.

In the case of substance (which will be very briefly considered here) the concept has not so much been criticized as used in differing ways to express differing metaphysical views of the world. On one measure, at least, the extremes in this respect are represented by the philosophies of Spinoza and Leibniz. Spinoza gave what he claimed was an a priori demonstration that there could be only one substance (*Deus sive Natura*, God or Nature); this was intrinsically neither material nor mental, these distinctions arising (as noted above) only at the level of the different attributes of this same substance. Essential to Leibniz's outlook, on the other hand, was an infinite set of substances, the monads, each of them different from all the others. In their character, although there are difficulties of interpretation on this point, they are more of a mental than of a material kind.

On the question of causality an important stream in the history of rationalism stems from the problem left by Descartes, concerning the causal interaction of mind and matter. Descartes's own view, which postulated simple efficient causation as holding between the two types of substance, failed to appeal to even the most ardent Cartesians, and their attention was particularly directed to this question, although difficulties about the meaning of causation even between material bodies also were considered. The natural tendency in the Cartesian tradition was to move toward attributing all causal power to God, and this movement of thought culminated in the doctrine of occasionalism—that both physical and mental events in the world are occasions for the application of God's power, which itself directly produces what would normally be called the effects of those events. This doctrine is most thoroughly expressed in the writings of Malebranche. Similar views, however, are to be found in Louis de la Forge (*Le traité de l'esprit et de l'homme*, 1666) and Géraud de Cordemoy (*Le discernement du corps et de l'âme*, 1666), whose work was known to Malebranche.

The theory of occasionalism can be usefully contrasted with Berkeley's empiricist account of causation. For both the only genuine activity was spiritual. For Berkeley the effects of such activity were also spiritual (mind can affect only mind), and indeed there was no other type of substance. The occasionalists retained material substance and did not find it unintelligible that mind can act upon matter; however, they held that the only mind for which such action is intelligible is the infinite mind of God. Here, as elsewhere, the questions of the gulf between mind and matter and of causation as activity emerge as of common concern to both rationalist and empiricist metaphysics, the influence of Descartes being clearly discernible in both.

Another writer who inclined to occasionalism was Arnold Geulincx (*Ethics*, 1665; 2nd ed., 1675); however, he also suggested a different model for causality, in which God did not, as in occasionalism, make a constant series of miraculous interventions into the natural order but had established ab initio a series of coordinated developments, the relations between which are what is taken for causal interaction. In this connection Geulincx introduced the example of the two clocks, perfectly adjusted to keep the same time, one of which strikes when the other shows the hour; the appearance of causal connection between them is only a result of precise prearrangement.

This same analogy was frequently employed by Leibniz in explaining his own very thoroughgoing version of this thesis, in which all appearance of causal interaction is an instance of the preestablished harmony between the several developments of the monads. Here again there is a notable contrast with and a similarity to empiricism: Both Leibniz and Hume, each representing the culmination of one of the two traditions in its classical form, deny the existence of "transeunt action" between different things and see what is called causation as a correlation between phenomena. Leibniz, however, emphasized some kind of spontaneous activity within the monad, while for Hume neither such activity, nor any notion of a substance, such as a monad, was acceptable. The views of these two philosophers are also worthy of comparison on other subjects, such as space and time; and the points of contact between them are the more significant in the light of the radical and very obvious differences in the spirit, method, and presuppositions of their two philosophies. These differences in the two culminating figures constitute a paradigm, almost a caricature, of the divergent styles of thought associated with rationalism and empiricism, while at the same time similar pressures in the his-

tory of thought produced partly parallel developments in each.

See also Alembert, Jean Le Rond d'; A Priori and A Posteriori; Augustinianism; Bacon, Francis; Berkeley, George; Carroll, Lewis; Cartesianism; Chomsky, Noam; Condorcet, Marquis de; Cordemoy, Géraud de; Descartes, René; Diderot, Denis; Empiricism; Encyclopédie; Enlightenment; Experience; Geulincx, Arnold; Herbert of Cherbury; Humanism; Hume, David; Innate Ideas; Kant, Immanuel; Knowledge, A Priori; Leibniz, Gottfried Wilhelm; Locke, John; Malebranche, Nicolas; Neoplatonism; Propositions; Reason; Scientific Method; Scotism; Spinoza, Benedict (Baruch) de; Spinozism; Tschirnhaus, Ehrenfried Walter von; Voltaire, François-Marie Arouet de.

Bibliography

For an understanding of rationalism, writings by and about particular rationalist philosophers should be read; references to these will be found in the bibliographies of the appropriate entries.

GENERAL WORKS

Bouillier, Francisque. *Histoire de la philosophie cartésienne.* 3rd ed. Paris, 1868.

Cassirer, Ernst. *The Philosophy of the Enlightenment*. Translated by Fritz Koelln and James Pettegrove. Princeton, NJ: Princeton University Press, 1951. As the title indicates, this is not directly on seventeenth-century rationalism but relates to it.

Cottingham, John. *Rationalism*. London: Paladin, 1984.

Doney, Willis. "Rationalism." *Southern Journal of Philosophy*, Supp. 21 (1983): 1–14.

Lecky, W. E. H. *History of the Rise and Influence of the Spirit of Rationalism in Europe*. London: Longman, Green, Longman, Roberts, and Green, 1865; rev. ed., London, 1910.

Santillana, Giorgio de, and Edgar Zilsel. "The Development of Rationalism and Empiricism." In *International Encyclopedia of Unified Science.* Vol. II. Chicago: University of Chicago Press, 1941.

Windelband, Wilhelm. *A History of Philosophy*. Translated by J. H. Tufts. New York, 1901. Part IV, Ch. 2; Part V, Ch. 1. Cites the works and—summarily—the views of many philosophers, including minor ones.

PHILOSOPHY OF SCIENCE

Burtt, E. A. *The Metaphysical Foundations of Modern Physical Science*, rev. ed. New York, 1952; paperback reprint, Garden City, NY: Doubleday, 1954.

Dijksterhuis, E. J. *The Mechanization of the World Picture*. Translated by C. Dikshoorn. Oxford: Clarendon Press, 1961.

Gillispie, C. C. *The Edge of Objectivity*. Princeton, NJ: Princeton University Press, 1960.

INNATE ELEMENTS IN LANGUAGE ACQUISITION

Fodor, J. A., and J. J. Katz, eds. *The Structure of Language*. Englewood Cliffs, NJ: Prentice-Hall, 1964. Chs. 21 and 22. Papers by Noam Chomsky and E. H. Lenneberg, respectively.

OTHER RECOMMENDED TITLES

Bealer, George. "Modal Epistemology and the Rationalist Renaissance." In *Conceivability and Possibility*, edited by Tamar Szabo Gendler. Oxford: Oxford University Press, 2002.

BonJour, Laurence. *In Defense of Pure Reason*. Cambridge, U.K.: Cambridge University Press, 1998.

BonJour, Laurence. "A Rationalist Manifesto." *Canadian Journal of Philosophy*, Supp. 18 (1992): 53–88.

Brink, David. "Kantian Rationalism: Inescapability, Authority, and Supremacy." In *Ethics and Practical Reason*, edited by Garrett Cullity. New York: Clarendon Press, 1997.

Carruthers, P. *Human Knowledge and Human Nature*. Oxford: Oxford University Press, 1992.

Gendler, Tamar Szabo. "Empiricism, Rationalism, and the Limits of Justification." *Philosophy and Phenomenological Research* 63 (2001): 641–648.

Markie, Peter. "Rationalism vs. Empiricism." Stanford Encyclopedia of Philosophy. Available from http://www.plato.stanford.edu.

Peacocke, Christopher. *The Realm of Reason*. Oxford: Clarendon Press, 2004.

Peacocke, Christopher. "Three Principles of Rationalism." *European Journal of Philosophy* 10 (2002): 375–397.

Bernard Williams (1967)
Bibliography updated by Benjamin Fiedor (2005)

RATIONALISM IN ETHICS (PRACTICAL-REASON APPROACHES)

Practical-reasoning theory is a kind of metaethical view—alongside noncognitivism and other cognitivisms such as naturalism and rational intuitionism—that aims to understand ethics as rooted in practical reason.

Tradition divides the faculty of reason into two parts: theoretical and practical. Theoretical reason concerns what we should believe, practical reason what we should do. Beliefs aim to represent reality and are mistaken or in error when they do not. Theoretical reason's task, therefore, is to discover what is true of the independent order of fact to which belief is answerable. But what about practical reason? What could make it the case that an action is something a person ought to do?

Plainly, ethical convictions also aim at a kind of objectivity. If Jones thinks he should devote all his resources to conspicuous consumption but Smith thinks that Jones should donate some to help the poor, their convictions conflict. Only one, at most, can be true.

Practical-reasoning theories aim to explain the objective purport of ethical conviction, but in a way that respects a fundamental distinction between theoretical and practical reason. Like noncognitivism, these theories sharply distinguish between ethics and those theoretical disciplines that aim to represent some independent reality, whether the order of nature or some supersensible metaphysical realm. They therefore reject both naturalism and rational intuitionism. But they also deny noncognitivism, since they hold that ethical propositions can be true or false.

According to practical-reasoning theories, objectivity consists not in accurate representation of an independent order, but in demands that are universally imposed within an agent's own practical reasoning. What marks ethics off from science is its intrinsically practical character, its hold on us as agents. It is because there is such a thing as practical reason, a form rational agents' deliberations must take, that there is such a thing as ethics.

But what form does rational deliberation take? Uncontroversially, practical reasoning includes reasoning from ends to means. The interesting debates concern what else it involves, if anything, and how instrumental reasoning is itself to be understood. Humeans maintain that means-end reasoning exhausts practical reason and that instrumental reason can be reduced to the use of theoretical reason in discovering means to ends. They tend not to be practical-reasoning theorists, however, since they argue that ethics fundamentally concerns what engages human sympathy or moral sentiment rather than what it is rational for a person to do. By contrast, practical-reasoning theorists deny that practical reason can be reduced to theoretical reason. As Christine Korsgaard has argued, even instrumental practical reason directs an agent who has already used theoretical reason in determining that *B* is the only means to his end *A* to undertake *B* (or to give up *A* as an end). In this way instrumental practical reasoning parallels the structure of *modus ponens* in theoretical reasoning (the move from "*p*" and "if *p*, then *q*" to "*q*").

Pursuing the analogy with theoretical reasoning (while insisting on irreducibility) further suggests that instrumental reasoning cannot exhaust practical reason. When we reason from our beliefs—for example, with *modus ponens*—we reason from their contents, not from the fact of our believing them. We reason from *p* and if *p*, then *q*, not from the facts that we believe that *p* and that we believe that if *p*, then *q*. Similarly, when we adopt an end, we do not simply select it by sheer fiat. Rather, we choose it as something (we think) there is some reason to do. Thus, when we reason from our ends, we do not reason from the fact that they are our ends but from our commitments to them as things it makes sense to do. That is why instrumental rationality is so uncontroversial. As R. M. Hare argued, it is questionable at best that it follows from the facts that a person's end is to kill someone in the most grisly possible way and that using a cleaver is such a way that the person ought, or has some normative reason, to use a cleaver. What is uncontroversial is simply that the support of reasons transfers from end to means, other things being equal, and from not taking the (only available) means to renouncing the end, other things being equal. It follows only that a person ought to use a cleaver or give up my end.

On grounds such as these, practical-reasoning theorists tend to hold that instrumental rationality cannot exhaust practical reason. But how are we to deliberate about ends? What makes something a reason for adopting an end? Since they hold that reasons for action are necessarily connected to the agent's deliberative perspective, practical-reasoning theorists generally adopt what Korsgaard has called the internalism requirement, according to which a reason must be something the agent could, in principle, be moved by in deliberation and act on. This makes it a necessary condition of something's being a reason for an agent that she would be moved by it insofar as she deliberated rationally.

But what then is rational deliberation? Practical-reason theorists are loath to derive a deliberative ideal by independently specifying paradigm reasons for acting and holding that deliberation is rational when it responds appropriately to them. That would theorize practical reason too much on the model of theoretical reason. Rather, they maintain that rational deliberation must be understood formally, so that reason for acting is a status consideration inherit when it is such that it would move an agent who formed her will in accordance with that deliberative ideal.

The aspects that have been considered so far are relatively common among practical-reasoning theories, although not, perhaps, universal. Within these theories, however, there is a major division between neo-Hobbesians and neo-Kantians. Although nothing on the surface of practical-reasoning theory might suggest this result, it is notable that both camps attempt to vindicate the commonsense idea that moral obligations are supremely authoritative. Both argue that (at least some central) moral demands are demands of practical reason.

NEO-HOBBESIANISM

Recent versions of this view have their roots in ideas advanced by Kurt Baier in the late 1950s and attempt to address a significant problem faced by Baier's early view. Baier argued that reasons for acting must ultimately connect with the agent's interests. This does not reduce all practical reasoning to prudential reasoning, since other forms may advance agents' interests also. Specifically, Baier argued that morality may be viewed as a system of practical reasoning that is in the interest of everyone alike. Since it is mutually advantageous for everyone to regard moral obligations as supremely authoritative, Baier concluded that they actually do create overriding reasons for acting.

David Gauthier objected to Baier's theory that, while it is in the interest of each that all regard interest-trumping moral reasons as supreme, it is unclear how this can show that an individual agent should so regard them, since it will still most advance her interest to act prudentially when morality conflicts with self-interest. Why, then, might it not be true that instrumental and prudential reasoning exhaust practical reason, even if a person should hope to live in a world in which other people view things differently and (mistakenly) treat moral reasons as authoritative?

Gauthier is himself responsible for the major recent neo-Hobbesian practical-reasoning theory. Like Baier, Gauthier begins from the premise that practical reasoning must work to advance the agent's interests, although here his account is more nearly "internalist," since he understands a person's interests to consist in what she would herself prefer were she to be fully informed. Also like Baier, Gauthier argues that the fact that mutual advantage may require individuals to constrain their pursuit of self-interest can be used to show that practical reason counsels this constraint. However, it is not enough that it be true that everyone would do better if everyone so constrained his or her prudential reasoning. The crucial point for Gauthier is that individuals can do better if they constrain self-interest by a willingness to abide by mutually advantageous agreements.

Two agents who appear to each other to be unconstrained pursuers of self-interest simply cannot make agreements, however mutually advantageous the agreements might be, if these agreements would require the agents to act contrary to their own interests. In what have come to be known as prisoner's dilemma situations, therefore, mutually advantageous rational agreement between such persons is impossible. If each believes the other will rationally defect from the agreement on the condition that doing so is in her interest, then neither can rationally make the agreement.

Personal advantage therefore counsels presenting oneself to others as someone who is not an unconstrained maximizer of self-interest. Of course, it is possible, theoretically, for someone to do this while still deliberating as an unconstrained prudential reasoner. But it may not be practically possible, Gauthier argues, at least not for normal human beings. Human motivation may be sufficiently translucent—through involuntary response, for example—so that the least costly way of appearing to others as someone who can be relied upon to keep mutually advantageous, interest-constraining agreements is actually to be such a person. If that is so, then instrumental and prudential reason will not support themselves as principles to guide rational deliberation. On the contrary, they will recommend that agents deliberate in terms of an alternative conception of practical reason that counsels keeping mutually advantageous agreements, even when this is contrary to self-interest.

As a practical-reasoning theorist, Gauthier believes that reasons for acting cannot be understood except in relation to what should guide a rational agent in deliberation. And he believes that a rational agent is someone whose dispositions of choice and deliberation serve her best and most advance her interest. But just as indirect forms of ethical consequentialism, such as character- and rule-consequentialism, face the objection that they are unstable and threaten to collapse into either act-consequentialism or deontology, Gauthier's indirect consequentialist theory of rationality may face the same objection. What motivates the move away from unconstrained prudence, on the grounds that it cannot support itself in the agent's practical thinking, is a view about the role a principle of rational conduct must be able to play in the deliberations of an autonomous rational agent that may be more Kantian than Hobbesian in inspiration.

NEO-KANTIANISM

This contemporary tradition may be held to date from Thomas Nagel's *The Possibility of Altruism* (1970) and John Rawls's reinvigoration of Kantian moral and political philosophy in *A Theory of Justice* (1971) and "Kantian Constructivism in Moral Theory" (1980). Nagel's book was read as having both a modest and a more ambitious agenda. His more modest goal, suggested by his title, was to show how such "objective" (or, as he later termed them, "agent-neutral") considerations as "that acting would be relative *someone's* pain" can be genuine reasons for acting. A consideration can be rationally motivating, he argued,

even if the agent lacks any relevant desire for acting on it other than one that is motivated by the awareness of that very consideration. A person may be moved, for instance, by considering long-term interests. And if motivation at a distance is possible with prudence, it can happen with altruism as well. Altruistic and other agent-neutral considerations can be rationally motivating.

Nagel's more ambitious agenda was to argue that practical reasoning is subject to a formal constraint that effectively requires that any genuine reason for acting be agent-neutral. Stressing the "motivational content" of genuine practical judgments, Nagel argued that avoiding a kind of solipsism is possible only if an agent is able to make the same practical judgment of himself from an impersonal standpoint as he does from an egocentric point of view. Since accepting practical judgments from one's own point of view normally motivates, Nagel maintained, making the same judgment of oneself from an impersonal standpoint should normally motivate also. But this will be so only if the reasons for acting that ground practical judgments are agent-neutral. So it is a necessary condition for avoiding practical solipsism that agents take considerations such as that something will advance their own ends or interests as reasons only if they regard them as instantiating more general, agent-neutral reasons, such as that acting will advance someone's ends or interests. Nagel later retreated from this strong claim in a direction that is arguably even more Kantian. Autonomous agency, he later argued, involves an agent's acting on reasons she can endorse from an objective standpoint, and such a set of reasons will include both agent-relative and agent-neutral ones.

Neo-Kantian practical-reasoning theories have been put forward by a number of philosophers, including Alan Gewirth, Stephen Darwall, and Christine Korsgaard. Korsgaard's sympathetic reconstruction of Immanuel Kant's own arguments in a series of papers has been especially influential. Common to all these neo-Kantian approaches has been the idea that the practical reasoning of an autonomous agent has a formal structure, with its own internal standards and constraints, and that these provide the fundamental truth and objectivity conditions for ethical thought and discourse. Thus, Gewirth maintains that fundamental moral principles are derivable from propositions to which a rational agent is committed from within the deliberative standpoint in acting. And Korsgaard argues that even instrumental theorists are committed to the "hypothetical imperative" as a practical norm. Since, however, we regard ourselves to be free as agents to adopt and renounce ends, practical reason cannot possibly be exhausted by any mere consistency constraint, such as the hypothetical imperative. It follows, the neo-Kantians argue, that practical reason requires norms to regulate the choice of ends no less than to guide the choice of means. In choosing ends for reasons we commit ourselves implicitly to principles of choice as valid for all. But such a commitment is not, they claim, a hypothesis about some independently existing order of normative fact to which we might have cognitive access. That, after all, is precisely the difference between theoretical and practical reason. So the standards to which deliberation is subject must ultimately be based on some formal principle of impartial endorsement that is internal to free practical reasoning itself. And this will be so, they conclude, only if practical reasoning is regulated by some such principle as the categorical imperative, which requires that one act only on principles that one can will to regulate the deliberation and choices of all. If moral demands are ultimately grounded in the categorical imperative also, it will follow that moral demands are demands of practical reason.

See also Baier, Kurt; Consequentialism; Decision Theory; Gewirth, Alan; Hare, Richard M.; Intuitionism; Kant, Immanuel; Metaethics; Nagel, Thomas; Naturalism; Noncognitivism; Practical Reason; Rationality; Rawls, John; Reason.

Bibliography

Baier, K. *The Moral Point of View.* Ithaca, NY: Cornell University Press, 1958.

Darwall, S. *Impartial Reason.* Ithaca, NY: Cornell University Press, 1983.

Falk, W. D. *Ought, Reasons, and Morality.* Ithaca, NY: Cornell University Press, 1986.

Gauthier, D. "Morality and Advantage." *Philosophical Review* 76 (1967): 460–475.

Gauthier, D. *Morals by Agreement.* Oxford: Clarendon Press, 1986.

Gewirth, A. *Reason and Morality.* Chicago: University of Chicago Press, 1978.

Korsgaard, C. "The Source of Normativity." In *The Tanner Lectures on Human Values,* edited by G. Peterson. Salt Lake City: University of Utah Press, 1994.

Nagel, T. *The Possibility of Altruism.* Oxford: Clarendon Press, 1970.

Nagel, T. *The View from Nowhere.* New York: Oxford University Press, 1986.

Rawls, J. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77 (1980): 515–572.

Rawls, J. *A Theory of Justice.* Cambridge, MA: Belknap Press of Harvard University Press, 1971.

Stephen Darwall (1996)

RATIONALISM IN ETHICS [ADDENDUM]

Moral rationalism, like many philosophical "isms," is an umbrella term for a variety of constituent claims. Not all moral rationalists endorse all of these claims, but the central ones that have been associated with moral rationalism are:

1) The metaphysical thesis: Basic moral requirements are constituted by the deliverances of sound practical reason.

2) The epistemological thesis: Humankind's basic moral requirements are knowable a priori.

3) The normative thesis: Moral requirements entail excellent reasons for action.

The metaphysical thesis is at the heart of practical reason approaches to ethics, such as those of Thomas Hobbes and Immanuel Kant, and, in modern times, David Gauthier, Christine Korsgaard, and Michael Smith. Though the views of these thinkers differ in many important respects, they agree that our basic moral duties are a function of sound practical reason. This means that something is a moral duty for a person just because it would be regarded as rationally authoritative were that person (or some idealized counterpart thereof) to reason soundly from his or her most important commitments.

What distinguishes adherents of the metaphysical thesis from one another are their views about the nature of our fundamental commitments, and about what sound practical reasoning looks like. Kantians, such as Korsgaard, will consider some of our fundamental commitments as intrinsic to human nature, and so shared by all human moral agents. Successful reasoning on this shared basis will yield a set of universal moral duties. Hobbesians, such as Gauthier, deny that humans have any essential or intrinsic ends. Yet Gauthier believes that sound reasoning—which, for him (unlike Kantians), is restricted to instrumental reasoning designed to maximize self-interest—will also yield a set of common moral duties. This is because each of us is likeliest to do the best by adhering to (and acquiring a disposition to conscientiously adhere to) mutually beneficial rules that sometimes mandate self-sacrifice, and these are just what Gauthier thinks moral rules are.

Smith's view is a kind of ideal observer theory. What we have reason to do is a matter of what an ideal adviser would want us to do. This ideal advisor shares our fundamental desires, but is fully factually informed, has flawless reasoning abilities, and is possessed of a fully coherent set of desires. Though such advisers obviously do not really exist, they play a crucial role in determining the content of our moral duties. Smith endorses the Kantian tenet that sound reasoning is more than self-interested instrumental reasoning. He endorses the Humean view that all of our fundamental commitments are contingent. But he believes that sound reasoning on the basis of disparate commitments will nevertheless yield a universal set of rationally authoritative moral requirements: The ideal advisers would converge in all of their basic recommendations.

What all of these thinkers share is a belief that the content of our basic moral duties conceptually depends very importantly on the outcome of sound practical reasoning. Perhaps the nature of this dependence can best be seen by contrasting their metaphysical thesis with that of moral realism. Moral realists, some of whom consider themselves rationalists because of their endorsement of the epistemological and normative theses, nevertheless deny the metaphysical thesis. Whichever picture of the ideal reasoner and the ideal reasoning process we adopt, the metaphysical thesis tells us that something is our moral duty just because a process of sound reasoning would ratify it. Moral realists, by contrast, might allow for the possibility of inerrant reasoners who infallibly identify all moral duties, but realists will insist that the content of our duties is fixed in a way that does not depend on the outcomes of such reasoning. Ideal reasoners will discover the moral truths there are to be discovered, rather than creating, through the exercise of sound practical reasoning, the content of our basic moral duties.

Though moral realists and those who endorse practical reason approaches to ethics are divided in their opinion of the metaphysical thesis, those among them who accept the rationalist moniker may still agree on the epistemological thesis. Moral rationalists in this domain believe that at least some moral knowledge is a priori. While experience may be needed to obtain an understanding of moral concepts, and empirical premises might be needed to determine what our all-things-considered duty in a given case might be, knowledge of fundamental moral principles can be had without relying on contingent, empirical premises for evidential support. Rationalists here believe, for instance, that one needs nothing other than a sound understanding of the proposition to be justified in thinking that, prima facie, it is immoral to rape or torture people solely for personal enjoyment.

Both moral realists and adherents of practical reason approaches to ethics are divided on the availability of a

priori moral knowledge. Among moral realists, the sides are drawn by reference to whether the emerging view takes the form of ethical naturalism or ethical nonnaturalism. Naturalists will see all knowledge, and so all moral knowledge, as a posteriori, available only by applying the same methods of discovery and justification as are utilized in the natural sciences. Ethical nonnaturalists, by contrast, will reject the attempted assimilation of ethics to natural science, and so will leave room for the possibility of a priori moral knowledge. Indeed, the historically most prominent moral realists, up to and including the British intuitionists of the early twentieth century (Henry Sidgwick, G. E. Moore, and W. D. Ross), all endorsed the idea that fundamental moral principles were self-evident: knowable solely on the basis of adequately understanding the content of the relevant principle. Robert Audi, the contemporary philosopher who has done most to develop this view, agrees with his historical forbears in attributing to reason the power to discern certain fundamental truths about reality—in this case, moral reality.

Like moral realists, adherents of practical reason approaches to ethics (and so of the rationalists' metaphysical thesis) disagree among themselves about the tenability of a priori moral knowledge. Kantians, who have championed the possibility of synthetic a priori knowledge in other areas, likewise endorse its possibility in ethics. But others, such as Gauthier, are avowed naturalists who believe that knowledge of an agent's moral duties depends crucially on knowledge of her contingent, fundamental commitments, and so cannot be a priori.

The normative thesis, which states that moral requirements entail excellent reasons for action, is accepted by all practical reason approaches to ethics. Indeed, a major attraction of such theories is that they are able to explain the intimate connection that might obtain between one's moral duty and one's reasons for action. We will always have reason to do as morality says, because morality is constructed from our own deepest commitments. If moral duties are rational extensions of our fundamental commitments, then, to the extent that we are rational, we cannot be alienated from what morality requires of us.

Moral realists have had a more difficult time explaining the reason-giving power of morality. Indeed, some, such as Peter Railton and David Brink, have given up on the idea that there is any necessary connection between moral demands and reasons for action. Such thinkers believe that something qualifies as a practical reason only if it furthers an agent's ends (all of which, in their view, are contingent). They believe that moral demands may fail to serve an agent's ends. It follows that there may be no good reason to abide by the demands of morality.

If one retains this popular view of reasons, then the moral rationalist's normative thesis can be sustained in only one of two ways. The first involves rejecting moral realism, and taking up a practical reason approach that identifies our moral duties with the rational extensions of our commitments. Here, both our reasons and our moral duties will rely on these commitments, and this can explain the perfect alignment of morality and reasons for action. The second, more controversial way of defending both rational egoism and the normativity thesis, is to accept moral realism, and then to insist that fulfillment of our moral duties will always, of necessity, further our ends. Plato argued this way, and some theists who are moral realists do so as well. This is the less traveled path, however, because its defense requires that we posit a set of objective human ends that are invariably furthered by moral conduct. Such ends might be renounced by apparently coherent and rational individuals, and this has led to a great deal of suspicion about their existence.

The more common strategy for moral realists who want to vindicate the normativity thesis is to reject the view that reasons for action must always further the agent's ends. A misanthrope, for instance, might be morally bound to rescue another, if he can do so at little or no inconvenience to himself. According to the normativity thesis, there is therefore excellent reason for him to do so, even if none of the misanthrope's ends are furthered as a result. Given who the person is, he will reject the existence of such a reason. It will seem to the misanthrope an alien demand, of only spurious rational authority. It will play no role in explaining the actions he undertakes. Still, says the realist who embraces the normativity thesis, the misanthrope's reactions in such a case do not immunize him either from the moral duty, or from the practical reason that he ignores in his neglect of the victim he might have aided.

An adherent of the practical reason approach to ethics will insist that reasons be able to engage agents who are reasoning well from their fundamental commitments. And realists cannot secure this guarantee, since they do not make the content of moral demands dependent on the outcomes of sound practical reasoning. Realists will either concede the point, as Railton and Brink do, or affirm their allegiance to the normative thesis. In the latter case, they will deny that morality must engage all who are able to reason efficiently about securing their ends. The plausibility of such a denial is the subject of much contemporary metaethical discussion.

See also Categorical Imperatives; Error Theory of Ethics; Ethical Subjectivism; Ideal Observer Theories of Ethics; Internalism and Externalism in Ethics; Intuitionism, Ethical; Kantian Ethics; Moral Realism; Practical Reason; Response-Dependence Theories.

Bibliography

Audi, Robert. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton, NJ: Princeton University Press, 2004.

Brink, David. *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press, 1989.

Gauthier, David. *Morals by Agreement*. New York: Oxford University Press, 1986.

Korsgaard, Christine. *The Sources of Normativity*. New York: Cambridge University Press, 1996.

Railton, Peter. *Facts, Values, and Norms: Essays toward a Morality of Consequence*. New York: Cambridge University Press, 2003.

Smith, Michael. *The Moral Problem*. New York: Oxford University Press, 1994.

Russ Shafer-Landau (2005)

RATIONALITY

Philosophers have, at least characteristically, aspired to possess "rationality" but have not thereby sought exactly the same thing. Portrayed vaguely, rationality is reasonableness, but not all philosophers take rationality as dependent on reasons; nor do all philosophers have a common understanding of reasons or of reasonableness. Some theorists consider rationality to obtain in cases that lack countervailing reasons against what has rationality; they thus countenance rationality as, in effect, a default status. In ordinary parlance, persons can have rationality; so, too, can beliefs, desires, intentions, and actions, among other things. The rationality appropriate to action is practical, whereas that characteristic of beliefs is, in the language of some philosophers, theoretical.

Many philosophers deem rationality as instrumental, as goal oriented. You have rationality, according to some of these philosophers, in virtue of doing your best, or at least doing what you appropriately think adequate, to achieve your goals. If ultimate goals are not themselves subject to assessments of rationality, then rationality is purely instrumental, in a manner associated with David Hume's position. Rationality, according to this view, is a minister without portfolio; it does not require any particular substantive goals of its own but consists rather in the proper pursuit of one's ultimate goals, whatever those goals happen to be. Many decision-theoretic and economic approaches to rationality are purely instrumentalist. If, however, ultimate goals are susceptible to rational assessment, as an Aristotelian tradition and a Kantian tradition maintain, then rationality is not purely instrumental. The latter two traditions regard certain rather specific (kinds of) goals, such as human well-being, as essential to rationality. Their substantialist approach to rationality lost considerable influence, however, with the rise of modern decision theory.

When relevant goals concern the acquisition of truth and the avoidance of falsehood, so-called epistemic rationality is at issue. Otherwise, some species of nonepistemic rationality is under consideration. One might individuate species of nonepistemic rationality by the kind of goal at hand; moral, prudential, political, economic, aesthetic, or some other. Some philosophers have invoked rationality "all things considered" to resolve conflicts arising from competing desires or species of rationality; even so, there are various approaches to rationality "all things considered" in circulation. The standards of rationality are not uniformly epistemic, then, but epistemic rationality can play a role even in what some call nonepistemic rationality. Regarding economic rationality, for instance, a person seeking such rationality will, at least under ordinary conditions, aspire to epistemically rational beliefs concerning what will achieve the relevant economic goals. Similar points apply to other species of nonepistemic rationality. A comprehensive account of rationality will characterize epistemic and nonepistemic rationality, as well as corresponding kinds of irrationality (e.g., weakness of will).

Taking rationality as deontological, some philosophers characterize rationality in terms of what is rationally obligatory and what is merely rationally permissible. If an action, for instance, is rationally obligatory, then one's failing to perform it will be irrational. Other philosophers opt for a nondeontological evaluative conception of rationality that concerns what is good (but not necessarily obligatory) from a certain evaluative standpoint. Some of the latter philosophers worry that, if beliefs and intentions are not voluntary, then they cannot be obligatory. Still other philosophers understand rationality in terms of what is praiseworthy, rather than blameworthy, from a certain evaluative standpoint. The familiar distinction between obligation, goodness, and praiseworthiness thus underlies three very general approaches to rationality.

Following Henry Sidgwick, William Frankena has distinguished four conceptions of rationality: (1) an egoistic conception implying that it is rational for one to be

or do something if and only if this is conducive to one's own greatest happiness (e.g., one's own greatest pleasure or desire satisfaction); (2) a perfectionist conception entailing that it is rational for one to be or do something if and only if this is a means to or a part of one's moral or nonmoral perfection; (3) a utilitarian conception implying that it is rational for one to be or do something if and only if this is conducive to the greatest general good or welfare; and (4) an intuitionist conception implying that it is rational for one to be or do something if and only if this conforms to self-evident truths, intuited by reason, concerning what is appropriate. The history of philosophy represents, not only these conceptions of rationality, but also modified conceptions adding further necessary or sufficient conditions to one of (1)–(4).

Given an egoistic conception of rationality, one's being rational will allow for one's being immoral, if morality requires that one not give primacy to oneself over other people. Rationality and morality can then conflict. Such conflict is less obvious on a utilitarian conception of rationality. In fact, if morality is itself utilitarian in the way specified (as many philosophers hold), a utilitarian conception of rationality will disallow rational immorality. A perfectionist conception of rationality will preclude rational immorality only if the relevant perfection must be moral rather than nonmoral; achieving nonmoral perfection will, of course, not guarantee morality. As for an intuitionist conception of rationality, if the relevant self-evident truths do not concern what is morally appropriate, then rational immorality will be possible. An intuitionist conception will bar conflict between rationality and morality only if it requires conformity to all the self-evident truths about what is morally appropriate that are relevant to a situation or person. So, whether rationality and morality can conflict will depend, naturally enough, on the exact requirements of the conception of rationality at issue.

Richard Brandt has suggested that talk of what it would be rational to do functions to guide action by both recommending action and by making a normative claim that evaluates the available action relative to a standard. An important issue concerns what kind of strategy of using information to choose actions will enable one to achieve relevant goals as effectively as any other available strategy. Brandt has offered a distinctive constraint on such a strategy: A rational decision maker's preferences must be able to survive their being subjected to repeated vivid reflection on all relevant facts, including facts of logic. This constraint suggests what may be called (5) a relevant-information conception of rationality: Rationality is a matter of what would survive scrutiny by all relevant information.

A relevant-information conception of rationality depends, first, on a clear account of precisely when information is relevant and, second, on an account of why obviously irrational desires cannot survive scrutiny by all relevant information. Evidently, one could have a desire caused by obviously false beliefs arising just from wishful thinking, and this desire could survive a process of scrutiny by all relevant information where the underlying false beliefs are corrected. In any case, a relevant-information conception of rationality will preclude rational immorality only if it demands conformity to all relevant moral information.

The egoistic, perfectionist, utilitarian, and relevant-information conceptions of rationality are nonevidential in that they do not require one's having evidence that something is conducive to self-satisfaction, perfection, general welfare, or support from all relevant information. Many philosophers would thus fault those conceptions as insufficiently sensitive to the role of relevant evidence in rationality. If relevant evidence concerns epistemic rationality, we again see the apparent bearing of epistemic rationality on rationality in general. The latter bearing deserves more attention in contemporary work on nonepistemic rationality.

Philosophers currently divide over internalism and externalism about rationality. If rationality demands reasons of some sort or other, the dispute concerns two senses of talk of a person's having a reason to perform an action. An internalist construal of this talk implies that the person has some motive that will be advanced by the action. An externalist construal, in contrast, does not require that the person have a motive to be advanced by the action. Bernard Williams, among others, has suggested that any genuine reason for one's action must contribute to an explanation of one's action and that such a contribution to explanation must be a motivation for the action. He concludes that externalism about rationality is false, on the ground that external reasons do not contribute to explanation of action in the required manner. Externalism about rationality does allow that reasons fail to motivate, but this, according to externalists, is no defect whatever. Externalists distinguish between merely motivating reasons and justifying reasons, contending that only the latter are appropriate to rationality understood normatively; what is merely motivating in one's psychological set, in any case, need not be justifying. Perhaps, then, disputes between internalists and externalists

will benefit from attention to the distinction between justifying and merely motivating reasons.

Modern decision theory assumes that, in satisfying certain consistency and completeness requirements, a person's preferences toward the possible outcomes of available actions will determine, at least in part, what actions are rational for that person by determining the personal utility of outcomes of those actions. In rational decision making under certainty one definitely knows the outcomes of available actions. In decision making under risk one can assign only various definite probabilities less than 1 to the outcomes of available actions. (Bayesians assume that the relevant probabilities are subjective in that they are determined by a decision maker's beliefs.) In decision making under uncertainty one lacks information about relevant states of the world and hence cannot assign even definite probabilities to the outcomes of available actions. Acknowledging that rationality is purely instrumental (and thus that even Adolf Hitler's Nazi objectives are not necessarily rationally flawed), Herbert Simon has faulted modern decision theory on the ground that humans rarely have available the facts, consistent preferences, and reasoning power required by standard decision theory. He contends that human rationality is "bounded" in that it does not require utility maximization or even consistency. Rather, it requires the application of a certain range of personal values (or preferences) to resolve fairly specific problems one faces, in a way that is satisfactory, rather than optimal, for one. Simon thus relies on actual human limitations to constrain his account of rationality.

Contemporary theorists divide over the significance of human psychological limitations for an account of rationality. The controversy turns on how idealized principles for rationality should be. This raises the important issue of what exactly makes some principles of rationality true and others false. If principles of rationality are not just stipulative definitions, this issue merits more attention from philosophers than it has received. Neglect of this metaphilosophical issue leaves the theory of rationality as a subject of ongoing philosophical controversy.

See also Aristotle; Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Decision Theory; Hume, David; Kant, Immanuel; Sidgwick, Henry; Utilitarianism.

Bibliography

Audi, Robert. *Action, Intention and Reason*. Ithaca, NY: Cornell University Press, 1993.

Audi, Robert. *The Architecture of Reason: The Structure and Substance of Rationality*. New York: Oxford University Press, 2001.

Benn, S. I., and G. W. Mortimore, eds. *Rationality and the Social Sciences*. London: Routledge and Paul, 1976.

Brandt, R. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Brown, Harold. *Rationality*. London: Routledge, 1988.

Elster, J., ed. *Rational Choice*. New York: New York University Press, 1986.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Frankena, W. "Concepts of Rational Action in the History of Ethics." *Social Theory and Practice* 9 (1983): 165–197.

Fumerton, Richard. "Audi on Rationality: Background Beliefs, Arational Enjoyment, and the Rationality of Altruism." *Philosophy and Phenomenological Research* 67 (2003): 188–193.

Gauthier, D. *Morals by Agreement*. Oxford: Clarendon Press, 1986.

Harman, Gilbert. *Change in View: Principle of Reasoning*. Cambridge, MA: MIT Press, 1986.

Hollis, M., and S. Lukes, eds. *Rationality and Relativism*. Cambridge, MA: MIT Press, 1982.

Jeffrey, Richard. *The Logic of Decision.* 2nd ed. Chicago: University of Chicago Press, 1983.

Katz, Jerrold. *Realistic Rationalism*. Boston: MIT Press, 1998.

Mele, A. *Irrationality*. New York: Oxford University Press, 1987.

Mele, A., and P. Rawlings, ed. *The Oxford Handbook of Rationality*. Oxford: Oxford University Press, 2004.

Moser, P., ed. *Rationality in Action*. Cambridge, U.K.: Cambridge University Press, 1990.

Nozick, R. *The Nature of Rationality*. Princeton, NJ: Princeton University Press, 1993.

Rescher, N. *Rationality*. Oxford: Clarendon Press, 1988.

Sidgwick, H. *The Methods of Ethics.* 7th ed. London, 1907.

Simon, Herbert. *Reason in Human Affairs*. Stanford, CA: Stanford University Press, 1981.

Skyrms, Brian. *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard University Press, 1990.

Slote, M. *Beyond Optimizing: A Study of Rational Choice*. Cambridge, MA: Harvard University Press, 1989.

Williams, B. "Internal and External Reasons." In *Rational Action*, edited by R. Harrison. Cambridge, U.K.: Cambridge University Press, 1979.

Paul K. Moser (1996)
Bibliography updated by Benjamin Fiedor (2005)

RAVAISSON-MOLLIEN, JEAN GASPARD FÉLIX
(1813–1900)

The French spiritualist philosopher and art historian Jean Gaspard Félix Ravaisson-Mollien was born in Namur, Belgium. He received his philosophical training in Munich under Friedrich von Schelling and took a degree

in Paris in 1838 under Victor Cousin. His philosophical work began with his prize essay, *Essai sur la métaphysique d'Aristote,* and a short teaching career at Rennes in 1838. In 1840 he was appointed inspector general of libraries, a post that he held until 1860, when he became inspector general in the department of higher education. Meanwhile, as a semiprofessional painter he had become interested in classical antiquities, and in 1870 he was made curator in the department of antiquities in the Louvre. The fruit of this was his well-known set of reconstructions of the Venus de Milo.

The most influential of Ravaisson's publications was his *Rapport sur la philosophie en France au XIX^{e} siècle,* made at the request of the imperial government in 1867. At this time the school of Cousin was in the ascendancy in France, and it was difficult, indeed practically impossible, for a man who was not an eclectic to get an appointment in the university system. Ravaisson's purpose in his report was to show that there was a continuity in the French philosophical tradition and that French philosophers had always presupposed metaphysical principles that implied what he called spiritualism. This tradition, he maintained, always swung between sensationalism, phenomenalism, and materialism, on the one hand, and idealism, on the other. But spiritualism really began in the nineteenth century with Maine de Biran, who used as his starting point the human will and who held that the will is independent both of sensations and of ideas. This viewpoint, Ravaisson argued, was not only the proper beginning of a philosophy but also the only one that could unify the opposing tendencies of empiricism and idealism.

Such a conclusion was in clear contradiction to the tenets of Cousin's eclecticism, which aspired to fuse "the best in each philosopher." Ravaisson tried to show that such a fusion in reality consists in refuting those philosophies which displease the eclectic and retaining those which please him. In classifying all philosophies under the headings of sensualism, idealism, mysticism, and skepticism, Cousin accepted only that philosophy which he called idealism but which, said Ravaisson, was really a simple mixture of the Scottish philosophy of common sense with a few ideas from Maine de Biran. The eclectics, moreover, failed to understand these ideas. Ravaisson claimed for himself the credit of introducing the true thought of Maine de Biran to his contemporaries. Readers of this report were thus informed that the de facto official philosophy of the French universities was not only a foreign importation but also untrue.

Ravaisson was not satisfied with undermining eclecticism. He also felt it important to point out the weaknesses of positivism. These weaknesses, he claimed, arose from the identification of philosophical method with the methods of science. Science, which admittedly studies the external world, can never tell us anything about the internal world of thoughts, aspirations, desires, and dreams; and when it attempts to do so, it transforms them into quasi-external objects. This inevitably leads to materialism, for the laws of matter are the only laws that science can formulate. Science's basic categories are space and quantity, and its basic method is analysis. But the phenomena of consciousness are never spatial or quantitative, and to attempt to categorize them in these terms is to change their essential nature.

Ravaisson's report reviewed all the contemporary schools of thought and all the contemporary philosophers. It was a model of patience and thorough investigation and has become the primary source of information about individuals who are obscure and in some cases forgotten. It did not stop at professional philosophers but looked into the presuppositions of scientists, such as the physiologist Claude Bernard and the psychiatrist Albert Lemoine. In every case, Ravaisson found either too strong an emphasis on the dependence of the "spirit" upon material causes or an identification of ideas with strictly logical, hence analytical, reason. Whereas one set of philosophers tried to explain the mind in terms that were inappropriate, the other failed to ask the central question of why the mind operated as it did. Neither group could explain our undeniable feeling of being active causes; neither could see why the spirit needs both analysis and synthesis.

Whether the object of our thinking is the external or the internal world, it will be found that we have to use two absolutely general metaphysical principles, that of an infinite reality and that of limitation. The dialectical reason for this is that every analytical sentence distinguishes between parts of a whole and no whole can be discussed except by reference to its constituent parts. But Ravaisson did not rest his doctrine on this dialectical argument. On the contrary, he believed that history had shown that every philosopher presupposes these principles, whether he knows it or not. The tendency of the history of philosophy is toward the progressive realization of this truth. It is implicit in all philosophy and is steadily becoming explicit. Ravaisson's report thus presented not only an exposition of contemporary French philosophies but also a theory about the history of philosophy.

In a shorter study, *De l'habitude,* written as a thesis at the Sorbonne in 1838, Ravaisson returned to the problem raised in Maine de Biran's prize-winning essay on the influence of habit on thinking. Ravaisson's study is of special historical interest since it forms the nucleus of the philosophy of Henri Bergson.

At the beginning of his argument Ravaisson laid down a fundamental distinction between the roles played by space and time in our lives. "Space," he said, "is the most obvious and elementary condition and form of stability or permanence; time the universal condition of change." Corresponding to these two basic principles are matter and life respectively. In matter there is no individuality and no possibility of habit, a point that Ravaisson probably encountered in his study of Aristotle. Life, on the contrary, forms a world of its own, a world that is internal to the living being. A set of oppositions follows, that of necessity (matter) versus that of "nature" (life), a set that echoes the two realms of necessity and freedom elaborated by Schelling. The repetition of a change modifies "nature," and the living being swings between the limitations of its material conditions and its own inner freedom. As the forms of life develop, their power of spontaneous action becomes greater, so that although the inorganic is timeless, life implies a "definite continuous *durée.*" As we move up from vegetable to animal to human life, we find that whereas sensory impressions become weaker when repeated, our powers of movement become stronger and stronger.

Corresponding to these dualities is another. Within the human soul are the two powers of understanding and of activity. The understanding sees everything under the aspects of diversity, quantity, and space; the power of activity appears primarily in our feeling of effort, which is gradually reduced by habit. Habit transforms voluntary movements into instinctive movements. Voluntary movements could not be made if there were no resistance from without, but for them to be made at all requires that somewhere there be an undetermined center of activity, which is the will. And when one asks what the will is seeking, the answer is that it seeks the good, or God. It is not difficult to see in these views both the influence of Schelling and the anticipation of Bergson.

See also Aristotle; Bergson, Henri; Bernard, Claude; Cousin, Victor; Empiricism; Idealism; Maine de Biran; Materialism; Mysticism, Nature and Assessment of; Phenomenalism; Positivism; Schelling, Friedrich Wilhelm Joseph von; Sensationalism; Skepticism; Space; Time.

Bibliography

WORKS BY RAVAISSON

Essai sur la métaphysique d'Aristote, 2 vols. Paris, 1837–1846.

De l'habitude. Paris: H. Fournier, 1838. Reprinted in *Revue de métaphysique et de morale* 2 (1894): 1–35.

Rapport sur la philosophie en France au XIX^e siècle. Paris, 1867.

Testament philosophique et fragments. Paris, n.d. Published posthumously, with bibliography.

WORKS ON RAVAISSON

Bergson, Henri. "Notice sur la vie et les oeuvres de M. Félix Ravaisson-Mollien." *Comptes-rendus de l'Académie des sciences morales et politiques* (1904). Reprinted in Bergson's *La pensée et le mouvant* (Paris: Alcan, 1934) as well as in the *Testament philosophique* of Ravaisson.

Dopp, Joseph. *Félix Ravaisson, la formation de sa pensée d'après des documents inédits.* Louvain: Editions de l'Institut Superieur de Philosophie, 1933.

Gunn, J. Alexander. *Modern French Philosophy,* 73–75 and passim. London: Unwin, 1922.

George Boas (1967)

RAWLS, JOHN
(1921–2002)

John Rawls is widely regarded as one of the most significant political philosophers of the twentieth century. Educated at Princeton University, he taught at Cornell University and the Massachusetts Institute of Technology before joining the faculty of Harvard University in 1962. Rawls's *A Theory of Justice* (1971) revitalized political theory as an academic discipline and rejuvenated interest in the substantive social issues that had long been neglected by academic philosophers. Rawls continued to refine and defend his theory in a series of articles and lectures, the most important of which he revised and collected in his 1993 work *Political Liberalism*. In 1999 *The Law of Peoples* extended his theory to questions of international relations, and in the next two years, despite declining health, he published *Lectures of the History of Moral Philosophy* (2000) and *Justice as Fairness: A Restatement* (2001).

JUSTICE AS FAIRNESS

The primary objective of Rawls's political theory is to articulate and defend a conception of justice for a modern democratic regime. The theory begins with the idea of society as a fair system of cooperation between free and equal persons. The principles of justice for such a society characterize its fair terms of cooperation by specifying its citizens' basic rights and duties and by regulating the distribution of its economic benefits. To formulate his particular conception of justice, Rawls

invokes the familiar theory of the social contract, according to which the legitimate rules for a society are arrived at by the autonomous agreement of its members. Rawls's version of the contract theory is distinctive, however, in its insistence on the essential fairness of the point of view from which the agreement itself is conceived. This enables Rawls to appeal to the justificatory force of *pure procedural justice*, the idea that the fundamental fairness of a procedure can ensure the justice of its outcome provided that there is no independent criterion for the justice of that result. Fairness thus characterizes both the terms of the contractual agreement and the conditions in which that agreement is made. Rawls appropriately names the resulting theory *justice as fairness*.

Rawls's contractarian or *constructivist* theory represents this fundamental ideal of fairness by situating the contracting parties in a hypothetical *original position*. The most important feature of this theoretical model is the *veil of ignorance*, which denies to the parties any knowledge of their actual natural endowments, their social position, or even their conception of what makes for a good life. As a consequence, the parties cannot determine how proposed principles would affect their interests personally. The veil of ignorance thereby reflects our conviction that it would be patently unreasonable to allow principles that favored any individuals or groups merely in virtue of their possession of morally arbitrary attributes such as their race or sex, or because they happened to affirm a particular religious or philosophical doctrine.

THE PRINCIPLES OF JUSTICE

Though deprived of knowledge of their particular ends, attachments, and aspirations, the parties in the original position are still rationally motivated to further their conception of the good, whatever it is. They also have *higher-order interests* in developing and exercising the two *moral powers* that they share as free and equal beings: (1) the capacity to understand and act from a sense of justice and (2) the capacity to form, revise, and rationally pursue a conception of the good. The parties will therefore seek for themselves the best possible package of *primary goods*, those all-purpose, socially regulable opportunities and resources needed to advance those interests. Rawls's enumeration of these primary goods includes basic rights and liberties, the powers and prerogatives of offices and positions, income and wealth, and the social bases of self-respect. Assuming that a society has reached a minimal level of economic development, Rawls argues that the following two principles for allocating the primary goods would be selected:

> a. Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all.
>
> b. Social and economic inequalities are to satisfy two conditions. First, they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greatest benefit of the least advantaged members of society. (Rawls 1993, p. 291)

Since the first principle is given absolute priority over the second, Rawls argues that the basic liberties guaranteed by it, such as freedom of religion or the right to run for political office, cannot be sacrificed for any amount of personal or collective economic benefit. Such liberties can be limited only to protect the central range of application of other conflicting liberties, as when the right to a fair trial necessitates some restrictions on the freedom of the press. Specific rights are included in the protection of the first principle if agents in the original position would rationally require them. For example, freedom of religion would be insisted on by the parties, for they could not risk the possibility that their religion, should they have one, would be a minority faith subject to repression by a dogmatic majority.

The second principle deals with economic and social primary goods such as income and wealth. Its second condition, the so-called difference principle, stipulates that any departures from equality of resources can be justified only if the resulting inequality benefits the least advantaged members of society. Thus, positions that require the development of talents and the expenditure of extraordinary effort might deserve greater economic rewards, but only if the increased productivity generated by such a differential would improve the condition of the least well off. Rawls argues that this requirement would be the reasonable and rational choice of individuals who, because of the fairness conditions imposed in the original position, did not know their natural and social endowments and therefore could not determine their actual position in the social order.

The first part (of the second principle) stipulates that even the limited inequalities that would satisfy the difference principle are permissible only if the positions that give rise to them are open to all under conditions of *fair equality of opportunity*. This strong requirement goes beyond mere prohibition of discrimination based on arbitrary features such as gender or race. It demands that all individuals of like natural ability and similar motivation should have the same opportunities throughout their entire lives, a requirement that obviously necessi-

tates equal access to education, health care, and other social resources.

STABILITY

A viable political theory, Rawls insists, must be practical. The well-ordered society that it mandates must be feasible and stable given realistic economic, cultural, and psychological assumptions. In *A Theory of Justice* Rawls argues that a society regulated by justice as fairness would be stable since the laws of moral psychology show that its members would tend to acquire and maintain a common comprehensive moral doctrine that would sustain it. In *Political Liberalism*, however, he admits that a liberal, nonauthoritarian regime would be characterized by a plurality of reasonable, though incompatible comprehensive religious and moral doctrines. Nonetheless, he believes that the requirement of stability can be met by justice as fairness if it is understood as a political theory. As such, it regulates only the *basic structure* of society: the background institutions that specify political and civil rights and that determine entitlements to other socially regulated goods. Members of a well-ordered society may therefore hold deeply conflicting comprehensive religious and moral views, yet still endorse a common political conception of justice as the focus of an *overlapping consensus*. Moreover, Rawls stresses that this consensus can be more than a mere modus vivendi, a practical compromise based on a tenuous balance of power. Rather, it can express a genuine moral commitment that reflects ideas and values implicit in the society's political culture, such as its conception of the citizen as a free and equal person and its willingness to rely on reasonable standards of *public reason* in the conduct of its political affairs.

THE LAW OF PEOPLES

Rawls applies his principles of justice to individual, self-contained societies and not to humanity at large. He does not think, for example, that the difference principle should be applied globally. Rather, he argues for a *law of peoples*, a more limited set of obligations on just societies, first, to obey some traditional canons of international law (such as to wage war only in self-defense and to honor basic human rights) and, second, to aid peoples that lack sufficient resources to support just social institutions. These duties, he argues, would be agreed to in a second original position, populated now by representatives of just or "decent" peoples who are behind a veil of ignorance with respect to the particular societies that they represent.

REFLECTIVE EQUILIBRIUM

Rawls's methodology has been as influential and as controversial as his substantive views. Declining to ground his views on any deep metaphysical or other philosophical truths, Rawls maintains that political theory should formulate a coherent set of principles that accounts for the considered convictions that we actually hold. The process goes beyond mere summarization of particular considered judgments, however, for it also postulates theoretical models, mediating ideas, and principles at all levels of generality. All judgments and principles are held open to revision in light of other aspects of the theory, until no further changes are needed to develop a compelling and coherent view. The resulting theory is then said to be in *reflective equilibrium*. It is also *objective*, Rawls contends, because it would gain the assent of all reasonable individuals on due reflection.

See also Justice; Good, The; Liberty; Moral Psychology; Political Philosophy, History of; Political Philosophy, Nature of; Rights; Social and Political Philosophy; Social Contract.

Bibliography

WORKS BY RAWLS

A Theory of Justice. Cambridge, MA: Belknap Press of Harvard University Press, 1971. A somewhat revised 1975 version has been translated into several languages.

Political Liberalism. New York: Columbia University Press, 1993. Contains revised versions of several important papers published between 1974 and 1993. The revised paperback edition of 1996 has an added preface.

The Law of Peoples. Cambridge, MA: Harvard University Press, 1999.

Collected Papers, edited by Samuel Freeman. Cambridge, MA: Harvard University Press, 1999.

Lectures on the History of Moral Philosophy. Edited by Barbara Herman. Cambridge, MA: Harvard University Press, 2000.

Justice as Fairness: A Restatement, edited by Erin Kelly. Cambridge, MA: Harvard University Press, 2001.

WORKS ABOUT RAWLS

Arneson, Richard J., et al. "Symposium on Rawlsian Theory of Justice: Recent Developments." *Ethics* 99 (4) (1989): 695–710.

Barry, Brian. "The Liberal Theory of Justice: A Critical Examination of the Principal Doctrines." In *A Theory of Justice*, by John Rawls. Oxford, U.K.: Clarendon Press, 1973.

Blocker, H. Gene, and Elizabeth Smith, eds. *John Rawls' Theory of Social Justice: An Introduction*. Athens: Ohio University Press, 1980.

Daniels, Norman, ed. *Reading Rawls: Critical Studies on Rawls' Theory of Justice*. New York: Basic Book, 1975.

Freeman, Samuel, ed. *The Cambridge Companion to Rawls*. New York: Cambridge University Press, 2003. An

outstanding series of original articles and an exhaustive bibliography.
Kukathas, Chandran, and Philip Pettit. *Rawls: A Theory of Justice and Its Critics*. Cambridge, U.K.: Polity University Press, 1990.
Martin, Rex. *Rawls and Rights*. Lawrence: University Press of Kansas, 1985.
Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
Pogge, Thomas W. *Realizing Rawls*. Ithaca, NY: Cornel University Press, 1989.
Richardson, Henry, and Paul Weithman, eds. *The Philosophy of Rawls: A Collection of Essays*. 5 vols. New York: Garland, 1999.
Sandel, Michael J. *Liberalism and the Limits of Justice*. New York: Cambridge University Press, 1982.
Wellbank, J. H., Denis Snook, and David T. Mason. *John Rawls and His Critics: An Annotated Bibliography*. New York: Garland, 1982.

Alan Fuchs (1996, 2005)

REALE, MIGUEL

(1910–)

Miguel Reale, the Brazilian philosopher of law, historian of ideas, and politician, was a professor of law and a rector at the University of São Paulo, where he founded the Instituto Brasileiro de Filosofia and its journal, *Revista brasileira de filosofia.* Reale is a prolific author, and his books embrace the full range of his concerns, although his greatest contribution lies in the philosophy of law.

Reale has developed an analytical method (derived from German phenomenology and Italian historicism) that he calls "critical ontognoseological historicism." Rejecting both traditional realism and idealism, he locates the transcendental conditions of human experience and knowledge in a fundamental and inseparable correlation of subject and object. These conditions are mutually implicit and reciprocally necessary and are comprehensible only as moments in a polar dialectical process. Man's being emerges only through his own historicity, as values are realized in time through his conduct. The person finds his essence (*ser*) in what he ought to be (*dever-ser*), and he is the source of all values. Values are possible only where there are persons, and personality consists in conduct that is comprehensible only with reference to ends and values. A phenomenological description of human action reveals its essential orientation toward ends that represent values determining action and serving as the foundation of the "ought-to-be" in which man finds his essence. Reale interprets human history as a process through which values are converted into ends, accompanied by cultural crises whenever a new generation refuses to recognize the value of traditional ends.

Legal phenomena are basic to the realization of values in common. In law, two persons are joined in a polar nexus of common needs. Reale distinguishes three traditional approaches to the understanding of the nature of law. Sociologism interprets law as a positive fact and explains it in sociological and historical terms. Neopositivism interprets law as the expression of the operative norms of a given society and analyzes its function therein. Culturalism interprets law as axiological in nature and investigates the transcendental conditions that make it possible. Reale rejects all three as merely partial interpretations. Fact, norm, and value, in his view, are dialectically unified and not merely juxtaposed. Law is a fact through which values are made concrete in history and through which intersubjective relations are normatively ordered.

See also Historicism; Idealism; Ideas; Latin American Philosophy; Philosophy of Law, History of; Philosophy of Law, Problems of; Realism; Value and Valuation.

Bibliography

WORKS BY REALE

Teoria do direito e do estado (Theory of law and the state). São Paulo: Martins, 1972.
Filosofia do direito (Philosophy of law). São Paulo: Edição Saraiva, 1962.
Pluralismo e liberdade (Pluralism and liberty). São Paulo: Edição Saraiva, 1963.

WORKS ON REALE

Revista brasileira de filosofia 11, Fasc. 42 (April–June 1961). Dedicated to Reale on the occasion of his fiftieth birthday, this issue contains five expository articles as well as other material about him, his thought, and activity.

Fred Gillette Sturm (1967)

REALISM

In the early history of philosophy, particularly in medieval thought, the term *realism* was used, in opposition to nominalism, for the doctrine that universals have a real, objective existence. In modern philosophy, however, it is used for the view that material objects exist externally to us and independently of our sense experience. Realism is thus opposed to idealism, which holds that no such material objects or external realities exist apart from our knowledge or consciousness of them, the whole universe thus being dependent on the mind or in some sense mental. It also clashes with phenomenalism,

which, while avoiding much idealist metaphysics, would deny that material objects exist except as groups or sequences of sensa, actual and possible.

THE POLEMIC AGAINST IDEALISM

At the close of the nineteenth century, idealism was the dominant Western philosophy, but with the opening of the twentieth century, there was an upsurge of realism in Britain and North America, associated in the former with G. E. Moore, Bertrand Russell, and Samuel Alexander and in the latter with William James (despite his pragmatism), the new realists, and later the critical realists. Before a discussion of realist doctrine, a brief survey may be given of its attack on idealism.

The claim that material objects cannot exist independently of mind had been made on various grounds. First, the analysis of perception, especially of illusions, was held to show that our knowledge was limited to groups of sensations "in the mind" or to products of the synthesis or interpretation of sensory data. Later idealists, under the slogan "all cognition is judgment," stressed the role of judgment and interpretation in perception, concluding that objects as we know them must be largely or even wholly the work of the mind. Second, physical objects cannot exist independently of the mind, for whatever is known is relative to the mind that knows it. This is the "egocentric predicament"—that one can never eliminate the "human mind" from knowledge and discover what things are like apart from one's consciousness or, indeed, whether they exist when they are not known, for the discovery itself involves consciousness and thus would be knowing. This may also be stated in terms of the doctrine of internal relations—that the nature of anything is grounded in and constituted by the relations it has with other things; no two related things could be what they are if the relation between them did not exist, and so, as a special case of this, physical objects could not be as they are apart from their relation to the mind that knows them.

STATUS OF THE OBJECTS OF PERCEPTION. Concerning the analysis of perception, realist philosophers have devoted considerable attention to showing that in perception we obtain knowledge of external physical objects either directly or by means of sensa. Their accounts of perceiving and their solutions to the problems raised by illusions and other facts of perception differ greatly, but they agree in rejecting the view that things cannot exist unperceived. G. E. Moore's influential "Refutation of Idealism" consisted in an attack on this thesis, which, following George Berkeley, he stated as "*esse* is *percipi*" ("to be is to be perceived"). He claimed that in maintaining this the idealists had failed to distinguish between the act and the object in sensation. They had confused the sensation of blue with its object blue or, when claiming to distinguish them, inconsistently treated them as identical.

Sensations are alike in being acts of awareness but differ in what they are awareness of. Once the object is distinguished from the awareness of it, there is no reason to deny its existence unperceived. Further, in no other situation have we a better claim to be aware of something distinct, so that if sensations are not cases of awareness of objects, no awareness is ever awareness of anything, and we cannot be aware of other persons or even of ourselves and our own sensations. Fundamentally, Moore's thesis concerning sensations rested on introspection; it has been denied on a similar introspective appeal by upholders of the adverbial analysis of sensing, and Moore himself later had grave doubts about it. Commonsense realists would say that he conceded too much in talking of sensations and interpreting "being perceived" (*percipi*) as "being sensed" (*sentiri*); the proper starting point is our awareness of material objects. But Moore was no doubt accepting the usual conclusions from the argument from illusion. From his analysis arises the question: "What is the object of sensation?" The answer, "A sense datum," posed the problem, which he never solved, of the relation between sense data and material objects. It was met by others with some form of representative realism or, more usually, phenomenalism. Phenomenalism, however, particularly if coupled with the adverbial analysis of sensing, means the abandonment of realism. The idealist stress on judgment in perception was at first little discussed, but critical realism and the sense-datum theory later offered more plausible alternatives.

THE EGOCENTRIC PREDICAMENT. The realist attack on the egocentric predicament involved considerable discussion, particularly in the United States, and led to some close argument—for example, in attempts to show that the idealist principle led to self-contradiction or circularity when developed. The egocentric predicament was claimed to have no idealist implications. To infer from it that nothing exists outside consciousness is simply fallacious—that one cannot discover *X* does not mean that *X* does not exist or even that it is unreasonable to suppose that *X* exists. Indeed, if it were true that things could not exist apart from a person's consciousness of them, neither, presumably, could other persons; the predicament would imply an incredible solipsism.

Nor is there any evidence of the lesser conclusion that objects outside consciousness would be quite different. No conclusion about the degree of distortion introduced by our consciousness follows from its ubiquity, and it may be negligible; one can only try to discover the degree by comparing various methods of knowing. (Distortion by the method of observation may be serious in atomic physics, but the same argument that establishes distortion there shows it to be negligible for objects larger than atoms.)

The predicament is sometimes stated in terms of the privacy of experience—a person can never know anything that is not a content of his private experience. This, however, is question-begging in that it simply denies the ordinary assumptions that we are aware of other persons and external public objects. There may be grounds for denying these assumptions in certain cases, but such grounds rest on evidence of causal processes and of illusions, evidence that is largely obtained from other persons, or with the aid of public objects, or from comparisons with perceptions of public objects. Further, though more dubiously, Wittgenstein has argued that if we had only private experiences, not only would they be incommunicable, but also we could not describe or speak about them even to ourselves, for the use of language implies rules that are communal and have to be established and checked with respect to public objects.

Against the doctrine of internal relations it was claimed that relatedness is compatible with independence, that the same thing can enter into a variety of relations without losing its identity. This seemed so obvious that James confessed to finding it "weird" to have to argue for it. (Anticipating a contemporary approach, he accused the idealists of confusing linguistic or conceptual differences with factual ones; in referring to two relations of an object, our phrases and thoughts differ, but there is no corresponding difference in the object itself.) As the realists were defending what in their eyes was obvious, they were forced into detailed criticism rather than into the kind of positive thesis that can be readily summarized.

This battle was certainly won by the realists in that few English-speaking philosophers in the twentieth century espoused idealism. Indeed, to anyone coming from contemporary discussions, the controversy has an air of unreality. Partly this is because in a climate of thought that respects common sense and science, realism seems so obvious a starting point that it is difficult to explain how the idealist view ever seemed plausible; partly it is because current idioms, issues, and logical presuppositions are so different from earlier ones. Granted, however, that material objects exist independently of our perception, the difficulties facing a realist account of this perception still remain and cause serious divisions among realists.

DIRECT REALISM

Direct realism is the general view that perception is a direct awareness, a straightforward confrontation (or in touch, contact) with the external object. It may be further subdivided according to the various attitudes then taken toward illusions and hallucinations. In contrast, there are the various types of indirect or dualist realism, which claim that perception is primarily of mental representations of the external object, as in traditional representative realism, or that our perception of the external object is by means of private, mental sensa.

NAIVE REALISM. Naive realism is the simplest form of direct realism and is usually alleged by philosophers to be an innocent prejudice of the average person that has to be overcome if philosophical progress is to be made. It is normally stated in terms of sensible qualities or sensa. When we look around us, we can distinguish various colored, shaped expanses that we suppose to be the surfaces of material objects, we may hear various sounds that we suppose to come from such objects, we may feel something smooth and hard that we suppose to be a table top, and so on. Naive realism claims that these suppositions are all correct—that the shapes, colors, sounds, and smooth, hard expanses (the sensible qualities) are always the intrinsic properties of material objects and in sight and touch are their surfaces.

Such a claim can easily be shown to be erroneous by the argument from illusion. When *A* looks at the table from above, he sees a round expanse; when *B* looks at it from a distance, he sees an elliptical one. Without self-contradiction, however, the round and elliptical shapes cannot both be the surface of the table—that is, an intrinsic property. Similarly, when *C*, who is color-blind, looks at a red book, he sees a black shape that, again, cannot be the surface of that red book; when *D*, a drunkard, sees snakelike shapes on the bed, they are not real snakes. Such examples may be multiplied indefinitely and dispose of naive realism as thus stated, but commonsense realists would say that the doctrine misrepresents the views of the average person and that philosophical discussions of it beg the question in favor of dualism by speaking of sensible qualities or sensa as distinct from physical objects.

NEW REALISM AND THE SELECTIVE THEORY. The new realists—E. B. Holt, W. T. Marvin, W. P. Montague, R. B. Perry, W. B. Pitkin, and E. G. Spaulding—are notable chiefly for a common realist platform published in 1910 and expanded in 1912 and for their polemic against idealism. Their realism was carried to the Platonic extreme of claiming real existence for logical and mathematical entities, and they had difficult and conflicting views about consciousness. Without, however, pursuing these, we may note their main attempt (by Holt) to deal with illusions, which is a version of what is often called the selective theory. The essential points of this theory are, first, all the various appearances of an object are its intrinsic, objective properties and are directly apprehended by the percipient. For example, the table that looks round to *A* and elliptical to *B* is intrinsically both round and elliptical; the mountain that looks green close up and blue in the distance is both green and blue. There is nothing private or mental about such appearances, for they can be photographed, as can mirror images and various optical illusions. Second, the function of the nervous system and of the causal processes in perception is to select and reveal to the percipient one property from each set of properties, for example either the elliptical or the round shape of the table.

One difficulty in this is that it does not really account for error. If we are always directly aware of actual characteristics of objects, what sense does it make to talk, as we do, of illusions, mistakes, or misperceptions? Another lies in the weakness of the selective theory compared with the generative theory, adopted by dualist realism, which states that the sensible qualities, or sensa, are "generated," by the action of the object on the sense organs and nervous system and thus are not intrinsic properties of external objects. The usual reasons for preferring the generative theory are, on the one hand, that it is self-contradictory to say the table is intrinsically both round and elliptical or the mountain is intrinsically both green and blue. Furthermore, objects must be incredibly complex if they are to possess all these shapes and colors, plus, presumably, qualities corresponding to the queer appearance of objects when one has taken mescaline or suffers from giddiness or double vision. On the other hand, it is not clear how the nervous system specifically responds to or selects one of the various shapes, colors, and so on. This is particularly so in such cases as color blindness, drugs, and double vision, where the different appearances are the result of differences in the percipient and where the pattern of light waves can be detected as already differentiated for the shape and color normally perceived.

The generative theory, however, fits the facts of the causal processes quite well; it is natural to suppose that the generation of the sensory experience and its sensum occurs at the end of the causal chain that extends from object to brain by way of sense organ and nerves. This is confirmed by the reproduction of such experiences in mental imagery (presumably because the appropriate brain activity recurs), by the sensations resulting from electrical stimulation of the brain, and by the time lag that may occur between an event and our perception of it—all things that the selective theory cannot explain. Also, the generative theory can explain how voluntary selection occurs. When we turn our head to look at *X* rather than *Y*, we are allowing light from *X* rather than *Y* to strike our eyes and thus bring into being the sensa appropriate to *X*. As to photographing appearances, the photograph corresponds to the retinal image, not the sensum—that is, it reproduces not the perceived appearance but an intermediate cause of it; to enter into human experience, it must, in turn, be perceived by generating sensa.

PERSPECTIVE REALISM AND THEORIES OF APPEARING. The first objection to the selective theory—that it makes objects possess contradictory qualities—might be met by stressing that shapes, colors, and other qualities are not intrinsic but relative properties. The table is round from here, elliptical from there; the mountains are green in this light, blue in that light, and so on. This idea has been coupled with direct realism in a number of similar theories: perspective realism (E. B. McGilvary), objective relativism (A. E. Murphy), or the theory of appearing. (This last name was given by H. H. Price to a view put forward by H. A. Prichard. Roderick M. Chisholm, however, uses it more widely, and it is convenient to class all these views as theories of appearing.) Their central point is that direct realism can deal with illusions, or at least perceptual relativity, by saying that sensible qualities are not possessed by the object *simpliciter* but are always relative to some point of view or standing conditions. We always perceive sensible qualities in some perspective—spatial, even temporal (we see the distant star as it is from here and now), or illuminative (the object as it is in this light). (In such theories the shape, color, and so on are possessed by the object at its own location but are perceived subject to perspective, meaning from a viewpoint. In contrast, Bertrand Russell had a phenomenalistic theory of "perspectives" that were spread through space as possible sensa and actualized by or in the percipient.)

Such perspective-realist statements as "The table is round from here" sound forced, for the natural word to

use is *looks*, not *is*, and it is possible to express this kind of direct realism in terms of looking or appearing. Physical objects simply are such that they appear different from different positions, and we see them as they appear from a viewpoint or in certain conditions. Thus, we may see the round table looking elliptical from here, but even so it is still the table that we see.

Thus far the theory is trite and does little more than state the situation in a way that dualists could accept and then claim to analyze. To be distinctive, it must, as its essential characteristic, separate directness and incorrigibility. Sense-datum theory links the two, assuming that if we see an object directly, we must see it as it actually is. Thus, when the round table looks elliptical, we do not see it directly; what we see directly is an elliptical datum belonging to it.

In contrast, theories of appearing must simply claim that seeing an object directly is compatible with variation or even error in perception, so that we still see it directly when according to viewpoint, lighting, and similar factors, it appears really different from what it is. (Some might object that the theory cannot admit that perceiving is ever erroneous. Perspective realism treats all properties as relative and all perspectives as equal—the table is round from here, elliptical from there, but not round in itself; similarly all appearances should be treated as equally valid. Nevertheless, it seems more plausible to treat some appearances as privileged; in some conditions we see the real shape, the round object appearing as it is—that is, round. It may be considered a weakness of the perspective theory that it does not take into account the fact that objects do seem to have real [measured] shapes and volumes absolutely, not relative to a viewpoint.)

The approach of theories of appearing may deal plausibly with perspectival and similar variations, but it has two main defects. First, not all variations are of this nature. In double vision or mescaline illusions there seems to be existential appearing—there may appear to be two or even many tables when we look at one table. Price has argued that this cannot really be a case of directly seeing one table, for it differs significantly from seeing something merely with different properties, such as seeing a brown table instead of a black one. Also, many illusions are the result of subjective factors, so that it is difficult to say that one has a genuine perspective.

Talk of physiological perspectives is little help. "The bottle from here" is not on a par with "the bottle as it is to someone who has taken mescaline," for mescaline may cause a range of different experiences. Similarly, when a sentry at night is convinced he sees the enemy approaching but only a shadow is there, is he directly seeing the shadow in some special perspective, such as "the way it is to an anxious sentry" or "looking like a man"? Another anxious sentry might see it as a shadow and say it does not look like a man. And in a full hallucination there is no object at all. Second, theories of appearing cannot deal plausibly with the causal processes in perception since they have to adopt the selective theory. Further, we do know with varying degrees of completeness why things suffer perspectival distortion or how they cause illusion. The explanations concerned are often in terms of the causal processes and so seem to call for the generative theory and the abandonment of direct realism.

COMMONSENSE REALISM. In the tradition of Thomas Reid, revived by G. E. Moore, many twentieth-century British philosophers defended what they took to be a commonsense view of perception. Moore's defense was primarily of the certainty of such simple perceptual statements as "This is a hand"; he argued that denial of these statements leads to inconsistency in beliefs and behavior and that the grounds for their denial involve propositions less certain than they are. However, his analysis of such statements in terms of sense data led away from direct realism and the commonsense view of the nature (as opposed to the reliability) of perception.

Defense of common sense became particularly associated with the Oxford linguistic analysts. Strong critics of the sense-datum theory (unlike Moore), they also reject the traditional naive realism as unfair to common sense—after all, we do not think that everything we see is the surface of a physical object (certainly not lightning flashes or rainbows) and are quite ready to admit that we often see things looking different from what they are. Although quarreling with the common philosophical uses of *appear, direct*, and *real*, they maintain a direct realism not unlike the theories of appearing and attempt to show in detail that in so-called illusions, including reflection and refraction, we do actually see the physical object concerned. Criticism has been made of the view that hallucinations are indistinguishable from normal perception, and more positively it may be claimed that hallucinations are mental images confused with perceptions owing to such special circumstances as drugs or fever. It is doubtful whether this can explain all the cases, and the role of the psychological processes—for example, in attention or in the influence of expectation and past experience—throws doubt on the directness of perceiving.

Some attempt has also been made to deal with the causal processes, but not very convincingly. Attacks have been made on the dualist interpretation for making it seem that we perceive something in our heads and not external objects and for the view that perceiving involves awareness of sensations. But linguistic analysts have said little of a positive nature; their main attitude is that the causal processes are at most only the conditions of perception and are the concern of the scientist but that the philosopher is concerned with perception itself, which is a skill or instantaneous achievement, not a physical process or the final stage of one. Unfortunately, scientists generally claim that the study of the causal processes requires representative realism, and even if the average person does not bother about them, an adequate philosophical theory cannot ignore the causes and conditions of perceiving, particularly since the explanation of illusions depends on them.

INDIRECT OR DUALIST REALISM

Many realists are persuaded by the argument from illusion and by their study of the causal and psychological processes in perception to reject direct realism and to distinguish between external material objects as the causes and ultimate objects of perceiving and private sensa that are the mental effects of brain processes due to the action of those objects on the sense organs. The classic form of this general view was the representative realism (also called the representative or causal theory) of René Descartes and John Locke, which is still maintained in principle by many scientists. From Berkeley on it suffered much criticism, and its defects led to its being unpopular among philosophers. Modern attempts have been made, however, to remedy these defects and to propose an acceptable theory. The resultant position we shall discuss as critical realism. Although they start from an analysis of perceptual experience and do not argue from the causal processes underlying it, supporters of the sense-datum analysis who are not phenomenalists are forced into one of these kinds of dualist realism.

REPRESENTATIVE REALISM. In what is loosely called "seeing a table," light rays reflected from the table strike the eye, cause chemical changes in the retina, and send a train of impulses along the optic nerve to the brain. The resultant brain activity is then said to cause the mind of the percipient to be directly aware of private sensa (Locke called them "ideas") that represent the shape, color, and other visual properties of the table. A similar account is given for the other senses. The essential point is that perceiving proper is the direct awareness of sensa; perceiving external objects is redefined as perceiving sensa caused by them, and so all our awareness is strictly limited to sensa. "Represent" is usually interpreted in accordance with the doctrine of primary and secondary qualities—that is, the sensa resemble the object in spatiotemporal properties but not insofar as colors, sounds, smells, and other secondary qualities are concerned. Modern analogies of "representing" are the relation between a map or radar screen and the region they cover or between television or movies and the studio events reproduced.

Merits of representative realism. Representative realism has important merits. It is the easiest inference from the scientific account of the causal processes up to the brain in all perceiving and fits other scientific evidence. Thus, color blindness and deafness are the result of defects in the sense organs that so affect all subsequent stages in the causal transmission that the resultant sensa are different from normal. That electrical stimulation of the brain causes sensations of color, smell, and so on, according to location, seems to confirm the theory, and it can easily accommodate the time lag in perception. Further, by holding that representation does not amount to resemblance in the case of secondary qualities, it can be made to fit the distinction between the world as we see it (that is, the sensa grouped as ostensible objects) and the scientific account of material objects, which is in terms of colorless, tasteless, and smell-less elementary particles.

Representative realism also accounts for illusions, dreams, images, hallucinations, and the relativity of perception. Relativity and many illusions result from changes in the stimulation of the sense organs because of distance, medium, angle of sight, and other relevant factors; such changes affect all that follows and so vary the sensa caused. Other illusions are the result of misinterpretation of sensa. In imagery and dreams the brain activity that occurred in corresponding perceptions is reactivated as the result of internal causes and so brings about the recurrence of similar sensa. (The reactivation may be only partial, and the resultant data may be consciously or unconsciously altered by the mind.) Hallucinations are also imagery. Since the images are of a similar character to normally perceived data and are the result of a similar immediate cause in the brain, it is easy to see how they may merge in integrated or triggered hallucinations or how perception may be imaginatively supplemented. The standard explanation of phantom limbs—that they are sensations caused by irritation at the stump of nerves normally coming from the amputated limb—is also accommodated. As perception is confined

strictly to the effects of the causal chain, interference with it en route may readily deceive us.

Finally, representative realism has also traditionally been part of the widely accepted interactionist or dualist account of the relation of mind and body: The body affects mind in perception, mind affects the body in voluntary action. Not all who accept that theory realize that they are saddled with representative realism.

Defects of representative realism. Despite its merits, representative realism has some serious defects. If, as it claims, our perceiving is strictly awareness of the mental ideas or sensa, it is difficult to see how we can break out of the circle of sensa and observe external objects. How can we tell what these objects are like; indeed, how do we know that there are such objects? If we try to verify the existence of the table by touching it, we simply obtain more sensa—tactile ones—and if we see our hands touching the table, we are just having visual sensa. Whenever we try to peer over the barrier of sensa, we just get more sensa. This difficulty undermines the analogies used in the theory. Representation is conceived of as something like mapping or photographing, but we know a map represents or a photograph resembles an object because we can observe both and compare them; ex hypothesi, however, we can never strictly observe both objects and sensa to compare them. Observing objects is just observing sensa, so we do not know that objects and sensa resemble each other in primary but not in secondary qualities.

It is often said that representative realism not only leads to skepticism but is also self-refuting, cutting off the branch on which it sits. Its premises and evidence assume that we discover the action of the objects on the sense organs by observing them. Its conclusion—all our perception is of sensa—denies that we can do this. However, there would be self-refutation only if the conclusion contradicted the premises, which it need not do if carefully stated. The theory may be regarded as really distinguishing two types of perceiving: perception in its everyday meaning, which is discovering about external objects by means of the senses, and perception proper—direct awareness of sensa. It is saying that the first type really amounts to or, better, is really effected by the second type. Thus, granted that by perceiving sensa we do discover the nature of objects (at least insofar as their primary qualities are concerned) and their interaction, the first type of perception and the evidence it gives still hold good, and there is no self-refutation. Nevertheless, the skepticism remains, for since our direct awareness is limited to sensa, we do not know that there are objects or what they are like; we only suppose or guess that and what they are.

Even though representative realism need not be self-refuting, it is open to the charge of circularity if considered as an attempt to explain perceiving. It appears simply to transfer perceiving as ordinarily conceived (a face-to-face confrontation) from outside to inside the person; perceiving external objects is now put forward as perceiving private replicas of them, for we look at maps and television pictures in the same way that we look at the countryside. Even if we say perceiving objects is achieved by perceiving sensa, there is the same duplication of perceiving, which is thus explained in terms of itself.

Representative realism's view of the mind is rather crude, for it tends to speak almost as if the self or mind were a little person in the head looking at pictures of the outside world. It is not clear how sensa can exist in an unextended mind, since they apparently possess shape and size; nor is any serious attempt made to fit the psychological processes of perception into the general scheme.

There are special difficulties for those versions of the theory that claim that in perceiving objects we infer the existence or nature of external objects from our sensa. Apart from the inevitable dubiety of such inference, the main objection is that we are never conscious of these inferences nor are we aware of sensa as such—that is, as private mental data. If we were, it is difficult to see how the notion of publicly observable causes would occur to us. But the representative theory may simply say that the sensa seem to be external (or externally caused) from the start and that any inference is justificatory to deal with skeptics. (This seems to have been Locke's view in his *Essay concerning Human Understanding*, Bk. IV, Ch. xi, Sec. 2.)

CRITICAL REALISM. Critical realism is the name primarily given to the views expressed by the American authors of *Essays in Critical Realism*—namely, that the data in perception (that is, what is intuited, what we are directly aware of) are not actually part of external objects but are "character-complexes … irresistibly *taken*, in the moment of perception, to be the characters of existing outer objects" (p. 20). In veridical perception these characters are the characters of external objects; in illusions they are not. The authors were unfortunately divided over the nature of this datum or character complex, Durant Drake, A. K. Rogers, George Santayana, and C. A. Strong claiming that it was not a mental existent or any kind of existent, but only an essence, a mere logical entity or uni-

versal, whereas A. O. Lovejoy, J. B. Pratt, and R. W. Sellars held that it was a mental existent, a content of sensory experience. It is difficult to grasp what the datum can be if it is not a mental content or existent, and so the second version is the more plausible and is adopted here. Although clearly dualist, it should not be confused with representative realism; in fact, it provides remedies for representative realism's main faults.

The critical realists held that the root of the troubles of representative realism lay in its failure to analyze perceiving or perceptual knowledge. Accepting the ordinary notion of perceiving as intuiting, which means a direct awareness or confrontation, and finding that because of the causal processes and of illusions such awareness was not of external objects, Locke concluded that it must be of intramental ideas and so imprisoned us in the circle of such ideas. The more reasonable conclusion, however, would be that this ordinary notion of perceiving is wrong and that a more careful analysis is needed. This will show that an essential feature of perceiving, even as ordinarily understood, is that it is the way we discover the existence and nature of external objects—that it is, in fact, a claim, often justified, to knowledge. If we appreciate this from the start, we shall not be tempted by the apparently intuitive character of perceiving into an analysis that limits it to ideas, and if we remember that this knowledge claim is not always justified—that is, that there are illusions and errors—we shall avoid the other pitfall of direct realism, in which error becomes inexplicable.

The next step is to realize that though it involves an intuition or direct awareness, perceiving is much more than this. It also involves an active external reference, as is implied by the knowledge claim; we refer this intuited mental content or character complex to an external object—that is, we explicitly judge that it is, or is the character of, an external object or we unreflectingly take it to be this or we immediately react to it as if it were an external object. These modes of reference are differently stressed by different writers, but the point seems to be that they occur in varying degrees according to circumstances. Our perception is sometimes an explicit identification or judgment, or at least it immediately issues in one—for example, we say, "Here's our bus" or "There's Tommy"; more often we just see that it is Tommy without formulating any judgment, or our perception that it is our bus and our starting to go and catch it seem indistinguishable, for the reference to the external object is manifest in an immediate physical response.

All the same, in contrast to the behaviorists, the critical realists stressed that there was an intuited mental content, the character complex of which we were directly aware. Attempts were made to fit the analysis in with current psychology by explaining how this external reference arose in childhood—the apparent externality of the content was with us from the beginning of perceptual discrimination, largely because the external reference was founded in physical response to the object.

There is some similarity between this "reference of an intuited datum to an external object" and the "taking for granted that a sense datum belongs to a material object" of Price's sense-datum theory, especially since both stress that no distinction between datum and object is drawn by the percipient at the time. But there is a difference in starting point and emphasis. Price began with sense data, treating them as distinct existents and willing to allow that material objects consisted of them. This branch of critical realism began with knowledge of external objects, but, being mental, the content or datum distinguished within it was not regarded as capable of distinct existence and was very difficult—much more so than Price thought—to isolate even subsequently from the associated reference. Also, reference covered a wider set of activities than taking for granted, for it also involved the bodily reactions. In order to stress the relative subordination of the datum, some critical realists spoke of perceiving external objects by means of, guided by, or mediated by, the datum.

Since critical realism can agree that the datum is generated, it is free from the difficulties of the selective theory and can share in the advantages of representative realism. In this version it seems able to avoid the latter's worst faults. There is no self-refutation, for from the start perceiving is always perception of external objects by means of the intuited data, an analysis that does not deny that we perceive such objects. There is no duplication or circularity, for the direct awareness of the datum is not a replica of perceiving; insofar as it can be distinguished at all, it is much less complex than perceiving, for it involves no identification with external objects and is not in itself directed on them—hence, the map and movie analogies are essentially faulty. Common sense is not being offered an explanation of perceiving in terms of perceiving; it is being shown that perceiving is far more complex than common sense supposes, involving not only causal processes that bring about the datum or mental content but also the psychological processes of reference or response.

Moreover, there need be no skepticism. True, in perceiving we only take the datum to be an external object or its properties, and this may, of course, be erroneous. In a

sense it is always erroneous in that the datum or content is never the object, but normally the taking or reference is correct to the extent that we are perceiving an external object and that the intuited characters also do characterize the external object insofar as primary qualities are concerned; to that extent we are perceiving actual properties or at least projections of them. In general, the claim that perceiving is thus far veridical and amounts to knowledge is said to be the best hypothesis to explain the order and nature of our sense experiences. The realist claim is simply that once ordinary errors and illusions are ruled out by comparing the evidence of different senses or of different persons, the simplest explanation of the situation is that there are external objects causing the sense data or contents and corresponding to them in primary qualities. And this is plausible because if we dismiss as incredible solipsism the view that only oneself and one's own sense experiences exist, then the only real alternative is phenomenalism, a view that has fatal weaknesses and really amounts to proposing a series of deceptive coincidences.

Critical realism is not fully satisfactory, however, particularly if regarded as a theory of perceptual consciousness—that is, as an account of the mental activity that goes on in perception. Thus, the alleged datum or character complex suggests a group of sense data and invites the objections discussed under the entry Sensa. A closer examination is required not only of the concepts of datum and reference but also of the general relation of mind and body presupposed in perception and of the nature of mental contents; above all, the theory must take full account of the numerous quasi-interpretative activities that modern psychology has found to be involved in perception.

See also Illusions; Sensa.

Bibliography

GENERAL

A clear, simple introduction to the philosophical problems of perception is Bertrand Russell, *Problems of Philosophy* (London: Williams and Norgate, 1912), and a fuller one is W. P. Montague, *The Ways of Knowing* (New York: Macmillan, 1928). A good account of modern positions on perception is given by T. E. Hill, *Contemporary Theories of Knowledge* (New York: Ronald Press, 1961). Detailed summaries of many realist works are given by W. H. Werkmeister, *A History of Philosophical Ideas in America* (New York: Ronald Press, 1949). Less detailed but more intelligible is John Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1957); also see Rudolf Metz, *A Hundred Years of British Philosophy* (New York: Macmillan, 1938). Many of the works listed below deal with the topics of more than one section.

THE REALIST POLEMIC AND NEW REALISM

The main source for new realism is E. B. Holt and others, *The New Realism* (New York: Macmillan, 1912), but new realism owed much to William James; see the papers (dating from 1904) collected in his *Essays in Radical Empiricism* (London and New York: Longman, 1912). For a useful collection of articles from this early period see Roderick M. Chisholm, *Realism and the Background of Phenomenology* (Glencoe, IL: Free Press, 1960). Among other important and often closely reasoned articles are R. B. Perry, "The Ego-centric Predicament," *Journal of Philosophy* 7 (1) (1910): 5–14; Bertrand Russell, "On the Nature of Truth," *PAS* 7 (1906–1907): 28–49; and G. E. Moore, "The Refutation of Idealism," in his *Philosophical Studies* (London: Routledge, 1922), to which compare W. T. Stace's counterattack, "The Refutation of Realism," *Mind* 43 (170) (1934): 145–155. For a general summing up in favor of realism, see D. C. Williams, "The A Priori Argument for Subjectivism," *Monist* 43 (1933): 173–202, "The Inductive Argument for Subjectivism" and "The Inductive Argument for Realism," *Monist* 44 (1934): 80–107, 186–209. For a once influential direct-realism treatment of perceptual problems, see T. P. Nunn, "Are Secondary Qualities Independent of Perception?," *PAS* 10 (1909–1910): 191–218. Ludwig Wittgenstein's language arguments are in his *Philosophical Investigations* (Oxford: Blackwell, 1958), Secs. 256ff.; for criticisms see Carl Wellman, "Wittgenstein and the Ego-centric Predicament," *Mind* 68 (270) (1959): 223–233, or the symposium "Can There Be a Private Language?," *PAS*, Supp. 28 (1954).

PERSPECTIVE REALISM AND ALLIED THEORIES

Perspective realism and allied theories are stated by Evander Bradley McGilvary in *Toward a Perspective Realism* (La Salle, IL: Open Court, 1956) and in "Perceptual and Memory Perspectives," *Journal of Philosophy* 30 (1933): 310ff. Older versions are by Samuel Alexander, "On Sensations and Images," *PAS* 10 (1909–1910): 1–35, and H. A. Prichard, *Kant's Theory of Knowledge* (Oxford: Clarendon Press, 1909), Ch. 4. Despite its title, G. Dawes Hicks, *Critical Realism* (London: Macmillan, 1938), gives a theory of appearing. Such theories are lucidly discussed by Roderick M. Chisholm, "The Theory of Appearing," in *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950). C. D. Broad, *The Mind and Its Place in Nature* (London: Kegan Paul, 1925), and H. H. Price, *Perception* (London: Methuen, 1932), criticize these theories carefully from a sense-datum standpoint, though Price's article "Illusions," in *Contemporary British Philosophy*, edited by H. D. Lewis (London, 1956), Vol. 3, defends a limited perspective realism.

COMMONSENSE REALISM

For commonsense realism see G. E. Moore's "A Defense of Common Sense" and "Proof of an External World," which are in his *Philosophical Papers* (London: Allen and Unwin, 1959), but also see his *Some Main Problems of Philosophy* (London: Allen and Unwin, 1953). The staunchest more recent defender of common sense against the argument from illusion is J. L. Austin in his lucid and lively *Sense and*

Sensibilia (London: Oxford University Press, 1962), and somewhat similar views are clearly and concisely expressed by Anthony M. Quinton, "The Problem of Perception," *Mind* 64 (253) (1955): 28–51. Gilbert Ryle's "Sensations," in *Contemporary British Philosophy*, edited by H. D. Lewis (London, 1956), Vol. III, and Ryle's *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954) try to deal also with the causal argument in a nontechnical manner. D. M. Armstrong, *Perception and the Physical World* (London: Routledge and Paul, 1961), defends direct realism but in so doing is driven toward behaviorism.

REPRESENTATIVE REALISM (OR THE CAUSAL THEORY)

For early statements of representative realism, see René Descartes, *Principles of Philosophy*, Pt. IV, and John Locke, *Essay concerning Human Understanding.* Bk. 2, Ch. 8.

Representative realism is assumed by many modern neurologists, though often not under its philosophical title. Walter Russell Brain states and discusses it as "physiological idealism" in "The Neurological Approach to the Problem of Perception," *Philosophy* 21 (79) (1946): 133–146, reprinted with further consideration of perception in his *Mind, Perception and Science* (Oxford: Blackwell, 1951); he gives a further defense of his position in his *The Nature of Experience* (London: Oxford University Press, 1959). J. C. Eccles, *The Neurophysiological Basis of Mind* (Oxford: Clarendon Press, 1953), pp. 279–281, outlines the theory as if it were fact. J. R. Smythies, *Analysis of Perception* (London: Routledge and Paul, 1956), puts forward an improved form of it closer to critical realism.

CRITICAL REALISM

The main source for critical realism is Durant Drake et al., *Essays in Critical Realism* (London: Macmillan, 1920), which reveals the differences as well as agreements; see also other works by the essayists, especially R. W. Sellars's comprehensive *The Philosophy of Physical Realism* (New York: Macmillan, 1932), his general apologia, "A Statement of Critical Realism," *Revue internationale de philosophie* 1 (1938–1939): 472–498, and A. O. Lovejoy's impressive general defense, *The Revolt against Dualism* (La Salle, IL: Open Court, 1930). R. J. Hirst, *The Problems of Perception* (London, 1959), also discusses commonsense and representative realism and reaches a somewhat similar position. A primarily pragmatist view that has affinities to and criticism of critical realism is C. I. Lewis, *Mind and the World Order* (New York: Scribners, 1929).

OTHER RECOMMENDED TITLES

Almeder, Robert. *Blind Realism*. Lanham, MD: Rowman & Littlefield, 1992.

Alston, William. *A Realist Conception of Truth*. Ithaca, NY: Cornell University Press, 1996.

Benacerraf, P. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–680.

Brink, D. *Moral Realism and the Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Carnap, Rudolf. "Empiricism, Semantics, and Ontology." In his *Meaning and Necessity.* 2nd ed. Chicago: University of Chicago Press, 1956.

Churchland, Paul M. *Scientific Realism and the Plasticity of Mind*. Cambridge, U.K.: Cambridge University Press, 1979.

Devitt, Michael. *Realism and Truth*. Oxford: Blackwell, 1984.

Dummett, M. A. E. "Realism." In *Truth and Other Enigmas*. London: Duckworth, 1963.

Goodman, Nelson. *Ways of Worldmaking*. Indianapolis: Hackett, 1978.

Moser, Paul K. *Philosophy after Objectivity*. New York: Oxford University Press, 1993.

Putnam, Hilary. *The Many Faces of Realism*. La Salle, IL: Open Court, 1987.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Quinton, Anthony. *The Nature of Things*. London: Routledge and Kegan Paul, 1977.

Searle, John R. *The Construction of Social Reality*. New York: Free Press, 1995.

Sellars, Wilfrid. *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Press, 1991.

Smart, J. J. C. *Philosophy and Scientific Realism*. London: Routledge and Kegan Paul, 1963.

van Fraassen, Bas. *The Scientific Image*. Oxford: Clarendon Press, 1980.

Wright, C. J. G. *Realism, Meaning, and Truth*. Oxford: Blackwell, 1987.

Wright, Crispin. *Truth and Objectivity*. Cambridge, MA: Harvard University Press, 1992.

R. J. Hirst (1967)
Bibliography updated by Benjamin Fiedor (2005)

REALISM [ADDENDUM]

Contemporary philosophical "realism" is not a single thesis but rather a diverse family of positions, unified chiefly by their invocation of certain characteristic images and metaphors. The realist about a region of discourse typically holds, for example, that our central commitments in the area describe a world that exists anyway, independently of us; that cognition in the area is a matter of detection rather than projection or constitution; and that the objects of the discourse are real things and not just linguistic or social constructions. Debates over realism defined in terms such as these persist in nearly every philosophical subdiscipline: from ethics and the philosophy of mind to the philosophy of science and the philosophy of mathematics. (Although it is common to describe a philosopher as a realist or nonrealist *tout court*, realism in one area is generally independent of realism in another, and advocates of global realism and its opposite number, global nonrealism, are comparatively rare.) Contemporary discussion is concerned in part with the evaluation of these discipline-specific realist theses. But it is also concerned (and increasingly so) with the more basic question of how exactly the realist's distinctive imagery is to be understood.

We may epitomize the realist's stance by saying that to be a realist about a region of discourse is to regard it as describing a genuine domain of objective fact. But what is it for a discourse to describe a "domain of fact"? And what is it for a domain of fact to be "objective"? These questions are usefully approached by attempting a taxonomy of the alternatives to realism. The nonrealist rejects the realist's rhetoric of objectivity. But this rejection can take a number of more determinate forms, and their variety sheds considerable light on what realism requires.

The realist's most basic commitment is to the view that statements in the target area purport to describe a world—to say how things stand with some distinctive range of objects or facts. This claim is often glossed as the minimal requirement that statements in the area be capable of truth or falsity. Realism is thus opposed at this most basic level to nonfactualism (also called irrealism or noncognitivism): the view that declarative statements in the target area cannot be evaluated as true or false and so cannot serve a descriptive function. Nonfactualist theses have been advanced mainly in moral philosophy, where it has been suggested that moral utterances serve to express emotional attitudes (emotivism: Blackburn, 1984; Gibbard, 1990) or to endorse or proscribe certain courses of action (prescriptivism: Hare, 1963; cf. Geach, 1963). But they have occasionally been proposed in other areas. Formalism in the philosophy of mathematics (the view that mathematics is a game with meaningless marks, manipulated according to formal rules) and instrumentalism in the philosophy of science (according to which theoretical statements function as uninterpreted tools for deriving predictions about future experience) are further examples of this kind of nonrealism.

To say that a region of discourse purports to describe a world is to say more than that its central commitments are apt for truth. It is to say, in addition, that they are aimed at truth—that they are typically put forward as genuine assertions about how things stand with their ostensible subject matter. Realism is thus opposed at this second level to fictionalism, the view that seeming assertions in the target area, though capable of truth, are in fact designed only to provide representations that are somehow "good" or "interesting" or "useful" for certain purposes. Fictionalist approaches have been developed mainly in the philosophy of science, where Bas van Fraassen's constructive empiricism provides a useful example (van Fraassen, 1980; cf. Churchland and Hooker, 1985). Van Fraassen agrees with the scientific realist, against the instrumentalist, that theoretical statements possess definite truth conditions and so constitute genuine representations of unobservable structures. However, he further maintains, this time against the realist, that the truth-value of a theory is irrelevant to its acceptability from the standpoint of science. The aim of science on van Fraassen's view is empirical adequacy: the correct description of the observable world. Theories may posit unobservable things. But a good scientific theory—one that satisfies to some high degree all of the aspirations implicit in the scientific enterprise—may be largely false in its account of such matters, so long as it is a reliable guide to the observable world. In advancing a theory in what seems to be the assertoric mode, the scientist shows only that he accepts it as empirically adequate. Van Fraassen's fictionalism thus consists centrally in his contention that the endorsement of a scientific theory does not involve the belief that it is true or that the unobservables it posits exist. Generalizing, we may say that realism involves, in addition to the semantic thesis of truth aptitude, the pragmatic thesis of truth directedness, according to which the target discourse aims at truth, and the endorsement of a claim is normally an expression of one's belief that it is true. (See Field, 1980, for a fictionalist approach to the philosophy of mathematics.)

Before we have a position that is recognizable as realist we must add one further ingredient. It is not enough that our central commitments aspire to truth. They must also be true, or at least not wildly mistaken. Realism is thus opposed at this third level to a conception of the target area as involving a fundamental mistake about what the world contains. This "error-theoretic" alternative to realism is typified by J. L. Mackie's view of morality (Mackie, 1977). According to Mackie, ethical discourse purports to describe a range of objective prescriptions, constraints on action that are somehow built into the fabric of nature. But since it can be shown (Mackie held) that there are no such items, it follows that morality is based on a mistake—the entities it purports to describe do not exist; the properties it trades in are not instantiated—and hence that moral discourse demands reconstrual, if not outright rejection. A more familiar instance of the error-theoretic approach is atheism, the view that theological discourse is vitiated by the mistaken supposition that God exists. Agnostic versions are also possible, though in fact they have played no significant role outside the philosophy of religion.

A philosopher who holds that our core commitments in an area succeed in providing a true account of their intended subject matter may be called a minimal realist about that area. It is sometimes suggested that there is nothing more to realism than this minimal view

and hence that once the questions of truth aptitude, truth directedness, and truth have been settled, there is no further space for debate about whether the discourse is to be understood "in a realistic fashion." There are, however, at least two reasons to resist this claim.

The first concerns the classification of reductionist positions. The behaviorist thesis that psychological statements can be reduced without remainder to claims about overt bodily movements and the like is clearly compatible with minimal realism about the mind. And yet the view that there is nothing more to being in pain than exhibiting "pain behavior" has generally been regarded as a clear alternative to a robust realism about mental states. It has thus become customary to insist that the realist's commitment to the truth of our views in the target area be a commitment to their truth on a literal or face-value construal (Blackburn, 1984, chap. 5; cf. Wright, 1983). The behaviorist translation of a simple psychological statement such as "Nadja is dreaming of Paris" will typically be a long conjunction of conditional claims describing the outward behavior Nadja would exhibit if prompted by various stimuli. But this paraphrase has a very different "surface form" from the psychological claim whose meaning it is meant to capture. And this suggests that on the behaviorist's account, the correct interpretation of psychological statements is not a face-value interpretation and hence that while he may endorse a version of minimal realism about the mental, the behaviorist should not be classed as a realist without qualification.

The second and more serious reason to resist the identification of realism with minimal realism is that minimal realism by itself involves no commitment to the mind independence or objectivity of the disputed subject matter. Immanuel Kant's transcendental idealism has generally been regarded as a paradigmatic alternative to full-blown realism about the external world; and yet it is fully compatible with minimal realism as defined above. Objects in space and time are real, for Kant, in the sense that much of what common sense and science have to say about them is literally true. And yet there is another sense in which they are not fully real. The structure of the spatiotemporal world is "conditioned" for Kant by the structure of the mind that experiences it. Empirical investigation is therefore not addressed to a domain of fact that is altogether "independent of us." Clearly, Kant's position should not be described as a species of realism without qualification, its consistency with minimal realism notwithstanding.

Much of the most important work on realism has been devoted to explicating the commitment to objectivity that seems a necessary component of any fully realist position. The most natural thought is to identify objectivity directly with a straightforward sort of mind independence. A state of affairs will then count as objective if it would have obtained (or could have obtained) even if there were no minds or mental activity. But this precludes realism about the mind itself and also about any discourse in the social sciences that concerns itself with the products of human thought and action. And this is implausible. It should be possible to be a realist about psychology, for example, while conceding that the facts it describes are obviously mind dependent in the sense that they would not have obtained if there were no minds.

One influential approach to this problem is due to Michael Dummett, whose work is largely responsible for the current prominence of realism as a theme in Anglophone philosophy (Dummett, 1978; cf. Wright, 1992). On Dummett's view the dispute over realism, though ultimately an issue in metaphysics, is best approached by recasting it as a dispute within the philosophy of language about how to construct a theory of meaning for the target discourse. A theory of meaning in Dummett's sense is a representation in propositional form of what a competent speaker knows in virtue of which he understands his language. Dummett identifies realism with the view that a meaning theory must take the form of a classical two-valued semantic theory: an assignment of truth conditions to sentences that respects the principle of bivalence, according to which every sentence is determinately either true or false. Realism's slogan is: To understand a sentence is to know its truth condition. The leading alternative—sometimes called semantic antirealism—holds instead that to understand a sentence is to know the conditions under which it is correctly asserted. A view of this sort assigns each declarative sentence a class of "verification conditions," each of which must be the sort of condition a competent human being can in principle recognize as obtaining. A semantic theory constructed upon such a basis will generally fail to respect bivalence. The only notion of truth it makes available will be epistemically constrained: Truth will be identified with knowable truth, and falsity with knowable falsity. On a view of this sort we shall not be entitled to say in advance that every well-formed question must have an answer, or that every statement of the form "*p* or not-*p*" must be true. This rejection of bivalence (and the closely related law of excluded middle) is the hallmark of semantic antirealism. To suppose that the only notion of truth we possess for a region of discourse is an epistemically constrained one is to suppose that the facts in the area are (as it were) cut to fit our intellectual capacities. Conversely, to insist that bivalence

must hold regardless of our cognitive limitations is to conceive of the facts at which our thought is directed as obtaining (in one sense) independently of us.

A closely related proposal has been advanced by Hilary Putnam, (1978, 1987). Putnam identifies full-blown "metaphysical" realism directly with the view that truth is epistemically unconstrained. As Putnam frames the issue, the metaphysical realist's characteristic thought is that an ideal theory might be false, where an ideal theory is one that satisfies perfectly every criterion we normally employ in deciding what to believe in the target area. In the scientific case, for example, an ideal theory would be one that supplies accurate predictions of experimental outcomes while simultaneously displaying every internal theoretical virtue that scientists consider in the context of theory choice: simplicity, elegance, explanatory power, "intrinsic plausibility," and the like. It is natural to suppose that such a theory could be false. After all, the theoretical virtues that provide our only grounds for choice among empirically equivalent hypotheses seem importantly subjective. A theory that strikes us as particularly powerful because it provides informative answers to interesting questions might strike creatures with different interests as unacceptably silent on important matters; a theory that strikes us as "intrinsically plausible" might strike creatures with different histories or cultures as strange and unlikely. The thought that an ideal theory can be false thus seems a natural expression of an appropriate human modesty, according to which we can have no guarantee in advance that our contingent, biologically, and historically conditioned sense of theoretical virtue must be a reliable guide to the facts about the physical world.

Putnam rejects this natural thought. Metaphysical realism presupposes a concept of truth that is radically divorced from our notion of correct assertion. But according to Putnam such a concept is unattainable. Putnam's case for his view, like Dummett's, defies simple summary; but in rough outline it proceeds as follows: The only serious effort to explain an epistemically unconstrained notion of truth is a version of the correspondence theory of truth. This approach proceeds in two stages. First, subsentential expressions such as names and predicates are associated with objects and properties as their referents. Then truth as a feature of sentences is defined recursively according to a scheme well known to logicians. Putnam's central contention is that there is no credible account of the first stage. Every attempt to explain in realist terms how a word manages to refer to one object rather than another—that is, every attempt to explain how language "hooks on" to the world—is either plainly unsatisfactory or implies a radical indeterminacy of reference.

Putnam's alternative is to identify truth directly with "ideal acceptability," a position he calls internal realism. The position is realist, not simply because it is compatible with what has here been called minimal realism, but also because it eschews reductionism while remaining compatible with all of the ordinary denials of mind dependence that are part of our scientifically informed worldview. Since it is plainly correct by ordinary standards to assert that mountains exist even when no one is aware of them, the internal realist will agree that mountains do not depend in this literal sense on our thought and are therefore in that sense objective. Still, the view does imply an internal connection at the global level between the way the world is and the way we are disposed to conceive of the world in what Charles Sanders Peirce called "the ideal limit of inquiry." According to the internal realist, we should not say (as the idealist would) that the mind somehow constructs the world but rather that "the mind and the world together make up the mind and the world" (Putnam, 1978).

It remains uncertain whether the efforts of Dummett, Putnam, and others to describe a plausible alternative to realism on the matter of objectivity can succeed. It is to be noted that the arguments they provide indict any epistemically unconstrained notion of truth whatsoever, and hence that if they succeed at all they imply a global antirealism according to which every region of human thought that satisfies the condition of minimal realism is directed at a region of fact that is somehow constituted in part by our thought about it. But this can be rather hard to believe. The difficulty emerges most dramatically when we consider discourse about the past. Most of us are inclined to believe that every (nonvague) question about the past must have an answer. There is a fact of the matter, we suppose, as to whether Genghis Khan was right-handed, even if we cannot in principle obtain any pertinent evidence. But it is likely that neither "Genghis Khan was right-handed" nor its denial is assertible. Any view according to which this implies that the statement is neither true nor false is therefore bound to strike us as initially incredible. Perhaps more important, there is reason to doubt whether a commitment to an epistemically constrained notion of truth always implies a rejection of the realist's rhetoric of objectivity and independence. It is conceivable, for example, that a moral realist for whom the demands of morality are entirely independent of our passions and interests might nonetheless insist that the

moral facts—because they represent rationally compelling demands on human action—must be accessible in principle to human beings. Moral truth would then be epistemically constrained; and yet the realist's rhetoric of objectivity and independence would not be undermined.

To be a realist about a region of discourse is to hold at a minimum that our core commitments in the area are largely true when interpreted "at face value." However, this minimal characterization fails to capture the realist's commitment to the objectivity or mind independence of his subject matter. In some cases this further commitment can be understood as the requirement that the concept of truth appropriate to the target area be epistemically unconstrained. It remains unclear, however, whether this characterization is adequate to every case. The search for a fully general account of the realist's commitment to objectivity is perhaps the central open question in this part of philosophy.

See also Atheism; Dummett, Michael Anthony Eardley; Idealism; Kant, Immanuel; Mackie, John Leslie; Meaning; Metaphysics; Noncognitivism; Peirce, Charles Sanders; Philosophy of Language; Philosophy of Mind; Philosophy of Science, History of; Philosophy of Science, Problems of; Putnam, Hilary; Reference; Tarski, Alfred; Truth; Van Fraassen, Bas.

Bibliography

Blackburn, S. "How to Be an Ethical Anti-Realist." In *Essays in Quasi-Realism.* New York: Oxford University Press, 1993.

Blackburn, S. *Spreading the Word.* Oxford: Clarendon Press, 1984.

Churchland, P., and C. Hooker, eds. *Images of Science.* Chicago: University of Chicago Press, 1985.

Devitt, M. *Realism and Truth.* Cambridge, MA: Blackwell, 1991.

Dummett, M. *The Seas of Language.* Oxford: Clarendon Press, 1993. Especially the essays "What Is a Theory of Meaning II" and "Realism."

Dummett, M. *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press, 1978. Especially the essays "Realism," "Truth," and "The Philosophical Basis of Intuitionistic Logic."

Field, H. *Science without Numbers.* Princeton, NJ: Princeton University Press, 1980.

Geach, P. "Ascriptivism." *Philosophical Review* 69 (1960): 221–225.

Gibbard, A. *Wise Choices, Apt Feelings.* Cambridge, MA: Harvard University Press, 1990.

Hare, R. M. *Freedom and Reason.* Oxford: Clarendon Press, 1963.

Honderich, T., ed. *Morality and Objectivity: A Tribute to J. L. Mackie.* London: Routledge and Kegan Paul, 1985.

Mackie, J. L. *Ethics: Inventing Right and Wrong.* New York: Penguin, 1977.

Putnam, H. *The Many Faces of Realism.* La Salle, IL: Open Court, 1987.

Putnam, H. *Meaning and the Moral Sciences.* Boston: Routledge and Kegan Paul, 1978.

van Fraassen, B. *The Scientific Image.* Oxford: Clarendon Press, 1980.

Wright, C. *Frege's Conception of Numbers as Objects.* Aberdeen: Aberdeen University Press, 1983.

Wright, C. *Truth and Objectivity.* Cambridge, MA: Harvard University Press, 1992.

Gideon Rosen (1996)

REALISM, LEGAL

See *Legal Realism*

REALISM AND NOMINALISM

See *Universals, A Historical Survey*

REALISM AND NATURALISM, MATHEMATICAL

Many versions of realism in mathematics are intimately related to versions of naturalism. The purpose of this article is to explore relationships between the various views, and, briefly, the main opposition to them. The focus here is exclusively on mathematics. So, for example, "Platonism" is to be read as "Platonism about mathematics." This entry does not claim to do justice to the subtle and detailed works of everyone who works in the philosophy of mathematics, or even everyone who defends versions of realism and/or naturalism. Instead, this entry seeks to provide a useful road map of an important part of the territory.

In broad terms, *realism* is the view that mathematics is objective: independent of the lives, customs, language, and form of life of mathematicians. This statement is deliberately indeterminate. What aspects of mathematics are being discussed? What, exactly, is it independent of? And what is it to be independent? What is it to be objective? In philosophy there is little that one can take for granted.

REALISM

There are at least two forms of realism: realism in ontology, which concerns mathematical objects, and realism in truth value, which concerns mathematical truth. *Realism in ontology* is the view that mathematical *objects*, such as numbers, sets, functions, and geometric points exist independently of the mathematician. Prima facie, these mathematical objects do not occupy physical space; they exist eternally and are not created or destroyed; and they do not enter into causal relationships with either each other or with physical objects. Because Platonic forms share these features, realism in ontology is sometimes called "Platonism" or, as Geoffrey Hellman (1989) dubs it, "objects Platonism." This sort of Platonism is sometimes written with a lowercase "p," perhaps to mark some distance from Plato. For the realist in ontology, mathematical propositions are taken at face value, as statements about mathematical objects. The theorem that 101 is a prime number just is the statement that a given object, the number 101, enjoys a certain property, primeness. The sentence "101 is a prime number" has the same logical form as "Socrates is Greek." Most versions of realism in ontology have it that mathematical truth is necessary, in a deep metaphysical sense: If the subject matter of mathematics is as these realists say it is, then typical propositions about mathematical objects—the principles of pure mathematics, for example—do not suffer from the contingencies of science or ordinary statements about ordinary physical objects.

Probably the most difficult problems associated with realism in ontology are in epistemology (see Benacerraf 1973). The realist declares that mathematics is about a realm of prima facie abstract, causally inert, and eternally existing objects. How can human beings ever come to know anything about these objects? How can humans have reliable, justified beliefs about such objects? The way people come to know things about physical objects typically involves some sort of causal contact between people, the knowers, and the objects (e.g., seeing them). This is ruled out with mathematical objects. Presumably, most of the beliefs that mathematicians have about mathematical objects are true. Mathematicians are reliable indicators of how things are in the mathematical realm. How does one explain this reliability (see Field 1989, essay 7)?

One resolution to these problems is to postulate a special faculty that humans have, an intuition, that links humans to the mathematical realm. Such was Plato's own solution to the analogous problem concerning Forms. Some of the logician Kurt Gödel's (1944, 1964) remarks can be interpreted along these lines:

> Despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that axioms force themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception … It should be noted that mathematical intuition need not be conceived of as a faculty giving an *immediate* knowledge of the objects concerned. Rather, it seems that, as in the case of physical experience, we *form* our ideas also of those objects on the basis of something else which is immediately given … It by no means follows … that the data of this … kind, because they cannot be associated with actions of certain things upon our sense organs, are something purely subjective … Rather they … may represent an aspect of objective reality, but, as opposed to the sensations, their presence in us may be due to another kind of relationship between ourselves and reality. (Gödel 1964, p. 484)

A philosopher who is inclined this way has the task of trying to square the presence of mathematical intuition with the current scientific view of a human being as a thoroughly physical organism in a physical universe.

LOGICISM. Another strategy for epistemology comes from *logicism*, the view that mathematical truth is a species of logical truth. The epistemology for mathematics is thus the epistemology for logic. The most detailed developments are those of Gottlob Frege (1884, 1893) and Alfred North Whitehead and Bertrand Russell (1910). Of those, Frege was a realist in ontology, at least for arithmetic and analysis. So for Frege, logic has an ontology—there are "logical objects." Numbers are constructed out of logical objects.

In attempting to define the natural numbers and the general notion of *natural number*, Frege (1884, §63) proposed the following principle, which has become known as "Hume's principle":

> For any concepts *F*, *G*, the number of *F*'s is identical to the number of *G*'s if and only if *F* and *G* are equinumerous.

Two concepts are equinumerous if they can be put in one-to-one correspondence. Frege showed how to define equinumerosity without invoking natural numbers. In the end, he balked at taking Hume's principle as the ulti-

mate foundation for arithmetic, and went on to provide an explicit definition of the natural numbers in terms of concepts and their extensions. The number two, for example, is the extension (or collection) of all concepts that hold of exactly two elements. Unfortunately, the inconsistency in Frege's theory of extensions, as shown by Russell's paradox, marked a tragic end to Frege's logicist program.

Variations of Frege's approach are vigorously pursued in the early twenty-first century, in the work of Crispin Wright, beginning with (1983), and others such as Bob Hale (1987) and Neil Tennant (1997). The idea is to bypass the treatment of extensions and to work with Hume's principle, or something like it, directly. On this *neo-Fregean* approach, Hume's principle is taken to be an explanation of the concept of "number." It is an implicit definition, true by stipulation. Frege's own technical development shows that the Peano postulates can be derived from Hume's principle in a standard, higher-order logic. Indeed, the only essential use that Frege made of extensions was to derive Hume's principle—everything else concerning numbers follows from that.

HYPOTHETICAL-DEDUCTIVE APPROACH. Another popular strategy for epistemology comes from an overarching hypothetical-deductive approach. The argument begins with the observation that virtually all of science is formulated in mathematical terms. One cannot believe in the truth of physics, say, without also accepting the mathematics that occurs in it. Thus, mathematics is confirmed to the extent that science is. In short, because mathematics is indispensable for science, and because science is well-confirmed and (approximately) true, one can conclude that mathematics is well-confirmed and true as well. This "indispensability argument" is attributed to W. W. O. Quine; a clear articulation is found in Hilary Putnam's *Philosophy of Logic* (1971, ch. 5) (see also Colyvan 2001).

STRUCTURALISM. According to *structuralism*, the subject matter of arithmetic, for example, is the pattern common to any infinite system of objects that has a distinguished initial object, which plays the role of zero, and a successor relation or operation that satisfies the induction principle. The arabic numerals exemplify this *natural number structure*, as does an infinite sequence of distinct moments of time, an infinite sequence of discrete points in space, and so on. Similarly, real analysis is about the real number structure, set theory is about the set-theoretic-hierarchy structure, and topology is about topological structures. According to the *ante rem* version of this view, the natural number structure, for example, exists independently of whether it has instances in the physical world, or any other world for that matter (see Shapiro 1997, Resnik 1997, also Parsons 1990). This is an ontological realism. The number six, for example, is a place in the natural number structure, the seventh place (if one begins with zero). Because, on the view in question, the structure exists objectively, then so do its places. Structuralists have proposed various epistemological strategies, ranging from pattern recognition, linguistic abstraction, implicit definition (much like neo-logicism), and postulation via indispensability (with the Quinean). One line, shared with the full-blooded platonism articulated by Mark Balaguer (1998), holds that the realm of structures is so robust that every coherent axiomatization is true of at least one structure. So the sticky problem concerning knowledge of mathematical objects reduces to knowledge of the coherence of an axiomatization.

THE OPPOSITION: ANTIREALISM IN ONTOLOGY. Speaking logically, the opponents of realism in ontology fall into two camps. One group holds that numbers, functions, sets, points, and the like exist, but not objectively. Mathematical objects are not independent of the mind, language, conventions, or the form of life of the mathematician or the mathematical/scientific community. According to traditional intuitionism, for example, mathematical objects are mental constructions (e.g., Brouwer 1912, 1948; Heyting 1956). This is an idealism of sorts. Some intuitionists have explicitly Kantian roots, tying mathematical construction to the forms of pure intuition (typically of time). Another ontological antirealist view sees mathematical objects as social constructions.

The other way to reject ontological realism is to hold that there are no distinctive mathematical objects at all. There simply are no numbers, sets, functions, points, and so on. This is called *nominalism*. Again, it comes in two varieties. On one of them, mathematical assertions keep a straightforward, face-value reading. So the statement that every natural number is prime is vacuously true, because there are no natural numbers. "Seven is prime" is either false or lacks truth-value, depending on how nondenoting singular terms are handled. On this view, mathematical objects are likened to characters and objects in fiction. The sentence that seven is prime is of a piece with "Miss Marple is nosy." Of course, fictionalists do not recommend that mathematicians settle their questions via the literal, face-value reading of their assertions. Either they advert to a "truth in mathematics" akin to "truth in the story" for fiction, or else they provide some other purpose for mathematics beyond seeking mathematical truth (see

Field 1989). Fictionalism is an error-theory about mathematics.

The other variety of nominalism provides alternate, non-face-value readings of mathematics. So the statements of mathematics come out true or false, without presupposing a mathematical ontology. The modal structuralist, for example, reads a statement such as "there are infinitely many prime numbers" as "any exemplification of the natural number structure has infinitely many places, each of which satisfied the property of being prime in that structure." Charles Chihara (1990, 2004) provides versions of various mathematical theories in terms of possible linguistic constructions. One interesting issue concerns the relationship between the "nominalized" assertions and their original counterparts (see Burgess and Rosen 1997).

REALISM IN TRUTH-VALUE. These nominalistic programs lead to another major type of realism concerning mathematics. Georg Kreisel is often cited as suggesting that the important questions in the philosophy of mathematics do not concern the existence of mathematical objects, but rather the objectivity of mathematical assertions. Let us define *realism in truth-value* to be the view that mathematical statements have objective and nonvacuous truth-values independent of the minds, languages, and conventions of mathematicians.

Once again, the opponents to this view logically fall into two categories, depending on what is being denied. The radical opposition holds that mathematical statements have no nonvacuous truth-values at all. The fictionalist, noted above, is the primary and perhaps only occupant of this category. It is difficult to conceive of a projectivism or expressivism concerning mathematics.

The less radical versions of truth-value irrealism allow that mathematical statements have truth-values, but these are not independent of the minds, languages, and conventions of mathematicians. The traditional intuitionists, as described above, fit this bill. Because, for them, mathematical objects are mental constructions, mathematical assertions relate to the activity of construction. Contemporary intuitionists, following Michael Dummett (1977, 1978) also fit this bill, holding that all truths are knowable, on broadly semantic grounds.

Realism in ontology is naturally allied with realism in truth-value. To get from the former to the latter, one just insists that the sentences of mathematics be read literally, at face value. If, for example, "seven" is a genuine singular term, and the sentence "seven is prime" is objectively true, then, it seems, "seven" denotes something, namely, the number seven. And it exists objectively. Conversely, a realist in ontology gets to realism in truth-value by insisting that the typical propositions concerning the interrelations of the mind-independent mathematical objects are themselves objective.

Nevertheless, the connections between these realisms are not forced by logical connections that are obvious to all. As noted, many nominalists are realists in truth-value. They reject the face-value reading of mathematical assertions. At least one prominent philosopher of mathematics goes in the opposite direction. Neil Tennant (1987, 1997) holds that mathematical objects exist objectively, of necessity, and yet he adopts a Dummettian antirealism concerning truth-value.

NATURALISM

Unfortunately, the word "naturalism" has become something of a term of art, and it is hard to find a common theme that underlies every view that goes by that name. Perhaps most of them share a certain deference to the natural sciences. Quine characterizes naturalism as "the abandonment of first philosophy" and "the recognition that it is within science itself … that reality is to be identified and described" (Quine 1981, p. 72; see also 1969). The idea is to see philosophy as continuous with the sciences, not prior to them in any epistemological or foundational sense.

QUINEAN NATURALISM. The naturalist accepts the existence of the theoretical entities, such as forces and electrons, that occur in the most up-to-date scientific theories. Current science describes the world in such terms, and it runs against the theme of naturalism to reject them on philosophical grounds, adopting some sort of instrumentalism or constructive empiricism. When it comes to mathematics, however, naturalists differ. As seen with the aforementioned indispensability argument, Quine himself accepts mathematics to the extent—but only to the extent—that it is needed in science. It is impossible to do physics, or just about any other science for that matter, without invoking real analysis. So the theorems of real analysis are confirmed to the extent that the various scientific theories are confirmed, and these theories are the best ones available. So Quine accepts the truth of real analysis. Moreover, some of the traditional, Platonic themes have naturalistic counterparts. For example, the eternity of mathematical objects corresponds to the fact that mathematical assertions are not inflected with tense.

Naturalized epistemology is the application of Quinean naturalism to the study of knowledge. The

philosopher sees the human knower as a thoroughly natural being within the physical universe. Any faculty that the philosopher invokes to explain knowledge must involve only natural processes amenable to ordinary scientific scrutiny.

This theme exacerbates the epistemic problems with realism. Platonic apprehension of a detached mathematical universe is ruled out from the start, as a nonnatural process. The challenge to the ontological realist is to show how a physical being in a physical universe can come to know about *abstracta* such as mathematical objects. There may be no refutation of realism in ontology, but there is a deep challenge to it. The advocate of indispensability cites the role of mathematics in science. The idea is that mathematics is known the same way that science is. However, it is not enough to leave it at that. The advocate of realism in ontology should delimit the exact role that mathematics plays in science. How, for example, is it possible for a casually isolated realm of abstracta to shed light on the interactions of physical matter? An answer to this would go a long way toward solving the epistemological puzzles.

Notice that, at best, the indispensability argument delivers the *truth* of the principles of real analysis. If one assumes that science is objective, then there is realism in truth-value. It is not clear that the Quinean naturalist is also committed to realism in ontology, despite Quine's own tendencies in that direction. This depends on whether naturalism requires the philosopher to accept the pronouncements of mathematical science at face value. Quine famously calls for regimentation of ordinary and scientific discourse, to clean up the ontological commitments. One can see some of the aforementioned nominalistic programs in this spirit. Some of them show (or try to show) how mathematics can be true without presupposing the existence of distinctively mathematical objects (Hellman 1989). And this truth is all that is needed in science, or so the argument goes.

Other nominalists take issue with the indispensability argument itself. They show how science could proceed without mathematics, or at least without mathematics as it is standardly understood (Field 1980, Chihara 2004). This is also perhaps in the spirit of naturalism.

Quine's own realism extends to real analysis, functional analysis, and perhaps a bit more. But it stops there. Quine does not accept the truth of the higher reaches of set theory unless and until it finds application in science. In fact, Quine goes so far as to recommend the adoption of a restrictive axiom in set theory (V=L), because it simplifies higher-set theory, noting that simplicity is a criterion of theory acceptance in science. This is despite most set-theorists' rejection of this axiom. It is ironic that Quine, the naturalist, feels comfortable dictating something to mathematicians on philosophical grounds.

OTHER VERSIONS OF NATURALISM. Penelope Maddy's (1997) and John Burgess's and Gideon Rosen's (1997) versions of naturalism defend a deferential attitude towards mathematics much like the one Quine shows toward science. They note, first, that mathematics has its own methodology, distinct from so-called scientific method, and that this methodology has proven successful over the centuries. The success of mathematics is measured in mathematical, not scientific terms. Moreover, if mathematicians gave serious pursuit only to those branches known to have applications in natural science, much of the mathematics known in the twenty-first century would not exist, nor would the *science*. The history of science is full of cases where branches of pure mathematics eventually found application in science (see Steiner 1997). That is to say, the overall goals of the scientific enterprise have been well-served by mathematicians pursuing their own disciplines with their own methodology, ignoring science if necessary. Thus, one does not need a direct inferential link between a piece of mathematics and sensory experience before accepting the mathematics as a legitimate part of the web.

On general naturalistic grounds, Burgess and Rosen adopt a realism in ontology for mathematics. For them, the convenience of the face value reading of mathematical propositions counts in its favor. Someone who proposes a nominalistic reconstruction must defend their account on accepted scientific, or mathematical grounds. That is, they must show that the ontology-free versions of mathematics are better mathematics and/or better science. Foregoing philosophical puzzles concerning epistemology do not count. Maddy is more circumspect, arguing that naturalism does not demand a realist interpretation of mathematics.

The varieties of naturalism treated here might be dubbed *methodological* because they focus on the methods of science, adopting those to traditional philosophical questions. Nominalism, as construed here, is an expression of another, ontological variety of naturalism. The thesis is that the only things that exist are the material objects of science, and the only properties people need to consider are the material properties of those objects. Alternately, the only objects in which people are licensed to believe are those with which they causally interact. Mark Colyvan (2001, ch. 3) calls this the eleatic

principle. Another issue that separates naturalists—or at least philosophies that go by that name—is whether all legitimate knowledge is empirical. In the spirit of radical empiricism, in the manner of John Stuart Mill, Quine has launched a sustained attack on a priori knowledge. Not every contemporary naturalist follows suit. Bernard Linksy and Edward Zalta (1995) argue that the proper interpretation of science requires a more traditional Platonism, according to which mathematical propositions are synthetic a priori. Clearly, an article such as this can do no more than scratch the surface of these rich and wonderful topics.

See also Mathematics, Foundations of; Nominalism, Modern.

Bibliography

Balaguer, M. *Platonism and Anti-Platonism in Mathematics*. Oxford: Oxford University Press, 1998. Account of realism in ontology and its rivals.

Benacerraf, P. "What Numbers Could Not Be." *Philosophical Review* 74 (1965): 47–73. Reprinted in Benacerraf and Putnam, *Philosophy of Mathematics*, 272–294. One of the most widely cited works in the field; argues that numbers are not objects, and introduces an eliminative structuralism.

Benacerraf, P. "Mathematical Truth." *Journal of Philosophy* 70 (1973): 661–679. Reprinted in Benacerraf and Putnam, *Philosophy of Mathematics*, 403–420. Another widely cited work; argues that realism in ontology has formidable epistemological problems.

Benacerraf, P., and H. Putnam, eds. *Philosophy of Mathematics*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983. A far-reaching collection containing many of the central articles.

Brouwer, L. E. J. "Consciousness, Philosophy and Mathematics." 1948. In *Philosophy of Mathematics: Selected Readings*. 2nd ed., edited by P. Benacerraf and H. Putnam, 90–96. New York: Cambridge University Press, 1983.

Brouwer, L. E. J. *Intuitionisme en Formalisme*. Groningen, Germany: Noordhof, 1912. Translated as "Intuitionism and Formalism." In *Philosophy of Mathematics: Selected Readings*. 2nd ed., edited by P. Benacerraf and H. Putnam, 77–79. New York: Cambridge University Press, 1983.

Burgess, J. "Why I Am Not a Nominalist." *Notre Dame Journal of Formal Logic* 24 (1983): 93–105. Early critique of nominalism.

Burgess, J., and G. Rosen. *A Subject With No Object: Strategies for Nominalistic Interpretation of Mathematics*. Oxford: Oxford University Press, 1997. Extensive articulation and criticism of nominalism.

Chihara, C. *Constructibility and Mathematical Existence*. Oxford: Oxford University Press, 1990. Defense of a modal view of mathematics, and sharp criticisms of several competing views.

Chihara, C. *A Structural Account of Mathematics*. Oxford: Oxford University Press, 2004. Sequel to Chihara, *Constructibility and Mathematical Existence*, with particular focus on the application of mathematics.

Colyvan, M. *The Indispensability of Mathematics*. Oxford: Oxford University Press, 2001. Elaboration and defense of the indispensability argument for ontological realism.

Dummett, M. *Elements of Intuitionism*. Oxford: Oxford University Press, 1977.

Dummett, M. "The Philosophical Basis of Intuitionistic Logic." In *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, 1978. Reprinted in Benacerraf and Putnam, *Philosophy of Mathematics*, 97–129; in Hart, *The Philosophy of Mathematics*, 63–94. Influential defense of intuitionism.

Dummett, M. *Elements of Intuitionism*. Oxford: Oxford University Press, 1977. Detailed introduction and defense of intuitionistic mathematics.

Field, H. *Science Without Numbers*. Princeton, NJ: Princeton University Press, 1980. A widely cited defense of fictionalism, by attempting to refute the indispensability argument.

Field, H. *Realism, Mathematics and Modality*. Oxford: Blackwell, 1989. Reprints of Field's articles on fictionalism.

Frege, G. *Die Grundlagen der Arithmetik*. Breslau, Germany: Koebner, 1884. Translated by J. Austin as *The Foundations of Arithmetic*. 2nd ed. New York: Harper, 1960. Classic articulation and defense of logicism.

Frege, G. *Grundgesetze der Arithmetik*. Vol. 1. Hildescheim, Germany: Olms, 1893. More technical development of Frege's logicism.

Gödel, K. "Russell's Mathematical Logic." 1944. In *Philosophy of Mathematics*, edited by P. Benacerraf and H. Putnam, 447–469. New York: Cambridge University Press, 1983. Much-cited defense of realism in ontology and realism in truth value.

Gödel, K. "What is Cantor's Continuum Problem." 1964. In *Philosophy of Mathematics*, edited by P. Benacerraf and H. Putnam, 470–485. New York: Cambridge University Press, 1983. Much-cited defense of realism in ontology and realism in truth value.

Hale, Bob. *Abstract Objects*. Oxford: Basil Blackwell, 1987. Detailed development of neo-logicism, to support Wright (1983).

Hart, W. D., ed. *The Philosophy of Mathematics*. Oxford: Oxford University Press, 1996. Collection of articles published elsewhere.

Hellman, G. *Mathematics Without Numbers*. Oxford: Oxford University Press, 1989. Articulation and defense of modal structuralism.

Heyting, A. *Intuitionism: An Introduction*. Amsterdam: North Holland, 1956. Readable account of intuitionism.

Hodes, H. "Logicism and the Ontological Commitments of Arithmetic." *Journal of Philosophy* 81 (1984): 123–149. Another roughly Fregean logicism.

Linksy, B., and E. Zalta. "Naturalized Platonism versus Platonized Naturalism." *Journal of Philosophy* 92 (1995): 525–555.

Maddy, P. *Naturalism in Mathematics*. Oxford: Oxford University Press, 1997. Lucid account of naturalism concerning mathematics, and its relation to traditional philosophical issues.

Maddy, P. *Realism in Mathematics*. Oxford: Oxford University Press, 1990. Articulation and defense of realism about sets.

Maddy, P. "Three Forms of Naturalism." In Shapiro, *Oxford Handbook of Philosophy of Mathematics and Logic*, 437–459. Contrast with Quinean naturalism.

Parsons, C. "The Structuralist View of Mathematical Objects." *Synthese* 84 (1990): 303–346. Reprinted in *The Philosophy of Mathematics*, edited by W. D. Hart, 272–309. Oxford: Oxford University Press.

Putnam, H. *Philosophy of Logic*. New York: Harper Torchbooks, 1971. Source for the indispensability argument for ontological realism.

Quine, W. V. O. *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Quine, W. V. O. *Theories and Things*. Cambridge, MA: Harvard University Press, 1981.

Resnik, M. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press, 1997. Full articulation of a realist-style structuralism.

Schirn, M., ed. *Philosophy of Mathematics Today*. Oxford: Oxford University Press, 1998. Proceedings of a conference in the philosophy of mathematics, held in Munich in 1993; coverage of most of the topical issues.

Shapiro, S., ed. *Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press, 2005. Up-to-date articles covering the discipline.

Shapiro, S. "Philosophy of Mathematics." In *Philosophy of Science Today*, edited by Peter Clark and Katherine Hawley, 181–200. Oxford: Oxford University Press, 2003.

Shapiro, S. *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press, 1997. Extensive articulation and defense of structuralism.

Shapiro, S. *Thinking about Mathematics: The Philosophy of Mathematics*. Oxford: Oxford University Press, 2000. Popularization and textbook in the philosophy of mathematics.

Steiner, M. *The Applicability of Mathematics as a Philosophical Problem*. Cambridge, MA: Harvard University Press, 1997.

Tennant, N. *Anti-Realism and Logic*. Oxford: Oxford University Press, 1987. Articulation of antirealism in truth value, realism in ontology; defends intuitionistic relevance logic against classical logic.

Tennant, N. *The Taming of the True*. Oxford: Oxford University Press, 1997. Detailed defense of global semantic antirealism.

Weir, A. "Naturalism reconsidered." In *Oxford Handbook of Philosophy of Mathematics and Logic*, edited by S. Shapiro, 460–482. Oxford: Oxford University Press.

Whitehead, A. N., and B. Russell. *Principia Mathematica*. Vol. 1. Cambridge, U.K.: Cambridge University Press, 1910.

Wright, C. *Frege's Conception of Numbers as Objects*. Aberdeen, Scotland: Aberdeen University Press, 1983. Revival of Fregean logicism.

Stewart Shapiro (2005)

REALITY

See *Appearance and Reality; Being*

REASON

In English the word *reason* has long had, and still has, a large number and a wide variety of senses and uses, related to one another in ways that are often complicated and often not clear. However, there is one particular sense of the word in which it, with its synonyms or analogues in other languages, has figured prominently in philosophical controversy. This is the sense, sometimes distinguished typographically by an initial capital, in which the term is taken to designate a mental faculty or capacity—in which reason might, for example, be regarded as coordinate with, but distinguishable from, sensation, emotion, or will.

QUESTIONS TO BE EXAMINED

The question that has been chiefly debated by philosophers might be expressed succinctly, but far from clearly, as "What can reason do?" However, there has also been discussion of the question whether the faculty of reason is peculiar to humanity (and presumably to "higher" beings, if there are any), or whether its possession and exercise in some degree can also be ascribed to "lower" animals. It should perhaps be added that in recent years there has been much debate as to whether machines can, or in principle ever could, properly be said to think; for if an affirmative answer were to be given to this question, then there is a quite common sense of *reason* in which it would follow that that faculty could be exercised by a machine. Only the first of these questions is dealt with here.

The short but unclear question "What can reason do?" is peculiarly liable to give rise to theoretical dissension. The question may, however, be transformed with advantage into a question not directly about the "faculty" of reason itself but about those beings to whom this faculty is attributed. What, we may ask, are human beings in a position to do, in virtue of their possession of the faculty of reason? What, by means of reasoning, are we in a position to achieve? In this form it becomes very clear that the question raises at least two highly disputable issues. First, it is far from immediately clear what reasoning is—on what occasions, in what activities or processes, reason is exercised. And second, if we determine—probably with some degree of arbitrariness—what reasoning is, it may very well remain highly disputable whether this or that can or cannot be achieved by reasoning. One should, indeed, distinguish further at this point between two radically different kinds of dispute that may arise; if it were held that, for instance, knowledge of God cannot be

attained by reasoning, there would plainly be an important further distinction between holding this to be true in fact and true in principle. It might be maintained that the reasoning necessary for knowledge of God is, as a matter of fact, too difficult for frail and mortal human beings to manage; or it might be maintained, quite differently, that the kind of conclusion capable of being established by reasoning excludes in principle that kind, if there is any such, to which knowledge of God must belong. This sort of distinction can be seen as differentiating the positivism preached by Comte in the nineteenth century from the logical positivism of recent philosophy.

MANY SENSES OF *REASON*

What, then, is reason? Alternatively, what is reasoning? It seems scarcely possible to maintain that these questions can be given definite answers. The definitions, implicit or explicit, of the relevant terms that have been employed by philosophers and other writers vary widely and significantly; and while some may be judged preferable to others, or may adhere more closely than others to senses which the terms may bear in ordinary discourse, there seems to be no basis secure enough to support a pronouncement that a particular meaning, and hence a particular answer to the question, is exclusively correct. In any case, what is important to the understanding of philosophical writing on this topic is not that one should know what *reason* means but, rather, that one should discern, so far as possible, what meaning is attached to *reason* by an author.

CONTRASTS WITH OTHER TERMS

Here it seems particularly important and helpful to consider with what reason is contrasted, or from what it is distinguished. There is, for example, a large body of literature in which reason stands essentially in contrast with faith. In this context, what we can achieve by reason is taken to embrace the entire field of knowledge and inquiry in which, with varying degrees of skill and success, we produce or seek reasons for our views, proofs of or evidence for our conclusions, and grounds for our opinions. This whole field is set in contrast with another, in which supposedly we may—or should or must—accept certain propositions or doctrines without any grounds but rather on authority or perhaps on unreasoned conviction.

There is another large body of literature in which reason stands in contrast with experience. In this context, what we can achieve by reason is much more narrowly circumscribed; here a distinction is being made between, roughly, what we can discover or establish by merely sitting and thinking, and what we can discover or establish only by the use of our senses, by observation or by experiment. It will be observed that there are, corresponding to these wider and narrower senses of *reason*, also wider and narrower senses of the term *rationalist*; a rationalist in the one sense is concerned with denying or belittling the claims or the role of faith, and in the other with denying or belittling the role, in the acquisition of knowledge, of experience. There is no particular reason why one who is a rationalist in either one of these senses should be expected to be a rationalist in the other sense also; the two positions are quite independent of one another.

THE OBJECTS OF REASON

There is, then, no universally agreed or uniquely correct sense of *reason*. This is obvious enough, perhaps; but it is not unimportant. Clearly, even though philosophers may use this term in diverse senses without being wrong, the fact that they do so must, if unobserved by them or their readers, generate confusion and argument at cross purposes. Further, as was noted above, even if we avoid confusion at this point, many problems as to the "scope" or the "powers" of reason remain. They are, in fact, some of the major and central problems of philosophy.

Suppose that, following Brand Blanshard in his *Reason and Analysis*, we define *reason* as "the faculty and function of grasping necessary connections." We may feel that this is not a very good definition, since it seems excessively restrictive. For example, a judge arguing his way to a decision, or a meteorologist setting forth his grounds for a weather forecast, would in this sense not be exercising the faculty of reason; the argument in each case is nondemonstrative—that is, it does not set out or rely on strictly necessary connections. However, waiving that point, the definition is at least a clear one. But notwithstanding its possession of the important virtue of clarity, the question of what reason can do is not thereby settled.

In order to settle this question, we must decide what necessary connections there are and in what cases or what fields there are necessary connections to be grasped; and the determination of this question raises, or might very well raise, almost every problem of philosophy. Are we to hold, with Plato, that no necessary connections are to be discerned in the everyday world, but only in an intelligible world of Forms? Or are we to hold, with David Hume and many others, that strictly necessary connections are to be found only in the formal, abstract relations between our concepts or ideas? Was Immanuel Kant right in supposing that the moral law can be demonstrated a priori,

and is therefore necessary? Or, on the contrary, was Hume correct in holding that in the field of moral judgment "reason is the slave of the passions"? Are causal relationships cases of necessary connection? Are they perhaps, as John Locke seems to have held, really cases of necessary connection that in practice, however, we are inveterately unable to grasp as such? And so on.

BASIC QUESTIONS

The point that emerges here is simply this: Whatever particular definition of the faculty of reason we may, implicitly or explicitly, adopt, it seems unavoidable that it will be attempted thereby to distinguish this faculty from others as being that by the exercise of which we can perceive, or arrive at, truths of some particular kind or kinds; and this kind of truth, or these kinds of truths, will in turn be distinguished from other kinds on logical or epistemological grounds. If so, then the question of what we can actually achieve or come to know by reason unavoidably becomes the question of what propositions are of that kind or those kinds; and this is precisely the question about which, in any field, philosophical controversy may, and characteristically does, arise. Thus the apparently simple question "What can reason do?" is not a neutral question on which otherwise dissentient philosophers may expect to be in agreement. On the contrary, it is very likely that their disagreement consists precisely in their diverse answers to this question. It may further be felt, with justice, that if this innocent-looking question unavoidably raises major philosophical issues concerning the logical and epistemological analysis and classification of propositions, it would probably be advantageous to raise those questions directly and overtly rather than as an only half-acknowledged corollary of a discussion that is ostensibly concerned with a faculty of the mind. There are few modern philosophers who would naturally cast their discussions in this latter idiom.

One final risk of confusion is worth pointing out. It is probably true that in recent philosophy there has been a persistent tendency to narrow the field in which necessities, strictly speaking, are admitted to be found; and also, perhaps more significantly, a persistent tendency to take the awesomeness out of necessity by attempts, more or less successful in various fields, to exhibit necessity as fundamentally derived from the unpuzzling, and perhaps unimposing, phenomenon of tautology. In this sense, then, it can be said that there has been some tendency both to narrow the scope conceded to reason and perhaps also to make reason itself seem less mysterious and grand. In some, this tendency has occasioned considerable distress: As Bertrand Russell has expressed it, "My intellectual journeys have been, in some respects, disappointing.… I thought of mathematics with reverence, and suffered when [Ludwig] Wittgenstein led me to regard it as nothing but tautologies" (*The Philosophy of Bertrand Russell*, edited by P. A. Schilpp, Evanston, IL, 1946, p. 19).

EXAMINATION OF REASON'S POWERS

There are several instances in which Russell's sense of distress has been expressed in curiously bellicose terms. Books have been written in defense of reason, and exponents of the contemporary trend have been castigated as reason's enemies. But this latter charge, even if there is some sense in which it might be well founded, is peculiarly liable to mislead, and very commonly has misled, those who urge it. One thinks, naturally and rightly, of an enemy of reason as one who is opposed or hostile to the exercise of reason. Such a person might be, for instance, a religious bigot, fearful that reason might shake the obscure foundations of his bigotry; he might be a political or racial fanatic, hostile to the careful weighing of arguments and evidence because he is half conscious that his program or doctrine lacks reasonable grounds; or he might, less malignantly, hold some doctrine about the merits of unreflecting spontaneity, disliking the slow pace, the qualifications and hedging, of rational thought. It is obvious, however, that scarcely any philosopher is, or ever has been, an enemy of reason in this sense.

Nor, to mention a group not uncommonly arraigned on the same charge, is the psychoanalyst. It is a tenet of psychoanalytic theory that reason, the dispassionate consideration of arguments and evidence, is a less conspicuous and influential determinant of the beliefs and the conduct of men than has often been supposed, or than most people might like to admit; but the psychoanalyst does not, as would an enemy of reason, rejoice in this circumstance or seek to aggravate it. Quite the contrary: Recognizing the state of the case as being what, in the light of his evidence, he takes it to be, he deploys his art in the attempt to enable people to become more rational than they would otherwise be. He may be mistaken in his theory and unsuccessful in his practice, but in any case neither in theory nor in practice does he display the least enmity toward reason.

Somewhat similarly, the philosopher who produces an argument against high traditional claims for, or traditional characterizations of, reason is, in so doing, exercising reason to the best of his ability; nor does it occur to him to question the desirability of doing so. Thus, to dissent from rationalism as a philosophical doctrine is cer-

tainly not to disparage reason; the man who values, and shows that he values, reason is not he who merely pitches reason's claims exceptionally high but, rather, he who attempts, by painstaking reasoning, to determine how high those claims may justifiably be pitched. Philosophers, whose work consists mostly in sitting and thinking, have often enough and naturally enough been prone to estimate very highly the range and significance of the results that can thereby be achieved. However, this propensity is scarcely an indication of devotion to reason; rather, it is an indication, if of anything, of pardonable self-importance.

See also Blanshard, Brand; Comte, Auguste; Faith; Hume, David; Locke, John; Logical Positivism; Plato; Positivism; Practical Reason; Rationalism; Russell, Bertrand Arthur William; Thinking; Wittgenstein, Ludwig Josef Johann.

Bibliography

Bermudez, Jose Luis, and Alan Millar, eds. *Reason and Nature: Essays in the Theory of Rationality*. Oxford: Clarendon Press, 2002.

Blanshard, Brand. *Reason and Analysis*. London: Allen and Unwin, 1962.

BonJour, Laurence. *In Defense of Pure Reason*. New York: Cambridge University Press, 1997.

Brewer, Bill. *Perception and Reason*. Oxford: Clarendon Press, 2002.

Chisholm, Roderick. *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice-Hall, 1966; 2nd ed., 1977; 3rd ed., 1989.

Ewing, A. C. *Reason and Intuition*. London: H. Milford, 1942.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Freud, Sigmund. *The Future of an Illusion*. Translated by W. D. Robson-Scott. London: Hogarth Press, 1928.

Friedman, Michael. "Philosophy as Dynamic Reason: The Idea of a Scientific Philosophy." In *What Philosophy Is: Contemporary Philosophy in Action*, edited by Havi Carel. London: Continuum, 2004.

Howson, Colin, and Peter Urbach. *Scientific Reasoning: The Bayesian Approach.* 2nd ed. Chicago: Open Court, 1993.

Moser, Paul K., ed. *Rationality in Action*. Cambridge, U.K.: Cambridge University Press, 1990.

Murphy, Arthur E. *The Uses of Reason*. New York: Macmillan, 1943.

Nagel, Ernest. *Sovereign Reason*. Glencoe, IL: Free Press, 1954.

Papineau, David. *The Roots of Reason: Philosophical Essays on Rationality, Evolution, and Probability*. Oxford: Clarendon Press, 2003.

Peacocke, Christopher. *The Realm of Reason*. Oxford: Clarendon Press, 2004.

Russell, Bertrand. *Skeptical Essays*. New York: Norton, 1928.

Santayana, George. *The Life of Reason*, rev. ed. New York: Scribners, 1954.

Stich, Stephen. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Walsh, W. H. *Reason and Experience*. Oxford: Clarendon Press, 1947.

Whitehead, A. N. *The Function of Reason*. Princeton, NJ: Princeton University Press, 1929.

G. J. Warnock (1967)
Bibliography updated by Benjamin Fiedor (2005)

REASONING

See *Thinking*

REBIRTH

See *Reincarnation*

RECURSION THEORY

See *Computability Theory*

REDUCIBILITY, AXIOM OF

See *Russell, Bertrand (section on logic and mathematics)*

REDUCTION

A cursory glance at the history of science reveals a continuous succession of scientific theories of various areas or domains. For example, since ancient times theories of the cosmos have been proposed to account for the observed behavior of the heavenly bodies. The geocentric Ptolemaic theory was, for instance, succeeded by the heliocentric theory of Copernicus. Another example concerns the nature of light. Corpuscular theories were succeeded by wave theories of light. Wave theories, in turn, have been followed by the quantum theories of electromagnetic radiation.

This entry concerns the nature of certain relations that may obtain between different pairs of theories in such sequences. A radical or extreme view of those relations is that of Thomas Kuhn. Kuhn (1970) famously argues that across scientific revolutions there is a radical disconnect between theories. One can find a similar argument in Paul K. Feyerabend (1962). On such a view, no

rational relations can obtain between a theory and its predecessor. However, it is fair to say that most philosophers have held, contrary to this extreme position, that there are, indeed, interesting and contentful relations between various pairs of scientific theories. One such relationship is that of reduction. It is often claimed that successor theories reduce those that they succeed. Such a relation may involve the idea that the successor or reducing theory explains or otherwise absorbs the successful features of the reduced theory. However, getting clear about exactly how the notion of reduction should be understood has been and continues to be a difficult philosophical problem.

This entry begins with a discussion of what may be called the received view of theory reduction and examines how that view has evolved as the result of various criticisms. Work on intertheoretic relations from 1997 through 2005 is then considered.

NAGELIAN REDUCTION

The *locus classicus* for contemporary discussion is Ernest Nagel's presentation of a model for theory reduction in *The Structure of Science* (1961). Nagel takes reduction to be an explanatory relation between theories where explanation is understood to involve deductive logical relations between statements characterizing the explanans and the statement characterizing the explanandum in accordance with the Hempelian (Hempel 1965) deductive-nomological model. Nagel holds that "[r]eduction … is the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory usually though not invariably formulated for some other domain" (1961, p. 338). The idea here is that a theory T reduces a theory T' just in case one can derive (and thereby explain) the laws of T' from the laws of T.

Nagel realizes that for some intuitive cases of theory reduction such derivations would not be immediately possible. If the vocabulary of the reduced (succeeded) theory contains terms referring to entities or properties that are not mentioned in the vocabulary of the reducing (successor) theory, then it will be impossible to derive the laws of the reduced theory containing those terms from the laws of the reducing theory. Reductions involving theories with distinct vocabularies are called heterogeneous by Nagel. By contrast, homogeneous reductions are taken by him to be rather straightforward and unproblematic.

This view of homogeneous reductions is somewhat naive. Lawrence Sklar (1967) points out that homogeneous reductions, in fact, are rare. Instead, what one has typically is the derivation of an approximation to the reduced theory and not of the reduced theory itself. An example discussed by both Nagel and Sklar concerns the homogeneous reduction of the Galilean theory of free fall to Newtonian mechanics and gravitational theory. Sklar notes that there really is no strict derivation of the Galilean theory, although no terms appear in the Galilean theory that do not also appear in Newton's theory.

The example of a heterogeneous reduction Nagel discusses is the apparent reduction of thermodynamics to statistical mechanics. This example has become paradigmatic of intertheoretic reduction in the general philosophical literature. (In actual fact, the reduction of thermodynamics to statistical mechanics is much more complex than Nagel's discussion allows. Sklar [1993] provides a detailed discussion of various difficulties involved in the reduction of thermodynamics to statistical mechanics.) Thermodynamics contains terms referring to properties such as temperature and entropy. Such terms are completely lacking in the vocabulary of statistical mechanics. To effect the (supposed) derivational reduction, one must connect these thermodynamic terms with terms occurring in the vocabulary of statistical mechanics.

Nagel introduces two necessary formal conditions for such heterogeneous reductions:

- Connectability: "Assumptions of some kind must be introduced which postulate suitable relations between whatever is signified by 'A' [a term appearing in the reduced but not the reducing theory, such as 'temperature'] and traits represented by theoretical terms already present in the primary [reducing] science."
- *Derivability*: "With the help of these additional assumptions, all the laws of the secondary [reduced] science, including those containing the term 'A,' must be logically derivable from the theoretical premises and their associated coordinating definitions in the primary [reducing] discipline." (1961, pp. 353–354)

The connectability requirement is vague as it stands. What is the exact nature of the required "suitable relations"? In the literature such relations of connectability are typically called bridge laws or bridging hypotheses and their status is a matter of debate. Nagel allows that such bridge laws need not have the form of universally quantified biconditionals for theory reduction to be possible. They might, he holds, have the form of one-way conditionals. It is this possibility that renders the requirement of derivability not superfluous (Nagel 1961, p. 355

note). With the aid of bridge laws, Nagel thinks that the reducing theory would be able to fully explain the laws of the reduced theory.

However, even having universal biconditionals as bridge laws may not itself be sufficient for reduction. Many examples exist where correlatory laws may be established—where the biconditionals are true and apparently lawlike—yet, where nothing resembling reduction can take place. Sklar (1967) offers the example of the Wiedemann-Franz law expressing a correlation between the thermal conductive properties of a material and its electrical conductivity properties. Such a law does not allow one to reduce the theory of thermal conductive properties of the material to a theory of its electrical conductive properties. Something more than mere correlation is required.

That something more is usually taken to be some kind of empirically established identity claim. For example, the reduction of physical optics to the theory of electromagnetic radiation is accomplished by noting the identity of one class of entities—light waves—with (part of) another class— electromagnetic radiation. As Sklar notes, "Light waves are not correlated with electromagnetic waves, for they are electromagnetic waves. There are not two classes of entities, but only one" (1967, p.120). Another classic example is the reduction of Mendelian genetics to molecular genetics via the identification of genes with DNA molecules.

The idea that the bridge laws must express necessary identifications between entities or classes of entities has much to recommend it. However, in many cases of apparent intertheoretic reduction such identity relations are not available. In the paradigmatic case of the reduction of thermodynamics to statistical mechanics, one sees that terms such as *temperature* and *entropy*, occurring in thermodynamics but not in statistical mechanics, refer to properties possessed by thermodynamic systems. Still, it is not at all clear what properties of statistical systems can be identified with the thermodynamic properties. For example, the standard claim that temperature is just (identical to) mean molecular kinetic energy is deeply problematic. Again, see Sklar (1993) for a detailed discussion of some of these problems.

One way of emphasizing the difficulty here is in terms of questions about the meaning of terms appearing in the distinct theories. In orthodox thermodynamics, for example, the term *entropy* gets its meaning (on one view of how theoretical terms acquire meaning) at least in part by the role the term plays in the theory. (A classic presentation of orthodox thermodynamics explicitly exhibiting the roles of the terms is by A. B. Pippard [1957].) One sees that such terms refer to unvarying and nonstatistical properties of systems. Nevertheless, in the apparent reduction of thermodynamics to statistical mechanics the concept of entropy changes to one that explicitly allows for statistical variation and fluctuation. In what sense can one identify here? Feyerabend (1962), for one, takes this to be evidence that reduction (understood as Nagelian derivation with bridge laws) must fail.

NEO-NAGELIAN REDUCTION. In contrast to Feyerabend's (1962) pessimistic conclusion many philosophers hold that some sort of reductive relation still obtains even in the face of problems of heterogeneity. In fact, it is often noted that in the process of reducing one theory to another, the reduced theory gets emended. One sees textbooks with titles referring to statistical thermodynamics, indicating that the orthodox thermodynamic conceptions of entropy and temperature have been changed to allow for (observable and observed) fluctuations in those quantities. The explicit recognition that the reduced theory is often changed as a result of reduction or attempted reduction takes one beyond the Nagelian conception of reduction as a relatively straightforward explanatory derivation.

Kenneth Schaffner's (1967, 1976) model of reduction deserves mention here as a sophisticated attempt to incorporate this aspect of theoretical change into a Nagelian-type framework. Schaffner explicitly includes the corrected reduced theory in the model. On this view a theory T reduces a theory T' just in case there is a corrected version of the reduced theory, T'^* such that

(1) The primitive terms of T'^* are associated with various terms of T via bridge laws or reduction functions

(2) T'^* is derivable from T when supplemented by these bridge laws

(3) T'^* corrects T in that it makes more accurate predictions than does T'

(4) T' is explained by T in that T' and T'^* are strongly analogous to one another, and T says why T' works as well as it does in its domain of validity.

It is clear that work must be done to explicate the intuitive notion of strong analogy playing a role in this model of reduction. See William C. Wimsatt (1976) for some suggestions along these lines.

OBJECTIONS TO NAGELIAN REDUCTIONS

A number of influential objections have been raised against Nagelian models of reduction. Most of these concern the possibility of providing the appropriate bridge laws. As a result they can be seen as telling also against more sophisticated models such as Schaffner's (1967, 1976). Additionally, it has been objected that even if such bridge laws can be provided, there remains an explanatory question about their status as laws. Consider the second objection first.

EXPLANATORY QUESTIONS ABOUT BRIDGE LAWS. In those cases where bridge laws express the identification of classes of entities, to ask why those bridge laws hold is to ask a question that can be trivially answered. The reason the bridge laws hold is because the entities in question are one and the same. "Why should I believe that light waves are electromagnetic radiation?" Answer: "They just are. Period, end of story." By contrast, in cases where bridge laws express some kind of (perhaps, nomologically) necessary coextensivity between properties appearing in two theories, such a question may seem legitimate and answers may be hard to come by. Jaegwon Kim (1998) forcefully argues that this poses a serious problem for Nagelian reduction understood as attempting to effect an explanatory relation among pairs of theories.

Kim discusses the attempted Nagelian reduction of psychology (the science of the mental) to physical theory, say, neurophysiology. One can suppose that one discovers empirically a nomological correlation between being in pain and having one's C-fibers firing. A statement characterizing this correlation is taken to be a bridge law necessary for Nagelian reduction. In this case it seems reasonable to ask: "Can we understand why we experience pain when our C-fibers are firing, and not when our A-fibers are firing? Can we explain why pains, not itches or tickles, correlate with C-fiber firings?" (Kim 1998, p. 95). Kim's point is that if Nagelian reduction is supposed to provide an explanation of the reduced theory in terms of the reducing theory, then surely one must demand an explanation of the bridge laws employed in the explanatory derivation. "*For it is the explanation of these bridge laws, an explanation of why there are just these mind-body correlations, that is at the heart of the demand for an explanation of mentality*" (p. 96, emphasis in the original).

MULTIPLE REALIZABILITY. A different argument due to Jerry Fodor (1974) has been used to block attempts at almost every Nagelian reduction of a given (special science) theory to more basic (physical) theory. This argument has come to be called the multiple realization argument. It depends on the assumption that properties appearing in the special (to-be-reduced) science may have diverse and "wildly heterogeneous" realizers in the reducing physical theory. As Fodor puts it, "The problem … has been that there is an open empirical possibility that what corresponds to the natural kind predicates of a reduced science may be a heterogeneous and unsystematic disjunction of predicates in the reducing science" (p. 108). Thus, to continue the psychology example so prevalent in the literature, pain—a property appearing in the science of psychology whose predicate (perhaps) appears in its laws—may be realized by distinct physical or neurophysiological properties in humans, in reptiles, and possibly even in inorganic robots.

This has the consequence no one neurophysiological state can be correlated or identified with the psychological property pain. In humans it may be C-fibers firing; in reptiles it may be D-fibers firing; and in robots it may be the activation of some particular integrated circuit. The heterogeneous nature of the distinct realizers also makes it unlikely that a disjunction of those realizers will be a natural kind term in the reducing theory. Given this, and if laws relate natural kinds to natural kinds, it is unlikely that there can be anything lawlike about the bridge laws. This argument has been applied to many functionally defined properties such as being a thermostat or being a heart—properties that can be realized in many different ways in different systems or organisms.

One response, due to Kim (1992), to the realization argument is to note that while the argument may block a kind of global reduction of the special science to the lower-level physical theory, it may be possible to have (local) species or structure specific reductions. Thus, for instance, one might be able to locally reduce human pain to human neurophysiology, reptilian pain to reptilian neurophysiology, and robot pain to robot neurocircuitry. Kim (1998, chapter 4) develops an alternative functional model of reduction appropriate to this response.

Another approach (Batterman 2000) asks for an account of what makes the multiple realizability possible. Typically, multiple realizability is simply assumed and applied via the multiple realization argument to block Nagelian reductions. For instance, Fodor (1974) cites it simply as an open possibility. However, it seems reasonable to ask whether one can explain, from the point of view of the supposed reducing theory, the possibility of multiple realizability. If so, this may lead to a kind of explanation without reduction. Cases where such expla-

nations are indeed possible can be found in the physics literature where attempts are made to explain surprising universal features of various systems. *Universality* means identical or similar behavior in physically distinct systems and is, therefore, a term essentially synonymous with *multiple realizability*. For details about how such an approach to multiple realizability will go, see Robert W. Batterman (2000, 2002).

REDUCTION IN THE OTHER DIRECTION

There is an interesting terminological ambiguity that infects the term *reduction* as it is typically used in the philosophical literature and as it is used in the physics literature. Philosophers will discuss the reduction of thermodynamics to statistical mechanics, the reduction of classical mechanics to quantum mechanics, and the reduction of the ray theory of light to the wave theory. The succeeded theories are all reduced to their successors. Physicists, when they talk about theory reduction at all, tend to put things the other way around. They will say that statistical mechanics reduces to thermodynamics in the limit as the number of degrees of freedom goes to infinity. They will say that quantum mechanics reduces to classical mechanics in some kind of correspondence limit. Furthermore, they will say that the wave theory reduces to the ray theory in the limit as the wavelength of light approaches zero. That there are such different senses of intertheoretic reduction was first noted by Thomas Nickles in the paper "Two Concepts of Intertheoretic Reduction" (1973).

Interestingly, the physicists' sense of reduction appeals to limiting relations between the pair of theories and is not concerned with derivation in logical/syntactic sense that has primarily concerned philosophers following in Nagel's footsteps. In other words, there is no explicit concern, say, with the derivation of the laws of thermodynamics from the laws of statistical mechanics. Instead, the interest is in the potential emergence of those laws as some sort of mathematical limit is asymptotically approached. Thus, while the philosophical tradition focuses on the schema according to which theory T' reduces to theory T just in case one can derive (and thereby explain) the laws of T' from the laws of T, this other sense of reduction focuses on a schema of the following form (1):

$$\lim_{\epsilon \to 0} T = T'$$

in which theory T reduces to T' in the regime where a parameter $\in$ appearing in theory T takes on a limiting value. For instance, quantum mechanics contains a constant (Planck's constant) that plays no role in classical mechanics. As a result, one may be motivated to study the limit of quantum mechanics in which Planck's constant approaches zero. This is a kind of correspondence limit.

The two schemas are related to one another at least in the following way. Should the equality in (1) hold for two theories T and T', then it is reasonable to expect that the laws of T' are derivable from those of T. That is to say, it is likely that one will be able to find the appropriate connections that will allow something like a Schaffner-style neo-Nagelian reduction. On the contrary, if the equality in 1fails to obtain for the pair of theories, then such neo-Nagelian reduction will not be possible. It will be impossible to form the relevant corrected reduced theory T'^*.

One case for which the schema (1) does obtain is in the relationship between (certain aspects of) classical Newtonian mechanics (NM) and the special theory of relativity (SR). In the limit in which velocities are slow compared with the speed of light ($(v/c) \to 0$), SR reduces to NM. The limit exists and the formulas of SR smoothly (that is uniformly) approach those of NM.

However, far more often than not pairs of theories will be related to one another by so-called singular limits. Singular limits arise when the behavior as the limit is approached (no matter how small $\in$ becomes) is qualitatively different from the behavior at the limit (when $\in = 0$). In such cases the equality in schema (1) fails to obtain: There will be no smooth approach of the formulas of theory T to those of theory T'.

In fact, it is fair to say that for most theory pairs of interest, schema (1) will fail. Important examples include those mentioned earlier: quantum mechanics and classical mechanics; wave theory and ray theory; and statistical mechanics and thermodynamics. Certain formulas in each of these theory pairs are related by singular limits, and it is best, perhaps, to give up on speaking of reductive relations between the theories or at least between those features characterized by the singularly related formulas (see Berry 1994, 2002; Batterman 1995, 2002).

It is important to stress that if this is correct, and so physicists' reductions are genuinely few and far between, this does not mean that there is no reason to study the singular limiting relationships between theories. In fact, the opposite is true. There is much of interest to study in the borderland between the theories. Michael V. Berry

notes the importance of the failure of reduction due to singular limiting relations between the theories, "[M]any difficulties associated with reduction arise because they involve singular limits. These singularities have both negative and positive aspects: They obstruct smooth reduction of more general theories to less general ones, but they also point to a great richness of borderland physics between theories" (2001, p. 42).

The "great richness" of this borderland is fertile ground for studying certain aspects of emergence, a philosophical topic related to the failure of reduction. Emergent phenomena are typically taken to be novel in certain respects where this novelty is often understood as resulting from the failure of the more basic theory to explain or otherwise account for the phenomena. There has been considerable interest in the controversial issues surrounding the nature and existence of emergent phenomena. Two related approaches with the same starting point—the singular nature of limiting intertheoretic relations—are examined by Batterman (2002) and Alexander Rueger (2000a, 2000b). However, another related approach can be found in the work of Hans Primas (1998). For a different, more metaphysically motivated attempt to understand emergence, see Paul Humphreys (1997). This is currently an active area of research.

See also Copernicus, Nicolas; Fodor, Jerry A.; Galileo Galilei; Hempel, Carl Gustav; Kuhn, Thomas; Laws and Theories; Multiple Realizability; Nagel, Ernest; Newton, Isaac; Philosophy of Science; Properties; Scientific Theories.

Bibliography

Batterman, Robert W. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press, 2002.

Batterman, Robert W. "Multiple Realizability and Universality." *British Journal for the Philosophy of Science* 51 (2000): 115–145.

Batterman, Robert W. "Theories between Theories: Asymptotic Limiting Intertheoretic Relations." *Synthese* 103 (1995): 171–201.

Berry, Michael V. "Asymptotics, Singularities, and the Reduction of Theories." In *Logic, Methodology, and Philosophy of Science*, edited by Dag Prawitz, Brian Skyrms, and Dag Westerståhl, 597–607. Amsterdam, Netherlands: Elsevier Science B. V., 1994.

Berry, Michael V. "Chaos and the Semiclassical Limit of Quantum Mechanics (Is the Moon There When Somebody Looks?)." In *Quantum Mechanics: Scientific Perspectives on Divine Action*, edited by Robert John Russell et al., 41–54. Rome, Italy: Vatican Observatory CTNS Publications, 2001.

Berry, Michael V. "Singular Limits." *Physics Today* (May 2002): 10–11.

Feyerabend, Paul K. "Explanation, Reduction, and Empiricism." In *Minnesota Studies in the Philosophy of Science. Vol. 3, Scientific Explanation, Space, and Time*, edited by Herbert Feigel and Grover Maxwell. Minneapolis: University of Minnesota Press, 1962.

Fodor, Jerry. "Special Sciences, or the Disunity of Sciences as a Working Hypothesis." *Synthese* 28 (1974): 97–115.

Hempel, Carl G. *Studies in the Logic of Explanation: In Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press, New York, 1965.

Humphreys, Paul. "How Properties Emerge." *Philosophy of Science* 64 (1) (1997): 1–17.

Kim, Jaegwon. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press, 1998.

Kim, Jaegwon. "Multiple Realization and the Metaphysics of Reduction." *Philosophy and Phenomenological Research* 52 (1) (1992): 1–26.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press, 1970.

Nagel, Ernest. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace, 1961.

Nickles, Thomas. "Two Concepts of Intertheoretic Reduction." *Journal of Philosophy* 70 (7) (1973): 181–201.

Pippard, A. B. *Elements of Classical Thermodynamics*. Cambridge, U.K.: Cambridge University Press, 1957.

Primas, Hans. "Emergence in Exact Natural Sciences." *Acta Polytechnica Scandinavica, Ma* 91 (1998): 83–98.

Rueger, Alexander. "Physical Emergence, Diachronic and Synchronic." *Synthese* 124 (2000a): 297–322.

Rueger, Alexander. "Robust Supervenience and Emergence." *Philosophy of Science* 67 (2000b): 466–489.

Schaffner, Kenneth. "Approaches to Reduction." *Philosophy of Science* 34 (2) (1967): 137–147.

Schaffner, Kenneth. "Reductionism in Biology: Prospects and Problems." In *PSA 1974*, edited by A. C. Michalos et al., 613–632. Dordrecht, Netherlands: D. Reidel, 1976.

Sklar, Lawrence. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. New York: Cambridge University Press, 1993.

Sklar, Lawrence. "Types of Inter-theoretic Reduction." *British Journal for the Philosophy of Science* 18 (1967): 109–124.

Wimsatt, William C. "Reductive Explanation: A Functional Account." In *PSA 1974*, edited by A. C. Michalos et al., 671–710. Dordrecht, Netherlands: D. Reidel, 1976.

Robert W. Batterman (2005)

REDUCTIONISM IN THE PHILOSOPHY OF MIND

See Appendix, Vol. 10

REFERENCE

"Reference" is usually conceived as the central relation between language or thought and the world. To talk or think about something is to refer to it. Twentieth-century philosophy found such relations particularly problematic. One paradigm of reference is the relation between a proper name and its bearer. On a more theoretical conception all the constituents of an utterance or thought that contribute to determining whether it is true refer to their contributions (as, for example, a predicate refers to a property). In analytic philosophy discussion of reference was dominated until the 1960s by the views of Gottlob Frege and Bertrand Russell and modifications of them (such as those by P. F. Strawson). Criticisms of assumptions common to those views then provoked a revolution in the theory of reference. The alternatives include causal and minimalist theories.

OBJECTIONS TO DESCRIPTIVISM

One model of reference is that of descriptive fit. The paradigm is a definite description (such as "the tallest tree") that refers to whatever it accurately describes. Frege and Russell assimilated the reference of ordinary proper names to this case by supposing that speakers associate them with descriptions. Similar accounts were later given of mass terms (such as "blood"), natural-kind terms ("gorilla"), and theoretical terms in science ("inertia"). It was conceded that most terms are associated with vague and context-dependent clusters of descriptions and that reference might be to whatever they least inaccurately described, but such liberalizations did not challenge the underlying idea that descriptive fit determines reference. However, Keith Donnellan, Saul Kripke, and Hilary Putnam proposed counterexamples to that idea. Suppose, for instance, that speakers associate the name "Jonah" with the Bible story. Traditional descriptivism concludes that the sentence "Although Jonah existed, those things happened only to someone else" is untrue. For if one person satisfied the relevant descriptions, "Jonah" would refer to him. But then descriptivism proves too much, for philosophical reflection cannot show that the Bible story is not a mere legend that grew up about a real person; if those things really happened to someone else, of whom no word reached the biblical writer, the name "Jonah" would still refer to the former, not the latter. Similarly, traditional descriptivism permits someone who thinks of gorillas primarily as ferocious monkeys to conclude falsely that the sentence "Gorillas exist, but they are not ferocious monkeys" is untrue.

A second criticism was this. Say that a term *t* rigidly designates an object *x* if and only if *t* designates (refers to) *x* with respect to all possible circumstances (except perhaps for circumstances in which *x* does not exist). Most descriptions designate nonrigidly: "the tallest tree" designates one tree with respect to present circumstances, another with respect to possible circumstances in which the former is outgrown. The descriptions that traditional descriptivists associated with names were nonrigid. However, names designate rigidly: Although we can envisage circumstances in which the Danube would have been called something else instead, we are still using our name "Danube" to hypothesize circumstances involving the very same river. Thus, most descriptions do not behave like names.

The second criticism was met by a modification of descriptivism. The descriptions associated with a name were rigidified by a qualifying phrase such as "in present circumstances." "The tallest tree in present circumstances" rigidly designates what "the tallest tree" nonrigidly designates. The first criticism is less easily met. Some descriptivists used deferential descriptions such as "the person referred to in the Bible as 'Jonah.'" A more general strategy is to exploit the success of any rival theory of reference by building that theory into the associated descriptions. However, such moves jeopardize the connection between reference and speakers' understanding (a connection that descriptivism was intended to secure) as the descriptions that speakers supposedly associate with names become less and less accessible to the speakers themselves.

It is in any case clear that, as Russell recognized, not all reference is purely descriptive. If the sentence "It is hot now" is uttered at different times in exactly similar circumstances, associated with exactly the same descriptions, those descriptions are not what determines that it changes its reference from one time to the other. The reference of a token of "now" is determined by the time of its production and the invariant linguistic meaning of "now," the rule that any such token refers to the time of its production. Similarly, the presence of an object to the speaker or thinker plays an ineliminably nondescriptive role in the reference of demonstratives such as "this."

NONDESCRIPTIVISM

THE KRIPKE-PUTNAM PICTURE. Kripke and Putnam proposed an alternative picture. Something *x* is singled out, usually demonstratively ("this river," "this kind of animal"). A name *n,* proper or common, is conferred on *x* ("Danube," "gorilla"). The name is passed on from one

speaker to another, the latter intending to preserve the former's reference. Such intentions are self-fulfilling: *n* continues to refer to *x.* The beliefs that speakers would express in sentences containing *n* play no role in making *n* refer to *x,* so it can turn out that most of them are false. The picture involves two kinds of deference. Synchronically, there is division of linguistic labor: Ordinary speakers defer to experts (as in deciding which animals "gorilla" refers to). Diachronically, later speakers defer to earlier ones in a historical chain. Thus, reference typically depends on both the natural environment of the initial baptism (to fix the demonstrative reference) and the social environment of the later use. An individual speaker's understanding plays only a minor role. The account may be generalized (as to many adjectives and verbs).

The picture needs qualification. Gareth Evans pointed out that a name can change its reference as a result of misidentification, even if each speaker intends to preserve reference. What matters is not just the initial baptism but subsequent interaction between word and object. Such concessions do not constitute a return to descriptivism.

CAUSAL THEORIES. The Kripke-Putnam picture is often developed into a causal theory of reference, on which for *n* to refer to *x* is for a causal chain of a special kind to connect *n* to *x.* Such a theory goes beyond the original picture in at least two ways. First, although that picture required later uses of *n* to depend causally on the initial baptism, it did not require the initial baptism to depend causally on *x.* Kripke allowed reference to be fixed descriptively (not just demonstratively), as in "I name the tallest tree 'Albie'"; he merely insisted that the description did not give the meaning of the name. There is no causal connection between the name "Albie" and the tree Albie. Second, Kripke and Putnam did not attempt to define the notions they used in causal terms; the notion of an intention to preserve reference is not obviously causal.

Causal theories are often motivated by a desire to naturalize linguistic and mentalistic phenomena by reducing them to the terms of physical science. Such theories are therefore not restricted to proper names. Causal theorists will postulate that our use of the words "tall" and "tree" is causally sensitive to tallness and trees respectively, hoping thereby to explain the reference of "Albie." One problem for causal theories is that any word is at the end of many intertwined causal chains with different beginnings. It is extremely difficult to specify in causal terms which causal chains carry reference. For this reason, causal theories of reference remain programmatic.

DIRECT REFERENCE. Consonant with the Kripke-Putnam picture, but independent of causal theories of reference, is the theory of direct reference developed by David Kaplan. A term *t* directly refers to an object *x* in a given context if and only if the use of *t* in that context contributes nothing to what is said but *x* itself. For Kaplan, proper names, demonstratives, and indexicals such as "now" refer directly. Ruth Barcan Marcus had earlier made the similar suggestion that proper names are mere tags. The reference of a directly referential term may be determined relative to context by its context-independent linguistic meaning, as for "now"; the claim is that what "now" contributes to the proposition expressed by an utterance of "It is hot now" is not its invariant linguistic meaning but the time itself.

Although all direct reference is rigid designation, not all rigid designation is direct reference: "the square of 7" rigidly designates 49, but the reference is not direct, for the structure of the description figures in the proposition expressed by "The square of 7 is 49." On one view all genuine reference is direct, sentences of the form "The *F* is *G*" being quantified on the pattern of "Every *F* is *G*" (as Russell held); "the *F*" is neither a constituent nor a referring term.

If "Constantinople" and "Istanbul" have the same direct reference, the proposition (C) expressed by "Constantinople is crowded" is the proposition (I) expressed by "Istanbul is crowded," so believing (C) is believing (I), even if one would not express it in those words. Similarly, when a term of a directly referential type fails to refer, sentences in which it is used express no proposition. The view is anti-Fregean. In suitable contexts Frege would attribute different senses but the same reference to "Constantinople" and "Istanbul" and a sense but no reference to an empty name; for him the sense, not the reference, is part of what is said or thought. Russell held that logically proper names are directly referential but concluded that ordinary names are not logically proper. The challenge to defenders of the direct-reference view is to explain away the appearance of sameness of reference without sameness of thought and absence of reference without absence of thought, perhaps by postulating senselike entities in the act rather than the content of thought. The theory of direct reference concerns content, not the mechanisms of reference.

MINIMALISM. Traditional theorizing about reference is ambitious; the possibility of a broad and deep theory such as it seeks has been questioned by Richard Rorty, Robert Brandom, Paul Horwich, and others. The following schema constitutes a minimal account of reference ("*a*" is replaceable by singular terms):

(R) For any *x*, "*a*" refers to *x* if and only if $x = a$.

"London" refers to London and nothing else. A minimalist account adds to (R) the claim that (R) exhausts the nature of reference.

Some qualifications are necessary. First, if anything but a singular term replaces "*a*" in (R), the result is ill formed, for only singular terms should flank the identity sign. If expressions of other syntactic categories refer, those categories will require their own schemas. The schema for predicates might be:

(R') For any *x*, "*F*" refers to *x* if and only if $x =$ *F*ness.

Second, the notion of a singular term must be explained (can "my sake" replace "*a*"?). Third, (R) does not say which singular terms refer. When "*a*" does not refer, (R) may not express a proposition. Fourth, (R) cannot be generalized by the prefix "In all contexts": "today" used tomorrow does not refer to today. Rather, (R) should be understood as instantiated by sentences in different contexts (for instance, uttered tomorrow with "today" for "*a*"). Fifth, when one cannot understand the term "*a*," one cannot understand (R). Thus, one will find many instances of (R) unintelligible.

One's grasp of the minimal theory is not a grasp of each of many propositions; it is more like one's grasp of a general pattern of inference. For (R) the pattern is in the sentences that express the propositions, not in the propositions themselves (it is not preserved when a synonym replaces the unquoted occurrence of "*a*"). This generality does not satisfy all philosophers. Many accept the minimal theory but reject minimalism, because they postulate a deeper (for instance, causal) theory of reference that explains (R) and (R'). Although the reductionist demand for strictly necessary and sufficient conditions for reference in more fundamental terms may be overambitious, a good picture of reference might still reveal more than (R) and (R') without meeting that demand.

See also Frege, Gottlob; Indexicals; Kaplan, David; Kripke, Saul; Marcus, Ruth Barcan; Philosophy of Language; Proper Names and Descriptions; Putnam, Hilary; Rorty, Richard; Russell, Bertrand Arthur William; Sense; Strawson, Peter Frederick.

Bibliography

Almog, J., J. Perry, and H. Wettstein, eds. *Themes from Kaplan.* New York: Oxford University Press, 1989.

Brandom, R. "Reference Explained Away: Anaphoric Reference and Indirect Description." *Journal of Philosophy* 81 (1984): 469–492.

Devitt, M. *Designation.* New York: Columbia University Press, 1981.

Evans, G. *The Varieties of Reference.* Oxford: Clarendon Press, 1982.

Fodor, J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge, MA: MIT Press, 1987.

French, P., T. Uehling, and H. Wettstein, eds. *Contemporary Perspectives in the Philosophy of Language.* Midwest Studies in Philosophy, Vol. 5. Minneapolis: University of Minnesota Press, 1979.

Horwich, P. *Truth.* Oxford: Blackwell, 1990.

Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.

Lewis, D. "Putnam's Paradox." *Australasian Journal of Philosophy* 62 (1984): 221–236.

Neale, S. *Descriptions.* Cambridge, MA: MIT Press, 1990.

Putnam, H. *Philosophical Papers,* Vol. 2: *Mind, Language, and Reality.* Cambridge, U.K.: Cambridge University Press, 1975.

Récanati, F. *Direct Reference: From Language to Thought.* Oxford: Blackwell, 1993.

Rorty, R. *Philosophy and the Mirror of Nature.* Princeton, NJ: Princeton University Press, 1979.

Schwartz, S., ed. *Naming, Necessity, and Natural Kinds.* Ithaca, NY: Cornell University Press, 1977.

Timothy Williamson (1996)

REFLECTIVE EQUILIBRIUM

Reflective equilibrium is a coherence method of philosophical justification or inquiry. Nelson Goodman (1955) introduced reflective equilibrium, although not under that name, to contemporary philosophy in a discussion of deductive and inductive logic. It is arguable, however, that philosophers have employed something such as reflective equilibrium to inquire into a wide range of topics since ancient times.

Goodman maintained that we justify an inference by showing that it conforms to the rules of either deduction or induction. But for the inferences to be justified, these rules must be valid. Goodman held that we justify rules of inference by showing that they accord with judgments we make about which particular inferences are acceptable and which are unacceptable. Goodman addressed the obvious objection to such a procedure as follows:

> This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general

> rules are justified by their conformity to valid inferences. But this circle is a virtuous one. A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement thus achieved lies the only justification needed for either.
>
> (GOODMAN 1955, P. 67)

It is possible to read Goodman's proposal as purely methodological or as more epistemological. According to the epistemological reading, when we complete the process of mutual adjustment Goodman describes, thereby bringing our judgments regarding the particular inferences and the rules of inference we accept into a state of reflective equilibrium, these rules and particular judgments are by definition justified. According to this reading, being justified consists in being part of a system of beliefs, including rules and particular judgments, that has the sort of coherence that reflective equilibrium represents.

The methodological understanding of reflective equilibrium accepts that a philosophical inquiry into inductive or deductive inference is properly conducted by a process of mutual adjustment of the kind Goodman describes; it agrees that this is the best we can do in an attempt to justify the inferences we make and the rules of inference we accept. But the methodological reading is not definite about the epistemic status of the particular and general judgments we manage to bring into reflective equilibrium. It leaves open what positive epistemic status, if any, principles and judgments that are in reflective equilibrium might have. In addition, whereas the epistemological reading is committed to a coherentist account of justification, the methodological reading leaves open how best to account for the precise epistemic status (or statuses) attained by judgments that are in reflective equilibrium. Although Goodman probably intended an epistemological reading, let us adopt a methodological understanding of reflective equilibrium in order to keep as many epistemological options open as possible.

We owe the term reflective equilibrium to John Rawls (1971), who developed the method further and applied it to moral inquiry. If we think of the method as something to be applied in a stepwise fashion, which is useful even if not entirely accurate, then an inquirer, S, begins with a large set of initial moral judgments. These judgments will be intuitive for S in the sense that they are cognitively spontaneous; they might concern propositions that are either particular or general. The first step on the road to reflective equilibrium, according to Rawls, is for S to eliminate certain initial moral judgments. For example, judgments that are not stable over time or in which S has little confidence should be dropped, as should judgments formed when S is emotionally distraught. In general, S eliminates those judgments formed in circumstances where there is some obvious reason for suspecting error. The remaining judgments will be S's considered moral judgments.

S's next task is to formulate a moral theory, that is, a set of moral principles that accounts for S's considered moral judgments. The coherence element of the method comes into play at this stage because S will grant neither considered moral judgments nor moral theory a privileged status. S will make revisions on both sides in the attempt to forge a coherent system of moral beliefs. When the emerging theory is found to conflict with central, very confidently made considered judgments, S must revise the theory. But if a well-confirmed element of the theory that is independently plausible is found to conflict with less firmly held considered judgments, then S will revise these judgments. S's decision regarding what to revise is made for each case on the basis of what seems most likely to be true or correct to S upon due consideration (there are various other ways of expressing this idea: We might, for example, say that S is to decide on the basis of S's degrees of belief or commitment or on the basis of what seems most plausible or acceptable to S).

To this point, Rawls's method corresponds with Goodman's, but Rawls does not allow S to stop here, at a point of narrow reflective equilibrium. According to Rawls, S must next consider alternatives to the moral theory that S accepts in narrow equilibrium along with philosophical arguments for and against S's own theory and the various alternatives S is considering. In his important work on reflective equilibrium, Norman Daniels (1979) argues that we can think of this as an attempt to attain coherence between the considered moral judgments and theory that S accepts in narrow reflective equilibrium and the background theories S accepts. The idea is that the philosophical arguments that S constructs will use premises drawn from among S's broader background beliefs, which might include such things as sociological views regarding the role of morality in society and philosophical or psychological theories regarding rational decision or the nature of persons.

An argument in favor of an alternative to the moral theory that S accepts in narrow reflective equilibrium that

is successful in the sense that S finds it compelling would, in effect, show that S's moral theory, considered moral judgments, and background beliefs are not coherent. As before, S is not bound to favor any type of belief when responding to such an argument; S must decide whether to revise considered moral judgments and moral theory or background beliefs on the basis of what seems most likely to be true to S after thorough reflection. When S attains a coherent system of considered moral judgments, moral theory, and background beliefs, S will have reached a state of wide reflective equilibrium.

An inquirer can certainly move from narrow to wide reflective equilibrium in the way that Daniels maintains, but this is not the only or most interesting way things can go. Consideration of alternative moral views and the relevant philosophical arguments can provide an occasion for a more radical type of revision of belief. Daniels seems to suppose that the only way in which considering alternatives to one's own view can force one to revise beliefs is by revealing that something else one believes, and believes more strongly, conflicts with one's prior view but coheres with the alternative. But it is quite clear that consideration of alternative moral and broader philosophical views can also lead one to revise beliefs in a way that is not dictated by one's prior beliefs and degrees of commitment. It is possible for an inquirer to find an alternative view attractive in its own right, even though it conflicts with everything the inquirer previously thought; and if upon reflection the inquirer finds the new alternative sufficiently attractive, he or she might well respond by accepting the alternative and revising his or her previous views.

Wide reflective equilibrium is best understood in a way that allows for this radical type of belief revision. According to this understanding, achieving wide reflective equilibrium is not simply a matter of rooting out conflicts among the beliefs one already holds and forging general principles that coherently account for one's considered moral judgments. It crucially involves exposing one's self to alternative moral and philosophical views with the knowledge that reflection upon such alternatives might lead one to make a radical break with one's previous views.

On this understanding, the ideal of wide reflective equilibrium is not defined merely as achieving coherence among all of one's beliefs: considered moral judgments, a moral theory, and background beliefs. The ideal crucially involves attaining a kind of reflective stability. Inquirers who have attained reflective equilibrium are, in effect, immune to threats from the inside and the outside. There will be no conflicts within such inquirers' systems of belief, and in addition they can be confident that there are no alternatives to their own systems of belief that they would find more compelling than their own upon due reflection. The first sort of reflective stability is provided by reflective equilibrium on either understanding, the second only if reflective equilibrium is understood in a more radical way.

It is important to recognize that the essential feature of a belief revision that is radical, in the strict sense here at issue, is not the number or range of beliefs that are altered but rather the fact that the alteration is not continuous with the things that one previously believed. When a belief is revised in a way that is not strictly radical, the change is required in order to attain coherence among one's beliefs, and the alteration is dictated by other things that one believes more firmly than the belief that is revised. Particularly if the belief that is revised in this way concerns a general principle, the change can require revisions to a large number of other beliefs; such a revision would likely be called radical in common parlance, but it would not be radical in the strict sense.

When a change is one that counts as radical in the strict sense, the new beliefs come to seem compelling to one on their own, apart from their logical or evidential relations to one's previous beliefs. Indeed, the new belief will likely contradict things that one previously believed very strongly. Such a change may involve many beliefs or only a few. Philosophers, who might be guilty of considering such matters only abstractly, may find it difficult to accept the possibility of such radical changes in belief, but it is easy to find descriptions in novels, biographies, and autobiographies of people altering their views in ways that seem to be radical in the strict sense.

Because reflective equilibrium grants the inquirer's considered moral judgments a crucial role in inquiry, it has been widely criticized as a sophisticated version of intuitionism, making it an unreliable and extremely conservative method. Daniels (1979) sought to rebut the charge of intuitionism by arguing that reflective equilibrium is not compatible with foundationalism, which is a characteristic of intuitionism. Daniels's basic idea is that the inquirer's considered moral judgments do not function as intuitions because reflective equilibrium allows for such extensive revision of these judgments. This response seems to rely upon too narrow a conception of foundationalism; in particular, it seems to suppose that the beliefs that serve the foundational role must be identifiable in advance of inquiry and also be unrevisable, or at least relatively unrevisable.

Given these suppositions it is natural to think that if reflective equilibrium is a foundationalist method, the inquirer's initial considered moral judgments must be the foundations, and then conclude that it cannot be a foundationalist method because these judgments are subject to way too much revision. However, if one supposes instead that the foundations might emerge through a course of inquiry, it is unclear that reflective equilibrium does not constitute a version of foundationalism. During a person's inquiry there will be various relatively strongly held judgments that determine the course of the inquiry and the views the person comes to hold in reflective equilibrium. It is unlikely that all these judgments will be drawn from among the inquirer's considered moral judgments, but almost certainly many will. Others might come from among the person's background beliefs, some might concern moral principles, and perhaps some will be about which member of a conflicting set of beliefs should be revised.

But that does not really matter. The fact remains that at the end of inquiry, it will be possible to identify a set of judgments that provide a psychological basis for the rest of the beliefs the person holds in reflective equilibrium. Many, although perhaps not all, of these judgments will probably be intuitive in the sense that they are cognitively spontaneous. It is possible, therefore, that these intuitive judgments serve as an epistemological foundation for the rest of the beliefs the person holds in reflective equilibrium. It remains to be seen whether this constitutes a ground for objection to reflective equilibrium.

Reflective equilibrium fares better when it comes to the charge that it is extremely conservative. The extensive revisability of considered moral judgments may show that it can be construed as a version of intuitionism, but it surely shows that the method is not guaranteed to produce nothing more than a cleaned up, systematized version of conventional morality. The method has the potential to, and indeed is likely to, lead many inquirers to make extensive changes to their moral views. A particular inquirer might, of course, end up holding very conventional views in reflective equilibrium. Indeed, this is just what will happen if the inquirer is more strongly committed to enough elements of conventional morality than he or she is to anything that conflicts with them, and retains these commitments through the course of reflection upon alternatives to and criticisms of conventional morality. But it is not clear that a method of moral inquiry is inadequate unless it absolutely excludes this possibility.

More worrisome is a general fact illustrated by the possibility just considered: Given the right (or perhaps one should say wrong!) moral judgments and background beliefs held strongly and tenaciously enough, it would seem to be possible for an inquirer to end up holding virtually any moral view, even a bizarre or repugnant view, in reflective equilibrium. The worry is not confined to reflective equilibrium when used as a method of moral inquiry. No matter what a person might use this method to inquire about, given the right intuitive beliefs held with sufficient strength and tenacity, the person could end up holding virtually any view one could imagine in reflective equilibrium: extreme skepticism, solipsism, nihilism, anarchism, totalitarianism, atheism, or theism—you pick whatever views you think are beyond the pale. How then could anyone take reflective equilibrium to be an acceptable approach to moral inquiry, or philosophical inquiry more generally? As various critics have put the point: The method clearly leads an inquirer to the coherent position he or she finds most acceptable, the position that best preserves beliefs to which he or she is most strongly committed, but why think this position is anything more than that, in particular, why think it is true or likely to be true?

One might have once thought that this fundamental objection is really pressing only against reflective equilibrium when used to inquire into morality and other such things, where it is all too obvious that different people can hold, and hold very strongly, very different and incompatible considered judgments. One might have presumed that for such purposes as working out valid rules for deductive or inductive inferences, which is what Goodman originally proposed that the method be used for, there is no real problem because there just is not the same sort of diversity and conflict between the strongly held considered judgments of different people. But as Stephen Stich most particularly has stressed, empirical work has shown that many of the inferences ordinary people find intuitively acceptable are in fact fallacious. Hence, we can foresee that the rules of inference these people accept in reflective equilibrium will not be valid. So we cannot even trust reflective equilibrium to be an acceptable method of inquiry for those areas where it was originally proposed.

It might seem, therefore, that Rawls's early critics were right to argue that unless we can find some reason for trusting the reliability of the intuitive judgments that play such a crucial role in the method of reflective equilibrium, this method of inquiry cannot be acceptable. Daniels (1979) was perhaps right when he claimed, in response, that it is unreasonable to expect such a reason to be provided before we begin our inquiries and that

such a reason should only be expected to emerge as part of the overall system of beliefs accepted in reflective equilibrium. But surely it seems unreasonable to hope that it will ever be possible to offer even such an internal defense of reflective equilibrium as a method that is reliable for all who might employ it. Such a defense would require that all, or nearly all, inquirers employing the method converge on the same theory, and this seems rather unlikely. In addition, this sort of defense seems to underestimate the obstacles facing reflective equilibrium: It is not just that there is no reason to think the method is reliable, but also none for suspecting that it is unreliable—there are fairly strong reasons for believing that the method is unreliable, that it is not the case that a very high proportion of those who employ it will be led to accept a system of beliefs that is largely correct.

Nevertheless, it is possible to offer a defense, albeit a modest defense, of reflective equilibrium. The first step is to recognize that there are a number of different positive epistemic statuses. For simplicity, let's distinguish only two. The first is the positive epistemic status that plays the major role in distinguishing knowledge from mere true belief. This status is most commonly referred to as justification or warrant. Attempts to account for justification in terms of the reliable formation of belief have been popular and influential. Even if such attempts fail, the majority of epistemologists would still maintain that there is some sort of strong connection between justification and truth: justified beliefs must, in some sense, be likely to be true. The second positive epistemic status is the sort of subjective rationality that Richard Foley has stressed. A belief is rational in this sense when it satisfies the believer's own epistemic standards, that is, when the believer would consider the belief likely to be true after due reflection. Unlike justification, there seems no reason to suppose that beliefs that are rational in this sense are likely to be true.

Having distinguished these two positive epistemic statuses, it should be fairly clear that reflective equilibrium can, in fact, be guaranteed to lead the inquirer to hold rational beliefs. It should also be easy to see that an inquirer who deviated from reflective equilibrium would be led to hold some beliefs that are not rational because, in order to deviate, the inquirer would have to resolve some conflict by rejecting a belief that, upon reflection, the inquirer considers more likely to be true than the belief being retained. What reflective equilibrium cannot guarantee every inquirer is justified beliefs. If it followed that by employing the method of reflective equilibrium an inquirer was sure to form rational beliefs but equally sure to form unjustified beliefs, and hence fail to attain knowledge, the method would indeed be unacceptable. But it would be hasty to infer that no inquirer employing reflective equilibrium can be led to hold justified beliefs simply because the method cannot guarantee justified beliefs to all inquirers.

In an influential paper on the method of moral inquiry written years before he advocated reflective equilibrium, Rawls (1951) argued that we should construct a moral theory by formulating principles that account for the considered moral judgments of competent moral judges. Whereas the notion of considered moral judgments is used in describing reflective equilibrium, the notion of the competent moral judge has fallen by the wayside. But suppose, as we clearly do in our ordinary lives, that some people are competent moral judges, whereas others are not. We ordinarily suppose that the intuitive moral judgments of competent moral judges are reliable. We might be wrong, and of course we might also be right, that those people who have the characteristics we commonly associate with moral competence, in fact, make a person a reliable moral judge. So let us understand competent moral judges as those whose intuitive moral judgments are reliable. The beliefs competent moral judges would hold in reflective equilibrium obviously would be reliable.

If one condition for being justified is that a person must not only be reliable, but be able to prove that he or she is, then perhaps even the beliefs held by competent moral judges in reflective equilibrium are not justified. For competent judges will not be able to prove to incompetent judges that they are reliable. But this condition for being justified is almost certainly too strong: If we were to apply it across the board, we would know little or nothing. If what is necessary for justification is only that one is reliable, not that one be able to prove that one is, then the beliefs competent moral judges hold in reflective equilibrium may, for all that has been said so far, be justified.

This is not, of course, all that we might have wanted in the way of a defense of reflective equilibrium. We cannot prove that anyone is or is not a competent judge. We cannot prove which characteristics make for competent judges and which make for incompetent judges. Perhaps there are no competent judges. Perhaps there are, but even they do not know that they are. But it is not outlandish to think that there are competent moral judges and that most of us know something about who they are and what they are like. But for the sake of this argument, suppose only that it is possible that there are competent

moral judges. If there are any competent judges, then the beliefs they hold in reflective equilibrium are justified. And this suggests that these beliefs, or at least many of them, might count as knowledge.

So we can say this much about reflective equilibrium. It is the only rational method of inquiry and it is possible that, by employing this method, a person will be lead to hold justified beliefs and to attain knowledge. This is certainly less than one would like to be able to say in support of a method of philosophical inquiry, but it is sufficient to show that reflective equilibrium is an acceptable method for ethics, and philosophy more generally.

See also Applied Ethics; Goodman, Nelson; Logic, History of: Modern Logic: From Frege to Gödel; Metaethics; Moral Epistemology; Rawls, John.

Bibliography

Bonevac, Daniel. "Reflection without Equilibrium." *The Journal of Philosophy* (2004).

Brandt, Richard. *A Theory of The Good and The Right*. Oxford: Clarendon Press, 1979.

Brink, David O. *Moral Realism and The Foundations of Ethics*. Cambridge, U.K.: Cambridge University Press, 1989.

Daniels, Norman. *Justice and Justification: Reflective Equilibrium in Theory and Practice*. New York: Cambridge University Press, 1996.

Daniels, Norman. "On Some Methods of Ethics and Linguistics." *Philosophical Studies* 37 (1980): 21–36.

Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76 (1979): 256–282.

DePaul, Michael R. *Balance and Refinement: Beyond Coherentism in Moral Inquiry*. London: Routledge, 1993.

DePaul, Michael R., and William Ramsey. *Rethinking Intuition: The Psychology of Intuitions and Their Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield, 1998.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Goodman, Nelson. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press, 1955.

Hare, R.M. "Rawls' Theory of Justice—I." *The Philosophical Quarterly* 23 (1973).

Harman, Gilbert. "Moral Philosophy, and Linguistics." In *Proceedings of The Twentieth World Congress of Philosophy*. Vol. I: *Ethics*, edited by Klaus Brinkmann. Bowling Green, Ohio: Philosophy Documentation Center, 1999.

Harman, Gilbert. *The Nature of Morality*. New York: Oxford University Press, 1977.

McNaughton, David A. *Moral Vision*. Oxford: Oxford University Press, 1988.

Nussbaum, Martha. *Love's Knowledge*. Oxford: Oxford University Press, 1990.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Rawls, John. "Outline of a Decision Procedure for Ethics." *The Philosophical Review* 60 (1951): 177–197.

Sencerz, Stefan. "Moral Intuitions and Justification in Ethics." *Philosophical Studies* 50 (1986): 77–95.

Singer, Peter. "Sidgwick and Reflective Equilibrium." *Monist* 58 (1974): 490–517.

Sosa, Ernest. "Equilibrium in Coherence?" In *Knowledge in Perspective: Selected Essays in Epistemology*. Cambridge, U.K.: Cambridge University Press, 1991.

Stich, Steven. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Stich, Stephen, and Richard Nisbett. "Justification and The Psychology of Human Reasoning" *Philosophy of Science* 47 (1980): 188–202.

Timmons, Mark. "Foundationalism and The Structure of Ethical Justification." *Ethics* 97 (1987): 595–609.

Weinberg, Jonathan, Shaun Nichols, and Stephen Stich. "Normativity and Epistemic Intuitions." *Philosophical Topics* 29 (2001): 429–460.

Michael R. DePaul (2005)

REFORMATION

In the narrower and probably most common sense, "Reformation" is the name given to the spiritual crisis of the sixteenth century that resulted in the permanent division of the Western church. The birthdate of the Reformation is traditionally given as 1517, the year in which Martin Luther posted his Ninety-five Theses on the door of the Castle Church in Wittenberg; the termination of the period may be assigned to the 1550s, by which time an ecclesiastical stalemate between the Protestants and the Roman Catholics appeared unavoidable. Sometimes the Reformation is extended backward to include such early reform movements as Lollardy or forward to include the religious conflicts, lasting into the seventeenth century, that sought to resolve the Catholic-Protestant stalemate forcibly or to readjust the divisions between the various Protestant groups. Reformation describes the aspirations of the age rather than its achievements. The Protestants did not succeed in reforming the church but only in splitting it into rival groups, each of which claimed for itself the fulfillment of the old dream of reformation in head and members.

THE AGE OF REFORMATION

The Protestant movement was not the only attempt to bring the dream into reality. It can, indeed, be correctly interpreted only in relation to other reform movements even if we determine not to include these under the same general descriptive label. The sixteenth century was the age of reformation (or of reformations, in the plural), not

just of the Reformation, and this is a fact of some importance in assessing the impact of the spiritual crisis on Western intellectual history. We should distinguish four reform groups in the sixteenth century, each of which left its own distinctive mark on Western culture.

HUMANISTIC REFORMERS. The humanists were not merely (as Luther himself thought) forerunners who prepared the way for the Protestants. They developed a reform program of their own that did not lead to the formation of independent institutions but continued, even after the appearance of Luther, to exercise influence from within both of the two main confessional groups. The foremost humanistic reformer in northern Europe was Desiderius Erasmus, who wished to purify the church by returning to its primitive sources—the New Testament and the writings of the Fathers. His "philosophy of Christ" minimized the dogmatic and the institutional and treated Christ mainly as a teacher of virtue and Christianity as an ethical affair not essentially different from the pagan philosophies. Although not less critical of ecclesiastical abuses than was Luther, Erasmus deplored any action that might disrupt the unity and peace of Christendom, and this was one of the reasons that he remained aloof from the Protestant Reformation.

RADICAL REFORMERS. "Radical reformers" is a general term for a variety of groups and individuals who felt that the Protestant leaders had not gone far enough and that reform could not be brought about without abandoning the old idea of the state church (the *corpus Christianum*). Of these radical or left-wing reformers, the Anabaptists (Swiss Brethren, Hutterites, and Mennonites) were biblical literalists who sought to establish voluntary associations of the regenerate on the New Testament pattern. The spiritualists (Andreas Carlstadt, Thomas Münzer, Sebastian Franck, Caspar Schwenckfeld), appealing to the Spirit who caused the Scriptures to be written, laid claim to immediate converse with God. The rationalists (notably the two Socinus) read the Bible in the light of reason even when reason led them to deny Christ's full deity and atoning sacrifice. A few of the radicals (for example, the leaders of the Münster uprising in 1534) were revolutionaries who brought total destruction upon themselves; many, like Michael Servetus, were free spirits who founded no school, but the influence of others, despite brutal persecution by Roman Catholics and Protestants alike, still survives in some present-day denominations and sects.

CATHOLIC REFORMERS. The Roman Catholics rejected the Protestant reform as essentially a revolt against the church, and they sought renewal of the church by the twofold means of fostering a churchly piety and taking an official stand on the administrative and dogmatic demands of the "heretics." Two of the greatest landmarks of the Catholic reformation were the establishment of the Jesuit order under the leadership of Ignatius Loyola and the work of the Council of Trent (1545–1563). The council, not without political and theological difficulties, sought to repudiate Protestant errors on authority, justification, and the sacraments. Yet the Tridentine fathers opposed many of the practical abuses and even theological inadequacies that had first provoked the Protestant movement. Preoccupation with Protestant errors, together with the militant campaign of suppression that followed the council, make it not inappropriate to speak of the Catholic reformation as the Counter-Reformation, though it was not merely this and had its roots in pre-Lutheran piety.

THE PROTESTANT REFORMATION. The Protestant leaders (the reformers in the narrower sense) were themselves not strictly a single group. Protestantism took three distinctive, though fundamentally related, forms.

Lutheranism, rooted in the religious struggles of Luther and his revolt against the papacy, prevailed in most of Germany and was wholly victorious in the Scandinavian countries. It was the Lutheran princes and cities represented at the Imperial Diet of Speyer in 1529 who, by making their historic protest, gave the Lutheran movement its nickname Protestantism. The classic formulation of Lutheran belief is the Augsburg Confession of 1530.

The so-called Reformed churches grew up first in Switzerland (under Huldrych Zwingli and John Calvin); won majorities in Scotland, Holland, and parts of Germany; and maintained strong pockets of influence in France, England (where they were called Presbyterians), and eastern Europe. From their beginning they were a less homogeneous group than the Lutherans and produced a variety of national confessions rather than a single statement comparable to the Lutheran Augsburg Confession. Nevertheless, the Lutheran interpretation of the Gospel exercised a decisive influence over the Reformed confessions, and though Zwingli sought to affirm his relative independence from the Germans, Calvin was one of Luther's staunchest admirers.

The Anglican reformation proceeded slowly, largely for political reasons. The repudiation of papal authority by Henry VIII, though not intended to alter Catholic doc-

trine, left the door open to Protestant reform in the reign of his son Edward VI, and the Romanizing reaction under Mary only temporarily reversed the trend. The Thirty-nine Articles of Religion (Latin 1563, English 1571), adopted under Elizabeth I as the official doctrinal standard of the reformed Church of England, are largely a compilation of Continental Protestant ideas. Parts of the Lutheran confessions of Augsburg and Württemberg are reproduced verbatim, and the articles on predestination and the Eucharist are clearly indebted to Reformed (Calvinistic) theology.

ESSENTIAL PROTESTANT DOCTRINES

In all three of its branches the Protestant Reformation was inextricably bound up with social and political factors, so that its triumph was always, in the final analysis, contingent on governmental support. Nevertheless, it was essentially a religious movement and its theological ideas have left their mark on European intellectual history—sometimes, however, because they have been misinterpreted or interpreted too one-sidedly. Three beliefs are particularly associated with the Protestant movement: the authority of the Word, justification by faith alone, and the priesthood of all believers. These beliefs have frequently been explained as the advent of individualism in the religious sphere, as though the intention were to regard the individual as his own priest with immediate access to God, to leave him in solitude with his conscience and his Bible, or to make each man his own pope in the interpretation of Scripture.

Fundamentally, however, the original Protestant reformers were suspicious of "immediate access to God," which they associated with the spiritualists, and they sought, rather, to replace the medieval notion of institutional means with a concept of the Christian fellowship as the locus of God's Word. The Word of God was understood chiefly as an effective proclamation of the Gospel, based on the Scriptures, which evokes faith and sustains a fellowship of believers each of whom is priest to his brothers. The heart of this proclamation is the promise of free forgiveness (justification) through Christ, which needs only to be accepted by the faith that is awakened through the proclamation itself. We may perhaps add a fourth idea of great religious and even social consequence: vocation—that the good works required of the justified man are not so much special religious acts as the thankful performance of his calling for the good of his neighbor. These four ideas were held in common by all three Protestant groups, and their formulation may be traced to Luther himself. Characteristic differences among the groups also developed; for example, the Reformed differed from the Lutherans, as is well known, on the manner in which the benefits of Christ's Passion are received in the Eucharist.

THE REFORMATION AND WESTERN THOUGHT

The Reformation's role in the making of the modern mind is a complex question that has ramifications in areas as diverse as social, economic, political, and artistic history as well as in the history of philosophy and science. Sometimes the Reformation has been represented as the great watershed between the medieval and modern worlds. This is, perhaps, partly because the individualism of Reformation thought has been overestimated and partly because certain isolated events in Luther's life—the burning of the papal bull, the defiant stand before the Diet at Worms—have deeply impressed themselves on the German imagination. In some respects, however, the Reformation can be better understood as a late phase of medieval history than as an early stirring of the modern mind. The fundamental concerns of Luther were medieval, and it may be argued that in giving fresh vitality to religious questions he merely postponed for a while the triumph of Renaissance secularism. Moreover, though the Protestant reformers spoke ideally of a communion of saints (believers), in practice they refused to abandon the medieval concept of a Christian society (that is, an authoritarian, church-dominated society).

Unquestionably, the very existence of the Protestant churches alongside the Roman Catholic Church weakened the authoritarian ideal. But this was an accidental product of the Reformation—a consequence, indeed, of its failure rather than of its cherished principles. It was the humanistic reformers, not the Protestants, who undermined the dogmatic conception of religion, and it was the radicals who broke with the old alliance between the spiritual and secular arms of the *corpus Christianum*. Similarly, if, as has been argued, Calvinistic ideas had revolutionary economic and political consequences, this was hardly the reformer's intention. On the other hand, the Reformation did, by its very nature, make a powerful impact on literature and music, education and scholarship; even its influence on the visual arts was not always uncreative.

REFORMATION AND SCIENCE

The chief contribution of the Reformation to the history of Western philosophy was no doubt the accidental one of helping philosophy toward autonomy by weakening

ecclesiastical domination. Attempts to establish the influence of Lutheran ideas on some of the German philosophers are often interesting but seldom of very great importance and sometimes farfetched. It might have been a service for philosophy, as it was for theology, that Luther shattered the medieval synthesis of Christianity and Aristotelianism, but the reformer's immediate successors reinstated the Greek philosopher, and the Christian faith was perilously entangled in an obsolete cosmology. (Ironically, Philipp Melanchthon repudiated Copernican astronomy on the ground that it represented merely a revival of outmoded theories that had already been rejected in the ancient world.)

Luther himself prepared the way for the conflict of theology and the modern worldview by refuting a scientific theory on theological grounds—if, indeed, the notorious passage from the *Table Talk*, "Joshua commanded the sun, not the earth, to stand still," is authentic. Yet an open clash of science and religion was not unavoidable until post-Reformation theologians in the age of Protestant scholasticism had reaffirmed the old partnership with Aristotelianism and had come to think of the Scriptures as containing a "biblical science" that could compete with Copernican science. Luther and Calvin themselves did not accept the Ptolemaic cosmology in defiance of scientific evidence since the weight of the evidence during their lifetimes was still against Nicolas Copernicus. In principle, they were not suspicious of scientific progress. On the contrary, Luther welcomed the stirring of the new science, in which he saw a partial recovery of Adam's lost dominion over nature, and Calvin envied the astronomer's closeness to the mind of the Maker. They were both interested in the Bible not as an encyclopedia of supernaturally communicated information but as the vehicle of Christ's presence to his church in the Gospel proclamation.

Luther had grasped clearly that theological and scientific interest in nature are two distinct things. For example, from the religious viewpoint the light of the moon was for him a symbol of divine care, but he recognized that the astronomer's concern was to show how the moon's light was borrowed from the sun. Similarly, Calvin argued that biblical observations on the heavenly bodies, such as those in Genesis and the Nineteenth Psalm, are not scientific statements but homely forms of speech accommodated to the unlearned. Luther understood even better than Calvin that theology's heaven is not the same as the astronomical heavens; hence, the celebrated *Dextera Dei est ubique* (God's right hand is everywhere). Elementary though they may seem today, such concessions and insights, had they not been neglected or expressly repudiated by Protestant orthodoxy, could have saved the Reformation churches from their warfare with science. Conversely, they might have prevented skeptics from drawing overhasty theological conclusions from natural science.

See also Aristotelianism; Calvin, John; Copernicus, Nicolas; Erasmus, Desiderius; Franck, Sebastian; Humanism; Luther, Martin; Melanchthon, Philipp; Servetus, Michael; Socinianism.

Bibliography

Studies of the Reformation include Roland H. Bainton, *The Reformation of the Sixteenth Century* (Boston: Beacon Press, 1952); Harold J. Grimm, *The Reformation Era* (New York: Macmillan, 1954); G. R. Elton, ed., *The New Cambridge Modern History*, Vol. II, *The Reformation 1520–1559* (Cambridge, U.K.: Cambridge University Press, 1958); Émile G. Léonard, *Histoire générale du protestantisme*, Vol. I, *La réformation* (Paris, 1961); H. A. Enno van Gelder, *The Two Reformations in the 16th Century: A Study of the Religious Aspects and Consequences of Renaissance and Humanism* (The Hague, 1961); George Huntston Williams, *The Radical Reformation* (Philadelphia: Westminster Press, 1962); and H. Daniel-Rops, *The Catholic Reformation*, translated by John Warrington (New York: Dent, 1962).

The Augsburg Confession, the Thirty-nine Articles, and a representative Reformed confession will be found in *The Faith of Christendom: A Source Book of Creeds and Confessions*, edited by B. A. Gerrish (New York, 1963). For the religious ideas of the reformers there are excellent bibliographies in the studies by Grimm and Léonard.

On the special question of the Reformation in relation to Western culture and thought, see Heinrich Bornkamm, *Luther's World of Thought*, translated by Martin H. Bertram (St. Louis: Concordia Publishing House, 1958); John Dillenberger, *Protestant Thought and Natural Science: An Historical Interpretation* (Garden City, NY: Doubleday, 1960); Karl Holl, *The Cultural Significance of the Reformation*, translated by Karl Hertz and Barbara Hertz and John H. Lichtblau (New York: Meridian, 1959); Jaroslav Pelikan, *From Luther to Kierkegaard: A Study in the History of Theology* (St. Louis: Concordia Publishing House, 1950); and Ernst Troeltsch, *Protestantism and Progress: A Historical Study of the Relation of Protestantism to the Modern World*, translated by W. Montgomery (London: Williams and Norgate, 1912).

OTHER RECOMMENDED WORKS

Brady, Thomas A., Jr. *The Protestant Reformation in German History*. Washington, DC: German Historical Institute, 1998.

Caponetto, Salvatore. *The Protestant Reformation in Sixteenth-Century Italy*. Translated by Anne C. Tedeschi and John Tedeschi. Kirksville, MO: Thomas Jefferson University Press, 1999.

Gonzalez, Justo L. *A History of Christian Thought*. Nashville: Abingdon Press, 1970–1975.

Gray, Madeleine. *The Protestant Reformation: Belief, Practice, and Tradition*. Brighton: Sussex Academic Press, 2003.
Harrison, Peter. "Original Sin and the Problem of Knowledge in Early Modern Europe." *Journal of the History of Ideas* 63 (2) (2002): 239–259.
Hillerbrand, Hans Joachim, ed. *The Protestant Reformation*. New York: Walker, 1968.
Johnston, Andrew. *The Protestant Reformation in Europe*. New York: Longman, 1991.
Kourvetaris, George A. *Social Thought*. Lanham, MD: University Press of America, 1994.
Painter, Mark A. *The Depravity of Wisdom: The Protestant Reformation and The Disengagement of Knowledge from Virtue in Modern Philosophy*. Brookfield, VT: Ashgate, 1999.
Spitz, Lewis William, ed. *The Protestant Reformation*. Englewood Cliffs, NJ: Prentice-Hall, 1966.
Spitz, Lewis William. *The Protestant Reformation, 1517–1559*. New York: Harper & Row, 1985.
Steinmetz, David Curtis. *Luther and Staupitz: An Essay in the Intellectual Origins of the Protestant Reformation*. Durham, NC: Duke University Press, 1980.
Thompson, Stephen P., ed. *The Reformation*. San Diego: Greenhaven Press, 1999.

B. A. Gerrish (1967)
Bibliography updated by Tamra Frei (2005)

RÉGIS, PIERRE-SYLVAIN
(1632–1707)

Pierre-Sylvain Régis was a student of the Cartesian physicist Jacques Rohault. Like Rohault, Régis expounded Cartesianism in public lectures. In 1680 François de Harlay de Champvallon, the archbishop of Paris, told Régis that King Louis XIV forbad public lectures for fear of uproar concerning the Cartesian explanation of transubstantiation. Régis continued to give private lessons, and by 1699 the conflict over Cartesianism had subsided, leading to his admission to the Academie des Sciences, along with Nicolas Malebranche, whose occasionalist philosophy was a response to problems of Cartesian dualism.

Régis's system is based on fourteen self-evident metaphysical principles derived from the *cogito* (Descartes' basic axiom: "I think, therefore I am").

(1) All properties belong to something, that is, nothing can have no properties

(2) All effects presuppose causes

(3) An effect can have no more perfection than does its total cause

(4) All changes in a subject proceed from an external cause

(5) All modes presuppose substances in which to exist

(6) A mode that modifies one kind of substance cannot modify any other kind of substance

(7) All that exists is either a substance or a mode

(8) Essences are indivisible; if anything is added to or subtracted from an essence, it is destroyed

(9) Privations and negations are known only by their opposites

(10) External things are known only by way of ideas; what one has no idea of is to one as though nonexistent

(11) All ideas, to the extent that they represent properties, depend on their objects as exemplary causes

(12) The exemplary cause of an idea contains formally all the perfections the idea represents

(13) Facts attested by many people of diverse times, nations, and interests as known in themselves and of which one cannot suspect conspiracy to support a lie should be accepted as constant and indubitable as though one had viewed them oneself

(14) Witness of infinitely powerful, wise, good, and truthful God should have as much persuasive force on one's mind as the most convincing reasons

The ontological principles are to the effect that whatever exists is either a substance or a modification of a substance and that modifications cannot belong to nothing. The basic epistemological principle is that external material bodies can be known only by way of representational ideas that are themselves mental modifications of the mind. The central principle is that all effects presuppose causes that must have as much or more perfection than their effects. All modifications are effects and thus are ultimately caused by a substance. All ideas (both sensory images and intelligible concepts) are also effects caused by some substance, and all of them represent the perfections of their causes. The basic problem with these assertions is that mind whose essence is active unextended thinking is essentially unlike matter whose essence is passive unthinking extension. Thus, all Cartesians must face the two questions first posed to René Descartes by Princess Elisabeth of Bohemia: Given that Cartesian mind and matter are essentially unlike one another, how can they interact causally? And how can a mind know extended matter by way of unextended sensations and ideas? Régis answers as Descartes does that one knows causal interaction takes place and God can make this hap-

pen, even if one does not understand how. The question of how mental ideas can represent material bodies that are essentially unlike mental things is also settled in the Cartesian way with the assertion that God makes it so, even if one does not understand how.

Régis maintains that his metaphysics, logic (which follows Antoine Arnauld's), and ethics (based on self-interest) are certain and complete. Physical explanations are also based on self-evident principles, but those humans give are only probable because, on the principles of Cartesian physics, several explanations are deducible for each event, and one does not know which one God chose. Régis says that the simplest is the most probable. If one had complete knowledge of the Cartesian deductive system as God does, however, one would have certain knowledge in physics, and one should keep that goal in mind. But the search must be made systematically. Régis, like Descartes, opposes "arbitrary hypotheses," explanations not deducible from self-evident principles within a system. Ad hoc explanations that are not part of a comprehensive theory are useless. Régis, like Descartes and all later Cartesians, believes the correct theory is mechanism.

The most distinctive feature of Régis's Cartesianism is his doctrine that man is a compound substance. In this union, eight conditions pertain:

(1) The soul always has the idea of extension

(2) Specific brain movements cause specific ideas

(3) Animal spirits cause motions in the brain that give rise to imagination, sensation, and memory of material objects

(4) Pleasure and pain are signs of bodies suitable and unsuitable to the human body

(5) Man has a penchant to love or hate, and pursue or flee, the objects of pleasurable or painful ideas

(6) Sentiments and passions lead to actions of the body toward self-preservation

(7) The soul thinks of particular bodies only when particular brain movements occur

(8) The union holds only so long as the body is alive and functions properly

Because ideas must have existing exemplary causes, Régis argues contrary to Descartes that one knows both the essence and existence of both mind and matter.

Man, Régis explains (as did Descartes' Dutch disciple Regius [Henry de Roy]), is an accidental union of mind or soul and body. Descartes himself adamantly opposes this view by insisting that the union is substantial, not accidental. If the union is only accidental, this makes the mind or soul a property of the body, not a substance on equal standing with the material substance. Then when the body dies, this accidental mind or soul would disintegrate with the other bodily properties. Régis argues that in fact the soul disintegrates and the mind survives, but it cannot think temporally—because this depends on bodily motions—and instead can contemplate only itself and God.

Like Descartes, Régis shows how operations of the body take place through actions of external bodies on sense organs to cause movements in the brain. He admits that causal interaction between mind and body is inexplicable and can be accepted only on faith. It is a brute fact that because of its union with a body, a mind or soul has the idea of extension and can cause that body to move. And because of the union, distinctive brain movements always give rise in the mind to distinctive sensations and concepts of the material objects affecting the brain.

Descartes asserts that one is born with innate ideas of mind, God, and matter. But Régis says that all ideas, even of God, depend on brain movements. Thus for Régis, after separation of mind and body at death, the mind no longer has the idea of extension, and no imagination or memory of, or power over, the material world.

Régis claims that Pierre-Daniel Huet thinks Descartes is a skeptic because Huet does not distinguish methodological from real doubt. Also, Huet is wrong to argue that Descartes' explanation of transubstantiation does not preserve the body of Christ in the sacrament and to claim that because Descartes believes God has the power to do anything, one has no certain knowledge of one's world. Régis shows that Malebranche's theory of seeing all things in God requires an impossible union of man with God. Jean Du Hamel is accused of failing to see that mental ideas that do not resemble their material objects still make these objects known. Régis insists that Benedict (Baruch) de Spinoza fails to see that God is not an ordinary substance and thus confuses the material world with God.

Like Rohault, Régis insists that reason and faith do not conflict. Reason is infallible in the order of nature; faith, in the order of grace. Events in one order cannot be explained with principles of the other. Thus, Régis offers no physical explanation of transubstantiation—as do Descartes, other Cartesians, and the scholastic physicists. He argues that transubstantiation is an event not in the order of nature but in the order of grace.

See also Arnauld, Antoine; Cartesianism; Descartes, René; Elisabeth, Princess of Bohemia; Essence and Existence; Huet, Pierre-Daniel; Malebranche, Nicolas; Regius, Henricus (Henry de Roy); Rohault, Jacques; Spinoza, Benedict (Baruch) de.

Bibliography

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Watson, Richard A. *The Breakdown of Cartesian Metaphysics*. Indianapolis: Hackett, 1998. Originally published as *The Downfall of Cartesianism, 1673–1712: A Study of Epistemological Issues in Late Seventeenth-Century Cartesianism* (The Hague: Nijhoff, 1965).

Richard A. Watson (1967, 2005)

REGIUS, HENRICUS (HENRI DE ROY)

(1598–1679)

Regius (Henri de Roy), a Dutch academic, was a major figure in disputes over Cartesianism in Utrecht, The Netherlands, during the seventeenth century. Regius received a medical degree from the University of Padua before returning in 1638 to his hometown of Utrecht to become a professor of medicine and botany at the university. Before his appointment there he gave private lectures based on the *Dioptrics* and *Meteors*, two of the essays published with René Descartes's *Discourse on the Method* (1637). What Regius found to be particularly congenial in these texts was the proposal there that observable phenomena be explained in terms of the mechanical properties of insensible material parts.

In 1641 Regius took advantage of his good relations with the new rector of the university, Gisbertius Voetius, to obtain permission to submit for discussion various "medical disputations." The first two disputations provide a mechanistic reinterpretation of Aristotelian notions, but in the third disputation Regius took the more aggressive tack of claiming that the union of the soul and body is not substantial, as the Aristotelians claimed, but accidental. Voetius responded with an appendix that defended Aristotelianism against the "new philosophy," and on the advice of Descartes, Regius offered a response that suggested that the Aristotelians had difficulty avoiding atheism. In 1642 the burgomasters of Utrecht ordered the confiscation of Regius's response and endorsed a statement by the faculty that condemned the teaching of the new philosophy. Descartes intervened by publishing attacks on Voetius in 1642 and 1644 that the burgomasters judged to be libelous. Fearful of imprisonment, Descartes sought the protection of the French ambassador, who succeeded in suppressing his arrest warrant.

To this point, Descartes had a favorable opinion of Regius. However, matters took a turn for the worse in 1645 when Regius sent Descartes a draft of his *Fundamenta physices*. Descartes was shocked by the assertion in one section of this text that it is impossible to prove that the soul is anything more than a mode of body. When Regius went ahead and published his text in 1646, Descartes denounced it in the preface to the French edition of the *Principles* (1647). A student of Regius published a broadsheet that highlighted Regius's rejection not only of a proof of immortality but also of innate ideas and the possibility of a proof of the existence of the material world. Descartes responded in 1648 with his *Notes against a Broadsheet*, and Regius replied that same year with his *Brevis explicatio mentis humanae*. Regius's text included a letter from Petrus Wassenaer defending Regius against the charge in the preface to the *Principles* that he had plagiarized portions of Descartes's unpublished treatise on animals. After Descartes's death Regius published second and third editions (1654 and 1661, respectively) of the *Fundamenta physices* with the new title *Philosophia naturalis*, in which he attempted to further defend the project of freeing mechanistic physics and physiology from dogmatic metaphysics, on which Descartes had attempted to found it. Regius's presentation there of the new science as a system of probable hypotheses is similar to the one found in the *Traité de physique* (1671) of the French physicist Jacques Rohault, perhaps the most influential defense of Cartesian physics in the century following Descartes's death.

See also Aristotelianism; Cartesianism; Descartes, René; Philosophy of Physics; Rohault, Jacques.

Bibliography

Dijksterhuis, E. J., et al. *Descartes et le cartésianisme hollandais*. Paris: Presses Universitaires de France, 1950.

WORKS BY REGIUS

Fundamenta physices. Amsterdam: Elsevier, 1646.

Brevis explicatio mentis humanae sive animae rationalis … a Notis Nobil. Cartesii vindicata. Utrecht, Netherlands: N.p., 1648.

Philosophia naturalis. Amsterdam: Elsevier, 1654.

WORKS ABOUT REGIUS

Rothschuh, Karl E. "Henricus Regius und Descartes: Neue Einblicke in die frühe Physiologie (1640–1641) des Regius." *Archives internationals d'Histoire des Sciences* 21 (82–83) (1968): 39–66.

Verbeek, Theo. *Descartes and the Dutch: Early Reactions to Cartesian Philosophy, 1637–1650*. Carbondale: Southern Illinois University Press, 1992.

Verbeek, Theo, ed. *Descartes et Regius: Autour de l'explication de l'esprit humain*. Amsterdam: Rodopi, 1993.

Tad M. Schmaltz (2005)

REHMKE, JOHANNES

(1848–1930)

Johannes Rehmke, the German epistemologist, ontologist, and ethical philosopher, was born at Elmshorn in Schleswig-Holstein. He studied evangelical theology and philosophy at Kiel and Zürich from 1867 to 1871, receiving his doctorate in philosophy at Zürich in 1873. After some years as a high school teacher at St. Gallen, Rehmke was appointed unsalaried lecturer in philosophy at the University of Berlin in 1884. The following year he became professor of philosophy at the University of Greifswald, where he taught until 1921.

THEORY OF KNOWLEDGE

Rehmke did not assume the existence of two worlds: a world, only indirectly knowable, of transsubjective objects, and an immediately knowable world, with intrasubjective perceptions and the like as contents. Rather, he asserted the existence of directly knowable real objects. This epistemological monism was a consequence of his ontological dualism of two essentially different kinds of being. Physical (material) beings are spatially extended and occupy a place; mental (immaterial) beings are not extended and have no place.

The nonspatial, placeless character of consciousness conflicts with the uncritical application to the subject of such concepts as "in" and "external," as exemplified in such terms as "intrasubjective" and "transsubjective"—in other words, "immanent" and "transcendent," or "content of consciousness" and "external object." Not only does consciousness not involve the having of a content; it does not involve any kind of having by means of a relation, in any event one that presupposes the existence of at least two realities separated from one another. On the contrary, knowing without any relation between diverse things is possible from the outset, as can be seen in self-consciousness. In self-consciousness only one thing is given, the particular knowing consciousness as knowing itself and as being known by itself. Thus, Rehmke's proposition "Knowing is having without a relation" expresses the immediacy of all knowledge, including knowledge of the so-called external world, the world of objects outside the body.

In his *Logik oder Philosophie als Wissenslehre* (Leipzig, 1918), Rehmke sought to demonstrate the importance of the general or universal for the movement of knowledge toward clarity. In accord with his proof of the immediacy of cognition, he rejected as false the notion that thinking is an internal, that is, intramental, activity and even rejected the notion of thought activity because the purported activity never produces a change in objects. Thinking is not a "doing" but a "finding." If, for example, someone makes the judgment "A boiled crayfish is red," this observation signifies that he as thinker finds anew in the object the red known before. What is thus discovered in the object is never something single, an individual being, but something repeated, a universal.

Because the universal forms part of each particular object, it is something objective. If red is found in the crayfish, the logical subject of the judgment is not simply "(boiled) crayfish," but "red boiled crayfish." Consequently, every judgment, with respect to the universal discovered in the particular object, is logically analytic. Grammatically, with regard to the joining of the linguistic signs into a sentence, it is synthetic.

In its function as predicate of a judgment, an objective universal is called a concept. Every concept is thus a universal. Because of its objectivity, the universal as concept, despite its relation to the thinking subject, cannot be merely subjective. It is equally erroneous to confuse or to equate the concept, which is always bound up with a particular word, with that word, that is, with the phonic structure as linguistic sign.

The objectivity of the universal as a possible concept reveals the error in the phrase "concept formation." A concept (for example, "tree") is not first constructed by comparing several objects (for example, pines, beeches, and alders) by means of an "internal activity" of thought. The concept is presupposed in the very selection of objects of the same kind. Concept formation is really conceptual clarification, the determination of which characteristics in union constitute a concept already given. Clarity is the guiding notion in Rehmke's logic. He claimed that, in any deepening of knowledge, the universal as logical predicate helps consciousness to obtain clarity, and ultimately unquestionable clarity.

Rehmke's conception of logic, that is, philosophy as theory of knowledge, is linked with his notion of philosophy as fundamental science, expressed in his *Philosophie als Grundwissenschaft* (Frankfurt, 1910). Both theory of knowledge and fundamental science are genuine sciences, directed toward that which is simply given, that is, toward objects regardless of their being real or unreal. They are also in equal measure philosophy because they deal with the totality of the given, in contrast with the particular sciences, each of which deals with only a particular section of the world. Theory of knowledge deals with the given as that which is thought (known); fundamental science deals with it in regard to its most universal character. But while logic presupposes the concept "universal," and each special science presupposes its own fundamental concepts, the task of philosophy as fundamental science is to elucidate without prejudice precisely the basic "that which is most universal."

THE TRADITIONAL ONTOLOGY

Theory of knowledge is not a fundamental science. Historically, it arose from an epistemological dualism, and as a consequence its form is faulty. In any case, it must presuppose the basic distinction between knower and other. Rehmke's painstaking ontological studies in *Philosophie als Grundwissenschaft* of the manifold "most universal" embrace five paired notions: (1) matter and consciousness, (2) the universal and the unique, (3) unity and simplicity, (4) the changeable and the unchangeable, (5) the real and the unreal. For Rehmke, of course, the first pair was primary. Beyond the merely negative description—immaterial, nonspatial, and place-less—the essence of the mental is completely determined by the concept of consciousness, or knowledge. Rehmke therefore opposed both materialism and idealism (spiritualism), as well as Spinozism.

Everything without exception proves to be either a unique thing (something that occurs only once, such as a unique tree) or a universal (something that is repeated, such as green or "treeness"). It follows that the unique and the universal do not exist without each other; indeed, objectively the universal belongs to the unique. Rehmke classified the unique into individuals (for example, individual trees) and units of individuals. He divided the latter into operational units (for example, an auto with a trailer) and living units (for example, a state). The universal is either a determination (such as angularity) or a relation (such as similarity). Rehmke attached great value to his recognition that many seemingly ontological concepts, such as space, time, being, and value, are merely relational ones.

In connection with the third of his five pairs, unity and simplicity, Rehmke distinguished between individuals that are composed of individuals (and hence are ephemeral, passing) and individuals that are absolutely simple (and hence are everlasting). Examples of the latter are elementary particles and consciousness. Denying the theory of substance, he held that the individual is a union of its determinations (a body, for instance, is a union of size, shape, and location). He also analyzed each specific determination into determination as such (for instance, shape as such) and particularity.

Rehmke equated the fourth relationship, the changeable and the unchangeable, with the distinction between individual and universal. In this context he pointed out that the concept of change refers only to exchange of individual characteristics, with the determination as such (for instance, the shape as such) remaining the same.

Rehmke treated in detail the relationship between the real and the unreal. He defined the real as consisting in relationship of action. This enabled him to do justice to such properties of things as sweetness, which are often dismissed as merely subjective.

PSYCHOLOGY AND ETHICS

In his *Lehrbuch der Allgemeinen Psychologie* (Frankfurt, 1894), Rehmke stressed that human consciousness (mind) is a simple, immaterial individual being, in a constant unity of action with an essentially different body. Thus, man is not a "double-beinged" individual. There are four general characteristics of consciousness: (1) determination of objects, each one directly perceived or imagined, even though the perception is mediated by the sense organs; (2) states (conditions), for example, delight or listlessness; (3) thought—either distinguishing (being aware of the distinct) or uniting (awareness of unity); (4) the subject, the determination of which establishes at the same time the unity of the ego. These determinations are not to be construed as mental activities.

Because of its intermittent character, volition, despite its relations with the above determinations, is not one of them. Rehmke's analysis of volition aided him in his solution of the problem of free will. He separated the problem into four parts, each of which is answerable: (1) Is an act of the will prevented or not? (2) Is the volition random or conditioned? (3) Is there a genuine possibility of choice, or is the will constrained? (4) Is the volition freely self-determined or not?

Rehmke's theory of the will constitutes the background for his ethics. He distinguished five forms of ethics—four false and one genuine. The ethics of shrewdness has to do with men "for themselves." The ethics of the unity of control expresses duty as an "ought." The ethics of the unity of life expresses duty as a "must" and comprises the ethics of society (in which unity as "being with one another" is a means to a selfish end) and the ethics of community (in which unity as "being for one another" is an end in itself). Finally, separating the merely social from the moral proper is the ethics of selfless love of one person as such "for another," arising from his knowledge of himself as at one with the other.

See also Being; Consciousness; Determinism and Freedom; Epistemology; Ethics, History of; Idealism; Knowledge and Belief; Matter; Ontology, History of.

Bibliography

ADDITIONAL WORKS BY REHMKE

Unsere Gewissheit von der Aussenwelt (Our certainty about the external world). Heilbronn, 1892.

Grundlegung der Ethik als Wissenschaft (Foundations of ethics as a science). Leipzig: Quelle and Meyer, 1925.

Die Willensfreiheit (The freedom of the will). Leipzig, 1925.

Gesammelte philosophische Aufsätze (Collected philosophical essays). Edited by Kurt Gassen. Erfurt: K. Stenger, 1928.

"Selbstdarstellung." In *Die Philosophie der Gegenwart in Selbstdarstellung,* edited by Raymund Schmidt, 7 vols., Vol. I, 177–200. Leipzig, 1921–1929.

For a full bibliography see *Grundwissenschaft. Philosophische Zeitschrift der Johannes-Rehmke-Gesellschaft* 1 (1919): 72–88, and 10 (1931): 36–44.

WORKS ON REHMKE

Heyde, J. E. *Grundwissenschaftliche Philosophie.* Leipzig, 1924.

Heyde, J. E. "Johannes Rehmke." *Grundwissenschaft. Philosophische Zeitschrift der Johannes-Rehmke-Gesellschaft* 10 (1931): 1–35.

Heyde, J. E. "Johannes Rehmke." *Zeitschrift für philosophische Forschung* (1947): 603–606.

Heyde, J. E. "Johannes Rehmke 1848–1930." *Philosophische Studien* 2 (1951): 260–271.

"Johannes Rehmke." In *Grundriss der Geschichte der Philosophie,* edited by F. Schneider, 319–329. Bonn, 1959.

Joh's Erich Heyde (1967)
Translated by Albert E. Blumberg

REICH, WILHELM

(1897–1957)

Wilhelm Reich was an Austrian psychiatrist and social critic. After serving in the Austrian army during World War I, Reich became a medical student. He obtained his M.D. from the University of Vienna in 1922 and worked for some time as assistant to Julius Wagner-Jauregg at the latter's psychiatric clinic. Even before his graduation Reich began practice as a psychoanalyst and soon came to occupy an influential position in the psychoanalytic movement. From 1924 to 1930 he conducted what came to be known as the Vienna Seminar for Psychoanalytic Therapy, the first organized attempt to devise a systematic and effective analytic technique.

Reich also founded and directed sex hygiene clinics among the industrial workers of Vienna and later, on a much larger scale, in Berlin and other German cities. During his years in Germany, Reich was a member of the Communist Party, and he attempted to integrate his work as a sex counselor within the broader revolutionary movement. Adolf Hitler's assumption of power forced Reich to flee to Denmark. His activities had always been viewed with suspicion by the leaders of the Communist Party, and Reich was finally expelled from the party after the publication of *Die Massenpsychologie des Faschismus* (Copenhagen, 1933), in which he repudiated the official communist theory about the nature of fascism and the factors leading to its victory in Germany. Also, by 1933 Reich's psychiatric views were so far removed from those of orthodox psychoanalysis that the Internationaler Psychoanalytischer Verlag handled and printed but did not "publish" (that is, refused its imprint to) the first edition of Reich's *Charakteranalyse*. The break with the psychoanalytic organization became official at the Lucerne conference of the International Psychoanalytic Association in 1934.

Attacks by orthodox psychiatrists made it necessary for Reich to leave Denmark for Sweden, but in Sweden too there was official hostility and suspicion. Reich therefore gladly accepted an invitation by the Norwegian psychologist and philosopher Harald Schjelderup to teach at the University of Oslo, where he also hoped to undertake various physiological experiments. Reich worked in Norway from 1934 to 1939. Among his students and patients at that time were the English educational reformer A. S. Neill, the American psychiatrist and pioneer in psychosomatic research T. B. Wolfe, and leading figures in Norwegian psychiatry, including Nic Hoel (Waal), Ola Raknes, and Odd Havrevold. The distinguished Norwegian novelist Sigurd Hoel was also closely associated with Reich at this time—in fact, he succeeded Reich as editor of the journal *Zeitschrift für politische Psychologie und Sexualökonomie.* In 1937 Reich became the victim of a campaign in sections of the Norwegian press. Although he

had a number of influential defenders and the government renewed his permit to stay in the country, he decided to move to New York City, where he resumed his psychiatric practice and trained numerous psychiatrists in the new technique that he had worked out during his stay in Scandinavia. Reich also lectured at the New School for Social Research from 1939 to 1941.

In the last years of his life Reich showed little interest in psychiatry, devoting all his energies to what he took to be his great discoveries in physics. In 1956 he was sentenced to two years' imprisonment for disobeying a government injunction. He died in Lewisburg Penitentiary in 1957. A brief account of the main events leading to Reich's imprisonment will be found in the last section of the present entry.

It will be convenient to distinguish three phases in Reich's career: (1) his work within the psychoanalytic movement, marked, however, by some significant departures from orthodox psychoanalysis—the rejection of symptom analysis in favor of what Reich called "character analysis," the orgasm theory, and the attempt to understand the social function of sexual repression and neurosis; (2) Reich's efforts to relate neurotic attitudes to their somatic foundation and the development of what he called "character-analytic vegetotherapy"—a technique that constituted a drastic departure from all that preceded it; and (3) his theories about orgone energy—Reich's claim to have discovered a form of energy that is found in the atmosphere and also in the living organism and which can be concentrated in various ways, including the "orgone accumulator." What Reich claimed during the third period is of no philosophical interest. If any of the assertions in question were true, they would be of great scientific interest; but, in fact, most professional physicists who have heard of the orgone theory have dismissed it as nonsense. In fairness to Reich it should be added that a really unbiased investigation of his physical theories remains to be undertaken.

We shall here be exclusively concerned with certain of the ideas advanced by Reich during the first two periods. Of interest to philosophers are Reich's views concerning the origin of religious and metaphysical needs, the relation between the individual and society and the possibility of social progress, and, above all, the implications of his psychiatry for certain aspects of the mind-body problem. It is regrettable that, partly because Reich's books and articles were not easily accessible and partly because the wild claims of his last years created widespread distrust of his entire work, the remarkable achievements of his second phase are relatively little-known. To those who are put off by the recent metaphysical and pro-religious trends in psychiatry, as exhibited in the vogue of existentialist psychoanalysis and in the metapsychological speculations of Carl Jung and various Freudian analysts, Reich's concentration on the somatic basis of neurotic disturbances and the sexual problems and longings of human beings will come as a pleasant and refreshing change.

THERAPEUTIC INNOVATIONS

The philosophically most interesting part of Reich's work is unquestionably what he called "the breakthrough into the vegetative realm," that is, his attempt to determine the physiological basis of neurotic phenomena. However, first we should briefly describe Reich's earlier psychiatric work. In the early 1920s Freudian psychiatrists practiced what in retrospect came to be known as "symptom analysis." Neurotic symptoms were regarded as foreign bodies in an otherwise psychologically healthy organism; they are expressions of a repressed infantile drive that has reappeared in a disguised form. The task of therapy is to eliminate the repression: The symptom is removed by bringing the repressed part of the personality into harmony with the rest of the ego. By his own account, Reich soon became dissatisfied with this approach. The traumatic experiences leading to repression and the repressed drives were to be elucidated by means of free association and dream interpretation, but in fact only very few patients were capable of giving their associations free rein. Furthermore, Reich was critical of the superficial criteria of "cure" current at that time. Patients were considered "cured" upon the disappearance or alleviation of the symptom of which they had complained. However, Reich believed that the elimination of symptoms is quite compatible with the continuation of a character disturbance. Also, he questioned the existence of "monosymptomatic neuroses"—neuroses with only one serious symptom. "There are no neurotic symptoms," he later observed, "without a disturbance of the total character. Neurotic symptoms are, as it were, nothing but peaks of a mountain chain representing the neurotic character" (*The Function of the Orgasm*, p. 16). It was Reich's contention that, unless the characterological basis of a symptom has been eliminated, it or some equally troublesome symptom is likely to reappear.

On the few occasions on which either Reich or his associates at the Vienna Seminar appeared to achieve impressive and lasting improvements, this was invariably the result of the release of powerful dammed-up emotions like rage and hatred. Some years earlier, while work-

ing in Wagner-Jauregg's clinic, Reich had been struck by a catatonic who suddenly abandoned his stupor. "It was one great discharge of rage and aggression," Reich writes. "After the seizure had subsided he was clear and accessible. He assured me that his explosion had been a pleasurable experience, a state of happiness. He did not remember the previous stuporous phase.… It was very impressive, and could not be explained on the basis of the psychoanalytic theory of catatonia" (*The Function of the Orgasm,* pp. 43–44). Neurotics, too, showed noticeable improvement only when, instead of merely achieving an intellectual recognition of a repression, the impulse or emotion in question could actually be experienced. Such "liberations" were, however, infrequent and, what is more, they occurred more or less accidentally. An effective therapy would have to bring them about in a controlled fashion.

THE "CHARACTER ARMOR." Something should be said at this stage about Reich's concept of the "character armor" that came to play a central role in the technique of character analysis with which he gradually replaced the technique of symptom analysis. This concept was originally introduced in connection with certain cases of compulsion neurosis. Sigmund Freud had shown that compulsion symptoms always bind anxiety. If such a symptom is disturbed, the anxiety frequently appears. It does not, however, always appear—anxiety cannot usually be released in this way either in compulsion neuroses of long standing or in cases of chronic depression. Such patients appeared quite inaccessible. "Emotionally blocked compulsive characters gave associations in great numbers freely, but there never was a trace of affect. All therapeutic efforts bounced back, as it were, from a thick, hard wall" (*The Function of the Orgasm,* p. 114). These patients were "armored" against any attack. Over the years they had developed a set of attitudes whose function was to protect the individual against external injury (such as being hurt or rejected by other human beings) and to protect him against feeling his own repressed emotions, especially (though not exclusively) various kinds of destructiveness.

Reich introduced the term *character armor* to refer to the totality of the typical or chronic attitudes of this kind characterizing a given individual. It is, writes Reich, "as if the affective personality put on an armor, a rigid shell on which the knocks from the outer world as well as the inner demands rebound. This armor makes the individual less sensitive to unpleasure, but it also reduces his libidinal and aggressive motility and, with that, his capacity for pleasure and achievement" (*Character Analysis*, p. 310). Patients who do not suffer from a severe compulsion neurosis (and indeed most people growing up in a repressive environment) also have a character armor, but in their cases it can usually be attacked or broken down more easily.

The technique used to attack the character armor emphasizes the so-called negative transference. According to Reich, every patient has a deep mistrust of the treatment and feels strong hostility to the psychiatrist. Although patients wish to be cured, they also resent any attempt to disturb their "neurotic equilibrium." It is tempting for the analyst to shy away from these negative reactions, since it takes a great deal of strength and composure to bear the often furious hatred that is released when the armor begins to "crack." Nevertheless, it is precisely this negative reaction that can and must be used as the foundation of the treatment. The patient must feel free to criticize the analyst, and any attitudes that mask his hostility have to be broken down. Reference to the case of a "passive-feminine young man with hysterical symptoms" may give some idea of what this technique is like. The patient was excessively polite and, because of his fears, extremely sly. He always yielded and produced abundant material, but without any inner conviction. "Instead of discussing this material," Reich reports,

> I only kept pointing out his politeness as a defense against me and any really affective insight. As time went on, his hidden aggression appeared increasingly in his dreams. As the politeness decreased, he became offensive. In other words, *the politeness had been warding off the hatred*. I let the hatred come out fully by destroying every defense mechanism against it. The hatred up to that time had been unconscious. Hatred and politeness were antitheses, and at the same time the over-politeness was a disguised manifestation of hatred. (*The Function of the Orgasm*, p. 117)

If in this way repressed emotions are released and the patient actually experiences them, it is unnecessary to persuade him that he "really," "unconsciously" feels this or that. "The patient no longer talked about his hatred, he felt it; he could not escape it as long as his armor was being correctly taken apart" (p. 146).

The armor, according to Reich, varies from patient to patient, depending on his individual history, and the technique of destroying it has to be fitted to the individual case. The armor may be viewed as consisting of several layers. These layers, in Reich's words, "may be compared to geological or archaeological strata which,

similarly, are solidified history. A conflict which has been active at a certain period of life always leaves its trace in the character, in the form of a rigidity" (pp. 121–122). The neurosis of each patient has a specific structure that corresponds to its historical development, but in reverse order: "that which had been repressed latest in childhood was found to lie nearest the surface" (p. 121).

Anger and hate are not the only emotions bound by the character armor. Although destructiveness has to be emphasized and liberated in the early stages of the treatment, eventually genuine love and tenderness that had to be suppressed will also be released. The destructiveness, in the last resort, is "nothing but anger about frustration in general and denial of sexual gratification in particular" (p. 124). Destructive tendencies are most frequently "reactions to disappointment in love or to loss of love." An organism that has been freed of its dammed-up destructiveness becomes once again capable of love. Reich referred to persons who are unarmored and who possess the capacity for love in the fullest sense as "genital characters"; and the goal of therapy is to change the patient's neurotic character into a genital structure. According to Reich, the "energy" that nourishes neurotic symptoms and various destructive attitudes can be adequately discharged only in fully satisfactory sexual intercourse. A person with a genital character, unlike the neurotic, possesses "orgastic potency." This Reich defined as "the capacity for surrender to the flow of energy in the orgasm without any inhibitions; the capacity for complete discharge of all dammed-up sexual excitation through involuntary pleasurable contractions … free of anxiety and unpleasure and unaccompanied by phantasies" (*The Function of the Orgasm*, p. 79).

An individual with a genital character has undisturbed contact with his own drives and with his environment and, as a consequence, he has no need for any of the endless variety of substitute contacts and substitute gratifications of the neurotic individual. He, too, may not succeed in achieving a happy existence, since this depends on a great many factors, not all of which are within his control, but he will at least not be hampered in his struggle for happiness by irrational and destructive emotions or by excessive respect for the institutions of a life-denying society.

Reich vigorously repudiated the suggestion that, either in his therapy or in his social philosophy, his goal was a world "containing nothing but pleasure." The function of the armor, he observed, is to protect against pain, and in breaking it down, Reich's therapy aimed at reestablishing the capacity to feel pain as well as pleasure. "Pleasure and joie de vivre are inconceivable without fight, without painful experiences and without unpleasurable struggling with oneself" (p. 173). The goal is not a positive "hedonic balance" that, for all one can prove to the contrary, might be more effectively achieved by a life of monasticism but "full vitality in all possible situations of life." The capacity to take happiness and to give love goes hand in hand with "the capacity of tolerating unpleasure and pain without fleeing disillusioned into a state of rigidity."

REPRESSIONS AND CHRONIC MUSCULAR RIGIDITIES. Reich was led to his study of what he calls the "physiological anchoring" of neurotic conflicts and traumatic experiences partly as a result of his fundamentally materialistic orientation and partly because of the special attention paid in his technique of character analysis to the *manner* in which patients talked and acted. It is a mistake, he said, to regard rage and love (or any other emotion) as events "in the mind." They *are* physiological processes, and if an emotion is repressed, there must be some physiological mechanism by whose means the energy in question is "bound." Furthermore, Reich was convinced that if an adult's neurotic character attitude is the result of childhood experiences, this can be so only if the person's organism has in some way been chronically altered. The employment of "theoretical" terms such as "Id" and "unconscious" can easily lead to pseudoexplanations in this context. To say, for example, that a repressed childhood conflict exerts its influence "from the unconscious" may call attention to a suspected causal relation between the childhood experience and the present difficulties of the individual, but beyond that it simply amounts to admitting that one does not know how the influence in question is exerted. On occasions, it is true, Freud himself said as much and expressed his hope that some day explanations in terms of unconscious conflicts would be given a physiological meaning. At other times, however, Freud treated his theoretical terms as if they designated real and eternally inaccessible entities; and many of Freud's followers, according to Reich, became metaphysicians whose theorizing was euphemistically labeled "metapsychology."

Perhaps of greater influence than these general reflections was Reich's interest in the "how" of the patient's communications. The infantile structure, Reich observes in one place, is "conserved" in what an individual does as well as in the way in which he acts, talks, and thinks (*Character Analysis*, p. 188). Elsewhere Reich explains that he made himself independent of the so-called fundamental psychoanalytic rule ("to say everything that comes to

mind"), since it was impracticable with most patients, and that instead he took as "point of attack not only what the patient said, but everything he presented, particularly the manner of his communications or of his silence. Patients who kept silent were also communicating, were expressing something that gradually could be understood and handled" (*The Function of the Orgasm*, p. 145). It became increasingly evident to him, Reich adds, that "the *form* of behavior and communications, was much more essential than what the patient related. Words can lie. *The mode of expression never lies.*"

Special attention to the "how" of a patient's behavior very naturally led to close observation of the changes in his organism during and after the release of repressed emotions. Reich's earlier clinical reports already contained remarks about the awkwardness and the rigid movements of certain types of patients. However, it was not until the early 1930s that he began to elucidate the precise role played by muscular rigidities in the binding of impulses and emotions that had to be suppressed. The following extracts describing the beginning of a treatment in 1933 will perhaps convey better than any definition what Reich meant by the "physiological anchoring" of affects:

> In Copenhagen in 1933, I treated a man who put up especially strong resistances against the uncovering of his passive-homosexual phantasies. This resistance was manifested in an extreme attitude of stiffness of the neck. … After an energetic attack upon his resistance he suddenly gave in, but in a rather alarming manner. For three days, he presented severe manifestations of vegetative shock. The color of his face kept changing rapidly from white to yellow or blue; the skin was mottled and of various tints; he had severe pains in the neck and the occiput; the heartbeat was rapid; he had diarrhea, felt worn out, and seemed to have lost hold. … *Affects had broken through somatically after the patient had yielded in a psychic defense attitude.* The stiff neck, expressing an attitude of tense masculinity, apparently had bound vegetative energies which now broke loose in an uncontrolled and disordered fashion. … It was the musculature that served this inhibitory function. When the muscles of the neck relaxed, powerful impulses broke through, as if propelled by a spring. (*The Function of the Orgasm*, pp. 239–240)

This and other cases led Reich to a systematic study of chronic muscular rigidities and their relation to neurotic character attitudes. He reached the conclusion that "every neurotic is muscularly dystonic and every cure is directly reflected in a change of muscular habitus" (*Character Analysis*, pp. 311–312). Chronic muscular rigidities or spasms are found all over the bodies of the patients: in the forehead, around the mouth and in the chin, in the throat, the shoulders, the chest, the abdomen, the pelvis and thighs, and many other places. The rigid expression in the eyes of many patients, their chronic "stare," is the result of a chronic rigidity in the lid muscles. The breathing of neurotic individuals is disturbed in comparison with the natural and free respiration of emotionally healthy people. Reich referred to the totality of these chronic muscular rigidities that an individual develops as the "muscular armor."

Reich emphasized that it is muscle groups rather than individual muscles that become spastic—muscle groups which jointly serve a certain function, for example, to suppress the impulse to cry. Not only do the lower lips become tense in this event but also "the whole musculature of the mouth, the jaw and the throat; that is, all the muscles which, as a functional unit, become active in the process of crying" (*The Function of the Orgasm*, p. 269). In discussing the spasms frequently found in the mouth, chin, and neck, Reich enlarges on the tensions set up by the stifling of impulses to cry:

> Many people have a mask-like facial expression. The chin is pushed forward and looks broad; the neck below the chin is "lifeless." The lateral neck muscles which go to the breastbone stand out as thick cords; the muscles under the chin are tense. Such patients often suffer from nausea. Their voice is usually low, monotonous, "thin." (*The Function of the Orgasm,* p. 271)

This is not the place to discuss in detail other of the typical rigidities that make up the muscular armor. The interested reader will find these described in various of the publications devoted to the new technique.

Upon discovering the muscular spasms and their relation to suppressed impulses and emotions, Reich devised various ways of attacking or "dissolving" them directly. In working on tensions in and around the eyes, for example, it is frequently possible to release a great deal of anxiety; in loosening up and encouraging the movement of certain muscles around the mouth, suppressed feelings of disgust can be liberated; by suitable work on the chin, it is possible, in Reich's words, "to set free an unbelievable amount of anger." Reich writes that he had

previously been able to bring about the release of repressed impulses and emotions by way of dissolving purely characterological inhibitions and attitudes. Now, however, "the break-through of biological energy was more complete, more forceful, more thoroughly experienced, and it occurred more *rapidly*. Also, it was accompanied in many patients by a spontaneous dissolution of the characterological inhibitions" (p. 241). Reich warns, however, that it is not possible to dispense with work on character attitudes. "Everyday practise soon teaches one," he writes, "that it is not permissible to exclude one form of work at the expense of the other" (p. 293). With some patients work on the muscular rigidities will predominate from the beginning; with others, work on the character attitudes; but in all cases work on the muscular armor becomes more important in the later stages of the treatment.

MIND-BODY PROBLEM. The facts he discovered about chronic muscular rigidities and their relation to character attitudes and repressed emotions, Reich maintained, required the abandonment of the dualistic theories about body and mind tacitly or explicitly accepted by many psychologists and most psychoanalysts. It is a mistake to regard the muscular rigidity as a mere accompaniment or as an effect of the corresponding character attitude: It is "its somatic side and the basis for its continued existence" (*The Function of the Orgasm*, p. 269). The rigidity of a muscle group and the corresponding attitude serve the same function, namely, that of holding back a repressed emotion. The muscular armor and the character armor may therefore be said to be "functionally" identical. The only tenable answer to the body-mind problem, according to Reich (who quotes Julien Offray de La Mettrie as anticipating his position), is a materialistic form of the identity theory.

Reich's identity theory is materialistic, not in the sense that introspection is regarded as illusory or as devoid of scientific value but in holding that a change in a person's character, or indeed any change in a human being, cannot come about without appropriate physiological changes. The notion, writes Reich, that "the psychic apparatus functions by itself and influences the somatic apparatus is not in keeping with the facts" (ibid., p. 313). Even an idea such as that of going to sleep will not "exert a somatic influence unless it is already the expression of a vegetative impulse" (ibid.). This conclusion, Reich insists, is not contradicted by the observation that a patient (or anybody) feels relieved when a previously repressed idea or impulse is allowed to become conscious. "We used to say," writes Reich, "that it is a matter of a discharge of psychic energy which previously was bound" (*Character Analysis*, p. 311).

Closer examination will show in such a case that both the tension and the relaxation are clearly observable somatic processes. What is introspectively felt as tension and as relief are in fact certain fairly typical rigidities and relaxations of muscles—in the forehead, in the eyes, and elsewhere in the body. Both Reich and his translator, T. P. Wolfe, insist that the issue between dualism and the identity theory is not merely a question of alternative languages but makes a difference to therapeutic practice and further research. Wolfe in particular claims that only a theory of "psychosomatic identity" makes sense of the vast array of facts that had accumulated in psychosomatic studies by 1940 and that only such a theory can provide a fruitful method of research (*The Function of the Orgasm*, pp. x and xiii).

There are two very different questions that may be raised about all of this. One may ask whether, granting that Reich has hit upon something interesting and important in connection with muscular rigidities, their origin, and their possible dissolution, an identity theory is the only philosophical position that can accommodate these facts. More fundamentally, one may raise the question of whether Reich's empirical claims about the muscular armor are true in the first place.

As to the first of these questions, it should be pointed out that when Reich speaks of the "*functional* identity" of the character and the muscular armor, he does not seem to mean by "identity" anything as strong as has been claimed by philosophical defenders of the identity theory. To say that a certain character attitude and a certain muscular rigidity have the same function, for example, that of binding anxiety or anger, is not anything that a dualist is required to deny. It is certainly compatible with, but it does not by itself imply, the claim that the character attitude and the muscular rigidity are two aspects of the same phenomenon. It might be argued that Reich's work on the connection between muscular rigidities and character attitudes, rather than proving any traditional version of the identity theory, shows the inadequacy of interactionistic forms of dualism. Interactionism, in allowing only for causal relations between physical and psychological phenomena, could not do justice to the intimate relations between muscular rigidities and character attitudes to which Reich has called attention. There is no reason to suppose, however, that a more open-minded form of dualism, which would not restrict the relations between body and mind to one simple type, could not accommodate the facts in question.

In the present entry we cannot attempt to answer the second of our two questions—whether Reich's empirical claims about the muscular armor can in fact be sustained. Perhaps, however, it is permissible to remark, especially since this part of Reich's work has received so much less publicity than his orgone theory, that psychiatrists and others who have some firsthand knowledge of it have generally been enthusiastic. This includes persons who have observed and treated children in the light of Reich's account of the muscular armor. Since the process of repression as well as the process of cure would, on almost any theory, be most readily observable in children, confirmations (or disconfirmations) here would seem to be of special significance.

CULTURE, SOCIETY, AND CHARACTER STRUCTURE

CULTURE, MORALITY, AND THE DEATH INSTINCT. On the basis of both his clinical observations and his very extensive social work, Reich maintained that there is nothing more deadly than to be subjected to the moralistic and authoritarian upbringing which is or which was until very recently the lot of the great majority of children all over the world. The preaching and the antisexual moralism of the religious home and the authoritarian character of the conventional school stifle every vital impulse in the child. Insofar as traditional education is successful, it produces human beings with a craving for authority, a fear of responsibility, mystical longings, impotent rebelliousness, and pathological drives of all kinds. The "morals" fostered by religious mysticism and slavishly followed by many who no longer believe in religion "create the very perverted sexual life which it presumes to regulate moralistically; and the elimination of these 'morals' is the prerequisite for an elimination of that immorality which it tries in vain to fight" (*The Mass Psychology of Fascism*, p. 156).

There is nevertheless an important element of truth in the contention of conservative ideology that if one were to "eliminate morals," the "animal instincts" would gain the upper hand, and that this would lead to social chaos. What is true in this contention is that the average person in our culture carries within himself an "unconscious inferno," and while his perverse and destructive impulses are not in most cases adequately controlled by moral inhibitions, they would presumably dominate personal and social life to an even greater extent in the absence of moral regulations. This fact makes it clear that any transition from an authoritarian to a rational self-governing society must be gradual and cannot be accomplished by simply telling people, *as they now are*, to live according to their impulses. It does not, however, provide a justification for an ascetic morality or for the usual conservative theory that maintains that culture is based on sexual repression.

The conservative theorist errs in assuming that the antisocial impulses are "absolute and biologically given" (*The Sexual Revolution*, p. 20). This view is advocated not only in the writings of religious moralists and others to whom Reich contemptuously referred as "uplifters" or "guardians of the higher values" but also in many of the later writings of Freud and those of Freud's followers who accepted the theory of the death instinct. Accordingly, Reich devoted much effort to a very detailed attack on the theory of the death instinct, especially as it is applied to human society and culture in Freud's *Civilization and Its Discontents*.

On Freud's view, Thanatos, or the striving for peace and extinction, is just as much biologically given as Eros, or the sexual strivings. Although the death instinct itself cannot be perceived, it manifests itself in a great many ways—in various forms of aggression, in self-destructiveness, and in the masochistic "need for punishment." It also accounts for the resistances put up by patients against getting well. According to Reich, however, both clinical experience and observation of children show that the phenomena which supposedly prove the death instinct are "secondary formations," the products of the neurosis, and not "primary" and "biological" like the sexual instinct or the need for food. Investigation reveals that suicide is either an unconscious revenge upon another person or a way of escaping the pressure of a situation that has become overwhelming. The neurotic fear of and concern with death that is frequently found in quite young people can in every case be reduced to a fear of catastrophe, and this, in turn, to genital anxiety. As for aggressiveness, Reich claimed that the proponents of the death instinct did not sufficiently distinguish between perfectly healthy forms and those which are sadistic and destructive. The former are intimately connected with life-affirming tendencies, and the latter are always reactions of the organism to the denial of the gratification of a vital need.

Reich equally denied that there is any evidence whatsoever for the theory of "primary masochism." All clinical observations support Freud's earlier theory that patients "had come to grief as a result of their *fear* of punishment for sexual behavior and not as a result of any *desire* to be punished for it" (*The Function of the Orgasm*, pp. 103–104). The theory of the death instinct, furthermore,

is therapeutically sterile and offers an excellent excuse for one's inability to handle a difficult resistance. In addition to providing an alibi for therapeutic failures, it serves the same function as the discredited biologistic theory of congenital criminality or the view of Magnus Hirschfeld that exhibitionism is due to special "exhibitionistic hormones": All such views shift problems from the social to the biological realm, where nothing can—and need—be done about them.

Conservative theorists who maintain that there is an antithesis between sexuality and work fail to distinguish between "compulsive-unpleasurable" work, which is indeed regarded as a burdensome duty, and "natural joyful work," which frequently requires discipline but which is nevertheless a pleasurable gratification of a need. Reich regards as especially significant his observations on patients who achieved sexual happiness. He reports that those who, because of neurotic disturbances, had not been working, began to feel a strong need for some vital work. Those who had been engaged in work that was intrinsically interesting now blossomed and gave full rein to their talents. In some cases, however, there was a complete breakdown of work. This at first seemed to confirm the view of the antisexual moralists, but closer inspection showed that these people had previously been driven by a compulsive sense of duty and that what they rebelled against was empty and mechanical work, and not work as such. Their aversion was to pleasureless work, and their impulses were by no means antisocial. Just as society rewards some highly antisocial activities with fame and honor, Reich remarks, so "there are highly valuable, even culturally important traits and impulses which have to be repressed for considerations of material survival" (*The Function of the Orgasm,* p. 150). If there were more human beings with a genital character, this would not result in the end of "civilization," but it would in all probability lead to radical changes in the ways in which the world's work is done.

Reich concluded that civilization and culture do not depend on instinctual repression. If authoritarian education were abolished and if children grew up in a sex-affirmative environment, people would be more, and not less, peaceful and cooperative. Some types of work, namely, those in which only a person with a compulsive character can take any interest, would indeed suffer, but the arts and sciences would in all likelihood flourish as never before. Reich was not an irrationalist in any sense of the word, and like Freud he favored "the primacy of the intellect," adding, however, that the full utilization of a person's intellectual capacities presupposes "an orderly libido economy." "Genital and intellectual primacy have the same mutual relationship as have sexual stasis and neurosis, guilt feelings and religion, hysteria and superstition" (*Character Analysis*, p. 170).

SOCIETY AND CHARACTER STRUCTURE. Freudian social theory, insofar as it existed at all when Reich began his elaborate critique of what he called "authoritarian" society, was vitiated by its "biologism" as well as its "psychological atomism," or, as Reich also called it, a "feudal individualistic psychology." By "biologism" Reich meant the tendency to treat as universal and biologically inevitable attitudes and impulses that were determined by cultural conditions. When he spoke of Freud's "psychological atomism," Reich referred to the tendency to treat individual patients and their families in isolation from the social environment that had in fact a great deal to do with their tribulations.

Rejecting Freud's biologism and accepting the early Freudian view that neurosis is basically the result of the conflict between instinctual needs and the reality which frustrates them, Reich naturally asked whether and how this frustrating reality could be significantly altered. His work at the sex hygiene clinics, furthermore, had convinced him that neuroses were by no means the fads of middle-class women who did not know what to do with their time but were emotionally crippling illnesses of almost epidemic proportions. Contrary to the assertions of the more doctrinaire and narrow-minded Marxists, there could be no doubt in Reich's view that "sexual repression, biological rigidity, moralism and puritanism are ubiquitous" and not confined to certain classes or groups of the population (*The Function of the Orgasm*, p. xxiii). The vast majority of people suffering from psychological disturbances cannot, however, be reached by individual therapy, disregarding here all the difficulties and limitations of such therapy when it is available.

If one is to do anything about this deplorable state of affairs, one must first achieve an understanding of the precise relations between society and the individual and, more specifically, between social institutions and neurotic disturbances. "Society," Reich writes, "is not the result of a certain psychic structure, but the reverse is true: character structure is the result of a certain society" ("Character and Society," p. 254). The ideology of a given society can anchor itself only in a certain character structure, and the institutions of the society serve the function of producing this character structure. If, as in all authoritarian societies, a minority holds economic and political power, it also has the power to form ideology and structure. As a

consequence, in authoritarian society, the thinking and the structure of the majority of people "corresponds to the interests of the political and economic rulers" (*The Sexual Revolution*, p. xx). The majority of human beings (Reich is writing in 1936) are "suppressed and exploited and spend most of their working hours doing monotonous and mechanical labor which they cannot help regarding as a loathsome duty." How is it possible that "people can bear it, that they are unable to change it, that they seem to endure in silence the suffering it imposes on them?" ("Character and Society," p. 252). They can bear their fate because the ruling economic system is "anchored in the psychic structure of the very people who are suppressed" (*People in Trouble*, p. 100).

The most important structure-forming institutions in authoritarian society are the authoritarian family, the authoritarian school, and religion. "From infancy on," writes Reich, "people are trained to be falsely modest, self-effacing and mechanically obedient, trained to suppress their natural instinctual energies" ("Character and Society," p. 252). In this way children become subservient to their parents and people in general "subservient to the authoritarian state power and capitalistic exploitation" (*People in Trouble*, p. 99). The most powerful instrument in achieving this mass structure is sexual repression, which is fostered in the home, in the school, and above all through the influence of religious moralism. The major mechanisms of sexual suppression in Christian countries are the prohibition of infantile masturbation, the prevention of sexual gratification in adolescence, and the institution of compulsorily lifelong monogamy, accompanied by the belief that the function of sexuality is procreation rather than pleasure. The parents who punish children for masturbating and who do their best to prevent adolescents from having a full sex life are unwittingly carrying out the purpose of the ruling powers.

There is something plausible about Reich's contention that an atomistic psychology, no matter how correctly it may determine the causes of mental health and illness, will not by itself explain why various institutions that are plainly inimical to life and happiness nevertheless flourish and receive the support of all the major official and unofficial agencies of society. However, it is not entirely clear what he means by his claim that character structure is the result of social structure and, more specifically, that the "function" of sex-denying institutions is to make the masses helpless and dependent. Although he occasionally uses the word *purpose*, Reich is presumably asserting the existence of a "latent" rather than a "manifest" function, to use the terminology introduced by R. K. Merton. While it may be plausibly argued that some rulers, like Joseph Stalin and certain church figures, have been aware of the connection between sexual suppression and such "desirable" traits as obedience and uncritical acceptance of the status quo, it would be farfetched to hold that either in capitalistic or in other societies the ruling circles deliberately support sex-denying institutions in order to perpetuate their power and privileges. But if the rulers are not conscious of the causal connection between sexual suppression and the submissive traits it produces, in what sense is a reference to their interest an explanation of the institutions in question? It is tempting to speak here of an "unconscious knowledge" or "unconscious realization" that sexual suppression produces submissiveness, but it is far from clear what these expressions would mean.

Reich's views about the relation between the ideology that prevails in a society and the interests of the holders of power has obvious affinities with Marxism, and in fact a number of Marxist writers of the late 1920s and early 1930s hailed his account of the social function of sexual repression as a valuable supplement to historical materialism. However, the most influential Marxist ideologists, socialist as well as communist, rejected Reich's account and also strongly opposed his work in his sex hygiene clinics. In his turn, Reich repudiated what he called the "economism" of Marxist theory as emphatically as he attacked the atomism of psychoanalysis. "Marxists again and again argued," he recalls, "that the sexual etiology of the neuroses was a bourgeois fancy idea, that only 'material want' caused neuroses … as if the sexual want were not a 'material' one: It was not the 'material want' in the sense of the Marxian theorists that caused the neuroses, but the neuroses of these people robbed them of their ability to do anything sensible about their needs, actually to do something constructive about their situation, to stand the competition on the labor market, to get together with others in similar social circumstances, to keep a cool head to think things out." (*The Function of the Orgasm*, pp. 56–57).

Moreover, just as it is wrong to think that neuroses are (except very indirectly) caused by economic hardships, so it is a mistake to suppose that the social and political actions of the working classes can be predicted on the basis of their economic interests alone. Factors such as mystical and sexual longings and perverse sadistic fantasies may exert very powerful influences, as Hitler, unlike the communist, socialist, and liberal politicians, understood only too well. Fascism, to take but one example, is very incompletely characterized as a movement

engineered by capitalists to prevent the establishment of socialism. At least the German variety of fascism differed from other reactionary movements in that it was "supported and championed by masses of people" (*The Mass Psychology of Fascism*, p. ix). Marxist theory, which assumes that with few exceptions the underprivileged will be guided by their rational economic interests, is incapable of accounting for such a phenomenon.

THE STRUGGLE AGAINST RELIGION

MYSTICAL FEELINGS AND SEXUAL INHIBITIONS. According to Reich, both Karl Marx and Freud made significant contributions to our understanding of religion. Patriarchal religions are always politically reactionary, and Marxists are perfectly right in pointing out that "in every class society they are in the service of the powers that be" (p. 124). Freud, too, was correct in his view that the idea of God derives from the idea of the father and, more generally, that "the psychic contents of religion stem from the infantile family situation" (ibid.). Granting all of this, there remains a question that is not answered by the Marxist or the Freudian account, or by any of the great eighteenth-century critics of religion. Indeed, it is a question that most of these writers did not even raise but which must be asked and answered if one is to have an adequate comprehension of religion. How are we to account for the fact that "religious ideas are invested with such intense feelings"? What explains the "enormous emotional power of mysticism" (p. 122)? Or, using Reich's favorite terminology, what is the "energy" that enables religions to gain such a firm hold on people? What is it that compels human beings not only to accept the idea of a pleasure-prohibiting, all-seeing God and the ideologies of sin and punishment, and "not to feel them as a burden but, on the contrary, to uphold and fervently defend them, at the sacrifice of their most primitive life interests?" (*The Mass Psychology of Facism,* p. 124).

Reich is strongly opposed to the tendency of "emancipated" unbelievers to dismiss religions as nothing more than the fancies of silly and ignorant people. He insists that a study of religious people—of the content of their emotions and beliefs, of the ways in which these are implanted, and of the function that they fulfill in their psychological economy—is highly rewarding. It sheds light on many other phenomena, including, for example, the psychological basis of fascism and of reactionary political movements. Such a study also explains why, by and large, free-thought propaganda is so unsuccessful in spite of the fact that from a purely rational point of view the positions defended by freethinkers are vastly superior to the religious claims—something that is not altogether unknown among believers. Above all, a happy life for the majority of humankind is impossible unless the power of religion is broken, unless one can prevent "the mystical infestation of the masses" (p. 161).

However, in order to be effective in "the relentless fight against mysticism," one must have a full comprehension of its origin and its psychological sources of strength so that one can meet its "artful apparatus … with adequate counter-measures" (p. 152). To suppose that mystical attitudes become anchored in human beings simply as a result of intellectual indoctrination is a naive and dangerous mistake. It should be noted that Reich sharply distinguishes mysticism from primitive animism. The latter is best regarded as bad science. Reich does not offer an explicit definition of "mysticism," but it seems clear from his various writings on the subject that mysticism in the "strict and wider sense" is characterized by the belief (or feeling) that the ordinary world of physical objects and human emotions is *not enough* and the related view that there are some grand truths which human beings can come to know by nonscientific or superscientific means. Various nontheological systems of metaphysics and ontology, as well as the standpoint of those who deny that psychology can properly be a natural science (Reich is specially scathing in his comments about Ludwig Klages and Karl Jaspers), are treated by him as forms of mysticism.

The most basic feature of what Reich variously calls "religious excitations" or "mystical feelings" is that they are "at one and the same time *anti*-sexual and a *substitute* for sexuality" (p. 125). Reich claims that this conclusion is borne out by the close observation of genuinely religious people (as contrasted with those who merely pretend belief for purposes of personal gain and advancement); by character-analytic treatment of religious individuals and patients having mystical feelings of any kind; by observation of children, especially those suffering from prayer compulsions; by the writings of the mystics themselves; and also by what is known about the changes that occurred when social organization passed from matriarchy to patriarchy and class society.

Biologically, the religious individual is subject to states of sexual tension like any other living being. However, as a consequence of his sex-negating upbringing and especially his fear of punishment, he has lost the capacity for normal sexual stimulation and gratification. The result of this is that he suffers from a chronic state of excessive somatic excitation. The more thorough his religious education has been, the more it appears to him that

happiness is not attainable for him in this world and, in the long run, it does not even seem desirable any more. However, he remains a biological organism and hence cannot completely renounce the goals of "happiness, relaxation and satisfaction." In these circumstances all he can do is seek "the *illusory* happiness provided by the religious *forepleasure* excitations" (*The Mass Psychology of Fascism,* p. 126).

The "somatic suffering" of the religious person creates in him the need for consolation and help from outside himself, particularly in his fight against what he terms the "evil instincts," which in turn are identified with the "evils of the flesh." His religious ideas enable him to attain a state of "vegetative excitation which resembles gratification but does not, in reality, bring about somatic relaxation" (p. 127). Not even religious ecstasies bring about anything comparable to the orgastic relief of a satisfying sexual experience. What the religious person calls his longing for "delivery from sin" is in fact a longing for relief from sexual tension. To people who cannot achieve sexual gratification, sexual excitation gradually and inevitably becomes something "torturing and destructive." In this way the religious conception of sex as evil and debasing has its foundation in real somatic processes. People who feel a disgust for their body quite naturally develop obsessive concepts of "purity" and "perfection" (p. 144).

It would lead too far afield to discuss here the various ways in which, according to Reich, the "mystical idea of God" becomes anchored in people. These mechanisms may vary in detail, but they all involve the implanting of sexual anxieties; and Reich concludes that from the point of view of energy, mystical feelings are "sexual excitations which have changed their content and goal." The energy of these emotions is the energy of natural sexuality that has become transformed and attached to mystical, psychic contents. Religious patients, upon establishing a fully satisfying sex life, invariably lose their God-fixation.

Once one comprehends the nature of "religious excitations," it becomes clear why the free-thought movement "cannot make itself felt as a counter-force" (p. 147). Aside from the fact that in many countries the churches enjoy the support of the state and that generally the mass information media are grossly biased in favor of religion and religious morality, the impact of free-thought propaganda is limited because it relies almost exclusively on intellectual arguments. These are not, indeed, a negligible factor, but they are no match for the "most powerful emotion" on which the mass-psychological influence of religious institutions is based: sexual anxiety and sexual repression. People with a religious upbringing who, as a result of the study of science and philosophy, have turned into unbelievers very frequently retain religious longings and emotions. Some of them even continue to pray compulsively. This does not prove, as some advocates of religion argue, that religious needs are "eternal and ineradicable." It does, however, show that "while the religious feeling is opposed by the power of the intellect, its sources have not been touched" (p. 152).

The fight against religion is nevertheless far from hopeless. Mysticism can be eradicated if, in addition to depriving the churches of their "evil right of preparing the children's minds for the reception of reactionary ideologies" (p. 148), one is guided in the struggle by one's knowledge that mysticism stems from inhibited sexuality. From this insight it follows incontrovertibly that "full sexual consciousness and a natural regulation of sexual life mean the end of mystical feelings of any kind, that, in other words, natural sexuality is the deadly enemy of mystical religion" (p. 152). Any social efforts that are directed toward making people affirm their sexual rights will ipso facto weaken the forces of mysticism. The most good can be done with children and adolescents. Reich gives numerous instances from his experience in Germany of the "burning interest" of children in sexual questions that made even the most enlightened adults ashamed of their prudishness and hesitation. "Once children and adolescents are reached on a mass basis through their sexual interests," there will be a "powerful counterweight against the reactionary forces" (p. 169). As for those people who are too old to have their structure basically altered, it is still all to the good to bring "silent suffering to the surface." They might then be less likely to become instruments in the process of maiming their own children, and they will not continue to support sex-repressive laws.

THE GREAT CULTURAL REVOLUTION. Reich never abandoned the conviction he had reached during his Marxist phase that individual therapy is socially insignificant and that "alteration of the social structure is a prerequisite for an alteration of the psychic structure on a mass scale" ("Character and Society," p. 255). However, after his separation from organized Marxism, he gradually came to the conclusion that political action was of little consequence and that it was a grave error to judge social developments primarily in terms of a rigid, clear-cut class war. If one is not blinded by the political slogans of an earlier age, one cannot help noticing that we are in the midst of a "deep-reaching revolution of cultural living" (*The Sexual Revolution*, p. xiv).

It is a revolution "without parades, uniforms, drum or cannon salutes," but, unlike the Russian Revolution of 1917, which was merely "politico-ideological," it is a "genuine social revolution" (*The Mass Psychology of Fascism*, p. 201). It is not a revolution by the proletariat against the bourgeoisie, and it remains to be seen what major economic changes will accompany it. What is happening is that "the senses of the animal, man, for his natural life functions are awakening from a sleep of thousands of years" (p. xiv). Ever since the beginning of the century, numerous social factors have been operating in the direction of freedom and health. These factors include the creation of huge industrial plants with vast armies of workers of both sexes and the gradual undermining of the authoritarian parental home. There has been a "thorough disintegration of the moralistic ascetic forms of living," and this "objective loosening of the reactionary fetters on sexuality cannot be undone" (p. 164), regardless of how vociferously the churches and their conscious or unconscious allies continue to preach the old morality.

This "great cultural revolution" is bound to be chaotic and to give rise to all kinds of grotesque developments. The disintegration of the old moralistic institutions and customs expresses itself at first as a rebellion that takes pathological forms, but it is not difficult to see that healthy forces are trying to break through in these pathological manifestations. At one time Reich envisaged a "powerful international organization" that would create an atmosphere of sex-affirmation and thus help to "guide the rebellion into rational channels" (*The Mass Psychology of Fascism*, p. 121). However, regardless of whether any organizations are brought into being which could accelerate the process and make it less painful, there is no reason to "fear for the final outcome." As yet, human beings, "moved by obscure, 'oceanic' feelings, dream instead of mastering their existence; and they perish from these dreams" (*Character Analysis*, p. 324). But when once they master their existence, when they become capable of giving and receiving love and when work will be a source of pleasure and not a burden, this will mean "the death-knell of all transcendental mysticism, of the 'absolute objective spirit,'" and of all the metaphysical and irrationalist philosophies that are "subsumed under mysticism in the … wider sense." An individual "who is sexually happy does not need an inhibiting 'morality' or a supernatural 'religious experience.' Basically, life is as simple as that. It becomes complicated only by the human structure which is characterized by the fear of life" (*The Sexual Revolution*, p. 269).

REICH'S LAST YEARS

It is not surprising that the ideas sketched in the preceding sections of this entry should have appealed to many who were dissatisfied with the conservative developments of psychoanalysis as well as to those who, disillusioned with the results of communism in Russia, nevertheless strongly believed in social progress. During his early years in the United States, Reich did in fact count among his followers or sympathizers a number of remarkably talented men, from the most varied walks of life, who saw the dawn of a new enlightenment in his psychiatry and in the implications of his theories for education and for the proper direction of social reform. It would be difficult to convey to anybody who was not actually living in New York at that time the enthusiasm that was felt for Reich personally and for what were regarded as his liberating insights. As was to be expected, communists and psychiatrists of other schools were violently hostile, but this only served to heighten people's admiration for Reich's independence and for his uncompromising integrity.

It was mentioned previously that Reich himself became less and less interested in psychiatry. He also gradually lost most of his concern to guide into rational channels the "great cultural revolution" that he had diagnosed in his writings. The publications of his last years do indeed contain numerous discussions of social topics, but, at least in the opinion of the present writer, most of what Reich now had to say was flat and trivial. He became increasingly obsessed with the evil conspiracies of "red fascism" (some of Reich's remarks during this period could be quoted with approval by members of the John Birch Society) and with the menace of the "emotional plague." This term was originally introduced to refer to the harmful activities of individuals who take out their sexual sickness and frustrations on the rest of humankind, usually under the pretense of promoting some worthy cause.

Reich's earlier description of emotional plague reactions and motives had been extremely perceptive, but now anybody who was in any way opposed to any of his ideas became automatically classified as an agent of the emotional plague. The writings of the last years are also filled, in a manner reminiscent of Friedrich Nietzsche's *Ecce Homo*, with hymns of self-praise (sometimes in the third person), and there is much evidence of extreme bitterness toward a world that did not accept or even pay attention to his theories. From the available accounts it appears that Reich had always been impatient and somewhat autocratic, but he had also been singularly compassionate and generous. Dr. Nic Waal, in her sketch of

Reich, describes him as "enormously stimulating and lovable" but adds that in his last years he "became less and less patient, less loving … and finally pathologically suspicious" (*Wilhelm Reich—A Memorial Volume*, p. 37).

If Reich became increasingly bitter, this was not without a good deal of justification. Right from the beginning, even while he was a psychoanalyst "in good standing," Reich was the victim of an extraordinary amount of spite and slander. Any study of the records will make it clear that he was treated outrageously by the officials of the Psychoanalytic Association both before and at the Lucerne Conference. We have already mentioned Reich's troubles in Scandinavia. In New York, he was arrested by the Federal Bureau of Investigation in December 1941 and held at Ellis Island for three weeks. The reasons for the arrest were never divulged. In 1947 an exceptionally vicious campaign was initiated in the *New Republic*, by the journalist Mildred E. Brady. There was not a paragraph in her article that did not contain a major distortion, but it was nevertheless quoted and reprinted all over America.

In an article ten years earlier, the German poet Stephan Lackner had expressed his indignation at the treatment that Reich had received and continued to receive from leading figures among the psychoanalysts and the left-wing parties. "It was not enough," wrote Lackner, "to expel Reich from their organizations"; in the struggle against this man and his disturbing ideas, "every kind of slander and distortion is a permissible weapon" (*Das neue Tagebuch*, February 1937, p. 140). This last remark applies, word for word, to the campaign instigated by Brady and her associates. In March 1954, the U.S. Food and Drug Administration obtained an injunction against Reich and his foundation, ordering the destruction of all orgone accumulators, all of Reich's journals, and some of his books; the books that were not destroyed were to be impounded. Among the works proscribed on the ground that they constituted "labeling" of the orgone accumulator were such books as *The Function of the Orgasm* and *Character Analysis*, in which the accumulator is not so much as mentioned.

Nobody except fanatical partisans of Reich can dispute the right of the Food and Drug Administration to intervene. When on the defensive, Reich denied that he had ever claimed any curative powers for the orgone accumulator, but the truth is that the literature is full of such claims. However, granting that the Food and Drug Administration had evidence to show the accumulator medically worthless (no such evidence has ever been published), the injunction is nevertheless a startling document constituting a blanket attack on Reich's character and his entire work.

Reich had two weeks in which to appeal, but to everybody's consternation he refused to appear in court. Instead, he wrote a letter to the judge in the case, declaring that a court of law was not the appropriate place for adjudicating scientific questions. For some months Reich obeyed the injunction, but in October 1954 he notified the authorities that he was about to resume all the activities of his institute, including the sale of books and periodicals. This led to a trial in 1956, at which Reich was given the maximum sentence of two years in a federal penitentiary. Reich died of a heart attack eight months after he had started serving his sentence. All journals published by Reich's institute that were seized by government agents were burned in two separate actions in 1956 and 1960, and his books were impounded until they began to be republished by a commercial house in 1960.

There is no doubt in the mind of the present writer that during his last years Reich was mentally ill. Some of those who were close to him deny this, and the prison psychiatrist who examined Reich certified him as sane. Nevertheless, if one reads the records of the trial or the brief that Reich filed in his appeal, one can hardly resist drawing the conclusion that a great man had broken down. Reich finally "went to pieces," observed Dr. Waal, "partly on his own—but mostly due to other people," adding that "a human being cannot bear cruelty and loneliness in the long run" (*Wilhelm Reich—A Memorial Volume,* pp. 38–39). It is worth recalling the words of Josef Popper-Lynkeus, whose ideas bear little resemblance to Reich's but who was also described as "mad" during the better part of his life. "I assure you," he told his biographer, "[that] of all experiences none is more painful than that of finding oneself described as mad as a consequence of having discovered something that is good and true: of all martyrdoms, this is perhaps the most terrible" (A. Gelber, *Josef Popper-Lynkeus*, p. 101).

See also Functionalism in Sociology.

Bibliography

The following abbreviations are used throughout; ZPS for *Zeitschrift für politische Psychologie und Sexualökonomie* and *IJSO* for *International Journal for Sex-Economy and Orgone Research*.

Several biographies of Reich have been announced, but none had been published by the time this entry went to press. The only published sketches of Reich are A. S. Neill's "The Man Reich" and Nic Waal's "On Wilhelm Reich," both in *Wilhelm Reich—A Memorial Volume*, edited by Paul Ritter (London, 1958). There is a good deal of autobiographical material,

especially on his relations with Freud, in Reich's *The Function of the Orgasm*, translated from the German manuscript by T. P. Wolfe (New York: Orgone Institute Press, 1942; paperback reprint, 1961). This book is a good introduction to all of Reich's theories discussed in the present entry. The reader should be warned, however, that in the 1961 reprint the very valuable introduction by Dr. Wolfe has been deleted. *People in Trouble* (Rangeley, ME: Orgone Institute Press, 1953) contains an account of Reich's work at his sex hygiene clinics and of his difficulties with communist functionaries in Germany and Denmark. Reich's attempt to organize an international movement in support of a sex-affirmative culture is described by him in two articles: "Zur Geschichte der Sexpol Bewegung," in *ZPS* 1 (1934): 259–269, and "Geschichte der deutschen Sexpol-Bewegung," in *ZPS* 2 (1935): 64–70. The only published account of Reich's troubles in Norway is Gunnar Leistikow, "The Fascist Newspaper Campaign in Norway," in *IJSO* 1 (1942): 266–273. This article also discusses Reich's troubles in Denmark. Its title is misleading in that many of Reich's opponents were not fascists. Reich's *Listen Little Man!*, translated from the German manuscript by T. P. Wolfe, with illustrations by William Steig (New York: Orgone Institute Press, 1948), is a moving outburst against the various people who harassed and defamed him.

The fullest published account of Reich's technique of vegetotherapy is *Orgasmusreflex, Muskelhaltung und Körperausdruck* (Copenhagen: Sexpol, 1937), parts of which are translated in Chapter 8 of *The Function of the Orgasm* and Chapter 15 of the third edition of *Character Analysis* (New York: Orgone Institute Press, 1949). Various aspects of Reich's new technique are also discussed in the following articles: Odd Havrevold, "Vegetotherapy," in *IJSO* 1 (1942): 65–87, written under the pseudonym Walter Frank; Ola Raknes, "The Treatment of a Depression," in *IJSO* 1 (1942): 163–170, and "Sex-Economy," in *IJSO* 3 (1944): 17–37, written under the pseudonym Carl Arnold. (These pseudonyms were necessary during the Nazi occupation of Norway.)

Child therapy is discussed in Felicia Saxe, "A Case History," in *IJSO* 4 (1945): 59–71, and "Armored Human Beings versus the Healthy Child," in *Annals of the Orgone Institute* 1 (1947): 35–72; and in Nic Waal, "A Case of Anxiety Neurosis in a Small Child," in *Bulletin of the Menninger Clinic* 12 (1948), and "A Special Technique of Psychotherapy with an Autistic Child," in *Emotional Problems of Early Childhood*, edited by G. Caplan (New York: Basic, 1955).

The fullest statement of Reich's views on religion is found in Chapters 6 and 7 of *Massenpsychologie des Faschismus* (Copenhagen: Verlag für Sexualpolitik, 1933), third edition translated by T. P. Wolfe as *The Mass Psychology of Fascism* (New York: Orgone Institute Press, 1946). Discussions of religion strongly influenced by Reich are J. H. Leunbach, "Religion und Sexualität," in *ZPS* 1 (1934): 70–72; Karl Teschitz, "Grundlagen der Religion," in *ZPS* 2 (1935): 103–129, and "Religiöse Ekstase als Ersatz der sexuellen Auslösung," in *ZPS* 4 (1937): 23–34; and Theodor Hartwig, "Religion und Sexualität," in *Der Freidenker* (Bern, April 1936), and "Der Sinn der 'religiös-sittlichen Erziehung,'" in *ZPS* 4 (1937): 203–205. Of philosophical interest are Reich's articles "Zur Anwendung der Psychoanalyse in der Geschichtsforschung," in *ZPS* 1 (1934): 4–16, and "Die Funktion der 'objektiven Wertwelt,'" in *ZPS*2 (1935): 32–43. For some years Reich considered himself a dialectical materialist. His attempt to give empirical meaning to the so-called dialectical laws can be found in *Dialektischer Materialismus und Psychoanalyse* (Berlin, 1929; 2nd ed., Copenhagen, 1934).

Reich's views concerning the relation between society and character structure are stated succinctly in "Charakter und Gesellschaft," in *ZPS* 3 (1936), translated by T. P. Wolfe as "Character and Society," in *IJSO*1 (1942): 247–256, and much more fully in *The Mass Psychology of Fascism* and in *Die Sexualität im Kulturkampf* (Copenhagen, 1936), translated by T. P. Wolfe as *The Sexual Revolution* (New York: Orgone Institute Press, 1945). Reich's claims about the "function" of sexual suppression are partly based on his anthropological theories, which are an extension of the work of Malinowski. The fullest statement of these theories is found in *Der Einbuch der Sexualmoral* (Berlin: Verlag für Sexualpolitik, 1932). There is a critical discussion of Reich's anthropology in a review of this book by Erich Fromm, in *Zeitschrift für Sozialforschung* 2 (1933): 119–122.

Critical discussions of Reich's character-analytic technique are found in Carl M. Herold, "A Controversy about Technique," in *Psychoanalytic Quarterly* 8 (1939): 219–243; and in Richard Sterba, "Clinical and Therapeutic Aspects of Character Resistance," in *Psychoanalytic Quarterly* 22 (1953): 1–20. The dispute over the existence of "primary masochism" is surveyed in C. C. Flugel, *Man, Morals and Society* (London: Duckworth, 1945). Flugel, after some hesitation, sides with Theodor Reik and Franz Alexander against Reich. Even sympathetic commentators have frequently expressed doubts about what they take to be the excessively simple and "Rousseauist" view concerning the "natural" man that is implicit in many of Reich's discussions. Reich's view on this subject is condemned as "a stale left-over of the eighteenth-century imagination" in Philip Rieff, "The World of Wilhelm Reich," in *Commentary* 38 (September 1964): 50–58. There are replies to Rieff in *Commentary* 39 (February 1965): 19–22. Perhaps the best-known attack on Reich is found in Chapter 21 of Martin Gardner, *Fads and Fallacies in the Name of Science* (New York, 1952; 2nd ed., New York: Dover, 1957). In the opinion of the present writer, Gardner gives an extremely distorted picture of Reich's significance, concentrating on the wild claims of his last years and doing scant justice to the ideas discussed in the present entry.

Various of Reich's theories are sympathetically discussed in Max Hodann, *A History of Modern Morals* (London: Heinemann, 1937); Stephan Lackner, "Ein moderner Ketzer," in *Das neue Tagebuch* (Paris) 5 (1937): 140–141; Harald K. Schjelderup, *Nervose Og Opdragelse* (Oslo, 1937); Neil McInnes, "An Examination of the Work of Wilhelm Reich," in *Hermes* 48 (1946): 26–29; Paul Goodman, "The Political Meaning of Some Recent Revisions of Freud," in *Politics* 2 (1945): 197–203, and "Dr. Reich's Banned Books," in *Kulchur* (1960), reprinted in Goodman's *Utopian Essays and Practical Proposals* (New York: Vintage, 1964), C. Berg, *Psychotherapy* (New York, 1948); Rudolf Brun, *General Theory of Neuroses*, translated by B. Miall (New York, 1951); and R. A. Wilson, "Wilhelm Reich and the Book Burners," in *Minority of One* 3 (February 1961): 6–7. There are constant references to Reich's therapy and to his social theories in the books written by A. S. Neill from 1939 on, such as *The*

Problem Teacher (London: Jenkins, 1939), *Hearts Not Heads in the School* (London: Jenkins, 1945), and *The Problem Family* (New York: Hermitage Press, 1949).

The *Orgone Energy Bulletin* 5 (1953): 1–137, contains a very extensive bibliography of writings by and about Reich up to 1952. Unfortunately all issues of this periodical, as well as all issues of *IJSO*, are among the publications that were burned by the U.S. government.

OTHER RECOMMENDED WORKS

Cohen, Ira H. *Ideology and Unconsciousness: Reich, Freud, and Marx*. New York: New York University Press, 1982.

Grunberger, Béla, and Janine Chasseguet-Smirgel. *Freud or Reich?: Psychoanalysis and Illusion*. New Haven, CT: Yale University Press, 1986.

Mann, W. E. *Orgone, Reich, and Eros: Wilhelm Reich's Theory of Life Energy*. New York: Simon and Schuster, 1973.

Mann, W. E., and Edward Hoffman. *The Man Who Dreamed of Tomorrow: A Conceptual Biography of Wilhelm Reich*. Los Angeles: J.P. Tarcher; Boston: Houghton Mifflin, 1980.

Reich, Ilse Ollendorff. *Wilhelm Reich: A Personal Biography*. New York: St. Martin's Press, 1969.

Reich, Wilhelm. *American Odyssey: Letters and Journals, 1940–1947: Beyond Psychology*. New York: Farrar, Straus and Giroux, 1999.

Reich, Wilhelm. *Beyond Psychology: Letters and Journals, 1934–1939*, edited by Mary Higgins. New York: Farrar, Straus, and Giroux, 1994.

Reich, Wilhelm. *Character Analysis*. 3rd ed. New York: Farrar, Straus and Giroux, 1972.

Reich, Wilhelm. *Early Writings*. New York: Farrar, Straus and Giroux, 1975.

Reich, Wilhelm. *Passion of Youth: An Autobiography, 1897–1922*, edited by Mary Higgins and Chester M. Raphael. New York: Farrar, Straus, Giroux, 1988.

Reich, Wilhelm. *Reich Speaks of Freud: Wilhelm Reich Discusses His Work and His Relationship with Sigmund Freud*. New York: Farrar, Straus and Giroux, 1967.

Reich, Wilhelm. *Selected Writings: An Introduction to Orgonomy*. New York: Farrar, Straus and Giroux, 1973.

Robinson, Paul A. *The Freudian Left: Wilhelm Reich, Geza Roheim, Herbert Marcuse*. New York: Harper & Row , 1969.

Sharaf, Myron R. *Fury on Earth: A Biography of Wilhelm Reich*. New York: St Martin's Press/Marek, 1983.

Paul Edwards (1967)
Bibliography updated by Michael J. Farmer (2005)

REICHENBACH, HANS

(1891–1953)

Hans Reichenbach was a leading philosopher of science and a proponent of logical positivism. He made important contributions to the theory of probability and to the philosophical interpretation of the theory of relativity, quantum mechanics, and thermodynamics.

LIFE

Reichenbach studied civil engineering, physics, mathematics, and philosophy at Berlin, Göttingen, and Munich in the 1910s. Among his teachers were neo-Kantian philosopher Ernst Cassirer, mathematician David Hilbert, and physicists Max Planck, Max Born, and Arnold Sommerfeld. Reichenbach received his degree in philosophy from the Friedrich-Alexander University of Erlangen-Nürnberg in 1915 with a dissertation on the theory of probability titled *Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit* (The Concept of Probability for the mathematical Representation of Reality), published in 1916. Between 1917 and 1920, while he was working as a physicist and engineer, Reichenbach attended Albert Einstein's lectures on the theory of relativity at Berlin. He was fascinated by the theory of relativity and in a few years published four books about this subject: *The Theory of Relativity and A Priori Knowledge* (1920), *Axiomatization of the Theory of Relativity* (1924), *From Copernicus to Einstein* (1927), and *The Philosophy of Space and Time* (1928). In 1920 he began teaching at the Technische Hochschule at Stuttgart as private docent.

With the help of Einstein, Planck, and Max von Laue, in 1926 Reichenbach became assistant professor in the physics department of Berlin University. In 1930 he undertook the editorship of the journal *Erkenntnis* (Knowledge) with Rudolf Carnap. In 1933, soon after Adolf Hitler became chancellor of Germany, Reichenbach was dismissed from Berlin University because his family had Jewish origin. He emigrated to Turkey, where he was appointed chief of the philosophy department of Istanbul University with a five-year contract. During his stay in Turkey he published *The Theory of Probability* (1935). In 1938 he moved to the United States, where he became professor at the University of California at Los Angeles. In the following years Reichenbach published *Experience and Prediction* (1938); *Philosophic Foundations of Quantum Mechanics* (1944); *Elements of Symbolic Logic* (1947); "The Philosophical Significance of the Theory of Relativity" in *Albert Einstein: Philosopher-Scientist* (1949), edited by Paul Arthur Schilpp; and *The Rise of Scientific Philosophy* (1951). Reichenbach died in 1953 while he was working on the nature of scientific laws and on the philosophy of time. The two books that came from this work, *Nomological Statements and Admissible Operations* (1954) and *The Direction of Time* (1956), were published posthumously.

COORDINATIVE DEFINITIONS

An important tool introduced by Reichenbach for the philosophical analysis of scientific theories is that of coordinative definitions. According to Reichenbach, a mathematical theory differs from a physical theory because the latter uses a specific type of definition, named coordinative definition, which coordinates (that is associates) some concepts of the theory with physical objects or processes. An example of a coordinative definition is the definition of the standard unit of length in the metric system, which connects the meter with a rod housed in the International Bureau of Weights and Measures in Sèvres, or with a well-defined multiple of the wavelength of a determined chemical element. Another example is the definition of the straight line as the path of a ray of light in vacuum. A scientific theory acquires a physical interpretation only by means of coordinative definitions. Without such type of definitions a theory lacks of a physical interpretation and it is not verifiable, but it is an abstract formal system, whose only requirement is axioms' consistency.

Geometry well illustrates the role of coordinative definitions. In Reichenbach's opinion, there are two different kinds of theories concerning geometry, namely mathematical geometry and physical geometry. Mathematical geometry is a formal system that does not deal with the truth of axioms, but with the proof of theorems—that is, it only searches for the consequences of axioms. Physical geometry is concerned with the real geometry in the physical world; it searches for the truth or falsity of axioms using the methods of the empirical science. The physical geometry derives from the mathematical geometry when appropriate coordinative definitions are added. For example, if the concept of a straight line is coordinated with the path of a ray of light in a vacuum, the theory of relativity shows that the real geometry is a non-Euclidean geometry. Without coordinative definitions, Euclidean and non-Euclidean geometry are nothing but formal systems; with coordinative definitions, they are empirically testable. Coordinative definitions are conventions, because it is admissible to choose a different definition for a concept of a theory. In the case of geometry, with a different definition for the straight line, Euclidean geometry is true. In a sense, choosing between Euclidean and non-Euclidean geometry is not a matter of facts, but a matter of convention.

RELATIVITY OF GEOMETRY

Reichenbach insists on the importance of the coordinative definitions in his philosophical analysis of the theory of relativity, especially in connection with the problem of determining the geometry of this world. In principle, scientists can discriminate between different geometry by means of measurements. For example, on the surface of a sphere, the ratio of the circumference of a circle to its diameter is less than π, whereas on the surface of a plane this ratio is equal to π. With a simple measurement of a circumference and of its diameter, we can discover we live on a sphere (the surface of Earth) and not on a plane. In the same way, using more subtle measurements, scientists can discover we live in a non-Euclidean space. However, there is a fundamental question: is measuring a matter of facts or does it depend on definitions? Reichenbach proposes the following problem, discussed in *The Philosophy of Space and Time*: is the length of a rod altered when the rod is moved from one point of space, say A, to another point, say B? We know many circumstances in which the length is altered. For example, the temperature in A can differ from the temperature in B. However, the temperature acts in a different way on different substances. If the temperature is different in A and in B, then two rods of different material, such as wood and steel, which have the same length in A, will have a different length in B. So we can recognize a difference in temperature and use suitable procedures to eliminate variations in measurement due to variations in temperature. In general, this is also possible for every differential force—that is, for every force that acts in a different way on different substances. But there is also another type of forces, called universal forces, which produce the same effect on all types of matter.

The best-known universal force is gravity, whose effect is the same on all bodies. What happens if a universal force alters the length of all rods, in the same way, when they are moved from A to B? By the very definition of universal forces, there are no observable effects. If we do not exclude universal forces, we cannot know whether the length of two measuring rods, which are equal when they are in the same point of space, is the same when the two rods are in two different points of space. Excluding universal forces is nothing but a coordinative definition. We can also adopt a different definition, in which the length of a rod depends on the point of space in which the rod stays. So the result of a measurement depends on the coordinative definition we choose. As a consequence, the geometrical form of a body, which depends on the result of measurements, is a matter of definition. The most important philosophical consequence of this analysis concerns the relativity of geometry. If a set of measurements supports a geometry G, we can arbitrarily choose a different geometry G' and adopt a different set of coordinative definitions so that the same set of meas-

urements supports G', too. This is the principle of relativity of geometry, which states that all geometrical systems are equivalent. According to Reichenbach, it falsifies the alleged a priori character of Euclidean geometry and thus falsifies the Kantian philosophy of space.

CAUSAL ANOMALIES

The principle of relativity of geometry is true for metric relationships—that is, for geometric properties of bodies depending on the measurement of distances, angles, and areas. The situation seems different when we are concerned about topology, which deals with the order of space—that is, the way in which the points of space are placed in relation to one another. A typical topological relationship is "point A is between points B and C." The surface of a sphere and the surface of a plane are equivalent with respect to metrics, provided an appropriate choice of the coordinative definitions, but they differ from a topological point of view.

Consider the following example presented by Reichenbach in *The Philosophy of Space and Time*. Intelligent beings living on the surface of a sphere can adopt coordinative definitions that, from a metric point of view, transform the surface of the sphere into the surface of a plane. However, there is an additional difficulty: Because the surface of a sphere is finite, it is possible to do a round-the-world tour, walking along a straight line from a point A and eventually returning to the point A itself. Of course this is impossible on a plane, and thus it would seem that these intelligent beings have to abandon their original idea that they are living on a plane and instead must recognize they are on a sphere. But this is not true, because another explanation is possible: They can assert that they had walked in a straight line to point B, which is different from point A but, in all other respects, is identical to A. They can also fabricate a fictitious theory of pre-established harmony—according to which everything that occurs in A immediately occurs in B—in order to explain the similarity between A and B. This last possibility entails an anomaly in the law of causality. We can reject causal anomalies, but only by means of an arbitrary definition. Thus topology depends on coordinative definitions, and the principle of relativity of geometry also holds for topology. According to Reichenbach, this example is another falsification of Kantian theory of synthetic a priori. In Kantian philosophy, the Euclidean geometry and the law of causality are both a priori, but if Euclidean geometry is an a priori truth, normal causality can be false; if normal causality is an a priori truth, Euclidean geometry can be false. We arbitrarily can choose the geometry, or we arbitrarily can choose the causality, but we cannot choose both.

QUANTUM MECHANICS

Quantum mechanics differs from the other scientific theories because in this theory there is no possibility to introduce normal causality. No set of coordinative definitions can give an exhaustive interpretation of quantum mechanics free from causal anomalies.

It is important to explain some concepts used by Reichenbach in *Philosophical Foundations of Quantum Mechanics*, his main work about quantum mechanics. Using a wider sense of the word "observable," some events occurring in quantum mechanics are observable; they are events consisting in coincidences between particles or between particles and macroscopic material, like the collision of an electron on a screen, signaled by a flash of light. Events between such types of coincidences are unobservable; an example is the path of an electron between the source and the screen on which it collides.

Quantum observable events are called, by Reichenbach, phenomena, whereas unobservable ones are called interphenomena. Reichenbach explains that there are three main interpretations concerning interphenomena: wave interpretation, according to which matter consists of waves; corpuscular interpretation, according to which matter consists of particles; and Bohr-Heisenberg interpretation, according to which statements about interphenomena are meaningless. The first two interpretations are called exhaustive interpretations, because they include a complete description of interphenomena. The last is a restricted interpretation, because it prohibits assertions about interphenomena. A normal system is an interpretation in which the laws of nature are the same for phenomena and interphenomena. This definition of a normal system is modeled on a basic property of classical physics: the laws of nature are the same whether or not the object is observed.

With these definitions, it is possible to formulate Reichenbach's principle of anomaly in quantum mechanics: there is no normal system. Thus causal anomalies cannot be removed from quantum mechanics. However, there is another peculiarity in quantum mechanics: for every experiment there exists an exhaustive interpretation—which is a wave or a corpuscular interpretation—that provides a normal system, although limited to this experiment. In other words, there does not exist an interpretation free from all causal anomalies, but for every causal anomaly there does exist an interpretation that ruled out this anomaly. For example, if we adopt the cor-

puscular interpretation, we have to face causal anomalies raising from some experiments, such as the two-slits experiment. In this experiment a beam of electrons is directed toward a diaphragm with two open slits and an interference pattern is produced on a screen behind the diaphragm; the probability that an electron, passing through an open slit, will reach the screen at a given point is depending on whether the other slit is open or closed—with the electron behaving as if it is informed about the state of the other slit.

This causal anomaly is eliminated if we adopt the wave interpretation, according to which the interference patterns are produced by waves in conformity with Huygens's principle. The wave interpretation is in turn affected by other anomalies raising from the so-called reduction of the wave packet: The wave originating from an open slit occupies a hemisphere centered on the slit, but when the wave hits the screen, a flash is produced in a point only and the wave disappears in all other points. Apparently all physical properties transported by the wave, such as momentum and energy, suddenly materialized in a single point, even if they were distant from this point. This situation is explained without anomalies by the corpuscular interpretation. According to Reichenbach, in every experiment about quantum mechanics we can adopt an interpretation free from causal anomalies, but we have to use a different interpretation in a different experiment. Only two interpretations are required: the wave and the corpuscular interpretation. This is the real meaning of the duality of wave and corpuscle in quantum physics. The possibility of eliminating causal anomalies from every quantum experiment is called, by Reichenbach, the principle of eliminability of causal anomalies.

The Bohr-Heisenberg restricted interpretation of interphenomena named after Danish physicist Niels Bohr and German physicist Werner Karl Heisenberg, states that speaking about values of unmeasured physical quantities is meaningless. Reichenbach criticizes the Bohr-Heisenberg interpretation on two points. First, Heisenberg's indeterminacy principle becomes a metastatement about the semantics of the language of physics; second, this interpretation implies the presence of meaningless statements in the language of physics.

Using a three-valued logic, in which admissible truth values are *truth*, *falsehood*, and *indeterminacy*, Reichenbach constructs another restrictive interpretation in which a statement about an unmeasured physical quantity can be neither true nor false, but indeterminate.

INTERPRETATIONS OF REICHENBACH'S PHILOSOPHY

An open question regards the relation between Reichenbach and conventionalism. His insistence on the major role played by the coordinative definitions, the relativity of geometry, the equivalence between wave and corpuscular interpretation of quantum mechanics has suggested that his philosophy can be ascribed to conventionalism. In Reichenbach's works there are some points corroborating this view. For example, he asserts that the philosophical meaning of the theory of relativity is that this theory proves the necessity of coordinative definitions, which are arbitrary, in situations in which empirical relations had been previously assumed. But there are also some elements against the conventionalist reading of Reichenbach's philosophy, as seen in the last paragraph of *The Philosophy of Space and Time*, in which Reichenbach affirms that the reality of space and time is an irrefutable consequence of his epistemological analysis; it is an assertion apparently incompatible with conventionalism. As an example of the debate about Reichenbach's attitude toward conventionalism, it is possible to mention the conventionalist interpretation of Reichenbach's philosophy developed by Adolf Grünbaum in *Philosophical Problems of Space and Time* (1973) and Hilary Putnam's counterarguments offered in "The Refutation of Conventionalism" (1975).

A different explication of Reichenbach's philosophy, based on an analysis of the role of the coordinative definitions in the light of Kantian philosophy, is advanced by Michael Friedman and exposed in *Reconsidering Logical Positivism* (1999). According to Friedman's interpretation, Reichenbach, in his first published work on the theory of relativity (*Theory of Relativity and A Priori Knowledge*), distinguishes two different meanings of synthetic a priori, which are united in Kantian philosophy. In the first meaning, a synthetic a priori judgment is necessary and thus not modifiable; in the second meaning, a synthetic a priori statement is constitutive of the object. The coordinative definitions are not necessary judgments, because we can make use of a different definition. Moreover, all coordinative definitions are subjected to changes with the evolution of knowledge, so they are modifiable. Thus they are not a priori in the first meaning present in Kantian philosophy. But the coordinative definitions are required to give an empirical interpretation to a theory and so they are constitutive of the object of knowledge. Thus they are synthetic a priori in the second meaning present in Kantian philosophy. Friedman calls this type of a priori judgment "constitutive, *rela-*

tivized a priori" (1999, p. 62), because they are a priori in the constitutive sense, relative to a given theory.

Surely Kantian philosophy exerts a great influence on Reichenbach. He professes admiration for Kant in his first works. In the article "Kant und die Naturwissenschaft" (1933, p. 626) he says, "There is no doubt that he [Kant] was one of the few thinkers whose work showed the way on which the contemporary philosophy of natural science continues to proceed." According to Reichenbach, Kantian philosophy of nature is a meaningful theory, although it is superseded by the outcomes of contemporary physics. Later, Reichenbach accentuates his departure from Kant, stressing his criticism of synthetic a priori and developing many arguments against Kantian philosophy.

See also Causation: Philosophy of Science; Philosophy of Statistical Mechanics; Time.

Bibliography

Barone, Francesco. *Il neopositivismo logico*. Bari, Italy: Laterza, 1986.

Friedman, Michael. *Dynamics of Reason*. Stanford, CA: Center for the Study of Language and Information, 2001.

Friedman, Michael. *Reconsidering Logical Positivism*. New York: Cambridge University Press, 1999.

Grünbaum, Adolf. *Philosophical Problems of Space and Time*. Dordrecht, Netherlands: Reidel, 1973.

Putnam, Hilary. "The Logic of Quantum Mechanics." In *Mathematics, Matter and Method. Philosophical Papers*. Vol. 1. Cambridge, U.K.: Cambridge University Press, 1975.

Putnam, Hilary. "The Refutation of Conventionalism." In *Mind, Language and Reality. Philosophical Papers*. Vol. 2. Cambridge, U.K.: Cambridge University Press, 1975.

Salmon, C. Wesley, and Gereon Wolters, eds. *Logic, Language, and the Structure of Scientific Theories: Proceedings of the Carnap-Reichenbach Centennial, University of Konstanz, 21–24 May 1991*. Pittsburgh: University of Pittsburgh Press, 1994.

Schilpp, Paul Arthur, ed. *Albert Einstein: Philosopher-Scientist*. Evanston, IL: Library of Living Philosophers, 1949.

Spohn, Wolfgang, ed. *Erkenntnis Orientated: A Centennial Volume for Rudolf Carnap and Hans Reichenbach*. Boston: Kluwer Academic Publishers, 1991.

REICHENBACH'S WORKS

"Kant und die Naturwissenschaft." *Die Naturwissenschaften* 21 (1933): 601–606, 624–626.

From Copernicus to Einstein. New York: Alliance Book Corp., 1942.

Philosophic Foundations of Quantum Mechanics. Berkeley: University of California Press, 1944.

Elements of Symbolic Logic. New York: Macmillan, 1947.

The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability. Translated by Ernest H. Hutten and Maria Reichenbach. Berkeley: University of California Press, 1949.

The Rise of Scientific Philosophy. Berkeley: University of California Press, 1951.

The Direction of Time. Edited by Maria Reichenbach. Berkeley: University of California Press, 1956.

The Philosophy of Space and Time. Translated by Maria Reichenbach and John Freund. New York: Dover Publications, 1957.

Axiomatization of the Theory of Relativity. Translated and edited by Maria Reichenbach. Berkeley: University of California Press, 1969.

Laws, Modalities, and Counterfactuals. Berkeley: University of California Press, 1976.

Mauro Murzi (2005)

REID, THOMAS

(1710–1796)

Thomas Reid was the founder of the Scottish "Common Sense" school of philosophy. A contemporary and critic of David Hume, he is best known for his staunch defense of common sense and trenchant opposition to the "way of ideas," the theory that the immediate objects of perception and other cognitive acts are always internal images or ideas, not external physical objects. His views exerted a good deal of influence until the mid-nineteenth century or so, when they began to be eclipsed by absolute idealism, pragmatism, and other philosophical movements, but they have been the subject of renewed interest from the 1970s on.

After being educated at Marischal College in Aberdeen, Scotland, Reid served for fifteen years as a parish minister in nearby New Machar. In 1752 he was appointed professor at King's College in Aberdeen, where he taught mathematics, physics, and philosophy, among other subjects. He tells us that in his youth he believed nearly the entire philosophy of George Berkeley but that a reading of Hume's *Treatise of Human Nature* (1739–1740) convinced him (by carrying Berkeley's philosophy to its logical conclusions) that there must be some original defect in it. Reid identified this defect as the theory of ideas, which he went on to challenge in college lectures, meetings of the Aberdeen Philosophical Society, and two books. In 1764 he published his first major work, *An Inquiry into the Human Mind on the Principles of Common Sense*, in which he set forth his reasons for opposing the theory of ideas and offered an alternative theory of how we gain knowledge by means of the various senses.

In the same year he accepted the chair in moral philosophy at Glasgow, succeeding Adam Smith. He lectured there until 1780, when he resigned to prepare his last two major works: *Essays on the Intellectual Powers of Man* (1785), devoted to the contributions of perception, memory, reason, and other cognitive powers to human knowledge, and *Essays on the Active Powers of the Human Mind* (1788), devoted to the nature of action, will, freedom, and morality.

This article provides a synopsis of Reid's main views more or less in the order in which he presented them in his three published books: the *Inquiry* (abbreviated as Inq.), the *Intellectual Powers* (abbreviated as IP), and the *Active Powers* (abbreviated as AP). Numbers separated by a period refer to chapter and section numbers in the *Inquiry* and to essay and chapter numbers in the two volumes of *Essays*.

CRITIQUE OF THE THEORY OF IDEAS

Almost alone among the great modern philosophers, Reid espoused a direct realist theory of perception. He repudiated the assumption that what is immediately present to the mind is never an external thing, but only an internal image, impression, representation, or (to use the most common eighteenth-century term) *idea*. Ideas were usually conceived of as mental entities that existed only as long as the mind was aware of them. Reid found the theory of ideas to be taken for granted in the work of most of his philosophical predecessors, including René Descartes, Nicolas Malebranche, Antoine Arnauld, John Locke, Berkeley, and Hume. Some of these philosophers (for example, Descartes and Locke) were realists, believing that ideas are caused in us by physical objects existing outside the mind. Others (notably, Berkeley) were idealists, repudiating the existence of a world outside the mind and believing that the things we call physical objects are simply bundles of ideas. In either case, we are cut off from direct perception of the physical world, either because there is no physical world to be perceived, or because our perception of it is indirect—not strictly perception at all, but inference based on what we do perceive, namely, ideas.

Reid makes at least three important points against the theory of ideas. First, the arguments in favor of the theory are weak and without cogency; second, the theory does nothing to explain how perception is possible; third, the theory stands in the way of our knowing or even being able to conceive of the physical world.

One of the arguments for the theory of ideas that Reid singles out for criticism is a version of the argument from perceptual relativity. Hume had claimed that the "universal and primary opinion of all men" that they perceive external objects directly is "destroyed by the slightest philosophy," offering the following argument in section 12 of the *Enquiry concerning Human Understanding*: "The table, which we see, seems to diminish as we remove further from it; but the real table, which exists independent of us, suffers no alteration. It was therefore nothing but its image which was present to the mind." Hume's argument may be cast into the following syllogism: (1) What I see diminishes in magnitude as I retreat from it; (2) the table itself does not diminish in magnitude as I retreat from it; (3) therefore, what I see is not the table itself (but only an image or idea).

Reid contends that Hume's premises are true only if we restate them as follows (IP 2.14, p. 182): (1) What I see diminishes in *apparent* magnitude as I retreat from it; (2) the table itself does not diminish in *real* magnitude as I retreat from it; (3) therefore, what I see is not the table (but only an image or idea).

The real magnitude of an object (for example, the edge of a table) is an intrinsic property of it, measured in feet or inches, whereas the apparent magnitude of an object is a relation between the object and a perceiver (or his vantage point), measured by the angle the object subtends at the eye. Reid takes the terminology of "real" versus "apparent" from the astronomy of his day; it is not necessarily implied that there is anything illusory about apparent magnitude. It is easy to see that apparent magnitude varies with the distance between object and perceiver (objects subtending smaller angles when further away) whereas real magnitude does not. Once we record these facts correctly, as in Reid's version of the syllogism, we see that the conclusion of the argument does not follow from the premises. Moreover, Reid would resist the thought that if O has greater apparent magnitude when seen from p than when seen from p', that is because it presents a larger image to the observer at p than to the observer at p'. Apparent magnitude is a strictly dyadic relation, involving only the object and the perceiver (or his vantage point) and no third thing such as a mental image.

Reid's second point against the hypothesis of ideas is "that ideas do not make any of the operations of the mind to be better understood" (p. 184). Ideas had been thought necessary to explain how we perceive things that are distant, remember things that are past, or imagine things that do not exist at all, but Reid argues that all such explanations are worthless. They presuppose that ideas themselves can somehow be of the remote, the past, or the

nonexistent. But if ideas can do that, what prevents the simple idealess acts of perceiving, remembering, and imagining from doing it as well? Moreover, our ability to be aware of ideas themselves is no less mysterious than our ability to be aware of things that are not ideas.

> It is as difficult to conceive how the mind perceives images in the brain as how it perceives things more distant. If any man will shew how the mind may perceive images in the brain, I will undertake to shew how it may perceive the most distant objects: for if we give eyes to the mind, to perceive what is transacted at home in its dark chamber, why may we not make these eyes a little longer-sighted? (Inq. 6.12, p. 121)

Reid's third point against the theory of ideas is "that the natural and necessary consequences of it furnish a just prejudice against it to every man who pays a due regard to the common sense of mankind" (p. 185). Chief among these consequences is that if we do not simply see or touch external objects, it becomes necessary to prove their existence by arguments. Descartes, Malebranche, and Locke all tried to muster such arguments, but none of the arguments is convincing. Reid thus thinks that skepticism about the material world is a built-in consequence of the theory of ideas. By contrast, if what we see and touch are not ideas but things in the external world, as in Reid's own view, this source of skepticism is eliminated.

SENSATION AND PERCEPTION

A sensation is an event that occurs in a sentient subject when he or she smells a rose or tastes a fig. It lacks figure and extension and other qualities of bodies, being entirely mental. Reid calls sensations "principles of belief," by which he means that when we have a sensation and attend to it, we cannot help believing that it exists, that a subject of it exists (ourselves), and that some external object (for example, some quality in the rose) exists as its cause.

Reid is among the first to distinguish between sensation and perception. He explains this distinction as follows:

> Thus, *I feel a pain*; *I see a tree*: the first denoteth a sensation, the last a perception. The grammatical analysis of both expressions is the same: for both consist of an active verb and an object. But, if we attend to the things signified by these expressions, we shall find, that in the first, the distinction between the act and the object is not real but grammatical; in the second, the distinction is not only grammatical but real. The form of the expression, *I feel pain*, might seem to imply, that the feeling is something distinct from the pain felt; yet, in reality, there is no distinction. As *thinking a thought* is an expression which could signify no more than *thinking*, so *feeling a pain* signifies no more than *being pained*. What we have said of pain is applicable to every other mere sensation. (Inq. 6.20, pp. 167–68)

When I perceive a tree, there is an object (the tree) apart from my act of seeing, but when I have a sensation, there is no object apart from the act of sensing. As he defines sensation in the *Intellectual Powers*, it is an act of the mind "which may be distinguished from all others by this, that it hath no object distinct from itself" (IP 1.1, p. 36). That formulation is ambiguous: Does an act of sensing have itself for its object, or does it have no object at all? Although Reid's language often suggests the former option, his proposal that *being pained* is the model for all sensation suggests the latter option. If we take Reid in the latter way, he is a precursor of "adverbial" theories of sensation, such as were developed by C. J. Ducasse and Roderick Chisholm two centuries later: to have a sensation of red is not to sense some*thing*, but is simply to sense some*how*—"redly," as the adverbial theory styles it.

Some critics of Reid have thought that his sensations are simply ideas under a new name, but there are important differences—especially if he holds an adverbial theory rather than a theory that divides sensation into act and object. If sensing required its own special objects, the argument from perceptual relativity against direct realism could be reinstated. The mountain that looks blue from a distance and green from close up would do so by generating first blue and then green sensory objects in my mind, and these variously colored objects would have to be distinct from the unchanging mountain. They would displace the mountain itself as my object of direct awareness. But Reidian sensations do not have objects to get in the way of direct perception of external things.

Although sensations do not *have* objects, they can *become* objects for us, in the sense that we can know through proper attention what sorts of sensations we are having. Indeed, Reid thinks that if we attend carefully to our sensations, we can know perfectly what they are like and can scarcely make any mistake about them. Yet typically we pay so little attention to them that we become almost oblivious to them; they serve as mere cues or signs from which our minds leap instantly to other things that they signify.

Our apprehension of that which sensations signify is *perception*. Reid's official characterization of perception involves three elements: conception, belief, and immediacy:

> If, therefore, we attend to that act of our mind which we call the perception of an external object of sense, we shall find in it these three things:—*First*, Some conception or notion of the object perceived; *Secondly*, A strong and irresistible conviction of its present existence; and *Thirdly*, That this conviction and belief are immediate, and not the effect of reasoning. (IP 2.5, p. 96; cf. Inq. 6.20, p. 168)

Note that this definition makes no mention of sensation. Although Reid says that sensation generally serves as the trigger for the conception and belief involved in perception, perception proper is just conception plus immediate belief. Reid thinks it possible that perception should occur in the absence of sensation, and he holds that there is one variety of human perception that actually does occur without any characteristic sensation: namely, the perception of visible figure. Reid thus deemphasizes the role of sensation in perception in a way that some contemporary theorists (for example, James J. Gibson) would applaud. By the same token, however, his threefold definition may strike others as leaving out precisely that by which a genuine perception of a snake in the path ahead is distinguished from the conception and immediate belief in it one may form as the result of a friend's warning. Here Reid's views may gain in plausibility if we reckon his "conception" as something like what Bertrand Russell called knowledge by acquaintance. It is not necessarily the exercise of a concept in mere thought.

REID'S NATIVISM

Reid thought that much of what he found alarming in Hume's philosophy stemmed from Hume's adherence to the empiricist maxim that we have no ideas or notions that are not derived from previous impressions or sensations. It is by this principle that Hume was led to conclude that we have no legitimate notions of objects existing unperceived, of causal connections amounting to more than constant conjunctions, and of a self that is the subject of various mental operations. Reid sought to overthrow Hume's philosophy by undermining its foundations, and for this purpose he tackled the empiricist principle head-on. He pointed to a notion that he thought Hume would surely concede that we possess—the notion of *extension*, or being spread out in space—and contended that this notion lacks a proper Humean birthright in our sensations. If it were once acknowledged that not even so uncontroversial a notion as extension can be extracted from our sensations or impressions, Reid thought, the way would be open for recognizing the legitimacy of other notions with no sensory origin, such as the ideas of agency, self, and an external world.

To back up his contention that the notion of extension is not derived from sensation, Reid offers a thought experiment he calls his *experimentum crucis* (Inq. 5.6 and 5.7, pp. 65–72). He asks us to imagine a being furnished with a progressively richer array of sensations, beginning with those caused by the prick of a pin, advancing to more complex sensations such as those caused by the pressure of a blunt object against his or her body, and culminating with the sensations accompanying the motion of his or her limbs. He asks at each step in the series whether those sensory materials would suffice to give a being who reflected upon them a conception of extension, and his answer is no. Positively, Reid's doctrine is that the conception of extension is innate—not in the sense that we have it from birth, but in the sense that it is triggered in us by certain sensations from which it could never have been derived from any process of abstraction or ratiocination. We are enabled to form the conception of extended things only because we are innately programmed to do so.

For further light on the import of Reid's nativism, we may restate it in terms of the threefold classification of natural signs he offers in sections 4.2 and 5.3 of the *Inquiry*. Reid first divides signs into the artificial and the natural. In the former class, the connection between sign and thing signified is established by compact or convention, as with the words of human language. In the latter class, the connection between sign and thing signified is established by nature, as with smoke and fire and other cases of effect and cause. Reid then further divides natural signs into three classes. In the first class, the connection between sign and thing signified is "established by nature, but discovered only by experience" (Inq. 5.3, p. 59), as in the example of smoke and fire already given. In the second class, the connection is "not only established by nature, but discovered to us by a natural principle, without reasoning or experience" (Inq. 5.3, p. 60).

Reid thinks that certain features of the human countenance are signs in this sense of thoughts and other mental states. For example, an infant is innately disposed to read a smile on its mother's face as a sign of approval without having to learn this connection through experience. Unless there were a basic repertoire of natural signs of this second class, Reid believes, the signification of arti-

ficial signs could never be agreed upon or learned. Finally, in the third class are those signs "which, though we never before had any notion or conception of the thing signified, do suggest it, or conjure it up, as it were, by a natural kind of magic" (Inq. 5.3, p. 60). Not only is the *connection* between sign and thing signified innately programmed into our constitution (as with signs of the second class), but also the very *notion* of the thing signified is innate in the sense that it is in no way derivable by abstraction from any of our sensations. Reid believes that the tactile sensations to which we respond with conceptions of extended bodies are natural signs belonging to this third class.

Reid takes his nativism to afford an answer to an argument for skepticism he finds embodied in the combined philosophies of Berkeley and Hume. He formulates the argument as follows (Inq 5.8, p. 75): (1) We can have no conception of anything but what either resembles or is deducible from our sensations; (2) nothing resembles or is deducible from sensations but other sensations; (3) therefore, we can have no conception of anything but sensations.

If the argument were correct in both its premises, it would follow that we cannot even *conceive* of, let alone have knowledge of, a world lying beyond our sensations. Reid thinks the second premise is correct, and he credits Berkeley with having made it evident. But he thinks the first premise—which states in Reid's language Hume's principle that all our ideas are copied from precedent impressions—is false. "That we have clear and distinct conceptions of extension, figure, motion, and other attributes of body, which are neither sensations, nor like any sensation, is a fact of which we may be as certain, as that we have sensations" (Inq. 5.8, p. 76).

THE MECHANICS AND GEOMETRY OF VISION

More than half of the *Inquiry* is devoted to vision, which Reid regards as the noblest of the senses. It informs us of the properties of objects far distant, such as the sun and the moon, and it can disclose in a glance the figure of a cathedral, whose delineation by touch would be the work of a lifetime.

Reid provided solutions to a number of puzzles about vision that lie today within the province of cognitive science rather than philosophy. For example, why do we see things upright despite having inverted retinal images of them? To explain this, Reid appeals to the law that an object will be seen in the direction of a straight line drawn from the point of retinal stimulation through the center of the eye and into ambient space. Why do we normally see objects single despite having two retinal images of them, yet under certain circumstances see them double? Reid's answer appeals to the law of corresponding retinal points: If rays from an object fall on points of the two retinas lying at equal distances and directions from their centers, the object will be seen as single, but otherwise as double.

One of Reid's more remarkable findings is that the visible figures of objects are governed by a non-Euclidean geometry. Reid believed that sight by itself (before we have learned any correlations with touch) informs us only of the two-dimensional spatial features of objects. Although the objects we see are at a distance from us (*pace* Berkeley), the eye is incapable of making any discriminations of depth. To an eye placed at the center of a sphere and looking out, great circles on the surface of the sphere (whose outward curvature is invisible to the eye) must appear as straight lines, and every figure seen by the eye must have the same geometrical properties as some figure drawn upon the sphere. In consequence of this, Reid argued that the geometry of visibles is what we would nowadays classify as a Riemannian geometry. A visible triangle, unlike a triangle perceived by touch, always has an angle sum at least slightly greater than 180 degrees, and no two visibly straight lines are ever strictly parallel.

MEMORY

Essay III of the *Intellectual Powers* is devoted to memory. Reid characterizes memory as "an immediate knowledge of things past" (IP 3.1, p. 253). There are two senses in which this is true. First, the object of memory is the very thing or event formerly perceived, not some present idea or simulacrum of it (Inq. 2.3, p. 28). Second, the knowledge one has by memory of this past object is noninferential; it does not rest on any reasoning from premises. Memory is thus like perception for Reid in involving both the conception of an object and an immediate belief in its existence; but it differs from perception because the object is, and is believed to be, past. Reid is severely critical of Hume's attempt to distinguish imagining, remembering, and perceiving solely in terms of the force and vivacity of their objects.

Reid criticizes Locke's view that memory is what constitutes personal identity—that person A is identical with person B who existed in the past if and only if A remembers what B did. He insists that memory is the *evidence* of personal identity, rather than that in which it consists (IP 3.4, p. 265). He also presents the famous

"brave officer" objection to Locke's theory, courtesy of his friend George Campbell: Suppose a man who has become a general late in life remembers capturing the enemy flag as a young officer; that as an officer he remembered being flogged as a boy for robbing an orchard; and that as a general he no longer remembers being flogged as a boy. It follows from Locke's theory that the general both is and is not the same person as the boy (IP 3.6, p. 276).

CONCEPTION AND ABSTRACTION

Essays IV and V of the *Intellectual Powers* are devoted to conception and abstraction. Conception, the most basic operation of the mind, is "that operation of the understanding which the logicians call *simple apprehension*," that is, the apprehension of a thing without any belief or judgment about it (IP 4.1, p. 295). "Judgment can be expressed by a proposition only, and a proposition is a complete sentence; but simple apprehension may be expressed by a word or words, which make no complete sentence" (IP 6.1, p. 408). The objects of conception expressed by words or subsentential phrases are either individuals or universals. What Reid calls conception should not be confused with conceptualization—that is, subsuming something under a concept—for the latter is judgment, and conception is more basic than judgment.

Reid holds that all the operations of our minds except sensation have objects distinct from themselves. "He that conceives, must conceive something" (IP 4.1, p. 311), and the same goes for perception and memory. It is a distinctive feature of conception, however, that its objects need not exist: "it is not employed solely about things which have existence" (IP 4.1, p. 310).

On this point, Reid is sometimes seen as a precursor of Alexius Meinong, who held that there can be cognitive relations to the utterly nonexistent and that a thing therefore need not exist in order to stand in relations. Meinong's view strikes many as paradoxical. Yet Reid makes it look like one more piece of common sense or, at any rate, a consequence of two pieces of common sense (IP 4.1, p. 311): (1) I can conceive of a centaur; (2) no centaur exists; (3) therefore, I can conceive of what does not exist. In case anyone objects that the truth in premise 1 is simply that I can conceive of the *idea* of a centaur, which *does* exist, Reid is ready with a reply: I know the difference between conceiving of a centaur and conceiving of the idea of a centaur, and I can assure you that I am doing the former rather than the latter (IP 4.2, p. 321).

Reid's view that the objects of conception may be nonexistent has an interesting application to the problem of abstract ideas, which pitted Locke against Berkeley and Hume. On this topic, Reid writes, "Mr. Locke and his two antagonists have divided the truth between them" (IP 5.6, p. 394). Locke saw clearly "that the power of forming abstract and general conceptions is one of the most distinguishing powers of the human mind," but he did not see "that this power is perfectly irreconcileable to his doctrine concerning ideas." Berkeley and Hume "saw this inconsistency; but instead of rejecting the hypothesis of ideas, they explain away the power of abstraction."

To see how Locke and his critics "divided the truth between them," consider the following inconsistent triad of propositions:

(1) We are sometimes aware in thought of the general and the abstract—in Reid's terminology, we have the power of forming abstract and general conceptions.

(2) We can only be aware of what exists: "in all of the operations of the understanding, there must be an object of thought, which really exists while we think of it" (IP 4.2, p. 312).

(3) General entities have no existence: "every thing that really exists is an individual" (IP 5.6, p. 393).

As Reid saw it, Locke accepted both 1 and 2 and was therefore driven to deny 3, despite his affirmation of it elsewhere. He posited "abstract general ideas," such as the infamous image of a triangle that is neither isosceles nor scalene, as merely generic entities existing in the mind. Berkeley and Hume, on the other hand, both accepted 2 and 3, and were thus led to reject 1. Not believing that entities such as Locke's merely generic triangle could exist even in the mind, they denied that we are ever aware of general entities. Thus were born their attempts to explain how we can think generally (for example, in proving propositions about all triangles) by means of ideas that are particular.

Reid's novelty is to deny proposition 2, which he castigates as one of the prejudices giving rise to the theory of ideas. It led all three of his predecessors in the British Empiricist tradition to affirm that the immediate object of awareness, in conception as well as in perception, must be an idea. By denying 2, Reid was enabled to uphold both 1 and 3, thus collecting together the truths his predecessors had divided between them.

FIRST PRINCIPLES

Essay VI of the *Intellectual Powers* contains an extensive and important treatment of what Reid calls first principles. A first principle is a self-evident proposition—a proposition that is evident to us without need of any rea-

sons to support it. Like Aristotle, Reid thinks that our knowledge must ultimately rest on first principles, for without them we would be faced with an infinite regress of supporting propositions. He may therefore be classified as an epistemological foundationalist. Reid believes there are first principles both of necessary and of contingent truths. The first principles of necessary truths include axioms of logic, mathematics, grammar, metaphysics, and morals. The first principles of contingent truths include principles pertaining to the deliverances of consciousness (Reid's term for introspection), perception, memory, inductive reasoning, and others of our faculties.

In Reid's enumeration of the first principles of contingent truths, there is a subtle ambiguity that greatly affects how his epistemology is to be interpreted. Here is how he formulates Principle 1, which gives us the first principle(s) regarding consciousness: "First, then, I hold, as a first principle, the existence of every thing of which I am conscious" (EIP 6.5, p. 470). Putting this in terms of truth rather than existence, he might just as well have said that he holds, as a first principle, the truth of every proposition to which consciousness testifies. The ambiguity in Principle 1 may then be brought out by the following two ways of symbolizing it, where "Cp" is short for "I am conscious that p":

> 1.1 It is a first principle that (p)(Cp -> p).(It is a first principle that for any proposition p, if I am conscious that p, then p.)
>
> 1.2 (p)(Cp -> it is a first principle that p).(For any proposition p, if I am conscious that p, then it is a first principle that p.)

The difference between the formulations is this: 1.1 says that it is a first principle that all the deliverances of consciousness are true. In other words, 1.1 gives us *one general proposition* as first principle. 1.2, on the other hand, says that each of the deliverances of consciousness—which may include propositions such as *I am now in pain*—are themselves first principles. So 1.2 gives us *many particular propositions* as first principles. A similar ambiguity holds in regard to the first principles of perception and memory.

How should Reid's first principles be understood—as general or particular? Perhaps the best overall interpretation of Reid's epistemology is provided by the particularist construction. At the very least, his epistemology must be understood as recognizing particular first principles even it recognizes general first principles as well.

If Reid's first principles are construed in the particularist way, he is not only a foundationalist in his epistemology but also (in one important sense) an *externalist.* Externalists hold that there are sources or factors that give a subject knowledge even if the subject does not know anything about how the factors work or whether they are reliable. On the particularist construction of Reid's principles, the mere fact that a proposition is a deliverance of perception, memory, or consciousness suffices to make that proposition evident (and thus, in favorable circumstances, known). To know that there is a tree over there, for example, one need only have a perception of a tree. It is not necessary for the subject to know anything about the reliability of sense perception, which may be a matter to which he has never given any thought. On the generalist construction of the principles, by contrast, the subject would presumably have to know that the general principles are true in order for knowledge of particular propositions to arise in accordance with them. In other words, he would have to know his faculties are reliable before they could be sources of any other knowledge. That puts an obstacle in the way of knowledge that skeptics might claim to be insurmountable.

It may be useful to summarize by drawing together the various things Reid has to say in response to skepticism about the material world. First, what skeptical philosophers profess cannot be believed and is not believed even by skeptics themselves. This is a point that Hume famously admitted, and it may be questioned what force it has against the truth or reasonableness of the skeptic's position. Second, the argument that we cannot even conceive of a material world is answered by Reid's nativism, according to which we are endowed by our constitution with conceptions of extended external objects.

Third, the argument that knowledge of the external world must be based on problematic inferences from our ideas is undercut by Reid's attack on the theory of the ideas. Fourth, the position of the "semiskeptic," who says we can be certain about the deliverances of our consciousness but not about anything else, is objectionably arbitrary. For who can prove that consciousness never errs? And what reason is there to believe the deliverances of consciousness that is not a reason for believing the deliverances of our other faculties as well? (Inq. 5.7, p. 71).

Finally, the position of the total skeptic, who refuses to accept the deliverance of any faculty unless its reliability is proven in advance, is irrefutable (Inq. 5.7, p. 71; cf. IP 6.5, p. 480). We cannot *show* that he or she is wrong without assuming something he or she would question.

But perhaps, for all that, we may *know* that he or she is wrong, if Reid's externalist approach to epistemology is correct.

CAUSATION AND FREEDOM

As noted above, Reid thinks we have many conceptions, such as that of a self or subject of mental operations, that we could not have on Humean principles. Among them is the conception of active power, or real efficacy in bringing about changes, to which Reid turns in his third book. He thinks we obtain a clear conception of such power when we are conscious of our own activity in bringing something about by an act of will. Active power is exercised only by agents or substances, not by events, so in the strictest sense of causation, only agents are causes for Reid. When we speak of one event causing another, Reid tells us, it would be more proper to speak of events related by lawful sequence or a relation of sign and thing signified.

That we sometimes act freely (or that we possess "moral liberty") is, according to Reid, a natural conviction, comparable to our belief in a material world. In the *Active Powers*, he offers three arguments that we really do possess such liberty. The first is based on the "naturalness" of our conviction in regard to it, the second on the notion of moral responsibility, and the third on our ability to secure ends by prosecuting a long series of means.

Reid rejects accounts of moral liberty such as those of Hobbes and Hume, who seek to make liberty compatible with determinism. He would reject Hume's suggestion that I did A freely if I did A as a result of willing to do it and would have done otherwise if I had willed otherwise. In a universe in which my willing was itself the end of a causal chain stretching back to the Big Bang, the conditions of this definition might be satisfied, yet I would not, according to Reid, have acted freely. It is a further requirement of liberty that my willing not have been determined by antecedent events in that way. But that is not to say that my willing must be random or uncaused. On the contrary, in a case of free action, it is caused by *me*, the agent. In this way Reid brings his theory of agent causation into his account of liberty, attempting to escape the dilemma that has determinism as one horn and arbitrary uncaused acts of will as the other. Reid believes that every event has a cause, but he holds that the cause of an event need not be another event—it may be an agent.

Agent-causation theories of human action inspired by Reid began to undergo a revival during the last third of the twentieth century, finding advocates in Roderick Chisholm and Richard Taylor, among others. Such theories offer a tantalizing glimmer of hope for resolving old problems yet face formidable problems of their own. If I am the cause of my willing to do A, mustn't there be such an event as my causing the willing? If so, what is the cause of *that* event? If it is *nothing*, we have fallen back on the randomness horn and violated Reid's professed belief that every event has a cause. If it is a *further event*, we are back on the horn of determinism. If it is the *agent*, we have taken the first step of an infinite regress in which I am the cause of my willing A, the cause of my causing of my willing A, and so on, *ad infinitum*.

MORAL PHILOSOPHY

Reid is often considered to be a member of the moral-sense school of philosophy, insofar as he holds that moral notions and moral determinations are the product of a moral faculty or sense. He insists, however, that the employment of the term *sense* is accurate only with the proviso that a sense can deliver judgments as well as feelings. In opposition to Hume, he holds that "moral approbation implies a real judgment" (AP 5.7, pp. 457–481), capable of being true or false, and is not merely an expression of feeling like "Hurrah!" (It must be said, however, that his criticisms of Hume sometimes convert the supposedly noncognitivist view he is attacking into a subjectivist form of cognitivism). In further opposition to Hume, he holds that reason is not merely the slave of the passions but has a real role to play in the selection of ultimate ends of action (AP, 5.3).

Reid also opposes another kind of view that sometimes goes under the rubric of moral-sense theory: the view that moral properties are analogous to secondary qualities, as in the suggestion that for an action to be right is for it to arouse favorable moral emotion in those who contemplate it. Reid protests that such accounts abolish the necessity of moral principles. It is necessary, according to him, that actions of certain types are right but contingent that they produce whatever effects they do in those who contemplate them (IP, 6.6, pp. 494–495). On the whole, Reid's views probably bear less resemblance to moral-sense theories than they do to the intuitionism of G. E. Moore. Much of what Reid says about *right* anticipates what Moore said about *good*: that it is indefinable, that we understand what it is by an original power of the mind, and that our moral faculty provides us with first principles about which types of acts are right and which wrong.

See also Aristotle; Arnauld, Antoine; Berkeley, George; Causation; Chisholm, Roderick; Common Sense;

Descartes, René; Ducasse, Curt John; Geometry; Hobbes, Thomas; Hume, David; Introspection; Locke, John; Malebranche, Nicolas; Meinong, Alexius; Moore, George Edward; Nativism, Innatism; Russell, Bertrand Arthur William; Smith, Adam.

Bibliography

WORKS BY REID

The Works of Thomas Reid, D. D., edited with notes by Sir William Hamilton. 2 vols. 8th ed. Edinburgh, U.K.: James Thin, 1895. Reprint Hildesheim: Georg Olms, 1983.

Essays on the Active Powers of Man, edited by Baruch Brody. Cambridge, MA: MIT Press, 1969.

An Inquiry into the Human Mind on the Principles of Common Sense, edited by Derek R. Brookes. Edinburgh, U.K.: Edinburgh University Press, 1996. Reprint University Park: Pennsylvania State University Press, 1997.

Essays on the Intellectual Powers of Man, edited by Derek R. Brookes. University Park: Pennsylvania State University Press, 2002.

The Correspondence of Thomas Reid, edited by Paul Wood. University Park: Pennsylvania State University Press, 2002.

WORKS ABOUT REID

Alston, William P. "Thomas Reid on Epistemic Principles." *History of Philosophy Quarterly* 2 (1985): 435–452.

Barker, Stephen F., and Tom L. Beauchamp, eds. *Thomas Reid: Critical Interpretations*. Philadelphia: Temple University Philosophical Monographs, 1976.

Broadie, Alexander. "Reid Making Sense of Moral Sense." *Reid Studies* 1 (1998): 5–16.

Cummins, Phillip. "Reid's Realism." *Journal of the History of Philosophy* 12 (1974): 317–340.

Dalgarno, Melvin, and Eric Matthews, eds. *The Philosophy of Thomas Reid*. Dordrecht, Netherlands: Kluwer, 1989.

De Bary, Philip. *Thomas Reid and Skepticism: His Reliabilist Response*. London: Routledge, 2002.

Daniels, Norman. *Thomas Reid's Inquiry: The Geometry of Visibles and the Case for Realism*. Palo Alto, CA: Stanford University Press, 1974.

Falkenstein, Lorne. "Reid's Account of Localization." *Philosophy and Phenomenological Research* 61 (2000): 305–328.

Gallie, Roger. *Thomas Reid and "The Way of Ideas."* Dordrecht, Netherlands: Kluwer, 1989.

Grave, S. A. *The Scottish Philosophy of Common Sense*. Oxford, U.K.: Clarendon Press, 1960.

Haldane, John. "Thomas Reid: Life and Work." *American Catholic Philosophical Quarterly* 74 (2000): 317–344. This issue contains eleven articles devoted to Reid.

Immerwahr, John. "The Development of Reid's Realism." *The Monist* 61 (1978): 245–256. This issue is entirely devoted to Reid and includes thirteen articles on his philosophy.

Lehrer, Keith. *Thomas Reid*. London: Routledge, 1989.

McKitrick, Jennifer. "Reid's Foundation for the Primary/Secondary Quality Distinction." *The Philosophical Quarterly* 52 (2002): 478–494. This issue includes a previously unpublished manuscript by Reid ("Of Power") and ten articles on his philosophy.

O'Connor, Timothy. "Thomas Reid on Free Agency." *Journal of the History of Philosophy* 32 (1994): 605–622.

Rowe, William L. *Thomas Reid on Freedom and Morality*. Ithaca, NY: Cornell University Press, 1991.

Van Cleve, James. "Thomas Reid's Geometry of Visibles." *The Philosophical Review* 111 (2002): 373–416.

Wolterstorff, Nicholas. *Thomas Reid and the Story of Epistemology*. Cambridge, U.K.: Cambridge University Press, 2001.

Woudenberg, Rene van, and Terence Cuneo, eds. *The Cambridge Companion to Reid*. Cambridge, U.K.: Cambridge University Press, 2004.

Yaffe, Gideon. *Manifest Activity: Thomas Reid's Theory of Action*. Oxford: Clarendon Press, 2004.

For complete bibliographies of Reid's own works and secondary sources, see *Reid Studies* 1 (1998): 71–81; 2 (1998): 57–78; and 4 (2000): 93–96.

The Center for Scottish Philosophy, based at the University of Aberdeen, maintains the following web site with information about events and publications relating to Reid: http://www.abdn.ac.uk/cssp/

James Van Cleve (2005)

REIMARUS, HERMANN SAMUEL

(1694–1768)

Hermann Samuel Reimarus, the German philosopher and theologian, was born in Hamburg and studied theology at Jena. After serving as a lecturer in Wittenberg and as director of a high school in Wismar, he became a teacher of oriental languages at the Johannes-gymnasium in Hamburg. He began writing very late in life, when he was about sixty. One of his most important works, *Apologie oder Schutzschrift für die vernünftigen Verehrer Gottes* (Apology for or Defense of the Rational Worshiper of God), was first published by Gotthold Ephraim Lessing—posthumously and only in part—as fragments of an allegedly anonymous manuscript found in the Wolfenbüttel Library, where Lessing was librarian ("Wolfenbüttler Fragmente eines Ungenannten," in *Beiträge zur Geschichte und Literatur,* 1774–1777).

Reimarus was originally a Wolffian, and Wolffianism was a lasting foundation for his thought; but he developed individual doctrines in both philosophy and theology as one of the "popular philosophers." He stressed the moral aim of philosophy, that is, the happiness and moral perfectibility of man. He dissented from Christian Wolff chiefly in his views of philosophical methodology. He wrote in a "popular," or nonscholastic, style; he asserted that philosophy can be neither as certain as mathematics nor mathematically demonstrated; he stressed the func-

tion of common sense in knowledge; and he tried to simplify logic. In metaphysics, his main points of divergence from Wolff were his admission of a real interaction of soul and body and his view that life cannot be mechanically explained, but that it is an effect of the soul.

Reimarus's most important work was in the field of animal psychology and in his classification of the instincts of animals. Humans, unlike animals, have only a very few instincts. This lack may be a disadvantage for material life, but it is the basis for morality.

Reimarus appeared in his lifetime to be a moderate advocate of natural religion who did not openly oppose Christian revelation. But in the posthumous *Apologie* he submitted Christian revelation to a radical criticism in the spirit of English deism. In this work, for the first time in Germany the traditional view of Christianity was attacked neither on a speculative plan nor through superficial historical arguments, but on the basis of sound historical scholarship. Reimarus pointed out discordances between the Old and the New Testaments and between the different sections of each. He refused to accept the Gospels as the word of God, but described them as being an exposition of theological views elaborated by Jesus' successors in the leadership of Christianity. He considered the accounts of miracles, and in particular the account of the resurrection of Jesus, to relate events that never happened and to be forgeries of the Apostles. This purely rationalistic criticism made a tremendous impression on late eighteenth-century Germany, and deeply influenced the subsequent evolution of German theology.

See also Animal Mind; Deism; Lessing, Gotthold Ephraim; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY REIMARUS

Abhandlungen von den vornehmsten Wahrheiten der natürlichen Religion. Hamburg, 1754.

Vernunftlehre. Hamburg arid Kiel, 1756.

Allgemeine Betrachtungen über die Triebe der Tiere. Hamburg, 1760.

Uebrigen noch ungedruckte Werke des Wolfenbüttelschen Fragmentisten, edited by C. A. E. Schmidt. Berlin, 1786. Contains further portions of the *Apologie.*

WORKS ON REIMARUS

Aner, Karl. *Die Theologie der Lessingzeit.* Halle: Niemeyer, 1929; reprinted Hildesheim, 1964.

Büttner, W. *H. S. Reimarus als Metaphysiker.* Würzburg, 1909.

Engert, Joseph. *Der Deismus und die Religions- und Offenbarungskritik des H. S. Reimarus.* Vienna, 1915.

Engert, Joseph. *H. S. Reimarus als Metaphysiker.* Paderborn, 1909.

Fittbogen, G. "Lessings Entwicklung bis zur Begegnung mit Reimarus Werk." *Preussischer Jahrbücher* 73 (1918).

Richardt, H. *Darstellung der moralphilosophischen Anschauungen H. S. Reimarus.* Leipzig, 1906.

Schettler, R. *Die Stellung des Philosophen H. S. Reimarus zur Religion.* Leipzig, 1904.

Strauss, D. F. *H. S. Reimarus und seine Schutzschrift für die vernünftigen Verehrer Gottes.* Leipzig, 1862.

Giorgio Tonelli (1967)

REINCARNATION

The doctrine variously called transmigration of souls, metempsychosis, palingenesis, rebirth, and "reincarnation" has been and continues to be widely believed. Although some of these terms imply belief in an immortal soul that transmigrates or reincarnates, Buddhism, while teaching rebirth, denies the eternity of the soul. The word *rebirth* is therefore the most comprehensive for referring to this range of beliefs.

In one form or another the doctrine of rebirth has been held in various cultures. It was expressed in ancient Greece (Pythagoras, Empedocles, Orphism, Plato, and later, Plotinus); among some Gnostics and in some Christian heresies such as the medieval Cathari; in some phases of Jewish Kabbalism; in some cultures of tropical Africa; and most notably in such Eastern religions as Jainism, Buddhism, Hinduism, and Sikhism. Some European philosophers, notably Arthur Schopenhauer and J. M. E. McTaggart, have incorporated the doctrine into their metaphysics. The origin of the doctrine of rebirth as a religious belief is obscure. There is evidence, both in Greece and India, that it was not characteristic of early Aryan cultures. It is virtually certain that in India it goes back to prehistoric times; it was then taken up by Brahmanic religion and appears as a new doctrine in the Upaniṣads.

Views vary about the scope and mechanism of rebirth. It is part of Indian thought, for instance—but not of African beliefs—that men can be reborn as animals and even as plants (not to mention as gods and spirits). Rebirth can take place not merely on Earth but also in a multiplicity of heavens and purgatories. Thus, although the prevalent belief is that rebirth occurs immediately upon death, this does not entail immediate earthly reincarnation, a feature that helps to make rebirth theory incapable of empirical disproof. In the Buddhist *Tibetan Book of the Dead,* however, a transitional period (*bardo*) of forty-nine days between death and rebirth is postulated. During this state the individual is translated to a

realm where he perceives the divine secrets; for the impure, these are so frightening that they flee back to earth and are reborn.

In Indian thought, there is a fairly large amount of speculation about the embryological mechanics of rebirth. Thus the Sāṃkhya school of Indian philosophy holds that the mental aspect of a person bears the impression of previous deeds (karma) and that it accordingly becomes associated with a particular fetus. But since during the period of fetal development the growing body is not capable of supporting the mental aspect, a "subtle" (unobservably refined) body is postulated. Thus the continuous element throughout rebirth and until liberation is the mental aspect associated with the subtle body.

In Buddhism it is held that the fetus results from the interaction of the sperm and material in the mother. These combine in a suitable way when associated with conscious states, as a further element in the process, to produce the right sort of individual to fit previous karma. Broadly speaking, then, rebirth theory implies that the genetic endowment of a person does not fully determine his early development but that a mental or spiritual factor associates itself with a suitable organism at conception. Thus karma is often taken to function through the homing of a soul upon a morally and physically appropriate fetus. McTaggart, in urging this, uses the analogy of chemical affinities.

A number of arguments in favor of the theory have been propounded; they can be classified as metaphysical, empirical, and theological. It is convenient to record here those arguments that do not depend too closely on metaphysical conclusions peculiar to particular philosophers, such as the argument for rebirth as accounting for knowledge of the Forms, as in Plato, and the complex metaphysical argument in McTaggart that depends in part on his theory of causation.

In Indian sources, two main metaphysical arguments have been employed. It may be noted that there has been relatively little explicit discussion of the issue in Indian philosophy, since no school was concerned with denying the doctrine, except the Materialist school, which was extinct by medieval times. (1) A Buddhist argument can be expressed as follows. All states have prior causes; some conscious states are not caused by bodily states; therefore the first physically uncaused state of an individual must have a prior nonphysical cause. But the existence of God is not admitted; hence there must be an empirical conscious state prior to conception and birth. This argument applies indefinitely in a backward direction through previous births. (It may be noted that the argument is consistent with the Buddhist denial of an eternal soul, since the mental states of an organism are no more permanent than the physical ones.) (2) There is a Hindu argument from the eternity of the soul, which has been used in modern times by Radhakrishnan. Souls are eternal, but the normal condition for a soul is to be associated with a body. Hence it is likely that the soul in the past and future has a virtually everlasting succession of bodies. Thus metaphysical arguments attempting to establish the eternity of the soul have been taken to imply preexistence as well as postexistence.

Empirical arguments are as follows. (3) Children have instinctive capacities, which suggests that there must be learning prior to birth. Similarly, it is sometimes argued that child geniuses, such as Wolfgang Amadeus Mozart, indicate prenatal training. (4) Some people claim to remember past births, as in the case of Bridey Murphy. This claim is commonly made in the East for yogis and persons of deep spiritual insight, such as the Buddha and Buddhist saints. (5) The déjà vu experience and claims to knowledge of people and places that are not based on previous experience in this life have been cited as indicating rebirth. A counterargument is used against the objection that most people have no memories of such previous lives: Death is a traumatic experience (and so is birth), likely to cause amnesia. (6) The soul is indivisible and thus cannot derive from the parents, since it would then have to be a combination of parts.

The three important forms of theological argument are as follows. (7) Hindu and other scriptures and theologians are reliable in other matters and so ought to be reliable with respect to the teaching of rebirth. (8) Rebirth, associated with karma, provides a solution to part of the problem of evil, since inequalities and sufferings are the result of people's past deeds. (9) The doctrine of rebirth provides the possibility of a long process of self-perfection, which harmonizes well with the religious vision of the world as a theater for moral striving.

The following are the objections that have been or can be brought against the arguments for reincarnation. Three objections to argument (1) are, first, the concept of emergent characteristics obviates the difficulty in explaining the cause of psychical states, although perhaps at the expense of being obscure. Second, the first premise (that all states have prior causes) is arguable, and it might be that nonphysically caused mental states are simply not caused. Third, the existence of God cannot be ruled out. (2) The plausibility of the argument depends on the plausibility of arguments for the eternity of the soul. Further, in Indian religious thought there is the possibility of

mokṣa, or nirvāṇa, a state of liberation in which there is no more rebirth. Consequently, it is inconsistent to hold that embodiment is necessary to souls. The Buddhist denial of a permanent self occasioned the criticism that there is nothing carried over to another life that would ensure individual continuity—the reply being that, on the Buddhist analysis, the individual in his present life is only a series of events, so that there is no essential difference in considering a succession of lives as constituting an individual series.

The following are objections to the empirical arguments. (3) Modern biology can sketch alternative explanations of instinct and genius in children. (4) Although some people seem to remember past lives, the evidence is not so unambiguous as to be conclusive; and if saintliness is a condition for remembering previous births, it would be difficult to verify such a memory—it would be hard to conduct an "experiment" in becoming a saint. (5) Similar problems arise with the evidence of déjà vu experiences. As to whether death is a traumatic experience, there is no evidence. (6) The creation of souls by God is compatible with the argument concerning the indivisibility of the soul; but in any case the argument depends on a soul-body distinction that may not be acceptable.

The objections to theological arguments are the following. (7) The validity of particular scriptures and theologies on matters of detail is especially suspect. (8) The argument that rebirth explains the existence of evil could not by itself be conclusive, since the problem of evil exists only for those who believe in a good God. (9) A similar consideration applies to the argument that rebirth allows the possibility of self-perfection.

Although believers in rebirth have scarcely touched on the matter, the theory of evolution also presents considerable difficulties to the traditional doctrine of a virtually infinite series stretching back into the past. In Indian mythological cosmology, however, there are periodic destructions of the cosmos, and during these periods embodied souls continue to exist latently; no doubt a similar assumption may deal with the above biological difficulties by arguing that before the emergence of life, souls existed latently, or in other parts of the cosmos. The problem remains, however, that this account would not be easily, if at all, checked by empirical evidence.

The hypothesis of reincarnation presents interesting problems about personal identity. If personal identity is analyzed in terms of memory, there would seem to be only a vacuous distinction between saying that *A* is reborn as *B* and that *A* and *B* are separate persons. C. J. Ducasse, however, has argued (*A Critical Examination of the Belief in a Life after Death,* p. 225) that memory of any given life may be regained at some time or other in the series, and this would hold the series together. If bodily identity were held to be necessary to personal identity, rebirth could scarcely be meaningful, as it involves causal action at a distance in the transition from *A*'s death to *B*'s birth or conception.

See also Buddhism; Ducasse, Curt John; Empedocles; Evil, The Problem of; Gnosticism; Immortality; Indian Philosophy; Kabbalah; Karma; McTaggart, John McTaggart Ellis; Nirvāṇa; Orphism; Plato; Plotinus; Pythagoras and Pythagoreanism; Schopenhauer, Arthur.

Bibliography

Brandon, S. F. G. *Man and His Destiny in the World's Religions.* Manchester, U.K., 1962.

Dasgupta, S. N. *History of Indian Philosophy,* 5 vols. Cambridge, U.K.: Cambridge University Press, 1922–1955.

Davis, Stephen, ed. *Death and Afterlife.* New York: St. Martin's Press, 1989.

Ducasse, C. J. *A Critical Examination of the Belief in a Life after Death.* Springfield, IL: Charles C. Thomas, 1961.

Edwards, Paul. *Reincarnation: A Critical Examination.* Amherst, MA: Prometheus, 1996.

Hick, John. *Death and Eternal Life.* New York: Harper and Row, 1976.

Kear, Lynn. *Reincarnation: A Selected Annotated Bibliography.* Westport, CT: Greenwood Press, 1996.

McTaggart, J. M. E. *Some Dogmas of Religion.* London: Arnold, 1906.

O'Flaherty, Wendy, ed. *Karma and Rebirth in Classical Indian Traditions.* Berkeley: University of California Press, 1980.

Radhakrishnan, S. *The Brahma Sūtra.* London, 1959. Translation with commentary plus a lengthy introduction to Vedānta.

Reichenbach, Bruce. *The Law of Karma.* London: Macmillan, 1990.

Smart, N. *Doctrine and Argument in Indian Philosophy.* London: Allen and Unwin, 1964. Ch. 12.

Stevenson, Ian. *The Evidence for Survival from Claimed Memories of Former Incarnations.* New York, 1961.

Ninian Smart (1967)
Bibliography updated by Christian B. Miller (2005)

REINHOLD, KARL LEONHARD
(1758–1823)

Karl Leonhard Reinhold, the Austrian philosopher, was educated by Jesuits until the dissolution of their order in 1773, when he entered the Catholic college of the Barn-

abites, where he also taught, from 1778 to 1783. In 1783 Reinhold left Vienna for Leipzig and in the same year abandoned Catholicism in favor of Protestantism. A year later he moved to Weimar, where he was invited by Christoph Martin Wieland to contribute to his *Teutscher Merkur*. Soon he was not only Wieland's closest friend but also his son-in-law. Reinhold's first article, "Gedanken über Aufklärung," in which he traced the emergence of Enlightenment thought, appeared in July 1784, just a few months before the publication of Immanuel Kant's famous essay "What Is Enlightenment?" In his article Reinhold pleaded for the fuller realization of such. Enlightenment aims as greater tolerance toward religious minorities, more widespread secularization of knowledge and its greater accessibility to all sections of the population, and, above all, for the right of the individual to seek and assert truth free from fear, according to his critical reason and moral convictions.

Although two years later (1786) he was to publish a series of articles in support of Kant's critical philosophy, his second article in the *Merkur* (1785) was directed against Kant's unfavorable review of Johann Gottfried Herder's *Ideen*. The article appeared anonymously, but Reinhold later admitted his authorship to Kant. The articles dealing with Kant's *Critique of Pure Reason*, published under the title "Briefe über die Kantische Philosophie" from 1786 to 1787, established Reinhold's reputation as the most skillful exponent of Kant's philosophy and resulted in his being offered the chair of philosophy at the University of Jena in 1787. Reinhold was no less successful as a university teacher, and soon after his arrival Jena became one of the chief centers of Kantian studies. He attracted many students to Jena, and so great was his popularity that he was repeatedly urged to refuse the appointment offered him at the University of Kiel. Reinhold hesitated at first but eventually decided to move to Kiel in 1794, where he remained until his death.

One of the reasons for his departure, perhaps the most decisive, is revealed in a letter to Wieland that Reinhold later published in a selection of essays (*Auswahl vermischter Schriften*, Jena, 1796), under the title "Ueber die teutschen Beurtheilungen der französischen Revolution." Reinhold became increasingly worried over his countrymen's reactions to the excesses of the French revolutionary tribunals. In Kiel, which was then under Danish rule, he hoped to find a calmer political climate. Without condoning the terror of the revolutionaries, he nevertheless deplored the inferences that were drawn from it by leading public figures in Germany. In particular he viewed with anxiety the introduction of repressive measures and the tendency to regard the French Revolution as a conspiracy of the philosophers. The French revolutionaries, he argued, may have been mistaken in attempting to deduce political rules from abstract principles that were often inadequately understood, but they were correct in their assessment of the desperate plight of their compatriots. If inferences were to be drawn, these would not suggest that philosophy presented a danger to orderly government but rather that disorderly government encouraged men to invoke philosophy in a manner unwarranted by its inherent limitations. Practical considerations such as these, no less than more strictly theoretical ones, prompted Reinhold to inquire more closely into the nature and scope of philosophical speculation.

Most of the works that he wrote at Kiel advanced a "fundamental philosophy" concerned with the basic presuppositions of scientifically valid thought. As the basic axiom of his "fundamental philosophy" Reinhold postulated the principle of consciousness, which he formulated in this way: By virtue of consciousness the perceiving (*erkennende*) subject is capable of distinguishing himself as something distinct from, while at the same time related to, the object of his consciousness, which, however, is not the object itself but rather the idea or notion (*Vorstellung*) of it. The consciousness itself constitutes a basic and irreducible fact, capable of neither proof nor further definition. It can only verify itself by reflecting upon itself. Reinhold was anxious to demonstrate that every thought process involves both a priori and a posteriori elements. The relation of the *Vorstellung* to the external object embodies its a posteriori material content (*Stoff*), whereas the subjective activity involved (*Vorstellungsvermögen*) in shaping the material content into a clear *Vorstellung* constitutes its a priori form (*Form*).

Reinhold stipulated three interconnected stages in the operation of consciousness: sense perception (*Anschauung*), which he classified as a receptive activity, and cognitive understanding (*Verstand*) and reflective reasoning (*Vernunft*), both of which he described as spontaneous activities. The product of these combined activities is the *Vorstellung*, which, Reinhold warned, must not be confused with an "image" or an "impression," for both terms suggest mere receptivity. Nor must it be identified with a "representation" of the object, since there is no way of either proving the identity of the *Vorstellung* with the object or even of comparing its similarity to the object. It follows that the object as such, no less than the subject as such, remains not only unknowable (as Kant realized) but also inconceivable. Both subject and object, therefore, as things-in-themselves are

pure abstractions. They are the residue of a *Vorstellung*, the thing minus the notion or conception of it.

Without denying the existence of things-in-themselves, Reinhold refused to commit himself as to the nature of their existence. He explicitly stated that he was merely anxious to determine the possibility and the limitations of cognition, not to inquire into its psychological origins or into the ontological nature of the objects of cognition. His declared aim was to provide a descriptive account, a phenomenology, rather than a theory of cognition, together with an analysis of the terminology commonly employed in this field. In spite of, or perhaps because of, Reinhold's deliberate delimitation of his theoretical undertaking, his works provided suggestively fertile starting points for subsequent Kantian research from Johann Gottlieb Fichte to Arthur Schopenhauer.

See also A Priori and A Posteriori; Consciousness; Enlightenment; Fichte, Johann Gottlieb; Herder, Johann Gottfried; Kant, Immanuel; Schopenhauer, Arthur; Toleration.

Bibliography

ADDITIONAL WORKS BY REINHOLD

Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens. Prague and Jena, Germany, 1789.

Beyträge zur Berichtigung bisheriger Missverständnisse der Philosophen. 2 vols. Jena, Germany, 1790 and 1794.

Ueber das Fundament des philosophischen Wissens. Jena, Germany, 1791.

Versuch einer Kritik der Logik aus dem Gesichtspunkt der Sprache. Kiel, 1806.

Grundlegung einer Synonymik für den allgemeinen Sprachgebrauch in den philosophischen Wissenschaften. Kiel, 1812.

Das menschliche Erkenntnisvermögen aus dem Gesichtspunkt der durch die Wortsprache vermittelten Zusammenhänge zwischen Sinnlichkeit und dem Denkvermögen. Kiel, 1816.

WORKS ON REINHOLD

Adam, Herbert. *Carl Leonhard Reinholds philosophischer Systemwechsel*. Heidelberg: Winter, 1930.

Bondeli, Martin, and Wolfgang H. Schrader, eds. *Die Philosophie Karl Leonhard Reinholds*. New York: Rodopi, 2003. In English and German.

Cassirer, Ernst. *Das Erkenntnisproblem*. Vol. III, 33–57. Berlin: Cassirer, 1923.

Klemmt, Alfred. *Karl Leonhard Reinholds Elementarphilosophie*. Hamburg: Meiner, 1958.

Kroner, Richard. *Von Kant bis Hegel*. Vol. I, 315–326. Tübingen, 1921.

Reinhold, Ernst. *Karl Leonhard Reinholds Leben und literarisches Wirken*. Jena, Germany: Frommann, 1825.

Roehr, Sabine. "Freedom and Autonomy in Schiller." *Journal of the History of Ideas* 64 (1) (2003): 119–134.

Selling, Magnus. *Karl Leonhard Reinholds Elementarphilosophie in ihrem philosophischen Zusammenhang*. Lund, 1938.

Solomon, Robert C. *Introducing the German Idealists: Mock Interviews with Kant, Hegel, Fichte, Schelling, Reinhold, Jacobi, Schlegel, and a Letter from Schopenhauer*. Indianapolis: Hackett, 1981.

Von Schonborn, Alexander. "Karl Leonhard Reinhold: Endeavoring to Keep up the Pace Mit Unserem Zeitalter." In *The Emergence of German Idealism: Studies in Philosophy and the History of Philosophy*. Vol. 34, edited by Michael Baur. Washington, DC: Catholic University American Press, 1999.

Frederick M. Barnard (1967)
Bibliography updated by Tamra Frei (2005)

RELATIONS, INTERNAL AND EXTERNAL

Common sense would seem to hold that if some properties of a thing were taken away from it, it would no longer be the same thing. Further, it seems to hold that this is not the case for all properties of the thing. This intuition is the basis of the distinction between essential and accidental properties of a thing. It is also the basis of the distinction between the internal and the external relations that that thing bears to other things. For if among the properties that are essential to a thing (for example, the state of Maine) are relational properties, properties whose characterization essentially involves reference to some other thing (for example, the property of being north of Boston), then we say that the relations in question (for example, the relation between Maine and Boston) are *internal* to that thing (Maine). If we think that the thing would be the same were it (for example) not north of Boston—as in the case of a railroad car traveling through Maine—then we say that the relation in question is merely *external* to that thing.

The most familiar sort of relations considered when the topic of internal relations is discussed are relations between two or more particulars. However, the same internal–external distinction may be drawn in the case of relations between universals and particulars and also in the case of relations between two or more universals. If one holds that for every property *P* that a particular *X* displays, there is a universal, *P*-hood, to which *X* stands in the relation of "exemplification," then *all* of *X*'s properties may be construed as relational properties. Some of these relations of exemplification may be regarded as internal to *X* and others as external. Again, one may say that a universal such as "manhood" stands in an internal relation to certain other universals (for example, "rationality") and

in an external relation to other universals (for example, "philosopherhood"). Here the internal relation in question will be entailment, in the sense of "entails" in which we say that a given property ("being a man") entails another property ("being rational"). In what follows, however, we shall confine ourselves as far as possible to relations holding between particulars, both because the philosophical literature has focused on such relations and because the notions of "exemplification of universals" and of "relations of entailment holding between universals" are sufficiently obscure and controversial to require detailed supplementary discussions. (Also, we shall not always trouble to distinguish between discussion of internal *properties* and of internal *relations,* since whatever doctrine a philosopher holds about the former will apply, mutatis mutandis, to the latter).

Two extreme positions have been put forward by philosophers who regard the internal-external distinction as unclear or incoherent. The first is that all of a thing's properties are essential to its being what it is (and, a fortiori, that all its relations are internal to it). This position is associated with idealism and monism, for reasons that will emerge as we proceed. It holds that the connections between each of a thing's properties (including its relational properties) and all of its other properties are so close that the deprivation of a single property would force us to say that, in a nontrivial sense, the thing is no longer what it was.

The second extreme position holds that none of a thing's properties are essential to it (and thus, a fortiori, that no relations are internal to it). This view is put forward by those who make a firm distinction between the thing itself and a description of it. These philosophers say that, although certain properties of the thing are such that a given description could no longer be correctly applied to it were these properties absent, the notion that "the thing would no longer be the same" if these properties were absent is either trivial or misleading. For, in the weakest sense of "same," the absence of any of its properties would make the thing no longer the same. Any stronger sense will, however, equate "being the same thing" with "being such that a given description correctly applies to it." But since for each thing there are an infinity of equally correct descriptions, and nothing in the thing itself determines which of these is the description, any specification of "essential properties" will be arbitrary.

Both positions hold that the traditional essence-accident distinction, which is drawn by common sense and was first formulated explicitly by Aristotle, must be abandoned. The second position holds that the notion of "essential property" must be seen as a purely conventional notion, without a ground in the nature of the thing itself. It therefore suggests that we replace the notion of a relation being internal to a thing with the notion of a given relational description of a thing (such as "being north of Boston") being internal to (that is, a necessary condition of) another description of the thing (such as "being in Maine"). The first position holds that the notion of "essential property" suggests, wrongly, that there is such a thing as a nonessential property. But since omniscience would see the universe as a seamless web (and, perhaps, as one single individual thing—the Absolute), this suggestion is misleading. Granted, they may say to representatives of the second position, that our present notion of "essential property" is a merely conventional one, we should not be led to conclude that things have no intrinsic natures. They do have intrinsic natures, but these can be known only sub specie aeternitatis, as facets of the Absolute. The commonsense essence-accident distinction is natural and inevitable, given the imperfect state of our knowledge. For omniscience, however, this distinction would be pointless.

This brief sketch of the opposing positions suffices to suggest how intimately the issues about internal relations are bound up with a whole range of other philosophical problems—problems about the notions of substance, of essence, and of "bare particulars," about "real" versus "nominal" definitions, about nominalism versus realism, about the way in which we refer to and identify particulars, and about the nature of necessary truth. It is perhaps not too much to say that a philosopher's views on internal relations are themselves internally related to all his other philosophical views.

THE VIEW THAT ALL RELATIONS ARE INTERNAL

The view that all relations are internal, in the form in which it has been discussed in the twentieth century, originated in the writings of the absolute idealist school in England and America in the period 1890–1920. In various forms it was held by F. H. Bradley, Josiah Royce, Bernard Bosanquet, and many others. Its most recent sustained defense is found in the work of Brand Blanshard, a follower of Bradley, notably in *The Nature of Thought* (1939). It has obvious historical connections with the doctrines of the seventeenth-century rationalists, notably Gottfried Wilhelm Leibniz's view that all truths are analytic and Benedict de Spinoza's assimilation of causal relations to logical relations. Its most important historical

antecedent, however, is the philosophy of G. W. F. Hegel. Hegel's insistence that the world was rational through and through, because Reason (or "Spirit") alone was real, was the principal inspiration of the philosophers who adopted the view that all relations were internal. For, if some relations were external, then the universe would be "impenetrable" to reason, in the sense that there would be brute particular facts not deducible from universal truths even by God himself.

A. C. Ewing, in *Idealism* (1934), provides a comprehensive account of the various meanings given to the term *internal* by exponents of absolute idealism and a critical analysis of those arguments in favor of the doctrine that all relations are internal that depend upon an ambiguous use of "internal." As Ewing points out, the meanings given to "internal" ranged from a very weak sense, in which to say that the relation *R* which *X* bore to *Y* was internal to *X* meant merely that "*R* makes a real difference to *X*," to a very strong sense, in which it meant that "from a knowledge of *Y* and *R* we could infer with logical necessity that *X* possesses a certain determinate or relatively determinate characteristic other than the characteristic of standing in the relation in question." Because such ambiguities permeate the discussion of the topic in such writers as Bradley and Royce, we shall not attempt an exegesis of their arguments. Instead, we shall attempt a reconstruction of two particularly persuasive arguments that seem to represent at least part of the common core of the absolute idealists' defense of their position on this subject. The two arguments to be examined by no means exhaust the repertoire of arguments that have been deployed in favor of the view that all relations are internal, but they are the arguments on which criticism of this view has chiefly centered.

ARGUMENT FROM THE NATURE OF SELF-IDENTITY. The first argument, which will be called here the argument from the nature of self-identity, was first clearly formulated by a critic rather than a proponent of the view that all relations are internal. G. E. Moore, in a classic attack on this view ("External and Internal Relations"), suggests that "one thing which is always implied by the dogma that 'All relations are internal' is that, in the case of every relational property, it can always be truly asserted of any term *A* which has that property, that any term which had not had it would necessarily have been different from *A*." The argument in favor of this view is simply that, as Moore puts it, "if *A* has *P*, and *x* has not, it *does* follow that *x* is other than *A*." In other words, it is unquestionably true that

(1) *A* has *P* entails that (*x* does not have *P* materially implies that *x* is other than *A*).

Contemplation of this truth, Moore suggested, led philosophers to say that "*A* could not be what it is (but would necessarily be something different) did it not have *P*."

Now, as Moore points out, the argument as it stands is fallacious. (1) does not permit the conclusion that

(2) *A* has *P* materially implies that (*x* does not have *P* entails that *x* is other than *A*).

Only (2) would permit the conclusion that *A* would necessarily be a different particular did it not have *P*. The difference between (1) and (2) may be put by saying that all that (1) tells us is that *A* cannot both have and not have the property *P*, whereas (2) tells us that *A* could not be *A* unless it had *P*. (1) is trivial, whereas (2) blurs the commonsense contrast between essential and accidental properties (and thus between internal and external relations). As Moore puts it, "(1) asserts that if *A* has *P*, then any term which has not, *must* be other than *A*; (2) asserts that if *A* has *P*, then any term which had not, would necessarily be other than *A*." Moore notes that to confuse the two propositions, "you have only to confuse 'must' or 'is necessarily' with 'would necessarily be.'" This confusion, in turn, will lead one to confuse the (physically necessary but logically contingent) fact that *A* has *P* with a statement about what is *logically* necessary for something to be *A*. While not attempting to cite examples of this fallacy in the writings of the absolute idealists, Moore claimed that much of their willingness to adopt the view that all relations are internal was due to their having confused (1) and (2). Whether or not this fallacy played the role in their thought that Moore thought it did is less important, from a historical point of view, than the influence exercised by Moore's diagnosis. Philosophers in general tended to agree with Moore that the absolute idealists had been guilty of this confusion, and his essay was a turning point in discussion of the topic. Defenders of the thesis of the internality of all relations who came after Moore were forced to produce arguments against the main presupposition of Moore's argument—that the commonsense distinction between logically contingent propositions and logically necessary propositions was unobjectionable. Crudely put, one may say that before Moore's essay, defenders of the view that all relations were internal felt able to argue that simple reflection on commonsense criteria for self-identity led to the conclusion they desired. After Moore's essay, they were forced to attempt to undermine common sense by claiming that the distinctions

Moore had drawn were, though commonsensical, philosophically indefensible.

ARGUMENT FROM THE NATURE OF CAUSALITY. The above was the strategy adopted by Blanshard in his *The Nature of Thought,* in which he presents the second, far more important and profound, argument in favor of the doctrine that all relations are internal. This may be called the argument from the nature of causality. Moore, like most philosophers in the tradition of British empiricism, had taken for granted a distinction between physical necessity and logical necessity, a distinction between the sense in which it is necessary, given the laws of nature and the past history of the universe, that a given particle be located at a given point in space at a given time, and the sense in which it is not necessary, *simpliciter.* Traditional rationalism, on the other hand, had questioned this distinction. Although earlier absolute idealists had also rejected the distinction between two kinds of necessity, they had done so *en passant.* They had treated it as simply one more consequence of empiricism's uncritical acceptance of a commonsense metaphysics that, they claimed to have shown, was fundamentally incoherent. Blanshard, approaching the matter epistemologically rather than metaphysically, brought forward a battery of arguments designed to show that the acceptance of this distinction was the result of a mistaken Humean analysis of knowledge. By weakening this distinction and claiming that causal necessity (by virtue of which *A* had *P*) could not be separated from logical necessity (by virtue of which *A* was self-identical), he was able to argue that what Moore had viewed as a simple confusion was at worst a confused formulation of a vitally important insight.

In examining this second argument, it will again be convenient to look to its critics rather than to its defenders. Ernest Nagel, in a critique of Blanshard's *The Nature of Thought* titled "Sovereign Reason," restates and criticizes Blanshard's views on internal relations in a way that brings out very clearly their connection with Blanshard's treatment of causality. Blanshard, in turn, has replied to Nagel in the later chapters (particularly Ch. 12) of his *Reason and Analysis* (1963). A summary of the Blanshard-Nagel controversy will serve two purposes. It will trace the most recent line of defense adopted by defenders of the view that all relations are internal, and it will lead us to an understanding of why some philosophers claim that no relations are internal.

Blanshard puts forward, and Nagel quotes as a basis for criticism, the following version of the doctrine that all relations are internal. Despite the ambiguities detected by Ewing, Blanshard holds that "the principal meaning" of this doctrine is clear and formulates it as follows:

> (1) that every term, i.e., every possible object of thought, is what it is in virtue of relations to what is other than itself; (2) that its nature is affected thus not by some of its relations only, but in differing degrees by all of them, no matter how external they may seem; (3) that in consequence of (2) and of the further obvious fact that everything is related in *some* way to everything else, no knowledge will reveal completely the nature of any term until it has exhausted that term's relations to everything else. (*Nature of Thought,* Vol. II, p. 452)

Nagel notes, and Blanshard would agree, that everything here turns on the notion of the "nature of a term." If the term's nature includes all its properties, then Blanshard is right. Nagel bases his general objections to Blanshard on the claim that this is a perverse use of "nature," since "it is quite clear that just what characters are included in an individual, and just where the boundaries of an individual are drawn, depend on decisions as to the use of language. These decisions, though motivated by considerations of practical utility, are *logically arbitrary*" (p. 275). Nagel, in other words, is saying that "the nature of *X*" consists of just those properties of *X* whose absence would cause us to cease using "*X*" to refer to *X* and that the selection of these properties is determined not by empirical study but by convention. The list of such properties is finite, whereas the list of the properties of *X* is potentially infinite. Nagel thus adopts what has become the standard empiricist view, first clearly formulated by A. J. Ayer in "Internal Relations," that to determine which properties of *X* are internal to it is merely a matter of determining which propositions about *X* are analytic and that determining *this* is simply a matter of consulting linguistic usage. To urge that the nature of a thing includes all its properties would, given this view, be to urge that *all* propositions about *X* are analytic. Both Nagel and Ayer treat this conclusion as a reductio ad absurdum.

In examining Blanshard's arguments, Nagel first takes up Blanshard's form of the argument from the nature of self-identity and disposes of it by drawing what is essentially Moore's distinction between the logically contingent fact that *A* has *P* and the logically necessary fact that anything that does not have *P* cannot be identical with *A*. His defense of this distinction is simply that unless the distinction is drawn, we shall wind up with the view that "the nature of *X*" is identical with *X* itself and

thus that "the nature of a thing, like the thing itself, would be something that is in principle indefinable and could not therefore be made the basis for bringing into systematic order any of the characters which the thing displays" (p. 276). But from Blanshard's point of view, this reply begs the question, since Blanshard would be quite willing to say that the nature of any given particular is indeed indefinable (by finite minds). For Blanshard the question is merely pushed back to the issue of whether a satisfactory epistemology can be constructed on the basis of the view that all logical necessity has its source in linguistic convention. But this latter issue is just the issue of whether causal relationships (which are agreed on all sides to be matters not of convention but of empirical inquiry) can, in the last analysis, be held to be distinct from logical relationships. If they cannot, then it would seem fair to say that although we must (unfortunately) work with the commonsense distinctions between necessary and contingent truths, essence and accident, physical and logical necessity, and the like, these distinctions are nevertheless mere pragmatic makeshifts (pertaining, in Bradleian terminology, to Appearance rather than to Reality). To invoke them to is to attend not to how things are but merely to how we are forced (by the limitations of our minds and of our everyday language) to talk about them.

Thus the battle between Blanshard and Nagel is truly joined only when Nagel takes up the question whether "logical necessity is involved in causal relations." Blanshard has, as Nagel notes, two principal arguments for the view that it is so involved. The first is that causal relations must be analyzed either in terms of "mere regularity of sequence" or in terms of "entailment." The failure of the regularity view will, in Blanshard's eyes, constitute a proof of the entailment view. But the entailment view is just that "*A* causes *B*" is a statement about a logical relation between *A* and *B.* Now if (as is not implausible) all true relational propositions about particulars are propositions that are true in virtue of causal relations between the particulars mentioned in these propositions, then it follows that all particulars are connected to all others by logical relations and that every such proposition would be seen (by omniscience) to entail a logical truth about every such particular.

Nagel has two objections to this argument. First, the "regularity" and "entailment" views do not exhaust the available analyses of causality; second, "the entailment view contributes *nothing* toward advancing the aims of specific inquiries into the causal dependencies of physical nature." The second objection can be dismissed by Blanshard as irrelevant, since he is quite willing to admit, with David Hume, that observation of regular sequence is our only method for determining what causal relations actually hold (except, perhaps, in the case of "direct insight" into certain relations between mental states or events). Blanshard need merely insist that regularity provides evidence of an underlying entailment but that the regularity and the entailment must not be confused. Blanshard offers no reply to Nagel's first objection, but one suspects that he would argue that all proposed via media analyses of causality in fact boil down to one of the two alternatives he has suggested. Even if this point is granted to Blanshard, however, the whole question of the validity of his attack on the regularity theory remains. We must leave the topic with the remark that Blanshard can, in attacking this theory, take full advantage of the embarrassment encountered by Rudolf Carnap, Nelson Goodman, and others in their attempts to construct an inductive logic on the basis of Neo-Humean "regularities." Further, recent work in inductive logic (such as Goodman's *Fact, Fiction and Forecast,* 1955) and the philosophy of science (the work of Hilary Putnam, Wilfrid Sellars, P. K. Feyerabend, and others) has made it apparent that the distinction between matters of convention and matters of fact is not so clear as Hume and the early positivists believed. This recent work is closely connected with W. V. Quine's skepticism about the analytic-synthetic distinction and related work in the philosophy of language. It is perhaps not too much to say that empiricism is presently in a state of crisis and that the crisis revolves precisely around the validity of the distinctions that empiricists have traditionally invoked against the thesis of the internality of all relations. We must conclude that the question of the validity of Blanshard's first form of the argument from the nature of causality must remain undecided until these issues have been further clarified.

Before leaving the Blanshard-Nagel controversy, however, we must take up the second of Blanshard's arguments in favor of the view that logical necessity is involved in causation. This argument is that philosophical reflection upon the nature of causality leads us to conclude that

> to say that *a* produces *x* in virtue of being *a* and yet that, given *a, x* might not follow, is inconsistent with the laws of identity and contradiction. Of course if *a* were a cluster of qualities abstracted from their relations, and its modes of causal behaviour were another set conjoined with the former externally, then one could deny the latter and retain the former with perfect con-

> sistency. But we have seen that when we say *a* causes *x* we do *not* mean that sort of conjunction; we mean an intrinsic relation, i.e., a relation in which *a*'s behaviour is the outgrowth or expression of *a*'s nature. And to assert that *a*'s behaviour, so conceived, could be different while *a* was still the same would be to assert that something both did and did not issue from the nature of *a.* (*Nature of Thought,* Vol. II, p. 513)

With this argument, as Nagel notes, we are back at the perplexing notion of "the nature of *a.*" Whereas the entailment analysis of the nature of causation can perhaps be stated without using the notion of the "nature of *A*" (although if it were, it might be difficult for Blanshard to infer the thesis of the internality of all relations from the truth of the entailment view), this present argument about the nature of causality makes essential use of this notion. At this point, therefore, Nagel returns to his general line of attack on Blanshard's formulation of the thesis of the internality of all relations and argues that what Blanshard says here is true only if "the nature of *X*" is defined as "all the properties of *X,*" a definition that, in Nagel's eyes, is both idiosyncratic and such as to trivialize Blanshard's claim.

The effectiveness of Nagel's reply can be judged only in the light of a general theory about the relation between thought, language, and reality. For, here again, Nagel is taking for granted the view that whether a given property is included within a thing's nature is a question about our language, rather than a question to be settled by further inquiry about the thing itself. Just as judgment of the validity of the first form of Blanshard's argument from the nature of causality must be postponed until certain general philosophical issues have been (at least) clarified, so also judgment of the validity of the second form of this argument must be deferred until questions about the standard empiricist doctrine that all "essences" are "nominal" and that "real essence" is an incoherent notion are settled. For Blanshard can insist that Nagel has begged these latter questions. In *Reason and Analysis* we find Blanshard arguing that Nagel's view that decisions about what characters are included in an individual are "logically arbitrary" leads to the view that, for example, Socrates's snub-nosedness is as good a candidate for an essential property of Socrates as his philosopherhood. Blanshard thinks this a reductio ad absurdum, but this rebuttal, once again, merely moves the argument one step further back. Nagel's point is not that we arbitrarily select which characteristics of an individual shall count as essential but that the criteria of selection are pragmatic, dictated by our present interests and the modes of classification that we have, in the past, found it convenient to adopt. Nagel would say that a choice about linguistic usage, which is, from a practical point of view, far from arbitrary, is nonetheless logically arbitrary, in the sense that a language with alternative conventions is, though inconvenient, perfectly possible.

Blanshard's basic disagreement with Nagel consists in his view that such pragmatic considerations are not the last word and his insistence that the goal of thought is the discovery of real essences. Such real essences would be discovered by discovering the chains of entailment that connect all the various universals that characterize (and, in Blanshard's metaphysics, constitute) a particular. In Blanshard's view, to say that analytic propositions are true by convention is thoroughly misleading, for such conventions are the results of attempts to discover such entailments. For Blanshard the identification of the nature of *X* with *X* itself, and of both with the totality of properties that characterize *X,* and of all of these with *X*-as-known-by-an-ideal-knower (one who could grasp the entailments between all of these properties), is not (as it is for Nagel) a series of confusions but is forced upon us by an analysis of what we mean by "knowing *X.*" The validity of Blanshard's second form of the argument from the nature of causality ultimately depends upon the validity of this analysis.

UNIVERSALS. The nature and depth of the issues involved in the controversy between Blanshard and Nagel may be further clarified by calling attention to one more area of disagreement between them. This concerns the nature and knowledge of universals. Blanshard views a particular as a congeries of universals and views the internal relations between particulars as reflecting the internal relations holding between the universals that constitute them. It is almost a cliché of recent analytic philosophy that to have knowledge of a universal is simply to know the meaning of a word; thus, to be acquainted with all the universals that characterize a particular would be merely to know the meanings of all the words correctly applicable to that particular. Such knowledge would obviously fall far short of telling us about the relations in which that particular stands to other particulars. For Blanshard, however, universals have natures that are not known to those who merely know the meanings of the words that signify those universals. To know the nature of a universal "fully and as it really is" would involve knowing its relations to all the universals that are exemplified in all the particulars that exemplify the first universal.

Thus, to know any universal "fully and as it really is" would be possible only for omniscience, just as, and for the same reasons that, knowledge of the real essence of a particular would be possible only for omniscience. Thus, resolution of the controversy about internal relations would require, at a minimum, a decision concerning the adequacy of a nominalistic account of universals. Blanshard views the current antagonism toward idealism (and a fortiori toward the thesis of the internality of all relations) as largely a result of analytic philosophy's "systematic confusion between thought and language," a confusion that leads philosophers such as Ludwig Wittgenstein to hold (1) that the notion of having a concept or being acquainted with a universal prior to the use of language is incoherent, and (2) that the notion of detecting internal relations between universals apart from considerations of linguistic usage is a relic of a radically mistaken analysis of mental events. If these latter tenets are accepted, clearly Blanshard's arguments cannot even get off the ground. Once again, we must conclude that the thesis of the internality of all relations cannot be profitably discussed until one has taken sides on the most fundamental issues in contemporary philosophy.

THE VIEW THAT NO RELATIONS ARE INTERNAL

When we turn to the view that no relations are internal, we turn from a controversy that reflects profound underlying disagreements concerning the analysis of knowledge to a controversy about much narrower issues concerning the analysis of naming and predication. Those who say that no particular is internally related to any other particular insist that the only entities that can be internally related to one another are characteristics of particulars. Following to its logical conclusion Nagel's claim that the assignment of a given description to a given particular is "logically arbitrary," they hold that to say that *X* would "not be what it is" unless it had *P* is merely to say that the particular could not be characterized in a given way unless it had this property. But since the particular is sublimely indifferent to how it is characterized, it "is what it is" regardless of whatever properties it may have. To speak of "logically necessary conditions for the self-identity of *X*" is, at best, to speak elliptically of "logically necessary conditions for correctly describing *X* as a *K*," where "*K*" signifies some kind of thing of which *X* is a representative, or (more generally) of "logically necessary conditions for correctly describing *X* as *C*," where "*C*" is some general characterization.

The whole notion of "properties (and, a fortiori, relations) such that *X* would cease to be what it is if they were removed" is thus either incoherent or misleading. For "being what it is" is simply too ambiguous a notion; there are indefinitely many kinds to which *X* belongs and indefinitely many characterizations that apply to it. "Being what it is" is incoherent if it suggests that one of these kinds or characterizations is *intrinsically* privileged and misleading if a user of the phrase has already picked out some such kind or characterization, thus making his choice "privileged" by stipulation. To philosophers who deny the internality of any relations, the whole notion of internal properties and relations is an unfortunate vestige of the Aristotelian notion that there are real essences of particulars to be discovered by empirical inquiry. These philosophers heartily agree with the seventeenth-century rationalists, and with Blanshard, that any Aristotelian attempt to divide intrinsically essential and intrinsically accidental properties is foolish. But whereas Blanshard, sticking to the quest for real essences, insists that this point merely shows that the real essence of an object must include *all* its properties, these philosophers take the point to show the incoherence of the notion of "real essence" and the notion of "internal property."

It may be useful to put the contrast between the roughly Aristotelian commonsense view and the two extreme views in yet another way. If we say that common sense holds that there are both particulars and properties of particulars, then we may say that common sense holds that each particular stands in a necessary relation to some of its properties and in a contingent relation to others. Blanshard dissolves the particular into a congeries of properties, and, because he believes (*a*) that properties (qua universals) have intrinsic natures to be discovered by inquiry (other than inquiry into linguistic usage) and (*b*) that such inquiry would, in principle, discover relations of entailment between all possible properties of all possible particulars, he holds that a particular stands in a necessary relation to all its properties. Philosophers who deny both doctrines and who assert (*c*) that "logical necessity" can only characterize relationships between universals, naturally emerge with the conclusion that the whole notion of logically necessary relations between particulars and their properties must be discarded. To put it picturesquely, Blanshard thinks that the dissolution of the traditional essence-accident distinction leaves us with the particular as a node in a network of internal relations between universals. His opponents think that this dissolution leaves us with "bare" particulars on the one hand (particulars that could logically have any properties) and with a network of entailments between universals on the

other (a network that is, however, much "looser" than Blanshard's, since between most universals no relations of entailment exist).

CONCEPT OF BARE PARTICULARS. As an illustration of the movement toward leaving particulars bare, we may cite Gilbert Ryle, who says, in his article "Internal Relations," that

> for this view [the thesis of the internality of all relations] to be true *or false,* it would have to be significant to predicate a logically proper name or designation of a logically proper name or designation; and it would have to be significant to assert or deny that *this* was *this*; and the question "is anything this?" would have to mean something.… "This" is not a predicate, and a sentence in which it pretends to function as one is meaningless. So there *could* be no such dispute as to whether this's being this does or does not depend on its being in one or other of its relations. (p. 165)

This line of thought suggests the general conclusion that there are no analytic propositions that ascribe properties to particulars. For example, it is misleading to call "Socrates was a Greek philosopher" analytic, for what this statement expresses is either (1) the contingent fact that certain features (snub-nosedness, being married to Xanthippe, and so on) were compresent with certain others (being Greek, being a philosopher), or (2) the contingent fact that the word *Socrates* is used to refer to an individual who exhibited certain features.

Even among philosophers who both reject (*a*) and (*b*) and accept (*c*), however, this general conclusion has been a matter of debate. In what follows, we shall consider an attempt to avoid the conclusion that there can be no analytic propositions that ascribe properties to particulars and an attempt to avoid the extreme position that no relations are internal to particulars by providing a "rational reconstruction" of the commonsense view. Such attempts are motivated, at least in part, by philosophical discomfort over the notion of "bare particulars." The nature of this discomfort may be illustrated by considering the question "What, then, *are* these particulars, apart from the properties we ascribe to them?" If particulars really are "bare," then any answer to this question is bound to be either wrong (if it lists some features that are criteria for particularity) or unhelpful (if it consists in saying simply "Well, particulars are just the kind of thing that properties can be ascribed to"). Although the realistic bent of contemporary analytic philosophy makes philosophers hesitate to accept the Bradleian-Blanshardian view that the whole category of (plural) "particulars" belongs to Appearance rather than to Reality, it nevertheless seems that having only bare particulars would be as bad as having no particulars at all.

INTERNAL PROPERTIES AS RELATIVE. The most explicit and comprehensive attempt to avoid Ryle's conclusion and still retain most of his premises is found in an article by Timothy Sprigge ("Internal and External Properties"); an examination of Sprigge's treatment of the problem will bring out the underlying issues concerning naming and predication upon whose resolution the present question depends. Sprigge notes that the strength of the Rylean position lies in the fact that

> in sentences expressing particular propositions where the subject word is a name, the subject word has no connotation. Therefore no predicate word can have a connotation which is incompatible with the connotation of the subject word. But a subject-predicate sentence could only express a necessary proposition if the connotation of the subject word were incompatible with the connotation of the negation of the predicate word. … Of course, this rests upon the questionable view that there may be naming words without connotation—and this indeed is basically the point at issue. (p. 204)

One reason why this latter point is disputable is, as Sprigge says, that "it seems that one must identify a thing by some description. Having been thus identified," he continues, "as answering to that description, is it not in effect defined as the thing having those properties, which properties therefore it necessarily has?" (p. 205). In other words, proper names could not be used unless their users could identify their referents, and how could the users do this save by having a description in mind? Must we not say that the notion that the logician's dogma that "proper names do not connote" is true only of such Russellian "logically proper names" as "this" (which cannot be used save in the presence of their referents)? Sprigge replies to this point by granting it but noting also that since the same particular can be identified by an indefinitely wide range of different descriptions, the point is useless if one is trying to defend the notion of internal properties. In the case of a predicate, rough agreement on criteria for its application is required if the term is to play a useful role. But there seems nothing to prevent every speaker of the language from having a different set of procedures for identifying a particular while nevertheless using the same

proper name for it. Too many connotations are, so to speak, as bad as no connotation at all for purposes of formulating necessary truths.

If we follow Sprigge here, we need not be troubled by the spectacle of bare particulars. Every particular we refer to will always be dressed in some description or other, so we need not worry about how they look when undressed. But since each particular can be dressed up in so many ways, we are as far as ever from understanding what an "internal property" might be, unless we relativize the notion and say that certain properties are internal to *X* relative to a person *S* whose personal criteria for identifying *X* include the presence of these properties. Relativizing the notion in this way is, in essence, the basis for Sprigge's "reconstruction" of the notion of internal property. As a sample of the sort of intuition upon which the commonsense distinction between internal and external properties is based, he notes that even though we are driven by the Rylean reasoning outlined above to call all subject-predicate statements about particulars synthetic, we find it hard to imagine the falsehood of, for example, "Scott was, at some time in his life, a man." But what is a synthetic proposition if not one whose falsehood can be imagined? Sprigge proposes that we simply face up to the fact that there is a class of propositions that, if we *must* choose between calling them synthetic or analytic, must be called synthetic, even though they do not have imaginable contradictories. Specifically, they are such that no program of empirical inquiry could be formulated that would lead us to decide between them and their contradictories. The point is most effectively made in the following passage:

> To ask whether a thing could have been quite different, from what it is, whether Scott could actually have had all the properties of Handel, is on a different level. The questions we have just been asking are all to some degree requests for further descriptions of Scott. But the present question is not one that calls for any investigation of Scott, and it is difficult to accept that a question which calls for no investigation of Scott, to which nothing about Scott is relevant, is really about Scott. (p. 209)

On the basis of these considerations, Sprigge makes the following proposal:

> I suggest that a property is internal to a particular to the extent that no information about that particular is conveyed by one who says that it might have lacked that property. I think that the distinction between internal and external properties is not exact.… Let *F* be any property of a thing *a.* Then *F* is an external property of *a* if something interesting and true may be said of the form "if such and such then not-*Fa.*" Otherwise *F* is an internal property. But as from different points of view different things are interesting, so from different points of view different properties are internal and external. (p. 210)

The notion of "internal" is thus not only made a matter of degree but also relativized to the interests and purposes of those who are discussing *X.* Conceivably, everyone might be interested in *X* for a widely different reason; in this case, it would be quite possible that everyone might identify *X* by means of a widely different, but equally true, description. Then there would be no agreement on internal properties, and an Aristotelian metaphysics would seem unintelligible to us. As it stands, however, we tend to be interested in things for roughly the same reasons and thus to group the same things into the same natural kinds (for example, to regard Scott as "essentially" a man, rather than as a collection of physical particles occupying a given stretch of space time, or as a colorful patch on the landscape of nineteenth-century Scotland). Given this agreement and given our natural taxonomical instincts (our tendency to turn differences of degree into differences of kind whenever possible in order to facilitate inquiry), we can explain the commonsensical character of the distinctions between essence and accident and between internal and external properties (and, a fortiori, internal and external relations).

As an account of the internal-external distinction that avoids both the arbitrariness of Aristotelianism and the counterintuitive character of absolute idealism, Sprigge's proposal is a happy solution. But, like all such solutions, it is no better and no more permanent than the conceptual framework within which it is constructed. There is, to put it mildly, no consensus among philosophers of language as to when a sentence is "about" a given particular, when two sentences are about the same particular, the proper analysis of the notion of "name," the reducibility of names to descriptions, the assimilation of demonstrative pronouns to proper names, the question of whether proper names can be said to have meanings, the utility of the analytic-synthetic distinction, the equation of "necessary truth" with "analytic truth," and a host of related issues. In the absence of a comprehensive philosophy of language in which these issues are clarified and resolved in a systematic way, Sprigge's proposal must be treated as a useful guideline, rather than as a definitive

resolution of the issue concerning internal relations. One can imagine, for example, a revivification of the Aristotelian doctrine of predication, according to which "Socrates is a man" exemplifies a radically different sort of predication from "Socrates is a Greek," such an Aristotelian philosophy of language would, when conjoined with a realistic, anti-instrumentalist philosophy of science, produce a view according to which it would make good sense to say that Socrates's humanity really was internal to him, not simply relative to our interests but absolutely and intrinsically. Such a view would argue that "man" signifies a natural kind and is thus naturally suited to be a predicate "in the category of substance," whereas "Greek" or "atoms located at *p* at *t*" is not, and that this is an empirical truth.

There probably would never have been a problem about internal relations were it not for the efforts of speculative metaphysicians, such as Parmenides, Spinoza, and Hegel, to undermine our commonsense conceptual framework. If one rejects such attempts out of hand, one will treat the adoption of monism and of the thesis of the internality of all relations as a reductio ad absurdum of the premises from which these views are derived. Since Moore, the vast majority of Anglo-American philosophers have rejected such attempts and have differed only in their diagnoses of the confusions of falsehoods that engendered metaphysical conclusions. As long as the dogma that logical necessity was a matter of linguistic convention remained unchallenged, a simple and elegant resolution of the problem of internal relations seemed possible. However, recent doubts about this dogma (combined with the realization that Aristotle's distinction between essential and accidental properties is not simply a philosopher's invention but is firmly grounded in common sense) have made the problem look more complex than it appeared in the days of Ayer's *Language, Truth and Logic.* Philosophers who wish, as P. F. Strawson has put it, to substitute a "descriptive" metaphysics for a "revisionary" one are now faced with the problem of reconciling (*a*) the existence of this commonsense distinction with (*b*) the standard empiricist view that knowledge of how we speak either does not reveal anything about the nature of the objects we refer to, or at least does so in a very different way than does empirical research directed to those objects themselves, (*c*) the fact that the meaning we assign to a term is in part a function of the amount of empirical knowledge we possess, and (*d*) the fact that common sense seems to require a realistic, rather than an instrumentalistic, view of what it is to "know the nature of an object."

If the difficulties of such a reconciliation prevent "descriptive" metaphysicians from carrying out their chosen task, then the door will be open once again to the two extreme views examined above. It may turn out that common sense is, if not as incoherent as Parmenides and Bradley thought it, at least sufficiently inconsistent as to require the adoption of paradoxical philosophical theses. Whether one then turns to the extreme represented by Ayer's radical conventionalism and instrumentalism, or to the extreme represented by Blanshard's idealistic monism, will be largely a matter of taste. Both views, as suggested above, are parts of internally consistent philosophical systems. Each system retains certain portions of our commonsense framework and insists on these at the expense of other portions. In the absence of a touchstone other than common sense, it is difficult to see how a rational choice between such systems can be made.

See also Absolute, The; Aristotle; Ayer, Alfred Jules; Blanshard, Brand; Bosanquet, Bernard; Bradley, Francis Herbert; Carnap, Rudolf; Common Sense; Goodman, Nelson; Hegel, Georg Wilhelm Friedrich; Leibniz, Gottfried Wilhelm; Moore, George Edward; Nagel, Ernest; Parmenides of Elea; Putnam, Hilary; Quine, Willard Van Orman; Realism; Royce, Josiah; Ryle, Gilbert; Sellars, Wilfrid; Spinoza, Benedict (Baruch) de; Strawson, Peter Frederick; Universals, A Historical Survey.

Bibliography

Allaire, Edwin B. "Bare Particulars." In *Essays in Ontology.* Iowa Publications in Philosophy, Vol. I, 14–21. The Hague, 1963.

Alston, William P. "Internal Relatedness and Pluralism in Whitehead." *Review of Metaphysics* 5 (1951–1952): 535–558.

Anscombe, G. E. M. "Aristotle." In *Three Philosophers,* by G. E. M. Anscombe and P. T. Geach. Ithaca, NY: Cornell University Press, 1961. Gives a sympathetic treatment of Aristotle's distinction between secondary substance and quality.

Ayer, A. J. "Internal Relations." *PAS,* supp., 14 (1935): 173–185. Reprinted in Ayer's *Language, Truth and Logic.* London: Gollancz, 1936. Ch. 8.

Blanshard, Brand. *The Nature of Thought,* 2 vols. London: Allen and Unwin, 1939; New York: Macmillan, 1940.

Blanshard, Brand. *Reason and Analysis.* La Salle, IL: Open Court, and London: Allen and Unwin, 1962

California, University of. *Studies in the Problem of Relations.* University of California Publications in Philosophy, Vol. XIII. Berkeley, 1928.

Chappell, Vere. "Sameness and Change." *Philosophical Review* 69 (1960): 351–362. On criteria of self-identity.

Church, Ralph W. "On Dr. Ewing's Neglect of Bradley's Theory of Internal Relations." *Journal of Philosophy* 32 (1935): 264–273.

Ewing, A. C. *Idealism.* London: Methuen, 1934. This contains the best exposition of the similarities and differences between the exponents of absolute idealism and also of the relation between the doctrine that all relations are internal and other idealistic doctrines.

James, William. *Essays in Radical Empiricism.* New York: Longman, 1912. Ch. 3, "The Thing and Its Relations." On Bradley.

Moore, G. E. "External and Internal Relations." *PAS* 20 (1919–1920): 40–62. Reprinted in Moore's *Philosophical Studies.* New York: Harcourt Brace, 1922.

Nagel, Ernest. "Sovereign Reason," in his collection of articles *Sovereign Reason.* Glencoe, IL: Free Press, 1954.

Rome, Sydney, and Beatrice Rome, eds. *Philosophical Interrogations,* 219–246. New York: Holt, Rinehart and Winston, 1964. Largely a continuation of the debate between Nagel and Blanshard.

Royce, Josiah. *The Spirit of Modern Philosophy.* Boston: Houghton Mifflin, 1892. Expounds the relation between early rationalism and Hegel, as it was conceived by the absolute idealists.

Russell, Bertrand. "The Monistic Theory of Truth." In his *Philosophical Essays,* 150–169. London: Allen and Unwin, 1910. Criticism of absolute idealism.

Ryle, Gilbert. "Internal Relations." *PAS,* supp., 14 (1935): 154–172.

Sellars, Wilfrid. *Science, Perception and Reality.* London, 1963. Ch. 9, "Particulars." On criteria for self-identity of particulars.

Sprigge, Timothy, "Internal and External Properties." *Mind* 71 (1962): 197–212.

Stace, W. T. *The Philosophy of Hegel.* London: Macmillan, 1924. Reissued New York: Dover, 1955. Ch. 2 discusses the relation between early rationalism and Hegel, as it was conceived by the absolute idealists.

Thompson, Manley. "On the Distinction between Thing and Property." In *The Return to Reason,* edited by John Wild. Chicago: Regnery, 1953. Another defense of Aristotle's distinction between "secondary substance" and "quality."

Will, Frederick. "Internal Relations and the Principle of Identity." *Philosophical Review* 44 (1940): 497–514.

Wollheim, Richard. *F. H. Bradley.* Harmondsworth, U.K.: Penguin, 1959. Ch. 3 contains an exegesis of Bradley's treatment of relations.

Richard M. Rorty (1967)

RELATIVISM IN ETHICS

See *Ethical Relativism*

RELATIVITY OF KNOWLEDGE

See *Functionalism in Sociology; Historicism; Skepticism, History of; Sociology of Knowledge*

RELATIVITY THEORY

From 1905 to 1915 Albert Einstein revolutionized the conception of space and time and gravity that had been central in physics since Isaac Newton. For a brief discussion of the history of the development of relativity see the entry "Einstein, Albert." This entry describes the content of the theories.

The special and general theories of relativity are, at heart, theories of spatiotemporal structure. They are not particularly about observers or reference frames or ways to synchronize clocks, although as fundamental physical theories they have implications about what observers will observe and what various physical procedures for coordinating clocks will accomplish. It is easy to fall under the impression that these theories are basically concerned about coordinate systems or reference frames because physical events are typically described by means of coordinates or reference frames, but that temptation ought to be avoided.

Perhaps the easiest way to understand special relativity is by analogy to Euclidean geometry. Euclidean geometry postulates a particular spatial structure and, beginning with the Euclid's *Elements,* the implications of that structure for geometrical figures were studied by purely geometrical methods. For two millennia, the study of Euclidean geometry made no use of coordinate systems or of numbers. The introduction of Cartesian coordinates allowed for the translation of geometrical objects into algebraic ones by means of assigning numbers as coordinates to points. There are all sorts of ways to lay down coordinates on a Euclidean space, such as polar coordinates or spherical coordinates, but the most familiar is the system of Cartesian coordinates. Cartesian coordinates are rectilinear and orthogonal; the coordinate curves are straight lines that intersect at right angles. Because of this feature, distances between points in a Euclidean space are easy to calculate from their Cartesian cooridnates: If point *p* has coordinates (x_p, y_p, z_p) and point *q* has coordinates (x_q, y_q, z_q), then the distance from *p* to *q* is:

$$\sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2}.$$

In most spaces, such as the surface of a sphere, Cartesian coordinates do not exist. It turns out that for a space to be Euclidean is just for the space to admit of Cartesian coordinates. That is, the distances between points in the space must be of just the right form for Cartesian coordinates to exist.

In order to grasp relativity, we have to think not of distances between points in a three-dimensional space, but of a fundamentally spatiotemporal distance between

points in a four-dimensional space-time. Points in the space-time correspond to instantaneous, localized events, such as the bursting of a bubble when it reaches the surface of a glass of champagne. Such events occur both at a place and at a time. To locate these events, we typically ascribe to them four numbers, such as a latitude, longitude, altitude, and time. It is in this uncontroversial sense that the space-time of classical physics and of relativity is four-dimensional.

What sorts of spatiotemporal relations are there between events? All of classical physics agreed on at least one point: There is a definite, objective, purely temporal relation between the events. Two events either take place at the same time, or one takes place a certain amount of time before the other. So the notion of there being a lapse of time between events, and the specific case of simultaneity of events, is inherent in the classical account of space-time structure.

The classical account of spatial structure is not so straightforward. Newton believed that a single three-dimensional Euclidean space persists though time, and that every event, whenever it occurs, takes place somewhere in that absolute space. So Newton thought that any pair of events, no matter whether they occur at the same or different times, have some spatial distance separating them. But consider the following case: On a train traveling along the tracks, there sits a glass of champagne. A bubble rises to the surface and pops, followed a minute later by a second bubble. How far was the first popping from the second?

According to a passenger on the train, the two events took place in close spatial proximity, within a few inches of each other. But according to a spectator watching the train go by, these two events would be considered yards apart because the train has moved in the intervening minute. Newton would insist that there is a true spatial distance between the events, even though no observation could reveal for certain whether the passenger or the spectator (or neither) is right. But a natural reaction is to reject the whole question: There may be definite spatial relations between simultaneous events, but there is no fact at all about the spatial distance between nonsimultaneous events. Thus we arrive at two classical space-time structures: Newtonian space-time, with temporal and spatial relations between every pair of events, and Galilean (or neo-Newtonian) space-time, with temporal relations between all events and spatial relations only between simultaneous events (Galilean space-time then needs to add a new spatiotemporal structure, called an affine connection, to distinguish inertial from non-inertial trajectories). Note that the classical accounts agree on the temporal structure, and particularly on the objective physical relation of simultaneity.

Special relativity postulates a four-dimensional space-time with a radically different spatiotemporal structure. Instead of having a pure temporal structure and a pure spatial structure, there is a single relativistic "distance" between events (the scare quotes around *distance* must be taken seriously, as the quantity is not at all like a spatial distance). How can this spatiotemporal structure be specified?

The easiest method, albeit a bit roundabout, is by means of coordinates. Here we will take the analogy with Euclidean geometry quite seriously. As we saw, even though Euclidean geometry has no need of coordinate systems, the spatial structure of a Euclidean space can still be specified in this way; a Euclidean space is a space that admits of Cartesian coordinates. More specifically, a three-dimensional Euclidean space has a structure of distance relations among its points such that each point can be given coordinates (x,y,z) and the distance between any pair of points is:

$$\sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2}.$$

In exactly the same way, we can specify the spatiotemporal structure of Minkowski space-time, the space-time of special relativity. Minkowski space-time is a four-dimensional manifold that admits of Lorentz coordinates (or Lorentz frames). A Lorentz frame is a system of coordinates (t, x, y, z) such that the relativistic spatiotemporal distance between any pair of events *p* and *q* is:

$$\sqrt{(t_p - t_q)^2 - (x_p - x_q)^2 - (y_p - y_q)^2 - (z_p - z_q)^2}.$$

Written this way, the similarity with the example of Cartesian coordinates on Euclidean space is manifest; the only difference is the minus signs in place of plus signs. The consequences of that small mathematical difference are profound.

Before investigating the nature of this spatiotemporal structure, we should renew some of our caveats. First, there is always the temptation to invest the coordinates with some basic physical significance. For example, it is very natural to regard the coordinate we are calling *t* as a time coordinate, and to suppose that it has something to do with what is measured by clocks. But as of yet, we have said nothing to justify that interpretation. The Lorentz coordinates are just some way or other of attaching numbers to points such that the quantity defined above is proportional to the spatiotemporal distance between events. Indeed, just as there are many ways to lay down Cartesian coordinates on a Euclidean plane, systems differ with respect to the origin and orientation of the coordinate grid, so there are

many ways to lay down Lorentz coordinates in Minkowski space-time. Different systems will assign different *t* values to the points, and will disagree about, for example, the difference in *t* value between two events. We do not invest these differences with any physical significance; because the various systems agree about the quantity defined above, they agree about all that is physically real.

A second caveat is in order. We have been speaking so far as if the spatiotemporal distance between events is itself a number (viz., the number that results when one plugs the coordinates of the events into the formula above). But it is easy to see that this is wrong even in the Euclidean case. Distances are only associated with numbers once one has chosen a scale, such as inches or meters. What exists as a purely geometrical, nonnumerical structure is rather a system of ratios of distances. Having chosen a particular geometrical magnitude as a unit, other magnitudes can be expressed as numbers (viz., the numbers that represent the ratio between the unit and the given magnitude). The Greeks had a deep insight when they divided mathematics into arithmetic (the theory of number) and geometry (the theory of magnitude). They recognized that the theory of ratios applied equally to each field, but kept the two subjects strictly separate. Our use of coordinates to associate curves in space with algebraic functions of numbers has blurred the distinction between magnitudes and numbers. To understand relativity, it is important to recognize the conventions employed to associate geometrical structure with numerical structure.

Holding these warnings in mind, let us turn to the relativistic spatiotemporal distance. What are the consequences of replacing the plus signs in the Euclidean distance function with minus signs?

One obvious difference between the Euclidean structure and the Minkowski structure is this: In Euclidean space, the distance between any two distinct points is always positive, and the only zero distance is between a point and itself. In mathematical terms, the Euclidean metrical structure is positive definite. But in the Minkowski structure, two distinct events can have zero distance between them. For example, the events with coordinates (0,0,0,0) and (1,1,0,0) have zero distance (where we list the coordinates in the order (t,x,y,z)). Of course, this does not mean that these two events are the same event; assigning the numerical value zero to this sort of distance is just a product of the conventions we have used for assigning numbers to the distances. But the fact that two events have a zero distance between them does show that they are related in a particular spatiotemporal way. In order to remind ourselves that these spatiotemporal distances do not behave like spatial distances, from now on we will call them spatiotemporal intervals.

If we choose a particular event, the popping of a particular champagne bubble, and call the event *p*, then we can consider the entire locus of events that have zero interval from *p*. There will be infinitely many such events. If *p* happens to be at the origin of a Lorentz frame, assigned coordinates (0,0,0,0), then among the events at zero interval from it are (1,1,0,0), (1,0,1,0), (5,0,-3,4), and (-6,4,-4,2). To get a sense of how these events are distributed in space-time, we draw a space-time diagram, but again one must be very cautious when interpreting these diagrams. The diagrams must repress one or two dimensions of the space-time, because we cannot draw four-dimensional pictures, but that is not the principle problem. The main problem is that the diagrams are drawn on a Euclidean sheet of paper, even though they represent events in Minkowski space-time. There is always the danger of investing some of the Euclidean structure of the representation with physical significance it does not have. Bearing that in mind, the natural thing to do is to suppress the *z* coordinate and draw the *x*, *y*, and *t* coordinates as the *x*, *y*, and *z* coordinates of three-dimensional Euclidean space.

Adopting these conventions, the points at zero interval from (0,0,0) will be points that solve the equation $t^2 - x^2 - y^2 = 0$, or $t^2 = x^2 + y^2$. The points that solve this equation form a double cone whose apex is at the origin. According to relativity, the intrinsic spatiotemporal structure associates such a double cone with every event in the space-time. This locus of points is called the light-cone of the event p, and divides into two pieces, the two cones that meet at p. These cones are called the future light-cone and the past light cone of p.

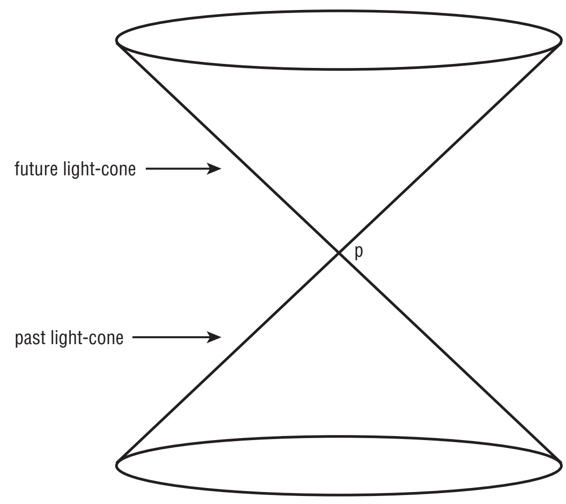

As the name light-cone suggests, we are now in a position to make contact between the spatiotemporal structure postulated by relativity and the behavior of physical entities. According to the laws of relativistic physics, any light emitted at an event (in a vacuum) will propagate along the future light-cone of the event, and any light that arrives at an event (in a vacuum) arrives along the past light-cone. So the tiny flash of light emitted when our champagne bubble pops races away from the popping event along its future light-cone. One can think of the ever-growing light-cone as representing the expanding circle (or, if we add back the *z* dimension, the expanding sphere) of light that originates at the bursting of the bubble.

Having associated the spatiotemporal structure with the behavior of an observable phenomenon such as light, we can now see how relativistic physics gains empirical content. For example, it is an observable fact that any pair of light rays traveling in parallel directions in a vacuum travel at the same speed; one light ray in a vacuum never overtakes another. This is not, of course, how material particles behave. One spaceship traveling in a vacuum can overtake another, or one electron in a vacuum can overtake another, because where a spaceship or an electron goes depends on more than the space-time location of the origin and direction of its journey. Two electrons can start out at the same place and time and set off in the same direction but end up in different locations because they were shot out at different speeds. Their trajectories depend on more than just the space-time structure. Light, in contrast, is intimately and directly tied to the relativistic space-time structure. Space-time itself, as it were, tells light in a vacuum where to go.

The assignment of zero relativistic interval between the origin of a light-cone and any event on it has one other notable consequence. We have already said that when we assign numbers to magnitudes, we want the ratios between the numbers to be identical to the ratios between the magnitudes. Because 0:0 is not a proper ratio, the relativistic interval does not license comparisons between the various intervals on a light-cone. If one light ray originates at (0,0,0,0) and travels to (1,1,0,0), and a second light ray originates at (0,0,0,0) traveling in some other direction, there is no fact about when the second light ray has gone as far as the first.

What other structure, beside the light-cone structure, does Minkowski space-time have? There is a well-defined notion of a straight line in the space-time, and this is accurately represented in our Euclidean space-time diagram: Straight lines in the Euclidean diagram correspond to straight trajectories in the space-time. Indeed, we have tacitly been appealing to the notion of a straight line all along; when we speak of the relativistic interval between two events, we mean the interval as measured along a straight line connecting the events, or, even more precisely, we mean the relativistic length of the straight line that connects the events. The straight-line structure (affine structure) of Minkowski space-time plays a central role in framing physical laws.

If a light ray is emitted from (0,0,0,0) into a vacuum, we already know that its trajectory through space-time will lie on the future light-cone of (0,0,0,0). But more than that, the trajectory will be a straight line on the light-cone. An analogous fact holds for material particles that travel below the speed of light. If a material particle is emitted from (0,0,0,0), its trajectory will lie entirely within the future light-cone of (0,0,0,0), which is to say that the particle can never travel at or above the speed of light. But more than that: If the particle is emitted into a vacuum, and is not subject to any forces, then its trajectory will be a straight line in space-time.

This law, in abstract form, enormously predates the theory of relativity. For this is just the proper space-time formulation of Newton's first law of motion: "Every body continues in its state of rest, or of uniform motion in a right line, unless compelled to change that state by forces impressed on it." The trajectory of a particle at rest or in uniform motion in Newtonian space-time is a straight line through the four-dimensional space-time. Newton's first law, stated in terms of space-time trajectories, also retains the same form in Galilean space-time, and can be taken over without change into Minkowski space-time. As we will see, in this abstract space-time formulation, Newton's first law also holds in the general theory of relativity. That is why we should try to formulate physical laws directly in terms of space-time structure.

Once we deal with material particles that travel below the speed of light, the relativistic interval takes on even greater significance. Consider a particle that travels from (0,0,0,0) to (5,4,0,0) along a straight trajectory (i.e., a particle emitted from the origin of the coordinate system that arrives at the event [5,4,0,0] without having any forces acting on it). The relativistic interval along its space-time trajectory is:

$$\sqrt{(5-0)^2-(4-0)^2-(0-0)^2-(0-0)^2}=3.$$

The size of this interval has direct physical significance; it is proportional to the amount of time that will elapse for a clock that travels along that trajectory. Clocks in the theory of relativity are like odometers on cars; they meas-

ure the length of the path they take. But *length* here means the interval, and *path* the space-time trajectory of the clock. Events in space-time separated by positive intervals are time-like separated.

It is not, of course, a further unanalyzable postulate of relativity that clocks measure the interval along their trajectory; clocks are physical mechanisms subject to physical analysis. But one can easily analyze how a simple clock will behave, such as a clock that counts the number of times a light ray gets reflected between two mirrors, and find that the reading on the clock will be proportional to the interval along the clock's trajectory.

With the clock postulate in hand, we can now analyze the notorious *twins paradox* of relativity. One of a pair of twins takes a rocket from Earth and travels to a nearby star. Upon returning to Earth, the twin has aged less than the stay-at-home sister, and the clocks in the twins' spaceship show less elapsed time than those that remained on Earth. Why is that?

To be concrete, suppose the event of the rocket leaving Earth is at the point (0,0,0,0) in our coordinate system, and the rocket travels inertially (without acceleration) to the point (5,4,0,0). The rocket immediately turns around, and follows an inertial trajectory back to Earth, arriving at the event (10,0,0,0). The interval between (0,0,0,0) and (5,4,0,0) is, as we have seen, 3. Suppose this corresponds to an elapse of three years according to the onboard clocks. The return trajectory from (5,4,0,0) to (10,0,0,0) also has an interval length 3, corresponding to another three years elapsed. So the astronaut twin arrives back having aged six years, and having had all the experiences that correspond to six years of life.

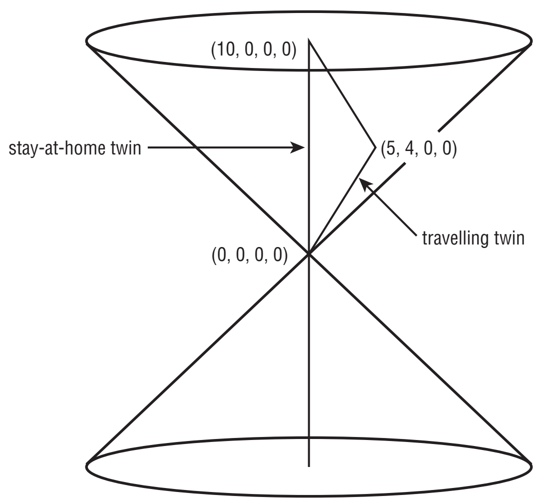

The stay-at-home twin, however, always remained at the spatial origin of the coordinate system. Her trajectory through space-time is a straight line from (0,0,0,0) to (10,0,0,0). So the interval along her trajectory is 10, corresponding to an elapse of ten years. She will have biologically aged ten years at her sister's return, and had four more years of experience than her twin.

The relativistic analysis of the situation is quite straightforward. It is really no more surprising, from a relativistic perspective, that the clocks of the twins will show different elapsed times from departure to return than it is surprising that two cars starting in the same city and end in the same city will show different elapsed mileage on their odometers, given that one took the freeway and the other a winding scenic route. The sense that there is a fundamental puzzle in the twins paradox only arises if one has mistaken views concerning the content of the theory of relativity.

In particular, it is often said that, according to the theory of relativity, all motion is the relative motion of bodies. If so, then there seems to be a complete symmetry between the twins: The motion of twin A relative to twin B is identical to the motion of twin B relative to twin A. But the relative motion of the twins plays no role at all in the physical analysis of the situation. The amount of time that elapses for twin B on her trip has nothing to do with what twin A is doing, or even if there is a twin A. The amount of time is just a function of the space-time interval along her trajectory.

It is also sometimes said that the theory of relativity gets rid of all absolute spatiotemporal structure; all facts about space and time are ultimately understood in terms of relations between bodies, so in a world with only one body there could be no spatiotemporal facts. This is also incorrect. The special theory of relativity postulates the existence of Minkowski space-time, whose intrinsic spatiotemporal structure is perfectly absolute, in whatever sense one takes that term. It is not a classical space-time structure, but it is not just a system of relations between bodies.

One occasionally also hears that the resolution of the twins paradox rests on facts about acceleration; the situation of the two twins is not exactly symmetric because the astronaut twin must accelerate (when she turns around to come home), whereas the stay-at-home twin does not. That is true, but irrelevant: The difference in elapsed time is a function of the intervals along the trajectories, not a function of the accelerations that the twins experience. Indeed, in the general theory of relativity we will be able to construct a twins scenario in which neither twin accelerates at all, but still they suffer different elapsed times between parting and reunion. It would be just as mis-

leading to attribute the difference in elapsed time to the accelerations of the twins as it would the difference in odometer reading to the accelerations of the cars, even if the car that took the longer route did accelerate more.

The paradoxical or puzzling aspect of the twins paradox really arises from the difference between Euclidean geometry and Minkowski space-time geometry. If we draw the trajectories of the twins in space-time, we get a triangle whose corners lie at (0,0,0,0), (5,4,0,0), and (10,0,0,0). The astronaut twin travels along two edges of this triangle, whereas the stay-at-home twin travels along the third. And in Euclidean geometry, the sum of the lengths of any two sides of a triangle are greater than the length of the remaining side. But in Minkowskian geometry, the opposite is true: The sum of the intervals of two sides is less than the interval along the remaining side. Indeed, for time-like separated events, a straight line is the longest path between the two points in space-time. This is one consequence of exchanging the plus signs in the Euclidean metric for minus signs in the Minkowski metric.

The relativistic clock postulate has been most strikingly checked using natural clocks: unstable particles whose decay rate displays a known half-life in the laboratory. The muon, a sort of heavy electron, is unstable and will decay on an average of 10^{-6} seconds after having been created. Muons can be created in the upper atmosphere by collisions between molecules in the air and high-energy cosmic rays. According to clocks on Earth, it should take the muon about 10×10^{-6} seconds to reach the Earth, so very few should survive the trip without decaying. Nonetheless, many more muons than that calculation suggests do reach the Earth's surface. Calculation of the interval along the muon's trajectory predicts this because that interval corresponds to less than 10^{-6} seconds.

If we idealize muons a bit, and imagine that they all decay in exactly 10^{-6} seconds (according to their own clocks), then we can use them to map out the geometry of Minkowski space-time. Suppose we create a swarm of muons in space and send them out in all directions. Their decays will provide a map in space-time of events that are all the same interval from the point of creation. If we choose units so that the size of the interval corresponds to seconds, and we choose the creation of the muons as the origin of the coordinate system, then the coordinates of the decay events will satisfy:

$$\sqrt{t^2 - x^2 - y^2 - z^2} = 10^{-6}$$

This is the equation of a hyperboloid of revolution that asymptotically approaches the light-cone, as depicted below.

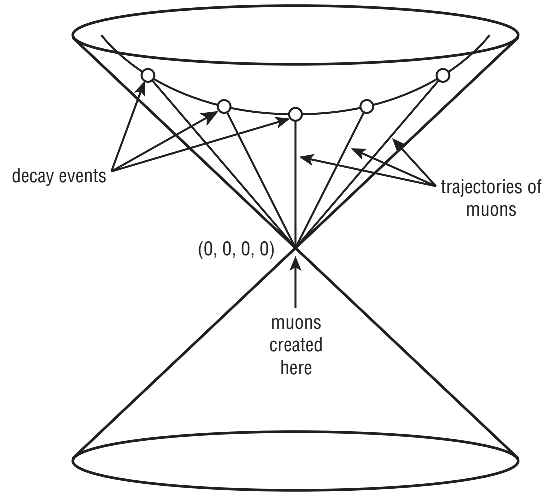

The hyperboloid represents events all at the same interval from (0,0,0,0), and so corresponds to a circle or sphere of fixed radius in Euclidean geometry. There would be a corresponding hyperboloid in the past light-cone, representing places from which a muon could have been sent that would have decayed at (0,0,0,0).

Indeed, we are now in a position to make a thoroughgoing analogy between the geometry of Minkowski space-time and Euclidean geometry that makes no reference to coordinates at all. Classical Euclidean geometrical proofs do not use coordinate systems of numbers, they use two instruments: the straightedge and the compass. The straightedge allows one to identify straight lines in the space, and the compass to draw the locus of points at a fixed distance from a given center. In Minkowski space-time, we can use light rays in a vacuum and inertially traveling particles as straightedges because their trajectories are straight lines in the space-time. Setting a Minkowski compass at interval zero and identifying a center should result in drawing the light-cone: the locus of points of interval zero from the center. So we can use light rays for this purpose. Setting the compass to draw points at a fixed positive interval should result in drawing hyperbola; we can use clocks for this just as the muons are employed above. In this way, we can free Minkowski geometry from coordinates altogether.

So far we have left one species of space-time relation out of account. All the points on the past or future light-cone of some event p are at zero interval from p. All the events inside the past or future light-cone are at positive interval from p (taking always the positive square root by

convention). What of points that are outside the light-cone altogether?

The point labeled (0,1,0,0) is outside the light-cone of the point (0.0.0.0). If we plug these coordinates into our formula, we find that the interval between the points is:

$$\sqrt{(0-0)^2-(1-0)^2-(0-0)^2-(0-0)^2}=\sqrt{-1}=i.$$

That is, according to the definition of the interval that we have given, the interval between these points is imaginary. What could this mean?

Once again, we have to recall that the assignment of numbers to the intervals is somewhat a matter of convention. In fact, some physics books define the interval as:

$$\sqrt{(x_p-x_q)^2+(y_p-y_q)^2+(z_p-z_q)^2+(t_p-t_q)^2},$$

Here the interval between time-like separated events becomes imaginary. Does this mean that a clock could measure an imaginary number? Of course it can: Just take a regular clock and paint a little *i* after all the numerals! The numbers we assign to intervals have no intrinsic significance; it is the ratios between the numbers that represent the ratios among the magnitudes. Events that lie outside each other's light cones, so-called space-like separated events, have intervals among them that also stand in ratios to each other. The set of events at fixed space-like separation from (0,0,0,0) forms another sort of hyperboloid of revolution, depicted below.

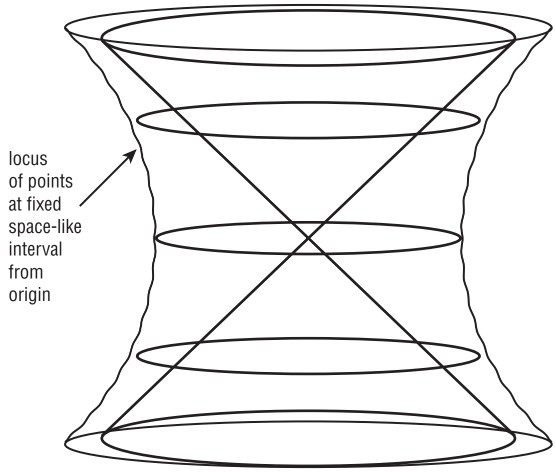

We now have a sense of the spatiotemporal structure of Minkowski space-time. A special relativistic physical theory must have laws that employ only this spatiotemporal structure. We could now go on to see how, for example, classical electromagnetic theory can be reformulated in this way, but that would take us too far from foundational issues.

It should be noticed that this account of special relativity has made no mention at all of several well-known features often associated with relativity, such as the constancy of the speed of light, the relativity of simultaneity, and the Lorentz-Fitgerald contraction. That is because all of these are frame-dependent (or coordinate system dependent) effects, and we have been presenting the theory in a frame-independent way. For example, we have no basis to discuss the relativity of simultaneity because we have had no ground, and no need, to introduce any notion of simultaneity at all. In classical physics, simultaneous events are events that take place at the same time, but we have no general notion of the time at which an event occurs, only the time that elapses on a clock following a certain trajectory. So the proper thing to say is not that special relativity implies the relativity of simultaneity, but that it implies the nonexistence of any objective notion of simultaneity. And we cannot discuss whether the speed of light is constant because we do not have any grounds to ascribe any speed to anything.

We have seen that a light ray can never overtake another light ray, but assessing a speed requires determining how far an object went in a given period of time. So far, we have not needed any notion of the distance an object travels, nor of the time that it takes to travel that distance. We can say how much time will elapse on a clock that follows a given trajectory, but that is evidently no use in defining a speed of light; no material clock can travel along with a light ray, and if it could, it would show no elapsed time for the journey. The notion of simultaneity requires a global time function, that is, an assignment of times to all events, so that there is a locus of events that are all assigned the same time. And the notion of a speed requires both the notion of the time that elapses between the start and the end of a journey, and the notion of the distance covered in that time. The relativistic space-time structure does not, per se, support either of these notions.

There is, however, a reasonably natural method for introducing both a global time function and a notion of spatial distance into Minkowski space-time. We begin with a family of co-moving inertial clocks (i.e., a family of clocks all moving on straight, parallel trajectories through space-time). There will be an infinitude of such families, corresponding to all the directions their trajectories can have. We begin by picking one such family.

We now want to "synchronize" the clocks. Scare quotes have to be put around the word since the classical notion of synchronization presupposes the notion of simultaneity: Synchronized clocks all show the same time at the same instant. But in relativity there is no such thing

as *the same instant*. So one must think of the method we are about to describe as a way to coordinate a family of clocks that we simply call synchronizing them.

Let us choose a single master clock from our family of co-moving clocks. The other clocks will coordinate with this master clock by the following method: Each clock sends a light ray to the master clock, noting the time of emission (according to the sending clock). When the light ray reaches the master clock, it is immediately sent back and shows the time reading on the master clock at the moment it arrived. When this return signal reaches the sending clock, the time reading on the sending clock is noted. The sender, then, has three bits of data: the time it sent the signal (according to the sending clock), the time it received the return signal (ditto), and the reading on the master clock when the signal got to it. On this basis, the sending clock synchronizes with the master clock by adjusting its time so that the time that the master clock read when the signal arrived corresponds to the event on the sending clock exactly midway between the moment the signal was sent and the moment the return signal arrived. All of these notions are relativistically well-defined, so this method of coordinating clocks can be carried out. Every event in space-time is now assigned a time (viz., the reading on that member of the family of clocks that passes through the event when it passes through the event).

We can now identify simultaneity according to this family of clocks as sets of events that are all assigned the same time by this family of clocks. Such a set is called a simultaneity slice through the space-time. The figure below shows one such simultaneity slice. Because all of the light signals that reach the master clock at noon lie on the past light-cone of the master clock showing noon, and because all of the return signals lie on the future light-cone of that event, it is easy to calculate the points at which all of the coordinated clocks will register noon. It is the flat plane in the middle.

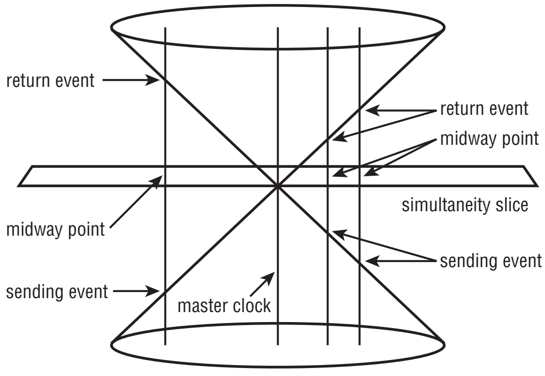

The simultaneity slice is a function of which family of co-moving clocks we choose. Choosing another family will give a different notion of simultaneity:

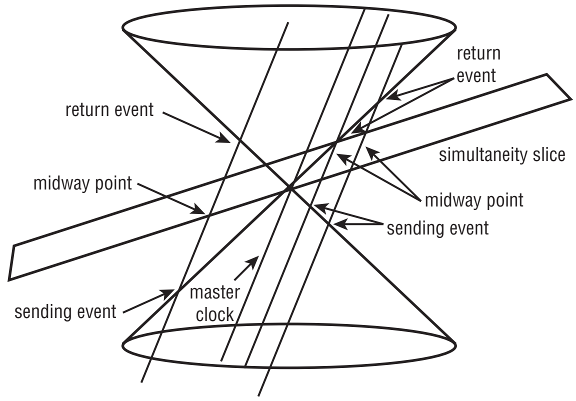

Each family of co-moving clocks determines its own notion of simultaneity, and these various notions render different judgments concerning which pairs of events happen at the same time. All the families will agree about the time order of time-like or light-like separated events, but for any pair of space-like separated events, some families will say that they happened at the same time, others that one happened first, and yet others that the other happened first. Each family introduces its own global time function. None of these functions is superior to the other, and none is needed at all to explicate the basic spatiotemporal structure.

What of spatial distance? Once a family of clocks has been synchronized, there is a simple way to assign a spatial distance between any pair of clocks. Send a light ray from one clock to the other. We can now understand the time of travel for the light ray as the difference between the time showing on the emitting clock at the emission event and the time showing on the receiving clock at the reception event. So we now have a definition of how long the light ray took to get from one clock to the other (again, this is not the time that a clock traveling along with the light ray would show elapsing). If we now define the speed of light to be a given constant, c, then we can say that the distance between the clocks is just c times the elapsed time of transmission. This will give us a structure of spatial distances between the clocks as defined by that particular family of clocks. Those spatial distances will, in special relativity, constitute a Euclidean space. Different families of clocks will disagree about the precise spatial distance between events, and about the spatial size of material objects, but each family will construct for itself a Euclidean spatial structure. Finally, if we allow such a

family of clocks to introduce Cartesian coordinates on its Euclidean space, then the family will assign each event four coordinate numbers: the three spatial coordinates and the global time function. These are exactly the Lorentz coordinate frames that we began with to express the relativistic metric, so we have come full circle.

The interconnection between the global time defined by a family of clocks and the spatial structure among events defined by that family resolves many of the intuitive puzzles in special relativity. We have seen that, according to clocks at rest on the Earth, a high-energy muon has a much longer lifetime than a muon at rest. That explains, from the point of view of the Earth frame, how the muon manages to make the trip to the surface. But of course, from the point of view of the muon, and clocks co-moving with it, the muon lifetime is the normal 10-6 seconds. From their point of view, the Earth is approaching them at high velocity. In that frame of reference, the muon is able to get through the whole atmosphere not because of any slowing down of their clocks, but because of the spatial contraction of the atmosphere. In the muon's frame of reference, the distance from the upper atmosphere to the Earth is much less than we on Earth take it to be.

The Lorentz contraction and time dilation effects of relativity then arise as disagreements that occur between the Lorentz frames about the amount of time that elapses between events and the spatial distance between events. Clocks in any frame will be seen to run slow according to the time function associated with any other frame. A meter stick at rest in one frame will be judged to be less than a meter long according to a frame in which the stick is moving. These are symmetric effects: From the point of view of any Lorentz frame, clocks at rest in any other frame run slow. We need to sharply distinguish these effects from the twins paradox. There, the difference in elapsed time for each twin is a consequence of the fundamental spatiotemporal structure, and has nothing do to with frames or families of clocks. The time dilation between frames results only from different ways of defining coordinates. In the latter case, there is no fact about which set of clocks is really running slower, but in the former case there is an objective fact about which twin is biologically younger when they are reunited.

GENERAL RELATIVITY

Special relativity is a theory that postulates a certain intrinsic spatiotemporal structure, and then formulates the laws of physics in terms of that structure. General relativity is the relativistic theory of gravity. It is also fundamentally a theory about spatiotemporal structure, and allows for different structures than special relativity. So the first question that arises when approaching general relativity is why gravity should particularly be connected to spatiotemporal structure. The special relativistic theory of electromagnetism, for example, simply accepts the Minkowski space-time and employs it in framing the electromagnetic laws. But gravity, in contrast, led to the rejection of special relativity in favors of a new theory. What is so special about gravity?

One sometimes hears that there needed to be a relativistic theory of gravity because Newton's gravitational theory postulates that gravity acts instantaneously between distant masses, but in relativity there is no available notion of instantaneous action (because there is no physical notion of simultaneity). But this observation does nothing to suggest that the theory of gravity should require any change from the special relativistic space-time. Classical electrostatics postulated that the coulomb force between distant charged particles acts instantaneously, but electromagnetic phenomena do not require changes to special relativity. Rather, relativistic electrodynamics simply rejects the claim that electric and magnetic forces act instantaneously. Electromagnetic influences are propagated along the light cones, at the speed of light, by electromagnetic waves. Similarly, one might think that the obvious way to deal with gravitation is simply to deny that it acts instantaneously. Let the gravitational effects also propagate along the light cones, and the special relativistic structure can be used to formulate the laws.

Adding such a delay in gravitational influence would, of course, modify the predictions of Newtonian gravity. One might even plausibly argue that Newton himself would have expected such a correction to his instantaneous gravity. For Newton thought that gravitational forces were mediated by particles exchanged between the gravitating bodies, and he would have expected the particles to take some time in traveling between the bodies. Of course, the fundamental cause of the gravitational force was a topic on which Newton refused to *fingere* any hypothesis, so we must be a bit speculative here. But it is worthwhile to note that if we modify classical Newtonian gravitational theory to allow gravitational influence to propagate along the light cones, we can exactly derive some famous relativistic effects, such as the anomalous advance in the perihelion of Mercury.

In order to understand why gravity is plausibly taken to be deeply connected to space-time structure, we need to look elsewhere. Consider again the family of co-moving inertial clocks we made use of in our discussion

of special relativity. Once set in motion, the family of clocks will move together, never approaching or receding from each other. That is because: a) the clocks are all traveling inertially, not subject to any force; b) according to the space-time version of Newton's first law, the trajectories of bodies subject to no forces will be straight lines in space-time; and c) the straight-line trajectories of the co-moving clocks form a family of parallel straight lines. Note that in giving this argument, we never had to mention the mass of any of the clocks. Because they are moving inertially, the trajectories of the clocks are determined by the intrinsic space-time structure, without the mass playing any role. It would not matter if some of the clocks were heavy and others light; they would still move together parallel to one another.

In Newtonian physics, the mass of a body only comes into consideration when a body is subject to a force and thereby deflected off its inertial trajectory. The inertial mass of a body is nothing but a measure of the body's resistance to being deflected by a force from its inertial trajectory: The more massive a body is, the harder it is to make its trajectory bend in space-time. Newton's second law, which we now render **F** = m**A**, tells us that the same force will only produce half the acceleration in a body that is twice as massive. So in the presence of forces, the trajectories of bodies will depend on their masses, whereas in the absence of forces the more and less massive bodies will move on parallel trajectories. Turning this observation around, we should find it very suggestive if there is a situation in which the trajectory of a body does not depend at all on its mass. It is natural to suspect that in such a situation, the mass of the body is playing no role because the body is not being subject to any force; it is moving inertially.

Recall Galileo at the top of the Leaning Tower of Pisa dropping a lighter and heavier object and seeing them hit the ground together. Here is a common situation in which the mass of an object does not affect its trajectory: The heavy and light follow the same space-time path. According to Newtonian gravitational theory, this is a rather fortuitous result. In that theory, both the heavy and the light object are subject to a force, the force of gravity, and so each is being deflected off its inertial trajectory. But, luckily, the gravitational force on each object is exactly proportional to its inertial mass. So the more massive object, which needs a greater force to be accelerated, is subject to a greater force than the less massive object. Indeed, the gravitational force on the more massive object is exactly as much larger as it needs to be to produce precisely the same acceleration as the lesser force of gravity produces on the less massive object. That, according to Newton, is why they fall together; they are both accelerated, but at exactly the same rate.

If we follow the hint above, though, we will be led to suspect a different account. Perhaps the two objects move together not because they are equally deflected off their inertial, straight-line trajectories, but rather because they are both following their inertial trajectories. Because the inertial trajectories are straight lines in space-time, this suggests a deep connection between gravity and fundamental spatiotemporal structure.

In this way we arrive at the general theory of relativity. According to general relativity, objects that are falling in a gravitational field or under the influence of a gravitational force are not being affected by any force at all. Gravity does not deflect objects from their inertial paths, it rather influences the very structure of space-time itself. The balls falling from the Leaning Tower of Pisa, or the planets orbiting the sun, are following straight trajectories through space-time.

To realize this theory, we must reject Minkowski space-time. Consider, for example, two satellites orbiting the Earth in opposite senses. The space-time diagram of the situation looks like this:

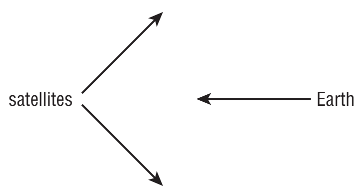

As the satellites orbit, their paths cross and recross in space-time. But in Minkowski space-time, as in Euclidean space, two straight lines can intersect only once at most. So the space-times of general relativity must have a different spatiotemporal structure than the space-time of special relativity.

An analogy with pure spatial geometry helps here. Euclidean geometry is just one of an infinitude of spatial geometries. Lines on the surface of a sphere, for example, do not satisfy Euclid's postulates. But even spherical geometry is highly regular compared to most geometries. Consider, for example, the surface of North America. In regions of the Great Plains, the geometry is nearly Euclidean (and even more nearly spherical), whereas in the Rocky Mountains the geometry of the surface varies wildly from place to place. We need new mathematical machinery to deal with this sort of situation.

The general mathematics needed is called differential geometry. Differential geometry is suited to deal with spaces whose geometrical structure varies from place to place. In some regions, a space may be locally Euclidean, in others non-Euclidean, so we have to be able describe the geometry region by region.

Euclidean spaces have a particularly uniform geometrical structure that allows them to admit of very convenient coordinate systems. As we have seen, a Euclidean space admits of Cartesian coordinates, in which the distances between points is a simple mathematical function of the coordinates of the points. Non-Euclidean spaces do not admit of such convenient systems. For example, points on a sphere can be coordinatized by latitude and longitude, but distances between the points on a sphere are not a simple function of their coordinate differences. If you are near the North Pole, you can change your longitude by several degrees just by taking a few steps; near the equator the same change of longitude would require traveling hundreds of miles. And even spherical coordinates are relativity simple and uniform.

To get a sense of a completely generic coordinate system, imagine walking down a road where each successive house has an address—one greater than the house before. You want to get to house number 200 and you are currently at house 100. How far must you walk?

There is no way to tell. If you go through a densely populated area, such as a small town, you will get to your destination quickly. If it is a sparsely built region, you may have to walk a long way. To know how far you have to go, you would need a complete listing of the distances between successive houses. If you have such a list, you can calculate the distance between any two houses, and so can reconstruct the geometrical structure of the region where the houses are built. In an analogous way, the general theory of spaces allows for the use of any arbitrary coordinate system. Accompanying the system is a metric that specifies the distances between nearby points. We do not have any general rule for calculating distances between distant points as a function of their coordinates, but we do not need one. The distance between faraway points is just the length of the straight path that connects them, and we can calculate the length of that path by knowing the distance between nearby points and adding up all the distances along the path. Thus we have the mathematical tools to deal with generic spaces of variable curvature that admit of nothing like Cartesian coordinates.

It is sometimes said that the general theory of relativity requires us to replace Euclidean space with a non-Euclidean space, but that is not a very useful, or accurate, explanation of the situation. As we have seen, even in special relativity the notion of spatial geometry is rather derivative and non-fundamental. The fundamental notion is the relativistic interval, which is a spatiotemporal object. It is only relative to a family of co-moving objects, such as clocks, that we can even define a spatial geometry. It turns out that, in special relativity, each such family will ascribe Euclidean geometry to its space, but that is somewhat fortuitous; there is no logical guarantee that the various families will agree on their findings. After all, in special relativity the various families will disagree about the exact spatial distance (and temporal gap) between a given pair of events. In general relativity, there will, in general, not exist families of co-moving inertial observers that maintain the same spatiotemporal relations to one another, and so there is no unproblematic way to define a spatial geometry at all.

In any case, it is simply incorrect to say that objects moving in a gravitational field trace out straight paths in a non-Euclidean spatial geometry. The orbits of the planets, for example, are nearly elliptical in any reasonably defined space for the solar system, and the ellipses are not (spatially) straight lines.

The proper account of general relativity rather employs an analogy. As the variably curved non-Euclidean spaces are to Euclidean space, so the variably curved space-times of general relativity are to Minkowski space-time. The orbits of the satellites depicted above are not straight paths in any spatial geometry, but they are straight paths in space-time. The effect of the Earth is not to produce a force that deflects the satellites off their inertial paths, it is to alter the space-time geometry so that it contains inertial paths that cross and recross.

On the Newtonian picture of gravity, when we sit on a chair we are not accelerated because we are acted on by counterbalancing forces: The gravitational force pulling us down and the force of the chair pushing us up. According to the general relativistic account, the force of the chair pushing up still exists, but it is unbalanced by any gravitational force. It follows that according to general relativity, as we sit we are constantly accelerated (i.e., constantly being deflected off of our inertial, straight-line trajectories through space-time). The inertial trajectory is that of an object unsupported by anything like the chair (i.e., an object in free fall).

The curvature of general relativistic space-time is partially a function of the distribution of matter and energy; that is why space-time near a massive object like the Earth is curved in such a way as to produce a gravitational field. This connection between the matter and

energy distribution and the spatiotemporal geometry is provided by Einstein's general relativistic field equations. But although the distribution of matter and energy influences the space-time geometry, it does not completely determine it.

The situation is similar to the relationship between the electromagnetic field and the electric charge distribution in classical physics. The presence of electric charges contributes to the electromagnetic field, but does not, by itself, determine it. For example, even in a space devoid of electric charges, there can be a nonzero electromagnetic field: electromagnetic waves (i.e., light) can propagate through the vacuum. Similarly, the general theory predicts the existence of gravitational waves—disturbances of the spatiotemporal geometry that can exist even in the absence of any matter or energy and that propagate at the speed of light. There are, for example, many vacuum solutions of the Einstein field equations. One solution is Minkowski space-time, but other solutions contain gravitational waves.

Because general relativity concerns spatiotemporal structure, and because the trajectory of light rays is determined by the light-cone structure, general relativity must predict the gravitational bending of light. It is not clear whether Newtonian physics would predict any gravitational effect on light because that would depend on whether light feels any gravitational force, but light certainly does propagate through space-time. The effect of gravity on light was dramatically confirmed in Arthur S. Eddington's 1919 eclipse expedition, but is even more strikingly illustrated in the phenomenon of gravitational lensing: A galaxy positioned between the Earth and a more distant light source can act as a lens, focusing the light of the distant source on the Earth. Two astronauts traveling inertially could experience a similar effect; they could take different straight paths that both originate at their home planet and both end on Earth, going different ways around an intervening galaxy. Because the relativistic interval along those paths could differ, such astronauts could illustrate the twins paradox without any acceleration; twins coming from the distant planet could have different biological ages when they reunite on Earth, even though neither suffered any acceleration.

The spatiotemporal geometry of general relativity accounts for familiar gravitational phenomena, but the theory also has dramatic consequence at the cosmological scale and in extreme physical conditions. When a massive star burns through its nuclear fuel and collapses, for example, the increasing density of matter causes ever greater curvature in space-time. If the star is sufficiently massive, the light-cone structure deviates enough from Minkowski space-time to form a trapped surface: a region from which light cannot escape. The event horizon around a black hole is such a trapped surface; an object falling through the horizon can never send light, or any other signal, back to the exterior region. Once the infalling matter of the star reaches this point, it is destined to become ever more compressed without limit, and the curvature of the space-time will grow to infinity. If the equations continue to hold, this results in a space-time singularity; the spatiotemporal structure cannot be continued beyond a certain limit and space-time itself comes to an end. Because the spatiotemporal structure itself has become singular, it no longer makes any conceptual sense to ask what happens after the singularity; no meaning could be attached to the term after in the absence of spatiotemporal structure.

In the opposite temporal direction, the general theory also contains models in which the universe as a whole arises out of such a singularity, the singularity we call the big bang. Indeed, if general relativity is not modified, the observed motions of galaxies require that the universe began at a singularity, and that space-time itself has been expanding ever since. There is equally no sense to be made of the question what happened before the big bang because the spatiotemporal structure needed to define temporal priority would not extend beyond the initial singularity.

It is, of course, possible that the equations of the theory will be modified in some way so as to avoid the infinities and singularities, but that takes us from the analysis of general relativity into speculations about the replacement of general relativity. The mathematical structure of general relativity also admits of models of the theory with very peculiar spatiotemporal structures. Some models, for example, admit closed time-like curves, that is, time-like trajectories that loop back through space-time and meet up with themselves. In such a model, a person could in principle continue going always locally forward in time, but end up (as an adult) back at the events of their childhood. There seems to be no way to physically test this possibility (that is, there is no physical mechanism to produce closed time-like curves through laboratory operations), so it is unclear whether the existence of these mathematical models proves the physical possibility of such time travel or rather the physical impossibility of space-times that correspond to these mathematical solutions. In any case, general relativity provides a means for considering spatiotemporal structures unlike any that occur in classical physics.

The special and general theories of relativity provide a rich source of novel concepts of great interest to metaphysics. The topics that could be informed by these theories are too long even to list, but the most obvious metaphysical implications of the theories are worthy of remark. The nature of space and time occupies a central place in Immanuel Kant's *Critique of Pure Reason*, where supposed a priori knowledge of spatial and temporal structure provided grounds for the conclusion that space and time have no existence outside the faculty of intuition. After all, how could one know anything a priori about space and time if they exist outside the mind?

The theories of relativity simply refute the claim that there is any a priori knowledge of spatiotemporal structure. Even if relativity ultimately proves to be incorrect, everything in our everyday experience of the world can be accommodated in the relativistic spatiotemporal account. For all we know at present, we could be living in a relativistic universe, in which there is no Euclidean space and in which even time need not have a universal linear order. The nature of space and time is a matter of empirical inquiry, not a priori proof.

The special and general theories are also relevant to the question of the nature of space and time: Are they entities in their own right (as Newton supposed) or just relations among material bodies (as G.W. Leibniz insisted)? Taken at face value, the theories posit an independent existence to the four-dimensional space-time manifold. Even in the absence of material bodies, there is a spatiotemporal structure among the points of space-time. As the twins paradox shows, the observable behavior of material objects is determined by that structure. And even more dramatically, in general relativity the space-time manifold takes on a life of its own; gravitational waves can exist even in the absence of any material objects, and the presence of material objects influences the structure of space-time.

Attempts have been made to reformulate general relativity in a more relationist manner, in terms only of relations among material objects without commitment to any spatiotemporal structure of vacuum regions. These attempts have not succeeded. One can, of course, simply declare that in the general theory, space-time itself counts as a material entity, but then the argument seems to be only over labels rather than ontology.

Like all empirical theories, relativity is supported but not proven by observation. The spatiotemporal structure cannot be directly observed, but theories of matter couched in terms of the relativistic structure yield testable predictions that can be checked. The general theory, for example, has been checked by flying an atomic clock around the world and comparing its reading with an initially synchronized clock that remained on Earth. Because the trajectories of the clocks have different relativistic intervals, one can predict that the traveling clock will show a different elapsed time from the clock that remained behind—which it does. There may be other ways to explain the effect, but it is a natural consequence of the relativistic account of space-time structure.

Challenges to the theory of relativity are more likely to come from considerations of the compatibility of the theory with other fundamental physical theories than from direct empirical problems. It is, for example, a still unsolved problem how to reconcile quantum physics with the pure relativistic space-time structure, and another unsolved problem of how to produce a quantum theory of gravity. Most particularly, the observable violations of John Bell's inequality for events at space-like separation are difficult to account for in any theory that has no preferred simultaneity slices in its space-time. So the metaphysician ought not to take the account of space-time provided by relativity as definitive; progress in physics may well demand radical revision of the account of spatiotemporal structure. Still, relativity illustrates how empirical inquiry can lead to the revision of the most seemingly fundamental concepts, even those that were once taken as preconditions for any experience at all.

See also Bell, John, and Bell's Theorem; Eddington, Arthur Stanley; Einstein, Albert; Energy; Galileo Galilei; Geometry; Knowledge, A Priori; Laws, Scientific; Leibniz, Gottfried Wilhelm; Matter; Motion, A Historical Survey; Newton, Isaac; Philosophy of Physics; Quantum Mechanics; Space; Time.

Bibliography

Bell, John S. "How to Teach Special Relativity." In *The Speakable and Unspeakable in Quantum Mechanics*. Cambridge, U.K.: Cambridge University Press, 1987.

Einstein, Albert. "On The Electrodynamics of Moving Bodies." In *The Principle of Relativity*. New York: Dover, 1952.

Einstein, Albert. "The Foundations of General Relaivity." *Annalen der Physik* 49 (1916).

Geroch, Robert P. *General Relativity from A to B*. Chicago: University of Chicago Press, 1978.

Maudlin, Tim. *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Oxford: Basil Blackwell, 1994.

Misner, Charles, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973.

Taylor, E. F., and J. A. Wheeler. *Spacetime Physics: Introduction to Special Relativity*. New York: W. H. Freeman, 1992.

Tim Maudlin (2005)

RELEVANCE (RELEVANT) LOGICS

The conditional, "if … then …" (→) has been a contentious topic throughout the history of Western logic, and numerous accounts of its behavior have been proposed. One recurrent account (usually called the *material conditional*) is that $A \rightarrow B$ is true just if the antecedent, A, is false or the consequent, B, is true. This account was built into the logic of Frege and Russell, and so came to assume orthodoxy throughout much of the twentieth century (at least where there are no subjunctive moods in the antecedent or consequent). The account has obvious problems, however. It entails, for example, that both of the following are true—which they do not appear to be: "If Melbourne is the capital of Australia, it is in China" (false antecedent), "If there is life on another planet, Canberra is the capital of Australia" (true consequent).

It is natural to suppose that in a true conditional the antecedent must be relevant to the consequent in some way. This idea is packed into the contemporary definition of a relevant logic. A propositional logic is a relevant/relevance (both words are used) logic just if whenever $A \rightarrow B$ is a logical truth A and B share a propositional parameter (variable). (A quantifier logic is relevant if its propositional part is.)

Relevant logics can be of several different kinds. However, one has come to dominate current work in the area. This is the Anderson/Belnap tradition. Axiomatizations of logics (or fragments of logics) of this kind were proposed by Ivan Orlov (1928), Alonzo Church (1951), and Wilhelm Ackermann (1956). But the subject took off with the work of the Pittsburgh school of Alan Anderson and Nuel Belnap in the 1960s and 1970s. Probably the most important system of relevant logic developed by the school was the logic *R* (though Anderson and Belnap themselves preferred the system *E*). This contained most of the intuitively correct principles concerning the conditional, but not "paradoxes" such as $(A \& \neg A) \rightarrow B$ and $A \rightarrow (B \rightarrow B)$.

Semantics of various kinds for relevant logics were produced about ten years later by, among others, J. Michael Dunn, Alasdair Urquhart, and Kit Fine. But perhaps the most versatile semantics for relevant logics are the world-semantics developed by the Canberra school of Richard Sylvan (né Routley) and Robert Meyer (who had also been a member of the Pittsburgh school).

The world-semantics of relevant logics may be thought of as extending the possible-world semantics of modal logic by adding a class of logically impossible worlds—though validity is defined in terms of truth-preservation at just the possible worlds. (This comes out most clearly in the simplified form of the semantics, as later developed by Graham Priest, Sylvan, and Greg Restall.) At a possible world, w, the truth conditions for → are the same as those for the strict conditional in the modal logic *S5*:

> $A \rightarrow B$ is true at w iff for all worlds, x (possible and impossible), when A is true at x, B is true at x.

At an impossible world, logical truths—for example, of the form $B \rightarrow B$—may fail. This is achieved by giving the truth conditions of → at such a world, w, in terms of a ternary relation, $\boldsymbol{R}$:

> $A \rightarrow B$ is true at w iff for all worlds x, y, such that $\boldsymbol{R}wxy$, when A is true at x, B is true at y.

These semantics give the base member of the family of logics, *B*. Other logics in the same family may be obtained by adding constraints on the relation $\boldsymbol{R}$. The Anderson/Belnap logic, *R*, is one requiring a number of such constraints. At the time of writing, the nature of $\boldsymbol{R}$, and so of plausible constraints on it, are still contentious issues.

Another important feature of the semantics of relevant logics is their handling of negation. If $(A \& \neg A) \rightarrow B$ is not to be a logical truth, there must be worlds at which $A \& \neg A$ holds (bringing out the connection between relevant logic and paraconsistent logic). This may be achieved in (at least) two ways. In the first (due originally to Dunn), formulas may take the values *true* and *false* independently (and so may take both or neither). The truth/falsity conditions for negation at a world, w, are then:

> $\neg A$ is true at w iff A is false at w
> $\neg A$ is false at w iff A is true at w

If A is both true and false at w, so is $\neg A$. So (given the natural semantics for &) $A \& \neg A$ is true (and false) at w.

The second way to handle negation is to treat truth and falsity as usual, but to use the "Routley *"—invented by Valerie Routley (later Plumwood) and Sylvan. For each world, w, there is a world w^* (usually taken to satisfy the condition that $w = w^{**}$.) The truth conditions for negation are:

> $\neg A$ is true at w iff A is false at w^*

If w is w^*, then exactly one of A and $\neg A$ holds at w. But if w is distinct from w^*, and A is true at w and false at w^*, then $A \& \neg A$ is true at w. Again, at the time of writing, the philosophical meaning of * is still a contentious issue.

However the semantics of negation is handled, there will be worlds where A and $\neg A$ hold; and so, assuming the standard behavior of disjunction, where $\neg A \vee B$ holds, for arbitrary B. It follows that the disjunctive syllogism (A, $\neg A \vee B \vdash B$) is invalid. This is significant because it shows that the ramifications of relevant logic spread much wider than may have been thought. In particular, the syllogism does not seem inherently dubious in the same way that the paradoxes of the material conditional are. The invalidity of the syllogism has therefore occasioned much of the criticism attracted by relevant logic. Defenders of relevant logic have replied in various ways.

Philosophical critiques aside, relevant logics have turned out to have a number of interesting mathematical properties. For example, *R* and some of the other stronger logics (though not the weaker ones) have the unusual property (for a propositional logic) of being undecidable (as shown by Urquhart). Relevant logics are intimately related with algebraic structures called De Morgan lattices, and can also be shown to fit in to the more general class of substructural logics.

See also Logic, Non-Classical; Modal Logic; Paraconsistent Logics.

Bibliography

Anderson, Alan, and Nuel Belnap. *Entailment: The Logic of Relevance and Necessity*. Vol. 1. Princeton, NJ: Princeton University Press, 1975.

Anderson, Alan, Nuel Belnap, and J. Michael Dunn. *Entailment: The Logic of Relevance and Necessity*. Vol. 2. Princeton, NJ: Princeton University Press, 1992.

Dunn, J. Michael. "Relevance Logic and Entailment." In *Handbook of Philosophical Logic*. Vol 3., 117–224, edited by Dov M. Gabbay and Franz Guenthner. Dordrecht, Holland: Kluwer, 1986. The article is revised, with Greg Restall as a co-author, as "Relevant Logic" in the second edition of the *Handbook*. Vol. 6, 1–128. Dordrecht, Holland: Kluwer, 2002.

Mares, Edwin. *Relevant Logic: A Philosophical Interpretation*. Cambridge, U.K.: Cambridge University Press, 2004.

Mares, Edwin D., and Robert K. Meyer. "Relevant Logic." In *The Blackwell Guide to Philosophical Logic*, edited by L. Goble, 280–308. Oxford: Blackwells, 2001.

Priest, Graham. *Introduction to Non-Classical Logic*. Chaps. 8–10. Cambridge, U.K.: Cambridge University Press, 2001.

Read, Stephen. *Relevant Logic: A Philosophical Examination of Inference*. Oxford: Blackwells, 1988.

Restall, Greg. *An Introduction to Sub-Structural Logics*. London: Routledge, 2000.

Routley, Richard, Val Plumwood, Robert K. Meyer, and Ross Brady. *Relevant Logics and Their Rivals*. Vol 1. Atascadero, CA: Ridgeview, 1982. Vol. 2, edited by Ross Brady, Aldershot, U.K.: Ashgate, 2003.

Graham Priest (2005)

RELEVANT ALTERNATIVES

To know a proposition, is it necessary that one is able to rule out every possibility of error associated with that proposition? Notoriously, infallibilism about knowledge—as defended, for example, in early work by Peter Unger (1975)—demands just this and argues on this basis for the skeptical conclusion that knowledge is rarely, if ever, possessed. Intuitively, however, the answer to this question is "no," in that in everyday life we only demand that knowers rule out those error-possibilities that are in some sense relevant. For example, to know that the bird before me is a goldfinch, I may be required to be able to rule out that it is not some other bird that could be in the area just now, like a jackdaw, but we would not normally demand (at least not without special reasons) that I be able to rule out the possibility that it is not a mechanical goldfinch made up to be an exact replica of the real thing.

If this line of thought is right, then this prompts a relevant alternatives (RA) theory of knowledge that demands that one only needs to be able to rule out all relevant error-possibilities in order to know, not that one is able to rule out all error-possibilities, even irrelevant ones. (A similar view could be applied to other epistemic notions, like warrant or justification. For simplicity, the focus here is on knowledge.) Such a position would thus be a form of fallibilism, which is directly opposed to infallibilism and which thereby counters those versions of skepticism that are based on infallibilist considerations. The task at hand for the RA theorist is to offer a principled account of what makes an alternative relevant.

RELEVANT ALTERNATIVES AND SENSITIVITY

One can find the beginnings of an RA theory of knowledge in the writings of such figures as Ludwig Wittgenstein and John Austin. The first worked out versions of an RA theory, however, can be found in the works of Fred Dretske (1970) and Robert Nozick (1981), who primarily understand knowledge in terms of the possession of beliefs that are sensitive to the truth in the following manner:

> *Sensitivity*
>
> An agent, S, has a sensitive belief in a true contingent proposition, p, if and only if, in the nearest possible worlds in which p is not true, S no longer believes p.

To illustrate this, consider again the example of the goldfinch discussed earlier. Given that the actual world is

roughly as we take it to be, gaining a sensitive belief in the proposition, P, that there is a goldfinch before one is relatively straightforward. All one needs is a true belief in this regard and, in the nearest possible worlds where P is no longer true—where, for example, the goldfinch has flown away leaving behind just an empty branch—one no longer believes that there is a goldfinch there, as presumably one does not. Notice that the relevant possible worlds here are limited and concern error-possibilities (e.g., that there is nothing at all on the branch rather than a goldfinch), which are easy to rule out. A theory of knowledge that treats sensitivity as the key requirement on the acquisition of knowledge is thus in a good position to capture the intuition that ordinarily we do not demand that agents are able to eliminate all possibilities of error before we count them as possessing knowledge.

Interestingly, however, the sensitivity-based approach does treat far-off possible worlds, and thus far-fetched error-possibilities, as sometimes being relevant to the possession of knowledge. Consider, for example, the hypothesis Q, that there is a mechanical goldfinch before one, constructed in such a way as to be indistinguishable to the naked eye from the real thing. When one is faced with what seems to be a goldfinch (and circumstances are, apparently, entirely normal), does one know not-Q? According to the sensitivity-based account of knowledge, this is unlikely because it is difficult to have a sensitive belief in not-Q. After all, to have a sensitive belief in this proposition it would be necessary to have a belief that was not only true in the actual world, but that was also no longer held in the nearest possible worlds in which not-Q is false—that is, those worlds in which Q is true, where one is at present looking at a mechanical goldfinch. The problem is, of course, that, ex hypothesi, one would continue to believe that one is looking at a real goldfinch even when one is faced with a mechanical goldfinch, at least unless one conducted special tests (such as capturing the "creature" and cutting it open). So while knowing P is relatively easy, knowing not-Q is hard. And notice that the reason this is the case is because the range of possible worlds, and thus the range of error-possibilities, that is relevant to the determination of one's knowledge is different in each case.

RELEVANT ALTERNATIVES AND NONCLOSURE

On the face of it, this rendering of the RA theory seems to capture our pretheoretical intuition that in normal circumstances we ought to be able to know that we are looking at a goldfinch even though we are unable to rule out (i.e., know to be false) the hypothesis that we are looking at a mechanical goldfinch. Nevertheless, this view does have a counterintuitive result, one that both Dretske and Nozick are prepared to accept. This is that the highly intuitive principle that knowledge is "closed" under known entailment ("closure") has to be rejected. We can roughly formulate closure as follows:

> *Closure for Knowledge*
>
> If an agent, S, knows a proposition, p, and S knows that p entails a second proposition, q, then S also knows q.

For example, if one knows P, then given that one also knows that P entails not-Q (as surely one does), it follows from closure that one must know not-Q. Conversely, of course, if one fails to know the latter proposition, which is what the sensitivity-based approach predicts, then one fails to know the former.

Closure is highly intuitive and yet, as we have just seen, if it holds it would appear to license a restricted form of infallibilism. For although closure does not demand that it is a precondition on knowledge possession that one is able to rule out all possibilities of error, it does demand that one is able to rule out (i.e., know to be false) all those error-possibilities that are known to be inconsistent with what one knows, and this set of error-possibilities, while smaller, is large enough. This point is important, since if the appeal of infallibilism rests on the appeal of closure, then the view is on far stronger ground that one might have initially supposed because of the obvious appeal of the closure principle.

Nevertheless, Dretske and Nozick argue that recognizing that sensitivity is a necessary condition for knowledge highlights why this principle must go, since there are clearly cases, such as the goldfinch example, where one knows one proposition (and thus has a sensitive belief in this proposition) and knows that this proposition entails a second proposition, and yet one lacks a sensitive belief in the entailed proposition and thus fails to know it.

RELEVANT ALTERNATIVES AND CONTEXTUALISM

Although the sensitivity-based proposal has been influential, it does face the problem that it denies the highly intuitive closure principle for knowledge, and this has led some commentators to try to see if there is a way of accommodating the general intuition behind the RA theory in a way that preserves this principle. One of the guiding considerations behind views that try to offer an RA thesis that is consistent with closure is that the Dretske-

Nozick treatment seems to incorporate the idea that closure fails because while sometimes knowledge is hard to attain, sometimes attaining it is relatively straightforward. This tends to suggest that an alternative way of approaching the issue could be to regard knowledge as in some sense context-sensitive, so that one knows both of the target propositions in the closure-based inference relative to one set of epistemic standards (the less demanding ones), but knows neither of them relative to another set of epistemic standards (the more demanding ones). We would thus get a view that incorporates a reading of the RA intuition—because it would remain that not every error-possibility is always relevant to the possession of knowledge—but which was also consistent with closure. This view—known as contextualism about knowledge—is hinted at in an early response to Dretske's denial of closure written by Gail Stine (1976), and has been developed by Stewart Cohen (1991), Keith DeRose (1995), and David Lewis (1996).

Consider again the goldfinch example. On the Dretske-Nozick view the class of possible worlds, and thus the class of error-possibilities, that is relevant to the determination of knowledge can differ depending on the content of the proposition at issue, which is why in this case the agent comes out as knowing P while failing to know not-Q, despite knowing that the former entails the latter. The reason for this is that when it comes to knowing P only nearby possible worlds are relevant, whereas knowing not-Q brings in farther out possible worlds. Suppose instead, however, that one simply treated the class of possible worlds as fixed in each context, so that the epistemic status of all beliefs—whatever their content—were in that context evaluated relative to those possible worlds. In normal contexts, then, only nearby possible worlds would be relevant, while in more demanding contexts far-off possible worlds would become relevant. This way of understanding knowledge would mean dropping sensitivity as a requirement on knowledge, of course, since there may be no nearby possible worlds in which the target proposition is false (this is, indeed, what we would expect to be the case when it comes to the hypothesis that one is at present looking at a mechanical goldfinch). Nevertheless, the guiding thought here is that so long as the agent's belief matches the truth in the relevant possible worlds—that is, where the agent believes that proposition, it is true; and where the proposition is not true, the agent does not believe it—then the agent's belief will be in the market to be counted as an instance of knowledge.

By contextualist lights, then, in contexts where the epistemic standards are low (and thus only nearby possible worlds count as relevant) one will tend to know both P and not-Q, since even one's belief in not-Q will tend to match the truth (i.e., one believes it in all nearby possible worlds and it is true in all nearby possible worlds). In contrast, in contexts where the epistemic standards are more demanding, and thus where farther out possible worlds become relevant, it will now no longer be the case that one will tend to know either of these propositions. After all, there will be possible worlds, such as the far-off world in which there is a sophisticated plot to deceive people about the presence of goldfinches, in which one's beliefs in P and not-Q no longer match the truth. Thus, as long as one consistently sticks to a specific epistemic standard then this construal of the RA intuition is not in conflict with closure since, depending on the context at issue, either one has knowledge of both of the target propositions or one has knowledge of neither of them.

RELEVANT ALTERNATIVES AND SAFETY

In more recent work, however, a third rendering of the RA thesis has come to the fore, one that is neither contextualist nor results in the denial of closure. This position—defended, for example, by Ernest Sosa (1999)—holds that far-off possible worlds are always irrelevant to knowledge, whatever the content of the target proposition or the context at issue. Accordingly, one is able to know, for example, both P and not-Q, whatever the context.

This view tends to hold that the key condition that a belief must meet if it is to count as knowledge is that it be safe. Safety can be roughly formulated as follows:

> *Safety*
>
> An agent, S, has a safe belief in a true contingent proposition, p, if and only if, in all nearby possible worlds in which S believes p, p is true.

Notice that contextualists will have to appeal to something like safety to explain how agents can know a proposition like not-Q in epistemically undemanding contexts where there are no nearby worlds in which what is believed is false. The point will be that while such beliefs are not sensitive, since there is no relevant Q-world for them to be sensitive to, they are safe, in that the agent's belief in not-Q is always true across the relevant possible worlds. What is important about safety for our purposes is that it simply specifies the class of possible worlds that is relevant and leaves the matter at that—there is no room

here for a shift in context that would in turn alter the class of possible worlds, and thus the class of error-possibilities, that is relevant to the determination of knowledge. Accordingly, it will not matter which context one is in. Just so long as one's beliefs that P and not-Q are both safe—as presumably they will be—then one is in a position to know both of these propositions and thus there is no tension with closure.

CONCLUDING REMARKS

There are thus three competing conceptions of the RA intuition in the literature. The first view treats relevance as being determined by the content of the proposition known, and as a result maintains that the closure principle for knowledge fails. The second view treats relevance as being determined by context, and thereby retains closure. Finally, the third view also retains closure, but does so by maintaining an invariant standard of relevance, regardless of the content of the target proposition or of the context at issue.

See also Austin, John; Contextualism; Dretske, Fred; Lewis, David; Nozick, Robert; Propositions; Sosa, Ernest; Wittgenstein, Ludwig Josef Johann.

Bibliography

Austin, J. L. "Other Minds." In *Philosophical Papers*. Oxford, U.K.: Clarendon Press, 1961.

Cohen, Stewart. "Skepticism, Relevance, and Relativity." In *Dretske and His Critics*, edited by Brian P. McLaughlin. Cambridge, MA: Basil Blackwell, 1991.

DeRose, K. "Solving the Sceptical Problem." *Philosophical Review*104 (1995): 1–52.

Dretske, Fred. "Epistemic Operators." *Journal of Philosophy* 67 (1970): 1007–1023.

Lewis, D. "Elusive Knowledge." *Australian Journal of Philosophy* 74 (1996): 549–567.

Nozick, Robert. *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.

Sosa, Ernest. "How to Defeat Opposition to Moore." *Philosophical Perspectives* 13 (1999): 141–154.

Stine, Gail. "Skepticism, Relevant Alternatives, and Deductive Closure." *Philosophical Studies* 29 (1976): 249–261.

Unger, Peter. *Ignorance: A Case for Scepticism*. Oxford, U.K.: Clarendon Press, 1975.

Wittgenstein, Ludwig. *On Certainty*. Translated by Denis Paul and G. E. M. Anscombe, edited by G. E. M. Anscombe and G. H. von Wright. Oxford, U.K.: Blackwell, 1969.

Duncan Pritchard (2005)

RELIABILISM

Reliabilism is an approach to the analysis of either knowledge or justified belief that makes, in some way or another, the reliability of belief-producing faculties or processes the key notion of epistemic assessment. An early version of a reliabilist theory of *knowledge* was proposed by David M. Armstrong (1973), who thought of knowledge in terms of a reliable thermometer that accurately indicates the correct temperature. A (noninferential) true belief amounts to knowledge, according to Armstrong, if its properties nomically (i.e., via the laws of nature) guarantee its truth. Closely related theories conceive of knowledge as resulting from a counterfactual guarantee of truth. For instance, according to Robert Nozick (1981), knowledge comes about when a subject's belief that p tracks the truth of p, which it does (focusing just on the core of Nozick's theory) if the following condition is met: S would not believe that p if p were false. Alternatively, Fred Dretske (1971, 1981) suggests that a true belief counts as knowledge if the subject possesses a conclusive reason for p. According to this proposal, S knows that p on the basis of a reason, R, if R would not be true unless p were true.

RELIABILIST THEORIES OF JUSTIFIED BELIEF

Turning to reliabilist theories of *justified belief*, there are two main versions to consider: reliable indicator theories, and process reliabilism. A version of the former was developed by Marshall Swain (1979): What makes S's belief that p justified is the belief's being a reliable indication (conceived of in terms of objective probability) of p's truth. In contrast, process reliabilism (reliabilism$_{pr}$), as advocated by Alvin Goldman (1979), focuses not on the belief itself, but its causal history, or the cognitive process by which it was formed. The basic proposal is that a belief that p is justified if and only if the process by which it was formed is (sufficiently) reliable. On the one hand, perception, memory, and introspection are reliable cognitive processes, typically resulting in justified beliefs and indeed knowledge. Hasty generalization, wild hunches, and wishful thinking, on the other hand, are unreliable processes and invariably produce unjustified beliefs and prevent the formation of knowledge. Moving beyond this initial characterization, an account of reliabilism$_{pr}$ requires refinement as to the question of precisely what determines the reliability of a cognitive process. According to one approach, a process's reliability is fixed by the truth ratio of its actual doxastic output: The greater the ratio of true over false beliefs within the set of beliefs that

make up a process's actual track record, the higher its degree of reliability. As an alternative to such a *track record* conception, William Alston (1995) recommends a *propensity* construal, according to which a process's reliability is determined by, not its actual track record, but what the truth ratio of its output *would be* in an appropriate range of cases.

According to reliabilism$_{pr}$, as well as other versions of reliabilism, whether S's belief that p is justified, or an instance of knowledge, does not depend—at least not exclusively—on S's evidence relative to p. Reliabilist theories of knowledge and justification must thus be viewed as intended alternatives to evidentialist theories (Chisholm 1989, Conee and Feldman 1985). Nevertheless, in Goldman's defense of reliabilism$_{pr}$, evidentialist considerations do not completely drop out of the picture. Suppose the following: (i) S forms the belief that p via process C; (ii) C is in fact reliable; and (iii) on the basis of professional testimony, S has reason to believe that C is unreliable. Intuitively, S's belief is not justified. Cases such as that make it doubtful that origination in a reliable process is by itself sufficient for justification. Accordingly, Goldman (1986) supplements his account with a nonundermining clause to the effect that S must not believe, or be in possession of evidence supporting the belief, that the relevant process is unreliable.

In its most radical and challenging manifestation, reliabilism$_{pr}$ asserts that reliable belief formation is both necessary and sufficient for a belief's justification. Alston (1989) defends a moderate version, according to which reliable belief formation is merely necessary for justification. An approach that may be viewed as an alternative to reliabilism$_{pr}$—virtue epistemology—shifts the focus away from reliable processes to reliable faculties or cognitive virtues, giving rise to the thought that justification and knowledge may be conceived of as resulting from the employment of virtuous faculties (Sosa 1991). A related approach is advocated by Alvin Plantinga (1993). Plantinga's view is that knowledge is generated by properly functioning faculties, where a faculty's proper functioning requires reliability, in addition to adequate design and an orientation towards truth and the avoidance of falsehood.

THE THREE MAJOR PROBLEMS WITH RELIABILISM

Reliabilism$_{pr}$ is confronted with three major problems. The first of these raises the issue of whether production via reliable processes is necessary for justification (Cohen 1984, Ginet 1985). Consider the beliefs of a subject who is deceived by an evil demon. Because the evidential situation of an evil demon victim is not relevantly different from that of a normal person who has (presumably) by and large justified beliefs, it is commonly agreed that such a victim's beliefs are, just like a normal person's, by and large justified. Alas, the victim's beliefs are, unlike those of a normal person, massively false. The challenge for reliabilism$_{pr}$, then, is this: The beliefs of an evil demon victim are justified although they are the result of *unreliable* cognitive processes.

There are four ways in which reliabilists can respond. First, they can—implausibly—deny that evil demon victims have justified beliefs. Second, they can deny the relevance of the counterexample, as Alston (1995) does. He argues that the cognitive processes to consider are to be restricted to those that would yield a high truth ratio over a wide range of situations *of the kind one typically encounters*. Third, advocates of reliabilism$_{pr}$ can try to accommodate the counterexample by modifying reliabilism$_{pr}$. For example, Goldman (1986) introduces normal worlds reliabilism, the basic idea of which is that a belief (in any possible world) is justified if and only if the process by which it was formed is reliable in normal worlds: worlds that correspond to our general beliefs about the actual world. Because perception is reliable in normal worlds, normal worlds reliabilism arguably yields the result that the perceptual beliefs of evil demon victims are justified.

Whether this response succeeds depends on whether the processes by which an evil demon victim forms beliefs can properly be characterized as perceptual, memorial, and so on. After all, at the end of the causal chains from which a deceived subject's beliefs originate, there is the evil demon: not exactly the kind of creature one finds in any normal worlds. So if the belief-generating processes are considered in their entirety, it is hard to tell what their truth ratios in normal worlds would be because such processes are not part of any normal worlds to begin with. Fourth, defenders of reliabilism$_{pr}$ can introduce different concepts of justification. For example, Goldman (1988) distinguishes between strong and weak justification. According to this proposal, the beliefs of evil demon victims are justified in the weak sense, whereas reliabilism is intended to be an analysis of strong justification, the kind of justification that is needed if a true belief is to count as an instance of knowledge.

THE SECOND PROBLEM. The second problem raises the issue of whether origination in a reliable process is sufficient for justification (BonJour 1985, ch. 4). Suppose the following: (i) Norman's belief that p is the result of clair-

voyance; (ii) Norman's faculty of clairvoyance is reliable; and (iii) Norman has no reason to believe that his belief that p originated in, or is sustained by, a reliable faculty. Reliabilism$_{pr}$ implies that Norman is justified in believing that p, whereas intuitively he seems to be unjustified. In response, reliabilists can deny the intuition underlying the objection—that is, insist that Norman's belief is justified. Alternatively, they can, once again, attempt to accommodate the example by devising a suitable modification. For example, they might consider *ignoring the absence of evidential support* an unreliable cognitive process, and suggest that beliefs whose causal origin includes that process are unjustified. The problem with this suggestion is that it threatens to rob reliabilism$_{pr}$ of its identity by letting it collapse into a disguised version of evidentialism.

THE THIRD PROBLEM. The third problem, known as the "generality problem," raises the issue of how to individuate the cognitive processes the reliability of which is supposed to determine whether a belief is, or fails to be, justified (Feldman 1985). Suppose a person sees, and thus believes, that the cat is lying on the couch. The process by which this belief is formed could plausibly be classified as perception. More specifically, it could be viewed as an instance of visual perception. Further specification yields further choices: visual perception at a distance of (say) eight feet; visual perception of a medium sized object at a distance of eight feet; visual perception of a medium sized object at a distance of eight feet under daylight illumination. In general terms, the point is that a particular *token* of a cognitive process instantiates many different process *types*. Some of them are reliable, some of them are not. Whether a belief whose justificational status is at issue comes out justified or unjustified will depend on which process type is made the basis of the assessment. The challenge advocates of reliabilism$_{pr}$ face is to give a principled account of how to select the right process type.

Consider perception, an obviously reliable process. But not all perceptual beliefs are justified. Nor are, for that matter, all memorial beliefs, or all visual beliefs, or all auditory beliefs. Sometimes, perceivers fail to take into account undermining evidence, and then beliefs produced by reliable processes fail to be justified. Hence individuating process types using broad categories such as perception, vision, or memory will often yield the wrong results. More specification is clearly required. But too much specification also yields the wrong results. At the extreme end of specification are process types instantiated by one and only one process token; one then encounters what Richard Feldman (1985) calls the "single case problem." If such a token results in a false belief, the result will be total unreliability, for the process type's output is false in all cases. If the process token in question results in a true belief, the result will be perfect reliability, for the process type's output is true in all cases. In the former case, the belief will be unjustified no matter what; in the latter case, it will be justified no matter what. This will result in clearly counterintuitive results. Suppose S is a paranoid schizophrenic. While riding on the bus, S's paranoia leads him to believe that the bus will blow up. Suppose further that that is in fact true. Let P* stand for a process type described in such a way that P* has *one and only one* instantiation: the process token that caused S's belief about the bus. (Such a description can easily be achieved by making reference to properties that uniquely pick out S and the circumstances under which S formed that belief.) Because P* is a (perfectly) reliable process, reliabilism$_{pr}$ implies—implausibly—that S is justified in believing that the bus will blow up. The problem for advocates of reliabilism$_{pr}$ is this: On the basis of what principled grounds can they claim that P* is not the process type the reliability of which determines the justificational status of S's belief?

Alston (1995) claims that there is a solution to the generality problem. Regarding the single case problem, he suggests that it does not arise when the track-record conception of reliability is replaced with a propensity conception. Consider again the belief of the paranoid subject that the bus will blow up. On the propensity conception of generality, the process type in question, having precisely one instantiation that led in fact to a true belief, nevertheless counts as unreliable when it is taken into account that beliefs resulting from paranoia tend to be false. Moreover, Alston argues that there are objective, psychological facts of the matter that determine, for each process token leading to a particular belief, which process type this token instantiates. According to Alston, every process token instantiates an input-output function. Each time a belief is formed, Alston claims, there is one and only one input-output pair that is psychological real. However, Earl Conee and Richard Feldman (1998) respond that, even if one accepts the constraints Alston places on the selection of legitimate process types, there will still be a wide range of process types going from narrow to broad characterizations. As a result, there will be cases of belief formation for which reliabilism$_{pr}$ will not yield a determinable implication about the belief's justificational status.

INTERNALIST AND EXTERNALIST THEORIES

It is common practice to distinguish between *internalist* and *externalist* theories of knowledge and justification. According to internalism about justification, the factors that determine a belief's justificational status (call them "J-factors") must be internal to the subject's mind. Typically, such internality is defined epistemically: An item x is internal to S if and only if S can, merely by reflecting on it, determine whether x is present or absent. Internalist theories of justification, then, usually demand that J-factors must be such that their presence or absence is always on reflection recognizable by the subject. As a result of this constraint, justification itself turns into something the presence or absence of which can be recognized upon reflection. According to externalism about justification, J-factors are not subject to any internality constraint. Reliabilism about justification, in its various manifestations, is an externalist theory, for the obvious reason that subjects are not always in a position to determine, solely on the basis of reflection, whether their beliefs are the result of reliable cognitive processes. Consider, again, the victim of an evil demon. Upon reflection, such a victim will think that her perceptual beliefs originate in reliable cognitive processes, when in fact they do not. However, the classification of reliabilism as an externalist theory should not be misunderstood to mean that, according to reliabilism, the reliability of our cognitive processes is completely beyond our ken. To the contrary, there is no reason why reliabilists should deny that, in typical situations when a subject forms, for example, perceptual or memorial beliefs, it should be knowable to the subject, on the basis of presently available evidence, that the beliefs in question have their origin in reliable processes or faculties.

The internalism/externalism issue presents itself in a different form when the object of the dispute is not justification but knowledge. There is broad agreement that knowledge is not, and indeed cannot be, internal in the way in which, according to some, justification is internal. Suppose the following: (i) S has a body of excellent evidence, E, in support of p; (ii) E is misleading: p is in fact false. Reflecting on whether she knows that p, S will of course conclude, mistakenly, that she does. Clearly, then, whether or not one knows cannot always be determined upon reflection. Thus it is beyond dispute that knowledge is external. Nevertheless, it would not be inaccurate to say that evidentialists defend an internalist conception of knowledge. According to evidentialists, S knows that p only if S has a good reason for p. But whether or not one has a good reason for p is something that is internal to the subject; it is something that can be determined merely by reflecting on one's evidence. Evidentialists, therefore, hold that one of the necessary conditions of knowledge is internal, and thus may be considered internalists about knowledge. Reliabilists, however, are externalists about knowledge, for they typically claim that reliable belief production, suitably qualified, is sufficient for making a true belief an instance of knowledge, thus advocating an account of knowledge without any internalist condition.

The reliabilist, externalist view that the employment of reliable processes or faculties is sufficient for making a true belief thus produced an instance of knowledge that can be supported by citing that very young children and animals possess knowledge, for neither the former nor the latter would seem to be capable of having good reasons in support of their beliefs. The evidentialist, internalist view that one knows only if one possesses a good reason can be defended by pointing out that, upon discovering that a person believes that p without having a good reason for p, we tend to judge that that person does not know that p. Reliabilists, on the one hand, need to come to terms with one kind of fact about our ordinary cognitive practice: people are reluctant to attribute knowledge in the absence of good reasons. Evidentialists, on the other hand, need to come to terms with another kind of fact about our ordinary cognitive practice: people do not hesitate to attribute knowledge to very young children and even such animals as cats and dogs.

CONCLUSION

Even though reliabilist theories are properly classified as externalist, there is no reason in principle why internalists should not acknowledge the relevance, or even fundamental importance of reliability. To begin with, internalists might agree that a true belief counts as knowledge only if it originates in a reliable faculty. Furthermore, Matthias Steup (2004) proposes *internalist reliabilism* as an answer to the question of why sense experience is a source of justification. According to Steup's proposal, what makes sense experience a source of justification is not *de facto* reliability, but *evidence* of reliability. Perception is a source of justification for a subject, S, if and only if S has, on the basis of track record memories, reason to believe that her perceptual faculties are reliable. According to this proposal, perception is a source of justification even in worlds in which it is in fact unreliable, as long as, on the basis of adequate evidence, it appears to be reliable. According to such an approach, issues of reliability lie indeed at the heart of epistemology, for S acquires

knowledge only if S (i) employs a faculty that is in fact reliable, and (ii) possesses evidence of that faculty's reliability.

See also Alston, William P.; Armstrong, David M.; Dretske, Fred; Epistemology; Epistemology, History of; Evidentialism; Goldman, Alvin; Nozick, Robert; Plantinga, Alvin; Virtue Epistemology.

Bibliography

Alston, William. "How to Think about Reliability." *Philosophical Topics* 23 (1995): 1–29.

Alston, William. "An Internalist Externalism." In *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca, NY: Cornell University Press, 1989.

Armstrong, D. M. *Belief, Truth, and Knowledge*. Cambridge, U.K.: Cambridge University Press, 1973.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press, 1985.

Chisholm, Roderick. *The Theory of Knowledge*. Englewood Cliffs, NJ: Prentice Hall, 1989.

Cohen, Stewart. "Justification and Truth." *Philosophical Studies* 46 (1984): 279–95.

Conee, Earl, and Richard Feldman. "Evidentialism." *Philosophical Studies* 48 (1985): 15–44.

Conee, Earl, and Richard Feldman. "The Generality Problem for Reliabilism." *Philosophical Studies* 89 (1998): 1–29.

Dretske, Fred. "Conclusive Reasons." *Australasian Journal of Philosophy* 49 (1971): 1–22.

Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

Feldman, Richard. "Reliability and Justification." *Monist* 68 (1985): 159–74.

Ginet, Carl. "Contra Reliabilism." *Monist* 68 (1985): 175–87.

Goldman, Alvin. "Epistemic Folkways and Scientific Epistemology." In *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press, 1991.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Goldman, Alvin. "Strong and Weak Justification." *Philosophical Perspectives* 2 (1988): 51–69.

Goldman, Alvin. "What Is Justified Belief?" In *Justification and Knowledge*, edited by George Pappas. Dordrecht, Netherlands: Reidel, 1979.

Nozick, Robert. *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.

Plantinga, Alvin. *Warrant: The Current Debate*. Oxford: Oxford University Press, 1993.

Sosa, Ernest. *Knowledge in Perspective*. Cambridge, U.K.: Cambridge University Press, 1991.

Steup, Matthias. "Internalist Reliabilism." *Philosophical Issues* 14 (2004).

Swain, Marshall. *Reasons and Knowledge*. Ithaca, NY: Cornell University Press, 1979.

Matthias Steup (2005)

RELIGION

This entry is not a survey of the various forms that "religion" has taken in human history; rather, it treats the nature of religion as a problem in the philosophy of religion. It will be concerned with attempts to develop an adequate definition of religion, that is, to make explicit the basic features of the concept of religion.

GENERAL DEFINITION AND CHARACTERISTICS

EXAMINATION OF DEFINITIONS. A survey of existing definitions reveals many different interpretations.

> "Religion is the belief in an ever living God, that is, in a Divine Mind and Will ruling the Universe and holding moral relations with mankind." —James Martineau

> "Religion is the recognition that all things are manifestations of a Power which transcends our knowledge."—Herbert Spencer

> "By religion, then, I understand a propitiation or conciliation of powers superior to man which are believed to direct and control the course of Nature and of human life."—J. G. Frazer

> "Religion is rather the attempt to express the complete reality of goodness through every aspect of our being."—F. H. Bradley

> "Religion is ethics heightened, enkindled, lit up by feeling."—Matthew Arnold

> "It seems to me that it [religion] may best be described as an emotion resting on a conviction of a harmony between ourselves and the universe at large."—J. M. E. McTaggart

> "Religion is, in truth, that pure and reverential disposition or frame of mind which we call piety."—C. P. Tiele

> "A man's religion is the expression of his ultimate attitude to the universe, the summed-up meaning and purport of his whole consciousness of things."—Edward Caird

> "To be religious is to effect in some way and in some measure a vital adjustment (however tentative and incomplete) to whatever is reacted to or regarded implicitly or explicitly as worthy of serious and ulterior concern."—Vergilius Ferm

If we take these definitions as attempts to state necessary and sufficient conditions for something to be a religion, it is not difficult to show that none of them is adequate. With respect to necessary conditions, consider Martineau's definition. It is clear that such a belief does not have to be present in a religion. No polytheistic religion recognizes a single divine ruler of the universe; and there are religions, such as Hinayana Buddhism, in which beliefs in personal deities play no role at all. Bradley and Arnold identify religion with morality, but there are primitive societies in which there is no real connection between the ritual system, with its associated beliefs in supernatural beings, and the moral code. The latter is based solely on tribal precedent and is not thought of as either originating with or sanctioned by the gods. If, as would commonly be done, we call the former the religion of the culture, we have a religion without morality. As for McTaggart and Tiele, it seems likely that if we specify "piety" or "feeling of harmony" sufficiently to give them a clear and unambiguous meaning, we will be able to find acknowledged religions in which they do not play an important role. It would seem that we could avoid this only by construing "piety," for example, to cover any state of feeling that arises in connection with religious activities. It does seem plausible to regard some of the definitions as stating necessary conditions, as in Caird and Ferm. However, it is doubtful that these are sufficient conditions. Does any "ultimate attitude" or any "vital adjustment" constitute a religion? As William James points out (*The Varieties of Religious Experience,* Ch. 2), it seems doubtful that a frivolous attitude toward life constitutes a religion, even if it is the fundamental attitude of a given person. And Ferm's overcarefully worded statement would seem to admit any attitude with respect to anything considered important to the ranks of the religious. This would presumably include one's attitude toward one's spouse, toward one's vocation, and, in many cases, toward one's athletic activities. At this point one wonders what has happened to the concept of religion. Many of the definitions are deficient on grounds of both necessity and sufficiency. To return to Martineau, it is quite conceivable that such a belief might be held purely as a speculative hypothesis, without affecting the believer's feelings and attitudes in the way that would be requisite for religious belief. And as for McTaggart, it seems clear that one could from time to time have such a sense of harmony without this being integrated into anything that we would call a religion.

It is noteworthy that most of these definitions stress one aspect or another of religion to the exclusion of others. Thus, Martineau and Spencer represent religion as some sort of belief or other cognitive state; Frazer, as ritual (conceived in a utilitarian fashion); Bradley and Arnold, as a kind of moral attitude and activity; and McTaggart and Tiele as a certain kind of feeling. One might attribute the failings of these definitions to their one-sidedness. One could hardly expect to get an adequate statement of the nature of so complex a phenomenon as religion, essentially involving, as it does, all these forms of human activity by restricting oneself to belief, feeling, ritual, or moral attitude alone. Caird and Ferm escape this particular failing by concentrating on a comprehensive term such as *attitude* or *adjustment,* which itself embraces belief, feeling, and moral attitude. But, as we have seen, these formulations do not come measurably closer to providing a set of necessary and sufficient conditions.

There are other ways of construing definitions of religion. Instead of taking the above statements as attempts to specify features that are common and peculiar to cases of religion, we might take each of them as an attempt to state the *essence* of religion, that central feature in terms of which all religious phenomena are to be understood. This approach to the matter is explicit in the following statements:

> "The essence of religion is a belief in the persistency of value in the world."—Harald Høffding
>
> "The heart of religion, the quest of the ages, is the outreach of man, the social animal, for the values of the satisfying life."—A. E. Haydon
>
> "The essence of religion consists in the feeling of an absolute dependence."—Friedrich Schleiermacher

There are two distinguishable interpretations of claims of this type. They might be interpreted genetically, as accounts of the origin of religion. The claim would then be that what is specified as the essence of religion is the original root from which all phenomena of religion have sprung. Thus, Julian Huxley, like Schleiermacher working with a conception of the essence of religion as a kind of feeling, says, "the essence of religion springs from man's capacity for awe and reverence, that the objects of religion … are in origin and essence those things, events, and ideas which arouse the feeling of sacredness" (*Religion without Revelation,* p. 111). Similarly starting with Høffding's formulation, we might try to show how typical religious doctrines, rites, and sentiments grew out of an original belief in the persistency of value. However, since we know virtually nothing about the prehistoric origins of religion, speculation in this area is almost completely unchecked by data, and it seems impossible to find

any rational basis for choosing between alternative genetic accounts.

However, we might also give a nongenetic interpretation. Saying that the essence of religion is a feeling of absolute dependence, for example, might mean that the full interrelatedness of the various features of religion can be understood only if we view them all in relation to a feeling of absolute dependence. This claim would be independent of any view of the origin of religion. The difficulty with this is that there would seem to be several different features of religion that could be taken as central—such as ritual, a need for reassurance against the terrors of life, or a need to get a satisfactory explanation of the cosmos—and it is illuminating to view the rest of religion as related to each of these. How is one to settle on a unique essence?

CHARACTERISTIC FEATURES OF RELIGION. Despite the fact that none of the definitions specifies a set of characteristics which is present when and only when we have a religion, or gives us a unique essence, it does seem that they contribute to our understanding of the nature of religion. It appears that the presence of any of the features stressed by these definitions will help to make something a religion. We might call such features, listed below, religion-making characteristics.

(1) Belief in supernatural beings (gods).

(2) A distinction between sacred and profane objects.

(3) Ritual acts focused on sacred objects.

(4) A moral code believed to be sanctioned by the gods.

(5) Characteristically religious feelings (awe, sense of mystery, sense of guilt, adoration), which tend to be aroused in the presence of sacred objects and during the practice of ritual, and which are connected in idea with the gods.

(6) Prayer and other forms of communication with gods.

(7) A worldview, or a general picture of the world as a whole and the place of the individual therein. This picture contains some specification of an overall purpose or point of the world and an indication of how the individual fits into it.

(8) A more or less total organization of one's life based on the worldview.

(9) A social group bound together by the above.

Interrelations of characteristics. Religion-making characteristics do not just happen to be associated in religion; they are intimately interconnected in several ways. Some of these connections have been indicated, but there are others. For example, the distinction between sacred and profane objects is based on other factors mentioned. It is not any intrinsic characteristic of a thing that makes it a sacred object; things of every conceivable kind have occupied this position—animals, plants, mountains, rivers, persons, and heavenly bodies. Certain objects are singled out as sacred in a given community because they typically arouse such feelings as awe and a sense of mystery, and thus the members of that community tend to respond to these objects with ritual acts. Again, the emotional reaction to sacred objects may be rationalized by conceiving the object to be the habitation or manifestation of a god. The awe aroused by the wild bull led to its being identified with the wild god of intoxication, Dionysus. The very special impression made by Jesus of Nazareth on certain of his contemporaries was expressed by calling him the Son of God. These examples make it sound as if emotional reactions to sacred objects come first and that these reactions are then explained by positing gods as their causes. But it can also happen the other way round. The acceptance of beliefs about the gods and their earthly habitations can contribute to the evocation of awe and other feelings in the presence of certain objects. The members of a religious community are taught to hold certain objects in awe by being taught various doctrines about the gods. Thus, Christians are taught to regard the cross and the consecrated bread and wine with reverence by being told of the Crucifixion and the Last Supper.

A similar reciprocal relationship holds between ritual and doctrine. A doctrine can be introduced as the justification of an already established ritual. Thus, the myth of Proserpine being carried off to the underworld and remaining there half the year seems to have been introduced as an explanation of a preexisting magical fertility cult, in which an ear of grain, perhaps called the corn maiden, was buried in the fall and raised sprouting in the spring. On the other hand, changes in doctrine can engender, modify, or abolish rituals. Beliefs about the divine status of Jesus Christ played an important role in shaping the Christmas festival.

Definition in terms of characteristics. If it is true that the religion-making characteristics neither singly nor in combination constitute tight necessary and sufficient conditions for something being a religion, and yet that each of them contributes to making something a religion, then it must be that they are related in some looser way to the application of the term. Perhaps the best way to put it

is this. When enough of these characteristics are present to a sufficient degree, we have a religion. It seems that, given the actual use of the term *religion,* this is as precise as we can be. If we tried to say something like "for a religion to exist, there must be the first two plus any three others," or "for a religion to exist, any four of these characteristics must be present," we would be introducing a degree of precision not to be found in the concept of religion actually in use.

Another way of putting the matter is this. There are cultural phenomena that embody all of these characteristics to a marked degree. They are the ideally clear paradigm cases of religion, such as Roman Catholicism, Orthodox Judaism, and Orphism. These are the cases to which the term *religion* applies most certainly and unmistakably. However, there can be a variety of cases that differ from the paradigm in different ways and to different degrees, by one or another of the religion-making characteristics dropping out more or less. For example, ritual can be sharply de-emphasized, and with it the demarcation of certain objects as sacred, as in Protestantism; it can even disappear altogether, as with the Quakers. Beliefs in supernatural beings can be whittled away to nothing, as in certain forms of Unitarianism, or may never be present, as in certain forms of Buddhism. And, as mentioned earlier, in certain primitive societies morality has no close connection with the cultic system. As more of the religion-making characteristics drop out, either partially or completely, we feel less secure about applying the term *religion,* and there will be less unanimity in the language community with respect to the application of the term. However, there do not seem to be points along these various dimensions of deviations that serve as a sharp demarcation of religion from nonreligion. It is simply that we encounter less and less obvious cases of religion as we move from, for example, Roman Catholicism through Unitarianism, humanism, and Hinayana Buddhism to communism. Thus, the best way to explain the concept of religion is to elaborate in detail the relevant features of an ideally clear case of religion and then indicate the respects in which less clear cases can differ from this, without hoping to find any sharp line dividing religion from nonreligion. (Cf. Ludwig Wittgenstein's notion of "family-resemblances" among the things to which a term applies.)

An adequate definition of religion should throw light on the sorts of disputes and perplexities that typically produce a need to define religion, such as disputes over whether communism is a religion, and whether devotion to science can be called a man's religion. So long as we are dealing with definitions of the simplistic type that we have criticized, these problems are not illuminated. Each party to the dispute will appeal to a definition suited to the position he is defending, and since none of these definitions is wholly adequate, there is an irreducible plurality of not wholly inadequate definitions to be used for this purpose. Person *A,* who claims that communism is a religion, will give, for instance, Caird's statement as his definition of religion, and person *B,* who denies this, will choose Martineau's. Obviously, the position of each is upheld by his chosen definition. Hence, it would seem that the only way to settle the dispute is to determine which is the correct definition. However, we have seen that this gets us nowhere; no such definition is wholly adequate.

At this point there is a temptation to brand the dispute purely verbal, a reflection of different senses attached to the word *religion.* It may seem that the disagreement can be dissolved by persuading all parties to use the word in the same sense. But this is a superficial reaction that does not adequately bring out how much the parties to the dispute have in common. In fact, Martineau and Caird represent two contrasting emphases within a common framework. Suppose that *A* and *B* begin with the same paradigm, orthodox Protestant Christianity. But *A* gives greatest weight to the moral-orientation–emotion elements in this paradigm. As long as anything strongly manifests these elements, as long as it serves as a system of life orientation for the individual who is bound to it by strong emotional ties, he will call it a religion. *B,* on the other hand, gives greatest weight to the belief in a personal God and the complex of emotions, ritual, and devotional acts that is bound up with that belief. Thus, although they have basically the same concept of religion, they will diverge in their application of the term at certain points. Once we realize that this is the true situation, we can state the problem in a more tractable form. We can enumerate the religion-making characteristics and determine which of them communism has and in what degree. Then we can proceed to the heart of the dispute—the relative importance of these characteristics. Insofar as there is a real issue between *A* and *B,* once both are in possession of all the relevant facts, it is whether communism is similar to clear cases of religion in the most important respects, that is, whether the respects in which it is like Protestant Christianity are more important than those in which it is different.

TYPES OF RELIGION

In the case of so complex a concept as religion, it is desirable to supplement the very general portrayal of basic fea-

tures with some indications of the varying emphases placed on them in different religions. To do this, we must develop a classificatory scheme.

William James has reminded us that in every religion there is some sort of awareness of what is called divine and some sort of response to this divinity. This being the case, a very fruitful way of classifying religions is to ask in the case of each: "Where is the divine (the object of religious responses) primarily sought and located, and what sort of response is primarily made to it?" In answering these questions for a given religion, the religion-making features most stressed in that religion will also come to light. According to this principle of division, religions fall into three major groups: sacramental, prophetic, and mystical.

LOCATION OF THE DIVINE. In sacramental religion the divine is sought chiefly in things—inanimate physical things like pieces of wood (relics of saints, statues, crosses), food and drink (bread and wine, baptismal water), living things (the totem animal of the group, the sacred cow, the sacred tree), processes (the movements of the sacred dance). This does not mean that the thing itself is responded to as divine, although this can happen in very primitive forms of sacramental religion, called fetishism. Usually the sacred thing is conceived to be the habitation or manifestation of some god or spirit. Thus, the ancient Hebrews treated the elaborate box that they called the Ark of God as the habitation of their god, Yahweh; the Hindus consider the river Ganges sacred to the god Shiva—they believe that Shiva is in some specially intimate relation to that river, and they bathe in its waters to benefit from his healing power. The Roman Catholic finds the presence of God concentrated in the consecrated bread and wine, which, he believes, has been transformed into the body and blood of Christ. At a more sophisticated level the material thing may be taken as a symbol of the divine rather than as its direct embodiment, as in the definition of a sacrament given in the Anglican Book of Common Prayer, "an outward and visible sign of an inward and spiritual grace."

In prophetic religion the divine is thought to manifest itself primarily in human society—in the events of human history and in the inspired utterances of great historical figures. It is not denied that nature issues from the divine and is under divine control, but it is not in nature that God is most immediately encountered. The divine reality is to be discovered in great historical events—the destruction of cities, the rise and fall of empires, the escape of a people from bondage. The hand of God is seen in these matters because God is encountered more immediately in the lives and the inspired words of his messengers, the prophets, who reveal in their utterances God's nature, his purposes and commands, and derivatively in the sacred books that contain the records of these revelations. Christianity, Judaism, and Islam, the three chief prophetic religions, are sometimes called religions of the book. Here the key term is not *sacrament* but *revelation.* Prophetic religion, unlike the others, stresses the word as the medium of contact with the divine. (An example is the opening of the Gospel of John.) For the ritualist, and still more for the mystic, whatever words he may use, the consummation of his endeavors is found in a wordless communion with the divine. In prophetic religion, however, the linguistic barrier is never let down; it is not felt as a barrier at all.

The center of mystical religion is the mystical experience, which at its highest development dominates the consciousness, excluding all awareness of words, nature, even of the mystic's own self. In this experience the individual feels himself pervaded and transformed by the divine, identified with it in an indivisible unity. The world and all its ordinary concerns seem as naught as the mystic is caught up in the ineffable bliss of this union. It is not surprising that those who have enjoyed this experience, and those who aspire to it, should take it to be the one true avenue of contact with the divine and dismiss all other modes as spurious, or at least as grossly inferior. Rituals and sacraments, creeds and sacred books, are viewed as paltry substitutes, which are doled out to those who, by reason of incapacity or lack of effort, miss the firsthand mystic communion; or else they are external aids that are of use only in the earlier stages of the quest, crutches to be thrown away when direct access to God is attained.

RESPONSE TO THE DIVINE. In sacramental religion, where the divine is apprehended chiefly in material embodiments, the center of religious activity will be found in ritual acts centering on these embodiments. The sacred places, animals, statues, and such, must be treated with reverence, approached and made use of with due precautions; and around these usages tend to grow prescribed rites. Since the sense of the divine presence in certain objects is likely to be enhanced by participation in solemn ceremonials centering on these objects, the religious activity becomes a self-perpetuating system, embodying what is currently called positive feedback.

In sacramental religion, the ritual tends to absorb most of the religious energies of the adherents and to

crowd the other elements out of the center of the picture. Primitive religion, which is strongly sacramental in character, is often unconcerned with moral distinctions; and we might speculate that the progressive moralization of religion is achieved at the expense of ritual preoccupations. We can see this conflict at many points in the history of religions, most notably in the denunciations that the Hebrew prophets directed against the ritual-minded religionists of their day, and in their exhortations to substitute thirst for righteousness for the concern for niceties of ceremony. Even in its highest developments, sacramental religion tends to slacken the ethical tension that is found in prophetic religion. Where sacramentalism is strong in a monotheistic religion, the natural tendency is to take everything in nature as a divine manifestation. If everything is sacred, then nothing can be fundamentally evil; and thus the distinction between good and evil becomes blurred. One of the elements in the Protestant Reformation was a protest against tendencies to blurring of this sort, which took place in the largely sacramental medieval form of Christianity.

The typical response of prophetic religion to the divine is also nicely coordinated with the chief form in which the divine is apprehended. The reaction naturally called for by a message from the divine is acceptance. This involves both an intellectual acceptance of its contents—belief that whatever statements it makes are true—and obedience to the commands and exhortations it contains. Hence, in prophetic religion faith is the supreme virtue, and affirmations and confessions of faith play an important role. This is illustrated by the insistence of such great Christian prophetic figures as Paul and Martin Luther on faith in Christ as both necessary and sufficient for salvation and by the Muslim practice of repeating daily the creed "There is no God but Allah, and Muḥammad is his prophet." It is important to realize that faith in this sense means far more than the intellectual assent to certain propositions. It also involves taking up an attitude on the basis of that affirmation and expressing that attitude in action. The Jewish prophet Micah expressed the essence of prophetic religion when he said, "What doth the Lord require of thee, but to do justly and to love mercy, and to walk humbly with thy God?" Thus, it would not be incorrect to say that the emphasis in the prophetic response is ethical, providing we do not separate ethics from the believing acceptance of the divine message that is its foundation.

To understand the typical response of mystical religion, we must remember that for the mystic, immediate identification with the divine is of supreme importance. Therefore he concentrates on an ascetic and contemplative discipline that will be conducive to the attainment and maintenance of that condition. He tends to become involved in abstentions and self-tortures designed to wean him from his attachment to things of this world, and in contemplative exercises designed to withdraw the attention from finite things, leaving the soul empty and receptive to influences from the divine. He will make use of ceremonies and will accede to moral principles insofar as he believes them to be efficacious in furthering his ultimate goal. But ultimately they must go; when union with God has been achieved, they are of no more significance. Thus, like sacramentalism, mysticism tends toward the amoral. Only rarely does either become completely amoral, and then for different reasons. For the sacramentalist, conventional moral distinctions may come to seem unimportant because he views everything as equally saturated with the divine; they seem unimportant to the mystic because every finite object or activity is outside the mystic union, and so all are, in the end, equally worthless. The righteous and the wicked are equally far from the true religious goal. While united with God, one does not act.

PLACE OF DOCTRINE. Finally, we may compare the three types of religion with respect to the status of beliefs and creeds. Since faith is central for prophetic religion and since the word is stressed as the primary medium of divine manifestation, it is not surprising that in prophetic religion, creed and doctrine are emphasized more than in the others. Mystical religion, at its purest, is indifferent to matters of belief and doctrine. The mystical experience and the divinity it reveals are often regarded as ineffable, not to be expressed in human language; hence, mystics tend to reject all doctrinal formulations as inadequate. At best, a mystic will admit that some formulations are less inadequate symbols of the unutterable than are others. Thus, in such predominantly mystical groups as the Sufis and the Quakers, little or no attempt is made to enforce doctrinal conformity. And in an extreme form of mysticism, like that of Zen Buddhism, any doctrinal formulation is discouraged. Sacramental religion occupies a middle ground in this respect. In its more primitive forms, it is often extremely indefinite about belief. It has been said that primitive man "dances out his religion." Certainly the elaboration of ritual in primitive religion far outstrips the associated theory. The primitive will often possess an incredibly detailed set of ritual prescriptions but have only the haziest idea of what there is about the nature or doings of the gods that makes them appropriate. In its more developed forms, sacramental theology

becomes more definite, but it is still true that to the extent that a religion is preoccupied with a sacramental approach to the divine, it is more impatient than prophetic religion with doctrinal subtleties.

We can coordinate this classification with the list of religion-making characteristics by pointing out that sacramental religion stresses sacred objects and ritual, prophetic religion stresses belief and morality, and mystical religion places chief emphasis on immediate experience and feeling.

CONCRETE APPLICATION. When we come to apply our scheme to particular cases, we must not suppose that any religion will fall completely in one class or another. In fact, it is better not to think of types of religions, but of religious tendencies that enter in varying proportions into the makeup of any actual religion. However, we can usually say that one tendency or another predominates in a given religion. Thus, Buddhism and philosophical Hinduism are predominantly mystical; Judaism, Islam, and Confucianism are primarily prophetic; and popular Hinduism, in company with all polytheistic and primitive religions, is primarily sacramental. Often a religion that begins with a definite bent will admit other elements in the course of its development. Islam, which began as the most severely prophetic of religions, has developed one of the world's most extreme group of mystics in the Sufis, who are completely out of harmony with the spirit of Muḥammad, no matter how they may continue to express themselves in his phrases. Again, in Tibet, Buddhism has undergone a development quite foreign to its founder's intentions, blossoming into an extremely elaborate sacramentalism.

Christianity furnishes a good opportunity to study the intermingling and conflict of the different tendencies. It began as an outgrowth of Jewish prophecy, but in the process of adapting itself to the rest of the Western world it took on a considerable protective coloration of both the sacramental and mystical, and these aspects have remained with it throughout its career. Christian mysticism presents a good example of an element existing in a religion that is dominated by another element. As the price of toleration, Christian mystics have had to pay lip service to the official theology and to the prophetic moral element; and as a result, mystic thought and practice in Christianity have seldom received the extreme development found in India. In those cases where the mystical spirit has burst the fetters, as with Meister Eckhart, official condemnation has often resulted.

Looking at Christianity today, it can be said that although it is predominantly a prophetic religion, as compared with Hinduism and Buddhism, with respect to its internal divisions the Catholic wing (both Roman and Greek) tends more toward the sacramental, while the Protestant is more purely prophetic, with mysticism appearing sporadically throughout. In Catholicism the elaborateness of prescribed ceremonies, the emphasis on the necessity of material sacraments for salvation, and the insistence on a special status for consecrated priests are all typically sacramental. In Protestantism the emphasis on the sermon (the speaking forth of the Word of God) rather than on ritual, the emphasis on the Bible as the repository of divine revelation, and the moral earnestness and social concern are all earmarks of the prophetic spirit."Religion" new copy p. 235:

See also Buddhism; Chinese Philosophy: Religion; Christianity; Creation and Conservation, Religious Doctrine of; Epistemology, Religious; Islamic Philosophy; Jewish Philosophy; Philosophy of Religion, History of; Philosophy of Religion, Problems of; Philosophy of Religion; Religion and Morality; Religion and the Biological Sciences; Religion and the Physical Sciences; Religion, Naturalistic Reconstructions of; Religion, Psychological Explanations of; Religious Experience, Argument for the Existence of God; Religious Experience; Religious Language; Theism, Arguments For and Against.

Bibliography

Friedrich Schleiermacher, *On Religion,* translated by John Oman (New York: Harper, 1958), and *The Christian Faith,* translated by H. R. Mackintosh (Edinburgh, 1956), contain classic statements of the view that religion is essentially a mode of experience. A more recent statement of this point of view that emphasizes mystical experience is in W. T. Stace, *Time and Eternity* (Princeton, NJ: Princeton University Press, 1952). The moral aspect of religion is stressed in Immanuel Kant, *Religion within the Limits of Reason Alone,* translated by T. M. Greene and H. H. Hudson (La Salle, IL: Open Court, 1934), and in John Baillie, *The Interpretation of Religion* (New York: Scribners, 1928). Conceptions of religion from the standpoint of philosophical naturalism are to be found in Auguste Comte, *A General View of Positivism* (London, 1865); Ludwig Feuerbach, *The Essence of Christianity,* translated by George Eliot (New York: Ungar, 1957); George Santayana, *Reason in Religion* (New York: Scribners, 1905); John Dewey, *A Common Faith* (New Haven, CT: Yale University Press, 1934); and Erich Fromm, *Psychoanalysis and Religion* (New Haven, CT: Yale University Press, 1950).

Stimulating discussions, not so easily classified, are Henri Bergson, *The Two Sources of Morality and Religion,* translated by R. Ashley Audra (New York: Holt, 1935); William James, *The Varieties of Religious Experience* (New

York: Longman, 1902); Josiah Royce, *The Sources of Religious Insight* (New York: Scribners, 1912); John Oman, "The Sphere of Religion," in *Science, Religion and Reality,* edited by Joseph Needham (New York: Macmillan, 1925); and Alfred North Whitehead, *Religion in the Making* (New York: Macmillan, 1926). Important discussions from the social sciences include Émile Durkheim, *The Elementary Forms of the Religious Life,* translated by J. W. Swain (London: Allen and Unwin, 1915), and Bronisław Malinowski, *Magic, Science, and Religion* (Garden City, NY: Doubleday, 1954). James H. Leuba, *A Psychological Study of Religion* (New York: Macmillan, 1912), and Julian Huxley, *Religion without Revelation* (New York: Harper, 1958), provide extensive critical discussion of a wide variety of definitions, as well as presenting original conceptions.

OTHER RECOMMENDED TITLES

Banton, Michael, ed. *Anthropological Approaches to the Study of Religion.* Reprint edition, London: Routledge, 1990.

Batson, C. Daniel et al. *Religion and the Individual: A Social-Psychological Perspective.* New York: Oxford University Press, 1993.

Berger, Peter. *The Social Reality of Religion.* Harmondsworth, U.K.: Penguin, 1967.

Boyer, Pascal. *Religion Explained: The Evolutionary Origins of Religious Thought,* Reprint edition. New York: Basic, 2002.

Braun, Willi, and Russell McCutcheon, eds. *Guide to the Study of Religion.* London: Cassell, 2000.

Byrne, Peter. *Natural Religion and the Nature of Religion.* London: Routledge, 1989.

Clarke, Peter, and Peter Byrne, eds. *Religion Defined and Explained.* New York: St. Martin's Press, 1993.

Derrida, Jacques, and Gianni Vattimo, eds. *Religion.* Stanford, CA: Stanford University Press, 1998.

Godlove, Terry. *Religion, Interpretation, and Diversity of Belief: The Framework Model from Kant to Durkheim to Davidson.* Cambridge, U.K.: Cambridge University Press, 1989.

Guthrie, Steward. *Faces in the Clouds: A New Theory of Religion.* New York: Oxford University Press, 1993.

Hick, John. *An Interpretation of Religion: Human Responses to the Transcendent.* New Haven, CT: Yale University Press, 1989.

Lawson, E. Thomas, and Robert McCauley. *Rethinking Religion: Connecting Cognition and Culture.* New York: Cambridge University Press, 1990.

Phillips, D. Z. *Religion without Explanation.* Oxford: Blackwell, 1976.

Smart, Ninian. *The Science of Religion and the Sociology of Knowledge.* Princeton, NJ: Princeton University Press, 1973.

Yinger, J. Milton. *The Scientific Study of Religion.* New York: Macmillan, 1970.

William P. Alston (1967)

Bibliography updated by Christian B. Miller (2005)

RELIGION, NATURALISTIC RECONSTRUCTIONS OF

In philosophy a naturalist is one who holds that there is nothing over and above nature. A naturalist is committed to rejecting traditional religion, which is based on beliefs in the supernatural. This does not necessarily carry with it a rejection of religion as such, however. Many naturalists envisage a substitute for traditional religion that will perform the typical functions of religion without making any claims beyond the natural world. We can best classify naturalistic forms of religion in terms of what they take God to be—that is, what they set up as an object of worship. In traditional religion the supernatural personal deity is worshiped because he is thought of as the zenith of both goodness and power. More generally, we can say that religious worship is accorded to any being because it is regarded as having a controlling voice in the course of events and at least potentially exercising that power for the good. This suggests that to find a focus for religious responses in the natural world, we should look for a basic natural source of value. Forms of naturalistic religion differ as to where this is located. Broadly speaking, achievements of value in human life are due to factors of two sorts: (1) man's natural endowments, together with the deposit of his past achievements in the cultural heritage of a society, and (2) things and processes in nonhuman nature on which man depends for the possibility of his successes and, indeed, his very life. Most naturalists locate their religious object primarily on one or the other side of this distinction, although some try to maintain an even balance between the two.

The first factor is stressed most by those who are called religious humanists. This group includes Ludwig Feuerbach and Auguste Comte in the nineteenth century and John Dewey and Erich Fromm in the twentieth. Of these men Comte has been the most influential.

COMTE

In Comte's view, it is to humanity that the individual man owes everything that he is and has. It is because he shares in the general biological and psychological capacities of human nature that he is able to live a human life. And the men of a given generation are able to lead a fully human life because of the labors of their predecessors in building up their cultural heritage. Moreover, according to Comte, the service of humanity, in the many forms this can take, is the noblest ideal that could be proposed to an individual; and humanity, unlike an omnipotent God, needs this service. Thus, Comte proposed to set up a religion of

humanity with man, viewed as a unitary though spatiotemporally scattered being, as the object of worship.

Unlike many naturalists Comte was not at all vague about the detailed functioning of his proposed religion. He was impressed with the ritual structure of Roman Catholicism and took it as his model. For example, in the analogue of baptism, the sacrament of presentation, the parents would dedicate their child to the service of humanity in an impressive public ceremony. Public observances were to be reinforced by the regular practice of private prayer, on which Comte laid the greatest stress. A person was to pray four times daily, with each prayer divided into a commemorative and a purificatory part. In the first part one would invoke some great benefactor of humanity; by reflecting gratefully on his deeds, one would be inspired to follow his example, and one's love of humanity would thus be quickened. The purificatory part would give solemn expression to the noble desires thereby evoked; in it the individual would dedicate himself to the service of humanity. Other rituals included a system of religious festivals and a calendar of the saints of humanity that provided the material for the prayers on each day of the year.

Some idea of the religious fervor generated in Comte by the contemplation of humanity may be gained from this quotation from *A General View of Positivism*:

> The Being upon whom all our thoughts are concentrated is one whose existence is undoubted. We recognize that existence not in the Present only, but in the Past, and even in the Future: and we find it always subject to one fundamental Law, by which we are enabled to conceive of it as a whole. Placing our highest happiness in universal Love, we live, as far as it is possible, for others: and this in public life as well as in private; for the two are closely linked together in our religion; a religion clothed in all the beauty of Art, and yet never inconsistent with Science. After having thus exercised our powers to the full, and having given a charm and sacredness to our temporary life, we shall at last be forever incorporated into the Supreme Being, of whose life all noble natures are necessarily partakers. It is only through the worship of Humanity that we can feel the inward reality and inexpressible sweetness of this incorporation. (p. 444)

Comte had considerable influence in his lifetime, and a few functioning parishes of his religion of humanity sprang up. They have not survived, however, and a revival in our time hardly seems likely. In the twentieth century, reeling under the impact of two world wars and the hourly expectation of the death knell of civilization, we are not inclined to grow misty-eyed over humanity. Recent humanists have tended to be more critical in their reverence. The latest trend is to single out the more ideal aspects of man—his aspirations for truth, beauty, and goodness—for religious worship. Or the emphasis shifts from man as he actually exists to the ideals that man pursues in his better moments. Thus, in his book *A Common Faith,* John Dewey defines God as "the unity of all ideal ends arousing us to desire and action" (p. 42).

DEWEY

Unlike Comte, Dewey has no interest in developing an organized naturalistic religion. It would seem that religious organization and religious ritual are too closely associated in his mind with the supernaturalism that he rejects. For Dewey the important thing is the religious quality that experience can assume under certain conditions. Any unification of the whole self around the pursuit of an ideal end is religious in quality. Dewey is emphatic in insisting that this is a quality, rather than a kind, of experience. Whenever a person is thoroughly committed to the pursuit of any ideal, be it scientific, social, artistic, or whatever, his experience attains the kind of fulfillment that has always been characteristic of what is most valuable in religion. According to Dewey, in traditional religion this quality has been encumbered and obscured by irrelevant trappings, particularly the theological dogma in terms of which it has been pursued. In the past, self-integration in the pursuit of the ideal has been thought of as service of God, unity with God, or submission to God's will. It is Dewey's conviction that the religious quality can be more effectively sought if the quest is not carried on under this banner. To reflective men, supernaturalistic dogma will always appear dubious at best. If the quest for self-integration in the service of the ideal is too closely tied to theology, it will be endangered when the theology is rejected as rationally groundless. Moreover, insofar as the theology is taken seriously, it diverts attention from the active pursuit of the ideal. Worse, the assurance that the good is already perfectly realized in the divine nature has the tendency to cut the nerve of moral effort; in that case it is not up to us to introduce the good into the world. Thus, Dewey's main concern as a philosopher of religion is to redirect religious ardor into the quest for a richer quality of human life rather than to construct a framework for a naturalistically oriented religious organization.

There is no developed naturalistic philosophy of religion that stresses the nonhuman side of the natural sources of value to the extent to which Comte stresses the human side. (Though we can find this in literature, notably in Richard Jeffries, who had a kind of religious intoxication with inanimate nature without, however, conceiving of it as suffused with a spiritual being or beings. This is a naturalistic counterpart of the nature worship of ancient Greece, just as Comte's religion of humanity is a naturalistic counterpart of an ethical monotheism like Christianity.) However, there is a marked tendency among contemporary naturalists to emphasize the nonhuman side much more than Comte or Dewey. Good examples of this are the liberal theologian Henry Nelson Wieman and the biologist Julian Huxley, who in his book *Religion without Revelation* has made the most coherent and comprehensive recent attempt to sketch out a naturalistically oriented religion.

HUXLEY

According to Huxley's conception, religion stems from two basic sources. One is man's concern with his destiny—his position and role in the universe and their implications for his activity; the other is the sense of sacredness. Following Rudolf Otto, Huxley thinks of the sense of sacredness as a unique kind of experience that is an intimate blend of awe, wonder, and fascination; this mode of feeling arises spontaneously in reaction to a wide variety of objects and situations. Religion, then, is a social organ for dealing with problems of human destiny. As such it involves a conception of the world within which this destiny exists, some mobilization of the emotional forces in man vis-à-vis the world thus conceived, some sort of ritual for expressing and maintaining the feelings and attitudes developed with respect to the forces affecting human destiny, and some dispositions with respect to the practical problems connected with our destiny. The sense of sacredness enters into the second and third of these aspects. As Huxley sees it, a way of dealing with problems of human destiny would not be distinctively religious if it did not stem from and encourage a sense of the sacredness of the major elements in its view of the world, man, and human life.

Huxley, as a thoroughgoing naturalist, holds that the supernaturalistic worldview in terms of which religion has traditionally performed its functions is no longer tenable in the light of modern scientific knowledge. Moreover, he thinks that it is possible to develop a full-blown religion on a naturalistic basis. As the intellectual basis for such a religion, Huxley puts forward "evolutionary naturalism," a view of the spatiotemporal universe, inspired by modern biology and cosmology, in which the universe is conceived of as an indefinitely extended creative process, always tending to higher levels of development, with all the sources and principles of this creativity immanent in the process. The basic role of man is to be the chief agent of this evolutionary advance on earth through the application of his intelligence to the problems of life on Earth and through the building of a harmonious and stable community. A religion based on these conceptions will be focused on an object of worship that is a construct out of all the forces affecting human destiny, including basic physical forces as well as the fundamental facts of human existence and social life. God, then, will consist of all these factors, held together by the feeling of sacredness with which they are apprehended. As a start toward conceiving this assemblage as a unified object of worship, Huxley presents a naturalistic version of the Christian doctrine of the Trinity. God the Father is made up of the forces of nonhuman nature. God the Holy Ghost symbolizes the ideals toward which human beings at their best are striving. God the Son personifies human nature as it actually exists, bridging the gulf between the other two by channeling natural forces into the pursuit of ideals. And the unity of all the persons as one God represents the fact that all these aspects of the divine are intimately connected.

Many thinkers, atheists as well as theists, take a dim view of all these proceedings. Since the theists' lack of enthusiasm stems from obvious sources, let us concentrate on the atheists. The issues here are normative or evaluative rather than factual. Comte and Huxley as philosophers of religion are not, with perhaps minor exceptions, making any factual judgments with which other naturalists might disagree because they are making no factual judgments at all beyond their basic commitment to naturalism. If a man like Bertrand Russell or Jean-Paul Sartre disagrees with Huxley, he differs about the value of what Huxley is proposing. His low evaluation may have different bases. First, he may feel that man or the basic forces of nature constitute too pallid a substitute for the God of theism to afford a secure footing for the distinctively religious reactions of reverence, adoration, and worship. A man like Huxley might, for his part, interpret this as a reflection of a suppressed hankering after the old supernatural deity. Second, Russell or Sartre may turn this charge on Huxley and maintain that one searches for an object of worship within nature only because he has not sufficiently emancipated himself from the old religious orientation and that this religion of evolutionary naturalism represents an uneasy compromise between religious and secular orientations. It seems clear

that there is no one objective resolution of such disputes. People differ in such a way that different total orientations will seem congenial to people with different temperaments and cultural backgrounds. It is perhaps unfortunate, on the whole, that many people need to find something fundamentally unworthy in every other religion in order to find a firm attachment to their own religious positions, although it is undoubtedly true that religious discussions are more lively than they would be if this were not the case.

See also Comte, Auguste; Dewey, John; Evolutionary Theory; Feuerbach, Ludwig; Human Nature; Naturalism; Otto, Rudolf; Russell, Bertrand Arthur William; Sartre, Jean-Paul.

Bibliography

Ames, E. S. *Religion.* New York, 1929.

Comte, Auguste. *A General View of Positivism.* Translated by J. H. Bridges. London, 1880. Ch. 6.

Dewey, John. *A Common Faith.* New Haven, CT: Yale University Press, 1934.

Feuerbach, Ludwig. *The Essence of Christianity.* Translated by George Eliot. New York: Harper, 1957.

Fromm, Erich. *Psychoanalysis and Religion.* London: Gollancz, 1951.

Huxley, Julian. *Religion without Revelation.* New York: Harper, 1957.

Russell, Bertrand. *Why I Am Not a Christian and Other Essays.* New York: Simon and Schuster, 1957.

Wieman, Henry Nelson. *The Source of Human Good.* Chicago: University of Chicago Press, 1946.

William P. Alston (1967)

RELIGION, NATURALISTIC RECONSTRUCTIONS OF [ADDENDUM]

CONTEMPORARY NATURALISTIC RELIGION

What should one contrast nature with? The supernatural, maybe? What is meant here by supernaturalism is the thesis that the divine is different in kind from familiar things and persons; and/or that there are divine interventions that are contrary to the laws of nature. If this is the relevant contrast then naturalistic religion requires merely that God be taken as either a person or a community of persons. God is then like humans, although infinitely more powerful, and acts in the world in whatever way people act when they exercise their freedom. Such antisupernaturalism is weaker than naturalism as understood by contemporary philosophers, who would balk at calling the dualist Richard Swinburne (2004), the idealists Timothy Sprigge (1983) and John Foster (2004), or even the nonreductive physicalist Peter Forrest (1996) naturalists. This suggests that naturalism is to be contrasted not merely with the supernatural but also with anthropocentric *Metaphysics*, which takes consciousness and agency as fundamental features of reality that may be used to explain but must themselves be accepted without explanation. Naturalism in this strong sense is unlikely to support the humanist attitudes of Auguste Comte or John Dewey, but coheres well with Julian Huxley's evolutionary naturalism.

DEEP ECOLOGY

The most widespread contemporary naturalistic movement with religious tendencies is deep ecology, which typically goes beyond an attitude of aesthetic appreciation of—and scientific interest in—life on earth, to attitudes of reverence and self-sacrifice (Naess 1989). Combined with a suitable metaphysical system this could be a genuinely naturalistic religion, although neopagan movements such as Wicca tend to incorporate belief in the supernatural. Two such metaphysical systems are process theology and pantheism.

PROCESS PHILOSOPHY

The process philosophy of Alfred North Whitehead, Charles Hartshorne, and most recently David Ray Griffin (2001) can support either a liberal theistic religion or deep ecology. Process philosophy counts as naturalistic because it is biocentric rather than anthropocentric, in that it relies on preconscious sensitivity to the environment (prehension) and final causation. For that reason the God of process philosophy is immanent in the processes of the natural world, resulting in something similar to, although less austere than, Huxley's evolutionary naturalism. A chief objection to process philosophy is that we no longer have a theoretical need for either prehension or final causes even in biology.

PANTHEISM

The universe as a whole or, perhaps better, the natural order is sufficiently awe-inspiring to ground some religious attitudes. So pantheism can form the basis of a naturalistic religion (Levine 1994). Like any religion this has metaphysical commitments: either the existence of the universe as a whole or the existence of laws of nature, but neither of these commitments would worry most naturalists.

THE AFTERLIFE

Much religious motivation (for good and ill) lies in the belief in an afterlife. Does naturalism cohere with this belief? Granted that if there is a God concerned about individuals then there is not much problem, for there are ways God could ensure an afterlife without miracles and without there being souls (van Inwagen 1992). However, a pantheist God that just is the natural order will not be concerned about individuals, whereas the God of process philosophy might well lack the power required to be providential. Frank J. Tipler (1996) has suggested that in the distant future sentient beings will be able to reconstitute all the lives of those who have died. In his version all possible lives seem to get reconstituted, which prevents any of them being the same as early twenty-first century people. But one might surmise that there are traces of actual lives that could be used to reconstitute only those who have actually lived. A less far-fetched naturalistic account of the afterlife is based on the many worlds interpretation of quantum theory. For if there are many parallel universes and every physically possible event occurs in some of them, then in some of them it seems humans survive anything (Price 1996, ch. 9; Lewis 2004) The chief problem with such scenarios is over-survival, that is, at each moment each person divides into millions of successors.

CONCLUSION

Not surprisingly the more narrowly naturalism is understood the more drastic a naturalistic reconstruction of religion must be. At one extreme, anti-supernaturalism sits comfortably with all but conservative religious movements. At the other, naturalists might reject even the biocentrism of process thought and be left with only a rather austere pantheism.

See also God, Concepts of; Naturalism; Pantheism; Physicalism; Whitehead, Alfred North.

Bibliography

Forrest, Peter. *God without the Supernatural*. Ithaca, NY: Cornell University Press, 1996.

Foster, John. *The Divine Lawmaker: Lectures on Induction, Laws of Nature, and the Existence of God*. Oxford: Clarendon Press, 2004.

Griffin, David Ray. *Reenchantment without Supernaturalism*. Ithaca, NY: Cornell University Press, 2001.

Levine, Michael P. *Pantheism: A Non-Theistic Concept of Deity*. New York: Routledge, 1994.

Lewis, David. "How Many Lives Has Schrödinger's Cat?" *Australasian Journal of Philosophy* 82 (2004): 3–22.

Naess, Arne. *Ecology, Community, and Lifestyle: Outline of an Ecosophy*. Translated and revised by David Rothenberg. Cambridge, U.K.: Cambridge University Press, 1989.

Price, Huw. *Time's Arrow & Archimedes Point: New Directions for the Physics of Time*. Oxford: Oxford University Press, 1996.

Sprigge, Timothy. *The Vindication of Absolute Idealism*. Edinburgh, Scotland: Edinburgh University Press, 1983.

Swinburne, Richard. *The Existence of God.* 2nd ed. Oxford: Clarendon Press, 2004.

Tipler, Frank J. *The Physics of Immortality: Modern Cosmology, God and the Resurrection of the Dead*. London: Pan, 1996.

van Inwagen, Peter. "The Possibility of Resurrection." In *Immortality*, edited by Paul Edwards, 242–246. New York: Macmillan, 1992. Reprinted from the *International Journal for the Philosophy of Religion* 9 (1978).

Peter Forrest (2005)

RELIGION, PSYCHOLOGICAL EXPLANATIONS OF

In the seventeenth and eighteenth centuries the chief impact of science on religion came from the revised picture of the cosmos that emerged from developments in astronomy and physics. In the nineteenth century the impact was from the changed view of the history of life on Earth that was presented by geology and evolutionary biology. In the twentieth century the social sciences had the greatest impact on religion, although of a different nature. Physics and biology worried theologians because they introduced theories about the cosmos, life, and man that were at variance with beliefs intimately bound up with the religious tradition, such as the special creation of man. The impact of the social sciences, on the other hand, comes not from theories that contradict basic religious doctrines but from explanations of religion itself that seem to rob it of its significance.

Since the nineteenth century numerous ideas have been put forward as to the psychological and sociological factors that are responsible for religion. The most important of these are (1) the Marxian theory that religion is one of the ideological reflections of the current state of economic interrelations in a society; (2) the similar, but more elaborately developed, theory of the sociologist Émile Durkheim that religious belief constitutes a projection of the structure of society; and (3) the Freudian theory that religious belief arises from projections designed to alleviate certain kinds of unconscious conflict. These are all scientific explanations in that they trace religion to factors wholly within the world of nature, and hence they

are, at least in principle, subject to empirical test. Concentration on one of these, the Freudian, will enable us to illustrate the philosophical problems raised by such explanations.

THE FREUDIAN EXPLANATION

The Freudian account begins with certain similarities between attributes of and attitudes toward a personal deity, on the one hand, and the small child's conception of and mode of relating to his father, on the other. In both cases the superior being is regarded as omnipotent, omniscient, inscrutable, and providential. In both cases the individual reacts to this superior being with utter dependence, awe, fear of punishment, and gratitude for mercy and protection. These parallels suggest, though they do not prove, that the original model for the conception of God is to be found in the infantile conception of one's parents, and that the almost universal inclination to believe in personal deities is to be traced to psychological remnants of the infantile situation. According to Sigmund Freud, these remnants are mostly the result of the Oedipal conflict. According to his theory, around the age of four the boy (restricting ourselves to the male for simplicity of exposition) comes to desire his mother sexually and to regard his father as a rival. Reacting more or less to actual indications, the boy becomes so afraid of the father's hostility, and also so afraid of losing his love, that he not only abandons his sexual aims but also represses the entire complex of desires, fears, and conceptions. This complex remains, in greater or lesser intensity, in the unconscious; and it is because a supernatural personal deity provides an external object on which to project it that men have as much inclination as they do to believe in such a being and to accept the attitudes and practices that go with this belief.

To understand what the projection does for the individual, we must recognize that the repressed material involves severe conflict between tendencies to rebel against the father and tendencies to submit to the father, and between the Oedipal desires and the standards that would be violated by satisfying those desires. Projection of this material onto an external deity reduces distress in several ways. First, the externalization of the problem provides some relief. Instead of being plagued by mysterious discomfort, the individual is faced with a clear-cut opposition between various desires of his own and a forbidding external person. Second, there is less conflict because the external figure is so powerful as to seriously weaken the rebellion, and he is so idealized as to render resentment and hostility less appropriate. Third, there are various mechanisms provided for dissipating the guilt over sexual desire for the mother and hostility toward the father. Confession, penance, and renunciations of various kinds afford socially approved means for relieving this guilt and counteracting its crippling influence.

People are more receptive to religious belief at some times than at others. Freud explains this in terms of the mechanism of regression. When a person encounters severe difficulties and frustrations at one stage of life, he tends to regress psychologically to an earlier stage at which these problems did not exist. Thus, when an adult is particularly hard pressed, there is generally some reinstatement of earlier modes of thinking, feeling, and relating to the environment. This means that the Oedipal material in the unconscious will become more intense and closer to the surface, while at the same time the person is more likely to engage in the childish practice of projection.

Thus, according to Freudian theory, an individual's tendency to accept belief in a supernatural personal deity (together with the other aspects of religious activity and involvement) is at least partly caused by a tendency to project a childhood father image existing in the unconscious, this projection normally following a regression set off by a current problem of adjustment and serving to alleviate unconscious conflicts and unconscious guilt. It is clear that, at best, this is only a partial explanation of religious belief. For one thing, it presupposes the prior existence of the religious ideas in the culture; at most, it is an explanation of the individual's readiness to accept these ideas when they are proffered.

Freud tried to supply this lack by developing a parallel theory of the development of religion in society. According to this theory, religion develops as a projection of a psychological complex that results from unconscious racial memories of a primal murder of the tyrannical father figure of a "primal horde." Cultural development is thus treated along the same lines as the development of the individual; something like a "collective unconscious" is posited in which psychic material can be transmitted in an unconscious form from one generation to another. However, these ideas have never won any considerable degree of acceptance, and in discussing Freud we can concentrate on his account of the psychological basis of religion in the individual.

CRITICISM OF FREUDIAN EXPLANATION

With respect to any scientific explanation of religion, there are two questions to be raised. (1) What reason is

there to accept it? (2) If it is true, what bearing does it have on the truth, value, or justifiability of religion? It is the second question that specially lies within the province of the philosophy of religion.

It is clear that the Freudian explanation does not imply that the beliefs of religion are false; Freud himself recognized this, though not all Freudians do. But it is often assumed that the success of any explanation of religion in terms of factors within the natural world would show that we do not need to bring anything supernatural into the explanation, and hence would seriously weaken religion's claims to credibility. However, this depends on how these claims were made. If religion is based solely on divine revelation, then the fact that we can give an adequate explanation of religion without bringing in divine activity, revelatory or otherwise, seriously affects—though it does not conclusively disprove—the claim that certain beliefs are true because they are communicated to man by God. But if rational arguments are advanced in support of religious doctrine, such as the classical arguments for the existence of God, then whatever force these arguments have is in no degree lessened by the fact—if it be a fact—that the psychological basis for religion is as Freud supposed. Of course, if the Freudian mechanisms constitute a necessary as well as sufficient condition of religious belief, then it follows that no one has any good reason for these beliefs. If anyone did have a good reason, that would itself be a sufficient condition of the belief, and this would show that it is possible to have the belief without needing to project an unconscious father image. However, it is almost inconceivable that we should show that projection is a necessary condition of belief. At most, we could hope to show that there is some correlation between degree of unconscious Oedipal conflict and firmness of religious belief. Showing that a certain set of natural factors is one of the things that can produce religious belief may well nullify certain ways of supporting the beliefs, but it could hardly show that no adequate rational grounds could be produced.

There is another way in which it has been thought that the Freudian theory of religion carries with it a negative evaluation of religion. The particular causal factors to which Freud traced religion are of a sort associated with undesirable patterns of organization. To regard religion as caused by these factors is to class it with neurotic and infantile modes of behavior, and as such it is hardly worthy of serious consideration. In this respect, too, the psychoanalytic explanation is typical. One can imagine an explanation that traces religious activity to evaluatively neutral natural factors, such as patterns of neural activity in the brain, but all the explanations in the field trace religion to states and activities that are more or less irrational, immature, or unworthy. Projection is involved in all the theories cited at the beginning of this article; the Marxist theory adds the point that religion is used by the dominant class to provide illusory consolations to those being exploited.

To be clear on this issue, we must distinguish the different forms these claims can take. Psychoanalytic literature is often simply an enumeration of similarities between religion and compulsion neuroses, such as firm attachment to rituals without having a rational explanation of the attachment. However, the similarity in itself proves nothing. A scientist "obsessed with an idea" also exhibits marked similarities to a compulsion neurotic, but this has no implications for the value of his work. The more important claim has to do with the causal factors said to underlie religion. Here, too, we must distinguish between (1) the claim that some neurotic condition is always or generally among the factors producing attachment to a religion, and (2) the claim that the causal basis of such attachment is markedly similar to the basis of recognized neuroses. There is no real evidence for the first claim. Controlled studies on the required scale have never been carried out. As for the second, we must ask how similar the causal basis is and what implications we are to draw from whatever degree of similarity exists. The mere fact that religion involves projection as a relief from unconscious conflict is not sufficient ground for labeling religion, in Freud's terms, "the universal obsessional neurosis of mankind." We must distinguish between pathological and healthy resolutions of unconscious conflict.

The anti-Freudian psychoanalyst Carl Jung, in terming religion an alternative to neurosis, expressed his belief that it is a healthy outcome. The basic issue involved here concerns the definition of "neurosis." If we define it in terms of a certain causal basis, then it may be that according to the Freudian theory, religion is, by its very nature, a form of neurosis. But then it remains an open question whether or not it is a desirable, justifiable, or realistic mode of activity. If neurosis is defined in this way, we may have to distinguish between good and bad neuroses. If, on the other hand, we accept common usage and build a negative evaluation into the definition of neurosis (by having as a necessary condition of neurosis that it make a satisfactory adjustment to one's environment difficult), then it would no longer be an open question whether religion, if neurotic, is a good thing. But with this concept of neurosis, we have a much stronger thesis, which calls for evidence that has not yet been pro-

vided. No one has shown that in general religious believers are less able to establish satisfying personal relations and less able to get ahead in their work than are nonbelievers. Even if this were shown, there would be further problems of a very sticky sort. The believer might complain that restricting "the environment" to the natural environment is question-begging. He would say that whatever the bearing of religious attachment on getting along in human society, it is essential to adequate adjustment to God and his demands. To ignore this aspect of "the environment" is to employ a criterion of adjustment that presupposes the falsity of religious beliefs.

Similar comments apply to the idea that the psychoanalytic theory implies that religion is infantile and hence unworthy of mature men. It is true that the way a religious man relates himself to God is in many ways similar to the way a small child relates himself to a father. But whether or not this is a mature, realistic mode of activity is wholly a function of whether there really is such a God. If there is, then this is the only reasonable stance to take. Hence, to condemn religion on these grounds is to presuppose the falsity of its beliefs.

Thus, there are many gaps in any line of reasoning that tries to derive a negative evaluation of religion from a causal explanation of religion in psychological or sociological terms. If a person does not feel that he has a firm basis for his religious beliefs, then looking at religion in a Freudian or Marxian light may well lead him to give up his beliefs. More generally, we can say that Freudian or Marxian theory does not provide an intellectual atmosphere in which one would expect religious belief to flourish; but it does not appear that these theories, as so far developed, are in any way logically incompatible with the truth, justifiability, and value of traditional religion.

See also Durkheim, Émile; Freud, Sigmund; Jung, Carl Gustav; Marxist Philosophy; Philosophy of Religion, Problems of; Popular Arguments for the Existence of God.

Bibliography

Important treatments of religion from a psychological point of view include Sigmund Freud, *Totem and Taboo,* translated by A. A. Brill (New York: Moffat, Yard, 1918), *The Future of an Illusion,* translated by W. D. Robson-Scott (New York: Liveright, 1928), and *Moses and Monotheism,* translated by Katherine Jones (New York: Knopf, 1939); Carl Jung, *Psychology and Religion* (New Haven, CT: Yale University Press, 1938), and *Modern Man in Search of a Soul,* translated by W. S. Dell and C. F. Baynes (New York: Harcourt Brace, 1933); Erich Fromm, *Psychoanalysis and Religion* (New Haven, CT: Yale University Press, 1950); A. T. Bosien, *The Exploration of the Inner World* (New York: Willett, Clark, 1952); and Theodor Reik, *Dogma and Compulsion,* translated by Bernard Miall (New York: International Universities Press, 1951). R. S. Lee, *Freud and Christianity* (New York: A. A. Wyn, 1949), and William P. Alston, "Psychoanalytic Theory and Theistic Belief," in *Faith and the Philosophers,* edited by John Hick (New York: St. Martin's Press, 1964), present discussions of the Freudian treatment of religion.

For a sociological point of view, see Émile Durkheim, *The Elementary Forms of the Religious Life,* translated by J. W. Swain (London: Allen and Unwin, 1915); Vilfredo Pareto, *The Mind and Society,* translated by Andrew Bongiorno and Arthur Livingston (New York: Harcourt Brace, 1935), Vol. III; V. F. Calverton, *The Passing of the Gods* (New York: Scribners, 1934); and G. E. Swanson, *The Birth of the Gods* (Ann Arbor: University of Michigan Press, 1960).

William P. Alston (1967)

RELIGION, PSYCHOLOGICAL EXPLANATIONS OF [ADDENDUM]

During the last few decades of the twentieth century scientific ability to explore the brain directly increased dramatically, so neuroscientific discoveries during the period resulted in a broadening of perspectives from which psychological explanations of religion may be given. First, the ideological impasse on method between behavioristic and psychoanalytic or introspective approaches in psychology yielded to more pragmatic heterophenomenological (Dennett 2003) or neurophenomenological (Varela, Thompson, and Rosch 1991) methods for investigating mental states. Second, Platonic and Cartesian views of emotion as inherently irrational and subversive of productive cognitive functioning were contested by studies that showed that absence of emotion produced a cognitively dysfunctional Phineas Gage, not a pure-minded Philosopher King (Damasio 1994). Third, the Enlightenment notion of a person as an isolated, autonomous rational optimizer, a "ghost" in a bodily machine, began to yield to a notion of a person as an embodied and interactive global workspace (Baars 1997) that is distributed across both interpersonal relationships (attachment theory; Panksepp 1998) and the environment (Clark 1999). Fourth, clinical, cognitive, and transpersonal psychologists (Wilber 1998) began to see the value of studying and using religion in their clinical practices to aid in communication, understanding, and healing.

The net effect of these shifts in perspective on psychological methodology and ontology has been a return to a Jamesian (James 1981) view of human psychology as consisting of a stream of variously conscious and unconscious processes, related to one another in modular ways (Fodor 1983; Weiskrantz 1997), and integrated somewhat haphazardly through the accidents of evolutionary history.

New avenues of exploration for religious psychological states, beliefs, and practices have been opened by these developments in cognitive neuroscience, as well as by new technology. Some of these include:

(1) Brain scans: studies of the brains of persons engaged in religious activities, through Positron Emission Tomography, Computed Axial Tomography, and Single Photon Emission Computed Tomography (SPECT) scans, and a comparison of the experimental results with base-line brain scans and scans of persons with known pathological conditions such as brain lesions, schizophrenia, and epilepsy.

(2) First-person methods: Without behaviorist presuppositions, methods for systematic and controlled introspection can be studied in a critical but open-minded way.

(3) Health and integration studies: Studies of the interpersonal and integrative effects of religious experiences, beliefs, or practices are being done in clinical settings.

While it is still possible that the pathologies attributed to religious experience, beliefs, and practices by eliminativists, Freudians, Marxists, and Durkheimians might be corroborated by the emerging twenty-first–century evidence, religious psychology is now at least open to the vindication of religion from charges of pathology. In what follows, samples of each of the previous lines of inquiry into the psychology of religion in cognitive science are cited.

BRAIN SCAN STUDIES

Brain scans of advanced-level meditators, persons suffering from hallucinations, and persons engaged in prayer or other religious ceremonies are being produced by researchers at a variety of universities and institutes. Michael A. Persinger (1993) induced hallucinations in laboratory subjects through stimulation of temporal lobes of the brain. Based on this evidence and reports of religious experience by schizophrenic and epileptic patients, he argues that religious experiences, as hallucinations, are a result of kindling, erratic neuronal stimulation that spreads through sections of the brain. He also reports that enhanced geomagnetic activity and limbic seizures produce religious senses of a "felt presence" and that meditation contributes to intrusive experiences.

In contrast, Eugene G. d'Aquili and Andrew B. Newberg (1999) offer SPECT scans of advanced-level meditators that show changes in regional cerebral blood flow as evidence that alternate circuits of brain activity are developed during meditation. D'Aquili and Newberg discovered that during meditation there is increased activity in the frontal lobes of the brain correlated with decreased activity in the posterior parietal lobes of the brain. They claim that the result is deafferentation of the outward orientation and association areas of the prefrontal cortex, resulting in senses of spacelessness, timelessness, and selflessness typically associated with religious experiences that they characterize as Absolute Unitary Being experiences.

FIRST-PERSON AND INTROSPECTIVE STUDIES

Neurophenomenologists are examining systematic approaches to introspection as a tool of study, using both Husserlian phenomenological techniques and meditative techniques developed in Asian religious traditions, to gain insight into the psychology of religious states of consciousness. The Mind and Life Institute, working at the Keck Laboratory at the University of Wisconsin, Madison, and at the University of Paris, engages in collaborative research between Buddhist meditators and Western neuroscientists, aimed at correlating the Buddhist first-person trained experience of focused attention, open attention, visualization, and compassion, with states of neural phase-symmetry detected on high-density electroencephalography, magnetoencephalography, and functional magnetic resonance imaging. The researchers hope to show that stabilized, trained, first-person experiences of focused attention, compassion, and so on can be systematically correlated to states of neural phase-synchrony that represent states of large-scale integration within the brain.

HEALTH-INTEGRATION STUDIES

Psychologists, such as Mihaly Csikszentmihaly (1997), are studying the relationship between happiness and peak experiences of the type outlined by Abraham Maslow, and are discovering that highly engaged attitudes and relationships, of the type long encouraged by religions, are productive of happiness, or Aristotelian eudaimonia.

In this research self-sacrificing and loving relationships to work and significant others are turning out to produce both happiness and physical and mental health, despite predictions to the contrary made by psychological survivalist and egocentrist theories.

Research groups such as the John Templeton Foundation, the Metanexus Institute, and Stephen G. Post's Institute for Research on Unlimited Love are using methodologies that could be characterized as heterophenomenological to explore the health and social effects of compassionate behavior on human thriving. Also, Division 36 of the American Psychological Association has been sponsoring conferences, several journals (i.e., *International Journal for the Psychology of Religion*, *Journal for the Scientific Study of Religion*, and *Review of Religious Research*), and a newsletter cataloging its members' study of a wide variety of issues related to the clinical and psychological roles of religion in the family, in coping with illness and death, in youth violence, in gender studies, in psychotherapy, in shaping values, and in sociological group formation, among many other topics.

See also Mysticism, Nature and Assessment of; Religious Experience; Religious Experience, Arguments for the Existence of God.

Bibliography

Alper, Matthew. *The "God" Part of the Brain: A Scientific Interpretation of Human Spirituality and God*, 4th ed. Brooklyn, NY: Rogue Press, 2001.

Andresen, Jensine, and Robert K. C. Forman, eds. *Cognitive Models and Spiritual Maps: Interdisciplinary Explorations of Religious Experience*. Thorverton, U.K.: Imprint Academic, 2002.

Baars, Bernard J. *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press, 1997.

Baars, Bernard J., William P. Banks, and James B. Newman, eds. *Essential Sources in the Scientific Study of Consciousness*. Cambridge, MA: MIT Press, 2003.

Clark, Andy. *Being There*. Cambridge, MA: MIT Press, 1997.

Csikszentmihaly, Mihaly. *Finding Flow: The Psychology of Engagement with Everyday Life*. New York: Basic Books, 1997.

Damasio, Antonio R. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam, 1994.

D'Aquili, Eugene G., and Andrew B. Newberg. *The Mystical Mind: Probing the Biology of Religious Experience*. Minneapolis, MN: Fortress Press, 1999.

Dennett, Daniel C. "Who's on First? Heterophenomenology Explained." *Journal of Consciousness Studies* 10 (9–10) (2003): 19–30.

Fodor, Jerry A. *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

Forman, Robert K. C., ed. *The Innate Capacity, Mysticism, Psychology, and Philosophy*. New York: Oxford University Press, 1998.

Freedman, Anthony, ed. *The Emergence of Consciousness*. Thorverton, U.K.: Imprint Academic, 2001.

Hood, Ralph W., Jr., et al. *The Psychology of Religion: An Empirical Approach*, 2nd ed. New York: Guilford Press, 1996.

James, William. *The Principles of Psychology*. Edited by Frederick Burkhardt. Cambridge, MA: Harvard University Press, 1981.

Lutz, Antoine, and Evan Thompson. "Neurophenomenology: Integrating Subjective Experience and Brain Dynamics in the Neuroscience of Consciousness." *Journal of Consciousness Studies* 10 (9–10) (2003): 31–52.

Metanexus Institute, http://www.metanexus.net, 2004.

Mind and Life Institute. "Training and Studying the Mind: Toward and Integration of Buddhist Contemplative Practices and Neurosciences." http://www.mindandlife.org/collaboration.html, 2002.

Newberg, Andrew, et al. "The Measurement of Regional Cerebral Blood Flow during the Complex Cognitive Task of Meditation: A Preliminary SPECT Study." *Psychiatric Research: Neuroimaging Section* 106 (2) (2001): 113–122.

Nielsen, Michael E. "Psychology of Religion in the USA." *Psychology of Religion*, http://www.psywww.com/psyrelig/USA.html, February 2000.

Núñez, Rafael, and Walter J. Freeman. *Reclaiming Cognition: The Primacy of Action, Intention, and Emotion*. Thorverton, U.K.: Imprint Academic, 1999.

Panksepp, Jaak, *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press, 1998.

Pargament, Kenneth I. *The Psychology of Religion and Coping: Theory, Research, Practice*. New York: Guilford Press, 1997.

Persinger, Michael A. "Transcendental Meditation and General Meditation Are Associated with Enhanced Complex Partial Epileptic-Like Signs: Evidence for Cognitive Kindling?" *Perceptual and Motor Skills* 76 (1) (February 1993): 80–82.

Varela, Francisco J., and Jonathan Shear. *The View from Within: First-Person Approaches to the Study of Consciousness*. Thorverton, U.K.: Imprint Academic, 2000.

Varela, Francisco J., Evan Thompson, and Eleanor Rosch. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press, 1991.

Weiskrantz, Lawrence. *Consciousness Lost and Found: A Neuropsychological Exploration*. New York: Oxford University Press, 1997.

Wilber, Ken. *The Essential Ken Wilber: An Introductory Reader*. Boston: Shambhala Publications, 1998.

Laura E. Weed (2005)

RELIGION AND MORALITY

Morality is closely associated with religion in the minds of many people. When religious leaders speak out on moral topics, their opinions are often treated with special

deference. They are regarded as moral experts. This raises the question of whether morality depends in some way on religion. Many philosophers have held that it does. John Locke, for example, argued that atheists could not be trusted to be moral because they would not consider themselves obliged even by solemn oaths, much less by ordinary promises. The answer to this question may be of considerable practical importance. If morality does depend on religion, the process of secularization, in the course of which religious belief and practice wither away, seems to pose a serious threat to morality. At one time many social theorists were confident that secularization was inevitable in modern and postmodern societies. Experience has undermined this confidence. Secularization no longer appears to be an inevitable consequence of modernization. Moreover, the process seems to occur at different rates in different modern societies. Thus secularization is more advanced in some Western European societies than it is in the United States. Nevertheless, it seems reasonable to be concerned about whether morality will decline to the extent that modern societies become more secular if it is the case that morality depends on religion.

This entry discusses several ways in which morality may depend on religion. It considers causal, conceptual, epistemological, and metaphysical dependency relations. It also explores the possibility that morality and religion may come into conflict. But a fruitful discussion of how two things are related must rely on some understanding of what those two things are. Hence the entry begins with characterizations of domains of morality and religion.

MORALITY AND RELIGION CIRCUMSCRIBED

Understood in broad terms, morality consists of answers to the general normative question of how one should live one's life. It covers a wide range of topics related to the conduct of human life. Morality concerns actions that should and should not be performed and rules of conduct that should and should not be followed. It also comprehends motives for actions that people should and should not have and character traits or habits that people should and should not try to develop. Another subject of moral concern is ideals of saintliness or heroism to which some people may properly aspire, even though not everyone is called upon to live up to these ideals. Yet another subject is social and political arrangements that people should and should not strive to create or to sustain. Thus understood, morality consists of a diverse array beliefs and practices, and it is probably not possible to give an illuminating definition of its scope. Philosophers often say that the realm of morality in this broad sense coincides with the realm of the ethical.

When philosophers reflect on the contents of the ethical, they find it useful to distinguish within it two domains, each characterized by a distinctive family of fundamental concepts. One is the axiological domain. Its basic concepts are goodness, badness, and indifference. The other is the deontological domain. Its basic concepts are requirement (obligation), permission (rightness), and prohibition (wrongness). Duty is the chief subject matter of the deontological domain. Some philosophers—Bernard Williams, for example—have proposed that morality be conceived narrowly as restricted to the deontological domain. On this conception, the domain of morality is a proper subdomain of the realm of the ethical.

Discussions of whether morality depends on religion frequently focus exclusively on the deontological domain. It is not hard to see why this occurs. Deontology consists of a system of requirements, permissions, and prohibitions. It is structurally similar to systems of law. Hence it is natural to think of deontology as the domain of moral law. Once this way of thinking has been adopted, the question arises as to whether moral law's binding force depends on the authority of a divine lawgiver. Most of the discussion in this entry will address the issue of whether moral requirements (obligations) and prohibitions (wrongness) depend on a deity of the sort to which the major monotheisms of Judaism, Christianity, and Islam are committed. However, some consideration will also be given to the topic of whether axiological goodness depends on such a deity. For this reason, the narrow conception of morality—which restricts it to the deontological domain—will not be adopted in this entry.

Religion, too, consists of beliefs and practices that exhibit great diversity. Most scholars who study it doubt that the concept of religion can be defined or analyzed in terms of necessary and sufficient conditions for being a religion. Some philosophers—for instance, John Hick—take the concept of religion to be a family-resemblance concept. On this view, religions resemble one another as members of a family resemble one another. For example, the ancient cults of Moloch, Christianity, and Theravada Buddhism may be classified as religions because they resemble one another in various respects, without supposing that all three of them satisfy a single set of necessary and sufficient conditions for being a religion. A more refined version of this view is provided by accounts developed in cognitive psychology of concepts organized

around examples that serve as prototypes or paradigms. As a result of complex patterns of similarity to—and difference from—the prototypes, other cases lie at various distances from the prototypes in a similar space. Cases near the prototypes fall under the concept; cases far enough away from the prototypes do not fall under the concept. In between there may be a gray area in which can be found borderline cases.

In attempting to define the concept of religion in terms of necessary and sufficient conditions, there is often disagreement about whether commitment—in theory or in practice—to superhuman beings is a necessary condition for being a religion. A celebrated debate in anthropology nicely illustrates such disagreement. Melford Spiro made the following proposal: "I shall define 'religion' as 'an institution consisting of culturally patterned interaction with culturally postulated superhuman beings'" (Spiro 1966, p. 96). However, there is an obvious objection to Spiro's proposal. In its purest form, Theravada Buddhism does not postulate superhuman beings. Yet most scholars think that pure Theravada Buddhism counts as a religion. So Spiro's proposal fails to provide an adequate necessary condition for being a religion. It is too narrow.

Clifford Geertz (1966) offered a more complex definitional proposal. According to Geertz, "a *religion* is: (1) a system of symbols which acts to (2) establish powerful, pervasive, and long-lasting moods and motivations in men by (3) formulating conceptions of a general order of existence and (4) clothing these conceptions with such an aura of factuality that (5) the moods and motivations seem uniquely realistic" (Geertz 1966, p. 4). Theravada Buddhism will count as a religion by this definition. But so too will the system of symbols characteristic of Nazism, although most scholars wish to classify Nazism as a secular political ideology rather than as a religion—or at least to insist that it is religious only in some extended or analogical sense. Thus Geertz's proposal fails to provide an adequate sufficient condition for being a religion. It is too broad. Disagreements of this kind fuel skepticism about whether it is possible to frame an illuminating definition of the concept of religion in terms of necessary and sufficient conditions.

For historical reasons, the monotheistic religions of Judaism, Christianity, and Islam are the prototypes of religion for people brought up within European and North American cultures. Discussion in this entry will focus almost entirely on the theism that is common to these paradigmatic religions.

CAUSAL DEPENDENCE

Morality would depend historically on religion if moral beliefs and practices were derived by causal processes from prior religious beliefs and practices. It is often imagined that early human societies had worldviews in which no distinctions were drawn between moral and religious beliefs and practices. All norms of human conduct were then religious in character; their authority was taken to rest on superhuman sources such as the prescriptions of gods. Independent moral beliefs and practices emerged from such religious worldviews in the course of cultural evolution as a result natural processes of functional differentiation. Rules governing the performance of religious rituals, for example, were distinguished from norms of ordinary human social interaction. The idea that all early human societies had tightly integrated worldviews dominated by religious concerns is, of course, highly speculative. There is little direct evidence that supports it. Perhaps studies of tribal societies by anthropologists during the nineteenth and twentieth centuries lend this idea some indirect evidential support. But the inference from structural features of the worldviews of those tribal societies to structural features of the worldviews of early human societies is problematic. After all, when anthropologists encountered them, the tribal societies they studied had themselves been evolving for a long time.

Moreover, even if something such as this story of the historical origins of morality were true, it would not have important philosophical consequences. It would not establish the conclusion that human beings would never have developed morality if there had been no antecedent religion because a function of large parts of morality is to make possible human cooperation for mutual benefit. People would have encountered problems of cooperation even in the absence of religious beliefs and practices. Given human ingenuity, therefore, it is plausible to suppose that some form of moral belief and practice would have arisen in the course of human history, even if religion had never existed. Nor would history show that the truth of moral beliefs depends on the truth of religious beliefs. In general, it is fallacious to infer from the premise that one belief grew out of another that the truth of the former depends on the truth of the latter. Though modern chemistry grew out of alchemy, it is believed that modern chemistry is true, whereas alchemy is viewed as mostly false.

Morality would depend psychologically on religion if religious beliefs were causally necessary to motivate general compliance with the demands of the moral law. If

human beings are sufficiently selfish, many of them will not behave morally when the moral law requires large sacrifices from them—unless they believe that it is in the long run in their self-interest to do so. The common theistic belief that in the afterlife God rewards those who obey the moral law and punishes those who do not will thus serve to motivate compliance with the demands of moral duty. Maybe this purpose can only be effectively served by a belief that morality has the backing of a system of divine rewards and punishments in the afterlife. If this is the case, people who lack a religious belief of this kind will also lack what it takes to cause or motivate them to live up to the demands of morality when the going gets tough.

However, there are compelling reasons to think that the view of human nature on which this line of thought rests is inaccurate. Living in a social world in which many people lack belief in an afterlife, experience shows that many people are motivated to comply with the most stringent demands of morality even though they lack any belief in a system of divine postmortem rewards and punishments. It was clear to thoughtful people who inhabited social worlds—worlds in which belief in heaven and hell was nearly universal—that belief in divine punishment in the afterlife all too often did not suffice to motivate people who did believe to obey the moral law.

What is more, according to some moral theories, morality requires not only that people comply with the moral law but also that their compliance be motivated by respect for the moral law itself. For example, Kantians hold that actions that are in compliance with the moral law but are motivated by hope for rewards or fear of punishment have no moral worth, even though they are legally correct. In other words, morality demands both that people do their duty and that they do it for duty's sake. They will do the right thing for the wrong reason if their obedience to the moral law is caused by the belief that obedience will be rewarded or the belief that disobedience will be punished. On a view of this sort, religious belief in rewards and punishments in the afterlife constitutes a danger to morality; such belief may tempt people to rely on motivational factors that will deprive their actions of moral value, even when they are the actions prescribed by morality.

CONCEPTUAL DEPENDENCE

Some philosophers have maintained that concepts of moral deontology contain religious content. In a seminal paper defending a modified divine command account of wrongness, Robert M. Adams (1987, 1999) proposed a theory in which being contrary to the commands of a loving God is part of the meaning of the term *wrong* in the discourse of some Jewish and Christian theists. And in her famous attack on modern moral philosophy, G. E. M. Anscombe (1981) recommended getting rid of the concepts of moral obligation and moral duty—and the concepts of moral right and wrong—because they belong to an earlier conception of ethics that no longer survives. The earlier conception she had in mind was a law conception. In it, according to Anscombe, the ordinary terms *should*, *needs*, *ought*, and *must* acquired a special sense by being equated in certain contexts with terms such as *is obliged*, *is bound*, or *is required*, in the sense in which one can be obliged or bound—or something be required—legally. She contends that "it is not possible to have such a conception unless you believe in God as a law-giver; like Jews, Stoics and Christians" (Anscombe 1981, p. 30). In the absence of this religious belief, the concepts of moral deontology have no reasonable sense; they are not really intelligible outside a divine law conception of ethics. Modern moral philosophers who lack belief in God would therefore do well to cease using the deontological concepts in their thinking.

Anscombe realizes, of course, that some nonreligious moral theorists will wish to retain a law conception of ethics without a divine legislator. In a Kantian conception of the moral law, for example, practical reason substitutes for God in the role of moral legislator. One's own practical reason engages in self-legislation; it is the authoritative source of moral obligations. Anscombe alleges that the idea of self-legislation is absurd. She remarks: "That legislation can be 'for oneself' I reject as absurd: whatever you do 'for yourself' may be admirable; but is not legislating" (Anscombe 1981, p. 37). However, she does not offer an argument to support the charge of absurdity. Hence Kantians are in a position to take issue with her cursory dismissal of the idea of moral self-legislation.

A deflationary approach to the deontological concepts provides another nonreligious alternative to the divine law conception. According to the account of this kind proposed by Williams (1983), obligations are not always prescriptively overriding; they do not always beat out ethical considerations of all other kinds. Instead, they are constituted by considerations to which some deliberative priority is granted in order to secure reliability in human social life. High deliberative priority is, in the case of some obligations, responsive to the basic and standing importance of the human interests they serve. Such obligations are negative telling people what not to do. In the

case of positive obligations, high deliberative priority is responsive to the demands imposed by emergencies. Williams thus indicates how it is possible for nonreligious moral theory to salvage at least deflated versions of the concepts of traditional moral deontology.

Anscombe's claim that the main concepts of traditional moral deontology have theistic content is intuitively plausible. However, moral belief and practice seem capable of surviving, almost unchanged, the replacement of such concepts by successors without religious content. And nonreligious moral theorists may even welcome the deflationary features of such a replacement if it is carried out along the lines envisaged by Williams.

EPISTEMOLOGICAL DEPENDENCE

Many religious believers hold that their moral convictions acquire some positive epistemic status, such as being justified or being warranted, and thereby count as moral knowledge, by virtue of being rooted in religious sources. Among the sources widely acknowledged in theistic religions are divine revelation recorded in sacred texts, divinely inspired prophetic utterances, and the teachers of divinely guided institutions. Frequently such sources purport to reveal divine commands by means of which God promulgates moral obligations. In addition, calls from God to perform particular actions or to enter into religious vocations are taken to be revealed in individual religious experience. Perhaps the most celebrated example in the history of Christianity comes from Augustine's *Confessions*. In retrospect, he took the childish voice he heard saying "Take and read" to be an indirect communication from God, because the biblical reading he did in response served providentially to trigger his conversion to Christianity. Because they hold that these sources are reliable—at least in certain circumstances—theists suppose that their deliverances, when properly interpreted, have positive epistemic status.

Religious diversity furnishes the grounds for an objection to this supposition. Survey the entire religious scene and it becomes evident that there is enormous disagreement among religious people about which sources are reliable, as well as how to interpret the deliverances of these various sources. Consequently, theists disagree among themselves about what God has commanded, and so they disagree about what is morally required or forbidden. Such disagreement undermines the claim that religious sources confer positive epistemic status on their deliverances. Positive epistemic status for one's moral convictions can only be derived from nonreligious sources, because only they can yield agreement. Jeremy Bentham clearly articulated the epistemic asymmetry implicit in the objection. He remarked: "We may be perfectly sure, indeed, that whatever is right is conformable to the will of God: but so far is that from answering the purpose of showing us what is right, that it is necessary to know first whether a thing is right, in order to know from thence whether it be conformable to the will of God" (Bentham 1948, p. 22). In other words, people do not first come to know, from religious sources, that actions are commanded by God and then, on that basis, come to know that they are morally obligatory. Rather, they first come to know, from nonreligious sources, that actions are morally obligatory and then, on that basis, come to know that they are commanded by God.

Religious disagreement clearly does have a negative impact on the degree to which moral beliefs derive positive epistemic status from religious sources. At least for those who are sufficiently aware of it, religious diversity reduces that degree to a significant extent. After all, moral convictions would acquire a higher degree of positive epistemic status from religious sources if all the sources produced exactly the same outputs. However, nonreligious sources also yield conflicting moral judgments in pluralistic societies that tolerate free inquiry into moral issues. Anyone who is familiar with the history of secular moral theory in the modern era is apt to think it unlikely that agreement on a single moral theory will ever be achieved under conditions of free inquiry. So unless people are prepared to live with extensive moral skepticism, they should be reluctant to think that moral beliefs derive no positive epistemic status at all from religious sources merely because those sources yield conflicting deliverances.

Few people who live in religiously pluralistic societies rely exclusively on religious sources for epistemic support of their moral beliefs. Most people think the moral beliefs they form when responding intuitively to their experiences or to works of imaginative literature—or those beliefs acquired from interaction with parents and peers outside of religious contexts—often have positive epistemic status bestowed on them by nonreligious sources of these kinds. Even the religious people who inhabit such societies typically find themselves with moral convictions that stem from a plurality of sources, some religious and others nonreligious. However, unless the religious worldviews that serve to accredit their religious sources are disqualified for rational acceptance—which would be difficult to establish—religious people seem to be entitled to trust those religious sources and to regard them as conferring positive epistemic status on

their deliverances. Hence the moral convictions of religious believers apparently can, in principle, derive positive epistemic status from both religious and nonreligious sources. Bentham's view is therefore one-sided. While religious believers in pluralistic societies may acquire knowledge of what God commands by first coming to know their obligations, they may also acquire knowledge of their obligations by first coming to know what God commands. At least some of the moral convictions of such people can be epistemologically dependent on their religious beliefs and yet possess positive epistemic status. Or, at any rate, this view is more plausible than Bentham's if moral and religious skepticism is ruled out.

METAPHYSICAL DEPENDENCE

Beginning in the last third of the twentieth century, interesting ideas about how morality might depend metaphysically on God were developed and defended in the work of proponents of divine command theories of morality. In an influential paper offering suggestions to divine command theorists, William P. Alston (1990) proposed that axiology and deontology depend on God in different ways. In the axiological domain, in Alston's view, God is the paradigm or supreme standard of goodness. An analogy to the situation helps to clarify Alston's suggestion. He maintained that the meter could be defined in terms of a certain metal bar kept in Paris. What then made a particular table a meter in length was its conformity to a certain existing individual. Similarly, according to Alston, "what ultimately makes an act of love a good thing is not its conformity to some general principle but its conformity to, or approximation to, God, Who is both the ultimate source of the existence of things and the supreme standard by reference to which they are to be assessed" (Alston 1990, p. 320). There is, to be sure, a disanalogy as well. While it is arbitrary which particular physical object was chosen to be the standard meter, Alston does not suppose that it is similarly arbitrary whether God or someone else serves as the standard of goodness. Thus understood, moral axiology depends metaphysically on the nature and character of God. By contrast, within the domain of deontology, moral obligations and moral wrongness depend metaphysically on God's commands, and ultimately on the divine volitions expressed by those commands.

Alston's suggestions have been developed into a framework for ethics by Robert M. Adams. According to his theistic Platonism, God plays the role that the Form of the Good plays in Plato's metaphysics. God is the Good Itself, the standard of goodness; and other things are good by virtue of resembling or being images of God in various ways. Modifying again his modified divine command theory of wrongness, Adams has claimed that wrongness bears the metaphysical relation of property-identity to contrariety to the commands of a loving God. He asserts: "My new divine command theory of the nature of ethical wrongness, then, is that ethical wrongness *is* (i.e., is identical with) the property of being contrary to the commands of a loving God" (Adams 1987, p. 139). And in presenting his framework for ethics, Adams sometimes says that an action's being obligatory consists in its being commanded by a loving God and that an action's being wrong consists in its being contrary to the commands of a loving God. The fundamental principle of obligation of a theory of this kind asserts that actions are obligatory if and only if, and solely because, they are commanded by a loving God. Its fundamental principle of wrongness claims that actions are wrong if and only if, and solely because, they are forbidden by a loving God. The metaphysical dependency of moral deontology on God is expressed in such principles by their requirement that actions are obligatory or wrong just because a loving God commands or prohibits them.

Of course, many philosophers have mounted objections to divine command theories of morality. Perhaps the most famous objection alleges that divine command theories render moral deontology arbitrary because God could have commanded absolutely anything. Thus, for example, God could have made cruelty for its own sake obligatory simply by commanding it. A defense against this allegation is available within the framework proposed by Alston and developed by Adams. God's nature and character, which constitute the standard of goodness, constrain what God can command. Though they may well leave some room for discretion in what God commands, God cannot command absolutely anything. If God is essentially loving and so could not be otherwise, it is impossible for God to command cruelty for its own sake. Hence, according to a divine command theory of this sort, it is likewise impossible for cruelty for its own sake to be obligatory.

Divine command theories have been defended against many other objections in work by Philip L. Quinn (1978) and Edward R. Wierenga (1989). As a result, it seems that these theories are good candidates for adoption by theists. If the larger theistic worldviews in which divine command theories are embedded are themselves rationally acceptable, an account of the metaphysics of morals, according to which morality depends on God, is a live option in moral theory.

CONFLICT THREATENED AND RESISTED

The arbitrariness objection to divine command theories suggests a threat that religion may—in some cases—pose to morality. It is the possibility of a religious obligation, imposed by divine command, coming into conflict with moral duties. The possibility is ominous because the historical record is full of crusades, inquisitions, and terrorist acts perpetrated in the name of theistic religions. Those who have done such things have often sincerely believed that they act in obedience to God's will. Within Jewish and Christian traditions, reflection on this possibility frequently focuses on the Hebrew Bible's story of the *akedah*, the binding of Isaac, narrated in Genesis 22. According to the story, God commands Abraham to offer his innocent son, Isaac, as a sacrifice, and Abraham shows that he is willing to perform this terrible deed of human sacrifice. As it turns out, an angel tells Abraham that he is permitted to substitute a ram for Isaac as the sacrificial victim, but the substitution is permitted precisely because Abraham has demonstrated to God his willingness not to withhold Isaac from being killed as a sacrifice.

Johannes de Silentio, the pseudonymous author of Søren Kierkegaard's *Fear and Trembling*, argues that the story of the *akedah* reveals a teleological suspension of the ethical. De Silentio conceives of the ethical in broadly Hegelian terms. People have prima facie duties to social groups of various size. If a duty to a smaller group conflicts with a duty to a larger group, the duty to the larger group is more stringent than—and hence overrides—the duty to the smaller group.

Thus, for example, Agamemnon's familial duty not to sacrifice his innocent daughter, Iphigenia, is overridden by his political duty to lead the Greek expedition to Troy. He is a tragic hero because he sacrifices Iphigenia. However, he remains within the ethical in doing so because he does so in order to fulfill his overriding political duty. Abraham is not a tragic hero. When he consents to sacrifice Isaac, he does not do so in order to fulfill some more stringent duty to a larger social group. Were he to carry out the sacrifice, he would be violating a duty that has not been overridden within the ethical. Yet Abraham lies under an absolute religious obligation to obey God. De Silentio regards Abraham's situation as a paradox that cannot be solved by mediation. He claims: "During the time before the result, either Abraham was a murderer every minute or we stand before a paradox that is higher than all mediations" (Kierkegaard 1983, p. 66). In other words, from the time he consents to sacrifice Isaac, Abraham is a murderer in his heart unless the ethical is suspended from the outside in his case. But Abraham, whom de Silento acknowledges to be the Father of Faith, is never a murderer. Therefore the divine command to Abraham must produce a suspension of the ethical.

Many theists do not wish to accept such a radical interpretation of the *akedah*. Kant is a notable example. In an often cited footnote in *The Conflict of the Faculties*, he insists: "Abraham should have replied to this supposedly divine voice: 'That I ought not to kill my good son is quite certain. But that you, this apparition, are God—of that I am not certain, and never can be, not even if this voice rings down to me from (visible) heaven'" (Kant 1996, p. 283). Kant's strategy of resistance to radical readings of the *akedah* carries with it an epistemological price. No matter how impressive the sound effects in the sky may be, they cannot confer on the claim that the voice commanding Abraham to kill Isaac actually came from God the exalted positive epistemic status of certainty. More generally, religious sources cannot confer epistemic certainty on claims about what God has commanded that conflict with epistemically certain moral judgments. On this Kantian view, therefore, there are limits on the extent to which claims about what God commands or wills can derive positive epistemic status from religious sources. No doubt this is a price many theists will be happy to pay in order to rule out certain sorts of conflict between their religious beliefs and the moral beliefs to which they are most deeply committed.

See also Atheism; Authority; Bentham, Jeremy; Deontological Ethics; Enlightenment; Ethics, History of; Hobbes, Thomas; Kant, Immanuel; Locke, John; Mill, John Stuart; Philosophy of Education, History of; Philosophy of Law, History of; Philosophy of Religion, History of; Rashdall, Hastings; Teleological Ethics.

Bibliography

Adams, Robert M. *Finite and Infinite Goods: A Framework for Ethics*. New York and Oxford: Oxford University Press, 1999. Develops a framework in which a divine command deontology is embedded in an axiological theistic Platonism.

Adams, Robert M. *The Virtue of Faith and Other Essays in Philosophical Theology*. New York and Oxford: Oxford University Press, 1987. Reprints the 1973 paper in which his modified divine command theory of wrongness was first proposed and the 1979 paper in which that theory was further modified.

Alston, William P. "Some Suggestions for Divine Command Theorists." In *Christian Theism*, edited by Michael D. Beaty. Notre Dame, IN, and London: University of Notre Dame Press, 1990. Suggests combining an axiology in which God is

the standard of goodness with a divine command deontology.

Anscombe, G. E. M. *Ethics, Religion and Politics*. Minneapolis: University of Minnesota Press, 1981. Reprints the 1958 paper that contains her famous attack on modern moral philosophy.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation* (1789). New York: Hafner, 1948. Argues that morality is not epistemologically dependent on religion.

Geertz, Clifford. "Religion as a Cultural System." In *Anthropological Approaches to the Study of Religion*, edited by Michael Banton. New York: Praeger, 1966. Presents a definition of religion according to which the postulation of superhuman beings is not a necessary condition for being a religion.

Hick, John. *An Interpretation of Religion: Human Responses to the Transcendent*. New Haven, CT, and London: Yale University Press, 1989. Endorses a family-resemblance account of the concept of religion.

Kant, Immanuel. *The Conflict of the Faculties* (1798). In Immanuel Kant, *Religion and Rational Theology*, edited and translated by Allen W. Wood and George di Giovanni. Cambridge, U.K.: Cambridge University Press, 1996. Comments on the *akedah* in the course of a discussion of the relations of theology and philosophy.

Kant, Immanuel. *Religion within the Boundaries of Mere Reason* (1793). In his *Religion and Rational Theology*, edited and translated by Allen W. Wood and George di Giovanni. Cambridge, U.K.: Cambridge University Press, 1996. Develops his account of moral religion, which contains reflections on the *akedah*.

Kierkegaard, Søren. *Fear and Trembling* (1843). In his *Fear and Trembling/Repetition*, edited and translated by Howard V. Hong and Edna H. Hong. Princeton, NJ: Princeton University Press, 1983. Interprets the *akedah* as involving a teleological suspension of the ethical.

Outka, Gene, and John P. Reeder eds. *Prospects for a Common Morality*. Princeton, NJ: Princeton University Press, 1993. Collects papers by scholars in philosophy and religious studies that address the question of whether religion can contribute to a common morality for a pluralistic world.

Outka, Gene, and John P. Reeder, eds. *Religion and Morality*. Garden City, NY: Doubleday, 1973. Collects papers by scholars in philosophy and religious studies on the relations of religion and morality.

Quinn, Philip L. *Divine Commands and Moral Requirements*. Oxford, U.K.: Clarendon Press, 1978. Defends divine command theories against a variety of objections.

Spiro, Melford E. "Religion: Problems of Definition and Explanation." In *Anthropological Approaches to the Study of Religion*, edited by Michael Banton. New York: Praeger, 1966. Presents a definition of religion according to which the postulation of superhuman beings is a necessary condition for being a religion.

Wierenga, Edward R. *The Nature of God: An Inquiry into Divine Attributes*. Ithaca, NY, and London: Cornell University Press, 1989. Formulates and defends a divine command theory in which deontology depends metaphysically on divine agent causation.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985. Distinguishes morality from ethics and offers a deflationary account of moral obligation.

Philip L. Quinn (2005)

RELIGION AND POLITICS

Is it morally appropriate for citizens in a liberal democracy like the United States to support or oppose public policies solely for religious reasons? Although regularly serving as grist for the mill of political theorists, that question is not the familiar fare of ordinary political discussion. It's not a question about, or at least directly about, which laws our government ought to enforce. We're all too familiar with such questions—about the moral propriety and practical wisdom of abolishing the death penalty, legalizing abortion, declaring war on Afghanistan, and so on. Rather, it is a question about the kinds of *justifications* citizens should or should not have for their political commitments.

The most common position on this issue calls for a general constraint on the political use of religious reasons. Proponents of this constraint argue that citizens must support public policies for secular reasons and therefore that they morally ought to restrain themselves from supporting public policies solely for religious reasons. So, for example, they argue that a citizen who lacks any secular reason to criminalize abortion or discourage homosexuality ought to refrain from supporting any such policy.

This entry presents the main lines of argument for and against this view that citizens should obey the *doctrine of restraint*: that they ought to restrain themselves from supporting or opposing public policies solely for religious reasons.

LIBERAL DEMOCRACY, RELIGIOUS PLURALISM, AND THE DOCTRINE OF RESTRAINT

In order to understand why a given philosophical commitment is significant, it's helpful to identify the problem that that commitment is supposed to solve. This is true of the doctrine of restraint: It is significant because it claims to solve a problem that naturally arises from the institutionalization of a liberal democracy's deepest normative commitments. What is that problem?

At the very least, a liberal democracy is a form of government that affirms and protects a citizen's rights—to private property, to freedom of association, to freedom

of conscience, and so on. The most fundamental of these rights, both morally and historically, is the right to religious freedom: Citizens are entitled to decide which religious creed or practice, if any, they wish to pursue. So a liberal democracy just is a kind of political system that provides citizens with considerable leeway to decide for themselves what they are to believe regarding religious matters.

Protection of the right to religious freedom has far-reaching social consequences, the most important of which is religious pluralism: A society that assiduously protects each citizen's right to religious freedom will find its citizens disagreeing among themselves as to which religion is true, how to please God, and so on. What explains this close connection between religious freedom and religious pluralism? Because our rational capacities are not powerful enough to produce widely convincing proofs (or refutations) of religious truth claims, even the flawless employment of our rational capacities will lead to disagreements about such matters. From this claim about the limited powers of our rational capacities, it follows that citizens who are free to decide which religious tradition to affirm will embrace a diversity of religious traditions. So, then, since well-functioning liberal democracies effectively protect the right to religious freedom, and since the effective protection of religious freedom results in a citizenry that is rationally committed to divergent religious traditions, it follows that well-functioning liberal democracies will be characterized by a citizenry that is rationally committed to divergent religious traditions.

This fact of religious pluralism raises a question of enormous moral and practical import: How are the multifariously committed citizens of a liberal democracy to make collective decisions about the laws with which each citizen must comply? For advocates of the doctrine of restraint, the pervasive pluralism of a liberal democracy renders obedience to the doctrine of restraint imperative: The morally appropriate response of citizens to religious pluralism is to refrain from resolving public matters solely for sectarian, and therefore for religious, reasons.

THE DOCTRINE OF RESTRAINT

Although agreed on a core prohibition of exclusively religious support for public policies, advocates of the doctrine of restraint diverge in their formulations of this doctrine. Some have argued that it constrains all political actors, including citizens, legislators, judges, and other public officials; others limit its scope to public officials. Some have argued that the doctrine enjoins restraint with respect to all public policies, whereas others have limited its scope to coercive public policies; others further delimit its scope to laws of fundamental and structural importance. Some have argued that the doctrine of restraint applies only to public political advocacy, whereas others contend that it should also apply to political decision-making the decision as to whether to vote for some candidate, for example.

A more important variation is between the inclusive formulation of the doctrine of restraint on which this entry focuses and a more demanding but less plausible cousin with which the inclusive version is sometimes confused. The inclusive version of the doctrine of restraint has been most effectively advocated by Robert Audi (2000), who insists that citizens should *include* secular arguments in their political practice, not that they should *exclude* religious reasons. According to Audi, citizens may support only those public policies for which they have secular reasons they regard as sufficiently weighty that they would continue to support the relevant policies absent corroborating religious reasons. But as long as citizens have, and are motivated by, adequate secular reasons for a given public policy, they are free to appeal to religious reasons as well. So inclusivists such as Audi advocate a kind of *limited* privatization: Citizens are free to rely on religious reasons for public policies so long as they have correlative secular reasons but must refrain from supporting public policies for which they have only religious reasons

Other advocates of the doctrine of restraint have demanded a complete privatization of religious reasons. This *exclusive* version of the doctrine of restraint, advocated by Richard Rorty (1994), demands that citizens refrain entirely from relying on religious reasons when supporting or opposing public policies. Its advocates expect that the real business of politics will be conducted exclusively on the basis of secular argument. But this complete privatization of religious reasons is gratuitously exclusionary: There is no good reason to stigmatize citizens who support a given public policy on religious grounds if those citizens also have and are sufficiently motivated by plausible secular reasons. Any serious evaluation of the doctrine of restraint, therefore, will pay due regard to the more moderate, inclusive version articulated by Audi.

AGAINST THE DOCTRINE OF RESTRAINT

Critics of the doctrine of restraint reject even its inclusive formulation. According to Eberle, (2002), Perry (2003), and Audi and Wolterstorff (1997), citizens need not be

morally criticizable in any respect for supporting public policies for religious reasons—even solely for religious reasons. These critics of the doctrine of restraint need not and do not typically license an "anything goes" approach. They may advocate *substantive* constraints on the reasons citizens have for their favored public policies—for example, they might argue that citizens ought not appeal to reasons that deny the dignity of their fellow citizens. They may also advocate *epistemic* constraints on the manner in which citizens support public policies, arguing, for example, that citizens should engage in critical reflection on their reasons. They may even argue that citizens should *try* to articulate secular reasons for their favored political commitments: Citizens ought to do what's within their power to speak to their fellow citizens in ways that their fellow citizens can take seriously and so ought to do what's in their power to articulate reasons that speak to their secular compatriots—presumably these will be secular reasons. Critics of the doctrine of restraint argue, however, that none of these constraints provide an adequate basis for a general constraint on religious reasons; so long as a citizen satisfies the appropriate substantive and epistemic constraints, and so long as he or she genuinely searches for a plausible secular rationale, then a citizen has no good reason not to support a given public policy solely for religious reasons.

These critics argue that advocates of restraint must discharge a heavy burden of proof: Absent sufficiently powerful reasons in favor of the doctrine of restraint, citizens may refuse without compunction to comply with that doctrine. The argument for this distribution of the burden of proof is short, direct, and powerful. We surely want and expect citizens to treat their compatriots as conscience dictates: A citizen ought to support or oppose public policies on the basis of what he or she sincerely and responsibly believes to be the just and decent thing to do. And sometimes what a citizen sincerely takes to be the just and decent thing to do will depend solely on religious beliefs. And in that case, the heavy presumption in favor of acting in accord with conscience translates into a heavy presumption permitting a citizen to decide, solely on religious grounds, to support or oppose some public policy. Consider, for example, a Christian pacifist who, after sober and competent reflection on the morality of war, concludes that the life and teachings Jesus Christ forbid the lethal use of force. In this case, our conviction that citizens should support those public policies that they actually believe to be morally correct should lead us to expect—indeed, encourage—Christian pacifists to oppose war, even though they have an exclusively religious rationale for that policy.

So critics of the doctrine of restraint will appeal to the very great good of a citizen's acting in accord with her conscience to establish a heavy presumption in favor of the moral propriety of that citizen's making political decisions solely on religious grounds. But that there is a presumption against the doctrine of restraint by no means implies that that doctrine is false. After all, presumptions can be overridden, and the burden of proof can be met—so long as advocates of the doctrine of restraint can marshal sufficiently powerful arguments.

THE ARGUMENT FROM RESPECT

Some advocates of the doctrine of restraint have argued that citizens should obey the doctrine of restraint out of respect for their compatriots. When a citizen supports a public policy, she is complicit in authorizing the government to coerce citizens. But her compatriots aren't mere playthings who may be forced to satisfy her whims; rather, they're rational persons who are fully capable and desirous of deciding for themselves how they will live their lives on the basis of reasons *they* find acceptable. And so if she is to respect her compatriots as persons, she must be committed to providing them with reasons that they find, or at least can find, acceptable. That requires a search for some common ground, premises that one might *share* with one's compatriots. Given the pervasive religious pluralism of a well-functioning democracy, this common ground will most likely be secular, not religious in content. On this view, advocated by Charles Larmore (1987), it is respect for the dignity and autonomy of our fellow citizens that requires us to abide by the doctrine of restraint.

The argument from respect is both popular and controversial. Critics have expressed doubt that reliance on religious reasons—even exclusive reliance—necessarily involves disrespect for other persons. Some argue that it is unclear why any disrespect can be imputed to citizens who affirm their compatriots' dignity, who are willing to engage in critical analysis of their favored public policies, and who provide their fellow citizens with sincerely held and carefully elaborated reasons that are, nevertheless, based on religious doctrines. Consider again the Christian pacifist: There is every reason to believe that her opposition to war is based on the kind of moral commitment that putatively underlies the doctrine of restraint: respect for *all* other persons. The argument from respect seems an entirely unpromising rationale for requiring a Christian Pacifist to exercise restraint—thereby casting doubt on the doctrine of restraint insofar as it constrains on religious reasons generally.

THE ARGUMENT FROM RELIGIOUS WARFARE

Religious wars have played a defining role in the history of liberal democracies; the commitment to religious freedom was formulated and defended in reaction to a century and a half of wars fought to "resolve" religious disagreements. The specter of religious warfare lingers on, and in some cases, that wariness motivates the argument from religious warfare.

Here is one way to formulate that argument. Religious wars are morally abhorrent: Military conflicts guided by religious aims are purely destructive, extraordinarily vicious, and utterly without redeeming value. If large numbers of citizens rely solely on religious reasons to direct state coercion, there is a glaring temptation to enlist the power of the state to force conversion and persecute heretics, thereby provoking armed conflict; hence only a policy of religious restraint can ward off the specter of sectarian bloodshed. In short, that citizens firmly commit to supporting only those public policies for which they have an adequate secular rationale is a crucial bulwark protecting us from confessional conflict.

It is, however, reasonable to deny that there is a *realistic* prospect that segments of the population of the United States will enter into armed conflict over religious matters. Religious warfare is not a realistic prospect in the contemporary United States because we have learned how to prevent it and have taken the appropriate measures: The proper preventive for religiously generated strife is constitutional and cultural, viz., effective protection of religious freedom on the part of the government and commitment to religious freedom on the part of citizens. This point has direct implications for the idea that obedience to the doctrine of restraint is necessary to prevent religious war. For it implies that what's essential in preventing religious war is that citizens are fully committed to religious freedom, not that they refrain from making use of that right to support public policies solely for religious reasons. So long as citizens are firmly committed to religious freedom, their willingness to support public policies solely for religious reasons has no realistic prospect of engendering religious warfare.

THE ARGUMENT FROM PUBLIC DISCOURSE

A third argument for the doctrine of restraint, advocated by Daniel Conkle (1993–1994) hinges on the following two claims: (1) that healthy public discussion of public policies is a great moral and political good and (2) that that good would be threatened by the refusal of large numbers of citizens to abide by the doctrine of restraint. On this view citizens should not support public policies without reflecting on those policies with their compatriots; as an implicit acknowledgment of human fallibility, the pursuit of such political discourse invites our compatriots to challenge our mistaken assumptions and inherited prejudices. Moreover, a commitment to public discourse about public policies affords those who hold a minority view the opportunity to convince other citizens of good will that their minority position is in fact correct. Hence this kind of public discourse is advanced as an important moral good.

In order to secure that good, citizens must abide by a number of constraints, especially that which requires citizens to support public policies on the basis of reasons open to rational evaluation and debate. Religious reasons, by contrast, are not subject to rational analysis and thus require a nonrational, subjective act of faith that can only be experienced, not rationally analyzed or debated. On this view compliance with the doctrine of restraint is a prerequisite for healthy discourse about public policies.

Critics have pointed to a number of problems in the argument from public discourse. Insofar as advocates of the doctrine of restraint depend heavily on the argument from public discourse, they seem to rely on controversial claims about the epistemic status of religious reasons. If religious reasons are not amenable to rational criticism by others, then it follows that religious reasons lack what many regard as an important epistemic desideratum. But this demotion of the epistemic status of religious reasons likely to trouble religious citizens. As this entry noted at the outset, the primary significance of the doctrine of restraint is that it putatively provides a morally attractive guideline for a pluralistically committed citizenry to follow when supporting public policies. This implies that the doctrine of restraint should be acceptable not just to secular citizens, but to all citizens and therefore to the religious citizens who are expected to comply with that doctrine. But many religious believers are likely to regard the epistemic assessment of religious claims that underpins the argument from public discourse as thoroughly objectionable; after all, it's dubious that we should place our trust in claims, whatever their content, that aren't amenable to rational criticism by others. So the argument from public discourse recommends that religious believers exercise restraint, but on grounds that many religious believers will find deeply objectionable. It seems likely, then, that the best argument advocates of the doctrine of restraint can muster will be anathema to the very citizens

who are expected to comply with that doctrine, thus emptying the doctrine of restraint of its primary significance.

CONCLUSION

The literature on the proper role of religious reasons in liberal politics is voluminous. And so as one might expect, the preceding discussion is far from definitive (or exhaustive for that matter). But this should hardly be surprising: The problem to which the doctrine of restraint responds rests on a pluralistic social reality that results from the successful implementation of a liberal democracy's defining commitments. That social reality is here to stay, as are the problems that it engenders, and so reflective people will continue to advocate for and criticize proposed solutions to those problems. The doctrine of restraint, and its critics, will be with us for the foreseeable future

See also Democracy; Liberalism; Philosophy of Religion, History of; Political Philosophy, History of; Rawls, John; Social and Political Philosophy.

Bibliography

Audi, Robert. *Religious Commitment and Secular Reason*. Cambridge, U.K.: Cambridge University Press, 2000.

Audi, Robert, and Nicholas Wolterstorff. *Religion in the Public Square: The Place of Religious Convictions in Political Debate*. Lanham, MD: Rowman and Littlefield, 1997.

Carter, Stephen. *The Culture of Disbelief*. New York: Basic Books, 1993.

Conkle, Daniel. "Different Religions, Different Politics." *Journal of Law and Religion* (1993–1994).

Eberle, Christopher. *Religious Convictions in Liberal Politics*. Cambridge, U.K.: Cambridge University Press, 2002.

Greenawalt, Kent. *Private Consciences and Public Reasons*. Oxford: Oxford University Press, 1995.

Greenawalt, Kent. *Religious Convictions and Politics Choice*. Oxford: Oxford University Press, 1988.

Larmore, Charles. *Patterns of Moral Complexity*. Cambridge, U.K.: Cambridge University Press, 1987.

Perry, Michael. *Love and Power: The Role of Religion and Morality in American Politics*. Oxford: Oxford University Press, 1991.

Perry, Michael. *Religion in Politics: Constitutional and Moral Perspectives*. Oxford: Oxford University Press, 1997.

Perry, Michael. *Under God? Religious Faith and Liberal Democracy*. Cambridge, U.K.: Cambridge University Press, 2003.

Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.

Rorty, Richard. "Religion as a Conversation-Stopper." *Common Knowledge*. (1994).

Weithman, Paul ed. *Religion and Contemporary Liberalism*. Notre Dame, IN: University of Notre Dame Press, 1997.

Weithman, Paul ed. *Religion and the Obligations of Citizenship*. Cambridge, U.K.: Cambridge University Press, 2002.

Quinn, Philip L. "Political Liberalism and their Exclusions of the Religious." In *Religion and Contemporary Liberalism*, edited by P. J. Weithman, 138–161. Notre Dame, IN: Notre Dame Press, 1977.

Christopher J. Eberle (2005)

RELIGION AND THE BIOLOGICAL SCIENCES

HISTORY

Plato and Aristotle recognized that understanding nature demands reference to factors—what Aristotle called "final causes"—that in some sense anticipate what will or should happen. In the *Timaeus*, Plato wrote, "From the combination of sinew, skin, and bone, in the structure of the finger, there arises a triple compound, which, when dried up, takes the form of one hard skin partaking of all three natures, and was fabricated by these second causes, but designed by mind, which is the principle cause with an eye to the future." He continued, "For our creators well knew … that many animals would require the use of nails for many purposes; wherefore they fashioned in men at their first creation the rudiments of nails. For this purpose and for these reasons they caused skin, hair, and nails to grow at the extremities of the limbs" (*Timaeus*, 76d–e).

Such adaptations, organic features that demand a final-cause understanding, are the basis for (what was to prove) a very popular and longstanding proof of God's existence. The forward-looking aspect of adaptations comes from the fact that they seem as if they were designed. They are like artifacts. Why? Quite simply because adaptations are artifacts—the artifacts of a deity. Just as a couch has a couch designer, so the hand and the eye must have a hand and eye designer. There is no necessary implication that there is just one designer, or that it has the attributes of the Judeo-Christian God—eternal, all powerful, all loving, creator of all from nothing—but this Greek argument (known as the "argument from design") was taken over by the great Christian philosophers and theologians, and became one of the main supports of the route to God through reason (natural theology).

This argument continued to enjoy great popularity and force right into the nineteenth century. Archdeacon William Paley in his book *Natural Theology* (1802) pro-

moted the argument: The eye is like a telescope; telescopes have telescope makers; therefore eyes must have eye makers—what one might call the Great Optician in the sky. By this time, however, the pendulum was starting to swing the other way, with biology giving theists cause for concern. The eighteenth century saw the rise of evolutionary speculations—hypotheses that organisms are the end results of long, slow, natural processes of development from very different and much simpler forms. At the most obvious level, evolutionary ideas challenge the Genesis story of creation. But though this was certainly a stumbling block for many, believers have long had resources to deal with problems caused by literal interpretations of the Bible.

Far more threatening to the theist was the connection between organic evolution and the doctrine of intellectual or cultural progress. As humans supposedly have risen up from ignorance and poverty in the cultural world to the sophisticated state in which we humans now find ourselves, so in the world of organisms, primitive forms have developed into humans. Cultural development points to biological evolution, which in turn reinforces cultural development. To quote an early evolutionist, Erasmus Darwin (a grandfather of Charles):

> Imperious man, who rules the bestial crowd,
> Of language, reason, and reflection proud,
> With brow erect who scorns this earthy sod,
> And styles himself the image of his God;
> Arose from rudiments of form and sense,
> An embryon point, or microscopic ens!
> (1803, 1, 295–314)

All of this progressivism was a direct challenge to the Christian notion of Providence. For the believer, because of Adam's sin, we are in a fallen state. To earn us salvation in this fallen state, God intervened in his creation, choosing freely to die on the cross. This means that our happiness comes not from our merits, but simply as the result of God's forgiveness and grace. Progress challenges this. It carries the central message that improvement is possible and due entirely to human intentions and labors. Success comes from our own efforts, not from those of others—including God. As part of the picture of progress, evolution was rightly seen as challenging conventional religious verities.

Although popular in some quarters, evolution was always somewhat of a pseudoscience. As Immanuel Kant pointed out in his third critique, *The Critique of Judgment*, there are difficulties with final causes. Such a complex, apparently intentional entity as the eye simply could have come about through blind law. Charles Darwin addressed this issue in *On the Origin of Species*, published in 1859. Committed to evolution, Darwin sought a cause that would speak to adaptation. This he found in the mechanism of natural selection. More organisms are born than can survive and reproduce. This brings on a struggle for existence. Organisms tend to vary naturally, and the winners in the struggle (the fit) have features not possessed by the losers (the unfit). Moreover, these features tend to be deciding factors in whether an organism is successful or unsuccessful. Hence, equivalent to the selection practiced by animal and plant breeders, there is a natural selection, where the winners pass on their favorable features. Over time this leads to full-blown evolution, a key feature of which is the development and perfection of adaptations.

Although he himself was never an atheist—at the time of writing the *On the Origin of Species* he was a deist and later turned to agnosticism—Darwin apparently drove a stake through the heart of the argument from design. The eye resulted from blind, unguided processes through natural selection. There is no need to invoke a designer. In the words of the contemporary English biologist Richard Dawkins (1986), only after Darwin was it possible to be "an intellectually fulfilled atheist." As expected, not everyone agrees that such a conclusion follows. Below are the different positions taken on the relation of biology and religion in the post-Darwinian era.

SEPARATION

One strategy is to separate science and religion, specifically, biology and Christianity. This means that biology cannot support religion, but then again neither can it refute it. A common suggestion is that biology can tell us how things occur—that humans came from apelike creatures, for example—but it cannot tell us why things occur—why there should be creatures with the conscious ability to tell good from evil. The great English theologian John Henry Newman, an Anglican convert to Catholicism, had no trouble at all with evolution. It was simply not something that bore on his faith. "I believe in design because I believe in God; not in a God because I see design." He continued, "Design teaches me power, skill and goodness—not sanctity, not mercy, not a future judgment, which three are of the essence of religion" (Newman 1973, 97).

This kind of reversal of the argument—design because of God, rather than God because of design—found much favor in the twentieth century, particularly in circles influenced by Karl Barth, another major critic of natural theology. In the opinion of such thinkers, often

labeled "neo-orthodox," evolution is true. But this does not prove anything affecting religion. In the language of the German theologian Wolfhart Pannenberg (1993), we must strive for a "theology of nature," where the beauties of the living world enrich our faith, rather than a "natural theology," where the living world is used as a substitute for faith. Thus, Dawkins is wrong not so much in thinking that one can be an intellectually fulfilled atheist, but in thinking that this is the end of the journey. It is the beginning. Darwin shows that there can be no proofs, and that is where faith begins.

INTERACTION

Not every post-Darwinian thinker has been so negative about natural theology. Many think that Darwin's work is the spur to find a new natural theology, a natural theology that accepts evolution and works with it rather than against it. Instead of rejecting progress, Christians should take it on board in some fashion, arguing that we humans should work with God to achieve our salvation. The rise of organisms, from slime to humans, "from monad to man," as it was traditionally put, is proof that not all is random and without purpose. It shows that God is working out his plan, and also that we are obligated to work with him.

The thinker who tried most fully to work out a theology that stayed true to conventional Christian belief and yet made the upward progressive message of evolutionism central was the French Jesuit and paleontologist Pierre Teilhard de Chardin. In his masterwork *The Phenomenon of Man*, Teilhard saw life evolving upward through the realm of life (the biosphere), to the realm of humans and consciousness (the noosphere), and then even further onward and upward to the Omega Point, which in some way he identified with the Godhead, with Jesus Christ. "An ever-ascending curve, the points of transformation of which are never repeated; a constantly rising tide below the rhythmic tides of the ages—it is on this essential curve, it is in relation to this advancing level of the waters, that the phenomenon of life, as I see things, must be situated" (p. 101).

One major problem with this whole approach is less whether the post-Darwinian Christian should accept the doctrine of progress than whether the post-Darwinian evolutionist should accept such a doctrine. If natural selection is true, then change is much relativized. Which species are fit? Not necessarily those at the top of an absolute scale. Intelligence might seem a good thing, but it has major costs, not the least of which is a constant supply of quality protein. In many circumstances, stupidity and strength might be a better biological strategy. Many evolutionists now reject progress entirely. The late Stephan Jay Gould (1989), paleontologist and science writer, argued that there is no genuine progress, and certainly no guarantee that if the tape of life were replayed, humans would inevitably emerge.

This is not the last word. Darwin himself believed in progress and thought that natural selection gives rise to what biologists of 2005 label "arms races," where one line of organisms competes and improves adaptations against the threat of other lines. Intelligence is an end result. Darwin has his supporters in the early twenty-first century, notably the English paleontologist Simon Conway Morris (2003), who argues that selection leads steadily to the conquering of one major ecological niche after another. Consciousness is the prize at the top, waiting to be grasped, and if not by humans, then by some other contender with outstretched paw.

DARWINIAN OPPOSITION TO THEISM

Dawkins is an atheist. He thinks that Darwinian evolution is hardly neutral. Although the argument from evil—that the bad things of this world are incompatible with an all-loving, all-powerful god—is not new with Darwin, his theory focuses on evil and makes it a central part of the evolutionary story. For Dawkins and others, this is confirmation that the Christian God does not exist, that other forms of deity are not worth entertaining, and hence that life has no meaning, that it just is. "In a universe of blind physical forces and genetic replication, some people are going to get hurt, other people are going to get lucky, and you won't find any rhyme or reason in it, nor any justice. The universe we observe has precisely the properties we should expect if there is, at bottom, no design, no purpose, no evil and no good, nothing but blind, pitiless indifference" (Dawkins 1995, 133).

Theists have standard counters to the problem of evil (Ruse 2001). Some theists separate moral evil (the extermination of Jews at Auschwitz) from physical evil (cancer). In the case of moral evil, it is better that humans have free will, even though they will do wrong, these theists argue, than that humans have no genuine choices at all. This may or may not be an adequate response, but if one argues for the philosophical position known as compatibilism—the position that freedom and natural law are not contradictory—then an evolutionist could in principle support this defense. The fact that we humans are the product of biological law and still subject to it does not in itself deny some dimension of freedom and ability to act on our own choices.

In the case of physical evil, recourse is often made to an argument of Gottfried Wilhelm Leibniz (1646–1716), namely that such evil is an unfortunate but unpreventable consequence of a world governed by natural law. Here too the evolutionist has a defense. Somewhat paradoxically, Dawkins himself supports this counter, for he argues that if organisms were created naturally, then adaptive complexity could have been achieved only through the action of natural selection. "The Darwinian Law … may be as universal as the great laws of physics" (Dawkins 1983, 423). One might still argue that given the consequent pain, it was a pity that God created at all, but this is a different claim totally independent of evolution. From the viewpoint of biology, if God did create and did so through natural law—and there may be good theological reasons for this—then Darwinism does not refute this, but shows rather why physical pain is bound to occur.

INTELLIGENT DESIGN

Notoriously, from the beginning many American evangelical Christians have rejected all forms of evolution. The best-known clash between such Christians and evolutionists occurred in 1925 in the state of Tennessee, when the young school teacher John Thomas Scopes was put on trial for teaching evolution. As it happened, although Scopes was found guilty, his penalty was overturned on appeal, and that was the end of so-called creationism for several decades. Yet thanks to a number of dedicated fundamentalists, people who insist on taking every verse of the Bible literally, opposition to evolutionism started to grow again, particularly after the publication in 1961 of *Genesis Flood*, a work by the biblical scholar John Whitcomb and the hydraulic engineer Henry Morris defending every verse of the Bible. This led to renewed efforts to get literal biblical teachings into publicly financed American schools, and again the matter ended in court, this time in Arkansas in 1981, where it was ruled that "creation science" is religion and as such has no place in school biology classes.

More recently, those who oppose Darwinian evolution on religious grounds have been promoting a more sophisticated form of creationism. Supporters of intelligent design argue that the organic world is just too complex and tightly functioning to have been produced by natural forces. The world, particularly at the micro level, exhibits what they call "irreducible complexity," and hence cannot possibly have been the result of something like natural selection. In the words of Michael Behe, author of *Darwin's Black Box*, an "irreducibly complex biological system, if there is such a thing, would be a powerful challenge to Darwinian evolution. Since natural selection can only choose systems that are already working, then if a biological system cannot be produced gradually it would have to arise as an integrated unit, in one fell swoop, for natural selection to have anything to act on" (p. 39).

As an example of something irreducibly complex, Behe turns to the micro world of the cell and of the mechanisms found at that level. Take bacteria that use flagella, driven by a kind of rotary motor, to move around. Every part is incredibly complex, and so are the various parts in combination. For example, the flagellin (the external filament of the flagellum) is a single protein that forms a kind of paddle surface contacting the liquid during swimming. Near the surface of the cell, one finds a thickening, just as needed, so that the filament can be connected to the rotor drive. The connector is a hook protein. There is no motor in the filament, so it has to be located somewhere else. And so on. Such an intricate mechanism is much too complex to have come into being in a gradual fashion. Only a one-step process will do, and this one-step process must involve some sort of designing cause. Behe and his supporters, including the mathematician-philosopher William Dembski, are careful not to identify this designer with the Christian God, but the implication is that the designing cause is a force beyond the normal course of nature. Biology works through "the guidance of an intelligent agent" (p. 96).

Evolutionists strongly deny that there are irreducibly complex phenomena, and they strive to show that the adaptations highlighted by intelligent-design theorists could in fact have been produced by natural selection. Of course, often mechanisms as we see them today could not function if a part were removed, but this is compatible with their coming into being through blind natural law. Perhaps formerly essential but now redundant parts have been removed. Think of a stone arch. Build it without supports, and the center keystones will fall before they are secured. Build supports and then build the arch, and the completed structure will stand even after the supports are removed.

In any case, argue critics of intelligent-design theory, there are significant theological problems with the theory, which is little more than Paley's natural theology brought up to date with some modern examples. If an intelligence intervened to produce the irreducibly complex, why does the intelligence not intervene to prevent life's simple but devastating occurrences? Sometimes a simple change in the structure of DNA can have horrific effects on an indi-

vidual. Why are these sorts of occurrences not prevented? One might say that the intelligence is not interested in doing everything, but if this is true, then it at least seems that the intelligence pointed to by intelligent-design theory is far removed from the traditional Christian conception of God.

CONCLUSION

There is more debate at the beginning of the twenty-first century than perhaps at any other time about the relationship between science and religion, and in particular between biology and Christianity. It is neither static nor philosophically uninteresting.

See also Religion and the Physical Sciences.

Bibliography

Behe, Michael. *Darwin's Black Box: The Biochemical Challenge to Evolution*. New York: Free Press, 1996.

Conway Morris, Simon. *Life's Solution*. Cambridge, U.K.: Cambridge University Press, 2003.

Darwin, Charles. *On the Origin of Species*. London: John Murray, 1859.

Darwin, Erasmus. *The Temple of Nature*. London: J. Johnson, 1803.

Dawkins, Richard. *The Blind Watchmaker*. New York: W. W. Norton, 1986.

Dawkins, Richard. *A River out of Eden*. New York: Basic Books, 1995.

Dawkins, Richard. "Universal Darwinism." In *Evolution of Molecules to Men*, edited by D. S. Bendall. Cambridge, U.K.: University of Cambridge Press, 1983.

Gould, Stephen Jay. *Wonderful Life: The Burgess Shale and the Nature of History*. New York: W. W. Norton, 1989.

Newman, John Henry. *The Letters and Diaries of John Henry Newman*. Vol. 25, edited by Charles Stephen Dessain and Thomas Gornall. Oxford, U.K.: Clarendon Press, 1973.

Paley, William. *Natural Theology* (1802). Vol. 4 of his *Collected Works*. London: Rivington, 1819.

Pannenberg, Wolfhart. *Towards a Theology of Nature*. Louisville, KY: Westminster/John Knox Press, 1993.

Ruse, Michael. 2001. *Can a Darwinian be a Christian? The Relationship between Science and Religion*. Cambridge, U.K.: Cambridge University Press.

Teilhard de Chardin, Pierre. 1959. *The Phenomenon of Man*. London: Collins.

Whitcomb, John C., and Henry M. Morris. *The Genesis Flood: The Biblical Record and Its Scientific Implications*. Philadelphia: Presbyterian and Reformed Publishing, 1961.

Michael Ruse (2005)

RELIGION AND THE PHYSICAL SCIENCES

This entry is concerned with philosophical questions arising from the interaction of religion and physical science. Here the focus is primarily upon Western religious monotheism, for this is the larger religious context in which modern science arose. And among the physical sciences, the focus is on astronomy and physics.

HISTORICAL ROOTS

The relationships between physical science and monotheism have deep roots in the history of Western thought. The simple assumption that religion and science have been and remain in conflict is falsified by the historical data. Rather, more complex and interesting connections hold between religious faith and scientific understanding in at least three domains: individual scientists and scholars, social institutions, and worldviews. At the individual level, the facts are too complex for one simple view to be true all the time, or even in a majority of cases. At the institutional level, the record of religion is at best one of indifference, and at worst outright opposition to physical science. At the level of worldviews, in contrast, Western religion has helped to make modern physical science possible.

The regular pattern of astronomical events traced by ancient Babylonian astrologers and the understanding of the physical world in Greek natural philosophy and astronomy gave currency to the idea that there must be a supreme god of some sort behind the universal patterns of causes and motions in heaven and earth. As Plato argued in *The Laws*, "If the whole path and movement of heaven and all its contents are of like nature with the motion, revolution, and calculations of wisdom, and proceed after that kind, plainly we must say that it is the supremely good soul that takes forethought for the universe and guides it along that path" (bk. 10, 897c). Both Plato and Aristotle were philosophical monotheists, a view based in part on their understanding of the workings of nature.

The tradition of Greek natural philosophy continued to develop in the monotheistic traditions of Christian, Jewish, and Arabic scholarship by means of commentaries on the physical works of Aristotle. What these philosophers had in common was what we might call a *macrodesign* scientific worldview: God created the whole cosmos and sustains the principles and laws of nature that regulate physical interaction and motion. The purpose of natural philosophy (as physical science was then

called) was to investigate the primary and secondary causes sustained by the first cause. Natural philosophy did not discuss God per se as the first cause, nor did it appeal to God as an explanation for the natural phenomena of the world. God's nature was the province of theology. This division of labor aided the development of the rationality of early modern science in the European universities in the thirteenth century, the later Middle Ages, and the Renaissance.

Important to this development was the influx of the "new" Aristotelian science from Arabic sources. Combined with a Platonic-Pythagorean tradition of mathematics, this Aristotelian tradition of empirical study was assisted by voluntarism in theology and nominalism in metaphysics. This complex tradition of inquiry formed the background to the development of early modern science and made sense of a quest for empirical, mathematical laws of nature grounded in the will of the Creator. A good historical example of this combination is Jean Buridan (c. 1292–1358), a natural philosopher and one of the most honored intellectuals in Europe, who was twice elected rector of the University of Paris. In his commentary on Aristotle's *On the Heavens*, he wrote, "In natural philosophy we ought to accept actions and dependencies as if they always proceed in a natural way" (bk. 2, ques. 9; p. 423 f.). In the same question, Buridan went on to attribute the existence and design of the universe to God as first cause, but he did not appeal to God in natural philosophy.

The scientific revolution was a genuine revolution in human thought. Despite some continuity with the Middle Ages, a whole new way of seeing the world was born. The contributions of Nicolaus Copernicus, Galileo Galilei, Johannes Kepler, and Isaac Newton, for example, gave rise to a new understanding of the physical cosmos. While a macrodesign worldview did assist in the development of early modern science, there was tension at the institutional level. The Catholic Church continued to insist upon its right to judge theological truth, including the proper way to interpret the Scriptures. The Catholic astronomer Galileo Galilei (1564–1642) ran into trouble with the Congregation of the Holy Office (the Inquisition) over exactly this point. In his famous "Letter to the Grand Duchess Christina" (1615/1957), he argued as an individual scholar that the Scriptures should be interpreted in a manner consistent with the new Copernican astronomy. The Counter-Reformation authorities in Rome soon banned the work of Copernicus "until corrected," and got Galileo to agree not to publish his views except as purely hypothetical theories. When Galileo broke this agreement by publishing his *Dialogue concerning the Two Chief World Systems*, he was suspected of heresy and forced to recant publicly. This became the most famous example of institutional religion suppressing the scientific quest for truth in the physical sciences.

For the most part the Christian churches have been unconcerned with science, focusing instead on spiritual truth and religious practices. Indeed, by creating the Western university and the hospital, the Church provided indirect support for scientific research.

METHODOLOGY

By way of contemporary issues of philosophical interest, the rise in the latter half of the twentieth century of theology-and-science debates has stimulated a number of methodological questions concerning both religion and science. The question of how we know in both disciplines has given rise to philosophical investigation into the nature and limits of knowledge in physical science and academic theology. Thomas Kuhn's *Structure of Scientific Revolutions* (1962), a revolutionary work in the philosophy of science, made a lasting contribution to the dialogue between theology and science. Science, according to Kuhn, is based on tradition and on "paradigms" of shared values, rationalities, and perspectives that gave shape to each of the scientific disciplines. It is thus based on epistemic values and metaphysical presuppositions that it owns but cannot justify. Far from being a complete worldview, science depends upon these larger perspectives for its working assumptions. This overarching view brought science into closer contact with philosophy and religion, since it was no longer the domain of purely objective, empirical fact derived from logic and evidence alone.

Investigations of the different methods of theology and science has also raised issues in the philosophy of language. How language is used in both physical science and theology has highlighted the importance of analogy and metaphor for both disciplines (Barbour 1974). This is especially true in subjects that study phenomena beyond human experience or full comprehension, for example, God and quantum reality. Yet both theology and quantum physics wish to make truth claims about their subjects, and this can only happen if we allow metaphorical truth and analogical predication.

MATHEMATICAL PERSPICUITY

Contemporary physical science, going back to the days of Galileo, constantly uses mathematics to model reality. Yet mathematics is a symbolic language that humans created

over centuries but never grounded in pure logic. Why should mathematics be such a powerful tool to describe physical reality? The physicist Eugene Wigner raised this question in his oft cited essay "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" (1960). The structures of mathematics and the deep structure of the physical universe share a feature that makes physics possible. Especially in the area of quantum physics, the ability of mathematics to predict the outcome of difficult and complex experiments is a striking example of this aspect of the universe. For theoretical physicists, the beauty and elegance of the mathematical formulas of a theory has become a key indication of the truth of the theory. But why should this be so? Is there any a priori reason to believe that the structures of mathematics should describe and predict the nature of the cosmos so well? Religious faith, especially monotheism, provides an answer to this question. The rational mind that designed the cosmos set it on a mathematically well behaved path (macrodesign again). Whether this is *the* answer to the question is a matter of serious dispute. A possible naturalistic answer might point to the evolution of the brain. Human consciousness (including the ability to create mathematics) is the ultimate product of the very laws and principles of nature that we study—a fact that makes their harmony seem more reasonable, perhaps.

ASTRONOMY AND COSMOLOGY

From mathematics we now turn to astronomy. Three areas of this science have especially drawn the attention of philosophers and theologians: the age and size of the universe, big-bang cosmology, and the fine tuning of certain physical constants in a way that allows for the evolution of stars, planets, and people. This discussion requires the distinction between a universe and the cosmos. Here "universe" refers to our space-time domain. A universe is a spatially related collection of objects under a set of natural laws and principles. "Cosmos" refers to all the universes that have ever been or ever will be.

Our universe is expanding, and this implies that it had a beginning, when the volume of space was zero and physical time first began. Along with this discovery, astronomers in the twentieth century discovered how vast the universe really is. We are a very small part of a gigantic system of planets, stars, galaxies, and galactic clusters whose vast reaches boggle the mind. Just our galaxy alone consists of 100 million stars, and many of them may well have planets. How can we think of the Earth or our species as special in any way? Philosophers and scientists alike have embraced a kind of Stoic defiance against a cold, dark, empty universe in which humanity has no special place. Somehow the vastness of space and time makes humans less significant, they argue. However, this ignores the fact that the God of traditional Western religion is both eternal and omnipresent. To an infinite, unbounded deity, what difference can it make how big or old the cosmos is? Any finite being will be the same relative to the creator, namely, of limited time and size. In biblical religion, the special status of human beings comes from their capacity for a personal relationship with God, not from how big, strong, or old they are. Still, the scientific conception of our universe has forced religious scholars to rethink the interpretation of the Scriptures and their understandings of the place of humans within creation. But nothing in the size and age of the universe actually falsifies the teachings of the great world religions.

The development of the concept of an initial singularity for the entire universe is one of the fascinating stories of twentieth-century physics. Suffice it to say that reluctantly, after several decades of debate, the physics community agreed that the general structure of space-time is dynamic. While such a conception of the beginning of the universe fits very poorly with the scientific materialism common in the physics community of the twentieth century, it does fit quite well with the older macrodesign view. The problem has to do with what caused the cosmos to come into existence. Even if space and time break down at the very earliest moments of space-time, we can still point to the first instances of time (which would not have any particular metric) and ask, Where did that come from? What caused it to be? Where do the structures and laws that allow such an event to take place come from? A macrodesign worldview has an answer to these questions—not a scientific one, but a religious one. The cosmos has a creator of some kind, who must be eternal, omnipotent, and omniscient (in fact, a necessary being). Note that this answer is not physical but metaphysical. It has implications for religion as well.

Philosophers who resist this implication, such as Quentin Smith and Adolf Grünbaum (2000), are forced to suggest either that (1) the earliest prematerial phase of the first quantum field that gave rise to the big bang sprang into being from nothing at all, or that (2) we can only ask questions about things that begin to exist when there is a space-time metric to measure temporality (Grünbaum), or that (3) the cosmos was just an accidental, random event in an infinite series of random events. None of these answers is especially cogent. First, a quantum field is, after all, a kind of order. Where did this order come from? If matter is structured energy, as quantum

theory teaches, the origin of structure is the key to the question of where matter comes from. The idea that all matter sprang out of an utter nothing at all—not simply no particles, but no laws, no fields, no energies of any kind—seems rather absurd. Second, to suggest that we can only think about why things come into being when there is a temporal metric to the time in question confuses physics and metaphysics. In metaphysics, it is still perfectly natural and rational to ask where the universe and its measurable temporal passage came from (and where it came from in the first place), even if there was no physical, measurable time prior to the first event of cosmic time. Finally, to suggest that the whole cosmos is purely random seems much more like an evasion of the problem than an attempt to answer the question. To postulate an infinite number of universes (or space-time domains) only to explain the design of this one is ad hoc and violates in the most extreme way Ockham's razor, or the principle of simplicity. We should not need to be reminded that this principle is important to the rationality of both physics and metaphysics. The existence and ultimate origin of the cosmos cry out for an explanation. This final issue, however, raises the question of design, and the possibility of a "multiverse," in the fundamental structures of physical reality.

FINE TUNING, DESIGN ARGUMENTS, AND THE MULTIVERSE HYPOTHESIS

In addition to the cosmological argument (the existence of the universe as evidence for the existence of God), in the 1990s there appeared a new and powerful version of the design argument that relies on certain fundamental constants in nature. It seems that for any intelligent life (including human life) to ever evolve anywhere in the universe, the exact values of some fundamental physical constants must be so precisely fine-tuned and balanced that it boggles the human imagination. For this reason John Barrow and Frank Tipler have called these physical constants "anthropic."

This quality of fine-tuning for anthropic purposes is widespread. Stephen Hawking, for example, estimates that the initial temperature of the universe at 10^{-43} second was fine-tuned to one part in a trillion. A tiny increase would have precluded galaxies from condensing out of the expanding matter; a tiny decrease would have resulted in the collapse of the universe. Such fine-tuning is also present in two constants in Einstein's equations for general relativity: the gravitational constant and the cosmological constant. It is also found in the fine-structure constant (which regulates electromagnetic interaction), the proton-to-neutron mass ratio, the weak nuclear force, the strong force, and so forth. According to Barrow and Tippler, a 50 percent decrease in the strength of the strong force, to take another example, would make all elements necessary for life unstable.

The initial response to this problem was to develop a number of inflationary models of the big bang. According to such models (and there are many of them), matter in the very early universe (10^{-35} second) expanded faster than the speed of light but then slowed down, and this resulted in a nearly flat curvature of space and the isolation of our relatively homogeneous space-time within a larger cosmos. We should remember that these models are highly speculative and as yet have no empirical support (that is, they are mathematical and theoretical constructions). On the basis of some inflationary models, theoretical physicists have gone even father and suggested that our cosmos may be a "multiverse." In such theories, which need much further investigation in both physics and metaphysics, our universe is one space-time domain in a vast cosmos that might contain a large number of other universes. No serious astronomer or physicist suggests that there are an infinite number of universes. But there could be an extremely large number of universes in the cosmos, and the number might be potentially infinite (that is, finite at any moment of time but open to an infinite future). If we assign laws and principles of nature randomly among all these universes in the cosmos, the fact that ours is so well fine-tuned for the evolution of intelligent life seems less surprising.

But is it less surprising? Stephen Barr (2003) has argued cogently that even if there are many, many universes in the cosmos, the fine-tuning needed across the whole range of principles and laws is so great that no finite number of universes would lower the "surprise" (the probability of our universe, against a background knowledge consisting only of the truths of reason). If Barr is even close to being right, then the multiverse hypothesis does very little to make our biologically friendly universe less surprising (or more probable). Perhaps some macrodesign scientific worldview is the most rational explanation of the order of the universe. Other options are possible, of course, for those uncomfortable with belief in some kind of creator. These options include the view that epistemic probabilities are purely subjective, and that the only real probabilities are physical ones, so that one simply cannot judge probabilities for the initial conditions of a universe. Another possibility is that our probability reasoning cannot apply to a whole universe: Any universe is just as improbable (and just as probable)

as the next one. We are extremely lucky that one universe in the cosmos of multiple space-time domains is capable of bearing life. Despite these options, or perhaps because of them, philosophers, scientists, and theologians continue to find the new fine-tuning arguments of great interest.

See also Religion and the Biological Sciences.

Bibliography

Barbour, Ian. *Myths, Models, and Paradigms*. New York: Harper, 1974.

Barr, Stephen M. *Modern Physics and Ancient Faith*. Notre Dame, IN: University of Notre Dame Press, 2003.

Barrow, John D., and Frank J. Tipler. *The Anthropic Cosmological Principle*. Oxford, U.K.: Clarendon Press, 1986.

Buridan, Jean. *Ioannis Buridani Expositio et Quaestiones in Aristotelis De Caelo*, edited by Benoit Patar. Louvain, Belgium: Peeters, 1996.

Craig, William Lane, and Quentin Smith. *Theism, Atheism, and Big Bang Cosmology*. Oxford, U.K.: Oxford University Press, 1993.

Davies, Paul. *The Accidental Universe*. Cambridge, U.K.: Cambridge University Press, 1998.

Denton, Michael. *Nature's Destiny*. New York: Free Press, 1998.

Earman, John, and Jesus Mosterin. "A Critical Look at Inflationary Cosmology." *Philosophy of Science* 66 (1999): 1–49.

Galilei, Galileo. *Dialogue concerning the Two Chief World Systems, Ptolemaic and Copernican*. Translated by Stillman Drake. Berkeley: University of California Press, 1953.

Galilei, Galileo. "Letter to the Grand Duchess Christina" (1615). In *Discoveries and Opinions of Galileo*. Translated by Stillman Drake. Garden City, NY: Doubleday, 1957.

Grünbaum, Adolf. "A New Critique of Theological Interpretations of Physical Cosmology." *British Journal for the Philosophy of Science* 51 (2000): 1–43.

Guth, Alan. *The Inflationary Universe*. New York: Helix Books, 1997.

Hawking, Stephen. *A Brief History of Time*. New York: Bantam, 1988.

Kaiser, Christopher. *Creational Theology and the History of Physical Science*. Leiden, Netherlands: Brill, 1997.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Leslie, John, ed. *Physical Cosmology and Philosophy*. New York: Macmillan, 1989.

Lindberg, David C., and Ronald L. Numbers, eds. *God and Nature: Historical Essays on the Encounter between Christianity and Science*. Berkeley: University of California Press, 1986.

Manson, Neil A., ed. *God and Design*. London: Routledge, 2003.

Padgett, Alan G. "Science and Theology." In *The Encyclopedia of Christianity*, edited by Erwin Fahlbusch et al., 4: 192–198. Grand Rapids, MI: Eerdmans, 2006.

Padgett, Alan G. *Science and the Study of God*. Grand Rapids, MI: Eerdmans, 2003.

Polkinghorne, John. *The Faith of a Physicist*. Princeton, NJ: Princeton University Press, 1994.

Shea, William, and Mariano Artigas. *Galileo in Rome*. Oxford, U.K.: Oxford University Press, 2003.

Swinburne, R. G. *The Existence of God*. 2nd ed. Oxford, U.K.: Clarendon Press, 2004.

Wigner, Eugene P. *Symmetries and Reflections*. Bloomington: Indiana University Press, 1967.

Wigner, Eugene. "The Unreasonable Effectiveness of Mathematics in the Natural Sciences." *Communications in Pure and Applied Mathematics* 13 (1) (1960).

Alan G. Padgett (2005)

RELIGIOUS EXPERIENCE

Most of the philosophical work on "religious experience" that has appeared since 1960 has been devoted to its phenomenology and epistemic status. Two widely shared assumptions help account for this—that religious beliefs and practices are rooted in religious feelings and that whatever justification they have largely derives from them.

The majority of the discussions of the nature of religious experience are a reaction to Walter Stace, who believed that mysticism appears in two forms. Extrovertive mysticism is an experience of nature's unity and of one's identity with it. Introvertive mysticism is an experience of undifferentiated unity that is devoid of concepts and images; it appears to be identical with what others have called "pure consciousness"—a state in which one is conscious but conscious of nothing.

R. C. Zaehner argued that Stace's typology ignores love mysticism in India and the West. There are two types of introvertive mysticism—monistic (pure consciousness) and theistic. The latter is a form of mutual love that unites God and the mystic in an experience without images and with very little, if any, conceptual content. The most effective defense of a position of this sort is Nelson Pike's. Pike argues that the principal forms of mystical prayer in Christianity (quiet, rapture, and full union) are phenomenologically theistic. He defends his analysis against William Forgie, who denies that the identification of the experience's object with God can be part of its phenomenological content.

Phenomenological analyses of religious consciousness presuppose that we can distinguish descriptions of religious experience from interpretations. Ninian Smart proposed two tests for distinguishing descriptions—that the accounts be autobiographical and that they be relatively free from doctrinal concepts. The question of crite-

ria remains vexed, however (see Wainwright, 1981, chap. 1).

Others have argued that, because religious experience is significantly constituted by the concepts, beliefs, expectations, and attitudes that the mystic brings to it, attempts to distinguish interpretation from description are misguided. For example, an influential article by Steven Katz contends that a mystic's experiences are largely shaped by his or her tradition. This has two consequences. First, there are no "pure" or "unmediated" mystical experiences and, second, there are as many types of mystical experiences as there are traditions.

Katz's "constructivism" has been attacked by Robert Forman and Anthony Perovitch among others. Since pure consciousness is devoid of content, it is difficult to see how it could be constituted by contents that the mystic brings to it. To argue that it must be mediated because all experience is mediated begs the question; on the face of it, pure consciousness is a counterexample to the thesis in question. Forman also argues that constructivism cannot adequately account for novelty—the fact that the mystic's experiences are often unlike what he or she expected.

Defenses of religious experience's cognitive validity have taken several forms. William Wainwright argues that mystical experiences are presumptively valid because they are significantly similar to sense experiences. Both experiences have what George Berkeley called "outness"—the subject has the impression of being immediately presented with something transcending his or her own consciousness. Corrigible and independently checkable claims about objective reality are spontaneously made on the basis of both types of experience. There are tests in each case both for determining the reality of the experience's apparent object and for determining the genuineness of apparent perceptions of it. The nature of the tests, however, is determined by the nature of the experiences' alleged objects. Since the apparent objects of religious experience and ordinary perceptual experience differ, so too will the tests for veridical experiences of those objects.

Richard Swinburne's defense of religious experience's cognitive validity is based on the principle of credulity, which roughly states that apparent cognitions are innocent until proven guilty. This is a basic principle of rationality; without it we would be unable to justify our reliance on memory, sense perception, and rational intuition. The principle implies that there is an initial presumption in favor of how things seem to us, although this presumption can be overridden. What is true of apparent cognitions in general is true of religious experiences. They too should be accepted in the absence of good reasons for thinking them deceptive. Swinburne argues that there are none.

The most sustained defense of religious experience's epistemic credentials is William Alston's. Whereas Wainwright and Swinburne concentrate on perceptual (or perception-like) experiences, Alston focuses on perceptual practices. Doxastic (belief-forming) practices are basic when they provide our primary access to their subject matter. The reliability of a basic doxastic practice like memory cannot be established without circularity; any attempt to justify it relies on its own outputs. Alston argues that sense-perceptual practice and "Christian mystical practice" are epistemically on a par. Since both doxastic practices are basic, neither's reliability can be established without circularity. Both practices are socially established, internally consistent, and consistent with the outputs of other well-established practices. They are also self-supporting in the sense that they have the outputs we would expect them to have if they were reliable (successful predictions in the first case, for example, and moral and spiritual improvement in the second). Alston concludes that it is unreasonable to engage in sense-perceptual practice while rejecting the rationality of engaging in Christian mystical practice. The rationality at issue, however, is not epistemic. Neither practice can be shown to be epistemically rational, since it is impossible to establish their reliability without circularity. Alston intends to show only that it is practically or pragmatically rational to engage in them, although it should be noted that engaging in them involves accepting their outputs as true and therefore believing that they are reliable. Alston concedes that the existence of competing mystical practices weakens his case but denies that it destroys it. Critiques of Alston's work have tended to focus on this point (see, for example, Hasker, 1986).

The most significant attacks on religious experience's cognitive validity to have appeared since 1960 are Wayne Proudfoot's and Richard Gale's. Proudfoot argues that an experience's noetic quality should be identified with its embedded causal judgment (that the experience is caused by a tree, for example, or by God) and this judgment's affective resonance. The incorporated causal judgment has no intrinsic authority; it is merely one hypothesis among others and should be accepted only if it provides a better overall explanation of the experience than its competitors'. While the causal hypotheses embedded in religious experiences could be correct, they are in fact suspect; they appear to be artifacts of the subject's religious or cultural tradition and not products of nonnatural causes.

Proudfoot's identification of an experience's noetic quality with an incorporated causal judgment and its affective resonance is more plausible in some cases than others. Given my background knowledge, I believe that a certain sort of pain in one's tooth is caused by cavities. Believing this, and having a pain of that sort, I spontaneously form the belief that my pain is caused by a cavity. While my pain is not noetic, the experience as a whole is, since it incorporates a causal judgment. But the experience lacks "outness." It thus differs from sense perception, which (because of this quality) seems to have an intrinsic authority that noetic experiences like my toothache lack. Religious experiences are also diverse. Some, like my toothache, involve spontaneous causal attributions and nothing more. Others, however, are perception-like and have the same claim to intrinsic authority that sense perceptions do.

Richard Gale, on the other hand, argues that religious experience lacks the authority of sense experience. The only way of establishing religious experience's cognitivity is by showing that the tests for it are similar to those for sense experience. Arguments for religious experience's cognitive validity fail because the dissimilarities are too great. Alston and Wainwright contend that these dissimilarities can be explained by differences in the experiences' apparent objects. Gale objects that explaining the disanalogies does not explain them away and that there is a "tension" or "inconsistency" in claiming that the tests are similar (as they must be if the defense of religious experience's cognitivity is to be successful) and yet different in nature. The first point is dubious. Only relevant disanalogies count. The point of Wainwright's and Alston's explanations is to show that the disanalogies are not relevant—that is, that the features that tests for sense experiences have and tests for religious experiences lack are not ones we would expect the latter to have if religious experiences were veridical perceptions of their apparent objects.

Gale's most original (and controversial) contribution is his contention that veridical experiences of God are conceptually impossible. The argument is roughly this: Talk of veridical experiences is in place only where it makes sense to speak of their objects as existing "when not actually perceived" and as being "the common object of different" experiences of that type. Sense experiences exhibit this feature because their objects are "housed in a space and time that includes both the object and the perceiver." Religious experiences do not exhibit this feature because there are no "analogous dimensions to space and time" that house both God and the perceiver. Gale attempts to establish this by refuting P. F. Strawson's claim that a "no space world … of objective sounds" is conceptually possible. We could neither reidentify sounds in such a world nor distinguish between numerically distinct but qualitatively identical ones. It would make no sense, therefore, to speak of sounds as the common objects of distinct auditory experiences or as existing when unperceived. Talk of veridical experiences of objective sounds would thus be out of place. A fortiori, talk of veridical experiences would be out of place in a nonspatial and nontemporal world. Therefore, since no common space (and, on some accounts, no common time) houses God and the mystic, talk of veridical perceptions of God is inappropriate.

A few general observations about discussions of religious experience since 1960 are in order. First, most defenses of religious experience's cognitive validity have been offered by theists. Stace is one of the few who has attempted to establish the veridicality of pure consciousness and other nontheistic experiences that lack intentional structure. Second, philosophical discussions of religious experiences tend to abstract them from the way of life in which they occur and thereby impoverish our understanding of them. Whether this penchant for abstraction adversely affects the discussion of phenomenological and epistemological issues, however, is more doubtful. Finally, a philosopher's assessment of the cognitive value of religious experience is affected by his or her metaphysical predilections. For example, those who assign a low antecedent probability to theism will demand stronger arguments for theistic experiences' cognitive validity than those who do not. One's assessment of religious experience cannot be separated from one's general assessment of the relevant religious hypotheses.

See also Alston, William P.; Berkeley, George; Constructivism and Conventionalism; Constructivism, Moral; Intuition; Memory; Mysticism, History of; Mysticism, Nature and Assessment of; Perception; Philosophy of Religion; Rationality; Stace, Walter Terence; Strawson, Peter Frederick.

Bibliography

Alston, W. P. *Perceiving God.* Ithaca, NY: Cornell University Press, 1991.

Davis, C. F. *The Evidential Force of Religious Experience.* Oxford: Clarendon Press, 1989.

Forgie, W. J. "Pike's Mystic Union and the Possibility of Theistic Experience." *Religious Studies* 30 (1994): 231–242.

Forgie, W. J. "Theistic Experience and the Doctrine of Unanimity." *International Journal for Philosophy of Religion* 15 (1984): 13–30.

Forman, R. K. C. "Introduction: Mysticism, Constructivism, and Forgetting." In *The Problem of Pure Consciousness,* edited by R. K. C. Forman. New York: Oxford University Press, 1990.

Gale, R. *On the Nature and Existence of God.* Cambridge, U.K.: Cambridge University Press, 1991. Chap. 8.

Gutting, G. *Religious Belief and Religious Skepticism.* Notre Dame, IN: University of Notre Dame Press, 1982. Chap. 5.

Hasker, W. "On Justifying the Christian Practice." *New Scholasticism* 60 (1986).

Katz, S. T. "Language, Epistemology, and Mysticism." In *Mysticism and Philosophical Analysis,* edited by S. T. Katz. London: Sheldon Press, 1978.

Perovitch, A. N., Jr. "Does the Philosophy of Mysticism Rest on a Mistake?" In *The Problem of Pure Consciousness,* edited by R. K. C. Forman. New York: Oxford University Press, 1990.

Pike, N. *Mystic Union: An Essay in the Phenomenology of Mysticism.* Ithaca, NY: Cornell University Press, 1992.

Proudfoot, W. *Religious Experience.* Berkeley: University of California Press, 1985.

Smart, N. "Interpretation and Mystical Experience." *Religious Studies* 1 (1965): 75–87.

Stace, W. T. *Mysticism and Philosophy.* Philadelphia: Lippincott, 1960.

Swinburne, R. *The Existence of God.* Oxford: Clarendon Press, 1979. Chap. 13.

Wainwright, W. J. *Mysticism: A Study of Its Nature, Cognitive Value, and Moral Implications.* Madison: University of Wisconsin Press, 1981.

Yandell, K. E. *The Epistemology of Religious Experience.* Cambridge, U.K.: Cambridge University Press, 1993.

Zaehner, R. C. *Concordant Discord.* Oxford: Clarendon Press, 1970.

William J. Wainwright (1996)

RELIGIOUS EXPERIENCE, ARGUMENT FOR THE EXISTENCE OF GOD

Arguments from Religious Experience show remarkable diversity, (*a*) in the sorts of experience taken as data for the argument, (*b*) in the structure of the inference itself, and (*c*) in the alleged conclusion, whether to a vague Presence, an Infinite Being, or the God of traditional Christianity.

The following exemplify some versions of the argument:

"At very different times and places great numbers of men have claimed to experience God; it would be unreasonable to suppose that they must all have been deluded."

"The real argument to God is the individual believer's sense of God's presence, the awareness of God's will in tension and conflict with his own will, the peace that follows the acceptance of God's command."

"Experiences of meeting God are self-authenticating: They involve no precarious chain of inference, no sifting of rival hypotheses. They make unbelief logically absurd."

"In itself, religious experience is neither theistic nor pantheistic, Christian nor Buddhist. All these distinctions are *interpretations* of the experience. By itself, religious experience testifies to something far less definite but still infinitely valuable—the insufficiency of all materialisms and naturalisms."

If we compare any of these arguments with the Ontological, Cosmological, and Teleological arguments, important differences in their logic and history can readily be shown. Arguments from Religious Experience are clearly not a priori, like the Ontological Argument, and whereas the Cosmological and Teleological arguments work from premises that affirm highly general facts about the world (that it exists, that it is purposefully ordered), Arguments from Religious Experience rely on far more particularized and elusive premises than these. Not all men have (or are aware of having) distinctively religious experiences, and to those that do have them religious experiences are apt to be short-lived, fugitive sets of events that are not publicly observable.

Despite this slipperiness, the Argument from Religious Experience has attracted some theologians who have been skeptical about the more rationalistic "proofs." In the course of the eighteenth century these proofs received formidable criticism from Immanuel Kant and David Hume. The Ontological Argument was shown to be radically confused over the logic of "existence," and (in Kant's account) the Cosmological and Teleological arguments themselves presuppose the Ontological. Even more important, Kant and Hume together produced a general weakening of confidence that any survey of the observable cosmos (including "the starry heavens above") could yield premises powerful enough to argue to an infinite, unconditioned, all-good deity. Kant turned to "inner" experience, to our awareness of the moral law, and argued that the moral life is intelligible only if we postulate God and immortality.

Although a number of writers followed Kant in arguing from inner moral experience, many others, while accepting the shift from outer to inner, based their inference on a distinctive class of religious experiences. If we describe this shift, in general terms, as a move from objective to subjective, from surveying the world at large for evidences of God to focusing attention on the personal

and existential, it clearly was a shift of the greatest moment and one that still helps to determine our contemporary climate of theological thought. We human beings are not stars or electrons—the argument goes—and we cannot experience or guess the role of star or electron in the divine economy. But we are persons, and we are directly aware (or some of us are) of a meeting of person with Person in religious experience.

Thinking back, however, to the post-Humean, post-Kantian period, the centering upon inner experience can be seen as one aspect of the romantic movement's protest against the Enlightenment, the new concern for subjectivity, the life of the emotions and intuitions of the individual. The most important and most seminal single figure here is Friedrich Schleiermacher (1768–1834), with his bold insistence on the primacy of religious feeling—particularly the feeling of utter creaturely dependence—and his distaste for religious doctrines or arguments entertained in a purely intellectual manner, as mere ideas, lacking the life and authority of experience.

OBJECTIONS TO THE ARGUMENT

Prima facie it seemed a reasonable and empirically sound enterprise to establish arguments for God upon claims to have actually experienced him, to have "seen" him, "met" him, encountered him in a personal relation. But there are in fact several directions from which it can be challenged.

Orthodox and neoorthodox theologians tend to object that the content of religious experience is too indeterminate to yield clear knowledge of the God of Christianity. The case for Christianity must not be allowed to rest on the deceptive and elusive emotions of religious people. It rests on the revealed Word of God, on the Person of Jesus Christ as disclosed in the Scriptures, not as constructed out of the assorted emotions of the devout. The working of the Holy Spirit cannot be correlated with the experiencing of peculiar feelings, even uplifting ones.

A second familiar objection is that although we certainly do have religious experience, we cannot employ it as the premise of an argument to God. The relationship between man and God—an I-Thou (in Martin Buber's phrase), personal relationship—is maintained by faith alone. The conception of superseding faith through a proof of God's existence forgets the irreducibly personal nature of encounter between man and God.

The objector may be making a religious claim, that it is religiously improper to attempt to replace faith by rational argument, or his point may be a logical one, that God—being "pure" person, having nothing bodily or thinglike in his nature—cannot be shown to exist in the way things can be shown to exist.

Suppose, again, we take the Argument from Religious Experience as an explanatory hypothesis; then a skeptical critic may deny that the existence of God is the likeliest, or simplest, or most intelligible, explanation of the experiences. We cannot be intellectually compelled to posit God if more economical and naturalistic explanations can be found—psychoanalytic accounts, it might be, or accounts in terms of individual suggestibility or the influence of religious expectations or tradition.

Last, a critic may concentrate on the conceptual difficulties in the idea of God, for if the argument as a whole is to be sound, its conclusion ("therefore God exists") must be intelligible and free of inner contradictions. This objection may bewilder and disappoint the arguer-from-experience. To him one of the chief apparent advantages in the argument is that its direct appeal to experience bypasses logical or metaphysical complexities. But some element of interpretation, and therefore some application of concepts, must take place when an experience is taken to be an encounter with God. Wherever concepts are handled, they can also be mishandled. Inner contradictions in the claim to experience God could invalidate the interpretation of the experience.

NATURE OF RELIGIOUS EXPERIENCES

What, more exactly, are religious experiences? Descriptions of religious experiences can be heavily loaded with doctrinal, even sectarian, interpretation or can be almost entirely free of it. Their impact may fix one's attitudes and evaluations for a lifetime or for only a brief period. They may not only be benign and optimistic, as we have so far assumed, but can also—with no less intensity—be pessimistic and grim. They may involve conversions to a religious orthodoxy or conversions away from one. Consider the following experiences, neither of which is more than minimally interpreted, and both of which are certainly in an important sense religious. The first is from Lev Tolstoy's *War and Peace,* at the point where Prince Andrew has been wounded in the Battle of Austerlitz.

> He opened his eyes, hoping to see how the struggle of the Frenchmen with the gunners ended.… But he saw nothing. Above him there was now nothing but the sky—the lofty sky, not clear yet still immeasurably lofty, with grey clouds gliding slowly across it. "How quiet, peaceful and solemn, not at all as I ran," thought Prince Andrew—"not as we ran shouting and fight-

> ing.… How was it I did not see that lofty sky before? And how happy I am to have found it at last! Yes! All is vanity, all falsehood except that infinite sky. There is nothing, nothing but that. But even it does not exist, there is nothing but quiet and peace. Thank God!"

The second is from Leonard Woolf's autobiographical work, *Sowing* (1960). At the age of eight, the author was sitting in a garden enjoying the fresh air after a train journey. He watched two newts basking in the sun.

> I forgot everything, including time, as I sat there with those strange, beautiful creatures, surrounded by blue sky, sunshine, and sparkling sea. I do not know how long I had sat there, when, all at once, I felt afraid. I looked up and saw that an enormous black thunder cloud had crept up and now covered more than half of the sky. It was just blotting out the sun, and, as it did so, the newts scuttled back into their hole.… I felt something more powerful than fear, once more that sense of profound, passive, cosmic despair, the melancholy of a human being eager for happiness and beauty, powerless in face of a hostile universe.

Turning to theistic types of experience, we can start from the very basic experience of wonderment, notably wonderment at there being any world at all. This may pass into the sense that the world owes its existence to, and is maintained in existence by, something "beyond," "outside" the world itself, a Being whose nature is utterly remote from the world, yet whose activity and energy are perceptible within the world, as a disturbing, awesome, and thrilling presence. Rudolf Otto's concept of the "numinous" gathers together these ingredients of mystery, dread, and fascination and emphasizes very properly the qualitative distinctiveness and elusiveness of such experience (*The Idea of the Holy,* passim). No set of categories can neatly contain it: The person who has never known it can barely understand the claims of the person who has.

Religious experiences can be generated by perceptions of individual objects (a grain of sand, a bird), by a train of events, by actions—for instance, the memorable account of Jesus setting his face to go to Jerusalem to his Passion. Even a passage of philosophical reasoning may do this, as when someone contemplates the incompleteness of all explanation, the intellectual opacity of space and time, and feels compelled—with a sense of mystery—to posit a divine completeness and unity.

Closer to the province of morality are experiences of divine discontent, interpreted as intimations of God's existence and call to moral endeavor, the conviction of sin correlative to a sense of God's own holiness, the sense of divine aid in the rectifying of one's moral life, and, in Christian evangelical terms, a sense that one has been redeemed or saved by God's action on man's behalf.

The overall impression is of the immense diversity of religious experiences. They are indeed linked by complex webs of "family resemblances" (to use Ludwig Wittgenstein's phrase)—resemblances of attitude, emotional tone, alleged content—but if we ask what all of them have in common, the answer must be meager in content: perhaps only a sense of momentous disclosure, the sense that the world is being apprehended and responded to according to its true colors. What is actually being observed or contemplated can never (logically) be the whole universe, yet the quality of religious experience is such that it does seem to imply something about the whole.

EPISTEMOLOGICAL STATUS OF RELIGIOUS EXPERIENCES

Our sampling of religious experiences may help to deliver us from the dangers of oversimplification, but it cannot by itself determine whether arguments to God based upon them are valid. Clearly, not all the experiences we have mentioned could yield data with which a theistic argument could start. Some, such as that of Woolf quoted earlier, are decidedly *anti*theistic. But there is a further set of differences among them that must be noted at this stage, differences of an epistemological kind.

When someone speaks of his religious experience, he may be using the word *experience* as it appears in such phrases as "business experience," "driving or teaching, etc., experience." He has found the religious pattern of life viable; he has interpreted a multiplicity of events in its categories, and these categories have proved durable. There is the suggestion that the person with religious experience in this sense has been confirming his faith by living it out over a substantial stretch of his life—furnishing data for a pragmatic proof of God's existence.

In other cases the experiences are of much shorter duration, often judgments or quasi perceptions accompanied by certain religious emotions, alleged cognitive acts or intuitions in which the necessity of God's existence is "seen" and an awesome emotive response is elicited simultaneously. Again, the language used may be nearer to that of perceiving—*seeing God* (not just seeing *that* God exists). There is a claim to knowledge of God by "acquaintance," rather than "description."

Some cases resemble the dawning of an aspect or interpretation, as when we recognize a person in a poor light or make out a pattern in what looked like a maze of lines. It can be like a sudden reading of the expression on a face, the face, as it were, of the universe, or like a realization of meaning, as when one sees the point of a poem with which one has long been verbally, but only verbally, familiar. In the light of this disclosure, a new orientation and purposeful organization of life may take place. Energies hitherto dissipated or in mutual conflict are rallied and integrated.

Feelings or emotions may predominate in religious experience, but even so, perception and judgment are almost always involved as well. Feelings are often "feelings *that* …," surmises, and in that sense feelings involve judgment, have an essential component of belief. Part of what it is to have an emotion is to see and appraise one's situation in a particular way. ("I feel remorse for doing *x*," for example, presupposes "I did *x* freely" and "*x* was morally wrong.") It is only with twinges, frissons, aches, and such like that no appraisal of the situation need (logically) be made; these, in any case, could furnish only very weak premises for a theistic argument. Their occurrence can be due to a great variety of causes immanent in one's own organism and one's environment, and they can hardly, without supplementation, force one to posit a transcendent cause.

Obviously the structure (and maybe the validity) of an Argument to God from Religious Experience will vary enormously according to what epistemological type of experience is taken as the starting point, and in the literature this is often hard to discern.

VERIFIABILITY OF RELIGIOUS EXPERIENCE

If someone claims to have discovered, perceived, become aware of an ordinary sort of object, we usually know what to do about checking his claim. If we are told that there is a frog in the garden pond, we know what it will be like to confirm this or to find it untrue. We know how to investigate whether it was Smith we saw in the dim light, whether we did hit the right answer to a sum or cried "Eureka" too soon. But when someone claims to have direct awareness of God, to encounter, see, or intuit the divine, we are not able to suggest a test performance of an even remotely analogous kind. The more developed and theologically sophisticated the conception of deity is, the more it eludes and resists any such check.

This being so, some critics have pointed out a disturbing resemblance between claims to experience God and a certain other range of statements that are not publicly testable—namely, psychological statements such as "I seem to hear a buzzing noise," or "I seem to see a patch of purple." If statements like these cannot be refuted, it is only because they make no assertions about what exists, beyond the experiences of the speaker at the moment he speaks. But the person who says he has direct and certain experience of God wishes to claim irrefutability and to affirm at the same time something momentous about what exists. Can this be done? Or would it take a far more elaborate and many-stranded apologetic to give effective backing to these claims—especially the claims to objectivity?

One might try to obtain this support by compiling records of numerous experiences of the same general kind and treating them as cumulative evidence for the truth of claims to experience God. Without doubt there is an impressive mass of such records within the Judeo-Christian tradition. Other religious traditions, however, can also produce their own very different records—of the various well-ordered phases in the quest for nirvana or for mystical union with a pantheistic object of worship.

Are these differences, however, real incompatibilities; do they correspond to genuinely different sorts of religious experience? Or are the experiences basically the same, though differently interpreted? On this it is extremely hard to give any confident answer. Part of the difficulty is that most of the developed religions contain several strands in their conceptions of the divine. Christianity, for instance, seeks to unite numinous and mystical views of God: God is "remote" and "other," yet also mystically "near." What can be said again is that any common elements must be very indeterminate in content and able to bear great variety of interpretation—to be taken, among other things, as the disclosure of a state or spiritual goal (nirvāṇa) or of a personal or suprapersonal God. We have seen how an experience may have a minimal—quite undoctrinal—interpretation and yet be religious in a broad sense. But from such an experience alone one can hardly infer anything so definite as the God of theism. Unfortunately, the interpretations that supplement the experience are conceptually intricate and involve all the uncertainty and fallibility of philosophical and theological speculation. In this region we are far removed from the ideals of immediacy, directness, and self-evidence.

Yet a critic who claimed that the Argument from Religious Experience was thereby refuted would be missing the mark. The theist could insist that a much too crude notion of "interpretation" has so far been used, one

that suggests, falsely, that there is a merely external and almost arbitrary relation between having and interpreting an experience.

The full impressiveness of the theistic case appears only when we survey the historical development of religious experience in the direction of Christian monotheism. As the idea of deity evolves, from finite and local numen to infinite and omnipresent Lord, from the god of a tribe or nation to the Ruler of all nature, from the deity concealed in holy tent or temple to the one God beyond all phenomena whatever, religious experience is itself simultaneously transformed. It is transformed not haphazardly but so as to produce a crescendo of numinous intensity, a constant refining away of merely superstitious and idolatrous awe at objects unworthy of worship, and the arrival of a distinctive, lofty note of adoration. Experience and interpretation here advance in indissoluble unity. It is argued that this historical development provides material for a more adequate argument to God—one in which the risks of fantasy and subjectivism are much reduced.

Impressive this is, and it may well be the truth of the matter. We must notice, however, that we are now looking at a much more complex piece of argument than the claims of individuals to have direct experience of God. New logical problems appear at several points. Can we be confident, for instance, that an intensification of numinous experience is necessarily a sign that we have a more adequate disclosure of God and not simply that we have constructed a more adequate and awesome idea of God? (This is the question that also calls in doubt any purely pragmatic philosophy of religion.) Again, sometimes an artist, or a school of artists, succeeds in progressively clarifying and intensifying an original vision or the expression of some distinctive emotion. But success in this ("now he has brought the theme to full explicitness," for example) is not necessarily correlated with a progress in discovery about the world. Can we be sure that the development of numinous awareness is different in this vital respect?

The person with theistic religious experience is assured that it is different. But the sense of assurance, the "Aha!" experience, the penny dropping, the light dawning—these are very unreliable guides to truth, validity, or value. Not the most tempestuous sense of poetic inspiration can guarantee that a good poem is being brought to birth, nor can any of these conviction-experiences by itself authenticate its related judgments. It is enough to recall how often incompatible judgments are made with equal assurance on each side. Yet it is not easy to formulate a version of the Argument from Religious Experience that does not rely crucially on a sense of conviction. Even when appeal is made to the pattern of development toward theism, and thus to a far wider range of phenomena than in any argument from the experience of an individual, still the issue of objectivity—that we are coming to know God, not simply an idea of God—seems to hang upon the fallible, illusion-prone assurance of the subjects. On the other hand, to point this out is to draw attention only to the risk, not to the certainty, of being wrong. A religious person may realize, and be prepared to accept, this measure of risk.

Could we escape the uncertainty, by claiming that genuine experience of God is necessarily followed by a godly life, whereas illusory experiences betray themselves by the absence of any practical fruits? Hardly; there might well be a positive correlation between genuine experiences and godliness, but in fact they are not necessarily related. Lapses, moral failures, are always open to human beings, and one cannot rule out by definition the possibility of a man's being both morally remarkable and atheistic.

But, one might argue, is the situation vis-à-vis God any worse in principle than the situation vis-à-vis material objects, such as tables and chairs? Our traffic with these consists in having actual experiences (visual, tactual, etc.) and ordered expectations of future and possible experiences. Where our experience has this sort of structure and can thus be the subject of intelligible discourse, we confer on it the status of objectivity without more ado. But theistic experience certainly occurs, and it too has its structure of expectations.

If we can bring out the difference between these cases (and the peculiar difficulty of the religious case), we shall be showing more clearly than hitherto that the Argument from Religious Experience is most intimately involved in problems of logic and meaning—problems that at first seem alien to its empirical appearance. With a material object (say, a cube) there are quite intricate but intelligible ways in which we come to see it as a single object out-in-the-world. It is given unity most obviously by possession of perceptible limits and boundaries and by the manner in which its several surfaces can be seen and felt to connect with one another. Moreover, we have mastered the laws of perspective and so can anticipate and understand variations in our perceptions of the object, owing to our own variable positions as observers. Such variations do not, therefore, impugn the assertion that the object exists in the world external to us.

With God, who is not a finite material object, there can be no inspecting of boundaries or surfaces. And if part of what we mean by "God" is "an infinitely and eternally loving Being," no conceivable experience or finite set of experiences could by itself entitle us to claim that we had experienced such a being. We might well report experiencing "a sense of immense benevolence toward us," "a sense of complete safety and well-being," but from their intensity alone one could not rigorously conclude, "Therefore I am in touch with an infinitely and eternally loving God." From the intensity of a human love one cannot infer, "This love will endure," and without bringing in a supplementary doctrine of God's attributes (not derived from experience) one could no more legitimately do so in the religious case.

Material objects, of course, are sometimes observed in unfavorable perceptual conditions—at a great distance, half-concealed, and so on. Imagination must "fill in" the perceptual gaps as best it can, until conditions improve. Analogous thought models are indeed employed in theological discourse, but they are peculiarly difficult to assess. The Christian theologian is normally most ready to admit that we can neither perceive nor imagine how the various attributes of God unite in a single being (if he is to be called "a being" at all). A fair measure of agnosticism here is compatible with full Christian belief. But it may not be compatible with a reliance upon an Argument to God from Religious Experience, if this is one's chief apologetic instrument. Unless the principles that confer unity and objectivity are able to be collected from the experiences themselves (which seems not to be possible), we have to look elsewhere for them, and the argument is in this respect shown to be inadequate. But it is not, on that account, proved useless, for if it cannot demonstrate the existence of God unaided, it might still function as a necessary auxiliary of other arguments—for example, the Cosmological Argument.

One might be forced to a deeper agnosticism than that to which we have just alluded—deeper in that it dares to affirm scarcely anything at all about the focus (or focuses) of religious experience, whether personal or impersonal. Yet with a minimal ontological commitment it might still set great value on certain religious experiences and seek after them. The attempt to work out a coherent and systematic theological interpretation would be quite abandoned.

This would save something, but assessing just how much to expect from a religious agnosticism like this would be a difficult task. The bigger the area of agnosticism, the smaller the area of legitimate religious expectations, such as that of ultimately seeing God "face to face" or of being received by him into glory. As we have more than once observed, the relation between experience and what the experience is taken to be is a most intimate one; the experiences of a Christian and those of a religious agnostic could both be valuable but could not be identical.

PSYCHOLOGICAL EXPLANATIONS

Is it not more enlightened, however, to deny that these experiences really disclose anything about the world? Psychoanalytic research has, after all, revealed many situations in which interior mental events are projected upon the world and are furnished with all the assurance of objectivity, the full sense of "givenness." One does not have to accept the entire Freudian account of religion to see plausibility in its central claim that early parent-child relations of "creaturely" dependence and reverence, with their tensions between love and fear, can yield the unconscious material from which experiences of God-man relations are fashioned. To accept this claim is not necessarily to reduce all religion to neurosis or worse. For it is absurd to class together the person who attains a stable religious solution to his conflicts and the person who retreats to genuine neurosis, developing, say, obsessions, compulsions, or delusions of persecution. Sigmund Freud certainly went further in his naturalistic explanation of religious experience, being prepared to reduce God to an illusory parent substitute. It may be possible, however, to invert the Freudian account of religious experience and, instead of seeing God as a father substitute, to see human fathers as God substitutes and the human experience of love as training for loving God. The close psychological relation between love of man and love of God would thus have its skeptical sting removed. It may be argued, again, that naturalistic and Christian explanations are compatible: God may elicit from us an effective response to his existence without making use of anything but our natural human equipment of senses, desires, emotions. Even mechanisms of projection can be involved and the projected image of deity be yet a trustworthy symbol of a God who does in fact exist. It is clear from all this that depth psychology does not provide a self-sufficient, decisive refutation of theism.

Nonetheless, depth psychology troubles and disturbs the Arguments from Religious Experience, and so do the very attempts to reconcile it with Christian belief. These virtually admit that the religious experiences might occur much as they actually do occur—without there being a God—in other words, that naturalistic explanations are

possible. There seems no way, at the experiential level, of settling the really urgent questions, most of all the following: Do we have in theistic experience mere projection? Or do we have a projection matched by an objectively existing God?

See also Agnosticism; Buber, Martin; Cosmological Argument for the Existence of God; Enlightenment; Hume, David; Kant, Immanuel; Mysticism, Nature and Assessment of; Ontological Argument for the Existence of God; Otto, Rudolf; Popular Arguments for the Existence of God; Religious Experience; Schleiermacher, Friedrich Daniel Ernst; Teleological Argument for the Existence of God.

Bibliography

WORKS ON THE PSYCHOLOGY OF RELIGION

The following works have important implications for the Argument from Religious Experience:

Freud, Sigmund. *The Future of an Illusion* (1927). Translated by W. D. Robson-Scott. London: Hogarth Press, 1928; paperback ed., Garden City, NY, 1957.

Freud, Sigmund. "Obsessive Acts and Religious Practices" (1907). In *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* translated by James Strachey, ed., in collaboration with Anna Freud. 24 vols. New York: Macmillan, 1953–1963.

Freud, Sigmund. *Totem and Taboo* (1913). Translated by A. A. Brill. New York: New Republic, 1927. Also translated by James Strachey. New York: Norton, 1952.

James, William. *Varieties of Religious Experience.* London: Longman, 1902.

Jung, C. G. *Psychology and Religion.* New Haven, CT: Yale University Press, 1938; paperback ed., 1960.

Philp, H. L. *Freud and Religious Belief.* London: Rockliff, 1956.

Thouless, R. H. *An Introduction to the Psychology of Religion.* Cambridge, U.K.: Cambridge University Press, 1923; reprinted, 1936.

RELEVANT THEOLOGICAL WORKS

Raven, C. E. *Natural Religion and Christian Theology.* Vol. II. Cambridge, U.K.: Cambridge University Press, 1953.

Schleiermacher, Friedrich. *Der christliche Glaube.* Berlin, 1821–1822. 2nd ed. translated by H. R. Mackintosh and J. S. Stewart as *Christian Faith.* Edinburgh: T. and T. Clark, 1928.

Schleiermacher, Friedrich. *Die Religion. Reden an die Gebildeten unter ihren Verächtern.* Berlin, 1799. Translated by John Oman as *On Religion: Speeches to Its Cultured Despisers.* London: K. Paul, Trench, Trubner, 1893; reprinted, Gloucester, MA, 1958; paperback ed., New York: Harper, 1958.

Tennant, F. R. *Philosophical Theology.* Vol. I. Cambridge, U.K.: Cambridge University Press, 1928.

PHILOSOPHICAL STUDIES

The following works explore the relation between religious experience and religious belief or contain relevant ideas and arguments:

Buber, Martin. *Ich und Du.* Leipzig, 1923. Translated by R. Gregor Smith as *I and Thou.* Edinburgh, 1937.

Campbell, C. A. *On Selfhood and Godhood.* London: Allen and Unwin, 1957. On Rudolf Otto see Lecture 16.

Hick, John, ed. *Faith and the Philosophers.* New York: St. Martin's Press, 1964.

Holland, R. F., and J. M. Cameron. "Religious Discourse and Theological Discourse." *Australasian Journal of Philosophy* 34 (1956): 147–163, 203–207.

Hook, Sidney, ed. *Religious Experience and Truth.* New York: New York University Press, 1961.

Lewis, H. D. *Our Experience of God.* London: Allen and Unwin, 1959.

Lewis, H. D., and C. H. Whiteley. "The Cognitive Factor in Religious Experience." *PAS,* Supp. 29 (1955): 59–92.

Martin, C. B. *Religious Belief.* Ithaca, NY: Cornell University Press, 1959.

Otto, Rudolf. *Das Heilige.* Breslau: Trewendt and Granier, 1917. Translated by John W. Harvey as *The Idea of the Holy,* 2nd ed. London, 1950.

Smart, Ninian. *Reasons and Faiths.* London: Routledge and Paul, 1958.

Smith, N. Kemp. "Is Divine Existence Credible?" *Proceedings of the British Academy* 17 (1931).

OTHER RECOMMENDED TITLES

Alston, William. *Perceiving God: The Epistemology of Religious Experience.* Ithaca, NY: Cornell University Press, 1991.

Bagger, Matthew. *Religious Experience, Justification, and History.* Cambridge, U.K.: Cambridge University Press, 1999.

Clouser, Roy. *Knowing with the Heart: Religious Experience and Belief in God.* Downers Grove, IL: InterVarsity Press, 1999.

Davis, Caroline. *The Evidential Force of Religious Experience.* New York: Oxford University Press, 1989.

Ellwood, Robert. *Mysticism and Religion.* Englewood Cliffs, NJ: Prentice Hall, 1980.

Forsthoefel, Thomas. *Knowing beyond Knowledge: Epistemologies of Religious Experience in Classical and Modern Advaita.* Burlington: Ashgate, 2002.

Gellman, Jerome. *Experience of God and the Rationality of Theistic Belief.* Ithaca, NY: Cornell University Press, 1997.

Griffith-Dickson, Gwen. *Human and Divine: An Introduction to the Philosophy of Religious Experience.* London: Duckworth, 2000.

Gutting, Gary. *Religious Belief and Religious Skepticism*. Notre Dame: University of Notre Dame Press, 1982.

Mackie, J. L. *The Miracle of Theism.* Oxford: Clarendon Press, 1982.

Miles, Thomas. *Religious Experience.* London: Macmillan, 1972.

Pine, Nelson. *Mystic Union: An Essay in the Phenomenology of Mysticism*. Ithaca, NY: Cornell University Press, 1992.

Proudfoot, Wayne. *Religious Experience.* Berkeley: University of California Press, 1985.

Rowe, William. "Religious Experience and the Principle of Credulity." *International Journal for Philosophy of Religion* 13 (1982): 85–92.

Smith, John. *Experience and God.* New York: Oxford University Press, 1968.

Swinburne, Richard. *The Existence of God.* Oxford: Clarendon Press, 1979.

Wainwright, William. *Mysticism.* Brighton: Harvester Press, 1981.

Yandell, Keith. *The Epistemology of Religious Experience.* Cambridge: Cambridge University Press, 1993.

Ronald W. Hepburn (1967)
Bibliography updated by Christian B. Miller (2005)

RELIGIOUS INSTINCT

See *Common Consent Arguments for the Existence of God*

RELIGIOUS LANGUAGE

Utterances made in religious contexts are of many sorts. In the performance of public and private worship men engage in acts of praise, petition, thanks, confession, and exhortation. In sacred writings we find historical records, dramatic narratives, proclamations of law, predictions, admonitions, evaluations, cosmological speculations, and theological pronouncements. In devotional literature there are rules of conduct, biographical narratives, and introspective descriptions of religious experience. Philosophical discussions of religious language have concentrated on a restricted segment of this enormous diversity, namely, theological statements, that is, assertions of the existence, nature, and doings of supernatural personal beings.

There are two reasons for this emphasis. First, the crucial problems about religious language appear in their purest form in theological statements. If we consider a petitionary prayer or a confession, what is puzzling about it is not the act of petition or confession, but the idea of addressing it to God, and God answering it. It is the concept of communication with a supernatural incorporeal person that seems unclear. And this lack of clarity is most apparent in the statement that there exists a God who communicates with men in various ways. We may say that the difficulties in understanding other forms of religious language all stem from obscurities in statements about God.

The second reason for philosophical concentration on theological statements lies in the fact that the philosophy of religion is primarily concerned with questions of justifiability, significance, and value. And it has generally been supposed that whether religion is a justifiable form of human activity largely depends on whether there are sufficient grounds for accepting the theological statements on which it is based. Christianity is a justifiable institution if and only if we are warranted in accepting the proposition that the world is created and governed by an omnipotent, perfectly good personal deity who has revealed himself to men in the Bible. Thus the philosophy of religion is largely taken up with examining the grounds of religious statements. And it is when we do this that we become most acutely aware of the puzzling aspects of religious language. When we make a determined effort to decide whether it is true that God created the physical universe, it is difficult to avoid realizing how unclear what we are saying is, what implications it has, what it logically excludes, and what would count for or against it. Thus the philosophical investigation of religious language focuses on those indeterminacies in theological statements that hamper attempts to find rational grounds for acceptance or rejection.

MEANING OF THEOLOGICAL PREDICATES

Most philosophers who have concerned themselves with the problem have located the difficulties of religious language in the predicates of theological statements. (What does "good" mean in "God is good"?) It may seem that we should start with the subject of the statement, with the concept of God. But there is really no alternative to starting with the predicates. For the only way to make clear what one means by "God" is to provide an identifying description, such as "the creator of the universe"; and to understand that phrase one must understand the predicate "created the universe" as applied to God. Theological predicates can be divided into negative (infinite, nontemporal, incorporeal) and positive. The positive predicates can be concerned either with attributes (good, wise, omniscient) or with actions (makes, forgives, speaks, watches over). Negative predicates present no special difficulty, but in themselves they are clearly insufficient to give any positive conception of the deity. Of the positive attributes we shall concentrate on attributions of action, partly because action terms pose more severe problems, partly because other attributes are dependent on them. (To say that God is wise is to say that he acts wisely; if we cannot understand what it is for him to perform one or another action, we cannot understand the attribution of wisdom to him.)

DERIVATION AND APPLICATION. When one reflects on the use of predicates in theological statements one comes to realize two fundamental facts: (1) this use is necessarily derivative from the application of the predicates to human beings and other observable entities; (2)

the theological use of predicates is markedly different from the application of predicates to human beings.

Theological predicates are derivative primarily because it is impossible to teach theological language from scratch. How would one teach a child what it means to say "God has spoken to me" without first making sure that the child knows what it is for a human being to speak to him? In order to do so one would have to have some reliable way of determining *when* God was speaking to him, so that when this happens one could say to him, "*That* is what it is for God to speak to you." And even if we admit that God does speak to people from time to time, there is no way for one person to tell when God is speaking to another person unless the other person tells him, which would require that the other person have already mastered the theological use of language. Hence there is no alternative to the usual procedure of teaching the theological use of terms by extension from their application to empirically observable objects.

As for the difference in the use of predicates as applied to God and to human beings, there are many ways of seeing that the terms cannot have quite the same meaning in both cases. If, as in classical Christian theology, God is conceived of as not in time, then it is clear that God's performance of actions like speaking, making, or comforting is something radically different from the temporally sequential performance of actions by human beings. St. Thomas Aquinas in his famous discussion of this problem based the distinction between the application of predicates to human beings and the application of predicates to God on the principle that God is an absolute unity and that, therefore, various attributes and activities are not distinguishable in God as they are in men. But even if we allow God to be temporal and straightforwardly multifaceted, we are left with the corporeal-incorporeal difference. If God does not have a body, it is clear that speaking, making, or comforting cannot be the same thing for God as for man.

This leaves us with a serious problem. We must show how the theological use of these terms is derived from their nontheological use. Until we do, it will be unclear just what we are saying about God in such utterances. The usual way of dealing with this problem is by cutting out the inapplicable portions of the original meaning of the terms, leaving the remainder for theology. Thus, since God is incorporeal, his speaking cannot involve producing sounds by expelling air over vocal cords. What is left is that God does something that results in the addressee having an experience of the sort he would have if some human being were speaking to him. The nature of the "something" is deliberately left vague. Since God is a pure spirit, it will presumably be some conscious mental act; perhaps an act of will to the effect that the addressee shall have the experience of being told such-and-such. More generally, to attribute any interpersonal action to God is to attribute to him a purely mental act that has as its intended result a certain experience, like the one that would result from such an action on the part of a human being.

This account may throw some light on the content of statements about God, but religious thinkers have become increasingly dissatisfied with it. For one thing, it represents theological statements as metaphysical speculations and does little to illuminate the ways they fit into religious activity. Having postulated a pure immaterial substance performing mental acts that, miraculously, have effects in human experience, how do we go about getting into communication with this immaterial substance? Why should it be worshiped at all, and if it should, why in one way rather than another? Moreover, this line of reasoning is not helpful in our efforts to verify theological statements. It offers no hints on how we might determine whether our statements are true, or even whether there is such a being that performs the actions in question.

VERIFIABILITY OF THEOLOGICAL STATEMENTS

Recent discussions have concentrated on the problem of verifiability. In the last few decades a great many philosophers have come to accept some form of the "verifiability theory of meaning," according to which one is making a genuine factual assertion, a real claim as to the way the world is, only if it is possible to conceive of some way in which what he is saying can be shown to be true or false by empirical observation. Applying this theory to theology, it has been argued that since an empirical test is in principle impossible to carry out for statements about a supernatural incorporeal personal deity, these statements cannot be regarded as straightforward factual assertions, but must be interpreted in some other way.

John Wisdom in his influential essay, "Gods," analogizes the function of theology to the following situation. Two people return to a long-neglected garden and find some of the old flowers still surviving among the weeds. One suggests that some gardener has been caring for the plot, and the other expresses doubt about this. On investigation, it turns out that no one in the vicinity has ever noticed anyone working on the garden. Moreover they discover that gardens left to their own devices often take

this form. But the first man does not abandon his hypothesis. Instead he expresses his belief that someone who is not discernible by the senses comes and cares for the garden, carrying out designs he and his companion do not fully grasp. At this point the first man has modified his "gardener" hypothesis to the point at which it is no longer susceptible to empirical confirmation or refutation. No matter what is or is not discovered empirically, he will continue to hold it. In this case it seems plausible to say that he is no longer expressing a belief about actual objective events. If he were, he would be able to imagine, however inadequately, some way in which the existence or nonexistence of these events would be revealed to our experience. He is, rather, expressing a "picture preference." It is rewarding to him to think of the situation as if a gardener were coming to take care of the flowers. If beliefs about God are equally refractory to empirical test, it would seem to follow that they too must be interpreted otherwise than as straightforward matters of fact. (Wisdom, however, does not commit himself to this conclusion.)

In considering the "verificationist" challenge to theology, we must scrutinize both premises of the argument: (1) theological statements are not susceptible to empirical test; (2) if they are not empirically testable they cannot be construed as factual assertions that can be assessed as true or false.

ARE THEY EMPIRICALLY TESTABLE? The question of whether theological statements are subject to empirical test is quite complicated. If we rule out mystical experience as a means of observation, then it is clear that statements about God cannot be tested directly. But science is full of hypotheses about unobservable entities—electromagnetic fields, social structures, instincts—which verificationists accept as meaningful because they can be tested indirectly. That is, from these hypotheses we can draw implications that can themselves be tested by observation. The question is whether directly testable consequences can be drawn from theological statements. We can phrase this question as follows: Would we expect any possible observations to differ according to whether there is or is not a God? It would clearly be unreasonable to require of the theologian that he specify a set of observations that would conclusively prove or disprove his assertions. Few, if any, scientific hypotheses could meet that requirement. The most that could reasonably be demanded is that he specify some observable states of affairs that would count for or against his assertions.

One thing that makes this problem difficult is the fact that on this point religious belief differs at different times and places. Supernatural deities have often been thought of as dealing in a fairly predictable way with contingencies in the natural world and human society. Thus in many primitive religions it is believed that the gods will bring abundant crops or victory in battle if they are approached in certain ways through prayer and ritual. Even in as advanced a religious tradition as the Judeo-Christian, it is believed that God has certain fixed intentions that will result in prayers being answered (when made in the right spirit and under proper conditions) and will result in the final victory of the church on earth.

It would seem that such expectations provide a basis for empirical test. Insofar as they are fulfilled, the theology is confirmed; insofar as they are frustrated it is disproved. However, things are not that simple. Even in primitive communities such tests are rarely allowed to be decisive; the empirical implications are hedged around with a variety of escape clauses. If the ritual dances are held and still the crops fail, there are several alternatives to abandoning traditional beliefs about the gods. Perhaps there was an unnoticed slip somewhere in the ritual; perhaps devils were conducting counterrituals. More sophisticated explanations are employed in the more advanced religions. For example, God will answer prayers, but only when doing so would be for the true good of the supplicant.

Moreover, as science develops, religion comes to be more concerned with the personal life of the worshiper and less concerned with prediction and control of the course of events. Among religious intellectuals today such predictions as are still made are clearly not testable in practice, because of their lack of specificity ("all things will work together for the good for those who love God"), their enormous scope ("everything in the world contributes to the development of moral personality"), or their inaccessibility ("after death we shall see God face to face"). Nevertheless, it seems that within religion there are strong barriers to completely divorcing belief in God from the expectation of one event rather than another; and so long as there is some connection of belief with testable predictions, however tenuous, it would be a mistake to think of religious statements as absolutely unverifiable in principle.

ARE THEY ASSERTIONS OF FACT? As to whether a statement that cannot be empirically tested must not be construed as an assertion of fact, a theologian might well challenge the application of the verifiability theory to the-

ology. If God is supernatural, we should not expect his behavior to be governed by any laws or regularities we could hope to discover. But then we could never be certain that, for example, the statement that God loves his creatures would ever imply that a war should have one outcome rather than another. This would mean that, according to the verifiability theory, it would be impossible for us to make any statements, even false ones, about such a being. But a theory that would prevent us from recognizing the existence of a certain kind of entity, if it did exist, would be an unreasonable theory.

NONASSERTIVE INTERPRETATIONS

Be that as it may, a number of philosophers have been so impressed by these difficulties over verifiability that they have tried to construe theological utterances as something other than straightforward factual assertions. Attachment to the verifiability theory is not the only motivation behind the development of such theories. There are those, like George Santayana, who, without holding that theological sentences are factually meaningless, are convinced that as factual assertions they are false, but still are unwilling to abandon traditional religious discourse. They feel that somehow it has a valuable function in human life, and in order to preserve it they are forced to reinterpret it so that the unwarranted factual claims are expunged. Still another motivation is the hope that this will contribute to the resolution of the problem mentioned earlier, that of specifying the way predicates are used when they are applied to God. As we saw, attempts to give an illuminating definition of theological predicates have not been wholly successful, and this can be taken to indicate that a different sort of approach is needed.

One such line of investigation takes sentences as its units rather than words. It focuses on the kind of linguistic act performed when theological sentences are uttered, rather than on the meaning of words in theological contexts. Instead of asking what "forgives" means when applied to God, we ask what linguistic action is performed when one uses the sentence "God forgives the sins of those who truly turn unto him." It is this sort of question one is asking when one wonders whether theological sentences make factual assertions and, if not, what they are used to do. If we could answer this question we would have made sufficiently clear how words are being used in theological sentences without having to define special senses for constituent words.

Nonassertive interpretations can be divided into four groups. Statements about God have been interpreted as (1) expressions of feelings of various sorts; (2) symbolic presentations of a variety of vital aspects of experience, from natural facts to moral ideals; (3) integral elements in ritualistic worship; (4) a unique kind of "mythical" or "symbolic" expression, not reducible to any other use of language.

EXPRESSIONS OF FEELING. Theological utterances have been interpreted as expressions of feelings that arise in connection with religious belief and activity. Thus we might think of "God made the heavens and the earth" as an expression of the sense of awe and mystery evoked by grandeurs of nature; of "God has predestined every man to salvation or damnation" as an expression of a pervasive sense of helplessness; and of "God watches over the affairs of men" as an expression of a sense of peace, security, at-homeness in the world. This is "poetic" expression rather than expression by expletives. It is like expressing a sense of futility by saying "life's a walking shadow" rather than like expressing futility by saying "Ah, me." That is, the feeling is expressed by depicting a situation that might naturally evoke it; a sense of security, for instance, is evoked by some powerful person looking after one.

SYMBOLIC PRESENTATIONS. Symbolic interpretations of religious doctrines have been common for a long time. The story of Noah and the Flood has been regarded by many Christian thinkers not as an account of actual historical occurrences, but rather as a symbolic way of presenting certain religiously important points—that God will punish the wicked, but will also, under certain conditions, show mercy. Many of the traditional ways of speaking about God have to be taken as symbolic. God cannot literally be a shepherd or a rock. The shepherd functions as a symbol of providence and the rock as a symbol for God's role as a refuge and protection in time of trouble. A symbol in this sense is some (relatively) concrete object, situation, or activity that can be taken to stand for the ultimate object of discourse through some kind of association, usually on the basis of similarity. We speak symbolically when what we literally refer to is something that functions as a symbol.

In the traditional use of symbolic interpretation it is, necessarily, only a part of theological discourse that is taken as symbolic. For if we are to hold that the symbolic utterances are symbolizing facts about God, we will have to have some way of saying what those facts are; and we cannot make that specification in symbolic terms, on pain of an infinite regress. But we are now considering views according to which all theological discourse is symbolic, which means that if we are to say what is being

symbolized it will have to be something in the natural world that can be specified in nontheological terms. The most common version of such a view is that theological utterances are symbolic presentations of moral ideals, attitudes, or values. This position has been set forth most fully and persuasively by George Santayana, and in a more up-to-date form by R. B. Braithwaite. According to Santayana every religious doctrine involves two components: a kernel of moral or valuational insight, and a poetic or pictorial rendering of it. Thus the doctrine that the physical universe is the creation of a supremely good personal deity is a pictorial rendering of the insight that everything in the world is potentially usable for the enrichment of human life. The Christian story of the incarnation, sacrificial death, and resurrection of Jesus Christ is a way of making the point that self-sacrifice for others is of supreme moral value. It is worthwhile embodying these moral insights in theological doctrine because this vivid presentation, together with the systematic cultivation of feelings and attitudes that accompanies it, provides a more effective way of getting across the insights than would a bald statement.

The way in which interpretations of the first two kinds throw light on the theological use of predicates is analogous to the way in which one explicates the use of words in poetic metaphors. If we consider the metaphor in "sleep that knits up the raveled sleeve of care," it is clear that "knit" is not used simply to refer to a certain kind of physical operation. This utterance has quite different kinds of implications from "she knit me a sweater," in which *knit* does have its usual sense. In the metaphoric statement, *knit* is used in its usual sense to depict a certain kind of situation that, as a whole, is presented as an analogue of the effect of sleep on care. The only way of effectively getting at the function of the word *knit* is by seeing how the whole phrase "knits up the raveled sleeve" is used to say something indirectly about sleep.

In the first two of the four kinds of nonassertive interpretation we are examining, theological statements are essentially metaphors. And if they are correctly so regarded, we get nowhere if we extract the word *made* from the sentence "God made the heavens and the earth" and try to say what it means by itself. What we have to do is take the picture presented by the whole sentence and see how it functions as a way of expressing a feeling of security, or as a way of presenting the insight that everything in the world can be used to enrich human life.

RITUALISTIC INTERPRETATION. The ritualistic interpretation of theological discourse can best be introduced by citing the reply of an intellectually sophisticated high-church Anglican to a question from an agnostic friend. The question was, "How can you go to church and say all those things in the creed?" The reply: "I don't say them; I sing them." In the view under consideration, the corporate practice of worship is the native soil from which talk about God springs. Talk about the attributes, doings, and intentions of a supernatural personal being has meaning as a part of the practice of worship and is puzzling only when it is separated from that context. If we think of an utterance like "God made the heavens and the earth" as the expression of a belief about the way things in fact originated and then wonder whether it is true or false, we will be at a loss. To understand it we have to put it back into the setting where it (or rather a second-person correlate, such as "Thou, who hast made the heavens and the earth") does its work. In that setting, these words are not being used to explain anything, but to do something quite different.

Unfortunately, proponents of this view have never been very clear about what this "something different" is. The clearest suggestion they give is that the talk about God serves to provide an imaginative framework for the conduct of worship. It articulates one's sense that something important is going on, and it helps to indicate the appropriateness of one response rather than another. In speaking of the sacrament of communion as the reenactment of the self-sacrifice of an omnipotent personal deity who took on human form, and in conceiving of it as a cleansing and renewing incorporation of the substance of such a deity, one provides for the activity a pictorial framework that records and nurtures the felt solemnity of the occasion and the attitudes and aspirations kindled by the ceremony. This position presupposes, contrary to the usual view, that ritual worship has an autonomous value, apart from any theological foundation. It is generally supposed that a given ritual has a point only if certain theological doctrines are objectively true. But in the ritualistic interpretation, theological doctrines are not regarded as statements about which questions of truth or falsity are properly raised. Since these doctrines depend for their significance on the ritual, it is supposed that the ritual has some intrinsic value in forming and giving expression to valuable sentiments, feelings, and attitudes.

MYTHS. Ernst Cassirer has developed the notion that the basis of religious discourse lies in a unique "symbolic form" that he terms "mythical." He maintains that it is found in purest form in the myths of primitive peoples and is based on a way of perceiving and thinking about the world that is radically different from our accustomed

mode. In the "mythical consciousness" there is no sharp distinction between the subjective and the objective. No clear line is drawn between symbol and object, between wish and fulfillment, between perception and fantasy. Again, no sharp distinction is made between the object itself and the emotional reaction it evokes; emotional response is taken to be an integral part of the environment. As a result none of our familiar standards of truth or objectivity are applicable. What is most real is what arouses the greatest intensity of emotional response and, particularly, what is felt as most sacred. (The sacred-profane distinction is the fundamental contrast.) The mythical consciousness carries its own special organizations of space and time. For example, there is no distinction made between a position and what occupies it; every spatial position is endowed with a qualitative character and exerts influence as such.

It is the view of Cassirer, and of followers such as Susanne Langer, that sophisticated theology represents an uneasy compromise between mythical and scientific modes of thought, and as such cannot be understood without seeing how it has developed from its origins. It is basically a mythical view of the world, given a "secondary elaboration" in a vain attempt to make it acceptable to the rationalistic consciousness; judged by rationalistic standards it is not only groundless, but meaningless.

Mysticism. Philosophers and theologians in the mystical tradition have put forward versions of this fourth kind of interpretation that do not regard theology as a manifestation of cultural lag. To the mystic the only way to communicate with God is through mystical experience, and this experience reveals God to be an ineffable unity. He can be directly intuited in mystical experience, but since there are no distinctions within the absolute unity of his being, and since any statement we can make predicates of him one thing rather than another, for example, wisdom as distinguishable from power, no statement can be true of him. The most we can do in language is to direct our hearers to the mode of experience that constitutes the sole means of access. Proponents of this view sometimes speak of theological language as "symbolic," but this differs from our second type of theory in that here there is no way to make explicit what it is that the theological utterances symbolize, and it is therefore questionable whether we should use the term *symbol.* A symbol is always a symbol of something. In fact it is difficult to make clear just what, on this view, religious utterances are supposed to be doing. They are said to "point to," "adumbrate," or "indicate" the ineffable divine reality, but all too often these expressions remain uninterpreted.

In recent years two interesting attempts have been made to develop this position further. W. T. Stace, in his book *Time and Eternity* (1952), considers the chief function of religious language to be the evocation of mystical experience, or faint echoes thereof. This seems at first to be a subjectivist account, with the deity omitted, but, as Stace correctly points out, it is an axiom in the mystical tradition that no difference can be found in mystical experience between subject and object, and on these grounds Stace refuses to make the distinction. Although Stace goes along with the mystical tradition in regarding mystical experience as ineffable, he departs from this official position to the extent of giving some indications of the aspects of this experience that different theological utterances evoke. "God is truth" evokes the sense of revelatoriness, "God is infinite" the sense of all-inclusiveness, "God is love" the blissful, rapturous character of the experience, and "God is one" the absolute unity of the experience and the sense of the dissolution of all distinctions.

Paul Tillich, although not squarely in the mystical tradition, is faced with similar problems in the interpretation of religious language. He holds that theological doctrines "symbolize" an ultimate reality, "being-itself," about which nothing can be said literally except that it is metaphysically ultimate. In attempting to clarify the function of religious language, Tillich develops the notion that it is an expression of "ultimate concern," a complex of devotion, commitment, and orientation, focused on something nonultimate—a human being, a nation, or a supernatural deity. Religious statements, which literally refer to such relatively concrete focuses of ultimate concern, express the sense of the sacredness such objects have as "manifestations" of being-itself. But just what it is for such an object to be taken as a "manifestation" or "symbol" of being-itself, Tillich never makes clear.

The basic weakness in these mythical and mystical interpretations is the failure to present any clear hypothesis concerning the function of religious language. Even Cassirer's ideas on "mythical thought" have never been developed to the point of clarifying what contemporary religious believers mean when they talk about God. The other positions are more intelligible, and they all base themselves on important aspects of the use of language in religion. But it seems that each, by inflating its chosen aspect to sole authority, has killed the goose that lays the golden eggs. There is no doubt that in talking about God, religious people express feelings of various sorts, present moral ideals, and articulate what is going on in ritual. But it is not at all clear that they would be using this kind of

language if they were not convinced of the truth of the statements they make. Why should I express a feeling of security by saying "God made the heavens and the earth" unless I believe, or at least have some tendency to believe, that as a matter of objective fact the physical universe owes its existence to the creative activity of a supernatural personal deity? Still more, why should I take on the complex of attitudes and activities that goes along with this assertion unless I believe it to be true?

The statement-making function is the cornerstone on which all the other functions depend. And if one is convinced that theological statements are either false or meaningless and still wants to hold to traditional religious formulations, one may *propose* a reinterpretation of theological utterances as expressions of feeling or symbolizations of natural facts. But a proposal for adopting a certain interpretation must be distinguished from a claim that the proposed interpretation correctly reflects the way doctrines are commonly understood.

It would seem that talk about God is much more complex than is recognized by any of the existing theories. The brief discussion given above of empirically testable implications illustrates this point. Theological sentences perform a great many closely interrelated linguistic functions. In saying "God, who created the world, watches over the affairs of men," the believer is committing himself to a certain general view of the ultimate basis of the world, giving voice to certain, perhaps very indefinitely specified, expectations as to how things will ultimately turn out, expressing a basic sense of security in life, committing himself to approach God in prayer and ritual in one way rather than another. And these functions are intimately dependent on each other. What is needed is a description of the relationships among these functions, one sufficiently complex to match the complexity of the subject matter.

See also Braithwaite, Richard Bevan; Cassirer, Ernst; Mysticism, History of; Philosophy of Religion, Problems of; Propositions, Judgments, Sentences, and Statements; Santayana, George; Stace, Walter Terence; Subject and Predicate; Tillich, Paul; Verifiability Principle; Wisdom, (Arthur) John Terence Dibben.

Bibliography

St. Thomas Aquinas's historically important discussion of the "analogical" character of theological terms is found in Question XIII of Part I of the *Summa Theologiae.* For further discussions in the Thomist tradition see Cajetan, *The Analogy of Names,* translated by E. A. Bushinski and H. J. Koren (Pittsburgh: Duquesne University, 1953), and E. L. Mascall, *Existence and Analogy* (London, 1949). Recent discussions by practitioners of analytical philosophy are to be found in Antony Flew and Alasdair MacIntyre, eds., *New Essays in Philosophical Theology* (New York: Macmillan, 1964), and in Basil Mitchell, ed., *Faith and Logic* (London: Allen and Unwin, 1957). See also Alasdair MacIntyre, "The Logical Status of Religious Belief," in *Metaphysical Beliefs* (London: SCM Press, 1957); John Wisdom, "Gods," in *Essays in Logic and Language,* edited by Antony Flew, first series (Oxford, 1951); C. B. Martin, *Religious Belief* (Ithaca, NY: Cornell University Press, 1959); and I. T. Ramsey, *Religious Language* (London: SCM Press, 1957). Various kinds of symbolic interpretations of religious statements are presented in George Santayana, *Reason in Religion* (New York: Scribners, 1905); R. B. Braithwaite, *An Empiricist's View of the Nature of Religious Belief* (Cambridge, U.K., 1955); W. M. Urban, *Language and Reality* (New York, 1939); Edwin Bevan, *Symbolism and Belief* (Boston: Beacon Press, 1957); and Philip Wheelwright, *The Burning Fountain* (Bloomington: Indiana University Press, 1954). Ernst Cassirer's views on mythical language are set forth most completely in Vol. II of *Philosophie der symbolischen Formen,* 3 vols. (Berlin: Cassirer, 1923, 1925, 1929), translated as *The Philosophy of Symbolic Forms* by Ralph Manheim (New Haven, CT: Yale University Press, 1952, 1955, 1957); and in more concentrated form in *Sprache und Mythos* (Leipzig: Teubner, 1925), translated as *Language and Myth* by S. K. Langer (New York: Harper, 1946). For other versions of the view that religious language constitutes an autonomous mode of discourse, see W. T. Stace's *Time and Eternity* (Princeton, NJ: Princeton University Press, 1952) and Paul Tillich's *Systematic Theology* (Chicago: University of Chicago Press, 1951–1963) and *Dynamics of Faith* (New York: Harper, 1957).

William P. Alston (1967)

RELIGIOUS LANGUAGE [ADDENDUM]

Two significant contributions to recent discussions of religious language are offered by Janet Soskice and William P. Alston. In *Metaphor and Religious Language* (1985) Soskice offers as a working definition "metaphor is that figure of speech whereby we speak about one thing in terms which are seen to be suggestive of another" (p.15). The minimum unit in which a metaphor is established is semantic. A satisfactory theory of metaphors "should regard metaphors neither as a simple substitution for literal speech nor as strictly emotive. Metaphors should be treated as fully cognitive and capable of saying that which may be said in no other way. It should explain how metaphor gives us "two ideas for one," yet do so without lapsing into a comparison theory" (p.44). Noncognitive accounts of metaphor are rejected because "we cannot conceive of emotive 'import' apart from a

cognitive content which elicits it" (p. 27). The "two ideas for one" feature involves a metaphor having a "unity of subject matter" that "draws upon two (or more) sets of associations, … characteristically, by involving the consideration of a model or models" (p. 49). A model is "an object or a state of affairs … viewed in terms of its resemblance, real or hypothetical, to some other object or state of affairs" (p. 100).

Models come in two types: paramorphic (the source and subject differ, as in billiard ball movement serving as a model for the properties of gases); or homeomorphic, where the subject is the source (e.g., a dummy used to teach lifesaving skills). Models, in both theology and science, are essential to theories because they carry their explanatory force. "The fertility of a theory lies in its ability to suggest possibilities of explanation which, while not inconsistent with, are more than simply the logical extensions of mathematical formulas … this suggestive capacity … constitutes the fruitfulness of a theory, and gives the theory the predictive nature which is its *raison d'etre*" (p. 114). We do not describe God but point to God through effects, and beyond them to him. We refer without defining. "This is the fine edge at which negative theology and positive theology meet, for the apophatic insight that we say nothing of God, but only point to Him … this separation of referring and defining is at the very heart of metaphorical speaking.…" (p. 140).

Nothing in Soskice's account of metaphor entails that language about God must be nonliteral. The claim *All language about God is metaphorical* is not metaphorical. The idea that no metaphor can be translated into or replaced by literal terms is false. Consider Soskice's example of an expression of hope that a soldier will be pardoned eliciting, "That's blowing on cold coals." "There's no chance of that" is a literal translation. "God is a rock" seems replaceable by "God is utterly reliable." If it is not, this is a matter of the associations of "rock" in biblical and theistic literature being multiple. It does not follow that any of the things that "rock" suggests are nonliteral. It just suggests that there are a variety of possibilities, more perhaps than we can list, each of which may be perfectly expressible without remainder in literal fashion.

A basic assumption is that no literal description can be true of God. As is typical, we are referred to certain ideas: We cannot comprehend (know all there is to know) about God; descriptions of God based on religious experience are defeasible; certainty about claims concerning God is unattainable; and it is always possible that we will have to modify our concept of God. But there are an infinite number of truths concerning a golden retriever, seeing the golden is defeasible, certainty about it is unavailable, and we may have to revise our concept thereof. But it is not beyond literal description. Further, God can be misdescribed (e.g., "God is a cantaloupe"), which even the most deluded of empiricist positivists presumably will recognize as false. But then what, in principle, precludes God from being correctly described?

William P. Alston's major essays concerning religious language are collected in *Divine Nature and Human Language* (1989). In "Irreducible Metaphors in Theology" he says that "in the typical metaphorical statement the speaker is 'building on' the relevant meaning of the predicate term in two ways … he is presenting the thing to which the term literally applies as a model of the subject [and] … he has in mind one or more resemblances between model and subject and he abstracts from these resemblances what he means to be attributing to the subject" (1989, p. 23). The resemblance may be either general or specific. Everything resembles everything else in some way. Any metaphor based on this fact corresponds to a literal way of expressing the similarity. Regarding metaphors intended to express truths, he writes: "Though irreducible metaphors seem to promise a way of combining the denial of predication in theology with the preservation of significant theological truth claims, this fair promise dissipates on scrutiny like mist before the morning sun. Either the panmetaphoricist abandons the aspiration to significant truth claims or he revokes the ban on literal predication" (p. 37).

"Can We Speak Literally of God?" considers predicates that apply to personal agents ("P-predicates") in their application to God. These include mental and action predicates. These have been understood on a private paradigm model (one knows what "depression" means by being depressed) and functionally ("being depressed" refers to a state that functions efficaciously in a causal system to yield a distinctive range of behavior). The idea of basic actions that involve no bodily movements, and of nonbasic actions that involve only mental actions that bring about effects, are both intelligible and applicable to incorporeal beings. "Literal" does not mean "empirical."

"Functionalism and Theological Language" and "Divine and Human Action" consider functional accounts of mental concepts to argue that these concepts can apply to God. We can "form the conception of a being (a 'system') in which some factors depend on their relations to others for being what they are, even though there are no temporally successive processes for formation of any subjection to laws. More specifically, we are to think

of God as realizing a complex structure of attitudes, knowledge, tendencies, executive intentions, and volitions in the 'eternal now' …" (p. 99). The stability of this system is to be understood, not by way of there being laws that hold regarding it, but by way of essential properties of the system. But this gives us only a description of God as a system of items that bear various dependence or causal relationships, not of a personal agent. Insofar as the relevant concepts are strictly functionalist, they do not entail even consciousness. When we turn to religious discourse about God, the functionalist account is not nearly enough: "For the religious life, we need to go beyond that in ways that launch us into the still not sufficiently charted seas of the figurative and the symbolic" (p. 103).

"Referring to God" distinguishes between direct reference and reference by description. Reference by description offers a description that is true only of the referent; direct reference names an object of one's experience. Direct reference to God can occur only if someone experiences God (Alston takes it that some people do). Others who do not themselves experience God can then refer to the being that others have referred to; reference thus spreads throughout a religious community. Direct reference is more basic than descriptive, because if one refers to a being both descriptively and directly, and one learns that the description is false of the being directly referred to, it is the latter that determines what was the actual object of reference. Nonetheless, Alston admits that "reference could always take place via a description (p. 107). A consequence is that it is possible that someone who thinks of God as an omnipotent, omniscient spirit, and one who thinks of God as an impersonal force, may refer to the same being. Alston says that it may be that both are "worshiping the one true God" (p. 116). If so, worship does not require much by way of actual knowledge of God.

There was never any reason to think that a causal theory of reference wedded to a functionalist account of P-predicates would yield significantly positive results regarding description of God. It seems fair to say that in spite of the sophisticated and helpful discussions provided by Soskice and Alston, accounts of religious language that are philosophically articulate and allow for seriously realistic accounts in theology remain more matters on the agenda than they are accomplishments of current work in the field.

See also Alston, William P.; Metaphor; Subject and Predicate.

Bibliography

Alston, William P. *Divine Nature and Human Language: Essays in Philosophical Theology*. Ithaca, NY: Cornell University Press, 1989.

Alston, William P. *Illocutionary Acts and Sentence Meaning*. Ithaca, NY: Cornell University Press, 2000.

Soskice, Janet M. *Metaphor and Religious Language*. Oxford: Oxford University Press, 1985.

Keith E. Yandell (2005)

RELIGIOUS PLURALISM

The fact that there is a plurality of religions is significant in different ways from different points of view. From a skeptical point of view their different and often incompatible beliefs confirm the understanding of religion as delusion. Thus, Bertrand Russell wrote that "It is evident as a matter of logic that, since [the great religions of the world] disagree, not more than one of them can be true" (1957, xi). From the point of view of an exclusive and unqualified commitment to any one religion the fact of religious plurality is readily coped with by holding that all religions other than one's own are false, or false insofar as their belief systems differ from one's own. But from a point of view that sees religion as a worldwide phenomenon that is not to be dismissed in toto as delusion but as the human response to a divine/transcendent/ultimate reality, the fact of plurality poses a major philosophical problem. On the one hand, the "great world religions" seem—to many impartial observers, at any rate—to affect human life for both good and ill to more or less the same extent. But on the other hand their respective belief systems, although having important similarities, also include starkly incompatible elements. According to some the Real (a term at home in the Judeo-Christian tradition and corresponding to the Sanskrit *sat* and the Arabic *al-Haqq*) is personal but according to others not personal. And within each group of religions there are wide differences. Is the ultimate Person the Christian Trinity or the Qurʾanic Allah, or the Adonai of Judaism, or Vishnu, or Shiva? Is the nonpersonal Ultimate the Brahman of advaitic Hinduism, or the Dao, or the Dharmakaya or Void or Nirvāṇa of the Buddhist traditions? And how could the Real be all of these at once? The logic of religious difference here is in fact very complex, as is shown by William Christian's analysis (1987).

The problem is particularly acute for a major form of religious apologetic that became prominent in the 1980s and 1990s. This holds that the basic empiricist principle that it is rational, in the absence of specific overriding

considerations, to base beliefs on experience should be applied impartially to all forms of putatively cognitive experience, including religious experience—unless, again, there are specific overriding considerations to the contrary. This has been argued directly by William Alston (1991) and others and indirectly by Alvin Plantinga (in Plantinga and Wolterstorff 1983), whose defense of the rationality of holding "properly basic" religious beliefs presupposes religious experience as their ground.

Most of the philosophers who employ this kind of apologetic have applied it only to specifically Christian beliefs. But it is evident that precisely the same argument is available for the belief systems of other religions. If Christian religious experience renders it epistemically justifiable (subject to the possibility of specific reasons to the contrary) to hold Christian beliefs, then Buddhist religious experience renders it epistemically justifiable, with the same qualification, to hold Buddhist beliefs, Muslim religious experience to hold Muslim beliefs, and so on. Thus, anyone who maintains that the Christian belief system is true, but that the belief systems of Buddhists, Muslims, and so on are false insofar as they differ from it, has implicitly reversed the original apologetic and is presenting Christian religious experience as the sole exception to the general rule that religious experience gives rise to false beliefs!

Alston, recognizing the challenge posed by the fact of religious diversity to the experiential apologetic, has responded by saying that in this situation it is proper for the Christian to continue within her own belief system, despite the existence of other equally well-justified alternatives, while, however, she seeks "a way to show in a noncircular way which of the contenders is correct" (1991, p. 278).

An alternative use of the experiential apologetic rejects the assumption that only one of the different religious belief systems can be true. This approach (Hick 1989) distinguishes between, on the one hand, the ultimate religious reality, the Real, beyond the scope of our (other than purely formal) human conceptualities, and, on the other hand, the range of ways in which that reality is humanly conceived, and therefore humanly experienced, and therefore humanly responded to within the different religiocultural ways of being human. The epistemology operating here is one that, in the Kantian tradition, recognizes an important contribution by the perceiver to the form a reality is perceived to have. As Thomas Aquinas wrote, "Things known are in the knower according to the mode of the knower" (*Summa Theologiae,* II/II, 1, 2). And in religious knowing the mode of the knower differs from religion to religion. From this point of view the fact of religious diversity does not constitute a challenge to the experiential apologetic but rather a series of examples of its valid application.

Other philosophical responses to the fact of religious plurality, not specifically related to the experiential apologetic, include the "perennial philosophy" (e.g. Schuon 1975, Smith 1976), which distinguishes between the essence (or esoteric core) of religion and its accidental (or exoteric) historical forms. In their esoteric essence all the great traditions converge in a transcendental unity, the Absolute Unity that is called God. Experientially, this sees the mystics of the different religions as participating in an identical experience, although they articulate it in the different ways provided by their traditions. This view is opposed by those (e.g., Katz 1978) who hold that all experience is concept laden and that mystical experience accordingly takes different forms within the different traditions.

There is also the view of John Cobb (in Kellenberger 1993) that the religions are directed toward different ultimates, particularly the personal reality worshiped in the theistic religions and the nonpersonal process of the universe experienced in Buddhism. Yet other constructive suggestions include those of Joseph Runzo (1986), James Kellenberger (1989), and the authors included in the symposium *Inter-Religious Models and Criteria* (Kellenberger 1993).

See also Philosophy of Religion; Philosophy of Religion, History of; Philosophy of Religion, Problems of; Religious Experience; Russell, Bertrand Arthur William; Thomas Aquinas, St.

Bibliography

Alston, W. P. *Perceiving God.* Ithaca, NY: Cornell University Press, 1991.

Basinger, David. *Religious Diversity: A Philosophical Assessment.* Burlington: Ashgate, 2002.

Christian, W. A. *Doctrines of Religious Communities.* New Haven, CT: Yale University Press, 1987.

Godlove, T. F. *Religion, Interpretation, and Diversity of Belief.* Cambridge: Cambridge University Press, 1989.

Griffiths, Paul. *Problems of Religious Diversity.* Oxford: Blackwell, 2001.

Hick, J. *An Interpretation of Religion.* New Haven, CT: Yale University Press, 1989.

Katz, S. T., ed. *Mysticism and Philosophical Analysis.* London: Sheldon Press, 1978.

Kellenberger, J. *God: Relationships with and without God.* New York: St. Martin's Press, 1989.

Kellenberger, J., ed. *Inter-Religious Models and Criteria.* New York: St. Martin's Press, 1993.

Krieger, D. J. *The New Universalism.* Maryknoll, NY: Orbis, 1991.
McKim, Robert. *Religious Ambiguity and Religious Diversity.* Oxford: Oxford University Press, 2001.
Plantinga, A., and N. Wolterstorff, eds. *Faith and Rationality.* Notre Dame, IN: University of Notre Dame Press, 1983.
Quinn, Philip, and Kevin Meeker, eds. *The Philosophical Challenge of Religious Diversity.* New York: Oxford University Press, 2000.
Runzo, J. *Reason, Relativism, and God.* New York: St. Martin's Press, 1986.
Russell, B. *Why I Am Not a Christian.* London: Allen and Unwin, 1957.
Schuon, F. *The Transcendent Unity of Religions,* rev. ed. New York: Harper and Row, 1975.
Smith, H. *Forgotten Truth.* New York: Harper and Row, 1976.
Vroom, H. M. *Religions and the Truth.* Amsterdam: Rodopi, 1988.

John Hick (1996)
Bibliography updated by Christian B. Miller (2005)

RENAISSANCE

"Renaissance" is the term customarily employed to designate a cultural movement that began in Italy in the middle of the fourteenth century and spread throughout the rest of Europe. Although the term is well established in the writings of historians, its usefulness has been challenged. Indeed, there has grown up around the concept of the Renaissance an extensive controversy that sometimes threatens completely to divert the attention of scholars from the historical facts. In part, this controversy is simply an acute form of the general problem of periodization in history. The concept of the Renaissance, however, arouses particularly strong opposition because it involves a disparagement of the preceding period, the Middle Ages (*medium aevum*), from which culture presumably had to be awakened.

The idea of a rebirth of literature or of the arts originated in the period itself. Petrarch in the fourteenth century hoped to see an awakening of culture, and many later writers expressed their conviction that they were actually witnessing such an awakening in their own time. Latin was generally the language used by cultivated men to discuss such matters, but no single Latin term or phrase became the standard name for the whole cultural epoch. One of the earliest historians of philosophy in the modern sense, Johann Jakob Brucker, in 1743 referred to the Renaissance only as the "restoration of letters" (*restauratio literanum*), and wrote of the "recovery of philosophy" (*restitutio philosophiae*): Even in an earlier German work he used such Latin phrases. Scholars who wrote in Latin never used *rinascentia* as the name for the cultural epoch as a whole. It was the French word *renaissance* that finally acquired this status and was then adopted or adapted into other languages. During the seventeenth century, and fitfully before, French scholars used the phrase *renaissance des lettres* for the humanists' *restitutio bonarum literarum*, taking over in the process the humanist periodization of history. Other writers translated the Latin phrase or phrases into their own vernacular: Edward Gibbon (1787) spoke of the "restoration of the Greek letters in Italy," while Heinrich Ritter, in his history of philosophy (1850), remarked that the *Wiederherstellung der Wissenschaften* derived its name from philology.

Various French authors used the term *renaissance* in titles of their works before Jules Michelet devoted one of his volumes on sixteenth-century France to *la Renaissance* (1855). However, Michelet gave only the sketchiest characterization of the period, and hardly deserves to be credited (if indeed any one person can be) with having "invented" the concept of the Renaissance. Michelet did coin one memorable phrase: He remarked that two things especially distinguished the Renaissance from previous periods—"the discovery of the world, the discovery of man." This phrase was also used by the Swiss cultural historian Jakob Burckhardt for the title of a chapter in his famous work, *The Culture of the Renaissance in Italy* (1860). At his hands, the concept of the Renaissance received what was to become its classic formulation; all subsequent discussion of the concept invariably focuses upon Burckhardt's description of the essential features of life during the Renaissance. Burckhardt, taking the term in its narrow sense of a literary revival of antiquity, conceded that there had been earlier "renaissances" in Europe; but he insisted that a renaissance in this sense would never have conquered the Western world had it not been united with the "already-existing spirit of the Italian people" (*italienischen Volksgeist*). Not until the time of Petrarch, so Burckhardt held, did the European spirit awake from the slumber of the Middle Ages, when the world and man lay "undiscovered."

The relation of the Renaissance to the era that preceded it has been much studied because defenders of medieval culture quickly came to the rescue of their period, stressing its continuity with, or even its superiority to, the Renaissance. However, little has been done to clarify the relation of the Renaissance to the Enlightenment. This is rather surprising, for there was an issue that ran straight through the thought of both these eras: "Can we modern men hope to equal or even excel the achievements of antiquity?" This issue is known to literary histo-

rians as the "quarrel of the ancients and moderns." We think of Bernard Le Bovier de Fontenelle in the seventeenth century as the main champion of the moderns, who had science and truth on their side, as against those writers, with their inflexible rules, who favored the ancients. However, much the same attitude as Fontenelle's is found in the *De Disciplinis* of the Renaissance humanist Juan Luis Vives, who wrote in the early sixteenth century. The Renaissance itself had championed the moderns even before modern science had arisen to prove their case. Renaissance confidence in men's powers was based on art and literature rather than on science, but it was strong nevertheless. Men could respect classical excellence and yet strive to outdo the ancients in every field, including vernacular literature.

CHRONOLOGICAL LIMITS

Various events have been taken as marking the beginning of the Renaissance: the crowning of Petrarch as poet laureate of Rome in 1341; the short-lived triumph of Cola di Rienzi in setting up a republican Rome in 1347, an attempt to revive Rome's former greatness; the arrival in Italy of Greek *émigrés* (which actually antedated by a few years the much publicized fall of Constantinople in 1453); the opening up of new trade routes to the East. Each choice represents the selection of a particular field as central in the history of the period: art, architecture, religion, politics, economics, trade, or learning. In certain fields it is hard to maintain any sharp break between conditions in, let us say, 1300 and those in 1350. However, few students of the history of art or of literature are prepared to deny completely the start of new trends in the fourteenth century (at least in Italy). In literature, Petrarch's enthusiasm for Greek antiquity must surely be accepted as inaugurating, in the eyes of men in the fourteenth century, a fresh start. In painting, there is little hesitation about ascribing a similar place to Petrarch's contemporary, Giotto; this ascription dates from the earliest attempt at a history of art, that of Giorgio Vasari (1550). No such figures can plausibly be singled out to mark new beginnings in economic or political history.

Difficulties also surround the choice of an event to mark the end of the Renaissance: the sacking of Rome in 1527, the hardening of the Counter-Reformation via the Council of Trent in 1545, the burning of Giordano Bruno in 1600, or Galileo Galilei's setting of experimental physics on its true path around 1600—any of these might be selected. Once again, however, a periodization that is useful in one field may prove useless in another field. Generally speaking, the era from 1350 to 1600 will include most of the developments commonly dealt with under the heading "Renaissance."

GEOGRAPHICAL LIMITS

The shifting locale of the Renaissance presents problems similar to those of its chronological limits. Burckhardt's description focused exclusively on Italy; he implied that the Renaissance, after it had been taken over by the Italian *Volksgeist*, moved on to the rest of Europe. The movement to France is usually said to have resulted from the French invasion of Italy in 1515, which gave the French nobility their first glimpse of the glories of the Italian Renaissance. No comparable event can be singled out for the bringing of the Italian Renaissance to England, unless it be the return from Italy to their native land of the classical scholars William Grocyn, Thomas Linacre, and John Colet in the last decade of the fifteenth century, or perhaps Desiderius Erasmus's arrival there about the same time. Clearly England did enjoy a renaissance, but it is not easy to fix its dates: English literary historians prefer to discuss the Elizabethan age or the age of the Tudors, thus sidestepping the question of the relation of the English Renaissance to that of the Continent. Still less clear is the coming of the Renaissance to the German lands: German historians treat the sixteenth century as the "time of the Reformation," and tend to discuss the Renaissance chiefly in terms of its impact upon individual reformers.

The Renaissance is sometimes called the "age of adventure." It is not at all clear, however, that the spirit behind men's daring and adventurous actions was entirely new: The two chief incentives toward voyages of discovery, for instance, were commercial acquisitiveness and religious zeal—attitudes by no means foreign to medieval men. It was the shutting off of Venetian trade routes through the Mediterranean by the Turks that forced Europeans to search for new routes to the East, not a new desire for scientific knowledge of geography. The Spanish *conquistadores* may have thirsted for glory, but such a thirst was characteristic of medieval knights as well as of Renaissance humanists. The motives of the Franciscan missionaries were clearly religious and medieval in spirit. Moreover, in the field of domestic trade, the resurgence of economic activity in the fifteenth century that formed the basis for the cultural developments of the Renaissance was less a matter of suddenly effective acquisitiveness than of normal recovery from the slump brought about by the Black Death in 1348.

THE NEW LEARNING

Historians may without hesitation ascribe a rebirth of classical knowledge to the Renaissance period. The discovery of old manuscripts and the invention of printing combined to make the heritage of ancient Greece and Rome available to a far wider audience. The humanists of the fourteenth and fifteenth centuries discovered and preserved many ancient texts that had been neglected for centuries. Of these perhaps the most significant from a philosophical point of view was Lucretius's *De Rerum Natura*, but many other newly discovered texts helped to enrich men's general familiarity with antiquity and to present in full view the setting in which Greek and Roman philosophy originated.

The collecting of manuscripts could be indulged in only by noblemen or well-to-do scholars, but the invention of printing made possible a broader social base for intellectual interests. With the production of vast numbers of newly discovered texts, self-education became a real possibility, as did institutional education on a broad scale. Peter Ramus in France and Philipp Melanchthon in Germany urged the educating of the people, chiefly with the idea of promoting intelligent Christian piety.

SCIENCE

Developments in technology and science indirectly provided material for philosophical reflection. The increased use of firearms and cannon in war, for example, made necessary the mathematical study of ballistics; and the scientific work of Benedetti and Galileo drew upon the practical experience of foundries and arsenals. However, Renaissance philosophy of science still took its cue largely from Aristotle: Francis Bacon, dissatisfied with Aristotelian logic and methodology of science, found a replacement not in the actual practices of mechanics and craftsmen but in the rhetorical method derived from Aristotle and applied to the questioning of Nature.

The most spectacular and far-reaching scientific development during the Renaissance was the heliocentric theory advanced by Nicolas Copernicus, who found hints about Pythagorean cosmology in ancient works. The Copernican theory was surely the most significant revolution ever to take place in science. Far less conspicuous, but still important, were the developments in pure and applied mathematics. Modern notation (such as the use of the "equals" sign) began to be adopted, bringing with it the possibility of greater attention to logical form.

SOCIAL VALUES

There have been many attempts, beginning with Michelet and Burckhardt, to capture *the* mind or spirit of Renaissance man. All such attempts seem doomed to failure, for they are bound to oversimplify complex social facts. We may, however, single out four sets of social ideals that were characteristic of various groups during the Renaissance.

The ideals of the feudal nobility, medieval in origin, persisted through the Renaissance among the ruling class, although they underwent considerable refinement. The rude military virtues of camp and field gave way to the graces of the court, which were set forth most admirably in Baldassare Castiglione's book *The Courtier* (1528), one of the most influential treatises on manners ever written. In Castiglione's ideal courtier we may recognize the ancestor of our "gentleman." Works of this sort are presumably also the source of the "universal man," a concept closely associated in modern minds with the Renaissance. In the heroic life idealized by the feudal tradition, love of glory and concern for one's reputation were strong social motives. The humanists' thirst for glory, which Burckhardt emphasized, merely continues this concern but applies it to the achievements of a nonwarrior class, the "knights of the pen." The urban middle class chose, as usual, to emulate the style of life of their superiors: the modern gospel of work as a raison d'être, shaping the whole of life, hardly existed during the Renaissance. Few social theorists extolled the virtues of commercial activity until Martin Luther stressed the sanctity of all callings, provided they benefited one's fellow men.

Religion provided the second set of ideals, which centered upon moral salvation and involved a willingness to relinquish the world and all its goods. This mood, exacerbated in some individuals by the terror of imminent death or of eternal damnation, continued unabated throughout the Renaissance; and the entire Reformation movement has been called the "last great wave of medieval mysticism." Although such a religious concern is usually associated by modern secular critics with contempt for this world and with pessimism, it is equally compatible with a cheerful resignation in the face of unavoidable misfortunes and gratefulness for such morally harmless pleasures as life affords. A genuine tension often resulted from the opposing pulls of these religious values and of secular attitudes and this-worldliness: Aristotelian philosophers as well as humanists felt this tension during the Renaissance.

A third set of ideals, that of the ancient sage (Platonic or Stoic), was consciously adopted by Renaissance

humanists as an adjunct to Christian exhortation, for many of them felt that Christians could learn much from pagan expounders of virtue. Rarely, if ever, did a humanist attempt to replace the Christian ideal altogether: Burckhardt undoubtedly overstressed the "paganism" of the humanists.

Finally, there was the ideal of a return to nature, a flight from the complexities of sophisticated urban life to pastoral pleasures. This theme has ancient antecedents in the poetry of Theocritus and Vergil, but it emerges into new prominence with Petrarch, who also stressed the benefits of solitude. Passive delight in the beauties of nature can hardly ever be totally lacking in human beings, of course, but during the Renaissance we find an interest in such activities as gardening, the collecting of strange plants and animals, and strolling through woods and fields. Petrarch's famous excursion to the summit of Mont Ventoux turned into an occasion for Christian self-reproach, to be sure, but his letters also abound in references to his gardening and to lone promenades in the countryside near Vaucluse.

HUMANISM

A major role in the culture of the Renaissance was played by the humanists. All sorts of people call themselves "humanists" today, but in the early days of the Renaissance the name had a clear occupational meaning. During the fourteenth century, the traditional subjects of grammar, rhetoric, and poetry had begun to be called, after a phrase of Cicero, the *studio humanitatis*. The term *umanista* was coined (on the analogy of *artista*, also a product of university slang) to designate a teacher of these subjects in Italian universities. Such studies were by no means new in the fourteenth century; in fact, the humanists were the heirs of a less ambitious but old and respectable medieval profession, that of the *dictator* or teacher of the art of letter-writing (*ars dictaminis*). The Renaissance teachers of "humanities" placed a greater emphasis on ancient models than had the *dictatores*, but their teaching had much the same; objective. Their students often became official letter-writers or speechmakers for popes and princes. Coluccio Salutati and Leonardo Bruni, two of the most influential humanists of the fifteenth century, were chancellors of Florence. The study of Greek philosophy owes much to these two men.

Renaissance humanists did not propound a distinct philosophy but took over from Cicero and Aulus Gellius the ancient ideal of a civilized and urbane way of life that could be formed through acquaintance with Greek literature. With such a program in mind, the humanists began to concern themselves with moral and political philosophy, and this brought them into conflict with the philosophers who taught ethics or politics in the universities. The humanists regarded the Aristotelian Schoolmen as derelict in the performance of their duties, since their teaching (so the humanists claimed) made no differences in the lives of students. The scholastic teachers, in return, regarded the humanists as dilettantes and upstarts, meddling in subjects beyond their depth. The feud of humanists with philosophers began with Petrarch's invective against the secular Aristotelians, the so-called Averroists of his day, and continued through the seventeenth century.

We still tend to see Renaissance Aristotelianism (and medieval Scholasticism as well) through the eyes of these Renaissance humanists. Their bias has crept into most histories of philosophy, largely because the first writers of histories of philosophy shared some of the humanist attitudes. One such early historian was Brucker, whose *Critical History of Philosophy* (1742–1744) has already been mentioned. Brucker presented the Renaissance as a time when human thought emerged slowly into the light (a standard metaphor) from the tiresome labyrinths of medieval Scholasticism. He divided his treatment into various sections, dealing with schools of Greek philosophy that were "restored" during the Renaissance. In spite of his scorn for "more recent Aristotelian-scholastic philosophers," Brucker had great respect for the philosophers who followed the "genuine philosophy of Aristotle": Pietro Pomponazzi, Simon Porta, Jacopo Zabarella, and others. Few modern historians of philosophy pay much attention to these writers. They do, however, characteristically devote lengthy sections to Paracelsus, Jakob Boehme, Robert Fludd, and other "theosophers." According to Brucker, these theosophers "condemn all use of reason in understanding the nature of things," and hence do not belong to the history of philosophy; he includes them only because they have commented incidentally on philosophical matters.

Whatever his own philosophical competence may have been, Brucker had one clear advantage over most later historians: He had actually read the Renaissance writers he discussed. Much of Renaissance philosophy still awaits reevaluation based upon such actual reading of texts.

The general framework of Brucker's treatment of Renaissance philosophy remains a useful way of dealing with most of the thought of the period. The various sects of Greek philosophy were indeed "reborn" during the Renaissance; few of them escaped some sort of revival.

There was even what might be called a genuine rebirth of Aristotle, if we mean by this what Brucker probably meant: an Aristotelianism based directly upon the Greek texts rather than upon Latin or Arabic commentators.

ARISTOTELIANISM. It cannot be too strongly emphasized that the main stream of philosophical inquiry during the Renaissance continued to be Aristotelian. The terms employed in philosophical discussion, the problems posed, and the characteristic solutions remain, in basic outline, Aristotelian. Almost all Renaissance philosophers show the influence of their Aristotelian school training, even when they are trying most strenuously to break the shackles of that tradition. The technical terms of philosophy (such as *propositio, entitas, realis, materia, forma, essentia* and many others) originated or became naturalized in the Aristotelian school-tradition, and persisted even in the writings of the most daring innovators, such as Bruno. The Aristotelian tradition, for reasons already in part suggested, remains the least known and most maligned of all Renaissance schools. Elements of the critical spirit of later medieval philosophy (Scotist and Ockhamist) formed part of the school philosophy of the Spaniard Francisco Suárez and of the Scotsman John Major.

PLATONISM. Platonism took on new life during the Renaissance, after having been known for centuries chiefly through Aristotle's attacks on it. There was more acquaintance with Plato during the medieval period than is generally recognized, but it is still true that Marsilio Ficino's translations into Latin (first published in 1484) gave the main impetus to the spread of Plato's doctrines. Later editions of Plato often contained Ficino's translations of Proclus and Porphyry, together with his own commentaries, which were strongly colored by his Neoplatonism. Hence, the Platonism that emerged during the Renaissance cannot be distinguished easily from Neoplatonism, for it tends to be otherworldly and religious in tone. The cultural influence of Florentine Platonism emanated from the famous academy founded by Ficino in direct imitation of Plato's school. The society that grouped itself around Ficino aimed at moral improvement and resembled in character certain lay religious societies common in Italy at that time. The whole movement of natural religion was set in motion by Florentine Platonism, as was the renewed study of Pauline theology by such men as John Colet.

Florentine Platonism is well known, by name at least, to most students of the Renaissance. Much less well known is a tradition of reconciling Plato with Aristotle, which also found expression during the period. Byzantine scholars had brought with them to Italy an old battle over the superiority of Plato or Aristotle. During the late Renaissance this battle resolved itself into a truce, with many books written to show that Plato and Aristotle agreed on fundamentals and differed only on words or nonessentials.

STOICISM. Only a few late Renaissance thinkers, such as Justus Lipsius and Guillaume du Vair, committed themselves explicitly to Stoicism, but the influence of Stoic philosophy may be seen at work directly and indirectly (largely via Cicero, Seneca, and the Greek commentators on Aristotle) even during the early Renaissance. Pomponazzi's rigorous moral doctrine, for example, is strongly tinged with Stoic attitudes.

EPICUREANISM. Rejected with horror by medieval thinkers, who saw him through the eyes of the Church Fathers, Epicurus began to be more sympathetically known as a result of humanist activity in the fifteenth century. Previous to this time, anyone who believed that the soul perished with the body was called an Epicurean, whether he held to any other Epicurean tenet or not. Now it was no longer possible to apply this label so casually. Lucretius's great poem won immediate favor because of its sturdy poetic qualities, but, until Pierre Gassendi in the seventeenth century, no one adopted the system of Epicurus in its entirety. Nevertheless, Epicurean influence prior to Gassendi's time did foster a climate less hostile to the concepts of pleasure and utility.

SKEPTICISM. The direct influence of philosophical skepticism in a technical sense began with the first publication of Sextus Empiricus in 1562, from which time skepticism exercised an important influence upon European thought and literature. The religious factionalism or warfare of the sixteenth century had brought about a widespread distrust of dogmatism and fanaticism on the part of such sophisticated minds as Erasmus and Michel Eyquem de Montaigne, whose writings may have contributed to the growth of that spirit of toleration usually associated with the Enlightenment.

THE OCCULT TRADITION. The Renaissance was immensely receptive (perhaps more so than the Middle Ages) to occult and secret lore of all kinds, especially if it claimed to come from the most ancient times and to incorporate the wisdom of the Egyptians, Chaldeans, and Hebrews. When the fashion for reviving ancient thought was at its height, the spurious treatises of "thrice-great

Hermes," the so-called Hermetic writings, enjoyed great prestige and blended easily with various other secret teachings, such as that of the Jewish Kabbalah.

Toward the end of the Renaissance, the vogue for reviving past philosophies began to subside: Instead, there began to appear "new" philosophies and "new" systems of thought proudly announced as such, for instance, the *Nova de Universis Philosophia* offered by Francesco Patrizzi or the *Great Instauration* (explicitly opposed to a "restoration") of Francis Bacon. However, most of these efforts at original creation clearly bear the stamp of some ancient sect or sects of philosophy. Even Nicholas of Cusa, the most original systematic mind of the Renaissance, could be called (and indeed once called himself) a Pythagorean. Philosophers hardly ever make a complete break with the past, even when they most loudly claim to be doing so. The great merit of the Renaissance was that thinkers learned what they could from the school of Athens and brought what they learned to bear with fresh vigor upon the problems of human life.

CARDANO

No individual completely typifies his age, yet it may be useful to focus for a moment on the way in which the various philosophical traditions converged in a single person. As a case history of this sort, we may take the thought of Girolamo Cardano (1501–1576), an Italian medical man and mathematician. Cardano lived in the late, mature stage of the Renaissance, when the dialogues of Plato and the works of Aristotle were known in their entirety, as were Galen and Hippocrates. The Greek commentators on Aristotle were just being recovered and translated. These works were well known at the universities where Cardano studied: Pavia, a stronghold of humanist learning, and Padua, a center of science and medicine. At Padua the biological and logical aspects of Aristotle's thought were stressed in connection with medical training. Cardano studied under Joannes Montesdoch, a Spaniard, whom he mentions in his writings. There were quite a few such Iberian philosophers studying and teaching in Italy at this time. Aristotelian philosophy was clearly a common European heritage and knew no national boundaries.

A considerable number of Renaissance philosophers were, like Cardano, medical men, and of these quite a few dabbled in mathematics (Galen had urged them to study mathematics for the sake of the training it gave them in sound demonstration). Cardano was, of course, far more successful than most in mathematics: No matter what the true story of his relations with Niccolò Tartaglia may be, there can be no questioning of Cardano's grasp of algebra, as shown by his solution of cubic equations. Cardano wrote works on medicine, astrology, and mathematics, but his philosophical reputation must rest primarily on two works in natural philosophy: *De Subtilitate Libri XXI* (On subtlety; 1550) and its sequel, *De Rerum Varietate* (On the variety of things; 1557). *De Subtilitate* attempted a total reconstruction of natural philosophy.

Since other philosophers of the period were inspired to embark on similar projects, it is clear that there was widespread dissatisfaction with Aristotle's philosophy of nature even before the attacks of Galileo or René Descartes. Aristotle's physical system was to be threatened dramatically by Copernican heliocentrism, which upset the conceptual scheme on which Aristotle's analysis of motion was based. This threat was not explicitly posed, however, until the next century, with Galileo's *Two Chief World Systems*. A Renaissance philosopher such as Cardano did not specifically base his criticisms of Aristotle on the findings of Copernicus or Vesalius: Instead, he reproached Aristotle in a general way for having built up "certain general propositions that experiment teaches to be false." Cardano presumably intended to remedy this defect, although it must be confessed that his empiricism is not worked out in philosophical detail. This observation would apply with equal force to most Renaissance nature philosophers, few of whom gave more than perfunctory attention to epistemology.

In developing his own system, Cardano started out by taking as his central category something called "subtlety," which he described as "a certain reason by which sensibilia are with difficulty comprehended by the sense, and intelligibilia by the intellect." Cardano soon abandons this unpromising concept in favor of a revised Aristotelian terminology in which matter, form, soul, principle, and element play roles somewhat analogous to those they play in Aristotle's philosophy. For example, Cardano retains the notion of elements but reduces their number from the traditional Aristotelian four to three by eliminating fire, which he classifies as an "accident." Matter and motion—those central concepts of mechanism—are regarded by Cardano as principles, but they must share this status with form, place, and soul. The last addition puts Cardano into the class of hylozoists, those who believe that all matter is somehow animated, a rather characteristic Renaissance doctrine borrowed largely from Neoplatonism.

Cardano's writings must have appealed to his Renaissance readers: They are lively, detailed, and full of medical and factual information and misinformation. His style

contrasts sharply with the dry, logically structured argument of the medievals, which can still be found early in the century in the work of a man such as John Major. Cardano obviously delighted in mathematics and in machinery, in this respect, at least, anticipating Galileo in the generation that followed. The amount of superstitious nonsense incorporated in Cardano's work, however, is still distressingly high, and one can easily understand the impatience of later figures such as Gassendi, Thomas Hobbes, and Galileo with their Renaissance predecessors. Cardano wrote a painfully candid autobiography, which appeared in Paris with an evaluation by the French writer Gabriel Naudé (1643). Naudé's judgment on Cardano's character is quite severe. This illustrates a general trend in scholarship: The information current today about many Renaissance thinkers, especially the Italians, comes to us by way of generally hostile French writers of the seventeenth century (Pierre Bayle is exceptional in his lack of polemical intent). If we approach Cardano with the distaste of a Naudé, for example, we too might be inclined to dismiss his work *On Consolation* (1542) as a piece of moralizing cant, when in fact a more humane scholar might consider it a noble document in the light of Cardano's wretched life. Or again, Cardano's passion for gambling could be presented as a despicable and mercenary motive for his interest in games of chance.

But a less censorious approach, such as that of Oystein Ore in his *Cardano, the Gambling Scholar* (Princeton, NJ, 1953), will give Cardano the credit he deserves for anticipating the modern conception of probability as the proportion of favorable outcomes to total possible outcomes. Finally, the mere fact that there was enough interest in Cardano's thought still lingering in seventeenth-century France to justify the publication of his entire work (*Opera Omnia*, 10 vols., Lyons, 1663), shows that Naudé's attitude was by no means universal. This comment could also be made of many other Renaissance philosophers who continued to be read in the seventeenth century, even if not all students of that century were as receptive to Renaissance thought as was Gottfried Wilhelm Leibniz.

See also Florentine Academy; Hermeticism; Humanism.

Bibliography

Everyone interested in the Renaissance should begin by reading two masterpieces of historical writing: Jacob Burckhardt, *Civilization of the Renaissance in Italy* (1st German ed., Basel, 1860), and Johan Huizinga, *The Waning of the Middle Ages* (1st Dutch ed., Haarlem, 1919), both available in various English editions. These works complement each other: Huizinga deals with France and the Low Countries; Burckhardt deals only with Italy and apologizes for having even mentioned Rabelais. No works of comparable standing in cultural history exist for the Renaissance as it affected England, the German lands, or other European countries.

On the concept of the Renaissance, see Wallace K. Ferguson, *The Renaissance in Historical Thought* (Boston: Houghton Mifflin, 1948), and Franco Simone, *Il rinascimento francese* (Turin: Società Editrice Internazionale, 1961), which adds new evidence for the seventeenth century in France. Invaluable evidence is contained in Herbert Weisinger's articles, especially "The Self-Awareness of the Renaissance," in *Papers of the Michigan Academy of Science, Arts, and Letters* 29 (1944): 561–567. See also Augusto Campana, "The Origin of the Word 'Humanist,'" in *Journal of the Warburg and Courtauld Institutes* 9 (1946): 60–73, and Hans Baron, "The *Querelle* of the Ancients and the Moderns as a Problem for Renaissance Scholarship," in *Journal of the History of Ideas* 20 (1959): 3–22. See also Federico Chabod, "The Concept of the Renaissance," in *Machiavelli and the Renaissance* (New York: Harper Torchbooks, 1965).

For general historical background, Edward M. Hulme, *The Renaissance, the Protestant Revolution, and the Catholic Reformation in Continental Europe* (New York: Century, 1914), is convenient. Excellent summaries may be found in *The New Cambridge Modern History*, Vol. I, *The Renaissance, 1493–1520* (Cambridge, U.K.: Cambridge University Press, 1957).

On humanism, two older works are still basic: Georg Voigt, *Die Widerbelebung des classischen Alterthums*, 2 vols., 3rd ed. (Berlin, 1893), and J. E. Sandys, *A History of Classical Scholarship*, 3 vols. (Cambridge, U.K.: Cambridge University Press, 1903–1938). All can profit from reading Erwin Panofsky, *Renaissance and Renascences in Western Art* (Stockholm: Almqvist and Wiksell, 1960).

Coming more particularly to philosophy, a useful guide is Paul O. Kristeller and J. H. Randall Jr., "The Study of the Philosophies of the Renaissance," in *Journal of the History of Ideas* 2 (1941): 449–496. Two leading mid-twentieth-century historians of Renaissance philosophy are Paul O. Kristeller and Eugenio Garin. Kristeller's major books are *The Classics and Renaissance Thought* (Cambridge, MA: Oberlin College by Harvard University Press, 1955), revised and enlarged in *Renaissance Thought* (New York, 1961); *Studies in Renaissance Thought and Letters* (Rome: Edizioni di Storia e Letteratura, 1956), which brings together many valuable articles; and *The Philosophy of Marsilio Ficino* (New York: Columbia University Press, 1943). Kristeller's work is solidly based on careful reading of the sources. Garin has dealt with philosophers of the Italian Renaissance extensively in his *La filosofia*, 2 vols. (Milan, 1947), and has compiled useful anthologies, including *Filosofi italiani del quattrocento* (Florence, 1942) and *Il rinascimento italiano* (Milan, 1941). An invaluable anthology, edited by Ernst Cassirer, Kristeller, and Randall, is *The Renaissance Philosophy of Man* (Chicago: University of Chicago Press, 1948).

For French thought during the Renaissance, see Augustin Renaudet, *Préréforme et humanisme à Paris pendant les premières guerres d'Italie* (1494–1517), 2nd ed. (Paris: Librairie d'Argences, 1953). For Germany, see Peter Petersen, *Geschichte der Aristotelischen Philosophie im*

protestantischen Deutschland (Leipzig: Meiner, 1921), and Max Wundt, *Die deutsche Schulmetaphysik des 17. Jahrhunderts* (Tübingen, 1939). On Spain see Carlo Giacon, *La seconda scolastica*, 3 vols. (Milan: Fratelli Bocca, 1944–1950). On Italy, besides Garin's works, see Giuseppe Saitta, *Il pensiero italiano nell'umanesimo e nel rinascimento*, 3 vols., 2nd ed. (Florence: Sansoni, 1961), and Hans Baron, *The Crisis of the Early Italian Renaissance*, 2 vols. (Princeton, NJ: Princeton University Press, 1955). Important studies are those of Ernst Cassirer, especially *The Individual and the Cosmos in Renaissance Philosophy*, translated by Mario Domandi (New York: Harper, 1964).

The Renaissance is a favorite topic for symposia, and the results are sometimes useful: See Tinsley Helton, ed., *The Renaissance* (Madison: University of Wisconsin Press, 1961), with contributions by various authorities. Two specialized studies are Carlo Angeleri, *Il problema religiosa del rinascimento* (Florence, 1952), and Georg Weise, *L'ideale eroico del rinascimento e le sue premesse umanistiche* (Naples: Edizioni Scientifiche Italiane, 1961). Useful articles appear in *Renaissance Studies* and in *Bibliothèque d'humanisme et de renaissance*.

OTHER RECOMMENDED WORKS

Artz, Frederick B. *Renaissance Humanism 1300–1500*. Kent, OH: Kent State University Press, 1966.

Coates, Willson Havelock, Hayden V. White, and J. Salwyn Schapiro. *The Emergence of Liberal Humanism: An Intellectual History of Western Europe*, Vol. 1. New York: McGraw-Hill, 1966–1970.

Hankins, James. *Plato in the Italian Renaissance*. Leiden: Brill, 1991.

Kahn, Victoria. *Rhetoric, Prudence, and Skepticism in the Renaissance*. Ithaca: Cornell University Press, 1985.

Kelly, Donald R. *Renaissance Humanism*. Boston: Twayne, 1991.

King, Margaret L. *Humanism, Venice, and Women: Essays on the Italian Renaissance*. Burlington, VT: Ashgate, 2005.

Kristeller, Paul Oskar, and Mooney, Michael, eds. *Renaissance Thought and Its Sources*. New York: Columbia University Press, 1979.

Mahoney, Edward P. *Two Aristotelians of the Italian Renaissance: Nicoletto Vernia and Agostino Nifo*. Burlington, VT: Ashgate/Variorum, 2000.

Moyer, Ann E. *Musica Scientia: Musical Scholarship in the Italian Renaissance*. Ithaca: Cornell University Press, 1992.

Palisca, Claude V. *Humanism in Italian Renaissance Musical Thought*. New Haven, CT: Yale University Press, 1986.

Parkinson, G. H. R., ed. *The Renaissance and Seventeenth Century Rationalism*. New York: Routledge, 1993.

Stinger, Charles L. *Humanism and the Church Fathers: Ambrogio Traversari (1386–1439) and Christian Antiquity in the Italian Renaissance*. Albany: State University of New York Press, 1977.

Trinkaus, Charles. *In Our Image and Likeness: Humanity and Divinity in Italian Humanist Thought*. Chicago: University of Chicago Press, 1970.

Neal W. Gilbert (1967)
Bibliography updated by Michael J. Farmer (2005)

RENAN, JOSEPH ERNEST

(1823–1892)

Joseph Ernest Renan, the French critic and historian, was born in Tréguier, Brittany. He studied for the priesthood at seminaries in Paris but left the seminary of Saint-Sulpice in 1845 to devote himself to secular teaching and writing. He contributed to the *Revue des deux mondes* from 1851 and the *Journal des débats* from 1853. He received a *docteur ès lettres* in 1852, was elected a member of the Académie des Inscriptions in 1856, and was elected to the Académie Française in 1878. He was appointed professor of Hebrew at the Collège de France in 1862, but the course was then immediately suspended until 1870. In 1884 he became administrator of the Collège de France.

Renan's abandonment of his priestly calling was largely determined by the doubts engendered by his philological study of the Bible. After leaving the seminary, he was strongly influenced by Marcelin Berthelot, the chemist, with whom he maintained a lifelong friendship. Another major influence was German idealism, particularly that of G. W. F. Hegel.

In one sense Renan's life's work can be seen as an attempt to expand the horizons of scientific rationalism by incorporating into it what was valid in idealist philosophy—principally the theme of development, particularly the theme of spontaneous evolution of the human mind. It was the historical aspect and the historical emphasis of Hegel's thought that appealed to Renan, for the cast of his own mind was fundamentally historical, not philosophical. Philosophy for him is not a discipline in its own right, and it is history, not philosophy, that should dominate science; "History is the necessary form of the science of the future." It is evident that Renan used the word *science* in the original sense of "knowledge"; "science" is not to be equated with the natural sciences. On the other hand, his philological and historical method is rationalistic and critical. He was interested, above all, in the evolution of languages and religions as manifestations of the development of the human mind, which is in turn the key to the universe. These manifestations and the universe itself, however, are concrete realities to be discovered through observation, experiment, criticism, and disciplined imagination. They are susceptible to this approach because they are the products of the interplay of natural causes according to constant laws. Renan denied in principle that there is any mystery in the world; what seemed mysterious would yield before the advancing frontiers of knowledge. This is the case in the human no less than in the natural

sciences. Renan, in contact with working scientists, rejected the simplistic notions of natural science characteristic of the positivism of Auguste Comte. He maintained that progress in the natural as well as in the human sciences depends on human judgments of the balance of probabilities on the evidence. He further maintained that all reality is in some degree historical, that the natural sciences (paleontology, for example) reveal the remote parts of history, and that the human and natural sciences can and must therefore be of mutual help.

Just as he banished all traditional metaphysics from philosophy, Renan rejected any supernatural content in religion. The true religion of humankind, in the sense of "a belief accompanied by enthusiasm which crowns conviction with devotion and faith with sacrifice," is that of science (that is, knowledge). Renan's argument runs as follows: The universe is characterized by change according to "laws of progress" under which the human mind becomes increasingly conscious of itself and the ideal is increasingly manifested amid the real: "The goal of the world is the development of mind." At the end of the process God, in the sense not of a creative providence but of an immanent ideal, will be realized. Since this ideal consists in the complete development of consciousness and in the attainment by that consciousness of the full measure of beauty and morality of which it is capable, science must be the great task of humankind. This task must be approached in the spirit neither of mere curiosity nor of mere utilitarianism but in the true religious spirit, seeking revelation of the divine.

The above sketch of Renan's thought is based mainly on his youthful work, *L'avenir de la science,* written in 1848 but first published in 1890. In his later philosophical writing he modified, but did not abandon, the fundamental position adopted there. Political and social events in France, in particular, damped his optimism and strengthened his skeptical and ironical streak. He began to have doubts about the "religion of science" to which he had turned when he abandoned Roman Catholicism. He became less sure that men had the capacity to attain adequate knowledge, and some of his own writing became tentative, cast at times in the form of dialogue. In his professional historical work, however, which always remained his chief concern, he stood fast by his views on the development of rationality out of instinct and on the progressive realization of God on Earth. Even in the new preface that he added to *L'avenir de la science* on its publication late in his life, Renan declared that his religion was still "the progress of reason, that is to say, of science." He had been too sanguine, too anthropocentric, and not entirely emancipated from Catholicism; the growth of knowledge had not, in fact, clarified human destiny. He confessed that he did not see how humankind could maintain its ideals if deprived of its illusions, but he retained his faith in knowledge as the supreme pursuit.

See also Hegel, Georg Wilhelm Friedrich; Idealism; Modernism; Natural Law; Rationalism.

Bibliography

WORKS BY RENAN

Averroes et l'averroïsme. Paris: Durand, 1852.

Essais de morale et de critique. Paris: Calmann-Lévy, 1859.

Vie de Jésus. Paris: Michel Lévy, 1863. Translated by C. E. Wilbour as *The Life of Jesus.* New York: Carleton, 1864.

Questions contemporaines. Paris: n.p., 1868.

Dialogues et fragments philosophiques. Paris: Calmann-Lévy, 1876.

Souvenirs d'enfance et de jeunesse. Paris: Calmann-Lévy, 1883. Translated by C. B. Pitman as *Recollections of My Youth.* Boston and New York: Houghton Mifflin, 1929.

L'avenir de la science. Paris: Calmann-Lévy, 1890. Translated by A. Vandam and C. Pitman as *The Future of Science.* Boston: Roberts Brothers, 1891.

WORKS ON RENAN

Allier, Raoul. *La philosophie d'Ernest Renan.* Paris, 1895.

Berthelot, René. "La pensée philosophique de Renan." *Revue de métaphysique et de morale* 30 (1923): 365–388.

Charlton, D. G. *Positivist Thought in France during the Second Empire.* Oxford: Clarendon Press, 1959. See Ch. 6.

Cresson, André. *Ernest Renan: Sa Vie, son oeuvre, avec un exposé de sa philosophie.* Paris: Presses Universitaires de France, 1949.

Mott, Lewis Freeman. *Ernest Renan.* New York and London: Appleton, 1921.

Séailles, Gabriel. *Ernest Renan: Essai de biographie psychologique.* Paris: Perrin, 1895.

W. M. Simon (1967)

RENOUVIER, CHARLES BERNARD
(1815–1903)

The French critical philosopher Charles Bernard Renouvier was born in Montpellier and was educated at the École Polytechnique, where he specialized in mathematics and natural science. At the school he came under the influence of the work of Antoine Cournot and of Auguste Comte, who at that time was an instructor (*répétiteur*) in higher mathematics there. In 1848 Renouvier published in Paris his *Manuel républicain de l'homme et du citoyen*, a volume addressed to schoolteachers, which urged the

preaching of socialism. But his political views were frustrated by the coup d'état of Louis Napoleon, and he retired from active participation in politics to write philosophy. Renouvier never held an academic position but worked as a private individual, producing one of the longest series of philosophical works in French history. In 1867 he began the publication, with his friend and collaborator François Pillon, of *L'année philosophique*, a monthly that propagated Renouvier's philosophical doctrines. These doctrines were chiefly expounded in a series of books, constantly revised by Renouvier, the *Essais de critique générale*, the final edition of which appeared in 1897. He continued writing up to the time of his death, his last work being *Le personnalisme* (1903). Though his pluralism and his personalism anticipated some philosophical doctrines of the early twentieth century, his main influence was upon his French contemporaries.

NEOCRITISM

Renouvier's general position is called neocriticism, because it took the method of Immanuel Kant's critical philosophy as its starting point. But though Renouvier started with Kant's method, he did not accept Kant's conclusions but used them rather as a basis from which to launch a set of ideas often critical of Kant.

Renouvier laid down as an integral part of his philosophy what he called the "law of numbers," according to which every cardinal number is an ultimate individual, finite and irreducible. Mathematics is the paradigm of thinking, and the law of contradiction is more clearly manifested in mathematical operations than anywhere else. But the term *mathematics*, as Renouvier used it, was restricted to arithmetic, and he derived the nature of numbers exclusively from the cardinal numbers. This led him to deny the existence of any infinite, for he maintained—unable to anticipate the work of Georg Cantor—that an infinite number was a contradiction in terms. Renouvier extended his criticism of the notion of infinity beyond numbers to deny the infinity of space and time as well.

Renouvier recognized that knowledge is relative to its premises and to the person who laid down the premises; nevertheless he could not accept the relativity of logical processes. There is a distinction involved here between logic and the psychology of thought. Just as each number is a distinct and separate entity, so is each human being. And just as the characteristics of each number—duality, triplicity, and so on—can never be reduced to, or "reconciled" with, the characteristics of any other number, so each human being is not exactly like any other and cannot be merged into a general group-consciousness or absorbed into an absolute mind. Knowledge is always the property of individual knowers, and the distinction between knowledge and belief disappears. What an individual knows is what seems reasonable to him, and his contribution to knowledge can never be subtracted. The subtraction can be made verbally, to be sure, but to do so is to alter the character of cognition, which is essentially judgment.

PHENOMENA

Renouvier also differed from Kant in his doctrine of phenomena. Phenomena are not the appearances of anything other than themselves. They are neither illusions nor purely subjective beings. They are sui generis, being whatever we perceive or whatever we make judgments about. He granted that the name is unfortunate except insofar as it indicates appearances. Because there are no things-in-themselves, Renouvier criticized Kant's antinomies, which hold good only if there are noumena. His attack on the first antinomy, for example, was based on its use of the concept of infinity. Since infinity is an inherently inconsistent idea, Renouvier asserted that the world must have had a beginning in time and that space is limited. The domination of the number concept as a conceptual model appears here in full force. For Renouvier, the numbers begin with one, since zero and negative numbers are not really numbers, and spaces are the spaces of individual discrete beings, there being no such entity as number-in-itself or space-in-itself.

There exists within the number series the category of relation. For the numbers are ordered, and order is a kind of relation. All other categories are, for Renouvier, forms of relation, but of relation as discovered within the framework of an individual's consciousness. There turn out to be nine categories—relation, number, position, succession, quality, becoming (*devenir*), causality, purposiveness, and personality. Each has its thesis, antithesis, and synthesis; and all are rooted in the phenomenal world as judged by us. It is uncertain whether Renouvier attempted to derive his categories in the manner of Maine de Biran from personality—our acting as a cause, our seeking ends, our sensory discriminations (which might produce the separateness of quantity and quality), spatial positions, moments in time, and the intervals between them—or whether his assertion that personality is one of the categories is derived from his premise of the law of numbers. In any event, just as each number has its own distinctive quality, its own position in the numeral order, and its many relations to other numbers, determined not

only by its own character but also by that of the other numbers, so the human being has his own personality and displays the other categories not only as a distinct entity but also as a perceptive consciousness.

The parallelism between the ways in which a man judges, perceives, and knows and the ways in which he as a person differs from other beings pervades Renouvier's writings. Thus, because one acts to achieve one's purposes, it follows that both causality and purposiveness exist within the human being and must likewise be combined in the phenomenal world. A cause determines the path of an event, but the direction of that event is determined by that which participates in it.

Since no two events are exactly alike, the deterministic factor in nature is mitigated by chance. Renouvier probably got this argument from Cournot, who also insisted upon the probabilistic element in nature. To frame a law or a generalized description depends upon our ability to discover absolutely homogeneous phenomena or groups of phenomena. If this is impossible, then generalizations are at most only probable. But at the same time, each individual phenomenon contributes something to the events of which it is a part, and that contribution in the very nature of things cannot be predicted.

INDETERMINISM

The problem of causation arises with regard to human beings in the form of the antithesis between free will and determinism. Since every act of consciousness is a relation between a perceiving subject and that which is perceptible, then as soon as a conscious act is formulated and made clear to the perceiving mind, it will be organized in terms of the categories. But there is a choice among the various categories to be applied, for we are not forced either to quantify or qualify, to count or to locate, to assign a date or to recognize a cause. The categories limit the possibilities of judgment but have no inherent order of predominance. In other words, Renouvier held that when we see a phenomenon, for example, a tree, we are not forced first to judge it as green, then as distant, old, fan-shaped, simple, or what you will. The order of judgment is determined by us, and we are free, within the range of possible categories, to judge it as we will. The selection of a category or group of categories depends on our free choice in accordance with our interests at the moment of judging.

Freedom cannot be proved, nor can determinism. Both are assumptions utilized in view of their consequences. These consequences may be purely intellectual or may be moral or practical. But freedom itself rests upon the inherent individuality of the human will, an individuality which cannot be completely absorbed into any larger class of beings. Insofar as any being is unique, to that extent it is undetermined or self-determined. And insofar as it is identical with other beings, to that extent the homogeneity of its class accounts for the regularity of its behavior. In short, individuality and freedom are synonymous terms, and Renouvier even called freedom the principle of individuation. The consequence is that just as the personal equation enters into all judgments, so the only certainty we have is the certainty of our judgments. Renouvier put it as follows:

> Certitude is not and cannot be absolute. It is a condition and act of man—not an act and a condition in which he grasps immediately that which could not be immediate, i.e., facts and laws external and superior to present experience, but rather one in which he posits his awareness as it exists and as he maintains it. Strictly speaking, there is no certitude; there are only men who are certain. (*Traité de psychologie rationnelle*, Paris, 1912, Vol. I, p. 366)

But indeterminism is not limited to human judgments. It extends also to history. For since history is in part made up of human behavior, human decisions must be included in its scope, and there is no way of eliminating them. One can, of course, describe the environment of human life, its stability, and its mutability; but if it remains stable, that is because human beings have not changed it, and if it changes, that is due to human acts as much as to natural disasters. People modify their living conditions, not as a group acting as one person, but as a collection of individuals. Their reasons for doing so may vary, as is inevitable, and of course they are not able to modify their conditions completely. But Renouvier emphasized the importance of human decisions for the way in which individuals will live, since the ability of human beings to make choices makes it impossible to lay down either a law of universal progress toward the good or one of constant degeneration. Hence Renouvier rejected historical laws, such as those of Comte and G. W. F. Hegel, though he was attracted to meliorism.

ETHICS

If there is no historical law dooming humankind to move in any predetermined direction and if history only records actual change, the question arises of the relation of history to ethics. People make moral judgments and act so as to achieve what they believe to be right. Morals, then, are not the result of history, though what happens

in history reflects our moral judgments. Morals are rather the source of historical changes, and if we are to appraise historical events, we shall have to do so in moral terms. This clearly requires a definition of good and evil, and in view of the radical individualism of Renouvier this might seem an insurmountable task. But he identified evil with conflict, conflict both between persons and between groups of persons. For warfare is in essence the prevention by one or more persons of the fulfillment of the volitions of others. Hence tyranny, slavery, and conquest are to be condemned. This assumes that it is possible for a group of enlightened people to respect the individuality of their fellows and for all to live in peace.

In his fictional account of what history might have been, *Uchronie* (1876), Renouvier claimed that the secret of human happiness lies in our recognition of the individual's freedom. If at any epoch people had accepted individual freedom wholeheartedly, he argued, universal peace and harmony would have prevailed. Religious, economic, and national wars would have ended at once; for everyone would have taken it for granted that each person has a right to his own religious views, to the satisfaction of his own economic interests, and to his own national loyalties. Renouvier held that education alone could bring this about, though he had no illusions that proper education was ever likely to be instituted. The dogma of historical determinism has had too firm a hold on human will power and has brought about acquiescence, sloth, injustice, and ignorance.

The basic premises of Renouvier's *Science de la morale* (1869) are that human nature is rational and that people believe themselves to be free. Their belief in freedom leads individuals to act for what they judge to be better, and their rationality guides them in their choice of ends. To act morally is to act rationally. By doing so we rise above the beasts; we recognize the humanity in our fellows and respect it. For this reason Renouvier became a bitter opponent of the Catholic Church and of monarchy and urged his readers to turn to Protestantism as the religion of individual conscience. To him Protestantism was the religion of a personal God—not an absolute and unchanging Being, omniscient and omnipotent, but finite, limited, free, and the guarantor of our freedom. God's existence is not proved, but it is a reasonable hypothesis drawn from the existence of our moral objectives. Running through Renouvier's many works are the premises that the plurality of existing things is irreducible; that chance is real and is reproduced in individual freedom of choice; that time and novelty really exist; and that no absolutes or infinites exist.

See also Comte, Auguste; Cournot, Antoine Augustin; Determinism and Indeterminism; Hegel, Georg Wilhelm Friedrich; History and Historiography of Philosophy; Infinity in Mathematics and Logic; Kant, Immanuel; Maine de Biran; Neo-Kantianism.

Bibliography

WORKS BY RENOUVIER

Essais de critique générale. 4 vols. Paris, 1854–1864.

Science de la morale. 2 vols. Paris, 1869.

Uchronie, l'utopie dans l'histoire. Paris, 1876.

Esquisse d'une classification systématique des doctrines philosophiques. 2 vols. Paris, 1885–1886.

Philosophique analytique de l'histoire. Paris, 1896–1897.

Le personnalisme. Paris, 1903.

Lettres de Renouvier à Henneguy (1861–1899), edited by Marcel Méry. Gap: Editions Ophrys, 1969.

Manuel républicain de l'homme et du citoyen (1848). Paris: Garnier, 1981.

Les dilemmes de la métaphysique pure. 2nd ed. Paris: Presses universitaires de France, 1991.

Victor Hugo, le philosophe. Paris: Maisonneuve et Larose, 2002.

WORKS ON RENOUVIER

Blais, Marie-Claude. *Au principe de la République: Le Cas Renouvier*. Paris: Gallimard, 2000.

Fedi, Laurent. *Le problème de la connaissance dans la philosophie de Charles Renouvier*. Paris: L'Harmattan, 1998.

Foucher, Louis. *La jeunesse de Renouvier et sa première philosophie*. Paris: Vrin, 1927.

Grande, Sonia Kase. *Charles Renouvier and the Search for Intellectual and Social Synthesis in Nineteenth-Century France*. 1975.

Gunn, J. Alexander. *Modern French Philosophy*. London: Unwin, 1922. Contains an excellent bibliography.

Hamelin, Octave. *Le système de Renouvier*. Paris: Vrin, 1927.

Logue, William. *Charles Renouvier, Philosopher of Liberty*. Baton Rouge: Louisiana State University Press, 1993.

Milhaud, Gaston. *La philosophie de Charles Renouvier*. Paris, 1927.

Séailles, Gabriel. *La philosophie de Charles Renouvier*. Paris: Alcan, 1905.

Turlot, Fernand. *Le personnalisme critique de Charles Renouvier: Une philosophie française*. Strasbourg: Presses Universitaires de Strasbourg, 2003.

George Boas (1967)
Bibliography updated by Michael J. Farmer (2005)

RENSI, GIUSEPPE
(1870–1941)

Giuseppe Rensi was an Italian skeptical philosopher and professor of philosophy at the universities of Messina and

Genoa. Rensi first upheld a religiously or theistically oriented idealistic philosophy, defending it in a number of essays and fostering it through his translations of the works of Josiah Royce. He contrasted his theistic "constructive idealism" with the "immanentistic idealism" of Benedetto Croce and Giovanni Gentile; he regarded the latter as a temporary position that, if developed coherently, would have led to constructive idealism. According to Rensi, an idealism that does not arrive at God subtracts reality both from the external world, which then becomes a set of ideas, and from the human spirit, which is then resolved into a set of ideas without a subject.

After World War I, regarded by Rensi as proof of the fundamental irrationality of the world, he began to defend a radical skepticism based on the multiplicity, irreducibility, and irreconcilability of opinions, the reasons used to justify them, and some aesthetic tastes and moral ideas. Rensi held that the traditional objection to skepticism—that it contradicts itself by asserting that there is no truth while dogmatically asserting its own truth—was a purely verbal objection, because the skeptic holds his position against any doctrine taken in itself by showing the contradictions and shortcomings of that doctrine. Therefore the skeptic does not assert that there is no truth but instead that a particular doctrine that claims to possess truth does not and cannot possess truth. Skepticism, in other words, shows the disagreement of reason with itself both within the views of one man and between the views of different individuals. War, the conflict of rights and of political powers, and the contradictory character of philosophies are, according to Rensi, proofs of the intrinsic contradiction in reason. Skepticism does not exclude faith but stems from the preservation of faith. The skeptic is skeptical not because he does not believe but because others believe differently than he; that is, they believe that which he considers absurd.

Rensi had been a socialist in his youth but later came to defend authority. He wished to give to power (and even violence) the function of helping man escape from the chaos of opinions and contrasting interests and of forming a people into an economic, political, and spiritual unity. Authority need not base itself on reason, because it creates for itself the reason of all that it wishes. Although these ideas seem close to those of fascism, Rensi quickly declared himself opposed to fascism and remained so until his death.

According to Rensi, skepticism implies atheism in the field of religion. The refinement of religion that leads to regarding God as inaccessible to the senses and to human powers makes God a nonbeing, the pure and simple negation of every reality accessible to man. From this point of view, both negative theology and mysticism demonstrate atheism. Atheism is still a religion because it is an answer—even if a negative one—to the problem of supreme reality. Unlike other religions, atheism is absolutely disinterested because it contains no egoistic motive and because it places man before the mystery of the All without his being able to expect from the All any help for his own needs.

After 1922, when the absolute idealism of Croce and Gentile assumed the status of an official or semiofficial philosophy in Italy, Rensi accentuated his polemic against it and affirmed the theses most opposed to those of idealism: materialism and pessimism. The Kantian system, considered to be idealistic by the idealists, seemed to Rensi to justify materialism because the Kantian forms of intuition and of thought that condition phenomena, and therefore the totality of nature, are not created by the self but constitute "consciousness in general," which is the intelligibility of the things themselves. According to Rensi, the Kantian doctrine is, therefore, that nature gives reality and knowability to natural things, that things generate of themselves, and of themselves are spatial, temporal, perceptible, and representable; in one word, they are material.

Rensi held that materialism implies pessimism because a material nature deprived of any finality offers man no guarantee and necessarily includes evil, error, and conflicts. For a man who lives in such a nature, morality, when not based on an egoistic calculus or subjected to an imposed code, is a disinterested recognition of evil and a protest against it. It is therefore pure folly. Nevertheless, all of Rensi's works contain a mystical and religious strain, a sense of mystery and of a force that, the triumph of evil in nature and in history notwithstanding, reveals itself in the interiority of man. Rensi condensed this feeling into the phrase "Atoms and the void—and the divine in me."

See also Atheism; Croce, Benedetto; Gentile, Giovanni; Idealism; Materialism; Pessimism and Optimism; Royce, Josiah; Skepticism.

Bibliography

WORKS BY RENSI

Il genio etico e altri saggi. Bari, 1912.

La trascendenza. Turin, 1914.

Lineamenti di filosofia scettica. Bologna, 1919.

La filosofia dell'autorità. Palermo, 1920.

La scepsi estetica. Bologna: N. Zanichelli, 1920.

Introduzione alla scepsi etica. Florence, 1921.
Apologia dell'ateismo. Rome, 1925.
Realismo. Rome, 1925.
Il materialismo critico. Rome, 1927.
La morale come pazzia. Modena: Guanda, 1942.

WORKS ON RENSI

Morra, G. *Scetticismo e misticismo nel pensiero di Giuseppe Rensi.* Palermo, 1958.
Nonis, P. *La scepsi etica, di Giuseppe Rensi.* Rome: Studium, 1957.

Nicola Abbagnano (1967)
Translated by Nicholas Spadaccini

REPRESENTATIVE REALISM

See *Realism*

REPUBLICANISM

Republicanism is one of the great traditions of Western political thought. To say that republicanism is a "tradition" of political thought is to say that distinctively republican ideas about politics have been championed by a number of authors in the history of political theorizing, and that many of the later authors who championed those ideas consciously drew on and developed the work of earlier ones. This continuity of reference and influence makes it possible to trace a republican strand in Western political writing. But what ideas about politics are distinctively republican? What ideas define the republican tradition?

The republican tradition is often associated with the claims that citizens can only be free in a free society, that the opposite of freedom is a state of dependence akin to slavery, that societies are most likely to enjoy freedom and to realize their common good when they are governed by politically engaged citizens who act from the civic virtues, and that the pursuit of the common good is undermined when citizens' virtues are corrupted by selfishness, luxury, and ambition. These claims turn up consistently in the writings which make up the republican tradition from Rome at least through the eighteenth century.

The classic texts of the republican tradition were produced in political circumstances very different than those of the early twenty-first century. These texts commend ways of life, such as a life of politically active citizenship, that are open to relatively few citizens of large, modern societies such as England and the United States. The political threats in the face of which these texts were produced were quite different than the threats to liberty and equality posed by the modern states of late capitalism. Republicans' emphasis on civic virtues raises the possibility that republican politics would be difficult to sustain under conditions of moral pluralism. It is therefore not immediately clear that republicanism can provide guidance to modern politics. Even if republican ideas can provide some guidance, it is far from clear that republicanism alone can provide sufficient guidance. Perhaps republican ideas about politics are most useful as supplements to political theories that belong to other traditions of political thought and that are explicitly framed for current conditions.

Republicanism has enjoyed a revival in legal and political philosophy since the 1980s. Those who have revived republicanism in these disciplines have tried to apply the insights and arguments of the tradition to contemporary politics. But some participants in the revival themselves seem to raise questions about the sufficiency and distinctiveness of republicanism. They move easily between republicanism and democratic liberalism and seem content to describe themselves as both republicans and liberals. The fact that they do so raises questions about whether any version of republicanism that is of more than historical interest is faithful to ideas that have distinguished the republican tradition. It also raises questions about whether versions of republicanism that bear on contemporary politics are part of a strand that ought to remain distinct from other movements of political thought. Perhaps the insights of republicanism are best absorbed into liberalism, a tradition which has its origins in the early modern period.

Late twentieth-century work on republican views of liberty has, however, changed the way historians and political theorists think about republicanism. It has already shed new light on some of the classic texts in the republican tradition. It promises to show how some of the claims characteristic of republican writing can be systematically united and given a theoretical basis. And it promises to illuminate deep and interesting differences between republicanism and liberalism. If these promises can be made good, then republicanism's claim to be a distinctive family of political thought—and one of continuing relevance—can be vindicated.

ROMAN REPUBLICANISM

The origins of the republican tradition lie in the writings of Roman political thinkers, such as Cicero and Sallust, who lamented and analyzed Rome's transformation into an empire just before the beginning of the common era.

They came to be called "republicans" in part because the form of government they favored was that of pre-imperial Rome—a regime for which Cicero popularized the name "the republic." They are also called "republicans" because of the features of that government that they seized upon when arguing that rule by the republic's government was superior to imperial rule. The words "republic" and "republicanism" derive from the Latin phrase *res publica*, which means "public matter."

According to these thinkers, the republic was better suited to advance the common good of the Roman people than the empire was because, unlike the empire, its government was participatory. It was governed by public-spirited citizens—in particular, public-spirited citizens serving in the Roman senate—who devoted themselves to the pursuit of public matters rather than to the pursuit of their own wealth and ambition. In the "Dream of Scipio," Cicero famously claimed that those who dedicate themselves to the preservation of the republic would enjoy an eternal reward. Such devotion to the republic, he thought, required civic virtue. Republican writers claimed that the Roman republic was subverted by corruption. It was subverted, they said, by those who sought and used political power to further their own ends rather than the common good of the Roman people.

RENAISSANCE AND EARLY MODERN REPUBLICANISM

Republicanism had little impact on the political thought of the Middle Ages, though some of the writings of Cicero were certainly known to such great medieval philosophers as Augustine and Thomas Aquinas. But the writings of the Roman republicans were important sources for political thinkers in Renaissance Italy who wanted to maintain the freedom of city-states against internal and external threats. They were also important sources for thinkers in seventeenth-century England who opposed the absolutist tendencies of the Crown. These writers located themselves in the tradition of Roman republicanism. They drew on republican claims about the importance of political participation, the need for a virtuous citizenry and the threats posed by corruption and self-interest, even as they adapted those claims to their own situations.

Among Italian thinkers, the greatest was undoubtedly Machiavelli, especially the Machiavelli who wrote *The Discourses*. Machiavelli believed that the citizens could only enjoy freedom if their city was free. One of the most significant threats to a city's freedom, he argued, was an internal threat: the threat posed to a city's good government by factions that would pursue their own interests once in power. Inspired by the political ideas of the humanist tradition and the writings of Roman republicans, Machiavelli argued that the dangers of factionalism could best be averted by a government of citizens committed to the common goods of civic wealth, glory, and independence.

English republicans such as John Milton and James Harrington were less concerned with the threat of faction than they were with what they regarded as the absolutist tendencies of the monarchy. They maintained that absolute power corrupted, but it did not corrupt only the monarch. A powerful court, they thought, was one that corrupted courtiers and politicians by encouraging their dependence upon royal favor. English republicans stressed the importance of the civic virtues, among which they numbered independence and frugality. They held up as models of good government the republics of Renaissance Italy, and the Roman republic. Historians sometimes call them "commonwealth men" because of their support for the Puritan commonwealth.

THE REPUBLICAN REVIVAL

The last third of the twentieth century saw a resurgence of scholarly interest in republicanism, primarily but not exclusively in the English-speaking world. The resurgence of interest among American Constitutional lawyers in the 1980s and 1990s came to be known as the "republican revival." That term can be stretched to encompass the contemporaneous revival of interest in republicanism among political philosophers. The republican revival among lawyers and philosophers was preceded by and drew upon work by historians of Renaissance political thought and by historians of the American founding. Indeed it was because of the resurgence of interest among historians that so much has been learned about early modern republicanism. It is useful to begin a survey of the republican revival with a look at some of the historical work that preceded and influenced legal philosophers responsible for the revival.

In the 1950s Louis Hartz articulated what was for a time the received orthodoxy about the intellectual foundations of the American Revolution and founding. According to Hartz, the revolutionaries and founders owed their greatest intellectual debts to the classical liberalism Hartz ascribed to John Locke. In the 1960s historians of the American founding and its intellectual antecedents, notably Bernard Bailyn and Gordon Wood, raised serious challenges to this orthodoxy. Bailyn and Wood argued powerfully that the intellectual underpin-

nings of the revolution and the founding period were in large part republican, drawn from the English commonwealth tradition of the previous century. John Pocock, who traced the origins of the commonwealth thought to Renaissance Italy, provided an even longer genealogy for American republicanism.

Bailyn and Wood mined the pamphlets and popular literature of early America for evidence of republican political thinking. The expressions of republicanism they found there included pervasive emphasis on the need for citizens to dedicate themselves to the common good and on the deleterious effect of faction and the elevation of private over public interest, concern with the corrupting effects of various forms of dependence upon Britain (including dependence on its monied and manufacturing interests), and the description of American dependence as a condition of slavery. Pocock, Bailyn, and Wood all maintained that the republicanism of the American founding was only gradually eclipsed by other forms of political thought in the years or decades that followed.

The question of whether and to what extent the American founders were republicans is a question of some importance for legal philosophy. The founding period of the United States was the period in which the body of the American Constitution and the Bill of Rights were written. The conclusion that the founders owed deep intellectual debts to republicanism arguably has profound implications for how the Constitution and the Bill of Rights should be read and applied. The argument that it has such implications seemed especially pressing to legal scholars at a time when some were defending originalist canons of Constitutional interpretation. In the 1980s Constitutional scholars began to draw on the historical work of Pocock, Bailyn, Wood, and others, and initiated the republican revival in legal scholarship.

The leading figures of this revival, such as Cass Sunstein and Frank Michelman, emphasized the participatory strain of republicanism. Republican government, according to these thinkers, is government by citizens who participate in politics. The politics in which they are to participate is to be deliberative: citizens of a republican regime are to participate in collective deliberations about public matters. Such public deliberation, they argued, promises to combat the factionalism and self-interest that republicans had traditionally seen as undermining good government. It does so because the process of deliberating with others is not one of bargaining in which parties try to satisfy the preferences they have formed before public deliberation begins. Rather, it is to be a process of reasoning with others about how to advance the common good. When citizens reason together about the common good, they are forced to rethink whatever self- and group-interested preferences they may bring into public deliberation.

Republican accounts of politics had previously been addressed to societies much smaller than the democracies of the late twentieth century. Framing a version of republicanism adequate for such large societies required imagining institutional forms through which republican government could be exercised within them. The leaders of the republican revival in the law offered republican readings of the American constitution and drew out the implications of those readings for a host of questions in public law, from environmental law to campaign finance reform.

The republicanism offered by republican revivalists in the legal academy—like the republicanism uncovered by historians of the American founding—emphasized the value of political participation, the importance of their commitment to the common good and the threat posed by citizens' unregulated pursuit of self- and group-interested preferences. Because of these emphases, republicanism seemed to offer a healthy corrective to the individualism, self-interest, acquisitiveness, and withdrawal from public life that some thinkers, such as Michael Sandel, have alleged that liberalism encourages. Yet the republicans in the legal academy saw significant continuities between their own views and some forms of liberalism, particularly between their own views and the version of liberalism developed and refined by John Rawls from the 1960s until his death in 2003. By a decade after the republican revival began in American law schools, some of its leading figures had ceased to insist that there was anything distinctively republican about their views. Even some who continued to describe themselves as republicans, such as Sunstein, also described themselves as liberals and "deliberative democrats" as well. The development of a republicanism that was explicitly contrasted with liberalism had to await the republican revival in constructive political philosophy.

According to some leading republican political philosophers, the differences between republicanism and liberalism lie, not in the former's emphasis on political participation and civic virtue or in the latter's emphasis on individual rights, but in the very different conceptions of political freedom associated with each. Liberalism, as its name suggests, is a political philosophy that values liberty. The liberty that liberals are sometimes said to value is what has come to be called "negative liberty." Someone enjoys negative liberty to the extent that she can act as she

likes, without external impediments. Political liberty is the freedom citizens enjoy in political society. Those who identify political liberty with negative liberty must think that even the best law is an external impediment to action, and so interferes with citizens' political liberty. If they also think, as liberals do, that it is the job of government to promote and secure political liberty, then they must also think government should rely as little on these impediments as possible. It should secure as much negative liberty for citizens as is compatible with the enforcement of laws needed to maintain public order.

Some of the most prominent republicans who have contrasted their views with liberalism have contrasted them with versions of liberalism which equate political liberty with negative liberty. They have introduced another kind of freedom, which they call "freedom as nondomination." They have argued either that political liberty includes both negative liberty and liberty as nondomination, or that it consists in liberty as nondomination alone. To appreciate the differences these republicans see between liberalism and their own views, it is necessary to see what it is for one agent to dominate another.

One agent dominates another just in case the former is in a position to interfere arbitrarily with the choices of the latter. An agent is in a position to interfere arbitrarily with another's choices just in case that agent is able to interfere with the other's choices without having to take the latter's interests into account. This way of characterizing domination implies that there are two important differences between liberal views which identify political liberty with negative liberty and republican views which either equate political liberty with liberty as nondomination, or which claim that political liberty includes liberty as nondomination.

One difference is that, according to the latter, not all laws restrict citizens' freedom. When political authorities take account of the interests of citizens in the enactment and enforcement of law, they do not dominate citizens. They do not dominate them because, though the laws may interfere with citizens' freedom of action, they do not do so arbitrarily. Therefore, though these authorities compromise citizens' freedom on liberal accounts which equate liberty with negative liberty, they do not do so on republican accounts.

The other difference is that republicans think one person can restrict another's liberty just by being in a position to interfere with him arbitrarily, even if she never actually interferes with him at all. Thus a political authority who can exercise power without accountability, but who chooses to enact laws which further the common good, still dominates citizens. These citizens therefore lack political freedom on republican accounts but not on liberal ones.

The account of liberty as nondomination has been stated and defended most notably by Philip Pettit, beginning in the mid-1990s. Pettit labels that account of freedom a "republican" account because he claims that it is found in the seminal texts of the republican tradition. He argues quite convincingly that, by taking freedom understood as nondomination as the supreme political value, he can account for why republicans have valued political participation and why they have maintained that citizens are free only in free societies. Quentin Skinner, the historian of Renaissance republicanism, also has claimed to have found a distinctive conception of liberty in the republican tradition. In response to Pettit's work, Skinner has argued that republican political liberty includes both negative liberty and liberty as nondomination. Whether Pettit's or Skinner's view of political liberty is more faithful to the texts remains a matter of scholarly debate. What is beyond debate is that Pettit's conceptual work on republican liberty has greatly influenced historical work on republicanism, and that Pettit and Skinner have taken the republican revival to a new level of philosophical sophistication. Questions remain, however, about exactly where their versions of republicanism differ from prominent forms of liberalism which do not equate political liberty with negative liberty.

The liberalism developed by Rawls in the last third of the twentieth century has been enormously influential. Rawls argued for principles of justice which, he maintained, would be agreed to in what he called "the original position." The original position is, like the state of nature in Locke's work, a condition appropriate for writing a social contract. Thus the principles of justice Rawls defends are principles citizens would choose for themselves under the conditions appropriate for such a choice. Rawls calls a society which is regulated by those principles a "well-ordered society." When citizens live in a well-ordered society and when their own plans are in accord with the principles of justice, they live under and act from principles they would give to themselves. To be autonomous is, literally, to give oneself a law. Citizens who live in a well-ordered society and act from principles they would give themselves therefore enjoy an important form of autonomy, which Rawls calls "political autonomy."

One question republicans need to answer is how their view of political liberty differs from a view of political liberty like Rawls's, according to which political lib-

erty includes political autonomy as an important ingredient. Another is whether they think Rawls's well-ordered society would include injustices or obstacles to freedom—for example, instances of domination—that a republican regime would not. This question will require a complicated answer because, while Rawls thinks that political freedom includes political autonomy, he does not equate the two. He insists that, in a well-ordered society, the liberties exercised in political participation will have what he calls "fair value." Citizens' possession of political liberties which have fair value may be required by their political autonomy, but it seems to be distinct form of political freedom. Moreover, it is at least arguable that when citizens enjoy this form of political freedom—when these liberties have fair value—much of the political domination that concerns republicans will be eliminated. Finally, republicans need to ask whether a republican society would allow injustices that a well-ordered society would eliminate.

Republicanism has undoubtedly been a philosophically interesting tradition of thought which exercised great influence at important points in Western political history. Since the late twentieth century it has been revived with brilliance and ingenuity. But until contemporary republicans answer these questions, it will be difficult for them to maintain that republicanism is superior to all forms of liberalism. Until they answer them, it will also be unclear whether republicanism can stand on its own as a theory of contemporary interest, or whether the insights that have been systematized by its most sophisticated exponents are better incorporated into some version of liberal theory.

See also Cosmopolitanism; Libertarianism; Multiculturalism.

Bibliography

Bailyn, Bernard. *The Ideological Origins of the American Revolution*. Cambridge, MA: Harvard University Press, 1992.

Cicero, Marcus Tullius. *De re Publica*. Translated by Clinton Walker Keyes. Cambridge, MA: Harvard University Press, 1959.

Harrington, James. *"The Commonwealth of Oceana" and "A System of Politics."* Cambridge, U.K.: Cambridge University Press, 1992.

Hartz, Louis. *The Liberal Tradition in America*. New York: Harcourt, 1991.

Machiavelli, Niccolo. *The Discourses*. New York: Penguin, 1984.

Michelman, Frank. "Law's Republic." *Yale Law Journal* 97 (1988): 1493–1538.

Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Cambridge, U.K.: Cambridge University Press, 1997.

Pocock, J. G. A. *The Machiavellian Moment: Florentine Political Thought and the Atlantic Republican Tradition*. Princeton, NJ: Princeton University Press, 1975.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1999.

Rogers, Daniel T. "Republicanism: The Career of a Concept." *The Journal of American History* 11 (1992): 11–38.

Sandel, Michael J. *Democracy's Discontent: America's Search for a Public Philosophy*. Cambridge, MA: Harvard University Press, 1996.

Skinner, Quentin. "The Idea of Negative Liberty: Philosophical and Historical Perspectives." In *Philosophy in History: Essays on the Historiography of Philosophy*, edited by Richard Rorty, J. B. Schneewind, and Quentin Skinner, 193–221. Cambridge, U.K.: Cambridge University Press, 1984.

Skinner, Quentin. *Liberty Before Liberalism*. Cambridge, U.K.: Cambridge University Press, 1998.

Skinner, Quentin, ed. *Machiavelli and Republicanism*. Cambridge, U.K.: Cambridge University Press, 1990.

Sunstein, Cass. "Beyond the Republican Revival." *Yale Law Journal* 97 (1988): 1539–1590.

Wood, Gordon. *The Creation of the American Republic*. Chapel Hill: University of North Carolina Press, 1998.

Paul Weithman (2005)

RESCHER, NICHOLAS

(1928–)

Born in Hagen, Germany, where his father had established a law practice after serving as a German army officer in the First World War (1939–1945), Rescher's family emigrated to the United States in 1938, and he was educated there, receiving his PhD from Princeton University in 1951 at the age of 22. Since 1961 he has taught at the University of Pittsburgh, where he serves as University Professor of Philosophy and also as vice chairman of the Center for the Philosophy of Science. He has published more than 300 articles in scholarly journals, has contributed to many encyclopedias and reference works, and has written more than 100 books in various areas of philosophy, including epistemology, metaphysics, value theory and social philosophy, logic, the philosophy of science, and the history of logic.

In various publications Rescher has developed a detailed and systematic theory on the nature and limits of human knowledge along with its central implications for metaphysics and for the theory of values and ethics broadly conceived. Best viewed as an analytic pragmatist, Rescher has sought primarily in many books and essays to revive and refurbish the idealistic tradition in epistemology and metaphysics. Although he has written extensively on metaethics and issues of value and justice, his primary

efforts in epistemology and metaphysics constitute the central focus, and his approach to philosophy is comprehensively expressed in a trilogy titled *A System of Pragmatic Idealism*.

Generally, Rescher affirms the centrality of the natural sciences as the privileged source of understanding the nature of the empirical world and as directing our actions within it. He emphasizes, however, that the presuppositions of the natural sciences cannot be directly defended in natural science without circularity; such presuppositions, therefore, fall into the realm of metaphysics and are to be defended philosophically, philosophy being distinct from, but inextricably dependent on the deliverances of, natural science.

Rescher also sees the scientific method(s) as the product of an evolutionary process of rational selection, which leaves us with only those methods that have been proven to work by way of providing reasonably precise predictions of our sensory experience. In short, the methods of natural science, as well as their presuppositions, find their justification ultimately in the fact that we have a deep need for the products of natural science and epistemology, thereby underscoring the deeply practical or pragmatic nature of the whole of the cognitive enterprise, and whatever theoretical conclusions we reach therein (Rescher 1992–1994, 2001).

Regarding foundational beliefs or basic knowledge, Rescher affirms that basic beliefs, like all factual beliefs, are fallible and hence subject to revision in the light of ongoing evidence. Such beliefs begin as working presumptions about how things generally are, and are accepted as true until experience requires their rejection, but until experience forces such rejection they qualify for acceptance as items of human knowledge and serve as evidence for other beliefs, nonbasic beliefs. On the question of nonbasic knowledge, or scientific knowledge, he has consistently argued in *Methodological Pragmatism* and elsewhere that while particular scientific theses established by the inductive methods of science may be false (although we must presume them to be true when strongly confirmed), rationality requires us to use such a method because they generally tend to produce more effectively supplementable beliefs about the physical world than any other methods available.

Rescher construes truth in terms of any classical formulation of the correspondence theory of truth satisfying Alfred Tarski (1902–1983) biconditionals, and he argues that the criterion for it is fully warranted, assertible belief. The satisfaction of this criterion in any given case, however, does not entail logically that the proposition is true rather than our best estimate of, or approximation to, truth; and it would be irrational to ask (as skeptics do) for anything more in the pursuit of truth, for nothing more can be had (Rescher 2003).

At no time, then, can we be sure of having the truth, rather than a reliable, but fallible, estimate of how things are, and it is this essentially fallibilistic conclusion that leads to Rescher's antirealistic view that we cannot be sure at any given time whether science actually succeeds in correctly describing an external world, although indeed we have good reason in this fallibilism to suppose that there is an external world. And this same fallibilism leads, with the support of various arguments, to the essential incompletability of our knowledge of the world (Rescher 1978, 1999, 2000a, 2001, 2003). His idealism is consciously not an idealism affirming that all properties are linguistic in nature, but it does emphasize the fact that all systems of knowledge are the products of pervasive and profound human cognitive construction (Rescher 2001).

By way of philosophical methodology Rescher adopts a view he calls philosophical standardism. He thinks, for example, that human knowledge is fundamentally and standardly a matter of justified true belief. Prevalent counterexamples to the classical definition of knowledge as justified true belief are maximally distortive of the fact that philosophical explanations are based on limited generalizations that are subject to revision and we seek what is normally and typically the case rather than what is unexceptionally and necessarily the case (Rescher 1994, 2003). For Rescher, then, traditional philosophy is too much given over to abstract necessities of general principle that do not capture our understanding of the world as it is actually experienced, and the price we pay for his more modest construal of philosophical generalizations is to acknowledge the essential open-endedness of our philosophically relevant concepts.

By way of compensation for this less demanding view of philosophical methodology, Rescher urges that we can resolve a host of philosophical problems, such as the Gettier problem that have lingered too long because of the mistaken and pervasive belief that philosophical generalizations will be adequate only if they do not admit of exception in any context (Rescher 1994, 2003). Rescher argues, then, for a different view of conceptual analysis, a view allowing us to resolve the Gettier problem as well as the problem about the concept of meaningfulness in empiricism. For example, he argues that the classical definition of knowledge in terms of justified true belief has led to a hopeless set of counterexamples and counter-definitions simply because people think mistakenly that

counterexamples refute a concept defined, and this mistaken belief roots in a faulty concept of analysis that is traditional and aprioristic.

Rescher's proposal is that we construe the relationship between knowledge and justified true belief not as a definition, but as a merely standardistic or generalized linkage under which "Standardly, knowledge is justified true belief" is a perfectly acceptable generalization, not only plausible but largely unproblematic. For Rescher, in the context of an epistemological standardism, interpreting such generalizations in a standardistic way does not allow the definition to be annihilated by counterexample. After all, as he says, knowledge is pretty standardly justified true belief. This same approach he applies to the empiricist criterion of meaning.

Although Rescher ascribes a certain primacy to induction and the methods of natural science because they are the products of the evolutionary process, he has not argued that the only legitimately answerable questions are those that admit of answer under the methods of science. He in fact has argued against that view when, among other arguments, he defends metaphysics as that philosophical venture seeking to examine and criticize the presuppositions of natural science, which natural science cannot do without viciously circular reasoning. He also has claimed that such presuppositions find their ultimate justification in the ultimate consequences, formal and material, of accepting them and their capacity to satisfy human needs for practical adaption.

On the question of scientific progress Rescher has aggressively argued in various places that unto eternity science is progressive and revolutionary, meaning thereby that there will never be a time when we would be justified in believing that we had answered all answerable questions about the world. Owing to an inevitable exponential decay in our economic capacity to fund scientific technology, scientific progress will accordingly slow, without stopping, to increasingly infrequent theoretical and factual advances. But it will always be an open-ended and unfinishable affair (Rescher 1978, 2000a, 2000b, 2003).

With regard to scientific realism, Rescher advances a cautious form of scientific instrumentalism without endorsing instrumentalism as a whole on the issue of factual knowledge. For Rescher, commonsense beliefs (those beliefs so obviously true that we cannot even imagine factual conditions under which they would be false) do succeed in standardly describing the physical world because such beliefs are not in any way likely to suffer truth value revision (Rescher 2003). Scientific beliefs, however, have no such property and must, for that reason, be regarded as instrumentally reliable beliefs that we can plausibly presume true when strongly confirmed.

Otherwise, Rescher's fundamental metaphysical view on the question of reality originates in what he calls his *pragmatic idealism*, which he also sees as an antirealist implication of fallibilism and is an idealism only to the extent that it emphasizes the constructive role of cognitive processes in structuring our beliefs about an external and independent world about which we have knowledge in terms of our capacity to estimate the truth in the light of available evidence and in terms of what we can reasonably ascertain as typically and generally the case. But, it is not an idealism denying the existence of an external world. That such a world exists fundamentally roots in the essentially fallibilistic limitations and incompletability of our knowledge of the world, as has been demonstrated time and again in the history of science and elsewhere (Rescher 1992–1994, 2000a, 2001, 2003).

In several cases philosophers have associated Rescher's name with a particular concept or principle broadly discussed, most notably in *Rescher's Law of Logarithmic Returns*, *The Rescher Quantifier*, *Rescher's Effective Average Standard* in the theory of distributive justice, *The Dienes-Rescher Inference Engine* in nonstandard logic, and *The Rescher–Manor Mechanism* in non-monotonic reasoning theory.

See also Epistemology; Metaphysics; Social Epistemology.

Bibliography

PRIMARY WORKS

Methodological Pragmatism: A Systems-Theoretic Approach to the Theory of Knowledge. Oxford: Basil Blackwell, 1977.

Scientific Progress: A Philosophical Essay on the Economics of Research in Natural Science. Pittsburgh, PA: University of Pittsburgh Press, 1978.

Scientific Realism: A Critical Appraisal. Boston; Dordrecht, Netherlands: D. Reidel; Norwell, MA: Klumer Academic Publishers, 1987.

A System of Pragmatic Idealism. Princeton, NJ: Princeton University Press, 1992–1994.

Philosophical Standardism: An Empiricist Approach to Philosophical Methodology. Pittsburgh, PA: University of Pittsburgh Press, 1994.

The Limits of Science. 2nd ed. Pittsburgh, PA: University of Pittsburgh Press, 1999.

Nature and Understanding: The Metaphysics and Method of Science. Oxford: Clarendon Press, 2000a.

Realistic Pragmatism: An Introduction to Pragmatic Philosophy. Albany: State University of New York Press, 2000b.

Cognitive Pragmatism: the Theory of Knowledge in Pragmatic Perspective. Pittsburgh, PA: University of Pittsburgh Press, 2001.

Epistemology: An Introduction to the Theory of Knowledge. Albany: State University of New York Press, 2003.

Robert Almeder (2005)

RESPECT

The ideas that people should be treated with respect and that individuals should respect themselves are important elements of everyday morality and moral philosophy. Some theories treat respect for persons as the basis of morality or the hallmark of a just society, while self-respect is often viewed as a core moral duty or something that social institutions must support. There is disagreement, however, about whether things other than persons, such as animals or the environment, are appropriate objects of respect.

Most generally, respect is acknowledgement of an object as having importance, worth, authority, status, or power. As its Latin root *respicere* (to look back) indicates, to respect something is to pay attention or give consideration to it. As the etymology also suggests, respect is responsive: the object is regarded as due, deserving, or rightly claiming acknowledgement. Respect can be an unmediated emotional response, but it typically involves a conception of certain forms of acknowledgement as appropriate in virtue of some feature of or fact about the object, which is the basis of respect. Respect thus differs from attitudes such as liking, which are based in the agent's interests. Respect also typically involves behaving in ways that show regard for the object or refraining from certain conduct out of respect for it. We can respect rules by obeying them, dangerous things by taking precautions, and authorities by deferring to them; but respect is commonly thought to involve appreciating the value of the object. Valuing respect can be akin to admiration, awe, or honor, but contrasts with valuing modes such as maximizing and using. We can respect things we do not approve of, but regarding something as worthless or irrelevant is incompatible with respecting it.

There are many types of respect. Consider the well-mannered respect children should show parents and teachers, the great respect one might have for accomplished or morally exemplary individuals, the just respect people demand for their rights, the wary respect a prudent hiker has for rugged backcountry, the pro forma respect of standing for the judge entering a courtroom, and the basic respect many believe we owe people simply as people. These can be understood in terms of Stephen Darwall's (1977) now-standard distinction between two fundamentally different kinds of respect: recognition respect and appraisal respect.

Recognition respect is a disposition to take something appropriately into account in deliberations about action. A diversity of things, including laws, rights, hazards, opinions, social institutions and positions, nature, and people can be objects of different forms of recognition respect. What recognition respect involves in various cases depends on the reasons why objects of that sort should be taken into account. Recognition respect is a moral attitude if the object is regarded from a moral point of view, for example, as having moral worth or as morally constraining actions. By contrast, we have appraisal respect (which some call evaluative respect) only for people, either as persons or in some role or activity, or for their qualities or achievements. Like esteem, it is based on a positive assessment of an individual's merits and admits of degree; but whereas any valued feature can be a basis of esteem, appraisal/evaluative respect concerns the moral quality of an individual's character. In addition, some philosophers regard the feeling of reverential respect as a distinct third kind of respect.

Whereas everyday discourse tends to use "respect" in the evaluative sense, as thinking highly of someone, philosophical attention focuses chiefly on moral recognition respect for persons. Individuals can be owed recognition respect in virtue of their social position (for example, as an elder or judge); such respect involves conforming to conventions for appropriate behavior. However, respect for persons commonly means recognition respect that all persons are morally owed solely because they are persons, regardless of social positions or individual qualities.

The moral philosophy of Immanuel Kant contains the most influential discussions of respect for persons. Kant holds that all and only persons, by virtue of their rational autonomy, are "ends in themselves" and have a special, unconditional worth called "dignity." Respect is the only fitting response to dignity; consequently, we have a fundamental and absolutely binding moral obligation to respect persons as ends in themselves. Moreover, all persons are equal in dignity and moral status with other persons, so each has a right to respect from all as well as a duty to respect themselves. Kant expresses this idea in *Groundwork of the Metaphysics of Morals* (1785) in one version of the categorical imperative, which is the supreme principle of morality: "Act so that you treat humanity, whether in your own person or the person of any other, never simply as a means but always at the same time as an end." In *The Metaphysics of Morals* (1797) Kant

explicates specific ethical duties of respect for others and self-respect.

Kant's account prompts numerous questions. Is rational autonomy indeed what gives persons the unconditional claim to moral recognition respect? Some thinkers argue that this is too thin a view of what matters morally about persons. Are all humans owed respect? What about those who lack rationality, such as profoundly mentally disabled individuals or human fetuses and embryos? Must persons always be respected regardless of moral merit, or can recognition respect be forfeited, for example, by evildoing? Some contend that remorseless evildoers warrant no respect; others hold that while they deserve punishment, they must still be respected as persons. What attitudes and conduct express respect or disrespect for persons? Humiliation, coercion, and enslavement are quintessential forms of disrespect; what positive measures (e.g., helping others pursue their ends, listening to their points of view) does respect require? What does respect imply for issues such as assisted suicide, pornography, poverty, and political rights for cultural minorities? Theorists also ask whether respect for persons is the foundation of all other moral duties and rights or simply one important moral consideration among others, and whether non-Kantian ethical approaches such as utilitarianism can accommodate the idea that persons are unconditionally owed respect.

A rich debate concerns whether things other than persons, such as other living things or the natural environment, which are often valued merely as means serving human interests, have a moral status that demands respect. Some thinkers argue that the basis of morally required respect is wider than rationality and can be possessed by nonpersons. Others hold that there are levels of respect such that while persons are owed maximal respect, other things may be due a lower level of respect that nevertheless rules out certain treatment, such as destroying them for trivial reasons. Widespread acknowledgment of duties of respect to nonpersons could entail significant changes in many human activities, such as eating, land and energy use, and biomedical research.

Self-respect, important in its own right, involves due appreciation of one's morally significant worth: worth one has either as a person or in some position or activity (recognition self-respect), or worth earned through the quality of one's character and conduct (evaluative self-respect). Both kinds of self-respect include an engaged understanding of the implications of having worth for directing one's life and interacting with others. Respecting oneself contrasts with, among other things, servility, acquiescence to disrespect, shamelessness, chronic irresponsibility, self-destruction, and self-contempt. Evaluative self-respect is distinguishable from self-esteem. The former involves regarding one's character and conduct as coming up to scratch; it is lost if one comes to regard oneself as morally intolerable. The latter is enhanced or diminished through believing that one has or lacks any highly prized quality.

Self-respect is regarded both as morally required and as essential to the individual's well-being. It is thus strong criticism to say that a person does what no self-respecting person would do or that a social institution undermines people's self-respect. For Kant, individuals have a moral duty to respect their equal dignity as persons and to do nothing that would degrade or disavow it. In *A Theory of Justice* (1971) John Rawls maintains that because the ability of individuals to respect themselves is significantly affected by their social and political circumstances and because self-respect is vital to individual well-being, justice requires that sociopolitical institutions support self-respect. Connections between self-respect and, for example, responsibility, self-identity, forgiveness, prostitution, oppression, and education are also of philosophical interest.

See also Kantian Ethics; Moral Sentiments; Rights.

Bibliography

Buss, Sarah. "Respect for Persons." *Canadian Journal of Philosophy* 29 (4) (1999): 517–550.

Darwall, Stephen L. "Two Kinds of Respect." *Ethics* 88 (1) (1977): 36–49.

Dillon, Robin S., ed. *Dignity, Character, and Self-Respect*. New York: Routledge, 1995.

Hill, Thomas E., Jr. *Respect, Pluralism, and Justice: Kantian Perspectives*. New York: Oxford University Press, 2000.

Hudson, Stephen D. "The Nature of Respect." *Social Theory and Practice* 6 (1) (1980): 69–90.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals* (1785). In *Immanuel Kant: Practical Philosophy*. Translated and edited by Mary Gregor. New York: Cambridge University Press, 1996.

Kant, Immanuel. *The Metaphysics of Morals* (1797). In *Immanuel Kant: Practical Philosophy*. Translated and edited by Mary Gregor. New York: Cambridge University Press, 1996.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Taylor, Paul W. *Respect for Nature: A Theory of Environmental Ethics*. Princeton, NJ: Princeton University Press, 1986.

Robin S. Dillon (2005)

RESPONSE-DEPENDENCE THEORIES

The term *response-dependent* was introduced by Mark Johnston (1989) for concepts, such as *red*, that support an a priori biconditional on roughly the following lines: "X is red if, and only if, X is such that it would look red under normal conditions." Any concept of the intended kind will apply to something just in case the object has a property in virtue of which it would elicit a relevant response, on a par with the appearance of redness, under relevantly favorable conditions; it will be akin to the concept of a secondary quality, traditionally conceived. The response to be elicited will involve a cognitive impression, so that the object looks, seems, or presents itself in a certain manner. And the conditions under which that response is guaranteed will have to be capable of independent specification; they cannot be defined just as whatever conditions will provide the guarantee.

Response-dependent concepts in this sense are meant to contrast with response-independent concepts whose application to an object depends solely on the nature of that thing in itself, not on the cognitive impression that the object makes on human beings. As Crispin Wright (1992) has emphasized in ongoing reflections around the theme, there must be a sense in which the object is of the conceptualized kind because it elicits that response, and not (or not just) the other way around; there must be a sense in which an object is red because it normally looks red.

The interest of the notion of response-dependence lies in the prospect of illuminating the character of a variety of concepts: for example, concepts of an evaluative, affective, or aesthetic kind; concepts associated with practices such as praise and blame or intervening causally in the world; concepts that are anthropocentric in any such manner; or perhaps all concepts that are mastered ostensively, without reliance on prior definition.

There are two very different theories of response-dependence in the literature. The biconditional associated with response-dependence, so all sides assume, does not hold because people's relevant cognitive impressions never miss or misrepresent anything. So what makes certain concepts response-dependent, assuming that some concepts are indeed of this kind? What underpins the truth of the biconditional that governs them? The two theories diverge on that question.

One theory, explored by Johnston himself, would say that certain concepts are response-dependent because the properties they designate are dispositions in things to evoke the relevant responses. Under this account we use a term like *red* to apply to those things that are such as to look red in suitable conditions; we think of the property of redness as the higher-order property of things that have a lower-order property, maybe this, maybe that, which makes them look red in suitable conditions. According to this theory, the concepts are response-dependent because the properties are defined by reference to responses; the a priori biconditionals hold, because they reflect the character of the properties conceptualized.

This theory has the disadvantage that, as Johnston (1993, 1998) himself has argued, few of our concepts are response-dispositional in this sense. With concepts such as *red*, we want to say that something looks red because it is red, where this is a causal explanation. But it is not clear that that claim remains available if redness is construed as a disposition; looking red will be a manifestation of the disposition, not a contingent effect. The issue has been a focus of controversy (Menzines and Petit 1993, Miller 2001).

The alternative theory of response-dependence would avoid this difficulty (Jackson and Petit 2002, Petit 2002). While allowing that there may be response-dispositional concepts, it says that other concepts may be governed by an a priori biconditional, too (or, being partly response-dependent, by at least an "only if" conditional). That will not be because they are paired with anthropocentric dispositions, but because the explanation of why they are paired with their particular, response-independent referents is that those properties have certain anthropocentric effects. On this account *red* may refer, not to the disposition to look red, but to a perfectly physical property, such as a certain profile of surface spectral reflectances. The reason why it will refer to that property is that it is the one that elicits the appearances on which speakers rely in learning to use the term. And so a connection will remain in place between the presence of the property and the looks-red response. Response-dependence will become salient, not at the level of semantics where we pair off terms with items in the world but, to invoke a distinction made by Robert Stalnaker (2004), at the level of meta-semantics where we try to explain the pairings that obtain between words and world.

Suppose that people generally rely on appearances in using the term *red*. Suppose that they intend to refer to a common, objective property in using the term; they are not content to go their idiosyncratic ways. And suppose

that because of that intention they seek to coordinate their usage, discounting some of the appearances of redness on which they divide. If this enterprise of coordination is to have objective significance, then there must be an objective reason for speakers to discount some appearances and not others; equivalently, there must be an objective reason to treat certain factors as perturbing or limiting influences on appearances. Why should speakers indict some influences as perturbing or limiting, then, but not others? According to this approach, it will be right to indict factors such that by privileging situations of usage where they are absent—by treating those conditions as favorable—speakers can optimally satisfy their joint intention to pick out the same objective property in things (Pettit 2002); the associated practice will do best in helping them to triangulate on a common, presumptively objective feature.

This approach will make it a priori, for any property such as redness that is available to be named in our language, that X counts as red if, and only if, it is such that it would look red under favorable conditions, with *favorable* defined by reference to the practice of discounting. The basis of the response-dependence will now lie, not in the nature of the property, but in the requirements that must be fulfilled for the property to deserve the name of *red*; that is why the biconditional is restricted to properties available to be named.

This theory of response-dependence allows us to say that while things may be conceptualized as red because of the associated appearances, still their redness is causally responsible for such appearances. It can mark out certain concepts as special, on a par with the concept of redness, particularly if the concepts are ineradicably response-dependent. And yet it can allow us at the same time to be realists about redness and similar properties, even holding that a term like *red* refers rigidly to an actual-world property (Haukioja 2001). If one wants to espouse response-dependence without too deep a revision of common sense, this is the way to go. If one aspires to be revisionary, the other theory of response-dependence is the better option.

See also Ethical Naturalism; Metaethics; Philosophy of Language; Primary and Secondary Qualities; Semantics.

Bibliography

Haukioja, J. "The Modal Status of Basic Equations." *Philosophical Studies* 104 (2001): 115–122.

Jackson, F., and P. Petit. "Response-Dependence without Tears." *Philosophical Issues* 12 (2002): 97–117.

Johnston, M. "Are Manifest Qualities Response-Dependent?" *Monist* 81 (1998): 3–43.

Johnston, M. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* 63 (1989): 139–174.

Johnston, M. "Objectivity Refigured: Pragmatism with Verificationism." In *Reality, Representation, and Projection*, edited by J. Haldane and C. Wright. Oxford: Oxford University Press, 1993.

Menzies, P., and P. Petit. "Found: The Missing Explanation." *Analysis* 53 (1993): 100–109.

Miller, A. "The Missing Explanation Argument Revisited." *Analysis* 61 (2001): 76–86.

Petit, P. *Rules, Reasons, and Norms: Selected Essays*. Oxford: Oxford University Press, 2002.

Stalnaker, R. "Conceptual Truth and Metaphysical Necessity." *Philosophical Studies* 18 (2004).

Wright, C. *Truth and Objectivity*. Cambridge, MA: Harvard University Press, 1992.

Philip Pettit (2005)

RESPONSIBILITY, MORAL AND LEGAL

The term *responsibility* or one of its variants figures in moral discussion in many different ways. Philosophers have traditionally been especially interested in the concept of moral or personal responsibility. It is with the problems connected with this notion that the following discussion is primarily concerned.

JUDGMENTS OF PERSONAL RESPONSIBILITY

F. H. Bradley once claimed that "for practical purposes we need make no distinction between responsibility and liability to punishment." Although it is true that discussions of responsibility have often turned quickly to discussions of blameworthiness and liability to punishment, there is little justification for Bradley's claim. For responsibility is equally relevant to many other forms of social treatment—among others, praise, reward (including special honors such as honorary degrees or titles), legal punishment, legal liability. And, of course, the topic is intimately related to the theological issue of salvation, the allocation of divine rewards and punishments.

Judgments of personal responsibility pertain to this range of practices in a very special way. Unless a person is judged personally responsible for some act or outcome, he would not normally be thought to deserve blame, praise, reward, punishment, and so on. Personal responsibility is generally regarded as a necessary condition of the justice of a person's receiving what he deserves. Yet

Bradley's error is repeated in many contemporary discussions of "freedom and responsibility" that start with some unilluminating remarks about "responsibility," then move swiftly to examination of blame or punishment. Discussion of responsibility is theoretically fundamental, not ancillary, to accounts of such practices.

MEANING OF "MORAL RESPONSIBILITY"

Persons are normally judged morally responsible for their actions. But they may be judged responsible for almost anything—events, processes, their own psychological characteristics. Thus, a person may be judged morally responsible for his firm's loss of a contract, the Napoleonic wars, his bad temper, a technique for maintaining the fertility of land, or his friend's divorce. Under what conditions is a person responsible for one of his acts or for some other occurrence? If we can state the necessary and sufficient conditions for judgments of moral responsibility, we shall, in the process, be assigning a sense to the expression. "Moral responsibility," like so many other terms of moral discourse, is inevitably defined persuasively, for one is bound to be influenced in defining it by convictions about the requirements for deserved blame, praise, and so on. That is, one is bound to be influenced by convictions, explicit or implicit, about the requirements of justice in such matters.

Most persons, however, would accept the following form of definition, although those with different moral outlooks would complete it differently: A person is regarded as morally responsible for some act or occurrence *x* if and only if he is believed (1) to have done *x,* or to have brought *x* about; and (2) to have done it or brought it about freely. The completion of this formulation depends on what is meant by a human action; what would count as bringing some outcome about; and, above all, in what sense the terms *free, freely,* or *freedom* are employed. All these conceptions are problematic in ways that lead to very different theories of responsibility. Philosophers have too often supposed that the concept of "freedom" essential to moral responsibility can be fixed independently of what it is to be responsible, and that only after the meaning of *freedom* is specified can we determine whether, and under what conditions, a person is responsible. But in fact what a person means by *free, freely,* or *freedom* will reflect his moral convictions, and especially his views about justice, in the same way and for the same reasons that his conception of "moral responsibility" will reflect these views. As Harald Ofstad put it, "Ethical systems may determine the sense of 'freedom' we select as relevant" (*Freedom of Decision,* p. 279). One need add only that they not only may, they do.

FREEDOM AND MORAL RESPONSIBILITY

In his *Nicomachean Ethics,* Aristotle tried to analyze the concept of "voluntary action." Nowhere in his discussion did he clearly take account of the problems that arise if all our decisions and actions are determined by circumstances beyond our control. But he did claim that actions are compulsory "when the cause is in the external circumstances and the agent contributes nothing." It is difficult to say whether, in this and other passages, Aristotle intended to claim that the fact that the cause of action is external implies that the agent contributes nothing and is therefore not free in the sense relevant to responsibility. But from the beginning of the Christian era, the view that if decisions and actions are so determined, then persons are not free in the relevant sense, has been forcefully advanced and denied by countless numbers of theologians and philosophers.

The earliest form of the controversy arose in the context of Christian doctrine. In particular the fourth-century Christian theologian Pelagius argued that the doctrines of original sin and grace, and of divine omnipotence and foreknowledge, led to morally repugnant conclusions, primarily the conclusion that although a person's tendencies, decisions, and actions are in no way the fault of the agent, he is nevertheless morally culpable for his actions and, in consequence, justly suffers the torments of hell. If these doctrines are true, Pelagius argued, God is not just. But as God is certainly just, these doctrines must be false. Pelagius insisted that man is possessed of free will in that he has the power of "contrary choice." This power makes it possible for men to sin. In the fifth century St. Augustine countered Pelagius's attack on orthodox doctrine with the claim that though God knows and wills all, he grants to each person who has faith freedom of choice. Though God knows what a man will do, he wills only hypothetical claims, of the form "*If* this man sins, then he *shall* be punished." Divine decrees of this kind are consistent with freedom of the will. But what about the possession of faith—is this in a man's power? St. Augustine insisted that it was; for to have faith is to believe, and "belief is simply consenting to the truth of what is said, and consent is necessarily an act of will. It follows that faith must be in our power."

Although the terms are often different, the issues generated by this exchange persist. The doctrine of scientific determinism, and not the doctrine of divine

omnipotence, is today more commonly thought to pose the chief difficulties. Scientific determinists maintain that external conditions specified in scientific laws are sufficient to produce each human choice and action. But the nature of free choice, the nature of human power and ability, the relevance of necessity to freedom, the role of choice and deliberation, the very possibility of human choice, and many other issues suggested by or actually crystallized in the debate between Pelagius and Augustine are still vigorously debated.

DILEMMA OF MORAL RESPONSIBILITY

Efforts to solve the problem of freedom of the will are conveniently considered against the background of the following dilemma.

> If determinism is true, then all events, including any person's decisions and actions, are fully determined by circumstances that are ultimately beyond that person's control. If this is so, then that person could not have decided or acted differently. Hence the person was not free.
>
> If determinism is false, then there are at least some events that are not fully determined by antecedent circumstances. To the extent that human decisions and actions are among those events which are not fully determined, those decisions and actions occur by pure chance. But what occurs by pure chance is not within a person's control. Therefore, to the extent that decision and action are not determined, the person is unfree.
>
> But determinism is either true or false. Hence a person is never free with respect to decisions, actions, or the results of actions.
>
> But, it is claimed, a person is morally responsible for an action or occurrence only if he is free in that respect.
>
> Therefore, no one is ever morally responsible for any decision, action, or outcome.

FREEDOM AS THE LACK OF CONSTRAINT. Some philosophers have argued that determinism does not imply that a person's actions are beyond the person's control. They argue that there is a perfectly clear, ordinary sense to "being able" or "being free" to do something that is compatible with determinism. As Jonathan Edwards, the great American theologian, put it in *Freedom of the Will,* the most sustained, penetrating defense of this position: "Let the person come by his volition or choice how he will, yet, if he is able, and there is nothing in the way to hinder his pursuing and executing his will, the man is fully and perfectly free, according to the primary and common notion of freedom" (Paul Ramsey, ed., 1957, p. 164). The central assumption of Edwards's argument is that the ordinary sense of statements like "Eisenhower could have ordered his troops to take Berlin before the Russians arrived" and "Kennedy was able to call off the invasion of Cuba, but he decided not to do so" is such that these statements are perfectly consistent with determinism. In David Hume's terms, there is an important distinction between an action being caused or determined by antecedent circumstances, and its being constrained or compelled or coerced by antecedent circumstances. Only when an action that is determined is also in some way constrained or compelled is the actor not morally responsible for that act.

Other philosophers have found this position unacceptable for a variety of reasons. Some have argued that the ordinary use of such expressions as "was free to," "could have," and "was able to" involves more than lack of constraint. They argue that careful analysis reveals that determinism is indeed inconsistent with statements of the form "*X* could have done such-and-such." Others have argued that freedom of the will depends upon freedom of decision, not freedom of action; and that if decisions are determined then it surely cannot be the case that one could have decided other than he did. Still others have claimed that there is no reason to accept the authority of common sense or ordinary language in these matters; that it is the philosopher's job to subject our common opinions to the test of careful, reasoned scrutiny, in the manner of Socrates.

MORAL JUDGMENTS AND RESPONSIBILITY. Among those who reject common sense as reflected in ordinary language as a basis for philosophical opinion are those who nevertheless endorse the distinction between constrained and nonconstrained causally determined actions but defend it on explicitly moral grounds. Thus, certain philosophers have argued that the aim of holding someone morally responsible should be to influence future behavior in desirable ways—that, indeed, moral responsibility consists in the ability to be influenced by moral judgments. If a judgment of responsibility will not affect behavior in desirable ways, then there is no moral point in holding that person responsible. On this view, most customary excuses will still be acceptable. For it will not, in general, be possible to exert beneficial influence on a person if he did what he did either unintentionally or because no other course of action was possible.

One difficulty with this position is that we are, after all, concerned with persons other than the one whose responsibility is being judged. This concern can be accommodated by taking into account all of the consequences of a given judgment of moral responsibility, and determining whether the consequences are good, or best on the whole. But such a position seems to imply that a person believed to be innocent of an offense might be held morally responsible and be blamed or convicted on the general grounds that it would be socially beneficial to do so. And this seems to conflict with deeply held convictions about the requirements of justice in our commerce with other human beings. Considerations of this sort led Immanuel Kant to warn against the "serpent-windings" of utilitarianism. Utilitarianism seems to many to imply just such an unqualified appeal to social consequences.

Many thinkers feel that a related consideration has great importance in assigning moral responsibility. They have argued that the claims of justice are satisfied if we justify the rules according to which a person is judged to be morally responsible and blameworthy on the basis of the principle that social utility ought to be maximized, but then apply these rules to particular cases in a way that precludes any further appeal to this principle of utility. In this way, the claims of justice may be satisfied and the problem of freedom bypassed. This view, usually called "rule utilitarianism," has been vigorously discussed by many contemporary moral philosophers. One criticism of it is that the restriction placed on the relevance of the principle of utility cannot itself be justified on utilitarian grounds, and that therefore the principles of justice cannot be explained or defended on a purely utilitarian basis.

FREEDOM AS SELF-DETERMINATION. Another gambit directed against the first argument of the dilemma rests on the distinction between self-determined action and action determined by circumstances external to the agent. Thus, Bradley argued that it is the self that may determine action and that, to the extent that this is so, the person is morally responsible for his actions. He argued that self-determinism does not imply that actions are predictable; actions are, in fact, not predictable, provided that the determining conditions are not entirely "materialistic" because they include "spiritual" or, perhaps, mental causes. The difference between the views of freedom as self-determination and freedom as absence of external constraint is that, although the latter allows that nonconstraining circumstances may be bodily causes external to the agent, the former view rules out this possibility. However, even if one could formulate a clear notion of the self that determines action, there seems to be no reason to suppose that that self, or its determining characteristics, are themselves not determined by circumstances external to the agent. And if this is so, then the action would seem to be determined by circumstances that are ultimately beyond the person's control. In reply to this objection it has been suggested that determinism does not imply that determinants occur before that which is determined—and that in the case of human decisions and action, the causal determinants occur simultaneously with the decision that in turn accounts for the action. Thus, the action is determined by a decision that is not itself the result of circumstances beyond the person's control. For, as the determinants are concurrent conditions, in principle they can be affected by prior action. But it is not clear that this view rests on anything more than an ad hoc assumption needed to establish the possibility of self-determinism. There is, moreover, much psychological evidence for the view that if one's decisions and actions are determined, then the determinants are circumstances temporally prior to them and external to the agent who decides and acts.

INDETERMINISM. Philosophers have been equally fertile in rebutting the second argument of the dilemma. Those who believe that only if determinism is false can a person be morally responsible, and thereby are impelled to attack this second argument, are usually called "libertarians" because they believe that the will itself is free in the sense of being undetermined. Libertarians claim that the fact that a decision or action is not fully determined by antecedent conditions does not imply that it occurred by "chance" or "accident" in a way that confers exemption from moral responsibility. But this argument does not refute the claim that an undetermined event is a matter of chance in a way that implies that it occurred by chance or by accident, in the sense of those terms that *is* relevant to moral responsibility. For example, the difference between knocking a flowerpot off a shelf as the result of the fully determined but accidental motion of someone's arm or as a result of an undetermined motion of that arm seems irrelevant to a judgment of responsibility. The two events seem equally to void the responsibility of the agent. Both occurrences seem *accidental* in the relevant sense.

Others—J. D. Mabbott, for instance—claim that the first argument of the dilemma is sound, but it is inconceivable that moral responsibility is inapplicable to the human situation and, therefore, the second argument of the dilemma must be unsound. However, this is hardly an argument; it is rather a dogmatic affirmation of the point at issue. Still another argument is that human beings are so constituted that they necessarily hold others responsi-

ble for their actions and necessarily employ concepts in doing so that presuppose indeterminism. This conclusion would seem to rest on dubious psychological assumptions. In any event, if one could develop an account of moral responsibility that does not presuppose that determinism is false, which is morally defensible, and acceptance of which is psychologically possible, this view would be refuted.

HARD DETERMINISM. There have been other ingenious efforts to escape the toils of the dilemma. But it has also been argued that persons are indeed never morally responsible. According to this view, which has been called "hard determinism," determinism is true and the first argument of the dilemma is sound. Hard determinists allow that blame and punishment may be useful, but they deny that they are ever morally deserved. Persons who blame or punish should do so only when engaged in moral education or social engineering; and blame and punishment have no special moral significance when they are justified as effective aids in these tasks. As the blame is not moral blame, there is no need to establish that it is deserved in virtue of the fact that the person is morally responsible. As John Hospers put it: "When we view other people's frailties and shortcomings in the light of this perspective, we shall no longer say, 'He deserves what he's getting.' Instead, we shall say, 'There, but for the grace of God (and a favorable early environment) go I'" (*Human Conduct,* p. 521).

Hard determinists forget, however, that the claim that someone deserves what he is getting is not necessarily an expression of moral indignation. It may instead be an expression of the belief that all of the requirements of justice have been satisfied. If it is defensible to suppose that "freedom," used in some sense consistent with determinism, is a requirement of justice, then hard determinism is unacceptable.

The general defect of the dilemma is that it presupposes that the relevant sense of "freedom" can be specified independently of a specific moral outlook, and particularly of a conception of justice. This defect reverses the proper order of moral reflection. The sense in which one can be said to have "acted freely," and therefore to be morally or personally responsible and to deserve blame or praise or punishment or reward, should be specified in the light of one's moral outlook—not independently of it.

LEGAL RESPONSIBILITY AND PUNISHMENT

Many philosophers regard the legal context as paradigmatic for the discussion of moral responsibility. It seems clear that the unfortunate tendency to identify moral responsibility with blame and punishment derives partly from this fact. Nevertheless, the assessment of legal responsibility is so closely related to the assessment of moral responsibility, and legal experts have given such sustained and imaginative attention to the task of articulating criteria that are applicable to complex cases, that a careful study of the relevant aspects of the law will certainly assist the development of an adequate account of moral responsibility. Though problems pertaining to responsibility occur in all branches of the law, criminal law has received the most attention; the topics most frequently discussed in this connection are mens rea and criminal insanity.

MENS REA. The doctrine of mens rea requires a certain "mental element" to have been present when the offense was committed. This mental element is usually, but misleadingly, described as "guilty mind." The characterization is misleading, first, because it is generally supposed that the offender need not be aware that he is committing an offense ("ignorance of the law is no excuse"); and, second, because many advocates of mens rea do not even require that the offender be morally culpable. On this second point there is, in fact, considerable disagreement. Some argue that unless an offender is morally blameworthy for his offense, he does not deserve to be convicted. Others insist on the distinction between moral responsibility and moral blameworthiness, arguing that a person may be morally responsible and may deserve to be convicted and punished for a crime even though his actions were not blameworthy. Broadly speaking, then, those who subscribe to the doctrine of mens rea believe at least that only persons who are morally responsible for their offense deserve conviction and punishment.

Discussions of mens rea usually take for granted the possibility of resolving the philosophical perplexities described above. Certain assumptions, generally unexamined, are made, and the work of articulating criteria appropriate to the criminal law goes forward. Those who accept the doctrine of mens rea in any of its forms believe that the requirement is satisfied if the offender has committed his offense intentionally. Some also claim that unintentional actions that are performed recklessly or negligently involve the necessary mental element. In general, the person who commits an offense is thought to

have satisfied the doctrine of mens rea if he knew what he was doing at the time or if he would have known what he was doing had he proceeded with reasonable care and deliberation. The extent to which an offender is able to or actually does exercise deliberate control over his actions and their results seems to be central to the way in which moral responsibility as a condition of deserved conviction and punishment is incorporated into the criminal law. This point is, however, more general than the doctrine of mens rea itself—it being possible for someone to have acted intentionally while, by reason of mental defect, not possessing deliberate control over his actions. Before going on to this point, two criticisms of the doctrine of mens rea should be considered.

Objections to mens rea. There are those who argue that, at least for certain criminal offenses, the requirement of mens rea ought to be abandoned and that strict liability ought to prevail. That is, for certain offenses it does not matter that the act was unintentional and it does not matter that reasonable care was taken. There are various arguments for strict liability, but, in general, the case for it is specific to the offense.

Though the agent's state of mind would seem not to enter into legal deliberations where strict liability prevails, this is not quite so. For example, it has been held that a bank director is strictly liable for borrowing money in excessive amounts from his own bank. In *State v. Lindberg*, 258 U.S. 250 (1922), the director pleaded that he had been assured that the money borrowed did not come from his own bank. Though the director did not borrow the money from his bank intentionally, the act of borrowing was itself intentional. A person cannot be said to have borrowed money that he accepted as a gift; his own intentions as well as the intentions of the donor are controlling. Though borrowing does, therefore, require a certain state of mind, the absence of the "mental element" involved in intentionally borrowing from one's own bank would be sufficient to discharge a person from moral responsibility. Insofar as the doctrine of mens rea is designed to satisfy the requirement that only a person who is morally responsible for some act or its result deserves to be held legally responsible and punished, strict liability conflicts with it.

Criteria of mens rea. The second criticism does not so much repudiate the requirement of mens rea in establishing responsibility as it criticizes the effort to develop criteria for mens rea. H. L. A. Hart argued that the practical meaning of mens rea is given in what is allowed as excuse or mitigation within the law. In order to determine whether mens rea is established, Hart argues, "it is necessary to refer back to the various defenses; and then these general words (like 'mistake,' 'accident,' and so on) assume merely the status of convenient but sometimes misleading summaries expressing the absence of all the various conditions referring to the agents' knowledge or will which eliminate or reduce responsibility." In other words the general "rules" summarize accepted excuses, and there just are no general principles in terms of which we can account for the acceptance of specific excuses. Hart then generalizes his discussion of mens rea to pertain equally to the assessment of responsibility in nonlegal contexts.

This thesis encounters many difficulties. For one thing Hart neglects to distinguish adequately between exemption from responsibility and exemption from blame or legal responsibility. Thus, if a person defends himself against moral criticism of his having hit someone else by claiming that he was acting in self-defense, he is in effect accepting responsibility but rejecting blame on the grounds that he was justified in what he did. Second, if proposed as a purely descriptive thesis about our actual use of the language of "excuses," Hart's position begs the prescriptive claim that a general rationale of excuse and mitigation ought to be given—that otherwise the acceptance of a certain excuse is morally arbitrary. Those who defend mens rea try to meet this obligation by focusing on the element of awareness of what we are doing when we choose and act. Indeed, Hart becomes his own best critic when, in a later essay, he argues that the main rationale for excuse and mitigation within law is respect for "the claims of the individual as such, or at least as a *choosing being.*"

CRIMINAL INSANITY. A person might intend to kill a particular person after careful deliberation, and do so; and this would be sufficient to satisfy mens rea. But if the offender suffered from extravagant delusions of having been persecuted by the person killed, he would normally be thought to be entitled to exemption from criminal liability on grounds of insanity.

The criterion of legal insanity generally adopted within Anglo American law is the M'Naghten Rule. This rule was formulated by the judges of England in 1843 in response to the public outcry that resulted when Daniel M'Naghten was acquitted, on grounds of criminal insanity, of murdering Sir Robert Peel's private secretary. M'Naghten had mistaken the secretary for Peel and had killed that unfortunate man while suffering from persecutory delusions about Peel's intentions toward him. The judges attempted to provide a morally sound, legally workable criterion for determining whether a person was

entitled to acquittal on grounds of criminal insanity. They affirmed that:

> to establish a defence on the ground of insanity it must be clearly proved that, at the time of committing the act, the party accused was laboring under such a defect of *reason,* from disease of the mind, as not to know the *nature and quality of the act he was doing,* or, if he did know it, that he did not *know he was doing what was wrong.* … The question has generally been, whether the accused at the time of doing the act *knew the difference between right and wrong.* [Italics added.]

The rule has been the object of vigorous attack and defense ever since its formulation. One type of criticism roughly follows the line of argument expressed by various parts of the dilemma formulated earlier. For example, Barbara Wootton, arguing from a determinist position, claimed that no acceptable criterion of criminal insanity can be formulated; that efforts to formulate an adequate criterion of mental defect, and, in the final analysis, of responsibility itself, shatter on the rock of the first argument of the dilemma. Consequently, all efforts to assess moral responsibility should be abandoned within the criminal law. The law should be concerned solely with treating the offender. It is clear that this "reform theory" approach to the criminal law would sweep away not only the insanity plea, but mens rea as well. Thomas Szasz, by contrast, argued that there is no such thing as a mental illness, that the insanity plea is never a valid excuse, and that, therefore, it ought to be abandoned. This argument leads to the same conclusion on policy as that reached by the reform theorists with respect to the insanity plea, but leaves mens rea intact. Szasz is not skeptical of moral responsibility as such. Others, like David Bazelon (in his Isaac Ray Award Lecture, "Equal Justice for the Unequal"), criticize the M'Naghten Rule as being too narrow—as not embracing all those defects of mind that entitle an offender to exemption on grounds of not having been morally responsible for his offense. It seems clear that many of the issues generated by this debate, as well as those that concern the doctrine of mens rea, await an adequate philosophical theory of moral responsibility.

AN APPROACH TO A THEORY OF MORAL RESPONSIBILITY

An adequate theory of moral responsibility cannot identify moral responsibility with liability to blame or punishment. Moreover, any such theory must explicitly recognize what is, in any event, generally the case: that the meaning assigned to the key concepts in the theory, particularly "freedom," reflects the moral outlook of its author.

The second point is of particular importance. Suppose one reflectively endorses a conception of justice according to which a person deserves blame or praise, reward or punishment, and so on, only if that person's decisions or actions are not determined. Then one should define "freedom" in such a way that "*P* decided (acted) freely" implies "*P*'s decision (action) was not determined." Correspondingly, suppose one endorses a conception of justice according to which a person deserves blame, and so on, only if his decisions or actions have some property that may or may not be causally determined by circumstances beyond his control. Then "freedom" ought to be defined in such a way that the meaning of "*P* decided (acted) freely" is consistent with determinism. It is our practical aims and interests that should govern the shape of our language, and not unreflected-upon linguistic habit that should govern the shape of our moral outlook.

Thus, a theory of justice is the essential foundation for a theory of moral responsibility. In this connection it should be remembered that just acts are not always right. (Would it be right to refrain from punishing an innocent person if the consequence was the destruction of human civilization?) Moreover, acts of blame, praise, reward, and punishment that are not just may sometimes be right. (One may be justified in blaming or praising an infant in order to influence his future behavior, but there would be no justice in it.)

See also Action; Aristotle; Augustine, St.; Bradley, Francis Herbert; Consequentialism; Determinism and Freedom; Edwards, Jonathan; Hart, Herbert Lionel Adolphus; Justice; Kant, Immanuel; Pelagius and Pelagianism; Philosophy of Law, History of; Punishment; Socrates; Utilitarianism.

Bibliography

Some of the more interesting classical and contemporary discussions (in a vast literature) are cited below.

GENERAL

For broad treatments of responsibility, see Harald Ofstad's *Freedom of Decision* (Oslo, 1960) and Austin Farrar's *The Freedom of the Will* (London: A. and C. Black, 1958). A number of important papers can be found in the anthologies *Free Will,* edited by Sidney Morgenbesser and James Walsh (Englewood Cliffs, NJ: Prentice-Hall, 1962); *Determinism and Freedom,* edited by Sidney Hook (New York: New York University Press, 1958); *Freedom and the Will,* edited by D. F. Pears (New York: St. Martin's Press,

1963); and *Freedom and Responsibility,* edited by Herbert Morris (Stanford, CA: Stanford University Press, 1961). The Morris collection contains many papers particularly relevant to legal responsibility and has an extensive bibliography.

Historically, the most important work is Aristotle's *Nicomachean Ethics.*

DETERMINIST AND LIBERTARIAN DEFENSES

The best traditional defenses of the view that moral responsibility and determinism are compatible are Jonathan Edwards's *Freedom of the Will* (1754), edited by Paul Ramsey (New Haven, CT: Yale University Press, 1957), and David Hume's *An Enquiry concerning Human Understanding,* Ch. 8. Important contemporary statements of the same position have been made in C. L. Stevenson's *Ethics and Language* (New Haven, CT: Yale University Press, 1944), Ch. 14; Moritz Schlick's *Problems of Ethics* (Englewood Cliffs, NJ: Prentice-Hall, 1939), Ch. 7; and P. H. Nowell-Smith's *Ethics* (New York: Philosophical Library, 1957), Chs. 19–21.

The most influential recent defenses of the "libertarian" position are C. A. Campbell's "Is 'Free-Will' a Pseudo-Problem?," in *Mind* 60 (1951): 441–465, and Stuart Hampshire's *Thought and Action* (London: Chatto and Windus, 1959). See also J. D. Mabbott's "Free Will and Punishment," in *Contemporary British Philosophy,* edited by H. D. Lewis, 3rd series (London, 1956).

HARD DETERMINISM

For hard determinism, see Holbach's *Système de la nature* (London, 1770) and John Hospers, "Free Will and Psychoanalysis," in *A Modern Introduction to Philosophy,* edited by Paul Edwards and Arthur Pap. 2nd ed. (New York: Free Press, 1965), pp. 75–85. This volume contains an extensive bibliography of material bearing on the topic of the present entry on pp. 99–108.

ETHICAL THEORY

Some recent texts that place the problem of moral responsibility in the general context of an ethical theory are Richard Brandt, *Ethical Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1959); A. C. Ewing, *Ethics* (London: English Universities Press, 1953); W. Frankena, *Ethics* (Englewood Cliffs, NJ: Prentice-Hall, 1963); and John Hospers, *Human Conduct* (New York: Harcourt Brace, 1961).

PROBLEMS IN THEORY-MAKING

Of the many technical articles dealing with various conceptual problems involved in the development of an adequate theory of moral responsibility, the following are of some interest.

On the difficulties of analyzing such terms as *can* and *ability,* see John Austin, "Ifs and Cans," in *Proceedings of the British Academy* 42 (1956): 109–132; Richard Taylor, "I Can," in *Philosophical Review* 59 (1960): 78–89, and A. S. Kaufman, "Ability," *Journal of Philosophy* 60 (September 12, 1963): 537–551.

On the difficulties of analyzing such terms as *intention* and *decision,* see Carl Ginet, "Can the Will Be Caused?," in *Philosophical Review* (1962): 49–52; Stuart Hampshire and H. L. A. Hart, "Decision, Intention, and Causality," in *Mind* 67 (1958): 1–12; G. E. M. Anscombe, *Intention* (Oxford: Blackwell, 1957), and A. S. Kaufman, "Practical Decision," in *Mind* 75 (1966): 25–44.

QUESTIONS IN CRIMINAL LAW

For works especially pertinent to the relevance of moral responsibility to the criminal law, see, in addition to the Morris collection (*Freedom and Responsibility*): J. D. J. Edwards, *Mens Rea in Statutory Offenses* (London, 1955); Joel Feinberg, "Problematic Responsibility in Law and Morals," *Philosophical Review* 71 (1962): 340–351; and G. L. Williams, *Criminal Law: The General Part* (London: Stevens, 1953), pp. 28–45, 77–81. See also Henry Hart's "The Aims of the Criminal Law," in *Law and Contemporary Problems* 23 (1958): 405–441, which presents a general defense of the claim that moral blameworthiness is an essential condition of criminal liability. For a contrary view, see H. L. A. Hart's "Legal Responsibility and Excuses," in Hook's collection, *Determinism and Freedom*; this article also in part constitutes an amendment to his own earlier essay, "The Ascription of Responsibility and Rights," *PAS* 59 (1949): 171–194, where he argues against the possibility of a general rationale of excuse and mitigation.

On the topic of strict liability, see R. A. Wasserstrom's "Strict Liability in the Criminal Law," *Stanford Law Review* 12 (1960): 730–745.

The literature dealing with criminal insanity is vast; see Barbara Wootton's *Social Science and Social Pathology* (London: Allen and Unwin, 1959), Ch. 8; Thomas Szasz's *Law, Liberty, and Psychiatry* (New York: Macmillan, 1963); David Bazelon's "Equal Justice for the Unequal" (The Isaac Ray Award Lecture). Unfortunately this last work is available only in mimeographed form.

Among scientific writings that have a special relevance are M. Gluckman, *The Judicial Process among the Barotse of Northern Rhodesia* (Manchester, U.K.: Manchester University Press, 1955), and B. F. Skinner, *Science and Human Behavior* (New York: Macmillan, 1953), Chs. 12 and 22.

Arnold S. Kaufman (1967)

REVELATION

The notion of "revelation" is central to three of the major world religions: Judaism, Christianity, and Islam. Through Christianity in particular it has long been an important element in the religious thought of the West, and the present entry will treat it in this context, especially that of Christian theology.

During the twentieth century, but beginning in the nineteenth century, many—especially Protestant—theologians radically revised their conception of revelation. The view that was virtually axiomatic for all schools of thought in the mid-nineteenth century and that still remains the majority position (for it continues both in Roman Catholicism and in sections of conservative Protestantism) may be called the prepositional view of revelation.

THE PROPOSITIONAL CONCEPT

In the prepositional view, that which is revealed is a body of religious truths capable of being expressed in propositions. Because a knowledge of these truths is necessary for man's salvation, God has supernaturally made them known. Accordingly, in the words of the *Catholic Encyclopedia,* "Revelation may be defined as the communication of some truth by God to a rational creature through means which are beyond the ordinary course of nature" (Vol. XIII, p. 1).

The fuller significance of this prepositional understanding of revelation appears when we view it in relation to three other basic theological categories with which it is closely connected. A particular conception of the nature of revelation involves a particular conception of the nature of faith, as man's response to revelation; of the Bible and its inspiration, as a medium of revelation; and of the character of theological thinking, as thought that proceeds on the basis of revelation.

When revelation is conceived as the divine disclosure of religious truths, faith is necessarily understood as the obedient believing of these truths. Thus faith was defined by the First Vatican Council (1870) as a supernatural virtue whereby "with the inspiration and help of God's grace, we believe that what he has revealed is true, not because its intrinsic truth is seen with the natural light of reason, but because of the authority of God who reveals it" (*Enchiridion Symbolorum,* edited by H. J. D. Denzinger, 29th ed., Freiburg, Germany, 1952, No. 1789).

The Bible finds its place in this system of thought as the book in which divinely imparted truths are written down and thereby made available to all humankind. Indeed, throughout considerable periods of Christian thought the Scriptures have been called the Word of God and have been virtually identified with revelation. The Bible is accordingly thought of as being ultimately of divine authorship; it has been written by human beings, but in the writing of it, their minds were directed by the Holy Spirit. Thus, the First Vatican Council said of the Scriptures that "because they were written as a result of the prompting of the Holy Spirit, they have God for their author" (*Deum habent auctorem*; Denzinger, *Enchiridion Symbolorum,* No. 1787); and in a similar vein, in the twentieth century, the Protestant evangelist Dr. Billy Graham said, "The Bible is a book written by God through thirty secretaries."

The propositional conception of revelation has also been integral to an understanding of the structure of theology that until recently has held unquestioned sway in Christian thought since it was established by Thomas Aquinas in the thirteenth century. This hinges upon the distinction between natural and revealed theology. Natural theology comprises all those truths about God, and about the created universe in its relation to God, that can be arrived at by human reasoning without benefit of divine revelation. Accordingly, the core of natural theology consists in the traditional philosophical arguments for the existence of God. Revealed theology, on the other hand, comprises those truths about God, and about the created universe in its relation to God, that are not accessible to right reasoning as such and that can be known to men only because God has chosen to reveal them. (For example, it is held that while the existence of a supreme being is a tenet of natural theology, the further fact, stated in the Trinitarian dogma, that this being is "three Persons in one" belongs to revealed theology.) These various truths constitute the materials with which the theologian works, his primary task being to bring them together into a systematic body of doctrine.

These conceptions of faith, the Bible, and theology are linked together by the propositional character of revelation, with which they are all concerned. The revelation that is imparted by God, believed by men, published in the holy Scriptures, and systematized in the church's dogmas is a body of theological knowledge. This propositional conception of revelation began to form soon after the end of the New Testament period; reached its fullest development in medieval scholastic thought; was largely abandoned by the first Reformers in the sixteenth century, particularly Martin Luther, but became reestablished in the Protestant scholasticism of the seventeenth and eighteenth centuries; began to be questioned in the later nineteenth century; and was finally set aside by considerable sections of Protestant thought in the twentieth century.

THE *HEILSGESCHICHTLICH* CONCEPTION

The fundamental premise of the propositional view has no place in the nonpropositional conception of revelation that was widely adopted by Christian theologians in the twentieth century. This view maintains that revelation consists not in the promulgation of divinely guaranteed truths but in the performance of self-revealing divine acts within human history. The locus of revelation is not propositions but events, and its content is not a body of truths about God but "the living God" revealing himself in his actions toward man. The nonpropositional view thus centers upon what has come in recent theology to be

known as *Heilsgeschichte* (salvation history) identified as the medium of revelation.

It is not supposed that God has marked his presence by performing a series of miracles, if "miracle" is taken to mean an event that compels a religious response by eluding all natural explanations. It is not characteristic of those theologians who think of revelation in nonpropositional terms to regard the biblical miracles as constituting theistic proofs. Rather, the *Heilsgeschichte* is the way in which a certain segment of human history—beginning with the origins of the national life of Israel and ending with the birth of the Christian community as a response to Jesus—was experienced by men of faith and became understood and remembered as the story of God's gracious dealings with his people. What Christianity (and, confining itself to the Old Testament, Judaism) refer to as the story of salvation is a particular stream of history that was interpreted by prophets and apostles in the light of a profound and consistent ethical monotheism. They saw God at work around them in events that accordingly possessed revelatory significance. The *Heilsgeschichte* is thus a portion of history seen "from the inside" by the illumination of a particular religious faith. The publicly observable series of events forming its basis belongs to secular world history and is capable of a variety of political, economic, psychological, and other analyses besides that of theistic faith. As a central instance of this capacity of history to be construed both nonreligiously and religiously, Jesus of Nazareth, who has been seen by those outside the Christian community in various ways—for example, as rabbi, prophet, or political revolutionary—is seen by Christian faith as the divine Son incarnate in a human life, seeking to draw men into a new life in relation to God.

Revelation, understood in this way, presupposes faith as its correlate. That God is at work in a certain situation, which accordingly serves a revelatory purpose, is always a judgment of religious faith. The part played by faith is thus integral to the total event of revelation, if we use "revelation" to refer to the completed communication that occurs when God's approach has met with a human response. In the words of William Temple, whose formulation of this conception of revelation has become classic, "there is event and appreciation; and in the coincidence of these the revelation consists" (*Nature, Man and God,* p. 314).

As in the case of its older rival, the fuller significance of what may be called the *heilsgeschichtlich* conception of revelation can best be indicated by sketching its implications for the understanding of faith, the Bible, and theological thinking. Clearly, in this view faith is not primarily the believing of revealed propositions, but is rather (in its cognitive aspect) a mode of discernment or interpretation in which men are convinced that they are conscious of God at work in and through certain events of both their personal experiences and world history.

The Bible is not a collection of divine oracles, but a record of the events through which God has revealed himself to a special group, a record that itself functions as a further medium of God's self-revelation beyond that group. It has not been written at the dictation of the Holy Spirit, but has been composed by many different writers at different points within the period of the thousand years or so that it documents. It is distinguished from secular records of the same sequence of events by the fact that it is written throughout from the standpoint of faith. The Old Testament is dominated and unified by the God-centered interpretation of Hebrew history taught by the great prophets, in the light of which the story of the nation came to be understood and celebrated and its chronicles edited. The New Testament is dominated and unified by the witness of Jesus' first disciples and of the Christian communities that grew up around them to the life, death, and resurrection of Jesus, whom they had received as the Christ. The faith by which alone the several writers could produce this particular literature constitutes the "inspiration" that has presided over its production.

Finally, there is no body of divinely authoritative theological propositions. Religious doctrines are not revealed, but represent human—and therefore fallible—attempts to understand the religious significance and implications of the revelatory events depicted in the Scriptures. Theologians who regard revelation in this manner have generally abandoned the traditional natural theology, with its theistic proofs, and base their doctrines instead upon faith as it responds to the scriptural records.

SOME QUESTIONS

One of the questions that Christian theologians have repeatedly discussed is whether there is both general and special revelation. Are nature and history as a whole—including the whole religious history of humankind—revelatory of God, as well as the special occasions of the biblical *Heilsgeschichte*? Many theologians of all communions today hold that God is indeed universally active and that his activity always discloses something of his nature, even though his fullest personal self-revelation has occurred only in the person of Christ.

Another question that has at times been hotly disputed is whether there is an image of God (*imago dei*) in man that constitutes an innate capacity to respond to divine revelation (Emil Brunner) or whether, on the contrary, human nature is so totally corrupted by the Fall that in revealing himself to men God has to create in them a special capacity for response (Karl Barth).

The main philosophical question that arises concerns the criteria by which revelation claims may be judged. For proposition-centered religious thought the answer is provided by natural theology considered as a preamble to revelation. This establishes the existence of God and points, by means of miracles and fulfillments of ancient prophecy, to Christ and the Scriptures as the sources of revealed truth, supplemented in Roman Catholicism by the church as its divinely appointed guardian. For those theologies, on the other hand, that find God at work in historical events whose significance is discerned only by faith, there can be no proof of revelation. Such theologies arise within a community of faith (whether Jewish or Christian) that lives on the basis of what it believes to be an experience of divine revelation. It embodies in its life and literature the "memory" of momentous events in which God has opened a new and better life to humankind. The form of apologetic appropriate to this view is one that defends the right of the believer, as a rational being, given the distinctively religious experience out of which his faith has arisen, to trust that experience and to proceed to live upon the basis of it.

See also Barth, Karl; Brunner, Emil; Faith; Liberation Theology; Luther, Martin; Miracles; Thomas Aquinas, St.

Bibliography

The biblical basis of the idea of revelation is discussed in H. W. Robinson, *Inspiration and Revelation in the Old Testament* (Oxford: Clarendon Press, 1946; 4th ed., 1956), and E. F. Scott, *The New Testament Idea of Revelation* (London and New York: Scribners, 1935).

R. Garrigou-Lagrange, *De Revelatione per Ecclesiam Catholicam Proposita,* 2 vols. (4th ed., Rome, 1945), is a classic Roman Catholic exposition. See also G. H. Joyce, "Revelation," in *Catholic Encyclopedia,* edited by C. G. Herbermann and others, 15 vols. (New York: Appleton, 1907), Vol. XIII.

John Baillie, *The Idea of Revelation in Recent Thought* (New York: Columbia University Press, 1956), traces the modern move in Protestantism to a nonpropositional conception of revelation. B. B. Warfield, *Revelation and Inspiration* (New York: Oxford University Press, 1927), provides a conservative, and C. H. Dodd's *The Authority of the Bible* (London: Nisbet, 1928), a liberal, Protestant view of the place of the Bible in divine revelation. J. K. S. Reid's *The Authority of Scripture* (London: Methuen, 1957) represents a postliberal Protestant point of view.

The nonpropositional conception of revelation is expounded in William Temple, *Nature, Man and God* (London: Macmillan, 1935), Ch. 12; Emil Brunner, *Offenbarung und Vernunft* (Zürich, 1941), translated by Olive Wyon as *Revelation and Reason* (Philadelphia: Westminster Press, 1946); H. R. Niebuhr, *The Meaning of Revelation* (New York: Macmillan, 1941); Karl Barth, "The Christian Understanding of Revelation," in *Against the Stream,* edited by R. G. Smith (New York: Philosophical Library, 1954); and H. Thielicke, *Offenbarung, Vernunft und Existenz* (Gütersloh, 1936). John Baillie and Hugh Martin, eds., *Revelation* (London: Faber and Faber, 1937), is a useful symposium.

For a contemporary Jewish discussion, see Abraham Heschel, *God in Search of Man* (New York: Meridian, 1956). For the Islamic conception of revelation, not touched upon in this entry, see I. Goldziher, *Die Richtungen der islamischen Koranauslegung* (Leiden: E.J. Brill, 1920), or A. J. Arberry, *Revelation and Reason in Islam* (New York: Macmillan, 1957); and for Hindu conceptions, K. S. Murty, *Revelation and Reason in Advaita Vedanta* (New York: Columbia University Press, 1959).

For contemporary skepticism concerning revelation, see Julian Huxley, *Religion without Revelation* (London: Watts, 1941; 2nd ed., London: Parrish, 1957), and Ronald Hepburn, *Christianity and Paradox* (London: Watts, 1958).

OTHER RECOMMENDED TITLES

Abraham, William. *Divine Revelation and the Limits of Historical Criticism.* Oxford: Oxford University Press, 1982.

Avis, Paul, ed. *Divine Revelation.* Grand Rapids, MI: Eerdmans, 1997.

Dulles, Avery. *Models of Revelation.* New York: Doubleday, 1983.

Mavrodes, George. *Revelation in Religious Belief.* Philadelphia: Temple University Press, 1988.

Swinburne, Richard. *Revelation: From Metaphor to Analogy.* Oxford: Clarendon Press, 1992.

Trembath, Kern. *Divine Revelation: Our Moral Relation with God.* Oxford: Oxford University Press, 1991.

Ward, Keith. *Religion and Revelation.* Oxford: Clarendon Press, 1994.

Wolterstorff, Nicholas. *Divine Discourse: Philosophical Reflections on the Claim That God Speaks.* Cambridge, U.K.: Cambridge University Press, 1995.

John Hick (1967)

Bibliography updated by Christian B. Miller (2005)

REVERSE MATHEMATICS

Reverse mathematics has its origins in Harvey Friedman's 1974 address to the International Congress of Mathematicians. In it Friedman asked two fundamental questions: "What are the proper axioms to use in carrying out proofs of particular theorems, or bodies of theorems, in mathematics?" and "What are those formal systems which

isolate the essential properties needed to prove them?" Reverse mathematics was developed as an attempt to answer these questions, and since 1974 many logicians (especially Friedman and Stephen Simpson) have contributed to this project.

The goal in reverse mathematics is to find the minimal collection *S* of set theoretic axioms which suffices to prove a given theorem *T*. Because Zermelo-Frankel set theory is too powerful to provide this type of delicate analysis, second order arithmetic is used as the axiomatization of set theory. The formal language of second order arithmetic contains the symbols $+, \cdot, <, 0, 1, \in$, and $=$, as well as two types of variables: number variables (denoted by lower case letters and intended to range over natural numbers) and set variables (denoted by upper case letters and intended to range over sets of natural numbers). In this formalization, sets of numbers are referred to using the set variables, but there are no variables that range over sets of sets of numbers. Thus, unlike Zermelo-Frankel set theory, second order arithmetic has to treat collections of sets as formal classes. One potential point of confusion concerning second order arithmetic is that despite its name, it is not a form of full second order logic. Second order arithmetic uses first order predicate logic, but allows two distinguished kinds of variables to separate its notation for numbers from its notation for sets. Therefore, the usual tools of first order logic such as compactness and the Lowenheim-Skolem theorems apply to second order arithmetic and its subsystems. In particular, there are countable models of second order arithmetic.

The axioms for second order arithmetic fall into three categories. First, there is a finite number of axioms stating the basic relationships between $+, \cdot, <, 0$, and 1 in the natural numbers. Two examples of these axioms are that for all m, $m + 0 = m$ and that for all m and n, $m \cdot (n + 1) = (m \cdot n) + m$. Technically these axioms are exactly the noninduction axioms from Peano arithmetic. Second, there is an induction axiom for sets which says that from the assumptions that 0 is an element of X and that for all n, if n is an element of X, then $n + 1$ is an element of X, we can conclude that every n is an element of X. This axiom captures the fundamental inductive nature of the natural numbers. Third, there is an infinite collection of axioms called the comprehension scheme. For each formula $\varphi(x)$ in the language of second order arithmetic (allowing additional free variables as parameters), there is an axiom stating that there exists a set whose members are exactly the numbers n for which $\varphi(n)$ holds. Because second order arithmetic does not allow the formation of sets of sets, these axioms do not give rise to a version of Russell's Paradox concerning the set of all sets which are not members of themselves.

The first step in analyzing a theorem *T* in reverse mathematics is to formalize the statement of *T* in second order arithmetic. This formalization can be done for most theorems in areas of mathematics that can be captured in some countable manner, such as classical geometry, number theory, real and complex analysis, countable algebra, and countable combinatorics. However, one of the limitations of using second order arithmetic as the underlying form of set theory is that it is not well suited to formalizing theorems from subjects such as general topology that depend heavily on the twentieth-century development of abstract set theory.

The second step is to find a subsystem *S* of second order arithmetic that is strong enough to prove *T*. This step often involves translating a classical proof of *T* into the formal system of second order arithmetic and letting *S* be the collection of axioms needed in the proof.

The third step is to show that *T* can prove each of the axioms in *S* over a suitably weak base theory (described below). This process of proving the axioms from the theorem is called a reversal, and it gives rise to the name reverse mathematics. If the system *S* contains axioms which are too powerful then the third step may be not possible. For example, a classical proof may use more axioms than are necessary. Therefore it is not uncommon to return to the second step and to try to find a different proof of *T* which uses weaker axioms. The third step is then repeated to see if these weaker axioms are provable from *T*. Once the third step is realized, the equivalence of the theorem *T* with the axioms *S* shows that *S* is a minimum collection of axioms which suffices to prove *T*.

Because the comprehension axioms (which state that for any formula $\varphi(x)$, there is a set whose elements are exactly the numbers n for which $\varphi(n)$ holds) are the only ones which explicitly state the existence of sets, the subsystem *S* is often formed by restricting the types of formulas allowed in this scheme. For example, the base theory over which one typically proves the reversals is denoted RCA and is called Recursive Comprehension Axiom.

Roughly, RCA restricts the comprehension scheme to those formulas which define sets whose membership can be calculated by a finite algorithmic procedure.

Many of the theorems analyzed in reverse mathematics as of this writing fall into one of five categories. They are either provable in RCA or equivalent to one of four standard subsystems: WKL (which is RCA plus an

axiom stating that every infinite binary branching tree has an infinite path); ACA (which allows sets to be defined by formulas that do not contain quantifiers ranging over sets, but may contain quantifiers ranging over numbers); ATR (which allows sets to be defined by transfinite recursion); and Π^1_1-CA_O. (which allows sets to be defined by formulas containing arbitrarily many quantifiers ranging over numbers and at most one quantifier ranging over sets). A small number of examples do not fit neatly into these systems, but there is little evidence that this number will not grow as more theorems are analyzed.

One of the philosophical applications of reverse mathematics is that it provides a general framework for showing the necessity of impredicative methods in particular areas of mathematics. (Roughly, a set *A* is predicative if it can be defined as the set of all numbers satisfying a predicate for which the truth value does not depend on the existence of *A*.) On the one hand, if a theorem *T* is equivalent to a set of axioms that contains impredicative axioms (such as Π^1_1-CA_O.) then *T* cannot be established without the use of impredicative methods. On the other hand, if *T* can be proved inside a system which contains only predicative axioms (such as ACA), then this proof shows that *T* constitutes a piece of predicative mathematics. Some of the other subsystems have similar foundational connections. For example, because RCA restricts comprehension to formulas defining sets whose membership can be decided by a finite algorithmic procedure, proofs in RCA have a number of similarities to constructive proofs. However, this analogy is not perfect because most varieties of constructivism differ sharply with RCA over the treatment of induction and over the law of excluded middle.

A second philosophical application of reverse mathematics is to give a partial realization of Hilbert's program. In order to eliminate concerns over set theoretic paradoxes and to establish the consistency of infinitary methods in mathematics, Hilbert tried to find two formal systems to capture mathematical reasoning. The first system would be foundationally secure but would only capture finitary reasoning. The second system would be large enough to encompass all of mathematics including the general methods of infinitary reasoning. His goal was to use the first system to show the consistency of the second system, thus justifying the use of infinitary methods once and for all. Gödel's Second Incompleteness Theorem is widely viewed as showing that Hilbert's program cannot succeed as it was originally conceived. However, it is still possible that a reasonable fragment of mathematics could be developed within a formal system that could be shown to be relatively consistent in a finitistic formal system. Hilbert did not provide a strict definition for his notion of finitary, but it has been argued that the system of primitive recursive arithmetic satisfies his intuitive concept. Furthermore, since primitive recursive arithmetic can prove the relative consistency of WKL, the (substantial amount of) mathematics that can be developed in WKL is finitely reducible in Hilbert's sense and provides a partial realization of Hilbert's program.

See also First-Order Logic; Geometry; Hilbert, David; Logic, History of: Modern Logic: From Frege to Gödel; Logic, History of: Modern Logic: Since Gödel: Friedman and Reverse Mathematics; Second-Order Logic; Set Theory.

Bibliography

Feferman, Solomon. "Systems of Predicative Analysis." *Journal of Symbolic Logic* 29 (1964): 1–30.

Friedman, Harvey. "Some Systems of Second Order Arithmetic and Their Use." *Proceedings of the 1974 International Congress of Mathematicians. Volume 1*. Canadian Math Congress, 1975: 235–242.

Simpson, Stephen G. "Partial Realizations of Hilbert's Program." *Journal of Symbolic Logic* 53 (1988): 349–363.

Simpson, Stephen G. *Subsystems of Second Order Arithmetic.* Berlin: Springer, 1998.

Tait, William W. "Finitism." *Journal of Philosophy* 78 (1981): 524–546.

Peter Cholak (2005)
Reed Solomon (2005)

RIBOT, THÉODULE ARMAND
(1839–1916)

Théodule Armand Ribot, the French psychologist, was a professor of psychology at the Sorbonne and from 1889 was the director of the psychological laboratory at the Collège de France. A philosophical disciple of Hippolyte Taine and Herbert Spencer (whose *Principles of Psychology* he translated), Ribot, with Taine, initiated the study in France of a positivistic and physiologically oriented psychology. His interest in philosophy was inseparable from his interest in concrete psychological problems and persisted throughout his life. He founded and edited the *Revue philosophique de la France et de l'étranger,* one of the first French philosophical journals. Ribot influenced not only French positivists and physiological psycholo-

gists but even some thinkers who, like Henri Bergson, rejected his epiphenomenalism.

Ribot's work falls into three main periods, but he remained loyal throughout his life to the program expounded in the introduction to his first book, *La psychologie anglaise contemporaine* (Paris, 1870). He insisted that psychology must be liberated from "the yoke of metaphysics" and stressed the need for an empirical, biological approach to psychology and the limitations of an exclusive reliance on introspection. However, although he insisted on excluding metaphysics from the empirical sciences, he did not dismiss it altogether. The works of Ribot's first period were mainly expository and historical. *La psychologie anglaise contemporaine* surveyed English associationist psychology from David Hartley to Samuel Bailey. In *La psychologie allemande contemporaine* (Paris, 1879) he introduced the work of Gustav Fechner, Wilhelm Wundt, Hermann Helmholtz, and others to the French public. *La philosophie de Schopenhauer* (Paris, 1874) foreshadowed Ribot's later emphasis on the affective and instinctive basis of personality.

Ribot's second period, characterized by an interest in psychopathology, produced three classic works: *Les maladies de la mémoire* (Paris, 1881), *Les maladies de la volonté*; (Paris, 1883), and *Les maladies de la personnalité* (Paris, 1885). Despite a wealth of clinical, empirical material, the underlying motive of these works was philosophical—a positivistic distrust of such reified abstractions as "memory," "will," and "self." These abstractions had played a prominent role in French speculative psychology and in Victor Cousin's eclectic idealism. Ribot showed that the simplicity of such abstract words hides the complexity of the phenomenon named, a complexity revealed by the dissociation found in mental diseases. Ribot was among the first to study dissociations of personality, and his law of regression—that amnesia affects the most recent and least organized impressions and reactions first—was a lasting contribution to psychology.

In Ribot's third period, which began with his *La psychologie de l'attention* (Paris, 1888), his interest shifted to normal psychological phenomena, particularly to affective phenomena. The major work of this period, *La psychologie des sentiments* (Paris, 1896), reflects Ribot's biological approach and his epiphenomenalism. Physiological drives underlie our elementary feelings of pleasure and pain, and more complex and evolved stages of these drives underlie more complex emotions. Organic sensibility evolved prior to consciousness, and feelings prior to intellect. Ribot's last work, *La vie inconsciente et les mouvements* (Paris, 1914), interpreted various manifestations of subconscious activity in terms of motor activity.

See also Bergson, Henri; Fechner, Gustav Theodor; Hartley, David; Helmholtz, Hermann Ludwig von; Psychology; Spencer, Herbert; Taine, Hippolyte-Adolphe; Wundt, Wilhelm.

Bibliography

Dugas, Laurent. *La philosophie de Théodule Ribot.* Paris, 1924.

Faguet, Émile. *Propos littéraires,* fourth series. 1903.

Krauss, S. *Théodule Ribots Psychologie.* Jena, 1905.

Lamarque, G. *Théodule Ribot: Choix des textes et étude de l'oeuvre.* Paris, 1913.

Taine, Hippolyte. *Derniers Essais de critique et d'histoire,* 6th ed. Paris: Hachette, 1923.

Milič Čapek (1967)

RICHARD OF MEDIAVILLA

(d. c. 1300)

Richard of Mediavilla, or Richard of Middleton, *doctor solidus,* was a Franciscan philosopher, theologian, and canon lawyer. Although his date of birth and country of origin are unknown, scholars are generally agreed that he was either French or English. We are certain that in 1283 he was appointed as one of the judges of the works of Peter John Olivi, and we possess three of his sermons, preached in Paris in 1281 and 1283. He was a master of theology in Paris during 1284–1285. In 1288, Richard was one of the tutors of the exiled Prince Louis, son of King Charles II of Sicily and later bishop of Toulouse. Richard's last writings seem to date around 1295, when he completed his commentary on the fourth book of the *Sentences* of Peter Lombard. After 1295 we lose all trace of Richard of Mediavilla.

Richard was a scholar in the tradition of Bonaventure and John Peckham. He seems to have had a flair for clear and orderly presentation and to have enjoyed wide popularity among his Franciscan confreres. Like many of his fellow Franciscans, he regarded Bishop Tempier's condemnation of 219 propositions in 1277 as definitive. As a result, he set himself to defend, clarify, and organize a philosophy and theology that would vindicate and establish the doctrines contrary to the condemned propositions. He differs from most of his fellow Franciscans, however, in that he is more sympathetic to the Thomistic theory of knowledge.

Richard was one of the first Franciscans to reject the Augustinian theory of divine illumination. For Richard our ideas are solely the result of abstraction from sensible things, though as universals they are strictly intramental. In metaphysics he held that being is predicated analogically, not univocally, of God and creatures. Because every effect somehow bears the trademark of the first cause, God's existence can be proved from the world of nature. Richard found the so-called a priori argument of Anselm unacceptable; he adopted Henry of Ghent's position that essence and existence are only intentionally, not really, distinct. His doctrine of universal hylomorphism—that is, that all creatures are composed of matter and form—coincides with that of Bonaventure. Richard's theory of one substantial form's consisting of multiple grades constitutes the most complete and well-ordered doctrine of the plurality of forms in the Middle Ages. Richard argues to the soul's spirituality from the immateriality of universal concepts. The faculties of intellect and will are not accidents of the soul, nor do they add to its essence; they merely constitute a new relation between the essence of the soul and its acts and objects. Liberty is formally in the will. In common with his Franciscan confreres, Richard asserted that the will is a more noble faculty than the intellect.

Conservative by nature, Richard of Mediavilla was not one to shrink from speaking out. In one remarkable statement we catch a glimpse of his spirit in the search for truth and goodness: "We must start a good war. It is better to fight against falsehood and malice with a certain amount of discord, than, by dissimulating, to give way to malice and falsehood for the sake of harmony" (*Quodlibeta* III, 22).

See also Anselm, St.; Augustinianism; Bonaventure, St.; Henry of Ghent; Illumination; Medieval Philosophy; Olivi, Peter John; Peckham, John.

Bibliography

WORKS BY RICHARD

Commentarius in Sententias, 4 vols. Brescia, 1591.

Quaestiones Disputatae. Vaticani Latini No. 868. Folios 105–116.

Quodlibeta. Brescia, 1591.

WORKS ON RICHARD

Hocedez, Edgar. *Richard de Middleton. Sa vie, ses oeuvres, sa doctrine.* Louvain: "Spicilegium sacrum lovaniense" Bureaux, 1925.

Zavalloni, Roberto. *Richard de Mediavilla et la controverse sur la pluralité des formes.* Louvain, 1951. Has excellent bibliography.

Girard J. Etzkorn (1967)

RICHARD OF MIDDLETON

See *Richard of Mediavilla*

RICHARD OF SAINT VICTOR

See *Saint Victor, School of*

RICKERT, HEINRICH

(1863–1936)

Heinrich Rickert, the German neo-Kantian philosopher, was born in Danzig and received his degree in 1888 from the University of Strasbourg. In 1891 he began lecturing at Freiburg, succeeding Alois Riehl as professor in 1894. In 1916 he went to Heidelberg as successor to Wilhelm Windelband.

Rickert belonged to the southwestern school of neo-Kantianism. His main efforts were devoted to a study of the logical and epistemological foundations of the natural sciences and to the historical disciplines in the hope of arriving at a "unity of reality and values." He departed from Wilhelm Dilthey in his criticism of Dilthey's subjective approach to the understanding of historical reality and in his attempt to find a set of more objective criteria; his departure from Windelband consisted in rejecting Windelband's separation of natural and historical disciplines and offering instead a theory that considered all reality to be historical.

PHILOSOPHY AND NATURAL SCIENCE

In his early work, particularly in *Der Gegenstand der Erkenntnis* (Freiburg, 1892), Rickert raised the question of the relationship between philosophy and the natural sciences. He denied the universal validity of the method of the natural sciences and attempted to establish the primacy of practical reason as the foundation of his epistemology. He believed that only the Kantian critical method is adequate for explaining the epistemological presuppositions and limitations of the various sciences. While phenomenology may provide a method for describing the contents of consciousness, it fails to account for their intelligibility and relationship to objective reality. Hegelianism, on the other hand, in identifying the real with the rational, leaves out of account or distorts

the pluralistic character of reality. Only critical philosophy yields knowledge that is both universal and necessary; it alone can explain the pluralistic, dynamic, and yet rational character of society and history. In view of the lack of philosophical attention to the historical disciplines and because the then prominent philosophical problems of *Weltanschauung* seemed to hinge most directly on distinguishing scientific thinking from historical thinking, Rickert devoted himself thereafter primarily to the problem of historical conceptualization (*Begriffsbildung*).

INDIVIDUALIZING AND GENERALIZING THOUGHT

On the basis of Windelband's distinction between nomothetic (universal) and ideographic (particular) judgments, Rickert developed his logic of the historical disciplines. At both the scientific and the prescientific stages of conceptualization, he claimed, there are two ways of grasping reality: individualizing and generalizing. Individualizing thought is proper to historical thinking. Instead of fabricating a copy of a historical phenomenon in its complex totality, it establishes the essential relationships that bind the phenomenon to its environment and traces the various stages of its development. Philosophy studies the concept of development, while the objects of historical study are unique developments. Generalizing thought, therefore, is proper to the natural sciences but is inapplicable to history. "Reality," Rickert claimed, "becomes nature if we consider it in regard to what is general; it becomes history if we consider it in regard to the particular or individual" (*Kulturwissenschaft und Naturwissenschaft,* 5th ed., p. 63).

Historical method for Rickert is highly selective, and in the selection of data, value judgments are operative from the very outset. This being the case, the determination of value criteria (*Wertbegriffe*) becomes the primary concern of historical understanding. Generalizing thought is logically free of values (*wertfrei*) because it constructs universally valid concepts. The particular objects to which they apply are interchangeable, and each object, abstracted from all its other relationships, functions only to illustrate the general law. Although in generalizing thought a selective process is at work to determine the common character of a group of particulars, it is the common character, expressed in a formula, that is essential. The aim of generalizing thought is precisely to free its objects from relations of value (*Wertverbindungen*).

KULTURWISSENSCHAFT

Although history is a science of values, this does not mean that the historian may organize his inquiry arbitrarily; in that case history would be mere propaganda. In order for history to be objective, its values (state, law, art, religion) must be universal. The universality of historical values must be established epistemologically, and the relevance of the various social phenomena with respect to these values must be demonstrated empirically. Because history is written by, about, and for civilized men, social activity must be its subject matter. Since social activity can be grasped only by individualizing thought in terms of its significance for universal values, the historian's criterion must be culture, because social activity and value most nearly converge in culture. Culture is most directly concerned with the realization of universal values: "Culture is the common affair in the life of the nations; it is the possession with respect to the values of which the individuals sustain their significance in the recognition of all peoples, and the cultural values which adhere to this possession are therefore those which guide historical representation and conceptual formation in the selection of what is most essential" (*Die Grenzen der naturwissenschaftlichen Begriffsbildung,* 2nd ed., p. 509). Thus, believing that his method made of history a logically valid discipline that deals with objective reality, Rickert called the historical sciences *Kulturwissenschaft* (cultural science) in preference to Dilthey's term, *Geisteswissenschaft* (science of the mind or spirit).

UNIVERSAL HISTORY

Far from being a contradiction, universal history is not only possible but is the logical outcome of the search for the value principles (*Wertprinzipien*) according to which the historical process as a whole may be viewed. "The system of values provides the possibility of systematization, and the relationship [of history] to the system of values permits of individualizing treatment" ("Geschichtsphilosophie," p. 400). But precisely because the evaluation of the whole of history is involved, the system of value principles must be purely formal. "We would need something timeless in order to extract an objective sense from the temporal course of history" (ibid., p. 418). Like Immanuel Kant, Rickert proposed three stages in the development of civilization: dogmatism, skepticism, and criticism, the last of which was the achievement of German idealistic philosophy. While this periodization cannot be verified empirically, it is an example of the critical approach to the question of the unity of historical development. Although it is purely theoretical, it nonetheless

gives an axiological grounding to the results of empirical research. In the last analysis, the problem of universal history is to introduce a method whereby the real and the ideal may be theoretically synthesized.

CRITICISM

The principal criticism brought against Rickert is that the introduction of a transcendental system of values is unhistorical and leads to the reification of existing values (*Wertabsolutierung*). In isolating universality by viewing it as a distinct realm of thought rather than as a function of all thought, Rickert actually confirmed the positivism and cultural relativism he had sought to overcome. In radically separating the universal from the particular, he was compelled to regard historical data as being identical with those of science, a series of discrete facts that differ only in the relationships in which they are observed. Nevertheless, the fruitfulness of Rickert's theory is borne out by his influence on such contemporaries as Ernst Troeltsch, Friedrich Meinecke, and Max Weber.

See also Dilthey, Wilhelm; Geisteswissenschaften; Hegelianism; Historicism; Meinecke, Friedrich; Neo-Kantianism; Phenomenology; Philosophy of History; Riehl, Alois; Troeltsch, Ernst; Weber, Max; Windelband, Wilhelm.

Bibliography

WORKS BY RICKERT

Die Grenzen der naturwissenschaftlichen Begriffsbildung. Tübingen: Mohr, 1896–1902; 2nd ed., Tübingen, 1913.

Kulturwissenschaft und Naturwissenschaft. Tübingen, 1899; 5th ed., Tübingen, 1921.

Die Probleme der Geschichtsphilosophie. Leipzig: Duncker and Humblot, 1905; 3rd ed., Heidelberg: C. Winter, 1924.

"Geschichtsphilosophie." In *Die Philosophie im Beginn des zwanzigsten Jahrhunderts, Festschrift für Kuno Fischer,* edited by Wilhelm Windelband. Heidelberg, 1907.

Kant als Philosoph der modernen Kultur. Tübingen: Mohr, 1924.

WORKS ON RICKERT

Becher, E. *Geisteswissenschaften und Naturwissenschaften.* Munich: Duncker and Humblot, 1921.

Cohen, M. R. "The Insurgence against Reason." *Journal of Philosophy* 22 (1925): 120–123.

Collingwood, R. G., A. E. Taylor, and F. C. S. Schiller. "Are History and Science Different Kinds of Knowledge?" *Mind* 31 (1922): 426–466.

Faust, A. *Heinrich Rickert und seine Stellung innerhalb der deutschen Philosophie der Gegenwart.* Tübingen: Mohr, 1927.

Federici, F. *La filosofia dei valori de Heinrich Rickert.* Florence, 1933.

Gurvitch, Georges. "La théorie des valeurs de Heinrich Rickert." *Revue Philosophique* 124 (1937): 80–85.

Troeltsch, Ernst. "Über den Begriff einer historischen Dialektik. Windelband-Rickert und Hegel." *Historische Zeitschrift* 119 (1919): 373–426.

Robert Anchor (1967)

RICOEUR, PAUL

(1913–2005)

Paul Ricoeur is widely regarded as among the most important French philosophers of the twentieth century. He had contributed to most of the major philosophical movements from the 1940s to the present, including existentialism, phenomenology, hermeneutics, structuralism, critical theory, narrative theory, philosophy of religion, ethical theory, political philosophy, and philosophy of law. Ricoeur was a prolific author of twenty-seven books and more than 500 articles as of 2004. His works tend to focus on theories of interpretation and the philosophy of human nature, examining the limits on our ability to understand the world and to know ourselves. If there is a guiding thread that runs through Ricoeur's career it would be an attempt to develop a philosophical anthropology of human capability, in particular our capacities to act, understand, communicate, and be responsible.

Born in 1913 in Valence, France, Ricoeur studied classics and philosophy at the University of Rennes and at the Sorbonne. After holding a number of teaching positions in provincial colleges, he was drafted into the French army in 1940. He was soon captured and spent the next five years in a German prison camp. While in prison, Ricoeur translated Edmund Husserl's book *Ideas* (1913) into French and coauthored a study on Karl Jaspers with fellow inmate Mikel Dufrenne. After he was freed in 1945, Ricoeur taught at the University of Strasbourg (1948–1956), the Sorbonne (1956–1966), and the University of Paris, Nanterre (1966–1987). In 1970 he succeeded Paul Tillich as the John Nuveen Professor of Philosophical Theology at the University of Chicago where he held a joint appointment at the School of Theology and Department of Philosophy until his retirement in 1992. Ricoeur continued to publish works on hermeneutics, moral-political philosophy, and theology until his death in May 2005.

EXISTENTIAL-PHENOMENOLOGY

Ricoeur's early works were devoted to a phenomenological study of the human will. He sought to combine the

existentialist themes of Gabriel Marcel (incarnate existence) and Karl Jaspers (limit situations, such as birth, war, and death) with the methodological rigor of Husserlian phenomenology. The result is a proposed three-volume, systematic "philosophy of the will" that includes *Freedom and Nature: The Voluntary and the Involuntary* (1950), *Fallible Man* (1960), and *Symbolism of Evil* (1960). These works form the core of Ricoeur's early philosophical anthropology. The third volume was to be on the "poetics of the will" but was never written.

In *Freedom and Nature*, Ricoeur employs the Husserlian method of eidetic analysis to the spheres of the will, affection, and volition. The goal is to describe the structures of voluntary action to uncover our fundamental possibilities of existence. Ricoeur retains from Husserl the central insight into the intentionality of consciousness and the methodological technique of bracketing, while recognizing that phenomenology must be supplemented with non-phenomenology given the limits placed on knowledge by the body. A phenomenology of action reveals the full extent to which consciousness is embodied and tied to involuntary functions, thus known both phenomenologically (as a subject of the will) and empirically (as an object for the will). The experience of our own bodies is never direct and unmediated; instead we interpret the involuntary aspects of our bodies as signs or symptoms for the will. These signs are read indirectly through one's will as indications of the involuntary for the voluntary. We find consciousness in the body and the body in consciousness.

Ricoeur shows how the act of willing is both the realization of freedom and the reception of necessity. The act of willing has three moments, each inextricably related to the involuntary. The three parts are: 1) "I decide"; 2) "I move my body"; and 3) "I consent." Each part has an object (or intentional correlate): A) the decision or project; B) the action or motion; C) the acquiescence or consent. Finally, each correlate is itself related to the different modes of the involuntary: a) motives, needs, values; b) skills, emotions, habits; c) character, unconscious, life. There is a fundamental reciprocity of voluntary decision, choice, and action with involuntary bodily functions, which act as a vehicle for the will. The involuntary necessity of the body both limits and enables human freedom. Yet the unity of mind and body, voluntary and involuntary, is never fully realized. Rather it is a regulative idea for understanding how humans are both free and constrained. Embodied freedom and unifying the will free from conflict is something we can only hope for but never completely realize. A poetics of the will in the proposed third volume was to be Ricoeur's attempt to show how imaginative and creative uses of language can suggest ways to reconcile the dualism of mind and body.

After having described the eidetic structures of the will as incarnate freedom, in the second volume of the philosophy of the will Ricoeur seeks to uncover the actual conditions of existence through an "empirics of the will." One of Ricoeur's aims is to overcome the tendency among existentialists to overvalue human transcendence and devalue human finitude. He believes that the mistake made by Jean-Paul Sartre and Martin Heidegger is to equate finitude (our inevitable and necessary limitations) with guilt (an undesirable experience of limitation). In *Fallible Man*, Ricoeur examines the conditions under which the will confronts its finitude and chooses evil and sin. The reason one chooses evil stems from the divided will. There is a disproportion between our finite limitations and infinite possibilities. This gap between our limited bios (our bodies, passions, and desires) and unlimited logos (our reason capable of grasping universals) renders us fragile and fallible. The fractured will, or fault in our existence (like a geological fault), opens the way for temptation, evil, and sin. We are not evil by nature but we have the capacity to be thanks to the disproportion in our will. Following Immanuel Kant, Ricoeur analyzes the existential significance of our fallibility in terms of imagination (the limits of knowledge), character (the limits of the will), and feeling (the limits of our emotion). The ineliminable conflict within human beings constitutes our capacity both for good and evil.

In the second volume of the empirics of the will, *The Symbolism of Evil*, Ricoeur continues to examine our capacity for evil by considering the various ways humans are already guilty, sinful, and fallen. He conducts a "phenomenology of confession" that describes the way we experience the transition from fallibility to fault (from our potential for evil to actually being evil). According to Ricoeur, confession arises from three sources: defilement, sin, and guilt. Defilement is interpreted as an objective state of impurity, sin as a social state, and guilt as a psychological interiorization of sin. Ricoeur then shows how this progressive "fallenness" is reproduced in four basic types of myth: myths of creation, myths of tragedy, myths of the fall, and myths of exile. Each type of myth is a symbolic expression of our experience of evil.

HERMENEUTICS

Through his analysis of myths, Ricoeur began to shift away from phenomenology to hermeneutics as he became more interested in the symbolic systems that

relate us to the world and impose an indirect or interpretive approach to knowledge and self-understanding. Symbols are double-meaning expressions with an apparent, first-order, meaning and a hidden, second-order, meaning. Symbols must be interpreted rather than merely perceived in order to be understood. Ricoeur contends that if language is taken to be the medium for thought and experience, it is impossible to realize a prelinguistic and presuppositionless realm of consciousness. As a result, we can never have the kind of unmediated knowledge that phenomenologists have traditionally hoped to attain. As such, the mediation of self-understanding by signs, symbols, and language requires an interpretive, hermeneutic philosophy. Ricoeur often speaks of the detour self-understanding must take through language. The idea of a detour as a hermeneutical technique for reading signs of experience through something else is one of Ricoeur's favorite metaphors that reappears throughout his career.

In *Freud and Philosophy* (1965), Ricoeur develops a hermeneutic philosophy by contrasting Husserlian phenomenology and Freudian psychoanalysis. According to Ricoeur, Sigmund Freud introduces a model for understanding the relationship between experience and desire, as well as a technique for uncovering the relation of a latent, unconscious meaning to a manifest, conscious meaning. The unconscious is an interplay of language and desire that reveals and conceals, thus shaping and distorting how we understand ourselves and others. Freud removes the illusion of a subject that ostensibly is immediately transparent to itself, thereby frustrating the aim of phenomenology to describe experience faithfully. For Ricoeur, Freud's contribution to hermeneutics is a theory of interpretation geared toward unmasking and decoding symbolic expressions. Dreams and symbols are models of the complexity of language in which meanings are both given and hidden. This symbolic language requires an interpretation in terms of rules and law-like regularities to understand how it mediates experience. Psychoanalysis, on this account, is a hermeneutic technique for interpreting the semantics of desire, that is, the interrelations among language, experience, bodily desires, and culture.

In *Freud and Philosophy*, Ricoeur contrasts two opposing kinds of hermeneutics: the hermeneutics of belief and the hermeneutics of suspicion. The hermeneutics of belief is geared toward recovering and recollecting lost or forgotten meanings. Understanding a religious symbol involves a hermeneutics of belief because to understand its full meaning one must already have the prior belief that it is sacred. The hermeneutic situation is that we must believe in order to know, yet know in order to believe.

By contrast, the hermeneutics of suspicion is geared toward unmasking, demystifying, and removing the illusions of symbols, which not only reveal but conceal meaning. Ricoeur draws on the "masters of suspicion," Karl Marx, Friedrich Nietzsche and Freud, each of whom posit a false consciousness in place of an immediate, self-transparent consciousness, and deception or delusion in place of the experience of participation. The hermeneutics of suspicion decipher meanings hidden and distorted by literal and apparent meanings. Ricoeur argues that self-understanding involves a dialectic of belief and suspicion: We must have a clear understanding of our past that is shaped by a projection of what we hope we can become.

In *The Conflict of Interpretations* (1969), Ricoeur further develops a hermeneutic philosophy through his confrontation with structuralist semiotics. Like psychoanalysis, structuralist semiotics calls into question the primacy of consciousness as the privileged, self-evident home of meaning. Ricoeur retains the insight of structuralists, such as Ferdinand de Saussure and Claude Levi-Strauss, that language has objective characteristics best understood as an empirical science and that meaning is a function of a different agency than consciousness. Yet Ricoeur maintains this aspect of language without rejecting the fundamental intentionality of consciousness and role of the individual as a bearer of meaning. Language has expressive meanings that must be understood from the perspective of the first person as well as objective meanings that must be understood from the perspective of the third person. Ricoeur tries to integrate a structuralist method of objective explanation into an interpretive theory for understanding spoken and written language.

As a result of the confrontation with structuralism, Ricoeur develops a theory of language as discourse. In *Interpretation Theory* (1971), discourse is defined as a dialectic of event and meaning, sense and reference. Discourse takes place as an event but has an ideal, repeatable meaning that allows what is said to be repeated, identified, and said differently. As an event, discourse is referential (about something), self-referential (said by someone), temporal (said at some moment), and communicative (said to someone).

As a meaning, discourse is both what the speaker means and what the utterance means. The dialectic of event and meaning, sense and reference constitutes writ-

ten discourse as well. But where spoken discourse is addressed to someone, in a particular dialogical situation, written discourse is addressed to an indefinite number of absent readers. The task of hermeneutics is to understand the matter of the text that is autonomous with respect to the intentions of the author, its original addressee, and the context in which it was written. The matter of the text discloses a proposed world (of real and imaginary references) one could possibly experience, inhabit, verify, criticize, and so on. To appropriate the meaning of a text we must first let go and relinquish the illusion that subjectivity alone confers meaning. Now the text, not the symbol, is Ricoeur's model for the linguistic mediation of experience. Self-interpretation is mediated by textual interpretation; conversely, textual interpretation results in self-interpretation.

The hermeneutics of texts also applies to actions. In his article *The Model of the Text as Meaningful Action* (1971), Ricoeur argues that actions, like texts, are readable, with a meaning that is independent of the intentions of the actors and subject to conflicting interpretations. In the same way that a text becomes detached from its author, an action is detached from its agent and may take on unintended meanings of its own. The meaning of an action is then open to an indefinite number of interpretations by an indefinite number of possible readers.

Ricoeur believes that if human action can be read and interpreted like written works, then the methods and practices of textual interpretation can function as a paradigm for the interpretation of action for the social sciences. Ricoeur accepts Wilhelm Dilthey's distinction between two forms of inquiry: scientific explanation of the natural world, and historical understanding of the social world. Yet Ricoeur maintains that hermeneutics is a dialectic of explanation and understanding. Texts and actions have underlying structures to be explained as well as social meanings to be understood.

METAPHOR AND NARRATIVE

Throughout his career Ricoeur has examined how imaginative and creative uses of language improve our ability to express ourselves and extend our understanding of the world. Symbols, myths, metaphors, and fiction can capture experience in ways that ordinary, descriptive language cannot. Ricoeur maintains that the reference of creative language is divided or split, meaning that such writing points to aspects of the world can only be suggested and referred to indirectly. Creative language refers to such aspects of the world as if they were real and as if we could be there. In *The Rule of Metaphor* (1975), Ricoeur develops his thesis that the split-reference of creative discourse discloses a possible way of being in the world that remains hidden from ordinary language and first-order reference. A metaphor is a heuristic fiction" that redescribes reality by referring to it in terms of something imaginative or fictitious, allowing us to learn something about reality from fiction. Heuristic fictions help us to perceive new relations and new connections among things, broadening our ability to express ourselves and understand ourselves.

Like all discourse, a metaphor is a communicative utterance that is produced as event, but understood as meaning. Yet, only a live metaphor is at the same time both event and meaning. A dead metaphor has lost its event character when it becomes a commonplace expression, such as, for example, to describe someone who is nervous as having butterflies in their stomach. A live metaphor contains a metaphorical twist that produces a new, surprising meaning. The meaning results from a tension in the way something is described metaphorically and how we normally understand it to be. In order to grasp the differences and resemblance that constitute a metaphor, we must see through the first-order, ostensive reference to the second-order, creative reference to understand how it relates the world. To understand what a metaphor means is to see that it is similar to and different from an ordinary description. The tension in a living metaphor between literal and imaginative must be preserved, not overcome, to be understood. Ricoeur argues that living metaphors create new interpretations that may potentially transform the way we understand and act.

In his three-volume *Time and Narrative* (1983, 1984, and 1985), Ricoeur continues to develop the themes of semantic innovation and the ability of poetic discourse to disclose new ways to see and to be in the world. The basic unit of a narrative is a plot, which unifies the elements of a story, including the reasons, motives, and actions of characters with events, accidents, and circumstances together into a coherent unity. A plot synthesizes, integrates, and schematizes actions, events, and, ultimately, time into a unified whole that says something new and different than the sum of its parts. The thesis of *Time and Narrative* is that connection exists between the temporal character of human experience and the act of narrating a story. Temporal experience is expressed in the form of a narrative, as a narrative is able to reflect our social reality because it expresses temporal experience.

The circularity of time and narrative is mediated by three senses of representation: *mimesis*$_1$, *mimesis*$_2$, and *mimesis*$_3$. *Mimesis*$_1$, or prefiguration, represents the aspect

of the imitation of action that draws on our pre-understanding of the difference between human action and physical activity. It has three aspects—structural, symbolic, and temporal—that form the cognitive and practical background that determines how we interpret human action. *Mimesis*$_2$, or configuration, is the pivot of the analysis of the relationship between time and narrative in which actions are configured into a story by means of a plot. An action becomes an event in relation to a plot of a story. In turn, a story is more than a succession of actions but rather the organization of events into an intelligible whole. *Mimesis*$_3$, or refiguration, refers to the act of reading that changes our practical understanding according to the configuration of the story. The act of telling and interpreting stories links narration with the practical transformation of the world. Hermeneutics for Ricoeur is now construed as the telling, writing, and understanding of fictional and nonfictional stories, in effect, linking time, narratives, and history.

In *Time and Narrative*, Ricoeur introduces the idea of a narrative identity. His thesis is that we understand a person's identity as we would a character in a fictional or historical narrative. One's self-identity is constituted by means of *emplotment*, which configures and synthesizes diverse and multiple elements of a life into a unified whole. Just as the story of a life unfolds like a narrative, the identity of a character also unfolds in a narrative. One's identity is constituted by the stories told about oneself, as well as the stories told by other significant figures in our lives such as parents, spouses, friends, and enemies. A personal identity is also tied to larger group identities, which, similar to a personal identity, are partly chosen, partly inherited, and constituted by the stories we tell about it. The identity of a group, culture, or nation requires that its members are convinced of the truth and rightness of their story. To be effective, these narratives have to shape how the members understand themselves as a part of the group. Ricoeur is particularly interested by stories of founding events that establish and sustain communities, and form our individual and group identities.

ETHICS AND POLITICS

Ricoeur is also interested in the role narratives play in moral deliberation. The ethical implications for personal or narrative identity is that an agent must maintain some kind of continuity of time in order to be accountable for one's actions. Identity is constitutive of accountability; narrative is constitutive of identity. In addition to attributing actions to agents, narrative discourse also attributes moral obligations to agents who have the power to act and who are capable of being acted upon. Narration further mediates between description and prescription by providing a context and characters in ethical questioning. We tell stories as a part of the thought experiments we conduct, which allow us to test moral judgments in imaginary cases. Narration thus forms not only our moral ideals but also the stories we tell of ourselves and each other that help us determine if we have achieved it. How we individually and collectively remember events is crucial to the way we hold others accountable for their actions.

The fourth set of studies in *Oneself As Another* (1990), following the studies on speaking, acting, and narrating, form what Ricoeur ironically calls his "little ethics," an ambitious attempt to mediate between an Aristotelian, teleological conception of the ethical aim and a Kantian, deontological conception of the moral norm. Ricoeur's notion of practical wisdom incorporates the idea from the Aristotelian heritage that ethics is the practice of becoming a good person as a member of a political community. The good life consists in developing the virtues, habits, and practices that enable us to develop ourselves, sustain interpersonal relationships, and create a life of happiness together.

Ricoeur also incorporates the idea from the Kantian heritage that morality is defined by the obligation to respect universal moral norms. In this tradition, moral actions must be motivated solely out of duty to the moral law. Morality consists in obeying moral laws that are binding on everyone, respecting the dignity of other people, and acting as an autonomous member of a moral community. For Ricoeur, practical wisdom is the art of mediating the particular requirement of the ethical aim and the universal requirement of the moral norm geared toward acting appropriately and justly in order to achieve happiness with others in a good and just society.

In *Oneself As Another*, Ricoeur proposes three theses with respect to ethics and morality: 1) the primacy of ethics over morality; 2) the necessity that the ethical aim be mediated by the moral norm; 3) morality must seek recourse in ethics to resolve conflicts and aporias. Ethics encompasses morality—but while it is subordinate to ethics, morality is a necessary, deontological moment of the actualization of ethics. The final recourse to ethics (informed by morality) is a form of practical wisdom geared toward the appropriate application of universal norms in particular situations. The reason why ethics needs morality is to ensure that ethical life respects the autonomy and dignity of everyone.

The reason why morality needs ethics is twofold: 1) Without ethics morality would be empty; it is founded on and presupposes our desire to live well together with others; 2) When deontological norms produce conflicting obligation—as they inevitably do—we must refer back to the ethical aim of a particular good life in order to figure out what to do. Sometimes there is no right answer to moral problems. If moral judgment were simply a matter of balancing the ethical aim and moral norm, there would be no room for the tragedy of action, exemplified in stories similar to that of Sophocles's *Antigone*. It is in these intractable situations that the art of practical wisdom helps us make decisions and act justly and appropriately in the face of tragic situations.

Ricoeur's main contribution to political philosophy is his notion of the political paradox. He maintains that, on one hand, political authority is legitimate if it comes from the rational consent of the governed; on the other hand, political practice is often coercive, even violent, which is something, in principle, to which individuals cannot consent. The paradox of political authority is permanent. Ricoeur agrees with Hannah Arendt that it is necessary to distinguish legitimate power-in-common from illegitimate power-over, but he agrees with Max Weber that political institutions are in fact often characterized by domination. Consequently, we should recognize that political power and political discourse always teeters at the edge of violence and illegitimacy. The political sphere is a fragile balance between authority and force, reason and tradition, ideology and utopia.

In *The Just* (2000) Ricoeur argues that coping with political power is an exercise in practical wisdom, a mediation of our desire to live together in communities with the requirement of justice and the rule of law. Social justice not only requires democratic political and economic institutions that respect human rights, treat people equally, protect our liberties, and allow for full political participation, but it should aim to foster a good life for communities, emphasize the membership of citizens for whom political participation matters, and recognize the plurality of social goods and historic values that make us who we are.

Ricoeur's more recent work, *Memory, History, and Forgetting* (2004), examines the role memories play in our ability to represent the past and make present something that is absent. The first part is a meticulous phenomenology of memory, examining the object of memory, the act of remembering, and the nature of personal and collective memories. The second part is an epistemology of history that examines the documentary phase of archiving eyewitnesses, the explanatory phase where historical explanations occur, and the representative phase where history takes its written or literary form. The third part is a hermeneutics of our historical conditions that examines the limits to our historical knowledge, the existential and temporal conditions of our historical knowledge, and the role of forgetting in relation to memory and history.

The work concludes with a plea for forgiveness as the best way to remember events in order to right past wrongs and to restore social bonds. Individual and groups must learn to remember events differently if they wish to achieve recognition and potentially reconcile. The political implications of memory involve policy considerations for the just allotment of memory to redress excesses of both memory and forgetting. Forgiveness, however, goes beyond justice and approaches the realms of charity and gift-giving. To ask for forgiveness is to recognize that a crime may be unforgivable. Yet, Ricoeur maintains that forgiveness is the best way to remember events to permit more hopeful futures together.

In the 1950s, Ricoeur rivaled Sartre in popularity in France. By the end of the 1960s his popularity waned, along with other phenomenologists, as a new generation of intellectuals dominated the French scene. Ricoeur spent much of the 1970s in the United States, writing in English and assimilating the work of Anglo-American philosophers. His French readers, however, were perplexed by his turn to analytic philosophy and, although his works continued to be read by theologians, his reputation suffered among philosophers. However, with the publication of the highly acclaimed *Time and Narrative* in the 1980s, Ricoeur was again recognized as among France's leading intellectuals. By then he had outlived and surpassed the generation of postmodernist philosophers from the late 1960s, taking his own creative, literary turn without ever abandoning his conviction that philosophy and reason are synonymous. Over his career, Ricoeur received honorary degrees from approximately forty universities, delivered the prestigious Gifford Lectures in 1986, and was awarded numerous significant international prizes, including the 2004 John W. Kluge Prize for Lifetime Achievement in the Human Sciences. After his retirement, he resided outside of Paris and continued to publish well into his nineties.

See also Continental Philosophy; Ethics; Hermeneutics; Metaphysics; Philosophy of Language; Philosophy of Religion; Social and Political Philosophy.

Bibliography

PRIMARY WORKS

Gabriel Marcel and Karl Jaspers. Philosophie du mystère et philosophie du paradoxe. Paris: Temps Present, 1948.

Freedom and Nature: The Voluntary and The Involuntary (1950). Translated by Erazim Kohak. Evanston, IL: Northwestern University Press, 1966.

History and Truth (1955). Translated by Charles A. Kelbley. Evanston, IL: Northwestern University press, 1965.

Fallible Man (1960). Translated by Walter J. Lowe. New York: Fordham University Press, 1986.

Freud and Philosophy: An Essay on Interpretation (1965). Translated by Denis Savage. New Haven, CT: Yale University Press, 1970.

The Symbolism of Evil (1960). Translated by Emerson Buchanan. New York: Harper and Row, 1967.

Husserl: An Analysis of His Phenomenology. Translated by Edward E. Ballard and Lester G. Embree. Evanston, IL: Northwestern University Press, 1967.

The Conflict of Interpretations: Essays in Hermeneutics (1969). Translated by Willis Domingo. Evanston, IL: Northwestern University Press, 1974.

Political and Social Essays. Translated by Donald Stewart. Athens: Ohio University Press, 1974.

The Rule of Metaphor: Multi-Disciplinary Studies in The Creation of Meaning in Language (1975). Translated by Robert Czerny with Kathleen McLaughlin and John Costello. London: Routledge and Kegan Paul, 1978.

Interpretation Theory: Discourse and The Surplus of Meaning. Fort Worth: Texas Christian Press, 1976.

The Philosophy of Paul Ricoeur: An Anthology of His Work, edited by Charles E. Reagan and David Stewart. Boston: Beacon Press, 1978.

Hermeneutics and The Human Sciences: Essays on Language, Action, and Interpretation. Translated by John B. Thompson. Cambridge, U.K.: Cambridge University Press, 1981.

Time and Narrative. 3 vols. (1983, 1984, and 1985). Translated by Kathleen McLaughlin and David Pellauer. Chicago: University of Chicago Press, 1984, 1985, and 1988.

Lectures on Ideology and Utopia. Translated by George H. Taylor. New York: Columbia University Press, 1985.

From Text to Action: Essays in Hermeneutics II (1986). Translated by Kathleen Blamey and John B. Thompson. Evanston, IL: Northwestern University Press, 1991.

À l'école de la philosophie. Paris: J. Vrin, 1986.

Le mal: Un défi à la philosophie et à la théologie. Geneva: Labor et Fides, 1986.

Oneself as Another (1990). Translated by Kathleen Blamey. Chicago: University of Chicago Press, 1992.

A Ricoeur Reader: Reflection and Imagination, edited by Mario J. Valdes. Toronto: University of Toronto Press, 1991.

Lectures I: Autour du politique. Paris: Seuil, 1991.

Lectures II: La contrée des philosophes. Paris: Seuil, 1992.

Lectures III: Aux frontières de la philosophie. Paris: Seuil, 1994.

The Just (1995). Translated by David Pellauer. Chicago: University of Chicago Press, 2000.

Thinking Biblically: Exegetical and Hermeneutical Studies. Chicago: University of Chicago Press, 1998.

Critique and Conviction (1995). Translated by Kathleen Blamey. New York: Columbia University Press, 1998.

What Makes Us Think? A Neuroscientist and A Philosopher Argue About Ethics, Human Nature, and The Brain. Translated by M. B. DeBevoise. Princeton, NJ: Princeton University Press, 2000.

La mémoire, l'histoire, l'oubli. Paris: Seuil, 2000.

Le Juste II. Paris: Esprit, 2001.

Parcours de la reconnaissance. Paris: Stock, 2004.

Sur la traduction. Paris: Bayard, 2004.

SECONDARY WORKS

Abel, Olivier. *Paul Ricoeur: la promise et la régle*. Paris: Michalon, 1996.

Bourgeois, Patrick. *Extension of Ricoeur's Hermeneutic*. The Hague: Martinus Nijhoff, 1975.

Cohen, Richard A. and James L. Marsh, ed. *Ricoeur As Another: The Ethics of Subjectivity*. Albany: State University of New York Press, 2002.

Dauenhauer, Bernard. *Paul Ricoeur: The Promise and Risk of Politics*. Lanham, MD: Rowman & Littlefield Publishers, 1998.

Dosse, Francois. *Paul Ricoeur: Les sens l'une vie*. Paris: La Découverte, 1997.

Greisch, Jean. *Paul Ricoeur: L'intiniraire du sens*. Paris: Jérome Millôn, 2001.

Hahn, Lewis Edwin, ed. *The Philosophy of Paul Ricoeur*. Chicago: Open Court, 1987.

Ihde, Don. *Hermeneutic Phenomenology: The Philosophy of Paul Ricoeur*. Evanston, IL: Northwestern University Press, 1971.

Jervolino, Domenico. *The Cogito and Hermeneutics: The Question of The Subject in Ricoeur*. Dordrecht; Boston: Kluwer Academic Publishers, 1990.

Kaplan, David M. *Ricoeur's Critical Theory*. Albany: State University of New York Press, 2003.

Kearney, Richard. *On Paul Ricoeur: The Owl of Minerva*. Burlington, VT: Ashgate Publishers, 2004.

Klemm, David E. *The Hermeneutical Theory of Paul Ricoeur*. Lewisberg, PA: Bucknell University Press, 1983.

Leeuwen, T.M. van.*The Surplus of Meaning: Ontology and Eschatology in The Philosophy of Paul Ricoeur*. Amsterdam: Rodopi, 1981.

Reagan, Charles E. *Paul Ricoeur: His Life and His Work*. Chicago: University of Chicago Press, 1996.

Schweiker, William ed. *Paul Ricoeur and Contemporary Moral Thought*. New York: Routledge, 2003.

Thompson, John B. *Critical Hermeneutics: A Study in The Thought of Paul Ricoeur and Jürgen Habermas*. Cambridge, U.K., and New York: Cambridge University Press, 1981.

Van den Hengel, J. *The Home of Meaning: The Hermeneutics of The Subject of Paul Ricoeur*.Washington, DC: University Press of America, 1982.

Venema, Henry Isaac. *Identifying Selfhood: Imagination, Narrative, and Hermeneutics in The Thought of Paul Ricoeur*. Albany: State University of New York Press, 2000.

Wood, David ed. *On Paul Ricoeur: Narrative and Interpretation*. New York: Routledge, 1991.

David M. Kaplan (2005)

RIEHL, ALOIS
(1844–1924)

Alois Riehl, the Austrian neo-Kantian philosopher, was born in Bolzano. Riehl was consecutively Privatdozent (1870), extraordinary professor (1877), and professor (1878) at the University of Graz. He moved to the University of Freiburg in 1882, to Kiel in 1895, to Halle in 1898, and to Berlin in 1905.

Riehl's first philosophy was a realistic metaphysics based on Johann Friedrich Herbart and indirectly on Gottfried Wilhelm Leibniz, and it is of interest, just as in the case of Immanuel Kant, to study the relation between Riehl's precritical and critical writings. Between 1870 and 1872 Riehl made his first realistic, monistic, evolutionist decisions within that dogmatic framework. His *Realistische Grundzüge* (Graz, 1870) centered on the problem of sensation, which he originally conceived as a polycentric reciprocal matrix of consciousness and movement. In *Über Begriff und Form der Philosophie* (Berlin, 1872) he advocated a critical, rational requirement and the scientific character of philosophy, to which he assigned the historical task of leading to ideal ends. In *Moral und Dogma* (Vienna, 1872) he defended the independence of positive morality from beliefs.

A profound study of Kant freed Riehl from his metaphysical dogmatism. The first volume of his *Der philosophische Kritizismus* (1876) marked an important date in the history of the new Kantianism. This work highlighted the hold on Kant of the spirit of the new positive science (not so much through the influence of René Descartes as through that of John Locke and David Hume). Combating psychological and idealistic "misconceptions" of Kant's views, Riehl proposed that the evolution of Kant's thought be studied, and successive editions of *Der philosophische Kritizismus* benefited from previously unpublished writings of Kant discovered by Kant philologists. Kant, according to Riehl, clarified the method of philosophy; in abandoning metaphysics but not identifying itself with science, philosophy shows itself to be theory of knowledge and the methodology of the natural sciences. It is false, however, to eliminate the thing-in-itself and the presupposition of realism common to the sciences, as Hermann Cohen did. Kant distinguished form from content and sought to determine the formal a priori of nature in general and not the particular laws of nature evident in the real experience of the sciences.

In the second and third volumes of *Der philosophische Kritizismus* (1879 and 1887) Riehl reassessed and amplified his own views. It was not easy: To do so he had to fight with Kant himself (whom Eugen Dühring had blamed for having "two centers of gravity"), even reduced to the first *Critique* alone. In Riehl's view, neither dogmatic realism nor idealism, whether phenomenalist, or absolute, or positivistic, was adequate. Riehl sought to bring Kant up to date concerning the "sensible and logical foundation of knowledge" by surveying the great scientific innovations since Kant's day, such as Robert Mayer's principle of the conservation of energy and the Darwinian theory of evolution. Only then could Riehl critically resume his own realistic monism centered on perception. But perception, the first cognition, is not, in Riehl's judgment, the first reality. The two aspects of perception—the mechanical, which can be made objective and is quantitatively determinable by positive science, and the qualitative, which is subjectively immediate and the sole revealer of the real universal reciprocity—are both phenomenon (*Erscheinung*), although not merely appearance (*Schein*); neither of the two aspects makes up "nature in itself." The monistic propensity, leading to the threshold of metaphysics, comes upon reefs the critique must steer clear of. For example, he desires that his identification of the physical and the psychical should not be confused with materialism, or monadism, or universal psychophysical correspondence, or Spinozistic panpsychism. Again, although Riehl saw mental life as a product of natural evolution, he denied the evolutionary genesis of logical and mathematical concepts.

In 1883, in his inaugural lecture at Freiburg, "Über wissenschaftliche und nichtwissenschaftliche Philosophie," Riehl turned to other fields of philosophy with a progressive valuation (compare a lecture at Princeton, 1913: "Der Beruf der Philosophie in der Gegenwart"). Even in *Der philosophische Kritizismus,* confined to the naturalistic horizon, he had apologized for glancing at "the field of practical philosophy" but had at the end intimated that beyond the realm of science lay the realms of moral action and artistic production (to which he later added religion). It may be asked whether there could be a philosophy of these things if "theoretical" is identical with "scientific" (*wissenschaftliche*). In his later years Riehl struggled with this problem, surrounded too by the other neo-Kantian movements. "Feeling," which he had acknowledged as another side of experience, might be available for that theoretical purpose, but to be so available its evaluations must be freed from practical empiricism. Heinrich Rickert, who had frequent contact with Riehl, later sought to show Riehl's increasing interest in the world of values, until Riehl finally acknowledged that

the role of philosophy is "to raise to conceptual clarity our knowledge of values and their system."

See also Cohen, Hermann; Descartes, René; Dühring, Eugen Karl; Herbart, Johann Friedrich; Hume, David; Idealism; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Neo-Kantianism; Realism; Rickert, Heinrich; Value and Valuation.

Bibliography

WORKS BY RIEHL

Der philosophische Kritizismus, 3 vols. Leipzig, 1876–1887; 2nd ed, Leipzig: W. Engelmahn, 1908–1926; 3rd ed., Vol. I, Leipzig: Kröner, 1924.

Friedrich Nietzsche: Der Künstler und der Denker. Stuttgart, 1897.

Zur Einführung in die Philosophic der Gegenwart. Leipzig: Teubner, 1903.

"Helmholtz in seinem Verhältnis zu Kant." *Kantstudien* 9 (1904): 261–285.

"Logik und Erkenntnistheorie." *Kultur der Gegenwart* 1 (1907): 73–102.

Philosophische Studien aus vier Jahrzehnten. Leipzig, 1925. Essays.

WORKS ON RIEHL

Hofmann, Paul. "Riehls Kritizismus und die Probleme der Gegenwart." *Kantstudien* 31 (1926): 330–343.

Jaensch, Erich. "Alois Riehl Der Mann und das Werk." *Kantstudien* 30 (1925): III–XXXVI.

Maier, Heinrich. "Alois Riehl." *Kantstudien* 31 (1926): 563–579.

Rickert, Heinrich. "Alois Riehl." *Logos* 13 (1924–1925): 162–185.

Siegel, Carl. *Alois Riehl: Ein Beitrag zür Geschichte des Neukantianismus. Festschrift der Universität Graz.* Graz: Leuschner and Lubensky, 1932.

Spranger, Eduard. "Alois Riehl." *Forschungen und Fortschritte* 20 (June 1944): 129–130.

Mariano Campo (1967)
Translated by Robert M. Connolly

RIGHTS

Although ancient ethics used the concepts of property and justice, each of which presupposes something similar to the concept of a right, the concept of a right in the modern sense developed only later. The first philosopher to define a moral right was most likely William of Ockham (c. 1285–1347), who noted that *jus* sometimes refers to the power to conform to right reason. Thus, he integrated the legal concept of dominium or property into the moral theory that the law of nature determines right action. Hugo Grotius (1583–1645) adopted the resulting theory of natural rights—rights conferred by the law of nature—and made it the basis for his theory of international law. Hobbes and Locke used their conceptions of natural rights to explain the grounds and limits of political obligation. Hobbes (1588–1679) conceives of a right as a liberty of action that is the absence of any contrary obligation. Because the law of nature requires only that one seek peace, in a state of nature where there is no peace one has a natural right to do anything one desires. He infers that citizens can have a duty to obey the sovereign only if they give up most of their natural rights.

Locke (1632–1704), however, uses his theory of inalienable natural rights to limit the authority of the sovereign over the members of a society. He conceives of natural rights in the traditional way as powers of acting in conformity with the natural law and assumes that the law of nature also imposes obligations upon others not to prevent one from exercising these rights. The Lockean theory of inalienable and imprescriptible natural rights inspired the American *Declaration of Independence*, the French *Declaration of the Rights of Man and the Citizen*, and the subsequent development of constitutional law in much of Europe and North America.

Jeremy Bentham (1748–1832) subjected the doctrine of natural rights to severe criticism. The notion that there are natural rights—rights not created by human legislation—is conceptual nonsense because a right of one party implies a corresponding obligation of some second party, and an obligation exists only when commanded by some law enforced with coercive sanctions. Granting the existence of natural rights is morally perverse, for one could appeal to natural rights to justify any arbitrary action because without any law to define their content they would be indeterminate. Finally, the doctrine of inalienable and imprescriptible natural rights is politically dangerous because it would justify anarchy, for the individual's absolute natural right to liberty or the pursuit of happiness could neither be given up by the citizen nor extinguished by coercive legislation. After the 1950s, few moral philosophers adopted a traditional natural rights theory because of skepticism about both the existence of a natural law expressing the will of God or a Cosmic Reason as well as one's ability to know what it might command or forbid by the natural light of reason.

JURISTIC THEORIES

Skeptical as one may be of natural rights, one can hardly doubt the existence of legal rights. A central concern of modern jurisprudence has been to explain the nature of these rights. Although juristic theories are primarily the-

ories about the nature of legal rights, they can be extended to moral rights if one conceives of morality in terms of something like the moral law. John Austin held that a legal right is equivalent to a relative legal duty. Legal duties are imposed upon persons by laws that command some act or forbearance and constitute obligations because they are enforced by sanctions in the event of disobedience. Although some duties, such as the duty to pay one's taxes, are absolute in the sense that they are owed to society in general and not to any assignable individual, a relative duty is owed to one or more determinate second parties. Thus, the creditor's contractual right to be repaid by the debtor is simply the debtor's duty of repayment owed to the creditor viewed from the latter's point of view.

John Salmond (1920) developed the view of Rudolph von Jhering that a legal right is a legally protected interest. The object of any right is the thing—not necessarily a material object—in which the right holder has an interest. Not all of one's interests constitute legal rights, however. Only those protected by the law. For example, the object of one's legal right not to be killed is one's life, probably the most fundamental interest of every individual. This interest is protected by laws that prohibit killing and that punish murderers. Paul Vinogradoff (1928) maintained that there are three elements of any legal right. (1) A legal right must be claimed by some individual or state; (2) this claim must be recognized by organized society as justified from the public point of view; and (3), this declaration of right must be enforced by the legal authorities. Hence, a legal right is a legally protected claim. One's right not to be injured, and, if injured, to sue for damages, illustrates this conception of rights. Jhering and Vinogradoff set the stage for the ongoing contest between interest and will theories of rights. The former hold that the essential function of rights is to protect some interest of the right holder; the latter insist that the function of rights is to give the will of the right holder some privileged legal status.

Karl Llewellyn (1962) defined a right as a future judicial remedy. To say that Jones has a property right to his car is simply to predict how the courts will decide any cases concerning that car. For example, if someone steals the car, a court will hold that the vehicle shall be returned to Jones. Or if someone damages that car and Jones sues for damages, a court will decide in favor of Jones. This theory of rights reflects the rule skepticism of American legal realists. They argued that it is a mistake to identify the law with the general principles or rules written in the law books. What these general statements mean is left open to the interpretation of the courts and may or may not be applied in practice depending upon how judges choose to decide particular cases. Hence, what the law really is consists of the decisions made by the courts.

Scandinavian legal realists are not rule skeptics. In fact, Alf Ross (1957) identified a legal right with a set of legal rules such that any one of a number of facts legally implies all of a variety of legal consequences. For example, if someone has purchased a thing *or* if someone has inherited a thing *or* if someone has earned a thing, then if another person steals that thing it shall be returned to the person who purchased or inherited or earned it, *and* if a second party damages that thing the second party shall compensate the person who purchased or inherited or earned that thing, and so on. What the legal right to ownership really amounts to is simply the preceding set of legal rules. Thus, the language of rights is a convenient technique of summing up a complex set of connections in the law. If legal rights seem to be some mysterious sort of ideal entities, this is merely because the language of rights has an emotive meaning that seems to give rights a magical power. This theory of rights reflects the concern of Scandinavian legal realists to reject the metaphysical idealism prevalent in continental legal philosophy.

Wesley Newcomb Hohfeld (1919) examined the writings of judges and jurists and concluded that they use "a right" indiscriminately to express any of four very different fundamental legal conceptions. This ambiguity suggests invalid legal reasoning because claims, liberties, powers, and immunities are different legal relations. A legal claim of X against Y is logically correlative with a corresponding legal duty of Y to X. For example, X's legal right not to be struck by Y is the logical correlative of Y's legal duty not to strike X. A legal liberty or privilege of X in the face of Y is simply the absence of any corresponding duty of X to Y. Thus to say that X has a legal liberty to phone Y after midnight is to deny that X has a legal duty not to phone Y after midnight. A legal power of X over Y is the ability to change some legal position of Y by some voluntary action of X. For example, X's power to give Y a book is X's ability to confer ownership of that book upon Y by handing it to Y and saying "I hereby give this book to you." A legal immunity of X against Y regarding some legal consequence C is Y's lack of legal power over X regarding C. Thus, X has a legal immunity against Y that Y not extinguish X's ownership of a book by Y's act of saying "I hereby take ownership of the book away from X." Hohfeld argued that in the strict sense, only legal claims are rights because an essential feature of the concept of a right is that rights and duties are logically cor-

relative. Every right implies a duty with a corresponding content, and every duty implies a corresponding right. He admitted that lawyers often speak of multiple sets of claims as a legal right, but insisted that clarity and precision require conceiving of a legal right as a single legal claim of one individual against one second party.

CONCEPTUAL ANALYSES

Joel Feinberg (1980) agreed with Hohfeld that rights are claims, but denied that rights and duties are logically correlative. Although every right implies some duty, not every duty implies any corresponding right. For example, one's legal duty to obey the orders of a police officer is imposed by the impersonal law and not owed to the officer. Similarly one's moral duty to sacrifice some of one's wealth to assist those in need does not imply any right of this or that needy individual to one's charity. Even when a right does imply some corresponding duty, it is a mistake to reduce the right to that duty. What is distinctive and most valuable about rights is that they put one in a position to claim and to demand—and not merely request or beg—performance of the duty owed to one. What confers this status of claimant upon the right holder is some set of rules—legal rules in the case of legal rights and moral principles for moral rights. Hence, a right is a valid claim, a claim justified by some appropriate set of rules.

H. J. McCloskey (1959, 1979) denied that rights are claims *against* and argued that they are entitlements *to* do, have, enjoy, or have done. Having purchased a car and obtained a driving license, one has a legal right to drive one's car. This is a right to do something—to drive on public thoroughfares; it is not primarily a claim against policemen and magistrates not to interfere. One possesses and exercises rights; one makes claims but does not possess or exercise them. Although one's moral right to life gives rise to duties of others not to kill one, it is primarily a right to live and preserve one's life. A hermit's right to life is the hermit's right to do whatever is necessary to sustain the hermit's life—including killing and eating animals—although there is no one else on or near the isolated island against whom the hermit could possibly claim the right to do so. Admittedly, the creditor's right to be repaid does hold against the debtor, but this is a special sort of right and not typical of rights in general.

H. L. A. Hart (1982) agreed with Hohfeld that "a legal right" is used to refer to four very different legal relations, but did not conclude that this makes the expression ambiguous. He explained what liberty-rights, claim-rights, power-rights, and immunity-rights have in common. They all consist of one or more bilateral liberties protected by a perimeter of duties. For example, at the center of one's liberty-right to look over one's garden fence at one's neighbor is one's legal liberty either to look over one's fence at one's neighbor or not to do so. This right does not impose upon one's neighbor any logically correlative legal duty to allow herself to be looked at; she is legally permitted to erect a higher fence or hide behind a screen. Still, this bilateral liberty is protected by a number of duties against interference. One's neighbor has legal duties not to climb over one's fence and assault one or to blind one with a chemical spray. What distinguishes one species of rights from another is the kind of bilateral liberties at their center. Thus, central to one's power-right to contract is one's liberty either to exercise one's legal power to accept an offer or to refrain from accepting it. And central to the creditor's claim-right to be repaid are the liberties to cancel or refuse to cancel the debt and, in the event of nonpayment, to sue or refrain from suing for payment. Thus, a legal right is an individual choice respected by the law. Presumably Hart thought of a moral right as an individual choice respected by the rules of morality. His view that moral rights concern the proper distribution of freedom strongly suggests some such theory.

Hart's respected choice theory of rights is a will theory of rights, but it is a mistake to assume that all will theories are option theories. Feinberg's claim theory of rights is also a will theory because he argues that to have a right is to be in a position to make a claim, to demand performance of some corresponding duty. Yet, his theory does not place any bilateral liberty at the center of every right. He even recognizes mandatory legal rights, such as the right to vote in Australia—where voting is a legal duty—so that one does not have any legal liberty to refrain from voting. As one would expect, Hart argued against interest theories of rights. For one thing, to hold that a right consists in an interest protected by a duty reduces rights to duties that benefit some second party. This renders the concept of a right redundant, for one can say everything one needs to say in the language of beneficial duties. But what is distinctive of rights correlative with duties are powers such as the right holder's power to cancel or enforce performance of that duty. Also, when a right benefits some third party, the right holder is not the party whose interest is protected by the law. Thus, when a parent purchases life insurance and names a child as beneficiary, it is the child's interest that is protected by the insurance company's legal duty to pay, but the right holder is the parent rather than the child.

Carl Wellman (1985, 1995) agreed with Hart that what is distinctive and important about rights is the way in which they allocate freedom and control upon the right holder. He defined a right as a complex of Hohfeldian positions that, if respected, confer dominion over some defining core upon the right holder in the face of one or more second parties. For example, at the core of the creditor's legal right to repayment is the creditor's legal claim against the debtor that the debtor repay the contracted amount at or before the due date. But Hohfeld was mistaken in identifying this legal right with a single legal claim. The creditor's claim would not hold against the debtor unless the right holder also had a legal immunity against the debtor's extinguishing the claim merely by saying "I hereby cancel my debt to you." And it also includes additional associated legal positions, such as the power to sue for repayment in the face of a recalcitrant debtor, the power to cancel the debt if one so chooses, and the legal liberties of exercising these powers. Although this is a modified version of Hart's will theory of rights, Wellman did not put a legal liberty—much less a liberty to choose—at the core of every right. The defining core of a legal right can be a legal claim, power, immunity, or even a liability. He also extended his dominion theory of rights to moral rights by arguing that there are moral liberties, claims, powers and immunities analogous to the legal relations Hohfeld identified.

Joseph Raz (1986) defined rights by their role in practical reasoning rather than in terms of Hohfeld's fundamental legal conceptions. It is a mistake to identify rights with interests, even protected interests, because rights serve as intermediate reasons linking interests to duties. At the same time, a right cannot be reduced to some correlative duty because a right is logically prior to any duty, and a single right can imply more than one duty or various duties under varying circumstances. To say that someone has a right is to say that—other things being equal—an aspect of an individual's well-being (one of that individual's interests) is a sufficient reason for holding some other person or persons to be under a duty. However, what makes one's interest of sufficient importance to ground duties need not be merely the value of that interest to oneself. For example, Abel's right to free speech is based on Abel's interest in speaking freely together with the public interest in allowing citizens to speak their minds without unjustified restrictions. Thus, rights are interest-based reasons for duties. This is an interest theory of rights, but more complex than the traditional protected interest theories of Jhering and Salmond.

Rex Martin (1993) rejected Feinberg's view that rights are valid claims. For one thing, not every right implies some corresponding duty as claim-rights do. The logical correlative of the constitutional right to free speech is a disability of Congress to enact statutes limiting speech. More importantly, moral or legal justification is not sufficient to establish a claim as a right. A slave's claim to freedom, no matter how thoroughly justified by moral principles, would be infirm as a right in any society where others could disregard it with impunity. Even if the slave's claim to freedom were justified by the legal rules of that society, it would be merely a nominal right if public officials, including judges, failed or refused to act in accordance with these rules. Real rights—moral rights as well as legal rights—presuppose the social practices of recognition and maintenance. Thus, rights are established ways of acting or being treated— for example, the civil right to the free exercise of one's religion or the moral right to be rescued from imminent danger. Because rights must be established by social practices, no right can exist independently of the institutions of a society. Thus, human rights are best understood as morally justified civil rights.

Judith Jarvis Thomson (1990) doubted that one can base a theory of rights on a definition of "a right" or general description of the nature of rights. Therefore, she used a conceptual analysis different from those previously described. She suggested that to attribute a right might be to talk about permissible and impermissible actions, but in a way which groups them to bring whole clusters of cases to bear on each other. Hence, to learn what the moral or legal significance of having some right is, one must discover the moral or legal consequences for the right holder and others of that right. For example, Jill has a moral right that Jack not break her nose. This implies—other things being equal—at least that Jack ought not to break Jill's nose, that it is morally permissible for Jill to defend herself against any attempt by Jack to break her nose, and that if Jack does break Jill's nose he ought to pay her medical expenses. Thus, Thomson analyzed rights in terms of what the right holder and others may or ought to do. This is an analysis resting upon the judgments of particular cases, not one derived from general principles.

POSSIBLE POSSESSORS

A conceptual analysis of rights usually implies something about the necessary conditions for the possession of any right. Thus, Hart's respected choice conception of rights implies that it is idle and misleading to ascribe rights to young children who have not yet developed the ability to

choose. Neil MacCormick (1982) suggested that children's rights are a test case for any theory of rights. He thought it clear that these children do have moral rights, including the rights to be nurtured, cared for, and if possible, loved. Because will theories cannot explain these rights, they must be rejected. In their place, he proposed a protected interest theory of rights. Because even neonates do have interests, this theory can explain how it is possible for them to be right holders.

Feinberg's theory also seems to imply that very young children could not be right holders. He maintained that to have a right is to be in a position to make a claim, to demand something as one's due. Infants seem incapable of claiming in this performative sense. But Feinberg believed that clearly even wee babies do make claims, not in their own persons but through parents or guardians who act as their representatives. These representatives are claiming on their behalf, acting in their interests. Because children do have interests from the day they are born, they are capable of being represented and, therefore, are possible holders of legal and moral rights. Thus, Feinberg combined a will theory of the nature of rights with an interest theory of possible right holders.

Feinberg extended his theory of possible right holders to the more controversial debate about animal rights. Because many non-human animals do have interests in food, shelter, and freedom from pain, human beings can represent them and make claims on their behalf. Hence, animals are also possible possessors of rights. R. G. Frey (1980) challenged the assumption that animals can have interests in the relevant sense. One can take an interest in something only if one can desire or want that thing, and this requires that one believe something about that thing. But because animals lack any language adequate for believing, they cannot possibly have the desires or wants presupposed by the interest theory of possible right holders.

H. J. McCloskey denied that it is the capacity to have interests that makes one a possible right holder because one may, on moral grounds, choose to exercise one's rights contrary to one's interests. The notion of exercising—acting on the basis of—one's rights is central to the concept of a right. And to exercise or refrain from exercising some moral right requires that one make a moral choice. Hence, it is the capacity for moral autonomy, for self-direction, and self-determination, that is required for the possession of rights. Because animals lack this capacity, they are not possible right holders. Robert Elliot (1987) agreed that the capacity to exercise one's rights is necessary for the possession of rights, yet denied that this requires full moral autonomy—the ability to consider moral reasons and choose on those grounds. A human being could exercise one's moral right to self-defense simply by unreflectively defending oneself against an attacker. An animal could do the same. Hence, animals are capable of acting in a sense robust enough to enable them to possess rights.

Tom Regan (1983) accepted the Kantian view that human beings have moral rights because of their inherent value, but argued that it is arbitrary to restrict inherent value to moral agents. It is being the subject-of-a-life—having a life that goes better or worse for one—that confers inherent and not merely instrumental value upon one. Because at least the higher animals are also subjects-of-a-life, they are also moral right holders.

Are human fetuses capable of possessing human rights, including the right to life? Mary Anne Warren (1973) granted that unborn children are human in the genetic sense of being members of the same biological species as adult human beings, but denied that this is relevant to whether they are members of the moral community of right holders. It is because normal adult human beings are persons that they possess moral rights. The traits that are central to personhood are consciousness, rationality, self-motivated action, the capacity to communicate, and self-awareness. Although it may be uncertain how many of these and in what degree are required for personhood, it is clear that fetuses possess few of these traits in any significant degree. Therefore, a human fetus cannot possess any significant right to life.

Advocates of the right to life often argue that the human fetus has the capacity to develop into an adult person; this potentiality gives it the moral right to life. The standard reply to this argument is that the potentiality to become a person implies only the capacity to acquire rights in the future, not the capacity to possess them before birth. However, Francis C. Wade (1975) argued that if a kernel of seed corn has the capacity to grow into a stalk of corn, this must be because of something in the present nature of that kernel, an active tendency to grow. Similarly, the human fetus's potentiality of full humanity in the morally relevant sense is an active tendency to develop personhood already existing in the fetus. This explains how fetuses can now possess moral rights, including the human right to life.

MORAL RIGHTS

Although Bentham (1962) rejected the existence of moral rights independent of the law, John Stuart Mill (1969) defended their existence partly as moral grounds for

judging that some law is unjust. A duty is a kind of action that a person may rightfully be compelled to perform and that one may be punished for not performing. Legal duties ought to be enforced with legal sanctions; moral duties are obligations one should be compelled to perform by public opinion or the internal sanction of one's conscience. A moral right is the logical correlative of a relative moral duty, a duty owed to the right holder because that is the one who would be harmed by its nonperformance. Why ought society to defend one in the possession of one's moral rights? Mill's answer is that this will promote the general utility—the greatest well-being—of all the members of the society.

Ronald Dworkin (1977) argued that no utilitarian theory of rights can take moral rights seriously. Sometimes to say that someone has a right to do something is to say merely that to do so would not be to act wrongly, but to say that someone has a right in the strong sense is to assert that it would be wrong to prevent one from so acting. Thus, one may say that Jones has a moral right to spend his money gambling, although he ought to spend it in a more worthwhile way. The moral rights of the citizen against the state—such as the rights to free speech or to freedom from unreasonable searches—are worth taking seriously only if they are rights in the strong sense. If they were grounded on utility, the government would be justified in infringing them whenever it would be useful to do so. But this would undermine their moral purpose: to give the individual special protection against political interference. Therefore, a theory can take moral rights seriously only if it grounds them either on the human dignity of the individual person or on the ideal of political equality. Thus, the most fundamental moral right must be the right to equal concern and respect.

David Lyons (1994) responded that a utilitarian can take moral rights seriously. Dworkin failed to notice Mill's distinction between expediency and morality. An expedient act—one that has the best consequences—can be morally wrong because it violates a moral obligation. What makes an act a moral obligation is not its utility, but that society would be justified in imposing sanctions—either the disapproval of others or of one's own conscience—upon agents who fail or refuse to act. And enforcing moral obligations is justified by its social utility. For someone to have a moral right to something is for others to have a corresponding moral obligation at least not to injure and perhaps to promote the right holder's interest in that thing. This sort of indirect grounding of moral rights on the utility of the enforcement of the correlative duties does not imply that a right may permissibly be violated whenever it would be expedient to do so.

L. W. Sumner (1987) agreed with Bentham that there are no natural rights, but did not infer that there are no moral rights. Although he rejected Hart's will theory of rights and held that the function of rights is to protect some interest of the right holder, he adopted Hart's view that rights presuppose social practice rules. Legal rules are made and upheld by the officials in some legal system; the rules of the moral code of a society are constituted by the practices of its members. But not every conventional right has moral force. A moral right is a morally justified conventional right, either an existing conventional right that it would be morally justified to retain or one that it would be justified to introduce into the conventional morality. And what justifies moral rights is the valuable consequences of maintaining the social practice rules that confer them upon moral agents. Although moral rights are grounded upon their contribution to human welfare, they can be taken seriously because the rules that confer them often constrain the direct pursuit of social utility.

Jeffrie G. Murphy (1977) argued that there are two very different kinds of moral rights. Autonomy rights mark out the special kind of treatment required to respect the dignity of autonomous rational persons. As Kant recognized, persons are ends—and not means only—and ought not to be sacrificed or used without their consent as instruments or resources for achieving the ends of others. Autonomy rights are grounded on the inherent moral value of autonomous rational agents. But Mill recognized a different function of moral rights: to pick out those moral claims that ought to be protected by society, especially by the law. However, Murphy rejected Mill's utilitarian justification of moral rights and argued that they are grounded on a hypothetical social contract of the sort described by John Rawls (1971). Thus, an individual has a social contract right to X only if a law guaranteeing X to the individual would be unanimously chosen by rational agents who are not aware of what their special circumstances would be in their society.

HUMAN RIGHTS

Although jurists usually identify human rights with the universal human rights recognized in international law, philosophers tend to view human rights as fundamental moral rights one possesses by virtue of being human. It is this latter sort of view that is most relevant here. Gregory Vlastos (1962) defined a just act as one prescribed exclusively by regard for the rights of all those it affects substantially. Although it is often just to distribute goods

unequally according to the merit of the recipients, equalitarian justice respects the equal human rights of everyone affected. Human rights are necessarily equal because they are grounded on the equal human worth of all persons, however different their individual merits. And what gives all human beings equal worth is the equal intrinsic value of their well-being and freedom. Hence, there are two classes of human rights: rights to goods required for human well-being and rights to fundamental human freedoms.

Alan Gewirth (1982) also grounded human rights upon freedom and well-being, but by a different argument. Human rights are primarily moral claim-rights of individual human beings that entail correlative moral duties of other individuals and organizations. All human beings are actual or potential agents, and human action consists in the voluntary pursuit of goals one values. Anyone who engages in action must presuppose that one has a right to the necessary conditions of prospective purposive action. And one cannot claim this and without self-contradiction deny that all other agents have the same rights. The two necessary conditions of human action are freedom and well-being. Hence, the various human rights to basic freedoms and goods are implied by the necessary presuppositions of human action.

James Griffin (2001) argued that the best account of human rights is one that preserves—but goes beyond the insights of—the traditional natural rights theories. It conceives of human rights as protections of personhood. Personhood should be understood in terms of the various strands of agency. These are autonomy (or making one's own decisions), forming a conception of the good life and being able to pursue it, and freedom from interference from others. Hence, there are three classes of human rights: autonomy rights, welfare rights, and liberty rights. The abstract human rights grounded on personhood are made more determinate by practicalities that spell out what is necessary, given the circumstances, to protect personhood. Thus, there are two grounds of human rights: personhood and practicalities.

Contemporary human rights documents reaffirm the traditional civil and political rights, such as the rights to life, free speech, and a fair trial. Yet they also assert social and economic rights, such as the rights to work, social security, and an adequate standard of living. Maurice Cranston (1967) argued that these supposed welfare rights are not genuine human rights. If they were universal human rights, they would impose upon every society the duties to provide employment, old-age pensions, and all necessary food and medical care for all their citizens. But many societies lack the resources to provide such welfare benefits to all, and there can be no moral duty to do the impossible. Hence, there can be no human rights that would imply such duties. Civil and political rights, however, require only the appropriate legislation. James W. Nickel (1987) replied that civil and political rights also face the problem of scarce resources. No society can afford a police force adequate to secure the right to life of every citizen and the right to a fair trial can be real only where there is an expensive system of courts and adequate legal assistance for all. Moreover, there are ways to realize human welfare rights in societies with varying levels of affluence. A society can introduce programs to achieve progressively full employment or adequate medical care, prune a welfare right to achieve what is most important in it, or if necessary sacrifice more costly human rights in order to secure those that are affordable.

Cranston also insisted that a human right is something of which no one may be deprived without a grave affront to justice, but that it is not a grave injustice for human beings to lack old-age pensions or the medical care they need. Hence, social security and an adequate standard of living may be moral ideals, but they are not genuine human rights. Robert Nozick (1974) went even further and argued that there can be no basic welfare rights because their implementation would violate justice. Welfare rights would require programs such as Aid to Families with Dependent Children and Medicaid. But to fund these or similar programs, the state would have to tax the affluent and redistribute their wealth to the poor. This would be unjust because it would violate the human right to property of those whose wealth is taken from them without their consent. Nozick bases his argument on his entitlement theory of justice. John Rawls (1971), however, advanced a different theory of justice. According to Rawls, the morally justified principles of social justice are those that rational persons would unanimously choose were they in an original position of equality and unaware of their particular circumstances. He argued that one of these principles is that social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all. This would require some redistribution of wealth to overcome great economic inequalities in a society.

Nozick also proposed a concept of rights that seems to exclude rights to welfare benefits. A moral right is a side-constraint on the pursuit of individual or social goals. For example, the right to property makes it morally wrong for one person to become wealthy by stealing from

anyone who owns something one desires, and the right to liberty implies that it would be morally wrong for the state to suppress opposition by imprisoning its critics. Moral rights reflect the Kantian idea that individuals are inviolable because they are ends in themselves and ought not to be used to achieve the goals of others. On this view, moral rights are negative; they constrain the actions of others by imposing only negative duties not to mistreat right holders in morally impermissible ways such as injuring them or interfering with their freedom of action. But any imagined welfare right would impose positive duties to provide welfare benefits such as old-age pensions or payments for medical care to individual right holders. There can be no human rights of this sort if fundamental moral rights are side-constraints that impose only negative duties.

Henry Shue (1980) rejected the view that the traditional civil and political rights are purely negative and that social and economic rights are positive. Every basic moral right imposes three sets of duties: duties to forbear from depriving right holders of the substance of their right, duties to protect right holders against the deprivation of the substance of their right, and duties to aid right holders in obtaining or regaining the substance of any right of which they have been deprived. For example, the basic right to liberty implies that the state has the duties not to imprison innocent persons, to protect individual persons from being kidnapped, and to aid anyone who has been unjustly imprisoned or kidnapped to obtain her release. Because some of these duties are negative and others positive, it is a mistake to argue that there cannot be any welfare rights, such as the right to subsistence benefits, simply because this would impose positive duties to aid those who lack the means of subsistence. As one would expect, the philosophy of rights remains as controversial today as it has been during the past few centuries.

See also Social and Political Philosophy.

Bibliography

Austin, John. *Lectures on Jurisprudence, or the Philosophy of Positive Law.* 5th edition, edited by Robert Campbell. London: John Murray, 1885.

Bentham, Jeremy, *Anarchical Fallacies*. In *Works of Jeremy Bentham.* vol. 2, edited by John Bowring. New York: Russell & Russell, 1962.

Cranston, Maurice. "Human Rights, Real and Supposed." In *Political Theory and the Rights of Man*, edited by D. D. Raphael. Bloomington, IN: Indiana University Press, 1967.

Dworkin, Ronald. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press, 1977.

Elliot, Robert. "Moral Autonomy, Self-Determination and Animal Rights." *The Monist* 70 (1) (1987): 83–97.

Feinberg, Joel. *Rights, Justice, and the Bounds of Liberty*. Princeton, NJ: Princeton University Press, 1980.

Finnis, John. *Natural Law and Natural Rights*. Oxford: Clarendon Press, 1980.

Frey, R. G. *Interests and Rights*. Oxford: Clarendon Press, 1980.

Gewirth, Alan. *Human Rights*. Chicago: University of Chicago Press, 1982.

Glendon, Mary Ann. *Rights Talk*. New York: Free Press, 1991.

Griffin, James. "First Steps in an Account of Human Rights." *European Journal of Philosophy* 9 (3) (2001): 306–327.

Grotius, Hugo. *The Rights of War and Peace*. Translated by A. C. Campbell. New York: M. W. Dunne, 1901.

Hart, H. L. A. *Essays on Bentham*. Oxford: Clarendon Press, 1982.

Hobbes, Thomas. *Leviathan,* edited by C. B. Macpherson. Hammondsworth, U.K.: Penguin, 1968.

Hohfeld, Wesley Newcomb. *Fundamental Legal Conceptions*. New Haven, CT: Yale University Press, 1919.

Jhering, Rudolph von. *Geist des römischen Rechts*. Leipzig, Germany: Breitkopf und Hartel, 1877.

Llewellyn, Karl. *Jurisprudence*. Chicago: University of Chicago Press, 1962.

Locke, John. *Two Treatises of Government,* edited by Peter Laslett. London: Cambridge University Press, 1967.

Lomasky, Loren E. *Persons, Rights, and the Moral Community.* New York: Oxford University Press, 1987.

Lyons, David. *Rights, Welfare, and Mill's Moral Theory*. New York: Oxford University Press, 1994.

MacCormick, Neil. *Legal Rights and Social Democracy*. Oxford: Clarendon Press, 1982.

Martin, Rex. *A System of Rights*. Oxford: Clarendon Press, 1993.

McCloskey, H. J. "Moral Rights and Animals." *Inquiry* 22 (1–2) (1979): 23–54.

McCloskey, H. J. "Rights." *Philosophical Quarterly* 15 (2) (1959): 115–127.

Melden, A. I. *Rights and Persons*. Oxford: Basil Blackwell, 1977.

Mill, John Stuart. *Utilitarianism*. In *Collected Works of John Stuart Mill*, vol 10, edited by J. M. Robson. Toronto: University of Toronto Press, 1969.

Murphy, Jeffrie G. "Rights and Borderline Cases." *Arizona Law Review* 19 (1) (1977): 228–241.

Nickel, James W. *Making Sense of Human Rights*. Berkeley, CA: University of California Press, 1987.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Ockham, William of. *Opus nonaginta Dierum*. In *Opera politica*, vol. 2, edited by R. F. Bennett. Manchester, U.K.: Manchester University Press, 1963.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Raz, Joseph. *The Morality of Freedom*. Oxford: Clarendon Press, 1986.

Regan, Tom. *The Case for Animal Rights*. Berkeley, CA: University of California Press, 1983.

Ross, Alf. "Tû-Tû." *Scandinavian Studies in Law* 1 (1957): 139–153.

Salmond, John William. *Jurisprudence*, 6th edition. London: Sweet and Maxwell, 1920.
Shue, Henry. *Basic Rights*. Princeton, NJ: Princeton University Press, 1980.
Sistare, Christine, Larry May, and Leslie Francis, eds. *Groups and Group Rights*. Lawrence, KS: University Press of Kansas, 2001.
Steiner, Hillel. *An Essay on Rights*. Oxford: Blackwell, 1994.
Sterba, James P. "The Welfare Rights of Distant People and Future Generations: Moral Side-Constraints on Social Policy." *Social Theory and Practice* 7 (1) (1981): 99–119.
Sumner, L. W. *The Moral Foundation of Rights*. Oxford: Clarendon Press, 1987.
Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge, MA: Harvard University Press, 1990.
Tuck, Richard. *Natural Rights Theories*. Cambridge, U.K.: Cambridge University Press, 1979.
Vinogradoff, Paul. *Jurisprudence*. In *Collected Papers of Paul Vinogradoff*, vol 2. Oxford: Clarendon Press, 1928.
Vlastos, Gregory. "Justice and Equality." In *Social Justice*, edited by Richard B. Brandt. Englewood Cliffs, NJ: Prentice-Hall, 1962.
Wade, Francis C. "Potentiality in the Abortion Discussion." *Review of Metaphysics* 29 (2) (1975): 239–255.
Waldron, Jeremy. *Liberal Rights*. Cambridge, U.K.: Cambridge University Press, 1993.
Warren, Mary Anne. "On the Moral and Legal Status of Abortion." *The Monist* 57 (1) (1973): 43–61.
Wellman, Carl. *Real Rights*. New York: Oxford University Press, 1995.
Wellman, Carl, ed. *Rights and Duties*, 6 vols. New York: Routledge, 2002.
Wellman, Carl. *A Theory of Rights.* Totowa, NJ: Rowman and Allenheld, 1985.

Carl Wellman (2005)

RIGNANO, EUGENIO

(1870–1931)

Eugenio Rignano was an Italian positivist philosopher and founder (1907) and lifelong editor of the scientific journal *Scientia.* Rignano's first works were sociologically oriented, but he later turned to biology and philosophical biology. His major work, *Psicologia del ragionamento* (1920), places the activity of memory at the basis of all biological and psychic phenomena. Memory is an activity that, through the specific accumulation of concepts, makes possible the progressive adaptation of the organism to the environment, the formation of instincts and emotions, and, in higher organisms, of reasoning. According to Rignano, reasoning is "a series of operations or experiences merely thought out simply"; in other words, a series of operations performed in imagination. The results of these operations are also imagined and are assumed as the conclusions of the reasoning itself. This conception of reasoning, which Rignano derived chiefly from Ernst Mach, was later applied by him to explain the various kinds of reasoning: intuition, reduction, mathematical and mathematico-logical reasoning, intentional reasoning (dialectical or metaphysical reasoning), and pathological forms of reasoning as well. Rignano stressed the distinction between constructive and intentional reasoning. Constructive reasoning is motivated by a desire to discover the truth, and intentional reasoning by a desire to confirm a truth that a person believes he already possesses. Both types of reasoning utilize the same syllogistic form, but constructive reasoning is characteristic of the positivist scientist and intentional reasoning of the metaphysician. Rignano did not distinguish clearly between logical and psychological considerations; rather, he assumed the psychological mechanism as the basis of the logical validity of reasoning processes. The result is that Rignano's account is not very convincing either as logic or as psychology.

Despite his distaste for metaphysics, Rignano in subsequent works elaborated a kind of biological metaphysics based on the hypothesis that at the foundation of life and its evolution there is a "nervous energy" able to mold organic matter and direct it toward an increasing development and a growing adaptation to the environment. According to Rignano, life in its entirety shows a finalistic aspect that would be inexplicable if it were the product of physicochemical forces. This finalism can be explained, however, by assuming that life is a product of psychic, mnemonic energy, which on the basis of past experience envisions ends of future experience and adapts organic material to those ends. It is a kind of vitalism or animism that, according to Rignano, guarantees to evolution a progressive significance. The progress of evolution continues beyond organic life into moral life. The purpose of moral life is to guarantee to all individuals the satisfaction of their needs and to coordinate these needs in harmonious forms that gradually eliminate conflicts.

See also Cybernetics; Italian Philosophy; Mach, Ernst; Memory; Metaphysics; Philosophy of Biology; Positivism; Vitalism.

Bibliography

WORKS BY RIGNANO

Di un socialismo in accordo con la dottrina economica liberale. Turin: Fratelli Bocca, 1901.
La sociologica nel corso di filosofia positiva di Augusto Comte. Palermo, 1904.

Psicologia del ragionamento. Bologna, 1920. Translated by W. A. Holl as *Psychology of Reasoning.* London: K. Paul, Trench, Trubner, 1923.
La vita nel suo aspetto finalistico. Bologna, 1922.
La memoria biologica. Bologna, 1922. Translated by E. W. MacBride as *Biological Memory.* London: K. Paul, Trench, Trubner, 1926.
Che cosa è la vita? Nuovi saggi de sintesi biologica. Bologna, 1926. Translated by N. Mallinson as *The Nature of Life.* London, 1930.
Man Not a Machine. London: K. Paul, Trench, Trubner, 1926.
Problemi della psiche. Bologna, 1928.

Nicola Abbagnano (1967)
Translated by Nicholas Spadaccini

RILKE, RAINER MARIA (RENÉ)

(1875–1926)

The German poet Rainer Maria (René) Rilke was born in Prague, the son of a minor railway official. His mother, who was of upper-middle-class origin, encouraged him in his early ambition to become a poet. The years 1886–1891, which Rilke spent at military academies in Moravia and Austria, had a traumatic effect on him, and not until 1920 was he able to come to terms with his unhappy childhood and family background. His first volume of poetry, *Leben und Lieder,* appeared in Prague in 1895. Desultory studies, mainly in the history of art, at the universities of Prague, Munich, and Berlin were followed by two journeys to Russia in 1899 and 1900 in the company of Lou Andreas-Salomé, a German-Russian to whom Friedrich Nietzsche had proposed marriage and who later became a follower and friend of Sigmund Freud. During the second of these journeys he met Lev Tolstoy. On his return Rilke joined an art colony in Worpswede near Bremen, and early in 1901 he married the sculptress Clara Westhoff, one of its members. They had a daughter, but the short-lived marriage was only an interlude in Rilke's essentially solitary and unsettled life. For the next few years, Rilke's attention was centered on Paris and on Auguste Rodin, to whose work he devoted a monograph in 1903. Although his job as Rodin's private secretary ended in a quarrel, Rilke never ceased to acknowledge the very direct inspiration he received from close daily contact with the sculptor. The first collection of poems that bears the authentic stamp of greatness, *Neue Gedichte I* (Leipzig, 1907), represents Rilke's aim to render in words the immediacy, concreteness, and intensity ("the inward reality") that he discerned in Rodin's work.

With a single-mindedness that has rarely been paralleled in modern literature, Rilke devoted his whole existence to the poetic task he felt called upon to accomplish, subordinating to it all personal and public considerations. The long list of his patrons, most of whom belonged to the aristocracy of central Europe and a few to the German and Swiss patrician bourgeoisie, testifies to the restlessness of his life, and so do his journeys to Sweden (in 1904, on the invitation of Ellen Key), Italy, north Africa (1910–1911), Spain (1913), and repeatedly to France. The long list of his friends (mainly female) and correspondents, among them Paul Valéry and André Gide, includes surprisingly few German writers. Two places were of major importance for the fruition of his poetry: Duino (1910 and 1912), a castle on the Adriatic that belonged to the Princess Marie von Thurn und Taxis-Hohenlohe, where the first Duino Elegy was written, and the little castle of Muzot in the Swiss canton of Valais. It was at Muzot, in February 1922, as the guest of Werner Reinhart, that Rilke, in a storm of inspiration, wrote most of the fifty-five *Sonette an Orpheus* and several smaller collections of poems; and it was there, above all, that he completed his greatest work, which had been interrupted by World War I—the cycle of ten *Duineser Elegien,* several of which were written in the span of a few days. Rilke died at Valmont, Switzerland, after a protracted and painful illness that was diagnosed as leukemia.

Rilke's mature poetry, written after 1907, displays a consistency of attitude and a coherence of poetic devices that make it representative of a whole era of European thought. Following in the wake of Nietzsche, this poetry is informed by an acute historical consciousness. We live in an age when a "religion" that is based on separating transcendence from immanence is no longer viable:

> All of the living
> Make the mistake of drawing too sharp distinctions.
> Angels (it is said) would be often unable to tell
> Whether they moved among living or dead.
>
> (FIRST ELEGY)

Our impoverished state is marked by our awareness that "we are not very reliably at home in the interpreted world." Hence, in order to regain for ourselves something that would equal the spiritual and existential fervor that characterized the ages of faith, we must take upon ourselves the task of endowing the world (which, for Rilke, is the world of things and of intimate personal relations) with the inwardness of feeling that other ages directed toward a divinity. Joy, love, and above all suffering and

pain should not be diffuse sensations accompanying an unending series of vague hopes and regrets; they must become the objects of a total commitment. Thus, since we are "not yet" strong enough to give ourselves totally in love, we had better follow the example of the lover ("Gaspara Stampa") who drew her strength from an unrequited, "uninterrupted" feeling, or indeed of Narcissus, who used the natural world as a magnifying mirror of his feeling. *Les saltimbanques,* the traveling *artistes* of Pablo Picasso's "blue period" paintings, celebrated in the Fifth Elegy, most fully symbolize our condition. In a world in which all actions are liable to remain uncompleted ("We, though, while we're intent upon one thing, / can feel the cost and conquest of the other"), suffering—the fullest possible realization and appropriation by the self of what is inflicted from without—will be the greater virtue:

> Killing merely is a form of our wandering sadness …
> Pure in the spirit serene
> Is what we ourselves endure.

History, for Rilke, is not a social phenomenon but a pageant of situations and persons in whom the ideal of completion and strength of feeling was realized, just as the contemporary world is a series of images that portray our deprivations and stunted responses. To make something of one's fate, of one's experiences, is to give them the permanence (essentially poetic) of a moment of intensity. Similarly, the supreme task, set by the imminence of death, is to repair the adventitiousness of death by drawing it into my life, by making of it "my own death."

The immensity of the task of creating a new spirituality is betrayed by the complex, and quite conscious, ambiguities of Rilke's images of transcendence, chief of which is the image of the Angel, as he appears in the Elegies. He is a messenger (*angelos*) from another sphere; hence, there must be one who sent him. But the Angel comes upon us with a terrible majesty and strength which, to us who are weak, is all his own. In many such astonishing images Rilke expresses the "pure [=necessary] contradiction" that he sees as the root of our being: only by living in total commitment to "the Earth," the here and now, can man transform it into "the heart's inner space," and thus wrest some eventual transition into a "soundless" Beyond—wrest it from he knows not whom. The most accomplished practitioner of such transformations is Orpheus, the poet-maker who, in the creative act, stills all strife by transforming it into song, eternalizes the moment by making of it a monument of inwardness, and transfixes suffering into the eternally valid image of "Lament" (Tenth Elegy).

In a world yearning for the security of faith and finding it in ideology, Rilke's vertiginous images were reduced to prosy precepts for living, becoming thus at once esoteric and banal. Rilke's poetry is not necessarily esoteric, and the creative activity he extolled is closely related to the poetic; but he addressed himself to the single individual. The social sphere of modern life is branded as wholly inauthentic (Rilke either ignored or briefly satirized it); all concerted action is an escape from defective selfhood. He understood and expressed velleities supremely well; his poetry hardly offers a nostrum to cure them.

See also Freud, Sigmund; "My Death"; Nietzsche, Friedrich; Valéry, Paul.

Bibliography

WORKS BY RILKE

Only the most recent four-volume edition, *Gesammelte Werke,* edited by Ernst Zinn (Frankfurt, 1955–1961), shows the magnitude of Rilke's work. In addition, some eight collections of his letters are of major importance. The two-volume edition of *Selected Works* (London, 1954–1960) includes J. B. Leishman's brilliant translations of all of Rilke's major poetry.

WORKS ON RILKE

No English biography has appeared since E. M. Butler, *Rainer Maria Rilke* (Cambridge, U.K.: Cambridge University Press, 1941). H. E. Holthusen, *Rainer Maria Rilke in Selbstzeugnissen und Bilddokumenten* (Hamburg: Rowohlt, 1958), is a succinct and discriminating account of the poet's life. See also N. Purtscher-Wydenbruck, *Rilke: Man and Poet* (London, 1949), and E. Buddeberg, *Rainer Maria Rilke—Eine innere Biographie* (Stuttgart: Metzler, 1955). A brilliant assessment of Nietzsche's influence on Rilke is Chapter 5 of Erich Heller, *The Disinherited Mind* (Cambridge, U.K.: Bowes & Bowes, 1952). E. C. Mason has devoted many studies to a critique of Rilke's work, among them *Lebenshaltung und Symbolik bei R. M. Rilke* (Weimar, Germany: H. Böhlaus Nachf, 1939) and *Der Zopf des Münchhausen* (Einsiedeln, Germany, 1949). Two exegetic works of major importance are Jacob Steiner, *Rilke's Duineser Elegien* (Bern: Francke, 1962), and Romano Guardini, *Rilke's Duino Elegies: An Interpretation,* translated by K. G. Knight (London, 1961).

J. P. Stern (1967)

RINTELEN, FRITZ-JOACHIM VON
(1898–1979)

Fritz-Joachim von Rintelen, the German philosopher of value, was born in Stettin. He received a doctorate in philosophy in 1924 from the University of Munich, where he began to lecture in 1928. Von Rintelen was appointed professor at the University of Bonn in 1933 and at Munich in 1936, but he was suspended on political grounds in 1941. In 1947 von Rintelen became professor of philosophy, psychology, and pedagogy at the University of Mainz. He was a visiting professor at the Universidad Nacional de Córdoba, Argentina, from 1951 to 1952 and at the University of Southern California in 1957.

Von Rintelen gave both a systematic and a detailed historical interpretation of the problem of value. Prior to his dismissal from his professorship at Munich, he had explicated in detail a theory of value (*Wert*) and of meaning (*Sinn*) and had built a philosophical anthropology upon it. After World War II he applied this theory to an analysis and penetrating criticism of the irrationalistic, nihilistic, and pessimistic currents in contemporary European philosophy and literature, showing how the theory resolves the conflicts and paradoxes that he reveals in these currents.

His doctrine of values and of personality is rooted in the realistic tradition of Platonism and Scholasticism but also shows the influence of German idealism. The chief direction of his thought was set in his two academic dissertations, a criticism of the pessimistic philosophy of religion of Eduard von Hartmann and an attempt to extend Ernst Troeltsch's efforts, in the later years of his life, to overcome historical relativism through a theory of values and their operation in history.

Two points in von Rintelen's criticism are particularly salient. The first is his attack upon all dualisms of intellect and will or of mind and life (Max Scheler), all subordination of the rational to a more inclusive irrational, and every combination of an idealistic theory of scientific knowledge with a realism in metaphysics. To these distinctions he opposes an ontological interpretation of value by which these dualistic tensions can be resolved. He rejects von Hartmann's teleology of self-destruction as an ontological impossibility and an aesthetic misreading of the tragedy of our culture; this tragedy cannot be denied, but it implies a transcendent normative meaning to be attained through the reflective transformation of our actions. Thus, there is an inclusive ontological meaning, not attainable through scientific logic but through the value experiences of life, which sustains the human spirit and human life.

Von Rintelen's ontological theory of real value (*Realwert*) was constructed in opposition to psychological, positivistic, and phenomenological definitions. Real value is an objective context of meaning that can be particularized and made concrete through conscious or unconscious strivings. Each actualized value possesses an intrinsic worth varying in intensity with its degree of meaning and a relational worth by virtue of which it enters into a wider order of values. Thus every real value is vertically capable of degrees of normative validity and historically capable of individualization within larger contexts of culture and of personal action. Values are individualized in two spheres corresponding to life and mind in man. In the sphere of nature, objects are primarily existent and only secondarily valued; in that of mind or personality, objects are primarily mental and only secondarily grounded in concrete existence. From this viewpoint human history can be understood as a continually renewed effort to actualize values in terms of a personal regulative ideal of the highest possible fulfillment and in relation to an ultimate *summum bonum,* God, the unique, autonomous, and inclusive real value.

In 1932 von Rintelen published the first volume (ancient and medieval) of a historical study in which the development of this theory of value was to be traced in European thought. This work was left incomplete but was supplemented by specialized historical and systematic articles.

In his later critiques of existentialism and other contemporary intellectual currents, von Rintelen analyzed the plight of man as portrayed in modern philosophy and literature and, by correcting the subjectivism and finitism implicit in this portrayal through his own doctrine of value transcendence, points out the way to "a rewon security of spirit." Outstanding among these works are *Philosophie der Endlichkeit* (1951), which includes analyses of Martin Heidegger, Rainer Maria Rilke, Gabriel Marcel, and Jean-Paul Sartre, and *Der Rang des Geistes* (1955), a thorough and distinguished study of Johann Wolfgang von Goethe as philosopher, in which the inner tensions or polarities of the poet's thought are examined and Goethe's movement from an eclecticism to a rationally justified theism and an operative human ideal of rational freedom and love is portrayed. In these books von Rintelen shows himself not merely as a constructive philosopher but also as an able critic of literature and culture.

Von Rintelen's thought may thus be considered as a reconstruction of the Christian intellectual tradition in which the inevitable tragedy that inheres in the polarities of human existence may be overcome through a transcendent order of values in which meaning and impulse are harmonized.

See also Existentialism; Goethe, Johann Wolfgang von; Hartmann, Eduard von; Heidegger, Martin; Idealism; Intrinsic Value; Marcel, Gabriel; Philosophical Anthropology; Platonism and the Platonic Tradition; Rilke, Rainer Maria (René) ; Sartre, Jean-Paul; Scheler, Max; Troeltsch, Ernst; Value and Valuation.

Bibliography

Works by von Rintelen include *Pessimistische Religionsphilosophie der Gegenwart: Untersuchung zur religionsphilosophischen Problemstellung bei Eduard von Hartmann und ihre erkenntnistheo-retischmetaphysischen Grundlagen* (Munich, 1924); "Der Versuch einer Überwindung des Hitstorismus bei Ernst Troeltsch," in *Vierteljahresschrift für Literaturwissenschaft und Geistesgeschichte* 8 (1930): 324–372; *Der Wertgedanke in der europäischen Geistesentwicklung,* Part I, *Altertum und Mittelalter* (Halle, 1932); *Dämonie des Willens* (Wiesbaden, 1948); *Von Dionysos zu Appollon. Der Aufstieg im Geiste* (Wiesbaden, 1948): *Philosophie der Endlichkeit als Spiegel der Gegenwart* (Meisenheim am Glan, Germany: Hain, 1951), partly translated by Hilda Graf as *Beyond Existentialism* (London: Allen and Unwin, 1963); and *Der Rang des Geistes. Goethes Weltverständnis* (Tübingen, 1955).

Other works by von Rintelen are *Der Aufstieg im Geiste*, 2nd ed. (Frankfurt: Metopen-Verlag. 1968); *Johann Wolfgang von Goethe: Sinnerfahrung und Daseinsdeutung* (Munich: E. Reinhardt, 1968); *Contemporary German Philosophy and Its Background* (Bonn: Bouvier, 1970); *Values in European Thought* (Pamplona: Universidad de Navarra, 1972–); *Philosophie des lebendigen Geistes in der Krise der Gegenwart: Selbstdarst* (Frankfurt [Main]: Musterschmidt, 1977); *Beyond Existentialism* (Westport, CT: Greenwood Press, 1978).

For further bibliography and evaluation, see Richard Wisser, "Wertwirklichkeit und Sinnverständnis. Gedanken zur Philosophie von Fritz-Joachim von Rintelen," in *Sinn und Sein: Ein philosophisches Symposion F. J. v. Rintelen gewidmet,* edited by Richard Wisser (Tübingen: Niemeyer, 1960), pp. 611–708.

L. E. Loemker (1967)

Bibliography updated by Michael J. Farmer (2005)

RITSCHL, ALBRECHT BENJAMIN
(1822–1889)

Albrecht Benjamin Ritschl, the German theologian, was born in Berlin and studied theology at Bonn, Halle, Heidelberg, and Tübingen. He taught theology at Bonn from 1846 to 1864, and at Göttingen for the remainder of his career. Ritschl reexamined Christianity in the light of neo-Kantianism and historicist principles. After 1875 his influence was widespread in a number of German universities and led to increased interest in religious psychology, comparative religion, and related fields. However, his school came under sharp criticism from orthodox, pietist, and liberal quarters.

Ritschl undertook to establish Christian theology as an autonomous and systematic discipline. To do this he had first to purge German religious thought of pietism, Hegelian speculative theism, and the pantheism of Friedrich Schleiermacher and then to apply the techniques and results of contemporary literary and historical criticism. On the basis of Immanuel Kant's ascription of priority to practical reason over theoretical reason and his separation of philosophy and religion, Ritschl distinguished between value judgments and theoretical judgments. Unlike Kant, however, Ritschl accorded primacy to religion over philosophy on the grounds that spirit (the noumenal) takes precedence over matter (the phenomenal); also unlike Kant, he accorded moral primacy to the community (the nation) over the individual.

Ritschl believed that the deep-rootedness and continuity of religion, as expressed in dogmas and institutions, testifies to the reality and superiority of the religious need of practical reason in human nature. This need arises out of a basic contradiction between nature and spirit in human nature. The value of religion and particularly of Christianity, Ritschl thought, can be verified by history, which shows that this contradiction seeks a resolution in some form of redemption in the world. The Kantian elements in Ritschl's thinking, in combination with this positivist tendency, led him to believe that history does not merely provide material in support of some arbitrary, nonhistorical preconception but reveals an essential structure of human consciousness and the intrinsic historicity of Christianity.

In attempting to satisfy both the requirements of history and the claims of practical reason, Ritschl adopted the dogmas of redemption and the kingdom of God as embodied in the life of Christ as the pivots of his religious theory. Man seeks to realize his destiny here on earth by

leading an ethically self-conscious life, which is the core of religiosity. The acts of love that he performs, the content of the ethical life, represent the human counterpart to redemption, and the community required for their performance represents the terrestrial counterpart to the kingdom of God. God's purpose is thus manifest in history. Sin, which is only the result of ignorance, is pardonable because it is only a transitory opposition to this purpose. Ritschl therefore rejected the dogma of original sin as unhistorical and hence unverifiable.

Biblical exegesis led Ritschl to believe that the community is both logically and chronologically prior to the church. Only in the community can man find justification and reconciliation in God. Christ was founder of a community and can be comprehended historically only through our knowledge of how that community conceived him.

From his conviction that religious consciousness is universal and characterized by its quest for redemption, Ritschl concluded that Christianity is the superior expression of that consciousness. History, rather than dogma, verifies Christianity, but its validity is thereby strengthened, not relativized.

Although the community takes precedence over the individual, the individual is not thereby depreciated but is provided with a field within which he is able to realize his personality. While the community is prior to the church, this does not devalue the church's interests but emphasizes its actual efficacy as the ecclesiastical form of the community's organization. Religious truths are established in practice rather than by their appearance in the New Testament, but its authority is thereby strengthened, not subverted. Martin Luther is the most significant religious figure since Christ, not because he modernized Christianity but because he recaptured and restored an understanding of the original Christian attitude.

See also Historicism; Kant, Immanuel; Neo-Kantianism; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

WORKS BY RITSCHL

Die christliche Lehre von der Rechtfertigung und Versöhnung, 3 vols. Bonn, 1870–1874. Vols. I and III translated (Vol. I by J. S. Black, Vol. III by H. R. Mackintosh and A. B. Macauley) as *Critical History of the Christian Doctrine of Justification and Reconciliation.* Edinburgh and New York, 1872–1900.

Geschichte des Pietismus, 3 vols. Bonn: Marcus, 1880–1886.

Theologie und Metaphysik. Bonn, 1881.

WORKS ON RITSCHL

Barth, Karl. *Die protestantische Theologie im 19. Jahrhundert.* Zürich: Evangelischer Verlag, 1947. Translated by B. Cozens as *Protestant Thought: From Rousseau to Ritschl.* New York: Harper, 1959. Pp. 390–399.

Ecke, G. *Die theologische Schule Albrecht Ritschls,* Berlin: Reuther and Reichard, 1897–1904.

Flügel, O. *Ritschls philosophische Ansichten.* Langensalza, 1886; 3rd ed., 1895.

Garvie, Alfred E. "Ritschlianism." In *Encyclopedia of Religion and Ethics,* edited by James Hastings, Vol. X, pp. 812–820. Edinburgh and New York, 1919.

Häring, Theodor. *Zu Ritschls Versöhnungslehre.* Zürich, 1888.

Mackintosh, R. *Albrecht Ritschl and His School.* London: Chapman and Hall, 1915.

Ritschl, Otto, *Albrecht Ritschls Leben,* 2 vols. Freiburg and Leipzig: Mohr, 1892–1896.

Schoen, H. *Les origines historiques de la théologie de Ritschl.* Paris, 1893.

Weber, W. *Die Frage der Rechtfertigung in der Theologie Albrecht Ritschls.* Heidelberg, 1940. Dissertation.

Wendland, J. *Albrecht Ritschl und seine Schüler.* Berlin, 1899.

Robert Anchor (1967)

ROBINET, JEAN-BAPTISTE-RENÉ

(1735–1820)

Jean-Baptiste-René Robinet, the French littérateur and speculative philosopher, was born in Rennes. He started to become a Jesuit, but withdrew from the order and went to Holland to devote himself to letters. There he published his principal work, *De la nature* (4 vols., Amsterdam, 1761–1768), and in 1768, *Considérations philosophiques de la gradation naturelle des formes de l'être, ou les Essais de la nature qui apprend à faire l'homme* (2 vols., Amsterdam and Paris). He eked out an existence by hackwork, translating English novels and giving English lessons. He became embroiled with Voltaire by selling the manuscript of *Lettres secrétes* for publication without Voltaire's permission. He went to Paris in 1778 when he was made royal censor and secretary to one of the king's ministers. During the Revolution he returned to Rennes, where he lived quietly. In addition to many minor pieces, he published a translation of David Hume (*Essais de morale, ou Recherches sur les principes de la morale,* 1760) and edited a vast compilation, *Dictionnaire universel des sciences morale, économique, politique et diplomatique* (London, 1777–1783, 30 vols. in quarto).

De la nature caused some stir because of its strange ideas. When it was attributed to François-Vincent Toussaint, Denis Diderot, and Claude-Adrien Helvétius, Robi-

net admitted his authorship in a letter to the *Journal des savants.* The many quotations in *De la nature* testify to its author's vast readings; his thinking, however, is original. It is characterized by a curious mélange of mysticism and scientific spirit. *De la nature* touches on many subjects, but its announced theme is a modern version of Manichaeanism: There is an equilibrium of good and evil in all substances and their modes. Robinet's purpose is to exculpate God and establish the necessity of evil. Embracing Benedict de Spinoza's principle that all possibles exist, he attacks Gottfried Wilhelm Leibniz by asserting that, therefore, there can be only one world and that God had no choice in the matter. "God no more had the power to modify the nature of the world than his own nature."

Robinet argued that behind the apparently random distribution of pleasure and suffering in the world there lies a fluid but fixed order. "The physical economy is such that good and evil are engendered with equal fecundity. They flow naturally from the depth of essences." God can in no way remove evil, for omnipotence does not extend to impossibles or contradictions. The suppression of evil implies contradiction, for good without evil would be infinite. The total quantity of good and evil is at every moment equal. Thus the harmony of the world is always the same, and progress is a myth or an illusion. Despite this equilibrium, God is good and his justice is seen in his not having favored one species at the expense of the others; for man is not king of the universe, as Buffon had claimed, and nothing has been created especially for his use. For human beings, life is a balance of happiness and unhappiness, and they should therefore console themselves by the enjoyment of pleasures. Moderation is the best path in all areas of life. The lower classes must be kept in ignorance, for their own benefit and that of the state; slavery is justifiable. Human nature being what it is, equality and fraternity are impossible.

The universe, for Robinet, is animate. All forms of being, including planets and stars, have the power of reproduction. The individual is unimportant, an instrument nature uses for its procreative purposes; only the species endures. Robinet speculates that nature has developed variations on a single prototype; from stones to men, there is a natural gradation of beings. The "prototype" is "a germ which tends naturally to develop itself.… Its energy cannot be repressed.… The germ develops, then, and each degree of development gives a variation of the prototype, a new combination of the original plan." The only difference between stone, plant, and animal is "the measure in which they participate in that essence.… A stone, an oak, a horse are not men; but they can be regarded as more or less rough types in their relation to a single primitive design." We must consider the succession of individuals "as so many steps of being [advancing] toward humanity."

Robinet draws close to an evolutionary hypothesis in his concept of nature as experimenting and as developing toward greater complexity; he also considers all species as related. It is not a true evolutionism, however, inasmuch as each trial in the ascending scale of complexity is made de novo from the relatively unorganized stage of the original prototype. Species do not themselves have a history but are fixed once they are spewed forth. Robinet also pictures a biological struggle for existence and a natural balance, but does not relate these to transformism. Robinet's work influenced both Johann Gottfried Herder and G. W. F. Hegel and was considered of interest in the former Soviet Union.

See also Evolutionary Theory; Diderot, Denis; Evil; Hegel, Georg Wilhelm Friedrich; Helvétius, Claude-Adrien; Herder, Johann Gottfried; Hume, David; Leibniz, Gottfried Wilhelm; Mani and Manichaeism; Spinoza, Benedict (Baruch) de; Voltaire, François-Marie Arouet de.

Bibliography

Damiron, M. *Mémoires pour servir á l'histoire de la philosophie au XVIII^e siècle,* Vol. II. Paris: Ladrange, 1858.

Rosso, C. *Moralisti del bonheur.* Turin, 1954.

L. G. Crocker (1967)

ROCHEFOUCAULD, DUC FRANÇOIS DE LA

See *La Rochefoucauld, Duc François de*

ROHAULT, JACQUES

(1620–1672)

Jacques Rohault was a mechanistic Cartesian experimental physicist. He was born in Amiens, France, and earned his MA in Paris in 1641. There, he became Claude Clerselier's Cartesian disciple and son-in-law. He was Pierre-Sylvain Régis's teacher and converted him to Cartesianism. In the 1650s Rohault was a private tutor in Paris, and his "Cartesian Wednesday" evening lectures, complete with laboratory table demonstrations, were

attended by many members of the noble class, women as well as men, and did a great deal toward popularizing Cartesianism. His *Traite de physique* (Paris, 1671) was a standard text for nearly fifty years. John Clarke and Samuel Clarke, rather than writing a Newtonian physics, translated Rohault's work into Latin (1697) and English (1723) and added Newtonian footnotes to correct Rohault's Cartesian mistakes. The *Traite* contains descriptions of explanations and experiments in support of Cartesian mechanistic physics. Like René Descartes, Rohault holds that these explanations are only probable because absolute certainty is unattainable by humans.

Also in Paris in 1671 Rohault published his *Entretiens sur la philosophie*, in which he defends the thesis that Cartesian principles and Christian doctrines do not conflict because each pertains to a separate and distinct realm of truth and knowledge. The book was popular, but Rohault's position was generally viewed as heretical by the Catholic Church.

Rohault opposes Nicolas Malebranche's occasionalism and presents his own mechanistic Cartesianism based on eight axioms he takes to be self-evident:

(1) Nothing (that which has no existence) has no properties

(2) Something cannot possibly be made of nothing, that is, nothing cannot become something

(3) No thing or substance can be annihilated, that is, something cannot be reduced to nothing

(4) Every effect presupposes some cause

(5) If one does not cause an effect, that effect necessarily depends on some othercause

(6) Everything endeavors to continue in the state in which it is (an early Cartesian rendering of a principle of inertia)

(7) Every alteration is made by some external cause, that is, in opposition to Aristotle, no material thing can alter itself through an inner power, force, or form

(8) Every alteration is proportional to the force of the causal agent

Certain propositions follow logically from these axioms, but Rohault says these truths of reason remain purely formal and have no application if there are no existents. Thus, the first task in understanding the world is to seek out existents. In strict Cartesian order one knows first one's own self, whose existence Rohault proves syllogistically:

(a) From principle (1) above, whatever has properties is something

(b) Thinking is a property

(c) Whatever thinks, therefore, exists as something because it has the property of thinking

(d) I think

(e) Therefore, I exist

Reasoning with these principles about ideas and sensations leads to knowledge of the essences of mind, God, and matter and to proofs of the existence of God and of matter. The essence of mind is thought; of God, necessary existence; and of matter, extension. Rohault states that mind and matter are completely different but that God so created the human mind or soul such that motions caused by material impressions on the sense organs and in the brain of the body with which it is united give rise in the soul to sensations and ideas. Neither sensations nor ideas resemble material things, and so resemblance is not necessary for knowledge. It is simply the nature of sensations to give knowledge of the existence of material things, and the nature of some ideas is to give knowledge of the place, situation, distance, magnitude, figure, number, and motion or rest of material things.

Rohault's method in physics is to reason mathematically about experiments before conducting them. His goal is to explain the sensible effects of material things. For this only the primary material properties of size, figure, motion, and arrangement of divisible, impenetrable particles in a plenum are needed; occult qualities such as Aristotelian forms are unnecessary.

In *Entretiens de philosophie* (Paris, 1671), the companion volume to the *Traite*, Rohault explains in mechanical terms Cartesian opinions on animal machines and transubstantiation. Animal behavior, he claims, can be explained if animals are completely material; human behavior, however, requires a rational soul that is immaterial, hence indivisible, hence immortal. For Cartesians, the sensible qualities as they exist in material things are not seen, tasted, and so on as one sees and tastes them, but are merely the powers bodies have, determined by the size, figure, motion, and arrangement of their particles, to cause sensations in the mind. There is no further explanation of these powers beyond the fact that God made the correlations between bodily movements and one's sensory experience.

Transubstantiation, then, is the point-by-point replacement of bread and wine by Christ's flesh and blood. Therefore, the flesh and blood of Christ that occu-

pies the places (is bound by the surfaces) formerly occupied by bread and wine causes sensations exactly like those that the bread and wine formerly caused. Consequently, real accidents or Aristotelian forms subsisting separately from Aristotelian matter as postulated in the scholastic explanation of transubstantiation are unnecessary. There are further physical explanations and assurances that Cartesian principles do not contradict Catholic doctrine in *Oeuvres posthumes de Rohault* (Paris, 1682).

Overall, Rohault disclaims metaphysics and says that although the substitutions are miraculous, even his mechanist explanation of transubstantiation is only a solution to a problem in physics. His work illustrates the strong empiricist stress on observation and experiment toward probable mechanistic explanations in physics so prominent in many Cartesian philosophers. Finally, use of Rohault's *Traite* as a physics textbook merely with addition of Newtonian footnotes constitutes a major shift to nonmetaphysical, explanatory concerns in science.

See also Aristoteliasm; Aristotle; Cartesianism; Clarke, Samuel; Descartes, René; French Philosophy; Malebranche, Nicholas; Régis, Pierre-Sylvain.

Bibliography

Schmaltz, Tad M. *Radical Cartesianism: The French Reception of Descartes*. New York: Cambridge University Press, 2002.

Watson, Richard A. *The Breakdown of Cartesian Metaphysics*. Indianapolis, IN: Hackett, 1998. Originally published as *The Downfall of Cartesianism, 1673–1712: A Study of Epistemological Issues in Late Seventeenth-Century Cartesianism* (The Hague: Nijhoff, 1965).

Richard A. Watson (1967, 2005)

ROMAGNOSI, GIAN DOMENICO

(1761–1835)

Gian Domenico Romagnosi was born in Salsomaggiore, near Parma, and studied at the Collegio Alberoni in Piacenza. Through the teaching of Giovanni Antonio Comi, a follower of Gottfried Wilhelm Leibniz and Christian Wolff, Romagnosi became acquainted with the doctrines of Étienne Bonnot de Condillac and with the writings of Charles Bonnet, which had a decisive influence on him. After his graduation in 1786, he conceived his best-known work, *Genesi del diritto penale* (Genesis of Penal Law; completed in 1789 and published in Pavia in 1791), in which he claimed that the fundamental right to punish belongs to society. Society alone, and not the individual, can mete out "that amount of evil that is necessary to preserve the well-being of our fellow men" and can oppose the "criminal impulse" with a "moral counterimpulse."

Named mayor of Trent in 1791, Romagnosi remained in that office for ten years, during the period of the French Revolution and the rise of Napoleon Bonaparte. During this time he published, among other works, his *Cosa è l'eguaglianza* (What Is Equality?; Trent, 1792) and *Cosa è libertà* (What Is Freedom?; Trent, 1793). After a brilliant political career under the Napoleonic government, he became professor of natural and public law at Parma (1802), but after the restoration he was dismissed from his position and was arrested. The Austrian government also prevented him from accepting a post at the Ionian University at Corfu offered to him by Frederick North, Lord Guilford. Regarded as a master by Italian patriots, Romagnosi died, after a sad but active old age, in Milan. His major works, in addition to the *Genesi,* are considered to be the *Introduzione allo studio del diritto pubblico universale* (Introduction to the Study of Universal Public Law; Parma, 1805), the *Assunto primo della scienza del diritto naturale* (A First Thesis on the Science of Natural Law; Milan, 1820), and a series of essays on *incivilimento* (civilizing, or the process of civilization) in 1832.

Although he was influenced by the Enlightenment, Romagnosi remained attached to the historicism of Giambattista Vico and followed a "positive" method of research, advocating the activity of the human spirit rather than sensationalism and substituting for the abstractness of the isolated human individual the concreteness of the nation as the subject of the historical action. In epistemology he refused to reduce all cognitions to "transformed sensations," but at the same time he denied that intelligence is independent of sensitivity: In reality, "discernment" is already present in "feeling." The mind acts by means of its own "rational signs." These cannot be regarded as preexisting ideas but, rather, as manifestations of mental activity, which, along with the sensory datum, gives form to experience. By contrast, the correspondence of our prior judgments with the actual signs of things, that is, with experience, constitutes the criterion of the truth of our knowledge, which is sought and found pragmatically.

Romagnosi's civil philosophy is the most interesting part of his work: Man is real only in a historically determined society—the "collective person of the society"—which is in a state of constant civilizing progress and whose characteristic traits, elements, and laws Romagnosi

sought to define. Romagnosi's doctrine of *incivilimento* constituted a philosophy of history faithful to the concrete development of real events, in contrast with that of G. W. F. Hegel, which Romagnosi opposed as "ultrametaphysical." Society develops through the synthesis of national character (tradition) with stimulation—spontaneous, free, and renewing—according to a law of convenience, with all parts of the nation tending toward an equilibrium of force and utility through the balance of interests and powers. This dialectic of civilization is a work of art, even the highest work of art of a humanity striving for perfection.

See also Bonnet, Charles; Condillac, Étienne Bonnot de; Enlightenment; Hegel, Georg Wilhelm Friedrich; Historicism; Leibniz, Gottfried Wilhelm; Punishment; Vico, Giambattista; Wolff, Christian.

Bibliography

Editions of Romagnosi's works are *Opere,* 19 vols. in 24 (Florence, 1832–1837), and *Opere,* 8 vols. in 16 (Milan: Perelli e Mariani, 1841–1848; actually 1841–1852), both edited by Alessandro de Giorgi; *Opere scelte,* 3 vols. (Rome: Reale accademia d'Italia, 1936–1937); and *Opere di G. D. Romagnosi, C. Cattaneo, G. Ferrari,* edited by Enrico Sestan (Milan: Ricciardi, 1957).

Works on Romagnosi are L. G. Confalerione, *G. D. Romagnosi. Notizie storiche e biografiche. Bibliografia e documenti* (Carate Brianza, 1928); Achille Norsa, *Il pensiero filosofico di G. D. Romagnosi* (Milan, 1930); G. A. Belloni, *Romagnosi, profilo storico* (Milan, 1935); E. di Carlo, *Bibliografia Romagnosiana* (Palermo, 1936); Francesco Orestano, *Giandomenico Romagnosi* (Rome, 1935); Giacomo Perticone, *La filosofia "civile" di G. D. Romagnosi* (Florence, 1935); Alessandro Levi, *Romagnosi* (Rome, 1935); Lorenzo Caboara, *La filosofia politica di Romagnosi* (Rome: Istituto di filosofia del diritto, 1936); and Loris Ricci Garotti, "G. D. Romagnosi nella critica recente," in *Società* 14 (1959): 109–140.

Eugenio Garin (1967)
Translated by Robert M. Connolly

ROMAN PHILOSOPHY

See *Cicero, Marcus Tullius; Lucretius; Marcus Aurelius Antoninus; Seneca, Lucius Annaeus; Stoicism*

ROMANTICISM

"Romanticism" and "romantic" are protean words, the despair of a rigorous semanticist. They designate a generally accepted period, especially in literature and the arts, of Western cultural history, roughly from the late eighteenth to the mid-nineteenth century. They embrace a cluster or syndrome of ideas about the true, the good, the beautiful, philosophical ideas both in the popular and in the technical sense, ideas endlessly debated in the last few centuries. Although the behavioral scientists groping to establish a rigorous classification of human personality generally eschew the word, *romantic* remains in common use to describe a temperament or personality often, perhaps usually, held to be a constitutional element of an individual and at least in part independent of cultural fashion. In all these senses "romanticism" and "romantic" cover a multitude of particulars that in a given combination can appear very different, if not mutually incompatible. Hence so good a historian of ideas as Arthur Lovejoy urges the use of the plural, *romanticisms,* and can write of the "Chinese origins of *a* romanticism"; and W. T. Jones insists that romanticism can only be understood as a very complex syndrome of "biases" in the direction of what he calls the dynamic, the disordered, the continuous, the soft-focused, the inner, the this-worldly.

THE ROMANTIC TEMPERAMENT

Sensitive, emotional, preferring color to form, the exotic to the familiar, eager for novelty, for adventure, above all for the vicarious adventure of fantasy, reveling in disorder and uncertainty, insistent on the uniqueness of the individual to the point of making a virtue of eccentricity, the typical Romantic will hold that he cannot be typical, for the very concept of "typical" suggests the work of the pigeonholing intellect he scorns. Though his contempt for this world of reason and commonsense calculation may push him toward otherworldliness, the Romantic is too much a man of words and sensations to make a good mystic. He may admire the mystic, especially the exotic mystic from the East, but he himself is a good Westerner. In fact, the difficulties of reconciling the often contradictory particulars of romanticism in respectable generalization come out in any attempt to isolate a romantic personality. William Blake has most of the marks of the Romantic, from the positive one of extreme transcendental yearning to the almost universal romantic negative one of contempt for the "meddling intellect"; yet in his quite otherworldly drawings his symbolic, mystical figures are delineated with a draftsmanship of classical solidity and of firm this-worldliness. There is nothing fuzzy, nothing Turner-like, in Blake's art. William James has the full romantic love for the struggling, the unestablished, the untried; but he cannot be accused of what he

himself called "tender-mindedness," of idealistic distrust of the instrument of thought. Friedrich Nietzsche, who used "romantic" as a term of reproach, who said of Richard Wagner's music that it sweats, and called Mme. de Staël "that prolific ink-yielding cow," shared all the romantic hatreds for the shopkeeper's world of grubbing common sense and above all had the Romantic's desire for *etwas mehr,* the something more of Percy Bysshe Shelley's "desire of the moth for the star."

However difficult the romantic personality may be to isolate in analysis, it can be recognized all through Western cultural history, and indeed in the active life of enterprise and politics. Euripides and Catullus were surely Romantics. The *Odi et amo* (I hate and I love) of Catullus is a classic assertion of romantic ambivalence; the *rumoresque senum severiorum/omnes unius aestimemus assis* (Let us regard all the gossip of censorious old men as not worth one penny) is a fine assertion of one of the minor marks of romanticism, contempt for the Philistine decencies of the old in spirit. François Villon and François Rabelais were Romantics, even though they were Frenchmen who, as Frenchmen, so nineteenth-century English and German romanticists thought, should have been incapable of transcending the petty ways of *mesure* and *la raison raisonnante.* In our own day, the romantic temperament crops up everywhere—in artists and poets of course, but also in philosophers. Henri Bergson was a Romantic. But so too, it may be argued, was A. N. Whitehead; and there are scientists not untouched by the desire of the moth for the star. In active life, Alexander the Great and Napoleon Bonaparte were Romantics; Frederick the Great and Otto von Bismarck were classicists.

There are then, in our Western civilization, presumably always born romanticists and born classicists—or born Dionysians and born Apollonians, to use an expressive dualism especially popular with the Germans from Gotthold Ephraim Lessing through Nietzsche to Oswald Spengler. (The Germans usually classify themselves as the great Dionysian force in the West.) We can but guess at the distribution of these two types in a general population. Probably the well-defined or extreme temperaments are limited in numbers always; most human beings can adapt to the fashion of their age. In one age, say Vergilian and Horatian Rome, or the France of Louis XIV, the Apollonian is dominant, the Dionysian subdued, even silent. Sometimes in Apollonian ages, however, the Dionysian is the rebel, the man out of tune with his times; Giambattista Vico, perhaps, should be so listed in the Apollonian early eighteenth century. In another age, and notably in the Romantic Age here considered, the Dionysian is dominant and the Apollonian repressed, sometimes tempted, as was the quite unecstatic J. S. Mill, to romantic depths of understanding.

ROMANTICISM AND THE ENLIGHTENMENT

One type can be dominant, but not in sole and exclusive possession. To the cultural historian, the early and mid-eighteenth century and the early nineteenth can stand for two great antithetical styles or fashions: the first, classical or enlightened; the second, romantic. The years from about 1770 to the first decade of the nineteenth century are obviously years of transition. In a graph, the rising lines of Romanticism cross the descending lines of classicism somewhere in the 1770s in Germany (with the heyday of "Sturm und Drang"), 1798 in England (with the publication of the *Lyrical Ballads*), and 1820 in France (with the publication of *Méditations* by Alphonse-Marie-Louis de Prat de Lamartine). But even after the triumph of Romanticism as a cultural fashion, individuals and groups continued to display the tastes and attitudes associated with the classicism and rationalism of the eighteenth-century Enlightenment J. S. Mill tells us in his autobiography that he was influenced by the lyricism and even the transcendentalism of the Lake poets, notably Samuel Taylor Coleridge; but the influence seems not to have weaned him away from the fundamentals of Benthamite thought. In France the thought of such men as Comte de Saint-Simon, Louis Blanc, Auguste Comte, though some of the externals of romantic fashion are visible among them, is, on the whole, along with that of the French Left generally, true to the traditions of the *philosophes.* Even in Germany, a philosopher such as Ludwig Feuerbach asserts the unromantic doctrines of materialism; and Marxism itself, though it shows romantic marks—the concept of the dialectic, derived of course from G. W. F. Hegel, is essentially romantic in its insistence on change as an overcoming of contradictions—is nonetheless committed to an optimistic and very eighteenth-century stand on the rational organization of man and society.

The romantic generation was indeed very conscious of breaking sharply with its parents and grandparents. Few breaks between cultural generations in the West have been more vigorously asserted than this one. The romantic youth absorbed in the depths of William Wordsworth's *Prelude,* or Vicomte Chateaubriand's *Génie du Christianisme,* or Johann Wolfgang von Goethe's *Faust* felt nothing but contempt for the abstract ideas and the confined tastes of his shallow Voltairean grandfather. To a

surprising extent, the fashionable Romantic was—or claimed to be—in all things the opposite of the Enlightened. Yet our own generation can hardly avoid holding that the romantic rebellion against its parent was in itself a proof of the filial relation between Romanticism and Enlightenment. Not only were the ideas of men like Jean-Jacques Rousseau, Vico, Lessing, and even Denis Diderot, all of whom lived at the height of the Enlightenment, seminal to all later Romanticism, but both Enlightenment and Romanticism shared much—a belief in process, change, if not actually progress, a belief in the possibilities of manipulating the environment, indeed a fundamental and very modern relativism never really transcended in the search for eternal verities. Both, whatever their metaphysical position on the problem of determinism, in practice displayed a firm conviction that things not only change, but that they can be changed by human effort. Of many specific doctrines—primitivism, for instance, or individualism in ethics and politics—it is hard to decide whether they are more characteristic of enlightened or of romantic thought.

SOME SPECIFIC ROMANTICISMS: ART AND LETTERS

The romantic touch is extremely visible in all the arts, from painting through architecture to interior decoration. Bright colors, or soft and fuzzy ones; exotic themes, Oriental scenes; crowded and action-filled historical paintings—concretely, almost any canvas by Eugène Delacroix—set romantic painting off from the sculptured Roman figures of David. And yet, to point up the coexistence of the romantic and the classical throughout the period, the sharp outlines, the measured realism—the Romantic would hold, the conventionality—of the portraits by Jean-Auguste-Dominique Ingres, who survived until 1867, outdo David's in classical firmness. The great romantic style in architecture was the neo-Gothic, itself a manifestation of the romantic rehabilitation of everything medieval that had been held in contempt by the Renaissance and Enlightenment. Yet Neo-Gothic was never a dominant style, not even in the Nordic lands; moreover, it soon fell into a most unmedieval and unromantic regularity and repetitiveness of detail. But Romanticism did rescue from the neglect in which they had long been left the great medieval cathedrals. In the decorative arts romantic tastes were extremely eclectic, fond of the exotic, addicted to rich dark woods and, in the climax of the Victorian drawing room, to a clutter of display wholly dependent on the existence of inexpensive domestic labor. In music, the romantic at its extreme went in for program music, birdcalls and thunderstorms, vast orchestras, and appropriate dissonances. The difference between the music of Joseph Haydn or Wolfgang Amadeus Mozart and that of Hector Berlioz or Wagner, like that between the painting of David and that of Delacroix, is obvious to the most untutored.

Poetry, the novel, and history are the great romantic literary genres, and in all of them the romantic syndrome is readily recognized. Although Goethe was a complex personality who was frequently in conflict with contemporary representatives of the romantic movement, his *Faust* is in itself a masterly summary of romantic themes: revolt against the dullness, the narrowness of rationalism ("gray dear friend is all theory, green only life's eternal tree"), striving for *etwas mehr,* for the infinite (the essential theme of Faust's bargain with Mephistopheles); contempt for the Philistine, the literal-minded ordinary man (the walk with Wagner); primitivism (Gretchen's innocence); ambivalence ("Two souls, alas, live in my breast"); and much else, right on to the final *chorus mysticus* of Part II. Indeed, this last is a fine touchstone; anyone who finds it nonsense or at least unpalatable is definitely not Romantic:

> Alles Vergängliche
> Ist nur ein Gleichnis;
> Das Unzulängliche
> Hier wird's Ereignis;
> Das Unbeschreibliche
> Hier ist's getan;
> Das Ewig-Weibliche
> Zieht uns hinan.

The three English Lake Poets, Wordsworth, Coleridge, and Robert Southey, together pretty well cover the romantic range; and Wordsworth's "The Tables Turned" ("One impulse from a vernal wood," "We murder to dissect," "Enough of science and of art;/Close up those barren leaves") states the central position of the romantic *Weltanschauung* almost as neatly as Goethe's *Gefühl ist alles.* One more figure, one more complex of themes, is needed to round out our concept of romantic poetry: This is the unhappy, misunderstood, heroic Promethean figure, half Shelley and all Lord Byron. In terms of sheer educated fashion, Byron and his whole train of European congeners (imitators would be an unfair word here)—Alfred de Musset, Alfred-Victor de Vigny, Giacomo Leopardi, José de Espronceda, Mikhail Lermontov, and the rest—may stand for the romantic poet.

Forerunners of the romantic novel are clear in the eighteenth-century "Gothic" novel, such as those of Ann Radcliffe (so charmingly satirized by the nonromantic

Jane Austen in *Northanger Abbey*); in the sentimental novel, such as Rousseau's *Nouvelle Héloïse*; and in the psychological novel of disturbed and disturbing love, such as Pierre-Ambroise-François Choderlos de Laclos's *Liaisons dangereuses* and the novels of the Marquis de Sade. The psychological novel reaches its best in the work of Stendhal, whose heroes foreshadow a long line of adventurers of soul and body, a line by no means extinct today. Yet in terms of the wider public of romantic fashion, Walter Scott's Waverley novels were the great success of their day. They carried their audience back into a simpler, more varied, more interesting past than the present of the Industrial Revolution. They exemplified that other inheritance from the German side of the Enlightenment, the theme, best marked in Johann Gottfried Herder, of organic historic growth of a folk spirit, a folk character, a product of time, not a product of the planning, present-bound intellect. One lost one's self in Sir Walter's pages, became one with one's own best past. We are a long way from Henry St. John Bolingbroke's definition of history as "philosophy teaching by examples."

HISTORY AND POLITICAL THOUGHT

The writing of serious history received a great impetus from the romantic movement, and in particular from Scott's work. Augustin Thierry, Jules Michelet, the Heidelberg school in Germany; in England Henry Hallam, indeed T. Macaulay, by no means a Romantic in temperament; and in the United States the great New England school of W. H. Prescott, J. L. Motley, and Francis Parkman wrote history for a wide reading public, history with narrative force and movement, history with a message of patriotism, of identification with a folk, yet also history carefully reconstructed by painstaking research. The historian and the critic of art and literature insisted on one of the great romantic themes: continuity, the continuity of life and flow, growth, development; a process, to the Romantic, always denatured, indeed destroyed, by the dividing analytical mind ("We murder to dissect").

The complexities and difficulties of generalizing about Romanticism come out most clearly, perhaps, in the field of political thought. You can, of course, always construct a pair of Procrustean beds: a conservative bed for the Romantics; a liberal, progressive, or radical bed for the Enlightened. Edmund Burke and Scott can be squeezed into the first, Thomas Paine, W. Godwin, Thomas Jefferson into the second. But the trouble is that you can quite plausibly switch the beds, putting the Romantics into the liberal or progressive bed, the Enlightened into the conservative bed. Shelley, Byron, Benjamin Constant can go into the first; Voltaire (surely no democrat), John Adams, the *idéologues* who rallied to Napoleon can go into the second. But Victor Hugo would have to be divided, his younger self put into the conservative, his older self into the liberal bed.

Critics have indeed tried to fix Romanticism on one side or the other in politics, and—given their premises—not without some success. Probably in the balance Romanticism has worked toward the growth of modern democracy, toward a belief in progress and toward "liberty, equality, fraternity," toward the open society—toward much, in fact, that gets its start from the rationalists of the Enlightenment. Yet the Burkean belief in human fallibility, human blindness of passion, and in tradition-enshrined institutional dikes to restrain these anarchic thrusts (dikes not to be tampered with by the intellect), as well as belief in the folk, in an organic society not the product of planning, is surely also congruous with much of Romanticism. So too is the anti-intellectual strain that comes out much later in theories of racism, elitism, *Blut und Boden,* in Nazism and Fascism.

PHILOSOPHY

Romanticism is more than a fashion in arts and letters, more than an approach to political problems: It is a philosophy, or better, a set of philosophies loosely tied together if only by their common rejection of eighteenth-century rationalism, of refusal to line up, shall we say, on the Locke-Hume axis. Arthur Schopenhauer is the arch Romantic, the extreme Romantic, among formal philosophers. The world of phenomena, of sense perception, is to him unreal; the will that moves the universe is real enough, but certainly is not rationally knowable by those it moves; this will is blind, shapeless, evil; life, merely phenomenal though it is, is still for us all painful, wearisome, a long unhappy voyage (note the metaphors of movement); Schopenhauer seems at times to hold that a nirvana of surcease is perhaps attainable; at any rate, this life is hopeless.

Romantic pessimism is not, however, the central theme of philosophy in these years. Hegel, at bottom an optimist, is much more central. In a sense, the great romantic philosophers, most of them Germans, go back to Immanuel Kant, who always thought of himself as firmly enlightened, and whose brief *Was ist Aufklärung?* is one of the landmarks of the century of prose and reason. The romantic seedling in Kant, however, is his distinction between the noumenal and the phenomenal, and his resolution of the dualism by what amounts to intuition or faith. Johann Gottlieb Fichte and Friedrich Schleierma-

cher and the rest developed this essentially romantic reliance on a "faculty" transcending common calculating logic. Hegel accepted, and gave his own turn to, this very old dualism of spirit-matter, real-unreal, and sought to bring them together by his famous and influential concept of the dialectic of thesis–antithesis–synthesis. The dialectic in all its forms displays a most nineteenth-century and romantic general bias toward historicism, process, development—but such a process seen teleologically as an end, a purpose. For Schopenhauer, there was no end save extinction. But for Hegel there was an end, a vague one, a Germanic eternal peace in which change somehow turns out to be, in the workings of the World-Spirit, the real form of permanence.

These philosophers, trained and subtle professionals whom we have no doubt traduced in this brief account, are less definitely to be associated with Romanticism as a broad cultural movement than the popularizers, the essayists, the preachers. To many devotees of Thomas Carlyle, Ralph Waldo Emerson, and John Ruskin, and to those who listened to bumblers like Bronson Alcott, romantic philosophy became fashionable transcendentalism, an agreeable summary of the less difficult phases of romantic thought—contempt for the rationalist side of the eighteenth century (indeed, blindness to the existence of any other side of that century), exaltation of intuition, spirit, sensibility, imagination, faith, the unmeasurable, the infinite, the wordless—or at least, only the noblest sounding words. This sort of Romanticism was indeed a solace and an escape, an escape from the difficult and unlovely works that science, technology, and industry were building. But it is by no means the whole of Romanticism, which as a spiritual spur to precisely the kind of invention, adventure, and enterprise, to the preoccupation with change and growth, that was building the new world of the nineteenth century, must be seen as having played, and as continuing to play, an essential part, along with the rational and scientific inheritance from the eighteenth-century Enlightenment, in building our own world of today.

See also Art, Expression in; Bergson, Henri; Blake, William; Bolingbroke, Henry St. John; Burke, Edmund; Carlyle, Thomas; Chateaubriand, François René de; Coleridge, Samuel Taylor; Comte, Auguste; Diderot, Denis; Emerson, Ralph Waldo; Enlightenment; Feuerbach, Ludwig Andreas; Fichte, Johann Gottlieb; Godwin, William; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hume, David; James, William; Jefferson, Thomas; Leopardi, Count Giacomo; Lessing, Gotthold Ephraim; Locke, John; Lovejoy, Arthur Oncken; Mill, John Stuart; Neo-Kantianism; New England Transcendentalism; Nietzsche, Friedrich; Paine, Thomas; Pessimism and Optimism; Rousseau, Jean-Jacques; Ruskin, John; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Schleiermacher, Friedrich Daniel Ernst; Schopenhauer, Arthur; Shelley, Percy Bysshe; Spengler, Oswald; Staël-Holstein, Anne Louise Germaine Necker, Baronne de; Vico, Giambattista; Whitehead, Alfred North.

Bibliography

Romanticism is not only a complex cluster of ideas; it is one that arouses strong feelings among critics and historians, and that has had its ups and downs in the estimation of the various cultural generations since the late eighteenth century. The following should set the reader on his way through these thickets of critical and philosophical discussion of Romanticism.

Howard Hugo, ed., *The Romantic Reader* (New York: Viking Press, 1957), is an admirable anthology with a good bibliography of works in English and a useful prologue, "What the Romantics Said about Romanticism." W. T. Jones, *The Romantic Syndrome* (The Hague: Nijhoff, 1961), presents a very suggestive analysis, helpful for all study of the history of ideas. Jacques Barzun, *Classic, Romantic and Modern* (Boston: Little, Brown, 1961), contains the ablest defense of Romanticism; see the section "Romantic—A Sampling of Modern Usage" (pp. 155–168). G. A. Borgese, "Romanticism," *Encyclopaedia of the Social Sciences* (New York, 1934), Vol. XIII (VII), a remarkably rich brief account, with full bibliographies up to 1934 in all Western tongues, is sympathetic. Irving Babbitt, *Rousseau and Romanticism* (Boston: Houghton Mifflin, 1919), is still the sharpest attack on Romanticism. A. O. Lovejoy, *Essays in the History of Ideas* (Baltimore: Johns Hopkins University Press, 1948), contains several pertinent essays, especially one titled "On the Discrimination of Romanticisms." Sir Maurice Bowra, *The Romantic Imagination* (Cambridge, MA: Harvard University Press, 1949), is a graceful essay by a distinguished English scholar and critic. Walter Jackson Bate, *From Classic to Romantic* (Cambridge, MA: Harvard University Press, 1946), is one of the best studies of the complex interweaving of classic and romantic in English literature. Ricarda Huch, *Blüthezeit der Romantik,* 12th ed. (Leipzig, 1922), and *Ausbreitung und Verfall der Romantik,* 10th ed. (Leipzig, 1922), are sympathetic and graceful accounts of the German Romantics. Pierre Lasserre, *Le Romantisme français* (Paris, 1907), is an unsympathetic account of the French Romantics.

Crane Brinton (1967)

ROMANTICISM [ADDENDUM]

When romanticism is understood broadly, as referring to a major development in European thought and culture

since the turn of the nineteenth century that shows itself distinctly in the spheres of art, historical writing, and political thought, the concept has only a limited role to play in the history of philosophy: Certain very general notions—an emphasis on agency, expression, the cognitive dimension of affect, and the potential of human beings to become genuine wholes—can be described as manifestations of romanticism in philosophy, but the term does not serve to pick out any more determinate set of philosophical commitments.

Here, as with modernism, is a category that is indispensable for general intellectual history, but lacks equivalent value in the history of philosophy. Where the concept does achieve significant purpose in the history of philosophy is in its much narrower application to the group of thinkers based in Jena at the very end of the eighteenth century known as the (early) German romantics, or *Frühromantik*, whose activity centered on production of the *Athenäum*, a journal whose historical importance far exceeds its short life span. Friedrich von Schlegel and Novalis (the pen name of Friedrich von Hardenberg) comprise the philosophical core of German romanticism, with F. W. J. von Schelling and F. D. E. Schleiermacher in close, albeit temporary and qualified, association. J. C. F. Hölderlin—like Novalis, a major German lyric poet—did not belong to the group in Jena but is considered properly as belonging to the same philosophical tendency as Schlegel and Novalis.

Philosophical understanding of the German romantics has been obstructed by the fragmentary form of much of their output, and the literary concern of the movement taken as a ground for assuming its importance to lie outside philosophy, but more recent work, above all by Manfred Frank (b. 1945), Ernst Behler (b. 1928), and Frederick Beiser (b. 1949), has revealed the distinctiveness and importance of the philosophical outlook formulated by the German romantics in the context of the problems and issues facing post-Kantian philosophy. The problems of Immanuel Kant's legacy revolved in the first place around the perceived incompleteness of Kant's transcendental or critical philosophy, which was considered to have opened up a new range of intellectual possibilities and yet to require further development for it to fulfil its emancipatory promise and thereby meet the demands of the age. Johann Gottlieb Fichte's attempt to do exactly this in his *Wissenschaftslehre* held the attention of, but failed to convince, the German romantics, who accepted the rational necessity of seeking to construct a self-grounding philosophical system but believed themselves to have achieved insight into the reasons why this ideal cannot be realized and must remain an infinite task.

Schlegel's original and influential conception of irony as not merely a literary trope, but rather a corollary of the structure of reflection that, having achieved critical freedom, cannot bring itself to a halt, was developed in part to rationalize this complex attitude toward the ideal of philosophical systematicity. The reorientation proposed by the German romantics in place of Fichte's *Wissenschaftslehre* centered on a novel and very high valorization of art and the aesthetic. This move, far from signalling an aestheticist turning away from the philosophical tasks and the social and political realities that occupied Kant and Fichte, was envisaged as engaging with the full spectrum of philosophical, practical, and cultural problems. The key to the importance ascribed by the German romantics to art—at least, to that art which possesses the qualities of what they called *Poesie* (*romantische*)—lay in its supreme exemplification of true (organic) unity, its synthetic relation to the metaphysical oppositions that structure human existence and reality at large, and its embodiment or symbolization of freedom.

Both this conception of art and the romantics' claim for its practical importance show the influence of Schiller's *Letters on Aesthetic Education*, but the German romantics projected the concept of art and the aesthetic, in a way that Schiller had not, well beyond the sphere of works of art in the strict sense. Schelling's account in the final part of his *System of Transcendental Idealism* of art as what he calls the only true organ and document of philosophy provided one formulation of the German romantic idea that art is philosophically preeminent, and the *Naturphilosophie* that Schelling developed in the late 1790s, which attributes organic status to nature as a whole, and disputes the primacy of mechanism over teleology maintained in modern philosophy even by Kant, falls equally into line with the German romantic program.

In the ethical and political sphere, the German romantics sought to achieve recognition for the claims of personal individuality while at the same time urging the pursuit of organic wholeness in collective life, in opposition to Kant's ethical universalism and political atomism, yet without any intention of contradicting Kant's modern affirmation of freedom. Schleiermacher's ethical theory, though it was composed some years after the dissolution of the romantic circle, may be regarded as giving systematic shape to at least some German romantic ethical insights, in the same way that his earlier, highly successful

work, *On Religion*, stood in close accord with the German romantic intention to recreate religious forms.

German romanticism has affinities with positions that had been developed earlier by J. G. Hamann, J. G. Herder, and F. H. Jacobi in their critique of the German Enlightenment, and some commentators have suggested that it also prefigures deconstruction and postmodern philosophy, on account of its skepticism regarding the attainability of final philosophical truth. This view runs a risk of anachronism, however, for while it is true that German romanticism diverges from the three great developments of German idealism, it nevertheless remains committed to an ideal of rationality and retains many of the idealistic, not to say Platonistic, elements that are present in the philosophies of Fichte, Schelling, and Hegel. Indicative in this regard is the fact that the sharp criticism made by Hegel of German romanticism—which collapses, Hegel believes, into hyper-subjectivism and arbitrariness—is premised on an understanding of the movement as having grasped, without giving adequate form to, important philosophical truths.

See also Enlightenment; Fichte, Johann Gottlieb; German Philosophy; Hamann, Johann Georg; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hölderlin, Johann Christian Friedrich; Jacobi, Friedrich Heinrich; Kant, Immanuel; Modernism and Postmodernism; Novalis; Schelling, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

COMMENTARY

Behler, Ernst. *German Romantic Literary Theory*. Cambridge, MA; New York: Cambridge University Press, 1993.

Beiser, Frederick. "Absolute Idealism. " In *German Idealism: The Struggle Against Subjectivism 1781–1801*. Cambridge, MA: Harvard University Press, 2002.

Beiser, Frederick. *The Romantic Imperative: The Concept of Early German Romanticism*. Cambridge, MA: Harvard University Press, 2003.

Craig, Edward. "The Metaphysics of the Romantic Era. " In *The Mind of God and the Works of Man*. Oxford, U.K.: Clarendon Press, 1987.

Frank, Manfred. *The Philosophical Foundations of Early German Romanticism*. Translated by Elizabeth Millán-Zaibert. Albany: State University of New York Press, 2004.

Frank, Manfred. *"Unendliche Annäherung": Die Anfänge der philosophischen Frühromantik*. Frankfurt am Main: Suhrkamp, 1997.

Henrich, Dieter. *Course of Remembrance and Other Essays on Hölderlin*, edited by Eckart Förster. Stanford, CA: Stanford University Press, 1997.

Lacoue-Labarthe, Philippe, and Jean-Luc Nancy. *The Literary Absolute: The Theory of Literature in German Romanticism*. Translated by Philip Barnard and Cheryl Lester. Albany: State University of New York Press, 1988.

Richards, Robert J. *The Romantic Conception of Life: Science and Philosophy in the Age of Goethe*. Chicago: University of Chicago Press, 2002.

Seyhan, Azade. *Representation and Its Discontents: The Critical Legacy of German Romanticism*. Berkeley: University of California Press, 1992.

WRITINGS IN TRANSLATION

Beiser, Frederick, ed. *The Early Political Writings of the German Romantics*. Cambridge, MA; New York: Cambridge University Press, 1996.

Bernstein, J. M., ed. *Classic and Romantic German Aesthetics*. Cambridge, U.K.; New York: Cambridge University Press, 2003.

Hölderlin, Friedrich. *Friedrich Hölderlin: Essays and Letters on Theory*. Translated by Thomas Pfau. Albany: State University of New York Press, 1988.

Novalis. *Fichte Studies/Novalis*. Cambridge, MA: Cambridge University Press, 2003.

Novalis. *Philosophical Writings* Translated by Margaret Mahony Stoljar. Albany: State University of New York Press, 1997.

Schlegel, Friedrich von. *Philosophical Fragments/Friedrich Schlegel*. Translated by Peter Firchow. Minneapolis: University of Minnesota Press, 1991.

Schulte-Sasse, Jochen, ed. *Theory as Practice: A Critical Anthology of Early German Romantic Writings*. Translated by Haynes Horne. Minneapolis: University of Minnesota Press, 1997.

Wheeler, Kathleen, ed. *German Aesthetic and Literary Criticism: The Romantic Ironists and Goethe*. Cambridge U.K.: New York: Cambridge University Press, 1984.

Willson, A.L., ed. *German Romantic Criticism*. New York: Continuum, 1982.

Sebastian Gardner (2005)

ROMERO, FRANCISCO

(1891–1962)

Francisco Romero, the Argentine philosopher of transcendence, was born in Seville, Spain, but moved to Argentina as a child. After military and literary careers he turned to philosophy, joining the faculty of the University of Buenos Aires in 1928 and of La Plata in 1929. He renounced his academic posts in 1946 in protest against the government of Juan Perón but resumed them in 1955. Because of his conceptual discipline, scope, originality of thought, and limpid clarity of style, Romero is considered one of the ablest and most satisfying of Latin American philosophers.

The idea of transcendence dominates and unifies Romero's metaphysics and theories of knowledge and

values. Transcendence implies at least the diversity achieved by passing beyond a given condition or limit and suggests a universal impetus or agency of such passage, an agency that may be purposive. Opposed to transcendence is immanence, which implies identity and containment within, or return to, a limit. Of the two major forms of transcendence, one is that relation of parts to each other in a structural whole by which novel characteristics emerge that were only latent in the parts considered separately. The other form of transcendence is change and, in particular, evolution in the creative and vitalistic sense of Henri Bergson. Its immanent reduction occurs in the mechanistic evolutionary views of Charles Darwin and Herbert Spencer.

Romero identified reason with immanence; experience, in a broad sense, is related to transcendence. Reason may be either intuitive or discursive. In either case it demands identity and transparency. Identity is found in homogeneity and in permanence; it leads reason to the mechanistic conception of atoms that are similar in kind, endure in time, and are governed by causal laws that presuppose the identification of effects with their causes. Transparency, or clarity, is found in forms emptied of content and in the space in which atoms move and with which they tend to be identified.

Reason is formal only and has no avenue of its own to reality and concrete fact. It is not identical with intelligence, which may criticize it. Where reason fails, experience succeeds. Experience supplies a datum by which knowing must be guided. The objects of experience are not sense data and perceptual objects alone, but also essences and values. In addition, Romero held open the possibility of a metaphysical experience of something ultimate and noumenal but subject to connection with ordinary experience and its phenomenal objects.

Romero divided phenomena into four strata, of which each level is a ground for the next and has greater scope for transcendence than the preceding level. The physical level, that of space and moving atoms, is most pervaded with immanence, but the shift in physical theory from the rigid corpuscle to the *foco activísimo* means a greater emphasis on the role of transcendence even on this level. The vital level is characterized by true duration, a factor of transcendence. The psychical level involves consciousness, which intends, or transcends toward, an object, but there is a countering immanence in the egocentric tendency of the human individual to absorb the object into his own forms and needs. On the spiritual level, the human person, rising above his egocentric needs and attaining a universal subjectivity, contemplates the object disinterestedly in the sphere of knowing and conducts himself altruistically and with regard to general principles in the sphere of action. On the spiritual level transcendence becomes absolute. The person is transcendence incarnate and unqualified. Each level contains and is supported by transcendence, but each is unique and irreducible.

Romero, proceeding cautiously and with an air of hypothesis, proposed that Arthur Schopenhauer and Bergson were not wrong in positing a metaphysical datum, but that they misconstrued it. Schopenhauer's will and Bergson's vital impulse are forms of transcendence, which is a more general and basic being than either. Romero did not try to sketch the nature of this being, but he appears to have thought of it as a universal impulse at work in every level of phenomenal transcendence, an impetus that is the essence of reality, the source of value, and possibly the spirit's point of flower, which this being intended from the beginning.

See also Bergson, Henri; Change; Darwin, Charles Robert; Latin American Philosophy; Nature, Philosophical Ideas of; Schopenhauer, Arthur; Spencer, Herbert.

Bibliography

WORKS BY ROMERO

Lógica. Written with E. Pucciarelli. Buenos Aires, 1938.

Filosofía de la persona. Buenos Aires, 1944.

Papeles para una filosofía. Buenos Aires: Editorial Losada, 1945.

Teoría del hombre. Buenos Aires: Editorial Losada, 1952. Translated by W. F. Cooper as *Theory of Man.* Berkeley: University of California Press, 1964.

Historia de la filosofía moderna. Mexico City: Fondo de Cultura Económica, 1959.

Arthur Berndtson (1967)

RORETZ, KARL

(1881–1967)

Karl Roretz, the Austrian epistemologist, philosopher of culture, and aesthetician, was born at Schloss Breiteneich. He studied law, and later philosophy, at the University of Vienna, receiving his doctorate in 1906 with the dissertation "The Problem of Empathy in Modern Aesthetics." In 1922 Roretz became a *Privatdozent* at the university and taught history of modern philosophy until 1938, when he ceased lecturing after the Nazi takeover of Austria. He resumed lecturing in 1945 and continued until his retire-

ment in 1951.

As an epistemologist, Roretz espoused a "critical positivism," a philosophy whose foundation is both scientific and, in Immanuel Kant's sense, criticist. The outstanding features of his thought are critical reflection, skeptical rationality, intellectual honesty, and independence of mind. He rejected dogmatism and unsupported metaphysical speculation. Like Hans Vaihinger, he regarded metaphysical concepts as self-contradictory fictions. Thus, Roretz held, metaphysics lacks any purely logical meaning.

Roretz's major work, *An den Quellen unseres Denkens* (Vienna, 1937), contains his most acute epistemological analyses. In this monograph he studied "vital concepts," concepts in whose formation an element of will or an element of value plays a decisive part and whose definition is therefore preceded by a decision. Among such concepts are those of art, of ethics, of popular education, and of the slave trade.

Roretz's elegant and penetrating psychological analyses of culture and his critical analyses of values deserve particular consideration. The decline of spiritual values, he contended, is due to internal degeneration or disintegration within the person and the society, and only seldom to external pressure. He also studied the genesis and structure of mass psychological phenomena ("mass, illnesses," *Massenerkrankungen*) in religion, politics, economics, art, fashion, and sports—notably such extremely dangerous religious and other spiritual "epidemics" as belief in vampires and devils, witch-hunting, and racial persecution.

As a philosopher of culture, Roretz felt most akin to Friedrich Nietzsche. Like Nietzsche, he believed in life with a deep conviction. But Roretz's view of life was Kantian, and the meaning of life for him consisted in working at the problems life poses. He advocated a philosophy that interpreted reality from an aesthetic point of view. Such a philosophy, he held, provides an orientation toward life and the world that is biologically optimal. The world appears, in this view, as a drama without metaphysical supramundane or transmundane galleries to which it must play. Roretz professed a deep joy in the variegated splendor of the world. "The meaning of the world," he wrote, "is an aesthetic meaning."

In his studies of what he called intellectual-aesthetic values—aesthetic effects bound up with specific achievements of thought, as in mathematics, strategy, or chess—Roretz made important contributions to aesthetics itself.

His interest in ethical problems was equally great. A convinced humanist and democrat, he supported the Ethical Culture movement and strove for a secular ethics independent of any metaphysical or religious assumptions.

See also Aesthetics, History of; Culture and Civilization; Epistemology; Kant, Immanuel; Nietzsche, Friedrich; Value and Valuation.

Bibliography

Additional works by Roretz include *Zur Analyse von Kants Philosophie des Organischen* (Vienna, 1922); *Religiöse Epidemien* (Munich, 1925); *Die Metaphysik—eine Fiktion* (Vienna, 1927); and *Ziele und Wege philosophischen Denkens* (Wien: F. Deuticke, 1976).

Also see Franz Austeda, *Dem österreichischen Philosophen K. Roretz zum achtzigsten Geburtstag* (Vienna, 1961), which contains a complete bibliography.

Franz Austeda (1967)
Translated by Albert E. Blumberg
Bibliography updated by Michael Farmer(2005)

RORTY, RICHARD
(1931–)

An American philosopher and pragmatist, Rorty is among the most widely discussed and controversial philosophers at the turn of the twenty-first century. A New Yorker by birth, Richard Rorty was educated at the University of Chicago (1946–1952) and at Yale (1952–1956) where he received his doctorate in philosophy. After brief flirtations with Platonism and the work of A. N. Whitehead, Rorty's more mature interests began to form at the end of his military service in 1958, at which point he began serious study of the philosophers who would later number among his chief influences: Wilfrid Sellars, Ludwig Wittgenstein, Martin Heidegger, John Dewey, and W.V. O. Quine.

EARLY PERIOD

Rorty's early work in analytic philosophy, sometimes thought to represent a completely distinct period, is in fact touched by two themes that resurface throughout his career. The first theme is anti-Cartesianism about the mind and knowledge. In a series of papers written during the 1960s Rorty was the first to develop a subsequently contentious theory in the philosophy of mind—eliminative materialism, which holds that the mind and mental

states are theoretical, and hence dispensable, constructions.

The second theme is an abiding concern with the function and importance of philosophy. Again, this theme appears early on, particularly in Rorty's 1967 introduction to *The Linguistic Turn*, a collection of essays on analytic philosophy of language. In his introduction, Rorty praised analytic philosophy for knocking the entire philosophical tradition on its heels—a sentiment that he would later characterize as naïve. In subsequent work, Rorty came to believe that mainstream Anglo-American philosophy of language makes many of the same mistakes as the intellectual traditions he had earlier taken it to supplant.

ANTIREPRESENTATIONALISM

This latter sentiment first emerges in Rorty's seminal book, *Philosophy and The Mirror of Nature* (hereafter: PMN). Originally published in 1979, and at once hailed and denounced as a critique of analytic philosophy, the book brings together Rorty's hostility to Cartesianism with a positive vision about the nature and limits of philosophy.

At its core, PMN is a sustained attack on representationalism. According to representationalism—which Rorty argues is a largely unquestioned assumption of Cartesian and Kantian philosophy—the mind is a device for representing the world and knowledge is accurate representation. Rorty holds the representationalism responsible for two major philosophical mistakes: a false dualism of mind and body, and a bankrupt foundationalist picture of knowledge, which holds that all knowledge must sit on the foundation stones of intrinsically accurate privileged representations. Moreover, Rorty takes representationalism to paint a misleading picture of philosophy's importance—as a master discipline that judges whether the claims of science, morality, or art can represent reality.

In Rorty's view, twentieth-century linguistic philosophy continued to assume representationalism (and its mistakes) in a linguistic guise—an opinion he sees as shared by philosophers as diverse as Dewey, Quine, Sellars, Wittgenstein, and Donald Davidson. Following Quine and Wittgenstein, for example, he argues that the notions of meaning and analyticity are mere linguistic shadows of the privileged mental representations of the early moderns. And with Sellars, he rejects "the given,"or theoretically innocent sense-experience, as a myth. His moral: language, like the mind, should not be understood as a device for representing a ready-made world.

In opposition to representationalism, Rorty suggested in PMN that we should adopt what he called *epistemological behaviorism*, and explain epistemic rationality and justification in terms of what our society will let us say, rather than the other way around. The thought is that there is no mystery about how the mind represents the world. The very idea of such representation is a fable; what claims we accept as knowledge depends not on how well they mirror the world but on how well they hang together with what else we already accept. Accordingly, Rorty concluded, philosophy has nothing distinctive to offer about knowledge. To learn why we accept what we accept we must turn to biology, psychology, and sociology; Charles Darwin will have more to teach us about the mind than René Descartes. The philosopher's role was instead therapeutic—to cure of us intellectual maladies—and revisionary so as to convince us to engage in new forms of conversations.

PRAGMATISM AND TRUTH

Rorty has continually emphasized that his view is a form of pragmatism—particularly the pragmatism of Dewey. And much like the classical pragmatists before him, he sees his debate with the representationalist as coming down to a debate over how to understand truth.

Yet Rorty's own views on truth shifted in subtle ways over the course of his career. He always rejected the correspondence theory of truth, according to which a statement is true just when it corresponds to the facts. But in his earlier work, Rorty was tempted to follow the classical pragmatists and define truth in terms of justification or warranted assertibility. Truth, on his version of the view, simply is what we are justified in believing in light of our cultural practices. But in later works, Rorty has come to see this position as another misguided attempt to uncover the secret nature of truth. The contemporary pragmatist, Rorty argues, should instead simply reject the idea that truth has any nature at all. Truth is not the sort of thing that can be defined—not because its nature is mysterious or ineffable, but because there is nothing general and informative one can say about what is in common between "Snow is white," "Two and two are four" and "Democracy is a better form of government than tyranny." There simply is no metaphysically substantive property of truth that some propositions have and others lack.

Rorty argues that adopting this attitude toward truth has several important consequences. In his later work, for example, he has particularly emphasized that, for the pragmatist, truth is not a goal of inquiry. According to

Rorty, something can only be a goal if we can recognize when we have reached it. But whenever we check to see whether our beliefs are true, we can only discover whether they are justified or unjustified. Thus we should give up on the idea that we are aiming at truth; instead, Rorty says, echoing Davidson, we should see ourselves as aiming only at honest justification. And for Rorty, justification is a practical matter—what beliefs we find justified depends on whether we can use them in achieving the aims of our culture. Nonetheless, truth is not reducible to what our immediate community finds useful because one important function of the word *true* in our language is to remind us that what may be practically justified to some audiences may not be justified to all.

Rorty's views on truth have drawn considerable criticism. He is often derided as advocating a naïve form of cultural relativism. But Rorty insists that it is as misleading to describe him as a relativist as it would be to describe him as a realist. In Rorty's eyes, the realist and the relativist commit linked sins: the realist by taking the world to be ready-made, the relativist by thinking it is made by us. From the Rortyian perspective, we should instead take truth making—whether understood in a realist or relativist fashion—as simply a metaphor that should be given up. Consequently, Rorty might be better described as advocating a form of quietism about metaphysics and epistemology.

DEMOCRACY AND PHILOSOPHY

Rorty takes the failure of representationalism as linked with the failure of another enlightenment project: the project of grounding our political ideals in a common human nature. For Rorty, democratic, liberal government is a great achievement, but it is not an achievement whose value can be given a philosophical justification. Rather than trying to justify liberal democratic ideals philosophically, we should instead seek to ground our philosophical ideals in our democratic values. Thus we should stop searching for objective foundations and instead aim for solidarity with our fellows.

Rorty describes his positive political position as *liberal ironism*. It is liberal because it takes self-creation and freedom as central values. Individuals should be free from suffering and cruelty, but also free to create and live their own vision of the good life. But the Rortyian liberal also takes an ironic stance toward his own liberal commitments. He realizes that his values are contingent reflections of his own time and place, and not reflections of the values of the world itself. To those critics who protest that this position is too weak to offer sufficient defense against the tyrant, Rorty responds that philosophy is of no use against tyrrany anyway, and that those who believe that all is lost without appeal to the world's own true values are much like those nineteenth-century intellectuals who believed that without God, everything was permitted.

Rorty has sometimes been charged with no longer doing philosophy. And that charge is fair if one takes philosophy to be in the job of representing the world as it is in itself. But Rorty's own views encourage a different view of philosophy, according to which the job of the philosophy is not so much to discover the world's own language as he sometimes put it, but to invent new vocabularies and means of description. In this sense, Rorty's stance toward philosophy is Marxian: The goal of the philosopher is not to map the landscape as it is, but change how we see the world; to paint new landscapes, new pictures.

See also Epistemology; Philosophy of Mind; Pragmatism; Social and Political Philosophy.

Bibliography

PRIMARY WORKS

Philosophy and The Mirror of Nature. Princeton, NJ: Princeton University Press, 1979.

Consequences of Pragmatism. Minneapolis: University of Minnesota Press, 1982.

Contingency, Irony, and Solidarity. Cambridge, MA: Cambridge University Press, 1988.

Objectivity, Relativism, and Truth: Philosophical Papers I. Cambridge, MA: Cambridge University Press, 1991.

Essays on Heidegger and Others: Philosophical Papers II. Cambridge, MA: Cambridge University Press, 1991.

The Linguistic Turn. 2nd ed. Chicago: University of Chicago Press, 1992.

Truth and Progress: Philosophical Papers III. Cambridge, MA: Cambridge University Press, 1998.

SECONDARY WORKS

Brandom, Robert. *Rorty and His Critics*. Oxford: Blackwell Publishers, 2000.

Michael P. Lynch (2005)

ROSCELIN
(1050–1120)

Despite much scholarly effort, little is known about Roscelin of Compiègne. The only work that we can safely attribute to him is a letter he sent to Peter Abelard around 1119–1120. In this ill-tempered piece of writing Roscelin sets out to distinguish his position on the Trinity from

that which Abelard was developing in his *Theologia Summi Boni*.

The problem for Roscelin and Abelard is to give an account of the distinction between the persons of the Trinity compatible with the unity of God. Roscelin notes that he has to navigate here between two heresies: Sabellianism, requiring such a unity in the singular substance of God that the distinction between the persons can be only verbal, and Ariansm, which distinguishes the persons as greater and lesser so as to constitute three distinct gods. Roscelin, in effect, accused his former student of Sabellianism, and so contributed to Abelard's being called before the Council of Soissons in 1121 and required to burn a copy of his *Theologia Summi Boni* and confirm his orthodoxy.

Roscelin's position is that the names *Father*, *Son*, and *Holy Spirit* pick out distinct items but that these exist in God with the unity of likeness and equality. Unfortunately he invokes authority rather than arguing for his theory. It has recently been suggested by author Constant Mews, however, that Roscelin depends on an account of the semantics of names developed in contemporary writing on grammar.

Our remaining information about Roscelin comes from two unsympathetic reporters. Abelard, in a letter written around 1119, complains about Roscelin's attack on him and recalls that long before, at the Council of Soissons in 1092, Roscelin had been charged with tritheism. From Saint Anselm we have two letters from 1090–1092, and the treatise *De incarnatione verbi*, written after the Council, in which Roscelin is said to have maintained that the persons of the Trinity are as separate from one another as three angels, or three souls. This is certainly not what he claims in his letter to Abelard, which may thus represent a refinement of his theory in response to Anselm's objections. In *De incarnatione verbi* Anselm characterizes Roscelin as a heretic in logic who holds that universal substances are nothing more than "puffs of air made with the voice," who cannot distinguish a body from its color, or a soul from its wisdom, and cannot understand how human beings are one in species.

In the middle of the twelfth century Roscelin was said to have been the first to have upheld the doctrine of words (*sententia vocum*). From the information given, however, it is impossible to recover anything of his theory. We are told by John of Salisbury that he held that utterances themselves were genera and species. This is the position advocated in the *Dialectica* of Garlandus Compotista, written around 1115, which may provide our best guide to the views of Roscelin and those referred to at the time as the *Vocales*.

Appealing to Garlandus, Abelard's early writings, and various other texts, author John Marenbon has argued that vocalism in general and perhaps Roscelin's views in particular, developed out of what he calls the *in voce* reading of Porphyry's *Isagoge* and Aristotle's *Categories*. In this exegetical procedure, he suggests, theoretical commitment was suppressed in favor of reading the texts as simply about the relations between words.

Abelard confirms in his letter that Roscelin held that universals are in some sense words, and parodies him by saying he would have to read Scripture as claiming that Christ ate the expression *broiled fish* rather than the fish itself. It is unfortunately impossible to tell whether Abelard is constructing or reporting an argument when he reports, in his *Dialectica*, that as well as holding that species are words, Roscelin claimed that things do not have parts—so a wall is not part of a house. Perhaps what was really at issue were the questions that seem to have exercised Roscelin throughout his career: What counts as a thing and what is the nature of unity?

See also Abelard, Peter; Anselm, St.; Aristotle; Arius and Arianism; John of Salisbury; Medieval Philosophy; Porphyry; Universals, A Historical Survey.

Bibliography

Marenbon, John. "Life, Milieu, and Intellectual Contexts." In *The Cambridge Companion to Abelard*. Cambridge, U.K.: Cambridge University Press, 2004.

Mews, Constant. *Reason and Belief in the Age of Roscelin and Abelard*. Aldershot U.K.: Ashgate, 2002.

Christopher J. Martin (2005)

ROSENKRANZ, JOHANN KARL FRIEDRICH

(1805–1879)

Johann Karl Friedrich Rosenkranz, the German Hegelian philosopher, was born in Magdeburg. He entered the University of Berlin in 1824. Although he was to become G. W. F. Hegel's most devoted disciple, Rosenkranz was first drawn to Friedrich Schleiermacher; he heard only an occasional lecture by Hegel and was unimpressed. He began reading Hegel as a student at Halle in 1826 and the following year came under the influence of Karl Daub (1765–1836), a Hegelian theologian at Heidelberg. As a *Privatdozent* and extraordinary professor at Halle,

Rosenkranz participated actively in the Hegelian circle there. Called to Berlin, he struck up a friendship with Hegel and joined his birthday celebration a few weeks before Hegel died of cholera in 1831. Rosenkranz himself was stricken almost fatally with the disease, reflecting, as he later reported, that this was carrying discipleship entirely too far. In 1833 he succeeded Johann Friedrich Herbart as professor of philosophy at the University of Königsberg, where he remained until his death except for a brief political career in Berlin during the revolutionary crisis of 1848/1849.

Rosenkranz wrote over forty substantial works, on systematic philosophy, aesthetics, theology, logic, psychology, literary history, pedagogics, philosophical history and biography, and political and social theory. He also composed poetry and contributed articles on current issues to the newspapers.

Rosenkranz defended the Hegelian system as the authentic expression of the German spirit and the fulfillment of German philosophy. He attacked the "one-sidedness" of the Hegelian left-wing and denied that there was any irreconcilable conflict between Hegel and other major German thinkers, such as Schleiermacher and Immanuel Kant. Other Hegelians charged that Rosenkranz had interpreted Hegel in a Kantian way, maintaining the duality between thought and being and between the ideal and the actual. Certainly in his view the ideal was always in tension with existing conditions, although it constituted their *telos* and guiding norm. In practice, for example, he held that the church should be independent of the state; because Christianity embodies the highest ideal, the church must be free to hold before the culture its most ideal possibilities. He argued on similar grounds for the freedom of the university from political control.

Underlying religious, political, and intellectual life alike, however, was the *Volksgeist* ("spirit of a people"), interpreted more romantically than in Hegel. It is not the result of the cultural process but the distinctive psychic root of a particular people that gives the people unity as a nation and seeks expression in a total cultural life. A people is free to the extent that it fully embodies this spirit; genuine "public opinion" is the self-understanding of a free people. As a consequence, although Rosenkranz gave humankind precedence over the nation in principle and affirmed the Kantian vision of universal peace, he opposed the supranationalism of the left-wing Hegelians; moreover, he regarded their revolutionary aims as empty abstractions, without relevance to "realities" or to the concrete aspirations of any people, and productive only of despotism. He advocated German unification, under a constitutional monarchy and through Prussian initiative, but only under a constitution that would express the German spirit. Although he vigorously opposed revolutionary change in the Prussian form of government, he just as vigorously, and at personal risk, attacked the repressive policies of its administration. For example, he defended the freedom of the press as the organ of "public opinion"; the local press, in turn, hailed him as "the most popular and liberal man in Königsberg."

See also Hegel, Georg Wilhelm Friedrich; Hegelianism; Herbart, Johann Friedrich; Kant, Immanuel; Schleiermacher, Friedrich Daniel Ernst.

Bibliography

WORKS BY ROSENKRANZ

Encyklopädie der theologischen Wissenschaften. Halle, 1831.

Kritik der Schleiermacherschen Glaubenslehre. Königsberg, 1836.

Psychologie. Königsberg: Gebrüder Bornträger, 1837.

Studien. 5 vols. Berlin and Leipzig: Jonas Verlagsbuchhandlung, 1839–1848.

Kritische Erläuterungen des Hegelschen Systems. Leipzig: Gebrüder Bornträger, 1840.

Geschichte der Kantischen Philosophie. Leipzig, 1840. Vol. XII of Kant's *Werke,* edited by Rosenkranz and Schubert.

Schelling. Danzig: F.S. Gerhard, 1843.

G. W. F. Hegels Leben. Berlin: Duncker and Humblot, 1844. Supplement to Hegel's *Werken.*

System der Wissenschaft: Ein philosophisches Enchiridion. Königsberg: Gebrüder Bornträger, 1850.

Aesthetik des Hässlichen. Königsberg: Gebrüder Bornträger, 1853.

Wissenschaft der logischen Idee. 2 vols. Königsberg: Gebrüder Bornträger, 1858–1859.

Hegel als deutscher Nationalphilosoph. Leipzig: Duncker and Humblot, 1870.

Neue Studien. 4 vols. Leipzig, 1875–1878.

Politische Briefe und Aufsatze 1848–1856, edited by Paul Herre. Leipzig: Dieterich, 1919.

Die Hegelsche Rechte, edited by Hermann Lübbe. Stuttgart and Bad Cannstatt, 1962. Contains several essays, mostly political, and complete bibliography.

WORKS ON ROSENKRANZ

Esau, Lotte. *Karl Rosenkranz als Politiker.* Halle: M. Niemeyer, 1935.

Jonas, Richard. *Karl Rosenkranz.* Stuttgart: Weicher, 1906.

Metzke, Erwin. *Karl Rosenkranz und Hegel.* Leipzig: Heims, 1929.

Stephen D. Crites (1967)

ROSENZWEIG, FRANZ
(1886–1929)

Franz Rosenzweig, the religious existentialist, was born in Cassel, Germany. From 1905 to 1912 he studied natural sciences, modern history (under Friedrich Meinecke), and philosophy (under Heinrich Rickert) at the universities of Göttingen, Munich, Freiburg, and Berlin. At Berlin he earned a doctor of philosophy degree in 1912 with a dissertation on G. W. F. Hegel's political doctrines; later, he expanded this study. In the fall of 1913, after a spiritual crisis, he turned to religious, especially Judaic, philosophy. In 1918–1919 he wrote *Der Stern der Erlösung* (*The Star of Redemption*), a three-part religio-philosophical system; in 1920 he founded the Freies Jüdisches Lehrhaus (Independent House of Judaic Studies) in Frankfurt. Two years later he was appointed lecturer for Jewish religious philosophy and ethics at the University of Frankfurt, but the onset of progressive paralysis prevented him from accepting the appointment. Despite his affliction, he continued his scholarly work until his death in Frankfurt.

Hegel und der Staat (Hegel and the state), completed in 1914, for which Rosenzweig used both published and unpublished materials, analyzes the development of Hegel's concept of the state and its place in the philosopher's system. For Rosenzweig, the reasons motivating the successive changes in Hegel's political theories are to be found in the philosopher's intellectual progression.

In "Das älteste Systemprogramm des deutschen Idealismus" (The earliest systematic program of German idealism; written in 1914), Rosenzweig established that young Friedrich von Schelling was the author of a treatise preserved in Hegel's handwriting. This treatise is Schelling's sole attempt at formulating a unified system, a feat most perfectly realized by Hegel.

Rosenzweig's own philosophy may be defined as religious existentialism. *The Star of Redemption* begins with a critique of the Western philosophic tradition and, especially, of Hegel. Rosenzweig rejected as contrary to experience the attempt to reduce to one basic essence the three elements of reality: God, the world, and man.

In German idealism it is human consciousness and thought from which both God and world are deduced. In addition, consciousness is understood as "consciousness in general," which reduces to insignificance the individual being and his separate consciousness. But thought, Rosenzweig argued, is only one of the components of existence; it does not precede existence. The significance of the individual man stems from his being alive; he is more than a part of nature or the world. In this affirmation of the concrete person in his particularity Rosenzweig resumed the anti-Hegelian revolt of Arthur Schopenhauer, Ludwig Feuerbach, Søren Kierkegaard, and Friedrich Nietzsche, with its concern for the individual. The experience (*Erfahrung*) of the thinker, intent upon the value and significance of things, must guide him in confronting existence. Experience offers knowledge of God, the world, and man.

Under the influence of the later Schelling, and, to a certain degree, of Hermann Cohen, Rosenzweig links his theory of experience with a theory of conceptual construction; this linkage helps him to discover the interrelationship and interaction of the elements of God, world, and man. By way of an intricate logical construct he arrives at the following statement of relationships in terminology borrowed from theology: creation denotes the action of God upon the world; revelation, the encounter of God and man; and redemption, the relation of man to the world.

In pagan imagery God, the world, and man are separated and independent of each other. The hero of Greek tragedy is isolated from men and alien to the gods; the plastic cosmos is unrelated to man and the gods, who, in turn, have no concern for the world or man. Only biblical religion teaches the interaction of the elements of reality; in this concept, added to what he calls experience, lie the roots of Rosenzweig's existentialism. According to this view, creation is the process through which God, hitherto hidden in the mythical beyond, appears to give the world reality. But creation implies transitoriness, finiteness, death; the process of creation is renewed and perfected in revelation, through which God, in his love, turns to man; the experience of this love evokes in man the consciousness of being a self and accords man reality. Now his original isolation and dumbness are overcome; his response to God's love is his own love. Man translates his love for God into love for his "neighbor," and by so doing participates in leading the world toward redemption. Through the deeds of love the temporality of life and the finality of death are overcome. Ultimate redemption is anticipated, and a sense of eternity in time experienced, primarily in the rhythm of the days that constitute the sacred calendar in the religions based on revelation, Judaism and Christianity. Both these religions represent, under the aspect of faith, authentic, though different, manifestations of reality, and both are concerned with the existential situation of individual man.

The ideal representative of the "new thinking," as Rosenzweig called his view, is a philosopher-theologian who, while maintaining scholarly objectivity, accepts the

subjective, unique self as the new point of departure. The new theology should be existentially orientated, and theological problems should be translated into human terms. In contradistinction to abstract, timeless, purely logical, solitary thinking, the new existential thinking is "grammatical": Human language, the word, the name, dialogue, are keys to the understanding of reality; the speaking thinker thinks *for* someone and speaks *to* someone. In such language-bound thinking, utmost importance is accorded to time; past, present, and future are actively involved in the process of thought, a notion found also in Martin Heidegger's philosophy.

In the Judaic field, Rosenzweig advocated a revaluation of the thought of classical Judaism. With Martin Buber he undertook to translate the Old Testament, faithfully transposing into German the style of the original.

See also Buber, Martin; Cohen, Hermann; Consciousness; Existentialism; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Heidegger, Martin; Idealism; Jewish Philosophy; Kierkegaard, Søren Aabye; Meinecke, Friedrich; Nietzsche, Friedrich; Revelation; Rickert, Heinrich; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur.

Bibliography

WORKS BY ROSENZWEIG

Hegel and der Stoat. 2 vols. Berlin: Oldenbourg, 1920.

Der Stern der Erlösung. Frankfurt am Main: Kauffmann, 1921; 3rd ed., Heidelberg, 1954. An English version is *The Star of Redemption,* translated by William W. Hallo. New York: Holt, Rinehart and Winston, 1971.

Briefe, edited by Ernst Simon and Edith Rosenzweig. Berlin: Schocken, 1935. A selection of letters.

Kleinere Schriften. Berlin: Schocken, 1937. A collection of papers, including "Das älteste Systemprogramm des deutschen Idealismus" and "Das neue Denken" (1925), a discursive "epilogue" to *Der Stern.*

Das Büchlein vom gesunden und kranken Menschenverstand. Dusseldorf, 1964. Written in 1921. Includes a popular presentation of the chief theme of *Der Stern.* An English version is *Understanding the Sick and the Healthy,* edited by N. N. Glatzer. New York: Noonday Press, 1953.

WORKS ON ROSENZWEIG

Altmann, Alexander. "Franz Rosenzweig on History." In *Between East and West.* London, 1958.

Buber, Martin. "Franz Rosenzweig." *Kant-Studien* 35 (4) (1930): 517–522.

Freund, Else. *Die Existenzphilosophie Franz Rosenzweigs.* 2nd ed. Hamburg, 1959.

Glatzer, Nahum N. *Franz Rosenzweig: His Life and Thought.* New York: Farrar, Straus and Young, 1953.

Guttmann, Julius. "Franz Rosenzweig." In *Philosophies of Judaism.* Translated by D. Silverman, 367–398. New York, 1964.

Löwith, Karl. "M. Heidegger and F. Rosenzweig, or Temporality and Eternity." *Philosophy and Phenomenological Research* 3 (1) (1942): 53–77.

Nahum Norbert Glatzer (1967)

ROSENZWEIG, FRANZ [ADDENDUM]

Rosenzweig drew heavily from the lectures of his teacher in modern Jewish philosophy, Hermann Cohen, to construct his own highly original revaluation of the thought of classical Judaism on the model of Judah Halevi's philosophy and poetry within the framework of the post-Hegelian, post-rationalist, German Romantic philosophy of the early twentieth century. Rosenzweig expressed his Jewish thought through many forms, including new German translations of the Hebrew Scriptures, essays on Jewish education, and his personalized administration of a nonaccredited school for Jewish studies at the University of Frankfurt. No Jewish theologian has had a more lasting impact on the subsequent development of Jewish philosophy than has Franz Rosenzweig. It is not an exaggeration to say that with very few exceptions every important Jewish religious thinker in the second half of the twentieth century was either his student or a student of his students in the United States, in Israel, and in western Europe.

See also Cohen, Hermann; German Philosophy; Halevi, Yehuda; Jewish Philosophy; Rosenzweig, Franz.

Bibliography

PRIMARY WORKS

Buber, Martin, and Franz Rosenzweig. *Schrift und ihre Verdeutschung* [Scripture and Translation]. Translated by Lawrence Rosenwald with Everett Fox. Bloomington, IN: Indiana University Press, 1984.

Rosenzweig, Franz. *Cultural Writings of Franz Rosenzweig.* Translated by Barbara E. Galli. Syracuse, NY: Syracuse University Press, 2000.

Rosenzweig, Franz. *Die "Gritli" – Briefe an Margrit Rosenstock-Huessy*. Tübingen, Germany: Verlag Bilam, 2002.

Rosenzweig, Franz. *Der Mensch und sein Werk: Gesammelte Schriften*. The Hague, Netherlands: Martinus Nijhoff, 1976.

Rosenzweig, Franz. *Ninety-Two Poems and Hymns of Yehuda Halevi*, edited by Richard A. Cohen. Translated by T. Kovach, E. Jospe, and G.G. Schmidt. Ithaca, NY: State University of New York Press, 1999.

Rosenzweig, Franz. *On Jewish Learning*. New York: Schocken Books, 1995.

Rosenzweig, Franz. *Der Stern der Erlösung*. The Hague, Netherlands: Martinus Nijhoff, 1976. Translated into English by William Hallo as *The Star of Redemption*. Notre Dame, IN: Notre Dame Press, 1985. Translated into Hebrew by Yehosha Amir as *Kokhav Ha-Geulah*. Jerusalem: Mosad Bialik, 1970.

Rosenzweig, Franz. *Understanding the Sick and the Healthy: A View of World, Man, and God*. Cambridge, MA: Harvard University Press, 1999.

SECONDARY WORKS

Anckaert, L., and Casper, B. *Franz Rosenzweig: A Primary and Secondary Bibliography*. Leuven, Belgium: Bibliotheek Van De Faculteit Der Godgeleerdheid Van De K. U. Leuven, 1990.

Batnitzky, Leora. *Idolatry and Representation: The Philosophy of Franz Rosenzweig Reconsidered*. Princeton, NJ: Princeton University Press, 2000.

Cohen, Richard A. *Elevations: The Height of the Good in Rosenzweig and Levinas*. Chicago: University of Chicago Press, 1994.

Freund, Else Rahel. *Franz Rosenzweig's Philosophy of Existence: An Analysis of the Star of Redemption*. The Hague, Netherlands: Martinus Nijhoff, 1979.

Funkenstein, Amos. "The Genesis of Rosenzweig's 'Stern der Erlösung': 'Urformel' and 'Urzelle.'" *Jahrbuch des Instituts für Deutsche Geschichte* Beiheft 4 (April, 1982).

Galli, Barbara Ellen. *Franz Rosenzweig and Jehuda Halevi*. Montreal: McGill-Queen's University Press, 1995.

Gibbs, Robert. *Correlations in Rosenzweig and Levinas*. Princeton, NJ: Princeton University Press, 1992.

Glatzer, Nahum Norbert, ed. *Franz Rosenzweig: His Life and Thought*. New York: Farrar, Straus & Young, 1953.

Gordon, Peter Eli. *Rosenzweig and Heidegger: Between Judaism and German Philosophy*. Los Angeles: University of California Press, 2003.

Mendes-Flohr, Paul R. ed. *The Philosophy of Franz Rosenzweig*. Hanover, NH: University Press of New England, 1988.

Miller, Ronald H. *Dialogue and Disagreement: Franz Rosenzweig's Relevance to Contemporary Jewish-Christian Unerstanding*. Lanham, MD: University Press of America, 1989.

Mosès, Stéphane. *System and Revelation: The Philosophy of Franz Rosenzweig*. Translated by Catherine Tihanyi. Detroit: Wayne State University Press, 1992.

Oppenheim, Michael D. *What Does Revelation Mean for the Modern Jew?: Rosenzweig, Buber, Fackenheim*. Lewiston/Queenston, U.K.: The Edwin Mellon Press, 1985.

Rotenstreich, Nathan. *Jews and German Philosophy: The Polemics of Emancipation*. New York: Shocken, 1984.

Rubinstein, Ernest. *An Episode of Jewish Romanticism: Franz Rosenzweig's The Star of Redemption*. Albany, NY: SUNY, 1999.

Samuelson, Norbert M. "The Philosophy of Rosenzweig." Chaps. 27–29 in *Jewish Philosophy: An Historical Introduction*. London: Continuum, 2003.

Samuelson, Norbert M. *User's Guide to Franz Rosenzweig's Star of Redemption*. Richmond, U.K.: Curzon, 1999.

Santner, Eric L. *On the Psychotheology of Everyday Life: Reflections on Freud and Rosenzweig*. Chicago: University of Chicago Press, 2001.

Turner, Joseph. *Faith and Humanism: A Study in Franz Rosenzweig's Religious Philosophy*. In Hebrew. Tel Aviv, Israel: Hakibbutz Hameuchad, 2001.

Vogel, Manfred. *Rosenzweig on Profane/Secular History*. Atlanta: Scholars Press, 1996.

Norbert Samuelson (2005)

ROSMINI-SERBATI, ANTONIO
(1797–1855)

The Italian philosopher, educator, and statesman Antonio Rosmini-Serbati was born in Rovereto, then part of the Austrian Tyrol. The families of both his parents held patents of nobility under the Holy Roman Empire. A private education begun at an early age and directed to the priesthood established a firm foundation for his later work. Finding Austrian rule oppressive, Rosmini moved to the freer region of Piedmont. He started his career by founding the Institute of Charity, devoted to education and missions. He began to publish prolifically in philosophy, literature, and pedagogy. In politics he became an active exponent of the principles of Neo-Guelphism and reached the peak of his public career as counselor to Pius IX during the period from 1848 to 1853; at the end of this period, more conservative forces came to power. Retiring to private life, Rosmini continued his writing and assumed the direction of his institute. The present article restricts itself to Rosmini's philosophical work.

Although developed in a large number of works, Rosmini's philosophical thought presents a high degree of unity. This unity has two sources: the historical and apologetic intentions that sustain it and the internal development of certain germinal ideas. Rosmini's overt intention was to create a Christian-Catholic apologetics that would meet the demands of modern philosophical thought while remaining faithful to the core of traditional Christian philosophy. Since Augustinian and Thomistic realism predominated in Christian philosophy, Rosmini endeavored to anchor his thought in that tradition, exhibiting an affinity to the Augustinian strain. At the same time, he sought to meet the demands of rationalism and empiricism, and especially of the Kantian attempt at a resolution of the tension between the two. The effort to meet these conditions imparted to Rosmini's thought a high degree of complexity and sophistication.

The point of departure of the Rosminian system is his *Nuovo saggio sull'origine delle idee* (1830), a work of elaborate synthesis. The controlling principle of the synthesis is basically Augustinian, but the work develops around three centers: the idea of being, intellectual perception, and the origin of ideas.

THE IDEA OF BEING

Following Immanuel Kant, Rosmini accepted a dual order of a posteriori and a priori in the process of knowledge and identified the ground of science with a priori principles of knowledge. Whereas Kant distinguished diverse orders or forms of a priori synthesis, Rosmini reduced that plurality to a single form, the idea of being. Only the idea of being can be thought without reference to any other idea, and only that idea is thought, at least implicitly, in the thinking of any other idea. The idea of being is not the product of the subject, whether empirical or transcendental; it is a datum offered immediately by God to the intelligent subject; it is, moreover, ontologically and functionally constitutive of that subjectivity. The idea of being is both a category and a transcendental operation. It is a category, for the subject knows through the process of the existential judgment, in which being as given in the idea of being is predicated of things. This judgment establishes the subsistence of the object as present and known in the judgment.

As a category, the idea of being is the irreducible "other" to any specific content of thought or knowledge. It must also either be a product of the empirical subject or be truly objective. In the first case, the idea of being would be subjective and would render all knowledge subjective; in the second, its objectivity would seem to require the postulation of a "transcendental" subject. Rosmini accepted neither horn of this ostensible dilemma. He held that the human subject is empirical but also capable of a transcendental operation by which it can secure universal and necessary knowledge. It performs this operation through the idea of being; more accurately, this operation is one with the idea of being. As a transcendental operation, the idea of being constitutes the knowing subject ontologically and existentially; it secures the realm of universal and necessary knowledge. Finally, it is transcendent, for it is not the product of the subject, whether empirical or transcendental, but a datum that must be referred to the action of God. It is this last point that relates Rosmini's view to that of Augustine.

INTELLECTUAL PERCEPTION

Although no knowledge is possible except through the idea of being, that idea does not suffice for the effective knowledge of the actual world of determinate forms of subsistence. This world can be known only if sensation has entrance into the realm of the idea of being and vice versa. Sensation is the vehicle of the multiple forms of determinate subsistence of the real world, but it does not present them as being; for them to be presented as being, sensation must be infused by the idea of being. This infusion is achieved concretely in an operation that Rosmini called intellectual perception.

Intellectual perception is rooted in man's fundamental constitution, for he is both sentient and intelligent. Every concrete act of knowing is structured by sensation and intelligence, related in a radical unity. There is neither pure sensation nor pure intellection, or intellectual vision. Intellectual perception, in which these pure elements occur in vital union, places man in authentic contact with the concrete real world. This operation is perception because by it the subject sensibly lays hold of reality, which actually stands before it, as subsistent. It is intellectual because the sensible perception evokes in the indeterminate being, which is already present to the subject in the idea of being, determinations by which the ideas of particular things arise. Intellectual perception is not, manifestly, the synthesis of two antecedently existing elements; it is the complex term of a complex, concrete operation, rooted in the fact that man is a complex principle and subject, both intelligent and sensitive.

ORIGIN OF IDEAS

On the basis of the foregoing points, Rosmini addressed the problem of the origin of ideas. Ideas, except the idea of being, arise through the process of abstraction. Empiricists and sensationists confuse intellectual perception with sensation when they speak of the formation of ideas out of the elements of sensation through abstraction and reflection. The act of reflection is not performed on the simple sense datum but upon objects already known and present through intellectual perception. By noting certain characteristics and averting attention from others, abstraction forms ideas of various degree up to the most general. The idea of being is alone excluded from this account; for it is the presupposition, not the result, of intellectual perception.

SUBJECTIVE REALISM

Rosmini proceeded in *Psicologia* (1850) to consider the subject, which is the locus of the process of knowledge.

Here again his doctrine reflects his concern to meet both empirical and idealist claims by passing beyond them. He refused to resolve the subject into the transcendental process, as he claimed idealists did, or into the process of sensation, as he said empiricists did. Instead, he offered a "subjective realism" or, better, a "realism of the subject." Its basis is the theory of the "fundamental sentiment," the immediate analogue for which is intellectual perception. The soul, while retaining its classical status as the active principle of vital operations and psychic phenomena, takes on a new dimension; it is the substance-sentiment, the intuitive sense of immanent being that generates subsistence. The reality of the subject is constituted by this immediate, nonobjective, and synthetic sense of self, which draws into a subsistent unity all aspects—sensitive, intelligent, and volitional—of the subject's complex life. This fundamental sentiment is the first and the continuous experience that man has of himself. It always involves, moreover, a relationship to a corporeal term, the body. This specific aspect of the fundamental sentiment, the corporeal sentiment, is characteristic of human nature. By it Rosmini justified the classical doctrine of man's composition out of body and soul. All other sensations are accidental to this fundamental sentiment; it is primitive and incommunicable and constitutes the subject in its unity and complexity.

Rosmini was also able to offer a fresh form of the classical doctrine of the immortality of the soul. Immortality has its basis in the fundamental sentiment of the idea of being; through the corporeal sentiment, the body shares immortality.

THE PERSON

Important both in itself and for its function in his social and political thought is Rosmini's central doctrine of the person. A subject has two aspects, nature and person. A subject's nature is the complex and sum of the activities of which the subject is agent. The perfection of the subject in the order of nature is the perfection, in number and in quality, of these activities. "Person" designates the directive unity of these activities and hence is associated in a special way with the will. The will is fundamental because it directs and organizes the activities of a person's nature, and in so doing it exhibits the basic deontic character of the person, its orientation toward a norm, toward the ought. The person emerges as the unique and incommunicable unity of the activities of the nature through a unique and unrepeatable activity of the will. It is not merely an operational or structural unity but a deontic one, basically oriented toward the world of values and norms and hence constitutively moral. The central effort of life is the realization of the developed or explicit person, which is achieved through the exercise of moral decision within the context of nature and its diverse activities. This effort is the basis of Rosmini's distinction between *vita direta* and the *vita riflessa,* which is central to his moral philosophy. The central effort of the moral life is the practice of the *vita riflessa,* the examined life in a creative sense.

The elaboration of the notion of the person gives structure to Rosmini's moral philosophy; his philosophy of right, law, and state; and his theory of education.

ETHICS

Personalism enabled Rosmini to overcome the formalism of Kantian ethics. The idea of being is the criterion of the good as well as of truth. In the intimate unity of the person, the speculative act of intellectual perception immediately translates itself into a practical judgment that becomes the legislative principle of action. The truth of being that intellectual perception presents inevitably involves the assenting activity of the will. The will seeks the being of all things in the idea of being, revealed in the deontic order as the good. Rosmini, on Kant's model, tried to distill this insight into a rule: Recognize in action or practice what you have recognized speculatively. The essence of the moral life resides in this act of recognition, reflected prismatically in all the concrete situations in which the agent discovers himself. The obligatoriness of the rule springs from the fact that a hiatus between the speculative and the practical orders, between universal recognition and individual action, is intolerable. The psychological expression of this intolerance is remorse, the characteristic state of a person who deviates from this imperative. The true form of Rosmini's moral philosophy is embodied in another imperative: Be faithful to being; specifically, to the being that is revealed in the idea of being and which is the ground of all.

Fidelity to being was immediately translated by Rosmini into a rule of justice. The idea of being contains all the grades of being. The realm of being thus constitutes at the same time a hierarchy of values. Fidelity to being demands that the rule of justice, "Give to each its due," be interpreted in terms of this hierarchy. How is this hierarchy of values to be apprehended? Rosmini's reply is that it is to be apprehended through spontaneous recourse to the intellectual light, the constitutive presence in the subject of the idea of being.

POLITICAL PHILOSOPHY

The concepts of person and justice provide the bases of Rosmini's political philosophy. Abstractly, right is the property of being, for being demands to be recognized and in doing so establishes the moral and the juridical orders. Concretely, right has its locus in the person, because of the person's ontological status as subject. In the person, right becomes a capacity to act eudaemonically, a capacity that is protected by the moral law; the same law imposes on others the obligation to recognize this capacity. Rosmini sought to bring right under the moral law in order to oppose those who would make it rest on force; he made it an endowment of the person to oppose those who would assign its origin to any other source, such as organized society in any of its forms. He distinguished innate natural rights, derived connatural rights, and acquired rights. Property, by means of which the person acquires physicomoral dominion over objects, is the chief acquired right.

While property defines the relation of the person to objects, the social bond relates him to other persons. The basis of the bond of sociality among a plurality of persons is their participation in a common intelligent principle, ideal being. Rosmini placed the forms of social life on a continuum between the terms of the most rudimentary and inclusive—membership in the human race—and the most intimate and exclusive—the conjugal relationship. Civil society falls midway on this continuum. Civil society has only a functional and not a substantive character: It does not originate rights but simply regulates the mode of their enjoyment and exercise. This provides Rosmini with his definition of the state and of government: The state is a regulatory principle of the modality of human rights. A just state achieves a balance between the common good (the good of the members distributively considered) and the public good (that of the social body considered as an organism). Abstractly, the common good is to be preferred to the public good, so as to preclude justification of acts of the state by recourse to the doctrine of "reason of state"; concretely, this preferential status is less determinate.

BEING AND GOD

In two extensive works, the *Teosofia* (posthumously published, 1859–1874) and the *Teodicea* (1845), Rosmini drew the widest possible conclusions from his personalistic premises. The theme of the *Teosofia* is the unity of being as prior to any of its modes (the absolute metaformality of being). Being, in this sense, is the basis of all the actual and determinate forms of being and contains within itself all of the principles of that determination *in abstracto* or *virtualiter.* It is not, however, the creative principle by which those forms are reduced to actuality. The need for a creative principle opens the way for the argument of the *Teodicea,* that God necessarily exists. God is the creative principle by which the virtuality of the order of primal being is realized in the actual and concrete modes of existence.

EDUCATIONAL THEORY

The culmination of Rosmini's thought is considered by many to be his pedagogical theory. The guiding principle of this theory is a summary of his entire philosophy: respect for the human person as the vehicle of divine light and ideal being. Rosmini stressed the unity of educational process and also methodology. The person is the principle of integrity; education is the process of the realization of the person in this sense. The principle of this integration is religion, which gives unity of purpose, unity of doctrine, and unity of powers to the educational process. The supreme law of method, the principle of gradation, ensures the conformity of the process of education to that of life. The process of growth and integration according to this law is from the universal to the particular. The object of the entire process is the free and realized person who fulfills himself in free association with other persons in all social forms and whose freedom rests ultimately upon his foundation in ideal being.

See also Augustine, St.; Augustinianism; Being; Empiricism; Ideas; Kant, Immanuel; Personalism; Philosophy of Education, History of; Rationalism; Realism; Thomism.

Bibliography

WORKS BY ROSMINI

Opere. Edited by Enrico Castelli, 28 vols. Rome, 1934–. One hundred volumes planned.

Saggio sull'unità dell educazione. Milan, 1826.

Nuovo saggio sull'origine delle idee. Rome, 1830.

Teodicea. Milan, 1845. Translated as *Theodicy.* 3 vols. New York, 1912.

Psicologia. 2 vols. Novara, 1850. Translated as *Psychology,* 3 vols. 1884–1888.

Del principio supremo della methodica e di alcune sue applicazione in servizio della umana educazione. Turin, 1857.

Teosofia. 5 vols. Turin, 1859–1874.

WORKS ON ROSMINI

Bibliographies

Antologia rosminiana, Vol. 1 (1955).

Caviglione, Carlo. *Bibliografia delle opere di Antonio Rosmini.* Turin: Paravia, 1925.

Palhories, Fortunat. *La philosophie de Rosmini.* Paris, 1908.

Biographies

Bertoletti, C. G. *Vita di Antonio Rosmini.* Turin, 1957.

Leetham, Claude. *Rosmini.* London and New York: Longmans, Green, 1957.

Pagani, G. B. *Vita di Antonio Rosmini,* 2 vols. Turin, 1897. Revised by G. Rossi and G. Bozzetti. Rovereto, 1957. Translated as *The Life of Antonio Rosmini-Serbati.* London: Routledge & Sons, 1907.

Philosophical Studies

Bruno, G. F. "Rosmini's Contributions to Ethical Philosophy." In *Archives of Philosophy.* New York: Columbia University Press, 1916.

Bulferetti, L. *Antonio Rosmini nella Restaurazione.* Florence: F. Le Monnier, 1942.

Capone Braga, Gaetano. *Saggio su Rosmini: Il mondo delle idee,* 2nd ed. Florence, 1924.

Caviglione, Carlo. *Il Rosmini vero.* Voghera, Italy, 1912.

Chiavacci, C. *Il valore morale nel Rosmini.* Florence, 1921.

Franchi, A. *Saggio sul sistema ontologico di Antonio Rosmini.* Milan, 1953.

Galli, Gallo. *Kant e Rosmini.* Città di Castello: S. Lapi, 1914.

Gentile, Giovanni. *Rosmini e Gioberti,* 2nd ed. Florence: Sansoni, 1955.

Gonella, G. *La filosofia del diritto secondo Rosmini.* Rome: Studium, 1934.

Honan, U. *Agostino, Tommaso e Rosmini.* Milan: Sodalitas, 1955.

Morando, Dante. *La pedagogia di Antonio Rosmini.* Brescia: La scuola, 1948.

Morando, Giuseppe. *Esame critico delle XL proposizione rosminiani condannate.* Milan: Cogliati, 1905.

Nicola, G. B. *Saggi di scienza politico di Antonio Rosmini.* Turin, 1933.

Pignoloni, E. *Il reale nei problemi della Teosofia di Antonio Rosmini.* Milan: Sodalitas, 1955.

Prini, Pietro. *Introduzione alla metafisica di Antonio Rosmini.* Milan: Sodalitas, 1953.

Riva, C. *Il problema dell'anima intelletiva secondo Antonio Rosmini.* Domodossola: Sodalitas, 1956.

Rovea, Giuseppe. *Filosofia e religione in Antonio Rosmini.* Milan: Sodalitas, 1951.

Sciacca, M. F. *La filosofia morale di Antonio Rosmini.* Rome: Bocca, 1938.

Collections of Articles

Atti del congresso internazionale di filosofia "Antonio Rosmini," 2 vols. Florence, 1957.

Crisis (1955), no. 6.

Giornale di metafisica (1955), nos. 4 and 5.

Humanitas (1955), nos. 9 and 10.

Rivista della filosofia neoscolastica (1955), nos. 4 and 5.

Rivista rosminiana (1955), nos. 3 and 4. Contains an extensive bibliography.

Teoria (1955), nos. 3 and 4.

A. Robert Caponigri (1967)

ROSS, WILLIAM DAVID

(1877–1971)

William David Ross was a British Aristotelian scholar and moral philosopher. Sir David Ross was born in Scotland and was educated at the Royal High School in Edinburgh, Edinburgh University, and Balliol College, Oxford, where he took firsts in classical moderations and "greats." He was a fellow of Merton College from 1900 to 1902, when he was elected a fellow and tutor of Oriel. He was provost of Oriel from 1929 until his resignation in 1947.

Ross was prominent in academic and public life. He was vice-chancellor of Oxford University (1941–1944), pro-vice-chancellor (1944–1947), president of the Classical Association (1932), and president of the British Academy (1936–1940). He was chairman of Council of the Royal Institute of Philosophy continuously since 1940. In 1947 he served as president of the Union Académique Internationale.

Ross was awarded the Order of the British Empire for his work in the ministry of munitions and as a major on the special list during World War I. He was knighted in 1938. During World War II he was a member of the appellate tribunal for conscientious objectors and after the war was honored by the governments of Norway and Poland. Among his many public services were the chairmanships of three government departmental committees (1936–1937) and of the civil service arbitration tribunal (1942–1952). From 1947 to 1949 he was chairman of the important Royal Commission on the Press.

The qualities that made Ross successful in public life are those to which he owes his distinction as a philosopher. He was not only an Aristotelian scholar, but he also had an Aristotelian frame of mind—moderate, critical, balanced, thorough, and, above all, judicious. He valued and possessed what Aristotle called "practical wisdom" no less than speculative ability.

Ross edited the Oxford translations of Aristotle, published between 1908 and 1931. He translated the *Metaphysics* and the *Ethics* himself, and he published definitive editions of a number of Aristotle's works. His *Aristotle* (London, 1923) is mainly expository, each chapter being concerned with a major aspect of Aristotle's work; this is still the best all-round exposition in English.

Ross was the leading opponent of the view of John Burnet and A. E. Taylor that the Socrates of Plato's dialogues is never a mouthpiece for Plato's own doctrines. In *Plato's Theory of Ideas* (Oxford, 1951), Ross rejected their contention that the theory of Ideas was originally the work of Socrates and not of Plato. This book traces the

development of the theory of Ideas through Plato's thought. It includes a detailed discussion of Plato's cryptic doctrine of "ideal numbers," using Aristotle's account in the *Metaphysics* as a guide to the interpretation of the doctrine.

Ross's main contribution to philosophy, as distinct from philosophical scholarship, is in the field of ethics. In *The Right and the Good* (Oxford, 1930), he argued the case for intuitionism with a lucidity and thoroughness that made the book a classic. For some ten years it was the center of ethical controversy. In his *Foundations of Ethics* (Oxford, 1939) Ross restated his case and replied to his critics.

Ross's approach to ethics is Aristotelian. "The moral convictions of thoughtful and well-educated people are the data of ethics, just as sense-perceptions are the data of a natural science" (*The Right and the Good*, p. 41). He appeals to what we mean by rightness and goodness and assumes that this guarantees the existence of what is meant and is a sure guide to its nature.

The germ of Ross's position is to be found in an article by H. A. Prichard, "Does Moral Philosophy Rest on a Mistake?" (*Mind* 21 [1912]: 21–152; reprinted in *Moral Obligation*, Oxford, 1949, pp. 1–17). Prichard was a pupil of John Cook Wilson, who also influenced Ross directly, an influence that appears in Ross's opposition to reductionism and in his view that knowledge and opinion are distinct in kind. The other main debt acknowledged by Ross is to G. E. Moore, whose arguments against ethical subjectivism he endorses, although he rejects Moore's "ideal utilitarianism."

Right and good are for Ross distinct, indefinable, and irreducible objective qualities. Rightness belongs to acts, independently of motives; moral goodness belongs to motives. Ross uses "act" for what is done and "action" for the doing of it. Thus, the doing of a right act may be a morally bad action—that is, a right act can be done from a morally bad motive; the inverse also holds. Nor can it ever be morally obligatory to act from a good motive. There are four kinds of good things—virtue, knowledge, pleasure, and the allocation of pleasure and pain according to desert. No amount of pleasure equals the smallest amount of virtue. In *Foundations of Ethics* Ross argued that virtue and pleasure are not good in the same sense—virtue is "admirable," pleasure only "a worthy object of satisfaction." What alone is common to the two senses is that they express a favorable attitude.

Ross's two main targets are ethical subjectivism and "ideal utilitarianism," which "ignores, or at least does not do full justice to, the highly personal character of duty" (*The Right and the Good*, p. 22). Specific duties are of three kinds—reparation, gratitude, and keeping faith. The "plain man" (to whom Ross, as a good Aristotelian, frequently appeals), in deciding what he ought to do, thinks as often of the past (a promise made, a debt incurred) as of future consequences. Ross does, however, admit among duties the utilitarian general duty of beneficence when it does not conflict with a specific duty. And "even when we are under a special obligation the tendency of acts to promote general good is one of the main factors in determining whether they are right" (p. 3a).

Conflict of duties is one of the main problems facing an intuitionist, who cannot accept the utilitarian's "Do what will produce the most good." Ross says: "Do whichever act is more of a duty." To make sense of "more of a duty," he draws a distinction between prima-facie and actual duties and holds that conflict can only arise between prima-facie duties. An act is a prima-facie or "conditional" duty by virtue of being of a certain kind (for instance, the repaying of a debt) and would be an actual duty if it were not also of some other morally important kind or did not conflict with another more important prima-facie duty. Thus, if I have promised to lend money to a friend in need, I have a prima-facie duty to hand over the money. But suppose that before I have done so, I find that I need it for the legal defense of my son, charged with a crime of which I believe him innocent. I recognize a conflicting prima-facie duty to help him. Ross maintains that (*a*) one, and only one, of these two prima-facie duties is my actual duty; (*b*) I know each of them to be a prima-facie duty—this is self-evident; (*c*) I can have only an opinion about which is "more of a duty" and therefore my actual duty.

See also Aristotelianism; Aristotle; Ethical Subjectivism; Ethics, History of; Intuitionism; Intuitionism and Intuitionistic Logic; Moore, George Edward; Plato; Socrates; Taylor, Alfred Edward; Utilitarianism.

Bibliography

OTHER WORKS BY ROSS

Kant's Ethical Theory: A Commentary on the Grundlegung zur Metaphysik der Sitten. Oxford, Clarendon Press, 1965.

The Right and the Good, edited by Philip Stratton-Lake. Oxford: Clarendon, New York: Oxford University Press, 2002.

Some works on Ross are J. H. Muirhead, *Rule and End in Morals* (London: Oxford University Press, H. Milford, 1932); H. J. McCloskey, "Ross and the Concept of Prima Facie Duty," in *Australasian Journal of Philosophy* 41 (1964): 336–345; G. E. Hughes, "Motive and Duty," in *Mind* 53

(1944): 314–331, W. A. Pickard-Cambridge, "Two Problems about Duty," in *Mind* 41 (1932): 72–96, 145–172, 311–340; P. F. Strawson, "Ethical Intuitionism," in *Philosophy* 24 (1949): 23–33.

A. K. Stout (1967)
Bibliography updated by Michael J. Farmer (2005)

ROUGIER, LOUIS

(1889–1982)

Louis Rougier, the French philosopher, was a pupil of Edmond Goblot. Rougier taught philosophy at the universities of Besançon and Caen. In 1935 he organized and presided over the Paris International Congress of Scientific Philosophy, where the leading spokesmen for logical empiricism, at the time little known in France, presented their views in a body.

From the start, Rougier's thought had been marked by the contemporary upheavals in the sciences of physics, mathematics, and logic. To these developments he devoted several of his early books, including *La philosophie géométrique d'Henri Poincaré* (Paris, 1920), *La structure des théories déductives* (Paris, 1921), *La matiére et l'energie selon la théorie de la relativité et la théorie des quanta* (Paris, 1921), and *En Marge de Curie, de Carnot et d'Einstein* (Paris, 1922).

In his view, the upsets in the sciences reinforced the closely pressed critique which he had directed in his doctoral thesis, *Les paralogismes du rationalisme* (Paris, 1920), against the theory academic philosophers call "rationalism." This is an a priori rationalism, quite different from scientific and experimental rationalism. It asserts the existence of a universal, immutable reason and of eternal, necessary truths, with all the theological, ontological, and epistemological implications that such a thesis requires. According to Rougier, the body of notions and principles that constitute "reason" in the classic sense is simply the characteristic of a certain mental structure, the ontological or metaphysical temperament, which is also the subject of his detailed study *La scolastique et le thomisme* (Paris, 1925). Besides the temperament dominated by "reason," history discloses other temperaments—animistic, symbolic, scientific—having command of other types of explanation. The human mind possesses an infinite plasticity; it is able to take delight in quite varied forms of intelligibility, without any internal necessity having compelled it to evolve in just the direction that it has. If the laws of logic are necessary truths, it is only because they are tautologies in the sense of Ludwig Wittgenstein; that is, they are devoid of any information about the universe and hence stripped of any ontological import. Even this logical necessity, as is shown by the existence of a plurality of logics, is relative to a given system of axioms and rules.

This rejection of all a priori synthesis, this radical separation between logico-mathematical statements and empirical statements, and the condemnation it entails of all metaphysics as victim of the imperfections of our natural languages (*La métaphysique et le langage,* Paris, 1960), closely ally Rougier's philosophy to that of logical empiricism. His long *Traité de la connaissance* (Paris, 1955) offers analyses illustrated with abundant examples from the past and contemporary history of the sciences; in style and ideas it is probably closer than any other French book to the majority of central European and American works on epistemology. Nevertheless, certain features testify to his originality in comparison with the logical empiricism of the Vienna circle. Rougier rejects the physicalist reduction and upholds a plurality of languages. Nor does he agree that all basically unsolvable problems must by their nature alone be regarded as devoid of meaning; besides, meaninglessness is a notion relative to the language chosen. Further, several of the ideas he developed in works other than the *Traité,* for example his thesis of the diversity of mental structures and the plasticity of the intellect, do not strictly belong to the common stock of the school of logical empiricism, but have been added to it.

Although epistemology and the critique of knowledge are at the center of Rougier's philosophy, he wrote in two other fields. One is the history of scientific, philosophical, and religious ideas, to which he devoted *Celse ou le conflit de la civilisation antique et du christianisme primitif* (Paris, 1926) and *La religion astrale des Pythagoriciens* (Paris, 1959). The other is contemporary political problems; he dealt critically with the democratic and egalitarian ideology of the "men of 1789" and their successors in such works as *Les mystiques politiques et leurs incidences internationales* (Paris, 1935), *Les mystiques économiques* (Paris, 1949), and *L'erreur de la démocratie française* (Paris, 1963).

Rougier systematically omitted these two aspects of his thought from the account he himself gave of his "philosophical itinerary" (*La revue libérale,* no. 33, [1961]: 6–79), an account which can well serve as an overall study of his theory of knowledge.

See also A Priori and A Posteriori; Epistemology; Logical Positivism; Rationalism.

Bibliography

There are no studies of Rougier, but the reader may wish to consult the brief account by A. D. Dandieu in *Anthologie des philosophes français contemporains* (Paris: Éditions du Sagittaire, 1931) followed by a detailed bibliography and two extracts from Rougier's writings.

WORKS BY ROUGIER

Celse discours vrai contre les Chrétiens. Paris: J.-J. Pauvert, 1965.

Histoire d'une faillite philosophique: La Scolastique. Paris: J.-J. Pauvert, 1966.

Le culte des images et la primitive Église. Paris: (7e), Cercle Ernest-Renan, 3, rue Récamier, 1970.

The Genius of the West. Los Angeles, CA: Nash. 1971.

La genèse des dogmes chrétiens. Paris: A. Michel 1972.

La métaphysique et le langage; Essai. Paris: Denoël 1973.

Le conflit du christianisme primitif et de la civilisation antique. Paris: Grece, 1974.

Celse contre les chrétiens: La réaction païenne sous l'empire romain. Paris: Copernic, 1977.

De Gaulle contre De Gaulle. Paris: Éditions du Triolet 1978.

Du paradis à l'utopie. Paris: Copernic, 1979.

Astronomie et religion en Occident. Paris: Presses universitaires de France, 1980.

La mystique démocratique: Ses origines, ses illusions. Paris: Editions Albatros, 1983.

La religion astrale des pythagoriciens. Monaco: Editions du Rocher, 1984.

Poincaré, Henri. Scientific Opportunism: An Anthology = L'opportunisme scientifique. Compiled by Louis Auguste Paul Rougier; edited by Laurent Rollet. Boston: Birkhäuser, 2002 .

Robert Blanché (1967)
Translated by Albert E. Blumberg
Bibliography updated by Michael J. Farmer (2005)

ROUSSEAU, JEAN-JACQUES
(1712–1778)

Jean-Jacques Rousseau, the philosopher, essayist, and novelist, was born at Geneva. His mother having died a few days after his birth, he was brought up by an aunt and an erratic father who taught him to read through the medium of sentimental novels and Plutarch's *Lives.* He had little formal education. After staying for about two years with a country minister at Bossey, he returned to Geneva and lived with an uncle. He was then apprenticed in turn to a notary and an engraver, the latter of whom treated him so brutally that in 1728 he left Geneva to seek his fortune elsewhere.

Rousseau was protected and befriended by Mme. de Warens, a convert to Roman Catholicism, who had left her native canton of Vaud to live at Annecy in Savoy, with financial support from the king of Sardinia and the ecclesiastical authorities. Rousseau's subsequent attachment to her was a decisive factor in his conversion to Roman Catholicism as well as in his emotional development. He made a formal abjuration of Protestantism at the hospice for catechumens at Turin. He then served for a time as a lackey, finally returning to Mme. de Warens in 1729. Thereafter, he led an unsettled life, restless travel alternating with a more stable existence at Chambéry, where Mme. de Warens had established herself. Intellectually, the most important event of this phase of his life was a protracted spell of enthusiastic study under his own direction. A brief experience as a private tutor at Lyons in 1740 helped to create a lifelong interest in education and at the same time convinced Rousseau that he had no aptitude for this profession. As he had acquired some musical competence at Annecy, he set out hopefully for Paris in 1742 with a new system of musical notation. Although this did not bring him the success he hoped for, he was introduced to a number of influential people, including the wealthy Mme. Dupin and her stepson M. de Francueil.

In 1743, Rousseau was appointed secretary to the French ambassador at Venice, M. de Montaigu, but he lost this post the following year because of a quarrel with him. On his return to Paris, Rousseau increased his difficulties by an irregular union with an ignorant servant girl, Thérèse Le Vasseur, in 1745; by her he probably had five illegitimate children, who were all sent to a foundlings' home. He also met Denis Diderot, Jean Le Rond d'Alembert, and other *philosophes* and was invited to contribute musical articles to the *Encyclopédie.*

Rousseau's literary career began in 1750 with the publication of the *Discours sur les sciences et les arts,* which had previously won a prize at the Academy of Dijon. However, his first real success came with the performance of his opera *Le devin du village* before Louis XV at Fontainebleau, but his refusal to allow himself to be presented to the king lost him any chance of securing a royal pension. A journey to Geneva in 1754 led to a reconciliation with the republic and a formal return to Protestantism.

After the publication in 1755 of his *Discours sur l'origine de l'inégalité,* Rousseau felt increasingly unhappy in Paris, and in 1756 he installed himself in a small country house, called the Hermitage, which belonged to a rich friend, Mme. d'Épinay. There followed a comparatively short but intense period of literary activity that saw the publication of the *Lettre à d'Alembert sur les spectacles* (1758), *Julie, ou la Nouvelle Héloïse* (1761), *Émile* (1762), and the *Contrat social* (1762).

During this time Rousseau's relations with the Encyclopedists became increasingly strained, with intellectual differences, especially on the subject of religion, being aggravated by personal quarrels with former friends such as Diderot and the Baron von Grimm. In 1762 the condemnation of *Émile* by the Paris Parlement forced him to flee from France and settle in Neuchâtel under the protection of the king of Prussia. In the *Lettre à M. de Beaumont* (1763) Rousseau vigorously defended the "Profession de foi du vicaire savoyard," which had been included in the fourth book of *Émile,* against its condemnation by the archbishop of Paris; this was followed in 1764 by another polemical work, the *Lettres écrites de la montagne,* provoked by increasing opposition from the Genevan authorities to his political and religious views. Alarmed by local hostility, Rousseau decided to leave the Neuchâtel region in 1765, and he accepted an invitation from the philosopher David Hume to make his home in England. His arrival in that country in 1766 and his subsequent residence in Derbyshire were disturbed by the appearance of abnormal emotional and mental reactions, culminating in the irrational conviction that Hume's invitation had been a mere pretext for Rousseau's defamation. After quarreling violently with Hume (who riposted by publishing an account of the affair), Rousseau fled panic-stricken to France in 1767. For the next few years he moved from place to place, oppressed by the thought of universal persecution. He eventually settled in Paris in 1770 and died suddenly on July 2, 1778, less than two months after he had gone to live on the estate of the marquis de Girardin at Ermenonville.

The chief literary activity of Rousseau's last years was the composition of a remarkable series of personal works, the *Confessions,* on which he had worked intensively during his stay in England; the dialogues known as *Rousseau juge de Jean-Jacques,* a curiously pathological document illuminated by some pages of remarkable brilliance and insight; and the beautiful but unfinished *Rêveries du promeneur solitaire.* These writings are remarkable for their lyrical power and sustained efforts at self-analysis.

THOUGHT

From the very first Rousseau's work betrayed the strongly personal emphasis of a writer who felt that he did not truly belong to his immediate environment. Being of Genevan origin, largely self-taught, and endowed with a particularly sensitive temperament, he could never bring himself fully to accept the social and moral implications of French culture, even though he never ceased to admire French taste. In 1749, as he was on his way to Vincennes to visit his imprisoned friend Diderot, he saw in a copy of the *Mercure de France* the subject of the prize essay set by the Academy of Dijon: whether the restoration of the arts and sciences has contributed to the purification of manners. In the *Confessions* he writes that at that moment he experienced a sudden "illumination" and "inspiration," the dazzling vision of a "new universe," which impelled him to answer the academicians' question with an emphatic "No!" Although this viewpoint was already familiar to a certain type of traditional Christian moralist, Rousseau struck a new personal note remarkable for its deeply felt sincerity; he always refused to consider himself as a professional man of letters and stressed his role as an independent writer with a message for humanity.

NATURE AND SOCIETY. Rousseau's early works (the two discourses and the *Lettre à d'Alembert*) developed the fundamental antithesis that he deemed to exist between contemporary society and the nature of man. European civilization was indicted for having sacrificed the moral demands of human nature to the superficial allure of a purely intellectual culture and thus for having replaced natural by artificial needs. The artificial uniformity of behavior that society imposes on people causes them to ignore "the duties of man and the needs of nature," so that appearance and reality are constantly at variance in modern social life, as for example in the case of an excessive regard for politeness and convention concealing the most ruthless and calculating egoism. Likewise, insisted Rousseau, the sciences and the arts, in spite of their brilliance, are not the genuine expression of fundamental human needs but the result of pride and vanity. The rapid growth of luxury and idleness serves merely to increase the corruption of the contemporary situation. Consequently, as culture appears to attain an ever increasing splendor, genuine human relationships become steadily weaker. Man is alienated from his original nature and prevented from being his real self; a perpetual prey to inner contradictions, he vainly grasps at objects outside himself as he neglects the true lessons of nature in order to pursue the illusions of opinion.

To "society" Rousseau opposed "virtue"—a constant theme of his early works. Virtue confers stability and unity upon human existence because it subordinates idle speculation to the active needs of the moral life. Unlike mere reflection, it induces "strength and vigor of soul," allowing full expression to man's genius and conferring on his existence a solidity and permanence that are quite unlike the ephemeral brilliance of contemporary culture. Whereas society forces man to assume the mask of

hypocrisy and deceit as a means of satisfying his selfish interests, virtue, "the sublime science of simple souls," gives him an authentic openness and innocence that allow him to reveal himself to others as he truly is.

A particularly serious feature of modern society is the prevalence of an unnatural inequality based on power and wealth. In the *Discours sur l'origine de l'inégalité* Rousseau examines this phenomenon in the light of man's evolution from the primitive state to his present existence as a political being and concludes that modern conditions represent a fall from happiness into misery. In spite of its historical form, this discourse, as the author himself admits, is a purely hypothetical and imaginative reconstruction that deliberately ignores facts, whether historical or theological, in order to concentrate on the nature of man as it is revealed to Rousseau's intuitive perception. If the state of nature can never be known as a historical fact, it at least serves as a useful concept that enables him to distinguish man's original qualities from fortuitous historical accretions.

Limited and instinctive though the life of primitive man may have been, it was at least a happy one inasmuch as the savage knew how to live in accordance with his own innate needs. Leading an isolated existence in the forests, satisfying his basic appetite for food and sex without difficulty, untouched by modern man's anxiety before illness and death, he was largely self-sufficient; the primordial urge toward self-preservation was effectively counterbalanced by an innate feeling of natural pity that prevented him from inflicting needless pain upon his fellow men. Man was from the outset endowed with free will and perfectibility, but these became active only when the first rudimentary social communities, based mainly on the family, were established, a period that Rousseau treats as the golden age of humanity since it lay halfway between the brutishness of primitive existence and the corruption of political societies. The discovery of agriculture and metallurgy and the distinction between "mine" and "thine" meant that people had to work together, and this inevitably led to the establishment of property. Men then became divided into rich and poor and, later, into powerful and weak, so that the inequality of the social system was at last made permanent through the institution of laws and political organization. In Rousseau's opinion the historical process will culminate in the triumph of despotism, which makes all men once again equal because all have become slaves of one master.

Whereas the early discourses dealt mainly with general principles, the publication of d'Alembert's article "Geneva" in the seventh volume of the *Encyclopédie* in 1757, with its suggestion that the Genevans would benefit from the establishment of a theater, led Rousseau to deal with a specific aspect of his criticism of society. In his various replies to early critics he had already insisted that man, having once left the primitive state, could never return to it; he also maintained that it was the large states, especially the monarchies of Europe, which had traveled furthest on the road to perdition. However, small republics like Geneva, though no longer close to nature, still retained a relative simplicity and innocence and could be protected against further corruption. To introduce the theater into Geneva was, in Rousseau's eyes, to bring an evil product of society into a comparatively unspoiled community. The *Lettre à d'Alembert* also provided him with an opportunity of examining not only the general characteristics of the theater but also the whole question of what amusements are best suited to man's true nature.

Starting from the assumption that all valid entertainment must "derive from man's work, relationships and needs," Rousseau insists that it must be an integral part of man's daily life, different from his work and yet inspired by the same spirit. The theater, however, is primarily an artificial entertainment, the product of idleness and vanity and the fomenter of dangerous passions and emotions; it is always subservient to the impulses that create it and remains incapable of directing people toward moral activity. The theater is typical of a large city like Paris, with its reversal of natural values. Whereas for Rousseau woman is naturally modest and self-effacing, the theater makes her a shameless figure who transforms love into a public spectacle; the very existence of actresses also sets the example of a completely unfeminine way of life that is characteristic of a society in which women set the tone and rule the *salons,* reducing men to a condition of abject and effeminate dependence. By contrast, Rousseau extols the simplicity of the Montagnons, the simple, industrious mountain dwellers whom he remembered from his youth and recalls with heartfelt enthusiasm. Unlike modern men such people relied upon their own creative resources for their work and entertainment. The Genevans, too, through their "societies" and "circles," wisely allowed men and women to indulge in their own separate pastimes. The *Lettre* ends with a remarkable evocation of the kind of national entertainment that, in Rousseau's opinion, would be suitable for a small homogeneous community like Geneva. The Genevans should actively participate in a joyous public entertainment that takes place "beneath the sky" and in the presence of their fellow citizens; in this way the whole community would be inspired by feelings that are both social and human.

Perhaps one of the gravest general aspects of society's harmful influence on the nature of man is its constant tendency to transform *amour de soi* (self-love) into *amour-propre* (pride). Although this antithesis was not peculiar to Rousseau, who had already noted its existence in Vauvenargues, it does, occupy a particularly significant position in his social criticism. *Amour de soi* is always good and, in its purest state, quite spontaneous because it expresses the real essence of human existence. It is an absolute feeling or passion that serves as the source of all genuinely natural impulses and emotions; already revealing itself at the instinctive level as the desire for self-preservation, it assumes a much nobler expression as soon as it is combined with reason. Being in complete uniformity with the principle of order, it will affect all the main aspects of human existence as it brings the individual into contact with his own inner self, his physical environment, and his fellow men. Unfortunately modern society has changed this natural *amour de soi,* which makes a man what he truly is, into *amour-propre,* an artificial reaction originating in an anxious reflection that induces a man to be forever comparing himself with others and even finding his sole pleasure in their misfortune or inferiority; through *amour-propre* he is taken outside himself into the realm of illusion and opinion and so prevented from being a complete person.

EDUCATION. Having diagnosed the malady of modern civilization, Rousseau was faced with the task of suggesting a cure, and this led him into the domain of education and politics, activities that are, or should be, rooted in man's moral nature. Rousseau was convinced that man's original nature is good, but that it has been corrupted mainly by the historical accident of society. It therefore seemed quite consistent to affirm that men are wicked but that man himself is good. To be good is to exist in accordance with the intrinsic potentialities of one's nature, and *Émile* seeks to trace the natural development of a human being brought up in the country away from the nefarious influence of contemporary social life. From this point of view the work is not just a manual of education but, as Rousseau himself points out, a philosophical treatise on the goodness of human nature. It is less concerned with laying down the practical details of a specific pedagogic method than with describing the fundamental principles that underlie the whole of man's development from infancy to maturity. Rousseau's ultimate object is to teach the art of living, for man's first duty, he says, is to be human.

The educator must realize that "vice and error, alien to man's constitution, are introduced into it from outside"; his first task will be to keep away harmful influences from the young child. This is why *Émile* is set in a rural environment that allows the child to grow in accordance with his own nature. Early education is therefore largely negative insofar as it is mainly concerned with removing obstacles that might hinder this development.

From the first Rousseau stresses the importance of a progressive education: Each stage of the process must be carefully adapted to the individual's developing needs and so follow "the natural progress of the human heart." In this respect Rousseau uses in his own way the genetic method of contemporary thinkers like d'Alembert, Condillac, and Comte de Buffon, who, in turn, had taken over the notion of the genealogy of ideas developed by John Locke in his famous *Essay concerning Human Understanding.* In *Émile,* however, as in the *Discours sur l'origine de l'inégalité,* Rousseau does not strive to establish an inductive law based on the empirical examination of facts but starts from a fundamental principle (man's natural goodness) that is derived initially from personal intuition, though he believes it to be subsequently verifiable by observation and psychological analysis. *Émile* therefore involves certain metaphysical elements, but these are referred back to the concrete aspects of human nature.

Rousseau maintains that a truly progressive education will recognize that the child has his own special needs as a being who exists in his own right. "Nature wants children to be children before being men.… Childhood has its own ways of seeing, thinking and feeling." Since the child's needs are largely physical, negative education "tends to perfect the organs, instruments of our knowledge." Incapable of dealing with abstractions, the child must be educated through contact with things. To him dependence on things will be natural and inevitable; acknowledging only the "heavy yoke of necessity," he will escape the tyranny of any human will. Unlike the despotic power of men necessity is quite compatible with properly controlled freedom since it lets the human being exercise his powers within the limits prescribed for him by nature. "The truly free man" wishes to do no more than this. Well-regulated freedom thus provides the only valid basis and aim of sound education.

Early education, being based primarily on the senses, ignores bookish learning for direct contact with the physical world. Learning through a process of trial and error, the child experiments, as it were, through the medium of facts rather than words. (The sole book Rousseau will allow the child is Daniel Defoe's *Robinson Crusoe* and that only because it describes a man's reliance on his own

ingenuity and resourcefulness.) Freed from the tyranny of human opinion, the child identifies himself effortlessly with the requirements of his immediate existence; content to be himself and completely absorbed in his present being, he leads a kind of self-sufficient, timeless existence that knows no anxious concern for the future, none of that tormenting foresight that causes modern man to be so unhappily "outside himself." The child is happy because he is unaware of artificial needs or of any serious disproportion between capacity and desire, power and will, and in this respect he resembles the happy savage.

Rousseau recognizes that even at the stage of greatest inner harmony, the child must be prepared for the future, for in him there is a reservoir of potential energy that he does not immediately need. The educator's task is to hold back this energy, this "superfluous aspect of his immediate being," until it can be effectively used. It is particularly important to avoid any precocious excitement of the imagination that may be the source of future unhappiness. These dangers will be largely averted if, after the lesson of necessity, the child learns that of utility, his developing reason being applied only to what interests and helps him. That is why his early judgment must be formed not through words or abstractions but through sensations and feelings.

A truly positive education begins only when the child becomes aware of his relationships with other people, although these early social lessons will be based on sensibility rather than reason, in particular on the innate feeling of pity, with its later concomitants of love and aversion. There are no good or bad passions, says Rousseau. All are good when they are under our control; all are bad when they control us. Through the force of our passions we are impelled beyond ourselves; through the "superabundance of our strength" we are induced to "extend our being." With the growth of sensibility, reason, and imagination the child leaves the self-sufficiency of the primitive stage for a fuller life involving relations with the physical realm of nature and the world of human beings. The educational process must therefore be carefully timed and controlled so that the various potentialities of the human being are brought to fulfillment in an orderly and harmonious manner.

It is clear from the last book of *Émile* that man must be educated for society, though not necessarily for society in its present form. Man's nature is not fully mature until it becomes social. However, the natural man in the state of nature and the natural man in the social state cannot be identical, for whereas the former is predominantly an instinctive, primitive creature living on the spontaneous expression of his innate vitality, the latter is a rational, moral being aware of his obligations to other people, a man called upon to subordinate the impulse of "goodness" to the demands of "virtue." Therefore, only in society can a genuinely human morality become possible. If by "nature" is meant the merely primordial responses of the presocial man, then it is true to say that "good institutions denature man" inasmuch as they raise him up from the absolute self-sufficiency of the isolated primitive state to the level of a moral, relative existence based on an awareness of the common good and the need to live in harmonious relationship with his fellow men. Since morality inevitably involves the problem of man's life as a social being, it is impossible to separate morality and politics, and Rousseau states most emphatically that "those who want to treat morality and politics separately will never understand anything about either." This is a most important aspect of his political thinking. If "nature" intended man for a moral existence, then it also intended him for social life; indeed, only through the individual's participation in the "common unity" can full personal maturity become possible. "Nature" is still the norm, but one that has to be re-created, as it were, at a higher level, conferring on man a new rational unity that replaces the purely instinctive unity of the primitive state.

POLITICAL THEORY. There appears to be no valid reason for finding, as some critics have done, any fundamental contradiction between *Émile* and the *Contra social.* Such a difficulty arises only when anachronistic attempts are made to explain Rousseau's thought in purely individualist or collectivist terms. If at first sight Émile seems to be an isolated individual, this is mainly because Rousseau wanted to stress the importance of the human being's natural development, and it in no way excludes the idea that all true education must eventually be for society.

In itself the particular form of education, like that of government, must be determined by specific historical and physical conditions, but Rousseau was less concerned with this question than with that of the fundamental principles on which all true education and all true government must be based. In this respect *Émile* and the *Contrat social* are similar since each is a theoretical, normative work. Rousseau points out in his correspondence that the *Contrat social* is a philosophical discussion of political right (the work is actually subtitled *Principes du droit politique*) rather than an examination of any existing form of government. As he says in the introduction to his work, he is taking "men as they are" and "the laws as they *can* be." He seeks to reconcile "what right permits

with what interest prescribes, so that justice and utility are not divided." In Rousseau's eyes this is what distinguishes his approach to political problems from Baron de Montesquieu's. Whereas Montesquieu is concerned with "the positive right of established governments," Rousseau, as the theorist of political right, examines the philosophical basis of all legitimate government.

Although the *Contrat social* has often been described as the forerunner of totalitarianism, this interpretation is certainly not consistent with Rousseau's conscious intention, for from the very outset his overriding preoccupation is the same as it was in *Émile*—the problem of freedom. No doubt, just as the concept of nature undergoes a radical transformation when it is applied to society, so the natural freedom enjoyed by man in the state of nature differs in important respects from the civic freedom of the social state; both, however, are natural to man at different stages of his development. Man living in society faces a problem that does not affect primitive man—namely, the possible tyranny of his fellow men. Now, a true and just society can never be based on sheer force, for right can never be equated with might. Rousseau vigorously repudiates traditional views that seek to justify the right of conquerors to subject the vanquished to permanent enslavement; no society founded on such a principle can ever be legitimate. Man's participation in society must be consistent with his existence as a free and rational being. Society is therefore unthinkable without a freedom that expresses man's most fundamental attribute. "To give up freedom is to give up one's human quality: to remove freedom from one's will is to remove all morality from one's actions." Moreover, it is with the emergence of society that man comes into possession of his freedom and thus attains the status of a moral being. The institution of any genuine political society must be the result of a social pact, or free association of intelligent human beings who deliberately choose to form the type of society to which they will owe allegiance; this is the only valid basis for a community that wishes to live in accordance with the requirements of human freedom.

However, there still remains the problem of finding a form of association that will continue to respect the freedom that brought it into being. Although man is naturally good, he is constantly threatened by forces that not only alienate him from himself but also transform him into a tyrant or a slave. From this point of view the political problem is not dissimilar from the pedagogic one. How is man to be protected from the tyranny of the human will? Just as the child has to be liberated from dependence upon human caprice in order to confront necessity, so the individual is to be preserved from tyranny by "an excessive dependence" of all citizens on a new kind of necessity, on something that is greater than the citizen himself and yet in one sense a part of his life. Rousseau seeks a form of association in which "each one uniting with all obeys, however, only himself and remains as free as before." In other words, "each one giving himself to all gives himself to nobody." The possibility of inequality and injustice will be avoided through the "total alienation of each associate, with all his rights, to the community"; if such alienation were less than total, it would expose the individual to domination by others. As it is, the citizen does not obey some sectional interest but the general will, which is a "real force, superior to the action of any particular will." Nor, in Rousseau's view, need this arouse any apprehension, for unlike the individual will which concerns itself with specific and perhaps selfish interests, the general will is always directed toward the general good.

Moreover, total alienation involves equality in another way; the general will is not simply an external authority that the citizen obeys in spite of himself but the objective embodiment of his own moral nature. In accepting the authority of the general will, the citizen not only belongs to a collective, moral body but also achieves true freedom by obeying a law that he has prescribed for himself. Through the law he escapes from the bondage of appetite in order to follow, as an intelligent being, the dictates of reason and conscience. Submission to a will possessing an "inflexibility which no human force could ever overcome" leads to a freedom that "keeps a man exempt from vice" and to "a morality which lifts him up to virtue." The individual is thereby invested with another kind of goodness, the genuine virtue of the man who is not an isolated being but part of a great whole. Liberated from the narrow confines of his own being, he finds fulfillment in a truly social experience of fraternity and equality with citizens who accept the same ideal.

This conception of political right is essentially democratic insofar as the source of all political authority and, therefore, of all true sovereignty must always lie with the people as a whole. Moreover, such sovereignty is both inalienable and indivisible since, as the basis of freedom itself, it is something that can never be renounced by the people or shared with others. However, Rousseau establishes an important distinction between sovereignty and government. The sovereign, or subjects (for "sovereign and subjects are simply the same people in different respects"), may delegate the executive function of the state to the prince, or government, which thus becomes

the agent, or officer, of the people; this is true whatever the form of any particular government, whether monarchy, aristocracy, or republic. If every legitimate government is democratic in essence, this does not mean that democracy, as a definite political institution through which the people themselves carry on the government by assembling as a body, is either possible or desirable in modern conditions. Any specific form of government, as Rousseau was to show very clearly in his *Projet de constitution pour la Corse* (1765) and his *Considérations sur le gouvernement de la Pologne* (probably written about 1770–1771), will depend on a variety of historical and geographical factors.

Law, as the act of the general will and the expression of sovereignty, is of vital importance, for the establishment of sound laws can determine the whole destiny of the state. As Rousseau observes, only the gods themselves would be capable of giving good laws to the human race. That is why the legislator has such an important role in the *Contrat social*; he is invested with a remarkable, almost divine quality. It is from him that the citizen "receives in some way his life and his being"; through the legislator's actions he experiences a genuine transformation of his personal life, forsaking the "physical, independent existence he received from nature" for a moral existence as a social being. This new mode of existence is not something imposed upon him from the outside but a possibility elicited from the depths of his inner self. The legislator is in one respect an almost godlike figure, but his purpose is to serve the essential needs of human nature.

At the end of the final version of the *Contrat social* (though not in the original draft), Rousseau seems to acknowledge that an even more powerful sanction may be required to ensure complete political stability, for he proposes to introduce into the state a kind of civil religion or civic profession of faith to which every citizen, having once given his free assent, must remain obedient under pain of death. This is an aspect of Rousseau's political thought that many commentators have found either shocking or inconsistent. However, it will already be clear that Rousseau is no liberal in the classical political sense since he does not believe in the possibility of any rigid separation of the individual and the state; the development of a full moral life is inconceivable without active participation in society, and the unity and permanence of the state depend, in turn, upon the moral integrity and undivided loyalty of its citizens. This civic profession of faith is deliberately restricted to the "few simple dogmas" that, according to Rousseau, every rational, moral being ought readily to accept: belief in a supreme being, the future life, the happiness of the just, and the punishment of the wicked, together with a "single negative dogma, the rejection of intolerance." Anybody repudiating these principles would presumably be, in Rousseau's opinion, little more than a criminal who, by forfeiting his right to be considered as a responsible human being, threatens the state with anarchy and dissolution. The practical implications of this view may still sound alarming to a modern liberal, but they are not necessarily inconsistent with Rousseau's ideas.

RELIGION. If the chapter "Civil Religion" seems to strike a new note in the *Contrat social,* it is certainly not incompatible with the religious emphasis of Rousseau's thought, for religion had always played an important role in his work, as the "Profession de foi du vicaire savoyard" made clear. Nature itself must be understood in the widest sense, as the whole realm of being originally created by God, who guarantees its goodness, unity, and order. Rousseau offended traditional Christian orthodoxy with his belief that man needs no intermediary between himself and God and is able to attain salvation by his own efforts. (In spite of his great respect for the figure of Jesus and the message of the Gospels, Rousseau could not accept the notion of the Incarnation as a solution to the problem of human sin.) But Rousseau never doubted the importance of accepting God's existence; man, he believed, is impelled toward God by the evidence of both feeling and reason, for apart from the presence of intelligence in the universe there is also the sensitive man's deep "feeling for nature" and the inescapable conviction of a real bond uniting his immortal soul with the spiritual order that underlies the outward appearance of the physical world.

As is well known, Rousseau was the eighteenth-century writer who gave particularly eloquent expression to this aspect of the "feeling for nature." Furthermore, apart from the testimony of reason and sensibility there is also that of the all-powerful conscience, the "divine instinct" or "voice of the soul" which forms the basis of man's moral existence. In moments of doubt and perplexity, when all else fails man, he can always rely for guidance on the promptings of his conscience. This does not mean that reason is thereby excluded, for reason is to be condemned only when it becomes the instrument of blind passion or selfish reflection—in other words, when it fails to recognize its dependence upon other essential elements of human nature. Conscience, however, is an even more important attribute; it is a fundamental feeling that is strikingly effective when reason may be impotent.

Even so, conscience, reason, and freedom are all integral elements of man's natural endowment, potentialities that it is his right and duty to develop, for God gave him "conscience to love the good, reason to know it and freedom to choose it." It is only through the harmonious development of all man's faculties that he can come to a full understanding of his own nature and the place allotted to him by God in the universal order.

At first sight Rousseau's philosophy seems to retain many characteristics of the traditional metaphysical outlook, and several critics have stressed his great admiration for Plato and Nicolas Malebranche. In Rousseau's eyes the universe still possesses a rationality, order, and unity that reflect the wisdom and intelligence of its creator. Yet this cannot be known by reason alone, for although reason has a function in all reflection about the meaning of the world, the heart may often provide surer insights into the ultimate mystery of creation. Moreover, Rousseau's system took the form of a series of basic intuitions that he subsequently linked together into a unified whole. His thought, therefore, is imbued with a strongly personal element that excludes any purely abstract or rationalistic speculation about the ultimate meaning of reality. What concerns him is that part of reality which is identified with the nature of man. The nature of man is, of course, inseparable from nature in the wider sense, but sensibility and feeling, rather than mere reason, are probably the most effective means of penetrating this wider objective realm of being. The thinker concerned with fundamental truths will do well, in Rousseau's view, to concentrate on what is of interest to him, "interest" here being defined not in any narrowly pragmatic or empirical sense but as indicating those matters that appertain to man's original nature. This means that Rousseau finally emerges as a moralist rather than as a traditional metaphysician.

Since reflection on the nature of man involves the ability to distinguish between reality and appearance, between the genuinely original and the merely artificial aspects of existence, the thinker's first task must be to abandon the illusions of opinion for the truths of nature. This explains both the negative, critical aspects of Rousseau's views of modern society and his more positive, constructive efforts to elaborate a philosophy of man. If his interpretation of nature seemed too optimistic to satisfy the demands of contemporary religious orthodoxy, it was also too religious to please the advocates of philosophical skepticism or materialism. Of one thing Rousseau felt quite certain: To ignore or reject the profound moral and spiritual aspects of human existence could have only the most disastrous consequences for the welfare of humanity. The discovery of truth requires an active renewal of the whole man and a reawakened moral consciousness that acknowledges the full implications of man's situation in the universe; the genuine possibilities of human life cannot be separated from the universal order of which they are a part, and man's ultimate felicity is to feel himself at one with a God-created "system in which all is good."

Like so many of his contemporaries Rousseau considered happiness to be the legitimate goal of human endeavor, but he insisted that "enjoyment" must not be interpreted in a shallow or selfish manner. Happiness consists of being oneself and of existing according to one's own nature, but a nature that has been purified of all extraneous artificial elements. When truly fulfilled, man will experience satisfaction with himself and a sense of being identified with the pure "feeling of existence"; this, in turn, presupposes the ability to find a true personal unity and plenitude. No doubt, Rousseau's efforts to realize this ideal in his own life were not free from ambiguity and contradiction, as an examination of his personal writings well shows, but his didactic works are consistent in their main objective.

In a corrupt society the recovery of a full human existence can never take the form of a mere return to nature, for the nature of man cannot be equated with the primordial state of nature. Although Rousseau was often nostalgically drawn to the innocence and simplicity of early times, he also treated nature as a dynamic, forward-looking concept. Starting from man as he is, the movement toward nature must be constantly sustained by the vision of what man might be. The achievement of this goal requires a radical transformation of human existence, the rediscovery and re-creation of a new nature. At the same time Rousseau did not believe in the need for any kind of supernatural grace to help man to carry out this task, since nature represented an innate possibility that could be realized through the wise exercise of human freedom alone.

Rousseau's powerful influence on later generations was partly due to this vision of a regenerated human nature, but unlike merely utopian thinkers he seemed to promise a transfiguration of everyday existence, not the pursuit of a hopeless chimera. Indeed, his philosophy revealed a striking, if often elusive, combination of idealistic and realistic elements that constantly seemed to open up the possibility of a better world. Moreover, this optimistic outlook was transmitted through a particularly eloquent and persuasive style, rich in emotional and musical overtones, giving the impression of intense sin-

cerity and convincing the humblest of men that he need never feel ashamed to call himself a human being.

See also Alembert, Jean Le Rond d'; Analytical Feminism; Authority; Buffon, Georges-Louis Leclerc, Comte de; Condillac, Étienne Bonnot de; Diderot, Denis; Encyclopédie; Equality, Moral and Social; French Philosophy; General Will, The; Human Nature; Hume, David; Malebranche, Nicolas; Philosophy of Education, History of; Plato; Plutarch of Chaeronea.

Bibliography

WORKS BY ROUSSEAU

For bibliographical information about the various editions of Rousseau's works, though not about works on Rousseau, the reader is referred to Jean Sénelier, *Bibliographic générale des oeuvres de J.-J. Rousseau* (Paris, 1949), and Théophile Dufour, *Recherches bibliographiques sur les oeuvres imprimées de J.-J. Rousseau,* 2 vols. (Paris: L. Giraud-Badin, 1925).

Collections and Other Editions

Of the collected editions of Rousseau's works, *Oeuvres complétes de J.-J. Rousseau,* 13 vols. (Paris, 1865–1870), and *Oeuvres et correspondence inédites de J.-J. Rousseau,* edited by G. Streckeisen-Moultou (Paris, 1861), will be superseded by *Oeuvres complétes de Jean-Jacques Rousseau,* edited by Pléiade Bernard Gagnebin and Marcel Raymond (Paris, 1959–), five volumes projected. Similarly, the *Correspondance générale de J.-J. Rousseau,* edited by Théophile Dufour and P.-P. Plan, 20 vols. (Paris: Colin, 1924–1934), with the *Table de la correspondance générale,* edited by P.-P. Plan and Bernard Gagnebin (Geneva, 1953), superseded by the *Correspondance compléte de Jean-Jacques Rousseau,* 51 vols., edited by R. A. Leigh (Geneva, 1965–1991).

Although the new Pléiade edition of the collected works will henceforth be authoritative, the following separate editions are still important: *Political Writings,* edited by C. A. Vaughan, 2 vols. (Cambridge, U.K.: Cambridge University Press, 1915); the critical edition of *La nouvelle Héloïse,* edited by Daniel Mornet, 4 vols. (Paris, 1925); *La profession de foi du vicaire savoyard,* edited by P. M. Masson (Paris, 1914); *Les réveries du promeneur solitaire,* edited by J. S. Spink (Paris, 1948). For the *Confessions* see also the edition by Jacques Voisine (Paris, 1964).

WORKS ON ROUSSEAU

Some general bibliographic information about studies of Rousseau's work is to be found in Albert Schinz, *État présent des travaux sur J.-J. Rousseau* (New York: Modern Language Association of America, 1941); Jacques Voisine, "État des travaux sur J.-J. Rousseau," in *L'information littéraire* 16 (May–June 1964): 93–102; and Peter Gay, *The Party of Humanity, Studies in the French Enlightenment* (New York: Knopf, 1964), Ch. 8, "Reading about Rousseau."

A very pleasant introduction to Rousseau is Bernard Gagnebin's *A la Rencontre de Jean-Jacques Rousseau* (Geneva, 1962), a picturesque collection of texts, documents, and illustrative material. Sound general introductions to Rousseau's thought are Daniel Mornet, *Rousseau l'homme et l'oeuvre* (Paris, 1950); E. H. Wright, *The Meaning of Rousseau* (Oxford: Oxford University Press, H. Milford, 1929), which corrects earlier misinterpretations of Rousseau's idea of nature; and J. H. Broome, *Rousseau: A Study of His Thought* (New York: Barnes and Noble, 1963), to which should be added *Rousseau par lui-même,* edited by Georges May (Paris: Éditions du Seuil, 1962), the pioneering article by Gustave Lanson, "L'unité de la pensée de Jean-Jacques Rousseau," *Annales de la société Jean-Jacques Rousseau* 8 (1912), and Ernst Cassirer, *The Question of Jean-Jacques Rousseau,* translated and edited with an introduction by Peter Gay (New York: Columbia University Press, 1954).

For Rousseau's biography see Jean Guéhenno, *Jean-Jacques,* 3 vols. (Paris: Grasset, 1948–1952), of which a new edition, titled *Jean-Jacques, histoire d'une conscience,* 2 vols., was published in 1962. See also F. C. Green, *Jean-Jacques Rousseau: A Critical Study of His Life and Writings* (Cambridge, U.K.: Cambridge University Press, 1955).

More detailed discussions of various aspects of Rousseau's philosophy are to be found in C. W. Hendel, *Jean-Jacques Rousseau: Moralist,* 2 vols. (New York and Oxford: Oxford University Press, 1934; 2nd ed., 1 vol., New York, 1962), an important examination of Rousseau's intellectual development; Albert Schinz, *La pensée de Jean-Jacques Rousseau* (Paris, 1929), which stresses and perhaps exaggerates the basic conflict between the "romantic" and the "Roman" Rousseau but is nevertheless a significant study; Pierre Burgelin, *La philosophie de l'existence de J.-J. Rousseau* (Paris: Presses Universitaires de France, 1952), an important modern synthesis of Rousseau's thought; Robert Derathé, *Le rationalisme de Jean-Jacques Rousseau* (Paris, 1948), a helpful corrective to earlier sentimentalist interpretations of Rousseau; and Jean Starobinski, *Jean-Jacques Rousseau, la transparence et l'obstacle* (Paris: Plon, 1957), an original and important study of certain key themes in Rousseau's work.

Still important for a study of Rousseau's Genevan background is Gaspard Vallette, *Jean-Jacques Rousseau genevois* (Paris, 1908), although the discussion of the Genevan aspects of Rousseau's thought has been modified by more recent criticism, especially by J. S. Spink, *Rousseau et Genéve* (Paris, 1934), and François Jost, *Jean-Jacques Rousseau suisse,* 2 vols. (Fribourg: Éditions Universitaires, 1961).

Indispensable for any serious study of Rousseau's religious thought is P. M. Masson's *La religion de J.-J. Rousseau,* 3 vols. (Paris, 1916), in spite of some exaggeration of both its Roman Catholic and sentimental elements; for a corrective see Albert Schinz, *La pensée religieuse de J.-J. Rousseau et ses récents interprètes* (Paris, 1927). On Rousseau's political thought see Alfred Cobban, *Rousseau and the Modern State* (London: Allen and Unwin, 1934; 2nd ed., 1964), and Robert Derathé, *Jean-Jacques Rousseau et la science politique de son temps* (Paris: Vrin, 1950), which sets Rousseau's political thought in its contemporary philosophical context. For the difficult question of Rousseau's psychology and personality see Louis Proal, *La psychologic de J.-J. Rousseau* (Paris, 1930), and Suzanne Elosu, *La maladie de J.-J. Rousseau* (Paris, 1929). More recent discussions of this problem and its bearing on the personal writings are to be found in Ronald Grimsley, *J.-J. Rousseau: A Study in Self-Awareness* (Cardiff: University of Wales Press, 1961), and

Marcel Raymond, *J.-J. Rousseau: La Quête de soi et la réverie* (Paris: Librairie J. Corti, 1962).

The year 1962, being the 250th anniversary of Rousseau's birth, was marked by three important international conferences whose proceedings were published: *Jean-Jacques Rousseau et son oeuvre, Colloque de Paris, 16–20 Octobre,* 1962 (Paris: Librairie C. Klincksieck, 1964); *Annales de la société Jean-Jacques Rousseau,* Vol. 35, *Entretiens sur J.-J. Rousseau* (Geneva, 1962); and *Études sur le contrat social de J.-J. Rousseau* (Paris: Belles Lettres, 1964), proceedings of Dijon conference of May 1962.

It has not been possible to include in this bibliography many important articles on Rousseau. For further information on this and other subjects, the reader is referred to the indispensable *Annales de la société Jean-Jacques Rousseau* (Geneva), published from 1905 on, which contain not only original articles but a full review or Rousseau literature.

Ronald Grimsley (1967)

ROUSSEAU, JEAN-JACQUES [ADDENDUM]

The writings of Jean-Jacques Rousseau continue to attract a wide range of readers throughout the world. Persistent questions concerning nationalism, political legitimacy, and the social costs of technological progress sustain an ongoing interest in Rousseau's major political writings (*The Social Contract, Considerations on the Government of Poland*, the first and second discourses). Controversies over child-rearing, the nature of language, and the role of the media in public life keep alive his educational and cultural writings (*Emile, Essay on the Origin of Languages, Letter to d'Alembert on the Theater*). Speculations about psychology and the arts of autobiography draw readers to Rousseau's personal writings (*The Confessions, Reveries of a Solitary Walker, Rousseau Judge of Jean-Jacques*). And new attitudes regarding love, marriage, and eroticism provoke reconsideration of his romantic novel (*La nouvelle Héloïse*). As the editors of a 1978 issue of *Daedalus* commemorating the bicentennial of Rousseau's death observed, Rousseau anticipated many of the moral, political, social, and aesthetic concerns that continue to preoccupy us today.

Three intellectual currents have contributed significantly to a growing body of scholarship on Rousseau. Feminist studies have offered fresh interpretations of his notoriously controversial writings about the nature, education, and status of women (see esp. *Emile*, book 5). Some feminist theorists (e.g., Okin, 1979) argue that Rousseau's advocacy of sexually differentiated social and political roles contradicts his egalitarian principles and undermines the logic and validity of his political theory. Others (e.g., Weiss, 1994) maintain that sexual differentiation constitutes a necessary social construct undergirding the unity of his entire system. At issue in many of these debates are fundamental questions about the usefulness for modern feminism of any theory that posits a close connection between a woman's essential "nature" and her moral role in society.

Deconstruction has also affected the content and direction of Rousseau criticism, especially among scholars in language and literature departments. The French philosophers and literary critics who originated this movement in the 1960s and 1970s gave prime place to Rousseau in the development of their ideas (see, e.g., Derrida, 1976). In seeking to expose the indeterminacy of the meaning of Rousseau's texts by examining details that are commonly overlooked (e.g., footnotes, metaphors, his choice of particular terms), deconstructionist critiques illuminate the multilayered quality of his prose and show that even an author committed to the truth may produce writings fraught with artifice.

A third important source of Rousseau criticism has been the legacy of Leo Strauss (1899–1973)—a political philosopher who is as well known for the habits of close textual analysis he passed on to his students as for the ideas put forth in his own writings (see, e.g., Strauss, 1953). Straussian interpretations take seriously Rousseau's claims that his political thought forms a single coherent system; they also emphasize his debt to classical sources. Most important, perhaps, the Straussian legacy includes a substantial number of English translations of Rousseau's work (e.g., by Allan Bloom, Victor Gourevitch, Christopher Kelly, Judith R. Bush, and Roger D. Masters)—thus making him more accessible to the general reader in North America.

Rousseau specialists have benefited from the publication of Rousseau's *Oeuvres complètes* and *Correspondance complète*, from the appearance of scholarly journals and associations devoted to Rousseau studies (*Annales de la Société Jean-Jacques Rousseau, Études Jean-Jacques Rousseau*, and the *Proceedings* of the North American Association for the Study of Rousseau), and from the publication of papers delivered at various conferences held in 1978 to commemorate his death and in 1989 to mark his relationship to the French Revolution.

See also Deconstruction; Derrida, Jacques; Love; Nationalism; Rousseau, Jean-Jacques.

Bibliography

WORKS BY ROUSSEAU

Oeuvres complètes. 5 vols, edited by B. Gagnebin and M. Raymond. Paris: Gallimard, 1959–.

Correspondance complète de Jean-Jacques Rousseau. 51 vols, edited by R. A. Leigh. Geneva: Institut et Musée Voltaire, 1965–1991.

Emile: or, On Education. Translated by A. Bloom. New York: Basic, 1979.

The Collected Writings of Jean-Jacques Rousseau, edited by R. D. Masters and C. Kelly; translated by J. R. Bush, C. Kelly, and R. D. Masters. Hanover, NH, 1990–.

WORKS ON ROUSSEAU

Affeldt, Steven G. "The Force of Freedom: Rousseau on Forcing to Be Free." *Political Theory* 27 (3) (1999): 299–333.

Bloom, Harold, ed. *Jean-Jacques Rousseau*. New York: Chelsea House, 1988.

Boyd, Richard. "Pity's Pathologies Portrayed: Rousseau and the Limits of Democratic Compassion." *Political Theory* 32 (4) (2004): 519–546.

Charvet, John. *The Social Problem in the Philosophy of Rousseau*. Cambridge, U.K.: Cambridge University Press, 1974.

Cooper, Laurence D. *Rousseau, Nature, and the Problem of the Good Life*. University Park: Pennsylvania State University Press, 1999.

Cranston, Maurice William and Richard S. Peters, eds. *Hobbes and Rousseau: A Collection of Critical Essays*. Garden City, NY: Anchor Books, 1972.

Cranston, Maurice William. *Jean-Jacques: The Early Life and Work of Jean-Jacques Rousseau, 1712–1754*. New York: Norton, 1982.

Cranston, Maurice William. *The Noble Savage: Jean-Jacques Rousseau, 1754–1762*. Chicago: University of Chicago Press, 1991.

Cranston, Maurice William. *The Solitary Self: Jean-Jacques Rousseau in Exile and Adversity*. Chicago: University of Chicago Press, 1997.

Dascal, Marcelo. "'Aporia' and 'Theoria': Rousseau on Language and Thought." *Revue Internationale et Philosophie* 32 (1978): 214–237.

Dent, N. J. H "Jean-Jacques Rousseau." In *The Blackwell Guide to the Modern Philosophers from Descartes to Nietzsche*, edited by Steven Emmanuel. Cambridge, MA: Blackwell, 2001.

Dent, N. J. H. *Rousseau: An Introduction to His Psychological, Social and Political Theory*. New York: Blackwell, 1988.

Dent, N. J. H. *A Rousseau Dictionary*. Cambridge, MA: Blackwell Reference, 1992.

Ferrara, Alessandro. *Modernity and Authenticity: A Study in the Social and Ethical Thought of Jean-Jacques Rousseau*. Albany: State University of New York, 1993.

Fralin, Richard. *Rousseau and Representation*. New York: Columbia University Press, 1978.

Grimsley, Ronald. *The Philosophy of Rousseau*. New York: Oxford University Press, 1973.

Hall, John C. *Rousseau, An Introduction to his Political Philosophy*. London: Macmillan, 1973.

Havens, George Remington. *Jean-Jacques Rousseau*. Boston: Twayne, 1978.

Hedman, Carl G. "Rousseau on Self-Interest, Compassion, and Moral Progress." *Revue de l'universite d'Ottawa* 49 (1979): 430–447.

Hiley, David R. "The Individual and the General Will: Rousseau Reconsidered." *History of Philosophy Quarterly* 7 (2) (1990): 159–178.

Hoffman, Piotr. *Freedom, Equality, Power: The Ontological Consequences of the Political Philosophies of Hobbes, Locke, and Rousseau*. New York: Lang, 1999.

Horowitz, Asher. *Rousseau, Nature, and History*. Toronto: University of Toronto Press, 1987.

Howard, Dick. "Rousseau and the Origin of Revolution." *Philosophy and Social Criticism* 6 (1979): 349–370.

Johnston, Steven. *Encountering Tragedy: Rousseau and the Project of Democratic Order*. Ithaca, NY: Cornell University Press, 1999.

Kain, Philip J. "Rousseau, the General Will, and Individual Liberty." *History of Philosophy Quarterly* (1990): 315–334.

Kavanagh, Thomas M. *Writing the Truth: Authority and Desire in Rousseau*. Berkeley: University of California Press, 1987.

Kukla, Rebecca. "The Coupling of Human Souls: Rousseau and the Problem of Gender Relations." In *Political Dialogue: Theories and Practices*, edited by Stephen L. Esquith. Amsterdam: Rodopi, 1996.

Lange, Lynda. *Feminist Interpretations of Jean-Jacques Rousseau*. University Park: Pennsylvania State University Press, 2002.

Lauth, Reinhard. "Le systeme des pensees de Rousseau selon a Philonenko." *Archives de Philosophie* 50 (1987): 23–53.

Leigh, R. A., ed. *Rousseau after Two Hundred Years: Proceedings of the Cambridge Bicentennial Colloquium*. New York: Cambridge University Press, 1982.

Levine, Andrew. *The General Will, Rousseau, Marx, Communism*. New York: Cambridge University Press, 1993.

Levine, Andrew. "Beyond Justice: Rousseau against Rawls." *Journal of Chinese Philosophy* 4 (1977): 123–142.

Masters, Roger D. *The Political Philosophy of Rousseau*. Princeton, NJ: Princeton University Press, 1968.

Miller, Jim. *Rousseau: Dreamer of Democracy*. New Haven, CT: Yale University Press, 1984.

Morgenstern, Mira. *Rousseau and the Politics of Ambiguity: Self, Culture, and Society*. University Park: Pennsylvania State University Press, 1996.

Morris, Christopher W. *The Social Contract Theorists: Critical Essays on Hobbes, Locke, and Rousseau*. Lanham, MD: Rowman and Littlefield, 1999.

O'Hagan, Timothy, ed. *Jean-Jacques Rousseau and the Sources of the Self*. Brookfield, MA: Avebury, 1997.

O'Hagan, Timothy. *Rousseau*. London: Routledge, 1999.

Orwin, Clifford and Nathan Tarcov, eds. *The Legacy of Rousseau*. Chicago: University of Chicago Press, 1997.

Perkins, Merle L. *Jean-Jacques Rousseau on the Individual and Society*. Lexington: University Press of Kentucky, 1974.

Qvortrup, Mads. *The Political Philosophy of Jean-Jacques Rousseau: The Impossibility of Reason*. New York: Manchester University Press, 2003.

Rapaczynski, Andrzej. *Nature and Politics: Liberalism in the Philosophies of Hobbes, Locke, and Rousseau*. Ithaca, NY: Cornell University Press, 1987.

Reisert, Joseph R. *Jean-Jacques Rousseau: A Friend of Virtue*. Ithaca, NY: Cornell University Press, 2003.
Riley, Patrick, ed. *The Cambridge Companion to Rousseau*. New York: Cambridge University Press, 2001.
Rosenblatt, Helena. *Rousseau and Geneva: From the "First Discourse" to the "Social Contract," 1749–1762*. New York: Cambridge University Press, 1997.
Ryan, Bruce A. "Jean-Jacques Rousseau and Behavior Control: The Technology of a Romantic Behaviorist." *Behaviorism* 4 (1976): 245–256.
Scholz, Sally J. *On Rousseau*. Belmont, CA : Wadsworth/ Thomson Learning, 2001.
Schwartz, Joel. *The Sexual Politics of Jean-Jacques Rousseau*. Chicago: University of Chicago Press, 1984.
Shaver, Robert. "Rousseau and Recognition." *Social Theory and Practice* 15 (1989): 261–283.
Simon Ingram, Julia. "Rousseau and the Problem of Community: Nationalism, Civic Virtue, Totalitarianism." *History of European Ideas* 16 (1–3) (1993): 23–29.
Starobinski, Jean. *Jean-Jacques Rousseau, Transparency and Obstruction*. Translated by Authur Goldhammer. Chicago: University of Chicago Press, 1988.
Steinberg, Jules. *Locke, Rousseau, and the Idea of Consent: An Inquiry into the Liberal-Democratic Theory of Political Obligation*. Westport, CT: Greenwood Press, 1978.
Still, Judith. *Justice and Difference in the Works of Rousseau*. New York: Cambridge University Press, 1993.
Strong, Tracy B. *Jean Jacques Rousseau: The Politics of the Ordinary*. Lanham, MD: Rowman and Littlefield, 2002.
Trouille, Mary Seidman. *Sexual Politics in the Enlightenment: Women Writers Read Rousseau*. Albany: State University of New York Press, 1997.
Viroli, Maurizio. *Jean-Jacques Rousseau and the 'Well-Ordered Society.'* Translated by Derek Hanson. New York: Cambridge University Press, 1988.
Vulliamy, Colwyn E. *Rousseau*. Port Washington, NY: Kennikat Press, 1972.
Weirich, Paul. "Rousseau on Proportional Majority Rule." *Philosophy and Phenomenological Research* 47 (1986): 111–126.
Weiss, Penny A. *Gendered Community: Rousseau, Sex, and Politics*. New York: New York University Press, 1993.
Wokler, Robert. *Rousseau*. New York: Oxford University Press, 1995.
Wokler, Robert. *Rousseau: A Very Short Introduction*. New York: Oxford University Press, 2001.

Grace G. Roosevelt (1967)
Bibliography updated by Tamra Frei (2005)

ROYCE, JOSIAH

(1855–1916)

Josiah Royce, the American idealist philosopher, was born in Grass Valley, California. He received his AB degree from the University of California in 1875 and his doctorate from Johns Hopkins University in 1878. In the intervening years he studied in Germany at Leipzig and Göttingen, where he attended the lectures of Hermann Lotze. Royce returned to the University of California in 1878 as an instructor of English. Four years later, with the help of William James and George Herbert Palmer of the Harvard department of philosophy, he was invited to Harvard, where he taught for two years as a replacement for men on leave; in 1885 he received a regular appointment as assistant professor. Until his death Royce was one of the mainstays of the philosophy department in its so-called golden period. During that time he carried on his friendly debate with William James about the merits and demerits of absolute idealism, supervised the doctoral work of George Santayana, and delivered the Gifford Lectures at the University of Aberdeen in Scotland. Royce was a prolific writer and was much in demand as a public speaker.

PHILOSOPHICAL ORIENTATION

Royce's philosophy is a unique synthesis of the rationalist metaphysic we associate with the system builders in the Western philosophical tradition and the appeal to experience and practice that has been dominant in American philosophy since 1875. Royce is the best American representative of absolute idealism, although there are voluntaristic elements in his position that distinguish it from both the Hegelian position and the systems of the British idealists. Royce's theory of the will and his conception of its role in the knowledge process introduced novel features into the tradition of rationalistic idealism. Royce was aware of this fact and hence called his position absolute voluntarism or absolute pragmatism.

Royce's thought revolves around the problems raised by a religious view of reality. He sought to resolve them through a metaphysical system constructed with the aid of concepts drawn from a wide range of thought and experience. Basic to his position is the concept of the self, an idea that he elucidated in several forms. In his earlier thought the self appears as the Absolute Knower, grasping all truth in one synoptic vision *totum simul*. Later, however, Royce put more emphasis on mediation and on the idea of system. Ultimately, he arrived at the community of interpretation, or social theory of reality, according to which all selves are joined in a Universal Community whose goal is to possess the truth in its totality.

THE NATURE OF BEING

In large measure Royce's idealism consists in his having given to the process of knowing a privileged position in the definition of reality. The nature of Being is to be

determined through the elucidation of the process of being known.

ARGUMENT FROM ERROR. The pattern of the approach through knowing was established early in Royce's development. In a paper, "Kant's Relation to Modern Philosophic Progress" (1881), he argued that the proper task of philosophy is to study the nature of experience, especially the role played by the forms of intellectual activity in knowing. In later works he returned repeatedly to the task of defining the relation between sense and understanding, between the perceptual and conceptual poles in experience and knowledge. Strongly influenced by Immanuel Kant, Royce sought to discover the exact relation between the knowing activity and its matter. He asked how the function of judgment transforms the sensible starting point of all experience into knowledge. Whereas Kant had argued that the past moment and its datum can be brought into the present through the activity of the transcendental subject, Royce regarded the past and future as projections from the present. Knowledge starts with immediate data of sense; these data, as present, are beyond the control of judgment (this is the realistic element in Royce's idealism), but the whole of experience involving reference to a past, a future, and a public object is to be built up from the momentary consciousness. In order to accomplish this construction, judgment and principles of transcendence are required.

Dissatisfied with the view that assigns the status of postulates to the principles needed for transforming immediate data into knowledge, Royce sought to justify those principles. His theory of the Absolute Knower, which he developed in the well-known chapter "The Possibility of Error" in *The Religious Aspect of Philosophy* (Boston, 1885), was intended to show that the conditions for both knowledge and error must themselves be actual; what is actual cannot be explained or justified by what is merely possible or postulated. The argument that is presented for the existence of God or the Absolute Knower may be summarized as follows. Error actually exists; erroneous judgments cannot be made erroneous by finite knowers. In order to be in error, a judgment must fail to agree with its intended object. Yet if the intended object is wholly and completely defined by the isolated judgment, it is difficult to see how the judgment can fail to be true. Royce's central contention is that a judgment can have its own object and at the same time fail to agree with it only if the judgment is not isolated as an entirely enclosed fact but is, instead, part of a system of judgments or an organized body of thought. The isolated judgment cannot have within itself the distinction between its truth and falsity; for that we need an inclusive thought capable of relating the isolated judgment to all other actual and possible judgments about the intended object. In finding error as a fact that we cannot create, we are actually involved in the Infinite Thought. Without that Thought, error is either impossible or unintelligible. This ingenious argument assumes, among other things, that the real individual at which knowledge aims can be identified only at the end of the knowledge process. However, as Charles Peirce and others have shown, there is no need to make this assumption, although without it the argument fails.

THOUGHT AND REALITY. Royce continued to approach the problem of Being—the problem of defining the basic nature of the real—through concentration on the knowledge process. He was also trying to retain critical philosophy and neutralize its negative judgment on the possibility of ontology. His solution was to say that a theory of Being is possible if we can discover the true relation between our ideas and the real world. In *The World and the Individual* (New York, 1901–1902) Royce posed the problem of Being as one of explaining what thought and reality must be like if the former is to attain genuine knowledge of the latter. By means of an extended dialectical argument, Royce examined three classical theories of Being (in his language, theories of "the ontological predicate")—realism, mysticism, and critical rationalism. In subjecting them to critical analysis, he tried to show the element of truth and error in each. From this analysis Royce's own voluntaristic idealism emerged; it was designed to avoid the errors of the other positions while preserving their truth in a new and more comprehensive system that defined Being in terms of purpose fulfilled.

For Royce realism is the doctrine that to be is to be independent of being known. According to realism, the real is just what it is apart from the knower and his acts of knowledge. Royce, however, aimed at exposing this position and hence placed a narrow construction on the term *independent*. To be independent is taken to mean that the idea and object are totally externally related. If the idea and object are thus disconnected, he argued, then knowledge becomes inexplicable, and reality is severed from truth. Peirce, among others, objected to this statement of the realist position, describing it as one-sided.

Mysticism is defined as the thesis that to be is to be immediate. Here again, the real is understood as that which falls effectively beyond the power of analytical reason.

Royce's exposition of critical rationalism, which he defined somewhat cryptically with the formula "to be is to be valid," has been charged with ambiguity; John Dewey claimed that Royce's entire argument was vitiated by his having confused "possible experience" and "validity" in his presentation of the position. Dewey's claim is not without warrant; Royce combined several ideas under one heading, and it is not clear that they are compatible. Nevertheless, Royce's argument is clear enough in its main outline. The critical rationalist does not accept the independent objects of either realism or common sense and still less allows the immediacy of mysticism. Instead, he defines the real as that which gives warrant or validity to our ideas. To be real in this instance means that an object conforms to certain universal forms or conditions—causal sequence, temporal succession, spatial relations, numerical identity, and so on—that are marked out in advance as the general structure of all experience. For Royce the merit of this position is that it comes closer to defining reality in terms of truth than was possible with either realism or mysticism. Critical rationalism, however, is inadequate because it can define or anticipate only the universal form of experience and cannot reach the determinate individual. Royce's point is that the determinate individual cannot be defined in terms of universal conditions of possible experience alone; in order to have knowledge of an individual, we must appeal to actual, sensible experience. But it is just the need for this appeal that marks the defect of the position; a completed rationalistic idealism would show us how to pass from the idea to its fulfillment in the individual object without having to appeal to a brute, sensible experience that is "given." Critical rationalism, however, is forced to rest with "possible experience," by which Royce meant the universal conditions that any proposed object of knowledge would have to satisfy in order to be an object of experience at all. It is important to notice that the entire discussion is dialectical, in the sense that Royce expounds and criticizes the alternative theories only in relation to his own final view. Competing theories fail or succeed precisely to the extent that they are incompatible with, or contribute to the development of, his voluntaristic idealism.

VOLUNTARISTIC IDEALISM. Royce's own view can be summed up in the thesis that to be is to be the individual or determinate fulfillment of a purpose. Distinguishing between the internal and external meaning of ideas, Royce defined an idea as a purpose (internal meaning) seeking its object, or other (external meaning). An idea intends, and thus selects, its object; the object, as the full realization of the idea, must be the determinate individual that allows no other of its kind if it is to be the unique fulfillment of the purpose expressed by the original idea. If we say that Socrates is snub-nosed, our ideas (internal meaning) aim at, or intend, the unique and unduplicable individual Socrates (external meaning). Our ideas are not about just anyone or anything but only about the individual intended; the internal meaning selects the object (external meaning) by reference to which it can be judged true or false. The voluntarism of the position lies in the idea that the other at which all ideas aim is itself the expression of the absolute will or purpose. For Royce it is only in this way that we can explain how an idea can correspond with an object other than itself while that object remains other and yet is the object intended by the idea.

The entire theory is recognizable as a modern version of an ancient doctrine of self-knowledge. We start with an idea that is fragmentary and imperfectly understood, and we seek to find its true meaning in the object that is its individual fulfillment. The object intended exceeds the fragment with which we began; we can discover the true nature of the object and the truth or falsity of our idea only when we have reached the total individual reality that fulfills our purpose. Royce developed this conception of Being into a comprehensive system embracing a doctrine of man, nature, and God. The rational will and its purpose mark the ultimate reality; all finite individuality is what it is in virtue of its fulfilling the purpose of the Absolute Self.

The reality of the infinite. In the essay "The One, the Many and the Infinite" appended to *The World and the Individual*, Royce introduced the topic that was to occupy much of his later thought—the reality of the infinite. He attempted to refute the claim, made by F. H. Bradley in *Appearance and Reality* (1893), that we cannot express in clear concepts the detail of the many facts constituting the Absolute. Since such a claim, if true, would have rendered Royce's entire project pointless, he felt called upon to refute it. To explain how the many develop out of the one, Bradley argued, always leads to an actual infinity, and this is self-contradictory. In the Absolute all is one, but according to Bradley, we are unable to comprehend the unity. Royce denied that an actual infinite is self-contradictory. Through the concept of a self-representative system based on what would now be called a recursive function, he developed a modern version of the actual infinite. The form of the self-representative system was construed as a purpose or an ordering plan and defines once and for all an actual infinity of members. A self-representative system is one that represents itself with all else that it represents. A mirror of the entire universe, for

example, would have to include itself among the represented items. By the form of the system, Royce meant the principle or purpose behind it, which in the above example would be mirroring. From the one form or purpose there comes, by the recurrent or self-representative operation, an infinity of detail such that nothing less than that infinity will serve to express all that was meant by the original form. Understanding the self as having the form of a self-representative system, Royce claimed that the multitude of details constituting the concrete individuality of the real world is an expression of that self. Reality is an actual infinite, a unity of one and many. Royce's later doctrine of the community of interpretation represents his final attempt to elaborate the theory.

Logic and mathematics. It is important to note that Royce took very seriously the development of mathematical logic and studies in the foundation of mathematics. He was fond of criticizing pragmatism for neglect of what he took to be a doctrine of absolute truth implied in the new logic of Gottlob Frege, Bertrand Russell, Giuseppe Peano, and Ernst Schröder. Maintaining that "order is the fundamental category of exact thought about facts," Royce argued for the validity of using technical logical concepts in the construction of a metaphysical theory. Two examples will clarify the point. In the analysis of discrimination, he used the concept of between, arguing that discrimination and comparison are possible because, given any two conceptions, we are always able to find a third conception that is between the other two and expresses some relation in which they stand. This point was later expressed through the logic of triadic relations and the theory of interpretation. An even more striking illustration is found in the use of the limit concept to define the nature of the real as individual. The reality at which the process of knowledge is directed is said to be the "limit" of a series of attempts to apprehend the object. Royce understood "limit," not in the sense of an end term that we can approach at will, but in the sense of a least upper bound, which, in the series $1 + \frac{1}{2} + \frac{1}{4} \cdots$, for example, is the least number that lies beyond the sum of the series—namely, 2. Thus, the real, individual reality is what is immediately beyond the whole series of efforts to know it.

ETHICAL AND RELIGIOUS DOCTRINES

Royce contributed ideas worthy of consideration to almost every branch of philosophy, not least in ethics.

LOYALTY. Royce's *Philosophy of Loyalty* (New York, 1908) is still one of his best-known books. In it he developed the principle of loyalty to loyalty as the basic moral law. He regarded his principle as superior to both Kant's categorical imperative and J. S. Mill's principle of utility. Loyalty, by which is meant a freely chosen and practical devotion to a cause or goal, is the highest virtue. Royce was well aware of the existence of evil causes and of the fact that not every cause aims at the loyal spirit. Hence, he argued that loyalty in the ethical sense means devotion to causes that extend the spirit of loyalty and do not contribute to deception, dishonesty, racial and social strife, and so on. Every cause involves some loyalty, but not all causes involve loyalty to loyalty. It is only through loyalty to loyalty itself, the virtue that makes all social life possible, that the self can solve the basic problem of ethics, which is to find a good that is at once objective, in the sense that it constrains our purely individual and subjective interests, and freely chosen, so that the self can acknowledge its obligatory character. Royce followed G. W. F. Hegel in finding the good in a form of self-realization, and he followed Kant in upholding the autonomy of the will.

PHILOSOPHY OF RELIGION. Royce's interest in the philosophy of religion was a basic factor in the shaping of his philosophical position. Religious issues constitute the foundation of his thought, starting with *The Religious Aspect of Philosophy* (Boston, 1885) and continuing to his last major work, *The Problem of Christianity* (New York, 1913). Royce had a twofold aim in the philosophical treatment of religion. First, he sought to reinterpret classical religious ideas through contemporary experience and current language; second, he attempted to assess their validity by comparing them with the results of metaphysical analysis. Both aims are clearly present in *The Problem of Christianity*, in which he developed an original interpretation of the Christian religion, first, by uncovering the experiential roots of three central ideas—the church, sin, and atonement—and, second, by seeking support for these ideas in his metaphysic of interpretation and community.

Starting with the view that neither perception nor conception alone, nor any indeterminate combination of the two, is able to yield knowledge of selves, Royce went on to develop the theory of interpretation, according to which all our knowledge is mediated through signs. From this view it follows that the human self is not known (either by itself or another) intuitively as a particular datum or as a universal character but only as the goal of an infinite process of interpretation. In requiring comparison with other selves, this process necessitates a community if there is to be self-knowledge. Persons are involved in, and linked together by, a number of different

communities—political, legal, economic, moral, religious—each of which is defined by its purpose or the goal for which it exists. The religious or Beloved Community has the special purpose of redeeming man from sin (a moral burden) and from the consequences of the self-centered deeds by which he endangers the community through disloyalty. The three central ideas of Christianity (the church, sin, and atonement) are linked together. The Beloved Community is the locus of the love (in Royce's terms, loyalty) exemplified by the atoning deed of Jesus; the church exists to overcome, through love, the self-centeredness of the individual and to transmute the evil consequences of treachery by a constant renewal of the community of many selves devoted to the cause of charity.

The novel feature of Royce's reinterpretation of Christianity is his attempt to rework the much neglected doctrine of the Spirit, or Third Person, of the ancient Trinitarian tradition. God now appears as the Spirit or Interpreter, linking together a multiplicity of distinct selves in a spiritual unity of love. The Beloved Community, founded by the sacrificial or atoning deed of Jesus, becomes the ultimate instrument of the redemptive process.

Unlike William James, Royce was clearly dissatisfied with a purely practical basis for religious belief. Instead, he made the validity of religion dependent on a metaphysical system. He set forth one such system in *The World and the Individual*, and he returned to the task in *The Problem of Christianity*, in which he dealt with specifically Christian ideas. In the intervening years Royce fell under the influence of Peirce's thought, and he freely acknowledged an indebtedness to Peirce's theory of signs, his analysis of triadic relations, and the idea of the community of knowers engaged in interpreting the meaning of things through an infinite system of signs.

The continuation of the logical and epistemological aspects of Royce's philosophy is to be found mainly in the work of C. I. Lewis, and its metaphysical aspects are developed in the thought of W. E. Hocking. The strong current of pragmatism on the American scene, however, carried philosophical thinking away from the speculative realm and directed it into other channels.

See also Absolute, The; American Philosophy; Being; Bradley, Francis Herbert; Dewey, John; Frege, Gottlob; Hegel, Georg Wilhelm Friedrich; Hocking, William Ernest; Idealism; James, William; Kant, Immanuel; Lewis, Clarence Irving; Lotze, Rudolf Hermann; Loyalty; Mathematics, Foundations of; Mill, John Stuart; Mysticism, Nature and Assessment of; Peano, Giuseppe; Peirce, Charles Sanders; Pragmatism; Realism; Relations, Internal and External; Russell, Bertrand Arthur William; Santayana, George; Self-Knowledge.

Bibliography

WORKS BY ROYCE

The Basic Writings. Edited by John J. McDermott. Chicago: University of Chicago Press 1969.

The Conception of God: A Philosophical Discussion concerning the Nature of the Divine Idea as a Demonstrable Reality. St. Clair Shores, MI: Scholarly Press, 1971.

The Conception of Immortality. New York: Greenwood Press, 1968.

Lectures on Modern Idealism. New Haven, CT: Yale University Press 1967.

The Letters of Josiah Royce. Edited by John Clendenning. Chicago: University of Chicago Press 1970.

The Philosophy of Loyalty. Nashville, TN: Vanderbilt University Press, 1995.

WORKS ON ROYCE

The Problem of Christianity: Lectures Delivered at the Lowell Institute in Boston, and at Manchester College, Oxford. Hamden, CT: Archon Books, 1967.

Race Questions, Provincialism, and Other American Problems. Freeport, NY: Books for Libraries Press 1967.

The Religious Aspect of Philosophy: A Critique of the Bases of Conduct and of Faith. Gloucester, MA: Peter Smith, 1965.

The Religious Philosophy of Josiah Royce. Westport, CT: Greenwood Press, 1976.

The Sources of Religious Insight. New York: Octagon Books, 1977.

The Spirit of Modern Philosophy: An Essay in the Form of Lectures. Norton, 1967.

William James and Other Essays on the Philosophy of Life. Freeport, NY: Books for Libraries Press 1969.

For a complete list of Royce's writings, published and unpublished, it is necessary to combine the bibliographies in the following works:

Cotton, J. H. *Royce on the Human Self*, 305–311. Cambridge, MA: Harvard University Press, 1954.

Loewenberg, Jacob. *Philosophical Review* 26 (1917): 578–582.

Rand, Benjamin. *Philosophical Review* 25 (1916): 515–522.

Smith, J. E. *Royce's Social Infinite*, 171–173. New York: Liberal Arts Press, 1950.

WORKS ON ROYCE

Cady, Linell Elizabeth. *The Conceptions of Love and the Self in the Thought of Soren Kierkegard and Josiah Royce: A Thesis*. PhD diss. Cambridge, MA: Harvard, 1981.

Clendenning, John. *The Life and Thought of Josiah Royce*. Madison: University of Wisconsin Press, 1985.

Costello, Harry Todd, and Grover Cleveland Smith. *Josiah Royce's Seminar, 1913–1914: As Recorded in the Notebooks of Harry T. Costello*. Westport, CT: Greenwood Press, 1981.

Cotton, James Henry. *Royce on the Human Self*. Cambridge, MA: Harvard University Press, 1954; New York: Greenwood Press, 1968. Well-focused; based on manuscript as well as

printed works; aimed at interpreting the whole of Royce's thought in terms of his theory of the self.

Cunningham, G. Watts. *The Idealistic Argument in Recent British and American Philosophy*. London and New York: Century, 1933. A clear and helpful account of Royce's thought, focusing on his attempt to establish the truth of idealism by the argument from error.

Fuss, Peter. *The Moral Philosophy of Josiah Royce*. Cambridge, MA: Harvard University Press, 1965. The most incisive account of Royce's ethical philosophy in relation to both his theory of knowledge and the doctrine of the community.

Galligan, Michael W. E. *God and Experience: A Study of Josiah Royce's "The World and the Individual."* PhD diss. New Haven, CT: Yale University, 1973.

Imbelli, Robert Peter. *Man's Quest for Salvation in the Thought of Josiah Royce*. PhD diss. New Haven, CT: Yale University, 1973.

Jarvis, Edward A. *The Conception of God in the Later Royce*. The Hague: Martinus Nijhoff, 1975.

Kegley, Jacquelyn Ann Kovacevic. *Josiah Royce's Theory of Knowledge*. New York: 1971.

Kuklick, Bruce. *Josiah Royce: An Intellectual Biography*. Indianapolis: Bobbs-Merrill, 1972.

Mahoney, James Michael. *Josiah Royce: Moral Reasoning*. PhD diss. Notre Dame, IN: University of Notre Dame, 1975.

Mahowald, Mary Briody. *An Idealistic Pragmatism: The Development of the Pragmatic Element in the Philosophy of Josiah Royce*. The Hague: Nijhoff, 1972.

Marcel, Gabriel. *La métaphysique de Royce*. Paris: Aubier, 1945. Translated by V. Ringer and G. Ringer as *Royce's Metaphysics*. Chicago: Regnery, 1956. Four articles on Royce's thought that originally appeared in 1918–1919. The entire range of Royce's philosophy is considered, with indications of Marcel's own existentialism and the resources for it that are to be found in Royce's theory of knowledge and the individual.

Marcel, Gabriel. *Royce's Metaphysics*. Westport, CT: Greenwood Press, 1975.

Muirhead, John H. *The Platonic Tradition in Anglo-Saxon Philosophy*. London: n.p., 1931. A fairly complete, but not especially penetrating, account of all of Royce's works from the earliest essays to *The Problem of Christianity*.

Oppenheim, Frank M. *Bibliography of the Published Works of Josiah Royce*, 1967.

Oppenheim, Frank M. *Royce's Mature Ethics*. Notre Dame, IN: University of Notre Dame Press, 1993.

Oppenheim, Frank M. *Royce's Mature Philosophy of Religion*. Notre Dame, IN: University of Notre Dame Press, 1987.

Powell, Thomas F. *Josiah Royce*. New York: Washington Square Press, 1967.

Skinner, John E. *The Logocentric Predicament: An Essay on the Problem of Error in the Philosophy of Josiah Royce*. Philadelphia: University of Pennsylvania Press, 1965.

Smith, John E. *Royce's Social Infinite*. New York: Liberal Arts Press, 1950. A study of Royce's later thought, stressing the idea of the community and of its relation to Peirce's theory of signs. Considerable emphasis is placed on the logic of interpretation and its use in Royce's final statement of the religious issues.

John E. Smith (1967)
Bibliography updated by Michael J. Farmer (2005)

ROYER-COLLARD, PIERRE PAUL

(1762–1845)

Pierre Paul Royer-Collard, the French statesman and professor of philosophy, was born at Sompuis, a village in what is now the department of the Marne. He represented this department in the Chamber of Deputies from 1815 to 1839, usually in the opposition. He is best known as the leader of the *Doctrinaires,* a group whose members derived their political views from what they believed to be immutable and self-evident principles. These principles led to a compromise between absolute and constitutional monarchy, and though the principles were supported by Louis XVIII, they were rejected by his brother and successor, Charles X.

Royer-Collard had little, if any, philosophical training. Nevertheless, from 1811 to 1814 he was professor of philosophy and dean at the Sorbonne. He lectured first on Thomas Reid and later on his own views. Just as his political views were a compromise, so in philosophy he sought a compromise between the left wing of sensationalism and the right wing of authoritarian traditionalism. He found it in the philosophy of Reid. Royer-Collard rejected sensationalism on the ground that it could not account for judgment, which is always something contributed to sensory material by the active mind. Since the individual mind is active and capable of making judgments, there is no need of a supernatural authority to dictate to it. In place of such an authority he substituted common sense, which is a consolidation of the judgments of all men. But this did not imply a return to tradition except insofar as tradition itself is an expression of common sense. On the contrary, every man has within him the ability to distinguish between right and wrong, truth and falsity, by a power that resembles the natural light of medieval philosophy. If this faculty did not exist, he maintained, one would be stranded in solipsism, for there would be no reason to believe that one man's conclusions would be harmonious with another's.

Common sense, however, does not operate entirely without the guidance of reason. In reaching its decisions, reason uses two principles of argument, that of causality and that of induction. The search for causes is intrinsic to thinking itself and will inevitably lead back to the idea of a First Cause. For, following Isaac Newton, Royer-Collard believed that one must never accept more causes than are necessary to explain phenomena. However, he does not

seem to have had any clear idea of the nature of a causal explanation.

The principle of induction is a necessary accompaniment to that of causality, for it is by induction that one discovers the essential similarities among phenomena that permit one to group them in a single class. It is man's nature to look for these similarities, as it is his nature to look for causes.

Following Reid, Royer-Collard maintained that the distinction between sensation and perception is all-important. Sensation is simply the pleasure found in experience and is purely subjective. Perception is the apprehension of an external object as external. The externality of the object is not proved by reasoning; it is judged by a spontaneous act of the human mind, as in the twentieth-century epistemology of G. E. Moore.

Though only fragments of Royer-Collard's philosophy exist, collected by his admirer Théodore Jouffroy, it is probable that he saw the philosophy of common sense as a support for his political views. Common sense is the basis of communal life; it provides stable theses of morality and religion; it has all the authority of natural law; and to those who accept it, it is incontrovertible. It is, however, generally admitted that the main contribution of Royer-Collard to French philosophy was the introduction into France of Scottish philosophy.

See also Common Sense; Induction; Jouffroy, Théodore Simon; Medieval Philosophy; Moore, George Edward; Newton, Isaac; Reid, Thomas; Sensationalism; Solipsism; Traditionalism.

Bibliography

WORKS BY ROYER-COLLARD

Les fragments philosophiques de Royer-Collard, edited by André Schimberg. Paris, 1913.

Discours prononcé à l'ouverture du cours de l'histoire de le philosophie. A very rare pamphlet in the Bibliothèque de l'Institut.

WORKS ON ROYER-COLLARD

Boas, George. *French Philosophies of the Romantic Period,* 157–164. Baltimore: Johns Hopkins Press, 1925.

Garnier, Adolphe. "Royer-Collard." In *Dictionnaire des sciences philosophiques,* edited by Adolphe Franck. Paris, 1875.

Spuller, Eugène. *Royer-Collard.* Paris: Hachette, 1895.

Taine, Hippolyte. *Les philosophes classiques de XIX*[e] *siècle en France.* Paris, 1857. To be read with caution.

George Boas (1967)

ROZANOV, VASILII VASIL'EVICH

(1856–1919)

Vasilii Vasil'evich Rozanov, the Russian critic and philosopher, was born in Vetluga, Russia, and attended secondary schools in Simbirsk and Novgorod before entering Moscow University as a student in the faculty of history and philology. After his graduation from the university in 1881, he taught history and geography in a succession of secondary schools in provincial towns and began the writing on religious and philosophical themes that was to gain him a reputation as a brilliant if erratic critic of contemporary culture, both secular and religious. In 1893 a minor government post in St. Petersburg brought him to the center of Russian literary life, and in 1899 he retired to devote full time to writing. He published numerous books and contributed many articles to the Russian reviews of the day, particularly the reactionary *Novoe vremia* (New times). During the Russian Revolution he took refuge with the religious philosopher Father Pavel Florenskii in Sergiev Posad, near Moscow, where he died.

Rozanov's first major writing and his only strictly philosophical work was an elaborate scholarly treatise titled *O ponimanii* (On the understanding), in which he developed a conception of understanding as a unifying mode of cognition that reconciles science and philosophy. He first won public acclaim with his critical study of Fëdor Dostoevsky, *Legenda o Velikom Inkvizitore* (The legend of the grand inquisitor). In a number of impressionistic, aphoristic works written from 1911 to 1918 he developed most fully the critique of Christianity and the "metaphysics of sex" for which he is best remembered. Chief among these later works are *Opavshie list'ia* (*Fallen leaves*), *Uedinennoe* (*Solitaria*), and *Apokalipsis nashego vremeni* (The apocalypse of our time).

Rozanov's mature worldview was a mystical theism based on the sanctification of sex. Emphasizing the generative power of sexuality, Rozanov saw in it the aspect of man that relates him most intimately to God. Sexuality is man's "noumenal aspect," of which his other qualities and capacities are manifestations. Rozanov vigorously attacked Christianity for its denial of the flesh in preaching celibacy and fasting and for its failure to recognize the holiness of elementary animal processes. He preferred the religion of the Old Testament because of what he regarded as its greater acceptance of life and greater humanitarianism, and he called for renewed worship of the vital biological forces enfeebled by Christianity.

See also Dostoevsky, Fyodor Mikhailovich; Florenskii, Pavel Aleksandrovich; Mysticism, Nature and Assessment of; Philosophy of Sex; Russian Philosophy.

Bibliography

WORKS BY ROZANOV

O ponimanii. Moscow: Tip. E. Lissiera i IU. Romana, 1886.

Legenda o Velikom Inkvizitore. St. Petersburg, 1894.

Religiia i kul'tura (Religion and Culture). St. Petersburg, 1899.

Priroda i istoriia (Nature and History). St. Petersburg, 1900.

Temnyi lik: metafizika khristianstva. St. Petersburg, 1911.

Uedinennoe. St. Petersburg, 1912. Translated by S. S. Koteliansky as *Solitaria.* London: Wishart, 1927. The English edition contains an abridged account of Rozanov's life by E. Gollerbach, other biographical material, and matter from "The Apocalypse of Our Time."

Opavshie list'ia. St. Petersburg, 1913. Translated by S. S. Koteliansky as *Fallen Leaves.* London: Mandrake Press, 1929.

Izbrannoe (Selections). Edited by George Ivask. New York, 1956.

WORKS ON ROZANOV

Edie, James M., James P. Scanlan, and Mary-Barbara Zeldin, eds. *Russian Philosophy,* 3 vols. Chicago: Quadrangle, 1965.

Poggioli, Renato. *Rozanov.* New York: Hillary House, 1962.

Zenkovsky, V. V. *Istoriia russkoi filosofii,* 2 vols. Paris, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy,* 2 vols. New York: Columbia University Press, 1953.

James P. Scanlan (1967)

ROZANOV, VASILII VASIL'EVICH [ADDENDUM]

Since the late Soviet period and especially since the collapse of the USSR, Vasilii Rozanov has been one of the most popular and influential thinkers of the religious-philosophical movement of the early twentieth century. His playful and artful texts, at once pungent and profoundly self-conscious, have provided an example of Russian postmodernism *avant là lettre*. He has had a noticeable impact on recent Russian literature (most notably Andrei Siniavskii [aka Abram Terts], Venedikt Erofeev, and Viktor Erofeev) and philosophy (Vladimir Bibikhin). A collection of Rozanov's works, numbering eighteen volumes by the end of 2004, includes many of his published books, scores of uncollected essays, and a wealth of previously unpublished material, including several volumes in the genre of "fallen leaves" (*Sakharna*, 1913; *Mimoletnoe* [Transitory things], 1914 and 1915; *Poslednie list'ia* [Final leaves], 1916 and 1917) and the greater part of texts written for *Apokalipsis nashego vremeni* (The apocalypse of our time), left unpublished at his death.

Rozanov's flouting of conventional philosophical methods and genres has generated challenging insights in the face of rapidly changing circumstances and intellectual currents. In particular, his phenomenological observations and reflective self-analysis provide an incisive commentary on the interaction between the self and the modern world in such areas as nature, women's fashion, sex, Christianity, and Judaism. In his essay "O sladchaishem Iisuse i gor'kikh plodakh mira" (On sweet Jesus and the bitter fruits of the world, 1907) Rozanov scandalized Russian orthodoxy by condemning its denial of worldly values. In *Ital'ianskie vpechatleniia* (Italian impressions, 1909) Rozanov used the persona of an estranged and bemused Russian tourist to analyze the legacy of roman antiquity and the place of religion in modern Western civilization. Even his first book *O ponimanii* (On understanding), a spectacular flop upon publication in 1886, has been revisited for its more systematic exposition of ideas that Rozanov later developed in his prolific journalistic work.

There has been considerable interest in Rozanov's Aristotelian concepts of potentiality and teleology, and his quest for a holistic form of knowledge more closely attuned to living reality. For Rozanov, cognition cannot be strictly separated from the reality it constructs; long before existentialism, Rozanov declared existence to be prior to essence. He maintains that the philosopher's task is to describe the interaction between the active percipient and his object in all of its complex existence; only in this way can one access and activate essences. Of especial interest is Rozanov's insistence that a particular kind of sympathetic attention is needed to unleash the potential of the inert forms of life. These themes—reality as dependent upon human interaction and history as the process of human understanding—associate Rozanov with the hermeneutic tradition in European philosophy and inspired his turn to an original, narrative style of philosophy in his books of the 1910s. The significance of Rozanov's thought for philosophical aesthetics and the philosophy of religion has only recently begun to be explored.

Rozanov's writings on sexuality (including homosexuality) and Judaism have become only more controversial with time. While Rozanov verged on outright anti-semitism in his political commentary, he also professed profound admiration for Jewish traditions of kinship, coupled (especially in his final writings) with increased hostility toward Christian asceticism. His veneration of

fertility caused him to oppose homosexuality, although his frank discussion of the matter rankled with traditionalist allies. Some attempt has been made to cast these writings as "carnivalistic" and "dialogic" (both terms stem from the discourse philosophy of Mikhail Bakhtin): Rozanov is seen as exploring various voices and inverting social conventions in order to overcome the manifest conflicts in society at the level of discourse. Continuing this theme of dialogism, Rozanov's correspondence (which he often cited in his essays) and his broader intellectual exchange with such leading figures as Konstantin Leont'ev and Pavel Florenskii shed much light on the broader philosophical discourse of his day. As new texts come to light and his entire project is more fully understood, interest in Rozanov's work will likely continue to grow.

See also Aesthetics, History of; Aristotelianism; Bakhtin, Mikhail Mikhailovich; Existentialism; Florenskii, Pavel Aleksandrovich; Hermeneutics; Leont'ev, Konstantin Nikolaevich; Possibility; Philosophy of Religion; Teleology.

Bibliography

WORKS BY ROZANOV IN RUSSIAN

Apokalipsis nashego vremeni. Moscow: Respublika, 2004.

Kogda nachal'stvo ushlo, edited by A. N. Nikoliukin. Moscow: Respublika, 1997.

Literaturnye izgnanniki: Vospominaniia. Pis'ma, edited by G. R. Evgrafov. Moscow: Agraf, 2000.

Mimoletnoe, edited by A. N. Nikoliukin. Moscow: Respublika, 1994.

O ponimanii. Opyt issledovaniia prirody, granits i vnutrennego stroeniia nauki kak tsel'nogo znaniia. Moscow: Tipografiia E. Lissnera i Iu. Romana, 1886. Repr. Moscow: Tanais.

Poslednie list'ia, edited by A. N. Nikoliukin. Moscow: Respublika, 2000.

Religiia i kul'tura, edited by E. V. Barabanov. 2 vols. Moscow: Pravda, 1990.

Sakharna, edited by A. N. Nikoliukin. Moscow: Respublika, 1998.

Sredi khudozhnikov, edited by A. N. Nikoliukin. Moscow: Respublika, 1994.

V temnykh religioznykh luchakh, edited by A. N. Nikoliukin. Moscow: Respublika, 1994.

WORKS BY ROZANOV TRANSLATED INTO ENGLISH

Dostoevsky and the Legend of the Grand Inquisitor. Translated by Spencer E. Roberts. Ithaca: Cornell University Press, 1972.

Four Faces of Rozanov: Christianity, Sex, Jews, and the Russian Revolution. Translated by Spencer E. Roberts. New York: Philosophical Library, 1978.

Selections from Fallen Leaves. Translated by Boris Jakim and Laury Magnus. Chicago: The Variable Press, 2003.

WORKS ABOUT ROZANOV

Clowes, Edith W. *Fiction's Overcoat: Russian Literary Culture and the Question of Philosophy*. Ithaca and London: Cornell University Press, 2004.

Crone, Anna Lisa. *Rozanov and the End of Literature: Polyphony and the Dissolution of Genre in* Solitaria *and* Fallen Leaves. Würzburg: jal-Verlag, 1978.

Fateev, V. A., ed. *V. V. Rozanov: Pro et contra. Antologiia*. 2 vols. St. Petersburg: Izdatel'stvo Russkogo Khristianskogo gumanitarnogo institute, 1995.

Hutchings, Stephen C. "Breaking the Circle of the Self: Domestication, Alienation, and the Question of Discourse Type in Rozanov's Late Writings." *Slavic Review*52 (1993): 67–86.

Lebedev, Iu. V., ed. *Chteniia, posviashchënnye 80-letiiu pamiati V. V. Rozanova*. Kostroma: Kostromskoi filial Rossiiskogo fonda kul'tury, 1999.

Nikoliukin, A. N. *Golgofa Vasiliia Rozanova*. Moscow: Russkii put', 1998.

Rudnev, Pavel. *Teatral'nye vzgliady Vasiliia Rozanova*. Moscow: Agraf, 2003.

Siniavskii, A. D. *Ocherki russkoi kul'tury. 1. Opavshie list'ia Vasiliia Vasilievicha Rozanova*. Paris: Sintaksis, 1982.

Robert Bird (2005)

RÜDIGER, ANDREAS

(1673–1731)

Andreas Rüdiger, the German physician and philosopher, was born in Rochlitz, Saxony. Poverty and bad health allowed him to study only irregularly. In 1692 he served as a tutor in the home of Christian Thomasius. He was compelled to interrupt his studies completely in 1695; not until 1697 could he enter the University of Leipzig, where he studied law and medicine, receiving a master's degree in 1700. He received a doctorate in medicine from the University of Halle in 1703, but he continued to lecture at the University of Leipzig. From 1707 to 1712 he practiced medicine and lectured in Halle, and from 1712 until his death he did so in Leipzig.

The development of Rüdiger's philosophy was greatly influenced by his teachers Christian Thomasius and Franz Budde. However, he soon developed individual views within the Thomasian school. His medical studies centered his interests on natural philosophy and gave his thought a practical bent. Like Budde's, Rüdiger's mind was more systematic than Thomasius's.

Rüdiger's most important work, *Philosophia Synthetica* (1706–1707), is divided into three sections: "Wisdom," "Justice," and "Prudence." The section on wisdom embraces logic and natural philosophy, that on justice

covers metaphysics and natural law, and that on prudence covers ethics and politics.

Rüdiger's logic had a clear psychological orientation. He was mainly interested in the origin and development of our ideas, which, he held, come into our minds through the senses, although there are some innate mental elements, too. He criticized René Descartes, discussed Pierre Gassendi, and drew some inspiration from John Locke. Rüdiger stressed the passive element of the mind; reflection, or *sensio interna,* is (contrary to Locke) a passive fact. The standard of truth lies in man's consciousness, in a *recta ratio,* which is not common sense but something that can be acquired only through instruction in logic (*lumen acquisitum*). Logic was therefore more important for Rüdiger than for the other members of the Thomasian school. He developed a refined syllogistic theory, formalizing his acceptance of the mathematical method in philosophy. However, he conceived the mathematical method quite differently from Christian Wolff, as a method for deducing facts from given facts rather than as the drawing of possible conclusions from abstract principles. Rüdiger's philosophy, like that of the Thomasian school generally, was based in large part on the notion of reality and appealed mainly to the senses and to experience, both interior and exterior. He defined "truth" in connection with the possibility of perceiving and "existence" in connection with being perceived—again in the tradition of Thomasian subjectivism.

In natural philosophy, Rüdiger tried to combine the Thomasian and Pietistic animistic or spiritualistic physics with mechanism, but the spiritualistic element predominates. He held that we have no certain knowledge of nature, and generally he refrained from choosing between different hypotheses, for instance, between the Copernican and the biblical astronomical theories.

The practical bent of Rüdiger's philosophy explains why he discussed metaphysics under the heading of justice. His metaphysical discussions were largely devoted to theology and to man's duties toward God; his discussions of natural law were devoted to our duties toward other men. Metaphysics is the science of reality, and in particular of the *ens realissimum,* rather than the science of possibility. However, according to Rüdiger, we cannot penetrate the essence of things in metaphysics; we can only establish, by means of experience, that things exist and how they exist.

Rüdiger's section on prudence constitutes, in the Thomasian tradition, a kind of anthropology, both private and public. Ethics provides precepts for reaching happiness on Earth, and politics provides precepts for governing a commonwealth.

Through his pupil A. F. Hoffmann, Rüdiger exerted a strong influence on the development of the philosophy of Christian August Crusius, and through Crusius on the whole development of German philosophy.

See also Crusius, Christian August; Thomasius, Christian.

Bibliography

WORKS BY RÜDIGER

Disputatio Philosophica de Eo, Quod Omnes Ideae Oriantur a Sensione. Leipzig, 1704.

Philosophia Synthetica. Tribus Libris de Sapientia, Justitia et Prudentia. Methodo Mathematicae Aemula Breviter et Succinte in Usam Auditorum Comprehense. Leipzig, 1707. Reprinted at Halle, 1711, and at Frankfurt, 1717, as *Institutiones Eruditionis* and at Leipzig, 1723, as *Philosophia Pragmatica. Methodo Apodictica et Quoad Eius Licuit Mathematica Conscripta.*

De Sensu Veri et Falsi. Halle, 1709.

Physica Divina Recta Ira ad Ultramque Hominis Felicitatem Naturalem Atque Moralem Ducens. Frankfurt, 1716.

Klugheit zu Leben und zu Herrschen. Leipzig, 1722.

Christian Wolffens Meinung von dem Wesen der Seele und eines Geistes überhaupt, und Andreas Rüdigers Gegen-Meinung. Leipzig, 1727. A pamphlet attacking Wolff.

WORKS ON RÜDIGER

Carls, W. *Rüdigers Moralphilosophie.* Halle, 1894.

Schepers, Heinrich. *A. Rüdigers Methodologie und ihre Voraussetzungen.* Cologne, 1959.

Wundt, Max. *Die deutsche Schulphilosophie im Zeitalter der Aufklärung,* 82–97. Tübingen: Mohr, 1945.

Giorgio Tonelli (1967)

RUFUS, RICHARD
(?–after 1259)

Richard Rufus, a thirteenth-century philosopher and theologian, was among the first European medieval authors to study Aristotelian metaphysics, epistemology, and natural philosophy. His lectures on the so-called *libri naturales* date from a period shortly after the effective lapse of the ban on teaching them in 1231 and are among the earliest European commentaries on those works. In 1238, after writing treatises against Averroes and lecturing on Aristotle—at greatest length on the *Metaphysics*—he joined the Franciscan Order, left Paris, and became a theologian.

Rufus's lectures on Peter Lombard's *Sentences* were the first presented by an Oxford bachelor of theology. Greatly influenced by Robert Grosseteste, Rufus's Oxford lectures were devoted in part to a refutation of Richard Fishacre, the Dominican master who first lectured on the *Sentences* at Oxford.

Rufus's Oxford lectures were employed as a source by St. Bonaventure, whose lectures on the *Sentences* were vastly influential. Returning to Paris shortly after Bonaventure lectured there, Rufus took Bonaventure's lectures as a model for his own Parisian *Sentences* commentary. Rufus's Paris lectures made him famous. According to his enemy Roger Bacon, when he returned to Oxford after 1256 as the Franciscan regent master, his influence increased steadily. It was at its height forty years later in the 1290s, when John Duns Scotus was a bachelor of theology. Early versions of many important positions developed by Duns Scotus can be found in Rufus's works.

NATURAL PHILOSOPHY

Among the first medieval European philosophers to encounter Aristotle's arguments for the eternity of the world, Rufus also presented some of the most cogent counterarguments. One argument is based on a contradiction between the definitions of *past* and *infinity*. It is impossible to traverse an infinity, but it pertains to the nature of the past to have been traversed; therefore, past time cannot be infinite. In 1235 Rufus presents the argument with characteristic brevity: "Having been traversed" is incompatible with the definition of infinity, but "having been traversed" belongs to the definition of the past. Therefore, being past is incompatible with the definition of infinity.

This argument, first presented in late antiquity by John Philoponus, is now associated with Immanuel Kant; in medieval philosophy it is ordinarily ascribed to Bonaventure, who advanced it in 1250 or 1251. It occurs in different versions, some more persuasive than others. Grosseteste, for example, mistakenly seeks to apply it to the future as well as to the past, claiming that the argument can be used to show that time could not be infinite *a parte post*. Rufus sees even more clearly than Philoponus that the direction of time is an important part of this argument. He notices that the argument must be based on the fact that the whole of past time has been traversed, rather than on the claim that the whole of the past and the future will have been traversed. In his later work he seeks to force his opponents to see that they are committed to the claim that some past days are not now and never were present. By contrast, Philoponus sees this as an argument about the impossibility of completely counting an infinite series, with no particular focus on the direction of time.

Rufus's version of another of Philoponus's arguments is based on the concept of priority. If the number of days before today is infinite, and the number of days before tomorrow is infinite, then the number of days before today is not less than the number of days before tomorrow. Consequently, today does not arrive sooner than tomorrow, which is absurd. Rufus assumes here that unequal infinities are impossible. Following Georg Cantor, modern mathematicians reject this assumption. However, Rufus needs only the uncontroversial claim that mappable infinities are equal: If one postulates beginningless time, the number of days before today and the number of days before tomorrow are mappable infinite series. Rufus might still argue that if the world has no beginning, then one must give up the belief that less time elapses before earlier events than before later events.

Philoponus's original version of this argument is not based on the claim that more time transpires before later events than before earlier events. The absurdities he asks one to reject are mathematical: that it is possible to add to an infinity, or that one infinity can be multiplied by another, so that one infinity would be greater than another by a determinate proportion. By contrast, the absurd conclusion Rufus asks one to reject is that "today does not come sooner than tomorrow"; he emphasizes the unique properties of time.

THEORY OF KNOWLEDGE

The fullest statement of Rufus's epistemological views now known is a treatise titled *Speculum animae* (A mirror of the soul), probably written to explain problems in Aristotelian philosophy to his Franciscan confreres. This treatise addresses the question: What does Aristotle mean when he says that "in some manner the soul is every thing"? In the *Speculum* Rufus develops and changes his views; he rejects the view his predecessors based on patristic authorities: The soul is everything because it shares being with rocks, life with animals, and understanding with angels—a view Rufus states without comment in the last lectures he gave before becoming a Franciscan, when expounding *Metaphysics* Lambda. Since Rufus also rejects a literal interpretation of the dictum, he must explain in what sense the soul becomes an object when it understands or senses that object.

Rufus has to face two related questions: Why does the soul not become green when it perceives something green? If reception of species produces apprehension in the soul, why does the presence of such species not have

this effect in other subjects? One element of Rufus's reply is constant. He postulates that sensible and intelligible species are nonnatural and different in their mode of being from external objects. Such species are not described by Aristotelian categories; they are neither substances nor accidents. Accordingly, their reception does not produce the object sensed but the sensation or cognition of the relevant object; the sensitive soul does not become green; it senses the color green. In his *De anima* commentary Rufus describes the direct objects of sensation as spiritual beings, and he holds that plants do not sense colors since spiritual beings do not act on them. Spiritual being, a concept Rufus owes to Averroes, is the key to Rufus's exposition of a phrase in Aristotle, who says that the senses are susceptible of sensible species "without matter" (2.12.424a18–19). In his *Contra Averroem* Rufus confronts the objection that accidents that are not spiritual also act without matter—in producing heat, for example. He replies by claiming that Aristotle was contrasting species or intentions with "materiated species" designed to perfect matter rather than to produce cognition. In his last *Metaphysics* commentary Rufus contrasts spiritual with material reception, repeating terminology from his *In De anima*, but omitting the *De anima* commentary's reference to spiritual being.

In his *Contra Averroem* Rufus also makes it harder to answer the second question in another respect. Lecturing on *De anima*, Rufus claims that wood, for example, apprehends nothing because its matter receives only the natural form of wood, not its species (a similitude of the whole). In *Contra Averroem* he cites passages that convince him that the objects one senses are not mere similitudes of sensed objects but really the same as them. Rufus's response to the problem this presents in the *Speculum animae* (and subsequently in his Oxford theology lectures) is to argue that what is really identical may be formally distinct. Since species exist nonnaturally, they can be really the same as, but not formally identical with, or predicable of, the objects of apprehension. This safeguards the claim that what one apprehends is really the same as external objects; in some sense the soul really is all things.

METAPHYSICS

Postulating a kind of identity that permits real but not formal predication is a conceptual tool that Rufus employs when discussing a variety of philosophical topics—for example, the problem of individuation. Like Duns Scotus in his *Metaphysics* commentary, Rufus postulates individual forms to explain individuation. Individual forms are really, but not formally, the same as specific forms. Specific forms are principles of shared identity; they pertain to common natures capable of instantiation (*multiplicabilis*). By contrast, individual forms pertain to the same natures as they are actually instantiated (*actu multiplicata*).

Rufus's arguments against alternative theories were initially more influential than his own views. He holds that the cause of individuation cannot be an accident or an aggregation of accidents, since individual primary substances are ontologically prior to accidents. Though he allows a role for matter as an occasional cause of individuation, he argues that even determinate matter could not by itself be the principle of individuation. Being an individual means being distinct and united, both of which are functions of form, the active principle of substance, not matter, the passive principle.

Holding that individual forms added to an aggregate of matter and specific form must be the principle of individuation, Rufus denies that the ultimate constituents of individuals are knowable. He is not sure whether what is added to the common nature can be located within an Aristotelian category. He suggests that, strictly speaking, the cause of individuation may be neither a substance nor an accident. Identifying individual forms as perfections of the specific form, he suggests that they may be substantial without being substances. Specific and individual forms provide different degrees of unity: Specific unity is less than individual unity and greater than generic unity.

Like Rufus's views on individuation, his argument for the existence of God was accepted and modified by Duns Scotus. Rufus rejected St. Anselm's famous ontological argument as sophistical (though subtle). In its place he advanced a modal argument based on the concept of God as an independent being (*a se et non ab alio*). The existence of independent beings is either necessary or impossible. Therefore, if an independent being can exist, it does exist. Rufus employs logically sophisticated arguments to show that an independent being can exist.

At the opposite end of the cosmological scale, Rufus also makes an original contribution to the problem of elemental composition—positing elemental forms incompletely actualized in the compounds they comprise. Rufus sets out to explain how elements can retain their identity in a compound, so that we can correctly say that a compound is composed of elements, and yet also explain the unity of the compound with a distinct identity. Rufus has to describe how flesh and bone can be composed of elements that can be separated out when the compound breaks down, without immediately being dis-

solved by their component elements' actions on one another—the hot heating the cold, for example. If there is to be a compound at all, the elements cannot exist in the compound in quite the same way that they do when separated. Supposing that the elements are substances, Rufus argues contrary to Averroes that elemental forms are in no sense accidents, either in the compound or outside it, though they are subject to intension and remission. His solution to the difficulty is to postulate that the elemental forms can be more or less actual; they exist in compounds in accidental or proximate potential, prevented from complete actuality by the presence of the contrary elements. The resulting mixture Rufus describes as having the unity of fusion, intermediate between absolute unity and unity of aggregation.

INFLUENCE

Rufus's importance has long gone unrecognized, in part because he preferred not to take credit for his own work and in part because, unlike his contemporaries, he provided long quotations of the positions he treated seriously. Since his own views were often stated briefly, historians who overlooked his critical bent saw him as a derivative figure. Now that Bonaventure's borrowing from Rufus has been discovered, and scholars are beginning to appreciate the significance of citations by Grosseteste (to Magister Richardus), Duns Scotus (to *Doctor antiquus*), and Franciscus de Marchia (to Richardus), the question of Rufus's influence will have to be reconsidered.

See also Anselm, St.; Bacon, Roger; Bonaventure, St.; Duns Scotus, John; Eternity; Grosseteste, Robert; Philoponus, John.

Bibliography

WORKS BY RUFUS

In Physicam Aristotelis, edited by Rega Wood. New York: Oxford University Press, 2003.

[Pseudo-Peter of Spain]. *Expositio libri De anima*. In Pedro Hispano, *Obras filosóficas*. Vol. 3, edited by M. Alonso. Instituto de Filosofía "Luis Vives," Serie A, no. 4. Madrid: n.p., 1952.

WORKS ABOUT RUFUS

Brams, J. "Le premier commentaire médiéval sur le 'Traité de l'âme' de Aristote?" *Recherches de théologie et philosophie médiévales* 68 (2001): 213–227.

Brown, S. F. "The Eternity of the World Discussion at Early Oxford." In *Mensch und Natur im Mittelalter*. Vol. 1, edited by Albert Zimmermann and Andreas Speer, 1: 259–280. Berlin and New York: De Gruyter, 1991.

Dales, R. C. "The Influence of Grosseteste's Hexaëmeron on the Sentences Commentaries of Richard Fishacre O.P. and Richard Rufus of Cornwall O.F.M." *Viator* 2 (1971): 271–300.

Gàl, G. "Commentarius in 'Metaphysicam' Aristotelis cod. Vat. lat. 4538 fons doctrinae Richardi Rufi." *Archivum franciscanum historicum* 43 (1950): 209–242.

Gàl, G. "Opiniones Richardi Rufi Cornubiensis a censore reprobatae." *Franciscan Studies* 35 (1975): 136–193.

Gàl, G. "Viae ad existentiam Dei probandam in doctrina Richardi Rufi OFM." *Franziskanische Studien* 38 (1956): 177–202.

Karger, E. "Richard Rufus on Naming Substances." *Medieval Philosophy and Theology* 7 (1998): 51–67.

Karger, E. "Richard Rufus's Account of Substantial Transmutation." *Medioevo* 27 (2002): 165–189.

Libera, A. de. "La littérature des abstractiones et la tradition logique d'Oxford." In *The Rise of British Logic: Acts of the Sixth European Symposium on Medieval Logic and Semantics, Balliol College, Oxford, 19–24 June 1983*, ed. P. Osmund Lewry, 63–114. Papers in Mediaeval Studies, no. 7. Toronto, Canada: Pontifical Institute of Mediaeval Studies, 1985.

Little, A. G. "The Franciscan School at Oxford in the Thirteenth Century." *Archivum franciscanum historicum* 19 (1926): 803–874.

Long, R. J. "The First Oxford Debate on the Eternity of the World." *Recherches de philosophie et théologie médiévales* 65 (1998): 54–98.

Long, R. J. "Of Angels and Pinheads: The Contributions of the Early Oxford Masters to the Doctrine of Spiritual Matter." *Franciscan Studies* 56 (1998): 239–254.

Noone, T. B. "An Edition and Study of the Scriptum super Metaphysicam, Bk. 12, Dist. 2: A Work Attributed to Richard Rufus of Cornwall." PhD diss., University of Toronto, 1987.

Noone, T. B. "Richard Rufus of Cornwall and the Authorship of the Scriptum super Metaphysicam." *Franciscan Studies* 49 (1989): 55–91.

Noone, T. B. "Richard Rufus on Creation, Divine Immutability, and Future Contingency in the Scriptum super Metaphysicam." *Documenti e studi sulla tradizione filosofica medievale* 4 (1993): 1–23.

Noone, T. B. "Roger Bacon and Richard Rufus on Aristotle's Metaphysics: A Search for the Grounds of Disagreement." *Vivarium* 35 (1997): 251–265.

Noone, T. B., and R. J. Long. "Fishacre and Rufus on the Metaphysics of Light: Two Unedited Texts." In *Roma, magistra mundi: Itineraria culturae medievalis: Mélanges offerts au père L. E. Boyle à l'occasion de son 75e anniversaire*, 3 Vols., edited by Jacqueline Hamesse, 2:517–548. Textes et études du moyen âge, no. 10. Louvain-la-Neuve, Belgium: Fédération des Instituts d'Etudes Médiévales, 1998.

Pelster, F. "Der älteste Sentenzenkommentar aus der Oxforder Franziskanerschule: Ein Beitrag zur Geschichte des theologischen Lehrbetriebs an der Oxforder Universität." *Scholastik* 1 (1926): 50–80.

Pelster, F. "Richardus Rufus Anglicus O.F.M. (c. 1250), ein Vorläufer des Duns Scotus in der Lehre von der Wirkung der priesterlichen Lossprechung." *Scholastik* 25 (1950): 549–552.

Plevano, R. "Richard Rufus of Cornwall and Geoffrey of Aspall: Two Questions on the Instant of Change." *Medioevo* 19 (1993): 167–232.

Raedts, P. *Richard Rufus of Cornwall and the Tradition of Oxford Theology*. Oxford Historical Monographs. Oxford, 1987.

Weisberg, M., and R. Wood. "Richard Rufus's Theory of Mixture: A Medieval Explanation of Chemical Combination." In *Chemical Explanation: Characteristics, Development, Autonomy*, edited by J. E. Earley, Sr., 282–292. Annals of the New York Academy of Sciences, vol. 988. New York, 2003.

Wood, R. "Distinct Ideas and Perfect Solicitude: Alexander of Hales, Richard Rufus, and Odo Rigaldus." *Franciscan Studies* 53 (1993): 7–46.

Wood, R. "The Earliest Known Surviving Western Medieval Metaphysics Commentary." *Medieval Philosophy and Theology* 7 (1998): 39–49.

Wood, R. "Early Oxford Theology." In *Mediaeval Commentaries on the Sentences of Peter Lombard: Current Research*, edited by G. R. Evans, 1: 289–343. Leiden: E.J. Brill, 2002.

Wood, R. "Individual Forms: Richard Rufus and John Duns Scotus." In *John Duns Scotus: Metaphysics and Ethics*, edited by L. Honnefelder, R. Wood, and M. Dreyer, 251–272. Leiden: E.J. Brill, 1996.

Wood, R. "Richard Rufus: Physics at Paris before 1240." *Documenti e studi sulla tradizione filosofica medievale* 5 (1994): 87–127.

Wood, R. "Richard Rufus and the Classical Tradition: A Medieval Defense of Plato." In *Néoplatonisme et philosophie médiévale: Actes du Colloque international de Corfou 6–8 octobre 1995 organisé par la Société Internationale pour l'Étude de la Philosophie Médiévale*, edited by L. G. Benakis, 229–251. Rencontres de philosophie médiévale, no. 6. Turnhout, 1997.

Wood, R. "Richard Rufus of Cornwall and Aristotle's Physics." *Franciscan Studies* 52 (1992): 247–281.

Wood, R. "Richard Rufus of Cornwall on Creation: The Reception of Aristotelian Physics in the West." *Medieval Philosophy and Theology* 2 (1992): 1–30.

Wood, R. "Richard Rufus's De anima Commentary: The Earliest Known, Surviving, Western De anima Commentary." *Medieval Philosophy and Theology* 10 (2001): 119–156.

Wood, R. "Richard Rufus' 'Speculum animae': Epistemology and the Introduction of Aristotle in the West." In *Die Bibliotheca Amploniana: Ihre Bedeutung im Spannungsfeld von Aristotelismus, Nominalismus und Humanismus*, edited by A. Speer, 86–109. Miscellanea mediaevalia, vol. 23. Berlin, 1995.

Wood, R. "Roger Bacon: Richard Rufus' Successor as a Parisian Physics Professor." *Vivarium* 35 (1997): 222–250.

Wood, R., and R. Andrews. "Causality and Demonstration: An Early Scholastic Posterior Analytics Commentary." *The Monist* 79 (1996): 325–356.

Wood, R., and M. Weisberg. "Interpreting Aristotle on Mixture: Problems about Elemental Composition from Philoponus to Cooper." *Studies in History and Philosophy of Science* 35 (2004): 681–706.

Rega Wood and Jennifer R. Ottman (2005)

RULE FOLLOWING

In 1982 Saul Kripke published *Wittgenstein on Rules and Private Language* and ushered in a new era of Ludwig Wittgenstein interpretation. Although elements of Kripke's view of Wittgenstein could be found in the preceding literature (notably in Robert Fogelin's *Wittgenstein*), nothing had captured attention like his presentation of the "rule-following considerations."

Kripke presented his essay as a reconstruction of the problems Wittgenstein is addressing between around §140 and §203 of the *Philosophical Investigations.* These issue in the form of a paradox—that there can be no such thing as the meaning of a word; no fact of the matter that entails that a word is used according to a rule, whereby some applications of it are determined to be correct and other applications incorrect. In §201 Wittgenstein wrote "This [is] our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer [is] if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here."

The paradox is developed by Kripke through the figure of a "bizarre skeptic." The defender of common sense, here the view that words do indeed have meanings and obey rules, is challenged to show what this meaning consists in. The facts he or she can adduce typically include past applications and present dispositions to apply words in new cases. They may also include flashes of consciousness—for instance, if we associate a particular image with a term. But, Kripke's skeptic argues, these are not the kinds of facts that can determine the actual rule that governs the meaning of a word. The skeptic adduces three kinds of problems. First, our dispositions are finite, whereas a rule can cover a potential infinity of new cases. Second, our dispositions sometimes fail to match the relevant rules: This is precisely what happens when we mistakenly apply words to things to which they do not in fact apply. Third, the existence of a rule has normative implications. It determines correctness and incorrectness of application of the term it governs. Our dispositions, by contrast, have no such implication. There is nothing intrinsically wrong about bending our dispositions from moment to moment, in the way that there is about applying a term in a way that fails to accord with its meaning. Finally, the addition of flashes of consciousness is unlikely to help, for, as Wittgenstein himself said, any such fact itself stands in need of interpretation. A flash of consciousness cannot comprehend all the possible appli-

cations of a term and sort them into those that are correct and those that are not.

Kripke illustrates these points with the case of a strange arithmetical operator, "quus." For two numbers *n* and *m, n* quus *m* is identical with *n* plus *m* for sufficiently small or common numbers, but the two results (or calculations) diverge when *n* and *m* are greater than a certain value (the function is therefore reminiscent of Nelson Goodman's predicate, "grue"). We do not mean *n* quus *m* when we talk of *n* plus *m.* But our dispositions with "plus" might match those of people who in fact use the term to mean quus; we might give the answer *n* quus *m* when we attempt to add *n* and *m,* since we make mistakes; and finally there is nothing right or wrong about having one disposition or another.

The conclusion is paradoxical, since nothing seems more certain than that we do succeed in attaching reasonably determinate meanings to terms. It may be true that the "open texture" of terms suggests that meanings are never fully precise, capable of determining their application in any circumstances, however outlandish. Nevertheless, over an indefinite normal range of cases, there is no doubt that some applications are correct and others not, and any interpretation of us according to which we mean something along the lines of the "quus" function is incorrect. Yet so long as the skeptic wins, we have no conception of our right to say such things. Kripke's own solution to the paradox is that the skeptic wins on his chosen ground. There is indeed no fact of the matter whether one rule rather than another governs the use of a term. But we can advance a "skeptical solution" (David Hume's phrase from a different context) to the doubts. What there is instead is a practice of regarding ourselves and others in certain lights. We dignify each other as meaning one thing or another by our terms, and this ongoing practice is all that there is.

Kripke's work generated enormous interest and a variety of responses in the literature. Some outraged students of Wittgenstein argued that it was not at all his intention to produce a paradox but to lay bare the oversimplifications, or desire for a simple theory, that trap people into finding rule following problematic (Baker and Hacker 1984). Many writers queried whether Wittgenstein could consistently have been content with a "non-truth-conditional" account of rule following, which is what Kripke offers him, since Wittgenstein's abhorrence of theory and his belief that philosophy leaves everything as it is would make it impossible for him to say that it is not strictly speaking true that the application of words is correct or incorrect. Some (McDowell 1981) detected a mischievous dislike of soft, humanly oriented facts in the setting up of the paradox and argued that a proper appreciation of the human constitution of rule following had wide implications for the notion of objectivity, as it occurs in domains such as aesthetics or ethics. Some (McGinn 1984) found that Kripke had not looked hard enough for natural facts with which to identify the obtaining of a rule; others (Blackburn 1985) embraced the thought that since the loss of a normative element in meaning was the main problem underlying the paradox, and since naturalistic theories of normativity have been proposed in many guises, a more generous sense of how to talk about facts solves the paradox. Paul Boghossian (1989) provided a summary of the state of the debate and a controversial contribution to it.

See also Goodman, Nelson; Hume, David; Kripke, Saul; Philosophy of Language; Wittgenstein, Ludwig Josef Johann.

Bibliography

Baker, G. P., and P. M. S. Hacker. *Scepticism, Rules, and Language.* Oxford: Blackwell, 1984.

Blackburn, S. "The Individual Strikes Back." *Synthese* 20 (1985).

Boghossian, P. "The Rule-Following Considerations." *Mind* 98 (1989).

McDowell, J. "Non-Cognitivism and Rule-Following." In *Wittgenstein: To Follow a Rule,* edited by S. Holtzman and C. Leich. London: Routledge and Kegan Paul, 1981.

McGinn, C. *Wittgenstein on Meaning.* Oxford: Blackwell, 1984.

Wittgenstein, L. *Philosophical Investigations.* Oxford: Blackwell, 1953.

Simon Blackburn (1996)

RULE FOLLOWING [ADDENDUM]

Paul Boghossian (1990) summarizes four components of what has become the received interpretation of Saul Kripke's book *Wittgenstein on Rules and Private Language* (1982) (*WRPL*): (i) Kripke's meaning-skeptic argues that meaning is normative and from descriptive facts one cannot derive normative claims, so descriptive facts cannot explain or reduce meaning; (ii) nonreductive accounts of meaning are not vulnerable to the meaning-skeptic's argument; (iii) Kripke's skeptical solution is a kind of nonfactualism about meaning; and (iv) Kripke's arguments do not show that an isolated individual cannot follow rules. Contributions to the literature about

rule-following published after 1989 include criticisms of the meaning-skeptic's argument, investigations of the nature and coherence of nonfactualism about meaning, revisionary interpretations of *WRPL*, and interpretations of Wittgenstein that challenge Kripke's and others' understanding of rule-following.

Ruth Garrett Millikan (1990) argues that the normativity of meaning can be explained in terms of biological purposes shaped and sustained by natural selection. Philip Pettit (1990) argues that the normativity of meaning is explained by one's dispositions to question or revise a judgment when one has reason to think it was the result of interfering factors, such as poor lighting or intoxication. According to Paul Horwich (1995), there is a sense of "determine" in which facts about the use of a linguistic expression *E* can determine the meaning of *E*, even if the meaning of *E* cannot be "read-off" from those facts, as Kripke's meaning-skeptic argues.

Boghossian (1990) argues that nonfactionalism about meaning is incoherent because it simultaneously *presupposes* and *conflicts with* deflationism about truth. Robert Kraut (1993) responds to this argument by describing a kind of deflationism about truth that seems compatible with nonfactualism about meaning. Crispin Wright (1992) and Scott Soames (1999) also identify what they regard as errors in Boghossian's argument.

George Wilson (1992) criticizes (iii), citing passages of *WRPL* that suggest that ascriptions of meaning can be factual even if they are not determined by independently specifiable facts. Against (i), José Zalabardo (1997) argues that the heart of the meaning-skeptic's argument is not that from descriptive facts one cannot derive normative claims, but that no facts of which we can be immediately aware can determine that our application of a word to a new case is justified. Gary Ebbs (1997) criticizes both (i) and (ii) by reconstructing Kripke's skeptical reasoning in a way that does not presuppose reductionism about meaning. Donald Davidson (1992) qualifies (iv) by supplementing Kripke's account of meaning-ascriptions with the premise that meaning requires triangulation.

Donna Summerfield (1990) reads Wittgenstein as presenting a positive account of representation that does not apply to an individual in isolation. Edward Minar (1991) presents an alternative to Kripke's skeptical solution and to the communitarian view that rule-following is constituted by linguistic interactions between speakers. David Bloor (1997) defends a communitarian interpretation of Wittgenstein. David Finkelstein (2000) suggests that to understand Wittgenstein one must come to see that there is not always a gulf between a rule and its application.

See also Davidson, Donald; Kripke, Saul; Meaning; Millikan, Ruth; Private Language Problem; Wittgenstein, Ludwig Josef Johann.

Bibliography

Bloor, D. *Wittgenstein: Rules and Institutions*. London: Routledge, 1997.

Boghossian, P. "The Status of Content." *The Philosophical Review* 99 (1990).

Davidson, D. "The Second Person." *Midwest Studies in Philosophy* 17 (1992).

Ebbs, G. *Rule-Following and Realism*. Cambridge, MA.: Harvard University Press, 1997.

Finkelstein, D. "Wittgenstein on Rules and Platonism." In *The New Wittgenstein*, edited by A. Crary and R. Read. London: Routledge, 2000.

Goldfarb, W. "Kripke on Wittgenstein on Rules." *Journal of Philosophy* 83 (1985).

Horwich, P. " Meaning, Use, and Truth." *Mind* 104 (1995).

Kraut, R. "Robust Deflationism." *Philosophical Review* 102 (1993).

Kripke, S. *Wittgenstein on Rules and Private Language*. Cambridge, MA.: Harvard University Press, 1982.

McDowell, J. "Wittgenstein on Following a Rule." *Synthese* 58 (1984).

Miller, A., and C. Wright. *Rule-Following and Meaning*. Montreal: McGill-Queen's University Press, 2002.

Millikan, R. "Truth, Hoverflies, and the Kripke-Wittgenstein Paradox." *Philosophical Review* 99 (1990).

Minar, E. "Wittgenstein and the 'Contingency' of Community." *Pacific Philosophical Quarterly* 72 (1991).

Pettit, P. "The Reality of Rule-Following." *Mind* 99 (1990).

Soames, S. *Understanding Truth*. New York: Oxford University Press, 1999.

Summerfield, D. "*Philosophical Investigations* 201: A Wittgensteinian Reply to Kripke." *Journal of the History of Philosophy* 28 (1990).

Wilson, G. "Kripke on Wittgenstein on Normativity." *Midwest Studies in Philosophy* 17 (1992).

Wright, C. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press, 1992.

Zalabardo, J. "Kripke's Normativity Argument." *Canadian Journal of Philosophy* 27 (1997).

Gary Ebbs (2005)

RULE-UTILITARIANISM

See *Utilitarianism [addendum]*

RUSKIN, JOHN
(1819–1900)

John Ruskin, the English critic of art and society, was born in London, the son of a wine merchant. He began writing while at Oxford and in 1843 published, in London, the first volume of *Modern Painters,* four more volumes of which were published during the next sixteen years. In 1849 he published *The Seven Lamps of Architecture* and between 1851 and 1853 *The Stones of Venice* (3 vols.). The major part of his work as a young man was criticism of art and architecture, and his subsequent ethical and social writing grew from this root. The beginnings of this important extension of his range can be seen in the famous chapter "The Nature of Gothic" in *The Stones of Venice*; the important connection established there, between art and "the right kind of labour," is developed in *The Political Economy of Art* (printed as *A Joy for Ever,* 1857), *Unto This Last* (1862), *The Crown of Wild Olive* (1866), and *Munera Pulveris* (1863 and 1872). Meanwhile Ruskin continued his criticism of art and architecture, notably in *The Two Paths* (1859) and in his lectures as Slade professor of art at Oxford, between 1870 and 1879 and in 1883/1884. A volume of essays, *Sesame and Lilies,* appeared in 1865 and an unfinished autobiography, *Praeterita,* between 1885 and 1889. He also published letters on social questions, notably in *Time and Tide* (1867) and *Fors Clavigera* (8 vols., 1871–1884).

Ruskin's social and ethical teaching, though deeply influenced by the work of Thomas Carlyle, followed from his understanding of the nature of art. The artist's function is to reveal aspects of the universal truth, which is also beauty. Any corruption of the moral nature of the artist is an inevitable corruption of this revelation, but it is impossible, finally, for an artist to be good if his society is corrupt. The art of any society is, correspondingly, "the exact exponent of its social and political virtues." Where there is a lack of "wholeness" in art (wholeness being a full and deep response to the organic life of the universe), there is a corresponding lack of "wholeness" in society; to recover the one men must recover the other. Just as the beauty of art is the expression of the essential nature of the universe—what Ruskin called "typical beauty"—so the goodness of man is the "exertion of perfect life," which, in comparable relation to the grand design of the created universe, is no more and no less than the "felicitous fulfillment of function" in all living things.

From his work on Venice, Ruskin developed a comparative historical approach to the social conditions in which the "exertion of perfect life" can be fostered or damaged. In particular, following the English romantic writers and the architectural critic A. W. Pugin, he saw nineteenth-century industrial civilization as the enemy of wholeness in its rampant individualism, its substitution of "production" for "wealth," and its basic misunderstanding of the nature of work. This kind of social criticism came in many respects to resemble the ideas of some philosophical socialists, and Ruskin's work had an important formative influence on the British labor movement, both directly and through his influence on William Morris, who united Ruskin's ideas with a direct commitment to socialism.

Ruskin's opposition to individualism as a social principle and to competition as a method of political economy was based on his idea of function, the fulfillment of each man's part in the general design of creation. This required a social order based on intrinsic human values, whereas the existing social order, based on the supposed laws of supply and demand, tended to put the economy above men—indeed, to reduce them to mere "labor"—and, by separating work from the pursuit of human perfection, to separate the work from the man, producing only an alienated and fragmented being. Wherever value is understood as "exchange value," rather than as the "intrinsic value" derived from function in the universal design, this corruption of man to a mere tool or machine is inevitable. In particular, the confusion about the nature of value leads to false definitions of both wealth and labor. Labor is degraded whenever it is anything other than the "exertion of perfect life," a creative activity comparable to that of the artist. Wealth is degraded whenever it is confused with mere production, for the meaning of wealth is human well-being, which in material terms is "the possession of useful articles *which we can use.*" Even if the existing system always produced useful articles, the kind of society that it also produced made just distribution and wise consumption impossible. Much actual production, and its widespread misuse, could more properly be called "illth" than "wealth," for if it possessed only exchange value and not intrinsic value it corrupted its makers and its users.

The most remarkable aspect of Ruskin's work, then, is the development of a philosophy of art into a moral critique of industrial capitalism. It is a very individual achievement, but it is also part of a general movement of nineteenth-century English thought and has evident connections with William Wordsworth, Percy Bysshe Shelley, Samuel Taylor Coleridge, and Carlyle, as well as with Morris and the Guild Socialists whom Ruskin so notably influenced.

See also Aesthetics, History of; Beauty; Carlyle, Thomas; Coleridge, Samuel Taylor; Shelley, Percy Bysshe; Socialism.

Bibliography

WORKS BY RUSKIN

Works. Edited by E. T. Cook and Alexander Wedderburn, 39 vols. London: Allen, 1903–1912.

WORKS ON RUSKIN

Evans, Joan. *John Ruskin.* London: Cape, 1954.

Hobson, J. A. *John Ruskin, Social Reformer.* Boston: Estes, 1898.

Hough, Graham. *The Last Romantics.* London: Duckworth, 1949.

Quennel, Peter. *John Ruskin, Portrait of a Prophet.* London, 1949.

Whitehouse, J. H. *Ruskin the Prophet.* London: Allen and Unwin, 1920.

Williams, Raymond. *Culture and Society,* Part 1, Ch. 7. New York: Columbia University Press, 1958.

Raymond Williams (1967)

RUSSELL, BERTRAND ARTHUR WILLIAM

(1872–1970)

Bertrand Arthur William Russell, the British philosopher, mathematician, and social reformer, was born in Trelleck, Wales. He was the grandson of Lord John Russell, who introduced the Reform Bill of 1832 and later twice served as prime minister under Queen Victoria. John Stuart Mill, a close friend of Russell's parents, was his godfather in an informal sense. Russell's parents died when he was a little child. Both of them had been freethinkers, and his father's will had provided that he and his brother were to have as their guardians friends of his father's who shared the latter's unorthodox opinions. As the result of litigation the will was set aside by the Court of Chancery and the two boys were placed in the care of their paternal grandparents. Lord John Russell died two years later, and it was the boys' grandmother who determined the manner of their upbringing. Russell was not sent to school but received his early education from a number of Swiss and German governesses and, finally, English tutors. He entered Cambridge University in October 1890 and studied mathematics and philosophy at Trinity College from 1890 to 1894. He was a fellow of Trinity College from 1895 to 1901 and lecturer in philosophy there from 1910 to 1916. In 1916 Russell was dismissed by Trinity College because of his pacifist activities. He was reinstated in 1919 but resigned before taking up his duties.

What is generally considered Russell's most important work in philosophy was done between 1900 and the outbreak of the first world war. From 1916 until the late 1930s Russell did not hold any academic position and supported himself by writing and public lecturing. During this period he wrote some of his most influential books on social questions, including *Marriage and Morals* (London, 1929) and his two books on education—*On Education, Especially in Early Childhood* (London, 1926) and *Education and the Social Order* (London, 1932). These views were put into practice in Russell's experimental school, the Beacon Hill School, which he started with his second wife, Dora, in 1927. Russell left the school in 1934 after he and Dora were divorced (the school itself continued until 1943). Russell returned to more concentrated work in philosophy around 1936. He moved to the United States in 1938, teaching first at the University of Chicago and then at the University of California at Los Angeles. In 1940 he accepted an invitation from the Board of Higher Education of New York City to join the department of philosophy at City College. However, he never had an opportunity to take up this appointment, having been found unfit for this position in a remarkable opinion by a judge who felt he had to protect "public health, safety and morals." From 1941 until 1943 Russell lectured at the Barnes Foundation in Philadelphia (these lectures were later expanded into *A History of Western Philosophy*). Dr. Albert Barnes, the head of this foundation, dismissed Russell in January 1943, on three days' notice. In this instance Russell successfully brought action for wrongful dismissal. In 1944 he returned to Cambridge where he had been reelected to a fellowship at Trinity College.

Russell was a candidate for Parliament on three occasions and was defeated each time: In 1907 he ran at Wimbledon as a candidate of the National Union of Women's Suffrage Societies, in 1922 and 1923 he stood as the Labour Party candidate for Chelsea. Russell was twice jailed—in 1918 for six months on a count of an allegedly libelous article in a pacifist journal and in 1961, at the age of eighty-nine, for one week, in connection with his campaign for nuclear disarmament.

In 1908 Russell was elected a fellow of the Royal Society. He became an honorary fellow of the British Academy in 1949, and in the same year he was awarded the Order of Merit. Russell twice served as president of the Aristotelian Society and was for many years president of the Rationalist Press Association. In 1950 he received the

Nobel Prize for literature. In making the award, the committee described him as "one of our time's most brilliant spokesmen of rationality and humanity, and a fearless champion of free speech and free thought in the West."

Russell had three children and was married four times. In 1931, upon the death of his brother, he became the third earl Russell.

Writing in 1935 the German historian Rudolf Metz referred to Russell as "the only British thinker of the age who enjoys world-wide repute." At that time his works could not circulate in Germany, Italy, or Russia. Now they are available in every major and a great number of minor languages (a truncated version of *A History of Western Philosophy* was allowed to circulate even in the Soviet Union). It is safe to say that not since Voltaire has there been a philosopher with such an enormous audience. Russell also shares with Voltaire a glittering and graceful prose style and a delicious sense of humor. It is perhaps Russell's humorous irreverence as much as the substance of his heretical opinions that has so deeply offended several generations of moralists and religious conservatives.

In the following section we shall briefly recount some of the highlights and formative influences in Russell's eventful life and sketch his views on political and social issues. Although these views are certainly logically independent of his more technical work as a philosopher, they deal with questions that have traditionally been discussed by philosophers, and they also help one to understand the basic motives inspiring Russell's thought.

LIFE AND SOCIAL THEORIES

Russell's childhood and adolescence were unhappy. The atmosphere in his grandmother's house was one of puritan piety and austerity, and his loneliness, he recalls, was almost unbearable. Only virtue was prized—"virtue at the expense of intellect, health, happiness, and every mundane good." At the age of five Russell reflected that if he lived to be seventy, he had endured only a fourteenth part of his life, and he felt the long-spread-out boredom ahead of him to be unendurable. In adolescence, he remarks, he was continually on the verge of suicide, from which, however, he was "restrained by the desire to know more mathematics." His grandmother had gradually moved from Scottish Presbyterianism to Unitarianism. As a child Russell was taken on alternate Sundays to the parish church and to the Presbyterian Church, while at home he was taught the tenets of Unitarianism. When he was fourteen he began to question theological doctrines and in the course of four years abandoned successively belief in free will, immortality, and God, the last as the result of reading John Stuart Mill's *Autobiography.* For some time, however, Russell had metaphysical attachments that served as substitutes for religion, and it was not until the end of the first world war that he became a militant opponent of all forms of supernaturalism.

EARLY PLATONISM AND HEGELIANISM. Under the influence of J. M. E. McTaggart and F. H. Bradley, Russell came, in his early years at Cambridge, to believe "more or less" in the Absolute and the rest of the apparatus of British Hegelianism. "There was a curious pleasure," Russell wrote in retrospect, "in making oneself believe that time and space are unreal, that matter is an illusion, and that the world really consists of nothing but mind." In a "rash moment," however, he turned "from the disciples to the Master." G. W. F. Hegel's remarks in the philosophy of mathematics he found "both ignorant and stupid," and in other ways Hegel's work appeared a "farrago of confusions." After that Russell was converted by G. E. Moore to a "watered down" version of Plato's theory of Ideas, regarding the subject matter of mathematics as eternal and unchanging entities whose exactness and perfection is not duplicated anywhere in the world of material objects. Eventually Russell abandoned this "mathematical mysticism" as "nonsense." Following Ludwig Wittgenstein he came to believe "very reluctantly" that mathematics consists of tautologies. As to the timelessness of mathematics, Russell now regarded this as resulting from nothing more than that the pure mathematician is not talking about time. Aside from this, it became emotionally difficult for him to remain attached to "a world of abstraction" in the midst of the slaughter of the Great War. "All the high-flown thoughts that I had had about the abstract world of ideas," he wrote later, "seemed to me thin and rather trivial in view of the vast suffering that surrounded me." The nonhuman world, he added, "remained as an occasional refuge, but not as a country in which to build one's permanent habitation." After his abandonment of Platonism, Russell wrote, he was not able to find religious satisfaction in any philosophical doctrine that he could accept.

PACIFISM. Russell was interested in social questions throughout his life. He was an early member of the Fabian Society and for some time in the 1890s, under the influence of Sidney and Beatrice Webb, championed imperialism and supported the Boer War. In 1901 he had a quasi-religious experience. He became "suddenly and vividly aware of the loneliness in which most people live" and felt the need to find ways of "diminishing this tragic isolation." In the course of a few minutes he changed his

mind about the Boer War, about harshness in the education of children and in the administration of the criminal law, as well as about fierceness in personal relations. This experience led him to write his famous essay "A Free Man's Worship" (1903). Although Russell became a pacifist right then, for another ten years or more he was preoccupied with work in mathematical logic and theory of knowledge. It was not until the war that he became passionately concerned about social issues. It is probable, he observed later, that "I should have remained mainly academic and abstract but for the War." The war, however, "shook him" out of many prejudices and made him reexamine a number of fundamental questions. He recalled:

> I had watched with growing anxiety the policies of all the European Great Powers in the years before 1914, and was quite unable to accept the superficial melodramatic explanations of the catastrophe which were promulgated by all the belligerent governments. The attitude of ordinary men and women during the first months amazed me, particularly the fact that they found a kind of pleasure in the excitement. (*Selected Papers of Bertrand Russell,* p. xi)

He decided that he had been quite mistaken in believing the claims of pacifists that wars were the work of devious tyrants who forced them on reluctant populations. Although he was not then familiar with the theories of psychoanalysis, Russell concluded that the majority of human beings in our culture were filled with destructive and perverse impulses and that no scheme for reform would achieve any substantial improvement in human affairs unless the psychological structure of the average person was suitably transformed.

Russell recalls that his decision to oppose the war was made particularly difficult by his passionate love of England. Nevertheless, he had no doubt as to what he had to do. "When the war came I felt as if I heard the voice of God. I knew that it was my business to protest, however futile protest might be. My whole nature was involved. As a lover of truth, the national propaganda of all the belligerent nations sickened me. As a lover of civilisation, the return to barbarism appalled me" (*Portraits from Memory,* p. 27). Russell remarks that he never believed much tangible good would come from opposition to the war, but he felt that "for the honor of human nature," those who "were not swept off their feet" should stand their ground. He patiently argued in lectures and books that the slaughter of millions of men was not justified by any of the possible gains of a defeat of the Central Powers. Russell's pacifism was not mystical. It was not then and had not been his contention at any time that the use of force is always wrong, that war can never possibly be justified. He maintained that *this* war in *these* circumstances was not worth all the pain and misery, and the lying of all the parties. Consistently with his general position, Russell favored the Allies during World War II on the ground that the defeat of the Nazis was essential if human life was to remain tolerable. The Kaiser's Germany by contrast was "only swashbuckling and a little absurd," allowing a good deal of freedom and democracy.

Prior to the war there had been strong pacifist sentiment in all the major Western countries, especially among the intellectuals and the powerful socialist and liberal parties. When war came only a tiny minority of these pacifists remained true to its convictions. Overwhelmed by their need to conform and in many cases by what Russell would have regarded as their own primitive impulses, many of them became the most violent jingoists. Russell was bitterly attacked for his pacifist activities not only, as one might have expected, by conservatives and professional patriots but also by many of his erstwhile friends. H. G. Wells, for example, publicly heaped abuse on Russell when he was already in trouble with the authorities. Russell's political philosophy, according to Wells, amounted to a "tepid voluntaryism," and he (unlike Wells) had no right to speak for British socialism. Wells even abused Russell's work as a mathematical philosopher. Russell, he wrote, is that "awe-inspiring" man who "objected to Euclid upon grounds no one could possibly understand, in books no one could possibly read" (preface to P. H. Loyson, *The Gods in the Battle,* London, 1917).

At Cambridge, Russell's teacher and friend McTaggart led a move for his ouster. Meetings addressed by Russell were broken up by violent mobs without any police interference. Eventually he was prosecuted by the government. For writing a pamphlet on the case of a conscientious objector he was fined £100. When he would not pay the fine the government sold parts of his library, including rare books on mathematics that Russell was never able to recover. In 1918 he was sentenced to six months' imprisonment for an article in the *Tribunal,* a pacifist weekly, in which he had written that "unless peace comes soon … the American garrison, which will by that time be occupying England and France, … will no doubt be capable of intimidating strikers, an occupation to which the American army is accustomed when at home." In a fierce denunciation which accompanied the sentence, the magistrate, Sir John Dickinson, referred to Russell's offense as "a very despicable one" and added that Russell "seems to

have lost all sense of decency." It should be added that as the result of the intervention of Arthur Balfour, Russell was treated with consideration while in prison—he finished there his *Introduction to Mathematical Philosophy* and began work on *The Analysis of Mind.*

Attitude toward the Soviet Union. Russell's isolation was not ended with the return of peace. This was due to his failure to support the Bolshevist regime in Russia. Like many Western socialists he at first welcomed the news of the revolution, but, wanting to see things for himself, he visited Russia in 1920 and came back totally disillusioned. Some of Russell's friends argued that any criticism of the revolution would only play into the hands of the reactionaries who wanted to reestablish the old order. After some hesitation Russell decided to publish the truth as he saw it. Russia, he later wrote, "seemed to me one vast prison in which the jailors were cruel bigots. When I found my friends applauding these men as liberators and regarding the regime that they were creating as a paradise, I wondered in a bewildered manner whether it was my friends or I that were mad."

The little book in which he recorded his views of the Soviet Union, *The Theory and Practise of Bolshevism* (1920), was remarkable for, among other things, its prescience. Long before most Westerners had heard of Joseph Stalin, Russell predicted, point by point, the reactionary features that came to characterize the Soviet system under Stalin—its militarism and nationalism, the hostility to free art and science, its puritanism, and the gradual ascendancy of bureaucrats and sycophants over the early idealists. Russell was able to reprint the book in 1947 without a single alteration. His isolation after his return from Russia was even greater than during the war. The patriots had not yet forgiven him his opposition to the war, while the majority of his former political friends denounced him for his opposition to the Soviet regime. But Russell has never played to the galleries. As on many other occasions he acted in accordance with his favorite biblical text—"Thou shalt not follow a multitude to do evil."

Education and sexual morality. Probably the most controversial of Russell's opinions are those relating to education and sexual morality. These were closely connected with his observations of the joy people took in the fighting and killing during the war. Russell wrote that he thought he saw the inward and outward defeats that led to cruelty and admiration of violence and that these defeats were, in turn, largely the outcome of what had happened to people when they were very young. A peaceful and happy world could not be achieved without drastic changes in education. In sexual matters, although not only in these, irrational prohibitions and dishonesty were exceedingly harmful. "I believe," he wrote in *Marriage and Morals,* "that nine out of ten who have had a conventional upbringing in their early years have become in some degree incapable of a decent and sane attitude towards marriage and sex generally" (p. 249). Conventional education was judged to be at fault in a great many other ways as well. Its general tendency was to cramp creative impulses and to discourage a spirit of critical inquiry. While a certain amount of discipline is necessary, very much of the coercion traditionally employed cannot be justified. The child who is coerced "tends to respond with hatred, and if, as is usual, he is not able to give free vent to his hatred, it festers inwardly, and may sink into the unconscious with all sorts of strange consequences throughout the rest of life."

Although puritanical moralists were or professed to be violently shocked by Russell's views on sex and education, it is worth emphasizing that his recommendations are not extreme and that unlike his opponents he stated his position temperately and without recourse to personal abuse. Russell may be characterized as a "libertarian" in education, but he was strongly opposed to the view of other educational pioneers who played down the importance of intellectual training and encouraged originality without insisting on the acquisition of technical skill. Similarly, although he may quite fairly be called a champion of free love, it is grossly misleading to describe Russell as an advocate of "wild living." On the contrary, he disavowed any such intentions. He wrote:

> The morality which I should advocate does not consist simply in saying to grown-up people or adolescents: "follow your impulses and do as you like." There has to be consistency in life; there has to be continuous effort directed to ends that are not immediately beneficial and not at every moment attractive; there has to be consideration for others; and there should be certain standards of rectitude. (*Marriage and Morals,* p. 243)

But this does not mean that we should be "dominated by fears which modern discoveries have made irrational." Russell could see nothing wrong in sexual relations before marriage, and he advocated temporary, childless marriages for most university students. This, he wrote, "would afford a solution to the sexual urge neither restless nor surreptitious, neither mercenary nor casual, and of such a nature that it need not take up time which ought to be given to work" (*Education in the Modern*

World, pp. 119–120). It would be wrong to regard Russell as an enemy of the institution of marriage. He did indeed object to keeping a marriage going when no love is left, and, what shocked people a great deal, he remarked that a "permanent marriage" need not exclude "temporary episodes," but he also emphatically affirmed that "marriage is the best and most important relation that can exist between two human beings … something more serious than the pleasure of two people in each other's company" (*Marriage and Morals,* p. 115).

Russell's views on sexual morality featured prominently in the New York City case of 1940. When his appointment was announced, Bishop Manning of the Episcopal Church wrote an inflammatory letter to all New York City newspapers in which he denounced Russell's subjectivism in ethics and his position on religion and morality. It was unthinkable that "a man who is a recognized propagandist against both religion and morality, and who specifically defends adultery" should be held up "before our youth as a responsible teacher of philosophy." The bishop's letter was the beginning of a campaign of vilification and intimidation unsurpassed in a democratic nation in recent times. The ecclesiastical journals, the Hearst press, and numerous Democratic politicians joined in the chorus of abuse. Russell was described as "the Devil's minister to men," as an advocate of "the nationalization of women," as "the mastermind of free love and of hatred for parents," and also, needless to say, as an exponent of communism.

The climax of the campaign was a taxpayer's suit by a Mrs. Jean Kay of Brooklyn demanding that Russell's appointment be annulled. The case was heard before Justice McGeehan, who had previously shown his notions of tolerance by trying to have a portrait of Martin Luther removed from a courthouse mural illustrating legal history. In a startling decision, which was bitterly criticized by legal experts as in many respects grossly improper, McGeehan voided Russell's appointment on three grounds: First, Russell had not been given a competitive examination; second, he was an alien and there was no reason to suppose that the post in question could not be competently filled by an American citizen; and, finally, the appointment would establish "a chair of indecency." Elaborate arguments were adduced in behalf of this last claim. Among other things it was maintained that Russell's doctrines would tend to bring his students "and in some cases their parents and guardians in conflict with the Penal Law." In some fashion not explained by the judge, Russell's appointment would lead to "abduction" and rape. Russell's opposition to the laws that make homosexuality a crime was misread as advocacy of a "damnable felony … which warrants imprisonment for not more than 20 years in New York State." Evasive actions of the mayor of New York, Fiorello La Guardia, prevented any effective appeal against this monstrous decision, and Russell was never able to take up his position at City College. In 1950, shortly after receiving the Nobel Prize, he returned to New York to deliver the Machette Lectures at Columbia University. He received a rousing reception that those who were present were not likely to forget. It was compared with the acclaim given Voltaire in 1784 on his return to Paris, the place where he had been imprisoned and from which he had later been banished. As for McGeehan, it is safe to say that he will go down in history as a minor inquisitor who used his one brief moment in the limelight to besmirch and injure a great and honest man.

McGeehan did not pass judgment on Russell's competence as a philosopher, but other opponents of the appointment were not so restrained. Thus, Joseph Goldstein, attorney for Mrs. Kay, described Russell as "lecherous, libidinous, lustful, venerous, erotomaniac, aphrodisiac, irreverent, narrow-minded, untruthful, and bereft of moral fiber." After a few gratuitous lies about Russell's private life, he concluded:

> He is not a philosopher in the accepted meaning of the word; not a lover of wisdom; not a searcher after wisdom; not an explorer of that universal science which aims at the explanation of all phenomena of the universe by ultimate causes … all his alleged doctrines which he calls philosophy are just cheap, tawdry, worn-out, patched-up fetishes and propositions, devised for the purpose of misleading the people.

In the present encyclopedia a somewhat different view is taken of the value of Russell's philosophy. Some of his most important theories in epistemology and metaphysics will be discussed in the next section, his contributions to logic and the foundations of mathematics will be covered in the following section, and his views on ethics and religion will be dealt with in the last section. However, a number of Russell's most interesting ideas are not at all or only briefly discussed in the present entry. Many of these are treated elsewhere in the encyclopedia.

EPISTEMOLOGY AND METAPHYSICS

Russell exercised an influence on the course of Anglo American philosophy in the twentieth century second to that of no other individual. Yet, unlike many influential thinkers, he neither founded nor attached himself to any

definite movement. Although he wanted above all to be empirical, he always had reservations of one sort or another to the proposition that all acceptable beliefs can be derived from purely empirical premises, and although his stress on analysis as the proper philosophical method is one of the chief sources of the analytical bent that philosophy currently has in English-speaking countries, he never accepted the view that philosophy is nothing but analysis.

EARLY REALISM. Russell's first distinctive philosophical work was colored by a violent reaction against the absolute idealism then dominant in England, which was ultimately based on the thought of G. W. F. Hegel and whose outstanding British exponent was F. H. Bradley. According to Bradley if we try to think through the implications of any fact whatever, we will inevitably be forced to conclude that everything that there is constitutes a single, immediate unity of consciousness. In Russell's view the main weapon used to bludgeon people into submission to this result was the "doctrine of internal relations," according to which any relational fact—for example, that *x* is above *y*—is really a fact about the natures of the terms involved. This doctrine in effect refuses to take relations as ultimate.

It follows from this position that whenever *x* and *y* are related, each "enters into the nature of the other." For when *x* is above *y,* then being above *y* is part of the nature of *x* and being below *x* is part of the nature of *y.* Hence, *y* is part of the nature of *x* and *x* is part of the nature of *y.* Since everything is related to everything else in one way or another, it follows that everything else enters into the nature of any given thing, which is just another way of saying that there is no "other thing" relative to a given thing. In other words, the only thing that exists is one all-comprehensive entity. From the related principle that when we are aware of something, that something enters into the nature of the awareness or of the mind which has the awareness, it follows that it is impossible to conceive of anything which is not included within consciousness. Thus, the one all-comprehensive entity is a unity of consciousness.

Although in his youth Russell, with most of his philosophical contemporaries, was caught up in this philosophy, he and G. E. Moore became disenchanted with it shortly before the turn of the twentieth century. Russell came to hold that in sense perception we are as immediately aware of the relations between things as of the things themselves and therefore that any philosophy which denied ultimate reality to relations must be mistaken. Moreover, he came to think that mathematics would be impossible if we held that every relation enters into the nature of its terms; for in mathematics we must understand what our units are before we can know anything about their relations to other units. Russell therefore argued for a "doctrine of external relations," according to which relations have a reality over and above the terms they relate and do not enter into the definition of the terms they relate. This led him to a kind of philosophical atomism that thenceforth was characteristic of his philosophy. We may think of the basic core of atomism, which runs through all the shifts in Russell's later philosophizing, as being constituted by the following principles:

(1) There are nonmental facts that are what they are whether or not any mind ever becomes aware of them. This does not follow from the doctrine of external relations, but that doctrine enabled Russell to reject the idealistic argument based on the doctrine of internal relations and thus left him free to hold his native realist convictions with a good conscience.

(2) A particular proposition (for example, that my car is in the garage) can be unqualifiedly true "in isolation." This follows from the thesis that facts are "atomic" in the sense that any given fact could hold, whatever is the case with the rest of the world, together with the correspondence theory of truth—that what makes a true proposition true is its correspondence with an objective fact. Hegelians, on the other hand, had argued that since one could not adequately think about any particular fact without inflating it into the absolute totality of being, whenever one is saying something short of everything, what he is saying is not quite true in any absolute sense.

(3) An important corollary of (2) is the usefulness of analysis as a method in philosophy. If it is possible to get an adequate grasp of the parts of a totality without considering their place in the whole, then it is possible to give an illuminating account of something complex by showing how its simple parts are related to form the whole. Hegelians had argued that analysis cannot get started because we cannot understand what any part is without already seeing how it fits into the whole, which means already knowing everything about the whole. The conviction that analysis is the proper method of philosophy has remained the most prominent strand in Russell's thought.

Intoxicated by his release from idealism, Russell, as he later put it, tended to accept as objectively real anything that the absolute idealists had not succeeded in showing to be unreal. Numbers, points of space, general

properties like roundness, physical objects as they appear to sense perception, were all regarded as having an independent existence. Under the influence of Alexius Meinong this extreme realism was reinforced by an extreme form of the referential theory of meaning, the view that in order for a linguistic expression to have a meaning there must be something that it means, something to which it refers. In this stage of Russell's thought, represented most fully by *The Principles of Mathematics* and to a lesser extent by *The Problems of Philosophy,* Russell was inclined to think that the meaningfulness of the sentence "The car is in the garage" required that there be objectively existing referents not only for the words *car, garage,*and *in* but even for the words *the* and *is.* An objectively existing "isness" soon proved to be too much for Russell's self-proclaimed "robust sense of reality." He came to think that terms belonging to the logical framework of sentences, such as "the," "is," "or," could perform their function without each being correlated with extralinguistic referents. Nevertheless, a modified form of the referential theory of meaning continued to dominate Russell's thinking.

LOGICAL CONSTRUCTIONISM. Russell's decisive shift away from the full-blooded realism of *The Principles of Mathematics* came with the development of logical constructionism. The theory can be generally stated as follows. We start with a body of knowledge or supposed knowledge which we feel strongly inclined to accept but which has the following drawbacks: (1) the knowledge claims do not seem to be adequately justified, (2) there are unresolved problems about the natures of the entities involved, and (3) we feel uncomfortable about committing ourselves to the existence of such entities. If we can show that this body of knowledge could be formulated in terms of relations between simpler, more intelligible, more undeniable entities and that when so formulated there is a decisive justification for it, we will have made a philosophical advance. We will have converted the problematic to the unproblematic, the obscure to the clear, the uncertain to the certain. Russell called this technique logical constructionism because the problematic entities were said, in a possibly misleading metaphor, to be "constructed" out of the simpler ones.

Reduction of mathematics to logic. The technique of logical constructionism was first employed in the theory of mathematics worked out by Russell and A. N. Whitehead and published in *Principia Mathematica* (3 vols., 1910–1913). In the *Principia* the authors set out to show that all of pure mathematics can be stated in terms of logic, using no undefined terms other than those required for logic in general—for example, implication, disjunction, class membership, and class inclusion. In the course of carrying out this reduction, various more or less problematic mathematical entities were "constructed" out of what were thought to be less problematic entities. Thus, numbers were defined as classes of classes: Zero is the class of all empty classes. The number 1 is the class of all classes each of which is such that any member is identical with any other member. The number 2 is the class of all classes each of which is such that it includes a member not identical with another member and such that any member is identical with one or the other of these. If one is puzzled about what sort of entity a number is (it does not seem to be in space or time and is not perceivable by the senses) or is uncomfortable about assuming that such queer entities exist, he will presumably be reassured by the discovery that he can think of numbers as classes of classes of familiar, unproblematic entities. Of course analogous problems may arise with respect to the entities made use of in this first reduction—for example, classes. And in fact various difficulties in doing mathematics in terms of classes led Russell to try to "construct" classes out of "prepositional functions." (See the section on logic and mathematics, below.) Starting from a given point we may well have to perform a series of reductions before we get down to maximally intelligible, indubitable entities.

Construction of physical objects. After *Principia Mathematica,* Russell applied the technique of logical constructionism to our knowledge of physical objects, both in physical science and in common sense. Physical theories are formulated in terms of a variety of unperceivable entities—electromagnetic fields, protons, energy quanta, forces exerted at a point, and so on. There are serious problems in the philosophy of science both about the content of our concepts of such entities and about the basis for our accepting their existence. We can try to show that such entities can be inferred from what we know about perceivable entities, but how could we get an empirical basis for a principle correlating observed and unobserved entities? Or we can try to show that unobserved entities have to be postulated in order to give an adequate explanation of observed happenings, but it seems impossible to show conclusively that no adequate explanation could be given purely in terms of observables. If we apply the constructionist principle, "Whenever possible, substitute constructions out of known entities for inferences to unknown entities," to this problem, we shall try to show that electromagnetic fields can be construed as complexes of less problematic entities related in various ways. Russell devoted a large proportion of his philosophical energy to trying to show that sci-

entific entities can be constructed out of undeniable data of perception. But it will be easier to illustrate this kind of analysis by taking ordinary physical objects like trees and buildings, for Russell thought that they raise analogous problems, although in less obvious ways.

There is a long tradition, dominant since the time of René Descartes, according to which common sense is mistaken in supposing that we directly perceive physical objects. According to this tradition what we are directly and indubitably aware of in sense perception is something private to the individual observer. There are several sources of this view, the most important of which are, first, the fact that the content of one's perception can change with, for example, changes in perspective, lighting, and physiological condition of the observer, without there being any change in the physical object which, according to common sense, one is perceiving, and, second, the fact that in dreams and hallucinations one can have experiences which are intrinsically indistinguishable from those one has when one is "really" seeing a tree, but in these cases no tree is present. In dreams and hallucinations one is really aware of something that is not a physical object and is not perceivable by anyone else. And since these experiences are intrinsically just like those in which a physical object is present, one must be perceiving these private objects in the latter cases as well. This consideration is reinforced by the first, which is designed to show that even where a physical object admittedly is involved, I am often aware of different things without the physical object's undergoing any change.

The conclusion of these arguments is that the colors, shapes, sounds, and so on, of which we are directly aware in sense perception (sense data) are private objects that must be distinguished from the entities in the physical world (if any) which we suppose ourselves to be perceiving. This conclusion inevitably gives rise to the question how, if at all, I can start from the private objects of whose existence I can be certain and show that public, physical objects like trees exist. No generally accepted solution to this problem has emerged in several centuries of discussion. Here again Russell tries to avoid the necessity for an inference by showing that the public physical objects can be construed as a complex structure of data of immediate experience. At first Russell aimed at a solipsistic reduction in which a given physical object would be constructed out of the actually experienced data of a single observer, but he soon came to lower his aspiration and to admit into the construction data experienced by others, as well as data which would have been experienced by others if they had been in a certain place. The view, then, is that a tree can be regarded as a system of all the actual and possible sense experiences that would be regarded as figuring in perceptions of that tree. This is a form of the position known as phenomenalism, and it is subject to the difficulties to which that position is notoriously subject, particularly the apparent impossibility of specifying which experiences go into defining a particular physical object without referring to that physical object or others in the specification.

Construction of mind. Until about 1920 Russell was a mind-matter dualist. As we have just seen, physical objects were regarded as complex structures of data of the sort given in sense perception. Now, although the mind might be partly constituted by data which are given to "inner sense"—that is, things which are the objects of introspective awareness, such as images and feelings—it seemed to Russell, as it had to most philosophers, that in any act of awareness, be it directed to the external or to the internal world, there is in addition to the data of which one is aware a subject or self which has the experience or which performs the act of awareness. But as the spirit of logical constructionism took increasing hold of Russell, he came to feel that there was no real warrant for believing in a subject of awareness which performs acts. He became convinced that one cannot really find any such constituent of the experience; its apparent obviousness is a reflection of the grammar of the sentences in which we speak about such matters—we say "*I* saw a flash of light" rather than "A flash of light occurred." As it presents itself, a minimal piece of consciousness does not involve a relation between two components. It is a unitary whole. Only the flash of light is given. The "I" and the "saw" are added interpretations. If we have no real basis for accepting a subject or mind as an ultimate entity, then the logical constructionist will try to show that it can be exhibited as a complex of entities of which we are directly assured by our experience. Here Russell followed the lead of William James, who had earlier formulated a view known as neutral monism, according to which both mind and matter consisted of the data of immediate experience, the difference between them lying in the grouping of the constituents.

Thus, if I am looking at a tree the visual datum (an irregularly shaped green splotch) of which I am directly aware is both part of my mind and part of the tree. When grouped together with other experiences from this and other perspectives that would be said to be experiences of that tree, it goes to make up a tree; when grouped with other data bound together in a single conscious field, along with other data related to these by memory, it goes

to make up a mind. If this theory is acceptable, traditional puzzles about the mind-body relation are dissolved. We are faced not with two radically distinct kinds of stuff but with two different kinds of arrangement of the same elementary components. (That is, some of the components are the same. Russell considers images to be peculiar to mind.) It is in the light of this theory that one should consider Russell's notorious view that what one perceives is always his own brain. Whenever I have any sense perception whatever, I do so because a certain kind of physical activity is going on in my brain. This activity, as a physical process, is to be regarded, like all physical processes, as a construction out of the sort of data given in immediate experience. And since whatever may be the case otherwise, my brain is always active when I perceive, the data of which I am aware enter into the constitution of my brain, whatever other entities they may enter into. Hence the paradoxical view that whenever one is conscious he is aware of his own brain.

When Russell abandoned the subject of experience as an ultimate constituent of the world he rejected sense data and thenceforth spoke simply of sense experiences. But he would have represented his view more clearly by saying that he had given up belief in anything other than sense data. For in the old paradigm of *subject aware of sense data,* it was the subject exercising awareness that was abandoned. In *The Analysis of Mind* Russell set out to construct the conscious mind out of sensations and images. (Insofar as facts regarded as mental do not consist of consciousness, Russell's strategy is to give a behavioristic analysis. Thus, desire, belief, and emotions can be regarded as made up, at least in part, of dispositions to behave in one way rather than another in certain circumstances.) The results are admittedly equivocal. Russell has always been too honest to overlook glaring deficiencies in his analyses. One that has particularly bothered Russell is this: On a commonsense basis it seems clear that one must distinguish between simply having a sensation and taking that sensation as an indication of a tree, and there seems to be an important difference between simply having an image and employing that image in, for example, thinking about a forthcoming election. If this analysis of mind is to be made to work, one must give an account of the reference of perception and thought in terms of the interrelations of data. Thus, we might hold that to take a sensation as an indication of a tree is to be disposed to have the sensation of surprise if certain other sensations were to follow. But apart from difficulties about the nature of these dispositions, which are themselves neither images nor sensations, this is all extremely difficult to work out in detail, and it is equally difficult to make sure that one has shown that it can be done.

It is clear that logical constructionism is based on a tendency opposite to that of the realism briefly espoused by Russell in his youth. Logical constructionism wields Ockham's razor with a heavy hand. We begin with those entities whose existence is indubitable because they are given in immediate experience, and we then try to show that anything we might wish to say about anything else can be stated in terms of relations between these indubitable entities. In other words, anything we want to say about something else is not really about something else. Thus, we try to represent all our knowledge as having to do with as few kinds of entities as possible, thereby reducing the possibility of error.

LOGICAL ATOMISM. Thus far we have concentrated on the epistemological side of logical constructionism, its concern with reducing the number of assumptions we make and with exhibiting clearly the basis for what we claim to know. But it also has a metaphysical side, although Russell wavers about this. Sometimes he talks as if his constructionism is metaphysically neutral. At such times he says that in showing that minds can be constructed out of sensations and images we do not show that there is no ultimate, irreducible subject of awareness; we show merely that everything we know about minds can be expressed without assuming the existence of such an entity. At other times, however, he claims that by showing that minds can be constructed out of sensations and images we have shown what minds really are—we have revealed their metaphysical status. And by carrying through constructions of everything that can be constructed out of simpler entities we will have developed a complete metaphysical scheme.

Ideal language. The most systematic presentation of this metaphysical side of logical constructionism is found in the set of lectures *The Philosophy of Logical Atomism,* which Russell gave in 1918. Here Russell makes explicit the principle on which a metaphysical interpretation of logical constructionism depends—namely, isomorphism of the structure of an ideal language and of the structure of reality. If we can determine in outline how the world would be described in an ideal language, we will have, in outline, an account of what the world is like. The restriction to an "ideal" language is essential. Since there are alternative ways of stating the same body of facts, it could not be the case that all these ways reflect the real structure of the world. In this approach to metaphysics the basic metaphysical commitment is to the identity of structure

between reality and an ideal language, and one shows one's hand metaphysically by choosing one rather than another set of criteria for an ideal language.

For Russell the most important requirement for an ideal language is an empiricist one, formulated in the "principle of acquaintance": "Every proposition which we can understand must be composed wholly of constituents with which we are acquainted." In other words, we can understand a linguistic expression only if it either refers to something we have experienced or is defined by other expressions which are so used. This principle plays a part in the constructions we have been surveying, as do the considerations we have already made explicit. That is, Russell holds not only that if physical objects were not defined in terms of sense experiences we would have no way of *knowing* anything about them but also—and even more important—we would not be able to *understand* talk about them. In logical atomism this principle is reflected in the requirement that the expressions which figure in the "atomic" sentences in terms of which everything is to be expressed must get their meaning through direct correlation with experience. They will, therefore, be names of particular sense data and terms for properties of sense data and relations between sense data. Russell is forced to exclude the logical framework of sentences from this requirement ("is," "the," etc.), but he is recurrently uneasy about this exclusion and recurrently disturbed by the question how, in that case, we can understand them.

In addition to the need for its undefined terms getting their meaning through correlation with immediately experienced items, the ideal language will have to satisfy some more strictly logical requirements. These will include the absence of vagueness and having one and only one expression for each meaning. But the most important restriction concerns the form of the basic sentences. An atomic sentence will be one that contains a single predicate or relational term and one or more than one name, the whole sentence asserting that the entity named has the indicated property ("This is white") or that the entities named stand in the indicated relation ("This is above that"). If a sentence (1) has this form, (2) contains only terms that get their meaning through correlation with experienced items, and (3) has to do with entities that cannot be analyzed into anything simpler, then it is an atomic sentence. It is clear that for Russell the sentences which satisfy these requirements will all state a minimal fact about a momentary content of sense experience.

Logical atomism can then be presented as the thesis that all knowledge can be stated in terms of atomic sentences and their truth-functional compounds. A truth-functional compound of two sentences is one whose truth or falsity is a determinate function of the truth or falsity of the components. Thus, "I am leaving and you are staying" is a truth-functional compound of "I am leaving" and "You are staying." For the compound is true if and only if both its components are true. There is an empiricist motivation for maintaining this thesis. Atomic sentences, in the sense specified above, can be conclusively verified or falsified by a single experience, and as long as we are dealing only with truth-functional compounds of these no further problem can arise concerning their truth or falsity. Consider a "contrary-to-fact conditional," such as "If I had offered him more money, he would have accepted the job." As it stands this sentence is not a truth-functional compound of its constituents. For in saying it we are presupposing that both its constituents are false, yet this does not settle the question whether the whole statement is true or false. There is a corresponding puzzle about what empirical evidence would settle the question. Obviously I cannot go back in time and offer him more money and see what he will do. If we could find some way to restate this as a (very complicated) truth-functional compound of atomic sentences, it would become clear which experiences would verify or falsify it.

Pluralism and knowledge by acquaintance. The metaphysical correlate of this sketch of the ideal language brings together two of Russell's deepest convictions, the logical independence of particular facts (pluralism) and the dependence of knowledge on the data of immediate experience. In this view reality consists of a plurality of facts, each of which is the sort of fact which could be infallibly discerned in a single moment of experience and each of which could conceivably be what it is even if nothing else were in existence. All the familiar and seemingly relatively simple objects in the world of common sense are really extremely complicated complexes of atomic facts of these sorts.

Russell was well aware that logical atomism in this extreme form was untenable. For example, he insisted that generalizations could not be truth-functional compounds of atomic sentences. The most promising way of so construing them would be to take, for example, "All lemons are yellow" as a conjunction of a large number of atomic sentences of the form "This lemon is yellow," "That lemon is yellow," …. But as Russell points out, even if it were possible to list *all* the lemons, the conjunction would say the same thing as the original universal generalization only if we added the conjunct "and that is all the lemons there are." And this last addition is not an atomic

sentence. Moreover, Russell had doubts about so-called intensional contexts, such as "Smith believes that the White Sox will win," where the truth or falsity of the compound is clearly independent of the truth or falsity of the components. Whether Smith has this belief does not in any way depend on whether the White Sox win. Russell has always hoped that neutral monism would help him to get out of this difficulty. If we could construct beliefs out of sensations and images we might be able to restate this fact as some truth-functional derivative of atomic sentences.

Later doubts. In the mid-twentieth century Russell came to have more fundamental doubts about logical atomism, including doubts concerning the very notion of a logical atom. How can we ever be sure that we are dealing with something that cannot be further analyzed into parts? How can one be sure that yellowness is an absolutely simple property? More basically, what makes a property *logically* simple? Does the fact that one can explain the word *yellow* to someone by saying "Something is yellow if it has the same color as the walls of your room" show that being the same color as the walls of your room is logically a *part* of yellowness? If so, then yellowness is not absolutely simple. If not, what does count against logical simplicity? Moreover, if there are alternative minimum vocabularies, then a simple, undefined term in one mode of formulation may turn out to be definable in another. Thus, on one systematization "pleasure" might be defined as the satisfaction of desire, whereas on a different systematization "desire" would be defined as the belief that something is pleasant. Russell gave up the belief that we can know that we have gotten down to ultimate simples and even the belief that there must be absolute simples. He became disposed to think, in more relativistic terms, of a class of things that can be taken as simple at a given stage of analysis. In those terms he still tended to fall back on sense experiences that are as apparently simple as anything we can find. Such experiences, even if not absolutely simple, can be regarded as being independent of anything except their possible components.

LATER DEVELOPMENTS. Despite Russell's frank admissions that logical atomism does not work as a depiction of the structure of an ideally adequate language, he did not develope an alternative metaphysics. On the principle of isomorphism, if one cannot represent general statements as functions of atomic statements, then one must admit general facts as ultimate constituents of the world. This metaphysical implication did not seem to bother Russell as it once had. This is partly because he became less preoccupied with metaphysics in his later years and partly because the principle of isomorphism became so heavily qualified as to remove most of the cutting edge. In his major philosophical works of the 1940s, *An Inquiry Into Meaning and Truth* and *Human Knowledge,* he is more concerned with the nature of atomic facts thought of as the ultimate pieces of empirical data and the kinds of inferences required to get from these to the rest of what one wants to count as knowledge than he is with inferring a metaphysical structure from the logical form which an adequate statement of our knowledge would assume.

In these works there is a major shift in his view of the structure of atomic facts. Russell had earlier interpreted the word *this* in "This is red" as referring to a particular, something which has qualities and stands in relations but is not itself a quality or relation or set of qualities or relations. This is the traditional concept of substance as the substratum of properties, which was still alive in the realm of sense data even after physical objects and minds were no longer taken to be substances. But eventually the sense datum as substratum of properties went the way of physical objects, minds, and numbers. Here, too, Russell became convinced that there is no empirical warrant for assuming the existence of any such thing. In sense experience I am aware of a variety of qualities and their interrelations, but I am not also aware of something which *has* qualities. The bearer of qualities turns out to be the shadow of the usual grammatical form of the sentences used to report atomic facts. (There is a subject of the sentence—for example, "this"—which does not refer to any quality.) Russell's latest position was that the subject of qualities is simply a construction out of a set of compresent qualities. Thus, in the ideal language "This is red" would be restated as "Red is compresent with …," where in place of the dots we have a specification of the other properties involved in that experience, for example, being round, being in the middle of the visual field, having ragged edges. It might be thought that this necessarily involves giving up the idea of absolute simples, for what takes the place of things in this view is bundles of qualities. But in this theory qualities themselves are regarded as the ultimate particulars (possibly simple) of which the world consists. Thus, in "Red is compresent with …," "red" does not refer to a particular exemplification of redness. If we took that line we would have to suppose that there is something which distinguishes *this* exemplification from other exemplifications of just the same color, and that would have to be something as unempirical as a substratum. Instead, it is taken to refer to the color conceived as a "scattered particular," something which can

exist in a number of different places at the same time. And such a particular might well be simple.

Russell continued to think of commonsense physical objects and the entities of physics as constructions out of entities of the sort that are given in sense experience. But he came to require less similarity to sense data in the elements of these constructions. His later view was that although all ultimate entities have basic structural similarities to sense experiences, they need not involve only qualities that are given in sense experience. They may have qualities that it is impossible for us to be aware of. This uncertainty does not carry with it any serious gap in our knowledge, since for physical science it is the structure of external events that is important. In the 1940s Russell became increasingly concerned with the principles that are required to justify inferences from sense experience to unexperienced events and complexes of unexperienced events. The simplest form this takes is, for example, the inference that my desk has continued to exist in my office throughout the night, when no one was observing it. On Russell's view this is an inference from certain sense experiences to structurally similar events spatiotemporally connected with them in certain ways. He felt that the principle of induction by simple enumeration (the more often one has observed *A* and *B* to be associated, the more it is likely that they are invariably correlated) is insufficient to justify such inferences. What is needed, he thought, is a set of assumptions having to do with spatiotemporal connections of events of like structures. In *Human Knowledge* he presents a set of such assumptions. He does not claim that they can be *known* to be true. His point is a Kantian one: We must accept these assumptions if we are to accept the inferences to unobserved events that we all do accept in the course of our daily life.

Russell's entire philosophical career was dominated by the quest for certainty. In the middle decades of the twentieth century he was driven to admit that it is less attainable than he had hoped, but nevertheless the desire to approximate it as much as possible continued to shape his thinking about knowledge and the nature of the world. Because of this desire he was continually preoccupied with the problem of how to formulate those pieces of knowledge that are rendered indubitable by experience. And because of it he consistently attempted to analyze anything that appears dubitable into constituents about which there can be no doubt. Even where he was forced to admit that inferences beyond the immediately given are inevitable, he strove to reduce the principles of such inferences to the minimum. Russell is distinguished from other seekers after absolute certainty chiefly by the ingenuity of his constructions and by the candor with which he admits the failures of the quest.

LOGIC AND MATHEMATICS

REDUCTION OF MATHEMATICS TO LOGIC. Russell's main work in logic and mathematics was concerned with the problem of bringing the two together and with the interpretation of mathematics—arithmetic in particular—as a simple extension of logic, involving no undefined ideas and no unproved propositions except purely logical ones. Russell achieved this synthesis at the beginning of the twentieth century, a little later than Gottlob Frege, but independently of him; in working it out in detail he had the collaboration of A. N. Whitehead. By current standards Russell's work lacks rigor, and in this respect it compares unfavorably with that of Frege; at an early stage, however, Russell did notice a difficulty that had escaped Frege's attention, the paradox about the self-membership of classes, which will be examined later. Because of its complexity it will be best to treat Russell's picture of the logical foundations of mathematics systematically rather than historically, with occasional comments about the actual development of his thought. We shall also separate from the outset two elements of Russell's treatment of his and other paradoxes, the theory of "types" and the theory of "orders," which Russell himself ran together, and thereby give a slightly clearer picture of his intention than his own writings immediately furnish.

Definition of "similarity." Russell took over from Giuseppe Peano the reduction of all other arithmetical notions to complications of the three arithmetically undefined ideas of "zero," "number," and "successor" and defined these in terms of the theory of logical relations between classes or sets. In particular, he defined a number as a class of classes with the same number of members; for example, he defined the number 2 as the class of pairs. This procedure may seem unnatural (do we really mean by "2" the class of two-membered classes?) and circular. To the charge of unnaturalness Russell's answer was that his definition (together with the definitions of addition, etc.) gives all the ordinary results (2 + 2 = 4, for example) and that for a pure mathematician this is enough; another answer can be given only after it has been made clearer what Russell means by a class. With regard to the charge of circularity, Russell defines the complex "having the same number of members," or "similarity," as he calls it, not in terms of "number" (or of his definition of "number") but in other terms altogether.

At this point some notions from the logic of relations have to be introduced. A relation is said to be one to one if whatever has that relation to anything has it to one thing only and if whatever has anything standing in that relation to it has one thing only standing in that relation to it. (In strictly monogamous countries, "husband of" is a one-to-one relation in this sense.) Here the phrase "one only" does not presuppose the notion of the number 1. The sentence "*x* stands in the relation *R* to one thing only" means "For some *y,* whatever *x* stands in the relation *R* to is identical with *y.*" The domain of a relation is the set of objects that stand in that relation to anything (the domain of "husband of" is the class of all husbands); the relation's converse domain is the set of all objects to which anything stands in that relation (the converse domain of "husband of" is the class of individuals that have husbands—that is, the class of wives). A class *A* is similar to (that is, has the same number of members as) another class if there is some one-to-one relation of which the first class is the domain and the second the converse domain.

One can see that in a monogamous country the class of husbands will be similar in this sense to the class of wives, but one might think that two sets of objects could have the same number of members without there being any relation at all that pairs them off in the way that "husband of" does in our example. This, however, is a mistake when the term *relation* is understood as widely as it is by Russell. A relation in Russell's sense is, roughly, anything that can be expressed by a sentence with two gaps in it where names might go, and this covers not only obvious relating expressions like "______ shaves (   )" or "______ is the husband of (   )" but also ones like "Either ______ is identical with *A,* or *B* is identical with (   )." Take any set of two objects *C* and *D.* The relation "Either ______ is identical with *A* and (   ) with *C,* or ______ is identical with *B* and (   ) with *D*" (where all dashes must be replaced by the same name, and similarly with the bracketed blanks) will be a one-to-one relation in which *A* stands to *C* alone and *B* to *D* alone and in which *C* has *A* alone standing to it and *D* has *B* alone—that is, it will be a one-to-one relation of which the class with *A* and *B* as sole members is the domain and the class with *C* and *D* as sole members the converse domain; there are analogous relations in the case of larger classes. (Where the classes are infinitely large these relations will not be expressible in a language with only finite expressions, and perhaps that means that they will not be expressible in any language. Some philosophers would regard this as a serious difficulty; others would not.)

Axiom of infinity. Similarity, then, or having-the-same-number-of-members, is defined in terms of notions from the logic of relations: one to one, domain, and converse domain. The number-of-members of a given class is the class of classes similar to it, and a class of classes is a number (strictly, a cardinal number) if there is some class of which it is the number-of-members. This last step gives rise to another difficulty: Suppose there are (as there might well be) no more than a certain number *n* of objects in the universe. Then there will be no classes with more than *n* members and so, by the above definition, no cardinal numbers greater than *n.* This makes a great part of arithmetic (for example, the principle that every number has a successor different from itself) subject to the hypothesis (sometimes called the axiom of infinity) that there are an infinite number of objects.

Russell came to accept this last consequence of his definitions, but at an earlier stage he had thought that the axiom of infinity was provable, as follows: If we assume that every property demarcates a class, we must admit that some classes are empty (have no members), for example, there are no objects not identical with themselves. (The number 0 is precisely the class of classes with no members.) Thus, even if the universe contains no ordinary objects at all, there will still be at least one object of a more abstract sort, the universe itself considered as an empty class. And if there is this object there will also be two further objects of a still more abstract sort: the class of classes that has the first empty class as its one member and the empty class of classes. That makes three objects, call them *A, B,* and *C.* In addition to these there will be four classes of classes of classes—the class with *B* as its sole member, that with *C* as its sole member, that containing both *B* and *C* as members, and the empty class of classes of classes. And so on ad infinitum.

Russell paradox and the theory of types. Russell was led to abandon the above demonstration (which, as he said, has anyway "an air of hocus-pocus about it") by his discovery of the paradox of self-membership, mentioned earlier. If we can concoct classes with some members that are themselves classes, some that are classes of classes, and so on as we please (if, in other words, we can treat classes, classes of classes, etc., as so many sorts of classifiable "objects"), we can, it seems, argue as follows: The most obvious classes do not contain themselves as members—for example, the class of men is not itself a man and so is not itself a member of the class of men (that is, of itself). On the other hand, the class of non-men *is* a non-man (is one of the things that are not men) and thus *is* a member of itself. We can therefore divide classes into two broad

classes of classes—the class of classes that are members of themselves and the class of classes that are not. Now take the class of classes that are not members of themselves: Is it a member of itself or not? If it is, it must possess the defining property of this class to which ex hypothesi it belongs—that is, it must be not-a-member-of-itself. (Thus, if it is a member of itself, it is not a member of itself.) And if it is not a member of itself, ipso facto it possesses its own defining property and so *is* a member of itself. (If it is not, it is.) Let *p* be the proposition that our class is a member of itself; it follows even from the attempt to deny it, so it *must* be true—but it entails its own denial, so it *must* be false. There is clearly something wrong here.

Russell thought the error lay in treating a class seriously as an object. Perhaps it is an object in a sense, but not in the same sense in which genuine individuals are objects—and classes of classes are different again. They are, as he put it, of different "logical type." In particular, in an intelligible sentence you cannot replace an individual name by a class name or a class name by the name of a class of classes, or vice versa, and still have the sentence make sense. If "Russell is dead" makes sense, "The class of men is dead" does not, and if "The class of men is three-membered" makes sense (even if false), "Russell is three-membered" does not. And where a sentence makes no sense (as opposed to being merely false), its denial makes no sense either. Since "The individual *I* is a member of the class-of-individuals *C*" makes sense, "The class-of-individuals *C* is a member of the class-of-individuals *C*" does not and neither does "The class-of-individuals *C* is not a member of the class-of-individuals *C*"—and so on at higher points in the hierarchy. This being granted, the paradox with which we began simply cannot be intelligibly formulated and thus disappears from the system.

At this point it would be wise to remove a possible source of confusion. The relation of class membership is different from the relation of class inclusion. One class is included in another if all the members of the former are members of the latter; for example, the class of men is included in the class of animals—all men are animals. But the class of men is not a member of the class of animals; that is, the class of men is not an animal (or, more strictly, "The class of men is an animal" is nonsense). The class of men is a member, rather, of the class of classes-of-animals—it is a class of animals. And the class of classes of animals is included in (but is not a member of) the class of classes of living things—any class of animals, in other words, is a class of living things. Inclusion thus relates classes of the same logical type; membership, on the other hand, relates an entity with another entity of the logical type one above its own. The membership of an individual in a class of individuals is membership in a sense different from the membership of a class of individuals in a class of classes, and similarly for inclusion—there is a hierarchy not only of classes but also of membership and inclusion relations.

All this, besides solving a technical problem, is not without some attraction for philosophical common sense. Even apart from paradoxes it seems an artificial "multiplication of entities" to suppose that in addition to the individual objects which form the members of the lowest type of classes there are classes, classes of classes, and so on, and Russell devoted some attention to the problem of showing how what appears to be talk about these rather strange objects is in reality just more and more oblique talk about quite ordinary ones. To see just how he shows this it is necessary to look more closely at what might be called his "straight" language, into which this talk of classes, etc., does not enter and into which, once this talk *has* been introduced, it can always be "translated back."

LOGIC. From what has been said so far, it is clear that the "logic" to which Russell reduced arithmetic covered, implicitly or explicitly, such subjects as class membership and class inclusion, identity, and some sort of theory of relations. This is that part of logic that we first encounter when we work back to logic from arithmetic. We must now try and work forward to the same point from the fundamentals of logic.

Russell thought of logic as being at bottom "the theory of implication" (to quote the title of one of his early papers). And from the first he considered it important to distinguish implication from inference. He objected to the view that logic is primarily about thinking—conception, judgment, and inference, as some of the traditional logic texts put it. The connection of logic with inference is rather that logic is concerned with that in the real world that makes inference justified, and this is implication. "Where we validly infer one proposition from another," he wrote in 1903, "we do so in virtue of a relation which holds between the two propositions whether we perceive it or not: the mind, in fact, is as purely receptive in inference as common sense supposes it to be in perception of sensible objects" (*Principles of Mathematics,* p. 33).

Material implication. Even in Russell's purely objective, nonpsychological sense "implication" is ambiguous. Implication may be a relation between complete propositions, in which case it is called "material" implication and

holds whenever it is not the case that the implying proposition is true and the implied proposition false. Before enlarging and commenting upon this account, certain grammatical and metaphysical clarifications are in order. Russell originally believed that sentences symbolized abstract objects called "propositions" and that material implication was a relation between these objects in exactly the same sense that marriage might be a relation between two people. He later dropped this view and regarded propositions, like classes, as mere "logical constructions," but he still used the old forms of words (as being, no doubt, accurate enough for practical purposes). In particular, the partly symbolic form "*p* implies *q*" (or "*p* materially implies *q*") freely occurs in all his writings, and we ought to be clear about what he means by it. Generally it is simply a variant of "If *p* then *q*," or completely symbolically "$p \supset q$" ("*p* hook *q*"), where the phrase "If ______ then ( )"—or the hook—is not a transitive verb expressing a relation between objects but a conjunction, or, as we now say, a "sentential connective." "If *p* then *q*" is thus not a statement about two objects symbolized by "*p*" and "*q*" but rather a complex statement about whatever the statements represented by "*p*" and "*q*" are about. For example, "If James is going to come, John will stay away" is not about two objects symbolized by "James is going to come" and "John will stay away," nor is it about these subordinate sentences themselves; rather, it links these two sentences to make a more complex statement about James and John. And if we say "That James is going to come implies that John will stay away," this is just a verbal variant of "If James is going to come then John will stay away"; that is, the linking expression "That ______ implies that ( )" has the same meaning as the conjunction "If ______ then ( )." The general form "That *p* implies that *q*" thus has the same sense as the form "If *p* then *q*" or "$p \supset q$," and Russell's "*p* implies *q*" is thus just a loose way of saying "That *p* implies that *q*." In a similar way Russell often uses "*p* is true" and "*p* is false" as variants of "It is the case (is true) that *p*" and "It is not the case (is false) that *p*"; although sometimes he may really be talking about sentences in such a way that the sentence "John will stay away" may be described as true if and only if John will stay away and as false if and only if he will not, and the sentence "James is going to come" may be said to "imply" the sentence "John will stay away" if and only if the sentence "If James is going to come then John will stay away" is true.

The assertion that an implication is true if and only if it is not the case that the implying statement (antecedent) is true and the implied statement (consequent) false is not intended as a definition of the form "If *p* then *q*." It is simply an informal attempt to fix our attention on the relation (or quasi relation) that Russell intends. In his earliest works, like Frege and C. S. Peirce before him, Russell took this relation to be indefinable, and "the discussion of indefinables—which forms the chief part of philosophical logic—is the endeavour to see clearly, and to make others see clearly, the entities concerned, in order that the mind may have that kind of acquaintance with them which it has with redness or the taste of a pineapple" (*Principles of Mathematics,* 1st ed., preface; 2nd ed., p. xv). Later he preferred to take as undefined the conjunction "or" and the negative prefix "it is not the case that" (or just "not") and to define "If *p* then *q*" as an abbreviation of "Either not *p* or *q*"; later still he followed H. M. Sheffer and Jean Nicod in using the stroke form "$p \mid q$" (which is true if and only if the component statements are not both true) and defined "if," "not," and "or" in terms of it. But for Russell the central part of logic has always been the study of implication, whether taken as undefined or not.

Since the form "If *p* then *q*" as understood by Russell is true as long as it is not the case that the antecedent is true and the consequent false, it is automatically true if the antecedent is false (for then it is not the case that the antecedent is true and thus not the case that the-antecedent-is-true-and-the-consequent-false) or the consequent true (for then it is not the case that the consequent is false and thus not the case that the-antecedent-is-true-and-the-consequent-false). In other words, a false proposition materially implies, and a true one is materially implied by, any proposition whatever. But implication is supposed by Russell to justify inference, and the mere fact that "Grass is pink" is false would not seem to justify us in inferring the 25th proposition of Euclid from it, and the mere fact that Euclid's proposition is true would not seem to justify us in inferring it from "Grass is green"—geometry would be much easier if we could do this. Russell's explanation is that the first of these inferences cannot be performed because we cannot get it started (the premise not being true) and that the second inference is justified but we cannot know it to be so unless we already know the conclusion, so that we will not need it. In other words, "Infer a true proposition from anything at all" is a rule with no practical use, but this does not make it logically wrong.

Formal implication and propositional functions. Implications are of practical use when we know their truth without knowing either the falsehood of their antecedents or the truth of their consequents, and this happens most often when a material implication is an

instance or particularization of an implication in the second of Russell's senses, a "formal" implication.

Formal implication is not (to use Russell's "realistic" language) a relation between propositions but one between what he calls "propositional functions." One might say roughly that formal implication is a relation between properties and that one property formally implies another if it is never present without the other; for example, being human formally implies being mortal (nothing is human without being mortal). Formal implication is clearly involved in the notion of class inclusion—*A* is included in *B* if being a member of *A* formally implies being a member of *B.* But the notion of a propositional function is wider than that of a property. It is what is meant by an "open sentence," a sentence in which some expression—say, a name—has been replaced by a variable. "Socrates is a man" expresses a proposition; "*x* is a man" expresses a propositional function. Sometimes, more simply, Russell uses the term *propositional function* for the open sentence itself. And the proposition that Socrates is a man may be said to be the *value* of the propositional function "*x* is a man" for the value "Socrates" of the argument *x.* The propositional function "*x* is a man" formally implies "*x* is mortal" if *x*'s being a man materially implies that *x* is mortal whatever *x* may be—that is, if we have "For any *x,* if *x* is a man then *x* is mortal." Russell writes this sort of implication as "$\varphi x \supset_x \psi x$." At one stage he treated this notion, for systematic purposes, as undefined, but even then he regarded it as complex in meaning, being built up from material implication together with the prefix "for any *x,*" called a quantifier. Writing this last as "(x)," we may spell out the sense of a formal implication by writing it as "$(x) : \varphi x \supset \psi x$." It should be noted that whereas a propositional function is not a proposition, a formal implication between such functions is a proposition. The propositional function "*x* is a man" is neither true nor false; only its various values are true or false. But "For any *x,* if *x* is human then *x* is mortal" is as it were complete and is as it happens true. The quantifier is said here to "bind" the variable *x,* or, in the terminology Russell took over from Peano, *x* is in this context not a "real" but an "apparent" variable.

A propositional function may also have more than one expression in a proposition replaced by a variable, as in "*x* shaves *y,*" "*x* gives *y* to *z,*" and "If *x* shaves *y* then *x* does not shave *z.*" In such cases the function corresponds to a relation (two-termed or many-termed) rather than to a property, and such functions may again be linked by formal implication, as in "For any *x* and *y,* if *x* is a child of *y* then *x* detests *y*"—that is, "All children detest their parents." Symbolically, we have here the form "$\varphi xy \supset_{x,y} \psi xy$," or "$(x,y) : \varphi xy \supset \psi xy$." Again, formal implication may link a propositional function and a complete proposition, as in "If anything is in that box I'm very much mistaken," which is of the form "For any *x,* if φx then *p*" or "$\varphi x \supset_x p$." Moreover, the expression whose place is taken in a propositional function by a variable need not be a name. It might, for example, be a sentence—"If *p* then *q*" is a propositional function of which "If James is going to be there then John will not come" is the value when "James is going to be there" is the value of the argument *p* and "John will not come" the value of the argument *q.* If we prefix quantifiers to forms of this sort we obtain further formal implications, including the laws of propositional logic themselves—for example, "For any *p, q,* and *r,* if *p* implies *q* then if *q* implies *r, p* implies *r,*" which may be written "$(p,q,r) : (p \supset q) \supset ((q \supset r) \supset (p \supset r))$" or "$(p \supset q) \supset_{p,q,r} ((q \supset r) \supset (p \supset r))$."

A further case of special interest is that in which a variable replaces a verb or equivalent expression, as in "φ(Socrates)," where φ stands indifferently for "is a man," "smokes," "is running," etc. With appropriate quantifiers this function will yield such formal implications as "$(\varphi) : \varphi a \supset \varphi b$," "$\varphi a \supset_\varphi \varphi b$" (roughly, "Whatever *a* does, *b* does," or "Whatever goes for *a* goes for *b*"). However, Russell says not that "φ(Socrates)" and "If φ(Socrates) then φ(Plato)" are functions of the verb or predicate φ but that they are functions of the *function* φx or, as he writes it in this type of context, the function $\varphi \hat{x}$ (the significance of this accenting or "capping" will be indicated later). His aim here is in part to bring out what Peirce and Frege called the "unsaturatedness" of verbs: The function of verbs can be understood only in relation to names and sentences; we use verbs to *make statements* about objects, not to name a special sort of object. The additional associated variables also enable one to represent unambiguously such complexes as "shaving oneself"—if φ is "shaves," shaving oneself is $\varphi \hat{x}\hat{x}$, as opposed to simply shaving ($\varphi \hat{x}\hat{y}$). But Russell was hampered by not having the word *functor* to designate what makes a function out of its argument; it is more natural to speak of "Socrates is a man" as a propositional function of "Socrates," of "*x* is a man" as the same propositional function of "*x,*" and of "is a man" as the functor which forms this function in both cases than to speak of "$\hat{x}$ is a man" as a propositional function and to treat it in practice as a functor (Frege and W. E. Johnson were more accurate here, although they, too, lacked the term *functor*).

This part of Russell's "philosophical grammar" can now be set out fairly straightforwardly: Sentences may be

built out of other units in various ways—out of other sentences by connectives, as in "$p \supset q$ and out of names by verbs, as in "φx," "φxy," and "$\varphi x \supset \psi x$" (which may be conceived of as constructed out of the subsentences "φx" and "ψx" by the connective "$\supset$" or out of the name "x" by the complex verb "$\varphi \hat{x} \supset \psi \hat{x}$," "$\psi$'s-if-it-$\varphi$'s"). The rest of the hierarchy goes on from here—there are, for example, functors that form sentences out of verbs (that is, out of functors that form sentences out of names) and functors that form sentences out of these again, and so on ad infinitum. Functors may require one or more than one argument to make a sentence (the difference between "is a man" and "shaves," in the transitive sense), and when more arguments than one are required they may or may not be of the same type (for example, "If $\hat{x}$ is a man then $\hat{p}$ " requires a name and sentence).

Quantification. Of functors that form sentences from verbs, the most important are quantifications, such as "$(x)\hat{\varphi}x$" (which makes a sentence out of the verb whose place in the sentence is kept by "φ"), represented in English by such words as "everything." "Everything" is, or is constructed out of, the universal quantifier; there are many other quantifiers. Russell distinguished one other basic quantifier, "something." "Something is a man" expands in his language to "For some x, x is a man," or symbolically $(\exists x)$ (x is a man)."

Given the quantifier "something" and negation we can construct the complex "It is not the case that (for some x (x is a man))" or "For no x is x a man," Here we have the philosophical beginnings of the number series. The number 0 makes its appearance as part of a quantification, for we could write the preceding form as $(0x)(x$ is a man)." And the series can be continued. "Some" means "*At least* one, and "*At most* one thing is a man" is "For some, x, if anything is human it is identical with x"—that is, "For some x: for any y, if y is human y is identical with x." The combination of "At least one thing is a man" with "At most one thing is a man" gives us "Exactly one thing is a man" that is, "$(1x)(x$ is a man)." "$(2x)(x$ is a man)" is, similarly, "At least two things are men, and at most two things are men"—that is, "(For some x and for some y, x is a man, y is a man, and y is not identical with x) and (for some x and for some y: for any z, if z is a man z is either identical with x or identical with y)," Apparent occurrences of numbers as objects can be analyzed away in terms of this primary sense; "1 and 1 is 2" for instance, becomes "For any φ and for any ψ, if exactly one thing φ's, exactly one thing ψ's, and nothing does both, then exactly two things either-φ-or-ψ." Numbers are inseparable components of functors of functors of names, or, as Russell would say, functions of functions, but the naturalness of this analysis is disguised in his own work by the fact that before he brings arithmetic into the picture he introduces the language of classes and defines numbers in terms of classes. (The notation "$(0x)\varphi x$," etc., is not Russell's.)

DESCRIPTIONS. Before going on to Russell's discussion of classes, we should note that "$(x)\hat{\varphi}x$," "$(\exists x)\hat{\varphi}x$," and also "$(0x)\hat{\varphi}x$," "$(2x)\hat{\varphi}x$, and so on, are *functions* of functions of names, not *arguments* of such functions—that is, they are not names. "Something," "nothing," exactly one thing," etc., are not names, although, like names, they go with verbs to make sentences. They go, so to speak, on the other side of verbs: They "govern" the verbs; the verbs do not govern them. And although Russell's hierarchy of types of functors or "functions" provides innumerable ways of constructing sentences (and so of constructing functions), it provides no way of constructing genuine names. It is of the essence of the expressions represented by Russell's variables of lowest type (x, y, z, etc.)—that is, individual names—that they are logically structureless; they pick out individuals, and that is all. But in common speech and in mathematics we do seem to construct names, or at least ways of designating objects, out of expressions of other types: For example, "the man who broke the bank at Monte Carlo" seems to function as a name, yet it seems to be constructed from the verb "broke the bank at Monte Carlo." On Russell's view this appearance is illusory, and sentences in which such apparent names occur can always be replaced by paraphrases expressed entirely in Russell's language of structureless names, functions of functions, etc. However, he regarded it as useful for logical symbolism to reproduce at this point, although with greater precision, some of the devices of common speech and to have, as it were, a secondary language imposed on the primary one.

(Some account of Russell's handling of descriptions—that is, expressions of the form "The so-and-so"—and of other points raised below is given in the entry Existence.) "There φ-er," or "The thing that φ's," when it occurs as the apparent subject of a further verb—that is, in a context of the form "The φ-er ψ's"—is in reality a functor, in some way like a quantifying expression, of which the verbs "φ's" and "ψ's" are arguments; in fact "The thing that φ's ψ's" amounts precisely to "Exactly one thing φ's, and whatever φ's ψ's," whose first component has been analyzed above and whose second component is a simple formal implication. Expressions of this kind are especially important in mathematics when the contained functor $\varphi$is relational in form, as in "The φ-er of y"—that is, "The thing that φ's y," "The square root of y" (the num-

ber that yields y when multiplied by itself) is such an expression. Russell called expressions of this kind "descriptive Functions." They include most "functions" in the ordinary mathematical sense. It is a little inaccurate, of course, to use name symbols like "y" for numbers, which on Russell's view are not genuine individuals, but once the devices that yield class language and number language have been worked out, Russell's analysis of descriptive functions can be reproduced at the new level in a transposed form. This language of classes and numbers, to which we shall now turn, is itself a case of a secondary language containing apparent names (like "The class of persons that shave themselves") that disappear from the primary-language paraphrase.

CLASSES, FUNCTIONS, AND PROPERTIES. Russell represented the form "The class of things that φ" as "$\hat{x}(\varphi x)$"—usually read as "the x's such that φx—and represented "y is a member of the class of φ-ers" as "$y \in \hat{x}(\varphi x)$." Alternatively we may read "$\hat{x}(\varphi x)$" simply as "φ-er" and "$y \in \hat{x}(\varphi x)$" as "y is a φ-er." The expression "y is a φ-er" is true if and only if y φ's. One can in fact simply define "$y \in \hat{x}(\varphi x)$" as "φy." Given this definition, other concepts associated with class theory are easily introduced. For example, as noted earlier, "The class of φ-ers is included in the class of ψ-ers" amounts to the formal implication "For any x, if x is a φ-er then x is a ψ-er."

Classes of classes are related to functions of functions as classes are related to functions. To say that a given class—$\hat{x}(\varphi x)$, for example—is a member of the class of two-membered classes (or, as Russell would write it, "$\hat{x}(\varphi x) \in 2$") is just to say that exactly two things φ—i.e., the class of classes that Russell identifies with the number 2 is just the correlate in the class hierarchy of the function of functions $(2x)\hat{\varphi}x$.

Counting classes. There are two difficulties in Russell's views concerning classes. One is that classes, and, for that matter, numbers, can themselves be counted, as can individuals, but a number of classes would have to be not a class of classes but a class of classes of classes, and a number of numbers would similarly have to be a class of classes of classes of classes. This means that when we say "The number of numbers between 2 and 5 is 2," the first "2" has a sense quite different (belongs to a place quite different in the type hierarchy) from the second; and this seems a little implausible. Russell at this point is content to speak of the "systematic ambiguity"" of the key expressions of his symbolic language. Given the proof of "1 + 1 =2," for instance, considered as a statement about numbers of individuals, an analogous proof can always be constructed for the analogous statements about numbers of classes, numbers of numbers, etc., so that in practice it does not matter at which place in the type hierarchy we are working, provided we keep the types going up in order.

Ludwik Borkowski has suggested what may be a better solution: Suppose we always express quantification by a sign followed by a variable; for Russell's "(x)" we might put "$(\forall x)$," by analogy with "$(\exists x)$." We might then use the term *quantifier* not for this expression as a whole but for the initial sign, which can then be described as a functor that constructs a sentence out of a variable followed by a sentence, usually an "open" sentence in which the variable just mentioned occurs. We might then say that the initial sign "$\forall$" or " $\exists$"—or in the case of numerical quantifiers "0" or "1" or "2," etc.—is of the same logical type whatever the type of the variable that comes between it and the sentence following it. For counting properties (and, therefore, classes), we would have prefixes like "(2φ)"—for example, "$(2\varphi)\varphi$(Socrates)" would mean "Socrates has exactly two properties" or, better, "Exactly two things are true of Socrates"; and "(2φ)" is different from "$(2x)$," but the "2" is exactly the same in both contexts.

Counting functions. The other difficulty in Russell's theory is that classes dissolve into functions, but we do not count classes and functions in quite the same way. We would say, for example, that any two-membered class has four subclasses, in the sense that there are four ways of selecting members from such a class (both members, the first only, the second only, and neither). The corresponding theorem about functions would seem to be this: If exactly two things φ, then for exactly four ψ's, whatever ψ's φ's. But in fact there will always be vastly more than four ψ meeting this condition. Suppose, for example, that there are just two men in a room—i.e., $(2x)$(x is a man in the room)—and that one of them wears spectacles, spats, spotted socks, a red tie, and striped trousers; this much alone gives us five ψ's (namely, "______ is a man in the room wearing spectacles," "______ is a man in the room wearing spats," etc.), such that whatever ψ's is a man in the room. The key point here is simply that we count classes as being the same when they have the same members, but we do not count propositional functions as being the same merely because they are satisfied by the same arguments, and all the numerical concepts that are built up from the concept of identity must be similarly adjusted. For instance: "At most one class is a sub-class of $\hat{x}(\varphi x)$" does not mean "For some ψ: for any χ, if $\hat{x}(\chi x)$ is a subclass of $\hat{x}(\varphi x)$, then χ-ing is the same as ψ-ing," but rather it means "For some ψ: for any χ, if $\hat{x}(\chi x)$ is a sub-

class of $\hat{x}(\varphi x)$, then whatever χ's ψ's and whatever ψ's χ's." It is the same when we move up a type and count numbers themselves. If we write "$(0x)(\varphi x)$" for "It is not the case that (for some x, φx)" and "$(0'x)(\varphi x)$" for "For any x, if x φs then x is not identical with itself," we may say that these are different functions of functions—but whatever function either of them applies to the other applies to also; thus, they determine a single class of classes and a single "number," 0. The class and number language that Russell superimposes on his basic one is such that this is the way these quasi entities are counted.

Extensionality. One very radical way of simplifying this whole problem (one that Russell has considered from time to time) is to say that functions (properties, relations, etc.) are to be counted in just the same way that classes are; that is, that if $\varphi\hat{x}$ and $\psi\hat{x}$ characterize precisely the same objects (are formally equivalent), they are the same function. This is called the principle or law of extensionality; it in effect simply identifies a function with its "extension"—that is, with the class that it determines. The objection to this principle is simply its extreme implausibility in particular cases. For example, it seems obvious that even when two individuals and these two only are the men in a certain room wearing spats and the men in that room wearing spectacles, being a man in the room with spats is something different from being a man in it with spectacles.

Quine's criticism. Logicians such as W. V. Quine, following Ernst Zermelo and John von Neumann, have developed systems in which classes, classes of classes, and so on, are treated not as logical constructions but as genuine objects, and Russell's paradox is dealt with not by saying that "x is (is not) a member of x" is meaningless but by denying that "$x\varphi$'s" always implies that x is a member of the class of φ-ers. This account runs into difficulty when we try to handle certain nonmathematical properties of these supposed objects. Russell's view seems to have the advantage of not unnecessarily "multiplying entities," but Quine argues that Russell succeeds in dispensing with classes only by making genuine objects of properties or functions. This is said on the ground that in the course of his treatment of classes and numbers Russell is compelled to quantify over predicate variables—that is, to employ such quantifiers as "$(\exists\varphi)$" (for example, in defining "Exactly as many things ψ as χ as "*For some relation,* φ whatever φ's anything ψ's and vice versa, whatever is φ'd by anything χ's and vice versa, and whatever φ's or is φ'd by anything φ's or is φ'd that thing only"). This, Quine says, is to make properties and relations (like φ-ing) the "values of bound variables," and to do this is to treat them as existing.

This amounts to saying that to generalize an expression by quantifying over it is ipso facto to make it a name of an object; but this claim may be contested. We do not elucidate "He must have killed him somehow" by translating it "There must be some way in which he killed him" (which, taken literally, suggests that there are objects called "ways") but rather vice versa: We understand "somehow" directly as a generalization of qualifications like "with a knife," and the "way" line of talk is merely a variant of this. Even "something" is often to be understood as a generalized adjective rather than as a generalized individual name—for example, when I say "I am something that Jones is not—logical." It seems more plausible to interpret "I *have* something that Jones has not—logicality" as a verbal variant of the preceding sentence than to say that the latter alone brings out what I am really doing. And the logical rules for such higher-order quantifications are simple—we proceed from the specific case to the generalization, from "I am logical and Jones is not" to "For some A, I am A and Jones is not A," exactly as we do from "I am logical but not intelligent" to "For some individual x, x is logical but not intelligent."

Elimination of abstract terms. Russell might more plausibly be said to "hypostatize" or "reify" abstractions on the ground that there are some contexts from which it seems impossible to eliminate from his basic language his symbols for "abstracts," that is, φx, etc. This part of his system is developed more tidily in Alonzo Church's calculus of λ-conversion, in which the property of φ-ing is represented not by "$\varphi\hat{x}$ but by "$\lambda x\varphi x$," and of "ψ-ing if one φ's" by "$\lambda x\,.\,\varphi x \supset \psi x$." The basic rule of this calculus is that the application of $\lambda x\varphi x$ to an object a, symbolized by $(\lambda x\varphi x)a$, is equivalent to the plain φa, and similarly $(\lambda x\,.\,\varphi x \supset \psi x)a$ is equivalent to $\varphi a \supset \psi a$, And where we have a function of functions f, we can in general similarly replace $f(\lambda x\varphi x)$ by $f(\varphi)$—but not always. For instance, it is an obvious law that any such function f which holds for any φ whatever will hold for χ-ing-if-one-ψ's, as in formula F:

$$(f) : (\varphi)\ f(\varphi)\ .\supset .\ f(\lambda x\,.\,\psi x \supset \chi x.$$

Here the expression with λ seems uneliminable. We cannot replace it with "$\psi \supset \chi$," for this is meaningless—the hook joins sentence forms, not predicate forms. Where we have a specific f the elimination is again possible; for example, if f is the function "applying to exactly two objects," then $f(\lambda x\,.\,\psi x \supset \chi x)$ will amount to $(2y) : (\lambda x\,.\,\psi x \supset \chi x)y$ and thus to $(2y)(\psi y \supset \chi y)$. But where the f

itself is a variable, as it is in formula *F,* nothing of this sort is done. We could indeed (following Stanisław Leśniewski) introduce a symbol for the predicate "χ-ing-if-one-φ's" by a special definition; for example

$$[\supset \psi\chi]x =_{\text{Df}} \varphi x \supset \psi x$$

and so replace *F* with *G*:

$$(f) : (\varphi)f(\varphi) \,.\supset .\, f([\supset \psi\chi]),$$

but then it would be impossible to eliminate the defined symbol from *G* in favor of the symbols by which it is defined, and it seems an odd sort of definition that would be thus limited. (Church's use of λ can in fact be regarded as simply a generalization of Leśniewski's procedure.)

The uneliminability of "abstracts" from these contexts is an odd and perhaps awkward fact, but it need not be taken to imply that there are abstract objects, for "abstracts" need not be regarded as a kind of name. In expositions of the λ-calculus it is often said that the form $\lambda x \varphi x$ corresponds to the ordinary-language quasi noun "φ-ing," but this is not strictly correct, as may be seen from the fundamental equation $(\lambda x \varphi x)a = \varphi$" If "$\varphi$" here represents not a name but a verb ("φa" means "a φ's"), then so must "$\lambda x \varphi x$" ("$(\lambda x \varphi x)a$" also means "a φ's"), so that if f in $f(\varphi)$ is a function with not names but predicates as arguments, so it must be in "$f(\lambda x \varphi x)$."

RAMIFIED THEORY OF TYPES. We may now describe the added feature that makes Russell's own presentation of his theory of types more complex than the presentation so far given here. Russell divides functions into types not only according to the types of argument that they take but also according to whether they do or do not involve an internal reference to all functions of (what appear to be) their own type. For example, the function $\hat{x}$ has all the qualities of a great general" has individual-name arguments, just as "$\hat{x}$ is brave" does, but unlike "$\hat{x}$ is brave" it has a "for all φ" within itself—it amounts to "For all φ, if whoever is a great general φ's, then $\hat{x}$ φ's." Russell therefore regards it as of a different type, or, as he often says, of a different order, from "$\hat{x}$ is brave." Functions that do not thus involve a reference to all functions of (what appear to be) their own type he calls "predicative" functions and symbolizes them by putting an exclamation mark or "shriek" after the symbol, as in "$\varphi!x$," Functions cannot in fact (on Russell's view) strictly contain references to all functions of *their own* type or order but references only to ones of orders below their own. A function of individuals, which contains a reference to all predicative functions of individuals, is not itself predicative and cannot be regarded as being among the functions to which it implicitly refers. Having all the properties of a great general, for example, is not itself a property of a great general, at least not in the same sense of "property"—it is a second-order property.

What this means in practice might be illustrated as follows: It seems that if there were no facts about *x*—that is, if for no φ, φx—then there would be at least one fact about *x,* namely the fact that there are no facts about it, and hence it cannot be that there are no facts about *x.* In symbols, from

(1) $$\psi x \supset (\exists\varphi)(\varphi x)$$

it seems possible to obtain

(2) $$\sim(\exists\varphi)(\varphi x) \,.\supset (\exists\varphi)(\varphi x$$

by letting $\psi\hat{x}$ in (1) be, in particular, $\sim(\exists\varphi)(\varphi\hat{x})$; and from (2) it follows by a kind of reductio ad absurdum that for any given x we have $(\exists\varphi)(\varphi x)$. But on Russell's view this proof will not do, for (1) ought to have been written

(3) $$\psi!x \supset (\exists\varphi)(\varphi!x)$$

and here $\sim(\exists\varphi)(\varphi!\hat{x})$, not being itself predicative, is not a permissible substitution for $\psi!\hat{x}$. It is worth noting, however, that our final conclusion, $(\exists\varphi)(\varphi!x)$, can be proved from (3) in a different way—by letting our $\psi!\hat{x}$ be $\chi!\hat{x} \supset \chi!\hat{x}$ ("$\hat{x}$ χ's-if-it-χ's"), which *is* predicative and is true of any *x,* so that what it implies must be true of any *x* also. (The new argument is as follows: There is always some fact about *x,* since at least it is a fact that *x* is red-if-it-is-red, square-if-it-is-square, etc.)

Axiom of reducibility. Russell lumps together all his type and order restrictions under the general head of avoiding "vicious circles," and the theory of types with the theory of orders worked into it is called the "ramified" theory of types. One trouble with it is that it vitiates certain essential arguments in the higher reaches of mathematics, and to save these Russell introduced an "axiom of reducibility," that to every function of any order there corresponds a predicative function that is formally equivalent to it—that is, which holds for exactly the same arguments as the given function. This means that any argument like our allegedly invalid proof of $(\exists\varphi)(\varphi x)$ above, where it is worth saving, can in principle be replaced by one like our second and valid one; the axiom of reducibility does not itself enable us to find this valid argument but entitles us to proceed as if we had it. It is, however, an intuitively dubious principle and can be dispensed with if we can content ourselves with the theory

of types in the "simple" form in which it has been stated in earlier sections.

Semantic paradoxes. It was pointed out by F. P. Ramsey that those paradoxes which Russell lists and which cannot be eliminated (as can, for example, the paradox of the class of all classes not members of themselves) by the "simple" theory of types always contain some implicitly or explicitly "semantic" feature; that is, they all have to do with the relation of language to what it is about and all involve conceptions like truth and meaning. A typical example is the paradox of the liar, of the man who says "What I am now saying is false" and says nothing else but this, so that what he says is true if it is false and false if it is true. Such paradoxes are now generally dealt with by assuming not only a hierarchy of "parts of speech" in one's basic language (this is what the simple theory of types amounts to) but also a hierarchy of languages—a basic language, a "metalanguage" in which we discuss the meaning and truth of expressions in the basic language, a "metametalanguage" in which we deal similarly with the metalanguage, and so on.

It is both easy and necessary to criticize Russell's theories concerning the logical and semantic paradoxes, and his work in logic and the foundations of mathematics generally, but he remains, more than any other one person, the founder of modern logic.

ETHICS AND THE CRITIQUE OF RELIGION

ETHICS. Much of Russell's life, as we saw in an earlier section, was devoted to the advocacy of certain moral and political ideals. In this sense of the word *moralist,* in which it has no derogatory implications, Russell was certainly a moralist and frequently a very passionate one at that. Unlike many other moralists he was also concerned with what are now referred to as "metamoral" or "metaethical" issues. He repeatedly addressed himself to questions about the status of moral principles—what, if anything, they mean, what kind of disagreement there is between people who support opposite moral positions, and whether inferences from nonmoral premises to a moral conclusion can ever be valid. In discussing Russell's ethics, we will be concerned only with his metamoral theories.

Early views. In his first important essay on this subject, "The Elements of Ethics" (1910), Russell defended a position closely akin to that of G. E. Moore in *Principia Ethica.* "*Good* and *bad,*" he wrote, "are qualities which belong to objects independently of our opinions, just as much as *round* and *square* do; and when two people differ as to whether a thing is good, only one of them can be right, though it may be very hard to know which is right." The goodness or badness of a thing cannot be inferred from *any* of its other properties. "Knowledge as to what things exist, have existed, or will exist, can throw absolutely no light upon the question as to what things are good." Russell was by no means unaware at this time of the wide appeal of the familiar arguments for subjectivism—the "divergence of opinion" on moral questions and the difficulty of "finding arguments to persuade people who differ from us in such a question" ("The Elements of Ethics," in *Readings in Ethical Theory,* edited by Wilfrid Sellars and John Hospers, New York, 1952, pp. 6–7). But he did not then regard these arguments as having any logical force. "Difficulty in discovering the truth," he wrote, "does not prove that there is no truth to be discovered" (p. 6). Like Moore, he argued that if subjectivism were true it would follow that in a moral dispute there is never really any "difference of opinion" between the disputing parties. If when A says x is good and B says x is bad, A and B were really talking about their respective feelings or desires, they might well both be right at the same time and "there would be no subject of debate between them." At that time Russell regarded this as plainly false. "As a matter of fact," he observed, "we consider some tastes better than others: we do not hold merely that some tastes are ours and other tastes are other people's." When "The Elements of Ethics" was reprinted in 1952 in *Readings in Ethical Theory,* the anthology mentioned above, Russell added a footnote in which he explained that "not long after publishing this paper [he] came to disagree with the theory that it advocates." He explains that the change in his views was originally due to George Santayana's criticisms in his *Winds of Doctrine,* but he adds that he "found confirmation" for his later position "in many other directions." Russell's later position was first mentioned very briefly in a 1921 preface to a paperback reprint of "A Free Man's Worship"; it was explained in some detail in *What I Believe* (1925) and in *The Outline of Philosophy* (1927), and it received its fullest formulations in *Religion and Science* (1935), *Power* (1938), "Reply to My Critics" (in P. A. Schilpp, ed., *The Philosophy of Bertrand Russell,* 1944), and *Human Society in Ethics and Politics* (1955).

The subjectivity of values. Except on one basic issue, Russell's later position is a point-by-point denial of the earlier theory. "Good" and "bad" are no longer regarded as qualities belonging to objects, and in this respect they are now explicitly contrasted with "square" and "sweet": "If two men differ about values, there is not a disagree-

ment as to any kind of truth, but a difference of taste" (*Religion and Science,* pp. 237–238); "There are no facts of ethics" (*Power,* p. 257); "I see no property analogous to truth that belongs or does not belong to an ethical judgment" ("Reply to My Critics," p. 723). "Taste" in the first of these passages is used in a very broad sense to cover all kinds of psychological states and attitudes, including desires. Russell does not, of course, deny the plain fact that people regard some tastes as better than others and some desires as higher than other desires, but now he is willing to maintain that this merely means that the tastes or desires are their own. "What we 'ought' to desire is merely what someone else wishes us to desire" (*What I Believe,* p. 29).

Russell is quite ready to have his later theory classified as a form of "the doctrine of the subjectivity of values" (*Religion and Science,* p. 237), but it differs in some significant respects from the older theories that have gone by that name. If somebody maintains that pleasure, for example, or the love of God, is intrinsically good, or good "on its own account," this must not be taken to be equivalent to the statement that he approves of it or in some way desires it. Like the advocates of the so-called emotive theory of ethics, Russell maintains that intrinsic moral judgments, grammatical appearances notwithstanding, are not statements or assertions at all but *expressions* of desire. "A judgment of intrinsic value," he writes in *Power,* "is to be interpreted, not as an assertion, but as an expression of desire concerning the desires of mankind. When I say 'hatred is bad,' I am really saying: 'would that no one felt hatred.' I make no assertion; I merely express a certain type of wish" (*Power,* p. 257).

Both here and in his capacity as a reformer Russell places much emphasis on the distinction between purely personal and what he calls "impersonal" desires. A hungry man's desire for food or an ambitious man's desire for fame are examples of the former; a desire for the abolition of the death penalty or the end of racial discrimination, independently of whether the person in question stands to gain from these changes, are examples of the latter. In moral judgments we express certain of our impersonal desires. A king who says, "Monarchy is better than republican forms of government," is using the word *better* in its properly moral sense if he is expressing not just his desire to remain king but a desire that nations have monarchical systems regardless of his own personal position. Russell occasionally writes as if the desire expressed by moral judgments must be a second-order desire—that is, a desire that everybody have a certain first-order desire—but as several of his own examples make clear, this is not part of his position. What is essential is that the desire be impersonal. In this connection he also observes that the philosophers who stressed the "universality" of moral principles were in a sense quite right. This universality, however, does not consist in any a priori character or logical necessity. What is universal is the *object of the desire* expressed by a moral judgment. "The wish, as an occurrence, is personal, but what it desires is universal.… It is this curious interlocking of the particular and the universal which has caused so much confusion in ethics" (*Religion and Science,* p. 236).

As we shall see, Russell had a tendency to overestimate the scope of application of his subjectivism, but in a number of places he points out quite explicitly that large classes of everyday moral judgments and disputes do not come within the purview of the theory. "Ethical controversies are very often as to means, not ends" (*Power,* p. 259). "The framing of moral rules, so long as the ultimate Good is supposed known, is matter for science" (*Religion and Science,* p. 228). It follows from this that if human beings could agree about ultimate ends, all moral disputes would in principle be decidable by an appeal to facts even though the intrinsic judgments would still be not bona fide propositions but expressions of wishes. In fact, however, Russell insists, there is no such agreement about ends. In "The Elements of Ethics" he had conceded that there were *some* ultimate ethical differences but had maintained that people in fact "differ very little in their judgments of intrinsic value." Many of the commonly observed differences are wrongly regarded as ultimate because what are really disagreements about means are mistaken for disagreements about ends. In his subjectivist phase Russell seems to think that differences about ends are not at all uncommon. Behind such disputes as, for example, the subjection of women or the persecution of religious minorities, which do involve questions of means, he writes, "there is generally a difference as to ends," and this sometimes becomes "nakedly apparent," as in Friedrich Nietzsche's criticisms of Christian ethics. In Christianity, all men are valued equally, but for Nietzsche the majority exists only as means to the superman. This, Russell maintains, is an example of a dispute about ends, and "it cannot be conducted, like scientific controversies, by appeals to facts" (*Power,* p. 259).

In "The Elements of Ethics" Russell had quite properly observed that the mere existence of widespread ethical disagreement (if it is indeed widespread) does not establish any form of subjectivism. Although he has evidently come to believe that ethical disagreement is more widespread than he had thought earlier, he does not offer

this as evidence for his new theory. What he does offer as evidence is the undecid*ability* of ethical disputes. He writes:

> [The chief ground for adopting this view] is the complete impossibility of finding any arguments to prove that this or that has intrinsic value.… We cannot *prove,* to a color-blind man, that grass is green and not red. But there are various ways of proving to him that he lacks a power of discrimination which most men possess, whereas in the case of values there are no such ways … since no way can be even imagined for deciding a difference as to values, the conclusion is forced upon us that the difference is one of taste, not one as to any objective truth. (*Religion and Science,* p. 238)

If three men argue, one saying "The good is pleasure," the second "The good is pleasure for Aryans and pain for Jews," and the third "The good is to praise God and glorify him forever," they cannot, as people engaged in a scientific dispute, "appeal to facts," for facts, it seems obvious, "are not relevant to the dispute" (*Power,* p. 257).

Russell's later view agrees with the earlier position on only one significant point, its opposition to naturalism. By "naturalism" is here meant the theory that there *is* a logical connection between some moral judgments and factual premises where the latter are not necessarily confined to empirical statements but may also include metaphysical doctrines. We saw how in "The Elements of Ethics" Russell had insisted that from statements concerning what exists nothing can be inferred about "the goodness of anything." "It is logically impossible," he repeated in the course of expounding his later position, "that there should be evidence for or against" a moral judgment, but now this is maintained because a moral judgment "makes no assertion" and hence possesses neither truth nor falsehood (*Religion and Science,* pp. 236–237).

"Incredibility" of Russell's subjectivism. Rather than attempt a detailed critical evaluation of Russell's subjectivism, we will discuss one objection that has been urged by a number of his critics and which, in one form or another, has been leveled against nearly all forms of subjectivism. It has been argued that a subjectivist cannot consistently make moral judgments. All he can say is that some people have one kind of feeling or attitude while other people feel differently. More specifically, how can Russell's subjectivism be reconciled with his judgments as a moral critic and reformer?

It may be replied that as a matter of pure logic there is no inconsistency between holding that moral judgments are expressions of taste and using moral language to express one's own tastes. Russell, it might be said, would be inconsistent only if he claimed that *his* moral judgments, unlike those of his opponents, are more than expressions of taste. Then he would indeed be like the man who, in the course of an argument about the value of a piece of music, remarked to his opponent "It is all a matter of taste, except that my taste is better than yours." However, while this answer is valid as far as it goes, it does not meet the heart of the objection. For Russell seems to be saying—or at least he would like to be able to say—that his moral judgments (for example, his judgment that democracy is a better system than totalitarianism or that the sexual code advocated in *Marriage and Morals* is superior to that associated with orthodox religion) are in some sense rational or right or well-grounded while the judgments of his opponents are irrational, wrong, or unsupported by the evidence.

Russell apparently did not, when he first advanced his subjectivism, see any serious problem here, but in the 1940s and 1950s he repeatedly expressed dissatisfaction with his own theory on this ground. Thus, in "Reply to My Critics" he writes:

> What are "good" desires? Are they anything more than desires that you share? Certainly there *seems* to be something more. Suppose, for example, that some one were to advocate the introduction of bull-fighting in this country. In opposing the proposal, I should *feel,* not only that I was expressing my desires, but that my desires in the matter are *right,* whatever that may mean. As a matter of argument, I can, I think, show that I am not guilty of logical inconsistency in holding to the above interpretation of ethics and at the same time expressing strong ethical preferences. But in feeling I am not satisfied. (*The Philosophy of Bertrand Russell,* p. 724)

To this he adds: "I can only say that, while my own opinions as to ethics do not satisfy me, other people's satisfy me still less." More than a decade later Russell expressed himself even more strongly. In a letter to the *Observer* (October 6, 1957) he comments on Philip Toynbee's review of *Why I Am Not a Christian*: "What Mr. Toynbee says in criticism of my views on ethics has my entire sympathy. I find my own views argumentatively irrefutable, but nevertheless incredible. I do not know the solution."

It is doubtful whether in such comments Russell is really fair to his own subjectivism. Let us recall that the

theory was never meant to apply to anything other than what are variously called intrinsic or fundamental value judgments and differences. The questions whether happiness is better than unhappiness and love better than hate are frequently cited as such ultimate moral issues, but it would be hard to find anybody who seriously maintains that suffering is good on its own account or that hate is better than love, although of course people have often maintained that in certain situations and for certain reasons suffering and hate are preferable to enjoyment and love. However, on occasions there do appear to be real value differences of an ultimate kind. Thus, some people would maintain that dignity is "more important" or "nobler" than happiness. Many who do not despise happiness at all would maintain without hesitation that a man who chose to suffer a great deal rather than compromise his integrity (where it is assumed that he would in fact have suffered much less if he had not stood his ground) lived a better life than he would have if he had made the opposite choice. Or, again, there is sometimes disagreement as to whether a person suffering from a fatal illness should be told the truth, although there may be full agreement about the consequences of both telling and not telling him the truth. Russell's subjectivism does apply to this kind of intrinsic moral disagreement, and in such situations he could not, consistently with his theory, claim that the moral judgment he endorses is "more rational" or better supported than that of his opponents.

However, the examples Russell offers when expressing dissatisfaction with his subjectivism are not at all of this ultimate kind, and this applies to all or nearly all the positions he has advocated in his social and political writings. The man who says that the good is pleasure for Aryans and pain for Jews, if he is willing to engage in moral argument at all—if he is not, the problem does not arise—presumably does not *just* say this but proceeds to make all kinds of factual claims about the psychological and physical qualities of Aryans and Jews, respectively, about the laws of heredity, and about various other matters that he regards as justifying his moral position. Similarly, the man who maintains that "the good is to praise God and glorify him forever" presupposes that there is a God, and a God of a certain kind, probably also that he has revealed himself in certain ways, and, if challenged (or perhaps even without being challenged), he will make claims about the hollowness of all earthly satisfactions and the greater reliability, intensity, and duration of the satisfactions derived from glorifying God. Again, a man, who advocates the introduction of bullfighting into the United States would not *just* advance this proposal but would give reasons having to do, perhaps, with the benefits to be derived from engaging in dangerous sports and the special thrills experienced by the spectators. All these supporting factual claims are discussable, and it may be possible to show that they are mistaken or highly implausible. If so, it might well be possible to regard the case of one side in such a dispute as well supported and the other as unsupported by the evidence. In all cases in which the person is willing to support his moral judgment by factual premises, it is perfectly consistent for Russell to assert that one position is "more rational" than the other, where "more rational" does not merely mean that Russell shares the attitude of the person taking this position.

What seems to be amiss here is not Russell's subjectivism but his view (which is not logically implied by it) that the theory applies to cases like the dispute about bullfighting. In his later period Russell seems to be guilty of a gross overestimate of the prevalence of ultimate moral disagreements. It is true, as he observes in *Power,* that behind disagreements about means there is frequently disagreement about ends, but it is very doubtful that the ends in question are in most cases *ultimate* ends. To give a simple illustration of a very common type: Two people may offer conflicting moral judgments about a bill to legalize abortion. The man who opposes the legislation may give as his reason (or as one of his reasons) that it would remove one of the conditions restraining unmarried people from engaging in sexual intercourse, whereas the other man might offer this as his reason (or one of his reasons) for supporting the legislation. Although the disagreement may in the immediate context be properly described as one about an end, it is clearly not about an ultimate end. In all likelihood the parties to the dispute would differ about the effects of a freer sex life on personal happiness, on society at large, on the future of religious institutions, and many other things. It is doubtful that either of them would maintain that suffering as such is better than happiness or that hate is better than love.

Even people who advocate what by most contemporary standards would be regarded as "outlandish" moral positions can usually be seen to share many of the intrinsic value judgments of the rest of humankind. Thus, Arthur Schopenhauer and other champions of asceticism recommend the suppression of desires, including those that to most human beings seem the most natural and the most innocent, but they do so *not* because in their opinion suppression of these desires would make people unhappy but, on the contrary, because it would enable them to achieve greater happiness or at least because it would reduce suffering to a minimum. In Norman Mailer's bizarre novel *An American Dream* the main char-

acter offers a defense of murder, but this unusual position is justified by the argument that "murder offers the promise of vast relief. It is never unsexual." It is accompanied by "exhilaration" that must come "from possessing such strength." It should be noted that murder is here justified not because it causes suffering but because, according to the character, it leads to "exhilaration." In other writings Mailer tells us that the "modern soul marooned in … emptiness, boredom and a flat, dull terror of death" would be well advised to pass through "violence, cannibalism, insanity, perversion" and other states and activities that are usually considered highly undesirable, but these recommendations are offered not for their own sake but because they will lead the person "back to life."

As for the really intrinsic clashes of the kind mentioned earlier, to which Russell's subjectivism would apply, one wonders if the consequences of the theory are there really so paradoxical. No doubt people do in such disputes regard their position as superior to that of their opponents—the man who admires integrity will feel contempt for the "cowardly" compromiser, and the compromiser will think the man who chooses to suffer a fool. Here, however, unless there are some *hidden* differences concerning matters of fact, it seems not at all incredible to maintain that calling one position superior simply amounts to expressing one's own preference for it.

None of the above is meant to prove that Russell's subjectivism is a correct account of the logical status of moral judgments, but it would indicate that the favorite objection of his critics can be disposed of without much difficulty.

CRITIQUE OF RELIGION. No such doubts as Russell has expressed about his subjectivism in ethics mark his views on religion. Unlike many academic philosophers whose position is very similar to his, Russell did not hesitated to express his convictions publicly and without equivocation or compromise. Ever since he abandoned the Platonic theory of ideas, Russell was a forthright opponent of religion in more senses than one: He regards the basic doctrines of (supernaturalistic) religions as intellectually indefensible, he argues that religious belief has not on balance been a force for good but quite the opposite, and he hopes and believes that religion will eventually die out. "I am myself," he wrote in 1922, "a dissenter from all known religions, and I hope that every kind of religious belief will die out.… I regard religion as belonging to the infancy of human reason and to a stage of development which we are now outgrowing" (*Sceptical Essays,* p. 101). In a television interview thirty-seven years later he slightly qualified this prediction. If great wars and great oppressions continue so that many people will be leading very unhappy lives, religion will probably go on, but "if people solve their social problems religion will die out" (*Bertrand Russell Speaks His Mind,* p. 31).

God. Russell wavered between calling himself an agnostic and describing himself as an atheist. He evidently did not attach too much importance to this distinction, but he had made it clear that if he is to be classified as an agnostic, it would have to be in a sense in which an agnostic and an atheist are "for practical purposes, at one." In the television interview mentioned earlier the interviewer asked Russell, "Do you think it is certain that there is no such thing as God, or simply that it is just not proved?" "No," Russell answered, "I don't think it is certain that there is no such thing—I think that it is on exactly the same level as the Olympic gods, or the Norwegian gods; they also may exist, the gods of Olympus and Valhalla. I can't prove they don't, but I think the Christian God has no more likelihood than they had. I think they are a bare possibility" (*Bertrand Russell Speaks His Mind,* pp. 24–25). He explained his views more fully in an interview published in *Look* magazine in 1953. An agnostic, in any sense in which he can be regarded as one, Russell said, "may hold that the existence of God, though not impossible, is very improbable; he may even hold it so improbable that it is not worth considering in practice" (Leo Rosten, ed., *A Guide to the Religions of America,* New York, 1955, p. 150).

Immortality. On survival, Russell's position is similarly negative. All the evidence indicates that what we regard as our mental life is "bound up with brain structure and organized bodily energy." There is every reason to believe that mental life ceases when the body decays. Russell admits that this argument is "only one of probability" but adds that "it is as strong as those upon which most scientific conclusions are based" (*Why I Am Not a Christian,* p. 51). It is conceivable that evidence from psychical research might change the balance of probability some day, but, writing in 1925, Russell considered such evidence far weaker "than the physiological evidence on the other side." He did not later see any reason to modify this judgment.

Russell's views on the body-mind problem are known as "neutral monism," and it would be inaccurate to call him a materialist. However, he always emphasized that as a theory about man's place in the universe his philosophy is closely akin to materialism. "Emotionally," he wrote in 1928, "the world is pretty much the same as it would be if the materialists were in the right" (*In Praise of*

Idleness, p. 143). The opponents of materialism, he adds, have been actuated by the desire to prove that the mind is immortal and that the "ultimate power" in the universe is mental and not physical. On both these points, Russell makes clear, he agrees with materialism. When he returned to the subject in 1959 he had not changed his opinion at all. "I still think," he wrote then, "that man is cosmically unimportant, and that a Being, if there were one, who could view the universe impartially, without the bias of *here* and *now,* would hardly mention man, except perhaps in a footnote at the end of the volume" (*My Philosophical Development,* p. 213).

Objections to fideism. Although, needless to say, Russell rejected the traditional arguments for the existence of God and immortality, he greatly preferred the rationalistic theology of such philosophers as Thomas Aquinas and Descartes to the fideism of Blaise Pascal, Jean-Jacques Rousseau, Søren Kierkegaard, and their numerous modern followers. "The rejection of reason in favor of the heart," he writes, "was not, to my mind, an advance." He remarks that "no one thought of this device so long as reason appeared to be on the side of religious belief" (*A History of Western Philosophy,* p. 720). There are two fatal objections to the practice of justifying religious belief by an appeal to the emotions of the heart. To begin with, the heart says different things to different men and to the same man at different times, but even if the heart said the same thing to all men this would still not be evidence for the existence of anything outside our emotions, and the fideists, no less than the rationalistic believers, mean to make claims about objective fact, not merely about their own emotions. At bottom, Russell concludes, the only reason offered for the acceptance of the new theology is "that it allows us to indulge in pleasant dreams. This is an unworthy reason, and if I had to choose between Thomas Aquinas and Rousseau, I should unhesitatingly choose the Saint" (*My Philosophical Development,* p. 721).

Some unbelievers have gone out of their way to praise the greatness of Jesus and to admit that religious belief, although perhaps not true, is at least of great value to individual believers and to society. Russell makes no such concessions. Although he grants that some of Christ's maxims were indeed admirable (especially those consistently disregarded by Christian dignitaries) he finds much in the teachings of Jesus to be defective, in particular his doctrine of eternal damnation. "Either in the matter of virtue or in the matter of wisdom," Russell concludes, Christ does not "stand as high as some other people known to history"—for example, Buddha and Socrates (*Why I Am Not a Christian,* p. 19).

Harmfulness of religious belief. Russell's views about the nature of the emotions that inspire religious belief ("it is based, primarily and mainly, upon fear") and also about the harmful influence of religious organizations are very similar to those of David Hume, Baron d'Holbach, and other eighteenth-century freethinkers. He did, however, devote rather more attention to the bad effects of the habit of accepting propositions on faith—in the absence of or even in opposition to the evidence. It is an error, Russell contends, to suppose that a person who does not form his beliefs on the basis of evidence in one domain can remain open-minded and scientific in another. Furthermore, somebody holding comfortable beliefs on faith dimly realizes that they are myths and "becomes furious when they are disputed." Such a person will therefore do his best to suppress all critics who might remind him of the feeble backing of his beliefs. Russell makes it clear that in this context he is not criticizing Christianity only. "The important thing," he writes, "is not what you believe, but how you believe it." The objections to "faith" do not depend on what the faith in question may be. "You may believe in the verbal inspiration of the Bible or of the Koran or of Marx's *Capital.* Whichever of these beliefs you entertain, you have to close your mind against evidence; and if you close your mind against evidence in one respect, you will also do so in another, if the temptation is strong." The person who bases his belief on reason will support it by argument rather than by persecution and will abandon his position if the argument goes against him. If, however, his belief is based on faith, he will conclude that argument is useless and will "therefore resort to force either in the form of persecution or by stunting and distorting the minds of the young whenever he has the power to control their education" (*Human Society in Ethics and Politics,* pp. 207–208).

"The world is horrible." Russell never denied that in some respects a "godless" philosophy like his has to be gloomy. The beginning of wisdom, he teaches, is acceptance of the fact that the universe does not care about our aspirations and that happiness and unhappiness are not meted out in accordance with what people deserve. "The secret of happiness," he observed during a television program commemorating his ninety-second birthday, "is to face the fact that the world is horrible." What Russell meant by this becomes clear from a story related by his biographer, Alan Wood. Wood's wife had expressed her opinion that it seemed horribly unjust that the young men who had been killed in the war should not somehow or somewhere have a second chance to achieve happiness. "But the universe *is* unjust," Russell replied, "the secret of happiness is to face the fact that the world is horrible,

horrible, *horrible* … you must feel it deeply and not brush it aside … you must feel it right here"—hitting his breast—"and then you can start being happy again" (*Bertrand Russell: The Passionate Sceptic,* p. 237). Once a person has stopped looking at the universe in terms of anthropomorphic demands, he can concentrate on what is attainable and not waste his time in self-pity and cosmic complaints. For those whose philosophy is shaped not by a respect for facts but by their wishes Russell was always scathing in his contempt. He expressed his amazement that courage is praised in all types of situations but not when it comes to forming a view about the world. "Where traditional beliefs about the universe are concerned," he writes, "craven fears … are considered praiseworthy, while intellectual courage, unlike courage in battle, is regarded as unfeeling and materialistic." Writing in 1957, he notes that this attitude is perhaps less widespread than it was in his youth, but he adds that it "still inspires vast systems of thought which have their root in unworthy fears." "I cannot believe," he concludes, that there can ever be any good excuse for refusing to face the evidence in favor of something unwelcome. It is not by delusion, however exalted, that mankind can prosper, but only by unswerving courage in the pursuit of truth" (*Fact and Fiction,* p. 46).

See also Absolute, The; Asceticism; Analysis, Philosophical; Balfour, Arthur James; Bradley, Francis Herbert; Church, Alonzo; Correspondence Theory of Truth; Descartes, René; Epistemology, History of; Ethical Subjectivism; Existence; Frege, Gottlob; Hegel, Georg Wilhelm Friedrich; Hegelianism; Holbach, Paul-Henri Thiry, Baron d'; Hume, David; Infinity in Mathematics and Logic; James, William; Kierkegaard, Søren Aabye; Logical Paradoxes; Logic, History of; Logic, Modern; Logic, Traditional; Luther, Martin; Mathematics, Foundations of; McTaggart, John McTaggart Ellis; Memory; Metaethics; Mill, John Stuart; Mind-Body Problem; Modal Logic; Moore, George Edward; Neumann, John von; Nietzsche, Friedrich; Number; Pascal, Blaise; Peano, Giuseppe; Peirce, Charles Sanders; Plato; Platonism and the Platonic Tradition; Pluralism; Proper Names and Descriptions; Propositions; Quantifiers; Quine, Willard Van Orman; Ramsey, Frank Plumpton; Realism; Rousseau, Jean-Jacques; Santayana, George; Schopenhauer, Arthur; Sellars, Wilfrid; Socrates; Thomas Aquinas, St.; Types, Theory of; Voltaire, François-Marie Arouet de; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

BIOGRAPHY

There is a good deal of autobiographical material in Russell's *Portraits from Memory and Other Essays* (New York: Simon and Schuster, 1956); in *Fact and Fiction* (New York: Simon and Schuster, 1962); in his introduction to *Selected Papers of Bertrand Russell* (New York: Modern Library, 1927); in "My Religious Reminiscences," in *Rationalist Annual* 55 (1938): 3–8; in "My Mental Development," in *The Philosophy of Bertrand Russell,*edited by P. A. Schilpp (Evanston and Chicago: Open Court, 1944); and in *My Philosophical Development* (New York: Simon and Schuster, 1959). Alan Wood, *Bertrand Russell: The Passionate Sceptic* (London : Allen and Unwin, 1957), is the only full-length biographical study of Russell. H. W. Leggett, *Bertrand Russell* (New York, 1950), is a short pictorial biography.

G. H. Hardy, *Bertrand Russell and Trinity* (Cambridge, U.K.: Cambridge University Press, 1942), traces the controversy between Russell and the fellows of Trinity College over his pacifist activities during World War I. *Rex versus Bertrand Russell, Report of the Proceedings before the Lord Mayor* (London, 1916), gives the text of the first of Russell's trials.

D. H. Lawrence, *Letters to Bertrand Russell* (New York: Gotham Book Mart, 1948), reproduces Lawrence's letters to Russell during World War I; Russell's letters to Lawrence have not been preserved.

Russell's part in the Beacon Hill School is most fully described in Joe Park, *Bertrand Russell on Education* (Columbus: Ohio State University Press, 1963). The Park volume also contains a complete list of Russell's writings on educational topics. Details about the City College case of 1940 can be found in *The Bertrand Russell Case,* edited by John Dewey and Horace M. Kallen (New York: Viking Press, 1941); in a publication by the American Civil Liberties Union titled *The Story of the Bertrand Russell Case—The Enlightening Record of the Obstruction by Courts and Officials of the Appointment of Bertrand Russell to a Professorship at the College of the City of New York* (New York: American Civil Liberties Union, 1941); and in Paul Edwards, "How Bertrand Russell Was Prevented from Teaching at City College," which is an appendix to Russell's *Why I Am Not a Christian and Other Essays on Religion and Related Subjects* (London: Allen and Unwin, and New York: Simon and Schuster, 1957).

EPISTEMOLOGY AND METAPHYSICS

Principles of Mathematics (Cambridge, U.K.: Cambridge University Press, 1903) was Russell's first major philosophical work. Its position is one of Platonic realism. In the preface to the second edition (1937) Russell sets forth his later disenchantment with this position. For a nonmathematical exposition of Russell's early realism, see "Meinong's Theory of Complexes and Assumptions," in *Mind* 13 (1904): 204–219; 336–354; 509–524. Russell's criticisms of the idealist theory of truth are to be found in "The Monistic Theory of Truth," in *Philosophical Essays* (New York: Longman, 1910), a revised version of "The Nature of Truth," in *Mind* 15 (1906): 528–533. *Philosophical Essays* also contains two influential essays by Russell attacking the pragmatist theory of truth.

The shift from realism to logical constructionism can be followed in a number of articles, the most important of

which is "On Denoting," in *Mind* 14 (1905): 479–493. This, together with other important but otherwise largely unavailable essays, is reprinted in Russell's *Logic and Knowledge,* edited by R. C. Marsh (London: Allen and Unwin, 1956). Russell's "On the Relations of Universals and Particulars," in *PAS* 12 (1911–1912): 1–24, reprinted in *Logic and Knowledge,* is a classic presentation of the largely Platonic theory of universals Russell still held at that time. *Problems of Philosophy* (New York: Holt, 1912) gives an excellent semipopular account of the general state of Russell's thinking then. Russell's early attempts to represent physical objects as logical constructions can be seen in *Our Knowledge of the External World* (Chicago: Open Court, 1914) and in two essays, "The Ultimate Constituents of Matter," in *Monist* 25 (1915): 399–417, and "The Relations of Sense-Data to Physics," in *Scientia* (4) (1914), both reprinted in *Mysticism and Logic* (London: Allen and Unwin, 1918). Other important essays in this collection are "On Scientific Method in Philosophy" (1914); "On the Notion of Cause," originally published in *PAS* 13 (1912–1913): 1–26; and "Knowledge by Acquaintance and Knowledge by Description," originally published in *PAS* 11 (1910–1911): 108–128. See also "The Philosophy of Logical Atomism," in *Monist* 28 (1918): 495–527; 29 (1919): 32–63, 190–222, and 345–380; reprinted in *Logic and Knowledge* (see above). The analysis of basic concepts and principles of physical science is pushed further in *The Analysis of Matter* (New York: Harcourt Brace, 1927). Logical constructionism is applied to mental phenomena in *The Analysis of Mind* (New York: Macmillan, 1921). Russell's increasing concern with psychological aspects of meaning can be traced in "On Propositions, What They Are and How They Mean," in *PAS,* supp. 2 (1919): 1–43, reprinted in *Logic and Knowledge,* in Ch. 10 of *The Analysis of Mind*; and in Russell's most extensive work on meaning and empirical data, the rich but chaotic *An Inquiry into Meaning and Truth* (New York: Norton, 1940). Russell's later thoughts on meaning and various other problems concerning empirical knowledge, particularly in the physical sciences, are given a relatively systematic presentation in *Human Knowledge, Its Scope and Limits* (New York: Simon and Schuster, 1948).

In several works Russell summarized his philosophy and/or its development. The most important of these are "Logical Atomism," in *Contemporary British Philosophy,* edited by J. H. Muirhead, first series (London: Allen and Unwin, 1924), reprinted in *Logic and Knowledge* (see above); "My Mental Development," in *The Philosophy of Bertrand Russell,* edited by P. A. Schilpp (see above); and the very interesting recent work *My Philosophical Development* (New York: Simon and Schuster, 1959). The last-named work also contains some of Russell's polemics against Oxford philosophers and their criticisms of his views. Russell's *A History of Western Philosophy* (New York: Simon and Schuster, 1946) and *The Wisdom of the West* (New York: Doubleday, 1959), aside from their intrinsic interest, are of great value to students of Russell's thought in showing us his mature evaluations of the great philosophers of past ages.

The critical literature on different aspects of Russell's epistemology and metaphysics is vast. *The Philosophy of Bertrand Russell* (see above) contains a number of excellent discussions, together with Russell's replies. Special mention should also be made of C. A. Fritz, *Bertrand Russell's Construction of the External World* (London: Routledge & K. Paul, 1952); Erik Götlind, *Bertrand Russell's Theories of Causation* (Uppsala: Almqvist and Wiksells, 1952); J. O. Urmson, *Philosophical Analysis: Its Development between Two World Wars* (Oxford: Clarendon Press, 1956); and G. J. Warnock, *English Philosophy since 1900* (London: Oxford University Press, 1958). The books by Urmson and Warnock contain detailed appraisals of Russell's logical atomism. Russell's logical atomism as well as his neutral monism and his theories about truth and induction are sympathetically discussed by D. J. O'Connor in Ch. 26 of his *Critical History of Western Philosophy* (New York: Free Press of Glencoe, 1964). *Rivista critical di storia della filosofia* 8 (2) (1953): 101–335, and several articles in *Philosophy* 35 (January 1960): 1–50, are devoted to Russell's philosophy, including Anthony Quinton's useful sketch of the development of Russell's ideas in epistemology and metaphysics, "Russell's Philosophical Development," 1–13.

LOGIC AND MATHEMATICS

Of Russell's own works on logic and mathematics, see *Principles of Mathematics* (Cambridge, U.K.: Cambridge University Press, 1903; 2nd ed., London, 1937); *Principia Mathematica,* 3 vols., written with A. N. Whitehead (Cambridge, U.K.: Cambridge University Press, 1910–1913; 2nd ed., 1927); *Introduction to Mathematical Philosophy* (London: Allen and Unwin, 1919); and the papers "On Denoting" (1905), "Mathematical Logic as Based on the Theory of Types" (1908), "The Philosophy of Logical Atomism" (1918), and "Logical Atomism" (1924), all of which are reprinted in *Logic and Knowledge* (see above).

On Frege's parallel work, see his *Grundlagen der Arithmetik* (Breslau, 1884), translated by J. L. Austin as *The Foundations of Arithmetic* (Oxford: Blackwell, 1950); and P. T. Geach and Max Black, eds., *Translations from the Philosophical Writings of Gottlob Frege* (New York: Philosophical Library, 1952).

Important critical discussions of Russell's work occur in W. E. Johnson, *Logic,* Pt. II (Cambridge, U.K., 1922), Chs. 3 and 6; F. P. Ramsey, *The Foundations of Mathematics* (London, 1931), papers I and II; W. V. Quine, *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953), essays I, V, and VI; and G. E. Moore, *The Commonplace Book of G. E. Moore,* 1919–1953, edited by Casimir Lewy (New York: Humanities Press, 1963), Notebook II, item 4, and Notebook V, item 13.

On formal implication, see A. N. Prior, "The Theory of Implication," in *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9 (1963): 1–6. On simplifications of type theory, see Alonzo Church, "A Formulation of the Simple Theory of Types," in *Journal of Symbolic Logic* 5 (1940): 56–68, and Ludwik Borkowski, "Reduction of Arithmetic to Logic Based on the Theory of Types," in *Studio Logica* 8 (1958): 283–295.

ETHICS AND RELIGION

Russell's early views on ethics are in "The Elements of Ethics," Ch. 1 of *Philosophical Essays* (New York: Longman, 1910); it has been reprinted in *Readings in Ethical Theory,* edited by Wilfrid Sellars and John Hospers (New York: Appleton-Century-Crofts, 1952), pp. 1–34. The fullest statements of his later position are in Ch. 9 of *Religion and Science* (New York: Holt, 1935) and in *Human Society in Ethics and Politics* (New York: Simon and Schuster, 1955). There are critical

discussions of Russell's views in Lillian W. Aiken, *Bertrand Russell's Philosophy of Morals* (New York: Humanities Press, 1963); in Justus Buchler, "Russell and the Principles of Ethics," in *The Philosophy of Bertrand Russell* (see above); and in D. H. Monro, "Russell's Moral Theories," in *Philosophy* 35 (1960): 30–50.

Russell's earlier views on religion are in "The Essence of Religion," in *Hibbert Journal* 11 (1912): 46–62. His first published discussion of the arguments for the existence of God is contained in Ch. 15 of *A Critical Exposition of the Philosophy of Leibniz* (Cambridge, U.K.: Cambridge University Press, 1900; 2nd ed., London and New York, 1937). His later views are expounded in several of the essays in *Why I Am Not a Christian* (see above) and in Pt. II, Ch. 7, of *Human Society in Ethics and Politics* (see above). The BBC debate with Father F. C. Copleston (1948), "The Existence of God," is available in the British edition, but not in the American edition, of *Why I Am Not a Christian,* but it has been reprinted in *A Modern Introduction to Philosophy,* edited by Paul Edwards and Arthur Pap, 2nd ed. (New York, 1965), and in *The Existence of God,* edited by John Hick (New York: Macmillan, 1964). Several chapters in *The Scientific Outlook* (London and New York, 1931) and in *Religion and Science* (see above) contain criticisms of the attempts of certain scientists to derive theological conclusions from physics and biology. Russell's objections to the fideistic position are found in Ch. 12, Bk. 3, of *A History of Western Philosophy* (see above). His objections to William James's defense of religion are contained in Ch. 29, Bk. 3, of the same work and in Ch. 5 of *Philosophical Essays* (see above). Russell's views on religion are criticized in H. G. Wood, *Why Mr. Bertrand Russell Is Not a Christian* (London, 1928); C. H. D. Clark, *Christianity and Bertrand Russell* (London: Lutterworth Press, 1958); G. S. Montgomery, *Why Bertrand Russell Is Not a Christian* (New York, 1959); and E. S. Brightman's contribution to the Schilpp volume, "Russell's Philosophy of Religion," pp. 537–556.

SOCIAL AND POLITICAL THEORY

In addition to the works mentioned in the first section of the present entry, the following among Russell's books dealing with social and political questions have been influential: *Principles of Social Reconstruction* (London: Allen and Unwin, 1916); *Roads to Freedom: Socialism, Anarchism and Syndicalism* (London: Allen and Unwin, 1918); *The Problems of China* (New York: Century, 1922); *Power: A New Social Analysis* (London: Allen and Unwin, 1938); *Authority and the Individual* (London: Allen and Unwin, 1949); and *New Hopes for a Changing World* (London: Allen and Unwin, 1951). Ch. 17 of *New Hopes* contains a moving discussion of the problems of growing old and facing death. Russell's fullest discussion of Marxism can be found in *Freedom and Organization 1814–1914* (London: Allen and Unwin, 1934; as *Freedom versus Organization,* New York, 1934), which is in effect a history of the main social and intellectual forces of the nineteenth century.

OTHER WRITINGS

Philosophical discussions sooner or later crop up in most of Russell's writings. Some of his most delightful occasional pieces have been collected in *Sceptical Essays* (London: Allen and Unwin, 1927); in *In Praise of Idleness* (London: Allen and Unwin, 1935); and in *Unpopular Essays* (London: Allen and Unwin, 1950). The last of these contains his "Auto-obituary," which was first published in 1936. *Bertrand Russell Speaks His Mind* (London: Barker, 1960) is a most interesting volume containing the unedited text of a series of television interviews, dealing with a great variety of topics, which took place in the spring of 1959.

The Basic Writings of Bertrand Russell, 1903–1959, edited by R. E. Egner and L. E. Dennon (New York, 1961), is a very useful anthology of writings by Russell. The Schilpp volume contains an extremely comprehensive bibliography up to 1944.

RECENT WORKS

Coffa, Alberto. "The Humble Origins of Russell's Paradox." *Russell* 33 (1979): 31–37.

Eames, Elizabeth R. *Bertrand Russell's Theory of Knowledge*. London: George Allen and Unwin, 1969.

Griffin, Nicholas. *Russell's Idealist Apprenticeship*. Oxford: Oxford University Press, 1991.

Hylton, Peter W. *Russell, Idealism, and the Emergence of Analytic Philosophy*. Oxford: Oxford University Press, 1990.

Irvine, A. D. "Epistemic Logicism and Russell's Regressive Method." *Philosophical Studies* 55 (1989): 303–327.

Irvine, A. D., ed. *Bertrand Russell: Critical Assessments*. 4 vols. London: Routledge, 1999.

Jager, Ronald *The Development of Bertrand Russell's Philosophy*. London: George Allen and Unwin, 1972.

Kaplan, David. "What Is Russell's Theory of Descriptions?" (1970). In Pears (1972), 227–244.

Landini, Gregory. *Russell's Hidden Substitutional Theory*. New York and Oxford: Oxford University Press, 1998.

Linsky, Bernard. *Russell's Metaphysical Logic*. Stanford: CSLI, 1999.

Lycan, William. "Logical Atomism and Ontological Atoms." *Synthese* 46 (1981): 207–229.

Pears, David F. *Bertrand Russell and the British Tradition in Philosophy*. London: Collins, 1967.

Pears, David F., ed. *Bertrand Russell: A Collection of Critical Essays*. Garden City, NY: Anchor Books, 1972.

Putnam, Hilary. "The Thesis that Mathematics Is Logic." In Schoenman (1967), 273–303.

Ryan, Alan. *Bertrand Russell: A Political Life*. New York: Hill and Wang, 1988.

Sainsbury, R. M. *Russell*. London: Routledge, 1979.

Savage, C. Wade, and C. Anthony Anderson, eds. *Rereading Russell: Essays on Bertrand Russell's Metaphysics and Epistemology*. Minneapolis: University of Minnesota Press, 1989.

Schoenman, Ralph, ed. *Bertrand Russell: Philosopher of the Century*. London: George Allen and Unwin, 1967.

Schultz, Bart. "Bertrand Russell in Ethics and Politics." *Ethics* 102 (1992): 594–634.

Paul Edwards (1967)
(Life and Social Theories, Ethics and Critique of Religion)
William P. Alston (1967)
(Epistemology and Metaphysics)
A. N. Prior (1967)
(Logic and Mathematics)
Bibliography updated by Ernest Sosa (2005)

RUSSIAN PHILOSOPHY

In the broad sense the words "Russian philosophy" refer to all schools of philosophical thought pursued in Russia, regardless of differences among them. In the narrower sense the terms describe the religious-philosophical trend that flourished in the late nineteenth and early twentieth centuries. Both uses have value: The first embraces the variety of interests among Russian philosophers, whereas the second points to their most distinctive contribution to philosophy in general. But even on the broadest level, the common preoccupations that were typical of major Russian thinkers shaped the physiognomy of Russian philosophy as a whole.

Philosophy in Russia developed in a variety of forms. Philosophical ideas permeated religious, political, and literary debates throughout the country's history. For a long time they were not articulated in what counted as philosophical parlance in the West, largely because of unfavorable historical conditions. But when these conditions changed, as they did, for example, in the late nineteenth and especially in the early twentieth centuries, there emerged a vibrant philosophical scene. This flourishing had been prepared within Russian culture, among other things, by its religious, literary, and scientific thought. Thus it should not be surprising that some theologians, novelists, and scientists are relevant to the history of Russian philosophy.

TYPICAL FEATURES

For various reasons Russian philosophy has been dominated, not to say oppressed, by pragmatic concerns. Realistic or utopian, philosophical thought in Russia is expected to be engaged. It is not an accident that Marxism, for which social practice is the criterion of theoretical truth, has had such a firm grip on the Russian polity. Even when Russian philosophy did reach the heights of speculation—as in the thought of Vladimir Solov'ëv (1853–1900)—it still bore the mark of "theurgic restlessness," in Vasilii Zen'kovskii's (1881–1962) words—that is, the desire to transfigure life. Still, when conditions were right, and sometimes despite harshly adverse conditions, Russian thinkers have achieved reflexive insights of uncommon depth.

Closely related to this is Russian philosophy's realist *ontologism*; that is, the tendency to value the reality of *being* over and above the truths of abstract *understanding*. Nikolai Berdyaev (1874–1948) noted that the Russian mind strongly doubts whether the creation of culture is justified in the face of life's problems. This doubt was typical of Lev Tolstoy (1828–1910) who disparaged art in contrast with the peasant's work. Paradoxically, this tendency was also responsible for the seriousness with which Russians have treated the arts and philosophy. Likewise Russian thinkers often sought *justice* more eagerly than *truth* because the former seemed more tangible and urgent than the latter.

Many commentators have insisted that Russian philosophy is also inherently religious and personalistic. While the aggressively atheist and collectivist Soviet Marxism is an inescapable counterexample, it cannot be denied that the themes of religion and personhood have occupied and continue to occupy a prominent place in Russian philosophical discourse. Fëdor Dostoevsky's (1821–1881) persistent interest is only the more familiar, especially to the West, among many manifestations of these themes.

Russian thought has a marked predilection for viewing things holistically. Russian philosophers have often been preoccupied with global, wide-ranging problems and visions of all existence as an integral whole. In metaphysics this trait is responsible for Solov'ëv's doctrine of all-unity. On the opposite end, this holism transmogrified into totalitarianism for which Stalinism stands as the most ominous example.

The evolution of philosophical ideas in Russia has been shaped by the persistent Slavophile-Westernist dichotomy; that is, tension between the impulses, on the one hand, toward national uniqueness and, on the other, toward closer affiliation with the West. However, from the earliest time these tendencies were so closely intertwined with each other that any attempt at a simple delineation is misleading.

And, finally, there is in Russian thought what Berdyaev called the "eschatological" orientation that can also be described as striving toward limits—in particular, the limits of thinking and of intelligibility of things. Like all the other features, this one also has had two opposite consequences. On the one hand, it makes Russian thought philosophically inclined in general, for it pushes rational enquiry to dwell persistently on ultimate questions. On the other hand, such a passion for limits could encourage, as it did in Berdyaev's own case, impatience with careful argumentation.

HISTORICAL EVOLUTION

Russian philosophical thought cannot be properly understood apart from its historical development. Its constant and eager immersion in cultural, social, and political con-

texts, as well as its stubborn continuity, make a historical perspective necessary for grasping both the problems that it grappled with and the solutions that it proposed.

KIEVAN PERIOD. Philosophical ideas, properly so called, first appeared in Russia when Christianity was introduced in 988 by the Kievan Prince Vladimir. The prior, polytheistic view of the world was partially replaced with the Christian outlook, resulting in the fertile amalgam of Eastern Orthodoxy and Slavonic paganism called "dual faith" (*dvoeverie*).

Universities and academic philosophy did not appear in Russia until the eighteenth century, nor was there a direct engagement with ancient Greco-Roman thought of the sort that shaped western medieval learning. Nascent Russian literature absorbed from Byzantium a number of early patristic writings, particularly those of the Cappadocian Fathers, in the form of religious-dogmatic texts translated from Greek into Church Slavonic. Anthologies comprising the writings of John Chrysostom, Gregory of Nazianzus, Basil the Great, Gregory of Nyssa, and John of Damascus introduced Russians to Christian Neoplatonist cosmology, metaphysical anthropology, and allegorical exegesis. Kievan Rus also imported the veneration of Sophia Divine Wisdom that found expression in architecture, icon-painting, and hymns.

In the mid-eleventh century this learning began to bear fruit when the first Russian Metropolitan of Kiev Ilarion described in his "Sermon on Law and Grace" history in terms of contrast between the law of the Old and the grace of the New Testaments, and argued the equal standing of Kievan Rus among Christian nations. Moral ideas were disseminated through "instructions" (*poucheniia*) for righteous living that often contained philosophical ideas derived from ancient and Byzantine thought. Throughout the premodern period philosophy in Russia was viewed primarily as *ancilla fidei* and a path toward religious illumination. At the same time it was understood in broad terms: Plato, Fathers of the Church, and even certain icon painters were considered "philosophers."

Around the mid-thirteenth century this early flourishing was interrupted by the Mongol invasion. Bishop Serapion, who witnessed the sack of Kiev in 1240, was a proponent of the view that history was a series of catastrophes visited by God upon humanity for its sins. With Kiev devastated by the invasion, the center of religious and cultural life shifted to Vladimir and Moscow in the forested northeast that was less vulnerable to attack from the steppes.

MUSCOVITE PERIOD. In the fourteenth century the influence of hesychasm became pronounced, especially through the activities of St. Sergii of Radonezh (1314/22–1391/92). The Trinity-Sergius Monastery near Moscow that he founded soon rivaled Kiev's Monastery of the Caves as Russia's main religious center. St. Sergii's popularity and influence signaled the rebirth of Russian culture around the Grand Duchy of Moscow that in 1380 successfully challenged the Mongol rule. The icon painter Andrei Rublev (d. c. 1430), whose art had a marked contemplative quality, was another representative of this cautious revival. In 1371 the translation of the Areopagitic corpus appeared that had a lasting impact on medieval Russian thought. (More than seventy copies of this work dating to the sixteenth and seventeenth centuries are extant.)

Russia's final emancipation from the Mongol yoke followed soon after Byzantium's fall in the mid-fifteenth century to Ottoman Turks. Both events affected Russians deeply. Soon monk Filofei proposed that Moscow was the third Rome (after Rome proper and Constantinople) and "there will be no fourth" (cf. Zen'kovskii 1953, pp. 34–35). The idea resonated with Muscovite rulers who sought to establish themselves on the European scene. According to this doctrine, Byzantium had fallen because it departed from the true faith and Russia now inherited its mission.

Two major debates convey the atmosphere of the time. Led by the hesychast Nil Sorskii (1433–1508), the so-called "Nonpossessors" (*nestiazhateli*) condemned accumulation of wealth by monasteries and sumptuous church ritual. Their opponent Iosif Volotskii (1439–1515) argued for economically strong monasteries that could help the unfortunate and have a part in social and political affairs. Nil Sorskii was, incidentally, among the first in Russia to refer to the "natural rights" of a person—a theme that gained currency in sixteenth-century religious and political polemics there. (As peasants were being enserfed, some religious writers argued passionately—but to no avail—against slavery as a violation of Christian principles.) The other dispute was the correspondence between Ivan the Terrible and Prince Andrei Kurbskii. The latter argued in favor of a sustained role of traditional aristocracy in government. The tsar's course, however, was to assert his absolute authority with the help of a new gentry that completely depended on his favor. Conducted with ostentatious cruelty the policy did solidify Ivan's autocracy but at a price: By the end of the sixteenth century Russia was in the throes of a major crisis.

On Nil Sorskii's side was Maksim the Greek (1470–1556), the most remarkable intellectual in Russia during that period. Originally a humanist scholar in Florence who later took monastic vows, he was invited in 1518 from Mount Athos to Moscow to assist in translating theological works. While a controversial figure for Russian ecclesiastical authorities, Maksim was nonetheless a scholar of European stature who helped spread philosophical knowledge in Russia.

With the seventeenth century came the painful "time of troubles": Russia's medieval complexion began to change into a modern one. By the middle of the century political and religious tensions erupted in a major schism (*raskol*), which resulted in the separation from the Church of a large group of the so-called "Old Believers." Given the western leanings of their opponents, the energetic Patriarch Nikon and Tsar Alexis (reigned 1645–1676), the schism has been viewed as a struggle between medieval fideistic and modern rationalistic outlooks.

Early in the century Petr Mogila established a spiritual academy in Kiev, fashioned after Polish (Jesuit) models. Secular schools began to appear in Moscow and in 1678 the first institution of higher learning was founded there: the Hellene-Greek Academy. The curricula of these schools included logic, psychology, and physics. The budding academia was occupied by the controversy between the "Graecophiles" faithful to the Byzantine roots of their learning and the "Latinists" influenced by western scholasticism.

THE ENLIGHTENMENT. Inaugurated by the reforms of Peter the Great (reigned 1696–1725), the eighteenth century became the time of a rapid assimilation of western European thought. Philosophical ideas from Europe were absorbed along with progress in the arts, secular education, and science. With the establishment of the Academy of Sciences and universities philosophy attained an official secular status. From translation, publication, and dissemination of foreign literature in the beginning, Russian Enlighteners eventually moved to creating their own works.

The most urgent task for the new educated elite was the development of a secular national ideology. The medieval ideal of "Moscow the Third Rome" was being replaced with the secular ideal of the Russian Empire. The first modern Russian historian, Vasilii Tatishchev (1686–1750) saw, in the Hobbesian vein, the basis of monarchy as the agreement between the sovereign and his subjects rather than in the sovereign's divine right. He argued, in the proto-utilitarian spirit, that "the desire of well-being is inexorable in man and stems from God." (Zen'kovskii, p. 79). His younger fellow-historian, Prince Ivan Shcherbatov (1733–1790), sharply criticized the established church—even as his political sympathies remained on the side of landed aristocracy. Tatishchev and Shcherbatov differed on the most burning moral question of that era, the freedom of the serfs, but both saw the well-being of the nation, rather than its religious mission, as the chief goal of the state.

The ideas of the Encyclopaedists circulated widely among the educated Russian society. Empress Catherine the Great (reigned 1762–1796) was an attentive reader of Charles Montesquieu's treatise *L'esprit des lois* (The Spirit of the Laws, 1748) and maintained correspondence with Voltaire, Diderot, and d'Alambert. Her friendship with *philosophes* doubtless stimulated Voltaire's near-cult status among educated Russians. Unavoidably this interest had much to do with a facile imitation of the West but it also had its serious side. Playing the part of an enlightened monarch, Catherine undertook a relatively progressive, if halting and ultimately unfinished, governmental and legal reform.

The accelerated development of the arts and sciences in this period was epitomized by the polymath and poet Mikhail Lomonosov (1711–1765). A fisherman's son from a northern province, he became the first Russian scientist of European stature and was instrumental in promoting scientific research and higher education in his country. To him belonged the famous prophecy, in verse, that combined the zeal of an Enlightener with national pride: "The Russian land can give birth to its own Platos and quick-witted Newtons."

Simultaneously the traditional line of Orthodox theology was carried on by Paisii Velichkovskii (1722–1794) and St. Tikhon Zadonskii (1724–1783). Velichkovskii was a spiritual elder, the type best known from Zosima, a character in Dostoevsky's novel *The Brothers Karamazov*. Arguing against the alleged sanctification of the created world in secular thought, St. Tikhon taught that the external world had to be transfigured rather than accepted on its own terms. Concentrated on righteous living and one's personal connection with the Absolute, this theology was a welcome reprieve, as Zen'kovskii notes, from the burden of justifying Russian state messianism.

A counterpoint woven of both secular rationalism and religious mysticism was created by the most remarkable philosopher of the Russian eighteenth century, the Ukrainian Grigorii Skovoroda (1722–1794). A "Nonpos-

sessor" and itinerant philosopher of a Socratic mold, Skovoroda expounded an original doctrine that was inspired by ancient sources, patristic thought, and modern European philosophy. There was a Christian Neoplatonist note in his belief that man's proper purpose was an "erotic" ascent to divinity, as well as in his self-written epitaph: "The world tried to catch me, but has failed." His influence on the contemporary philosophical scene was, sadly, almost nonexistent; his works were not published during his lifetime and began to attract serious attention only in the nineteenth century.

A different quest for spirituality outside the Church was evident in the movement of Freemasons that started in Russia in the second third of the eighteenth century. In the 1770s there emerged among them a group led by Nikolai Novikov (1744–1818) and Johann Schwarz (1751–1784). Novikov's contribution was mostly as an editor and publisher: from 1779 to 1792 he published almost nine hundred titles that included, aside from Russian authors, translated works of Jacob Boehme, Voltaire, John Locke, G. E. Lessing, and Novikov's favorite, Blaise Pascal. These Freemasons combined respect for natural science with the primacy of morality over the intellect.

Alongside modern scientific realism the nascent Russian intelligentsia absorbed western utopianism. As in the West, however, utopia often served as a vehicle for social criticism. Vasilii Trediakovskii in his *Tilemakhida* (1766), a verse translation of François Fénelon's novel *Les Aventures de Télémaque*, described the torment of monarchs in Tartarus: they looked at their own monstrous images in the "mirror of truth." From Trediakovskii's poem came the epigraph to Aleksandr Radishchev's (1749–1802) *Journey from St. Petersburg to Moscow* (1790): "The monster is opulent, impudent, enormous, hundred-mouthed, and barking." The main target of Radishchev's moral sermon from the standpoint of natural rights was the inhumanity of the institution of serfdom. It was the most striking fruit of the Russian Enlightenment, and cost the author dearly: he was exiled to Siberia. Novikov was likewise imprisoned in St. Petersburg. Frightened by the French Revolution, the aging Empress was now perturbed by the liberties her subjects were taking.

The turn from the Enlightenment to conservatism among Russian intellectuals was vividly exemplified by the historian and writer Nikolai Karamzin (1766–1826). A proto-Westernist, he was originally attracted to Locke and Rousseau but his views evolved from a vague empiricism and tolerant sentimentalism to defending the expedience, for the stability of the state, of "enslaving people rather than prematurely freeing them." The French Revolution was the key factor in this striking change. Karamzin initially hailed it as "the triumph of reason" but then, as terror struck, condemned it as the collapse of the Enlightenment. He was among the first to give Russians a serious perspective on their own history. The poet Aleksandr Pushkin (1799–1837) compared his discovery of Russia's past to Columbus's discovery of America.

The Russian Enlightenment drew to a close when, after Catherine's death, Novikov was freed only to live out the remainder of his life in obscurity, and Radishchev, a few years after his release, committed suicide. But its ideas became an integral part of Russia's intellectual makeup. Its complex legacy contained mutually intertwined, conflicting themes, such as national identity and universal humanism, secularism and religious tradition, scientific cognition and mysticism, art and morality, theoretical quest for truth vis-à-vis social practice.

THE GOLDEN AGE. Although rooted in a long-standing cultural and spiritual tradition, Russian philosophy proper was born in the nineteenth century. As it matured, it underwent several waves of foreign influence: idealist (especially German) in the 1830s and 1840s, positivist in the 1860s, Marxist in the 1880s and 1890s—to mention only the most poignant ones. Once it appeared, each strand remained an active factor in the continuing philosophical debate. Russian mentality has been described as inclined toward extremes, and the reception of Western ideas in Russia bears out this observation: their assimilation often meant radicalization. This was true of the "Nihilists" of the 1860s who developed a cult of natural science, and later of Vladimir Lenin (1870–1924), who stripped Marxism down to its bare essentials and ruthlessly pursued his vision. Solov'ëv, by contrast, strove to synthesize diverse strands into a holistic idealist vision.

The famous phenomenon of the intelligentsia arose in this century. Recruited mostly from the middle class, the new educated elite developed a degree of self-consciousness one rarely finds in its Western counterparts. The idea of its "debt to the people," articulated in Petr Lavrov's (1823–1900) *Istoricheskie pis'ma* (Historical letters, 1868–69), shaped the ethos of this group. From the very beginning, though, the intelligentsia was torn by internal conflict and contradictions. Its admirers saw in it the "conscience of the nation," its critics an intolerant "monastic order" of political radicalism, and many of its members were convinced that the two were synonymous. In the meantime such major thinkers as Solov'ëv, Dosto-

evsky, and Tolstoy resisted being included among its ranks. In the early twentieth century philosophers of religious orientation subjected the intelligentsia's atheist outlook to an unflattering critique. They were noisily rebuked both by radicals and liberals. Left-wing intelligentsia played a crucial part in bringing about the revolutionary turmoil of the early twentieth century—the turmoil that led to its own dispersal in the thin air of history. Originally the flag-bearer for social progress and against despotism, in the Soviet period it became an evanescent specter. Its relation to the so-called "Soviet intelligentsia" was too problematic to warrant a continuum between them.

Early developments in philosophical education were not auspicious. Organized on Wolffian principles, academic philosophy had enjoyed steady growth since the middle of the eighteenth century. From 1817 and until the mid-nineteenth century, however, it suffered from a crisis precipitated by a conservative turn in Alexander I's policy and then exacerbated by the oppressive rule of Nicolas I (reigned 1825–1855). The teaching of philosophy was abolished for long periods in gymnasia and universities. A senior official summed up the government's view of it: "Utility is doubtful, whereas harm is obvious" (Radlov, *Ocherk istorii russkoi filosofii* [Essay on the history of Russian philosophy], 1920, p. 7). To circumvent restrictions some professors taught philosophy under the guise of other disciplines, such as history or geology. Philosophical instruction continued uninterrupted, however, in religious seminaries and academies but it was not until the second half of the century that the situation of academic philosophy began to be more or less normalized. Yet even as conditions improved, Russian thought retained much of its nonacademic character. For various, mostly political, reasons prominent thinkers—be it Aleksandr Herzen (1812–1870), Solov'ëv, or Nikolai Chernyshevskii (1828–1889)—worked outside universities.

In the 1820s the first philosophical circle appeared; its members called themselves by the Russian equivalent of *philosophes*—*liubomudry*, "lovers of wisdom." The group's leader, Prince Vladimir Odoevskii (1804–1869), presented a Schellingian view of Russia's future in his utopian dialogue-novel *Russian Nights* (1844) in which he gave a modern version of Russian messianism. History moved, he rhapsodized, toward "a holy triunity of faith, science, and art." Anticipating Dostoevsky, he claimed that Russia was destined to accomplish this universal synthesis because of her "all-embracing multifaceted spirit."

Such optimism, however, was in sharp contrast to the somber skepticism of Petr Chaadaev's (1794–1856) *Philosophical Letters*. Chaadaev saw the West as the ideal of civilization; all other societies were, in his opinion, mere approximations to it, with Russia falling outside the category altogether. Chaadaev's bitterness was cast against the background of two recent events: Russia's victory over Napoleon in 1812 that encouraged hopes for the nation's greatness, and the crushing defeat of the 1825 Decembrists's uprising that extinguished hopes for reform and liberty. He was inspired in large part by Joseph de Maistre and Friedrich Schelling. He later fine-tuned his position to argue that Russia was called upon to resolve the contradictions that still plagued the West. The evolution of Chaadaev's views became typical for Westernists: from adulation of the West to disillusionment to seeing Russia's potential in her backwardness. The conviction that Russia was a "virgin soil" whose lagging behind could be turned to advantage as "the possibility of choice" became the cornerstone of Westernist constructions from Herzen to Vladimir Lenin (1870–1924). Chaadaev's caustic but profound outburst brought into existence two opposite trends, the Westernists and Slavophiles, whose mutual rivalry has since shaped, and continues to shape, the evolution of Russian thought.

Because of their intertwined destinies "Slavophiles" and "Westernists" come close to being the worst misnomers in the history of Russian thought. Both groups were deeply dissatisfied with the current conditions in Russia. Contrary to the xenophobic connotation of their name, many Slavophiles respected European learning and culture and kept abreast of recent Western philosophical thought. For their critique of the West they often borrowed ammunition from the West itself. Conversely, the Westernists' professed cause was to save Russia, and many of them even believed, such as Herzen, that Russia held the key to saving the West from the West's own woes. For both, the goal of "enlightening" Russia was of paramount importance, although they were divided on the possibility of "national science." Slavophiles defended the idea (without defining it clearly), whereas Westernists rejected it in favor of universal rationality.

And yet their differences were not trivial. Slavophiles believed that, enviably advanced as it was, Europe had come to an impasse and Russia had to avoid a similar fate. The West's original sin, according to Slavophilism, consisted in the rationalistic tendency of Roman Catholicism that was codified in the *filioque*; that is, the dogma that the Holy Spirit proceeded from *both* the Father *and* the Son. Both early Slavophiles, such as Aleksei Khomiakov (1804–1860) or Ivan Kireevksii (1806–1856), and their later followers, such as Sergei Bulgakov (1871–1944) or

Nikolai Losskii (1870–1965), accused Catholic theology of replacing the mystery of the Holy Trinity with a hierarchical scheme in which the Holy Ghost was subordinated to the other two persons. This eventually led, via scholasticism, to Protestantism and thence to modern secularism. The decline of the authority of the Church in turn weakened, Slavophiles believed, the foundations of communal life and created the West's atomistic individualism. Russia, they claimed, could offer an alternative because its culture still contained the original wholesome elements, unspoiled by the westernization of the previous two centuries.

Against rationalism in epistemology the Russian mind could offer, Kireevksii argued, the ideal of integral knowledge in which rational thinking and divine revelation would be properly balanced. Against individualism in social philosophy it could offer *sobornost'*—the concept that amalgamates "togetherness" with "conciliarism" (from "church council") and projects the ideal of an humanity united by love and faith, where the freedom of the individual is in harmony with the common cause. Khomiakov found its manifestation in the Orthodox Church and Konstantin Aksakov (1817–1860) in the Russian village commune. Russia's historical task was understood as universal, although it remained unclear how other nations, who had their own traditions, were supposed to accept Eastern Orthodoxy. Slavophiles' concern, however, was to outline Russia's potential place in the "family of nations" rather than to develop a specific strategy for attaining it. The mankind of the future was perceived, in Aksakov's terms, as a "choral person"—the notion that in the twentieth century was assimilated by Lev Karsavin (1882–1952) into his doctrine of humanity as a "symphonic person."

Westernists, on the contrary, insisted that Russia needed to join advanced European nations in pursuing economic, social, and political progress. Where Slavophiles envisioned *sobornost'*, Westernists insisted on the legal rights of the individual. If Slavophiles found pristine purity in pre-Petrine Russia, Westernists blamed the country's slow progress on xenophobic medieval Russian tsardom. Their sharpest difference from Slavophiles, however, consisted in their hostility toward religion. In Herzen's words, there was an "ecclesiastic wall" between him and his opponents. The common limitation of both was their utopianism: One idealized Russia's past and the other, the West's present. Furthermore, for neither of them philosophy had independent value but was merely an instrument for achieving goals other than knowledge and understanding.

The reception of Schelling and Hegel casts a helpful light on the manner in which philosophy's tasks were conceived. Schelling's philosophy enjoyed a warmer reception—at least in the religious segment of Russian thought. In fact, there is some truth to Arsenii Gulyga's (1921–1996) remark that "Russian philosophy is a Schellingian." With Hegel Russians tended to distance themselves, even as they respectfully learned from him; in Schelling they found a kindred spirit. The view of the world as an organic whole has had more followers and fewer detractors in Russia than in the West; it retains importance there to this day. Schelling's doctrine of intellectual intuition proved particularly attractive to Russian thinkers. From Odoevskii to Solov'ëv they embraced the notion of an immediate meeting of consciousness with both inner and outer reality; in the twentieth century it inspired a whole intuitivist school. Chaadaev was deeply affected by Schelling's philosophy of revelation; Kireevskii and Solov'ëv, by his epistemology; Odoevskii and Bulgakov, by his *Naturphilosophie*; and Aleksei Losev (1893–1988), by his aesthetics and philosophy of myth. Many of them found in Schelling's thought inspiration for viewing art and religion as (extrarational) sources of rational thinking.

Russian liberal thought was, by contrast, at its inception primarily Hegelian. Vissarion Belinskii's (1811–1848) and especially Herzen's engagement with Hegel's philosophy were typical. Both embraced Hegelianism in the beginning but then rejected what they perceived as its abstract universalism. Belinskii, on the one hand, got most of his Hegel via Mikhail Bakunin (1820–1900) who at the time was an overenthusiastic Hegelian. Herzen, on the other hand, attentively studied Hegel's writings firsthand. The result was, however, more or less similar. "(Hegelian) reason does not know," Herzen impugned, "*this* person but only the necessity of a person in general …" (Zen'kovskii, pp. 285–6). The main point of Herzen's dissatisfaction was the same as Karl Marx's: life is not merely about thinking, he insisted, but chiefly about acting in the world. Virtually all Russian philosophers turned away from Hegel upon initial acquaintance. Those consumed by revolutionary causes, such as Bakunin, blamed him for excessive contemplativeness, whereas Slavophiles and religious philosophers rejected his doctrine of a rationally cognizable absolute. Various parts of Hegel's system were adopted but only the rarest exceptions, such as Boris Chicherin (1828–1904), accepted its essential core, the doctrine of the absolute concept. Characteristically, Herzen found in Hegel's dialectic "the algebra of revolution"—a description that was later eagerly endorsed by Lenin. This appropriation

epitomized the political pragmatism that was imposed on the German philosopher's speculative method.

Philosophers' concerns for "the concrete person" were nourished by the burgeoning Russian realist literature that paraded, in an intensely empathetic light, a series of characters whose suffering was a condemnation of a social order in which human dignity was out of place. Conversely, Russian thinkers frequently offered their insights in literary form. In fact, the most burning of the "cursed questions" that preoccupied the intelligentsia throughout its existence were articulated as titles of literary works: Herzen's 1847 novel *Who Is to Blame?* and Nikolai Chernyshevsky's (1828–1889) 1863 socialist utopia *What Is to Be Done?* The latter query proved particularly haunting: Leo Tolstoy in 1883 and Lenin in 1902 each wrote a work bearing similar titles.

Hegel and Schelling were soon replaced by Ludwig Feuerbach and Left Hegelians as socialist ideas were spreading among educated Russians. In the 1860s materialism propounded by Ludwig Büchner and others was added to the mix; it was embraced by the so-called "Nihilists" whose leading figures were Dmitrii Pisarev (1840–1868), Nikolai Dobroliubov (1836–1861), and Chernyshevskii. Pisarev's crude materialism, however, was not so much a philosophical position as a propagandistic means of destabilizing old religious and social values. Calculated to outrage, his maxim that "boots are more valuable than Shakespeare" was, in fact, a call to social activism as opposed to the aesthetic hedonism of the leisure classes. It was also a message about the utility of science and technology; that is, the business of the newly emerging class of physicians and engineers, contrasted with the aristocratic art of the previous era. The most articulate thinker of the "Nihilist" camp, Chernyshevskii, by contrast, argued for genuine art that would be a life-transforming praxis rather than idle entertainment. The rise of Nihilism marked the radicalization of Herzen's intellectually broad and humane liberalism, and the beginning of the latter's transmogrification into fanatical revolutionism.

In the late 1860s and 1870s the earlier materialism was absorbed into the broad social, cultural, and ideological movement called "Populism" (*narodnichestvo*). Its intellectual leaders, Lavrov and Nikolai Mikhailovskii (1842–1904), combined positivist epistemology and materialist metaphysics with an evolutionist view of history. The Populists' goal was socialism in Russia, on the basis of the village commune. Their views about both the goal and the ways of achieving it, however, varied from the anarchism of Bakunin and Petr Kropotkin (1842–1921) to the conspiratorial terrorism (with a Marxist tinge) of Petr Tkachev (1844–1886). The Populists' main philosophical difficulty consisted in reconciling the individual's agency with positivist determinism. Like their materialist predecessors, however, these thinkers did not embrace a particular philosophy of nature or history for its intellectual merits but were interested primarily in using it for social change. It was Mikhailovskii who pointed out, memorably, the conflation of "truth" and "justice" in the Russian word *pravda* that has since come to signify one of the most pervasive features of the Russian philosophical mindset. It was also Mikhailovskii whose "subjective method" in sociology was intended to enhance the ability of "critically thinking individuals," as Lavrov called them, to influence the course of history. Populism later evolved into the political party of Socialist Revolutionaries, the Bolsheviks' most powerful left-wing rival, and its ideas continued to exercise their influence well beyond its final collapse in the 1920s.

Less influential was the moderate liberal thought of such thinkers as Konstantin Kavelin (1818–1885) and Boris Chicherin (1828–1904) who defended, from an Hegelian position, the ideals of the law-governed state in political theory and the universal "higher synthesis" of religion and philosophy in epistemology. As the earlier Westernism was radicalized, so too the original, rather moderate Slavophilism was producing its own increasingly radical offshoots. Konstantin Leont'ev (1831–1891) offered a scathing critique, on aesthetic grounds, of contemporary Western society. Unlike Friedrich Nietzsche with whom he is frequently compared, Leont'ev ended not with a call for a proud Overman, but with a return to an ascetic Orthodoxy. Nikolai Danilevskii's (1822–1885) theory of "cultural-historical types" advanced a cyclical model of history in which the tired Romano-Germanic civilization was about to yield its place to a younger Pan-Slav one. Danilevskii's ideas had an impact on the "back-to-the-soil" group of authors (*pochvenniki* from *pochva*, the Russian for "soil"), whom Dostoevsky lent his not insignificant authority.

Dostoevsky was, incidentally, one of the first Russian thinkers who had a marked influence on Western philosophy. His explorations of the religious, moral, and psychological dimensions of the human condition made a deep impression on both contemporaries such as Nietzsche and later figures such as Albert Camus. Inside Russia Dostoevsky's ideas reverberated in the religious-philosophical school of the early twentieth century.

The more liberal patrimony of Slavophilism, however, was cultivated by Russia's first truly great philosopher Solov'ëv. Solov'ëv's philosophy was an impressive attempt to fuse together positivism, idealism, and mysticism. His early critique of positivism evolved into the assimilation of Auguste Comte's ideas into his own view of history as divine will unfolding toward "free theocracy." Comte's *Grand Être* was likewise absorbed, along with Gnostic, Cabalistic, Eastern Orthodox, and German Romantic ideas, into Solov'ëv's neoplatonist metaphysics of Sophia Divine Wisdom. Later Solov'ëv performed a similar operation on Chernyshevskii's positivist aesthetics by interpreting it in the light of his own doctrine of art as theurgy: that is, humanity's continuation of divine creation. Yet his syntheses were not eclectic but rested on a broad conceptual foundation and formed a more or less coherent system—the first created by a Russian philosopher. With his more eager, ecumenical acceptance of the West Solov'ëv modified earlier Slavophilism and worked to reconcile it with Westernism. Above all, however, his most lasting contribution consisted in the apologia of philosophical idealism. Solov'ëv and Dostoevsky remained lonely voices among the intelligentsia during their lifetime but by the time of Solov'ëv's death a reaction had already begun among a new generation of philosophers against secular ideologies and in favor of a serious engagement with religion.

While the rebirth of philosophical idealism was only dawning, however, its antipode was vigorously gaining ground. Marxism was known in Russia since the late 1840s but in its early stages it was only one among several currents of socialist thought. Nevertheless, it soon attracted significant interest: In 1869 Bakunin published (abroad) his translation of the *Communist Manifesto*, and three years later Russian became the first foreign language in which the first volume of *Das Kapital* appeared. By the end of the century Marxism became the most influential political doctrine among the intelligentsia. It established itself in competition with earlier socialist theories, primarily Populism. In contrast to Populists who wished Russia to avoid capitalism and leap, via village commune, directly into socialism, Marxists viewed capitalism as a stepping stone to socialist revolution. The abolition of serfdom in 1861 by Alexander II gave a strong impetus for the development of capitalist enterprise and, as the number of factory workers grew, socialist theorists began to pin their hopes on the new class. The key figure in the transition from Populism to Marxism was Georgii Plekhanov (1856–1918). His main concern seems to have been to elaborate a philosophical system based on Marxist precepts, while guarding the original doctrine against misinterpretation and revisions. A significant feature of Plekhanov's reception of Marx's ideas was their refraction through Frederic Engels's work. Russian Marxists did not always take care to distinguish Marx from Engels and often argued—in fact, often they simply assumed—the unity of the two founders' respective positions.

In the last quarter of the century Russian academic philosophy finally became the key factor on the philosophical scene. The generation of Solov'ëv and Mikhailovskii was receding into the past and most leading thinkers now taught at universities. Chicherin gradually developed his own system with an emphasis on the philosophy of right and of history. A Leibnizian revival was evident in the trend started by Aleksei Kozlov (1831–1931) that stimulated the development of personalism in Russian thought. The latter had an exceptionally far-reaching impact on such thinkers as Berdyaev, Losskii, and Lev Shestov (1866–1938). This was also the time when Kant's presence in Russian thought finally came to match that of Schelling and Hegel. The leading neo-Kantian Aleksandr Vvedenskii (1856–1925) concentrated on logic and philosophical psychology. Advocated by a number of scientists and philosophers, such as Vladimir Vernadskii (1863–1945) and especially Vladimir Lesevich (1837–1905), neopositivist thought was another major current in academic philosophy. It was concerned almost exclusively with the philosophy of science and empirical epistemology. Vernadskii's ideas later played an important part in what became known as Russian cosmism. The original tenets of this loosely defined trend were formulated by the (nonacademic) Nikolai Fedorov (1828–1903) whose eccentric hybrid of positivism and Christian eschatology aimed at the physical resurrection of all past generations.

THE SILVER AGE. The flourishing of the arts and philosophy, roughly, from 1890 to 1925 is often referred to as the "Silver Age." It was marked by the rise of Symbolist poetry, modernist music, avant-garde art, and a general invigoration of cultural life. The Silver Age unfolded against the background of growing capitalism and a relative liberalization of political life, punctuated by wars and revolutionary turmoil. New developments in the arts underscored expectations of tectonic shifts in political history. The theme of an impending catastrophe—hailed as a purifying storm by some and feared as a fatal calamity by others—haunted artists and philosophers alike. Russia's humiliating defeat in a war with Japan precipitated the first, abortive popular uprising in 1905. The tsarist government agreed to halfhearted parliamentary reforms but they were undermined by the outbreak of

World War I in 1914 and then annulled altogether by the Bolshevik revolution of October 1917.

Russian philosophy matured during this period. From the 1890s on government restrictions were loosening and in the early 1900s the autonomy of universities finally began to materialize. In 1889 the first professional philosophical journal, *Voprosy filosofii i psikhologii* (Questions of philosophy and psychology) was founded, followed in the first decade of the new century by several other publications specializing in philosophy. In 1897 the St. Petersburg University Philosophical Society was established and a few years later it was joined by the Religious-Philosophical Society in Memory of Vladimir Solov'ëv in Moscow and the Religious-Philosophical Society in St. Petersburg.

Contacts with European philosophy reached a high point. Russian philosophy was now fully integrated, if still as a minor partner, into the European philosophical culture. The most recent developments in Western thought were quickly assimilated by Russian thinkers; empiriocriticism and phenomenology were only the more notable among such new trends. The growing influence of Kant was mentioned above. Nietzsche's impact on the Russian thought of this period was profound and pervasive.

Two opposite, unequal trends dominated the scene during this time: Marxism and religious philosophy. The former was philosophically unimpressive but politically influential, whereas the latter, on the contrary, was politically insignificant but philosophically fertile. Their complex mutual interactions, ranging from antagonism to fusion, were the manifestations of a dynamic and visionary rather than rigorous *Zeitgeist*. Scientific positivism and political liberalism also continued, adding to the increasingly vibrant philosophical life.

The brand of Marxism that emerged as a result of Plekhanov's efforts and was now endorsed in the main by Lenin included the following basic components. It was founded on a materialist ontology; that is, the view that matter constitutes the source of all existence. Materialism was enhanced by a positivist epistemology that held modern science to be the only legitimate source of knowledge. Marxism considered itself a true—in fact, the only true—doctrine *because* it was a modern scientific theory. Its next key component, historical materialism, was the result of synthesizing the first two with Hegel's philosophy of history. And finally the whole was held together by dialectical materialism, also a permutation of Hegelian dialectics adapted to fit materialism and positivism. (Needless to say both Hegel's philosophy of history and his dialectics were drastically deformed in these hybrids.) Materialist orientation also dictated that all social and political phenomena be viewed as determined by a society's economic base. The latter developed, according to the theory, over periods of gradually accumulating quantitative changes leading up to abrupt moments of revolutionary qualitative change. The result was the view that history was logical progress from one socioeconomic formation to another, culminating in communism as the most rational system. There was no room for divine authority in this picture; militant atheism was an indelible feature of Russian Marxism. In an apparent contradiction to its own economic determinism, the key factor in the "inevitable" socialist revolution was Marxism itself as a doctrine of "scientific socialism." Further, despite being the most revolutionary class, the proletariat had to be educated; as Lenin argued, "scientific socialism" had to be instilled in its consciousness.

In ethics universal moral values were rejected as products of "abstract bourgeois humanism" in favor of the view that all values were determined by class interest. The corollary was that, as the revolutionary vanguard of society, the proletariat held values that were superior to those of any other class. In aesthetics a similarly class-based criterion was adopted: judgment about art was determined by which class interest it promoted. Leo Tolstoy's oeuvre, for example, was famously described by Lenin as "the mirror of the Russian revolution." These principles received a less stark complexion once they were combined with a dialectical view of history according to which new eras partially reject but also partially absorb the achievements of previous ones. Thus the proletariat was supposed to have inherited the best that world civilization had developed prior to socialist revolution. But the ultimate authority on all issues belonged to the proletariat's own vanguard, the Communist Party. Likewise Lenin's unabashedly utilitarian, ideological aesthetic eventually replaced Plekhanov's earlier, more nuanced attitude as the official "partisan principle" in evaluating art.

There soon evolved two currents in Russian Marxism: radical and moderate. Lenin and his fellow Bolsheviks (the term "Bolshevik" literally means "a member of the majority") promoted the former, whereas the so-called "legal Marxists" that included Petr Struve (1870–1944), Berdyaev, and Bulgakov, advocated the latter. The radical trend absorbed from extremists such as Tkachev revolutionary voluntarism and justification of terror as a means of political change.

The main controversies that divided these currents had to do with whether Russia could bypass an extensive phase of capitalism and bourgeois democracy, and proceed directly to a socialist revolution. Lenin answered in a resounding affirmative, whereas his opponents, including Plekhanov, favored a less precipitous path. They feared that the dictatorship of the proletariat, which the Bolsheviks envisioned as the key instrument of transition from a semifeudal to socialist society, would be as oppressive as the tsarist regime. The question of eschewing a prolonged capitalist phase was also bound up with whether Russia could pursue the socialist path alone among nations. In classical Marxism progress toward socialism had been envisioned as an international process because capitalism was itself an international system, too entrenched for the proletariat of one country to overpower it. Russian Marxists split on the issue: The moderate wing laid stress on international cooperation and advocated waiting for ripe conditions in advanced European nations, whereas the radical wing insisted that it was possible to establish socialism in one country.

The main philosophical difficulty for Marxists stemmed from the materialist foundation of their doctrine and consisted in explaining how purely physical, unconscious matter could generate movement and, ultimately, consciousness. The argument that matter evolved in accordance with the laws of nature only raised questions about the origin of these laws themselves. Plekhanov and Lenin asserted that science disclosed what matter was but this claim lost its persuasiveness as new conceptions of matter were developed in physics and the hypothetical nature of these views became increasingly apparent. Lenin's statement that "matter is objective reality given us in sensations" was vague enough to accommodate idealism and thus created more problems than it solved. Similar problems haunted Marxist ethics. The critics of the dogmatic trend, such as Struve, complained that class interest did not provide a firm foundation for morality and, further, dissolved individual agency in socioeconomic forces. The dismissal of art as an activity with a distinct purpose was also problematic. Nor did philosophy itself fare better. "From Marx's and Engels' point of view," wrote Lenin in his essay "The Economic Meaning of Populism" (1894), "philosophy has no right to independent existence and its subject-matter divides itself [literally 'disintegrates,' *raspadaetsia*] among several branches of positive science."

In evaluating its claims, however, it is critical to realize that Russian Marxism was first and foremost a doctrine of political action. Its logic, philosophy of history, social philosophy, epistemology, and even materialist ontology were adopted under the pressure of a specific sociopolitical ideal. It was the ideal of a strictly secular, modern society aimed at assuring the fullest realization of the immanent human potential by rationalizing the production and distribution of material wealth. The "superstructure" was to align itself with, and serve the achievement of, this goal. Hence Lenin's relentless defense of materialism, insistence on the scientific nature of Marxism, and uncompromising atheism. In Lenin's thought Russian Marxism's ideological pragmatism reached its apogee. Scant and unimpressive at best, his philosophical writings were all occasioned by topical debates and aimed at ensuring the resolve of the Bolshevik party. The motivation for his most extensive philosophical work, *Materializm i empiriokrititsizm* (Materialism and empiriocriticism, 1909), for example, was to rein in his comrades Lunacharskii and Aleksandr Bogdanov (1873–1928) who had strayed into "God-building" and "empiriomonism." The only exception was Lenin's *Filosofskie tetradi* (Philosophical notebooks, 1914) in which a more serious engagement with Hegel was evident, but these were private ruminations published only posthumously. Materialism and dialectics were meaningless for Lenin unless they were employed for the communist cause. "Materialism," he wrote, "includes partisanship (*partiinost'*)." (*Collected Works*, Vol. 1, 1960, p. 401).

The moderate branch of Russian Marxism was more in earnest about resolving the philosophical difficulties of the doctrine but attempted solutions led to revisions of its original materialist, positivist, and deterministic tenets. The "legal Marxists" Struve, Berdyaev, and Bulgakov eventually abandoned orthodox Marxism in favor of philosophical idealism.

This was a sign of the opposite trend that became evident in the emergence of neo-Kantianism and especially religious idealism. The return to Kant was chiefly a development in academic philosophy, whereas the turn to religion swept along academics, independent thinkers, and artists. In later literature the appearance of a group of philosophers who drew inspiration from religion was described as a "religious-philosophical renaissance." The writer Dmitrii Merezhkovskii's (1865–1941) quest for a "new religious consciousness" was a more popular manifestation of this trend. Merezhkovskii initiated Religious-Philosophical Meetings in 1901–1903 as an attempt at a rapprochement between the church and the intelligentsia. The participating sides were ill at ease with each other and after Vasilii Rozanov's (1856–1919) character-

istically shocking call upon the clergy to sanctify physical sex the meetings were stopped on government's orders.

Three publications mark the evolution of this trend in the first two decades of the century. The 1902 anthology *Problemy idealizma* (Problems of idealism) was an initial attempt to revive idealism as a viable contemporary philosophical position, followed by the *Vekhi* (Landmarks, 1909), a cutting critique of the intelligentsia's ideological dogmatism, atheism, and social isolation, and, finally, by *Iz glubiny* (De profundis, 1918), a reaction to the Bolshevik revolution as an anti-Christian act prepared by the spiritual, cultural, and moral crisis of the previous two decades. (This indictment was echoed by Rozanov who called the Revolution "the apocalypse of our time.")

The crucial problem that these thinkers confronted was the reconciliation of philosophy with religion. The impulse to embrace religion came as a result of recoiling from materialism and positivism. Many religious philosophers began as Marxists in their younger years and then underwent an idealist conversion. But an attempt to reconcile religion and philosophy led to the choice between fideism and rationalism. Like their predecessors, Slavophiles and Solov'ëv, Russian religious philosophers ultimately leaned toward the former. Ernest Radlov (1854–1928) even claimed in his 1920 *Ocherk istorii russkoi filosofii* (A survey of the history of Russian philosophy) that the tendency among Russian thinkers toward a mystical solution of ethical and epistemological questions was a "national trait." A closely related task that these philosophers pursued was defense of idealism. In many cases such defense involved rethinking the relation between ideas and empirical reality and resulted in a number of constructs: the "concrete idealism" of Sergei Trubetskoi (1862–1905), "ideal-realism" of Losskii, and "mystical realism" of Berdyaev. The "abstract" thought of German Idealism often served as a contrasting foil for these attempts to bring idealism closer to life.

At the same time the religious-philosophical school argued for a secular culture and philosophy informed by the Orthodox faith—domains that had been neglected, in their opinion, by the Russian Orthodox Church. Florenskii's 1914 classic *Stolp i utverzhdenie istiny* (The pillar and ground of truth) was perhaps the most monumental attempt to fuse together a modernist philosophical and aesthetic sensibility with Orthodox faith. The most important sources of inspiration for them included the thought of the early Slavophiles and especially of Vladimir Solov'ëv. In epistemology they questioned both extreme rationalism and extreme fideism but their attitudes varied widely. In method their approaches ranged from Losskii's strict adherence to formal logic to Semen Frank's (1877–1950) moderate dialectics to Berdyaev's aphoristic impressionism. In metaphysics many of them followed and further developed Solov'ëv's doctrines of all-unity and Sophia Divine Wisdom. Their views on philosophy of history encompassed Florenskii's admiration of the Middle Ages, at one pole, and Berdyaev's progressivist Christian socialism, at the other. In political philosophy they were likewise diverse: Ivan Il'in (1883–1954) rigidly advocated monarchism, whereas Viacheslav Ivanov (1866–1949) vaguely evoked mystical anarchism. The only thing that united them was the conviction that modern secularism had exhausted itself and the reinvigoration of philosophy and culture in general was to be sought in a union with religion.

A particularly notable contribution by this group was their writings on the history of Russian philosophy. Evgenii Trubetskoi's (1863–1920) classic study on Solov'ëv, Berdyaev's essay on Khomiakov, Gustav Shpet's (1879–1937) hypercritical survey, and Radlov's work mentioned above were part of this self-examination. A special place in this literature belongs to works on the "Russian Idea." Rooted in the writings of Dostoevsky and Solov'ëv, this trope grew into a body of literature created by several generations of philosophers. On the broadest level, it referred to the unique Russian type of consciousness, culture, historical destiny, and place among the peoples of the world. After the Revolution this tradition was further elaborated in Eurasianism and culminated in Berdyaev's classic *Russkaia ideia* (The Russian idea, 1946). It eventually reemerged in post-Soviet thought where it took on still other interpretive hues.

Along with metaphysical, epistemological, and political issues, an exceptionally preeminent concern for this group was art, which they viewed as a conduit for religious enlightenment. Evgenii Trubetskoi, Florenskii, Bulgakov, and Berdyaev all dedicated to art some of the most inspired pages of their philosophical prose. Their insights into icon-painting (which Trubetskoi described as "theology in color"), liturgy (which Florenskii interpreted as the Orthodox *Gesamtkunstwerk*), and artistic creativity in general remain to this day exemplary in their subtlety and depth.

Even more than before Russian philosophy evolved during this time in an intense dialogue with the arts. The Russian avant-garde was often inspired by, and inspired in turn, the volatile mix of philosophical ideas. Among artistic movements Symbolism stood out, both in terms of its artistic influence and engagement with philosophy.

Poets Andrei Belyi (1880–1934), Aleksandr Blok (1880–1921), and especially Ivanov keenly explored the philosophical dimensions of their art. Symbolists came to the view, rooted in Romanticism and Solov'ëv's theurgy, that art provided access to the "more real" plane of being and was a path toward spiritual or even cosmic transfiguration. Both philosophers and artists were fascinated with the limits of art. A wide array of artistic movements was driven by a desire to break down the barrier between art and life raised by Kantian disinterested aesthetic contemplation. The pivotal event of this period, the Bolshevik Revolution, did not initially stop this feverish activity but marked a watershed that inaugurated a new phase in the history of Russian philosophy.

THE SOVIET PERIOD. Two major processes were under way in the 1920s: the decline of the Silver Age and the rise of Soviet ideology. The new government sought a total submission of philosophy to state ideology and the means by which this was assured ranged from administrative pressure to exile to physical annihilation of dissenting thinkers. Berdyaev's Free Academy of Spiritual Culture and the Free Philosophical Association founded by Radlov, Losskii, and others in St. Petersburg were short-lived attempts to continue prerevolutionary activity. Philosophers associated with both were expelled from the country in 1922 among a large number of thinkers and scholars unsympathetic to the Bolshevik regime. In 1921 the teaching of non-Marxist philosophy was banned and in 1923 philosophy was replaced by dialectical materialism in higher education.

The tasks of Soviet philosophy consisted in "developing" Lenin's patrimony (which meant strictly adhering to its key tenets), combating domestic and foreign "bourgeois idealism," justifying the Party's political decisions, and supplying methodology to the sciences. Formulated even before Soviet philosophy as such was in existence, these tasks remained unchanged throughout the Soviet period. As state ideology Soviet Marxism was based on the Plekhanov-Lenin interpretation of Marx and Engels's views that was soon branded "Marxism-Leninism."

The debate during the 1920s between the so-called "mechanists," such as Nikolai Bukharin (1888–1937), and "dialecticians," led by Abram Deborin (1881–1964), was "resolved" by a ukase from the Communist Party. The episode served to solidify the typical Soviet way of "philosophizing": The last appeal was not to logic and reason, but to the recorded opinion of the "classics of Marxism-Leninism." The highest authority in interpreting the latter belonged, in turn, to the leadership of the Party. The debate highlighted the paradox encapsulated in the expression "Soviet philosophy." On the one hand, Soviet ideology was based on a *philosophical* theory; on the other hand, this theory was *dogmatically* accepted as the final word in all ultimate matters. As a result, Soviet philosophy was implicitly burdened with the impossible task of reconciling the internal contradictions of Marxism—but only by appeal to Marxist principles themselves. The basic contradiction of the doctrine consisted in the fact that it insisted on the ontological primacy of matter over spirit but at the same time wished to be a theory (i.e., spirit) that changed the material world.

The untenable nature of this exercise did not escape contemporaries. Losev, whose eight volumes published between 1927 and 1930 were the swan song of the philosophical Silver Age, publicly called dialectical materialism a crying absurdity and challenged Soviet Marxists to acknowledge that their professed scientistic rationalism was at bottom as mythological as any theology. His was a lonely voice, however, and it was silenced forthwith by an arrest, confinement at labor camps, and a ban on publishing upon release.

The repressions of the 1930s were the lowest point in the history of philosophy in Russia. The pre-Soviet intelligentsia was either intimidated or physically annihilated. In 1937 Florenskii and Shpet were executed in the Gulag. Russia was being purged of its philosophy. To train the new cadre was the task of the recently established Institute of Philosophy in Moscow. There were some attempts to simulate philosophical activity but they were crude and tendentious beyond redemption. Stalin's chapter on dialectical materialism in the 1938 *Kratkii kurs istorii KPSS* (History of the Communist Party of the Soviet Union [Bolsheviks]; short course) was not a philosophical work; it merely sealed the reduction of philosophy to ideological indoctrination for which epistemological or logical concerns were irrelevant. During Stalin's time the voluntaristic (i.e., ultimately terrorist) component in Russian Marxism overshadowed its other aspects. The three-volume *Istoriia filosofii* (History of philosophy) that appeared in 1940 brought to a simplistic pitch a tradition of interpretation established already by Lenin. The entire history of philosophy was presented as a struggle between "progressive" materialism and "reactionary" idealism. In 1947 Georgii Aleksandrov (1908–1961) published his *Istoriia zapadnoevropeiskoi filosofii* (History of Western-European Philosophy), based on similar principles of analysis.

This *History* figured prominently in Andrei Zhdanov's speech the same year, in which he announced

to the new generation of philosophers the Party's orders to be "more creative." Zhdanov's admonitions had a certain positive effect: For the first time since the 1920s the history of Russian philosophy, for example, became a legitimate subject. To bolster Marxism-Leninism's pedigree Soviet authors ingeniously discovered materialism in the ideas of Russian thinkers. Radishchev, Herzen, and Belinskii were recruited into the ranks of Lenin's precursors. Pisarev, Dobroliubov, and Chernyshevskii were, somewhat more justifiably, painted as "revolutionary democrats" and their materialism as a spontaneous discovery of truth prefiguring "scientific socialism." Tendentious as it was, this work was a step forward from the previous period of forced oblivion.

In the meantime philosophers of non-Soviet orientation continued to write privately "into the drawer." After his release from the camps Losev wrote treatises on ancient mythology and aesthetics, as well as philosophical prose. His fellow-survivor from the Silver Age Mikhail Bakhtin (1895–1975) worked on his theories of literature and culture; and Vernadskii developed his doctrine of the noosphere. The ideas of these authors became known only decades later when their works contributed to the intellectual ferment of the 1960s–1980s.

After Stalin's death in 1953 and especially after Nikita Khrushchev's 1956 official condemnation of Stalin's "personality cult" a "thaw" began during which ideological constraints on philosophy were gradually loosened. Khrushchev made an attempt to boost slipping enthusiasm for communism by adopting a new program for the Party but the effect of its exorbitant promises was cynicism rather than renewed optimism. Leonid Brezhnev and the new generation of leaders who came to replace Khrushchev were even less capable of reviving the decaying ideology and from the late 1960s a period of ever deepening disillusionment set in that eventually led to the fall of the Soviet Union in 1991.

In the 1950s and 1960s Soviet philosophy became an increasingly complex agglomeration of disciplines and approaches. The list of permissible themes gradually expanded. The precept, for example, of sacrificing the individual to the needs of the socialist state began to be revised as the human person was cautiously explored as a philosophical subject. Debates on the nature of philosophy ended, thankfully, in an ambiguity as to whether it was a science, theory of action, or world view. The discussion of materialism and dialectics likewise led to a number of diverging positions that included even disagreement with Lenin. Restrictions were still in place and it was impossible to challenge official orthodoxy directly but attempts to solve its problems objectively tended to water down and sometimes even to dissolve its basic precepts. Some philosophers sought refuge from ideology in such relatively neutral areas as philosophy of science, logic, and other formal pursuits that became possible since the late 1940s. Formal logic was somewhat buttressed by the growing prestige of science and technology. Although difficult and limited, exchanges with the outside world gradually expanded through translations, visits, and conferences. Conversely, the work of some Soviet philosophers, such as the semiotician Iurii Lotman (1922–1993) and his colleagues in the so-called "Moscow-Tartu School," found international recognition.

From the mid-1950s on some pre-Stalin figures reemerged. In Losev's prodigious output from 1953 to the time of his death in 1988 the partial truths of Marxism found their place among the broader principles of a phenomenologically modified Christian neoplatonism. Bakhtin's dialogic theories of culture, literature, and the (moral) self similarly rested on philosophical foundations that were sufficiently deep not to be perverted by adaptation to Soviet censorship. Unlike Losev who remained virtually unknown outside Russia, Bakhtin has become a towering presence in the western humanities.

The reappearance of these and other authors demonstrated that communism had not destroyed the continuity of the Russian intellectual tradition. This was largely due to Russian classical literature that remained even in the worst of times the backbone of all humanistic learning and education in Russia. The other key factor was the "Aesopian" writing, stemming from nineteenth-century polemics, by which philosophers masked (transparently enough for the reader to grasp) the true principles behind their critique of philosophy, art, religion, and culture. Losev delivered, for example a blistering critique of modernity in his *Estetika Vozrozhdeniia* (The aesthetics of the renaissance, 1978) that was tacitly based on an Eastern Orthodox view. A similar line of thought was pursued by younger philosophers such as Piama Gaidenko (b. 1934), Iurii Davydov (b. 1929), and Sergei Averintsev (1937–2004).

Characteristically, the most gifted among the newer generation of philosophers had to abandon classical Marxist materialism. Eval'd Il'enkov (1924–1979) and Merab Mamardashvili (1930–1990) exemplified opposing positions on the dialectical method, almost Hegelian in Il'enkov's case and almost openly neo-Kantian in Mamardashvili's. Yet another, mathematical-formalist, argument against the officially accepted dialectical materialism was developed by Aleksandr Zinov'ev (b. 1922),

who eventually had to emigrate and became a well-known writer.

In the 1970s and early 1980s censorship became more lax, allowing—although not without a struggle—the works of, and about, such authors as Solov'ëv and Fedorov to be published. Alongside Vernadskii's ideas about the noosphere, Fedorov's doctrine of the "common cause" served as an inspiration for the loosely defined, nonofficial movement of cosmism. The latter was merged with Eurasianism by Lev Gumilev (1912–1992) who proposed a theory of ethnogenesis as a process affected by cosmic energy. All this signified a halting but perceptible expansion of the boundaries of philosophical discourse that increasingly weakened the hegemony of dogmatic Marxism.

Following Mikhail Gorbachev's reforms from the mid-1980s and until the dissolution of the USSR the hegemony of Marxism rapidly evaporated. One of the leading authors in official Marxism, Ivan Frolov (b. 1929) admitted in his study *Chelovek, nauka, gumanizm: novyi sintez* (Man, science, and humanism: A new synthesis, 1986) that the truths of Marxism were not, after all, absolute. The admission was an attempt to preserve the relevance of the doctrine in the new situation. The history of Soviet Marxism came to an end when the floodgates that held back previously suppressed philosophical literature, both Russian and foreign, finally opened. The return of the works of prerevolutionary and émigré Russian philosophers was the most remarkable part of this revival.

RUSSIAN PHILOSOPHY ABROAD. With the emigration after the 1917 Revolution and the expulsion of a large group of thinkers in 1922 Russian philosophy split into two strikingly unequal branches: the one inside and the other outside the country. The Bolshevik government's intolerance proved to be a blessing in disguise. While all independent philosophical thought was brutally suppressed in the Soviet Union, many of Russian philosophers abroad created the largest and the best part of their oeuvres. This was true of Berdyaev, Frank, Bulgakov, Shestov, and Il'in, as well as of the younger generation of philosophers among whom Georgii Florovskii (1893–1979) and Karsavin deserve special note. Russian thinkers in exile collectively created a body of literature that fulfilled the promise of the Silver Age as the Russian "religious-philosophical renaissance." A comprehensive evaluation of this literature remains a task for the future.

Among the diverse trends that existed in Russian philosophy abroad two seem particularly notable from today's point of view: religious-philosophical and Eurasianist. The first was the continuation of the prerevolutionary religious idealism, whereas the second became yet another refraction of the old theme of Russia's destiny in a new situation created by the Bolshevik revolution. Berlin and then Paris were the centers of the first trend and Prague (as well as, briefly, Sofia), of the second.

Russian religious philosophy continued its preexile themes: critique of (Western) rationalism and the quest for integral knowledge; metaphysics of all-unity and sophiology; Russia's historical destiny cast in religious-idealist terms; and religious foundations of personhood.

The study of the history of Russian philosophy by this group became the culmination of the work begun in Russia. Zen'kovskii's two-volume *Istoriia russkoi filosofii* (History of Russian Philosophy, 1948–1950), Losskii's book of the same title (1951), Berdyaev's aforementioned essay on the Russian Idea, and Florovskii's *Puti russkogo bogosloviia* (Ways of Russian Theology, 1939) were towering achievements supplemented by numerous articles and essays by other authors. Their work was, collectively, the most important philosophical attempt to make sense of the Russian experience and especially of its last, vastly tragic phase. It is surprising how little would need to be changed, for example, in Frank's essay *Krushenie kumirov* (The Collapse of Idols, 1923), created before Stalin's repressions and World War II, if it were to be rewritten today.

Eurasianism began as a distinct movement with the publication of a collection titled *Iskhod k Vostoku* (Exodus to the East; Prague, 1921). It viewed Russia as straddling Europe and Asia in the geographic, geopolitical, and cultural-historical sense and enhanced the traditional Slavophile critique of the West by the Spenglerian sense of the "twilight" of Europe. Postcolonialist critique of Europe was prefigured in Eurasianism's claim that the western view of history merely promoted the West's ulterior interests under the guise of objective truth. The mistrust of the West was supplemented by the affirmation of the positive significance of the Asian element in Russian history and culture. Eurasianism had both a religious and a secular branch. The former was represented by such authors as Petr Savitskii (1894–1968) and Petr Suvchinskii (1892–1985), the latter by Florovskii and Karsavin. In Karsavin's case Eurasianism had a close affinity with the Solov'ëvian school. Nevertheless, for most Eurasianists religion was important only as a cultural-historical factor that contributed to the formation of Russia as a Eurasian entity. The religious theme in Eurasianism weakened especially after Florovskii left the movement. Some of his

secular opponents went so far as collaborating with the Bolshevik government that they saw as the heir to the cause of great Russian statehood. Those who returned to Russia, however, perished eventually in Stalin's concentration camps. Eurasianism as a political movement declined in the mid-1930s with the rise of National Socialism in Germany. Many of its members made significant contributions to the social and human sciences: George Vernadsky (1887–1973) in history, Nikolai Trubetskoi (1890–1938) and Roman Jakobson (1896–1982) in linguistics. Suvchinskii was a prominent musical critic. The political influence of Eurasianist ideas was restored to life in the post-Soviet period when they became a source of inspiration for a widely divergent spectrum of ideological schools of thought, ranging from nationalists dreaming of a new Russian Empire to Soviet-style Communists.

POST-SOVIET PERIOD. Rather than being resolved, philosophical questions were merely suspended by the ideological freeze during the Soviet period and once constraints fell old divisions quickly reemerged. During the early and mid-1990s Russian philosophers were primarily occupied with bringing back formerly suppressed patrimony and rejoining the international philosophical dialogue. Berdyaev, Bulgakov, and Florenskii's writings were particularly favored during this period. But the list was quickly expanded to include the entire galaxy of Silver Age thinkers.

The second tendency—that is, restoration of contacts with the outside world—has by now resulted in a full spectrum of western and nonwestern influences without any apparent restrictions. Like several times earlier in history, Russian philosophers eagerly acquaint themselves with foreign philosophy: phenomenology, analytic philosophy, psychoanalysis, critical theory, poststructuralist thought, and a variety of nonwestern wisdom traditions. The old controversy between Slavophiles and Westernists was also apparently merely suppressed and has again become a notable factor in Russians' debates about their past, present, and future. The theme of the Russian Idea has returned in all of its prior permutations and now has been co-opted, among others, by communist authors who try to breathe new life into a doctrine that has lost much of its appeal. While the tradition of nonacademic philosophizing remains strong, the academy is now the backbone of philosophical life in Russia.

Soviet institutions, such as the Institute of Philosophy of the Russian Academy of Sciences and the journal *Voprosy filosofii* (Questions of Philosophy), have survived their original ideological functions. The teaching of philosophy in higher education occupies the same place as elsewhere in the world and occurs without any ideological constraints. Literature for instruction in philosophy figures prominently among philosophical publications. The current output of academic philosophers embraces all disciplines of philosophy and represents all shades of opinion one finds elsewhere. Numerous works are published on the history of Russian philosophy; they include both special studies and historical surveys. Another notable feature is the striking decentralization of philosophical life that is no longer confined to "the capitals" but is active in many centers of higher learning in the country. There are no overwhelming political parties among Russian thinkers of the early twenty-first century. Neither the surviving communism nor the revived nationalism seem to hold commanding heights. If there is a threat to philosophy today it comes not from the state or radical ideology but from different quarters. Russian philosophy has joined contemporary western and nonwestern philosophical traditions in surviving the onslaught of mass culture. The new freedom and the rich intellectual, artistic, and literary legacy encourage hope, however, that Russian philosophy will rediscover not only its roots, but also the creative inspiration from which it first sprang.

See also Bakhtin, Mikhail Mikhailovich; Bakhtin Circle, The; Bakunin, Mikhail Aleksandrovich; Belinskii, Vissarion Grigor'evich; Berdyaev, Nikolai Aleksandrovich; Bulgakov, Sergei Nikolaevich; Chernyshevskii, Nikolai Aleksandrovich; Chicherin, Boris Nikolaevich; Dostoevsky, Fëdor Mikhailovich; Fëdorov, Nikolai Fëdorovich; Florenskii, Pavel Aleksandrovich; Florovskii, Georgii Vasil'evich; Frank, Semën Liudvigovich; Herzen, Alexander Ivanovich; Ivanov, Viacheslav Ivanovich; Karsavin, Lev Platonovich; Kavelin, Konstantin Dmitrievich; Khomiakov, Aleksei Stepanovich; Kireevskii, Ivan Vasil'evich; Kozlov, Aleksei Aleksandrovich; Kropotkin, Pëtr Alekseevich; Lavrov, Pëtr Lavrovich; Lenin, Vladimir Il'ich; Leont'ev, Konstantin Nikolaevich; Losev, Aleksei Fëdorovich; Losskii, Nikolai Onufrievich; Lotman, Iurii Mikhailovich; Lunacharskii, Anatolii Vasil'evich; Mamardashvili, Merab Konstantinovich; Mikhailovskii, Nikolai Konstantinovich; Pisarev, Dmitri Ivanovich; Plekhanov, Georgii Valentinovich; Rozanov, Vasilii Vasil'evich; Shestov, Lev Isaakovich; Shpet, Gustav Gustavovich; Skovoroda, Grigorii Savvich; Solov'ëv (Solovyov), Vladimir Sergeevich; Tolstoy, Lev Nikolaevich; Trubetskoi, Evgenii Nikolaevich; Trubetskoi,

Nikolai Sergeevich; Trubetskoi, Sergei Nikolaevich; Zen'kovskii, Vasilii Vasil'evich.

Bibliography

PRIMARY SOURCES

Edie, J. M., J. P. Scanlan, and M. B. Zeldin, eds. *Russian Philosophy*. 3 vols. Knoxville: University of Tennessee Press, 1976.

Leatherbarrow, W. J., and D. C. Offord, eds. *A Documentary History of Russian Thought: From the Enlightenment to Marxism*. Ann Arbor, MI: Ardis, 1987.

Russian Studies in Philosophy (1962 to 1992 *Soviet Studies in Philosophy*). A quarterly translation journal.

HISTORIES

Berdyaev, N. A. *The Russian Idea*. Translated by M. French. Hudson, NY: Lindisfarne Press, 1992.

Copleston, F. *Philosophy in Russia: From Herzen to Lenin and Berdyaev*. Notre Dame, IN: Notre Dame University Press, 1986.

Florovsky, G. *Ways of Russian Theology*. Translated by Robert Nichols. Belmont, MA: Nordland, 1979.

Kuvakin, V., ed. *A History of Russian Philosophy: From the Tenth to the Twentieth Centuries*. Buffalo: Prometheus Books, 1994.

Lossky, N. O. *History of Russian Philosophy*. New York: International Universities Press, 1951.

Masaryk, T. G. *The Spirit of Russia: Studies in History, Literature and Philosophy*. 3 vols. Translated by E. Paul and C. Paul, with additional chapters and bibliographies by J. Slavik. London: Allen and Unwin, 1955.

Scanlan, J. P., ed. *Russian Thought after Communism: The Recovery of a Philosophical Heritage*. Armonk, NY: M. E. Sharpe, 1994.

Walicki, A. *A History of Russian Thought from the Enlightenment to Marxism*. Translated by H. Andrews-Rusiecka. Stanford, CA: Stanford University Press, 1979.

Zen'kovskii, V. V. *A History of Russian Philosophy*. Translated by George Kline. New York, Columbia University Press, 1953; London: Routledge, 2003.

REFERENCE LITERATURE

Alekseev, P. V., ed. *Filosofy Rossii XIX–XX stoletii: biografii, idei* [The philosophers of Russia in the 19th–20th centuries: Biographies, ideas]. Moscow: Akademicheskii Proekt, 1999.

de Lazari, A., ed. *Idee w Rosji. Idei v Rossii. Ideas in Russia*. Warsaw: Semper, 1999.

Maslin, M., ed. *Russkaia filosofiia: Slovar'* [Russian philosophy: A dictionary]. Moscow: Terra, 1999.

Vladimir L. Marchenkov (2005)

RUYSBROECK, JAN VAN
(1293–1381)

Jan van Ruysbroeck, the Flemish mystic, was born in the village of Ruysbroeck, near Brussels. He stood in close relation to German contemplatives of the period, notably Meister Eckhart. In 1343 Ruysbroeck, together with two others, established a community at Groenendael that ultimately came under Augustinian rule. He was the prior of this community.

Ruysbroeck was not a trained theologian and had an imperfect knowledge of Latin. Though he made use in his mystical writings of language drawn from Eckhart, such as the "birth of Christ in the soul" and the "eternal Now," he was sensitive to the kind of allegations of pantheism encountered by Eckhart and in fact directed against Ruysbroeck by Jean de Gerson. In his later writings in particular Ruysbroeck made it clear that he did not believe in the identification of the soul with God in the mystical state, and he criticized those contemplatives who gave up the active life and lapsed into quietism. He thus evolved a practical account of contemplation that connected it with good works.

Ruysbroeck distinguished between different phases of the good life, which should be practiced together. First, there is the active life of doing good works. This by itself will not bring blessedness, since it can mean moral self-reliance rather than dependence on God's grace. But good works are a necessary part of the purification of the soul. Second, there is the practice of the inner virtues—faith, hope, and love. Third, there is the contemplative life, through which the soul may gain union with God. Those who attain this last condition are called "God-seeing." They are not continually immersed, as it were, in this inner blessedness, but find themselves impelled to practice love and good works as a result of it. The practice of good works, suffused by the knowledge of God gained in the state of contemplative union, is what Ruysbroeck referred to as "the common life." This, the ideal he tried to realize in his own monastic community, was interpreted as a reflection of the life of the Trinity, which was united in a common fruition analogous to that enjoyed by the mystic but was also outward-going through the creative power of God, analogous to the work of the monk in serving the society around him.

In order to illustrate the relation of union, yet difference, between the soul and God, Ruysbroeck made use of analogies drawn from human love, as the title of his major work, *The Adornment of the Spiritual Marriage*, indicates. Thus one should "rest in Him whom one enjoys.… There love has fallen in love with the lover, and each is all to the other, in possession and in rest" (*The Sparkling Stone,* 13). The love analogy had a certain aptness in bringing out both the sense of union and the necessary theistic distinction between the soul as creature

and the Creator. Ruysbroeck also made use of the Neoplatonic doctrine of eternal archetypes or forms, existing in God. Thus the ground of the soul is man's eternal archetype, and in realizing it in its purity and nakedness, the contemplative finds union with God. In this, Ruysbroeck, like other mystics of the period, exhibited the influence of Pseudo-Dionysius. He thus made use too of the notion that creatures proceed from God through the process of creation and return to him through contemplation. But since the creature needs to reflect the love displayed by God in the work of creation, so likewise the mystic must combine his return to God with the outgoing work of love.

Ruysbroeck's works were closely studied by those who belonged to the movement known as the Brethren of the Common Life, started in the latter part of the fourteenth century by Gerhard Groot, who knew Ruysbroeck. Thomas à Kempis belonged to this confraternity. Despite contemporary criticisms of his language as not always squaring with orthodox theology, Ruysbroeck was beatified by the Roman Catholic Church.

See also Augustinianism; Eckhart, Meister; Gerson, Jean de; Mysticism, History of; Pseudo-Dionysius; Thomas à Kempis.

Bibliography

The definitive edition is B. Ponkens and others, eds., *Ruusbroec Werken,* 2nd ed., 4 vols. (Cologne, 1950). English translations include a one-volume edition of *The Adornment of the Spiritual Marriage, The Sparkling Stone,* and *The Book of Supreme Truth,* translated by C. A. Wynschenk Dom and edited with an introduction and notes by Evelyn Underhill (London: Watkins, 1951), and *The Spiritual Espousals,* translated by E. Colledge (London: Faber and Faber, 1952). See also Ray C. Petry, *Late Medieval Mysticism* (Philadelphia: Westminster Press, 1957).

Ninian Smart (1967)

RYLE, GILBERT

(1900–1976)

Gilbert Ryle, the British philosopher, was born in Brighton. Having read Classical Honour Moderations and the Final School of Literae Humaniores (Greats) he went on to read the then newly established School of Philosophy, Politics and Economics at the Queen's College, Oxford. He became a lecturer at Christ Church in 1924 and in the following year a student and tutor, and he remained there until his appointment as professor at the end of World War II. He was the Waynflete professor of metaphysical philosophy in the University of Oxford from 1945 to 1968. Ryle was largely responsible for the institution of the new degree of bachelor of philosophy at Oxford. He served as the editor of *Mind,* after the retirement of G. E. Moore, from 1947 until 1971.

Ryle's philosophical writings covered a wide range of topics. They fall mainly within the fields of philosophical methodology, philosophical logic, and the philosophy of mind, but the total spread is very wide and includes some work on the history of philosophy, especially on Plato. Only the fields of moral philosophy, political philosophy, and aesthetics are comparatively neglected. Much of his writing takes the form of articles addressed to the solution of quite specific issues, and it is impossible to discuss here seriatim his "Negation," "Plato's *Parmenides,*" "Conscience and Moral Conviction," and "Heterologicality," to mention the titles of only four such papers.

Probably the best approach to Ryle's philosophical work is through his views on the nature and method of philosophy, which have developed in a consistent way after the end of a short and early flirtation with phenomenology. Many of his articles on specific topics seem to have a clear subordinate aim of illuminating these questions, while such important writings as "Systematically Misleading Expressions," his inaugural lecture, *Philosophical Arguments,* and the book *Dilemmas* are explicitly devoted to them. That *The Concept of Mind* can be regarded as an illustration of his views on philosophical method is a tribute to the consistency of his theory with his practice, though it would be an injustice to treat it merely as such.

Ryle's well-known article "Systematically Misleading Expressions" is important as being easily the first, although incompletely worked out, version of a view of philosophy closely akin to that which Ludwig Wittgenstein was then beginning to work out independently, and which is often spoken of as having been first suggested by Wittgenstein. This view treats philosophy as the activity of removing fundamental conceptual confusions that have their source in our overreadiness to construe grammatical similarities and differences as indicative of logical similarities and differences. For example, since either unpunctuality or the unpunctual Smith may, with grammatical similarity, be said to be reprehensible, some philosophers are inclined to conclude that similar things are being said of two objects, Smith and unpunctuality; hence, the world is thought to be populated by two kinds of objects, universals and particulars. Again, since "Mr. Baldwin is a statesman" is grammatically similar to "Mr.

Pickwick is a fiction," philosophers have been tempted to suppose that the world contained fictions alongside of statesmen.

However, Ryle's view is not fully worked out at this stage. Writing in a climate of opinion in which philosophy was widely regarded as the activity of analysis by which the true logical form of facts was explicitly displayed and the test of adequate language was taken to be a one-to-one correspondence with the form of facts, he did not entirely free himself from its influence. As a result, he cannot regard the reformulation of statements in a way that removes misleading grammatical similarities as merely a useful expedient for making ourselves aware of important differences between them; the reformulation is still thought of as the revelation of the true form of the fact, so that "Baldwin is a statesman" is, in an absolute sense, a correct form of utterance, while "Pickwick is a fiction" is incorrectly formulated.

This anomalous relic of logical atomism caused Ryle uneasiness even then, and it does not appear again. If we neglect it, we may regard "Systematically Misleading Expressions" as an exposition of a view that Ryle never abandoned, although he did refine it. One such refinement is found in *Dilemmas.* Here it is claimed that many philosophical problems, if not all, immediately present themselves in the form of dilemmas: We find ourselves holding, without the possibility of sincere repudiation, two or more opinions that seem to be incompatible (that, for example, we often choose responsibly what to do, and that we are what we are through our natural endowment as modified by environment—the problem of free will). Such dilemmas must be overcome by showing that the apparent conflict is a consequence of conceptual confusion rather than by choosing one horn on which to be impaled.

The emphasis is somewhat different in *Philosophical Arguments.* While in "Systematically Misleading Expressions" and in *Dilemmas* the emphasis is on the activity of freeing ourselves from conceptual errors and puzzlement, in *Philosophical Arguments* the more constructive side of the procedure is stressed. By methodically determining what can and what cannot be said without absurdity, which inferences are valid and which are invalid, which grammatical parallels are likely to mislead and which are not, we come to see better the "logical geography" of our conceptual system—how different concepts are related to each other and what are the different roles that they play. There is no essential conflict between the view of Ryle's philosophical procedures as "removing conceptual roadblocks" and "freeing conceptual traffic jams," to echo the metaphor employed in *Dilemmas,* and the more constructive view of them. Thus, it would be idle to ask whether, or at which stages, *The Concept of Mind* is correctly viewed as exposing the confusion of "the ghost in the machine," into which we are led by grammatical analogies, or as mapping the extension and boundaries of such interrelated concepts as "will," "intelligence," "imagination," "thought," and the like; the two aspects are not thus separable.

Ryle often expressed this view of philosophy in terms of the notion of a category mistake, as in *The Concept of Mind.* A category mistake occurs when something is taken to belong to a different category from its true one. Neither in *The Concept of Mind* nor elsewhere is any serious attempt made by Ryle to give a rigorous account of the notion of a category, although there is a historical discussion of it in "Categories," and Ryle sees this notion as akin to Bertrand Russell's notion of type. Although this is a gap, it is probably of little direct importance to the argument of *The Concept of Mind.* The essential thesis here is that there is a special kind of confusion that can be illustrated by that of taking team spirit as an element in a game as being on equal footing with serving or receiving, of taking a division as a military formation as being on equal footing with its component regiments, of taking Oxford University as an institution as being on equal footing with its component colleges. Ryle then goes on to claim that traditional Cartesian dualism treats the mind as an entity on equal footing with the body and mental activities as being on equal footing with bodily activities, and that this is a confusion of the same kind as those in the three illustrative cases. The language of category mistakes is not essential; Ryle could have used his terminology of 1931 and said that just as the grammatical similarity of "Jones gave an exhibition of dribbling" and "Jones gave an exhibition of ball control" could mislead us into thinking that Jones was giving two independent and simultaneous exhibitions, so the grammatical similarities between our talk of mental and bodily activities could mislead us into thinking that they were independent and simultaneous activities.

Such a misconception Ryle calls the dogma of the ghost in the machine. He attempts to show its falsity in a series of chapters on the main aspects of mental life, in which the arguments fall into two main classes. On the one hand he tries to show that the dogma of the ghost in the machine fails in its explanatory task and is logically incoherent, leading to such logical evils as vicious infinite regresses. On the other hand he tries to show that a satisfactory positive account of mental phenomena can be

given, without invoking the ghost, in terms of such things as style of performance, dispositions to certain characteristic performances, and acquired skills. Thus, if a person does a physical action while thinking about what he is doing, we must take it not that the ghost discursively thinks and the bodily machine moves but that the person performs bodily in an appropriate way, while being disposed to perform other actions if the occasion arises.

One chapter in the book is a restatement of the argument of the paper "Knowing How and Knowing That," published in 1946, and it is a plausible inference that this paper was the germ from which the larger enterprise sprang. In that article Ryle suggested that philosophers commonly take it that knowing how to do something is knowing the truth of certain principles and applying them to an activity. He pointed out that although a given cook may learn to cook from a cookbook, the principles of cookery are logically a distillation from the practice of those who know how to cook, just as the principles of valid argument are a distillation from the practice of those who know how to argue. Thus, knowing how to do things, being able to perform intelligently, is logically independent of any interior theorizing; therefore it involves a display of intelligence that others can witness, rather than a mechanical event from which we have to infer a piece of unwitnessable ghostly theorizing. *The Concept of Mind* attempts to extend the same line of thought to other mental phenomena.

It should be noted that Ryle is not content with the "weaker" thesis that overt human actions must not be analyzed as mechanical events brought about by nonphysical, ghostly activities. In fact, he adopts the far stronger thesis that all references to the mental must be understood in terms of, in principle, witnessable activities. We must not only avoid ascribing the skill of a skillful driver to a ghostly "inner" driver, but we must also explain all mental life, including emotion and feeling, in terms of the witnessable. Certainly it is this feature of his book that has led many, with considerable plausibility, to class Ryle as a philosophical behaviorist, though he repudiated this label in advance. Ryle, indeed, sometimes refers to "twinges," "throbs," "flutters," and "glows" in his characterization of feelings in a way hard to reconcile with behaviorism, but it is notoriously difficult to see how such terms are not a relic of the essentially private in Ryle's public world. By adopting this stronger thesis Ryle avoids well-known difficulties about knowledge of other minds and privacy; however, it is not clearly required for the basic program, and much that he has to say is independent of it.

Much of the interest of this modern classic is independent of the question whether Ryle succeeds in demonstrating any general thesis. The detailed discussions of thinking, knowledge, will, emotion, sensation, intellect, and the like have great independent interest. In the course of these discussions Ryle introduces a number of philosophical distinctions, such as those of "task and achievement," "avowal," and "mongrel-categorical," that have become the common tools of modern philosophical discussion. The whole character of philosophical discussion of the mind has been decisively changed, even in quarters where Ryle's conclusions are strongly challenged, by the appearance of *The Concept of Mind.*

Another set of problems to which Ryle devoted a number of papers are those concerned with the concept of meaning. Here his review of Rudolf Carnap's *Meaning and Necessity* in *Philosophy,* his "The Theory of Meaning," published in *British Philosophy in the Mid-century,* and his contribution to the symposium "Use, Usage and Meaning" in the *PAS* supplementary volume for 1961 deserve special mention. One main contention in these articles is that it is words that are the bearers of meaning, and whose meanings have to be taught and learned, rather than sentences. To learn a language is to acquire a vocabulary and a syntax; this language is then used in speech, which is an activity that one performs by means of a language. The sentence is a unit of speech, not of language. The theory of meaning is therefore concerned primarily with words, not with sentences; but this theory, Ryle holds, has been often vitiated by a simple model of meaning that he calls the "'Fido'-Fido" theory, one that seeks always to find as the meaning of a word something that stands to that word rather as the dog Fido stands to the name "Fido." J. S. Mill partly emancipated himself from the theory by distinguishing connotation from denotation, but he continued to say that meaning was connotation and denotation. In the review of Carnap mentioned before, Ryle attempts to show that the "'Fido'-Fido" theory is still not an outworn fallacy but something that continues to vitiate much sophisticated modern work.

It is notable that the bulk of Ryle's philosophical writing avoids, rather than lacks, any historical discussion. There is the very minimum of reference to even recent learned controversy, and the great philosophers are rarely given even a casual mention. In *The Concept of Mind* the expression "Cartesian dualism" is a nickname for a kind of view that Ryle had once held and to which he thinks many are prone rather than a genuine historical reference. However, this is a policy of segregating the his-

tory of philosophy from the treatment of problems, not a sign of lack of interest in the history of philosophy. Ryle's historical interests, though eclectic, are wide. They have, however, centered on Plato; in addition to already published articles on Plato, Ryle devoted much work to problems arising from the Platonic dialogues, and further publications in that field may be expected.

In conclusion, a word should be said about Ryle's highly individual style, for it is of more than literary interest. It is peculiarly his own, so that it would be impossible for anyone familiar with it not to recognize his work from even a few sentences. One hallmark is the freshness of the vocabulary; although he liberally coined technical terms when he needed them, he always avoided the well-worn counters of philosophical exchange. Another hallmark is that although the general style is informal, the choice of words is literary rather than colloquial; this is achieved by the use of a vocabulary more novelistic than learned. Although there is much close argument in his writing, the importance of the fresh language, the bold metaphor, and the terse epigram in giving the problem a striking presentation, in bringing down pretentious castles of learned jargon, and in making his own contention memorable is very great indeed.

See also Analysis, Philosophical; Artificial and Natural Languages; Behaviorism; Carnap, Rudolf; Categories; Meaning; Mill, John Stuart; Moore, George Edward; Language, Philosophy of; Philosophy of Mind; Plato; Propositions, Judgments, Sentences, and Statements; Russell, Bertrand Arthur William; Thinking; Type Theory; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY RYLE

Review of Martin Heidegger's *Sein und Zeit. Mind* 38 (1929).
"Negation." *PAS* supp. 9 (1929).
"Are There Propositions?" *PAS* 30 (1929–1930).
"Systematically Misleading Expressions." *PAS* 32 (1931–1932).
"Phenomenology." *PAS* supp. 11 (1932).
"Imaginary Objects." *PAS* supp. 12 (1933).
Locke on the Human Understanding. Oxford: Oxford University Press, 1933.
"About." *Analysis* 1 (1933).
"Internal Relations." *PAS* supp. 14 (1935).
"Mr. Collingwood and the Ontological Argument." *Mind* 44 (1935).
"Unverifiability-by-Me." *Analysis* 4 (1936).
"Induction and Hypothesis." *PAS* supp. 16 (1937).
"Taking Sides in Philosophy." *Philosophy* 12 (1937).
"Categories." *PAS* 38 (1937–1938).
"Plato's *Parmenides.*" *Mind* 48 (1939).
"Conscience and Moral Convictions." *Analysis* 7 (1940).
Philosophical Arguments. Oxford: Clarendon Press, 1945. Inaugural lecture.
"Why Are the Calculuses of Logic and Mathematics Applicable to Reality?" *PAS* supp. 20 (1946).
Review of Marvin Farber's *Foundations of Phenomenology. Philosophy* 21 (1946).
"Knowing How and Knowing That." *PAS* 46 (1946–1947).
Review of Carnap's *Meaning and Necessity. Philosophy* 25 (1949).
The Concept of Mind. London: Hutchinson, 1949.
"'If,' 'So' and 'Because.'" In *Philosophical Analysis,* edited by Max Black. Ithaca, NY: Cornell University Press, 1950.
"Logic and Professor Anderson." *Australasian Journal of Philosophy* 28 (1950).
"Thinking and Language." *PAS* supp. 25 (1951).
"Feelings." *Philosophical Quarterly* 1 (1951).
"Heterologicality." *Analysis* 11 (1951).
"The Verification Principle." *Revue international de philosophie* 5 (1951).
"Ludwig Wittgenstein." *Analysis* 12 (1951).
Contribution to *The Physical Basis of Mind,* edited by Peter Laslett. Oxford: Blackwell, 1950.
"Ordinary Language." *Philosophical Review* 62 (1953).
"Thinking." *Acta Philologica* 9 (1953).
"Pleasure." *PAS* supp. 28 (1954).
"Proofs in Philosophy." *Revue internationale de philosophie* 8 (1954).
Dilemmas. Cambridge, U.K.: Cambridge University Press, 1954.
Introduction to *The Revolution in Philosophy,* edited by D. F. Pears. New York: St. Martin's Press, 1956.
"Sensation." In *Contemporary British Philosophy,* edited by H. D. Lewis, third series. London, 1956.
"The Theory of Meaning." In *British Philosophy in the Mid-century,* edited by C. A. Mace. London: Allen and Unwin, 1957.
"Predicting and Inferring." *Colston Papers* 9 (London, 1957).
"On Forgetting the Difference between Right and Wrong." In *Essays in Moral Philosophy,* edited by A. I. Melden. Seattle: University of Washington Press, 1958.
"A Puzzling Element in the Notion of Thinking." *Proceedings of the British Academy* 44 (1958).
Article on Wittgenstein in *Scientific American* (1959).
"Letters and Syllables in Plato." *Philosophical Review* 59 (1960).
Review of Richard Wollheim's *F. H. Bradley. Spectator* (1960).
"Comment on Mr. Achinstein's Paper." *Analysis* (1960).
"Use, Usage and Meaning." *PAS* supp. 35 (1961).
"La phénomenologie contre le *Concept of Mind.*" In *La philosophie analytique,*edited by H. Bréra. Paris, 1962.
"Thinking Thoughts and Having Concepts." *Institut Internationale de Philosophie* (1962).
"Abstractions." *Dialogue* 1 (1962).

J. O. Urmson (1967)

RYLE, GILBERT [ADDENDUM]

From Gilbert Ryle's death in 1976, and through the late 1990s, his views were not the focus of much philosophical attention. Studies of his thought have been published (e.g., Lyons 1980; Stroll 2001) and his character and role in Oxford have been illuminated by the memoirs of others (e.g., Mabbott 1986).

Ryle's own approach to the understanding of psychological concepts was superseded by the emergence of the psycho-physical identity theory, mainly because his analytical concentration on (behavioral) dispositions, of whatever complexity, seemed not to confer a sufficiently real status on lots of psychological processes, for example, feeling a pain or occurrent thinking.

His most lasting intellectual legacy has been the supposed distinction between knowing how and knowing that, which has remained part of philosophical folklore since he propounded it. According to Ryle knowing how to F is distinct from any knowledge that a proposition is true, amounting rather to a capacity to do the action in question. Ryle's distinction has been relied on by those (e.g., David Lewis) who have tried, when answering certain antimaterialist arguments, to give a practical, nonfactualist account of knowing what an experience is like. However, considerable skepticism is being generated about Ryle's distinction (e.g., in Stanley and Williamson 2001; Snowdon 2003). Thus, someone who is injured can know how to do something even though he or she is unable to do it, and much knowhow seems to be knowledge that some way is the way to act. These criticisms have been resisted (e.g., Koethe 2002; Rumfit 2003). There is considerable debate and it remains to be seen whether this aspect of Ryle's thought will suffer the fate of the rest.

See also Behaviorism; Lewis, David.

Bibliography

Koethe, J. "Stanley and Williamson on Knowing How." *Journal of Philosophy* 94 (2002): 325–328.

Mabbott, J. D. *Oxford Memories*. Oxford: Thorntons, 1986.

Rumfit, I. "Savoir Faire." *Journal of Philosophy* 100 (2003): 158–166.

Snowdon, Paul F. "Knowing How and Knowing That: A Distinction Reconsidered." *Proceedings of the Aristotelian Society* 104 (1) (2003): 1–29.

Stanley, Jason, and Tim Williamson. "Knowing How." *Journal of Philosophy* 98 (2001): 411–444.

WORKS ON RYLE

Lyons, William. *Gilbert Ryle: An Introduction to His Philosophy*. Brighton, U.K.: Harvester Press, 1980.

Stroll, Avrum. "Gilbert Ryle." In *A Companion to Analytic Philosophy*, edited by A. P. Martinich and David Sosa. Malden, MA: Blackwell, 2001.

Paul F. Snowdon (2005)

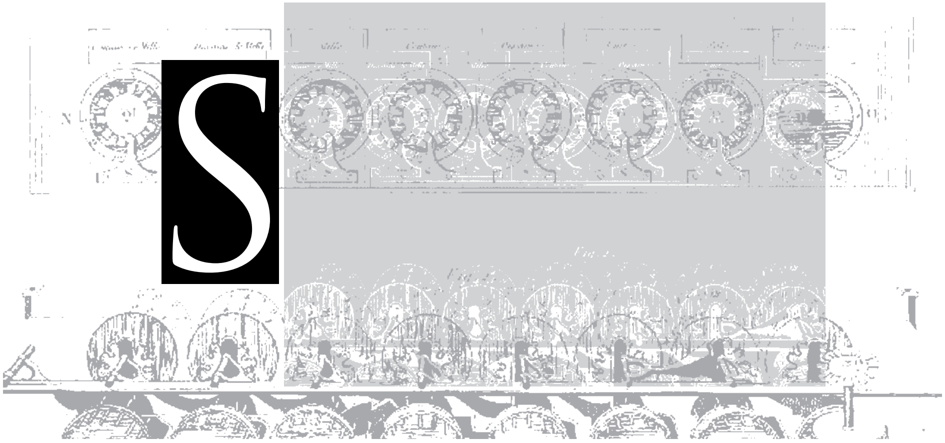

S

SAADYA

(882–942)

Saadya, sometimes called al-Fayyumi from the section of Upper Egypt in which he was born, had a brilliant career as the most distinguished intellectual leader of Jewry in his age. He was twenty-three when he left his Egyptian home to play his part on the wider stage of Palestine, Syria, and Babylonia. By this time he had already composed the first known Hebrew dictionary and an important treatise refuting the views of Anan ben David, the founder of the rationalistic Karaite sect. In 921, the rabbis of Babylonia challenged the authority of the Palestinian rabbis to fix the Hebrew calendar. Saadya's defense of the position of the Babylonian rabbis was most effective; he was rewarded by appointment to the rabbinical academy at Sura in Babylonia; and a few years later, in 928, he was the first non-Babylonian ever to be named as the head (*gaon*) of the academy. His tenure of this position was neither calm nor prolonged. Disputes with the exilarch of the Babylonian Jewish community led to the removal of Saadya and his retirement from active participation in the life of the community. His last years saw a burst of literary creativity.

The writings of Saadya truly signalized the birth of a new creative period in Jewish life. He was a pioneering student and productive scholar in many fields of Jewish concern, including Hebrew grammar and philology, biblical exegesis, and Jewish liturgy. The early attacks on the views of Anan were followed by a long series of writings against Anan's fellow sectarians; since Karaism, a movement that rejected rabbinical and Talmudic law, was at this time the major internal threat to the unity of Jewish life, Saadya's anti-Karaite polemics continued throughout his career. The primary activity of Saadya's public life was in the legal field, and here his contributions were outstanding. In addition to commentaries on Talmudic treatises, Saadya wrote at least ten systematic monographs on a variety of Jewish legal subjects; one of these, *Inheritance,* is preserved in its entirety in the Bodleian Library at Oxford. It was published in 1897 under the editorial care of Joel Mueller. Fragments of others still exist. Saadya was the first to translate the Old Testament into Arabic; this translation, still in use, is notable for its use of paraphrase where a literal translation would have been subject to censure for anthropomorphism. He also composed the earliest known commentary on *Sefer Yetzira* (The book of creation), an important work of the Jewish mystical tradition.

Thus his major philosophical work, *The Book of Beliefs and Opinions* (Arabic title, *Kitab al-ʿamanat waliʾ tikadat*; Hebrew title, *Sefer ha-emunoth weha-deoth*), probably completed in 933, is but one of a long list of eminent contributions for which Saadya is remembered. He was probably impelled toward a systematic consideration of the relation between the religious beliefs of Judaism and the opinions arrived at through rational investigation both by the comparable activities of Muslim philosophers—the *kalam* and other schools—and by the quasi-rational approach characteristic of most of the Karaite spokesmen. In the intellectual milieu of the tenth century, the philosophical issues with which Saadya was concerned were widely and thoughtfully debated. Muslim philosophers of this age had far more of the corpus of Greek philosophical literature available to them than had their compeers in the Christian West. Saadya's *Book of Beliefs and Opinions* may best be described, therefore, as a philosophical apologetics for rabbinite Judaism. The Muʿtazilite school of Muslim philosophers generally presented their systematic treatises in the form of theodicies, treating first of the unity of God and then of his justice. Saadya's philosophical work is similarly patterned but assigns a rather larger share of the discussion to the second, ethical part then to the first, more purely metaphysical and theological one.

Prefaced to the ten sections into which the body of the work is divided is an introductory treatise in which Saadya justifies his engaging in this sort of philosophical enterprise. Here he enters into questions of the sources of human knowledge, the relations of belief and doubt, and the prevalent view that rational speculation necessarily leads to heresy. He argues that not the use of reason, but exclusive dependence on human reason is undesirable. Properly used, in combination with revelation, rational speculation supports revealed religion. From this discussion Saadya moves, in the first major section, to a proof of the doctrine of creation out of nothing and a refutation of twelve contrary views. In the second major section of his book, Saadya discusses the unity of God and demonstrates how the Christian doctrine of the Trinity is based upon a misinterpretation of certain scriptural verses. Treatise three defends the idea of a divine law for God's creatures as a necessary demand of reason and urges the need for prophecy and prophets as the vehicle by means of which the divine law is transmitted to men.

From the fourth treatise to the end of the work, Saadya's concern is more with ethical questions and the consequences for men's future redemption of their obedience or disobedience to the divine precepts delivered by the prophets. In these sections, he defends on rational grounds all of the major doctrines of the Jewish tradition. The tenth and last treatise is of slightly different character; it presents an ethic of the middle way as the proper guide to man's conduct in the affairs of daily life. Thus we may say that Saadya concluded his work on religious philosophy with a secular ethic.

See also Jewish Philosophy; Mysticism, History of.

Bibliography

Les oeuvres complètes de Saadia was edited by J. Derenbourg (6 vols., Paris, 1893–1896). Saadya's *Book of Beliefs and Opinions* may be found in a translation from the Arabic by Samuel Rosenblatt in the Yale Judaica Series, Vol. I (New Haven, CT: Yale University Press, 1948).

For material on Saadya, see Jacob B. Agus, *The Evolution of Jewish Thought* (New York and London: Abelard-Schuman, 1959); Joseph L. Blau, *The Story of Jewish Philosophy* (New York: Random House, 1962); Julius Guttman, *Philosophies of Judaism,* translated by D. Silverman (New York: Holt, Rinehart and Winston, 1964); Isaac Husik, *History of Mediaeval Jewish Philosophy* (New York: Macmillan, 1916); Henry Malter, *Saadia Gaon: His Life and Works* (Philadelphia: Jewish Publication Society of America, 1921); and *Three Jewish Philosophers. Philo, Saadya Gaon, Jehuda Halevi,* selections with introductions by H. Lewy, A. Altman, and I. Heinemann, eds. (New York: Meridian Books, 1960).

J. L. Blau (1967)

SAADYA [ADDENDUM]

882–942

Saadya's intellectual career was characterized by polemic and defense of rabbinic Judaism. As head (*gaon*) of the rabbinic academy of Sura in Babylonia in the first third of the tenth century, Saadya felt compelled to respond to the Karaites, those who challenged the authority of rabbinic Judaism. Saadya's major philosophical work, written in Arabic, *Kitāb al- Āmānāt wa'l-Iʿ tiqādāt* (The book of doctrines and beliefs) should be understood in large part as a defense against the Karaites. Following the Muslim Muʿtazilites, who emphasized divine justice and unity, Saadya's *Āmānāt* focuses on creation, divine unity, divine law and justice, and reward and punishment in this world and posthumously.

Prefacing the particular discussions just noted, Saadya outlines the sources of human knowledge and understanding, and of error. The senses can be unreliable, reason may be derailed if inferential skills are lacking, and overarching all of this is the propensity to impatience in

inquiry. For Saadya, the sources of error and doubt are both intellectual and moral. Human beings are frail creatures, and as a result, out of benevolence God provided humankind with a superhuman source of insight and true belief—the prophetic tradition—encapsulated in scripture and in the oral tradition of the rabbis. In advance of people's (slowly and laboriously) discovering the truth unaided, God has provided them the answers through his prophets. In this way revelation and reason, revealed religion, and rational speculation coincide over time. Revelation provides an anchor for humans as they quest for knowledge. As Alexander Altmann puts it on behalf of Saadya: "Revelation is not essentially superior, but historically prior to Reason and has an educational function in the evolution of humanity" (1946, p. 18 [2002]). Maimonides viewed Saadya's project as less than philosophical, assuming conclusions that ought to be proved. But Saadya's defense against this would be that revelation of the truth provides just the starting point for patient inquiry.

See also Jewish Philosophy.

Bibliography

WORKS BY SAADYA

Altmann, Alexander, tr. *Saadya Gaon, The Book of Doctrines and Beliefs*. Abridged edition. Oxford: East and West Library, 1946. Reprinted with a new introduction by Dan H. Frank. Indianapolis: Hackett, 2002.

WORKS ON SAADYA AND THE BACKGROUND OF HIS THOUGHT

Altmann, Alexander. "Saadya's Conception of the Law." *Bulletin of the John Rylands Library* 28 (1944): 320–339.

Altmann, Alexander. "Saadya's Theory of Revelation: Its Origin and Background." In *Saadya Studies*, edited by E. I. J. Rosenthal, 4–25. Manchester, England: Manchester University Press, 1943.

Brody, Robert. *The Geonim of Babylonia and the Shaping of Medieval Jewish Culture*. New Haven, CT: Yale University Press, 1998.

Davidson, Herbert. *Proofs for Eternity: Creation and the Existence of God in Medieval Islamic and Jewish Philosophy*. 95–106. New York: Oxford University Press, 1987.

Davidson, Herbert. "Saadia's List of Theories of the Soul." In *Jewish Medieval and Renaissance Studies*, edited by Alexander Altmann, 75–94. Cambridge, MA: Harvard University Press, 1967.

Davidson, Herbert. "The Study of Philosophy as a Religious Obligation." In *Religion in a Religious Age*, edited by S. D. Goitein, 53–68. Cambridge, MA: Association for Jewish Studies, 1974.

Goodman, Lenn E., tr. *The Book of Theodicy: Translation and Commentary on the Book of Job by Saadiah ben Joseph al-Fayyumi*. New Haven, CT: Yale University Press, 1988.

Goodman, Lenn. "Maimonides' Responses to Saadya Gaon's Theodicy and Their Islamic Background." In *Studies in Islamic and Judaic Traditions II*, edited by W. M. Brinner and S. D. Ricks, 3–22. Atlanta: Scholars Press, 1989.

Goodman, Lenn. "Rational Law/Ritual Law." In *A People Apart: Chosenness and Ritual in Jewish Philosophical Thought*, edited by D. H. Frank, 109–200. Albany: State University of New York Press, 1993.

Goodman, Lenn. "Saadiah Gaon's Interpretive Technique in Translating the Book of Job." In *Translation of Scripture Jewish Quarterly Review*, 47–76. Philadelphia, 1990.

Stroumsa, Sarah. "Saadya and Jewish *kalam*." In *The Cambridge Companion to Medieval Jewish Philosophy*, edited by D. H. Frank and Oliver Leaman, 71–90. Cambridge, U.K.: Cambridge University Press, 2003.

Wolfson, Harry A. *The Philosophy of the Kalam*. Cambridge, MA: Harvard University Press, 1976.

Wolfson, Harry A. *Repercussions of the Kalam in Jewish Philosophy*. Cambridge, MA: Harvard University Press, 1979.

Daniel H. Frank (2005)

SABATIER, AUGUSTE
(1839–1901)

Auguste Sabatier was perhaps the Protestant theologian most influential in the early twentieth century. Many Catholic modernists as well as Protestant liberals believed that his philosophy of religion had achieved its object, a reconciliation between the essential verities of Christian experience and the demands of science. Sabatier was a professor of reformed dogmatics at Strasbourg and Paris and a sometime journalist and literary critic. He ended his career as dean of the Theological Faculty of Paris.

Sabatier described his theory of religious knowledge as "critical symbolism." By this he meant to indicate that religious doctrine and dogma are attempts to symbolize the primary and eternal religious experience (or consciousness) of the believer. He taught that the doctrines of historical religions are secondary, temporal, and transient symbols of this central religious experience. Christian dogmas, then, are necessarily inadequate attempts to "express the invisible by the visible, the eternal by the temporal, spiritual realities by sensible images." Christ and his disciples through the ages have experienced the divine presence of God the loving Father and with it a sense of moral repentance and an inner energy of the spirit. As with all personal experience, no symbolic structure can act as substitute. Such structures are, in every field, merely hypothetical attempts to grasp experience.

Correspondingly, Sabatier held that the cosmologies, legends, dogmas, and statements about the world and man propagated by historical religions in an attempt to

express and communicate the fact of religious experience can claim only derivative and relative validity. Moreover, they are conditioned by the state of science and philosophy as understood by those who create such religious symbolism. And just as science and philosophy do not give absolute and final truth, neither does religious dogma—hence the decline of older religious symbolism with the progress of science. God lives in man's consciousness, not in dogmas and cosmologies. Man's need for and experience of God's presence prove his existence. Science and philosophy are masters of their own proper domain. Thus, "God is the final reason of everything, but the scientific explanation of nothing."

Sabatier's critical symbolism was exceedingly Protestant in that it rejected Catholic dogmatic absolutism for the absolutism of justification by faith. It appealed to many modern religionists of his day because it seemed to retain valid science and yet avoid positivistic nihilism and agnostic defeatism. Putting personal experience above theories about experience, Sabatier's approach was found congenial in an age that produced Henri Bergson and William James. Like them, Sabatier seemed to give moral claims and value judgments a renewed truth. To know a thing religiously, Sabatier held, is to experience the sovereignty of spirit and to estimate the object known as a means or obstacle to the true moral life of the spirit. Teleology is reintroduced along with objective value, and the meaning of life, as well as the will's freedom to choose good or evil, is made manifest. Sabatier's theories could easily be adapted to the neo-Kantian and neoidealist tendencies at work in philosophy, social science, political ideology, literature, and art in the new century. His continued influence seems assured, for by basing the truth of religion on the personal experience of the believer, he joined the long line of "crisis" and existential theologians of our time.

See also Bergson, Henri; James, William; Life, Meaning and Value of; Neo-Kantianism; Teleology.

Bibliography

WORKS BY SABATIER

Esquisse d'une philosophie de la religion. Paris: Fischbacher, 1897. Translated by T. A. Seed as *Outlines of a Philosophy of Religion.* New York: Harper, 1957.

Essai d'une théorie critique de la connaissance religieuse. Paris, 1899.

Les religions d'autorité et la religion de l'esprit. Paris, 1903. Translated by L. Seymour as *Religions of Authority and Religions of Culture.* London: McClure, Phillips, 1904.

WORKS ON SABATIER

Michalescu, J. *Darlegung und Kritik der Religionsphilosophie Sabatiers.* Bern: Buchdruckerei Scheitlin, Spring 1903.

Vienot, J. *Auguste Sabatier.* Paris, 1927. Contains the best available bibliography.

John Weiss (1967)

SAINT-HYACINTHE, THÉMISEUL DE

(1684–1746)

The real name of Thémiseul de Saint-Hyacinthe, the French freethinker, was Hyacinthe Cordonnier. Born at Orléans, he was unjustly reported to be the son of Jacques Bénigne Bossuet. His ambitious mother induced him to change his name and to become a cavalry officer. Later he devoted himself to the study of ancient and modern languages in Holland, from which he had to flee because of a jealous husband and to which he later returned because he had seduced one of his pupils. He became an editor of the new *Journal littéraire* (1713) and wrote in favor of the moderns. In 1714 his anonymous *Le chef-d'oeuvre d'un inconnu,* a satire of pedantry, won him notoriety. He eloped to London in 1722 with the daughter of a nobleman. He stayed there for twelve years, became a member of the Royal Society, and began a long and gratuitous quarrel with Voltaire, whom he offended in a satirical play (*Déification d'Aristarchus Masso,* 1732). He returned to Paris in 1734 and later moved to Holland, where he died in 1746.

Three of Saint-Hyacinthe's writings are worthy of mention. The first book, *Le chef-d'oeuvre d'un inconnu,* is a bizarre work that could easily be a satire on the *explication de texte* method, as it is practiced in some milieus. His last book, *Recherches philosophiques sur la nécessité de s'assurer par soi-même de la vérité* (1743), is a defense of the power of reason to find truth and of its right to do so. He also argues for the moral-sense theory, with which he probably became familiar during his stay in England. His discussion of words as signs of ideas points toward linguistic analysis. Other chapters deal with demonstration and evidence, matter and the soul.

In between these two works, Saint-Hyacinthe wrote his interesting *Lettres écrites de la campagne* (1721). This potpourri is a long conversation treating of many subjects, moral and epistemological. He discusses truth in the light of John Locke's definition; evidence for certitude, following the Cartesian *cogito* and the principle of contradiction. He proposes a methodology for discovering

the truth that is also Cartesian. Most interesting is his recognition of the nihilistic challenge to moral values that was becoming more vigorous at the time. The longest section of the book expounds the argument that moral nihilism is justified and that all moral values disappear if God does not exist. Saint-Hyacinthe's real purpose was to urge men to believe in God, but the effect of his argument was more likely to lead them to immoralism, for he expounds that doctrine forcefully and endeavors to make it an impregnable position except in the face of God's existence. These little-known pages are notable as the most systematic exposition of moral nihilism before the Marquis de Sade. The *Lettres* had some success, and were translated under the title *Letters Giving an Account of Several Conversations Upon Important and Entertaining Subjects* (2 vols., London, 1731).

Among Saint-Hyacinthe's other publications are the *Lettres à Mme. Dacier* (1715, concerning the *querelle d'Homère*); "Lettre à un ami, touchant le progrès du déisme en Angleterre" (in his edition of *Mémoires concernant la théologie et la morale,* 1732); and the novel *Histoire du prince Titi* (1735).

L. G. Crocker (1967)

SAINT LOUIS SCHOOL, THE

See *Harris, William Torrey*

SAINT-SIMON, CLAUDE-HENRI DE ROUVROY, COMTE DE

(1760–1825)

The French social philosopher, Claude-Henri de Rouvroy, Comte de Saint-Simon, the founder of French socialism, was the eldest son of an impoverished nobleman. He was educated privately by tutors, among them the encyclopedist Jean Le Rond d'Alembert. Beginning a military career at the age of seventeen, he took part in the American Revolution and was wounded at the naval battle of Saintes in 1782. Despite subsequent disclaimers, Saint-Simon actively supported some of the measures introduced by the French Revolution of 1789. He renounced his title; he also drew up the *cahier* of his locality for the Estates General and presided at the meeting at which his commune elected a mayor. Although his revolutionary zeal earned him two certificates of *civisme,* his activities were not wholly disinterested. He took advantage of the sale at low prices of church and *émigré* property by making considerable purchases. He was arrested in 1793, but since it transpired that a mistake had been made, he was released the following year. He was active in political life under the Directory, among other things participating in the peace negotiations with the English at Lille.

Saint-Simon finally retired from governmental and financial activity and embarked on the career of writer and prophet that continued until the end of his life. He first studied physics for three years, at the same time forming friendships with a number of leading scientists and writers whom he helped to support. Later he traveled extensively, especially in Germany, England, and Switzerland. It was not until 1814, however, when he found an able and enthusiastic collaborator and disciple in the future historian Augustin Thierry, that his writings began to reach a wide public, particularly among the managers and businessmen who had risen to positions of influence during the Napoleonic era. The list of subscribers for his publication *L'industrie,* the first number of which appeared in 1816, included various prominent industrialists and bankers. The next year Saint-Simon's partnership with Thierry ended, and he began an association with Auguste Comte—an event of considerable significance, for it was in Comte's later work that some of Saint-Simon's fundamental conceptions were given more systematic and trenchant expression than their originator had been able to achieve. The collaboration between these two forceful personalities lasted for seven years but was finally broken by a quarrel in 1824, the year before Saint-Simon's death.

IDEALS AND REALITY

"The philosophy of the eighteenth century has been critical and revolutionary: that of the nineteenth century will be inventive and destructive" (*Oeuvres complètes,* Vol. XV, p. 92). This remark accurately reflects the position that Saint-Simon envisaged himself as occupying in the history of political and social ideas. He in no way wished to underestimate the achievements of his Enlightenment predecessors the *philosophes,* who by their bold attacks upon the traditional frameworks of thought and their criticisms of existing institutions had prepared the way for the vast upheaval of the French Revolution. Saint-Simon saw in the writings of such men as Étienne Bonnot de Condillac and the Marquis de Condorcet anticipations of his own belief that human affairs should be approached in a scientific, Newtonian spirit of inquiry,

and he sympathized with their contention that religious dogmas had over the centuries become the means by which the mass of the people had been held in ignorant and superstitious servitude to their rulers. He also shared the humanitarian and internationalist ideals that had inspired the work of his predecessors. (His subscription to these ideals, apparent in all his main publications, was perhaps most distinctively expressed in *Nouveau Christianisme* [Paris, 1825], an essay that appeared at the very end of his life.)

On the other hand, Saint-Simon's work also pointed forward to the quite new ways of conceptualizing and interpreting social relations that were later to gain wide currency through the writings of Karl Marx. In particular, Saint-Simon exhibited a far firmer grasp of the conditions that determine and mold historical change than had earlier thinkers, and this profoundly affected the form taken by his own practical recommendations. Sincerely held utopian ideals, even when carefully worked out in detailed political programs, were by themselves quite useless, he held, if they did not take account of these conditions. Utopian changes, if put into effect, were likely to result in a vacuum that would eventually be filled by forces as undesirable as those which had been expelled. The destruction of outdated institutions was one thing; their replacement by others of lasting validity, adapted to the technological, economic, and social requirements of the time, was another. This was surely the lesson of the French Revolution. Had not the high hopes and aspirations that marked its beginning ultimately foundered in atrocities, suffering, and tyranny?

HISTORICAL CHANGE

Despite the importance he assigned to it, Saint-Simon never set out his conception of historical change and development in a precise or systematic form. Like his other contributions to social theory, it was put forward in a somewhat disjointed and piecemeal fashion. Nevertheless, an outline of his view can be extracted from various works, notably from his writings in the periodical *L'organisateur* (Paris, 1819–1820). Saint-Simon spoke as if he had discovered a necessary law of evolution valid for all societies at all times, but the kernel of what he had to say was actually based upon a single instance, the transformation that had overtaken European society since the feudal period. The chief originality and importance of his analysis of how this change came about lay in his recognition of the role played by the emergence and conflict of classes and of the way in which such conflict issues in new forms of political organization and of ideology adapted to the interests of the socially and economically dominant class. The institutions and beliefs of the Middle Ages fulfilled a perfectly intelligible, and indeed necessary, function from the point of view of the stage of development society had at that time reached (it is notable that Saint-Simon's approach to medieval history was considerably more sympathetic than that of either his Enlightenment predecessors or his liberal contemporaries).

Only later, with the enfranchisement of the communes, the emergence of a class of independent producers, and the subsequent growth of an industrial system of production under the impact of scientific and technological advances, did feudal organization become evidently anachronistic. Then the very features of the framework that had provided medieval society with the protection and unity of purpose it required impeded the free development of the new forces germinating within it. Thus, the seventeenth and eighteenth centuries witnessed the culmination of two major developments. On the one hand, there were increasingly effective attacks by the commons against privileges and institutions that had outgrown their social utility; on the other, the doctrines of the church, which during the Middle Ages had performed valuable services but which had been rendered obsolete by scientific discoveries, were subjected to a series of unanswerable criticisms. The net result was "the ruin of the old system in its parts and as a whole" (*Oeuvres complètes,* Vol. XX, p. 104).

ECONOMIC AND POLITICAL PROGRAM

The lessons Saint-Simon drew from previous developments for his own time were far-reaching. Although the old order was in a general condition of dissolution, it had still not been wholly superseded. Many of the chief centers of power and influence remained in the hands of "more or less incapable bureaucrats" (ibid., pp. 17–26), idlers, and ignoramuses who owed their positions to the accident of birth or inherited wealth and who were in effect no better than destructive parasites. To a considerable extent "men still allow themselves to be governed by violence and ruse" (ibid.). In order to remedy this state of affairs, Saint-Simon appealed directly to the leaders of the new class of *industriels,* claiming that the hour had arrived for them to take into their own hands the management of society and thereby complete the revolution that had been maturing for so long. Only if this were done could society be reorganized in a way that would ensure its direction by efficient administrators, men who would see that those who could make a genuinely pro-

ductive contribution to its advance and prosperity were no longer ignored or exploited and received, instead, their appropriate reward.

Yet despite his insistence on the need for social justice, Saint-Simon had little faith in political democracy. He envisaged a hierarchical system, characterized by equality of opportunity rather than equality of wealth and run on explicitly elitist lines. The central administration of the community would consist of three chambers—the chamber of invention, the chamber of examination, and the chamber of deputies. Of these the first was to consist of artists and engineers who would propose plans, the second of scientists who would critically assess the proposals and also control education, and the third of captains of industry whose function would be executive and who (Saint-Simon somewhat optimistically assumed) would give just consideration to the interests of all members of the industrial class, workers and managers alike. Saint-Simon appears to have thought that in the type of society he had in mind, which would be rationally planned in a manner advantageous to all, there would be little or no need for the use of force to compel obedience to law and that government in the traditional sense would no longer be required. There is a clear anticipation of the Marxian conception of the withering away of the state.

ETHICS AND RELIGION

Saint-Simon was always conscious of the importance of moral and social ideals in helping to promote harmony and a sense of purpose in human communities. In medieval times the Christian religion had performed this role, and he thought that there was a place for a comparable system of beliefs, adapted to contemporary knowledge and interests, in any viable modern society. For the creation of such a system he initially looked to philosophy, but in his later years he recommended a return to the fundamental tenets of Christian teaching. The ethical doctrines of Christianity, he held, retained their validity even if the theological and metaphysical dogmas associated with them are no longer acceptable.

INFLUENCE

It is impossible in a short space to do justice to the fertility and originality of Saint-Simon's thinking on what he called social physiology. An untidy, impatient, and inelegant expositor of his own ideas, he nonetheless understood the central issues of his time better than many of his contemporaries and exhibited a keener insight into the economic and technical realities that lie beneath the surface of political arrangements and change. Marx indisputably owed a significant debt to him, but Marx was only one among a host of nineteenth-century thinkers who profited in one way or another from Saint-Simon's perceptive and imaginative mind.

See also Alembert, Jean Le Rond d'; Comte, Auguste; Condillac, Étienne Bonnot de; Condorcet, Marquis de; Enlightenment; Marx, Karl; Social and Political Philosophy.

Bibliography

WORKS BY SAINT-SIMON

Oeuvres complétes de Saint-Simon et Enfantin, 47 vols. Paris, 1865–1876.

Lettres d'un habitant de Genève à ses contemporains. Edited by A. Pereire. Paris, 1925. Written in 1803.

Textes choisis. Edited by J. Dautry. Paris, 1951.

Selected Writings. Translated, with an introduction, by F. M. H. Markham. Oxford: Blackwell, 1952.

Henri Saint-Simon (1760–1825): Selected Writings on Science, Industry and Social Organisation. Translated and edited by Keith Taylor. New York: Holmes and Meier, 1975.

The Political Thought of Saint-Simon. Edited by Ghita Ionescu. Oxford: Oxford University Press, 1976.

WORKS ON SAINT-SIMON

Berlin, Isaiah. "Saint-Simon." In *Freedom and Its Betrayal: Six Enemies of Human Liberty,* edited by Henry Hardy. Princeton, NJ: Princeton University Press, 2002.

Carlisle, R. B. *The Proffered Crown. Saint-Simonianism and the Doctrine of Hope.* Baltimore: Johns Hopkins University Press, 1987.

Dondo, M. M. *The French Faust, Henri de Saint-Simon.* New York: Philosophical Library, 1955.

Durkheim, Émile. *Le socialisme.* Edited by Marcel Mauss. Paris: Alcan, 1928. Translated by Charlotte Sattler as *Socialism and Saint-Simon.* Yellow Springs, OH: Antioch Press, 1958.

Ferrarotti, Franco. "Claude-Henri de Saint-Simon." In *An Invitation to Classical Sociology: Meditations on Some Great Social Thinkers.* Lanham, MD: Lexington, 2003.

Iggers, Georg G. *The Cult of Authority; The Political Philosophy of the Saint-Simonians,* 2nd ed. The Hague: Nijhoff, 1970.

Manuel, F. E. *The New World of Henri de Saint-Simon.* Cambridge, MA: Harvard University Press, 1956.

Plamenatz, J. P. *Man and Society.* London: Longmans, 1963. Vol. II, Ch. 2.

Patrick Gardiner (1967)
Bibliography updated by Philip Reed (2005)

SAINT VICTOR, SCHOOL OF

The Augustinian house of canons at St. Victor in Paris was founded in 1108 by William of Champeaux, the cele-

brated logician and theologian who retired there from the schools of Paris after undergoing a religious conversion and after Peter Abelard's attacks on his realism. The abbey survived until the French Revolution, but in the twelfth and early thirteenth centuries it was especially famous for its public school and for the distinction of the masters and canons who resided and taught there. From William, St. Victor derived high religious ideals, a leaning toward the conservative theological tradition of the school of Anselm of Laon, and an active interest in the work of other Parisian schools. Its masters mediated between the theological orthodoxy and strictness of the Cistercians—Bernard of Clairvaux was a friend to St. Victor—and the intellectual adventurousness of such secular masters as Abelard. St. Victor in the twelfth century combined Scholasticism and mysticism and exerted a most powerful influence upon the development of both philosophical and theological thought in that century. Not only did it possess among its canons some of the ablest writers of the age but it also attracted as long-staying guests such celebrated teachers as Peter Lombard and Robert of Melun. Besides producing a wealth of literature, its leaders also contributed to the fall of Abelard, to the damping of the enthusiasms of the Chartrains, to the containment of Gilbert of Poitiers, and to the correction of the Christological errors that abounded in the mid-twelfth century.

HUGH OF ST. VICTOR

St. Victor, unlike Chartres, was not devoted to the liberal arts. No commentary upon a nontheological text is known to have been written there, and purely literary writings were even relegated by the greatest Victorine, Hugh (d. 1141), to the position of mere appendices to the liberal arts. The Victorines did not encourage profane studies for their own sakes. The extreme, fanatical Walter (d. circa 1180) intemperately denounced the Aristotelian spirit of Abelard, Gilbert of Poitiers, William of Conches, and Peter Lombard. Absalon (d. 1203), too, warned against the dangers found in Aristotle.

Hugh vigorously challenged his humanist contemporaries who in the first half of the twelfth century thought more often of pagan philosophy than of Christ and his saints. Against the Chartrains he insisted upon the disparity between the cosmogony of Plato's *Timaeus* and Christian truth. Nonetheless, in his *Didascalicon* (which contains a program of Christian education) Hugh shows that he was thoroughly immersed in secular studies as the preliminary to divine science, for he considered the arts an indispensable aid to the understanding of Scripture. Hugh sought to pass through knowledge to wisdom and to promote that participation in the divine Wisdom for which man was made. Similarly, Godfrey (d. after 1194) also affirmed that the liberal arts and theology were inseparable and that together they offered a complete education.

Philosophical elements are found scattered in the writings of the Victorines. Inheriting the Boethian-Aristotelian theory of abstraction, Hugh appreciated the necessity for logic without exalting it as highly as did Abelard. In physics Hugh maintained the atomic theory of matter and accepted the principle of the conservation of matter. His psychology was Augustinian, and he found the proof for the existence of the immaterial soul in the fact of its self-consciousness.

RICHARD OF ST. VICTOR

Both Hugh and his disciple Richard (d. 1173) describe the ascent of the soul in contemplation; Richard especially is the theorist of the degrees of love. But whereas Hugh insisted upon the inadequacy of reason and the necessity for faith, Richard, who rivaled Hugh as a spiritual writer, was more scholastic and laid a stronger emphasis upon dialectic to supplement the traditional scriptural and patristic authorities. Inheriting from Anselm of Canterbury his zeal to search for the "necessary reasons" of faith and for an understanding of belief, he accounted for the trinity of persons in God in abstract style with a very original dialectic of mutual love; he was also the first medieval thinker to provide, in one of the great speculative achievements of the period, an empirical basis in the principle of causality for a proof of God's existence.

VICTORINE THEOLOGY

Essentially the Victorines provided a theology for contemplatives within the cloister rather than for the schools. Hugh was a systematizer of theology on Augustinian lines, using dialectic when needed. Richard became the mystical doctor of the later Middle Ages. Both Hugh and Richard were biblical exegetes and spiritual writers, and it is for this that they and such other Victorines as Andrew (d. 1175) and Gamier (d. 1170) and the poet Adam were best known in the Middle Ages. Godfrey was more pronounced in his humanism, combining Chartrain Platonism and Aristotelian dialectic with Victorine spirituality. Achard (abbot 1155–1160) also mingled Augustinian theology with Chartrain Platonism, but all the Victorines concurred in wishing to turn knowledge into wisdom and the reader of the profane sciences into a contemplative. They always returned to the internal and external experi-

ences of the soul, and frequently to the use of allegory and symbolism in the penetration of divine truths. In the early thirteenth century the influence of Pseudo-Dionysius, which had been powerful upon Hugh, prevailed again upon Thomas Gallus, who was a forerunner of the mysticism of the later Middle Ages.

See also Abelard, Peter; Aristotle; Augustinianism; Bernard of Clairvaux, St.; Gilbert of Poitiers; Peter Lombard; Plato; Platonism and the Platonic Tradition; Pseudo-Dionysius; William of Champeaux; William of Conches.

Bibliography

Baron, R. *Science et sagesse chez Hugues de Saint-Victor.* Paris, 1957.

Bonnard, Fourier. *Histoire de l'abbaye royale de Saint-Victor.* 2 vols. Paris, 1904.

Delhaye, P. *Le microcosmus de Godefroy de Saint-Victor.* 2 vols. Lille and Gembloux, 1951. Text and study.

Dumeige, G. *Richard de Saint-Victor et l'idée chrétienne de l'amour.* Paris: Presses Universitaires de France, 1952.

Ebner, J. *Die Erkenntnislehre Richards von St. Viktor.* Beiträge zur Geschichte der Philosophie des Mittelalters, Vol. XIX, no. 4. Münster in Westfalen: Aschendorff, 1917.

Ethier, A. M. *Le de trinitate de Richard de Saint-Victor.* Paris: J. Vrin, 1939.

Fritz, G. "Richard de Saint-Victor." In *Dictionnaire de théologie catholique,* Vol. XIII, Cols. 2676–2695. Paris, 1937.

Geyer, B. *Die patristische und scholastische Philosophie,* 261–272. Basel, 1927.

Godfrey of St. Victor. *Fons Philosophiae,* edited by P. Michaud-Quantin. Namur, Belgium, 1956.

Hugh of St. Victor. *Didascalicon.* Translated by J. Taylor. New York: Columbia University Press, 1961.

Kleinz, J. P. *The Theory of Knowledge of Hugh of St. Victor.* Washington, DC: Catholic University of America Press, 1944.

Ostler, H. *Die Psychologie des Hugo von St. Victor.* Beiträge zur Geschichte der Philosophie des Mittelalters, Vol. VI, no. 1. Münster in Westfalen: Aschendorff, 1906.

Richard of St. Victor. *De Trinitate,* edited by J. Ribaillier. Paris: J. Vrin, 1958.

Vernet, F. "Hugues de Saint-Victor." In *Dictionnaire de théologie catholique,* Vol. VII, Cols. 240–308. Paris, 1922.

David Luscombe (1967)

SALMON, WESLEY
(1925–2001)

The American philosopher of science Wesley Charles Salmon was born August 9 in Detroit, Michigan, and died April 22 near Madison, Ohio. He pursued undergraduate studies at Wayne University and the University of Chicago Divinity School, received an MA in philosophy from the University of Chicago in 1947, and a PhD in philosophy from the University of California at Los Angeles in 1950. His principal academic appointments were at Brown University, Indiana University, the University of Arizona, and the University of Pittsburgh; he retired from this last institution in 1999. At UCLA his dissertation advisor was the philosopher of science Hans Reichenbach and much of Salmon's subsequent work was influenced by Reichenbach's philosophy. A lifelong defender of empiricism, Salmon made significant contributions to a wide range of topics, primarily in explanation, causation, inductive inference, and the philosophy of probability.

WORK

Beginning in 1971, Salmon developed a widely discussed alternative to Carl Hempel's covering law model of scientific explanation. The key element of Salmon's statistical relevance model was its insistence that explanatory factors must be statistically relevant to the occurrence of the event to be explained. This undermined in two ways Hempel's view that an explanation must lead people to expect the explanandum to occur. It showed that this condition was not necessary because events with low probability, such as the occurrence of lung cancer, can be explained in terms of statistically relevant factors such as cigarette smoking. It also showed that Hempel's model did not provide sufficient conditions for an explanation because irrelevant factors such as a man's taking birth control pills, when included in a Hempelian explanation, undermine the effectiveness of an explanation of his not getting pregnant.

In the course of developing the statistical relevance model, Salmon began to stress the importance of the causal relevance, rather than the statistical relevance, of explanatory factors; his 1984 book *Scientific Explanation and the Causal Structure of the World* contains an account of probabilistic causality grounded in an "at-at" theory of causation within which spatiotemporally continuous markable processes connecting cause and effect play a central role. The aim was to provide an account of causation free of appeals to counterfactuals and thus acceptable to an empiricist, yet different from Hume's in stressing the importance of connecting processes. In the light of criticisms that the markability criterion required tacit appeals to counterfactuals, Salmon abandoned it in the early 1990s and adopted a position where the transmission of conserved quantities was what distinguished causal from non-causal processes.

This appeal to causal processes and the distinction between genuine processes and pseudo-processes meshed with Salmon's interests in space-time theories. In that area he defended a causal theory of space and time within which the direction of time was to be grounded in causal asymmetries. He also maintained a long-term interest in conceptions of synchrony and in defending a conventionalist approach to simultaneity relations.

In his 1984 book, Salmon argued for a form of scientific realism based on the principle of the common cause. This principle states that if an association is observed between two types of event then, in the absence of a direct causal connection between instances of the events, there exists a common cause responsible for generating the association. This principle is general and can be used to argue for the existence of unobserved entities. It lies behind the reasoning used by Bertrand Russell in inferring the continued existence of a cat from its occasional observed appearances and was employed, Salmon claimed, by Jean Perrin in using the similarity of values obtained from different experimental techniques to determine Avogadro's number to argue for the reality of atoms. Despite its appeal, the principle does have its limitations. It is inapplicable in certain quantum mechanical situations where there are no hidden variables. It is also easy to find cases where two properties each increase over time but there is no common cause underlying the two. Nevertheless, Salmon's emphasis on this principle has led to an important new way of thinking about scientific realism.

Much of Salmon's early work concerned issues in probability and induction. For many years he defended Reichenbach's pragmatic vindication of induction, which argues that inductive inferences, more specifically the "straight rule" that projects the existing relative frequency of an event's occurrence into the future, is at least as likely to be successful as any other rule. In the light of criticisms by Ian Hacking, Salmon tempered his advocacy of this approach while continuing to insist on the importance of linguistic invariance for inductive rules.

Throughout his career Salmon defended a relative frequency interpretation of probability, including its use in his accounts of causation. For Salmon, the correct frequency to attribute to an event was the frequency within the broadest homogenous reference class to which the event belongs—that is, that class of events for which no further statistically relevant factors exist. Although he occasionally displayed sympathy for a propensity approach to probabilities and appreciated the role played by logical probabilities in Carnap's inductive logic, he developed important criticisms of both. In many of his writings, Salmon argued that an objective form of Bayesian inference could illuminate a number of issues in the philosophy of science. Most notable, Salmon's insistence on preserving the context of discovery/context of justification distinction and on using Bayesian methods in the latter area led him to argue that Kuhn's account of theory choice could be made more objective by employing Bayesian techniques of theory justification.

In addition to his philosophical contributions, Salmon was an outstanding teacher and was much admired for his personal qualities. His introductory logic book was widely used as an undergraduate text, went into three editions, and was translated into five foreign languages. The exemplary clarity of his writing is evident in all of his publications.

Bibliography

PRIMARY WORKS

The Foundations of Scientific Inference. Pittsburgh: University of Pittsburgh Press, 1967.

Space, Time, and Motion. 2nd ed. Minneapolis: University of Minnesota Press, 1981.

Scientific Explanation and the Causal Structure of the World. Princeton, NJ: Princeton University Press, 1984.

Four Decades of Scientific Explanation. Minneapolis: University of Minnesota Press, 1990.

Causality and Explanation. New York: Oxford University Press, 1998.

Reality and Rationality, edited by M. Salmon and P. Dowe. New York: Oxford University Press, 2005.

SECONDARY WORKS

Fetzer, J., ed. *Probability and Causality: Essays in Honor of Wesley C. Salmon*. Dordrecht, Netherlands: D. Reidel, 1988. Contains a comprehensive annotated list of his publications through 1988.

McLaughlin, R., ed. *What? Where? When? Why?: Essays on Induction, Space and Time, Explanation*. Dordrecht, Netherlands: D. Reidel, 1982. Contains a brief autobiographical essay.

Paul Humphreys (2005)

SANCHES, FRANCISCO
(c. 1551–1623)

Francisco Sanches, a philosopher and physician, was born on the Spanish-Portuguese border, either in Tuy or Braga, of Marrano or New Christian parents. His family had moved to Portugal and then to southern France to escape religious and political persecution. The young Sanches studied at the Collège de Guyenne in Bordeaux, the same

school that his distant cousin, Michel Eyquem De Montaigne, attended. Sanches studied in Rome and then went to the University of Montpellier, where he received a degree in medicine in 1574. He was appointed professor of philosophy in 1585 and professor of medicine in 1612 at the University of Toulouse, where he had a successful career until his death in 1623.

One of Sanches's first philosophical writings that has survived is a letter to the Jesuit mathematician, Father Christopher Clavius, who had just edited Euclid's works and whom Sanches had met in Rome. Sanches offered a skeptical attack on the possibility of attaining genuine truth in mathematics. This was followed by his most famous writing, *Quod nihil scitur* (*That Nothing Is Known*). He soon thereafter wrote a critical examination of the astrological interpretations of the comet of 1577, *Carmen de Cometa*, published in 1578, and some commentaries on portions of Aristotle's writings, as well as many medical works. Sanches criticized various Renaissance naturalistic views, such as those of Girolamo Cardano, and may have actually debated Giordano Bruno in person in Toulouse.

In the letter to Christopher Clavius, Sanches attacked a form of the Platonic theory of knowledge. We cannot gain knowledge of things through mathematical study, because the objects studied by mathematics are not the natural, real ones encountered in human life. Rather, these objects are ideal, or maybe even impossible ones, such as points and lines. The mathematical relations that are demonstrated about such objects do not help explain anything in nature or experience, unless we happen to know independently that the experienced objects have mathematical properties, and also know that the principles of mathematics are in fact true. As far as we can tell, mathematics is just conjectural or hypothetical until we can independently determine the nature of things.

Sanches's *Quod nihil scitur* was written in 1576 and published in 1581. In it he develops his skepticism by means of a critique of Aristotelianism. He begins by asserting that he does not even know if he knows nothing. Then he proceeds to analyze the Aristotelian conception of knowledge to show why this is the case.

Every science begins with definitions, but definitions are nothing but names arbitrarily imposed upon things in a capricious manner, having no relation to the things named. The names keep changing, so that when we think we are saying something about the nature of things by means of combining words and definitions, we are just fooling ourselves. On the one hand, if the names assigned to an object such as man, such as "rational animal," all mean the same thing, then they are superfluous and do not help to explain what the object is. On the other hand, if the names mean something different from the object, then they are not the names of the object. By means of such an analysis, Sanches worked out a thoroughgoing nominalism.

Sanches went on to examine the Aristotelian notion of science. Aristotle defines science as "disposition acquired through demonstration." But what does this mean? This is explaining the obscure by the more obscure. The particulars that one tries to explain by this science are clearer than the abstract ideas that are supposed to clarify them. The particular, Socrates, is better understood than something called "rational." Instead of dealing with the real particulars, these scientists argue about a vast number of abstract notions and fictions. "Do you call this science?" Sanches asked, and then replied, "I call it ignorance."

The method of Aristotelian science, demonstration, is next attacked. A demonstration is supposed to be a syllogism that produces science, but this involves a vicious circle rather than engendering any new information. To demonstrate that Socrates is mortal, one argues from "all men are mortal" and "Socrates is a man." The premises, however, are built up from the conclusion: the particular, Socrates, is needed to have a concept of man and mortality. The conclusion is clearer than the proof. Also, the syllogistic method is such that anything can be proven by starting with the right premises. It is a useless, artificial means, having nothing to do with the acquisition of knowledge.

Sanches concludes that science cannot be certitude acquired by definitions, neither can it be the study of causes, for if true knowledge is to know a thing in terms of its causes, one would never get to know anything. The search for its causes would go on ad infinitum as one studied the cause of the cause, and so on.

For Sanches, true science is the perfect knowledge of a thing—"SCIENTIA EST REI PERFECTA COGNITIO." Genuine knowledge is immediate, intuitive apprehension of all the real qualities of an object. Thus, science will deal with particulars, each somehow to be individually understood. Generalizations go beyond this level of scientific certainty, and introduce abstractions, chimeras, and so on. Sanches's scientific knowledge consists, in its perfect form, of experiential apprehension of each particular in and by itself.

Sanches showed that, strictly speaking, human beings were incapable of attaining certainty. The science

of objects known one by one cannot be achieved, partly because of the nature of objects and partly because of the nature of humankind. Things are all related to one another and cannot be known individually. There are an unlimited number of things, all different, so they could never all be known. And still worse, things change so that they are never in such a final or complete state that they can be truly known.

On the human side, Sanches devoted a great deal of time to presenting difficulties that prevent people from obtaining true knowledge. Our ideas depend on our senses, which only perceive the surface aspects of things, the accidents, and never the substances. From his medical information, Sanches was also able to point out how unreliable our sense experience is, how it changes as our state of health alters. The many imperfections and limitations, with which God has seen fit to leave us, prevent our senses and our other powers and faculties from ever attaining any true knowledge. The conclusion of all this is that the only truly meaningful scientific knowledge cannot be known. All that humans can achieve is limited, imperfect knowledge of some things that are present in their experience through observation and judgment.

Sanches's claim that *nihil scitur* is argued for on philosophical grounds, on a rejection of Aristotelianism and an epistemological analysis of what the object of knowledge and the knower are like. His totally negative conclusion is not the position of Pyrrhonian skepticism, the suspense of judgment as to whether anything can be known, but rather the negative dogmatism of the Academics. A theory of the nature of true knowledge is asserted, and then it is shown that such knowledge cannot be attained. The Pyrrhonists, with their more thoroughgoing skepticism, could neither assent to the positive theory of knowledge, nor to the definite conclusion that *nihil scitur*.

Sanches put forward a procedure, not to gain knowledge, but to deal constructively with human experience. This procedure, for which Sanches introduced the term, for the first time, of *scientific method*, "Método universal de las ciencias," consists in careful empirical research and cautious evaluation of observable data. In advancing this limited or constructive view of science, Sanches was the first Renaissance skeptic to conceive of science in its modern form, as the fruitful activity about the study of nature that remained after one had given up the search for absolutely certain knowledge of the nature of things.

Sanches was influential in his own day and throughout the seventeenth century. *Quod nihil scitur* was reissued several times up to 1665. Late in the seventeenth century two refutations appeared in Germany. People have seen possible influences not only on Descartes, but also on Pierre Gassendi, Marin Mersenne, Spinoza, and Leibniz, among others, although it is hard to delineate his exact influence as different from that of Montaigne, Sextus Empiricus, Cicero, Charron and other available skeptical sources who were read by most intellectuals of the time. It may be that Sanches's formulation of the skeptical problem is closer to the modern idiom than that of any of his contemporaries, including Montaigne, and in terms of how philosophy developed, reads more like a precursor of Bacon or Descartes.

See also Scientific Method.

Bibliography

WORKS BY SANCHES

Que nada se sabe. Buenos Aires: Editorial Nova, 1944. Spanish translation of *Quod Nihil Scitur*, with an introduction by Marcelino Menédez y Pelayo.

Opera Philosophica, edited by Joaquim de Carvalho. Coimbra, Portugal: 1955. Introduction in Portuguese by Carvalho.

Tratados filosoficos. 2 vols. Lisbon, Portugal: Univ. de Coimbra, 1955. Latin-Portuguese edition.

WORKS ABOUT SANCHES

Coralnik, A. "Zur Geschichte der Skepsis. I. Franciscus Sanchez." *Archiv für Geschichte der Philosophie* 27, Neue Folge vol. 20 (1914).

Franck, Adolphe, ed. *Dictionnaire des sciences philosophiques*. Paris: Hachette, 1875. See "Sanchez, François."

Iriate, J. *Kartesischer oder Sanchezischer Zweifel*. Bottrop: W. Postberg 1935.

Limbrick, Elaine, and Douglas F. S. Thomson, eds. *That Nothing Is Known*. Cambridge, U.K.: Cambridge University Press, 1988.

Menédez y Pelayo, Marcelino. "De los orígenes del criticismo y del escepticismo y especialmente de los precursors españoles de Kant." In his *Ensayos de critica filósofica*. Madrid: V. Suárez, 1918.

Moreau, Joseph. "Doubte et savoir chez Francisco Sanches." In *Aufsätze zur portugiesischen Kulturgeschichte*. 1st ser., vol. 1: *Portugiesische Forschungen der Görresgesellschaft*. Munster in Westfalen, Germany: 1960.

Moreira de Sá, Artur. *Francisco Sanches filósofo e matematico*. 2 vols. Lisbon, Portugal: Oficinas graficas casas portuguesas 1947.

Owen, John. *The Skeptics of the French Renaissance*. London: S. Sonnenschein, 1893.

Popkin, Richard H. *The History of Skepticism from Savonarola to Bayle*. New York: Oxford University Press, 2003.

Senchet, Émilien. *Essai sur la méthode de Francisco Sanchez*. Paris: V. Giard and E. Brière,1904.

Richard Popkin (1967, 2005)

SANCTIS, FRANCESCO DE

See *De Sanctis, Francesco*

SANTAYANA, GEORGE

(1863–1952)

George Santayana, the philosopher and man of letters, was born in Madrid. His parents separated within a few years of his birth, and his mother went to live in Boston, Massachusetts, with the children of a previous marriage. Santayana grew up in Ávila under his father's care, but at the age of eight he joined his mother in Boston. He was educated at the Boston Latin School and at Harvard College. After graduating from Harvard in 1886, he studied in Germany for two years and then returned to take his doctorate at Harvard, for which he wrote a thesis on Rudolf Lotze. He subsequently joined the department of philosophy and remained a member of the Harvard faculty until 1912, when a small inheritance permitted him to retire. He lived in England for a number of years and then in Paris, but in 1925 he finally settled in Rome. During World War II, he took refuge in the convent of an order of English nuns in Rome, and he continued to live there until his death.

CULTURAL BACKGROUND

Both Santayana's personal life and his philosophical development were decisively influenced by his peculiar position as a Spanish Catholic living and teaching in a predominantly Protestant society with a philosophical and cultural tradition that he felt to be in many respects deeply alien to his own personality. He was always proud—rather defiantly so—of his Catholicism and his Latinity, despite the fact that he was not a believer and was not notably attached to Spain or to Spanish culture. These loyalties expressed instead a deeply rooted hostility to the commercial and democratic ethos of modern industrial society and an equally deep aspiration toward a radically different style of life and thought that, for Santayana, was best exemplified in the classical Mediterranean world. Philosophically, he felt his truest affinities to be with the Greeks and perhaps the Hindus, and among the moderns, with Benedict de Spinoza, rather than with the empiricism and idealism of German and Anglo American philosophy. In fact, however, his points of affiliation with the European and American philosophy of the modern period are both numerous and obvious, and it would appear that his debt to the post-Cartesian tradition in modern philosophy is much greater than he was inclined to think. What chiefly set his work apart from the mainstream of twentieth-century philosophy was his highly personal and literary mode of writing and his rather disdainful lack of interest in the methodological questions that were of central importance to the development of phenomenology on the Continent and analytic philosophy in the English-speaking world. When one considers the substantive doctrines to which he was committed, however, and, in particular, the ontological distinctions on which his "Realms of Being" rest, his philosophy emerges as a highly idiosyncratic doctrine of transcendental subjectivity that would scarcely be conceivable apart from the very tradition of modern philosophy which he so violently criticized.

PHILOSOPHICAL DEVELOPMENT

Santayana's philosophical career falls naturally into two main periods. The first of these is the period in which he published *The Sense of Beauty* (1896) and *The Life of Reason* (1905–1906); its chief distinguishing feature is Santayana's disposition at that time to conceive of philosophy as a kind of descriptive psychology of the higher mental functions. He assumed the broad truth of the doctrine of biological evolution and its relevance to the understanding of mental phenomena, and while he held all knowledge to be representational in nature, he did not question "our knowledge of the external world," nor did he feel the need for any initial withdrawal of belief in such a world in the Cartesian manner. "Mind" is placed firmly in its biological context, and such independence as it enjoys is due not to any special ontological status, but rather to its capacity for giving an ideal and aesthetic meaning to its natural setting and functions.

In the second period, during which he wrote *Scepticism and Animal Faith* (1923) and *Realms of Being* (1927–1940), Santayana came to feel the need for a greater systematic rigor in the exposition of his views and for a purified and nonpsychological mode of stating the fundamental distinctions on which his philosophy rested. In particular, he felt that in *The Life of Reason* he had not made clear enough that the "nature" described there as having been "drawn like a sponge, heavy and dripping from the waters of sentience" was the idea of nature, not nature itself. He now tried to correct this error by means of a set of ontological—that is, nonpsychological—distinctions between the different kinds of being that are the objects of different kinds of mental activity. Thus, imagination, for example, must be defined by reference to the essences or abstract characters that Santayana now recognized as having a distinct ontological status, rather than

the other way around. In carrying out this revision of his earlier views, Santayana was in some measure aligning himself with similar antipsychologistic tendencies at work in the logical realism of Bertrand Russell, as well as in the phenomenology of Edmund Husserl, which he regarded as having a certain affinity to his own views.

Some commentators have felt that this shift from what they describe as Santayana's earlier naturalism to his later "Platonism" amounted to a fundamental change in his general philosophical perspective. Santayana's own statements, however, make it clear that the system presented in *Realms of Being* is to be understood as the ultimate philosophical basis of the naturalistic *Weltanschauung* sketched out in *The Life of Reason,* in which he had paid relatively little attention to technical philosophical issues. It must be admitted that the moral atmosphere of the two works differs, and that in the later one Santayana seems even more the detached spectator of the noncontemplative phases of the "life of reason" than he had before. But this is as much a personal as a philosophical matter, and there is no good reason for denying the fundamental unity of Santayana's thought during the two main periods of its development.

AESTHETICS

Santayana's first important philosophical work was *The Sense of Beauty* (1896). In it he attempted to state a complete aesthetic theory, which he later developed further in *Reason in Art* (1905), Volume IV of *The Life of Reason.* In the earlier book, aesthetic theory is characterized as a psychological inquiry whose data are aesthetic judgments considered as "phenomena of mind and products of mental evolution"; the inquiry is to be distinguished both from the actual exercise of critical judgment and from the historical investigation of the evolution of the various art forms. Santayana argued that this inquiry must be carried out independently of metaphysical issues and the "interests of the moral consciousness," and that it must make clear the bases of aesthetic experience in human nature as conceived by natural science and in particular evolutionary biology. To this end, Santayana sketched out a theory of value according to which all preference is an essentially irrational expression of vital interest and the standard of value is the enjoyment or pleasure procurable through different courses of action. Morality is concerned with negative values, namely, the avoidance of pain and suffering, while aesthetic value is concerned with positive enjoyment and stands in the same relation to morality as play does to work.

The pleasure that is distinctively aesthetic, however, must be further qualified as intrinsic (or immediate) and as "objectified," in the sense of being experienced as a quality of a thing and not as an affection of the organ which apprehends it. Santayana denied that it must have the disinterested character attributed to it by Immanuel Kant and that it must be universally shared. He defined beauty as "pleasure objectified."

MEDIUM, FORM, EXPRESSIVENESS. Santayana added to this definition of beauty a threefold distinction between the materials of a work of art, its form, and its expressiveness. Of these, the first two are intrinsic features of the work of art, which thus consists of sensuous elements that have varying degrees of aesthetic value by themselves, and a form or arrangement by means of which these elements are unified and which has its own distinctive value. This synthesis, which constitutes form, is "an activity of the mind." While Santayana throws out suggestions as to how the nature of our perceptual apparatus may determine *which* forms give pleasure, these suggestions are never developed, and there is a heavily mentalistic cast to his whole account of aesthetic experience. This is particularly true of his treatment of expression, which is the power of a work of art to suggest images and ideas that, by becoming associated with it, enhance its value. These associated values may be aesthetic, practical, or moral; or they may be intellectual, as they are in the case of those forms of art, for example, tragedy, which present the ugly as well as the beautiful, and whose value thereby consists in satisfying our desire to know life as a whole. In the end, however, while these distinctions of materials, form, and expression have the validity proper to their spheres, the experience of beauty remains, according to Santayana, unique and unanalyzable.

FUNCTION OF ART. In *Reason in Art* Santayana was concerned with the place of art, as one good among many, within the moral economy of the life of reason. He distinguished between the practical arts and the fine arts and explained the emergence of the latter from the former through the gradual growth of an appreciation of the intrinsic value of what originally had merely instrumental value. Applying this principle, Santayana described the development of music, poetry, and the plastic arts, and in each case attempted to relate the special features of the artistic medium to the mode of abstraction and selectivity that is peculiar to a given art form. He treated all works of art as more or less abstract symbolizations of the natural environment and interests of human beings, and as being animated by an internal "dialectic" of their own

through which the moral and dramatic unities of our experience are indirectly expressed. There can be no absolute or universal principles for criticizing works of art, since our critical judgments are simply the corrections or modifications that our aesthetic preference undergoes in the wake of experience; and there is no a priori guarantee that these corrections must be convergent. The ultimate justification of art is simply that it adds greatly to human enjoyment, and thus to human happiness.

THE LIFE OF REASON

Santayana intended *The Life of Reason; or The Phases of Human Progress* (1905–1906) as a naturalistic biography of the human mind, but as he himself pointed out, it was at least partially inspired by G. W. F. Hegel's *Phenomenology of Mind.* What appealed to Santayana in that work and similar ones in the idealistic tradition was the idea of sympathetically espousing the changing perspectives—scientific, moral, religious, and aesthetic—by which the mind progressively defines its relationship to its natural milieu. By beginning with *Reason in Common Sense,* he hoped to avoid the fundamental error of the idealists, which was to lose all sense of the dependency of this evolution upon a nonmental nature and of its responsiveness to the strains and stresses of our animal being. For the fraudulent dialectical necessity that Hegel had imposed on human history, Santayana proposed to substitute an appraisal—in the broad sense, a moral appraisal—of the contribution made by each of these phases of human development to the ideal of a rational and happy life.

REASON AND IMAGINATION. In *Reason in Common Sense,* the discovery of natural objects is described as the first and irreversible achievement of human reason operating upon the materials of sense experience. Knowledge of these objects is inevitably representative and indirect, and the relationship of thought to reality must be conceived as an ideal correspondence and not as a material appropriation. Coordinate with these "concretions in experience" are "concretions in discourse," or concepts which sustain among one another all manner of "dialectical" relationships; and the active elaboration of these is the generic activity of imagination. Imagination becomes understanding when, almost by accident, some of its structures prove to be faithful transcriptions of a sequence of natural events; but even when the understanding is most successful, there remain unassimilable traits of experience which, at best, have a tangential relation to the natural order.

Toward the free creative activity of the imagination itself, Santayana maintained a dual attitude. It must not, he said, be allowed to impose itself as a literal rendering of what exists, as it all too often attempts to do. When it is allowed to do so, it can only produce a fantastic physics in which dramatic and moral unities are substituted for unities of fact and real process. In another sense, however, the life of reason is the life of the imagination, and its function of idealization and symbolic transformation yields the highest and purest enjoyments of the mental life. Even when the imagination becomes practical, as it does in science, it is the intrinsic aesthetic value of its creations, and not their ulterior practical use, which gives them a place within the life of reason. But at the same time that he praised the imagination, Santayana continually warned against the tendency to confer substantial reality upon the essences it elaborates and to assign to them a causal efficacy within the order of nature. The only power that Santayana was willing to attribute to consciousness itself was that of conferring meaning and ideal unity upon events, and it is in this sense that he described himself as being a materialist.

RELIGION. If Santayana's theory of the imagination finds its most natural application in his treatment of art, an area in which the claim to any literal validity is reduced to a minimum, the case of religion, which he considers in *Reason in Religion,* Vol. III of *The Life of Reason,* is somewhat different. Religion, Santayana said, is a poetic transformation of natural life in the interest of the moral ordering of that life, even though each religion is typically regarded by its followers as embodying a literal truth. Religion is myth, and it presents "an inverted image of things in which their moral effects are turned into their dramatic antecedents." Because it is myth, religion must not be judged by the inappropriate standard of literal truth, but on the basis of the imaginative richness and comprehensiveness of its reorganization of our moral experience. One's religion is in fact something like one's language or nationality—a native idiom of the moral life which may have its imperfections, but which is both difficult and unwise wholly to abandon. Mystical religions are those that effect vast simplifications of the moral life by excluding all but one element in the natural life, while fanatical religions are those that suppress, on the authority of their own unique truth, all forms of moral poetry other than their own. In Santayana's view, both are inimical to the true value of religion, which is the encouragement it gives us to live in the imagination. True religion stimulates both piety, which Santayana defined as "man's reverent attachment to the sources of his being and the

steadying of his life by that attachment," and spirituality, which liberates us from the harsh realities of animal need and desire by interposing an ideal meaning—one that assigns to the goods of this world their proper and subordinate place.

What is paradoxical in Santayana's philosophy of religion is the fact that while he treated all religions as having, at best, a symbolic or expressive truth, he severely condemned the liberals and "modernists" who have attempted, while remaining within the church, to substitute for the literalistic dogmatism of the past a view of religion that in many respects resembles the one held by Santayana himself. It seems inconsistent to deny that a claim to literal truth is essential to religion and at the same time to require that those who surrender this claim must leave the church. This is perhaps a special case of a general paradox resulting from the fact that while Santayana declared "spirit" to be wholly inefficacious, it is an intrinsic feature of the life of reason that spirit should view itself as having efficient power. One may also speculate as to whether Santayana's distaste for views resembling his own, when they become more than the private insights of detached and passive observers and are applied to the task of modifying some institution such as a church, did not itself express a social attitude and a partisanship that cannot claim any special philosophical justification.

SOCIAL THEORY. Santayana's theory of society is stated in *Reason in Society,* Volume II of *The Life of Reason,* and also, in expanded form, in *Dominations and Powers* (1949), his last major work. In the main, social life is assigned a subordinate role within the life of reason. Its principal task and justification is the generation of, and care for, human beings, and it serves ideal ends only incidentally. Society originates in the reproductive instinct, and while this instinct lends itself readily to imaginative development, it finds its ultimate fruition in institutions (the family, the army, the state) that are predominantly practical in nature and, at best, capable of a retrospective idealization. It is, of course, possible for individuals to become associated with one another outside the disciplinary framework of these primary institutions, and when they do so freely, on the basis of a common allegiance to an ideal, they form what Santayana called a "free," or "rational," society. Patriotism is the loyalty they feel to such societies; but the deepest loyalties of the life of reason are not to anything actual, but to the ideal presences of which, Santayana said, our human partners in the pursuit of the ideal, as well as we ourselves, are at best imperfect symbols. Thus it turns out that the true society—the only society that is a perfect instrument of the life of reason—is the society of the mind and of the essences it entertains.

If Santayana's theory of society expresses, as indeed it does, a profound lack of interest in the practical concerns by which any human society is principally animated, he was nevertheless not without his own strong preferences with regard to a certain ordering of society. A pervasive animus against democracy and liberalism runs through all his discussions of society and is perhaps most noticeable in *Dominations and Powers* (1949). Human society, Santayana argued, is necessarily aristocratic and hierarchical, and egalitarian democracy, which would put an end to the injustice that social inequality so often generates, succeeds only in destroying the interest of life by denying or attempting to suppress our inevitable human diversity. An authentic and "natural" aspiration to some good expresses itself in the form of an authoritative direction of the more passive members of a society and shapes their lives in the light of this aspiration's own moral vision. Accordingly, Santayana frequently tended to identify strong authoritarian government with the natural bent of a self-assertive vitality and uniformly treated liberalism as an incoherent and sterile principle of dissolution, roughly comparable in its inspiration and effect to the Protestant principle in the province of religion. Both liberalism and the Protestant principle are expressions of that romantic individualism that Santayana was willing to tolerate as a kind of playful self-deception of the "inner life," but which he abominated whenever it took itself seriously and became a principle of action directed toward correcting the "natural" order of things.

MORALITY. Strangely enough, it is in *Reason in Science,* Volume V of *The Life of Reason,* that Santayana's fullest exposition of his views on morality is to be found. In this work he distinguished between "rational" morality and the morality that is either "prerational" or "postrational." Rational morality is no longer the straightforward hedonism of *The Sense of Beauty,* for Santayana now recognized that there must be a principle of selective preference among possible enjoyments. But he still regarded our adoption of such an ideal standard as a matter of temperament and natural inclination; and even the attempt to achieve a comprehensive integration of diverse satisfactions, which is what distinguishes rational morality, is presented as just one possible attitude toward life. Rational morality and the moral philosophy associated with it, Santayana argued, are concerned with what is really good, and they require a highly developed capacity for sympathetic understanding and assessment of all

competing goods; but in the end, what is really good can only be what genuinely expresses some vital bias of our natures. By contrast, prerational morality is the unreflective life of primary impulse, which cannot conceive the possibility of alternative goods nor support the discipline entailed by a principled organization of the moral life. Postrational morality, finally, is an essentially religious abandonment of the hope for a rational ordering of human life in favor of some otherworldly ideal. Its sole strength, as Santayana observed, lies in the remnant of natural assertiveness that survives in its condemnation of the works of the natural man and the desperate energy with which a single and exclusive regimen of life is proclaimed to be the sole means of salvation.

SCIENCE. Santayana's attitude toward science, as one phase of the life of reason, was an inconsistent mixture of hospitality and indifference. Convinced as he was that all causal efficacy belongs to physical nature, he was strongly inclined to accept the claim of science to exclusive authority in the determination of what is really true. Natural science is at once an extension of common sense and a uniquely successful application of "dialectics," that is, the logical elaboration of terms of thought, or "concretions in discourse," to the study of the physical world. The ideal of such a science would be a closed, mechanistic, and materialistic system, and Santayana believed that progress in the sciences of man, notably psychology, required the adoption of this ideal. But beyond this recognition of the authority of science, Santayana had no detailed interest in its findings and only a very limited belief in its power to contribute to those ideal values that are the true substance of the life of reason. It deals, after all, with only one of many possible worlds; and while the discipline of fact to which it subjects the mind is infinitely preferable to the projection upon the world of some moral fable of our own devising, the highest form of intellectual freedom is still to survey the field of ideal possibilities without any sense of an obligation to describe or a fear of misdescribing any actual state of affairs.

SCEPTICISM AND ANIMAL FAITH

In *Scepticism and Animal Faith* (1923), Santayana undertook the extensive recasting of his whole system of thought; to which reference has been made above. The reformulation was to consist in the substitution of a set of ontological distinctions for the introspective psychology of his earlier writings. Properly speaking, this work is an introduction to, and a partial summary of, the main doctrines of *Realms of Being* (1927–1940). It begins with an attempt to radicalize, and thus to overcome, the idealistic skepticism concerning the existence of an external world that has been a central theme of Western philosophy since René Descartes. The argument is that if we limit ourselves to what is immediately given (and therefore incapable of being doubted), not only our belief in an external world, but also our belief in the existence of the self, of other selves, and of a past and a future is undermined. All that remain are certain characters or essences that bear no relationship to things or events and cannot properly be said to "exist."

Santayana's point is that a genuine skepticism, pushed to its logical extreme, is just as fatal to the "mind" of the idealists as it is to the matter they were prepared to abandon. In a positive sense, the upshot of such skepticism is to reveal essence as the primary and incontestable mode of being; but it is practically and psychologically impossible for human beings to recognize only essential being. "Animal faith" thus supervenes upon the intuition of essence and posits the existence of a world of things and events that transcends immediate intuition. In one sense this belief is quite baseless, since there cannot, in a strict sense, be proof that anything exists; but in another sense this belief is the beginning of wisdom. In this conception there is no great shift away from the view set forth in *The Life of Reason*. The chief difference, however, is that in *Scepticism and Animal Faith* the commitment to existence and to substance (which in the earlier work was presented retrospectively as the first great achievement in the history of consciousness) is first dramatically revoked and then reinstated by the individual mind. But with respect to the logical status and practical necessity of this belief, Santayana's views would not appear to have undergone any significant change.

THE REALMS OF BEING

The Realms of Being is a detailed characterization of the four major modes of being or basic categories that emerge from the skeptical self-interrogation of consciousness. The modes of being consist of essence and matter, as noted above, and two derivative modes, truth and spirit.

ESSENCE. The being of essence is first carefully distinguished from certain adventitious notions that have been associated with it in the history of Western philosophy. Among these are the views that attribute causal efficacy or some special moral or aesthetic status to essences as such, and also the views that envisage essence only in the context of some mental activity such as "abstraction" or "imagination." Properly conceived, the being of essence

consists simply in the self-identity of its character. Since this intrinsic character involves no reference to any location in space or time, essences are universal and repeatable. They are infinite in number and yet collectively compose one absolute essence in "Pure Being," which is common to all essences. Essences are logically discrete and individual, and one essence can "imply" another only if it is first stipulated that the relationship is that of a whole to one of its parts and that no logical necessity governs the constitution of such wholes. Essences may be exemplified in the realm of matter, but they need not be; and even when they are, the things and events that are the bearers of these ideal characters have a quite different mode of being.

MATTER. The "indispensable properties" of the material mode of being are spatial extension and temporal process. Matter exists contingently and is therefore unstable and evanescent; but it also maintains a dynamic continuity, through change and can in this sense be called "substance." It is external to and independent of consciousness; and it is ultimately unknowable, since we know it only through the essences it exemplifies—and these are radically incapable of representing the element of process and diffusion that is peculiar to the realm of matter. Organisms are part of that realm, and the psychological histories (as distinct from the pure consciousness) of human beings can be understood only by reference to the behavioral unity that Santayana calls the "psyche."

TRUTH. Originally, Santayana had intended to establish only three "Realms of Being," and in fact the Realm of Truth that he later added has obvious affinities with both essence and matter. Truth is the truth about matter, or what exists, and yet it is independent of existence both because "no fact can be a description of itself" and because even if nothing existed, it would still be true that nothing did exist or that just such and such things had existed in the past. Truth is "the sum of all the propositions," and as such it represents a certain selection from the infinite essences or character that things might have had. Truth is timeless and independent of all beliefs. There are no necessary truths, and even the propositions of mathematics are only contingently true since it is simply an accident if they correctly describe the material world.

SPIRIT. Spirit, as Santayana used the term, is simply pure transcendental consciousness, and as such it must be distinguished from its physical basis (the "psyche") and from particular mental events. The only criterion of the existence of spirit is internal; and it exists contingently. It is entirely passive in its relation to physical nature, and its sole function is pure intuition, which, Santayana says, is "the direct and obvious possession of the apparent without commitments of any sort about its truth, significance, or material existence." The unities of intuition are simply individual essences and are not the product of any mental machinery. By itself, intuition is not cognitive. Considered simply as a skein of meanings, the life of intuition may acquire a unity and a life and even a kind of freedom that lacks the power to intervene in the world but is nevertheless the highest and purest human good.

To some extent, *The Realms of Being* effects a clarification of Santayana's earlier views, although it may be doubted whether he was ever in much danger of being taken for an idealist. Unfortunately, the style of the later book is even more luxuriant than that of *The Life of Reason,* and Santayana's unwillingness to argue technical philosophical issues was still as strong as ever. If what he hoped to present in *Realms of Being* was, as he says, a language in which the great distinctions to which we all have recourse would be clearly marked out, his success must be judged to be only very partial. All doctrines of transcendental subjectivity, including Santayana's, engender immense difficulties which cannot be resolved unless the philosopher is more inclined to meet criticism on some ground other than the assumed truth of his own views. In *Realms of Being,* there are very few signs, of such a disposition on Santayana's part.

CRITICAL WORKS

Santayana was not just a philosopher in his own right but also a critic, both philosophical and aesthetic. Several of his books, among them *Interpretations of Poetry and Religion* (1900), *Three Philosophical Poets* (1910), *Winds of Doctrine* (1913), *Character and Opinion in the United States* (1920), *Platonism and the Spiritual Life* (1927), and *Obiter Scripta* (1936), are made up of critical studies of systems of thought as diverse as the pragmatism of William James and the atomism of Lucretius; and in many ways, Santayana was at his best as a critic; and in many ways, Santayana was at his best as a critic. In spite of the severity of his judgments and his tendency to use both philosophers and imaginative writers as stalking horses for his own philosophical purposes, he seldom failed to make some telling observation or incisive criticism that had a validity independent of his own special point of view. At the same time, it must be noted that in his critical essays he too often affected an Olympian manner that only partially concealed the strongly personal

character of his tastes and distastes both for individuals and ideas.

See also Aesthetic Judgment; Aesthetics, History of; Art, Expression in; Beauty; Descartes, René; Essence and Existence; Husserl, Edmund; Kant, Immanuel; Lotze, Rudolf Hermann; Realism; Russell, Bertrand Arthur William; Skepticism, History of; Spinoza, Benedict (Baruch) de; Value and Valuation.

Bibliography

WORKS BY SANTAYANA

Major Philosophical Works

The Sense of Beauty. New York: Scribners, 1896.

The Life of Reason; or The Phases of Human Progress. 5 vols. New York: Scribners, 1905–1906; 1 vol. ed., New York, 1954.

Scepticism and Animal Faith. New York: Scribners, 1923.

Realms of Being. 4 vols. New York: Scribners, 1927–1940; 1 vol. ed., 1941.

"Three Proofs of Realism." In *Essays in Critical Realism.* New York, 1941.

Dominations and Powers. New York: Scribners, 1949.

Critical Writings

Interpretations of Poetry and Religion. New York: Scribners, 1900.

Three Philosophical Poets. Cambridge, MA: Harvard University Press, 1910.

Winds of Doctrine. New York: Scribners, 1913.

Egotism in German Philosophy. New York: Scribners, 1915.

Character and Opinion in the United States. New York: Scribners, 1920.

Platonism and the Spiritual Life. New York: Scribners, 1927.

Some Turns of Thought in Modern Philosophy. New York: Scribners, 1933.

Obiter Scripta, edited by J. Buchler and B. Schwartz. New York: Scribners, 1936.

The Idea of Christ in the Gospels. New York: Scribners, 1946.

Other Works

The Last Puritan. New York: Scribners, 1936. Santayana's only novel, which is also a "memoir" of considerable interest.

Persons and Places. 3 vols. New York: Scribners, 1944–1953. Autobiography.

The Letters of George Santayana, edited by Daniel Cory. New York: Scribners, 1955.

WORKS ON SANTAYANA

Duron, Jacques. *La pensée de George Santayana.* Paris: Nizet, 1950. A massive and detailed study of Santayana's whole philosophy.

Howgate, George W. *George Santayana.* Philadelphia: University of Pennsylvania Press, 1938. A less technical study of Santayana, mainly as a critic and aesthetician.

Munitz, Milton. *The Moral Philosophy of Santayana.* New York: Columbia University Press, 1939.

Schilpp, P. A., ed. *The Philosophy of George Santayana.* Vol. II in the Library of Living Philosophers. Evanston, IL: Northwestern University Press, 1940. A collection of critical essays by a group of philosophers, among them Bertrand Russell, with an autobiographical statement by Santayana, a detailed reply to his critics, and a detailed bibliography. The most useful single work on Santayana's philosophy.

Singer, Irving. *Santayana's Aesthetics.* Cambridge, MA: Harvard University Press, 1957.

Frederick A. Olafson (1967)

SARTRE, JEAN-PAUL
(1905–1980)

Jean-Paul Sartre, French existentialist philosopher and author, was born in Paris where he attended prestigious *lycées* and then the École Normale Supérieur from 1924 to 1928. After passing his *agrégation* the following year, he taught in several *lycées* both in Paris and elsewhere. In 1933, he succeeded Raymond Aron (1905–1983) as a research stipendiary for a year at the Institut Français in Berlin, where he immersed himself in phenomenology, concentrating on Edmund Husserl but also reading Max Scheler and some Martin Heidegger. In the years following his return to France, he published several phenomenological works as well as the philosophical novel *La nausea* (*Nausea*) (1938) that brought him public recognition. He resumed his teaching till conscripted into the French Army in 1939. After serving ten months as a prisoner of war chiefly in Trier, where he taught Heidegger's *Being and Time* (1962) to several imprisoned priests and continued writing his masterwork caps for *L'etre* (*L'être et le néant*) (Being and nothingness) (1943), he returned to Paris for three more years of *lycée* teaching. Soon he was able to make his living from his writing and would never teach again. He was involved in a short-lived resistance movement of intellectuals that included Maurice Merleau-Ponty and Simone de Beauvoir, the latter his lifelong companion. With these and several others, he founded the journal *Les temps modernes* (Modern times), its first issue appearing in October 1945, which quickly became the voice of existentialism and remains a leading literary and political publication to this day.

In the aftermath of the war, Sartre emerged as the leader of the existentialist movement, the quasi manifesto of which he delivered in a famous address subsequently published as *L'existentialisme est un humanisme* (1946). By then, he was world famous. He used his celebrity to promote political and social causes of the Left in accord with the theory of *committed literature* introduced in a series of essays published as *Qu'est-ce que la littérature?* (What is literature?) in *Les temps modernes* (1947). He wrote a number of short stories, novels, and plays as well

as several studies of the lives of famous authors, including his autobiography, *Les mots* (The words), for which he was awarded the Nobel Prize in Literature (1964) and which he declined. After an unsuccessful association with an incipient noncommunist nonparty of the Left, he abandoned organized politics. His relations with the Communist Party ran hot and cold. Initially vilified by the party as a bourgeois individualist, he gradually became a fellow traveler, using different standards with which to judge the East and the West during the Cold War. But after the Soviet occupation of Budapest in 1956 and the crushing of the Prague Spring in 1968, he turned against the French Communist Party and moved farther Left, titling one interview: "*Les Communistes ont peur de la revolution*" (The Communists are afraid of revolution) (*Situations*, VIII, 1969). In 1960 he published his second major philosophical work, the first volume of *Critique de la raison dialectique, précédé de questions de méthode* (The critique of dialectical reason) preceded by a kind of preface *Questions de méthode* (Search for a method) that had appeared in *Les temps modernes* in 1957. This marked his theoretical shift from a philosophy of consciousness and subjectivity to one of dialectical praxis (human activity in its socioeconomic milieu). Many see this as the theoretical basis for the student revolt known as *the events of May, 1968* that constituted a turning point in French cultural life.

Throughout these years of political turmoil and despite his proclaimed abandonment of imaginative literature in favor of political action, Sartre continued to labor on his multivolume study of Gustave Flaubert's life and times, *L'idiot de la famille; Gustave Flaubert de 1821–1857* (The family idiot: Gustave Flaubert de, 1821–1857) (1971–1972). After a number of strokes in the 1970s left him almost totally blind, he began a series of interviews with former Maoist activist Benny Lévy (1945–2003), then serving as his secretary, that he announced would leave none of his earlier positions unchanged. The proposed elements of an *ethic of the* '*We*', as he called it, appeared in three issues of the weekly magazine *Le nouvel observateur*. Titled "*L'espoir maintenant*" (Hope now) these interviews constitute his last publication during his lifetime. After his death on April 15, 1980, the funeral cortege was joined by thousands of people in the largest spontaneous demonstration Paris had seen since the death of France's president Charles De Gaulle (1890–1970). France had lost "the conscience of his time," proclaimed the lead essay in a major journal (*Magazine littéraire*, September 1981) and the immense crowd of mourners seemed to agree.

A PHILOSOPHER OF THE IMAGINATION

Starting with his thesis for the diplôme d'études supérieures titled "The Image in Psychological Life: Role and Nature" (1926) Sartre exhibited a strong interest in the realm of the imaginary. This becomes the object of two of his early publications, *L'imagination* (1936), a reworking of the earlier thesis, and the more important *L'imaginaire* (The imaginary) (1940), in many ways the key to his subsequent thought. For what he attributes to imaging consciousness in the latter—namely, that it is the locus of possibility, negativity, and lack—is precisely how he will later characterizes being-for-itself or consciousness in *Being and Nothingness*. Imaging consciousness becomes the paradigm of consciousness in general for Sartre.

From this follow several characteristic features of his aesthetics, ethics, and political theory as well as the *choice* of the imaginary on the part of the subjects of his existentialist biographies or *psychoanalyses*. It also explains the ease with which he employed the method of *free imaginative variation of examples* (eidetic reduction) from Husserlian phenomenology in constructing his philosophical position. Many of his *arguments* are descriptive in nature, exhibiting Husserl's remark that the point of phenomenology is not to explain but to get us to *see*. Moreover, the matching of imaging consciousness with conceptual analysis in Sartre's works serves to bridge the commonly perceived distance between philosophy and imaginative literature, helping us better appreciate the philosophical approach to literature and the literary approach to philosophy that mark his writings. His novel *Nausea*, for example, anticipates, and his play *No Exit* (1944) applies, theses and themes of *Being and Nothingness* in concrete fashion.

Sartre remained faithful to the descriptive method of phenomenology throughout his career. Even when he introduced the dialectical *progressive-regressive* method in *Search for a Method*, it was to be preceded by a phenomenological description of the situation at hand. But he was not an uncritical reader of Husserl. In a major essay, "Transcendence of the Ego," composed while in Berlin but published in 1937, Sartre defends what Aron Gurwitsch called *a nonegological conception of consciousness*. The *of* in the title denotes both a subjective and an objective genitive: The transcendental ego of Husserlian phenomenology has been rendered unnecessary (transcended) whereas the empirical ego (the subject of our reflective knowledge and scientific study) *transcends* consciousness in the sense that it is *other* than the conscious-

ness one has of it. This allows Sartre to distinguish between an autonomous, *prereflective* consciousness that is impersonal or *prepersonal* and the realm of reflective awareness that constitutes our psychological life, which he will call the *Psyche*. He wrote a lengthy manuscript on the latter, only a portion of which was ever published—*Esquisse d'une théorie des émotions* (Sketch for a phenomenological theory of the emotions) (1939).

One of the core theses of phenomenology is the claim that all consciousness is consciousness *of* an other-than-consciousness. Consciousness simply is this aiming at or *intending* an object. This is Husserl's famous thesis of *intentionality* as the defining characteristic of the mental. Perhaps no other phenomenologist has pursued the thesis of intentionality with such consistency as Sartre, even to the point of accusing Husserl, rightly or not, of having betrayed this principle by his understanding of mental images as simulacra *inside* the mind. Sartre will insist that if images are conscious, then they, too, are ways of intending the world as are our emotions. The challenge is to articulate the distinguishing features of these various ways of being *in-the-world*, an expression Sartre adopts from Heidegger.

In *The Imaginary* Sartre undertakes the task of describing the defining characteristics of the image. Relying on the evidence from his reflective description of our prereflective awareness, he identifies four essential features of the image:

1. *The image is a consciousness* rather than an object *inside* consciousness (Sartrean consciousness has no inside; it is essentially *outside*, in-the-world). The image is a relationship to an object. Hence, it is more accurate to speak of *imaging consciousness* than of *images*. The latter term suggests miniatures that we project outside the mind, an example of what Sartre terms *the illusion of immanence*, which is contrary to the intentionality of consciousness.

2. In contradistinction to perception, which must grasp its object in *profiles* that it synthesizes into a perceptual judgment of identity (these are profiles of one and the same cube that cannot all be given simultaneously) imaging consciousness presents its object all at once (we see in the object only what we place there; the image teaches us nothing). Whereas the perceived object *overflows* our perception of it and invites further investigation, in the case of imaging consciousness, what you imagine is what you get. The *studying* of an imagined object is actually the sequential *viewing* of a series of imagings. Sartre calls this the phenomenon of *quasi-observation*. I can synthesize the series into the object of flesh and blood (my friend Peter, for example) that I could perceive, were he available for perception, but ex hypothesi, as imagined, he is unavailable.

3. *Imaginative consciousness posits its object as a nothingness*. Sartre describes this as making its object *present-absent*, that is, present but out of the circuit of my perceptual beliefs that define the *real*. The realm of the imaginary is what Sartre designates the *irreal* as distinct from the *unreal*, which could apply to the perceptual or the conceptual realm. Following Husserl, Sartre allows for just four types of presence-absence: One can imagine the object as nonexistent (unicorns), as absent (Peter as not here), as existing elsewhere (Peter in Berlin), or in a neutral mode that simply prescinds from its existence (as with ideal objects, for example). This is what distinguishes my awareness of the imagined tree from that of the perceived one, which is grasped as present in its *materiality*. Sartre will elaborate this *nothingness* when he describes the *othering* or *nihilating* nature of consciousness in general in *Being and Nothingness*.

4. Imaging consciousness is *spontaneous*, another feature that Sartre will later extend to consciousness sans phrase. This characteristic denotes the prereflective and implicit (Sartre calls it *nonthetic*) awareness that imaging consciousness has of its creative power as it sustains the object in presence-absence. Sartre will speak of this as an awareness of freedom, which he already extends to prereflective consciousness across-the-board and which he will later liken to Descartes's notion of God's power to conserve in existence the created world.

Much of Sartre's aesthetic theory turns on this idea of the image, which he defines as: "an act that aims in its corporeality at [intends] an absent or nonexistent object, through a physical or psychic content that is given not as itself, but in the capacity of analogical *representative* of the object aimed at [intended]" (Sartre 1940/2003, p. 22). As intentional, consciousness has no contents but it does have objects. In the case of aesthetic objects such as the portrait of Charles VIII or the playing of the Appassionata Sonata, the artifact, say the physical painting or the musical performance, serves as *analogon* for the creative imagination of artist and public alike. By our assuming the *aesthetic* attitude, that is, by *derealizing* the perceptual object, the artifact serves as analogon for making present-absent (re-presenting) this particular aesthetic object. Sartre emphasizes that the imaging act is a synthesis of cognitive and emotional intendings. But his analysis

attends chiefly to the primary role of imaging consciousness in this *derealizing* act.

To indicate the pervasiveness of imagination in Sartre's thought, it suffices at this point to mention the role reserved in his existentialist ethic for the *image* of the kind of person I want to be that is implicit in my moral choices, a clear reference to the phenomenological ethics of Scheler. Nor should we overlook the guiding ideal of the *City of ends* throughout Sartre's political philosophy. And when we recall its character as the locus of negativity, possibility, and lack, the presence of the imagination appears as far-ranging as consciousness itself.

AN ONTOLOGIST

Sartre remarked late in his career that what distinguished him from the Marxists was that he raised the class question starting with being, which is wider than class, whereas they do not. He elaborates his ontology in two major works.

BEING AND NOTHINGNESS. The subtitle of *Being and Nothingness* is "A Phenomenological Ontology." Like Heidegger, whose presence is palpable in this work as was that of Husserl in the earlier ones, Sartre begins his study with the being for whom being is a problem, namely, *human reality* (Heidegger's *Dasein*). By accepting this translation of that basic Heideggerian term, Sartre already seems to be following the anthropological track that Heidegger sought to move beyond. But, in fact, Sartre, too, is concerned with gaining access to being in order to delineate its fundamental modes. Still, his point of access is the immediate experience of the *phenomenon of being* in experiences of boredom, nausea, and the like.

In his novel *Nausea*, Sartre's protagonist experiences the sheer contingency of the tree root that captures his attention, its gratuitous *existence*—and his own: "Every existing thing is born without reason, prolongs itself out of weakness and dies by chance"(*Nausea* 1964, p. 113). Sartre's formal ontology in *Being and Nothingness* will follow from the descriptive analysis of that phenomenon of the being of things. Against idealism, against those who succumb to the *illusion of immanence*, Sartre insists on the *transphenomenal* character of being, that is, its irreducibility to appearances. Showing himself as much the pupil of Henri Bergson as of Heidegger in this regard, Sartre appeals to a *revealing intuition of the phenomenon of being*. But this being is not some Kantian thing-in-itself standing behind the appearances; the phenomenon of being is coterminous with, though irreducible to, the being of the phenomena. The phenomena that the eidetic reduction yields are the objects of knowledge; for example, the kind of knowledge that we gain about the nature of imaging consciousness. Such phenomena are reflective and our awareness of them cognitive. The phenomenon of being is prereflective and noncognitive. It follows that knowledge cannot give an account of transphenomenal being. To attempt to do so Sartre calls *metaphysics*, to which he gives short shrift toward the end of the book.

Using the phenomenological method of descriptive analysis, Sartre discovers three irreducible modes of being, namely, being in-itself, or the inert; being-for-itself, or the spontaneous (consciousness); and being-for-others, or the interpersonal. Though he claims that the for-others is as fundamental as the for-itself, it is clear that being-for-others is inconceivable without the other two, which are conceivable without it. So having distinguished between being and the phenomena, Sartre's descriptive analysis now reveals two radically different regions of being: the transphenomenal *being* of the prereflective cogito or *I think* that precedes and sustains any reflective awareness such as Descartes's *Cogito* or any other phenomena insofar as they are consciousness-relative, on the one hand, and the transphenomenal being of the objects of consciousness, revealed in the experiences of nausea, boredom and the like, on the other.

Pursuing this analysis, Sartre discovers that consciousness, which he will soon call *being for-itself*, simply is the transphenomenal dimension of nonbeing, which he calls *nothingness* (*le néant*), the nothingness of *Being and Nothingness*, whereas *being-in-itself* denotes the dimension of transphenomenal being of the object of consciousness. Each region bears distinctive features. Being in-itself, in Sartre's metaphorical discourse, is thing-like in its solidity and identity. An inert plenum, the in-itself simply is what it is. This region includes any aspect of experience that manifests these properties; for example, substances or the temporal past or any of the *givens* of our experience that Sartre, borrowing from Heidegger, calls our *facticity*. Once other subjects enter the scene and a third, irreducible. dimension emerges, which Sartre calls *being-for-others* (*l'être-pour-autrui*), the scope of facticity expands to include such givens as our reputations, social institutions, and cultural phenomena generally. These, too, are forms of being-in-itself.

Being-for-itself bears contradictory features. As the nothingness of *Being and Nothingness*, the for-itself is the internal negation, or *nihilation*, of the in-itself. Sartre agrees with Heidegger that negativity is not simply a property of propositions but that it is introduced into the world by human reality itself. As evidence, Sartre cites a

whole series of *negativities* (*négatités*), such as our experience of the *fragility* of entities, of absence, of distance, of distraction, of regret, and of lack. (Recall his characterization of imaging consciousness). The for-itself is an exception to the Parmenidean rule of self-identity: Consciousness is nonself-identical. It is always *other* than itself, which is an ontological expression of its intentionality.

That *inner distance* that separates consciousness from itself accounts for three major characteristics of *human reality* (which is the human being as a composite but not a synthesis of these two ontological regions, related as thing and no-thing). First, it gives rise to the three dimensions of original, *ekstatic* temporality whereby human reality *stands out* from the other and from its very self, namely, the past as facticity, the future as existence or project, and the present as presence-to. This is another way of parsing the nonself-identity of the for-itself. A second consequence of this *gap* or *time lag* that consciousness introduces is the ontological freedom that characterizes our existence. Human reality is free, Sartre insists, because it is not a self but a presence-to-self. Part of Sartre's political endeavor after the war is to pursue the kind of *concrete freedom* that completes this abstract freedom as the definition of the human. Finally, it is this nonself-coincidence that accounts for the paradoxical discourse that Sartre adopts with regard to human reality. Besides the traditional paradoxes of temporality that he inherits, the chief paradox is that human reality *is what it is not* (its possibilities) and *is not what it is* (its facticity as *nihilated* by consciousness). On this account, whatever I am, be it my previous choices or the labels others have affixed to me, I am in the manner of *not-being* them, that is, with the possibility of changing my particular stance in their regard. For the quasi motto of Sartrean humanism is that *you can always make something out of what you've been made into*. This is both the burden of our responsibility and the source of our hope.

With the advent of another subject into my world comes another realm of being as well—being-for-others. Ontologically, this gives rise to an additional set of characteristics that belong to the interpersonal dimension of our existence. The existence of the other subject cannot be deduced; it must be encountered. The most dramatic argument for the existence of other subjects is Sartre's eidetic reduction of shame consciousness. His descriptive analysis centers on the experience one has of being caught in the act of looking at a couple through a keyhole. The feeling of shame that registers in bodily changes such as the face turning red is stronger evidence for the existence of other minds, Sartre believes, than any argument from analogy. As he unpacks the experience, in one and the same moment, I become aware of the vulnerability of my embodiedness to the look of the other. In other words, what is revealed in this instant is my prereflective consciousness of being objectified by that gaze of another subject. My experience of objectification is simultaneously my experience of the other as subject. Even if on this occasion I happen to be mistaken about the source of the sound I hear behind me, the experience is indicative of being seen by another.

Though Sartre admits that other, derivative modes of access to being-for-others are available (for example, the existence of cultural objects such as directional signs or language itself), he insists that *the look* (*le regard*) is the basic form of interpersonal relation, and he interprets this gaze as objectifying and alienating. "Conflict is the original meaning of being-for-others" (Sartre 1943/1956, p. 364). The interpersonal is like a game of mutual staredown, each trying to objectify the other. The only type of social philosophy that one can expect from such a thesis is a Hobbesian *war of all against all*. In a famous footnote Sartre concedes that "an ethic of deliverance and salvation" is possible but that this can be achieved only after "a radical conversion" which, he insists, cannot be discussed in that work (Sartre 1943/1956, p. 412). In fact, the elements of an ethic of authenticity are sketched in his posthumously published *Cahiers pour une morale* (Notebooks for an Ethics) composed in 1947–1948, where the basics of this *conversion* are discussed.

Human reality is being-in-situation. *Situation* is composed of facticity and freedom as transcendence; that is, the given that we are always surpassing in our projects. Though he insists that the situation is an ambiguous phenomenon because the precise contribution of each component cannot be determined, it is clear that, as Sartre's sense of social conditioning increases with his shift from abstract to concrete freedom, his respect for the force of circumstance in our situations grows apace. At this stage of his thought, he seems ambivalent as to the limiting and conditioning role of facticity in our actions. But later in life Sartre's sense of what Max Weber called *objective possibility* will heighten and, with it, the claim that fundamental changes in our socioeconomic system are required for abstract freedom to be made concrete. Thus, he will note shortly after the end of the war that "it is the elucidation of the new ideas of 'situation' and of 'being-in-the-world' that revolutionary behavior specifically calls for" ("Materialism and Revolution," Michelson 1962, p. 253).

It is in the context of situation that the concept of bad faith arises. Bad faith is a kind of self-deception, a sort of lying to oneself about the truth of one's situated being. Its most common form consists in collapsing our transcendence (our freedom) into our facticity by appeal to a type of determinism or by simply confessing: *That's just the way I am*. It is a denial of the possibility that consciousness brings to every situation. A related version of this type appeals to the image I wish to present to others or the one they have of me. That, too, is part of my facticity with which I seek to identify in self-deception as if my consciousness did not resist any attempt at full identity. A less common form of bad faith *volatilizes* our facticity into transcendence by choosing to ignore the givens of our situation. This is the bad faith of the dreamer or of the person who flees their past as if it were not part of their situation. But the possibility for self-deception arises from the *dividedness* of our consciousness as prereflective and reflective such that one can be prereflectively aware of more than one knows at the reflective level. Not that one is dealing with two consciousnesses: This deception occurs within the unity of one and the same consciousness.

Since Sartre denies the existence of the Freudian unconscious as he understands it because of its incompatibility with the ontological freedom of human reality, this notion of bad faith cannot appeal to unconscious drives or complexes. What Sartre calls *existential psychoanalysis* aims at dealing with such phenomena as bad faith and fundamental project without appealing to unconscious motives. Its basic premise is that "*man is a totality and not a collection*" (Sartre 1943/1956, p. 568). In other words, at the base of human reality is a fundamental, unifying *choice* that establishes the criteria for all subsequent selections.

We come on the scene having already made that choice, which Sartre believes is guided by the ruling value to consciously be self-identical, that is, to be in-itself-for-itself—an ontological impossibility. This is the meaning of Sartre's famous claim that humankind is a *futile passion*. But how each one lives out that self-defining choice is revealed in the subsequent choices that define a life. "There is not a taste, a mannerism, or a human act," Sartre insists, "which is not revealing" (Sartre 1943/1956, p. 568). The task of psychoanalysis is hermeneutical: to interpret the specific nature of that fundamental choice, that is, the way one acquiesces in or resists that futile passion, by deciphering the symbols of a person's life. What he calls the possibility of *conversion* is the constant threat of altering this basic choice, which haunts our lives. Echoing Danish philosopher Søren Kierkegaard, Sartre calls this the anguish that accompanies the experience of our radical freedom. Admitting that this psychoanalysis has yet to find its Freud, and with a nod toward the work that will occupy a good part of his remaining years, Sartre finds the intimations of such psychoanalysis in certain *successful biographies*.

In many ways, one can read *Being and Nothingness* as an argument moving from the highly abstract (nihilating consciousness, being in-itself and for-itself) to increasingly concrete phenomena such as my *concrete relations with others*, and culminating in the hermeneutic of our particular actions in order to determine the fundamental choice that defines the unity of our lives. Existentialist psychoanalysis both brings this undertaking to a close and opens the door for its application in the several *biographies* that will occupy Sartre's attention over the following decades.

THE *CRITIQUE OF DIALECTICAL REASON*. It was during the war, Sartre insists, that he discovered the philosophical significance of social relations. *Being and Nothingness*, with its emphasis on the *looking/looked-at* model of interpersonal relations, was incapable of explaining the positive reciprocity, collective action, and unintended consequences that a social philosophy requires. In fact, *Being and Nothingness* describes the *we subject* as a "purely subjective *Erlebnis* (experience)" (Sartre 1943/1956, p. 420). Sartre breaks the barrier that confined *Being and Nothingness* to the psychological by introducing the concepts of dialectical praxis, the practico-inert and the mediating third. Together, they account for the dialectical enrichment of individual praxis by group praxis that bears properly social predicates such as rights/duties, power, and function while preserving the freedom and responsibility of the individual, which is a defining characteristic of existentialist thought.

Praxis supplants *consciousness* in the lexicon of the *Critique*. It denotes human activity in its sociohistorical context. Praxis is dialectical in the sense that it both negates and conserves aspects of its object in a totalizing action that advances toward a more comprehensive viewpoint. Thus, the negative reciprocity of two boxers in a match, in Sartre's example, when viewed dialectically, is realizing an enveloping social whole called *professional boxing*, which itself invites a still broader contextualization in various socioeconomic systems, such as racism, colonialism, and capitalism. In Sartre's view dialectical thinking is holistic; unlike analytical reason, it welcomes properly social phenomena as irreducible to purely atom-

istic, usually psychological, relations. While admitting the validity of analytical reason within its domain, at a certain level of abstraction, he notes, the class struggle can be seen as the conflict of rationalities.

Sartre reserves a threefold primacy for free organic praxis in his social ontology: ontological, epistemic, and moral. Ontologically, there are only individuals and real relations among them. Praxis is the constitutive dialectic of social phenomena, which are relational entities constituted by individual praxes. This is true even of group praxis, which is the synthetic enrichment of individual praxes in relation, mediated by each member as *third* to every other. Epistemically, the intelligibility of the group and of other social units is a function of the intelligibility of individual praxis, which is its foundation. Sociohistorical intelligibility is dialectical, and the dialectic is grounded in individual praxis. In other words, Sartre denies the existence of a collective consciousness or subject except insofar as it can be seen as a quality of individuals-in-relation.

Sartre speaks of *comprehension* as the *translucidity of individual praxis*. It assumes the clarity that Sartre has reserved for the prereflective cogito in *Being and Nothingness*. The moral primacy of individual praxis follows from the other two forms. Sartre is intent on preserving the moral responsibility of the group members as well as of those he describes as *serialized* by the mediation of worked matter, such as the television-viewing audience or the crowd waiting for a bus. In either case, whether *the same* in group activity and concern or *other* through the separation effected by the mediation of material things, individuals retain moral and not just causal responsibility for the praxis that sustains such relations.

The second basic component of Sartre's social ontology is what he calls the practico-inert. This complex term introduces aspects of being-in-itself into the realm of action. Sartre describes it as "simply the activity of others in so far as it is sustained and diverted by inorganic inertia" (Sartre 1960/1985, p. 556). Not raw nature, but the *practico*-inert is this mediating factor. It includes the sedimentation of prior praxes whether in the form of socioeconomic systems such as colonialism and capitalism or as alienating forms of thought and behavior such as racism, which Sartre calls a *serial idea*. It constitutes the material memory of a society.

Sartre allows for two fundamental kinds of social reality: the active group constituting the *common field* and the effectively separated though ostensibly united (*serialized*) individuals forming the *practico-inert field*. The practico-inert constitutes *fundamental sociality*. Since Sartre conceives the group as arising through an essential negation of practico-inert seriality, he characterizes the practico-inert ensemble as *the matrix of groups and their grave*. This rich concept is amenable to analytic reason since it is atomistic in nature. But insofar as it occasions counterfinality in the sense that it sustains the boxer's feints and jabs, the conspirator's traps, and the unintended consequences of historical projects, its very antidialectic plays a role in dialectical rationality, conveying the experience of what Sartre calls *dialectical necessity*. Perhaps Sartre's best example of such counterfinality is the flooding and resultant soil erosion caused by Chinese peasants' deforestation undertaken to conserve their land.

But the concept of the mediating third is the key that opens the door to properly group praxis in Sartre's social ontology. There was a concept of the third in *Being and Nothingness*, but this third exercised an objectifying and an alienating function in accord with the looking/looked at model. That concept continues in the *Critique*, where it generates the alienating relations of serial individuals and collectives. But the *mediating* third is a functional concept denoting the group member who is *the same* as the others in common interest and action. As such, it does not objectify or diminish but enriches the responsibility of each in a common practice. Sartre refers to this ternary relation as a *free, interindividual reality*. Simply put, where the practico-inert mediates, human relations are serial; where praxis mediates, these relations are free. And where the practico-inert is modified by material scarcity, Sartre argues, this mediation becomes violent. Such is his bridge between social ontology and history as we know it.

AN EXISTENTIALIST BIOGRAPHER AND HISTORIAN

In *Search for a Method*, reprinted as a kind of preface to the *Critique* but more properly its sequel, Sartre introduces the progressive-regressive method for investigating social phenomena. This hybrid of existentialist psychoanalysis and historical materialism serves as the model for his later biographies, especially his multivolume study of Gustave Flaubert's life and times, *The Family Idiot*. Sartre studies the socioeconomic and cultural structures of Flaubert's life, particularly as these conditioned the choices available to a would-be literary artist in the second and third quarters of the nineteenth century (the *regressive* movement), the better to chart the spiral of Flaubert's *personalization* as artist, novelist, and finally author of *Madame Bovary* (1956) (the progressive stage). The approach is dialectical in its emphasis on the factors that *mediate* these abstract conditions toward their con-

cretization in Flaubert's *choice* of the imaginary, that is, of an artist's life. Indicative of Sartre's increasingly nuanced opposition to the Freudian unconscious is his remark that "everything took place in childhood … a childhood we never wholly surpass" (Barnes 1968, p. 59–60 and 64).

The dialectic expands to include the *objective spirit* of the age, which Sartre characterizes as *culture as practico-inert*. Using an expression that Aron had employed to describe narrative history in general, Sartre calls *The Family Idiot a novel that is true* (*un roman vrai*). Its dialectical interlacing of history and biography render it a properly existentialist approach to historical understanding.

A MORALIST

If Sartre was a philosopher of the imagination and an ontologist, he was above all a moralist in the French tradition of Duc François de La Rochefoucauld and François-Marie Arouet de Voltaire. His earlier philosophy of consciousness, as well as the primacy of praxis in the social ontology of the *Critique*, are conceived to preserve freedom and responsibility that are the hallmarks of vintage existentialist thought in the midst of impersonal forces, and what Louis Althusser (1918–1990) called *structural causality*. In the hyperbolic mode that he favored, Sartre insisted that *we are without excuse*.

In the course of his life, Sartre developed one ethical theory, sketched a second, and gestured toward a third, in that order. The first and best known is his ethic of authenticity. He describes *authenticity* briefly in *Réflexions sur la Question Juive* (Anti-Semite and Jew) (1946) as "having a true and lucid consciousness of the situation, in assuming the responsibilities and risks that it involves, in accepting it in pride or humiliation, sometimes in horror and hate" (Becker 1995, p. 90). This seems to yield an ethical style rather than a content. It stresses doing rather than being in the sense of embracing my ontological condition, namely, that whatever I am, I am in the manner of not-being it, that is, in nihilating it. I am its *creative unveiling*, with the anguish and joy that accompanies that prereflective awareness.

The ethical content emerges in his novels, stories, plays, and biographies, especially his biography of Jean Genet, *Saint Genet: Comédien et martyr* (Saint Genet, actor and martyr) (1952) and is elaborated in his posthumously published *Notebooks for an Ethics*, which discusses such concepts as good faith, generosity, and positive reciprocity. Maximizing concrete freedom of choice and action becomes an increasingly important moral precept as Sartre's social sense confronts exploitative systems and oppressive practices after the war.

Exchanging the vocabulary of *Being and Nothingness* for the discourse of the *Critique* in the notes for two sets of lectures and a collection of unpublished reflections from the same period, Sartre sketches a second, *dialectical* ethics that promotes the value of *integral humanity*. This value includes the moral imperative to satisfy human needs by harnessing the practico-inert. Elsewhere, Sartre envisions a socialism of abundance and the new, currently inconceivable, philosophy of freedom that will follow upon it. These lecture notes seem to turn this ideal into an obligation based on the nonnegotiability of basic human needs. In his last discussions with Lévy, he speaks of an ethic of the *we* that will revise many of his previous claims in this regard. However, these recorded remarks were published only in part, and what is available thus far, despite suggestive insights, does not constitute a coherent moral theory. They remain chiefly of biographical interest.

CONCLUDING OBSERVATIONS

One of the strengths of Sartre's philosophical thought is its insight into the psychological and moral life of individuals and societies. That same gift for imaginative interpretation that fits so well with descriptive phenomenology and makes him a prize-winning novelist and playwright is suspect in the court of conceptual analysis. And once Sartre turns to historical dialectic, the suspicion is compounded. Much of this is simply philosophical bias, which Sartre attempted to address with his distinction between dialectical and analytical reason and their respective logics. But some of it is a reasonable distrust of a lack of rigor evidenced by what Iris Murdoch called Sartre's *great inexact equations*. And then there are his rather extreme political positions and their accompanying moral ascriptions. While one cannot help but admire Sartre's outrage at social injustice and hypocrisy, a remark once reportedly made about Bertrand Russell could be extended to Sartre in this regard: He has the uncanny ability to hit the bull's-eye on the first shot but undermined by a tendency then to splatter all over the target in exaggeration.

Still, Sartre's observations on bad faith and authenticity are now staples in the ethical discourse of our day. And the basic concepts of his social ontology, namely, praxis, the practico-inert, and the mediating third, make a significant contribution that merits the close scrutiny that the prolixity of the *Critique* has denied them. The Cartesian dualism often attributed to Sartre is misap-

plied. His is not a two-substance ontology since only the in-itself is substantial. But a survey of his social ontology in the *Critique* suggests that his dualism is best described as one of spontaneity and inertia, which sends us back to imaging consciousness once more. Perhaps nowhere is the relation between philosophy and imaginative literature more acutely problematized than in Sartre's work. That, too, deserves close attention. Finally, the lessons of Sartrean existentialism speak directly to the renewed interest among our contemporaries in philosophy as a way of life.

See also Beauvoir, Simone de; Bergson, Henri; Cartesianism; Descartes, René; Epistemology; Existential Psychoanalysis; Existentialism; Gurwitsch, Aron; Heidegger, Martin; Historical Materialism; Husserl, Edmund; Kant, Immanuel; Kierkegaard, Søren Aabye; La Rochefoucauld, Duc François de; Marxist Philosophy; Merleau-Ponty, Maurice; Murdoch, Iris; Ontology; Parmenides of Elea; Phenomenology; Russell, Bertrand; Scheler, Max; Voltaire, François-Marie Arouet de; Weber, Max.

Bibliography

For a complete annotated bibliography of Sartre's works, see *The Writings of Jean-Paul Sartre*, edited by M. Contat and M. Rybalka. Evanston, IL: 1973, updated in *Magazine littéraire* 103–104 (1975): 9–49, and by Michel Sicard in *Obliques* 18–19 (1979): 331–347. Rybalka and Contat have compiled an additional bibliography of primary and secondary source published since Sartre's death in *Sartre: Bibliographie, 1980–1992*. Paris: CNRS éditions, 1993.

Bulletin du Groupe d'Études Sartriennes, the *Yearbook of Sartre Studies*, has published annually since 1987 a list of recently published works of Sartre as well as secondary source material in several languages. Information about this publication, its bibliographies, and the work of the *Groupe* can be found at www.jpsartre.org.

WORKS

L'imagination. Paris: Félix Alcan, 1936. Translated by Forrest Williams as *Imagination, A Psychological Critique*. Ann Arbor: University of Michigan Press, 1962.

"La transcendence de l'ego. Esquisse d'une description phénoménologique." In *Recherches philosophiques* 6 (1936): 65–123. Reprinted in edition by Sylvie le Bon, Paris, 1965. Translated by Forrest Williams and Robert Kirkpatrick as *Transcendence of the Ego: An Existentialist Theory of Consciousness*. New York: Octagon Books, 1972.

La nausée. Paris: Gallimard, 1938. Translated by Lloyd Alexander as *Nausea*. New York: New Directions, 1964.

Esquisse d'une théorie des émotions. Paris: Hermann, 1939. New edition, 1960. Translated by Bernard Frechtman as *The Emotions. Outline of a Theory*. New York: Philosophical Library, 1948.

L'imaginaire. Paris: Gallimard, 1940. Translated by Jonathan Webber as *The Imaginary: A Phenomenological Psychology of the Imagination*. London: Routledge, 2003.

L'être et le néant. Paris: Gallimard, 1943. Translated by Hazel E. Barnes as *Being and Nothingness; An Essay on Phenomenological Ontology*. New York: Philosophical Library, 1956.

Réflexions sur la question Juive. Paris: P. Morihien, 1946. Translated by George J. Becker as *Anti-Semite and Jew*. New York: Schocken, 1995.

L'existentialisme est un humanisme. Paris: Gallimard, 1946. Présentation et notes par Arlette Elkaïm-Sartre. Paris: Gallimard, 1996. Translated by Philip Mairet as *Existentialism and Humanism.* London: Meuthen, 1957. Also translated by Bernard Frechtman as *Existentialism*. Secaucus, NJ: Citadel 1957.

Situations. 10 vols. Paris: Gallimard, 1947–1976. Contains reprints of important essays and interviews such as "La Liberté Cartésienne" (Vol. 1, pp. 289–307) and "Matéralisme let Révolution" (Vol. 3, pp. 135–223). Translated by Annette Michelson as "Cartesian Freedom" and "Materialism and Revolution." In *Literary and Philosophical Essays*. New York: Collier, 1962.

Qu'est-ce que la littérature? Les temps modernes 17–22 (1947): no. 17, pp. 769–805; no. 18, pp. 961–88; no. 19, pp. 1194–1218; no. 20, pp. 1410–1429; no. 21, pp. 1607–1641; no. 22, pp. 77–114. Reprinted in *Situations* 2. Paris: Gallimard, 1948. pp. 55–330. Translated by Bernard Frechtman et al. as *What Is Literature? And Other Essays*. Cambridge, MA: Cambridge University Press, 1988.

Saint Genet: Comédien et martyr. Paris: Gallimard, 1952. Translated by Bernard Frechtman as *Saint Genet: Actor and Martyr*. New York: G. Braziller, 1963.

Critique de la raison dialectique, précédé de questions de méthode. Vol. 1, *Théorie des ensembles pratiques*. Paris: Gallimard, 1960. Reprinted in a new annotated edition, 1985. The prefatory essay to the *Critique*, "*Questions de méthode,*" appeared in *Les temps modernes* 139–140 (1957): no. 139, pp. 338–417; no. 140, pp. 658–698. Translated by Hazel E. Barnes as *Search for a Method*. New York: Random House, 1968. An annotated edition of *Questions de méthode* by Arlette Elkaïm-Sartre appeared in 1986. Paris: Gallimard, collection Tel. The *Critique* itself was translated by Alan Sheridan-Smith as *Between Existentialism and Marxism (Essays and Interviews, 1959–70)*. Translated by John Mathews. London: NLB, 1974.

Les mots. Paris: Gallimard, 1964. Translated by Bernard Frechtman as *The Words*. New York: G. Braziller,1964.

L'idiot de la famille; Gustave Flaubert de 1821–1857. Vols. 1–3, Paris: Gallimard, 1971–1972. Revised edition of Vol. 3, 1988. Translated by Carol Cosman as *The Family Idiot: Gustave Flaubert, 1821–1857*. 5 vols. Chicago: University of Chicago Press, 1981–1993.

On a raison de se révolter: Discussions. With Philippe Gavi and PierreVictor. Paris: Gallimard, 1974.

Critique of Dialectical Reason. Vol. 1, *Theory of Practical Ensembles*. London: NLB, 1976. A new edition based on the revised French edition of 1985 appeared in 2004. London: Verso.

Sartre on Theater, edited by Michel Contat and Michel Rybalka, translated by Frank Jellinek. New York: Pantheon, 1976.

Situations. Vol. 10. Paris: Gallimard, 1976. Translated by Paul Auster and Lydia Davis as *Life/Situations: Essays Written and Spoken*. New York: Pantheon, 1977.

Oeuvres romanesques. Paris: Gallimard, 1981.

Cahiers pour une morale. Paris: Gallimard, 1983. Translated by Davie Pellauer as Notebooks for an Ethics. Chicago: University of Chicago Press, 1992.

Les carnets de la drôle de guerre. Paris: Gallimard, 1983. Translated by Quintin Hoare as *The War Diaries of Jean-Paul Sartre: November 1939–March 1940*. New York: Pantheon, 1984.

Scénario Freud. Paris: Gallimard, 1984. Translated by Quintin Hoare as *The Freud Scenario*. Chicago: University of Chicago Press, 1985.

Critique de la raison dialectique. Vol. 2, *L'intelligibilité de l'histoire*. Paris: Gallimard, 1985. Translated by Quintin Hoare as *Critique of Dialectical Reason*. Vol. 2, *The Intelligibility of History*. London: Verso, 1991.

Vérité et existence, edited by Arlette Elkaïm Sartre. Paris: Gallimard, 1989. Translated by Adrian van den Hoven as *Truth and Existence*, edited by Ronald Aronson. Chicago: University of Chicago press, 1992.

L'espoir maintenant: Les entretiens de 1980. Lagrasse: Verdier, 1991. Translated by Adrian van den Hoven, as *Hope Now: The 1980 Interviews*. Chicago: University of Chicago Press, 1996.

Théâtre complet, edited by Michel Contat. Paris: Gallimard, 2005.

SELECTED STUDIES

Anderson, Thomas C. *Sartre's Two Ethics: From Authenticity to Integral Humanity*. Chicago: Open Court, 1993.

Aronson, Ronald. *Sartre's Second Critique*. Chicago: University of Chicago Press, 1987.

Barnes, Hazel E. *Sartre & Flaubert*. Chicago: University of Chicago Press, 1981.

Bell, Linda A. *Sartre's Ethics of Authenticity*. Tuscaloosa: University of Alabama Press, 1989.

Busch, Thomas. *The Power of Consciousness and the Force of Circumstances in Sartre's Philosophy*. Bloomington: Indiana University Press, 1990.

Catalano, Joseph. *A Commentary on Jean-Paul Sartre's Being and Nothingness*. Chicago: University of Chicago Press, 1980.

Catalano, Joseph. *A Commentary on Jean-Paul Sartre's Critique of Dialectical Reason*. Vol. 1, *Theory of Practical Ensembles*. Chicago: University of Chicago Press, 1986.

de Beauvoir, Simone. *Adieux: A Farewell to Sartre*. Translated by P. O'Brian. New York: Pantheon, 1984

de Beauvoir, Simone. *The Force of Circumstances*. Translated by Richard Howard. New York: Putnam's, 1964–1965.

Detmer, David. *Freedom as a Value: A Critique of the Ethical Theory of Jean-Paul Sartre*. La Salle, IL: Open Court, 1988.

Dobson, Andrew. *Jean-Paul Sartre and the Politics of Reason*. Cambridge, U.K.: Cambridge University Press, 1993.

Fell, Joseph P. *Heidegger and Sartre: An Essay on Being and Place*. New York: Columbia University Press, 1979.

Flynn, Thomas R. *Sartre and Marxist Existentialism: The Test Case of Collective Responsibility*. Chicago: University of Chicago Press, 1984.

Flynn, Thomas R. *Sartre, Foucault and Historical Reason*. Vol. 1. *Toward an Existentialist Theory of History*. Chicago: University of Chicago Press, 1997.

Flynn, Thomas R. *Sartre, Foucault and Historical Reason*. Vol. 2. *A Poststructuralist Mapping of History*. Chicago: University of Chicago Press, 2005.

Howells, Christina, ed. *The Cambridge Companion to Sartre*. Cambridge, U.K.: Cambridge University Press, 1992.

Jeanson, Francis. *Sartre and the Problem of Morality*. Translated by Robert Stone. Bloomington: Indiana University Press, 1981.

McBride, William Leon, ed. *Sartre and Existentialism*, 8 Vols. New York: Garland, 1997.

McBride, William Leon. *Sartre's Political Theory*. Bloomington: Indiana University Press, 1991.

Santoni, Ronald E. *Bad Faith, Good Faith and Authenticity in Sartre's Early Philosophy*. Philadelphia: Temple University Press, 1995.

Santoni, Ronald E. *Sartre on Violence: Curiously Ambivalent*. University Park: Pennsylvania State University Press, 2003.

Schilpp, Paul Arthur, ed. *The Philosophy of Jean-Paul Sartre*. La Salle, IL: Open Court, 1981.

Schroeder, William. *Sartre and His Predecessors: The Self and the Other*. Boston: Routledge & Kegan Paul, 1984.

Silverman, Hugh J. *Inscriptions: Between Phenomenology and Structuralism*. London: Routledge & Kegan Paul, 1987.

Stone, Robert, and Elizabeth Bowman. "Dialectical Ethics: A First Look at Sartre's Unpublished 1964 Rome Lecture Notes." *Social Text* 13–14 (1986): 195–215.

Stone, Robert, and Elizabeth Bowman. "Sartre's 'Morality and History': A First Look at the Notes for the Unpublished 1965 Cornell Lectures." In *Sartre Alive*, edited by Ronald Aronson and Adrian van den Hoven, 53–82. Detroit, MI: Wayne State University Press, 1991.

Taylor, Charles. *The Ethics of Authenticity*. Cambridge, MA: Cambridge University Press,1991.

Verstraeten, Pierre, ed. *Sur les écrits posthumes de Sartre*. Brussels: University of Brussels, 1987.

Thomas R. Flynn (2005)

SAVAGE, LEONARD
(1917–1971)

Leonard James Savage was the most influential Bayesian statistician of the second half of the twentieth century. Born November 20, 1917, in Detroit, Michigan, Savage received his PhD in mathematics at the University of Michigan in 1941. He then spent a year serving as John von Neumann's assistant at the Institute for Advanced Study in Princeton, where he was exposed to von Neumann's ideas on game theory and the mathematical modeling of human behavior, topics that became a central focus of Savage's research. In his next position at Columbia University's wartime Statistical Research Group—whose members included such luminaries as Abraham

Wald, Milton Friedman, Harold Hotelling, Fredrick Mostler, and Abraham Girshick—Savage developed an interest in statistics and became convinced that the subject should be grounded on a "personalist" conception of probability. After Columbia, Savage went on to hold academic positions at Chicago, Michigan, and Yale.

Savage's research focused on the mathematical analysis of rational belief and desire, and the advancement of Bayesianism in statistics. His masterpiece, *The Foundations of Statistics* (1954), pursued both these projects by first developing what has come to be the canonical version of subjective expected utility theory, and then attempting to recast all of statistical methodology along subjectivist Bayesian lines.

SAVAGE'S CONTRIBUTIONS TO DECISION THEORY

Savage's most notable contributions to the study of rational behavior were his construction of a general framework for modeling decisions under uncertainty, his systematic defense of subjective expected utility maximization as the hallmark of rational choice, and his innovative account of the role of "personal" probabilities in decision making.

Savage portrays decision making as being a matter of using beliefs about possible *states of the world* to choose *actions* that provide the optimal means of producing desirable *consequences*. Actions are identified with functions from states to consequences, and the agent is assumed to have a preference ranking over all acts at her disposal. Influenced by the behaviorism that dominated the social sciences of his day, Savage interpreted preferences operationally, so that an agent may be said to prefer one act *f* to another *g* if and only if she would be disposed to freely choose *f* over *g*. Overt choices thus function as "observables" in decision theory, and talk about the underlying beliefs and desires that cause them is rendered scientifically respectable by showing how they can be operationally defined in terms of preferences. (Savage's behaviorism remains controversial, but some commentators, e.g., Joyce (1999), regard it as inessential to his overall account of rationality.)

Following Frank Ramsey (1931) and Bruno de Finetti (1937), Savage invoked the hypothesis of subjective expected utility maximization to forge a link between empirically measurable preferences and hidden beliefs and desires. Given a probability function *P* defined over states of the world, and a utility function *u* defined over consequences, the expected utility of an act *f* is the probability-weighted average of the utilities of *f*'s consequences. When there are finitely many states, $s_1, s_2, \ldots, s_n$, this expected utility is defined as $Exp_{P,u}(f) = P(s_1)u(f(s_1)) + P(s_2)u(f(s_2)) + \ldots + P(s_n)u(f(s_n))$. Savage maintained that an agent's preferences can only be deemed rational to the extent that they can be represented as ranking acts according to increasing subjective expected utility.

To establish this conclusion, Savage proposed that any rational preference ranking should satisfy a specific system of axiomatic constraints. The central axiom is the *sure-thing principle*, which states, roughly, that for any acts *f* and *g*, and any event *E*, if *f* is preferred to *g* both conditional on *E* and conditional on not-*E* then *f* is preferred to *g* outright. Savage went on to prove that any preference ranking satisfying his axioms implicitly defines a unique subjective probability *P*, which represents the agent's degrees of confidence in various states, and a utility *u*, which gauges the strength of her desires for consequences. The agent prefers *f* to *g* just in case $Exp_{P,u}(f) > Exp_{P,u}(g)$. In this way, the hypothesis of expected utility maximization allows us to extract degrees of belief and desire from rational preferences.

Many objections to Savage's theory misinterpret it as a descriptive account, but it was clearly meant to be prescriptive. The most serious doubts about the theory's normative import concern the status of the sure-thing principle, which some critics see as improperly prohibiting certain sorts of rational aversions to risk or uncertainty. Savage always regarded such worries as misguided, and steadfastly defended the principle's normative credentials. Many people agree with him, as evidenced by the fact that Savage's theory, or its close variants, remain central to treatments of rational decision making across the social sciences.

SAVAGE'S CONTRIBUTIONS TO STATISTICS

Savage maintained that the subjective or "personal" probabilities that figure into decision making should serve also as the basis for statistical reasoning. He implacably opposed the frequentist paradigm that had come to dominate statistics during the 1930s and 1940s. In *Foundations* Savage had tried to incorporate the methods of frequentist statisticians, like Ronald A. Fisher and Jerzy Neyman, into his personalist framework, but by the end of his career he had entirely "lost faith in the devices of the frequentist schools" (Savage 1954). In the second edition of *Foundations* (1972), written six months before his death, he rejects as "ill-founded" such frequentist devices as minimax rules, confidence intervals, tolerance intervals, significance tests, and fiducial probabilities. To take

their place he advocated a thoroughgoing Bayesianism in which all question of statistical reasoning boil down to the choice of a prior personal probability and the use of Bayes's rule to alter personal probabilities in light of evidence.

Savage made many contributions to the development of Bayesian statistics, of which the most significant are these: He proved a "washing-out" theorem that shows how, under fairly unrestrictive conditions, Bayesian agents with diverse prior probabilities will eventually converge to the same posterior given a sufficiently long run of shared observations. In a highly influential paper, written with Ward Edwards and Harold Lindeman (1963), he established the principle of stable estimation, which specifics conditions under which the value of a posterior probability will be independent of its prior. In one of his last papers, he developed an elegant general method for eliciting personal probabilities using proper scoring rules (1971). Savage died November 1, 1971, in New Haven, Connecticut, after having made lasting and seminal contributions to statistics, decision theory, psychology, and economics.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Decision Theory; Statistics, Foundations of.

Bibliography

de Finetti, Bruno. "La prévision: Ses lois logiques, ses sources subjectives." *Annals de l'Institut Henri Poincaré* 7 (1937): 1–68. Translated as "Foresight: Its Logical Laws, Its Subjective Sources." In *Studies in Subjective Probability*, eds. Henry Kyburg and Howard Smokler. New York: John Wiley, 1964.

Edwards, Ward, Harold Lindeman, and Leonard J. Savage. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70 (1963): 193–242.

Lindley, Dennis. "L. J. Savage—His Work in Probability and Statistics." *The Annals of Statistics* 8 (1980): 1–24.

Joyce, James M. *The Foundations of Causal Decision Theory*. Cambridge, U.K.: Cambridge University Press. 1999.

Ramsey, Frank P. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*, edited by Richard Braithwaite. London: Kegan Paul, 1931.

WORKS BY SAVAGE

The Foundations of Statistics. New York: Wiley, 1954. 2nd revised edition, New York: Dover, 1972.

"The Elicitation of Personal Probabilities and Expectations." *Journal of the American Statistical Association* 66 (1971): 783–801.

James M. Joyce (2005)

SAVIGNY, FRIEDRICH KARL VON

(1779–1861)

Friedrich Karl von Savigny, the founder of historical jurisprudence, was born in Frankfurt, Germany, into a family that had moved there from Lorraine. Left an orphan at thirteen, Savigny was brought up by a friend who educated him in ways that recall the experience of young John Stuart Mill. At seventeen Savigny entered the University of Marburg; after studying at other universities, he returned to Marburg for his doctor's degree in 1800 and began a long, influential, and distinguished teaching career. At the age of twenty-four he published *Das Recht des Besitzes* (The Right of Possession; Giessen, 1804), and in the following year he began to tour libraries in search of manuscripts for his historical work. In 1810 he accepted a teaching post at the newly founded University of Berlin, which he helped organize and where he became rector. He did much to raise the standards of German universities and to help them achieve a dominant position in the world of scholarship. While teaching, writing, and assisting in the administration of the university until 1842, he also performed judicial tasks, and from 1842 to 1848 he was chancellor of Prussia.

In his stress on continuity and tradition Savigny may have been influenced by Edmund Burke, and in his understanding of the methods and aims of historical research he may have been influenced by Barthold Georg Niebuhr, who also took part in the founding of the University of Berlin and was an admirer of Roman institutions.

Savigny's two *magna opera* were the *Geschichte des römischen Rechts in Mittelalter* (7 vols., Heidelberg, 1815–1834) and the *System des heutigen römischen Rechts* (8 vols., Berlin, 1840–1849). In 1850 his miscellaneous writings, *Vermischte Schriften,* were published at Berlin in five volumes, and in 1851 and 1853 his two-volume work *Das Obligationenrecht als heute römischen Rechts* was published. He was cofounder, in 1815, of the *Zeitschrift für geschichtlichen Rechtswissenschaft.* His massive work on Roman law in the Middle Ages became the source of subjects for countless historical monographs. His students, and their students in turn, dominated historical and legal scholarship and teaching for several generations, and he was universally acknowledged as one of the most influential thinkers and scholars of the nineteenth century.

The main thrust of Savigny's jurisprudential thought, however, is not found in his monumental his-

torical treatises but in a polemical tract published at Tübingen in 1814, *Vom Beruf unserer Zeit für Gesetzgebung und Rechtswissenschaft.* This pamphlet was in rebuttal to A. F. J. Thibaut's *Civilistische Abhandlungen* (Heidelberg, 1814), in which a plan for a single code of laws for all German states was urged.

Savigny argued that law has no abstract origin in nature or mind but is organically connected with the people of a nation and is an expression of its *Volksgeist,* or collective genius. Fundamentally, law is formed by custom and popular faith, "by internal, silently operating powers, not by the arbitrary will of a lawgiver." The "real law" is always "the proper will of the people." Like language and manners, law has movement and development; it grows with a people and dies with it.

In earliest historical times, Savigny claimed, law was no more separable from a people than was its language or its manners. Rights and duties were created and extinguished by symbolic acts, which were the "true grammar" of law in this period. As social existence became more complex and sophisticated, law came to be expressed in abstract forms; jurists became a professional class, and law perfected its language and took a scientific direction. Instead of existing in the consciousness of the people, it now existed in the consciousness of the jurists, who became the representatives of the community, the voice of its *Volksgeist.* Now the law had a twofold existence: the "political" element, or the connection of the law with the general existence of the people, and the "technical" element, or the abstract and scientific existence of the law. From this it follows that the jurist needs a twofold spirit: the historical sense, with which to seize "the peculiarities of every age and every form of law," and the systematic sense, with which to see "every notion and every rule in lively connection and cooperation with the whole" legal order. Through these senses the jurist will acquire mastery over a body of law, obtain for that law a thorough grounding in history, and discover its organic principle. He will be able to separate that which still has life from that which is lifeless "and only belongs to history," and in this way he will arrive at a truly national law—a "living customary law."

Savigny's views contributed in varying degrees to a number of significant results: (1) They helped bring to an end the dominant natural law philosophy that looked to pure reason as the source of law. (2) They delayed the movement for codified legal systems that had started with the Napoleonic codes. (3) They established the historical school of jurisprudence. (4) They laid the basis for the sociological school of legal thought. (5) They retarded the development and acceptance of legislation as a source of law. (6) They contributed to an exaggerated stress on nationalism and to a disparagement of the idea of a common law of humankind as an expression of *Menschengeist.* Perhaps Savigny's most enduring influence is to be found in his idea that law must not be isolated into an autonomous science but must be treated as an aspect of social life, development, and order—as a social, historically conditioned phenomenon.

See also Burke, Edmund; Historical School of Jurisprudence; Mill, John Stuart; Philosophy of Law, History of; Philosophy of Law, Problems of.

Bibliography

PRIMARY WORKS

Of the Vocation of Our Age for Legislation and Jurisprudence. New York: Arno Press, 1975.

System of the Modern Roman Law. Translated by William Holloway. Westport, CT: Hyperion Press, 1979.

SECONDARY SOURCES

Kantorowicz, Hermann. "Savigny and the Historical School of Law." *Law Quarterly Review* 53 (1937): 326ff.

Montmorency, J. E. G. "Friedrich Carl von Savigny." In *Great Jurists of the World,* edited by J. Macdonell and E. Manson. Boston: Little Brown, 1914.

Stoll, Adolf. *Friedrich Karl von Savigny,* 3 vols. Berlin, 1927–1939.

Wolf, Erik. *Grosse Rechtsdenker der deutschen Geistesgeschichte.* Tübingen, 1951; 4th ed., 1963. Ch. 12.

Zwilgmeyer, Franz. *Die Rechtslehre Savignys.* Leipzig, 1929. Bibliography on pp. ix–xii.

Milton R. Konvitz (1967)
Bibliography updated by Philip Reed (2005)

SCEPTICISM

See *Skepticism, History of*

SCHELER, MAX

(1874–1928)

A pioneering German phenomenologist, ethicist, and social philosopher, Max Scheler was born in Munich in 1874. His father was Lutheran, his mother was Jewish; Scheler himself, ever independent, embraced Catholicism. After studying with Wilhelm Dilthey and Georg Simmel, he earned his doctorate in 1897 under Rudolf Eucken in Jena, where he taught until 1906. From 1907 he taught in Munich, where he met Franz Brentano and sev-

eral disciples of Edmund Husserl, the father of the phenomenological movement. He soon became acquainted with a growing circle of phenomenologists from Munich and Göttingen, including Moritz Geiger (1880–1937), Dietrich von Hildebrand (1889–1977), Alexander Pfänder, Adolf Reinach (1883–1917), Edith Stein, and others. But as early as 1901, when he first met Husserl, Scheler had already taken an independent phenomenological direction of his own.

In 1910 Scheler lost his post in Munich after a divorce alienated him from the Catholic university administration. In 1912, he married Märit Furtwängler, sister of the noted conductor. From 1910 to 1919, he freelanced as an independent scholar, publishing a prolific number of works, particularly on ethics, but also on political issues of the day, including war, capitalism, feminism, the psychology of resentment, and various social issues. He served on diplomatic missions to Switzerland and the Netherlands. After World War I, he actively promoted the causes of international reconciliation, moral renewal, pacifism, and European reunification based on ideals of Christian socialism. It was not until 1919 that Scheler received a full professorship, in Cologne, where his focus turned to religion, anthropology, metaphysics, and sociology of knowledge. By 1922 he had fallen away from Catholicism in favor of a pantheistic conception of divine self-realization in history. He died on the eve of assuming his final post in Frankfurt in 1928, after repeatedly warning against the rise of German Nazism and Italian Fascism. His writings were suppressed by the Nazis in Germany from 1933 to 1945.

Scheler's impact on the phenomenological movement was considerable, despite ambivalent relationships with Husserl and Martin Heidegger. Many prominent thinkers have acknowledged their debt to him, including Jean-Paul Sartre, Maurice Merleau-Ponty, Gabriel Marcel, Nicolai Hartmann, Roman Ingarden, Hans-Georg Gadamer, Martin Buber, and José Ortega y Gasset. Pope John Paul II wrote a doctoral dissertation on him. Scholars in the Spanish-speaking world, Japan, and Russia were well acquainted with Scheler long before he was known in the English-speaking world.

Scheler's most important phenomenological works were published during his prolific middle period. These include *Formalism in Ethics and Non-Formal Ethics of Values* (1913–1916), his seminal critique of Immanuel Kant's ethics and outline of his own phenomenological ethics based on a theory of values. His 1916 essay *Ordo Amoris* develops his Pascalian conception of a faculty of cognitive feeling independent of reason, which apprehends a hierarchical array of values in its pure incontrovertible immediacy. Between 1912 and 1913 he also published phenomenological studies of sympathy, love, and hate in *The Nature of Sympathy*, and a study of resentment and impotence in modern bourgeois morality in *Ressentiment* (1994) [1964])—a brilliant transmutation of Friedrich Nietzsche's claim that Judeo-Christian morality stems from resentment, eliciting Ernst Troeltsch's famous characterization of Scheler as "the Catholic Nietzsche."

While initially collaborating with Husserl, Scheler criticized Husserl's "Cartesianism" and for giving inordinate primacy to reason. By contrast, Scheler insisted on the primacy of feeling and its independence from reason in apprehending values, which he considered the primordial phenomena of consciousness. Scheler did not use Husserl's terms *noesis* and *noema* to distinguish the act of thinking from the object of thought, yet he recognized that this polarity within consciousness, first investigated by Brentano, allows for two approaches in investigation. Thus he distinguished *act-phenomenology* from *phenomenology of facts*, the former focusing on persons as the source of the unifying intention animating acts, the latter analyzing three types of facts—natural, scientific, and phenomenological.

The preeminent phenomenological facts overlooked by Kant, according to Scheler, are values. Kant rightly denies that moral obligation can be defined by reference to empirical objects of desire without subordinating it to the relativizing contingencies of particular whims, ends, and purposes. But he fails to discern the distinctive nature of values as pure qualities or essences, distinct from empirical entities or objects of desire that might serve as their bearers. Just as colors can be conceived independently of any colored surfaces or bearers, values can be intuited as pure, independent essences. Furthermore, values exhibit an objective hierarchical ranking, furnishing a material basis for ethics, in contrast to Kant's empty formalism. Accordingly, Scheler distinguishes four basic ranks of values. From highest to lowest, these include the (1) *religious*, such as the sacred and profane; (2) *cultural*, such as the true, right, and beautiful; (3) *vital*, such as the noble and common; and (4) *sensory*, such as the pleasant and painful. Scheler's criteria for this classification are reminiscent of Jeremy Bentham's hedonic calculus, including relative duration, depth of satisfaction, and so forth. He also held that this ranking reflects an a priori "logic of preference."

As in teleological theories generally, Scheler defines moral values in terms of the nonmoral value realized or intended through an act. Accordingly, moral good is

achieved as a by-product of realizing or intending a positive or comparatively higher nonmoral value, such as sacrificing the lower value of physical comfort for the higher value of one's children's education. His ethic, unlike Kant's, is based not on "blind duty," but on positive insight into the nature of values.

Scheler is unabashedly objectivist and absolutist in his value theory, but acknowledges the relativity of actual value judgments among societies and individuals. Someone suffering a pathological urge to sacrifice does not have the same obligation to be selfless as the self-centered egoist. Differences of cultural ethos are also significant. Recognition of such relativities inform Scheler's theories of virtue, conscience, and obligation, as well as his concepts of types of *exemplary acts* and *exemplary persons*—such as saints, geniuses, and heroes—that he proposes as vehicles for moral education. Yet he steadfastly maintains that such relativities do not undermine the absolute objectivity of values themselves.

See also Phenomenology.

Bibliography

WORKS BY SCHELER

Ressentiment, edited with introduction by Lewis A. Coser. Translated by William W. Holdheim. New York: Free Press of Glencoe, 1961; Milwaukee, WI: Marquette University Press, 1994.

Formalism in Ethics and Non-Formal Ethics of Values. Translated by Manfred S. Frings and Roger L. Funk. Evanston, IL: Northwestern University Press, 1973.

Max Scheler on Feeling, Knowing, and Valuing, edited by Harold J. Bershady. Chicago: University of Chicago Press, 1992. Contains a listing of English translations.

SECONDARY WORKS

Blosser, Philip. *Scheler's Critique of Kant's Ethics*. Athens: Ohio University Press, 1995.

Frings, Manfred S. *The Mind of Max Scheler: The First Comprehensive Guide Based on the Complete Works*. Milwaukee, WI: Marquette University Press, 1997.

Kelly, Eugene. *Structure and Direction: Studies in the Phenomenological Philosophy of Max Scheler*. Boston: Kluwer, 1977.

Nota, John H. *Max Scheler: The Man and His Work*. Translated by Theodore Plantinga and John H. Nota. Chicago: Franciscan Herald Press, 1983.

Spader, Peter H. *Scheler's Ethical Personalism: Its Logic, Development, and Promise*. New York: Fordham University Press, 2002.

Philip Blosser (2005)

SCHELLING, FRIEDRICH WILHELM JOSEPH VON

(1775–1854)

Friedrich Wilhelm Joseph von Schelling, the German idealist philosopher, was born at Leonberg in Württemberg, the son of a learned Lutheran pastor, Joseph Friedrich Schelling. From his earliest years, he was destined by his family for the ministry. He was educated at the cloister school of Bebenhausen and, from 1790 to 1792, at the theological seminary at Tübingen. There he became friendly with two older students who were to play significant roles in his own life, as well as in cultural history: G. W. F. Hegel and J. C. F. Hölderlin, the great romantic poet. The three young men were keen partisans of the French Revolution, and they also enthusiastically discussed the ideas of the philosophers, especially Benedict de Spinoza, Immanuel Kant, and Johann Gottlieb Fichte.

For several years Schelling held a position as tutor of the sons of a noble family. Then, in 1798, at the unusually young age of twenty-three, he was called to a professorship at Jena. There the famous Fichte, the leading philosopher in Germany at the time and the idol of Schelling's youth, became his colleague and friend. In 1802 and 1803 Schelling and Hegel jointly edited the *Kritisches Journal der Philosophie*. At that time, though Hegel was five years older than Schelling, he was generally considered to be Schelling's disciple, and Hegel's first book was a comparison of Fichte's and Schelling's philosophies.

In nearby Weimar, Johann Wolfgang von Goethe and Friedrich Schiller were at the peak of their careers. Schelling met them both and became friendly with Goethe. Jena was now the center of German romanticism, and the ideas and personalities of this movement made a profound and lasting impression on Schelling. The romantic movement was, of course, also influenced by his philosophy. In its stress on the importance of the individual and the supreme value of art, and in its antirationalism, organicism, and vitalism, Schelling's transcendental idealism is the epitome of German romantic philosophies.

His friends among the romantics included Ludwig Tieck, who interested Schelling in folklore and mythology; the brilliant young poet Novalis; and August and Friedrich von Schlegel, whose translations of William Shakespeare made the English playwright one of the main shaping forces of German literature. Schelling was particularly intimate with August and his charming, intellectually gifted wife Caroline. Soon he became infor-

mally engaged to Auguste Böhmer, the sixteen-year-old daughter of Caroline by a previous marriage, but she died in 1800 before they could marry. It was rumored at the time that Schelling's amateur medical attentions contributed to her death. Certainly he was impetuous and self-confident to a point that some felt bordered on irresponsibility. This was a personal pattern common among the romantics, who sometimes defended themselves with the words of Schelling, "The beginning and end of all philosophy is—freedom."

In 1803 Caroline divorced August Schlegel and married Schelling. In keeping with the romantic creed, the three remained friends. It seems to have been an ideal marriage in every way. Schelling produced his most successful works during these years, and when Caroline died in 1809 he was grief-stricken; from then on he seemed unable to put his ideas together in a way that satisfied him. He never published another book as long as he lived, though he continued to write and lecture for many years. In 1812 he married Pauline Gotter, a friend of Caroline's.

From 1803 to 1806 Schelling taught philosophy at the new University of Würzburg, and in 1806 he was called to Munich as an associate of the Academy of Sciences and as secretary of the Academy of Arts. He later became secretary of the philosophical section of the Academy of Sciences. These positions were government sinecures that afforded him abundant leisure and also allowed him to lecture at Stuttgart and, from 1820 to 1827, at Erlangen. In 1827 he became a professor at Munich. In 1841 the Prussian authorities, in the hope that he would serve as a counterbalance to the powerful influence of the radical Young Hegelians, appointed him to the position of Prussian privy councilor and member of the Berlin Academy, and he lectured for the next five years at the University of Berlin. He died at the age of seventy-nine at Bad Ragaz, Switzerland.

Of all the major German philosophers, Schelling is the least known in the English-speaking world. His name is familiar as the historic link connecting Kant and Fichte with Hegel, but this description fits only his earlier work. Through his personal association with some of the German romantic writers and his doctrinal influence on the entire German romantic school, as well as through the direct influence of his aesthetics on Samuel Taylor Coleridge and, through Coleridge, his indirect influence on other English poets of the period, he is also known as the philosopher of romanticism. In his last phase, which was partly a conscious reaction to Hegel, he anticipated some of the central ideas of the existentialists, and for this reason there has been a revival of interest in his later writings.

The development of Schelling's philosophy can be conveniently divided into four stages—subjective idealism, the philosophy of nature, the philosophy of identity, and the philosophy of the opposition of the negative and the positive. The stages are logically connected with one another, but also are clearly separate, so much so that their author was often accused of inconsistency. For example, Hegel wrote, "Schelling carried on his philosophic education before the public and signaled each fresh stage of his advance with a new treatise."

SUBJECTIVE IDEALISM

In the first stage Schelling was gradually working himself free from Fichte's subjective idealism to an independent position of his own. The major works of this phase were *Vom Ich als Prinzip der Philosophie, oder über das Unbedingte im menschlichen Wissen* (Tübingen, 1795), in which he posited the ego as the supreme, unconditioned element in human knowledge, and *Philosophische Briefe über Dogmatismus und Kritizismus* (in *Philosophisches Journal*, 1796), in which he compared Spinoza and Fichte. There is little that is original in these works other than the style and the tone. However, Schelling's style is important because its eloquence, its sense of emotional urgency, and its relative freedom from technical jargon—a rare trait in the writings of German idealists—all point to his affinity with the romantic movement and his unique philosophic stress on the importance of aesthetics.

PHILOSOPHY OF NATURE

The second stage, the philosophy of nature, was the most famous and the most influential of Schelling's philosophies and remained so until recent years. The first important work of this stage was *Ideen zu einer Philosophie der Natur* (Leipzig, 1797). Against Fichte's conception of the world as the construction of the ego, Schelling now insisted that the world of nature is just as real and just as important as the world of the ego. In fact, it is nature, the objective, that gives to consciousness what consciousness reproduces anew. Originally, consciousness and nature are one and infinite; but consciousness limits itself and presents itself to itself as finite, as different from nature. The essence of the ego is spirit, and the essence of nature is matter, but the essence of matter is force; that is, attraction and repulsion. In force, Schelling finds the common ground of nature and ego. As attraction it is objective, it is nature, it is matter; as repulsion it is subjective, it is ego, it is spirit. This duality also governs human perception:

As attraction to the self, force governs the streaming of the outer world into the inner world of sensation, and this internal experience of movement constitutes the a priori basis of time; as repulsion, pushing out into the world, force constitutes the a priori basis of space.

PHYSICAL SCIENCES. In *Von der Weltseele* (Hamburg, 1798) Schelling dealt with the philosophic problems of the physical sciences. He believed that the fundamental aim of the sciences was the interpretation of nature as a unity, and therefore the proper study of all sciences was force. He tried to show that mechanical, chemical, electrical, and vital forces were all different manifestations of the same underlying force. In the following year, in *Erster Entwurf eines Systems der Naturphilosophie* (Jena and Leipzig, 1799) and in *Einleitung zu dem Entwurf eines Systems der Naturphilosophie oder über den Begriff der spekulativen Physik* (Jena and Leipzig, 1799), he depicted this force as "pure activity." He saw nature as an infinite self-activity, realizing itself in finite matter but forever unexhausted, forever short of completely realizing itself. He felt that he had thus found a parallel in the physical universe for Kant's idea of the moral universe as practical reason forever striving toward an unattainable ideal. He further developed this phase of his thought in "Allgemeine Deduktion des dynamischen Prozesses" (in *Zeitschrift für spekulative Physik*, Vol. 1, 1800); *Über den wahren Begriff der Naturphilosophie; Darstellung meines Systems der Philosophie* (Jena and Leipzig, 1801); and *Bruno, oder über das göttliche und natürliche Prinzip der Dinge* (Berlin, 1802).

KNOWLEDGE. In the *System des transzendentalen Idealismus* (Tübingen, 1800), his most systematic and mature statement, Schelling applied to the philosophy of nature the insights gained from the Kantian and Fichtean philosophy of knowledge. His technique for deriving the world of objects from the world of the ego was to turn consciousness upon itself as the only object of which we have immediate firsthand knowledge. Thus, he found that when we abstract from all objects of knowledge, both within ourselves and in the outside world, we arrive at the pure activity of abstracting, which is pure self-activity. Seen in this light, the consciousness of the not-self is the limit of self-activity, just as the things-in-themselves are at the limits of knowledge in *The Critique of Pure Reason*.

On this foundation, Schelling built a theory of three stages of knowledge, which he described as progressing from sensation to perception, from perception to reflection, and from reflection to will. At first, consciousness of a limit, of the not-self, is felt as a sensation. The limit, where the sensation is felt, is the meeting place of self-consciousness pushing outward and the force of the consciousness of external objects streaming inward. Therefore, all sensation is a feeling of myself as limited. Here we become aware of gravity, of the force of the real objective world in space, and also of intensity, which is the immediate consciousness of the self and its own activity in time. From the perception of the outside world comes reflection, and from reflection on the internal world comes will.

In this way Schelling felt that he had established links among Kant's categories, schemata, and objects of perception. Aside from the technical question of the correctness of this linkage—certainly Kant would have disputed it—it has great historical importance, because this is perhaps the only area in which Schelling decisively influenced the fully matured philosophy of Hegel, who used this reasoning to connect the dialectic of thesis-antithesis-synthesis with Kant's triadic formulation, though the dialectic itself was borrowed by Hegel from Fichte.

Schelling argued that the separation of knowledge from its object occurs only in abstraction. In reality, concepts have no existence apart from their objects, since knowledge is the meeting of objects and self. Therefore, the self is not merely one of the objects of knowledge; it is the condition of all knowledge. And since the essence of the self is pure self-activity, knowledge ultimately derives from willing, which is the action of the self.

OTHER MINDS. Schelling now asks two fundamental questions. How do I know there are other intelligences? And how can they act on me? He answers that our consciousness of limitations implies the existence of other selves that act as limiting factors. (Here he takes issue with Kant's teaching that intelligence is limited by something not itself.) But the other selves can act on me only indirectly, through my representation of their acts. Their action does not compel mine, but limits it; and such limitation is compatible with my freedom. It is the community of interacting intelligences that constitutes the historical life of man. And while nature exists when not perceived by me, it exists then only because it is perceived by other human beings. Objectivity is intersubjectivity.

WILL AND IMAGINATION. Although perception is necessary and limited, will is free and unlimited. The imagination and its ideas mediate between perception and will. As opposed to the conceptions of the understanding, which are finite, the ideas of the imagination are both

finite and infinite. An idea's relation to its object is finite, but the activity of the imagination in this relation is infinite. Each idea is subsumed under an ideal, as conceptions are subsumed under their schemata in Kant. The function of the will is to idealize the imagination's ideas. The contradiction thus engendered gives rise to impulse, defined as the desire to restore destroyed identity. Through impulse, there is constant realization of ideals, but the ideas of the imagination are constantly striven after and never attained.

WILL AND KNOWLEDGE. The distinction between will and intelligence thus is relative, not absolute. From a higher point of view, they are identical. In intelligence, the I that acts and the I that knows are one. The acting I is an object for itself, while the knowing I merely perceives other objects. In action there is no transition from the world of nature to the world of mind, for the subject has become an object to itself. Any change in the outer world is received as a perception, but every action causes such a change; therefore action is perception. (Here, as elsewhere, Schelling anticipates Gestalt psychology.) Self-determination is the primary condition of all consciousness.

JUSTICE. The object of impulse, which always acts to restore the lost identity of the self and the world, is happiness. But an impulse that transcends its proper limits acts against itself and must be prevented by a sanction not found in nature—a sanction of the will. This sanction of the will is thus the basis of justice, and the law of justice is a second nature that our will sets above the first nature.

THE NATURE OF HISTORY. The process of history is the gradual realization of law; history can be described as the development of human freedom, as an eternal progress toward the perfect state—a sovereign world federation of all sovereign states—in which all men would be citizens. Thus, history is the realization of freedom through necessity. There is an absolute identity between freedom and necessity, but this identity is forever unconscious, never the object of knowledge but always the object of faith. God is neither personal nor objective, but the revelation of the divine in man. This revelation is never complete. History is a drama in which human beings are not merely the actors, but also the authors.

ART AND AESTHETICS. If history is a drama for Schelling, nature is a work of art. Like Kant in *The Critique of Judgment*, Schelling believed that organisms and works of art are alike in that they can be properly understood only teleologically; that is, as entities in which the parts serve the whole and the whole is itself purposive. The main difference between art and organisms, according to Schelling, is that in organisms the activity of the organizing intelligence lies hidden or unconscious, manifest only in the product—the organism itself; but in the work of art the productive activity is conscious whereas the product, the true art work, is unconscious and infinite. The artist never fully understands his art. The purpose of art is neither utility, nor pleasure, nor morality, nor knowledge, but beauty—the realization of the infinite in the finite.

In his aesthetics, which is elaborated in the *System des transzendentalen Idealismus* and his lectures on the philosophy of art, *Über das Verhältniss der bildenden Künste zu der Natur* (Munich, 1807), Schelling is at his most personal, his most impassioned, his most characteristic, and his most original. He held that in art, intelligence for the first time becomes completely self-conscious. In philosophy, it is abstract and limited in the expression of its potential infinity. But in art, which is completely free from abstraction in this sense, intelligence fully realizes its infinite nature. (It is pertinent that Hans Arp, the abstract artist, has written that the works usually called "abstractions" are more accurately referred to as "concretions.") Thus art is the goal toward which all intelligence moves. Art is the true philosophy, because in it nature and history are forever reconciled; but the artist is not therefore a philosopher, since he often lacks a theoretical understanding of his own creation. The theoretical intelligence merely contemplates the world, and the practical intelligence merely orders it; but the aesthetic intelligence creates the world.

PHILOSOPHY OF IDENTITY

The third stage of Schelling's thought was the philosophy of identity, first expounded at length in *Vorlesungen über die Methode des academischen Studiums* (Tübingen, 1803), appropriately written in Spinoza's geometric mode. Here Schelling said that the philosophy of nature and the philosophy of knowledge, taken together, constitute only half the truth and need to be completed by the other half, which unites nature and knowledge in an undifferentiated identity. The production of reality does not rest on the opposition of intelligence and nature, subject and object, but in the identity of all reality as it rises from the absolute. The absolute identity of nature and intelligence is found in their common neutral source, reason. Reason is one and infinite, embracing things-in-themselves and knowledge of things. In reason there is no

object, no subject, no space, no time. Its supreme law is the law of identity, $A = A$, which is true regardless of all spatial or temporal considerations. In the formula $A = A$, the distinction between subject and object is formal and relative. Subject and object here concern only the form, and are indifferent as to essence. It was this phase of Schelling's thought that Hegel wittily called "the night in which … all cows are black."

PANTHEISM. The philosophy of identity was a kind of pantheism, but it stressed the aliveness of nature in contradistinction to Spinoza's dead, materialistic, deterministic pantheism. Although Spinoza's influence is evident, it is filtered through the vitalistic interpretations of Johann Gottfried Herder and Goethe and tempered by the parallel influence of Giordano Bruno's vitalistic pantheism. Schelling believed that life was the basis of the inorganic world, and not vice versa. Nature is inseparable from God, but distinguishable from him. God is not to be comprehended rationally, because his essence is will and he can be apprehended only through the will, in action. For the most part, Schelling's thought here draws from Jakob Boehme, and reintroduces Protestant mysticism into the mainstream of Western philosophy.

GOD AND EVIL. In *Philosophische Untersuchungen über das Wesen der menschlichen Freiheit* (Landshut, 1809; translated as *Of Human Freedom*, Chicago, 1936) Schelling, like Boehme, distinguishes between God as ground of being and God as perfection. Evil is explained as the ground eliciting the self-will of man in order to awaken him to the distinction between good and evil, which originally were united in one identity. Thus, evil is a necessary stage in the progress toward the total realization of good. Imperfection in being is perfection in the process of becoming. There is a dark ground or negative principle in God, but it exists so that he can become separate from it as a personality.

POSITIVE PHILOSOPHY

After 1809, the year of his first wife's death, Schelling made the given situation of existence his predominant concern. This final existentialist phase of his philosophy was first propounded in *Die Weltalter* (written in 1811 but not published in Schelling's lifetime), consummated in his lectures at the University of Berlin, and saved for posterity in three volumes, *Einleitung in die Mythologie, Philosophie der Mythologie*, and *Philosophie der Offenbarung*, which were published posthumously in the *Sämmtliche Werke*. In these works he sought to erect a positive philosophy based on the evolution of the divine principle in human history, especially in myths and religions, which he felt opposed and thus completed his own earlier, negative, merely rational philosophy. However, rather than representing a sharp break with his past, this last phase can be considered as the flowering of tendencies he showed as early as 1795, when he wrote, "The main function of all philosophy is the solution of the problem of the existence of the world." It is significant that while the prolific and influential writings of his first three periods were crowded into fourteen brief years, from 1795 to 1809, his last period, during which his rate of production slowed and his influence waned, lasted from 1809 to his death in 1854.

GOD. The root of existence is now found in nonbeing, in God as the ungrounded, the abyss, the eternal nothing. Only against the ungrounded can the ground arise, because nothing can become evident without resistance. Thus God is "eternal contrariety," forever alienating himself from himself. This alienation creates the possibility of the fall. As only the Absolute is real, finite things, which are not real, can exist only in a removal, in a fall from reality. The Absolute creates its own counterpart, freedom, which is both the cause of the fall and the last trace of divinity things bear after the fall. Because of this progression through opposites, Schelling called this fourth phase of his thought the opposition of negative and positive philosophy.

MAN. As the creature in whom the fall, and the state of things before the fall, both rise for the first time into consciousness, man is the crown of creation and the most interesting and rewarding object of philosophic attention. Man is free creative activity, the essence of the world. Thus, in his last phase, Schelling was led to a kind of philosophic anthropology, seeking for the essence of man in what he thought was his deepest activity, myth-making and religion. Despite the profoundly mystical flavor of his thought in this period, he still kept contact with his Kantian heritage. In *Philosophie der Mythologie* he explained mythology as a symbolic system of ideas with its own a priori structure as necessary for its functioning as, according to Kant, the a priori structure of the understanding is necessary for logical thought. Ernst Cassirer's neo-Kantian formulation of mythology as just such a conceptual structure owes a great deal to Schelling, a debt fully acknowledged in the second volume of *The Philosophy of Symbolic Forms*.

RESEMBLANCE TO EXISTENTIALISM. What has made this last phase of Schelling's thought most apposite to

modern existential philosophy is another question rising from his consideration of man's being in the world. As he put it, "Just he, man, impelled me to the final desperate question: Why is there anything at all? Why not nothing?" It is this question, described as "dreadful" rather than "desperate," that Martin Heidegger took for his central theme in *Being and Time*.

Schelling's resemblance to the modern existentialists is suggestive rather than substantive, but the suggestion is inescapable. Like them, he emphasized that philosophy must deal not only with the "what" of the world, which explains its nature, but also with the "that" of the world—the fact of its existence, of its being there. And like Søren Kierkegaard (who attended some of his lectures in Berlin but was not impressed), Friedrich Nietzsche, Heidegger, and Jean-Paul Sartre, Schelling tried to express the inexpressible pathos of existence in oracular utterances halfway between poetry and metaphysics, the quality of which can be conveyed only by quotation. The world and God have as common ground "the incomprehensible basis of reality." "Existence is self-affirmation." God is "the infinite affirmation of himself." The objective world is the unconscious poetry of the spirit creating itself. Finally, there is a striking formulation of the existential anxiety, which is also an anticipation of the psychoanalytic doctrine of resistance: "The philosopher who knows his calling is the physician who … seeks to heal with gentle, slow hand the deep wounds of human consciousness. The restoration is all the more difficult since most people do not *want* to be healed at all and, like unhappy patients, raise an unruly outcry if one even approaches their wounds."

So the problems posed by Schelling in the nineteenth century are still very much alive in the philosophic and literary world of today. At that time his main influence in England was in aesthetics, and his lectures on the philosophy of art were translated as *The Philosophy of Art* in 1845. The continuing, perhaps growing contemporary interest in him is demonstrated by the fact that the first translations into English of any of his books since then—significantly, both from his last, existentialist phase—were published in America in 1936 and 1942.

Bibliography

WORKS BY SCHELLING

Grundlegung der positiven Philosophie. Münchener Vorlesungen WS 1832/33 und SS 1833, edited by H. Fuhrmans. Turin: Bottega d'Erasmo, 1972.

Einleitung in die Philosophie, edited by W. E. Ehrhardt. Stuttgart: Fromman-Holzboog, 1989.

System der Weltalter, edited by S. Peetz. Frankfurt: Klostermann, 1990.

Urfassung der Philosophie der Offenbarung, edited by W. E. Ehrhardt. Hamburg: Meiner, 1992.

Collected Works

Sämmtliche Werke. 14 vols, edited by K. F. A. von Schelling. Stuttgart and Augsburg: Cotta, 1856–1861.

Werke. 8 vols, edited by M. Schröder. Munich, 1927–1956.

Translations

The Ages of the World. Translated by F. de Wolfe Bolman. New York: Columbia University Press, 1942. Translation of *Über die Gottheiten von Samothrake*, Stuttgart and Tübingen, 1815, a part of *Die Weltalter*, which is mentioned in the text. *The Ages of the World* includes a long, informative introduction on Schelling's philosophy, especially its last phase.

Of Human Freedom. Translated by J. Gutmann. Chicago: Open Court, 1936. Translation of *Philosophische Untersuchungen über das Wesen der menschlichen Freiheít*.

The Philosophy of Art: An Oration on the Relation between the Plastic Arts and Nature. Translated by A. Johnson. London: Chapman, 1845. Translation of *Über das Verhältniss der bildenden Künste zu der Natur*.

System of Transcendental Idealism. Translated by P. Heath. Charlottesville: University Press of Virginia, 1978.

The Unconditional in Human Knowledge: Four Essays, 1794–1796. Translated by Fritz Martl. Lewisburg, PA: Bucknell University Press, 1980.

Bruno, or On the Natural and the Divine Principles of Things. Translated by Michale Vater. Albany: State University of New York Press, 1984.

Ideas for a Philosophy of Nature: An Introduction to the Study of This Science. Translated by E. E. Harris and Peter Heath. Cambridge, U.K.: Cambridge University Press, 1988.

On the History of Modern Philosophy. Translated by Andrew Bowie. Cambridge, U.K.: Cambridge University Press, 1994.

WORKS ON SCHELLING

Beach, E. *The Potencies of God(s): Schelling's Philosophy of Mythology*. Albany: State University of New York Press, 1994.

Bowie, A. *Schelling and Modern European Philosophy*. London: Routledge, 1993.

Bréhier, Émile. *Schelling*. Paris: Alcan, 1912.

Cassirer, Ernst. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*, Vol. III: *Die Nachkantischen Systeme*. Berlin: Cassirer, 1920. Includes a comprehensive, clear, and sympathetic account of Schelling's epistemology.

Copleston, Frederick C. *History of Philosophy*. Vol. VII. London, 1963. Includes the best account of Schelling's philosophy in English.

Courtine, J.-F. *Extase de la raison: Essais zur Schelling*. Paris: Galilée, 1990.

Esposito, J. *Schelling's Idealism and Philosophy of Nature*. Lewisburg, PA: Bucknell University Press, 1977.

Fischer, Kuno. *Geschiehte der neueren Philosophie*. Vol. VII. 3rd ed. Heidelberg, 1902. Probably the best full-length study of Schelling's life and work.

Frank, M. *Der Unendliche Mangel am Sein: Schellings Hegelkritik und die Anfänge der Marxschen Dialektik*. Frankfurt: Suhrkamp, 1975.

Hartmann, Nicolai. *Die Philosophie des deutschen Idealismus*. 2nd ed. Berlin, 1960. An exhaustive and authoritative work on German idealism.
Heidegger, M. *Schelling's Treatise on the Essence of Human Freedom*. Athens: Ohio University Press, 1985.
Hirsch, Eric D. *Wordsworth and Schelling*. New York, 1962. Reviews Schelling's aesthetics and its parallels in William Wordsworth.
Jaspers, Karl. *Schelling: Grösse und Verhängnis*. Munich: Piper, 1955. A profound and searching analysis from an existentialist point of view.
Kaspar, W. *Das Absolute in der Geschichte*. Mainz: Grünewald, 1965.
Kroner, Richard. *Von Kant bis Hegel*. 2 vols. Tübingen: Mohr, 1921–1924.
Marx, W. *The Philosophy of F. W. J. Schelling: History, System, and Freedom*. Bloomington: Indiana University Press, 1984.
Noack, Ludwig. *Schelling und die Philosophie der Romantik*. Berlin: Mittler, 1859.
Read, Herbert. *The True Voice of Feeling*. London: Faber and Faber, 1953. Concerned with Schelling's influence on Coleridge. Includes translation of a chapter on philosophy of art by Schelling.
Schneeberger, G. *Schelling: Eine Bibliographie*. Bern, 1954. An extensive bibliography.
Schulz, W. *Die Vollendung des deutschen Idealismus in der Spätphilosophie Schellings*. Pfüllingen: Neske, 1975.
Snow, D. E. *Schelling and the End of Idealism*. Albany: State University of New York Press, 1996.
Tillich, P. *The Construction of Historical Religion in Schelling's Positive Philosophy*. Lewisburg, PA: Bucknell University Press, 1974.
Tilliette, X. *L'Absolu et la philosophie: Essais sur Schelling*. Paris: Presses Universitaires de France, 1987.
Watson, John. *Schelling's Transcendental Idealism*. Chicago: Griggs, 1882. A clear, careful, scholarly paraphrase of Schelling's *System des transcendentalen Idealismus*.
White, A. *Schelling: An Introduction to the System of Freedom*. New Haven, CT: Yale University Press, 1983.

Adam Margoshes (1967)
Bibliography updated by Thomas Nenon (2005)

SCHILLER, FERDINAND CANNING SCOTT
(1864–1937)

Ferdinand Canning Scott Schiller, the British pragmatist philosopher, was born in Schleswig-Holstein and studied at Rugby and at Balliol College, Oxford. After teaching German at Eton, he returned to Oxford for his MA. In 1893 he went to Cornell University as an instructor and graduate student. In 1897, without receiving a doctorate, he returned to Corpus Christi College, Oxford, where he was successively assistant tutor, tutor, senior tutor, and fellow and where he received a DSc in 1906. He served as treasurer of the Mind Association and president of the Aristotelian Society (1921), and he was elected a fellow of the British Academy in 1926. From 1926 on, Schiller spent part of each year at the University of Southern California as visiting lecturer and then as professor; in 1935 he moved there permanently.

PRAGMATISM

Schiller's views, which he called at various times humanism, voluntarism, and personalism, as well as pragmatism, were strongly influenced by William James; and Schiller paid James great tribute, although he claimed to have arrived at his opinions independently. There was, however, an important difference of emphasis between them: James stressed the purposive aspect of thinking, and Schiller, the personal. James also accepted the independence of what is objectively given, whereas Schiller regarded all knowledge, even of "facts," as relatively subjective. Both Schiller and John Dewey were strongly influenced by G. W. F. Hegel and took the process of knowing as central to reality, but the influence of idealism was much stronger on Schiller than on Dewey. And whereas Schiller pursued the subjective and individual aspects of James's psychology, Dewey built upon its objective and social aspects. C. S. Peirce thought that Schiller's philosophy was intermediate between James's and his own.

Schiller's views may best be understood in terms of his opposition to the dominant absolute idealism of the British Hegelians, F. H. Bradley (Schiller's particular bête noire), J. M. E. McTaggart, Bernard Bosanquet, and T. H. Green. To Schiller the absolutism, monism, authoritarianism, rationalism, and intellectualism that these thinkers espoused ignored the basic insight of Protagoras that man is the measure of all things.

Schiller was convinced that all acts and all thoughts are irreducibly the products of individual human beings and therefore inescapably associated with the needs, desires, and purposes of humans. Such terms as *reality* and *truth* denote nothing complete and absolute; rather, they are intertwined with human intentions and deeds. Schiller emphasized the effective creativity of the human mind in organizing the universe of human experience and thus in making or remaking "reality." Man makes his truth along with his other values, Beauty and Goodness. Our axioms are never God-given but are human-made; they are not a priori verities but postulates, or working hypotheses, whose truth grows or diminishes within our experience. The logic we employ in gathering knowledge is dynamic and functional rather than eternally fixed. Our data are not "the given" but "the taken." Thus, in

Schiller's view, human activity is focal both to epistemology and to metaphysics, and there is genuine novelty in our growing universe and no theoretical limit to human freedom.

MAKING REALITY

The absolute idealists maintained that reality is a seamless logical unity, not a mere disjointed plurality; that in the Absolute all separateness vanishes; that nothing finite, nothing that changes, is ever quite real, not even human personality; and that there is something makeshift, transitory, and unsatisfactory about the bits of matter we see, the individual acts we perform, and the private thoughts we think. But, Schiller pointed out, that is all that exists for us. An independent or absolute reality that does not enter into our experience, or explain our knowledge, is irrelevant to us. "Reality" for us is piecemeal, incomplete, and plastic. It is idle to ask "What is real?" Rather, the only question we can answer is "What can I know as real?"

The reality revealed by our actual active procedures of knowing is not rigid but malleable, not completed but evolving. Because it responds, at least to some extent, to our working and probing, it must somehow be not unrelated to our needs and purposes. The process of knowing, Schiller said, is "*never* one of bringing the mind into relation with a fundamentally alien reality, but always one of improving and extending an already existing system which we know." What we call real is that which, for our own reasons, we evaluate as important. It is the result of the kind of selection by which we reduce the chaos about us to order.

Schiller's critics found intolerable the thesis that we make reality. Bertrand Russell, for example, wrote, "Dr. Schiller says that the external world was first discovered by a low marine animal he calls 'Grumps,' who swallowed a bit of rock that disagreed with him, and argued that he would not have given himself such a pain, and therefore there must be an external world. One is tempted to think that … many people … had not yet made the disagreeable experience which Grumps made. Meanwhile, whatever accusations pragmatists may bring, I shall continue to protest that it was not I who made the world" ("Professor Dewey's 'Essays in Experimental Logic,'" *Journal of Philosophy* 16 [January 1919] 26).

Schiller found it hard to meet two particular objections to the theory of the making of reality: The world obviously preceded the existence of humans, and there are patent limits to human powers. In his later writings Schiller therefore reluctantly accepted the distinction between "finding" and "making" the real, although he reiterated the meaninglessness of the "real-as-it-is-in-itself." He revived the Greek term *hule* to refer to the indeterminate, formless chaos, to whatever may be beyond man's ability to perceive or manipulate, to the raw malleable material of the cosmos.

Despite its drawbacks, the doctrine of the making of reality provided Schiller with the basis for certain important conclusions. In his view, it provided a perfect accommodation for Darwinian evolution; it supported a belief in the existence of genuinely new things and situations (always a problem for the absolute idealists because they regarded reality as a self-contained whole); it legitimized human progress; it provided a suitable conceptual scheme for the view, which Schiller ascribed to Albert Einstein and other scientists, that to posit "the real" independently of our sensations is to make an intellectual construction; and, most significantly, it was a firm foundation for man's freedom.

OTHER METAPHYSICAL VIEWS

Schiller's other metaphysical views may be briefly stated. The function of philosophy, he thought, was to preserve the grand synoptic vision, to be an ultimate synthesis of the special sciences. Metaphysical systems, he held, are quasi ethical, or even aesthetic, in character; they reflect personality and temperament. Because the individual human person was an ontological ultimate for Schiller, he was a personalistic pluralist. He was also a hylozoist, asserting that all matter is more or less alive.

TRUTH

Many theories of truth have been propounded through the centuries, but none has been entirely satisfactory. Schiller pointed out the shortcomings of some, particularly the correspondence and coherence theories. Pragmatists agree that no statement wears its truth like a badge; its truth can be determined only by what follows from it in the course of experience. Truth is only a potential, a valuation applied as the result of a procedure called verifying, or making true. Truth is relative to the evidence and to the purpose of the investigator; no degree of verification will ever establish the absolute truth of a statement. Schiller held that truth is personal and particular, dynamic and progressive, not eternal or absolute but the best solution found so far for any problem. That which thwarts or defeats the purpose of an inquiry we call false; that which furthers it we call true. "Truth is that manipulation of [objects] which turns out upon trial to be useful, primarily for any human end, but ultimately for that

perfect harmony of our whole life which forms our final aspiration" (*Humanism*, p. 61).

Nevertheless, Schiller thought the conversion of "The truth is useful" to "The useful is true" to be malicious. Therefore, in Chapter 8 of *Logic for Use* he distinguished seven kinds of truth claims. (1) A postulate is a statement that is "desirable if true," whose truth we try to establish. (2) "A fully verified postulate which serves as principle for a fully established science" and "rests securely on the solid mass of scientific fact it has been instrumental in eliciting" is an axiom. (3) A methodological assumption (determinism, for example) is any guiding principle that appears to be useful in analyzing the flux of events. (4) An assumption of limited usefulness, such as the use of Euclidean geometry in cartography, is a methodological fiction. Finally, truth claims may be, or are, made in (5) fictions, (6) jokes, and (7) lies. Lies are deliberately untrue but may be useful, as in propaganda.

Thus, Schiller held, to claim that all truths work for us in some way and that there is no useless knowledge is far from saying that whatever is useful is true. However, he was aware of difficulties concerning the status of past truth, the usefulness of some parts of pure mathematics, and such questions as whether truth is equivalent to survival value or to social acceptance.

LOGIC

Since the true is what is true for us as seekers for it, Schiller deplored the divorce of logic from the empirical sciences and from psychology. He criticized traditional formal logic for having been a word game and for having been allied to metaphysics rather than to the empirical sciences and to psychology. For Schiller, as for Dewey, thought arises as an element in the solution of a problem. Thus the activity of reasoning has a biological matrix, and it is conditioned by such factors as interest, purpose, emotion, and satisfaction. Schiller was concerned with showing that meanings had been misunderstood and ignored by logic. Meanings, he pointed out, are acquired only in use; they are plastic and personal, and they occur only in contexts. Traditional logic regarded them as purely verbal and as fixed; it believed that one meaning corresponded to one form, and vice versa.

Schiller thought that logic had made the two mistakes of "etherealizing" and "depersonalizing" truth. In its search for formal validity, it had made three fatal abstractions; from actual thinking processes (psychology); from purpose, truth, or utility; and from meaning, matter, and context. In two books, *Formal Logic* (1912) and *Logic for Use* (1929), Schiller made an exhaustive study of formal logic, including terms, propositions, definitions, the syllogism, and fallacies. He showed that, even on its own terms, logic was not free from ambiguity—how can there be novelty in the conclusion of a syllogism? What is the precise import of the copula in a proposition? Moreover, logic appealed at several crucial points to such psychological notions as the "necessity" of implication and the "certainty" or "self-evidence" of propositions. Schiller thought that logic should become a systematic evaluation of actual knowing, a study continuous with the sciences. His resolute experimentalism led him to assert, in "Axioms as Postulates" (1902), that even the laws of thought (identity, contradiction, excluded middle) are not principles of being or rules of logic but postulates.

SCIENTIFIC METHOD

In analyzing the procedures of science, Schiller made several noteworthy contributions. He showed that the concept of "fact" is ambiguous. The "facts" of the scientist are the result of a process of selection, segregation, and evaluation; they are relative to the state of the science, the methods and instruments used, and the aims and bias of the scientist. They are also relative to the hypothesis used, to our own senses, to our memory, and to our words. Schiller also said, "The impossibility of 'breaking' a Law of Nature proves nothing but our determination to uphold a phraseology we have found convenient" (*Formal Logic*, p. 328).

ETHICS AND RELIGION

Schiller carried his pragmatic approach into ethics and religion. There are no abstract values, he said, but only acts of personal valuation. Moral principles are not a priori presuppositions of right conduct; they are its results. The statements of religion are likewise postulates. (James spoke of the will to believe; Schiller, of the right to postulate.) God is a pervasive principle of goodness, not infinite but finite, struggling to develop; the actions of men therefore make a difference. Man's freedom is correlative to the postulate that man is responsible for his acts and is an agent in the full sense of the term. Schiller shared with James and Henri Bergson an interest in psychical research that stemmed from his desire to examine the methods of science at its periphery and from his postulate of immortality. Schiller was also keenly interested in eugenics. This led him to oppose democracy as a "sham" (*Problems of Belief*, p. 81) and to praise the British fascist Oswald Mosley. His social opinions were generally regarded by his philosophic supporters as a vagary.

Schiller was a prolific writer, a sprightly stylist, and a spirited polemicist who maintained a role of philosophic *enfant terrible* through hundreds of essays and books. He edited and wrote most of a parody of *Mind*, which he called *Mind!*—one of the rare examples of philosophic humor.

See also Humanism.

Bibliography

Schiller's major books are two collections of essays, *Humanism: Philosophical Essays* (London and New York: Macmillan, 1903) and *Studies in Humanism* (London and New York: Macmillan, 1907). Two later collections are *Must Philosophers Disagree?* (London: Macmillan, 1934) and *Our Human Truths* (New York: Columbia University Press, 1939).

His major works in logic are *Formal Logic: A Scientific and Social Problem* (London: Macmillan, 1912) and *Logic for Use: An Introduction to the Voluntarist Theory of Knowledge* (London: G. Bell, 1929).

His essay "Axioms as Postulates" appeared in the anthology *Personal Idealism*, edited by Henry Sturt (London: Macmillan, 1902), pp. 47–133. An early, prepragmatic metaphysics, titled *Riddles of the Sphinx* and written under the pseudonym of A. Troglodyte (London and New York, 1891), was offered unsuccessfully to Cornell University as a PhD thesis. The essays "Scientific Discovery and Logical Proof" (1917) and "Hypothesis" (1921) were contributed to *Studies in the History and Method of Science*, edited by Charles J. Singer (Oxford: Clarendon Press, 1917–1921), and may be found in Vol. I, pp. 235–289, and Vol. II, pp. 414–446, respectively. Another work by Schiller is *Problems of Belief* (London: Hodder and Stoughton, 1924).

A critical study of Schiller's philosophy, with an exhaustive bibliography, is Reuben Abel, *The Pragmatic Humanism of F. C. S. Schiller* (New York: King's Crown Press, Columbia University, 1955).

See also Reuben Abel, ed., *Humanistic Pragmatism; The Philosophy of F. C. S. Schiller* (New York: Free Press, 1966); and *A Bibliography of the Works of F. C. S. Schiller, with an Introduction to Pragmatic Humanism* (San Diego, CA: San Diego State College Press, 1969).

Reuben Abel (1967)
Bibliography updated by Michael J. Farmer (2005)

SCHILLER, FRIEDRICH
(1759–1805)

Friedrich Schiller, a famed dramatist, poet, and essayist, was born in Marbach, a small town in southwest Germany, to Elisabeth Kodweiss and Johann Kaspar Schiller, a lieutenant in the army of the Duke of Württemberg. Though tutored in Latin at an early age by his local pastor to prepare him for theological studies, Schiller was mandated by the duke to attend the duke's new military academy, Karlsschule. Schiller later related how his rebellion against the suffocating rigidity and isolation of Karlsschule paradoxically fostered his love of poetry. He remained at the school for eight years, focusing first on law, then on medicine. After his second medical dissertation, "On the Connection of the Animal Nature of Man with his Spiritual Nature," was accepted, he became a regimental physician in Stuttgart. There, he completed his first drama, *The Robbers*, the staging of which a year later (1782) in Mannheim brought him immediate acclaim and confirmation of his literary gifts. When the duke forbade him to write anything but medical treatises, Schiller fled Württemberg. For most of the rest of his life he would suffer considerable financial hardship and extremely poor health. Nevertheless, from 1782 to 1787 he managed to complete three plays (*Fiesco*, *Intrigue and Love*, and *Don Carlos*), to compose several poems (e.g., "Ode to Joy") and essays (e.g., "Theater Considered as a Moral Institution" and "Philosophical Letters"), and to found the journal *Rheinische Thalia*—all of which helped cement his reputation as a member of the *Sturm und Drang* (Storm and Stress) literary movement of the time.

While Schiller's literary output as a critic continued unabated in the ensuing years, his attention over the next decade (1787–1796) turned from the stage to the study of history and to an increasing preoccupation with philosophical treatments of morals and the arts. His *History of the Revolt of the Netherlands* (1787), which celebrated religious tolerance, won him a professorship (albeit unsalaried) in history at the University of Jena in 1789, and over the next two years he produced the enormously successful *History of the Thirty Years War*. His inaugural lecture, "What Does 'Universal History' Mean and to What End Is It Studied?" (1789) contains reflections, fairly conventional at the time, on history's progressive character. This progressive view of history collided, however, with a longing for a lost harmony that he thought art alone can provide (compare his nostalgic elegy of 1788, "The Gods of Greece," with his stirring, forward-looking call to his caste in the 1789 poem "The Artists").

This collision converged with a burgeoning interest in Immanuel Kant's moral and aesthetic writings. Following his marriage to Charlotte von Lengefeld in 1790 and an almost fatal bout with pneumonia a year later, Schiller was given the opportunity to pursue these interests in earnest thanks to a three-year pension provided by Prince Friedrich Christian von Schleswig-Holstein-Augustenburg of Copenhagen. Over the next four years Schiller composed several essays on aesthetics. The organ for

many of these essays was the journal *Die Horen*, founded by Schiller with the help of many of the leading figures in German letters at the time, among them Johann Wolfgang von Goethe and Wilhelm von Humboldt, with whom Schiller developed close friendships that had a lasting influence on his work.

Following this academic and philosophical interlude, and with Goethe's increasing encouragement, Schiller turned his attention back to the theater where he crowned his fame as a playwright with several historical plays: the *Wallenstein* trilogy (1798–1799), *Maria Stuart* (1800), *The Maid of Orleans* (1801), *The Bride of Messina* (1802), and *Wilhelm Tell* (1804).

CRITICAL APPROPRIATION OF KANT'S PHILOSOPHY

Though philosophical concerns are apparent in Schiller's earliest publications, he makes his most influential philosophical contributions in essays composed between 1792 and 1796. The common feature of the first group of these essays is their critical engagement with Kant's philosophy. The aborted project of the "Kallias-Letters" (1793; published 1847) attempts in Kantian terms to establish something Kant declared impossible: "an objective concept of beauty" and, indeed, one that unites the realms of nature and freedom. In the "Kallias-Letters" Schiller accordingly construes beauty as "freedom in the appearance" of something, an appearance that is the natural or artistic, dynamic counterpart to moral autonomy. In "On Grace and Dignity" (1793) Schiller takes further aim at Kant's dualism, in particular, his account of an obligatoriness that is independent of grace ("the expression of a beautiful soul, where sense and reason harmonize").

Schiller's remarks provoke an exchange of letters and a public response in Kant's *Religion within the Limits of Reason Alone* (1793), where Kant suggests that any apparent disagreement can be resolved by distinguishing duty, the dignity of which is necessarily independent of grace, from virtue, which is not. Though Schiller accepts the suggestion in correspondence with Kant, he ultimately finds the distinction unpersuasive. Nevertheless, Schiller utilizes themes from Kant's aesthetics to develop a conception of tragedy in other essays from this period, notably, "On the Reason for Taking Pleasure in Tragic Subjects" and "On the Art of Tragedy" in 1792 and "On the Pathetic" in 1793. In particular, in Kant's notion of the dynamically sublime, the aesthetically pleasing displays of human beings' moral capacity to defy nature's otherwise all-powerful sway over them, Schiller finds the key to explaining the point of tragedy, though he invests art with a purpose beyond the confines of Kant's aesthetics. As Schiller puts it in the opening lines of "On the Pathetic," "Portrayal of suffering—as mere suffering—is never the end of art, but as a means to this end it is of the utmost importance to art. The ultimate purpose of art is to depict what transcends the realm of the senses and the art of tragedy in particular accomplishes this by displaying morality's independence, its freedom, in the throes of passion, from nature's laws" (1993 [1793] p. 45).

THE AESTHETIC LETTERS

Schiller's most influential work on aesthetics is *On the Aesthetic Education of Man in a Series of Letters* (1795). In this work (hereafter *Letters*) Schiller frames an argument for the necessity of an aesthetic education against the backdrop of a dire assessment of contemporary culture. Echoing Jean-Jacques Rousseau and anticipating Karl Marx, the assessment emphasizes the stupefying fragmentation and lifeless mechanism of society. Still, neither reason nor politics, Schiller argues, provides an answer to humanity's plight. The French Revolution had demonstrated only too well the failure of political reform without a moral transformation of the citizenry, that is, a transformation of individuals into citizens. As for reason, if it is the answer, Schiller asks, why in an "enlightened age" are we still barbarians? With art as the sole remaining alternative Schiller announces his central thesis, "If man is ever to solve the problem of politics in practice he will have to approach it through the problem of the aesthetic, because it is only through beauty that man makes his way to freedom" (1993 [Letter 2, 1795], p. 90). Though Schiller sometimes (e.g., Letter 14) ascribes freedom and morality solely to the rational side of human nature, the overriding sense of freedom at work in the *Letters* is freedom as self-mastery, equally liberated from the tyranny of nature and the tyranny of ideas. (In a footnote to Letter 19 Schiller acknowledges the possible misunderstandings caused by these two notions of freedom.)

Though the example set by the Greeks, Schiller submits, provides reason not to despair, he is well aware that experience and the historical record seem to speak volumes against the thesis. Still, they do so only if there is no transcendental path to a nonempirical, purely rational concept of beauty. Schiller accordingly proposes just such a path that takes its bearings from "the sheer potentialities" of human nature, potentialities that he juxtaposes with "what is absolute and unchanging" and the "necessary conditions" of human life. Though he feels no need to justify the considerable presuppositions built into this precarious move, what no doubt justifies it in his mind is

a fundamental analogy running throughout the *Letters*, namely, the analogousness of individual and political self-production to artistic production. In each case the reality in question can be conceived as the product of shaping something natural provided by experience, according to an idea that is, at least in regard to the initiative in question, irreducible to the respective experience of nature.

On the basis of this same analogy, the integrity of the reality (the production) in question demands that both nature and the idea—or, analogously, feeling and principle, the human condition and the human person—be given their due. Corresponding to this dual necessity are two basic laws of human nature, namely, "to externalize all that is within it, and give form to all that is outside it," and two basic drives: a sensuous drive toward the material content of individual, momentary sensations, and a formal drive toward freedom in the form of universal, eternal laws. While the sensuous drive acts as a physical constraint and the formal drive as a moral constraint, the "task" of culture, Schiller submits, is to amplify each drive to the point where they have a moderating effect on one another. Departing from Kant and appropriating Johann Gottlieb Fichte's accounts of a dialectical unity, Schiller declares that freedom requires, not the subordination of one drive to the other, but their coordination.

Schiller acknowledges the utopian character of the task. Still, he submits that there are moments in life when feeling and thinking merge, when human beings are able to realize both drives in a complementary way. These are the moments when human beings play. As Schiller famously puts it, "[M]an only plays when he is in the fullest sense of the word a human being, and he is*only fully a human being when he plays*" (1993 [Letter 15,1795] p. 131). (In Letter 27 Schiller gives a genealogy of play, from the physical play of an overflowing nature to the free play of human fantasy and association, culminating in aesthetic play with the capacity to transform sexual desire.) The play drive, as Schiller calls it, reconciles the otherwise competing sensuous and formal drives through its preoccupation with an object that combines their respective objects, life and form. In this way Schiller introduces his definition of beauty as a living form that is the object of the play drive. Precisely by yielding these moments of play, beauty is both a regenerative means to and a symbol of the consummate freedom that is, in his eyes, the destiny of humankind. Beauty here is not an empty (purposeless) form and the experience of it is not merely a matter of taste or the play of human faculties. Instead, it is a living form that embodies in a concrete, autonomous way the unity of feeling and principle, of sense and reason. So conceived, beauty has a vitality that transcends human subjectivity without leaving it behind and yet, for this reason, holds an incomparable historical promise for humanity.

The already mentioned tension in Schiller's conception of moral freedom takes on a new twist as Schiller describes freedom as the point where the sensuous and rational drives, far from being coordinated and facilitated, are said to be "canceled" (Letter 19). Further complicating matters, he gives an account of an "aesthetic condition" as a necessary means of predisposing human beings to a moral condition, "Man in his *physical* condition merely suffers the dominion of nature; he emancipates himself from this dominion in the *aesthetic* condition, and he acquires mastery over it in the *moral*" (1993 [Letter 24, 1795] p. 156). Still, if the aesthetic condition is now depicted as necessary for the transition to morality, its necessity is not something that one can leave behind. Beauty continues to be living proof "that a human being need not flee matter in order to manifest herself as spirit" (1993 [Letter 25, 1795], p. 165).

The transition from the aesthetic condition to the moral condition is supposedly far easier than the transition to the former from the physical condition. Hence, Schiller devotes his final remarks (Letters 26–28) to the role of "aesthetic semblance" in the former transition. Basic needs must be met, he notes, before aesthetic semblance can be indulged, though such indulgence is also a natural development of seeing and hearing. These two senses do not simply receive but help produce their objects. In the process, the play-drive develops, as people find enjoyment in mere semblance, as does the mimetic drive to shape and form this or that semblance into something relatively self-sufficient (though only relatively since it is a human product and subject to human dictates). As these drives develop, the realm of beauty expands but also gives further definition to the boundaries between semblance and reality. Moreover, only in this world of semblance does the artist enjoy sovereign rights. What makes the artist an artist and renders semblance aesthetic is a certain honesty (no pretense of being real) and autonomy (dispensing with all support from reality).

In the end, the aesthetic semblance is self-reflexive and self-redeeming. In an important respect art is the semblance of semblance, the illusion of illusion. The aesthetic education overturns a deficient, actual stage of human nature because art is capable of articulating ever higher human possibilities. Moreover, these are possibilities at the crossroads of the individual and the species. In

contrast to a strictly private sensual pleasure the enjoyment of semblance is a pleasurable activity that is inherently shareable, though not through some dictate of a *volonté générale*. Herein lies yet another side to the promise of beauty discussed earlier. Only in an aesthetic state (*Staat*) can we confront each other, not as enforcers of our respective rights ("the fearful kingdom of forces") or as executors of our wills ("the sacred kingdom of laws"), but as free and equal citizens, "the third joyous kingdom of play and of semblance" (1993, [Letter 27, 1795], p. 176).

POETS, PHILOSOPHY, AND PSYCHOLOGY

While Schiller concentrates in the *Letters* on art's prospects of overcoming modernity's alienating effects on humanity at large, his final major study, *On Naive and Sentimental Poetry* (1795–1796), turns to those effects on writers themselves. Naive poets, typified by ancient authors such as Homer, write effortlessly in a straightforward way without intruding themselves onto the scene, whereas "sentimental" (self-conscious) poets, so typical among modern writers like Ariosto, express their feelings about the scenes they depict. Characterizing the difference in terms of nature, Schiller explains, "The poet either is nature or will seek it. The former constitutes the 'naive,' the latter the 'sentimental' poet" (1993 [1795], p. 200). Thus, sentimental poets, in contrast to naive poets, are acutely aware of the difference between reality and their ideas and idealizations. Thus conflicted in their mode of feeling, they either mock reality in pathetic or playful satires, mourn the absence or loss of the ideal in elegies, or—most difficult of all—celebrate its future realization in idylls. Schiller's use of the terms *naive* and *sentimental* is idiosyncratic; *naive* does not mean simplistic but direct, and *sentimental* does not mean maudlin but reflective.

Moreover, he construes the difference between these notions at times historically, at other times theoretically, to designate antithetical kinds of poetic consciousness in some contexts, and contrary traits within a single poet in others. For example, Goethe is a modern naive poet who is nonetheless capable of treating a theme "sentimentally," as in his 1774 novel *Sorrows of the Young Werther*. (The contrast between naive and sentimental is in fact motivated, some argue, by Schiller's attempt to come to terms with what he takes to be the difference between Goethe's natural genius and his own more reflective, labored approach to writing.)

Nevertheless, in the first two parts of the essay, Schiller manages to accord each of these divergent literary modes its due, while conceding "that neither the naive nor the sentimental character, considered in itself, can completely exhaust the ideal of beautiful humanity, an ideal that can only emerge from the intimate union of both" (1993 [1796], p. 249) That union itself is, Schiller adds, present only in "a few, rare individuals" since the difference between the naive and the sentimental poet is, he maintains, rooted in a broader difference as old as culture itself. Accordingly, in the third and final part of the essay, Schiller inscribes the difference between naive and sentimental poetry in a psychological profile of the difference between realists and idealists, that is, those who allow themselves to be determined in the end by nature or reason, respectively, be it in the form of the competing theoretical demands of common sense and speculation or the rival practical demands of happiness and nobility.

In "Concerning the Sublime" (first published in 1801 but begun around 1795) Schiller argues that sublimity must come to the aid of beauty in completing an aesthetic education, not least because nature's intransigence defeats philosophy's attempts to bring "what the moral world *demands* into harmony with what the real world *does*"(1993 [1801], p. 81). According to some critics, besides signaling a departure from the more optimistic (idealist) chord struck in the *Letters* and even the vestiges of a rationalist idea of harmony in *Naive and Sentimental Poetry*, this emphasis on philosophy's limitations, with the grounding of realism and idealism in a "psychological antagonism," explains why Schiller's philosophical reflections on art largely come to a halt and he turns his attention once again to the stage. (One particularly noteworthy exception is the criticism of naturalism in the preface to the book version of the *Bride of Messina* in 1803, titled "On the Use of the Chorus in Tragedy," an essay utilized by Friedrich Nietzsche in *The Birth of Tragedy* in 1871).

The influence of Schiller's writings on German idealists and romantics is enormous. Shortly after the appearance of the *Letters*, Georg Wilhelm Friedrich Hegel writes Schelling that they are a "masterpiece," and Johann Christian Friedrich Hölderlin makes plans to write his own "New Letters" on the same topic. Shaken in his neoclassicist beliefs by Schiller's deft counterpoint of naive and sentimental poetry, Friedrich von Schlegel famously reconstrues them as "Classical" and "Romantic" poetry. Twenty years later, in his lectures on aesthetics, Hegel pays tribute to Schiller's "great service of having broken through the Kantian subjectivity and abstraction and having dared to go beyond it, grasping unity and recon-

ciliation as the truth intellectually and realizing it artistically"(Hegel 1970 [1835], p. 89).

See also Aesthetics, History of; Beauty; Fichte, Johann Gottlieb; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hölderlin, Johann Christian Friedrich; Humboldt, Wilhelm von; Kant, Immanuel; Marx, Karl; Rousseau, Jean-Jacques; Schlegel, Friedrich von; Tragedy.

Bibliography

WORKS BY SCHILLER

Schillers Werke. 43 vols., edited by Julius Petersen et al. Weimar, Germany: Böhlaus, 1943–.

Naive and Sentimental Poetry, and On the Sublime: Two Essays. Translated by Julius A. Elias. New York: Ungar, 1966.

On the Aesthetic Education of Man, edited and translated by Elizabeth M. Wilkinson and L. A. Willoughby. Oxford, U.K.: Clarendon Press, 1967.

Friedrich Schiller. "Über die ästhetische Erziehung des Menschen in einer Reihe von Briefen: Text, Materialien, Kommentar, edited by Wolfgang Düsing. Munich, Germany: Hanser, 1981.

Essays, edited by Walter Hinderer and Daniel O. Dahlstrom. New York: Continuum, 1993.

WORKS ON SCHILLER

Hegel, Georg Wilhelm Friedrich. *Werke*. Vol. 13. Frankfurt am Main: Suhrkamp, 1970.

Martin, Nicholas. *Nietzsche and Schiller: Untimely Aesthetics*. Oxford, U.K.: Clarendon Press, 1996.

Reed, T. J. *Schiller*. New York: Oxford University Press, 1991.

Sharpe, Leslie. *Friedrich Schiller: Drama, Thought and Politics*. New York: Cambridge University Press, 1991.

Sharpe, Leslie. *Schiller's Aesthetic Essays: Two Centuries of Criticism*. Columbia, SC: Camden House, 1995.

Daniel O. Dahlstrom (2005)

SCHLEGEL, FRIEDRICH VON

(1772–1829)

Friedrich von Schlegel, a critic and philosopher, whose writings spearheaded early German Romanticism, started out as a devotee of Greek poetry. Born to an illustrious literary family in Hanover and classically trained, Schlegel was an unhappy and unfocused student of law at Göttingen and Leipzig from 1790 to 1793, all the while piling up enormous gambling debts. Fleeing creditors and abandoning his legal studies, he moved in 1794 to Dresden where, inspired by Caroline Böhmer, his future sister-in-law, he launched his literary career with essays extolling ancient poetry's superiority to modern poetry. In "On the Study of Greek Poetry" (completed 1795, published 1797), he echoes Johann Joachim Winckelmann by attributing the greater unity, objectivity, and naturalness of ancient works to the Greeks' single-minded pursuit of idealized beauty.

PHILOSOPHY, CRITICISM, AND THE ROMANTIC TURN

Schlegel eventually wrote the *History of the Poetry of the Greeks and Romans*, but by the time the only volume was published in 1798, his view of modern poetry had changed. Already in his 1795 essay his admiration for William Shakespeare seems to belie his insistence on Sophocles' superiority. His politics, too, though inspired by the ancients, were decidedly unconventional, as evidenced by his defense of the legitimacy of insurrection in his "Essay on the Concept of Republicanism" (1796), itself a critical review of Immanuel Kant's "Toward Perpetual Peace" (1795). But it was chiefly Friedrich Schiller's *On Naive and Sentimental Poetry* (1795–1796)—with its balanced judgment of the comparable virtues of ancient, "naive" and modern, "sentimental" (self-conscious) poetry and its reference to an even loftier poetry—that challenged Schlegel to reconsider his earlier views. Also like Schiller, Schlegel began to embrace Johann Gottlieb Fichte's dialectical vindication of human dignity in the face of the threats posed to it by empiricism and mechanistic materialism. A growing awareness of Fichte's impoverished view of nature eventually tempered this enthusiasm. Contrasting the "consistent empiricist" for whom everything sacred is "nonsense" with "mystics" as the real source of philosophy, Schlegel declares "Spinoza the best mystic known to us before Fichte" (*Kritische Ausgabe*, Vol. 18, p. 5). At the close of the eighteenth century Schlegel searched, much like Schelling and Georg Wilhelm Friedrich Hegel, for a philosophical path combining Benedict (Baruch) de Spinoza's pantheistic naturalism with Fichte's idealism.

Still, neither in Spinoza nor in Fichte did Schlegel find the sort of historical sensibility already exhibited in his early neoclassicist phase. This sensibility was accentuated in 1796 when, further signaling his departure from classicism, Schlegel begrudgingly accepted Kant's argument that there are no objective rules for aesthetic judgments. Schlegel proposed that critics compensate for this lack of rules by being as comprehensively informed as possible of not only a writer's but also an entire culture's literary repertoire. At the same time he insisted that "criticism compares a work with its own ideal"(*Literary Note-*

books, p. 1135). This joint concern for a work's context and its sui generis character (exemplified by Schlegel's essays on Friedrich Heinrich Jacobi's *Woldemar* [1779], Georg Forster's works, and Gotthold Ephraim Lessing in 1797) would profoundly influence the development of hermeneutics by Wilhelm Dilthey and others.

In 1797 Schlegel moved to Berlin where close friendships with Friedrich Daniel Ernst Schleiermacher, Johann Ludwig Tieck, and Novalis (Friedrich Leopold von Hardenberg) gave rise to the new literary and philosophical movement eventually known as Romanticism. Its chief organ, the journal *Athenäum*, edited by Schlegel and his brother, August Wilhelm, contained Schlegel's most influential contributions to Romantic theory: "Fragments" and an essay on Johann Wolfgang von Goethe's *Wilhelm Meister* in 1798, and "Ideas" and "Dialogue on Poetry" in 1800. Another important source for Schlegel's theory is "Critical Fragments," printed in Johann Friedrich Reichardt's *Lyceum der schönen Künste* (1797). The form of fragments is itself a testament to the new theory's defiance of traditional literary theory. In memorable fashion Schlegel contrasts "Classical" with "Romantic" poetry, which disregards the traditional insistence on preserving purity of genre (epic, drama, and lyric). The novel (*Roman* in German) is, at least at first, paradigmatic for this theory that applauds the highly imaginative, genre-mixing fantasies (often with a love interest) typified by such Romance language writers as Dante Alighieri, Petrarch, and Miguel de Cervantes (but also by Shakespeare). In 1799 Schlegel provides his own example of a Romantic novel: *Lucinde*, a celebration of a complete but extramarital love, notoriously based on his affair with his future wife, the divorcée Dorothea Veit (Moses Mendelssohn's daughter).

THE THEORY OF ROMANTIC POETRY

In *Athenäums-Fragment 116*, Schlegel's most influential account of Romantic poetry, he deems it "progressive universal poetry" because it aims not only to reunify all genres and connect poetry with philosophy but also to mingle and fuse "poetry and prose, genius and criticism, the poetry of the educated and the poetry of the people, to make poetry alive and social and to make life and society poetic, to poeticize wit, to fill and saturate the forms of art with matters of genuine cultural value" (*Athenäum I*, p. 220). To this end, a Romantic work is supposed to present sentimental but actual historical material in a witty, fantastic form ("an artfully ordered confusion") that is a synthesis of Eros and chaos, infinite unity and infinite fullness, mirroring nothing less than the universe as a divine manifestation. The universe itself is conceived as a poem of the Godhead at this intersection of metaphysics and literary aesthetics.

In a good poem, as in reality, everything seems capricious and instinctive, though it is in fact necessary and deliberate. So, too, the Romantic artist must combine deadly seriousness with playfulness in a "constant self-parody," as Schlegel puts it. The model here is Socratic irony, a sense of the limitlessness of things and one's own limited capacity to express them, combined with the utter necessity of doing so.

In the final volume of *Athenäum* the emphasis on criticism and universality in the "Fragments" gives way to an enthusiasm for religion ("the all-animating world-soul of culture") and mythology (how "religion must appear in the world of language"(*Athenäum II*, p. 734, 740). In his "Ideas," which is deeply influenced by Schleiermacher, Schlegel touts the religious complementarity of poetry and philosophy that he also counterposes as realism and idealism, respectively. While claiming that "logic can develop into philosophy only through religion" and that "only someone who has his own religion can be an artist," Schlegel also insists paradoxically that "there is as yet no religion"(*Athenäum II*, p. 736, 751). Returning to this theme in "Dialogue on Poetry," he attributes the isolation of modern poets to their lack of a focal point such as ancient mythology provided ancient poets. He accordingly calls for the creation of a new mythology. This new mythology, like the ancient, would represent nature symbolically, though now against the background of philosophical idealism and the new physics (Schilling's philosophy of nature) and with an openness to the mythologies of the Orient.

THE LATER WORKS

After 1800 Schlegel's fortunes initially took a turn for the worse. He failed as a lecturer on transcendental philosophy at the University of Jena and as a playwright, his collaboration with his brother ended with the publication of *Characterizations and Criticisms* (1801), and his relationships to other members of the Romantic movement deteriorated. Isolated and financially strapped, Schlegel moved in 1802 to Paris, where he published the periodical *Europa*, in which he influentially opposed classicism again, this time by championing the symbolism of early modern religious painters. Vainly looking for a professorship, Schlegel moved to Cologne in 1804, where he helped rediscover German Gothic architecture and published *On the Language and Wisdom of India* (1808). Seminal for the development of Sanskrit studies, comparative linguistics,

and Indian philosophy, this work also contained attacks on pantheism that introduced Schlegel's final, Catholic phase of thinking.

Following his conversion to Catholicism, Schlegel moved to Vienna and worked for the Austrian government (serving as Prince Klemens von Metternich's representative at the Diet of Frankfort) but also found time to give well-received lectures: *On Modern History* (published 1811) and the monumental *History of Ancient and Modern Literature* (published 1815). From 1820 to 1823 he published the periodical *Concordia*, to which he contributed "Signature of the Age," a plea for an "organic" state headed by a strong monarchy and animated by "corporations," most prominently, the Church. In lectures on the philosophy of life, history, and language in his final years (published from 1828 to 1830), Schlegel challenged reigning philosophical systems—deduced, in his view, from merely a part of human consciousness—with a "Christian philosophy" grounded in the total, personal experience of a thinker as a believer.

See also Dante Alighieri; Dilthey, Wilhelm; Fichte, Johann Gottlieb; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Jacobi, Friedrich Heinrich; Kant, Immanuel; Lessing, Gotthold Ephraim; Novalis; Petrarch; Romanticism; Schiller, Friedrich; Schleiermacher, Friedrich Daniel Ernst; Spinoza, Benedict (Baruch) de; Winckelmann, Johann Joachim.

Bibliography

WORKS BY SCHLEGEL

Kritische Friedrich Schlegel Ausgabe. 35 vols., edited by Ernst Behler, et al. Munich: F. Schöningh, 1958–2002.

Athenäum, edited in two parts by Bernhard Sorg. Dortmund: Harenberg, 1989.

Dialogue on Poetry and Literary Aphorisms. Translated by Ernst Behler and Roman Struc. University Park: Pennsylvania State University Press, 1968.

Literary Notebooks 1797–1801, edited by Hans Eichner. London: Athlone, 1957.

Lucinde and the Fragments. Translated by Peter Firchow. Minneapolis: University of Minnesota Press, 1971. Partially reprinted in *Philosophical Fragments*. Translated by Peter Firchow. Minneapolis: University of Minnesota Press, 1991.

"Essay on the Concept of Republicanism." In *Early Political Writings of the German Romantics*, edited and translated by Frederick C. Beiser. New York: Cambridge University Press, 1996.

On the Study of Greek Poetry. Translated by Stuart Barnett. Albany: SUNY Press, 2001.

WORKS ON SCHLEGEL

Behler, Ernst. *German Romantic Literary Theory*. New York: Cambridge University Press, 1993.

Beiser, Frederick C. "Friedrich Schlegel's Absolute Idealism." In *German Idealism: The Struggle against Subjectivism, 1781–1801*, 435–461. Cambridge, MA: Harvard University Press, 2002.

Eichner, Hans. *Friedrich Schlegel*. New York: Twayne, 1970.

Frank, Manfred. *The Philosophical Foundations of Early German Romanticism*. Translated by Elizabeth Millán-Zaubert. Albany: SUNY Press, 2004.

Daniel O. Dahlstrom (2005)

SCHLEIERMACHER, FRIEDRICH DANIEL ERNST
(1768–1834)

Friedrich Daniel Ernst Schleiermacher was nineteenth-century Protestantism's great systematic theologian. It was he who marked the points of the compass for much of subsequent theology and philosophy of religion. Like St. Augustine, Schleiermacher desired to know God and the soul, and his place in the history of philosophy is due largely to the fact that he was able to state in modern language and concepts the great Augustinian conviction that religious faith is native to all human experience. Therefore, the knowledge of God and the knowledge of the soul are two orders of knowledge that must be distinguished but cannot be separated.

LIFE

Schleiermacher was first and foremost a preacher and theologian, a church statesman, and an educator. He carried out his work as a philosopher in the context of the great idealist systems of Friedrich von Schelling, Johann Gottlieb Fichte, and G. W. F. Hegel, but instead of attempting to imitate these men he applied himself to the critical analysis of religion, both in its personal and societal manifestations, without reducing such experience to some form of philosophic intuition. The upbringing that his father, a Reformed clergyman, gave him and his early education in Moravian institutions set Schleiermacher upon this course. After studying at the university in Halle and taking his examinations for ordination in 1790, he served briefly as a private tutor to the family of Count Dohna in East Prussia and as a minister in the Prussian town of Landsberg. In 1796 Schleiermacher settled in Berlin as a preacher, became a close friend of Friedrich von Schlegel, and emerged as an interpreter of religion to the romantic worldview that Schlegel himself epitomized. *On Religion: Speeches to Its Cultured Despisers* (1799) gave Schleiermacher a national reputation at the age of thirty. The following year another publication, *Soliloquies,*

attested to Schleiermacher's thorough absorption of the spirit of romanticism, but at the same time it indicated the direction that his ethical interests were to take in the future, as in his *Grundlinien einer Kritik der bisherigen Sittenlehre* (Outline of a critique of previous ethical theory; 1803).

The relation between the religious and ethical dimensions of life constituted a major preoccupation of Schleiermacher's maturity, and it is here that his indebtedness to and divergence from Immanuel Kant are clearly evident. Of decisive importance during his Berlin sojourn was his embarking upon the translation of Plato, in the course of which his mind became imbued with the philosophy of the author of the *Republic.* By 1804 Schleiermacher was teaching philosophical ethics (philosophy of culture), theology, New Testament, and hermeneutics at Halle. By 1810 he was lecturing as professor of theology at the University of Berlin, where for the remainder of his life he taught dogmatic theology, New Testament theology and criticism, hermeneutics, practical theology, history of philosophy, ethics, and dialectics, to name only the more important of the wide variety of subjects with which he dealt. Concomitantly he held an appointment as preacher at the Dreifaltigkeitskirche, to which he attracted persons from all sections of Berlin, and from this pulpit he wielded a powerful moral influence on the nation. In ecclesiastical politics he labored for the union of the Lutheran and Reformed churches in Prussia, and in national politics he worked not only for stiffer resistance to French expansionism under Napoleon Bonaparte but for internal social reform.

The Christian Faith (*Der christliche Glaube nach den Grundsätzen der evangelischen Kirche im Zusammenhange dargestellt*) appeared in 1821–1822 and in revised form in 1830–1831. Together with the *Brief Outline of the Study of Theology* (1st edition, 1811) and the two open letters concerning the revised edition of *The Christian Faith* which Schleiermacher wrote to a close friend (*Sendschreiben über seine Glaubenslehre an Dr. Lücke,* 1829), *The Christian Faith* gives us not only Schleiermacher's thought on Christian doctrine and substantive theological issues but also his conception of the organization of the theological disciplines and of systematic theology itself. Schleiermacher made Protestant theology methodologically self-conscious.

PHILOSOPHY OF CULTURE

Schleiermacher criticized Kant for tacitly making ethics into a "highest science" that ignored and devaluated the particular and idiosyncratic in human nature. Ethics, Schleiermacher argued, is the discipline that has for its object "reason in history." Reason never appears except in historical personality—in the personalities of both individual persons and corporate persons. This position leads to a significant relaxation of the Kantian separation between practical reason, on the one hand, and the inclinations, temperament, talent, etc., on the other. Schleiermacher viewed these "accidents" and, indeed, the entire spatial, temporal embodiment of reason—apart from which we have no self-consciousness and hence no access to reason—not merely as the "place" of reason in its practical and theoretical functions but also as the organ of reason, by which reason itself is conditioned. The notion of a pure, universal reason could, therefore, be only a regulative concept for Schleiermacher.

Insofar as we consider reason in its practical capacity, as a willing or organizing activity, it is not the quest of virtue and autonomous assent to a self-imposed universal law that is foremost in view., but rather the sight of an ethical agent acting according to his own individuated rational nature. Moreover, the individuation of the ethical agent is accomplished not only by the "natural" accidents of time and place but also by the communities, societies, and institutions of which the individual person is the offspring. Schleiermacher presents the ethical agent as an end in himself, that is, as a good, who produces goods according to the peculiar law of his own unique nature. The doctrine of the highest good is formulated through the delineation of the relations of community and reciprocity in which such agents stand to each other, inheriting and endowing, receiving and bestowing. The primary forms in which these relations appear are the family, the nation, the church, the institutions of learning, and what Schleiermacher calls free sociality (*Geselligkeit*).

Nature and society affect reasoning in its theoretical as well as practical operations. When we think, we are conscious of engaging in an activity that is common to all men; nevertheless, our thinking, even at the most abstract level, as in thinking about thought itself, is in actuality predicated upon the specific organization of the physical means of sensation as well as upon the prior existence of a particular system of communication. The speculative activity of reason is thus conditioned by the natural medium in which it is individuated and shaped by the historical, moral character of the primary media (for example, a particular language) through which it maintains itself. Discourse is the means for the sociality of thinking, as Schleiermacher liked to say, and thinking is the inner side of speaking. He defined dialectic as the principles of correctly conducting a dialogue in the realm

of pure thinking and taught that all thinking proceeds in the form of dialogue or colloquy. On these grounds, Schleiermacher ruled out the possibility of an intuition of the absolute or of a highest science; the ideal and the real appear only as already informed by each other; pure spirit and matter lie outside of experience. Consequently, the ideal of a universal philosophy, for example, is nullified by the lack of a universal language and the impossibility of such.

The person, as the subject of the activities of thinking/knowing and of willing/doing, is more than a being composed of mind and body, individuated by time and space. A person not only is differentiated from others by nature and history but inwardly differentiates himself and acknowledges such an inward differentiation in all other human beings. That by virtue of which the person makes this inward differentiation is the proprium (*Eigenthümlichkeit*). It is this property in each man that endows him with a life unity, an inalienable identity. Schleiermacher described this proprium as the peculiar organization that reason assumes for itself in each man. However, the life unity, or identity, of the individual person can never come to direct and full expression either in thinking/knowing or in willing/doing, although it accompanies and informs each of these rational activities. The self-consciousness that this sense of identity requires is a self-consciousness to be distinguished—though not isolated—from the forms of self-consciousness in which the subject is responding to or acting upon external objects.

Schleiermacher appropriates the word *feeling* for this form of self-consciousness, whose content is the given identity and unity of the self, incapable of being derived from others or surrendered to them. Feeling, thinking, and doing thus make up the three forms of consciousness that constitute the self-consciousness which distinguishes persons. Correspondingly, every person must be seen as a participant in the life of society in both his practical and theoretical functions, but he is also one whose proprium is wholly original. In a person whose feeling form of self-consciousness remains latent or inchoate, the sense of personal identity is deficient and personal consciousness is confused or immature. Such a person fails to contribute to the common or highest good; he is an inert reflection of his world, not one who moves and enriches it; he is a person in the formal sense but is destitute of spiritual life. Since, for Schleiermacher, religion is the most highly and fully developed mode of the feeling form of self-consciousness, all of human culture ultimately depends upon the cultivation of the religious life.

RELIGION

In his earliest published work, the *Speeches,* Schleiermacher made ample use of the romantic preoccupation with the nature and value of individuality, but he qualified the world view of German romanticism in two important respects. First, an individual comes to self-knowledge only in the presence of other persons; hence the need to know and to express the self can be fulfilled only by observing and cultivating the morality of human community and communication. Second, the individual's cultivation of his own humanity—which the romantic accepted as a self-evident imperative—requires that he acknowledge his religious nature, as well as his aesthetic, scientific, and moral nature, and that he cultivate this side of his nature, or self-consciousness, by seeking out religious community. Schleiermacher's thesis, from 1799 to his death, was that man is a religious being. But since the individual must always appropriate his humanity in a fashion that is at once concordant with his generic identity and accordant with his own peculiar identity, religion is as much a problem for the individual as it is a natural endowment. In his mature thinking, as he came to align himself theologically with Augustine and John Calvin, Schleiermacher stressed not only the fact that man is a religious being but also the fact that the most fundamental, pervasive confusion inhibiting human consciousness is religious confusion. Thus, in his Christian theology, he described sin as the failure to maintain a clear distinction between that upon which men are entirely dependent, God, and that upon which men are only relatively dependent, namely, objects within the world.

In *The Christian Faith,* Schleiermacher stated that religion is a determination of feeling. More narrowly defined, it is a feeling of being absolutely dependent, and this feeling, he believed, is one and the same thing with consciousness of being in relation with God. A number of elements in this characterization need to be distinguished if Schleiermacher is to be understood. (1) The feeling of being absolutely dependent is also the feeling of identity, through which the individual is conscious of his inner uniqueness; in describing this feeling as one of being absolutely dependent, Schleiermacher was calling attention to the fact that the identity, or life unity, of the individual is an endowment which cannot be derived from any of the intellectual or volitional relations in which the self stands to other persons and forces, taken either singly or together. In this sense, the individual is utterly dependent, for the particular constitution of his existence, on a "whence" that cannot be rendered conceptually. Hence, the feeling of absolute dependence is not expressive of a

felt deficiency or of awe, as it is according to the interpretation of Rudolf Otto in *The Idea of the Holy*; nor is it wholly the same as Paul Tillich's conception of faith as being ultimately concerned about that which concerns us ultimately, since this concern is aroused in part by what Tillich called "nonbeing." (2) The feeling of being absolutely dependent—or "immediate self-consciousness" or "God-consciousness"; Schleiermacher regarded all three terms as equivalent—is discernible only because self-consciousness also involves thinking and willing, which are forms of rational relation between the person and his world, forms involving consciousness of "relative dependence" and "relative freedom." The feeling of being absolutely dependent is distinguishable from the feeling of relative dependence by virtue of the fact that in the latter a person stands in the relations of community and reciprocity with nature and society, while in the feeling of absolute dependence there is no reciprocity present. Consequently, there can be no consciousness of being in relation to God, apart from consciousness of being in relation to the world. (3) The original meaning of the word *God* is not a concept of perfect being, or the like, but the felt relation of absolute dependence. Hence, religion arises not in ideas, nor—for that matter—in willing, but in the immediate consciousness of what Schleiermacher described to Lücke as "an immediate existence-relationship." (4) In fact, then, religion is more than a determination of feeling; it is the name Schleiermacher gives to the personal self-consciousness in which the feeling of absolute dependence and consciousness of the world coexist and must achieve or receive a living, stable order.

The religion that Schleiermacher described in this way is a purely formal and abstract religion, which exists nowhere in actuality. In conformity with the principles we have outlined above, he insisted that religion always appears in a particular social and historical form. The great religions are religions bearing the stamp of their founders, and he defined Christianity as a monotheistic faith of the teleological variety in which everything is related to the redemption accomplished by Jesus of Nazareth. Everything in the outward, social, and institutional aspect of Christianity is related to its founder, and similarly, everything pertaining to the inner piety of the Christian is related to the historical figure of the redeemer. Thus, while Christianity is, without question, the religion on the basis of which Schleiermacher formed his understanding of all other religions, what is of more importance is that he was the first among modern theologians to perceive that Christianity is historical in two senses. Not only does it have a history, but each Christian becomes a Christian by appropriating to his total self-consciousness the relation to Jesus Christ. Christ must become a part of the self-consciousness, or inner history, of the Christian. There is no part of the relation to God, Schleiermacher stated, in which the relation to Christ is not also actively present. Hence, Schleiermacher revived in his conception of the feeling of being absolutely dependent the Augustinian notion of the inseparability of the knowledge of the soul and the knowledge of God; at the same time he originated the distinctive form of modern Protestant theology—Christocentrism, or Christ as the center of the individual's inner religious consciousness.

See also Augustine, St.; Augustinianism; Calvin, John; Faith; Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; Kant, Immanuel; Otto, Rudolf; Philosophy of Religion, History of; Plato; Romanticism; Schelling, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von; Tillich, Paul.

Bibliography

WORKS BY SCHLEIERMACHER

Collected Works and Letters

Aus Schleiermachers Leben in Briefen. 4 vols. Berlin, 1860–1863.

Sämtliche Werke. Berlin, 1835–1864. Published in three divisions: I, theological; II, sermons; III, philosophical and related subjects.

Schleiermachers Briefwechsel mit J. Chr. Gass. Berlin, 1852.

English Translations

Brief Outline of the Study of Theology. Translated by W. Farrer. Edinburgh, 1850.

The Christian Faith. Translated by H. R. Mackintosh and J. S. Stewart. Edinburgh, 1948; New York: Harper and Row, 1963.

Christmas Eve: A Dialogue on the Celebration of Christmas. Translated by W. Hastie. Edinburgh, 1890.

The Life of Schleiermacher as Unfolded in His Autobiography and Letters. 2 vols. Translated by F. Rowan. London, 1860.

"On the Discrepancy between the Sabellian and Athanasian Method of Representing the Doctrine of a Trinity in the Godhead." Translated by Moses Stuart. *Biblical Repository and Quarterly Observer* 6 (1835): 1–116.

On Religion: Speeches to Its Cultured Despisers. Translated by John Oman. London, 1893; New York, 1958.

Selected Sermons of Friedrich Schleiermacher. Translated by Mary F. Wilson. New York, n.d.

Soliloquies. Translated and edited by Horace Friess. Chicago, 1926.

Critical Editions

Der christliche Glaube, edited by Martin Redeker. 2 vols. Berlin, 1960. Contains appendices and Schleiermacher's own marginal notations.

Grundriss der philosophischen Ethik, edited by A. Twesten and F. M. Schiele. Leipzig, 1911.

Hermeneutik, nach den Handschriften, edited by Heinz Kimmerle. Heidelberg, 1959.

Kurze Darstellung des theologischen Studiums, edited by Heinrich Scholz. Hildesheim, 1961.

Monologen, edited by F. M. Schiele and Hermann Mulert. Leipzig, 1914.

Reden über die Religion, edited by G. C. B. Pünjer. Braunschweig, 1879.

Schleiermachers Sendschreiben über seine Glaubenslehre an Lücke, edited by Hermann Mulert. Giessen, 1908.

Weihnachtsfeier, edited by Hermann Mulert. Leipzig, 1908.

WORKS ON SCHLEIERMACHER

Barth, Karl. *From Rousseau to Ritschl.* London: SCM Press, 1959. A sophisticated chapter on Schleiermacher by his foremost theological critic.

Brandt, Richard. *The Philosophy of Friedrich Schleiermacher.* New York: Harper & Brothers, 1941. The most detailed examination of Schleiermacher's philosophy in English.

Dilthey, Wilhelm. *Leben Schleiermachers.* Berlin, 1870; 2nd ed., edited by H. Mulert. Berlin, 1922. The classic biographical work on Schleiermacher, but extends only to the period at Halle. Should be supplemented by Dilthey's articles in his *Gesammelte Schriften,* 2nd ed. Stuttgart, 1959–1960. Vols. IV and XII.

Niebuhr, Richard R. *Schleiermacher on Christ and Religion.* New York: Scribners, 1964. An account of the central ideas in Schleiermacher's theology against the background of his philosophy.

Süskind, Hermann. *Der Einfluss Schellings auf die Entwicklung von Schleiermachers System.* Tübingen: Mohr, 1909. Valuable account of the early development of Schleiermacher's thinking.

Wehrung, Georg. *Die Dialektik Schleiermachers.* Tübingen: Mohr, 1920. A scholarly examination of the development of Schleiermacher's lectures on this subject.

Richard R. Niebuhr (1967)

SCHLEIERMACHER, FRIEDRICH DANIEL ERNST [ADDENDUM]

In the past forty years there has been an explosion in research on Friedrich Daniel Ernst Schleiermacher in regard to both philosophical and theological dimensions of his thought. This entry is limited to discussing three issues of significance to philosophers: religious epistemology and the problem of religious pluralism, hermeneutics, and the question of the influence of Gottfried Wilhelm Leibniz, Immanuel Kant, and others on Schleiermacher's thought as a whole.

RELIGIOUS EPISTEMOLOGY AND RELIGIOUS PLURALISM

Schleiermacher's contribution to the question of religious pluralism lies in his religious epistemology, which is developed in the first twenty-two chapters of *The Christian Faith* (1821–1822, second edition 1830/1999) as well as in *On Religion* (1799/1996; other editions followed in 1806, 1821, and 1831). In both, he offers a comprehensive theory of the nature of religion grounding it in experience. In *On Religion* he grounds religion in an original unity of consciousness that precedes the subject-object dichotomy, and in *The Christian Faith* the feeling of absolute dependence is grounded in immediate self-consciousness.

In *The Christian Faith* Schleiermacher explains that doctrines are expressions of this fundamental experience: Christian doctrines are "accounts of the Christian religious affections set forth in speech" (p. 76, § 15). This view has been labeled *experiential expressivism*. Christian doctrines are not a set of truth claims that are to be judged in virtue of their correspondence with reality, but are rather a human attempt to express in symbols the experience of absolute dependence. This original experience is immediate and is not itself conceptually structured, for any conceptual structure presupposes the subject-object dichotomy and thereby also one's counter-influence on that which is posited. God cannot be "given as an object exposed to our counter-influence, however slight this may be" (1999, p. 18; §4.4). As such, theological concepts and symbols are only indirect representations of one's consciousness of God. Given such an understanding of Christian doctrines, it is possible that two religions with differing symbols both adequately express the feeling of absolute dependence. Nevertheless, while the feeling of absolute dependence is not itself conceptually structured, it determines the way that one represents and knows oneself and the world around one. Hence, Schleiermacher states that "the world will be a different thing to a man according as he apprehends it from the standpoint of a God-consciousness completely paralyzed or of one absolutely paramount" (p. 267; §64.2).

HERMENEUTICS

Schleiermacher's hermeneutics has also received a good deal of philosophical scrutiny. There are two diametrically opposed positions on the question of how it is possible to interpret a text or utterance. The first is the structuralist position: The meaning of any given utterance is determined by the publicly available meanings of the words that constitute it. Schleiermacher calls this the

"grammatical" element of language. According to the intentionalist position the meaning of an utterance lies in the intention of the speaker. The history and inner life of the speaker is of decisive importance in determining its significance. Schleiermacher's hermeneutics, especially when understood in the context of his *Dialektik* (2001), offers a fruitful way to move beyond this impasse. According to Schleiermacher, one cannot strictly separate receptivity and spontaneity because both share a single underlying root. This plays a crucial role at several levels, the first being how one moves from sense-data to the ordinary world of tables and chairs. How the sense-data is organized will depend on the interpretive work of language: there is no bedrock given in receptivity. Ludwig Josef Johann Wittgenstein's famous "duck-rabbit" is a useful example of this. Similarly, just as sense data provide no bedrock "given," neither do the publicly available meanings of words. While language users begin from there, their own mental activity is important in shaping and sometimes even recasting those publicly available meanings. The level of the subject's activity in shaping these meanings will vary from activity to activity, from high in aesthetic endeavors to low in scientific ones. For Schleiermacher, hermeneutics is "the art of understanding … the … discourse of another person correctly" (*Hermeneutics and Criticism*, p. 3). Both grammatical and psychological elements are vital to this task.

RECEPTION OF THE PHILOSOPHICAL TRADITION

Lastly, a good deal of scholarship explores the systematic character of Schleiermacher's thought and how it relates to preceding philosophical thought. In what ways was Schleiermacher influenced by the systems of Plato, Leibniz, Spinoza, and Kant? For instance, Schleiermacher's *Dialektik* has received a good deal of scrutiny. Several scholars point to Schleiermacher's Leibnizian heritage and its relation to Schleiermacher's reception of foundational Kantian ideas. Specifically, Schleiermacher's adoption of Leibniz's complete concept, which contains all the predicates applicable to an individual, does not square with another idea essential to Schleiermacher's system, namely that one is both a spontaneous and receptive being. Schleiermacher agreed with Kant that what is given to one through sensation is necessary, although not sufficient for knowledge. But if this is true, all true judgments cannot be analytic, as the Leibnizian tradition assumed. There is an important class of judgments that are synthetic: they are true in virtue of some third thing that one becomes aware of through one's receptivity.

See also Religious Experience.

Bibliography

WORKS BY SCHLEIERMACHER

The Christian Faith. Translated by H. R. Mackintosh and J. S. Stewart. Edinburgh, 1948; New York, 1963.

Kritische Gesamtausgabe. 24 vols. to date. Edited by Hans-Joachin Birkner et al. Berlin: Walter de Gruyter, 1980–.

On Religion: Speeches to Its Cultured Despisers (1799). Translated with an introduction by Richard Crouter. New York: Cambridge University Press, 1996.

Hermeneutics and Criticism and Other Writings. Translated and edited by Andrew Bowie. New York: Cambridge University Press, 1998.

Dialektik, edited by Manfred Frank. Frankfurt am Main: Suhrkamp, 2001.

Lectures on Philosophical Ethics. Translated by Louise Adey Huish, edited by Robert B. Louden. New York: Cambridge University Press, 2002.

WORKS ON SCHLEIERMACHER

Frank, Manfred. *Das Individuelle Allgemeine: Textstrukturierung und Interpretation Nach Schleiermacher*. 1 Aufl. Ed. Frankfurt am Main: Suhrkamp, 1989.

Helmer, Christine, ed., in cooperation with Marjorie Suchoki, John Quiring, and Katie goetz. *Schleiermacher and Whitehead: Open Systems in Dialogue*. Berlin: Walter de Gruyter, 2004.

Lamm, Julia. *The Living God: Schleiermacher's Theological Appropriation of Spinoza*. University Park: Pennsylvania State University Press, 1996.

Mariña, Jacqueline. *The Cambridge Companion to Friedrich Schleiermacher*. Cambridge, U.K.: Cambridge University Press, 2005.

Jacqueline Mariña (2005)

SCHLICK, MORITZ
(1882–1936)

Moritz Schlick, one of the founders of modern analytical philosophy and a guiding spirit of the Vienna circle of logical positivists, was born in Berlin. He was a direct descendant on his mother's side of Ernst Moritz Arndt, the famous German patriot and political leader of the war of liberation against Napoleon Bonaparte. At the age of eighteen, Schlick entered the University of Berlin to study physics under Max Planck. He received his doctorate in 1904 with a dissertation on the reflection of light in a nonhomogeneous medium.

Schlick's familiarity with the methods and criteria of research in the natural sciences left him dissatisfied with the epistemological notions both of neo-Kantianism, which then dominated the German universities, and of

Edmund Husserl's phenomenology, which had already become widely known. Instead, Schlick's starting point was the analyses carried out by Ernst Mach, Hermann von Helmholtz, and Henri Poincaré of the basic concepts and presuppositions of the individual sciences. His central interest at the time was the fundamental question of what is to be understood by knowledge.

From 1911 to 1917, Schlick served as lecturer and associate professor at the University of Rostock. In this period he published a series of works, among them his *Allgemeine Erkenntnislehre* (1918; 2nd ed., 1925). These works were devoted partly to a logically precise critical discussion of traditional philosophical conceptions and partly to an elaboration of new criteria for scientific knowledge which attracted considerable attention. In these publications Schlick already presented a first systematic account of his philosophical views.

In 1921 Schlick was named to a professorship at Kiel, and a year later he accepted a call to a chair in philosophy at the University of Vienna. These two years may thus be seen in retrospect as a kind of turning point in the history of philosophy. In 1921 Ludwig Wittgenstein had published his *Tractatus Logico-Philosophicus*, and in these same years the first writings of Rudolf Carnap appeared. Under the influence of Wittgenstein and Carnap, Schlick's philosophical views underwent a profound modification, which he later characterized by saying that he no longer saw the goal of philosophy as acquiring knowledge and presenting it as a system of propositions but, rather, as the application of a method. In applying its method, philosophy must take as its aim the discovery and understanding of the meaning of the statements, concepts, and formulations of problems of the special sciences, of philosophy, and of everyday life. When philosophy is understood in this manner, as Schlick emphasized in his French essay "L'école de Vienne et la philosophie traditionelle" (*Travaux du IX^ième Congrès International de Philosophie*, Paris, 1937), it resembles the method of Socrates, who constantly strove in his conversations to clarify the concepts, assertions, traditional notions, and ordinary modes of expression found in both the philosophy and the practical life of his time.

Schlick taught at the University of Vienna from 1922 until his death in 1936. During these years he twice made trips to the United States as a visiting professor. While in Vienna, Schlick published *Fragen der Ethik* (*The Problems of Ethics*, 1930), as well as numerous papers, most of which were later collected in various volumes. But his views were disseminated most effectively, perhaps, through the discussion society that he founded and that acquired a worldwide reputation as the *Wiener Kreis*. Besides professional philosophers, regular participants in the meetings of the Vienna circle included primarily mathematicians and natural scientists but also psychologists and sociologists. They published a profusion of writings of their own, in which they applied the methods—constantly refined in discussion—of the new Vienna philosophy to the fundamental problems of scientific research.

Schlick was responsible for Carnap's appointment as lecturer at the University of Vienna. Another member of the Vienna circle was Kurt Gödel, who in this period published his famous proofs of the completeness of first-order logic and of the incompletability of formal arithmetic. Numerous scholars from Germany, Poland, England, Norway, Sweden, and the United States visited the sessions of the Vienna circle and took part in its discussions. Conflicting views frequently were championed, but the application of the most rigorous logical tools to the positions under consideration was common to all the deliberations. These discussions thus turned out to be a genuine symposium in the classical sense of the term, and the international exchange of views that took place worked a transformation in the philosophical thought of the American and European universities.

On June 22, 1936, while on the way to his lecture in the main building of the University of Vienna, Schlick was fatally wounded by a deranged student. The motives for this act have never been fully clarified. The assailant had been under psychiatric observation for some time because of a previous attempt on Schlick's life. With the death of Schlick, the meetings of the Vienna circle came to a sudden end. The Austrian Ministry of Education, for its part, now embarked on a reactionary cultural policy that barred representatives of scientific, analytic philosophy from all official chairs in the universities. With few exceptions, the participants in the Vienna circle immigrated to England and America. The rigorous scientific requirements of the Vienna philosophy met with widespread sympathy in the West and in Poland and Scandinavia; as a result, philosophy as the "logic of knowledge" experienced a fruitful further development abroad.

In Austria, however, the philosophical movement initiated by Schlick encountered the uncompromising hostility of the state authorities. After the interruption caused by World War II, all the official chairs in the Austrian universities were systematically filled by speculative philosophers generally committed to a theological outlook. Only exceptionally was a representative of scientific philosophy able to qualify as a lecturer. But since lectur-

ers and associate or titular professors, unlike regular professors, are not paid a salary in Austria, the authorities had an effective economic means of compelling the unwanted logical analysts of knowledge to turn elsewhere. In practice, this resulted in a suppression of scientific philosophy that continues to exist to this very day. The necessary consequence of a policy so harmful to science has been a shocking decline in the level of scholarship. Psychologically, the only explanation for this reactionary course of isolating research from the rigorous demands of modern scientific philosophy is the fear that logico-mathematical or empirical scientific analysis might endanger some ideological position. In support of this view is the fact that the eastern European countries, which profess a diametrically opposed ideology, also keep Viennese logical positivism away from their chairs of learning out of the same medieval anxiety that prevails in Austria.

CRITIQUE OF KANTIANISM

In his early work *Raum und Zeit in der gegenwärtigen Physik* (1917), Schlick presented a critical examination of the synthetic a priori character that Kantian transcendental philosophy attributed to propositions about space and time. Methodologically following the work of Poincaré and von Helmholtz, he based his thought primarily on the changes introduced by the theory of relativity into certain of the definitions and principles of classical physics. In conformity with scientific opinion of his time, Immanuel Kant had sought to establish the absolute validity of Newtonian mechanics by means of the theory of transcendental forms of intuition and of understanding. He regarded the presuppositions and basic principles of classical mechanics as necessary truths about empirical reality, that is, as synthetic a priori propositions. This conception had first been shaken by investigations of mathematicians. In consequence, doubt had also arisen regarding the synthetic a priori character of the general laws of physics. The theory of relativity made a final break with the synthetic a priori characterization of the foundations of Newtonian physics. According to relativity theory, statements about physical states (including propositions about physical space and physical time) are, as a consequence of the methods used by the natural sciences, empirical in character. That is, they are synthetic a posteriori propositions. Meanwhile, Poincaré had pointed to the possibility of interpreting general laws of nature, such as statements about physical space, as conventions or analytic propositions. Thus he had made evident the conventional nature of certain steps in the methodology of empirical research.

This systematic critique, confined at first to the foundations of mathematics and the natural sciences, was generalized by Schlick to all the basic problems of human knowledge. It thus became the basis of his philosophy in this initial period. In the *Allgemeine Erkenntnislehre* (1918), he made a critical study of all the propositions to which Kant and his followers had ascribed a synthetic a priori character. Schlick concluded that in all cases these propositions, where precisely formulated as logically necessary truths, are analytic in character; when, on the other hand, they are interpreted as statements with real content, they are empirical or synthetic a posteriori. There are no synthetic a priori propositions. Later, in his examination of foundational theories in logic and mathematics and of David Hilbert's formalism in particular, Schlick conceded that the possibility of synthetic a priori propositions in the realm of logico-mathematical forms must be left open. We are in no position to come to a final decision on this question. But even if necessarily valid propositions with content do exist—perhaps in the sense of the mathematical intuitionists—in the domain of logic and mathematics, they could never, Schlick stressed, be interpreted as absolutely valid statements about the empirically real world.

CRITICAL REALISM

Schlick's view was that epistemology, in investigating the criteria of reality, is not obliged in the first instance to ask for absolutely true knowledge of reality. The Cartesian method of doubt leads merely to immediate data of experience, the establishment of which in no way suffices to answer the question "What is real?" Instead of seeking absolutely certain knowledge, we must address ourselves to the systems of propositions by the aid of which science seeks to describe reality, and through a critical examination expunge from these systems all propositions that are demonstrably false. The system that remains will then portray reality just as it is. Here, when we speak of the reality depicted by the natural sciences, we mean those phenomena described by true spatiotemporal propositions. Schlick identified the objects of empirical knowledge, thus characterized, with the Kantian thing-in-itself; he called his own philosophical position "critical realism."

According to Schlick, the method by which we arrive at knowledge of the spatiotemporally ordered world has the feature that whereas the truth of propositions about objective, empirical reality can in principle be established only hypothetically, the falsity of such propositions can in

some cases be demonstrated beyond question. It is interesting to note that Karl Popper's asymmetrical confirmation theory, which did not appear until some twenty years later, likewise attributes a kind of certainty to the disconfirmation of natural laws in contrast with the fact that full verification is unattainable.

In this first period of his philosophical development, Schlick regarded the controversy between idealism and realism as a factual issue which philosophical reflection could resolve. He believed that critical realism provided the correct answer, and he sought to substantiate this answer by a more precise characterization of what is to be understood by empirical knowledge. Knowledge is "knowledge of sameness." Something is cognized as something else, for example, a whale as a mammal. An especially important form of the knowledge of sameness is recognition. Memory outputs over short spans of time are a constitutive element of consciousness. Knowledge of sameness includes not only establishing the sameness or similarity of sense data, memory images, imagined ideas, and the like but also the rediscovery of certain conceptual orderings known, say, from mathematics in the relationships of empirical phenomena. Schlick did not consider the possibility that the study of empirical relationships might lead to the construction of new, hitherto unknown mathematical orders and that in such a case one might arrive at knowledge descriptive of reality that is not knowledge of sameness.

LANGUAGE AND KNOWLEDGE

The problem of knowledge and its criteria had led Schlick to a further question: How is it possible to express knowledge linguistically? Scientific knowledge and insights, whether logico-mathematical or empirical, are presented in the form of sentences of some language. What conditions must be satisfied by these combinations of linguistic signs if they are to count as analytic or empirical sentences? In this earlier period Schlick's answer was the following: The languages employed in the sciences are designed to make possible the construction of unambiguous expressions that can be true or false. But this property of language presupposes the choice and establishment of rules according to which the linguistic signs are to be employed and to be strung out into expressions and sentences. If in using a language one does not heed the logical and linguistic rules set up for it, sign combinations will occur which, although they may appear on the surface to be sentences with a subject and a predicate, actually violate the rules for combining signs. Consequently, they have no meaning and cannot be either true or false.

Applying this notion to philosophy, Schlick held that the theses of metaphysical systems are just such sequences of signs put together in a way that violates the logical rules of language. For this reason metaphysics is to be denied the status of scientific knowledge. But why does metaphysics disregard the logical rules of scientific languages in its linguistic formulations? Schlick thought the reason lay in the fact that whereas metaphysics endeavors to know reality, it does not seek to know the relations between the magnitudes characterizing states of affairs but strives to obtain knowledge of the content of phenomena. However, according to Schlick, only relations can be the object of knowledge—relations that reproduce the order of the phenomena and which include particulars on the number, sameness, similarity, and succession of the empirical data, as well as functional connections between measured quantities. The content of phenomena cannot be grasped by means of ordering relations, which are all that are at the disposal of the understanding. In Schlick's opinion, it is only through an intuitive, emotional experience that we can become acquainted with the actual content of reality. Metaphysics desires to know the "content" of real things, and it therefore finds itself compelled to use expressions from scientific languages in a manner contrary to the rules. For this reason the theses of metaphysics cannot have the character of meaningful propositions.

Schlick arrived at these views under the influence of the writings of Bertrand Russell and Hilbert, both of whom had by this time extensively treated the logical and linguistic foundations of mathematics. They clearly held that in mathematics questions about the logical and linguistic conditions for unambiguous statements must be put with special precision and exactness, but that these questions also affect the foundations of all scientific language systems and hence of scientific knowledge in general. Schlick was the first person to draw, on the basis of these insights into the foundations of logic and mathematics, consequences for epistemology as a whole and to undertake, by logical and linguistic means, the demarcation of a boundary between science and metaphysics.

PHILOSOPHY AND REALITY

During his teaching career in Vienna, Schlick subjected the philosophical views he had published before 1922 to a fundamental reexamination. Influenced by Wittgenstein and Carnap, he no longer saw the task of philosophy as the acquisition of knowledge. Instead, philosophy,

through the application of logical analysis to the concepts, propositions, and methods of the separate sciences, should aim at reaching an understanding of knowledge as found in the individual disciplines and of its presuppositions. Schlick no longer treated realism and idealism as factually contradictory theses but, rather, as alternative ways of speaking; at most, one could ask which permits a simpler, more easily understood way of talking about the world of experience and about purely conceptual relationships. But if realism and idealism are interpreted as statements about something that exists, the realism-idealism antithesis becomes a "pseudo problem" to which neither a true nor a false answer can be given.

This conception was carried over by Schlick to certain problems in the foundations of physics. In his essay "Die Kausalität in der gegenwärtigen Physik" (1931, reprinted in *Gesammelte Aufsätze* and in *Gesetz, Kausalität und Wahrscheinlichkeit*), he cited the answer given by Werner Heisenberg when he was asked to what extent particles are real or unreal. Heisenberg had replied that whether or not one wished to label particles as really existing was simply a matter of taste (*Die physikalischen Prinzipien der Quantentheorie*, Leipzig, 1930, p. 15). In the systems of propositions that constitute physics, we speak only about the data of observation and the regularities they display, or we construct hypotheses and predictions about the occurrence of observable phenomena. Whether the terms *real* and *unreal* are applied to the observational data, to the hypothetical constituents, or to any other elements of the theories is, so far as the content of the system of propositions is concerned, of no consequence at all. Descriptions in terms of "real" and "unreal" can be omitted without any loss of asserted content. Whether one wishes to make use of these terms is merely a matter of convenience and simplicity in expression.

PHILOSOPHICAL METHOD

Schlick generalized his analysis of modes of speech and ways of formulating questions into a philosophical method. Viewed from his new epistemological standpoint, numerous questions, especially in philosophy, turn out to be anchored in ordinary or scientific forms of speech, or in forms artificially created by metaphysics. The first step in Schlick's method of analyzing knowledge consists in finding out the logical and linguistic rules governing the use of the expressions that occur in the problems, propositions, and forms of speech under study. Such a logical and syntactical critique may show that a certain expression, ordinarily assumed to have an unambiguous meaning, is being applied in accordance with different rules in different contexts and therefore is being used in different senses. A striking example is the concept of space. For a long time only one meaning was attributed to it, and the assumption was that the term *space* as employed in mathematics, physics, and psychology has the same meaning. The logical critique of language reveals that mathematical geometries represent analytic systems of relations, whereas physical space is described by means of a system of empirical laws that have as their content the order schema of possible positions and motions of physical bodies. Empirical sentences with different content describe the geometrical and metrical properties of psychological spaces—visual space, auditory space, tactile space, and the like. Similarly, in the case of such terms as *real, ideal, actual*, and *imaginary*, syntactical analysis yields different meanings corresponding to the different rules that govern the use of these expressions on various occasions. Failure to notice such differences of meaning often gives rise to philosophical problems which are then regarded as insoluble.

Thus the first step in the logical analysis of knowledge is to ascertain the rules for the linguistic use of the expressions under consideration. The second step is to study what meaning is to be ascribed to these expressions in a given complex of questions or system of propositions. Schlick called this the "interpretation" of the expressions, concepts, propositions, questions, or theories. If, for example, the first step in the analysis has shown that the word *real* is used in several senses, then the interpretation must determine which particular meaning the word has in, for instance, the sentence "Only that is real which is immediately experienced," or in the sentence "The real is that which leaves traces behind," or "The real is that which can be described by means of conjugate measured quantities." The connection between the two steps in the method is manifest: The clarification of the possible meanings of an expression must precede the interpretation of it in a given context. According to Schlick, the understanding gained through interpretation is the insight for which philosophy strives.

Schlick applied his philosophical method, among other things, to the physical concepts of causality and energy and to the principles of causality and of the conservation of energy, which were still regarded as synthetic a priori propositions. Interpretation requires that in the case of "universally valid" sentences one must always ask whether one can conceive of conditions under which these sentences would have to be regarded as false. If they can be so regarded, then the empirical character of the sentences in question has been recognized. Schlick was

able to specify circumstances whose empirical confirmation is conceivable and under which both the principle of causality and the principle of the conservation of energy (as they are used within physics) would be termed invalid. Accordingly, he expressed the view—at a time when physicists were not yet of this opinion—that the two principles admitted of empirical testing. Subsequent research in physics has confirmed this view. At the same time, Schlick recognized that the concepts of causality and energy can also be defined in such a way that the principles of causality and of the conservation of energy become analytic sentences. It is this possibility that conventionalism exploits when it declares that general forms of laws are absolutely valid by convention. In a further application of his method, Schlick subjected Hans Driesch's vitalism and the general propositions both of psychology and of Husserl's phenomenology to an analytical critique. He arrived at the general conclusion that if the expressions these theories contain are precisely and properly clarified, the sentences in question take on either an analytic or an empirical character, but they never at one and the same time express synthetic and a priori propositions.

One criterion of meaning Schlick used in his analytical procedure was the criterion of verification that Schlick and others attributed to Wittgenstein. By this criterion, general laws of nature can have no significant content because they are not verifiable (or, as it is usually put, are not fully verifiable). This problem gave rise to wide-ranging discussions that went far beyond the Vienna circle. Essentially, Schlick supported Wittgenstein's view that natural laws are not themselves propositions but are to be understood as directives regarding the kind of sentences to be constructed in order to describe or predict individual cases of empirical phenomena. Directives cannot be true or false, so that on this interpretation the verification criterion is not applicable to the laws of nature. On several occasions Schlick characterized this interpretation of natural laws as not entirely satisfactory. But he did not find the opportunity for a definitive exposition of his own position.

PRESUPPOSITIONS AND CONFIRMATION PROCEDURES

Schlick replied to certain criticisms of the philosophy of the Vienna circle. Doubt was expressed that the criteria of the analysis of knowledge are sufficient for distinguishing between analytic and empirical sentences or for drawing a boundary between metaphysics and the individual sciences. Extreme skeptics even questioned the possibility of making such sharp distinctions at all. One argument used by critics concerned the presuppositions that are required whenever one attempts to specify the conditions for determining unambiguously the meaning of concepts and propositions or for deciding unambiguously the truth of analytic and empirical sentences. These presuppositions evade any formal characterization or any determination of their validity, and consequently they have a metaphysical character. Even if these ineluctable presuppositions are limited to the minimal performances of memory necessary for recognizing in a subsequent moment what meaning we have previously assigned to a given expression, the knowledge by recollection we thus presuppose is intuitive in kind and as impossible to check as the theses of metaphysics. Because of these problematical presuppositions, the logical positivist distinctions between analytic and empirical propositions and between scientific and metaphysical propositions cannot possess any validity.

Schlick analyzed these criticisms of recollections that cannot be checked but yet must be presupposed if consciousness, language, thought, and knowledge are to exist. The real problem of the logic of knowledge, he argued, consists in the fact that despite the inexact presuppositions of our methods of knowledge, we nevertheless do obtain exact scientific knowledge. It is wrong to conclude that because the recollections presupposed are unanalyzable and intuitive, the formal logico-mathematical derivations, concept formations, and principles or the empirical criteria of meaning and judgment are inaccurate. The exactness of scientific methods is anchored in proof procedures that guarantee an undeniable advance of knowledge in all the sciences. These procedures distinguish exact scientific knowledge from unverifiable metaphysical speculation. There are no such confirmation procedures for metaphysics, nor does it permit the application of scientific (logical or empirical) criteria of confirmation to its theses and methods. Consequently, in metaphysics there is no such thing as progress of knowledge. Thus the decisive criterion of exactness for the sciences is the advance in knowledge that can be gained through the process of testing, a criterion not satisfied by the speculative methods of metaphysics.

ETHICS AND VALUE THEORY

Schlick also applied the method of the analysis of knowledge to problems of ethics and the theory of value. He concluded that the a priori arguments for absolute values do not fulfill the logical criteria of meaning. Only the value-ascribing forms of behavior actually found among

people, relative assignments of relative values, can be taken as the basis for ethical and other value systems. In Schlick's view, this sort of value analysis leads to a new kind of empirical foundation for eudaemonism. In his *Fragen der Ethik*, Schlick offered as the fundamental principle of an ethics so based the maxim "Increase your happiness" (*Mehre deine Glückseligkeit*).

Schlick's ethics has been widely criticized as superficial, on the ground that there can be morally objectionable happiness. To understand it correctly, one must take into account how he characterized the happiness that one should strive to increase. By happiness he meant the quiet, joyous assent that accompanies our actions when we carry out for its own sake some activity springing from our talents. This is the kind of activity that is to be evaluated as ethically worthwhile behavior. The joy in such activity resembles the joy of a child at play, and it should be regarded generally as the criterion for emotional and intellectual youthfulness. This youthfulness is not tied to physical age. Anyone who has found the activity proper to himself, and has thus experienced this quiet, joyous happiness, has realized the highest attainable ethical goal and will keep his youthfulness throughout his entire life. On this basis, Schlick rejected all varieties of ethical rigorism, including the Kantian system. No ethical worth can be attributed to actions undertaken from a mere sense of duty when such actions inspire only distaste and annoyance both beforehand and afterward. On the contrary, acting out of a sense of duty is ethically valuable only if a quiet satisfaction accompanies the action. Moral value, Schlick used to emphasize, attaches only to vital action; the sign of life is youthfulness, but we are young only when we act from joy. When the quiet, inner joyous assent accompanies our action, we fulfill the requirements of the highest principles of ethical value.

See also Critical Realism; Kantian Ethics; Logical Positivism; Neo-Kantianism.

Bibliography

WORKS BY SCHLICK

Books

Raum und Zeit in der gegenwärtigen Physik. Zur Einführung in das Verständnis der Relativitats- und Gravitationstheorie. Berlin: Springer, 1917; 2nd ed., 1919. Translated by Henry L. Brose as *Space and Time in Contemporary Physics: An Introduction to the Theory of Relativity and Gravitation*. Oxford: Oxford University Press, 1920.

Allgemeine Erkenntnislehre. Berlin: Springer, 1918; 2nd ed., 1925. Translated as *General Theory of Knowledge*. New York: Springer, 1974.

Vom Sinn des Lebens. Berlin, 1927.

Fragen der Ethik. Vienna: Springer, 1930. Translated by David Rynin as *Problems of Ethics*. New York: Prentice Hall, 1939.

The Future of Philosophy. Publications in Philosophy of the College of the Pacific, No. 1. Stockton, CA, 1932. Reprinted in *Gesammelte Aufsätze*.

Gesammelte Aufsätze 1926–36 (Collected papers 1926–1936). Vienna: Gerold, 1938.

Gesetz, Kausalität und Wahrscheinlichkeit (Law, causality and probability). Vienna: Gerold, 1948.

Grundzüge der Naturphilosophie, edited by W. Hollitscher and J. Rauscher. Vienna: Gerold, 1948. Posthumous papers. Translated by Amethe Smeaton as *Philosophy of Nature*. New York: Philosophical Library, 1949.

Aphorismen. Vienna, 1962.

Schlick also edited, with Paul Hertz, the *Schriften zur Erkenntnistheorie* of Hermann von Helmholtz. Berlin: Springer, 1921.

Articles

"Naturphilosophie." In *Lehrbuch der Philosophie*, edited by Max Dessoir, Vol. II, *Die Philosophie in ihren Einzelgebieten*. Berlin: Ullstein, 1925.

"Gibt es ein materiales Apriori?" *Wissenschaftlicher Jahresbericht der philosophischen Gesellschaft an der Universität zu Wien für das Vereinsjahr 1930/31*. Translated by Wilfrid Sellars as "Is There a Factual A Priori?," in *Readings in Philosophical Analysis*, edited by Herbert Feigl and Wilfrid Sellars. New York: Appleton-Century-Crofts, 1949.

"Die Wende der Philosophie." *Erkenntnis* 1 (1930–1931): 4–11. Reprinted in *Gesammelte Aufsätze*. Translated by David Rynin as "The Turning Point in Philosophy," in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

"Die Kausalität in der gegenwärtigen Physik." *Die Naturwissenschaften* 19 (1931). Reprinted in *Gesammelte Aufsätze* and *Gesetz, Kausalität und Wahrscheinlichkeit*. Translated by David Rynin as "Causality in Contemporary Physics." *British Journal for the Philosophy of Science* 12 (1961): 177–193 and 281–298.

"Positivismus und Realismus." *Erkenntnis* 3 (1932–1933): 1–31. Reprinted in *Gesammelte Aufsätze*. Translated by David Rynin as "Positivism and Realism." In *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

"Philosophie und Naturwissenschaft." *Erkenntnis* 4 (1934): 379–396.

"Über das Fundament der Erkenntnis." *Erkenntnis* 4 (1934): 79–99. Reprinted in *Gesammelte Aufsätze*, Vienna: Gerald, 1935. Translated by David Rynin as "The Foundation of Knowledge," in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

"Facts and Propositions." *Analysis* 2 (1935). Reprinted in *Philosophy and Analysis*, edited by Margaret MacDonald. Oxford, 1954.

"Unanswerable Questions?" *Philosopher* 13 (1935). Reprinted in *Gesammelte Aufsätze*. Vienna: Gerald, 1938.

"Gesetz und Wahrscheinlichkeit." *Actes du Congrès International de Philosophie Scientifique, Paris 1935*. Actualités scientifiques et industrielles, no. 391, 46–57. Paris, 1936. Reprinted in *Gesammelte Aufsätze* (Vienna: Gerald,

1938), and in *Gesetz, Kausalität und Wahrscheinlichkeit* (Vienna: Gerald, 1948).

"Sind die Naturgesetze Konventionen?" *Actes du Congrès International de Philosophie Scientifique, Paris 1935*. Actualités scientifiques et industrielles, no. 391, 8–17. Paris, 1936. Reprinted in *Gesammelte Aufsätze* (Vienna: Gerald, 1938), and in *Gesetz, Kausalität und Wahrscheinlichkeit* (Vienna: Gerald, 1948). Translated by Herbert Feigl and May Brodbeck as "Are Natural Laws Conventions?," in *Readings in the Philosophy of Science*, edited by Herbert Feigl and May Brodbeck. New York: Appleton-Century-Crofts, 1953.

"Meaning and Verification." *Philosophical Review* 45 (1936). Reprinted in *Gesammelte Aufsätze* (Vienna: Gerald, 1938), and in Feigl and Sellars, *Readings in Philosophical Analysis* (New York: Appleton-Century-Crafts, 1949).

"Quantentheorie und Erkennbarkeit der Natur." *Erkenntnis* 6 (1936–1937): 317–326. Reprinted in *Gesammelte Aufsätze* (Vienna: Gerald, 1938), and in *Gesetz, Kausalität und Wahrscheinlichkeit* (Vienna: Gerald, 1948).

"Form and Content: An Introduction to Philosophical Thinking. Three Lectures, Delivered in the University of London in Nov. 1932." In *Gesammelte Aufsätze*.

WORKS ON SCHLICK

Feigl, Herbert. "Moritz Schlick." *Erkenntnis* 7 (1937–1939): 393–419.

Popper, Karl. *The Logic of Scientific Discovery*. Translated by the author, with the assistance of Julius Freed and Lan Freed. London: Hutchinson, 1959.

Reichenbach, Hans. "Moritz Schlick." *Erkenntnis* 6 (1936): 141–142.

Rynin, David. "Remarks on M. Schlick's Essay 'Positivism and Realism.'" *Synthese* 1 (1948–1949).

OTHER RECOMMENDED WORKS

Gadol, Eugene T., ed. *Rationality and Science: A Memorial Volume for Moritz Schlick in Celebration of the Centennial of His Birth*. Wien; New York: Springer-Verlag, 1982.

Haller, Rudolf. *Schlick und Neurath—Ein Symposion: Beitr‰ge zum Internationalen philosophischen Symposion aus Anlass der 100. Wiederkehr der Geburtstage von Moritz Schlick (14.4.1882–22.6.1936) und Otto Neurath (10.12.1882 - 22.12.1945), Wien, 16.-20. Juni 1982*. Amsterdam: Rodopi, 1982.

Helmholtz, Hermann von. *Epistemological Writings: The Paul Hertz/Moritz Schlick Centenary Edition of 1921 with Notes and Commentary by the Editors*. Dordrecht, Holland; Boston: D. Reidel, 1977.

Lewis, Joia A. *Positivism and Realism in the Writings of Moritz Schlick*. PhD diss., Indiana University, 1989.

McGuinness, Brian. *Moritz Schlick*. Dordrecht; Boston: D. Reidel, 1985.

McGuinness, Brian. *Zur¸ck zu Schlick: Eine Neubewertung von Werk und Wirkung*. Wien: Hˆlder-Pichler-Tempsky, 1985.

Sarkar, Sahotra, ed. *Logical Empiricism at Its Peak: Schlick, Carnap, and Neurath*. New York: Garland, 1996.

Schlick, Moritz. *Causality in Everyday Life and in Recent Science*. New York: Johnson Reprint, 1969.

Schlick, Moritz. *General Theory of Knowledge:*. Wien; New York: Springer-Verlag, 1974.

Schlick, Moritz. *Philosophical Papers*. Edited by Henk L. Mulder and Barbara F. B. van de Velde-Schlick. Dordrecht; Boston: D. Reidel, 1979.

Schlick, Moritz. *The Problems of Philosophy in Their Interconnection: Winter Semester Lectures, 1933–34*. Edited by Henk L. Mulder, Anne J. Kox, and Rainer Hegselmann. Dordrecht; Boston: D. Reidel, 1987.

Schlick, Moritz, and Bernd Philippi. *Philosophische Logik*. Frankfurt am Main: Suhrkamp, 1986.

Waismann, Friedrich. *Ethics and the Will: Essays*. Edited by Brian McGuinness and Joachim Schulte. Dordrecht; Boston: Kluwer Academic Publishers, 1994.

Waismann, Friedrich. *Wittgenstein und der Wiener Kreis*. Oxford: B. Blackwell, 1967.

Béla Juhos (1967)
Translated by Albert E. Blumberg
Bibliography updated by Michael J. Farmer (2005)

SCHOLASTICISM

See *Augustinianism; Averroism; Medieval Philosophy; Ockhamism; Scientia Media and Molinism; Scotism; Thomism*

SCHOLZ, HEINRICH

(1884–1956)

Heinrich Scholz, the German theologian and logician, was born in Berlin. He professed an outspoken Platonism based on a profound knowledge of the history of metaphysics and of the logical works of Gottfried Wilhelm Leibniz, Bernard Bolzano, and Gottlob Frege. Scholz identified philosophy, in its original Platonic sense as the striving for universal knowledge, with the study of the foundations of mathematics and science. Thus, in *Was ist Philosophie?* (1940; *Mathesis Universalis,* pp. 341–387) he concluded, from Plato's demand for knowledge of geometry and a mathematical astronomy, that the axiomatic method is required for universal knowledge. He regarded mathematical logic as developed by Leibniz, Bolzano, Frege, Bertrand Russell, and others as the "*epochale Gestalt*" of *metaphysica generalis.* He opposed formalism in logic because it failed to provide for the semantics of formal languages, and he opposed constructivism because of its arbitrary anthropocentric limitations of logic.

Scholz's devotion to logic arose from a concern with metaphysics in theology. He studied theology at Berlin and philosophy at Erlangen, receiving a doctorate in philosophy from Erlangen with a dissertation on Friedrich Schleiermacher. He held the chair of systematic theology and philosophy of religion at Breslau from 1917 to 1919,

and then a chair of philosophy at Kiel. In his main systematic theological work, *Religionsphilosophie* (Berlin, 1921), he rejected subjective and existential foundations for religion. God is a transsubjective datum whose being is independent of any "leap of faith"; otherwise truth would be irrelevant to religion: "nothing remains but either to give up the solution to the problem of truth or to enter upon an entirely new course" (*Mathesis Universalis,* p. 13). By a "lucky accident," the discovery of A. N. Whitehead and Russell's *Principia Mathematica* in the library at Kiel, Scholz found his new course. From 1923 to 1928 he immersed himself in the study of logic, mathematics, and physics, and of their histories. His thoughts on metaphysics were galvanized, and he developed an enthusiasm for logical calculi rare even among mathematicians; it infused his later lectures and doubtless alienated those readers in Germany who were not quite convinced of the need to analyze Plato and other classical metaphysicians logically.

In 1929, his metamorphosis into a logician complete, Scholz assumed a chair of philosophy at Münster, which was transferred to the mathematical faculty in 1943 when he founded the Institut für mathematische Logik und Grundlagenforschung. This institute was inspired by the Warsaw school under Jan Łukasiewicz (whom Scholz later rescued from a Nazi concentration camp). But Scholz did not renounce theology. In "Das theologische Element im Beruf des logistischen Logikers" (1935; *Mathesis Universalis,* pp. 324–340) he likened his motives for undertaking *Grundlagenforschung* to the motives of an Augustinian theologian in search of illumination from the eternal forms. He undertook logical investigations of Anselm's ontological argument ("Der Anselmische Gottesbewies," 1950; *Mathesis Universalis,* pp. 62–74) and of Augustine's arithmetical proof ("Der Gottesgedanke in der Mathematik," 1950; *Mathesis Universalis,* pp. 293–312).

Scholz wrote one of the first competent histories of logic, *Abriss der Geschichte der Logik* (Berlin, 1921; translated by Kurt F. Leidecker as *Concise History of Logic,* New York, 1961), based on the pioneering studies of Louis Couturat and Łukasiewicz. He exhibited what may be called a coincidence of logic and metaphysics through several works that together constitute in effect the first logically competent history of metaphysics. His "Logik, Grammatik, Metaphysik" (1944; *Mathesis Universalis,* pp. 399–438) discusses metaphysics in Aristotle, Leibniz, and Immanuel Kant. "Die mathematische Logik und die Metaphysik" (*Philosophisches Jahrbuch der Görres-Gesellschaft* 51 [1938]: 257–291), a 1938 lecture intended to convince a meeting of German Thomists of the importance of mathematical logic, discusses scholastic philosophy, Plato, and Aristotle. He discusses the fundamental importance of the axiomatic method for metaphysics in "Die Axiomatik der Alten" (1930; *Mathesis Universalis,* pp. 27–44), on Aristotle's *Posterior Analytics*; in *Was ist Philosophie?*; and in *Die Wissenschaftslehre Bolzanos* (1937; *Mathesis Universalis,* pp. 219–267). Scholz regarded the *mathesis universalis* of René Descartes, Blaise Pascal, and Leibniz as of special importance in the history of metaphysics. He developed Leibniz's metaphysical doctrines of identity and possibility in *Metaphysik als strenge Wissenschaft* (Cologne, 1941), a thorough treatment of the logic of identity, and in *Grundzüge der mathematischen Logik,* written in collaboration with Gisbert Hasenjaeger (Göttingen, 1961). In *Grundzüge,* logical truth is defined as that which is identical throughout all possible worlds. Scholz used this definition to explain the a priori (the pre-Kantian *Transzendentale*): Possible (not necessarily actual) worlds constitute the logical frame for any description of the real world. Scholz's "Einführung in die Kantische Philosophie," a series of lectures given in 1943 and 1944 (*Mathesis Universalis,* pp. 152–218), was the first systematic treatment of Kant's logical, mathematical, and physical doctrines to call upon both mathematical logic and physics. Of particular interest is Scholz's account of how Kant came to reject the *mathesis universalis* because of Christian Wolff's garbled presentation of Leibniz's mathematical philosophy.

Scholz greatly admired the work of the Vienna circle, particularly that of Rudolf Carnap. However, he held that Platonism, especially in the form of classical mathematics, has been more useful to science than positivism, since it permits theoretical constructions more powerful than any offered by positivism. Positivism retards scientific growth. Thus, according to Scholz, modern relativity theory, even though positivistic tendencies helped lay its observational foundation, is Platonist because of its use of classical analysis. According to Scholz, the logic of Frege and Russell was adequate evidence that Platonism is feasible, and Alfred Tarski's noneffective method of proof and his semantic definition of truth proved that Platonism can be given an absolutely rigorous foundation.

Scholz held that competence in metaphysics requires knowledge of mathematical logic, but he failed to convince most German metaphysicians. His works were ignored, and irrationalism exercised virtual hegemony in Germany during the Nazi era. (Even in the United States, his work was mentioned only in the *Journal of Symbolic Logic.*) Scholz saw language being employed as a poorly controlled, quasi-literary means of expression rather than

as a logical tool for grasping objective truth. He therefore engrossed himself in his technical work, the crowning achievement of which was the posthumously published *Grundzüge der mathematischen Logik.* This work deals extensively with the elements of logic; develops propositional logic, quantificational logic, and type-theoretical logic (this last is called "Russell-revised Platonism" because it functions as an ontological foundation for mathematics) in formalized syntactic and semantic metalanguages; and examines the questions of completeness and independence with respect to both effective and noneffective proof methods.

See also Anselm, St.; Aristotle; Augustine, St.; Bolzano, Bernard; Carnap, Rudolf; Couturat, Louis; Descartes, René; Frege, Gottlob; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Łukasiewicz, Jan; Mathematics, Foundations of; Metaphysics, History of; Ontological Argument for the Existence of God; Pascal, Blaise; Plato; Platonism and the Platonic Tradition; Russell, Bertrand Arthur William; Schleiermacher, Friedrich Daniel Ernst; Tarski, Alfred; Thomism; Whitehead, Alfred North; Wolff, Christian.

Bibliography

See *Mathesis Universalis,* edited by H. Hermes, F. Kambartel, and J. Ritter (Basel: B. Schwabe, 1961), a selective anthology with a comprehensive bibliography; *Glaube und Unglaube in der Weltgeschichte* (Leipzig: J. C. Hinrichs'sche Buchhandlung, 1911); *Die Religionsphilosophie des Als-ob* (Leipzig, 1921).

An article on Scholz, accompanied by a photograph, is Hans Hermes, "Heinrich Scholz zum 70. Geburtstage," in *Mathematisch-Physikalische Semesterberichte* n.f. 5 (1955): 165–170.

Eckehart Köhler (1967)

SCHOOL OF QOM, THE

The School of Qom refers to the tradition of theological institutions of Shiʿa learning in Qom, a city in southern Iran. Along with Meshhad in North Eastern Iran and Karbala and Najaf in Iraq, Qom is a major center of Shiʿism, which houses the golden domed shrine of Fatimah, the holy site named for the sister of the eighth Imam who died in 816 in Qom and was buried there. Recently a few speculative theologians of this school proffered the theoretical foundation for a theocracy commonly labeled as an "Islamic Republic"; their views have become a cause célèbre in the Muslim world in a challenge of and a confrontation with the European cultural, economic, and political dominance in many predominately Muslim states. In spite of its political charisma, Qom continues to be the source of research in the scholarship of the Shiʿa philosophical heritage and exports a number of both young and seasoned scholars to the most prestigious European academic centers.

BACKGROUND

(a) Following the teachings of Nāṣir Khosrow (b. 1003–4) and Nāṣir al-Dīn al-Ṭūsī (1201–1274), traditional curricula of the school of Qom and its approach to Islamic studies integrate religious studies with philosophy and mysticism.

(b) In their ethics of self-realization and in their social philosophy, members of the school of Qom focus on philosophies of intentional processes and on analyses of mystical virtues (instead of "the golden mean"); they praise the use of archetypal memory (*dhikr*) and empathetic intimacy (*uns*) with the ultimate being. A key explicit view is the rejection of the Aristotelian depiction of time as an accident and the replacement of substance-event metaphysics with the so-called "process" ontology, expressed in Mullā Ṣadrā's so-called theory of "substantial motion."

(c) Members of the Qom school are actively engaged in an encounter with Europe; they do appreciate the development of European science as a continuation of Islamic sciences, and have mastered the art of application of computer technology to the humanities, such as the scanning of basic Shiʿa literature and internet communications.

MAJOR FIGURES AND CONTRIBUTIONS

A majority of Iranian theologians were educated and taught in Qom. Salient doctrines of four major thinkers of this school follow:

HUSSEIN TABATABA'I (1903–1981). The most prominent thinker of this school is ʿAllameh Seyyed Muhammad Hussein Tabataba'i (hereafter "Tabataba'i"), a scholar of Shiʿa theology and teacher of recent major thinkers of the school of Qom. Although not directly involved in politics, his writings established the school's theocratic agenda with key political implications.

A major aim of human actions is happiness, according to Tabataba'i—a desire that is only partially achievable in societal contexts that are regulated by laws that focus on external interrelations among human beings. Religious beliefs connecting the internal, intentional, and

spiritual bases of persons to the cosmos create an intimacy with the creator that complements the deficiencies of secular laws by providing total fulfillment of persons' needs, balancing tolerance with praxis. According to Tabataba'i, Islamic society goes beyond tolerance in recognizing the religious practices of other peoples of the book (Jews, Christians, and Zoroastrians) who live under the rule of an Islamic state and contribute to it through their taxes. However, the Muslim community is committed to jihad against those who knowingly reject the principle of unity, namely against rebellious Muslims, against enemies of the faith, and against those who transgress against Islam by occupying the Muslim homeland by force. Following Islamic traditional belief, Tabataba'i considers homicide a major sin against all humanity. For pragmatic purposes, truth may be hidden (the principle of *taqiyya*) when the expression of truth endangers the cause of religion. While speculation about religion is normally not recommend, Islam is open to new visions, inviting learned scholars to make innovations (*ijitahd*) in deducing philosophical points from archetypal monotheistic truths.

MURTTAZA MUTAHHARI (1920–1979). Mutahhari's major achievement was the dissemination of a clear, rational justification of the political views of the school of Qom to the Iranian Shiʿa masses, as his books were printed by the tens of thousands and circulated as textbooks in many schools. Although Mutahhari is known primarily for his plan to refute communism, no less well known is his open acceptance of the advancement of European science and his caricature of the claim of supremacy of European secular philosophy. His research focuses on a number of politically important reforms. For example, he criticized the literal interpretation of sacred texts and advocated the rational adoption of religious archetypes to solve contemporary problems; he advocated the education of women; and he was receptive to the progress of science and technology. Significantly, he preferred political action as a specific innovative application of religious precepts carried out under the guidance of an exemplary political leader, and he eternalized the ethos of the martyrdom of Karbala, and in so doing providing an energizing rationale for the Islamic revolution.

ROHALLAH KHOMEINI (1902–1989). More than any member of this school, Khomeini permanently influenced the course of history of the Islamic world, claiming execution of basic Islamic principles such as Prophet Muhammad's agenda in the transformation of persons from a tribal self to a member of a community of the faithful. In this tenor, a salient feature of Khomeini's political theory was his emphasis on how persons need to feel an intimate existential allegiance to Islam's spiritual nature in order to experience their religious societal self. The faithful are guided by the juridical authority in their participation in a revolution that has the following agenda: (a) to create a continuous confrontation with secular nationalism; (b) to challenge European secular capitalistic political and military imperialism that supports European puppet regimes in predominately Muslim countries; (c) to issue directives against the lives of those who transgress against Islam in any place in the world; and (d) to constitute—in the absence of an Imam (spirited leader)—a juridical authority that supercedes the authority of monarchs or even that of the elected president of a country.

PRESENT STATUS

Qom remains the major center of academic Shiʿa research, where, in addition to Islamic studies, both male and female students study the works of philosophers such as Bertrand Russell, Ludwig Wittgenstein, and Martin Heidegger, and where mastery of computer technology applications to the humanities is expected of students.

See also Islamic Philosophy.

Bibliography

Dabashi, Hamid. *Theology of Discontent: The Ideological Foundation of the Islamic Revolution in Iran*. New York: New York University Press, 1993.

Khomeini, Rohollah. *Islam and Revolution*. Translated by Hamid Algar. Berkeley, CA: Mizan Press, 1981.

Misbah Yazdi, Muhammad Taqi. *Philosophical Instructions: An Inquiry to Contemporary Islamic Philosophy*. Translated by M. Legenhausen and A. Sarvdalir. Binghamton, NY: Binghamton University, 1999.

Mutaharri, Mortiza. *Fundamentals of Islamic Thought: God, Man, and Universe*. Translated by R. Campbell. Berkeley, CA: Mizan Press, 1985.

Tabataba'i, ʿAllameh Sayyed Mohamad Hosayn. *Islamic Teachings: An Overview*. Translated by R. Campbell. New York: Moustazafan Foundation, 2000.

Parviz Morewedge (2005)

SCHOPENHAUER, ARTHUR

(1788–1860)

Arthur Schopenhauer was a German philosopher of pessimism who gave the will a leading place in his meta-

physics. He was born in Danzig. His father, a successful businessman of partly Dutch ancestry, was an admirer of Voltaire and was imbued with a keen dislike of absolutist governments. When Danzig surrendered to the Prussians in 1793, the family moved to Hamburg and remained there until the father's death (apparently by suicide) in 1805. Schopenhauer's mother was a novelist who in later years established a salon in Weimar, which brought him into contact with a number of literary figures, including Johann Wolfgang von Goethe. His relations with his mother, however, were bitter and antagonistic and eventually led to a more or less complete estrangement.

EDUCATION

Schopenhauer's early education was somewhat unconventional. He spent two years in France in the charge of a friend of his father, and for another period he accompanied his parents on a prolonged tour of France, England (where he attended school in London for several months), Switzerland, and Austria. After his father's death he was tutored privately in the classics for a time and then entered the University of Göttingen as a medical student, studying, among other subjects, physics, chemistry, and botany. At Göttingen he first read Plato and Immanuel Kant, and the powerful and lasting impression their writings made upon him directed his interests decisively toward philosophy. In consequence he left Göttingen in 1811 for Berlin, which was at that time the chief philosophical center in Germany, and worked there for two years, attending the lectures of Johann Gottlieb Fichte and Friedrich Schleiermacher (both of whom he found profoundly disappointing) and making preparatory notes for a doctoral thesis. When the uprising against Napoleon Bonaparte led to the closing of the university, Schopenhauer, for whom nationalistic sentiment held little appeal, retired to Rudolstadt to write his thesis, subsequently published there in 1813 under the title of *Über die vierfache Wurzel des Satzes vom zureichenden Grunde* (*On the Fourfold Root of the Principle of Sufficient Reason*).

EARLY CAREER

Apart from producing a short book on the perception of color, *Über das Sehn und die Farben* (Leipzig, 1816), which was inspired by a previous essay on the same subject by Goethe, Schopenhauer employed the next four years writing his principal work, *Die Welt als Wille und Vorstellung* (*The World as Will and Idea*). From the very first stages of the composition of this work, Schopenhauer believed that the ideas he was striving to express were of major importance, and when it was published at Leipzig in 1818 (dated 1819), he was confident that its significance would immediately be recognized. In this expectation he was to be quickly disappointed; the scanty reviews his book received were generally tepid in tone, and the number of copies sold was small. Nevertheless, its publication helped him to obtain the post of lecturer at the University of Berlin, where he chose to give lectures at the same hours as G. W. F. Hegel, who was then at the height of his reputation and popularity. From the start, Schopenhauer advertised his opposition to Hegelian conceptions. He spoke of sophists who, having arisen after Kant, "first exhausted the thinking power of their time with barbarous and mysterious speech, then scared it away from philosophy and brought the study into discredit," and he made it clear that he regarded his own mission as one of repairing the damage that had been done. Schopenhauer's lectures, however, were a failure; Hegel's authority was too firmly established to be undermined in this manner, and Schopenhauer's audience dwindled away.

LATER CAREER

Schopenhauer made no further attempt to establish himself academically. From then on he lived a solitary life, profoundly resentful at the lack of the recognition he felt to be his due and confirmed in his opinion that the dominant Hegelian philosophy was the product of a charlatan who, by an artful combination of sophistry and rhetoric, had succeeded in corrupting the intellects of an entire generation. Despite his disappointment, however, Schopenhauer continued to write, producing books that were in effect elaborations and developments of themes already adumbrated in his main work. He published an essay titled *Über den Willen in der Natur* (Frankfurt, 1836); and a volume on ethics and the problem of free will, *Die beiden Grundprobleme der Ethik* (Frankfurt, 1841), which contained the two essays "Über die Freiheit des Willens" (1839) and "Über die Grundlage der Moral" (1840). In 1844 he brought out a second edition of *Die Welt als Wille und Vorstellung*, greatly expanded by the addition of fifty supplementary chapters. He also contemplated translating Kant's *Critique of Pure Reason* into English and David Hume's *Dialogues concerning Natural Religion* (a work he greatly admired) into German. There can be little doubt that he would have performed both of these tasks well, for his knowledge of English was excellent; but unfortunately nothing came of either project. Finally, Schopenhauer published a collection of essays and aphorisms called *Parerga und Paralipomena* (2 vols., Berlin, 1851), and with this work he began to be widely

known. Discussions of his ideas appeared in foreign as well as in German periodicals, and his system was made the subject of lectures in a number of major European universities. By the time of his death in Frankfurt, he had a growing circle of admirers in England, Russia, and the United States, while nearer home the influence of his writings was soon to show itself in the work of such thinkers as Friedrich Nietzsche and Jakob Burckhardt.

CHARACTER

Schopenhauer's personality, which is reflected in much of his writing, was complex and compounded of curiously diverse elements. Although intellectually self-assured to the point of arrogance, he had a brooding, introspective disposition, and he betrayed an extreme susceptibility to irrational fears and anxieties. Thus, he always slept with a loaded pistol near him, and he took compulsive precautions against disease; he once remarked that if nothing alarmed him, he grew alarmed at this very condition—"as if there must still be something of which I am only ignorant for a time." His manner could be truculent and overbearing; as many of his aphorisms make clear, his view of others was colored by a deep suspiciousness and cynicism, and his general outlook on life and existence was unrelievedly pessimistic. Yet this did not prevent him from taking pleasure in many things—art and music, good food and wine, travel, and, despite his notorious essay on the subject, women. And while he detested bores, in company that he found sympathetic he appears to have been a lively and entertaining talker, displaying a sharp, satirical wit.

THE NATURE OF PHILOSOPHICAL THINKING

Schopenhauer's philosophy is best approached from a position that clearly recognizes his indebtedness to Kant, whom he believed to have been indisputably the greatest thinker of modern times. Schopenhauer's chief charge against his own philosophical contemporaries in Germany (Friedrich von Schelling, Fichte, and Hegel)—was that under the pretense of carrying forward and developing Kantian ideas, they had in fact attempted to philosophize in a fashion that Kant himself had ruled out as wholly inadmissible. For if Kant had shown anything, it was that metaphysical speculation in the old "transcendent" sense was useless as a means of achieving knowledge of what lay beyond all human experience. Such knowledge is in principle unattainable, and it followed that any philosopher, whatever his procedure might be, who tried to establish such things as the existence of God and the immortality of the soul was engaged in a hopeless quest.

Rationalist metaphysicians like René Descartes had employed deductive a priori arguments in an endeavor to prove certain fundamental propositions of theology, and Kant had sufficiently exposed the inadequacy of these arguments by a series of devastating refutations. Yet according to Schopenhauer, Kant's strictures had not prevented some of his self-appointed successors from speaking as if they had mysterious access to truths necessarily outside the range of human cognition—a "little window opening on to the supernatural world," as it were. He suggested, too, that writing in this way appeared more expedient to many academic teachers of philosophy than the honest alternative of expounding truthfully and directly the antidogmatic theses contained in the *Critique of Pure Reason*.

While he accepted Kant's reasons for rejecting metaphysical theorizing in the sense described above, Schopenhauer was nevertheless far from wishing to claim that all philosophical speculation concerning the ultimate nature of the world must be deemed illicit and misconceived. The impulse to seek some general interpretation of reality and of the place of human existence within it was too deeply embedded in the human mind to be totally ignored or set aside. Man, Schopenhauer held, is an *animal metaphysicum*, a creature who cannot avoid wondering at the existence of the world and raising questions concerning its fundamental character and significance—questions that empirical science is unable adequately to resolve, for they lie beyond its sphere. Religion, it is true, attempts in its own way to meet this pervasive need, although not in a manner susceptible to rational justification or certification. For the tenets and concepts of religious faiths, whatever those who subscribe to them may believe to the contrary, can never be more than "allegories" or imaginative figures, and treating them as if they represented literal truths about a higher order of things leads straightway to manifest absurdities and contradictions.

By contrast, the concern of philosophical thinking is not with the metaphorical intimation of ideas that are beyond the grasp of the human intellect; rather, such thinking aims at truth *sensu proprio*. It follows, therefore, that any solution of "the riddle of the world" that philosophy purports to provide must not be one that involves overstepping the boundaries within which all human knowledge is set and confined. The determination of exactly where these boundaries lie is accordingly of pri-

mary importance as a preliminary to all philosophical inquiry.

PERCEPTION AND THOUGHT

Schopenhauer's theory of knowledge may be said to start with Kant's distinction between *phenomena* (what appears to a perceiving mind) and *noumena* (things as they are in themselves). In our perceptual consciousness of the world, we are in fact aware of it only as mediated through our sense organs and intellect—a point Schopenhauer expressed by saying that, so conceived, the world is "idea" or "representation" (*Vorstellung*). Moreover, everything that presents itself to us in perception necessarily conforms to a certain formal and categorial framework that underlies and finds expression in all departments of our commonsense and scientific knowledge.

Thus Schopenhauer was at one with Kant in holding that the human mind cannot (as the British empiricists had suggested) be envisaged as a mere passive recipient of sense impressions, but on the contrary plays an essentially active part in shaping and organizing the sensory material. It is the structure of the intellect, comprising "sensibility" and "understanding," which ensures that this sensory material apprises us of a realm of external objective phenomena, spatially and temporally ordered and standing in determinate causal relations both with one another and with ourselves as percipients. Space and time as forms of sensibility, together with causality considered as the sole category of the understanding (here Schopenhauer diverged from Kant), are therefore "subjective in origin," while at the same time they are necessary conditions of our knowledge of the world as idea. According to Schopenhauer, it is also the case that their valid employment is restricted to this sphere; they have no application to anything not given, or that could not be given, in sense experience.

Schopenhauer distinguished a further class of ideas, namely, what he termed "ideas of Reflection," or sometimes "ideas of ideas" (*Vorstellungen von Vorstellungen*). It is in terms of these that we think about and communicate the contents of our phenomenal experience. In other words, they are the general concepts by virtue of which we can classify phenomena according to common features that are of interest or importance to us, forming thereby a conceptual structure or system that may be said to mirror or copy the empirical world. The function of this system is essentially a practical one; it provides a means of memorizing, and generalizing from, our observations of how things behave under varying conditions, and hence of putting to use what we learn from experience.

Schopenhauer insisted, moreover, that this system cannot legitimately be separated from the foundation of empirical reality upon which it is based, and he claimed that concepts and abstract notions that cannot be traced back to experience are comparable to bank notes "issued by a firm which has nothing but other paper obligations to back it with." Consequently, metaphysical theories that pretend to offer an account of the world purely a priori, and that in doing so employ terms or propositions not susceptible to empirical interpretation, are empty of cognitive content; they "move in the air without support." Indeed, such theories often represent no more than the development, by laborious deductive steps, of the implications of a small group of initial axioms or definitions, yielding systems of empty tautologies.

Thus far, Schopenhauer would appear to have placed fairly stringent limits upon the scope of human inquiry. Attempts to transcend these limits by appealing to the resources of deductive reasoning alone are necessarily impossible, since they involve fundamentally wrong ideas concerning the nature of logical inference. These ideas can never provide us with information of which we were not previously cognizant, for such inference merely makes explicit what is already implicitly asserted in the premises from which it proceeds. Equally, there can be no justification for trying to extend the use of nonlogical, formative principles like the principle of causality in order to establish matters of nonempirical fact, after the manner of some earlier metaphysicians. Schopenhauer even accused Kant of inconsistency in this matter, on the ground that he wrote as though the existence of things-in-themselves, which for Kant are by definition incapable of being experienced, could be validly inferred from the phenomenal data, thereby disregarding his own prohibition. Nonetheless, Schopenhauer considered that the Kantian notion of the thing-in-itself remained a fertile one. Properly conceived, it offered the needed clue to the discovery of a legitimate and correct philosophical interpretation of existence.

THE WILL

According to Schopenhauer, it is not true that the thing-in-itself, the noumenal reality that underlies the world of phenomenal appearances, is beyond the range of all possible human experience. To realize this, it is necessary to take account of the facts of self-consciousness, that is, our own intimate knowledge of ourselves. Self-awareness has two distinct aspects. From one point of view, namely, the

standpoint of ordinary perception, I cannot avoid regarding myself as an "object," as much a physical entity as a building or a tree is. In this sense, I necessarily conform to the conditions that constitute the "world as idea" in general; I am a body that occupies space, endures through time, and causally responds to stimuli.

INDIVIDUAL WILL. My inner experience also assures me that I am nevertheless more than "an object among objects," for I do not appear to myself under this aspect alone. I am also aware of myself from within as a self-moving, active being whose overt perceptible behavior directly expresses my will. This inner consciousness that each one of us has of himself as will is primitive and irreducible. Thus, Schopenhauer claimed that the will reveals itself immediately to everyone as the "in-itself" of his own phenomenal being and that the awareness we have of ourselves as will is quite different from the awareness we have of ourselves as body. At the same time, however, he emphatically denied that the operations of a man's will and the movements he makes with his body are two distinct series of events—events of the first kind being thought of as causally productive of events of the second kind. Schopenhauer believed that dualistic conceptions of the relation of will and body, deriving largely from Descartes, had wrought havoc in philosophy, and he argued instead that a man's body is simply the "objectification" of his will as it appears under the conditions of external perception; what I will and what in physical terms I do are one and the same thing, but viewed from different standpoints.

THE WILL IN NATURE. What has just been discussed represents the cornerstone of Schopenhauer's metaphysic. For it was his contention that we should not assume the above distinction between the phenomenal appearance and the thing as it is in itself to apply only insofar as we ourselves are concerned. On the contrary, just as my own phenomenal being and activity is ultimately intelligible as the expression of my inner will, so may the rest of the phenomenal world be understood to share the same fundamental character that we recognize to be ours. Here was the "great extension" of the concept of will whereby Schopenhauer claimed that all phenomena—human and nonhuman, animate and inanimate—might be interpreted in a way that gave the world as a whole a new dimension of significance and that at the same time was not open to the insuperable objections vitiating traditional metaphysical doctrines.

The latter claim may reasonably be doubted. Schopenhauer often displayed considerable perspicacity in detecting errors and inconsistencies in the theories of other philosophers, but he did not always show a comparable critical acumen with regard to his own ideas. Even so, the picture he drew of the world, in accordance with his conception of its inner essence, is not without a certain novelty and horrific fascination, standing as it does at the opposite pole from all those metaphysical systems that have, in one way or another, endeavored to present ultimate reality as if it were the incarnation of rational or moral order.

For Schopenhauer, the real was not the rational (as Hegel, for instance, implied that it was); on the contrary, "will" was for him the name of a nonrational force, a blind, striving power whose operations are without ultimate purpose or design. So portrayed, nature in all its aspects, ranging from the simplest physical structures to the most complex and highly developed organisms, takes on the character of an endless, and in the last analysis meaningless, struggle for existence, in which all is stress, conflict, and tension. The mechanistic models, the rationalistic schemes and constructions, in terms of which we find it useful to try to systematize the phenomenal data for scientific and practical purposes, merely serve to disguise from view the true nature of the underlying reality; the proper task of philosophy lies, not in seeking (as so many previous thinkers had sought) to reinforce these misconceptions by consoling and sophistical arguments, but rather in removing the veil of deception and setting the truth in a clear light.

HUMAN NATURE. As indicated above, Schopenhauer took as the starting point of his theory of the world the nature of man himself, regarded as the embodiment of will. Man is the microcosm in which all that is fundamental to reality as a whole (the macrocosm) may be plainly discerned. And it is in connection with what he wrote about human nature that Schopenhauer's doctrine of the will can perhaps be most profitably considered. For this doctrine, far from being merely an extravagant philosophical fantasy, foreshadows much that was central to the later development of psychological theory; it represents a highly significant contribution with genuinely revolutionary implications.

Will and intellect. What Schopenhauer had to say on the subject of human nature revolved about his conception of the role of the intellect in human behavior. We like to suppose that in principle, everything we do lies within the province of our reason and is subject to our control; only if this is so can we deem ourselves to be truly our own masters. Traditionally, philosophers have given their

support to such beliefs; according to Schopenhauer, however, the situation is quite the reverse. For the will is not, as Descartes and others have taught, a sort of instrument or component of the intellectual faculty, mysteriously controlling our actions from on high by means of independent acts of rational choice. As has already been seen, Schopenhauer argued that will and body are simply the same thing viewed under different aspects, and he further claimed that the intellect, far from being the original source and spring of the will and the master of the body, is in fact no more than the will's servant and appendage. From an epistemological point of view, this governance of the intellect by the will manifests itself in the forms of knowledge under which the world appears to us—for example, as a causally governed system. To see things as causes or effects is to see them in terms of their potential uses, that is, as possible means to the gratification of the will.

Motivation. According to Schopenhauer, however, the primacy of will exhibits itself in a number of other important ways. Thus he gave various illustrations, drawn from everyday experience, of the manner in which we are often quite unaware of the true import and significance of our responses to circumstances and situations. Believing ourselves to be activated by some consideration that we find acceptable on moral or other grounds, we miss the real motive and might well be shocked or embarrassed if we knew it. Although we are inwardly and immediately aware of ourselves *as* will, our own consciously formulated conceptions of what we desire or what we are intending are, in fact, a highly unreliable guide when the question under consideration is *what* we will. Sometimes, indeed, Schopenhauer seems to have been making the extreme claim that conscious acts of choice never really determine behavior at all. He suggested in a number of instances that our conduct is not ultimately decided by resolves intellectually arrived at after weighing the pros and cons of alternative courses of action; the real decision is made by the will below the level of rationally reflective consciousness, the sole role of the intellect being to put before the will the various possibilities that lie open to the agent and to estimate the consequences that would ensue upon their actualization. In this sense, we never really form more than a "conjecture" of what we shall do in the future, although we often take such conjectures for resolves; what we have decided to do becomes finally clear to us only a posteriori, through the deed we perform. As it stands, this doctrine gives rise to obvious difficulties. Some cases doubtless occur that we should be inclined to describe in some such manner as Schopenhauer recommends, but it does not follow that every case of deliberate action can be so characterized. Indeed, it may be claimed against all positions of this sort that it is only in virtue of our knowledge of what it is to act in accordance with consciously formed choices that the explanation of certain actions in terms of secret or concealed determinations of the will becomes intelligible.

Unconscious mental activity. The above-mentioned difficulties do not invalidate Schopenhauer's exceptionally perceptive and shrewd observations regarding much human motivation. These observations retain their importance even if the more bizarre speculations he based upon them are rejected; and Schopenhauer in fact connected them with a wider theory of human nature that, considering the time in which he wrote, manifested an astonishing prescience. According to this theory, the entire perspective in terms of which we are disposed to view our characters and doings is distorted. We customarily think of ourselves as being essentially free and rational agents, whereas in fact the principal sources and springs of our conduct consist in deep-lying tendencies and drives of whose character we are often wholly unaware. "Consciousness," Schopenhauer wrote, "is the mere surface of our mind, of which, as of the earth, we do not know the inside but only the crust," and in consequence we often put entirely false constructions upon the behavior in which these basic impulses are expressed. He suggested, moreover, that the ignorance we display, the rationalizations which in all innocence we provide, may themselves have a motive, although not one we are aware of. Thus, he frequently wrote of the will as preventing the rise to consciousness of thoughts and desires that, if known, would arouse feelings of humiliation, embarrassment, or shame. Another example of the same process is to be found in instances of memory failure. It is not a mere accident that we do not remember certain things, since there may be powerful inducements for us not to do so; events and experiences can be "completely suppressed," becoming for us as if they had never taken place, simply because unconsciously we feel them to be unendurable. And in extreme cases this can lead to a form of insanity, with fantasies and delusions replacing what has thus been extruded from consciousness.

Sexuality. Sigmund Freud himself recognized the similarity between ideas like those above and some of the leading conceptions of psychoanalytical theory. Certainly there are striking parallels, and perhaps most obviously between what Schopenhauer had to say about the sexual instinct and the Freudian account of libido. For instance, Schopenhauer claimed that the sexual urge represents the "focus of the will." Apart from the instinct to survive, it is

the most powerful motive of all and exercises a pervasive influence in every area of human life. Yet despite this, the amount of attention sexuality had received from most philosophers and psychologists had been remarkably small; it is as though a veil had been thrown over it, through which, however, the subject kept showing through. Nevertheless, Schopenhauer was far from extolling the operations of the sexual drive. Although he thought it necessary to expose honestly the stark reality that human beings seek to hide by falsely romanticizing and idealizing their primitive passions, he also made it clear that he considered sexuality to be a source of great mischief and suffering. Thus he referred to it as a "demon" that "strives to pervert, confuse and overthrow everything," and spoke of sexual desires as being inherently incapable of achieving lasting satisfaction; according to Schopenhauer, the end of love is always disillusion. In other words, here, as elsewhere, conformity to the dictates of the will ultimately results in unhappiness, which is the universal condition of human existence.

PESSIMISM AND ANTIRATIONALISM. In sum, Schopenhauer's doctrine of the will constituted, in a variety of ways, a reaction against the then dominant eighteenth-century, or "Enlightenment," conceptions of human nature. He not only rejected the Cartesian belief in the primacy of intellect or reason in man, but also, by implication, repudiated the "mechanistic" model according to which writers like Hume sought to explain human personality and motivation in terms of the combination and association of atomistically conceived impressions and ideas. In place of this model, he substituted one of dynamic drive and function that was oriented toward the biological rather than the physical sciences and that stressed the importance of unconscious rather than conscious mental processes. Furthermore, Schopenhauer's writings represent a complete departure from the strain of optimism that underlay so much eighteenth-century thinking about history and society. Schopenhauer utterly rejected such ideas as the inevitability of human progress and the perfectibility of man and replaced them with a picture of humankind in general as doomed to an eternal round of torment and misery. Radical changes in the social structure, however "scientifically" applied, would solve nothing, for the evil condition of life as we find it is merely the reflection of the aggressive and libidinous urges rooted in our own natures. All that can usefully be employed are certain palliatives in the form of social and legal controls that give the individual minimal protection against the incursions of his neighbors; and with such measures men have long been familiar.

ART AND AESTHETIC EXPERIENCE

The preeminent position that Schopenhauer assigned to art (certainly no other major philosopher has elevated it to a higher status) is not difficult to understand in the light of his general theory. In this theory, our modes of knowledge and understanding, as well as the activities in which we normally engage, are regarded as being determined by the will. Scientific inquiry was the supreme instance of this, since (Schopenhauer believed) its essential function was one of providing, through the discovery of empirical uniformities, practical techniques for satisfying our wants and desires.

THE AESTHETIC ATTITUDE. The artist's concern, however, is not with action, or the possibility of action, at all, but with what Schopenhauer termed "contemplation" or "will-less perception." This type of perception must not be confused with perception of the ordinary everyday kind, wherein things are looked at from the standpoint of practical interest and appear under the aspect of particular phenomenal objects. For it is the mark of aesthetic contemplation that in the enjoyment of artistic experience "we keep the sabbath of the penal servitude of willing"; the world is seen in abstraction from the various aims, desires, and anxieties that accompany our normal apprehension of it, with the result that it presents itself to us in a completely different light.

It is a further consequence of such detachment (and on this point Schopenhauer followed Kant) that all judgments of taste or aesthetic value are disinterested: They cannot have as their basis some titillation of sensual appetite, for instance; nor can they be grounded upon considerations of social utility, or even of moral purpose. To speak of a natural scene or of a work of art or literature as "beautiful" is to judge it in and for itself, and quite outside the framework of cause and consequence within which our ordinary perceptual judgments have their natural place and from which they derive their significance.

THE AESTHETIC OBJECT. The claim that aesthetic awareness presupposes a distinctive attitude of mind and attention is clearly separable from the contention, also advanced by Schopenhauer, that in such awareness the content of our experience is of a radically different kind from that involved in ordinary sense perception. Surprising as it may seem in the light of some of his earlier pronouncements, Schopenhauer held that the subjective conditions that define and universally determine our perception at the everyday level are wholly in abeyance in the case of aesthetic apprehension, and that to this complete

"change in the subject" there is a corresponding change in the object. As aesthetic observers, we are no longer confronted with a multiplicity of individual things and events that are spatiotemporally and causally interrelated, but instead are presented with the "permanent essential forms of the world and all its phenomena," which Schopenhauer termed the "Ideas" (*Ideen*). This conception of fundamental Ideas, which Schopenhauer adapted from Plato to serve the purposes of his own, very different, theory of art, helps us to understand why he regarded art not merely as a kind of knowledge, but as a kind of knowledge vastly superior to any found in the sphere of the natural sciences. In his view, the natural sciences can never do more than discover regularities at the stage of phenomenal appearance, whereas works of genuine art exhibit to the beholder the nature of the archetypal forms of which the particular phenomena of sense perception are necessarily incomplete and inadequate expressions. Artistic productions may, in fact, be said to be the vehicles through which the artist communicates his profound discoveries and insights and thereby enables others to share his vision.

The notion that the proper objects of artistic perception are Platonic Ideas in the sense described above gives rise to obvious objections. It certainly fits somewhat uneasily into Schopenhauer's system insofar as that originally seemed to be based upon the postulate that phenomenal representation and noumenal will between them exhaust the field of possible human knowledge. And quite apart from this, the theory of Ideas raises problems on its own account. It appears paradoxical, for instance, to suggest that a picture of, say, apples in a bowl is not a picture of things of the sort we can all see and touch in the ordinary way, but of something set mysteriously apart from these and situated in a realm beyond the range of normal vision.

Even so, it is at least to Schopenhauer's credit that he recognized some of the difficulties presented by much that we are prone to think and say about artistic portrayals of experience. The concept of perception, for instance, seems to play a significantly different role in the context of aesthetic appraisal and criticism from the role it plays in other contexts. Again, the specific sense in which certain art forms (painting, for example) are concerned with "representing" reality is notoriously difficult to analyze. The claim that the artist sees something literally distinct from what we ordinarily see is, no doubt, hard to defend; on the other hand, the (different) claim that he sees and is able to portray ordinary things in unfamiliar ways, and under fresh and revealing aspects, appears to contain an obvious truth.

Schopenhauer himself never clearly distinguished between these two claims. Theoretically he subscribed to the first, but much that he said in his discussion of concrete cases accords better with the second. Not only did he often stress the particularity of the artist's observation of phenomena; he also suggested that the artist's unique mode of presenting individual objects, scenes, or situations succeeds in illuminating for us whole ranges of our experience to which we have previously been blind. He argued, however, that it would be a mistake to suppose that we can ever convey by verbal description what we learn from our direct acquaintance with particular works of art. For what these works communicate will in the end always elude anything we try to say about them. "The transition from the Idea to the concept," he wrote, "is always a fall."

MUSIC. Schopenhauer thought that all forms of artistic activity—with one important exception—could be understood and explained in terms of his theory of Ideas. The exception was music. Music is not concerned with the representation of phenomena or the fundamental forms that underlie phenomena, but has as its subject the will itself, the nature of which it expresses directly and immediately. Thus, of all the arts, music stands closest to the ultimate reality of things that we all bear within ourselves and speaks "the universal imageless language of the heart." Schopenhauer's ideas, in this instance and in general, produced a deep impression upon Richard Wagner, who in his opera *Tristan und Isolde* tried to realize in musical form the leading conceptions of Schopenhauer's theory of the world. It is a curious irony that Schopenhauer, far from reciprocating Wagner's admiration, spoke of his music with actual distaste.

ETHICS AND MYSTICISM

Although the world, viewed from a purely contemplative standpoint, presents a spectacle that can be aesthetically enjoyed, it does not follow that the operations of the agency which underlies all that we perceive can afford us any kind of moral guidance or solace. On the contrary, the ethical significance of existence lies in its ultimate horror. Unlike many other metaphysicians, Schopenhauer concluded from his system, not that we should gratefully seek to make our lives conform to the pattern implicit in the nature of reality, but rather that true salvation consists in a total rejection of this pattern. The moral worth of individuals lies in their capacity to liberate themselves from the pressures and urges of the rapacious will.

INALTERABILITY OF CHARACTER. It is not altogether easy to see how liberation is possible. Schopenhauer had claimed that human beings, like everything else in nature, are in essence expressions of will. How, then, can they become otherwise? Furthermore, he insisted upon a strictly deterministic interpretation of human character and action, one that makes the type of freedom of choice postulated by traditional libertarian doctrines inconceivable. What a person does is always and necessarily a manifestation of his inner disposition, which remains fixed and is unalterable by any resolutions he may form to be different. The individual discovers what he is really like by observing his behavior over the course of his life. He will find that this behavior conforms to certain invariant patterns of reaction and response, so that if the same circumstances recur, his conduct in the face of them will be the same as it was before. Such consistent behavior patterns are the outward manifestation of the individual noumenal essence, or timeless character, which each man is in himself—a conception Schopenhauer claimed to have derived from Kant's discussion of the foundations of moral responsibility, though the consequences he drew from it were in fact far removed from any drawn by Kant. Nor can some of these consequences be said to have been logically very happy; for instance, Schopenhauer seems to have employed the notion of a man's character so elastically that it ruled out the possibility of any imaginable state of affairs falsifying his thesis concerning its innate and unchangeable nature.

ETHICAL VARIATION. Schopenhauer's claim that a man cannot change his character at will does not, however, commit him to the view that the dispositions of different individuals do not show significant ethical variations. For an explanation of the fact that there are good as well as evil persons in the world, he returned to the fundamental tenets of his metaphysic. It is a feature of the good, as contrasted with the self-centered or egotistical, individual that he comprehends himself and his relations with others from a "higher" standpoint, which enables him to recognize, however obscurely or inarticulately, the common unitary nature shared by all things. Egoism rests upon the assumption that the individual is a self-sufficient unit, to which all else is foreign. But the individual appears to be set apart from his fellows by an impassable gulf only when apprehended in accordance with the spatiotemporal scheme that informs our everyday "will-governed" way of looking at things.

A profounder insight, such as is exhibited intuitively in the behavior of the just and compassionate man who "draws less distinction between himself and others than is usually done," involves awareness of the illusory character of the phenomenal world. Those who possess this awareness no longer see their fellow creatures as alien objects to be overcome or manipulated in pursuit of their own egocentric aims, but rather as "themselves once more," homogeneous with their own being and nature. Thus, in the last analysis, the distinction between virtue and vice has its source in radically different modes of viewing those around us; and this distinction could, Schopenhauer believed, be adequately explicated and justified in the terms provided by his own philosophical system.

DENIAL OF THE WILL. Schopenhauer frequently quoted the Brahman formula, *tat tvam asi* ("that thou art"), when discussing the metaphysical unity of things that underlies the realm of appearance. Indeed, all his writings on ethical and related subjects show affinities with the doctrines advanced in the Upaniṣads and in Buddhist texts—affinities that he freely acknowledged. Like the Indian teachers, he considered all human life to be enmeshed in suffering, and following them, he often used the word *māyā* to refer to the illusory phenomenal world to which, as empirical individuals, we belong. Total release from the enslavement of the will, as compared with the identification of himself with others that is displayed in the conduct of the morally good man, in fact occurs only when a person finally ceases to feel any attachment to earthly things and when all desire to participate in the life of the world completely vanishes. Such an attitude of mind, which Schopenhauer attributed to ascetics and mystics of all times, becomes possible when a man's will "turns and denies itself," and when what in the eyes of ordinary men is the very essence and substance of reality appears to him as "nothing."

But Schopenhauer was insistent that this "turning of the will," which is a highly mysterious process, is not something a man can bring about through his own deliberate volition, since the process involves the complete "abolition" of his previous personality. This "turning of the will" comes to him, as it were, "from outside" and springs from an insight that wholly transcends the will and the world. Such mystical insight, moreover, is necessarily incommunicable and indescribable; all knowledge, including that attainable by philosophy, here reaches its limit, and we are left with only "myths and meaningless words" that express no positive content. "The nature of things before or beyond the world, and consequently beyond the will," Schopenhauer declared at the close of his main work, "is open to no investigation." The end of philosophy is silence.

IMPORTANCE AND INFLUENCE

Schopenhauer's critics have not failed to draw attention to discrepancies and inconsistencies in his system. These certainly exist, and his natural clarity of expression, which contrasts so sharply with the obscure and cloudy terminology favored by his philosophical contemporaries in Germany, makes them comparatively easy to detect. On the other hand, these discrepancies should not be allowed to stand in the way of a proper appreciation of what was important and influential in Schopenhauer's thought. The nineteenth century witnessed a decline in the fascination that achievements in physics and mathematics had previously exercised over philosophy, and there was a tendency in speculative thought to explore new ways of interpreting and conceptualizing human life and experience. In this development Schopenhauer played a central role. Both through his theory of will, with its psychological implications, and also through the new metaphysical status he gave to art, he helped to bring about a profound shift in the intellectual and imaginative climate.

In this connection, the impression made by his ideas upon novelists such as Lev Tolstoy, Joseph Conrad, Marcel Proust, and Thomas Mann is particularly noteworthy. Among philosophers, the impact of Schopenhauer's thought was weaker and certainly never approached that produced by Hegel's writings; while in more recent times, when philosophical speculation in general has been at a discount, he has attracted little interest. Yet such neglect is undeserved, and the significance of his contribution should not be underestimated. He realized more fully than the majority of his contemporaries the implications of the Kantian critique of traditional metaphysics, and some of the things he himself had to say about the nature of a priori knowledge have a strikingly modern ring. Again, it is worth emphasizing his "instrumentalist" view of human thinking, which anticipated William James and the American pragmatist school, and also his highly perceptive attacks upon the Cartesian theory of personality and self-consciousness, which in important respects foreshadowed present-day approaches to problems in the philosophy of mind. (In particular, his theory of the double knowledge we have of ourselves as agents in the world has interesting contemporary analogues.)

Finally, it should be remembered that possibly the greatest philosopher of modern times, Ludwig Wittgenstein, read Schopenhauer and was influenced by him. The extent of this influence appears most clearly in the notebooks Wittgenstein kept during World War I (*Notebooks 1914–1916*, translated by G. E. M. Anscombe, Oxford, 1961), but signs of it are also to be found in the *Tractatus Logico-Philosophicus* (translated by D. F. Pears and B. F. McGuiness, London, 1961), particularly in the sections on ethics and the limits of language in the latter part of the work.

See also Kant, Immanuel; Pessimism and Optimism.

Bibliography

WORKS BY SCHOPENHAUER

German Editions

Sämtliche Werke, edited by Arthur Hübscher. 7 vols. Wiesbaden: F. A. Brockhaus, 1972; 4th edition, 1988.

Der handschriftliche Nachlaß, edited by Arthur Hübscher. 5 vols. Frankfurt-on-Main: Kramer, 1970.

Philosophische Vorlesungen. Aus dem handschriftlichen Nachlaß, edited by Volker Spierling. 4 vols. Munich and Zurich: Piper, 1984–1986.

English Translations

Schopenhauer's Early Fourfold Root. Translated by F. C. White. Aldershot: Avebury, 1997.

On Vision and Colors. Translated by Eric F. J. Payne. Oxford: Berg, 1994.

The World as Will and Representation. Translated by Eric F. J. Payne. 2 vols. New York: Dover, 1969.

On the Will in Nature, edited by David E. Cartwright. Translated by Eric F. J. Payne. New York and Oxford: Berg, 1992.

Prize Essay on the Freedom of the Will, edited by Günter Zöller. Translated by Eric F. J. Payne. Cambridge, U.K.: Cambridge University Press, 1999.

On the Basis of Morality. Translated by Eric F. J. Payne. Providence and Oxford: Berghahn, 1995.

On the Fourfold Root of the Principle of Sufficient Reason. Translated by Eric F. J. Payne. LaSalle, IL: Open Court, 1974.

Parerga and Paralipomena. Translated by Eric F. J. Payne. 2 vols. New York: Dover, 1966.

Manuscript Remains. Translated by Eric F. J. Payne. 4 vols. Oxford: Berg, 1988.

WORKS ON SCHOPENHAUER

Atwell, John E. *Schopenhauer on the Character of the World: The Metaphysics of Will*. Berkeley: University of California Press, 1995.

Atwell, John E. *Schopenhauer: The Human Character*. Philadelphia: Temple University Press, 1990.

Copleston, Frederick. *Arthur Schopenhauer: Philosopher of Pessimism*. London: Barnes and Noble, 1975.

Fox, Michael, ed. *Schopenhauer: His Philosophical Achievement*. Brighton: Harvester Press, 1980.

Gardiner, Patrick. *Schopenhauer*. Middlesex: Penguin Books, 1967; repr. Bristol: Thoemmes Press, 1997.

Hamlyn, D. W. *Schopenhauer*. London: Routledge & Kegan Paul, 1980.

Hübscher, Arthur. *The Philosophy of Arthur Schopenhauer in Its Intellectual Context: Thinker Against the Tide*. Translated by Joachim T. Baer and David E. Cartwright. Lewiston, NY: Edwin Mellen Press, 1989.

Janaway, Christopher, ed. *The Cambridge Companion to Schopenhauer*. Cambridge, U.K.: Cambridge University Press, 1999.

Janaway, Christopher. *Schopenhauer*. Oxford: Oxford University Press, 1994.

Janaway, Christopher. *Self and World in Schopenhauer's Philosophy*. Oxford: Clarendon Press, 1989.

Magee, Bryan. *The Philosophy of Schopenhauer*. Oxford: Clarendon Press, 1983; 2nd ed., 1997.

Malter, Rudolf. *Arthur Schopenhauer. Transzendentalphilosophie und Metaphysik des Willens*. Stuttgart and Bad Cannstatt: Frommann-Holzboog, 1991.

Malter, Rudolf. *Der Eine Gedanke. Hinführung zur Philosophie Arthur Schopenhauers*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1988.

Neymeyr, Barbara. *Ästhetische Autonomie als Abnormität: Kritische Analysen zu Schopenhauers Ästhetik im Horizont seiner Willensmetaphysik*. Berlin: Walter de Gruyter, 1996.

Safranski, Rüdiger. *Schopenhauer and the Wild Years of Philosophy*. Translated by Ewald Osers. London: Weidenfeld and Nicolson, 1989.

Simmel, Georg. *Schopenhauer and Nietzsche*. Translated by Helmut Loiskandl, Deena Weinstein, and Michael Weinstein. Amherst: University of Massachusetts Press, 1986.

Spierling, Volker, ed. *Materialien zu Schopenhauers "Die Welt als Wille und Vorstellung."* Frankfurt-on-Main: Suhrkamp, 1984.

von der Luft, Eric, ed. *Schopenhauer: New Essays in Honor of His 200th Birthday*. Lewiston, NY: Edwin Mellen Press, 1988.

White, F. C. *On Schopenhauer's Fourfold Root of the Principle of Sufficient Reason*. Leiden: E. J. Brill, 1992.

Young, Julian. *Willing and Unwilling: A Study in the Philosophy of Arthur Schopenhauer*. Dordrecht: Martinus Nijhoff, 1987.

Patrick Gardiner (1967)
Bibliography by Günter Zöller (2005)

SCHRÖDINGER, ERWIN
(1887–1961)

Erwin Schrödinger was born in Vienna, Austria. After his years at the Gymnasium, where he was given a strong education in classics and in science, he studied physics and mathematics at the university of Vienna from 1906. His major teachers were the successors of Ludwig Boltzmann: Franz Exner and Fritz Hasenöhrl. Schrödinger's early interest for philosophy is evident in several manuscripts of this period, which contain reflections about Greek and Indian thought and British empiricism. He was then awarded the D. Phil. Degree in 1910 and became assistant experimental physicist in Exner's laboratory in 1911. From this date until 1922, he worked on several subjects, including atmospheric radioactivity, statistical physics, psycho-physics of sensations, general relativity, and atomic physics.

At the end of World War I, in which Schrödinger served as an artillery officer in the Austrian army, he devoted one year to studying philosophy. He wrote down his philosophical reflections later, during the summer of 1925, in an essay that became part one of his book *My View of the World*. After brief appointments in various German universities, he became full professor of theoretical physics at the university of Zurich in 1922. In the autumn of 1925, he formulated wave mechanics, construed as a development and alteration of Louis de Broglie's ideas. In 1926 Schrödinger published classic papers in which he formulated and solved the "Schrödinger equation" and demonstrated the empirical equivalence between wave mechanics and Werner Heisenberg's matrix mechanics.

Schrödinger then succeeded Max Planck at the prestigious chair of theoretical physics of Berlin in 1927. But he left Berlin to go to Oxford in mid-1933, a few months after Hitler's rise to power. The same year, he shared the Nobel Prize for physics with Paul Dirac. During his two years at Oxford, he wrote several important papers about the interpretation of quantum mechanics, presenting for the first time the concept of "entanglement" of states and the "cat paradox." Schrödinger then accepted an appointment at the University of Graz in Austria in 1936. But with the advent of the "Anschluss" in 1938, Schrödinger had to flee once more from the Nazis. After spending some time in Vatican and in Belgium, he received an appointment in 1940 at the Dublin Institute for Advanced Studies. The Institute was founded mainly for him by Eamon de Valera, then prime minister of the Irish Republic. During his stay in Ireland, Schrödinger devoted his work to unified field theories (in Einstein's spirit), to renewed reflections on the interpretation of quantum mechanics, and also to conferences for a broader audience. His well-known books *What Is Life?*, *Nature and the Greeks*, and *Mind and Matter* arose from these conferences. In 1956, Schrödinger returned to Austria, where he retired in 1958. He died in Vienna in January 1961.

PHILOSOPHY OF PHYSICS

The key to Schrödinger's philosophy of physics (especially quantum mechanics) is contained in a letter to Arthur Eddington of March 22, 1940. There, Schrödinger insists that Ernst Mach's radical empiricism and Ludwig Boltzmann's taste for rational "pictures" are not mutually exclusive strategies. He regarded Mach's empiricism as a good guide to *tabula rasa* whenever unwarranted old intellectual constructs hinder a proper understanding of new physical phenomena. But this is only the first step of

research. Boltzmann's urge to picture must be the second step. Indeed, "forming absolutely clear, almost naively clear and detailed "pictures" allows one "to be quite sure of avoiding contradictory assumptions." (*The Interpretation of Quantum Mechanics*, 1995, p. 121).

Schrödinger used both methods. He was clearly inspired by Mach's method when he criticized vehemently the old-fashioned concept of "particle" construed as a small permanent material body. He formulated his criticism as early as 1913, when he first heard of Bohr's model of the atom, and then refined it throughout his career. According to Schrödinger, the concept of an object is constructed out of actual observations complemented with appropriately selected virtual observations. But if the interpolation of arbitrarily numerous virtual observations is not allowed by the most advanced predictive theory, then the very process of construction collapses, and the corresponding object *cannot be said to exist*. For elementary particles, "Observations are to be regarded as discrete, disconnected events. Between them there are gaps which we cannot fill in" (*Science and Humanism*, p. 27). We cannot fill them in according to a trajector pattern, because of Heisenberg's uncertainty relations. But if there is no trajectory, the discrete events cannot be tied up into a spatio-temporal continuant. Therefore, the idea that these scattered events reveal some permanent being is a sort of kinetic *illusion*: particles *do not exist*.

After this Machian preliminary move, however, Schrödinger activated the Boltzmannian side of his philosophy of physics. To him, without a precise picture, scientific thought is threatened with ambiguity. Yet the picture must not be taken as mere mimicry of "things out there." It is nothing more than the most efficient *mental tool* we have, with no ontological implications. This is the status Schrödinger ascribed to his *wave function* in the 1950s, after having apparently held a naively realist belief in the existence of ψ-waves in 1926. His mature view of wave functions was expressed in *Science and Humanism* (p. 40): "We do give a complete description, continuous in space and time … a description of something. But we do not claim that this 'something' is the observed or observable facts; and still less do we claim that we thus describe what nature … really is." Yet the description, or picture, must be taken seriously in view of its epistemological value. Its continuous evolution according to the Schrödinger equation and the entanglement between wave functions must be allowed to develop throughout without any sudden "reduction of the state." The only constraint to be exerted on this picture is that it must have some connection with experimentally observable events. But to secure this connection, it is sufficient to use either a rule about expectation values of observables or Born's probabilistic rule: *no reduction, no "quantum jump," no collapse of the wave packet, is needed*. This is Schrödinger's "solution" (or rather "dissolution") of the *measurement problem*of quantum mechanics.

METAPHYSICS

Schrödinger was usually careful to separate his metaphysics from his scientific work. He held that Western science arose from the act of "objectivation"—the act of withdrawing oneself from the domain under study. By this objectivation, we push aside color, pain, esthetic judgment, and ethical values, and restrict our interest to that which is common to all: numbers and structures. But, Schrödinger argues, there is no real duality between ourselves and the objects we have thus posited. Furthermore, our personal selves are identical with the one all-comprehending universal self. Whereas science is only concerned with the relations between objectified entities, metaphysics ventures to say something about the one that comes before any objectification has taken place. This nondualist conception (which Schrödinger called the "identity theory") was overtly borrowed from the Indian *Advaita Vedânta* and was remarkably similar to Schopenhauer's earlier views. The arguments Schrödinger presents in favor of this view are as follows: (i) The truth of the "identity theory" is somehow directly experienced; (ii) The "identity theory" provides us with a coherent picture of the world as a whole, including the vexing mind-body problem; and (iii) The "identity theory" has a potentially high ethical value, because it cuts egocentrism at its root. The only point of contact between Schrödinger's metaphysics and philosophy of physics is negative. In *Mind and Matter*, Schrödinger sharply criticized Heisenberg's suggestion that quantum mechanics had weakened the Cartesian dichotomy between *res cogitans* and *res extensa*. After all, Schrödinger wrote, "Subject and object are only one. The barrier between them cannot be said to have broken down as a result of recent experience in the physical sciences, for this barrier does not exist."

See also Einstein, Albert; Heisenberg, Werner; Quantum Mechanics.

Bibliography

PRIMARY WORKS

Science and Humanism. Cambridge, U.K.: Cambridge University Press, 1951.

Nature and the Greeks. Cambridge, U.K.: Cambridge University Press, 1954.

Science, Theory, and Man. Mineola, NY: Dover Publications, 1957.

My View of the World. Cambridge, U.K.: Cambridge University Press, 1964.

What Is Life? and *Mind and Matter.* Cambridge, U.K.: Cambridge University Press, 1967.

Gesammelte Abhandlungen: Verlag der österreichischen Akademie der Wissenschafte. Vienna: Friedrich Wievweg & Sohn, 1984.

The Interpretation of Quantum Mechanics, edited by M. Bitbol. Woodbridge, CT: Ox Bow Press, 1995.

SECONDARY WORKS

Bitbol, M. *Schrödinger's Philosophy of Quantum Mechanics.* Dortrecht: Kluwer, 1996.

Moore, W. *Schrödinger: Life and Thought.* Cambridge, U.K.: Cambridge University Press, 1989.

Scott, W. T. *Erwin Schrödinger: An Introduction to his Writings.* Cambridge, MA: University of Massachussetts Press, 1967.

Michel Bitbol (2005)

SCHULTZ, JULIUS

(1862–1936)

Julius Schultz, the German philosopher, dramatist, historian, and philologist, was born in Göttingen. From 1888 until 1927 he taught at a high school in Berlin. Among Schultz's numerous writings dealing with philosophy, the most important are *Die Maschinentheorie des Lebens* (1909) and *Die Philosophie am Scheidewege* (1922).

Schultz's starting point is the question How must we conceive of consciousness, on the one hand, and the object, on the other, if we wish to understand from their combined action the world of phenomena? To answer the psychological part of this question, Schultz first studied the axioms and categories of ordinary and of scientific thinking in order to see what attitude toward the phenomena is forced upon our understanding by its own innermost essence. At the same time he found a solution to the epistemological problem, namely, that if we desire not only to describe the world scientifically but also to understand it uniformly and completely, we must reduce all qualitative differences to quantitative ones. Accordingly, we must interpret the world of sense as a world of motion and explain all the happenings in the world in a mechanistic-dynamistic manner.

In epistemology, Schultz acknowledged special indebtedness to Immanuel Kant and Hans Vaihinger, whose views he interpreted and developed in a psychologistic fashion. His philosophy of nature is characterized by the attempt to outline a thorough and systematic causal-mechanistic worldview. The nucleus of this view is a "machine theory of life," which Schultz developed on a broad scientific basis. The theory explains the phenomena of life with the help of the postulate of "biogenes." These are defined as unobservable molecules of submicroscopic size, which are not themselves alive but which build up the living forms. Schultz conceived of the "biogenes" in such a manner that from their joint action one can understand all the processes of life in their goal-directedness and wholeness, and thus both the forms as well as the functions of organisms. In this biomechanistic conception, organic forms are extremely complicated physicochemical systems. The goal-directed course of living processes arises out of the meaningful arrangement of these systems, and their structure and behavior are explained by strictly causal natural laws, making unnecessary the assumption of immaterial vital forces.

Schultz also contributed to the typology of philosophical thought. He sought to reduce all philosophical standpoints to two basic conceptions of the world and of life, corresponding to two different types of men. The first type pays homage to the value of conservation and prefers purposeful, useful activity; as a thinker, this practical-minded man professes an ethics of duty and believes in progress and in the efficacy of metaphysical forces. The second type prefers the value of formation and as an aesthete or theorist playfully seeks a sympathetic understanding of forms, which he desires to behold in their abundance. He professes an ethics of character, or ethics of the beauty in life, and believes in an eternal recurrence of coming into being and ceasing to be. As an advocate of determinism and causality, he envisages a mechanistic picture of the world in order to understand it in its depth. Schultz himself preferred the second standpoint, which determined his attitude in the philosophy of history. In particular, he took a pessimistic view of the future development of culture. He feared that man would become part of a machine, a socialized worker-ant—organized for common work down to the last detail, but, as in the early ages, without a history.

See also Consciousness; Determinism, A Historical Survey; Kant, Immanuel; Vaihinger, Hans.

Bibliography

Schultz's main philosophical works are *Psychologie der Axiome* (Göttingen, 1899); *Die Bilder von der Materie* (Göttingen, 1905); *Die drei Welten der Erkenntnistheorie* (Göttingen, 1907); *Die Maschinentheorie des Lebens* (Göttingen, 1909; 2nd ed., Leipzig: Meiner, 1929); *Die Philosophie am*

Scheidewege (Leipzig: Meiner, 1922); *Leib und Seele* (Berlin, 1923); *Das Ich und die Physik* (Leipzig: Meiner, 1935).

His chief historical work is *Wandlungen der Seele im Hochmittelalter,* 2 vols. (Breslau, 1936; 2nd ed., 3 vols. Breslau, 1940).

For Schultz's philosophical autobiography see *Die Philosophie der Gegenwart in Selbstdarstellungen,* edited by Raymund Schmidt (Leipzig: Meiner, 1922), Vol. III, pp. 177–198. For a complete biography and bibliography see Erwin Ditz, *Julius Schultz' Maschinentheorie des Lebens,* Vol. XIV in *Studien und Bibliographien zur Gegenwartsphilosophie* (Leipzig: Hirzel, 1935).

Franz Austeda (1967)
Translated by Albert E. Blumberg

SCHULZE, GOTTLOB ERNST

(1761–1833)

Gottlob Ernst Schulze, the skeptic and critic of Kantian philosophy, was born in Heldrungen, Thuringia. He was professor at Wittenberg and Helmstedt and later at Göttingen, where one of his students was Arthur Schopenhauer. His influence is due chiefly to his writings, in which he developed his critical-skeptical position. Schulze's main work, and the one that made him famous, was *Aenesidemus*. In this work, which first appeared anonymously and without the place of publication, Schulze presents objections to the Kantian critique and to K. L. Reinhold's intended vindication of the critical philosophy. Schulze's arguments against the critical philosophy led him to share David Hume's skepticism, of which he gave a concise presentation.

The *Aenesidemus* tries to show that Hume's skepticism has not been refuted by the critical philosophy. However, Schulze's position is not that of absolute skepticism: The validity of formal logic and the principles of identity and contradiction are not subject to doubt. He defined skepticism as the doctrine "that philosophy can establish neither the existence nor the non-existence of things-in-themselves and their qualities. Also the limits of our cognitive capacity cannot be fixed and ascertained on the basis of generally valid principles. … But the reality of presentations and the certitude of mental events immediately given through consciousness no skeptic has ever doubted" (*Aenesidemus*, p. 24). On the other hand, "skepticism does not declare the metaphysical questions to be eternally unanswerable and in principle not liable to a solution" (p. 24). Through progressive development it is possible to approach a solution of the problems concerning the existence or nonexistence of things-in-themselves and the limits of our cognitive capacities.

Thus the possibility of perfecting human cognition so as to attain clarity and certitude in particular metaphysical questions was not denied by Schulze. However, his objection to the critical philosophy was not limited to the question concerning the possibility of progress in metaphysics; he also attempted to show the self-contradictory nature of Immanuel Kant's critical philosophy. The critical philosophy argues that since general and necessary knowledge is possible only through synthetic a priori judgments, such judgments must represent reality. Furthermore, such judgments are possible only on the assumption of a pure capacity of understanding; hence, such a capacity must exist. In interpreting Kant, Reinhold generalized this mode of argumentation, formulating the fundamental principle that the presentation of any object implies the distinction between consciousness of the subject, of the object, and of the relation obtaining between them. From these indispensable components of the presentation Reinhold concluded the reality of corresponding objects.

However, from the fact that presentations always contain the notions of subject, of object, and of their relation to each other it is illegitimate, according to Schulze, to conclude the objective reality of corresponding objects. The transition from thought to being is grounded in ontological thinking, which Kant himself showed to be defective in his criticism of the classic proofs for the existence of God and of dogmatic metaphysics. Since one cannot argue from the conditions of thought to the reality of objects, the problem of philosophy is, according to the critical philosophy, to search for the competence and the legitimacy of our thought to determine objects of reality. The task of the Kantian critique is to show the objective validity of our judgments. However, the indispensable conditions of thought constitute subjective necessity, from which objective validity cannot be derived.

Furthermore, "it is presupposed that each part of human cognition must be grounded in reality as its cause. Without such an assumption the doctrine of the *Critique* concerning the origin of the necessary judgments has no meaning whatsoever" (ibid., pp. 137f.). The conclusion from the necessary judgments in our consciousness as to the reality of objects is based on the principle of causality. Existing objects constitute the causes of our cognition. The category of causality is thus employed with reference to noumena. Also, in the conception of sensibility as a faculty of receptivity, the existence of things-in-them-

selves that have the capacity to affect our sensibility is presupposed. Here again the concept of causality is applied to noumena, while, according to the critical philosophy, causality as a category of understanding is confined to the realm of phenomena. Reinhold's doctrine that things-in-themselves, although not cognizable, are nonetheless thinkable, is untenable. Since the things-in-themselves are thought to be the cause of cognition, they are cognized as having the capacity to affect the knowing and thinking capacity. The thing-in-itself must be cognizable, or it cannot be considered as a cause of cognition.

Likewise, the concept of causality cannot be employed for proving the reality of the subject as a thing-in-itself. Schulze understood the Kantian solution of the question "How are synthetic a priori propositions possible?" as consisting in the derivation of these propositions from the subject as their cause: "The *Critique* derives the necessary synthetic propositions from the subjective mind (*Gemüth*) and its a priori determined cognitive processes … by the application of the principle of causality, which does not harmonize with its own principles delimiting the area of application of the categories" (ibid., pp. 153f.). Moreover, the conclusion from the propositions to the reality of a capacity in the mind does not explain the process of cognition. Nothing is gained by proposing that the perception of the material given is due to a receptive capacity, for a problem is not explained by reducing it to something unknown. The problem of cognition of experience is not solved by a reduction of cognition to a receptive capacity that is no less problematical.

Schulze considered the a priori concepts as existing in time "prior" to the cognition of objects. This account of the a priori concepts as innate ideas and as inherent qualities of the subjective mind is a misunderstanding of the Kantian position that has been common to numerous interpreters of Kant until the present. Schulze thus failed to understand the essence of the critical philosophy, which does not aim at deriving the synthetic propositions from the subject as a thing-in-itself. Kant was not concerned with the psychological process of cognition but with objective cognition, as manifested in the scientific process. The problem is how synthetic a priori propositions are possible in mathematics and science, and not how the human mind as a subject conceives such propositions. The objectivity of the judgments is vouchsafed by the scientific laws determining objects that arise through these laws. This is implied in the Kantian principle of the "possibility of experience." Scientific experience is possible only through synthetic propositions. Since without synthetic propositions there would be no scientific experience at all, their legitimacy is vouchsafed by the function they fulfill for experience. Furthermore, Schulze held the difference between synthetic and analytic propositions was not an objective distinction, and, psychologically considered, it depends on subjective circumstances whether a proposition is synthetic or analytic for a particular individual at a certain moment.

Schulze's criticism of Kant implied the notion of the subject and predicate of the proposition as individually given and fixed entities, so that the synthetic proposition connects elements that can be thought of in themselves. Hence, the concepts of the subject and the predicate must be thought of as separately given, and the question is how their connection can be of a necessary nature. Schulze did not realize that for Kant the concept of the subject arises by its determination through the synthetic proposition. In the proposition "*S* is *P*," *S* is an unknown before its determination through *P*. The investigation of *S* is a "doubt-inquiry process" (John Dewey's expression); *S* acquires determination only through the predicate. Schulze's criticism is thus predicated upon an understanding of the critical philosophy as subjective idealism with the notion of a priori concepts as innate ideas, which leads to dogmatic assumptions concerning the application of the concept of causality to things-in-themselves. The a priori concepts in the critical philosophy are not to be understood as constituent features of the subjective human mind but as creative functions of thought in the process of ordering experience.

Schulze was also critical of Kant's conception of moral theology. He raised objections to the Kantian doctrine of the postulates (God, freedom, immortality) as formulated in the *Kritik der praktischen Vernunft*. From the sense of the moral command in us, the categorical imperative, there can be no conclusion as to the reality of a most perfect being. As ideas of reason, God, freedom, and immortality are endless tasks for human activity, but by the conception of these ideas as postulates their real existence as objects is posited. "The Kantian moral theology postulates *more* than what practical reason demands for the satisfaction of its requirements" (ibid., pp. 440ff.). In his criticism of the postulates Schulze has partly anticipated the neo-Kantianism of the Marburg school. Hermann Cohen, for example, although motivated by different considerations, has pointed out that the regulative ideas of reason do not require the support of the doctrine of postulates.

Schulze's contribution to the development of Kantian idealism consists in his exposing the contradictions

and inconsistencies involved in both dogmatic-realistic and subjective-idealistic interpretations of the critical philosophy, but his attempt at a vindication of Hume's skepticism proved ineffective for further development of Kantian idealism. Philosophical thought took the course not back to Hume but to a more consistent critical idealism eliminating the concept of a thing-in-itself (as in Salomon Maimon) and to speculative idealism as it developed in the post-Kantian metaphysical systems. However, by his valuable criticism of the doctrine of the faculties of the soul Schulze anticipated Johann Friedrich Herbart and influenced Friedrich Eduard Beneke (1798–1854).

According to Schulze, a phenomenon of the life of the soul is not explained by attributing it to a "faculty." Such an attribution does not explain, but merely gives another name to the same thing. The task of psychology as a science is, rather, a detailed description of actual mental occurrences and their systematic classification. By such a method, general concepts of psychological phenomena can be attained; but they should not be attributed to "faculties" of the soul, which is a metaphysical concept.

See also A Priori and A Posteriori; Beneke, Friedrich Eduard; Cohen, Hermann; Dewey, John; Herbart, Johann Friedrich; Hume, David; Innate Ideas; Kant, Immanuel; Maimon, Salomon; Neo-Kantianism; Reinhold, Karl Leonhard; Skepticism.

Bibliography

WORKS BY SCHULZE

Grundriss der philosophischen Wissenschaften. 2 vols. Wittenberg, 1788–1790.

Aenesidemus, oder über die Fundamente der von dem Herrn Professor Reinhold in Jena gelieferten Elementarphilosophie. Nebst einer Vertheidigung des Skepticismus gegen die Anmassungen der Vernunftkritik. 1792. Reedited by the Kantgesellschaft, in the series of rare philosophical works edited by Arthur Liebert. Berlin: Reuther and Reichard, 1911.

Kritik der theoretischen Philosophie. 2 vols. Hamburg, 1801.

Enzyklopädie der philosophischen Wissenschaften. Göttingen, 1814.

Über die menschliche Erkenntnis. Göttingen, 1832.

WORKS ON SCHULZE

Cassirer, Ernst. *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit*. Vol. III. Berlin: Cassirer, 1920.

DiGiovanni, George. "The First Twenty Years of Critique: The Spinoza Connection" In *The Cambridge Companion to Kant*, edited by Paul Guyer. New York: Cambridge University Press, 1992.

Erdmann, Johann Edward. *Versuch einer wissenschaftlichen Darstellung der Geschichte der neueren Philosophie*. Vol. III. Riga and Dorpat, 1848–1853.

Fischer, Kuno. *Geschichte der neueren Philosophie*. 3rd ed., vol. VI. Heidelberg, 1900.

Kroner, Richard. *Von Kant bis Hegel*. 2 vols. Tübingen: Mohr, 1921–1924.

Rosen, Michael. "From Kant to Fichte: A Reply to Franks" In *Transcendental Arguments: Problems and Prospects*, edited by Robert Stern. Oxford: Clarendon Oxford Press: 1999.

Wiegershausen, Heinrich. *Aenesidem-Schulze, der Gegner Kants und seine Bedeutung im Neukantianismus*. Berlin, 1910. This is *Kantstudien*, Supp. No. 17.

Samuel Atlas (1967)
Bibliography updated by Tamra Frei (2005)

SCHUPPE, ERNST JULIUS WILHELM

(1836–1913)

Ernst Julius Wilhelm Schuppe, the German philosopher, was born in Brieg, Silesia. He studied at the universities of Breslau, Bonn, and Berlin, and he took his doctorate at Berlin in 1860. He taught at grammar schools in Silesia and then held a chair of philosophy at the University of Greifswald from 1873 to 1910.

EPISTEMOLOGY

In his main work, *Erkenntnistheoretische Logik* (Bonn, 1878), largely anticipated by his earlier book *Das menschliche Denken* (Berlin, 1870) and summarized in his later *Grundriss der Erkenntnistheorie und Logik* (Berlin, 1894; 2nd ed., Berlin, 1910), Schuppe was concerned with the epistemological bases of knowledge generally and of logic in particular. Schuppe held that a theory of knowledge should avoid hypotheses such as the transcendent reality postulated by realists and metaphysicians, but that it should equally avoid one-sided objective or subjective foundations of knowledge, whether materialist, positivist, or idealist.

In keeping with these requirements, Schuppe developed the notion of conscious immanence (*Immanenz, Bewusstsein, Ich*) in which subject and object form a unity. This immanence of consciousness, or ego, is a fact (*Tatsache*) that is given with certitude and can therefore serve as a starting point for epistemology. Only abstractly can the ego be divided into subject and object; concretely it is a correlation of the two. This is not to say that the object is a psychic entity, but merely that there is no being not related to a subject. To ignore the correlation would

be to incur a contradiction because a supposed unthought entity is nevertheless implied in the thought of the epistemologist.

To account for the distinction and division of consciousness and content (*Inhalt*), and of contents among themselves (subjective elements such as acts are distinct from objects of acts, however much both may have to be considered contents for an abstract subject), Schuppe presented a theory of "common" content: Objective content is a given that can be shared by several, whereas subjective content (sensation, for example) is unique and private. The need for this division led to a theory of consciousness in general (*Bewusstsein überhaupt*) as distinguished from the consciousness of a concrete individual subject. Individuality is based on content not shared by others. Other minds, which are presupposed by the notion of consciousness in general, are known, Schuppe claimed, by inference mediated by one's own body; but he also asserted that they can be regarded as immediately perceived. Schuppe denied the claim that other minds are immanent contents of one's mind—like any other object—as being tantamount to solipsism. Schuppe drew upon the ontic fact of a plurality of minds as a basis for consciousness in general.

Schuppe held that thought is also a "component" of the content of consciousness, along with the sense component; it "accedes" to perceptual data. Accordingly, objects of cognition can be considered as constituted by an interaction of an original given of sense, by itself an abstraction, with performances of thought (*Denkarbeit*). In fact Schuppe came to regard thought as the central function of consciousness: To think is to appropriate content, to receive an impression in its positive determinacy, to fixate it as identical. This primary performance of appropriation is thought-in-general, which is prior to judgment. Schuppe argued that at this stage there is only one datum to be appropriated but that for judgment two contents, subject and predicate, are required. (Here Schuppe was influenced by a grammatical notion of judgment.) Continuing to develop his notion of content, Schuppe introduced an analysis of content in which thought stands for the identification of two contents (an instance of the principle of identity, with the principles of contradiction and limitation as corollaries) and, somewhat surprisingly, for the establishment of causal connection between them. Identity and causality are the categories that constitute objective content. (Here Schuppe was guided by a metagrammatical or transcendental notion of judgment, interpreting the category as the predicate of the unified contents.)

ONTOLOGY

With this basis of transcendental thought, Schuppe's "espistemological" logic was not so much concerned with the "forms" of formal logic as with the establishment of a priori truths about the object of knowledge. Thus the logic constitutes a theory of objects, an ontology. Schuppe analyzed the given into its elements (temporal and spatial determinateness, sense impression) and conceptual moments (genera and species), and distinguished several kinds of union (*Zusammengehörigkeit*) among them. In a transcendental progression Schuppe established number, space region, thing, organism, and artifact; and genera, species, and matter. He avoided any reference to a transcendent cause. Understandably, he presented a coherence theory of truth.

LOGIC

Schuppe sought a transcendental genealogy as a basis for logic. This project involved a certain deviation from the traditional understanding of formal logic. He rejected the isolation in logic of a purely formal realm, denying in fact that purely formal theorems are significant. He regarded propositions as assertions of categorial unification. Logic must be concerned with the realm of material content in which unity is asserted and must examine the various types of union of content, that is, the "real" genera of content, which, in the case of objects of appearance, are grounded in the causal context. This doctrine has ramifications in many areas of logic, for example, in the theory of definition.

Schuppe's theoretical philosophy can be regarded as a doctrine of the constitution of knowledge and its objects by transcendental synthesis. In view of its intuitive starting point and its analysis of given content, however, it seems to be a compromise between a logico-transcendental theory and a theory of reflective intuition. The agency responsible for the grounding of objectively constituted content is both a transcendental principle and an existent consciousness. The normative element of a transcendental theory is merged with the factual basis of a subjective ontology. Schuppe's philosophy thus stands between transcendental critique and ontological philosophy of immanence. Although it leans heavily on Immanuel Kant, it anticipates much of Edmund Husserl's phenomenology and constitutes an example for a theoretical understanding of the interplay of factuality and logico-transcendental thought.

PRACTICAL PHILOSOPHY

Schuppe's *Grundzüge der Ethik und Rechtsphilosophie* (1881; reprinted 1963) offers an independent compromise between a normative position, based on the will as a form of consciousness in general, and a eudaemonistic one, based on pleasure. He also wrote several studies in the philosophy of law, such as *Der Begriff des subjektiven Rechts* (Breslau, 1887), and joined the philosophical discussion concerning the new German civil code (*Das Gewohnheitsrecht,* Breslau, 1890; *Das Recht des Besitzes,* Breslau, 1891).

See also Coherence Theory of Truth; Epistemology; Epistemology, History of; Husserl, Edmund; Kant, Immanuel.

Bibliography

In addition to the works by Schuppe already cited, see *Allgemeine Rechtslehre mit Einschluss der allgemeinen Lehren vom Sein und vom Wissen,* edited and with an introduction by Wilhelm Fuchs (Berlin: Verlag für staatswissenschaften und geschichte, 1936).

For works on Schuppe, see R. Hermann, *Schuppes Lehre vom Denken* (Greifswald, 1894); Gunther Jacoby, *Wilhelm Schuppe,* no. 45, *Greifswalder Universitätsreden* (Greifswald, 1936); L. Kljubowski, *Das Bewusstsein und das Sein bei Wilhelm Schuppe* (Heidelberg, 1912); Rudolf Zocher, *Husserls Phänomenologie und Schuppes Logik* (Munich, 1932).

Klaus Hartmann (1967)

SCHUTZ, ALFRED

(1899–1959)

Alfred Schutz was born in Vienna on April 13, 1899. He studied law and social sciences at the University of Vienna from 1918 until 1921, where he completed a doctorate in law and then continued his studies in the social sciences until 1923. Equally as important for his intellectual development as his studies at the university was his participation in the informal academic life of Vienna, in which he also cultivated his philosophical interests. After completing his academic studies and in addition to his ongoing scholarly activities, Schutz held the full-time job of a bank lawyer—a dual life that lasted until he retired from the bank in 1956. Following the annexation of Austria to the Third Reich, Schutz and his family escaped via Paris to New York. There he became affiliated with the graduate faculty at the New School for Social Research in New York, where he taught from 1943 until 1952 as a lecturer, then afterwards as a full professor of sociology and from 1956 as professor of both sociology and philosophy. He died on May 20, 1959, in New York.

Schutz is regarded as the founder of the phenomenological approach in sociology. Influenced by Max Weber, Henri Bergson, Edmund Husserl, the Austrian School of Economics, and pragmatism, he sought to give a philosophical foundation to interpretative social sciences. As a critical follower of Husserl, he developed his own *mundane phenomenology* of the life-world and its structures, showing how actors produce and understand social reality in everyday interactions and communication.

Schutz begins with Max Weber's view of social reality as a meaningful sociocultural world and shares his concept of meaning-oriented social action, but he criticizes Weber for neglecting to inquire into the constitution of meaning in general. In order to analyze how the meaning attached to action is revealed, Schutz refers to the philosophical concepts developed by Henri Bergson and Edmund Husserl. He adopts the Bergsonian idea of the stream of consciousness, but he later comes to recognize difficulties in Bergson's intuitivism and turns to Husserl's phenomenology.

In 1932 Schutz writes his masterpiece "Der sinnhafte Aufbau der sozialen Welt" where he develops his basic concept of the constitution of the social world and formulated his own phenomenological position. Influenced by pragmatism—which was mediated to him by Henri Bergson, Max Scheler, and William James—in this work he proceeds beyond the realm of consciousness and perception as analyzed by Husserl and considers both human action and interactions as well as acts of consciousness as factors in the constitution of meaning. Leaving behind the transcendental philosophical approach, he develops his own "mundane" phenomenology that analyzes the constitution of a meaningful world within "mundane" social relationships in the everyday world. He adopts the results of the Husserlian analyses of the temporality of consciousness, of the intentional structure of lived experience, as well as of the meaning constitution based on embodiment as preconditions on which the social shaping of experience patterns is based, but he rejects Husserl's assumption that intersubjectivity in the sense of understanding of the others could be a result of the acts of consciousness alone (Schutz 1966). Rather, he shows how the schemes of experience are shaped by influencing (*Wirken*) and by relationships of influence (*Wirkensbeziehung*) which consist of interaction and communication.

Schutz understands communication as a process where two subjective streams of consciousness are coordinated within a relationship of mutual influence and where the meaning of one ego's action consists in the intention to evoke a reaction on the part of the other. Actions have here the function of signs that are mutually indicated and interpreted. Because the final meaning of one person's action is revealed in the reaction of the other and vice versa, communication provides a common stock of shared patterns of interpretation that allows for mutual understanding, even if each of the agents always refers to his or her own schemes of experience. In this concept of understanding based on interaction, Schutz offers his own solution to the problem of intersubjectivity posed by Husserl.

In his later work Schutz (1962, 1964, 1966) determines this communicatively created social reality as the world of everyday life whose typical patterns are taken for granted and represent the intersubjective common core of the reality in which people live. He also discloses further structural characteristics of this everyday core of the life-world: Its typical structure depends considerably on the pragmatic orientation of action selecting the areas where typification processes take place (1962, 1966, 1970). Both typicality and this selection based on *systems of relevance* represent two generative principles of order in the everyday world. This everyday reality is nevertheless not identical to the life-world as a whole. By suspending his pragmatic interest, the agent is able to modify his or her everyday experiences and perceive them as objects of a game, fantasy, art, science, or as a dream. All those modifications represent different provinces of meaning that transcend the everyday world and constitute the *multiple realities* (Schutz 1962) of which the life-world is composed. The different strata of meaning in the life-world are integrated by semiotic systems whose structure allows the contents of one province of meaning to be symbolized by another through appresentation.

By considering communication as a substantial constitutive mechanism of social reality, Schutz (1962) stresses the role of language in this process. On his view, language maintains relevances and typifications unique to specific cultures and to social groups and is thus crucial for the constitution of the life-world as a cultural one (Schutz and Luckmann 1989).

The methodological rule that Schutz derives from his approach is expressed in his *postulate of adequacy* (Schutz 1962, 1964) between everyday and scientific typifications. This postulate holds that higher-order interpretative types employed by social sciences have to be constructed in correspondence to the structure of the everyday typifications (first-order types). Thus the structure of the life-world that guides everyday actions also represents the methodological framework within which the social and cultural sciences have to proceed.

The Schutzian phenomenological approach represents one of the main paradigms in the area of interpretative social and cultural sciences. In philosophy his theory led to a critical assessment of the Husserlian view of intersubjectivity, to conceptions of a worldly phenomenology and theory of the cultural sciences (Embree 1988), and to a philosophy of modern anonymity (Natanson [1962–1995] 1986), as well as to new insights into intercultural hermeneutics (B. Waldenfels 1997, 1998, 1999). It also influenced the philosophy of gender (E. List 1993). In a modified form, Schutz's concept of the life-world was also integrated into the social philosophy of Jürgen Habermas.

See also Phenomenology.

Bibliography

WORKS BY SCHUTZ

Der sinnhafte Aufbau der sozialen Welt. Wien, Austria: Springer, 1932. Translated by George Walsh and Frederick Lehnert as *The Phenomenology of the Social World* (Evanston, IL: Northwestern University Press, 1967).

Collected Papers. 4 vols., edited by Maurice Natanson. The Hague: M. Nijhoff, 1962–1995. Vol. 1, *The Problem of Social Reality* (1962); vol 2, *Studies in Social Theory*, edited by A. Broderson (1964); vol. 3, *Studies in Phenomenological Philosophy*, edited by I. Schutz (1966); vol. 4, [without special title], edited by Helmut Wagner and George Psathas with Fred Kersten (1995).

Reflections on the Problem of Relevance, edited by Richard M. Zaner. New Haven, CT: Yale University Press, 1970.

The Structures of the Life-World. With Thomas Luckmann. Evanston, IL: Northwestern University Press, 1973. Revised edition, 1989.

Life Forms and Meaning Structure. Translated by Helmut Wagner. London: Routledge, 1982.

Werkausgabe. 12 vols. Konstanz, Germany: Un[n]iversitätsverlag Konstanz, 2003–.

WORKS ON SCHUTZ

Barber, Michael D. *The Participating Citizen: A Biography of Alfred Schutz*. Albany: State University of New York Press, 2004.

Embree, Lester, ed. *Worldly Phenomenology: The Continuing Influence of Alfred Schutz on North American Human Science*. Washington, DC: University Press of America, 1988.

List, Elisabeth. *Die Präsenz des Anderen: Theorie und Geschlechterpolitik*. Frankfurt am Main, Germany: Suhrkamp, 1993.

Natanson, Maurice. *Anonymity: A Study in Philosophy of Alfred Schutz* Bloomington: Indiana University Press, 1986.

Srubar, Ilja. *Kosmion: Die Genese der pragmatischen Lebensweltheorie von Alfred Schütz und ihr anthropologischer Hintergrund*. Frankfurt am Main, Germany: Suhrkamp, 1988.

Wagner, Helmut R. *Alfred Schutz: An Intellectual Biography*. Chicago: University of Chicago Press, 1983.

Waldenfels, Bernhard. *Studien zur Phänomenologie des Fremden*. 4 vols. Frankfurt am Main, Germany: Suhrkamp, 1997, 1998, 1999.

Ilja Srubar (2005)

SCIACCA, MICHELE FEDERICO

(1908–1975)

Michele Federico Sciacca was a founder of the Gallarate movement, professor of theoretical philosophy at the University of Genoa, and the founder and editor of the journals *Giornale di metafisica* and *Humanitas*. He started as a historian of ideas, writing important works on Reid (1935), Plato (1939), and St. Augustine (1939); a massive review of Italian thought, *Il XX secolo* (2 vols., Milan, 1941); and a review of contemporary European thought, *La filosofia oggi* (Milan, 1945).

Although Sciacca studied under Antonio Aliotta, his major stimulus came from Giovanni Gentile, from whom Sciacca derived his basic axiom that concrete being must be act, never fact. Sciacca developed this principle in his own fashion under the influence of Plato, St. Augustine, Antonio Rosmini-Serbati, and Maurice Blondel.

Sciacca's position was one of "integralism." The central notion of integralism is interiority, according to which the ground of all forms of being and existence lies in the activity of the subject. Sciacca asserts that the existent, or act, cannot be a fact among facts; its existence resides wholly in its own self-generative actuality. Against existentialism he asserts that the being of the existent cannot be pure possibility or nothingness; it must be being. The whole concern of integralism is to establish the character of the being that the existent is. Sciacca holds this being to be objective interiority, which he delineates in his most original speculative work, *Interiorità oggettiva* (Milan, 1951). Interiority is the positing by the existent of itself as act. So defined, it cannot be conceived as purely immanent, in the manner of Gentile. It must posit itself with reference to a transcendent and objective reality and define itself within this horizon. The basic structural principle of interiority is truth, or the subject's affirmation of the ground of its existence in the very act of existing. The immanent ground of the subject and of all existence is a transcendent being, not abstract but more concrete and existentially real than the subject—God. In affirming the existence of God, the subject also affirms its own being, the innermost character of its own act of existing.

Sciacca's basic insight is thus that the being of the subject cannot be mere possibility, nothingness, or facticity but must be act; that this act is the affirmation of its own actuality through the affirmation of its transcendent ground; and that the absolute existent is present in concrete human existence. It is this presence of the Absolute that establishes the human existent as a person. In *Morte ed immortalità* (Brescia, 1954), Sciacca holds that the affirmation of God within human existence that constitutes the human subject cannot be a merely transitory relationship and that immortality is therefore the logical extension of interior objectivity.

See also Gentile, Giovanni.

Bibliography

Works on Sciacca include the following: Sante Alberghi, "Sul pensiero di M. F. Sciacca," in *Giornale critico della filosofia italiana*, Series III, 4 (1950): 841–893; A. Robert Caponigri, "The Existence of God: Proof from Truth," in *Saggi in onore di M. F. Sciacca* (Milan, 1960), pp. 53–78; Cornelio Fabro, "La verità integrale dell'uomo integrale," in *Divus Thomas*, Series III, 27 (1950): 511–519; Augusto Guzzo, Romeo Crippa, and A. Arata, eds., *Michele Federico Sciacca* (Turin: Edizioni di "Filosofia," 1951), which contains a complete bibliography of Sciacca's writings and of writings about him; and "Note critiche sul neo-spiritualismo," in *Giornale di metafisica* 9 (2) (1956), which contains contributions by various authors.

WORKS BY SCIACCA

L'interiorità oggettiva. 4th ed. Milano: Marzorati, 1965 .

La libertà e il tempo. Milano: Marzorati-editore, 1965.

La filosofia e la scienza nel loro sviluppo storico. Per i licei scientifici. Roma, Cremonese, 1965–1967.

Dallo spiritualismo critico allo spiritualismo cristiano (1939–1951). Milano: Marzorati, 1965.

Lettere dalla campagna e scritti affini. 2nd ed. Milano: Mazarati, 1966.

Opere complete. Saggi, scritti inediti, opere minori, recensioni. Milano: Marzorati 1966.

Lettere dalla compagna e scritti affini. 2nd ed. Milano: Marzorati, 1966.

Platone. 2nd ed. Milano: Marzorati, 1967.

Il Pensiero moderno (secoli XVII–XVIII). Milano: Marzorati, 1968.

Morte e immortalità. 3rd ed. Milano: Marzorati, 1968.

Studi sulla filosofia moderna. 4th ed. Milano: Marzorati, 1968.

Filosofia e antifilosofia. Milano: Marzorati, 1968.

La filosofia morale di Antonio Rosmini. 5th ed. Milano: C. Marzorati, 1968.

Il problema di Dio e della religione nella filosofia attuale. 5th ed. Milano: Marzorati, 1968 .
In spirito e verità: pensieri e meditazioni. 7th ed. Milano: Marzorati, 1969.
Pagine di critica letteraria. (1931–1935). Milano: Marzorati, 1969 .
Autobiografia e risposte al Giornale de' letterati d'Italia. Milano: Marzorati, 1969.
Gli arieti contro la verticale. Milano: Marzorati, 1969.
La Chie sa e la civiltà moderna. 2nd ed. Milano: Marzorati, 1969.
Cultura e anticultura. Torino: Borla, 1969 .
La filosofia, oggi. Milano: Marzorati, 1970.
L'oscuramento dell'intelligenza. Milano: Marzorati, 1970.
Interpretazioni Rosminiane. 3rd ed. Milano: Marzorati, 1971.
Dialogo con Maurizio Blondel. 2nd ed. Milano: Marzorati, 1971.
Studi sulla filosofia antica. Con un'appendice sulla filosofia medioevale. 2nd ed. Milano. Marzorati, 1971.
Il chisciottismo tragico di Unamuno e altre pagine spagnole. Milano: Marzorati, 1971.
Ontologia triadica e trinitaria. Stresa: Centro internazionale di studi rosminiani, 1972.
Pagine di pedagogia e di didattica. Milano: Marzorati, 1972.
Il pensiero italiano nell'età del risorgimento. 3rd ed. Milano: Marzorati, 173.
Pascal. 8th ed. Milano: Marzorati, 1973.
L'ora di Cristo. 2nd ed. Milano: Marzorati, 1973.
L'estetismo, Kierkegaard, Pirandello. Milano: Marzorati, 1974.
Pirandello, I giganti della montagna. Stresa: Centro internazionale di studi rosminiani, 1974.
Figure e problemi del pensiero contemporaneo. Milano: Marzorati, 1974
Prospettiva sulla metafisica di San Tommaso. Milano: Città Nuova editrice, 1975.
Il magnifico oggi. Roma: Città nuova, 1976.
La filosofia di M.F. Sciacca: saggi. Bologna: Tip. editrice compositori, 1976.
Il pensiero della rinascenza e della riforma: protestantesimo e riforma cattolica. Milano: Marzorati, 1977.
L'interiorità oggettiva. Palermo: L'epos, 1989.
Il chisciottismo tragico di Unamuno. Palermo: L'epos, 1989.
Pascal. Palermo: L'epos, 1989.
Dialogo con Maurizio Blondel. Palermo: L'epos, 1989.
La filosofia morale di Antonio Rosmini. Palermo: L'epos, 1990.
Platone. Palermo: L'epos, 1990.
L'estetismo, Kierkegaard, Pirandello. Palermo: L'epos, 1990.
Atto ed essere. Palermo: L'epos, 1991.
Sant'Agostino. Palazzo Speciale: Epos, 1991.
Prospettiva sulla metafisica di S. Tommaso. Palermo: L'epos, 1991.
La filosofia italiana nel secolo XX. Palermo: L'epos, 1997–.
Filosofia e antifilosofia. Palermo: L'epos, 1998.
Come si vince a Waterloo. Palermo: L'epos, 1999
La casa del pane. Palermo: L'epos, 1999.
La libertà e il tempo. Palermo: L'epos, 1999.
Gli arieti contro la verticale. Palermo: L'epos, 1999.
L'uomo questo squilibrato. Palermo: L'epos, 2000.
L'oscuramento dell'intelligenza. Palermo: L'epos, 2000.
Filosofia e metafisica. Palermo: L'epos, 2002.
I principi della metafisica rosminiana. Palermo: L'epos, 2003.

A. Robert Caponigri (1967)
Bibliography updated by Michael J. Farmer (2005)

SCIENCE, RESEARCH ETHICS OF

The idea that ethics is important in scientific research is not new. In 1830 Charles Babbage (1791–1871) admonished British scientists for engaging in dishonest research. In 1912 researchers discovered the fossil skull of a missing link between humans and apes at the Piltdown quarry in Sussex, England. After four decades of controversy, several scientists proved that the skull was a hoax.

At the beginning of World War II, prominent physicists believed that it was their moral obligation to help defeat Nazi Germany. Albert Einstein wrote a letter to President Franklin D. Roosevelt (1882–1945) urging the United States to develop the atomic bomb. J. Robert Oppenheimer (1904–1967) directed the Manhattan Project, a $1 billion effort to build the first nuclear weapons. After the United States dropped two bombs on Japan in the summer of 1945, many scientists who worked on the bomb also led the Atoms for Peace movement, which helped to establish the International Atomic Energy Commission. During the Nuremberg Trials (1949–1949), the international community adopted a code of conduct for human experimentation, the Nuremberg Code (1947), in response to the horrific experiments on human subjects conducted by Nazi researchers at Nuremberg.

In 1961 Rachel Louise Carson (1907–1964) alerted the public to the toxic effects of the pesticide DDT on animal species and helped to launch the environmentalist movement. In 1966, Henry Knowles Beecher (1904–1976) published an article describing twenty-two ethically problematic medical experiments, including the Tuskegee Syphilis Study. In this experiment, which took place from 1932 to 1972 in Tuskegee, Alabama, researchers withheld medical treatment from African American subjects with advanced syphilis, even after an effective treatment, penicillin, became available in the 1940s. The study continued until the media brought it to the public's attention in 1972, prompting Congress to hold hearings on biomedical research and adopt new laws pertaining to research on human subjects. In 1975, philosopher Peter Singer published a book that challenged the moral legitimacy of most experiments on animals and helped to energize the growing animal rights

movement. In that same year scientists held a conference at Asilomar, California, on the risks of genetically engineered microorganisms.

Interest in the ethics of research increased dramatically in the mid-1980s due to at least two factors. First, there were many highly publicized allegations of data fabrication (making up data), falsification (changing data), and other unethical activities in federally funded research. Second, the academic research enterprise became much more commercialized due to changes in intellectual property laws and the expansion of the pharmaceutical and biotechnology industries. In the 1980s, patent offices began awarding patents on many different biological products and processes, such as DNA, cell lines, and genetically modified organisms. The U.S. Congress also passed several laws encouraging the transfer of technology from the public to the private sector. In response to these changes in the law, universities began aggressively pursuing and protecting intellectual property. Academic researchers also took a greater interest in intellectual property and in forming start-up companies to commercialize new inventions and discoveries. The pharmaceutical industry increased its spending on research and development, and the biotechnology industry, which emerged in the late 1970s following the development of gene sequencing, splicing, and copying techniques, did the same. By the beginning of the twenty-first century, private industry accounted for more than sixty percent of all research and development expenditures in the United States.

As research became more commercialized, financial ties between academic and government scientists and private companies, for example, ownership of stock or patents and gifts or consulting arrangements, increased. These financial interests created a conflict of loyalties for scientists and universities and threatened the objectivity and trustworthiness of research. Scientists, ethicists, and journalists presented evidence that some researchers and private companies were biasing data analysis and interpretation, research design, and publication practices to produce results favorable to those companies. Financial interests (and pressures) in research also were linked to fabrication, falsification, and other ethical problems.

From the late-1980s to the early twenty-first century, many different organizations took steps to promote ethics in research. The National Institutes of Health (NIH) mandated that all students on Public Health Service training grants and all intramural researchers receive instruction in responsible conduct of research. Universities incorporated ethics education into the graduate curriculum to meet NIH requirements and to minimize the risk of the legal liability and public embarrassment from ethical misconduct in research. The NIH and the National Science Foundation adopted a common definition of *research misconduct* as well as policies and procedures for investigating and adjudicating misconduct allegations. The National Academy of Sciences (NAS) published several reports concerning ethics in research. Many different professional organizations and scientific journals adopted or revised codes of conduct in research.

Research ethics has become multidisciplinary field of scholarship, education, and policy, encompassing the humanities, the social sciences, and the natural sciences. Some of the key topics in the field include: the foundations of research ethics; ethical decision-making in research; recording, storing, and sharing data; honesty and objectivity in research; scientific misconduct; authorship and publication; collaboration; mentoring; intellectual property; ownership of research materials; conflicts of interest; diversity in science; research on human and animal subjects; research in genetics and biotechnology; scientific freedom; social responsibility in research; research funding; and legal and regulatory aspects of research.

See also Einstein, Albert; Philosophy of Social Sciences; Singer, Peter.

Bibliography

"About ORI." Historical Background. Office of Research Integrity. Available from www.ori.dhhs.gov/.

Angell, Marcia. *The Truth about the Drug Companies: How They Deceive Us and What to Do about It*. New York: Random House, 2004.

Babbage, Charles. *Reflections on the Decline of Science in England, and Some of Its Causes* (1830). New York: Augustus Kelly, 1970.

Beecher, Henry. "Ethics and Clinical Research." *New England Journal of Medicine* 274 (1966): 1354–60.

Broad, William, and Nicholas Wade. *Betrayers of the Truth: Fraud and Deceit in the Halls of Science*. New York, Simon and Schuster, 1983.

Brody, Baruch. *The Ethics of Biomedical Research: An International Perspective*. New York: Oxford University Press, 1998.

Bulger, Ruth Ellen, Elizabeth Heitman, and Stanley Joel Reiser, eds. *The Ethical Dimensions of the Biological Sciences*. New York: Cambridge University Press, 2002.

Carson, Rachel. *Silent Spring*. Boston: Houghton Mifflin, 1962.

Dresser, Rebecca. *When Science Offers Salvation: Patient Advocacy and Research Ethics*. New York: Oxford University Press, 2001.

Emanuel, Ezekial, Robert Crouch, John Arras, et al. *Ethical and Regulatory Aspects of Clinical Research: Readings and*

Commentary. Baltimore, MD: Johns Hopkins University Press, 2003.

"Guidelines for the Conduct of Research in the Intramural Research Programs at NIH." Available from www.nih.org.

Hart, Lynette A., ed. *Responsible Conduct with Animals in Research*. New York: Oxford University Press, 1998.

Institute of Medicine. *Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct*. Washington, DC: National Academy Press, 2002.

Jones, Anne, and Faith McLellan, eds. *Ethical Issues in Biomedical Publication*. Baltimore, MD: Johns Hopkins University Press, 2000.

Jones, James. *Bad Blood: The Tuskegee Syphilis Experiment*. New York: Free Press, 1981.

Kevles, Daniel. *The Baltimore Case: A Trial of Politics, Science, and Character*. New York: Norton, 1998.

Krimsky, Sheldon. *Science in the Private Interest: Has the Lure of Profits Corrupted Biomedical Research?* Lanham, MD: Rowman & Littlefield, 2003.

LaFollette, Hugh, and Niall Shanks. *Brute Science: Dilemmas of Animal Experimentation*. New York: Routledge, 1996.

Levine, Robert. *Ethics and the Regulation of Clinical Research*. 2nd ed. New Haven, CT: Yale University Press, 1988.

Macrina, Francis L. *Scientific Integrity: An Introductory Text with Cases*. Washington, DC: American Society for Microbiology Press, 2000.

Millar, Ronald. *The Piltdown Men*. New York: St. Martin's Press, 1972.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine. *Advisor, Teacher, Role Model, Friend: On Being a Mentor to Students in Science and Engineering*. Washington, DC: National Academy Press, 1997.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine. *On Being a Scientist: Responsible Conduct in Research*. Washington, DC: National Academy Press, 1995.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine. *Responsible Science, Vol. 1: Ensuring the Integrity of the Research Process*. Washington, DC: National Academy Press, 1992.

"NIH Policy on Instruction in the Responsible Conduct of Research." National Institutes of Health. Available from www1.od.nih.gov/.

Nuremberg Code. Available at: www.hhs.gov/ohrp/references/nurcode.htm.

Penslar, Robin Levin (ed.). *Research Ethics: Cases and Materials*. Bloomington: Indiana University Press, 1995.

"Requirement for the Instruction in Responsible Conduct of Research in National Research Service Award Training Grants." *National Institutes of Health Guide* 21 (43) (1992). Available from http://grants.nih.gov/.

Resnik, David B. "Social Epistemology and the Ethics of Research." *Studies in the History and Philosophy of Science* 27 (1996): 565–86.

Resnik, David. *The Ethics of Research*. New York: Routledge, 1998.

Resnik, David B. "From Baltimore to Bell Labs: Reflections on Two Decades of Debate about Scientific Misconduct." *Accountability in Research* 10 (2003): 123–135.

Shamoo, Adil E. and David B. Resnik. *Responsible Conduct of Research*. New York: Oxford University Press, 2003.

Shrader-Frechette, Kristin. *Ethics of Scientific Research*. Lanham, MD: Rowman & Littlefield, 1994.

Sigma Xi. *Honor in Science*. New Haven, CT: Sigma Xi, The Scientific Research Society, 1986.

Sigma Xi. *The Responsible Researcher: Paths and Pitfalls*. Research Triangle Park, NC: Sigma Xi, The Scientific Research Society, 1999.

Singer, Peter. *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: Random House, 1975.

Steneck, Nicholas. *ORI Introduction to Responsible Conduct of Research*. Bethesda, MD: Office of Research Integrity, Department of Health and Human Services, 2004.

"Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication, 2004." International Committee of Medical Journal Editors. Available from www.icmje.org/.

Veatch, Robert M. *The Patient as Partner: A Theory of Human Experimentation Ethics*. Bloomington: Indiana University Press, 1987.

Whitbeck, Caroline. *Ethics in Engineering Practice and Research*. New York: Cambridge University Press, 1998.

David B. Resnik (2005)

SCIENCE AND PSEUDOSCIENCE

Since the rise of modern science in the sixteenth and seventeenth centuries, attempts to adjudicate the difference between science and pseudoscience have always been more than an exercise in academic debate. The religious, political, and social implications of how science is defined, who defines it, and who and what is left out of the definition has been a contentious one. Today, the term *pseudoscience* is often employed by those in the scientific community to disparage claims to scientific credibility that, in fact, lack evidence or fail to employ the methods of science. *Pseudoscience* is only one term used to contrast with science; others include, on the neutral side, *nonscience*, *protoscience*, *prescience*, *frontiers science*, and *borderlands science*; and on the pejorative side, *pathological science*, *junk science*, *voodoo science*, *crackpot science*, and *bad science*.

With the ascendancy of science in the seventeenth century other knowledge traditions began to employ the empirical methods of science to gain respectability. The study of demons, witches, and spirits, for example, took a decidedly empirical turn in the early modern period, out of religious concerns that atheism might ascend to social respectability along with science. One observer wrote, "Atheists abound in these days and witchcraft is called

into question. If neither possession nor witchcraft (contrary to what has been so long generally and confidently affirmed), why should we think that there are devils? If no devils, no God" (Walker, pp. 71–72). By the nineteenth century the study of such quasi-scientific ideas as phrenology, mesmerism, and spiritualism was organized through scientific societies, such as the Society for Psychical Research, founded in London in 1882, and the American Society for Psychical Research, founded in Boston in 1885, both of which included as active members prominent scientists.

THE DEMARCATION PROBLEM

In the twentieth century the philosophy of science developed into a viable academic discipline, out of which grew attempts to delimit science and nonscience traditions. In *The Logic of Scientific Discovery*, for example, the philosopher of science Karl Raimund Popper identified what he called "the problem of demarcation," that is "the problem of finding a criterion which would enable us to distinguish between the empirical sciences on the one hand, and mathematics and logic as well as 'metaphysical' systems on the other" (1934, p. 27). Most scientists and philosophers use induction as the criterion of demarcation—if one reasons from particular observations or singular statements to universal theories or general conclusions, then one is doing empirical science. Popper's thesis was that induction does not actually provide empirical proof—"no matter how many instances of white swans we may have observed, this does not justify the conclusion that all swans are white" (p. 34)—and that, de facto, scientists actually reason deductively, from the universal and general to the singular and particular. But in rejecting induction as the preferred (by others) criterion of demarcation between science and nonscience, Popper was concerned that his emphasis on deduction would lead to an inevitable fuzziness of the boundary line. If a scientific theory can never actually be proven, then is science no different from other knowledge disciplines?

Popper's solution to the problem of demarcation was the criterion of *falsifiability*. Theories are "never empirically verifiable," but if they are falsifiable then they belong in the domain of empirical science. "In other words: I shall not require of a scientific system that it shall be capable of being singled out, once and for all, in a positive sense; but I shall require that its logical form shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical scientific system to be refuted by experience" (1934, p. 70). The theory of evolution, for example, has been accused by creationists as being nonscientific because no one was there to see it happen and biologists cannot observe it in the laboratory because it takes too long. But, in fact, by Popper's criterion of falsifiability, the theory of evolution would be doomed to the trash heap of bad science if, say, human fossil remains turned up in the same geological bedding planes as 300-million-year-old trilobites. No such falsification of evolution has ever been found, and although by Popper's criterion this does not mean that the theory has been proven absolutely, it does mean that it has yet to be falsified, thus placing it firmly in the camp of solid empirical science.

SCIENCE DEFENDED, SCIENCE DEFINED

The evolution-creationism controversy, in fact, has provided both scientific and legal forms of demarcation between science and pseudoscience. It is one thing for academic scientists and philosophers to debate the definition of science; it is another matter when the U.S. Supreme Court weighs in on the issue. Because evolution could not be excluded from public school science classrooms in the early twentieth century, and because the teaching of religious tenets was deemed unconstitutional in a number of state trials in the middle of the twentieth century, in the latter part of the century creationists began to call their doctrines creation-science. Since academic openness calls for a balanced treatment of competing ideas, they argued, creation-science should be taught side by side with evolution-science. In 1982 creationists succeeded in getting passed the Louisiana Balanced Treatment for Creation-Science and Evolution Science Act. In 1985 the law was struck down in the Federal Court of Louisiana, a decision that was appealed to the U.S. Court of Appeals for the Fifth Circuit. In 1986 the U.S. Supreme Court agreed to hear the case, leading to the publication of a remarkable document that clearly and succinctly adjudicated (literally in this case) the difference between science and pseudoscience.

The document was an amicus curiae brief submitted to the court on behalf of seventy-two Nobel laureates in science, seventeen state academies of science, and seven other scientific organizations. The amicus brief begins by offering a general definition: "Science is devoted to formulating and testing naturalistic explanations for natural phenomena. It is a process for systematically collecting and recording data about the physical world, then categorizing and studying the collected data in an effort to infer the principles of nature that best explain the observed

phenomena." Next, the scientific method is discussed, beginning with the collection of "facts," the data of the world. "The grist for the mill of scientific inquiry is an ever increasing body of observations that give information about underlying 'facts.' Facts are the properties of natural phenomena. The scientific method involves the rigorous, methodical testing of principles that might present a naturalistic explanation for those facts" (1986, p. 23).

Based on well-established facts, testable hypotheses are formed. The process of testing "leads scientists to accord a special dignity to those hypotheses that accumulate substantial observational or experimental support." This "special dignity" is called a "theory" that, when it "explains a large and diverse body of facts" is considered "robust" and if it "consistently predicts new phenomena that are subsequently observed" it is "reliable." Facts and theories are not to be used interchangeably or in relation to one another as more or less true. Facts are the world's data. Theories are explanatory ideas about those facts. "An explanatory principle is not to be confused with the data it seeks to explain." Constructs and other nontestable statements are not a part of science. "An explanatory principle that by its nature cannot be tested is outside the realm of science" (pp. 23–24).

It follows from the nature of scientific method that no explanatory principles in science are final. "Even the most robust and reliable theory … is tentative. A scientific theory is forever subject to reexamination and—as in the case of Ptolemaic astronomy—may ultimately be rejected after centuries of viability." Scientists encounter uncertainty as a regular and natural part of their work. "In an ideal world, every science course would include repeated reminders that each theory presented to explain our observations of the universe carries this qualification: 'as far as we know now, from examining the evidence available to us today'" (1986, p. 24). Science also seeks only naturalistic explanations for phenomena. "Science is not equipped to evaluate supernatural explanations for our observations; without passing judgment on the truth or falsity of supernatural explanations, science leaves their consideration to the domain of religious faith" (p. 23). According to the amicus any body of knowledge accumulated within the guidelines previously described is considered scientific and suitable for public school science education; and any body of knowledge not accumulated within these guidelines is not considered scientific.

On June 19, 1987, the U.S. Supreme Court, by a vote of 7 to 2, held that the Louisiana Act "is facially invalid as violative of the Establishment Clause of the First Amendment, because it lacks a clear secular purpose" and that "[t]he Act impermissibly endorses religion by advancing the religious belief that a supernatural being created humankind" (*Edwards v. Aguillard*, 1987). The Louisiana trial in general, and the amicus brief in particular, had the effect of temporarily galvanizing the scientific community into defining science as a body of knowledge accumulated through a particular scientific method, as defined by the leading members of the scientific community themselves. Science is as scientists do.

DELIMITING THE BOUNDARIES BETWEEN SCIENCE AND PSEUDOSCIENCE

Creation-science (and its most recent hybrid, intelligent design theory) is just one of many claims that most mainstream scientists reject as pseudoscience. But what about those claims to scientific knowledge that are not so obviously classified as pseudoscience? When encountering a claim, how can one determine whether it constitutes a legitimate assertion as scientific? What follows is a list of ten questions that get to the heart of delimiting the boundaries between science and pseudoscience:

(1) How reliable is the source of the claim? All scientists make mistakes, but are the mistakes random, as one might expect from a normally reliable source, or are they directed toward supporting the claimants' preferred beliefs? Scientists' mistakes tend to be random; pseudoscientists' mistakes tend to be directional.

(2) Does this source often make similar claims? Pseudoscientists have a habit of going well beyond the facts, and so when individuals make many extraordinary claims, they may be more than iconoclasts. What one is looking for here is a pattern of fringe thinking that consistently ignores or distorts data.

(3) Have the claims been verified by another source? Typically, pseudoscientists make statements that are unverified or are verified by a source within their own belief circle. One must ask who is checking the claims and even who is checking the checkers.

(4) How does the claim fit with what is known about how the world works? An extraordinary claim must be placed in a larger context to see how it fits. When people claim that the pyramids and the Sphinx were built more than 10,000 years ago by an advanced race of humans, they are not pre-

senting any context for that earlier civilization. Where are its works of art, weapons, clothing, tools, and trash?

(5) Has anyone made an effort to disprove the claim or has only confirmatory evidence been sought? This is the confirmation bias or the tendency to seek confirmatory evidence and reject or ignore disconfirmatory evidence. The confirmation bias is powerful and pervasive. This is why the scientific method, which emphasizes checking and rechecking, verification and replication, and especially attempts to falsify a claim, is critical.

(6) Does the preponderance of evidence converge on the claimant's conclusion or a different one? The theory of evolution, for example, is proved through a convergence of evidence from a number of independent lines of inquiry. No single fossil or piece of biological or paleontological evidence has the word *evolution* written on it; instead, there is a convergence from tens of thousands of evidentiary bits that adds up to a story of the evolution of life. Creationists conveniently ignore this convergence, focusing instead on trivial anomalies or currently unexplained phenomena in the history of life.

(7) Is the claimant employing the accepted rules of reason and tools of research or have those rules and tools been abandoned in favor of others that lead to the desired conclusion? UFOlogists exhibit this fallacy in their continued focus on a handful of unexplained atmospheric anomalies and visual misperceptions by eyewitnesses while ignoring that the vast majority of UFO sightings are fully explicable.

(8) Has the claimant provided a different explanation for the observed phenomena or is it strictly a matter of denying the existing explanation? This is a classic debate strategy: Criticize your opponent and never affirm what you believe to avoid criticism. This strategy is unacceptable in science.

(9) If the claimant has proffered a new explanation, does it account for as many phenomena as does the old explanation? For a new theory to displace an old theory, it must explain what the old theory did and then some.

(10) Do the claimants' personal beliefs and biases drive the conclusions or vice versa? All scientists have social, political, and ideological beliefs that potentially could slant their interpretations of the data, but at some point, usually during the peer-review system, those biases and beliefs are rooted out or the paper or book is rejected for publication.

THE ENCHANTED GLASS OF SCIENCE

At the dawn of the scientific revolution in the early seventeenth century, the English philosopher Francis Bacon sought to turn away from the scholastic tradition of logic and reason as the sole road to truth, as well as reject the Renaissance quest to restore the perfection of ancient Greek knowledge. In his 1620 work *Novum Organum* (New Tool, contrary to the opinion of Aristotle's *Organon*), Bacon portrayed science as humanity's savior that would inaugurate a restoration of all natural knowledge through a proper blend of observation and logic, data and theory. Bacon understood, however, that there are significant social and psychological barriers that interfere with one's understanding of the natural world, "For the mind of man is far from the nature of a clear and equal glass, wherein the beams of things should reflect according to their true incidence; nay, it is rather like an enchanted glass, full of superstition and imposture, if it be not delivered and reduced" (p. 53). In the end, thought Bacon, science offers the best hope to deliver the mind from such superstition and imposture. Today, science continues to deliver on that hope.

See also Evolutionary Theory (Natural Selection); Philosophy of Science, Problems of; Popper, Karl Raimund.

Bibliography

Amicus Curiae Brief of 72 Nobel Laureates, 17 State Academics of Science, and 7 Other Scientific Organizations, in Support of Appellees, Submitted to the Supreme Court of the United States, October Term, 1986, as *Edwin W. Edwards, in His Official Capacity as Governor of Louisiana, et al., Appellants* v. *Don Aguillard et al., Appellees*, 1986.

Bacon, Francis. *Francis Bacon: A Selection of His Works*, edited by Sidney Warhaft. Toronto: Macmillan, 1965.

Carroll, Robert Todd. *The Skeptic's Dictionary: A Collection of Strange Beliefs, Amusing Deceptions, and Dangerous Delusions*. Hoboken, NJ: Wiley, 2003.

Charpak, Georges, and Henri Broch. *Debunked!: ESP, Telekinesis, and Other Pseudoscience*. Translated by Bart K. Holland. Baltimore, MD: Johns Hopkins University Press, 2004.

Cohen, I. Bernard. *Revolution in Science*. Cambridge, MA: Belknap Press of Harvard University Press, 1985.

Dawkins, Richard. *Unweaving the Rainbow: Science, Delusion, and the Appetite for Wonder*. Boston: Houghton Mifflin, 1998.

Eddington, Arthur Stanley. *The Philosophy of Physical Science*. Ann Arbor: University of Michigan Press, 1958.

Feyerabend, Paul. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: NLB, 1975.

Feyerabend, Paul. *Science in a Free Society*. London: NLB, 1978.

Frank, Philipp. *Philosophy of Science: The Link between Science and Philosophy*. Englewood Cliffs, NJ: Prentice-Hall, 1957.

Gardner, Martin. *Fads and Fallacies in the Name of Science*. New York: Dover Publications, 1957.

Gardner, Martin. *Science: Good, Bad, and Bogus*. Buffalo, NY: Prometheus Books, 1981.

Gilovich, Thomas. *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. New York: Free Press, 1991.

Harré, Rom. *The Philosophies of Science*. 2nd ed. New York: Oxford University Press, 1985.

Holton, Gerald. *The Advancement of Science and Its Burdens: The Jefferson Lecture and Other Essays*. New York: Cambridge University Press, 1986.

Kuhn, Thomas S. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 1977.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Medawar, Peter. *Pluto's Republic*. New York: Oxford University Press, 1984.

National Science Foundation. "Science Fiction and Pseudoscience." *Science Indicators Biennial Report*, 2002. http://www.nsf.gov/statistics/seind02/c7/c7h.htm.

Nickell, Joe. *The Mystery Chronicles: More Real-Life X-Files*. Lexington: University of Kentucky Press, 2004.

Olson, Richard. *Science Deified and Science Defied: The Historical Significance of Science in Western Culture*. Vol. 1. Berkeley: University of California Press, 1982.

Oppenheim, Janet. *The Other World: Spiritualism and Psychical Research in England, 1850–1914*. New York: Cambridge University Press, 1985.

Park, Robert L. *Voodoo Science: The Road from Foolishness to Fraud*. New York: Oxford University Press, 2000.

Piattelli-Palmarini, Massimo. *Inevitable Illusions: How Mistakes of Reason Rule Our Minds*. Translated by Massimo, Piattelli-Palmarini, and Keith Botsford. New York: Wiley, 1994.

Planck, Max. *The Philosophy of Physics*. Translated by H. W. Johnston. New York: Norton, 1936.

Plait, Philip C. *Bad Astronomy: Misconceptions and Misuses Revealed, from Astrology to the Moon Landing "Hoax."* New York: Wiley, 2002.

Popper, Karl. *The Logic of Scientific Discovery*. New York: Harper and Row, 1934.

Randi, James. *Flim-Flam!: Psychics, ESP, Unicorns, and Other Delusions*. Buffalo, NY: Prometheus Books, 1982.

Sagan, Carl. *The Demon-Haunted World: Science as a Candle in the Dark*. New York: Random House, 1995.

Schick, Theodore, Jr., and Lewis Vaughn. *How to Think about Weird Things: Critical Thinking for a New Age*. Mountain View, CA: Mayfield, 1995.

Shermer, Michael. *The Borderlands of Science: Where Sense Meets Nonsense*. New York: Oxford University Press, 2001.

Shermer, Michael, ed. *Encyclopedia of Pseudoscience*. Santa Barbara, CA: ABC-CLIO, 2002.

Shermer, Michael. *Science Friction: Where the Known Meets the Unknown*. New York: Times Books, 2005.

Shermer, Michael. *Why People Believe Weird Things: Pseudoscience, Superstition, and Other Confusions of Our Time*. New York: W. H. Freeman, 1997.

Taubes, Gary. *Bad Science: The Short Life and Weird Times of Cold Fusion*. New York: Random House, 1993.

Vyse, Stuart A. *Believing in Magic: The Psychology of Superstition*. New York: Oxford University Press, 1997.

Walker, D. P. *Unclean Spirits: Possession and Exorcism in France and England in the Late Sixteenth and Early Seventeenth Centuries*. Philadelphia: University of Pennsylvania Press, 1981.

Williams, William F., ed. *Encyclopedia of Pseudoscience*. New York: Facts on File, 2000.

Michael Shermer (2005)

SCIENCE AND RELIGION

See *Religion and the Biological Sciences; Religion and the Physical Sciences*

SCIENCE POLICY

Science policy deals with how society supports science and how science is utilized in society. The philosophy of science policy considers both interactions from the perspectives of logic, epistemology, ethics, political philosophy, metaphysics, and ontology. Its domain is broader than the philosophy of science, which emphasizes logical and epistemological questions and goes deeper than the descriptive analyses of science, technology, and society (STS) studies.

The central issues in the philosophy of science policy may be distinguished in terms of its two constituent terms: the structure and proper influence of policy on science, and the structure and proper role of science in public policy. Propaedeutic is the question of the nature of policy itself.

WHAT ARE POLICIES?

What is known as the demarcation problem in the philosophy of science analyzes science as a special form of knowledge. What are known as boundary issues in STS studies describe the distinctive practices of the science-society interface. By contrast, the phenomenon of policy has been subject to little conceptual examination either as knowledge or as practice.

The term *policy* does not occur in traditional political philosophy. There is no word in either Plato's *Republic* or Aristotle's *Politics* that translates as *policy*. Neither

does it occur in Thomas Hobbes's *Leviathan* (1651) or Jean-Jacques Rousseau's *Social Contract* (1762). Indeed, the term is somewhat peculiar to the English language. *Policy* is translated into French as *politique* and into Spanish, depending on context, as *política* or *norma*. In German it can be rendered by *Politik* and a host of other terms.

In English policies are associated with legal documents such as insurance contracts and guidelines for corporate or governmental behavior. Corporations have policies for the treatment of customers or employees, and governments and government agencies debate military, fiscal, educational, healthcare, and environmental policies. Although a policy has sometimes been defined simply as a decision, this seems inadequate if for no other reason than that one can talk about "policy decisions" and "decision policies."

Reframing Ludwig Wittgenstein's famous question, one may ask what is the difference between my arm going up, me raising my arm (*Philosophical Investigations* §161)—and a policy for raising my arm. The comparison suggests the concept of policy as a guideline for action justified by some kind of analysis. Policies fall in a middle range between decisions about individual actions and general principles for actions. Policies are also to be distinguished from laws and rules.

Since the key difference between my hand going up and me raising my hand is the presence of an intention, a policy might be seen as a particular kind of intention. G. E. M. Anscombe (1957) maintains that for a person to have an intention is to have both a desire to do X and a belief that he or she will do X. On this account intention becomes a secondary rather than a primary phenomenon. In like manner, policies would become secondary phenomena, derivative of desires and beliefs, with the beliefs being justified by scientific evidence or analysis that X will provide results satisfying the desire.

Anscombe's view is criticized by Donald Davidson (1978) and Michael E. Bratman (1987). Davidson believes that intentions are best described as pro-attitudes or evaluative judgments. According to Bratman the most effective way for human beings living in association with others to become effective agents is to have plans, the elements of which are intentions. For Davidson, then, policies might be defined as group pro-attitudes regarding types of actions. For Bratman policies would be closely associated with group plans.

The field of policy studies forms part of a general twentieth-century effort to extend scientific rationality into group planning, especially in institutional contexts. (Having or making policies applies more to groups of people than to individuals, except for individuals in positions of power who set policies for others.) In this sense all policy is science policy, since it commonly involves scientific justifications of action plans, whether these are for military, fiscal, educational, healthcare, or environmental contexts.

POLICIES FOR SCIENCE

Following Harvey Brooks (1968), the philosophy of science policy explores two domains: the philosophical aspects of (1) policies for the funding and governance of science, and (2) ways that science can contribute to and/or impede the political process.

For fifty years after World War II, the basic principle underlying U.S. policy for science was that the government should provide no-strings attached funding to scientists, on the grounds that autonomous scientific research invariably benefits society by making contributions to military power, healthcare, and economic competitiveness further down the road (Bush 1945). There were arguments around the margins regarding how much independence to give scientists (e.g., national security required some limits) and about what constituted a well-balanced investment in mathematics, physics, chemistry, biology, and the social sciences. But no debate altered the basic policy: Give money to scientists and let them make their own decisions about how to spend it, because this will eventually rebound to the good of society.

The end of the cold war and increasing budget pressures allowed questions to surface about this basic policy and its foundational justification, the linearity thesis—the belief that autonomous scientific research produces social benefits in an automatic and linear way: more science, more benefit. As historical and sociological analyses of science have shown, however, the linearity thesis applies more to a few highly qualified special cases than as a general rule.

Reassessment of this policy approach has taken multiple forms. In one instance, in response to cases of research misconduct, it has been argued that conscious efforts are needed to promote collaboration between scientists and stakeholders (Guston 2000). Others have asked whether additional knowledge may overwhelm, getting in the way of the reflection needed about alternatives (Mitcham and Frodeman 2002). More generally, STS studies have argued the sociopolitical construction of scientific knowledge, thus challenging the ideal of scientific autonomy.

Taking these reassessments in a political philosophical direction, Philip Kitcher (2001) argues for a modification of linearity policy. Although a moderate realist who sees scientific knowledge as true, Kitcher is not willing to accept existing institutional arrangements for science as the best imaginable. Moreover, given the limitations of public funding, any one scientific research program is necessarily pursued at the expense of others, so that there is a proper place for extrascientific influence on the selection of publicly funded research priorities. Creating the proper policy for science depends on an understanding of what constitutes "well-ordered science" under such conditions.

SCIENCE IN POLICY MAKING

Several positions have been staked out in terms of how science properly contributes to policy making. In many quarters (both scientific and nonscientific) there has been a strong presumption that science can "answer" policy questions with the definitive account and/or solution to a problem. Although most policy analysts and many scientists now reject any simple version of this belief, it continues to influence the policy-making process. Two basic issues here concern the extent to which science can serve as an assessor or provider of means for nonscientifically determined ends, and whether or not science can assess ends as well as means.

The advancement of external ends has been a vision of modern science since its origins in the work of Francis Bacon, Galileo Galilei, and René Descartes. However, there has been little systematic examination of assumptions about whether in particular cases science is the best way to achieve certain goals. Does increased scientific knowledge or enhanced technologically efficient action always promote social or personal goods? Information overload can, for instance, actually inhibit decision making, and the excitements of technology have been known to skew appreciation of other goods.

At the end of the twentieth century a cadre of scientists and social scientists began to argue that science policy should go beyond the assumption of linearity. Daniel Sarewitz (1996), Donald E. Stokes (1997), and others proposed to examine the publicly stated goals of science funding and then scrutinize whether end-benefit outcomes have been or are likely to be achieved. While this new science policy is a substantial improvement over the old, it nevertheless limps in one important respect: It accepts whatever social goals may have been given a rhetorical blessing by the existing body politic. The philosophical analysis of methods for assessing connections between scientific effort and assumed end-benefits deserves attention, but it does not reconsider the worthiness of the proposed ends themselves. Ends must be reflected on as well as means—which is where philosophy has a significant role to play.

The most philosophically expansive approach to policy research is what Harold D. Lasswell called the *policy sciences*. In the course of his long, interdisciplinary career, Lasswell sought to develop a method for the systematic analysis of any policy problem (see Lerner and Lasswell 1951, Lasswell 1971). Influenced by the Chicago school pragmatism of such thinkers as George Herbert Mead and Charles E. Merriam, Lasswell's method centers around five intellectual tasks: clarification of goals; descriptions of trends; analysis of conditions; projection of future developments; and invention, evaluation, and selection of alternatives. These tasks are necessary to address intelligently any number of policy issues, whether public or private, from those associated with taxation or warfare to problems of manufacturing and marketing.

Despite Lasswell's achievements, however, there are evident opportunities for further philosophical criticism of method in the policy sciences, the practice of which depends on a prior commitment to goods such as human rights and democracy. Illustrations of deeper reflections on ends can be found in such diverse work as Daniel Callahan (2003); Alan Lightman, Daniel Sarewitz, and Christina Desser (2003); and Leon Kass et al. (2003). Callahan questions what he calls the *research imperative* that seems to take every social problem as an opportunity for more scientific research. Lightman, Sarewitz, and Desser undertake a collective reflection on "living with the genie" of scientific and technological productivity. Kass and the President's Council on Bioethics philosophically assess contemporary aspirations to turn therapy into enhancement.

PHILOSOPHICAL CRITICISM

In what sense is philosophy of science policy genuine philosophy? Philosophy may be subdivided along two major axes. The first axis is defined by the fundamental questions that constitute philosophical reflection, of which it is common to distinguish at least logic, epistemology, ethics, political philosophy, and metaphysics or ontology. A second axis is constituted by the particular fields or topoi where such questions are deployed. This axis yields an indefinite series of regionalizations such as the philosophy of science, of art, of religion, of law, of language, and more. The philosophy of science is characterized by the prominence of logical and epistemological issues, with

only subsidiary attention to ethics, political philosophy, or metaphysics. But, in fact, there are also important questions of the logic of science policy arguments, the character of science policy knowledge, the ethics of science policy decision making, and the political philosophy of science. Because every science policy makes assumptions about the status of science itself, the philosophy of science policy must consider not just the epistemological status of scientific knowledge but also the justice and ontological boundaries of science as a human activity and of its various institutions.

As a regional expression of philosophy, the philosophy of science policy explores a spectrum of concerns. Policy methods deserve logical analysis. The epistemological strengths and weaknesses of models and simulations, not just in physics or climatology but also in policy analysis, call for critical reflection. Behaviors within the professional scientific community and in relations between scientists and the public, including those of policy analysts, require philosophical assessment. Policies in and for scientific communities and those mediating between science and society, along with the role of scientific expertise in a democratic state, are subject to political philosophical scrutiny. Questions related to scientific institutions and their manifold boundary organizations are ontological as well as sociological; the distinction between policy for science and science in policy may be less sound than is commonly assumed.

Discussions that move from interest group power and economic efficiency to questions of truth, goodness, and beauty can make science policy work richer and more robust—and thus, in other than technical or economic senses, more effective. The philosophy of science policy holds out the promise of promoting science policies that are less incomplete, distorted, and unconscious than might otherwise be the case.

See also Anscombe, Gertrude Elizabeth Margaret.

Bibliography

Anscombe, G. E. M. *Intention*. Oxford, U.K.: Blackwell, 1957.

Bratman, Michael E. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.

Brooks, Harvey. *The Government of Science*. Cambridge: MIT Press, 1968.

Bush, Vannevar. *Science: The Endless Frontier*. Washington, DC: U.S. Government Printing Office, 1945.

Callahan, Daniel. *What Price Better Health? Hazards of the Research Imperative*. Berkeley: University of California Press, 2003.

Davidson, Donald. "Intending" (1978). In *Essays on Actions and Events*, 83–102. New York: Oxford University Press, 1980.

Frodeman, Robert, and Carl Mitcham, eds. "Toward a Philosophy of Science Policy: Approaches and Issues." *Philosophy Today*, vol. 48, no. 5 (supplement 2004).

Guston, David H. *Between Politics and Science: Assuring the Integrity and Productivity of Research*. New York: Cambridge University Press, 2000.

Kass, Leon R., et al. *Beyond Therapy: Biotechnology and the Pursuit of Happiness*. New York: Regan Books, 2003.

Kitcher, Philip. *Science, Truth, and Democracy*. New York: Oxford University Press, 2001.

Lasswell, Harold D. *A Pre-view of Policy Sciences*. New York: American Elsevier, 1971.

Lerner, Daniel, and Harold D. Lasswell, eds. *The Policy Sciences: Recent Developments in Scope and Method*. Stanford, CA: Stanford University Press, 1951.

Lightman, Alan, Daniel Sarewitz, and Christina Desser, eds. *Living with the Genie: Essays on Technology and the Quest for Human Mastery*. Washington, DC: Island Press, 2003.

Mitcham, Carl, and Robert Frodeman. "The Plea for Balance in the Public Funding of Science." *Technology in Society* 24 (1–2) (2002): 83–92.

Sarewitz, Daniel. *Frontiers of Illusion: Science, Technology, and the Politics of Progress*. Philadelphia: Temple University Press, 1996.

Stokes, Donald E. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings Institution Press, 1997.

Carl Mitcham (2005)
Robert Frodeman (2005)

SCIENCE STUDIES

The phrase "science studies" is sometimes used as an umbrella term referring to work in history of science, philosophy of science, research ethics, and so on. But it can also designate a new interdisciplinary approach to the study of science, technology, and society, one that challenges traditional views about the epistemic basis of scientific knowledge and the proper role of science in society. It is this intellectual movement called Science and Technology Studies (STS) that will be discussed here.

Science Studies in the STS sense discards almost all of the distinctions common in traditional philosophy of science, such as the demarcation between the context of discovery and the context of justification, prescriptions versus descriptions, and theory versus observation. Instead, it looks at science as a social activity that cannot be usefully understood in isolation from either technology or society at large. More important than analysis is contextualization. Case studies of local scientific practices are the preferred route to understanding.

What we now call Science Studies (or STS) has a surprising variety of geographical roots. The Strong Program began in Edinburgh in the 1970s. Harry Collins's study of the experimenters' regress was done in England, where Steve Woolgar is also located. Bruno Latour's actor-network theory was developed in Paris. Karin Knorr Cetina worked in Bielefeld, Germany. Feminist standpoint epistemology stems from Sandra Harding's work in America. The 1980s saw a plethora of influential books and articles centered on the notion of social construction, some of it directly influenced by the pioneers such as that of Andrew Pickering, who earned his PhD at Edinburgh, and Trevor Pinch, who worked with Harry Collins at Bath. Systematic criticism of both the historical work and the philosophical claims of STS scholars resulted in the so-called Science Wars, which was triggered by a hard-hitting critical study in 1994 by Paul Gross and Norman Levitt.

Contributors to Science Studies (or STS) draw on resources from diverse disciplinary backgrounds, such as sociology, anthropology, political science, economics, and linguistics. But all the studies share starting points based on interpretations of work in more contemporary philosophy of science. It is to these common philosophical presuppositions that we now turn.

PHILOSOPHICAL ROOTS

Although Thomas Kuhn himself repudiated some of the most radical extensions of his theory of normal science and scientific revolutions, Science Studies is certainly part of his legacy. There are frequent STS references to his thesis that scientific knowledge is embodied in the practices of a community. Because observation is theory-laden, it is assumed that empirical studies can never be used as a neutral arbiter between rival paradigms. The logical fact that any universal theory is underdetermined by the evidence in support of it is taken to mean that there can be no rational basis sufficient to justify the choice of one theory over another. The point, noted by Pierre Duhem and W.V. Quine, that it is always logically possible to save a theory from refutation by altering auxiliary hypotheses is deployed to advance the skeptical conclusion that philosophical accounts of scientific method and scientific rationality cannot explain why scientists prefer some theories to others and assign great epistemic weight to certain claims and not to others.

Most philosophers of science in the early twenty-first century would agree that there is no ironclad defense against these kinds of skeptical arguments—indeed philosophers are the ones who undermined the quest for unrevisable foundations in the first place! It is a truism today that there is no instant rationality in science. Instead, philosophers look for the fallible canons of comparative rationality that underlie scientific judgments. For example, in his theory of Scientific Research Programmes, Imre Lakatos argued that most of the history of science could be understood in terms of his methodology. It was only when the evaluations of actual scientists differed from his normative account that one should invoke extra-scientific considerations such as ideology, personal rivalries, or economic interests. Earlier historians of science also drew a distinction between internal history, a narrative of the mostly rational development of scientific ideas through the experimental method, and external history, the story of scientific institutions and their interaction with the larger society. And when sociologists such as Robert Merton studied those scientific institutions, they looked for the operation of special norms that would show why the output of science was superior to the knowledge produced by theologians or grocers!

Adherents to the Science Studies find little use for the internal or external distinction: Trying to isolate ideas from the people who hold them impedes understanding. Similarly, they would invert Lakatos's order of analysis. Instead of looking for the intellectual problems motivating researchers and their attempts to bring evidence to bear on proposed solutions, they begin with a rich sociological description of a scientific episode. What are the lines of authority and collaboration? Through what mechanism is expertise awarded? An important focus of the analysis should be the various rhetorical stratagems employed in communication: How do they reflect the conflicting interests or differing cognitive resources of the participants (for example, theoreticians versus experimenters, policy makers versus scientists, and funding committees versus referees)? The factors that philosophers of science or internalist historians find of interest would be embedded within the STS account, but they would never be privileged. Merton's norms of objectivity and organized skepticism are viewed primarily as part of the rhetorical arsenal of scientists, not as fundamental guides to behavior.

THE STRONG PROGRAM

Let us now take a brief look at some detailed proposals of the Science Studies approach, bearing in mind that there are disputes within this loosely-knit, interdisciplinary field. We begin with an early, very influential initiative—the Strong Programme of the Edinburgh school. David

Bloor describes the program as a sociology of scientific knowledge (SSK) that is based on four tenets.

First, our accounts of science should be causal (one is reminded of Marxian attempts to explain the content of scientific theories). Secondly, they should be impartial with respect to truth or falsity, rationality or irrationality—both are in need of explanation (many philosophical accounts take rational inference as self-explanatory—in fact to go further may lead to an infinite regress). Furthermore, the explanations of true and false beliefs should be symmetrical; that is, the types of causal factors invoked should be the same (if ideology or political interests are invoked to explain false beliefs, they also need to be brought into a story of the origins of true beliefs). Finally, our approach should be reflexive: The claims made from within the Strong Program are to be analyzed and explained in exactly the same manner as are episodes in the history of science.

As laid out above, the Strong Programme is a variety of philosophical naturalism, a position that underlies most work in cognitive science. Its most controversial aspect is the symmetry thesis. Why should we posit a priori that the causal chain leading to hallucinations should contain exactly the same elements as the process that produces ordinary visual experiences? Or compare the stories of adherence of Soviet scientists to T. D. Lysenko's theory of acquired characteristics with that of the acceptance of Dmitri Mendeleev's Periodic Table in Czarist Russia: Undoubtedly both situations involved elements of nationalism and the striving of scientists for recognition. But are not the asymmetries more significant? In one case a major factor was the coercion to conform with the wishes of Joseph Stalin; in the other, chemists followed their noses to arrive at a workable classification of chemical phenomena.

Many of the case studies produced by adherents of the Strong Programme focus on scientific controversies. The general pattern is to look at a wide range of social factors, such as class, political pressures, disciplinary commitments, and power structures within the profession. They then argue that these kinds of interests have a strong influence on the conclusions that scientists reach about which of the competing theories is deserving of their allegiance. Their account of the resolution of a controversy does not privilege appeals to epistemic considerations such as predictive accuracy or theoretical coherence.

SOCIAL CONSTRUCTIONISM

Much work in Science Studies is based on the tenet that all scientific entities are socially constructed. In certain instances, such a claim is nontrivially true: John Searle uses the example of money. A metal disc with Sacagawea's portrait stamped on it does not count as money without the construction of a vast social network that turns it into legal tender. One can also make sense out of the assertion that there were no homosexuals until the late nineteenth century, by adding a gloss to the effect that the term homosexual is to be read as connoting a historically specific, medical-psychological category. But how are we to understand Latour's claim that there was no anthrax before Louis Pasteur's research or that TRH, a product of the hypothalamus, was invented in a certain California laboratory?

It is important to appreciate the difficulty scientists face in isolating natural products, especially when the process involves new sorts of instruments or laboratory procedures. But what is gained by blurring the distinction between a new concept, which certainly is a social construction, a chemical or biological entity, which either existed in nature or was synthesized, and the development of scientific consensus about the match between concept and the object described? It perhaps helps a little bit to understand why Latour and Woolgar would make such perplexing claims if we note that they set out to apply ethnographic approaches to life in the laboratory. Anthropologists who are studying an exotic culture dutifully describe the behavior of the people they are studying. An ethnomethodologist's thick description of a rain dance need say nothing about whether rain actually ensues. In a similar fashion STS scholars can describe the interactions and assertions of scientists without saying anything about whether the object the scientists claim to be studying actually exist. However, radical social constructionists go on to say that social constructions exhaust reality—there is no underlying strata that is being more or less accurately represented.

Even more startling is the so-called Actor-Network Theory (ANT) developed by Latour and Michel Callon, which posits a symmetry between humans and nonhuman entities, such as scallops (Callon's example) or technological devices. These so-called actants form networks in which their competing interests are negotiated. The result is a complex ecological system in which ideas and artifacts, scientist, and resources form an ontology based on what they call relational materiality. Difficult to understand, ANT has generated considerable critical discussion among STS practitioners. Some draw the line at

assigning agency to scallops; others object to attributing stability to inscription devices and the implication that scientific findings involving instruments can be exported from one lab to another without intervening social constructions by the local community.

FEMINIST STUDIES

A branch of Science Studies that has generated wide interest is the large corpus of feminist writings looking for the effect of gender ideology on both the content and practice of science. These range from significant, but relatively uncontroversial, empirical analyses of the social factors that lead to the attrition of women at every stage of their professional careers to radical claims about the intrinsically sexist nature of the science of mechanics. A central claim, and one that fits in well with some of the STS approaches described above, declares that scientific concepts of sex and gender have historically been strongly influenced by biases inherent in patriarchal societies. One recalls Aristotle's association of the active form with maleness, whereas females were the bearers of passive matter. Anne Fausto-Sterling argues that similar nonscientific influences have entered into the modern study of sex hormones. These are typical examples of social constructionist analyses: Scientific results are held to be strongly influenced by the social milieu; they are not simple reflections of empirical studies.

Feminists posit the influence of gender ideology on the content of science in areas increasingly distant from the study of reproduction. Londa Schiebinger claims that Carl Linneaus's characterization of the class mammalia was influenced by political debates about the propriety of wet nursing. Some have argued that the development of the science of hydrodynamics was delayed because men were uncomfortable dealing with material that was moist and yielding and that the interest in mechanical interactions between hard, rigid bodies that characterized the beginning of the Scientific Revolution was a masculinist preoccupation. And what about the scientific prejudice in favor of linear theories, reductionism, and simplicity? Relying on object-relations theory from psychology, it has even been claimed that the traditional norm of objectivity, of distancing oneself from phenomena, is a reflection of the process by which male children are psychically separated from the mother.

All of these studies follow the STS pattern of trying to show the radical contingency of scientific developments: If social circumstances had been different, the content of science would have been different. Feminists accompany this descriptive analysis with prescriptions for changing science. Sandra Harding calls for what she calls strong objectivity: If present science is distorted by the predominance of male perspectives, would not science become more objective by the deliberate inclusion of views from the standpoint of women, minorities, workers, and any other group that is underrepresented in today's scientific community? Helen Longino advocates a sort of affirmative action for approaches to understanding the world that are anti-reductionist, nonhierarchical, and unabashedly politically progressive. If science is always socially constructed anyway, why not deliberately construct scientific inquiry in a humanitarian fashion?

REACTIONS TO SCIENCE STUDIES

The above descriptions of the leading STS approaches give an indication of why their underlying philosophical posits might be viewed as tendentious. Many of their case studies of the factors affecting the acceptance of scientific theories have also generated historiographic skepticism. Critics argue that, contrary to STS claims, Pasteur's religious views, Robert Boyle's preoccupation with chastity, or Karl Pearson's upwardly mobile class interests had a negligible effect on their scientific positions. As an explanation of the acceptance or rejections of scientific hypotheses, STS accounts are not satisfactory.

Ironically, however, the academic reaction to work in Science Studies cannot be understood purely in terms of its intellectual merits—or demerits. Instead, we must also invoke the sorts of interests and ideological factors that STS brings to the forefront! Members of the so-called academic left found the rhetoric of STS very congenial. Already suspicious of the authority of science and troubled by the pace of technological change, they eagerly took up slogans to the effect that science was a creature of the military-industrial complex and a handmaiden to imperialist regimes. The title of a popular textbook by Harry Collins and Trevor Pinch sums it up nicely: *Science as Golem*. There were calls for science for the people, feminist science, and postcolonial science.

Scientists and other intellectuals concerned about the level of funding for scientific research and the general low level of scientific literacy in America mounted a vigorous response. Paul Gross and Norman Levitt's *Higher Superstition: The Academic Left and Its Quarrels with Science* set off intense debates both in the media and in universities. Science Studies practitioners were sometimes lumped in with postmodernists, new age mystics, and so-called scientific creationists. Although STS people sometimes protested that they were not antiscience per se, but only objecting to what they considered to be overly adu-

latory accounts of science as a hyperrational activity, they did not make a concerted effort to disassociate themselves from their more radical fellow travelers. Some critics took their ambivalent response to a hoax perpetrated by Alan Sokal, who succeeded in publishing a factually absurd, but politically correct, paper in a leading journal called *Social Text*, as an indication of a weak commitment to traditional scholarly norms.

Science Studies in the STS sense has spurred the attempts of historians, philosophers, and sociologists of science, who favor more traditional approaches to science studies, to provide accounts of the development of science that give us more understanding of the social dimensions of scientific inquiry (see Philip Kitcher's calls for a new socio-historico-philosophical approach). It has also highlighted the importance of developing a more detailed and realistic picture of scientific inquiry. This project is nicely described in the title of Susan Haack's book: *Defending Science—Within Reason: Between Scientism and Cynicism*.

See also Feminist Epistemology; Kuhn, Thomas.

Bibliography

Barnes, Barry, David Bloor, and John Henry. *Scientific Knowledge: A Sociological Analysis*. Chicago: University of Chicago Press, 1996.

Biagioli, Mario, ed. *The Science Studies Reader*. New York: Routledge, 1998.

Bloor, David. *Knowledge and Social Imagery*. 2nd ed. Chicago: University of Chicago Press, 1991.

Callon, Michel. "Some Elements of A Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay." In *Power, Action, and Belief: A New Sociology of Knowledge?*, edited by J. Law. London: Routledge and Kegan Paul, 1986.

Collins, Harry M. *Changing Order: Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press, 1992.

Fausto-Sterling, Anne. *Myths of Gender: Biological Theories about Women and Men*. 2nd ed. New York: Basic Books, 1992.

Gross, Paul R. and Norman Levitt. *Higher Superstition: The Academic Left and Its Quarrels with Higher Education*. Baltimore, MD: Johns Hopkins Press 1998.

Haack, Susan. *Defending Science—Within Reason: Between Scientism and Cynicism*. Amherst, NY: Prometheus Books, 2003.

Hacking, Ian. *The Social Construction of What?* Cambridge, MA: Harvard University Press, 2000.

Harding, Sandra. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca, NY: Cornell University Press, 1991.

Kitcher, Philip. *Science, Truth, and Democracy*. New York: Oxford University Press, 2003.

Knorr Cetina, Karin D. *The Manufacture of Knowledge: An Essay on The Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press, 1981.

Koertge, Noretta, ed. *A House Built on Sand: Exposing Postmodernist Myths about Science*. New York: Oxford University Press, 1998.

Latour, Bruno. *The Pasteurization of France*. Cambridge, MA: Harvard University Press, 1993.

Latour, Bruno. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press, 1988.

Latour, Bruno, Steve Woolgar, and Jonas Salk. *Laboratory Life*. Princeton, NJ: Princeton University Press, 1986.

Longino, Helen. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Merton, Robert K. *The Sociology of Science*. Chicago: University of Chicago Press, 1973.

Parsons, Keith. *The Science Wars: Debating Scientific Knowledge and Technology*. Amherst, NY: Prometheous, 2003.

Pickering, Andrew. *Constructing Quarks: A Sociological History of Particle Physics*. Chicago: University of Chicago Press, 1984.

Pinch, Trevor. *Confronting Nature: The Sociology of Solar-Neutrino Detection*. Dordrecht: Reidel, 1986.

Pinnick, Cassandra L., Noretta Koertge, and Robert F. Almeder, eds. *Scrutinizing Feminist Epistemology: An Examination of Gender in Science*. New Brunswick, NJ: Rutgers University Press, 2003.

Potter, Elizabeth. *Gender and Boyle's Law of Gases*. Bloomington: Indiana University Press, 2001.

Schiebinger, Londa. *Has Feminism Changed Science?*. Cambridge, MA: Harvard University Press, 1999.

Searle, John R. *The Construction of Social Reality*. New York: The Free Press, 1995.

Seismondo, Sergio. *An Introduction to Science and Technology Studies*. Oxford: Blackwell, 2004.

Sokal, Alan D., and Jean Bricmont. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: Picador USA, 1998.

Noretta Koertge (2005)

SCIENTIA MEDIA AND MOLINISM

The *scientia media* is a key term in the theology of Luis de Molina (1535–1600) and in the variants of his teaching introduced by the later Jesuits, especially Robert Bellarmine, Leonard Lessius, Francisco Suárez, and Gabriel Vasquez, in the attempt to resolve the apparent contradiction between the doctrines of grace and of free will.

Molina, a Spanish Jesuit who taught at Coimbra and Evora in Portugal, published his famous *Liberi Arbitrii cum Gratiae Donis, Divina Praescientia, Providentia, Praedestinatione et Reprobatione Concordia* at Lisbon in

1588. The publication of the *Concordia*, as it came to be called, soon led to a controversy that divided the theologians and philosophers of Spain. Generally, the position of Molina was enthusiastically supported by members of his own order and just as vigorously denounced by the Thomists.

For Molina the essential problem was to maintain both human freedom and the efficacy of grace. Given the fact of God's foreknowledge, Molina wished to preserve such a foreknowledge without lapsing into determinism, to show that although God knows infallibly what an individual will freely do, such an infallible knowledge in no way determines the will of the individual. Molina argued that there is a cooperation or concursus of human free will with the divine grace, in contrast to the Thomist view that man's will was physically predetermined to act freely by God. Molina held that this was only a disguised form of determinism. The Thomists maintained that Molina denied the universal divine causality.

The central point in Molina's solution of this problem is based upon the *scientia media*. This, according to Molina, is a form of the divine knowledge that lies between the two forms of God's knowledge that Thomas Aquinas had described in the *Summa*. Thomas maintained that God's knowledge may be one of "vision," a knowledge of that which exists, has existed, or will exist. Alternatively, God's knowledge may consist of the purely possible, a knowledge of "simple understanding," of things and events that have not existed, do not exist, and will not exist. The *scientia media* for Molina is a mean between these two forms of knowledge and is the knowledge that God has of conditional future contingent events; thus, God foreknows from all eternity what an individual would do under certain circumstances if offered his grace. Thomas held that nothing lies outside the divine causality and that God's knowledge, or vision, of the future free acts of the individual entails an act of will by God that predetermines that our acts are free. Molina insisted that God's knowledge is prior to the decree of his will and that his foreknowledge does not predetermine our free acts. God, knowing infallibly what an individual will do under certain circumstances if offered his grace, decrees the circumstances and the grace necessary to effect the cooperative action of the individual. Hence, the infallibility and efficacy of grace is due to the infallibility of God's knowledge, the *scientia media*, not to anything in the grace itself.

The distinction between sufficient and efficacious grace throws further light on these contrasting positions. Like the Thomists, Molina accepted the necessity of grace for salvation, the absolute gratuity of grace, and that sufficient grace is given to all people. However, Molina denied the need for any distinction between sufficient and efficacious grace. He claimed that sufficient grace becomes efficacious if the will of the individual accepts it. Thus, God foreknew St. Paul's consent before he decreed the grace necessary for conversion. The concurrence of the simultaneous act of the individual and the grace of God replaced the notion that the decree of God is prior to the act of the individual and predetermines it. Thomists objected that this made the efficacy of the divine grace dependent on man rather than on God. Molina declared that the efficacy of grace was unimpaired, for its efficacy or infallibility was extrinsic to the act of the individual and intrinsic in God's foreknowledge. In effect, Molina endeavored to preserve more fully the freedom of the individual without destroying the power of grace; the Thomists were more concerned with preserving the power of grace without destroying the freedom of the individual.

LATER MOLINISM

Later Molinism is identified largely with what is termed *congruism*, a theological doctrine reflecting especially the views of Bellarmine and Suárez. Congruism retains the principal features of Molina's theology but modifies it in certain respects. Efficacious grace is equated with *gratia congrua* and sufficient grace with *gratia incongrua*. This distinction emphasized more strongly that grace was efficacious when it was congruous with those circumstances and the disposition of the individual that would enable him to will a certain act freely but infallibly. Grace was inefficacious when it was not congruous with the circumstances and disposition of the individual. The efficacy of the *gratia congrua* is intrinsic to the *scientia media* and extrinsic to the will of the individual. *Gratia incongrua* is grace that is sufficient for a salutary act but which the individual will reject.

On predestination the Molinists agreed with the Thomists that God wishes all people to be saved and that he extends sufficient grace to all, that contrary to Pelagianism predestination is wholly gratuitous, and that some individuals are elected in preference to others solely as God wills. However, they tended to modify the Thomist view of an absolute predestination to glory irrespective of foreseen merits. Many of the Molinists argued that predestination is conditional upon the future actions of the individual and becomes absolute only with the foreseen merits of the individual. In contrast to the Thomists, who argued for the priority of predestination

to glory to the predestination of efficacious grace, the Molinists held that God foresees in the *scientia media* that some will cooperate with his grace and predestined them to glory by offering them his grace.

The differences between Molina and his successors are more often subtle than essential. Although the debate on Molinism has continued for more than three centuries, Molinism is clearly compatible with faith and continues to have many supporters. Like Thomism it has its difficulties and its critics. The difference between the two schools remains essentially one of the relative emphasis to be placed upon grace or freedom.

See also Bellarmine, St. Robert; Molina, Luis de; Suárez, Francisco; Vasquez, Gabriel.

Bibliography

Beilby, James K., and Paul R. Eddy. *Divine Foreknowledge: Four Views*. Downers Grove, IL: InterVarsity Press, 2001.

Bonet, A. *La filosofia de la libertad en los controversias teologicas del siglo XVI. y primera mitad del XVII*. Barcelona, 1932.

Boyer, C. "Providence et liberté." *Gregorianum* (1938): 194–208.

Craig, William Lane. "'No Other Name: A Middle Knowledge Perspective on the Exclusivity of Salvation Through Christ." In *The Philosophical Challenge of Religious Diversity*, edited by. Philip L. Quinn and Kevin Meeker. New York: Oxford University Press, 2000.

Craig, William Lane. "Creation, Providence and Miracles." In *Philosophy of Religion*, edited by Brian Davies. Georgetown University Press: Washington, 1998.

Dekker, Eef. *Middle Knowledge*. Leuvens: Peeters, 2000.

Flint, Thomas P. *Divine Providence: The Molinist Account*. Ithaca: Cornell University Press, 1998.

Flint, Thomas. "Middle Knowledge and the Doctrine of Infallibility." In *Philosophical Perspectives, 5: Philosophy of Religion*, edited by James E. Tomberlin. Ridgeview: Atascadero, 1991.

Flint, Thomas P. "The Multiple Muddles of Maverick Molinism." *Faith and Philosophy* 20 (1) (2003): 91–100.

Flint, Thomas P. "The Possibilities of Incarnation: Some Radical Molinist Suggestions." *Religious Studies* 37 (3) (2001): 307–320.

Hasker, William. "The Antinomies of Divine Providence." *Philosophia-Christi* 4 (2) (2002): 361–375.

Hasker, William. "A New Anti-Molinist Argument." *Religious Studies* 35 (3) (1999): 291–297.

Hasker, William. "Providence and Evil: Three Theories." *Religious Studies* 28 (1) (1992): 91–105.

Kleinhappl, Johann. *Der Staat bei Ludwig Molina*. Innsbruck: Rauch, 1935.

Linville, Mark. "Ockhamists and Molinists in Search of a Way Out." *Religious Studies* 31 (4) (1995): 501–515.

Myers, David B. "Exclusivism, Eternal Damnation, and the Problem of Evil: A Critique of Craig's Molinist Soteriological Theodicy." *Religious Studies* 39 (4) (2003): 407–419.

O'Connor, Timoth. "The Impossibility of Middle Knowledge." *Philosophical Studies* 66 (2) (1992): 139–166.

Pegis, Anton C. "Molina and Human Liberty." In *Jesuit Thinkers of the Renaissance*, 75–132. Milwaukee: Marquette University Press, 1939.

Perszyk, Kenneth J. "An Anti-Molinist Argument." *Philosophical Studies* 90 (3) (1998): 215–235.

Perszyk, Kenneth J. "Free Will Defence With and Without Molinism." *International Journal for Philosophy of Religion* 43 (1) (1998): 29–64.

Perszyk, Kenneth J. "Molinism and Compatibilism." *International Journal for Philosophy of Religion* 48 (1) (2000): 11–33.

Perszyk, Kenneth J. "Molinism and the Consequence Argument: A Challenge." *Faith and Philosophy* 20 (2) (2003): 131–151.

Perszyk, Kenneth J. "Molinism and Theodicy." *International Journal for Philosophy of Religion* 44 (3) (1998): 163–184.

Rasolo, Louis. "Le dilemme du concours divin." *Analecta Gregoriana* 80 (1956).

Romeyer, B. "Libre arbitre et concours selon Molina." *Gregorianum* (1942): 169–201.

Stegmüller, Francis, ed. *Geschichte des Molinismus*, I: Neue Molinaschriften. Münster: Aschendorff, 1935.

Vansteenberghe, E. "Molinisme." In *Dictionnaire de Théologie Catholique*, Vol. X, cols. 2184–2187. Paris, 1929.

Werther, David. "Open Theism and Middle Knowledge: An Appraisal of Gregory Boyd's Neo-Molinism." *Philosophia Christi* 5 (1) (2003): 205– 215.

Zimmerman, Dean. "Richard Gale and the Free Will Defense." *Philo* 6 (1) (2003): 78–113.

John A. Mourant (1967)

Bibliography updated by Tamra Frei (2005)

SCIENTIFIC EXPLANATION

See *Explanation*

SCIENTIFIC LAWS

See *Laws, Scientific*

SCIENTIFIC METHOD

What follows is a description of various views on inductive inference and methods for inferring general theories as they have developed from the scientific revolution to modern times. Later, the development of methods for discovering causal relationships will be discussed.

MODERN METHODOLOGY. A strong influence on contemporary methodology is interdisciplinary research. In

the twentieth century, the question of how we can use observations to attain empirical knowledge became the subject of research in a number of disciplines, such as statistics, econometrics, and computer science. Modern philosophy of method continues to contribute to and draw on developments in related disciplines.

Another strong influence on contemporary methodology arises from studies of the history of science, which captured the attention of philosophers because of the groundbreaking work of Thomas Kuhn (1922–1996) on the *Structure of Scientific Revolutions*. Kuhn argued that scientific textbook accounts of the history of science as a wholly progressive series of discoveries are false for scientific revolutions. His work has suggested that changes of method across revolutions undercut attempts to apply common standards to evaluate prerevolution and postrevolution theories.

Kuhn also criticized the methodological ideas of Karl Popper (1902–1994). Popper had asked the question of what distinguishes ("demarcates") scientific hypotheses from nonscientific hypotheses. He emphasized that science proceeds by testing hypotheses against empirical data, and thus located the characteristic of scientific hypotheses in their empirical testability. Popper's basic view of testing a hypothesis against data was to derive predictions from the hypothesis and see if they matched the data (conjectures and refutations). If the data does not match the predictions, they falsify the hypothesis.

This led Popper to postulate that scientific hypotheses must be falsifiable. Popper's falsifiability criterion has been very influential, arguably more outside of the philosophy of science than inside. Kuhn objected to the falsifiability concept because, according to him, history shows that scientists do not subject major scientific theories (or paradigms) to falsification. Instead, scientists view a mismatch between theory and data as an anomaly, a puzzle to be resolved by further research. Many philosophers of science took Kuhn's moral to be that logic-based analyses of scientific method cannot capture the dynamics of major scientific change. Scientific revolutions would instead be determined by complex sociopolitical processes within the scientific community, played out within the specific historical context. Modern methodologists aim to avoid both the extremes of a context-free universal scientific logic on the one hand, and an entirely context-specific study of particular historical episodes on the other.

METHOD IN THE SCIENTIFIC REVOLUTION

Two topics of inquiry held center stage during the scientific revolution: the traditional problems of astronomy, and the study of gravity as experienced by bodies in free fall near the surface of the earth. Johannes Kepler (1571–1630) proposed that the predictive empirical equivalence between geocentric and heliocentric world systems that holds in principle could be offset by appeal to physical causes (Jardine 1984). He endorsed the appeal by Nicolas Copernicus (1473–1543) to the advantage offered his system from agreeing measurements of parameters of the earth's orbit from several retrograde motion phenomena of the other planets (1596/1981). In his classic marshaling of fit to the impressive body of naked eye instrument observation data by Tycho Brahe (1546–1601), Kepler appealed to this advantage as well as qualitative intuitions about plausible causal stories and intuitions about cosmic harmony to arrive at his ellipse and area rules (1609/1992). He later arrived at his harmonic rule (1619/1997). His Rudolphine Tables of 1627 were soon known to be far more accurate than any previously available astronomical tables (Wilson 1989).

Galileo Galilei (1564–1642) described his discovery of Jupiter's moons and exciting new information about our moon in the celebrated report of his telescope observations (1610/1989). His later observations of phases of Venus provided direct observational evidence against Ptolemy's system, though not against Tycho's geoheliocentric system. This was included in his argument for a Copernican heliocentric system in his famously controversial *Dialogue* (1632/1967).

Galileo's study of gravity faced the challenge that because of complicating factors such as air resistance one could not expect the kind of precise agreement with measurement that was available in astronomy. In his celebrated *Two New Sciences* (1638/1914), Galileo proposed uniformly accelerated fall as an exact account of idealized motion that would obtain in the absence of any resistant medium, even though the idealization is impossible to actually implement. He argues that the perturbing effects of resistance are too complex to be captured by any theory, but that the considerations he offers, including inclined plane experiments that minimize the effects of resistance, support his idealized uniformly accelerated motion as the principal mechanism of such terrestrial motion phenomena as free fall and projectile motion.

An important part of what distinguishes what we now characterize as the natural sciences is the method exemplified in the successful application of universal

gravity to the solar system. Isaac Newton (1642–1727) characterizes his laws of motion as accepted by mathematicians and confirmed by experiments of many kinds. He appeals to propositions inferred from them as resources to make motion phenomena measure centripetal forces. These give systematic dependencies that make the areal law for an orbit measure the centripetal direction of the force maintaining a body in that orbit, that make the harmonic law for a system of orbits about a common center, and that make the absence of orbital precession (not accounted for by perturbations) for any such orbit, measure the inverse square power for the centripetal force. His inferences to inverse-square forces toward Jupiter, Saturn, and the sun from orbits about them are inferences to inverse-square centripetal acceleration fields backed up by such measurements.

Newton's moon-test shows that the length of a seconds pendulum at the surface of the earth and the centripetal acceleration of the moon's orbit count as agreeing measurements of a single earth-centered inverse-square acceleration field. On this basis Newton identified the force maintaining the moon in orbit with terrestrial gravity. His first two rules endorse this inference. Rule number one states "no more causes of natural things should be admitted than are both true and sufficient to explain their phenomena" (Newton 1726/1999, p. 794). Rule number two adds that, therefore, "the causes assigned to natural effects of the same kind must be, so far as possible, the same" (Newton 1726/1999, p. 795).

Newton argues that all bodies gravitate toward each planet with weights proportional to their masses. He adduces a number of phenomena that give agreeing measurements of the equality of the ratios of weight to mass for bodies at equal distances from planets. These include terrestrial pendulum experiments and the moon-test for gravitation toward the earth, as well as the harmonic laws for orbits about them for gravitation toward Saturn, Jupiter, and the sun. They also include the agreement between the accelerations of Jupiter and its satellites toward the sun, as well as between those of Saturn and its satellites and those of the earth and its moon toward the sun.

His third rule endorses the inference that these all count as phenomena giving agreeing measurements of the equality of the ratios of weight to mass for all bodies at any equal distances from any planet whatsoever. Rule number three states that "those qualities of bodies that cannot be intended and remitted (i.e., qualities that cannot be increased and diminished) and that belong to all bodies on which experiments can be made should be taken as qualities of all bodies universally" (Newton 1726/1999, p. 795).

Newton's fourth rule added that "In experimental philosophy propositions gathered from phenomena by induction should be considered either exactly or very nearly true notwithstanding any contrary hypothesis until yet other phenomena make such propositioins either more exact or liable to exceptions" (Newton 1726/1999, p. 796).This rule was added to justify treating universal gravity as an established scientific fact, notwithstanding complaints that it was unintelligible in the absence of a causal explanation of how it results from mechanical action by contact.

Newton's inferences from phenomena exemplify an ideal of empirical success as convergent accurate measurement of a theory's parameters by the phenomena to be explained. In rule four, a mere hypothesis is an alternative that does not realize this ideal of empirical success sufficiently to count as a serious rival.

Rule four endorses provisional acceptance. Deviations count as higher order phenomena carrying information to be exploited. This method of successive corrections guided by theory mediated measurement led to increasingly precise specifications of solar system phenomena backed up by increasingly precise measurements of the masses of the interacting solar system bodies.

This notion of empirical success as accurate convergent theory mediated measurement of parameters by empirical phenomena clearly favors the theory of general relativity of Albert Einstein (1879–1955) over Newton's theory (Harper 1997). Moreover, the development and application of testing frameworks for general relativity are clear examples of successful scientific practice that continues to be guided by Newton's methodology (Harper 1997, Will 1986 and 1993). More recent data such as that provided by radar ranging to planets and lunar laser ranging provide increasingly precise post Newtonian corrections that have continued to increase the advantage over Newton's theory that Newton's methodology would assign to general relativity (Will 1993).

HYPOTHETICO-DEDUCTIVISM

In the preface to his *Treatise on Light*, Christian Huygens (1629–1695) provided a nice characterization of the hypothetico-deductive (H-D) alternative to Newton's method:

> There will be seen in it demonstrations of those kinds which do not produce as great a certitude

> as those of Geometry, and which even differ very much therefrom, since whereas the Geometers prove their Propositions by fixed and incontestable Principles, here the Principles are verified by the conclusions to be drawn from them; the nature of these things not allowing of this being done otherwise. It is always possible to attain thereby to a degree of probability which very often is scarcely less than complete proof. To wit, when those things which have been demonstrated by the Principles that have been assumed correspond perfectly to the phenomena which experiment has brought under observation; especially when there are a great number of them, and further, principally, when one can imagine and foresee new phenomena which ought to follow from the hypotheses which one employs, and when one finds that therein the fact corresponds to our prevision.
>
> (HUYGENS 1690/1962, P. VI AND VII)

Thus H-D method construes empirical success as success in prediction. The limitation of empirical success to prediction alone has suggested to some philosophers of science that distinguishing between theories that agree on predictions would have to be based on nonempirical criteria.

PREDICTED FIT TO FUTURE DATA

Given plausible assumptions about errors in data, a model that fits a given body of data too closely is likely to be tracking random errors in the data in addition to the lawlike phenomenon under investigation. Statisticians refer to this as "overfitting the data." They have designed many criteria to reveal cases where a simpler model has better expected fit-to-future data generated by repetitions of an experiment than a more complex model that better fits the data so far. Among philosophers of science, Malcolm Forster and Elliott Sober have appealed to the Akaike Information Criterion to challenge the assumption that fit-to-past data exhausts the criteria for scientific inference. This criterion is not sufficient to recover Newton's method (Myrvold and Harper 2002). The extent to which other such proposals can recover Newton's method is an open question.

BAYESIAN METHODS

Central to the Bayesian methods is epistemic probability, a rational agent's degree of belief. A number of arguments have been put forward to defend the probability axioms as coherence conditions for rational degrees of belief, in analogy to the way logical consistency can be taken as a coherence condition for rational acceptance. Dutch book arguments have shown that degrees of belief violating the probability axioms would assign positive expectations to each bet in a system of bets and conditional bets that would result in sure loss if they were all made together. A number of other arguments for this synchronic condition on rational degrees of belief have been advanced (particularly by Frank Plumpton Ramsey, Leonard J. Savage, Abner Shimony, Bas van Fraassesn, Richard T. Cox, Irving John Good, and J. Aczel).

David Lewis (1941–2001) provided a diachronic Dutch book argument (published in Teller 1976) to defend the Bayesian conditionalization learning model, according to which assigning new degrees of belief given by $P'(B) = P(B\&A)/P(A)$ is the appropriate response to a learning experience in which the total relevant empirical input is to accept A as new evidence. In 1984 van Fraassen (1941–) extended this diachronic Dutch book argument to defend a condition he called reflection. His proposal to treat the reflection condition as a constraint on degrees of belief that could be counted as rational has led to much controversy.

One central Bayesian theme has been to investigate conditions under which evidence leads to convergence of opinion. Bruno de Finetti (1906–1985) specified conditions that would lead Bayesian agents, who update by repeated conditionlization on the outcomes of the same observations, to converge toward agreement in their degrees of belief, however otherwise divergent their prior degrees of belief may have been (1937/1980). Brian Skyrms (1990) has given what is probably the most general possible version of de Finetti's condition for convergence.

In 2003 Wayne Myrvold (1963–) argued that, for Bayesians, the degree to which a hypothesis unifies phenomena contributes to the degree to which these phenomena support the hypothesis. This suggests that Bayesians can recover important aspects of Newton's method. It may well be that investigating the representation of Newton's method of provisional acceptance in a Bayesian model will result in enriching the Bayesian framework to make it offer more resources for illuminating scientific method.

CAUSATION, CORRELATION, EXPERIMENTATION

In his famous methods (1843), John Stuart Mill (1806–1873) combined ideas about causal inference previously

proposed by John Duns Scotus (1265/66–1308), William Ockham (1280–1349) and Francis Bacon (1561–1626). The work of twentieth century statisticians such as Jerzy Neyman (1894–1981), Karl Pearson (1857–1936), and Ronald A. Fisher (1890–1962) addressed two major shortcomings of Mill's method.

First, Mill assumed that we would observe deterministic causal relationships: Given the cause, the effect must follow every time. However, in a complex situation we typically do not have a complete specification of all operative causes, so we expect to observe trends rather than necessary relationships. For example, although smoking causes lung cancer, it does not do so in every person, because people's physiology varies. Rather, what we observe is a strong association between smoking and lung cancer: Among smokers, the incidence of lung cancer is much higher than among nonsmokers. To define precisely the intuitive notion of "strong association," statisticians developed the concept of correlation, which defines degrees of association (DeGroot 1975).

A second deficiency in Mill's methods is that they fail in the presence of common causes (*confounders* in statistical terminology). For example, suppose we observe that children who play violent video games are more prone to aggressive behavior than children who do not. Mill's logic would lead us to infer that playing violent video games causes aggressive behavior. But another possibility is that the correlation is because of personality traits: that children with an aggressive nature are drawn to violent video games and tend toward aggressive behavior; a preference for violent video games does not cause the behavior, but is merely a symptom of preexisting aggressive tendencies. If this alternative explanation is true, then Mill's methods lead us to the wrong conclusion. The policy implications are significant: If there is a direct causal relationship between video games and aggressive behavior, we expect to reduce aggressive behavior by restricting the availability of video games. But if personality is the underlying common cause of both, restricting access to video games should not decrease aggressive behavior.

A great advance for the problem of unobserved common causes was Fisher's revolutionary idea of the randomized experiment. Suppose that we have the ability to randomly assign half of a group of children to playing violent video games (the treatment group) and the other half to playing something else (the control group). For example, we might flip a coin for each participating child to make the assignment. Then we expect that personality traits, such as a tendency to aggression, would be randomly distributed in each half so that the children playing the video games would, on average, have no more aggressive personalities than the children playing something else. Under those circumstances, if we still find that significantly more of the video game players engage in aggressive behavior than the children playing something else, we can infer a direct causal relationship.

The idea of using randomization to rule out unobserved common causes has been applied in countless practical problems of causal inference, from clinical studies of the effectiveness of medical treatments to experiments for agricultural methods. It has been a most effective tool for addressing the problem of unobserved common causes that besets many of the traditional philosophical proposals for causal inference.

The power of randomization is available only when we have the ability to experimentally create the conditions we wish to investigate. In many settings of interest, we cannot perform experiments but can only passively gather data (these are called "observational studies" in statistics). A prominent physical science based on passive observation is astronomy. Many examples occur in the social sciences and economics. For instance, an economist cannot randomly assign inflation rates to various countries to study how inflation affects employment. A recent set of examples comes from computer science: While many companies gather vast amounts of data about their customers and the transactions they engage in, they rarely have the ability to assign customers randomly to various conditions (e.g., household income).

Philosophers continued to refine their understanding of the relationship between correlation and causation in nonexperimental settings. The work of Hans Reichenbach (1891–1953), published in 1956, was seminal. Reichenbach expounded the common cause principle: roughly, for every correlation between two events A and B, there is some causal explanation that posits either that one is a cause of the other (e.g., A causes B) or that A and B share a common cause. Reichenbach argued that the assumption that significant associations or correlations have causal explanations is deeply ingrained in our scientific and everyday reasoning. Another important concept of Reichenbach was the notion of screening off. The purpose of this concept is to capture the distinction between immediate and intermediate causes in terms of correlations.

For example, suppose that tar content in lungs is the direct cause of cancer, while smoking directly causes tar to accumulate in the lungs, and thereby indirectly causes lung cancer. Then we would observe a correlation between smoking and lung cancer; but knowing the tar

content of the lung would make smoking irrelevant to lung cancer. By contrast, even if we knew whether a subject smokes, the tar content of one's lungs would still be relevant to, or correlated with, the subject getting lung cancer. In Reichenbach's terms, information about tar content screens off information about smoking from conclusions about lung cancer. Because tar content screens off smoking from lung cancer, but not vice versa, Reichenbach suggested that such evidence rules out smoking as a direct cause of lung cancer, and allows us to infer that the effects of smoking are mediated through tar in the lungs.

The philosophers of science—Peter Spirtes, Clark Glymour, and Richard Scheines—developed Reichenbach's ideas about the relationships between correlation and causation using the framework of causal graphs or diagrams (Spirtes 1993). A causal graph is an intuitive representation of causal relationships, in which direct causes are connected with their effects by arrows pointing from cause to effect.

Using the language of causal graphs, Spirtes, Glymour, and Scheines gave a precise formulation of Reichenbach's precept that direct causes screen off indirect ones, known as the Markov condition (I-map in computer science terminology). The common cause principle—that there is no correlation without causation—can be formulated as another principle about diagrams, termed faithfulness (perfect I-map in computer science terminology). Given these principles relating causation and correlation, it is possible to characterize when valid inferences about causal relationships can be drawn from passive observation of associations. The theory is powerful and precise enough to develop computer programs that perform these inferences automatically (the TETRAD system, for instance). With such a program, we can analyze the kind of large datasets that we find in practice, realizing the vision of Bacon and Mill of applying causal inference methods to extensive observation histories.

In computer science, causal diagrams (often called Bayes Nets) have been firmly established as a scheme to capture and reason about associations and causal relationships, giving rise to thriving commercial developments with many practical applications (Pearl 1988, 2000). Econometrics, the study of statistical methods for economic problems, has a rich tradition of developing methods for nonexperimental causal inference going back to the early twentieth century (path diagrams and structural equation models). It turns out that many of these ideas and techniques can be seen as instances of causal diagram methods (Pearl 2000). While the theory of causal inference from passive observation is not yet as firmly established as the methodology based on randomization, at the beginning of the twenty-first century we see a common framework emerging shared and sustained by philosophy, computer science, and economics.

See also Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Philosophy of Statistical Mechanics; Scientific Revolutions.

Bibliography

Aczel, J. *Lectures on Functional Equations and Their Applications*. New York: Academic Press, 1966.

Cox, R. *The Algebra of Probable Inference*. Baltimore, MD: John Hopkins Press, 1961.

de Finetti, B. "Foresight: Its Logical Laws, Its Subjective Sources" (1937). Translated by H. E. Kyburg and H. Smokler. In *Studies in Subjective Probability*. Huntington, NY: Kreiger, 1980.

DeGroot, Morris H. *Probability and Statistics*. Reading, MA: Addison-Wesley, 1975.

Earman, J. *Bayes or Bust?: A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT press, 1992.

Forster, M., and E. Sober. "How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions." In *British Journal for The Philosophy of Science* 45 (1994): 1–35.

Galilei, G. *The Sidereal Messenger* (1610). Translated by A. Van Helden. Chicago: University of Chicago Press, 1989.

Galilei, G. *Dialogue concerning the Two Chief World Systems* (1632). Translated by Stillman Drake. Los Angeles: University of California Press, 1967.

Galilei, G. *Two New Sciences*. Translated by H. Crew and A. De Salvio. New York: Dover, 1914.

Good, I. J. *Probability and The Weighing of Evidence*. London: Griffin, 1950.

Harper, W. L. "Isaac Newton on Empirical Success and Scientific Method." In *The Cosmos of Science: Essays of Exploration*, edited by J. Earman and J. D. Norton. Pittsburgh, PA: University of Pittsburgh Press, 1997.

Harper, W. L., and C. A. Hooker, eds. *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Vol. 1. Dordrecht; Boston: D. Reidel, 1976

Huygens, C. *Treatise On Light* (1690). Translated by S. P. Thompson. New York: Dover, 1962.

Jardine, N. *The Birth of History and Philosophy of Science: Kepler's "A Defense of Tycho against Ursis*. Cambridge, MA: Cambridge University Press, 1984.

Jeffreys, H. *Scientific Inference*. 3rd ed. Cambridge, MA: Cambridge University Press, 1973.

Kepler, J. *The Harmony of The World* (1619). Translated by E. J. Aiton, A. M. Duncan, and J. V. Field. Philadelphia: American Philosophical Society, 1997.

Kepler, J. *New Astronomy* (1609). Translated by W. H. Donahue. New York: Cambridge University Press, 1992.

Kepler, J. *The Secret of the Universe* (1596). Translated by A.M. Duncan. New York: Abaris Books, 1981.

Kuhn, T. S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Kyburg, H. E. *Science & Reason*. New York: Oxford University Press, 1990.

Myrvold, W. C. "A Bayesian Account of The Virtue of Unification." *Philosophy of Science* 70 (2003): 399–423.

Myrvold, W. C., and W. Harper. "Model Selection, Simplicity, and Scientific Inference." *Philosophy of Science* 69 (2002): 135–149.

Newton, I. *Mathematical Principles of Natural Philosophy*. 3rd ed. (1726). Translated by I. B. Cohen and A. Whitman. Los Angeles: University of California Press, 1999.

Pearl, J. *Causality*. San Mateo, CA: Morgan Kaufmann, 2000.

Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann, 1988.

Popper, K. R. *Conjectures and Refutations: The Growth of Scientific Knowledge* (1962). New York: Harper & Row, 1963.

Popper, K. R. *The Logic of Scientific Discovery* (1959). New York: Harper & Row, 1965.

Ramsey, F. P. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*. London: Kegan Paul, 1931.

Reichenbach, H. *The Direction of Time*. Berkeley, CA: University of Los Angeles Press, 1956.

Reichenbach, H. *The Theory of Probability*. London: Cambridge University Press, 1949.

Savage, L. J. *Foundations of Statistics*. New York: Wiley, 1954.

Shimony, A. *Search for a Naturalistic World View: Scientific Method and Epistemology*. Vol. 1. Cambridge, U.K.: Cambridge University Press, 1993.

Skyrms, B. *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard, 1990.

Spirtes, P., C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. New York: Springer Verlag, 1993.

Teller, O. "Conditionalization, Observation, and Change of Preference." In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, edited by W. L. Harper and C. A. Hooker. Dordrecht; Boston: D. Reidel 1 (1976): 205–257

van Fraassen, B. "Belief and The Will." *Journal of Philosophy* 81 (1984): 235–256.

van Fraassen, B. "Calibration: A Frequency Justification for Personal Probability." In *Physics, Philosophy, and Psychoanalysis: Essays in Honor of Adolf Grünbaum*, edited by R. S. Cohen and L. Laudan. Dordrecht: Reidel, 1983.

Will, C. M. *Theory and Experiment in Gravitational Physics*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1993.

Will, C. M. *Was Einstein Right? Putting General Relativity to the Test*. New York: Basic Books, 1986.

Wilson, C. "Predictive Astronomy in The Century after Kepler." In *The General History of Astronomy: Planetary Astronomy from the Renaissance to the Rise of Astrophysics, Part A: Tycho Brahe to Newton*. Vol. 2, edited by R. Taton and C. Wilson. Cambridge, U.K.: Cambridge University Press, 1989.

Willam Harper (2005)
Oliver Schulte (2005)

SCIENTIFIC REALISM

Scientific realism is a philosophical view about science that consists of three theses:

> *The metaphysical thesis*: The world has a definite and mind-independent structure.
>
> *The semantic thesis*: Scientific theories should be taken at face value. They are truth-conditioned descriptions of their intended domain, both observable and unobservable. Hence, they are capable of being true or false. The theoretical terms featured in theories have putative factual reference.
>
> *The epistemic thesis*: Mature and predictively successful scientific theories are well confirmed and (approximately) true of the world. So the entities posited by them, or entities very similar to those posited, inhabit the world.

METAPHYSICS

Let us call the first thesis of scientific realism *metaphysical realism*. What exactly is involved in the claim of mind-independence? One way to construe the opposite claim that the world is mind-dependent, along the lines of traditional idealism and phenomenalism, is to argue that the world consists of mental entities, be they ideas or actual and possible sense-data. Thus understood, mind-dependence is a thesis about the kind of stuff that makes up the world. The insistence of scientific realism on metaphysical realism might be thought of as opposing this idealist or phenomenalist doctrine. It might be seen as a declaration that there is nonmental stuff in the world and, in particular, that the entities posited by scientific theories are material. This view is certainly part of the realist construal of mind-independence, but there is more.

There is another, more complicated and interesting, way to construe the claim that the world is mind-dependent. This way centers not on what *types* of entity exist (whether they are material or mental or what have you) but rather on what is involved in claiming that they *exist*. There is a long antirealist philosophical tradition according to which it does not make sense to assert the existence (or reality) of some entities unless we understand this assertion to mean that … , where the ellipsis is filled with a suitable epistemic/conceptual condition. Much like realism, these views (call them varieties of verificationist antirealism) oppose idealism and phenomenalism. They entail the position (or at least are consistent with the

claim) that material objects are real (be they the middle-sized entities of common sense or unobservable entities).

The substantive disagreement between this antirealist tradition and realism is the *sense* of existence. Verificationist antirealism makes the world (or a set of entities) mind-dependent in a more sophisticated sense: What there is in the world is determined by what can be known to exist (verified to exist, rationally accepted as existing, or the like). Hence it forges a logical-conceptual link between what there is in the world and what is affirmed as existing on the basis that it satisfies suitable epistemic conditions. Accordingly, the realist claim of mind-independence should be understood as logical or conceptual independence: What the world is like does not logically or conceptually depend on the epistemic means and conceptualizations used to get to know it. Scientific realism allows for the possibility of a *divergence* between what there is in the world and what is issued as existing by a suitable set of conceptualizations and epistemic conditions. Verificationist antirealism precludes this possibility of divergence a priori by advancing an epistemic conception of truth. No matter what the details of this conception are, the key idea is that truth is conceptually linked with epistemic conditions so tightly that a theory cannot be false even though epistemically justified (because it meets the relevant epistemic condition, for example, being under ideal circumstances theoretically justified or warrantedly assertable). Typically, realists honor the possibility of divergence by adopting a non-epistemic conception of truth (the standard candidate for which is the correspondence theory of truth).

Why should scientific realism incorporate the claim of mind-independence? Why, that is, cannot someone who accepts the reality of unobservable entities but regards them as mind-dependent (in the above sense) be a scientific realist? Ultimately at stake in the debate over scientific realism is a robust sense of objectivity, that is, a conception of the world as the arbiter of our changing and evolving conceptualizations of it. Scientific realism honors this conception by claiming that the world is mind-independent. The kernel of its metaphysical thesis is that science is in the business of discovering what a world that is not of our making is like. This thesis implies that if the natural kinds posited by theories exist at all, they exist objectively, that is, independently of our ability to be in a position to know them, verify them, recognize them, etc., and hence that natural kinds, if anything, make scientific theories true. This robust sense of objectivity contradicts verificationist antirealism. It also blocks a number of projectivist or social constructivist views about science from being realist. In the view of scientific realism, scientific theories and scientific theorizing in general, instead of projecting (or worse, socially constructing) the structure of the world, discover and map out an already structured, mind-independent world.

SEMANTICS

Let us call the second thesis of scientific realism, the view that scientific theories should be taken at face-value, *semantic realism*. This view too was motivated by problems with verificationism.

Verificationism, at least in its traditional form as defended by the logical positivists, runs together two separate issues: the evidential basis for the truth of an assertion and the semantic relation of reference or denotation. It thereby conflates the issue of what constitutes evidence for the truth of an assertion with the issue of what makes the assertion true. This conflation was the product of concerns about the meaning of theoretical terms. Some empiricists thought that since the meaning of theoretical terms is not given directly in experience, these terms are semantically suspect. Hence, empiricists (even hard-core positivists like Ernst Mach) sought to show that theoretical statements and terms are parasitic on observational statements and terms.

This line of thought led to *reductive empiricism*, which treats theoretical statements as being disguised talk about observables and their actual (and possible) behavior. Interestingly, this view is consistent with the claim that theoretical statements have truth-values, but it understands their truth-conditions reductively: Their truth-conditions can be fully captured in an observational vocabulary. Hence, theoretical statements are ontologically innocuous: They do not refer to unobservable entities, and so imply no commitments to unobservable entities. Despite the heroic efforts of many empiricists (including the early Rudolf Carnap), all attempts to translate theoretical terms into observational terms have patently failed. As a result, empiricism became liberal. It admitted that theoretical terms and statements have excess content that cannot be fully captured by anyreference to observable entities and phenomena.

If evidence-conditions and truth-conditions are kept apart, verificationism loses its bite. Semantic realism, simply put, says that there should not be two semantic standards, one for observational statements and another for theoretical ones. Observational statements, as well as theoretical statements, are true if and only if their truth-conditions obtain. Hence, theoretical terms, no less than observational terms, have putative factual reference. If

theoretical statements cannot be given truth-conditions in an ontology that dispenses with theoretical entities, a full and just explication of scientific theories simply requires commitment to irreducible unobservable entities, no less than it requires commitment to observable entities.

Instrumentalism claims that theories should be seen as (useful) instruments for organizing, classifying, and predicting observable phenomena. So the "cash value" of scientific theories is fully captured by what theories say about the observable world. Faced with the semantic realist challenge that theoretical assertions are meaningful and purport to describe an unobservable reality, instrumentalism took refuge in Craig's theorem and claimed that theoretical commitments in science are dispensable: Theoretical terms can be eliminated en bloc without loss in the deductive connections between the observable consequences of the theory. If this is so, then the very question of whether theoretical terms can refer to unobservable entities evaporates. This challenge led Carl Hempel (1958) to formulate what he called "the theoretician's dilemma." If the theoretical terms and the theoretical principles of a theory do not serve their purpose of a deductive systematization of the empirical consequences of a theory, then they are dispensable (unnecessary). But by Craig's theorem, even if they do serve their purpose, they can still be dispensed with. Hence, the theoretical terms and principles of any theory are dispensable.

Is the theoretician's dilemma compelling? Note first that the very idea of this dilemma rests on a sharp distinction between theoretical terms and observational ones. This dichotomy was severely challenged in the 1960s, when Pierre Duhem's view that all observation is theory-laden resurfaced. Along with it came the view that, strictly speaking, there are no observational terms. But even if the dichotomy is accepted, instrumentalism based on Craig's theorem collapses. It is implausible to think of theories as establishing only a deductive systematization of observable phenomena. Theories also offer inductive systematizations in the sense that theories can be used to establish inductive connections among observable phenomena: They function as premises in inductive arguments and, together with other premises concerning observable phenomena, yield conclusions that refer to observable phenomena. Seen as aiming to establish inductive connections among observables, theories are indispensable. There followed a battery of indispensability arguments, fostered by Sellars (1963) and Quine (1960) among others, suggesting that theoretical terms are indispensable in any attempt to formulate a powerful and efficacious system of laws and to explain why observable entities obey the empirical laws they do.

Semantic realism opposes both instrumentalism and reductive empiricism. It renders scientific realism an "ontologically inflationary" view. Understood realistically, theories admit of a literal interpretation, that is, an interpretation according to which the world is populated by a host of unobservable entities and processes. Semantic realism is not contested any more. All sides of the debate take theoretical discourse to be irreducible and contentful. It should be clear from the above discussion, however, that making semantic realism the object of philosophical consensus was no trivial feat.

EPISTEMOLOGY

Let us call the third thesis of scientific realism *epistemic optimism*. Its thrust is that science can and does deliver theoretical truth no less than it can and does deliver observational truth. One can grant semantic (even metaphysical) realism and yet remain epistemically skeptical or agnostic toward scientific theories. This agnostic stance has appealed to empiricists who have come to terms with the collapse of instrumentalism and reductive empiricism. An argument for the realist interpretation of scientific theories is not ipso facto an argument for *believing* in the existence of the entities those theories posit and in the truth of what they say of them.

Can the epistemic thesis be avoided? Some realists, notably Alan Musgrave (1999), think that scientific realism is an exclusively axiological thesis: Science aims for true theories. There is clear motivation for this axiological approach: Even if *all* theories scientists ever came up with were false, scientific realism would not thereby be threatened. There are, however, inevitable philosophical worries about the axiological characterization of realism. First, it seems rather vacuous. Realism is rendered immune against the serious criticism stemming from the empirical claim that science has a poor record in tracking the truth. Second, aiming at a goal (truth) whose achievability by the scientific method is left unspecified makes the supposed regulative role of the goal totally mysterious. Finally, we lose all the excitement of the realist claim that science engages in a cognitive activity that pushes back the frontiers of ignorance and error. Other realists, notably Jarrett Leplin (1997), do take the epistemic thesis to be part of scientific realism, but argue for a minimal or thin version of it: There are possible empirical conditions that would warrant attributing some measure of truth to theories. The problem with this minimal account is that, in the end, it cannot provide a rational or warranted basis

for belief in the unobservable entities posited by science (and the assertions made about them).

Naturally, the scope of the epistemic thesis need not (and should not) be universal. Scientific realists need not take current science uncritically. They need not commit themselves to everything that current theories assert. They can have a differentiated attitude toward the theoretical constituents of modern science: Some of them are better supported by the evidence than others; some play an indispensable explanatory role, while others do not; some contribute to the successes of theories, while others do not. But we should not lose sight of the general philosophical issue at stake, which is this: Are there good reasons to believe that science cannot achieve theoretical truth? That is, are there good reasons to believe that, given that we understand the theoretical statements of scientific theories as genuine propositions, we can never be in a warranted position to claim that they are true (or at least, more likely true than false)? The *epistemic thesis* denies that there are such good reasons and defends the claim that the ampliative-abductive methods of science are reliable and can justify/support theoretical assertions. Hence, science has succeeded in tracking truth. To be sure, this success requires a certain amount of epistemic luck: It is not a priori true that science has been, or has to be, successful in truth tracking. If science does succeed in truth tracking, this is a radically contingent fact about how the world is and how science and its method have managed to latch onto it.

The prime argument in favor of the epistemic thesis has come to be known as "the no-miracles argument." It is an abductive argument, or inference to the best explanation. Jack Smart (1963) argued against instrumentalists that they must believe in cosmic coincidence. On the instrumentalist view of theories, a vast number of ontologically disconnected observable phenomena are "connected" only by a purely instrumental theory: These phenomena just *happen* to be related to one another in the way suggested by the theory. Scientific realism, in contrast, leaves no space for a cosmic-scale coincidence: It is *because* theories are true and *because* the unobservable entities posited by them exist that the phenomena are related to one another as they are. Smart's key point was that scientific realism (and its concomitant view of science) should be accepted because it offers the best explanation of why the observable phenomena are as scientific theories predict them to be.

Hilary Putnam (1975) and Richard Boyd (1973) argued that inference to the best explanation is the very method scientists use to form and justify their beliefs in unobservable entities, and that realism should be seen as an overarching empirical hypothesis deriving support from the fact that it offers the best explanation of the success of science. The no-miracles argument found pithy expression in Putnam's encapsulation: "The positive argument for realism is that it is the only philosophy that does not make the success of science a miracle" (1975, p. 73). A key element of the realists' epistemic optimism comes from the fact that some theories, because they yield novel predictions, can serve as "prophets for us," as Duhem put it. Only on a realist understanding do novel predictions about phenomena come as no surprise.

How exactly does the no-miracles argument support the epistemic thesis? Though this issue has been extensively debated, the role of the no-miracles argument in the realism debate is quite complex. To a good approximation, the argument should be seen as a grand inference to the best explanation. It is a philosophical argument that aims to defend the reliability of scientific methodology in producing approximately true theories. The argument proceeds in two steps. The first is that we accept as approximately true the theories that are implicated in the (best) explanation of the *instrumental* reliability of first-order scientific methodology. The second step is that since these theories have typically been arrived at by means of inference to the best explanation, such inference is reliable. The main strength of the no-miracles argument rests on the first part of the argument. Coming after more concrete types of explanatory reasoning that occur all the time in science, the argument suggests that it is reasonable to accept certain theories as approximately true, at least as concerns their components that guided predictions. These successful instances of explanatory reasoning in science provide the basis for the grand abductive argument. However, the no-miracles argument is not just a generalization over the scientists' abductive inferences. Although itself an instance of the method that scientists employ, it aims at a much broader target, specifically, to defend the thesis that inference to the best explanation (a *type* of inferential method) is reliable. This relates to the second step of the argument. What makes the no-miracles argument distinctive as an argument for realism is that it defends the claim that theoretical truth is achievable. The second step of the argument seeks to secure this claim. It is reasonable to believe that abductive reasoning is reliable, since it tends to generate approximately true theories.

There are two challenges to scientific realism. The first relies on the claim that the evidence underdetermines theories and is discussed in a separate entry. The

second argument is the so-called pessimistic induction. As Larry Laudan (1984) pointed out in developing this argument, the history of science is replete with theories that were once considered empirically successful and fruitful but that turned out to be false and were abandoned. If the history of science is a wasteland of aborted best theoretical explanations of the evidence, then it might well be that current best explanatory theories will travel the route to this wasteland in due course. The best defense of realism against the pessimistic induction has been to try to reconcile the historical record with some form of realism. To do this, realists need to be more selective in what they are realists about.

A claim that emerged with some force in the 1990s is that theory-change is not as radical and discontinuous as the opponents of scientific realism have suggested. Realists such as Philip Kitcher (1993) and Stathis Psillos (1999) have sought to ferret out the theoretical components of abandoned scientific theories that essentially contributed to their successes, separate them from other idle components, and demonstrate that the components making essential contributions to the empirical success of the theories were retained in subsequent theories of the same domain. In such a scenario, the fact that our current best theories may be replaced by others does not necessarily undermine scientific realism. All that such evolution shows is that we cannot get at the truth all at once, and that our judgments from empirical support to approximate truth should be more refined and cautious in that they should commit us only to the theoretical components that enjoy evidential support and contribute to the empirical successes of the theory. Realists ground their epistemic optimism on the fact that newer theories incorporate many theoretical components of their superseded predecessors, especially those components that have led to empirical successes. The substantive continuity in theory-change suggests that a rather stable network of theoretical principles and explanatory hypotheses has emerged, survived revolutionary changes, and become part and parcel of our evolving scientific image of the world.

Faced with the challenge of the pessimistic induction, other realists have sought to weaken realism. There have been two prominent strategies for weakening realism. The first is to opt for structural realism, and the second is to opt for entity realism. Structural realism, defended by John Worrall (1989), capitalizes on the fact that despite the radical changes at the theoretical level, successor theories have tended to retain the *mathematical structure* of their predecessors. It argues that theories can successfully represent the structure of the world even when they are wrong about the entities they posit. Despite its initial appeal, it turns out that this particular position is very difficult to defend. For one, the distinction between the mathematical structure of the theory and its theoretical content is not as clear-cut as it initially seems. For another, even if a sharp distinction is granted, it turns out that structural realism collapses the difference between the claim that a theory is true and the claim that it is empirically adequate.

Entity realism, defended by Nancy Cartwright (1983) and Ian Hacking (1983), accepts the existence of all sorts of unobservable entities but denies the truth of the theories in which descriptions of these entities are embedded. A major motivation for entity realism comes from laboratory life. Experimenters have good reasons to believe in specific unobservable entities, not because they accept the relevant theories, it is claimed, but rather because they *do* things with these entities. If these entities did not exist, the phenomena of the laboratory would be inexplicable. But can one be a realist about theoretical entities without also being a realist about the theories? In a sense, one can. For posited entities survive theory-change. For instance, scientists accept the existence of electrons even though their theoretical views about what electrons are have changed. So it appears that we can know *that* the electron is, even though we may not know *what* it is. But this cannot be fully right. We cannot assert that electrons are real, that is, that electrons are part and parcel of the furniture of the world, without also asserting that they have *some* of the properties attributed to them by our best scientific theories. So entity realism cannot be fully divorced from theory realism. In any case, the very same inferential process (inference to the best explanation) is involved in accepting the reality of an entity and in accepting the approximate correctness of some theoretical description of it.

SCIENTIFIC REALISM AND EMPIRICISM

Bas van Fraassen (1980) fostered a rivalry between scientific realism and empiricism with his influential doctrine of constructive empiricism. According to this view about science, (a) science aims at empirically adequate theories, and (b) acceptance of scientific theories involves belief only in their empirical adequacy (though acceptance involves more than just belief; it also involves commitment to the theory). Van Fraassen took realism to be, by and large, an *axiological* thesis: The aim of science is true theories. He supplemented it with a *doxastic* thesis:

Acceptance of theories implies belief in their truth. Seen in this way, realism and constructive empiricism *are* rivals. But, of course, a lot depends on whether an empiricist ought to be a *constructive* empiricist. There is no logical obstacle impeding an empiricist (who thinks that all knowledge ultimately stems from experience) from fostering methods that warrant belief in the truth of theories in a way that goes beyond belief in their empirical adequacy, and hence from being a scientific realist. Similarly, there is no logical obstacle impeding an empiricist from being stricter than constructive empiricism, for instance, by claiming that (a') the aim of science is unrefuted theories and (b') acceptance of a theory involves the belief only that it is unrefuted.

Constructive empiricism does set the boundaries of experience much farther afield than strict empiricism, and since what empiricism is, is not carved in stone, there is no logical obstacle to setting the boundaries of experience (that is, the reach of legitimate applications of scientific method) even farther afield, as realists demand. Indeed, as Hans Reichenbach (1938) noted, the key question is what kinds of methods are compatible with empiricism. Even if we grant, as we should, that all factual knowledge starts with experience, the boundaries of experience depend on the warrants of the methods employed. It is perfectly compatible with empiricism to accept ampliative methods and to accept the existence of unobservable entities on their basis. So there is no incompatibility between being an empiricist and being a scientific realist.

Van Fraassen tied empiricism to a sharp distinction between observable and unobservable *entities*. This, to be sure, is a step forward from the more traditional empiricist distinction between observational and theoretical terms and predicates. Drawing the distinction in terms of entities allows the description of observable entities to be fully theory-laden. Yet, van Fraassen insisted, even theoretically described, an entity does not cease to be observable if a suitably placed observer *could* perceive it with the naked eye.

Long before van Fraassen, Grover Maxwell (1962) denied this entity-based distinction, arguing that observability is a vague notion and that, in essence, *all* entities are observable under suitable circumstances. He based this view on the claim that "observability" is best understood as detectability by some means. If observability is thus understood, there are continuous degrees of observability, and hence there is no natural and nonarbitrary way to draw a line between observable and unobservable *entities*. Rebutting Maxwell's argument requires that naked-eye observations (which are required to tell us which entities are *strictly* observable) form a special kind of detection qualitatively set apart from any other way of detecting the presence of an entity (for example, with a microscope). Be that as it may, the issue is not whether the distinction between observable and unobservable entities can be drawn but what its epistemic relevance is: Why should the observable/unobservable distinction define the border between what is epistemically accessible and what is not?

In the end, scientific realism is better than constructive empiricism because (1) it does nor rely on a distinction of dubious epistemic significance, specifically, the observable/unobservable distinction, (2) it offers a better explanation of the empirical successes of science, and (3) it tallies better with actual scientific practice.

See also Realism; Underdetermination Thesis, Duhem-Quine Thesis.

Bibliography

Boyd, Richard. "Realism, Underdetermination, and the Causal Theory of Evidence." *Noûs* 7 (1973): 1–12.

Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford, U.K.: Clarendon Press, 1983.

Devitt, Michael. *Realism and Truth*. 2nd, rev. ed. Oxford: Blackwell, 1991. 1st ed., 1984.

Duhem, Pierre. *The Aim and Structure of Physical Theory*. Translated by Philip P. Wiener. Princeton, NJ: Princeton University Press, 1954. First published in 1906.

Fine, Arthur. *The Shaky Game*. Chicago: University of Chicago Press, 1986.

Hacking, Ian. *Representing and Intervening*. Cambridge, U.K.: Cambridge University Press, 1983.

Hempel, Carl. "The Theoretician's Dilemma: A Study in the Logic of Theory Construction." In *Concepts, Theories, and the Mind-Body Problem*, edited by Herbert Feigl, Michael Scriven, and Grover Maxwell. Minnesota Studies in the Philosophy of Science, no. 2. Minneapolis: University of Minnesota Press, 1958.

Kitcher, Philip. *The Advancement of Science*. Oxford, U.K.: Oxford University Press, 1993.

Laudan, Larry. *Science and Values*. Berkeley: University of California Press, 1984.

Leplin, Jarrett. *A Novel Defence of Scientific Realism*. Oxford, U.K.: Oxford University Press, 1997.

Maxwell, Grover. "The Ontological Status of Theoretical Entities." In *Scientific Explanation, Space, and Time*, edited by Herbert Feigl and Grover Maxwell. Minnesota Studies in the Philosophy of Science, no. 3. Minneapolis: University of Minnesota Press, 1962.

Musgrave, Alan. *Essays on Realism and Rationalism*. Amsterdam: Rodopi, 1999.

Psillos, Stathis. *Scientific Realism: How Science Tracks Truth*. London: Routledge, 1999.

Putnam, Hilary. *Mathematics, Matter and Method*. Cambridge, U.K.: Cambridge University Press, 1975.

Quine, Willard Van Orman. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Reichenbach, Hans. *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press, 1938.

Sellars, Wilfrid. *Science, Perception, and Reality*. Atascadero, CA: Ridgeview, 1963.

Smart, J. J. C. *Philosophy and Scientific Realism*. London: Routledge and Kegan Paul, 1963.

Van Fraassen, Bas C. *The Scientific Image*. Oxford, U.K.: Clarendon Press, 1980.

Worrall, John. "Structural Realism: The Best of Both Worlds?" *Dialectica* 43 (1989): 99–124.

Stathis Psillos (2005)

SCIENTIFIC REVOLUTIONS

Largely as the result of Thomas Kuhn's work, the concept of scientific revolution gains an importance in post-positivist philosophy of science that it lacks in the dominant logical empiricist tradition of the twentieth century. Kuhn's notion of scientific revolution becomes wedded to a historical relativism concerning scientific knowledge that many have sought to refute, or overcome with new accounts of knowledge that go beyond positivism and relativism.

THE CONCEPTION OF SCIENTIFIC REVOLUTION IN TRADITIONAL PHILOSOPHY OF SCIENCE

To set the context for these debates, it is useful to begin with the ordinary concept of scientific revolution and understand why it lacks fundamental epistemological significance in traditional philosophy of science. In ordinary parlance, a scientific revolution is a large-scale change in the fundamental concepts, theories, or methods that scientists in some area of inquiry emply to understand the course of nature (e.g., the Copernican revolution in astronomy). Such a change is also thought to be revolutionary in so far as it provokes similarly dramatic alterations in the way laypeople see the world. As such, the notion is obviously important to historians of science and popular culture. On the other hand, scientific revolution is not a central topic for the tradition of logical positivism (more broadly, logical empiricism) that generates the key figures, problems, and models of philosophy of science for most of the twentieth century.

In this tradition, the aim of philosophy of science is to provide analyses of the standards most vital to science as the best exemplar of empirical knowledge: the standards of scientific method, confirmation, prediction, falsification, explanation, truth, progress, observation, law, and theory. The philosopher's analyses are supposed to be timeless, normative, universal, non-historical, and non-empirical. To this end, logical empiricists employed the tools of logic and semantics to illuminate the a priori formal structure of all genuine scientific knowledge (such as explanation and confirmation). Science is identified with its most successful theories, which in turn are represented as finished bodies of propositions linked by logical and inferential relations connecting sense experience to the higher reaches of law and theory.

From this perspective, scientific revolutions alter the content of successful theories, but not the logic of scientific rationality and knowledge. Indeed, the empiricist's logical standards (e.g., Carl Gustav Hempel's deductive-nomological model of explanation, prediction and confirmation) provide the grounds for evaluating the scientific revolutions of Copernicus, Galileo, Johannes Kepler, Sir Isaac Newton, and Albert Einstein. This entire development could be reasonably represented as a logical, cumulative progress. On the philosopher's standards, this progress is one in which, for example, better confirmed theories of wider explanatory scope replace lesser predecessors, whose errors are corrected, and whose sound results are preserved and extended by their successors. The history of the best science(s) illustrates but does not alter the logic of scientific knowledge. So understood, the rationality of science makes it possible for humankind's best theories to converge on the truth concerning lawlike regularities in the world of observed phenomena and, perhaps, the underlying, unobservable entities and mechanisms causally responsible for these regularities.

These achievements of logical empiricism gain one of their last, most lucid and systematic reformulations in Hempel's *The Philosophy of Natural Science*. This work appeared in 1966 four years after Kuhn's *The Structure of Scientific Revolution (SSR)*. Of course many philosophers besides Kuhn challenge one or more of the presuppositions of traditional philosophy of science and reshape the debates in the post-positivist period (e.g. William Van Orman Quine, Wilfred Sellars, Norwood Hanson, Stephen Toulmin, Michael Scriven, Nelson Goodman, Paul Feyerabend, Mary Hesse, etc.). But Kuhn's challenge in *SSR* is probably unique in the avalanche of criticisms, rebuttals, and new approaches to the history and philosophy of science that it has provoked for decades. Much of this response focuses on Kuhn's notion of scientific revo-

lution and the incommensurability, relativism, and irrationalism it is taken to imply.

KUHN'S CONCEPTION OF SCIENTIFIC REVOLUTION

In effect, Kuhn mobilizes a new conception of the history of science, in which scientific revolution is fundamental and its nature contradicts the formal rationality, normativity, universality, logicism, and progressive cumulativity sought by logical empiricists, and still embraced in new forms by contemporary philosophers (e.g., scientific realists). The philosophical thrust of Kuhn's notion of scientific revolution can be tersely expressed as the following claim. It is in the very nature of (a) science that it undergo not simply changes in the content of its theories, but more fundamentally changes in the very language, problems, goals, and standards that (re)define science, the criteria of scientific knowledge, and membership in the scientific community. This sort of change is what Kuhn's conception of scientific revolution implies, an epistemological change in the requirements of scientific knowledge, explanation, proof, and confirmation. The claim that the essence of science is to generate scientific revolutions, in its own epistemological self-definition, seems like a general philosophical claim. But it is not an a priori claim, for Kuhn. Rather the claim is supposed to be justified by showing that it provides the best explanation of the actual development of science, which opens it up to criticism on this score. In any case, this argument gives history a central role in the evaluation of a philosophy of science.

Kuhn's view of scientific development turns on its division into periods of normal science marked by a normative consensus in the scientific community concerning how to conduct inquiry; and periods of scientific revolution, marked by the breakdown of this consensus. Revolutionary periods typically end when the scientific community is redefined on the basis of a new consensus that creates a different framework for normal science. The normative consensus required by normal science involves the existence of a paradigm that all experts accept as the basis of their research. A scientific revolution implies the dissolution of one paradigm and its eventual replacement by another. A paradigm is a concrete solution (e.g., Lavoisier's account of combustion) to a particular problem (why do some substances gain weight in combustion) that members of a scientific community commonly recognize as an exemplar of how to pursue inquiry in a wider domain of phenomena (chemical reactions); phenomena that may prove to be of the same or similar kind as the paradigm first treated. A group of inquirers only becomes a scientific community when their research generates a paradigm. As the central object of normative consensus, the paradigm guides practitioners in commonly recognizing what counts as a legitimate problem or phenomenon-to-be-explained in the domain of their science. It tells them what concepts, techniques, mechanisms, measurements, and standards must be present for a legitimate solution to the problem, a bona fide scientific explanation of it. Normal science is the research undertaken to articulate and extend the paradigm by solving a host of puzzles that arise in the attempt to reduce ever-wider phenomena to its terms.

In this process, the shared commitments of the scientific community grow and encompass the formulation of theories, laws, basic equations, standards of proof, mathematical techniques, and experimental procedures. In some contexts, Kuhn refers to this entire body of commitments as the paradigm. Normal science allows a cumulative progress of scientific knowledge, but it is progress within the paradigm, relative to its standards of puzzle solving and explanation. Normal science breaks down when the paradigm confronts anomalies. Anomalies are problems that it ought to be able to resolve, but over time cannot, and that motivate some practitioners to represent the problem, or attempt solutions in ways that abandon basic components of the paradigm and the normative consensus underlying the research tradition defined by it. For Kuhn, one of the best examples of scientific revolution is the abandonment of the premodern chemistry of the phlogiston theory and the theory of elective affinity, due to Lavoisier's oxygen theory of combustion and the new compositional paradigm of Daltonian chemistry.

KUHN'S CONCEPTS OF INCOMMENSURABILITY

Phlogiston chemistry succeeded in explaining many qualitative phenomena with a paradigm that posits the existence and properties of phlogiston (the presence of phlogiston solves the problem of why the metals have common metallic qualities lacking in their ores). But the phlogiston theory explained the combustion of a substance as a loss of phlogiston that implied weight loss. The phenomenon of weight gain in combustion constituted an anomaly for phlogiston theory because despite serious attempts, no phlogiston chemist succeeded in accounting for it within the constraints of this paradigm. As inquirers abandoned different components of the phlogiston paradigm, in order to accommodate the phenomena of combustion, the road was paved for a revolutionary transformation in the very concepts, language,

questions, techniques, data, values, aims, and standards at the heart of chemistry. In *SSR*, Kuhn stresses the discontinuities marked by scientific revolution and advances his most controversial claim that these discontinuities imply incommensurability between the paradigms or theories separated by scientific revolutions. Incommensurability seems to imply that pre- and postrevolutionary theories cannot be compared because there is no common measure to ground comparison. Such a view is at opposite poles from the project of logical empiricists and their heirs to establish a framework of concepts and standards external to particular theories and their history, and capable of grounding critical evaluation, and judgments of cognitive progress.

But there are different lines of argument in Kuhn concerning the sources and implications of incommensurability. Rival theories are said to be incommensurable because (1) they do not share the same language, or conceptual scheme, and the language of one is not translatable into the language of the other, or a neutral observation language; (2) they do not perceive or recognize the same observational data; (3) they do not address or acknowledge the same problems; (4) they do not embrace the same standards of theory-evaluation or the same interpretations of standards; and (5) they do not live in the same world. While all of these claims are present in Kuhn's argument, which of these sources of incommensurability is most basic, or most defensible? How much room does it leave for continuity and commensurability at the other levels of scientific development? These questions raise the issue of what role reasoning plays in Kuhn's conception of scientific revolution, and how large a role is played by psychological and sociological processes.

Kuhn's very notion of a paradigm and a paradigm-change is sociological in so far as it involves the collective mechanisms through which a scientific community builds up and protects a shared allegiance to its norms and social control over who is and is not a member. He characterizes scientists' embrace of a new paradigm in psychological terms as a gestalt-switch, a leap of faith, and a conversion experience. What role, if any, is left for reason (confirmation, proof, prediction, falsification) in scientists' acceptance of (1) a new conceptual scheme; (2) a new domain of observational data; (3) a different agenda of problems; (4) different standards of theory-evaluation; or (5) a novel world? Which of these is the most basic source of incommensurability? Kuhn's readers and critics focus on different strands of this account of scientific revolution and in response, move philosophy of science in different directions.

THE FIRST WAVE OF CRITICS: INCOMMENSURABILITY AS TOTAL MEANING CHANGE AND EXTREME RELATIVISM

The first influential line of criticism (Scheffler 1967, 1972; Shapere 1964, 1966, 1971) takes Kuhn's notion of scientific revolution to rest on a radical, holistic conceptual relativism and an implausible view of systematic meaning-variance between paradigms and theories. In essence, on this reading, the first alleged source of incommensurability, paradigms' unique untranslatable language of science, is taken to imply all the others, incommensurabilities of data, problems, standards, and worlds. Each scientific paradigm is imprisoned within its own framework of theoretical concepts whose internal relations determine the unique meaning of each concept and all observation terms employed by the paradigm. On this reading of Kuhn, scientific revolutions change the meaning of all concepts employed by the exponents of a paradigm (e.g., planet in the Copernican revolution) and no translation is possible between the rival languages of science.

With no language in common, it is easy to see why Kuhn would also hold that rival paradigms cannot share common observational data, problems, standards, or worlds. But in that case the advocates of rival paradigms cannot communicate or argue and thus their commitments (beliefs, values, etc.) must be explained by nonrational psychological and sociological processes. Furthermore, retrospective evaluations of theories of the sort grounded in the criteria of traditional philosophy of science (degree of confirmation, explanatory scope, etc.) will be impossible; because there will be no neutral language that permits comparisons of their empirical content. Thus Kuhn's concept of scientific revolution leads to a radical incommensurability and extreme relativism, on which every paradigm, or research tradition, is justified on its own terms, and none is any better than another (better confirmed, etc.).

For the first wave of Kuhn's critics, the resulting position of Kuhn's analysis is incoherent and a "reductio" of its own premises. If rival paradigms cannot be compared or communicate in a common language, in what sense are they rivals? With no common subject matter, there is nothing for them to disagree about. In that case, there would be no difference between a shift of paradigms (or scientific revolution) within a scientific discipline (Cartesian to Newtonian physics) and the movement of inquirers from one area of inquiry into an entirely different one (physicists becoming neuroscientists).

Furthermore, Kuhn's notion of anomalies implies that rival paradigms share some common observational data about which they disagree, and which allow comparisons of their empirical content and success. In that case, they must share some concepts or language, undermining the thesis of radical conceptual incommensurability. Finally the holistic conception of scientific meaning depends on a failure to distinguish sense and reference, among other flaws. Even if the reference of a concept changes ("planet" from Ptolemy to Copernicus; "mass" from Newton to Einstein), there may be sufficient stability of connotation to yield commensurability. On the other hand, when the connotation of observational concepts (temperature of a gas) changes, there is often sufficient stability of reference to allow comparison of paradigms' empirical contents. The development of causal theories of reference reinforced the arguments for continuity of reference (Kitcher 1978, Psillos 1999).

This entire line of criticism located the failure of extreme relativism and radical incommensurability within the terrain of philosophy of language and Kuhn's false starts there. It convinced many philosophers of science that whatever its problems, the tradition of logical empiricism had little reason to worry about Kuhn's notions of scientific revolution and incommensurability.

INCOMMENSURABILITY AS SHIFTS-IN-STANDARDS

A second reading of Kuhn shifts the focus to the strain of argument that bases incommensurability not on language, but rather on shifts in the epistemic standards or values that accompany scientific revolutions (Doppelt 1978, 1980; Zammito 2004). Such changes transform the criteria of theoretical knowledge and successful inquiry, for the field and scientific community in question. An allegiance to the new standards implicit in a paradigm shift typically involves a redefinition of the domain of problems and observational phenomena most important for any adequate theory to explain. These shifts sometimes generate losses in the problem-solving capacity and explanatory power of science, though the epistemic importance of these losses is evaluated differently on the disparate standards implicit in rival paradigms.

The premodern chemistry of the phlogiston theory and the theory of elective affinity generated solutions to a large number of problems that are eliminated from the domain of phenomena-to-be-explained by the modern chemistry instigated by Antoine-Laurent Lavoisier and John Dalton. It could account for the observable properties of a number of substances, solving the problem of why metals exhibited common metallic qualities, lacking in their ores, and why metals take on acidic qualities as a result of chemical reactions. While such questions could still be formulated in the nineteenth century, the failure to answer them, to explain the observed qualities of compounds, is not taken as a cognitive defect in Daltonian chemistry; even though empirical success with these phenomena was a central criterion of theoretical knowledge for premodern chemistry. Of course the modern chemistry of Lavoisier and Dalton succeeded in solving a whole range of problems (concerning weight relations and proportions in chemical reactions) that were largely unknown until their work. Still, given Kuhn's "loss-of-data" and "shift-in-standards" arguments concerning scientific revolution, on his view, the Daltonian paradigm is not well characterized as simply offering a better, truer, or more rational account of chemical phenomena than its predecessor. For, the premodern and modern paradigms provided explanations of different sorts of observed phenomena, in accordance with different problem-sets, and in line with different standards of adequacy for chemical theory.

Reading Kuhn's argument in this way generates a more moderate notion of scientific revolution, incommensurability, and relativism than the initial critics identified. The argument is compatible with considerable continuity and overlap across paradigms concerning language, observational data, problems, and even standards. The existence and role of anomalies exhibits such overlap. More generally, this reading is compatible with Kuhn's clear recognition that new paradigms often try to, and succeed at, treating many of the phenomena at the heart of their predecessors, and satisfying some of their standards, as well as their own. What, then, is left of incommensurability and relativism, in moderate form? Is there a moderate form of these doctrines?

On the moderate version of Kuhn, advocates of rival paradigms present good reasons and arguments to one another. But because their disagreement is about the standards of their science, and the strength of reasons is relative to such standards, paradigm debates and shifts (scientific revolution) are often marked by an absence of compelling reasons. Equally scientific and rational inquirers can weight the balance of good reasons in contradictory ways that favor the standards and achievements implicit in their rival paradigms. This moderate notion of incommensurability of reasons generates a distinctive Kuhnian version of the underdetermination of theory by evidence. Antirealists often argue that the observational implications of a theory do not confirm the

truth of the theory. Because one can always imagine another theory T1, incompatible with T, with the same confirmed observational implications; the two theories are empirically equivalent but cannot both be true. Realists reply that evidence and confirmation involve more than the mere logical consequences derivable from the theory. Confirmation of a theory by evidence for many realists requires that the theory provide the best explanation of the evidence, in virtue of its simplicity, accuracy, explanatory scope, fruitfulness, plausibility, and unification. Kuhn acknowledges the universality of such epistemic values in science. But he argues that shifts-in-paradigms change the criteria governing their application and their relative importance in determining the best explanation. Premodern and modern chemistry both valued unifying explanation, but embraced different standards concerning what sorts of phenomena required unified explanation.

If theory, or paradigm choice, is underdetermined by evidence, and good reasons, due to Kuhn's shift-in-standards claims, reason (scientific method) alone will not explain scientific revolution. Without glorifying irrationalism or mystical conversion, Kuhn can vindicate the relevance of psychological and sociological factors to explain which particular scientific considerations, in an ocean of conflicting reasons, prove compelling to the practitioners who accept a new paradigm, and why. Moderate relativism thus asserts that scientific development involves revolutions in which a new paradigm triumphs, even though it entails some losses in problem-solving capacity, and is no more rational to accept than its predecessor(s), given the different standards at play in the historical context.

CRITICS OF MODERATE RELATIVISM

This more moderate version of Kuhn's conception of scientific revolution moves its evaluation away from the philosophy of language onto the terrain of epistemological argument. Various critics of Kuhn's shift-of-standards relativism advance arguments based on the existence of external standards, piecemeal bootstrap scientific rationality, naturalist epistemology, and scientific realism (discussed, in turn, below). In the spirit of logical empiricism, some critics argue that Kuhn's emphasis on internal paradigm-specific standards is fully compatible with the existence of external, universal, and non-relative standards of scientific rationality and progress; such as predictive accuracy, explanatory scope, simplicity, completeness, empirical success, unifying power, and the like (Scheffler 1967). Isn't the existence of such independent standards what makes rational debate between exponents of rival paradigms possible and indeed intelligible as such to us today (Siegel 1980, 1987)?

Kuhn fully embraces the existence of such universal epistemic considerations (empirical success, etc.) in science. But he argues that they function as broad, abstract values of scientific inquiry, whose actual contents are transformed by scientific revolutions. In effect, he takes a moderate relativism of internal standards to imply a relativity of external standards to paradigms. But this is not supposed to be an a priori claim about scientific development. Kuhn's studies of normal science, revolution, and scientific debate are supposed to show that exponents of rival paradigms apply the aforementioned epistemic values in very different ways, yielding concretely different standards of explanation, simplicity, unification, and even accuracy (what counts as an acceptable measure of experimental deviation of prediction from observed result).

But does Kuhn's moderate relativism concerning the role of reasons and standards in scientific revolution imply any relativism concerning long-run scientific progress? The tradition of logical empiricism concerns the context of justification, not discovery. As long as there are external standards of theory-assessment sufficient to establish that science overall attains cognitive progress, Kuhnian short-run losses in problem-solving and standards need not imply any global relativism. As Kuhn himself observes these losses are often recouped in the long run. Though the chemical revolution initiated by Lavoisier abandons the effort to explain the qualities of compounds, these problems are taken up and resolved in twentieth century science. Newtonians first accepted and later abandoned the Aristotelian and Cartesian standards requiring a mechanical explanation of motion, thus gravity (no action-at-a-distance). Einsteinian physics produces an explanation of gravity without any loss to the data and problems handled by Newtonian science.

Kuhn explicitly claims that scientific development exhibits progress in the sense that there are dramatic increases in the number, range, variety, and accuracy of its problem-solutions (even if it is not consistently cumulative, step by step). Another critic seizes on problem-solving effectiveness as the way to accommodate Kuhn's historical insights while overcoming his relativism concerning scientific rationality and equivocations about cognitive progress (Laudan 1977). He seeks to establish an external standard of problem-solving effectiveness with a theory-neutral calculus for identifying, counting, and weighing the various empirical and conceptual prob-

lems tackled, solved, and unsolved in rival or successive research traditions. This account follows Kuhn's historicism in allowing that rival research traditions (a looser, more flexible concept than paradigm) are often committed to different problems, different standards of solution, different criteria for individuating, counting, and weighing important kinds of conceptual and empirical problems.

By accepting the historical relativity of problems, solutions and standards, the externalist model of maximal problem-solving effectiveness runs the risk of collapsing into a Kuhnian moderate relativism concerning the rationality of scientific change and cognitive progress. For example, on the externalist model, the objective importance of a problem (how much it affects a tradition's problem-solving effectiveness) is elevated if rivals tackle and solve it. Against this criterion, the problem solutions taken to be most important in establishing the chemistry first, of Lavoisier, and later, of Dalton, address phenomena that were largely unknown to premodern chemists (e.g., the alchemists) and thus should enjoy less epistemic weight than they were accorded and needed in the making of the chemical revolution. Once external standards are historicized, relativism threatens.

A second critique of Kuhn's notion of scientific revolution follows Kuhn in rejecting self-sufficient external standards and embracing a historicized account of scientific rationality, but one without relativist implications. These critics argue that there are typically good reasons for altering the standards and goals of scientific inquiry, internal to the historical context of shared beliefs in which the change occurs. If the context of shared belief can provide inquirers with a justification for preferring some standards over others, then paradigm change is in principle entirely rational and explainable by the reasons in its favor, without recourse to psychological and sociological dynamics (Siegel 1980, 1987). Some philosophers adopt a multilevel, piecemeal, and gradualist model of scientific change to show precisely how and why the background context of scientific change provides inquirers with good reasons to make these changes (Laudan 1984, Shapere 1984).

The gradualist model directly challenges Kuhn's holistic historiography of normal and revolutionary science. Normal science is supposed to be ordered by a global framework of tightly interwoven concepts, problems, theories, standards, and aims, such that change of any one component implies alterations in all the others. Scientific revolutions are supposed to imply something like a sudden and wholesale break with the entire framework (extreme incommensurability and relativism), or at least its alleged foundational standard(s) (moderate relativism), and the acceptance of a wholly new one.

Gradualist critics argue that if this is what scientific revolutions are supposed to be, then either there are not any, or very few. The process of rebuilding the framework of scientific inquiry is piecemeal and gradual. Change at one level—whether it is theoretical beliefs, empirical observation, methodological standards, or broad cognitive aims—does not dictate change at all other levels; and no one level is foundational for all the rest (Laudan 1984). On the other hand, change on any one of these levels can be justified by elements of continuity and agreement at other levels, even if we accept the Kuhnian view that there are no sacrosanct or permanent aims and standards with which to anchor justification (Shapere 1984).

To take a well-known example, consider the decision of inquirers during the nineteenth century to abandon an exclusive commitment to the Newtonian standard of inductive generalization, which ruled out the epistemic rationality of using observation to support inference to unobservable entities and processes. The strict empiricist inductive standard of proof was widely thought to be responsible for the great Newtonian achievement and its decisive methodological break with the vacuous, speculative hypotheses of Cartesian physics. But in the eighteenth and nineteenth centuries, scientific practitioners became increasingly interested in explaining well-known electrical, chemical, magnetic, gravitational, optical, and other sorts of observed phenomena.

This set of aims took their inquiries beyond the strictures of the Newtonian empiricist standard. The most successful theories (George Lesage, David Hartley, Roger Boscovitch) of these phenomena posited the existence of an unobservable ether(s) in order to account for them. The scientific credibility of these problem-solutions, turned on a new standard of theory-assessment, the method of hypotheses (hypothetico-deductive reasoning). Scientists like Lesage defended this standard as a sound route to genuine knowledge, alongside inductive empiricism. Some members of the scientific community became increasingly committed to the aim of explaining these phenomena, outside the privileged domain of Newtonian physics, and to the aether theories that realized this aim. These shared commitments provided good reason to defend the method of hypothesis and abandon inductivism as the sole standard of genuine knowledge (Laudan 1981, 1984).

The other theorists are neither practitioners of normal Newtonian science nor participants in a revolution-

ary break with it. They do not question the Newtonian achievement and do not reject the standards and aims associated with it. By justifying a wider standard of inference than Newtonians allowed, aether theorists grounded the empirical success of their theories and enhanced the internal consistency of their commitments. Scottish natural philosophers like Thomas Reid stuck to the Newtonian standard and thus argued that the ether theories could not embody genuine scientific knowledge. If one thinks of the parties to these debates as members of the scientific community, then it is much more loosely structured that the notion of a paradigm implies. Its members have different levels of commitment to the disparate components of the framework of scientific inquiry at the time. The framework itself may exhibit tensions or inconsistencies that different inquirers seek to resolve in different ways. The gradualist model of scientific change exploits cases like this to show how the historical context provides inquirers with good reasons for embracing a new standard of scientific knowledge.

Some philosophers press the gradualist model further to argue for a historical conception of progressive scientific rationality on which reasoning over time produces dramatic improvements in the standards, methods, and goals of good reasoning itself. For example, ether theories are ultimately discredited, and the method of hypothesis is supplanted by more demanding criteria of abduction (e.g., William Whewell's consilience of inductions). Nonetheless, the ether theorists' defense of an inference to unobservables, to account for observed phenomena, improved subsequent scientists' understanding of how knowledge can be achieved and what form it might take.

Scientific development can thus be understood as a process of learning how to learn, one in which reasoning generates progressive historical improvements in the very goals, methods and standards of good reasoning itself (Briskman 1977, Brown 1977, Laudan 1984, Nickles 1993, Shapere 1984, Zammito 2004). Such accounts of scientific rationality are characterized as bootstrap rationality, the internalization of reasons, evolutionary epistemology, or nonrelativist historicism, depending on which version is at issue. This dialectical growth in scientific rationality itself accounts for a feature of science that Kuhn himself acknowledges—the extraordinary increase in the power of science, what it can do, by way of problem-solving effectiveness, prediction, explanation, and control. If scientific development implies the enlargement of one's very capacity to reason, this account blunts the epistemological force of Kuhn's notion of scientific revolution (shifts-in-standards, moderate relativism).

THE TURN TO NATURALISM AND REALISM

A closely related development is the emergence of naturalized epistemology. The project of naturalistic epistemologists is to characterize scientific knowledge and its methods on the basis of empirical inquiry, not historical narrative of any sort. Scientific method can be characterized as whatever processes of inference are in fact most effective and reliable means to the ultimate aims of science. Some normative naturalists treat the history of science as a body of empirical evidence that can be used to determine which scientific aims are in fact realizable, and which methods are most effective in realizing them (Laudan). Reliabilist naturalists appeal to our best current sciences in order to determine which methods or mechanisms of belief-formation are most reliable in producing true beliefs (Goldman 1988). From the naturalists' standpoint, scientific change and new standards are not evaluated by the internal reasons provided by the historical context to the inquirers reasoning in that context. Rather naturalists appeal to external empirical knowledge in order to determine whether reliable and effective methods have been followed and this determination does not depend on the reasons or standards that inquirers themselves employ. From this standpoint, rational change and progress in science are evidenced by increases in the reliability of its methods and theories in generating true beliefs.

This naturalistic turn provides another way of circumventing Kuhn's notion of scientific revolution and the historical relativism (of reasons) it implies. One problem for naturalist epistemology arises from the plurality of aims or values in scientific inquiry, a central point in Kuhn's work. The naturalist cannot be expected to identify effective and reliable methods, or processes, of scientific inquiry, if its aim is left indeterminate. Is the aim explanation or prediction, maximal accuracy or unification, simplicity or completeness, etc? Even if one settles on a unitary aim such as truths about the world (as reliabilists hold), this does not settle the methodological debate between realists and empiricists, or instrumentalists.

If the aim is theoretical truths concerning the unobservable causes of observational regularities, as scientific realists argue, then they may also be correct in treating inference-to-the-best explanation as the most effective and reliable method. If the only realizable aim is exclu-

sively truths at the observational level itself, or instrumental reliability (as empiricists stress), then other methods may be more effective. Indeed, the debate between empiricists and realists is precisely over the reliability of inference-to-the-best explanation as a method of confirming the truth of theories. While there are good arguments on both sides, they are not mainly the sorts of purely empirical considerations that naturalist epistemology speaks to. They are closer to the normative and conceptual disagreements brought to light by Kuhn's conception of scientific revolution (Dopplet 1986, 1990, 2001).

Indeed Kuhn's conception places him squarely on the side of instrumentalists. His conception allows that science exhibits cognitive progress in the sense that our best current theories possess vastly more empirical success, instrumental reliability, and problem-solving effectiveness than their predecessors. For scientific realists, the great empirical success of our best current theories provides compelling evidence that they are true. If they weren't true, so realists argue, their great success would be a miracle (Boyd 1973, 1984, 1992; Putnam 1975, 1978; Psillos 1999). The realist view of theories provides the best explanation of their success. On the other hand, Kuhn takes his conception of scientific revolution to support an uncompromising antirealism. He sometimes claims that a scientific revolution alters the world, or more weakly, the aspects of the world central to scientific perception and inquiry (Hoynigen-Heune 1993). In addition, scientists' standards of success and truth shift in scientific revolution. For these reasons, scientific revolution is supposed to preclude the cognitive progress of theories toward the truth concerning the underlying, unobservable structure of reality.

Between Kuhn's virulent antirealism, and the argument of current scientific realists, there is a fundamentally different view of which features of science are most important to account for. Kuhn's notion of scientific revolution focuses on shifts in standards and aims. Scientific realists emphasize the remarkable success of our best science in realizing the ambitious standards and aims it has. If what is most important to explain is not how science arrived at its current standards and aims, but rather why the best current theories are so successful in realizing them, then scientific realists' account offers a powerful antidote to Kuhn's relativism.

Yet, scientific realists have not been entirely immune to Kuhn's historicism. One of the most influential criticisms of scientific realism stems from a careful consideration of past science (Laudan 1984). The realist appeals to the truth, or approximate truth of our best theories, to explain their empirical success. But how will the realist explain the fact that many outdated theories (e.g. the luminiferous ether theory of the propagation of light) were also empirically successful but false, to the best of our knowledge. Indeed, doesn't this record of false but successful theories constitute good inductive evidence that our currently most successful theories are also probably false? In response, scientific realists have turned to these historical cases and provided realist accounts of their successes and failures (Psillos 1999). Taking stock of the history, realists seek to narrow the range of truly successful theories, limit the components of theories confirmed by their success, and secure a greater continuity of reference than Kuhnian revolutions allow.

However its merits are finally judged, Kuhn's conception of scientific revolution drove a very fruitful wedge between traditional philosophy of science and historicism. It realigned the relation of philosophy of science both to the history of science, and studies of specific scientific practices, theories, and controversies. This realignment helped bring a fuller range of sciences such as biology into the purview of philosophy of science, where physics once reigned supreme. The debates inspired by Kuhn's work helped generate the new approaches to scientific method, rationality, and progress previously described. All told, there is more than a little irony in the fact that some of the most vocal and relentless critics of Kuhn's notion of scientific revolution ended up learning, and teaching, the most from it. What first appeared to many as Kuhn's revolution of irrationality, later proves to be a central component in a larger process of rethinking the aims and methods of philosophy of science itself.

See also Kuhn, Thomas; Scientific Method.

Bibliography

Barnes, Barry. *T. S. Kuhn and Social Science*. London: Macmillan, 1982.

Bird, Alexander. *Thomas Kuhn*. Princeton, NJ: Princeton University Press, 2000.

Boyd, Richard. "The Current Status of Scientific Realism." In *Scientific Realism*, edited by J. Leplin, 41–82. Berkeley: University of California Press, 1984.

Boyd, Richard. "Constructivism, Realism and Philosophical Method." In *Inference, Explanation and other Frustrations: Essays in the Philosophy of Science*, edited by J. Leplin, 131–198. Berkeley: University of California Press, 1992.

Boyd, Richard. "Realism, Underdetermination, and a Causal Theory of Evidence." *Nous* 7 (1973).

Briskman, Leonard. "Historical Relativism and Bootstrap Rationality." *Monist* 60 (1977): 453–472.

Brown, Harold. "For a Modest Historicism." *Monist* 60 (1977): 540–555.

Brown, Harold. *Perception, Theory and Commitment. The New Philosophy of Science*. Chicago: University of Chicago Press, 1977.

Devitt, Michael. "Against Incommensurability." *Australasian Journal of Philosophy* 57 (1979): 29–50.

Devitt, Michael. *Realism and Truth*. Oxford: Basil Blackwell, 1984.

Doppelt, Gerald. "Kuhn's Epistemological Relativism: An Interpretation and Defense." In *Relativism: Cognitive and Moral*, 113–146. Notre Dame: University of Notre Dame Press, 1978.

Doppelt, Gerald. "A Reply to Siegel on Kuhnian Relativism." *Inquiry* 23 (1980): 117–123.

Doppelt, Gerald. "Laudan's Pragmatic Alternative to Positivism and Historicism." *Inquiry* 24 (1981): 253–271.

Doppelt, Gerald. "Relativism and Recent Pragmatic Conception of Scientific Rationality." In *Scientific Explanation and Understanding: Essays on Reasoning and Rationality in Science,* edited by N. Rescher, 106–142. London: Center for Philosophy of Science Publications in Philosophy of Science, University of Pittsburg and University Press of America, 1983.

Doppelt, Gerald. "Relativism and the Reticulation Model of Scientific Rationality." *Syntheses* 69 (1986): 225–252.

Doppelt, Gerald. "The Philosophical Requirements for an Adequate Conception of Scientific Rationality." *Philosophy of Science* 55 (1986): 104–133.

Doppelt, Gerald. "The Naturalist Conception of Methodological Standards." *Philosophy of Science* 57 (1990): 1–19.

Doppelt, Gerald. "Incommensurability and the Normative Foundations of Scientific Knowledge." In *Incommensurability and Related Matters*, edited by Paul Hoyningen-Huene and Howard Sankey. Dordrecht: Kluwer Academic, 2001.

Friedman, Michael. "Remarks on the History of Science and the History of Philosophy." *World Changes: Thomas Kuhn and the Nature of Science*, edited by Paul Horwich. Cambridge, MA: MIT Press, 1993.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1988.

Giere, Ronald. "History and Philosophy of Science: Intimate Relationship or Marriage of Convenience?" *British Journal for the Philosophy of Science* 24 (1973): 282–297.

Hempel, Carl. *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall, 1966.

Horwich, Paul, ed. *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: MIT Press, 1993.

Hoyningen-Huene, Paul. *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Translated by Alexander T. Levine. Chicago: University of Chicago Press, 1993.

Kitcher, Phillip. "Theories, Theorists and Theoretical Change." *The Philosophical Review* 87 (1978): 519–547.

Kordig, Carl. *The Justification of Scientific Change*. Dordrecht: Reidel, 1971.

Kuhn, Thomas. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge, MA: Harvard University Press, 1957.

Kuhn, Thomas. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 1977.

Kuhn, Thomas. "Logic of Discovery or Psychology of Research?" In *Criticism and the Growth of Knowledge*, edited by I. Lakatos and A. Musgrave, 1–25. Cambridge, U.K.: Cambridge University Press, 1970.

Kuhn, Thomas. "The Relation between History and the History of Science." In *The Essential Tension*, 127–161. Chicago: University of Chicago Press, 1971.

Kuhn, Thomas. *The Structure of Scientific Revolution*. 2nd ed. Chicago: University of Chicago Press, 1970.

Kuhn, Thomas, J. L. Heilbrom, P. Forman, and L. Allen. *Sources for the History of Quantum Physics: An Inventory and Report*. Philadelphia, 1967.

Lakatos, Imre. "Falsification and the Methodology of Scientific Research Programmes." In *Criticism and the Growth of Knowledge*, 91–197. Cambridge, U.K.: Cambridge University Press, 1970.

Lakatos, Imre, and Alan Musgrave, eds. *Criticism and the Growth of Knowledge*. Cambridge, U.K.: Cambridge University Press, 1970.

Laudan, Larry. *Beyond Positivism and Relativism*. Boulder, CO: Westview, 1996.

Laudan, Larry. "The Demise of the Demarcation Problem." In *Physics, Philosophy and Psychoanalysis*, edited by Robert S. Cohen and Larry Laudan, 111–127. Dordrecht: Reidel, 1983.

Laudan, Larry. *Progress and Its Problems*. Berkeley: University of California Press, 1977.

Laudan, Larry. "Progress and Rationality? The Prospects for Normative Naturalism." *American Philosophical Quarterly* 24 (1987): 19–31.

Laudan, Larry. *Science and Hypothesis*. Dordrecht: Reidel, 1981.

Laudan, Larry. *Science and Values*. Berkeley: University of California Press, 1984.

Laudan, Larry. "Two Dogma of Methodology." *Philosophy of Science* 43 (4) (1976): 585–597.

McMullin, Ernan. "Philosophy of Science and Its Rational Reconstructions." In *Progress and the Rationality of Science,* edited by Gerald Radnitzky, Gunnar Andersson, Robert S. Cohen, and Marx W. Wartofsky, 221–252. Dodrecht: Reidel, 1978.

Meiland, John. "Kuhn, Scheffler, and Objectivity in Science." *Philosophy of Science* 41 (1974): 179–187.

Musgrave, Alan. "Logical versus Historical Theories of Confirmation." *British Journal for the Philosophy of Science* 25 (1974): 1–23.

Newton-Smith, William. *The Rationality of Science*. London: Routledge, 1981.

Nickles, Thomas. "Scientific Discovery and the Future of Philosophy of Science." In *Scientific Discovery, Logic and Rationality*, 1–59. Dordrecht: Reidel, 1987.

Nickles, Thomas. *Taking the Naturalist Turn, or How Real Philosophy of Science Is Done*. Chicago: University of Chicago Press. 1993.

Psillos, Stathis. *Scientific Realism: How Science Tackles Truth*. London and New York: Routledge. 1999.

Putnam, Hilary. *Mathematics, Matter and Method*. Vol. 1. Cambridge, U.K.: Cambridge University Press, 1975.

Putnam, Hilary. *Meaning and the Moral Sciences*. London: Routledge, 1978.

Radnitzky, Gerald, and W. W. Bartley, ed. *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*. La Salle, IL: Open Court, 1987.

Rosenberg, Alex. "Normative Naturalism and the Role of Philosophy." *Philosophy of Sciences* 57 (1990): 34–43.

Sankey, Howard. *The Incommensurability Thesis*. Brookfield, VT: Ashgate, 1994.

Sankey, Howard. "Methodological Pluralism, Normative Naturalism and the Realist Aim of Science." In *After Popper, Kuhn and Feyerabend: Recent Issues in Theories of Scientific Method*, edited by R. Nola and Howard Sankey, 211–229. Dordrecht: Kluwer. 2000.

Sankey, Howard. *Rationality, Relativism and Incommensurability*. Brookfield, VT: Ashgate, 1997.

Scheffler, Israel. *Science and Subjectivity*. Indianapolis, IN: Bobbs-Merrill, 1967.

Scheffler, Israel. "Vision and Revolution: a Postscript oh Kuhn." *Philosophy of Science* 39 (1972): 366–374.

Shapere, Dudley. "The Structure of Scientific Revolution." *The Philosophical Review* 73 (1964): 383–394.

Shapere, Dudley. "Meaning and Scientific Change." In *Mind and Cosmos: Essays in the Contemporary Science and Philosophy*, edited by R. Colodny, 41–85. Pittsburgh: University of Pittsburgh Press, 1966.

Shapere, Dudley. "The Concept of Observation in Science and Philosophy." *Philosophy of Science* 49 (1982): 485–525.

Shapere, Dudley. "The Paradigm Concept." *Science* 172 (1971): 706–709.

Shapere, Dudley. *Boston Studies in Philosophy of Science.* Vol. 78: *Reason and the Search for Knowledge*. Dordrecht. Reidel, 1984.

Siegel, Harvey. "Epistemological Relativism In Its Latest Form." *Inquiry* 23 (1980): 107–117.

Siegel, Harvey. *Relativism Refuted*. Dordrecht: Reidel, 1987.

Thagard, Paul. *Conceptual Revolutions*. Princeton, NJ: Princeton University Press, 1992.

Zammito, John. *A Nice Derangement of Epistemes*. Chicago: University of Chicago Press, 2004.

Gerald Doppelt (2005)

SCOT, MICHAEL

(fl. 1217–c. 1240)

Michael Scot was an astrologer, alchemist, and translator of Arabic and Hebrew works into Latin. Born in Scotland late in the twelfth century, he spent most of his active life in Toledo, Palermo, and mainland Italy—perhaps at Rome. He first appears with any degree of certainty at Toledo in 1217, when he finished a translation of al-Bītrogī's (Alpetragius's) *Liber Astronomiae* (On the spheres). The next certain date is 1220, when he is reported to have completed a Latin translation of Aristotle's *Historia Animalium,* probably at Toledo. He seems to have become favorably known at the papal court, for he was offered the archbishopric of Cashel in Ireland in 1225. He refused the office because of his ignorance of Gaelic. Probably during this period he produced the translation of Aristotle's *De Caelo et Mundo,* along with several other physical works of Aristotle with their Arabic commentaries by Averroes. It was these commentaries that were to be so influential among the Schoolmen for the next several generations. About 1228, as nearly as can be judged, Scot entered the service of Emperor Frederick II in Sicily, or at his court at Palermo, as his official astrologer. While there, he wrote his compendious *Liber Introductorius,* a general survey of the whole science of astrology, and the *Liber Particularis,* similar in content but much briefer, intended for popular use. He also composed a *Physiognomia,* a general handbook of physiological science. All three works were dedicated to the emperor and brought Scot a wide reputation. From this second, Sicilian period of his life comes the *Abbreviatio Avicenne de Animalibus,* probably done in 1231, in answer to Frederick's request for more scientific information about the animal kingdom. It was also during this period that Scot wrote *De Arte Alchemie* in which he reported having witnessed and himself verified alchemical experiments performed by Arabs, Jews, Spaniards, and north Africans.

Because of his renown many other works have been ascribed to him, such as a commentary on John of Holywood (Sacrobosco) titled *De Sphera* and a Latin translation of Maimonides' *Guide of the Perplexed,* but these attributions lack any proof or, indeed, likelihood. Scot's great contribution remains his work of translation from Arabic and Hebrew sources of Aristotle's zoological works, the work of al-Bītrogī, the commentaries of Averroes on Aristotle, and the zoological work of Avicenna. Dante Alighieri consigns him to hell as a magician.

See also Aristotle; Averroes; Avicenna; Dante Alighieri; Maimonides.

Bibliography

The biography by J. Wood Brown, *Life and Legend of Michael Scot* (Edinburgh, 1897), must be used with care. For more recent and reliable treatments, see Lynn Thorndike, *History of Magic and Experimental Science during the First Thirteen Centuries of Our Era,* 2 vols. (New York: Macmillan, 1923–1958), Vol. II, pp. 307–337; C. H. Haskins, *Studies in the History of Mediaeval Science* (Cambridge, MA: Harvard University Press, 1927), pp. 272–298; and G. Sarton, *Introduction to the History of Science* (Baltimore: Carnegie Institution of Washington, and Williams and Wilkins, 1931), Vol. II, Ch. 2, pp. 579–582, with a bibliography of editions.

S. Harrison Thomson (1967)

SCOTISM

Scotism refers variously to (1) a loosely identified body of views thought to be original with or characteristic of John Duns Scotus, (2) a tradition of texts, doctrines, and approaches that traces back to him, and (3) a *via* ("way" or perhaps "school") that had an institutional presence in the universities of the fifteenth through the seventeenth centuries. In the first sense scholars today speak of "logical Scotism" in the work of authors who perhaps have never heard of Scotus. In both the first and the second senses scholars ask whether Charles Sanders Peirce, who had read Scotus with care but was a fiercely independent thinker, was a Scotist. In the third sense scholars inquire about the presence of Scotism in the arts and theology faculties of particular universities in the sixteenth century and investigate its fortuna. Confusion can arise (and has arisen) from running these together and one should take special care to distinguish them all from the influence of Scotus—which was so pervasive in the later Middle Ages that almost every metaphysician and theologian of note felt obliged to locate himself with respect to the Subtle Doctor.

Scotus died young (perhaps as young as forty-two) and left an enormous number of fertile ideas in various stages of development. His immediate students and followers, particularly those at Paris, among whom one might add William of Alnwick (c. 1275–1333), Antonius Andreas, Nicholas Bonet (?–1360), Francis of Marchia (1290–1344), Francis Meyronnes (c. 1285–after 1328), and Petrus Thomae (c. 1280–c. 1337), took up those they found congenial and developed them in somewhat different directions. Within twenty years of Scotus's death there had also grown up a number of different explicitly critical responses to his teaching exemplified in the work of Petrus Aureoli (1280–1322) on the one hand and William of Ockham on the other. At least four elements of Scotus's thought became identified with him in particular: In metaphysics the view that there was an isomorphism between the structure of concepts and the structure of things and the associated postulation of *formalitates*; on the borderline of metaphysics and theology a distinctive argument for the existence and infinity of God; in theology the doctrine that Mary had been conceived immaculately, that is, without the stain of original sin; and also the view that the divine will was the ultimate cause of the truth of all contingent truths. It is not at all clear that any of these doctrines was entirely original with Scotus and so one should be cautious in locating someone who does not self-identify as a follower of Scotus as a Scotist in either the second or the third senses of the word simply because the person maintains some of them.

In the first half of the fourteenth century it seems to have been the metaphysical doctrines just mentioned that received the most attention. The key concept of a *formalitas* and the closely associated notion of an *haeceitas* as a formal principle of individuation attracted the attention of most of the metaphysicians of the period. There even grew up a distinctive genre of treatise *De Formalitatibus* that studied these notions. Scotus's argument for the existence and infinity of God as developed both in his *Ordinatio* and the treatise *De Primo Principio* became celebrated soon after his death Thomas Bradwardine devoted his enormous *De Causa Dei* to correcting, elaborating, and refining it and there was considerable controversy about it throughout the century. Scotus's distinctive views about the role of the divine will in the truth of contingent truths also attracted considerable attention. Much of this attention was hostile, but it was intense for all that. In the fifteenth century the doctrine of the immaculate conception, which had been rejected by Thomas Aquinas but maintained by many thinkers including Scotus, Ockham, and Pierre d'Ailly, became associated with Scotus more particularly and by the middle of the sixteenth century, as other alternatives to Thomism faded from the theological scene, it became thought characteristic of Scotism.

The earliest references to a Scotist school or at least to a group of thinkers whom one can identify as such, are, as is quite typical in the Middle Ages, by figures who see themselves as opposed to it. In 1331 Adam Wodeham, no friend of the view, identifies the isomorphism between things and concepts as characteristic of an unnamed group of thinkers who hold it to be the fundamental principle of metaphysics. By 1400 Jean de Gerson (1363–1429) identified a group holding this view as the *formalizantes* and set himself vigorously against it. In the fifteenth century one finds thinkers like John Foxoles both self-identifying as Scotists and attempting to work out histories of the movement with which they identified. Peter of Candia (c. 1340–1410) is a particularly interesting thinker of this period much influenced by Scotus whose work has received modern study.

Scotus's works were intensively studied throughout the Franciscan order during the fifteenth through the seventeenth centuries and the fortunes of that order considerably influenced his reception. Scotism as a via (school) reached its zenith in the seventeenth century. The Irish Franciscans claimed Scotus as their own (in the middle of the seventeenth century the prominent philosopher-the-

ologian John Ponce [1603–1670] even wrote "Scotus Hiberniae restitutus" to prove the point) and under the leadership of Luke Wadding (1588–1657) a team at the Irish college of St. Isidore in Rome prepared an edition of Scotus's works (Lyons 1639) that has been foundational for all subsequent editions. The considerable intellectual resources of the Franciscan order in the seventeenth century led to interesting philosophical development and debate of which the most celebrated instance is that between Ponce and Bartholomew Mastrius (1602–1673) over the nature of possibility.

See also Alexander of Hales; Augustinianism; Bonaventure, St.; Duns Scotus, John; Medieval Philosophy; Peter Lombard; Thomas Aquinas, St.; Thomism.

Bibliography

Catto, Jeremy I. "Theology after Wycliffism." In *The History of the University of Oxford*. Vol. 2, *Late Medieval Oxford*, edited by Jeremy I. Catto and Ralph Evans, 263–280. New York: Oxford University Press, 1992.

Coombs, Jeffrey S. "The Possibility of Created Entities in Seventeenth-Century Scotism." *Philosophical Quarterly* 41 (1993): 447–459.

Courtenay, William *Schools and Scholars in Fourteenth Century England*. Princeton, NJ: Princeton University Press, 1987.

Cross, Richard. "Medieval Theories of Haeceity." In *The Stanford Online Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University, 2005.

Dumont, Stephen "The Scotist of Vat lat 869." *Archiivum franciscanum historicum* 81 (1988): 254–283.

Gabriel, Astrik, L. "'Via Antiqua' and 'Via moderna' and the Migration of Paris Students and Masters to the German Universities in the Fifteenth Century." *Miscellanea Mediaevalia* 9 (1974): 439–483.

Gelber, Hester. *Logic and the Trinity*. Madison: University of Wisconsin Press, 1974.

Hoenen, Maarten. "Scotus and the Scotist School: The Tradition of Scotist Thought in the Medieval and Early Modern Period." In *John Duns Scotus (1265/6–1308)*, by E. P. Bos, 197–210. Amsterdam, Netherlands: Rodopi, 1998.

Kaluza, Zenon. *Les Querelles doctrinales a Paris*. Bergamo, Italy: Lubrina, 1988.

Martin, Christopher J. "Theories of Inference and Entailment in the Middle Ages." PhD diss., Princeton University, 1999.

Calvin G. Normore (2005)

SEARLE, JOHN
(1932–)

John R. Searle was born in Denver Colorado in 1932. He attended the University of Wisconsin (1949–1952), then Oxford (1952–1959) as a Rhodes Scholar. He earned his PhD (Oxford) in 1959 and went to the University of California Berkeley, where he remained, and where he is Mills Professor of the Philosophy of Mind and Language. Over the past forty years, Searle has been working on a selection of problems in philosophy at three levels of description: mind (the basic level), language (the middle level), and society (the highest level). In each case Searle can be seen as following a certain pattern: he proposes analyses of facts at one level of description in which they cause, are realized in, or constitute, facts at another higher level. Brute facts can count as institutional facts, and some objective brute facts can cause and realize other, subjective, brute facts. Like phenomenological analyses, Searle's approach is not classically reductive, but there is an explanatory asymmetry: higher level phenomena often are to be explained in terms of lower level phenomena (explaining is not explaining away). However, as contrasted with phenomenology, this procedure does not require that conditions revealed by analysis be revealed in experience.

MIND, COGNITIVE SCIENCE AND RATIONALITY

Searle (1981) presents the "Chinese Room" argument against "strong artificial intelligence," the view that mental states are and can be explained by programs running on the brain, by claiming that programs will give you at best the syntax or structure of thoughts, but not their semantics, their intentionality (aboutness). Searle (1985) schematized such intentional states *S(r)*, where *S* is a psychological mode, such as believing, and *r* is a representational or propositional content: that snow is white. *S* typically determines the "direction of fit" of the intentional state: beliefs have a mind-to-world direction of fit, intentions and desires have a world-to-mind direction of fit. Together, *S* and *r* fix conditions of satisfaction. For beliefs this is a truth-condition, for intentions and desires it is a fulfillment-condition. Some intentional states, such as perception, memory, intention, have the added feature of causal self-reference, in that their conditions of satisfaction make reference to their own causal role. All intentional states are linked in a causal and logical network, and function against a background of nonintentional capacities and abilities.

Consciousness, Searle (1992) argued, is not only a unified qualitative experiential state, it is a natural biological phenomena caused by and realized in the brain. Furthermore, according to the "connection principle," all mental states are either conscious, or available in principle

to consciousness. This principle if correct would rule out many of the kinds of preconscious mental states favored by cognitive science, including linguists' "cognized" principles of language, vision theorists's algorithms for the computation of stereopsis, and philosophers's "functionalist" analysis of intentional states. Furthermore, "cognitivism," the view that brains are computers (digital or connectionist) is mistaken, because being a computer is an observer-relative fact and not an intrinsic feature of the neuroscience of brains. Mental states are ontologically subjective in that they depend on a mind to exist, but they are epistemically objective in that claims about them are true or false independently of opinion.

Searle (2001) claims that human agents can act rationally because they have free choice. There are three potential "gaps" or decision points in the chain leading to free, voluntary action: a gap between having reasons and forming a prior intention to act; a gap between the prior intention and the intention-in-action that causes the movement that counts as the action; and the gap between segments of temporally extended activities—continuing to act. Acting freely involves selecting a reason to act on, and that reason cannot be causally sufficient for the action.

LANGUAGE, SPEECH ACTS, AND SOCIETY

According to Searle (1958), Frege was almost right: the use of proper names is backed by descriptive content, not by any particular one, but by a cluster. No particular predication on a name is necessary, but the disjunction of contents is. This doctrine is the target of Saul Kripke's attack on description theories of names. Searle (1969, 1979, 2001) elaborates and defends the idea that speaking a language is a form of rule-governed behavior, and that the semantics of a natural language is to be given in terms of "constitutive" rules for performing speech acts. These rules "regulate" antecedently existing forms of behavior, or "count as" the creation of a new form of behavior, or both. Illocutionary acts, such as asserting that snow is white, typically have the structure *F(P)*, where *F* is the illocutionary force (assertion) and *P* is the propositional content (that snow is white). Sentences typically encode this distinction in devices for indicating the force, *F*, of the utterance, and devices for indicating the propositional content, *P*, and these devices are governed by constitutive rules for performing the relevant illocutionary and propositional acts. Each illocutionary act has a distinct (illocutionary) point or purpose, which can be used to taxonomize such acts. Searle and Daniel Vanderveken (1985) propose an illocutionary logic in which relations between illocutionary acts and forces are captured formally.

Many illocutionary acts can be performed explicitly with the performative formula ("I hereby adjourn the meeting"), in which case the speaker makes a self-referential, self-guaranteeing declaration. Illocutionary acts can also be performed indirectly, and nonliterally (metaphor), and the theory of these forms of communication need not appeal to any special principles beyond the constitutive rules for speech acts, general rationality, and Gricean principles of conversation. Viewed from the lower level of intentional states, the performance of a speech act is the mental imposition of conditions of satisfaction on an utterance, which itself satisfies the intention in action to produce that utterance. Hence Searle's recurrent slogan that all meaning involves "imposing conditions of satisfaction on conditions of satisfaction" (Searle 2001, p. 53). Searle (1995) argues that institutions, and social facts in general, are created when agents collectively impose a new status on things that antecedently do not have it, and go on to attach certain functions to that status. Thus, a piece of paper or metal becomes money when exchange value is assigned to it and accepted. The general form of the creation of such institutional facts is: People collectively accept that X has the power to do A. Such status-functions can be nested within one another creating tangled hierarchies of social facts and organization—money can pay mortgages for property, and that property can then be inherited. Such "collective intentionality" is basic and cannot be reduced to individual or mutual intentionality.

See also Chinese Room Argument.

Bibliography

PRIMARY SOURCES

"Proper Names." *Mind* 67 (1958): 166–173.

Speech Acts: An Essay in the Philosophy of Language. Cambridge, U.K.: Cambridge University Press, 1969.

Expression and Meaning: Studies in the Theory of Speech Acts. Cambridge, U.K.: Cambridge University Press, 1979.

"Minds, Brains and Programs." *The Behavioral and Brain Sciences* 3 (1981): 417–424.

Foundations of Illocutionary Logic. With Daniel Vanderveken. Cambridge, U.K.: Cambridge University Press, 1985.

The Rediscovery of the Mind. Cambridge, MA: MIT Press, 1992.

The Construction of Social Reality. Glencoe, IL: Free Press, 1995.

Rationality in Action. Cambridge, MA: MIT Press, 2001.

Consciousness and Language. Cambridge, U.K.: Cambridge University Press, 2002.

SECONDARY WORKS

Grewendorf, Guenther, and Georg Meggle, eds. *Speech Acts, Mind and Social Reality*. Dordrecht, Netherlands: Kluwer, 2002.

Lepore, Ernest, and Robert van Gulick, eds. *John Searle and His Critics*. Malden, MA: Blackwell, 1991.

Smith, Barry, ed. *John Searle*. Cambridge, U.K.: Cambridge University Press, 2003.

Robert M. Harnish (2005)

SECONDARY QUALITIES

See *Primary and Secondary Qualities*

SECOND ORDER LOGIC

Second-order logic is the extension of first-order logic obtained by introducing quantification of predicate and function variables. A first-order formula, say *Fxy*, may be converted to a second-order formula by replacing *F* with a dyadic relation variable *X*, obtaining *Xxy*. Existential quantification yields $\exists X\, Xxy$, which may be read "*there is a relation* that *x* bears to *y*." In general relation variables of all adicities are admissible. Similarly, quantifiable function variables may be introduced.

SEMANTICS FOR THE SECOND-ORDER LOGIC

A structure, with non-empty domain *D*, for a second-order language includes relation domains $\mathrm{Rel}_n(D)$ and function domains $\mathrm{Func}_n(D)$. In general $\mathrm{Rel}_n(D)$ C $\mathbf{P}(D^n)$, where $\mathbf{P}(D^n)$ is the power set of D^n. Similarly, the function domains $\mathrm{Func}_n(D)$ are subsets of the collection of *n*-place total functions on *D*. Such second-order structures are called *Henkin* or *general* structures. If *X* is an *n*-place relation variable, a formula $\exists X\varphi(X)$ is *true* in a Henkin structure M if there is an *n*-place relation $R \in \mathrm{Rel}_n(D)$ such that $\varphi(X)$ is true in M when *X* has the value *R*. There is a similar definition for formulas of the form $\forall X\varphi(X)$ and for formulas with quantified function variables. A formula φ is a *Henkin semantic consequence* of a set Δ of formulas if φ is true in all Henkin models of Δ.

The relation domain $\mathrm{Rel}_n(D)$ need not contain all subsets of D^n. If $\mathrm{Rel}_n(D) = \mathbf{P}(D^n)$ for each *n*, we say that each relation domain is *full* (similarly for function domains) and that the structure is *full*, *standard* or *principal*. Second-order logic restricted to full structures is called *full* or *standard* second-order logic. A formula φ is a *full semantic consequence* of a set Δ if φ is true in all full models of Δ. A formula is *valid* if it is true in all full structures.

In Henkin semantics, the Completeness, Compactness and Löwenheim-Skolem Theorems hold because Henkin structures can be reinterpreted as many-sorted first-order structures. This yields Henkin's Completeness Theorem: There exists a deductive system DS such that if φ is a Henkin consequence of axioms Δ then there is a *deduction* of φ from Δ using the rules of DS. For further details, see Shapiro 1991, Shapiro 2001, or van Dalen 1994.

EXPRESSIVE POWER

Following Gottfried Leibniz, we may define "$x = y$" as "any property of *x* is a property of *y*." The corresponding second-order definition $\forall x \forall y(x = y \leftrightarrow \forall X(Xx \rightarrow Xy))$ is valid. In contrast with first-order logic, there are *categorical* second-order theories with infinite models: All full models are isomorphic. For example, let Δ be the theory with axioms $\forall x(s(x) \neq 0)$, $\forall x\forall y(s(x) = s(y) \rightarrow x = y)$ and $\forall X[(X0 \wedge \forall x(Xx \rightarrow Xs(x))) \rightarrow \forall x Xx]$. Any full model of Δ is isomorphic to the structure $(\mathbf{N}, 0, S)$, where $\mathbf{N}$ is the set of natural numbers and *S* the successor operation. So, the Löwenheim-Skolem Theorems fail in full second-order logic. Consider the theory $\Delta \cup \{c \neq 0, c \neq s0, c \neq ss0, \ldots\}$, with *c* a constant. This theory has no full model, but any finite subset of it has a full model. So the Compactness Theorem fails, too.

Extending Δ with the recursion axioms for addition and multiplication, we obtain the theory PA_2 whose unique full model up to isomorphism is the natural number structure (N, 0, *S*, +, x). Similarly there is an axiom system whose unique full model up to isomorphism is the ordered field of real numbers, $(\mathbf{R}, 0, 1, +, \mathrm{x}, <)$. More generally there exist second-order formulas expressing cardinality claims inexpressible in first-order logic. The most striking example concerns the Continuum Hypothesis (CH), which says that there is no cardinal number between $\aleph_0$ and $2^{\aleph_0}$. Results due to Kurt Gödel and Paul Cohen imply that the Continuum Hypothesis is independent of standard axiomatic set theory (ZFC). But there is a second-order formula CH* which is valid just in case CH is true.

If we augment PA_2 with inference rules for the second-order quantifiers and the monadic comprehension scheme $\exists X \forall x(Xx \leftrightarrow \varphi)$, we obtain axiomatic second-order arithmetic, Z_2. (See Simpson 1998 for a detailed investigation of Z_2 and its subsystems.) One may construct a Gödel sentence G, true just in case G is not a theorem of Z_2. Now, all full models of Z_2 are isomorphic to

(N, 0, *S*, +, x). So an arithmetic sentence φ is true just in case φ is a full semantic consequence of Z_2. G is thus a full semantic consequence of Z_2 but not a theorem of Z_2. The Completeness Theorem therefore fails; there is no sound and complete, recursively axiomatized, deductive system for full second-order logic. Indeed the set of second-order validities is not recursively enumerable. For further details see Shapiro 1991, Shapiro 2001, or Enderton 2001.

IS SECOND-ORDER LOGIC LOGIC?

Second-order comprehension has the form $\exists X \forall x_1 \ldots \forall x_n (Xx_1 \ldots x_n \leftrightarrow \varphi)$. Should such existential axioms count as logical? Does this violate the *topic-neutrality* of logic? W. V. Quine argued that second-order logic is "set theory in sheep's clothing" because "set theory's staggering existential assumptions are cunningly hidden … in the tacit shift from schematic predicate letter to quantifiable variable" (Quine 1970, p. 68). Another reason for not counting second-order logic as logic is that the full semantic consequence relation does not allow a complete proof procedure.

In reply George Boolos pointed out that the obvious translation from second-order formulas to first-order set-theoretic formulas does not map valid formulas to set-theoretic theorems. For example $\exists X \forall y Xy$ is valid, while $\exists x \forall y (y \in x)$ is refutable in axiomatic set theory. Furthermore $\exists X \exists x \exists y (Xx \wedge Xy \wedge x \neq y)$ is not valid, and so "second-order logic is not committed to the existence of even a two-membered set" (Boolos 1975 [1998], pp. 40–41). Furthermore first-order logic does have a complete proof procedure, but the set of first-order validities is undecidable (Church's Theorem), while the monadic fragment is decidable. So why is completeness used to draw the line between logic and mathematics rather than decidability?

THE INTERPRETATION OF SECOND-ORDER VARIABLES.

George Boolos (1984, 1985) has provided monadic second-order logic with a novel interpretation: the plural interpretation. Certain natural language locutions that receive monadic second-order formalizations are perhaps better analysed as instances of plural quantification. For example the Geach-Kaplan sentence, "Some critics admire only one another," may be formalized as $\exists X (\exists x Xx \wedge \forall x \forall y (Xx \wedge Axy \rightarrow x \neq y \wedge Xy))$. This formula is non-first-orderizable (not equivalent to a first-order formula containing just the predicates *A* and =). According to the usual interpretation, its truth implies the existence of a collection. The plural interpretation reads "There are some [critics] such that, for any *x* and *y*, if *x* is one of them and admires *y*, then *y* is not *x* and *y* is one of them." Rather than asserting the existence of a collection, this is a plural means of referring to individuals. Second-order logic can also be applied to set theory. In this context we can interpret monadic second-order quantification over sets as plural quantification.

See also Computability Theory; First-Order Logic; Gödel, Kurt; Leibniz, Gottfried Wilhelm; Logic, History of: Modern Logic: From Frege to Gödel; Mathematics, Foundations of; Proof Theory; Quine, Willard Van Orman.

Bibliography

Boolos, George. *Logic, Logic and Logic*. Cambridge, MA: Harvard University Press, 1998.

Boolos, George. "Nominalist Platonism." *Philosophical Review* 94 (1985): 327–344. Reprinted in Boolos 1998.

Boolos, George. "On Second-Order Logic." *Journal of Philosophy* 72 (1975): 509–527. Reprinted in Boolos 1998.

Boolos, George. "To Be Is to Be a Value of a Variable (or to Be Some Values of Some Variables)." *Journal of Philosophy* 81 (1984): 430–449. Reprinted in Boolos 1998.

Enderton, Herbert. *A Mathematical Introduction to Logic*. San Diego, CA: Harcourt/Academic Press, 2001.

Shapiro, Stewart. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press, 1991.

Shapiro, Stewart. "Second-Order Logic." In *Blackwell Guide to Philosophical Logic*, edited by Lou Goble. Oxford: Blackwell, 2001.

Simpson, Stephen G. *Subsystems of Second-Order Arithmetic*. Berlin: Springer, 1998.

Quine, W. V. O. *Philosophy of Logic*. Cambridge, MA: Harvard University Press, 1970.

van Dalen, Dirk. *Logic and Structure*. Berlin: Springer-Verlag, 1994.

Jeffrey Ketland (2005)

SELF

In its normal use the English expression "self" is not even quite a word, but something that makes an ordinary object pronoun into a reflexive one (e.g., *her* into *herself*). The reflexive pronoun is used when the object of an action or attitude is the same as the subject of that action or attitude. If I say Mark Twain shot *himself* in the foot, I describe Mark Twain not only as the shooter but as the person shot. In this sense "the self" is just the person doing the action or holding the attitude that is somehow in question. "Self" is also used as a prefix for names of

activities and attitudes, identifying the special case where the object is the same as the agent: self-love, self-hatred, self-abuse, self-promotion, self-knowledge.

"The self" often means more than this, however. In psychology it is often used for that set of attributes that a person attaches to himself or herself most firmly, the attributes that the person finds it difficult or impossible to imagine himself or herself without. The term *identity* is also used in this sense. Typically, one's sex is a part of one's self or one's identity; one's profession or nationality may or may not be.

In philosophy the self is the agent, the knower and the ultimate locus of personal identity. If the thought of future reward or punishment is to encourage or deter me from some course of action, I must be thinking of the person rewarded as me, as myself, as the same person who is now going to endure the hardship of righteousness or pass up the enjoyments of sin in favor of this ultimate reward. But this same self comes up in much more mundane transactions. If I pick up the cake and shove it in this mouth rather than that one, is it not because I think it will be me, the very same person who picks up the cake, that will have the pleasure of tasting it?

A straightforward view of the self would be that the self is just the person and that a person is a physical system. This view has been challenged on two fronts. First, the nature of freedom and consciousness has convinced many philosophers that there is a fundamentally nonphysical aspect of persons. The second challenge stems from puzzling aspects of self-knowledge. The knowledge we have of ourselves seems very unlike the knowledge we have of other objects in several ways, and this has led some philosophers to rather startling conclusions about the self. In his *Tractatus,* Ludwig Wittgenstein tells us that "I am my world" and that "'the world is my world'" (1961, 5.63, 5.641). This should lead us to the rather surprising conclusion that I am the world, or that at least Wittgenstein was. He draws at least one conclusion that would follow from this: "at death the world does not alter, but comes to an end."

The contemporary philosopher Thomas Nagel has been led to a possibly less radical but still quite dramatic view. According to Nagel, when he says "I am Tom Nagel," at least in certain philosophical moods, the "I" refers to the "objective self," which is not identical with but merely contingently related to the person Tom Nagel. This self could just as well view the world from the perspective of someone other than him (Nagel, 1983). We need to discuss the puzzling features of self-knowledge that give rise to such views.

SELF-KNOWLEDGE

"Self-knowledge" seems to have a straightforward meaning: cases of knowledge in which the knower and the known are identical. But this does not seem sufficient. The philosopher Ernst Mach once got on the end of a bus and saw a scruffy, unkempt, bookish-looking sort of person at the other end. He thought to himself,

(1) That man is a shabby pedagogue.

In fact, Mach was seeing himself in a large mirror at the far end of the bus. He eventually realized this and thought to himself:

(2) I am that man.

(3) I am a shabby pedagogue.

Now consider Mach at the earlier time. Did Mach have self-knowledge? In our straightforward sense it seems that he did. He knew that a certain person was a shabby pedagogue and, furthermore, that person was him. The knower and the known were the same. But this is not what we mean by self-knowledge. Self-knowledge is something Mach really had only when he got to step (3), when he would have used the word *I* to express what he knew.

Self-knowledge seems peculiar. First, it seems "essentially indexical." Statement (3) expresses self-knowledge because of the word *I*; it is hard to see how Mach could have expressed self-knowledge without using the first person. If he said "Mach is a shabby pedagogue," he would be claiming to know only what everyone else may have known. It does not seem that there is any objective characterization *D* of Mach, such that knowing that *he* is a shabby pedagogue amounts to knowing that *D* is a shabby pedagogue (Castañeda, 1966, 1968; Perry, 1990, 1993).

Secondly, we seem immune to certain sorts of misidentification with respect to self-knowledge. If we learn, in certain ways, that someone is in pain, then we cannot miss the fact that it is we who are in pain. That is, if Mach discovers that he has a headache in the ordinary way that a person discovers she has a headache, he can scarcely be wrong about *who* has the headache, if the range of choices is "I/you/that man," and so forth. Of course he can be wrong if the range of choices is "Mach/Freud/Wittgenstein," and so on, for he might not realize which of those people he is (Shoemaker, 1984).

Third, self-knowledge seems to play a unique cognitive role. If Mach desires that *he* do so and so, and believes that *he* can do so and so by executing such and such a

movement, then he will execute that movement without further ado (Perry, 1990).

AGENT-RELATIVE KNOWLEDGE

At least some of these peculiarities of self-knowledge can be explained by taking self-knowledge to be a species of agent-relative knowledge. There are two quite different ways of cognizing objects (people, things, places, and times). We can think of them via their relationship to us, the role they are playing in our lives at the moment of thought: the object I see; the present moment; the place I'm at; the person I'm talking to. We need to think about things in the first way, when we are picking up information about them perceptually or interacting with them, since ways of knowing and acting are tied to these agent-relative roles. I can learn about the here and now by looking; I can learn about the person I am talking to by asking questions, and so forth.

But these agent-relative roles cannot be our only ways of thinking about objects of more than passing interest to us. Different objects play the same agent-relative roles at different times, and at any given time many of the objects we wish to retain information about will not be playing any agent-relative role for us. And we cannot accumulate information along such roles. Suppose I am in Tokyo on Tuesday but return to Palo Alto on Friday. From the facts that on Tuesday I truly thought "Japanese is the official language *here*" and on Friday I truly thought "Senator Stanford used to live near *here*" it does not follow that there is some place where Japanese is the official language and near which Senator Stanford used to live.

In order to retain and accumulate information about objects, to construct and maintain a coherent picture of the world, we need to have a way of conceiving of objects as existing independently of us, as occupying and then ceasing to occupy various agent-relative roles. That is, we need objective ways of thinking about objects. We keep track of them by names or descriptions that do not depend on their relationship to us: Cordura Hall, 4 p.m., June 23, 1995, the southernmost town in Santa Clara County, Aurora Fischer. These serve as our fundamental ways of thinking about those objects. Recognition consists in connecting our objective ways of thinking of objects with the roles those objects play at a given moment. Consider the knowledge I might express with "Today is July 4." This is knowledge that a certain day, objectively conceived ("July 4"), is playing a certain role in my life; it is the present day, the day on which the thinking and speaking take place. This kind of knowledge, "knowing what day it is," is quite crucial to successful application of other, more objective knowledge. If I know that the party is on July 4 and know that today is July 4, then I will form the right expectations about what the day will be like.

Similarly, I may be in Kansas City and know that Kansas City is a good place for a steak dinner. But if I do not know that I am in Kansas City, if I do not realize that Kansas City is playing the "here" or "this city" role in my life at this moment, I will not be able to apply the knowledge that Kansas City is a good place for a steak dinner.

And again, I may know that Aurora Fischer has important information about my schedule, but unless I realize that the person I am talking to is Aurora Fischer, I will not apply this information and say, "Can *you* tell me where this afternoon's meeting is?"

These kinds of knowledge are, like self-knowledge, "essentially indexical." We use *now* and *today* to express our knowledge of what time it is and *here* to express our knowledge of where we are. These locutions are not reducible to names or objective descriptions, just as *I* was not. I cannot express what I say when I say, "The meeting starts right now" by saying "the meeting starts at *D*" for any description *D* of the present moment.

We are also immune to certain sorts of misidentification when we use certain methods of knowing. There is a way of finding out what is going on around one, namely opening one's eyes and looking (Evans, 1985). Now when one learns what is going on in this way, one can hardly fail to identify the time at which this is happening as now and the place as here. And finally, the forms of thought we express with *now* and *here* seem to have a unique motivational role. If I want to do something here and now, I will simply do it.

SELF-KNOWLEDGE AS AGENT-RELATIVE KNOWLEDGE

"Self" is really the name of such an agent-relative role, that of identity. As with other agent-relative roles, there are special ways of knowing and acting that are associated with identity. If Mach had wished to know, during the interval while he was confused, if the shabby pedagogue he was seeing had lint on his vest, he would have had to walk over to him and look. If Mach had wanted to know if he himself had lint on his vest, he could have simply lowered his head and looked. Had he done this, he would have had no doubt about whom the lint was on. If Mach found lint and wanted to brush it off, he would engage in self-brushing, a quick movement of the hand across one's

front that each of us can use to remove lint from our own vest and no one else's.

Unlike the other agent-relative roles, identity is permanent. I will talk to many people, be in many places, live through many times in the course of my life. But there is only one person I will ever be identical with, myself. Hence, accumulation along "I" is valid, unlike accumulation along "here" or "now" or "that man."

Earlier we rejected the straightforward account of self-knowledge, as knowledge about a person by that very person. Now we can put forward an alternative. Self-knowledge is knowledge about a person by that very person, with the additional requirement that the person be cognized via the agent-relative role of identity. This agent-relative role is tied to normally self-informative methods of knowing and normally self-effecting ways of acting. When these methods are employed, there will be immunity of misidentification as to who is known about, or who is acted upon.

This role can serve as a person's fundamental concept of himself or herself. In this way our self-conceptions have structures that are different from our conceptions of other individuals of importance to us. If we understand the special way in which a person's self-knowledge is structured, we do not need to postulate anything but the person himself or herself for the knowledge to be about.

See also Consciousness; Freedom; Identity; Indexicals; Mach, Ernst; Nagel, Thomas; Personal Identity; Philosophy of Mind; Reduction; Self-Knowledge; Wittgenstein, Ludwig Josef Johann.

Bibliography

Barnes, A. *Seeing through Self-Deception.* Cambridge, U.K.: Cambridge University Press, 1997.

Bermúdez, J. *The Paradox of Self-Consciousness.* Cambridge, MA: MIT Press, 1998.

Brewer, B. "Bodily Awareness and the Self." In *The Body and the Self,* edited by J. Bermúdez, A. Marcel, and N. Eilans. Cambridge, MA: MIT Press, 1995.

Cassam, Q. *Self and World.* Oxford: Oxford University Press, 1997.

Cassam, Q., ed. *Self-Knowledge.* Oxford: Oxford University Press, 1994.

Castañeda, H.-N. "'He': A Study in the Logic of Self-Consciousness." *Ration* 8 (1966): 130–157.

Castañeda, H.-N. "On the Logic of Attributions of Self-Knowledge to Others." *Journal of Philosophy* 65 (1968): 439–456.

Evans, G. "Understanding Demonstratives." In *Collected Papers.* Oxford: Clarendon Press, 1985.

Evans, G. *The Varieties of Reference.* Oxford: Clarendon Press, 1982.

Gallois, A. *The Mind Within, The World Without.* Cambridge, U.K.: Cambridge University Press, 1996.

Gertler, B., ed. *Privileged Access: Philosophical Accounts of Self-Knowledge.* Aldershot, U.K.: Ashgate, 2003.

Hill, C., ed. *Philosophical Topics: Introspection* 28 (2000).

Lowe, E. J. *Subjects of Experience.* Cambridge, U.K.: Cambridge University Press, 1996.

Moran, R. *Authority and Estrangement: An Essay on Self-Knowledge.* Princeton, NJ: Princeton University Press, 2001.

Nagel, T. "The Objective Self." In *Knowledge and Mind,* edited by C. Ginet and S. Shoemaker. New York: Oxford University Press, 1983.

Neisser, U., and D. Jopling, eds. *The Conceptual Self in Context: Culture, Experience, Self-Understanding.* Cambridge, U.K.: Cambridge University Press, 1997.

Perry, J. *The Problem of the Essential Indexical.* New York: Oxford University Press, 1993.

Perry, J. "Self-Notions." *Logos* 11 (1990): 17–31.

Shoemaker, S. *The First-Person Perspective and Other Essays.* Cambridge, U.K.: Cambridge University Press, 1996.

Shoemaker, S. *Identity, Cause, and Mind.* Cambridge, U.K.: Cambridge University Press, 1984.

Strawson, G. "The Self." *Journal of Consciousness Studies* 4 (1997): 405–428.

Wittgenstein, L. *Tractatus Logico-Philosophicus.* Translated by D. F. Pears and B. F. McGuinness. London: Routledge and Kegan Paul, 1961.

Wright, C., B. Smith, and C. Macdonald, eds. *Knowing Our Own Minds.* Oxford: Clarendon Press, 1998.

John Perry (1996)
Bibliography updated by Alyssa Ney (2005)

SELF-CONSCIOUSNESS

See *Consciousness*

SELF-DECEPTION

If weakness of will is a pathology of agency, then it is natural to regard self-deception as a pathology of cognition. Self-deception is a species of motivated believing in which the cognition of a subject is driven by desire towards the embrace of some proposition—typically, "in the teeth of the evidence." Here we may think of the alcoholic, the terminal cancer patient, or the anorexic, who, even while in possession of compelling evidence of his condition, insists, sincerely, that it is just not so. Many investigators require that, more than this, the self-deceiver must be understood to bring about his deception intentionally and knowingly in pursuit of the doxastic embrace of some motivationally or affectively favored proposition. Were this so, self-deception would seem to involve the sort of deep or internal irrationality distinc-

tive of weakness of will. For just as the weak-willed individual knowingly and intentionally acts against her judgment of what she takes herself to have best or sufficient reason to do, so the self-deceiver, on this picture of the phenomenon, knowingly violates her own norms or standards of reasoning—she comes to believe what she also believes there is insufficient reason to believe.

Producing a coherent account of how this is so has proved a vexing matter. Other investigators have argued that self-deception can be fully explicated without appeal to a subject's intentionally aiming to bring about her own deception against her current regard for the facts, and therefore without implicating this sort of deep irrationality. Notwithstanding these disputes, it seems clear that when we charge a subject with self-deception, we aim to offer both an explanation of how it is that a subject came to hold or retain a belief and a negative appraisal of the subject's belief-forming behavior.

In a quite literal way, the impetus behind the philosophical problem of self-deception springs from the force and puzzlement attached, in certain circumstances, to the question "How *could* he believe that?" We are all familiar with various unpleasant features of our cognitive lives, and there is no doubt that we do reason in ways that, as a matter of fact, violate epistemic norms that we endorse (the term 'epistemic,' meaning of or relating to knowledge, is derived from the Greek, "epistēmē"). The sources of such failures are many: we are subject to a profound confirmation bias, prone to be taken in by the vividness and salience of data (Nisbett and Ross 1980), forgetful and subject to fatigue, and so forth. Very plausibly, self-deception raises more pressing difficulties. In such cases, securing an answer to the question "How *could* he believe that?" compels us to reflect upon such issues as the nature of belief, doxastic agency, the unity of the self, and epistemic rationality and irrationality, among many others.

THE PHENOMENON

As suggested above, much controversy surrounds the effort to characterize the process of self-deception, the nature of the phenomenon itself, and the sort of irrationality characteristic of the phenomenon. Notwithstanding this disagreement, clear instances of what we call "self-deception" come readily to mind. The stock and shopworn example of the husband who, even though in possession of compelling evidence of his wife's infidelity, nonetheless insists upon her faithfulness is a case in point. Our husband may generate richly ornamented stories the apparent aim of which is to explain away the, by our lights, dispositive evidence of the fact of his wife's affairs. He may focus upon the occasions on which his wife has displayed great solicitousness and affection towards him, and he may well regard these data as clear and compelling evidence of her continued love for him. Moreover, he may subject evidence that strongly points towards his wife's infidelity to sustained and withering critical scrutiny, while precipitately embracing data indicative of her continued faithfulness. In short, our hapless husband repeatedly searches for reassuring evidence and probes various hypotheses in a sustained and continuing fashion in order to arrive at and then to retain the favored belief against various threats. Core cases of self-deception would, then, appear to involve a subject engaging in strategies the aim of which is the embrace of some proposition(s).

TRADITIONALISM ABOUT SELF-DECEPTION

How are we to characterize and explain such behavior? An approach to such cognitive misadventures that we can term "traditionalism" aims to assimilate the dynamics of self-deception to those of interpersonal deception. As Mary Haight writes: "[I]f A deceives B, then for some proposition(s) p, A knows that p; and either A keeps or helps to keep B from knowing that p, or A makes or helps to make B believe that ~p, or both" (1980, p.8). These lexical considerations (Mele 2001) for the traditionalist view, then, make it perfectly natural to characterize our self-deceived husband as knowing or believing that his wife is unfaithful and as aiming and ultimately succeeding in bringing it about that he comes to believe that she is, in fact, a loyal spouse. On such a model, the husband is not simply credulous, not merely stupid or epistemically careless; nor is he simply seduced by the salience or vividness of various data, or taken in by the confirmation bias. He is not, then, in the view of the traditionalist, merely a wishful thinker or believer. He aims at his own deception; he works hard to deceive himself. How else, we may ask ourselves, can he possibly believe that his wife is faithful? Why does he engage in such byzantine strategies, the apparent point of which is to avoid the implications of the evidence? Because he knows, or at the least strongly suspects, the truth—that she is unfaithful.

A typical traditionalist, then, will hold that our husband:

1. Believes that his wife is unfaithful (or believes that he ought rationally to believe that his wife is unfaithful).

2. Engages in intentional activity the aim of which is the acquisition of the belief that his wife is faithful.

3. Believes, at least for a time, both the belief adverted to in (1) and the belief adverted to in (2).

Donald Davidson, in an extremely influential essay titled "Deception and Division," embraced these three conditions. As he puts it in a much-cited passage:

> The acquisition of a belief will make for self-deception only under the following conditions: A has evidence on the basis of which he believes that p is more apt to be true than its negation; the thought that p, or the thought that he ought rationally to believe that p, motivates A to act in such a way as to cause himself to believe the negation of p. The action involved may be no more than an intentional turning away from the evidence in favor of p, or it may involve the active search for evidence against p. All that self-deception demands of the action is that the motive originate in a belief that p is true … and that the action be performed with the intention of producing belief in the negation of p. Finally, and this is what makes self-deception a problem, the state that motivates the self-deception and the state it produces co-exist. (1985, p. 145)

It should be noted that Davidson's rationale for a contradictory or inconsistent belief requirement for self-deception is not—as it is on some accounts of self-deception (Demos 1960)—that the self-deceiver literally *lies* to himself. Davidson takes, very plausibly, the project of lying to oneself to require a self-defeating intention. Rather, Davidson takes the philosophical problem of self-deception to be a matter of our being forced to come to grips with a continuing and synchronous irrational or inconsistent state—a state he takes to be distinctive of self-deception. Davidson characterizes self-deception as a condition brought about by my intentionally causing myself to believe against what I also believe and *continue* to believe to be the weight of the evidence. As a result, and not surprisingly, Davidson argues that the characterization of such a state requires the postulation of mental partitions, divisions in the mind.

THE PUZZLES OF SELF-DECEPTION. Still, whatever the attractions of such an account, it is difficult to fathom just how this sort of mental gymnastics can be carried off. Immanuel Kant was, for example, clearly puzzled by the looming difficulties here; as he put it in his *Metaphysical Principles of the Virtues*, "Since a second person is required when one intends to deceive, deceiving oneself deliberately seems in itself to contain a contradiction" (cited in Darwall 1988, p. 411).

In a bit more detail, traditionalism has been taken by many to give rise to two difficulties. First, there is what Alfred Mele has termed the "static puzzle" (1987, 2001). The very state of mind of the self-deceiver might strike us as deeply puzzling. How can it be that the self-deceiver believes that p and also believes that not–p? There is no doubt, of course, that human beings often harbor inconsistent beliefs, where one of the beliefs is repressed or otherwise not currently or fully available to a subject's awareness. What is harder to understand is a case in which both such beliefs are fully available to a subject.

Second, such an account makes for a strategic puzzle. Annette Barnes puts a version of the difficulty so: if I am to be self-deceived, I must "as deceived, be taken in by a strategy that, as deceiver I know to be deceitful" (1997, p. 18). The self-deceiver might well, as Davidson suggests, intentionally turn his attention away from evidence supportive of the threatening belief and seek out evidence of the favored belief with the aim of inducing in himself the latter. But if this plan is to succeed, it is not easy to see how the self-deceiver could fail to be wholly taken in by his ruse. That is, a condition of success of such a project would appear to be that the conviction that he ought rationally to believe the epistemically sanctioned proposition be exiled or come to be regarded as epistemiclly undermined *before* he can come to accept that his favored proposition is true. This is, however, very near to the sort of gambit recommended by Blaise Pascal in order to induce belief in the existence of God. There is no doubt that we can intentionally bring about conditions the result of which is that we come to believe what, at the time we brought about those conditions, we took ourselves to have no good reason to believe. This, however, does not appear to make for the deep and synchronous irrationality stalked by Davidson.

It should be noted that more modest traditionalists, while rejecting the contradictory belief requirement, have argued that the cognitive biasing in self-deception must be intentional (Talbott 1995), or that the self-deceiver need only actively avoid troubling recalcitrant evidence (Bach 1997).

Notwithstanding the difficulties to which traditionalism about self-deception has been alleged to be prey, its attractions and allure are clear. It works admirably to capture some very powerful vernacular (and philosophical) intuitions about the phenomenon. Traditionalism would sharply distinguish self-deception from putatively less puzzling phenomena such as wishful believing, for the

self-deceiver knowingly and actively brings about her deception, while the wishful believer is merely duped. Insofar as the self-deceiver succeeds in getting herself to believe what she also believes is not so, she would appear to be guilty of a profound form of epistemic irrationality. In addition, the sort of doxastic tension, instability, and fragility the traditionalist aims to describe has seemed to many the hallmark of self-deception. Lastly, insofar as the self-deceiver intentionally and knowingly brings about her deception, she is clearly blameworthy.

Predictably, perhaps, the modeling of self-deception upon interpersonal deception has tended to provoke three sorts of response. The first is outright skepticism about the phenomenon. As Mary Haight puts it: "[S]elf-deception is literally a paradox. Therefore it cannot happen" (1980, p. 73). The second response is a reconceptualization of self-deception as less a purely cognitive or doxastic affair and more an existential (or "actional") matter. Herbert Fingarette's pioneering work, *Self-Deception* (1969), is notable example of this tack. He writes of the self-deceiver that he "is one who is in some way engaged in the world but who disavows the engagement, who will not acknowledge it even to himself as his. That is, self-deception turns upon the personal identity one accepts rather than the beliefs one has" (p. 66). In this respect, Fingarette's is a powerful development and reworking of themes from Jean-Paul Sartre's famous discussion of "bad faith." Finally, in the third response one can cleave to the interpersonal model in literal fashion but seek to avoid the difficulties via a very robust partitioning or homuncularist account of self-deception. This is David Pears's account. He writes that cases of self-deception are to be explicated by appeal to a "subsystem" or homunculus that "is built up around the wish for the irrational belief [e.g. the husband's belief that his wife is faithful]. Although it is a separate centre of agency within the whole person, it is, from its own point of view, entirely rational. It wants the main system to form the irrational belief, and is aware that it will not form it, if the [belief that there is no good reason to so believe] is allowed to intervene. So with perfect rationality it stops its intervention" (1984, p. 87; see also Pears 1986). Mark Johnston (1988) develops a series of powerful objections (e.g., "Why should the deceiving subsystem be interested in the deception" (p. 64)) to homuncular explanations of self-deception.

DEFLATIONIST ACCOUNTS OF SELF-DECEPTION

A second family of accounts, "deflationism," aims to circumvent many of the difficulties the traditionalist regards as fundamental to the posing of the problem of self-deception. Alfred Mele, Mark Johnston, and Annette Barnes have all developed noteworthy deflationist accounts. According to deflationists, self-deception is a matter of coming to believe that p as a consequence of biased cognitive processing that is itself the product of the various motivational states of the subject. Such accounts very often take their cue from a rejection of the lexical considerations in favor of traditionalism (Mele 1987, 1997, 2001; Barnes 1997; Johnston 1998). So, for example, it is plausibly argued that there are many clear cases of interpersonal deception that involve neither the deceiver's knowledge of the proposition the deceived comes to believe, nor intentional deception. But if this is so, there is no obvious reason to require these conditions when it comes to the characterization and, ultimately, the explanation of self-deception. Rather, if, for example, the process of self-deceiving oneself must be understood to be mediated by the subject's intention to come to believe the favored and epistemically suspect proposition, this must be established by appealing to the fact that an explanation of particular features of the phenomenon itself *requires* such intentional activity. This is what deflationists deny. Core cases of self-deception, it is insisted, can be fully explained without appeal to the psychological exotica characteristic of many versions of traditionalism.

Alfred Mele's is the most influential of deflationist accounts. According to Mele, the following conditions are jointly sufficient for a subject's entering self-deception in acquiring a belief that p.

> 1. The belief that p which S acquires is false.
>
> 2. S treats data relevant, or at least seemingly relevant, to the truth value of p in a motivationally biased way.
>
> 3. This biased treatment is a nondeviant case of S's acquiring the belief that p.
>
> 4. The body of data possessed by S at the time provides greater warrant for not-p than for p.
>
> (2001, p. 51; see also Mele 1987, p. 127)

The account is notable for what it does not include. There is no requirement that the subject must intentionally bring about his deception, nor is there a contradictory belief requirement. It should be noted, as well, that the motivational states mentioned in (2) will typically be desires for states of affairs; for example, our husband's desire that his wife be faithful. This is to be distinguished from familiar traditionalist accounts according to which our husband not only desires that his wife be faithful but,

in addition, desires that he believe (or come to believe) that his wife is faithful; it is by virtue of the possession of this latter desire that, by the lights of the traditionalist, the husband comes to self-deceive himself. (Dana Nelkin [2002] has argued that, on pain of counting cases that do not involve self-deception as self-deception, the deflationist, like the traditionalist, must appeal to a subject's desire to believe.)

Mele, in particular, has emphasized the ways in which the motivational states of a subject can harness various sources of cognitive bias. Our husband's desire that his wife is faithful may trigger positive misinterpretation of data, negative misinterpretation of data, and selective evidence gathering and attention. Moreover, familiar "cold" or unmotivated sources of bias may also be triggered by motivation. That our husband desperately wants his wife to be loyal may make data indicative of her faithfulness more vivid as well as more salient. (We do, after all, tend to think about the objects of our desires.) Additionally, it seems clear that motivation will influence the selection of which hypotheses we begin testing with and so may trigger the confirmation bias.

DIFFICULTIES FOR DEFLATIONISM. Needless to say, it has been argued that various features of core cases of self-deception render the deflationist account implausible. William Talbott (1995), for example, has argued that not only is intentional self-deception possible in a single coherent self but, additionally, that we must appeal to an agent's intention to bias her cognition in favor of a particular proposition regardless of the truth of that proposition, if we are to explain various distinctive features of the phenomenon.

First, the process of self-deception might be regarded as too complex, too light-fingered and strategic to be the result of a non-intentional mechanism or process. Indeed, in core cases of self-deception—cases like that of our husband—the subject explains away just what needs to be explained away, he searches for just the evidence he needs in order to come to believe the favored proposition, he does not look just where he must not look, and so forth. This is just the sort of behavior characteristic of means-end rationality, and so of intentional behavior. Moreover, if the processes mediating self-deception are nonintentional, if such processes are "launched" as a simple result of our inhabiting various motivational states, why is it that human beings do not invariably bias their cognition in the direction of motivationally favored propositions? Happily, we do not always become self-deceived that p when we powerfully desire that p. Self-deception is in this sense "selective." Again, it would seem that an extremely plausible explanation of why it is that I do come to bias my cognition when I do is that I intend to do so. (It is to be emphasized that Talbott takes our self-deceptive intentions to bias our cognition to be *unconscious* intentions. Annette Barnes [1997] and Ariela Lazar [1999] have developed a number of powerful objections to the notion that unconscious intentions play a crucial role in the explanation of self-deception.)

FACING THE QUESTION: "P OR NOT-P?" Does the deflationist have the resources to respond to these difficulties? Much recent discussion of these issues has drawn on the social psychological investigation of lay-hypothesis testing. Consider the task of any hypothesis tester—including the prospective self-deceiver. He faces questions of the form: "p or not-p?" The effort to settle any such question will involve costs to the agent in the form of time, and energy spent in the task of hypothesis testing. What is central to this "pragmatic" account of hypothesis testing is another sort of cost involved in the settling of such questions: the cost of anticipated errors as noted by Friedrich (1993) and Trope and Liberman (1996). In aiming to settle a question, a subject aims to end her uncertainty, to reach her "confidence threshold" at which time hypothesis testing is ended. As such, there will be costs associated with settling the question in favor of p, when p is false (false positives), and costs associated with settling the question in favor of not-p, when p is true (false negatives). In brief, what is crucial to this account is that with regard to many such questions, the costs associated with such errors will be asymmetric rather than symmetric. As such, there will be what James Friedrich calls a "primary error," an error that the subject is preponderantly motivated to avoid. This error, not surprisingly, is fixed by the values, aims, and interests of the cognizer. Such asymmetric error costs, in turn, fix asymmetric confidence thresholds. The result is biased hypothesis testing and the striking appearance of intentional guidance toward the doxastic embrace of a favored proposition. As Friedrich (1993) puts it: "Lay hypothesis testers are always motivated by accuracy, in the sense that they want to detect and minimize particularly costly errors" (p. 357).

Consider the case of our husband. He must settle the question, "Does she or doesn't she?" His primary error is fixed by his desires and interests. As such, we can easily imagine that his primary error, the error he is most powerfully motivated to avoid, is the error of believing that his wife is unfaithful when she is not. This, then, generates asymmetric confidence thresholds. As a result, he will demand powerful and compelling evidence if he is to

accept that she is unfaithful, while requiring relatively little data to accept that she is faithful. As this is so, the model predicts that our husband will subject data suggestive of her infidelity to powerful critical scrutiny whereas he accepts data suggestive of her fidelity without serious investigation. The account promises a nonintentional explanation of the apparently strategic behavior of core cases of self-deception. It should not be forgotten, of course, that hypothesis testing is typically an amalgam of the intentional and non-intentional. *Any* hypothesis tester who faces the question “p or not-p?” does aim to settle that question. She knows, as well, of the means of which she must avail herself (seeking evidence, asking questions of those “in the know,” etc.) if she is to resolve her uncertainty. So the issue, it seems, is not whether the self-deceiver engages in any intentional behavior in coming to believe as she does. Rather, the issue is whether she *must* be understood to possess an intention to settle her question in some particular direction.

Moreover, it seems that the pragmatic account of hypothesis testing offers an explication of why it is that we do not invariably come self-deceptively to bias our cognition in favor of what it is that we anxiously desire to be so and, so, promises at least a tentative response to the selectivity problem. Again, whether an individual engages in biased hypothesis testing will be determined by the full range of the subject’s interests. So, for example, Talbott (1995) notes that hurtling down a steep mountain road and hearing unfamiliar and frightening noises when I depress my car’s brakes, I am not likely to come to believe that there is nothing amiss, even though there is no doubt that I very much want it to be the case that my brakes are just fine. Indeed, given that the error costs associated with believing my brakes are in working order when they are not are terrifically vivid, I may be likely to come to believe in biased fashion that my brakes are failing. (For skepticism concerning whether a pragmatic account of hypothesis testing holds an answer to the selectivity problem in its full generality see Jose Bermudez [2000].)

This last example indirectly raises the problem of “twisted” or “unwelcome” cases of self-deception (Mele 1999, 2001; Barnes 1997; Lazar 1999; Scott-Kakures 2000). It is indeed a striking fact that self-deception is not always a matter of coming—in biased fashion—to believe just what is desired (directly or indirectly) to be so. Indeed, overprotective parents come in strikingly biased ways to believe that their children are suffering from grave illnesses. Some subjects come to believe, on the basis of scant evidence, that their spouses are *un*faithful. And, of course, we all have our favorite hypochondriac. Though the matter is much disputed, such cases would appear to constitute at least a presumptive difficulty for familiar accounts of self-deception. Such cases do, however, appear to be explicable by appeal to the pragmatic account of hypothesis testing. Consider: A busy executive, driving to her work, is nearly hit by a careless motorist as she nears her freeway on ramp. As a result, it may be that she comes, later in her commute, to conclude that many drivers she passes are careless and so constitute a danger. This is not surprising, as she has been made vividly aware of the very high cost of failing to conclude that x is a bad driver if he is. As a result of these asymmetric error costs and the associated asymmetric confidence thresholds, she is apt to demand overwhelming evidence before concluding that x is a safe driver, and she is likely to require very little evidence to bring her to the conclusion that x is a bad driver.

According to the deflationist, then, the irrationality present in self-deception is not an irrationality that requires us to appeal to the traditionalist’s psychological machinery. Indeed, the irrationality present in self-deception is an irrationality with which we are all very familiar—it is a matter of biased reasoning. In this sense, self-deception, according to the deflationist, is not the cognitive pathology it has historically been understood to be. Much of the appeal to traditionalism springs from the intuition that only some distinctive cognitive pathology could explain the self-deceiver’s turning away from the proper aim of belief: truth. In this way, it may well be that, for the deflationist, the price of making self-deception appear more familiar is that what we are apt to regard as “normal” hypothesis testing will come to seem more suspect and less familiar.

See also Weakness of the Will.

Bibliography

Bach, Kent. “Thinking and Believing in Self-Deception.” *Behavioral and Brain Sciences* 20 (1997): 105.

Barnes, Annette. *Seeing Through Self-Deception*. Cambridge, U.K.: Cambridge University Press, 1997.

Bermudez, Jose. “Self-Deception, Intentions and Contradictory Beliefs.” *Analysis* 60 (2000): 309–319.

Darwall, Stephen. “Self-Deception, Autonomy, and Moral Constitution.” In *Perspectives on Self-Deception*, edited by Amélie Oksenburg Rorty and Brian McLaughlin. Berkeley: University of California Press, 1988.

Davidson, Donald. “Deception and Division.” In *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest LePore and Brian McLaughlin. Oxford: Basil Blackwell, 1985. Reprinted in his *Problems of Rationality*. Oxford: Oxford University Press, 2004.

Davidson, Donald. "Paradoxes of Irrationality." In *Philosophical Essays on Freud*, edited by Richard Wollheim and James Hopkins. Cambridge, U.K.: Cambridge University Press, 1982. Reprinted in his *Problems of Rationality*. Oxford: Oxford University Press, 2004.

Demos, Raphael. "Lying to Oneself." *Journal of Philosophy* 57 (1960): 588–595.

Fingarette, Herbert. *Self-Deception*. Berkeley: University of California Press, 1969. Reprint 2000.

Friedrich, James. "Primary Error Detection and Minimization (PEDMIN) Strategies in Social Cognition: A Reinterpretation of Confirmation Bias Phenomena." *Psychological Review* 100 (1993): 298–319.

Haight, Mary. *A Study of Self-Deception*. Atlantic Highlands, NJ: Humanities Press, 1980.

Holton, Richard. "What Is the Role of the Self in Self-Deception?" *Proceedings of the Aristotelian Society* 101 (2001): 53–69.

Johnston, Mark. "Self-Deception and the Nature of Mind." In *Perspectives on Self-Deception*, edited by Amélie Oksenburg Rorty and Brian McLaughlin. Berkeley: University of California Press, 1988.

Lazar, Ariela. "Deceiving Oneself or Self-Deceived? On the Formation of Beliefs 'Under the Influence'." *Mind* 108 (1999): 265–290.

Mele, Alfred R. *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. New York: Oxford University Press, 1987.

Mele, Alfred R. "Real Self-Deception." *Behavioral and Brain Sciences* 20 (1997): 90–102.

Mele, Alfred R. *Self-Deception Unmasked*. Princeton, NJ: Princeton University Press, 2001.

Nelkin, Dana. "Self-Deception, Motivation, and the Desire to Believe." *Pacific Philosophical Quarterly* 83 (2002): 384–406.

Nisbett, Richard, and Lee Ross. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice Hall, 1980.

Noordhof, Paul. "Self-Deception, Interpretation and Consciousness." *Philosophy and Phenomenological Research* 67 (2003): 75–100.

Pears, David. "The Goals and Strategies of Self-Deception." In *The Multiple Self*, edited by Jon Elster. Cambridge, U.K.: Cambridge University Press, 1986.

Pears, David. *Motivated Irrationality*. Oxford: Clarendon Press, 1984.

Sartre, Jean-Paul. *Being and Nothingness*. Translated by Hazel Barnes. New York: Philosophical Library, 1956.

Scott-Kakures, Dion. "Motivated Believing: Wishful and Unwelcome." *Nous* 34 (2000): 348–375.

Talbott, William. "Intentional Self-Deception in a Single Coherent Self." *Philosophy and Phenomenological Research* 55 (1995): 27–74.

Trope, Yaacov, and Akiva Liberman. "Social Hypothesis Testing: Cognitive and Motivational Mechanisms." In *Social Psychology: Handbook of Basic Principles*, edited by E. Tory Higgins and Arie W. Kruglanski. New York: Guilford, 1996.

Dion Scott-Kakures (2005)

SELF IN INDIAN PHILOSOPHY

The human phenomenological experience of the universe consists fundamentally of the self or subject encountering a world of objects. Thus the two main objects of philosophy are the subject or the self—its nature and constitution—on the one hand, and the universe, along with its nature and constitution, on the other. Indian philosophy is no exception to this rule.

This experiencing self is referred to by several terms in Indian philosophy, the one most widely used being ātman. The word is usually derived from the root *an*, which means "to breathe"; apparently the fact that the perceiving self is an animate being who faces other animate beings and inanimate objects is central to its emergence as the marker of the self. It is called *purua* when its distinction from inanimate nature or *prakti* is emphasized, and it is called *jīva* when the *ātman* is viewed as caught up in the cycle of *sasāra* or birth and death, freedom from which becomes a goal of this empirical self (*jīva*). In many systems this freedom is attained when the *jīva* or empirical self discovers its true relationship to the *ātman* or metaphysical self. This is the essential theological structure of the school of Indian philosophy known as Vedānta. But virtually each school of Indian philosophy possesses its own conception of the self or *ātman*, which must now be examined. Such an examination is facilitated by a review of the conception of the self in each of the nine schools of Indian thought. Although this standardization is relatively recent (Halbfass 1988, p. 353) it is worth employing because it enables one to present the concept of the self across the various schools with some measure of coherence. These nine schools, usually listed in order, are the Cārvāka (of Lokāyata), Jaina, Bauddha, Nyāya, Vaiśeika, Sānkhya, Yoga, Mīmāsā, and Vedānta.

CĀ RVĀ KA

According to the Cārvāka school, the body itself constitutes the self (*deha eva ātmā*); of course, what is meant is that the conscious body constitutes the self. However, this immaterial element of consciousness in the body is considered an epiphenomenon of the material components of the body, in a manner reminiscent of scientific materialism. The Cārvāka school would establish the plausibility of the emergence of a property not contained in the elements by their coming together on the analogy of water, which possesses the quality of wetness, a property not possessed by the two gases of which it is composed. There is no question then of postmortem survival

according to this school, as consciousness perishes with death. It therefore emphasizes making the most of life, with a pleasant death serving as the counterpart of salvation. Thomas McEvilley (2002) notes that these doctrines are similar to the ones Plato attributes to the *physiologoi.*

JAINA

According to the Jaina school, the self consists of the soul or *jīva* which occupies the body. The soul is formless but it can occupy a body just as light might occupy a room. It is a striking feature of the Jaina view of the self that this *jīva* is said to be coextensive with the body. In view of the fact that the soul occupies the body, it can be said to occupy space, as the body does, and may even be said to be capable of extension, as when the body grows. The whole range of existence, including plants and minerals along with insects and so on, possesses a conscious soul, and if such consciousness—which is characteristic of the soul—is not apparent, it is because it is dormant under the influence of karma. In Jainism karma is considered a very fine material substance that can permeate a soul, just as motes of dust might permeate light. Jaina soteriology consists of ridding the *jīva* of such matter, which keeps it weighed down in *sasāra*, so that, freed from it, it can rise to the top of the "universe" and be free forever. According to Jainism, knowledge is the natural attribute of the *ātman*, which is kept in check by *ajīva* or inanimate components of our being. "The eyes, for example are viewed here not as an aid to seeing, but as a check in the absolute sight of the soul" (Hiriyanna 1949, p. 61).

BAUDDHA

While the Cārvāka school does not believe in an *ātman* and denies anything like liberation, and the Jaina school believes in both, Buddhism denies the existence of a self or *ātman* while upholding liberation from rebirth in the usual Indic sense. According to Buddhism, continuity is possible without identity; hence there is no need to postulate a self that is reborn, for the next birth can be viewed as being caused by the present in the process of coming to an end, like an echo. Nirvāṇa brings silence to the re-echoing chamber of *sasāra*. The Buddhists seem to create many apparent logical difficulties for themselves by denying a permanent self or *ātman* but according to them the other systems create their own existential problems by believing in one. The Buddhist critique of a substantial ontology is very thoroughgoing; according to this critique, nothing whatsoever in this world possesses a permanent substratum (*sabbe dhammā anattā*). The permanence or lack of it in the self has been a major issue in the Hindu-Buddhist interface (Chakrabarti 1999, chapter 5, appendix).

NYĀYA AND VAIŚEIKA

The concepts of the self in the Nyāya and the Vaiśeika schools have much in common and hence are presented together. According to the Nyāya and the Vaiśeika school, the soul or *ātman* is eternal, but consciousness is not its inherent property. Consciousness arises when the self or *ātman* is conjoined with manas or the mind, which is, however, by itself inert. The soul or *ātman* differs from other atomic or all-pervasive objects in that, unlike them, it is potentially capable of consciousness. The selves are numerous and all-pervading but remain distinct in the state of release because of the property of *viśea*, which accounts for things being different that are in other respects all alike—for example, two atoms that are otherwise identical are not numerically one. The self has no consciousness in the state of release because such a state involves the absence of manas. The *ātman* in Nyāya is a unique substance that possesses the attributes of cognition, emotion, and conation and the qualities of desire, aversion, pleasure, pain, volition, and knowledge. The Vaiśeika school provides a longer list (Organ). As these are not perceived by the external senses and are not physical, they must belong to a nonphysical substance such as the soul. However, although consciousness or knowledge is an attribute of the *ātman*, it is not inseparable from it. The soul is thus an independent substance, but consciousness is an accidental property of it. In order for conscious states to arise, manas must come into play, hence the otherwise cryptic remark that "the true self is broken up here, we may say, into two 'selfless elements'" (Hiriyanna 1949, p. 91). Scholars such as McEvilley (2002) note parallels here with Aristotelian thought.

SĀNKHYA AND YOGA

The concepts of the self in the Sānkhya and the Yoga schools are also sufficiently similar to be treated together. In Sānkhya the self is called *purua* or soul and represents pure consciousness, in opposition to *prakti*, which represents matter. The self loses its inherent consciousness by mistakenly identifying itself with the body as involved in the process of *sasāra;* the self is utterly passive and merely a spectator but mistakes itself for an actor and thus undergoes the ups and downs of the cosmic drama. Although the word *purua* is often used in the singular, in reality the system allows for a plurality of *puruas*, all consisting of pure consciousness, but distinct from each other and *prakti* or matter. The *purua* in Sākhya and Yoga

is an uncaused, eternal, all-pervading, and changeless reality, which witnesses change as a transcendent subject distinguished by pure consciousness that can itself never become an object of knowledge. Salvation consists of this discrimination (*viveka*) that one is pure spirit and not the mind with whose derivative reality one identifies oneself. The system of Yoga with its eight limbs or constituent elements is meant to guide one, through a series of meditations, to the realization of this ultimate transcendent witnessing subject as distinct from the mind, body, and ego just as the surface of the mirror is totally independent of the objects that are reflected in it but appear included in it.

MĪMĀSĀ

The concept of the self in Mīmāsā is broadly similar to that found in Nyāya and Vaiśeika, but there are some differences. The list of specific qualities characterizing the self is similar but not identical, with Mīmāsā dropping those of dharma and adharma and adding that of *śakti* or potency. The most significant difference however consists of the fact that while according to the Nyāya-Vaiśeika school knowledge is a quality of the self, according to Mīmāsā it is an activity of the self.

VEDĀNTA

The conception of the self or *ātman* in Vedānta needs to be presented in accordance with the school of Vedānta involved—whether it is Advaita Vedānta, Viśiādvaita Vedānta, or Dvaita Vedānta. Thus the exact conception of the *ātman* depends on whether we are dealing with the "non-dualism of the qualified" (Advaita) or dualism (Dvaita). Prior to identifying the self in these three schools of Vedānta, however, it might be useful to indicate the concept of the *jīva* they all share in common on the basis of their reliance on the same foundational texts. Another aspect of the issues relating to the self or *ātman*, which receives relatively greater treatment in Vedānta than in other systems, is its relationship to *Brahman*, or the ultimate reality. It will therefore be useful to begin the discussion of the self in the three Vedantic schools with the conception of it they all share and conclude it with their views on the nature of the relationship of this *ātman* to *Brahman*.

The description of the human person as found in the *Taittirīya Upaniṣad* (II, 1–5) became paradigmatic in later Vedānta. According to this description a person consists of five sheaths within which the *ātman* lies enclosed. Starting from the outside, the first sheath consists of the body made of food (*annamaya-kośa*); within it are the vital airs that comprise the second sheath (*prāamaya-kośa*). The mind comprises the third sheath (*manomaya kośa*), consciousness the fourth (*vijñānamaya kośa*) and bliss the fifth (*ānandmaya*). In Advaita the self consists of self-effulgent consciousness (*svaprakāsa caitanya*), which is rather than has consciousness. It is one and the same in all human subjects (unlike Sākhya) and eternally free. Later Vedānta also developed a doctrine of the three bodies that comprise a human being, which ostensibly seems to possess only one body. These are the *sthūla-śarira* (or gross body) which corresponds to the *annamaya kośa*; the *sūkma-śarīra* (or subtle body), which corresponds to the *prāamaya* —the *manomaya* —and the *vijñānamaya kośa* and the *kāraa-sarīra* (or casual body), which corresponds to the *ānandamaya kośa*. The true self—the *ātman* —lies beyond all the five sheaths and the three bodies or may be said to constitute their nucleus depending on how one chooses to describe it (Kesarcodi-Watson 1994).

According to Advaita Vedānta, the *ātman* is one's true self and is identical with *Brahman*. Any differences between the two are adventitious, caused by *upādhis* or superimpositions. A popular metaphor illustrates the point as follows: Different jars of different shapes and sizes may contain jar-space. The space enclosed by these jars may appear distinct, but if one breaks the jars, all space becomes one and the same. It was, however, one and the same to begin with—the jars only created ultimately artificial and unreal differences. Thus the selves of all are identical with each other and with *Brahman*.

The *ātman* per se is of the nature of pure consciousness according to Viśiādvaita Vedānta. The self is not pure consciousness, as maintained by Advaita Vedānta, but "a conscious subject called the ego or the 'I.'" The *ātman* is the self, but this self both *is* and *has* consciousness. Moreover, the self may mistakenly identify with the objects of the world, but it is identical neither with them nor with *Brahman*. It has lost sight of its true nature, one of utter dependence on God or *Brahman*. *Moka* consists in being properly aligned with God through devotion and grace. One important difference between Advaita Vedānta and Viśiādvaita Vedānta is that whereas the *ātman* is infinite in its true nature according to Advaita Vedānta, it is considered atomic or infinitesimal in size in Viśiādvaita Vedānta, but it is able to have knowledge beyond itself through the fact that it not only is but possesses consciousness called *dharmabhūta jñāna*. *Jīvas* or empirical beings are infinite in number according to both the schools, but because of its metaphysical non-dualism, Advaita ultimately concedes only one reality: *ātman* = *Brahman*.

According to Dvaita Vedānta, the *ātman*s are infinite in number. The reason given to justify this is the obvious differences in their experiences (which are considered ultimately only empirical in Advaita Vedānta). They are atomic in size, and, as pointed out, differ from each other. They also differ from God, and the distance posited between them and God is somewhat greater in Dvaita Vedānta than in Viśiādvaita, as indicated by the very designations of these systems. Viśiādvaita Vedānta accepts the "monism of the qualified," of God as qualified by the *ātman*s, but Dvaita Vedānta is frankly dualistic. Salvation results from the grace of God.

An utterance found in the Chāndogya Upaniṣad famously states "that thou art." The *that* here is usually taken to relate to *Brahman* and the *thou* to *ātman*, and the interpretation of this seminal utterance in the three schools of Vedānta—the Advaita, the Viśiādvaita, and the Dvaita—is instructive of the differences in the concept of the *ātman* as it is understood in the three schools. According to Advaita Vedānta it means that *ātman* and *Brahman* are identical. "The identity of the denotation of the two terms" has to be realized "while their connotations are different" (Hiriyanna 1949, pp. 163–164). According to Viśiādvaita Vedānta it is to be interpreted as follows: "'That' finally denotes God as having the entire universe as his body; and 'thou,' God having the individual soul as his body" (p. 184). According to one interpretation offered by Dvaita Vedānta, the identity here really implies resemblance, for *ātman* "have features like sentience and bliss (though qualified) common with God" (p. 192). The precise idea of the self differs in virtually every system of Indian thought beyond the ones discussed here (see K. P. Sinha 1991).

See also Atomic Theory in Indian Philosophy; Brahman; God in Indian Philosophy; Knowledge in Indian Philosophy; Liberation in Indian Philosophy; Meditation in Indian Philosophy.

Bibliography

Chakrabarti, Kisor Kumar. *Classical Indian Philosophy of Mind: The Nyāya Dualist Tradition*. Albany: State University of New York Press, 1999.

Chatterjee, Satischandr,a and Dhirendramohan Datta. *An Introduction to Indian Philosophy*. Calcutta: University of Calcutta, 1950.

Halbfass, Wilhelm. *India and Europe: An Essay in Understanding*. Albany: State University of New York Press, 1988.

Hiriyanna, Mysore. *The Essentials of Indian Philosophy*. London: George Allen & Unwin, 1949.

Kesarcodi-Watson, Ian. *Approaches to Personhood in Indian Thought*, edited with a foreword by J. G. Arapura. Delhi: Sri Satguru Publications, 1994.

McEvilley, Thomas.*The Shape of Ancient Thought: Comparative Studies of Greek and Indian Philosophers*. New York: Allworth Press, 2002.

Organ, Troy Wilson. *The Self in Indian Philosophy*. The Hague: Mouton, 1964.

Sinha, K. P. *The Self in Indian Philosophy*. Calcutta: Punthi Pustak, 1991.

Arvind Sharma (2005)

SELF-INTEREST

Aristotle claims in the *Nicomachean Ethics* that it is the virtuous person "more than any other sort of person who seems to be a self-lover. … he awards himself what is finest and best of all" (1168b28–30). Aristotle's thought is that if one pursues things such as pleasure and wealth, one pursues what is base, injuring oneself. Contrast this with the implication of the recommendation "Look out for number one." This advice is not taken to mean that one should pursue virtue. Rather, the idea is that the interests of others should take second place to one's own. Virtue is not usually seen as the path of self-interest, especially because it can often involve self-sacrifice. This conflict suggests that effective pursuit of self-interest, or the interests of others, requires an account of the nature of well-being. (Henceforth, I will often use the term *well-being* rather than *self-interest* since that term is used more often in philosophical discussions of self-interest.) In the first part of this article, the major theories are discussed. In the second part, the focus is the importance (or lack thereof) of having an account of well-being for ethics.

THEORIES OF WELL-BEING

The three dominant types of theory regarding well-being are hedonism, desire theory, and objective-list theories. This classification needs refinement, but it is a useful starting point. Take hedonism first. Jeremy Bentham (1970) was probably the most notorious proponent of hedonism. He espouses a type of hedonism that Derek Parfit dubs "narrow hedonism." Bentham holds that pleasure is what is good for humans; pain is bad. He says, in *An Introduction to the Principles of Morals and Legislation*, that pleasures are homogenous sensations. How well one's life is going depends on quantity of pleasure—the more the better.

One major objection to this outlook is that there is no felt sensation in common among the experiences that

people find pleasurable. For example, L.W. Sumner (1986) asks us to imagine the difference between the pleasure of going for a walk in the woods and the pleasure of completing a difficult task. Both are pleasures but they have no felt sensation in common.

Henry Sidgwick's form of hedonism, "preference hedonism," avoids this difficulty. He observes, in *The Methods of Ethics*, that "the only common quality [among pleasures] … seems to be the relation to desire and volition expressed by the general term 'desirable'" (1907, p. 127). Sidgwick says that the judgments of the individual about which feelings are desirable must be taken as final. So pleasures, on this view, are those mental states that are desired by the individual. Some have noted that it strains the meaning of "pleasure" to call all of the mental states that we desire "pleasures." James Griffin's (1986) example citing Freud's desire to be mentally aware, but in horrible pain, rather than take opiates for his cancer pain, is such a case. Perhaps the name of the theory should be modified (as Shelly Kagan suggests) to "preference mental statism."

One strength of preference hedonism is that it respects the authority of the individual in determining which experiences make his or her life go better. Narrow hedonism says that a life of pleasurable sensation is better for the person even if one does not prefer it. Preference hedonism's weakness is that there are some desirable states of affairs that seem to contribute to well-being yet are not, strictly speaking, experiences. Probably, the most famous illustration of this problem is Robert Nozick's (1974) Experience Machine. Nozick asks us to imagine a machine that will give us all of the experiences we desire. He suggests that people would not choose to enter the machine because the experiences would have no relation to reality. Take another case. Imagine that someone is happy because she believes, falsely, that she has devoted friends. Now imagine the same person with happiness resulting from a true belief in devoted friends. Some think that the second is clearly the better life, especially if the second is preferred. The implication is that the fulfillment of desires for things other than mental states contributes to well-being. So, it seems that preference hedonism should be abandoned in favor of desire theory.

If desire theory is unrestricted, then it says that the fulfillment of any desire contributes to self-interest. That this is implausible is nicely shown by Parfit's (1984) case of the stranger: I meet a stranger with a supposedly incurable disease. I desire a cure for him; later, he is cured, though I never know this. It seems ludicrous to say that I am better off when the stranger is cured. This shows that the desires that should count as contributing to a person's well-being have to be restricted. Parfit suggests that the desire has to be a desire about one's own life. This encompasses, for example, the desire not to be deceived, if it is a desire about one's own life.

But it may be unclear what qualifies as a desire about my own life. Is my desire to live in a just world, or in a world without starvation, a desire about my own life or not? Shelly Kagan (1992) argues that for a state of affairs to matter to my well-being, it has to affect me. My subjective experience is the same whether I am deceived or not. Kagan concludes that it may be that we should restrict the class of desires that are relevant to well-being to desires about mental states. This would mean a return to some form of preference hedonism. Whatever account is better, it is clear that a successful desire theory needs a plausible way to restrict the class of desires that impact well-being.

Now, consider the third type of theory: objective list. According to these theories certain things are good for people, even if they do not want them or have a negative attitude toward them. Consider John Rawls's (1999) famous example of the talented mathematician who wants to spend his life counting blades of grass. Some think that such a life cannot be good for him because the activity is worthless.

However, the difficulty for objective-list theories lies in giving an account of which activities are objectively worthwhile. One prominent account is Aristotle's Function Argument. The function of a flautist is to play the flute, and the flourishing flute player plays the flute well. The function of a human being is to engage in rational activity in accordance with virtue. A good example of a flute player plays the flute well, and a good example of a human engages in virtuous rational activity.

One major worry for the argument, noted by Peter Glassen (1957), is that even if it gives a correct account of human excellence, the inference that it must be good for a human to be a good example of his or her kind is fallacious. It is easy to imagine cases in which the excellent thing fails to be good for the agent. There are prosperous, sensible knaves, and sometimes the good die young. Some form of desire theory is now most commonly thought to be a correct account of well-being.

WELL-BEING AND ETHICS

It may be thought that it is obvious why having a theory of self-interest is important for ethics. If moral theories yield principles about people's duties, and if their duties

include benefiting themselves and others, people need to know what counts as a benefit and what counts as a harm. The classical utilitarians—Jeremy Bentham, John Stuart Mill, and Henry Sidgwick—all think that the only thing that is intrinsically good is welfare or well-being and that the ultimate principle of morality is to perform the action that maximally benefits people.

Other theorists think that there are other goods besides well-being. W. D. Ross (1930) imagines two worlds in which there are equal amounts of happiness and equal amounts of virtue and vice. In the first the virtuous are happy, in the second the vicious are. Ross thinks that the first world is clearly better because of the distribution, even though they contain equal amounts of happiness. G. E. Moore holds, in *Principia Ethica* (1903), that it is good for beauty to exist even if it never affects anyone's conscious life. The deontologists writing in the Kantian tradition think that there is a duty of beneficence, although what is unconditionally good is the good will. However, there are some moral theorists who think that issues about well-being have little importance for ethics. For example, T. M. Scanlon (1998) argues both that individuals do not use the concept of well-being much in their deliberations about their own lives and that moral and political philosophers focus on just distributions of things such as primary goods, resources, or capabilities, rather than well-being. And he thinks that we do not have a general duty of beneficence. Notice, however, that one of the main reasons for the focus on primary goods or resources is the problem of expensive tastes. For example, Ronald Dworkin (1981) imagines a person who needs ancient claret and plover's eggs to be satisfied. Another person might reach an equal level of well-being with something much cheaper such as beer. To equalize welfare would require giving more resources to the first person. Dworkin and other theorists think that would be unjust, so they reject the idea of distributing welfare. The rejection might be correct, but it would be impossible to make the argument without a conception of human well-being.

See also Aristotle; Bentham, Jeremy; Dworkin, Ronald; Egoism and Altruism; Ethical Egoism; Eudaimonia; Freud, Sigmund; Happiness; Hedonism; Mill, John Stuart; Moore, George Edward; Nozick, Robert; Parfit, Derek; Pleasure; Rawls, John; Ross, William David; Sidgwick, Henry.

Bibliography

Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis, IN: Hackett Publishing, 1985.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*, edited by J. H. Burns and H. L. A. Hart. London: Athlone Press, 1970.

Dworkin, Ronald. "What is Equality? Part I: Equality of Welfare." *Philosophy and Public Affairs* 10 (1981): 185–246.

Glassen, Peter. "A Fallacy in Aristotle's Argument About the Good." *Philosophical Quarterly* 7 (1957): 319>en>322.

Griffin, James. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press, 1986.

Kagan, Shelly. "The Limits of Well-Being." *Social Philosophy and Policy* 9 (1992): 169–189.

Mill, John Stuart. "Utilitarianism." In *Essays on Religion, Ethics, and Society. The Collected Works of John Stuart Mill*. Vol. 10, edited by J. M. Robson. Toronto: University of Toronto Press, 1969.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.

Rawls, John. *A Theory of Justice*. 2nd ed. Cambridge, MA: The Belknap Press of Harvard University, 1999.

Ross, W. D. *The Right and the Good*. Oxford: Oxford University Press, 1930.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University, 1998.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1907.

Sumner, L. W. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press, 1986.

Joyce L. Jenkins (2005)

SELF-KNOWLEDGE

Legend has it that when Chilan of Sparta asked, "What is best for man?" Apollo replied, "Know thyself." Thus, carved into the lintel of the Oracle of Apollo at Delphi were the Greek words "*gnothi seauton*"—"Know thyself" (Parke 1933). We can try to follow this Delphic injunction because we are self-conscious beings, capable of self-reflection.

Sigmund Freud (1923) maintained that we have unconscious beliefs, desires, motives, and intentions, and that extensive use of psychoanalytic techniques is often required to uncover them. Whether there is a Freudian unconscious is controversial, as is whether or not there is suppression or repression in the psychoanalytic senses. Nevertheless, our mental lives can be dissociated. And self-reflection can be as biased as reflection on any topic. Too charitable an attitude towards ourselves can leave us overly sanguine about the strength of our characters or the goodness of our intentions. Too uncharitable an attitude can lead to an exaggerated view of our frailties: We

may see ourselves as more selfish, less kind, and less well-intentioned than we really are. We can engage in wishful thinking, believing something about ourselves on less than adequate evidence because we want it to be true; evasive thinking, which involves turning our attention to other matters when thoughts about ourselves arise that conflict with our self-image; and skeptical thinking, in which we construct hypotheses on the fly to explain away evidence that conflicts with our self-image. Arguably, we can practice self-deception about our motives and reasons, but, in any event, we can certainly unintentionally mislead ourselves about them (McLaughlin and Rorty 1986).

Social psychologists have found that we have a tendency to confabulate. Asked to explain our decisions or actions, we sometimes fabricate explanations, not with an intent to deceive, but apparently with an eye toward justifying or making sense of those decisions or actions; the end result is we are taken in by our own fabrications. Evidence for this proclivity has led some philosophers to maintain that confabulation is so pervasive that our self-reports are just unreliable stories that we tell about ourselves for a variety of aims (Dennett 1991). Many philosophers, however, deny that our fallibility warrants skepticism about the possibility of self-knowledge. And many follow Socrates in holding that there is virtue in heeding the Delphic injunction. The quest for self-knowledge is a way of taking responsibility for ourselves. But it should, of course, be balanced with other activities of value. Yet another Delphic injunction is "Everything in moderation." Narcissism or self-hatred can result in self-absorption, which is a vice.

Other people are better judges of certain aspects of our character than we are. They sometimes read our emotional state better than we do or remember our view about a certain topic better than we do. Moreover, they can tell us when we are confabulating or not being honest with ourselves. Nevertheless, it seems that each of us is able to know some things about ourselves in ways unavailable to others. Gilbert Ryle denied this claim, arguing that "the sorts of things that I can find out about myself are the same as the sorts of things that I can find out about other people, and the methods of finding them out are much the same. … John Doe's ways of finding out about John Doe are the same as John Doe's ways of finding out about Richard Doe" (1949, p. 155). He further claimed that "our knowledge of other people and ourselves depends on noticing how they and we behave" (1949, p. 181). According to behaviorism, we can know our own mental states only by observing our own behavior or relying on the testimony of others who have. Of course, often others are better positioned to observe our behavior than we are. Hence, the joke "One behaviorist meeting another on the street said, 'You feel fine! How do I feel?'" (Ziff 1958). The behaviorist view of self-knowledge seems untenable. We need not rely on observations of our behavior to know whether we are in pain, or are visualizing a red sunset, or are just now thinking to ourselves that behaviorism is untenable.

René Descartes (1985) drew attention to an area of mental life to which we seem to have first-person privileged access and with respect to which we seem authoritative: namely, our current conscious states. Conscious states include bodily sensations (aches, pains, itches, tickles, and the like), sense experiences (visual, auditory, and so on), mental imagery, felt emotions (feelings of fear, and so on), felt urges, and occurrent thoughts. We seem able to know our current conscious states in a way different from the way in which we know those of others. Indeed it seems that to know whether we are in a certain conscious state, we need only turn our attention to whether we are. To know whether we are in pain, for instance, it seems that we need only turn our attention to whether we are in pain.

Of course, we are by no means omniscient about such matters. Beliefs about what conscious states we are in involve the exercise of concepts (Sellars 1963, sec. 62); and we may lack the requisite concepts to know that we are in a conscious state of a certain sort. Even when we have the requisite concepts, we can fail to know simply because of lack of attention to the matter. Moreover, our concepts of types of conscious states are vague. Over the course of a morning we may move gradually from feeling cold to feeling warm, being unable to discern a difference in our thermal sensations from one moment to the next. On route to feeling warm we will pass through borderline cases of feeling cold; and in such cases we cannot know whether we feel cold. Theories of vagueness differ over why we cannot know in such borderline cases. According to semantic theories, the reason is that there is no fact of the matter whether the concept of feeling cold applies, and so they present no limitation to self-knowledge. But given our inability to discriminate cases falling very near the borderline from cases on it, our ability to know whether we feel cold may stop short of the borderline (Williamson 2000).

We can make verbal mistakes in our reports of our conscious states (Broad 1925), and even perhaps conceptual mistakes in our judgments about what conscious states we are in because of less than a full mastery of a rel-

evant concept (Burge 1979). But it has been held that if one has mastered the relevant concepts, one's belief that one is in a certain conscious state will be infallible and incorrigible (Ayer 1940). A belief or thought is infallible just in case it cannot be false; incorrigible just in case it cannot be shown to be false. Descartes (1985) argued, "*Cogito ergo sum*"—" I think, therefore I am"—taking his first-person thought that he thinks to be infallible, incorrigible, and indeed indubitable, such that it cannot be rationally doubted.

"Cogito-thoughts" such as that I am now thinking, and that I am now thinking that P, are indeed infallible, and hence incorrigible: they are true by virtue of my thinking them (Burge 1988). (Similarly, the belief that one has beliefs is true by virtue of one's having it.) Our infallibility in these cases, however, is not due to privileged access to the mental acts of thinking in question. If I write in English that I am writing in English, then what I write is true by virtue of my so writing it; even though I lack privileged access to whether I am writing in English, or even to whether I am in fact writing at all. Indeed, there are scenarios in which I am writing that I am writing in English but in which I fail to know that I am writing in English. The cogito-thought that one is thinking that P is (normally) an expression in consciousness of one's belief that one is thinking that P. But one can believe that one is thinking that P, when the only thought one is having is the thought that P; indeed that is the typical case (McLaughlin and Tye 1998a).

Beliefs to the effect that one is thinking that P are not true by virtue of one's having them. Moreover, they are fallible. To note just one reason: the longer it takes one to occurrently think that P, the more demand is put on short-term memory, and so the less reliable is one's belief that one is thinking that P (Armstrong 1963). Even the belief that we are in pain is fallible. Someone mesmerized by his guru might mistakenly believe that he is in pain solely on the basis of his guru's testimony to that effect. To take a more mundane case, upon hearing the start of the dentist's drill, one might momentarily mistake a feeling of pressure for a feeling of pain (Goldman 2002). (See also the "fraternity initiation" case described in Hill 1991, pp. 128–129.)

The term *introspection* is sometimes used very broadly to cover nearly any first-person, nonconsciously inferential avenue to knowledge of what mental states we are in. But on a more restricted usage (one to be followed here), introspecting a mental state is supposed to be a kind of direct act of awareness of the state. According to introspectionism, we can attend to our current conscious states by introspecting them (Locke 1690; Broad 1925; Armstrong 1963; Hill 1991; Lycan 1996; Macdonald 1998, 1999; McLaughlin 2000, 2001, 2003c; Sturgeon 2000; Goldman 2002). The term *introspection* derives from the Latin *spicere*, which means "look," and the Latin *intra*, which means "within." But the etymology is misleading. Introspectionists do not hold that we literally look within. There is no "mind's eye" by which we observe our visual experiences, no "mind's ear" or "mind's toe" by which we observe, respectively, our auditory experiences and tactile experiences.

It is widely held that we see the scenes before our eyes by having visual experiences caused by them. We are not, however, aware of our visual experiences by having visual experiences caused by them. We do not see our visual experiences; they do not look any way to us. (Nor do they look any way to an internal homunculus; an untenable view that leads to an infinite regress of sighted homunculi embedded within sighted homunculi.) Introspective access is direct in a way perceptual access is not. We experience our experiences, not by having experiences of them, but by having them. We can have them without introspecting them. But when we introspect, our attentional access to them is direct in that it is unmediated by any experiential states. Experiences are in that sense self-presenting (Chisholm 1977). If this view is correct, then we are immune to a certain kind of error. When our perceptual experiences are illusory, when things are not as they appear, we can be misled into believing that they are as they appear. If, however, our conscious states are self-presenting, then there is no appearance/reality distinction that pertains to them. We thus cannot be misled about them by their appearing to us some way that they are not.

Some introspectionists maintain that an act of introspective awareness of a conscious state is direct in yet another sense: it is unmediated by any causal mechanism. If, however, introspective awareness of a conscious state involves believing something of the state (for example, that it is a pain), the question arises as to whether this of-ness connection requires causation. It seems like mystery-mongering to maintain that it is a primitive, fundamental relation. One view is that the relation is part-whole rather than causal: The conscious state is a constituent of the introspective belief. But there are constituents of the belief that the belief bears no of-ness relation to, for example, the concepts involved in it. So, the constituency must be of a special sort. Proponents of this view are under an obligation to explicate it. There is also the issue of whether such an account can allow for mistaken intro-

spective beliefs. These remain topics of investigation (see Chalmers 2003).

The more common view is that an introspective belief and the state introspected are linked by a causal mechanism. Causes and effects, however, must be "distinct existences," and so capable of independent existence (Armstrong 1963). This causal view thus seems to entail that there could be a being with beliefs that it is in conscious states of various sorts on various occasions yet is never in such states. But perhaps there could be a silicon-based robot that is such a being—possessed of the relevant concepts but entirely devoid of sentience (McLaughlin and Tye 2003b). The shock of such a possibility is somewhat lessened if primary possession of concepts of conscious states requires acquaintance with such states (Peacocke 1998), so that the robot could possess them only in a secondary way—by communicative interaction with conscious beings that possess them in a primary way.

Another "independent existence" concern with the causal view is that it entails the possibility of beings who are in conscious states but lack the capacity to be introspectively aware of them, and so who are "self-blind" with respect to them (Shoemaker 1984b, 1984c). Introspectionists, however, maintain that introspective awareness of a conscious state consists of a belief that one is in the state, a belief formed by direct acquaintance with the state. Animals seem self-blind in the sense in question: they do not form beliefs about what conscious states they are in, for they lack the requisite concepts to do so. Indeed, animals do not introspect their conscious states; they are conscious, but not self-conscious. So, this sort of self-blindness may seem not to count against introspectionism. Nevertheless, there is a sense in which animals are aware of their pains or itches, for instance; that is why the dog yelps or scratches. Indeed it seems that their attention might be riveted on their sensation. It remains an open question whether the relevant mode of attention can be captured by a model of introspective attention as belief-formation or whether further distinctions are called for.

It has been claimed that when we try to direct our attention to our visual experience in order to introspect it, we seem to find ourselves only inspecting the scene before our eyes (Moore 1903; Harman 1990, Dretske 1994, 1999; Sturgeon 2000; Tye 2000). It is thus claimed that visual experience is phenomenologically "transparent" or "diaphanous." And some philosophers claim that all conscious states are diaphanous (Tye 2000). The phenomenological thesis of transparency seems most plausible for visual experiences and least plausible for bodily sensations. But it is maintained that even when we attend to a toothache, our attention seems focused on a feature of the tooth itself, however alarming we may find that feature.

In the light of these phenomenological considerations, a "displaced-perception model" of first-person knowledge of experience has been proposed (Dretske 1994, 1999; Tye 2000). The leading idea in the visual case is this: when we are attentively aware that we are having a visual experience, our "awareness-that" is not based on direct awareness of the experience but rather on awareness of the scene before our eyes. Our awareness of the experience is indirect, because we are aware of it by being aware of the scene. Nevertheless, if we have mastered the concept of visual experience, we can come to be aware that we are having a certain visual experience, without recourse to consciously drawing inferences.

Hallucination seems to pose no problem for the phenomenological transparency thesis itself: Perhaps, whenever we visually hallucinate, we seem to be aware only of a scene. But hallucination poses a problem for the displaced-perception model if, when we (completely) hallucinate, we are not actually aware of any scene at all. If there are sense data (Ayer 1940), then we will actually be aware of a scene, even when we completely hallucinate, for sense data would constitute a scene. But the leading proponents of the displaced-perception model are physicalists and so deny that there are sense data. Proponents of the model have tried to accommodate hallucination by maintaining that in such a case one is aware of a type of scene, despite not being aware of any actual instance of it. Whether this model applies to visual experience and all conscious states remains a topic of controversy.

Our ordinary epistemic practices seem to rely not only on the presumption that our (sincere) first-person ascriptions of conscious states (for example, 'I am in pain') are prima facie true but also on the presumption that our first-person ascriptions of beliefs (for example, "I believe that P'), desires, and intentions are prima facie true. It has been claimed that the social-psychological data about confabulation shows the latter presumption to be unfounded. But, arguably, the data seem to show only that we have a tendency to confabulate when under pressure to explain how we arrived at our propositional attitudes or made choices; thus the data seems not to raise an unanswerable challenge to first-person authority. In any case, many contemporary philosophers claim that whatever role introspection may play in explaining our first-person authority as self-ascribers of conscious states, it

has little to do with our first-person authority concerning our propositional attitudes (Davidson 1984, 1986; Gallois 1996; Moran 2001).

Even if we indeed introspect conscious states (as these were characterized earlier), we do not introspect our beliefs, desires, or intentions. Indeed, we do not even introspect our attitudinal emotions (fear that P, anger that Q, relief that R, and so on). Such states can count as conscious, but only in the sense that they can have characteristic manifestations in consciousness; and (at best) we introspect only conscious states that manifest them. Thus, we may introspect an impulse, but not a desire; a feeling of anger, but not an attitude of anger; a thought that P, but not a belief that P. Indeed, to be aware of one's belief that P is just to be aware that one believes that P; and similarly for the other cases (Shoemaker 1994b). Just as we can typically know what we believe without observing our behavior, we can typically know what we believe without introspecting.

Moreover, although we sometimes know that we believe something as a result of assessing evidence that we do, such a case seems atypical. When we ask ourselves whether we believe that P, want X, or intend to A, we usually do not reflect on evidence concerning whether we believe that P, want X, or intend to A. Of course, we sometimes do that. But in response to the questions we typically reflect, respectively, on whether P, whether X has some attractive feature, and whether we ought to do A (Evans 1981, Gallois 1996, Moran 2001). Although we typically do that, reasons for believing that P is true are not reasons for believing that one believes that P; and reasons for believing that one ought to A are not reasons for believing that one intends to A (similarly for the desire case). Rather, they are, respectively, reasons to believe that P and reasons to intend to do A. So, the question of how such reflection leads to knowledge of our beliefs, desires, and intentions persists.

Philosophers who seek a role for introspection here will claim that, when we engage in such deliberative reflective reasoning, we can be introspectively aware of our occurrent thoughts. Philosophers who reject any role for introspection here will claim that even if we can indeed introspectively observe manifestations of propositional attitudes in consciousness and so have more "observational data" than others who can only observe manifestations of our attitudes in our overt verbal and nonverbal behavior, the fact that we have such additional observational data will not explain our first-person authority about our attitudes. Moreover, occurrently thinking that P is a mental act—indeed a basic mental act: something we do, but not by doing something else (Moran 2001). Our knowledge of what we are occurrently thinking is knowledge of something that we are doing. Our distinctively characteristic knowledge of our basic actions may not be introspective. What explains first-person authority about our propositional attitudes and basic actions remains an open issue.

Many philosophers have related first-person authority about attitudes and actions to the fact that attitudes and actions (unlike bodily sensations, imagery, or sense experiences) can be rational or irrational. One view is that our practice of attributing propositional attitudes is essentially an interpretive practice governed (in part) by constitutive principles of rationality, and the presumption of first-person authority is required for interpretation to be possible (Davidson 1984, 1986). Another view is that the functional organization required to be a rational agent guarantees that a rational agent will, for the most part, be reliable in his or her beliefs about what propositional attitudes and experiences he or she has (Shoemaker 1994a, 1994b, 1994c). Yet another view seeks to explain our first-person authority in terms of rational commitment and first-person deliberation (Moran 2001). There are other very influential views (Burge 1988, 1993).

Belief, desire, intention, and occurrent thought are modes of intentionality; states of these (and other intentional) types have representational content. One issue is how one knows which of these (or other intentional) types a given intentional state falls under; another issue is how one knows what the content of the state is. Thus, there is, for instance, the issue of how one knows that one's belief that P is a belief (rather, than, say, a desire); and there is the issue of how one knows that one's belief is a belief that P (rather than a belief that something else is the case).

The leading contemporary theories of mental content are externalist theories, according to which the content of a mental state fails to supervene on intrinsic states of the subject (Putnam 1975, Burge 1979). On these views, two intrinsic duplicates (for example, an inhabitant of Earth and her doppelgänger on Twin Earth) could be in mental states with different contents. Some externalist theories hold that content depends on historical context (Dretske 1988), and according to others, it depends on social context (Burge 1979). There has been extensive debate about whether content externalism is compatible with our having first-person authority or privileged first-person knowledge concerning what we think. Some philosophers argue for incompatibilism (for

example, Boghossian 1989, 1997; McKinsey 1991). Some argue for compatiblism (for example, Davidson 1984, 1986; Burge 1988, 1993; Brucekner 1992; Heil 1992; Gallois 1996; Peacock 1998; Davies 1998; Gibbons 1996; Falvey and Owens 1994; McLaughlin and Tye 1998a, 1998b; McLaughlin 2000, 2003a; Brown 2004).

Here is an example of one of the leading incompatibilist lines of argument (McKinsey 1991, Boghossian 1997). For any of the content-externalists theories in question, there will be some contingent environmental proposition E such that E can be known only on the basis of empirical evidence, yet the theory will entail that it is a conceptual truth that if we are thinking that P, then E. Thus, if we could have privileged first-person knowledge that we are thinking that P, it follows that we would be able to infer that E and thereby come to know it on some basis other than empirical evidence. Some compatibilists have responded that the relevant contingent environmental propositions will be ones that can thereby be known on a basis other than empirical evidence, however surprising that might be (Warfield 1998, Sawyer 1998). But by far the more prevalent compatibilist response is to try to show that combinations of the relevant content-externalist and privileged self-knowledge theses do not lead to this result (Brueckner 1992; Davies 1998; McLaughlin and Tye 1998a, 1998b; 2003a).

See also Behaviorism; Consciousness; Descartes, René; Freud, Sigmund; Introspection; Intuition; Memory; Perception; Personal Identity; Ryle, Gilbert; Self; Socrates; Unconscious.

Bibliography

Alston, W. P. “Varieties of Privileged Access.” *American Philosophical Quarterly* 8 (1971): 223–241.

Armstrong, D.M. “Is Introspective Knowledge Incorrigible?” *Philosophical Review* 62 (1963): 417–32.

Ayer, A. J. *The Foundations of Empirical Knowledge*. London: Macmillan, 1940.

Boghossian, P. “Content and Self–Knowledge.” *Philosophical Topics* 1 (1989): 5–26.

Boghossian, P. “What the Externalist Can Know A Priori.” *Proceedings of the Aristotelian Society* 97 (1997): 161–175.

Broad, C. D. *The Mind and Its Place in Nature*. London: Routledge and Kegan Paul, 1925.

Brown, J. *Anti-Individualism and Knowledge*. Cambridge, MA: MIT Press, 2004.

Brueckner, A. “What an Anti–Individualist Knows A Priori.” *Analysis* 52 (1992): 111–118.

Burge, T. “Content Preservation.” *The Philosophical Review* 102 (1993): 457–488.

Burge, T. “Individualism and Self Knowledge.” *The Journal of Philosophy* 83 (1988): 649–663.

Burge, T. “Individualism and the Mental.” *Midwest Studies in Philosophy* 4 (1979): 73–121.

Chalmers, D. “The Content and Epistemology of Phenomenal Belief.” In *Consciousness: New Philosophical Essays*, edited by Q. Smith and A. Jokic. Oxford, 2003.

Chisholm, R. *Theory of Knowledge*. 2nd ed. Prentice Hall, 1977.

Davidson, D. “First Person Authority.” *Dialectica* 38 (1984): 101–111.

Davidson, D. “Knowing One’s Own Mind.” *Proceedings of the American Philosophical Association* 60 (1986): 441–458.

Davies, M. “Externalism, Architecturalism, and Epistemic Warrant.” In *Knowing Our Own Minds*, edited by C. Wright, B. Smith, and C. Macdonald. Oxford, 1998.

Dennett, D. *Consciousness Explained*. Boston: Little, Brown & Company, 1991.

Descartes, R. *The Philosophical Writings of Descartes.* Vols. 1 and 2. Translated by J. Lottingham, R. Stoothoff, and D. Murdoch. Cambridge, U.K.: Cambridge University Press, 1985.

Dretske, F. “Introspection.” *Proceedings of the Aristotelian Society* 94 (1994): 263–278.

Dretske, F. “The Mind’s Awareness of Itself.” *Philosophical Studies* 95 (1999): 103–124.

Evans, G. *The Varieties of Reference*. New York: Oxford University Press, 1982.

Falvey, K. and Owens, J. “Externalism, Self–Knowledge, and Skepticism.” *The Philosophical Review* 103 (1994): 107–137.

Freud, S. *The Interpretation of Dreams*. 3rd ed. Translated by A. A. Brill. London: Allen and Unwin, 1923.

Gallois, A. *The World Without, the Mind Within: An Essay on First-Person Authority*. Cambridge, U.K.: Cambridge University Press, 1996.

Gertler, B. “Self–Knowledge.” *The Stanford Encyclopedia of Philosophy*, Spring 2003 edition, edited by E. N. Zalta. Available from http://plato.stanford.edu/archives/spr2003/entries/self–knowledge/.

Gibbons, J. “Externalism and Knowledge of Content.” *The Philosophical Review* 105 (1996): 287–310.

Goldman, A. *Pathways to Knowledge: Private and Public*. New York: Oxford University Press, 2002.

Harman, G. “The Intrinsic Quality of Experience.” In *Philosophical Perspectives*, edited by J. Tomberlin. Atascadero, CA: Ridgeview, 1990.

Heil, J. *The Nature of True Minds*. Cambridge, U.K.: Cambridge University Press, 1992.

Hill, C. S. *Sensations*. New York: Cambridge University Press, 1991.

Locke, J. *An Essay concerning Human Understanding*. Oxford, 1690/1975.

Ludlow, P., and Martin, N. eds. *Externalism and Self-Knowledge*. Stanford, CA: CSLI Publications, 1998.

Lycan, W. G. *Consciousness and Experience*. Cambridge, MA: MIT Press, 1996.

Macdonald, C. “Externalism and Authoritative Self–Knowledge.” In *Knowing Our Own Minds*, edited by C. Wright, B. Smith, and C. Macdonald. Oxford: Oxford University Press, 1998.

Macdonald, C. “Shoemaker on Self–Knowledge and Inner Sense.” *Philosophy and Phenomenological Research* 59 (1999): 711–738.

McKinsey, M. “Anti-Individualism and Privileged Access.” *Analysis* 51 (1991): 99–121.

McLaughlin, B. P., and A. Rorty, eds. *Perspectives on Self-Deception*. Berkeley: University of California Press, 1988.

McLaughlin, B. P., and M. Tye. “Color, Consciousness, and Color Consciousness.” In *Consciousness: New Philosophical Essays*, edited by Q. Smith and A. Jokic. Oxford: Clarendon Press, 2003.

McLaughlin, B. P., and M. Tye. “Externalism, Twin Earth, and Self-Knowledge.” In *Knowing Our Own Minds*, edited by C. Wright, B. Smith, and C. Macdonald. Oxford: Clarendon Press, 1998a.

McLaughlin, B. P., and M. Tye. “Introspecting Thoughts.” *Facta Philosophica* 3 (2001): 77–83.

McLaughlin, B. P., and M. Tye. “Is Content Externalism Compatible with Privileged Access?” *The Philosophical Review* (1998b): 349–380.

McLaughlin, B. P., and M. Tye. “McKinsey’s Challenge, Warrant Transmission, and Skepticism.” In *New Essays on Semantic Externalism and Self-Knowledge*. Edited by S. Nuccetell, 79–96. Cambridge, MA: MIT Press, 2003a.

McLaughlin, B. P., and M. Tye. “A Naturalist Phenomenal Realist Response to Block’s Harder Problem.” *Philosophical Issues* 13 (2003b): 163–204.

McLaughlin, B. P., and M. Tye. “Skepticism, Externalism, and Self–Knowledge.” *The Aristotelian Society*, supp. vol. 74 (2000): 93–118.

Moore, G. E. “The Refutation of Idealism.” In *Philosophical Papers*. Routledge and Kegan Paul, 1903.

Moran, R. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press, 2001.

Parke, H. W. *A History of the Delphic Oracle*. Oxford: B. Blackwell, 1939.

Peacocke, C. “Conscious Attitudes, Attention, and Self-Knowledge.” In *Knowing Our Own Minds*, edited by C. Wright, B. Smith, and C. Macdonald. Oxford: Clarendon Press, 1998.

Putnam, H. “The Meaning of ‘Meaning.’” In *Mind, Language and Reality*. New York: Cambridge University Press, 1975, 215–271.

Ryle, G. *The Concept of Mind*. New York: Hutchinson’s University Library, 1949.

Sawyer, S. “Privileged Access to the World.” *Australasian Journal of Philosophy* 76 (1998): 523–533.

Sellars, W. “Empiricism and the Philosophy of Mind.” In *Perception, and Reality*, edited by W. Sellars. London: Routledge and Kegan Paul, 1963.

Shoemaker, S. *Identity, Cause, and Mind: Philosophical Essays*. Cambridge, U.K.: Cambridge University Press, 1994a.

Shoemaker, S. “Introspection.” in S. Guttenplan ed. *A Companion to the Philosophy of Mind. Cambridge*, 1994c, 395–400.

Shoemaker, S. “Self–Knowledge and ‘Inner Sense.’” *Philosophy and Phenomenological Research* 54 (1994b): 249–314.

Sturgeon, S. *Matters of Mind: Consciousness, Reason, and Nature*. New York: Routledge, 2000.

Tye, M. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press, 2000.

Warfield, T. “A Priori Knowledge of the World: Knowing the World by Knowing Our Minds.” *Philosophical Studies* 92 (1998): 127–147.

Williamson, T. *Knowledge and Its Limits*. Oxford: Oxford University Press, 2000.

Wright, C. “Self–Knowledge: The Wittgensteinian Legacy.” In *Knowing Our Own Minds*, edited by C. Wright, B. Smith, and C. Macdonald. Oxford: Clarendon Press, 1998.

Ziff, P. “About Behaviorism.” *Analysis* XVIII (1958).

Brian P. McLaughlin (2005)

SELF-PREDICTION

In recent years philosophers have produced arguments designed to prove that not all human behavior can be predicted or otherwise known in advance, and these arguments have been taken to be relevant to the problem of freedom of the will as well as to the question whether there can be genuine behavioral sciences. Specifically, it is argued that in certain circumstances it is logically impossible that one should come to know decisions, and actions for whose occurrence decisions are necessary conditions, in advance of the occurrence of such decisions. This has been interpreted as a refutation of determinism.

Two antipredictive arguments will be presented separately, and later their import when taken together will be discussed. The first concerns the scientific defectiveness of predictions that influence the predicted event, and the second concerns the logical impossibility of a person’s knowing now what he will decide only at some future time.

INFLUENCE OF PREDICTIONS

It is a familiar fact that some prophecies and predictions are self-fulfilling in the sense that the prediction itself produces the predicted event—for example, when all the stock market tip sheets predict that stock *x* will drop sharply in the next few weeks. We also know, for similar reasons, that some predictions are self-defeating. For example, Jones predicts that he will, as usual, take the easy way out of a difficulty, but then, to prove to himself that he can do better, he does just the opposite. This prediction affected his deliberation and caused him to make a decision opposite to the one he had predicted. Now, the argument that follows does not maintain that a person’s predictions of his own future decisions are necessarily or always self-defeating; instead, it maintains that it is logically impossible that by considering causes a person

should come to know that his final prediction of what he will decide is not self-defeating, and it maintains that the attempt to achieve such knowledge involves an infinite regress. In other words, this antipredictive argument purports to prove that predictions of one's own future decisions on the basis of antecedent causal conditions cannot possibly be scientifically complete.

It is necessary to state some assumptions and restrictions required by the argument. The first assumption is that decisions are events and hence are the sorts of things that can be caused; many philosophers would reject this assumption. Second, the argument concerns only causal knowledge of future decisions, by which is meant predictions derived with scientific adequacy from what one knows to be all the relevant antecedent causes of the decision, as distinct from predictions not known to be based on all the relevant causes and which consequently yield only a likelihood of the decision's occurrence. Finally, the argument aims to prove only that it is logically impossible for a person to have causal knowledge of his own decision in advance of making such a decision.

Let us assume, then, that some set of circumstances *C* is causally sufficient for a person *S* to make decision *D* and that *S* has unlimited knowledge of past circumstances and relevant causal laws. Can *S* come to know that *C* is sufficient for *D*? *S* may come to make a prediction *P* that past circumstances *C* are sufficient for *D*. We have supposed that as a matter of fact *C* is causally sufficient for *D*, but *S* nevertheless cannot know that this is so unless he also knows that there are no contrary causes. That is, before *S* can know that *C* is sufficient for *D* he must also know that there is no other circumstance which, together with *C*, is sufficient for not-*D*. One such probable cause of not-*D* is the prediction itself. Therefore, *S* cannot know that *C* is sufficient for *D* unless he knows that it is false that

(1) *C* plus *P* are causally sufficient for not-*D*.

S has been allowed unlimited knowledge of past circumstances and relevant causal laws, hence *S* can know that (1) is false, that is, he can know that making the prediction will not cause him to make a different decision. It does not follow, however, that *S* now can know that *C* is sufficient for *D*, for the same problem recurs: *S*'s knowledge that (1) is false, which we will call P_1, is a new datum and is itself a possible cause of not-*D*. Therefore, *S* cannot know that *C* is sufficient for *D* unless he knows that it is false that

(2) *C* plus P_1 are causally sufficient for not-*D*.

And *S*'s knowledge that (2) is false, or this knowledge plus his feelings or attitudes toward (2), constitute a further possible contrary cause, P_2. Thus, an infinite regress arises, within which the agent's prediction on the basis of some evidence *C* or his revision of the prediction or his final thoughts about the prediction are relevant data in addition to the data upon which the prediction was based. *S*'s calculating of causes cannot possibly "catch up" with the number of possible causes that must be examined if the prediction is to be scientifically complete, for the final results obtained cannot themselves also be part of the basis of one's prediction.

When one attempts to predict a supernova, it is true that in this case, too, the final prediction arrived at is necessarily excluded from the data upon which the prediction is based. However, although it is logically possible that predictions or thoughts about predictions can produce or impede a supernova, it is not scientifically possible that they do so. Therefore, the infinite regress argument is no obstacle to knowledge of, for example, scientific laws or stellar events but concerns only particular events that can be produced or prevented by human agency. And it is clearly applicable to attempted predictions of one's own decisions because we know that speculations and predictions about what one is likely to decide are always among the conditions most likely to be determinative of what one will in fact decide.

COUNTERARGUMENTS FAVORING DETERMINISM. The view that this first antipredictive argument casts doubt on determinism may be challenged in a number of ways:

(*a*) The argument presents no obstacle to the existence of a complete causal explanation of one's own past decisions.

(*b*) There is no logical obstacle to a person's predicting a future decision of someone other than himself, although such prediction does confront a methodological difficulty. That is, suppose that *A* predicts a future decision of *B*'s and resolves not to tell *B* the prediction. Then it appears that *A* must also predict something about himself; namely, that he will not later decide to revoke his past decision and tell *B*, after all—and this, according to the infinite regress argument, *A* cannot possibly do. One complication here is the question whether the regress argument precludes *A*'s predicting that he will make no decisions at all during a certain future period; if the regress argument does not preclude this, then *A* can predict that he will not change his mind and tell the original prediction to *B*. But in any case the solution seems to lie in

having *A* make his prediction of *B*'s decision from a dungeon or a distant planet or in such a way that he has no time to communicate with *B* in advance of *B*'s making his decision; that is, perhaps it is sufficient that it be physically (although not logically) impossible that *A* should ruin the impeccable scientific basis of his prediction by telling *B*.

(*c*) The regress argument shows no peculiarity of human or even of sentient beings. For it is easy to imagine a simple machine, for which no one would dream of claiming free will or moral responsibility, the behavior of which could not possibly be predicted in circumstances similar to those previously described. We need only suppose that the machine can do two things, *x* and *y*, that a prediction of either of these things, punched into a card, can be inserted in the machine, and that we announce our predictions of what the machine will do by inserting appropriately punched cards into the machine. The machine is built to do *x* when fed the prediction "machine will do *y*" and to do *y* when fed the prediction "machine will do *x*." The situation in which a prediction of a person's decision is defective is fully as artificial as this, and in each situation the prediction is defective for the same reason. In each case, given the causal hypothesis, one can in principle make a scientifically impeccable prediction of what will occur only if neither the person nor the machine is allowed to be influenced by the prediction. Meaning "*y*" when one inserts the card saying "machine will do *x*" into the machine is equivalent to telling a person he will decide not-*D* when one knows that telling him this will cause him to decide *D*.

It can be argued that the first antipredictive argument shows only that given the causal hypothesis, it is still possible to make predictions competently and incompetently and that one of countless ways in which one can make predictions incompetently is to allow one's prediction to disturb the system that one is trying to predict. However, although it may be the case that the self-defeating prophecy and the self-fulfilling prophecy are equally explicable and, in general, equally avoidable phenomena, it appears that the special situation in which the self-defeating prophecy is unavoidable is important to us—namely, the situation in which we attempt to predict our own decisions. The regress argument also poses a methodological problem for social scientists who wish to circulate predictions of human behavior, but it does not show that there is any event that in principle cannot be predicted.

LOGICAL IMPOSSIBILITY OF SELF-PREDICTION

The second antipredictive argument appears to follow from the analytic truth that one cannot know now what, by hypothesis, one will not know until some later time. Thus, one form of this argument (see Karl Popper, "Postscript: After Twenty Years") maintains that exact historical prophecy is incompatible with the fact of advancing knowledge. That is, it is impossible to predict the future decisions and actions of people because these future decisions and actions will be formed and done on the basis of knowledge that, by hypothesis, no one now possesses.

Another form of the argument maintains that it is logically impossible for a person to know what he will decide to do before he actually makes his decision (see Stuart Hampshire, *Thought and Action*; Carl Ginet, "Can the Will Be Caused?"; and D. F. Pears, *Freedom and the Will*). It is claimed that if a person knows or thinks he knows what he will try to do tomorrow, then either he has already decided what he will try to do or he believes that what he will try to do is not up to him. In neither of these two cases can he decide what he will try to do, for in each case there is nothing for him to decide. Decision is making up one's mind about what one will try to do or about what one will acquiesce in; therefore, to say that one will decide tomorrow appears to entail that there is something one will know then and which, by hypothesis, one does not know now.

However, there is a difficulty here. What is it that one knows as a result of decision and that one cannot know prior to the decision? From the fact that a person has decided to do something, it does not follow that he knows what he will do or try to do in the future. Decision does not give one knowledge of anything that will occur in the future because the mere fact that a person has decided does not ensure that he will not falter, change his mind, or die tomorrow. Hence, it appears to be mistaken to assume that because decision entails ignorance prior to decision, this ignorance is of something which one will know later as a result of decision; what one comes to know when one decides is nothing in addition to the decision itself and not any fact about the future. The reason for this appears to be that "decision" is an intentional concept.

Sometimes a person claims to know what a future decision of his will be, and various explanations of his supposed mistake can be made: (*a*) He has already decided, and he confuses with the act of decision itself some future reaffirmation, announcement, or implementation of his decision. (*b*) He has tentatively decided and plans at the last moment to reappraise his decision, but he

thinks that he knows the result of that reappraisal because of his tentative decision. In this case, if he does not deliberate again at the last moment, then he merely reaffirms what he has already decided, and if he does deliberate again, then it is impossible that he should know in advance the result of his deliberation, even though this new decision agrees with his earlier tentative decision. (*c*) He construes a future reaffirmation of a decision already made to be a new decision because its time, place, or context differs from that in which he first decided. (*d*) He confuses a guess, likelihood, or probability with knowledge of his future decision.

It has also been claimed (for example, by Richard Taylor, in "Deliberation and Foreknowledge") that if a person knows or thinks he knows what he will do in the future, then it is impossible for him to deliberate about what he will do, for deliberation also presupposes ignorance. "Jones is deliberating whether to do *x*" entails "Jones does not know whether or not he will do *x*." But here a distinction must be made between the agent's belief or knowledge that he will do a particular act in the future and the agent's belief or knowledge that this particular act he will do is in some sense not up to him. If a person believes that he will do *x*, he cannot deliberate whether to do *x*, even though he believes that he will do *x* freely, that what he does is up to him. On the other hand, if a person believes that what he will do is not up to him, then he cannot deliberate whether to do *x*, even though he lacks knowledge or belief about what he will do. Hence, although it has been claimed that both foreknowledge and lack of freedom preclude deliberation and decision, these claims nevertheless require separate argument, and only foreknowledge is relevant to self-prediction and the paradoxes thereof.

It might be thought that the two antipredictive arguments are not truly distinct, and indeed some philosophers have written as though these arguments were but two approaches to the same logical point. But they are distinct, except insofar as they can be put to similar purposes. The first argument applies to all predictions that can causally influence the events predicted, whether these events happen to be decisions, revolutions, or stock market trends. It is thus broader in scope and does not require that the event also be of that special sort which, in certain circumstances, is logically impossible to know in advance. The second argument attacks the very idea of foreknowledge, however obtained, of occurrences that entail prior ignorance and does not, as does the first argument, attack the scientific adequacy of predictions that can influence the predicted events.

LOGICAL IMPOSSIBILITY OF CAUSING DECISIONS

Many philosophers would maintain that if some set of antecedent conditions is causally sufficient for the occurrence of an event, then it is logically possible that the event be predicted or known prior to its occurrence. From this claim, together with the second antipredictive argument, can be constructed the following argument that attempts to prove that it is logically impossible that decisions have causes (see Ginet, op. cit.): If it is logically possible for a decision to be caused, then it is logically possible for a person to know what his own decision will be before he makes his decision; it is not logically possible for a person to know what his own decision will be before he makes his decision; therefore, it is not logically possible for a decision to be caused.

This argument is, in the following way, of more apparent relevance to the traditional problem of freedom of the will and in particular to a theory of human agency: Let us suppose that decisions are necessary conditions for the occurrence of certain actions, and let us suppose further that decisions are part of the causes of such actions. If so, then any set of causes sufficient for the occurrence of such an action must include a decision as part of the set, for whatever is sufficient for something to occur must include everything necessary for that thing to occur. But the decision, by the preceding argument, is uncaused, and therefore no set of causes existing prior in time to the decision can be sufficient for the occurrence of the action. The decision can thus be viewed as a partial, uncaused cause of the action, which, together with ordinary causes, is sufficient for the occurrence of the action.

Difficulties of the following sort have been raised against the argument that maintains that it is impossible that decisions be caused: First, it has been doubted that it follows from the causal hypothesis that it is possible for a person to predict his own decisions; for the possibility of predictability in principle need not include the possibility of predictability in all possible circumstances (see A. J. Stenner, "On Predicting Our Future"). As we have seen, it is not obvious that paradoxes arise when we suppose someone to predict decisions of persons other than himself. Second, a premise of this argument maintains that from the hypothesis that decisions are caused, it follows that one could in principle make a scientifically adequate prediction, based on knowledge of antecedent causes, of one's own future decision. But the first of the two antipredictive arguments claims that this does not follow at all, because it is impossible to establish that one's prediction has no contrary influence on the predicted event.

That is, the first antipredictive argument, if sound, shows that the causal hypothesis does not entail the apparent absurdity that in principle one could, by considering antecedent conditions and relevant causal laws, come to know one's own decisions in advance.

Bibliography

Canfield, J. "Knowing about Future Decisions." *Analysis* 22 (6) (1962): 127–129.

Fain, Haskell. "Prediction and Constraint." *Mind* 67 (1958): 366–378.

Ginet, Carl. "Can the Will Be Caused?" *Philosophical Review* 71 (1) (1962): 49–55.

Hampshire, Stuart. *Thought and Action*. London: Chatto and Windus, 1960.

Hampshire, Stuart, and H. L. A. Hart. "Decision, Intention and Certainty." *Mind* 67 (265) (1958): 1–12.

Herbst, P. "Freedom and Prediction." *Mind* 66 (1957).

Lucas, J. R. "Freedom and Prediction." Part 1. *PAS*, Supp. 41 (1967): 163–172.

Lyon, A. "The Prediction Paradox." *Mind* 68 (1959): 510–517.

MacKay, D. M. "Brain and Will" (BBC Third Program, 1957). Printed in *A Modern Introduction to Philosophy*, edited by Paul Edwards and Arthur Pap. New York: Free Press, 1965.

MacKay, D. M. "On the Logical Indeterminacy of a Free Choice." *Mind* 69 (1960).

Moran, Richard. *Authority and Estrangement: An Essay on Self-Knowledge.* Princeton, NJ: Princeton University Press, 2001.

Moran, Richard. "Self-Knowledge: Discovery, Resolution, and Undoing." *European Journal of Philosophy* 5 (1997): 141–161.

Murdoch, Iris, Stuart Hampshire, Patrick Gardiner, and D. F. Pears. "Freedom and Knowledge." In *Freedom and the Will*, edited by D. F. Pears, 80–104. London, 1963.

Perry, David. "Prediction, Explanation, and Freedom." *Monist* 49 (1965): 234–247.

Popper, Karl. "Postscript: After Twenty Years." In his *Logic of Scientific Discovery*. London: Hutchinson, and New York: Basic, 1959.

Rabinowicz, Wlodek. "Does Practical Deliberation Crowd Out Self-Prediction?" *Erkenntnis* 57 (2002): 91–122.

Schlesinger, George. "The Unpredictability of Free Choices." *British Journal for the Philosophy of Science* 25 (1974): 209–221.

Stenner, A. J. "On Predicting Our Future." *Journal of Philosophy* 61 (14) (1964): 415–428.

Stocker, Michael. "How to Prevent Self-Prediction." *Journal of Philosophy* 65 (1968): 475–476.

Taylor, Richard. "Deliberation and Foreknowledge." *American Philosophical Quarterly* 1 (1) (1964).

Andrew Oldenquist (1967)
Bibliography updated by Benjamin Fiedor (2005)

SELLARS, ROY WOOD
(1880–1973)

Roy Wood Sellars, the American critical realist, taught philosophy at the University of Michigan. Although he was never as well known outside philosophical circles as some of his contemporaries, after the publication of his first book, *Critical Realism,* in 1916, Sellars maintained a substantial reputation among his fellow philosophers as a vigorously independent thinker. His thought was rigorous and critical; he never yielded to the fashionable movements of the day but steadfastly pursued his own original insights into basic philosophical problems.

The core of Sellars's philosophy is epistemological. He is concerned with showing that the critical realism of the philosopher is related to the "natural realism" of the "plain man." The philosopher reflects on the plain man's uncritical view of knowledge, which he clarifies and refines so that it is philosophically justifiable, but he does not vitiate its essential insistence upon the independence of the object of knowledge. The most significant element in Sellars's vindication of realism is his revision of the theory of perception, which he describes as a process of interpretation of sensa, as mediated by factors both external and internal to the perceiving subject. This view of perception avoids both the simplistic claim of natural realism that things reveal themselves directly in perception and the subjectivist claim that the objects of perception are ideas rather than things. Knowledge, too, is a complex process and occurs at various levels of complication. Its ultimate biological source is to be found in the adjustment of the organism to its environment; its ultimate outreach is in scientific knowledge, which replaces the relativity of individual perspectives by close approximations to exact measurement. Whether on the implicit organic level or on the highly explicit and self-critical scientific level, we know that we know when the content of our beliefs corresponds to the externally observed state of affairs.

Working from this epistemological position, Sellars developed an evolutionary cosmology and a materialistic ontology, carrying on his insight that there are levels, or "gradients," of being. Even the higher levels like life and mind, which emerge under most favorable conditions, are, however, physical systems. Sellars's materialism is nonreductive, but he insists that "life is not a nonnatural force coming from outside, but a term for the new capacities of which nature has found itself capable." On the valuational side, Sellars argues from these positions to a humanistic theory of ethics and religion (he was one of

the major contributors to the composition of the Humanist Manifesto of 1933) and to a politics of democratic socialism.

See also Critical Realism.

Bibliography

WORKS BY SELLARS

Critical Realism. Chicago: Rand McNally, 1916.

The Next Step in Democracy. New York: Macmillan, 1916.

The Next Step in Religion. New York: Macmillan, 1918.

Evolutionary Naturalism. Chicago: Open Court, 1922.

Religion Coming of Age. New York: Macmillan, 1928.

The Philosophy of Physical Realism. New York: Macmillan, 1932.

The Philosophy of Physical Realism. New York: Russell & Russell, 1966.

Lending a Hand to Hylas. N.p., 1968.

Reflections on American Philosophy from Within. Notre Dame, IN: University of Notre Dame Press, 1969.

The Principles, Perspectives, and Problems of Philosophy: An Exploration in Depth. New York: Pageant Press International, 1970.

Patterns and Political Horizons. Nashville, TN: Aurora, 1970.

Social Philosophy and the American Scene: Materialism and Human Knowing. New York: Oriole Editions, 1970.

Principles of Emergent Realism: Philosophical Essays. St. Louis, MO: W.H. Green, 1970.

Neglected Alternatives: Critical Essays. Lewisburg: Bucknell University Press, 1973.

WORKS ON SELLARS

Bahm, Archie, R. M. Chisholm et al. "A Symposium in Honor of Roy Wood Sellars." *Philosophy and Phenomenological Research* 15 (1954): 1–97.

Reck, Andrew. *Recent American Philosophy.* New York: Pantheon, 1964.

J. L. Blau (1967)
Bibliography updated by Michael J. Farmer (2005)

SELLARS, WILFRID
(1912–1989)

Wilfrid Stalker Sellars, an American philosopher and teacher, was born in Ann Arbor, Michigan, the son of Roy Wood Sellars, the American critical realist who taught at the University of Michigan. Wilfrid Sellars's early education took place in the United States and in France, where he attended the lycées Montaigne and Louis le Grand; it was continued at the University of Michigan (BA, 1933), the University of Buffalo (MA, 1934), and Oxford University, where he was a Rhodes scholar and received a BA with first-class honors in philosophy, politics, and economics. He received an MA from Oxford in 1940. After a year at Harvard University he began his career as a teacher of philosophy in 1938 at the University of Iowa. During the war he spent several years as an officer in the Naval Reserve, and in 1946 he went to the University of Minnesota, where he eventually became professor of philosophy, chairman of the philosophy department, founding co-editor of the journal *Philosophical Studies*, and a member of Herbert Feigl's Minnesota Center for the Philosophy of Science. In 1959 he joined the faculty of Yale University, and in 1963 he moved to the University of Pittsburgh, where he became University Professor of Philosophy and Research Professor of the Philosophy of Science. Apart from numerous interludes as a visiting professor at other institutions, he remained at Pittsburgh until his death.

Although Sellars became an extremely prolific writer, in the early years of his career he had great difficulty putting his ideas on paper. His first scholarly essay, third in his list of publications, was "Realism and the New Way of Words"; it underwent seventeen major revisions, Sellars said in his "Autobiographical Reflections" (1975), before it finally appeared in print. In spite of its striking originality, his early work was strongly influenced by the logical empiricist movement, particularly by the work of Rudolf Carnap; in one essay, "Epistemology and the New Way of Words," he declared that philosophy "is properly conceived as the pure theory of empirically meaningful languages." From the vantage point of the early twenty-first century, perhaps the most significant of his early essays would be "Concepts as Involving Laws and Inconceivable without Them" (1948) and "A Semantical Solution of the Mind–Body Problem" (1953). Both show him to have been well ahead of his time in analytic philosophy. In the former he offered a clarification of necessity and natural law that anticipated the treatment of these notions in recent possible-world semantics, and in the latter he developed a distinctly functionalist view of intentional states. (The early essays discussed here are included in the volume *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, edited by J. F. Sicha.)

Sellars's best-known philosophical work is the lengthy essay "Empiricism and the Philosophy of Mind," included in Sellars' *Science, Perception, and Reality*. This essay originated in lectures that Sellars gave in 1956 attacking what he called "the myth of the given." The cluster of ideas making up this doctrine was, he thought, the source of important errors in both the theory of knowledge and the philosophy of mind;by exposing the doc-

trine as a myth, he hoped to lay the groundwork for an acceptable form of empiricism and for a proper understanding of mental and sensory phenomena. The basic epistemic error prompted by the myth was the idea that empirical knowledge rests on a foundation of certain truth that is simply given to the mind—that is, knowable without inference—and provides the ultimate evidence for anything knowable by inference. The root error in the philosophy of mind prompted by the myth was the conviction that, merely by having sensory experiences and conscious thoughts, people gain theoretically satisfactory conceptions of those experiences and thoughts. These corresponding errors are related by the belief, commonly held by those who accept the myth, that foundational empirical knowledge concerns the sensory and psychological items, the mere having of which supposedly results in their being adequately conceived of or understood.

In attacking the errors he saw in the myth Sellars defended the view that empirical knowledge cannot have a foundation—that the supposedly basic knowledge of psychological fact presumed by the myth cannot exist independently of general knowledge relating psychological experience to linguistic and other behavior—and that theoretically adequate conceptions of anything can be obtained only by a process of learning and can be known to be adequate only by reference to scientific theorizing about the sensory and cognitive capabilities of human beings. He argued that "empirical knowledge … is rational not because it has a *foundation* but because it is a self-correcting enterprise which can put *any* claim in jeopardy, though not *all* at once" (1991, pp. 127–196). As for commonsense sensory and psychological concepts, he argued that it is illuminating to think of them as resulting from an attempt to explain intelligent, nonhabitual human behavior by postulating appropriate "inner episodes" in substantially the way that theoretical scientists explain facts about observable objects by postulating unobservable microcauses. In arguing this point he added that, when concepts of such inner episodes are developed, people can learn to use them in making first-person reports of what they are experiencing. Seen this way, psychological concepts are fundamentally intersubjective rather than private, and they are as subject to revision as any concept of theoretical science.

In "Philosophy and the Scientific Image of Man" (1960), also included in *Science, Perception, and Reality*, Sellars developed the thesis that, although theoretical science is a natural development of commonsense thought about the world, it is not evidentially dependent upon it. Like David Hume, Sellars thought that scientific thinking yields a theoretical picture of humans in the world that is incompatible with the commonsense—or, as he called it, the "manifest"—image of the same reality. These clashing images are not on a par, he thought; in purely descriptive respects, the scientific image is an improvement upon the manifest image, containing "successor concepts" to commonsense counterparts. (Water, on this view, is not identical with H_2O; the technical concept of H_2O applies to a common ingredient in most puddles, wells, clouds, and seas—one that is not accurately singled out by any commonsense concept.) A philosophically adequate picture of humans in the world is not fully descriptive, however; it is partly normative. Working out such a picture is an important philosophical task that has yet to be accomplished: the scientific image is not yet complete, and serious problems exist about how some normative matters can be incorporated into a significantly different image.

In later writings Sellars worked out highly original ideas on most central fields of philosophy. He produced, as Johanna Seibt (1990) observed, a unique scheme of "full scope nominalism," which purports to demonstrate the expendability of abstract entities for all their supposed explanatory functions; he worked out (he was the first to do so) a sophisticated "conceptual role" semantics: he developed a neo-Kantian view of moral obligation and the moral point of view; and he had original things to say about central figures and issues in the history of philosophy. At a time when systematic philosophy was decidedly out of fashion, Sellars pursued the synoptic vision of humans in the world that Plato spoke of in the Republic. In parody of Kant he liked to tell his students that in philosophy analysis without synthesis must be blind.

See also Carnap, Rudolph; Empiricism; Functionalism; Philosophy of Mind.

Bibliography

WORKS BY SELLARS

"Autobiographical Reflections." In *Action, Knowledge, and Reality: Critical Studies in Honor of Wilfrid Sellars*, edited by H.-N. Castañeda. Indianapolis: Bobbs-Merrill, 1975.

Philosophical Perspectives: History of Philosophy. Atascadero, CA: Ridgeview, 1977.

Philosophical Perspectives: Metaphysics and Epistemology. Atascadero, CA: Ridgeview, 1977.

Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars, edited by J. F. Sicha. Atascadero, CA: Ridgeview, 1980.

Science, Perception, and Reality. Atascadero, CA: Ridgeview, 1991.

Naturalism and Ontology. Atascadero, CA: Ridgeview, 1992.

Science and Metaphysics. Atascadero, CA: Ridgeview, 1992.

WORKS ON SELLARS

Delaney, C. F., et al. *The Synoptic Vision: Essays on the Philosophy of Wilfrid Sellars*. South Bend, IN: Notre Dame, 1977.

DeVries, Willem A., and Timm Triplett. *Knowledge, Mind, and the Given: A Reading of Sellars' "Empiricism and the Philosophy of Mind."* Indianapolis: Hackett, 2000.

Pitt, Joseph C., ed. *The Philosophy of Wilfrid Sellars*. Boston: Reidel 1978.

Seibt, J. *Properties as Processes: A Synoptic Study of Sellars' Nominalism*. Atascadero, CA: Ridgeview, 1990.

Bruce A. Aune (1996, 2005)

SEMANTICS

Semantics is the study of meaning. More specifically, semantics is concerned with the systematic assignment of meanings to the simple and complex expressions of a language. The best way to understand the field of semantics is to appreciate its development through the twentieth century. In what follows, that development is described. As will be seen, advances in semantics have been intimately tied to developments in logic and philosophical logic.

Though there were certainly important theories, or proto-theories, of the meanings of linguistics expressions prior to the seminal work of the mathematician and philosopher Gottlob Frege, in explaining what semantics is it is reasonable to begin with Frege's mature work. For Frege's work so altered the way language, meaning and logic are thought about that it is only a slight exaggeration to say that work prior to Frege has been rendered more or less irrelevant to how these things are currently understood.

In his pioneering work in logic *Begriffschrift, eine der arithmetischen nachgebildete Formalsprache des reinen Denkens*, which was published in 1879, Frege literally revolutionized the field. It is well beyond the scope of the present entry to describe Frege's achievements in this work. But it should be said that one of his most important contributions was to achieve for the first time a clear understanding of the semantic functioning of expressions of generality, such as 'every,' 'some' and so on. This made it possible to understand, again for the first time, how sentences containing multiple expressions of generality, such as 'Every skier loves some mountain,' manage to mean what they do. In a series of papers written in the late 1800s, Frege articulated a novel theory of meaning for languages that was to be very influential. These papers included "Function and Concept" (1891), "On Concept and Object" (1892) and most famously "On Sense and Reference" (1892).

Frege made a fundamental distinction between expressions that are *unsaturated* or *incomplete* and expressions that are *complete*. The former he called *concept words* (perhaps *concept expressions* would be better) and the latter he called *proper names*. A sentence like:

1. Frege runs.

can be split up into the part that is unsaturated, the concept word 'runs,' and the complete part, the proper name 'Frege.' All expressions, Frege thought, are associated with a *sense* and a *reference*. These both have some claim to be called *the meaning* of the expression in question, and so it is probably best to think of Frege as claiming that there are two components to the meaning of an expression. The referent of an expression can be thought of as the thing in the world the expression stands for. Thus, the referent of the proper name 'Frege' is Frege himself. And the referent of the concept word 'runs' is a *concept*, which Frege took to be a function from an object to a truth value. So the concept 'runs' refers to maps an object o to the truth value true iff o runs. Otherwise, it maps the object to false. By contrast the sense of an expression Frege thought of as a way or mode in which the referent of the expression is presented. So perhaps Frege can be "presented" as the author of *Begriffschrift*. Then the sense of the name 'Frege' is the descriptive condition *the author of Begriffschrift*. It is perhaps more difficult to think of senses of concept words, but it helps to think of them as descriptive conditions that present the concept that is the referent in a certain way.

Now Frege thought that the sense of an expression determines its referent. So the sense of 'Frege' is a mode of presentation of Frege, a descriptive condition that Frege uniquely satisfies in virtue of which he is the referent of 'Frege.' Further, in understanding a linguistic expression, a competent speaker grasps its sense and realizes that it is the sense of the expression.

Of course complex linguistic expressions, such as 1 above, also have senses and references. Frege held that the sense of a complex expression is determined by the senses of its parts and how those parts are combined. (Principles of this general sort are called *principles of compositionality*, and so it could be said that Frege held a principle of compositionality for senses.) Indeed, Frege seems to have held the stronger view that the sense of a complex expression is literally built out of the senses of its parts. In the case of 1, its sense is the result of combining the sense of

'runs' and of 'Frege.' Frege believed that just as the expression 'runs' is unsaturated, so its sense too must be unsaturated or in need of completion. The sense of 'Frege,' by contrast, like the expression itself, is whole and complete (not in need of saturation). The sense of 1 is the result of the whole sense of 'Frege' saturating or completing the incomplete/unsaturated sense of 'runs.' It is the unsaturated sense of 'runs' that holds the sense of 1 together, and this is true generally for Frege. Frege called the sense of a declarative sentence like 1 a *thought*. Thus in "On Concept and Object" (p. 193) Frege writes:

> For not all the parts of a thought can be complete; at least one must be unsatured or predicative; otherwise they would not hold together.

Similarly, Frege held that the reference of a complex expression is determined by the references of its parts and how they are put together (i.e. he held a principle of compositionality for referents). In the case of 1, the referent is determined by taking the object that is the referent of 'Frege' and making it the argument of the function that 'runs' refers to. This function maps objects to the True or the False depending on whether they run or not. Thus, the result of making this object the argument of this function is either the True or the False. And whichever of these is the result of making the object the argument of the function is the referent of 1. So sentences have thoughts as senses and truth values (the True; the False) as referents.

Concerning Frege's account of sentences containing *quantifiers*, expressions of generality such as 'every,' 'some' etc., consider the sentence

2. Every student runs.

The words 'student' and 'runs' are both concept words. Thus they have unsaturated senses and refer to concepts: functions from object to truth values. Now Frege thought that a word like 'every' was doubly unsaturated. To form a whole/complete expression from it, it needs to be supplemented with two concept words ('student' and 'runs' in 2). The sense of 'every' is also doubly unsaturated. Thus the sense of 2 is a thought, a complete sense, that is the result of the senses of 'student' and 'runs' both saturating the doubly unsaturated sense of 'every' (in a certain order). By contrast, the referent of 'every' must be something that takes two concepts (those referred to by 'student' and 'runs' in 2) and yields a referent for the sentence. But as we have seen, a sentence's referent is a truth value. Thus the referent of 'every' must take two concepts and return a truth value. That is, its referent is a function from a pair of concepts to a truth value. In essence, 'every' refers to a function that maps the concepts A and B (in that order) to the True iff every object that A maps to the true, B maps to the true (i.e. iff every object that falls under A falls under B).

Above it was mentioned that Frege thought that the referent of a complex expression was a function of the referents of its parts and how they are combined (compositionality of reference). Some examples seem to show that this is incorrect. Consider the following:

3. Chris believes that snow is white.

3a. Chris believes that Mt. Whitney is more than 14,000 feet high.

These sentences may well have different referents, that is, truth values. But the embedded sentences ('snow is white'; 'Mt. Whitney is more than 14,000 feet high') have the same referents (the True) and the other parts of the sentences have the same referents as well. But then it would seem that compositionality of reference would require that 3 and 3a have the same reference/truth value. Frege famously gets out of this apparent problem by claiming that 'believes' has the effect of shifting the referents of expressions embedded with respect to it. In 3 and 3a, the shifted referents of the embedded sentences are their usual senses. So in these environments, the sentences have different referents because they express different thoughts outside of contexts involving 'believes' and related devices.

Frege's doctrine of sense and reference constitutes a semantical theory of languages, because it claims that the meanings of linguistic expressions have these two components, and it gives an account of what the senses and referents of different kinds of linguistic expressions are.

Shortly after Frege had worked out his semantical theory of sense and reference, the English philosopher and mathematician Bertrand Russell was working out a theory of the meanings, or information contents of sentences. While Frege had held that the thought expressed by a sentence, which captures the information the sentence encodes, consisted of senses, Russell (1903) held that the information encoded by sentences were *propositions*, where the constituents of propositions, far from being Fregean senses, where roughly (and for the most part) the things the propositions is about. Thus, whereas Frege held that 1 expressed a thought containing a mode of presentation of Frege and a mode of presentation of the concept of running, Russell held that the proposition expressed by 1 contained Frege himself and the concept of running (though Russell thought of concepts differently from the way Frege did). This contrast has more

than historical significance, because current semanticists are classified as Fregean or Russellian depending on whether they hold that the information contents of sentences contain the things those information contents are about (objects, properties and relations—Russellian) or modes of presentation of the things those information contents are about (Fregean).

In the early part of the twentieth century, the philosophical movement known as *Logical Positivism* achieved dominance, especially among logically minded philosophers who might have been interested in semantics. The Positivists thought that much of traditional philosophy was literally nonsense. They applied the (pejorative) term "*metaphysics*" to what they viewed as such philosophical nonsense. The Positivists, and especially Rudolf Carnap, developed accounts of meaning according to which much of what had been written by philosophers was literally meaningless. The earliest and crudest Positivist account of meaning was formulated by Carnap (1932). On this view, the meaning of a word was given by first specifying the simplest sentence in which it could occur (its *elementary sentence*). Next, it must be stated how the word's elementary sentence could be verified. Any word not satisfying these two conditions was therefore meaningless. Carnap held that many words used in traditional philosophy failed to meet these conditions and so were meaningless.

Carnap called philosophical statements (sentences) that on analysis fail to be meaningful *pseudo-statements*. Some philosophical statements are pseudo-statements, according to Carnap, because they contain meaningless terms as just described. But Carnap thought that there is another class of philosophical pseudo-statements. These are statements that are literally not well formed (Carnap gives Heidegger's "We know the nothing." as an example).

The downfall of the Positivist's theory of meaning was that it appeared to rule out certain scientifically important statements as meaningless. This was unacceptable to the Positivists themselves, who were self consciously very scientifically minded. Carnap heroically altered and refined the Positivists account of meaningfulness, but difficulties remained. Hempel (1950) is a good source for these developments.

At about the same time Carnap was formulating the Positivists' account of meaning, the Polish logician Alfred Tarski was involved in investigations that would change forever both logic and semantics. It had long been thought that meaning and truth were somehow intimately connected. Indeed, some remarks of Wittgenstein's in his *Tractatus Logico-Philosophicus* ("4.024. To understand a proposition means to know what is the case, if it is true.") had led many to believe that the meaning of a sentence was given by the conditions under which it would be true and false. However, the Positivists had been wary of the notion of truth. It seemed to them a dangerously metaphysical notion, (which is why they "replaced" talk of *truth* with talk of *being verified*).

Against this background, Tarski showed that truth ('true sentence') could be rigorously defined for a variety of formal languages (languages, growing out of Frege's work in logic, explicitly formulated for the purpose of pursuing research in logic or to be used to precisely express mathematical or scientific theories). Though earlier papers in Polish and German contained the essential ideas, it was Tarski (1935) that alerted the philosophical world to Tarski's important new results.

Tarski himself despaired of giving a definition of *true sentence of English* (or any other naturally occurring language). He thought that the fact that such languages contain the means for talking about expressions of that very language and their semantic features (so English contains expressions like 'true sentence,' 'denotes,' 'names,' etc.) meant that paradoxes, such as the paradox of the liar, are formulable in such languages. In turn, Tarski thought that this meant that such languages were logically inconsistent and hence that there could be no correct definition of 'true sentence' for such languages.

Nonetheless, Tarski's work made the notion of truth once again philosophically and scientifically respectable. And it introduced the idea that an important element, perhaps the sole element, in providing a semantics for a language was to provide a rigorous assignment to sentences of the language the conditions under which they are true. (Tarski's 1935 paper for the most part gave definitions of *true sentence* for languages with fixed interpretations. The now more familiar notion of *true sentence with respect to a model* was introduced later. See Hodges [2001] for details.)

Carnap's *Meaning and Necessity* (1947) is arguably the first work that contemporary semanticists would recognize as a work in what is now considered to be semantics. Following Tarski, Carnap distinguishes the languages under study and for which he gives a semantics, *object languages*, from the languages in which the semantics for the object languages are stated, *metalanguages*. The object languages Carnap primarily considers are a standard first order language (S_1), the result from adding 'N' ("a sign for logical necessity") to that language (S_2), and ordinary English. Carnap does not give detailed descriptions of any of these languages, noting that the book

> "… is intended not so much to carry out exact analyses of exactly constructed systems as to state informally some considerations aimed at the discovery of concepts and methods suitable for semantical analysis" (p. 8).

The heart of Carnap's semantics for these languages is given by *rules of designation* for predicates and individual constants, *rules of truth* for sentences and *rules of ranges* for sentences. The rules of designation state the meanings of the predicates and individuals constants using English as the metalanguage. So we have (p. 4):

> 's' is a symbolic translation of 'Walter Scott'
>
> 'Bx'—'x is a biped'

The rules of truth simply provide a Tarski style definition of truth for sentences of the language, (the definition assumes fixed meanings given by the rules of designation for predicates and individual constants). In order to specify the rules of range, Carnap introduces the notion of a *state-description*. For a language, say S_1, a state description in S_1 is a set that contains for every atomic sentence of S_1, either it or its negation, but not both; and it contains no other sentences. Carnap comments (p. 9):

> … it [a state-description in S_1] obviously gives a complete description of a possible state of the universe of individuals with respect to all properties and relation S expressed by predicates of the system. Thus the state-descriptions represent Leibniz' possible worlds or Wittgenstein's possible states of affairs.

Next Carnap gives a recursive characterization of *a sentence holding in a state-description*. An atomic sentences holds in a state-description iff it is a member of it. A disjunctions holds in it iff one of its disjuncts holds in it, etc. The characterization of *holding in a state description* is designed to formally capture the intuitive idea of the sentence being true if the possible world represented by the state-description obtained (i.e. if all the sentences belonging to the state-description were true). Given a sentence S, Carnap calls the class of state-descriptions in which S holds its *range*. Thus the clauses in the characterization of *holding in a state-description* Carnap calls *rules of ranges*. Regarding these rules of ranges, Carnap writes (p. 9–10):

> By determining the ranges, they give, together with the rules of designation for the predicates and the individual constants …, an *interpretation* for all sentences of S_1, since to know the meaning of a sentence is to know in which of the possible cases it would be true and in which not, as Wittgenstein has pointed out.

Thus, Carnap regards the rules of ranges together with the rules of designation as giving the meaning of the sentences of S_1 (the connection with truth and the rules of truth is that there is one state-description that describes the actual world, and a sentence is true iff it holds in that state-description).

Using these resources, Carnap defines his well known *L concepts*. We here concentrate on *L-truth* and *L-equivalence*. Before getting to that, we must say something about Carnap's notion of *explication*. Carnap believed that one of the main tasks for philosophers was to take a "vague or not quite exact" concept, and replace it by a more exact concept that one had clearly characterized. This new concept, called by Carnap the *explicatum* of the old concept, was intended to be used to do the work the old concept was used to do. Carnap thought that the notion of *L-truth* was the explicatum of the vague notions of "logical or necessary or analytic truth" (p. 10).

A sentence is L-true in a semantical system (e.g. S_1) iff it holds in every state description in that system. Carnap regarded this as a precise characterization of Leibniz's idea that necessary or analytic or logical truths hold in all possible worlds. Next, Carnap defines the notion of *L-equivalence* for sentences, predicates and individual constants. Effectively, two names, predicates or sentences are L-equivalent (in a semantical system—e.g. S_1) iff they have the same extension at every state-description in that system, (so L-equivalent names must name the same individual at every state description, L-equivalent predicates must be true of the same individuals at every state description, etc.).

The importance of Carnap's notion of *L-equivalence* is that he uses it to sketch a semantics for belief ascriptions. In order to do this, Carnap extends his notion of *L-equivalence* in several ways. First, he extends it so that expressions of different "semantical systems" (roughly, formal languages) may be L-equivalent (in effect, expressions e of system 1 and e' of system 2 are L-equivalent just in case the semantical rules of the two systems together suffice to show that the expressions have the same extension, p. 57). Second, he extends the notion of *L-equivalence* to apply to sentential connectives, variables (they are L-equivalent iff they have the same range of values) and to quantifiers (they are L-equivalent iff they are quantifiers of the same sort [universal, existential] whose variables are L-equivalent, p. 58). Third, he defines what it is for two expressions of the same or different semantical systems (again, roughly formal languages) to be *intensionally isomorphic*. Roughly, expressions are intensionally isomorphic just in case they are built up in the same

way out L-equivalent parts. With these tools in hand, Carnap writes (p. 61–62):

> It seems that the sentence 'John believes that D' in *S* [a fragment of English—see p. 53] can be interpreted by the following semantical sentence:
>
> 15-1. 'There is a sentence $\mathfrak{S}_i$in the semantical system S' such that (a) $\mathfrak{S}_i$ is intensionally isomorphic to 'D' and (b) John is disposed to an affirmative response to $\mathfrak{S}_i$

Though Carnap's semantics for belief ascriptions was criticized by Alonzo Church (1950), many philosophers were influenced by Carnap's idea that the objects of belief are structured entities built up in the same way out of entities with the same intensions. See, for example, Lewis (1970).

The final important feature of *Meaning and Necessity* was its semantic treatment of modality. Carnap begins his discussion of modality by mentioning the work of C. I. Lewis (presumably he had in mind especially Lewis and Langford [1932]) in constructing various systems of modal logic. As mentioned above, Carnap considered as an object of semantical investigation a language that was the first order predicate logic (S_1) supplemented with the sign 'N' "for logical necessity." He called the resulting language S_2. Syntactically, prefixing 'N' to a matrix (either a sentence or a formula with free variables) results in a matrix. A detailed discussion of Carnap's semantics for this modal language would go beyond the scope of the present entry. However, a couple points are worth making. First, if we just consider the case in which 'N' fronts a *sentence* (formula with no free variables) ϕ, to get the rules of range for S_2 we would simply add to the rules of range of S_1 the following:

> N(ϕ) holds in every state-description if ϕ holds in every state description; otherwise N(ϕ) holds in no state-description.

This is a consequence of Carnap's idea that 'N' is the sign for logical necessity, and the notion of L-truth is the explicatum of the vague notion of logical necessity. Thus a sentence fronted by 'N' should hold at a state description iff the sentence it embeds holds at every state-description. But then if the sentence fronted by 'N' holds at a state-description, it holds at every state-description. Thus, the above.

But of course since 'N"could front a matrix with free variables, one could then attach a quantifier to the result. Letting '..u..' be a matrix containing the variable 'u' free, we get things like

> (u)N(..u..)

That is, we get quantifying into the sign 'N' for logical necessity. However, Carnap's treatment here results in the above being equivalent to (indeed, L-equivalent to)

> N(u)(..u..).

The important point, however, is that Carnap had sketched a semantics for quantified modal logic.

Though virtually all of the crucial analyses and explications in *Meaning and Necessity* were eventually significantly modified or rejected (the explication of "logical necessity" by the notion of L-truth, understood in terms of holding at all state-descriptions; the treatment of 'N,' the sign of "logical necessity"; and the semantics for belief ascriptions), the work was nonetheless very important in the development of semantics. It provided a glimpse of how to use techniques from logic to systematically assign semantic values to sentences of languages, and began the project of providing a rigorous semantics for recalcitrant constructions like sentences containing modal elements and verbs of propositional attitude.

In the 1950s and early 1960s Carnap's ideas on the semantic treatment of modal logic were refined and improved upon. The result was the now familiar "Kripke style" semantics for modal logic. Kripke's formulations will be discussed here, but it is important to understand that similar ideas were in the air (see Hintikka [1961], Kanger [1957], and Montague [1960a]). Though these works were in the first instance works in logic, as we will see, they had a profound effect on people who were beginning to think about formal semantics for natural languages.

We will concern ourselves with the specific formulations in Kripke (1963). What follows will be of necessity slightly technical. The reader who is not interested in such things can skip to the end of the technical discussion for informal remarks. Assume that we have a standard first order logic with sentential connectives ~,& and □ (the first and third one-place, the second two-place), individual variables (with or without subscripts) x,y,z, … ; n-place predicates P^n, Q^n, … (0 place predicate letters are *propositional variables*), and universal quantifier (for any variable x_i, (x_i)). A *model structure* is a triple ⟨G, K, *R*⟩, where K is a set, G ε K and *R* is a *reflexive relation* on K (i.e. for all H ε K, H *R* H). Intuitively, G is the "actual world" and the members of K are all the possible worlds. *R* is a relation between worlds and is usually now called the *accessibility relation*. Intuitively, if H*R* H' (H' is accessible from H), then what is *true* in H' is *possible* in H.

Again intuitively, the worlds accessible from a given world are those that are possible relative to it.

Putting conditions on *R* gives one model structures appropriate to different modal logics. If *R* is merely reflexive, as required above, we get an *M* model structure. If *R* is reflexive and *transitive* (i.e. for any H, H', H'' ε K, if H *R* H' and H' *R* H'', then H *R* H''), we get an *S4* model structure. Finally, if *R* is reflexive, transitive and *symmetric* (i.e. for any H, H' εK, if H *R* H', then H' *R* H), we get an *S5* model structure. (It should be recalled that for Carnap, state-descriptions, which represented possible worlds, were each accessible for every other—in effect because there was no accessibility relation between state-descriptions; thus translated into the present framework Carnap's "models" would be S5 models. Also, in Kripke's semantics, possible worlds (members of K) are primitive; in Carnap's, of course, they are explicated as state descriptions.) A *quantificational model structure* is a model structure ⟨G, K, *R*⟩ together with a function ψ that assigns to every H in K a set of individuals: the *domain* of H. Intuitively this is the set of individuals existing in the possible world H. Of course, this allows different worlds (members of K) to have different domains of individuals. This formally captures the intuitive idea that some individuals that exist might not have, and that there might have been individuals that there aren't.

Given a quantificational model structure, consider the set *U* which is the union of ψ(H) for all H in K. Intuitively, this is the set of all possible individuals (i.e. the set U of individuals such that any individual in the domain of any world is in U). Then U^n is the set of all n-tuples whose elements are in U. A *quantificational model* on a quantificational model structure ⟨G, K, *R*⟩ is a function φ that maps a zero-place predicate and a member of K to T or F; and for n>0, an n-place predicate and a member of K to a subset of U^n. We extend φ by induction to assign truth values to all formula/world pairs *relative to a function assigning members of U to variables*:

1. *Propositional Variable*: Let f be a function assigning elements of U to all individual variables. Let P be a propositional variable. Then for any H in K, φ(P, H)=T relative to f iff φ(P, H)=T; otherwise φ(P, H)=F relative to f.

2. *Atomic*: Let f be as in 1. For any H in K, $\varphi(P^n x_1,\ldots,x_n, H)$=T relative to f iff $\langle f(x_1), \ldots ,f(x_n)\rangle$ ε $\varphi(P^n, H)$; otherwise $\varphi(P^n x_1, \ldots ,x_n, H)$=F relative to f.

(Note that 2 allows that an atomic formula can have a truth value at a world relative to an assignment to its variables, where some or all of its variables get assigned things not in the domain of the world, since f assigns elements of *U* to free variables; and φ assigns subsets of U^n to P^n!)

3. *Truth functional connectives*: Let f be as in 1. Let *A* and *B* be formulae. For any H in K, φ(*A*&*B*, H)=T relative to f iff φ(*A*, H)=T relative to f and φ(*B*, H)=T relative to f; otherwise φ(*A*&*B*, H)=F relative to f. (Similarly for ~)

4. *Modal operator*: Let f be as in 1. Let *A* be a formula. φ(□*A*, H)=T relative to f iff φ (*A*, H')=T relative to f for all H' ε K such that H *R* H'; otherwise φ(□*A*, H)=F relative to f.

(Note that according to 4, whether a formula □A is true at a world (relative to f) depends only on whether A is true at all worlds *accessible* from the original world.)

5. *Quantifier*: Let f be as in 1. Let $A(x, y_1, \ldots y_n)$ be a formula containing only the free variables x, $y_1,\ldots,y_n$. For any H in K, and any function g (assigning elements of U to free variables), suppose $\varphi(A(x, y_1, \ldots ,y_n), H)$ relative to g is defined. Then $\varphi((x)\, A(x, y_1, \ldots y_n), H)$=T relative to f iff for every f' such that f'(x) ε ψ (H) and f' differs from f at most in that f'(x) is not f(x), $\varphi(A(x, y_1, \ldots y_n), H)$ =T relative to f'; otherwise, $\varphi((x)\, A(x, y_1, \ldots y_n), H)$ =F relative to f.

(As Kripke notes, that in 5 we consider only functions f' such that f'(x) ε ψ(H) means that quantifiers range over only the objects that exist at the world where the quantified sentence is being evaluated.)

Now having gone through Kripke's semantics for quantified modal logic in some detail, let us step back and ask why it was important in terms of thinking of the semantics of natural language. People like Richard Montague, who we will discuss below, were clearly influenced in their thinking about the semantics of natural language by Kripke's semantics for modal logic, (recall too that Montague [1960a] itself contained ideas related to Kripke's). Since at least Carnap's *Meaning and Necessity* (and perhaps before), philosophers had thought of sentences as semantically associated with *propositions* and of n-place predicates as semantically associated with n-place *relations* (properties being one-place relations). Further, they had thought of these propositions and relations as determining truth values and extensions for the sentences and predicates expressing them relative to a "possible world" (which, of course, Carnap represented by a state description).

Now in Montague (1960b), it is suggested that an n-place relation just is a function from possible worlds to a set of n-tuples (intuitively, the set of n-tuples whose ele-

ments stand in the relation in question from the stand point of the world in question); and that a proposition just is a function from possible worlds to truth values. Generalizing these ideas leads straightforwardly to possible worlds semantics for natural languages discussed below. Further, Montague claims this way of understanding relations and propositions (which Montague calls *predicates*, one-place predicates, then, are properties; and zero-place predicates are propositions) is to be found for the first time in Kripke (1963). This, in turn, means that at least Montague saw the seeds of possible worlds semantics for natural languages in Kripke (1963).

This initially seems at least a little bit strange, since nowhere in Kripke (1963) does one find the identification of propositions with functions from possible worlds to truth values or relations with functions from possible worlds to sets of n-tuples. However, it is easy to see why a logician like Montague would see those ideas in Kripke (1963). Consider again a model on a quantificational model structure, forgetting for the moment about functions f that are assignments to free variables and that the domains of members of K can vary, (essentially, this means we are considering a model on a *propositional* model structure). A *model* φ on a (M/S4/S5) model structure ⟨G, K, *R*⟩ assigns to a propositional variable (a zero-place predicate—an atomic formula without any variables) and a member of K either T or F. Now consider a particular propositional variable P. Consider the function f_P defined as follows:

$$\text{For any H in K, } f_P(H) = T \text{ iff } \varphi(P, H)=T\text{; otherwise } f_P(H) = F$$

f_P is a function from worlds to truth values and so can be thought of *a la* Montague as the proposition expressed by P (in the model φ on the model structure ⟨G, K, *R*⟩)! That is, propositions, understood as functions from worlds to truth values, are trivially definable using Kripke's models. Similar remarks apply to n-place relations, understood as functions from possible worlds to sets of n-tuples of individuals. It seems likely that this is why a logician like Montague would take Kripke to have introduced them. Montague, after making the attribution to Kripke, does add (p.154): "… Kripke employs, however, a different terminology and has in mind somewhat different objectives."

These functions from worlds to truth values or sets of n-tuples are now generally called *intensions*. Their values at a world (truth values; sets of n-tuples) are generally called *extensions* (at worlds). The idea that the primary job of semantics is to assign to expressions of natural languages intensions and extensions of the appropriate sort very much took hold in the wake of work by Kripke and others in the semantics of modal logic.

With the resources Kripke and others had made available in hand, researchers thinking about the semantics of natural languages eagerly made use of them. Thus the late 1960s and early 1970s saw dizzying progress in natural language semantics as the techniques for modal logic were applied. Two works from that era that particularly capture the spirit of the times are Lewis (1970) and Montague (1973). The latter will be discussed here, since it is probably the most sophisticated and influential of the works of that period. The particular semantic phenomena Montague was concerned to understand were the workings of verbs of propositional attitudes like 'believes,' the workings of intensional verbs like 'worships' and related phenomena (see p. 248 where Montague lists some of his concerns).

We saw above that both Frege and Carnap were also concerned with understanding the semantics of verbs like 'believes.' We are now in a position to say more about why such expressions attract the attention of semanticists. Consider the expression 'It is not the case' in sentences like

4. It is not the case that snow is white.

4a. It is not the case that Mt. Whitney is more than 14,000 feet high.

Whether a sentence fronted by 'It is not the case' is true or false depends only on the extension/truth value of the embedded sentence. Since both the embedded sentences are true, 4 and 4a are both false. Let's put this by saying that 'It is not the case that' creates *extensional contexts*. As we saw above, 'believes' doesn't create extensional contexts. 3 and 3a can differ in truth value even though the embedded sentences are both true. Let's say that 'believes' creates *nonextensional contexts*. The same is true of 'Necessarily.' The following differ in truth value even though the embedded sentences have the same extensions/truth values:

5. Necessarily, everything is identical to itself.

5a. Necessarily, Aristotle is a philosopher.

Finally, intensional verbs like 'worship' exhibit similar behavior and we could extend our characterization of creating nonextensional contexts so as to include such verbs. For even though 'Samuel Clemens' and 'Mark Twain' have the same extension (a certain individual), the following two sentences apparently may differ in extension/truth value:

6. Lori worships Samuel Clemens.

6a. Lori worships Mark Twain.

Now semanticists have been puzzled as to how to think of the semantics of expressions that create nonextensional contexts. But the work of Carnap and Kripke suggested the way to understand 'Necessarily.' In particular,

> Necessarily S is true at a world w just in case the intension of S maps every world (accessible from w) to true.

In other words, whereas 'It is not the case' looks at the extension of the sentence it embeds to determine whether the entire sentence containing it is true, 'Necessarily' looks at the intension of the sentence it embeds to determine whether the entire sentence containing it is true. And given Kripke's semantics, intensions were well defined, respectable entities: functions from worlds to extensions. This made it appear to many that a semantics that assigned intensions to expressions could treat all expressions creating nonextensional contexts. Certainly, Montague had a version of this view.

As indicated above, Montague (1973) wanted to provide semantic treatments of verbs of propositional attitude such as 'believes,' intensional verbs such as 'worships,' and other phenomena. We will concentrate on these phenomena as well as Montague's treatment of quantification. Montague (1973) provides a syntax for a fragment of English. The fragment includes common nouns ('woman'; 'unicorn'), intransitive verbs (including 'run' and 'rise'), transitive verbs, (including both intensional transitives and "normal" transitive verbs like 'love'), ordinary names and pronouns, adverbs (including 'rapidly' and 'allegedly'), prepositions, verbs of propositional attitude and modal sentence adverbs ("adsentences"—'necessarily'). The fragment allows the formation of relative clauses (though they employ the somewhat stilted 'such that,' so that we get things like 'man such that he loves Mary') and so complex noun phrases, as well as prepositional phrases and quantifier phrases ('Every woman such that she loves John'). Thus, Montague's syntactic fragment includes sentences like:

7. Every man loves a woman such that she loves him.

8. John seeks a unicorn.

9. John talks about a unicorn.

10. Mary believes that John finds a unicorn.

11. Mary believes that John finds a unicorn and he eats it.

It should be noted that many sentences of Montague's fragment had non-trivially different syntactic analyses: that is, distinct syntactic analyses that are interpreted differently semantically. So, for example, 8 above has an analysis on which 'a unicorn' is the constituent last added to the sentence and an analysis on which 'John' is the last constituent added. The latter has an interpretation on which it may be true even if there are no unicorns and so John is seeking no particular one. The former requires John to be seeking a particular unicorn. Thus, it is really syntactic analyses of sentences, and not the sentences themselves, that get semantic interpretations.

The next aspect of Montague's semantic treatment of his fragment of English is his intensional logic. Montague's intensional logic is typed. In particular, e and t are the basic types; and whenever a and b are types, $\langle a,b \rangle$ is a type. Finally, for any type a, $\langle s,a \rangle$ is a type. For each type, there will be both constants and variables of that type (and hence quantifiers of that type). The key to understanding the syntactic interactions of the expressions of various types is to know that if α is of type $\langle a,b \rangle$ and β is of type a, then $\alpha(\beta)$ is of type b. Interpretations assign expressions of the logic various denotations (relative to an assignment of values to variables). Expressions of type e get assigned individuals (possible individuals); expressions of type t get assigned truth values. Expressions of type $\langle a,b \rangle$ get assigned as denotations functions from denotations of type a to denotations of type b. Finally, expressions of type $\langle s,a \rangle$ get assigned functions from a world/time pair to a denotation of type a ("an intension of a type a expression"). To take some examples, expressions of type $\langle e,t \rangle$ get assigned functions from individuals to truth values (the denotations can alternatively be thought of as sets of individuals: those that get assigned to true). Expressions of type $\langle s,e \rangle$ are assigned functions from world/time pairs to individuals. Such functions Montague called *individual concepts*. Expressions of type $\langle\langle s,e \rangle,t \rangle$ are assigned functions from individual concepts to truth values (alternatively, sets of individual concepts). Expressions of type $\langle s,t \rangle$ are assigned functions from world/time pairs to truth values. As indicated above, Montague thought of these as propositions.

The way Montague provided a semantic interpretation of his syntactic fragment of English was to provide an algorithm for translating English sentences (really, syntactic analyses of English sentences) into his intensional logic. Then the interpretation of the English sentences was given by the interpretation of its translation in intensional logic. Recall again that sentences like 8 above can be true even if there are no unicorns. Thus, a verb like

'seeks' could not have as its denotation (really, its translation into intensional logic could not have as its denotation) a relation between individuals (or a function from individuals to a function from individuals to truth values).

In order to get the proper results, Montague decided to assign to common nouns and intransitive verbs as their denotations sets of individual concepts rather than sets of individuals. Verbs like 'believes' have as their denotations functions from propositions to sets of individual concepts. Since individual concepts essentially function as individuals in Montague's semantics (recall that common nouns like 'man' have as denotations sets of individual concepts), this treatment essentially amounts to holding that verbs of propositional attitude denote relations between individuals and propositions. Quantifiers such as 'Every man' denote sets of properties of individual concepts (functions from world/time pairs to sets of individual concepts). Roughly, 'Every man walks' is true at a world and time ⟨w,t⟩ just in case the property of individual concepts that determines the correct set of individual concepts denoted by 'walks' at every world and time is in the set of properties of individual concepts denoted by 'Every man' at ⟨w,t⟩. 'Necessarily' denotes at a world/time ⟨w,t⟩ a set of propositions: those that are necessary at ⟨w,t⟩.

Finally, a transitive verb denotes a function from properties of properties of individual concepts (denotations of expressions of type ⟨s,⟨⟨s,⟨⟨s,e⟩,t⟩⟩,t⟩⟩—functions from world/time pairs to sets of properties of individual concepts) to sets of individual concepts. Again, recalling that individual concepts essentially stand in for individuals in Montague's framework, this means that transitive verbs in effect denote relations between individuals and properties of properties of individuals. Note that this means that for 8 to be true at a world/time pair ⟨w,t⟩ is for John to stand in a relation to the property of being a property possessed by a unicorn. This can be the case even if there are no unicorns.

Montague chose to treat all expressions of a given syntactic category the same way semantically. This means that transitive verbs like 'loves' get the odd denotation required by 'seeks' to get 8 right. But don't we want 'John loves Mary' to be true at world/time pair iff the individual John stands in a relation to the individual Mary? Surely this shouldn't require instead that John stands in a relation to the property of being a property possessed by Mary. Where's the love (between individuals)? Montague essentially requires interpretations to make true meaning postulates for "ordinary" verbs like 'loves,' and these end up insuring that 'John loves Mary' is true at ⟨w,t⟩ iff John and Mary themselves are properly related.

Montague's semantic account here was very influential. He showed that the resources Kripke and others developed for the semantics of modal logic could be rigorously applied to natural languages, and arguably treat such recalcitrant expressions as 'believes,' 'necessarily,' and 'seeks.' Montague's basic approach was picked up by many philosophers and linguists and much work in semantics through the 1980s and beyond was conducted in this framework. Indeed, much work is still done in this and closely related frameworks.

At about the same time Montague was doing his pioneering work on formal semantics for natural languages, Donald Davidson was developing a very different approach to semantics. Davidson (1967) begins with the idea that a theory of meaning for a natural language must specify how the meaning of a sentence is determined by the meanings of the words in it, and presumably how they are combined (in other writings, Davidson puts the point by saying that the meaning of sentence must be a function of a finite number of features of the sentence—presumably, one is its syntax). Davidson thought that only a theory of this sort could provide an explanation of the fact that on the basis of mastering a finite vocabulary and a finite number of syntactic rules, we are able to understand a potentially infinite number of sentences. More specifically, Davidson thought a theory of meaning should comprise an axiomatized theory, with a finite number of axioms, that entails as theorems (an infinite number of) statements specifying the meaning of each sentence of the language. Davidson thought that grasping such a theory would allow one to understand all the sentences of the language. Further, as suggested above, such a theory would explain how creatures like us are capable of understanding an infinite number of sentences. It would only require us to grasp the axioms of the theory of meaning, which are finite in number.

It might be thought that the theorems of a theory of meaning of the sort discussed would be all true sentences of the form 's means that p,' where 's' is replaced by a structural description of a sentence of the language and 'm' is replaced by a term referring to a meaning. Further, it might be thought that a theory would have such theorems in part by assigning meanings to the basic expressions of the language (such assignments being made by axioms). However, Davidson thinks that we have not a clue as to how to construct such a theory, mainly because we have no idea how the alleged meanings of simpler expressions combine to yield the meanings of complex

expression of which they are parts. Thus, Davidson concludes, postulating meanings of expressions gets us nowhere in actually giving a theory of meaning for a language.

Davidson's counterproposal as to what a theory of meaning should be like is radical. A theory of meaning must be a finite number of axioms that entail for every sentence of the language a true sentence of the form 's is true iff p,' where 's' is replaced by some sort of description of a sentence whose theory of meaning we are giving, and 'p' is replaced by some sentence. Henceforth, we will call such sentences *T-sentences*. Recalling our discussion of Tarski, the language we are giving a theory of meaning for is the object language and the theory of meaning is given in the metalanguage. Thus, the formulation just given requires the metalanguage to have some sort of (presumably standardized) description of each sentence of the object language (to replace 's'); if we imagine 'p' to be replaced by the very sentence that what replaces 's' describes (as Davidson sometimes supposes) the metalanguage must also contain the sentences of the object language. In short, Davidson held that to give a theory of meaning for a language is to give a Tarski-style truth definition for it.

Tarski thought that a condition of adequacy for a theory of truth for a (in his case, formal) language L was that the theory has as consequences all sentences of the form 's is true (in L) iff p', where 's' is replaced by a structural description of a sentence of the object language and 'p' is replaced by a translation of it. Here Tarski clearly seemed to think that for one sentence to translate another is for them to share a meaning. However in characterizing what is to replace 'p' in his T-sentences, Davidson cannot require 'p' to be replaced by a translation of the sentence the thing replacing 's' describes, assuming anyway that for one sentence to be a translation of another is for them to share the same meaning. For Davidson eschews meanings. After all, a theory of truth was supposed to *be* a theory of meaning; it would hardly do, then, to appeal to meanings in constructing one's theory of truth. Thus Davidson famously merely requires the T-sentences to be true. But this requirement is very weak, for 'iff' is truth functional in Davidson's T-sentences, and so the sentences require for their truth only that the two sides share a truth value. But then there is nothing in principle yet to prevent having a theory of truth for English that yields not:

> 12. 'Snow is white' is true (in English) iff snow is white.

but instead

> 13. 'Snow is white' is true (in English) iff grass is green.

After all, 13 is true! Davidson was aware of this consequence of his view, and explicitly discussed it. He claimed that by itself, the fact that a theory of truth yields 13 as a theorem instead of 12 doesn't cut against it. However, the theory has to get all the other T-sentences coming out true, and Davidson thought it was unlikely that it could do that and yield 13 as a theorem.

Of course, the picture sketched so far needs to be complicated to account for contextually sensitive expressions. It won't do to have as theorems of one's truth theory things such as:

> 14. 'I am hungry' is true (in English) if I am hungry.

Davidson himself thought that the way to deal with this was to relativize truth to e.g. a speaker and a time (to handle tense). Others have suggested that a theory of truth for a language containing such contextually sensitive words must define truth for utterances of sentences. For example, see Weinstein (1974).

Further complications are required as well. Natural language contains devices not contained in the relatively austere formal languages for which Tarski showed how to define truth. Natural languages contain verbs of propositional attitude ('believes'), non-indicative sentences and other features. Davidson attempted to provide accounts of many such devices in other papers. Davidson (1968) for example takes up verbs of propositional attitude.

One sometimes hears model theoretic approaches to semantics contrasted with those that offer an absolute truth theory. The contrast is illustrated by comparing Montague and Davidson, since each is perhaps the paradigmatic case of one of these approaches. As we saw, Montague gives a semantics for English sentences by associating them with formulae of intensional logic. He then gives a semantics for the formulae of intensional logic. Now the latter includes a definition of truth relative to an interpretation (and other parameters as well). As discussed, expressions of Montague's intensional logic only have denotations (and intensions) relative to interpretations, which are also sometimes called *models*. Roughly, then, a model theoretic semantics is one that defines truth *relative to models or interpretations*. By contrast, as we have seen, Davidson wants a theory of truth simpliciter (actually, truth for L, but truth isn't relativized to models). Thus, Davidson's approach is sometimes called an absolute truth theory approach. I believe it is

fair to say that most semanticists today use a model theoretic approach.

The 1960s and 1970s saw an explosion in the sort of model theoretic semantics pioneered by Montague, Lewis and others. Some of the important developments had to do with evolving notions of an *index of evaluation*. As we saw above, in Montague's intensional logic, expressions are assigned extensions/denotations at world/time pairs (under an interpretation relative to an assignment of values to variables—this will be suppressed in the present discussion for ease of exposition). In particular, formulae are assigned truth values at a pair of a world and time.

Since expressions of Montague's English fragment receive semantic interpretations by being given the interpretation assigned to the expressions of intensional logic they are translated into, exactly similar remarks apply to English expressions and sentences. We shall call these elements at which expressions are assigned extensions (in this case, world/time pairs) *indices*. (Terminology here varies: Montague called these things *points of reference*; Lewis [1970] called them *indices*, which is probably the most common term for them.) It should be obvious why sentences are assigned truth value at worlds. The reason Montague included times in his indices was that his intensional logic included tense operators in order that he could capture the rudimentary behavior of tense in English. Semantically, such operators work by shifting the time element of the index. Thus, where P is a past tense operator, φ a formula, w a world and t a time, $P\varphi$ is true at $\langle w,t \rangle$ iff φ is true at $\langle w,t' \rangle$ for some t' prior to t. Similarly, modal operators shift the world element of the index: Necessarily φ is true at $\langle w,t \rangle$ iff φ is true at $\langle w',t \rangle$ for all w'.

So the truth values of formulae of Monatgue's intensional logic, and so of the English sentences they translate, depend on (or vary with) both a world and a time. Of course, it was noticed that the truth values of some English sentences vary with other features as well, such as who is speaking (if the sentence contains 'I'); who is being addressed (if the sentence contains 'you'); where the sentences is uttered (if the sentence contains 'here') and so on. A natural thought was to build into indices features for all such expressions, so that indices would contain all the features that go into determining extensions of expression. Thus, indices would be n-tuples of a world, time, place, speaker, addressee and so on. Lewis (1970) is a good example of an "index semantics" with indices containing many features. However, a number of developments resulted in such approaches being abandoned or at least significantly modified.

Hans Kamp (1971) discovered that in a language with standard feature-of-index shifting tense operators and contextually sensitive expressions that are sensitive to that same feature, such as 'now,' one needs two temporal coordinates. The point can be illustrated using a sentence in which 'now' occurs embedded under e.g. a past tense operator (assume 'one week ago' is a past tense operator):

15. One week ago Sarah knew she would be in Dubrovnik now.

When this sentence is evaluated at an index, there must be a time in the index for 'one week ago' to shift. The embedded sentence ('Sarah knew she would be in Dubrovnik now') is then evaluated relative to an index whose time feature has been shifted back one week. But then if 'now' takes that time as its value, we predict that 15 means that one week ago Sarah knew she would be in Dubrovnik *then*. But the sentence doesn't mean that. So the index must contain a second time, in addition to the one shifted by 'one week ago,' that remains unshifted so that the embedded occurrence of 'now' can take it as its value.

Kamp's requirement of there being two time coordinates is sometimes called the requirement of *double indexing*. I emphasize again that the requirement stems from there being in the language an operator that shifts a certain feature (time, in our case) and a contextually sensitive expression that picks up as its value the same feature. The argument above given for double indexing of times, then, assumes that temporal expressions ('One week ago') are index shifting operators. Many, including the present author, doubt this claim. (See King [2003] for discussion.) But similar arguments (involving 'actual' and 'Necessarily') could be given for double indexing of worlds.

At any rate, on the basis of such considerations, it was thought that minimally, one needed two indices, each of which contained (at least) a world and a time. However it was Kaplan (1989) (written in the early 1970s and circulated for years in mimeograph form) that provided the proper theoretical understanding of double indexing. Kaplan forcefully argued that not only do we need two indices for the reasons Kamp suggested as well as others (see section VII of 'Demonstratives'), but we need to recognize that the indices are representing two very different things, with the result that we need to recognize two different kinds of semantic values. One index represents *context of utterance*. This is the index that provides values for contextually sensitive expressions such as 'I,' 'now,' 'here' and so on. The intuitive picture is that a sentence taken relative to a context of utterance has values assigned

to such contextually sensitive expressions. This results in the sentence having a *content, what is said by the sentence*, taken in that context.

So If I utter 'I am hungry now' on June 12, 2006, the content of the sentence in that context, what I said in uttering it then, is that Jeff King is hungry on June 12, 2006. Now that very content can be evaluated at different *circumstances of evaluation*, which are what the other index represents. For simplicity, think of circumstances of evaluation as simply possible worlds. Then we can take the sentence 'I am hungry now' and consider its content relative to the context of utterance described above. That content, or proposition, can then be evaluated for truth or falsity at different circumstances of evaluation (possible worlds). It is true at worlds in which Jeff is hungry on June 12, 2006 and false at those where he is not.

This distinction between context and circumstance, which the two indices represent, gives rise to a distinction between two kinds of semantic value (here we confine ourselves to the semantic values associated with sentences). On the one hand, the sentence 'I am hungry now' has a meaning that is common to utterances of it regardless of speaker or time. It is this meaning that determines what the content of that sentence is taken relative to contexts with different speakers and times. So this meaning, which Kaplan called *character*, determines a function from contexts to propositional content or what is said. By contrast, there is a sense in which the sentence 'I am hungry now' uttered by me now and Rebecca tomorrow means different things. This is because the sentence has different *contents* relative to those two contexts. So content is the other kind of semantic value had by sentences. Contents are true or false at worlds, so contents determine functions from worlds to truth values. In summary, character determines a function from context to content; content determines a function from worlds to truth values. Kaplan's distinction between context and circumstance and the corresponding distinction between character and content has been hugely influential and widely accepted.

Another important feature of Kaplan's (1989) work is his argument that both *demonstratives* (contextually sensitive words whose use requires the speaker to do something like demonstrate (point at) who she is talking about: 'he,' 'she,' 'this,' 'that') and *pure indexicals* (contextually sensitive words that don't require such demonstrations: 'I,' 'today,' etc.) are *devices of direct reference*. If we think of contents of sentences, propositions, as structured entities having as constituents the individuals, properties and relations that are the contents (relative to a context) of the expressions in the sentence, a view Kaplan likes, we can understand the claim that indexicals and demonstratives directly refer as the claim that these expressions contribute to propositions (relative to a context) the individuals they refer to (in the context). Thus, when I say: 'I am hungry,' the indexical 'I' contributes me to the proposition expressed by that sentence in that context.

Historically, the importance of this direct reference account of indexicals and demonstratives is its anti-Fregean thrust. Recall that for Frege, expressions generally, even those that refer to individuals, contribute to propositions senses that pick out their references and not the references themselves. In claiming that indexicals and demonstratives contribute individuals to propositions rather than senses that pick out those individuals, Kaplan was proposing a radically anti-Fregean account of indexicals and demonstratives. Kaplan's arguments here complemented the anti-Fregean arguments of one of the most influential works in philosophy of language of the twentieth century: Saul Kripke's (1980) *Naming and Necessity*.

Among other things, Kripke (1980) provided powerful arguments against what he sometimes calls *the description theory* of names. On the description theory, names are held to be both synonymous with definite descriptions and (more weakly) to have their references fixed by definite descriptions. So, for example, 'Aristotle' might be thought to be synonymous with 'the teacher of Alexander,' and whoever satisfies this description is the referent of 'Aristotle.' Frege's view was thought to be a version of the description theory, since Frege seems to say that the sense of a proper name can be expressed by a definite description (Frege [1892a] note B), in which case the name and descriptions would be synonymous. Kripke argued very compellingly that descriptions were neither synonymous with, nor determined the reference of, proper names. As to synonymy, Kripke pointed out that whereas

16. The teacher of Alexander taught Alexander.

expressed (nearly) a necessary truth,

17. Aristotle taught Alexander.

expresses a highly contingent truth. But if the name and description were synonymous, the two sentences should be synonymous and so both should be contingent or both should be necessary. But they aren't. Indeed, the name and description seem to function very differently semantically. As Kripke famously noted, whether 17 is true at any possible world depends on the properties of Aristotle at that world. This because 'Aristotle' is what Kripke called a *rigid designator*: the expression designates Aristotle at

every world where he exists, and never designates any individual other than Aristotle. Hence evaluating the sentence at a world always requires us to check Aristotle's properties there. By contrast, 'the teacher of Alexander' presumably designates different individuals at different worlds, depending on who taught Alexander there. Thus, this expression is *non-rigid*.

As to descriptions determining the referents of names, Kripke adduced a number of considerations but perhaps the most persuasive was the following. Consider a name and any description that allegedly fixes the referent of the name, say 'the man who proved the completeness of arithmetic' fixes the referent of 'Godel.' If we imagine that in fact some man Schmidt satisfies the description, we do not conclude that he is the referent of 'Godel.' Quite the contrary, we conclude that the referent of 'Godel,' that is, Godel, fails to satisfy the description. But then the description does not fix the referent of the name (i.e. the referent is not whoever satisfies the description).

The arguments of Kaplan (1989) and Kripke (1980), together with arguments given by Donnellan, Marcus, Putnam and others turned semantics in a very anti-Fregean direction from the 1970s on. This anti-Fregean strain as applied to singular terms is sometimes called *the new theory of reference*.

As we saw above, Kaplan claimed that indexicals and demonstratives were directly referential and contributed their referents (relative to a context) to the propositions expressed by sentences in which they occur (interestingly, this is not reflected in Kaplan's [1989] formal system, which makes use of unstructured propositions that have no constituents corresponding to the words in the sentences that express the propositions; but his informal remarks make clear his intent). By contrast, though Kripke (1980) argued against the descriptive theory of names, he cautiously made no positive claims about what names contribute to propositions (the preface to Kripke [1980] makes clear that this caution was intended—see pp. 20–21). In a series of works in the 1980s, most famously Salmon (1986) and Soames (1987), Scott Soames and Nathan Salmon offered powerful arguments in favor of the view that names too were devices of direct reference and contributed only their bearers to propositions expressed by sentences in which they occur. Both Soames and Salmon defended the view that sentences (relative to contexts) express structured propositions, with names (and indexicals and demonstratives) contributing the individuals to which they refer to propositions. Salmon and Soames both also thought that attitude ascriptions such as the following:

18. Nathan believes that Mark Twain is an author.

assert that the subject (Nathan) stands in a certain relation (expressed by 'believes') to a structured proposition (expressed by the embedded sentence). If that is right and if names contribute only individuals to propositions expressed by sentences in which they occur, then (assuming a simple principle of compositionality) 18 expresses the same proposition as

19. Nathan believes that Sam Clemens is an author.

Thus, on the Soames-Salmon view 18 and 19 cannot differ in truth value. Though this seems counterintuitive, Soames (1987) and Salmon (1951) offer spirited defenses of this result. Soames (1987) also offers extremely compelling arguments against the view that propositions are unstructured sets of worlds (or circumstances). Some version of the Soames/Salmon view is widely considered to be the standard direct reference view in semantics. Views such as theirs, which make use of structured propositions and endorse direct reference for names, demonstratives and indexicals, are often called *Russellian*.

About the same time the new theory of reference was becoming prominent, quite different developments were taking place in semantics. In pioneering work first presented in the late 1960s (as the William James Lectures at Harvard; later published in Grice [1989] as Essay 2), Paul Grice sought to give a (somewhat) systematic account of (as we would now put it) how the production of a sentence with a certain semantic content can convey further information beyond its semantic content. To give an example from Grice, suppose A and B are planning their itinerary for a trip to France and both know A wants to visit C if doing so wouldn't take them too far out of their way. They have the following exchange:

A: Where does C live?

B: Somewhere in the south of France.

Since both are aware that B offered less information than is required for the purposes at hand, and since B can be presumed to be attempting to cooperate with A, B conveys that she doesn't know where C lives, though this is no part of the semantic content of the sentence she uttered. Grice gave an account of how such information (not part of the semantic content of any sentence asserted) can be conveyed. The account depended on the claim that conversational participants are all obeying certain principles in engaging in conversation. The main idea, as illustrated above, is that conversational partici-

pants are trying in some way to be cooperative, and so to contribute to the conversation at a given point what is required given the purpose and direction of the conversation. Grice's central theoretical idea was that certain types of information exchange and certain types of regularities in conversations don't have purely semantic explanations. The study of how information gets conveyed that goes beyond the semantic content of the sentences uttered falls in the field of *pragmatics*, (which is why, though Grice's work is extremely important, it hasn't been discussed more in an entry on semantics).

In a series of papers that (for our purposes anyway) culminated in Stalnaker (1978), Robert Stalnaker, consciously following Grice, was concerned with ways in which in conversations information can be conveyed that goes beyond the semantic contents of sentences uttered as a result of conversational participants obeying certain principles governing conversation. More specifically, Stalnaker developed an account of how context of utterance and semantic contents of sentences (relative to those contexts) produced in those contexts can mutually influence each other.

Of course, how context influences the semantic content of sentences relative to those contexts was already fairly well understood. As discussed above, for example, context supplies the semantic values relative to those contexts for contextually sensitive expressions such as 'I.' Stalnaker sought to understand how the content relative to a context of a sentence uttered can affect the context. Stalnaker began by introducing the notion of *speaker presupposition*. Stalnaker understood the proposition expressed by a sentence (relative to a context) to be a set of possible worlds (the set of worlds in which the sentence taken in that context is true). Very roughly, the propositions a speaker presupposes in a conversation are those whose truth he takes for granted and whose truth he thinks the other participants take for granted too.

Consider now the set of possible worlds that are compatible with the speaker's presuppositions (the set of worlds in which every presupposed proposition is true). Stalnaker calls this *the context set*, and it is for him a central feature of a context in which a conversation occurs. (Strictly, every participant in the conversation has his own context set, but we will assume that these are all the same—Stalnaker calls this a *non-defective context.*) They contents of sentences (relative to a context) affect the context in the following way: if a sentence is asserted and accepted, then any world in the context set in which the sentence (taken in that context) is false is eliminated from the context set. In short, (accepted) assertions function to reduce the size of the context set, or eliminate live options.

Stalnaker uses this idea to explain a variety of phenomena, including how the utterance of sentences with trivial semantic content (relative to a context) can nonetheless be informative. It is important to see that Stalnaker, like Grice, took his account here to be not part of semantics, but rather to be something that presupposed the semantics of sentences (taken in contexts). In short, like Grice's work, it was work in pragmatics. However, Stalnaker's idea that the information conveyed by the utterance of multiple sentences in a discourse can go beyond anything countenanced by traditional semantics and that it is important to understand the dynamics of conversation to understand how information is conveyed influenced others who went on to develop *semantic* theories that capture the dynamics of conversation, (Lewis [1979] was another important early influence to the same effect).

In the early 1980s, Irene Heim (1982) and Hans Kamp (1981) independently arrived at very similar semantic accounts that were intended to apply to multisentence discourses. Kamp's view is called *Discourse Representation Theory* (*DRT*) and Heim's view is sometimes called that or *File Change Semantics* (*FCS*). To take a simple example of the sort that DRT and FCS were designed to handle, consider a (short) discourse such as:

20. Alan owns a donkey. He beats it.

Using Kamp's formulation, the *discourse representation structure* (*DRS*) associated with the first sentence of 20 would (roughly) look as follows:

x_1 x_2

x_1=Alan

donkey(x_2)

x_1 owns x_2

After the utterance of the second sentence, the DRS associated with the entire discourse would looks as follows, where we have simply added one more line (*a condition*) to the DRS for the first sentence of 20 (we assume that 'He' is anaphoric on 'Alan' and 'it' on 'a donkey'):

x_1 x_2

x_1=Alan

donkey(x_2)

x_1 owns x_2

x_1 beats x_2

Note that expressions like 'a donkey' introduce variables (called *discourse referents*) and predicates ('donkey') into DRS's and not existential quantifiers. Again very roughly, this DRS (and hence the original discourse) is true in a model iff there is an assignment to the variables of the DRS that results in all its conditions being true in that model. It is the requirement that *there is* such an assignment that results in default existential quantification of free variables. So though indefinites like 'a donkey' are not existential quantifiers on this view, they have existential force (in this case, anyway) due to default existential quantification of free variables. Aside from the desire to apply semantics at the level of discourse instead of sentence, much of the motivation for DRT and FCS came from cases such as 20 (and others) in which a pronoun is anaphoric on another expression (see entry on *anaphora*).

DRT and FCS led directly to the development of other semantic accounts designed to capture the dynamics of conversation. In the paper that initiated what is now often called *dynamic semantics*, Groenendijk and Stokhof (1991) make clear that they see their account as a descendent of DRT and throughout the paper they compare their *Dynamic Logic* (*DL*) account to DRT. The basic idea of DL is that instead of thinking of expressions as having "static" meanings, think of meanings as things that given inputs, produce outputs. A bit more formally, think of the meanings (in models) of formulae of first order logic as given by the sets of assignments to variables that satisfy the formulae. So for example, the meaning of 'Fx' in a model M is the set of all assignments such that they assign to 'x' something in the extension of 'F' in the model M. Dynamic logic claims that the meaning of a formula in first order logic is a set of *pairs* of assignments: the first, the input assignment; the second, the output assignment. For "externally dynamic" expressions (e.g. conjunction, existential quantifiers), these can differ and the result is that interpreting these expression can affect how subsequent expressions get interpreted. For since the output assignments can be different from the input assignments for these dynamic expressions, and since the output of these expressions may be the input to subsequent expressions, the interpretation of those subsequent expressions may be affected.

There is currently much research being done within the framework of dynamic semantics, particularly among linguists. Muskens, van Benthem and Visser (1997) provide a good general overview.

There are many important topics in semantics that could not be covered in the present article. These include the theory of generalized quantifiers, the semantics of conditionals, the semantics of non-declarative sentences, the semantics of metaphor and two dimensional semantics. Interested readers are encouraged to pursue these matters on their own.

See also Carnap, Rudolf; Conversational Implicature; Davidson, Donald; Frege, Gottlob; Grice, Herbert Paul; Heidegger, Martin; Hempel, Carl Gustav; Hintikka, Jaakko; Kaplan, David; Kripke, Saul; Lewis, Clarence Irving; Lewis, David; Logical Positivism; Marcus, Ruth Barcan; Meaning; Modality, Philosophy and Metaphysics of; Montague, Richard; Pragmatics; Putnam, Hilary; Reference; Russell, Bertrand Arthur William; Syntax; Tarski, Alfred; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ayer, A.J., ed. *Logical Positivism*. Glencoe, IL: Free Press, 1959.

Beaney, Michael, ed. *The Frege Reader*. Oxford U.K.; Cambridge, MA: Blackwell, 1997.

Bennett, Michael. "Some Extensions of a Montague Fragment." PhD diss, UCLA, 1974.

Carnap, Rudolf. "The Elimination of Metaphysics through The Logical Analysis of Language" (1932). In *Logical Positivism*, edited by A.J. Ayer. Glencoe, IL: Free Press, 1959.

Carnap, Rudolf. *Meaning and Necessity, A Study in Systematics and Modal Logic*. Chicago: University of Chicago Press, 1947.

Church, Alonzo. "A Formulation of the Logic of Sense and Denotation." In *Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer*. New York: Liberal Arts Press, 1951.

Church, Alonzo. "On Carnap's Analysis of Statements of Assertion and Belief." *Analysis* 10 (1950): 97–99.

Davidson, Donald. "On Saying That." *Synthese* 19 (1968): 130–146

Davidson, Donald. "Truth and Meaning." *Synthese* 17 (1967): 304–323.

Frege, Gottlob. "Function and Concept" (1891). In *The Frege Reader*. Oxford U.K.; Cambridge, MA: Blackwell, 1997.

Frege, Gottlob. "On Concept and Object" (1892). In *The Frege Reader*. Oxford U.K.; Cambridge, MA: Blackwell, 1997.

Frege, Gottlob. "On Sense and Reference" (1892). In *The Frege Reader*. Oxford, U.K.; Cambridge, MA: Blackwell, 1997.

Grice, Paul. *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press, 1989.

Groenendijk, J., and M. Stokhof. "Dynamic Predicate Logic." *Linguistics and Philosophy* 14 (1991): 39–100.

Heim, Irene. "The Semantics of Definite and Indefinite Noun Phrases." Doctoral thesis, University of Massachusetts, 1982.

Hempel, Carl. "The Empiricist Criterion of Meaning" (1950). In *Logical Positivism*. Glencoe, IL: Free Press, 1959.

Hintikka, Jaakko. "Modality and Quantification." *Theoria* 27 (1961): 110–128.

Hodges, Wilfrid. "Tarski's Truth Definitions." *The Stanford Encyclopedia of Philosophy*, winter 2001 ed. Available from http://plato.stanford.edu/archives/win2001/entries/tarski-truth/.

Kamp, Hans. "Formal Properties of 'Now.'" *Theoria* 37 (1971): 227–273

Kamp, Hans. "A Theory of Truth and Semantic Representation." In *Formal Methods in the Study of Language*, edited by J. Groenendijk and M. Stokhof. Amsterdam: Mathematical Centre, 1981.

Kanger, Stig. *Provability in Logic*. Stockholm: Almqvist & Wilksell, 1957.

Kaplan, David. "Demonstratives." In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press, 1989.

King, Jeffrey C. "Tense, Modality, and Semantic Values." *Philosophical Perspectives* 17 (2003): 195–245.

Kripke, Saul. "A Completeness Theorem in Modal Logic." *The Journal of Symbolic Logic* 24 (1) (1959): 1–14.

Kripke, Saul. "Semantical Considerations on Modal Logic" (1963). In *Reference and Modality*, edited by Leonard Linsky. London: Oxford University Press, 1971.

Kripke, Saul. *Naming and Necessity* (1972). Cambridge, MA: Harvard University Press, 1980.

Lewis, C. I., and C. H. Langford. *Symbolic Logic*. New York: Century, 1932.

Lewis, David. "General Semantics." *Synthese* 22 (1970): 18–67.

Lewis, David. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1979): 339–359.

Montague, Richard. "Logical Necessity, Physical Necessity, Ethics, and Quantifiers" (1960a). In *Formal Philosophy; Selected Papers of Richard Montague*, edited by Richmond Thomason. New Haven, CT: Yale University Press, 1974.

Montague, Richard. "On the Nature of Certain Philosophical Entities" (1960b). In *Formal Philosophy; Selected Papers of Richard Montague*, edited by Richmond Thomason. New Haven CT: Yale University Press, 1974.

Montague, Richard. "The Proper Treatment of Quantification in Ordinary English" (1973). In *Formal Philosophy; Selected Papers of Richard Montague*, edited by Richmond Thomason. New Haven, CT: Yale University Press, 1974.

Muskens, Reinhard, Johan van Ventham, and Albert Visser. "Dynamics." In *Handbook of Logic and Language*, edited by Johan van Bentham and Alice ter Meulen. Cambridge, MA: MIT Press, 1997.

Russell, Bertrand. *Principles of Mathematics*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1903.

Salmon, Nathan. *Frege's Puzzle*. Cambridge, MA: MIT Press, 1986.

Soames, Scott. "Direct Reference, Propositional Attitudes, and Semantic Content." *Philosophical Topics* 15 (1987): 47–87.

Stalnaker, Robert. "Assertion" (1978). In *Context and Content; Essays on Intentionality in Speech and Thought*. New York: Oxford University Press, 1999.

Tarski, Alfred. "Der Wahrheitsbegriff in den formalisierten Sprachen." *Studia Philosophica* 1 (1935): 261–405.

Weinstein, Scott. "Truth and Demonstratives." *Nous* 8 (1974): 179–184.

Jeffrey C. King (2005)

SEMANTICS, HISTORY OF

The scope of this article is in part determined by the following restrictions. (1) Although the development of semantics in the twentieth century equals or surpasses all that was done earlier, it receives very little attention here because the major theories and theorists of this period are thoroughly discussed in other articles. (2) The only semantic theories considered are those developed by Western philosophers; thus, no account is taken of the theories of meaning propounded, for example, by ancient Hindu philosophers or by European grammarians or linguists. (3) Since semantic theories concerning nonlinguistic signs tend to involve considerations of theories of knowledge generally, they are not discussed here except as they may occasionally bear directly on a theory of linguistic meaning. On the other hand, much of what philosophers have had to say about language is discussed here, whether or not it can be precisely described as semantics.

The contents of this article are arranged as follows.

ANTIQUITY

THE COSMOLOGISTS

Since the earliest Greek philosophers were primarily cosmologists, their views on language are not the most fully developed (or best preserved) of their doctrines. Sources very late in antiquity attributed to Pythagoras (fl. 530 BCE) the view that although the soul assigned names to things, it did so not arbitrarily but on the basis of a natural connection between them, somehow like that between mental images and their originals. Modern historians sometimes credit Heraclitus (fl. 500 BCE) with having thought a great deal about language, but most of the fragments offered in evidence have to do with the *logos*, which surely is to be interpreted as the guiding principle of nature rather than as word or language. While we have nothing of his explicitly on language, it seems likely that Heraclitus did attach philosophical significance to the puns or contradictions in terms on which some of his paradoxical remarks depend.

Semantic theory seems to have made its first definite appearance in philosophy in the monism of Parmenides (fl. 475 BCE), who maintained that only what was true was expressible. He evidently based this remarkable doctrine on the argument that a statement is false if and only if it contains a false name, but a false name is by definition a name lacking a real bearer and hence a name that names or expresses nothing. (His monism of course entailed that there was only one real name-bearer.) Thus he described several words, such as "becoming" and "perishing," as "mere names that mortals have established, believing them to be true"—that is, believing that there really are such processes, which Parmenides denied.

THE SOPHISTS

Language first became a subject of specialized inquiry among the Sophists, who, unlike their philosophical predecessors, were more interested in man than in the cosmos. That orientation alone would probably have drawn them to the study of language, but there was also the fact that they earned their livings teaching people to speak well. Economic as well as philosophical considerations therefore probably played a part in leading them to include at least grammar as an important part of their work. Protagoras (fl. 445 BCE), the first of the Sophists, may also be considered the first grammarian. He distinguished the tenses and something like grammatical moods (classifying sentences as answers, questions, commands, and wishes), and he classified nouns as masculine, feminine, and "inanimate" (a division based on semantical rather than syntactic considerations, since it depended on the particular sex or lack of sex in the things the nouns were used to name). Grammar developed rapidly among the Sophists. Among the more philosophically interesting parts of grammatical theory to be found in Plato, who doubtless learned much of it from the Sophists, are distinctions between subject and predicate, between substantive and adjective, between the active and passive voices, and among types of discourse—political, rhetorical, conversational, dialectical, and technical.

The Sophists originated semantical as well as grammatical inquiries. Prodicus (fl. 435 BCE), who Plato thought was the best of the Sophists on language, seems to have operated on the hypothesis that there were no genuine synonyms, that where there were two words, there were two meanings. In Plato's dialogues Prodicus is depicted drawing instructive distinctions between "enjoyment" and "pleasure," "esteem" and "praise," "fearlessness" and "courage," for example; and he insisted on the study of "the right use of words" as the beginning of education. Protagoras, Prodicus, and Hippias (fl. 435 BCE) are all credited with treatises on "the correctness of names," and Socrates (d. 399 BCE) is depicted discoursing on that subject in Xenophon's *Memorabilia* (III, xiv, 2–7).

Semantics may have become a theoretical issue for the first time in the paradoxical arguments propounded by Gorgias (fl. 435 BCE) in support of his third nihilistic thesis. The three theses were (1) nothing exists; (2) even if something existed, it would be unknowable; (3) even if something existed and were knowable, it would be incommunicable. Gorgias gave four arguments for thesis (3) along the following lines. Suppose there really are things and they can be perceived by our senses. Then (*a*) some of those things will be perceivable by one sense only and others by another sense only; and since one sense cannot perceive objects proper to another sense, a system of audible signs will not permit communication regarding things perceivable only by sight, and so on for the other senses. In any case, (*b*) those supposed things are not identical with any signs one might use to communi-

cate about them, and so one could never convey the things themselves to another person but only the signs.

Moreover, (*c*) even if one could produce signs exactly representing those supposed things, he could not communicate those signs to another person, for the signs themselves are things, and no one can have in his mind the same things that someone else has in *his* mind at the same time. Finally, (*d*) any signs we might use would have to be formed as a result of our perception of those supposed things, but since genuine knowledge of a cause cannot be gained from its effect, no knowledge of those things could be communicated by means of any signs. Occasionally in arguments (*a*), (*b*), and (*c*) Gorgias seems, like Jonathan Swift's Laputans, to have sophistically confused talking about things with handing them around; but not all his paradoxes of communication are transparent, and some passages in Plato and Aristotle suggest that Gorgias's arguments may have helped to shape their semantic theories.

CONVENTIONALISM AND NATURALISM

The oldest surviving arguments in support of a particular semantic theory may be those attributed very late in antiquity to Democritus (fl. 420 BCE), perhaps presented originally in his book *On Words*. He is supposed to have offered the following four considerations in support of his position that the relation between names and things named is conventional (*θέσει*) rather than natural (*φύσει*): (*a*) the occurrence of homonyms, that is, one and the same name for things different in nature; (*b*) the occurrence of synonyms, that is, different names for one and the same thing; (*c*) the occurrence of name-changes while the thing named remains the same in nature; (*d*) the nonoccurrence of verbal analogies corresponding with real analogies, for instance, there is a verb analogous to the noun "understanding" but none analogous to "justice."

In all probability no philosopher ever held a thoroughgoing semantic naturalism, although there are traces of tendencies in that direction in the doctrine attributed to Pythagoras and in the assumptions that appear to underlie the work of Prodicus and Gorgias. The opposition of naturalism and conventionalism as semantic theories forms the point of departure for the development of Plato's semantics of names in the *Cratylus*. Much of the significance of the *Cratylus* and of ancient philosophy of language generally has been obscured, from antiquity onward, by the confusion of this semantic issue with a dispute over the origin of language in which "naturalism" and "conventionalism" were the principal doctrines. In that dispute, however, it was not the naturalist but the conventionalist position that was preposterous, conventionalism in that context being the claim that language first arose as a result of agreements among men or because some especially powerful individual compelled those around him to use his names for things. There are, of course, implications for semantics in theories about the origin of language, but neither Plato nor any other ancient philosopher of the first rank failed to distinguish between the two inquiries.

PLATO

The oldest surviving work of any kind on language is Plato's *Cratylus* (probably written about 388 BCE). The main topic of this dialogue is the nature of the relation between names and things named.

THE *CRATYLUS*. At the beginning of the *Cratylus* a kind of semantic naturalism is attributed to Cratylus and a kind of semantic conventionalism to Hermogenes. All that is said about naturalism at the outset is that it seems unintelligible, and the first serious undertaking is a discussion of the conventionalism advanced by Hermogenes in these words:

> I cannot be persuaded that there is any correctness of name other than convention [*ξυνθήκη*] and agreement [*ὁμολογία*]. For it seems to me that whatever name anyone gives to a thing is the correct one, and if someone changes that name for another, the later one is no less correct than the earlier—just as when we change the names of our slaves. For no name has been generated by nature for any particular thing, but rather by the custom [*νόμῳ*] and usage [*ἔθει*] of those who use the name and call things by it. (384 C–D)

There is nothing in this conventionalism we have not already seen in the Democritean arguments except the claim that "whatever name anyone gives to a thing is the correct one," and Socrates immediately asks whether this claim is intended to apply to private persons as well as to nations (385A). Hermogenes fails to appreciate the difference, and when, as a result, Socrates is on the point of showing that this subjectivist claim destroys the possibility of distinguishing between true and false statements, Hermogenes tries to salvage it by suggesting an analogy between arbitrary individual name-giving and different natural languages (385D–E). The picture presented is that of a conventionalist who recognized that the existence of different autonomous natural languages was strong con-

firmation of his position and was then so carried away as to produce a doctrine of autonomous idiolects, evidently reasoning as follows: Just as the Greek word for horse is no more and no less correct than the Persian word for horse, so there is no basis for correcting a Greek who should decide to use the Greek word *anthropos* where other Greeks use *hippos* and vice versa.

The conventionalism presented as a basis for discussion in the *Cratylus* is entirely plausible except for the obviously untenable doctrine of autonomous idiolects. One consequence of the doctrine is that at any given time a given thing (or type of thing) has just as many correct names as there are people who name it differently (385D). This suggests some sort of Protagorean skepticism in its author, but Hermogenes is ready to agree that "things have some fixed reality of their own, not relative to us or caused by us" (386D). Socrates uses this admission to show the necessity of objectively correct names. There are real things, he says, and real things are not subject to our whims. We recognize that we cannot do certain jobs involving real things simply by fiat. We must make the correct moves, using the correct tools, and the correct tools for a given job cannot be generally described as the first ones anyone may choose (386E–387B). Now in the use of language, names are our tools, and we employ those tools in doing two essential jobs plainly involving real things: "teaching" (communicating the truth) and "classifying things according to their natures" (387B–388B). If "whatever name anyone gives to a thing is the correct one," we clearly have no chance of succeeding in communicating the truth to one another or in developing classification schemes that will "carve reality at the joints."

The destruction of the doctrine of autonomous idiolects leaves a gap in conventionalism, a gap that was there in any case but that would not have been so easily seen if Plato had not thus deliberately marred this conventionalism in order to call attention to it. Not just anyone is an arbiter of the correctness of names; but then "who *does* provide us with the names we use?" (388D). The answer is derived from the sounder portions of Hermogenes's conventionalism, in which he claimed that custom or law generates our names for things. This suggests that the arbiter of custom, or the lawgiver (*νομοθέτης*), may be identified as the name-maker (*ὀνοματουργός*) (388D–389A). The "law-giver" is Plato's personification of a recognized stipulative linguistic authority, more nearly like the French Academy or the *Oxford English Dictionary* than like an individual—Solon, for instance.

This refurbished conventionalism is adequate as far as questions of pronunciation, word order, and usage are concerned; these can be settled by having recourse to the recognized authority. The question raised by the criticism of autonomous idiolects, however, was not, "how do we determine which names are *accepted?*" but, rather, "how do we determine which names are *correct?*" Plato took the two questions to be distinct and made his most important contribution to the semantics of names in answering the second of them. The development of his answer may be traced out as follows.

If the refurbished conventionalism is to do any more than offer an account of the phenomena of a language, it must be augmented by part of Cratylus's naturalism, which was originally stated in these three claims: "(*a*) for each of the things that really exist there is a correctness of name that has been produced by nature; (*b*) that is not a name which some people agreeing together to give as a name do give as a name, uttering a bit of their voice in accordance therewith; but (*c*) there is a kind of correctness of names that is the same for all, both Greeks and barbarians" (383A–B).

At the beginning of the dialogue this position was taken to be unintelligible because it was thought to be in competition with conventionalism as an account of the phenomena of a language. Claim (*b*) does seem to justify the view that the theory is just a wrongheaded account of that kind. Temporarily ignoring claim (*b*), Socrates proceeds to show that this naturalism makes sense as an account of the conceptual underpinnings of all languages. The fact that the word for horse in Greek is "hippos" and in Latin "equus" shows that different linguistic authorities are operative in different natural languages. Both those words are perfectly acceptable, intertranslatable names for horse; and what makes them so is the fact that each of them embodies in different marks (or sounds) a single "ideal name," which belongs to horse "by nature," whose correctness has been produced by nature, and which is the same for all, both Greeks, who say "hippos," and "barbarians," who say "equus" (389C–390A). That single ideal name cannot be the type of which occurrences of "hippos" or of "equus" are the tokens, since it is "the same for all." Nor can it be identified with what Plato called the form of horse, for although the form of horse may be the ideal *horse*, there is nothing of which it could conceivably be a *name*. Instead, the ideal name embodied as well in "equus" as in "hippos" is the correctly framed concept *horse*, and the difference between the two words is merely the difference between two equally good notations. To say that the concept is

framed correctly is to say that it is the concept of the form rather than of individuals participating in that form; to say that its correctness has been produced by nature is to say that it somehow resembles the form. The correctly framed concept *horse* is a logically proper name of the form of horse; it is the ideal name for which all the words correctly translatable into English as "horse" are various notations.

Plato goes on to develop and apply this theory along the following lines. If we should come across a natural language the speakers of which owned horses and cows but had only one name for both species, or had no single name for horse, using instead an indifferently ordered string of names for legs, head, tail, and so on, then we should have a genuine case of incorrect names. The speakers of that language would be laboring under the influence of incorrectly framed concepts, concepts that fail to carve reality at the joints. Thus, we avoid incorrect names (such as "phlogiston") to the extent to which philosophy and science (personified by Plato as "the dialectician") have provided us with a correct conceptual schema (390C–E).

But the embodiment even of correctly framed concepts in the evolving phenomena of a natural language will sooner or later lead to the development of homonyms and synonyms, which, although not incorrect, are infelicitous for the purposes of science and philosophy. Such infelicities could be avoided if we were to construct a precise, consciously designed concept-notation for the use of philosophers and scientists (421E–423E, 424D–425A). And even if we do not or cannot actually construct it, the notion of a perfectly systematic embodiment of correctly framed concepts may serve as an ideal against which to measure the adequacy of technical language (435C). Thus, the frequently recurring project of an ideal language is to be found for the first time in the very first extant treatise on language.

Perhaps the single most unusual feature of this remarkable semantic theory is the doctrine of the ideal name. Within the *Cratylus* itself the identification of the ideal name with the correctly framed concept is not explicit, although it is clearly implied. That implication is strengthened by the many passages in other dialogues in which Plato did treat concepts as a kind of name—for instance, *Theaetetus* 189E, 206D; *Sophist* 263E; *Philebus* 38E–39A; *Phaedrus* 276A.

Cratylus's naturalism and Hermogenes's conventionalism are so expressed in the dialogue as to give every appearance of being simply Plato's devices for raising semantic questions. Each of them contains an obvious, completely gratuitous overextension. (Later in the dialogue [428A ff.] Cratylus's claim [*b*] goes the way of Hermogenes's autonomous idiolects.) Neither position alone is remotely plausible or likely to have been actually held by any philosopher, but each of them contains an essential ingredient of Plato's own semantics of names.

THE *PARMENIDES* AND THE *SOPHIST*. Plato's other major contributions to semantics occur in the later dialogues *Parmenides* and *Sophist*, in which he goes beyond the doctrine of the *Cratylus* in undertaking the connected tasks of (1) giving an account of the semantics of such names as lack existent bearers, (2) refuting the Parmenidean doctrine that false statements express nothing, and (3) giving an account of the semantics of simple statements.

(1) In *Parmenides* 160B–161A there is an attempt to state three necessary conditions for the meaningfulness of a denial of existence. (The example actually employed is the hypothesis "if a One does not exist," which is eminently generalizable.) If we are meaningfully to say of *x* that it does not exist, then (*a*) "there is knowledge of" *x* (since "otherwise the very meaning of … '[*x*] does not exist' would be unknown"); (*b*) *x* is "something different from other things"; (*c*) "this non-existent [*x*] has the characters of being *that*, and *something*, and of *being related to this*, or *to these*, and all other such characters.… If it does not exist, there is nothing against its having many characters; indeed it must [have many characters] if it is *this* [*x*], and not another, that does not exist. If what is [said] not to exist is neither *the* [*x*], nor *this*, and the statement is about *something else*, we ought not so much as to open our lips." These three interdependent conditions do not seem inconsistent with the semantics in the *Cratylus*, and much of what was to be brought out later in the *Sophist* is already implicit in them—for instance, the distinction between existential and predicational occurrences of "is" ("if it does not exist, there is nothing against its having many characters").

(2) When Parmenides or Plato speaks of expressing nothing, he means saying nothing meaningful, rather than saying nothing at all. This is implied in Parmenides' fragments and is quite plain in Plato, when he says, for example, "Must we not assert that [a man] is not even expressing anything when he sets about uttering the words 'a thing that is not'?" (*Sophist* 237E; *Cratylus* 429E). Those words constitute what Parmenides called a false name. A true, or meaningful, name is one having an identifiable existent bearer, a name that signifies something real; and there is no sharp semantical distinction between

true names and true statements—"If we are speaking the truth, evidently the things we are speaking of must be" (*Parmenides* 161E).

Thus the Parmenidean doctrine is that false statements are meaningless, or that truth and meaningfulness are indistinguishable. Although its scope was never restricted, the doctrine makes most sense when applied to statements of the form "*x* exists," with which Parmenides was preoccupied. Such a statement, he would say, either is true or expresses nothing. In order to preserve the possibility of falsity, even in the limiting case of such statements, Plato had to question the Parmenidean dictum and establish that *what is not* has being in some respect (*Sophist* 241D), which he does in the complex, important doctrine of the interweaving of the Forms (252E–259C). However, his most direct answer to the Parmenidean doctrine is developed in his semantics of statements, an account based directly on the ontological theory just mentioned, since "any discourse we can have owes its existence to the weaving together of Forms" (260A).

(3) "Now, remembering what we said about Forms, … let us consider words in the same way.… Words that when spoken in succession signify something, do fit together, while those that mean nothing when they are strung together do not" (261D–E). "Now a statement never consists solely of names spoken in succession, nor yet of verbs apart from names" (262A). Thus "the simplest and shortest possible kind" of statement is exemplified in "Theaetetus sits" or "Theaetetus flies," "because … it gives information about facts or events; … it does not merely *name* something but *gets you somewhere* by weaving together verbs and names. Hence we say it *states* something" (262D). "Whenever there is a statement, it must be *about* something" (262E). Both the examples above are about one and the same existent thing, the bearer of the name "Theaetetus," but the second is a combination of name and verb in which "what is different is stated as the same or [as is actually the case in this example] what is not as what is," and anything "answering to that description finally seems to be really and truly a false statement" (263D).

In the *Parmenides* and *Sophist*, then, Plato not only extended semantics for the first time beyond the consideration of names to that of statements but, in doing so, also distinguished between meaningfulness and truth, showing for the first time that truth depends not merely on names but on certain syntactically regular combinations of verbs and names. It should be noted, however, that he does seem to have taken meaningfulness as the necessary and sufficient condition of grammaticalness.

Plato's semantics of statements may be better appreciated against the background of the semantical doctrines of his contemporaries Antisthenes the Cynic (fl. 390 BCE) and Stilpo the Megarian (fl. 340 BCE). Beginning with the familiar "two names, two bearers" view, Antisthenes managed to reject all predication, on the grounds that what the subject named was one thing and what the predicate named was quite another, and to accept only identity statements of the form "*x* is *x*" or analogies of the form "*x* is like *y*." Stilpo, too, rejected predication, perhaps on ontological grounds, since he insisted on "the unity of being" and may have thought that this could be expressed only in strict identity statements.

ARISTOTLE

Aristotle's primary interest in language was naturally that of a logician, and while his writings contain many passages on semantic questions, there is relatively little developed theory. His semantics of words (he treats of more than just names) is like Plato's in many respects and is to be found mainly in *De Interpretatione*, Chapters 1–3. There he presents, with little or no argument, the following account of signification.

Although there are different natural languages, the people who use them are confronted with the same extramental things. The mental modifications arising from that confrontation are likenesses (*ὁμοιώματα*) of the things, and they are thus the same for all men too. Within a given natural language, written words are conventional symbols (*σύμβολα*) of spoken words. (Aristotle was no doubt unaware of ideographic notations.) The spoken words are, in turn, related to the mental modifications, first of all as symptoms, or natural signs (*σημεῖα*), of them—that is, of the presence of mental modifications in the speaker. More important, the spoken words are related to the mental modifications in the same way that written words are related to spoken words, as symbols of them. Just as written words constitute a conventional notation for (or embodiment of) spoken words, so do spoken words for mental modifications. Discussions of these passages have almost invariably failed to recognize the first of the two relations between spoken words and mental modifications as distinct and have confused the second relation with that of name to bearer.

It seems that, according to this account, words signify things in virtue of serving as symbols of mental modifications resembling those things. What sorts of "things" can words thus be made to signify? Not much is said on that question in *De Interpretatione*, but in *Categories* (Ch. 5) and *Sophistical Refutations* (Ch. 22), for example, vari-

ous words are said to signify (σημαίνειν) "a certain *this*," "a qualification," "a substance of a certain qualification," "passivity," "a certain relation to something else," "a quantity," and so on. More important, "'man' and every common name signifies not a certain *this*, but a quality or a relation or a mode (or something of the sort)" (*Sophistical Refutations* 178b38).

Ambiguity, Aristotle maintained, is theoretically unavoidable, for since "names and the sum-total of formulas [λόγοι] are finite while things are infinite in number … the same formula and a single name must necessarily signify a number of things." This will, however, give us no trouble unless "we think that what happens in the case of the names happens also in the case of the things, as people who are counting think of their counters," which are in a one-to-one correspondence with the things counted (*Sophistical Refutations* 165a5). Although this passage is part of a warning against sophisms of ambiguity, when taken together with the preceding passage it seems to constitute an injunction against seeking the bearer of a common name, as Plato and so many of Aristotle's successors did. A single individual is the bearer of many names in that they are all correctly predicable of it, but "we do not identify having one meaning with being predicable of one thing, since on that assumption even 'musical' and 'white' and 'man' [all of which are predicable of Socrates] would have one [and the same] meaning" (*Metaphysics* 1006b15).

The principal kinds of words recognized by Aristotle were the name (ὄνομα) and the verb (ῥῆμα—"predicate" is possibly a more accurate translation). He described them both as the smallest conventionally significant units, incapable of being true or false independently. A name without a bearer, such as "unicorn," is neither "false" (as some of his predecessors had claimed) nor nonsignificant; and a name combined with "is," "was," or "will be" always produces something true or false. A verb uttered by itself is a name, but it additionally signifies time and "some combination, which cannot be thought of without the components." Because of the latter additional signification, a verb "is always a sign (σημεῖον) of things being said of something else" (*De Interpretatione* 16b24, 16b7).

"Non-man" names nothing definite and so is not strictly a name; analogously, "does not walk" holds indifferently of all sorts of existents and nonexistents. These negated words Aristotle put into the separate categories of "indefinite names" and "indefinite verbs." "Inflections," such as "man's," are not names either, since they produce nothing true or false when combined with "is," "was," or "will be"—nor is "walked" a verb; it is an "inflection," because it signifies additionally "a time outside the present." In "complex names," such as "lifeboat," the parts are significant, but not independently, since, for example, "life" in this occurrence cannot be given an ordinary interpretation (*De Interpretatione*, Chs. 2 and 3). Finally, there are "connections" (σύνδεσμοι), words and phrases that "make many things one" (*Rhetoric* III, 12; 4), which seem to include particles, conjunctions, prepositions, and idiomatic phrases of several sorts and which in one passage of doubtful authorship are said to be nonsignificant (ἄσημοι) (*Poetics*, Ch. 20). (The "connections" are almost certainly the direct ancestors of Priscian's "syncategoremata," which figured prominently in medieval semantics.) This loosely organized classification, vaguely consistent at best, is based on a tangle of semantic and syntactic considerations, but it does contain important advances—for instance, in the treatment of names without bearers and complex names.

Aristotle's semantics of sentences is concentrated in but by no means confined to *De Interpretatione*, Chapters 4–8. Names have no significant parts and complex names no independently significant parts, but a sentence (λόγος) must have independently significant parts. (This is surprising in view of the fact that in a highly inflected language such as Greek there are frequent occurrences of one-word sentences—"I-walk," "he-walks," and so on.) "Every sentence is significant—not as a tool but … by convention," he maintained (16b33), apparently dissociating his view from Plato's in *Cratylus* 386D ff. Plato, however, was talking about names, not sentences, and Aristotle here seems to have gratuitously set aside an insight into the semantics of sentences that was later to be developed by the Stoics. Some sentences, such as "prayers" and future contingents, are neither true nor false according to Aristotle, and he set the pattern for nearly all logicians thereafter when he put such sentences aside and attended solely to the always true or false "statement" (λόγος ἀποφαντικός).

Aristotle maintained that among the independently significant parts of a statement there must be either a name or an indefinite name and a verb or an inflection of a verb arranged in such a way that the whole "signifies something about something." It is only in such a combination that there is truth or falsity, and, as Aristotle put it in the early chapters of *De Interpretatione*, it looks as if he took the combination in question to be one of words. In *Metaphysics* 1027b23, however, he said that "falsity and truth are not in things … but in *thought*; while with regard to *simple* concepts and essences falsity and truth

do not exist even in thought … but … the combination and the separation are in thought and not in things," and he suggested something similar in *De Interpretatione*, Chapter 14 (and elsewhere), as well (cf. Plato, *Republic* 382B). There is no evidence that Aristotle distinguished consistently or clearly between sentences and what later philosophers called propositions or judgments, but such passages indicate at least his sense of the difficulty in locating truth in strings of words, or in a direct relation between strings of words and arrangements of things.

Aristotle seems sometimes to have considered the communicative capacity or public character of a locution as a criterion of its having independent significance. Thus in *Metaphysics* 1006a21 he remarked that if a man "really is to *say anything*," he must "say something that is significant *both for himself and another*"; and in support of his claim that when a verb is uttered by itself it is really a name and signifies something, he noted that on such an occasion "the speaker arrests his thought and the hearer pauses."

THE STOICS

The nature of the Stoics' philosophy of language is the most tantalizing problem in the history of semantics. We know enough of it to say that it was by far the most intricate and probably the most insightful theory of its kind in antiquity and for centuries afterward; but we cannot be certain what its details were, and even its leading principles are sometimes obscured by vague or conflicting testimony. Those Stoics who had most to say about language were, naturally, the logicians, and the difficulty of determining the exact character of what they had to say stems from the fact that none of the many works of the Stoic logicians is extant. The best surviving sources (which date from almost five hundred years after the period of greatest development in Stoic logic and semantics) are Sextus Empiricus, *Outlines of Pyrrhonism*, Book II, and *Adversus Mathematicos*, Book VIII; and Diogenes Laërtius, Book VII. Under these circumstances it is seldom possible to assign a particular doctrine to a particular Stoic, but much of the best of their logic and semantics is very likely to be the work of Chrysippus (c. 280–206 BCE).

Under the Stoic division of philosophy into physics, ethics, and logic, logic was divided into rhetoric and dialectic, and dialectic further divided into an account of language (*περὶ τῆς φωνῆς*) and an account of things signified (*περὶ τῶν σημαινόμενων*). Both these subdivisions contain material relevant to semantics. In their account of language the Stoics distinguished vocal sound generally, "which may include mere noise," from the sort that is articulate (*ἔναρθρος*), that is, capable of being embodied in written symbols (*ἐγγράμματος*). Articulate sound, in turn, may be nonsignificant—for instance, "blityri"—or significant (*σημαντική*); but for any articulate sound to be considered a sentence (*λόγος*) it must be significant and a product of someone's reason (Diogenes Laërtius 7.55–57).

Within that same branch of their dialectic the Stoics recognized five kinds of words and distinguished their semantic or syntactic functions. They were the first who clearly separated (1) names, such as "Socrates," from (2) appellatives (*προσηγορίαι*), such as "man." (Cf. Aristotle's similar but significantly different distinction in *De Interpretatione*, Ch. 7.) A name "*points out* a kind proper to an individual," while an appellative "*signifies* a common kind." (3) A verb "*signifies* a predicate"; (4) a conjunction "*binds together* the parts of a sentence"; (5) an article (possibly also what would now be called a relative pronoun) serves to "*distinguish* the gender and number of nouns" (Diogenes Laërtius 7.58). Thus the function of conjunctions and articles is purely syntactic, the semantic function of (proper) names is different from that of appellatives (or common names), and the appellative and the verb—the standard ingredients of the simplest kind of logicians' sentence—have one and the same kind of semantic function. The appellative occurring in a sentence signifies a subject and the verb a predicate or "something attachable (*συντακτόν*) to the one or more subjects."

Obviously the division between the accounts of language and of things signified was not exclusive, but the transition from the one account to the other as the Stoics conceived of them may be seen in the claim that all we *utter* (*προφέρειν*) is sounds, while what we *express* (*λέγειν*) is matters of discourse (*πράγματα*), or *lekta*—"expressibles" (Diogenes Laërtius 7.57). It is the doctrine of the *lekton* around which the Stoics organized their account of things signified. In its novelty, importance, and difficulty that doctrine overshadows all the considerable remainder of their philosophy of language.

THE LEKTON. Probably the clearest introduction of the notion of the lekton is the one to be found in these passages from Sextus:

> The Stoics … said that three things are linked together: (1) what is conveyed by the linguistic sign [*τό σημαινόμενον*], (2) the linguistic sign itself [*τὸ σημαῖνον*], and (3) the object or event [*τὸ τυγχάνον*]. Of these the linguistic sign is the sound—e.g., "Dion"; what is conveyed by the

> sign is the matter of discourse indicated thereby, which we apprehend over against and corresponding to our thought (while the barbarians do not understand, although they do hear the sound); and the object or event is the extramental entity—e.g., Dion himself. Two of these are corporeal—viz. the sound and the object or event—and one is incorporeal—viz. the matter of discourse conveyed by the linguistic sign, the lekton. (*Adversus Mathematicos* 8.11–12)
>
> They also say that the lekton comes into being as corresponding to a rational presentation [*λογικὴν θαντασάν*], and that a rational presentation is one presenting something that can be set forth in a sentence. (*Adversus Mathematicos* 8.70)

The kind of lekton associated with the name "Dion" was said to "stand in need of completion," and the only categories cited for such completable lekta were subjects and predicates. In order to be "set forth in a sentence," the completable lekta must enter into the composition of a lekton "complete in itself." The kind of complete lekton regularly associated with a standard subject-predicate sentence was called a statement (*ἀξίωμα*), and truth or falsity was ascribed to it, not to the sentence. Statements naturally received most attention from the Stoic logicians, but they recognized many other varieties of complete lekta as well. The fact that they did so strongly suggests that they had developed other categories of completable lekta too, for most of the other complete lekta cannot be analyzed into subject and predicate. Among the other varieties were commands, prohibitions, yes-no questions, questions requiring more than "yes" or "no," curses, prayers, doubts ("Can it be that life and pain are akin?"), and quasi statements ("How like to Priam's son the cowherd is!") (*Adversus Mathematicos* 8.71–73; Diogenes Laërtius 7.65–68).

Since these are categories of incorporeal lekta rather than of sentences, they cannot be identified with strictly grammatical categories. Moreover, although some of the distinct lekta do correspond to grammatically distinct sentences—for instance, the two kinds of questions—many of them do not. The Stoics' own example of the kind of lekton called a doubt was expressed in what is grammatically a yes-no question; commands and prohibitions get expressed in declarative as well as in imperative sentences, and occasionally both may be expressed in one and the same sentence, for example, "Abstain from strong drink." Thus Plutarch reports, in his attack on the Stoics, that "they themselves maintain that those who forbid *say* one thing, *forbid* another, and *command* a third. For he who says 'you ought not to steal' forbids stealing and commands not stealing at the same time as he says you ought not to steal" (*On the Contradictions of the Stoics* 1037d).

As many as three different complete lekta may, then, be associated with a single sentence, and those lekta are obviously not to be identified as thoughts or intentions on the part of the speaker or hearer. Nor does it seem likely, despite Plutarch's way of presenting the doctrine, that all the complete lekta associated with a given sentence must be expressed whenever the sentence is uttered. Besides being far-fetched, that requirement would ignore the sense of express*ibility* built into the Stoics' technical term "lekton." Instead, the Stoic doctrine seems to be that a number of distinct linguistic jobs—such as stating, commanding, prohibiting—can be performed by means of a single sentence, depending on which of the complete lekta associated with that sentence is actually communicated on a given occasion of its use. Thus the three lekta associated with the example given by Plutarch may be presented as (1) the statement that one ought not to steal, (2) the command not to steal, and (3) the prohibition of stealing. It seems to be a discovery of the Stoics (and their greatest contribution to semantics) that the explication of meaning involves not only the things we talk about and the thoughts we express but also the jobs we do by means of language alone.

THE EPICUREANS

Of the Stoic semantic triad—linguistic sign, what is conveyed thereby (the lekton), and external object or event—the Epicureans accepted only the first and third, ascribing truth and falsity directly to spoken sentences (Sextus Empiricus, *Adversus Mathematicos* 8.13). This rejection of the lekton is typical of the Epicureans' mistrust of any doctrine that went beyond the evidence of the senses. Plutarch describes them as "completely doing away with the category of lekta, leaving only words and objects and claiming that the intermediate things conveyed by the signs simply do not exist" (*Adversus Coloten* 1119F), but there is also a vague suggestion that they may have found it convenient to provide "lekta" as dummy referents in one important kind of case. "They deprive many important things of the title of 'existent,' such as space, time, and location—indeed, the whole category of lekta (in which all truth resides); for these, they say, are not *existents* [*ὄντα*], although they are *something* [*τινά*]" (*Adversus Coloten* 1116B).

In stating their atomist metaphysics, the Epicureans were of course obliged to use such words as "space" and "time," and it looks as if they may have clumsily attempted to provide referents for them by associating only such words, or sentences containing such words, with lekta. However, even if they did maintain that there are two kinds of referents for words, real things and lekta, the latter to be invoked only in case the former are unavailable and the words are indispensable, there is nothing of the *Stoic* lekton in their doctrine.

Aside from this putative special use of special lekta, the Epicureans' philosophy of language seems to have remained remarkably faithful to their fundamental sensationalism. Epicurus (341–270 BCE) had originally stressed the importance of beginning the study of physics (one of the main branches of Epicurean, as of Stoic, philosophy) by ascertaining the ultimate referents (*ὑποτεταγμένα*) of words, "so that our proofs may not run on untested indefinitely nor the terms we use be empty. The primary intent (*ἐννόημα*) of every term employed must be clearly seen and ought to need no explication" (Diogenes Laërtius 10.37–38); and he went on to claim that these ultimate referents must then be "our sensations," "present impressions," "actual feelings." These are always veridical since their immediate causes are the *eidola*, and thus "the agent productive of each of them is always entirely presented and, as being presented, it is incapable of being productive of the presentation without being in very truth as it appears.… Thus the visible object not only appears but actually is as it appears.… The presentations that occur are, then, all true" (Sextus Empiricus, *Adversus Mathematicos* 7.203–204). The square tower in the distance appears round, but its round appearance is itself a physical object, an *eidolon* detached from the tower and impinging on the apparatus of sight. If I say, then, "the tower is round," I may (and in this case I shall) be mistaken, since the tower is not the immediately presented object. But if I say "the appearance (or presentation) of the tower is round," I cannot be mistaken (at least not in that same way). Although this is a move in the direction of protocol sentences, it does not rest on a distinction between sense-datum and physical-object sentences, since for the Epicureans the protocol sentence was only a more correctly framed physical-object sentence.

No full account of this Epicurean reductivism is extant, but its principle is clearly operative not only in their physics but in their ethics as well, where one pervasive maxim for the avoidance of fear is to reduce the mysterious (for example, in natural phenomena) to what is actually presented and to describe it in terms precisely associated with the features of the actual presentation (see, for instance, Diogenes Laërtius 10.78 ff.).

Epicurus's followers evidently took the nature of the relation between words and sensations as a major topic for psychological theory. The notion of *prolepsis* is at the center of the Epicurean psychology, and in at least one of its many guises prolepsis seems to be the act of associating a word with a *typos*, or outline left in the mind as the result of repeated similar presentations. One example of prolepsis is the identification "such and such a thing is a man"—"for no sooner is the word 'man' uttered than we think of the typos of man in accordance with the prolepsis, the senses having led the way. As a result, the immediate referent of every name is apparent.… Nor would we have given a name to anything if we had not first come to know the typos of it in accordance with prolepsis" (Diogenes Laërtius 10.33).

The typos, then, is the immediate referent (*τὸ πρώτως ὑποτεταγμένον*) of every name. When a name is used and understood, an act of prolepsis at once brings the corresponding typos to mind. (Since nothing but sensation can produce a typos, the need for some other sort of referent in the case of words such as "space" and "time" is apparent.) If this was indeed the core of the Epicurean semantics of words, it must be judged inferior to many other theories of its kind in antiquity.

Epicurus himself and the Epicureans generally had a good deal to say about the origin of language, and what they said usually makes better sense than most such accounts in antiquity. Lucretius (99–55 BCE) is especially good on this topic, which he treated at some length in Book V of his poem. Among his more novel and interesting achievements is an extended series of arguments against the theoretical possibility that language (as distinguished from *a* language) might have been *invented* (*De Rerum Natura* 5.1041 ff.).

THE MIDDLE AGES

ST. AUGUSTINE

Most of what St. Augustine (354–430) had to say about language and meaning was said not for its own sake but in support or elucidation of some theological doctrine. Partly for that reason, perhaps, his semantic doctrines had less effect on philosophy of language in the Middle Ages than might be expected, considering his enormous influence on medieval philosophy in general.

The short treatise *Principia Dialecticae* (probably written around 384, when Augustine was a professor of rhetoric) contains what may be the only instance of a semantic inquiry pursued by Augustine without a motive. In it he distinguishes four principal semantic elements: (1) the word (*verbum*), a spoken articulate sound, classifiable as a vocable of some language; (2) the expressible (*dicibile*), "whatever is sensed in the word by the mind rather than by the ear and is retained by the mind"; (3) the ordinary use of the word (*dictio*), (opposed, for instance, to the use of the word as a sign for itself), which involves "both [1] the word itself and [2] that which occurs in a mind as the result of the word" when the word occurs "not on its own account but on account of something else that is to be signified"; (4) the signified thing (*res*), which may be "something understood, or sensed, or inapprehensible"—the last category reserved for, for instance, God and formless matter (Ch. 5). Of these four elements, (2) and (3) together seem to represent different aspects of the Stoic lekton; but whatever their origin, their inclusion here indicates a level of sophistication in semantics that was not to be attained again for at least eight hundred years.

Chapter VII of the *Principia Dialecticae* is devoted expressly to the "import" (*vis*) of words. In it Augustine maintains that "the import of a word is that whereby the extent of its efficacy is recognized [*qua cognoscitur quantum valeat*], and it is efficacious to the extent to which it can affect a hearer." Import is a broader notion than signification and includes several sorts of effects a given word may have because of its sound, its degree of familiarity to the hearer, its degree of admissibility into polite conversation, its being recognized by the learned hearer as a dactylic foot or as some particular part of speech, and so on. The paradigm case of signification is described as occurring "on an occasion when a sign has been comprehended through a word [and] the mind regards [*intuetur*] nothing other than that very thing the sign it comprehends is a sign of. Suppose, for example, that Augustine has been named and someone to whom I am known thinks of nothing other than myself, or that some other man named Augustine comes to mind if the name happens to be heard by someone who does not know me but knows that other man." The most remarkable and apparently novel features of this brief account are (*a*) the extension of the notion of meaning in Augustine's doctrine of "import" and (*b*) the orientation of his account of meaning around the effects words have on their hearers. The remainder of the treatise deals with simple and conjoined words and sentences, etymology, and various types of ambiguity and obscurity.

The longest of Augustine's discussions of semantic questions occurs in the dialogue *De Magistro* (389), which is designed ultimately to support the Augustinian doctrine of "divine illumination" as the sole genuine source of truth. Thus the first 11 chapters are supposed to show that "we learn nothing through those signs that are called words" (Ch. X), while chapters XI–XIV develop the thesis that Christ, the truth, teaches us inwardly while men by their use of outward signs merely prompt us to raise questions. The argument in support of the negative conclusion is an outstanding example of overemphasis on the word as the unit of signification. "When words are uttered, either we know or we do not know what they signify. If we know, then we do not learn but are reminded. If, on the other hand, we do not know, then we are not even reminded (though we may be prompted to ask)" (Ch. XI). Therefore, "we learn nothing through those signs that are called words"—as if one's knowing what the words mean in "armadillos are mammals" precluded one's learning anything through hearing that sentence uttered. At this crucial juncture in the dialogue Augustine's ulterior motive seems to have distorted his judgment.

Perhaps the most interesting point in the early chapters of *De Magistro* is one that bears on the best-known Augustinian passage on language, the description of his learning to speak in *Confessions* (397), 1.8, made famous by Ludwig Wittgenstein's use of it in *Philosophical Investigations* (Sec. 1). The passage in the *Confessions* can hardly be considered a theoretical statement at all, since Augustine's main aim in it is to describe a milestone on his descent "into the stormy fellowship of human life," but it does contain a brief, uncritical account of one way in which a child might be shown "the things of which words are signs."

That this account cannot be considered important in the context of Augustine's own views on language is plain from the fact that he had already criticized just such an account on theoretical grounds in *De Magistro*. In an attempt to refine the original suggestion of the dialogue that a sign cannot be a sign "unless it signifies something" (Ch. II), Augustine asks to be shown "that one thing itself, whatever it is, which is signified by these two words," *ex* and *de*, Latin prepositions there taken to be synonymous. After several obviously unsuitable suggestions, the tentative conclusion is reached that not only in these problematic cases but also in every case of attempting to show "the thing signified," all that can be shown is further signs, such as other words, pointing, pantomiming. This criticism is of course not the same as Wittgenstein's, nor is it

particularly far-reaching in its own right, since it is soon modified to allow that we can in certain cases display the very thing signified without the use of further signs—for instance, if the thing signified is something we are able to do (such as walk) and we are not in the act of doing it when asked to display the thing signified (Ch. III).

A rather fully developed semantic theory appears in the *De Trinitate* (399–419), especially in 9.7–12 and 15.10–16, although it is presented no more for its own sake than is the theory in *De Magistro*. The theory appears as the explanatory half of an ingenious analogy designed to clarify (*a*) the relation between the First and Second Persons of the Trinity, (*b*) the two natures of the Second Person, and (*c*) the identification of the Second Person as the Word. The analogical points may be ignored for present purposes and the semantic doctrine sketched as follows. "Word" has at least two senses. "In one sense, those things are called words that occupy intervals of time with syllables, whether they are pronounced or only thought. In another sense, everything that is known is said to be a word impressed on the mind as long as it can be brought out of memory and defined" (9.10). (Augustine actually introduces a third sense involving the love of what is known, but it seems pointless except for purposes of the analogy.) The second kind of "word," which Augustine describes more generally as a "locution" when the demands of the analogy are not uppermost in his mind, occupies the central position in the doctrine. "The word that sounds outwardly is a sign of the word that gives light inwardly, and the name 'word' is better suited to the latter; for what is uttered by the mouth of the flesh is the articulate sound of the word [*vox verbi*]; … [thus] our word becomes an articulate sound … by taking on [articulate sound], not by consuming itself so as to be changed into it" (15.11).

The doctrine of the inward locution sometimes bears a striking resemblance to Plato's doctrine of ideal names in the *Cratylus*, although a direct historical connection seems unlikely. "For of necessity, when we say what is true—i.e., say what we know—the knowledge itself, which we retain in memory, gives birth to a word that is altogether of the same kind as the knowledge from which it is born. For the thought formed by the thing that we know is a word that is neither Greek nor Latin nor of any other language. But since it is necessary to convey it into the knowledge of those with whom we speak, some sign is adopted by which it is signified" (15.10; cf. *Sermo* 225.3). According to this doctrine, then, it seems that one's saying "armadillos are mammals" embodies in sounds one's inward locution to that effect, which itself differs from one's knowledge that armadillos are mammals only in being brought out of memory into conscious thought. Augustine sometimes suggests that the inward locution, then, is not itself verbal; words used in the mind are not essentially different from words outwardly pronounced, as Augustine's first division claims. Indeed, the inward locution is evidently less a mental entity than the state of consciousness into which a mental entity, namely, a known truth, must be brought if it is to be given verbal expression.

BOETHIUS

As an original contributor to semantics, Boethius (480–524) is much less interesting than Augustine. Since, however, his translations and commentaries constituted the sole source of Aristotelian logic for the medievals until the twelfth century, Boethius's influence over the development of semantics in the Middle Ages is powerful where Augustine's is slight.

Most medieval semantic theories take as their starting point Boethius's translation of the rudimentary account in Aristotle's *De Interpretatione*, Chapter 1. No doubt the traditional misreading of those passages during and after the Middle Ages is largely the result of the fact that in his otherwise faithful rendering Boethius obliterated the Aristotelian distinction between symbols and symptoms, translating both *σύμβολα* and *σημεῖα* as *notae*. Another of the principal difficulties in Aristotle's account—the apparent interposition of "mental modifications" between words and things—had been discussed at least as early as the third century by Alexander of Aphrodisias, whose confusing resolution of the difficulty was transmitted to the medievals in Boethius's second commentary on the *De Interpretatione*. Alexander had asked whether Aristotle's account forces us to consider the mental modifications as names of things. In order to avoid that consequence he had developed the view that although "a name is imposed on a thing" and "although spoken words are names of things, nevertheless we use spoken words not in order to signify things, but in order to signify those mental modifications that are produced in us as a result of the things. Therefore, since spoken words are uttered for the purpose of signifying those entities, he [Aristotle] was right to say that they are primarily the signs [*notas*] of those entities" (413A–B; all references in this section are to *Patrologia Latina*, edited by J. P. Migne, Vol. 64).

Perhaps the most influential doctrine (at least in the late Middle Ages) that can be traced directly to Boethius's treatment of *De Interpretatione*, Chapter 1, is that of the

three discourses: written, spoken, and mental. Citing Porphyry (c. 233–c. 305) as his authority, Boethius reported that "among the Peripatetics there were said to be three discourses [*orationes*]—one written in letters, another uttered in speech, and a third put together in the mind. Now if there are three discourses, the parts of discourse are no doubt likewise threefold; for since the noun and the verb are the principal parts of discourse, there will be some nouns and verbs that are written, others that are spoken, and still others that are silent and employed by the mind" (407B–C). Here, as in his transmission of the Aristotelian account itself, the vagueness of Boethius's presentation is as important historically as its content. Are there two completely different sets of nouns and verbs, one for writing and one for speech? And is this mental discourse nothing more than silently running over a sentence in Latin or English, Or is it a nonverbal operation, reminiscent of Augustine's "inward locution"? The fact that mental discourse is said to have nouns and verbs of its own suggests the former view, if either; but since Aristotle had maintained that the mental modifications were the same for all (regardless of their native tongue), and since Boethius offers this doctrine of the three discourses in explanation of Aristotle's account, there is some basis for the second view as well. These were among the difficulties discussed in the medieval development of the doctrine.

The medieval distinction between words of first and second "imposition," a genuine prefiguring of the twentieth-century distinction between object language and metalanguage, also has its roots in Boethius's transmission of older doctrines. In his commentary on Aristotle's *Categories* he presents the distinction very much as he found it in Porphyry's *Expositio* of the same work (A. Busse, ed., pp. 57–58). "The first imposition [*positio*] of a name is made with respect to the *signification* of the word, the second with respect to its *form*" (159C). Thus, whenever some extralinguistic entity is called a man, it is a case of first imposition. "But when the word 'man' itself is called a noun, no reference is made to the signification of the word [Boethius has 'noun'], but to its form, in virtue of which it admits of inflection by means of [grammatical] cases" (159B–C). Thus "noun" in its ordinary use is a word of second imposition.

In this primitive form the distinction seems to apply only to the grammarian's kind of interest in discourse. Boethius, however, took the position that "the whole art of logic is concerned with discourse" (161C–D). How does the philosopher's interest in language differ from the grammarian's? Very much as the economist's interest in money differs from the numismatist's, for Boethius compares the signification of a word to the buying power of a coin and its grammatical form to the "bronze stamped with a design." Consider "an utterance that designates nothing, such as 'gargulus.' Although the grammarians, considering its form, contend that it is a noun, philosophy does not recognize it as a noun unless it is imposed in such a way as to designate a conception belonging to a mind (in which same way it can signify some real thing)" (408C–D). Apparently, then, second imposition needs to be more broadly conceived, or a philosopher's kind of second imposition must be added to the kind described by Boethius. The resolution of such difficulties was among the goals of the later doctrine of the impositions and "intentions" of words.

By far the most influential of Boethius's bequests to the Middle Ages was his formulation of the problem of universals in his second commentary on Porphyry's *Isagoge*. Needless to say, a great many semantic issues were discussed in the long controversy over universals, and a few of the more important ones will be noted below. Boethius's formulation of the problem, however, was oriented around questions of metaphysics rather than of semantics and so may be passed over here.

ST. ANSELM

One of the semantic problems recognized by the early medievals in the few logical works of Aristotle available to them was the problem of paronyms, or denominatives. Its principal source is the following passage in the *Categories*, Chapter 8 (10a27 ff.). "These, then, that we have mentioned are qualities, while things called paronymously because of these or called in some other way from them are qualified. Now in most cases, indeed in practically all, things are called paronymously, as the pale man from paleness, the grammatical from grammar, the just from justice, and so on."

St. Anselm (1033–1109) remarks at the end of his dialogue on denominatives—*De Grammatico*—that the semantics of denominatives was a favorite topic among dialecticians of the eleventh century, evidently because of the difficulty of developing a satisfactory account of denominative words that occur both as concrete nouns and as adjectives. (Anselm's chief example is *grammaticus*, but because the English word "grammatical" is not a denominative of this sort, "illiterate" will be used here.) Thus, the opening question of Anselm's dialogue is whether "illiterate" signifies a substance or a quality. This seems to be a narrow, perhaps artificial problem, but

under his characteristically ingenious treatment it leads to results of general importance.

The superficially most plausible solution to the problem is that such a word sometimes signifies a substance—as in "not every illiterate is stupid"—and sometimes a quality (illiteracy)—as in "not every illiterate person is stupid." This solution is shown to fail, however, at least in its second half, for if we tried to use "illiterate" alone in speaking about the quality—as in "illiterate is a deplorable condition"—"not only the grammarians would be upset, but even the peasants would laugh" (Ch. XI). "Illiterate," we must recognize, "does not signify a person and illiteracy as a unit [*ut unum*] but signifies illiteracy directly [*per se*] and a person indirectly [*per aliud*]" (Ch. XII). Another way to put the distinction between the two kinds of signification is to say that "illiterate" is *significative* of illiteracy and *appellative* of a person. "I now describe a name as appellative of each thing itself that is called [*appellatur*] by that name in the speaker's usage; for there is no speaker's usage in which 'illiteracy is illiterate' occurs, … but rather 'the person is illiterate'" (Ch. XII).

The remainder of the dialogue refines and generalizes this account in the course of dealing with various objections to it. Anselm's most original and important contributions seem to be those developed mainly in the last two chapters (where the discussion centers around *albus*—"white"—rather than *grammaticus*). The Master of the dialogue has suggested that "white" signifies (rather than appellates) nothing but being in possession of whiteness (*habens albedinem*). This is disturbing to the Student, who feels the need of a signified *thing*. "White," he is willing to grant, "does not determinately signify this or that possessing entity, such as a body," but he wants to insist that it "indeterminately signifies something possessing whiteness." His principal argument is that "'white' signifies either something possessing whiteness or nothing; but one cannot conceive of nothing as possessing whiteness; therefore it is necessary that 'white' signify something possessing whiteness" (Ch. XX).

In reply Anselm takes the position that while it may always be the case that what is signified somehow depends for its being on some real thing, it cannot always be the case that what is signified is a thing. His arguments for this position display an interesting use of the principle of substitutivity.

If "white" signified a thing at all, it would signify something white. Now the signification of a word is what its definition presents, and what is presented by the definition may be substituted for the word itself. "So wherever 'white' is used it is taken correctly as 'something white'" (Ch. XXI). Then "Socrates is white" may be rewritten as "Socrates is something white." But "wherever 'something white' is used it is also correctly said twice—'something something white'—and wherever it is said twice, there also three times, and so on indefinitely" (Ch. XXI). Thus the plausible "Socrates is something white" would become the nonsensical "Socrates is something something white" and would ultimately lose all semblance of a statement.

Instead, in "Socrates is white," "white" appellates something white—Socrates himself—but what it signifies is being in possession of whiteness. Nor will it do to introduce a signified thing at this point, for if we take something in possession of whiteness to be what "white" signifies, we shall have to grant that something in possession of whiteness is that which is white. "If, therefore, 'white' is 'that which is white,' it is also 'that which is that which is white'; and if it is that it is also 'that which is that which is that which is white,' and so on indefinitely" (Ch. XXI). The nonsense-engendering substitutions cannot be made within "being in possession of whiteness," however, since the denominative "white" does not itself occur in it. Thus "it is clear enough that 'white' does not signify something in possession of whiteness …, but only being in possession of whiteness—i.e., [the categories] *quality* and *possession* [and not the category *substance*]—and quality and possession by themselves make up no *something*" (ibid.). This argument is described as holding good for all single words that, like "white," signify "a plurality [of categories] out of which no one thing is made up" (Ch. XXI).

Although the special consideration of denominatives apparently lost its vogue soon after Anselm, many of the problems dealt with in his *De Grammatico* remained current and can be found two centuries afterward at the center of the theory of the properties of terms (see below).

ABELARD

The extensive logical writings of Peter Abelard (1079–1142) are best known for the theory of universals developed in them. That theory is important in the history of semantics because (*a*) it explicitly approaches the problem of universals as a semantic rather than a metaphysical problem and because (*b*) in doing so it introduces many of the elements of the semantic theories developed by the terminist logicians of the thirteenth and fourteenth centuries.

Regardless of how the problem of universals is approached, it involves a consideration of the semantics of words, especially of common names. Nevertheless,

many of the countless medieval theories, in their preoccupation with the Porphyrian-Boethian questions about the existential status of genera and species, slighted or ignored the semantic issues. Abelard, on the other hand, began by adding a new semantic question to the three traditional metaphysical questions. "Could a universal consist of the signification of a concept [*significatione intellectus*] when the things named were destroyed, as [in winter] when there are no roses to which the name 'rose' is common?" (*Logica "Ingredientibus*," edited by B. Geyer, p. 8).

Having associated universals with words, Abelard asked "whether they are associated *only* with words or with things as well" (p. 9). Applying the Aristotelian criterion *predicability of more than one thing*, he showed in a series of elaborate arguments that a universal cannot be identified as (1) a single thing or (2) a collection of things (pp. 10–16). His principal objection really avoids the issue of whether or not it makes sense to speak of a thing or a collection as predicable at all and concentrates instead on the impossibility of predicating a thing or a collection of more than one thing. Thus "it remains for us to ascribe universality of that sort to words [*vocibus*] alone" (p. 16). As Abelard came to realize, words considered as utterances or inscriptions are themselves things. Accordingly he eventually distinguished between utterances [*voces*] and words [*sermones*] and organized his theory of universals around words in this strict sense—*sermo = vox + significatio*—which he described as products of human arrangements rather than mere natural effects (*Logica "Nostorum Petitioni Sociorum*," edited by B. Geyer, p. 522).

The only kind of word to which universality can conceivably be ascribed is the kind of word apparently predicable of more than one thing, that is, a common name in the nominative case. But that ascription cannot mean that the common name has some universal thing as its bearer, for, as he had shown, "universal thing" is a contradiction in terms. Nor can some particular thing be picked out as its bearer, for although it *may be* Socrates alone of whom the statement "a man is sitting in this house" is true, we cannot *infer* from it that Socrates is sitting in this house (p. 18). These considerations led Abelard to base the ascription of universality not on what the words name (*nominare*)—for example, "man" names each and every individual thing that is a man—but on their "mode of signification"; for although they name things that are discrete, they do so not "discretely and determinately" but "confusedly" (p. 29).

Abelard's explanation of this notion of confused naming, which was to play an important part in thirteenth- and fourteenth-century theories of the properties of terms, seems incomplete but runs along the following lines. "To signify is to establish [*constituere*] a concept" (p. 136), and "when I hear 'man' … I do not recall all the natures or properties that are in the things subject [to that name]; instead, as a result of [hearing] 'man' I have a conception of animal, rational, and mortal, though not of subsequent accidents as well, [a conception that is] confused rather than discrete" (p. 27). Thus Abelard's answer to his additional semantic question is a qualified "yes." In winter the name "rose" lacks universality in that there are no things of which it is predicable, that is, "it is devoid of nomination" (*nominatione*). "Nevertheless, it is still significative then in virtue of the concept [*ex intellectu*]; … otherwise there would not be the proposition 'no rose exists'" (p. 30).

Other medieval theories of universals, such as William of Ockham's, center on semantic doctrines; but Abelard's "sermonism" was perhaps the most important medieval influence on the development of semantics during the succeeding two centuries of the high Middle Ages. Topics and terminology remained relatively stable in that remarkable period in the history of semantics, although many philosophers and every logician contributed to the discussions. For that reason the remaining material on the Middle Ages is oriented mainly around *topics* in medieval semantics, and no attempt is made to mention every man who discussed them.

IMPOSITIONS AND INTENTIONS

The pervasive medieval distinctions between two levels of signification have attracted some attention in the twentieth century because of their resemblance to the object language–metalanguage distinction. Historically there were two such distinctions, both based on the observation, found already in Porphyry, that while some signs signify nonsigns, others are signs of signs.

The original distinction was drawn with respect to conventional signs, specifically, with respect to names (nouns and adjectives) in a natural language. Such signs acquired their signification only as a result of having been imposed by the users of the language. The primary, or first, imposition was on extralinguistic entities, and names such as "man" and "white" were classified as names of first imposition. As the language developed, other conventional signs were imposed on conventional signs as such; thus, names such as "noun" and "plural" are of second imposition. Names such as "utterance" and "mark"

do signify conventional signs, but not as such (since there are of course nonsignificant utterances and marks); they are therefore names of first imposition. Most medieval logicians, presumably in avoidance of an infinite regress, were careful not to define names of second imposition as names of names of first imposition. Thus "name of second imposition" is itself a name of second imposition.

But even those who, like Abelard (*Logica "Ingredientibus*," edited by B. Geyer, p. 112), did define second imposition in terms of first seem never to have recognized a "third" imposition. (The imposition distinction, therefore, cannot reasonably be described as prefiguring a hierarchy of types.) The use made of the imposition distinction was apparently rather meager. Aristotle's categories were, for example, often said to be names of first imposition, while the subject matter of his *De Interpretatione* was described as names of second imposition. The distinction, although it was refined and discussed well into the fifteenth century, seems to have acquired what importance it had mainly from its connection with the later and better known of the two distinctions between levels of signification.

Concepts in their capacity as natural signs were called intentions and described in the doctrine of the three discourses as mental terms. It was only natural, then, to distinguish levels of signification among intentions as among conventionally significant extramental terms. This distinction, probably stemming from Avicenna (see Carl Prantl, *Geschichte der Logik im Abendlande*, 2.328), classified as first intentions all those naturally significant of entities other than intentions as such, while those that did naturally signify intentions as such were second intentions. The concept *humanity* is of course a first intention, but so is the concept *mental entity*. The concepts of the predicables—*genus, species, differentia, property, accident*—are second intentions, as is the concept *predicable* itself; no "third" intentions were ever recognized.

Thus first and second intentions and impositions were fundamentally parallel distinctions in separate domains. However, their development in the thirteenth and early fourteenth centuries was complicated (and sometimes confused) by two factors. First, there were, of course, extramental terms imposed on first and second intentions, such as *humanity* and *genus*. Such names were all of first imposition (since no intention was a conventional sign), but they were sometimes further described as names of first or second intention. Second, even more complicating was the fact that the first and second intentions themselves were considered to be terms in mental propositions. Thus, while in the written proposition "*animal* is a genus" the subject and predicate terms are both of first imposition, in the corresponding mental proposition that *animal* is a genus the subject term is a first intention and the predicate term is a second intention.

Of the two distinctions between levels of signification, the intention distinction had much more philosophical importance. The confusing interrelations of the two distinctions are perhaps best exhibited in William of Ockham (d. 1349), particularly in *Summa Logicae*, 1, 11–12. Logicians after William—for instance, Albert of Saxony (d. 1390), Pierre d'Ailly (1350–1421), Paul of Venice (d. 1428), and Paul of Pergula (d. 1451)—exhibited a tendency to simplify them by reverting to the treatment of impositions and intentions as strictly separate, parallel distinctions. Postmedieval scholastics—for example, John of St. Thomas (1589–1644)—were inclined to apply the intention distinction indifferently to extramental as well as to mental terms and to ignore the imposition distinction; it is in this simplified form that the "medieval" distinction between levels of signification is usually discussed in recent literature.

SCIENTIA SERMOCINALIS

Almost everything genuinely novel in medieval logic is to be found in the theories of the properties of terms and of the functions of syncategorematic words developed by the logicians of the high Middle Ages. One reason why logic set off along that line of logicosemantic inquiries is that medieval logicians, especially through the formative period ending about 1250, thought of their subject as the science of language (*scientia sermocinalis*).

That classification itself marked a break with the Aristotelian-Boethian tradition in that it was precise where the tradition had been vague. The notion of predication was unquestionably an essential part of the subject matter of logic, but Aristotle and Boethius had treated it in ways that often suggested that predicates might be extralinguistic and even extramental entities. This crucial vagueness, which was to some extent also the source of the medievals' concern with universals, left open the possibility that logic might be essentially a science of reality, resembling or subsumed under metaphysics.

However, in the earliest complete European logic we have after Boethius—the *Dialectica* of Garland the Computist (d. before 1102)—that possibility was already noted and explicitly ruled out. Predication, Garland maintained, occurs only in a proposition, and the only constituents of propositions are utterances; thus, only

utterances may be predicated. The five predicables (genus, species, differentia, property, accident), the elementary subject matter of medieval logic, are, in virtue of being predicables, utterances and no more; and the ten categories (substance, quality, and so on) are likewise categories of utterances only—for instance, noun and adjective (*Dialectica*, edited by L. M. de Rijk, p. 3).

The attempt to establish logic as a science of linguistic entities only may be called sermocinalism. During the years 1150 to 1250, when medieval logic was acquiring its distinctive character, sermocinalism held undisputed sway as the philosophy of logic, but it did so in the refined and strengthened form given it in the writings of Abelard. Garland had attempted to make utterances (*voces*) the elements of logic, which he thought of as the science of language. Abelard, recognizing that utterances are physical events that are, as such, of no interest to logicians, replaced the overly simple utterance with what he called the *sermo*, defined as the utterance taken together with its signification. Logicians in the second half of the twelfth century seem to have been unanimous in their adoption of this refinement. An anonymous *Dialectica seu Logica* supported the rejection of utterances as the elements of logic with the following interesting argument, somewhat reminiscent of Aristotle's doctrine of complex names. "Some utterances are significant; some are not.… This division … is exhaustive but seems not to be exclusive, since the same utterance may be both significant and not significant. For example, the utterance 'king' [*rex*] is significant as a word, but since it is also part of a word, a syllable of a word—as in 'smoking' [*sorex*, shrew]—it is in that case and on that account not a significant utterance" (Martin Grabmann, *Bearbeitungen*, Berlin, 1937, p. 30).

Having more precisely identified the elements of logic as linguistic entities, Abelard suggested that logic as the science of language should determine significations on the basis of the application of utterances, determining the proper application of utterances on the basis of the investigations of the natural sciences (*Dialectica*, pp. 286–287). One reason for this suggestion seems to have been his concern with propositions true *gratia terminorum*, analytically true on semantic rather than on syntactic grounds—for instance, "if there is paternity, there is filiation" or "if it is a body, it is corporeal" (see pp. 284–286).

To most medieval philosophers Abelard's emphasis on the importance of signification as well as of utterance might have suggested that mental entities of some sort were to be considered the elements of logic. He explicitly rejected this possibility, however, and in doing so made his most important contribution to sermocinalism. He argued that a proposition true *gratia terminorum* could not be verified by an appeal to the status of mental entities. "When we say 'if it is man it is animal,' if we refer to the connection of the *understanding* of the propositions, as if we were concerned with the *concepts*, there is no truth to the conditional, since the one *concept* may occur entirely without the other" (p. 154). What we are concerned with, Abelard maintained, is the connection between the term *animal* and the definition of the term *man*—namely, the inclusion of the term *animal* within the string of terms making up the definition of the term *man*. As a result of this move, sermocinalism was directed not only against the notion of logic as a science of reality (*scientia realis*) but evidently also against the notion of it as the science of reason (*scientia rationalis*).

The philosophy of logic that eventually challenged sermocinalism concentrated its opposition on this last point. Since it was explicitly drawn from the philosophy of Avicenna, the rival doctrine may conveniently be called Avicennianism. Although as many of Avicenna's writings as were available to the medievals had been translated into Latin around the middle of the twelfth century, Avicennianism as a philosophy of logic seems not to have come into prominence until Albert the Great (1193–1280) adopted it around the middle of the thirteenth century. By that time, however, medieval logic was firmly committed to its distinctive line of development as the *scientia sermocinalis*. As a result, the main impact of Avicennianism as an alternative to sermocinalism was felt less on the work of the logicians than on the metaphysicians' discussions of the nature of logic.

The central doctrine of Avicennianism is presented in the frequently quoted passage from Avicenna's *Philosophia Prima*: "The subject matter of logic, as you know, is intentions understood *secondarily*, which are applied to intentions *primarily* understood" (I, 2, f70vA). Logic was the science of reason, Avicenna claimed, for "the relation of this doctrine [logic] to internal thought, which is called internal speech, is just like the relation of grammar to outward signification, which is called speech" (*Logica* f3rA). Thus grammar, not logic, was the sermocinal science, according to Avicennianism, and the rise of speculative grammar that was to follow may in part be attributed to this point of view.

THE PROPERTIES OF TERMS

Until about the middle of the twelfth century the subject matter of medieval logic was drawn from Aristotle's *Categories* and *De Interpretatione*, together with a set of

books by Porphyry and Boethius that were centered more or less closely on those two books of Aristotle. Later in the Middle Ages this collection of books, or the kind of logic these books contained, became known as *logica vetus*, the old logic. When the remaining four books of Aristotle's Organon began to circulate in western Europe during the twelfth century, they, or their contents, became known as *logica nova*, the new logic. The only completely new kind of material in the *logica nova* was the treatment of fallacy in Aristotle's *Sophistical Refutations*, which excited a tremendous interest in *sophismata*, fallacies resulting from the misuse of or natural ambiguities in various devices of ordinary discourse. Largely because of this lasting interest, medieval logicians of the late twelfth and early thirteenth centuries gradually developed an original logicosemantic inquiry. In order to distinguish this genuinely medieval contribution from Aristotle's contributions to logic, thirteenth-century philosophers began to speak of it as the *logica moderna*, lumping the *logica vetus* and *logica nova* together as *logica antiqua*. Perhaps the *logica moderna* was aimed originally at nothing more than providing ad hoc rules of inference to cover problematic locutions in ordinary discourse, but, although it retained that aim throughout its three-hundred-year history, its principal aim soon became the development of a reasonably general account of the different ways in which words are used to stand for things and to operate on other words.

The earliest known fully developed productions of the "modernist" or "terminist" logicians are the logical treatises of William of Sherwood (d. 1266/1272), Peter of Spain (d. 1277), and Lambert of Auxerre (fl. 1250), evidently written at Paris about the middle of the thirteenth century. At that time the *logica moderna* seems to have been thought of as having two branches, an account of "the properties of terms" (*proprietates terminorum*) and an account of the signification and function of "syncategorematic words" (*syncategoremata*). The two branches naturally differed in detail, but both accounts employed the same principles of explanation and had the same aims. Most nouns, pronouns, verbs, participles, and adjectives were considered to be categorematic words, words capable of serving as terms (that is, as subjects or predicates); and the syncategorematic words were those which can occur in a statement only together with categorematic words. The two branches of the *logica moderna* were thus theoretically exhaustive of the kinds of words occurring in various roles in statements.

The modernists of the thirteenth century regularly recognized four properties of terms: (1) signification—the word's meaning, broadly conceived, or the range of conventional uses of the word (a property of every categorematic); (2) supposition—the conventional interpretation of a word on a particular occasion of its use, a modification of its signification resulting from its syntactic context, if any, and other considerations (a property only of nouns, pronouns, and "substantive expressions," that is, other categorematics employed as substantives and particularly as subjects); (3) copulation—virtually the same as supposition, except that it is a property only of verbs, participles, and adjectives, especially when they occur as predicates; (4) appellation—"the present correct application of a term" (Sherwood), a property only of nouns, adjectives, and participles; for instance, in 2004 Chicago was an appellatum of "city" but Nineveh was not.

Obviously these four properties are not on an equal footing. The supposition, copulation, or appellation of a term was considered a function of its signification; Vincent of Beauvais (d. 1264) even designated signification the genus of which the other three are species (Carl Prantl, *Geschichte der Logik*, Vol. III, p. 83, n. 319). Moreover, copulation and appellation are of distinctly secondary importance. By the middle of the fourteenth century only signification and supposition were regularly recognized as properties of terms, and throughout the history of the *logica moderna* it was the supposition (*suppositio*) of terms on which the inquiry centered. For that reason the best way of quickly acquiring a broad but accurate idea of the modernists' account of the properties of terms is to examine their divisions of supposition. (The recognition and treatment of the divisions of course differed from one modernist to another, but the following selection includes all the major divisions and many of the more interesting minor divisions.)

The supposition of a term was divided initially into *proper* and *improper* supposition. A term had improper supposition when it was used figuratively, and several varieties of improper supposition were distinguished: antonomastic, synecdochic, metaphoric, ironic, and metonymic. The proper supposition of a term was divided into *formal* and *material* supposition, the latter being the use of a term to refer to itself, either as type or as token—for instance, "*man* is a noun," "'*man is an animal*' is a true statement," "*man* is a monosyllable," "*man* has three letters." Formal supposition was *personal* if the term was used to refer to individuals bearing the form signified by the term, *simple* if the reference was to the form itself, as in "*man* is a species." The initial division of personal supposition was sometimes based on the divi-

sion of terms as *common* or *discrete*, depending on the possibility of using them to refer to more than one individual at a time. Thus, the subjects of the statements "*Socrates* is running" and "*that man* is running" have discrete personal supposition.

The portions of supposition theory dealing with the divisions of the personal supposition of common terms were more fully developed than the rest, not only because they were intrinsically more interesting but probably also because they provided the most points of contact between the *logica antiqua* and the *logica moderna*. In those portions of supposition theory, far more than in the others, the emphasis lay on the application of the theory to the evaluation of inferences, especially such inferences as involved Aristotle's four categorical propositions (or near relatives of them) but could not be adequately evaluated within the *logica antiqua*.

A common term was said to have determinate personal supposition when it was used to refer to some one individual without identifying the supposited individual, as in "a *man* is running" or in "some *man* is running." A statement including a "distributive sign" (such as "every" or "no") was, on the other hand, bound to include one or more common terms having *confused* personal supposition, terms used to refer to more than one individual at once or to one individual many times (as in "every man sees a *man*," where the second occurrence of "man" has confused supposition even if it is being used to refer to only one individual). If the confused supposition included each and every individual bearing the form signified by the common term, it was designated *distributive*, as in "every *man* is an animal," "no *man* is an *ass*." If the confused supposition did not plainly include that totality, it was designated *merely confused* (*confusa tantum*), as in "every man is an *animal*," "every man sees a *man*."

The modernists observed that in many cases of distributive confused supposition it was possible to make a "descent" under the term having such supposition, instantiating as in "every man is an animal, therefore this man is an animal"; "no man is an ass, therefore no man is this ass." They described such cases as *mobile* but paid at least as much attention to the *immobile* distributive supposition produced by the use of "exclusives" or "exceptives" together with distributive signs. Thus, in "only every man is running" the distributive supposition of the common term is immobilized by the exclusive "only," so that one cannot infer "therefore only this man is running." It was also recognized that the inclusion of the distributive sign within the scope of the exclusive or exceptive was not always dependent on their relative positions in the statement, since from "every man except Socrates is running" one cannot infer "therefore this man except Socrates is running" (although this unacceptable inference is uninterpretable rather than invalid).

Supposition theory was an attempt to develop a unified treatment of a great number of semantical and logical topics that are still of interest, although now for the most part they are treated in separate inquiries. It is therefore especially intriguing and difficult to discover just what that unifying notion—supposition—amounted to. One broad description that plainly holds good for most of the divisions of supposition is that they are syntax-dependent referential functions of a term's signification. Any description in terms of syntax and semantics will, however, fail to cover all the divisions of supposition and will miss what is distinctive in it. In the case of improper supposition, for example, while the circumstances under which a term is used clearly do determine whether or not it is being used figuratively, it will not do to limit those circumstances to the syntactic context of the term's use.

Again, supposition theorists frequently remarked on the fact that the supposition apparently determined by the syntax often differs from the supposition intended by the framer of the statement. The man who visits his friend's garden and says "this plant grows in my garden" says what is false "with respect to discourse" (*de virtute sermonis*), but the circumstances of his utterance show that he intends to use the word *plant* as in the statement "such a plant grows in my garden." The correct analysis of the supposition of "plant" takes it to have the supposition determined for it *not* by that syntactic context but by the clearly discernible "intention of the framer" (*intentio ponentis*). Finally, among later supposition theorists it was a matter of controversy whether terms occurring in statements written in a closed book had any supposition at all, and those few who held that they did then have supposition seem to have been motivated by a misguided concern that otherwise certain true statements in a closed book, such as "God exists," might cease to be true while the book remained unread.

The consensus of the modernists seems to have been that a term had one or another kind of supposition only on an occasion of its actually being used in referring (or understood to have been used in referring) to some entity or entities, the particular kind of supposition being determined by a number of the circumstances of the occasion and its syntactic context being the most important but not in itself the decisive circumstance.

SYNCATEGOREMATA

Within the *logica moderna* the investigation of syncategorematic words complemented the investigation of the categorematic words under the doctrine of "the properties of terms." Something closely resembling the modernists' notion of syncategoremata seems to have been operative in Aristotle and the Stoics, but the medievals evidently acquired the technical term and the rudiments of the notion directly from Priscian (fl. 500), who had reported that "according to the dialecticians [not identified], there are two [principal (?)] parts to a sentence—the noun and the verb—since they alone and of themselves make a sentence; but they called the other parts *syncategoremata*—that is, consignificants" (*Institutio de Arte Grammatica*, M. Hertz, ed., in H. Kiel, *Grammatici Latini*, Leipzig, 1855, Vol. II, p. 54).

Interest in the syncategoremata as such began in connection with the twelfth-century interest in fallacies of ambiguity as it became plain that the crucial ambiguity was often located elsewhere than in subjects and predicates. The grammatical basis of distinction provided by Priscian seems to have been adopted at first and occasionally even narrowed so that only the "indeclinables"—prepositions and conjunctions—were considered to be syncategoremata (see the anonymous late twelfth-century *Fallacie Parvipontane*, edited by L. M. de Rijk, in *Logica Modernorum*, Vol. I, p. 559; also see Abelard, *Dialectica*, edited by L. M. de Rijk, pp. 118–121). The notion of syncategoremata that became important in the *logica moderna*, however, was not founded on a strictly grammatical distinction. Abelard's treatment of "alone" in "a man alone is capable of laughter" (*Logica "Ingredientibus*," edited by B. Geyer, p. 483) prefigured the pattern that was to be followed by the modernists. He pointed out that if "alone" is taken to be part of the subject, the statement is about a man who happens to be by himself, while if "it is attached to *the predication* it denies the capacity for laughter to all non-men, as if to say: 'a man is capable of laughter in such a way that nothing else is capable of laughter.'"

When the investigation of syncategorematic words appeared as a separate inquiry in the treatises on syncategoremata by William of Sherwood and Peter of Spain (first half of the thirteenth century), the distinguishing characteristic of syncategorematic words was the fact that they had some effect on the relation between the categorematic words—that is, the predication or the "composition"—or on the relation between two predications or compositions. Thus Sherwood's inventory of syncategoremata, most of which became standard, included the verbs "begins" and "ceases" and the noun "nothing" as well as grammarians' syncategoremata such as "every," "both," "except," "alone," "is," "not," "necessarily," "if," "unless," "and," and "or." The standard syncategoremata, then, cannot be completely described as a selection of logical operators, although they plainly included such a selection. Nor was the investigation of them aimed primarily at uncovering their strictly formal properties. Many of the rules put forward in connection with one or another syncategorema were rules of inference—such as "When there are two distributions over one and the same part of a locution, the first immobilizes the second" (Sherwood)—but just as many were semantic rules—such as "The sign 'every' or 'all' requires that there be at least three appellata [for the term to which the sign is attached]" (Sherwood)—and there seems to have been no clear distinction drawn between the two sorts of rules.

The modernists' treatises on syncategoremata presupposed the doctrine of the properties of terms, as is shown by the rules given just above, and much of their discussion of the function of such words is in terms of the various modifications of supposition produced within the scope of one or another syncategorema. (The problem of determining the scope of syncategoremata, especially in contexts including more than one, was particularly important in these investigations.) In this way the syntactic and semantic questions about syncategoremata were essentially interconnected. Sherwood at least among the older modernists was sometimes concerned to discuss the *signification* of syncategoremata—"every," for example, was said by him to signify universality—but that seems to have been a feature of his unusual doctrine that in order to be significant a word had to signify some *form*. Most writers on syncategoremata took up Priscian's really unjustified translation of the Greek *syncategoremata* as "consignificants" and used it as the basis for their view that "strictly speaking, a syncategorema signifies nothing, but when added to another word it makes that word signify something, or makes it suppose for something or some things in some definite way, or exercises some other function having to do with a categorema" (William of Ockham, *Summa Logicae*, I, 4; cf. the remarks of John of Salisbury in *Metalogicon*, Book I, Ch. 16).

The initial impetus to the study of syncategoremata came from the twelfth-century interest in fallacies, and the investigation continued to be associated with fallacies or with sophismata throughout its development. Observations about syncategoremata were only incidental in the twelfth-century treatises on fallacies, but the novel

emphasis of the *logica moderna* on the syncategoremata themselves is evident in the development of a new sort of treatise—"On Exponibles"—in the first half of the thirteenth century. "An exponible proposition is a proposition having an obscure sense that stands in need of exposition because of some syncategorema located in it explicitly or implicitly or in some word" (*Tractatus Exponibilium*, doubtfully ascribed to Peter of Spain, in *The Summulae Logicales of Peter of Spain*, edited by J. P. Mullally, p. 104). The exponibles did not involve fallacious arguments nor were they, strictly speaking, ambiguous. They were simply subjects for analysis, an analysis that was to explicate the force of some syncategorema in some particular context. For example, "'Man *inasmuch as* [*inquantum*] he is rational is not capable of braying'—that is, [1] no man is capable of braying and [2] every man is rational and [3] no rational entity is capable of braying and [4] because an entity is rational it is not capable of braying" (p. 115).

SOPHISMATA. The sophismata that played an increasingly important role in investigations of syncategoremata from the thirteenth through the fifteenth centuries may be characterized as falling somewhere between fallacies of ambiguity and exponible propositions. In the independent treatises on syncategoremata prevalent in the thirteenth century, the sophismata served as the illustrations of the principles uncovered in the investigation and characteristically took the form of an assertion (the sophisma proper) followed by a proof, a counterargument, and an adjudication of the apparent paradox by an appeal to the principles. For example, "Suppose that exactly one individual of each species of animal is running. Then [*a*] every animal is running. (Proof: a man is running; a lion … ; a goat … ; and so on with respect to the individuals; therefore every animal is running.) But [*b*] every man is an animal; therefore every man is running. [Solution:] [*a*] is ambiguous since [because of 'every'] the word 'animal' can distribute for the remote parts (or the individuals belonging to genera)—in which case it is false, since it is then distributed for all its individuals—or for the genera of the individuals (or for the proximate parts)—in which case the minor [*b*] is plainly not accepted" (William of Sherwood, *Syncategoremata*, edited by J. R. O'Donnell, in *Medieval Studies*, Vol. III, p. 49).

The continuity of the development of the *logica moderna* was enhanced by the fact that from the twelfth century through the fifteenth century the same sophismata were treated from varying points of view, but at the same time the number and intricacy of the sophismata were constantly increasing. As a result the modernists of the fourteenth century frequently produced treatises titled *Sophismata* in which large numbers of them were grouped according to the syncategoremata at issue in them, and the investigations that had begun in separate treatises on syncategoremata were pursued in the *Sophismata* and *Exponibilia* of the late Middle Ages.

SPECULATIVE GRAMMAR

The notion that grammar and philosophy were intimately related was one of the most pervasive of the assumptions that determined the character of medieval thought. It is probably to be explained by the facts that grammar was one of the very few inquiries to survive antiquity intact and that the only ancient philosophy available during the early Middle Ages was Aristotle's *Categories* and *De Interpretatione*, works of a decidedly grammatical cast. The usual view of the connection between the two subjects was the one expressed most memorably by John of Salisbury—"Grammar is the cradle of all philosophy" (*Metalogicon*, Book I, Ch. 13)—and the *logica moderna*, by far the most impressive medieval contribution to semantics, is a clear example of the influence of grammar on philosophy. The influence ran the other way, however, in the development of "speculative grammar" (*grammatica speculativa*), or the doctrine of the "modes of signifying" (*modi significandi*), a movement that began somewhat later and subsided somewhat earlier than the *logica moderna*. Although there were some connections between the two movements—for instance, Roger Bacon (1214/1220–1292), one of the first of the speculative grammarians, or "modists," also contributed to the *logica moderna*—they tended to be mutually independent and to some extent theoretically opposed developments in the history of semantics.

The most important single factor in the rise of speculative grammar in the early thirteenth century was the enthusiasm for the notion of a *science*, then being rediscovered in the *Posterior Analytics* of Aristotle and in his Arabic commentators. For a time it was the aim of every study to achieve the status of an Aristotelian science, a body of necessary knowledge deductively demonstrated, and two facts seemed to stand in the way of certifying grammar as a science. For one thing, as it had been presented by Priscian and Donatus, grammar was simply a set of observations about correct constructions without any attempt at explanation of the correctness; but only knowledge "by causes" qualified as scientific. For another, even Peter Helias (fl. 1150), who in his com-

mentaries on Priscian had been the first medieval to attempt explanations of grammatical facts, had maintained that there were as many grammars as there were languages; but a unified subject matter was a prerequisite of a science.

Thus a science of grammar was not to be found in the grammatical authorities, and it seemed one never would be found as long as grammar was conceived of as something to be discerned only in the investigation of actual languages. Robert Kilwardby (d. 1279) set the stage for speculative grammar when he argued that "since a science remains the same for all men and its subject matter remains the same, the subject matter of grammar must remain the same for all men. But grammatically ordered speech or articulate utterance that can be put into a grammatical pattern is *not* the same for all men, and for that reason it will not be the subject matter of grammar [as a science]" (Commentary on Book I of Priscian's *Ars Minor*). No science of languages was possible, but grammar might become the science of language, the *scientia sermocinalis*, if the variable external trappings were ignored and one concentrated on the conceptual underpinnings—which, as Aristotle had pointed out, *were* the same for all men. Thus Roger Bacon was led to proclaim that "with respect to its *substance* grammar is one and the same in all languages, although it does vary *accidentally*" (*Grammatica Graeca*, edited by E. Charles, p. 278), and this became the often repeated fundamental assumption of the speculative grammarians.

As it developed, speculative grammar took the form of an attempt to provide an Aristotelian ontology of language, finding analogues in the various parts of speech for matter, form, substance, process, and so on. As Siger of Courtrai (d. 1341) put it, "grammar is the *scientia sermocinalis*, which considers discourse and its properties [*passiones*] in general for the purpose of expressing principally concepts of the mind by means of interconnected discourse" (*Summa Modorum Significandi*, edited by G. Wallerand, p. 93). Siger then cited Avicenna for the Aristotelian doctrine that concepts are the same for all men because they are the result of experiencing extramental entities, which are the same for all men. Thus, the ontology that applies to the extramental entities must apply as well to the concepts derivative from them (if they adequately copy the extramental entities) and, in turn, to the discourse employed to express those concepts (if it is to be adequate for that purpose). "Therefore *modes of being*, or properties of things …, precede a *mode of understanding* as a cause precedes an effect" (ibid.). In the same way a mode of designating (*modus signandi*) follows a mode of understanding, "since a thing is understood and also conceived of before it is designated by means of an *utterance* [*vox*], for utterances are signs of passions, as is said in *De Interpretatione*, Ch. 1" (p. 94). When the understanding assigns a given concept to an utterance, the merely physical utterance becomes a word (*dictio*).

Up to that point the semantic theory underlying speculative grammar might fairly be described as a technical restatement of Aristotle. It was only with the introduction of its "modes of signifying" that the theory acquired its novelty and notoriety. (It was repeatedly attacked and ridiculed by logicians and grammarians of the late Middle Ages and even more strongly assailed by the Renaissance humanists.) As an utterance becomes a word by means of a mode of designating, so a word becomes one or another part of speech by means of a mode of signifying. The modes of signifying, however, are not modes of the utterance of the word but are "certain concepts of the understanding itself" (ibid.). The kind of concept in question seems to be one that links the word to some Aristotelian mode of being. Thus, the kind of concept involved in the mode of designating is the kind that supplies a *significatum* for the utterance "horse," transforming it from a mere sound into a word, while the mode of signifying consists in the recognition that it is substance that is signified by the word *horse*. And when the understanding adds to that general mode of signifying—substance—the specific mode of signifying—quality—then *horse* has been transformed in turn from a mere word to a substantive and from a mere substantive to a noun (pp. 94–95). Along these same lines, the utterance *horse* will eventually be accounted for as a common concrete noun, and similar patterns of modes of signifying are invoked in order to account for the other parts of speech.

Aristotelian ontology was employed by the speculative grammarians not only as the link between grammatical forms and modes of extralinguistic being but also as a picture of intralinguistic relations. Thus, verbs stood at the pinnacle of the linguistic microcosm because just as the other animals are submissive to man, so the inflections of the other parts of speech in a sentence are ultimately submissive to the verb. The infinitive of a verb, however, was analogous to primary matter in substances. And just as the organisms capable of fewest adaptations are ranked lowest in the kingdom of nature, so the indeclinables, the syncategoremata, are the most inferior parts of speech (pp. [52]–[54]).

THE RENAISSANCE AND ENLIGHTENMENT

SEMANTICS, LOGIC, AND EPISTEMOLOGY

As the Middle Ages gave way to the Renaissance in the late fifteenth century, logic (on which semantics had been centered) first lost its medieval attainments and then subsided into inactivity until the middle of the nineteenth century. What little there was in the way of logical inquiry from about 1450 to about 1850 was carried on under the view of logic as the art (or science) of reason, the idea of *scientia sermocinalis* having been ridiculed into oblivion by the Renaissance humanists. Aside from the work of late Scholastics, such as the *Ars Logica* of John of St. Thomas (1589–1644), and an occasional deliberate attempt at revival, such as the *Logica Fundamentis Suis a Quibus Hactenus Collapsa Fuerat Restituta* (1662) of Arnold Geulincx (1624–1669), there were no further developments of the logicosemantic theories of the *logica moderna*.

Philosophers retained their interest in semantics, however, after losing interest in and even all knowledge of the kind of logic with which it had been associated. Epistemology dominated the philosophy of the Renaissance and the Enlightenment (for present purposes, roughly 1500–1800) as logic had dominated medieval philosophy, and the development of semantics during this period centered on epistemology. As a consequence, much of the development took place in the context of discussions of nonlinguistic signs, such as representative ideas, and will not be directly considered here.

Perhaps partly because logic had lost its identity as an inquiry into language, the interest of philosophers in language was more intense and diversified during this period (and especially in the eighteenth century) than at any earlier time. Some of this interest was manifested in widespread speculation about the origin of language and in projects for a universal language or a "real characteristic." Although works on these subjects are typical of the period and often contain material of value for the history of semantics, they can be considered here only as they bear directly on a theory of meaning or philosophy of language selected for discussion.

BACON

Francis Bacon (1561–1626) produced comparatively little that can be described as philosophy of language, but the occasional novel insights and the programmatic character of what he did produce helped to give it a considerable influence over philosophy of language in the Enlightenment. Almost everything of his that is relevant to the history of semantics is to be found in the "Art of Elocution or Tradition" (in the *Advancement of Learning* and the *De Augmentis Scientiarum*) and the doctrine of the "Idols of the Market Place" (in the *Novum Organum* and the *De Augmentis Scientiarum*).

The first of these is plainly Bacon's revised version of the medieval *trivium*—grammar, logic, rhetoric—although he nowhere says so. In the later Middle Ages these subjects had sometimes been designated the *artes sermocinales*, and in the *De Augmentis Scientiarum* Bacon said that the subject matter of the *ars tradendi* was *sermo*. This inquiry into "tradition"—that is, discourse or communication—had three branches, concerning "the organ," "the method," and "the illustration" of tradition; and most of the work of the three branches was explicitly associated with grammar, logic, and rhetoric, respectively. For present purposes the first of these three branches is much more important than the other two.

In his scheme of "Human Philosophy" the Art of Tradition occurred as "the fourth kind of Rational Knowledge" (Spedding, Ellis, and Heath, eds., 3.383–4), because reason was "as it were the soul of discourse," according to Bacon. "Nevertheless, in treating of them reason and discourse ought to be separated, no less than soul and body" (1.651). He began his separate treatment of discourse by identifying speech and writing as the most familiar organs of discourse and stressing their connection with reason by citing with approval the traditional version of Aristotle's doctrine: "Words are the images of cogitations, and letters are the images of words" (3.399; but cf. 3.284, 3.85–86). But his interest in less familiar organs of discourse prompted him to frame a set of general conditions for an organ of discourse: "Whatever can be broken down into differences sufficiently numerous for explicating the variety of notions (provided those differences are perceptible to sense) can become a vehicle of cogitations from one man to another" (1.651; cf. 3.399). An organ of discourse can be used to *communicate* nothing but notions, but it will contain elements that *express* not only notions but also things.

In the most familiar arrangement of organs of discourse, words (by which Bacon meant only articulate sounds [2.411–412]) are expressed by letters—that is, phonograms. Letters, in turn, may be expressed by ciphers—that is, cryptograms—and both letters and ciphers may be designated "nominal characters." But he recognized another kind of "notes of things, which signify

things *without* the aid or intervention of words," either "on the basis of congruity" or "arbitrarily." As examples of the former sort he cited hieroglyphics and gestures, gestures being "transitory hieroglyphics," the "words" for which hieroglyphics may be the "letters," and he classified them together as "emblems"—that is, sensible images to which intellectual conceptions could be reduced by analogy (1.652–653; 649). As examples of the latter sort he cited "real characters" such as Chinese ideograms, which "have nothing emblematic in them, but are simply surds, no less than the elements of letters themselves; … there ought to be as many of them as there are radical words" (1.653).

Despite that disadvantage, real characters could and, Bacon thought, did function as an organ of discourse beyond the limits of a single natural language just because they signified "things and notions" without the intervention of words (1.652). Although he was convinced that there were no more convenient organs of discourse than words and letters, Bacon listed the study of the notes of things among his desiderata (1.653). Acting on this suggestion, the Royal Society commissioned some of its members to look into the project of a universal real character, the eventual result being John Wilkins's *Essay towards a Real Character and a Philosophical Language* (1668), one of many such attempts during this period.

As another part of the inquiry into organs of discourse Bacon proposed a "philosophical grammar," and this desideratum likewise had an extensive but un-Baconian influence. Some of what he had to say about philosophical grammar was reminiscent of the medieval speculative grammar—for instance, it was to be "a kind of grammar that would carefully inquire not into the analogy of words to one another, but into the analogy between words and things or words and reason" (1.654)—and this is what seems to have caught the imagination of his many successors in the Enlightenment who produced works in philosophical or "universal" grammar. What Bacon really had in mind was probably something more nearly like the comparative philology characteristic of the nineteenth century: "But the noblest kind of grammar would, I think, result if someone well taught in many languages, learned as well as vulgar, would treat of the various properties of *languages*, showing in what respects each excels and in what respects it is deficient" (ibid.; cf. 3.230, 3.401). He did, however, go on to suggest that one might combine all the best properties uncovered in that analysis into "a very finely formed image and remarkable model of speech itself for expressing the mind's meanings aright" (1.654).

In his sketch of a philosophical grammar Bacon emphatically disapproved of what he believed Plato had been attempting in the *Cratylus*, an inquiry into "the imposition and original etymology of names" (ibid.; cf. 3.531), but his own concern in the doctrine of the "Idols of the Market Place" closely parallels Plato's real concern in the *Cratylus*, that is, distinguishing between correct and incorrect names. "The idols imposed on the understanding through words are of two kinds. Either they are names of things that are not (for just as there are things that lack a name because they have not been observed, so there are names that lack things, resulting from a fantastic supposition); or they are names of things that are, but confused, ill-defined, and rashly and irregularly abstracted from the things" (1.171). As an example of the first he gave "prime mover"; his example of the second kind was "humid," which, as his discussion of it shows, is less objectionable on these grounds now than it was in seventeenth-century English.

HOBBES

Thomas Hobbes (1588–1679) conceived of his systematic philosophy as beginning with an investigation into language and produced different versions of the investigation in *Human Nature* (1650), Chapters 5 and 13; *Leviathan* (1651), Chapters 4–7; and *Elementa Philosophiae Sectio Prima: De Corpore* (1655; English 1656), Part I, "Computatio Sive Logica." (The latest of those versions is also in most respects the fullest and is used as the basis of the following account.)

Philosophy, Hobbes observed, depends on ratiocination, or "computation" (Molesworth edition, 1.3). In reasoning regarding particular things "we add and subtract in our silent thoughts, without the use of words" (ibid.; see 3.32); but in most instances, and certainly in philosophizing, "men owe all their true ratiocination to the right understanding of *speech*" (1.36), such ratiocination being "nothing but reckoning, that is adding and subtracting, of the consequences of *general names*" (3.30). In the second chapter of his *Logic*, devoted specifically to "names," Hobbes produced a novel combination of several elements in the Aristotelian-Scholastic account of the semantics of names. Ratiocination of every kind depends on memory, and the intelligent use of memory requires what Hobbes called "*marks*, namely, sensible things taken at pleasure, that, by the sense of them, such thoughts may be recalled to our mind as are like those thoughts for which we took them" (1.14).

It is possible, Hobbes thought, for a man to "spend all his time partly in reasoning and partly in inventing

marks for the help of his memory, and advancing himself in learning"—that is, to devise and profitably use a private language—but if science and philosophy are to develop, there must be "certain *signs* by which what one man finds out may be manifested and made known to others" (1.14). Signs that do "signify the cogitations and motions of our mind" are "*words* so and so connected," or what Hobbes called "*speech*, of which every part is a *name*" (1.15). The use of names as marks, he held, was logically prior to their use as signs, since "names, though standing singly by themselves, are marks …; but they cannot be signs otherwise than by being disposed and ordered in speech" (1.15). He recognized the syntactic disposition of names in speech as necessary but not sufficient for "declaring our conceptions to others." Speech cannot "perform that office alone without the help of many circumstances," such as "time, place, countenance, gesture, the counsel of the speaker" (2.274). We must "consider the drift, and occasion, and contexture of the speech, as well as the words themselves" (4.23).

When names are ordered in speech so as to be signs rather than marks, "it is manifest they are *not* signs of the things themselves" but signs only of our conceptions (1.17). Hobbes seems to have been following Aristotle's lead here, but more faithfully than most, since he went on to say, "That the sound of this word 'stone' should be the sign of a stone, cannot be understood in any sense but this, that he that hears it collects that he that pronounces it thinks of a stone" (1.17). Thus, even though indirectly and only in virtue of signifying that the speaker is thinking of a stone, the name "stone" ordered in speech *is* a sign of a stone. At any event, Hobbes nowhere suggested that "stone" occurring in speech was a *name* of some mental entity. On the contrary, in going on to show that "it is not at all necessary that every name should be a name of some thing," Hobbes began by pointing out that " 'man,' 'tree,' 'stone' are *names* of things themselves" (1.17), though they may be used as *signs* of our conceptions of men, trees, and stones and as *names* of "fictions and phantasms of things," such as images in dreams. "Moreover, that which neither is, nor has been, nor ever shall, or ever can be, has a name, namely, 'that which neither is, nor has been,' &c.; or more briefly this, 'impossible'" (1.17). For "a name is not taken in philosophy, as in grammar, for one single word, but for any number of words put together to signify one thing" (1.23), Hobbes having decided "to apply the word 'thing' to whatsoever we name; as if it were all one whether that thing be truly existent, or be only feigned" (1.18).

Much of Hobbes's investigation of names was presented in the form of discussions of traditional classifications of names. His treatment of them sometimes presents the half-understood remnants of complex medieval theories—for instance, his treatment of names of first and second intention (1.20–21)—but there are occasional interesting novelties as well. In his discussion of common and proper names he put forward his strict nominalism: "this word 'universal' is never the name of any thing existent in nature, nor of any idea or phantasm formed in the mind, but always the name of some word or name" (1.20); at another point he remarked that the univocal-equivocal distinction "belongs not so much to names as to those that use names" (1.23); and he based the distinction between simple and compound names not on appearances but on considerations of analyzability, so that in the context of a discussion of man "body" is a simple name while "man" is a "more compounded name," being equivalent to "animated rational body" (1.23–24).

Hobbes encountered important difficulties in his discussion of names of "certain and determined" and of "uncertain and undetermined" signification (1.21–23), which is evidently a badly distorted remnant of supposition theory. In the course of that discussion Hobbes was led to claim, for example, that *particular* names—such as "some man"—"are of *uncertain* signification, because the hearer knows not what thing it is the speaker would have him conceive" (1.22), as if the "uncertainty" in, say, "some man will marry my daughter" were the sort that could always be resolved by asking the speaker "*which* man?" Even worse confusion resulted from his attempt to show that such quantifiers as "every" and "some" were unnecessary for purposes of reasoning. Such words, he maintained, "which denote universality and particularity, are not names; so that 'every man' and 'that man which the hearer conceives in his mind' are all one; and 'some man' and 'that man which the speaker thought of' signify the same. From whence it is evident, that the use of signs of this kind, is not for a man's … getting of knowledge by his own private meditation (for every man has his own thoughts sufficiently determined without such helps as these) but … for the teaching and signifying our conceptions to others" (1.22).

In his treatment of propositions Hobbes sometimes spoke as if only such propositions as "Cicero is Tully" were true—for instance, "that proposition only is true in which are copulated two names of one and the same thing" (1.57)—but usually his description of a true proposition was more moderately and more accurately expressed along such lines as these: "A *true* proposition is

that, whose predicate contains, or comprehends its subject, or whose predicate is a name of every thing, of which the subject is a name" (1.35; cf. 4.23–24). He produced a detailed analysis of falsity as reducible to combinations of names of different sorts of entities (1.57–62). His truth theory was, however, quite radical in other respects. The "first truths," he claimed, "were arbitrarily made by those that first of all imposed names upon things, or received them from the imposition of others. For it is true [for example] that *man is a living creature*, but it is for this reason, that it pleased men to impose both those names on the same thing" (1.36). This suggests an identification of the proposition with a particular sequence of words, but Hobbes elsewhere gave the impression of having been on the point of drawing a clear distinction between propositions and the vehicles of their expression—for instance, "every proposition may be, and uses to be, pronounced and written in many forms.… And therefore, whensoever they [students of philosophy] meet with any obscure proposition, they ought to reduce it to its most simple and categorical form" (1.39).

Hobbes rejected the analysis of contingent categorical propositions into their corresponding hypothetical forms, pointing out that while this analysis was allowable for *necessary* categoricals, "in contingent propositions, though this be true, 'every crow is black,' yet this, 'if any thing be a crow, the same is black' [i.e., '$(x)(Cx \supset Bx)$'], is false" (ibid.). In several places Hobbes discussed the various uses of speech—for example, *Human Nature*, Chapter 13—and at one point argued against the notion that a promise simply by its form of words creates an obligation (2.18–20).

THE *PORT-ROYAL LOGIC*

René Descartes (1596–1650) had very little to say about language, but Antoine Arnauld (1612–1694) took an avowedly Cartesian approach to semantic questions in the *Port-Royal Grammar* (*Grammaire générale et raisonnée*, with Claude Lancelot, 1660) and the *Port-Royal Logic* (*Logique ou l'art de penser*, with Nicole, 1662). The latter book had a tremendous influence; it marks, better than any other, the abandonment of the medieval doctrine of an essential connection between logic and semantics. Disdain for medieval theories was emphatically expressed in it—"No one, thank God, is interested in … second intentions" (*Premier Discours*)—but at several points the theories were still employed (for instance, in 1.2 and 2.10) and elsewhere in the book they were supplanted by innovations that sometimes obscured what had been clear in the *logica moderna*.

In words reminiscent of Hobbes's on this point, Arnauld remarked that if logic considered only an individual's reflections on his ideas, the investigation of language would form no part of it. But we must use "exterior signs" for communication, "and since this custom is so strong that even when we think by ourselves things are presented to our mind only together with the words with which we are accustomed to adorn them in speaking to others, it is necessary in logic to consider the ideas joined to words and the words joined to ideas" (introduction; cf. Descartes, *Principles*, Part I, Principle 74). Arnauld of course argued (1.1) against Hobbes's anti-Cartesian suggestion that reasoning might be "nothing more than the uniting and stringing together of names or designations by the word 'is,'" so that all we could ever conclude is "whether or not there is a convention (arbitrarily made about their meanings) according to which we join these names together" (*Objections to Descartes's Meditations*, 3.4).

Signs and signification were frequently discussed in the *Port-Royal Logic*, sometimes with interesting results; the most fundamental questions were, however, treated with the kind of inattention to detail that came to characterize most of the many semantic theories of the Enlightenment. "The sign," said Arnauld, "comprises two ideas—one of the thing that represents, the other of the thing represented—and its nature consists in exciting the second by means of the first" (1.4). In the case of words, the "thing represented" was identified as a "thought" or an "idea." Even proper names were defined as those "that serve to mark … the ideas that represent only one single thing," and "general words" were said to be those "that are joined to universal and general ideas" (1.6). The doctrine is so far consistent and recognizably Cartesian, even if crude. But it is complicated, no doubt inadvertently, by many suggestions of a different sort of signification for words. Thus, on a single page Arnauld began by calling words "sounds that are intended to signify *ideas*," went on to speak of "things and modes" as "the *objects* of our thoughts," and ended by defining names as "the words intended to signify both *things and modes*" (2.1); and he nowhere provided an account that might justify this extended use of "signify." He may have been assuming a transitivity of signification—words signifying ideas representing things—but John Locke was the first to attempt to spell out such a theory.

Arnauld warned against the "great equivocation in the word 'arbitrary' when we say that the signification of words is arbitrary," pointing out that while "it is purely arbitrary to join one idea to one sound rather than to

another," nevertheless the ideas, "at least those that are clear and distinct," are not arbitrary. The result of correct reasoning is "a solid, effective judgment regarding the nature of things based on the consideration of the ideas of them a man has in mind, which ideas it has pleased men to mark by means of certain names" (1.1). But, on the other hand, one of his reasons for rejecting Aristotle's categories was that they were "arbitrary names that form no clear and distinct idea in the mind" (1.3).

Arnauld also explicitly rejected Aristotle's definition of a verb, putting in its place one that not only captured the essence that Aristotle missed but was also much simpler: "a word that signifies affirmation" (2.2). Evidently he did not mean that it signified the idea of affirmation (as would the noun "affirmation"), but he did not work out the definition in a way that tied it to the rest of his signification theory.

Like so many other philosophers of the period, Arnauld believed that "the best means of avoiding the confusion of words to be found in ordinary languages is to make a new language and new words that would be attached only to the ideas we want them to represent." He differed from most, however, in suggesting that this be accomplished simply by a conscientious, systematic use of precise nominal definitions attached to already extant vocables of ordinary languages (1.12).

One of the more interesting notions in Arnauld's doctrine of signification was introduced in his observation that "it often happens that besides the principal idea (which is regarded as its *proper* signification) a word excites several other ideas that may be called *accessory*." Sometimes the accessory ideas are attached to the words "as the result of common usage," as in "you lied," the *proper* signification of which is the idea that you knew the contrary of what you said, the ideas of contempt and outrage being *accessory* (1.14). In some respects this is reminiscent of Augustine's doctrine in *Principia Dialecticae*, especially when Arnauld uses it to argue (against Cicero) that certain words may, in virtue of their accessory ideas, be described as unchaste (1.14). Accessory ideas may also be attached for the purpose of a single use of a word, and on that basis Arnauld attempted an explanation of the varying signification of the demonstrative pronoun "this," here as elsewhere in the book applying his semantic doctrines to the elucidation of the formula of transubstantiation—"this is my body" (1.15).

Arnauld's notion of accessory ideas might have been (but was not) used to advantage in his discussion of problems of identity of reference, where he argued that when the mind frames the proposition "that Rome, which was of brick before the time of Augustus, was of marble when he died, the word 'Rome,' which appears as only one subject, nevertheless marks two subjects that are really distinct but reunited under a confused idea of Rome that prevents the mind from perceiving the distinction of subjects" (2.12). The suggestion is that the proposition should be rejected by anyone having a clear and distinct idea of Rome, which is preposterous. Even if the *proper* signification of "Rome" was taken to be only the idea of buildings, surely such *accessory* ideas as location, population, and institutions could have been invoked to warrant the continuing use of the single proper name.

Like Hobbes, Arnauld recognized complex terms expressed in a single word, but instead of Hobbes's criterion of analyzability Arnauld employed the notion of accessory ideas attaching to the word under certain circumstances. Thus, the term *king*, which is simple "*in expression*," was "a term complex *in sense*" when uttered in seventeenth-century France, "because in pronouncing the word 'King' we not only have in mind the general idea corresponding to that word; we also mentally join to it the idea of Louis XIV, who is now King of France. There is an infinity of terms in ordinary human discourse that are complex in this respect" (1.8).

Arnauld's analysis of the semantics of sentences clearly illustrates the importance of the loss of supposition theory. In one badly confused but typical passage he claimed that "when one says that men are animals, the word 'animal' no longer signifies all animals, but only the animals that are men" (2.17). Not only does this transform predication into identity, it also violates his own doctrine of signification. Again, in discussing "some man is just" Arnauld maintained that "just" there "signifies only the justice that is in some man," the result being that "some man is identified with some just [thing]" (2.18).

A complete chapter of the *Port-Royal Logic* is devoted to the discussion of propositions such as "this is Alexander" (pointing to his portrait), which he described as "expressions in which one uses the name of the thing to mark the sign," seldom if ever causing any difficulty in actual use, and propositions such as Joseph's explanation of Pharaoh's dream—"the seven full sheaves are seven full years of plenty"—"expressions in which, the sign being marked by its own name or by a pronoun, one affirms of it the signified thing." One result of this novel approach to metaphor is his formulation of a rule governing the appropriateness of the second sort of proposition: "the mind of those to whom one speaks must already regard the sign as a sign and be concerned to know of what it is a sign" (2.14).

"COMPREHENSION" AND "EXTENSION." Certainly the most influential semantic doctrine of the *Port-Royal Logic* was Arnauld's introduction of a distinction between the "comprehension" and the "extension" of a term. (In the nineteenth century Sir William Hamilton renamed the former "intension.") The principle of such a distinction had been employed in the medieval distinction between simple and personal supposition, and even the distinction itself had occasionally been anticipated in one form or another (for instance, by Cajetan), but Arnauld's introduction of it seems to have been original and certainly was the first instance of a systematic use of it.

It is difficult, however, to say exactly what Arnauld intended by the distinction, for its exposition is obscured by his generally confused account of signification. He first advanced the distinction as one pertaining to "universal ideas" (or terms). "I call the *comprehension* of the idea the attributes it comprises in itself that cannot be removed from it without destroying it, as the comprehension of the idea of the triangle comprises extension, figure, three lines, three angles, the equality of those three angles to two right angles, etc. I call the *extension* of the idea the subjects with which that idea agrees, … as the idea of triangle is extended to all the various species of triangle." And Arnauld went on to say that the idea could be restricted in its extension by "applying it to only some of the subjects with which it agrees, without thereby destroying it," for example, by attaching to it "an indistinct and indeterminate idea of a part, as when I say 'some triangle'" (1.6). If, however, the extension consists of species and not of the individuals, which is what Arnauld maintained, then such a device for restricting extension is always to be read as "some (species of) triangle," which produces an absurdity. Because of his theory of signification there would be theoretical difficulties for Arnauld in simply identifying the term's extension with the individuals in question, but for the most part he seems to have had that identification in mind rather than the one he laid down.

Individual terms, too, were said to have comprehension and extension. In the phrase "Julius Caesar, the greatest commander the world has ever seen," the comprehension of that individual term is "explicated" in one of countless possible ways. But the extension of an individual term cannot be restricted, Arnauld maintained, and thus every singular proposition is universal (1.8, 2.3).

LOCKE

In the third book of his *Essay Human Understanding* (1690) John Locke (1632–1704) produced the first modern treatise devoted specifically to philosophy of language. No work had a greater influence over the development of semantics during the Enlightenment than did Book III of this work, "Of Words"; yet its semantic theories were neither novel in principle nor clearly and thoroughly developed. To go no further back, many of its principles had been anticipated in Kenelm Digby's *Two Treatises* (1664), in Richard Burthogge's *Organum Vetus et Novum* (1678), and in Hobbes's works. Of course Locke's "Of Words" acquired importance simply by being a part of the enormously influential *Essay*, but the source of its special influence lay in the fact that Locke had expressly connected semantic inquiry with theory of knowledge. He had set out to investigate "our knowledge," and along the way he found himself unexpectedly compelled to investigate "the force and manner of signification" of words (3.9.21), having discovered that "there is so close a connexion between ideas and words … that it is impossible to speak clearly and distinctly of our knowledge, which all consists in propositions, without considering, first, the nature, use, and signification of Language" (2.33.19). The new epistemological orientation of semantic inquiries, apparent even in the logic books of the period, was first explicitly established in Locke's *Essay*.

Locke evidently thought of the material of Book III as serving two purposes in his philosophy. On the one hand, he characterized his new "way of ideas" as nothing more than "the old way of speaking intelligibly" (third letter to Stillingfleet), which he reduced to a few commonsensical maxims for the avoidance of "jargon," very much in the spirit of Bacon's treatment of the "Idols of the Market Place." The semantic theory in Book III was developed in part as a support for these "remedies of the … imperfections and abuses of words" (3.11), and Locke's preoccupation with that practical aim may help to explain some of the imprecision and inconsistency in his theoretical statements. He did, however, clearly recognize a more strictly theoretical purpose in the semantic inquiries of Book III, one which he summarized in his description of the third branch of science—"*Σημειωτικὴ*, or *the doctrine of signs*" (4.21.4), the consideration of ideas as the signs of things and of words as the signs of ideas.

Locke's account of words as the signs of ideas shows little of the sensitivity to the complexities of language that had characterized the work of many of his predecessors, including Hobbes. Except for one very short, cryptic chapter on "particles" (by which he evidently meant syncategorematic words but perhaps also verbs), the seman-

tics of words in Book III is exclusively a semantics of "names"—names of "simple ideas," of "mixed modes," and of "natural substances"—with no suggestion that anything has been left out of consideration (3.4.1).

The development of his fundamental thesis regarding the signification of these names begins with his observing that "words being voluntary signs, they cannot be voluntary signs imposed by [a man] … on things he knows not." Now what a man knows is in his mind, but all that is in a man's mind is his own ideas. Therefore, "words, in their primary or immediate signification, stand for nothing but *the ideas in the mind of him that uses them*" (3.2.2). This is not markedly different from the starting point of many earlier semantic theories, but Locke's initially uncompromising development of it led to some extreme consequences. Men, he observed, "suppose their words to be marks of the ideas also of other men" or "to stand also for the reality of things." Faithful to his fundamental thesis, Locke nevertheless insisted that "it is a perverting the use of words, and brings unavoidable obscurity and confusion into their signification, whenever we make them stand for anything but those ideas we have in our own minds" (3.2.4–5).

Thus, the basic semantic relation in Locke's account of language is that of a word used by some speaker as a proper name for some idea in that speaker's mind. It seems to follow from this doctrine that as long as one does use words in this (the only approved) way, one cannot misuse them; and Locke does sometimes suggest that in the early chapters of Book III (see, for instance, 3.2.3). Those chapters indeed present a classic formulation of what Wittgenstein was later to criticize as the notion of a "private language."

Establishing words as proper names of ideas in the speaker's mind fulfills the first of Locke's two principal conditions "for the perfection of language" (3.1.3). The second was the devising of "general words," which he thought men accomplished by using words "for signs of general ideas." It was evident, Locke observed, that general words "do not signify barely one particular thing; for then they would not be general terms but proper names." His account of the signification of general words is, however, severely damaged by the inclusion in those same passages of his declaration of a thoroughgoing nominalism: "things themselves … are all of them particular in their existence, even those words and ideas which in their signification are general" (3.3.11–12).

Although many of his most careful theoretical statements ruled out any extension of the signification of a word beyond an idea in the speaker's mind, Locke here (and frequently in the later chapters of Book III) was apparently assuming that by virtue of signifying an idea, a word also (secondarily and indirectly, perhaps) signified whatever the idea signified. However, he never examined that assumption or even recognized it to be one. When he came to apply his theory to the discussion of various sorts of names, he often relaxed or ignored the strictures laid down in the general theory developed in the first three chapters. Thus, in his chapter on the "names of our ideas of substances" he found it convenient to say "By the word *gold* here, I must be understood to design [that is, designate] a particular piece of matter; v.g., the last guinea that was coined. For if it should stand here in its ordinary signification, for that complex idea which I or anyone else calls gold, i.e. for the nominal essence of gold, it would be jargon" (3.6.19).

When, on the other hand, Locke did apply his semantic theory strictly, he was likely to produce such surprising results as his doctrine that every generalization about a substance, such as "all gold is fixed," means either "that fixedness is a part of the definition, i.e., part of the nominal essence the word gold stands for; and so this affirmation 'all gold is fixed,' contains nothing but the signification of the term gold. Or else it means, that fixedness, not being a part of the definition of the gold, is a property of the substance itself, in which case it is plain that the word gold stands in the place of a substance.… In which way of substitution it has so confused and uncertain a signification that, though this proposition—'gold is fixed'—be in that sense an affirmation of something real, yet it is a truth will always fail us in its particular application [since we know only our idea of gold and not 'the real essence' of gold], and so is of no real use or certainty" (3.6.50; compare his interesting treatment of "trifling propositions" in 4.8).

Locke's strictly subjectivist, nominalist theory of signification in the opening chapters of Book III, which gave him so much trouble in its application, may represent nothing more than his overzealous attempt to state precisely such characteristically commonsensical observations as can be found in his *Conduct of the Understanding*, Section 29, where he advised "those who would conduct their understanding right, not to take any term … to stand for anything, till they have an *idea* of it. A word may be … used as if it stood for some real being; but yet if he that reads cannot frame any distinct idea of that being, it is certain to him a mere empty sound without a meaning."

(Locke's influence is frequently discussed in the remainder of this entry. See, for instance, the sections on

Gottfried Wilhelm Leibniz, George Berkeley, and Étienne Bonnot de Condillac and on universal grammar.)

LEIBNIZ

Gottfried Wilhelm Leibniz (1646–1716) developed some of his views on language specifically as criticisms of Locke in his *Nouveaux Essais sur l'entendement* (finished after 1709; first published 1765). One example of this *ad hoc* development is his rejection of Locke's account of "general words" as no more than devices for avoiding the proliferation of proper names. Leibniz argued that they were necessary ingredients in the "essential constitution" of languages and went so far as to claim, in an exact reversal of Locke's position, that "it is certain that all *proper* or individual names were originally *appellative* or general" (3.1.3; see 3.3.5). Even in the *Nouveaux Essais*, however, most of Leibniz's views on language can be traced to considerations that lie at the center of his own philosophy.

UNIVERSAL CHARACTERISTIC. Perhaps the most important of the central considerations of Leibniz's philosophy is his lifelong preoccupation with the idea of a "universal characteristic," which cannot be examined here except as it bears directly on his philosophy of language. Leibniz's earlier doctrine of the characteristic (c. 1679) was "that a kind of alphabet of human thoughts can be worked out and that *everything can be discovered and judged by a comparison of the letters of this alphabet and an analysis of the words made from them*" (Gerhardt edition 7.185). Descartes, by contrast, had maintained that such a language (he never knew of Leibniz's scheme, of course) depended on the prior establishment of "the true philosophy" (letter to Marin Mersenne [1629], in Adam and Tannery edition 1.76).

Leibniz's initial response was that while the establishment of the characteristic "does depend on the true philosophy, it does not depend on its completion"; for as long as we have the true "alphabet of human thought" to begin with, we can complete the true philosophy simply by correctly manipulating the characteristic (Couturat edition, pp. 27–28). (The many artificial languages projected during the Enlightenment may be classified as "Cartesian" or "Leibnizian," depending on whether they were put forward solely as devices for recording and communicating knowledge or also as heuristic devices. It is the Leibnizian rather than the Cartesian projects that bear a significant resemblance in principle to the formalized languages for logic developed after the middle of the nineteenth century.) Writing some years later (1697) and in a context where the issue between his own and Descartes's views was not explicit, Leibniz did nevertheless acknowledge that "genuinely real, philosophic characters must correspond to the analysis of thoughts. It is true that such characters would presuppose the true philosophy, and it is only now [when he believed himself to have discovered the principles of the true philosophy] that I should dare to undertake the construction of them" (Gerhardt edition 3.216).

By a "real" characteristic Leibniz meant a symbolism that was in some important respect naturally (rather than conventionally) associated with what it symbolized. Although a thoroughly real characteristic could be developed only in an artificial language, Leibniz observed that natural languages were in certain respects real characteristics. It was on the basis of that observation that he became the first major philosopher after Epicurus to suggest an appeal to ordinary language as a philosophical technique. His general attitude is expressed in the *Nouveaux Essais*: "I truly think that languages are the best mirror of the human mind and that an exact analysis of the signification of words would make known the operations of the understanding better than would anything else" (3.7.6). Part of what he meant by "exact analysis" closely resembled Plato's use of etymology in the *Cratylus*.

In his preface to a 1670 edition of Nizolius (in which he has a great deal to say about language) Leibniz argued that "the good grammarian and the philosopher too can, so to speak, deduce the use of a word from its origin by means of an unbroken sorites of metaphors" (Gerhardt edition 4.140). But he also viewed ordinary language in its unanalyzed state as having a special philosophic value: "Whatever cannot be explicated by means of popular terms (unless like many kinds of colors, odors, and tastes, it consists in immediate sensation) is nothing, and should be kept away from philosophy as if by a kind of purifying incantation" (4.143). Not every ordinary language was equally valuable as a touchstone for philosophy: "No language in Europe is better suited than German for this certifying trial and examination of philosophical doctrines by means of a living language, for German is richest and most nearly complete in real characters [*in realibus*], to the envy of all other languages.… On the other hand, the German language is easily the least well suited for expressing fabrications [*commentitia*]" (4.144; cf. Duclos edition 6.2.10 ff.). Leibniz's praise of German for its high proportion of real characters was very likely based simply on the fact that it contains words of Germanic origin where English and the Romance languages are likely to have words of Greek and Roman origin—for instance,

Unabhängigkeit and "independence"—a feature of the German language which no doubt does provide its native speakers with comparatively easy access to many abstract notions.

Leibniz also recognized a more pervasive kind of "realness" in natural languages that might be called *syntactic*, in contrast with the historically more familiar kind just discussed. It constituted the essential ingredient in his doctrine of "expression" and thus formed part of his metaphysics (monads *express* the universe) as well as of his philosophy of language. In *What Is an Idea?* (1678) he offered this account:

> That is said to *express* a given thing in which there are relations [*habitudines*] that correspond to the relations belonging to the thing expressed. But these expressions are of different kinds—e.g., the model of a machine expresses the machine itself; the projective delineation of a thing in a plane expresses the solid; discourse expresses thoughts and truths; characters express numbers; an algebraic equation expresses the circle (or some other figure)—and what is common to these expressions is the fact that we can pass from the mere consideration of the expressed relations to a knowledge of the corresponding properties of the thing being expressed.

Leibniz drew the conclusion that "it is clearly not necessary that that which expresses be similar to that which is expressed as long as a certain analogy of [internal] relations is preserved" (Gerhardt edition 7.263–264). What he was proposing, however, was clearly a novel approach to resemblance as a basis for semantic relations, suggesting for the first time that in complex signs the "realness" of the symbolism may consist in the resemblance between the *schemata* of the expression and of what is expressed rather than in a resemblance between the *elements* of those two schemata. This was brought out most clearly in his *Dialogue* (1677)—for example, in the observation that "even if the *characters* are arbitrary, still *the use and interconnection* of them has something that is not arbitrary—viz. a certain proportion between the characters and the things, and the relations among different characters expressing the same things. This proportion or relation is the foundation of truth" (7.192).

In describing this schematic resemblance as the foundation of truth, Leibniz stated the principal thesis in his novel doctrine of propositions as extralinguistic, extramental schemata. Although such a notion of propositions had been hinted at by Hobbes, Leibniz was evidently the first to make it explicit; and, as it happened, he developed his doctrine in conscious opposition to Hobbes's view of truth as dependent on words and hence arbitrary. It had been standard philosophical usage from the beginning of the Middle Ages to use the word *proposition* for whatever was either true or false, and the principal refinement of this usage before Leibniz had been the medieval distinction of "mental" propositions from propositions spoken or written. Leibniz's first objection against what he called Hobbes's "super-nominalism" might be interpreted as going no further than that, as in his observation that "truths remain the same even if the notations vary" (*Preface to Nizolius* [1670], in Gerhardt edition 4.158). He subsequently recognized, however, that those "truths" could not be identified with true propositions that had been, were, or would be actually in someone's mind—for instance, in the *Dialogue* of 1677: "*A*. … Do you think that all propositions are thought? / *B*. I do not. / *A*. You see, therefore, that truth does belong to propositions or thoughts, but to *possible* [propositions or thoughts], so that this at least is certain, viz. that *if* anyone should think in this [or a contrary] way, his thought would be true [or false]" (7.190).

Once Leibniz had distinguished propositions from actual thoughts and from combinations of words, he was in a position to reject the traditional account of truth as "the conjunction or separation of signs according as the things themselves agree or disagree among themselves," in which account "by 'the conjunction or separation of signs' one must understand what is otherwise called a proposition." Leibniz's attack on this tradition contrasted its technical terminology with ordinary usage in order to show that it concealed rather than resolved problems:

> An epithet—e.g., "the wise man"—does not make a proposition, and yet it is a *conjunction* of two terms. Negation, moreover, is something different from *separation*, for saying "the man" and after an interval pronouncing "wise" is not to deny. Finally, *agreement* [or *disagreement*] is not, strictly speaking, what one expresses by means of a proposition; two eggs have agreement and two enemies have disagreement. The manner of agreeing [or disagreeing] at issue here is quite extraordinary [*toute particulière*]. Thus I think this definition completely fails to explicate the point at issue. (*Nouveaux Essais* 4.5.2)

And Leibniz went on from this criticism of the traditional doctrine of propositions to present once again his own

view of them as entities distinguishable both from words and from actual ideas (*Nouveaux Essais* 4.5.2).

"**LEIBNIZ'S LAW**." Leibniz's famous principle of substitutivity, known in recent literature as Leibniz's Law, was frequently used as a starting point by twentieth-century writers on semantics. Leibniz employed the principle as part of the primitive basis of his logical calculus and put forward several versions of it in papers written from 1679 through the early 1690s. The various versions may be accurately synthesized as follows: Those entities are the same, one of which may be everywhere substituted for the other, preserving the truth(-value) (see 7.219; 7.228; 7.236).

Although Leibniz did not identify the entities in question and sometimes discussed the principle as if it applied, for example, to geometrical figures, the context generally makes it plain that its principal intended application was to terms in propositions actually expressed in some notation. His discussion of the principle in the papers in which he applied it took no account of contexts in which the principle does not apply, but at least one passage in his later writings shows that he had by then recognized that cases of what the medievals had called material supposition did not fall under the principle. "Indeed, one sometimes speaks of words *materially* without being able in that context [*cet endroit-là*] to substitute in place of a word its signification, or its relation to ideas or to things. This occurs not only when one speaks as a grammarian but also when one speaks as a lexicographer, in giving the explication of a name" (*Nouveaux Essais* 3.2.6). Recent criticism of Leibniz's Law has often begun with the complaint that he failed to notice just such exceptions.

Berkeley. Locke had argued that a word was significant solely in virtue of standing for an idea in the mind of the user of the word. When George Berkeley (1685–1753) began philosophizing, he accepted that doctrine as axiomatic. In several early entries in his private *Philosophical Commentaries* (1707–1708) he presented it as part of the basis of his otherwise anti-Lockean position—for instance, "All significant words stand for Ideas" (Luce and Jessop edition 1.45; see 1.39, 1.43, 1.53). Even before ending the *Commentaries*, however, Berkeley had rejected Locke's semantics too and had begun to replace it with a doctrine of great importance in the development of his own philosophy and in the history of semantics.

The actual turning point was apparently reached in his discovery that some words that should have been paradigm cases for Locke's semantics had no precisely correspondent ideas. "Qu: How can all words be said to stand for ideas? The word Blue stands for a Colour without any extension, or abstract from extension. But we have not an idea of Colour without extension; we cannot imagine Colour without extension" (1.62). In this passage Berkeley questioned for the first time not only Locke's semantics but also (indirectly) his doctrine of abstract ideas. He very soon saw that the connection between the two was essential. Given Locke's semantics, together with the facts that a general word was significant and that no concrete particular idea corresponded to it, one was forced to introduce a Lockean abstract idea simply in order to give a general word something to stand for. (As Berkeley pointed out [2.36], Locke had virtually admitted as much in the *Essay* [3.6.39].) Berkeley's alternative account in the Introduction to *The Principles of Human Knowledge* (1710), Section 11, was that "a word becomes general by being made the sign, not of an abstract general idea but, of several particular ideas, any one of which it indifferently suggests to the mind" (2.31; see 2.127). Berkeley's account thus involved abandoning Locke's semantic principle that there be a single idea to serve as the name-bearer for each significant word.

In the history of semantics, however, as in the history of philosophy in general, Berkeley's rejection of abstract ideas is more important than his alternative account of the signification of general words. The rejection was based not only on the well-known exposition of the internal inconsistency—as in *Principles*, Introduction, Section 13 (2.32–33)—but also on his many and varied attacks on their semantic foundation. Since Locke's commitment to the view that each word had to stand for one idea in order to be significant was what had compelled him to introduce abstract ideas, Berkeley set out to show, by means of various sorts of counterinstances, "that words may be significant, although they do not stand for ideas. The contrary whereof having been presumed seems to have produced the doctrine of abstract ideas" (3.292–293; see 1.70).

He seems to have found at least four sorts of counterinstances, the first and most obvious consisting of words that stand for something other than ideas. Words such as "volition," "I," "person," and the "particles" (or syncategorematic words) are significant in virtue of standing for "spirits" or their activities (the particles standing for "the operations of the mind") (1.65, 1.80, 1.81; see 3.292). But in Berkeley's immaterialism there were no entities other than spirits and ideas for which words could stand, and so there could be no other counterinstances consisting simply of words that stood for

nonideas. He was thus led to investigate the relation "stands for" more closely than its relata. In a move reminiscent of supposition theory he attacked Locke's account of "understanding propositions by perceiving the agreement or disagreement of the ideas marked by their terms," claiming that when he asserted of a particular dog "Melampus is an animal" he had not two ideas but "only one naked and bare idea, viz. that particular one to which I gave the name Melampus." Nor does "animal" in that proposition "stand for any idea at all. All that I intend to signify being only this, that the particular thing I call Melampus has a right to be called by the name animal." But it would not do, he pointed out, to say that "animal" here stood for the same idea as did "Melampus," since that would make the proposition a tautology (2.136–137; cf. 1.69, 8).

The principal effect of this second sort of counterinstance was to raise some serious doubt regarding the nature of the relation "stands for," and Berkeley's remaining counterinstances took the almost unprecedented step of suggesting that that relation was not always an essential ingredient in significance. Words that might in certain occurrences be said to stand for ideas are very often used in reasoning and in ordinary conversation as uninterpreted (but interpretable) "counters." A word used in that way does *not* in each of its occurrences stand for an idea in the mind of the user or, for that matter, raise a corresponding idea in the mind of the hearer or reader (2.37, 3.291–292, 8.25, 8.27).

Finally, a word sometimes occurs in a context such that one would miss rather than grasp its significance by taking it to stand for the idea to which it is customarily attached. "For example, when a Schoolman tells me *Aristotle hath said it*, all I conceive he means by it, is to dispose me to embrace his opinion with the deference and submission which custom has annexed to that name" (2.38). What is more, a word may occur in a context that precludes the possibility of taking it to stand for an idea without thereby being rendered insignificant—for example, the subject term in "the good things which God hath prepared for them that love him are such as eye hath not seen nor ear heard nor hath it entered into the heart of man to conceive."

It was Berkeley's view that the significance of propositions such as these last two was to be found not in the ideas the words might otherwise be said to stand for but in the purpose, or "design," of the proposition. The design of this last example cannot be "to raise in the minds of men the abstract ideas of thing or good nor yet the particular ideas of the joys of the blessed. *The design is to make them more cheerful and fervent in their duty*" (2.137 [italics added]; see 2.293, 3.292). Words, he held, "have other uses besides barely standing for and exhibiting ideas, such as raising proper emotions, producing certain dispositions or habits of mind, and directing our actions" (3.307). Thus, in his attacks on the semantic foundation of Locke's doctrine of abstract ideas Berkeley came nearer than anyone since the Stoics to abandoning, or at least supplementing, the attempt to account for all linguistic meaning in terms of the relation between names and their bearers.

As for Locke's semantics, Berkeley had reduced it to the unexceptionable principle of common sense that had no doubt prompted Locke's theoretical claims, namely, we ought not to use words without knowing their meaning (1.78; 2.76). But he took Locke's call for a new "doctrine of signs" quite seriously, summarizing his own (mostly anti-Lockean) semantic theory under that heading in *Alciphron* (1732), Dialogue VII, Section 14 (3.307). Like Bacon, Hobbes, and Locke before him, Berkeley thought of himself as providing philosophical remedies for the abuse of words, but he differed from them in making this the core of his philosophy, announcing that "the chief thing I do or pretend to do is only to remove the mist or veil of Words" (1.78; see 2.40). He set out explicitly to do just that at many points throughout his writings, but nowhere in a more concentrated form than in the introduction to the *Principles*.

In keeping with that aim Berkeley frequently urged his readers to contemplate ideas apart from words, maintaining that "if men would lay aside words in thinking 'tis impossible they should ever mistake save only in Matters of Fact" (1.84; see 2.40). He felt, therefore, that it was "absurd to use words for the recording our thoughts to ourselves: or in our private meditations" (1.62) and introduced his "Solitary Man" for the purpose of examining that pristine state of mind in a concrete example, "to see how after long experience he would know without words" (1.71; see 2.141–142). Such passages taken together suggest an anticipation of Wittgenstein's attack on the notion of a private language, but Berkeley had second thoughts about the absurdity of the private use of words and seems to have concluded that "the Solitary Man would … find it necessary to make use of words to record his Ideas if not in memory or meditation yet, at least, in writing without which he could scarce retain his knowledge" (1.75).

Berkeley's ingenious linguistic analogy in his account of sense experience, the "Universal Language of Nature," was first put forward in his *New Theory of Vision* (1709) and developed in several later works. Speaking strictly, it

belongs to his theory of knowledge rather than to his philosophy of language, but in the course of developing the analogy he often made interesting observations about language conceived in the ordinary sense (see, for instance, 1.228–233, 1.264–265).

MAUPERTUIS AND HIS CRITICS

In the latter half of the eighteenth century philosophical interest in language was concentrated among French philosophers. Under the influence of Condillac and the British empiricists they eventually came to consider the analysis of signification their most important task. Among the earliest figures in this development was Pierre-Louis Moreau de Maupertuis (1698–1759), who first published his brief *Réflexions philosophiques sur l'origine des langues et la signification des mots* in 1748. To some extent his position resembled those taken by Berkeley and by Condillac in his first book, *Essai sur l'origine des connaissances humaines* (1746); but Maupertuis seems to have written his *Réflexions* before he knew their work. Partly because of the author's fame as a scientist, the *Réflexions* attracted considerable attention and was commented on by Baron de L'Aulne Turgot (1750), Condillac (1752), Denis Diderot (1753), Voltaire (1753), and Maine de Biran (1815), among others.

Maupertuis conceived of the question of the origin of language very much as philosophers since Descartes had been conceiving of the question of the origin of knowledge. It was intended to give rise not to speculations about prehistoric man but rather to an analysis of the hypothetical circumstances of a man with fully developed faculties who has suddenly been deprived of all his memories and of all human society. Would such an individual frame a language at all? If he did so, what would be the stages of its development? By asking and answering such questions as these within the framework of his "metaphysical experiment," Maupertuis expected to gain insight into the nature of language and its relation to the acquisition of knowledge. He began by imagining himself in the condition of the adult newborn. As soon as he had had two perceptions,

> I should see that the one was not the other, and I should try to distinguish between them. And since I should have no ready-made language, I should distinguish between them by means of any marks whatever and might be satisfied with the expressions "A" and "B" as standing for the same things I now mean when I say "I see a tree," "I see a horse." Receiving new perceptions afterwards, I could designate them all in that way. (*Réflexions*, Sec. 7)

It is not clear whether his saying "A" to himself in this protolinguistic context is really separable from his act of individuating the perceptual event of his seeing a tree, but Maupertuis did consider "A" and "B" as signs of his perceptions and thus presented a nearly classic case of what Wittgenstein later described as a "private language." The first development beyond those initial "signs" was recognized by Maupertuis as sufficiently radical to be described as "another language." In Section 8 he wrote:

> For example, in the preceding perceptions I should recognize that each of the first two had certain characteristics that were the same in both and that I could designate those by a single sign. Thus I should change my first simple expressions "A" and "B" into these: "CD" and "CE," which would differ from the first only in that new convention, and which would correspond to the perceptions I now have when I say "I see a tree," "I see a horse."

Maupertuis's analysis proceeded in this way until he had introduced devices for discriminating kinds of perception, numbers of objects perceived, remembered and anticipated perceptions, and so on. His purpose in doing so, however, was to provide the background for a new philosophic method, which he applied most notably in his analysis of "the force of the proposition 'there is ….'" Although in saying "there is a tree" I may seem to be making a claim that goes beyond the evidence of my perceptions, once language has been reconstructed on the basis of my perceptions alone, I am in a position to see that "there is a tree" is no more than an abbreviation for "I shall see a tree every time I go to that place." This latter proposition in turn is reducible to the sequence "I was in a certain place," "I saw a tree," "I returned to that place," "I saw that same tree again," and so on (Secs. 24–28).

Eight years after writing the *Réflexions* and having meanwhile read the French translation of Berkeley's *Dialogues* (1750), Maupertuis readily admitted the similarity of his metaphysics to Berkeley's and rested his claims of independent importance on having introduced an analysis of language as the means to that end. "The point is that this philosopher [Berkeley] attacks the system of our errors only by parts. He demolishes the structure at the top, and we undermine its foundations. This is a structure quite different from that famous tower the erection of which on the Plains of Shinar was prevented by the confusion of tongues; this one is not erected *except by* abusing or forgetting the meaning of words" (*Reply to*

Boindin, Sec. II, in *Oeuvres*, 1756 ed.). Berkeley's own attitude was, of course, much the same; but the *Dialogues* alone among his major works fails to bring that out.

Several of Maupertuis's critics, most notably Turgot, attacked the hypothesis on which he rested his inquiry: "A solitary man such as Maupertuis imagines … would not try to find marks with which to designate his perceptions. It is only when confronted with other people that one looks for such marks. From this there follows what is obvious in any case, that the first purpose and first step of language are to express objects and not perceptions" (*Remarques critiques*, Sec. 7). Only Maine de Biran among Maupertuis's critics defended his use of the private-language hypothesis (*Note sur les Réflexions*, Sec. V).

CONDILLAC

Étienne Bonnot de Condillac (1715–1780) wrote his first book, *Essai sur l'origine des connaissances humaines* (1746), in an effort to do what he felt Locke might have done if he had not "realized too late" the importance for epistemology of the material in Book III of his *Essay*, "Of Words" (Locke's *Essay* 3.9.21). Locke had "treated only in his third Book what should have been the subject matter of the second" (*Essai*, edited by G. Le Roy, 1.5a). Condillac acknowledged its historical value: Locke seemed to him to have been "the first to have written on this material as a genuine philosopher. I felt, nevertheless, that it had to form a considerable portion of my own work, both because it can be viewed in a novel, more extended way and because *I am convinced that the use of signs is the principle that discloses the source* [*développe le germe*] *of all our ideas*" (1.5b; italics added). Condillac thus became the first modern philosopher to found his theory of knowledge, and consequently his entire philosophy, on considerations of signification and language, considerations that occupied him throughout his career and that shaped French philosophy for at least fifty years afterward.

Like Locke, Condillac denied that the ideas produced in sensation alone constitute a kind of knowledge, but he began his divergence from Locke in his account of the acquisition of knowledge on the basis of such ideas. "The *sole* means of acquiring knowledge is to trace our ideas back to their origin, to observe their generation, and to compare them under all possible relations. This is what I call *analysis*" (1.27a). Analysis consists in discriminating and ordering elements that are presented confusedly and simultaneously and thus requires the introduction of interrelatable signs for those elements. On these observations Condillac based his leading principle that "every language is an analytic method and every analytic method is a language" (2.419a).

This has the look of a vicious circle. Analysis is said to be a necessary condition of knowledge, and language to be necessary for analysis; but surely knowledge is also necessary for the formation of a language. Condillac attempted to break this circle by introducing the notion of an innate language, which he called the language of action. "The elements of the language of action are born with man, and those elements are the organs given us by the author of our nature. Thus there is an innate language although there are no innate ideas. Indeed, it was necessary that the elements of some sort of language, prepared in advance, should precede our ideas, since without signs of some kind it would be impossible to analyze our thoughts" (2.396b).

In its most rudimentary form this "language" consists simply in overt reactions: "our external conformation is set to represent everything that takes place in our soul" (ibid.). Involuntary expressions of fear, pain, desire, and so on are not elements of analysis for the individual producing them, but observers of his responses can, as a result of observing the order of events making up his responses, see analyzed for them what is simply gross experience for the respondent. "Men begin to 'speak' the language of actions as soon as they feel anything, and they speak it then without having any plan of communicating their thoughts. They form the plan of speaking it in order to make themselves understood only when they notice that they have been understood" (2.397a). The usefulness of results gained by this means stimulates a natural feedback process of development on "the principle of analogy," and the language of action is made more effective by the gradual transformation of "natural" and "accidental" signs into "signs of institution," the most convenient of which are articulate sounds (1.60b–62a). The origin of language, discussed as an independent topic by many of his contemporaries and successors, is thus an essential consideration in Condillac's epistemology.

Signs of institution, including, of course, words, are themselves natural in the sense that as a language develops, they are framed on analogy with more primitive elements in that same language (and ultimately with elements of the language of action). The principle of analogy is in fact a necessary ingredient in any usable language (compare Bacon's doctrine of "emblems"). "Imagine an absolutely arbitrary language, such that analogy had determined neither the choice of words nor their various senses. That language would be an ununderstandable gibberish" (2.471a). If the principle of analogy

remained unimpaired in ordinary languages, "we would reason as nature teaches us to reason, moving effortlessly from discovery to discovery." But every ordinary language has been impaired to some extent by the intrusion of words that have the roots of their analogy in other languages. (A similar line of reasoning had led Leibniz to praise the German language as a natural "philosophical characteristic.") Perhaps, then, the principle of analogy can be retained in a perfectly unadulterated form only in a highly artificial language, such as algebra, which Condillac describes as "the language of mathematics."

Since language of some sort is a necessary condition of knowledge, it is a mistake to maintain, as Locke had done, that the primary purpose of language is to communicate knowledge. "The primary purpose of language is to analyze thought. In fact we cannot exhibit the ideas that coexist in our mind successively to others except in so far as we know how to exhibit them successively to ourselves. That is to say, we know how to speak to others only in so far as we know how to speak to ourselves" (1.442a). It is a consequence of this view that the art of thinking, or logic, reduces to the art of speaking.

> Although a thought is not a succession in the mind, it has a succession in discourse, where it is decomposed into as many parts as there are ideas making it up. Then we can observe what it is we do when we think, we can give an account of it, we can, consequently, learn how to conduct our reflective thought. In this way thinking becomes an art, and that art is the art of speaking. (1.403b)

Condillac's view of the connection of thought and language was reinforced by his observations on "abstract general ideas." "When, for example, I think about *man*, I cannot consider anything in that word except a common denomination, in which case it is perfectly plain that my idea is in some way circumscribed in that name, that it extends to nothing beyond the name, and that, consequently, it is only that name itself" (2.401b). Thus the clarity and precision of abstract ideas "depends entirely on the order in which we have produced the denominations of classes. Therefore, there is only one means of determining ideas of this sort, and that is to produce a well-made language" (ibid.). Abstract general ideas, however, are the principal ingredients of reasoning, and Condillac was even ready to say that "to speak, to reason, to produce abstract or general ideas for oneself, are at bottom one and the same thing" (2.402a). His consideration of abstract ideas, then, was one more "proof that we reason well or badly only because our language is well or badly made" (ibid.).

All intellectual progress, on Condillac's view, depended on and in part consisted in establishing a "well-made language," and "a science, properly treated, is nothing other than a well-made language" (1.216a). The one perfectly well-made language so far established, he thought, was mathematics, which he examined from this point of view in his last book, *La langue des calculs* (1798). One reason why Condillac was prepared to identify a science with a well-made language is to be found in his doctrine of propositions. All that remains to be done in a science once the appropriate language has been established is the mechanical exposition of the truths proper to that science. The exposition is mechanical because "a proposition is only the unfolding of a complex idea in whole or in part," and since a proposition "in which one and the same idea is affirmed of itself" is an identical proposition, "every truth is an identical proposition" (2.748a). An identical proposition may, however, be instructive for some persons, namely, those who observe "for the first time the relation of the terms out of which it is formed.… Thus a proposition may be identical for you and instructive for me" (2.748b). Nevertheless, "if in all the sciences we could equally trace the generation of the ideas and everywhere apprehend the true system of the things, we should see one truth give birth to all the rest, and we should find the abridged expression of all we know in this identical proposition: *the same is the same*" (2.749b).

Condillac's influence extended not only to philosophers but also to the great chemist Antoine Lavoisier (1743–1794), who in his *Méthode de nomenclature chimique* (with Louis-Bernard Guyton de Morveau, 1787) and *Traité élémentaire de chimie* (1789) wholeheartedly adopted Condillac's notion of a science as a well-made language. Operating under this notion, Lavoisier introduced such technical terms as "phosphoric acid" and "sulphuric acid" in a successful attempt to initiate the development of the language of modern chemistry on Condillac's principle of analogy.

LAMBERT, HAMANN, AND HERDER

In the century between Leibniz and Wilhelm von Humboldt, philosophy of language in Germany was concentrated in the writings of three men: Johann Heinrich Lambert (1728–1777), Johann Georg Hamann (1730–1788), and Johann Gottfried Herder (1744–1803).

Lambert was a distinguished mathematician and the first man to follow Leibniz's lead in his contributions to

logic, the most important of which was the earliest attempt at a calculus of relations. His work in the philosophy of language appeared in his *Neues Organon* (1764), especially Part III, "Semiotik, oder die Lehre von der Bezeichnung der Gedanken und Dinge."

In philosophy of language, as in logic, the principal influence on Lambert was that of Leibniz, as may be seen in his preoccupation with the effect of language on thought and knowledge and with the possibility of controlling and improving that effect. Various natural languages impose various structures on our knowledge, but every natural language is fundamentally the product of prephilosophical, prescientific humankind. When we attempt to use such a language in advanced intellectual activities, we must submit our thought to the tyranny of usage (III, 1). We are thus led to seek an artificial language that from its inception could be entirely subjected to the needs of the intellect.

Lambert's attitude toward such artificial languages differed, however, from Leibniz's and constitutes a significant development in this line of thought. Great men, he observed, have worked at the project of a simple, perfectly regular and precise rational language, but without notable success. In any case, the adoption of such a language would be practically impossible (III, 2, 330). If we then revert to natural languages, however, we find that, strictly speaking, we cannot adopt any single one of them as a foundation for knowledge. There are, in the first place, conflicting usages even within a single natural language, some of which would have to be more or less arbitrarily ruled out; and, in the second place, any set of usages finally adopted would inevitably continue to undergo changes within the natural language of which they were a part.

Once we recognize that we do thus necessarily deviate to some extent from any given language in adapting it to intellectual purposes, it is apparent that we ought to do so consciously and under the guidance of preexamined criteria. The criteria developed and employed by Lambert were, he observed, the sort that might have served as the operative rules of a philosopher's artificial language. In fact, he seems to have elevated Leibniz's projected "universal characteristic" to the status of an ideal language, the principles of which are approximable to varying degrees but never fully realizable. He described his detailed examination of language as one that made a point of not distinguishing sharply between "actual and possible languages," meaning thereby that his approach to natural language was a mixture of description and prescription in which he attempted to point out those aspects of the actual language which were already accommodated to certain requirements of the ideal and to suggest ways in which those aspects might be enhanced and extended without introducing radical reforms that had little chance of acceptance.

The fundamental criterion employed by Lambert in his evaluation of sign systems in general and of natural languages in particular was the interchangeability of "the theory of the signs" and "the theory of the objects" signified, the degree of interchangeability marking the extent to which the signs approximated the fundamental ideal of being "scientific" (III, 23–24)—he cited musical notation as an example of a particularly close approximation. It seems evident that this fundamental criterion, which with its many corollaries pervades Lambert's philosophy of language, constituted his adaptation of Leibniz's doctrine of "expression." Besides systematic general chapters on various aspects of language, Lambert's "Semiotik" includes specific examinations of the character and function of various parts of speech and of the philosophical significance of etymological and syntactic interrelations among words.

Lambert and Hamann shared the conviction that the character of language was a topic of the greatest importance for philosophy, but they differed in almost every other important respect. Hamann's writings are undisciplined, obscure, and strongly colored by religious mysticism. Philosophically he was a forerunner of romanticism and existentialism, consciously rejecting most of the attitudes of the Enlightenment.

To the extent to which Hamann's philosophy exhibits a structure, it centers on his views on language, so much so that he himself called it verbalism (*Schriften*, edited by C. H. Gildemeister, 5.493–495). In almost everything he had to say about language, however, he opposed his contemporaries—Lambert (and the Leibnizian tradition) implicitly, Herder explicitly. The fundamental thesis of Hamann's verbalism is that ordinary natural language does and should take philosophical precedence over all technical or abstract language. Occasionally he wrote as if his basis for this claim was that God had employed such language as the instrument of revelation (*Schriften*, edited by F. Roth and G. A. Wiener, 1.85–86, 1.99), but he seems to have had more generally evaluable reasons for it as well. He evidently felt that the opposition between the rationalists and the empiricists of the Enlightenment was irresolvable largely because of the reliance of both parties on introspection. The special importance of ordinary language in this connection was that it constituted a medium in which the operations of reason and experi-

ence were united and made publicly accessible. The operations of reason, indeed, consisted entirely in linguistic operations (Roth and Wiener, 6.15 and 6.25; Gildemeister, 5.515, 7.9). Philosophy, however, had traditionally adulterated what should have served as its principal source and instrument. Hamann brought this out in a characteristic attack on Immanuel Kant's abstract, technical language in the first *Critique*:

> While geometry fixes and fictionalizes the ideality of its concepts of points without parts, of lines and planes conforming to ideally divided dimensions, by means of empirical signs and figures, metaphysics misuses all the word-signs and figures of speech of our empirical knowledge as mere hieroglyphs and types of ideal relations and as a result of this learned mischief transforms the straightforwardness [*Biederkeit*] of language into such a senseless, whirling, unsteady, indeterminable something (= *x*), that nothing remains but a windy murmuring, a magical shadow-play, at best … the talisman and rosary of a transcendental superstition regarding *entia rationis*, their empty sacks and slogan. (Roth and Wiener, 7.8)

It was not only Kant's misuse of language that attracted Hamann's criticism but more especially his utter neglect of language as a topic for inquiry, which from Hamann's point of view vitiated Kant's claim to have provided a critique of reason. To point up the folly of such neglect, he tried to show that at various crucial junctures in his argument (as in the deduction of the categories) Kant had uncritically relied on certain linguistic conventions and that what he had called paralogisms and antinomies of reason really had their roots in the misuse of language.

Hamann based his doctrine of the preeminence of ordinary language not only on its value as a subject for philosophical inquiry but also on the fact that it alone among types of language was "objectively given." As such it served as the "womb" of reason and of all specialized, abstract languages designed to aid the operations of reason. Moreover, since ordinary language thus constituted the ultimate link between language and reality, all such abstract languages must be held finally accountable to it, that is, translatable into it.

As a philosopher of language, Herder is best known for his prize essay on the origin of language (1771), a topic with which this entry is generally not concerned. Herder's essay, however, occupies a position of special importance among hundreds of similar productions by eighteenth-century philosophers, for it began the trend away from the speculative problem of origin and toward the scientifically more accessible problems of the development of language. (It was praised for that reason by several of the great linguists of the nineteenth and twentieth centuries, such as Grimm, Theodor Benfey, Edward Sapir, and Otto Jespersen.)

Ostensibly Herder was adjudicating between two rival accounts of the origin of language, but his real purpose was to dismiss the problem as senseless. The two theories at issue were those of special divine creation and of deliberate human invention of language, the former as represented in J. P. Süszmilch's work and the latter associated primarily with Jean-Jacques Rousseau's second *Discours*. Herder took reason to be the defining characteristic of man and argued in support of Hamann's position that the operation of reason and the use of language were inseparable. He then drew the obvious conclusion that if God had created what was genuinely man, He had created a language-using animal and no special divine creation of human language was conceivable. By the same token, animals correctly describable as men could not conceivably have invented language. Thus the question of how and when humans came to use language was misconceived, although the question of how primitive human languages developed was well worth considering. Hamann ridiculed Herder's argument, with justification. It seems probable, however, that the argument was intended as irony, to deflate the pretensions of the theorists rather than to refute the theories.

THE "IDÉOLOGUES."

Antoine Louis Claude Destutt de Tracy (1754–1836) devised the name *idéologie* for the new "science of the analysis of sensations and ideas." One section of the Second Class of the Institut National (founded in 1795) was devoted to that science in lieu of the prescientific inquiry known as metaphysics, and Destutt de Tracy and other philosophers associated with the work of that section became known as *idéologues*. For about eight years, until Napoleon Bonaparte abolished the Second Class of the Institut in the reactionary atmosphere of the First Empire, the *idéologues* were the dominant philosophical group in France. They thought of themselves as working in a field that had been opened by Locke and first thoroughly explored by Condillac, whose most original contribution was considered to be his discovery that language was as essential to the more fundamental processes of thought and analysis as it was to communication. Although part of at least Destutt de Tracy's interest in lan-

guage was directed specifically to *grammaire générale* (see below), the *idéologues* followed Condillac in considering language a topic of importance in every area of philosophical inquiry.

The *idéologues* resented being thought of as disciples of Condillac, however, and while they did, professedly, share some of his broad philosophical convictions, much of what they had to say about language (as about other topics) involved substantial revision or outright rejection of Condillac's specific doctrines. There seems to have been some tendency for the revisions to take the form of a generalization of Condillac's doctrines, notably from a concern with language to a concern with signs of all sorts. Thus, in a representative passage of his *Rapports du physique et du moral* (delivered in 1796), Pierre-Jean-Georges Cabanis (1757–1808) purported to be defending and explicating Condillac's central claim that every language was an analytic method by arguing that "one distinguishes among sensations only by attaching to them signs that represent and characterize them; one compares them only in so far as one represents by signs either their resemblances or their differences." Of course, Cabanis pointed out, taking this account as explicative of Condillac's claim required taking "language" in "the broadest sense," as meaning "the methodological system by means of which one pins down [*fixe*] one's own sensations" (*Oeuvres*, edited by Claude Lehec and Jean Cazeneuve, p. 157).

The question of the nature and epistemological function of signs (including linguistic signs) took on critical importance for the *idéologues* as a result of Destutt de Tracy's *Mémoire sur la faculté de pensée* (delivered in 1796), prompting them to set "the influence of signs on the faculty of thought" as the subject for the first essay competition sponsored by the section on the analysis of sensations and ideas. The best entries were *Des Signes envisagés relativement à leur influence sur la formation des ideés* by Pierre Prévost, *Introduction à l'analyse des sciences, ou de la génération, des fondements, et des instruments de nos connaissances* by P.-F. Lancelin, and *Des Signes et de l'art de penser, considérés dans leurs rapports mutuels* by Marie-Joseph Degérando (1772–1842). Degérando's essay won the prize and was published in an expanded four-volume version in 1800. In it some of the principal issues in the philosophy of language of the seventeenth and eighteenth centuries were subjected to a final scrutiny, partly as a result of the historically apt questions provided by the *idéologues* as a guide for the essayists:

> (1) Is it really the case that sensations can be transformed into ideas only by means of signs? Or, what comes to the same thing, do our first ideas depend essentially on signs? (2) Would the art of thinking be perfect if the art of signs were perfected? (3) In those sciences in which there is general agreement as to the truth, is this because of the perfection of the signs employed in them? (4) In those branches of knowledge that provide inexhaustible fuel for dispute, is the division of opinion a necessary effect of the inexactitude of the signs employed in them? (5) Is there any means of correcting badly made signs and of rendering all sciences equally susceptible of demonstration? (*Mémoires de l'Institut National des Sciences et Arts. Sciences morales et politiques*, 1.i–ii)

Question (1) was on a thesis of *idéologie* itself, as may be seen in the passage from Cabanis quoted above. Degérando's answer was complex, but it was sufficiently affirmative to mark him an *idéologue*. On the one hand he felt that the mind needed no signs but merely an act of attention in order to pin down its sensations. On the other hand, "I shall give the name ['sign'] to every sensation that excites an idea in us in virtue of the association obtaining between them. Note carefully that it is not the sensation as such to which the name is given; it gets the name only in respect of the function it performs. Thus I shall say, for example, that the smell of a rose is the sign [not of the rose but] of the ideas of color and of form that the smell excites" (1.62–63). He distinguished between such prelinguistic signs and linguistic signs by pointing out that while the former "excite" ideas in us but attract attention to themselves, the latter lead our attention away from themselves to the ideas they have been made to signify, a formulation that constituted a refinement of the traditional distinction between natural and conventional signs.

In his detailed answers to questions (2) and (5) Degérando carefully criticized the many attempts at universal characteristics, calculi of reason, and philosophical languages that had been made by philosophers of the Renaissance and the Enlightenment. He laid down five criteria for such systems: (*a*) unambiguous relations between signs and signified ideas; (*b*) relations among signs exactly analogous to relations among signified ideas; (*c*) simplicity, that is, minimum number of primitives (*conditions premières*), each sign as abbreviated as possible, perspicuity of the sign system as a whole; (*d*) distinctness among signs of different sorts and among

syntactic relations of different sorts; (*e*) as many distinct sorts of signs as distinct sorts of ideas to be signified (4.353–355). The only hopes of satisfying such criteria, he maintained, lay along four different lines, or "systems for philosophical language." Having examined each in detail, he concluded that all were in some respects unacceptable. Like Lambert, however, he suggested that a judicious application of the principles of such systems might produce some improvement in natural languages for philosophical purposes (4.355–415).

Perhaps even more important historically than his arguments against the feasibility of such artificial languages was his attack on the attitudes underlying them. It had usually been assumed that such a language would be international, and, indeed, if it were not, it would fail to achieve a good part of its purpose. But, Degérando maintained, there was no feasible means by which to establish it internationally, and even if it were established, it would soon be modified into separate dialects in various localities (3.557). Worst of all, the notion is pernicious, for such a language could at best be the instrument of communication exclusively among the learned and would thereby tend to separate them further from those they ought to instruct (3.572).

The fundamental mistake giving rise to all such schemes, according to Degérando, is the confusion of "the *method of reasoning* employed by the mathematicians with the *mechanical processes* of their calculations. Their method, as I have shown …, they do have in common with the metaphysicians," but the mechanical processes of their perfectly satisfactory artificial languages are the result of "the relative simplicity of the ideas on which they operate" (4.447–451). Other *idéologues*, particularly Destutt de Tracy, joined in this thoroughgoing repudiation of artificial languages for philosophy (see, for instance, *Mémoires de l'Institut*, Vol. III).

Questions (3) and (4) together called for an examination of Condillac's contention that a science was to be identified with a well-made language. Cabanis, himself a scientist, and Destutt de Tracy had frequently made significant use of this doctrine, but Degérando rejected it in a way that seems symptomatic of the end of the Enlightenment conception of a science. Some of the basis for his answers to these questions is evident in his answers to questions (2) and (5). "A well-made language," he maintained, "proclaims and presupposes a science that is already well advanced," thus adopting the Cartesian position on this issue rather than the Leibnizian. "We shall say that the great art of perfecting a science consists above all in making better observations and only *then* adopting a better language—i.e., one that is better suited to the observations that have been assimilated" (3.150–151). "The nomenclature of a science is related to the science itself as monuments are related to history: it preserves what is, but it can neither predict what is not yet nor unfold the future" (3.199).

Degérando resembled other *idéologues* more closely, however, in his view that improvements in philosophy—that is, in the analysis of sensations and ideas—did depend on a thorough examination of the natural language in which it was carried out. His own rather novel, never-realized scheme for accomplishing this was the construction of a philosophical dictionary.

> It has been recognized that we can have clear ideas only in possessing a well-made language, and that a language can be well-made only in so far as we have reformed the most familiar operations of the mind from the very outset, only in so far as we have grasped the relation that interconnects them all. That being the case, we have felt the need of remaking the language in its entirety and, in some sense, recommencing the education of the human mind. The surest and perhaps the most truly efficacious means of accomplishing this great project would be, I think, the formation of a philosophical dictionary truly worthy of the name—one, that is to say, that would in some sense be a genealogical tree of our ideas and of the signs we use. Such a dictionary would be a sequence of definitions strictly bound to one another. Each notion would be defined in it by showing how it was acquired, or at least how it should have been acquired. The mind would find itself naturally led to *create* the words rather than seeking merely to *explain* them to itself.… The dictionary I propose would have an aim altogether different [from that of ordinary dictionaries]. In it one would seek to explain not so much how we speak as how we think; the conventions of the language would be presented in it as *results*, not as *principles*. … [This dictionary would not be arranged in alphabetical order but] it would be a book, a history. The order of facts would be the only order observed in it. It would not … be designed to be consulted occasionally, but it would have to be the object of a connected reading.… A definition would never be offered in it except in accordance with one general rule—that of determining an idea by means of tracing

> it back to the ideas that must have preceded it in the age when language was instituted among men.… The dictionary would thus in some sense embrace the history of mankind and would serve as a natural introduction to all the sciences. The study of it would be necessary for all who wished to think well, and its formation would be one of the noblest undertakings of philosophy. (4.80ff.)

What little influence the *idéologues* had is somewhat more noticeable in British than in Continental philosophy of the nineteenth century, partly because of the interest of some of the Scottish commonsense philosophers in their work.

UNIVERSAL GRAMMAR

Among the most important and distinctive influences on the philosophy of language of the Enlightenment was the development of universal grammar. In the broadest sense of the term it was, as defined by James Harris (1709–1780), "that grammar which without regarding the several idioms of particular languages only respects those principles that are essential to them all" (*Hermes, or a Philosophical Inquiry concerning Language and Universal Grammar*, 1751, Book I, Ch. 2). Although it resembles the speculative grammar of the Middle Ages in some of its basic assumptions, universal grammar seems to have had an independent origin that may with reasonable accuracy be dated 1660, the year in which Arnauld and Claude Lancelot published the *Port-Royal Grammar—Grammaire générale et raisonnée*.

Lancelot was a grammarian in the scholastic (rather than humanist) tradition who provided the subject matter that Arnauld presented in accordance with Descartes's method, believing that he was thereby "developing in grammar a branch of Cartesianism" (p. 137). In grammar, as in every subject, the method consisted fundamentally in "beginning with the most general and simplest matters in order to proceed to the least general and most complex," and in the study of language one therefore had to begin with principles and elements common to all languages in order to proceed to the study of one's own and other particular languages. Thus, universal grammar, as Arnauld (and many of his successors) conceived of it, was an investigation of language (*langage*) designed as a propaedeutic to the study of languages (*langues*). In practice, however, the elements and principles of universal grammar tended to be those of traditional Latin grammar, and thus many of the so-called universal grammars, at least before Condillac, are of little value either to linguists or to philosophers.

More important than the content of the early treatises on universal grammar, however, is the connection they established between grammar and philosophy, especially in France, where universal grammar dominated linguistic studies for 150 years following the *Port-Royal Grammar*. César Chesneau Dumarsais (1676–1756), the foremost of the universal grammarians between Arnauld and Condillac, maintained that "grammarians who are not philosophers are not even grammarians" (*Véritables Principes de la grammaire*, 1729, 1.201), and men engaged in the inquiry at that time styled themselves *grammairiens-philosophes*. Dumarsais seems often to have thought of "philosophy" in connection with grammar as no more than a certain scientific attitude (see, for instance, his article "Grammairien" in the *Encyclopédie*), but he also held some of the views that were to serve as the basis for a more strictly philosophical grammar, as can be seen in his explanation that "grammar has a necessary connection with the science of ideas and reasoning because grammar treats of words and their uses and words are nothing but the signs of our ideas and our judgments" (1.201).

The *grammairiens-philosophes* began to be more markedly philosophers than grammarians beginning with the articles on grammatical topics in the *Encyclopédie*. Although Dumarsais was in general charge of them, several were written by philosophers such as Voltaire, Diderot, and Turgot; and the articles as a group contain less information on the announced topics than they do discussions of philosophical questions more or less vaguely associated with those topics. Nicolas Beauzée (1717–1789), one of the authors of those articles, defended the new approach to grammar in his *Grammaire générale* (1767): "Why should one think metaphysics out of place in a book on universal grammar? Grammar ought to expose the foundations—the general resources and the common rules—of language, and language is the exposition of the analysis of thought by means of speech. No aim is more metaphysical or abstract than that" (*Préface*, p. xvii).

In the works of Condillac, the emphasis was no longer on the propriety of taking a philosophical approach to grammar but rather on the fundamental importance of universal grammar as an inquiry serving the purposes of philosophy itself. In the "Motif des études" introducing his course of studies for the prince of Parma, Condillac described grammar as "a system of words that represents the system of ideas in the mind

when we wish to communicate them *in the order and with the interrelations we apperceive*." Consequently, as he remarked in the "Objet de cet ouvrage" preceding his *Grammaire* (1775), he regarded grammar "as the primary division of the art of thinking. In order to discern the principles of language we must observe how we think; we must seek those principles in the analysis of thought. But *the analysis of thought is quite complete in discourse*, with more or less precision depending on the greater or less perfection of languages and the greater or less exactness of mind on the part of those who speak them. This is what makes me think of languages as so many analytic methods."

Condillac's elevation of universal grammar to the status of a fundamental philosophical inquiry marked the beginning of a new phase in the development of universal grammar, but his view of it was no more than a natural consequence of the then well-established belief that the construction of an absolutely universal grammar for all languages was a feasible undertaking. If language depicted thought and all languages shared a set of elements and principles, then the study of those common elements and principles would provide a science of human thought.

For Condillac's successors, the *idéologues*, who took the analysis of sensations and ideas to be the whole of philosophy, universal grammar became *the* philosophical method. As Destutt de Tracy put it, "this science may be called *idéologie* if one attends only to the subject-matter, *universal grammar* if one has reference only to the method, or *logic* if one considers only the goal" (*Élémens d'idéologie*, 1801, 1.5). It was in accordance with this conception of philosophy that the traditional chairs of logic and metaphysics in the *écoles centrales* of France were replaced in 1795 with chairs of universal grammar, which "by offering instruction in the philosophy of language would serve as an introduction to the course in private and public morality" (*Élémens, Préface*, p. xxiii).

Destutt de Tracy devoted the second part of his *Élémens d'idéologie*, some 450 pages, to a presentation of universal grammar suited to the purposes of the new course. Although he did occasionally cite parallel examples in other European languages and stress the value of knowing several languages, his principal interest was in what might fairly be described as the analysis of ordinary French. As an *idéologue* he was committed to provide analyses that would in every case disclose the signified idea (or sensation). His first step in establishing the conditions for such an analysis was to insist that the unit of signification was not the word or phrase, no matter how complex, but only the proposition, the linguistic device expressive of a judgment. If one simply utters the words "Peter," "to be not tall," we say that it means nothing, it makes no sense, although if one merely changes the form of the verb so that one says "Peter is not tall," thereby expressing a judgment, we can discern in what he says signs of his having an idea of Peter and an idea of his height (2.29–33).

Thus Destutt de Tracy committed himself to providing ideological analyses only if the linguistic entity to be analyzed occurs within a proposition. Even so, locating the signified idea sometimes required considerable ingenuity, as in his attempt to analyze all "conjunctions" in such a way as to show not only two propositions related by each conjunction but also the idea signified by each.

> One can say as much regarding the conjunctions we use in asking questions, even though they might at first seem not to connect two propositions, because the first is suppressed. Thus when I say "how did you get in again?," "why did you leave?," I am really expressing these ideas: "I want to know [*Je demande*] *how* you got in again," "I want to know *why* you left." And when we unfold the sense of those conjunctions the result is: "I want to know *a thing that is the manner in which* you got in again," "I want to know *a thing that is the reason for which* you left." (2.136)

In Destutt de Tracy considerations of grammar were entirely subject to the demands of ideological analysis, as is plain not only in the structure of his final analyses above but also in his readiness, quite unusual even among *grammairiens-philosophes*, to revise the classifications of traditional grammar (treating the "adverbs" *comment* and *pourquoi* as "conjunctions") when philosophical considerations seemed to call for their revision. When his volume devoted to universal grammar appeared in 1803, the experimental substitution of universal grammar for logic and metaphysics in the schools had already been abandoned, along with the *idéologues*' highest hopes for revolutionizing philosophy.

In Germany, Christian Wolff and Lambert had taken notice of universal grammar, but the movement had no appreciable impact on philosophy. In England it affected mainly the work of Harris, of James Beattie (1735–1803), and of John Home Tooke (1736–1812), all of whom developed universal grammar far less as a philosophical than as a philological inquiry. By the beginning of the nineteenth century, universal grammar was rapidly going out of fashion as a branch of philosophy, even in France, despite the last efforts of the *idéologues*.

THE NINETEENTH AND TWENTIETH CENTURIES

BENTHAM

Jeremy Bentham (1748–1832) did almost all his work in philosophy of language during the last twenty years of his life, primarily under the influence of Locke, the *idéologues*, and the universal grammarians. In a passage distinctly reminiscent of Locke's call for the development of "the doctrine of signs," Bentham expressed his own conviction that "a demand exists for an entirely new system of *Logic*, in which shall be comprehended a *theory of language*, considered in the most general point of view" (*Works*, edited by John Bowring, 8.119–120). His belief in the importance of a theory of language within a system of logic seems to have made an impression on J. S. Mill and may mark the beginning of the return to a view of the interrelations of logic and language more like that prevailing in the later Middle Ages than like that of the eighteenth century.

Universal grammar constituted a part of Bentham's plan for fulfilling the demand he had recognized, and his account of its subject matter is modeled explicitly on what he considered the "pioneering" work of Tooke (see, for instance, 8.187–188). Unlike Tooke, however, Bentham was inclined to consider it a branch of philosophical rather than philological inquiry and echoed the *idéologues* in his claim that within "the field of universal grammar it is not enough for a man to look into the books that are extant on the subject of grammar, whether particular or universal—he must look into his own mind" (10.193). He also followed the *idéologues*, Degérando in particular, in rejecting Condillac's view that languages and analytic methods were identifiable. On the one hand, he held, the analysis of experience on the most primitive level was dependent on the prelinguistic faculty of attention; on the other hand, "every name, which is not, in the grammatical sense, a *proper* name, is the sign and result" of an act of *synthesis* rather than of analysis (8.75; 8.121–126). Bentham did, however, cite Lavoisier's Condillac-inspired reform of the language of chemistry as a prime example of the practical value of the philosophy of language (3.273).

On more strictly semantic questions Bentham occasionally wrote as if he had simply absorbed and to some extent clarified the doctrines of Locke's Book III, but when his most distinctive refinements of Locke are brought together, they mark a genuine advance in the history of semantics. He was in general agreement with Locke that "language is the sign of thought, an instrument for the communication of thought from … the mind of him by whom the discourse is uttered [to another mind].… The immediate subject of a communication made by language is always the state of the speaker's mind." The crucial doctrine of immediate signification, which in Locke had been obscured by his vacillating treatment of it, was explicated by Bentham as follows:

> In both these cases ["I am hungry," "That apple is ripe"], an object other than the state of my own mind is the subject of the discourse held by me, but in neither of them is it the *immediate* subject. In both of them the *immediate* subject is no other than the state of my own mind—*an opinion entertained by me in relation to the ulterior object or subject*. … [Language] may be the sign of … other objects in infinite variety, but of this object [the utterer's state of mind] it is always a sign, and it is only through this that it becomes the sign of any other object. (8.329–331)

Since, however, "communication may convey information purely, or information for the purpose of excitation" (8.301), the immediately signified state of the speaker's mind may be either "the state of the *passive* or *receptive* part of it, or the state of the *active* or *concupiscible* part" (8.329). Bentham described the use of language as a medium of communication as its "transitive" use. "By its *transitive* use, the collection of these signs is only the vehicle of thought; by its *intransitive* use, it is an instrument employed in the creation and fixation of thought itself." Consequently the transitive use of language "is indebted for its existence" to the intransitive use (8.228–229, 8.301).

Partly because he had begun with "thoughts" rather than Lockean ideas as the immediate significata of linguistic signs, and perhaps also because of the similar position taken in Destutt de Tracy's universal grammar, Bentham recognized not words but propositions as the elements of significance.

> If nothing less than the import of an entire proposition be sufficient for the giving full expression to any [but] the most simple thought, it follows that, no word being anywhere more than a fragment of a proposition, *no word is of itself the complete sign of any thought*. It was in the form of entire propositions that when first uttered, discourse was uttered.… *Words may be considered as the result of a sort of analysis*—a chemicological process for which, till of a

> comparatively much later period than that which gave birth to propositions, the powers of the mind were not ripe. (8.320–323; italics added)

"In language, therefore, the *integer* to be looked for is an entire proposition" (8.188).

Many of Bentham's predecessors, but especially Locke, had inveighed against the philosophers' tendency to "take words for things." Bentham's refinement and extension of this notion into a doctrine of "linguistic fictions" is his most distinctive contribution to philosophy of language. He began by taking the evidently unprecedented step of defining extralinguistic elements in terms of the functions of certain elements of language.

> An *entity* is a denomination in the import of which every subject matter of discourse, for the designation of which the grammatical part of speech called a noun-substantive is employed, may be comprised.… A *real* entity is an entity to which, on the occasion and for the purpose of discourse, existence is really meant to be ascribed.… A *fictitious* entity is an entity to which, though by the grammatical form of the discourse employed in speaking of it, existence be ascribed, yet in truth and reality existence is not meant to be ascribed.

Thus the noun-substantive "motion" in "that body is in motion" is the name of a fictitious entity, since "this, taken in the literal sense, is as much as to say—Here is a larger body, called a motion; in this larger body, the other body, namely, the really existing body, is contained." While he insisted that linguistic fictions stood in need of what he called exposition, he also maintained that they were contrivances "but for which language … could not have existence" (8.195–199).

The mode of exposition to which linguistic fictions were to be subjected was called paraphrasis, which "consists in taking the word that requires to be expounded—viz the name of a *fictitious* entity—and, after making it up into a *phrase*, applying to it another phrase, which, being of the same import, shall have for its principal and characteristic word the name of the corresponding *real* entity" (8.126–127). Since all words designative of nonphysical entities involved linguistic fictions, most of the work of philosophy, Bentham thought, would consist in such exposition of language.

In his *Principles of Morals and Legislation*, Chapter X, Bentham recommended a method of starting philosophical inquiry that was later to be employed and advocated by J. L. Austin. "I cannot pretend," Bentham said of his catalogue of motives in that chapter, "to warrant it complete. To *make sure* of rendering it so, the *only* way would be to turn over the dictionary from beginning to end; an operation which, in a view to perfection, would be necessary for more purposes than this" (italics added).

HUMBOLDT

The special historical importance of the work of Wilhelm von Humboldt (1767–1835) lies in the fact that it incorporates the transition from the eighteenth-century philosophy of language to the nineteenth-century science of linguistics. It does so not only in respect of the philosophical doctrines presented in it but also because Humboldt coupled those doctrines with empirical investigations of the sort he considered to be demanded by his philosophy of language.

His most important work—*Ueber die Kawi-Sprache auf der Insel Jawa* (published 1836–1839)—begins with the lengthy philosophical essay "Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einflusz auf die geistige Entwickelung des Menschengeschlechts." In it he developed his single most influential and original notion—that language is to be viewed not as a finished product but as a continuous process (*Sie selbst ist kein Werk, ergon, sondern eine Tätigkeit, energeia*), as the totality of instances of speech (or of the understanding of speech). Written words constitute language only when they are read and to the extent to which they are understood (*Gesammelte Schriften*, edited by A. Leitzmann, 7[No. 1].46 ff.). The rules of syntax and the individual words of a language are, then, the products of analysis, having real existence only insofar as they are embodied in instances of actual speech. Thus, as Destutt de Tracy and Bentham had observed from other points of view, "we cannot possibly conceive of language as beginning with the designation of objects by words and thence proceeding to their organization. In reality, discourse is not composed from words that preceded it. On the contrary, the words issued from the totality of discourse" (7[No. 1].72 ff.; cf. 7[No. 1].143).

The essential role played by language in fixing and organizing thoughts had been recognized long before Humboldt, but he extended that recognition into the bold new doctrine that language activity was the medium of contact between the mind and reality. "Man lives with the world about him principally, indeed … exclusively, as language presents it to him." Humboldt felt that this conception of language held the solution to the post-Kantian problems regarding subjectivity and objectivity.

> In speech the energy of the mind breaks a path through the lips, but its product returns through our own ears. The idea is [thus] translated into true objectivity without being withdrawn from subjectivity. Only language can do this.… [Moreover,] just as the particular sound mediates between the object and the man, so the whole language mediates between him and the nature that works upon him from within and without. He surrounds himself with a world of sounds in order to assimilate the world of objects. (7[No. 1]. 55ff.)

Somewhat as Hamann had done, Humboldt thus believed that philosophy reduced to the philosophy of language, and that he had "discovered the art of using language as a vehicle by which to explore the heights, the depths, and the diversity of the whole world" (letter to Wolfe, 1805).

The differences among natural languages were philosophically as well as scientifically important in Humboldt's view, and he was opposed to the prevailing eighteenth-century type of universal grammar, which achieved its universality at the expense of linguistic differences that happened not to fit the grammatical schema adopted by the grammarian-philosopher. He proposed instead, and provided examples of, a genuinely comparative grammar, insisting that the comparative grammarian avoid adopting the grammar of Latin or of his native language as the schema within which to organize the forms of other languages ("Ueber das Entstehen der grammatischen Formen und ihren Einflusz auf die Ideenentwickelung," in *Gesammelte Schriften* 4.285 ff.). Each natural language, he believed, was characterized by its own "inner form," expressive of the psyche of the nation within which it had developed and which it bound together. The distinctive inner form manifested itself in the root words as well as in the patterns of word combinations peculiar to the language. This doctrine, which powerfully influenced the development of linguistics, was in Humboldt's presentation of it little more than a consequence of the traditional semantic doctrine that speech reflected not objects but man's view of objects, coupled with the novel romanticist conviction that the reactions of men to the world around them were not everywhere the same. Not only were the grammatical differences among languages to be respected and studied in their own right, but the separate vocabularies were also to be reexamined with a view to discovering not interlinguistic synonymy (which, strictly speaking, was illusory) but the nuances of meaning that gave expression to different world views (7[No. 1].59 ff.; 89ff.; 190ff.).

Humboldt's immediate influence was not on philosophers but on other founders of the science of linguistics, particularly on Franz Bopp (1791–1867). What influence his work eventually had on philosophy of language, at any rate outside Germany, seems to have been transmitted indirectly through the work of nineteenth- and twentieth-century linguists.

JOHNSON

Alexander Bryan Johnson (1786–1867), the earliest American philosopher of language, was an isolated figure in the history of semantics. Locke and the Scottish common-sense philosophers strongly influenced his work, and he had learned something of the *idéologues* through Dugald Stewart's account of them. He seems, however, to have had little or no knowledge of his other predecessors and contemporaries. Johnson's work on language, published in three successive versions and under various titles in 1828, 1836, and 1854, went unnoticed for a hundred years and has had no appreciable influence since its republication during the 1940s. As the circumstances of his work would lead one to expect, it was unusual for its time both in its insights and in its mistakes.

The mistake that led to most of the others and to some of his principal insights as well occurred in his account of the semantics of words, in which he identified the signification(s) of a word with the thing(s) to which the word is applied. "Every word," he argued, "is a sound, which had no signification before it was employed to name some phenomenon." Consequently, "words have no inherent signification, but as many meanings as they possess applications to different phenomena. *The phenomenon to which a word refers, constitutes in every case, the signification of the word*" (*Treatise on Language*, Lectures VI and V; italics added).

The phenomena available as referents (or meanings) were exhaustively divided by Johnson into "sights, sounds, tastes, feels, smells, internal feelings, thoughts, and words" (Lecture XI). The word *table*, for example, signifies both a sight and a feel, "two distinct existences" bearing a single name. In this way "language implies a oneness to which nature conforms not in all cases," and men are prone to "make language the expositor of nature, instead of making nature the expositor of language" (Lecture III). Johnson made this common human failing his constant theme and provided several examples of philosophical and scientific difficulties that he felt were obviated by exposing a confusion of this sort as their source.

Philosophers, he suggested, might append to every "nominal unit that aggregates objects generically different" a capital letter—for instance, *S*, sight; *F*, feel—indicative of the phenomenon signified on each occasion. By that means David Hume, for example, might be seen to be announcing an "unconscious quibble," when he says, 'The table (*S*) which we see, seems to diminish (*S*) as we recede from it, but the real table (*F*) suffers no diminution (*F*).' The whole zest of the proposition consists in the sensible duality of each of the nominal units table and diminution.… We play bo-peep with words, by neglecting to discriminate the intellectually conceived oneness of diminution, and its physical duality" (*The Meaning of Words*, pp. 89–92).

In his account of the semantics of propositions Johnson remained faithful to the identification of meaning and referent with disastrous results, the most obvious of which was the confusion of meaningfulness (or meaninglessness) with truth (or falsity). "No proposition," he held, "can signify more than the particulars to which it refers" (Lecture VIII). He saw that one difficulty with this doctrine was that under it "the proposition that all men must die seems equivalent only to the proposition that all men have died." In his attempt to preserve the "universal application" of such general propositions, he adopted the indirect criterion of the failure of their negations to refer to any sensible particular. Thus "the proposition that all men will die, possesses a universal application for the reason that to say, some men will not die, *refers to no sensible particulars, and hence is insignificant*" (Lecture IX; italics added).

It was, however, this same approach to the semantics of propositions that led Johnson to develop and make critical use of a verifiability criterion of meaningfulness. Chemists, he remarked, had an indisputable right to "say simply that they can produce hydrogen gas, and oxygen, from water, and vice versa," but what they say instead is "that water is nothing but a combination of these gases. The assertion is true, *so long as it means* [*merely*] *the phenomena to which it refers*; but it produces wonder, because we suppose it has a meaning beyond the phenomena" (Lecture VII; italics added). Similarly, "if you inquire of an astronomer whether the earth is a sphere, he will desire you to notice what he terms the earth's shadow in an eclipse of the moon, the gradual disappearance of a ship as it recedes from the shore, &c. After hearing all that he can adduce in proof of the earth's sphericity, *consider the proposition* ['the earth is a sphere'] *significant of these proofs. If you deem it significant beyond them, you are deceived by the forms of language*" (Lecture VIII; italics added). In his verifiability criterion of meaningfulness and in his related discrimination of significant and insignificant questions (Lectures XIX ff.), Johnson anticipated some of the fundamental semantic principles of the pragmatists and positivists.

MILL

Many of the remarkable developments in semantics in the late nineteenth and early twentieth centuries took place under the influence of or in reaction against the doctrines of John Stuart Mill (1806–1873). He presented his "philosophy of language" (a designation he seems to have made current) in Book I and Chapters III–VI of Book IV of *A System of Logic* (1843), acknowledging the influence of the medieval logicians and Hobbes in particular, but also of Locke, Dugald Stewart, and others in the tradition of British empiricism. Like most of his empiricist predecessors in France and England, Mill believed that a philosophical inquiry into language had a high therapeutic value for philosophy itself, viewing metaphysics as "that fertile field of delusion propagated by language" (1.7.5).

By way of explaining his return to the practice of associating semantical inquiries with logic, Mill argued that since "language is an instrument of thought," not only in the reasoning process proper but in the antecedent operations of classification and definition, "logic … includes, therefore, the operation of Naming" (introduction, Sec. 7). It is not clear whether Mill intended to *identify* the ratiocinative use of language with naming or to claim that all language *stems from* the operation of naming, but he did revert to the tradition of considering "names" as the elements of his semantic theory. And since he took it to be obvious that "a proposition … is formed by putting together two names" (1.1.2), it seemed equally obvious that "the import of words [or names] should be the earliest subject of the logician's consideration: because without it he cannot examine into the import of propositions" (1.1.1).

NAMES. In his account of the import of names, Mill began by taking the unusual tack of defending "the common usage" against the view of "some metaphysicians," arguing that words are "names of things themselves, and not merely of our ideas of things." (Although there are passages in Hobbes and Locke, for example, that can be interpreted as expressions of that view, neither they nor, it seems likely, anyone else held quite the view Mill was criticizing.) "It seems proper," Mill claimed, "to consider a word as the *name* of that … concerning which, when we

employ the word, we intend to give information." When, however, "I use a name for the purpose of expressing a belief, it is a belief concerning the thing itself, not concerning my idea of it," even when the belief in question is one concerning some idea of mine (1.2.1).

A name, in Mill's adaptation of scholastic terminology, was said to *denote*, individually and collectively, the things of which it was the name, "the things of which it can be predicated." But, as Mill observed, "by learning what things it is a name of, we do not learn the *meaning* of the name" (1.2.5). A name happens "to fit" a given thing "because of a certain *fact*. … If we want to know what the fact is, we shall find the clue to it in the *connotation*" of the name (1.5.2). The connotation of the name is the "attribute" or set of attributes possession of which by a given thing is the fact in virtue of which the name fits the thing, and "the *meaning* of all names, except proper names [which have no meaning] and that portion of the class of abstract names which are not connotative [such as 'squareness,' which *denotes* a single attribute], resides in the connotation" (1.5.2). Mill recognized the connection of this distinction with the doctrine of denominatives (see the discussion of Anselm above), and in a note to 1.5.4 he indicated its relations to Hamilton's intension-extension distinction (see below).

CONNOTATION AND DENOTATION. Mill believed that the connotation-denotation distinction was "one of those which go deepest into the nature of language" (1.2.5). He made considerable use of it himself, and it played an important part in philosophical discussions for at least seventy-five years afterward. It is, however, a notoriously unclear distinction, especially in Mill's own treatment of it. With regard to denotation, for example, he claimed that a "concrete general name" such as "man" denotes Socrates—that is, is a name of, is predicable of that individual—but he claimed also that it denotes *the class* of which that individual is a member, which (at best) introduces a crucial ambiguity into the notion of denotation. With regard to connotation, the most serious difficulty centers on the notion of "attributes," which Mill suggested at one point was to be identified with what medieval logicians meant by "forms" (1.2.5n). In an evidently more careful account he declared that "the meaning of any general name is some outward or inward phenomenon, consisting, in the last resort, of feelings; and these feelings, if their continuity is for an instant broken, are no longer the same feelings, in the sense of individual identity.

What, then, is the common something which gives a meaning to the general name? Mr. [Herbert] Spencer can only say, it is the similarity of the feelings: and I rejoin, the attribute is precisely that similarity.… The general term *man* does not connote the sensations derived once from one man.… It connotes the general type of the sensations derived from all men, and the power … of producing sensations of that type" (2.2.4n). The only plausible interpretation of this doctrine seems to bring it very close to Hobbes's (or Locke's) actual account of words as signs of our ideas (despite Mill's attack on its weakest version: Words are *names* of our ideas), for in the end Mill's semantics of words appears to be founded on the familiar view that words are signs of extramental entities (denotation) only in virtue of being signs of mental entities of some sort (connotation). Mill surely would have recoiled at the suggestion that his doctrine of the connoted attribute as a "general type" of sensations committed him to an acceptance of extramental metaphysical entities.

After a detailed, ingenious investigation of the semantics of many-worded connotative concrete individual names (1.2.5), frequently discussed by his successors, Mill turned to the semantics of propositions. His account of the meaning of names and his view that the meaning of a proposition is a function of the meanings of the names that serve as its terms led naturally to his view that "when … we are analyzing the meaning of any proposition in which the predicate and the subject, or either of them, are connotative names, it is to the *connotation* of those terms that we must exclusively look, and not to what they *denote*." The view of Hobbes—that the predicate term is to be considered a name of whatever the subject term names—"is a mere consequence of the conjunction between the two attributes," the connotations of the two terms, and is adequate only in case both terms are nonconnotative names (1.5.2).

Thus, "all men are mortal" asserts that "the latter set of attributes *constantly accompany* the former set." And on the basis of the account of attributes introduced above, "we may add one more step to complete the analysis. The proposition which asserts that one attribute always accompanies another attribute, really asserts thereby no other thing than this, that one phenomenon always accompanies another phenomenon; in so much that where we find the latter, we have assurance of the existence of the former." He was, however, careful to note that "the connotation of the word *mortal* goes no farther than to the occurrence of the phenomenon at some time or other" (1.5.4).

When he came to discuss "real" (as opposed to "verbal") propositions, however, Mill disclosed that with respect to real propositions the account just cited was only one of "two formulas" in which "their import may be conveniently expressed." The account in terms of companion sets of attributes is suited to the view of real propositions "as portions of speculative truth." But they may be viewed also "as memoranda for practical use," and Mill's consideration of them in this light prefigured some elements of pragmatist theories of meaning. "The practical use of a proposition is, to apprise or remind us what we have to expect in any individual case which comes in the assertion contained in the proposition. In reference to this purpose, the proposition, All men are mortal, means that the attributes of man are *evidence of*, are a *mark of*, mortality; … that where the former are we … [should] expect to find the latter." The two formulas for expressing the import of real propositions are, Mill maintained, "at bottom equivalent; but the one points the attention more directly to what a proposition means, the latter to the manner in which it is to be used" (1.6.5).

Mill agreed with the majority of his philosophical contemporaries in deploring attempts to devise a formalized language for philosophy and suggesting that philosophers reform the natural languages for their uses. He was in a minority, however, in urging philosophers to have a healthy respect for natural languages. One of the "inherent and most valuable properties" of a natural language is "that of being the conservator of ancient experience"—"Language is the depository of the accumulated body of experience to which all former ages have contributed their part, and which is the inheritance of all yet to come." Consequently, "it may be good to alter the meaning of a word, but it is bad to let any part of the meaning drop" (4.4.6). Mill was emphatic about the special respect with which words of uncertain connotation were to be treated, and he laid down as a principle for the guidance of philosophers that "the meaning of a term actually in use is not an arbitrary quantity to be fixed, but an unknown quantity to be sought" (4.4.3). The attitude toward natural languages enjoined on philosophers by Mill was in part the attitude adopted by J. L. Austin and other twentieth-century philosophers of ordinary language.

PEIRCE AND THE PRAGMATISTS

In a tradition stemming from Locke, Charles Sanders Peirce (1839–1914) characterized logic "in its general sense" as "*semiotic* (σημει ωτική), the quasi-necessary, or formal, doctrine of signs" (*Collected Papers* 2.227) and went much further than anyone before him had tried to go toward the development of a completely general theory of signs. (Insofar as Peirce's semiotic deals with nonlinguistic signs, it lies outside the scope of this article, but his elaborate, varying terminology makes it difficult to present a single standard version of even that portion of the theory which is directly relevant to his treatment of linguistic meaning.)

Peirce seems sometimes to have thought of semiotic as a generalized version of the medieval trivium, describing its three branches as "pure grammar," "logic proper," and "pure rhetoric" (2.228–229). The first branch was to investigate the necessary conditions of meaningfulness, the second was to investigate the necessary conditions of truth, and the third was "to ascertain the laws by which in every scientific intelligence one sign gives birth to another, and especially one thought brings forth another" (2.229). These branches, with their subject matter somewhat differently described, were to become well known in twentieth-century philosophy under the designations "syntactics," "semantics," and "pragmatics" respectively—designations introduced by Charles W. Morris (*Foundations of the Theory of Signs*, 1938) and used extensively by Rudolf Carnap and others.

"Semiosis" was Peirce's name for an instance of signification, which he described as involving three principal elements: the *sign*, "something which stands to somebody for something in some respect or capacity" (its "ground"); the *object*, that for which the sign stands; and the *interpretant*, another sign, equivalent to or "more developed" than the original sign and caused by the original sign in the mind of its interpreter (2.228). The notion of the interpretant is the distinctive element in Peirce's general account of signification and the one that played the central role in his pragmatism (or "pragmaticism"), which he often described as consisting entirely in "a method for ascertaining the real meaning of any concept, doctrine, proposition, word, or other sign" (5.6).

Some of Peirce's predecessors had already suggested that the meaning of a word could be determined only on a given occasion of its occurrence within a propositional context, but in Peirce's the traditional primacy of the semantics of words over the semantics of propositions was so thoroughly overturned that his theory of linguistic meaning is almost exclusively a theory regarding the meaning of whole propositions. According to that theory, a proposition, like every other sign, has an object—some state of affairs, factual or otherwise. The *meaning* of a proposition, however, he identified not with its object but with one particular kind of effect of the proposition on

an interpreter, namely, its "*logical*" (as opposed to "emotional" or "energetic") *interpretant* (5.476).

Peirce's definitive account of the logical interpretant appeared in the 1905 paper "What Pragmatism Is" (5.411–434), in which he attempted as well to explain the distinctive and often misinterpreted "futuristic" aspect of pragmatist meaning theory.

> The rational meaning of every proposition lies in the future. How so? The meaning of a proposition [that is, its logical interpretant] is itself a proposition. Indeed, it is no other than the very proposition of which it is the meaning: it is a translation of it. But of the myriads of forms into which a proposition may be translated, what is that one which is to be called its very meaning? It is, according to the pragmaticist, that form in which the proposition becomes applicable to human conduct, … that form which is most directly applicable to self-control under every situation and to every purpose. This is why he locates the meaning in future time; for future conduct is the only conduct that is subject to self-control.

The only form of the proposition that would satisfy all these conditions was "the general description of all the experimental phenomena which the assertion of the proposition virtually predicts. For an experimental phenomenon is the fact asserted by the proposition that action of a certain description will have a certain kind of experimental result; and experimental results are the only results that can affect human conduct." Thus, as Peirce finally conceived of it, the meaning of a proposition is evidently to be explicated in the form of a true conditional with the original proposition as antecedent and, as its consequent, a conjunction of propositions constituting "the general description of all the experimental phenomena which the assertion of the [original] proposition virtually predicts."

Among the more striking problems in this account are (1) the difficulty of applying it to propositions other than those which occur within the context of an experimental science and (2) the fact that the meaning of a proposition is said to consist in other propositions, the meanings of which are presumably explicable in the same fashion, ad infinitum. Peirce was aware of both these problems. His response to (1) was generally to minimize the differences between the context of an experimental science and other contexts within which propositions occur, although he did occasionally, especially in his later writings, acknowledge the perhaps insuperable difficulties in employing this as a completely general theory of linguistic meaning. With regard to (2) Peirce was at first inclined to claim that a proposition (or any other sign) was, indeed, imperfectly significant if the series of its interpretants was finite ("Sign," in Baldwin's *Dictionary*).

Later, however, the notion of "the *ultimate* logical interpretant" was introduced. "The real and living logical conclusion" of the series of logical interpretants is an expectation (on the interpreter's part) of certain phenomena "virtually" predicted by the assertion of the original proposition. This expectation Peirce frequently referred to as "habit." "The deliberately formed, self-analyzing habit—self-analyzing because formed by aid of analysis of the exercises that nourished it—is the living definition, the veritable and final logical interpretant" (5.491; cf. 5.486). Habit, which Peirce sometimes described as a "readiness to act in a certain way under certain circumstances and when actuated by a given motive," was not itself a sign and so stood in no need of interpretants of its own.

It was on this very point that Peirce thought his own doctrine differed from that of William James (1842–1910). "In the first place," he wrote, "there is the pragmatism of James, whose definition differs from mine only in that he does not restrict the 'meaning,' that is, the ultimate logical interpretant, as I do, to a habit, but allows percepts, that is, complex feelings endowed with compulsiveness to be such" (5.494). James's own definition of "pragmatism" in Baldwin's *Dictionary* identified it as "the doctrine that *the whole meaning* of a conception *expresses itself in practical consequences* either in the shape of conduct to be recommended or in that of experiences to be expected, if the conception be true" (italics added), but in doing so he evidently believed he was promulgating "Peirce's principle … that *the effective meaning* of any philosophic proposition can always be *brought down to* some *particular consequence*, in our *future practical experience*, whether active or passive" (*Collected Essays and Reviews*, edited by R. B. Perry, p. 412; italics added). James's conception of pragmatism as a theory of meaning (and of truth) was, however, unquestionably broader and less carefully qualified than Peirce's and may fairly accurately be summarized in his own characteristic observation that concepts and propositions "have, indeed, no meaning and no reality if they have no use. But if they have any use they have that amount of meaning. And the meaning will be true if the use squares well with life's other uses" (*Pragmatism*, p. 273).

Pragmatism first became generally known in the form given it by James and in the still wider "humanism"

of F. C. S. Schiller, and it was in those forms that it was subjected to intense criticism at the beginning of the twentieth century by F. H. Bradley and G. E. Moore, among others. Peirce's more intricate and interesting theory of meaning was not really considered in its own right until some years afterward, perhaps beginning with the publication of C. K. Ogden and I. A. Richards's very influential *The Meaning of Meaning* in 1923, in which some ten pages were devoted to an exposition of Peirce's semiotic.

At the same time a pragmatist theory of meaning more complex and no less broad than James's was being developed in the "instrumentalism" of John Dewey (1859–1952). Dewey discussed meaning of every imaginable sort and in countless contexts, with the result that it is difficult to elicit from his many writings a genuinely representative doctrine specifically of linguistic meaning. Perhaps the least misleading single source is Chapter 5 of his *Experience and Nature*, first published in 1925. His position there was as follows:

> The sound, gesture, or written mark which is involved in language is a particular existence. But as such it is not a *word*, and it does not become a word by declaring a mental existence; it becomes a word by gaining meaning; and it gains meaning when its use establishes a genuine community of action.… Language and its consequences are characters taken on by natural interaction and natural conjunction in specified conditions of organization.… Language is specifically a mode of interaction of at least two beings, a speaker and a hearer; it presupposes an organized group to which these creatures belong, and from whom they have acquired their habits of speech. It is therefore a relationship, not a particularity.… The meaning of signs moreover always includes something common as between persons and an object. When we attribute meaning to the speaker as *his* intent, we take for granted another person who is to share in the execution of the intent, and also something, independent of the persons concerned, through which the intent is to be realized. Persons and thing must alike serve as means in a common, shared consequence. This community of partaking is meaning.

Even when, as in these passages, Dewey seems to have been considering linguistic meaning specifically, there is a real possibility that his intentions were much broader, for his conception of language was itself considerably broader than that of most philosophers. He was, for example, prepared to say that "because objects of art are expressive, they are a language. Rather they are many languages. For each art has its own medium and that medium is especially fitted for one kind of communication.… The needs of daily life have given superior practical importance to one mode of communication, that of speech" (*Art as Experience*, 1935, p. 106).

Pragmatist theories of meaning, beginning with Peirce's 1878 paper "How to Make Our Ideas Clear," are alike in little more than their tendency to associate the meaning of a proposition with the conditions of its verification, but in that respect they may be said to have inaugurated twentieth-century developments of empiricist and operationalist theories of meaning.

FREGE

The contributions of Gottlob Frege (1848–1925) to logic, philosophy of mathematics, and semantics were largely unappreciated at the time of their publication, primarily during the last quarter of the nineteenth century. Their influence (direct or indirect) on recent philosophy has been so great, however, that Frege might fairly be characterized as the first twentieth-century philosopher. In his *Begriffsschrift* (1879) he developed "a formalized language of pure thought modeled on the language of arithmetic," which has been recognized as the first really comprehensive system of formal logic. In his other two major works, *Die Grundlagen der Arithmetik* (1884) and *Die Grundgesetze der Arithmetik* (1893–1903), he tried to show "that arithmetic is founded solely upon logic."

Philosophical problems encountered by Frege in those highly technical undertakings were explored by him in several papers that have had a wider influence than his books have had. As the topics of the books might lead one to expect, his philosophical papers are concerned almost exclusively with one or another aspect of systems of signs. Of these papers, the one that has had most effect on the development of semantics is "Ueber Sinn und Bedeutung" (1892, translated in P. T. Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege*, 1952), although the doctrine presented in it may prove to be historically less important than the doctrine of "functions" developed in other papers.

There is a broad and not wholly misleading similarity between Frege's distinction of sense (*Sinn*) and reference (*Bedeutung*) and such distinctions as comprehension-extension (Arnauld), intension-extension (Hamilton), connotation-denotation (Mill), depth-breadth (Peirce). It seems possible, however, that Frege

developed his distinction independently; in any case, the details of his doctrine are quite novel. Most important, perhaps, was his discovery of special contexts rendering the application of any such distinction problematic.

SENSE AND REFERENCE. Frege's development of the doctrine of sense and reference began, characteristically, in a consideration of the relation of identity: "=." He noted that "$a = a$ and $a = b$ are obviously statements of different cognitive value," which they would not be if we were to take the relation to hold "between that which the names '*a*' and '*b*' designate" or refer to (Geach and Black, p. 56). Consequently, "it is natural now, to think of there being connected with a sign (name, combination of words, letter), besides that to which the sign refers, which may be called the reference of the sign, also what I should like to call the *sense* of the sign, wherein the mode of presentation is contained.… The reference of 'evening star' would be the same as that of 'morning star,' but not the sense" (p. 57).

Frege first applied his distinction to proper names, by which he meant any "designation of a single object." In keeping with Arnauld's similar distinction but in opposition to Mill's, Frege ascribed sense as well as reference to such designations. "A proper name (word, sign, sign combination, expression) *expresses* its sense, *stands for* or *designates* its reference. By means of a sign we express its sense and designate its reference" (p. 61). The sense of "Aristotle" "might, for instance, be taken to be the following: the pupil of Plato and the teacher of Alexander the Great" (p. 58n).

Certain expressions, such as "the least rapidly convergent series," have a sense, he maintained, but no reference at all. "In grasping a sense, one is not certainly assured of a reference" (p. 58). An expression that has a reference "must not be taken as having its ordinary reference" when "standing between quotation marks" (pp. 58–59). Such observations had been made before, but Frege seems to have been the first to try to show what that extraordinary reference might be and, more important, to recognize that many different linguistic contexts affected the reference of expressions included within them, especially indirect discourse and subordinate clauses following such verbs as "hear," "conclude," "perceive," and "know." He claimed, for example, that "in reported speech, words … have [not their customary but] their *indirect* reference," and that "the indirect reference of a word is … its customary sense" (p. 59). His account of the effect of such contexts on reference has not been widely accepted, but the problems raised by it have stimulated the widespread interest of twentieth-century philosophers in such now familiar topics as synonymy, opacity of reference, Leibniz's Law, and what, following Franz Brentano (see below), have come to be called intentional contexts.

Frege was concerned with saying what sort of entities sense and reference were. In the case of a proper name his description of the reference was relatively unproblematic: "a definite object (this word ['object'] taken in the widest range)"—(p. 57)—so wide that "2 + 2" and "4," for example, were two proper names with one and the same "object" as their reference. Regarding the sense of a proper name, he found it easier to say what it was not: "The reference of a proper name is the object itself which we designate by its means; the idea, which we have in that case, is wholly subjective; in between lies the sense, which is indeed no longer subjective like the idea, but is yet not the object itself" (p. 60). Thus, there is "an essential [subjective-objective] distinction between the idea and the sign's sense." Frege seems not to have completely depsychologized the notion of the sense of the sign, however, since he suggested that it may be an element in humankind's "common store of thoughts which is transmitted from one generation to another" rather than "a part or a mode of the individual mind" (p. 59).

In Frege's discussion of the sense and reference of declarative sentences, the doctrine of the sense was relatively straightforward while the account of the reference became problematic. A sentence, he held, "contains a thought," and by "a thought" he meant "not the subjective performance of thinking but its objective content, which is capable of being the common property of several thinkers" (p. 62 and note). The two sentences "the morning star is a planet" and "the evening star is a planet" contain different thoughts, as may be seen from the fact that "anybody who did not know that the evening star is the morning star might hold the one thought to be true, the other false. The thought, accordingly, cannot be the reference of the sentence, but must rather be considered its sense" (p. 62). We are content to consider only the sense of sentences as long as we are not concerned to judge of their truth or falsity, but "in every judgment, no matter how trivial, the step from the level of thoughts to the level of reference (the objective) has already been taken" (p. 64). What we seek in judgment is the truth-value of the sentence. "We are therefore driven into accepting the truth-value of a sentence as constituting its reference.… Every declarative sentence concerned with the reference of its words is therefore to be regarded as a proper name,

and its reference, if it has one, is either the True or the False" (p. 63).

As his use of the phrase "driven into accepting" indicates, Frege was well aware that this was a startling doctrine of the semantics of sentences. Much of the remainder of his paper on sense and reference was devoted to considerations that he felt tended to support it, among them Leibniz's Law; for "what else but the truth-value could be found, that belongs quite generally to every sentence if the reference of its components is relevant, and remains unchanged by substitutions of the kind in question?" (p. 64).

The doctrine that sentences are proper names, whether or not of the True and the False, had an important negative effect in that its rejection by Wittgenstein (see, for instance, *Tractatus* 3.143) and by Bertrand Russell under Wittgenstein's influence (see, for instance, "Logical Constructions," Lecture I) helped to shape the course of philosophy of language in the twentieth century.

It is quite likely, however, that Frege's assimilation of declarative sentences to proper names was not quite so thorough or simple as his presentation of it in "Ueber Sinn und Bedeutung" suggests. Some of his remarks in an earlier paper, "Funktion und Begriff" (1891), at least raise the possibility that he may have denied proper-name status to sentences actually being used in making assertions (rather than considered as examples). In order to make what he took to be the indispensable "separation of the act from the subject-matter of judging" he introduced his assertion sign—"⊢"—"so that, e.g., by writing

$$\vdash 2 + 3 = 5$$

we *assert that* 2 + 3 equals 5. Thus here we are *not just writing down a truth-value*, as in

$$2 + 3 = 5,$$

but also at the same time *saying that it is the True*." And in a note to this passage he maintained that "'⊢ 2 + 3 = 5' does not *designate* [that is, refer to] anything; it *asserts* something" (p. 34; italics added).

MAUTHNER

Of the several late nineteenth-century philosophers writing in German whose work centered on a concern with language, the most unusual was Fritz Mauthner (1849–1923). His principal work, *Beiträge zu einer Kritik der Sprache*, fills three large volumes and went through three editions, the first in 1901–1902. In his thoroughgoing attempt to transform all philosophy into philosophy of language, in his criticisms of Kant, and in his penchant for paradox he resembled Hamann, whom he admired, and also, to some extent, Humboldt. He seems, however, to have been most powerfully influenced by the positivism of Ernst Mach and especially by Hume's skepticism, adopting as his philosophical watchword "Back to Hume!"

Part of what Mauthner meant by that is apparent in the epistemological doctrine on which he founded his critique of language: "Our memory [with which he identified our knowledge] contains nothing but what our poor fortuitous senses [*Zufallsinne*] have presented to it" (*Beiträge* 3.536). By calling our senses "fortuitous" he was calling attention to the fact that if we had been otherwise equipped with senses, we might have framed a very different view of the world. Language, however, depicts not the world but a world view. Therefore, any attempt to infer propositions regarding reality from facts of language is a form of "word-superstition."

Moreover, each man's individual senses present a world view unique in certain ultimately undeterminable respects, and so communication by means of language, even if it purports to be no more than an exchange of views, is fundamentally illusory. "No man knows the others.… With respect to the simplest concepts we do not know of one another whether we have the same representation associated with one and the same word."

From such avowedly Lockean observations Mauthner drew the typically paradoxical conclusion that "by means of language men have made it forever impossible to get to know one another" (1.54). Thus he characterized language as "nothing other than just the community or the mutuality of world-views." It is not a tool for the communication or acquisition of knowledge; indeed, it is not a tool or an object of any sort but merely a practice, a use. And "because it is no object of use but use itself, it perishes without use" (1.24). But of all Mauthner's many characterizations of language the one most suggestive of distinctively twentieth-century attitudes is this: "Language is merely an apparent value [*Scheinwert*], like a rule of a game [*Spielregel*], which becomes more binding as more players submit to it, but which neither alters nor comes into contact with [*begreifen*] the world of reality" (1.25).

Philosophy, in Mauthner's view, had to become a critique of language if it was to be anything at all, and in that guise its principal function was to be therapeutic. "Philosophy … cannot wish to be anything more than critical attention to language. Philosophy can do no more with

respect to the organism of language or of the human spirit than can a physician with respect to the physiological organism. It can attentively observe and designate the developments with names" (1.657). "If I want to ascend into the critique of language, which is at present the most important undertaking of thinking mankind, I must do away with language behind me and before me and in me, step by step—I must break in pieces each rung of the ladder as I tread on it" (1.1–2). It comes as no surprise to learn that the end of this therapeutic process was to be silence, the silence of mystical contemplation.

In the course of the long process, however, Mauthner found occasion to make many insightful observations on traditional problems of the philosophy of language. As several of his contemporaries and immediate predecessors had done, he recognized not the word but the sentence (*Satz*) as the unit of meaningfulness and described the meaning of the word as a function of its use in a given sentence. Another position that was not new but to which he gave an especially forceful presentation was the rejection of the view "that because there is a word, it must be a word *for* something; that because a word exists there must exist something real corresponding to that word." This form of word superstition he regarded as "mental weakness" (2nd ed., 1.159).

It is probably only coincidence, but the name theory of linguistic meaning against which Mauthner inveighed bears a strong resemblance in some respects to the theories of Alexius Meinong and Edmund Husserl then being published and to the early views of Russell and Wittgenstein. Mauthner also opposed efforts at universal grammar (such as some of Brentano's followers were then engaged in) and mathematical logic, maintaining that all formalization of language obliterated or obscured far more than it clarified. Thus, he noted that "if someone says 'cheese is cheese' … this utterance is *not* an instance of the general formula 'A = A'" (3.366), a formula "so empty that outside logic it must arouse the suspicion of insanity" (*Wörterbuch der Philosophie*, article "A = A").

Perhaps more than any other philosopher of language Mauthner had an appreciation of the history of the subject; at one time, in fact, he planned a fourth volume of his *Beiträge* that was to present the approach to the critique of language throughout the history of philosophy. Even as it stands, however, his work is filled with references to his predecessors and evaluations of their work from the viewpoint of the critique of language. Aristotle, for example, comes off badly, but Locke ranks very high. Indeed, Mauthner took "the English" to task for abandoning the work of Locke, for failing to see that "the content of their famous 'understanding' is simply the dictionary and grammar of human language" (*Beiträge* 3.535).

Mauthner's own effect on the history of the philosophy of language is still difficult to assess. Wittgenstein certainly knew of his work (see, for instance, *Tractatus* 4.0031). Whether or not Wittgenstein's turn in the direction of some of Mauthner's doctrines in *Philosophical Investigations* was coincidence or derived in part from Mauthner's influence remains an open question.

HUSSERL AND MEINONG

The students of Franz Brentano (1838–1917), among whom were Husserl, Meinong, Anton Marty, and Kazimierz Twardowski, were alike at least in taking Brentano's concept of intentionality as a point of departure in their own philosophizing. Brentano had introduced intentionality in his *Psychologie vom empirischen Standpunkt* (1874) as the differentia of "mental states," a characteristic "which the schoolmen of the middle ages called the intentional (or mental) inexistence of an object and which we … describe as *the relation to a content*, or *the direction to an object* (by which we need not understand a reality), or *an immanent objectivity*. Every mental state possesses in itself something that serves as an object, although not all possess their objects in the same way" (*Psychologie* 2.1.5; italics added).

The "intending" of an object by a mental state, the "directedness" of a mental state, bears a close enough resemblance to what is called significance in other contexts that much of what Brentano and his followers had to say in working out their central doctrine of intentionality has some relevance to semantics, broadly conceived. More specifically, the notion of intentionality underlies the considerable discussion in semantics of "intentional contexts," produced as a result of the ordinary use of such "intentional words" as "believe," "want," and "ascribe." For present purposes, however, our attention is confined to what Brentano's two best-known students, Husserl and Meinong, had to say expressly about language and linguistic meaning. The doctrines of both men passed through several stages of development and contain many complexities, only a few of which can be noted here.

The philosophy of language of Edmund Husserl (1859–1938) was developed at various places in his work but is concentrated in the first and fourth essays in his *Logische Untersuchungen* (1900–1901; rev. ed., 1913–1921). The first, titled "Expression and Meaning" (2.23–105), was designed partly as a general preparation for intensive work in phenomenology as conceived by

Husserl. It opens with an investigation of signs in general and proceeds to the consideration of expressions, signs that may be said to have meanings (*Bedeutungen*) and not merely to indicate something. The three ingredients of meaningfulness, or "the meaning-situation" are (1) a "meaning-endowing act," or "meaning-intention" on the part of the producer of the expression, which may be associated with a "meaning-fulfilling act" on the part of an interpreter of the expression; (2) the *content* of these acts, or the *meaning* of the expression; (3) the *object* of these acts, or, in Husserl's broader terminology, the *objectivity* that is meant by the expression. To talk about what is expressed by a given expression may be to talk about any one of these ingredients. (To some extent Husserl avoided the usual sort of technical distinctions among semantical relations, specifically rejecting Frege's sense-reference distinction as a violation of the ordinary use of the words *Sinn* and *Bedeutung*, words which Husserl used interchangeably [2.53].)

Somewhat more precisely, an expression used in ordinary circumstances for purposes of communication may be described as "manifesting" the psychical experience of its producer—that is, the meaning-endowing act—which is a necessary condition of its status as an expression. This manifesting function of an expression would, however, be lacking in the case of an expression used in an unoverheard monologue. The manifesting and the more strictly expressing functions differ also in that, for example, the expression "the three altitudes of a triangle intersect in a point" manifests a distinct mental state or act each time it is used in ordinary circumstances for purposes of communication, although what it expresses, in the stricter sense, remains the same on all occasions of such use.

Some of Husserl's main points in "Expression and Meaning" are summarized in sections dealing with "equivocations" associated with discussions of meaning and meaninglessness (2.52–61). "A meaningless expression is, properly speaking, not an expression at all." Thus "green is or" (Husserl's example) only gives the appearance of an expression (2.54). Meaningfulness, however, entails reference (*Beziehung*) to an object, regardless of whether that object exists or is "fictive." "Consequently, to use an expression with sense and to refer to an object (to present an object) are one and the same." Nevertheless, Husserl was careful to point out, the object of an expression is not to be confused with its meaning (2.54). As a result, "objectlessness" of an expression is not "meaninglessness" (where "objectlessness" indicates only the lack of a real object). Neither the name "golden mountain" nor the name "round square" is meaningless, although both are objectless, the second one necessarily so (2.55). After a rather obscure passage (2.56–57) in which Husserl was evidently criticizing (without mentioning) pragmatism for identifying meaning with meaning-fulfilling acts, he devoted an entire section to the criticism of Mill's doctrine of connotation and denotation, with particular attention to Mill's view of "non-connotative names" as meaningless. A proper name, Husserl objected, is not a mere sign but an expression. It can, like any other expression, function as a mere sign—for instance, in a signature—but it ordinarily does much more. He felt that if Mill's distinction between what a name denotes and what it connotes were carefully separated from the merely related distinction between what a name names and what it means, some of the confusion in Mill's doctrine would be dissipated (2.57–61).

In his fourth Logical Investigation, "The Distinction of Independent and Dependent Meanings and the Idea of a Pure Grammar" (2.294–342), Husserl pursued the analysis of meaning undertaken in the earlier treatise. Of most historical interest is his attempted refurbishing of the Enlightenment project of a universal grammar, an enterprise furthered by Anton Marty (1847–1914), another of Brentano's students, in his *Grundlegung der allgemeinen Grammatik und Sprachphilosophie* (1908). In his treatise Husserl developed a notion of "pure logic" as "the pure formal theory of *meanings*" and insisted that we could not understand the functioning of even our own language if we did not first construct a "pure-logical grammar," the subject matter of which would be the "ideal form" of language. At a later stage of his career, however, Husserl abandoned this "ideal-language" approach to considerations of semantics and syntax and urged the return to living history and actual speech—the return to the *Lebenswelt*—for the materials of philosophy.

Husserl's influence in all respects has been felt more strongly in Europe than in England and America. Some of his work in philosophy of language has been investigated and developed further by, among others, Maurice Merleau-Ponty.

Alexius Meinong (1853–1920) developed his theory of linguistic meaning as an integral part of the "theory of objects" in which he worked out his version of the doctrine of intentionality. His most complete presentation of it may be found in his *Ueber Annahmen* (1902; rev. ed., 1910).

Meinong began, in the traditional way, by developing a semantics of words. His assimilation of it to his theory of objects gave rise to no particularly novel features.

"Whoever happens to pronounce the word 'sun,'" he declared, "normally gives *expression* [*Ausdruck*] thereby, whether or not he wishes to do so, that a definite presentation—it may be a presentation of perception or one of imagination—is taking place within him. What kind of presentation it is is determined principally on the basis of what is presented in it—i.e., its object—and this object is precisely that which the word 'sun' *refers to* [*bedeutet*]" (*Ueber Annahmen*, pp. 19–20; italics added). He summed up his account of the "expression" and "reference" of words (which he presented explicitly as opposed to Husserl's doctrine of *Ausdruck* and *Bedeutung*) by saying that "a word always 'refers to' the object of the presentation that it 'expresses' and, conversely, expresses the presentation of the object that it refers to" (p. 20). (The obvious similarities to Frege's doctrine of sense and reference, even as to the same unusual use of *bedeuten*, may be coincidental. There are, of course, clear differences as well, especially as regards Frege's treatment of "sense" and Meinong's treatment of "expression.") Meinong concluded his rather brief account of the meaning of words by refining his original distinction to the point of recognizing a "secondary" as well as the "primary" expression and reference described above (pp. 20–23).

He then undertook to apply his "antithesis of expression and reference" to the semantics of sentences, and he first applied it in an effort to provide a more satisfactory criterion of sentencehood than that provided by traditional grammar. The phrase "the blue sky" and the sentence "the sky is blue" have, he maintained, one and the same object as their reference. If, however, "I say 'the sky is blue,' I thereby *express* an opinion [*Meinung*], a judgment, that can in no way be gathered from the words 'the blue sky'" (p. 25; italics added). The phrase expresses the kind of experience described by Meinong as the pure presentation or idea, the *Vorstellung* proper, while the sentence expresses a different fundamental kind of experience, the judgment (*Urteil*). The judgment differs from the pure presentation giving rise to it in two respects that might be described as "intentional"—conviction and a determinate position as regards affirmation and negation (p. 2). Sentences, he claimed, might also be used to express "assumptions" (*Annahmen*), which, because they have to be either affirmative or negative assumptions, share the second defining characteristic of judgments but lack the first, conviction (p. 4).

Meinong's most important contribution to semantics, partly because of its effect on the development of Russell's theory of descriptions, was his doctrine of "objectives," particularly in his application of it in the treatment of negative sentences, which he recognized to be crucial cases for his doctrine. Suppose that a magistrate judges that on a given occasion there was no disturbance of the peace. On the Brentano-Meinong view of mental states, there must be an object of that judgment. Putting it another way, there must be a reference for the sentence "there has been no disturbance of the peace," which expresses the magistrate's judgment. It cannot be the disturbance of the peace on the occasion in question, for, by the hypothesis, there is no such object. According to Meinong it can, however, be "the non-existence of a disturbance of the peace" or "that there has been no disturbance of the peace." Meinong held that it makes no sense to say that that nonexistence *exists*, but we may say that "it is the case." This entity, the being of which is being the case, is the "objective" to which the sentence refers. The objective *may* be a fact—if, for example, it is a fact that there has been no disturbance of the peace—but false judgments also have their objectives (2nd ed., p. 43). That regarding which the judgment is made—a disturbance of the peace—is the object (proper) of the judgment; what is judged in it—that there has been no disturbance of the peace—is its objective (2nd ed., p. 52).

The objectives of negative sentences and the objects of denials of existence, such as "a *perpetuum mobile* does not exist," "must have properties, and even characteristic properties, for without such the belief in non-existence can have neither sense nor justification; but the possession of properties is as much as to say a manner of being [*Sosein*]," which of course is not to be confused with existence. "In this sense 'there are' also objects that do not exist, and I have expressed this in a phrase that, while rather barbarous, I am afraid, is hard to improve upon—viz. 'externality [*Aussersein*] of the pure object'" (p. 79). Meinong believed that he had formulated an important principle in this doctrine, "the principle of the independence of manner of being from existence," which he illustrated and summarized in the following famous passage: "Not only is the often cited golden mountain golden, but the round square, too, is as surely round as it is square.… To know that there are no round squares, I have to pass judgment on the round square.… Those who like paradoxical expressions can therefore say: there are objects of which it is true that there are no objects of that kind" (*Ueber Gegenstandstheorie*, 1904, pp. 7ff.). Meinong's influence on the development of semantics is best exhibited in Russell's series of articles on him in *Mind*, 1899–1907.

See also Semantics, History of [Addendum].

Bibliography

There is at present no thorough general study of the history of semantics. Much relevant information may be found, however, in works on the history of logic, particularly in I. M. Bocheński, *A History of Formal Logic*, translated by Ivo Thomas (Notre Dame, IN, 1961); Philotheus Boehner, *Medieval Logic* (Manchester, U.K.: Manchester University Press, 1952); L. M. de Rijk, *Logica Modernorum*, Vol. I (Assen: Van Gorcum, 1962); William and Martha Kneale, *The Development of Logic* (Oxford: Clarendon Press, 1962); Benson Mates, *Stoic Logic* (Berkeley: University of California Press, 1953); E. A. Moody, *The Logic of William of Ockham* (New York and London: Sheed and Ward, 1935), and *Truth and Consequence in Mediaeval Logic* (Amsterdam: North-Holland, 1953); and Carl Prantl, *Geschichte der Logik im Abendlande*, 4 vols. (Leipzig: Hirzel, 1855–1870).

Surveys less thorough than this entry are contained in Ch. 7, "Language," of Janet and Séailles's *History of the Problems of Philosophy*, translated by A. Monahan and edited by H. Jones (New York: Macmillan, 1902), and in Ch. 1, "The Problem of Language in the History of Philosophy," of Vol. I, *Language*, of Ernst Cassirer's *The Philosophy of Symbolic Forms*, translated by R. Manheim (New Haven, CT: Yale University Press, 1953).

The following selection of books and articles is arranged in four parts corresponding to the four parts of the article.

ANTIQUITY

Allen, W. S. "Ancient Ideas on the Origin and Development of Language." *Transactions of the Philosophical Society* [of Great Britain] (1948): 35–60.

Billiesich, F. *Epikurs Sprachphilosophie*. Landskron, 1912.

De Lacy, Estelle. "Meaning and Methodology in Hellenistic Philosophy." *Philosophical Review* 47 (1938): 390–409.

De Lacy, Phillip. "The Epicurean Analysis of Language." *American Journal of Philology* 60 (1939): 85–92.

Demos, Raphael. "Plato's Philosophy of Language." *Journal of Philosophy* 61 (1964): 595–610.

Deuschle, Julius. *Die Platonische Sprachphilosophie*. Marburg, 1852.

Goldschmidt, Victor. *Essai sur le "Cratyle."* Paris: Champion, 1940.

Lersch, Laurenz. *Die Sprachphilosophie der Alten.* 3 vols. Bonn: H.B. König, 1838–1841.

Luce, J. V. "The Theory of Ideas in Plato's *Cratylus*." *Phronesis* 10 (1965): 21–36.

Moravcsik, Julius. "Being and Meaning in the *Sophist*." *Acta Philosophica Fennica*, Fasc. 14 (1962): 23–78.

Robinson, Richard. "A Criticism of Plato's *Cratylus*." *Philosophical Review* 65 (1956): 324–341.

Robinson, Richard. "The Theory of Names in Plato's *Cratylus*." *Revue internationale de philosophie* 9 (1955): 221–236.

Schoemann, G. F. *Die Lehre von den Redetheilen nach den Alten*. Berlin, 1862.

Steinthal, Hugo. *Die Geschichte der Sprachwissenschaft bei den Griechen und Römern*. Berlin: F. Dümmlers, 1863.

Weltring, G. *Das* Σημεῖον *in der Aristotelischen, stoischen, Epikureischen, und skeptischen Philosophie*. Bonn, 1910.

THE MIDDLE AGES

Arnold, Erwin. "Zur Geschichte der Suppositionstheorie, die Wurzeln des modernen Subjektivismus." *Symposion, Jahrbuch für Philosophie* 3 (1952): 1–34.

Boehner, Philotheus. "A Medieval Theory of Supposition." *Franciscan Studies* 18 (1958): 240–289.

Boehner, Philotheus. "Ockham's Theory of Signification." *Franciscan Studies* 6 (1946): 143–170.

Boehner, Philotheus. "Ockham's Theory of Supposition and the Notion of Truth." *Franciscan Studies* 6 (1946): 261–292.

Boehner, Philotheus. "Ockham's Theory of Truth." *Franciscan Studies* 5 (1945): 138–161.

Carreras y Artau, Joaquin. *De Ramón Lull á los modernos ensayos de formación de una lengua universal*. Barcelona: Consejo Superior de Investigaciones Científicas, Delegación de Barcelona, Instituto Antonio de Nebrija, 1946.

Grabmann, Martin. "Die Entwicklung der mittelalterlichen Sprachlogik (Tractatus de Modis Significandi)." In his *Mittelalterliches Geistesleben.* 3 vols. Munich: Hueber, 1926. Vol. I, Ch. 4.

Grabmann, Martin. *Thomas von Erfurt und die Sprachlogik des mittelalterlichen Aristotelismus; Sitzungsberichte der Bayerischen Akademie der Wissenschaften* (Philosophisch-historische Abteilung), Jahrgang 1943, Heft 2.

Heidegger, Martin. *Die Kategorien- und Bedeutungslehre des Duns Scotus*. Tübingen: Mohr, 1916.

Henry, D. P. "Saint Anselm's Nonsense." *Mind* 72 (1963): 51–61.

Manthey, Franz. *Die Sprachphilosophie des hl. Thomas von Aquin und ihre Anwendung auf Probleme der Theologie*. Paderborn, Germany, 1937.

Prior, A. N. "The Parva Logicalia in Modern Dress." *Dominican Studies* 5 (1952): 78–87.

Roberts, Louise Nisbet. "Supposition: A Modern Application." *Journal of Philosophy* 57 (1960): 173–182.

Rotta, P. *La filosofia del linguaggio nella patristica e nella scolastica*. Turin, 1909.

Saw, Ruth Lydia. "William of Ockham on Terms, Propositions, Meaning." *PAS* 42 (1941–1942): 45–64.

Thomas, Ivo. "Saint Vincent Ferrer's *De Suppositionibus*." *Dominican Studies* 5 (1952): 88–102.

Thurot, Charles. *Notices et extraits des manuscrits latins pour servir à l'histoire des doctrines grammaticales au moyen âge*. Vol. 22, Part 2 of Notices et extraits des manuscrits de la Bibliothèque Impériale. Paris, 1868.

Wallerand, G. "Étude sur Siger de Courtrai." In *Les oeuvres de Siger de Courtrai* (*Étude critique et textes inédits*). Vol. VIII of *Les philosophes belges*. Louvain: Institut supérieur de philosophie de l'université, 1913.

RENAISSANCE AND ENLIGHTENMENT

Acton, H. B. "The Philosophy of Language in Revolutionary France." *Proceedings of the British Academy* 45 (1959): 199–219.

Cohen, L. Jonathan. "On the Project of a Universal Character." *Mind* 63 (1954): 49–63.

François, Alexis. "La Grammaire philosophique." In *Histoire de la langue française des origines à 1900*, by F. Brunot. Paris: A. Colin, 1932. Vol. VI, Part 2, Book 2.

Friedrich, Hugo. "Die Sprachtheorie der französischen Illuminaten des 18. Jahrhunderts, insbesondere Saint-

Martins." *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 13 (1935): 293–310.

Funke, Otto. *Studien zur Geschichte der Sprachphilosophie: I. Zur Sprachphilosophie des 18. Jahrhunderts: J. Harris' "Hermes"; II. Zur Sprachphilosophie der Gegenwart*. Bern: A. Francke, 1927.

Funke, Otto. *Zum Weltsprachenproblem in England im 17. Jahrhundert: G. Dalgarno's "Ars signorum" (1661) und J. Wilkins' "Essay towards a Real Character and a Philosophical Language."* Anglistische Forschungen, no. 69. Heidelberg: C. Winter, 1929.

Furlong, E. J. "Berkeley's Theory of Meaning." *Mind* 73 (1964): 437–438.

Harnois, Guy. *Les théories du langage en France de 1660 à 1821*. Paris: Société d'édition "Les belles lettres," 1929.

Hunt, H. J. "Logic and Linguistics: Diderot as 'Grammairien-philosophe.'" *Modern Language Review* 33 (1938): 215–233.

Kuehner, Paul. *Theories on the Origin and Formation of Language in the Eighteenth Century in France*. Philadelphia: n.p., 1944.

Maynial, Édouard. "Les grammairiens-philosophes du XVIII^e siècle: La Grammaire de Condillac." *Revue Bleue*, 4th series, 19 (1903): 317–320.

Oesterle, John A. "Another Approach to the Problem of Meaning." *Thomist* 7 (1944): 233–263. Supposition theory in John of St. Thomas.

O'Flaherty, James C. *Unity and Language: A Study in the Philosophy of Johann Georg Hamann*. Chapel Hill: University of North Carolina, 1952.

Pott, August. "Zur Geschichte und Kritik der sogenannten allgemeinen Grammatik." *Zeitschrift für Philosophie und Philosophische Kritik* 43 (1863): 102–141.

Robert, Louis. *Les théories logiques de Condillac*. Paris, 1869.

Russell, L. J. "Note on the Term *Σημειωτική* in Locke." *Mind* 48 (1939): 405–406.

Sahlin, G. *César Chesneau du Marsais et son rôle dans l'évolution de la grammaire générale*. Paris: Presses Universitaires de France, 1928.

Salvucci, Pasquale. *Linguaggio e mondo umano in Condillac*. Urbino, Italy, 1957.

Trendelenburg, Adolf. "Ueber Leibnizens Entwurf einer allgemeinen Characteristik." In his *Historische Beiträge zur Philosophie*, Vol. III. Berlin, 1867.

Unger, Rudolf. *Hamanns Sprachtheorie im Zusammenhang seines Denkens*. Munich, 1905.

Wallace, Karl R. *Francis Bacon on Communication and Rhetoric*. Chapel Hill: University of North Carolina Press, 1943.

THE NINETEENTH AND TWENTIETH CENTURIES

Bronstein, Daniel. "Inquiry and Meaning." In *Studies in the Philosophy of Charles Sanders Peirce*, edited by P. P. Wiener and F. H. Young, 33–52. Cambridge, MA: Harvard University Press, 1952.

Buchler, Justus. "What Is the Pragmaticist Theory of Meaning?" In *Studies in the Philosophy of Charles Sanders Peirce*, edited by P. P. Wiener and F. H. Young, 21–32. Cambridge, MA: Harvard University Press, 1952.

Burks, Arthur W., and Paul Weiss. "Peirce's Sixty-six Signs." *Journal of Philosophy* 42 (1945): 383–388.

Dewey, John. "Peirce's Theory of Linguistic Signs, Thought, and Meaning." *Journal of Philosophy* 43 (1946): 85–95.

Farber, Marvin. *The Foundation of Phenomenology: Edmund Husserl and the Quest for a Rigorous Science of Philosophy*. Cambridge, MA: Harvard University Press, 1943.

Findlay, J. N. *Meinong's Theory of Objects*. Oxford: Oxford University Press, H. Milford, 1933.

Lovejoy, Arthur O. "What Is the Pragmaticist Theory of Meaning? The First Phase." In *Studies in the Philosophy of Charles Sanders Peirce*, edited by P. P. Wiener and F. H. Young, 3–20. Cambridge, MA: Harvard University Press, 1952.

Mohanty, J. N. *Edmund Husserl's Theory of Meaning*. The Hague: M. Nijhoff, 1964.

Morris, Charles W. "Signs about Signs about Signs." *Philosophy and Phenomenological Research* 9 (1948–1949): 115–133.

Ogden, C. K. *Bentham's Theory of Fictions*. London: K. Paul, Trench, Trubner, 1932.

Ogden, C. K., and I. A. Richards. "Some Moderns—1. Husserl, 2. Russell, 3. Frege, 4. Gomperz, 5. Baldwin, 6. Peirce." In *The Meaning of Meaning*. New York: Harcourt Brace, 1923. Appendix D.

Rynin, David. "Introduction" and "Critical Essay." In his edition of Alexander Bryan Johnson's *A Treatise on Language*. Berkeley and Los Angeles, 1947.

Walker, Jeremy D. B. *A Study of Frege*. Ithaca, NY: Cornell University Press, 1965.

Weiler, Gershon. "On Fritz Mauthner's Critique of Language." *Mind* 67 (1958): 80–87

Wienpahl, P. D. "Frege's Sinn und Bedeutung." *Mind* 59 (1950): 483–494.

The development of philosophy of language, like the development of philosophy generally in the twentieth century, has taken place at least as much in articles as in books, and the use of anthologies is consequently almost indispensable in the study of it. All the following anthologies contain material useful for the study of philosophy of language in the twentieth century: *Logical Positivism*, edited by A. J. Ayer (Glencoe, IL: Free Press, 1959); *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950); *Philosophy and Ordinary Language*, edited by C. E. Caton (Urbana: University of Illinois Press, 1963); *Ordinary Language*, edited by V. C. Chappell (Englewood Cliffs, NJ: Prentice-Hall, 1964); *Logic and Language*, edited by A. G. N. Flew (1st series, New York, 1951; 2nd series, New York: Philosophical Library, 1953); *The Structure of Language*, edited by J. Fodor and J. Katz (Englewood Cliffs, NJ: Prentice-Hall, 1964); *Classics in Semantics*, edited by D. E. Hayden and E. P. Alworth (New York: Philosophical Library, 1965); and *Semantics and the Philosophy of Language*, edited by L. Linsky (Urbana: University of Illinois Press, 1952).

William P. Alston, *Philosophy of Language* (Englewood Cliffs, NJ: Prentice-Hall, 1964), provides a good introduction to the subject as it developed in the first half of the twentieth century. John Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1957), provides an excellent history of this period and its immediate background, with considerable detail regarding developments in philosophy of language. More specialized and also to be recommended are G. J. Warnock, *English Philosophy since 1900* (New York and

London: Oxford University Press, 1958), and J. O. Urmson, *Philosophical Analysis: Its Development between the Two World Wars* (London: Clarendon Press, 1956).

Norman Kretzmann (1967)

SEMANTICS, HISTORY OF [ADDENDUM]

In the 1960s, the semantics in vogue in linguistics seems to have favored some kind of decompositional approach. Consider kinship terms. Taking *P* to mean "parent of" and *F* to mean *female* you can analyze most kinship terms using first-order predicate logic. So "*x* is *y*'s aunt" would come out as $\exists z \exists w(Pwy \,\&\, Pzw \,\&\, Pzx \,\&\, w \neq x \,\&\, Fx)$. Some linguists, notably George Lakoff (1971) and James McCawley (1972), championed what was called generative semantics, where the idea was that the base level was a semantic level of structures in something such as first-order logic, which could be converted to a surface level. Noam Chomsky's (1965) level of "deep structure" was thought to lie somewhere between what subsequently became known as the level of "logical form" and the surface level. There were then debates about how autonomous syntax is, and how much of it depends on semantic input. Chomsky himself in *Aspects of the Theory of Syntax* gave an important place to what were called selectional restrictions, one of whose jobs was to rule out sentences such as "Colorless green ideas sleep furiously."

None of this work however addressed the philosophical question of what semantics really is. For that, the input came from logic. At least since Ludwig Wittgenstein's *Tractatus Logico-Philosophicus* (1963) and Rudlolf Carnap's *Meaning and Necessity* (1947), the idea had been around that the clue to semantic understanding is in the notion of the truth conditions of a sentence. How might you tell the difference between someone who knows English and someone who does not? You point to a table and utter the sentence: "There is a pen on that table." First you utter it when the pen is on the table, and then you utter it when the pen has been taken away. You do not need to know English to know the difference between a situation in which there is a pen, and a situation in which there is not, but you do need to know English to know that that difference is correlated with the truth or falsity of a sentence of English. This insight had been at the heart of the languages of logic.

In particular, during the 1960s there emerged the possible worlds semantics for modal logic, whereby an interpretation to a logic provides a set of possible worlds, and each wff of the formal language has a truth value in each of these worlds. In modal logic the necessity operator written *L* or $\Box$ is so interpreted that $\Box\alpha$ is true in a world *w* iff α is true in every w' that is possible relative to *w*, where relative possibility is specified by a relation of accessibility between worlds. The truth at a world of every complex wff is determined by the semantic evaluation rules associated with each way of getting a more complex wff from its simpler parts—rules such as the one for $\Box$ just mentioned and the rule that $\sim\alpha$ is true in a world iff α is false in that world, and so on. A. N. Prior was at the same time developing interpretations that used moments of time as the semantical indices and interpreted accessibility as the temporal ordering. Such logics were developed for epistemic and deontic operators, though there were those who, mostly under the influence of W. V. Quine (1953), declared that none of this was genuine logic.

But at that time there was still a rift between the logicians, who argued that you could only discuss philosophical issues precisely and profitably in a language that was free from the vagueness and imprecision of natural language, and the ordinary language philosophers who argued that philosophical problems come up in ordinary language and that, as P. F. Strawson had famously claimed, "ordinary language has no exact logic." The key figure at this point was Richard Montague. During the 1960s Montague began to realize that the languages of intensional logic (logics that involved truth at indices such as worlds, times, and the like) could be used in the semantical study of natural language. Noting that, under Chomsky's influence, linguists were beginning to realize that the grammatical structures of English sentences could be produced by a set of formal rules, Montague was able to produce fragments of English in which the syntactical rules, such as subject-predicate combination, could be assigned semantic interpretations, which specified how the truth-conditional meaning of a complex structure derived by these syntactical rules could be obtained from the meanings of the simpler parts, from which the complex structure is constructed. Montague's work (collected in his *Formal Philosophy*) was taken up by Barbara Partee and her students first at UCLA and then at the University of Massachusetts, and has become the dominant tradition in semantics. Montague's work established a revolution in philosophy as well as linguistics because it became no longer tenable to maintain the distinction between formal logic and natural language.

An alternative tradition, which also developed during the late 1960s and the 1970s, was inspired by Donald

Davidson's *Truth and Meaning* (1967a). Davidson's insight was to apply Alfred Tarski's (1956) work on the semantics of the ordinary predicate calculus to natural language. Tarski had set the goal of semantics as the generation of what came to be called T-schemata. The best known example of a T-schema is "'Snow is white' is true iff snow is white," where the left-hand side says that a certain sentence is true, and the right-hand side states the conditions under which it is true by using the sentence itself. While this has an air of triviality, it is not trivial. The T-schema does not say that "Snow is white" is true iff it is true. It says that it is true iff snow is white, and in a language in which the words are used differently, "Snow is white" would not be true iff snow is white. Suppose that in English *snow* means what *grass* means in English, and vice versa, and that *white* means what *green* means in English. Then the corresponding T-schema in the new language would now be "'Snow is white' is true iff snow is white," but it would now say that "'Snow is white' is true iff grass is green."

Davidson's theory does not make use of worlds or times, and he uses various tricks to get their effect. His theory of indirect discourse (*On Saying That*) treats the *that* in "Galileo said that the earth moves" as being a demonstrative pronoun whose referent is a prior utterance of "The Earth moves," and where the utterer claims that that utterance make him and Galileo samesayers. Another area where the Davidson approach has been applied is that of adverbial modification. Sentential adverbs such as *possibly* or *unfortunately* are the kind that are studied by semantics in terms of worlds, times, and other such indices, but words such as *competently* also need analysis because even if those who drive and those who sing are the same people, those who drive competently need not be the same people as those who sing competently. Davidson's account of adverbial modification (*The Logical Form of Action Sentences*) treats events as individuals, and would analyze, say, "Alice sings competently" as $\exists x$(x is the singing of a song by Alice and x is competent). Davidson's use of events has been extended by Terry Parsons (1990).

The development of formal semantics raised questions about the connection between the semantical theories provided by logic and the practice of linguistic communication. One way of connecting language with communicative practice is via propositional attitudes, of which the principal ones are knowledge and belief. David Lewis (in *Languages and Language* [1975]) defines a formal model as the correct one for a population if they only produce sentences that are true according to the model when they have certain beliefs, and trust others to do the same. And Robert Stalnaker (1984) engages with Hartry Field (1978) on the question of whether meaning depends on beliefs, or whether beliefs can only be defined in terms of an internal language.

The connection between meaning and belief forms the focus of one of the most widely discussed issues in the philosophy of language, mainly because of the views of two influential philosophers, Saul Kripke and Hilary Putnam. Kripke (1972) talks about the fact that we can use the word *Aristotle* even though we may know almost nothing about him. Kripke also imagines a person named *Pierre* (A puzzle about belief) who has heard about a beautiful city called Londres, which he believes is pretty. But the part of London that he comes to live in is so dismal that he believes that the city he knows as London is not at all pretty. So, Kripke asks us, does Pierre believe that London is pretty? Putnam (1975) looks at our use of the word *water*. If this word means H_2O, then it would seem that those without a background in chemistry do not know the meaning of what they say.

Recent developments are too many and various to detail adequately here. One, for instance, has involved the claim that semantics should be dynamic. By this, it is meant that the meaning of a sentence should be thought of not so much in terms of its truth conditions as in terms of the potential to change truth conditions. Advocates of dynamic semantics tackle this by thinking of the utterance of a sentence as like a computer program. Instead of a sentence being true or false at an index, think of the first index as being how things are before the sentence is uttered and a second index as being how things are after the sentence has been uttered. One version of this thought appears in *Dynamic Predicate Logic*, found in Jeroen Groenendijk's and Martin Stokhof's 1991 paper. Other dynamic frameworks include those based on Hans Kamp's *Discourse Representation Theory* (or in the more or less equivalent writings of Irene Heim [1983]). A rather different version of a dynamic approach is found in the game-theoretical semantics developed by Jaakko Hintikka (1983).

If the meaning of a sentence is the set of indices at which it is true, we have the consequence that all sentences true at the same set of indices have the same meaning. So for instance, because all mathematical truths are true in all possible worlds, mathematical knowledge would be trivial. However, although the entities used in formal semantics have the job of delivering a set of indices as the final result, any interpreted sentence generates a semantic structure made up from the semantic val-

ues of its simple parts (Lewis 1972, Cresswell 1985). This structure can then provide the input to propositional attitude operators. The kind of indices that have been used in classical truth-conditional semantics are complete in the sense of deciding every sentence—or at least of every sentence that might be said to have a truth value at all. In the 1980s Jon Barwise and John Perry's *Situation Semantics* was based on the view that the entities used in semantics should not be complete. The meaning of a sentence such as "Sebastian laughs" would not be the set of indices at which Sebastian laughs, but would be referred to as a situation, which would be composed of Sebastian, a location, and the relation of laughing that holds between Sebastian and the location at which he laughs.

See also Semantics.

Bibliography

Bach, E. *Informal Lectures on Formal Semantics*. Albany: State University of New York Press, 1989.

Barwise, J., and J. Perry. *Situations and Attitudes*. Cambridge MA: MIT Press, 1983.

Benthem, J. F. A. K. van, and A. ter Meulen, eds. *Handbook of Logic and Language*. Amsterdam: Elsevier/MIT Press, 1997.

Cann, R. *Formal Semantics*. Cambridge, U.K.: Cambridge University Press, 1993.

Carnap, R. *Meaning and Necessity*. Chicago: University of Chicago Press, 1947.

Chierchia, G., and S. McConnell-Ginet. *Meaning and Grammar: An Introduction to Semantics*. Cambridge, MA: MIT Press, 1990.

Chomsky, N. *Aspects of The Theory of Syntax*. Cambridge, MA: MIT Press, 1965.

Cresswell, M. J. *Logics and Languages*. London: Methuen, 1973.

Cresswell, M. J. "Semantic Competence." In *Meaning and Translation*, edited by F. Guenthner and M. Guenthner-Reutter. London: Duckworth, 1978.

Cresswell, M. J. *Structured Meanings*. Cambridge, MA: MIT Press, 1985.

Davidson, D. "The Logical Form of Action Sentences." In *The Logic of Decision and Action*, edited by N. Rescher. Pittsburgh, PA: University of Pittsburgh Press, 1967b.

Davidson, D. "Truth and Meaning." *Synthese* 17 (1967a): 304–323.

Davidson, D. "On Saying That." In *Words and Objections: Essays on the Work of W. V. Quine*. Dordrecht: Reidel, 1969.

Davidson D., and G. Harman, eds. *Semantics of Natural Language*. Dordrecht: Reidel, 1972.

Dowty, D. R., R. Wall, and P. S. Peters. *Introduction to Montague Semantics*. Dordrecht: Reidel, 1981.

Field, H. "Mental Representation." *Erkenntnis* 13 (1978): 9–61.

Gamut, L. T. F. *Logic, Language, and Meaning*. Chicago: University of Chicago Press, 1991.

Groenendijk, J., and M. Stokhof. "Dynamic Predicate Logic." *Linguistics and Philosophy* 14 (1991): 39–100.

Heim, I. "File Change Semantics and the Familiarity Theory of Definiteness." In *Meaning, Use, and Interpretation of Language*, edited by R. Bäuerle. Berlin: de Gruyter, 1983.

Heim, I., and A. Kratzer. *Semantics in Generative Grammar*. Malden, MA: Blackwell, 1998.

Hintikka, K. J. J. *Anaphora and Definite Descriptions: Two Applications of Game-Theoretical Semantics*. Dordrecht: Reidel, 1985.

Hintikka, K. J. J., and J. Kulas. *The Game of Language*. Dordrecht: Reidel, 1983.

Kamp, J. A. W. "A Theory of Truth and Semantic Representation." In *Formal Methods in The Study of Grammar*, edited by J. A. G. Groenendijk. Amsterdam: Mathematische Centrum, 1983.

Keenan, E. L., ed. *Formal Semantics of Natural Language*. Cambridge, U.K.: Cambridge University Press, 1975.

Kripke, S. A. "Naming and Necessity." In *Semantics of Natural Language*, edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.

Kripke, S. A. "A Puzzle about Belief." In *Meaning and Use*, edited by A. Margalit Reidel: Dordrecht, 1979.

Lakoff, G. "On Generative Semantics." In *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, edited by D. D. Steinberg and L. A. Jakobovits. Cambridge, U.K.: Cambridge University Press, 1971.

Lappin, S., ed. *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell, 1996.

Lewis, D. K. *Counterfactuals*. Oxford: Basil Blackwell, 1973.

Lewis, D. K. "General Semantics." In *Semantics of Natural Language*, edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.

Lewis, D. K. "Languages and Language." In *Language Mind and Knowledge*, edited by K. Gunderson. Minneapolis: University of Minnesota Press, 1975.

McCawley, J. D. "A Program for Logic." In *Semantics of Natural Language*, edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.

Montague, R. M. *Formal Philosophy*. New Haven, CT: Yale University Press, 1974.

Parsons, T. *Events in The Semantics of English*. Cambridge, MA: MIT Press, 1990.

Prior, A. N. *Past, Present, and Future*. Oxford: Clarendon Press, 1967.

Prior, A. N. *Time and Modality*. Oxford: Clarendon Press, 1957.

Putnam, H. "The Meaning of 'Meaning.'" In *Language, Mind, and Knowledge*, edited by K. Gunderson. Minneapolis: University of Minnesota Press, 1975.

Quine, W. V. O. *A Logical Point of View; 9 Logico-Philosophical Essays*. Cambridge, MA: Harvard University Press, 1953.

Richard, M. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge, U.K.: Cambridge University Press, 1990.

Stalnaker, R. C. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Stalnaker, R. C. "A Theory of Conditionals." In *Studies in Logical Theory*, edited by N. Rescher. Oxford: Basil Blackwell, 1968.

Steinberg, D.D. and L.A. Jakobovits. *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. Cambridge, U.K.: Cambridge University Press, 1971.

Tarski, A. "The Concept of Truth in Formalized Languages." In *Logic, Semantics, and Metamathematics*. Oxford: Clarendon Press, 1956.

Wittgenstein, L. *Philosophical Investigations*. Oxford: Blackwell, 1953.

Wittgenstein, L. *Tractatus Logico-Philosophicus*. 2nd ed. Translated by D. F. Pears and B. F. McGinness. London: Routledge and Kegan Paul, 1963.

M. J. Cresswell (2005)

SEN, AMARTYA K.

(1933–)

Amartya K. Sen, an economist and philosopher, was born in Bengal in 1933. The memory of the Bengal famine of 1943, in which more than 2 million people died, drew him to work on economics and ethics. He studied economics at Presidency College, Calcutta, and Trinity College, Cambridge, where he received a doctorate in economics in 1959. After he taught at the Delhi School of Economics, the London School of Economics, he held the posts of Drummond Professor of Political Economy at Oxford University (also Fellow of All Souls College), and then Lamont University Professor and professor of economics and philosophy at Harvard University. His contributions to economics lie in the areas of social choice theory, theory of choice, development economics, labor economics, cost-benefit analysis, and the measurement of inequality and poverty. In 1998 he was awarded the Nobel Memorial Prize in Economics "for his contributions to welfare economics" and appointed Master of Trinity College, Cambridge.

INFORMATIONAL PERSIMONY OF SOCIAL WELFARE JUDGMENT

The large number of Sen's works in economics and philosophy are marked by tireless criticism of utilitarianism and the utilitarian foundations of welfare economics. According to Sen utilitarianism can be factored into three elements: act consequentialism (the goodness of an act is given by the goodness of its consequent states of affairs), welfarism (the goodness of a state of affairs is given by the goodness of utility information regarding that state), and sum-ranking (the goodness of utility information is given by the sum total of different people's utilities). These elements impose informational constraints on policy judgments and economic evaluation: Act consequentialism does not consider the intrinsic value of an act or the motivation underlying the act; welfarism rules out nonutility information such as violation of rights from influencing the goodness of an act; and sum-ranking excludes information about the state of people who are worse off. Sen holds that the informational basis for judgments of goodness should include nonutility information and information about the distribution of utility among different people.

Much of his philosophical standpoint originates in the close examination of Kenneth J. Arrow's (1921–) general impossibility theorem: there exists no collective decision-making rule that satisfies some seemingly uncontroversial axioms (unrestricted domain, weak Pareto principle, independence of irrelevant alternatives, and nondictatorship). In *Collective Choice and Social Welfare* (1970a) he scrutinizes the formal and philosophical reach of this theorem and points out the informational parsimony of Arrow's framework. Since Abraham Bergson (1914–2003) and Paul A. Samuelson (1915–) established the "new" welfare economics, individual preference orderings have been assumed to be ordinal and interpersonally incomparable, because there is supposed to be no scientific ground to compare one person's preference satisfaction with another's. Sen shows that if the informational basis is extended to include some kind of interpersonal comparability (e.g., the unit comparability or level comparability), there exist collective decision-making rules, including some egalitarian rules such as *maximin* and its lexicographic extension, *leximin* (as endorsed in John Rawls's difference principle).

INDIVIDUAL FREEDOM AND THE NOTION OF WELL-BEING

In "The Impossibility of a Paretian Liberal" (1970b) he further shows that two conditions in Arrow's theorem (unrestricted domain and weak Pareto principle) are inconsistent with individuals' minimal liberty. Although the weak Pareto principle (if everyone in the society strictly prefers x to y, x is socially preferred to y) is taken to be uncontroversial in economics, it is sufficient to spread the decisiveness of a certain group over all the pair of alternatives, even if the preference over the pair is a purely personal matter that the society should respect. While the same concern led Robert Nozick to his libertarian side-constraint theory, in "Rights and Agency" 1982), Sen adopts a broadly defined consequentialist theory called a goal-right system, according to which individual freedom should be promoted as an end by the society. This moves him to give individual freedom a central role in the evaluation of states of affairs.

In *Commodities and Capabilities* (1985a) he argues against the "opulence" view of well-being (e.g., real income and Rawls's primary goods) and the "utility" view

(e.g., happiness, desire-fulfillment, and the revealed preference theory in welfare economics), and proposes an alternative notion of well-being: the capability to function. He takes human life to consist in a combination of various doings and beings, which he calls functionings (e.g., moving, being well nourished, being in good health, and being socially respected). The capability to function refers to different combinations of functionings, and the capability of a person corresponds to freedom to choose one kind of life among others.

One advantage of this approach is that it takes account of the people's varying capacities to convert primary goods into abilities to pursue their ends. Each person's capability to function is influenced by internal factors such as disability, illness, age, and gender, as well as external factors such as climatic circumstances, educational arrangement, the prevalence of crime and violence, and the resource distribution within the family. What a disabled person can achieve from a larger set of goods may be much less than what an able-bodied person can achieve from a smaller set of goods. The capability approach offers the analytic ground to capture people's diverse needs. This approach changed not only the concept of well-being in ethics but also the paradigm of international development. It became the source of the Human Development Indicators of the United Nations Development Programme.

Through his empirical studies on famines, in *Poverty and Famines: An Essay on Entitlement and Deprivation* (1981), Sen maintains that there has never been a famine in a functioning multiparty democracy: The democratic poor countries such as India, Botswana, or Zimbabwe managed to avert famines despite serious crop failure, whereas the dictatorial countries had major famines. This is because, Sen claims, democracy would spread the penalty of famine to the ruling parties and political leaders, thus providing the political incentives to try to prevent any threatening famine. In subsequent works he champions the notion of human rights for their intrinsic importance, their consequential role in providing the political incentives for economic security, and their constructive role in the genesis of values and priorities.

See also Consequentialism; Philosophy of Economics; Rights; Utilitarianism.

Bibliography

WORKS BY SEN

Collective Choice and Social Welfare. San Francisco: Holden-Day, 1970a.

"The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78 (1970b): 152–157.

On Economic Inequality. Oxford, U.K.: Clarendon Press, 1973.

"Equality of What?" In *Tanner Lectures on Human Values*. Vol. 1., edited by S. McMurrin. Salt Lake City: University of Utah Press, 1980.

Poverty and Famines: An Essay on Entitlement and Deprivation. Oxford, U.K.: Clarendon Press, 1981.

"Rights and Agency." *Philosophy and Public Affairs* 11 (1) (1982): 3–39.

Commodities and Capabilities. New York: North-Holland, 1985a.

"Well-being, Agency, and Freedom: The Dewey Lectures 1984." *Journal of Philosophy* 82 (4) (1985b): 169–221.

Development as Freedom. New York: Knopf, 1999.

Iwao Hirose (2005)

SENECA, LUCIUS ANNAEUS
(4 BCE–65 CE)

A Roman adherent of Stoicism with a particular interest in ethics, Seneca had an extensive career in politics and literature. His *Moral Epistles*, two major treatises, and a series of essays including *On Anger* offer an engaging presentation of philosophical ideas and are an important source for earlier Stoic thought. Also extant are eight plays and a political lampoon.

Seneca was of provincial origin, having been born at Córdoba in southern Spain, but was brought to Rome at an early age. There he received an extensive education in public speaking and literary composition. His knowledge of philosophy came from the lecturers Papirius Fabianus and Sotion (both adherents of Sextian moral philosophy), from the Cynic Demetrius, and from the Stoic Attalus. He won considerable repute as an advocate, but his health was poor and he was in disfavor with the emperors Gaius and Claudius. Exiled by Claudius to Corsica, he was recalled in 49 to become tutor in rhetoric to the boy Nero.

Following Nero's accession he held a position of considerable influence, restraining the young ruler's excesses and composing important speeches for him; as Miriam Griffin has shown, however, his influence on administrative policy was much less than accounts of his career by Tacitus and Dio Cassius would lead one to believe. Late in life he withdrew from politics, transferred his large fortune to the imperial treasury, and devoted himself to philosophical study and writing. His suicide, after being implicated in the Pisonian conspiracy of 65, followed the model of enforced self-execution typical under the Roman emperors.

The major prose works treat a range of topics in ethics, including the theory of value and the human good, character and moral development, moral psychology, self-care and the management of emotion, friendship and political engagement, and practical morality. Occasional forays into logic and metaphysics (primarily in the *Epistles*) and physics (in the *Natural Questions*) are treated as subordinate to ethics. The intention of the mature works to represent the positions of the Stoic founders Zeno of Citium and Chrysippus is not seriously in doubt; there is little to support the claims of an older generation of scholarship for a deliberate program of eclecticism.

While Seneca claims intellectual independence, such claims are themselves in the Stoic spirit of intellectual inquiry. Doctrinal allegiance does not prevent him from studying the writings of Platonists and Epicureans, and he will occasionally endorse a point of doctrine. Epicurus himself he admires for his disciplined personal habits and effective use of various instructional methods. But it would not be accurate to describe him as amalgamating Stoic with Epicurean philosophy, for his criticism of Epicurus's hedonist foundations is as sharp as Cicero's. He is eager to claim common ground among philosophers on such points as do not require him to reverse his Stoic commitments; where definite doctrinal commitments are in evidence, however, his sympathies are decidedly Stoic.

Seneca's major work, the *Moral Epistles to Lucilius* (written between 62 and 65 CE), makes creative use of the epistolary format to present a sustained course of philosophical instruction. Intended for a wide circle of readers rather than for the sole addressee, the collection mingles scenes of daily life with a variety of topics in ethics, psychology, and occasionally metaphysics. The arrangement of ideas is deliberately unsystematic, with some topics treated in cursory fashion and others developed to considerable length. Discussions of particular interest include the non-utilitarian basis of friendship, Platonic ontology, the responsibilities of philosophers to the state, the status of moral rules, and the initial orientation of the human. A more sustained theme is the moral development of the individual as illustrated by that of the author himself and his addressee.

The earlier of Seneca's two full-length treatises, *On Benefits* (written between 56 and 64), is a study of social transactions based on similar works by the Stoics Chrysippus and Hecato. The giving and receiving of various benefactions is analyzed with rigorous attention to the motivation of the giver; there is considerable casuistic elaboration. Very different in character is the *Natural Questions* (written between 62 and 65), which offers rational explanations for a list of phenomena regularly grouped in antiquity under the heading of meteorology—weather events, comets, earthquakes, and other events whose causes were not directly observable. Like Epicurus, Seneca treats such phenomena as admitting of multiple explanations; this enables him to incorporate a wide range of competing theories into his work. On the whole, however, he maintains the Stoic position on cosmic design, which he sees as having ethical significance: by pondering the regularity of the heavens and the causes of natural events, one can rise above one's ordinary objects of concern and adjust one's thought to the standard of universal reason.

Preeminent among the essays is *On Anger* (complete by 52), in the tradition of Hellenistic anger-management treatises. A careful treatment of the psychology of anger adheres to the Stoic theory of emotions generally: anger is dependent on the rational being's capacity for assent; it is intractable once begun, but can be forestalled by the techniques of cognitive therapy. Among the shorter essays, three are consolatory works of a conventional nature; the remainder treat single topics: the superiority of the virtuous person to suffering (*On Providence*) and to injury and insult (*On Constancy*); the moral end and the supposed hypocrisy of philosophers (*On the Happy Life*); remedies for spiritual malaise (*On Tranquility*); the productive use of time (*On the Brevity of Life*); and the justification for scholarly retreat (*On Leisure*).

The level of doctrinal commitment varies considerably in the shorter essays: Some restrict themselves to a Stoic viewpoint, whereas others, notably *On the Brevity of Life*, are in the spirit of generalized philosophical protreptic. Although the title *Dialogi* is given to the essays in the major manuscript, none is a dialogue in the sense that Plato's works are dialogues. A second speaker is sometimes made to voice an objection, but neither that voice nor the named addressee is developed into a genuine interlocutor. The essay *On Tranquility* does, however, represent its addressee as offering a confessional description of his own moral struggles. This unusual device presages the more extensive experimentation with literary form in the *Moral Epistles*.

The essay *On Clemency* has the greatest political significance of any of Seneca's works. Circulated early in Nero's reign, it celebrates what had become a watchword of the new regime and offers the essentials of a theory of good government based on the character of the ruler. Clemency, or the justifiable mitigation of justifiable penalties, is distinguished both from leniency, which is

unjustifiable mitigation, and pity, which is emotional distress at another's misfortune. Seneca attempts, characteristically, to reform Nero's administration from within, preferring failure to immediate martyrdom; in this, he differs from the hard-line idealism of Thrasea Paetus and other philosophically-minded contemporaries.

Seneca's early rhetorical training manifests itself in his style of writing, which is enhanced with clever turns of phrase, metaphor, and wit. The narrative skills displayed in his eight tragedies and in his political farce, *The Pumpkinification of Claudius*, are put to use in the philosophical prose to give effectiveness to historical anecdotes and, especially in the *Moral Epistles*, to case studies from Seneca's own acquaintance. His interest in combining philosophical with literary achievement is evidenced in his decision to write in Latin rather than Greek, which was the usual language of philosophical discourse among educated Romans of the period. It has been argued by Nussbaum and others that his tragedies are themselves experiments in ethical suasion by literary means; there is some question whether, in that case, the values promoted could remain consistent with Stoic ethics, but certainly it is true that such elements of Stoic cosmology as the cyclical conflagration do appear with some regularity in those works.

See also Chrysippus; Epicurus; Eudaimonia; Hellenistic Thought; Kalon; Logos; Platonism and the Platonic Tradition; Stoicism; Zeno of Citium.

Bibliography

Barnes, Jonathan. *Logic and the Imperial Stoa*. Leiden; New York: Brill,1997.

Griffin, Miriam. *Seneca: A Philosopher in Politics*. 2nd ed. Oxford: Clarendon Press, 1991.

Hine, Harry, ed. *Naturales quaestiones*. Stuttgart: Teubner, 1996.

Inwood, Brad. "Natural Law in Seneca." *Studia Philonica* 15 (2003): 81–99.

Inwood, Brad. "Seneca and Psychological Dualism." In *Passions and Perceptions*, edited by J. Brunschwig and M. Nussbaum. Cambridge, MA: Cambridge University Press, 1993.

Inwood, Brad. "The Will in Seneca the Younger." *Classical Philology* 95 (2000): 44–60.

Nussbaum, M.C. *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press, 1994.

Reynolds, L.D., ed. *Dialogi*. New York: Oxford University Press, 1977.

Reynolds, L.D., ed. *Epistulae morales*. Oxford: Clarendon Press, 1965.

Rosenmeyer, T.G. *Senecan Drama and Stoic Cosmology*. Berkeley, CA: University of California Press, 1989.

Margaret Graver (2005)

SENSA

A distinction is often drawn in philosophy between two types of objects of awareness in perception. First, there are physical objects or substances (such as chairs, books, rocks, and water) and living organisms (animals, plants, and human beings insofar as they are perceptible, that is, their bodies). A common technical term for all these is *material objects*. Second, there are data of immediate awareness, which we shall refer to as "sensa" (singular, sensum), such as color patches or shapes, sounds, smells, and tactile feelings. This distinction is usually fourfold: (*a*) in status—material objects are external, located in physical space, and "public" (observable by different persons at once), while sensa are private and are usually held to have no external physical existence; (*b*) in extent—material objects may at one time correspond to several sensa and normally persist throughout the occurrence of many sensa; (*c*) in directness—the perception of material objects is indirect, that is, it involves inference from or interpretation of sensa that are "given" directly to consciousness; (*d*) in certainty—one is always certainly aware of sensa but not necessarily so of material objects.

There is no universally accepted term for sensa; *sensations* and *sense data* are commonest but indicate a further subdivision. *Sensation* is customarily used by scientists and psychologists and carries with it the suggestion that sensa are the immediate mental effects of brain activity resulting from the excitation of a sense organ by external stimuli. It and the less specialized term *sense impression* may be used interchangeably for the whole experience of awareness of sound, color, and the like, or for any sensum (such as a sound or a color patch) distinguished within it. The term *sense datum* (plural, *sense data*) apparently originated with G. E. Moore but was introduced in print by Bertrand Russell in 1912. It later became particularly associated with the sense-datum theory of Moore, C. D. Broad, and H. H. Price, while Russell developed different views and came to use other terms.

Sense data are not meant to carry any implications of causal theory, and awareness of them is called sensing (the term *sense datum* is used for the sensum only, not for the whole experience). With the development of the sense-datum theory, controversy arose between those who regarded sense data as objects distinguishable from the act of awareness of them (act/object analysis) and those who denied this and claimed that sensing is really of "sense contents" (adverbial analysis). But the terminology is generally fluid—for instance, some modern neurolo-

gists use the term *sense data* instead of *sensations* in causal contexts. Similar concepts are found in earlier writers, though their language is different. John Locke's "ideas of sense," George Berkeley's "ideas" or "sensible qualities," and David Hume's "impressions" are all forms of sensa.

SENSATIONS

It has often been maintained, by philosophers as well as by psychologists, that perceiving consists in the synthesis and interpretation of sensations. But it must be realized that the occurrence of sensations in all perception is only a hypothesis and not an obvious feature of experience. In ordinary language, one may speak of having or feeling sensations of thirst, cold, or pressure and may refer to itches or pains as sensations. But the technical use of the word *sensations* involves a considerable extension of meaning, since one then speaks of visual or auditory sensations (that is, colors or sounds), while such locutions have no place in ordinary speech. We do not have green sensations in our eyes, nor do we normally feel or have sounds in our ears. Admittedly we do have afterimages, spots before the eyes, or ringing in the ears; but these are special cases because, unlike the objects or data of normal perception, the images, spots, or ringing "follows us around" and cannot be avoided by moving the head, closing the eyes, or stopping the ears. Indeed, in normal perception we are conscious not of colored shapes or of sounds as such but of material objects, or at least of ostensible material objects. Admittedly we may sometimes be aware of sounds, smells, tastes, or feelings of pressure, as distinct from objects or object properties, but it is doubtful how far these can be said to be sensations.

Sounds and smells seem public and external: Two or more people may hear the same sound or smell the same smell and agree on its source; sounds travel, and a smell may fill a room. Tastes are a borderline case—private and in the mouth, yet in a sense external to the skin and membranes—while feelings of pressure or warmth are partly sensations proper and partly seem to be awarenesses of heavy or warm objects. However, colors and colored shapes normally seem quite external, public, and at a distance from us.

Sensations in this technical sense (private mental objects of immediate awareness) are thus mainly hypothetical occurrences. Their postulation can be justified only by its success in explaining the facts of perception, and it rests on two grounds. First, there is the causal argument—perception of objects depends on and is conditioned by a chain of causal processes; for example, light waves or sound waves stimulate the appropriate sense organ, causing impulses to travel along nerves to the brain and activate the appropriate receiving area. Perception cannot, therefore, be direct contact or confrontation with external objects—all immediate awareness must result from the causal process and be an awareness of mental sensations due to brain activity. Since they are thus separated from the external object in time and space, sensations cannot be identified with its properties, though they may resemble them.

Second, there is the psychological argument—many characteristics of perception show that it is not a direct intuitive awareness but involves interpretation of sensations. Thus, error and illusion are really misinterpretations; perception of motion, depth, and distance involves the use of sensory "cues"; and perceptual identification and discrimination are interpretative, not immediate, since they can be improved by learning and experience. (Both these arguments are discussed at greater length under the Perception entry. Here we may simply note some relevant difficulties.)

THE EXTENT OF SENSATIONS. Even if the causal argument forces us to distinguish between external material objects and the immediate objects of awareness caused by brain activity, it does not follow that the latter must be sensations, such as colors or sounds. They may be percepts, that is, mental contents that correspond to whole material objects, though here the psychological argument comes in, suggesting that percepts are the products of interpretation. Supporters of the theory of sensations, no doubt influenced by discoveries concerning the atomic structure of matter, at one time even claimed that the basic sensations are "atomic," that they are sensory point-elements, each corresponding to a different nerve cell—a patch of red color would thus be made up of many sensations of red. This view has now been completely abandoned, largely as the result of the experiments of the Gestalt psychologists, which show that our primary awareness is of organized wholes or figures (*Gestalten* in German), and not of elements into which these wholes might theoretically be analyzed. But even though sensations are not now thought of as minute elements that we synthesize, nonatomic sensations (colored patches of a larger size, or patterns of them, as well as sounds, smells, and so on) may still be regarded as data that we interpret in perception.

AWARENESS AND INTERPRETATION OF SENSATIONS. The awareness of sensations or, for that matter, of percepts must itself be explained; the danger is that it will be construed as analogous to perceiving; for example,

that seeing objects will be explained as seeing sensations caused by them, which is a circular explanation and can thus lead to an infinite regress—seeing sensations must require seeing further sensations, and so on. (Compare the duplication objection to representative realism in the Realism entry). It is therefore necessary to maintain that the awareness of sensations or percepts ("having sensations") is a special kind of direct awareness different from perceiving, an amendment explicitly adopted by the sense-datum theory.

The problems of the psychological argument are (*a*) that interpretation of anything would commonly be regarded as presupposing consciousness of what is interpreted, and we are normally conscious neither of having sensations (as opposed to perceiving objects) nor of interpreting them; and (*b*) that the nature of the interpretation of sensations is controversial—a range of theories is possible because it is not introspectable. The sensationalists (James Mill, J. S. Mill, and others who derived their inspiration from Hume) claimed that perceiving is the association of various sensations. *Association* is a vague term and was explained as the customary linking of ideas or sensations that are similar, contiguous in space and time, and so on. F. H. Bradley and other idealists successfully attacked the sensationalist view as inadequate to explain the facts of perception; instead, they claimed that the interpretation is an inference leading to a judgment, supposing that the possibility of error in perception required this. But this overintellectualized perceiving; inferences and judgments are not the only forms of mental activity liable to error.

ARGUMENTS FOR THE INTRODUCTION OF SENSE DATA

Since the start of the twentieth century, philosophers have made little use of the concept of sensation in their theories but have instead talked of sense data or sense contents. Though the same things—color patches, sounds, smells, and tastes—have been put forward as examples both of sensations and of sense data, the new terminology marks several changes. Recognition of the visual depth or stereoscopic qualities of sense data means that one visual sense datum or color patch is usually held to correspond to the whole of the visible part of an ostensible object (so that one may have striped or variegated sense data). Little detailed attention has been paid to psychological phenomena, except for discussion along traditional lines of error and illusion and their bearing on whether perceiving is a form of judgment. There has also been almost a revulsion from causal arguments, clearly influenced by their tendency to involve one in the notorious difficulties of representative realism. Instead, a fresh start has been made in the conviction that philosophy has its own distinct contribution to make in the logical and introspective analyses of perception and in the consideration of relevant epistemological issues, that is, of the extent to which perception provides knowledge of external reality. Nevertheless, with some adjustment the new arguments might be supplemented by and in turn supplement the causal and psychological arguments for sensations.

Sense data are defined as whatever is "given" or "directly present" in perceiving; they are the object of *sensing*, of "direct" or "immediate" or "actual" awareness in perception. The claim that this awareness occurs within perceiving is essential to the analysis. To most of its exponents it seems a clear fact of our experience as percipients, one revealed by reflective examination. "Direct" is explained by Price (in *Perception*) as meaning intuitive or "not reached by inference, nor by any other intellectual process." This formal definition was often supplemented by a kind of ostensive one: Moore, J. R. Smythies, and others gave instructions for looking at an object or scene and picking out the sense datum, such as a colored shape. (Misleadingly, afterimages were sometimes offered as examples of sense data, but their difference from normal perception has already been noted; misleadingly also, some talked of seeing or hearing sense data.)

This definition of sense data naturally raises the question "Why not say that tables, chairs, and other material objects are given or directly seen?" In answering this, these philosophers produce various arguments for distinguishing sense data from material objects.

THE CERTAINTY ARGUMENT. The certainty argument was stressed by Price and by Russell in his search for "hard data," though it is also found in other sources. Directness or givenness implies certainty—what is given must be limited to what we are absolutely certain of. But in any perceptual situation we cannot be sure that we are aware of any particular material object. For example, an object that seems to be a tomato may in fact be something quite different—a wax imitation, perhaps, or a reflected patch of light, or a hallucination (that is, not be a material object at all). Yet whatever the illusion may be, there can be no doubt, when we seem to see a tomato, that there is given a red, round, bulgy patch of color, a sense datum. Another version of this argument is the method of reduced claims; by confronting him with possible sources of error, you force the person concerned to reduce his

claim from "I see a tomato" to what he actually and directly sees or, rather, senses: "I see a red, round color patch."

THE PARTITIVE ARGUMENT. When we observe a tomato or a bell, what we "actually see"—the "objective constituent" of the situation, what is given or sensed—is the colored shape that seems to be its front surface. This is a sense datum. We assume that the object has other surfaces and has other characteristics, such as causal properties, three-dimensionality, and persistence in time; and if we loosely say that we see a bell, we imply that we are perceiving an object possessing these properties, although we do not directly see or sense them. This argument, which stresses extent of sense experience rather than certainty, was preferred by Broad and Moore but seems inferior in suggesting that sense data are those parts of an object that we "actually see" on a given occasion—which raises difficulties with respect to illusions.

THE ARGUMENT FROM THE CONTENT OF ILLUSIONS. When a drunkard sees a hallucinatory pink elephant or sees two bottles when only one is present, what is the elephant or second bottle if it is nothing material? The sense-datum theorist answers, "A private object of awareness, a sense datum," and applies this also to cases of the relativity of perceiving: For example, when a round plate looks elliptical to a person standing at one side, the elliptical appearance cannot be the plate, which is round; it is an elliptical sense datum private to that person. Indeed, it is argued that at all times we are directly aware only of sense data, since there is no qualitative jump between the cases where one cannot be directly aware of an object, and so must be sensing sense data, and the normal cases where we think we are directly aware of an object. This gradation or lack of jump is particularly clear in the case of relativity, as when we gradually move from where the plate looks round to where it looks elliptical, but it also applies to many hallucinations where the illusory sensa are integrated with a genuine background. In short, perceiving a material object involves sensing sense data related or "belonging" to it; when the plate looks round to me and elliptical to you, I am sensing a round sense datum belonging to it and you are sensing an elliptical one.

THE FULL SENSE-DATUM THEORY

The fundamental conception of sense data, as directly given elements of experience, spread far beyond epistemology. Both the atomic facts of the logical atomists and the supposedly incorrigible basic or protocol propositions of the logical positivists had as their prime examples simple statements about sense data (or sensa generally), such as "This is red." But the conception was also developed into a full theory of perception by consideration of the following topics, even though disagreements led to variant accounts.

THE GENERAL NATURE OF SENSE DATA. The arguments for the introduction of sense data, if valid, show that sense data are given and provide examples of them. Further alleged properties emerge from the discussion of illusions and relativity, namely, that sense data (1) are private, each sensed by only one percipient (see argument from the content of illusions); (2) are transitory existents, lasting only while they are sensed, so that they are usually claimed to be events rather than things or properties; (3) are distinct from the percipient and seem to be external (in contrast with sensations); (4) are without causal properties, for sounds (as opposed to sound waves) cannot act on other things, nor can colors or tastes, though the sensing of them may affect a person; (5) cannot be other than they appear to be, or the certainty argument is undermined.

Despite wide agreement on most of these points, a considerable divergence of view arose about (3) and (5). Point (3)—that sense data are distinct from the percipient and seem to be external—involves what came to be called the act-object analysis of sensing. Largely on phenomenological grounds—on how direct experience of color patches, sounds, and such seem to the person concerned—Price and others claim that sense data have distinct existence, that they are objects distinguishable from the act of awareness of them. But some philosophers maintain that the data are only "sense contents" and do not exist apart from the sensing of them any more than does a pain or sensation. This view is formulated in the so-called adverbial analysis of sensing, namely, that "I sense a red color patch" is properly to be regarded as a statement of how I sense or, to put it in a different way, "red color patch" is an internal accusative of the verb *sense*, just as "waltz" is an internal accusative of *dance* in "I danced a waltz."

There is agreement on point (5)—that sense data cannot appear to be what they are not, for example, sense data cannot appear elliptical when they are round. (Even this is dubious—an apparently pink expanse may, on examination, be found to consist of red dots on a white background.) But some say that sense data can fail to appear as they are (do not reveal their full properties at first sight); thus, one may see that a colored datum is

striped without noting how many or how thick the stripes are. Others deny this, claiming that a closer look results in a fresh sense datum. In fact, the theory cannot deal satisfactorily with the phenomenon of attention. A thing may look quite different on careful examination from the way it looks at a casual glance, and the difference seems to be a matter of how attentively we look, a matter of changes in our mode of observation. In line with this evidence, one should say that sense data may reveal their full properties only on a closer examination, but then one is suggesting that sensing may at times be casual and inattentive and is thus undermining the fundamental claim that sensing is certain and incorrigible.

THE RELATION OF SENSING TO PERCEIVING. The distinction between sensing and perceiving is threefold. First, perceiving is the awareness of some material object; except in certain kinds of illusion this awareness is the result of the object in question (or light or sound from it) acting on the percipient's sense organs. Sensing is the awareness of private sense data that differ from material objects and do not affect the sense organs. Second, sensing is claimed to be direct, immediate, and incorrigible, a form of knowing. Owing to illusions, perceiving cannot be this; it is fallible and indirect. Third, the indirectness of perceiving is said to consist in its being mediated by sensing; perceiving involves sensing, contains sensing within it.

Various views are possible about the nature of this mediation of perceiving by sensing, but they are best expressed as theories of perceptual consciousness. The same kind of consciousness of a tomato, for example, seems present in normal perception, when one sees a tomato as a tomato; in an illusion, when what one sees as a tomato is a piece of wax; and in a hallucination, when no corresponding material thing is there. The kind of consciousness present in these three cases may be called perceptual consciousness and is more conveniently discussed than perceiving, where the implication that there is an object acting on the sense organs complicates the issue.

Some, such as Brand Blanshard (*The Nature of Thought*, London, 1939, Ch. 2), claim that perceptual consciousness consists in sensing a datum and judging or inferring that it belongs to a material object. Price, however, argued that this is too intellectual and does not fit the facts. We unquestioningly accept or take for granted rather than infer or judge, and therefore he defined perceptual consciousness as sensing a sense datum (or data) and taking for granted that it (or they) belong to a material object. Others have said that we refer the sense datum to a material object, but *refer* is vague.

Two points of interest arise here. First, philosophers have most often said that we accept or judge that the sense datum belongs to a physical object. This seems obvious only about smells or tastes, and one would on first thought say we assume that the visual sense datum or color patch is the tomato. There is a reluctance on the part of sense-datum theorists to allow this, presumably because they are influenced by the partitive argument or by their knowledge that ex hypothesi the sense datum cannot possibly be the physical object. But there seems to be no reason why the ordinary person, whose mental processes are being described, may not mistakenly assume this; one would, for example, say "That patch of white over there on the hill is a sheep" (admittedly, the patch as "public" is hardly a sense datum, but it is the nearest one can get to a sense datum by ordinary examples).

Second, to say that we judge or infer that a sense datum belongs to (or is) a physical object is implausible, for it implies we are conscious of it first as a datum, which is not true to the facts: There is no passage of mind from datum to object, as in inference. Even to say we subconsciously judge or infer is unsatisfactory, for it seems extravagant to suppose that we constantly do subconsciously what we never do consciously. Price attempts to overcome this by maintaining that to take for granted that *A* is or belongs to *B*, one does not need to distinguish them at the time—indeed, the contrary is implied. Sensing thus comes to be regarded as a sort of sensory core within perceptual consciousness, surrounded, as it were, by the further activity of taking for granted. The two states of mind, sensing the red sense datum and consciousness of the tomato, arise together and simultaneously and can be distinguished only by subsequent analysis.

Even this account may be criticized on the grounds that it still does not do justice to the evidence of experience, namely, that perceptual consciousness is one unitary and unanalyzable state of mind, not two. No subsequent analysis of experience reveals sensing as an element within perceptual consciousness. Analysis or reflective examination can result in a "reduced" or critical phenomenological mode of observation in which one distinguishes sounds or colored shapes as such without attributing them to objects, but if this is sensing—and it seems to be the nearest one can get to it—then it is a quite different state of mind from normal perceiving. There is no ground for supposing that this, achievable only by an

effort of analysis, occurs as part of normal unconsidered perception. In general, therefore, the attempt to establish sensing sense data as an omnipresent basic element in perceiving faces the same difficulties that faced the claim that perceiving is the interpretation of sensations.

Another way of seeing the error is to consider the normal usage of "taking for granted." Price's analysis is at first sight closest to "*Y* saw the book and took for granted that it belonged to *B*," but then *Y* is referred to as conscious of the book, while the average percipient is not conscious of sense data as sense data; he is conscious only of the material object. This difficulty can be avoided by the formulation "*X* took for granted that *A* was *B*"; for example, that the piece of wax was a tomato, or that the visitor was the man he was expecting. In each case both *A* and *B* denote the same entity (the wax or the visitor). *A* describes this entity in a way that the speaker knows to be correct; *B* describes it as *X* saw it. Similarly, one might say, "He took for granted that the sense datum (*A*) was a material object (*B*)." But this will not really save the analysis in which the datum and the physical object are alleged to be two quite different entities; to fit the analysis the first phrase (*A*) must also be a description of the alleged object of awareness of *X*, not of the speaker. Price seems to be making the mistake of offering as a description of a percipient's actual mental content what is in fact a description of the situation that can be made only by someone correcting the percipient's error.

THE RELATION OF SENSE DATA TO PHYSICAL OBJECTS. One of the vaunted advantages of the sense-datum analysis of perception is its neutrality with respect to the traditional realist theories of knowledge. (Idealism was ruled out by the original claim that sense data are distinct from the sensing of them.) Indeed, sense data were even said to be neutral in that so far as the analysis is concerned, they can be mental or physical or neither. Consequently, it is possible to state the various theories of knowledge in terms of sense data. Naive realism reduces to the view that sense data are parts of the surface of material objects; representative realism would claim that sense data are mental existents caused or generated by cerebral activity ultimately due to material objects and that sense data resemble the properties of these objects. (The second view and, if not too naive, the first also, could admit "wild sense data"—hallucinations that are not part of or caused by physical objects.)

Moore at times toyed with supposing that sense data are parts of the surface of objects (and even seriously discussed whether they might be identical with objects), though this must have been due to his affection for the partitive argument. The other arguments for sense data and general considerations about illusion do not allow this; for example, a round dish cannot have an elliptical sense datum as part of its surface. Representative realism is a more likely possibility: Neurologists such as Smythies advocate this theory in terms of sense data, and Broad proposed something not unlike it. Most of the philosophers have, however, rejected it in view of its traditional difficulty—if our observation is limited to sense data while material objects are only assumed causes of sense data, then these objects are in fact never observed and therefore may, for all we know, not really exist.

A more common view is that sense data belong to material objects in the special sense that the latter are composed of "families" of sense data. This "family" relationship is not literally one of whole and part, as in naive realism; the material object is supposed to be a complex system or pattern of groups or sequences of sense data. But if a physical object is simply a family of sense data, then when no sense datum occurs—when the object is unobserved—the object must cease to exist. This is felt to be too paradoxical, and two main lines of development within this view have been put forward: (1) phenomenalism, in which the object is regarded as a family of actual and possible sense data—when unobserved, it consists solely of possible sense data; (2) a compromise theory put forward by Price in which the material object, while mainly such a family, contains a physical occupant that persists, even while it is unobserved, as the source of all its causal properties. The notion of a physical occupant has some analogies to Immanuel Kant's notorious thing-in-itself, and this view has not obtained widespread acceptance.

This divergence of view reflects a central dilemma in the sense-datum theory. If the theory maintains that sense data belong to material objects or that the latter in some way consist of them, then it is difficult to explain (*a*) the persistence of such objects when unobserved; (*b*) the privacy that all versions attribute to sense data—how can a public object be a family of private sense data?; (*c*) the conditioning or even generation of sense data by the sense organs and nervous system, which is required by the physiological facts, by the occurrence of hallucinations or color blindness, and by the effects of attention and learning on perception. (Most sense-datum theorists admit the generation as well as the conditioning.) But if one does not say that sense data belong to or constitute material objects, the distinctness and apparent depth of sense data (at least of visual ones) is difficult to explain;

and, more important, sense data tend to become mental entities like sensations. This, together with the privacy and the generation by the brain, leads one into representative realism.

One attempt to escape this dilemma is to say that sense data are extended and located in their own private "sensible" space along the lines first suggested by Russell in his *Mysticism and Logic* and *Our Knowledge of the External World*. There is one such sensible space, with its own extension and dimensions, for every point in physical space, and the latter in fact becomes the system of points at which sensible space occurs. A physical object is thus, as it were, spread over physical space in a series of "perspectives" or "unperceived aspects," in the special sense that from different points in physical space, granted that sense organs and brain function properly, sense data may occur in sensible space but also belong to the object as appearances of it and reproduce its characteristics in a way modified both by the viewpoint and by the nature of the sensory apparatus.

This theory is very complex, which means that any summary of it is necessarily garbled. Two of the complexities are that a special interpretation is needed of what we normally call the volume occupied by a physical object and that account must be taken of the different senses, for sight, sound, and touch at least each have their own specific spaces. (Russell later spoke of sensible space as a construct of these spaces, but a construct cannot be the space in which immediately given sense data are located.) A further difficulty is that a given sensible space cannot really be at a point. Not only are the hands, say, at some distance from the eyes, but the brain and the sensory activity associated with perception of an object at one time and place are also really spread over an area. However, the major objection is once again the causation and conditioning of sense data by sense organs and nervous system. How do they influence or produce data in sensible space, or modify the appearance in sensible space of an object in physical space? As soon as one tries to fit in the causal processes, it is difficult to avoid straightforward representative realism, in which all this elaboration becomes unnecessary; perspectives become otiose, except as mere possibilities, or turn into light waves and sound waves. Hence, Russell's later views gradually approach representative realism (for example, in *Human Knowledge*, 1948).

There does, in fact, seem to be no satisfactory way out of this dilemma for the sense-datum theory. Upholders of it must embrace one horn or the other—they must maintain pure phenomenalism or representative realism. Each has its well-known difficulties, but the second, though once thought hopeless, is now perhaps more easily made plausible than the first.

DIFFICULTIES CONCERNING SENSE DATA

A number of difficulties have been noted already in the full theory, but others lie even in the arguments for sense data.

THE CERTAINTY ARGUMENT. Various objections may be made to the certainty argument. First, so far as introspective examination is concerned, our awareness is, as we have mentioned, of putative objects, not of color patches—one sees a tomato or something looking like one. Awareness of color patches as such is a different kind of observation from normal perceiving, not a sensory core within it. One may more readily be said to be directly aware of sounds or smells as such; but even then, as we saw concerning sensations, one is aware of them as public and external, not private.

Second, the assumed link between immediacy and certainty is questionable. If immediacy is put forward as an introspective characteristic of the awareness of sense data, nothing follows about its certainty because any awareness we point to as direct may be mistaken. However, if immediacy and certainty are linked conceptually, as the premise of the certainty argument suggests—if they are defined in terms of each other—then it may be that what seems to be immediate, and hence certain, awareness is not immediate. This point may be illustrated in various ways. The certainty argument claims that sensing reveals existents—that when we look at an (apparent) tomato, we cannot doubt that something red and round and bulgy exists. Strictly speaking, however, we are certain only of something red-looking; it may in fact be orange that looks red in this light. Indeed, as J. L. Austin pointed out, even statements about how a thing looks may have to be retracted. Further, the controversy over whether sense data can fail to appear as what they are throws further doubts on the incorrigibility claim, and the alternative adverbial analysis, that sense data are only sense contents, challenges the view that something exists distinct from the percipient's experience of it.

Third, the certainty argument is too ready to deny that we see physical objects in cases of illusion and distortion and to assume that we are aware of the same kind of existent in both perception and hallucination. Both these assumptions may plausibly be denied. When we look at the putative tomato, even if it is a piece of wax or a reflection of a tomato or an image on a screen, we are

still seeing a material object—wax, or the tomato "in" (via) the mirror, or a screen illuminated in a certain way. There is no need to suppose that we are aware of something else, a sense datum. Contrastingly, the common explanation of hallucinations would be that they are unusually vivid mental images confused with perceptions. Such images, like afterimages, seem to be private, but one should not assume that they are identical with what we are aware of in normal perception. The sense-datum theory can, however, reply that hallucinations are normally quite indistinguishable by the victim from normal perception and may also be integrated with a perceived background—for instance, the apparition may walk across the room and cast shadows—so if the hallucinatory images are private, so must be the data of the background. Although two entities are not necessarily identical because they are generally indistinguishable, identity may be the most plausible explanation of their indistinguishability, and the integration is very difficult to explain except on the sense-datum theory or on some form of representative realism. All the same, the sense-datum theory, if treated as an explanatory hypothesis, has the disadvantage of being very uneconomical in postulating so many distinct entities (the sense data).

THE PARTITIVE ARGUMENT. The partitive argument can be dismissed quite briefly, apart from its other troubles already mentioned. From the fact that we do not actually see the whole of an object at once, it does not follow that we do not then see the object, any more than the fact that we cannot visit all of New York at once means that we cannot visit it at all. Consequently, there is no ground for regarding what we actually see of an object as something different from it (a sense datum) or the actual seeing as some special direct awareness (sensing).

THE ARGUMENT FROM THE CONTENT OF ILLUSIONS. The argument from the content of illusions presents problems similar to those of the certainty argument. The alternative to the sense-datum answer concerning what the drunkard sees in hallucinations is "a mental image," and in double vision "one bottle looking double." Neither answer is wholly satisfactory, since the first cannot explain the integration of the image with a real background, and the second has been accused of evading the issue—looking double is not like looking blue or looking elliptical, for it involves an extra apparent object, not a differing quality of the one object.

Ordinary cases of relativity are much more easily dealt with. When one sees a round dish that looks elliptical, one is simply seeing the dish and not some elliptical existent; the theory oddly assumes that things cannot look other than they are. This assumption is linked with the notion of immediacy: It is gratuitously supposed that in seeing the dish as elliptical, one is immediately aware of an elliptical existent. However, this begs the question by equating immediacy with incorrigibility, so that what looks elliptical is said to be elliptical. Furthermore, there is no cogent ground in experience or in the argument for supposing that nonhallucinatory sense data are private to a person: The elliptical shape of the plate or even the second bottle might also be sensed by others. The privacy is best supported by arguing that sense data are "generated" by brain processes (as in the causal argument for sensations).

OTHER DIFFICULTIES. Various other criticisms of sense data have been put forward, especially by Gilbert Ryle and J. L. Austin. First, sensing is either seeing under another name—in which case there is the reduplication or regress noted concerning sensations—or else it is a myth. The notion of a mistake-proof awareness, Ryle claimed, arises from misunderstanding the character of perception words, which are achievement words or indicate the scoring of an investigational success. One cannot perceive unsuccessfully any more than one can win unsuccessfully, but that is a linguistic or conceptual matter; it does not mean that if one looks or plays, one is bound to see or to win.

Second, the theory, in speaking of sense data as existents, is simply reifying (treating as things) the sounds, smells, or looks of things. Ryle claimed a linguistic origin for this: By wrongly speaking of "seeing looks" or "smelling whiffs," which are pleonastic usages like "eating nibbles," the theory tends to treat looks and whiffs as the sort of things we can see or smell—that is, as objects—and fails to see that the point of such words is to show how we are perceiving objects. (He could hardly condemn hearing sounds, even if the other examples are correct.)

Third, Austin attacked the tendency of Moore, A. J. Ayer, and others to distinguish different senses of the word *see*: the normal sense (seeing objects) and the restricted "direct" or "actual" seeing (sensing, which is incorrigible). He claimed that the second sense is a myth: The basic fact is that one may describe the object one sees in various ways, depending on how advertent one is; for example, as a tomato or as a red object. But in these two cases it is the same thing described in different ways, not two different things; nor does it follow that there are two kinds of seeing or two senses of the word *see*.

Austin had other alleged linguistic grounds for the theory's mistakes, such as confusion of illusion with delusion, but it is doubtful whether the several different linguistic origins that he and Ryle claimed for the theory are really genuine and important. The reflective examination of experience seems a more likely origin for the theory, in view of the stress laid on it by the sense-datum philosophers. They have been so struck by the apparent immediacy of perceiving, by its apparently direct confrontation with a "given," that they have readily assumed that it does involve such an immediate awareness or confrontation; and because (on account of illusions) they cannot identify immediate awareness with the perception of a physical object, they have supposed it to be an inner awareness of special data—the sensing of sense data.

FURTHER DEVELOPMENTS

SENSE CONTENTS. As we have seen, the adverbial analysis of sensing claims that sensa no more exist as entities distinct from the sensing of them than do itches or pains; consequently, they are often referred to in this analysis as sense contents. Important advocates of this approach have been C. J. Ducasse and Ayer, and under its influence Moore modified and Russell abandoned an earlier faith in an act—object analysis (that sense data are separate entities distinct from the act of sensing). Russell's conversion to the adverbial analysis was brought about by his conclusion that the subject of awareness is a logical fiction; since the act-object analysis presupposes a subject of the act of awareness, it had to be dismissed (*Analysis of Mind*, p. 141). Probably few would follow him on this; it is, at any rate, not clear that the adverbial analysis can dispense with the subject, nor is it clear why one should wish to. Moore's *Refutation of Idealism* relied on the act-object analysis, but he later had doubts about this. He tended to see the problem as whether sense data have any existence when unperceived, or rather, unsensed; that is, whether their *esse* is *percipi* or *sentiri*. He regarded this as an open question, producing various arguments on either side at different times. Actually, the two questions are not quite identical: The adverbial analysis implies that sense data or sense contents cannot exist unsensed, while the act-object analysis is neutral on this.

It seems clear that whether sense data exist unsensed is not a question that can be settled by sensing them. Consequently, some would say it is a purely conceptual matter, one of how sense data are to be defined or how the general theory is to be framed. But factual issues are relevant and present a dilemma similar to the one of the relation of sense data to physical objects. If one accepts that sense data are generated by the brain, then it seems that they cannot exist unsensed. Even if they are only conditioned by the nervous system, they must appear different from what they really are in the unconditioned, unsensed state, thus undermining the certainty argument. At the same time, to say that a physical object is a family of sense data is scarcely meaningful if sense data do not exist unsensed; therefore, Russell at one period claimed that they do exist unsensed, calling them sensibilia in this state. More usually, however, phenomenalism is maintained; sense data do not exist unsensed, but possible ones or possibilities of sensation do.

So far as introspection is concerned, decision between the analyses depends on which sense is considered. Visual sense data, such as color shapes, would seem clearly to be distinct and to require an act-object analysis. (Afterimages are more doubtful, but anyhow are a special case.) Much the same applies to sounds and smells, which are normally experienced as external: By contrast, tactile and other bodily (somatic) sense data, such as pains or feelings of warmth or pressure, and the sensations of movement (kinesthetic data) seem clearly adverbial, as perhaps is taste; but there are marginal cases. Explanation of this variation is difficult for the theory, which would be more plausible if it could give one account of all sense data; it is also difficult to square the distinctness claimed in the act-object analysis with the privacy always claimed. Another possible line, which seems required for dreams and mental images and for hallucinations where no distinct objects are present, would be to say that while sense data seem to the person to be distinct, they are actually contents of adverbial experiences, as are sensations. However, this would undermine the claim of the theory to rely on introspective analysis.

SENSE-DATUM LANGUAGE. One suggestion that has been made is that the sense-datum philosophers have not, as they at first thought, produced a new theory of perception; they have simply introduced a new and more convenient terminology for discussing the facts of ordinary perceiving. This was accepted for a time by those who sought to see all philosophy as dealing with language and by those who, impressed by the difficulties the sense-datum theory encountered, sought to salvage something from the wreck. It is not popular now, for those with a linguistic bias have turned to the examination of ordinary language rather than to the advocacy of new terminologies, while the general decline of support for sense data has proceeded beyond this halfway house. Another reason for supposing that the sense-datum theory was only a terminology was the view that theories must be verifi-

able by observation of predicted consequences, which the sense-datum theory is not, but this seems to confuse a philosophical theory with a scientific theory.

Considered simply as a terminology, the language of sensing and sense data was claimed to have certain advantages; for example, that it is (*a*) noncommittal—one can describe the contents of one's experience independently of the physical objects they are thought to refer to—and (*b*) neater than ordinary language, for one can avoid periphrases like "there appears to be a red, bulgy tomato-like object" merely by listing the data sensed. But these are only slight advantages, and it seems that they are far outweighed by the fact that a sense-datum language cannot be truly neutral. It has been so long associated with the sense-datum theory that it must inevitably beg the question by suggesting that the data are private, transitory existents; that one is not "actually seeing" physical objects; or that in describing the scene in terms of visual and tactile data, one has described the experiences of normal perception and not of the different "reduced" phenomenological observation.

See also Perception; Phenomenalism; Realism.

Bibliography

SENSATIONS

Readers are advised to consult the historical surveys and then follow up any particular treatment that interests them with the aid of the references given. On the psychological side, see Edwin Garrigues Boring, *Sensation and Perception in the History of Experimental Psychology* (New York: Appleton-Century, 1942) and *A History of Experimental Psychology*, 2nd ed. (New York: Appleton-Century-Crofts, 1950); on the philosophical side, most useful is David Walter Hamlyn, *Sensation and Perception* (New York: Humanities Press, 1961). A stimulating but idiosyncratic discussion is given in Bertrand Russell, *The Analysis of Mind* (London: Allen and Unwin, 1921). Also, many of the books listed below under sense data discuss sensations. Gilbert Ryle's article "Sensations," in *Contemporary British Philosophy*, edited by Hywel David Lewis, 3rd series (London, 1956), attacks the whole concept; a careful study of a limited field is David Malet Armstrong, *Bodily Sensations* (London: Routledge and Paul, 1962).

SENSE DATA

The following contain statements of the sense-datum theory.

Pioneer work is found in George Edward Moore, *Some Main Problems of Philosophy* (introductory lectures of 1910–1911; London: Allen and Unwin, 1953) and his more difficult *Philosophical Studies* (London: Routledge, 1922). He kept returning to the problems of perception, not always consistently; for the advanced student, Alan Richard White, *G. E. Moore* (Oxford: Blackwell, 1958), is a useful guide with a full bibliography. Bertrand Russell was another pioneer in his lucid elementary *Problems of Philosophy* (London: Williams and Norgate, 1912) and his more significant *Mysticism and Logic* (articles of 1914; London: Allen and Unwin, 1918) and *Our Knowledge of the External World* (London: Allen and Unwin, 1914), but he tended to subordinate the topic of sense data to his special perspective theory. Clear and systematic is Henry Habberley Price, *Perception* (London: Methuen, 1932), which develops a full sense-datum theory. Another well-known account is Charlie Dunbar Broad, *The Mind and Its Place in Nature* (London: Kegan Paul, 1925); earlier and fuller statements of his views, with more attention to causal problems are in his *Scientific Thought* (London: Kegan Paul, 1923) and *Perception, Physics and Reality* (Cambridge, U.K.: Cambridge University Press, 1914). Alfred Jules Ayer gives several clear discussions, mainly from a phenomenalist point of view, in *The Foundations of Empirical Knowledge* (London: Macmillan, 1940), *Philosophical Essays* (London: Macmillan, 1954), and *The Problem of Knowledge* (London: Macmillan, 1956). For a modern version of representative realism stated in terms of sense data, see John Raymond Smythies, *Analysis of Perception* (London: Routledge and Paul, 1956).

There are many criticisms of the sense-datum theory. A general introductory survey, chiefly of Price's version, and extended criticisms are given in Rodney Julian Hirst, *The Problems of Perception* (London: Allen and Unwin, 1959). Gilbert Ryle's criticisms are found in his *The Concept of Mind* (London: Hutchinson, 1949) and *Dilemmas* (Cambridge, U.K.: Cambridge University Press, 1954); John Langshaw Austin's are in his lively defense of a commonsense approach, *Sense and Sensibilia* (London: Oxford University Press, 1962); Roderick M. Chisholm has some succinct criticisms in his more technical *Perceiving: A Philosophical Study* (Ithaca, NY: Cornell University Press, 1957); Harold Arthur Prichard, *Knowledge and Perception*, edited by W. D. Ross (Oxford: Clarendon Press, 1950), attacks Russell and the view that sensing is a form of knowing. Martin Lean, *Sense Perception and Matter* (London: Routledge and Paul, 1953), attacks Broad's version; and valuable is H. H. Price's "The Nature and Status of Sense-Data in Broad's Epistemology," in *The Philosophy of C. D. Broad*, edited by Paul Arthur Schilpp (New York: Tudor, 1959). On Moore, besides A. R. White's book, senior students should see P. A. Schilpp, ed., *The Philosophy of G. E. Moore* (Evanston, IL: Northwestern University Press, 1942), particularly C. J. Ducasse, "Moore's Refutation of Idealism"; O. K. Bouwsma, "Moore's Theory of Sense-Data"; and Paul Marhenke, "Moore's Analysis of Sense Perception"; and Moore's replies to them.

Among many critical articles, notable are Roderick Firth, "Sense-Data and the Percept Theory," in *Mind* 58 (232) (1949): 434–465, and 59 (233) (1950): 35–56, attacking the sense-datum analysis of perceptual consciousness; Winston H. F. Barnes, "The Myth of Sense-Data," in *PAS* 45 (1944–1945); and Anthony M. Quinton, "The Problem of Perception," in *Mind* 64 (253) (1955): 28–51, criticizing the concept of sense data and the arguments for them; also a more advanced discussion of the earlier versions of the sense-datum theory in the symposia "The Status of Sense-Data," in *PAS* 14 (1913–1914): 355–406, and "The Nature of Sensible Appearances," in *PAS*, Supp. 6 (1926): 142–205.

SENSE-DATUM LANGUAGE

The linguistic interpretation of sense data is developed in George A. Paul, "Is There a Problem about Sense-Data?," in *PAS*, Supp. 15 (1936), and A. J. Ayer, "The Terminology of Sense-Data," in *Mind* 54 (216) (1945): 289–312, reprinted in his *Philosophical Essays* (London: Macmillan, 1954); there is a useful discussion of it in the symposium "Seeming," in *PAS*, Supp. 26 (1952): 195–252.

SENSE CONTENTS (ADVERBIAL ANALYSIS)

The most important advocate of this approach is C. J. Ducasse, *Nature, Mind and Death* (La Salle, IL: Open Court, 1951), largely identical with his contribution to *The Philosophy of G. E. Moore*. Brief statements may be found in A. J. Ayer, *Language, Truth and Logic* (London: Gollancz, 1946) and *Philosophical Essays* (London: Macmillan, 1954). A fuller discussion, with emphasis on the differences between the senses, is C. D. Broad, "Berkeley's Argument about Material Substances," in *Proceedings of the British Academy* 28 (1942): 127ff., and his *Scientific Thought* (London: Kegan Paul, 1923).

OTHER RECOMMENDED TITLES

Anscombe, G. E. M. "On Sensations of Position." *Analysis* 22 (1962): 55–58.

Armstrong, David Malet. *Perception and the Physical World*. London: Routledge and Kegan Paul, 1961.

Casullo, Albert. "A Defense of Sense-Data." *Philosophy and Phenomenological Research* 48 (1987): 45–61.

Chisholm, Roderick. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press, 1982.

Chisholm, Roderick. "The Problem of the Speckled Hen." *Mind* 51 (1942): 368–373.

Cornman, James, "Sellars, Scientific Realism, and Sensa." *Review of Metaphysics* 23 (1969): 417–451.

Fales, Evan. *A Defense of the Given*. Lanham, MD: Rowman & Littlefield, 1996.

Fantl, Jeremy, and Robert Howell. "Sensations, Swatches, and Speckled Hens." *Pacific Philosophical Quarterly* 84 (2003): 371–383.

Huemer, Michael. "Sense-Data." Stanford Encyclopedia of Philosophy. Available from http://www.plato.stanford.edu.

Jackson, Frank. *Perception: A Representative Theory*. Cambridge, U.K.: Cambridge University Press, 1977.

Lowe, E. J. "Indirect Perception and Sense Data." *Philosophical Quarterly* 31 (1981): 330–342.

Martin, Michael G. F. "Beyond Dispute: Sense-Data, Intentionality, and the Mind Body Problem." In *History of the Mind-Body Problem*, edited by Tim Crane. New York: Routledge, 2000.

Perkins, M. *Sensing the World*. Indianapolis: Hackett, 1983.

Sellars, Wilfrid. "Phenomenalism." In his *Science, Perception and Reality*. London: Routledge, 1963.

Swartz, R. J., ed. *Perceiving, Sensing, and Knowing*. New York: Doubleday, 1965.

R. J. Hirst (1967)

Bibliography updated by Benjamin Fiedor (2005)

SENSATIONALISM

"Sensationalism," the doctrine that all knowledge is derived from sensations, takes several closely related forms. As a psychological theory it stresses the origins of knowledge and the processes by which it is acquired; it seeks to reduce all mental contents to unitary sensations and has close connections with associationism. It is sometimes, as by its acute but sympathetic critic James Ward, called presentationism. As an epistemological theory it tends toward the view that statements purporting to describe the world are analyzable into statements concerning the relations between sensations and that this analysis elucidates the meanings of the original statements. It is sometimes regarded as a form of empiricism and adopted with antimetaphysical intentions.

Sensations are usually regarded as occurrences in us, either caused by external objects (Epicurus and John Locke) or not meaningfully attributable to external causes (James Mill and Étienne Bonnot de Condillac). By some they are explicitly likened to feelings or emotions (Anaxagoras and David Hartley), and by others to images (Ernst Mach); the more modern forms, however, probably depend, even if not explicitly, on taking them all as analogous to feelings.

There is a tendency to associate sensationalism with the nineteenth and twentieth centuries, as a development of the work of the empiricists of the seventeenth and eighteenth centuries, but it actually has a long history. A study of its development takes us back to the pre-Socratics, and although in its modern forms it usually leans heavily on the distinction between sensation and perception, there were views that can be called sensationalist long before the distinction was made (for example, Protagoras held them). The distinction between sensation and perception is used because it is believed that although perception involves interpretation and, thus, the possibility of error, sensation does not. Sensationalism is therefore sometimes looked upon as the end point of the empiricists' quest for certainty and a sure foundation for knowledge.

THE GREEKS

The Greeks had no linguistic means of distinguishing between sensation and perception, but they do not appear to have considered this a serious lack. The pre-Socratics were apparently interested in perception mainly from the physiological and physical point of view; they wanted to describe processes, which they tended to see as purely mechanical (this is especially true of Empedocles

and the atomists), involving the meeting of effluences from the sense organ and the external object. But Anaxagoras introduced a feature that has some significance for an understanding of sensationalism—namely, the idea that perception involves pain. This facilitates the assimilation of all sensations to feelings referred to below.

PROTAGORAS. Protagoras, accepting the Heraclitean view that all is change or becoming and having concluded that "man is the measure of all things," found it easy to regard our constantly changing sense experiences as the objects of knowledge and to hold that all the so-called qualities of things, not merely the secondary qualities as the atomists believed, were relative to the perceiver. This turned attention to epistemological questions connected with the nature of perception.

PLATO. Perhaps Plato and Aristotle were primarily reacting against this view of Protagoras in their discussions of perception. Plato's argument in the *Republic* is that sense experience does not give knowledge but only opinion, since knowledge must be certain and cannot be of what is constantly changing—that is, sensations or the sensible world. According to some scholars—D. W. Hamlyn, for example—another view can be extracted from the later dialogue the *Theaetetus*, but this is highly controversial. Protagoras was referring to knowledge of a familiar, everyday sort. The view allegedly to be found in the *Theaetetus* is that the senses can give us this rudimentary empirical knowledge; they give us direct acquaintance with the outside world and even without interpretation can therefore give us knowledge. There is no distinction to be made, as far as the sensible world itself is concerned, between what is and what appears. Because sense experience is caused by the external world, it can be regarded as infallible. But this step is suspect both on general grounds and in relation to Plato's own insistence that the categories of right and wrong are contributed by the mind. His thought seems to be that if judgment is made by the mind and if saying that something is wrong is making a judgment, then bare sense experience, being prior to judgments of it, cannot ever be said to be wrong. It should, of course, be added that it cannot be said to be right either.

ARISTOTLE. Aristotle, in attempting to refute the sensationalism of Protagoras, stressed the element of judgment in perception and almost arrived at the distinction between sensation and perception. At the same time he appears to admit an important feature of sensationalism. Each sense has its proper object or special sensible; the proper object of hearing is sound and that of sight is color. But there are also common sensibles, qualities of objects that are not specially related to any one of the five senses but that are related to the common nature of them all, which he referred to as the common sense. These qualities are, roughly, the primary qualities motion, rest, shape, size, and number. Because there is a necessary connection between each sense and its special sensible, it is impossible for the senses to make mistakes about them; for example, hearing cannot err about the fact that it is concerned with sound and not color. This, however, does not entail any incorrigibility in the deliverances of the senses as is required by sensationalism. It simply means that each sense is necessarily concerned with its special sensible. Aristotle's claims about incorrigibility probably arise, as Hamlyn says, from an unresolved conflict between his view of the senses as both active and passive. (The senses can make mistakes only if they are active and make judgments; as mere passive receptors, they cannot. If we fail to distinguish in this way, we may think of the senses as judging infallibly.) In *De Memoria et Reminiscentia* Aristotle outlined some principles of association that look forward to later accounts.

EPICURUS. Epicurus, who believed that sense perception is the source of all knowledge, held a causal theory of perception. He did not distinguish between sensation and perception and regarded what were later called sensations as incorrigible because caused. He was an atomist and attempted a mechanical account of perception. The Stoics opposed this account and again stressed the importance of at least rudimentary judgment in perception. Their conception of *phantasiae* roughly corresponds to the conception of sensations as images; they held that these were not necessarily veridical although some of them were intuitively certain.

THE SCHOLASTICS

Problems of perception were not central in medieval philosophy except as they bore on the relation between empirical and other varieties of knowledge.

AUGUSTINE. Augustine is important on the subject of perception perhaps only because he saw that it is not meaningful to talk of sensations as either true or false; these terms can be applied only to judgments. He simply assumed that sense impressions correspond to the external world but regarded the knowledge thus obtained as of the lowest kind.

THOMAS AQUINAS. Thomas Aquinas followed Aristotle in his views on perception to the extent of holding that it involves the reception of a sensible form without matter, but this produces a change in the soul, not merely, as for Aristotle, in the sense organ. Sensory images (*phantasmata*) are received passively, but they are images of external objects. They have the peculiarity that we are not aware of them. The mind abstracts universal qualities from these and uses them in making judgments. The senses and the intellect are closely connected: *Nihil est in intellectu quod non prius in sensu* (Nothing is in the intellect that was not first in the senses). Because our perceptions involve judgments, they may or may not be veridical, but the *phantasmata* are not appropriately called either. This, with the fact that the *phantasmata* are images of something, prevented Thomas from being a sensationalist, but he was very close to being one in spirit and utterance.

OCKHAM. Although William of Ockham differed from Thomas in many ways, he also distinguished a sensible and an intellectual element in cognition. Those cognitions that involve only immediate experiences are said to be perfect. Error arises in judgment, but when we are directly apprehending something, we are not in error.

THE SEVENTEENTH AND EIGHTEENTH CENTURIES

Sensationalism proper can perhaps be regarded as the product of a steady development of empiricist ideas from the seventeenth century to the nineteenth. Thomas Hobbes is sometimes credited with its inception, but his sensationalism is rudimentary. He did have some conception of the association of ideas and, of course, contributed to the foundations of empiricism.

Largely because of the climate of scientific opinion, involving as it did a growing belief in the importance of observation and experiment, the philosophers of the seventeenth century were much concerned with problems of perception. They were especially interested in the elimination of errors arising from sense experience and in the attempt to make our knowledge of the natural world as reliable as possible. The rationalists attempted to show that knowledge could be based on indubitable truths of reason, independent of sense experience. The empiricists sought a hard core of indubitable truths involved in sense experience upon which all knowledge could be based.

GALILEO. Galileo Galilei distinguished between primary and secondary qualities and thought that secondary qualities existed only as sensations in us. They are, however, caused by primary qualities in objects, especially by shape and motion.

HOBBES. Under the influence of Galileo, René Descartes, Marin Mersenne, and Pierre Gassendi, Hobbes developed the philosophy of motion into what must be the most thoroughgoing materialism there has ever been. For him all our inquiries must start from sense experience, but there are certain principles—for instance, that motion cannot be understood to have any other cause besides motion—which we know independently of sense experience and upon which other knowledge depends. Nothing exists but matter in motion, so sensations are material changes in us that somehow mediate between motions in the external world and the minute motions of our bodily parts. Hobbes assumed the existence of external motions causing our sensations; knowledge of these "objects" can come only through sensations. This does not entail the empiricist view that all knowledge is reducible to knowledge of sensations; Hobbes was in general rationalist, for he held that certain truths of reason are essential even for that knowledge of the natural world which depends upon sensation.

LOCKE. Locke's work marks the beginning of the growth of sensationalism proper, although he was not himself a sensationalist just because he did not develop his particular form of empiricism consistently. His "ideas of sensation" are close to what were later called simply sensations, but his representative theory of perception and his assertion of the existence of substance entail that in spite of explicit claims he relied on knowledge which did not come entirely through sensation.

BERKELEY. George Berkeley attempted to remove this inconsistency in his attack on material substance and representative perception. Whether we view his reliance on God as the unempirical importing of a concept merely for the purpose of filling an embarrassing gap—that is, to allow us to hold that objects continue to exist when no human being is perceiving them—or as the attempt to delineate a concept that is logically necessitated by our experience, Berkeley's account of ideas brings us very near sensationalism. There is no talk of external objects that are composed of any material different in kind from what we directly know—that is, ideas. Later sensationalism can be regarded as comparable to Berkeley's system without God, with all its problems as well as its advantages.

HUME. David Hume continued this development, in one direction by rejecting mental substance, which was retained by Berkeley, as well as material substance. The world for us, as far as we can justifiably say in philosophical contexts, consists of impressions and ideas, and knowledge is of relations between these. Hume was not, however, as great a skeptic as is often alleged. We have, naturally, certain beliefs—for example, in the external world and in causal efficacy—which cannot be rationally supported. When philosophy fails to provide this rational support, so much the worse for philosophy. If Hume had not been affected by the common view that knowledge implies certainty, he would no doubt have admitted these "natural" beliefs as knowledge and thus have been farther from sensationalism in his official theory than he actually was.

Sensationalism in its fullest sense is best seen in the works of Hume's lesser-known contemporaries Hartley and Condillac. Hartley's work was later developed by James Mill, and its most thoroughgoing exponent in the nineteenth century was perhaps Mach.

HARTLEY. Hartley was a medical man; his interests were largely physiological, and his work stimulated the development of a school of psychology. His basic concepts were sensations and the association of ideas, for which he admitted a debt to Locke and Isaac Newton. All mental occurrences originate in sensations caused by vibrations of minute particles of the brain set off by external stimulation. Simple ideas are "copies" of sensations—that is, physiologically they are tiny vibrations corresponding in character to the original vibrations and left behind by sensations when the stimulus is withdrawn. Complex ideas are built up from these by association according to certain discoverable principles. The vibrations occur in a subtle elastic fluid in the medullary substance of the nerves and brain. This mechanical account is reminiscent of Hobbes's view and admittedly owes a debt to Newton's mechanistic philosophy. The conception of the association of ideas springs from Locke, and the consequent contention that ideas are copies of sensations echoes Hume's account of impressions and ideas. Hartley's theory leads to the conclusion that we are aware only of occurrences within ourselves but that these depend for their character on the external world. There is a twofold correspondence, between ideas and sensations and between sensations and stimuli.

JAMES MILL. James Mill accepted Hartley's basic conceptions and developed the psychological side of the theory. Hartley had expressed in terms of vibrations two principal determinants of the strength of association—the vividness of the sensations and the frequency of their conjoint occurrence. Mill discussed these principles in some detail, without Hartley's preoccupation with vibrations, contrasting his principles of association with Hume's and using some rather unsatisfactory arguments for preferring his own. In place of Hume's contiguity in time and place, causation, and resemblance, Mill put synchronous order and successive order, which include causation as a special case, and vividness and frequency, which include resemblance as a special case. He went further than Hartley in considering the relation of sensations to the external world; external objects for him are "clusters of sensations." Most of our beliefs about them depend on sight and sensations of color, with which we associate the other properties we attribute to them.

CONDILLAC. While Hartley was writing in England, Condillac was developing similar ideas in France. He was a disciple of Locke, and his first book was largely an exposition of Locke's philosophy. In his *Traité des sensations*, he developed his own psychological theory, largely in opposition to the various current conceptions of innate ideas. He set out to show that all knowledge is "transformed sensation" and does not depend upon anything else, even, as Locke would have had it, reflection. He examined the nature and power of each of the senses by imagining a statue that has all the human faculties but has never had a sense impression. He then allowed its senses to be activated, one by one and in various combinations, and asserted that the results showed how all knowledge can gradually be constructed. He concluded that people consist of their experiences and that what they perceive is their own mental occurrences. Unlike Hartley, he did not try to give a mechanical account of these occurrences, being more concerned with psychology than physiology, and he admitted the reality of the soul. He had a considerable influence on the beginnings of British psychological thought through James Mill and J. S. Mill, Alexander Bain, and Herbert Spencer.

MACH AND TWENTIETH-CENTURY EMPIRICISM

Whereas Hartley, Condillac, and the Mills were interested in sensations mainly in relation to psychology, ethics and politics, Mach's interest sprang from an attempt to provide an analysis of the methods of the physical sciences. His sensationalism was associated with a search for a solid foundation for scientific statements and with a desire to free science of all metaphysics. He held that only state-

ments which are directly verifiable in sense experience can finally be accepted as conclusions in the sciences. He concluded that all scientific statements are analyzable into statements about the relations between our sensations and that nothing can be said, scientifically, about anything beyond this. In a sentence reminiscent of James Mill he said, "The world is my sensation." It follows, also, that the various branches of science do not differ in subject matter but only in their approach to the subject matter, which is—alike for all—sensations; this was the basis of the "unity of science" movement and the logical positivism of the Vienna circle.

Mach's work was very much in harmony with the spirit of his time, especially in relation to the physical sciences, and has had an important influence on later philosophical thought. He admitted a debt to Berkeley and Hume and a number of his philosophically minded scientific contemporaries. His idea that the world is composed of "elements" which can be regarded either as sensations or the constituents of physical objects has close connections with Bertrand Russell's neutral monism and logical atomism, and his description of the aims of science is similar to that of pragmatism and operationism. In one way or another, most empiricist thought about science during the twentieth century has been influenced by his work. Recent philosophical theories of perception involving sense data or sensa are in the direct line of descent insofar as they stress the mind dependence of sense data, our direct awareness of or acquaintance with them, and the alleged incorrigibility of certain sorts of statements about them. Such theories can be regarded as attempts to refashion Mach's form of sensationalism in order to avoid some of the obvious objections to it.

Sensationalism and related theories all suffer from one defect, which renders the whole approach suspect; under the heading "sensations" they class together things that it is important to distinguish—for example, such sensible qualities as colors and sounds; bodily aches and pains; desires and emotions; and such feelings as dizziness, anger, and jealousy. We would not normally be prepared to class all these as experiences, but certain empiricist contentions—for example, that we know colors only through their effects on us—can make it seem superficially plausible to call them all sensations. Just because this blurs the distinctions between various things included under the heading, sensationalism as a general theory gains plausibility. Toothaches and certain feelings have an air of immediacy and unmistakability that may lead us to suppose that color sensations, since they, too, are sensations, are ultimate and incorrigible data for the construction of a world picture. I can be certain that I have a toothache, and no one can be better justified than I in asserting or denying this. If color sensations can be assimilated to toothaches, there might seem to be some hope of arriving at incorrigible statements about the external world. Hence, the importance of the clue afforded by Anaxagoras's view that perception involves pain. A close examination of experiences of color and other sorts of experience reveals that the necessary assimilation is seriously misleading; moreover, it brings in its train enormous difficulties for an account of science. On one hand, incorrigibility can be achieved, if at all, only with the loss of the publicity of the statements concerned; on the other hand, it is difficult or impossible to show how scientific problems could ever arise if sensationalism were correct, since there is no reason that any particular combination of sensations should or should not follow any other.

In fact, the word *sensation* suffers from ambiguities similar to those involved in the word "idea" as used by Locke and Berkeley; as *sensation* must do even more work than *idea*, the ambiguities are correspondingly more serious. The view of science that springs from sensationalism, according to which science describes but does not explain, suffers further from insufficient consideration of the nature of description and its relation to explanation and from a failure to appreciate the difficulties involved in the idea of describing sensations.

See also Mach, Ernst; Pearson, Karl.

Bibliography

GENERAL

Armstrong, D. M. *Bodily Sensations*. London: Routledge and Paul, 1962. Consideration of the logic of sensations.

Boring, E. G. *Sensation and Perception in the History of Experimental Psychology*. New York: Appleton-Century, 1942. Useful discussions of some of the ideas of sensationalism from the point of view of psychology.

Copleston, F. C. *History of Philosophy.* 9 vols. Westminster, MD: Newman Press, 1946–1975.

Hamlyn, D. W. *Sensation and Perception*. New York: Humanities Press, 1961. Excellent historical survey with a useful bibliography.

Hirst, R. J. *The Problems of Perception*. London: Allen and Unwin, 1959. Account of philosophical problems connected with sensationalism.

Passmore, J. A. *A Hundred Years of Philosophy*. London: Duckworth, 1957.

THE GREEKS

Aristotle. *De Anima*. Translated by K. Foster and S. Humphries, with the commentary of St. Thomas Aquinas. London, 1951.

Aristotle. *Works.* 12 vols, edited by J. A. Smith and W. D. Ross. Oxford, 1910–1952. See Vol. VIII, *Metaphysics*, especially Γ, 5 (1928), and *De Memoria et Reminiscentia*, in Vol. III (1931). These works contain most of Aristotle's important views on sensation.

Burnet, John. *Early Greek Philosophy*. London, 1892.

Cornford, F. M. *Plato's Theory of Knowledge*. London: K. Paul, Trench, Trubner, 1935. Contains Plato's *Theaetetus*, which considers the place of perception in knowledge, with a commentary.

Gulley, Norman. *Plato's Theory of Knowledge*. London: Methuen, 1962. Authoritative and readable treatment.

Kirk, G. S., and J. E. Raven. *The Presocratic Philosophers*. Cambridge, U.K.: Cambridge University Press, 1957. This and Burnet, op. cit., give accounts of pre-Socratic views on sensation.

Plato. *The Collected Dialogues,* edited by Edith Hamilton and Huntington Cairns. New York: Pantheon, 1961. See the *Republic, Protagoras, Theaetetus*, and *Timaeus*.

Randall, J. H. *Aristotle*. New York: Columbia University Press, 1960.

THE SCHOLASTICS

Copleston, F. C. *Aquinas*. Harmondsworth, U.K., Penguin, 1955. Most easily accessible account of Thomas's thought.

Fremantle, Anne, ed. *The Age of Belief*. New York: New American Library, 1955. Selections from the writings of the Scholastics, with commentary.

Gilson, Étienne. *Christian Philosophy in the Middle Ages*. London, 1953. Valuable general account.

Thomas Aquinas. *Commentary on Aristotle's De Anima*. See Aristotle, *De Anima* (above).

Thomas Aquinas. *Summa Theologiae* (selections). In *Basic Writings of St. Thomas Aquinas*, annotated by Anton C. Pegis, ed. 2 vols. New York: Random House, 1945. Vol. I, Questions 78–86.

THE SEVENTEENTH AND EIGHTEENTH CENTURIES

Bacon, Francis. *Novum Organum* (1620). Translated by G. W. Kitchin. Oxford, 1855.

Bower, G. S. *Hartley and James Mill*. London, 1881.

Condillac, É. B. de. *Essai sur l'origine des connoissances humaines*. Amsterdam, 1746.

Condillac, É. B. de. *Traité des sensations* (1754). Translated by G. Carr as *Treatise on Sensations*. London, 1930. This and above work influenced British empiricism.

Galileo Galilei. *Il Saggiatore* (1623). Translated in part by Stillman Drake in *Discoveries and Opinions of Galileo*. Garden City, NY: Doubleday, 1957.

Hartley, David. *Observations on Man*. London, 1749. An early version of sensationalism that influenced both philosophers and psychologists.

Hobbes, Thomas. *De Corpore* (1655). Translated as *Elements of Philosophy concerning Body*. London, 1839.

Hobbes, Thomas. *Leviathan* (1651), edited by Michael Oakeshott. New York, 1947; paperback ed., 1962.

Locke, John. *Essay concerning Human Understanding* (1690). Edited by A. C. Fraser. Oxford, 1894.

Mill, James. *The Analysis of the Phenomena of the Human Mind,* edited by J. S. Mill. London, 1829. Development of Hartley's sensationalism.

THE NINETEENTH AND TWENTIETH CENTURIES

Alexander, Peter. *Sensationalism and Scientific Explanation*. London: Routledge and Paul, 1963. Criticism of the sensationalist account of science.

Ayer, A. J. *Foundations of Empirical Knowledge*. London: Macmillan, 1940. One version of the sense-datum theory.

Ayer, A. J., ed. *Logical Positivism*. Glencoe, IL: Free Press, 1959.

Ayer, A. J. *Philosophical Essays*. London: Macmillan, 1954.

Bain, Alexander. *The Senses and the Intellect*. London, 1855.

Broad, C. D. *Scientific Thought*. London: Kegan Paul, 1923. Another modern sense-datum theory.

Carnap, Rudolf. *The Unity of Science*. Translated by Max Black. London: Kegan Paul, 1934.

Clifford, W. K. *Lectures and Essays*. London, 1879.

Feigl, Herbert, and Wilfrid Sellars, eds. *Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts, 1949. Useful collection of papers, many of which represent modern developments from sensationalism.

Frank, Philipp. *Modern Science and Its Philosophy*. Cambridge, MA: Harvard University Press, 1949. Historical account of the development of logical positivism.

George, W. H. *The Scientist in Action*. London: Williams and Norgate, 1936. Science seen as the relating of pointer readings.

Herbart, J. F. *A.B.C. of Sense-Perception*. New York, 1896.

Herbart, J. F. *Lehrbuch zur Psychologie*. Königsberg, Germany, 1816. Translated by M. K. Smith as *Text Book of Psychology*. London, 1891.

James, William. *Essays in Radical Empiricism*. London: Longman, 1912.

Kraft, Victor. *The Vienna Circle*. New York: Philosophical Library, 1953.

Mach, Ernst. *Die Analyse der Empfindungen*. Jena, Germany, 1906. Translated by C. M. Williams as *The Analysis of Sensations*. Chicago, 1914.

Mach, Ernst. *Populärwissenschaftliche Vorlesungen*. Leipzig, 1894. Translated by T. J. McCormack as *Popular Scientific Lectures*. Chicago: Open Court, 1897. Mach's works are classics of modern sensationalism.

Mill, J. S. *An Examination of Sir William Hamilton's Philosophy*. London, 1865.

Moore, G. E. *Philosophical Studies*. London: Routledge, 1922.

Moore, G. E. *Some Main Problems of Philosophy*. London: Allen and Unwin, 1953. This and above work contain another variety of sense-datum theory.

Neurath, Otto, Rudolf Carnap, and C. W. Morris, eds. *International Encyclopedia of Unified Science*. Chicago: University of Chicago Press, 1938–1962.

Pearson, Karl. *The Grammar of Science*, 3rd ed. London, 1911.

Poincaré, Henri. *La science et l'hypothèse*. Paris: Flammarion, 1902. Translated by W. J. Greenstreet as *Science and Hypothesis*. London: Walter Scott, 1905.

Poincaré, Henri. *La valeur de science*. Paris: Flammarion, 1905. Translated by G. B. Halsted as *The Value of Science*. New York: Science Press, 1907.

Price, H. H. *Perception*. London: Methuen, 1932. Perhaps the most detailed and elegant version of the sense-datum view.

Russell, Bertrand. *The Analysis of Mind*. London: Macmillan, 1921.

Russell, Bertrand. *Our Knowledge of the External World*. London: Allen and Unwin, 1914.

Russell, Bertrand. "The Philosophy of Logical Atomism." In *Logic and Knowledge*, edited by R. C. Marsh. London: Allen and Unwin, 1956.

Schilpp, P. A., ed. *The Philosophy of Bertrand Russell*. La Salle, IL: Open Court, 1946.

Ward, James. *Psychological Principles*. Cambridge, U.K.: Cambridge University Press, 1918. Attack on sensationalism in psychology.

OTHER RECOMMENDED TITLES

Agassi, Joseph. "Sensationalism." *Mind* 75 (1966): 1–24.

Ayers, Michael. *Locke: Epistemology and Ontology*. New York: Routledge, 1991.

Bealer, G. "The Incoherence of Empiricism." *PAS*, Supp. 66 (1992): 99–138.

Bennett, Jonathan. *Locke, Berkeley, Hume: Central Themes*. Oxford: Oxford University Press, 1973.

Campbell, John. "Berkeley's Puzzle." In *Conceivability and Possibility*, edited by Tamar Szabo Gendler. Oxford: Clarendon Press, 2002.

Chisholm, Roderick. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.

Chisholm, Roderick. "The Problem of Empiricism." *Journal of Philosophy* 45 (1948): 512–517.

Dicker, Georges. *Perceptual Knowledge*. Dordrecht: Reidel, 1980.

Firth, Roderick. *In Defense of Radical Empiricism,* edited by John Troyer. Lanham, MD: Rowman & Littlefield, 1998.

Fumerton, Richard. *Metaphysical and Epistemological Problems of Perception*. Lincoln: University of Nebraska Press, 1985.

George, Rolf. "Kant's Sensationalism." *Synthese* 47 (1981): 229–256.

Moser, Paul. *Knowledge and Evidence*. Cambridge, U.K.: Cambridge University Press, 1989.

Pollock, J. *Knowledge and Justification*. Princeton, NJ: Princeton University Press, 1974.

Peter Alexander (1967)
Bibliography updated by Benjamin Fiedor (2005)

SENSATIONS

See *Sensa*

SENSE

"Sense" is the distinctive central notion in theories of thought and language inspired by the later work of Gottlob Frege ("sense" translates Frege's *Sinn*). For Frege what we think (not the act of thinking it) is a thought, an abstract object. Thoughts have quasi-syntactic structure. Any simple or complex constituent of a thought, even the thought itself, is a sense; thus, senses are abstract. Frege assumes that it is irrational to assent to a thought and simultaneously dissent from it. Since someone misled about astronomy may rationally combine assent to the thought that Hesperus is Hesperus with dissent from the thought that Hesperus is Phosphorus, the thoughts are distinct. Although the names "Hesperus" and "Phosphorus" have the same reference, they express different sense, two modes of presentation of one planet. The role of a sense is to present the thinker with a reference—that is, something on which the truth-value (truth or falsity) of the thought depends; if the sense fails to present a reference, the thought lacks a truth-value. For Frege the truth-value of a thought is independent of where, when, and by whom it is thought. Thus, what reference a constituent sense presents is independent of when, where, and by whom it is thought. Sense determines reference, not vice versa.

Frege used his notion of sense to analyze the semantics of thought attributions in natural language, as in the sentence "Someone doubts that Hesperus is Phosphorus." On Frege's account expression within such "that" clauses refer to their customary senses. This explains the presumed failure of the inference from that sentence and "Hesperus is Phosphorus" to "Someone doubts that Hesperus is Hesperus": The two names have different references within "that" clauses, for their customary senses are different. If sense determines reference, then the sense of "Hesperus" in "Someone doubts that Hesperus is Phosphorus" defers from its sense in "Hesperus is Phosphorus," since the reference differs. By appeal to iterated attributions such as "He doubts that she doubts that Hesperus is Phosphorus," it can be argued that Frege is committed to an infinite hierarchy of senses. His account involves the assignment of senses to natural-language expressions. However, in order to understand many words (e.g., proper names and natural-kind terms), there is arguably no particular way in which one must think of their reference; they do not express senses common to all competent speakers. Fregeans therefore distinguish sense from linguistic meaning but in doing so sacrifice Frege's original account of thought attributions.

Sense must also be distinguished from linguistic meaning for context-dependent expressions such as "I." Two people may think "I am falling" and each refer to themselves, not the other. Since the references are distinct and sense determines reference, the senses are distinct, even though the mode of presentation is the same. Others cannot think the sense that one expresses with "I"; they can only think about it. Communication here does not amount to the sharing of thoughts, and "You think that I am falling" does not attribute to the hearer the

thought that the speaker expresses with "I am falling." In contrast, the linguistic meaning of "I" is the same for everyone; it consists in the rule that each token of "I" refers to its producer. Unlike a sense, the rule determines reference only relative to context. Such cases reveal tensions within Frege's conception of sense. Sense cannot be both what determines reference and how it is determined. Since senses can be qualitatively identical but numerically distinct, they are not purely abstract objects, if qualitatively identical purely abstract objects must be numerically identical.

Although Fregeans distinguish sense from linguistic meaning, they still treat a given speaker on a given occasion as expressing senses in words. Frege gave the impression that the sense expressed by a word was a bundle of descriptions that the speaker associated with it: the word refers to whatever best fits the descriptions. However, this descriptive model of reference has fared badly for proper names and natural-kind terms. Nondescriptive models may also allow different routes to the same reference, but that is a difference in sense only if it is a difference in presentation to the thinker.

In spite of these problems a role for something like sense remains. An account is needed of the deductions that thinkers are in a position to make. When, for example, is one in a position to deduce "Something is black and noisy" from "That is black" and "That is noisy"? It is necessary but not sufficient that the two tokens of "that" refer to the same thing, for, even if they do, the thinker may lack evidence to that effect: Perhaps one refers through sight, the other through hearing. What is needed is more like identity of sense than identity of reference. Thus, the theory of rational inference may still require a notion of sense. It does not follow that thinkers are always in a position to know whether given senses are identical, for it is not obvious that they are always in a position to know what deductions they are in a position to make.

See also Frege, Gottlob; Proper Names and Descriptions; Reference; Semantics.

Bibliography

Burge, T. "Sinning against Frege." *Philosophical Review* 88 (1979): 398–432.

Campbell, J. "Is Sense Transparent?" *Aristotelian Society Proceedings* 88 (1987–1988): 273–292.

Dummett, M. *Frege: Philosophy of Language.* 2nd ed. Cambridge, MA: Harvard University Press, 1981.

Forbes, G. "The Indispensability of *Sinn.*" *Philosophical Review* 99 (1990): 535–563.

Kripke, S. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980.

McDowell, J. "On the Sense and Reference of a Proper Name." *Mind* 86 (1977): 159–185.

Peacocke, C. "Sense and Justification." *Mind* 101 (1992): 793–816.

Perry, J. *The Problem of the Essential Indexical and Other Essays.* New York: Oxford University Press, 1993.

Salmon, N. *Frege's Puzzle.* Cambridge, MA: MIT Press, 1986.

Timothy Williamson (1996)

SENSE DATA

See *Sensa*

SENSIBILIA

See *Phenomenalism*

SERVETUS, MICHAEL

(1511–1553)

Michael Servetus, the Spanish theologian and physician, was born in Spanish Navarre and was burned at the stake in Geneva. In the history of medicine he is remembered for having been the first to publish a description of the pulmonary circulation of the blood, and in the history of theology, he is noted for his systematic refutation of the Nicene doctrine of the Trinity. In philosophy, he developed a Christocentric pantheism that included elements from the Neoplatonic, Franciscan, and kabbalistic traditions. It should be pointed out, however, that he believed that natural philosophy should be grounded in empirical investigation.

After studying the three biblical languages as well as mathematics, philosophy, theology, and law at the universities of Zaragoza and Toulouse, Servetus, in the capacity of secretary, accompanied Juan de Quintana, the Franciscan confessor of Emperor Charles V, to the latter's coronation in Bologna. Breaking with the imperial court, he went on his own to Basel, where he sought out John Oecolampadius, and then went on to Strasbourg, where he had some contact with Martin Bucer and, in particular, Wolfgang Capito. In nearby Hagenau he had printed his *De Trinitatis Erroribus* (1531) and, in response to Bucer's critique, the more moderate and more Christologically oriented *De Trinitate* (1532). In Strasbourg Servetus met Kaspar Schwenkfeld, from whom he may

have taken over a heretical idea about the celestial flesh of Christ. In Strasbourg he may also have come in contact with the Anabaptists, whose views on baptism he was later to espouse. By way of Basel, where he tried to get Desiderius Erasmus's approval of his *De Trinitate,* he went to Lyons, where he worked as a proofreader and began his study of medicine under the Neoplatonizing Symphorien Champier. Next he went to Paris, where by chance he met John Calvin and got into trouble with the medical faculty over his views on astral influences. His *Apologetica Disceptatio pro Astrologia* (1538) marks an important turning point in Servetus's evaluation of the place of Greek philosophy. Whereas before he had regarded the influence of philosophy on theology as corrupting, he was now prepared to speak of "*divinus* Plato," on whose authority he defended astrology. After establishing himself in Vienne as physician to the archbishop, he engaged in correspondence with Calvin and composed the recently discovered and identified *Declaratio Jesu Christi Filii Dei* (c. 1540). Out of this grew his more massive *Restitutio Christianismi* (1553). Through the machinations of Calvin himself, Servetus was apprehended and tried for heresy, first in Catholic Lyons and then, after his escape, in Calvinist Geneva, where, after refusing to recant, he was burned at the stake.

Servetus's view of nature, history, and salvation was centered on the figure of Jesus Christ, whom he considered to be in a quite physical sense the Son of God. Servetus declined, however, to call the earthly Son eternal and declined to call either the Word or the Spirit *personae*; rather, he called them, neutrally, *res*—that is, in a modalist sense, the faces, forms, images, or manifestations of God. He mistakenly regarded the traditional *hypostasis* (*persona*) and *substantia* as equivalent, and hence, to avoid what he considered an unbiblical tritheism, he called the Father or Jehovah alone God. Before the Incarnation the Word was Elohim, or Uncreated Light. Indeed, this Light, or alternatively Christ (as distinguished from the earthly Son, Jesus), was also "the eternal sea (*pelagus*) of ideas." The Spirit has always been a Power of God, working outwardly in the world as his breath (*flatus*) and inwardly as the agitation, or motion, of the human spirit at regeneration.

The way in which the Uncreated Light became the Second Adam in Mary was for Servetus paradigmatic of the process by which creative Light was ever penetrating matter to form minerals, plants, animals, and all created things. For Servetus "even the treasures of natural science are hidden in Christ." Connected with his speculation on Light was Servetus's concept of the Shadow, according to which he was able to regard all of the Old Testament and all religion outside the Bible as a shadowing forth of the Son that was to be born of Mary. He cherished the old Law as a pregnant woman bearing the embryonic Christ until the fullness of time.

Servetus rejected post-Constantinian (post-Nicene) Catholicism because of its alleged tritheism and its use of political force in the realm of conscience. He also opposed the Reformation churches because of their use of force, their denial of free will in accepting redemptive grace, and their neglect of sanctification, which he understood as communicated in an almost physical sense through the believers' baptism at the age of thirty (in imitation of Jesus). Nevertheless, in common with the Spiritual Libertines and some Anabaptists, Servetus held to the provisional death of every soul with the body pending the general resurrection. Under the influence of Joachimite speculation, he believed that the true church would be restored in the year 1560.

Bibliography

WORKS BY SERVETUS

The two works on the Trinity have been translated by E. M. Wilbur as *Two Treatises of Servetus on the Trinity,* Harvard Theological Studies (Cambridge, MA: Harvard University Press, 1932), Vol. 16. See also Charles D. O'Malley, *Michael Servetus: A Translation of His Geographical, Medical and Astrological Writings* (Philadelphia: American Philosophical Society, 1953).

WORKS ON SERVETUS

For literature on Servetus, see Roland Bainton, *Hunted Heretic: The Life and Death of Michael Servetus, 1511–1553* (Boston: Beacon Press, 1953), which includes a bibliography; Wolrad Emde, "Michael Servetus als Renaissancephilosoph und Restitutionstheologe," in *Zeitschrift für Kirchengeschichte* 50 (1941): 96–131, especially good on *Restitutio Christianismi*; and Marcial Solana, *Historia de la filosofia española* (Madrid: Real Academia de Ciencias Exactas, Físicas y Naturales, 1941), Vol. I, pp. 629–681, especially good on *De Trinitate.*

George Hunston Williams (1967)

SETH, ANDREW

See *Pringle-Pattison, Andrew Seth*

SET THEORY

Set theory is a mathematical theory of collections, "sets," and collecting, as governed by axioms. Part of its larger significance is that mathematics can be reduced to set

theory, with sets doing the work of mathematical objects and their collections and set-theoretic axioms providing the basis for mathematical proofs. With this reduction in play, modern set theory has become an autonomous and sophisticated research field of mathematics, enormously successful at the continuing development of its historical heritage as well as at analyzing strong propositions and gauging their consistency strength.

Set theory arose in mathematics in the late nineteenth century as a theory of infinite collections and soon became intertwined with the development of analytic philosophy and mathematical logic. The subject was then developed as the logical distinction was being clarified between "falling under a concept," to be transmuted in set theory to "$x \in y$", x is a member of y, and subordination or inclusion, to be transmuted in set theory to "$x \subseteq y$", x is a subset of y. That set theory is both a field of mathematics and serves as a foundation for mathematics emerged early in this development.

In what follows, set theory is presented as both a historical as well as an epistemological phenomenon, driven forward by mathematical problems, arguments, and procedures. The first part describes the groundbreaking work of Georg Cantor on infinite sets analyzed in terms of power, transfinite numbers, and well-orderings. The next two parts describe the subsequent transmutation of the notion of set through axiomatization, a process to be associated largely with Ernst Zermelo. Next will come a description of the work of Kurt Gödel on the constructible sets, work that made first-order logic central to set theory, followed by a description of the work of Paul Cohen on forcing, a method that transformed set theory into a modern, sophisticated field of mathematics. The last section describes the modern investigation of relative consistency in terms of forcing, large cardinals, and inner models.

POWER, NUMBER, AND WELL-ORDERING

Set theory was born on that day in December 1873 when Cantor established that *the continuum is not countable*. The concepts here are fundamental: Taking infinite collections as unitary totalities, a set is *countable* if it is in one-to-one correspondence with the set of natural numbers {0, 1, 2, … }, and the *continuum* is the linear continuum regarded extensionally as a collection of points corresponding to the real numbers. In a 1878 publication Cantor investigated ways of defining one-to-one correspondences between sets. For sets of real numbers and the like, he stipulated that two sets have the same *power* if there is a one-to-one correspondence between them and that a set x has a *higher* power than a set y if y has the same power as a subset of x yet x and y do not have the same power. He managed to show that the continuum, the plane, and generally n-dimensional Euclidean space all have the same power, but at this point in mathematics there were still only the two infinite powers as set out by his 1873 proof. Cantor at the end of his 1878 publication conjectured:

> Every infinite set of real numbers either is countable or has the power of the continuum.

This was the *Continuum Hypothesis* (CH) in its nascent context. The *Continuum Problem* would be to resolve this hypothesis, and viewed as a primordial problem it would stimulate Cantor both to approach the real numbers in an increasingly arithmetical fashion and to grapple with fundamental questions of set existence.

In his magisterial *Grundlagen* of 1883, Cantor developed the *transfinite numbers* and the key concept of *well-ordering*. Investing the "symbols of infinity" of his early trigonometric series investigations with a new autonomy, Cantor conceived of the transfinite numbers as being generated by the operations of taking successors and of taking limits of increasing sequences. Extending beyond the finite 0, 1, 2, … , the progression of transfinite numbers could be depicted, in his later notation, in terms of natural extensions of arithmetical operations:

$$0, 1, 2, \ldots\ \omega, \omega+1, \omega+2, \ldots\ \omega+\omega(=\omega\cdot 2), \ldots\ \omega\cdot 3, \ldots\ \omega\cdot\omega(=\omega^2), \ldots\ \omega^3, \ldots\ \omega^\omega, \ldots$$

Definition. A binary relation $\prec$ is a *linear ordering* of a set a if it is transitive, that is, $x \prec y$ and $y \prec z$ implies $x \prec z$, and trichotomous, that is, for $x, y \in a$, exactly one of $x \prec y$, $x = y$, or $y \prec x$ holds.

A relation $\prec$ is a *well-ordering* of a set a if it is a linear ordering of the set such that every non-empty subset has a $\prec$-least element.

Well-orderings convey the sense of sequential counting, and the transfinite numbers serve as standards for gauging well-orderings. As Cantor pointed out, every linear ordering of a finite set is already a well-ordering and all such orderings are isomorphic, so that the general sense is only brought out by infinite sets. For these there could be non-isomorphic well-orderings. For example the set of natural numbers {0, 1, 2, …}, that is, the predecessors of ω, can be put into one-to-one correspondence with the predecessors of ω + ω by sequentially counting the evens before the odds. In fact all the infinite transfinite numbers in the above display are countable. Cantor

called the set of natural numbers the first number class (I) and the set of numbers whose predecessors are in one-to-one correspondence with (I) the second number class (II). Cantor conceived of (II) as bounded above according to a limitation principle and showed that (II) itself is not countable. Proceeding upward, Cantor called the set of numbers whose predecessors are in one-to-one correspondence with (II) the third number class (III), and so forth. In this way Cantor conceived of ever higher powers as represented by number classes and moreover took every power to be so represented. With this "free creation" of numbers, Cantor then propounded in section 3 of the *Grundlagen* a basic principle that was to drive the analysis of sets:

> It is always possible to bring any *well-defined* set into the form of a *well-ordered* set.

He regarded this as a "an especially remarkable law of thought which through its general validity is fundamental and rich in consequences." Sets are to be well-ordered and thus to be gauged via the transfinite numbers of his structured conception of the infinite.

The transfinite numbers provided the framework for Cantor's two approaches to the Continuum Problem, one through power and the other through definable sets of real numbers, these each to initiate two vast research programs. As for the first, Cantor in the *Grundlagen* established results that reduced the Continuum Problem to showing that the continuum and the second number class have the same power. However, despite several announcements Cantor could never develop a workable correlation, an emerging problem being that he could not define a well-ordering of the real numbers. As for the approach through definable sets of real numbers, Cantor showed that "CH holds for closed sets." Closed sets are a very simple kind of definable set of real numbers, and Cantor showed that a closed set either is countable or has the power of the continuum. He thus reduced the Continuum Problem to determining whether there is a closed set of real numbers of the power of the second number class. He could not do this, but he had established the first result of *descriptive set theory*, the definability theory for the continuum.

Almost two decades after his initial 1873 proof, Cantor in a short 1891 note gave his now celebrated *diagonal* argument. He proceeded in terms of functions, ushering in collections of arbitrary functions into mathematics, but we state and prove his result as is done nowadays in terms of the power set $\mathcal{P}(x) = \{y \mid y \subseteq x\}$ of a set x, the collection of all its subsets: *For any set x, $\mathcal{P}(x)$ has a higher power than x.*

First, the function associating each $a \in x$ with $\{a\}$, that subset of x with sole member a, is a one-to-one correspondence between x and a subset of $\mathcal{P}(x)$. Assume now to the contrary that there is a one-to-one correspondence F established between the members of x and all the members of $\mathcal{P}(x)$. Consider the "diagonal" set $d = \{a \mid a \in x$ and $a \notin F(a)\}$ consisting of those members a of x that do not belong to their corresponding subset $F(a)$. If d itself were a value of F, say $d = F(b)$ for some $b \in x$, then we would have the paradigmatic contradiction: $b \in d$ exactly when $b \notin d$. Hence, F was not a one-to-one correspondence after all!

Cantor had been shifting his notion of set to a level of abstraction beyond sets of real numbers and the like, and the casualness of his 1891 note may reflect an underlying cohesion with his earlier 1873 argument. Indeed the diagonal argument can be drawn out of the earlier argument, and the new result generalized the old since, with N the set of natural numbers, $\mathcal{P}(N)$ is in one-to-one correspondence with the continuum. With his new result Cantor affirmed that the powers of well-defined sets have no maximum, and he had proved for the first time that there is a power greater than that of the continuum. However, with his view that every well-defined set is well-ordered Cantor would now have had to confront, in his arbitrary function context, a general difficulty starkly abstracted from the Continuum Problem: *From a well-ordering of a set a well-ordering of its power set is not necessarily definable.* The diagonal proof called into question Cantor's very notion of set.

Cantor's *Beiträge*, published in two parts in 1895 and 1897, presented his mature theory of the transfinite. In the first part Cantor reconstrued power as *cardinal number*, an autonomous concept beyond being *une façon de parler* about one-to-one correspondence. He defined the addition, multiplication, and exponentiation of cardinal numbers primordially in terms of set-theoretic operations and functions. As befits the introduction of new numbers Cantor then introduced a new notation, one using the Hebrew letter aleph, $\aleph$. With $\aleph_0$ the cardinal number of the set of natural numbers Cantor showed that

$$\aleph_0 \cdot \aleph_0 = \aleph_0 \text{ and } 2^{\aleph_0} \text{ is the cardinal number of the continuum}$$

(and hence of $\mathcal{P}(N)$). With this he observed that the 1878 labor of associating the continuum with the plane and so forth could be reduced to a "few strokes of the pen" in his new arithmetic. Cantor only mentioned

$$\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_\alpha, \ldots,$$

these to be the cardinal numbers of the successive number classes from the *Grundlagen* and thus to exhaust all the infinite cardinal numbers.

Cantor then developed his theory of *order types*, "types" or abstractions of linear orderings. He defined the addition and multiplication of order types and characterized the order types of the rational numbers and of the real numbers. In the second *Beiträge* Cantor turned to the special case of well-orderings and reconstrued the transfinite numbers as their order types, newly calling the numbers the *ordinal numbers*. He then established their basic comparability properties by showing that given two well-orderings one is isomorphic to an initial segment of the other or vice versa. In this new setting he concentrated on the countable ordinal numbers, the new construal of the second number class, and provided an incisive structural analysis in terms of a new operation of ordinal exponentiation.

The two parts of the *Beiträge* were not only distinct by subject matter, cardinal number and the continuum vs. ordinal number and well-ordering, but also between them there developed a wide, insurmountable breach. In the first part nowhere is the 1891 result stated even in a special case, though it was now possible to express it as $\mathbf{m} < 2^{\mathbf{m}}$ for any cardinal number **m**, since in his arithmetic

$2^{\mathbf{m}}$ is the cardinal number of the power set of a set with cardinal number **m**.

Also, the second *Beiträge* does not mention any aleph beyond $\aleph_1$, nor does it mention the Continuum Hypothesis, which could have been stated as $2^{\aleph_0} = \aleph_1$. Every well-ordered set, through a corresponding ordinal number, has an aleph as its cardinal number, but how does $2^{\aleph_0}$ fit into the aleph sequence?

Thus the Continuum Problem was embedded in the very interstices of the early development of set theory, and in fact the structures that Cantor built, while now of great intrinsic interest, emerged out of efforts to articulate and solve the Continuum Problem. The tension uncovered by Cantor's diagonal argument between well-ordering and power set (or arbitrary functions) would soon be revisited by Zermelo. David Hilbert, when he presented his famous list of twenty-three problems at the 1900 International Congress of Mathematicians at Paris, made the Continuum Problem the very first problem and intimated Cantor's difficulty by suggesting the desirability of "actually giving" a well-ordering of the real numbers.

At the turn into the twentieth century the "logical" limits of set formation and existence were broached for sets being counterparts to "concepts" or properties. In correspondence with Hilbert and Richard Dedekind in the late 1890s Cantor became newly engaged with questions of set existence. He had earlier considered collections like all ordinal numbers or all alephs as leading out of his conceptual framework. These "absolutely infinite or inconsistent multiplicities," if admitted as sets, would lead to contradictions, and Cantor argued anew that every set can be well-ordered else it would in one-to-one correspondence with all the ordinal numbers and hence an inconsistent multiplicity. In this he anticipated later developments in set theory.

Bertrand Russell, a main architect of the analytic tradition in philosophy, focused in 1900 on Cantor's work. Russell was pivoting from idealism toward logicism, the thesis that mathematics can be founded in logic. Taking a universalist approach to logic with all-encompassing categories, Russell took the class of all classes to have the largest cardinal number but saw that Cantor's 1891 result leading to higher cardinal numbers presented a problem. Analyzing that argument, by the spring of 1901 he arrived at the famous *Russell's Paradox*. This paradox showed with remarkable simplicity that there are properties $P(x)$ such that the collection of objects having that property, the *class*

$$\{x \mid P(x)\},$$

cannot itself be an object: Consider $\{x \mid x \notin x\}$. If this were an object r in the range of possibilities, then we would have the contradiction $r \in r$ exactly when $r \notin r$. This paradox may have been critical for Russell's universalist approach to logic and for logicism, but it was less so for the development of set theory, which was emerging in mathematics. In any case the paradox did serve as a motivation for fashioning a consistent notion of set through axiomatization.

The first decade of the new century saw Zermelo make his major advances in the development of set theory. Already estimable as an applied mathematician, Zermelo turned to set theory and its foundations under the influence of Hilbert. Zermelo's first substantial result was his independent discovery of the argument for Russell's Paradox. He then established in 1904 the Well-Ordering Theorem, that every set can be well-ordered, assuming what he soon called the Axiom of Choice (AC). Zermelo thereby shifted the notion of set away from the implicit assumption of Cantor's principle that every well-defined

set is well-ordered and replaced that principle by an explicit axiom about a wider notion of set.

In retrospect Zermelo's *argument* for his Well-Ordering Theorem can be viewed as pivotal for the development of set theory. To summarize the argument, suppose that x is a set to be well-ordered, and through Zermelo's Axiom-of-Choice hypothesis assume that the power set $\mathcal{P}(x) = \{y \mid y \subseteq x\}$ has a choice function, that is, a function γ such that for every non-empty member y of $\mathcal{P}(x)$, $\gamma(y) \in y$. Call a subset y of x a γ-set if there is a well-ordering R of y such that for each $a \in y$,

$$\gamma(\{z \mid z \notin y \text{ or } z\, R\, a \text{ fails}\}) = a.$$

That is, each member of y is what γ "chooses" from what does not already precede that member according to R. The main observation is that γ-sets cohere in the following sense: If y is a γ-set with well-ordering R and z is a γ-set with well-ordering S, then $y \subseteq z$ and S is a prolongation of R, or vice versa. With this, let w be the union of all the γ-sets, that is, all the γ-sets put together. Then w too is a γ-set, and by its maximality it must be all of x and hence x is well-ordered.

Note that the converse to this result is immediate in that if x is well-ordered, say with a well-ordering $\prec$, then the power set $\mathcal{P}(x)$ has a choice function δ, namely for each non-empty member y of $\mathcal{P}(x)$, let $\delta(y)$ be the the $\prec$-least member of y. Not only did Zermelo's argument analyze the connection between well-ordering and choice functions, but it anticipated in its defining of approximations and taking of a union the proof procedure for von Neumann's Transfinite Recursion Theorem.

Zermelo maintained that the Axiom of Choice, to the effect that *every* set has a choice function, is a "logical principle" which "is applied without hesitation everywhere in mathematical deduction," and this is reflected in the Well-Ordering Theorem being regarded as a theorem. Cantor's work had served to exacerbate a growing discord among mathematicians with respect to two related issues: whether infinite collections can be mathematically investigated at all, and how far the function concept is to be extended. The positive use of an arbitrary function operating on arbitrary subsets of a set having been made explicit, there was open controversy after the appearance of Zermelo's proof. This can be viewed as a turning point for mathematics, with the subsequent tilting toward the acceptance of the Axiom of Choice symptomatic of a conceptual shift in mathematics.

AXIOMATIZATION

In response to his critics Zermelo published a second proof of the Well-Ordering Theorem in 1908, and with axiomatization assuming a general methodological role in mathematics he also published in 1908 the first full-fledged axiomatization of set theory. But as with Cantor's work, this was no idle structure building but a response to pressure for a new mathematical context. In this case it was not for the formulation and solution of a *problem* like the Continuum Problem, but rather to clarify a *proof*. Zermelo's motive in large part for axiomatizing set theory was to buttress his Well-Ordering Theorem by making explicit its underlying set existence assumptions. Effecting the first transmutation of the notion of set after Cantor, Zermelo ushered in a new abstract, prescriptive view of sets as solely structured by membership and governed by axioms, a view that would soon come to dominate.

The following are Zermelo's axioms, much as they would be presented today. They are to govern the connections between $\in$ and $\subseteq$ and to prescribe the generation of new sets out of old. The standard axiomatization would be the result of adding two further axioms and formalizing in first-order logic.

AXIOM OF EXTENSIONALITY. *Two sets are equal exactly when they have the same members.* Thus sets epitomize the *extensional* view of mathematics, it being stipulated that however sets are arrived at, there is a definite criterion for equality provided solely by membership.

AXIOM OF EMPTY SET. *There is a set having no members.* This axiom serves to emphasize the beginning with an initial set, the *empty* set, denoted $\emptyset$.

AXIOM OF PAIRS. *For any sets x and y, there is a set consisting of exactly x and y as members.* The posited set is denoted $\{x,y\}$ and is called the (unordered) pair of x and y. $\{x,x\}$ is denoted $\{x\}$, as we have already seen, and is called the *singleton* of x.

AXIOM OF UNION. *For any set x, there is a set consisting exactly of those sets that are members of some member of x.* The posited set is denoted $\cup x$ and is called the *union* of x. This "generalized" union subsumes the better known binary union, in that for any sets a and b,

$$a \cup b = \cup\, \{a,b\} = \{x \mid x \in a \text{ or } x \in b\}.$$

If a set x is structured as an indexed set $\{x_i \mid i \in I\}$, then $\cup\, x$ is often written as $\cup_{i \in I} x_i$ or just $\cup_i x_i$.

AXIOM OF POWER SET. For any set x, there is a set consisting exactly of the subsets of x. The posited set is denoted $\mathcal{P}(x)$ and is called the *power set* of x, as we have already seen.

AXIOM OF CHOICE. *For any set x consisting of non-empty, pairwise disjoint sets, there is a set c such that every member of x has exactly one element in c*. Thus, c acts like a choice function for x construed as a family of sets. This is a reductive way of positing choice functions.

AXIOM OF INFINITY. There is a set having $\emptyset$ as a member and such that whenever y is a member, so also is $y \cup \{y\}$. This has become the usual way of positing the existence of an infinite set, in light of the definition of ordinals. Zermelo actually stated his axiom with "$y \cup \{y\}$" replaced by "$\{y\}$," getting at a set describable informally as $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$.

AXIOM OF SEPARATION. *For any set x and definite property P, there is a set consisting exactly of those members of x having the property P*. Once a collection has been comprehended as a set, we are able to form a subset by "separating" out according to a property. Or, a subclass of a set is a set. Taking the property of being a member of a given set a, we have as a set the binary intersection

$$x \cap a = \{y \mid y \in x \text{ and } y \in a\}.$$

Taking the property of not being a member of a, we have as a set the set-theoretic difference

$$x - a = \{y \mid y \in x \text{ and } y \notin a\}.$$

As a further use of the axiom, consider for a set x the *intersection* of x:

$$\cap x = \{a \mid a \in y \text{ for every } y \in x\}.$$

This is a (property-specifiable) subclass of any member of x, and so we have as a theorem: *If $x \neq \emptyset$ then $\cap x$ is a set.* This is a "generalized" intersection, with the better known binary intersection being $\cap \{x, a\} = x \cap a$.

According to Zermelo a property is "*definite* if the fundamental relations of the domain, by means of the axioms and the universally valid laws of logic, determine without arbitrariness whether it holds or not." But with no underlying logic formalized, the ambiguity of definite property would become a major issue, one that would eventually be resolved only decades later through first-order formalization. In any case Zermelo saw that the Separation idea suffices for a development of set theory that still allows for the "logical" formation of sets according to property. Russell's Paradox is forestalled since only "logical" subsets are to be allowed; indeed, Zermelo's first theorem was that there is no universal set, a set that contains every set as a member, the *reductio* argument being the paradox argument.

Stepping back, Extensionality, Empty Set, and Pairs served to lay the basis for sets. Infinity and Power Set ensured sufficiently rich settings for set-theoretic constructions. Tempering the logicians' extravagant and problematic "all," Power Set provided the provenance for "all" for subsets of a given set, just as Separation served to capture "all" for elements of a given set satisfying a property. Finally, Union and Choice completed the encasing of Zermelo's proof(s) of his Well-Ordering Theorem in the necessary set existence principles.

Although Hilbert's axiomatization of geometry in his 1899 *Grundlagen der Geometrie* may have served as a model for Zermelo's axiomatization of set theory and Dedekind's 1888 essay *Was sind und was sollen die Zahlen?* on the foundations of arithmetic a precursor, there are crucial differences having to do with subject matter and proof. Both in intent and outcome Dedekind and Hilbert had been engaged in the analysis of fixed subject matter. Dedekind in particular had done a great deal to enshrine proof as the vehicle for algebraic abstraction and generalization. Like algebraic constructs, sets were new to mathematics and would be incorporated by setting down rules for their proofs. Just as Euclid's axioms for geometry had set out the permissible geometric constructions, the axioms of set theory would set out rules for set generation and manipulation. But unlike the emergence of mathematics from marketplace arithmetic and Greek geometry, sets and transfinite numbers were neither laden with nor bolstered by substantial antecedents. There was no fixed, intended subject matter. Like strangers in a strange land stalwarts developed a familiarity with sets guided step by step by the axiomatic framework. For Dedekind it had sufficed to work with sets by merely giving a few definitions and properties, those foreshadowing Extensionality, Union, and Infinity. Zermelo provided more rules: Separation, Power Set, and Choice.

Zermelo's 1908 axiomatization paper, especially with its rendition at the end of the Cantorian theory of cardinality in terms of functions cast as set constructs, brought out Zermelo's *set-theoretic reductionism*. Zermelo pioneered the reduction of mathematical concepts and arguments to set-theoretic concepts and arguments from axioms, based on sets doing the work of mathematical objects. Set theory would provide the underpinnings of mathematics, and Zermelo's axioms would resonate with

emerging mathematical practice. Zermelo's analysis moreover served to draw what would come to be generally regarded as set-theoretic out of the realm of the presumptively logical. This would be particularly salient for Infinity and Power Set and was strategically advanced by the segregation of property considerations to Separation. Based on generative and prescriptive axioms, set theory would become more combinatorial, less logical. With these features Zermelo's axioms indeed proved more than adequate to serve as a reductive basis for mathematics, at least for providing surrogates for mathematical objects; looking ahead it was for subsequent developments to bring out that set theory could also serve as a court of adjudication in terms of relative consistency.

Felix Hausdorff was the first developer of the transfinite after Cantor, the one whose work first suggested the rich possibilities for a mathematical investigation of the higher transfinite. A mathematician *par excellence* Hausdorff took the sort of mathematical approach to set theory and set-theoretic approach to mathematics which would come to dominate in the years to come. In a 1908 publication Hausdorff brought together his extensive work on *uncountable* order types, and in particular formulated the *Generalized Continuum Hypothesis* (GCH): For any infinite set x, there is no set of cardinal number strictly intervening between that of x and of its power set $\mathcal{P}(x)$; or in Cantor's later terms, for every ordinal number α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$. Hausdorff also entertained for the first time a "large cardinal" concept, of which more below. Hausdorff's classic 1914 text, *Grundzüge der Mengenlehre*, broke the ground for a generation of mathematicians in both set theory and topology. He presented Cantor's and Zermelo's work systematically, and of particular interest, he applied the Axiom of Choice to provide what is now known as Hausdorff's Paradox. The source of the later and better known Banach-Tarski Paradox, Hausdorff's Paradox provided an implausible decomposition of the sphere and was the first, and a dramatic, synthesis of classical mathematics and the new Zermelian abstract view.

In the *Grundzüge* Hausdorff defined an ordered pair of sets in terms of (unordered) pairs, formulated functions in terms of ordered pairs, and ordering relations as collections of ordered pairs. Hausdorff thus capped efforts of logicians by making their moves in mathematics, completing the set-theoretic reduction of relations and functions. In the modern setting, the definition of the ordered pair that has been adopted is not Hausdorff's, but one provided by Kazimierz Kuratowski in 1921:

$$\langle x,y\rangle = \{\{x\}, \{x, y\}\}.$$

This satisfies all that is operationally required of an ordered pair:

$$\langle x,y\rangle = \langle a,b\rangle \text{ exactly when } x = a \text{ and } y = b.$$

With this definition, a set r is a *relation* if it consists of ordered pairs. This objectification is often eased by reverting to the older conceptual notation $a\ r\ b$ for $\langle a,b\rangle \in r$. A set f is a *function* if it is a relation satisfying: If $\langle x,y\rangle \in f$ and $\langle x,z\rangle \in f$, then $y = z$. This objectification is eased by reverting to the older operational notation $f(x) = y$ for $\langle x, y\rangle \in f$, though the emphasis is on the generality and arbitrariness of f as just a relation with a univalency property. Finally the dynamic notation $f: a \to b$ specifies that f is a function such that every member of a is a first coordinate of an ordered pair in f, and that every second coordinate is a member of b.

AXIOMATIZATION COMPLETED

In the 1920s fresh initiatives structured the loose Zermelian framework with new features and corresponding developments in axiomatics, the most consequential moves made by John von Neumann with anticipations by Dimitry Mirimanoff in a pre-axiomatic setting. Von Neumann effected a Counter-Reformation of sorts that led to the incorporation of a new axiom, the Axiom of Replacement: The transfinite numbers had been central for Cantor but peripheral to Zermelo; von Neumann reconstrued them as *bona fide* sets, the ordinals, and established their efficacy by formalizing transfinite recursion, the method of sequential definition of sets based on previously defined sets applied with transfinite indexing.

Ordinals manifest the basic idea of taking precedence in a well-ordering simply to be membership:

> *Definitions*. A set x is *transitive* if $\cup x \subseteq x$, that is, whenever $a \in b$ and $b \in x$, then $a \in x$.
>
> A set x is a (von Neumann) *ordinal* if x is transitive and the membership relation restricted to $x = \{y \mid y \in x\}$ is a well-ordering of x.

For example, $\emptyset$ is transitive, but $\{\{\emptyset\}\}$ is not. Loosely speaking, transitive sets retain all their hereditary members. The first several ordinals are

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

and are newly taken to be the numbers 0, 1, 2, 3, … . If x is an ordinal, then so also is $x \cup \{x\}$, the *successor* of x, and this accounts for how the Axiom of Infinity was formulated in the previous section. It has become customary to use the Greek letters α, β, γ, … to denote ordinals. Von

Neumann, as had Mirimanoff before him, established the key instrumental property of Cantor's ordinal numbers for ordinals: *Every well-ordered set is order-isomorphic to exactly one ordinal with membership*. The proof made a paradigmatic use of Replacement, and so was the first proof to draw that axiom into set theory.

For a set x and property $P(v, w)$, the property is said to be *functional on x* if for any $a \in x$, there is exactly one b such that $P(a, b)$.

AXIOM OF REPLACEMENT. For any set x and property $P(v, w)$ functional on x, $\{b \mid P(a, b)$ for some $a \in x\}$ is a set.

This axiom allows for new sets that result when members of a set are "replaced" according to a property. If the functional property is given by a set, that is there is a function f, a *set* of ordered pairs, such that $P(v, w)$ exactly when $f(v) = w$, then Replacement is not needed. However, as in the case of the above-stated result correlating arbitrary well-orderings with ordinals, there are functional properties that are more general, typically formulated by recursion.

Replacement subsumes Separation. Suppose that x is a set and P is a (definite) property. If there are no members of x satisfying P, then we are done. Otherwise, fix such a member y_0. For any $a \in x$, let $P(a, a)$ hold if a satisfies P and $P(a, y_0)$ hold otherwise. Then the "replaced" set $\{b \mid P(a, b)$ for some $a \in x\}$ is the set of members of x satisfying P.

Von Neumann took the crucial step of ascribing to the ordinals the role of Cantor's ordinal numbers with their several principles of generation. Now, with ordinal numbers regarded as gauging well-orderings, that one is isomorphic to a proper initial part of another corresponds for ordinals to actual membership and can be rendered

$$\alpha < \beta \text{ exactly when } \alpha \in \beta.$$

For this reconstrual of ordinal numbers and already to define the arithmetic of ordinals von Neumann saw the need to establish the Transfinite Recursion Theorem, the theorem that validates definitions by recursion along well-orderings. The proof was anticipated by the Zermelo 1904 proof, but Replacement was necessary even for the very formulation, let alone the proof, of the theorem. With the ordinals in place von Neumann completed the restoration of the Cantorian transfinite by defining the *cardinals* as the *initial ordinals*, those ordinals not in one-to-one correspondence with any of its predecessors. The infinite initial ordinals are denoted

$$\omega = \omega_0, \omega_1, \omega_2, \dots, \omega_\alpha, \dots,$$

so that ω is to be the set of natural numbers in the ordinal construal, and the identification of different intensions is signaled by

$$\omega_\alpha = \aleph_\alpha$$

with the left being a von Neumann ordinal and the right being the Cantorian cardinal number. Every set x, with AC, is well-orderable and hence in one-to-one correspondence with an initial ordinal ω_α, and the *cardinality* of x is $|x| = \aleph_\alpha$. It has become customary to use the middle Greek letters $\kappa, \lambda, \mu, \dots$ to denote initial ordinals in their role as the cardinals. A *successor* cardinal is one of form $\aleph_{\alpha+1}$ and is denoted κ^+ for $\kappa = \aleph_\alpha$. A cardinal which is not a successor cardinal is a *limit* cardinal.

Replacement has been latterly regarded as somehow less necessary or crucial than the other axioms, the purported effect of the axiom being only on large-cardinality sets. Initially Abraham Fraenkel and Thoralf Skolem had independently in 1922 proposed the addition of Replacement to Zermelo's axioms, both pointing out the inadequacy of Zermelo's axioms for establishing that $E = \{Z_0, \mathcal{P}(Z_0), \mathcal{P}(\mathcal{P}(Z_0)), \dots\}$ is a set, where $Z_0 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ is Zermelo's infinite set from his Axiom of Infinity. However even $F = \{\emptyset, \mathcal{P}(\emptyset), \mathcal{P}(\mathcal{P}(\emptyset)), \dots\}$ cannot be proved to be a set from Zermelo's axioms: The union of E above, with membership restricted to it, models Zermelo's axioms yet does not have F as a member. Hence Zermelo's axioms cannot establish the existence of some simple countable sets consisting of finite sets and could be viewed as remarkably lacking in closure under finite recursive processes. If the Axiom of Infinity were itself modified to entail that F is a set, then there would still be many other finite sets a so that $\{a, \mathcal{P}(a), \mathcal{P}(\mathcal{P}(a)), \dots\}$ cannot be proved to be a set. Replacement serves to rectify the situation by allowing new infinite sets defined by "replacing" members of the one infinite set given by the Axiom of Infinity. In any case the full exercise of Replacement is part and parcel of transfinite recursion, and it was von Neumann's formal incorporation of this method into set theory, as necessitated by his proofs, that brought in Replacement.

Von Neumann (and before him Mirimanoff, Fraenkel, and Skolem) also considered the salutary effects of restricting the universe of sets to the well-founded sets. The *well-founded* sets are the sets that belong to some "rank" V_α, these definable through transfinite recursion:

$$V_0 = \emptyset;\ V_{\alpha+1} = \mathcal{P}(V_\alpha);\ \text{and } V_\delta = \bigcup\{V_\alpha \mid \alpha < \delta\} \text{ for limit ordinals } \delta.$$

V_ω consists of the "hereditarily finite" sets, $\omega \in V_{\omega+1}$, and $\mathcal{P}(\omega) \in V_{\omega+2}$, and so already in these beginning levels there are set counterparts for many objects in mathematics. That the universe V of all sets is the *cumulative hierarchy*

$$V = \cup\{V_\alpha \mid \alpha \text{ is an ordinal}\}.$$

is thus the assertion that every set is well-founded. Von Neumann essentially showed that this assertion is equivalent to a simple assertion about sets:

AXIOM OF FOUNDATION. $\forall x(x \neq \emptyset \longrightarrow \exists y \in x\,(x \cap y = \emptyset))$.

Thus non-empty well-founded sets have $\in$-minimal members. If a set x satisfies $x \in x$ then $\{x\}$ is not well-founded; similarly if there are $x_1 \in x_2 \in x_1$, then $\{x_1, x_2\}$ is not well-founded. Ordinals and sets consisting of ordinals are well-founded, and well-foundedness can be viewed is a generalization of being an ordinal that loosens the connection with transitivity. The Axiom of Foundation eliminates pathologies like $x \in x$ and through the cumulative hierarchy rendition provides metaphors about building up the universe of sets and the possibility of inductive arguments to establish results about all sets.

In a remarkable 1930 publication Zermelo offered his final axiomatization of set theory as well as a striking, synthetic view of a procession of models that would have a modern resonance. Proceeding in what we would now call a second-order context, Zermelo extended his 1908 axiomatization by adjoining both Replacement and Foundation. The standard axiomatization of set theory

ZFC, *Zermelo-Fraenkel with Choice*,

is recognizable, the main difference being that ZFC is a first-order theory (see the next section); "Fraenkel" acknowledges Fraenkel's suggestion of adjoining Replacement; and the Axiom of Choice is explicitly mentioned.

ZF, *Zermelo-Fraenkel*,

is ZFC without AC and is a base theory for the investigation of weak Choice-type propositions as well as propositions that contradict AC.

Zermelo herewith completed his transmutation of the notion of set, his abstract, prescriptive view stabilized by further axioms that structured the universe of sets. Replacement and Foundation focused the notion of set, with the first providing the means for transfinite recursion and induction, and the second making possible the application of those means to get results about *all* sets. It is nowadays almost banal that Foundation is the one axiom unnecessary for the recasting of mathematics in set-theoretic terms, but the axiom is also the salient feature that distinguishes investigations specific to set theory as an autonomous field of mathematics. Indeed it can be fairly said that modern set theory is at base a study couched in well-foundedness, the Cantorian well-ordering doctrines adapted to the Zermelian generative and prescriptive conception of sets. With Replacement and Foundation in place, Zermelo was able to provide natural models of his axioms and to establish algebraic isomorphism, initial segment, and embedding results for his models. Finally Zermelo posited an endless procession of his models, each a set in the next, as natural extensions of their cumulative hierarchies.

Zermelo found a simple set-theoretic condition, being an inaccessible cardinal, that characterizes the ordinal heights of his models, that is those ordinals ρ such that the predecessors of ρ are exactly the ordinals of a model.

Definitions. An infinite cardinal κ is *singular* if there is an $x \subseteq \kappa$ of smaller cardinality than κ which is cofinal in κ, that is to say for any $\alpha < \kappa$ there is a $\beta \in x$ with $\alpha \leq \beta$. An infinite cardinal which is not singular is *regular*

An infinite cardinal κ is a *strong limit* if for any cardinal $\beta < \kappa$, $2^\beta < \kappa$.

An infinite cardinal κ is *inaccessible* if it is both regular and a strong limit.

$\aleph_0$ is regular; $\aleph_1$, $\aleph_2$, … and generally, all successor cardinals are regular. The limit cardinal $\aleph_\omega$ is singular, since it has a countable cofinal subset $\{\aleph_0, \aleph_1, \aleph_2, \ldots\}$. Hausdorff in 1908 had initially entertained the possibility of having a regular limit cardinal. Inaccessible cardinals had later been considered to be a stronger version that arithmetically incorporated power sets, but Zermelo provided the first structural rationale for them, as the delimiters of his natural models.

Inaccessible cardinals are the modest beginnings of the theory of *large cardinals*, a mainstream of modern set theory devoted to the investigation of strong hypotheses and consistency strength. Large cardinal hypotheses posit structure in the higher reaches of the cumulative hierarchy, most often by positing cardinals that prescribe their own inaccessible transcendence over smaller cardinals, and were seen by the 1970s to form a natural hierarchy of stronger and stronger propositions transcending ZFC.

The journal volume containing Zermelo's 1930 publication also contained Stanisław Ulam's seminal paper on *measurable cardinals*, which became the most pivotal

of all large cardinals. For a set s, U is a (non-principal) *ultrafilter over s* if U is a collection of subsets of s containing no singletons; if $x \in U$ and $x \subseteq y \subseteq s$, then $y \in U$; if $x \in U$ and $y \in U$, the $x \cap y \in U$; and for any $x \subseteq s$, either $x \in U$ or $s - x \in U$. For a cardinal λ, an ultrafilter U is *λ-complete* if for any $D \subseteq U$ of cardinality less than λ, $\bigcap D \in U$. Finally an uncountable cardinal κ is *measurable* if there is a κ-complete ultrafilter over κ. Thus, a measurable cardinal is a cardinal whose power set is structured with a two-valued "measure" having a strong closure property. Measurability embodied the first large cardinal confluence of Cantor's two legacies, the investigation of definable sets of reals and the extension of number into the transfinite: The concept was distilled from measure-theoretic considerations related to Lebesgue's measure for sets of real numbers, and it also entailed inaccessibility in the transfinite.

FORMALIZATION AND MODEL-THEORETIC METHODS

Zermelo's 1930 publication was in part a response to Skolem's 1922 advocacy of the idea of framing Zermelo's 1908 axioms in first-order logic. First-order logic investigates the logic of formal languages consisting of formulas built up from specified function and predicate symbols using logical connectives and first-order quantifiers $\forall$ and $\exists$, these interpreted as ranging over the *elements* of a domain of discourse. (Second-order logic has quantifiers interpreted as ranging over properties, or collections of elements.) First-order logic had emerged in the 1917 lectures of Hilbert as a delimited system of logic potentially amenable to mathematical analysis. Entering from a different, algebraic tradition Skolem had established a seminal result for "metamathematical" methods with the Löwenheim-Skolem Theorem: *If a countable collection of first-order sentences has a model then it has a countable model.*

For set theory Skolem proposed formalizing Zermelo's axioms in the first-order language with $\in$ and $=$ as binary predicate symbols. Zermelo's *definite* properties were to be those expressible in this first-order language in terms of given sets, and the Axiom of Separation was to become a schema of axioms, one for each first-order formula that has variables allowing for set parameters. As a palliative for taking set theory as a foundation for mathematics, Skolem then pointed out what has come to be called the *Skolem Paradox*: Zermelo's 1908 axioms cast in first-order logic is a countable collection of sentences, and so if they have a model at all, they have a countable model. (Analogous remarks apply to the latterly adjoined Axiom of Replacement becoming a schema.) Thus we have the paradoxical existence of countable models for Zermelo's axioms although they entail the existence of uncountable sets. Zermelo found this antithetical and repugnant. However stronger currents were at work, leading to a further, subtler transmutation of the notion of set mediated by first-order logic and incorporating its relativism of set-theoretic concepts.

Gödel virtually completed the mathematization of logic by submerging metamathematical methods into mathematics. The main vehicle was the direct coding, "the arithmetization of syntax," in his celebrated 1931 Incompleteness Theorem, which worked dialectically against a program of Hilbert's for establishing the consistency of mathematics. But starting an undercurrent, the earlier 1930 Completeness Theorem for first-order logic clarified the distinction between the formal syntax and semantics (interpretations) of first-order logic, and secured its key instrumental property with the Compactness Theorem: *If a collection of first-order sentences is such that every finite subcollection has a model, then the whole collection has a model.*

Gödel's work showed that the notion of the consistency of a mathematical theory has a formal counterpart expressible in the first-order language with function symbols for addition and multiplication. Loosely speaking, a *theory* is a collection of sentences of some first-order language; that a sequence of formulas constitutes a *deduction* can be formalized; and a theory is *consistent* if from it no contradiction can be derived. Gödel's arithmetization of syntax codes all this into statements about the natural numbers and their arithmetic, yielding a formula

$$\mathrm{Con}(T)$$

asserting the formal consistency of T, at least for those theories whose sentences can be schematically defined. Gödel famously established through his Incompleteness Theorem that for consistent theories subsuming the arithmetic of the natural numbers, $\mathrm{Con}(T)$ itself cannot be deduced from T. However, one may be able to deduce relative notions:

Definitions. A sentence σ is *relatively consistent* with a theory T if $\mathrm{Con}(T)$ implies $\mathrm{Con}(T + \sigma)$.

A sentence σ is *independent* of a theory T if both σ and its negation are relatively consistent with T.

Two sentences σ_1 and σ_2 are *equi-consistent* over a theory T if $\mathrm{Con}(T + \sigma_1)$ is equivalent to $\mathrm{Con}(T + \sigma_2)$.

These assertions would be established over a weak base theory. For example, in the parlance, that a set-theoretic statement σ is relatively consistent with set theory generally means that Con(ZFC) implies Con(ZFC + σ), this itself deducible in (some weak version of) ZFC. *Consistency strength* in set theory can be discussed in these terms, typically for strong set theoretic statements not provable from ZFC: For two set-theoretic statements σ_1 and σ_2, the consistency strength of σ_1 is least that of σ_2 if Con(ZFC + σ_1) implies Con(ZFC + σ_2), and so σ_1 and σ_2 have equal consistency strength if σ_1 and σ_2 are equi-consistent over ZFC.

Tarski in the early 1930s completed the mathematization of logic by providing his "definition of truth," exercising philosophers to a surprising extent ever since. Tarski simply schematized truth as a correspondence between formulas of a formal language and set-theoretic assertions about an interpretation of the language and provided a recursive definition of the *satisfaction* relation, when a formula holds in an interpretation, in set-theoretic terms. This response to a growing need for a mathematical framework became the basis for *model theory*. The eventual effect of Tarski's mathematical formulation of semantics would be not only to make mathematics out of the informal notion of satisfiability, but also to enrich ongoing mathematics with a systematic method for forming mathematical analogues of several intuitive semantic notions. For coming purposes, the following specifies notation and concepts in connection with Tarski's definition:

Definitions. For a first-order language, an interpretation N of that language (i.e., a specification of a domain of discourse as well as interpretations of the function and predicate symbols), a formula $\varphi(v_1, v_2, \ldots, v_n)$ of the language with the variables as displayed, and $a_1, a_2, \ldots, a_n$ in the domain of N,

$$N \vDash \varphi[a_1, a_2, \ldots, a_n]$$

asserts that the formula φ is satisfied in N according to Tarski's recursive definition when v_i is interpreted as a_i.

A subset y of the domain of N is *first-order definable over* N if there is a formula $\psi(v_0, v_1, v_2, \ldots, v_n)$ and $a_1, a_2, \ldots, a_n$ in the domain of N such that

$$y = \{z \mid N \vDash \psi[z, a_1, \ldots, a_n]\}.$$

Set theory was launched on an independent course as a distinctive field of mathematics by Gödel's formulation of the model L of "constructible" sets, with which he established the relative consistency of the Axiom of Choice (AC) and the Generalized Continuum Hypothesis (GCH). L is a transitive class containing all the ordinals that, with the membership relation restricted to it, satisfies each axiom of ZFC as well as GCH. Through L Gödel established that Con(ZF) implies Con(ZFC + GCH) and thus attended to fundamental issues at the beginnings of set theory. In his first, 1938 announcement Gödel described L as a hierarchy "which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite orders." Indeed with L Gödel had refined the cumulative hierarchy of sets to a cumulative hierarchy of *definable* sets which is analogous to the orders of Russell's *ramified* theory. Gödel's further innovation was to continue the indexing of the hierarchy through *all* the (von Neumann) ordinals to get a model of set theory. In a 1939 note Gödel presented L essentially as it is presented today: For any set x let def(x) denote the collection of subsets of x first-order definable over x according to the previous definition. Then define:

$$L_0 = \emptyset;\ L_{\alpha+1} = \text{def}(L_\alpha),\ L_\delta = \bigcup\{L_\alpha \mid \alpha < \delta\}$$
$$\text{for limit ordinals } \delta;$$

and the *constructible universe*

$$L = \bigcup\{L_\alpha \mid \alpha \text{ is an ordinal}\}.$$

Gödel brought into set theory a method of construction and argument and thereby affirmed several features of its axiomatic presentation. First Gödel showed that def(x) and generally first-order definability over set domains is itself definable in set theory, so that in particular the definition of L can be effected in set theory via transfinite recursion. This significantly contributed to a lasting ascendancy for first-order logic which beyond its *sufficiency* as a logical framework for mathematics was seen to have considerable *operational efficacy*. Gödel's construction moreover buttressed the incorporation of Replacement and Foundation into set theory. Replacement was immanent in the arbitrary extent of the ordinals for the indexing of L and in its formal definition via transfinite recursion. As for Foundation, underlying the construction was the well-foundedness of sets, and significantly, Gödel viewed L as deriving its contextual sense from the cumulative hierarchy of sets regarded as an extension of the *simple* theory of types. In footnote 12 of his 1939 note he wrote, "In order to give A [that $V = L$] an intuitive meaning, one has to understand by 'sets' all objects obtained by building up the simplified hierarchy of types on an empty set of individuals (including types of arbitrary transfinite orders)." Some have been puzzled about how the cumulative hierarchy picture emerged in set-theoretic practice; although there was Mirimanoff,

von Neumann, and especially Zermelo, the picture came in with Gödel's method, the reasons being both thematic and historical: Gödel's work with L with its incisive analysis of first-order definability was readily recognized as a signal advance, while Zermelo (1930) with its second-order vagaries remained somewhat obscure. As the construction of L was gradually digested, the sense that it promoted of a cumulative hierarchy reverberated to become the basic picture of the universe of sets.

In a notable inversion, what has come to be regarded as the *iterative conception*, the conception of sets as being built up through stages of construction as schematized by the cumulative hierarchy, has become a heuristic for motivating the axioms of set theory generally. This has opened the door to a metaphysical appropriation in the following sense: It is as if there is some notion of set that is "there," in terms of which the axioms must find some further justification. But set theory has no particular obligations to mirror some prior notion of set, especially one like the iterative conception, arrived at *a posteriori*. Replacement and Choice for example do not quite "fit" the iterative conception, but if need be, Replacement can be "justified" in terms of achieving algebraic closure of the axioms, a strong motivation in the work of Fraenkel and the later Zermelo, and Choice can be "justified" as a logical principle as Zermelo had maintained.

Gödel's proof of the GCH in L, like Zermelo's proof of the Well-Ordering Theorem, was synthetic and pivotal for the development of set theory. Gödel actually established that if λ is an infinite cardinal and $x \in L_\lambda$, then for any $y \subseteq x$ in L, $y \in L_\lambda$. The Power Set Axiom was thus tamed in L leading to the relative consistency of GCH. Replacement played a crucial role not only by providing for the prior extent of ordinals, but also in allowing this first instance of model-theoretic reflection. Reflection properties, which in one form came to be seen as equivalent to Replacement, assert that various properties holding at one level of the cumulative hierarchy holds at an earlier level, and they have been a leading heuristic for motivating large cardinals. Gödel's proof also made a specific, positive use of the Skolem Paradox argument, as he used what are now known as *Skolem functions* to take a *Skolem hull*. Paradox became method, affirming the operational efficacy of first-order logic. Finally Gödel took for the first time what is now known as the *transitive collapse*. Andrzej Mostowski would later state in general terms the result, which is a generalization to well-founded relations and transitive sets of the Mirimanoff–von Neumann result, that every well-ordered set is order-isomorphic to exactly one ordinal with membership. While that result was basic to the analysis of well-orderings, the transitive collapse result grew in significance from specific applications and came to epitomize how well-foundedness made possible a coherent theory of models of set theory.

In all these ways Gödel's work promoted a further transmutation of, or at least a new relativism about, the notion of set as mediated by first-order logic. By the 1950s ZFC was generally taken to be a theory formalized in first-order logic. The relativism of set-theoretic concepts was brought to the fore, as well as new possibilities for constructions of models of set theory. Results even about definable sets of real numbers would turn on contingencies of relative consistency. Notably, Gödel himself held a "Platonistic" conception of set theory as descriptive of an objective universe schematized by the cumulative hierarchy; nonetheless, his work laid the groundwork for the development of a range of models and axioms for set theory.

Gödel's work with L stood as an isolated monument for quite a number of years, World War II no doubt having a negative effect on mathematical progress. On the crest of a new generation Dana Scott established a result in 1961 that would become seminal for the theory of large cardinals. Utrafilters gained prominence in model theory in the late 1950s because of the emergence of the *ultrapower* and more generally *ultraproduct* construction for building concrete models, when Scott made the crucial move of taking the ultrapower of the universe V itself by an ultrafilter as provided by a measurable cardinal. Such an ultrafilter provided well-founded ultrapowers, and the full exercise of the transitive collapse now led to an inner model M and an elementary embedding $j\colon V \to M$.

Definitions. M is an *inner model* if it is a transitive class containing all the ordinals that, with the membership relation restricted to it, satisfies each axiom of ZF.

A class function $j\colon V \to M$ from the universe V of sets into an inner model M is an *elementary embedding* if for any set-theoretic formula $\varphi(v_1, v_2, \dots, v_n)$ and sets $a_1, a_2, \dots, a_n$,

$$V \vDash \varphi[a_1, a_2, \dots, a_n] \text{ exactly when } M \vDash \varphi[j(a_1), j(a_2), \dots, j(a_n)].$$

(This suggests the general notion of elementary embedding in model theory; the notion cannot be formalized for V in ZFC, but sufficient schematic approximations can. Below, elementary embeddings are assumed not to be the identity function.) L is the paradigmatic inner model. Appealing to its definability Scott established: *If there is a measurable cardinal, then $V \neq L$.* Large cardinal

hypotheses thus assumed a new significance through a new proof construction, as a means for maximizing possibilities away from Gödel's delimitative universe. The ultrapower construction provided one direction and H. Jerome Keisler soon provided the other of a new characterization that established a central structural role for measurable cardinals: *There is an elementary embedding j: V → M for some inner model M exactly when there is a measurable cardinal.* Through model-theoretic methods set theory was brought to the point of entertaining elementary embeddings into well-founded models, soon to be transfigured by a new method for getting well-founded *extensions* of well-founded models.

FORCING

In 1963 Paul Cohen established the independence of the Axiom of Choice from ZF and the independence of the Continuum Hypothesis from ZFC. That is, complementing Gödel's relative consistency results with *L* Cohen established that Con(ZF) implies Con(ZF + the negation of AC) and that Con(ZFC) implies Con(ZFC + the negation of CH). These results delimited ZF and ZFC in terms of the two fundamental issues at the beginnings of set theory. But beyond that, Cohen's proofs were soon to flow into method, becoming the inaugural examples of *forcing*, a remarkably general and flexible method for extending models of set theory. Forcing has strong intuitive underpinnings and reinforces the notion of set as given by the first-order ZF axioms with conspicuous uses of Replacement and Foundation. If Gödel's construction of *L* had launched set theory as a distinctive field of mathematics, then Cohen's method of forcing began its transformation into a modern, sophisticated one. Cohen's particular achievement lies in devising a concrete procedure for extending well-founded models of set theory in a minimal fashion to well-founded models of set theory with new properties but without altering the ordinals. Set theory had undergone a sea-change, and beyond simply how the subject was enriched, it is difficult to convey the strangeness of it.

Cohen's approach was to start with a model *M* of ZF and adjoin a set G, one that would exhibit some desired new property. He realized that this had to be done in a minimal fashion in order that the resulting structure also model ZF, and so imposed restrictive conditions on both *M* and *G*. He took *M* to be a countable standard model, that is a countable transitive set that together with the membership relation restricted to it is a model of ZF. (The existence of such a model is an avoidable assumption in formal relative consistency proofs via forcing.) The ordinals of *M* would then coincide with the predecessors of some ordinal ρ, and *M* would be the cumulative hierarchy $M = \cup_{\alpha<\rho} (V_\alpha \cap M)$.

Cohen then established a system of terms to denote members of the new model, finding it convenient to use a ramified language: For each $x \in M$ let $\dot{x}$ be a corresponding constant; let $\dot{G}$ be a new constant; and for each $\alpha < \rho$ introduce quantifiers $\forall_\alpha$ and $\exists_\alpha$. Then develop a hierarchy of terms as follows: $\dot{M}_0 = \{\dot{G}\}$, and for limit ordinals $\delta < \rho$, $\dot{M}_\delta = \cup_{\alpha<\delta}\dot{M}_\alpha$. At the successor stage, let $\dot{M}_{\alpha+1}$ be the collection of terms $\dot{x}$ for $x \in V_\alpha \cap M$ and "abstraction" terms corresponding to formulas allowing parameters from $\dot{M}_\alpha$ and quantifiers $\forall_\alpha$ and $\exists_\alpha$. It is crucial that this ramified language with abstraction terms is entirely formalizable in *M*, through a systematic coding of symbols. Once a set G is provided from the outside, a model $M[G] = \cup_{\alpha<\rho} M_\alpha[G]$ would be determined by the terms, where each $\dot{x}$ is to be interpreted by x for $x \in M$ and $\dot{G}$ is to be interpreted by G, so that: $M_0[G] = \{G\}$; for limit ordinals $\delta < \rho$, $M_\delta[G] = \cup_{\alpha<\delta} M_\alpha[G]$; and $M_{\alpha+1}[G]$ consists of the sets in $V_\alpha \cap M$ together with sets interpreting the abstraction terms as the corresponding definable subsets of $M_\alpha[G]$ with $\forall_\alpha$ and $\exists_\alpha$ ranging over this domain.

But what properties can be imposed on *G* to ensure that *M*[*G*] be a model of ZF? Cohen's key idea was to tie *G* closely to *M* through a system of sets in *M* called *conditions* that would approximate *G*. While *G* may not be a member of *M*, *G* is to be a subset of some $Y \in M$ (with $Y = \omega$ a basic case), and these conditions would "force" some assertions about the eventual *M*[*G*] that is, by deciding some of the membership questions whether $x \in G$ or not for $x \in Y$. The assertions are to be just those expressible in the ramified language, and Cohen developed a corresponding *forcing relation* $p \Vdash \varphi$, "*p* forces φ", between conditions *p* and formulas φ, a relation with properties reflecting his approximation idea. For example, if $p \Vdash \varphi$ and $p \Vdash \psi$, then $p \Vdash \varphi \,\&\, \psi$. The conditions are ordered according to the constraints they impose on the eventual *G*, so that if $p \Vdash \varphi$, and *q* is a stronger condition, then $q \Vdash \varphi$. Scott made an important suggestion simplifying the definition for negation: $p \Vdash \neg \varphi$ if for no stronger condition *q* does $q \Vdash \varphi$. It was crucial to Cohen's approach that the forcing relation, like the ramified language, be definable in *M*.

The final ingredient is that the whole scaffolding is given life by incorporating a certain kind of set *G*. Stepping out of *M* and making the only use of its countability, Cohen enumerated the formulas of the ramified language in a countable sequence (shades of Skolem's

Paradox!) and required that G be completely determined by a countable sequence of stronger and stronger conditions $p_0, p_1, p_2, \ldots$ such that for every formula φ of the ramified language exactly one of φ or $\neg \varphi$ is forced by some p_n. Such a G is called a *generic* set. Cohen was able to show that the resulting $M[G]$ does indeed satisfy the axioms of ZF: Every assertion about $M[G]$ is already forced by some condition; the forcing relation is definable in M; and so the ZF axioms, holding in M, mostly crucially Power Set and Replacement, can be applied to derive corresponding forcing assertions about ZF axioms holding in $M[G]$.

The extent and breadth of the expansion of set theory described henceforth far overshadows all that has been described before, both in terms of the numbers of people involved and the results established. With clear intimations of a new and concrete way of building models, set theorists rushed in and with forcing were soon establishing a cornucopia of relative consistency results, truths in a wider sense, some illuminating classical problems of mathematics. Many different forcings were constructed for adding new real numbers and iterated forcing techniques were quickly broached.

Robert Solovay played a prominent role in the forging of forcing as a general method, and he above all in this period raised the level of sophistication of set theory across its breadth from forcing to large cardinals. Solovay proved a result already in 1964 remarkable for its sophistication: Suppose that κ is an inaccessible cardinal; then in an inner model of a forcing extension, κ becomes $\aleph_1$, the least uncountable cardinal, every set of real numbers is Lebesgue measurable, and Dependent Choices (a substantial form of AC for bolstering measure) holds. This model offered important insights into the possibilities of measure and the limits imposed by AC. The inaccessible cardinal was thought for some time to be an artifact of the proof, when in 1979 Saharon Shelah finally complemented Solovay's result by showing that if every set of real numbers is Lebesgue measurable and Dependent Choices holds, then $\aleph_1$ (in V) is inaccessible in the constructible universe L.

Through the 1970s and into the 1980s the forcing method was honed with sophisticated iterated forcing techniques, techniques that established new, more contextualized relative consistency results in the self-generating mainstreams of set theory, *infinitary combinatorics* and *cardinal invariants of the continuum*. Donald Martin formulated an instrumental "axiom," *Martin's Axiom* (MA), in terms of forcing notions, an axiom that became convenient and focal for relative consistency results. MA together with the failure of CH is relative consistent with ZFC via forcing, and MA directly implies many combinatorial statements in a way analogous to how CH had, and so relative consistency results can be established by drawing direct consequences fom MA. A culmination in this direction was the work of Shelah in the 1980s on *proper forcing*, a wide class of forcing notions. Corresponding to MA in this context is the *Proper Forcing Axiom*, an axiom requiring large cardinals to establish its relative consistency. An important barrier that has resisted many efforts is that starting with a model of CH, many iterated forcing constructions have established the relative consistency of various propositions with the continuum being $\aleph_2$, but corresponding relative consistencies with the continuum being at least $\aleph_3$ are not known. Can this be a limitation of forcing, or a delimitation imposed by ZFC?

LARGE CARDINALS AND INNER MODELS

A subtle connection quickly emerged, already in the 1960s and into the 1970s, between large cardinals and combinatorial propositions low in the cumulative hierarchy: Forcing showed just how relative the Cantorian notion of cardinality is, since one-to-one correspondence functions could be adjoined to models of set theory easily, often with little disturbance. In particular large cardinals, highly inaccessible from below, were found to satisfy substantial propositions even after they were "collapsed" by forcing to $\aleph_1$ or $\aleph_2$, that is correspondence functions were adjoined to make the cardinal the first or second uncountable cardinals respectively. Conversely such propositions were found to entail large cardinal hypotheses in the clarity of an L-like inner model, sometimes the very same initial large cardinal hypothesis. Thus, in a subtle synthesis, hypotheses of length concerning the extent of the transfinite were correlated with hypotheses of width concerning the fullness of power sets low in the cumulative hierarchy, sometimes the arguments providing *equi-consistencies*. Solovay's Lebesgue measurability result from inaccessbility when complemented by Shelah's result became an equi-consistency, albeit a sophisticated one bringing together Cantor's two legacies, the investigation of definable sets of reals and the extension of number into the transfinite. Other "weak" large cardinals were formulated, sometimes in response to the need of a large cardinal concept to gauge a set-theoretic proposition via equi-consistency. The complementarity also encompassed "strong" large cardinal hypotheses formu-

lated in terms of elementary embeddings and later, new canonical inner models.

Large cardinal hypotheses stronger than measurability were charted out in the late 1960s, motivated not only by the heuristics of generalization but also by those of *reflection*. The direct reflection heuristic is that various properties attributable to the class of all ordinals, since its extent is uncharacterizable, should be attributable already to some cardinal. This heuristic was already at work in Zermelo's 1930 paper and extends the closure provided by Replacement. The more subtle reflection heuristic is that strong large cardinal hypotheses posit elementary embeddings j: $V \to M$, and the closer the target inner model M is to V, the stronger the properties that translate and can be reflected between. The *supercompact* cardinals were thus formulated by Solovay and William Reinhardt as global generalizations of measurable cardinals; stronger than these were the *n-huge* cardinals; and the stronger hypotheses still were formulated. There is an ultimate delimitation in this direction that has framed the possibilities: Kenneth Kunen established in ZFC that there can be no elementary embedding j: $V \to V$ of the universe into itself. ZFC rallied at last to force a veritable Götterdämmerung for large cardinals.

The theory of these strong hypotheses was developed particularly to investigate the possibilities for elementary embeddings. But what really intimated their potentialities were new forcing proofs, especially from supercompactness, that established the relative consistency of strong existence assertions low in the cumulative hierarchy, at the very least lending these assertions an initial plausibility. The possibility of new complementarity was then brought about through the development of *inner model theory*, the mostly sophisticated part of the theory of large cardinals.

Gödel's L was the first inner model, and Ronald Jensen dramatically transformed its investigation in the 1960s by refining the first-order definability and Skolem hull arguments to a "fine structure" analysis, extracting important combinatorial principles and establishing new relative consistencies. Inner models of measurability were soon developed, and their interactions and fine structure investigated, and these models would be paradigmatic for inner models of large cardinals: They exhibited in their crystalline clarity akin to algebraic closure the minimal consequences of the large cardinal hypothesis and the maximal structural regularity. In the 1970s, Jensen and Anthony Dodd developed the *core model* for measurability, and this would be paradigmatic for core models of large cardinals: These were inner models that did not contain the large cardinal, but exhibited the maximal possibilities "up to" the cardinal. The ascent through the large cardinal hierarchy had begun, the inner and core models providing an abiding sense of structure for large cardinal hypotheses.

The development of core models, while quickly developing a life of its own, was initially triggered by work on the Singular Cardinals Problem. With the advent of forcing it had been quickly seen that ZFC imposed little control on the powers 2^κ of *regular* cardinals κ, successor or limit, since it became possible to extend a model of set theory by adjoining arbitrarily many subsets of such κ without adjoining any subsets of smaller cardinals. Thus Cantor's Continuum Problem and its generalization to regular cardinals were informed by a general manifestation of method. What about singular cardinals? Powers of singular cardinals seemed much less flexible with respect to forcing, and the *Singular Cardinals Problem* is the general problem of clarifying the possibilities for the function 2^κ for singular cardinals κ. Jensen, who found a seminal 1974 result of Jack Silver on powers of singular cardinals "shocking," was directly inspired by it to establish the *Covering Theorem* for L, easily the most important result of the 1970s in set theory. Very loosely speaking this theorem asserts that unless a surprisingly simple proximity criterion between V and L holds, a large cardinal transcendence over L ensues. It was efforts to extend this result that led to the core models. Through forcing and inner model analysis, results especially of Moti Gitik of the late 1980s established equi-consistency results for simple assertions about powers of singular cardinals and showed remarkable level-by-level connections with large cardinals that affirmed their central place in the investigation of the transfinite.

The extensive research through the 1970s and 1980s considerably strengthened the view that the emerging hierarchy of large cardinals provides *the* hierarchy of exhaustive principles against which all possible consistency strengths can be gauged, a kind of hierarchical completion of ZFC. First the various hypotheses, though historically contingent, form a *linear* hierarchy, one neatly delimited by Kunen's inconsistency result. Typically for two large cardinal hypotheses, below a cardinal satisfying one there are many cardinals satisfying the other, in a sense prescribed by the first. And second, a variety of strong propositions have been informatively bracketed in consistency strength between two large cardinal hypotheses: the stronger hypothesis implies that there is a forcing extension in which the proposition holds; and if the

proposition holds, there is an inner model satisfying the weaker hypothesis.

One of the great successes for large cardinals has to do with perhaps the most distinctive and intriguing development in modern set theory. Although the *determinacy of games* has roots as far back as a 1913 note of Zermelo, the concept of infinite games only began to be seriously explored in the 1960s when it was realized that it led to "regularity" properties for sets of real numbers like Lebesgue measurability.

With ω the set of natural numbers let ${}^{\omega}\omega$ denote the set of functions from ω to ω. For $A \subseteq {}^{\omega}\omega$, $G(A)$ denotes the following "infinite two-person game with perfect information": There are two players, *I* and *II*. *I* initially chooses an $x(0) \in \omega$; then *II* chooses an $x(1) \in \omega$; then *I* chooses an $x(2) \in \omega$; then *II* chooses an $x(3) \in \omega$; and so forth:

$$\begin{array}{llllll} I: & x(0) & & x(2) & & \ldots \\ II: & & x(1) & & x(3) & \ldots \end{array}$$

Each choice is a *move* of the game; each player before making each of his moves is privy to the sequence of previous moves ("perfect information"); and the players together specify an $x \in {}^{\omega}\omega$. *I* wins $G(A)$ if $x \in A$, and otherwise *II* wins. A *strategy* is a function from finite sequences of natural numbers to natural numbers that tells a player what move to make given the sequence of previous moves. A *winning strategy* is a strategy such that if a player plays according to it he always wins no matter what his opponent plays. A is *determined* if either *I* or *II* has a winning strategy in $G(A)$. The extent of the determinacy of games was investigated through hierarchies of definable sets of reals, and in 1962 the following sweeping axiom was proposed:

AXIOM OF DETERMINACY. Every $A \subseteq {}^{\omega}\omega$ is determined.

This axiom actually contradicts the Axiom of Choice, as one can get a counterexample A by "diagonalizing" through all strategies, and so the axiom was intended to hold at least in some inner model to establish regularity properties for sets of real numbers there. In the late 1960s initial connections were made between the Axiom of Determinacy and large cardinals by Solovay, who showed in ZF that the axiom implies that $\aleph_1$ is measurable, and by Martin, who showed in ZFC that if there is a measurable cardinal, then the analytic sets, the simplest significant sets of real numbers definable with quantifiers ranging over real numbers, are determined. Investigating further consequences of determinacy, a new generation of descriptive set theorists soon established an elaborate web of connections in the unabashed pursuit of structure for its own sake. Determinacy hypotheses seemed to settle many questions about definable sets of reals and to provide new modes of argument, leading to an opaque realization of the old Cantorian initiatives concerning sets of real numbers and the transfinite with determinacy replacing well-ordering as the animating principle. By the late 1970s a more or less complete theory for the "projective" sets of real numbers was in place, and with this completion of a main project of descriptive set theory attention began to shift to questions of overall consistency.

The investigation of the Axiom of Determinacy spurred dramatic advances in the theory of large cardinals and affirmed their central role in gauging consistency strength. In the 1970s the strength of the methods made possible by the axiom led to speculation that either the axiom was orthogonal to large cardinals or would subsume them in a substantial way. However, large cardinal hypotheses, first near Kunen's inconsistency and then around supercompactness, were shown to tame Determinacy. By looking at the workings of a proof, Hugh Woodin in 1984 formulated what is now known as a *Woodin cardinal*. Then Martin and John Steel showed that having more and more Woodin cardinals establishes the determinacy of more and more sets in the "projective hierarchy" of sets, sets of real numbers definable with quantifiers ranging over the real numbers. Finally Woodin established by 1992: *the existence of infinitely many Woodin cardinals is equi-consistent with the Axiom of Determinacy*. Woodin cardinals are weaker than supercompact cardinals, closer to measurable cardinals, and in subsequent developments the inner model theory was advanced to getting inner and core models of Woodin cardinals.

Woodin in the late 1990s built on the wealth of ideas surrounding Woodin cardinals and Determinacy and raising them to a higher level proposed a resolution of the Continuum Problem itself. This resolution features the use of arbitrarily many Woodin cardinals, the assimilation of new principles for sets of sets of real numbers, and an unresolved new conjecture about a new "logic" that would complete the picture. Thus structural ideas involving large cardinal hypotheses may circle back to effect an ultimate resolution of the original problem that stimulated the development of set theory.

What about the consistency of large cardinal hypotheses? As postulations for cardinals of properties of the class of all ordinals, they inherit substantial inaccessi-

bility properties from below, but even for large natural numbers given notationally, the meaning of a number is not conveyed by its dogged approach from below but by its mathematical postulation and the sense given it by proof and method. The inner model theory has fortified large cardinals up to Woodin cardinals by providing them with coherent inner models whose structure incisively exhibit their consistency. As for the hypotheses near Kunen's inconsistency, since that result was based on a combinatorial contingency, it could well be that a like inconsistency for a weaker hypothesis can be established. In any case these near-inconsistency hypotheses are less relevant, the forcing proofs applying them to get initial plausibilities having given way to more refined arguments from weaker hypotheses. Moreover the work of Woodin has shown that there is also quite a lot of structure near the Kunen inconsistency, analogous to the descriptive set theory of real numbers.

Stepping back to gaze at modern set theory, the thrust of mathematical research should deflate various possible metaphysical appropriations with an onrush of new models, hypotheses, and results. Shedding much of its foundational burden, set theory has become an intriguing field of mathematics where formalized versions of truth and consistency have become matters for manipulation as in algebra. As a study couched in well-foundedness ZFC together with the spectrum of large cardinals serves as a court of adjudication, in terms of relative consistency, for mathematical propositions that can be informatively contextualized in set theory by letting their variables range over the set-theoretic universe. Thus set theory is more of an open-ended framework for mathematics rather than an elucidating foundation. It is as a field of mathematics proceeding with its own internal questions and capable of contextualizing over a broad range that set theory has become an intriguing and highly distinctive subject.

See also Cantor, Georg; First-Order Logic; Gödel, Kurt; Gödel's Theorem; Hilbert, David; Logical Paradoxes; Logic, History of: Modern Logic; Mathematics, Foundations of; Model Theory; Neumann, John von; Russell, Bertrand Arthur William; Second-Order Logic; Tarski, Alfred; Truth.

Bibliography

Bartoszy'nski, Tomek, and Haim Judah. *Set Theory: On the Structure of the Real Line*. Wellesley, MA: A K Peters, 1995.

Cantor, Georg. *Contributions to the Founding of the Theory of Transfinite Numbers. Mathematische Annalen 46* (1895), 481–512; and II. *Mathematische Annalen* 49 (1897): 312–351; with introduction and notes by Philip E. B. Jourdain. Chicago: Open Court, 1915. Reprinted New York: Dover, 1965.

Dauben, Joseph W. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Cambridge, MA: Harvard University Press, 1979.

Ewald, William, ed. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Oxford: Clarendon Press, 1996.

Gödel, Kurt. *Collected Works*. Vol. 1. New York: Oxford University Press, 1986.

Gödel, Kurt. *Collected Works*. Vol. 2. New York: Oxford University Press, 1990.

Hallett, Michael. *Cantorian Set Theory and Limitation of Size*. Logic Guides #10. Oxford: Clarendon Press, 1984.

Jech, Thomas. *Set Theory*. Berlin: Springer Verlag, 2002.

Kanamori, Akihiro. *The Higher Infinite: Large Cardinals in Set Theory from their Beginnings*. 2nd ed. Berlin: Springer Verlag, 2003.

Kanamori, Akihiro. "Zermelo and Set Theory." *The Bulletin of Symbolic Logic* 10 (2004): 487–553.

Kunen, Kenneth. *Set Theory: An introduction to Independence Proofs*. Amsterdam: North-Holland, 1980.

Moore, Gregory H. *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*. New York: Springer-Verlag, 1982.

Moschovakis, Yiannis N. *Descriptive Set Theory*. Amerstadam: North-Holland, 1980.

Shelah, Saharon. *Cardinal Arithmetic*. Oxford: Clarendon Press, 1994.

Shelah, Saharon. *Proper and Improper Forcing*. 2nd ed. Berlin, Springer, 1998.

Van Heijenoort, Jean, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Cambridge, MA: Harvard University Press, 1967.

Woodin, W. Hugh. *The Axiom of Determinacy, Forcing Axioms, and the Nonstationary Ideal*. Berlin: Walter de Gruyter, 1999.

Woodin, W. Hugh. "The Continuum Hypothesis I and II." *Notices of the American Mathematical Society* 48 (2001): 567–576 and 681–690.

Akihiro Kanamori (2005)

SEXISM

The term *sexism* denotes any system of beliefs, attitudes, practices, social norms, or institutional forms that functions to create or perpetuate invidious social distinctions among persons on the basis of their actual or presumed sex. This characterization of sexism reflects a widespread consensus among feminist theorists and queer theorists that the phenomenon cannot be understood simply in terms of the prejudices or ill-intentioned behavior of individuals, but rather must be seen as involving wide-ranging social structures, structures that can affect both the meanings and consequences of the actions of individuals, even if such actions are otherwise benign.

Marilyn Frye (1983) has explained, in just these terms, the inadequacy of a definition of sexism as any act or policy involving an "irrelevant or impertinent marking of the distinction between sexes." She then bids us to consider an employer who refuses to hire a woman for a managerial position on the stated grounds that his employees would not accept the authority of a woman. We may suppose that the supervisor is right about his employees' attitudes; thus the woman's sex is, as things stand, relevant to her ability to do the managerial job. By this definition, then, the employer's act of discrimination would not count as sexist, yet sexism is surely at work, somehow, in this situation.

The problem, Frye argues, is that the definition fails to take account of the ways in which preexisting social conditions can make sex relevant in situations where it need not be, and should not be. In this case, myriad factors had already conspired to create conditions in which the male employees would not be inclined to trust a woman's judgment, or would not accept her possessing even limited authority over them. Such factors might include: 1) a history of explicit, de jure discrimination against women, limiting their participation in public life and their opportunities to assume authoritative social roles; 2) widespread belief (perhaps as a result of number one above) that women are incapable of carrying out the duties of a manager; 3) gender norms that would make it humiliating for men to submit to the authority of a woman. Against such a background, a woman's sex becomes relevant to the question of whether she can do the job at hand, no matter what other relevant qualifications she possesses.

Another important point illustrated by this example is that sexism does not require bad intentions. In Frye's case, we need not assume that the employer shares his workers' prejudices. He may believe the woman to be otherwise qualified, and regret that circumstances make it unwise for him to hire her. Whether he should be criticized for accommodating the sexist views of his employees is certainly a reasonable question, but it is not the same as the question whether the hiring decision was sexist. The point of recognizing sexism is not to indict and punish individuals, but rather to identify and alter all the factors that contribute to the subordination of women, where it is acknowledged that many of these will involve well-entrenched and perfectly mundane social practices.

Richard Wassserstrom (1977) has made the same point, distinguishing what he calls institutional racism and sexism from overt and covert racism and sexism. In the latter two cases, laws or policies are designed with the explicit intent of allocating unjustified burdens or unwarranted benefits to individuals on the basis of race or sex. In the overt cases, the categories of race and sex are explicitly mentioned, whereas in the covert cases, the categories are represented by surrogates. Jim Crow laws in the United States and the denial of the franchise to women in the United States prior to 1920 exemplify overt racism and sexism, respectively. The use of grandfather clauses after the U.S. Civil War (1861–1865) to disenfranchise formerly enslaved black men, and workplace safety rules that bar from certain jobs anyone capable of becoming pregnant exemplify covert racism and sexism.

Institutional racism and sexism, however, differ from both covert and overt forms in that there need be no intention on anyone's part to produce a racist or sexist consequence. Wasserstrom recognizes two sub-forms. The first involves regulations or practices that, while apparently neutral, operate against a social reality already configured by racism or sexism, so that their effects are to maintain or reinforce an unjust social hierarchy. This may be the case with the practice of assigning children to schools on the basis of their neighborhoods—if there is de facto segregation in housing, then the assignment policy will lead to or sustain racial segregation of schools, even if no one intended that outcome. The second subform of institutional racism or sexism involves practices that embody—at a level not readily accessible to consciousness—racist or sexist concepts and presuppositions. Let us consider in this context the question, raised frequently by critics of feminism, whether it is sexist for a man to offer a seat to a woman on a bus.

It is certainly true that many men sincerely regard gestures of this sort as courteous and respectful, and are offended by the suggestion that the gestures denigrate women. Still, if one probes the larger meaning of these customs, it becomes clear that they are part of a system of conventions that symbolically express and prescribe women's dependency upon and subordination to men. In the first place, there is no plausible moral or empirical rationale for making sex per se the criterion for the appropriateness of such a gesture. General moral considerations dictate that any able-bodied person—regardless of their sex—ought to offer a seat to anyone who is visibly incapacitated or subject to physical stress; a person on crutches, someone struggling with packages, or someone obviously exhausted, as might be the case for a woman in the late stages of pregnancy. But these conditions are not connected to sex. The suggestion that women must be accorded respect in virtue simply of their being female is simply peculiar.

There are obvious rationales for social conventions encoding such a stance toward the elderly: Age correlates with experience and wisdom and younger people are generally indebted in various ways to older people. But there is no estimable quality that correlates with femaleness per se. By contrast, the gestures and conventions we are considering make perfect sense in light of the prerequisites of a gender system that makes physical weakness normative for women (the "weaker sex") and physical strength and control normative for men. Such gestures express the relationship that ought to hold between men and women. In a system of social organization in which men hold, or are assumed to hold, real power over women, conventional acts of faux deference by men to women (such as a man's rising when a woman enters the room) function not as symbols of genuine respect, but rather as expressions of *noblesse oblige*.

Feminist theorists agree that the kinds of gender roles that exist in patriarchal societies are the raw material of sexism, but disagree about whether the elimination of sexism requires the complete dismantling of gender roles, or only their reform. Some feminist theorists argue that any way of attaching systematic social significance to biological sex will inevitably prove sexist. Theorists of this sort include so-called humanistic feminists, who hold that biological sex is a property accidental to, and thus morally irrelevant to one's humanity (De Beauvoir 1973, Nussbaum 1999, Antony 1998), and dominance theorists, who hold that gender differences are constructed ex post facto to mask or rationalize preestablished power disparities (Haslanger 2000, MacKinnon 1987). All such theorists point out that gender roles function to enforce both sexual dimorphism (the demand that one be clearly identifiable and self-identified as either male or female), and compulsory heterosexism (the requirement that one's erotic interest be focused exclusively on individuals of the opposite gender).

These restrictive social norms are not only deeply oppressive to transgendered and nonheterosexual people, but distorting and limiting for all members of a human society. Theorists who hold this position generally believe that a great deal of the content of gender roles is socially constructed—that there is no natural necessity linking the components of biological sex (morphology, endocrinology, and genetics) to the features of particular gender roles. But the issue of the naturalness of gender roles is in fact orthogonal to the question whether such roles should be socially enforced. Myopic people are biologically (and probably genetically) different from non-myopic individuals; nonetheless, we assign no social significance to this difference, and in fact acknowledge a social obligation to mitigate the natural consequences of poor eyesight.

Other feminist theorists hold that there is nothing inherently wrong with the existence of gender roles, and that such roles could, in a different social context, be liberatory and beneficial for all. On this view, sexism is constituted by two things: a) the gratuitous attachment of undesirable qualities, such as physical weakness, to the feminine gender role; and b) the widespread devaluation of central elements of that role, such as emotional sensitivity. According to these theorists, sometimes called *gynocentric* or *difference* feminists, facts about the female role in biological reproduction have inherent social significance, and so there is no serious prospect for eliminating social roles erected on the basis of reproductive difference (Young 1985). Central, then, to a feminine gender role will be the social role of mothering. Because the individuals who have been the predominant occupiers of this role are women, and because women have historically lacked social power, the virtues necessary for the proper performance of this role (e.g., empathy, cooperativeness, imaginativeness, nurturance, and altruism) have been devalued.

To dismantle sexism, these feminine virtues must be recognized as being as important to morality as masculine virtues such as impartiality (Gilligan 1982) and greater social support must be provided those who fulfill such typically feminine roles as tending children, caring for the sick, and managing social relationships (Ruddick 1989).Closely allied with gynocentric feminists are *ecofeminists*, who think that women's greater involvement with the bodily realities of birth, growth, and even death (in their roles as nurses and caregivers) create for women a more intimate relationship with the natural world than men have (Plumwood 1993). This, in turn, makes women more apt than men to strive for ways of life that are harmonious with nature, with nonhuman animals, and with other human beings. All these theorists agree that war and other forms of violence reflect the sexist devaluation of the feminine, and that a proper appreciation of feminine virtues is essential to producing peace.

This dispute within feminist theory about the nature of sexism has implications for social and legal policy. Gynocentric feminists charge that humanistic feminists are guilty of androcentrism—taking the male as the paradigm of the human. If laws and social institutions take no account of differences between men and women, then women will be forever socially and economically disadvantaged by policies and practices centered on male

needs, abilities, and interests. Humanistic feminists counter that they find the content of masculine gender roles just as objectionable as the content of feminine roles, and equally in need of elimination. The revalued gender roles envisioned by gynocentric feminists reflect a romanticized view of female experience, and threaten to legitimate a host of sexist stereotypes and prescriptions. Laws and policies should be based on parameters rationally related to the issue involved, parameters that will sometimes coincide with sex differences, but will more often not.

Feminist theorists have been increasingly concerned with understanding interactions among sexism and a host of other systems of oppressive social division, including racism, heterosexism, class oppression, ageism (invidious division on the basis of age), and ableism (invidious division on the basis of physical capacities), ethnocentrism, and jingoism. Critical legal theorist Kimberle Crenshaw has introduced the notion of intersectionality to capture the sui generis character of multidimensional oppression. Postmodern feminist theorists have appealed to this ever-increasing list of interacting parameters of identity to deconstruct categories such as sex and race, aruging that no one is simply or straightforwardly a woman or a black person, but that the self is essentially fragmented and fluid. However, they acknowledge the difficulty of making sense of oppression without appeal to such categories.

See also Affirmative Action; Analytical Feminism; Feminist Legal Theory; Feminist Philosophy of Science: Contemporary Perspectives; Feminist Social and Political Philosophy; Frye, Marilyn; Heterosexism; Racism; Violence.

Bibliography

Antony, Louise. "Back to Androgyny: What Bathrooms Tell Us about Equality." *Journal of Contemporary Legal Issues* 9 (1998): 1–20.

Crenshaw, Kimberlé. "Demarginalizing The Intersection of Race and Sex." *University of Chicago Legal Forum* (1989): 139–167.

De Beauvoir, Simone. *The Second Sex*. Translated by H. M. Parshley. New York: Bantam Books, 1973.

Frye, Marilyn. "Sexism." In *The Politics of Reality*. Freedom, CA.: The Crossing Press, 1983.

Gilligan, Carol. *In A Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press, 1982.

Haslanger, Sally. "Gender, Race: (What) Are They? (What) Do We Want Them to Be?" *Noûs* 34 (1) (2000): 31–55.

MacKinnon, Catherine. "Difference and Domination: On Sex Discrimination." In *Feminism Unmodified: Discourses on Life and Law*. Cambridge, MA: Harvard University Press, 1987.

Nussbaum, Martha. "Women and Cultural Universals." In *Sex and Social Justice*. Oxford: Oxford University Press, 1999.

Plumwood, Val. *Feminism and The Mastery of Nature*. London: Routledge, 1993.

Ruddick, Sara. *Maternal Thinking: Toward a Politics of Peace*. New York: Basic Books, 1989.

Wasserstrom, Richard. "Racism, Sexism, and Preferential Treatment: An Approach to the Topics" *UCLA Law Review* (1977): 581–615.

Young, Iris. "Humanism, Gynocentrism, and Feminist Politics" *Women's Studies International Quarterly* 3 (1985): 173–183.

Louise M. Antony (2005)

SEXTUS EMPIRICUS
(third century CE)

Sextus Empiricus was almost certainly, as his name suggests, a doctor of the empiricist school, which flourished from the third century BCE until at least the third century CE. His dates are very uncertain, but he probably lived and worked, perhaps in Rome, sometime early in the third century CE. He is mentioned as a prominent skeptic in Diogenes Laertius's *Lives of the Philosophers* (DL) 9.116, written probably in the third century; but the men Diogenes names as his teacher and pupil, Herodotus of Tarsus and Saturninus, are even more obscure. He was certainly known as the authoritative source of skeptical argument a century later, when Saint Gregory of Nazianzus blamed him for the "vile and malignant disease" of arguing both sides of a question that was infecting the church. How original he was is unclear—it is hard to tell partly because our other sources for skepticism are so exiguous—but he seems to have been more of a compiler than an original thinker; and in any case it is to his preservation of a large body of skeptical argumentation, whatever its provenance, that his importance in the history of philosophy is due.

His best-known work is *Outlines of Pyrrhonism* (PH), a digest of the arguments and general strategy of that version of skepticism named for Pyrrho of Elis. The precise relationship between the position outlined by Sextus and that of Pyrrho is unclear—Sextus himself says that his philosophy is called "Pyrrhonism" because "he seems to have applied himself to skepticism more thoroughly and conspicuously than any of his predecessors" (PH 1.7). PH summarizes what Sextus presents at much greater length in another work that also survives, at least in part. This consists of the five surviving books generally

(if misleadingly) known as *Against the Professors* (M) 7–11, subdivided as *Against the Logicians* (M 7–8), *Against the Physicists* (M 9–10), and *Against the Ethicists* (M 11). M 7–8 corresponds roughly to an expanded version of PH 2, while M 9–11 is summarized in PH 3. Whatever answered to the general treatment of skepticism in PH 1 is lost. The texts known as M 1–6 form a different treatise, written with a rather different aim, consisting of a series of essays directed against the practices (and practitioners) of six of the seven "liberal arts," in order, grammar, rhetoric, arithmetic, geometry, astrology, and music. Sextus also tells us that he wrote medical texts as well, but these are lost.

Book 1 of PH presents a general account of skeptical practice. Sextus is careful to avoid any suggestion that what he is presenting is a doctrine, or anything else with uncomfortably Dogmatic overtones. "Dogmatist" was the general name given by skeptics to their opponents who held positive, or even negative, views (the term also functions as an umbrella term to describe theoretically minded, as opposed to empirical, doctors; and was probably borrowed from medical terminology). At the beginning of PH, Sextus presents skeptics as starting out from the same position as all other inquirers: They seek to assuage their disquiet by finding out the truth about things. But in any search there are three possible outcomes: one may (a) claim that one has found what one was looking for; (b) deny that it can be found, saying that it is inapprehensible; or (c) simply keep on searching. Option (a) is the position of the Dogmatists (Stoics, Epicureans, Platonists, and Aristotelians). Option (b) Sextus ascribes to the Academic skeptics, claiming (controversially) that they positively deny that things are apprehensible (in the sense of certainly knowable) as to their actual natures, while also claiming that certain positions, in regard to both factual and evaluative matters are "more plausible" than others (PH 1 236–241); and both of these positions are unacceptably Dogmatic from a Sextan perspective. Moreover, even though both schools report that *epochē* (suspension of judgment) regarding things that are naturally nonevident is the natural conclusion of the inquiry, the Academics present this as a goal, and as a good thing, whereas the Pyrrhonist has no attitude to it at all—it is simply something that happens, although it seems to bring tranquility (the acquisition of which was the initial object of the search) in its train (PH 1.21–30). The Pyrrhonist follows option (c), and keeps on inquiring, following the appearances, but suspending judgment about what, if anything, might lie behind them (PH 1.13–24).

Sextus is acutely aware of the dangers of incoherence involved in this presentation of a life without commitment—he cannot consistently *recommend* it (since that would involve supposing it to be objectively good, or at least choiceworthy); and he cannot claim that, as a matter of fact, following it will have the desired therapeutic effect of removing anxiety. But he can (undogmatically) report his own experiences; it seems to him that this is how things have gone. Moreover, he is moved (so he tells us) by benevolence: Seeing the Dogmatists suffering from their vain pretensions to knowledge, he seeks to cure them (PH 3.280–281), not because he positively affirms that it is good to do so, but simply because he finds himself so moved. In the same vein, skeptics adhere (undogmatically) to "a quadripartite practice of ordinary life," since "we cannot remain wholly inactive": they follow "the guidance of nature, the constraint of the affections, the tradition of laws and customs, and the instruction of the arts" (PH 1.23). The skeptic has a "criterion of action"—the appearances—but no "criterion of truth" (PH 2.13–79). Sextus thus shows himself sensitive to the sort of objection made famous by David Hume, but anticipated in the Greek tradition, that skepticism is fatal.

PH 1.31–163 presents a version of the Ten Modes of Aenesidemus, but in a manner that betrays the later influence of Agrippa. The Modes are collections of considerations designed to (or rather, which have been found to) induce *epochē* on all nonphenomenal matters (the causal language is important: there can be no inference as such for the Pyrrhonist). They consist in the collection of "oppositions"—Sextus describes skepticism as "a capacity for oppositions": PH 1.8—cases where (apparently) x appears F to y (or in circumstances C), but not-F to z (or in circumstances C^*); since there can be no non-question-begging way of deciding on the superiority of any one of the opposed appearances over any other (that is, we have no criterion) "we are moved to suspend judgment" (PH 1.78; cf. 1.89, 99, 117). The skeptic will adduce considerations on both sides of any question to promote "equipoise of argument" (PH 1.8, 190), not in order to support or undermine one side or the other. If you claim p, skeptics will adduce reasons why not-p, not because they believe them *in propria persona*, but simply because their benevolence compels them to. All skeptical argument is dialectical. Thus when Sextus produces arguments against proof (PH 2.134–192), he does so not because he *believes*, inconsistently, in the capacity of proof, as the Dogmatists allege (PH 2.185–186); rather it is the Dogmatists who, insofar as they believe the canons of rational argument, must find those canons undermined from within. The skeptic has no beliefs about them

at all (PH 2.187–192). And it is in this manner that Sextus deploys the vast bulk of his argumentation in PH 2–3 and M 7–11, for and against particular philosophical and scientific positions.

In a similar vein, Sextus discusses the "skeptical slogans," such as "no more [so than not so]," *ouden mallon* (PH 1.188–191). The phrase *ouden mallon* had been used by earlier philosophers (including Democritus, Plato, Protagoras, and Pyrrho), but to signal, non-skeptically, that some things really *were* no more *F* than not-*F*. For the skeptic, it functions simply as a marker of a refusal to say, one way or another. Equally, when skeptics say "I determine nothing," or "all things are undetermined" (PH 1.197–199), they do not assert that nothing is determinable; these are merely expressions of how things seem. Indeed, the skeptical slogans apply to themselves: skeptics determine nothing, not even that they determine nothing (PH 1.206–209); here again they can avoid the charge of "negative Dogmatism" (option (b) above) they level at the Academics.

At PH 1.210–241, Sextus seeks to distinguish Pyrrhonism from other superficially similar philosophies by stressing the fact that all of them slide into Dogmatism. Thus Pyrrhonism is not relativism, at least if that positively affirms that everything is relative; the Pyrrhonist appeals to the relativity of appearances, but draws no ontological conclusion therefrom. Curiously, Sextus even distinguishes his practice from that of the empiricist doctors. The latter follow the appearances; and make use of the type of sign ("commemorative") that Sextus allows (PH 2.100–102), in which something evident is a sign of something else that is only temporarily nonevident, that is, whose existence can be confirmed by further investigation, as when smoke is a sign of (concealed) fire. (Sextus rejects "indicative signs," whereby Dogmatists seek to infer to the hidden internal structures of things on the basis of evident phenomena, on the grounds that there can be no noncontroversial inference of such a kind: PH 2.104–133; M 8 199–300). But they also developed a complex epistemology of reasonable expectation, based upon personal observation, reportage, and argument form analogy. And for Sextus, this strays too far toward a Dogmatic supposition that certain outcomes really are more likely than others. For this reason, he prefers the methodic school of medicine as a model, since this school also makes no affirmations, and simply follows the "quadripartite practice of ordinary life" (PH 236–141). In this vein, in M 1–6, Sextus allows that it is fine to practice some skill, as long as one does so undogmatically, that is without commitment to any supposed deep truths that the skill relies upon. This is the sense in which the skeptic may follow "the instruction of the arts," and how Sextus may consistently be a (type of) doctor.

See also Aenesidemus; Agrippa; Ancient Skepticism; Pyrrho.

Bibliography

PRIMARY WORKS

Sextus Empiricus. *Outlines of Skepticism*. Translated by Julia Annas and Jonathan Barnes. Cambridge, U.K.: Cambridge University Press, 1994.

SECONDARY WORKS

Bailey, Alan. *Sextus Empiricus and Pyrrhonean Scepticism*. Oxford: Oxford University Press, 2002.

Barnes, Jonathan. *The Toils of Scepticism*. Cambridge, U.K.: Cambridge University Press, 1990.

Hankinson, R. J. *The Sceptics*. London: Routledge, 1995.

R. J. Hankinson (2005)

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

SHAFTESBURY – ZUBIRI

9 *volume*

2nd edition

ENCYCLOPEDIA OF PHILOSOPHY

DONALD M. BORCHERT
Editor in Chief

MACMILLAN REFERENCE USA
An imprint of Thomson Gale, a part of The Thomson Corporation

THOMSON GALE

Detroit • New York • San Francisco • San Diego • New Haven, Conn. • Waterville, Maine • London • Munich

THOMSON
GALE

Encyclopedia of Philosophy, Second Edition

Donald M. Borchert, Editor in Chief

For more information, contact
Macmillan Reference USA
An imprint of Thomson Gale
27500 Drake Rd.
Farmington Hills, MI 48331-3535
Or you can visit our internet site at
http://www.gale.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of philosophy / Donald M. Borchert, editor in chief.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-02-865780-2 (set hardcover : alk. paper)—
ISBN 0-02-865781-0 (vol 1)—ISBN 0-02-865782-9 (vol 2)—
ISBN 0-02-865783-7 (vol 3)—ISBN 0-02-865784-5 (vol 4)—
ISBN 0-02-865785-3 (vol 5)—ISBN 0-02-865786-1 (vol 6)—
ISBN 0-02-865787-X (vol 7)—ISBN 0-02-865788-8 (vol 8)—
ISBN 0-02-865789-6 (vol 9)—ISBN 0-02-865790-X (vol 10)
1. Philosophy–Encyclopedias. I. Borchert, Donald M., 1934-

B51.E53 2005
103–dc22

2005018573

This title is also available as an e-book.
ISBN 0-02-866072-2
Contact your Thomson Gale representative for ordering information.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

contents

S

SHAFTESBURY, THIRD EARL OF (ANTHONY ASHLEY COOPER)

(1671–1713)

Anthony Ashley Cooper (the Third Earl of Shaftesbury) was born in London in the home of his grandfather, the first earl, a prominent Whig politician, who put his secretary and friend, John Locke, in charge of his grandson's education. Fluent at eleven in both Greek and Latin, Shaftesbury was an avid student of ancient philosophy, particularly Plato and the Stoics. In 1686, accompanied by a tutor, he embarked on a three-year tour of the Continent, learning French and acquiring a sophisticated taste for the arts. He was elected to Parliament in 1695 and served for three years, although asthma prevented him from standing for reelection. In 1698 he moved to Holland, where he met Pierre Bayle, an advocate for religious tolerance and one of the first to argue that it is possible for an atheist to be virtuous. After becoming the Third Earl of Shaftesbury in 1699, he attended meetings of the House of Lords until 1702, but once again ill health prevented him from continuing to serve and being more active in Whig causes. He married Jane Ewer in 1709; they had one son. His bad health forced him to move in 1711 to Italy, where he died in 1713.

BACKGROUND

Shaftesbury's first published work was an edited collection of the sermons of the Cambridge Platonist Benjamin Whichcote (1609–1683). In his preface Shaftesbury attacked Thomas Hobbes's conception of morality as a matter of law springing from the will of a sovereign, backed up by sanctions imposed on us to restrain our natural, selfish tendencies. His letters make clear, however, that he thought John Locke was an even greater threat to morality since he made Hobbes's views more respectable. Rejecting Locke's view that moral laws spring from the will of God and that morality requires sanctions, Shaftesbury complained that Locke not only "threw all order and virtue out of the world" but also made moral ideas "unnatural," without any "foundation in the mind" (1900, p. 403). In the Cambridge Platonists, however, he found doctrines that were both congenial to his own outlook and an antidote to those of Hobbes and Locke. Proposing a conception of morality that centered on love, the Cambridge Platonists emphasized the natural goodness and sociability of human beings and our ability to act virtuously without sanctions.

Shaftesbury's chief work is *Characteristicks of Men, Manners, Opinions, Times*, an anthology of his essays. It was first published in 1711 in three volumes; ten more editions were printed by 1790. *Characteristicks* includes "An Inquiry concerning Virtue or Merit," which John Toland originally published in 1699, although there is dispute about whether Shaftesbury authorized that version. He revised the "Inquiry" for inclusion in *Characteristicks*. The other four essays were written between 1705 and 1710 and cover a variety of topics in different genres. He discusses issues in morality, politics, religion, aesthetics, culture, and what he calls "politeness"—the conventions of good manners and refined conversation. The essays take different forms: the traditional treatise, as well as an epistle, a dialogue, and a soliloquy. He includes his own commentaries or "miscellaneous reflections" on each essay, which were written especially for the collection.

CONCEPTION OF PHILOSOPHY

Shaftesbury's unorthodox writing style goes hand in hand with his conception of philosophy as practical. He laments that philosophy "is no longer active in the world" (1711/1999, p. 232). On his view, philosophy should help people fashion themselves into moral and unified beings. Conceiving of moral self-transformation in Socratic terms as the pursuit of self-knowledge, he suggests that the best way to know yourself is by means of an inner dialogue. Dialogues and soliloquies, rather than lectures and sermons, are therefore the appropriate vehicles for inspiration and edification. His intended audience was cultivated readers rather than philosophers and other academics, so he thought his writing needed to be accessible—easy, smooth, and polite.

Shaftesbury's practical conception of the philosophical enterprise led him to reject metaphysical and epistemological studies on the grounds that they make people "neither better, nor happier, nor wiser" (1900, p. 269). He was largely indifferent to the successes in the natural sciences that were made during this period and opposed mechanistic conceptions of nature. In contrast to many eighteenth-century philosophers, he was uninterested in putting morality on a scientific footing. He preferred ancient philosophy to that of his contemporaries.

Shaftesbury is best read as a transitional figure, a bridge between the philosophical thinking of the ancients and the moderns, as well as between the seventeenth and the eighteenth centuries. Although he rejected the seventeenth-century natural law view of morality, he retained its Stoic conception of the universe as teleologically structured. The natural world is an integrated and harmonious whole composed of many subsystems, all of which are ordered to good ends. Each subsystem or species, including the human species, is designed to play specific functional roles in still larger systems, which together form the universal nature, the system of all things. The order and harmony in universal nature is a product of God's creative intelligence. As a reflection of God's intelligence, the universe itself embodies rational principles. Shaftesbury's teleological picture of the universe underwrites many of his views on religion, morality, and aesthetics.

ETHICAL THEORY

As Henry Sidgwick remarks in his *Outlines of the History of Ethics [for English Readers]* (1886), Shaftesbury's *Characteristicks* "marks a turning point" in the history of ethics, since he is the first to take "psychological experience as the basis of ethics" (p. 190). He makes morality dependent on the mind in two ways. First, first-order sentiments—the passions and affections that motivate people to act—and actions expressive of these sentiments—have moral value. Second, what gives these motives their value are reflective, second-order sentiments—sentiments we have about our own or other people's sentiments. Shaftesbury's inward turn was the inspiration for sentimentalist moral theories, especially Francis Hutcheson's and David Hume's, as well as Bishop [Joseph] Butler's electric theory.

Shaftesbury's best-known work today is his most traditional piece of writing, "An Inquiry concerning Virtue or Merit." The question that frames the "Inquiry" is whether virtue is able to support itself without the aid of religion. In the course of answering that question he explains both the nature of virtue and our obligation to it. Distinguishing between natural goodness and moral goodness, he defines natural goodness in a functional or teleological way. To say that something is naturally good is to say that it contributes to the good of the system of which it is a part. Where a subsystem is part of a larger system, judgments of natural goodness are relative to that larger system. He even says that something is "really" good or bad only if it benefits or hinders universal nature. However, when we judge the natural goodness or badness of a sensible creature, our judgments concern the structure or economy of its affections. Sensible creatures are good if their affections are adapted to contribute to the good of their species. Their goodness is a matter of being in a healthy state, one that enables them to realize their natural ends. Not surprisingly, Shaftesbury often equates the good with the natural and evil with the unnatural.

While sensible creatures are capable of natural goodness, Shaftesbury claims that only rational creatures are capable of moral goodness—virtue—because only they have the capacity to make their affections objects of reflection. When affections are "brought into the mind by reflection … there arises another kind of affection towards those very affections themselves and … now become the subject of a new liking or disliking" (1711/1999, p. 172). As rational creatures, human beings have second-order, reflective sentiments, sentiments about sentiments. Shaftesbury calls this reflexive capacity a "moral sense." He conceives of it in aesthetic terms—a sense of what is beautiful or harmonious, foul or dissonant in our sentiments. The harmony and proportion of the affections, like the natural beauty in the universe, is evidence of a creative designing mind: God. In feeling moral approval we are able to share in the divine intelligence that created the beauty in the universe. On Shaftesbury's view the moral sense is an active, intelligent, and creative power, not the passive faculty that Hutcheson took it to be.

Shaftesbury argues that what the reflective sense approves of, and so makes morally good, is our natural goodness. We are naturally good when our "natural" or social affections and our self-directed affections are balanced in such a way as to promote our own good and the good of our species. While he thinks that our concern for others may be too strong and our self-concern may be too weak, more typically people are vicious when their social affections are too weak or their self-directed affections too strong. Moral evil arises not only from an imbalance between the social and self-interested affections but also from such "unnatural" affections as malice, sadism, and "delight in disorder."

After explaining the nature of virtue Shaftesbury turns to the question of our obligation to virtue, which he takes to mean "what reason there is to embrace" a virtuous life (1711/1999, p. 192). He then proceeds to show that virtue and self-interest coincide. He begins by arguing that mental pleasures are superior to physical pleasures. He thinks that there are two kinds of mental pleasures: those that consist in the operation of first-order affections and those that result from second-order affections such as those of the moral sense. The first-order affections that are social are a superior source of pleasure since they are pleasant in themselves, never go stale, and enable us to share sympathetically in the pleasures of others. More important, virtuous people experience the pleasures of their own approval as well as the approval of others, while vicious people suffer the torments and pangs of their own disapproval and those of others. He concludes that what obligates us to the practice of virtue is that being virtuous makes us happy. Being a virtuous person is not only good but also good for you.

Returning to the topic of the relation between morality and religion, Shaftesbury argues that it is possible for an atheist to be virtuous and that superstitious or false religious beliefs do more harm than having none at all. He characterizes theism as the belief that the universe is designed by a benevolent God and ordered "for the best," whereas atheists deny that there is a natural order and believe that the universe is a product of chance. Theism is the "perfection and height of virtue," since the theist is attuned to the order and harmony of the universe (1711/1999, p. 192). As moral agents, this is an order and harmony to which we ought to aspire.

VIEWS ON POLITICS, AGENCY, AND AESTHETICS

In other essays Shaftesbury, like his grandfather, champions religious tolerance and liberty of thought. Tolerance and free discussion are the basis of moral and cultural improvement. The way to disarm religious fanatics or those who are superstitious is with "ridicule," lighthearted, good-mannered humor, and tolerance, rather than with punishment and persecution. Although highly critical of the enthusiasm that results from fanaticism or superstition, Shaftesbury argues for true or reasonable enthusiasm—a state of mind that raises people beyond their ordinary capacities and enables them to feel the divine presence. Shaftesbury's conception of reasonable enthusiasm informs his views on nature, religion, morality, and aesthetics.

Some commentators, notably Stephen Darwall, find Shaftesbury's thoughts on the self, its unity and self-government, to be suggestive even though his ideas on these topics are not developed in a systematical way. Shaftesbury thinks that soliloquy is necessary both for self-government and for an agent's unity and integrity. He describes soliloquy, a kind of self-analysis, as a process whereby we are able to divide ourselves into "two parties," an idea that foreshadows Adam Smith's conception of conscience. One part is the better self, the sage, demon, or genius—an ideal of character to which each person is committed. In dividing ourselves into two, we erect the better part as the "counsellor and governor" (1711/1999, p. 77). Soliloquy enables us to step back and critically assess our desires—scrutinizing their causes and their place in the scheme of our aims and concerns. Likewise,

soliloquy aims to make us unified agents, true to our ideals of character.

Shaftesbury has been described as the first great English aesthetician. Not only does he think of moral goodness as a species of the beautiful but he also thinks that moral and aesthetic taste amount to the same thing. Thus, he says that "the science of the virtuosi and that of virtue itself become, in a manner, one and the same." The real virtuoso understands and appreciates the inner harmony and order that constitute the goodness in works of art and in people's characters. The source of Beauty and what we ultimately find beautiful is the creative, intelligent mind. Thus, he says that "the beautifying, not the beautiful is the really beautiful" (1711/1999, p. 322). When we admire order and proportion in natural objects, we are really admiring the creator, God. Shaftesbury developed a concept of disinterested pleasure to explain the kind of pleasure we experience in a true apprehension of beauty.

Shaftesbury's *Characteristicks* was influential both in England and on the Continent during the eighteenth century. It is thought that virtually every educated man in the eighteenth century was acquainted with it. While the sentimentalists, Hutcheson and Hume, kept Shaftesbury's idea that moral goodness springs from second-order affections, they detached their accounts of natural goodness from his teleological picture of the universe. Thus, Hutcheson identifies natural goodness with pleasure. There has been renewed attention to Shaftesbury's work since the 1980s, not only by traditional philosophers interested in his moral and aesthetic views but also by those interested in literary theory and gender studies.

See also Aesthetics, History of; Bayle, Pierre; Cambridge Platonists; Locke, John.

Bibliography

WORKS BY SHAFTESBURY

Select Sermons of Dr. Whichcot. London: Awnsham and John Churchill, 1698.

Characteristics of Men, Manners, Opinions, Times (1711), edited by Lawrence E. Klein. New York: Cambridge University Press, 1999.

The Life, Unpublished Letters, and Philosophical Regimen of Anthony, Earl of Shaftesbury, edited by Benjamin Rand. London: Swan Sonnenschein, 1900.

Second Characters. Cambridge, U.K.: Cambridge University Press, 1914.

WORKS ABOUT SHAFTESBURY

Darwall, Stephen. "Shaftesbury: Authority and Authorship." In *The British Moralists and the Internal "Ought": 1640–1740*. New York: Cambridge University Press, 1995.

Gill, Michael B. "Shaftesbury's Two Accounts of the Reason to Be Virtuous." *Journal of the History of Philosophy* 38 (4) (2000): 529–548.

Grean, Stanley. "Self-Interest and Public Interest in Shaftesbury's Philosophy." *Journal of the History of Philosophy* 2 (1964): 37–46.

Grean, Stanley. *Shaftesbury's Philosophy of Religion and Ethics: A Study in Enthusiasm*. Athens: Ohio University Press, 1967.

Schneewind, J. B. *The Invention of Autonomy: A History of Modern Moral Philosophy*. New York: Cambridge University Press, 1998.

Sidgwick, Henry. *Outlines of the History of Ethics*. Indianapolis: Hackett Publishing Company, 1988.

Taylor, Charles. "Moral Sentiments." In *Sources of the Self: The Making of the Modern Identity*. Cambridge, MA: Harvard University Press, 1989.

Trianosky, Gregory W. "On the Obligation to Be Virtuous: Shaftesbury and the Question, Why Be Moral?" *Journal of the History of Philosophy* 16 (3) (1978): 289–300.

Voitle, Robert. *The Third Earl of Shaftesbury, 1671–1713*. Baton Rouge: Louisiana University Press, 1984.

Charlotte R. Brown (2005)

SHAME

Shame is the painful emotion occasioned by the realization that one has fallen far below one's ideal self—the person that one wants to be. Although shame no doubt originally involves a concern with being observed by others (its link with embarrassment), such observation need no longer be a part of shame once ideals of the self have been internalized.

SHAME AND GUILT

Shame is perhaps best understood initially by contrasting it with guilt. Both are painful emotions, but the relationship of shame to morality is more complicated than is the case with guilt. Guilt is necessarily a moral emotion, since it is essentially a painful negative self-assessment with a moral basis—namely, the belief that one has done something morally wrong. One may, of course, be mistaken about the actual moral status of what one has done—one may, for example, have mistaken moral beliefs—but this is a moral mistake. Even those feelings of guilt that we classify as irrational or neurotic are typically labeled as such because we believe that the person experiencing the guilt has made a moral mistake—for example, our belief that the conduct is in fact not wrong; or our belief that the person is assuming responsibility when not really responsible; or our belief that, even if the conduct is wrong, the guilt that one feels is radically disproportionate to the nature of the wrong. So we might classify great

guilt over, say, masturbation as irrational or neurotic. We surely would not, however, label as irrational or neurotic the Nazi death camp commandant who comes to feel great guilt over his evil acts.

Although shame may also have a moral dimension, this is not necessarily the case. Shame is best understood as the painful negative self-assessment that arises when it is brought to consciousness that one's actual self is radically at odds with the ideal that one has of oneself—what Freud called one's ego-ideal. Although shame typically involves an ideal self that is at least in part constructed from social norms, these norms are frequently not moral in nature; and thus it is quite common that one may feel great shame over aspects of oneself that are morally innocent and over which one may have little control. Examples are shame over one's appearance, weight, social awkwardness, or poverty. Although such shames can sometimes prompt people to do things that are good for them (e.g., diet), they can also be so destructive of self as to be properly labeled toxic. This does not make them moral, however. Not everything that is important—even very important—is moral in nature.

MORAL SHAME

Shame becomes a moral emotion when one's ideal self, one's ego ideal, is moral in nature. If one seeks to preserve an image of oneself as a decent person with largeness and generosity of spirit, for example, then one will feel great moral shame when it is brought to consciousness that one has revealed a nature that is in fact petty, grasping, and indifferent to the hurt that one may cause others in pursuit of one's own narrow interest.

The gnawing pain of bad conscience—the *agenbite of inwit*, some medieval writers called it—may be seen as guilt over the wrong that one has done, coupled with shame over something about oneself that the wrong has revealed: the kind of person that one is, and how far this person differs from the moral person one thinks one ought to be. "Shame creeps through guilt and feels like retribution," as the novelist William Trevor puts it.

Given these important differences between guilt and moral shame—the former directed primarily toward wronging others, the latter directed to flaws of the self—one can see why the agent's healing and restorative responses to the two feelings tend to be quite different as well. Guilt typically engages such responses as apology, atonement, restitution, and even the acceptance of punishment. Moral shame imposes an even more difficult burden, however: the construction of a different and better self.

Because of its potential for moral transformation, moral shame deserves more respect than it often receives. Critics of shame tend to focus on non-moral shame, and they are quite right to stress the potentially toxic nature of some instances of non-moral shame. Some instances can even be toxic, and quite literally so, to the body. Witness the large numbers of young women who get sick and even die of eating disorders (anorexia and bulimia) because they are ashamed of their body image.

SHAME AND PUNISHMENT

Respect even for moral shame should not lead to uncritical enthusiasm, however. We should be suspicious, for example, of the trendy movement in late twentieth-century American criminal law toward shaming punishments—for example, making prisoners wear signs saying "I molest children" or dressing them in black-and-white striped uniforms and putting them on public chain gang work details. However they may be described, such practices are often merely exercises in cruel and vindictive public humiliation—something more likely to harden the heart rather than transform it in morally admirable ways. (Shaming punishments have been given their most powerful defense by Dan Kahan [1996] and their most powerful critique by Toni Massaro [1991].) As John Braithwaite (1989) has pointed out, some impressive results have been achieved with shaming punishments in small homogeneous societies that provide for rituals of reintegration. The homogeneity guarantees that one is being shamed before a group in which one values membership and whose good opinion one values, and the rituals of reintegration provide a hopeful light at the end of the tunnel. It would be a fantasy to think that modern American criminal law satisfies either condition, however.

See also Guilt; Moral Sentiments; Punishment.

Bibliography

Braithwaite, John. *Crime, Shame and Reintegration*. Cambridge, U.K.: Cambridge University Press, 1989.

Freud, Sigmund. *The Ego and the Id*. Translated by Joan Riviere. Revised and edited by James Strachey. New York: W. W. Norton, 1960.

Kahan, Dan. "What Do Alternative Sanctions Mean?" *University of Chicago Law Review* 63 (1996): 591–653.

Massaro, Toni. "The Meanings of Shame: Implications for Legal Reform." *Psychology, Public Policy and Law* 3 (1997): 645–704.

Massaro, Toni. "Shame, Culture, and American Criminal Law." *Michigan Law Review* 89 (1991): 1880–1944.

Morris, Herbert, ed. *Guilt and Shame*. Belmont, CA: Wadsworth, 1971.

Murphy, Jeffrie G. "Shame Creeps through Guilt and Feels Like Retribution." *Law and Philosophy* 18 (1999): 327–344.

Taylor, Gabriele. *Pride, Shame and Guilt: Emotions of Self Assessment*. Oxford: Clarendon Press, 1988.

Trevor, William. *Death in Summer*. New York: Viking Press, 1998.

Williams, Bernard. *Shame and Necessity*. Berkeley: University of California Press, 1993.

Jeffrie G. Murphy (2005)

SHAO YONG

(1011–1077)

Shao Yong was a Chinese philosopher, historian, and poet born in 1011 (January 21, 1012, by European dating). He was the scion of a humble but educated family that had resided in northern China, near the modern-day national capital of Beijing, for several generations. However, the border conflicts that pitted the Chinese Song dynasty (960–1279) against various hostile and encroaching non-Chinese peoples forced the Shaos into a series of moves southward toward the safer center of the empire. Thus, in 1049, Shao relocated to nearby Luoyang, the secondary imperial capital and nascent cultural hub, where he lived until his death in 1077.

Shao was influenced early by teachers—among them his father Shao Gu (986–1064) and the scholar and minor official Li Zhicai (1001–1045). But his philosophical development was surely determined much less by any one person than it was by the singular divinatory text that constitutes one of the five works included in the vaunted corpus of ancient Chinese classics—the *Book of Change* or *Yijing*. Shao was unquestionably invested in the *Book of Change*. Nonetheless, he evinced an uncommon independence of mind in how he responded to it. In contrast to others who were similarly inspired by the classic, Shao diverged from his prominent contemporaries by never writing a separate commentary specifically on the *Book of Change*. Instead, one can rightly regard the magnum opus of Shao's own scholarly output—the *Book of Supreme World-ordering Principles* (*Huangji jingshi shu*)—as entirely an expansion on the seminal premises contained in the *Book of Change* and in related writings, including the remaining four classics. Moreover, as was customary among the Chinese educated elite, Shao composed poetry. His poems were collected as *Striking the Earth at Yi River* (*Yichuan jirang ji*); this work is also one in which his cardinal philosophical ideas are exhibited. Thus, the survival of Shao's only two verifiable writings permits us to divide his thought into its early- and late-emerging components.

EARLY THOUGHT

Shao is usually accorded a position in the movement called the "Learning of the Way" (*daoxue*, a term that Europeans equate with neo-Confucianism). But he is far more noteworthy for his unique departures from the solutions arrived at by this movement. The early *daoxue* movement was chiefly preoccupied with achieving consensus on a metaphysical "first principle" that would support a cosmogony and yet also account for the assumed ethical endowment of humankind. The concept settled upon was *li* (pattern or principle), which thinkers construed as the fundamental reality underlying both physical and human nature.

Shao, however, was alone in his advocacy of the concept of number (*shu*). For him, number—and not principle—became elemental, the foundation on which the universe rested and thus the key to uncovering its secrets. Shao's faith in the regulative power of number led him to proffer that the natural processes operative in the world were number-dependent—hence, his theme of "world ordering" (*jingshi*). His conviction that number was the basis of reality also led him to advance a kind of predictive knowledge that he promoted as "before Heaven" (*xiantian*) learning. This learning, he contended, is a priori in the sense that it has always existed, even prior to the formation of the universe.

LATER THOUGHT

The final component to emerge in Shao's philosophy was a concept of methodologically reflexive observation, the chief characteristics of which were its claims to ubiquity of application and the attainment of pure objectivity and gnosis. Shao called this concept the "observation of things" (*guanwu*). Its prescribed procedure of "reverse observation" (*fanguan*) purportedly empowered the observer to know or understand any and all animate or inanimate things objectively and yet also be able to apprehend them from their own distinctly individuated and particularized standpoints. Thus, through its putative capacity to observe each and every object fully in terms of the observed object itself, the "observation of things" promised its practitioners knowledge that was truly objective, universalistic, and omniscient in its perspective.

See also Cheng Hao; Cheng Yi; Confucius; Zhang Zai; Zhou Dunyi.

Bibliography

Arrault, Alain. *Shao Yong (1012–1077), Poète et Cosmologue*. Paris: Collège de France, Institut des Hautes Études Chinoises, 2002.

Birdwhistell, Anne D. *Transition to Neo-Confucianism: Shao Yung on Knowledge and Symbols of Reality*. Stanford, CA: Stanford University Press, 1989.

Wyatt, Don J. "Historicism, Contextualization, and the Western Reception of Master Shao." *Intellectual News: Review of the International Society for Intellectual History* 4 (5) (1999): 26–36.

Wyatt, Don J. *The Recluse of Loyang: Shao Yung and the Moral Evolution of Early Sung Thought*. Honolulu: University of Hawaii Press, 1996.

Smith, Jr., Kidder; Peter K. Bol, Joseph A. Adler, and Don J. Wyatt. *Sung Dynasty Uses of the I Ching*. Princeton, NJ: Princeton University Press, 1990.

Don J. Wyatt (2005)

SHARIATI, ALI

(1933–1977)

Ali Shariati did not live to see the Islamic Revolution in Iran of 1979, but he was definitely one of its intellectual authors. Like many Iranians in the twentieth century he combined an education in the traditional religious sciences in Iran with more modern ideas from a European context—in his case Paris. His connections with the anticolonialist movement in Paris led him to argue that Islam is a basically revolutionary and liberating doctrine; Shariati did not abandon religion as many of his fellow radical Iranians did, nor did he accept the reverence for the imam or spiritual leader so prevalent in Shiʿi Islam. This set him firmly aside from Khomeini and the ideology of the Islamic Revolution itself.

He was a great borrower of ideas that he then applied in his own way. Thus while he rejected the dialectical materialism of Marxism, he did use the notion of history having a direction and a pattern—albeit one based on divine will and class struggle by individuals progressively perfecting their consciousness. Islam is a religion based on liberation, and Shariati reads the Qur'an as a book representing a community struggling permanently to achieve social justice, a fraternal society, and freedom. Shariati was not impressed with the power of imported ideologies to generate political solidarity among the people against oppressive regimes. Like his distinguished Iranian predecessor, Jalal Al-e Ahmad, he recognized the importance of politicizing Islam as an ideology of emancipation and liberation of the Iranian people. Unlike another influence on him, Frantz Fanon, Shariati approved of religion, provided it is reinterpreted appropriately.

His version of Shiʿism placed emphasis on Imam ʿAli as a revolutionary leader as well as a religious thinker. This view of Shiʿism is different from that of the religious orthodoxy, especially as it places authority in the opinion of the individual, a vindication of *ijtihad* or independent judgment rather distant from normal understandings of the notion in Islam. Here he was undoubtedly influenced by Jean-Paul Sartre and the existentialist emphasis on the importance of authentic decisions being made by free agents. Shariati argued that Islam could be vindicated as a faith if it is seen as involving autonomous choices by individuals and a genuine progressive direction in both social and personal policies.

Bibliography

WORKS BY SHARIATI

On the Sociology of Islam: Lectures. Translated by Hamid Algar. Berkeley, CA: Mizan Press, 1979.

From Where Shall We Begin? Houston: Book Distribution Press, 1980.

Marxism and Other Western Fallacies: An Islamic Critique. Translated by R. Campbell. Berkeley, CA: Mizan Press, 1980.

Oliver Leaman (2005)

SHELLEY, PERCY BYSSHE

(1792–1822)

Percy Bysshe Shelley is usually thought of as a romantic and lyric poet rather than as a philosophical one. He was, however, the author of a number of polemical prose pamphlets on politics and religion; and both his prose and his poetry reflect a coherent background of social and metaphysical theory.

In general, Shelley's beliefs are those of the radical English intelligentsia of the period immediately before and after the French Revolution, and in particular of William Godwin, who became his father-in-law. It has often been said that Shelley was really antipathetic to Godwin's atheism and determinism and that he gradually threw off Godwin's influence in favor of a more congenial Platonic transcendentalism. This view, however, seems to rest on a misunderstanding of both Godwin and Shelley.

ATTACK ON CHRISTIANITY

In *The Necessity of Atheism,* for which he was expelled from Oxford in 1811, Shelley argued, on Humean lines, that no argument for the existence of God is convincing. He developed this position in *A Refutation of Deism* (1814), a dialogue that purports to defend Christianity against deism, but which actually presents a strong case against both and in favor of atheism. In both these works, and in some of his essays (many of which were not published in his lifetime), Shelley was concerned with what he later called "that superstition which has disguised itself under the name of the system of Jesus." In the longer *Essay on Christianity,* published posthumously, he explained what he thought that system really was: an allegorical expression of the virtues of sympathy and tolerance, and of an anarchistic belief in the equality of men and in the wickedness of punishment and all other forms of coercion. Christ, Shelley claimed, had "the imagination of some sublimest and most holy poet"; he was also a reformer who, like most reformers, practiced a little mild deception by pandering to "the prejudices of his auditors." The doctrine of a personal God, in particular, is not to be taken as "philosophically true," but as "a metaphor easily understood."

THE NATURAL AND THE MORAL ORDER

Shelley explained this coupling of poetry and religion, and the view that both are essentially allegory, in *A Defence of Poetry* (1821). It is the function of both poetry and religion to provide men with a coherent view of the world that will help them to understand both themselves and their fellow men, and to provide it in a form that will kindle the imagination as well as the intellect—that is, through metaphor. There is a natural order in the universe, which science and philosophy reveal; there is also a moral order, which men themselves must impose. The metaphor of a personal God is meant to impress this twofold order on men's minds. Since this metaphor had, unfortunately, been perverted by a superstitious interpretation, Shelley himself preferred such symbols as the World Soul or the Spirit of Intellectual Beauty.

ANARCHISM

The details of the moral order itself are made clear in Shelley's political pamphlets. Shelley began to write these when, as a youth of nineteen, he set out to settle the Irish question by instructing the Irish in the fundamental principles of Godwinian anarchism. Godwin's main thesis was that social institutions, and particularly the coercive ones imposed by governments, fasten blinkers on men's minds which prevent them from seeing their fellows as they really are. The ultimate solution is a community small enough for each member to know the other members as individuals. Such intimate personal knowledge will bring understanding and sympathy, so that men will be prepared to cooperate for the common good, without the coercion of law. As Shelley put it, "no government will be wanted but that of your neighbor's opinion." Men will indeed value their neighbors' opinions, but they will not take their neighbors' opinions on trust. To do so would be useless, because even a true opinion is of little value unless one understands the grounds for holding it. It is only when men see things as they are, in all their intricate interconnections, that they will feel the right emotions and thus lead happy and virtuous lives.

POLITICAL PAMPHLETEERING

In accordance with these general principles, Shelley urged the Irish not to seek emancipation by means of violence, but to agitate for freedom of assembly, freedom of the press, and parliamentary representation as the first steps toward the ideal society. It was also in accordance with these principles that Shelley wrote his *Letter to Lord Ellenborough* (1812), in which he protested vehemently against the sentence passed on the publisher of Thomas Paine's *Age of Reason.* Both this pamphlet and the *Address on the Death of Princess Charlotte* (1817), in which he suggested that Englishmen would do better to mourn for their lost liberties than for even the most beautiful and blameless of princesses, were eloquent attacks on judicial persecution and on the suppression of free speech. In another pamphlet, *On the Punishment of Death* (left unpublished), he opposed capital punishment. In the long essay *A Philosophical View of Reform,* another of the unpublished manuscripts found among Shelley's journals, he recapitulated the common radical objections to priests, kings, and the aristocracy, and gave his support to such measures as a more democratic suffrage and a capital levy on unearned wealth.

UNITY OF THE WORLD

Shelley's writings on politics and religion provide meanings for many of the symbols and metaphors to be found in his poetry. His frequent references to life and the world around us as "a painted veil," an illusion through which we must penetrate to the reality behind (this reality being the "one" that remains when "the many change and pass"), is probably to be interpreted as a Godwinian allegory. Godwin had said that men see life as if through a

veil—the veil of their own prejudices, which are imposed by social institutions. The constant theme of Godwin's novels was that men must transcend these prejudices in order to understand and love their fellow men. Shelley's idealization of love, which has been taken as a departure from Godwin, is actually his attempt to present this Godwinian theme in a form that will kindle the imagination. It is, moreover, quite in accord with Godwin's views to say that once the veil is removed, the world will be seen as a unity—both in the sense in which science may be said to be a unity (the truth about one field of study cohering with and illuminating the truth about another), and in the sense that a true understanding of our fellow men will give rise to virtuous behavior. This seems to be what Shelley had in mind when he spoke of "the indestructible order" that it is the business of poetry to reveal. There is no need to suppose that he thought of this order as being imposed upon the world by a moral being.

THE UNIVERSAL MIND

It is true that Shelley was also influenced by Plato, Benedict de Spinoza, George Berkeley, and (in spite of his derogatory remarks about Immanuel Kant in *Peter Bell the Third*) by the newer type of idealism that was beginning to be made fashionable by Samuel Taylor Coleridge. In *On Life* he suggested that there are no distinct individual minds, but one universal mind in which all minds participate. As early as 1812 he had identified this "mass of infinite intelligence" with Deity. In this, Shelley was certainly departing from the doctrine of materialists like Baron d'Holbach; but Godwin, although he was not an idealist, was hardly a materialist either. Godwin would certainly have said that when men see things as they are, they hold the same opinions and, in a sense, think the same thoughts. Each man, seeing things from his own point of view, grasps only part of the truth. He will come nearer to grasping the whole of the truth as he comes to understand and sympathize with the minds of other men. In a sense, the truth as a whole is the property not of any one mind but of the sum of all minds. Probably Shelley himself meant little more than this.

PROMETHEUS UNBOUND

Shelley's beliefs find expression in his poetry in a way that is seen fairly clearly in *Prometheus Unbound* (1820), which can be interpreted as a Godwinian allegory. Prometheus, chained to his rock, is suffering humankind, and as the discoverer of fire, he is also knowledge and the civilizing arts. These discoveries, in themselves, are not enough to liberate man from the oppressive rule of Jupiter, which is built "on faith and fear." Prometheus is freed when, instead of cursing his oppressor, he begins to pity and so to understand him. This reflects the favorite Godwinian theme that the oppressor, no less than the oppressed, is the victim of social institutions. A better order is possible only when men come to understand this fact and substitute mutual sympathy for recrimination and punishment. It is also necessary to understand the secrets of Demogorgon, who personifies the natural forces that control the universe, and to cooperate with the Hours, who, with their chariots, personify Godwin's conviction of the inevitability of gradualism.

See also Anarchism; Atheism; Berkeley, George; Coleridge, Samuel Taylor; Deism; Godwin, William; Holbach, Paul-Henri Thiry, Baron d'; Kant, Immanuel; Paine, Thomas; Plato; Spinoza, Benedict (Baruch) de; Wollstonecraft, Mary.

Bibliography

WORKS BY SHELLEY

The definitive edition of Shelley's writings is *The Complete Works of Shelley,* edited by R. Ingpen and W. E. Peck, 10 vols. (London and New York, 1926–1930). A useful collection is *Shelley's Prose, or The Trumpet of a Prophecy,* edited by D. L. Clark (Albuquerque: University of New Mexico Press, 1954).

WORKS ON SHELLEY

A detailed biography is N. I. White, *Shelley,* 2 vols. (New York: Knopf, 1940). Works dealing especially with his thought are C. Baker, *Shelley's Major Poetry* (Princeton, NJ: Princeton University Press, 1948); H. N. Brailsford, *Shelley, Godwin, and Their Circle* (New York: Holt, 1913); P. Butter, *Shelley's Idols of the Cave* (Edinburgh: Edinburgh University Press, 1954); A. M. D. Hughes, *The Nascent Mind of Shelley* (Oxford: Clarendon Press, 1947); D. King-Hele, *Shelley: His Thought and Work* (London: Macmillan, 1960); A. Sen, *Studies in Shelley* (Calcutta: University of Calcutta, 1936); and M. T. Wilson, *Shelley's Later Poetry* (New York: Columbia University Press, 1959).

D. H. Monro (1967)

SHEPHERD, MARY

(1777–1847)

Mary Shepherd was born in Scotland at her family's estate on December 31, 1777, the second daughter of Neil Primrose, Earl of Rosebery; she died in London on January 7, 1847. Relatively few details of her life and education are available. She married an English barrister, Henry Shepherd, in 1808. She published at least two works in philosophy, *An Essay upon the Relation of Cause and Effect*

(1824), and *Essays on the Perception of an External Universe and other Subjects Connected with the Doctrine of Causation* (1827). A third work, originally published anonymously in 1819, *Enquiry respecting the Relation of Cause and Effect*, has been credited to her, but it differs so significantly from her other work, both in style and content, as to make this attribution dubious. She was as well a participant in an exchange of views with a contemporary, John Fearn, which appeared in various venues.

Shepherd's work reflects the continued interest in the first quarter of the nineteenth century in developing alternative arguments to those of Hume, conceived largely skeptically. Her first work establishes the line of argument that was to direct her work. In it, she seeks to refute Hume's position on causality by arguing that Hume is mistaken in holding that we lack an intuitive understanding that events have causes. Shepherd reads Hume as holding that we cannot be intuitively certain that everything that begins to exist has a cause, and subjects to criticism the contained concept of a causeless beginning-to-be of some existence. Her argument is that this beginning is itself an action and hence must be a state of something that, by hypothesis, does not as yet exist until it has begun to be. Hence, she claims, the basic assumption of Hume's account is contradictory. Shepherd offers a realist account of cause as the productive principle of effects, themselves not subsequent to causes, but rather coexistent with the productive object. She uses her realist understanding of causation to criticize not only Hume, but also her own contemporaries, Thomas Brown and William Lawrence.

Shepherd's second work, *Essays on the Perception of an External World*, was originally intended as an appendix to her first work and consists primarily, although not exclusively, of an application of her ideas about causation to the question of the existence of an external world. By far the largest part is directed to providing an alternative answer to Hume's question about the sources of our idea of a continuous external existence. Appended are a series of essays about Berkeley, Reid, Stewart, Hume, and what Shepherd terms in the title of her work "various modern atheists." Shepherd argues, against Hume, that the possibility of causal reasoning, as demonstrated in her first book, makes such reasoning available to substantiate the existence of a continuously existing independent world. She feels it necessary, however, to give a different solution from that of Reid. This is because she thinks Reid failed to appreciate the importance of Berkeley's claim that an idea can only be like another idea. Shepherd takes this to mean that Reid is wrong to suppose that we can give content to our ideas of a mind-independent world. Thanks to the possibility of causal reasoning, however, we are able to assert the existence of causes responsible for our ideas. In particular, because our ideas change, there must be causes for these changes, independent of our ever-present mind. The variety we experience must be due to causes other than ourselves, whose nature, while unknown, must be, she thinks proportional to their effects.

Shepherd develops and clarifies these ideas further in her exchange of views with John Fearn, a retired naval officer and philosophical aficionado. This exchange is unusual as well as interesting, presenting one of the first occasions where a woman's ideas are attacked in print, and illustrating some of the different venues available to ordinary practitioners for publishing philosophy in the early nineteenth century. The first two parts of the exchange appear in 1828 in a volume loosely related to the clergyman, Samuel Parr, called *Parriana*, apparently supplied by Fearn to its compiler, Ernest Barker, and included by him, despite the lack of relevance to Parr. These consist of a four-page paper, critical of Fearn by Shepherd, apparently sent to him privately, and a longer defense of his views against Shepherd by Fearn. Shepherd was sufficiently concerned by this unauthorized use of her work that she published a rebuttal, "Lady Mary Shepherd's Metaphysics" in a well-known literary journal, *Fraser's Magazine*, in July 1832.

The exchange focuses on a disagreement over the idea of extension. It is Fearn's view that the content of the idea of extension is determined by our perception of it. There can be no extended external material cause of such an idea. The only possible cause consists of the energies of an extended mind, analogous to our own. Shepherd maintains that Fearn has not adequately distinguished the idea of extension from its unknown cause. On the one hand, Shepherd holds that extension can only apply to objects considered as causes, for it is as causes that they take up space and move. Ideas, on the other hand, neither move nor take up space, or we would be left with the ridiculous position that the idea of a fat man is itself fat. Shepherd, in defense of this claim, gives a fresh defense of her causal realism. While it is true that the mind perceives internal changes to its own states, it nevertheless reasons to the existence of external unperceived causes of these changes.

See also Berkeley, George; Brown, Thomas; Causation: Metaphysical Issues; Epistemology; Hume, David; Metaphysics; Reid, Thomas; Stewart, Dugald; Women in the History of Philosophy.

Bibliography

WORKS BY SHEPHERD

An Essay upon the Relation of Cause and Effect, controverting the Doctrine of Mr. Hume, concerning the Nature of the Relation; with Observations upon the Opinions of Dr. Brown and Mr. Lawrence, Connected with the Same Subject. London: printed for T. Hookham, Old Bond Street, 1824.

Essays on the Perception of an External Universe, and Other Subjects connected with the Doctrine of Causation. London: John Hatchard and Son, 1827.

"Observations of Lady Mary Shepherd on the 'First Lines of the Human Mind.'" In *Parriana: Or Notices of the Rev. Samuel Parr, L.L.D.*. Collected from various sources and in part written by E. H. Barker, esq., 624–627. 1828.

"Lady Mary Shepherd's Metaphysics." *Fraser's Magazine for Town and Country* 5 (30) (1832): 697–708.

Modern edition of all of the above: *The Philosophical Works of Lady Mary Shepherd*. Vols. 1–2, edited by Jennifer McRobert. Bristol, U.K.: Thoemmes Press, 2000.

WORKS ABOUT SHEPHERD

Atherton, Margaret. "Lady Mary Shepherd's Case against George Berkeley." *British Journal for the History of Philosophy* 4 (2) (1996): 347–366.

Atherton, Margaret. "Mary Shepherd." In *Dictionary of NineteenthCentury British Philosophers*. Bristol, U.K.: Thoemmes Press, 2002.

McRobert, Jennifer. Introduction to *The Philosophical Works of Lady Mary Shepherd*. Vols. 1–2. Bristol, U.K.: Thoemmes Press, 2000.

Margaret Atherton (2005)

SHESTOV, LEV ISAAKOVICH

(1866–1938)

Lev Isaakovich Shestov, the Russian philosopher and religious thinker, was born in Kiev. His real name was Lev Isaakovich Schwarzmann. Shestov studied law at Moscow University but never practiced it. He lived in St. Petersburg from the late 1890s until he migrated to Berlin in 1922; he later settled in Paris. He gave occasional lectures in Berlin, Paris, and Amsterdam and made two lecture tours in Palestine, but he held no regular academic position.

Shestov called William Shakespeare his "first teacher of philosophy"; in his later years he interpreted Hamlet's enigmatic "the time is out of joint" as a profound existential truth. Shestov apparently turned to philosophy relatively late, perhaps in 1895, when he reportedly underwent a spiritual crisis. He himself never referred to such a crisis; in general, his works are less confessional and autobiographical than those of most existential thinkers. However, they are neither impersonal nor unimpassioned; intensity and engagement (in a religious and moral rather than a political sense) are hallmarks of his thought.

Shestov was perhaps most strongly influenced by Blaise Pascal, Fëdor Dostoevsky, and Friedrich Nietzsche. He discovered Søren Kierkegaard quite late and found his position highly congenial, but he had worked out his own existentialist position independently of Kierkegaard. Shestov's philosophical works are written in an aphoristic, ironic, questioning style reminiscent of Pascal's *Pensées* and Nietzsche's *Beyond Good and Evil.* Shestov believed, with Kierkegaard, that subjective truth borders on paradox. "People seem shocked," he once wrote, "when I enunciate two contradictory propositions simultaneously.… But the difference between them and me is that I speak frankly of my contradictions while they prefer to dissimulate theirs, even to themselves.… They seem to think of contradictions as the *pudenda* of the human spirit" (quoted in de Schloezer, "Un penseur russe …," pp. 89–90).

Shestov was not a systematic thinker. He attacked the views of others, sometimes massively; but he was content to suggest or sketch his own position. His writings focus positively on the question of religion and morality or religiously based morality; negatively on the critique of theoretical and practical rationalism. Among the rationalists whom he attacked by name are Parmenides, Plato, Aristotle, Plotinus, Benedict de Spinoza, Immanuel Kant, G. W. F. Hegel, and Edmund Husserl.

The basic either/or of Shestov's thought is suggested by the title of his major work in philosophy of religion: *Afiny i Ierusalim* (Athens and Jerusalem). Athens is the home of reason, of a philosophical rationalism that insists on a neat and knowable cosmos ruled by eternal and unalterable laws. Jerusalem is the home of faith, of an existential irrationalism that stresses contingency, arbitrariness, mystery, and pure possibility. For God "all things are possible," even what René Descartes had called a logical absurdity, that is, causing what has in fact happened not to have happened.

Sometimes Shestov's attack on reason took the form of questioning reason's theoretical competence. Thus, he complained that theorists of biological and cosmic evolution, with their loose talk about "millions and billions of years" and about "eternal nature," were perpetrating a "monstrous absurdity."

More frequently Shestov made the rather different claim that rational knowledge neglects what is essential—

the individual, contingent, incomprehensible, and mysterious. "However much we may have attained in science," he wrote, "we must remember that *science cannot give us truth*…. For truth lies in the singular, uncontrollable, incomprehensible, … and 'fortuitous'" (*In Job's Balances,* p. 193). "We live," Shestov declared, "surrounded by an infinite multitude of mysteries" (*Afiny i Ierusalim,* p. 25).

Most frequently Shestov attacked the moral consequences of theoretical reason, its erosion and subversion of human values. Reason exhibits necessity and imposes nonfreedom. Faith assumes contingency and makes freedom possible. Rationalists recognize an eternal structure of being, a system of necessary laws that antedates any possible cosmic lawgiver. The necessity of such laws requires obedience. What is nonnecessary, whether contingent or arbitrary, admits of free decision and creativity. Shestov repudiated all obedience to necessity in the sense of acceptance of necessary evil, injustice, and inhumanity. There are scales, he declared, upon which human suffering weighs heavier than all the necessities of theoretical reason; such are "Job's balances."

In particular, Shestov rejected the Greek view, which he traced back to Anaximander, that coming to be (*genesis*) is a kind of affront to the gods, a cosmic hubris, justly rewarded by the punishment of passing away (*phthora*). He called this the "dreadful law which inseparably links death to birth." "In man's very existence," Shestov added, "thought has discovered something improper, a defect, a sickness, or sin, and … has demanded that this be overcome at its root [by] a renunciation of existence" (*Kirgegard ekzistentsial'naia filosofiia,* p. 8).

In such passages Shestov may appear to have confused natural (descriptive) laws with moral (prescriptive) ones. However, his point could be made in terms of such a distinction; descriptive laws, insofar as the regularities which they describe are universal and necessary and not merely local or statistical, demand unconditional acceptance and thus in a sense function prescriptively.

In any case, Shestov wished to assert that rationalists, in absolutizing theoretical truth, inevitably relativize human life. In yielding to "self-evidence," they accept the "horrors of human existence" as something necessary and legitimate. Shestov, in contrast, was quite prepared to relativize theoretical truth if that was the price to be paid for absolutizing moral and religious values and thus "redeeming" the existing individual.

The Nietzschean strain in Shestov's thought appears most clearly in his denial of the validity of universal norms. Such norms function to limit and repress creativity. "The fundamental property of life," he wrote, "is daring; all life is creative daring and thus an eternal mystery, irreducible to anything finished or intelligible" (*In Job's Balances,* p. 158). Under the tyranny of ethical rationalism (a part of the general tyranny of reason, which develops naturally out of the initial autonomy of reason), we come to fear chaos because it is a loss of order. But "chaos is not a limited possibility; it is an unlimited opportunity" (ibid., p. 226).

For Shestov the decisive either/or—reason and necessity or faith and freedom—is not a choice, as rationalists would claim, between sanity and insanity. It is a choice between two kinds of madness (the distinction is reminiscent of Kierkegaard's distinction between "objective" and "subjective" madness). The first kind of madness is that of theoretical reason, which takes as ultimate, eternal, and universally obligatory those objective truths which rationalize and legitimize the "horrors of human existence." The second kind of madness is the Kierkegaardian leap of faith which ventures to take up the struggle against rationalized and legitimized horror at the point where such struggle is "self-evidently" doomed to defeat. Between these two kinds of madness, Shestov's own choice is clear and final.

See also Anaximander; Aristotle; Descartes, René; Dostoevsky, Fyodor Mikhailovich; Existentialism; Hegel, Georg Wilhelm Friedrich; Husserl, Edmund; Kant, Immanuel; Kierkegaard, Søren Aabye; Nietzsche, Friedrich; Parmenides of Elea; Pascal, Blaise; Plato; Plotinus; Rationalism; Russian Philosophy; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY SHESTOV

Dostoevskii i Nietsshe: Filosofiia tragedii. St. Petersburg, 1903. Translated by R. von Walter as *Dostojewski und Nietzsche: Philosophie der Tragödie.* Cologne, 1924. Also translated into French as *Dostoevski et Nietzsche (La Philosophie de la tragédie).* Paris, 1926.

Apofeoz bespochvennosti: Opyt adogmaticheskogo myshleniia (The apotheosis of groundlessness: an essay in undogmatic thought). St. Petersburg, 1905. Translated by S. S. Koteliansky as *All Things Are Possible.* New York: R.M. McBride, 1920. With a foreword by D. H. Lawrence.

Na Vesakh Iova. Paris: Sovremennyia zapiski, 1929. Translated by A. Coventry and C. A. Macartney as *In Job's Balances.* London: J.M. Dent, 1932.

Kirkegard i ekzistentsial'naia filosofiia (Glas vopiiushchego v pustyne). Paris: Dom knigi i Sovremennyia zapiski, 1939. Translated by T. Rageot and Boris de Schloezer from the manuscript as *Kierkegaard et la philosophie existentielle (Vox Clamantis in Deserto).* Paris: J. Vrin, 1936; reprinted 1948.

English translation by Elinor Hewitt as *Kierkegaard and the Existential Philosophy.* Athens: Ohio University Press, 1966.

Afiny i Ierusalim. Paris: YMCA Press, 1951. Translated from the manuscript by Boris de Schloezer as *Athènes et Jerusalem: Essai de philosophie religieuse.* Paris: J. Vrin, 1938.

WORKS ON SHESTOV

Schloezer, Boris de. "Un penseur russe: Léon Chestov." *Mercure de France* 159 (1922): 82–115.

Zenkovsky, V. V. *Istoriia russkoi filosofii.* 2 vols. Paris: YMCA Press, 1948–1950. Translated by G. L. Kline as *A History of Russian Philosophy.* 2 vols., 780–791. New York: Columbia University Press, 1953.

George L. Kline (1967)

SHESTOV, LEV ISAAKOVICH [ADDENDUM]

Shestov has become the object of academic philosophical attention only since 1968. After the 1917 Russian Revolution Shestov became a significant voice in European philosophical existentialism, in his later life engaging with Blaise Pascal and Søren Kierkegaard, actively influencing the thought of Albert Camus, and corresponding with Martin Buber. Some of these philosophical relationships have received concentrated, though not exhaustive, critical attention (Maia Neto 1995). In addition, Shestov corresponded with and wrote an article on Edmund Husserl, which is the focus of one critical article.

Because of the Soviet ban on research and publication relating to Shestov, scholars inevitably found it difficult to define and establish Shestov as a philosopher. To begin with there was very little criticism outside the Paris émigré community. The two-volume biography on Shestov written by his daughter, Natalie Baranova-Shestova (1983), drew attention to the man and his work. Since the end of the Soviet Union Shestov has won renewed consideration among Russian philosophers.

Existentialist aspects of Shestov's thought have generally garnered the most critical attention and have generated other critical approaches. Some existentialist commentaries focus on the experience of suffering, isolation, and tragedy while others concentrate on the aspect of the absurd. Shestov has received attention as a religious thinker particularly in two contexts. First, scholars have viewed him in the context of the "Russian religious renaissance," a group of Russian religious philosophers of the early twentieth century who brought a personalist, antidogmatic and antirational approach to the question of religious experience and faith. Second, scholars have seen him as a major modern Jewish thinker. In the late twentieth century, philosophical research focused on two contrasting aspects of Shestov's thought, the paradoxical but invigorating interaction between skepticism and religious faith. His philosophy has been viewed together with that of Pascal and Kierkegaard, as part of the tradition of skeptical thought, Pyrrhonism, that goes beyond pure skepticism to employ reasoned doubt in a positive role within categories of faith.

Bibliography

WORKS BY LEV SHESTOV

Sobranie sochinenii. 6 vols. St. Petersburg: Izd. Shipovnik, 1911.

Athens and Jerusalem. Translated with an introduction by Bernard Martin. Athens: Ohio University Press, 1966. Originally published as *Afiny I Ierusalim.* Parizh: YMCA Press, 1951.

A Shestov Anthology, edited with an introduction by Bernard Martin. Athens: Ohio University Press, 1971.

WORKS ABOUT LEV SHESTOV

Baranova-Shestova, Natalie. *Zhizn' L'va Shestova: po perepiske i vospominaniiam sovremennikov* (The life of Lev Shestov from his correspondence and reminiscences of contemporaries). 2 vols. Paris: Presse Libre, 1983.

Clowes, Edith. "Philosophy as Tragedy." In *Fiction's Overcoat: Russian Literary Culture and the Question of Philosophy*. Ithaca: Cornell University Press, 2004, 130–154.

Kline, George L. *Religious and Anti-Religious Thought in Russia*. Chicago: University of Chicago Press, 1968.

Maia Neto, Jose R. *The Christianization of Pyrrhonism: Skepticism and Faith in Pascal, Kierkegaard, and Shestov*. Dordrecht: Kluwer Academic, 1995.

Martin, Bernard, ed. *Great Twentieth Century Jewish Philosophers: Shestov, Rosenzweig, Buber, with Selections from Their Writings*. New York: Macmillan, 1970.

Zakydalsky, Taras D. "Lev Shestov and the Revival of Religious Thought in Russia." In *Russian Thought after Communism: The Recovery of a Philosophical Heritage*, edited by James P. Scanlan, 153–164. Armonk, NY: M. E. Sharpe, 1994.

Edith Clowes (2005)

SHINRAN
(1173–1262)

Shinran, born Hino Arinori, is the foremost proponent of Japanese Pure Land Buddhism and is widely regarded as the founder of Jôdo-Shinshû, more commonly known outside of Japan as Shin Buddhism. Pure Land Buddhism has the largest following in East Asia (China, Korea, and Japan), and the Shin sect is the largest sect of Japanese Buddhism. As a development of Mahayana Buddhism,

the core of Shinran's thought is based on the twofold truth:

Conventional truth	Highest truth
Form	Emptiness
Distinctions	No distinctions
Words	Beyond words
suffering	liberation
samsara	nirvana
defiled world	Pure Land
blind passion	boundless compassion
self-power	other-power
foolish being	Amida Buddha
Namu	Amida Butsu

TWOFOLD TRUTH

These truths are twofold because they are like the two sides of the same coin. There is an aspect of truth defined conceptually by the discursive intellect, and there is a truth beyond words, beyond the grasp of the discursive intellect. In this view, all conceptual reality is nothing more than agreed on convention, hence the term *conventional truth*. When the mind is emptied of all preconceptions, the truth can be grasped for the first time with one's whole being. This is the highest truth, emptied of the concepts that act like an intervening smoke screen between subject and object. When the conceptual smoke screen is removed, the separation between subject and object also disappears. Paradoxically, this merging of subject and object does not mean the obliteration of perception. Rather, perception becomes more fluid, dynamic, and vivid. For example, when one is viewing a flower and is caught up in trying to determine its genus, species, and variety, one fails to see the vivid dynamism of the beautiful flower unfolding before one. However, when one lets go of one's obsession with grasping the flower taxonomically or conceptually, suddenly one feels that the flower is closer, more intimate, vivid, and fluid in its evanescence.

WORDS AND BEYOND WORDS

In Buddhism the problem does not lie with the categories or words themselves, such as *flower*, *peony*, and so on. Rather, it is the mind that becomes obsessed or attached to fixed conceptions of reality that causes one to become lost or separated from the dynamic flow of reality. Suffering, defiled perception, and blinded passion all result from this fixation. Conversely, words, properly used, can convey reality beyond words. They are like the words of a love poem. Although the individual words of a love poem cannot capture love itself, a beautiful love poem can nevertheless convey the sensibility of love. The words are no mere signs; they are vessels of a higher truth.

THE NAME OF AMIDA BUDDHA

In Shinran's Shin thought, the twofold truth is expressed through the Name of Amida Buddha, the Buddha of Infinite Light and Immeasurable Life. The practitioner of Shin invokes or chants the name Namu Amida Butsu. It originates in India and comes from the Sanskrit, Namas Amitâbha Buddha, meaning, "I entrust myself to the awakening of infinite light." When the practitioner, caught in the net of fixed ideas, is illuminated by the dynamic flow of reality, he or she is released from his or her blind passions and awakens to the light of emptiness, or the boundless oneness of reality.

The highest truth of reality is formless, without shape, definition, color, or scent. However, the experience of release from the ego-bonds of fixation and blind passion is one of illumination or light. This is neither merely symbolic nor merely material or physical. Similar to the experience of being relieved of a heavy mental burden, one's conscious awareness and field of vision become clearer, lighter, and more responsive.

According to Shinran the consciousness of an ego self-enclosed in its own solipsistic world works under the delusion of self-power (Japanese: *jiriki*), as though it sustains itself completely unrelated to the world around it. When the bonds of this delusion are exposed and illuminated by the dynamic unfolding of emptiness/oneness, the self awakens to the working of other-power (Japanese: *tariki*), so-called because it is other than (the delusory) ego.

However, one does not and cannot abandon the foolish delusions of the ego; as long as one lives in this limited body and mind, one will continue to suffer the ego's foolishness. Furthermore, it is this very foolishness, when one recognizes it, that connects the practitioner to his or her deepest humanity and that of others. For it is in the suffering of blind passion and foolishness that one finds the deepest bonds of humanity, and ultimately, with all sentient beings. In the illumination of Amida Buddha, the blind passion of the foolish being becomes the gateway to wisdom and compassion. Thus, Namu represents the foolish being who, in his very foolishness, is illuminated by Amida Butsu, infinite light and boundless compassion. The saying of the name Namu Amida Butsu embodies the realization of the oneness of foolish being and boundless compassion. Without the Namu, Amida Buddha is merely a cold abstraction; only when the practitioner

engages the vivid flow of reality by allowing his or her blindness to be illuminated does the reality of Amida Buddha come to life. For this reason, the real name of Amida Buddha is said to be Namu Amida Butsu.

SHINRAN'S SOCIAL VISION

Shinran's philosophical thought translated itself into an egalitarian social vision. According to him, no human being, even religious masters, were completely enlightened. Indeed, those who had engaged in intensive religious practice were considered particularly susceptible to the hubris of religious attainment. Shinran abandoned the monastic life, married openly, had four children, and lived among the farmers in outlying districts. Nevertheless, he and his wife, Esshinni, continued to wear religious robes and ministered to peasants and farmers until Shinran was about sixty. He describes himself as "neither monk nor layman" (Hirota, 289 [translation adapted]) and states, "I do not have even a single disciple"(Hirota, 664 [translation adapted]) since the power of compassion comes from Amida as the deepest reality of the self, and not from the finite human being Shinran.

He spent his final thirty years living in his brother's house, writing voluminously on his understanding of the wondrous working of Amida's boundless compassion, mythologically expressed as the working of Amida's Primal Vow. This is a way of expressing the relentless flow of reality that sooner or later breaks down and dissolves the brittle facade of self-power ego.

See also Buddhism; Buddhist Epistemology; Japanese Philosophy; Social and Political Philosophy; Truth.

Bibliography

Dobbins, James C. *Jōdo Shinshū: Shin Buddhism in Medieval Japan*. Honolulu: University of Hawaii Press, 2002.

Hirota. Dennis, trans. *The Collected Works of Shinran*., 2 vols. Kyoto, Japan: Jōdo Shinsū Hongwanji-Ha, 1997.

Keel, Hee-Sung. *Understanding Shinran: A Dialogical Approach*. Fremont, CA: Asian Humanities Press, 1995.

Unno, Taitetsu. *River of Fire, River of Water: An Introduction to the Pure Land Tradition of Shin Buddhism*. New York: Doubleday, 1998.

Mark T. Unno (2005)

SHĪRĀZĪ, SADR AL-DĪN

See *Mullā Ṣadrā*

SHOAH

See *Holocaust*

SHOEMAKER, SYDNEY

(1931–)

Sydney Shoemaker is the Susan Linn Sage Professor of Philosophy Emeritus at Cornell University. Before joining the Philosophy Department at Cornell in 1961, he taught at Ohio State University and he held the Santayana Fellowship at Harvard University. He also delivered the John Locke Lectures at Oxford University (1972) on "Mind and Behavior" and the Royce Lectures at Brown University (1993) on "Self-Knowledge" and "Inner Sense." He has pioneered work in a variety of areas in metaphysics and the philosophy of mind, particularly on the nature of mind, the nature of the self and of self-knowledge, and the nature of properties. Some of the most important of his contributions in these areas are charted in this entry.

Shoemaker's work on the topic of the self and self-knowledge is informed by a rejection of the Cartesian notion of an immaterial self and the accompanying view that self-knowledge involves a kind of "inner observation" of the contents of one's mind that is perception-like in certain characteristic ways. The nature of the self and self-knowledge forms the subject matter of his seminal *Self-Knowledge and Self-Identity* (1963). In this work Shoemaker conducts a sustained attack on the Cartesian view that the unity of the self, or personal identity, is due to or involves an immaterial unity. Shoemaker argues, against this, that personal identity involves both physical factors concerning persons' bodies and psychological factors concerning their memories and that although the primary criterion for such identity is bodily identity, a memory criterion is also applicable. His arguments make use of a distinctive methodological strategy that has come to be known as "the method of cases" (Johnston 1987), involving the use of thought experiments to determine answers to questions about personal identity (a method that John Locke [1985] used in his discussion of personal identity).

Shoemaker's examples and the style of argumentation in this work have been highly influential in discussions of personal identity. His views in this area are further developed in later work, such as his "Persons and Their Pasts" (1970), *Personal Identity* (1984; with Richard Swinburne), *The First-Person Perspective and Other Essays* (1996), and the Royce Lectures, where he revisits another

important theme found in *Self-Knowledge and Self-Identity*: the Cartesian, "inner sense" view of self-knowledge. In these lectures he argues that if self-knowledge were perception-like, and the object of such knowledge were the self, it would be possible to err in one's attempt to identify oneself, just as it is possible to err in one's attempt to identify the objects of ordinary perception. However, he claims, it is not possible to misidentify oneself in this way. He also argues that the "inner sense" model of self-knowledge, being a perceptual one, requires commitment to two conditions that are essential to a "broad" perceptual model, a causal condition and an independence condition, but that knowledge of one's own mental states does not meet these conditions. His own view is that self-knowledge is not based on evidence of any kind, whether this be from "outer" behavioral facts or from "inner" ones, as the perceptual model encourages one to suppose.

Shoemaker's arguments involve an appeal to a particular view of the nature of mind known as functionalism, a view that he has developed and defended extensively in several works, notably in "Functionalism and Qualia" (1975), "Some Varieties of Functionalism" (1981a), "Absent Qualia Are Impossible" (1981b), and "The Inverted Spectrum" (1981c). Functionalism in the philosophy of mind is, broadly construed, the doctrine that mental-state types or kinds can be exhaustively characterized and uniquely individuated by their functional properties—by the relations that they are apt to bear to certain characteristic kinds of physical stimuli, other mental states, and behavioral responses. Shoemaker (1981b, 1981c) defends this doctrine against two major objections, known as the inverted qualia and absent qualia objections. The inverted qualia objection supposes that two states—say, perceptual experiences—might vary in their visual qualia (one being reddish, perhaps, while the other is greenish) yet remain invariant with respect to their functional roles. Shoemaker agrees but argues that this possibility is compatible with the truth of functionalism. The absent qualia objection goes further and supposes that two states could be functionally identical yet differ to the extent that one has qualitative content while the other lacks it altogether. Shoemaker concedes that if this were a genuine possibility, it would show that functionalism is false, but it is not a genuine possibility.

A third area in which Shoemaker has done pioneering work, connected with his functionalist view of the nature of mind, concerns the nature of properties. In the case of mental states Shoemaker argues that their nature is causal-functional. In the case of properties Shoemaker is an advocate of what is known as the causal theory of properties (Armstrong 2000), a view championed in his influential "Causality and Properties" (1980). According to it, properties have causal powers essentially, rather than accidentally, in that it is in the nature of properties to bestow causal capacities on their instances or exemplifications. So, for example, it is in the nature of the property, pain, to confer on its instances, individual pains, the capacity to cause their subjects to believe that they are in pain, to wince, and so on. The view contrasts with a "categoricalist" one (Armstrong 2000), which takes properties to be contingently, rather than essentially, related to the capacities they bestow on their instances.

Although some have construed Shoemaker as holding the view that properties just are dispositions (rather than the weaker view that properties are essentially dispositional), which is a controversial and difficult view to defend, this is a mistake. Shoemaker argues that, strictly speaking, the dispositional/nondispositional distinction only applies to linguistic items, specifically, to predicates such as *soluble*, *fragile*, *round*, and so on. Some predicates (e.g., *soluble* or *fragile*) are dispositional whereas others (e.g., *round* or *red* are not. But all properties bestow causal capacities on their instances, for it is in their nature to do so. So, for example, the property round bestows on its instances in, say, marbles, the capacity to roll into round holes, but not triangular ones. Shoemaker argues that the identity conditions of properties can be given in terms of such capacities, that is, that properties are identical if and only if they bestow on their instances the same causal capacities or powers and that it follows that the relations that hold between properties are necessary rather than contingent, so that, if laws involve relations between properties, such laws are necessary rather than contingent.

See also Cartesianism; Philosophy of Mind; Qualia.

Bibliography

Shoemaker, Sydney, and Richard Swinburne. *Personal Identity*. Oxford, U.K.: Blackwell, 1984.

WORKS BY SHOEMAKER

Self-Knowledge and Self-Identity. Ithaca, NY: Cornell University Press, 1963.

"Persons and Their Pasts." *American Philosophical Quarterly* 7 (4) (1970): 269–285.

"Functionalism and Qualia." *Philosophical Studies* 27 (5) (1975): 291–315.

"Causality and Properties." In *Time and Cause: Essays Presented to Richard Taylor*, edited by Peter van Inwagen, 109–135. Dordrecht, Netherlands: Reidel, 1980.

"Some Varieties of Functionalism." *Philosophical Topics* 12 (1981a): 83–118.

"Absent Qualia Are Impossible." *Philosophical Review* 90 (1981b): 581–599.

"The Inverted Spectrum." *Journal of Philosophy* 74 (1981c): 357–381.

Identity, Cause, and Mind: Philosophical Essays. New York: Cambridge University Press, 1984.

"Self-Knowledge and Inner Sense." *Philosophy and Phenomenological Research* 54 (1994): 249–314.

The First-Person Perspective and Other Essays. New York: Cambridge University Press, 1996.

WORKS ABOUT SHOEMAKER

Armstrong, David. "The Causal Theory of Properties: Properties According to Shoemaker, Ellis, and Others." *Metaphysica* 1 (2000): 5–20.

Johnston, Mark. "Human Beings." *Journal of Philosophy* 84 (1987): 59–83.

Locke, John. *An Essay concerning Human Understanding*, edited by Peter H. Nidditch. Oxford, U.K.: Clarendon Press, 1985.

Cynthia Macdonald (2005)

SHPET, GUSTAV GUSTAVOVICH

(1879–1937)

In his most important phenomenological work, *Iavlenie i smysl* (Appearance and sense, 1914), Gustav Shpet took up Edmund Husserl's idea of pure phenomenology and developed it in the direction of a "phenomenology of hermeneutical reason." In this theoretical framework he formulated, between 1914 and 1918, hermeneutic and semiotic problems, which in the 1920s he elaborated more specifically within the fields of philosophy of language and theory of art. In doing so, he was combining Husserl's conceptions with ideas from other philosophical movements, particularly Wilhelm Dilthey's hermeneutics and Wilhelm von Humboldt's philosophy of language.

Shpet's reception and transformation of phenomenology must be seen in the context of Russian intellectual and cultural life during the first two decades of the twentieth century. The Platonic "Moscow Metaphysical School" (which included Vladimir Solov'ëv and Sergei Trubetskoi) provided the intellectual atmosphere in which Shpet's turn to Husserl's phenomonology took place. His ideas on theories of language and signs are close to those of contemporary Russian formalism. His phenomenological and structural theories influcenced Prague structuralism through the "Moscow Linguistic Circle," and his work is seen as a precursor to Soviet semiotics.

SHPET'S LIFE

Gustav Shpet was born in 1879 in Kiev. He studied there at Vladimir University from 1901 to 1905, completing his studies with a monograph entitled *Problema prichinnosti u Iuma i Kanta* (The problem of causality in Hume and Kant). In 1907 he moved to Moscow, and taught at Moscow University from 1910. During a stay in Göttingen (1912–1913), where he studied with Husserl, he turned to Husserl's transcendental phenomenology. His first phenomenological publication, *Iavlenie i smysl* (Appearance and meaning, 1914) marked the beginning of a productive reception of Husserl's phenomenology in Russia. In 1916 he defended his master's thesis *Istoriia kak problema logiki* (History as a problem of logic, Part I). In 1918 he finished *Germenevtika i eë problemy* (Hermeneutics and its problems), in which he discussed the problems of hermeneutics as they have been developed throughout history from antiquity (especially in Origen and Augustine) to modern times, thereby at the same time elaborating the basic outline of his "hermeneutical philosophy"—a philosophy that is caught in the field of tension exerted, on the one side, by Husserl's "Phenomenology of Reason" and, on the other, by Dilthey's "Philosophy of Life."

After the Revolution of 1917, Shpet was active in various fields of cultural and intellectual life. He received a professorship of philosophy at Moscow University. In 1920 he joined the "Moscow Linguistic Circle" (MLK), a center of Russian formalism, and in 1921 he was appointed director of the Institute for Scientific Philosophy, a new research institute at Moscow University. Expelled from the university in 1923 for political reasons, he concentrated his activities on the State Academy of the Arts (GAKhN), where he served as vice president until 1929, and where he temporarily chaired the Department of Philosophy. His most important contributions to the theory of art and language are his *Ėsteticheskie fragmenty* (Aesthetic fragments), published in 1922 and 1923 in Petrograd, and *Vnutrenniaia forma slova* (The internal form of the word) (1927).

Ėsteticheskie fragmenty includes a phenomenology of "living discourse" and an analysis of those rules that determine the constitution of meaning in poetic discourse. These phenomenological and structural analyses of language, which aim to construct a poetics, were further developed through a critical assessment of Wilhelm von Humboldt's philosophy of language in Shpet's last

substantial work, *The Internal Form of the Word* (1927). Following a "cleansing" of the GAKhN in 1929, Shpet was forced to retire from his academic post, and he subsequently worked as a translator, editor, and critic. It was during this period that he translated Dickens, Byron, and Shakespeare into Russian. In March 1935 he was arrested by the NKVD (People's Commissariat of Internal Affairs) and was charged with having led an anti-Soviet group during his time at the GAKhN in the 1920s. After a lengthy detention, he was exiled for five years to Eniseisk, and later to Tomsk. There, in 1937, he finished his Russian translation of Hegel's *Phenomenology of Spirit*. In October of that year he was arrested and shot by the NKVD.

SHPET'S DEVELOPMENT TOWARD PHENOMENOLOGY

Representative of Shpet's notion of philosophy before his turn to phenomenology, as well as expectations he held for a reform of philosophy and psychology, is his article *Odin put' psikhologii i kuda on vedët* (One way of psychology and where it leads), published in 1912. The article criticizes experimental and explanatory psychology for having replaced "living and concrete facts" with "empty schemata and abstractions." Only a descriptive psychology that focuses on the pure data of consciousness would be able to fathom psychic life in its concreteness and totality. He saw the basis for this new direction in psychological theory in Wilhelm Dilthey's *Ideas of a Descriptive and Analytical Psychology* (1894). Shpet argued for a philosophy that would take into account the totality of psychic life: a "realistic metaphysics," whose task it would be to grasp "the real in its true essence and its totality." Shpet thought that such a philosophy, which draws on the evident facts of "inner experience," had been realized in important movements of nineteenth- and early twentieth-century Russian philosophy. Philosophers of the "Moscow Metaphysical School" (especially Vladimir Solov'ëv and Sergei Trubetskoi) are cited as exponents of this trend in Russian thought.

Another, no less important, influence on Shpet's reception of Husserl was his interest in the logic of the historical sciences. During his stay in Göttingen (1912–1913) he discovered in Husserl's phenomenology the theory for which he had been searching, and his hermeneutical interest motivated him to try to develop Husserl's "Phenomenology of Reason," as outlined in *Ideas Pertaining to a Pure Phenomenology*, volume 1 (1913), into a theory of hermeneutic reason that focuses on the problem of understanding signs. Although the ideas Shpet encountered in Göttingen primarily concerned transcendental phenomenology—the seminar on "Nature and Spirit," which Shpet attended with other influential phenomenologists like Roman Ingarden and Hans Lipps, certainly met his hermeneutical interests—the ontological trend in the intellectual atmosphere among Husserl's fellow students in Göttingen also should be taken into account.

SHPET'S VERSION OF PHENOMENOLOGY

Shpet's encounter with Husserl's phenomenology, in light of Shpet's expectation of a reform of philosophy and psychology, leads to a singular notion of phenomenology, which is documented in *Iavlenie i smysl* (Appearence and sense, 1914). On the one hand, Shpet tries to reconstruct Husserl's noetic-noematic studies within the framework of an ontological inquiry, based on the Neoplatonism of the Moscow Metaphysical School; on the other hand, he demonstrates the incompleteness of Husserl's analyses of intentional objects, as presented in *Ideas*, volume 1, and completes these analyses with his own. The "noematic sense" intended in acts of consciousness, as presented by Husserl, presupposes for Shpet a class of intentional experiences hardly dealt with in *Ideas*: acts of consciousness through understanding, which play a role in the constitution of all classes of concrete objects. The structure of these "hermeneutic acts" is illustrated by a range of phenomena that are of only minor importance in *Ideas*: the mode of appearance of items of practical use, the specific character of historical sources, and the understanding of linguistic utterances. Thus Husserl's "Phenomenology of Reason" provides a basis for historical cognition in scientific logic, leading eventually to a grounding of the humanities throughan analysis of their conceptual framework and methodology.

Shpet's ensuing works on hermeneutics, philosophy of language, and theory of art, published or written between 1916 and 1927, can be seen as a further development of his hermeneutical phenomenology, the primary idea of which is the correlation of signs (as a combination of expression and meaning) and sign-interpreting consciousness. Shpet also characterizes his project as a semiotic "Philosophy of Culture" in which language, art, myths, and manners are to be described as systems of signs. He develops the basic model of a sign out of Husserl's concept of linguistic expression, which acts as a prototype for all other forms of signs. The idea of a "purely logical grammar," which formulates laws for the

grammatical meanings of natural languages, should be applied analogously to all other cultural systems.

HERMENEUTICS, PHILOSOPHY OF LANGUAGE, AND POETICS

The concrete form of Shpet's phenomenology of hermeneutical reason in his philosophy of language and his poetics was also much influenced by Dilthey's "Philosophy of Life." In Shpet's hermeneutical philosophy, as outlined in *Germenevtika i ee problemy* (Hermeneutics and its problems) (1918), he worked with Schleiermacher's, Boeckh's and Dilthey's theories of understanding. Above all, he tried to deepen and refine Dilthey's late grounding of the humanities—then the culmination in the development of hermeneutics—with insights in the domain of semiotic theories, which he found not only in Husserl's first *Logical Investigation on Expression and Meaning*, but also in other semantic works of the Brentano School (particularly Anton Marty and Alexius Meinong). A combination of Husserl's semantics with Dilthey's hermeneutics would be an enrichment for both sides, as Shpet wrote at the end of the manuscript. The theory of understanding could find a new answer to the question of the mutual relation of the different methods of interpretation, whereas semantics would experience in this combination a "philosophically lively and concrete embodiment."

This actualization of Husserl's philosophical semantics, with a hermeneutical intention, has left its traces in Shpet's *Ėsteticheskie fragmenty* (1922–1923), with which he entered contemporary discussions on literary theory, as initiated by Russian formalism. He was particularly concerned with the definition of the specific character of poetical discourse as opposed to others, be they scientific, rhetorical, or everyday discourses. If one puts this question phenomenologically, one has to ask under what conditions a linguistic utterance appears as artistic or poetic to a listener or reader. Since a poetic utterance is experienced only as a contrast to everyday use of language, one must first analyze the reception of everyday language. Shpet follows this procedure in the second part of *Ėsteticheskie fragmenty*. The difference between understanding the message and understanding its author plays a pivotal role in Shpet's description of the various forms and aspects of linguistic consciousness.

In contrast to such a phenomenological analysis of lingusitic consciousness, Shpet presents a structural analysis of linguistic expression as "ontology of the word," which he, in turn, subsumed under a general theory of semiotics. In this confrontation between a phenomenological inquiry, which is confined to the side of experience, and an ontology, which focuses on the object, the ever-increasing influence of Husserl's early concept of phenomenology on Shpet becomes visible. In Shpet's "ontology of the word" a particular concept of structure is of central importance. "The structure is a concrete construction whose individual parts can vary in their extent and even in their quantity, but not a single part of the whole *in potentia* can be removed without destroying this whole." (1922–1923, II, 11). By "structure of the word" Shpet did not mean the morphological, syntactic, or stylistic construction—in short, not the arrangement of linguistic units "in the plane," but "the organic, depth-wise, as it were, arrangement of the word—from the sensually conceivable wording to the eidetic object." The structure of the word, therefore, consists of the relations between phonemes and meaning, as well as of those between the word's meaning and "object," where the latter is ideal and ontologically distinct from concrete individual things.

When Shpet spoke of the structure of the *word*, he took it in the wide sense of the Russian expression for "word" *slovo*, which can mean sentences or combinations of sentences in discourse, as well as literary texts and even natural language in its entirety. Shpet used it with all these different meanings, yet was mainly concerned with the "communicating word": meaningful discourse able to convey something to another person. Thus Shpet took up Plato's definition of predicative statements, as "the shortest and most simple *logos*" (Sophistēs 262c). Shpet described its structure as follows: in a simple predication the subject denotes a concrete, individual object; the predicate indicates a property belonging to this object. In denoting, speakers refer to a thing; in predication, they say something about it. What can be said about this thing, and conversely, which predications are possible, is determined by the species to which the thing belongs. Therefore the act of intending a species, which Shpet called also the "eidetic object," is indispensable for the construction of a meaningful sentence.

With these definitions Shpet outlined the "word's structure," which is common to everyday and scientific communication, as well as to rhetorical and poetic discourse. In order to explain how this general structure manifests itself in the artistic usage of words, sentences, and discourse, Shpet developed a theory of linguistic functions that stems from a critical assessment of Husserl's and Marty's philosophy of language. He started from three different functions of language, each fulfilled by a particular type of discourse. These three communicative functions are the factual—the expressive, and the

poetic, the latter working through the creative formation of language. Depending on which of these three functions is dominant, discourse is either scientific (concerned with factual communication), rhetorical (concerned with influencing other people's emotions), or poetic (primarily concerned with the arrangement of linguistic expressions as such).

The predominance of one of the three functions implies in each case a different mutual relation between the above-mentioned parts of the word structure. Whereas, for example, in everyday language the arrangement on the level of expression aims primarily at structuring the expressed meaning, and thereby at the communication of facts, in poetic discourse all levels gain a relative importance of their own. The rhythmic forms and syntactic peculiarities of this discourse should attract attention as such. At the same time, the meaning expressed in poetic discourse is more dependent on the external forms of language: whereas the meaning of a factual—above all scientific—communication is not affected by each change of wording and syntactic arrangement.

FROM HUSSERL TO HUMBOLDT

By giving pure logic, which deals with the condition of the possibility of science, a phenomenological foundation, Husserl excluded important aspects of living discourse from his language analysis. Shpet's project was more extensive than Husserl's in that he analyzed scientific communication merely as one possible form alongside the poetic, rhetorical, and everyday discourses. This widening of the horizon entails a turning away from (not only a modification of) Husserl's concept of language. Shpet questioned, for example, one of the central presuppositions of *Logical Investigations*—that scientific discourse can be marked off from living discourse. These two ways of speaking are only tendencies, as Shpet emphasized; they are not fully realized in any empirical speech sample. "Figurativeness is not only a trait of 'poetry'… it is a general property of language, which belongs to scientific discourse as well." (1922–1923, III, 32).

The thesis of the irreducibility of figurative-ambivalent discourse has to do with Shpet's emphasis on the fact that thought is inseparably bound to language. With this concept of language as the "formative organ of thought," as outlined by Shpet in his interpretation of Humboldt in 1927, he turned away most clearly from Husserl's *Logical Investigations*, according to which "the fact of being expressed is arbitrary for the meaning."

See also Existentialism; Phenomenology.

Bibliography

WORKS BY SHPET

Ėsteticheskie fragmenty I–III (Aesthetic fragments). Petrograd: Kolos, 1922–1923. Reprinted in *Sochineniia* (Works), 343–472. Moscow: Pravda, 1989. Shpet's contribution to the theory of art and language.

"Germenevtika i eë problemy" (Hermeneutics and its problems, 1918). In *Gustav Shpet: Mysli Slovo: Izbrannye Trudy* (Thought and word: selected works), edited by Tatiana Shchedrina. Moscow: Rosspen, 2005. Translated into German as *Die Hermeneutik und ihre Probleme*, edited by R. Daube-Schackat and A. Haardt. Freiburg, Munich: Alber, 1993. Critical presentation of the problems of hermeneutics as they have developed throughout history.

Iavlenie i smysl: Fenomenologiia kak osnovnaia nauka i eëe problemy (1914). 2nd ed. Tomsk: Vodolei, 1996. Translated as *Appearance and Sense: Phenomenology as the Fundamental Science and Its Problems* by T. Nemeth. Dordrecht, Boston, London: Kluwer Academic Publishers, 1991. Shpet oulines his "phenomenology of hermeneutical reason."

Istoriia kak problema logiki (History as a problem of logic). Moscow, 1916. Shpet's master's thesis.

"Odin put' psikhologii i kuda on vedët" (One way of psychology and where it leads). In *Filosofskii sbornik L. M. Lopatinu*, 245–264. Moscow, 1912. Representative of Shpet's notion of philosophy before his turn to phenomenology.

Problema prichinnosti u Iuma i Kanta (The problem of causality in Hume and Kant). Kiev: Universitet Sv. Vladimira, 1907. Shpet's first work, a monograph.

Vnutrenniaia forma slova (The internal form of the word). Moscow: GAKhN, 1927. 2nd ed. Ivanovo: Ivanovskii gosudarstvennyi universitet, 1999. A critical assessment of Humboldt's philosophy of language.

WORKS ABOUT SHPET

Haardt, A. *Husserl in Rußland: Phänomenologie der Sprache und Kunst bei Gustav Špet und Alekseij Losev* (Husserl in Russia: Phenomenology of language and art in Gustav Shpet and Aleksei Losev). Munich: Fink, 1993.

Shchedrina, T. G. *"Ja pishu kak ėkho drugogo…": Ocherki intellektual'noi biografii Gustava Shpeta* ("I'm writing like an echo of other people…": Studies on the Intellectual Biography of Gustav Shpet). Moscow: Progress-Tradiciia, 2004.

Alexander Haardt (2005)

SIBLEY, FRANK

(1923–1996)

Frank Sibley was trained as a philosopher in postwar Oxford. His principal teacher was Gilbert Ryle, who, understandably, had a profound influence on Sibley's way of doing philosophical analysis—an influence that is as apparent in his last papers as in his first ones.

Sibley must be credited with inaugurating the renaissance in aesthetics and philosophy of art in the English-speaking world after World War II, a renaissance that is still in full cry. He did it in 1959, in an article that, in the years since, has never ceased being discussed and cited in the literature, and, at the time of its appearance, produced a veritable deluge of essays, and even books in response or defense, that completely reinvigorated the discipline.

"Aesthetic Concepts" (1959a), as Sibley titled his inaugural article, dealt, in a surprisingly few pages, with three of the most basic and difficult issues in the discipline: taste, criticism, and the distinction between the aesthetic and nonaesthetic. He began, with a sensitive ear for "ordinary language" that was to characterize to the end all of his work in aesthetics, by distinguishing between the kinds of things one says about works of art such as "that a novel has a great number of characters and deals with life in a manufacturing town" or "that a painting uses pale colors, predominantly blues and greens," and such remarks, in contrast, as "that a poem is tightly knit" or "that a picture lacks balance or has a certain serenity and repose." About these different kinds of remarks, Sibley claims, "It would be natural enough to say that the making of judgments such as these [latter ones] requires the exercise of taste, perceptiveness, or sensitivity of aesthetic discrimination or appreciation; one would not say this of my first group" (pp. 63–64).

Sibley calls the terms that he thinks require a perceptiveness, sensitivity, or taste beyond that of "normal eyes, ears, and intelligence," aesthetic concepts or terms. And it is the central, most controversial of his claims that aesthetic concepts or terms are, as he puts it, not condition-governed, which is to say, "There are no sufficient conditions, no non-aesthetic features such that the presence of some set or number of them will beyond question logically justify or warrant the application of an aesthetic term" (1959a, p. 67).

Sibley was, throughout his professional life, reticent to publish because of a deeply ingrained perfectionism. Even though his philosophical reputation stems mainly from the groundbreaking "Aesthetic Concepts," it is not the only one of his publications to have influenced the field. Particularly worthy of mention are "Aesthetics and the Looks of Things" (1959b) and "Aesthetic and Nonaesthetic" (1965), in both of which Sibley further explores the whole question of aesthetic reason-giving. As well, Sibley's work in aesthetics and philosophy of art is now likely to have a renewed influence on the field through the posthumous publication of essays he was in the process of preparing for the press at the time of his death. The range of subjects broached in these essays demonstrates that, to the last, Sibley was at the cutting edge of research and is likely to remain a potent philosophical force for many years to come.

See also Aesthetic Qualities.

Bibliography

WORKS BY SIBLEY

"Aesthetic Concepts." *Philosophical Review* 68 (1959a).

"Aesthetics and the Looks of Things." *Journal of Philosophy* 66 (1959b).

"Aesthetic and Nonaesthetic." *Philosophical Review* 74 (2) (1965): 135–159.

Perception: A Philosophical Symposium. London: Methuen, 1971.

Approach to Aesthetics: Collected Papers on Philosophical Aesthetics, edited by John Benson, Betty Redfern, and Jeremy Roxbee Cox. Oxford, U.K.: Clarendon Press, 2001.

WORKS ABOUT SIBLEY

Brady, Emily, and Jerrold Levinson, eds. *Aesthetic Concepts: Essays after Sibley*. Oxford, U.K.: Clarendon Press, 2001.

Cohen, Ted. "Aesthetic/Non-aesthetic and the Concept of Taste: A Critique of Sibley's Position." *Theoria* 39 (1973): 113–151.

Kivy, Peter. *Speaking of Art*. The Hague: Nijhoff, 1973.

Peter Kivy (2005)

SIDGWICK, HENRY
(1838–1900)

Henry Sidgwick, the English philosopher and educator, was born in Yorkshire and attended Rugby and Trinity College, Cambridge. After a brilliant undergraduate career, he was appointed a fellow at Trinity in 1859. He had already begun to have religious doubts, and in the years following 1860 he studied Hebrew and Arabic intensively, hoping to resolve these doubts through historical research. At the same time Sidgwick was teaching philosophy, and he had for many years been a leading member of the small group that met for philosophical discussions with John Grote. Gradually he came to think that if answers to his religious questions were to be found at all, they would be found through philosophy—but he never fully quieted his doubts. In 1869 he resigned his fellowship because he felt he could no longer honestly subscribe to the Thirty-nine Articles, as fellows were required to do. His college promptly appointed him to a lectureship, and when religious tests were dropped, he was reappointed fellow. In 1876 he married Eleanor Balfour, sister of Arthur Balfour. He succeeded T. R. Birks as Knight-

bridge professor of moral philosophy in 1883, and continued actively teaching in the moral sciences course until his death.

WORK AND ACTIVITIES

Philosophy was only one of Sidgwick's many interests—he also wrote on education, literature, political theory, and history of political institutions. He was active in the cause of women's education at Cambridge and had a large part in the founding of Newnham College for women, to which he devoted considerable time and money. Another main interest was psychical research—he performed some experiments with F. W. H. Myers as early as 1873, and in 1882 he helped found the Society for Psychical Research. He served twice as the society's president, and investigated and reported on many alleged psychical phenomena, very few of which, however, he believed to be both genuine and significant.

Sidgwick's most important work is *The Methods of Ethics* (1874). His other philosophical writings, although interesting for the light they throw on his moral philosophy, are too slight, too occasional, or too little original to be of independent significance; but the *Methods* has been held by C. D. Broad and other writers to be the greatest single work on ethics in English—and possibly in any language. Sidgwick's work in economics and political science is generally thought not to be of comparable importance.

PHILOSOPHICAL METHOD

The Methods of Ethics exemplifies Sidgwick's views on the nature of philosophy. The philosopher's aim is not to discover new truths; rather, it is to give systematic organization to knowledge that we already possess. Theoretical philosophy attempts to unify the knowledge obtained through the sciences, so that all of it may be seen as a whole and all the methods used in science may be seen as parts of one method. Practical philosophy has a similar task to perform with our common moral knowledge of what ought to be and what ought to be done, and with the methods we use in obtaining this knowledge.

In carrying out the task of practical philosophy, Sidgwick offered a resolution of a perennial controversy that had been particularly sharp in the middle years of the nineteenth century—that between utilitarians, such as J. S. Mill, and intuitionists, such as William Whewell. However, he found himself unable to reach a solution to another central controversy, that between those who held that morality is independent of religious belief and those who held that without religion no coherent morality is possible.

A brief summary of the course of the argument of *The Methods of Ethics* will make these points plain. Sidgwick took a method of ethics to be a reasoned procedure for reaching specific decisions about what one ought to do. The methods used by common sense, he argued, may be reduced to three. One method takes excellence or perfection as an ultimate goal, and claims that we have intuitive knowledge of a variety of independently valid moral principles and maxims. We reach specific conclusions by subsuming particular cases under the relevant principles. According to the other two methods, we are to infer the rightness or wrongness of acts from the amount of happiness they would cause. According to one method, we calculate the consequences to the agent alone. According to the other, we consider the consequences for everyone affected by the act. Moral rules and principles, for these two methods, are only useful indications of the effects that certain kinds of actions may generally be expected to have. After discussing some basic ethical concepts, Sidgwick examined each method separately and then considered their mutual relations. He concluded that the first method, intuitionism, and the third, utilitarianism, supplement one another, and that their conclusions form a systematic whole. Thus, it is reasonable to act as those conclusions dictate. The remaining method, egoism, can also be systematically developed, and it is reasonable to act according to its conclusions. Either of the two views thus reached dictates obligations that are binding quite independently of any religious sanctions.

However, empirical evidence alone does not show that the conclusions of the egoistic method will always agree with those of the intuitional-utilitarian method. Using methods that are perfectly reasonable, we are sometimes led to serious contradictions. Unless we can find some evidence for the existence of a moral power that will repay self-sacrifice and punish transgression, we will be unable to bring all our practical beliefs and methods into any coherent system. The mere fact that the existence of a power that rewards and punishes behavior is needed to make our practical beliefs coherent does not justify the assertion that there is such a power. Sidgwick personally held that the theistic view is natural for man, but he despaired of finding any evidence to support it and refused to use it in his philosophy. The consequence of the existence of these practical contradictions is (as Sidgwick put it in the melancholy concluding words of the first edition of the *Methods*) that "the prolonged effort of the human intellect to frame a perfect ideal of rational

conduct is seen to have been foredoomed to inevitable failure."

BASIS OF CLASSIFICATION

Sidgwick's classification of the methods implicit in commonsense morality rests on two considerations. First, the methods reflect two sides of human nature. Those taking happiness as the final end reflect the sentient side of man, the capacity for enjoying and suffering, while the method taking excellence as the final end reflects the fact that man is also an active being, with a need to do as well as a need to feel. Second, the classification indicates an epistemological distinction that Sidgwick constantly took as basic, the distinction between propositions that we are entitled to assert only because we have correctly inferred them from others that we know, and propositions that we are entitled to assert because we know them without any inference, directly or "intuitively." The intuitional method claims that we have noninferential knowledge of moral principles, while the other methods emphasize the ways in which moral rules and maxims are arrived at by inference.

NONINFERENTIAL TRUTH

If there is inferential knowledge, Sidgwick believed, there must be noninferential knowledge; and since he also held that there are no infallible sources of noninferential knowledge, the problem arises of how to test claims to possess noninferential truth or claims to have found self-evident propositions. Sidgwick proposed four tests that apparently self-evident propositions must pass before we can be justified in accepting them: (1) the terms in which they are stated must be clear and precise; (2) their self-evidence must be very carefully ascertained; (3) they must be mutually consistent; and (4) there must be general agreement of experts on their truth. Sidgwick argued at great length that commonsense moral principles, which according to traditional intuitionism are self-evident, fail to pass these tests. Hence, if they are true principles, as we all take them to be, they must be inferential and dependent, not self-evident and independent.

SELF-EVIDENT MORAL PRINCIPLES

What do commonsense moral principles depend on? There are four principles that do pass Sidgwick's tests and that he accepted as self-evident. (1) Whatever action anyone judges right for himself, he implicitly judges to be right for anyone else in similar circumstances. (2) One ought to have as much regard for future good or evil as for present, allowing for differences in certainty. (3) The good enjoyed by any individual is as important as the good enjoyed by any other. (4) A rational being is bound to aim generally at good.

PRINCIPLE OF BENEVOLENCE

From the principles that the good of each person is equally important and that a rational being must aim generally at good, Sidgwick deduced an abstract principle of benevolence. Commonsense morality, he argued, appeals to this principle to settle cases in which its usual rules give no answers, and allows its rules to be overridden by the principle if they conflict with it. These facts indicate that common sense considers its rules to depend for their validity on this principle. However, the abstract principle of benevolence is also at the center of utilitarianism, and commonsense morality—the stronghold of traditional intuitionists—is thus seen to be fundamentally utilitarian. The utilitarian, in turn, can have no objection to any of the self-evident principles, and the two methods can thus be completely synthesized. Even the egoist can accept three of the self-evident truths; his rejection of the fourth is an indication of the basic contradiction in the realm of practical reason.

CRITICISMS OF UTILITARIANISM

Sidgwick is usually considered a utilitarian, and he frequently referred to himself as one. However, his views differ considerably from those of the earlier utilitarians.

EMPIRICISM. Sidgwick rejected the empiricist epistemology that J. S. Mill developed and that seemed to underlie Jeremy Bentham's thought. Empiricism, as Sidgwick understood it, holds that the basic premises from which all knowledge is built are cognitions of particular facts and that these cognitions alone are infallible. Sidgwick argued that these cognitions are not infallible and that empiricism cannot give a satisfactory account of the principles of inference that guide the construction of knowledge from the basic data. Metaphysically, he rejected not only materialism but also the reductive sensationalism to which he believed the empiricist epistemology led. Following Thomas Reid, he held to what he called a commonsense dualism of mind and matter, although he found the connections between the two most obscure.

DEFINITION OF ETHICAL TERMS. Sidgwick also rejected what he took to be the traditional utilitarian attempt to define ethical concepts such as "good" and "ought" in terms of nonethical concepts such as "pleas-

ant" or "conducive to most pleasure" and in this way to justify the construction of a purely factual, scientific morality. No reduction of "ought" to "is," of ideal to actual, had yet been successful, he held, although he hesitated to say that no reduction could possibly succeed. However, he did affirm that it is impossible to make an ethical first principle true by definition. To define "good" as "pleasure" is self-defeating if you wish to hold, as a first principle that the good is pleasure, since what you hold as a principle would then be a tautology, and a tautology cannot be an ethical first principle. Recognition of these points, Sidgwick believed, would force the utilitarian to admit the need of a basic intuition in his philosophy.

MOTIVATION. Sidgwick rejected the motivational theories of Bentham and the Mills. He did not think that we always necessarily act to obtain what we take to be our own pleasure or our own good.

THE RELEVANCE OF PSYCHOLOGY. Sidgwick strongly objected to the tendency, which he attributed to Mill, to substitute psychological (or perhaps, with Auguste Comte, sociological) investigation into the origins of ideas and beliefs for properly philosophical investigation of their applicability or truth. Quite aside from his doubts as to the adequacy of the associationist psychology that the earlier utilitarians accepted, Sidgwick held that psychological discoveries about the antecedents and concomitants of ideas and beliefs are, in general, irrelevant to questions of their truth and validity—and psychology can tell us only about antecedents and concomitants. It cannot supersede the deliverances of direct introspective awareness on the question of what our ideas now are.

DETERMINISM. Sidgwick agreed with the earlier utilitarians that there seems to be overwhelming evidence in support of a deterministic view of human action. However, he held that this evidence must be balanced against the fact that in the moment of choosing between alternative actions we inevitably think ourselves free to choose either alternative. He argued that the issue is, therefore, not yet settled, but he held that it is not important for ethical theory that it should be.

INDEPENDENCE OF POLITICS. Sidgwick held that utilitarianism does not necessarily lead to reforming radicalism in politics. He pointed out the strong utilitarian element to such conservative thinkers as David Hume and Edmund Burke, and he argued at great length that a utilitarian would be extremely cautious in recommending important changes.

AGREEMENTS WITH UTILITARIANISM

Sidgwick's position was, of course, utilitarian in its major ethical aspects. He held that the only ultimate or intrinsic good is desirable or pleasant states of consciousness; that acts are objectively right only if they produce more good than any other alternative open to the agent; and that moral rules, such as those of truth-telling or promise-keeping, are subordinate to the principle of utility and are dependent on it for whatever validity they possess. He also held that the value of character and motive is derived from, and to be judged in terms of, the consequences of the actions to which they tend to lead. Sidgwick's disagreements with the traditional forms of utilitarianism are part of his attempt to show that the utilitarian view of morality is independent of metaphysical doctrines, psychological theories, and political platforms and therefore is capable of being what he argued it is—the position toward which commonsense morality in every age and in every society has tended.

See also Balfour, Arthur James; Bentham, Jeremy; Broad, Charlie Dunbar; Burke, Edmund; Common Sense; Consequentialism; Egoism and Altruism; Empiricism; Ethics, History of; Grote, John; Hume, David; Mill, James; Mill, John Stuart; Moral Principles: Their Justification; Pleasure; Reid, Thomas; Utilitarianism; Whewell, William.

Bibliography

WORKS BY SIDGWICK

The Methods of Ethics. London, 1874. Extensively revised for the editions of 1877, 1884, and 1890. The sixth edition (1901) was the last revised by Sidgwick; it contains an autobiographical sketch of great interest.

Principles of Political Economy. London: Macmillan, 1883, 1887, 1901.

Outlines of the History of Ethics. London, 1886. Many subsequent editions.

Elements of Politics. London: Macmillan, 1891 and 1897.

Practical Ethics. London: S. Sonnenschein, 1898.

Philosophy, Its Scope and Relations. London: Macmillan, 1902.

Lectures on the Ethics of Green, Spencer, and Martineau. London, 1902.

Development of European Polity. London, 1903.

Miscellaneous Essays and Addresses. London and New York: Macmillan, 1904.

Lectures on the Philosophy of Kant, edited by James Ward. London: Macmillan, 1905.

WORKS ON SIDGWICK

Bradley, F. H. "Mr. Sidgwick's Hedonism." In his *Collected Essays.* Oxford, 1935. Vol. I. Polemic against Sidgwick.

Broad, C. D. *Ethics and the History of Philosophy.* London: Routledge, 1952. This and the two following works are useful and lengthy discussions of Sidgwick.

Broad, C. D. *Five Types of Ethical Theory.* London: Kegan Paul, 1930.

Broad, C. D. *Religion, Philosophy, and Psychical Research.* London: Routledge, 1953.

Havard, W. C. *Henry Sidgwick and Later Utilitarian Political Philosophy.* Gainesville: University of Florida Press, 1959.

Hayward, F. H. *The Ethical Philosophy of Sidgwick.* London, 1901. Not very useful.

Sidgwick, A., and E. M. Sidgwick. *Henry Sidgwick; A Memoir.* London, 1906. The standard biography, written by his brother and his widow. It contains letters, unpublished papers, and a complete bibliography of his writings.

J. B. Schneewind (1967)

SIDGWICK, HENRY [ADDENDUM]

Henry Sidgwick is renowned for giving classical utilitarianism its most sophisticated dress and greatly advancing substantive ethical theory. Celebrated for his clarity and cool impartiality, he developed an approach to ethical theory that profoundly shaped influential philosophers from G. E. Moore and Bertrand Russell down to R. M. Hare, John Rawls, Marcus Singer, Derek Parfit, and Peter Singer. It was Sidgwick, rather than Moore, who set the course for twentieth-century debates over the ethics and metaethics of the utilitarian view that maximizing happiness is the ultimate normative demand—that is, over such matters as the conflict between egoistic and utilitarian reasons, the distinction between total and average utility, the role of commonsense in utilitarian reasoning, the meaning of *good*, and the moral standing of other beings that are not human. Yet Sidgwick himself had more comprehensive intellectual, religious, and cultural concerns than most of his later analytical admirers. He was haunted by the specter of skepticism in religion and morality, and if he turned utilitarianism into a respectable academic philosophy, he also reluctantly brought it into the crisis of the Enlightment.

Educated in classics and mathematics at Trinity College, Cambridge, Sidgwick spent his entire adult life at Cambridge, becoming Knightbridge Professor in 1883. Molded by the influential discussion society known as the Cambridge Apostles, he developed serious interests in theology, biblical criticism, poetry, education, ethics, political economy, jurisprudence, political theory, sociology, epistemology, metaphysics, and parapsychology (he was a founder and president of the Society for Psychical Research). He vastly influenced the Cambridge moral sciences curriculum and was a guiding force in the cause of women's higher education and the founding of Newnham College. In 1876 he married Eleanor Mildred Balfour, a force in her own right in psychical research and educational reform. Moreover, he was deeply involved with the work of his close friend John Addington Symonds, a pioneer of cultural history and gay studies.

Sidgwick's masterpiece, *The Methods of Ethics* (1874), was a sustained effort at independent, secular moral theory resulting from his decade of "storm and stress" over the defense and reform of Christianity. It also reveals that, however indebted Sidgwick was to his chief mentor, J. S. Mill, his hedonism was more consistently Benthamite, whereas his overall position was more eclectic, reconciling utilitarianism with arguments from Plato, Aristotle, Descartes, Immanuel Kant, Joseph Butler, Samuel Clarke, William Whewell, John Grote, F. D. Maurice, and T. H. Green. It rejects the empiricism and reductionism of earlier utilitarianism, and adheres to a sophisticated fallibilist intuitionism involving various tests for reducing the risk of error with respect to basic non-inferentially known propositions:

1) clarity and precision.

2) ability to withstand careful reflection.

3) mutual consistency.

4) consensus of experts.

The Methods of Ethics is largely a systematic critical comparison of the methods of ethical egoism, common sense or intuitional morality, and utilitarianism—for Sidgwick, the ongoing procedures for determining, on principle, what one ought to do (though he would later devote as much attention to idealism and evolutionism). He takes the notion of *ought* or *right* as fundamental and irreducible and, for the most part, gives an internalist account of moral approbation. But he also holds that it is a plausible and significant (not tautological) proposition that ultimate good is pleasure or desirable consciousness; egoism and utilitarianism hence reduce to egoistic and universalistic hedonism. He then shows that earlier utilitarians exaggerated the conflict with common sense, confused the utilitarian and egoist positions, and failed to give their view rational foundations. His exhaustive examination of commonsense morality, after the manner of Aristotle, reveals time and again that such principles as veracity, fidelity, justice, and benevolence are either too vague and indeterminate, or too conflicting and variably interpreted to form a system of rational intuitions.

Indeed, common sense is even unconsciously utilitarian because it is apt to resort to that view to complete its own system—for example, to settle conflicts between the duty to speak the truth and the duty to keep one's promises.

Thus, commonsense morality ends in utilitarianism, though utilitarianism grounded on philosophical intuitionism, and utilitarianism can in turn rationalize much of commonsense morality as the (indirect) means to the greatest happiness. But no such reconciliation of utilitarianism and egoism is forthcoming, each being, on reflection, equally defensible. Kantian universalizability, the essence of justice, comports with either egoism or utilitarianism and cannot decide between them, though it is another self-evident principle. Sidgwick dismally concludes that there is a dualism of practical reason rendering it incoherent. Without help from epistemology or theology, he has no rational way to settle conflicts between individual self-interest and universal good. Arguably, his demand that these be reconciled in a manner doing justice to the force of both means that his view is better described as dualist, rather than simply utilitarian.

Still, Sidgwick's other intellectual and reformist interests often radiated from his fears about the implications of the dualism of practical reason. His research in parapsychology was largely devoted to seeking evidence of personal survival of death, since such evidence, he believed, might bolster a theism affording the needed reconciliation. And although *The Principles of Political Economy* (1883) and *The Elements of Politics* (1891) tend rather to assume a utilitarian standpoint, they also bespeak his concern that human emotions be shaped in a more deeply altruistic direction, encouraging sympathetic, benevolent sentiments and reigning in narrow or materialistic egoistic ones.

Both his reformism and his philosophical and scientific pursuits were brought to bear on the potential for such societal evolution and the perhaps limited place of reason and religion within it. Never as sanguine as Comte, Mill, or Spencer, his concern for reform was tempered by fear that skepticism and crude egoism would lead to social deterioration. If Sidgwick was as good at defending an agent-relative egoism as an agent-neutral utilitarianism, this was scarcely the result he sought, unless some high-minded reconciliation could be effected as well.

But Sidgwick's views on civilization and its direction suggest both continuities and discontinuities with earlier utilitarianism. It would be hard to deny that troubling racist undercurrents can be found in his work, or that his educational and political writings and activities, in particular, reflected the pervasive late Victorian culture of imperialism. Sidgwick was a friend and colleague of such imperialist luminaries as Sir John Seeley, and went so far as to edit Seeley's posthumous *Introduction to Political Science*. Arthur Balfour, the future prime minister, was his student, brother-in-law, and colleague in psychical research, and also influenced his politics. Ironically, given the priority of politics in Benthamism, the political and economic dimensions of Sidgwick's utilitarianism have been comparatively neglected. This is doubly ironic because Sidgwick was, in fact, an influential economic and political theorist who shaped the views of Alfred Marshall, F. Y. Edgeworth, and other seminal figures in modern economics.

Only by reading *The Methods of Ethics* in the context of Sidgwick's other work and activities is there some hope of determining whether he was a true "government house" utilitarian, holding that the publicity of moral principles be subject to felicific calculations congenial to paternalistic governments, or a defender of the plain person's capacity for moral self-direction, as his focus on common sense and method might suggest. Still, it is clear that Sidgwick articulated a truly comprehensive practical philosophy, and a sophisticated metaethics and epistemology, one deeply informed by Kantianism and idealism as well as utilitarianism. He was not a naïve encyclopedist lacking any grasp of social theory or the historicity of his own philosophy. But whether he began, in his last decades, to doubt the philosophical quest for certainty enough to approximate the pragmatist via media is a very difficult question that has put Sidgwick back in the middle of debates over the imperialistic origins of contemporary political Liberalism.

See also Aristotle; Balfour, Arthur James; Bentham, Jeremy; Butler, Joseph; Clarke, Samuel; Comte, Auguste; Darwinism; Descartes, René; Enlightenment; Ethical Egoism; Green, Thomas Hill; Grote, John; Hare, Richard M.; Hedonism; Idealism; Justice; Kant, Immanuel; Liberalism; Mill, John Stuart; Moore, George Edward; Parfit, Derek; Plato; Rawls, John; Singer, Peter; Skepticism, History of; Utilitarianism; Whewell, William.

Bibliography

PRIMARY WORKS

The Ethics of Conformity and Subscription. London: Williams & Norgate, 1870.

The Methods of Ethics. London: Macmillan, 1874.

The Principles of Political Economy. London: Macmillan, 1883.

Outlines of the History of Ethics for English Readers. London: Macmillan, 1886.
The Elements of Politics. London: Macmillan, 1891.
Practical Ethics: A Collection of Addresses and Essays. London: Swan Sonnenschein, 1898.
Philosophy, Its Scope and Relations. London: Macmillan, 1902.
Lectures on the Ethics of T. H. Green, H. Spencer, and J. Martineau. London: Macmillan, 1902.
The Development of European Polity. London: Macmillan, 1903.
Miscellaneous Essays and Addresses. London: Macmillan, 1904.
Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays. London: Macmillan, 1905.
Sidgwick, Henry, A Memoir. London: Macmillan, 1906.
The Complete Works and Select Correspondence of Henry Sidgwick, edited by Bart Schultz. Charlottesville, VA: InteLex, 1996.
Essays on Ethics and Method by Henry Sidgwick, edited by Marcus G. Singer. Oxford: Oxford University Press, 2000.

SECONDARY WORKS

Crisp, Roger, and Bart Schultz, eds. "Sidgwick 2000." *Utilitas* 12 (3) (2000): 251–401.
Harrison, Ross, ed. *Henry Sidgwick*. Oxford: Oxford University Press, 2001.
Hurka, Thomas. "Moore in the Middle." *Ethics* 113 (2003): 599–628.
Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
Schneewind, J. B. *Sidgwick's Ethics and Victorian Moral Philosophy*. Oxford: Oxford University Press, 1977.
Schultz, Bart, ed. *Essays on Henry Sidgwick*. New York: Cambridge University Press, 1992.
Schultz, Bart. *Henry Sidgwick, Eye of the Universe: An Intellectual Biography*. New York: Cambridge University Press, 2004.
Shaver, Robert. *Rational Egoism*. New York: Cambridge University Press, 1999.
Williams, Bernard. "The Point of View of the Universe: Sidgwick and The Ambitions of Ethics." *Cambridge Review*, 7 (1982):183–91.

R. Barton Schultz (1996, 2005)

SIGER OF BRABANT

(c. 1240–c. 1281/1284)

Of Siger's life, we know very few facts for certain. His exact place of birth remains unknown, as well as the locale and circumstances of his death. (Did he die peacefully in Liege, Belgium, or was he assassinated in Italy at the Roman curia?) Even the chronology of his works is uncertain. Although they are thought to have been written between 1265 and 1277, the precise dates remain conjectural.

Concerning his university career, facts are again unclear. Although it is certain that he never left the faculty of arts for one of the higher faculties (theology, medicine, law), his role in the debates that shook the University of Paris and led to the statutes of 1272 remains the subject of discussion (Putallaz and Imbach 1997 versus Bianchi 1999). At the beginning of his career, he was one of Thomas Aquinas's most outspoken adversaries, but the question as to what degree he would have abandoned Averroism to adopt Thomist views remains open. Certain passages seem to support the view that he would have abandoned Averroism, while others are incompatible with this hypothesis (Van Steenberghen and Maurer defend the developmental interpretation, whereas Mandonnet and Bukowski defend the idea that Siger never changed his mind and was the strictest Averroist of his time, a philosopher who could without any guilt subscribe to heretical propositions).

All of these often radical oppositions about the interpretation of Siger's doctrines—whether metaphysical, psychological, ethical, or logical—illustrate the difficulty involved in understanding the complex thought of a Master of Arts who taught in a time as intellectually rich as it was eventful. Siger was influenced by the famous Dominican theologian Albert the Great, was directly attacked by another famous Dominican theologian, Thomas Aquinas, in his *De unitate intellectus*, was singled out by the condemnations of 1277 (although many of their propositions cannot be related to any of his works), was taken as a model for John of Jandun, later became one of the most important Averroists in the fourteenth century, and was placed by Dante in paradise beside Thomas Aquinas. Faced with this abundance of information, one must consider Siger's texts in themselves by situating them in their context, of course, but also by distinguishing what Siger said from what others say he said. It is well known that the opponents of a thesis tend to present it in a less than advantageous light to make it seem absurd and, in the Middle Ages, heretical. It is also important to take into account how Siger expresses his ideas. For example, Imbach (1996) showed clearly that Siger habitually took certain passages from Thomas Aquinas and twisted them from their original meaning to defend a thesis opposed to that of his illustrious opponent. Such a rhetorical procedure should not be surprising in the context of the condemnations. If we follow these methodological principles, we can draw a clearer and more nuanced portrait of Siger.

PRINCIPLE PHILOSOPHICAL THESES

Siger sought to be a career philosopher. At the end of the thirteenth century, this involved being autonomous from

theology and being independent from established philosophical authorities. This stance influenced Siger's philosophical thought.

Siger's claim that philosophy is independent of theology does not in any way involve a rejection of faith. Rather, it seeks to confine theology to the domain of revelation, where it is the supreme guarantor of truth, and only to where it applies there (Siger 1981/1983, VI, comm. 1). For example, we know through revelation that the world was created. However, revelation does not tell us whether the world was created in time or out of eternity. To decide this question, we would have to investigate the divine will, which is impossible. So we have a choice: either to believe the first thesis on the authority of Augustine, although it rests on no rational argument, or to believe, contrary to Aristotle, for whom the world was not created, the second thesis, a conclusion arrived at by means of natural reason (1972a, *QTDA*, q. 2; 1972a, *DEM*; 1981/1983, III). Between Aristotle, who opposes faith, and the theologians, who pretend to demonstrate their thesis in a philosophical manner that is false, Siger proposes an intermediary path that conforms to the demands of both faith and philosophy: creation out of eternity.

Siger sought to be independent of philosophical authorities, including Aristotle, as we have just seen, as well as Averroes. Indeed, he held that the philosopher must demonstrate for himself the proofs of his predecessors and oppose or correct them if they prove to be erroneous (1981/1983, IV, q. 34). Thus, even in his first work dedicated to noetic (philosophy-of-mind) questions (1972a, *QTDA*, written before 1270), where he is deeply influenced by Averroes, Siger never supported the monopsychist position that Thomas Aquinas attributed to Averroes, a position according to which all of humanity shares a single intellect. This position would imply that there is no individual thinking, as well as no individual immortality, no corporeal fires of hell, and no resurrection of the body.

The best evidence that Siger rejected monopsychism is Aquinas's introduction to his criticism of Siger's doctrine in the *De unitate intellectus* (On the unity of the intellect; written in 1270): "Some, seeing that on Averroes's position it cannot be sustained that this man understands, take another path and say that intellect is united to body as its mover" (III, sec. 66). This view of Siger's position also explains how, in his last work (1972b, presumably written in 1277), Siger could sincerely declare Averroes's noetic doctrine "absurd and heretical" without abandoning his previous doctrine (1972b, q. 27). Here too Siger takes a middle path. The intellect is not united to the body like a sailor to his boat (the error of Plato), nor is it united to the body like a mould to wax (the error of Alexander of Aphrodisias). Rather, the intellect functions intrinsically within the body. Siger held that the intellect is not a unique form completely separate from the body (the position of Averroes according to Aquinas, a position similar to Plato's, and a position against faith and individual morality). He also held that the intellect is not a multiple form completely immanent to the body (the position of Aquinas and Albert according to Siger, a position similar to Alexander's, and a position against philosophy). Rather, intellect, according to Siger, is a mixed form, separate from the body in substance, but joined with it in function (1972a, *QTDA*, q. 7; 1972a, *DAI*, III and VII; 1972b, q. 26).

With regard to morality, about which he wrote very little, as well as psychology, Siger resolutely defended the thesis that the intellect holds sway over the will (1974, *Quaestiones morales*; Ryan 1983), a position that many theologians of the time considered to be equivalent to determinism. In metaphysics, Siger held that there is no real distinction between existence and essence (1981/1983, I, qq. 7–8). He also held that universals, as such, are not substances; they exist only in the soul and are acquired by abstraction from the particular natures of things (1981/1983, III, qq. 15 and 28; 1972a, *DEM*).

See also Agent Intellect; Averroes; Averroism; Eternity; John of Jandun; Thomas Aquinas, St.

Bibliography

Although most significant texts have been critically edited, there are only very few monographs dedicated to Siger's thought. Most studies still more or less depend on Van Steenberghen's evolutionist interpretation and, not surprisingly, concentrate on one of the following subjects: the eternity of the world, the rational soul, and the relation between reason and faith. The most complete bibliography to date is available from http://www.mapageweb.umontreal.ca/pironetf.

WORKS BY SIGER

Quaestiones in tertium de anima. De anima intellectiva. De aeternitate mundi edited by Bernardo C. Baz án. Louvain, Belgium: Publications universitaires, 1972a. (Referred to as *QTDA*, *DAI*, *DEM* respectively.)

Les "quaestiones super librum 'De causis' " de Siger de Brabant edited by Antonio Marlasca. Louvain, Belgium: Publications universitaires, 1972b.

Écrits de logique, de morale et de physique edited by Bernardo C. Baz án. Louvain, Belgium: Publications universitaires, 1974. This volume contains *Sophisma omnis homo de necessitate est animal*; *Quaestio utrum haec sit vera: Homo est animal, nullo homine existente*; *Quaestiones logicales*;

Impossibilia; *Quaestiones morales*; *Quaestiones naturales*; *Compendium super "De generatione et corruptione"*; and *Quaestiones in physicam*.

Quaestiones in metaphysicam, edited by William Dunphy. Louvain-la-Neuve, Belgium: Éditions de l'Institut sup Érieur de philosophie, 1981.

Quaestiones in metaphysicam, edited by Armand Maurer. Louvain-la-Neuve, Belgium: Éditions de l'Institut sup Érieur de philosophie, 1983.

WORKS ON SIGER

Bianchi, Luca. *Censure et libert É intellectuelle à l'Universit É de Paris, XIIIe–XIVe siècles*, 165–201. Paris: Belles Lettres, 1999.

Bukowski, T. P. "Siger of Brabant vs. Thomas Aquinas on Theology." *New Scholasticism* 61 (1987): 25–32.

Dodd, Tony. *The Life and Thought of Siger of Brabant, Thirteenth-Century Parisian Philosopher: An Examination of His Views on the Relationship of Philosophy and Theology*. Lewiston, NY: Mellen Press, 1998.

Imbach, Ruedi. "Notice sur le commentaire du *Liber de causis* de Siger de Brabant et ses rapports avec Thomas d'Aquin." *Freiburger Zeitschrift für Philosophie und Theologie* 43 (1996): 304–323.

Mandonnet, Pierre. *Siger de Brabant et l'Averroïsme Latin au XIIIe siècle*. 2 vols. Louvain, Belgium: Institut sup Érieur de philosophie, 1908–1911.

Maurer, A. "Siger of Brabant and Theology." *Mediaeval Studies* 50 (1988): 257–278.

Putallaz, François-Xavier, and Imbach, Ruedi. *Profession, philosophe: Siger de Brabant*. Paris: Cerf, 1997.

Ryan, C. "Man's Free Will in the Works of Siger of Brabant." *Mediaeval Studies* 45 (1983): 155–199.

Van Steenberghen, Fernand. *Maître Siger de Brabant*. Louvain, Belgium: Publications universitaires, 1977.

Fabienne Pironet (2005)

SIGNIFICANCE TESTS

See *Statistics, Foundations of*

SIGWART, CHRISTOPH

(1830–1904)

Christoph Sigwart, the German philosopher and logician, was born and died in Tübingen. He studied philosophy, theology, and mathematics there and taught in Halle from 1852 to 1855, before joining the theological seminar in Tübingen in 1855. He accepted a professorship at Blaubeuren in 1859 and returned to Tübingen as professor of philosophy, a position he held from 1865 to 1903. His doctoral dissertation was on Giovanni Pico della Mirandola. He also wrote on Friedrich Schleiermacher, Benedict de Spinoza, Huldrych Zwingli, and Giordano Bruno, as well as on ethics. His most important work was the two-volume *Logik,* a comprehensive treatise on the theory of knowledge.

The aim of logic, Sigwart maintained, is normative rather than descriptive. Logic is a regulative science whose aim should be to present a useful methodology for the extension of our knowledge. It is "the ethics rather than the physics of thought" and concerns itself not with an account of psychological processes but with finding the rules in accordance with which thought may achieve objective validity. Like ethics, logic is concerned with the question "What ought I to do?" The adequacy of thought lies not in its correspondence with an antecedently objective reality but in its satisfaction of human purposes. The overriding purpose of reasoning is to reach ideas that are necessary and universal for us, for human beings. Objective validity is essentially a matter of intersubjective agreement. The possibility of discovering the rules for necessary and universally valid thinking, however, depends also on an immediate awareness of self-evidence, a property that is possessed by necessary judgments. The experience of self-evidence is a postulate beyond which we cannot inquire. Logic strives to disclose the conditions under which this feeling occurs.

In Sigwart's philosophy there is a voluntarist element combined with respect for natural science, both of which evidently impressed William James. (James quoted from Sigwart in his essay "The Dilemma of Determinism.") Sigwart held that an activity of free and conscious willing is presupposed not only by ethics and metaphysics but by logic as well. Free will is presupposed by any distinction between correct and incorrect reasoning, since thinking must be a voluntary activity and not necessitated. The will is supreme in the realm of theory as well as in that of practice. The ultimate presupposition of all experience, and therefore of all thinking too, is not merely Immanuel Kant's "I think," which can accompany all ideas, but also "I will," which governs all acts of thought.

Sigwart's classification of the forms of judgments and categories presents judging as the basic cognitive function. Judgments are divided into simple narrative judgments, expressive of an immediate recognition ("This is Socrates"), and complex judgments, presupposing twofold and higher syntheses ("This cloud is red"). The discussion of existential judgments agrees with Kant in denying that existence, or "to be," adds anything to the content of an idea.

Sigwart was also interested in the work of men outside his own country; for example, the *Logik* contains a lengthy discussion of J. S. Mill on induction. Sigwart's ethical and metaphysical views were somewhat conven-

tional: He held that progress in the development of the social order is an inevitable fact of history, and he argued that the attempt to make all our knowledge coherent inevitably leads to the idea of God.

See also Bruno, Giordano; Determinism and Freedom; Epistemology; Epistemology, History of; James, William; Kant, Immanuel; Mill, John Stuart; Pico della Mirandola, Count Giovanni; Schleiermacher, Friedrich Daniel Ernst; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY SIGWART

Spinozas neuentdeckter Traktat von Gott, dem Menschen und dessen Glückseligkeit. Gotha: R. Besser, 1866.

Logik. 2 vols. Tübingen, 1873 and 1878. 5th ed., Tübingen, 1924, contains Heinrich Maier's biography of Sigwart and a bibliography. Translated by Helen Dendy as *Logic.* 2 vols. London, 1890.

Vorfragen der Ethik. Freiburg, 1886.

WORKS ON SIGWART

Häring, T. L. *Christoph Sigwart.* Tübingen, 1930.

Levinson, R. B. "Sigwart's Logic and William James." *Journal of the History of Ideas* 8 (1947): 475–483.

Arnulf Zweig (1967)

SIMMEL, GEORG
(1858–1918)

Georg Simmel, the German philosopher and sociologist, was born in Berlin and resided there except for the last four years of his life. He was educated there, and in 1881 he received his doctorate from the University of Berlin. Three years later he began to teach at that university as a *Privatdozent* and from 1900 he was associate professor without faculty status. Although successful as a lecturer and a writer, he was never promoted to a full professorship at Berlin, nor was he able to secure such a position at any other leading German university. Only in 1914, when his career was almost ended, was he offered a chair in philosophy at the provincial University of Strasbourg. However, World War I disrupted university life there, so that Strasbourg benefited little from Simmel's teaching. Just before the end of the war, Simmel died of cancer.

Simmel's failure as an academic was connected with the nature of his interests, his style of lecturing and writing, and his philosophic position. He had many influential friends—he knew and corresponded with Max Weber, Heinrich Rickert, Edmund Husserl, Adolf von Harnack, and Rainer Maria Rilke—and his applications for openings were always well supported by the testimony of his crowded lecture halls and the success of his many writings, both technical and popular. However, from the straitlaced viewpoint of the German academic hierarchy Simmel was suspect. He seemed to be interested in everything: He wrote books or essays on Rembrandt and Johann Wolfgang von Goethe, on Michelangelo, Auguste Rodin, and Stefan George; on Florence, Rome, Venice, and the Alps; on the philosophy of money, adventure, love, landscapes, and the actor; on ruins, handles, coquetry, and shame; as well as on the more standard philosophic subjects of ethics, philosophy of history, Immanuel Kant, Arthur Schopenhauer, and Friedrich Nietzsche, and, at the end of his life, metaphysics.

Throughout his career Simmel made contributions of lasting importance to sociology, a subject that had not yet achieved academic respectability. His style, too, was not that expected of a professor of philosophy. It was insightful rather than expository; digressive rather than systematic; witty rather than solemn. Because Simmel's position on any particular point was frequently not easy to see, he was often considered to be a critic whose primary impulse was analytic, if not destructive. By some he was thought to have no philosophic position at all.

Other, more sympathetic, readers of his work called him a *Kulturphilosoph*, primarily on the basis of his preoccupation with the objects of culture. Yet because toward the end of his career Simmel began to sketch a philosophic position having a conception of human life at its center, he is also referred to as a *Lebensphilosoph*. Both of these activities, however, are but two sides of the same lifelong dual concern: to illuminate the objects of culture by showing their relation to human experience and to shed light upon the nature of human life by seeing it in relation to its products.

Simmel conceived of human life as being a process and as being, necessarily, productive. By calling life a process (which he expressed by partially defining life as "more-life"), Simmel sought to convey the view that life has the characteristics of what the Greeks called "becoming": It is continuous and continuously changing; strictly speaking, it can only be lived (experienced), not known. However, this same life produces objects that are not in constant flux, that have form and hence are intelligible. (In virtue of this productiveness of human life, Simmel completed his definition by saying that life is "more-than-life.") These products constitute the realm of culture and include not only works of science, history, and art, but social and political institutions and religious theories and practices as well. These objects stand in a twofold rela-

tionship to human life: Their genesis lies in human experience and, once in existence, they are independently subject to being experienced in various ways. Simmel's philosophy dealt in detail with both of these relationships.

To account for the existence of the objects of culture, Simmel made use, in his own particular way, of the categories of form and content. He posited a realm of contents (rather like George Santayana's realm of essences) as the material that enters into all experience. Contents, however, are not experienced as they are in themselves; they are shaped by the experiencing psyche. Experience (Simmel here followed Kant) is formative; to see how form arises thus requires an understanding of the natural history of experience.

Simmel conceived of a stage in human life in which all needs are instantly satisfied, in which there is no gap between desire and fulfillment. Such a stage of life would be prior to experience and hence prior to any differentiation of subject and object. In that stage there would be neither self nor sugar but only sweetness. However, the world is clearly not so organized that life could actually be lived in this way, and in the gap between need and fulfillment both experience and form are born. In becoming conscious, we distinguish between ourselves as subject and that which we experience as objects.

Experience, however, is not all of a piece: We experience in different modes. It is one thing to know an object, another to appreciate it as beautiful, and still another to revere it as an object of worship. In Simmel's view, the contents experienced in each of the three cases may be the same, although they are not the same in experience. The objects of the three experiences differ in that the contents are given shape—are objectified—by means of three different ways of experiencing. The same contents differ in form.

For the most part, people act to fulfill their needs. Their experience gives shape to contents only to the extent to which the immediate requirements of a situation demand it. In the scholastic language Simmel sometimes adopted, both the terminus a quo (the origin) and the terminus ad quem (the goal) of the objects produced by ordinary experience—of whatever mode—remain within the biography of the individual producing them. As a result of this subservience to the needs of individuals, form in ordinary experience is not pure, and the objects that are formed in this way are not yet properly the objects of culture. As long as life sets the goals of action (characteristic of the phase of life Simmel called teleological or pragmatic), knowledge is tentative and limited—not yet science; art is homespun and primitive—not yet fully aesthetic; religion is simple and sporadic—not yet embodied in a theology and in institutions. The form is proto-form and the objects are proto-culture.

However, the bonds of the teleology of life can be broken. The terminus ad quem of people's actions need not reside within their lives: They can act for the sake of a form, a type of action Simmel called free action. Instead of knowing for the sake of acting, some people act in order to know; instead of seeing for the sake of living, some people—artists—live in order to see. In acting for the sake of a form, experience in the relevant form is refined; the structure inchoate in ordinary experience is made explicit and worked out. Form proper is born and the objects of culture are produced.

There are many kinds of form; there is and can be no definitive list. Knowledge, art, religion, value, and philosophy are among the important forms (or "world forms," as Simmel called them) by means of which men have shaped the realm of contents. Reality, too, is only one such form and enjoys no privileged status; the objects of reality constitute the world of practice—those objects which we perceive and manipulate in our daily lives. There are other forms and other worlds, however; one of the tasks of the philosopher is to distinguish and analyze them.

Human life is not self-sufficient; it needs things outside itself to exist and to continue to exist. The objects life forms first come into being to meet its needs; but, because they are objects, they continue to exist independently of life and to make their demands upon the race that has produced them. Humans work out the forms implicit in the various modes of ordinary experience; they become artists, historians, philosophers, and scientists. But once works of art, history, philosophy, or science exist, they make a second demand upon humans: they are the objects by whose assimilation individuals become cultivated. Here Simmel saw a source of inevitable conflict. People differ from each other, and the way in which each person can fulfill himself is peculiar to him. Thus, to fulfill himself each person must utilize a different selection of already existing objects of culture. However, not every road, not just any selection, leads to the assimilation of these objects. To properly understand the objects an individual requires in order to become cultivated, he may need to learn to apprehend a vast number of other objects not so required. In order to serve life, his life, an individual may have to make his own needs subservient to those of forms. This is the tragedy of culture.

In his philosophic position Simmel attempted to do justice to the antitheses that have occupied philosophers since the pre-Socratics. Life as a process is the pole of flux and becoming; it can be lived, but not known. Form is stable and has structure; it is the pole of being and is intelligible. Life is one; experience in all modes is the experience of the same subject. Forms and worlds are many; they are severed from the life that produced them and take on existence independent of it. Neither Being nor Becoming, neither the One nor the Many, holds exclusive sway. The tension between the poles of these antitheses is a permanent feature of the world.

This position underlies the greatest part of Simmel's work. His writings in *Kulturphilosophie* are explorations into the nature of different forms and of different works, whether of philosophy or of art. They are investigations into the relationships between the lives and works of men like Rembrandt and Goethe. In sum, his essays in the philosophy of culture are a series of applications of his philosophy of life.

See also Experience; Philosophy of History.

Bibliography

IMPORTANT WORKS BY SIMMEL

Die Probleme der Geschichtsphilosophie. Eine erkenntnistheoretische Studie. Leipzig: Duncker and Humblot, 1892; 2nd ed. (completely revised), 1905.

Einleitung in die Moralwissenschaft. Eine Kritik der ethischen Grundbegriffe. 2 vols. Berlin: Hertz, 1892–1893.

Philosophie des Geldes. Leipzig: Duncker and Humblot, 1900.

Kant. Sechzehn Vorlesungen gehalten an der Berliner Universität. Leipzig: Duncker and Humblot, 1903.

Die Religion. Frankfurt am Main: Rütten and Loening, 1906. Translated by Curt Rosenthal as *Sociology of Religion*. New York: Philosophical Library, 1959.

Soziologie. Untersuchungen über die Formen der Vergesellschaftung. Leipzig: Duncker and Humblot, 1908. Partly translated, with other essays, by Kurt H. Wolff in *The Sociology of Georg Simmel*. Glencoe, IL: Free Press, 1950.

The Problems of the Philosophy of History: An Epistemological Essay. New York: Free Press, 1977.

The Philosophy of Money. London; Boston: Routledge & Kegan Paul, 1978.

With Otthein Rammstedt. *Gesamtausgabe*. Frankfurt am Main: Suhrkamp, 1989.

COLLECTIONS OF ESSAYS BY SIMMEL

Philosophische Kultur. Gesammelte Essays. Leipzig: Kröner, 1911.

Zur Philosophie der Kunst. Philosophische und kunstphilosophische Aufsätze, edited by Gertrud Simmel. Potsdam: Kiepenheuer, 1922.

Fragmente und Aufsätze aus dem Nachlass und Veröffentlichungen der letzten Jahre, edited by Gertrud Kantorowicz. Munich: Drei Masken, 1923.

Brücke und Tür. Essays des Philosophen zur Geschichte, Religion, Kunst und Gesellschaft, edited by Michael Landmann and Margarete Susman. Stuttgart: Koehler, 1957.

The Conflict in Modern Culture, and Other Essays. New York: Teachers College Press, 1968.

On Individuality and Social Forms: Selected Writings. Chicago:, University of Chicago Press, 1971.

Simmel on Culture: Selected Writings, edited by David Frisby and Mike Featherstone. London; Thousand Oaks, CA: Sage Publications, 1997.

WORKS ON SIMMEL'S PHILOSOPHY

Aron, Raymond. *Essai sur la théorie de l'histoire dans l'Allemagne*. Paris: J. Vrin, 1938. Includes a discussion of Simmel's philosophy of history.

Frisby, David. *Fragments of Modernity: Theories of Modernity in the Work of Simmel, Kracauer, and Benjamin*. Cambridge, MA: MIT Press, 1986.

Frisby, David. *Georg Simmel*. Rev. ed. London; New York: Routledge, 2002.

Frisby, David. *Georg Simmel: Critical Assessments*. London; New York: Routledge, 1994.

Frisby, David. *Simmel and Since: Essays on Georg Simmel's Social Theory*. London; New York: Routledge, 1992.

Frisby, David. *Sociological Impressionism: A Reassessment of Georg Simmel's Social Theory*. 2nd ed. London; New York: Routledge, 1992.

Gassen, Kurt. "Georg-Simmel-Bibliographie." In *Buch des Dankes an Georg Simmel*, edited by Kurt Gassen and Michael Landmann. Berlin: Duncker and Humblot, 1958. Excellent bibliography.

Jankélévitch, Vladimir. "Georg Simmel, philosophe de la vie." *Revue de métaphysique et de morale* 32 (1925): 213–257, 373–386.

Lawrence, Peter A. *Georg Simmel: Sociologist and European*. New York: Barnes & Noble Books, 1976.

Mandelbaum, Maurice. *The Problem of Historical Knowledge: An Answer to Relativism*. New York: Liveright, 1938. Includes a discussion of Simmel's philosophy of history.

Weingartner, Rudolph H. *Experience and Culture: The Philosophy of Georg Simmel*. Middletown, CT: Wesleyan University Press, 1962.

Weinstein, Deena, and Michael A. Weinstein. *Postmodern(ized) Simmel*. London; New York: Routledge, 1993.

Wolff, Kurt H., ed. *Georg Simmel, 1858–1918. A Collection of Essays with Translations and a Bibliography*. Columbus: Ohio State University Press, 1959.

Rudolph H. Weingartner (1967)
Bibliography updated by Michael J. Farmer (2005)

SIMON, RICHARD

(1638–1712)

The French biblical scholar Richard Simon was born in Dieppe, France, and studied with the Oratorians and the

Jesuits and at the Sorbonne, specializing in Hebrew and Near Eastern studies. Before being ordained a priest in 1670, he taught philosophy at an Oratorian college. He soon became one of the foremost experts in Hebrew, Judaism, and Eastern Church history. Influenced by Benedict (Baruch) de Spinoza's critique of the Bible and by the theory of his friend and fellow Oratorian, Isaac La Peyrère, that there were men before Adam, Simon began developing his views about the Bible and church doctrine. His first published work, a defense of the Jews of Metz (1670), attacked Christian anti-Semitism. It was followed by a study of the Eastern Church, another of Jewish ceremonies and customs, and an attack on the monks of Fécamp. His most important and revolutionary work, *Histoire critique du vieux testament*, was printed in 1678. Jacques Bénigne Bossuet caused it to be banned immediately, and almost all copies were destroyed. A few reached England, and the work was published in French with an English translation by Henry Dickinson in 1682. The scandal forced Simon to leave the Oratory and become a simple priest. Thereafter, he argued with various Protestant and Catholic thinkers and wrote many works on the history of religion and on the Bible, which culminated in his translation of the New Testament (1702). Bossuet caused this work to be banned also.

Simon's revolutionary contention was that no original text of the Bible exists, that the texts one possesses have developed and have been altered through the ages, and that it is therefore necessary to apply the method of critical evaluation to biblical materials to establish the most accurate human form of the revelation. This method involves philology, textual study, historical researches, and comparative studies. Protestants saw that Simon's claim that there is no perfect copy of scripture fundamentally challenged their position that truth is found only by examining the Bible. Catholics feared that he was undermining all bases of Judeo-Christianity by raising problems about all its documents and traditions. Simon contended that he was merely trying to clarify religious knowledge by showing its foundations and development and the need for a tradition to interpret and understand it. Whether intentional or not, Simon's method launched the whole enterprise of biblical higher criticism, which was often directed toward undermining confidence in the uniqueness and ultimate truth of the Judeo-Christian revelation.

See also Bossuet, Jacques Bénigne; Philosophy of Religion, History of; Revelation.

Bibliography

WORKS BY SIMON

Histoire critique du vieux testament. Paris: N.p., 1678.*Lettres choisies de M. Simon*. 4 vols. 2nd ed. Frankfurt am Main, Germany: Minerva G.N.B.H, 1967.

WORKS ABOUT SIMON

Bredvold, Louis I. *The Intellectual Milieu of John Dryden: Studies in Some Aspects of Seventeenth-Century Thought*. Ann Arbor: University of Michigan Press, 1956.

Hazard, Paul. *La Crise de la conscience européene (1680–1715)*. 2 vols. Paris: Boivin, 1935.

Mirri, F. Saverio. *Richard Simon e il metodo storico-critico di B. Spinoza: Storia di un libro e di una polemica sullo sfondo delle lotte politico-religiose della Francia di Luigi XIV*. Florence, Italy: F. Le Monnier, 1972.

Ranson, Patric. *Richard Simon, ou, Du caractère illégitime de l'augustinisme en théologie*. Lausanne, Switzerland: L'Age d'homme, 1990.

Steinmann, Jean. *Richard Simon et les origines de l'exégèse biblique*. Paris: Desclee de Brouwer, 1960.

Richard H. Popkin (1967, 2005)

SIMON MAGUS

Simon Magus, the earliest Gnostic leader known to us, was a native of the Samaritan village of Gitta. He is first mentioned in Acts (8:4–25), where he appears as a wonder-worker who had gained a considerable following in Samaria and who sought to augment his stock in trade by purchasing the power of conferring the Holy Spirit from the apostles. The identity of the Simon of the book of Acts and the founder of the Gnostic sect has been questioned, but Irenaeus, among others, has no doubt of it. According to Hippolytus, Simon died in Rome when he failed, in an abortive attempt at a miracle, to rise from the pit in which he had been buried alive. In the pseudo-Clementine literature Simon serves as the target for veiled Jewish-Christian attacks on Paul and Marcion. According to Origen, in his time the Simonians numbered only thirty, but Eusebius, years later, still knew of their existence.

The Simonian theory is of special interest not only as one of the earliest Gnostic systems but also as providing an illustration of the ways in which such systems developed and were modified. Assessment of the evidence is complicated by the meagerness of our sources and by various problems of evaluation and interpretation, but in general we may distinguish three main stages. Simon himself appears to have been a "magician" of the common Hellenistic type, who claimed to be a divine incarnation. His teaching would be not so much Gnostic in the

second-century sense (that is, the Gnosticism of the heretical Christian systems) but rather a form of syncretistic gnosis into which he sought to incorporate Christian elements. The accounts of Justin and Irenaeus introduce his companion, the ex-prostitute Helen, whom he declared to be the first conception (*Ennoia*) of his mind, emanating from him like Athena from the head of Zeus. A notable feature here is the blending of biblical elements with elements from Homer and Greek mythology.

Descending to the lower regions, Ennoia generated the angels and powers by whom this world was made but was then detained by them and compelled to suffer a round of incarnations (thus she is, inter alia, Helen of Troy) until Simon himself came to redeem her. The problem here is to know how much can be credited to Simon himself and how much to reflection among his followers.

A third and more philosophical stage is represented by the "Great Affirmation" preserved by Hippolytus, which probably has nothing to do with the original Simon but may be the work of later disciples attributed, as was often the case, to the master himself. Here the primal ground of being is fire, from which emanate three pairs of "roots," or Powers, which are the origin of all existence: Mind and Thought, Voice and Name, Reason and Desire (text in W. Völker, *Quellen zur Geschichte des christlichen Gnosis,* Tübingen, 1932, pp. 3ff.). In this scheme, elements from Greek philosophy (Heraclitus, Plato, Aristotle) are blended with biblical and Homeric elements into a thoroughly Gnostic system. It is of interest to note that Simonianism provides one of the sources of the later Faust legend.

See also Aristotle; Gnosticism; Heraclitus of Ephesus; Homer; Marcion; Origen; Plato.

Bibliography

GENERAL WORKS

Grant, R. M. *Gnosticism and Early Christianity,* 70ff. New York: Columbia University Press, 1959.

Jonas, Hans. *The Gnostic Religion,* 103ff. Boston: Beacon Press, 1958.

Leisegang, H. *Die Gnosis,* 60ff. Stuttgart, 1955.

Wilson, R. McL. *The Gnostic Problem,* 99ff. London: A.R. Mowbray, 1958.

SPECIALIZED WORKS

Casey, R. P. "Simon Magus." In *The Beginnings of Christianity,* edited by F. J. F. Jackson and Kirsopp Lake. Vol. V, 151ff. London, 1933.

Cerfaux, L. *Receuil Cerfaux.* 2 vols. Vol. I, 191ff. Gembloux, 1954.

Haenchen, E. "Gab es eine vorchristliche Gnosis?" *Zeitschrift für Theologie und Kirche* 49 (1952): 316ff.

Haenchen, E. *Die Apostelgeschichte,* 250ff. Göttingen, 1959.

Hall, G. N. L. "Simon Magus." In *Encyclopaedia of Religion and Ethics,* edited by James Hastings. Vol. XI, 514ff. Deals fairly extensively with the "Great Affirmation."

Headlam, A. C. "Simon Magus." In *Dictionary of the Bible,* edited by James Hastings. Vol. IV, 520ff. On p. 527 Headlam lists four points in the later Faust legend that point back to the legend of Simon Magus.

Lietzmann, H. "Simon Magus." In *Realencyclopädie der classischen Altertumswissenschaft,* edited by August Pauly and Georg Wissowa. 2nd series. Vol. III, Cols. 180ff.

R. McL. Wilson (1967)

SIMPLICIUS
(fl. c. 530)

Simplicius of Cilicia (in Asia Minor) tells us that he studied Platonic philosophy in Alexandria under Ammonius the son of Hermias (fl. c. 550). Afterward, he attended the lectures of Damascius, probably in Athens at the original and still flourishing school founded by Plato himself, the Academy. (An earlier scholarly opinion that there were doctrinal differences between the teachings on Plato in Alexandria and Athens is no longer held.)

All these figures were active neoplatonists, and Hermias and Damascius did in fact publish commentaries on various dialogues of Plato. But Ammonius and Simplicius (and to a lesser extent Damascius as well) devoted most of their writings to the explication of Aristotle's works. Simplicius, in addition to a commentary on Epictetus's Handbook (*Enchiridion*), wrote extensive commentaries on five of those works of Aristotle that most challenge philosophers: *Metaphysics* (no longer extant, although fragments are known), *Physics*, *Categories*, *De Anima*, and *De Caelo*, with the four extant commentaries totaling over 2,800 sizable pages in the series *Commentaria in Aristotelem Graeca*. (References in some modern books to a commentary by Simplicius on *Sophistici Elenchi* are mistaken.) In addition to the time obviously needed to complete these commentaries, a brief examination of Simplicus's learned exegeses shows that he also was in need of an extensive philosophical library, one that included not only Plato and Aristotle, their predecessors (the pre-Socratics) but also everything (it seems) ever written by an Academician or Peripatetic, as well as some Stoic texts.

Where could this library have been? An obvious answer is Athens, but one of the few hard facts concerning these philosophers is that, owing to increasing Christian hostility to pagan philosophizing, the emperor

Justinian in 529 forbade teaching by non-Christians, which gave Simplicius time to write his commentary of Epictetus, who as a philosopher struggling under tyranny could serve as a model for Simplicius and his colleagues. Agathias (c. 536–c. 582), the Christian historian (and epigrammatist), states that "Damascius of Syria, Simplicius of Cilicia, Eulamius of Phrygia, Prisican of Lydia, Hermias and Diogenes of Phoenicia, and Isidore of Gaza … concluded that, since Christianity was not to their liking [a euphemism], Persia was a better place for them." Unfortunately, the stories about King Chosroes I (reigned 531–579) that made him sound like a Platonic philosopher-king were greatly exaggerated. In time, even Greece, with all its dangers, seemed preferable; "and so all returned home," trusting in a treaty between Justinian I (483–565) and Chosroes that, among other things, stipulated that the philosophers could return to their homes and live there as long as they wished "on their own," this last vague phrase probably meaning that the treaty guaranteed them the freedom to congregate as philosophers and conduct themselves (mostly) as before (Agathias *Historiae* 2.30.3–31.4 Keydell).

Thus, although some scholars still believe that Simplicius chose to stay somewhere safe in the Persian Empire, probably in Haran, the explicit evidence of Agathias, who refers to these Academics as his (younger) contemporaries, strongly suggests that Simplicius returned to Athens. There, still denied the right to teach, he dedicated himself to scholarship.

For the most part, Simplicius's writings are straightforward analyses, lemma by lemma, of Aristotelian passages, a form of commentary designed for readers rather than for the students to whom he no longer could lecture. Here Simplicius not only dispassionately and at great length explains the meaning of selected passages he also attempts to harmonize or minimize the differences between Plato and Aristotle. Indeed, Simplicius often turns Aristotle into a neoplatonist, as when, for example, he argues that Aristotle's causes were six in number. The lemmas both explicate the meaning and summarize other scholars' views of the passage in question. In both aspects Simplicius is of immeasurable importance for the history of earlier Greek philosophy, for he, far more than any other commentator on Plato or Aristotle, took the trouble to go back both to the texts Aristotle quotes or alludes to as well as to the texts that comment on Aristotle.

Simplicius is thus the most important source for verbatim quotations of the pre-Socratics, Academics, Peripatetics, Stoics, and others. Time after time, where others comment on Aristotle's allusion to (say) Parmenides merely by elaborating on Aristotle's words, inferring from them what Parmenides meant, Simplicius, explicitly referring to the rarity of Parmenides' book, says that he will quote from it *in extenso*. By far the vast majority of the fragments of Parmenides, Empedocles, Zeno, Melissus, Anaxagoras, and Diogenes of Apollonia is known thanks to Simplicius alone. Earlier attempts to argue that he found these passages in Theophrastus's lost doxographical treatise on earlier thought falter when one looks at the extant *Metaphysics* and *De Sensibus* of Theophrastus, whose verbatim quotations of pre-Socratics are infrequent and not of great length, unlike many in Simplicius. In short, present-day knowledge of the actual words of the pre-Socratics would be halved or worse without him. It would doubtless be increased were a copy of his *In Metaphysica* found.

Similarly, Simplicius is now the only source for many of the earlier but now lost Aristotelian commentaries. Much of what is known of Theophrastus's *Physics* comes from Simplicius's commentary, and his quotations from John Philoponus's lost *Against Aristotle, on the Eternity of the World* are so extensive that they have been excerpted and published separately.

Although Simplicius is strictly neutral toward the pre-Socratics, he is capable of criticizing Aristotelian commentators of several centuries earlier, such as Alexander of Aphrodisias, whom he accuses, sometimes ironically, of not having considered all available sources, a virtue he explicitly declares necessary for the serious commentator in the beginning of *In Cat*, along with an ability to make dispassionate judgments. He is naturally more deferential to his teachers Ammonius and Damascius. He reserves his most critical if not contemptuous statements for Philoponus, who was also a student of Ammonius, but whose Christian interpretations, such as that the cosmos had a fixed beginning, he finds most abhorrent.

Since all of Simplicius's works are in the form of commentaries on two philosophers not of his own school, it is not easy to isolate beliefs and preoccupations that would distinguish him from other neoplatonists. Apart from his almost religious adoration of the Platonic Demiurge (reminiscent of Cleanthes' *Hymn to Zeus*), Simplicius writes very much in the tradition of Alexandrian and Athenian commentators on Aristotle who in place of sustained argument are more likely merely to state their interpretation of his text. Simplicius, then, a scholar like few before him, read every relevant text that would illuminate Aristotle, who he argued should be seen as a complement to Plato's noble philosophy.

See also Greek Academy; Neoplatonism; Peripatetics; Platonism and the Platonic Tradition; Stocism.

Bibliography

Cameron, Alan. "The Last Days of the Academy at Athens." *Proceedings of the Cambridge Philological Society* 195 (15) (1969): 7–29.

Hadot, Ilsetraut. *Le problème du néoplatonisme alexandrin: Hiéroclès et Simplicius*. Paris: Études Augustiniennes, 1978.

Hadot, Ilsetraut, ed. *Simplicius: Sa vie, son oeuvre, sa survie: Actes du colloque international de Paris*. Berlin: De Gruyter, 1987.

Hadot, Ilsetraut. "The Life and Work of Simplicius in Greek and Arabic Sources." In *Aristotle Transformed: The Ancient Commentators and Their Influence*, edited by Richard Sorabji, 275–304. London: Duckworth, 1990.

Hoffmann, Philippe. "Simplicius' Polemics." In *Philoponus and the Rejection of Aristotelian Science*, edited by Richard Sorabji, 57–83. Ithaca, NY: Cornell University Press, 1987.

Praechter, K. "Simplikios (10)." *Pauly-Wissowa, Real-Encyclopädie der classischen Altertumswissenschaft* 3A (1) (1927): 204–213.

Sorabji, Richard, ed. *Ancient Commentators on Aristotle*. London: Duckworth. In this series of translations the following works of Simplicius have appeared thus far, with others forthcoming: *In De Caelo* 1.1–1.4 (translated by R. J. Hankinson, 2002), 1.5–1.9 (translated by R. J. Hankinson), 1.10–1.12 (translated by R. J. Hankinson, 2003); *In Physica* 2 (translated by Barrie Fleet, 1997), 3 (translated by J. O. Urmson, notes by Peter Lautner, 2002), 4.1–4.5, 10–14 (translated by J. O. Urmson, 1992), 5 (translated by J. O. Urmson, notes by Peter Lautner, 1997), 4.6–4.9 (on the void; translated by J. O. Urmson, notes by Peter Lautner, 1994), 4, pp. 601.1–645.19 and 773.8–800.25 (corollaries on place and time, respectively; translated by J. O. Urmson, notes by L. Siorvanes, 1992), 6 (translated by David Konstan, 1989), 7 (translated by Charles Hagen, 1994), 8.6–8.10 (translated by translated by R. McKirahan, 2001), 8.10, pp. 1326.38–1336.34 (against Philoponus on the eternity of the world; translated by C. Wildberg, 1991); *In De Anima* 1.1–2.4 (translated by J. O. Urmson, notes by Peter Lautner, 1995), 2.5–2.12 (translated by Carlos Steel and J. O. Urmson, notes by Peter Lautner, 1997), 3.1–3.5 (translated by H. J. Blumenthal, 2000); *In Categorias* 1–4 (translated by M. Chase, 2003), 5–6 (translated by F. de Haas and Barry Fleet, 2002), 7–8 (translated by Barry Fleet, 2002), 9–15 (translated by R. Gaskin, 2000); *Handbook on Epictetus* 1–26 (translated by C. Brittain and T. Brennan, 2002), 27–53 (translated by C. Brittain and T. Brennan, 2002).

Thiel, Rainer. *Simplikios und das Ende der neuplatonische Schule in Athen*. Mainz, Germany: Abhandlungen der Geistes- und Sozialwissenschaftliche Klasse. Akademie der Wissenschaften und der Literatur, 1999. Nr. 8.

David Sider (2005)

SIMULATION THEORY

A prominent part of everyday thought is thought about mental states. We ascribe states like desire, belief, intention, hope, thirst, fear, and disgust both to ourselves and to others. We also use these ascribed mental states to predict how others will behave. Ability to use the language of mental states is normally acquired early in childhood, without special training. This naïve use of mental state concepts is variously called *folk psychology*, *theory of mind*, *mentalizing*, or *mindreading* and is studied in both philosophy and the cognitive sciences, including developmental psychology, social psychology, and cognitive neuroscience. One approach to mindreading holds that mental-state attributors use a naïve psychological "theory" to infer mental states in others from their behavior, the environment, or their other mental states, and to predict their behavior from their mental states. This is called the *theory theory* (TT). A different approach holds that people commonly execute mindreading by trying to simulate, replicate or reproduce in their own minds the same state, or sequence of states, as the target. This is the *simulation theory* (ST).

Another possible label for simulation is *empathy*. In one sense of the term, *empathy* refers to the basic maneuver of feeling one's way into the state of another, by "identifying" with the other, or imaginatively putting oneself in the other's shoes. One does not simply try to depict or represent another's state, but actually to experience or share it. Of course, mental life may feature empathic acts or events that are not deployed for mindreading. But the term *simulation theory* primarily refers to an account of mindreading that accords to empathy, or simulation, a core role in how we understand, or mindread, the states of others.

HISTORICAL ANTECEDENTS OF THE DEBATE

A historical precursor of the ST/TT debate was the debate between positivists and hermeneutic theorists about the proper methodology for the human sciences. Whereas positivists argued for a single, uniform methodology for the human and natural sciences, early-twentieth-century philosophers like Wilhelm Dilthey and R. G. Collingwood advocated an autonomous method for the social sciences, called *Verstehen*, in which the scientist or historian projects herself into the subjective perspective or viewpoint of the actors being studied. Contemporary ST, however, makes no pronouncements about the proper methodology of social science; it only concerns the prescientific

practice of understanding others. The kernel of this idea has additional historical antecedents. Adam Smith, Immanuel Kant, Arthur Schopenhauer, Friedrich Nietzsche, and W. V. Quine all wrote of the mind's empathic or projective propensities. Kant wrote:

> [I]f I wish to represent to myself a thinking being, I must put myself in his place, and thus substitute, as it were, my own subject for the object I am seeking to consider… (Kant 1787/1961, p. 336)

Nietzsche anticipated modern psychology in the following passage:

> To understand another person, that is to *imitate his feelings in ourselves*, we … produce the feeling in ourselves after the *effects* it exerts and displays on the other person by imitating with our own body the expression of his eyes, his voice, his bearing.… Then a similar feeling arises in us in consequence of an ancient association between movement and sensation. (Nietzsche 1881/1977, pp. 156–157.…)

Quine (1960) briefly endorsed an empathy account of indirect discourse and propositional attitude ascription. He described attitude ascriptions as an "essentially dramatic idiom" rather than a scientific procedure, and this encouraged him to see the attitudes as disreputable posits that deserve to be eliminated from our ontology.

THE BEGINNING OF THE DEBATE

It was in the 1980s that three philosophers—Robert Gordon, Jane Heal, and Alvin Goldman—first offered sustained defenses of ST as an account of the method of mindreading. They were reacting partly to functionalist ideas in philosophy of mind and partly to emerging research in psychology. According to analytic functionalism, our understanding of mental states is based on commonsense causal principles that link states of the external world with mental states and mental states with one another. For example, if a person is looking attentively at a round object in ordinary light, he is caused to have a visual experience as of something round. If he is very thirsty and believes there is something potable in a nearby refrigerator, he will decide to walk toward that refrigerator. By using causal platitudes of this sort, attributors can infer mental states from the conditions of an agent's environment or from his previous mental states. One might start with beliefs about a target's initial mental states plus beliefs in certain causal psychological principles, feed this information into one's theoretical reasoning system, and let the system infer the "final" states that the target went into or will go into. This TT approach assumes that attribution relies on information about causal principles, so TT is said to be a "knowledge rich" approach.

Simulationists typically doubt that ordinary adults and children have as much information, or the kinds of information, that TT posits, even at a tacit or unconscious level. ST offers a different possibility, in which attributors are "knowledge-poor" but engage a special mental skill: the construction of pretend states. To predict an upcoming decision of yours, I can pretend to have your goals and beliefs, feed these pretend goals and beliefs into my own decision-making system, let the system make a pretend decision, and finally predict that you will make this decision. This procedure differs in three respects from the theorizing procedure. First, it involves no reliance on any belief by the attributor in a folk-psychological causal principle. Second, it involves the creation and deployment of pretend, or make-believe, states. Third, it utilizes a mental system, here a decision-making system, for a non-standard purpose, for the purpose of mindreading rather than action. It takes the decision-making system "off-line."

Daniel Dennett (1987) challenged ST by claiming that simulation collapses into a form of theorizing. If I make believe I am a suspension bridge and wonder what I will do when the wind blows, what comes to mind depends on the sophistication of my knowledge of the physics of suspension bridges. Why shouldn't make-believe mindreading equally depend on theoretical knowledge? Goldman (1989) parried this challenge by distinguishing two kinds of simulation: theory-driven and process-driven simulation. A successful simulation need not be theory driven. If both the initial states of the simulating system and the process driving the simulation are the same as, or relevantly similar to, those of the target system, the simulating system's output should resemble the target's output, enabling the prediction to be accurate.

Heal (1994) also worried about a threat of ST collapsing into TT. If ST holds that one mechanism is used to simulate another mechanism of the same kind, she claimed, then the first mechanism embodies tacit knowledge of theoretical principles of how that type of mechanism operates. Since defenders of TT usually say that folk-psychological theory is known only tacitly, this cognitive science brand of simulation would collapse into a form of TT. This led Heal to reject such empirical claims about sub-personal processes. Instead, she proposed

(1998) that ST is in some sense an a priori truth. When we think about another's thoughts, we "co-cognize" with our target; that is, we use contentful states whose contents match those of the target. Heal has claimed that such co-cognition is simulation, and is an a priori truth about how we mindread.

Martin Davies and Tony Stone (2001) criticize Heal's proposed criterion of tacit knowledge possession. Yet another way to rebut the threat of collapse is to question the assumption that the integrity or robustness of simulation can be sustained only if it is not underpinned by theorizing. The assumption is that simulation is a sham if it is implemented by theorizing; ST implies that no theorizing is used. Against this, Goldman (2006) argues that theorizing at an implementation level need not conflict with higher-level simulation, and the latter is what ST insists upon.

TRANSFERENCE

According to the standard account, simulational mindreading proceeds by running a simulation that produces an output state (e.g., a decision) and "transferring" that output state to the target. "Transference" consists of two steps: classifying the output state as falling under a certain concept and inferring that the target's state also falls under that concept. Gordon (1995) worries about these putative steps. Classifying one's output state under a mental concept ostensibly requires introspection, a process of which Gordon is leery. Inferring a similarity between one's own state and a target's state sounds like an analogical argument concerning other minds, which Ludwig Wittgenstein and others have criticized. Also, if the analogy rests on theorizing, this undercuts the autonomy of simulation. Given these worrisome features of the standard account, Gordon proposes a construal of simulation without introspection or inference "from me to you."

Gordon replaces transference with "transformation." When I simulate a target, I "recenter" my egocentric map on the target. In my imagination, the target becomes the referent of the first-person pronoun "I" and his time of action, or decision, becomes the referent of "now." The transformation Gordon discusses is modeled on the transformation of an actor into a character he is playing. Once a personal transformation is accomplished, there is no need to "transfer" my state to him or to infer that his state is similar to mine. But there are many puzzling features of Gordon's proposal. He describes the content of what is imagined, but not what literally takes place. Mindreaders are not literally transformed into their targets (in the way princes are transformed into frogs) and do not literally lose their identity. We still need an account of a mindreader's psychological activities. Unless he identifies the type of his output state and imputes it to the target, how does the activity qualify as mindreading, that is, as believing of the target that she is in state M? Merely being oneself in state M, in imagination, does not constitute the mindreading of another person. One must impute a state to the target, and the state selected for imputation is the output state of the simulation, which must be detected and classified. First-person mental-state detection thereby becomes an important item on the ST agenda, an item on which simulationists differ, some, such as Harris (1992) and Goldman (2006), favoring introspection and others, such as Gordon (1995), resisting it.

Different theorists favor stronger or weaker versions of ST, in which "information" plays no role versus a moderate role. Gordon favors a very pure version of ST, whereas Goldman favors more of a hybrid approach, in which some acts of mindreading may proceed wholly by theorizing, and some acts may have elements of both simulation and theorizing. For example, a decision predictor might use a step of simulation to determine what he himself would do, but then correct that preliminary prediction by adding background information about differences between the target and himself. Some theory theorists have also moved toward a hybrid approach by acknowledging that certain types of mindreading tasks are most naturally executed by a simulation-like procedure (Nichols and Stich 2003).

What exactly does ST mean by the pivotal notion of a "pretend state"? Mental pretense may not be essential for simulational mindreading, for example, for the reading of people's emotional states as discussed at the end of this article. But most formulations of ST appeal to mental pretense. Mental pretense is often linked to imagining, but imagining comes in different varieties. One can imagine that something is the case, for example, that Mars is twice as large as it actually is, without putting oneself in another person's shoes. Goldman (2006) proposes a distinction between two types of imagining: suppositional-imagining and enactive-imagining.

Suppositional imagining is what one does when one supposes, assumes, or hypothesizes something to be the case. It is a purely intellectual posture, though its precise connection to other intellectual attitudes, like belief, is a delicate matter. Enactive imagining is not purely intellectual or doxastic. It is an attempt to produce in oneself a mental state normally produced by other means, where

the mental states might be perceptual, emotional, or purely attitudinal. You can enactively imagine seeing something—you can visualize it—or you can enactively imagine wanting or dreading something. For purposes of ST, the relevant notion of imagination is enactive imagination. To pretend to be in mental state M is to enactively imagine being in M. If the pretense is undertaken for mindreading, one would imagine being in M and "mark" the imaginative state as belonging to the target of the mindreading exercise.

Can a state produced by enactive imagining really resemble its counterpart state, the state it is meant to enact? And what are the respects of resemblance? Gregory Currie (1995) advanced the thesis that visual imagery is the simulation of vision, and Currie and Ian Ravenscroft extended this proposal to motor imagery. They present evidence from cognitive science and cognitive neuroscience to support these ideas, highlighting evidence of behavioral and neural similarity (Currie and Ravenscroft 2002). Successful simulational mindreading would seem to depend on significant similarity between imagination-produced states and their counterparts. However, perfect similarity, including phenomenological similarity, is not required (Goldman 2006).

PSYCHOLOGICAL EVIDENCE

Gordon's first paper on ST (1986) appealed to research in developmental psychology to support it. Psychologists Heinz Wimmer and Josef Perner (1983) studied children who watched a puppet show in which a character is outside playing while his chocolate gets moved from the place he put it to another place in the kitchen. Older children, like adults, attribute to the character a false belief about the chocolate's location; three-year-olds, by contrast, do not ascribe a false belief. Another experiment showed that older autistic children resemble three-year-olds in making mistakes on this false-belief task (Baron-Cohen, Leslie, and Frith 1985). This was interesting because autistic children are known for a striking deficit in their capacity for pretend play. Gordon suggested that the capacity for pretense must be critical for adequate mindreading, just as ST proposes. Most developmental psychologists offered a different account of the phenomena, postulating a theorizing deficit as the source of the poor performances by both three-year-olds and autistic children. It was argued that three-year-olds simply do not possess the full adult concept of belief as a state that can be false, and this conceptual "deficit" is responsible for their poor false-belief task performance.

ENDOWMENT EFFECT

The conceptual-deficit account, however, appears to have been premature. First, when experimental tasks were simplified, three-year-olds and even younger children sometimes passed false-belief tests. Second, researchers found plausible alternative explanations of poor performance by three-year-olds, explanations in terms of memory or executive control deficiencies rather than conceptual deficiencies. Thus, the idea of conceptual change—assumed to be theoretical change—was undercut. This had been a principal form of evidence for TT and, implicitly, against ST. It has proved difficult to design more direct tests between TT and ST.

Shaun Nichols, Stephen Stich, and Alan Leslie (1995) cite empirical tests that allegedly disconfirm ST. One of these types of empirical tests involves the "endowment effect." The endowment effect is the finding that when people are given an item, for example, a coffee mug, they come to value it more highly than people who do not possess one. Owners hold out for significantly more money to sell it back than do nonowners who are offered a choice between receiving a mug and receiving a sum of money. When asked to predict what they would do, before being in such a situation, subjects underpredict the price that they themselves subsequently set. Nichols, Stich, and Leslie argue that TT readily explains this underprediction; people simply have a false theory about their own valuations. But ST, they argue, cannot explain it. If simulation is used to predict a choice, there are only two ways it could go wrong. The predictor's decision-making system might operate differently from that of the target, or the wrong inputs might be fed into the decision-making system. The first explanation does not work here, because it is the very same system. The second explanation also seems implausible because the situation is so transparent. This last point, however, runs contrary to the evidence. Research by George Loewenstein and other investigators reveals countless cases in which self- and other-predictions go wrong because people are unable to project themselves accurately into the shoes of others, or into their own future shoes. The actual current situation constrains their imaginative construction of future or hypothetical states, which can obviously derail a simulation routine (Van Boven, Dunning, and Loewenstein 2000). So ST has clear resources for explaining underpredictions in endowment effect cases.

EMOTION RECOGNITION

One of the best empirical cases for simulation is found in a domain little studied in the first two decades of empir-

ical research on mindreading. This is the domain of detecting emotions by facial expressions. Goldman and Sripada (2005; also Goldman, 2006) survey findings pertaining to three types of emotions: fear, disgust, and anger. For each of these emotions, brain-damaged patients who are deficient in experiencing a given emotion are also selectively impaired in recognizing the same emotion in others' faces. Their mindreading deficit is specific to the emotion they are impaired in experiencing. ST provides a natural explanation of these "paired deficits": normal recognition proceeds by using the same neural substrate that subserves a tokening of that emotion, but if the substrate is damaged, mindreading should be impaired. TT, by contrast, has no explanation that is not ad hoc. TT is particularly unpromising because the impaired subjects retain conceptual ("theoretical") understanding of the relevant emotions.

By what simulational process could normal face-based emotion recognition take place? One possibility involves facial mimicry followed by feedback that leads to (subthreshold) experience of the observed emotion. In other words, normal people undergo traces of the same emotion as the person they observe. This resembles Nietzsche's idea, now supported by research showing that even unconscious perception of faces produces covert, automatic imitation of facial musculature in the observer, and these mimicked expressions can produce the same emotions in the self.

Another possible explanation of emotion recognition is unmediated mirroring, or resonance, in which the observer undergoes the same emotion experience as the observed person without activation of facial musculature. Such "mirror matching" phenomena have been identified for a variety of mental phenomena, in which the same experience that occurs in one person is also produced in someone who merely observes the first. Such mirror matching occurs for events ranging from action with the hands (Rizzolatti et al., 2001), to somatosensory experiences (Keysers et al., 2004), to pain (Singer et al., 2004). For example, if one observes somebody else acting, the same area of the premotor cortex is activated that controls that kind of action; if one observes somebody being touched on the leg, the same area of somatosensory cortex is activated that is activated in the normal experience of being touched on the leg; the same sort of matching applies to pain. This leads Vittorio Gallese (2003) to speak of a "shared manifold" of intersubjectivity, a possible basis for empathy and social cognition more generally. It is unclear whether mirror matching always yields recognition, or attribution, of the experience in question, so perhaps mindreading is not always implicated. But the basic occurrence of mental simulation, or mental mimicry, is strikingly instantiated.

See also Cognitive Science; Folk Psychology; Psychology.

Bibliography

Baron-Cohen, Simon, Alan Leslie, and Uta Frith. "Does the Autistic Child Have a 'Theory of Mind'?" *Cognition* 21(1985): 37–46.

Carruthers, Peter, and Peter K. Smith, eds. *Theories of Theories of Mind*. New York: Cambridge University Press, 1996.

Currie, Gregory. "Visual Imagery as the Simulation of Vision." *Mind and Language* 10(1995): 25–44.

Currie, Gregory, and Ian Ravenscroft. *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford: Oxford University Press, 2002.

Davies, Martin, and Tony Stone. "Mental Simulation, Tacit Theory, and the Threat of Collapse." *Philosophical Topics* 29 (2001): 127–173.

Davies, Martin, and Tony Stone, eds. *Folk Psychology*. Oxford: Blackwell, 1995.

Davies, Martin, and Tony Stone, eds. *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell, 1995.

Dennett, Daniel. "Making Sense of Ourselves." In his *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

Gallese, Vittorio. "The Manifold Nature of Interpersonal Relations: The Quest for a Common Mechanism." *Philosophical Transactions of the Royal Society of London B* 358 (2003): 517–528.

Goldman, Alvin. "Interpretation Psychologized." *Mind and Language* 4 (1989): 161–185.

Goldman, Alvin. *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. New York: Oxford University Press, 2006.

Goldman, Alvin, and Chandra Sripada. "Simulationist Models of Face-based Emotion Recognition." *Cognition* 94 (2005): 193–213.

Gopnik, Alison. "How We Know Our Minds: The Illusion of First-Person Knowledge of Intentionality." *Behavioral and Brain Sciences* 16 (1993): 1–14.

Gopnik, Alison, and Andrew N. Meltzoff. *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press, 1997.

Gordon, Robert. "Folk Psychology as Simulation." *Mind and Language* 1 (1986): 158–171.

Gordon, Robert. "Simulation without Introspection or Inference from Me to You." In *Mental Simulation*, edited by Martin Davies and Tony Stone. Oxford: Blackwell, 1995.

Harris, Paul. "From Simulation to Folk Psychology: The Case for Development." *Mind and Language* 7 (1992): 120–144.

Heal, Jane. "Simulation vs. Theory Theory: What Is at Issue?" In *Objectivity, Simulation and the Unity of Consciousness: Current Issues in the Philosophy of Mind*, edited by Christopher Peacocke. Oxford: Oxford University Press, 1994.

Heal, Jane. "Co-cognition and Off-line Simulation: Two Ways of Understanding the Simulation Approach." *Mind and Language* 13 (1998): 477–498.

Kant, Immanuel. *Critique of Pure Reason*. Translated by Norman Kemp Smith. London: Macmillan, 1961.

Keysers, C., B. Wicker, V. Gazzola, J.-L. Anton, L. Fogassi, and V. Gallese. "A Touching Sight: SII/PV Activation During the Observation and Experience of Touch." *Neuron*. 42 (2004): 335–346.

Nietzsche, Friedrich. "Daybreak" (1881). In *A Nietzsche Reader*. Translated by R. J. Hollingdale. Harmondsworth: Penguin, 1977.

Nichols, Shaun, and Stephen P. Stich. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press, 2003.

Nichols, Shaun, Stephen Stich, and Alan Leslie. "Choice Effects and the Ineffectiveness of Simulation: Response to Kuhberger et al." *Mind and Language* 10 (1995): 437–445.

Perner, Josef. *Understanding the Representational Mind*. Cambridge, MA: MIT Press, 1991.

Quine, Willard Van Orman. *Word and Object*. Cambridge, MA: Technology Press, 1960.

Rizzolatti, G., L. Fogassi, and V. Gallese. "Neurophysiological Mechanisms Underlying the Understanding and Imitation of Action." *Nature Reviews Neuroscience* 2 (2001): 661–670.

Singer, T., B. Seymour, J. O'doherty, H. Kaube, R. J. Dolan, And C. D. Frith. "Empathy for Pain Involves the Affective but not Sensory Components of Pain." *Science*, 303 (2004): 1157–1162.

Van Boven, L., D. Dunning, and G. Loewenstein. "Egocentric Empathy Gaps between Owners and Buyers: Misperceptions of the Endowment Effect." *Journal of Personality and Social Psychology* 79 (2000): 66–76.

Wimmer, Heinz, and Josef Perner. "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception." *Cognition* 13 (1983): 103–128.

Alvin I. Goldman (2005)

SINGER, PETER

(1946–)

Peter Singer is one of the most influential philosophers of the twentieth century. While other philosophers have been more important in the development of the discipline, none has changed more lives. *Newsweek* magazine observed that the modern animal rights movement may be dated from the publication of *Animal Liberation*. This book has sold more than 500,000 copies in sixteen languages thus far. Altogether Singer is responsible in whole or part for producing thirty-six books, and a vast number of articles and reviews in journals ranging from *The Philosophical Review* to the *New York Times*.

Peter Singer was born in Melbourne, Australia, on July 6th, 1946. His parents were Viennese Jews who escaped in 1938, shortly after the *Anschluss* incorporated Austria into the German Reich. He went on to Melbourne University, where as an undergraduate he studied law, history, and philosophy. In 1969 he received an MA in philosophy, writing a thesis on *Why Should I Be Moral?* A scholarship allowed Singer to complete his graduate studies in Oxford, where he received his bachelor's in philosophy in 1971 and served as Radcliffe lecturer from 1971 to 1973.

In 1972 Singer published *Famine, Affluence, and Morality* in the first volume of a new journal, *Philosophy and Public Affairs*. This article, which has been reprinted more than two dozen times, is important for several reasons. In terms of style it was an unconventional philosophical essay in that it was written in simple, direct prose, with few references to philosophical texts. Rather than beginning from Immanuel Kant, Aristotle, or a hypothetical moral question, it addressed events that were occurring as Singer was writing. The article began with these words: "As I write this, in November, 1971, people are dying in East Bengal from lack of food, shelter, and medical care." Singer went on to present his readers with a stark moral challenge. On the basis of some apparently simple, plausible premises, he argued that affluent people ought to transfer their resources to those who are worse off until they reach the point at which further transfers would hurt them more than they would benefit others. Singer was asking his readers to give up their opera tickets, their wine cellars, and private schools for their children—the accoutrements of the sophisticated, upper-middle-class life favored by many academics. Furthermore, Singer was completely unapologetic about making such demands: "The whole way we look at moral issues … needs to be altered, and with it, the way of life that has come to be taken for granted in our society."

In autumn 1973, Singer moved to the United States in order to teach at New York University. He was in America only sixteen months, but his visit had a large impact. He wrote most of *Animal Liberation* during his stay and, while working on the book, Singer presented draft chapters to philosophy departments around the country. Also during his time in New York, Singer wrote "Philosophers Are Back on the Job" for the *New York Times Magazine* (1974). This essay brought the practical ethics movement to the attention of a wide, non-professional audience.

In 1975 Singer returned to Melbourne where he remained until 1999, except to take up various visiting appointments in universities around the world. Since 1999 he has been the Ira W. DeCamp Professor of Bioethics at Princeton University.

Virtually all of Singer's work exemplifies the following three important characteristics. First, it is revisionary. The point of practical ethics is not simply to understand the world, but to change it. A second characteristic of Singer's work is that facts matter. Philosophy may begin where facts run out, as Singer wrote in "Philosophers Are Back on the Job" (p. 20), but it is hard to see what philosophy would be for Singer if it didn't start with a vivid appreciation of the way things are. Finally, Singer's work presupposes that individual action can make a difference. As his work has unfolded, Singer has increasingly addressed social policy dimensions of the problems that he considers, but he usually writes as one person in conversation with another. His goal is to change our attitudes and behavior because that is how one changes the world.

Although he has written widely, Singer is most closely associated with his defense of animals and his attack on the traditional ethic of the sanctity of human life. According to Singer, other things being equal, it is better to experiment on a profoundly brain-damaged human infant than on a normal chimpanzee. The normative theory that underwrites these judgments is utilitarianism. The good to be maximized, in the case of self-conscious creatures (persons), is satisfied preferences; in the case of non-persons, it is pleasure and the absence of suffering. In metaethics, Singer follows the universal prescriptivism of his teacher, R. M. Hare.

Singer's recent writing has ranged from practical ethics to work that is more personal. His most recent book, *The President of Good and Evil* (2004), takes President George W. Bush's moralism at face value, and subjects it to rigorous philosophical examination. His 2002 book, *One World*, is an ethical assessment of the environmental, economic, and legal dimensions of globalization. *Pushing Time Away* (2003) is the most personal of his books. It is a moving biography of Singer's maternal grandfather, David Oppenheim, a Viennese intellectual and teacher, who was murdered in the Holocaust. In recovering the life, thought, and sensibility of Oppenheim, Singer discovers strong affinities with his own thought and intellectual formation, perhaps because of a common source in the *Haskalah* (Jewish Enlightenment). Only in his late-50s as of 2005, Singer is likely to continue to produce important work in all areas of moral philosophy.

See also Animal Rights and Welfare; Moral Sentiments.

Bibliography

PRIMARY WORKS

Practical Ethics. 2nd ed. New York: Cambridge University Press, 1993.

Rethinking Life and Death: The Collapse of Our Traditional Ethics. New York: St. Martins Press, 1995.

Animal Liberation: A New Ethics for Our Treatment of Animals. 2nd ed. Ecco: New York, 2001.

SECONDARY WORKS

Jamieson, Dale, ed. *Singer and his Critics*. Oxford: Blackwelll, 1999.

Peter Singer. http://www.petersingerlinks.com/

Dale Jamieson (2005)

SKEPTICISM

See *Skepticism, Contemporary; Skepticism, History of*

SKEPTICISM, CONTEMPORARY

Skepticism regarding a subject matter is the view that knowledge about the subject matter is not possible. Many subject matters have come under skeptical attack. It has been argued, for example, that it is not possible to obtain knowledge about the external world, about as-yet-unobserved states of affairs, and about minds other than one's own. This entry will focus upon skepticism about knowledge of the external world.

THE CARTESIAN SKEPTICAL ARGUMENT

The following skeptical argument is suggested by Descartes's first Meditation. Consider the skeptical hypothesis SK: There are no physical objects; all that exists is my mind and that of an evil genius, who causes me to have sense experience just like that which I actually have (sense experience representing a world of physical objects). This hypothesis, says the skeptic, is logically possible and incompatible with propositions implying the existence of the external world, such as that I have hands. The skeptic then claims that (1) if I know that I have hands, then I know that not-SK. To justify premise (1), the skeptic points out that the proposition that I have hands entails not-SK, and he asserts this closure principle: If S knows that ϕ and S knows that ϕ entails ψ, then S knows that ψ. The skeptical argument's other premise is that (2) I do not know that not-SK. To justify this prem-

ise, the skeptic points out that, if SK were true, then I would have sense experience exactly similar to that which I actually have. Because my sensory evidence does not discriminate between the hypothesis that SK and the hypothesis that not-SK, this evidence does not justify me in believing not-SK rather than SK. Lacking justification for my belief that not-SK, I do not know that not-SK. From (1) and (2) it follows that I do not know that I have hands. A similar argument may be given for each external-world proposition that I claim to know.

Those who think that minds are physical in nature may well balk at the skeptic's claim that the evil-genius hypothesis is logically possible. Accordingly, the skeptic will replace that hypothesis with this updated version of SK: I am a brain in a vat connected to a computer that is the ultimate cause of my (thoroughly unveridical) sense experience.

To see how the foregoing pattern of skeptical reasoning may be extended to other subject matters, let the target knowledge claim be that there are minds other than my own, and let the skeptical hypothesis be that the complex patterns of bodily behavior that I observe are not accompanied by any states of consciousness. The analogue to premise (2) will in this case be supported by the claim that, if the skeptical hypothesis were true, then I would have behavioral evidence exactly similar to that which I actually have.

DENYING THE LOGICAL POSSIBILITY OF SK

Let us consider two radical responses to the Cartesian skeptical argument. The evil-genius and vat hypotheses both depend on the assumption that the external world is mind-independent in such a way that it is logically possible for sense experience to represent there to be a physical world of a certain character even though there is no physical world, or at least no physical world of that character. An idealist denies this assumption of independence. The idealist maintains that facts about physical objects hold simply in virtue of the holding of the right facts about sense experience, then denies that skeptical hypotheses such as SK are logically possible: any world in which the facts of sense experience are as they actually are is a world in which there is an external reality of roughly the sort people take there to be. Thus premise (2) is false: I know that not-SK in virtue of knowing the necessary falsity of SK.

The second radical response to the skeptical argument rests on a verificationist constraint on the meaningfulness of sentences. Like the idealist, the verificationist holds that the sentence "I am a victim of thoroughgoing sensory deception" fails to express a logically possible hypothesis. Given that the sentence fails to express a proposition for which sense experience could in principle provide confirming or disconfirming evidence, the verificationist counts the sentence as meaningless. Because the sentence expresses no proposition at all, it does not express a proposition that is possibly true.

The antirealist puts forward a similar view, maintaining that one's understanding of a sentence's meaning consists in a recognitional capacity manifestable in one's use of the sentence. Suppose that the conditions under which a sentence X is true transcend people's powers of recognition. Then one's understanding of X's meaning could not be identified with one's grasping of X's recognition-transcendent truth conditions (because such a grasping could not, in turn, be identified with a manifestable recognitional capacity). This conception may be applied to sentences that allegedly express skeptical hypotheses. If people cannot detect the obtaining of their truth conditions, then what is understood when skeptical sentences' meanings are understood must be something other than their truth conditions. Grasping such sentences' meanings must instead consist in grasping the detectable conditions under which they are warrantedly assertible. Thus, it would turn out that an allegedly problematic skeptical hypothesis fails to make any coherent claims about putative conditions in the world that outstrip the human capacity for knowledge.

ATTACKING PREMISE (1)

Premise (1) has come under attack by those who think that the skeptic has succeeded in stating a hypothesis that is genuinely logically possible and not known to be false. On this strategy the closure principle is denied. This opens up the possibility that I know that I have hands even though I do not know that not-SK. For example, one may deny closure by maintaining that knowing that ϕ requires knowing only that the *relevant* alternative hypotheses to ϕ do not obtain. Skeptical hypotheses, it is then said, are not relevant alternatives to the propositions involved in ordinary knowledge claims.

Another way of denying closure is to hold that S knows that P if and only if (i) S correctly believes that P, and (ii) S would not mistakenly believe that P if P were false. To satisfy the tracking condition (ii), S must not mistakenly believe that P in the possible worlds in which P is false that are *most similar* to the actual world, according to the standard semantics for counterfactuals. (Robert Nozick adds the further tracking requirement that S

believes that P in the possible worlds in which P is true that are most similar to the actual world.) Now suppose that some hypothetical normal believer S satisfies these conditions with respect to the proposition that he has hands (S correctly believes that he has hands and would not mistakenly believe that he has hands in the no-hands possible worlds most similar to his world, worlds in which, say, he has lost his hands in a terrible accident). Then S knows that he has hands. But in all the possible worlds in which not-SK is false (SK worlds), S mistakenly believes that not-SK (he mistakenly believes that he is not in a vat). So S does not know that not-SK, even though this proposition is entailed by the proposition that S has hands. This is a counterexample to the closure principle.

ATTACKING PREMISE (2)

Let us turn to antiskeptical strategies that do not challenge premise (1) and that accept that SK is indeed logically possible. On these strategies, premise (2) is attacked. For example, Kant tried to show via a transcendental argument that, in allowing knowledge of certain key features of one's own mind, the Cartesian is already committed to the possibility of knowledge of the external world. Kant argued (in "Refutation of Idealism" in *Critique of Pure Reason*) that, in order to have knowledge of one's own temporally-ordered inner states, one must also have knowledge of spatial objects outside one's mind, whose temporal ordering is related to that of one's inner states. A prima facie difficulty for the Kantian strategy is that arguing for a connection between knowledge of one's mind and knowledge of the external world seems to require the assumption of verificationism or idealism, which would render superfluous the rest of the transcendental argument.

The inference to the best explanation strategy relies on the idea that, even if two incompatible explanatory hypotheses are equally supported by the available evidence, I am still justified in rejecting one hypothesis if the other offers a better explanation of the evidence. It might be maintained that the ordinary hypothesis that the world is roughly as I take it to be offers a better explanation of my sensory evidence than does SK, in virtue of its greater simplicity. Thus, I can justifiably reject SK. The proponent of this strategy needs to specify the respect in which SK is more complex than the ordinary hypothesis and to make it plausible that hypotheses that are complex in the specified way are less likely to be true than simpler ones.

Another way to attack premise (2) is to adopt a reliabilist theory of knowledge, according to which knowing that ϕ is a matter of having a reliably produced true belief that ϕ. If reliabilism is correct, then in arguing that I do not know that not-SK, the skeptic would have the difficult burden of showing that there is in fact some flaw in the belief-producing mechanism that yields my belief that not-SK (thereby precluding that belief's amounting to knowledge).

Let us return to the skeptic's *defense* of his premise (2). To validate the premise, the skeptic needs to appeal to an epistemic principle that is (apparently) distinct from the closure principle. This is the underdetermination principle:

> (UP) If S's evidence for Φ does not favor Φ over a competing incompatible hypothesis Ψ, then S is not justified in believing Φ.

The skeptic maintains that one's perceptual evidence would be the same regardless of whether SK holds or not-SK holds. By (UP), then, one's perceptual evidence fails to justify one in believing that not-SK. Hence, one does not know that not-SK.

According to one response to this line of thought, experiences justify perceptual beliefs (such as that a cat is near) without providing evidence or reasons for these beliefs because evidence and reasons always come in the form of *beliefs* which inferentially justify other beliefs. Thus the skeptic cannot appeal to (UP) in the foregoing way. Some philosophers maintain that perceptual Perceptual experiences, some philosophers maintain, justify perceptual beliefs in virtue of having *propositional content*, although they are not themselves propositions. A visual perception, say, has the representational content expressible by the sentence "A cat is near," and accordingly justifies an associated perceptual belief's having that same content.

One problem for this view is that it is plausible to suppose that nonhuman animals have perceptual experiences with representational contents that are similar to those of humans (given the physiological similarities between the relevant perceptual systems). But the animals' perceptual representations do not possess *propositional* content. One may reply that experiences nevertheless justify perceptual beliefs by virtue of having *nonpropositional* representational content, such as that possessed by maps and pictures. This view is, in one way, less attractive than the propositional view, however, because it is easier to see how a belief-like state with propositional content can justify a perceptual belief than to see how a state with a nonpropositional content can perform the same justifying feat.

Further, both views about perceptual justification have the following difficulty. (UP) can be reformulated as:

> (UP*) If S's putative justifier for Φ does not favor Φ over a competing incompatible hypothesis Ψ, then S is not justified in believing Φ.

Now the skeptic may hold that one's nonevidential, perceptual putative justifier would be present regardless of whether SK holds or not-SK holds. Thus, one is not justified in believing not-SK, as the skeptic originally claimed.

Against this, it has been held that the perceptual states that one has when not-SK holds differ in their intrinsic nature from those that one has when SK holds. On this view, the veridical perceptual states possessed by a normal perceiver are *object-involving*, in that objects such as cats are constituents of their perceptual contents. This view might be put forward as a direct realist answer to skepticism, according to which our awareness of external objects is not mediated by awareness of our own experiences. But such direct realism has little antiskeptical force: the skeptic may maintain that even if veridical experience, should it occur, involves direct awareness of cats, it is nevertheless possible that all of one's experiences are unveridical, none possessing an object-involving perceptual content. When the object-involving view is put forward a little differently, however, there is a greater payoff. A disjunctive view challenges the skeptic's use of (UP*). Unlike a veridical perceptual experience of a cat, a nonveridical perceptual state of a brain in a vat is obviously *not* object-involving. The two states, then, are not tokens of a single perceptual state type; there is no common factor between the states. Because it is not true that the same putative perceptual justifier would be present regardless of whether SK holds or not-SK holds, (UP*) cannot be used to show that one lacks justification for believing not-SK. Thus, on the disjunctivist approach, premise (2) of the skeptical argument is not adequately supported.

One may use considerations from the philosophy of language and the philosophy of mind to argue that SK is in fact false. According to semantic externalism, the Cartesian commits an error in attempting to construct thought experiments involving massive deception. The Cartesian naively assumes that, starting with a subject S of thought and experience who is ensconced in a normal external environment, we may hold fixed the contents of S's thoughts and the meanings of his sentences while varying (in thought) S's external environment in such a way that S's thoughts about his environment come out to be predominantly false. According to the semantic externalist, the Cartesian fails to realize that the contents of one's thoughts and the meanings of one's sentences depend in certain ways on one's external environment.

For example, Donald Davidson argues that, when we interpret a speaker's sentences as expressing various beliefs that he holds, we are constrained to attribute beliefs to him that are by and large true of the environment with which he interacts (Davidson 1986). This is because there is no rational basis for preferring one interpretation that finds him to be massively mistaken in his beliefs over another such interpretation. It is constitutive of beliefs and of sentential meanings that they are what are correctly attributed in correct interpretation, on Davidson's view. Thus, it follows from the nature of belief and meaning that, contrary to what SK states, one can never be so massively mistaken.

To see another manifestation of this anti-Cartesian line of thought, consider Hilary Putnam's Twin Earth, a planet like Earth except for the circumstances that the clear, thirst-quenching liquid that the Twin Earthians call "water" is composed of XYZ molecules rather than H_2O molecules. The Twin Earthians' term "water" does not refer to water, but rather to the liquid on Twin Earth with which they interact. Hence, my Twin Earth counterpart's word "water" does not have the same meaning as my word, and when the Twin Earthian says "Water is wet," it is not to thereby express the thought that I think when I think that water is wet. Similarly, the semantic externalist maintains that, when my envatted twin in a treeless world uses the word "tree" in thought, it is not to refer to trees. Instead, the brain in a vat refers to those entities in the external environment that play a causal role with respect to his uses of "tree" analogous to that played by trees with respect to normal uses of "tree" in a tree-filled world. These entities may be states of the computer that systematically cause the brain in a vat to have "tree-like" sense experience. When the brain in the vat thinks the sentence "A tree has fallen," he does not thereby mistakenly express the thought that a tree has fallen. Instead, he expresses a thought about computer states, which may well be true of his environment. In general, then, the brain in a vat is not massively mistaken about the world, contrary to what the Cartesian maintains.

We may use these considerations, together with the assumption that I have knowledge of the contents of my own thoughts, against premise (2) in the following way: I am now thinking that a tree has fallen; if SK is true, then I am not now thinking that a tree has fallen; thus, SK is false. This argument, however, is powerless against ver-

sions of the skeptical hypothesis on which the brain in a vat is indirectly causally linked to ordinary objects. If, for example, there are programmers of the computer who refer to trees, then it becomes plausible to suppose that the brain does so as well. Further, there is a prima facie problem as to whether I may claim knowledge of the contents of my own thoughts, given semantic externalism. Such knowledge seems to require independent knowledge of the content-determining causal environment in which I am located, knowledge the antiskeptical argument was meant to provide.

AMBIVALENCE ABOUT THE SKEPTICAL ARGUMENT

Contextualism is a response to skepticism that is based upon a novel view of the semantics of knowledge-attributing sentences of the form "S knows that P." According to the contextualist, such sentences are like sentences of the form "X is flat." The truth value of the latter sort of sentence depends upon both (1) the shape of the pertinent object, and (2) contextually determined standards regarding contour. Relative to one conversational context (in which bicycle racing is under discussion, for instance), "The road is flat" can come out true; relative to another context (where inclined planes are under discussion), the sentence (concerning the same road) can come out false. Similarly, the truth value of an utterance of, say, "John knows that the bank is open this Saturday" depends upon both (1) John's epistemic situation (e.g., his evidential beliefs, his perceptual experience, whether the bank is indeed open), and (2) contextually determined epistemic standards (set by the interests, intentions, and expectations of the knowledge-attributing conversationalists).

Suppose that John's basis for claiming that the bank is open this Saturday is that he visited it on a Saturday a month ago. Suppose that my business partner and I wish to deposit a check this Saturday or some time the following week. Then my partner's utterance of "John knows that the bank is open this Saturday" may well be true, given John's epistemic situation and given the low stakes in our conversational context. Holding John's epistemic situation fixed, imagine a different case in which our business will go bankrupt if the check is not deposited on Saturday. In this case, my partner's utterance of "John knows that the bank is open this Saturday" may well be false, given the higher stakes in this context, in which evidence superior to John's may well be required for knowledge about the bank.

The contextualist claims that his view both (a) explains why the skeptical argument may seem compelling, and (b) implies that there is much ordinarily-attributed knowledge in the world. When skepticism and skeptical possibilities are under discussion, the conversational context is such that abnormally high epistemic standards are in place. Accordingly, an utterance of the argument's premise (2)—"I do not know that not-SK"—comes out true. According to the contextualist, utterances of the argument's closure-based premise (1) are true in all conversational contexts. Thus, relative to a skeptical context, an utterance of the argument's conclusion is true. However, in an ordinary, nonskeptical conversational context, the epistemic standards are lowered, and utterances of premise (2) are false. Thus, knowledge-attributions in ordinary conversational contexts are not threatened by the skeptical argument.

One problem for contextualism is that it is hard to coherently state the view. For example, I cannot now correctly say that Michael Jordan knows that he has hands, since I am currently involved in a skeptical (written) conversational context. What I must instead say is that neither I, nor anybody else, knows that he has hands. I cannot even justifiably say that some ordinary-context utterances of "Michael Jordan knows that he has hands" are true, relative to the low epistemic standards in effect in such contexts. This is because I, in my present context, do not know whether anyone has hands.

Another problem for contextualism is that it seems to imply that speakers are mistaken about the very meanings of their knowledge-attributing sentences. That is, suppose that I think that the skeptical argument is compelling and yet at the same time find its conclusion to be repugnant: it just can't be true that I do not know that I have hands. This means that I am failing to realize that the sentence stating the argument's conclusion is perfectly true as uttered in my current philosophical context. This betrays a misunderstanding of what my sentence means when used in the philosophical context.

See also Epistemology; Reliabilism; Verifiability Principle.

Bibliography

Ayer, Alfred Jules. *Language, Truth, and Logic*. New York: Dover, 1952. Verificationism.

Berkeley, George. *Three Dialogues between Hylas and Philonous*. Indianapolis: Bobbs-Merrill, 1954.

Brewer, Bill. *Perception and Reason*. Oxford: Oxford University Press, 1999. Perceptual beliefs.

Brueckner, Anthony. "Semantic Answers to Skepticism." *Pacific Philosophical Quarterly* 72 (1992): 200–201. Semantic externalism.

Brueckner, Anthony. "The Structure of the Skeptical Argument." *Philosophy and Phenomenological Research* 54 (1994): 827–835. *Underdetermination Principle*.

Cohen, Stewart. "How to be a Fallibilist." *Philosophical Perspectives* 2 (1988): 91–123. Contextualism.

Davidson, Donald. "A Coherence Theory of Truth and Knowledge." In *Truth and Interpretation*, edited by Ernest LePore. Oxford: Oxford University Press, 1986. Semantic externalism.

Descartes, René. *Meditations on First Philosophy*. In *The Philosophical Works of Descartes*. Vol. 1, translated by Haldane and Ross. New York: Dover, 1955. The Cartesian skeptical argument.

Dretske, Fred. "Epistemic Operators." *Journal of Philosophy* 67 (1970): 1007–1023. Denying the closure principle.

Dummett, Michael. "What Is a Theory of Meaning? (II)." In *Truth and Meaning*, edited by G. Evans and J. McDowell. Oxford: Oxford University Press, 1976. Antirealism.

Goldman, Alvin. "What Is Justified Belief?" In *Justification and Knowledge*, edited by G. Pappas. Dordrecht, Holland: Kluwer, 1979. Reliabilism.

Huemer, Michael. *Skepticism and the Veil of Perception*. Lanham, MD: Rowman and Littlefield, 2001. Perceptual beliefs; direct realism.

Kant, Immanuel. *Critique of Pure Reason*. Translated by N. K. Smith. New York: St. Martin's Press, 1965. Transcendental arguments.

McDowell, John. "Singular Thought and the Extent of Inner Space." In *Subject, Thought and Context*, edited by P. Pettit and J. McDowell. Oxford: Oxford University Press, 1986. Disjunctive view.

Nozick, Robert. *Philosophical Explanations*. Chap. 3. Cambridge, MA: Belknap Press of Harvard University Press, 1981). Denying the closure principle; tracking condition.

Putnam, Hilary. "Meaning and Reference." *Journal of Philosophy* 70 (1973): 699–711. Semantic externalism.

Putnam, Hilary. *Reason, Truth, and History*. Chap. 1. Cambridge, MA: 1981. Semantic externalism.

Schiffer, Stephen. "Contextualist Solutions to Skepticism." *Proceedings of the Aristotelian Society* 96 (1996): 317–333. Contextualism.

Stroud, Barry. "Transcendental Arguments." *Journal of Philosophy* 65 (1968): 241–256. Transcendental arguments.

Vogel, Jonathan. "Cartesian Skepticism and Inference to the Best Explanation." *Journal of Philosophy* 87 (1990): 658–666.

Anthony Brueckner (1996, 2005)

SKEPTICISM, HISTORY OF

Skepticism (also spelled "Scepticism") is the philosophical attitude of doubting knowledge claims set forth in various areas. Skeptics have challenged the adequacy or reliability of these claims by asking what they are based upon or what they actually establish. They have raised the question whether such claims about the world are either indubitable or necessarily true, and they have challenged the alleged grounds of accepted assumptions. Practically everyone is skeptical about some knowledge claims; but the skeptics have raised doubts about any knowledge beyond the contents of directly felt experience. The original Greek meaning of *skeptikos* was "an inquirer," someone who was unsatisfied and still looking for truth.

From ancient times onward skeptics have developed arguments to undermine the contentions of dogmatic philosophers, scientists, and theologians. The skeptical arguments and their employment against various forms of dogmatism have played an important role in shaping both the problems and the solutions offered in the course of western philosophy. As ancient philosophy and science developed, doubts arose about basic accepted views of the world. In ancient times skeptics challenged the claims of Platonism, Aristotelianism, and Stoicism, and in the Renaissance those of Scholasticism and Calvinism. After René Descartes, skeptics attacked Cartesianism and other theories justifying the "new science." Later, a skeptical offensive was leveled against Kantianism and then against Hegelianism. Each skeptical challenge led to new attempts to resolve the difficulties. Skepticism, especially since the Enlightenment, has come to mean disbelief—primarily religious disbelief—and the skeptic has often been likened to the village atheist.

VARIOUS SENSES AND APPLICATIONS

Skepticism developed with regard to various disciplines in which men claimed to have knowledge. It was questioned, for example, whether one could gain any certain knowledge in metaphysics (the study of the nature and significance of being as such) or in the sciences. In ancient times a chief form was medical skepticism, which questioned whether one could know with certainty either the causes or cures of diseases. In the area of ethics, doubts were raised about accepting various mores and customs and about claiming any objective basis for making value distinctions. Skepticisms about religion have questioned the doctrines of different traditions. Certain philosophies, like those of David Hume and Immanuel Kant, have seemed to show that no knowledge can be gained beyond the world of experience and that one cannot discover the causes of phenomena. Any attempt to do so, as Kant argued, leads to antinomies, contradictory knowledge claims. A dominant form of skepticism, the subject of this article, concerns knowledge in general, questioning whether anything actually can be known

with complete or adequate certainty. This type is called epistemological skepticism.

Kinds of epistemological skepticism can be distinguished in terms of the areas in which doubts are raised; that is, whether they be directed toward reason, toward the senses, or toward knowledge of things-in-themselves. They can also be distinguished in terms of the motivation of the skeptic—whether he or she is challenging views for ideological reasons or for pragmatic or practical ones to attain certain psychological goals. Among the chief ideological motives have been religious or antireligious concerns. Some skeptics have challenged knowledge claims so that religious ones could be substituted—on faith. Others have challenged religious knowledge claims in order to overthrow some orthodoxy. Kinds of skepticism also can be distinguished in terms of how restricted or how thoroughgoing they are—whether they apply only to certain areas and to certain kinds of knowledge claims or whether they are more general and universal.

ANCIENT SKEPTICISM

Historically, skeptical philosophical attitudes began to appear in pre-Socratic thought. In the fifth century BCE, the Eleatic philosophers, known for reducing reality to a static One, questioned the reality of the sensory world, of change and plurality, and denied that reality could be described in the categories of ordinary experience. On the other hand, the Ephesian philosopher of change Heraclites and his pupil Cratylus thought that the world was in such a state of flux that no permanent, unchangeable truth about it could be found; and Xenophanes, a wandering poet and philosopher, doubted whether man could distinguish true from false knowledge.

A more developed skepticism appeared in some of Socrates' views and in several of the Sophists. Socrates, in the early Platonic dialogues, was always questioning the knowledge claims of others; and in the *Apology*, he said that all that he really knew was that he knew nothing. Socrates' enemy, the Sophist Protagoras, contended that man is the measure of all things. This thesis was taken as a kind of skeptical relativism: no views are ultimately true, but each is merely one man's opinion. Another Sophist, Gorgias, advanced the skeptical-nihilist thesis that nothing exists; and if something did exist, it could not be known; and if it could be known, it could not be communicated.

ACADEMIC SKEPTICISM. Academic skepticism, so-called because it was formulated in the Platonic Academy in the third century BCE, developed from the Socratic observation, "All I know is that I know nothing." Its theoretical formulation is attributed to Arcesilas (c. 315–241 BCE) and Carneades (c. 213–129 BCE), who worked out a series of arguments, directed primarily against the knowledge claims of the Stoic philosophers, to show that nothing could be known. As these arguments have come down to us, especially in the writings of Cicero, Diogenes Laertius, and Saint Augustine, the aim of the Academic skeptical philosophers was to show, by a group of arguments and dialectical puzzles, that the dogmatic philosopher (that is, the philosopher who asserted that he knew *some* truth about the real nature of things), could not know with absolute certainty the propositions he said he knew. The Academics formulated a series of difficulties to show that the information we gain by means of our senses may be unreliable, that we cannot be certain that our reasoning is reliable, and that we possess no guaranteed criterion or standard for determining which of our judgments is true or false.

The basic problem at issue is that any proposition purporting to assert some knowledge about the world contains some claims that go beyond the merely empirical reports about what appears to us to be the case. If we possessed any knowledge, this would mean for the skeptics, that we knew a proposition, asserting some non-empirical, or trans-empirical claim, which we were certain could not possibly be false. If the proposition might be false, then it would not deserve the name of knowledge, but only that of opinion, i.e., that it might be the case. Since the evidence for any such proposition would be based, according to the skeptics, on either sense information or reasoning, and both of these sources are unreliable to some degree, and no guaranteed or ultimate criterion of true knowledge exists, or is known, there is always some doubt that any non-empirical or trans-empirical proposition is absolutely true and hence constitutes real knowledge. As a result, the Academic skeptics said that nothing is certain. The best information we can gain is only probable and is to be judged according to probabilities. Hence, Carneades developed a type of verification theory and a type of probabilism that is somewhat similar to the theory of scientific '"knowledge" of present-day pragmatists and positivists.

The skepticism of Arcesilas and Carneades dominated the philosophy of the Platonic Academy until the first century before Christ. In the period of Cicero's studies, the Academy changed from skepticism to the eclecticism of Philo of Larissa and Antiochus of Ascalon. The arguments of the Academics survived mainly through Cicero's presentation of them in his *Academica* and *De*

Natura Deorum, and through their refutation in St. Augustine's *Contra Academicos*, as well as in the summary given by Diogenes Laertius. The locus of skeptical activity, however, moved from the Academy to the school of the Pyrrhonian skeptics, which was probably associated with the Methodic school of medicine in Alexandria.

THE PYRRHONIAN SCHOOL. The putative father of Greek skepticism is Pyrrho of Elis (c. 360–c. 272 BCE) and his student Timon (c. 315–225 BCE). He avoided committing himself to any views about what was actually going on and acted only according to appearances. In this way he sought happiness or at least mental peace. The stories about Pyrrho that are reported indicate that he was not a theoretician, but rather a living example of the complete doubter, the man who would not commit himself to any judgment that went beyond what seemed to be the case. His interests seem to have been primarily ethical and moral, and in this area he tried to avoid unhappiness that might be due to the acceptance of value theories and to judging according to them. If such value theories were to any degree doubtful, accepting them and using them could only lead to mental anguish.

Pyrrhonism, as a theoretical formulation of skepticism, is attributed to Aenesidemus (c. 100–40 BCE). The Pyrrhonists considered that both the Dogmatists and the Academics asserted too much, one group saying, "Something can be known," the other that "Nothing can be known." Instead, the Pyrrhonians proposed to suspend judgment on all questions on which there seemed to be conflicting evidence, including the question whether or not something could be known.

Building on the type of arguments developed by Arcesilas and Carneades, Aenesidemus and his successors put together a series of "Tropes" or ways of proceeding to bring about suspense of judgment on various questions. In the sole surviving texts from the Pyrrhonian movement, those of Sextus Empiricus, these are presented in groups of ten, eight, five, and two tropes, each set offering reasons why one should suspend judgment about knowledge claims that go beyond appearances. The Pyrrhonian skeptics tried to avoid committing themselves on any and all questions, even as to whether their arguments were sound. Skepticism for them was an ability, or mental attitude, for opposing evidence both pro and con on any question about what was nonevident, so that one would suspend judgment on the question. This state of mind then led to a state of *ataraxia*, quietude, or unperturbedness, in which the skeptic was no longer concerned or worried about matters beyond appearances. Skepticism was a cure for the disease called Dogmatism or rashness. But, unlike Academic skepticism, which came to a negative dogmatic conclusion from its doubts, Pyrrhonian skepticism made no such assertion, merely saying that skepticism is a purge that eliminates everything including itself. The Pyrrhonist, then, lives undogmatically, following his natural inclinations, the appearances of which he is aware, and the laws and customs of his society, without ever committing himself to any judgment about them.

The Pyrrhonian movement flourished up to about 200 CE, the approximate date of Sextus Empiricus, and flourished mainly in the medical community around Alexandria as an antidote to the dogmatic theories, positive and negative, of other medical groups. The position has come down to us principally in the writings of Sextus Empiricus in his *Hypotyposes (Outlines of Pyrrhonism)* and the larger *Adversus mathematicos*, in which all sorts of disciplines from logic and mathematics to astrology and grammar are subjected to skeptical devastation. In his *Outlines of Pyrrhonism* and *Adversus mathematicos*, Sextus presented the tropes developed by previous Pyrrhonists. The ten tropes attributed to Aenesidemus showed the difficulties to be encountered in ascertaining the truth or reliability of judgments based on sense information, owing to the variability and differences of human and animal perceptions.

Other arguments raised difficulties in determining whether there are any reliable criteria or standards—logical, rational, or otherwise—for judging whether anything is true or false. To settle any disagreement, a criterion seems to be required. Any purported criterion, however, would appear to be based on another criterion, thus requiring an infinite regress of criteria, or else it would be based upon itself, which would be circular. Sextus offered arguments to challenge any claims of dogmatic philosophers to know more than what is evident; and in so doing he presented in one form or another practically all of the skeptical arguments that have ever appeared in subsequent philosophy.

Sextus said that his arguments were aimed at leading people to a state of *ataraxia* (unperturbability). People who thought that they could know reality were constantly disturbed and frustrated. If they could be led to suspend judgment, however, they would find peace of mind. In this state of suspension they would neither affirm nor deny the possibility of knowledge but would remain peaceful, still waiting to see what might develop. The Pyrrhonist did not become inactive in this state of suspense but lived undogmatically according to appearances, customs, and natural inclinations.

MEDIEVAL SKEPTICISM

Pyrrhonism ended as a philosophical movement in the late Roman Empire, as religious concerns became paramount. In the Christian Middle Ages the main surviving form of skepticism was the Academic, described in St. Augustine's *Contra academicos*. Augustine, before his conversion, had found Cicero's views attractive and had overcome them only through revelation. With faith, he could seek understanding. Augustine's account of skepticism and his answer to it provided the basis for medieval discussions.

In Islamic Spain, where there was more contact with ancient learning, a form of antirational skepticism developed among Muslim and Jewish theologians. Al-Ghazālī, an Arab theologian of the eleventh and early twelfth centuries, and his Jewish contemporary Judah ha-Levi (c. 1075/c. 1085–c. 1141), who was a poet and physician as well as a philosopher, offered skeptical challenges (much like those later employed by the occasionalist Nicolas Malebranche and by David Hume) against the contemporary Aristotelians in order to lead people to accept religious truths in mystical faith. This view that truth in religion is ultimately based on faith rather than on reasoning or evidence—what is known as fideism—also appears in the late Middle Ages in the German cardinal and philosopher Nicolaus of Cusa's advocacy of learned ignorance as the way to religious knowledge.

Another line of thinking that includes skeptical elements was that of the followers of William of Ockham (1285–1347) in the fourteenth century, who were exploring the consequences of accepting divine omnipotence and a divine source for all knowledge. They examined puzzles about whether God could deceive mankind, regardless of the evidence, and could make all human reasoning open to question.

MODERN SKEPTICISM

Modern Skepticism emerged in part from some of the Ockhamite views but mainly from the rediscovery of the skeptical classics. Very little of the Pyrrhonnian tradition had been known in the Middle Ages, but in the fifteenth century the texts of Sextus Empiricus in Greek were brought from the Byzantine Empire into Italy. Sextus' *Outlines of Pyrrhonism* was published in Latin in 1562, his *Adversus matematicos* in 1569, and the Greek texts of both in 1621. Interest in Cicero was revived and his *Academica* and *De natura deorum* were also published in the sixteenth century.

The voyages of exploration; the humanistic rediscovery of the learning of ancient Greece, Rome, and Palestine; and the new science—all combined to undermine confidence in man's accepted picture of the world. The religious controversy between the Protestants and Catholics raised fundamental epistemological issues about the bases and criteria of religious knowledge.

RENAISSANCE AND REFORMATION. Toward the end of the fifteenth century, there was a revival of interest in ancient skepticism among Florentine humanists. Politian was lecturing on philosophy using notes from Sextus with which he had recently become acquainted from manuscripts brought from Byzantium. Humanists, including Gianfrancesco Pico della Mirandola, were acquiring and studying Sextus' texts. Some of these manuscripts were deposited in the convent of San Marco where the Dominican friar and prophet Girolamo Savonarola was heading up an exciting intellectual forum in which ancient philosophies were being analyzed. Savonarola, who did not read Greek, asked two of his monks to prepare a Latin translation of Sextus from one of these manuscripts. This apparently was to be used as a weapon against philosophy independent of religion. Before Savonarola's project could be completed the convent was destroyed and he was executed.

Gianfrancesco Pico, one of Savonarola's disciples and the nephew of the great Pico della Mirandola, published the first work using skepticism as a way of challenging all of philosophy. Gianfrancesco Pico's *Examen Vanitatis* (1520) is the first work to present Sextus in Latin for the European audience. In 1562 Henri Estienne (Stephanus) published a Latin translation of the *Pyrrhoniarum Hypotyposes* in Paris, and in 1569 Gentian Hervet published a Latin translation of *Adversus Mathematicos* in Antwerp. The Greek texts were first printed at Cologne, Paris, and Geneva in 1621. Some texts of Sextus appeared in English in 1592 in a work attributed to Sir Walter Raleigh titled "The Skepticke." A full translation of Book One of Sextus appeared in 1659 in Thomas Stanley's *History of Philosophy*; instead of explaining skepticism he just presented the whole book to the readers. A French translation was started by Pierre Gassendi's disciple Samuel Sorbière but was never finished or published. The first complete French translation, by Claude Huart, did not appear until 1725.

RELIGIOUS CONTROVERSY: ERASMUS AND LUTHER. The skeptical issue became more central when raised in the debate between Erasmus and Martin Luther. Using Academic skeptical materials, Erasmus insisted that the

issues in dispute could not be resolved and that one should therefore suspend judgment and remain with the church. In 1524, Erasmus finally published a work, *De Libero Arbitrio*, attacking Martin Luther's views on free will. Erasmus' general anti-intellectualism and dislike of rational theological discussions led him to suggest a kind of skeptical basis for remaining within the Catholic Church. This contempt for intellectual endeavor was coupled with his advocacy of a simple, non-theological Christian piety. Theological controversies were not Erasmus' meat, and he states that he would prefer to follow the attitude of the skeptics and suspend judgment, especially where the inviolable authority of Scripture and the decrees of the Church permit. He says he is perfectly willing to submit to the decrees, whether or not he understands them or the reasons for them.

Scripture is not as clear as Luther would have us believe, and there are some places that are just too shadowy for human beings to penetrate. Theologians have argued and argued the question without end. Luther claims he has found the right answer and has understood Scripture correctly. But how can we tell that he really has? Other interpretations can be given that seem much better than Luther's. In view of the difficulty in establishing *the* true meaning of Scripture concerning the problem of free will, why not accept the traditional solution offered by the Church? Why start such a fuss over something one cannot know with any certainty? For Erasmus, what is important is a simple, basic, Christian piety, a Christian spirit. The rest, the superstructure of the essential belief, is too complex for a man to judge. Hence it is easier to rest in a skeptical attitude, and accept the age-old wisdom of the Church on these matters, than to try to understand and judge for oneself.

This attempt, early in the Reformation, at a skeptical "justification" of the Catholic rule of faith brought forth a furious answer from Luther, the *De Servo Arbitrio* of 1525. Erasmus' book, Luther declared, was shameful and shocking, the more so since it was written so well and with so much eloquence. *De Libero Arbitrio* begins with the announcement that the problem of the freedom of the will is one of the most involved of labyrinths. The central error of Erasmus' book, according to Luther, was that Erasmus did not realize that a Christian cannot be a skeptic. Christianity involves the affirmation of certain truths because one's conscience is completely convinced of their veracity. The content of religious knowledge, according to Luther, is far too important to be taken on trust. One must be absolutely certain of its truth. Hence, Christianity is the complete denial of skepticism. To find the truths, one only has to consult Scripture.

Of course there are parts that are hard to understand, and there are things about God that we do not, and perhaps shall not, know. But this does not mean that we cannot find the truth in Scripture. The central religious truth can be found in clear and evident terms, and these clarify the more obscure ones. However, if many things remain obscure to some people, it is not the fault of Scripture, but of the blindness of those who have no desire to know the revealed truths. Luther's view, and later that of Calvin, proposed a new criterion—that of inner experience—while the Catholics of the Counter-Reformation employed Pyrrhonian and Academic arguments to undermine the criterion. Following after Erasmus, H. C. Agrippa von Nettesheim, a stormy occult philosopher and physician, employed the skeptical arguments against Scholasticism, Renaissance Naturalism, and many other views to win people to the "true religion."

HERVET. Gentian Hervet, secretary to the Cardinal of Lorraine, and participant at part of the Council of Trent, linked his work on Sextus with what Gianfrancesco Pico had earlier done. During the 1560s, Hervet, a humanist, fought intellectually against the encroachments of Calvinism, challenging various Protestants to debate with him, and publishing many pamphlets against their views. He saw Sextus' work as ideal for demolishing this new form of heretical dogmatism, that of the Reformer. If nothing can be known, then, he insisted, Calvinism cannot be known. The only certainty we can have is God's Revelation. Skepticism, by controverting all human theories, will cure people from dogmatism, give them humility, and prepare them to accept the doctrine of Christ. Hervet's employment of Pyrrhonism against Calvinism was soon to be shaped into a skeptical machine of war for use by the Counter-Reformation. This view of Pyrrhonism, by one of the leaders of French Catholicism, was to set the direction of one of its major influences on the next three-quarters of a century.

MONTAIGNE AND SANCHES. The new concern with skepticism was given a general philosophical formulation by Michel de Montaigne and his cousin Francisco Sanches. Michel de Montaigne was the most significant figure in the sixteenth century revival of ancient skepticism. Not only was he the best writer and thinker of those who were interested in the ideas of the Academics and Pyrrhonians, but he was also the one who felt most fully the impact of the Pyrrhonian arguments of complete doubt—and its relevance to the religious debates of the

time. Montaigne was simultaneously a creature of the Renaissance and the Reformation. He was a thorough-going humanist, with a vast interest in, and concern with, the ideas and values of Greece and Rome, and their application to the lives of men in the rapidly changing world of sixteenth-century France. Montaigne was sent to the Collège de Guyenne in 1539 when he was six years old and was there for the next seven years. The college reflected the religious tensions of the time. Two of its leaders were André de Gouvea, a Portuguese New Christian, and George Buchanan, the Scottish Latin poet.

Montaigne's 1576 essay "Apologie of Raimond Sebond" unfolds in his inimitable rambling style as a series of waves of skepticism, with occasional pauses to consider and digest various levels of doubt, but with the overriding theme an advocacy of a new form of fideism—Catholic Pyrrhonism. The essay begins with a probably inaccurate account of how Montaigne came to read and translate the audacious work of the fifteenth century Spanish theologian, Raimond Sebond. Starting from a quibble about the validity of the arguments of Sebond, Montaigne moved to a general skeptical critique of the possibility of human beings understanding anything. In a rather back-handed manner, Montaigne excuses Sebond's theological rationalism by saying that although he, Montaigne, is not versed in theology, it is his view that religion is based solely on faith given to us by the Grace of God; true religion can only be based on faith, and any human foundation for religion is too weak to support divine knowledge. If human beings had the real light of faith, then human means, like the arguments of Sebond, might be of use. Montaigne explored the human epistemological situation and showed that man's knowledge claims in all areas were extremely dubious and so made pure faith the cornerstone of religion. Montaigne recommended living according to nature and custom and accepting whatever God reveals.

Sanches, in *Quod nihil scitur*, also written in 1576, advocated recognizing that nothing can be known and then trying to gain what limited information one can through empirical scientific means. In his book, Sanches develops his skepticism by means of an intellectual critique of Aristotelianism, rather than by an appeal to the history of human stupidity and the variety and contrariety of previous theories. Sanches begins by asserting that he does not even know if he knows nothing. Then he proceeds, step by step, to analyze the Aristotelian conception of knowledge to show why this is the case.

Every science begins with definition and definitions are nothing but names arbitrarily imposed upon things in a capricious manner, having no relation to the things named. The names keep changing, so that when we think we are saying something about the nature of things by means of combining words and definitions, we are just fooling ourselves. And if the names assigned to an object such as man, like "rational animal," all mean the same thing, then they are superfluous and do not help to explain what the object is. On the other hand, if the names mean something different from the object, then they are not the names of the object. By means of such an analysis, Sanches worked out a thorough-going nominalism.

Sanches' first conclusion was the usual fideistic one of the time—that truth can be gained only by faith. His second conclusion was to play an important role in later though: just because nothing can be known in an ultimate sense, we should not abandon all attempts at knowledge but should try to gain what knowledge we can, namely, limited, imperfect knowledge of some of those things with which we become acquainted through observation, experience and judgment. The realization that *nihil scitur* ("nothing is known") thus can yield some constructive results. This early formulation of "constructive" or "mitigated" skepticism was to be developed into an important explication of the new science by Marin Mersenne, Pierre Gassendi, and the leaders of the Royal Society.

THE SEVENTEENTH CENTURY

Montaigne's skepticism was extremely influential in the early seventeenth century. His followers, Pierre Charron in *De la Sagesse* (1601) and Jean-Pierre Camus in *Essay sceptique* (1603), became most popular in the early seventeenth century, especially among the avant-garde intellectuals in Paris. The so-called libertines, including Gabriel Naudé, Mazarin's secretary; Guy Patin, rector of the Sorbonne medical school; and François La Mothe Le Vayer, teacher of the dauphin, espoused Montaigne's attitude and were often accused of being skeptical even of fundamental religious tenets. Others, like François Veron, used the arguments of Sextus and Montaigne to challenge the Calvinist claim of gaining true knowledge from reading Scripture. French Counter-Reformers, by raising skeptical epistemological problems about whether one could determine what book is the Bible, what it actually says, what it means, and so on, forced Calvinists to seek an indisputable basis for knowledge as a prelude to defending their theological views.

GASSENDI AND MERSENNE. In the 1620s efforts to refute or mitigate this new skepticism appeared. Some authors simply stated that Aristotle would have resolved the difficulties by applying his theory of sense perception and knowledge to the problems raised. Others, like François Garasse, decried the irreligious tendencies they discerned in all this doubting. Still others, like Francis Bacon, tried to overcome the skeptical difficulties by appealing to new methods and new instruments that might correct errors and yield firm and unquestionable results. Herbert of Cherbury, in *De Veritate* (1624), offered an elaborate scheme for overcoming skepticism which combined Aristotelian and Stoic elements, and ultimately appealed to common notions, or truths known by all men, as the criteria by which reliable and indubitable judgment would be possible.

Perhaps the most forceful presentation of skepticism in the early seventeenth century is Pierre Gassendi's earliest work, *Exercitationes Paradoxicae Adversus Aristoteleos* (1624). A Christian Epicurean, Gassendi, himself originally a skeptic, challenged almost every aspect of Aristotle's view, as well as many other theories. He applied a battery of ancient and Renaissance skeptical arguments, concluding, "No science is possible, least of all in Aristotle's sense." In this work, Gassendi indicated in embryo what became his and Marin Mersenne's constructive solution to the skeptical crisis, the development of an empirical study of the world of appearances rather than an attempt to discover the real nature of things.

Mersenne, one of the most influential figures in the intellectual revolution of the times, while retaining epistemological doubts about knowledge of reality, yet recognized that science provided useful and important information about the world. Mersenne granted that the problems raised by Sextus could not be answered and that, in a fundamental sense, knowledge of the real nature of things cannot be attained. However, he insisted, information about appearances and deductions from hypotheses can provide an adequate guide for living in this world and can be checked by verifying predictions about futures experiences. Gassendi, in his later works, developed this constructive skepticism as a *via media* between complete doubt and dogmatism, and offered his atomic theory as the best hypothetical model for interpreting experience. Mersenne and Gassendi combined skepticism about metaphysical knowledge of reality with a way of gaining useful information about experience through a pragmatic scientific method. The constructive skepticisms of Gassendi and Mersenne, and later of members of the Royal Society of England like Bishop John Wilkins and Joseph Glanvill, thus developed the attitude of Sanches into a hypothetical, empirical interpretation of the new science.

DESCARTES. René Descartes offered a fundamental refutation of the new skepticism, contending that, by applying the skeptical method of doubting all beliefs that could possibly be false (due to suffering illusions or being misled by some power), one would discover a truth that is genuinely indubitable, namely, "I think, therefore I am" (*cogito ergo sum*), and that from this truth one could discover the criterion of true knowledge, namely, that whatever is clearly and distinctly conceived is true. Using this criterion, one could then establish: God's existence, that he is not a deceiver, that he guarantees our clear and distinct ideas, and that an external world exists that can be known through mathematical physics. Descartes, starting from skepticism, claimed to have found a new basis for certitude and for knowledge of reality.

REPLIES TO DESCARTES. Throughout the seventeenth century skeptical critics—Mersenne, Gassendi, the reviver of Academic philosophy Simon Foucher, and Pierre-Daniel Huet, one of the most learned men of the age—sought to show that Descartes had not succeeded and that, if he sincerely followed his skeptical method, his new system could only lead to complete skepticism. They challenged whether the *cogito* proved anything, or whether it was indubitable; whether Descartes' method could be successfully applied, or whether it was certain; and whether any of the knowledge claims of Cartesianism were *really* true. Nicolas Malebranche, the developer of occasionalism, revised the Cartesian system to meet the skeptical attacks only to find his efforts challenged by the new skeptical criticisms of Foucher and by the contention of the Jansenist philosopher Antoine Arnauld that Malebranchism led to a most dangerous Pyrrhonism.

Huet's *Censura Philosophae Cartesiana* (1689) and his unpublished defense of it raised doubts about each element of the proposition, "I think, therefore perhaps I may be." Gassendi, Huet, and others questioned whether Descartes' criterion could determine what was true or false. Could we really tell what was clear and distinct, or could we only tell that something appeared clear and distinct to us? Mersenne pointed out that even with the criterion we could not be sure that what was clear and distinct to us, and hence true, was really true for God. Hence, in an ultimate sense, even the most certain Cartesian knowledge might be false. Gassendi, in what Descartes called the "objections of objections," pointed out that for all anyone could ascertain, the whole Carte-

sian system of truths might be only a subjective vision in somebody's mind and not a true picture of reality. Huet argued that since all the fundamental Cartesian data consisted of ideas, and ideas are not real physical things, the Cartesian world of ideas, even if clear and distinct, cannot represent something quite different from itself.

FOLLOWERS OF DESCARTES. As Cartesianism was attacked from many sides, adherents modified it in various ways. The radical revision of Nicolas Malebranche, designed partially to avoid skeptical difficulties involved in connecting the world of ideas with reality, was immediately attacked by the skeptic Simon Foucher. The orthodox Cartesian Antoine Arnaud claimed that Malebranchism could only lead to a most dangerous Pyrrhonism. Foucher, who wished to revive Academic skepticism, applied various skeptical gambits to Malebranche's theory, one of which was to be important in subsequent philosophy. He argued that the skeptical difficulties which Descartes and Malebranche used to deny that sense qualities (the so-called secondary qualities—color, sound heat, taste, smell) were features of real objects, applied as well to the mathematically describable primary qualities like extension and motion, which the Cartesians considered the fundamental properties of things. These mathematical qualities, as perceived, are as variable and as subjective as the others. If the skeptical arguments are sufficient to cause doubt about the ontological status of secondary qualities, Foucher contended, they are also sufficient to lead us to doubt that primary ones are genuine features of reality.

ENGLISH SKEPTICISM. Various English philosophers, culminating in John Locke, tried to blunt the force of skepticism by appealing to common sense and to the "reasonable" man's inability to doubt everything. They admitted that there might not be sufficient evidence to support the knowledge claims extending beyond immediate experience. But this did not actually require that everything be doubted; by using standards of common sense, an adequate basis for many beliefs could be found.

This theory of limited certitude was articulated especially by two figures, John Wilkins and Joseph Glanvill. The theory is a development from the earlier solution to the skeptical problems advanced by Sebastian Castillio and William Chillingworth. Wilkins set forth the theory of limited certainty as both an answer to dogmatism and to excessive skepticism. Wilkins completely rejected the dogmatists' outlook, and then offered a way of defusing the potentially disastrous results of complete skepticism. In order to find a moderate skeptical stance from which religion and science could flourish, Wilkins felt it was necessary to analyze what kind of certainty human beings could actually attain. The highest level of certainty, absolute infallible certainty, which could not possibly be false, is beyond human attainment. Only God has such certainty. The highest human level Wilkins called conditional infallible certainty. This requires that "our faculties be true, and that we do not neglect the exerting of them."

Glanvill saw the reliability of our faculties as central for avoiding any ultimate and overwhelming skepticism. Glanvill, like Wilkins, saw that the kind of certainty we would need to be absolutely sure of our faculties is unattainable—"for it may not be absolutely impossible, but that our Faculties may be so construed, as always to deceive us in the things we judg most certain and assured." We may not be able to attain infallible certitude, but we can attain indubitable certitude—that our faculties are true. This is indubitable in two senses—one, that we find we have to believe them, and, two, that we have no reason or cause for doubting them. In terms of this distinction, Wilkins, Glanvill, and their colleagues built up a theory of empirical science and jurisprudence for studying nature and deciding human problems within the limits of "reasonable doubt." Their limited skepticism appears in the Anglo-American theory of legal evidence and in the theory of science of the early Royal Society. They believed that by applying their probabilistic empirical method to religious questions they could justify a tolerant, latitudinarian form of Christianity.

OTHER RESOLUTIONS OF SKEPTICISM. Other answers were offered to the skeptics and to their challenge of some of the basic tenets of the new philosophy. Thomas Hobbes had admitted the force of the problem of finding *the* criterion for judging what was genuinely true, and he insisted that the solution was ultimately political—the sovereign would have to decide. Blaise Pascal in his scientific works gave one of the finest expositions of the hypothetical probabilistic nature of science and mathematics. Pascal, who presented the case for skepticism most forcefully in his *Pensées*, still denied that there can be a complete skepticism; for nature prevents it. Lacking rational answers to complete skepticism, man's only recourse lies in turning to God for help in overcoming doubts. Spinoza, on the other hand, with his completely rational vision of the world, could not regard skepticism as a serious problem. If one had clear and adequate ideas, there would be no need or excuse for doubting. Doubt was only an indication of lack of clarity, not of basic philosophical difficulties.

The philosopher who took the skeptics most seriously was Gottfried Wilhelm von Leibniz, and he was regarded as a closer friend intellectually by the skeptics of his age than any of the other metaphysicians of the period. Leibniz, although certainly not a philosophical skeptic, agrees with some of the major contentions of the skeptics, and is willing to admit, unlike other metaphysicians of the seventeenth century, that there are general, and perhaps unanswerable, objections that can be raised against any philosophical theory. The skeptics and Leibniz could agree on the major failings of Cartesianism, although they were hardly in agreement as to what to do about them. Leibniz and the skeptics were all humanists and found great value in the tradition of man's effort to understand his universe; hence they rejected the Cartesian attitude towards the past. In his discussions, especially with Simon Foucher and Pierre Bayle, Leibniz agreed that there are first principles of philosophical reasoning that have not been satisfactorily demonstrated.

Leibniz was willing to regard metaphysics as a hypothetical enterprise, that is, as an attempt to present theories which agree with the known facts, which avoid certain difficulties in previous theories, and which give a satisfactory or adequate explanation of the world that is experienced. In the debate with Pierre Bayle over the article "Rorarius," in Bayle's *Dictionnaire historique et critique*, Leibniz does not argue for his theory as the true picture of reality, but rather as the most consistent hypothesis to explain the known scientific facts and the general conclusions of the "new philosophers" about the relation of the mind and the body, and to avoid the "unfortunate" complications or conclusions of the views of Descartes, Malebranche, or Spinoza. Leibniz was unwilling to see these limitations on our knowledge as a reason for skeptical despair or to see these points as constituting a radical skepticism that cast whatever knowledge we had in any serious doubt. For Leibniz, whatever merits the skeptical arguments had, they did not have to lead to negative or destructive conclusions. At best, skepticism should be a spur to constructive theorizing, and not a reason for doubting or despairing of the possibility of knowledge.

BAYLE AND THE ENLIGHTENMENT. The culmination of seventeenth-century skepticism appears in the writings of Pierre Bayle, especially in his monumental *Dictionnaire historique et critique* (1697–1702). Bayle, a superb dialectician, challenged philosophical, scientific, and theological theories, both ancient and modern, showing that they all led to perplexities, paradoxes, and contradictions. He argued that the theories of Descartes, Spinoza, Leibniz, and Malebranche, when skeptically analyzed, cast in doubt all information about the world, even whether a world exists. Bayle skillfully employed skeptical arguments about such things as sense information, human judgments, logical explanations, and the criteria of knowledge in order to undermine confidence in human intellectual activity in all areas. Bayle suggested that man should abandon rational activity and turn blindly to faith and revelation; he can therefore only follow his conscience without any criterion for determining true faith. Bayle showed that the interpretations of religious knowledge were so implausible that even the most heretical views, like Manichaeism—known for its cosmic dualism of good and evil—and Atheism made more sense. As a result Bayle's work became "the arsenal of the Enlightenment," and he was regarded as a major enemy of religion.

Bayle, in his later works, indicated that he held some positive views even though he presented no answers to his skepticism. There is still much scholarly debate as to what his actual position was, but he influenced many people in the eighteenth century. His skeptical arguments were soon applied to traditional religion by Voltaire and others. But in place of Bayle's doubts or his appeal to faith, they offered a new way of understanding man's world—that of Newtonian science—and professed an inordinate optimism about what man could comprehend and accomplish through scientific examination and induction. Though Bayle remained the heroic figure who had launched the Age of Reason by criticizing all the superstitions of past philosophy and theology, the leaders of the Enlightenment, both in France and Britain, felt that his skepticism was *passé* and only represented the summit of human understanding before "God said, Let Newton be, and all was light."

THE EIGHTEENTH CENTURY

Most eighteenth-century thinkers gave up the quest for metaphysical knowledge after imbibing Bayle's arguments. George Berkeley, an Empiricist and Idealist, fought skeptical doubts by identifying appearance and reality and offering a spiritualistic metaphysics. He was immediately seen as just another skeptic, since he was denying the world beyond experience.

HUME. Bayle's chief eighteenth-century successor was David Hume. Combining empirical and skeptical arguments, Hume, in the *Treatise of Human Nature and the Enquiry Concerning Human Understanding*, charged that neither inductive nor deductive evidence could establish the truth of any matter of fact. Knowledge could only

consist of intuitively obvious matters or demonstrable relations of ideas but not of anything beyond experience; the mind can discover no necessary connections within experience nor any root causes of experience. Beliefs about the world are based not upon reason or evidence, nor even upon appeal to the uniformity of nature, but only on habit and custom. Basic beliefs cannot be justified by reasoning. Belief that there is an external world, a self, a God is common; but there is no adequate evidence for it. Although it is natural to hold these convictions, they are inconsistent and epistemologically dubious. "Philosophy would render us entirely Pyrrhonian, were not Nature too strong for it." The beliefs that a man is forced to hold enable him to describe the world scientifically, but when he tries to justify them he is led to complete skepticism. Before he goes mad with doubts, however, Nature brings him back to common sense, to unjustifiable beliefs. Hume's fideism was a natural rather than a religious one; it is only animal faith that provides relief from complete doubt. The religious context of skepticism from Montaigne to Bayle had been removed, and man was left with only his natural beliefs, which might be meaningless or valueless.

THE *PHILOSOPHES*. The French Enlightenment philosophers, the *philosophes*, built on the skeptical reading of Locke and Bayle, and on their interpretation of Berkeley as a radical skeptic. While they produced vast accumulations of new forms of knowledge, they also placed this alongside a skepticism about whether one could ever establish that this knowledge was about an external reality. Perhaps the most skeptical of them was the great French mathematician Marquis de Condorcet who held that mathematics, physics, and moral philosophies were all just probable. He also raised the possibility that our present mental faculties by which we judged our knowledge might change over time and, hence, that what we found true today might not be so tomorrow.

REID AND THE COMMON-SENSE SCHOOL. The central themes in Hume's skeptical analysis—the basis of induction and causality, knowledge of the external world and the self, proofs of the existence of God—became the key issues of later philosophy. Hume's contemporary Thomas Reid hoped to rebut Hume's skepticism by exposing it as the logical conclusion of the basic assumptions of modern philosophy from Descartes onward. Such disastrous assumptions should be abandoned for commonsensical principles that have to be believed. When the conclusions of philosophy run counter to common sense, there must be something wrong with philosophy. Since nobody could believe and act by complete skepticism, the fact that this skepticism was the consistent issue of the Cartesian and Lockean way of ideas only showed the need to start anew. Reid offered his common-sense realism as a way of avoiding Hume's skepticism by employing as basic principles the beliefs we are psychologically unable to doubt.

Hume was unimpressed by Reid's argument. As Hume and Kant saw, Reid had not answered Hume's skepticism but had only sidestepped the issue by appealing to commonsensical living. This provided, however, neither a theoretical basis for beliefs nor a refutation of the arguments that questioned them. The Scottish common-sense school of Oswald, Beattie, Stewart, Brown, and others kept reiterating its claim to have refuted Hume's skepticism by appealing to natural belief, while at the same time conceding that Hume's fundamental arguments could not be answered. Thomas Brown, an early-nineteenth-century disciple of Reid, admitted that Reid and Hume differed more in words than in opinions, saying, "'Yes,' Reid bawled out, 'we must believe in an outward world': but added in a whisper, 'we can give no reason for our belief.' Hume cries out, 'we can give no reason for such a notion': and whispers, 'I own that we cannot get rid of it.'"

THE GERMAN ENLIGHTENMENT AND KANT. The Scottish school was perhaps the first to make Hume's version of modern skepticism the central view to be combated if philosophy was to make coherent sense of man's universe. The more fundamental attempt, for subsequent philosophy, to deal with Hume's skepticism was developed in Germany in the second half of the eighteenth century and culminated in Kant's critical philosophy. Such leaders of the Prussian Academy as Jean Henry Samuel Formey, Johann Bernhard Mérian, and Johann Georg Sulzer had long been arguing against Pyrrhonism. They were among the first to read, translate (into French and German), and criticize Hume's writings. They saw in the skeptical tradition up to Bayle and Huet, and in Hume's version of it, a major challenge to all man's intellectual achievements. Although their answers to skepticism were hardly equal to the threat they saw in it, these writers helped revive interest in and concern with skepticism in an age that thought it had solved, or was about to solve, all problems. Others in Germany contributed to an awareness of the force of skepticism: Johann Christoff Eschenbach by his edition of the arguments of Sextus, Berkeley, and Arthur Collier (Berkeley's contemporary) against knowledge of an external corporeal world; Ernst Platner by his skeptical aphorisms and his German edi-

tion of Hume's *Dialogues on Natural Religion* (1781); hosts of German professors by dissertations against skepticism; and the translators of the Scottish critics of Hume.

Kant saw that Hume had posed a most fundamental challenge to all human knowledge claims. To answer him, it had to be shown not that knowledge is possible but *how* it is possible. Kant combined a skepticism toward metaphysical knowledge with the contention that certain universal and necessary conditions are involved in having experience and describing it. In terms of these it is possible to have genuine knowledge about the forms of all possible experience, space and time, and about the categories in which all experience is described. Any effort to apply this beyond all possible experience, however, leads into contradictions and skepticism. It is not possible to know about things-in-themselves nor about the causes of experience.

SKEPTICAL REJOINDERS TO KANT. Though Kant thought that he had resolved the skeptical problems, some of his contemporaries saw his philosophy as commencing a new skeptical era. G. E. Schulze (or Schulze-Aenesidemus) a notable critic of Kantianism, insisted that, on Kant's theory, no one could know any objective truths about anything; he could only know the subjective necessity of his views. So Schulze, by insisting on the inability of the Kantian analysis to move from subjective data about what people have to believe to any objective claims about reality, contended that Kant had not advanced beyond Hume's skepticism, and that this failure of the Kantian revolution actually constituted a vindication of Hume's views.

Salomon Maimon contended that, though there are such things as a priori concepts, their application to experience is always problematical, and whether they apply can only be found through experience. Hence, the possibility of knowledge can never be established with certainty. Assured truth on the basis of concepts is possible only of human creations, like mathematical ideas, and it is questionable whether these have any objective truth. Thus Maimon developed a mitigated Kantianism (to some extent like that of the Neo-Kantian movement a century later) in which the reality of a priori forms of thought is granted but in which the relation of these forms to matters of fact is always in question. Knowledge (that is, propositions that are universal and necessary, rather than ones that are just psychologically indubitable) is possible in mathematics but not in sciences dealing with the world. Unlike the logical positivists, who were to claim that mathematics was true because it consisted only of vacuous logical tautologies, Maimon contended that mathematics was true because it was about creations of our mind. Its objective relevance was always problematical.

Maimon's partial skepticism exposed some of the fundamental limitations of Kant's critical philosophy as a solution to the skeptical crisis. Developing the thesis that human creativity is the basis of truth, Johann Georg Hamann posited a new way of transcending skepticism. Hamann accepted Hume's and Kant's arguments as evidence that knowledge of reality cannot be gained by rational means but only by faith. Hamann exploited the skeptical thought of these philosophies to press for a complete antirational fideism. He used Hume's analyses of miracles and of the evidence for religious knowledge to try to convince Kant of the futility of the search for truth by rational means. During the height of nineteenth-century positivism, materialism, and idealism, Hamann's type of fideism was revitalized by Kierkegaard and in France by Catholic opponents of the French Revolution and liberalism—like Joseph de Maistre and H.-F.-R. Lamennais, who used it as a critique of French liberal, empirical, and Enlightenment views and as a new defense of orthodoxy and political conservatism. Kierkegaard brilliantly combined themes from Sextus, Hume, and Hamann to attack the rationalism of the Hegelians, to develop a thoroughgoing skepticism about rational achievements, and to show the need for faith in opposition to reason. Fideism has become a major element in twentieth-century neo-orthodox and existentialist theology, which tries to show that the traditional skeptical problems still prevent us from finding an ultimate basis for our beliefs except by faith.

IDEALISM. In the mainstream of philosophy after Kant, although skepticism continues to play a vital role, few philosophers have been willing to call themselves skeptics. The German metaphysicians, from Fichte and Hegel onward, sought to escape from the skeptical impasse produced by Hume and Kant and to reach knowledge of reality through the creative process and the recognition of historical development. They attempted to portray skepticism as a stage in the awareness and understanding of the process of events. For Fichte, skepticism made one recognize the need for commitment to a fundamental outlook about the world. The commitment to see the world in terms of creative thought processes led to a revelation of the structure of the universe as an aspect of the Absolute Ego.

For Hegel skepticism was the nadir of philosophy, actually its antithesis. According to Hegel, human knowledge is a historically developing process. At each stage of the process both our knowledge and the world itself are limited and contain contradictions, which are overcome at the next stage. Only the final, Absolute stage, when no further contradictions can be developed, permits genuine knowledge that is not partly true and partly false. Then, presumably, skepticism is no longer possible. The English Hegelian F. H. Bradley, in his *Appearance and Reality* (1893), used the traditional skeptical arguments to show that the world was unintelligible in terms of empirical or materialistic categories, and hence that one had to go beyond the world appearance to find true knowledge.

RECENT AND CONTEMPORARY PHILOSOPHY

Irrational skepticism was developed into Existentialism by Søren Kierkegaard in the nineteenth century. Using traditional skeptical themes to attack Hegelianism and liberal Christianity, Kierkegaard stressed the need for faith. Only by an unjustified and unjustifiable "leap into faith" could certainty be found—which would then be entirely subjective rather than objective. Modern neo-orthodox and Existentialist theologians have argued that skepticism highlights man's inability to find any ultimate truth except through faith and commitment. Nonreligious forms of this view have been developed by Existentialist writers like Albert Camus, combining the epistemological skepticism of Kierkegaard and Leon Shestov with the skepticism regarding religion and objective values of Friedrich Nietzsche.

In his *Myth of Sisyphus*, Camus portrays man as trying to measure the nature and meaning of an essentially absurd universe by means of questionable rational and scientific criteria. Camus regards the skeptical arguments used by Kierkegaard and Shestov as showing decisively the contradictory nature of human rational attempts to understand the world, but he rejects their fideistic solution: overcoming the skeptical crisis by "a leap into faith." Instead, he accepts Nietzsche's picture of the ultimate meaninglessness of the world because "God is dead." The rational and scientific examination of the world shows it to be unintelligible and absurd but it is necessary to struggle with it. It is thus through action and commitment that one finds whatever personal meaning one can, though it has no objective significance. The mythological Sisyphus, eternally pushing a huge rock uphill, only to have to fall to the bottom again, typifies the human situation. He does not expect to find truth, nor does he expect to end his struggle. He finds no ultimate point or value in his situation, but he perseveres with a "silent joy," realizing that his struggle has meaning only for him, in terms of his human condition. The struggle is neither sterile nor futile for him, though it is meaningless in terms of understanding or possible achievement.

George Santayana, an American critical Realist, in *Scepticism and Animal Faith*, presented a naturalistic skepticism. Any interpretation of immediate or intuited experience is open to question. To make life meaningful, however, men make interpretations by "animal faith," according to biological and social factors. The resulting beliefs, though unjustified and perhaps illusory, enable them to persevere and find the richness of life. When the full force of complete skepticism is realized, Santayana claimed, one can appreciate what is in fact absolutely indubitable, the immediately experienced or intuited qualities that Santayana called "essences." The interpretation of these essences leads to various questionable metaphysical systems. A thoroughgoing skepticism makes one realize the unjustifiable assumptions involved in interpreting the realm of essences, and also that we do interpret them and thereby construct meaningful pictures of the world. Santayana called the process of interpretation "animal faith," which is consistent with complete skepticism and involves following natural and social tendencies and inclinations.

Types of skepticism also appear in logical positivism and various forms of linguistic philosophy. The attack on speculative metaphysics developed by the physicist and early Positivist Ernst Mach, Bertrand Russell, and Rudolf Carnap, a leader in the Vienna Circle, where logical positivism was nourished, incorporated a skepticism about the possibility of gaining knowledge beyond experience or logical tautologies. Russell and the important philosopher of science Karl Popper have further stressed the unjustifiability of the principle of induction, and Popper has criticized theories of knowledge based upon empirical verification. A founder of linguistic analysis, Fritz Mauthner, has set forth a skepticism in which any language is merely relative to its users and thus subjective. Every attempt to tell what is true just leads one back to linguistic formulations, not to objective states of affairs. The result is a complete skepticism about reality—a reality that cannot even be expressed except in terms of what he called godless mystical contemplation. Mauthner's linguistic skepticism bears some affinities to the views expressed in Ludwig Wittgenstein's *Tractatus Logico-Philosophicus*.

A different way of dealing with skepticism was set forth by the English philosopher, G. E. Moore at Cambridge. He contended that no matter what skeptical arguments may be they do not eliminate people's certitude about what they immediately perceive. There is a kind of "certain knowledge" that each of us has and can build on even though we know that it can be questioned in some theoretical way. Wittgenstein explored this kind of resolution in his essay *On Certainty* and sought to get beyond what Moore had done. Many contemporary philosophers are still writing and arguing about what constitutes knowledge and whether, in some way, we can find any basis for certainty.

POSTMODERNISM

A new, radical form of skepticism has developed in the last half century: postmodernism. This view challenges whether there can be any rational framework for discussing intellectual problems or whether the frameworks that people use are related to their life situations. Developing out of literary criticism and psychological investigations, the postmodernists have been undermining confidence in the investigation of the world in which we live by showing that the investigations are part of what needs to be scrutinized. Using ideas from Martin Heidegger, Michel Foucault, Jacques Derrida, Jean-François Lyotard, and Richard Rorty, they see philosophy and science as human activities to be judged in terms of their role in human life, rather than by some standard that can be said to be true or false. Rather than attempting to find a holistic truth or set of truths that are knowable and eternal, Postmodernists stress reflexivity, fragmentation, discontinuity, and ambiguity. Critics see this as a most dangerous development in that there will be no objective standpoint for evaluating theories. But that, of course, is part of the postmodernist outlook. Psychologists and sociologists have been adding to this view by stressing how intellectual outlooks vary according to sexual orientation, racial background, gender, and other fundamental features of human outlooks. Skepticism results from seeing that there is no objective standpoint from which to sort out the better or worse of these points of view.

CRITICISM AND EVALUATION. In Western thought, skepticism has raised basic epistemological issues. In view of the varieties of human experience, it has questioned whether it is possible to tell which are veridical. The variations that occur in different perceptions of what is presumed to be one object raise the question of which is the correct view. The occurrence of illusory experiences raises the question of whether it is really possible to distinguish illusions and dreams from reality. The criteria employed can be questioned and require justification. On what basis does one tell whether one has the right criteria? By other criteria? Then, are these correct? On what standards? The attempt to justify criteria seems either to lead to an infinite regress or to just stop arbitrarily. If an attempt is made to justify knowledge claims by starting with first principles, what are these based upon? Can it be established that these principles cannot possibly be false? If so, is the proof itself such that it cannot be questioned? If it is claimed that the principles are self-evident, can one be sure of this, sure that one is not deceived? And can one be sure that one can recognize and apply the principles correctly? Through such questioning, skeptics have indicated the basic problems that an investigator would have to resolve before he could be certain of possessing knowledge; that is, information that could not possibly be false.

Critics have contended that skepticism is both a logically and a humanly untenable view. Any attempt to formulate the position will be self-refuting since it will assert at least some knowledge claims about what is supposed to be dubious. Montaigne suggested that the skeptics needed a nonassertive language, reflecting the claim of Sextus that the skeptic does not make assertions but only chronicles his feelings. The strength of skepticism lies not in whether it can be stated consistently but upon the effects of its arguments on dogmatic philosophers. As Hume said, skepticism may be self-refuting, but in the process of refuting itself it undermines dogmatism. Skepticism, Sextus said, is like a purge that eliminates itself as well as everything else.

Critics have claimed that anyone who tried to be a complete skeptic, denying or suspending all judgments about ordinary beliefs, would soon be driven insane. Even Hume thought that the complete skeptic would have to starve to death and would walk into walls or out of windows. Hume, therefore, separated the doubting activity from natural practical activities in the world. Skeptical philosophizing went on in theory, while believing occurred in practice. Sextus and the contemporary Norwegian skeptic Arne Naess have said, on the other hand, that skepticism is a form of mental health. Instead of going mad, the skeptic—without commitment to fixed positions—can function better than the dogmatist.

Some thinkers like A. J. Ayer and J. L Austin have contended that skepticism is unnecessary. If knowledge is defined in terms of satisfying meaningful criteria, then knowledge is open to all. The skeptics have raised false problems, because it is, as a matter of fact, possible to tell that some experiences are illusory since we have criteria

for distinguishing them from actual events. We do resolve doubts and reach a state of knowledge through various verification procedures, after which doubt is meaningless. Naess, in his book *Scepticism*, has sought to show, however, that, on the standards offered by Ayer and Austin, one can still ask if knowledge claims may not turn out to be false and hence that skepticism has still to be overcome.

Skepticism throughout history has played a dynamic role in forcing dogmatic philosophers to find better or stronger bases for their views and to find answers to the skeptical attacks. It has forced a continued reexamination of previous knowledge claims and has stimulated creative thinkers to work out new theories to meet the skeptical problems. The history of philosophy can be seen, in part, as a struggle with skepticism. The attacks of the skeptics also have served as a check on rash speculation; the various forms of modern skepticism have gradually eroded the metaphysical and theological bases of European thought. Most contemporary thinkers have been sufficiently affected by skepticism to abandon the search for certain and indubitable foundations of human knowledge. Instead, they have sought ways of living with the unresolved skeptical problems through various forms of naturalistic, scientific, or religious faiths.

See also Aenesidemus; Agrippa von Nettesheim, Henricus Cornelius; al-Ghazālī, Muhammad; Ancient Skepticism; Antiochus of Ascalon; Aristotelianism; Aristotle; Arnauld, Antoine; Augustine, St.; Augustinianism; Austin, John Langshaw; Averroism; Ayer, Alfred Jules; Bacon, Francis; Bayle, Pierre; Beattie, James; Berkeley, George; Bradley, Francis Herbert; Brown, Thomas; Calvin, John; Camus, Albert; Carnap, Rudolf; Carneades; Cartesianism; Charron, Pierre; Cicero, Marcus Tullius; Collier, Arthur; Condorcet, Marquis de; Cratylus; Derrida, Jacques; Descartes, René; Diogenes Laertius; Enlightenment; Erasmus, Desiderius; Fichte, Johann Gottlieb; Fideism; Foucault, Michel; Foucher, Simon; Gassendi, Pierre; Glanvill, Joseph; Gorgias of Leontini; Greek Academy; Halevi, Yehuda; Hamann, Johann Georg; Hegel, Georg Wilhelm Friedrich; Hegelianism; Heidegger, Martin; Herbert of Cherbury; Huet, Pierre-Daniel; Hume, David; Kant, Immanuel; Kierkegaard, Søren Aabye; Lamennais, Hugues Félicité Robert de; La Mothe Le Vayer, François de; Leibniz, Gottfried Wilhelm; Locke, John; Logical Positivism; Luther, Martin; Lyotard, Jean Francois; Mach, Ernst; Maimon, Salomon; Maistre, Comte Joseph de; Malebranche, Nicolas; Mani and Manichaeism; Medieval Philosophy; Mersenne, Marin; Montaigne, Michel Eyquem de; Moore, George Edward; Moral Skepticism; Neo-Kantianism; Nicholas of Cusa; Nietzsche, Friedrich; Ockhamism; Pascal, Blaise; Philo of Larissa; Pico della Mirandola, Count Giovanni; Pico della Mirandola, Gianfrancesco; Platonism and the Platonic Tradition; Popper, Karl Raimund; Protagoras of Abdera; Pyrrho; Reformation; Reid, Thomas; Renaissance; Rorty, Richard; Russell, Bertrand Arthur William; Sanches, Francisco; Santayana, George; Schulze, Gottlob Ernst; Sextus Empiricus; Shestov, Lev Isaakovich; Skepticism, Contemporary; Socrates; Spinoza, Benedict (Baruch) de; Stewart, Dugald; Stoicism; Sulzer, Johann Georg; Timon of Phlius; William of Ockham; Wittgenstein, Ludwig Josef Johann; Xenophanes of Colophon.

Bibliography

The basic statements and arguments of various forms of skepticism are given in:

Academic skepticism: Cicero, *Academica* and *De natura deorum*, both with trans. by H. Rackham, Loeb Classical Library (1956).

Pyrrhonian skepticism: *Sextus Empiricus, Outlines of Scepticism*, ed. by J. Annas, J. Barnes, K. Ameriks, D. M. Clarke (2000); *The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism*, translated, introduction, and commentary by Benson Mates (1996); *Adversus Mathematicos*, with trans. by R. G. Bury, Loeb Classical Library: vol. 1–2, *Against the Logicians* and *Outlines of Pyrrhonism* (1933–1936); vol. 3, *Against the Physicists, Against the Ethicists* (1936); vol. 4, *Against the Professors* (1959–1960); and *Scepticism, Man, and God: Selections from the Major Writings of Sextus Empiricus*, ed. by P. Hallie, translated by S. G. Etheridge (1964); L. Floridi, *Sextus Empiricus: The Transmission and Recovery of Pyrrhonism* (2002).

Renaissance skepticism: Michel de Montaigne, "L'apologie de Raimond Sebond," in *Les essais de Michel de Montaigne*, edited by Pierre Villey, new ed. (1922); Fréderic Brahami, *Le scepticisme de Montaigne* (1997); Donald Frame, *Montaigne: A Biography* (1984); Elaine Limbrick and Douglas F. Thomson, *That Nothing Is Known: Francisco Sanches* (1988).

Skepticism and Fideism: Blaise Pascal, *Pensées*, edited by L. Brunschvicg (1951).

Skepticism in relation to modern philosophy: Pierre Bayle, *Dictionnaire historique et critique*, esp. the articles "Pyrrho" and "Zeno of Elea," both of which appear in Bayle's *Historical and Critical Dictionary: Selections*, translated and edited by Richard H. Popkin (1965); David Hume, *Dialogues concerning Natural Religion*, edited by Richard H. Popkin (1980); *Enquiries concerning the Human Understanding and concerning the Principles of Morals*, 2nd ed., edited by L. A. Selby-Bigge (1957), and *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton (2000); Sylvia Giocanti, *Penser l'irrésolution: Montaigne, Pascal, La Mothe le Vayer, Trois itinéraires sceptiques* (2001); John C. Laursen, *The Politics of Scepticism in the Ancients, Montaigne, Hume and Kant* (1992); Gianni Paganini, *The Return of Scepticism: From Hobbes and Descartes to Bayle* (2003); Antony Mckenna and Alain Mothu, *La philosophie clandestine a l'âge*

classique; José Maia Neto, *The Christianization of Pyrrhonism: Scepticism and Faith in Pascal, Kierkegaard, and Shestov* (1995); Gianluca Mori, *Bayle philosophe* (1999); Martin Mulsow, *Moderne aus dem Untergrund: radikale frühaufklärung in Deutschland 1680–1720* (2002); René Pintard, *Les libertinage érudits dans le première moitié du XVII siècle* (1943); Don Cameron Allen, *Doubt's Boundless Sea: Skepticism and Faith in the Renaissance* (1964); Richard H. Popkin, *The History of Scepticism from Savonarola to Bayle*, rev. ed. (2003) and *The High Road to Pyrrhonism* (1980); Popkin, Silvia Berti, and Françoise Charles-Daubert, *Heterodoxy, Spinozism, and Free Thought in Early-Eighteenth-Century Europe: Studies on the Traite Des Trois Imposteurs* (1996); Popkin and Arjo Vanderjagt, *Skepticism and Irreligion in the Seventeenth and Eighteenth Centuries* (1993); Popkin and José Maia Neto, *Skepticism in Renaissance and Post-Renaissance Thought: New Interpretations* (2004); Popkin, Ezequiel Olaso, and Giorgio Tonelli, *Skepticism in the Enlightenment* (1997).

The standard studies of ancient skepticism are: Victor Brochard, *Les Sceptiques grecs* (1887); Norman Maccoll, *The Greek Sceptics from Pyrrho to Sextus* (1869); Mary Mills Patrick, *The Greek Sceptics* (1929); Leon Robin, *Pyrrhon et le scepticisme grec* (1944); and Eduard Zeller, *The Stoics, Epicureans and Sceptics*, trans. by O. J. Reichel (1880); Charlotte L. Stough, *Greek Skepticism* (1969); Raoul Richter, *Der Skeptizismus in der Philosophie* (1904–1908); Cornelia De Vogel, *Greek Philosophy*.

For general studies of skepticism, see: Richard H. Popkin, "Skepticism," *Encyclopedia of Philosophy*, vol. 7, pp. 449–461 (1967); Arne Naess, *Scepticism* (1968); and Benson Mates, *Skeptical Essays* (1981).

Richard Popkin (1967, 2005)

SKINNER, B. F.

(1904–1990)

The name of B. F. (Burrhus Frederic) Skinner has become virtually synonymous with behaviorism. By introducing the concept of "operant conditioning" (in the late 1930s), Skinner fundamentally transformed behaviorist approaches to experimental psychology. Operant conditioning is based on the fact that the behavior of organisms (including people) typically has environmental consequences and is explained in important part by reference to them. Its fundamental principle is that the probability of occurrence of a specified kind of behavior is a function of the environmental consequences of previous occurrences of behavior of the same type, most notably, that the probability increases if the previous occurrences have been followed by "reinforcement." Skinner, surpassing older behaviorist "stimulus-response" approaches, inaugurated an experimental research program aiming to discover the laws of operant conditioning and, thus, generalizations concerning the three-term relation: discriminative stimulus-behavior-reinforcement.

The earliest laws of operant conditioning include generalizations about the relationship of the probability of a behavior's occurrence to its "schedule of reinforcement"—for example, to the conditions (discriminative stimulus) of its occurrence, the temporal duration between behavior and reinforcement, the proportion of behaviors that are followed by reinforcement, and whether these durations and proportions are fixed or variable. Later developments include generalizations about behavior that occurs under multiple schedules of reinforcement. The research program of operant conditioning constitutes Skinner's definitive and most lasting contribution. It also informs an applied program (of "behavioral technology"), based on the notion that behavior can be controlled by appropriate arrangement of the contingencies of reinforcement. The journals, *Journal of the Experimental Analysis of Behavior* (1958–) and *Journal of Applied Behavior Analysis* (1968–) are principally devoted, respectively, to publishing results of these and related programs.

Skinner considered his research program to underlie "radical behaviorism," a viewpoint that is distinct from the better-known (among philosophers) "logical behaviorism" and "methodological behaviorism." Unlike logical behaviorism, radical behaviorism does not hold that "mentalistic" terms—terms that may be taken to designate mental states or events (e.g., sensations, thoughts, memories, beliefs)—can be analyzed in terms of relations between behavior and the environment, or as referring to dispositions to behave in certain ways under specified environmental conditions. Unlike methodological behaviorism, it does not hold that any knowledge we may have about mental states and events is gained by means of inference (e.g., hypothetico-deductive) from knowledge of observed behavior, or that mental phenomena may be investigated by way of the behavioral phenomena causally linked with them. Radical behaviorism is not a philosophical thesis about meaning or about the epistemological primacy of behavior. It is a program aiming to "interpret" voluntary behavior (intentional action) in the light of the principle (in the most general terms) that voluntary behavior is under the control of environmental variables and the history of their relations with a person's behavior; or (more specifically) that it is explicable in terms of the history of contingencies of reinforcement to which a person has been exposed and the general laws (identified in the experimental program) of operant conditioning governing these contingencies. The philosophi-

cal journal *Behaviorism* (1972–1989) provided a forum for extensive discussion of radical behaviorism.

For Skinner, the philosophical impact of the program of radical behaviorism becomes apparent in the light of two proposals: (a) that adopting the program has the backing of scientific authority, and (b) that it is from science—rather than, say, from deploying ordinary intentional idiom—that we gain the best understanding of human phenomena. Regarding (a), he wrote a series of methodological articles (reprinted in Skinner 1969, 1972) arguing that the methodological and theoretical resources of the experimental program of operant (combined with respondent) conditioning at least match, and usually surpass, those of programs guided by methodological behaviorism. Thus, he concluded that theories that deploy mentalistic terms are unnecessary, and that a more complete account of behavior can be obtained within the framework of radical behaviorism. Regarding (b), in order to deal with the fact that language is integral to human behavior and that, in ordinary speech and communication acts, mentalistic terms are indispensable, he offered in *Verbal Behavior* (1957) a series of "interpretations" (speculative hypotheses) attempting to make it plausible that utterances containing these terms may be treated simply as instances of "verbal behavior," whose occurrences and other causal roles, can be explained (predicted and controlled) in terms of the principles of operant conditioning.

Radical behaviorism, applied to linguistic phenomena, had some influence on philosophical developments—for example, on the form of behaviorism adopted in W. V. Quine's *Word and Object* (1960), and on Quine's endorsement of "naturalistic epistemology." For the most part, however, philosophers are aware of *Verbal Behavior* mainly by way of Noam Chomsky's (1959) scathing review. Chomsky's most important criticism was that radical behaviorist "interpretations" are unable to encompass a number of fundamental aspects of linguistic phenomena: (e.g., the "creative" use of language, the rapidity and ease of the acquisition of language by children, and certain specific features of grammar, such as embedding of clauses). Furthermore, the linguistic phenomena cited by Chomsky became focal points of rival programs of experimental and theoretical psychology (psycholinguistics, cognitive psychology), which were designed to possess the theoretical resources needed to encompass them and to bypass Skinner's methodological objections. Chomsky, thus, rejected claim (a), that Skinner's program has the backing of scientific authority. Not so well known are behaviorist responses to Chomsky's arguments and further elaborations (and modifications) of Skinner's program (in, e.g., Place 1981), so much so that many philosophers consider Chomsky's review to have sounded the death knell of behaviorism.

Other critics questioned claim (b), that it is from experimentally based science that we get the best understanding of human phenomena. Barry Schwartz and Hugh Lacey (1982, 1987) argued against Skinner: (1) that his methodological criticism of the use of mentalistic terms in psychological theories does not apply at all to the use of intentional idiom in ordinary language; (2) that in fact human action cannot be reduced to behavior that is explicable in terms of laws (behaviorist or otherwise); and (3) that, using arguments that are formulated irreducibly in intentional idiom, the limits of applicability of radical behaviorist principles can be identified (Schwartz and Lacey 1982; Lacey and Schwartz 1987). These limits are ignored in *Verbal Behavior*, and also in Skinner's most controversial book *Beyond Freedom and Dignity* (1971). In the latter Skinner argued that fundamental notions of liberal democracy (freedom, dignity, autonomy) that are integral to standard defenses of civil rights are ill-founded and in conflict with the best scientifically grounded view of human nature. Such arguments suggested to his critics that the primary motivation for engaging in the program of radical behaviorism comes from commitment to the social value of the control of human behavior.

Although radical behaviorism ceased to have many high-profile adherents after the 1980s, and programs of cognitive psychology have become much more prominent than Skinner's experimental program in major universities, the residue of Skinner's contribution is deeply entrenched. The experimental program of operant conditioning continues at a high level of (increasingly mathematical) sophistication, exploring, for example, choices made under the influence of multiple contingencies of reinforcement in accord with the "matching law"; and Skinner's central theoretical term "reinforcement" has become a staple in practices that range from education to clinical psychology. In addition, newer behaviorist programs that are in continuity with Skinner's have emerged—for example, Howard Rachlin's (1994) "teleological behaviorism" and John Staddon's (1993) "theoretical behaviorism."

See also Behaviorism; Chomsky, Noam; Philosophy of Education, Epistemological Issues in; Psychology; Quine, Willard Van Orman.

Bibliography

Chomsky, Noam. "Review of B. F. Skinner's *Verbal Behavior*." *Language* 35 (1959): 26–58.

Lacey, Hugh, and Barry Schwartz. "The Explanatory Power of Radical Behaviorism." In *B. F. Skinner: Consensus and Controversy*, edited by S. Modgil and C. Modgil. New York: Falmer Press, 1987.

Place, U. T. "Skinner's *Verbal Behavior*." *Behaviorism* 9 (1981): 1–24, 131–152.

Quine, W. V. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Rachlin, Howard. *Behavior and Mind: The Roots of Modern Psychology*. New York: Oxford University Press, 1994.

Schwartz, Barry, and Hugh Lacey. *Behaviorism, Science, and Human Nature*. New York: Norton, 1982.

Staddon, J. E. R. *Behaviorism: Mind, Mechanism and Society*. London: Duckworth, 1993.

WORKS BY SKINNER

Verbal Behavior. New York: Appleton-Century-Crofts, 1957.

Contingencies of Reinforcement: A Theoretical Analysis. New York: Appleton-Century-Crofts, 1969.

Beyond Freedom and Dignity. New York: Knopf, 1971.

Cumulative Record: A Selection of Papers. 3rd. ed. New York: Appleton-Century-Crofts, 1972.

Hugh Lacey (2005)

SKOLEM-LÖWENHEIM THEOREM

See *Logic, History of: Modern Logic: From Frege to Gödel*

SKOVORODA, HRYHORII SAVYCH (GRIGORII SAVVICH)

(1722–1794)

Grigorii (Hryhorii) Savvich Skovoroda, the Ukrainian poet, fabulist, philosopher, and religious thinker, was educated at the Kiev Theological Academy. As a young man he traveled in eastern and western Europe and paid brief visits to St. Petersburg and Moscow, but eighteenth-century European culture left few traces on his thought. He taught, mainly literature, at Pereiaslavl' (Pereiaslavl'-Khmel'nitskii) about 1755 and at the Khar'kov (Khar'kiv) Collegium from about 1759 to 1765, but he fell out with his ecclesiastical superiors and was dismissed. He spent his last thirty years as a mendicant scholar and "teacher of the people."

Skovoroda's disciple, M. I. Kovalinski, has left an engaging account of Skovoroda's manner of life:

> He dressed decently but simply; … he did not eat meat or fish, not from superstitious belief but because of his own inner constitution; … he allowed himself no more than four [hours a day] for sleep; … he was always gay, good-natured, easy-going, quick, restrained, abstemious, and content with all things, benign, humble before all men, willing to speak so long as he was not required to …; he visited the sick, consoled the grieving, shared his last crust with the needy, chose and loved his friends for the qualities of their hearts, was pious without superstition, learned without ostentation, complaisant without flattery. ("The Life of Gregory Skovoroda," translated by G. L. Kline, in *Russian Philosophy,* Vol. I, p. 20)

Skovoroda aspired to be a "Socrates in Russia" both as a moralist, a gadfly provoking thoughtless and selfish men to scrutinize their lives, and as an intellectual forerunner, clearing the path for the more profound and systematic philosophizing of a future "Russian Plato." In many ways he was not only the last, but also the first, of the medievals in Russia. His metaphysics and philosophical anthropology are explicitly Christian and Neoplatonic, and his philosophical idiom is studded with Greek and Church Slavonic terms and constructions. He knew both German and Latin (he left over a hundred Latin letters and poems) and had some knowledge of Greek and Hebrew, but he wrote all of his philosophical works in Russian. As it happened, few of his own philosophic coinages were accepted by later Russian thinkers.

All of Skovoroda's philosophical and theological writings are in dialogue form. They are Socratic in method and in theme, genuinely dramatic and dialogic, written with wit, imagination, and moral intensity. They offer an acute critique of both ontological materialism and sense-datum empiricism, and they outline a dualistic cosmology with a pantheistic (or "panentheistic") and mystical coloring. One of Skovoroda's favorite metaphors for the relation of appearance to reality is that of a tree's many passive, shifting shadows to the firm, single, living tree itself.

In deliberate opposition to the Baconian summons to "know nature in order to master it," Skovoroda urged individuals to "know themselves in order to master themselves" and to put aside desires for comfort, security, fame, and knowledge. His position is thus Stoic as well as Socratic. Seneca, no less than Socrates, would have

savored the epitaph which Skovoroda wrote for himself: "The world set a trap for me, but it did not catch me."

See also Appearance and Reality; Neoplatonism; Pantheism; Plato; Russian Philosophy; Socrates; Stoicism.

Bibliography

WORKS BY SKOVORODA

Hryhori Skovoroda: Tvori v Dvokh Tomakh (Grigorii Skovoroda: works in two volumes), edited by O. I. Biletski, D. K. Ostryanin, and P. M. Popov. Kiev, 1961. Text in Russian and Latin; introduction, commentary, notes, and translation of Latin text in Ukrainian.

"A Conversation among Five Travellers concerning Life's True Happiness" (abridged translation by George L. Kline of "Razgovor pyati putnikov o istinnom shchastii v zhizni," *Tvori v Dvokh Tomakh,* Vol. I, 207–247). In *Russian Philosophy,* edited by James M. Edie, James P. Scanlan, Mary-Barbara Zeldin, and George L. Kline. 3 vols. Vol. I, 26–57. Chicago: Quadrangle, 1965.

WORKS ON SKOVORODA

Chyzhevsky, D. *Filosofiia H. S. Skovorody* (The philosophy of G. S. Skovoroda). Warsaw, 1934. In Ukrainian.

Ern, V. *Grigorii Savvich Skovoroda: Zhizn' i uchenie* (Grigorii Savvich Skovoroda: his life and teaching). Moscow, 1912.

Zenkovsky, V. V. *Istoriia russkoi filosofii,* 2 vols. Paris: YMCA Press, 1948–1950. Translated by George L. Kline as *A History of Russian Philosophy,* 2 vols., 53–69. New York: Columbia University Press, 1953.

George L. Kline (1967)

SKOVORODA, HRYHORII SAVYCH (GRIGORII SAVVICH) [ADDENDUM]

Hryhorii Savych Skovoroda's outdated language and literary style make it difficult to grasp his philosophical thought. He expresses his ideas mostly through images, symbols, proverbs, and stories instead of philosophical concepts and propositions, and he does not organize them logically into a system. Because of this he has been subject to many conflicting interpretations: He has been called both an eclectic and a strict rationalist, a mystic and a materialist, a theologian and a moral teacher. While some of his doctrines are obscure (the heart in man, personal immortality, the nature of matter), his ideas do fall into a logically coherent system that is intended to serve a practical purpose.

The aim of philosophy, according to Skovoroda, is to show people the way to happiness. This is why his moral teachings are articulated more fully than the other parts of his philosophy. His metaphysical, epistemological, and anthropological teachings are developed only to the extent that is necessary for grounding his moral principles. For Skovoroda happiness is not merely the absence of pain or a state of inner peace, but joy and gaiety, which are not free of tension. To attain happiness two things are necessary: to be content with everything and to fulfill one's true self. The first rests on a belief in a providential order that supplies each creature with whatever is necessary for its happiness. The Epicurean doctrine that what is necessary is easy and what is difficult is unnecessary liberates us from fear and anxiety. The other condition for happiness is the pursuit of one's God-given, innate, congenial task (*srodnyi trud*) in life. To work at one's natural task brings joy, while to work at an unnatural task brings misery regardless of the accompanying external rewards such as wealth and fame. Every congenial task corresponds to a necessary social role (e.g., ruler, teacher, soldier, farmer, and so on); hence, by fulfilling their natural potential people also ensure the harmonious and efficient functioning of society.

This moral teaching rests on a dualistic metaphysics. Skovoroda divides reality into three isomorphic worlds: the macrocosm or the all-encompassing universe, the human microcosm or man, and the symbolical microcosm or Bible. All three worlds have an inner and outer, spiritual and material, intelligible and sensible nature: in the macrocosm the two natures are called God and the physical universe; in man soul and body; and in the Bible the true and the apparent meaning. The inner principle in each world is the more important one: It sustains and rules the outer one and is eternal and immutable. Self-knowledge is the foundation of all knowledge: By delving into oneself one discovers the essential truths not only about one's own nature and one's congenial task, but also about the other two worlds. Skovoroda considered the Bible, his favorite book, to be a treasury of universal wisdom and a source of false beliefs for those who take its statements at their face value. His dialogues are largely discussions of its symbolic meaning.

Skovoroda's poetry, composed in a language close to the Ukrainian vernacular, became popular among the common people, while his dialogues circulated in manuscript within the narrow circle of his friends. The first collection of his works to appear in print (1861) contained only half of his dialogues. Fuller collections came out in 1894, 1912, and 1961 and the first complete collection did not appear until 1973. Skovoroda's ideas began to attract the attention of philosophers only at the end of the nineteenth century. Although Skovoroda's influence in Russ-

ian and Ukrainian philosophy has been negligible, his colorful and independent personality has served as an inspiration to Ukrainian writers during the cultural revival of the 1920s and 1960s.

See also Happiness; Macrocosm and Microcosm; Russian Philosophy; Self-Knowledge.

Bibliography

WORKS BY SKOVORODA

Hryhorii Skovoroda: Povne zibrannia tvoriv u dvokh tomakh (Hryhorii Skovoroda: The complete collection of his work in two volumes). Edited by V. Shynkaruk et al. Kyiv, Ukraine: Naukova dumka, 1973. Text in the literary Russian of his time and Latin. Introduction, commentary, notes, and translation of Latin texts in Ukrainian.

Fables and Aphorisms. Translated by Dan B. Chopyk. New York: Peter Lang, 1990.

WORKS ABOUT SKOVORODA

Special Issue on Hryhorii Skovoroda. *Journal of Ukrainian Studies* 22 (1–2) (1997).

Dva stolittia Skovorodiiany / Two Centuries of Skovorodiana: Bibliographical Guide. Project director, Leonid Ushkalov. Kharkiv, Ukraine: Acta, 2002.

Hryhorij Savyč Skovoroda: An Anthology of Critical Articles, edited by Richard M. Marshall and Thomas E. Bird. Edmonton: Canadian Institute of Ukrainian Studies Press. 1994.

Taras Zakydalsky (2005)

SMART, JOHN JAMIESON CARSWELL

(1920–)

John Jamieson Carswell Smart was born into an academic Scottish family on September 16, 1920. His father, W. M. Smart, was an astronomer in Cambridge until 1937 when the family moved to Glasgow. J. J. C. Smart entered the University of Glasgow in 1938.

War service interrupted Smart's education from 1940 to 1945, after which he rapidly completed his degrees at Glasgow, then proceeded to the University of Oxford, where he read for the newly established BPhil degree and came under the influence of Gilbert Ryle. After a short period at Corpus Christi College, he accepted, at the age of twenty-nine, the Hughes Professorship of Philosophy at the University of Adelaide.

Smart spent twenty-two years at the University of Adelaide, moving to La Trobe University in Melbourne in 1972. In 1976 he was appointed to a Chair in the Research School of Social Sciences of the Australian National University, which he held until his retirement in 1985. Since then he has continued to be active in philosophy at the Australian National University and in Melbourne.

Soon after his arrival in Australia Smart's thought moved away from its linguistic, Oxford orientation and began to take on its characteristic science-based form. Showing the influence of both eighteenth-century Scot David Hume and twentieth-century American W. V. Quine, Smart's mature philosophy has been consistently empiricist, taking human experience as the wellspring and touchstone of knowledge, giving primacy to statements of actual fact and treating modal claims regarding necessity or mere possibility as human artifacts, and embracing nominalism concerning universals. In the philosophy of science, he has upheld regularity views of causation and natural law. Unlike many empiricists, however—who regard imperceptible entities as human constructs—Smart has always been staunchly realist in his account of some theoretical entities, claiming that electrons, for example, are straightforwardly real components of the world.

Smart's ethics has been similarly consistent: He has defended a rather pure act-utilitarian consequentialism throughout. His major contributions to philosophy have involved three themes: in cosmology, four-dimensional physical realism; in the philosophy of mind, materialism; and in ethics, utilitarianism.

For forty years, culminating with *Our Place in the Universe* (1989), Smart has argued that the four-dimensional conception of space-time introduced by Minkowski for the interpretation of the theory of special relativity is superior to all others. This conception implies the equal reality of past, present, and future and rejects as unreal the flow of time that seems to underpin the human experience of time passing.

Smart's second major theme is materialism, the claim that there are no spiritual realities, and that in particular human minds are not spiritual. The mind—the organ with which one thinks—proves to be the brain. All the various states of mind are states, processes, or functions of the brain and its associated nervous system. This *central state materialism* emerged in its contemporary form from two landmark papers: Smart's colleague U. T. Place published his "Is Consciousness a Brain Process?" in the *British Journal of Psychology* in 1956; Smart's "Sensations and Brain Processes," which appeared in *The Philosophical Review* in 1959 (reprinted in *Essays, Metaphysical and Moral* [1987]), gave the view wide notoriety. The importance of Smart's paper consisted in his exposing the

inadequacy of the reasons then prevalent for holding that the mental and the physical belong to essentially incompatible categories. Smart expanded and defended materialism in subsequent discussions both of the general issue and of its implications for the secondary qualities, particularly color.

From *An Outline of a System of Utilitarian Ethics* (1961) onward, Smart has presented a utilitarian theory of moral judgment and action: What matters is not people's intentions, or character, nor any fixed set of moral rules, but the actual consequences of behavior. The consequences to be considered concern the happiness of all sentient beings, as judged from a natural, secular point of view. To adhere to a social or traditional rule of conduct, even in those cases where doing so would result in increased misery, Smart deprecates as "rule worship." He recognizes the notorious difficulties that questions of justice generate for any rigorously utilitarian theory; in *Ethics, Persuasion and Truth* (1984) discussing the enormity of accepting the idyllic happiness of many at the cost of the continuing torture of one lost soul. There is no definitive resolution in his ethical thought of this conflict between the claims of happiness and of justice.

Philosophy and Scientific Realism (1963) marked the first appearance of a line of thinking that continues through *Our Place in the Universe* (1989) and subsequent pieces: what is now known as the Argument to the Best Explanation. The issue is realism over theoretical entities such as electrons and quarks, which must forever be beyond any direct observational validation. Smart's position is that the complex, interlocking set of experimental results that have been obtained and validated about electrons, for instance, would constitute an incredible set of interlocking coincidences for which there could be no intelligible accounting, unless electron theory were (close to being) literally referentially correct.

In *Ethics, Persuasion, and Truth* (1984) Smart argues for a sophisticated subjectivist theory in metaethics. As an empiricist, Smart rejects the idea that moral judgments state some special kind of "moral fact," and develops a preference semantics and pragmatics for them. *Our Place in the Universe* (1989) presents a coherent naturalistic vision of the physical world and life on earth, suffused with a kind of natural piety or philosophic awe.

Since 1990, Smart has continued to write on all the major themes of his philosophy. In 1996 he joined with J. J. Haldane in a debate on the issue of atheism. In all his work, Smart argues for firmly held views with the calm, well-informed courtesy and candor that have made him one of the best loved, as well as most respected, of contemporary philosophers.

See also Colors; Consequentialism; Empiricism; Inference to the Best Explanation; Philosophy of Mind; Utilitarianism.

Bibliography

BOOKS BY SMART

An Outline of a System of Utilitarian Ethics. Melbourne: Melbourne University Press on behalf of the University of Adelaide, 1961.

Philosophy and Scientific Realism. London: Routledge & Kegan Paul, 1963.

Between Science and Philosophy: An Introduction to the Philosophy of Science. New York: Random House, 1968.

Utilitarianism: For and Against, with Bernard Williams. Cambridge, U.K.: Cambridge University Press, 1973.

Ethics and Science. Hobart, Australia: University of Tasmania, 1981.

Ethics, Persuasion, and Truth. London: Routledge & Kegan Paul, 1984.

Essays, Metaphysical and Moral: Selected Philosophical Papers. Oxford: B. Blackwell, 1987.

Our Place in the Universe: A Metaphysical Discussion. Oxford: Blackwell, 1989.

Atheism and Theism. With J. J. Haldane. Oxford: Blackwell, 1996.

BOOKS EDITED BY SMART

Problems of Space and Time: Readings. New York: Macmillan, 1964.

BOOKS ABOUT SMART

Petit, Philip, Richard Sylvan, and Jean Norman, eds. *Metaphysics and Morality: Essays in Honour of J. J. C. Smart*. Oxford: B. Blackwell, 1987.

Keith Campbell (1996, 2005)

SMITH, ADAM

(1723–1790)

Adam Smith, one of the most influential political economists of Western society, first became known as a moral philosopher. Smith was born in Kirkcaldy, Scotland. His father died shortly before he was born, and his mother's loss doubtless explains the lifelong attachment that flourished between her and her son. Smith entered the University of Glasgow in 1737, where he attended Francis Hutcheson's lectures. In 1740 he entered Balliol College, Oxford, as a Snell exhibitioner. He remained at Oxford for seven years and then returned to Kirkcaldy. In 1748 he moved to Edinburgh, where he became the friend of David Hume and Lord Kames (Henry Home). In 1751 he was elected professor of logic at the University of Glas-

gow, and in the next year he exchanged logic for the professorship in moral philosophy, an appointment that he held for the next ten years.

The *Theory of Moral Sentiments,* drawn from his course of lectures, was published in 1759. The work received wide acclaim and so impressed the stepfather of the young duke of Buccleuch that he invited Smith to become the duke's tutor, with the promise of a pension for life. Smith resigned his professorship at Glasgow and accompanied the duke on a visit to the Continent that lasted from 1764 to 1766. His tutoring duties ended, he returned again to Kirkcaldy, where he spent the next ten years in retirement at work on *The Wealth of Nations,* which was published in 1776 and for which he became famous. In 1778 he was appointed a commissioner of customs for Scotland. He died in 1790 and was buried in the Canongate churchyard, Edinburgh.

The greater part of the *Theory of Moral Sentiments* is an account of moral psychology. Only after he has settled the psychological questions does Smith turn, in the last seventh of the work, to moral philosophy. The mainstay of Smith's moral psychology is sympathy. Sympathy is our fellow feeling with the passions or affections of another person. Smith characterizes the mechanism of sympathy in this way: "Whatever is the passion which arises from any object in the person principally concerned, *an analogous emotion* springs up at the thought of his situation, in the breast of every attentive spectator." The important phrase here is "at the thought of his situation." Sympathetic feelings may seem to arise from our seeing the expression of a certain emotion in another person, but Smith argues that if the appearance of grief or joy, for example, arouses similar feelings in us, it is because these feelings suggest to us the general idea of some good or evil that has befallen the person in whom we observe them. What is more, there are some passions whose expression excites disgust rather than sympathy until we are acquainted with their cause. The furious behavior of an angry man, for example, is more likely to exasperate us against him than against his enemies. Thus, Smith concludes that sympathy does not arise so much from the view of the passion as from the view of the situation that excites it, and he reinforces this claim by noting that we sometimes feel for another a passion that he himself seems to be altogether incapable of, as when we feel embarrassed at someone's behaving rudely although he has no sense of the impropriety of his behavior.

Sympathy is the basis for our judgments of both the propriety and the merit of other people's feelings and the actions that follow from them. When the original passions of the principal person are in perfect accord with the sympathetic emotions of the spectator, the passions of the principal appear to the spectator as just and proper. Smith even goes as far as to say that to approve of the passions of another as suitable to their objects is the same as to observe that we entirely sympathize with them. Indeed, even though our own emotions may make it impossible for us to have on occasion a certain sympathetic emotion, we may "by general rules" recognize the appropriateness of some person's having a given emotion because, for example, we could sympathize with the other person's joy but for our own grief.

Although our sense of the propriety of some piece of conduct arises from our sympathy with the affections and motives of the agent, our sense of merit (that is, our sense of a certain action's making the agent worthy of a reward) stems from our sympathy with the gratitude of the person affected by the action. When we see someone aided by another, our sympathy with his joy at the receipt of the aid animates our fellow feeling with his gratitude toward his benefactor.

Having shown how sympathy gives rise to the senses of propriety and of merit in our judgment of the passions and conduct of others, Smith turns to showing how these sentimental mechanisms may be employed in our judgment of ourselves. We must take care to avoid a self-interested partiality in our judgments. According to Smith, impartiality can be achieved only if we look at our own behavior as though it were someone else's. Thus, we may judge ourselves from the same point of view that we judge others, and our approval or disapproval of our own conduct will depend on whether we can sympathize with the sentiments from which our actions flow. Conscience, "the judge within us," enables us to make a proper comparison between our own interests and the interests of others. With its aid we may approach the ideal of the man of perfect virtue, who is possessed of both a command of his own feelings and a sensibility for the feelings of others.

We may guard against self-deceit by keeping before us the general rules for what is appropriate in human conduct. These rules have their basis in the sentiments that certain kinds of behavior evoke, and our own respect for the rules should follow from the correspondence between them and our own feelings as we observe the conduct of others. Smith stresses that the rules are generalizations from particular instances in which conduct has excited the sense of propriety and merit in humankind. A just regard for these general rules is a sense of duty. By acting from a sense of duty, one can make up for any lack of the appropriate sentiment on a given occasion. Of all

the general rules, those that define justice have the greatest exactness.

Throughout his discussion of our moral psychology, Smith assumes the general acceptance of beneficence and justice as social virtues. He glides quickly over the problem of their description, and he introduces sympathy into his moral psychology as a kind of absolute without considering whether someone might sympathize with "wrong" affections.

In his moral philosophy Smith treats of two questions: Wherein does virtue consist? What power or faculty of the mind recommends virtue to us?

The different accounts of virtue may be reduced to three principles. First, virtue is the proper government and direction of all our affections (propriety). Second, virtue is the judicious pursuit of our own private interest (prudence). Third, virtue lies in the exercise of only those affections that aim at the happiness of others (benevolence). These principles make it evident either that virtue may be ascribed to all our affections when properly governed (as the principle of propriety implies) or that virtue is limited to one of two classes of our affections, either the prudent ones or the benevolent ones.

After surveying the various systems of morals, Smith offers the following conclusions. The systems based on propriety give no precise measure of it. Smith remedies this defect by pointing out that the standard of what is appropriate in sentiments and motives can be found nowhere but in the sympathetic feelings of the impartial spectator. The most that can be claimed for the definition of virtue as propriety is that there is no virtue without propriety, and where there is propriety, some approbation is due. But those who make propriety the sole criterion of virtue can be refuted by the single consideration that they cannot account for the superior esteem granted to benevolent actions. However, neither prudence nor benevolence can be allowed to be the sole criterion of virtue, for whichever we choose, we make it impossible to explain our approbation of the other. Smith's implied conclusion is that there can be no single criterion of virtue and that each of the three principles that he notes must be allowed its just scope.

When Smith turns to the question of what power or faculty of the mind recommends virtue to us, he remarks that this question is of purely speculative interest and has no practical importance whatsoever. Several candidates had been proposed by Smith's predecessors as the source of virtue, notably self-love, reason, or some sentiment. Smith rejects self-love as the ultimate basis of behavior, and hence as the basis of virtue, on the ground that its proponents have neglected sympathy as a cause of action. For Smith, sympathy is not a selfish principle. Smith also rejects reason as a source of the distinction between virtue and vice because reason cannot render any action either agreeable or disagreeable to the mind for its own sake. The first perceptions of right and wrong must be derived from an immediate sense of the agreeableness or disagreeableness of actions. Thus, Smith is left with the conclusion that there must be some sentiment that recommends virtue to us.

Smith considers the proposal that there is a special sense of virtue, the moral sense, as proposed by his former teacher Hutcheson. But Smith regards the moral sense as objectionable on two counts. First, no one seemed to be aware that he had a moral sense before the moral philosophers began to talk about it; and if the moral sense is a genuine sense, this state of affairs seems very odd indeed. Second, Smith finds that sympathy, a recognized human phenomenon, is the source of a range of feelings that provide a foundation for virtue. Therefore, since a sentimental basis for virtue is already provided by nature, there is no need to invent one in the form of a moral sense.

An Inquiry into the Nature and Causes of the Wealth of Nations is partly a description of the actual conditions of manufacture and trade in Smith's own time, partly a history of European economics, and partly recommendations to governments. Smith opposes the mercantilist beliefs that money is wealth and that the best economic policy for a country is the retention within its borders of as much gold and silver as possible. He argues, rather, that wealth is consumable goods and that the wealthiest country is one that either produces itself or can command from others the greatest quantity of consumable goods.

The development of a full-blown economic system requires some people in a society to possess a supply of either raw materials or manufactured goods greater than is required to fulfill their own immediate needs. The surplus stocks provide the opportunity for trade among people with various needs. Where the demand for a certain kind of thing is great enough to assure a producer that his other wants may be supplied in exchange for producing this certain good, he will specialize in its production. This kind of division of labor will continue, according to Smith, until some laborers are producing a very small part of a manufactured product because the master finds that a division of labor enables his workers to produce a greater quantity of goods in a shorter time.

Smith believes that the general welfare will be best served by permitting each person to pursue his own interest. Sympathy, which figured largely in Smith's account of moral psychology, is not mentioned in his economics. Self-interest is the motive required to explain economic action. Smith argues, "Every individual is continually exerting himself to find out the most advantageous employment for whatever capital he can command." Since the most advantageous employment of capital is to be found in producing and selling the goods that satisfy the greatest needs of a people, the capitalist is bound to work to satisfy those needs. Intending only his own gain, he contributes nonetheless to the general welfare. Thus, the capitalist is "led by an invisible hand to promote an end which was no part of his intention."

Smith was instrumental in bringing his contemporaries to see the modern European economic system for the first time, and we are the heirs of their vision. Of course, Smith is guilty of oversimplifications and omissions, but his work is nonetheless a model of both observation and systematization in the social sciences.

See also Ethics, History of; Ethics and Economics; Home, Henry; Hume, David; Hutcheson, Francis; Moral Sense; Philosophy of Economics; Virtue Ethics.

Bibliography

The edition of *The Wealth of Nations* prepared by Edwin Cannan (London: Methuen, 1904) is recommended. See also William R. Scott, ed., *Adam Smith as Student and Professor, with Unpublished Documents* (Glasgow: Jackson, 1937), a biographical account, and Eli Ginzberg, *The House of Adam Smith* (New York: Columbia University Press, 1934), which is the classic analysis of Smith's economics based on consideration of his predecessors and contemporaries. For Smith's moral philosophy, see James Bonar, *Moral Sense* (London: Allen and Unwin, 1930), Chs. X and XI.

Elmer Sprague (1967)

SMITH, ADAM [ADDENDUM]

Adam Smith's claim on the history of aesthetics lies in his essay, "Of the Nature of that Imitation which takes place in what are called the Imitative Arts," arguably the most logically acute and penetrating discussion of what one would call pictorial representation that eighteenth-century Britain produced. It was first published, posthumously, in 1795, in Smith's *Essays on Philosophical Subjects*.

The main thesis of Smith's account is that "the disparity between the imitating and the imitated object is the foundation of the beauty of imitation. It is because the one object does not naturally resemble the other, that we are so much pleased by it, when by art it is made to do so" (1795, p. 144).

Smith's most elaborately worked-out example concerns the contrast between painting and sculpture, much discussed in the eighteenth century. The idea is that statues represent three-dimensional objects in a three-dimensional medium, whereas paintings represent three-dimensional objects in two dimensions. Hence a higher level of resemblance would be required of a statue to its represented object than would be required of a painting to its, to achieve the same level of representational beauty. "The disparity between the object imitating, and the object imitated," Smith wrote, "is much greater in the one art than in the other; and the pleasure arising from the imitation seems to be greater in proportion as this disparity is greater" (1795, p. 137). Smith pays considerable attention in his essay, as well, to music and dance, concluding that "the imitative powers of Dancing are much superior to those of instrumental Music, and are at least equal, perhaps superior, to those of any other art" (Smith, p. 175).

See also Aesthetics, History of; Art, Expression in; Art, Representation in.

Bibliography

Smith, Adam. *Essays on Philosophical Subjects*. London, 1795.

Peter Kivy (2005)

SMITH, JOHN
(c. 1616–1652)

John Smith, the moral and religious philosopher of the Cambridge Platonist school, was born at Achurch, near Oundle, in Northamptonshire. Very little is known with certainty about his origins. It would seem that his father was a locally respected small farmer, that both of his parents were elderly when he was born, that he lost his mother in his early childhood and his father soon after. His short life was a continual struggle against poverty and ill health. In 1636 he was somehow enabled to enter Emmanuel College, where he came under the influence of Benjamin Whichcote. Although he was about the same age as his fellow Platonist Ralph Cudworth, Cudworth was already a fellow of Emmanuel before Smith took his

BA in 1640; Smith was very likely his pupil and certainly came under his influence. The influence may have been in some measure reciprocal.

Smith took his MA degree in 1644; the same year he was elected a fellow of Queen's College, Cambridge, having been declared by the London Assembly of Divines a suitable person to replace one of the fellows who had been ejected by the Puritan Parliament. He taught Greek, Hebrew, and mathematics. Like his master Whichcote he had gifts of character and personal warmth, which won for him not only the respect but also the affection of pupils and friends. The funeral sermon preached by Simon Patrick on his death on August 7, 1652, is, even allowing for the extravagance of phrase common to such occasions, an impressive tribute to his intellectual and personal gifts. He published nothing, but after his death a series of *Discourses* that he had delivered as dean of his college in the chapel of Queen's was collected, edited, and published by John Worthington. Another volume was promised but never appeared.

Matthew Arnold described Smith's *Discourses* as "the most admirable work left to us by the Cambridge School." This is the judgment of a man whose interests lay in religion and culture rather than in philosophy. As a philosopher Smith will not stand comparison with Cudworth or Henry More. Basically, he was an eloquent apologist for the liberal theology of the Cambridge school. The flow of that eloquence, however, is interrupted, in the Cambridge Platonist manner, by quotations in a variety of tongues from Plato, especially the *Phaedo* and the *Republic*, and the Neoplatonists, the Hebrew Scriptures, the Talmud, and, the sole contemporary, René Descartes. Smith's reasoning is by no means close. "It is but a thin, airy knowledge," he writes in the first *Discourse*, "that is got by mere speculation, which is ushered in by syllogisms and demonstrations." God's nature, he thinks, is to be understood by "spiritual sensation" rather than by verbal description; Smith's object is to arouse such a "spiritual sensation" in human souls, and philosophy is only ancillary to that task.

THOUGHT

The first six of the *Discourses* Smith composed as a continuous essay. They were to be the first segment of a book that he did not live to complete. As editor, Worthington broke up the essay into chapters and added, from Smith's papers, four sermons to act as a substitute for the unwritten segments of the essay. Smith's general thesis is the Platonic one that goodness and knowledge are intimately united; only the purified soul can achieve true knowledge. Every soul, he thought, has within it innate concepts of religion and morality. Ordinarily obscured by sensuality, they nevertheless act as a guide to the direction in which purification is to be sought. Such principles Smith thinks of as innate ideas. Knowledge, in his view, is derived by reflection of the character of our souls; it does not arise out of sensory experience. One can see why he admired the Neoplatonists and welcomed the teachings of Descartes. He did not live long enough to share in the revulsion against Descartes's teachings as mechanistic, which More and Cudworth were to exhibit; indeed, in his *Discourses* he draws on Descartes's physiology.

According to Smith, the three great enemies of religion are superstition, legalism, and atheism. Superstition consists of treating God as a capricious power who has to be cajoled by flattery, bribery, or magical spells. Legalism conceives of religion as laying down doctrines that have simply to be accepted as rules for governing our conduct. It can take a variety of forms, "Scripture-Christianity" is quite as legalistic as Jewish formalism if it consists of picking out of the Scriptures a set of doctrines on the acceptance of which salvation is supposed to depend. Smith attacks this sort of Christianity with particular vigor, especially in his Sermon "Pharisaical Righteousness" (*Discourses* VIII).

As for atheism, Smith, unlike Cudworth and More, did not have Thomas Hobbes to contend with. He knew of atheism only as it appears in the writings of the Epicureans; much of his (very brief) argument against atheism is directed against the Epicurean version of atomism. He regards the belief in God as a "natural belief" that scarcely needs to be defended. He is much more preoccupied with the belief in immortality, perhaps because Richard Overton in a notorious pamphlet, *Man's Mortality* (1643), published in London although as if from Amsterdam, had denied that humans are by nature immortal, arguing that the soul and the body are so compounded that they die and are resurrected together. Smith defends what Overton had rejected—the traditional distinction between soul and body—calling upon Descartes for support.

If people are led to doubt the immortality of their souls, Smith argues, this is only because they are conscious that their souls do not deserve to be immortalized. Once they improve the quality of their lives, they will come to be conscious of their souls as exhibiting a kind of goodness that is obviously destined to be eternal. Similarly, if questions arise about God's nature, these can be settled, as Plotinus had suggested, only by reflection on the workings of our own souls in their most godlike

moments. God is the perfect soul, the perfectly loving soul, the perfectly rational soul; that this is God's nature we see by reflection upon our own perfections and imperfections.

It is easy to see why men as different as John Wesley and Matthew Arnold expressed admiration for Smith and sought to introduce his writings to a wider audience. Smith's appeal to inwardness, to the capture of the soul by God, recommends him to the evangelical; his rejection of merely creedal religions, the moral emphasis of his teaching, recommends him to the liberal theologian.

See also Cambridge Platonists.

Bibliography

The *Select Discourses* were first published and edited by John Worthington in 1660 (London). The best edition is that edited by Henry Griffin Williams for the Cambridge University Press in 1859. Various extracts have been published, for example, by John Wesley in Vol. XI of his Christian Library (London, 1819–1827). A selection with the title *The Natural Truth of Christianity*, edited by W. M. Metcalfe (Paisley, U.K., 1882), includes an introductory commendation by Matthew Arnold.

For works on Smith, see bibliography to the "Cambridge Platonists" entry; Rufus Matthew Jones, *Spiritual Reformers in the Sixteenth and Seventeenth Centuries* (London: Macmillan, 1914), Ch. 16; the Address to the Reader prefixed to John Worthington's edition of the *Discourses* (reprinted in the 1859 edition); John K. Ryan, "John Smith, Platonist and Mystic," in *New Scholasticism* 20 (1) (1946): 1–25; J. E. Saveson, "Descartes' Influence on John Smith, Cambridge Platonist," in *Journal of the History of Ideas* 20 (2) (1959): 258–263; C. A. Patrides, *The Cambridge Platonists* (Cambridge, U.K.: Cambridge University Press, 1980); and Brenton J. Stearns, "Mediate Immediacy: A Search for Some Models," in *International Journal for Philosophy of Religion* 3 (1972): 195–211.

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

SMUTS, JAN CHRISTIAAN
(1870–1950)

Jan Christiaan Smuts, the South African statesman, soldier, and scholar, introduced the concept of "holism" into philosophy. Smuts was born on a farm near Riebeek West, Cape Colony (now Western Cape Province). He was graduated from Victoria College, Stellenbosch, in 1891 and from Cambridge in 1894, where he studied law. At both places his record was brilliant, but he had the reputation of being a bookish recluse who made few friends. Returning home in 1895, he was admitted to the bar, entered political life, and during the Boer War commanded a force against the British with the rank of general. However, when World War I broke out in 1914 he became a staunch defender of the Allied cause. In 1918 he published a pamphlet titled *The League of Nations: A Practical Suggestion,* which helped to form President Woodrow Wilson's ideas. From 1919 to 1924, and again from 1939 to 1948, he was prime minister of South Africa. In the intervening period he completed his only philosophical work, *Holism and Evolution* (New York, 1926). Smuts was a dominant figure in the politics of his country for over half a century and an influential figure on the world scene. His enemies considered him arrogant and ruthless, more interested in ideas than in people. Yet the theme of his politics, as of his philosophy, was the integration of parts into wholes.

This theme is central to *Holism and Evolution,* where it is used to integrate the results of the sciences, especially the biological sciences, and where it becomes the basis of "a new *Weltanschauung* within the general framework of Science." The background was supplied by the theory of evolution, so interpreted as to preclude mechanistic or materialistic formulations of it. Such formulations, Smuts held, are incompatible with the fact that evolution is creative, having successively brought into existence items that are genuinely novel and that were not even potentially existent before they appeared on the scene. These items he called "wholes." Their appearance was explained by postulating a primordial whole-making, or "holistic," factor in the universe. This factor he also called a "creative tendency or principle" operative throughout the history of nature.

Smuts apparently wished to distinguish wholes in the strict sense from mere aggregates, mechanical systems, and chemical compounds. In a true whole the parts lose forever their prior identity. In aggregates, mechanical systems, and chemical compounds, however, the identity of the parts or elements is not lost but is always recoverable. There are certain entities, such as biochemical systems, which appear to have an intermediate status. For they display "a mixture of mechanism and holism." These systems form "the vast ladder of life." At the bottom of the ladder, mechanistic features predominate; at the top, holistic features predominate. True wholes, free of any admixture of mechanism, are exemplified in minds or psychic structures, which first appear among higher organisms, and in human personality, "the supreme embodiment of Holism."

Smuts sometimes spoke of atoms and molecules as wholes, presumably using the term in other than the

strict sense he had defined. The broader use allowed him to affirm that the factor of holism is "responsible for the total course of evolution, inorganic as well as organic. All the great main types of existence are due to it." Long before organisms or minds arose, the holistic factor was producing elementary wholes of a purely physical kind. Later, through a series of "creative leaps," it became more fully embodied in biological structures, minds, and persons. Indeed, "it is in the sphere of spiritual values that Holism finds its clearest embodiment," for in this sphere love, beauty, goodness, and truth have their source.

Smuts nowhere attributed to the holistic factor any teleological orientation. Nor did he apply to it any personal or spiritual categories. It was represented as an ultimate principle, metaphysical rather than religious, at work and still working in the cosmos.

There is a considerable resemblance between Smuts's philosophical views and those of Henri Bergson and C. Lloyd Morgan. All three philosophers stressed the creativity of evolution, its engendering of novelties whose presence invalidates mechanistic materialism. All were critical of Darwinism and opposed it with arguments and assertions couched in highly general terms. Smuts differed from the other two philosophers in refusing to state explicitly that the holistic factor is spiritual or akin to mind. But at bottom it remains as inscrutable as Bergson's *élan vital* or Morgan's directing Activity.

See also African Philosophy; Bergson, Henri; Darwinism; Holism and Individualism in History and Social Science; Morgan, C. Lloyd.

Bibliography

McDougall, William. "The Confusion of the Concept." *Journal of Philosophical Studies* 3 (1928): 440–442.

Morgan, C. Lloyd. Review of *Holism and Evolution. Journal of Philosophical Studies* (1927): 85–89.

T. A. Goudge (1967)

SOCIAL AND POLITICAL PHILOSOPHY

It is generally agreed that the central task of social and political philosophy is to provide a justification for coercive institutions. Coercive institutions range in size from the family to the nation-state and world organizations, like the United Nations, with their narrower and broader agendas for action. Yet essentially, they are institutions that at least sometimes employ force or the threat of force to control the behavior of their members to achieve either minimal or wide-ranging goals. To justify such coercive institutions, we need to show that the authorities within these institutions have a right to be obeyed and that their members have a corresponding duty to obey them. In other words, we need to show that these institutions have legitimate authority over their members.

In philosophical debate at the beginning of the twenty-first century, a number of competing justifications for coercive institutions have been defended: (1) a libertarian justification, which appeals to an ideal of liberty; (2) a socialist justification, which appeals to an ideal of equality; (3) a welfare liberal justification, which appeals to an ideal of contractual fairness; (4) a communitarian justification, which appeals to an ideal of the common good; and (5) a feminist justification, which appeals to an ideal of a gender-free society. Each of these justifications needs to be examined in order to determine which, if any, are morally defensible.

LIBERTARIANISM

Libertarians frequently cite the work of F. A. Hayek, particularly his *Constitution of Liberty* (1960), as an intellectual source of their view. Hayek argues that the libertarian ideal of liberty requires "equality before the law" and "reward according to market value" but not " substantial equality" or "reward according to merit." Hayek further argues that the inequalities due to upbringing, inheritance, and education that are permitted by an ideal of liberty actually tend to benefit society as a whole.

In basic accord with Hayek, contemporary libertarians, like John Hospers (1971), Robert Nozick (1974), Tibor Machan (2004), and Jan Narveson (1998), define liberty negatively as "the state of being unconstrained by other persons from doing what one wants" rather than positively as "the state of being assisted by other persons in doing what one wants." Libertarians go on to characterize their social and political ideal as requiring that each person should have the greatest amount of liberty commensurate with the same liberty for all. From this ideal, libertarians claim that a number of more specific requirements, in particular a right to life, a right to freedom of speech, press, and assembly, and a right to property, can be derived.

The libertarian's right to life is not a right to receive from others the goods and resources necessary for preserving one's life; it is simply a right not to be killed. So understood, the right to life is not a right to receive welfare. In fact, there are no welfare rights according to the libertarian view. Correspondingly, the libertarian's

understanding of the right to property is not a right to receive from others the goods and resources necessary for one's welfare, but rather a right to acquire goods and resources either by initial acquisition or by voluntary agreement. By defending rights such as these, libertarians support only a limited role for coercive institutions. That role is simply to prevent and punish initial acts of coercion—the only wrongful actions for libertarians. Thus, libertarians are opposed to all forms of censorship and paternalism, unless they can be supported by their ideal of liberty.

Libertarians do not deny that it is a good thing for people to have sufficient goods and resources to meet their basic nutritional needs, but libertarians do deny that coercive institutions should be used to provide for such needs. Some good things, such as the provision of welfare to the needy, are requirements of charity rather than justice, libertarians claim. Accordingly, failure to make such provisions is neither blameworthy nor punishable.

SOCIALISM

In contrast with libertarians, socialists take equality to be the ultimate social and political ideal. In the *Communist Manifesto* (1848), Karl Marx and Friedrich Engels maintain that the abolition of bourgeois property and bourgeois family structure is a necessary first requirement for building a society that accords with the political ideal of equality. In *Critique of the Gotha Program*, Marx provides a much more positive account of what is required to build a society based upon the political ideal of equality. In such a society, Marx claims that the distribution of social goods must conform, at least initially, to the principle "from each according to his ability, to each according to his contribution." But when the highest stage of communist society has been reached, Marx adds, distribution will conform to the principle "from each according to his ability, to each according to his need." Contemporary socialists like Kai Nielson and Carol Gould continue to endorse these tenets of Marxism.

At first hearing, these tenets of Marxism might sound ridiculous to someone brought up in a capitalist society. The obvious objection is, how can you get persons to contribute according to their ability if income is distributed on the basis of their needs and not on the basis of their contributions? The answer, according to socialists, is to make the work that must be done in a society as much as possible enjoyable in itself. As a result, people will want to do the work they are capable of doing because they find it intrinsically rewarding. For a start, socialists might try to get people to accept presently existing, intrinsically rewarding jobs at lower salaries—top executives, for example, to work for $300,000, rather than $900,000 or more, a year. Yet ultimately, socialists hope to make all jobs as intrinsically rewarding as possible, so that after people are no longer working primarily for external rewards, while making their best contributions to society, distribution can proceed on the basis of need.

Socialists propose to implement their egalitarian ideal by giving workers democratic control over the workplace. They believe that if workers have more to say about how they do their work, they will find their work intrinsically more rewarding. As a consequence, they will be more motivated to work, because their work itself will be meeting their needs. Socialists believe that extending democracy to the workplace will necessarily lead to socialization of the means of production and the end of private property. By making jobs intrinsically as rewarding as possible, in part through democratic control of the workplace and an equitable assignment of unrewarding tasks, socialists believe people will contribute according to their ability even when distribution proceeds according to need. Liberation theology has also provided an interpretation of Christianity that is sympathetic to this socialist ideal.

Nor are contemporary socialists disillusioned by the collapse of the Soviet Union and the transformation of the countries in Eastern Europe. Judging the acceptability of the socialist ideal of equality by what took place in these countries would be as unfair as judging the acceptability of the libertarian ideal of liberty by what takes place in countries like Guatemala or Singapore, where there is a free market but very little political liberty. By analogy, it would be like judging the merits of college football by the way Vanderbilt's or Columbia s team play rather than by the way Florida's or USC's team play. Actually, a fairer comparison would be to judge the socialist ideal of equality by what takes place in countries like Sweden and to judge the libertarian ideal of liberty by what takes place in the United States. Even these comparisons, however, are not wholly appropriate because none of these countries fully conforms to those ideals.

WELFARE LIBERALISM

Finding merit in both the libertarian's ideal of liberty and the socialist's ideal of equality, welfare liberals attempt to combine both liberty and equality into one political ideal that can be characterized by contractual fairness. A classical example of this contractual approach is found in the political works of Immanuel Kant. Kant claims that a civil state ought to be founded on an original contract satisfy-

ing the requirements of freedom, equality, and independence. According to Kant, it suffices that the laws of a civil state are such that people would agree to them under conditions in which the requirements of freedom, equality, and independence obtain.

The Kantian ideal of a hypothetical contract as the moral foundation for coercive institutions has been further developed by John Rawls in *A Theory of Justice* (1971). Rawls, like Kant, argues that principles of justice are those principles that free and rational persons who are concerned to advance their own interests would accept in an initial position of equality. Yet Rawls goes beyond Kant by interpreting the conditions of his "original position" to explicitly require a "veil of ignorance." This veil of ignorance, Rawls claims, has the effect of depriving persons in the original position of the knowledge they would need to advance their own interests in ways that are morally arbitrary.

According to Rawls, the principles of justice that would be derived in the original position are the following: (1) a principle of equal political liberty; (2) a principle of equal opportunity; (3) a principle requiring that the distribution of economic goods work to the greatest advantage of the least advantaged. Rawls holds that these principles would be chosen in the original position because persons so situated would find it reasonable to follow the conservative dictates of the "maximin" strategy and maximize the minimum), thereby securing for themselves the highest minimum payoff. In his *Political Liberalism* (1993), Rawls explains how these principles could be supported by an overlapping consensus, and thus would be compatible with a pluralistic society whose members endorse diverse comprehensive conceptions of the good, and in his *The Law of Peoples* (1999), Rawls attempts to extend his theory of justice to the international realm.

COMMUNITARIANISM

Another prominent social and political ideal defended by contemporary philosophers is the communitarian ideal of the common good. As one might expect, many contemporary defenders of a communitarian social and political ideal regard their conception as rooted in Aristotelian moral theory. Alasdair MacIntyre in *After Virtue* (1981) sees his social and political theory as rooted in Aristotelian moral theory, but it is an Aristotelian moral theory that has been refurbished in certain respects. Specifically, MacIntyre claims that Aristotelian moral theory must, first of all, reject any reliance on a metaphysical biology. Instead of appealing to a metaphysical biology, MacIntyre proposes to ground Aristotelian moral theory on a conception of a practice. A practice, for MacIntyre, is "any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to and partially definitive of that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved are systematically extended" (1981, p.175). As examples of practices, MacIntyre cites arts, sciences, games, and the making and sustaining of family life.

MacIntyre then partially defines the virtues in terms of practices. A virtue, such as courage, justice or honesty, is "an acquired human quality the possession and exercise of which tends to enable us to achieve those goods which are internal to practices and the lack of which prevents us from achieving any such goods" (1981, p.178). However, MacIntyre admits that the virtues which sustain practices can conflict (e.g., courage can conflict with justice) and that practices so defined are not themselves above moral criticism.

Accordingly, to further ground his account, MacIntyre introduces the conception of a telos, or good of a whole human life conceived as a unity. It is by means of this conception that MacIntyre proposes to morally evaluate practices and resolve conflicts between virtues. For MacIntyre, the telos of a whole human life is a life spent in seeking that telos; it is a quest for the good human life and it proceeds with only partial knowledge of what is sought. Nevertheless, this quest is never undertaken in isolation but always within some shared tradition. Moreover, such a tradition provides additional resources for evaluating practices and for resolving conflicts while remaining open to moral criticism itself.

MacIntyre's characterization of the human telos in terms of a quest undertaken within a tradition marks a second respect in which he wants to depart from Aristotle's view. This historical dimension to the human telos that MacIntyre contends is essential for a rationally acceptable communitarian account is absent from Aristotle' s view. A third respect in which MacIntyre's account departs from that of Aristotle concerns the possibility of tragic moral conflicts. As MacIntyre points out, Aristotle only recognized moral conflicts that are the outcome of wrongful or mistaken action. Yet MacIntyre, following Sophocles, wants to recognize the possibility of additional conflicts between rival moral goods that are rooted in the very nature of things.

Initially, rather than draw out the particular requirements of his own social and political theory, MacIntyre defended his theory by attacking rival theories, and, by and large, he focused his attacks on liberal social and political theories; in this respect he shares common ground with contemporary deconstructionists. Thus, MacIntyre argues in his "Privatization of the Good" that virtually all forms of liberalism attempt to separate rules defining right action from conceptions of the human good. MacIntyre contends that these forms of liberalism not only fail but have to fail because the rules defining right action cannot be adequately grounded apart from a conception of the good. For this reason, MacIntyre claims, only some refurbished Aristotelian theory that grounds rules supporting right action in a complete conception of the good can ever hope to be adequate.

In his most recent book, *Rational Dependent Animals* (1999), however, MacIntyre's defense of the communitarian ideal of the common good has now moved in a socialist or Marxist direction. In this book, Macintyre argues that for independent practical reasoners, Marx's principle for a socialist society—to each according to his or her contribution—is appropriate, but between those capable of giving and those most dependent, it is Marx's principle for a communist society—from each according to his or her ability, to each according to his or her need—that is appropriate.

FEMINISM

Defenders of a feminist social and political ideal present a distinctive challenging critique to defenders of other social and political ideals. In *The Subjection of Women* (1869), John Stuart Mill, one of the earliest male defenders of women's liberation, argues that the subjection of women was never justified but was imposed upon women because they were physically weaker than men; later this subjection was confirmed by law. Mill argues that society must remove the legal restrictions that deny women the same opportunities enjoyed by men. However, Mill does not consider whether, because of past discrimination against women, it may be necessary to do more than simply remove legal restrictions: he does not consider whether positive assistance may also be required.

Usually it is not enough simply to remove unequal restrictions to make a competition fair among those who have been participating. Positive assistance to those who have been disadvantaged in the past may also be required, as would be the case in a race in which some were unfairly impeded by having to carry ten-pound weights for part of the race. To render the outcome of such a race fair, we might want to transfer the ten-pound weights to the other runners in the race, and thereby advantage the previously disadvantaged runners for an equal period of time. Similarly, positive assistance, such as affirmative action or preferential treatment programs, may be necessary if women who have been disadvantaged in the past by sexism are now going to be able to compete fairly with men. According to feminists, the argument for using affirmative action or preferential treatment to overcome sexism in society is perfectly analogous to the argument for using affirmative action or preferential treatment to overcome racism in society.

In *Justice, Gender and the Family* (1989), Susan Okin argues for the feminist ideal of a gender-free society. A gender-free society is a society in which basic rights and duties are not assigned on the basis of a person's biological sex. Being male or female is not the grounds for determining what basic rights and duties a person has in a gender-free society. Since a conception of justice is usually thought to provide the ultimate grounds for the assignment of rights and duties, we can refer to this ideal of a gender-free society as "feminist justice."

Okin goes on to consider whether John Rawls's welfare liberal conception of justice can support the ideal of a gender-free society Noting Rawls's initial failure to apply his "original position" concept to family structures, Okin is skeptical about the possibility of using a welfare liberal ideal to support feminist justice. She contends that in a gender-structured society like our own, male philosophers cannot achieve the sympathetic imagination required to see things from the standpoint of women. In a gender-structured society, Okin claims, male philosophers cannot do the "original position-type thinking required by the welfare liberal ideal because they lack the ability to put themselves in the position of women. According to Okin, the "original position" can only really be achieved in a gender-free society.

Yet at the same time that Okin despairs of doing "original position-type thinking in a gender-structured society, like our own, she herself purportedly does a considerable amount of just that type of thinking. For example, she claims that Rawls's principles of justice "would seem to require a radical rethinking not only of the division of labor within families but also of all the nonfamily institutions that assume it." She also claims that "the abolition of gender seems essential for the fulfillment of Rawls's criterion of political justice" (1989, p. 104).

PRACTICAL REQUIREMENTS

Unfortunately, unless we can show that either libertarianism, socialism, welfare liberalism, communitarianism, or feminism, or some combination of these ideals is most morally defensible, it will be difficult to know which practical requirements one should endorse. However, assuming we have obligations to distant peoples and future generations, it may be possible to show that the libertarian's own ideal of liberty leads to a right to welfare that is acceptable to welfare liberals, and that when this right is extended to distant peoples and future generations, it requires something like the equality that socialists endorse. This would effect a practical reconciliation of sorts among seemingly opposing social and political ideals.

There is also the question of whether we have obligations to animals and other nonhuman living beings. Until recently, there was very little discussion of whether humans have such obligations. It was widely assumed, without much argument, that we have obligations only to humans. However, this lack of argument has recently been challenged by defenders of animal rights on grounds of speciesism. Speciesism, they claim, is the prejudicial favoring of the interests of members of one' s own species over the interests of other species. Obviously, determining whether this charge of speciesism can be sustained is vital to providing a justification of coercive institutions, particularly the coercive institutions of animal experimentation and factory farming, and thus it is vital to fulfilling the central task of social and political philosophy as well.

See also Aristotle; Civil Disobedience; Communitarianism; Cosmopolitanism; Democracy; Engels, Friedrich; Feminist Social and Political Philosophy; Kant, Immanuel; Liberation Theology; Libertarianism; Liberty; MacIntyre, Alasdair; Marx, Karl; Mill, John Stuart; Multiculturalism; Nationalism; Nozick, Robert; Pluralism; Postcolonialism; Rawls, John; Republicanism; Socialism; Speciesism.

Bibliography

Gould, Carol. *Globalizing Democracy and Human Rights*. Cambridge, U.K.: Cambridge University Press, 2004.

Hayek, Friedrich A. *The Constitution of Liberty*. Chicago: University of Chicago Press, 1960.

Hospers, John. *Libertarianism*. Los Angeles: Nash, 1971.

Machan, Tibor. *Passion for Liberty*. Lanham: Rowman & Littlefield, 2004.

MacIntyre, Alasdair. *After Virtue*. Notre Dame, IN: University of Notre Dame Press, 1981.

MacIntyre, Alasdair. *Dependent Rational Animals*. Chicago: Open Court, 1999.

MacIntyre, Alasdair. "The Privatization of the Good." *Review of Politics* 52 (1990): 1–20.

MacIntyre, Alasdair. *Three Rival Versions of Moral Enquiry*. Notre Dame, IN: University of Notre Dame Press, 1990.

Marx, Karl, and Friedrich Engels. *The Communist Manifesto*. Originally published as *Manifest der Kommunistischen Partie* (1848). Translated by Paul Sweeney. New York: Monthly Review Press, 1998.

Mill, John Stuart. *The Subjection of Women* (1869). Orchard Park, NY: Broadview Press, 2000.

Narveson, Jan. *The Libertarian Idea*. Philadelphia: Temple University Press, 1988.

Nielson, Kai. *Liberty and Equality*. Totowa, NJ: Rowman and Littlefield, 1985.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Okin, Susan. *Justice, Gender, and the Family*. New York: Basic Books, 1989.

Rawls, John. *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.

Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Sterba, James P. *Justice for Here and Now*. New York: Cambridge University Press, 1998.

James P. Sterba (1996, 2005)

SOCIAL CONSTRUCTIONISM

Social constructionism (sometimes "constructivism") is a version of constructivism. The idea that human beings in some measure construct the reality they perceive can be found in many philosophical traditions. The pre-Socratic philosopher Xenophones, for instance, argued that humans construct gods in their own image (Fragment 16), a possibility that is also criticized in the Jewish, Christian, and Islamic religious traditions (among others). But the idea that human beings epistemologically construct the reality they perceive is first given extended philosophical articulation in the work of Immanuel Kant (1724–1804). In the nineteenth century a constructivism of sorts emerged as political theory in the work of Karl Marx (1818–1883) and others. Then, in the twentieth century, constructivism took new forms in psychology, in sociology, and in science, technology, and society (STS) studies.

CONSTRUCTIVISM IN PSYCHOLOGY

A root form of social constructionism is found in psychological constructivism. Illuminating research by the British psychologist Frederick Bartlett (1886–1969) revealed how humans use prior knowledge to make sense of new phenomena. In his landmark study *Remembering* (1932), Bartlett presented an unfamiliar indigenous American folk tale to students at Cambridge University. Later each subject was asked to recall the story in as much detail as possible. Bartlett was able to show how each retelling was a unique *reconstruction* of the story rather than a *reproduction* of the original. Subjects tended to replace unfamiliar elements of the story with objects drawn from their own experience. Bartlett concluded that in coming to understand the story, his students tended to make use of pre-existing mental structures or *schemata*, which proved essential both for originally comprehending the story and for subsequent recall.

The notion of schemata is central as well to Jean Piaget's (1896–1980) theory of intelligence. The Swiss psychologist undertook pioneering work on childhood intellectual development. From years of careful observations of and conversations with children and watching them function in problem-solving activities, Piaget argued that cognitive development is an adaptive process of schema correction by means of assimilation and accommodation. We assimilate new information by fitting it within existing cognitive structures. Where preexisting schema cannot incorporate a new experience, we adjust our mental structures to accommodate them. For Piaget, learning is not a passive activity of replication and data storage but an active process of invention and creation. Piaget's resultant genetic epistemology describes how increasingly complex intellectual processes are built on top of more primitive structures in regularly occurring stages.

Lev Semyonovitch Vygotsky (1896–1934), a Piaget contemporary, also studied the cognitive development of children in Soviet Russia during the Stalin years and noted how children engaged in a problem-solving activity invariably speak about what they are doing. This led to his theory of speech as a means for making sense of the activity. Although children's use of tools during their preverbal period is comparable to that of apes, as soon as speech and signs are incorporated into any action, the action becomes transformed and organized along entirely new lines.

Language is thus central to complex reasoning and higher order thinking. Intelligence is the readiness to use culturally transmitted knowledge and practice as prostheses of the mind, and learning is inherently social; learned social speech becomes inner speech through development. Vygotsky came to believe that speech precedes thought and that human thought is a social phenomenon that develops from society to the individual. The idea that cognition emerges out of social activity is central to Vygotsky's work. This is also a view that has become at once widely adopted—being applied especially in educational theory—and controversial, especially various forms of cognitive psychology.

SOCIAL CONSCTRUCTIONISM IN SOCIAL THEORY

The American social philosopher George Herbert Mead (1863–1931) took constructivism into sociology with a theory of self consciousness as originating from social interaction. In his posthumously published *Mind, Self, and Society* (1934), Mead argued that personal identity is constructed through social relationships. In the context of play, for instance, children take on the roles of others, eventually learning to view themselves from the standpoint of a "generalized other." Children's games thus function as instruments for personal and social development, especially when children adopt attitudes of those who in some sense control them or on whom they depend. For Mead the self is a dialectical conversation between the "me" and the "I"—"me" being the social self and "I" the creative self that responds to the "me" in multiple contexts to form, over time, the ontogenic, historical image of one's self.

The theorists Peter Berger and Thomas Luckmann cite Mead as a major source for their seminal sociological text *The Social Construction of Reality* (1966). In this treatise, Berger and Luckmann extend Mead's ontogenetic observations on the self to include all phenomena that we encounter in a social world. They describe the dialectic relationship between the subjective reality of the individual and the objective reality of society that emerges in a universe of discourse that is continuously under construction. Through interaction and conversation with others, knowledge is internalized, then externalized, becoming at once a subjective perception and an objective reality. From such a process of socialization we construct our daily lives.

Much social constructionism implies some degree of subjectivism. From an analysis of intentionality and how it plays out in a social context, however, the philosopher John Searle (1995) has argued that socially constructed reality exhibits its own distinctive type of objectivity. Searle's realism distinguishes between "brute facts" that

exist independently of what any humans think and "social facts" that depend on human thinking while being independent of what any one human thinks. Human beings construct a social reality through common intentions that assign functions to physical objects, as when a certain type of paper comes to be treated as money.

SOCIAL CONSTRUCTIONISM IN SCIENCE AND TECHNOLOGY

Epistemological constructivism has taken special forms in the development of cybernetics, evolutionary epistemology, and the philosophy of mathematics. But insofar as cybernetics moved from analyses of interactions between organisms and their physical environments to consideration of communication in a social environment, social cybernetics offered as well a science and a technology of social interactive constructions. Yet the cybernetic approach has been only marginally influential on social constructionism in general.

One of the most contested areas of social constructionism is not in science and technology but in studies about science and technology. Ludwik Fleck (1979) first proposed, in a controversial interpretation of the medical conceptualization of disease, that even some supposedly brute facts of science were socially constructed. This idea was picked up and developed by Thomas Kuhn (1962), which subsequently led to the development of a research program in the sociology of scientific knowledge (SSK). The sociology of scientific institutions, as initiated by Robert K. Merton (1910–2003) in the 1930s, came under increasing criticism in the 1970s for its idealization of science and its failures to treat the production of scientific truth and falsity in a symmetrical manner. Drawing on the ideas from the later Ludwig Wittgenstein about the influence of language games and forms of life on human understanding, David Bloor (1983) and others proposed that social factors influenced not only the production of falsehood (a weak SSK program) but also any consensus about truth (the strong SSK program).

The SSK program in conceptual and analytic criticism was quickly complemented by empirical studies of laboratory practices and how such practices themselves contribute to the production of scientific knowledge. Employing ethnographic approaches, Bruno Latour and Steve Woolgar (1979) and Karin Knorr Cetina (1981) argued that knowledge production is seldom the rational, linear process of hypothesis testing leading to article publication found in the standard image of science. Behind the scenes science is a mangle of practical skills, instrumental jiggering, personal relationships, interpretative debates, and consensus building that deploys a variety of rhetorical strategies to frame both problems and experimental results.

The full extent to which scientific knowledge is a social construction or laboratory production—and what this might imply for science, scientists, as science as a social institution—has been subject to extensive debate in the so-called "science wars" between scientists and their social scientific critics. Among the most philosophically astute assessments of this research program and ensuing debate has been Ian Hacking's *Social Construction of What?* (1995).

The program for a parallel analysis of the social construction of technology (SCOT) has been almost as controversial as social constructivism applied to science, but for different reasons. As Louis Bucciarelli (1994) has shown with his ethnographic examination of the engineering design process, social and personal factors of all sorts readily influence engineering products, processes, and systems. The question is whether this means that those such as Jacques Ellul (1954) or Hans Jonas (1984) who have raised ethical and political questions about the dominance of modern technology in human affairs are simply mistaken in their worries. For proponents of SCOT or one of its related programs such as actor-network theory, critics have too often criticized technology as a kind of "black box" that they failed to examine in sufficient detail. But critics such as Langdon Winner (1994) have responded that "opening the black box" can also be an exercise in avoidance of more fundamental questions.

Relations between social constructivism in psychology, sociology, and STS deserve further examination. Moreover, arguments concerning the social construction of science and technology exhibit unexplored affinities with the pragmatic epistemologies of the "fixation of belief" (C. S. Peirce), the merger of science and technology in the general category of tools (John Dewey), and criticisms of strict empiricism (Willard van Orman Quine). Indeed, social constructivism presents a broad philosophical interpretation of personal and public life, from the epistemological to the ethical, in ways that will likely continue to exercise considerable influence in twenty-first century thought.

See also Constructivism and Conventionalism; Critical Theory; Dewey, John; Feuerbach, Ludwig Andreas; Kant, Immanuel; Kuhn, Thomas; Marx, Karl; Mead, George Herbert; Peirce, Charles Sanders; Personal Identity; Piaget, Jean; Psychology; Quine, Willard Van Orman; Searle, John; Social and Political Philosophy;

Wittgenstein, Ludwig Josef Johann; Xenophanes of Colophon.

Bibliography

Bartlett, Frederick. *Remembering: A Study in Experimental and Social Psychology*. Cambridge, U.K.: Cambridge University Press, 1932.

Berger, Peter, and Thomas Luckmann. *The Social Construction of Reality*. New York: Doubleday, 1966.

Bloor, David. *Ludwig Wittgenstein: A Social Theory of Knowledge*. New York: Columbia University Press, 1983.

Bucciarelli, Louis L. *Designing Engineers*. Cambridge, MA: MIT Press, 1994.

Ellul, Jacques. *La technique, ou l'enjeu du siècle*. Paris: A. Colin, 1954. English translation: *The Technological Society*. Translated by J. Wilkerson. New York: Knopf, 1964.

Fleck, Ludwik. *Genesis and Development of a Scientific Fact*, edited T. J. Trenn and R. K. Merton, 1935. Translated by F. Bradley. Chicago: University of Chicago Press, 1979.

Hacking, Ian. *The Social Construction of What?* Cambridge, MA: Harvard University Press, 1999.

Jonas, Hans. *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Frankfurt: Insel Verlag, 1979. English translation: *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. Translated by H. Jonas and D. Herr. Chicago: University of Chicago Press, 1984.

Knorr Cetina, Karin. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press, 1981.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Latour, Bruno, and Steve Woolgar. *Laboratory Life: The Construction of Scientific Facts*. Beverly Hills, CA: Sage, 1979.

Mead, George Herbert. *Mind, Self, and Society*, edited by C. W. Morris. Chicago: University of Chicago Press, 1934.

Piaget, Jean. *Genetic Epistomology*. Translated by E. Duckworth. New York: Columbia University Press, 1970.

Piaget, Jean. *The Construction of Reality in the Child*. London: Routledge and Kegan Paul, 1955.

Searle, John. *The Construction of Social Reality*. New York: Free Press, 1995.

Vygotsky, Lev S. *Mind in Society: The Development of Higher Psychological Processes*, edited by M. Cole. Cambridge, MA: Harvard University Press, 1978.

Winner, Langdon. "Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology." *Science, Technology, and Human Values*, 18, 362–378.

Carl Mitcham (2005)
Martin Ryder (2005)

SOCIAL CONTRACT

"Social contract" is the name given to a group of related and overlapping concepts and traditions in political theory. Like other such aggregations in philosophy and intellectual history, it has at its center an extremely simple conceptual model, in this case that the collectivity is an agreement between the individuals who make it up. This model suggests that it is proper to ask whether the agreement was or is voluntary in character and whether, therefore, the individual can decide to withdraw either because he no longer agrees or because the conditions that are or were understood in the agreement are not being maintained. It suggests furthermore that the individual should be thought of as logically prior to the state or to society, and that it is meaningful to speculate on situations in which individuals existed but no collectivity was in being. From a historical point of view, it is therefore relevant to discuss periods during which no collectivity existed, when what is traditionally called a "state of nature" prevailed, and to contrast these periods with times when by agreement the collectivity had come into existence, that is, with what is traditionally called a "state of society."

The concept of a prepolitical state of nature that can be brought to an end by agreement can thus be applied to geographical areas of human society as well as to periods of time. Individuals in such areas must be considered, as Thomas Hobbes himself said, "to have no government at all and to live at this day in that brutish manner." Although this may seem to be the least persuasive of the elements belonging to the social contract, its parallel in relationships between politically constituted societies or states, that is to say, in the international state of nature, is perhaps the most useful and persistent. It seems still to command allegiance in the study of international relations. The actual process of agreeing ("contracting," "compacting," "covenanting") to end the state of nature and establish a state of society has been the subject of extensive analysis and elaboration by political and social theorists. Distinctions have been drawn, more precisely perhaps by academic commentators in modern times than by contractarian writers themselves, between a social contract and a governmental contract.

The social contract proper (*pactum societatis, pacte d'association, Gesellschaftsvertrag*) is thought of as bringing individuals together in society, and the governmental contract (*pactum subjectionis, pacte du gouvernement, Herrschaftsvertrag*) as establishing a formal government. As might be expected, the nature and form of the contract or contracts has been thought of in a variety of ways. In some systems the contract is a once-and-for-all, irrevocable act understood to have been performed in the remote past (Richard Hooker), but in others it appears as a continuing understanding that is perpetually being

renewed and is regarded rather as a trust than as a contract (John Locke). The parties to the various contracts differ also: Sometimes agreements are made between individuals only, sometimes between individuals and governments or sovereigns, sometimes between a body of individuals acting as a fictitious person (*persona ficta*) and either the sovereign or a member of the body. In such ways as these a whole set and succession of interrelated contractual agreements have occasionally been presumed, as in the case of the seventeenth-century German political theorist Samuel Pufendorf and his followers in the eighteenth century.

The theory of a social contract belongs with the individualist attitude to state and society; indeed the simple conceptual model of agreement for the collectivity in all its possible shapes seems to inform the entire individualistic outlook. Contractual political theory is, therefore, universally associated with the rights of the individual person, with consent as the basis of government, and with democratic, republican, or constitutional institutions. It has also been regarded as a part of early capitalist individualism, and in Victorian England a great watershed was held to exist between a condition in which status ruled relationships and one in which contract ruled them. Notwithstanding this assumption, the social contract is perfectly reconcilable with the most absolute of despotic rule and with the complete negation of constitutionalism or the rule of law. Hobbes is the classic case here, for his two alternative accounts of how society and government came simultaneously into being are designed to tie every citizen to unquestioning obedience to a supreme, irresistible, indivisible sovereign whose dictates are the law. Benedict de Spinoza makes a rather similar use of contractual principles, but the political theory of Jean-Jacques Rousseau, although expounded in contractual form, has collectivist tendencies, since it endows political society with the capacity to make people moral. Rousseau's major political work, *Du Contrat social,* must be looked upon as the point of departure of the quite separate and traditionally quite irreconcilable outlook whose model is the theory of the general will.

EXPLANATORY VALUE

If the collectivity is understood as embodying agreement, it does not necessarily follow that any such agreement between parties ever actually took place in historical time. Nor does it follow that there may be people in the world still living in a prepolitical, precontractual situation or that those now within constituted society could ever revert to the nonpolitical condition. A contractarian political theory, therefore, can be entirely hypothetical, analyzing state and society as if agreement must always be presumed. Such an argument can provide a penetrating critique of existing arrangements and of their rationale: It can be used in a reformist direction, to suggest what ought to be the aims and ends of statesmen. No reversion to a literal state of nature need be implied by criticism of this kind, only that this or that action or abuse requires a remedy in accordance with the suggested criterion of an assumed agreement.

In this hypothetical form the contract theory is still of importance to political philosophy. It has recently been used by John Rawls in his articles "Justice as Fairness" and "Distributive Justice" to develop an account of justice alternative to the utilitarian (previously assumed to have outmoded contractarianism). Contemporary appreciations of the great contractarian writers (for example, by Howard Warrender, C. B. Macpherson, and A. G. Wernham), especially of Hobbes but also of Locke, Spinoza, and David Hume, and even of Rousseau, have tended to insist that the classic theories are hypothetical, which makes it possible to free the theories to a surprising extent from the lumber that had attached to them—the unacceptable histories of the human race, the fanciful anthropology and sociology. Moreover, the assumptions of natural law can thus be put aside.

NATURAL LAW

The reinterpretation of social contract theory is an important example of the way in which past political theory can enter into present theoretical analysis independently of chains of influence and continuous traditions. Still, the reinterpretation may lead to a serious distortion of the truth about the actual contents of contractarian treatises on politics. All the many members of the school of natural law, including those named above, did in fact assume that their contractual claims were literal as well as hypothetical. They all made dogmatic statements about the history of humanity and the condition of savages. Moreover all of them, though here writers like Hobbes and Hume are in special categories, subscribed to the general system of natural law in one form or another.

The concept of natural law provided the fixed and enduring framework within which the contract ending the state of nature could be concluded, and subsequent breaches or revisions of the contract could be related to the original act. Therefore, natural law had to be assumed if the contract was to be taken at all literally. The duty to keep promises, on which any contract rests, could hardly come into being with the contract itself, and this duty

must persist should the contract be broken, if only to make a new one possible. When the Commons of England in January 1689 accused their former king, James II, of "breaking the original contract betwixt King and people," they did so in the secure belief that this was an offense that was and always would be punishable under natural law. It is understandable, then, that the history of the idea of a social contract has been largely the same as that of natural law itself.

HISTORY

The origins of social contract theory and of natural law can be sought in the Roman Stoicism of Cicero and in the system of Roman law. The development of social contract into a standard feature of the Western Christian attitude can be seen in the Middle Ages, and its apotheosis can be observed in the period between the Reformation and the eighteenth century. It is usual in fact to insist that the rise of the contractarian attitude to predominance in European political thought came about because of the Reformation. Certainly the justification of the right of a Protestant minority in a Catholic country, and of that of a Catholic minority in a Protestant country, to its own form of religious worship came about because of the gradual acceptance of contractarian notions by Reformation and post-Reformation political and legal thinkers and even by some politicians and sovereigns. The slow and hesitant growth of religious toleration would undoubtedly have been even more retarded if natural law and the social contract had not been at hand to provide a definition of the individual citizen, his individual rights, and the nature of his relationship to political authority. Accordingly, we find that the French religious wars of the 1560s, 1570s, and 1580s, together with the revolt of the Dutch against the throne of Spain, which began in 1568, brought about the elaboration of contractarian ideas. In both these cases embattled Calvinists were asserting their political as well as their religious rights against Catholic authorities, but in England at the same time it was the Catholics who needed contractarian justification for their rights, even finally their rights to resist government.

The *Monarchomachi* ("bearers of the sword against monarchs"), as the French writers were called, developed the contract between people and sovereign in various directions, and in the famous *Vindiciae Contra Tyrannos* (1579) it justified a recognizably revolutionary doctrine. In Holland the contract was codified further and became in the works of Johannes Althusius and Hugo Grotius an informing principle of political life as well as of the relations between sovereign and people. (Grotius's great work, however, the *De Jure Belli ac Pacis* of 1625, acquired and retains its fame because of its application of natural law and contractarian principles to international law.)

All these ideas and all these experiences—particularly the experience of religious separatism developing into civil war—can be seen at work in Hobbes, the most impressive of all contractarian theorists. In Hobbes's *Leviathan* (1651), the state of nature was a state of war, a propertyless anarchy brought to an end only by the contract of absolute submission. Hobbes made such devastating use of the destructive potentialities of the social contract in criticism of the conventional thinking about natural law that all succeeding systems can be looked upon to some extent as commentaries upon him. This is truest of Spinoza (*Tractatus Theologico-Politicus,* 1670; *Tractatus Politicus,* 1677) and until recently was thought to be true of Hobbes's eminent and enormously influential successor in England, Locke.

Locke's *Two Treatises of Government* (written 1679–1683, published 1689) are now known to have been written as an attack on Robert Filmer, not on Hobbes, and Locke's relatively peaceful and sociable state of nature, brought to an end by a very limited contract, has only a somewhat distant relationship with Hobbes's "war of all against all." It is interesting that Filmer should have been the most effective critic of the concept of a state of nature and of the possibility and relevance of contract and that his traditional, patriarchal authoritarianism was to a large extent immune from contractarian notions.

It was not traditionalism, however, which broke down contractarian assumptions within a generation of the death of Locke in 1704, but rather the rapid defeat of the natural law outlook by utilitarian criticism in England and by general will notions in France and elsewhere. Contract lost its persuasiveness as the rationalist outlook on the nature of law gave way to the historical outlook early in the nineteenth century. The development of observational anthropology and empirical sociology in more recent times makes it entirely unlikely that contract in anything but a strictly hypothetical form will ever be adopted again by political theorists.

This conventional account of the history of contract could be corrected and extended by reference to the simple model of the collectivity as agreement with which this entry began. This is so obvious an image that it can be found in some form in any political system, even in the refusal of Socrates to escape from his prison and avoid the poison on the ground that he owed obedience to his native city because of the benefits he had received as a citizen. It seems likely that every political theory must be

contractual, at least to some degree, in this very wide sense.

Nevertheless, since contract proceeds by abstracting the individual from society, and then by reassembling individuals again as society although they are by definition asocial abstractions, the general contractual social and political scheme seems incurably faulty, quite apart from the empirical objections to it on the part of contemporary social scientists.

See also Althusius, Johannes; Cicero, Marcus Tullius; Filmer, Robert; General Will, The; Grotius, Hugo; Hobbes, Thomas; Hooker, Richard; Hume, David; Locke, John; Natural Law; Philosophy of Law, History of; Pufendorf, Samuel von; Rawls, John; Reformation; Rousseau, Jean-Jacques; Socrates; Spinoza, Benedict (Baruch) de; State; Stoicism.

Bibliography

The standard account in English is J. W. Gough, *The Social Contract* (1st ed., Oxford: Clarendon Press, 1936; rev. ed., 1957). Ernest Barker analyzes classical notions of contract very succinctly in a compilation he edited for the World's Classics series titled *Social Contract: Essays by Locke, Hume, and Rousseau* (New York: Oxford University Press, 1948); he has also translated the relevant part of Otto Gierke's monumental general treatise, *Das Deutsche Genossenschaftsrecht* (1913), as *Natural Law and the Theory of Society,* 2 vols. (Cambridge, U.K.: Cambridge University Press, 1934). In addition, see A. G. Wernham, ed., *Spinoza, Political Works* (Oxford: Clarendon Press, 1958); Peter Laslett, ed., *John Locke, Two Treatises of Government* (Cambridge, U.K.: Cambridge University Press, 1960); Howard Warrender, *Political Philosophy of Hobbes* (Oxford: Clarendon Press, 1957); and C. B. Macpherson, *Political Theory of Possessive Individualism: Hobbes to Locke* (Oxford: Clarendon Press, 1962).

More recent theoretical analyses of contract are Margaret Macdonald, "The Language of Political Theory," in *Logic and Language,* edited by Antony Flew, first series (Oxford, 1956), and the essays in the collection *Philosophy, Politics and Society,* edited by Peter Laslett and W. S. Runciman (Oxford: Blackwell, 1957–), especially those by John Rawls.

Peter Laslett (1967)

SOCIAL CONTRACT [ADDENDUM]

Contemporary social contract theory is practically identified with the work of John Rawls (1921–2002). In his best known book, *A Theory of Justice*, Rawls attempts to generalize and carry to a higher level of abstraction the social contract theory of Locke, Rousseau, and Kant. In Rawls's version of social contract theory, people are to select the principles of justice they are to live by in imagined ignorance of whether natural or social contingencies have worked in their favor. His theory requires that we should choose as though we were standing behind an imaginary "veil of ignorance" with respect to most particular facts about ourselves, anything that would bias our choice or stand in the way of unanimous agreement. Rawls calls this choice situation "the original position" because it is the position we should start from when determining what principles of justice we should live by. Rawls explicitly argues that the principles of justice that would be selected are significantly different from the classical or average principle of utility.

Almost immediately, there was a utilitarian challenge to Rawls's theory led by R.M. Hare (2003) and Richard Brandt (1972), which maintained that the theory had the same practical consequences as utilitarianism. Soon after, there was a libertarian challenge led by Robert Nozick (1974), which claimed that Rawls's theory conflicted with an ideal of liberty, and later a communitarian challenge led by Michael Sandel (1982) and Michael Walzer (1983) contended that the theory ignored the situatedness of human beings, along with an Aristotelian challenge led by Alistair MacIntyre (1981) which objected to Rawls's theory for denying the priority of the good.

There was also a feminist challenged led by Susan Okin (1989), who, among others, maintained that Rawls's theory was biased against women, and a multicultural challenge led by a diverse array of Western and non-Western philosophers who maintained that the theory was biased against non-Western cultures. Since Rawls was reluctant to respond directly to his critics, these challenges created opportunities for others to step in and respond to them or to suggest ways in which Rawls's work needed to be modified to address these criticisms.

There was also the important question of the practical implications of Rawls's work for how we should live our lives individually and collectively. Rawls had always claimed to be developing primarily an ideal moral theory. *A Theory of Justice* only touched briefly on nonideal theory to provide an account of civil disobedience. But the farther removed one's society is from ideally just institutions, the greater is the need to spell out the practical requirements of justice for one's time, lest one stand accused of legitimating existing unjust institutions and practices. By deciding to focus his work on ideal moral theory, Rawls created opportunities for others either to

work out the practical implications of views developed in opposition or as a corrective to Rawls's view for the non-ideal world in which we live.

Rawls's second book, *Political Liberalism*, was written to correct a fundamental problem that Rawls perceived in *A Theory of Justice*. Rawls believed that his earlier book assumed a relatively complete Kantian conception of the good. In *Political Liberalism*, Rawls tries to ground his same theory of justice on a more minimal foundation—an overlapping consensus of reasonable comprehensive conceptions of the good. According to Rawls, citizens are to conduct their fundamental discussions within a framework of a conception of justice that everyone, irrespective of one's particular comprehensive conceptions of the good, could be reasonably expected to endorse. An important implication of Rawls's view is that religious considerations are generally excluded from public debate over fundamental issues in society. This feature of Rawls's view has engendered considerable debate, not only among philosophers, but also among theologians, political scientists, and lawyers, but it has not had any discernible effect on public policy, at least in the United States, where religious considerations continue to have an impact on public policy beyond anything that could be justified by a reasonable overlapping consensus.

Rawls's third major book, *The Law of Peoples*, attempts to extend his theory of justice to the international realm. Rejecting any straightforward application of his principles of justice to the international realm, Rawls favors more minimal obligations to other peoples. According to Rawls, there is virtually "no society anywhere in the world … with resources so scarce that it could not, were it reasonably organized and governed, become well-ordered." Rawls also allows for exceptions to international principles of justice, specifically a requirement of noncombatant immunity, in order to attain "some substantial good." At the same time he disallows any comparable exceptions to intersocietal principles of justice. Here again, Rawls's views have given rise to a wide-ranging discussion over possible exceptions to principles of justice, which has become even more important given the connection that exists between terrorisim and international justice.

Bibliography

Brandt, Richard. "Utilitarianism and the Rules of War." *Philosophy and Public Affairs*, 1 (1972) 145–165.

Hare, R. M. "Justice and Equality." In *Justice: Alternative Political Perspectives*, edited by James P. Sterba. 4th ed. Belmont: Wadsworth, 2003.

MacIntyre, Alasdair. *After Virtue*. Notre Dame: University of Notre Dame Press, 1981.

Nozick, Robert *Anarchy, State and Utopia*. New York: Basic Books, 1974.

Okin, Susan Moller. *Justice, Gender, and the Family*. New York: Basic Books, 1989.

Rawls, John. *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, 2001.

Rawls, John. *The Law of Peoples*. Cambridge. MA: Harvard University Press, 1999.

Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Sandel, Michael. *Liberalism and the Limits of Justice*. New York: Cambridge University Press, 1982.

Sterba, James P. *How to Make People Just*. Lanham: Rowman & Littlefield, 1988.

Walzer, Michael. *Spheres of Justice*. New York: Basic Books, 1983.

James P. Sterba (2005)

SOCIAL EPISTEMOLOGY

Since the early 1980s, social epistemology has become an important field in Anglo-American philosophy. It encompasses a wide variety of approaches, all of which regard the investigation of social aspects of inquiry to be relevant to discussions of justification and knowledge. The approaches range from the conservative acknowledgment that individual thinkers are aided by others in their pursuits of truth to the radical view that both the goals of inquiry and the manner in which those goals are attained are profoundly social.

Individualistic rather than social epistemologies have dominated philosophical discourse since at least the time of Descartes. The writings of Mill, Peirce, Marx, Dewey, and Wittgenstein, which began to develop social epistemologies, are among a few exceptions to individualistic approaches. They had little effect on epistemological work at the time they were published. Even the move to naturalism, taken by many epistemologists after W. V. Quine's polemics in its favor, persisted—quite unnecessarily—in individualistic assumptions about the nature of knowledge and justification. Quine argued in "Epistemology Naturalized" that epistemologists should attend to actual, rather than ideal, conditions of production of knowledge but he concluded that "epistemology … falls into place as a chapter of psychology," ignoring the sociology of knowledge altogether.

Movements outside of epistemology motivated and cleared the way for social epistemology. First and most

important, the proliferation of interdisciplinary research on social aspects of scientific change following the publication of Kuhn's *The Structure of Scientific Revolutions* pressured naturalistic epistemologists to take sociology of knowledge seriously. In particular, the skeptical and relativistic conclusions of sociologists and anthropologists of science—among them Barry Barnes, David Bloor, Steven Shapin, Simon Schaffer, Bruno Latour, Steve Woolgar, Harry Collins, Karin Knorr-Cetina, and Andrew Pickering—moved naturalistic epistemologists of science—including Ronald Giere, Larry Laudan, Philip Kitcher, and Paul Thagard—to take social accounts of scientific change seriously yet to draw their own epistemic conclusions. Second, influential work during the late 1970s in the philosophy of language and philosophy of mind—core fields of philosophy—by Hilary Putnam, Tyler Burge, and others eschewed individualism and began producing social accounts. A more general openness to social approaches in philosophy followed.

RANGE OF SOCIAL EPISTEMOLOGIES

Social epistemologies vary along several dimensions. First, they may emphasize either the *procedures* or the *goals* of inquiry. Whether the emphasis is on procedures or goals, the range here is as large as the range in epistemology as a whole: from consensus practices to critical engagement to truth to pragmatic success to socially constituted goals. Second, attempts to follow the procedures or attain the goals are evaluated for different units of inquiry. Some social epistemologists evaluate the attempts of individual human beings, assessing the influence of social processes on individual reasoning and decision making. Others evaluate the aggregate efforts of groups of people who may work together or separately.

Social epistemologies also tend to investigate particular domains and/or to work at particular levels of generality. Many (for example, Giere, David Hull, Kitcher, Helen Longino, Miriam Solomon, and Thagard) are social epistemologists of science rather than of ordinary knowledge or some other area of specialized knowledge. Feminist epistemologists (for example, Donna Haraway, Lynn Hankinson Nelson, and Naomi Scheman) look at the gender-relatedness of methodologies or assumptions in several fields, not only those explicitly dealing with sex or sex roles. Alvin Goldman (1992, 1999, 2002) works in the widest range of domains—from science to law to education to politics—and moves from the most general considerations of epistemics (in which he argues that truth is the ultimate epistemic goal) to the most concrete practical considerations (in which, for example, he argues that the common-law system is veritistically inferior to the Continental civil law system). Many social epistemologists work primarily at the general (abstract) level in their studies of areas such as testimony (Coady 1992), trust (Hardwig 1991), and knowledge (Kusch 2002).

Two journals are devoted to publishing material in social epistemology, *Social Epistemology* (1986–) and *Episteme* (2004–); many other journals publish special issues and individual articles in the area.

PROCEDURES OR GOALS OF INQUIRY

Longino's normative approach is to evaluate the procedures of a knowledge community. Her "critical contextual empiricism" (2002) evaluates four features of the knowledge community: the "tempered" equality of intellectual authority (equality moderated by deference for expertise), presence of forums for criticism, some shared norms (including empirical success in a scientific community), and responsiveness to criticism. Normative judgments will be of epistemic communities rather than of individuals and will be positive for communities following the appropriate procedures, irrespective of outcome.

Goldman, Kitcher, and Hilary Kornblith all take truth (or significant truth) to be the central goal of all kinds of inquiry. They assess various social processes and practices for their conduciveness to truth attainment. For example, Goldman (1992) shows that in some situations, such as some legal contexts, groups reach the truth more reliably when some true information is deliberately withheld from them—for example, misleading prejudicial information. So Goldman concludes that social epistemologists need to think about communication control, for paternalistic epistemic reasons. Goldman (1992, 1999) and Kitcher (1993) explore the consequences of intellectual rivalry and credit seeking in science. They both conclude that rivalry and credit seeking can lead scientists to distribute their cognitive effort well over the available research approaches, coming to a veritistic conclusion more quickly than they otherwise would. Kornblith (in Schmitt 1994) argues that the widespread practice of deference to experts may be reliable in one social setting and unreliable in another, depending on the institutions through which a society confers the title of "expert."

Some hold that, although truth is the ultimate epistemic goal, it is mediated by coherence of belief. They examine social processes for their conduciveness to coherence. For example, Keith Lehrer (1990) argues that individual reasoning yields more coherent belief if it makes use of all the information residing in a commu-

nity; Thagard (1993) argues that delays in the transmission of information across a community can be conducive to a good distribution of cognitive labor and thereby eventually to maximal explanatory coherence and truth.

Although most social epistemologists who employ normative goals regard truth as the most important epistemic goal, there is a range of other, less traditional, positions. Giere (1988), for example, claims that the goal of scientific inquiry is theories that model the world rather than directly correspond to it and that social practices such as credit seeking should be assessed for their conduciveness to producing good models. Solomon (2001) argues that scientific theories aim for empirical success. Steve Fuller (2002) writes of a range of epistemic goals espoused by scientific communities and argues that those goals should themselves be debated by scientists.

The most radical position on epistemic goals is one that claims that our social epistemic practices *construct* truths rather than discover them and, furthermore, negotiate the goals of inquiry rather than set them in some nonarbitrary manner. Work in the "strong program" in sociology of science during the 1970s and 1980s—notably by Barnes and Bloor, Latour and Woolgar, Shapin and Schaffer, Latour and Woolgar, and Collins and Trevor Pinch—was frequently guided by such social constructivism. (Recent work in the sociology of science is usually more philosophically sophisticated: See, for example, Shapin [1994].) Most contemporary social epistemologists in the Anglo-American philosophical tradition are motivated by their disagreement with the social constructivist tradition, and they argue for the less radical positions just described.

THE DISTRIBUTION OF COGNITIVE LABOR

The distribution of cognitive labor is a common theme in social epistemology and is a link between social epistemology and evolutionary epistemology. It is wasteful to duplicate the efforts of others, beyond the minimum required to check robustness of results. It is most efficient to have different individuals or research groups pursue different avenues of inquiry, especially when, as is usually the case, there is more than one promising direction to follow. Hull (1988), following the founder of evolutionary epistemology, Donald Campbell, was one of the first to apply this idea in the social epistemology of science, where he argued that new theories are like new organisms—produced by random variation on past theories—where only the fittest survive. And there is no way of knowing in advance which theory will be the fittest.

Others have also given accounts of how cognitive labor is distributed, although they have not emphasized the evolutionary analogy. Kitcher (1993) and Goldman (1992, 1999) have argued that the desire for credit leads to an effective division of cognitive labor; Thagard (1993) has argued that the same result is achieved by delays in dissemination of information; Giere (1988) thinks that interests and variation in cognitive resources distribute research effort; Solomon (2001) has argued that cognitive biases such as salience, availability, and representativeness can result in effective distribution of belief and thereby of research effort. (Not all these stories are, of course, true; some combination of them may be.) For all of the aforementioned thinkers, it is the *distribution* of cognitive labor across a community that is epistemically valuable rather than the decisions of any particular individual.

Cognitive labor can be divided not only for discovery and development of new ideas but also for storage of facts, theories, and techniques that are widely accepted. Just as books contain information that no individual could retain, information is also stored in communities in ways that are accessible to most or all members of that community but could not be duplicated within each head. One important way in which this is brought about is when people with expertise on different subjects—or with different experiences or techniques—increase the knowledge within a community. Knowledge and expertise is thus socially distributed. Edwin Hutchins's account of navigation (1995), in which skills and knowledge are distributed across the officers and enlisted men on board a naval vessel, is an example of this process.

A final way in which cognitive labor can be distributed is for the process of coming to consensus. In traditional philosophies of science, consensus is presented as the outcome of the identical decision of each member of a scientific community: A good consensus is the result of each scientist choosing the best theory through the same process, and a bad consensus is the result of each scientist choosing the wrong theory through the same inappropriate process. Of course, this is just the simplest model of group consensus formation, and it presumes the same starting point, the same endpoint, and the same processes of change. The only time that the members of the group may differ is during the period of dissent, when, as Hull (1998, p. 521) would say, a thousand theories may bloom. Giere, Hull, Kitcher, and others would also say that, when coming to consensus, each scientist picks the same theory for the same overriding good reasons. Other accounts of

consensus formation in which cognitive labor is distributed include that of Hussein Sarkar (1983), who finds that different scientists may select the same theory for different good reasons, and Solomon (2001), who finds that, although individual scientists may make biased and idiosyncratic decisions, there is a social perspective from which to evaluate the overall normativity of the decisions.

THE UNITS OF INQUIRY

Who knows? And who is justified in his or her knowledge? Nelson (1990) argues provocatively that only societies can really know. Some social epistemologists consider the outcomes of social epistemic processes for individuals and some for communities. The most conservative social epistemologies look only at the effects of social processes on individual reasoning and knowledge. For example, Kornblith (in Schmitt 1994) looks at those circumstances under which one scientist can judge that it is reasonable to rely on the expertise of another scientist. Coady's work on the role of testimony (1992) argues that individuals are typically justified in relying on the word of others. The claim is that individual human beings reason better when placed in favorable epistemic social situations. Epistemic terms such as "knows" and "is justified" are applied to individual human beings.

More radically, social groups can be understood as having emergent epistemic qualities that are due to something other than the epistemic properties of their members. Gilbert (1989) argues that group knowledge need have no coincidences with the knowledge ascribed to individual members of the group. Longino (1990, 2002) presents four conditions for objective knowledge that are satisfied by (some) knowledge societies rather than by individuals: tempered equality of intellectual authority, forums for criticism, responsiveness to criticism, and some shared values of inquiry. Nelson (1990) argues that communities set the standards of evidence and are the primary knowers. Kusch's "communitarian epistemology" (2002) argues for a similar conclusion through a performative analysis of testimony. Goldman (1999) shows that some kinds of social organization (for example, that of the American justice system) lead to poorer results than other kinds (for example, the Continental justice system). Schmitt (1994) argues that group justificatory processes can achieve, through interactions, more than the sum of individual justifications. Solomon (2001) shows that differently organized scientific communities make better and worse scientific decisions.

CONCLUSION

It is not surprising to find that the wide variety of social epistemologiesis connected to work in other disciplines. Economics, artificial intelligence (especially distributed computation), race and gender studies, sociology of science, anthropology, and European philosophical traditions (for example, Foucault and Habermas) are frequently cited, either for the data or for the methodologies that they supply.

When epistemologies are deeply social, recommendations for inquiry will often be applicable to communities or institutions rather than to individuals. Social epistemologists, especially those who are both naturalistic and applied, have begun to spell out these recommendations. The traditional focus on individual epistemic responsibility is being transformed by the addition of new, socially informed directions of inquiry.

See also Descartes, René; Dewey, John; Epistemology, History of; Foucault, Michel; Goldman, Alvin; Habermas, Jürgen; Kuhn, Thomas; Marx, Karl; Mill, John Stuart; Peirce, Charles Sanders; Putnam, Hilary; Quine, Willard Van Orman; Subjectivist Epistemology; Wittgenstein, Ludwig Josef Johann.

Bibliography

Coady, C. A. J. *Testimony: A Philosophical Study*. Oxford: Oxford University Press, 1992

Fuller, Steve. *Social Epistemology*. 2nd ed. Bloomington: Indiana University Press, 2002.

Giere, Ronald. *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press, 1988.

Gilbert, Margaret. *On Social Facts*. London: Routledge, 1989.

Goldman, Alvin I. *Knowledge in a Social World*. New York: Oxford University Press, 1999.

Goldman, Alvin I. *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press, 1992.

Goldman, Alvin I. *Pathways to Knowledge*. New York: Oxford University Press, 2002.

Haraway, Donna. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge, 1991.

Hardwig, John. "The Role of Trust in Knowledge." *Journal of Philosophy* 88 (1991): 693–704.

Hull, David. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press, 1988.

Hutchins, Edwin. *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.

Kitcher, P. *The Advancement of Science*. New York: Oxford University Press, 1993.

Kusch, Martin. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press, 2002.

Lehrer, Keith, *Theory of Knowledge*. Boulder, CO: Westview Press, 1990.

Longino, Helen E. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press, 2002.

Longino, Helen E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Nelson, Lynn Hankinson. *Who Knows: From Quine to a Feminist Empiricism*. Philadelphia, PA: Temple University Press, 1990.

Sarkar, Hussein. *A Theory of Method*. Berkeley and Los Angeles: University of California Press, 1983.

Scheman, Naomi. *Engenderings: Constructions of Knowledge, Authority, and Privilege*. New York: Routledge, 1993.

Schmitt, Fred, ed. *Socializing Epistemology*. Lanham, MD: Rowman and Littlefield, 1994.

Shapin, Steven. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago and London: The University of Chicago Press, 1994.

Solomon, Miriam. *Social Empiricism*. Cambridge, MA: MIT Press, 2001.

Thagard, Paul. "Collaborative Knowledge." *Nous* 31 (1997): 242–261.

Thagard, Paul. "Societies of Minds: Science as Distributed Computing." *Studies in the History and Philosophy of Science* (1993): 49–67.

Miriam Solomon (1996, 2005)

SOCIALISM

This entry is concerned with "socialism" from the time at which, so far as anyone knows, the word was first used in print to describe a view of what human society should be like. This was in 1827, in the English *Co-operative Magazine,* a periodical aimed at expounding and furthering the views of Robert Owen of New Lanark, generally regarded as the father and founder of the cooperative movement. (Owenite cooperation, incidentally, was an institution different from, and far more idealistic than, the distributive stores which in the Victorian age took over the name.) Some historians have traced the ancestry of socialism much further back: For example, to primitive communist societies, to the Jesuits of Paraguay, to the ideal communities described by Thomas More and others, to the Diggers of Cromwell's army, and even to Plato's *Republic.* Although there are elements of socialism to be found in all these, particularly in More's *Utopia,* the scope of this article is limited to socialism in modern times and to the sense in which the word is normally used, omitting both distant possible origins and, of course, bastard movements such as the National Socialism (Nazism) of twentieth-century Germany and Austria which, save for the bare fact that they enforced central control of social policy, had nothing of socialism in them.

ORIGIN OF SOCIALISM

The seedbed of socialism, as of so much else in modern thought, was the French Revolution and the revolutionary French thinkers who preceded it—Voltaire, Jean-Jacques Rousseau, and the Encyclopedists. Rousseau was no socialist, but from his cornucopia of seminal though sometimes unclear and inconsistent thought socialists drew the ideas of people born free but everywhere in chains, of a "general will" making for perfection in society, of the importance of education, and a host of others. From the Encyclopedists they learned to question all institutions in the light of reason and justice, and even from "Gracchus" Babeuf to demand equality for the downtrodden and to seek it by means of dedicated conspirators. Owen himself was no revolutionary; insofar as his ideas can be traced to anyone but himself, they probably came from early reading of the William Godwin who wrote *Political Justice*; Owen envisaged a society consisting of small, self-governing, cooperating communities, established by the free and rational consent of all, of whatever class or station. Originally, the word *socialism* appears to have laid particular emphasis on communal cooperation in contrast to the more-or-less liberalism that was coming to be the creed of the industrial revolution—hence Owen's rather contemptuous dismissal of Jeremy Bentham and the utilitarians. The idea of socialism came rapidly to fit the aspirations of the working classes and their radical champions not only in its country of origin but far beyond it.

SOCIALIST TENETS

Since its beginnings in the early 1800s, a period that has seen vast changes not merely in the industrial and political organization of society but also in people's minds, their modes of thought, and their interpretation both of themselves and of what they have seen around them, "socialism" has naturally borne many meanings, and dozens of views have been held and expressed about the form of society that socialists hope to see and about the means by which it should be attained and secured. Long before Karl Marx and Friedrich Engels introduced the great schism between what they called utopian and scientific socialism, there were wide differences of opinion; and the differences are no less wide today. George Bernard Shaw, for example, in *The Intelligent Woman's Guide to Socialism and Capitalism,* laid down absolute equality of money incomes as a sine qua non—a dictum

accepted by few of his fellow socialists, and not by Shaw himself in any practical sense. There are many other definitions that could be quoted. Nevertheless, the word is certainly not meaningless. It describes a living thing that grows and changes as it lives; and it is possible to discern certain beliefs that are fundamental to all who can be called socialists, as well as to note the divergences in what may be called secondary beliefs and to relate these, in part at least, to the conditions of the time.

CRITIQUE OF EXISTING SOCIETY. The first of the fundamental beliefs of socialists is that the existing system of society and its institutions should be condemned as unjust, as morally unsound. The institutions that are thus condemned vary from time to time and from place to place according to circumstances, the greatest stress being laid sometimes on landlordism, sometimes on factory industry, on the churches, the law, or the political government, or a combination of these (as William Cobbett, in an earlier century, denounced "The Thing"), depending on what seems to be the most potent engine or engines of oppression. This condemnation may be associated with the values of revealed religion, as in the case of the various forms of Christian socialism, or may positively repudiate those values, as Marx did; in either case the emphasis is on injustice. Pierre-Joseph Proudhon's dictum, "Property is theft," expresses this condemnation most concisely.

Many socialist movements, such as the Saint-Simonians in the 1830s and the Fabians half a century later, attacked the existing system for its economic and social inefficiency as well; but this criticism was less fundamental. Socialists such as François Marie Charles Fourier in France and William Morris in England laid much more stress on freedom, happiness, and beauty than on material wealth. Even the economists among them, however, long asserted that granted decent (that is, socialist) distribution of the product of industry and agriculture, there would easily be "enough to go round" and to provide everyone with a standard of living recognized to be reasonable. By the mid-twentieth century the enormous multiplication of potential demand, coupled with realization of the existence of hundreds of millions living far below European standards of life, had referred that type of prophecy to the far-distant future.

A NEW AND BETTER SOCIETY. The second fundamental of socialism is the belief that there can be created a different form of society with different institutions, based on moral values, which will tend to improve humankind instead of, as now, to corrupt it. Since it is living men who are to create the new institutions—men who must, therefore, recognize and follow the appeal of moral value—this belief is in effect an assertion of the perfectibility, or at least near-perfectibility, of man. It was most dogmatically stated by Owen, in books such as *A New View of Society*; and the history of socialism shows that it can survive innumerable disappointments. It is not the same as a belief in "progress," which has been held by many who were not socialists; it is more like *Magna est veritas et praevalebit* ("The truth is great and will prevail")—truth being here equated with justice.

Does justice, in social institutions, imply equality? Does it also imply democracy? For socialists, the answer to both these questions has generally been positive but the answer has not been absolute. Equality of rights—yes; equality before the law—yes, again. We have already observed, however, that complete equality of income was not a universal socialist tenet; and from the very earliest days there were sharp differences among socialists on the relationship between work and income. On the dictum "From each according to his ability" they more or less agreed. But some added "to each according to his needs"; others countered with "to each according to his effort—or his product." This debate, in which sides were taken, on the whole, in accordance with the temperament and/or environment of the individual and in which many intermediate positions were adopted, remained unresolved throughout the history of socialism—not surprisingly, since the problem of controlling the level of incomes has defeated all except completely static societies. On the question of democracy, again, the great majority of socialists have been democrats in the ordinarily accepted sense of the word. But some rejected any formal democratic process in favor of a communal consensus resembling the Quaker "sense of the meeting" (or Rousseau's general will). Owen, in practice, was an autocratic egalitarian; and post-Marxist socialism has evolved a procedure known as democratic centralism, which bears little relation to what any pre-Marxist would recognize as democracy.

Deep differences arose early on the kind of institution which would be best suited for a world devoted to justice. There was one main difference at first: Some put their faith in small communities of neighbors, as far as possible self-sufficient, cooperating freely with other similar communities in such functions as exchange of goods, and relying to the minimum on any regional or central authority for such necessities as defense and the supply of credit; others looked rather to a development of science, technology, and large-scale industrial production and

banking to increase rapidly the supply of material goods and thereby the prosperity of a socialist economy through centralized planning techniques. Of these two schools—whose views have necessarily been greatly simplified for the compass of this article—the first, or "utopian," is best known from the writings of Owen, Fourier, and Proudhon, and the second, or "scientific," from those of the Comte de Saint-Simon and his followers. The first clearly derived from rural society: Owen's villages of cooperation and Fourier's phalansteries were based upon small-scale agriculture, with such industrial and craft production as could conveniently be carried on in villages or small communities. This was the kind of society envisaged, much later, in William Morris's *News from Nowhere*; and much later still, there were curious echoes of it in V. I. Lenin's dreams of cheap electricity transforming the life of the Russian peasantry and even in the Chinese "great leap forward," with a piece of factory in every backyard.

The weakness of this school is that its fear of size, of external authority, and of the apparatus of the state and of central government, whatever concessions it may in theory make to "natural necessities," such as the conduct of a national railway system, are liable to lead in practice as well as in theory to anarchism and the repudiation of any government at all—which in the modern world means chaos. The second school, that of large-scale production and planning, was, from the beginning, in harmony with the way the world was tending. Its dangers are today only too obvious, and the recurrent malaise of large-scale industry in times of prosperity, the demands for "shares in control," and the like, show the vacuum created by the nonfulfillment of the utopian ideals of a just society.

REVOLUTION. Whatever form of institution the several schools of socialism envisaged for the future, all agreed that what was required was a fundamental transformation of society amounting to revolution, a program of action to effect such a transformation, and a revolutionary will so to transform it existing in the members of present-day society. This is the third fundamental socialist assumption; how it is to be put into effect has been the subject of much division of opinion. As socialism was generally believed to have a strong rational basis, it was natural that all schools of socialists should set great store by education, persuasion, and propaganda; Owen, indeed, carried the trust in rationality so far that he could not believe that anyone, whatever his condition or his preconceived opinions, could fail to be converted by "Mr. Owen's powers of persuasion," if only Mr. Owen could employ them sufficiently often and at sufficient length. Others, less confident, sought to achieve their end by preaching to and working upon groups already conditioned by the circumstances of their working lives to accept the whole or a part of the socialist gospel—the most obvious of these being, of course, the trade unions and other organizations of the working class. In this spirit Marx looked upon the British trade unions that supported the International Working Men's Association (the "First International") as "a lever for the proletarian revolution." Strikes, threats of strikes, and other forms of what much later came to be known as "direct action," supplemented persuasion by inducing the ruling classes to make concessions which could not otherwise have been wrung from them.

The practicability, either of persuasion or of group action, depended very largely on the political conditions of time and place. And although there was a running argument between gradualists, who believed that revolutionary change could be brought about peacefully and piecemeal, and revolutionaries, who thought head-on collision between the holders of power and their victims was inevitable in the long run if not immediately, the difference was not as absolute as was often supposed. In Britain, after the defeat of Chartism had registered the end of insurrectionism in any form, after the press had been freed and the franchise widened, the organizations of the working class leaned to peaceful evolution far more than to violence—the "inevitability of gradualism" was an accepted belief long before Sidney Webb put it into words in the 1920s. In tsarist Russia, at the other extreme, a generally authoritarian government, operating a police state, appeared to bar the door to anything but physical revolution. There were many possible in-between positions; and the role of the convinced individual socialist varied similarly, from that of open persuader, adviser, and organizer, like Keir Hardie at the end of the nineteenth century, to that of secret conspirator, like Auguste Blanqui in France after 1848 and organizers of communist cells in the twentieth century.

INTERNATIONALISM. One other characteristic should briefly be mentioned. Socialism was initially a world philosophy, not concerning itself with race or nation, not advocating the brotherhood of man so much as assuming it. The opening of the *Communist Manifesto,* "Workers of the world, unite," crystallized this into words; the nationalism of Poles, Irish, Italians, Hungarians, was only an aspect of the struggle against corrupt institutions. Later, of course, nationalism grew so strong that it clashed, sometimes violently, with other fundamentals of socialism; nevertheless, the idea remained potent for genera-

tions, and it may still be suggested that socialist movements that have become exclusively nationalist have ceased to be socialist at all.

MARXIAN SOCIALISM

The *Communist Manifesto* marks a great divide between pre-Marxian and post-Marxian socialism. Marx and Engels dismissed all their predecessors as utopians and formulated a system of socialism that they claimed was "scientific." There is no room here to expound Marxist philosophy or Marxist economics; but it must be pointed out that neither "utopian" nor "scientific" is an accurate description. Marxist socialism accepted the fundamentals as set out above; it differed from most of its forerunners in that it did not, save in a few very vague allusions, seek to describe the new, uncorrupt institutions that would appear after the revolution; it assumed—and what could be more utopian?—that after the proletariat had conquered, it would make all anew and "the government of man be replaced by the administration of things."

"Scientific," in Marxist language, meant not so much acceptance of technology and large-scale production—although this was included—as the proving, by logical argument and study of history, of two quite simple propositions: First, that under the existing capitalist system, the proletariat, the laboring class, is systematically and continuously robbed of its just share of the fruits of production; second, that "changes in the modes of production and exchange," and not any other factor, such as "man's insight into eternal truth and justice," are leading inevitably to a reversal of the system that will remove the bourgeois capitalist class from the seats of power and replace it by the organs of the proletariat. This is the base on which the whole enormous superstructure of Marxism is founded; it is not science, but messianic prophecy. It is easy to understand, however, the compelling effect that this fundamentally simple appeal had to the downtrodden at various times and in various places. At the same time, Marx's powerful and penetrating analysis, which discredited a great deal of current economic and historical theory, profoundly attracted many of the best brains among those who were dissatisfied with the human results of the existing system, and the teaching of the Marxists that morality in action was relative to the needs of the time, even if slightly inconsistent with their denunciation, on grounds of injustice, of slavery and wage slavery, gave their followers both the inspiration of those who were fighting a continuing battle and the sanction to use any and every method that could advance their cause. Marx did not invent the conception of classes, but Marxists fought the class war.

The work of Marx and Engels has had as great and lasting an effect on the thinking of non-Marxists, particularly after the Russian Revolution, as has that of Sigmund Freud on non-Freudians. This entry cannot deal with the developments in socialist thought, Marxist or non-Marxist, in the post-Marxian era. These are of enormous importance for the study of history and present-day politics; but they are concerned principally with method and strategy. The fundamental tenets of socialism as a view of society have remained substantially unaltered, although the process of translating them has been far more lengthy and complicated than the nineteenth century ever foresaw.

See also Anarchism; Bentham, Jeremy; Communism; Encyclopédie; Engels, Friedrich; Fourier, François Marie Charles; Freud, Sigmund; Godwin, William; Justice; Lenin, Vladimir Il'ich; Marx, Karl; Marxist Philosophy; More, Thomas; Plato; Proudhon, Pierre-Joseph; Rousseau, Jean-Jacques; Saint-Simon, Claude-Henri de Rouvroy, Comte de; Utilitarianism; Voltaire, François-Marie Arouet de.

Bibliography

By far the most complete work on socialism is G. D. H. Cole, *History of Socialist Thought.* 7 vols. (London, 1953–1960), last volume published posthumously. Each volume except the last has a full bibliography for every chapter.

Earlier books are Max Beer, *History of British Socialism* (London: G. Bell, 1929); Édouard Dolléans, *Histoire du mouvement ouvrier,* 2 vols. (Paris: A. Colin, 1936–1939); Sir Alexander Gray, *The Socialist Tradition: Moses to Lenin* (London: Longmans, Green, 1946); Élie Halévy, *Histoire du socialisme européen* (Paris: Gallimard, 1948); Thomas Kirkup, *History of Socialism,* rev. ed. by E. R. Pease (London: A. and C. Black, 1913); H. W. Laidler, *Social Economic Movements* (New York: Thomas Y. Crowell, 1944), first issued as *History of Socialist Thought* (New York: Thomas Y. Crowell, 1927); and Paul Louis, *Histoire du socialisme en France,* rev. ed. (Paris, 1950).

For documents, see Raymond Postgate, *Revolution from 1789 to 1906,* reissued, with a new introduction (Gloucester, MA: P. Smith, 1969).

Margaret Cole (1967)

SOCIALISM [ADDENDUM]

Socialism has seen enormous changes since the above entry was written. Its cachet has gone up and down and, after an all-time low during the early 1990s, is now per-

haps going to go up again. The socialist ideal fell on hard times when "actually existing socialism" collapsed in the Soviet Union and its satellite states in 1989 and somewhat later in Yugoslavia. The headlong rush of China towards free-market development has further deepened the crisis of contemporary socialism. Only Cuba, North Korea, and perhaps Vietnam and Laos remain as "actually existing socialisms."

There have been similar upheavals in socialist theory. Most Western socialists, including most Marxists, while not being cold warriors, did not regard these "actually existing socialisms" as genuinely socialist but as statist noncapitalist societies that were authoritarian, nondemocratic and excessively bureaucratic regimes parading as paradigms of socialist societies. Instead of the dictatorship *of* the proletariat (what was supposed to be the mass but democratic governing of the working class by the working class in the interim before "the withering away of the state" and the attaining of a classless society), there were what anarchist socialists (most notably Mikhail Bakunin) called the dictatorship *over* the proletariat, namely the rule over the proletariat by a small elite calling themselves communists.

Among most Western socialist theoreticians something like the following view became prevalent: The Soviet Union was not even a flawed socialism but an authoritarian statist postcapitalist society that had betrayed many of the most fundamental beliefs of socialism. Others, including Noam Chomsky, denied that it even had a somewhat progressive "postcapitalist" character at all, but actually the Soviet Union became a form of state capitalism; this latter claim is disputable as it is not for contemporary China. But state capitalist or postcapitalist, it became an authoritarian cumbersomely bureaucratic regime that betrayed many of the ideals of socialism. Both sorts of socialist intellectuals sought to reinvigorate socialist though and to help create a way to reinvigorate socialist practice. For them, in a standard sense of the word *democratic*, the term "democratic socialism" was a pleonasm.

In light of these historical realities, what it is to be a socialist has become more ambiguous than it was at the high tide of Marxism. Andrew Levine has well used "socialism" to designate those political tendencies and movements that, since the beginning of the nineteenth century, sought to deepen what the most radical of the French revolutionists began. Like the liberals, their tamer confreres on the Left, socialists always have been steadfast in their dedication to "liberty, equality, and fraternity." But, like their revolutionary forebears—and unlike liberals—they have usually favored radical, structural transformations, at least in principle. This broad characterization allows us to regard the more radical social democrats (for example, Jürgen Habermas), some anarchists (for example, Noam Chomsky), and Orthodox Marxists (for example, Bertell Ollman) as all socialists. Whether there is a spectrum here or some fundamental cleavage is a much debated matter.

The more Orthodox Marxists would take it to be axiomatic that a socialist of any sort is someone who favors public or social ownership, and at least indirect control, of at least the principal means of production. In such a society there would be no one because of this public and shared ownership who *simply* has to sell his or her labor. Public ownership in different forms of socialism can mean different things. For some it has meant state ownership and for others various schemes of worker ownership and control. For some social democrats socialism has meant a mixed-economy containing small-scale private ownership of the means of production but with larger-scale ownership being firmly public. Others would move so far from traditional conceptions of socialism as to not identify socialism necessarily with a distinctive form of ownership at all but with radical democracy and a thoroughly egalitarian-solidaristic conception of justice.

Some orthodox socialists would not regard a mixed economy at all; nor would they regard as socialist normative conceptions of socialism that identify it with radical democracy and egalitarian-solidaristic justice. They would classify the latter as social democratic and not genuinely socialist at all. A socialist society, on this view, must be a society without capitalism (or at least on the way to abolishing it). It would be a society in which everyone is either a worker, a potential worker (children), a former worker or person incapable of work (such as the retired or disabled), or someone soon destined to become a worker in a social order in transition to a classless society. Many Marxists believe that such a development would have to be global to be sustained.

Others would respond that contemporary society has too many strata doing various kinds of work to make "worker" a very useful category or class analysis the trenchant critical took that Marxists took it to be. Others insist on the centrality of a class analysis while arguing that in contemporary society classes have become more ramified than in Marx's time (Wright 1989).

Although the foregoing characterizations of socialism are matters of definition, they are not simply that. Each vision of a socialist future bears different implica-

tions for social policy and for the society and the world as a whole. At the far end of the social democratic spectrum capitalism would remain in place, but with much of its power curbed (or so the plan goes); on the more robustly socialist end of the spectrum, capitalism would have to be replaced with socialism either by the ballot box or by some form of revolution. And where socialism is identified with the ownership and control of the means of production, it matters considerably whether public ownership takes the form of state ownership or direct workers' ownership and control or some combination of both.

Socialism, taken in the more robust sense, is commonly thought to be tyrannical or authoritarian. But that claim has little merit. Contemporary socialists in the West have, like liberals, a commitment to liberty and democratic procedures, as did Marx and Engels, although the latter two seemed to have an unrealistically simplistic view of the implementation of radical democracy and paid little attention to procedures or to constitutional matters of protecting human rights. They thought that, as the dust of the socialist revolution settled, society would evolve in an ever more democratic direction. But contemporary socialists do not think that. Moreover, Marx and Engels thought that the socialist revolution would start in advanced capitalist societies, but when it arose in Russia instead, Rosa Luxemburg argued perceptively that if it did not quickly spread to the wealthy capitalist West, it would be doomed. But socialism did not spread westward from the Soviet Union; it originated in a backward authoritarian state with little in the way of a democratic tradition and without much in the way of developed productive forces.

Marxist socialists of whatever stripe are historical materialists and anticipated that socialism would piggyback on developed capitalism. No socialist society can succeed, they claim, without highly developed forces of production and a democratic tradition. Where those are absent, a socialist revolution will sour or collapse. But where these conditions obtain, Marxists claim, there is no fear of a socialist society succumbing to authoritarianism.

Another issue for contemporary socialist thought is whether socialism can work efficiently in the absence of markets. Only market societies have had a successful track record of providing consumer goods and services swiftly to a large portion of the population. The response by many contemporary socialists has been to propose market socialism. Alec Nove, John Roemer, and David Schweikart have proposed carefully worked out diverse models about how this hybrid could work. Market socialists (as in reality contemporary capitalists do) work with both market and plan. They disdain the Soviet command model, which regards markets (except in very limited domains) as dysfunctional, and are cautious about central planning.

Market socialism has been resisted by some Orthodox Marxists (for example, Mandel and Ollman). They believe that any market socialism will reproduce the inequalities and instabilities of large-scale capitalism. But Roemer responds that if markets are used solely to guide allocation, there is no reason why market socialism will lead to a society addicted to consumerism. Indeed some socialists believe that it might even surpass concentrated capitalist enterprises in meeting people's needs. The problem for others is rather a worry about what appears at least to be its *political* impossibility. A major worry is whether such an alternative could ever gain a serious hearing in societies dominated by capitalist states and by large capitalist media conglomerates.

See also Bakunin, Mikhail Aleksandrovich; Chomsky, Noam; Civil Disobedience; Cosmopolitanism; Engels, Friedrich; Marx, Karl; Marxist Philosophy; Postcolonialism; Republicanism.

Bibliography

Bernstein, Eduard. *Evolutionary Socialism*. New York: Schocken Books, 1961.

Cohen, G. A. *Karl Marx's Theory of History: A Defense*. Oxford, U.K.: Clarendon Press. 1978.

Cunningham, Frank. *Theories of Democracy*. London, U.K.: Routledge, 2002.

Dolgoff, Sam, ed. and trans. *Bakunin on Anarchy*. London: George Allen and Unwin, 1971.

Gay, Peter. *The Dilemma of Democratic Socialism*. New York: Octagon Books, 1983

Habermas, Jürgen. "What Does Socialism Mean Today? The Revolutions of Recuperation and the Need for New Thinking." In *After the Fall: the Failure of Communism and Future of Socialism*, edited by Robin Blackburn, 23–46. London: Verso. 1991.

Howard, Dick, ed. *Selected Political Writings: Rosa Luxemburg*. New York: Monthly Review Press, 1971.

Kolakowski, Leszek. *Main Currents of Marxism*. 3 Vols. Translated by P. S. Falla. Oxford: Clarendon Press. 1978.

Levine, Andrew. *Arguing for Socialism: Theoretical Considerations*. Boston: Routledge and Kegan Paul, 1984.

Mandel, Ernest. "The Myth of Market Socialism" *New Left Review* 169 (1988): 108–120.

Nove, Alec. *The Economics of Feasible Socialism*. London: George Allen & Unwin, 1983.

Ollman, Bertell, ed. *Market Socialism: The Debate Among Socialists*. London: Routledge, 1998.

Pierson, Christopher. *Socialism after Communism*. University Park: The Pennsylvania State University Press, 1995.

Roemer, John. *A Future for Socialism*. Cambridge, MA: Harvard University Press. 1984.

Schweikart, David. *Against Capitalism*. Cambridge, U.K.: Cambridge University Press, 1993.

Trotsky, Leon. *The History of the Russian Revolution*. New York: Anchor Foundation, 1980.

Tucker, Robert C., ed. *The Lenin Anthology*. New York: W. W. Norton, 1975.

Wright, Erik Olin. *The Debate on Classes*. London: Verso, 1989.

Kai Nielsen (2005)

SOCIAL SCIENCE EXPLANATIONS OF RELIGION

See *Religion, Psychological Explanations of*

SOCIAL SCIENCES, PHILOSOPHY OF

See *Philosophy of Social Sciences*

SOCIETY

A group of perennial problems in social philosophy arises from the concept "society" itself and from its relation to the "individual." What is the ontological status of a society? When one speaks of it as having members, is that to recognize it as a whole with parts, or is the relation of some different kind? Or is this a case of what Alfred North Whitehead called the fallacy of misplaced concreteness?

SOCIAL ACTION AND SOCIAL RELATIONS

"Society" is used both abstractly and to refer to entities that can be particularized, identified, and distinguished from each other as social systems or organizations. The phrase "man in society" is an instance of the more abstract use, for it refers neither to some particular form of association nor to a particular collectivity in which individuals find themselves. It refers, rather, to the social dimension of human action—to a certain generalized type of human relationship. Purely spatial or physical relations between human beings, like contiguity, are not social; for social relations give to human actions a dimension possessed neither by the mere behavior of things nor, indeed, of animals.

Max Weber defined a social action as one which, "by virtue of the subjective meaning attached to it by the acting individual (or individuals), … takes account of the behavior of others and is thereby oriented in its course" (*Theory of Social and Economic Organization,* p. 88). That is to say, the agent understands his own action as having a particular point, which in turn depends on an understanding of what another individual or other individuals have done in the past (as, for instance, in an act of vengeance), are doing now, or are expected to do in the future (as, for instance, in a proposal of marriage). So, said Weber, the efforts of two cyclists to avoid hitting one another would have a social character, whereas the collision between them would not.

An action would not be social merely because it was the effect on an individual of the existence of a crowd as such. For instance, laughing less inhibitedly in a crowd than one would when alone would not be an action oriented to the fact of the existence of the crowd "on the level of meaning"; while the crowd may be one of the causes of the action, the point or meaning of the action does not presuppose some conception of, say, the crowd's purposes or the reasons for its presence. Nor would merely imitative behavior be social; one could learn to whistle by imitating a man, a bird, or a whistling kettle. Learning and performance need neither an understanding of what is imitated as an action nor an orientation toward expected future action of the model. Nevertheless, says Weber, if the action is imitated because it is "fashionable, or traditional, or exemplary, or lends social distinction … it is meaningfully oriented either to the behavior of the source of imitation or of third persons or of both" (pp. 112–114). Weber then goes on to define "social relationship." This would exist wherever, among a number of actors, there existed a probability that their actions would be social actions.

Weber's concept of the "meaning" of an action is rather obscure. It may be a meaning "imputed to the parties in a given concrete case," or it may be what the action means "on the average, or in a theoretically formulated pure type—it is never a normatively 'correct' or metaphysically 'true' meaning" (p. 118). This concept is connected with Weber's much criticized conception of empathic understanding (*Verstehen*). But this connection is not strictly necessary, for the meaning we give to the actions of others depends not so much on an attempted reconstruction of what is in their minds as on a knowledge of the norms and standards regulating their behav-

ior in a given context. Thus I know what a man is about when he presents a bank teller with a signed paper of a certain size, shape, and color, not because I can reconstruct his state of mind in imagination but because I can recognize the procedures for cashing checks.

Weber insists that it is the probability itself of a course of social action that constitutes the social relation, not any particular basis for the probability. Yet we can rely on situational responses (like the bank teller's, for instance) very largely because we expect them to conform to norms and procedures, by which such responses are deemed appropriate or otherwise. Assuming, as many sociologists would, that even war is a social relation, the acts of opposing commanders are mutually oriented by an understanding of the aims and practices of warfare and by the supposition that the other's actions will be appropriate, not only in terms of means and ends but also in consideration of whatever rules of war may be current. Thus we can move from the concept of social relations as frameworks for interaction to Talcott Parsons's conception of a social system constituted by differentiated statuses and roles.

SOCIETIES AS ORGANIZATIONS

The concept of "*a* society" implies a system of more or less settled statuses, to each of which correspond particular patterns of actions appropriate to a range of situations. By virtue of qualifying conditions a man enjoys a status; in virtue of that status he has a role to play. These concepts, however, are meaningful only in the context of rules or norms of conduct—a man's role is not simply what he habitually does (for this may be no more socially significant than a tic), nor even what he is expected to do, if an expectation is only what one might predict about his future conduct from a knowledge of his past. His role is what is expected of him, in the sense of what is required of him by some standard. The role of secretary to an association, for instance, requires that he read the minutes of the last meeting, because the rules of procedure assign this action to whosoever enjoys this status. Less formally, a father's role may be to provide the family with an income, and failure to do so will be regarded not merely as falsifying predictions but also as disappointing reasonable or legitimate expectations—reasonable, because grounded on an understanding of the norms constituting the structure of the family. Indeed, though what we knew of some particular father might give us good grounds for predicting that he would neglect his role, that would not mean that its requirements did not apply to him. Of course, when we speak of "the family" or "the modern state," we commonly have in mind ideal types or paradigms. There may be significant deviations from these in practice. Any particular family may have its own standards, deviant from the social norm, according to which the role of father does not include providing the family income.

Looked at in these terms, a society is an aggregate of interacting individuals whose relations are governed by role-conferring rules and practices which give their actions their characteristic significance. Thus, to demand money with menaces is one thing if done by a common blackmailer or footpad, another if done by a tax collector.

Nevertheless, the act of John Smith, tax collector, is still the act of John Smith, who acts also in different roles in other situations—as father, member of Rotary, and so forth. So one may take two views of a society. On the one hand, one may see it, as a biographer might, as an aggregate of life histories of its individual members, each, in the course of his life, acting in a variety of roles that explain (but only partially) what he does. Or one may adopt the sociological standpoint. A society is then a pattern of roles, and what President Brown does is less important than that it instantiates the role of president.

INDIVIDUALIST AND HOLISTIC ACCOUNTS

Are there any statements about societies, or what Émile Durkheim termed "social facts," that are not ultimately reducible to statements about individuals? According to an extreme individualist or nominalist, such as Thomas Hobbes, social wholes have no substantial reality; propositions attributing properties or actions to a collectivity can be reduced, without residue, to a series of propositions about the relations and actions of individuals: "A multitude of men are made one person, when they are by one man, or one person, represented.… and *unity,* can not otherwise be understood in multitude" (*Leviathan,* edited by Michael Oakeshott, Ch. 16, p. 107). Karl Popper's methodological individualism is as uncompromising. So-called social wholes, he declares, are theoretical constructs; "social phenomena, including collectives, should be analysed in terms of individuals and their actions and relations" (*Conjectures and Refutations,* p. 341).

There is no agreement, however, on whether such analysis is possible. Some philosophers, while admitting that every action is the action of an individual, nevertheless deny that "statements which contain societal terms" can be reduced "to a conjunction of statements which only include terms referring to the thoughts and actions

of specific individuals" (Maurice Mandelbaum, "Societal Facts," p. 482). While the "societal fact" of cashing a check can be expressed in terms of what individuals do, nevertheless the description will always contain such societal terms as *bank* and *money,* which cannot themselves be translated without remainder into wholly individual terms. Furthermore, such societal facts, it is said, interact with individual behavior; a banking system can have an effect on a concrete individual. For it is clearly true that for every individual, the institutions and mores of his society present themselves as independent and external facts, just as much as his physical environment does. And if that is true for every individual, it is true for the totality of individuals composing the society. That is not to say that a totality is a thing independent of individuals or that it has a group mind; it is only to say that for any participant or for any observer of an individual's actions, it makes sense to talk of him confronting and confronted by independent social facts (Ernest Gellner elaborates this point). Moreover, the principle that social action can ultimately be explained by referring to the dispositions of individuals to behave in certain ways in given circumstances overlooks the possibility that these dispositions may themselves depend on social facts.

The view that social facts are not reducible to individual facts is commonly called holism. In its more extreme forms it relies heavily on biological organic analogies. An organism, it is said, is prior to its constituent parts in the sense that any understanding of their nature and function presupposes an understanding of the whole organism. The whole organism is more than the mere sum of its parts, since no account in terms of the parts considered separately could add up to some of the things that could be said about the whole. (The same might be said, however, of some of the properties of a triangle that arise from the three sides considered in relation to one another.) Just as the liver is a more significant object considered as an organ of a working body than as a detached piece of tissue, so the acts of individuals are significant or intelligible only when considered as the acts of role-bearers or as manifesting characteristics of their social or cultural environment. So drinking wine has a different range of social meaning in England from the one it has in France. The thought-experiment of the social contract theorists, who put man into an asocial state of nature the better to understand his real purposes in society, was radically misconceived, precisely because it abstracted man from the very context in which alone he would be a man but still attributed human properties to him.

According to the Hegelians (Bernard Bosanquet, for example), so far are we from being able to reduce social facts to individual facts that it is the individual himself who must be explained as an expression of the concrete social universal—an idea manifesting itself organically in its differentiated parts, as the idea of an oak tree is differentially but organically manifest in its leaves, bark, trunk, and so forth, all in a sense different from one another yet all linked by the idea of the oak and collectively differentiated thereby from the corresponding parts of an elm. "Man" is an abstraction—we are men as we are Germans, Englishmen, Frenchmen; that is, we instantiate the spirit of our own society.

Holistic organicism of this kind has laid great stress on history. Social wholes, it is said (by Friedrich Karl von Savigny, for instance), are not like mechanical wholes. Mechanical wholes can be understood by reducing them to their smallest constituent parts that conform in their behavior to general laws from which the varying behavior of the aggregates can be deduced. A social whole, on the contrary, is sui generis, to be understood not by analysis but by studying it *as* a developing whole. Consequently, there can be no general theory of social action, and history is the only legitimate mode of sociological inquiry.

According to Popper, these arguments are totally misconceived. There is simply no way of studying wholes as wholes; any attempt at understanding implies abstracting from a particular configuration of properties and circumstances those that seem significant for the particular study and relating them to general laws and hypotheses that are valid for all cases, irrespective of time, in which the stated initial conditions are satisfied. A law of development could be a statement about the general tendencies of certain types of society, given certain initial conditions; but it is a misunderstanding of the nature of both scientific and historical inquiry to propose a study of a society as a whole, partly because a social whole is a theoretical construct and partly because to attribute to it its own peculiar law of growth, in some sense true regardless of, or despite, any initial conditions whatsoever, is to make any explanatory statement about its behavior impossible.

COMMUNITY AND ASSOCIATION

The individualist account of social action is most persuasive when the form of social organization under consideration is a joint-stock corporation or a trade association. There is little temptation to attribute group personalities to such bodies, except in a strictly legal sense, and therefore little resistance to treating them as nothing but pro-

cedural forms. Their members and officials are clearly identified individuals with limited common interests. These interests explain their interaction, without suggesting that the association is anything more than a means for promoting them. Moreover, such interests remain intelligible even abstracted from the context of the society.

Ferdinand Tönnies distinguished this type of organization, which he called a *Gesellschaft* (association), from its polar opposite, the *Gemeinschaft* (community). Paradigms of the latter type are the family, the village, the tribe, and the nation. These are much less formally organized than a joint-stock company. They have no clearly defined, limited aim; qualifications for membership may be poorly defined, depending very largely on subjective criteria. Yet individuals do not deliberately join such bodies—more usually they are born into them or acquire membership by residence. At the same time, membership in such a community may mean much more to the individual. So far from his using the organization as a means for the pursuit of personal interest, privately conceived, what he conceives to be his interest may depend very much on the influence of the collectivity upon him. He may feel bound to it by ties and responsibilities not of his own choosing which nevertheless demand his respect. Moreover, such communities appear to have a lifespan greater than that of any generation of individual members, which cannot be explained, as might that of a corporation, by the continuities of constitutional procedures. It is, rather, that from generation to generation there passes an attachment to a common set of symbols and a common history, a participation in what Durkheim termed "collective representations" in a collective consciousness—a common culture, in short—which enables members to identify one another where other criteria are uncertain, which gives the society its cohesion, and which provides the standards by which its members' actions are regulated and assessed.

A FUNCTIONALLY INCLUSIVE COLLECTIVITY

"Boundary maintenance," to use Talcott Parsons's term, is a necessity for every society. To possess an identity, a society must furnish criteria whereby its members can identify one another, since their actions and attitudes toward one another will be different from those toward outsiders. But Parsons also conceives of boundary maintenance by social subsystems within a broader system. Thus he defines "a society" as a collectivity "which is the primary bearer of a distinctive institutionalized culture and which cannot be said to be a differentiated subsystem of a higher-order collectivity oriented to most of the functional exigencies of a social system" (*Theories of Society,* Vol. I, p. 44). Such a collectivity is organized by political, economic, familial, and similar subsystems. Parsons distinguishes polity and society, but he asserts that "the boundaries of a society tend to coincide with the territorial jurisdiction of the highest-order units of political organization" (p. 46). For, in Parsons's view, a society's existence depends so crucially on commitment to common values and on the maintenance of order between its individual and collective components that the political boundary tends to settle automatically the limits of the society.

The relation between state and society presented no problems for the Greeks. Political, religious, cultural, and athletic activities were largely undifferentiated and occurred within the single organizational structure of the polis. The first serious problems in this respect emerged with the Christian dichotomies between God and Caesar, church and state, the *Civitas Dei* and the *Civitas Terrena.* The medieval view was that, ideally, there was one universal community of humankind with two modes of organization, or "subsystems," church and empire. Reality never corresponded very closely to this ideal. It became irretrievably divorced from it with the rise of the nation-state and the Reformation. Since then, when people have talked of the society to which they belong, they have thought primarily (like Parsons) of the social order contained within the boundaries of a state and sustained by its organized power.

Nevertheless, liberal thinkers have striven hard to maintain the conceptual distinction between state, or polity, and society. One reason has been to resist the claim that the state could be the only focus of loyalty, competent by virtue of an overriding authority to lay down the terms on which other associations might function. On the other hand, there has emerged a new totalitarianism which identifies state and society. Every form of economic, religious, artistic, or scientific activity thereby acquires a political dimension, promoting or impeding the public good as embodied in state policy. G. W. F. Hegel provided a metaphysical justification for this kind of doctrine when he distinguished between, on one hand, civil society—a level of social organization including the market economy and the forces of civil order—and, on the other, the transcendent state—"the realized ethical idea or ethical spirit," "the true meaning and ground" of lower forms of social organization like the family and civil society (*Philosophy of Right,* Secs. 257, 256). By contrast, not only do liberals insist on the subordination of

the state to society; they have also tended, according to Sheldon S. Wolin, to depreciate the political and to attach increasingly to other social subsystems, like the business corporation or the voluntary association, concepts like statesmanship, authority, and legitimacy, which have been considered hitherto characteristic of the state. Meanwhile, Wolin argues, the concept of an organization directed to the most general interests of the community tends to get lost, to be replaced by a model of conflicting pressure groups operating within a very nebulously defined arena. If Parsons is right, our notion of a society as the most inclusive framework of social interaction depends on the political not only for its boundary maintenance but also for its very identity. There may be a danger that in pressing the antitotalitarian, pluralistic account so far that it dissolves the state, it will lose thereby its capacity to define the society.

See also Bosanquet, Bernard; Durkheim, Émile; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Holism and Individualism in History and Social Science; Popper, Karl Raimund; Savigny, Friedrich Karl von; Social Contract; Sovereignty; Weber, Max; Whitehead, Alfred North.

Bibliography

Bendix, Reinhard. *Max Weber—An Intellectual Portrait.* New York: Doubleday, 1960.

Black, Max, ed. *The Social Theories of Talcott Parsons.* Englewood Cliffs, NJ: Prentice Hall, 1961.

Bosanquet, Bernard. *The Philosophical Theory of the State.* 4th ed. London: Macmillan, 1923.

Durkheim, Émile. *Les règles de la méthode sociologique.* Paris: Alcan, 1895. Translated by S. A. Solovay and J. H. Mueller as *Rules of Sociological Method,* edited by G. E. G. Catlin. Chicago: University of Chicago Press, 1938.

Durkheim, Émile. *De la division du travail social.* 5th ed. Paris: Alcan, 1926. Edited and translated by George Simpson as *The Division of Labor in Society.* Glencoe, IL: Free Press, 1947.

Gellner, Ernest. "Holism versus Individualism in History and Sociology." In *Theories of History,* edited by Patrick Gardiner, 489–503. Glencoe, IL: Free Press, 1959.

Hegel, G. W. F. *Grundlinien der Philosophie des Rechts.* Berlin, 1821. Translated with notes by T. M. Knox as *Hegel's Philosophy of Right.* Oxford: Clarendon Press, 1942.

Hobbes, Thomas. *Leviathan.* London, 1651. Edited with introduction by Michael Oakeshott. Oxford, 1946.

MacIver, R. M., and C. H. Page. *Society: An Introductory Analysis.* New York: Rinehart, 1949.

Mandelbaum, Maurice. "Societal Facts." In *Theories of History,* op. cit., 476–488.

Parsons, Talcott. *The Structure of Social Action.* Glencoe, IL: Free Press, 1949. Includes extended discussions of the views of Alfred Marshall, Vilfredo Pareto, Émile Durkheim, and Max Weber.

Parsons, Talcott. *The Social System.* Glencoe, IL: Free Press, 1951. Parsons's intricate style and elaborate terminology will deter any but the most determined.

Parsons, Talcott, Edward Shils, K. D. Naegele, and J. R. Pitts, eds. *Theories of Society.* 2 vols. New York: Free Press, 1961.

Popper, Karl R. *The Poverty of Historicism.* London: Routledge, 1957.

Popper, Karl R. *The Open Society and Its Enemies.* 2 vols., 4th rev. ed. London, 1962.

Popper, Karl R. "Prediction and Prophecy in the Social Sciences." In his *Conjectures and Refutations.* London: Routledge, 1963. Ch. 16. Also in *Theories of History,* op. cit., 276–285.

Simmel, Georg. *The Sociology of Georg Simmel.* Translated and edited by K. H. Wolff. Glencoe, IL: Free Press, 1950. An anthology of Simmel's principal works on sociological theory.

Stark, Werner. *The Fundamental Forms of Social Thought.* London: Routledge and K. Paul, 1962. A study of the traditional organicist-mechanistic dichotomy and an attempt to overcome it.

Tönnies, Ferdinand. *Gemeinschaft und Gesellschaft.* Leipzig, 1887. Translated by C. P. Loomis as *Fundamental Concepts of Sociology.* New York: American Book, 1940. Reissued as *Community and Association.* London: Routledge and K. Paul, 1955.

Watkins, J. W. N. "Historical Explanation in the Social Sciences." In *Theories of History,* op. cit., 503–514.

Weber, Max. *Grundriss der Sozialökonomik—III Abt: Wirtschaft und Gesellschaft.* Tübingen: Mohr, 1925. Part I edited and translated by Talcott Parsons and A. M. Henderson (with introduction by Parsons) as *Theory of Social and Economic Organization.* New York: Oxford University Press, 1947.

Wolin, Sheldon S. *Politics and Vision.* Boston: Little, Brown, 1960.

Stanley I. Benn (1967)

SOCIETY [ADDENDUM]

Toward the end of the twentieth century, while earlier discussions of holism *versus* individualism did not die out, the interplay among three different but related notions of society—civil society, the corporation, and cosmopolitan society or the society of nations—an interplay adumbrated in the last two paragraphs above, began increasingly to dominate philosophical inquiry. The development that, more than any other, propelled the notion of civil society back into greater prominence late in that century was an ever more publicly articulated dissatisfaction with the totalitarian nature of the political regimes and their corresponding societies in Eastern Europe. It was widely contended that the suppressed elements of "civil society" in those countries needed to be regenerated

and kept independent of the state. Hence the eventual, generally peaceful dissolution of the governments in question was seen as a triumph of the ideals of civil society.

As in the past, so in the late twentieth and early twenty-first centuries, understandings of the meaning of "civil society" (as well as of "society") have varied widely. Some philosophers, such as Jürgen Habermas, have wished to exclude from the scope of civil society important aspects of the economic institutions that were so central to Hegel's use of the term and to focus on its informal, less easily quantifiable "life-world" elements. For others, the increasing power, in a world characterized by ever-accelerating "globalization," of transnational corporations—"*sociétés anonymes à responsabilité limitée*" in French or "*Gesellschaften mit beschränkter Haftung*" in German—with their essentially capitalist economic purposes and typically nondemocratic structures poses a threat to the viability of political, cultural, and other components of individual (national) civil societies; therefore, according to this line of thinking, corporations need to be treated as focal points in the philosophical analysis of the concepts of both "society" and "civil society." In addition, some have identified, and found great significance in, an emerging *global* civil society, exemplified especially by large transnational nongovernmental organizations (NGOs) that are not essentially profit-oriented, as well as by more informal institutions and practices with similar global concerns.

GLOBAL SOCIETY?

The idea of a global civil society implies that of a global, or cosmopolitan, society as such, contrary to the previously mentioned Parsonian insistence on "boundary maintenance." Resistance to the idea of a global society stems from both methodological and ethicopolitical considerations. John Rawls, for instance, explicitly took the self-contained "closed society"—that is, the nation-state or something similar—as the appropriate abstract entity within which to develop his original theory of justice, which advocates unequal distribution of goods only to the extent to which such distribution will benefit the least advantaged member of that society. This intentional limitation of scope was a methodological preference of his, as it had been of so many of his predecessors in social theory; but it also helped enable him, when he later undertook to analyze international issues in his *The Law of Peoples* (1999), to reject the application of his principles of justice to the world as a whole and to refrain from endorsing cosmopolitanism as a desirable or viable ethicopolitical ideal. (Rawls did, however, introduce the somewhat novel term "Society of Peoples" to refer to those existing "peoples," by no means all, who observe the principles and ideals specified in his book.) Others have used Rawls's theoretical framework in order to develop a more cosmopolitan viewpoint than his own, one that regards "global society" as the name of an emerging contemporary reality, its parts linked by the Internet and other technological innovations, its fate bound up with newly identified shared risks, such as global warming, that some of these innovations have exacerbated, and its extreme imbalances of wealth and poverty perpetuating injustice and instability.

In sharpest reaction to globalizing tendencies and their corresponding theories have been ideologies of resurgent nationalism and religious fundamentalism. The former have, by definition, insisted on the preeminence of individual societies characterized, most frequently, by a perceived common ethnic identity. But considerations of history and genetics alike indicate to how great a measure such perceptions are the products of a particular, time-limited collective imagination, rather than reflections of some underlying truths of social ontology. As for the religious fundamentalist notion that "societies" can be differentiated according to common religious beliefs, a notion shared by some Western writers who subscribe to the vague notion (with constantly shifting boundary definitions), of a global "clash of civilizations," the existence of numerous "warring sects" within the major world religions, combined with basic questions of hermeneutics (that is, how are the sacred scriptures to be interpreted?), casts strong doubt on this way of viewing and intellectually segmenting the world.

It was a British Prime Minister, Margaret Thatcher, rather than a professional philosopher, who is famously reported to have asserted, "There is no such thing as society." This seems a rather extreme claim concerning a supposed reality with references to which so many conversations in ordinary language are replete. It is rather the case, it would seem, that "society" is an exceptionally complex and multivocal term, the complexity and multivocity of which analyses by sociologists, such as Habermas's formalist, structuralist opponent in the broad Parsonian tradition, Niklas Luhmann, and by life-world- and *praxis*-oriented philosophers such as Habermas himself, the phenomenologist Alfred Schutz, and Jean-Paul Sartre in his late-life contribution to social theory, *Critique of Dialectical Reason* (1976), have served to underscore and articulate.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism.

Bibliography

Beck, Ulrich. *World Risk Society*. London: Polity, 1999.

Castoriadis, Cornelius. *The Imaginary Institution of Society*. Translated by K. Blamey. Cambridge, MA: MIT Press, 1987.

Cohen, Jean, and Andrew Arato. *Civil Society and Political Theory*. Cambridge, MA: MIT Press, 1992.

Habermas, Jürgen. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by W. Rehg. Cambridge, MA: MIT Press, 1996.

Jarvie, Ian Charles. *Concepts and Society*. London: Routledge, 1972.

Keane, John. *Global Civil Society?* Cambridge, U.K.: Cambridge University Press, 2003.

Luhmann, Niklas. *The Differentiation of Society*. Translated by S. Holmes and C. Larmore. New York: Columbia University Press, 1982.

Rawls, John. *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.

Sartre, Jean-Paul. *Critique of Dialectical Reason*. Vol. 1, *Theory of Practical Ensembles*, translated by A. Sheridan-Smith. London: New Left Books, 1976.

Schutz, Alfred. *The Phenomenology of the Social World*. Translated by G. Walsh and F. Lehnert. Evanston, IL: Northwestern University Press, 1967.

McBride, William L. (2005)

SOCINIANISM

"Socinianism," an evangelical rationalist movement, was one of the forerunners of modern Unitarianism. Three phases can be distinguished: (1) the thought of Laelius Socinus (1525–1562) and his nephew Faustus Socinus (1539–1604); (2) the thought and institutions of the Minor (Reformed) Church of Poland, especially as embodied in the Racovian Catechism (1605), which represented a fusion of Faustus's theology with that of the local anti-Trinitarian and partly Anabaptist Minor Church; and (3) the rationalist theology of the Socinianized Minor Church. This last phase was especially important after the Socinianized Minor Church was crushed in Poland in 1658 and the spirit of Socinianism became influential in the Netherlands among the Remonstrants; in the British Isles, in the seventeenth century, among certain Anglican divines and nonconformist intellectuals; and, in the eighteenth century, among the Arminian divines of New England, who were forerunners of the Unitarian congregationalists.

Socinian evangelical rationalism originated from an amalgam of the rationalist humanism of Juan de Valdés, Florentine Platonism, and Paduan Aristotelianism; in Poland it was augmented by certain Calvinist and Anabaptist ingredients. In all three phases Socinianism was characterized by (1) a rationalist interpretation of Scripture (which was nevertheless accepted as true and authoritative), with a predilection for the pre-Mosaic and the New Covenantal parts of the Bible; (2) an acceptance of Jesus as the definitive word or revelation of God but nevertheless solely a man, not divine but chosen by God to rule as king, priest, and prophet over the world and the church; (3) belief in the principle of pacific separation of church and state; (4) acceptance of the doctrine of the death of the soul with the body with, however, selective resurrection and immortality for all those who persevered "through the power of the Spirit" in observing all of Jesus' earthly commandments.

LAELIUS AND FAUSTUS SOCINUS

Laelius Socinus, born in Siena, was a well-to-do student with a wide and critical interest in theology. He established contact and became friendly with several reformers, notably Philipp Melanchthon, John Calvin, and Johann Bullinger, and also with the Rhaetian heretic Camillo Renato. Himself suspected of heresy, Laelius was obliged to prepare a Confession of Faith (in which, however, he reserved the right to further inquiry), one of the few extant documents from his hand. At his death he left his library, and perhaps some unpublished papers, to his nephew.

Faustus Socinus, born in Siena, was a student of logic and law, a member of the local academy, and an indifferent poet. He first clearly manifested his rejection of traditional Christian doctrines in a letter of 1563, in which he argued against the postulate of natural immortality. In 1570 he wrote his first major work, *De Auctoritate Sacrae Scripturae*, and in 1578 he issued his basic treatise on Christology and soteriology, *De Jesu Christu Servatore*. Because of the latter work he was invited to Transylvania to defend the legitimacy of prayer addressed to the ascended Christ against the faction in the Unitarian Reformed Church led by Francis Dávid. On the journey he was persuaded to make Poland his permanent home. There he became a major defender of the Minor Church, although he declined on principle to become a communicant member of it, refusing to submit to believers' baptism by immersion. Socinus was cocommissioned with local pastors to revise the Latin *Catechesis* (1574) of Racov, the communitarian settlement and spiritual center of the Minor Church, northeast of Kraków. The radical revision was published in Polish in 1605, a year after Soci-

nus's death, as the *Racovian Catechism*, the first Latin edition of which (1609) was dedicated to James I of England.

THE SOCINIANIZED MINOR CHURCH

The Socinianized Minor Church, centered in Racov, had an academy that at one time attracted a thousand students and a publishing house that turned out tracts and books in a score of languages; it became in fact more a school than a church. Among the faculty of the academy and the pastorate of the synod, which met annually in Racov, the most prominent were Socinus's own grandson, Andreas Wiszowaty (d. 1678), who wrote *Religio Rationalis*; Stanislas Lubieniecki (d. 1675), who wrote *Historia Reformationis Polonicae*; Samuel Przypkowski (d. 1670), who wrote *Vita Fausti Socini*; and quite a few converts from German Protestantism who resettled in Poland and were rebaptized: Christoph Ostorodt (d. 1611); Johann Völkel, who wrote *De Vera Religione* (1630); Johann Crell (d. 1631), who wrote *De Uno Deo Patre* and a defense of Socinus against Hugo Grotius, *De Satisfactione*; and Christoph Sand (d. 1680), who compiled the *Bibliotheca Antitrinitariorum*.

SPREAD OF SOCINIANISM

Well before the crushing of the Minor Church in 1658, Socinians were established in the Netherlands. At Amsterdam the basic works of the movement, the eight-volume *Bibliotheca Fratrum Polonorum*, edited by Wiszowaty, were printed in 1688. In England, Socinian rationality, latitudinarianism, Unitarianism, and mortalism (psychopannychism) variously appealed to Arminian prelates, Oxford rationalists (such as William Chillingworth), Cambridge Platonists (such as Benjamin Whichcote), philosophers and scientists (such as Isaac Newton and John Locke), and to the first avowed native Socinians, Paul Best, John Biddle ("the father of English Unitarianism"), and Stephen Nye, whose *History of Unitarianism commonly called Socinianism* set off the Trinitarian controversy in the Established church in 1687.

See also Arminius and Arminianism; Calvin, John; Cambridge Platonists; Grotius, Hugo; Locke, John; Melanchthon, Philipp; Newton, Isaac; Rationalism; Whichcote, Benjamin.

Bibliography

Bianchi, Daniela. "Some Sources for a History of English Socinianism: A Bibliography of 17th Century English Socinian Writings." *Topoi* 4 (1985): 91–120.

Chmaj, Ludwik. *Faust Socyn*. Warsaw: Ksiazka i Wiedza, 1963.

Edwards, John. *Some Thoughts concerning the Several Causes and Occasions of Atheism: Socinianism Unmask'd*. New York: Garland, 1984.

Kühler, W. J. *Socinianisme in Nederland*. Leiden, 1912.

McLachlan, H. John. *Socinianism in Seventeenth-Century England*. London: Oxford University Press, 1951.

Stewart, M. A. *English Philosophy in the Age of Locke*. New York: Oxford University Press, 2000.

Wilbur, Earl M. *A History of Unitarianism: Socinianism and Its Antecedents*. Cambridge, MA: Harvard University Press, 1945.

Yolton, John W., ed. *Philosophy, Religion and Science in the Seventeenth and Eighteenth Centuries*. New York: University of Rochester Press, 1990.

George Hunston Williams (1967)
Bibliography updated by Tamra Frei (2005)

SOCINUS, LAELIUS AND FAUSTUS

See *Socinianism*

SOCIOLOGY, FUNCTIONALISM IN

See *Functionalism in Sociology*

SOCIOLOGY OF KNOWLEDGE

The "sociology of knowledge" is concerned with determining whether human participation in social life has any influence on human knowledge, thought, and culture and, if it does, what sort of influence it is.

DEVELOPMENT

Although the term *sociology of knowledge* was coined in the twentieth century, the origins of the discipline date back to classical antiquity. Plato, for instance, asserted that the lower classes are unfit to pursue the higher kinds of knowledge, because their mechanical crafts not only deform their bodies but also confuse their souls. Plato also held the more refined doctrine of the correspondence of the knower (or more precisely, the faculties and activities of the knower's mind, which are in part determined by society) and the known. This latter theory became part of the Platonic tradition and ultimately stimulated some modern pioneers in the sociology of

knowledge, notably Max Scheler. Both theories anticipated an essential claim of the sociology of knowledge—that social circumstances, by shaping the subject of knowing, also determine the objects that come to be known.

In the Middle Ages, patterns of life were fixed and defined, and patterns of thought tended to be equally so; ideas appeared as absolute, and the factors that conditioned them remained hidden. As soon, however, as rifts developed in the social fabric, awareness of these factors reemerged. Niccolò Machiavelli's remark in the *Discourses* that the thought of the palace was one thing, the thought of the market place quite another, revealed this new awareness.

In the following centuries, the stream of ideas that was to lead to the modern sociology of knowledge was divided between rationalism and empiricism. The rationalists regarded mathematical propositions as the archetype of truth. As mathematical propositions do not change in content from age to age and from country to country, the rationalists could not concede that different societies might have different systems of knowledge, all equally valid. But if truth was one, error could be multiform, and its roots could be sought in social life—for instance, in the machinations of privileged classes in whose interest it was to keep the people in ignorance. Francis Bacon's doctrine of "idols," or sources of delusion, set forth in his *Novum Organum*, illustrates this tendency. The rationalists thus became the first "unmaskers" of "ideologies."

According to the empiricists, the contents of the mind depend on the basic life experiences, and as these are manifestly dissimilar in dissimilarly circumstanced societies, they almost had to assume that reality would offer a different face in each society. Thus, Giambattista Vico asserted that every phase of history has its own style of thought which provides it with a specific and appropriate cultural mentality. The treatment of the biblical account of creation by the two schools shows their contrast. Voltaire called it a piece of stultifying priestcraft that no rational person anywhere would accept: How could the light exist before the sun? Johann Gottfried Herder answered that for a desert nation like the ancient Hebrews the dawn breaks before the solar disk appears above the horizon. For them, therefore, the light *was* before the sun.

Though the problems of the genesis of error and the genesis of truth should be kept apart, the overly sharp distinction between them and the partisan handling of them before the end of the eighteenth century prevented any tangible progress. And even though Immanuel Kant achieved a synthesis of rationalism and empiricism, the sociology of knowledge failed to gain from his advances. Kant's whole approach prevented such a gain: The problem of knowledge arose for Kant from the meeting of the individual mind with the physical world. The social element was missing at either pole. The sociology of knowledge explains Kant's narrowness itself as socially determined. The decay of feudal society and the emergence of a class of independent producers (peasants and artisans) had created the desire to "liberate" man from the "artificial restrictions" of social life. A presocial, asocial, or antisocial type of man was thought possible and even superior to social man. The primacy of being was ascribed to the individual, and society was considered to be no more than a collection of individuals linked by contract. In these circumstances, no one could see the influence of social forces on the human mind.

The nineteenth century brought a strong reaction against this radical individualism. As the forces of social control reasserted themselves, man was once again conceived of as essentially a social creature. The result of this new trend was Karl Marx's mislabeled "materialistic interpretation of history." Marx wrote in his *Introduction to the Critique of Political Economy*: "It is not men's consciousness which determines their existence, but on the contrary their social existence which determines their consciousness." For Marx, the real "substructure" upon which the intellectual "superstructure" rests is a special set of human relationships. Though his definition of these relationships is too narrow, and though he has been variously interpreted, Marx's formulation provided the starting point in the development of the modern sociology of knowledge.

SOCIAL ORIGIN OF IDEAS

While there is general agreement among scholars in the field that social relationships provide the key to the understanding of the genesis of ideas, there are also far-reaching disagreements among several distinct schools, within which there are again individual differences. An attempt will be made here only to characterize the three most important basic attitudes.

MATERIALIST SCHOOL. A materialist group of writers emphasizes that human beings are creatures of nature before they are creatures of society and tends to see human beings as dominated by certain genetic drives, with decisive consequences for their emergent mentalities. Friedrich Nietzsche, for instance, ascribed to man an

elementary will to power; if this will is frustrated by a barrier, self-consolatory ideas are apt to appear. Christianity is one such idea; it is essentially a philosophy of "sour grapes," a "slave morality." It assures the defeated that they are really superior to those who have defeated them.

Vilfredo Pareto's *Trattato di sociologia generale* is the most elaborate statement of this position. According to Pareto, people act first and think of reasons for their action only afterward. These reasons he calls "derivations" because they are derived from, or secondary to, the "residues," or quasi instincts, which in fact determine human modes of conduct and, through them, human modes of thought as well. This school continued the line initiated by the rationalists. Theirs is a doctrine of ideologies that devalues thought while it accounts for its formation.

IDEALIST SCHOOL. A second group of writers asserts that every society has to come to some decision about the Absolute and that this decision will act as a basic premise that determines the content of the culture. Juan Donoso Cortés tried to explain the classical Greek worldview as the product of heathen preconceptions about the Absolute, and the medieval worldview as the product of Christian-Catholic preconceptions. An ambitious presentation of this theory is Pitirim Sorokin's *Social and Cultural Dynamics*. He distinguishes three basic metaphysics that, prevailing in given societies, color all their thinking. If a realm beyond space and time is posited as the Absolute, as in ancient India, an "ideational" mentality will spring up; if the realm inside space and time is posited as the Absolute, as in the modern West, a "sensate" mentality will come into being; and if, finally, reality is ascribed both to the here and now and to the beyond, as in the high Middle Ages, an "idealistic" mentality will be the result. Sorokin's doctrine is itself idealistic in character and finds its ultimate inspiration in a religious attitude.

SOCIOLOGISTS OF KNOWLEDGE. The third group of writers occupies the middle ground. These writers do not go beyond the human sphere but divide it into a primary and conditioning half and a secondary and conditioned one. There is, however, great diversity of opinion over exactly which social facts should be regarded as conditioning thought. Marx, for instance, held that relations of production, which themselves reflect still more basic property relations, were primary, but many other factors, such as power relations, have been singled out by other thinkers. Still others regard the social constitution as a whole as the substructure of knowledge, thought, and culture. A typical representative of this numerous group is W. G. Sumner. In his classic *Folkways*, he suggested that wherever individuals try to live together, they develop mutual adjustments that harden into a set of customs, supported and secured by social sanctions, which permanently coordinate and control their conduct. These habits of action have as their concomitants habits of the mind, a generalized ethos that permeates the mental life of the society concerned. This theory can be sharpened by formulating it in axiological terms. A society is a society because, and insofar as, it is attuned to certain selected and hierarchically ordered values. These values determine what lines of endeavor will be pursued both in practice and in theory.

This third group represents the sociology of knowledge in the narrower and proper sense of the word. The theory just summed up has received some empirical confirmation through the discovery that societies do gain mental consistency to the degree that they achieve better human coordination and integration.

RELATION OF A SOCIETY TO IDEAS EXPRESSED IN IT

The problem next in importance to the identification of the substructure of knowledge is the explanation of its relation to the superstructure. Here again there are three schools that may, but do not always, correspond to those already discussed. One tendency is toward causalism. The positivists Gustav Ratzenhofer and Hippolyte Taine, for example, expected of the future a science of culture no less deterministic than the sciences of matter. But though the term *determination* is frequently and generally used in all the literature of this school, it hardly ever means strict determination. While this first school concedes, in principle, no independence to the mind and its contents, a second, Platonic tendency ascribes complete independence to the mind. To Scheler, Florian Znaniecki, and others, thinking means participating in eternal preexistent ideas. If these ideas are to become active in the world, they must ally themselves to a social movement seeking appropriate ideas. Max Weber has called this doctrine the doctrine of elective affinity. The third theory argues in terms of interdependence and appears regularly in connection with functionalism. If society is to function as a unity, its modes of acting and thinking must be in, or on the way to, agreement. Neither substructure nor superstructure is given ontological priority, but there is a tendency to see thought in action as prior to thought as theory.

EXTENT OF INFLUENCE. Another problem concerns the extent of the influence of social factors on ideas. Here opinions range from the view that these factors influence only a few political slogans to the view that their influence is all-pervading. An important systematic dividing line separates the authors who assert that the categories of thought themselves are socially determined from those who deny that they are.

EPISTEMOLOGICAL SIGNIFICANCE

The main philosophical importance of the sociology of knowledge consists in its claim to supplement, if not to replace, traditional epistemology. If society partially or totally determines knowing and thinking, how does this affect their validity? All sociologists of knowledge are inclined to stress that initially the human mind is never aware of more than a sector of reality and that the selection of a sector to be investigated is dependent on the axiological system that a given society has made its own. From this point they diverge once again into three schools, and once more there is no simple correlation with the tendencies previously identified.

EFFECT OF SOCIAL FACTORS ON THOUGHT. Some writers, such as Pareto, hold that, in the last analysis, only the senses are reliable sources of knowledge. They tend to split the mental universe into a scientific and a nonscientific department and accord the ideas belonging to the latter at best conventional status, but no truth-value in the narrower sense of the term. The axiological system of society, insofar as it is not taken up with scientific and technological pursuits, appears as an opaque and distorting medium that interposes itself between the intellect and reality. The effect of society on the mind is thus something negative, to be regretted and, if possible, overcome.

Whereas this group denigrates the social element in human beings, and hence in human knowledge, another, including Émile Durkheim and Karl Mannheim, sees it as supreme. The latter group conceives the individual as the most likely source of error and society as the most reliable source of truth, if for no other reason than because personal blunders are neutralized in a common attitude. They regard society as the test of the validity of a belief: It is valid if those who hold it manage to operate smoothly within their social system. But if the true is what works and if different societies work differently (as manifestly they do), then truth is once again merely convention. At any rate, there can be no general truths.

The third group, including Weber and Scheler, considers that the social influence on mental activity consists essentially in giving directions. What knowledge will be sought in a society depends on the axiological system that reigns in that society. In its most radical form, this doctrine sees our very awareness of facts as socially determined: Only those aspects of reality that are marked by their possession of some value, social in origin, will be noticed and enter into the canon of knowledge. There appears, however, no cogent reason why a person should not see a thing thus selected for study on an axiological basis as what it really is. It can therefore be said that every society has its own truth, without giving the word a relativistic tinge. Any human being who integrates himself, factually or intellectually, with a certain society and accepts its constitutive values will have to agree that, from the chosen angle, the world does, and must, look as it is described by the searchers and thinkers of that society. Hence sociality is neither a truth-destroying nor a truth-guaranteeing, but merely a truth-limiting factor. The resulting limitations can, in principle, be overcome by combining the valid "aspectual" insights of all societies into a comprehensive whole.

KNOWLEDGE OF NATURE AND KNOWLEDGE OF CULTURE

An important distinction sometimes made is that between knowledge of nature and knowledge of culture. The facts of nature do not change from age to age and from country to country; the facts of culture do. Knowledge of the former, therefore, need not be marked by relativity. The Paretian theory, by making physical knowledge the model of all knowledge, does less than justice to the study of cultures; the theory of Mannheim and Durkheim, by making cultural knowledge the model of all knowledge, is apt to fall into the opposite mistake (though its best protagonists have managed to avoid this). The theory of Weber and Scheler escapes both weaknesses. In every society's axiological system, some interest in nature, especially in methods of dominating nature, will be present, and insights gained in the pursuit of this domination will be comparable, transferable, and absolute in the sense of binding on all human beings. Other values will vary from society to society; insights gained in pursuit of them will be correspondingly incomparable, nontransferable, and relative (even though they can all be fitted together as alternative actualized possibilities inherent in one creature, man).

Because people must take the facts of nature as they find them, while the facts of culture are their own work,

the social determination of knowledge will be different in the two instances. In scientific research, only the origin of an insight will be determined by the social factor (say, a pressing social need); in cultural studies, however, both the origin and the content will be socially determined. In the case of science, tendencies arising from the social sphere induce a person to open his eyes and see; in the case of cultural studies, they induce him to open his eyes and decide what he shall see. These considerations go far toward overcoming the conflict between the unduly negative and the unduly positive epistemological versions of the sociology of knowledge and show the superiority of the third approach.

SOCIOLOGY OF KNOWLEDGE AS A SCIENCE

In conclusion, it should be emphasized that the sociology of knowledge is not only a substantive philosophical discipline but also an analytical tool that can be used by the descriptive sciences concerned with the observable products of the mind. Because it can throw light on the genesis, and often on the content, of concrete thought structures, the sociology of knowledge may enable the historian or the anthropologist to achieve a deeper understanding of the facts before him. Considered from this angle, the sociology of knowledge appears, above all, as a hermeneutic method and need not become involved in the difficult ontological problems that the social "determination" of knowledge, thought, and culture is otherwise bound to raise.

See also Functionalism in Sociology.

Bibliography

DEVELOPMENT

Barth, Hans. *Wahrheit und Ideologie*. Zürich: Manesse, 1945. Important, particularly with regard to Nietzsche.

Grünwald, Ernst. *Das Problem der Soziologie des Wissens*. Vienna: Braumüller, 1934. First comprehensive critical survey.

Lukács, Georg. *Geschichte und Klassenbewusstsein*. Berlin: Malik, 1923. Most distinguished Marxist study of the twentieth century.

Stark, Werner. "The Conservative Tradition in the Sociology of Knowledge." *Kyklos* 13 (1960): 90–101.

Stark, Werner. "Die idealistische Geschichtsauffassung und die Wissenssoziologie." *Archiv für Rechts- und Sozialphilosophie* 46 (1961): 355–374. This paper and the preceding one attempt to balance the picture of the doctrine's development.

Stark, Werner. *Montesquieu: Pioneer of the Sociology of Knowledge*. London: Routledge and Paul, 1960.

MAIN TRENDS

Adler, Franz. "A Quantitative Study in Sociology of Knowledge." *American Sociological Review* 19 (1954): 42–48. From the viewpoint of logical positivism.

Alpert, Harry. *Emile Durkheim and His Sociology*. New York: Columbia University Press, 1939.

Eisermann, Gottfried. "Vilfredo Pareto als Wissenssoziologe." *Kyklos* 15 (1962): 427–464.

Lieber, Hans-Joachim. *Wissen und Gesellschaft*. Tübingen: Niemeyer, 1952. Good all around, mainly on Scheler and Mannheim.

Mannheim, Karl. *Essays on the Sociology of Knowledge,* edited by Paul Kecskemeti. London: Routledge, 1952.

Mannheim, Karl. *Ideologie und Utopie*. Bonn: Cohen, 1929. Translated by Louis Wirth and Edward Shils as *Ideology and Utopia*. London: Routledge, 1936. Basic; from the historicist point of view, with comprehensive bibliography.

Maquet, Jacques J. *Sociologie de la connaisance*. Louvain, 1949. Translated by John F. Locke as *The Sociology of Knowledge*. Boston: Beacon Press, 1951. On Mannheim and Sorokin.

Rüschmeyer, Dietrich. *Probleme der Wissenssoziologie*. Cologne, 1958. Strong bias toward naive empiricism.

Scheler, Max, ed. *Versuche zu einer Soziologie des Wissens*. Munich: Duncker and Humblot, 1924. Decisive pioneering effort.

Scheler, Max. *Die Wissensformen und die Gesellschaft*. Leipzig: Neue Geist, 1926. Basic; from the phenomenological point of view.

Znaniecki, Florian. *The Social Role of the Man of Knowledge*. New York: Columbia University Press, 1940.

PHILOSOPHICAL ASPECTS

Child, Arthur. "The Existential Determination of Thought." *Ethics* 52 (1941–1942): 153–185.

Child, Arthur. "The Problem of Imputation in the Sociology of Knowledge." *Ethics* 51 (1940–1941): 200–219.

Child, Arthur. "The Problems of Imputation Resolved." *Ethics* 54 (1943–1944): 96–109.

Child, Arthur. "The Theoretical Possibility of the Sociology of Knowledge." *Ethics* 51 (1940–1941): 392–418.

Geiger, Theodor. *Ideologie und Wahrheit*. Stuttgart: Homboldt-Verlag, 1953. Rationalistic approach.

Horowitz, Irving L. *Philosophy, Science, and the Sociology of Knowledge*. Springfield, IL: Thomas, 1961. Methodological problems.

Schelting, Alexander von. *Max Webers Wissenschaftslehre*. Tübingen, 1933. Basic.

Stark, Werner. "The Sociology of Knowledge and the Problem of Ethics." In *Transactions of the Fourth World Congress of Sociology*, Vol. IV. London, 1959.

GENERAL SURVEYS

De Gré, Gerard L. *Society and Ideology*. New York, 1943.

Merton, Robert K. "The Sociology of Knowledge." In *Twentieth Century Sociology*, edited by Georges Gurvitch and Wilbert E. Moore. New York: Philosophical Library, 1946.

Stark, Werner. *The Sociology of Knowledge*. London: Routledge and Paul, 1958. Attempt at a total evaluation.

OTHER RECOMMENDED TITLES

Alston, William P. "Belief-Forming Practices and the Social." In *Socializing Epistemology*, edited by Frederick Schmitt. Lanham, MD: Rowman & Littlefield, 1994.

Bloor, D. *Knowledge and Social Imagery*. Chicago: University of Chicago Press, 1991.

Brown, J. R. *Scientific Rationality: The Sociological Turn*. Dordrecht: Reidel, 1984.

Fuller, Steve. "Recent Work on Social Epistemology." *American Philosophical Quarterly* 33 (1996): 149–166.

Hollis, M., and S. Lukes. *Rationality and Relativism*. Oxford: Blackwell, 1982.

Katz, Jonathan. "Rational Common Ground in the Sociology of Knowledge." *Philosophy of the Social Sciences* 19 (1989): 257–271.

Kurzman, Charles. "Epistemology and the Sociology of Knowledge." *Philosophy of the Social Sciences* 24 (1994): 267–290.

Latour, B. *Science in Action*. Cambridge, MA: Harvard University Press, 1987.

Longino, H. *Science as Social Knowledge*. Princeton, NJ: Princeton University Press, 1990.

McMullin, Ernan, ed. *The Social Dimensions of Knowledge*. South Bend, IN: Notre Dame University Press, 1992.

Okasha, S. "The Underdetermination of Theory by Data and the 'Strong Programme' in the Sociology of Knowledge." *International Studies in the Philosophy of Science* 14 (2000): 283–297.

Rouse, J. *Knowledge and Power: Toward a Political Philosophy of Science*. Ithaca, NY: Cornell University Press, 1987.

Schmitt, F. F., ed. *Socializing Epistemology*. Lanham, MD: Rowman & Littlefield, 1994.

Susser, Bernard. "The Sociology of Knowledge and Its Enemies." *Inquiry* 32 (1989): 245–260.

Werner Stark (1967)
Bibliography updated by Benjamin Fiedor (2005)

SOCRATES

(c. 470–399 BCE)

Socrates is the first Western philosopher to have left to posterity any sense of his individual personality, and he is a central figure in the subsequent development of philosophy. Both of these aspects are due primarily to Plato. It is via his portrayal by Plato's literary genius that Socrates is a living figure for subsequent generations, and thereby an exemplar of the ideals of philosophy, above all dedication to truth and intellectual integrity. It was under the influence of Socrates that Plato applied systematic techniques of argument pioneered by Socrates and his contemporaries, the Sophists, to the fundamental questions of human nature and conduct that primarily interested Socrates, thereby placing ethics and psychology at the center of the philosophical agenda. But while Plato brings Socrates to center stage he also hides him; because Socrates wrote nothing himself we depend on others for our knowledge of him, and it is above all Plato's representation of Socrates that constitutes the figure of perennial philosophical significance. But that representation was itself the expression of Plato's understanding of an actual historical individual and the events of his life. It is necessary, therefore, to begin with a brief account of the little that is known of that individual and those events.

LIFE

Socrates was born in Athens around 470 BCE and lived in the city all his life, apart from military service abroad. Little is known of the circumstances of his life. His father, Sophroniscus, is said by some ancient sources to have been a stonemason, and in Plato's *Theaetetus* (149a) Socrates says that his mother, Phainarete, was a midwife. That may indeed be true, though the fact that the name literally means "revealing excellence" suggests the possibility that Plato has invented the story in allusion to Socrates' role as midwife to the ideas of others (*Theaetetus* 149–151). Because Socrates served in the infantry, who had to provide their own arms and equipment, his circumstances, at least initially, must have been reasonably prosperous, but Plato and other writers emphasize his poverty in later life, which they attribute to his spending all his time in philosophical discussion. The same sources stress that, unlike the Sophists, he never took payment for his philosophical activity, and he may have depended largely on support from wealthier friends. During his lifetime Athens became the principal center of intellectual and cultural life in Greece, attracting from all over the Greek world intellectuals who developed and popularized the tradition of natural philosophy begun by the Ionian philosophers of the previous century, together with exciting new argumentative techniques and radical questioning of traditional beliefs about theology, morals, and society.

Socrates was actively interested in most of these areas. Plato and others attest to his interest at one stage in questions of cosmology and physiology, though the sources agree that his interests subsequently shifted to fundamental questions of conduct. Socrates never engaged in formal philosophical instruction, or set up any school; his philosophical activity consisted in informal conversation, partly with a circle of mainly younger associates whom he attracted by the force of his intellect and personality, but also with others, including Sophists and prominent citizens. Some of his associates, including Plato and some of his relations, were opposed to the

Athenian democratic system, and it may be that Socrates shared that attitude to some extent.

Socrates married relatively late in life; at the time of his death at about the age of seventy his eldest son was an adolescent, and he had two more small sons, the younger probably a baby. His wife (who must have been at least thirty years younger than he) was Xanthippe. Her bad temper (attested by Xenophon and others, but not by Plato) became legendary; stories of her abuse of Socrates, and his equanimity in putting up with it, were a stock comic theme from antiquity to modern times. Thus Chaucer's Wife of Bath describes in the Prologue to her tale (727–732) how Socrates sat quietly while Xanthippe "caste pisse upon his heed," merely remarking mildly, "Before the thunder stops it comes on to rain." (The story goes back to Diogenes Laertius's life of Socrates, *Lives of the Philosophers* 2.36.) One element in this comic tradition is the story that Socrates had another wife, or possibly a concubine, while married to Xanthippe; stories of how the two women switched from quarrelling with one another to concerted assaults on Socrates afforded rich material. Ancient sources attribute the origin of this tale to Aristotle, but the supposed original source is lost, and the historical basis extremely dubious.

Nothing is known of specific events in Socrates' life till after the outbreak of the Peloponnesian War with Sparta in 432. He served with distinction in various campaigns, most notably the Battle of Delium in 424, where it was said (Plato, *Laches* 181b) that if everyone had behaved like Socrates the battle would not have been lost. By the 420s he had become sufficiently well known to be caricatured in several comic dramas. In the single example to survive complete, the *Clouds* of Aristophanes, first produced in 423, he appears as a representative of subversive contemporary tendencies, the head of a disreputable academy whose curriculum combines training in argumentative trickery with atheistic natural philosophy. Later, in his *Apology* (Defense of Socrates), Plato represents this portrayal as the origin of prejudice against Socrates that culminated in his condemnation on charges of impiety and corruption of the young (19a–19d); there is no reason to discount that evidence.

The only occasion on which Socrates is known to have intervened in public life took place in 406. After a naval engagement the Athenian commanders had failed to pick up survivors, and the popular assembly voted to try them collectively, instead of individually as required by law. At that period most civic offices were assigned by lot, and Socrates happened to be a member of the executive committee whose function was to prepare business for the assembly. In that capacity he was the only one to oppose the illegal proposal. A few years later when, after final defeat in the war, the democracy was temporarily overthrown by a junta known as the Thirty Tyrants, he showed the same adherence to legality and morality by refusing, at the risk of his own life, to obey an order from the tyrants to take part in the arrest of an innocent man. It is likely that he remained neutral during the civil war in which the tyranny was overthrown, because he had friends in both camps; in particular, two of the most prominent among the tyrants, Critias and Charmides, both relatives of Plato, were among his close associates.

It is probable that this was at least a contributory factor in the accusation brought against him under the restored democracy. The explicit charges were failure to recognize (or perhaps "to believe in") the gods of the state religion and the introduction of new divinities, coupled with corruption of the young. The case was tried early in 399, and the prosecution demanded the death penalty. There is no evidence of the detail of the prosecution's case. On the religious aspect the prosecutors may have sought to represent Socrates as the leader of an illegal private cult, and may have used his claim, amply attested by Plato, to be guided by a private divine sign or voice in support of that charge. It is highly likely that the charge of corruption centered on his associations with notorious enemies of the state, particularly the tyrants mentioned above, as well as Alcibiades, an intimate of Socrates who had instigated a disastrous invasion of Sicily in 415 and had later defected to Sparta. Knowledge of the trial is based on two versions of Socrates' defense, by Plato and Xenophon, each of whom, while preserving a core of fact, presents the defense in the light of his own agenda; Xenophon relies wholly on Socrates' adherence to conventional piety and morality, whereas Plato gives a radically unconventional picture of Socrates' philosophical activity as the fulfillment of a divine mission to perfect the souls of his fellow citizens by subjecting their basic beliefs and values to philosophical criticism.

Socrates was condemned to death. Plato's *Phaedo* gives a moving picture of his last hours, spent among his followers in discussion of the immortality of the soul and the task of philosophy to free it from the trammels of the body, followed by his tranquil death from self-administration of hemlock. While there is dispute about the relative degrees of realism and idealization in the description of the effects of the poison, there is little doubt that the primary aim of the whole work is less historical accuracy than depiction of the ideal philosophical death.

SOCRATIC LITERATURE

Besides Plato and Xenophon no fewer than nine associates of Socrates are reported by various ancient sources as having written imaginative accounts of Socrates' conversations, creating a body of literature collectively known as "Socratic conversations" (or "discourses") (*Sokratikoi logoi*). For the most part only the titles of these works survive, indicating that Socrates' relations with certain individuals, especially Alcibiades, who figures prominently in some Platonic dialogues—notably *Alcibiades* and *Symposium*—were a theme common to Plato and the other Socratic writers. Apart from Plato and Xenophon, the only Socratic writer of whose works any significant fragments survive is Aeschines of Sphettus; the fragments of his *Alcibiades* show Socrates using his characteristic critical method (see below) to convince Alcibiades of the vanity of his political ambitions. They thus provide evidence that the program of defending Socrates against the slanders occasioned by his associations with political undesirables was not confined to Plato and Xenophon, but they provide no evidence for Socrates' thought to complement those sources.

For information specifically about the thought of Socrates scholars are in fact almost wholly dependent on Plato, because the other principal source, Xenophon, focuses on the practical and moral import of Socrates' conversations, with comparatively little theoretical content, in keeping with his overall purpose (see above) of portraying Socrates as a good man and sound citizen. There is a systematic difficulty in determining which of the views attributed to Socrates in Plato's dialogues were actually held by the historical person, and scholarly opinion has embraced all possible positions. In the nineteenth century the dominant consensus (primarily on the part of German scholars) divided the Platonic writings into three broad groups, distinguished both chronologically and doctrinally. The first "early" group, including *Laches*, *Charmides*, *Protagoras*, and those dialogues dealing directly with the trial of Socrates (*Euthyphro*, *Apology*, and *Crito*), was generally held to give a veridical account of the personality, views, and philosophical activity of the historical Socrates.

Thereafter Plato's philosophy developed in directions independent of Socrates, and the importance of the dramatic figure of Socrates in the dialogues correspondingly declined, until its virtual disappearance in works such as the *Sophist* and the *Statesman* (which were taken to be late), and its total disappearance from the *Laws* (unfinished on Plato's death and generally regarded as his last work). This "developmental" model was supported by the stylometric studies of the later nineteenth century, in which a number of scholars, working largely independently of one another, converged on the identification of six dialogues—*Sophist*, *Statesman*, *Timaeus*, *Critias*, *Philebus*, and *Laws*—as a group distinct in various features of style and vocabulary from the rest of the Platonic corpus; these dialogues were fixed as late by the presence of the *Laws*. *Parmenides*, *Phaedrus*, *Republic*, and *Theaetetus*, which are by the same criteria closer in style to the late group than the rest of the dialogues, were identified as "middle" dialogues, and the remainder as "early."

While this developmental model, with its assumption that the early dialogues accurately represent the historical Socrates, is still highly significant in the twenty-first century, notably in the influential work of Gregory Vlastos and others, it has undergone challenge from two opposite extremes, on the one side the thesis maintained by John Burnet and A. E. Taylor in the early twentieth century that all the doctrines attributed by Plato to Socrates in the dialogues were actually maintained by the historical Socrates, and on the other side the views of those who, stressing that all information about Socrates derives from sources with their own literary and philosophical agenda, urge that the historical Socrates is inaccessible and should therefore disappear from the history of philosophy.

The Burnet/Taylor thesis has few if any adherents in the twenty-first century; not only does it present an implausible picture of a Plato who devoted the great part of his literary career to recounting the views of someone else, but it rests on an assumption about the nature of Plato's attitude to Socrates, namely that it would have been disrespectful to Socrates for Plato to do other than represent his views with historical accuracy, which seems totally foreign to the character of the dialogues themselves. It is clear from the dialogues that Plato's attitude to Socrates was that the latter's life and activity represented the paradigm of philosophy, and it is totally in keeping with that attitude that Plato should ascribe to Socrates what he (Plato) regards as the philosophical truth, whether or not Socrates himself had maintained it. What we may call the skeptical view of Socrates, however, is widely accepted today, and while its extreme versions are exaggerated and oversimplified, it is based on an important insight into the nature of our sources.

The insight is simply that all knowledge of Socrates is based on sources in which historical veridicality is at best one among the author's concerns, and generally not the principal concern. Oversimplification consists in the characterization of these sources as fiction, as opposed to

factual biography, and exaggeration in the conclusion that the historical Socrates is inaccessible. The dichotomy between biography and fiction seems inapplicable to the Socratic literature, including Plato's Socratic dialogues (and indeed this author doubts its appropriateness to most ancient biographical writing); Socratic conversation is a form of biography, but biography whose factual constraints are looser than is standardly the case in the modern world.

That is not to say that there are no factual constraints; Plato's dialogues do present an actual historical individual, some of the events in whose life are known, and they are no doubt faithful to the spirit and nature of the philosophical conversation that was that individual's principal activity. But when it comes to specific doctrines, while there are some doctrines found maintained by Plato's Socrates that it is virtually certain the historical Socrates did not maintain, there are none that is certain that he did. In the first class the paradigm case is the theory of separate Forms, (i.e., intelligible universal natures existing separately from their sensible instances) which we find maintained by Socrates in several dialogues, but which Aristotle (whose evidence this author regards as independent of the dialogues on this point) explicitly says Socrates did not hold (*Metaphysics* 1078b27–1078b32).

However, theses characteristically regarded as "Socratic"—for example, that Virtue is Knowledge (see below)—are not ascribed to Socrates by sources that are clearly independent of their appearance in the Platonic dialogues. They may in fact have been maintained by Socrates, or they may have been suggested to Plato, in the form in which they appear in the dialogues, by things that Socrates said. We cannot be sure, and in any case it is not of the first importance, because the philosophical significance of these doctrines consists in the role that they play, and the arguments by which they are supported, in the dialogues in which they appear. The brief account of Socrates' thought that follows is to be understood as based on that assumption. It identifies some central themes in the portrayal of Socrates in those dialogues, generally considered comparatively early compositions, in which the personality and argumentative style of Socrates are more prominent than in dialogues devoted to the more systematic exposition of Plato's own thought (see above). The attribution of any specific doctrine to the historical Socrates must be correspondingly tentative.

THOUGHT

DISAVOWAL OF WISDOM. In these dialogues Socrates is presented for the most part not as a systematic or authoritative teacher, but as a questioner and enquirer. His enquiries are all focused on questions of conduct, broadly understood, and frequently consist of attempts to reach an agreed definition of some fundamental value, such as courage, or goodness in general. Typically Socrates is depicted as engaged with one or more people in conversation on some specific, often practical topic, which leads on to the more general issues just mentioned. Socrates elicits the views of his interlocutors on these issues and subjects them to critical examination, conducted with a minimum of philosophical technicality, and utilizing other assumptions, usually of a commonsense kind, which the parties to the discussion agree on. Usually this procedure reveals inconsistency among the set of beliefs (including the general thesis or proposed definition) that the person examined holds, which is taken as requiring the abandonment of the thesis or definition. Frequently the dialogue ends with the acknowledgement by Socrates and the others that, having failed to settle the general issue raised, they are unable to proceed further; they thus end up in a state of *aporia*—that is, a state with no way out. This procedure of enquiry, rather than instruction, and its frequent aporetic outcome are in keeping with Socrates' denial (*Apology* 21b) that he possesses any wisdom (i.e., expertise). It is the mark of an expert to be able to define the concepts in the area of his expertise and to expound that area systematically, neither of which Socrates can do.

In later antiquity Socrates was regularly reported as having said that he knew nothing, or, paradoxically, that he knew nothing except that he knew nothing. Either formulation goes beyond anything found in Plato. Though Socrates frequently says in the dialogues that he does not know the answer to this or that particular question, he never says that he knows nothing, and occasionally makes emphatic claims to knowledge, most notably in the *Apology*, where he twice claims to know that abandoning his divine mission to philosophize would be bad and disgraceful (29b, 37b).

What he does disavow is having any wisdom. He seems to apply the notion of wisdom firstly to divine wisdom, a complete and perspicuous understanding of everything, that belongs to the gods alone, and is consequently unavailable to humans, and then to human expertise of the sort possessed by craftsmen such a builders and shoemakers, a systematic mastery of a technique that enables its possessor to apply it successfully and to expound and pass it on to others. The Sophists claimed to possess, and to teach to others, a practical expertise applying not to any specialized area of human

activity but to human life as such, mastery of which guaranteed overall success in personal and political life; this was "the political craft" (*Apology* 19d–20c, *Protagoras* 319a). Socrates rejects that claim, not on the ground that such expertise is not available to humans; but because the Sophists' activity fails to meet the ordinary criteria for human expertise, particularly that of being systematically learned and taught (*Protagoras*319d–320b, *Meno* 89c–94e). He denies that he possesses this expertise himself (*Apology* 20c), but does not say that it is impossible that he, or any human being, should possess it.

This disavowal of expertise is not incompatible with the claim to know particular things. The nonexpert can know some particular things, but not in the way the expert knows them; specifically the nonexpert is not able, as the expert is, to relate his or her particular items of knowledge to a comprehensive system that provides explanations of their truth by relating them to other items of knowledge and to the system as a whole. But that raises the problem of the source of Socrates' nonexpert knowledge of moral truths. Usually, nonexperts know some particular things because they have been told by an expert, or because they have picked them up from some intermediate source whose authority is ultimately derived from that of the expert. But Socrates does not recognize any moral experts, among human beings at any rate. So what is the source of his nonexpert knowledge? The dialogues provide no clear or uniform answer to this question. Sometimes he suggests that the application of his critical method is sufficient, not merely to reveal inconsistency in his interlocutor's beliefs, but to prove that some are false, and hence that their negations are true. Thus at the end of the argument with Callicles in the *Gorgias* he claims (508e–509a) that the conclusion that it is always better to suffer wrong than to do it has been established by "arguments of iron and adamant" (i.e., of irresistible force), while conjoining that claim with a disavowal of knowledge: "I do not know how these things are, but no-one I have ever met, as in the present case, has been able to deny them without making himself ridiculous."

This presents a contrast between expert knowledge, which Socrates disclaims, and a favorable epistemic position produced by repeated application of Socrates' critical method of argument. There are some propositions that repeated experiment shows no one capable of denying without self-contradiction. While it is always theoretically possible that someone might come up with a way of escape from this position, realistically the arguments establishing those propositions are so firmly entrenched as to be irresistible. While it is an attractive suggestion that Socrates considers the moral truths that he nonexpertly knows to be of this kind, it receives no clear confirmation from the dialogues. There is, for instance, no indication in the *Crito* that Socrates' unshakable commitment to the fundamental principle that one must never act unjustly (49a) is based on critical examination of his and Crito's moral beliefs. It has to be acknowledged that while Socrates indicates that critical examination is sometimes capable of establishing truth beyond at least the practical possibility of rebuttal, and sometimes suggests that he knows some moral truths on the strength of good arguments for them, he gives no general account of the grounds of his nonexpert moral knowledge.

RELIGION. One might perhaps speculate that the source of Socrates' nonexpert moral knowledge is supposed to be divine revelation, but though Socrates' attitudes to the divine are an important element in his portrayal by both Plato and Xenophon, neither in fact suggests that Socrates believed that his moral beliefs were divinely inspired. What he did believe, according to both writers, is that throughout his life he was guided by a private sign or voice that he accepted, apparently without question, as being of divine origin, but the content of that guidance appears to have been, not moral principles, but day-to-day practical affairs, and it had the peculiar feature that its guidance was always negative, warning Socrates against some course of action that he might otherwise have undertaken (Plato, *Apology* 31c–31d). Thus Xenophon reports him (*Apology* 4) as explaining his failure to prepare his defense because the divine sign had told him not to, while in Plato's *Apology* (40a–40b) he says that he is confident that his conduct at his trial has been correct because the divine sign has not opposed it.

Such a claim to continuous private divine guidance (as opposed to occasional private revelations, e.g., in dreams) was certainly unusual, and, as suggested above, it is likely that it at least contributed to the charge of religious unorthodoxy that was one of the grounds of his condemnation. The actual stance of the historical Socrates toward conventional religion is not altogether easy to reconstruct from the sources. Xenophon, as pointed out above, stresses his conventional piety, as measured by public observance and private conversation; for example, his demonstration to an irreligious acquaintance of the providential ordering of the world, down to such details as the design of the eyelashes to shield the eyes from the wind (*Memorabilia* 1.4). On that account it is difficult to see how the charge of impiety could have been brought at all.

Plato's presentation is more complex. He does indeed represent Socrates as concerned on occasion with prophetic dreams (*Crito* 44a–44b; *Phaedo* 60e–61b) and with ritual, most famously in his report of Socrates' last words: "Crito, we owe a cock to Asclepius; pay it and don't forget" (*Phaedo* 118a). But it is notable that all of these instances arise in the context of Socrates' imminent death. When Plato represents Socrates as praying on various occasions throughout his life he almost always makes him pray for nothing but wisdom and virtue, while in his most extensive discussion of piety, in *Euthyphro*, he suggests that Socrates thinks that what the gods require from humans is nothing other than moral virtue. That fits well with his *Apology*, where Socrates' rebuttal of the charge of impiety has nothing at all to say about ritual, consisting wholly in the claim that Socrates' life has been the fulfillment of a divine mission to promote the welfare, identified with the moral virtue, of his fellow citizens.

Plato's view of Socratic religion seems then to be that the essence of service to the gods is moral virtue, and that ritual fills its proper role, as in Socrates' life and death, as a complement to the fulfillment of that primary task. If that reflects Socrates' own view, then it is possible that it was seen by conservatively minded contemporaries as presenting a radical challenge to traditional ideas of the relations between gods and humans, which were founded on the belief that divine favor and protection for individuals and the community were secured by performance of the appropriate prayers and rituals, and thereby as justifying his condemnation for neglecting the state religion in favor of a new religion of his own.

DEFINITIONS. In the procedure of enquiry sketched in (i) above, the search for general definitions is central. This arises naturally from Socrates' search for expertise; the expert knows about his or her subject, and according to Socrates the primary knowledge concerning any subject is precisely knowledge of what that subject is. The general pattern of argument in the dialogues is that some specific question about a subject—for example, how is one to acquire goodness—is problematic in the absence of an agreed conception of what that subject is. Hence before the problematic question can be pursued, the definition of the subject must first be sought. The problematic question may be of various kinds; it may be, as in the example above (from the *Meno*) how goodness as such is to be acquired, or how a specific virtue is to be acquired (courage in the *Laches*), or whether a virtue is advantageous to its possessor (justice in the *Republic*). The *Euthyphro* exemplifies another pattern; it is disputed whether a particular action, Euthyphro's prosecution of his father for homicide, is an instance of piety or holiness, and Socrates maintains that the question will be settled when, and only when, the definition of piety is arrived at. This pattern has given rise to the accusation that Socrates is guilty of the "Socratic fallacy" of maintaining that in general it is impossible to tell whether anything is an instance of a property unless one already possesses a general definition of that property.

That general position would be methodologically disastrous for Socrates, because his approved strategy for reaching a definition is to consider what instances of the kind or property in question have in common, and it is impossible to do that if a person has to know the definition before he or she can even identify the instances from which the definition is to be derived. In fact the argument of the *Euthyphro* does not involve that fallacy; even if it is granted that there are some disputed cases where the question "Is this an instance of F?" cannot be settled without answering the prior question "What is F?" it does not follow that there are no undisputed cases where instances of F can be recognized without a definition. In the *Hippias Major*, however, Socrates does argue (286c–286e) that people cannot tell whether anything is fine or beautiful (*kalon*) unless they know—that is, can give a definition of—what fineness or beauty is; so though the Socratic fallacy is not a pervasive defect of Socrates' argumentative method, there does seem to be one instance of it in the dialogues.

The question "What is F?" can itself be understood in various ways; it may be a request for an elucidation of the linguistic meaning of the term "F," or a request for a substantive account of what the property of F-ness consists in, including, where appropriate, the decomposition of a complex property into its components (e.g., goodness consists of justice, self-control, etc.) and explanatory accounts of properties (e.g., self-control consists of the control of the bodily appetites by reason). The practical nature of the questions that often give rise to the search for definitions suggests that the latter kind of definition is what is sought. Someone who wants to know how virtue is to be acquired will not be helped by a specification of the meaning of "virtue" as "a property contributing to overall success in life"; what they are looking for is precisely an account of what it is that constitutes or guarantees success in life. That is confirmed by the fact that *Laches*, *Meno*, and *Protagoras*, all of which start from the practical question of how either a specific virtue or goodness in general is to be acquired, converge on the suggestion that courage (in *Laches*) and goodness (in *Meno* and *Protagoras*) are identical with knowledge, which is itself

part of a substantive theory of the nature of goodness (see next section). It must, however, be acknowledged that Plato shows no awareness of the theoretical distinction between a purely conceptual definition and the kind of substantive account that is favored by the structure of the dialogues just mentioned. Even in the *Meno*, the dialogue in which definition is treated in the greatest detail, he gives model definitions of either kind without any explicit differentiation. Substantive accounts are favored over conceptual definitions by his practice, not in the light of any theoretical discrimination between the two.

ETHICS. The picture of Socrates as a nonexpert enquirer outlined above needs to be qualified to this extent, that in some dialogues, specifically *Protagoras*, *Gorgias*, and *Meno*, he is represented as arguing positively, though not conclusively, in favor of certain propositions that amount to at least the outline of a theory of human nature and of human good. The basic theses of this theory are:

> (1) Every agent has a single overall aim, the achievement of a completely satisfactory life for him or herself.
>
> (2) Knowledge of what constitutes such a life is both necessary and sufficient for the achievement of it.
>
> (3) Such a life consists in the practice of the virtues of justice, self-control, courage, and holiness, which are identical with one another in that they are the application to different kinds of situation of the fundamental virtue of knowledge (of what the good for humans is and how it is to be achieved).

Thesis 2 is the famous thesis that "Virtue is knowledge," from which together with thesis 1 follows the still more famous thesis that "No-one does wrong willingly" (the latter two often referred to as "The Socratic Paradoxes"). The idea expressed in the second paradox is that, because everyone necessarily has the single aim of achieving the best life for him or herself, any action that does not in fact promote that aim must be explained by the agent's mistaken belief that it does promote it. Socrates is thus the first of a succession of philosophers throughout the ages to deny the possibility of acting against one's better judgment (often ascribed to weakness of will); that position remains as controversial in the twenty-first century as it was in antiquity. The identification of the conventional moral and social virtues as applications of the fundamental knowledge of what the human good is (with the implication that the virtues are identical with one another, conventionally labeled "The Unity of the Virtues"), though central to the prototheory, is never adequately argued for. It is supported at *Crito* (47e) by an analogy between virtue of soul and health of body; justice and injustice are respectively the health and sickness of the soul. So, just as it is not worth living with a diseased and corrupted body, it is not worth living with a diseased and corrupted soul. But that is not an argument. Even granted that health is an intrinsically desirable and disease an intrinsically undesirable state, the crucial claims that justice is the health of the soul and injustice its disease require defense, not mere assertion.

Plato supplies some arguments in the *Gorgias*, but they are weak. Socrates first agues that successful tyrants, who manifest the extreme of injustice, do not get what they really want—that is, the best life for themselves—because their injustice is bad for them. The crucial argument for that conclusion (473a–475c) starts from the premise, conceded by Socrates' opponent Polus, that acting unjustly, while good (i.e., advantageous) for the agent, is disgraceful. It is next agreed that whatever is disgraceful is so either because it is unpleasant or because it is harmful. Because acting unjustly is clearly not unpleasant, it must therefore be harmful. Hence the life of injustice is harmful to the unjust agent. This argument fails because it ignores the relativity of the concepts of unpleasantness and harmfulness. To be acceptable the first premise must be read as "Whatever is disgraceful to anyone is so either because it is unpleasant to someone or because it is harmful to someone." From that premise it clearly does not follow that because injustice is not unpleasant to the unjust person it must be harmful to that person. It could be harmful to someone else, and its being so could be the ground of its being disgraceful to the unjust person (as indeed people ordinarily think).

Later in the dialogue (503e–504d) Socrates argues against Callicles that because the goodness of anything, such as a boat or a house, depends on the proper proportion and order of its components, the goodness of body and soul alike depend on the proper proportion and order of their components, respectively health in the case of the body and justice and self-control in the case of the soul. The analogy of health and virtue, simply asserted in the *Crito*, is here supported by the general principle that goodness depends on the organization of components, but that principle is insufficient to ground the analogy, because the proper organization of components is determined by the function, point, or aim of the thing that those components make up. So in order to know which organization of psychic components is the appropriate one for humans we need a prior conception of what our aims in life should be. One conception of these aims may indeed identify the optimum organization as that defined

by the conventional moral virtues, but another may identify as optimum a different organization, say one that affords the maximum scope to certain kinds of self-expression, as exemplified by a figure such as the Nietzschean Superman. Socrates provides no argument to exclude that possibility.

In addition to the failure to establish that virtue is always in the agent's interest, the prototheory is more deeply flawed, in that it proves to be incoherent. This emerges when we consider Proposition 2, "Virtue Is Knowledge," and ask what virtue is knowledge of. The answer suggested by *Meno* and *Protagoras* is that virtue is knowledge of the best life for the agent; given the standing motivation to achieve that life, knowledge of what it consists in will be necessary if one is to pursue it reliably, and sufficient to guarantee success in that pursuit. But that requires that the best life for the agent is something distinct from the knowledge which guarantees that one will achieve that life. "Virtue is knowledge of the best life for the agent" will be parallel to "Medicine is knowledge of health," and the value of that knowledge will be purely instrumental and derivative from the intrinsic value of the success in life which it guarantees. But Socrates, as we have seen, treats virtue as analogous, not to medicine, but to health itself, and hence as intrinsically, not merely instrumentally valuable. Virtue is not, then, a means to some independently specifiable condition of life which we can identify as the best life, well-being, or happiness (in Greek, *eudaimonia*); rather it is a constituent of such a life, and one of the most difficult questions about Socratic ethics is whether Socrates recognizes any other constituents. That is to say, for Socrates a life is worth living either solely or at least primarily in virtue of the fact that it is a life of virtue.

The incoherence of the prototheory thus consists in the fact that Socrates maintains both that virtue is knowledge of what the agent's good is and that it is that good itself, whereas these two theses are inconsistent with one another. It could indeed be the case both that virtue is knowledge of what the agent's good is and that the agent's good is knowledge, but in that case the knowledge which is the agent's good has to be a distinct item or body of knowledge from the knowledge of what the agent's good is. So if Socrates is to maintain that virtue is knowledge, he must either specify that knowledge as knowledge of something other than what the agent's good is, or he must abandon the thesis that virtue is the agent's good. There are indications in the dialogues that Plato was conscious of this difficulty. In the *Euthydemus* he represents Socrates as grappling inconclusively with the problem, and in the *Republic* he offers a solution in a conception of human good as consisting in a state of the personality in which the nonrational impulses are directed by the intellect, informed indeed by knowledge, but by knowledge not of human good, but of goodness itself, a universal principle of rationality. This conception retains from the prototheory the thesis that human good is virtue, but abandons the claim that knowledge is virtue, because virtue is not identical with knowledge but directed by it, the knowledge in question being knowledge of the universal good.

Protagoras may plausibly be seen as exploring another solution to this puzzle, because in that dialogue Socrates sets out an account of goodness whose central theses are (i) virtue is knowledge of human good, and (ii) human good is a life in which pleasure predominates over distress. Whether Socrates is represented as adopting this solution in his own person, or merely as proposing it as a theory that ordinary people and Sophists such as Protagoras ought to accept (a question on which there has been much dispute), it represents a way out of the impasse that blocks the prototheory, though not a way that Plato was himself to adopt. Having experimented with this solution, which retains the identity of virtue with knowledge while abandoning the identity of virtue with human good, he settled instead for the *Republic*'s solution, which maintains the latter identity while abandoning the former.

The prototheory is not strictly inconsistent with Socrates' disavowal of wisdom or expertise, because it is presented in outline only, not established by conclusive argument as expertise requires. But the presentation of Socrates as even a prototheorist has at least a different emphasis from the depiction of him simply as a questioner and generator of *aporiai*. This author believes that it is impossible to tell how much of this theory is Plato's own and how much was actually held by Socrates; that it was at least suggested to Plato by certain ideas that had emerged in Socrates' conversations seems highly likely, but we are not justified in asserting more than that.

LATER INFLUENCE

The prototheory just sketched was an important element in the development of Plato's own ethical theory, and via Plato on those of Aristotle and the post-Aristotelian philosophical schools. With the exception of the Epicureans, each of the main schools adopted Socrates as, in effect, a patron saint, stressing aspects of his thought and personality congenial to its particular philosophical standpoint; the skeptics, especially those in the Platonic

Academy, which was converted to skepticism by Arcesilaus just over a century after its foundation and remained skeptical for two centuries, stressed Socrates' disavowal of wisdom and the undogmatic character of his questioning technique. The Cynics, whose doctrines and way of life derived from Antisthenes, one of Socrates' associates, claimed to emulate the austerity of his lifestyle and to accept his doctrine that virtue is sufficient for happiness. Via the Cynics, Socrates became a major influence on Stoicism, which combined the Cynic doctrine that happiness consists in living according to nature with the doctrine that for rational beings the life according to nature is the life in accordance with rationality. Accepting the essentials of the prototheory outlined above they drew the conclusion that moral virtue is the only good, everything else being indifferent—that is, neither good nor bad. A particularly significant figure in the Stoics' canonization of Socrates is Epictetus, who adopted Socrates at the exemplar of the philosophical life and reproduced in his protreptic discourses features of Socratic method such as elenctic and inductive arguments.

The influence of Socrates was not confined to the ancient philosophical schools. The second-century Christian apologist Justin claimed him as a forerunner of Christianity, a characterization that was revived by Renaissance Neoplatonists such as Marsilio Ficino. In Medieval Islam he was revered, though not well understood, as a sage and a defender of (and martyr for) monotheism against idolatry. In the Enlightenment era he was appropriated by rationalists such as Voltaire as an exemplar of natural virtue and a martyr in the struggle of rationality against superstition. In the nineteenth century Hegel, Kierkegaard, and Nietzsche identified him as a central figure in developments in the history of philosophy to which their own respective theories responded, and in the last quarter of the twentieth century he was a major influence on the later thought of Foucault.

The perennial fascination of Socrates owes less, however, to any specific doctrines than to Plato's portrayal of him as the exemplar of a philosophical life—that is, a life dedicated to following the argument wherever it might lead, even when it in fact led to hardship, poverty, judicial condemnation, and consequent death. Plato's depiction of how Socrates lived for philosophy would in any case have made him immortal; his presentation of how he died for it has given him a unique status in its history.

See also Plato; Sophists; Xenophon.

Bibliography

ANCIENT SOURCES

Aristophanes. *Clouds*, edited by K. J. Dover. Oxford: Clarendon Press, 1968. Translated by B. B. Rogers. London and Cambridge, MA: Loeb, 1924.

Diogenes Laertius. *Lives of Eminent Philosophers: Diogenes Laertius*. Loeb Classical Library. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1925. All volumes in the Loeb Classical Library contain the text in the original language with facing English translation.

Ferguson, John. *Socrates: A Source Book*. London: Macmillan for the Open University Press, 1970. This book contains a comprehensive collection of passages of ancient works (in English translation) referring to Socrates.

Plato. *Platonis Opera*, edited by John Burnet. 5 vols. Oxford: Clarendon Press, 1902–1906. Often called the Oxford Classical Text. New edition of Vol. 1, edited by E. A. Duke et al. (New York: Oxford University Press, 1995). New edition of *Republic*, edited by S. Slings (New York: Oxford University Press, 2003).

All the dialogues are available in numerous English translations.

Xenophon. *Apology*. Loeb Classical Library, edited by O. J. Todd. Cambridge, MA: Harvard University Press, 1961.

Xenophon. *Memorabilia*. Loeb Classical Library, edited by E. C. Marchant. London: Heinemann, 1923. Greek with facing English translation.

Xenophon. *Symposium*. Loeb Classical Library, edited by O. J. Todd. Cambridge, MA: Harvard University Press, 1961.

Xenophon, Conversations of Socrates, Translated by Hugh Tredennick and Robin Waterfield. Harmondsworth: Penguin Classics, 1990.

Minor Socratic writers. *Socratis et Socraticorum Reliquiae*. Edited by Gabriele Giannantoni. 4 vols. Naples: Bibliopolis, 1990. English translation of the principal fragments of Aeschines in *Plato and His Contemporaries*, by G. C. Field, chap. 11. London: Methuen, 1930.

MODERN WORKS

Comprehensive Survey

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 3, pt. 2. Cambridge, U.K.: Cambridge University Press, 1969. Published separately in 1971 under title *Socrates*.

Socratic Literature and the Problem of the Historical Socrates

Burnet, John. *Greek Philosophy: Thales to Plato*. London: Macmillan, 1914.

Burnet, John. "The Socratic Doctrine of the Soul." *Proceedings of the British Academy* 7 (1915–1916): 235–260.

Döring, Klaus. "Sokrates, die Sokratiker und die von ihnen begründeten Traditionen." In *Grundriss der Geschichte der Philosophie. Die Philosophie der Antike 2/1: Sophistik, Sokrates, Sokratik, Mathematik, Medizin*, edited by Helmut Flashar, 139–364. Basel, Switzerland: Schwabe, 1998.

Kahn, Charles H. *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*. Cambridge, U.K.: Cambridge University Press, 1996.

Patzer, Andreas, ed. *Der historische Sokrates*. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft, 1987.

Rutherford, R. B. *The Art of Plato: Ten Essays in Platonic Interpretation*. London: Duckworth, 1995.
Taylor, A. E. "Plato's Biography of Socrates." *Proceedings of the British Academy* 8 (1917–1918): 93–132.
Taylor, A. E. *Socrates*. London: Peter Davies, 1932.
Taylor, A. E. *Varia Socratica*. Oxford: James Parker, 1911.
Taylor, C. C. W. *Socrates*. Oxford: Oxford University Press, 1998.
Vander Waerdt, Paul A. *The Socratic Movement*. Ithaca, NY: Cornell University Press, 1994.

Critical and Analytical Studies, Primarily on Plato's Presentation of Socrates

Benson, Hugh H. *Socratic Wisdom: The Model of Knowledge in Plato's Early Dialogues*. New York: Oxford University Press, 2000.
Brickhouse, Thomas C., and Nicholas D. Smith. *The Philosophy of Socrates*. Boulder, CO: Westview Press, 2000.
Brickhouse, Thomas C., and Nicholas D. Smith. *Plato's Socrates*. New York: Oxford University Press, 1994.
Irwin, Terence. *Plato's Ethics*. New York: Oxford University Press, 1995.
McPherran, Mark L. *The Religion of Socrates*. University Park: Pennsylvania State University Press, 1996.
Santas, Gerasimos Xenophon. *Socrates: Philosophy in Plato's Early Dialogues*. London: Routledge, 1979.
Vlastos, Gregory. *Socrates, Ironist and Moral Philosopher*. Cambridge, U.K.: Cambridge University Press, 1991.
Vlastos, Gregory. *Socratic Studies*, edited by Myles Burnyeat. Cambridge, U.K.: Cambridge University Press, 1994.

Works on the Trial of Socrates

Brickhouse, Thomas C., and Nicholas D. Smith. *Socrates on Trial*. Princeton, NJ: Princeton University Press, 1989; and Oxford: Clarendon Press, 1989. This work contains a useful guide to modern literature on Socrates.
Brickhouse, Thomas C., and Nicholas D. Smith, eds. *The Trial and Execution of Socrates: Sources and Controversies*. New York: Oxford University Press, 2002.
Hansen, Mogens Herman. *The Trial of Sokrates: From the Athenian Point of View*. Copenhagen: The Royal Danish Academy of Sciences and Letters, 1995.
Stone, I. F. *The Trial of Socrates*. London: Jonathan Cape, 1988.

Collections of Articles

Benson, Hugh H., ed. *Essays on the Philosophy of Socrates* New York: Oxford University Press, 1992.
Gower, Barry S., and Michael C. Stokes, eds. *Socratic Questions: New Essays on the Philosophy of Socrates and Its Significance*. London: Routledge, 1992.
Prior, William J., ed. *Socrates: Critical Assessments*. 4 Vols. London: Routledge, 1996.
Smith, Nicholas D., and Paul B. Woodruff, eds. *Reason and Religion in Socratic Philosophy*. New York: Oxford University Press, 2000.
Vlastos, Gregory, ed. *The Philosophy of Socrates: A Collection of Critical Essays*. Garden City, NY: Doubleday, 1971.

Later Influence

Fitzpatrick, P.J., "The Legacy of Socrates." In Gower and Stokes, *Socratic Questions*.
Long, A. A. "Socrates in Hellenistic Philosophy." *Classical Quarterly* 38 (1988): 150–171.
Montuori, Mario. *Socrates: Physiology of a Myth*. Translated by J. M. P. Langdale and M. Langdale. Amsterdam: J. C. Gieben, 1981. Original Italian edition, Florence: G. C. Sanzoni, 1974.
Nehamas, Alexander. *The Art of Living: Socratic Reflections from Plato to Foucault*. Berkeley: University of California Press, 1998.

C. C. W. Taylor (2005)

SOLGER, KARL WILHELM FERDINAND

(1780–1819)

Karl Wilhelm Ferdinand Solger, the German romantic philosopher, was born in Schwedt. He studied jurisprudence, philology, and philosophy at the University of Halle and at Jena, where he heard Friedrich von Schelling lecture. After some time in the Prussian civil service, he lectured on philosophy at the University of Frankfurt an der Oder (1809), where he met Ludwig Tieck, the writer. From 1811 until his death he was a professor at the University of Berlin.

Like many romantics, Solger was preoccupied with the polarity of the finite and the infinite. Man is finite but filled with a desire for the infinite. The world in which he finds himself is fragmented. Grasping splinters of reality, common understanding operates in terms of polarities—concrete and universal, appearance and concept, body and soul, individual and nature. Only in the infinite Idea are polarities reconciled. Common understanding is tied to the finite. Man must escape from its rule if he is to recognize the infinite Idea. God made a sacrifice of himself to create the finite, and man must sacrifice himself and the phenomenal to return to the infinite. In this annihilation the Godhead reveals itself. The reconciliation of the finite and the infinite is the goal of the philosopher when he tries to capture truth in his systems; it is the duty of the moral man who confronts it as a task; it is achieved by the artist who, in creating the beautiful, reveals the Idea in the phenomenal.

The philosophy of art was at the center of Solger's philosophical program. Enthusiasm and irony are the two mainsprings of artistic creation. Enthusiasm, like Plato's Eros, ties man to the reality in which he has his ground. The enthusiast is possessed by the Idea. Irony recognizes the negativity of phenomenal reality and negates it. Thus it pushes away the veil that normally hides the Idea from us.

For Solger, as for Plato, philosophy is fundamentally conversation. It is a joint struggle for something that is

dimly apprehended and yet escapes adequate articulation. Truth is never a possession; it only reveals itself in the process of striving for it. Thus, the most adequate vehicle for the expression of philosophical thought is the dialogue.

See also Plato; Romanticism; Schelling, Friedrich Wilhelm Joseph von.

Bibliography

WORKS BY SOLGER

Erwin, Vier Gespräche über das Schöne und die Kunst. 2 vols. Berlin: In der Realschulbuchhandlung, 1815.

Philosophische Gespräche. Berlin, 1817.

Nachgelassene Schriften und Briefwechsel, edited by L. Tieck and F. von Raumer. Leipzig, 1826.

Vorlesungen über Aesthetik, edited by K. W. L. Heyse. Leipzig, 1829.

The Complete Correspondence (between Solger and Ludwig Tieck). Edited by P. Matenko. New York: Westermann, 1932.

WORKS ON SOLGER

Allemann, Beda. *Ironie und Dichtung.* Pfullingen, Germany: Neske, 1956.

Boucher, Maurice. *K. W. F. Solger, ésthetique et philosophie de la présence.* Paris: Stock, 1934.

Hartmann, Hans. *Kunst und Religion bei Wackenroder, Tieck und Solger.* Erlangen, 1916. A dissertation.

Hegel, G. W. F. *Sämmtliche Werke.* Vol. XII, pp. 100–106, 221; Vol. XX, pp. 132–202. Stuttgart, 1927–1930.

Heller, J. E. *Solgers Philosophie der ironischen Dialektik.* Berlin, 1928.

Kierkegaard, Søren. "Om Begrebet Ironi." *Samlede Vaerker,* 15 vols. Vol. XIII, pp. 376–387. Copenhagen, 1901–1906.

Müller, Gustav. "Solger's Aesthetics—A Key to Hegel." *Corona, Studies in Celebration of the 80th Birthday of Samuel Singer.* Durham, NC, 1941.

Schönebeck, Erich. *Tieck und Solger.* Berlin: H. Blanke's Spezial-druckerei für Dissertationen, 1910. A dissertation.

Walzel, Oskar. "'Allgemeines' und 'Besonderes.'" *Deutsche Vierteljahrsschrift* 17 (1939): 152–182.

Wildbolz, Rudolf. *Der philosophische Dialog als literarisches Kunstwerk.* Bern and Stuttgart, 1952.

Karsten Harries (1967)

SOLIPSISM

There are a number of importantly different views associated with the term *solipsism*. Its Latin roots—*solus*, meaning "alone," and *ipse*, meaning "self"—suggest the rough idea that a solipsistic doctrine is going to put some sort of emphasis on the self standing alone, but there are radically different ways in which a philosopher might develop that emphasis. In particular, we must distinguish an extreme metaphysical thesis, a view about the nature of mental states (sometimes misleadingly referred to as methodological solipsism), an epistemological/methodological thesis, and an ethical thesis.

METAPHYSICAL SOLIPSISM

The simplest and most radical of doctrines associated with solipsism is the puzzling doctrine that only the self exists. Stated in these terms, the doctrine is scarcely intelligible. The obvious question concerns whose self precisely it is that is supposed to be the only existing thing. It is easiest to state the doctrine from the first-person perspective. If I embrace solipsism, I am endorsing the view that I am the only existing thing. If you embrace solipsism, then you are endorsing the view that you are the only existing thing. If we both endorse solipsism, therefore, then we are both wrong. In asserting solipsism, the solipsist is usually not trying to deny the existence of *properties* exemplified by the self. So the self that exists may believe, fear, hope, plan, and so on. We can also distinguish the solipsist who intends only to deny the existence of other minds from the solipsist who denies the existence of all other objects, for example, physical objects. It would be odd, however, to hold the former without the latter for, as we shall see, the epistemological position that drives one to a skepticism about other selves often involves a skepticism with respect to the external world.

There is almost a comical aspect to the most extreme form of solipsism. It is certainly odd to hear any philosopher defending (to whom?) the view. One could certainly never take comfort in the fact that one succeeded in convincing anyone of the truth of the view. But in this respect solipsism is probably no worse off than any other extreme form of skepticism—say skepticism with respect to the past, the future, or the external world. In fact, solipsism is probably a view that one starts to take seriously precisely in the context of more general epistemological concerns. So, for example, while Descartes was no solipsist, he came perilously close to painting himself into a solipsistic corner.

In the *Meditations*, Descartes famously sought secure foundations for knowledge. To find those foundations he employed what is sometimes called the method of doubt. He tried to strip from his belief system all those beliefs that admit of the possibility of error. So, for example, he thought that no belief about the physical world belongs in the foundations of knowledge because our evidence for believing what we do about that world never gets any better than vivid sense experience. But the kind of sense

experience upon which we must rely is always compatible with our dreaming, or our being the victims of massive demon-induced hallucination. Since our knowledge of the existence of other people seems to rest critically on our knowledge of other bodies, skepticism with respect to the physical world might seem to entail a skepticism with respect to the existence of other selves. After rejecting a number of candidates for foundational truth, Descartes finally hit upon his own existence as one truth that he could not rationally doubt. No matter how hard he tried to convince himself that he did not exist, such efforts merely reinforced for him the fact that he did exist. One can only doubt one's own existence if one exists to do the doubting. "Cogito, ergo sum," Descartes concluded—I think, therefore I am.

While the exact nature of the evidence or justification to which Descartes appeals in claiming foundational knowledge of his own existence is a matter of some controversy, his attempt to begin a reconstruction of the rest of what he knows from this foundation is one that could have easily led him to a solipsistic conclusion. Descartes thought that he could find a way of legitimately inferring the rest of what he believes from knowledge of his own thoughts and experiences, but it is an understatement to suggest that his efforts did not meet with universal acceptance. Indeed, many contemporary philosophers are convinced that if we restrict ourselves to premises describing our own existence and the conscious states exemplified there, there is no path to knowledge of, or even justified belief in, the rest of what we commonsensically think we know.

The kind of radical foundationalism that Descartes embraced might naturally lead, then, to the conclusion that we can only know of our own existence and the perceptions and thoughts that reside there. And if one restricts one's metaphysical positions to what is licensed by knowledge, then one might be left affirming only one's own existence. Again, that claim is usually expanded, even by the solipsist, to include the conscious mental states exemplified by that self. When we discuss epistemological solipsism, we will say more about the epistemological assumptions that might lead one to take seriously the position of metaphysical solipsism. But let us first examine some influential criticisms leveled at the view.

CRITICISMS. One charge often leveled against metaphysical solipsism is the charge of self-refutation. There are a number of different ways in which a view might be self-refuting. The strongest form of self-refutation is logical—a self-refuting view entails that it is itself false. So, for example, the proposition that all claims are false is self-refuting in this sense. The claim entails its own falsehood. On the face of it, it is difficult to see how the solipsist's claim can be self-refuting in this way. Nevertheless, critics have claimed that for the solipsist's claim to be *meaningful* it must be false. Inspired by Wittgenstein, for example, some philosophers claim that language and meaning are essentially social; there can be no such thing as a private language or a private linguist.

Unfortunately, it is by no means easy to figure out just what the basis for this claim is. One crude characterization of the argument emphasizes the importance of rules in determining meaning. One uses a term meaningfully only if one uses it in accord with a rule that determines when the term is used correctly or incorrectly. If one is the sole arbiter of when a term is used correctly, the argument goes, one will be unable to make a mistake. But if one cannot make a mistake using the term, then it makes no sense to suppose that one is using the term correctly; correct use makes sense only against the possibility of incorrect use. It is only when there is a community of language users that one can understand the distinction between correct and incorrect use of language; incorrect use can be identified with divergence from standard or common use.

So to illustrate with an example, suppose that I see a creature I have never seen before and resolve to call it and anything relevantly like it a "gretl." One might initially suppose that I have successfully introduced a word into my own private language. Tomorrow, I see another creature—somewhat like the first, but also in many ways dissimilar. Is it a "gretl" or not? It seems that if I am the only one deciding whether it is enough like the first creature to count as a "gretl," then I cannot get it wrong—whatever I decide goes. Again, the Wittgensteinian will claim that where there is no possibility of error, there is no possibility of truth.

There are no uncontroversial interpretations of the private language argument, and a full evaluation of it would take us far afield. All versions of the argument, however, rest on highly controversial assumptions. It is not clear, for example, that judgment involves comparison to a paradigm. In any event, one must surely worry that the version stated above would rule out even the possibility of a *solitary* linguist—a sole language user. But it is hard to see how it could be impossible for there to exist one and only one person who was capable of both thought and language. We can imagine, for example, an infant who is the sole human survivor of a worldwide natural disaster and who, adopted by apes, somehow

manages to mature into an adult. In such a world, if that human being could formulate the thought that there are no other people, he or she would have formulated a true thought. And do we really want to argue that in the situation described it would be metaphysically impossible for the person to formulate either the thought or a language that could express the thought? If we reflect on the scenario just described, we might become suspicious of any argument that purports to show that the solipsist's doctrine that there exists only one self is in some sense unintelligible or necessarily false.

There is a more recent philosophical claim about the nature of thought that, like the earlier arguments inspired by Wittgenstein, might seek to cast doubt on the intelligibility of metaphysical solipsism. It is sometimes called semantic, psychological, or content externalism. The basic idea behind the view is encapsulated in Putnam's famous slogan that meanings are not "in the head," and its proponents sometimes seem to claim that one can only have thoughts about certain kinds of things if those kinds of things exist. If the view were true, one might be able to infer from the fact that one can form thoughts about physical objects (even the thought that there are no physical objects) that physical objects exist. Similarly, one might be able to infer from the fact that one can form thoughts about other people (even the thought that there are no other people) that other people exist. The view underlying this criticism of metaphysical solipsism is held in opposition to another thesis associated with solipsism, a thesis sometimes called "methodological solipsism."

METHODOLOGICAL SOLIPSISM

The term *methodological solipsism* was introduced by Hilary Putnam and made more familiar by Jerry Fodor. It is precisely the view rejected by the content externalist. The methodological solipsist (or internalist in the philosophy of mind) is convinced that psychological states (beliefs, desires, fears, pains, etc.) are entirely constituted by internal features of the person in those states. Two people cannot be in identical internal states while one of them has a certain desire, say, and the other does not. The externalist argues, somewhat paradoxically perhaps, that at least some of the conditions that constitute or determine your psychological states are factors that lie outside you—factors that include, for example, the causal origin of your internal states. So Putnam famously argued that two people could be in precisely the same internal states while one is thinking about water (the stuff with molecular structure H_2O) and the other is thinking about "twater" (something with an entirely different molecular structure). The difference in the content of their thoughts would be a function of the environments in which the respective internal states arise. In a much-discussed attempt to extend these considerations to issues involving skepticism, Putnam (1981) appeared to argue that if one were a brain in a vat whose experiences were produced by the machinations of some mad neurophysiologist, one could not even entertain that hypothesis. His idea is that without some sort of sensory interaction with the physical world, one could not even form a thought that was about a physical object like a brain or a vat.

If such an argument were successful it would not be hard to extend it as an attack on the intelligibility of the more extreme forms of metaphysical solipsism. When the solipsists make clear their views about what does not exist, their ability to form the thought, for example the thought that there is no external world, presupposes that there is one. Without interaction with external reality, no thought could be about such reality and one thus could not even coherently deny its existence. Since skepticism about the existence of others typically runs through skepticism about the external world, one will have undercut an argument for solipsism.

CRITICISMS. Content externalism is no less controversial than the various presuppositions Wittgenstein and others brought to their philosophical views about meaning. But even if we grant some of the basic tenets of the externalist's conception of the conditions necessary for thought, *careful* statements of the view will not take one very far toward interesting metaphysical conclusions about what there is. For one thing, the careful content externalist is going to radically restrict the view to a subclass of thoughts. No one thinks, for example, that in order for one to have thoughts about mermaids, one must have interacted in some way (or be connected with someone else who has interacted in some way) with actual mermaids. The most natural move, borrowed from the earlier empiricists who thought that all ideas are "copies" of prior impressions, is to make a distinction between complex ideas and simple ideas. The earlier empiricist conceded that the idea of a mermaid is not a copy of some prior impression or experience of a mermaid, but went on to claim that the idea is complex (the idea of woman's torso combined with a fish's tail), and the ideas out of which the complex idea is composed *are* copies of prior impressions. Of course, the idea of a torso itself might be complex, composed of still simpler ideas. The natural thought for both the earlier empiricist and

the content externalist is to restrict their thesis to the simple ideas that are the "building blocks" of other ideas.

The difficulty is that it is not clear what the best candidates are for the simple ideas out of which others are built. Suppose, for example, that I have the idea of a sensation. I might also be able to form the idea of causation. I can put those two ideas together to form the idea of that which causes the sensation. Arguably, in this way I can form the complex idea of an external object. But I have formed the idea in such a way that it might not correspond to anything—the sensation in question might have no cause. There seems to be nothing in the externalist's view that blocks the possibility of forming thoughts of this sort, thoughts that might well not correspond to anything. Consequently, it is not at all clear that the metaphysical solipsist would face any problems of self-refutation in framing various radical views about what does not exist.

EPISTEMOLOGICAL SOLIPSISM

The first two theses discussed above are metaphysical claims—claims about what exists. As we have just seen, one use of the expression "methodological solipsism" involves a claim about the nature of mental states. There was, however, an earlier use of the term "methodological solipsism" (by Hans Driesch, Rudolph Carnap, and others) expressing an epistemological thesis. Indeed, that earlier use of the expression is a much more natural way to describe what these philosophers had in mind—a *method* for arriving at truth. To avoid confusion, it is best to describe the view that I have in mind as epistemological solipsism.

The fundamental idea behind epistemological solipsism is the claim that in reaching conclusions about what exists, each of us is restricted to a foundation of knowledge about our own mental states. The foundationalist in epistemology is convinced that there must be some truths that are known or justifiably believed without their needing to be inferred from other different truths that are known or justifiably believed. This foundational knowledge is needed to block a vicious epistemological regress. To justifiably believe *P* by inferring it from *E1*, one would need, the argument goes, justification for believing *E1*. Some would argue that one would also need justification for believing that *E1* confirms *P*. But if the only way to justifiably believe something is to infer it from something else, then to justifiably believe *E1*, one would need to infer it from something else *E2*, which one would need to infer from something else *E3*, and so on, ad infinitum. Finite minds cannot complete infinitely long chains of reasoning. It is not even clear that infinite minds can complete infinitely long chains of reasoning. So if we are to justifiably believe anything at all, some of our beliefs must be non-inferentially justified—justified without inference.

The radical empiricist/epistemological solipsist is convinced that the only contingent truths that one can know without inference are truths about one's own existence and the thoughts and experiences contained there. Arguments for restricting the foundations of knowledge in this way depend, typically, on specific presuppositions about the nature of foundational knowledge. As we saw earlier, Descartes sought foundations in beliefs that are infallible. If one's justification for believing something is compatible with the belief's being false, then the belief is not a candidate for non-inferential knowledge. The radical empiricist was convinced that beliefs about the external world, the past, other minds, and the future, all fail this test for foundational knowledge. By contrast, one's beliefs that one exists, that one is in pain (when one is), that one has thoughts, all were supposed to pass the test.

A closely related version of foundationalism seeks to identify foundational knowledge with belief accompanied by direct acquaintance with facts that are the truth-makers for the belief. On this view, when one is in pain, for example, one's non-inferential justification consists in the fact that the pain itself is directly present to consciousness. Again, the claim is that objects in the physical world, other minds, facts about the past, and facts about the future, are never directly presented to consciousness in this way. Their existence must be inferred from what is known directly about present conscious states.

The epistemological solipsist's position was probably almost taken for granted by most prominent philosophers in the history of philosophy. The task of the philosopher is essentially egocentric. If one is to avoid begging questions, one has no choice but to begin one's search for truth with the various ways that things appear. This epistemological position does not entail metaphysical solipsism, but as we saw, there is the danger that one will be unable to reason oneself out from behind this "veil" of subjective appearance.

CRITICISMS. The version of foundationalism endorsed by the epistemological solipsist has come under sustained attack in the last several decades. In discussing metaphysical solipsism, we have already had occasion to examine Wittgenstein's worries about the possibility of a private language. To the extent that judgment involves categorizing things, categorizing things involves appeal to the correctness of following certain rules, and knowledge of

what rules sanction involves facts about communities of rule followers, one will have difficulty finding the kind of foundations sought by the epistemological solipsist. But as we saw, this criticism of private language and thought is by no means uncontroversial.

In the previous section, we also discussed a view about the nature of mental states that might also cast doubt on the radical version of foundationalism endorsed by many empiricists. If external reality is literally partially constitutive of mental states like belief, then it might seem to follow that our knowledge of mental states could be no more secure than the knowledge of that external reality upon which their content depends. That this follows from psychological externalism, however, is a matter of great dispute, and among those who take it to be an implication of externalism in the philosophy of mind, there are many who take this consequence of the view to be a *reductio* of the view. In any event, as we also noted, psychological externalism is no more uncontroversial than the presuppositions of Wittgenstein's argument against the possibility of knowing truths about a "private" experience.

There are other efforts to cast doubt on the claim that empirical knowledge begins (and perhaps even ends) with knowledge of one's inner mental states. In a famous attack on the radical empiricist's doctrine of what is "given," Wilfred Sellars (1963) claimed that it is an illusion to suppose that we can form thoughts about appearances that are independent of thoughts about objective reality. So suppose, for example, that the epistemological solipsist claims to know that something looks red to him, or that it appears as if something is red. That epistemological solipsist claims that knowledge that there actually is a physical object that is red is more tenuous, less secure, than knowledge of the subjective appearance presented by such an object. But Sellars wants to know precisely what it means to say that it looks as if something is red. Sometimes we use "seems"/ "appears" language to indicate tentative belief—R. M. Chisholm (1957) called this the epistemic use of "appears." But in its epistemic sense, the judgment that it appears as if *X* is red is just the tentative judgment that *X* is red—it is not a truth about an appearance to which one might appeal as evidence for the claim that there exists before one a red object.

There is another use of "appears," however—the comparative use. But it will not be of any use to the philosopher intent on restricting a knowledge claim to subjective experience. In the comparative sense, to judge that it appears to me as if *X* is red is just to judge that I am having the kind of experience that is usually caused by red things under normal conditions. It takes but a moment's reflection to realize that this thought about how things appear is not a thought confined to subjective reality at all. To know that it looks as if something is red, I would have to know something about objective reality—I would have to know how red things look under normal conditions—something that presupposes that I have had epistemic access to how things have been, not just how things appear.

If the only way that we could conceptualize experience was comparatively in the above sense, then it would be folly to suggest that our knowledge of reality begins with knowledge of subjective appearance. But, of course, it is not difficult to see how the epistemological solipsist should respond to the above criticism. The very characterization of the comparative use of "appears" seems to make reference to a "way" that red things look and the radical empiricist/epistemological solipsist thinks that we have no difficulty conceptualizing that way in terms of its intrinsic character. However the word "appears" is normally used, the epistemological solipsist can borrow that term to describe what Chisholm called the noncomparative intrinsic character of the experience (1957).

There are countless other attacks on the radical foundationalist's idea that all empirical knowledge rests on a foundation of knowledge about the character of subjective experience. Some, for example, argue that we must reject such a view because it will ultimately lead to a radical skepticism—perhaps even the metaphysical solipsism discussed earlier. The charge is that the foundations countenanced by such a view coupled with available epistemic principles simply will not allow us to get back the knowledge that we commonsensically take ourselves to have. To determine whether epistemological solipsism does lead to skepticism would take us too far afield, but one might wonder whether a commitment to the falsity of skepticism should rule philosophical thought.

Still others complain that the epistemological solipsist radically overintellectualizes the nature of our thought about external reality. Not only do we not *always* infer objective reality from subjective experience, we rarely pay attention to how things appear. As anyone who has tried to paint soon realizes, it takes a certain amount of learning and sophistication to see the world as it appears instead of as we take it to be objectively. But it is not clear what relevance this observation has for the epistemological solipsist's central thesis. To be sure, if the solipsist makes a claim about what we actually do know, the above observations might cast doubt on that claim by casting doubt on the question of whether we typically

form the required thoughts. But the careful epistemological solipsist might make a claim only about the possibility of knowledge. That epistemological solipsist might argue that whatever we think we know or justifiably believe, the only truths that *can* be known, at least directly and without inference, are truths about the character of subjective experience.

Just as there is an internalism/externalism controversy concerning the nature of mental states, so also there is an internalism/externalism controversy in epistemology. Many epistemological externalists argue that whether or not a belief is justified depends critically on the causal history of the belief—the way in which the belief was produced. Alvin Goldman (1979) advances one version of such a view—reliabilism. The reliabilist is a kind of foundationalist but argues that foundationally justified beliefs are just beliefs produced by reliable belief-producing processes that take as their input something other than beliefs. A belief-producing process of this sort is reliable when it does or would produce mostly true beliefs. Reflection on the reliabilist's criterion for non-inferential justification reveals that there can be no *a priori* restrictions on which beliefs might turn out to be non-inferentially justified. Against the traditional foundationalists, the reliabilist will argue that non-inferential justification has nothing to do with infallibility. A belief can be non-inferentially justified if it is just barely more likely to be true than false. Beliefs about the past, the physical world, and other minds all might be non-inferentially justified according to the reliabilist. Whether they are or not depends on empirical facts about the way in which such beliefs are caused.

It is certainly true that arguments for epistemological solipsism are challenged by contemporary versions of epistemological externalism. It is hardly the case, however, that philosophers agree on the success of externalist analyses of epistemic concepts. The epistemological solipsists have an array of weapons ready to deploy against the externalist. But underlying their criticism is often the common theme that the externalist's analysis of epistemic concepts has stripped them of their philosophical interest. The epistemological solipsist is likely to be convinced that satisfying the externalist's epistemic concepts does nothing to provide assurance of the sort the philosopher seeks. I may have a reliably produced belief. I may be evolutionarily programmed to believe reliably various truths about the world around me. But unless I have some reason to believe that the way in which my beliefs are formed is reliable, the mere fact of reliability does nothing to give me the kind of assurance I was looking for when I was interested in having justified beliefs.

EGOISM

Another quite different sort of doctrine that might be associated with the idea of the self standing alone is the ethical theory or theory of rational behavior known as egoism. A crude version of the theory is that rational people have only one goal or end in acting—their own happiness or well-being. Egoists can certainly take into account the well-being of others but only insofar as they have some reason to believe that the well-being of others impacts their own well-being.

Like other versions of solipsism, egoism has been accused by some of internal incoherence. As a theory, one argument goes, egoism must enjoin everyone to achieve his or her own well-being. But we can easily imagine a case in which my doing *X* maximizes my well-being, while *R*'s preventing me from doing *X* will maximize *R*'s well-being. I cannot coherently recommend, exhort, or want *R* to prevent me from doing what I want to do.

The above criticism presupposes that a principle of morality or rationality must be universalizable in certain respects. More specifically, it presupposes that if someone accepts the principle that everyone ought to seek only his or her own well-being, that commits that person to recommending such behavior to others, or acquiescing in such behavior on the part of others, or wanting others to behave in such ways. Such presuppositions are not uncontroversial even in the domain of morality, but are arguably downright implausible if the "ought" judgment in question is intended to assert only the rationality of egoistic behavior. There seems nothing at all inconsistent in my believing that it would be rational for all people to act egoistically while encouraging them not to so act and doing what I can to prevent them from acting that way. I know all too well what people ought to do to beat me in a game of tennis, but I never advise them concerning how to do it; I never want them to do it, and I do whatever I can to thwart them from behaving as they ought to.

As an ethical theory, the plausibility of egoism might in the end depend on metaphysical issues concerning the nature of ethical properties. G. E. Moore (1912) famously argued that if my happiness is objectively good, it is so in virtue of the property of being happy that I exemplify. But if objective goodness "supervenes" in this sense on the property of being happy, then it supervenes on that property no matter whose happiness we are talking about. An ethical egoist cannot recognize the goodness of his or her own happiness without recognizing the value

inherent in another person's being happy. But most egoists are not objectivists about value. On one view, diametrically opposed to ethical objectivism, something has intrinsic value for a person *S* only insofar as *S* subjectively values that thing for its own sake. And it is just a brute fact about most human beings, the egoist claims, that people care more about their own happiness than they do about the happiness of others.

That alleged empirical truth, however, is not uncontroversial. It might not be all that difficult for most parents, for example, to conclude that they value intrinsically the happiness of their children—perhaps even more than they value their own happiness. If they do, and if subjective valuing confers intrinsic value on that which is valued, then the egoist's view that rational people concern themselves only with their own well-being is implausible. It is worth noting, however, that the view according to which a thing's intrinsic value for a person is determined by that person's valuing it is itself a kind of solipsistic view. It is not egoism, because we might find ourselves valuing intrinsically the well-being of others, but it is still a view that makes the individual person the creator of the goals or ends that partially define for that person how life ought to be lived.

See also Augustine, St.; Ayer, Alfred Jules; Bradley, Francis Herbert; Bridgman, Percy William; Broad, Charlie Dunbar; Carnap, Rudolf; Descartes, René; Driesch, Hans Adolf Eduard; Egoism and Altruism; Epistemology; Fichte, Johann Gottlieb; Hamilton, William; Hume, David; Kant, Immanuel; Lewis, Clarence Irving; Locke, John; Mach, Ernst; Malcolm, Norman; Mill, John Stuart; Moore, George Edward; Other Minds; Pastore, Valentino Annibale; Private Language Problem; Royce, Josiah; Russell, Bertrand Arthur William; Santayana, George; Schiller, Ferdinand Canning Scott; Schuppe, Ernst Julius Wilhelm; Stace, Walter Terence; Stebbing, Lizzie Susan; Wittgenstein, Ludwig Josef Johann.

Bibliography

Ayer, A. J. *Privacy*. London, 1959; also in *Proceedings of the British Academy* 45 (1959): 43–65.

Ayer, A. J. *Problem of Knowledge*. London: Macmillan, 1956.

Bradley, F. H. *Appearance and Reality*. London: Swan Sonnenschien, 1893.

Carnap, Rudolf. *Der Logische aufbau der Welt*. Berlin: Weltkreis, 1928.

Carnap, Rudolf. "Die physikalische Sprache als Universalsprache der Wissenschaft." *Erkenntnis 2* (1932). Translated by Max Black as *The Unity of Science*. London: K. Paul, Trench, and Trubner, 1934.

Chisholm, R. M. *Perceiving*. Ithaca, NY: Cornell University Press, 1957.

Dretske, Fred. *Naturalizing the Mind*. Cambridge, MA: MIT Press, 1995.

Feigl, H. "Physicalism, Unity of Science and the Foundations of Psychology." In *The Philosophy of Rudolf Carnap*, edited by Paul A. Schilpp, 227–267. La Salle, IL: Open Court, 1963–1964.

Fodor, Jerry A. "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology." *The Behavioral and Brain Sciences* 3 (1) (1980): 63–72.

Fumerton, Richard. *Metaepistemology and Skepticism*. Lanham, MD: Rowman & Littlefield, 1995.

Fumerton, Richard. *Reason and Morality: A Defense of the Egocentric Perspective*. Ithaca, NY: Cornell University Press, 1990.

Gallie, W. B. "Solipsistic and Social Theories of Meaning." *PAS* 38 (1937–1938): 61–84.

Gauthier, David, ed. *Morality and Rational Self-Interest*. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Goldman, Alvin. "What Is Justified Belief?" In *Justification and Knowledge*, edited by George Pappas, 1–23. Dordrecht: Reidel, 1979.

Kalin, Jesse. "In Defense of Egoism." In *Morality and Rational Self-Interest*, edited by David Gauthier, 64–87. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Kripke, Saul A. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Lewis, C. I. *An Analysis of Knowledge and Evaluation*. La Salle, IL: Open Court, 1946. See Book II, especially chap. 7, sec. 11; chap. 9, sec. 8; chap. 11, sec. 5.

Ludlow, Peter, and Norah Martin, eds. *Externalism and Self-Knowledge*. Stanford, CA: CSLI Publications, 1998.

Malcolm, Norman. "Knowledge of Other Minds." *Journal of Philosophy* 55 (1958): 969–978.

Medlin, Brian. "Ultimate Principles and Ethical Egoism." *Australian Journal of Philosophy* 35 (1957): 111–118.

Moore, G. E. *Ethics*. London: Oxford University Press, 1912.

Moore, G. E. *Philosophical Papers*. London: Allen & Unwin, 1959. Especially "Defense of Common Sense" (1925) and "Proof of an External World" (1939).

Putnam, Hilary. "The Meaning of 'Meaning.'" In *Minnesota Studies in the Philosophy of Science*. Vol. 7, edited by K. Gunderson, 131–193. Minneapolis: University of Minnesota Press, 1975.

Putnam, Hilary. *Reason, Truth, and History*. Cambridge, U.K.: Cambridge University Press, 1981.

Rhees, Rush. "Can There Be a Private Language?" *PAS*, supp. vol. 28 (1954): 63–94. Symposium with Ayer.

Russell, Bertrand. *Human Knowledge*. New York: Simon & Schuster, 1948. See Part III, chap. 2.

Russell, Bertrand. *Our Knowledge of the External World*. Chicago: Open Court, 1914. Lectures III and IV.

Schlick, Moritz. "Meaning and Verification." *Philosophical Review* 45 (1936): 339–369. Reprinted in *Readings in Philosophical Analysis*, edited by H. Feigl and W. Sellars. New York: Appleton, 1949. See especially Sec. 5.

Sellars, Wilfred. *Science, Perception, and Reality*. London: Routledge and Kegan Paul, 1963.

Strawson, P. F. *Individuals.* Garden City, NY: Doubleday, 1963. See Part I, chap. 3.

Wisdom, John. *Other Minds.* Oxford: Blackwell, 1953.

Wittgenstein, Ludwig. *Tractatus Logico-philosophicus.* London: Routledge and Kegan Paul, 1922. See passages 5.6–5.641. Translated by D. F. Pears and B. F. McGuinness. London: Routledge and Kegan Paul, 1961. A survey of useful comments on the relevant passage in this work will be found in Max Black, *A Companion to Wittgenstein's Tractatus* (Cambridge, U.K.: Cambridge University Press, 1964), 307–311.

Wright, Crispin, Barry C. Smith, and Cynthia Macdonald, eds. *Knowing Our Own Minds.* Oxford: Clarendon Press, 1998.

Richard A. Fumerton (2005)

SOLOV'ËV (SOLOVYOV), VLADIMIR SERGEEVICH
(1853–1900)

Vladimir Sergeevich Solov'ëv was a Russian philosopher, poet, polemical essayist, and literary critic. His father, S. M. Solov'ëv, was an eminent historian and professor at Moscow University.

After graduating in 1873 from the historico-philological department of Moscow University, Solov'ëv studied for a year at the Moscow Theological Academy. In 1874 he defended his master's dissertation, *Krizis zapadnoi filosofii. Protiv pozitivistov* (The crisis of western philosophy: Against positivists) and was elected a docent of philosophy at Moscow University. During 1875–1876 he conducted research at the British Library, where he concentrated on mystical and Gnostic literature, including Jakob Boehme, Paracelsus, Emanuel Swedenborg, and the kabbalah.

Having a poetic and impressionable nature, Solov'ëv apparently possessed mediumistic gifts. Several times he had visions of Sophia, or the Eternal Feminine; he tells about one such vision, which he had in Egypt in 1875, in his poem "Three Meetings." After his return to Russia, he resumed lecturing at Moscow University; but in 1877, because of conflicts among the professors, he left the university and went to Petersburg to serve on the Scholarly Committee of the Ministry of National Education, meanwhile giving lectures at Petersburg University and at the Higher Courses for Women.

In 1877 he published the essay "Filosofskie nachala tsel'nlgo znaniia" (Philosophical principles of integral knowledge); during 1877–1880 he wrote the study *Kritika otvlechennykh nachal* (Critique of abstract principles); and in 1878 he began reading the cycle of *Chteniia o bogochelovechestve* (Lectures on godmanhood).

On March 28, 1881, after the assassination of Tsar Alexander II, Solov'ëv, in a public lecture on the incompatibility of capital punishment with Christian morality, called on the new tsar to refrain from executing the assassins. His lecture provoked a fierce reaction; the relations between the philosopher and the authorities were ruined, and he left public and academic service, becoming a professional writer.

In the 1880s his attention was focused on sociopolitical and religious questions. His most important works of this period were *Dukhovnye osnovy zhizni* (Spiritual foundations of life; 1882–1884), *Velikii spor i khristianskaia politika* (The great dispute and Christian politics; 1883), *Istoriia i budushchnost' teokratii* (The history and the future of theocracy; Zagreb, 1886), *Tri rechi v pamiat' Dostoevskogo* (Three speeches in memory of Dostoevsky; 1881–1883), *La Russie et l'Eglise Universelle* (Paris, 1889; Russian translation, 1911), and the cycle of essays *Natsional'nyi vopros v Rossii* (The national question in Russia; 1883–1891).

In the 1890s Solov'ëv returned to philosophical: work proper. He wrote the essay "Smysl liubvi" (The meaning of love; 1892–1894) and the treatise on ethics *Opravdanie dobra* (The justification of the good; 1894–1895); he proposed a new interpretation of the theory of knowledge in essays unified under the title *Pervoe nachalo teoreticheskoi filosofii* (The first principle of theoretical philosophy; 1897–1899); and his last significant work, *Tri razgovora* (Three conversations; 1899–1900), was devoted to the problem of evil. Excessive work and unsettled life ruined Solov'ëv's health, which had always been poor. He died near Moscow as a guest on the estate of his friends, the Princes Trubetskoi.

In his spiritual development, Solov'ëv experienced many influences that determined the orientation and character of his thought. In early youth he assimilated socialist ideas: the quest for social truth and faith in progress, which were characteristic for Russian thought and in fact for the nineteenth century in general. From the Slavophiles Solov'ëv assimilated the idea of "integral knowledge," which offered an answer to the question of the meaning of human existence, as well as to that of the goal of the cosmic and historical process. According to Solov'ëv the subject of this process is humanity as a single organism, a concept borrowed from Auguste Comte. This approach is based on Solov'ëv's belief in the reality of the universal, a belief formed under the influence of

Benedict (Baruch) de Spinoza and of German idealism, especially Georg Wilhelm Friedrich Hegel.

Solov'ëv was also greatly influenced by thinkers who attributed a metaphysical significance to the concept of the will: Immanuel Kant, Arthur Schopenhauer, Eduard von Hartmann, and especially Friedrich Wilhelm Joseph von Schelling. If Solov'ëv owes his dialectical method primarily to Hegel, his theology, metaphysics, and aesthetics bear the stamp of the influence of voluntaristic metaphysics. Solov'ëv converges with Schelling in his romantic aesthetic approach to problems of religion and in his erotic mysticism that culminates in the cult of the Eternal Feminine, the world soul. A significant role in the formation of Solov'ëv's views belonged to the Christian Platonism of P. D. Iurkevich, especially the latter's doctrine of the heart as the center of spiritual life. Solov'ëv creatively transformed these multifarious influences in his doctrine by developing a systematic philosophy, which, however, was not free of a number of difficulties and contradictions. In his works one finds a sober assessment and constructive critique of many philosophical conceptions that had previously contributed to forming his worldview.

BEING AND EXISTENCE

Solov'ëv constructs his philosophical system according to a schema of history as the development of the world spirit, that is, as a theo-cosmo-historical process. He rejects the secularism that permeates modern European philosophy and, following the early Slavophiles, seeks to attain integral knowledge that presupposes the unity of theory and practical activity. His goal is "to introduce the eternal content of Christianity into a new rational unconditional form proper to this content" (1908–1923, p. 2:89).

In other words, his goal is to justify this content by means of "theosophy," an investigation of the nature of God. Like the Slavophiles, Solov'ëv critiques abstract thought (particularly Hegel's idealism) from the vantage point of spiritualistic realism, which requires that thought, the thinking subject, and the thought content be separated into distinct elements—elements that coincide for Hegel in the absolute idea. According to Solov'ëv, that which genuinely exists is not a concept or an empirical given but a real spiritual entity, the subject of will, existent (*sushchee*). The bearers of power and volition, spirits and souls alone possess reality; following Kant and Schopenhauer, Solov'ëv considers the empirical world to be only a phenomenon and describes it as being, in contradistinction to existent. The first and supreme "existent," God, is defined by Solov'ëv in the spirit of neoplatonism and the kabbalah as a positive nothingness, which is the direct opposite of Hegel's negative nothingness—pure being obtained by abstraction from all positive definitions. Having defined existence as that which appears, and being as a phenomenon, Solov'ëv thus interprets the connection between God and the world as the connection between essence and phenomenon, establishing a relation of necessity between the transcendent foundation of the world and the world itself, which can be known by means of reason—with the aid of so-called organic logic.

However, there is a certain contradiction between Solov'ëv's mystical realism and his rationalistic method: If that which is, is a transcendent spiritual entity, one can have knowledge of it only on the basis of revelation. It is inaccessible to rational knowledge. However, Solov'ëv is convinced that the rationally unfathomable existent can be an object of mystical contemplation, of intellectual intuition understood in a special manner and identified by Solov'ëv with the state of inspiration. Following Schelling and the Romantics, Solov'ëv takes intellectual intuition to be akin to the productive capacity of the imagination and, accordingly, he takes philosophy to be akin to artistic creation, interpreting here the creative act by analogy with the passively mediumistic trance state. Solov'ëv considers the ecstatic inspired state to be the origin of philosophical knowledge:

> The action upon us of ideal entities, producing in us the intellectual or contemplative knowledge (and creation) of their ideal forms or ideas, is what is called inspiration. This action takes us out of our ordinary natural center and raises us to a higher sphere, thereby producing ecstasy. Thus … the directly defining principle of true philosophical knowledge is inspiration. (1911–1914, p. 1:294)

By identifying the direct action of transcendent entities on people with the intellectual contemplation of ideas, Solov'ëv removes the boundary between rational thought and mystical vision; and the removal of the distinction between mystically interpreted intellectual intuition and the productive capacity of the imagination leads to the confusion of artistic imagination with religious revelation and to a magical and occultist interpretation of art, characteristic not only of Solov'ëv but also of the symbolists whom he influenced. It is precisely in this manner that Solov'ëv understands the synthesis of philosophy, religion, and art. According to Solov'ëv the divine "That Which Is" is revealed directly, with the aid of sensation or emotion; and therefore no proofs of the existence of God are required: His reality cannot be logically derived from

pure reason but is given only by an act of faith. Nevertheless, the content of the divine "Existent" is revealed with the aid of reason.

ALL-UNITY

Solov'ëv describes the Absolute as the "eternal all-one" (1911–1914, p. 3:234), or as the "One and all." This means that all that which exists is contained in the Absolute: The all-unity is unity in multiplicity. According to Solov'ëv the one is independent of the all (the term *absolute* means "detached," "liberated"), and consequently it is defined negatively in relation to the other. But since it cannot have anything outside itself, it is defined positively in relation to the other. Thus, two poles or centers are eternally present in it: (1) independence of all forms, of all manifestation; and (2) the power that produces being, that is, the multiplicity of forms. The first pole is the One; the second is the potency of being, or the first matter, which, as in Boehme, is included in the Absolute as "its other," as the first substrate, or the "ground" of God.

Solov'ëv clarifies the concept of the first matter in terms of Schellingian-Schopenhaurean definitions—as power, attracting, striving, and originating in Boehme's doctrine of the "dark nature" in God, the doctrine of the unconscious depths of Divinity as the principle of evil. The inseparability of the two poles of That Which Exists signifies that the Absolute cannot appear except as actualized in matter, and matter cannot appear except as idea, as the actualized image of the One. In his *Critique of Abstract Principles* Solov'ëv describes the second pole of the all-unity, that is, the first matter (which is idea, or nature), as the becoming all-one, in contrast to the first pole, which is the existent all-one (1911–1914, p. 2:299). This means that the Absolute cannot exist except as actualized in its other. The pantheistic basis of this conception is obvious: This view of the relation between God and the world differs from the Christian idea of creation. The becoming all-one is the world soul, which, being the foundation of the entire cosmic process, only "in man first receives its proper inner activity, finds itself, is conscious of itself" (pp. 2:302–303).

SOPHIOLOGY

In his *Lectures on Godmanhood* Solov'ëv attempts to translate the self-sundering of the Absolute into the language of Christian theology, giving his own interpretation of the dogma of the Trinity. He distinguishes God as the absolute existent from His content (essence or idea), which appears in the person of the Son, or the Logos. The incarnation of this content is realized in the world soul, Sophia, the third person of the divine Trinity—the Holy Spirit. Distinguishing in God the active unity of the creative Word (Logos) and the actualized unity, His organic body, Solov'ëv views the latter as "the produced unity to which we have given the mystical name Sophia" (1911–1914, p. 3:111); it "is the principle of humanity, the ideal or normal man" (p. 3:111). Perfect humanity is not an empirical individual or man as a generic concept, but an eternal idea, a special kind of universal individuality, "the universal form of the union of material nature with Divinity … God-man-hood and Divine matter" (1911–1914, p. 8:231). The empirical world, where people appear as individuals, is "the somber and excruciating dream of a separate egotistical existence" (1911–1914, p. 3:120), an illusory and inauthentic world.

For Solov'ëv, as well as for Schopenhauer, the cause of this world is "the sin of individuation," producing the external, material existence of separateness and enmity. But if individuality is the source of evil and suffering, then in no wise can it be immortal: Salvation lies in the liberation from individual existence, not in its eternal continuation. Solov'ëv's philosophy of the last period is impersonalistic; it is not by chance that, on this question, there arose a polemic between him and Lev Mikhailovich Lopatin, who was convinced of the substantiality of the human self and of the immortality of the individual soul.

Solov'ëv sees the source of world evil in the meonic foundation of the divine all-unity. The world soul, Sophia, falls away from God, seeking to ground herself outside of Him, and "falls out of the all-one center of Divine being into the multiple periphery of creation, losing her freedom and her power over this creation" (1911–1914, p. 3:131). Meanwhile, the Divine Universe falls apart into a multiplicity of separate elements. The central personage of the theocosmic process—the eternally feminine principle in God, the body of Christ, the ideal humanity—acquires demonic characteristics. The image of the Eternal Feminine becomes dual. To eliminate this duality, Solov'ëv, in *Russia and the Universal Church*, introduces the distinction between Sophia on the one hand and the world soul on the other hand. The latter now appears as the antipode of Sophia, the Wisdom of God, who is a "radiant and heavenly entity." The essence of the cosmic process is the battle between the Divine Word and the infernal principle for power over the world soul, a battle that must end with the reunification of the fallen world soul with God and the restoration of the divine all-unity. The historical process leads with internal necessity to the triumph of good, to the victory of unity and love over disintegration and enmity. Solov'ëv's theodicy converges

not only with Hegel's teleological determinism but also with the evolutionism of the natural sciences.

PHILOSOPHY OF HISTORY

Solov'ëv's philosophy of history is an attempt to understand cosmic history as a series of free acts on the way to the restoration of the unity of God and humanity. At the first stage, that of natural revelation, humanity knows God as a natural entity: Such are the pagan beliefs of the ancient world and the materialistic doctrines of the modern period. At the second stage, God is revealed as the transcendent, extranatural principle; such are the Asian ascetic-pessimistic religions, especially Buddhism, which seek to overcome the active, personal principle. Finally, in the religion of the Old Testament, humanity received a positive revelation, whose full meaning was disclosed in Christianity. In Christ was manifested the synthesis of the religiously contemplative principle of Russia and the personal and human principle, which developed in the bosom of European culture.

However, the schism between the Eastern and Western Churches marked the epoch of a new disintegration, which now affected the Christian world because of the imperfection of "historical Christianity." Triumphant in Russia was the supraindividual divine principle that left no room for human freedom; whereas Europe was marked by excessive individualism and by freedom in its negative sense, which led to capitalism atheistic egotism. Russia had a messianic calling to unify the two separated sides and to realize the final act of the cosmic historical drama in which humanity will be reunited with God. In the 1880s Solov'ëv's philosophy of history takes the form of a utopian doctrine about a universal theocracy in which the secular power of the Russian tsar is joined with the spiritual power of the pope in Rome. The first step toward this theocracy was supposed to take place as the reunification of the Eastern and Western Churches.

ETHICS

Solov'ëv touched on the problems of ethics in many works, but he has one work that is specially devoted to moral philosophy: *The Justification of the Good*. In this work he critiques two extreme points of view: moral subjectivism, which asserts that only the person can be the bearer of good; and objectivism, which recognizes only social institutions as guarantors of moral conduct. According to Solov'ëv these two elements must complement each other. Here, he underscores the importance of the objective forms of moral life, taking as his point of departure the belief in the reality of the universal, that is, of Godmanhood as one organism. If in his early works Solov'ëv emphasized the dependence of ethics on religious metaphysics, he now insists on the autonomy of ethics, because as "in creating a moral philosophy, reason does nothing more than develop, on the basis of experience, the idea of the good that is originally inherent in it" (1911–1914, p. 7:29). Nevertheless, even if it is autonomous the philosophy of morality cannot be fully separated from metaphysics and religion, because only the doctrine of the cosmic divine-human process and of the final victory of the divine all-unity grounds morality—the reality of superhuman good.

Solov'ëv gives a deep analysis of moral emotions: shame, pity, piety, or veneration. Man is ashamed of that which constitutes his lower nature; characteristic in this respect is sexual shame. Human experience pity, that is, they empathize with the suffering of all living beings; as the source of altruism, pity is the basis of social relationships. Shame represents individual chastity, whereas pity represents social chastity. Finally, the sense of piety, that is, veneration of the supreme principle, is the moral foundation of religion. Examining the problem of the relation between morality and law, Solov'ëv sees their distinction in the fact that, in contrast to legal obligations, moral ones are unlimited, as well as in the coercive character of juridical laws. Law is the lower bound or minimum of morality, which is realized by means of compulsion. However, contrary to the common opinion, there is no contradiction between moral and juridical laws.

Although Solov'ëv does not have a work specially devoted to aesthetics, the theme of beauty permeates all of his works. For Solov'ëv, philosophical intuition converges, in the spirit of romanticism, with artistic creativity; and he sees in the latter a kinship with mystical experience and considers art to be a real power, illuminating and regenerating the world (see 1911–1914, p. 3:189). The supreme goal of art is theurgy, that is, the transformation of everyday reality into ideal, transfigured corporeality. Solov'ëv's aesthetics is connected with his sophiology and with his doctrine of Eros, to which his treatise *The Meaning of Love* is devoted. His aesthetic ideas were also expressed in his essays in the field of literary criticism devoted to the poetry of Aleksandr Pushkin, Fedor Tiutchev, Mikhail Lermontov, and Afanasy Fet.

Not long before his death, Solov'ëv became disenchanted with theocratic utopia and, in general, with the idea of progress. In his final work, *Three Conversations*, the central plane is occupied by the eschatological theme: The coming of the Kingdom of God is now conceived as the end of history. Solov'ëv had a powerful influence on

philosophical thought in Russia. The religious philosophy of the end of the nineteenth century and of the beginning of the twentieth century developed under his influence; this is true, in particular, of Sergei Trubetskoi and Evgenii Trubetskoi, Nikolai Losskii, S. L. Frank, Sergei Bulgakov, Pavel Florenskii, Nikolai Berdiaev, and so on. Just as significant was Solov'ëv's influence on Russian literature, especially on the symbolists Aleksandr Blok, Andrei Belyi, Viacheslav Ivanov, and so on. It is precisely from Solov'ëv that the Russian silver age got its mystical and Gnostic tendency, which was characteristic for the atmosphere of the spiritual life of the pre-Revolutionary period in Russia.

See also Absolute, The; Berdyaev, Nikolai Aleksandrovich; Boehme, Jakob; Bulgakov, Sergei Nikolaevich; Comte, Auguste; Florenskii, Pavel Aleksandrovich; Frank, Semën Liudvigovich; Gnosticism; Hartmann, Eduard von; Hegel, Georg Wilhelm Friedrich; Kabbalah; Kant, Immanuel; Losskii, Nikolai Onufrievich; Mysticism, Nature and Assessment of; Paracelsus; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Swedenborg, Emanuel; Trubetskoi, Evgenii Nikolaevich; Trubetskoi, Sergei Nikolaevich.

Bibliography

WORKS BY SOLOV'ËV

Pis'ma (Letters). Vols. 1–4. St. Petersburg: 1908–1923.

Sobranie sochinenii (Collected works). Vols. 1–10. St. Petersburg: 1911–1914.

Stikhotvoreniia i shutochnye piesy (Poems and comic plays). Leningrad: 1974.

Sochineniia (Works), 2 vols. Moscow: 1988.

Literaturnaia kritika (Literary criticism). Moscow: 1990.

WORKS BY SOLOV'ËV IN ENGLISH TRANSLATION

The Justification of the Good: An Essay in Moral Philosophy. Translated by Natalie A. Duddington. London: Constable, 1918. This was originally published under the title *Opravdanie dobra: Nravstvennaya filosofiia* in 1897.

God, Man, and the Church: The Spiritual Foundations of Life. Translated by Donald Attwater. Milwaukee, WI: Bruce Publishing, 1938; London: James Clarke, 1938, 1973. This was originally published in French under the title *Dukhovnye osnovy zhizni* in 1884.

Russia and the Universal Church. Translated by Herbert Rees. London: G. Bles, 1948. This was originally published in French under the title *La Russie et l'Eglise universelle* in 1888.

A Solovyov Anthology. Compiled by S. L. Frank. Translated by Natalie Duddington. London: SCM Press, 1950.

"Foundations of Theoretical Philosophy," Part 1. In *Russian Philosophy*, 3 vols. Translated by Vlada Tolley and James P. Scanlan; edited by James M. Edie, James P. Scanlan, and Mary-Barbara Zeldin, with the collaboration of George L. Kline, 99–134. Chicago: Quadrangle Books, 1965. This was originally published under the title "Osnovy teoreticheskoi filosofii" in 1897–1899.

The Meaning of Love. Translated by Thomas R. Beyer Jr. West Stockbridge, MA: Lindisfarne Press, 1985. This was originally published under the title *Smysl liubvi* in 1892–1894.

Three Conversations concerning War, Progress, and the End of History, Including a Short Story of the Antichrist. Hudson, NY: Lindisfarne Press, 1990. This was originally published under the title *Tri razgovora o voine, progresse, i kontse istorii, i kratkaia povest' ob Antikhriste* in 1900.

Lectures on Divine Humanity. Translated by Peter Zouboff; edited by Boris Jakim. Hudson, NY: Lindisfarne Press, 1995. This was originally published under the title *Chteniia o Bogochelovechestve* in 1877–1881.

The Crisis of Western Philosophy: Against the Positivists. Translated by Boris Jakim. Hudson, NY: Lindisfarne Press, 1996a. This was originally published under the title *Krizis zapadnoi filosofii: Protiv pozitivistov* in 1874.

Poems of Sophia. Translated by Boris Jakim and Laury Magnus. New Haven, CT: Variable Press, 1996b. Includes "Three Meetings."

WORKS ON SOLOV'ËV

Asmus, V. F. "V. S. Solov'ev: Opyt filosofskoi biografii" (V. S. Solov'ëv: An attempt at a philosophical biography). *Voprosy filosofii* 6 (1988).

Balthasar, Hans Urs von. *Studies in Theological Lay Styles, Vol. 3: The Glory of the Lord: A Theological Aesthetics*. Translated by Andrew Louth et al.; edited by John Riches, 279–352. Edinburgh: T. and T. Clark, 1986. This essay was originally published in German under the title *Herrlichkeit: Eine theologische Aesthetik* in 1962.

Gaidenko, P. P. "Chelovek i chelovechestvo v uchenii V.S. Solov'eva" (Man and humanity in Solov'ëv's doctrine). *Voprosy filosofii* 6 (1994).

Gaidenko, P. P. *Vladimir Solov'ev i filosofiia Serebrianogo veka* (Vladimir Solovyov and the philosophy of the silver age). Moscow: 2001.

George, M. *Mystische und religiöse Erfahrung im Denken Vladimir Solov'evs*. Göttingen: 1988.

Losev, A. F. *Vladimir Solov'ev*. Moscow: 1983.

Losey, A. F. *Vladimir Solov'ev i ego vremia* (Vladimir Solov'ëv and his time). Moscow: 1990.

Lossky, N. O. *History of Russian Philosophy*. New York: International Universities Press, 1951.

Madey, J. *W. S. Solowjew und seine Lehre von der Weltseele*. Düsseldorf, Germany: 1961.

Mochul'sky, K. V. *Vladimir Solov'ev. Zhizn' i uchenie* (Vladimir Solov'ëv: Life and teaching). 2nd ed. Paris: YMCA Press, 1951.

Radlov, E. L. *Vl. Solovyov. Zhizn' i uchenie* (Vl. Solov'ëv: Life and teaching). St. Petersburg: 1913.

Solov'ev, S. M. *Vladimir Solovyov. Zhizn' i tvorvcheskaia evoliustii* (Vladimir Solov'ëv: Life and creative evolution). Moscow: Respublika, 1997.

Strémooukhoff, D. *Vladimir Soloviev and His Messianic Work*. Translated by Elizabeth Meyendorff; edited by Phillip Guilbeau and Heather Elise MacGregor. Belmont, MA: Nordland, 1980. This was originally published in French

under the title *Vladimir Soloviev et son oeuvre messianique* in 1935 and 1975.

Sutton, Jonathan. *The Religious Philosophy of Vladimir Solovyov: Towards a Reassessment*. New York: St. Martin's, 1988.

Trubetskoi, E. N. *Mirosozertsanie Vl. S. Solovyova* (The world view of V. S. Solov'ëv). 2 vols. Moscow: 1913.

Utkina, N. F. "Tema vseedinstva v filosofii Vl. Solovyova" (The theme of all-unity in Solov'ëv's philosophy). *Voprosy filosofii* 6 (1989).

Velichko, V. L. *Vladimir Solovyov. Zhizn' i tvoreniia* (Vladimir Solov'ëv: Life and works). St. Petersburg: 1903.

Wenzler, L. *Die Freiheit und das Böse nach Vl. Solov'ev*. Freiburg-Munich: 1978.

Zen'kovskii, V. V. *A History of Russian Philosophy*. 2 vols. Translated by George L. Kline. New York: Columbia University Press, 1953. This was originally published under the title *Istoriia russkoi filosofii* in 1948–1950.

Piama Gaidenko (2005)
Translated by Boris Jakim

SOMBART, WERNER

(1863–1941)

Werner Sombart, the German economic and social theorist, was born in Ermsleben near the Harz Mountains. He was professor of economics at the University of Breslau from 1890 to 1906 and at Berlin University from 1906 to 1931. Sombart made a strong impact on German economic thought and policies; he played a leading role in the Verein für Sozialpolitik and the Deutsche Soziologische Gesellschaft, and he was joint editor with Max Weber and Edgar Jaffe of the journal *Archiv für Sozialwissenschaft und Sozialpolitik.*

Sombart's interests covered economic and social history and theory, sociology, and the methodology of the social sciences, although his contributions to methodology were more polemical than constructive. Together with Wilhelm Dilthey, Heinrich Rickert, Karl Jaspers, and Max and Alfred Weber, he helped to establish modern German historical and cultural sociology. Sombart was a highly prolific writer, and few of his writings are free from marks of careless workmanship, though nearly all sparkle with suggestive ideas.

STUDY OF CAPITALISM

Sombart concentrated on the study of the development and the structural makeup of European industrial society and in particular on the development of capitalism and the transition from capitalism to socialism. In his early work he was influenced by Karl Marx, but in his mature period he sought to go beyond Marx's theoretical and historical edifice and fundamentally to undermine the Marxist weltanschauung.

Sombart's magnum opus was *Der moderne Kapitalismus,* whose first and second versions (1902 and 1916–1927) both demonstrated methodological and substantive advances. In contrast to Max Weber's comparative-institutional approach, Sombart conceived of the European capitalist system as a "historical individual," that is, the collective expression of the values of the expansive "Faustian" spirit of enterprise and the acquisitive bourgeois spirit. He traced the development of capitalism through early, high (mature), and late periods, each representing different cultural attitudes and styles. The basic qualities of each period were seen as determined by its system of economic values (*Wirtschaftsgesinnung*)—which he understood as being in continuous interpenetration with the other areas of cultural and social activity; by the forms of its legal and social organization; and by its technology and methods. In a dialectical process of transition, one period generates another as its antithesis. His emphasis on the concrete historical elements caused Sombart to neglect the theoretical and analytical structure of economics, which he regarded as supplementary to his own kind of investigation. Thus, economists tend to regard Sombart's work as history, but historians do not.

Sombart supported his study of capitalism by a large number of sociological monographs on such subjects as the city, precious metals, the location of industry, Jews, fashion, advertising, the bourgeois, the proletariat, war and capitalism, and luxury and capitalism. Following the Russian and German revolutions at the end of World War I, Sombart sharply dissociated himself from Marxian socialism, which, like capitalism, he regarded as "uninhibited Mammonism," the victory of evil forces (utilitarianism and hatred) over idealism and love. He advocated "German socialism" or "anticapitalism," based on the rejection of materialism, "technomania," and belief in progress. His specific prescriptions became increasingly totalitarian.

SOCIAL PHILOSOPHY

In social and cultural philosophy Sombart stressed the idea of an "economic system" (*Wirtschaftssystem*) whose forms and organization are the creation of the mind and reflect the clusters of cultural values (*Wirtschaftsgesinnungen*) mentioned above. The concept of *Wirtschaftssystem* is related to that of structure and to Max Weber's "ideal types." Originally Sombart conceived of this con-

cept in terms of the early psychology of Dilthey and, like Weber, took account of the subjective intentions of historical agents. Later, however, he turned to an almost phenomenological interpretation of the "objective" meaning of cultural systems. Like Weber, Sombart regarded the "ideal type" both as a conceptual tool for evaluating historical processes and as a reflection of the essential structure of historicocultural reality. Sombart, however, emphasized the "realist" function and interpreted history as an expression of the national spirit rather than a multicausal sequence. In the first edition of *Der moderne Kapitalismus* this attitude led him to a naturalistic confusion of theory and history, which was assailed by Weber. Though Sombart was an economist by profession, he regarded economic laws as determined by the exigencies of the spirit of the age, and like Auguste Comte and the German historical school, he rejected the claim of economics to be an independent discipline. In his *Die drei Nationaloekonomien,* which he regarded as the theoretical key to his work, he distinguished between ethical (*richtende*), analytical (*ordnende*), and interpretive (*verstehende*) economics. He rejected the first because science should be ethically neutral, the second because it fastened on applied science only and opened the door to the mechanical methods of the natural sciences, which cannot lead to the required understanding of meanings, of cultural institutions, and of motivations (*Sinn-*, *Sach-*, and *Seelverstehen*). His insistence on the exclusion of value judgments, on the one hand, and on an intuition of essences, on the other hand, led Sombart into unresolved intellectual difficulties and caused him finally to stress the superiority of biased observation over the limited vistas of scientific thought. Sombart came to regard the dispute over methods as a contest between German (heroic-spiritual) and Western (utilitarian-mercenary) thought. He reproached Western philosophy for the "deconsecration of the mind," a destructive tendency to resolve the spiritual realm of ideas into their psychological and sociological elements.

Accordingly, Sombart saw sociology as more than a limited specialized discipline; to him, it was a universal discipline whose aim is to explain the whole of human relationships and cultural categories. He viewed society as a creation of the mind, and accordingly, his "noo-sociology" embraced religion, art, the law, and the state, as well as economics. In his final work, *Vom Menschen,* Sombart assigned the same universal function to philosophical (*geistwissenschaftliche*) anthropology, which was to be developed into a "basic science" coordinating all knowledge concerning human groups and peoples, both their structures and their origins. This work, a bitter indictment of civilization, was, however, merely programmatic.

Sombart exerted considerable influence upon a generation of German economists and sociologists, but his chief significance lies in his suggestive contributions to the morphology and genesis of capitalism and to the history of economic and social ideas.

See also Comte, Auguste; Dilthey, Wilhelm; Jaspers, Karl; Marx, Karl; Marxist Philosophy; Philosophical Anthropology; Philosophy of Economics; Rickert, Heinrich; Weber, Alfred; Weber, Max.

Bibliography

WORKS BY SOMBART

Sozialismus und soziale Bewegung im 19. Jahrhundert. Jena, Germany: Fischer, 1896; 10th ed. titled *Der proletarische Sozialismus,* 2 vols. Jena, 1924. Translated by M. Epstein as *Socialism and the Social Movement in the 19th Century.* London: Putnam, 1898.

Der moderne Kapitalismus, 2 vols., Munich, 1902; 7th ed., 6 vols., Munich: Duncker and Humblot, 1928. Condensed translation by F. L. Nussbaum as *A History of the Economic Institutions of Modern Europe.* New York: Crofts, 1933.

Das Proletariat. Frankfurt: Rütten and Loening, 1906.

Die Juden und das Wirtschaftsleben. Munich, 1911. Translated by M. Epstein as *The Jews and Modern Capitalism.* London: Unwin, 1913; paperback, with introduction by B. Hoselitz, New York, 1962.

Der Bourgeois. Munich: Duncker and Humblot, 1913. Translated by M. Epstein as *The Quintessence of Capitalism.* London: Unwin, 1915.

Studien zur Entwicklungsgeschichte des modernen Kapitalismus, 2 vols. Munich: Duncker and Humblot, 1913. Translated by W. R. Dittmar as *Luxury and Capitalism.* New York, 1938.

Die drei Nationaloekonomien. Munich: Duncker and Humblot, 1930.

Die Zukunft des Kapitalismus. Berlin: Buchholz and Weisswange, 1932. Translated as *The Future of Capitalism.* Berlin and London, 1932.

Deutscher Sozialismus. Berlin: Buchholz and Weisswange, 1934. Translated by Karl F. Geiser as *A New Social Philosophy.* Princeton, NJ: Princeton University Press, 1937.

Vom Menschen. Berlin: Buchholz and Weisswange, 1938.

Noo-Soziologie. Edited by Nicolaus Sombart. Berlin, 1956.

Why Is There No Socialism in the United States?, Translated by Patricia M. Hocking and C. T. Husbands; edited by C. T. Husbands. White Plains, NY: International Arts and Sciences Press, 1976.

WORKS ON SOMBART

Brocke, Bernhard vom. "Werner Sombart (1863–1941): Capitalism, Socialism, His Life, Works and Influence since Fifty Years," translated by Lode Vereeck. *Jahrbuch für Wirtschaftsgeschichte* (1) (1992): 113–182.

Mitzman, Arthur. "Werner Sombart (1863–1941)." In his *Sociology and Estrangement: Three Sociologists of Imperial Germany.* New York: Knopf, 1973.

Nussbaum, Frederick Louis. *A History of the Economic Institutions of Modern Europe; An Introduction to* Der moderne Kapitalismus *of Werner Sombart.* New York: A.M. Kelley, 1968.

Plotnik, Mortin J. *Werner Sombart and His Type of Economics.* New York: Eco Press, 1937. Has good bibliography.

Weippert, Georg. *Werner Sombarts Gestaltidee des Wirtschaftssystems.* Göttingen, 1953.

H. O. Pappé (1967)
Bibliography updated by Philip Reed (2005)

SOPHIA

See Appendix, Vol. 10

SOPHISTS

In English, the term *sophist* is most often used pejoratively, for one who argues with devious abuses of logic. The Greek *Sophistês* took on a similar sense in the fifth century BCE., but its original meaning is simply *expert* or *wise person*. In the study of Greek philosophy, the sophists denote a group of teachers and intellectuals of the fifth and fourth century BCE (the term is also used for later practitioners of their profession; this soon comes to be interchangeable with rhetoric or public speaking, as in the so-called Second Sophistic movement of the second century CE).

The sophists are perennially ambiguous and controversial figures, and it has long been debated whether they should be deemed philosophers. Two central points seem clear: First, the sophists did not constitute a philosophical school with a shared set of metaphysical and ethical positions; second, a number of them did develop serious, innovative, and influential ideas and arguments on a wide range of topics, and so demand inclusion in the history of ancient philosophy.

The sophists are best seen as an intellectual movement, comparable to the philosophies of the eighteenth century or the progressive thinkers of Victorian England (some of whom, such as George Grote, were champions of the ancient sophists). As always with such movements, it is debatable who should be counted as a member, and membership is in any case more a matter of shared interests and tendencies than common doctrines. The leading figures of the sophistic movement so understood include Protagoras, Gorgias, Hippias, Antiphon, and Prodicus. Gorgias was primarily a rhetorician (i.e., an expert in and teacher of public speaking), but the two professions must have overlapped widely, and his surviving texts are among the most important for reconstructing sophistic ideas. Socrates was often counted among the sophists by his contemporaries, and is used to represent the whole movement in Aristophanes' *Clouds*; in a number of dialogues Plato aims to show that he differs from them radically.

Sophistic ideas have also come from some important anonymous texts, such as the *Dissoi Logoi* and the *Anonymus Iamblichi* (a long discussion of virtue, apparently of sophistic origin, inserted by the Neoplatonist Iamblichus in his *Protrepticus*), or of contested authorship (notably the fragment on religion from the satyr play *Sisyphus*, attributed to both Critias and Euripides).

They can also be found in contemporary historical and medical texts (e.g., Thucydides' *Melian Dialogues*, the Hippocratic *On the Art*), as well as comedy and tragedy (especially Euripides). So there is no firm dividing line between sophistic thought and the broader fifth-century Greek culture around them, which was marked by a vigorous questioning of tradition and empirical, naturalistic researches into many subjects (*historiê*).

Sociologically, the sophists were professional teachers, the first in Greece to offer a higher education in the liberal arts. Sophists (who came from all over the Greek world) traveled from city to city presenting themselves to prospective students through public displays; this could involve giving a set speech (*epideixis*), performing feats of memory, undertaking to answer any question the audience might pose, or offering question-and-answer refutations of others. This practice of refutation, usually given the pejorative name *eristic*, is formally identical to the Socratic *elenchus*; to differentiate the two, Plato emphasizes that Socrates argues in pursuit of the truth and moral improvement, whereas sophists argue for victory and for money. Some sophists gave displays at the Olympic games, and the sophistic practices themselves were intensely agonistic.

Plato's *Protagoras* gives a vivid depiction of a gathering of sophists engaged in argument, banter, and competitive intellectual showing-off. Such sessions served as advertisements to the wealthy young men who made up the audience, encouraging them to sign on for further teaching. This would be an expensive proposition: The sophists (and above all Protagoras) seem to have charged far more than any other contemporary professionals, and became enormously rich from their teaching. Sophists also served on embassies for their native cities, drafted laws, and wrote books; they were famous and influen-

tial—and bitterly controversial—public intellectuals as well as teachers.

Most sophists are said to have claimed to teach virtue (*arête*), but their curricula and teaching methods varied. In the *Protagoras*, Protagoras chides Hippias for forcing students to study subjects like mathematics and astronomy; he himself claims to teach them good judgment (*euboulia*), enabling them both to manage their private affairs and to succeed in politics, and accepts that this amounts to the teaching of virtue. He also claims that the greatest part of education is the ability to analyze and criticize poetry. So the sophistic teaching of virtue was not a matter of moralistic indoctrination; rather, the sophists taught their students to reflect on traditional values, to analyze and criticize the literary texts that discussed them, and to apply this learning in a political career.

In practice, their teaching seems to have centered on rhetoric or public speaking (hence the blurriness of the line between rhetorician and sophist), which was the key skill for a political career. The connection between teaching rhetoric and teaching virtue is easier to understand if we bear in mind the traditional, Homeric sense of *arête* as excellence—that is, the skills and personal qualities that make a gentleman successful in his career and a valuable asset to his community. By teaching the arts of political success, the sophists *were* teaching virtue in a quite traditional sense. In doing so they prompted debate about just what virtue or excellence really consists in, and in particular about the status of qualities such as justice, *dikaiosunê*, which seem to benefit the community at the expense of their possessor.

The evidence for sophistic ideas is uneven and very defective. There are several brief, but substantial, works by Gorgias (*On Not Being*; *Defense of Helen*), and a few pages worth of Antiphon's *On Truth*; but for Protagoras, the leading figure of the movement, only a handful of brief fragments (that is, trustworthy-looking quotations in later authors) survive. Moreover, many of our texts are ambiguous or difficult to interpret. For instance, both the *Dissoi Logoi* and Antiphon's discussion of justice seem to argue for contradictory conclusions; perhaps they are exercises in *antilogikê*, opposing arguments, a sophistic genre associated with Protagoras. Gorgias's *On Not Being* and the *Defense of Helen* both seem to be exercises in defending the indefensible; whether they also have serious philosophical agendas is still debated.

A further difficulty is posed by the all-important evidence of Plato, who fixed forever the stereotype of the Sophist. Plato vividly depicts sophists in a number of dialogues (*Protagoras*, *Gorgias*, *Republic* [Thrasymachus], *Hippias Major* and *Minor*, and *Euthydemus*), and the *Sophist* is devoted to defining their nature. But Plato's evidence is not consistent: For instance, the *Protagoras* and the *Euthydemus* give very different pictures of sophistic argument, and the *Protagoras* and *Theaetetus* seem to give conflicting accounts of Protagoras's ethical views. Moreover, Plato's presentation of the sophists is sometimes warped by hostile prejudice (though, as Grote [1865] and T.H. Irwin [1995] have noted, he is not as uniformly hostile as scholars sometimes assume), and by his anxiety to distinguish them as sharply as possible from Socrates.

Unsurprisingly, given the focus of their teaching, sophistic thought seems to have centered on ethical and political topics. However, sophistic interests varied greatly; in some cases they were very broad, and several sophists are associated with ideas in mathematics or natural science. So the traditional scholarly contrast between the sophists and the pre-socratics, with their researches into natural science, is probably misguided or at least overstated. The sophists also were founders of what are now called the social sciences; they offered theories of the origins of human institutions such as law and religion, and took a particular scientific interest in language and the norms applicable to it. Here in the social realm, the closest thing to a unifying pattern in sophistic thought is found—their concern to distinguish *phusis* and *nomos* (i.e., the natural and the merely conventional or, as one might now say, socially constructed).

Surviving sophistic texts analyze a wide range of human institutions and values—above all, justice—in these terms, with the assumption that nature represents a deeper or more binding norm than convention. Combined with the sophists' recognition of the differing norms of various cultures (see the *Dissoi Logoi*) and their skepticism about traditional religion, this privileging of the natural could be seen as undermining the authority of moral tradition. However, hostility toward the sophists probably had less to do with their particular theories than with their teaching to all comers the ability to speak persuasively, and with it the power to manipulate both political assemblies and legal proceedings.

It is now generally recognized that it is wrong to describe the sophists collectively as moral skeptics, immoralists, or relativists (Bett 1989, 2002). Protagoras is presented as a relativist in Plato's *Theaetetus*, but not in his earlier and probably more historically accurate *Protagoras*. The *Dissoi Logoi* presents a wealth of evidence for the cultural relativity of values, but argues against relativistic conclusions as well as for them. Sophistic uses of

nomos and *phusis* were often in the service of conflicting ethical and political theories, and attempts to pin the sophists down to any common moral theory are doomed by the sheer diversity of sophistic thought. If anything, the sophists (as one would expect given the competitive character of their profession) tended to take up positions in opposition to each other—even if the battle lines are often now blurred by the incompleteness of evidence available.

On matters of natural science, metaphysics, and epistemology, it is still more difficult to identify shared sophistic positions. Antiphon's *On Truth* seems to have offered a complete cosmogony on natural science: Aristotle, in *Physics*, reports him as claiming that the true essence of a wooden bed is wood because if planted it would reproduce a tree rather than another bed. Presumably the force of the scientific part of the work was to spell out this kind of distinction between the underlying natures of things (the realm of *phusis*) and merely superficial human arrangements and projections, (*nomos*).

Gorgias's *On Not Being* seems intended to support a skeptical conclusion, at least as a critique of metaphysicians like Parmenides. His main criticism was: If beings do have a real nature independent of humans, it can neither be known or communicated. Plato's *Theaetetus* attributes a sophisticated relativism or subjectivism in epistemology (and ethics) to Protagoras: its slogan, "Man is the measure of all things," must go back to Protagoras's work *Truth*, but how much of the detailed theory presented by Plato that is genuinely Protagorean is uncertain. Even setting aside other sophistic views (where evidence is even scantier), the most these positions could be said to share is a critical orientation—a tendency to diagnose beliefs and perceptions (both everyday and scientific, or philosophical) as irreducibly subjective.

In keeping with their activities as teachers and writers, and their interest in the analysis of human conventions, the sophists were noted for their researches into language. Prodicus was celebrated for drawing fine distinctions in the meanings of words. It is thought that Protagoras analyzed the parts of speech, and claimed that the words for *wrath* and *helmet*, feminine in Greek, were properly masculine. The sophists are often associated with claims that falsehood and contradiction are impossible, but the evidence for this is unclear and confusing, and these claims are hard to square with the eristic practice of inducing contradictions in others. One might suspect that distinctive views about the nature of truth were entailed by Protagoras's *Measure Thesis*, and lay behind his practice of argument on both sides of a question (*antilogikê*); but attempts to reconstruct sophistic ideas on these questions are highly speculative.

See also Antiphon; Gorgias of Leontini; Nomos and Phusis; Protagoras of Abdera; Socrates.

Bibliography

Barney, R. "The Sophistic Movement." In *The Blackwell Companion to Ancient Philosophy*, edited by Mary Louise Gill and Pierre Pellegrin. Oxford: Blackwell, 2005.

Bett, R. "Is There a Sophistic Ethics?" *Ancient Philosophy* 22 (2002): 235–262.

Bett, R. "The Sophists and Relativism." *Phronesis* 34 (1989): 139–169.

Broadie, Sarah. "The Sophists and Socrates." In *The Cambridge Companion to Greek and Roman Philosophy*, edited by D.N. Sedley. Cambridge, U.K.: Cambridge University Press, 2003.

Diels, H. *Die fragmente der Vorsokratiker*, 6th ed. revised by W. Kranz. Berlin: Weidmann, 1951–1952.

Dillon, J., and T. Gergel. *The Greek Sophists*. Penguin, 2003.

Gagarin, M., and P. Woodruff. *Early Greek Political Thought from Homer to the Sophists*. Cambridge, U.K.: Cambridge University Press, 1995.

Grote, George. *A History of Greece*, 4th ed. London: John Murray, 1872.

Grote, George. *Plato and the Other Companions of Socrates*. London: John Murray, 1865.

Guthrie, W.K.C. *A History of Greek Philosophy*. Vol. 3. Cambridge, U.K.: Cambridge University Press, 1969.

Irwin, T.H. "Plato's Objections to the Sophists." In *The Greek World*, edited by C.A. Powell. London: Routledge, 1995.

Kerferd, G.B. *The Sophistic Movement*. Cambridge, U.K.: Cambridge University Press, 1981.

Morrison, J. S. "The Truth of Antiphon." *Phronesis* 8 (1963): 35–49.

Nehamas, A. "Eristic, Antilogic, Sophistic, Dialectic: Plato's Demarcation of Philosophy from Sophistry." *History of Philosophy Quarterly* 7 (1990): 3–16.

Pendrick, G. *Antiphon The Sophist: The Fragments*. Cambridge, U.K.: Cambridge University Press, 2002.

Robinson, T.M., ed. *Contrasting Arguments: an Edition of the Dissoi Logoi*. New York: Arno Press, 1979.

Sidgwick, Henry. "The Sophists." *Journal of Philology* (1872): 288–307.

Sidgwick, Henry. "The Sophists." *Journal of Philology* (1873): 66–80.

Striker, G. "Methods of Sophistry." In *Essays on Hellenistic Epistemology and Ethics*. Cambridge, U.K.: Cambridge University Press, 1996.

Waterfield, R. *The First Philosophers: The Presocratics and Sophists*. Oxford: Oxford Paperbacks, 2000.

Rachel Barney (2005)

SÔPHROSYNÊ

See Appendix, Vol. 10

SOREL, GEORGES

(1847–1922)

Georges Sorel, the French pragmatist philosopher and social theorist, was born in Cherbourg and was trained at the École Polytechnique. He served as an engineer with the French roads and bridges department for twenty-five years in Corsica, the Alps, Algeria, and Perpignan before retiring at the age of forty-five to devote himself to scholarship. In the following thirty years he produced a series of highly curious books on the philosophy of science, the history of ideas, social theory, and Marxism, of which one, *Réflexions sur la violence* (1908; *Reflections on Violence*), immediately became world famous. Before and after his retirement Sorel's life was quite uneventful, for despite his hatred of the bourgeois, his conduct was a model of provincial respectability. Nevertheless, he never married his lifelong companion, Marie David, to whom he dedicated his work after her death in 1897. Sorel's Roman ideas on the importance of chastity, marriage, and the family were no match for his family's objections to Marie's proletarian origins.

ECONOMICS AND POLITICAL VIEWS

Sorel's first books, on the Bible and the trial of Socrates, were written while he was still in charge of irrigation around Perpignan. They are works of erudition, marked by a streak of passionate eccentricity. Soon after retiring to the suburbs of Paris, Sorel discovered the work of Karl Marx and edited (1895–1897) a magazine, *Le devenir social,* that introduced theoretical Marxism to France. At the same time Sorel collaborated with Benedetto Croce and Antonio Labriola in propagating Marx's ideas in Italy. (Italy was always Sorel's second intellectual home, although he never visited it or even left French territory, and much of his work has been published only in Italian.) Sorel soon became dissatisfied with Marxism's scientific pretensions and joined with Croce, Eduard Bernstein, Tomáš Masaryk, and Saverio Merlino in precipitating the revisionist crisis. The other revisionists drew reformist conclusions from their critique of Marxism and abandoned revolutionary activity, but Sorel did the opposite. He transferred his interest from orthodox socialism to the most revolutionary wing of the French labor movement, the anarchosyndicalists. He argued that this was consistent because the syndicalists did not use Marxism as science but as myth. It was to account for this mythical character of extremist social doctrines that Sorel elaborated one of his most influential theories.

By the eve of World War I, Sorel had lost faith in syndicalism, and for a time he associated with such extreme right-wing groups as the monarchists and ultranationalists, as well as with groups of Catholic revivalists. Silent during the war, Sorel emerged after the Bolshevik Revolution to devote his last energies to the defense of the cause of V. I. Lenin, as he understood it. He supposed that it meant transfer of power away from central authority to the workers' and peasants' soviets and thus that it was in the federalist spirit of Pierre-Joseph Proudhon rather than in the spirit of Marx.

Years earlier, Sorel had predicted an important political career for Benito Mussolini, who, in turn, called *Reflections on Violence* his bedside book. Yet despite tenacious legend, Sorel had no influence over either fascism or communism. He himself disclaimed any part in Mussolini's nationalist doctrines, and Lenin denied drawing ideas from "that confusionist." Apologists of later revolutionary movements, notably African and Asian nationalism, have echoed Sorel's doctrines, and students of all such movements still find useful his conceptions of myth and violence. Croce said that Sorel and Marx were the only original thinkers socialism ever had.

PHILOSOPHY OF SCIENCE

Sorel accepted Jean-Joseph-Marie-Auguste Jaurès's scornful description of him as "the metaphysician of socialism," for he thought of himself as primarily a philosopher, though not of socialism alone. Socialism engaged no more of his attention than the philosophy of science or the history of Christianity. Sorel's philosophy of science was technological rationalism: Scientific laws were accounts of the working of experimental machinery into which a part of nature, after being purified to make it homogeneous with the manmade mechanism, had been incorporated. There was no cause to suppose that such machines were models of nature's hidden mechanisms, and in fact there was no sign that determinism of any sort operated in nature left to herself. Determinism existed only where men created it, in machines that did violence to nature by shutting out chance interference. Thus, science is concerned with "artificial nature," the manmade phenomena of experiment and industry. It has nothing to say about "natural nature," where hazard, waste, and entropy are uncontrolled, where our knowledge is limited to statistical probability and our intervention to rule of thumb. Sorel accepted the pessimistic conclusions often drawn at that time from the second law of thermodynamics, to the effect that there was absolute

chance in nature and that the universe was "running down" to heat-death.

It was against that malevolent nature of chance and waste that humanity struggled in a hopeless effort of "dis-entropy," seeking to establish regions of determinism (experimental science) and of economy of forces (productive industry). Being a professional engineer, Sorel could work out these ideas in great technical detail. He even applied them to mathematics, saying that geometry was about architecture, not nature.

SOCIAL THEORY

Sorel's social theory derived from his philosophy of science. There are "entropic" trends in society comparable to those in nature. Culture is constantly threatened by a relapse into barbarism and disorder that would make history sheer meaningless succession. Against perpetual decadence men struggle heroically to establish limited zones of law, order, and cultural significance. To succeed in this for a time, they must do violence to their own natures by imposing on themselves a hard discipline and accepting moral isolation amid their mediocre fellows. This means living in conformity to "the ethic of the producers" and seeing the good life to be a cooperative creative enterprise carried on in a self-reliant spirit. Against this ethic stands "the ethic of the consumers," which takes the good to be things to be obtained rather than a way of acting. In the consumers' view typical goods are welfare, prosperity, distributive justice, and the classless society, things to be aimed at for the future and enjoyed if secured. Sorel replied that enterprises undertaken in that spirit were based on envy and inevitably fell under the control of adventurers (usually intellectuals) who duped the masses. He cited as instances slave revolts, peasant wars, Jacobinism, anti-Semitism, and contemporary welfare-state socialism.

In contrast, producers' movements concentrated on building the independent institutions that embodied their morality of productivity and solidarity. Such movements might be concerned with religious, artistic, scientific, or industrial activities, and Sorel took capitalism and syndicalist socialism as successive and equally admirable types of an industrial producers' movement. The workers were in revolt against capitalism not because of exploitation or inequality of riches (such matters concerned consumers only) but because the bourgeoisie had become unenterprising, cowardly, hypocritical—in a word, decadent. Until some more youthful, vigorous movement wrested social preeminence from the bourgeoisie (and Sorel did not think that socialism was the only contender), Western history would be a meaningless sequence of parliamentary deals and predatory wars. All movements "ran down" in the end, as their nerve failed, even (or especially) without challenge from a new movement. This succession of periods of heroic creativity and decadent barbarism did not constitute a true historical cycle, but Sorel adopted the accounts of the heroic and decadent phases of society given by Giambattista Vico in his cyclical theory. Sorel and Croce stimulated the revival of interest in Vico, and Sorel regarded his own social theory as a Viconian revision of Marxism.

VIOLENCE

Sorel is remembered less for his general philosophical system than for two notions lifted from it, violence and myth. Sorel found the syndicalists using violence during industrial strikes, and he set out to answer the common charge that a movement that resorted to violence was ipso facto evil and retrograde. He pointed out that Christianity and French republicanism, for example, had welcomed violent confrontations in order to mark clearly their rejection of the social milieu and their refusal to compromise. In such cases violence was a sign of moral health that frightened away lukewarm supporters and gave notice of earnest determination to adopt a new way of life. Physical violence—head breaking and bloodshed—was only one extreme of a range of vehement attitudes of which the other extreme was "a violence of principles," such as parading the least acceptable part of one's doctrines (in the Christian religion, miracles) to discourage one's "reasonable" friends. Sorel's theory of violence was intended to cover that whole range of attitudes, and the only special stress on physical violence was the statement that without being at all typical of social relationships, physical violence is a logical extreme from which no rising movement will shrink in certain unfavorable circumstances. Such circumstances would be confrontation with the armed force of a state that preached pacifism and social unity while it sought to smother a rebellious minority. The classic case was primitive Christianity, which could have secured tolerance within Roman polytheism but enthusiastically courted violent persecution to mark its unbridgeable differences with paganism. Parliamentary democracy was an even greater threat to independent social movements than polytheism had been to Christianity, because it claimed to have devised, in parliament, a perfect market where all social demands could be reconciled by elected representatives, thus ensuring social harmony. A movement that refused to come to that market because it wanted things other

than parliamentary seats and budget subsidies would have to be unequivocal, vehement, and even violent to escape from the nets of democratic prejudice. Most shocking of all, violence might be exercised not only against supposed enemies but against the men of good will, the peacemakers sent to befriend the minority and corrupt it into conformity.

Sorel's theory of violence caused scandalized misunderstanding among respectable people and some morbid enthusiasm among protofascists. Yet Sorel had not defended indiscriminate violence. He had said that since violence is ubiquitous in society, in the form of war and the enforcement of law and order, one could not selectively deplore violence on the part of an opposition without first looking to see who that opposition was. One should ask if it were associated, as so often in the past, with a progressive and heroic morality obliged to be ruthless to force recognition of its independence and to signify its rejection of mediocrity. Sorel noted that such movements built up sanguinary legends about how much violence they had known. Just as strikers exaggerated police brutalities committed on "our martyred dead," so the early Christians had endured far too little persecution to justify the tradition that the church was nourished by the blood of martyrs. Such violent tales were only symbolically true; a few clashes that proved a willingness to go to extremes had revealed the Christian community to itself and its enemies.

Last, Sorel argued (in 1908, when the seeds of world war, Bolshevism, and fascism were germinating) that Edwardian democrats were deluding themselves in thinking that civilized men had progressed beyond the stage at which they would use violence to promote or oppose causes. Violence would never be outgrown (and if it were, that would not be progress) because it was not, absolutely and in itself, brutish. It could be lucid, noble, and applied to the defense of high purposes; it could mark the birth of a new civilizing agency. Of course, it could also be bestial and oppressive, in which case Sorel called it force.

MYTH

Sorel found that myth was being used by the syndicalists, and he recalled similar uses from history. In no sense did he urge political activists to adopt extremist beliefs they knew to be false. That ambiguity, of which Sorel was accused, was really in the sociological facts themselves, he said. One found movements uttering views about the future without trying to establish their prophecies as scientifically plausible, without even caring to argue whether the forecasts were sound. They cared for those visions of the future passionately, but they cared for them only as inspiring pictures of what the world would be like if the new morality won all men's hearts. Such visions were myths, a present morality stated in the future tense. The case in point was the general strike. Syndicalists said socialism would come if all workers went on strike at once, whereupon the capitalist state would be paralyzed. Parliamentary socialists replied, reasonably enough, that for the workers to strike all at once and successfully defy the state, they would have to be ardent socialists to a man and the regime ripe for overthrow. But in that event socialism would already have arrived, and the general strike would not be needed. It was not a means to anything because it presupposed that all the problems were solved. Precisely this, answered Sorel, is the social function of the general strike. It is the dramatic picture of a morality triumphant. It is not a plan or scientific forecast, and therefore rational criticism of it is pointless. Besides, intellect has nothing better to put in its place, because the future is radically unpredictable and there is no science of the unknowable. A myth, being the expression of the aspirations of an enthusiastic mass of men and women, could well foreshadow something like itself, at least something equally sublime, whereas scientific blueprints for the future foreshadowed nothing but disappointment, the rule of intellectual planners, and the spread of the consumer outlook among those who waited for the planned good time to start. Granted that prevision is impossible, there are only two sorts of attitude toward the future—myths and utopias. Myths command respect as the product of intense social wills that could achieve something in history; utopias deserve scorn as the divagations of solitary intellectuals.

Sorel's tolerant view of myths and his anxiety to protect their improbabilities from rational examination were dependent on his conviction (drawn from Henri Bergson's philosophy) that the future is undetermined and thus totally unknowable. Few philosophers accept that position, and they would thus feel entitled to be more critical of myths than Sorel allowed. Yet he provided social theory with a valuable new concept—the galvanizing mass faith about which even its own believers are ambivalent, half admitting it to be improbable and yet clinging to it as the dramatic epitome of the cause they live for.

See also Bergson, Henri; Continental Philosophy; Croce, Benedetto; Labriola, Antonio; Lenin, Vladimir Il'ich; Marx, Karl; Marxist Philosophy; Masaryk, Tomáš Garrigue; Myth; Nationalism; Philosophy of Science, History of; Philosophy of Science, Problems of; Political

Philosophy, History of; Proudhon, Pierre-Joseph; Socrates; Vico, Giambattista; Violence.

Bibliography

WORKS BY SOREL

Contribution à l'étude profane de la Bible. Paris: Ghio, 1889.

Le procès de Socrate. Paris: Alcan, 1889.

La ruine du monde antique. Paris: Jacques, 1901.

Essai sur l'église et l'état. Paris: Jacques, 1902.

Saggi di critica del marxismo. Palermo: Sandron, 1902.

Introduction à l'économie moderne. Paris, 1903.

Le système historique de Renan, 4 vols. Paris: Jacques, 1905–1906.

Insegnamenti sociali della economia contemporanea. Palermo: Sandron, 1907.

Réflexions sur la violence. Paris: Librarie de "Pages Libres," 1908. Translated by T. E. Hulme and J. Roth as *Reflections on Violence.* New York: Huebsch, 1914.

Les illusions du progrès. Paris: Rivière, 1908. Translated by John and Charlotte Stanley as *The Illusions of Progress.* Berkeley: University of California Press, 1969.

La décomposition du marxisme. Paris: Rivière, 1908.

La révolution dreyfusienne. Paris: Rivière, 1909.

Matériaux d'une théorie du prolétariat. Paris: Rivière, 1919.

Les préoccupations métaphysiques des physiciens modernes. Paris, 1921.

De l'utilité du pragmatisme. Paris: Rivière, 1921.

D'Aristote à Marx. Edited by Édouard Berth. Paris: Rivière, 1935.

Propos de Georges Sorel. Edited by Jean Variot. Paris: Gallimard, 1935. Sorel's conversations.

From Georges Sorel: Essays in Socialism and Philosophy. Edited by John L. Stanley. New York: Oxford University Press, 1976.

From Georges Sorel, Vol. 2: *Hermeneutics and the Science.* Edited by John L. Stanley; translated by John and Charlotte Stanley. New Brunswick, NJ: Transaction, 1990.

WORKS ON SOREL

V. Delesalle, *Bibliographic sorélienne* (Leiden, 1939), lists hundreds of articles and reviews by Sorel scattered through dozens of journals, including material of philosophical interest, and records the extensive literature on Sorel up to 1939; J. H. Meisel, *The Genesis of Georges Sorel* (Ann Arbor, MI: Wahr, 1951), adds later items. Of the many hundreds of studies, the most instructive are Richard Humphrey, *Georges Sorel, Prophet without Honor* (Cambridge, MA: Harvard University Press, 1951); Victor Sartre, *Georges Sorel* (Paris, 1937); Michael Freund, *Der revolutionäre Konservativismus* (Frankfurt, 1932); Giuseppe Santonastaso, *Georges Sorel* (Bari, Italy, 1932); Pierre Andreu, *Notre maître, M. Sorel* (Paris: Grasset, 1953); Jean Deroo, *Le renversement du matérialisme historique* (Paris: Rivière, 1942); Max Ascoli, *Georges Sorel* (Paris, 1921); Georges Guy-Grand, *La philosophie syndicaliste* (Paris: Grasset, 1911); Gaétan Piron, *Georges Sorel* (Paris: Rivière, 1927); Georges Goriely, *Le pluralisme dramatique de Georges Sorel* (Paris: Riviere, 1962); Fernand Rossignol, *La pensée de Georges Sorel* (Paris, 1948); various authors, special issue of *Fédération* (November 1947).

ADDITIONAL SOURCES

Berlin, Isaiah. "Georges Sorel." In *Against the Current: Essays in the History of Ideas,* edited by Henry Hardy. New York: Viking Press, 1980.

Jennings, Jeremy R. *Georges Sorel; The Character and Development of His Thought.* London: Macmillan, 1985.

Kolakowski, Leszek. "Georges Sorel: A Jansenist Marxism." In *Main Currents of Marxism,* Vol. 2, translated by P. S. Falla. Oxford: Clarendon Press, 1978.

Stanley, John L. *The Sociology of Virtue: The Political and Social Theories of Georges Sorel.* Los Angeles: University of California Press, 1981.

Vernon, Richard. *Commitment and Change: Georges Sorel and the Idea of Revolution.* Toronto: University of Toronto Press, 1978.

Vincent, Steven K. "Interpreting Georges Sorel: Defender of Virtue or Apostle of Violence?" *History of European Ideas* 12 (2) (1990): 239–257.

Neil McInnes (1967)
Bibliography updated by Philip Reed (2005)

SOSA, ERNEST
(1940–)

Ernest Sosa is Romeo Elton Professor of Natural Theology and Professor of Philosophy at Brown University and regular Distinguished Visiting Professor at Rutgers University. He received his Ph.D. from the University of Pittsburgh and has taught at Brown since 1964. Since 1983, he has been the editor of *Philosophy and Phenomenological Research* and since 1999, with Jaegwon Kim, the co-editor of *Nous*. Sosa has published essays on issues in a wide variety of philosophical areas such as metaphysics, logic, philosophy of mind, theory of action, and philosophy of language, but he has been most influential in epistemology, where he is known for advocating a virtue-based approach to the analysis of knowledge and justification with an emphasis on the importance of a reflective perspective.

What is distinctive of virtue epistemology is the order of explanation: A belief's epistemic status is to be understood in terms of the epistemic properties of the subject, which in turn are to be captured by employing the concept of an intellectual virtue. How is this concept to be understood? In pure virtue epistemology, construed in analogy to pure virtue ethics, the concept of an intellectual virtue is basic (Foley 1994). Sosa, however, conceives of an intellectual virtue as a stable disposition to form true beliefs in a certain field of propositions, F, under suitable circumstances, C. Thus his brand of virtue epistemology, which he has labeled virtue perspectivism,

is not an example of the pure kind but may be viewed as a form of reliabilism.

The two main elements of Sosa's virtue perspectivism are the concepts of an intellectual virtue and an epistemic perspective. As already indicated, Sosa conceives of intellectual virtues in terms of reliability. Reliably functioning faculties, such as vision, hearing, introspection, and memory are examples of intellectual virtues. Sosa calls beliefs that are grounded in the exercise of such virtues apt. Apt beliefs, if true, qualify as knowledge, or, more precisely, as animal knowledge, to be distinguished from reflective knowledge. With the distinction between these two kinds of knowledge, the second main element of virtue perspectivism comes to the fore: the concept of an epistemic perspective.

Let S refer to the subject whose beliefs we wish to evaluate. Suppose S's visual belief that p is true and, due to the reliability of S's vision, apt. Hence by employing her faculty of vision, S acquires animal knowledge that p. For S's belief to rise to the level of reflective knowledge, a further condition must be met: S must form a meta-belief to the effect that her belief and its being true have their origin in a reliable faculty. In general terms, if from S's epistemic perspective, a faculty is coherently viewed as reliable within field F and circumstances C, then by employing this faculty S can acquire reflective knowledge within field F and circumstances C.

Animal knowledge, then, results from external aptness: the exercise of faculties that are in fact reliable. Reflective knowledge also requires aptness, but, in addition, the adoption of an internally coherent perspective with respect to the reliability of one's faculties. Sosa's virtue perspectivism, then, combines both an externalist and an internalist element.

In the large body of work in which Sosa articulates and defends his approach to the philosophical explanation of knowledge and justification, he has addressed various problems that arise for virtue perspectivism. First, there is the problem of what a reliabilist should say about what are referred to as *evil demon victims*: subjects whose beliefs seem justified although, due to the massive deception to which the victims are subjected, their beliefs are grounded in unreliable faculties. Sosa responds that, whereas the demon victims' beliefs are *actual world* justified (as the victims' faculties would be reliable in the actual world, they are *same world* unjustified because the faculties the victims employ are unreliable in their own world (1994a). In more recent terminology, Sosa classifies the victims' beliefs as adroit though not apt (Sosa 2003).

Second, Sosa's reliability-grounded virtue perspectivism is challenged by what Sosa calls the *problem of meta-incoherence*, which arises from cases in which a subject's beliefs are produced by a faculty whose de facto reliability is not (or at least not yet) recognized by the subject. Since such subjects do not meet the perspectival condition of having formed reliability-attributing meta-beliefs about the relevant belief sources, Sosa judges that the beliefs in question are unjustified, or not reflectively justified (Sosa 1991).

Third, there is the generality problem, which for Sosa amounts to the challenge of finding the right level of specificity in describing field and circumstances. Here, Sosa's solution is to require that the relevant descriptions be useful within the subject's epistemic community and to the subject herself (Sosa 1991). Three further, important problems to which Sosa has articulated detailed solutions are the following: First, how can we distinguish between accidental and non-accidental reliability? Second, what justifies reliability-attributing perspectival meta-beliefs (Sosa 1994a)? Third, why is the process by which reliability-attributing meta-beliefs are formed (using, for example, perception to attest to the reliability of our perceptual faculties) not viciously circular (1994b and 1997)?

Recently, Sosa has also contributed important work on the following question: If a belief is to be an instance of knowledge, what modal link must there exist between the belief and its truth? According to some, knowledge requires sensitivity: S would not believe that p if p were false. Viewing this condition as too demanding, Sosa objects to it on the basis of the following case: Having dropped a trash bag in the garbage shoot, you believe the bag will momentarily reach its destination in the basement. This belief, Sosa suggests, amounts to knowledge even though it is not sensitive: if p (the bag will land momentarily) were false (because, say, the bag snagged in the shoot), you would still believe p. As an alternative, Sosa proposes safety: If S were to believe p, p would be true (or: Not easily would S believe incorrectly in believing that p). Though your belief that the bag will land momentarily is not sensitive, it is indeed safe, for possible worlds in which S believes that the bag will shortly arrive downstairs, but believes this mistakenly, are indeed remote (Sosa 1999).

The distinction between safety and sensitivity assumes particular significance for Sosa, for he appeals to it for the purpose of rejecting the contextualist solution to the puzzle of skepticism. Contextualists have argued that, when confronted with a skeptical argument, we face

a paradox because, although we find the premises plausible, we wish to reject the conclusion. According to the contextualist response, the puzzle is to be solved by appeal to the context-sensitivity of the word *know*. Sosa suggests an alternative solution: Skeptical arguments may (misleadingly) seem cogent because we fail to recognize that knowledge requires not sensitivity, but merely safety (Sosa 1999 and 2003).

See also Contextualism; Kim, Jaegwon; Moral Epistemology; Reliabilism; Skepticism, Contemporary; Virtue Epistemology.

Bibliography

PRIMARY WORKS

Knowledge in Perspective: Selected Essays in Epistemology. Cambridge,U.K.: Cambridge University Press, 1991.

"Philosophical Skepticism and Epistemic Circularity." *Aristotelian Society Supplementary Volume*, 68 (1994a): 263–290.

"Virtue Perspectivism: A Response to Foley and Fumerton."*Philosophical Issues*, 5 (1994b): 29–50.

"Perspectives in Virtue Epistemology: A Response to Dancy and BonJour." *Philosophical Studies* 78 (1995): 221–235.

"Reflective Knowledge in the Best Circles." *The Journal of Philosophy* 96 (1997): 271–293.

"How to Defeat Opposition to Moore." *Philosophical Perspectives* (1999) 141–153.

"Relevant Alternatives, Contextualism Included" *Philosophical Studies* 119 (2003): 35–65.

SECONDARY WORKS

BonJour, Laurence. *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtue*. Malden, MA: Blackwell, 2003

BonJour, Laurence. "Sosa on Knowledge, Justification, and 'Aptness.'" *Philosophical Studies* 78 (1995): 207–220.

Dancy, Jonathan. "Supervenience, Virtues, and Consequences." *Philosophical Studies* 78 (1995): 198–205.

Foley, Richard. "The Epistemology of Sosa." *Philosophical Issues* 5 (1994): 1–14.

Fumerton, Richard. "Sosa's Epistemology." *Philosophical Issues* 5 (1994): 14–27.

Greco, John. *Ernest Sosa and His Critics*. Malden, MA: Blackwell, 2004.

Matthias Steup (2005)

SOTO, DOMINIC DE

(1494–1560)

Dominic de Soto, the Dominican scholastic theologian, was born at Segovia, Spain, and died at Salamanca. He studied at Alcalá de Henares and became a professor of philosophy there after advanced studies at the University of Paris. Entering the Dominican order in 1524, Soto taught theology from 1525 onward at the University of Salamanca. He was very active in the deliberations of the Council of Trent. Soto's writings include two commentaries on Aristotle (*In Dialecticam Aristotelis,* Salamanca, 1543; *In Libros Physicorum,* Salamanca, 1545). Theological works containing some philosophical thought are *Summulae* (4 vols., Burgos, 1529); *De Natura et Gratia* (Venice, 1547); and the treatise *De Justitia et Jure* (Justice and the law; Salamanca, 1556).

One of the founders of the school of Spanish Thomism, Soto had his own opinions on many philosophical questions. Like John Duns Scotus, he denied the usual Thomistic distinction between essence and existence. In theory of knowledge, he also showed the influence of Scotism, teaching that the primary object of human understanding is indeterminate being in general. His psychology followed that of Thomas Aquinas, with strong emphasis on the intellectual functions: the intellect is a nobler power than the will. Soto is an important figure in the philosophy of law and politics. He violently criticized the theory of the state of pure human nature, as popularized by Cardinal Cajetan and Francisco Suárez. Unlike his teacher, Francisco de Vitoria, Soto taught that law stems from the understanding rather than from the will of the legislator; he clearly differentiated natural law, which depends on the real natures and relations of things, from positive law, which results from a decision of the legislator (*De Justitia* I, 1, 1). In political philosophy he represents a growing tendency toward democratic thinking in Renaissance scholasticism: Both civil and ecclesiastical power derive ultimately from God, but the civil power proceeds through the medium of society; the people concretize the authority received from God in the persons whom they designate as rulers. Soto is also regarded as one of the founders of the general theory of international law.

See also Aristotle; Cajetan, Cardinal; Philosophy of Law, History of; Philosophy of Law, Problems of; Renaissance; Scotism; Suárez, Francisco; Thomas Aquinas, St.; Thomism; Vitoria, Francisco de.

Bibliography

De Justitia et Jure has been translated into Spanish by Jaime T. Ripoll as *Tratado de la justicia y del derecho.* 2 vols. (Madrid, 1926).

For literature on Soto, see A. J. Carlyle, *History of Mediaeval Political Theory* (Edinburgh, 1950), Vol. VI, pp. 254–258; T. Davitt, *The Nature of Law* (St. Louis: Herder, 1951), pp. 161–177; and Beltrán de Heredia, "El maestro Domingo

(Francisco) de Soto," in *La ciencia tomista* 43 (1931): 357–373.

Vernon J. Bourke (1967)

SOUL

See *Immortality; Psychē; Psychology*

SOUND

"Sound" according to Aristotle's *De Anima* (418a12) and George Berkeley's *First Dialogue*, is the special, or proper, object of hearing. G. J. Warnock, in his *Berkeley*, interprets this as meaning that sound is the "tautological accusative" of hearing: Sounds can only be heard and must be heard if anything is heard.

Hearing receives attention in philosophy mainly for its differences from seeing. Two respects in which listening and hearing differ from looking and seeing are (1) that there is nothing analogous, in seeing, to hearing the sound of something, and (2) that, in telling where something is, there is nothing analogous, in listening, to our having to look in the right direction.

Warnock's explanation of the first of these differences is that we establish the presence and existence of an object by sight and touch, and then proceed to distinguish the object thus established from its smell and taste and the noises it makes. He mentions, as reasons for not ascribing such primacy to hearing, that inanimate objects often do not make any noises, that animate ones make them only intermittently, and that it is often difficult to tell where a sound is coming from. There would be a further reason if, as P. F. Strawson maintains (in *Individuals*, p. 65), a universe in which experience was exclusively auditory would have no place at all for spatial concepts. This reason would be decisive if in a nonspatial world there could be no concept of an object (*Individuals*, Ch. 2). Strawson asserts that we can discover some spatial features of things by listening (for instance, sounds seem to come from the left or right), but denies that such expressions as "to the left of" have any intrinsically auditory significance. In accordance with this, G. N. A. Vesey labels knowing where a sound comes from by listening "borrowed-meaning" knowledge. Berkeley makes use of the fact that we talk of hearing sounds caused by things, together with the principle that "the senses perceive nothing which they do not perceive immediately: for they make no inferences," to gain acceptance of the view that we cannot properly be said to hear the causes of sounds.

We can see directly (otherwise than by reflection) only what is on the same side of our heads as our eyes. Knowing in what position we have had to put our heads—in what direction we have had to look—to see an object, we know in what direction the object is. Hearing is not limited in this fashion, and so we identify the position of a merely seen object and a merely heard object very differently. Furthermore, if Strawson and Vesey are right about spatial expressions not having an intrinsically auditory significance, we cannot hear that one object is to the left of another as we can see that one object is to the left of another. It might be concluded that knowledge that the source of a sound is to one's left, gained by listening, must be mediated knowledge—that is, must have involved the making of an inference. To be valid, this conclusion would require the further premise that acquiring a perceptual capacity is invariably a matter of learning to interpret one thing as a sign of another. An alternative hypothesis would be that the only interpretation involved is at the physiological level; that is, that differences in the stimuli to the two ears which, in a person whose experience was exclusively auditory, would have no counterpart in experience, would, in a person who knew what it was to see and feel things as being on his left or right, subserve his hearing things as being on his left or right.

B. O'Shaughnessy ("The Location of Sound") asserts that hearing where a sound comes from is noninferential and immediate. He contends that the seeming mysteriousness of the fact that listening can tell us where a sound is coming from is the result of our thinking of what is heard as a complex of two elements, "the sound itself" and "its coming from the left" (defining "the sound itself" as what is auditory—evidence of a "metaphysical theory of the sensory substratum"), and then having to think of its coming from the left either as "part and parcel of the sound" or as something we experience "other than and additional to the sound itself" but somehow related to it. That the sound is coming from the left, O'Shaughnessy holds, is neither part of the sound, nor something else we experience; nor is it something "we simply know." The mistake lies in our thinking of what is heard as a complex, and O'Shaughnessy sees this as a result of our having "the idea that a thought or meaning is a complexity."

Sound is a Lockean secondary quality. Hylas, in Berkeley's *First Dialogue*, accordingly distinguishes between sound as it is perceived by us ("a particular kind of sensation") and sound as it is in itself ("merely a vibrative or undulatory motion in the air"). Consideration of this

philosophical position would not seem to raise issues peculiar to sound.

Bibliography

Grice, H. P. "Some Remarks about the Senses." In *Analytic Philosophy.* 1st series. Oxford: Blackwell, 1962.
Jonas, Hans. "The Nobility of Sight: A Study in the Phenomenology of the Senses." *Philosophy and Phenomenological Research* 14 (4) (June 1954): 507–519.
Malpas, R. M. P. "The Location of Sound." In *Analytic Philosophy*, edited by R. J. Butler, 131–144, 2nd series. Oxford, 1965.
O'Shaughnessy, B. "An Impossible Auditory Experience." *PAS* 57 (1956–1957): 53–82.
O'Shaughnessy, B. "The Location of Sound." *Mind* 66 (264) (October 1957): 471–490.
Pasnau, Robert. "What Is Sound?" *Philosophical Quarterly* 49 (1999): 309–324.
Strawson, P. F. *Individuals*. London: Methuen, 1959. Ch. 2.
Swartz, R., ed. *Perceiving, Sensing, and Knowing*. Berkeley: University of California Press, 1965.
Urmson, J. O. "The Objects of the Five Senses." *Proceedings of the British Academy* 54 (1968): 117–131.
Vesey, G. N. A. "Knowledge without Observation." *Philosophical Review* 72 (2) (April 1963): 198–212.
Warnock, G. J. *Berkeley*, 33–36. Harmondsworth, U.K., 1953.

G. N. A. Vesey (1967)
Bibliography updated by Benjamin Fiedor (2005)

SOUTH AMERICAN PHILOSOPHY

See *Latin American Philosophy*

SOVEREIGNTY

Analysis of "sovereignty" brings one into contact with nearly all the major problems in political philosophy. At least seven related concepts may be distinguished:

(1) A person or an institution may be said to be sovereign if he or it exercises authority (as a matter of right) over every other person or institution in the legal system, there being no authority competent to override him or it. For some writers, though not for all, this concept also implies unlimited legal competence; for, it is said, an authority competent to determine the limits of its own competence must be omnicompetent. (2) Difficulties arising from the first concept have led some writers to ascribe sovereignty to a constitution or basic norm from which all other rules of a system derive validity. (3) Sovereignty is sometimes ascribed to a person, or a body or a class of persons, said to exercise supreme power in a state, as distinct from authority, in the sense that their wills can usually be expected to prevail against any likely opposition.

The state itself is often said to be sovereign. This may mean any of at least four distinct (though possibly related) things: (4) that the state as an organized association will in fact prevail in conflict with any person or any other association in its territory; (5) that the rights of all such associations and persons derive from the legal order that is supported by the state or that (according to Hans Kelsen) *is* the state; (6) that the state is a moral order with claims to obedience and loyalty which have precedence over all others; (7) that the state is autonomous vis-à-vis other states; according to some theories, the state has only such obligations, whether in law or in morals, as it chooses to recognize.

CLASSICAL AND MEDIEVAL THEORIES

Aristotle regarded legislative authority as supreme in a state and classified states according to whether it was located in a monarch, in an oligarchical assembly, or in an assembly of the whole people. But to speak of a "supreme legislative authority" is a little misleading here; for the Greeks, legislation was the local application of a divinely ordained order, rather than the authoritative creation of new laws. The Roman concept of imperium was nearer sovereignty: The princeps (ruler) personally embodied the supreme authority of the Roman people. He was *legibus solutus* (not bound by the laws), at least in the sense that no one could question his enactments. Still, there were strong elements of natural law in Roman jurisprudence; the emperor was supreme because his function was to command what was *right* and for the public good.

There was rather less room for sovereignty in medieval political thought. According to Thomas Aquinas, for instance, the king was not only subject to divine and natural law but for most purposes to the custom of his realm as well. Medieval statutes commonly purported to restore laws that had been abused, rather than to innovate. In Thomas's view the Roman maxim *Quod principi placuit legis habet vigorem* (What pleases the prince has the force of law) was valid only if the prince's command was reasonable. According to Henry de Bracton, "the king ought to have no equal in his realm … [but] he ought to be subject to God and the law, since law makes the king … there is no king where will rules and not the law" (*De Legibus et Consuetudinibus Angliae,* edited by G. Woodbine, New Haven, CT, 1915–1942, Vol. II, pp. 32–33). Similarly, the *plenitudo potestatis* ascribed

to the pope usually meant that supreme ecclesiastical authority was undivided, or that he held a reserve jurisdiction in secular matters—not that he was *legibus solutus.*

Alongside the doctrine of royal supremacy was another that derived royal authority from the people corporately. According to Marsilius of Padua, supreme authority rested in the *legislator,* which was either the whole organized community or an assembly (not necessarily elected) that spoke for it. Marsilius's stress on legislation as the will of a supreme authority brought him closer than his predecessors to Jean Bodin and Thomas Hobbes.

BODIN: PARADOX OF LAWFUL SOVEREIGNTY EGOISM

Bodin's *Six livres de la république* (1576) is generally considered the first statement of the modern theory that within every state there must be a determinate sovereign authority. Writing during the French religious wars, he insisted that an ordered commonwealth must have a sovereign competent to overrule customary and subordinate authorities. Sovereignty is "a supreme power over citizens and subjects unrestrained by law"; it is "the right to impose laws generally on all subjects regardless of their consent." Law is "nothing else than the command of the sovereign in the exercise of his sovereign power." Accordingly the sovereign could be subject to no one else, for he makes the law, amends it, and abrogates it for everyone. Nevertheless, he is subject to the laws of God and of nature. For instance, he may not seize his subjects' property without reasonable cause and must keep his promises to them. Moreover, he must respect the fundamental laws of the constitution, like the succession law, for sovereignty, as a *legal* authority, stems from these.

In defining sovereignty as a supreme power unrestrained by law, while yet admitting these limitations, Bodin is not as inconsistent as he is commonly said to be. Within the legal system, sovereignty may be unlimited; yet the sovereign may be bound in morals and religion to respect the laws of God and nature. Bodin's suggestion that sovereignty can be limited by constitutional laws raises more serious difficulties; for if "law is nothing else than the command of the sovereign, in the exercise of his sovereign power," how can any law be beyond his power to amend? The qualification, "in the exercise of sovereign power," may be important. Constitutional laws seem to be what H. L. A. Hart calls "rules of recognition" (see his *Concept of Law*), that is, they are rules that lay down the criteria of validity for rules of substance; they constitute the sovereign office, designate who shall occupy it, and identify his acts as those of a sovereign authority. For the sovereign to interfere with them, Bodin said, would be for him to undermine his own authority. If the acts of the sovereign are those done "in the exercise of sovereign power," that is, in accordance with the rules of recognition, it would be logically impossible to act in a valid sovereign way inconsistently with these rules. Nevertheless, the sovereign could still amend them so long as he used the unamended procedures to do so. Yet Bodin regarded the rules constituting the sovereign office as unamendable in principle; should the prince infringe them, "his successor can always annul any act prejudicial to the traditional form of the monarchy since on this is founded and sustained his very claim to sovereign majesty" (all quotations from *Six Books,* Bk. I, Ch. 8).

Bodin's reasoning, though confused, bears closely on certain twentieth-century constitutional controversies in the United Kingdom and Commonwealth countries, which have hinged on the contention that a sovereign legislature, though admittedly competent to prescribe its own powers and procedures, must yet do so only by the procedures currently laid down. Such procedures, it is argued, are among the criteria for identifying the legislature and for determining what constitutes one of its acts. Bodin's analysis of sovereignty also suggests how an omnicompetent authority like the British Parliament can yet limit its omnicompetence, as it purported to do in the Statute of Westminster of 1931. In that statute it renounced supreme authority over the dominions by making their advice and consent part of the procedure for any future legislative acts affecting them.

HOBBES: SOVEREIGNTY AND SUPREME POWER

Where Bodin was concerned mainly with supreme legal authority, Hobbes was more concerned to show a necessary relation between order, political power, sovereign authority, and political obligation. Hobbes argued that since no man can safely rely on his own strength or wits alone, men's obligations under the law of nature to forbear from harming one another must be subject to mutual guarantees; otherwise, for anyone to forbear in the competitive struggle would be to endanger his life. There is no reliable guarantee unless all parties agree not to exercise their "natural right to all things," but to submit unconditionally to a sovereign authorized to act on behalf of each of them, with the power to make them keep their agreements. Mutual forbearance would then be a duty. Sovereignty, therefore, is necessary for a social order

among equals. Sovereignty cannot be made effectively subject to conditions without depriving it of its point; for on whom could be conferred the authority to judge whether such conditions had been violated? If on the individual subjects, no one individual could rely on the submission of any other. If on the sovereign, the conditions themselves would be merely formal. And there could be no independent arbiter, for any independent arbiter who could impose his ruling would himself be sovereign. Sovereignty is likewise indivisible, for if anyone had the power to mediate effectively in conflicts of authority, he would be sovereign. The united strength of all is therefore the sovereign's to use as he thinks fit. His duties under God and natural law are strictly God's business. The subject, having freely surrendered the right to interpret the law of nature for himself, must accept the sovereign's pronouncements on what is right and wrong. He could, however, be under no obligation to take his own life or to submit willingly if the sovereign should seek to kill him. Both commitments would be unnatural, being contrary to the supreme end, which is to avoid sudden death; and having no sanction in reserve, the sovereign would have no way of enforcing either obligation.

The sovereign remains one only so long as "the power lasteth, by which he is able to protect" his subjects. The purpose of submission is protection; protection requires overwhelming power; so overwhelming power is the actual condition for supreme authority. Conversely, supreme authority, brooking no rivals, commanding the power of everyone, wields supreme power. Further, natural law enjoins us to keep our covenants, above all the covenant establishing the civil order. In its concrete political expression, natural law is identical with the command of the sovereign and therefore with the civil law. So the sovereign authority is also the supreme moral authority.

JOHN AUSTIN AND THE IMPERATIVE THEORY OF LAW

The imperative theory of law expounded by Hobbes was developed by Jeremy Bentham to disarm opponents of legal reform who treated natural law and morality as built-in justifications of the unreformed common law. For if, as Bentham argued, law were simply whatever the sovereign commanded, or, in the case of the common law, what he chose not to rescind, then it might be reformed by command in accordance with rational principles of utility. In the hands of Bentham's disciple John Austin the theory of sovereignty became a tool for juristic analysis. "Law properly so-called" was distinguished from rules of other kinds as a "rule laid down for the guidance of an intelligent being by an intelligent being having power over him." Within any legal system there must be one supreme power, "a *determinate* human superior, *not* in a habit of obedience to a like superior (receiving) *habitual* obedience from the *bulk* of the society" (*Province*). His will was the ultimate validating principle of law; otherwise the quest for validity would lead to an infinite regress. Austin avoided it by resting sovereignty on the sociological fact of obedience.

The English Parliament, which is subject to legal limitation or restraint by no other authority is, prima facie, the paradigm of a sovereign legislature. Yet if its will is law, that is because law makes it so. Moreover, it is the law that defines the conditions for determining what that will is. For an institution has a will only by analogy; it is constituted by the decisions of individuals playing roles defined by rules. A change in the rules might change the will, though the individual decisions remained the same. Austin himself falters, admitting that to identify the members of the sovereign Parliament would require a knowledge of the British constitution. Habitual obedience, in short, may be rendered not to determinate individuals but to an institution, which is a legal creation. In the United States supreme legislative authority rests in the constitutional amending organ—composed of the two houses of Congress, each acting by a two-thirds majority, plus three-quarters of the states, acting through their legislatures or by conventions. So complex, discontinuous, and impersonal an authority cannot enjoy habitual obedience; its authority, like its very being, presupposes the law. To say that the law is what it commands, simply because it is formally competent to annul any rule, is to use "command" in a very strained sense.

In any case, there could be a constitution without an amending organ that nevertheless could allocate areas of competence to a number of organs. All authorities would then be limited. If one could still speak of sovereignty, it would be divided among them, with no "determinate human superior"; each would be supreme in its own sphere. The notion that sovereignty must be indivisible and omnicompetent is a corollary, then, of the false theory that every law is an enforceable command. Federal states retain their character not because their component institutions obey a sovereign authority able to enforce its will but because there is a general disposition to conform to accepted rules and in cases of dispute to accept the arbitration of the courts. The latter, however, being formally incompetent to legislate, cannot themselves be the requisite Austinian common superior.

The imperative theory was in part an attempt to determine the conditions that a legal system must satisfy if rules valid within the system are to be identifiable and conflicts of rules resolved. An alternative answer, however, is that every system must have what Hans Kelsen called a *Grundnorm* (a basic law), which is "the supreme reason of validity of the whole legal order" and which gives it its systematic unity. In these schematic analyses of legal systems, the basic law (usually a constitution) and the Austinian sovereign have very similar functions. Some writers indeed have transferred the concept of sovereignty from rulers to constitutions, thus abandoning the imperative theory. This either leaves a purely structural analysis of a legal order or it substitutes for Austin's "habitual obedience" respect for the constitution as the sociological starting point.

SOVEREIGNTY AND POLITICAL POWER

As Austinian analyses of sovereignty became metalegal and remote from political facts, attempts were made to split, not indeed the sovereign, but the concept of sovereignty into two types: legal and political (or practical). The first would be attributable to the supreme legislature; the second to the class or body in the society that "could make [its] will prevail whether with or against the law" (James Bryce) or "the will of which is ultimately obeyed by the citizens" (A. V. Dicey). In a democracy this would normally be the people, or the electorate.

The notion of sovereignty as supreme power in the latter sense, however, suggests certain problems. First, one must generally take account not only of what one can do by oneself but also of other people's possible resistance or cooperation. No one can ever do just what he wants; even the supreme army commander must keep the troops loyal. Every social choice is between only those alternatives that the powers of other men leave open. Political decisions reflect not only actual pressures but also those that might be anticipated were things decided differently. Again, a group may exercise very great power in that policy sphere in which it has an interest as a group; but in others its members' interests may be diverse and conflicting, and there may be quite different configurations of interests and pressures. This does not mean that there could never be a particular group strong enough to get its way regardless of counterpressures, and with group interests spanning most of the important areas of policy. Even so, many political scientists see decisions emerging not from the domination of any one particular will or group interest but rather from an interplay of interests and pressures. In their view, the concept of supreme power simply suggests the wrong model. At best the concept would mean that in the search for explanations one need not look outside the internal politics of the supreme group; other groups could safely be ignored.

SOVEREIGNTY AS MORAL SUPREMACY: ROUSSEAU

The transposition of the concept of sovereignty from the context of seventeenth-century and eighteenth-century despotisms to the modern, popularly based state accounts for many of the perplexing features of the concept. The sovereign was then a king by divine right who at his strongest was subject to very few restraints and no legal limitations and to whom, it was said, his subjects owed unconditional obedience as a moral and religious duty.

Jean-Jacques Rousseau shifted sovereignty from the king to the people, which was now to exercise supreme power, somewhat paradoxically, over itself. For Rousseau, the citizens of a state had put themselves freely but unconditionally "under the supreme direction of the general will." And he radically altered the emphasis of the old doctrine that the people is the source of supreme authority by suggesting that the general will would be authentic and binding only if every citizen participated equally in expressing it. Moreover, since its object was the common good, there could be no higher claim on the citizen; he realized his own highest ends in total submission to it. As a legislating participant and a beneficiary of the moral order sustained by the general will, he attained freedom, not in the unrestricted slavery of impulse and appetite, but in obedience to a moral law that he prescribed to himself. It is true that Rousseau did not identify the will of all with the general will. The latter would be expressed only if the citizens addressed themselves to the question Wherein does the common good lie?, not to the question What would suit me personally? Democracy, too, can be corrupt, and the state in decay.

From Rousseau on, to ascribe sovereignty to the people was not (or not only) to state a political fact or a legal theory but to make a moral claim. Moreover, Rousseau reshaped the whole conceptual order of politics when he wrote that "the public person" created by the act of political association "is called by its members *State* when passive, *Sovereign* when active, and *Power* when compared with others like itself. Those who are associated in it take collectively the name of *people,* and severally are called *citizens,* as sharing in the sovereign power, and *subjects,* as being under the laws of the state" (*Social Contract,* Bk. I, Ch. 7). It was the citizen, not the king, who might say,

henceforth, *L'état, c'est moi.* Consequently, the object of the state, if not corrupted by tyrants or by selfish sectional interests, was a good in which all its members might participate on terms of justice and equality. Its sovereignty amounted to a claim to override, in the name of the public interest, all lesser associations and interests.

THE STATE OF HEGELIAN IDEALISM

Rousseau was hostile to sectional associations as rivals to the general will; G. W. F. Hegel accepted them as partial expressions of, or vehicles for, the more inclusive Idea that was the state. The state's sovereignty lay in its moral preeminence over all other forms of human association. As the highest stage in the moral evolution of man, the state embodied concretely, as a living institution, man's autonomous, rational will. Man progressed dialectically through the conflict of states, the most vigorous and forward-looking state taking the leadership of humanity from the aging and debilitated and setting its own mark on a new age. The state was sovereign, therefore, in its relations with other states because it owed them nothing; its highest moral commitment was to its own survival as the agent of history, which alone could judge its works.

CRITICS. The Hegelian view of sovereignty was challenged early in the twentieth century by political and legal theorists and historians, such as Otto von Gierke, Hugo Krabbe, Léon Duguit, F. W. Maitland, J. N. Figgis, and H. J. Laski. They substituted a pluralistic for the monistic model of the state. They held that state and society must be distinguished; that society is made up of many associations, each serving its own range of human needs and interests. They denied that the state's moral purpose, whether ideal or actual, gives it a special claim on the allegiance of its members, overriding the churches' claim on those of them who are believers, or the unions' on those of them who are workers. In a given situation, a church might mean even more to believers than the state. Moreover, the suggestion that the corporate legal status and existence of associations depends on state recognition was vigorously repudiated. Associations came into existence to fulfill needs the state could not satisfy.

According to Duguit, the existence and corporate rights of associations and, indeed, law itself were social facts that the state simply registered; it did not create them. According to Figgis and Laski, the state's claim to regulate the constitutions, aims, and internal relations of other associations was an invasion of their corporate moral autonomy. Each was strictly sovereign in its own sphere. The pluralists conceded that the state must continue, but as an umpire, maintaining the minimal conditions of order, determining conflicts of jurisdiction, and protecting members of one association from the encroachments of another. Hobbes would certainly have interpreted this as an admission of the need for a single sovereign authority; for as arbiter, the state must have the power to judge what is an encroachment and therefore the powers of review and disallowance. Enjoying an overriding authority, the state could not be merely one among others. Despite Duguit, the law must ultimately be determined by state officials. For Kelsen, who identified state and law, corporations are necessarily subsystems within the state system, since their rules have legal effect only by the state's extending recognition to them. But, of course, the same could conceivably be said, in reverse, of other associations. For instance, the state could just as well be seen from a religious standpoint as encapsulated within the greater religious and moral order sustained by the church.

SOVEREIGNTY IN INTERNATIONAL RELATIONS

Is state sovereignty consistent with international law? In Hobbes's view, states confront one another in the posture of gladiators—lacking a common superior, they could not be subject to any law. Austin regarded international law as a kind of positive morality; without a sovereign, it could not be "law properly so-called." Attempts have been made to get around this difficulty by what Georg Jellinek termed *auto-limitation*: International law is binding because sovereign states have imposed it on themselves. The relation between international law and a municipal legal order can be expressed, in Kelsen's terms, as follows: Seen from the standpoint of a municipal legal order, international law is validated in a self-subsistent municipal legal system by the *Grundnorm* of that system, in other words, by being received into the system. Kelsen repudiated this conclusion, however, because he wanted to maintain that there is one all-inclusive world of law and that international law itself provides the principles validating the laws of so-called sovereign states as subsystems. But one could as well describe the one world of law from the standpoint of any legal system one chose, on the condition that it recognized other legal systems. For each system could encapsulate the rest, including international law.

Article 2 of the United Nations Charter claims that the organization is based on the sovereign equality of all members. This must surely mean that states are sovereign if, unlike colonies or trust territories, they are not liable to

have any binding obligations laid upon them by other states without their consent. If international law is really a legal system, however, it cannot mean that a state has obligations only if, and for as long as, it chooses. For then there is no law. The notions of unlimited competence or overriding authority associated with "sovereignty" in a state's internal relations are out of place here. A sovereign state in international law must therefore be a particular kind of legal personality, like corporations in municipal law, with characteristic powers, rights, immunities, and obligations, including those implied in the principle of equality—namely, freedom from interference in its domestic jurisdiction, and, in the absence of an international legislature, immunity from new obligations except by consent. Nevertheless, states are considered bound by the established law and custom of nations, and the obligations of new states date from their inception and do not wait upon any consent or deliberate act of acceptance.

Finally, the alleged equality of sovereign states is not, of course, equality in power. Sovereignty in law is consistent with a large measure of actual control over a state from outside, though a minimum of independence might be a qualifying condition for sovereign status. Even the most powerful state, however, cannot ignore altogether the need to placate its friends and to avoid provoking its foes to the point of inconvenient obstruction. Freedom to act is relative in international as in internal affairs.

See also Aristotle; Austin, John; Bentham, Jeremy; Bodin, Jean; Democracy; Hart, Herbert Lionel Adolphus; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Kelsen, Hans; Marsilius of Padua; Natural Law; Political Philosophy, History of; Social and Political Philosophy; Rousseau, Jean-Jacques; Society; State; Thomas Aquinas, St.

Bibliography

Dias, R. W. M. *A Bibliography of Jurisprudence.* London: Butterworth, 1964. A valuable annotated bibliography. For sovereignty, see Chs. 4 and 14.

HISTORY OF THE CONCEPT

Bennett, W. H. *American Theories of Federalism.* University: University of Alabama Press, 1964. Largely concerned with the concept of sovereignty in history of U.S. constitutional theories.

Cohen, H. E. *Recent Theories of Sovereignty.* Chicago: University of Chicago Press, 1937. Contains an extensive bibliography.

Galizia, Mario, *Teoria della sovranità dal medioevo alla rivoluzione francese.* Milan, 1951.

McIlwain, Charles H. *The Growth of Political Thought in the West.* New York: Macmillan, 1932.

Merriam, Charles E. *History of the Theory of Sovereignty since Rousseau.* New York: Columbia University Press, 1900.

Riesenberg, P. N. *The Inalienability of Sovereignty in Medieval Political Thought.* New York: Columbia University Press, 1956.

STUDIES DISTINGUISHING TYPES OF SOVEREIGNTY

Benn, S. I. "The Uses of 'Sovereignty.'" *Political Studies* 3 (1955): 109–122.

Bryce, James. *Studies in History and Jurisprudence.* Oxford: Oxford University Press, 1901. Vol. II.

Dicey, A. V. *Law of the Constitution,* edited by E. C. S. Wade, 10th ed. London: Macmillan, 1959.

Rees, W. J. "The Theory of Sovereignty Restated." In *Philosophy, Politics and Society,* edited by P. Laslett, first series. Oxford: Blackwell, 1956.

THE CONCEPT IN IMPERATIVE THEORIES OF LAW

Austin, John. *Lectures on Jurisprudence,* edited by R. Campbell, 5th ed. London: J. Murray, 1885.

Austin, John. *The Province of Jurisprudence Determined,* edited with an introduction by H. L. A. Hart. London: Weidenfeld and Nicolson, 1954. First published in 1832.

Bodin, Jean. *Six Livres de la république.* Lyon, 1576. Translated and abridged by M. J. Tooley as *Six Books of the Commonwealth.* Oxford: Blackwell, 1955.

Burns, J. H. "Sovereignty and Constitutional Law in Bodin." *Political Studies* 7 (1959): 174–177.

Hart, H. L. A. *Concept of Law.* Oxford: Clarendon Press, 1961.

Hobbes, Thomas. *Leviathan,* edited with an introduction by M. Oakeshott. Oxford: Blackwell, 1946. First published in 1651.

Kelsen, Hans. *General Theory of Law and State.* Cambridge, MA: Harvard University Press, 1945.

Kelsen, Hans. *Das Problem der Souveränität und die Theorie des Völkerrechts.* Tübingen, 1920.

Marshall, G. *Parliamentary Sovereignty and the Commonwealth.* Oxford: Clarendon Press, 1957.

Spinoza, Benedict. "Tractatus Politicus." In his *Opera Posthuma.* Amsterdam, 1677. This work and the one above may be found in translation in *Spinoza: The Political Works,* edited by A. G. Wernham. London: Clarendon Press, 1958.

Spinoza, Benedict. *Tractatus Theologico-Politicus.* Amsterdam, 1670.

Warrender, Howard. *Political Philosophy of Hobbes.* Oxford: Clarendon Press, 1957.

SOVEREIGNTY AS MORAL SUPREMACY

Bosanquet, Bernard. *Philosophical Theory of the State.* 4th ed. London, 1923. First published in 1899.

Green, T. H. *Lectures on the Principles of Political Obligation.* London: Longmans, 1941. First published in 1882.

Hegel, G. W. F. *Grundlinien der Philosophie des Rechts.* Berlin, 1821. Translated with notes by T. M. Knox as *Hegel's Philosophy of Right.* Oxford: Clarendon Press, 1942.

Rousseau, J.-J. *Le contrat social.* Amsterdam, 1762. Available in French in *Political Writings of Jean-Jacques Rousseau,* edited by C. E. Vaughn. Oxford, 1962. Translated and edited by F. Watkins in *Rousseau: Political Writings.* Edinburgh, 1953.

PLURALIST CRITICS OF SOVEREIGN-STATE THEORY

Duguit, Léon. *Les transformations du droit public.* Paris, 1913. Translated by H. J. Laski and F. Laski as *Law in the Modern State.* New York: B.W. Huebsch, 1919.

Figgis, J. N. *Churches in the Modern State.* London: Longmans, Green, 1913.

Gierke, Otto von. *Das deutsche Genossenschaftsrecht.* 4 vols. Berlin, 1868–1913. Part of Vol. III translated with an introduction by F. W. Maitland as *Political Theories of the Middle Ages.* Cambridge, U.K., 1900. Part of Vol. IV translated with an introduction by Ernest Barker as *Natural Law and the Theory of Society, 1500 to 1800.* Cambridge, U.K.: Cambridge University Press, 1934.

Krabbe, Hugo. *Lehre des Rechtssouveränität.* Groningen, Netherlands, 1906.

Krabbe, Hugo. *Die Moderne Staats-Idee.* The Hague, 1915. Edited and translated by G. H. Sabine and W. J. Shepard as *The Modern Idea of the State.* New York, 1922.

Laski, Harold J. *The Foundations of Sovereignty and Other Essays.* New York: Harcourt Brace, 1921.

Laski, Harold J. *Grammar of Politics,* 5th ed. London, 1948.

Laski, Harold J. *Studies in the Problem of Sovereignty.* New Haven, CT: Yale University Press, 1917.

Stanley I. Benn (1967)

SOVEREIGNTY [ADDENDUM]

Sovereignty is one of the central organizing concepts of modern Western political thought. To say that it is a concept central to the organization of political thought is not to say that it is one of the concepts on which political theorists have lavished the greatest amount of explicit attention. But it is to say that certain claims about sovereignty are crucial to the way philosophers in the modern period have modeled or pictured the political world about which they are theorizing. That way of picturing the political world gained currency following the Peace of Westphalia, which was brokered to end the wars of religion that wracked Europe after the Protestant Reformation. It can therefore be called the post-Westphalian model.

The Peace of Westphalia gave impetus and sanction to the emergence of national states in Europe. The post-Westphalian model is a model of the world of states as philosophers such as Thomas Hobbes and Jean Bodin thought that world should be. The most important claims made by proponents of the post-Westphalian model are that the world is (a) divided into states that (b) should be ruled by agents who exercise sovereignty within the boundaries of the states they govern, and (c) are themselves sovereign with respect to one another. Recall that according to (1) and (3) in Stanley Benn's entry above, to say that an agent exercises sovereignty is to say that that agent exercises political authority or power, and that there is no agent who is authorized to override the decisions of the agent to whom sovereignty is ascribed, or who can generally be expected to prevail against that agent. According to the post-Westphalian model, then, the political world (a) consists of states (b) each of which is ruled by an agent exercising supreme power or authority within that state's borders, and (c) those states are not themselves subject to an agent who exercises such authority or power over international relations.

The central elements of the post-Westphalian model raise a number of interesting and important philosophical questions. That the European political world seemed increasingly to conform to the post-Westphalian model in the modern period guaranteed that the questions raised by (a), (b), and (c) would have a high place on the agenda of Western political theory. It is because these questions are raised by the claims about sovereignty that lie at the heart of the post-Westphalian model that sovereignty has become a central organizing concept of political philosophy—a concept the analysis of which, as Benn said, brings one into contact with nearly all the major problems of the discipline.

While political philosophers continue to debate the details of the post-Westphalian model, it is widely agreed that the sovereignty, which the model ascribes to rulers and states, confers on the sovereign a presumption of control over a state's people, territory, and boundaries. To question the presumption of such control—by, for example, asserting that other states may interfere in a state's internal affairs at will—is to question the sovereignty of the ruler or the state in question. Contemporary developments in politics and philosophy have led to criticism of the post-Westphalian model. Critics proceed by questioning whether states are the only corporate agents of interest in the political world and whether rulers and states can or should enjoy the presumption of control—hence the sovereignty—the model is generally taken to imply.

Why question whether states can exercise the control presupposed by the post-Westphalian model? The increasing importance of non-state actors in international affairs, and the various processes that constitute what is often called globalization, make it increasingly difficult for governments to control their own affairs or their political agenda. The rise of international terrorism in the early twenty-first century clearly makes it difficult for states to pursue their security interests or to identify rival states that threaten them. The ability of individuals and private organizations to move goods, services, infor-

mation, and capital across national borders makes it increasingly difficult for contemporary nations to manage their own economies. The liability of some states to the environmental consequences of actions undertaken by other states and the corporations they house implies that there are important parts of a state—the quality of its air and water—that some governments cannot be presumed control.

Even when states are able significantly to control their economies or their environmental quality, they may think it wise to cede a certain amount of control over their economies, their environments, or the pursuit of their national security interests to multinational unions such as NATO and the European Union. Such surrender of control is a surrender of some of the powers of sovereignty. Thus are the increasing importance of non-state actors, globalization, and the emergence of economic, political, and military unions all thought to erode the sovereignty the post-Westphalian model ascribes to states.

Why question whether states should enjoy the sovereignty the post-Westphalian model ascribes to them? The sovereignty of a state is usually taken to imply that it has a very strong presumption of control over the natural resources that lie within its borders. According to this view, a state can extract, consume, or conserve those resources as it sees fit. But it is surely open to question whether states are morally entitled to deplete a resource the rest of the world needs, to control a river on which citizens of another state downstream depend, or to exacerbate global inequalities of wealth by profiting excessively from a resource it happens to possess. Furthermore, it is open to question whether states are morally entitled to control access to its resources and opportunities by forbidding or restricting the movement of people across its borders. So-called "failed states" may lack the capacity to address humanitarian crises that affect their citizens. They can also harbor terrorist and criminal organizations that threaten international order. The incapacities of failed states, and the dangers they pose, are sometimes thought to license foreign intervention even if such intervention entails a violation of state sovereignty.

Perhaps the most profound challenge to the post-Westphalian model is posed by growing international recognition of human rights. These rights are rights that people enjoy simply in virtue of their humanity. While the list and the philosophical foundations of human rights remains disputed, it is increasingly accepted that there are such rights, that they limit what governments may do to their people and that the gross and widespread violation of such rights by a government may give non-governmental organizations, other states, and international bodies the right to intervene. The easier it is to defeat the presumption of non-intervention in such cases, the greater the challenge a global regime of human rights poses to the post-Westphalian model and to the forms of sovereignty that model implies. With the rejection of the post-Westphalian model as descriptively or normatively inadequate, its displacement by another model of the political world, or the loosening of its hold on the imagination of political theorists, sovereignty would cease to be the central organizing concept it long has been.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism

Bibliography

Biersteker, Thomas, and Cynthia Weber, eds. *State Sovereignty as Social Construct*. Cambridge, MA: Cambridge University Press, 1996.

Camilleri, Joseph, and Jim Falk. *The End of Sovereignty?*. Hants, U.K.: Edward Elgar, 1992.

Goldstein, Judith, and Michael Doyle, eds. *Ideas and Foreign Policy*. Ithaca, NY: Cornell University Press, 1993.

Nickel, James. "Is Today's International Human Rights System a Global Governance Regime?" *The Journal of Ethics* 6 (2002): 353–371.

Tanguy, Joelle. "Redefining Sovereignty and Intervention." *Ethics and International Affairs* 17 (2003): 141–48.

Paul Weithman (2005)

SPACE

When men began to think about the nature of "space," they thought of it as an all-pervading ether or as some sort of container. Since a thing can move from one part of space to another, it seemed that there was something, a place or a part of space, to be distinguished from the material objects that occupy space. For this reason places might be thought of as different parts of a very subtle jellylike medium within which material bodies are located.

HISTORY OF THE CONCEPT OF SPACE

Some of the Pythagoreans seem to have identified empty space with air. For more special metaphysical reasons Parmenides and Melissus also denied that there could be truly empty space. They thought that empty space would be nothing at all, and it seemed to them a contradiction to assert that a nothing could exist. On the other hand,

there seems to be something wrong with treating space as though it were a material, which, however subtle, would still itself have to be *in* space. Democritus and the atomists clearly distinguished between the atoms and the void that separated them. However, the temptation to think of space as a material entity persisted, and Lucretius, who held that space was infinite, nevertheless wrote of space as though it were a container. Yet he seems to have been clear on the fact that space is unlike a receptacle in that it is a pure void. Since material bodies, in his view, consist of atoms, there must be chinks of empty space even between the atoms in what appear to be continuous bodies.

Plato's views on space have to be gotten mainly from the obscure metaphors of the *Timaeus*; he, too, appears to have thought of space as a receptacle and of the matter in this receptacle as itself mere empty space, limited by geometrical surfaces. If so, he anticipated the view of René Descartes, where the problem arises of how empty space can be distinguished from nonempty space. Even if, like Lucretius and other atomists, we make a distinction between the atoms and the void, what is this void or empty space? Is it a thing or not a thing?

ARISTOTLE. Aristotle tried to dodge the difficulty by treating the concept of space in terms of place, which he defined as the adjacent boundary of the containing body. For two things to interchange places *exactly,* they would have to be identical in volume and shape. Consider two exactly similar apples that are interchanged in this way. The *places* are not interchanged; rather, the first apple is now at the very same place at which the second apple was and vice versa. We seem, therefore, to be back at the notion of space as a substratum or ether, but it is probable that Aristotle was trying to avoid this and that he meant to define place by reference to the cosmos as a whole. Aristotle thought of the cosmos as a system of concentric spheres, and the outermost sphere of the cosmos would, on his view, define all other places in relation to itself. In the Aristotelian cosmology each of the various "elements" tends toward its own place. Thus, heavy bodies tend toward the center of Earth, and fire goes away from it. This is not, however, for any other reason than that the center of Earth happens to be the center of the universe; the places toward which the elements tend are independent of what particular bodies occupy what places. In more recent times we view these as two different and seemingly irreconcilable ways of thought—the notions of space as a stuff and of space as a system of relations between bodies.

DESCARTES AND LEIBNIZ. Descartes held that the essence of matter is extension, and so, on his view, space and stuff are identical, for if the essence of matter is to be extended, then any volume of space must be a portion of matter, and there can be no such thing as a vacuum. This raises the question of how we can distinguish one material object (in the ordinary sense of these words) from another. How, on Descartes's view, can we elucidate such a statement as that one bit of matter has moved relative to another one? In what sense, if matter just *is* extension, can one part of space be more densely occupied by matter than another? Descartes considered these objections but lacked the mathematical concepts necessary to answer them satisfactorily. We shall see that a reply to these objections can be made by denying that space is the same everywhere, and this can be done by introducing the Riemannian concept of a space of variable curvature.

As against Descartes, Gottfried Wilhelm Leibniz held a relational theory of space, whereby space is in no sense a stuff or substance but is merely a system of relations in which indivisible substances, or "monads," stand to one another. Few philosophers have followed Leibniz in his theory of monads, but in a slightly different form the relational theory of space has continued to rival the Cartesian, or "absolute," theory. The issue between the two theories has by no means been decisively settled, at least if we consider not space but space-time. It is still doubtful whether the general theory of relativity can be stated in such a way that it does not require absolute space-time.

KANT. In his *Prolegomena,* Immanuel Kant produced a curious argument in favor of an absolute theory of space. Suppose that the universe consisted of only one human hand. Would it be a left hand or a right hand? According to Kant it must be one or the other, yet if the relational theory is correct it cannot be either. The relations between the parts of a left hand are exactly the same as those between corresponding parts of a right hand, so if there were nothing else to introduce an asymmetry, there could be no distinction between the case of a universe consisting only of a left hand and that of a universe consisting only of a right hand. Kant, however, begged the question; in order to define "left" and "right" we need the notions of clockwise and counterclockwise rotations or of the bodily asymmetry which is expressed by saying that one's heart is on the left side of one's body. If there were only one hand in the world, there would be no way of applying such a concept as left or clockwise. The relationist could therefore quite consistently reply to Kant that if there were only one thing in the universe, a human hand,

it could not *meaningfully* be described as either a right one or a left one. (The discovery in physics that parity is not conserved suggests that the universe is not symmetrical with respect to mirror reflection, so there is probably, in tact, something significant in nature analogous to the difference between a left and a right hand.)

Later, in his *Critique of Pure Reason,* Kant argued against both a naive absolute theory of space and a relational view. He held that space is something merely subjective (or "phenomenal") wherein in thought we arrange nonspatial "things-in-themselves." He was led to this view partly by the thought that certain antinomies or contradictions are unavoidable as long as we think of space and time as objectively real. However, since the work of such mathematicians as Karl Theodor Wilhelm Weierstrass, Augustin-Louis Cauchy, Julius Wilhelm Richard Dedekind, and Georg Cantor, we possess concepts of the infinite which should enable us to deal with Kant's antinomies and, indeed, also to resolve the much earlier, yet more subtle, paradoxes of Zeno of Elea.

NEWTON'S CONCEPTION OF SPACE

Isaac Newton held absolute theories of space and time—metaphysical views that are strictly irrelevant to his dynamical theory. What is important in Newtonian dynamics is not the notion of absolute space but that of an inertial system. Consider a system of particles acting on one another with certain forces, such as those of gravitational or electrostatic attraction, together with a system of coordinate axes. This is called an inertial system if the various accelerations of the particles can be resolved in such a way that they all occur in pairs whose members are equal and lie in opposite directions in the same straight line. Finding an inertial system thus comes down to finding the right set of coordinate axes. This notion of an inertial system, not the metaphysical notion of absolute space, is what is essential in Newtonian dynamics, and as Ernst Mach and others were able to show, we can analyze the notion of an inertial system from the point of view of a relational theory of space. Psychologically, no doubt, it was convenient for Newton to think of inertial axes as though they were embedded in some sort of ethereal jelly—absolute space. Nevertheless, much of the charm of this vanishes when we reflect that, as Newton well knew, any system of axes that is moving with uniform velocity relative to some inertial system is also an inertial system. There is reason to suppose, however, that in postulating absolute space Newton may have been partly influenced by theological considerations that go back to Henry More and, through More, to cabalistic doctrines.

We can remove the metaphysical trappings with which Newton clothed his idea of an inertial system if we consider how in mechanics we determine such a system. But even before we consider how we can define an inertial system of axes, it is interesting to consider how it is possible for us to define any system of axes and spatial positions at all. As Émile Borel has remarked, how hard it would be for a fish, however intelligent, which never perceived the shore or the bottom of the sea to develop a system of geometrical concepts. The fish might perceive other fish in the shoal, for example, but the mutual spatial relations of these would be continually shifting in a haphazard manner. It is obviously of great assistance to us to live on the surface of an earth that, if not quite rigid, is rigid to a first order of approximation. Geometry arose after a system of land surveying had been developed by the Egyptians, who every year needed to survey the land boundaries obliterated by the flooding of the Nile. That such systems of surveying were possible depended on certain physical facts, such as the properties of matter (the nonextensibility of chains, for example) and the rectilinear propagation of light. They also depended on certain geodetic facts, such as that the tides, which affect even the solid crust of Earth, were negligible. The snags that arise when we go beyond a certain order of approximation were unknown to the Egyptians, who were therefore able to get started in a fairly simple way.

It might be tempting to say that it was fortunate that the Egyptians were unaware of these snags, but of course in their rudimentary state of knowledge they could not have ascertained these awkward facts anyway. When, however, we consider geodetic measurements over a wide area of the globe we need to be more sophisticated. For example, the exact shape of Earth, which is not quite spherical, needs to be taken into account. Moreover, in determining the relative positions of points that are far apart from one another it is useful to make observations of the heavenly bodies as seen simultaneously from the different points. This involves us at once in chronometry. There is thus a continual feedback from physics and astronomy. Increasingly accurate geodetic measurements result in more accurate astronomy and physics, and more accurate astronomy and physics result in a more accurate geodesy.

Such a geodetic system of references is, however, by no means an inertial one. An inertial system is one in which there are no accelerations of the heavenly bodies except those which can be accounted for by the mutual gravitational attractions of these bodies. It follows, therefore, that the directions of the fixed stars must not be

rotating with respect to these axes. In principle we should be able to determine a set of inertial axes from dynamical considerations, even if we lived in a dense cloud, as on Venus, and were unaware of the existence of the fixed stars. This may have influenced Newton to think of space as absolute. However, Newton was not on Venus, and he could see the fixed stars. It is therefore a little surprising that he did not take the less metaphysical course of supposing an inertial system to be determined by the general distribution of matter in the universe. This was the line taken in the nineteenth century by Mach and is referred to (after Albert Einstein) as Mach's principle. It is still a controversial issue in cosmology and general relativity.

Mach's principle clearly invites, though it does not compel, a relational theory of space, such as Mach held. The origin of the axes of an inertial system in Newtonian mechanics was naturally taken to be the center of gravity of the solar system, which is nearly, but not quite, at the center of the sun. In fact, it is continually changing its position with reference to the center of the sun. Now that the rotation of the galaxy has been discovered, we have to consider the sun as moving around a distant center. We shall here neglect the possibility that our galaxy is accelerating relative to other galaxies. In any case, once we pass to cosmological considerations on this scale we need to abandon Newtonian theory in favor of the general theory of relativity.

The philosophical significance of the foregoing discussion is as follows: When we look to see how inertial axes are in fact determined we find no need to suppose any absolute space. Because such a space would be unobservable, it could never be of assistance in defining a set of inertial axes. On the other hand, the complexities in the determination of inertial axes are such that it is perhaps psychologically comforting to think of inertial axes, or rather some one preferred set of such axes, as embedded in an absolute space. But Newton could equally have taken up the position, later adopted by Mach, that inertial systems are determined not by absolute space but by the large-scale distribution of matter in the universe.

SPACE AND TIME IN THE SPECIAL THEORY OF RELATIVITY

We have already noticed the dependence of space measurements on time measurements which sometimes obtains in geodesy. This situation is accentuated in astronomy because of the finite velocity of light. In order to determine the position of a heavenly body we have to make allowance for the fact that we see it in the position it was in some time ago. For example, an observation of a star that is ten light-years away is the observation of it in its position years ago. Indeed, it was the discrepancy between the predicted and observed times at which eclipses of the satellites of Jupiter should occur that led Olaus Rømer to assign a finite, and approximately correct, value to the velocity of light. The correction of position and time on account of the finite velocity of light presupposes in any particular case our knowing what this velocity is, relative to Earth. This would seem to depend not only on the velocity of light relative to absolute space (or to some preferred set of inertial axes) but also on Earth's velocity relative to absolute space (or to the preferred set of inertial axes). The experiment of Albert Abraham Michelson and Edward Williams Morley showed, however, that the velocity of light relative to an observer is independent of the velocity of the observer. This led to the special theory of relativity, which brings space and time into intimate relation with one another. For present purposes it is necessary to recall only that according to the special theory of relativity events that are simultaneous with reference to one inertial set of axes are not simultaneous with reference to another inertial frame. The total set of point-instants can be arranged in a four-dimensional space-time. Observers in different inertial frames will partition this four-dimensional space-time into a "space" and a "time," but they will do so in different ways.

Before proceeding further it is necessary to clear up a certain ambiguity in the word *space.* So far in this entry space has been thought of as a continuant. In this sense of the word *space* it is possible for things to continue to occupy space and to move from one point of space to another and for regions of space to begin or cease to be occupied or to stay occupied or unoccupied. Here space is something that endures through time. On the other hand, there is a different, timeless use of the word *space.* In solid geometry a three-dimensional space is thought of as timeless. Thus, if a geometer said that a sphere had changed into a cube, he would no longer be thinking within the conceptual scheme of solid geometry. In geometry all verbs must be tenseless. In this tenseless way let us conceive of a four-dimensional space-time, three of whose dimensions correspond roughly to the space of our ordinary thought whereas the other corresponds to what we ordinarily call time. What we commonly think of as the state of space at an instant of time is a three-dimensional cross section of this four-dimensional space-time.

Taking one second to be equivalent to 186,300 miles, which is the distance light travels in that time, any physical object, such as a man or a star, would be rather like a

four-dimensional worm—its length in a timelike direction would be very much greater than its spacelike cross section. Thinking in terms of space-time, then, two stars that are in uniform velocity with respect to each other and also with respect to our frame of reference will appear as two straight worms, each at a small angle to the other. An observer on either star will regard himself as at rest, so he will take his own world line—the line in space-time along which his star lies—as the time axis. He will take his space axes as (in a certain sense) perpendicular to the time axis. It follows that observers on stars that move relative to one another will slice space-time into spacelike cross sections at different angles. This makes the relativity of simultaneity look very plausible and no longer paradoxical. As Hermann Minkowski observed, the relativity of simultaneity could almost have been predicted from considerations of mathematical elegance even before the experimental observations that led to the special theory of relativity. Indeed, Minkowski showed that the Lorentz transformations of the theory of relativity can be understood as simply a rotation of axes in space-time. (In trying to picture such a rotation of axes it is important to remember that Minkowski space-time is not Euclidean but semi-Euclidean.) In Minkowski's words, "Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality." We must not forget that space-time is a space in the mathematical sense of the word *space,* not in the sense in which space is a continuant. Thus, certain objectionable locutions are often used in popular expositions. For example, we sometimes hear it said that a light signal is propagated from one part of space-time to another. The correct way to put the matter is to say that the light signal *lies* (tenselessly) along a line between these two parts of space-time. Space-time is not a continuant and is not susceptible of change or of staying the same.

EUCLIDEAN AND NON-EUCLIDEAN SPACE

Geometry, as we observed earlier, developed out of experiences of surveying, such as those of the ancient Egyptians. The assumptions underlying the surveying operations were codified by Greek mathematicians, whose interests were mainly theoretical. This codification was developed by Euclid in the form of an axiomatic system. Euclid's presentation of geometry shows a high degree of sophistication, though it falls considerably short of modern standards of rigor. Euclid's geometry was a metrical one. There are, of course, geometries that are more abstract than metrical geometry. The most abstract of all is topology, which deals with those properties of a space that remain unchanged when the space is distorted, as by stretching. Thus, from the point of view of topology a sphere, an ellipsoid, and a parallepiped are identical with one another and are different from a torus. Metrical geometry uses a bigger battery of concepts—not only such notions as those of betweenness and of being longer than (which itself goes beyond topology) but also those of being, say, twice or three and a half times as long as.

Euclid regarded one of his axioms as more doubtful than the others. This is the axiom that is equivalent to the so-called axiom of parallels. It will be more convenient to discuss the axiom of parallels than Euclid's own axiom. The axiom of parallels states that if *C* is a point not on an infinite straight line *AB,* then there is one and only one straight line through *C* and in the plane of *AB* that does not intersect *AB.* Geometers made many efforts to deduce the axiom of parallels from the other, more evident ones. In the seventeenth and eighteenth centuries Gerolamo Saccheri and J. H. Lambert each tried to prove the axiom by means of a reductio ad absurdum proof. By assuming the falsity of the axiom of parallels they hoped to derive a contradiction. They did not succeed; in fact, Saccheri and Lambert proved a number of perfectly valid theorems of non-Euclidean geometry, though they were not bold enough to assert that this was what they were doing.

János Bolyai and N. I. Lobachevski replaced the axiom of parallels with the postulate that *more than one* parallel can be drawn. The type of geometry that results is called hyperbolic. Another way to deny the axiom of parallels is to say that *no* parallel can be drawn. This yields elliptic geometry. (Some adjustments have to be made in the other axioms. For instance, straight lines become finite, and two points do not necessarily determine a straight line.) It is easy to prove (by giving a non-Euclidean geometry an interpretation within Euclidean geometry) that both hyperbolic and elliptic geometries are consistent if Euclidean geometry is. (And all can easily be shown to be consistent if the theory of the real-number continuum is.) A priori, therefore, there is nothing objectionable about non-Euclidean geometries. Unfortunately, many philosophers followed Kant in supposing that they had an intuition that space was Euclidean, and mathematicians had to free themselves from this conservative climate of opinion.

The question then arose whether our actual space is Euclidean or non-Euclidean. In order to give sense to this question we must give a physical interpretation to our

geometric notions, such as that of a straight line. One way of defining a straight line is as follows: Suppose that rigid bodies *A, B, C* have surfaces S_A , S_B , S_C , such that when *A* is applied to *B,* then S_A and S_B fit; when *B* is applied to *C,* then S_B and S_C fit; and when *C* is applied to *A,* then S_C and S_A fit. Suppose also that S_A , S_B , S_C can all be slid and twisted over one another—that is, that they are not like cogged gears, for example. Then S_A , S_B , S_C are all by definition plane surfaces. The intersection of two planes is a straight line. (In the above we have used the notion of a rigid body, but this can easily be defined without circularity.) With the above definition of a straight line and the like we can make measurements to tell whether the angles of a triangle add up to two right angles. If they make more than two right angles, space is elliptic; if less than two right angles, space is hyperbolic; and if exactly two right angles, space is Euclidean. However, such experiments could not determine the question to any high degree of accuracy. All that this method shows is that, as every schoolchild knows, physical space is *approximately* Euclidean.

To make measurements that could settle the question to any high degree of accuracy we should have to make them on an astronomical scale. On this scale, however, it is not physically possible to define straight lines by means of the application of rigid bodies to one another. An obvious suggestion is that we should define a straight line as the path of a light ray in empty space. One test of the geometry of space might then come from observations of stellar parallax. On the assumption that space is Euclidean, the directions of a not very distant star observed from two diametrically opposite points on Earth's journey round the sun will be at a small but observable angle. If space is hyperbolic, this angle, which is called the parallax, will be somewhat greater. If space is elliptic, the parallax will be less or even negative. If we knew the distance of the star, we could compare the observed parallax with the theoretical parallax, on various assumptions about the geometry. But we cannot know the distances of the stars except from parallax measurements. However, if space were markedly non-Euclidean, we might get some hint of this because the distribution of stars in space, calculated from parallax observations on Euclidean assumptions, would be an improbable one. Indeed, at the beginning of the twentieth century Karl Schwarzschild made a statistical analysis of parallaxes of stars and was able to assign an upper limit to the extent to which physical space deviates from the Euclidean.

A good indication that space, on the scale of the solar system at least, is very nearly Euclidean is the fact that geometrical calculations based on Euclidean assumptions are used to make those predictions of the positions of the planets that have so strongly confirmed Newtonian mechanics. This consideration points an important moral, which is that it is impossible to test geometry apart from physics; we must regard geometry as a part of physics. In 1903, Jules Henri Poincaré remarked that Euclidean geometry would never be given up no matter what the observational evidence was; he thought that the greater simplicity of Euclidean, as against non-Euclidean, geometry would ensure our always adopting some physical hypothesis, such as that light does not always travel in straight lines, to account for our observations. We shall not consider whether—and if so, in what sense—non-Euclidean geometry is necessarily less simple than Euclidean geometry. Let us concede this point to Poincaré. What he failed to notice was that the greater simplicity of the geometry might be bought at the expense of the greater complexity of the physics. The total theory, geometry plus physics, might be made more simple even though the geometrical part of it was more complicated. It is ironical that not many years after Poincaré made his remark about the relations between geometry and physics he was proved wrong by the adoption of Einstein's general theory of relativity, in which overall theoretical simplicity is achieved by means of a rather complicated space-time geometry.

In three-dimensional Euclidean space let us have three mutually perpendicular axes, Ox_1, Ox_2, Ox_3. Let *P* be the point with coordinates (x_1 , x_2 , x_3), and let *Q* be a nearby point with coordinates ($x_1 + dx_1$, $x_2 + dx_2$, $x_3 + dx_3$). Then if *ds* is the distance *PQ,* the Pythagorean theorem

$$ds^2 = dx_1^{\,2} + dx_2^{\,2} + dx_3^{\,2}$$

holds. In a "curved," or non-Euclidean, region of space this Pythagorean equation has to be replaced by a more general one. But before considering this let us move to four dimensions, so that we have an additional axis, Ox_4. This four-dimensional space would be Euclidean if

$$ds^2 = dx_1^{\,2} + dx_2^{\,2} + dx_3^{\,2} + dx_4^{\,2}.$$

In the general case

$$\begin{aligned} ds^2 = {} & g_{11}dx_1^{\,2} + g_{22}dx_2^{\,2} + g_{33}dx_3^{\,2} + g_{44}dx_4^{\,2} \\ & + 2g_{12}dx_1dx_2 + 2g_{13}dx_1dx_3 + 2g_{14}dx_1dx_4 \\ & + 2g_{23}dx_2dx_3 + 2g_{24}dx_2dx_4 + 2g_{34}dx_3dx_4. \end{aligned}$$

The *g*'s are not necessarily constants but may be functions of x_1, x_2, x_3, x_4. That it is impossible to choose a coordinate system such that for a certain region g_{12}, g_{13}, g_{14}, g_{23},

g_{24}, g_{34} are all zero is what is meant by saying that the region of space in question is curved. That a region of space is curved can therefore in principle always be ascertained by making physical measurements in that region—for instance, by testing whether the Pythagorean theorem holds. There is, therefore, nothing obscure or metaphysical about the concept of curvature of space. The space-time of special relativity, it is worth mentioning, is semi-Euclidean and of zero curvature. In it we have

$$g_{11} = g_{22} = g_{33} = -1, \; g_{44} = +1,$$

and g_{12}, g_{13}, g_{14}, g_{23}, g_{24}, g_{34} are all zero.

According to the general theory of relativity, space-time is curved in the neighborhood of matter. (More precisely, it has a curvature over and above the very small curvature that, for cosmological reasons, is postulated for empty space.) A light wave or any free body, such as a space satellite, is assumed in the general theory to lie along a geodesic in space-time. A geodesic is either the longest or the shortest distance between two points. In Euclidean plane geometry it is the shortest, whereas in the geometry of space-time it happens to be the longest. Owing to the appreciable curvature of space-time near any heavy body, a light ray that passes near the sun should appear to us to be slightly bent—that is, there should be an apparent displacement of the direction of a star whose light passes very near the sun. During an eclipse of the sun it is possible to observe stars very near to the sun's disk, since the glare of the sun is blacked out by the moon. In the solar eclipse of 1919, Sir Arthur Stanley Eddington and his colleagues carried out such an observation that gave results in good quantitative accord with the predictions of relativity. In a similar way, also, the general theory of relativity accounted for the anomalous motion of the perihelion of Mercury, the one planetary phenomenon that had defied Newtonian dynamics. In other cases the predictions of Newtonian theory and of general relativity are identical, and general relativity is, on the whole, important only in cosmology (unlike the special theory, which has countless verifications and is an indispensable tool of theoretical physics).

IS SPACE ABSOLUTE OR RELATIVE?

The theory of relativity certainly forces us to reject an absolute theory of space, if by this is meant one in which space is taken as quite separate from time. Observers in relative motion to one another will take their space and time axes at different angles to one another; they will, so to speak, slice space-time at different angles. The *special* theory of relativity, at least, is quite consistent with either an absolute or a relational philosophical account of space-time, for the fact that space-time can be sliced at different angles does not imply that it is not something on its own account.

It might be thought that the *general* theory of relativity forces us to a relational theory of space-time, on the grounds that according to it the curvature of any portion of space-time is produced by the matter in it. But if anything the reverse would seem to be the case. If we accept a relational theory of space-time, we have to suppose that the inertia of any given portion of matter is determined wholly by the total matter in the universe. Consider a rotating body. If we suppose it to be fixed and everything else rotating, then we must say that some distant bodies are moving with transitional velocities greater than that of light, contrary to the assumptions of relativity. Hence, it is hard to avoid the conclusion that the inertia of a body is partly determined by the local metrical field, not by the total mass in the universe. But if we think of the local metrical field as efficacious in this way, we are back to an absolute theory of space-time. Furthermore, most forms of general relativity predict that there would be a curvature (and hence a structure) of space-time even if there were a total absence of matter. Indeed, relativistic cosmology often gives a picture of matter as consisting simply of regions of special curvature of space-time. (Whether this curvature is the cause of the existence of matter or whether the occurrence of matter produces the curvature of space-time is unclear in the general theory itself.) The variations of curvature of space-time enable us to rebut the objection to Descartes's theory that it cannot differentiate between more and less densely occupied regions of space.

Nevertheless, there are difficulties about accepting such a neo-Cartesianism. We must remember that quantum mechanics is essentially a particle physics, and it is not easy to see how to harmonize it with the field theory of general relativity. One day we may know whether a particle theory will have absorbed a geometrical field theory or vice versa. Until this issue is decided we cannot decide the question whether space (or space-time) is absolute or relational—in other words, whether particles are to be thought of as singularities (perhaps like the ends of J. A. Wheeler's "wormholes" in a multiply connected space) or whether space-time is to be understood as a system of relations between particles. This issue can be put neatly if we accept W. V. Quine's criterion of ontological commitment. Should our scientific theory quantify over point-instants of space-time, or should we, on the other hand, quantify over material particles, classes of them,

classes of classes of them, and so on? The latter involves a commitment to particle physics, but if a unified field theory is successful, our ontology may consist simply of point-instants, classes of them, classes of classes of them, and so on, and physical objects will be definable in terms of all of these. So far neither Descartes nor Leibniz has won an enduring victory.

See also Aristotle; Atomism; Cantor, Georg; Cartesianism; Descartes, René; Eddington, Arthur Stanley; Einstein, Albert; Geometry; Kant, Immanuel; Lambert, Johann Heinrich; Leibniz, Gottfried Wilhelm; Leucippus and Democritus; Logical Paradoxes; Lucretius; Mach, Ernst; Melissus of Samos; More, Henry; Newton, Isaac; Parmenides of Elea; Philosophy of Physics; Plato; Poincaré, Jules Henri; Pythagoras and Pythagoreanism; Quantum Mechanics; Quine, Willard Van Orman; Relativity Theory; Time; Zeno of Elea.

Bibliography

An excellent, mainly historical, account of the philosophy of space is given in Max Jammer, *Concepts of Space* (Cambridge, MA: Harvard University Press, 1954; rev. ed., New York, 1960). Ch. 1, "The Concept of Space in Antiquity," is particularly valuable as a guide to the very scattered and obscure references to space in Greek philosophy. Also useful is John Burnet, *Early Greek Philosophy,* 3rd ed. (London: A. and C. Black, 1920). For Descartes, see his *Principles of Philosophy,* Part II, Secs. 4–21. For Leibniz, see especially his correspondence with Clarke, third paper, Secs. 3–6, and fifth paper, Secs. 32–124. The Leibniz-Clarke correspondence has been edited, with introduction and notes, by H. G. Alexander (Manchester, U.K., 1956). See also Bertrand Russell's *The Philosophy of Leibniz* (London, 1900). Newton's metaphysical views on space are to be found in the Scholium to the definitions of the *Principia* (*Mathematical Principles of Natural Philosophy,* translated by Florian Cajori, Berkeley: University of California Press, 1934). For Kant's example of the left hand and the right hand, see his *Prolegomena to Any Future Metaphysics,* translated by P. G. Lucas (Manchester, U.K., 1953), Sec. 13. Kant's most characteristic doctrines about space are to be found in his *Critique of Pure Reason,* in Secs. 2–7 of the "Transcendental Aesthetic" and in the "First Antinomy." In N. Kemp Smith's translation (London, 1929) these passages will be found on pp. 67–82 and 396–402. A criticism of Kant, Zeno, and other philosophers is to be found in Bertrand Russell's *Our Knowledge of the External World,* Lecture VI, "The Problem of Infinity Considered Historically" (London: Allen and Unwin, 1914). See also Adolf Grünbaum, "A Consistent Conception of the Extended Linear Continuum as an Aggregate of Unextended Elements," in *Philosophy of Science* 19 (1952): 288–306. For Mach's criticism of Newton, see especially Secs. 2–6 of his *Science of Mechanics,* translated by J. T. McCormack, 6th ed. (La Salle, IL: Open Court, 1960).

On a fairly elementary level, and although somewhat out of date in places, Émile Borel, *Space and Time* (New York: Dover, 1960)—a translation, by Angelo S. Rappoport and John Dougall, of the French edition published in 1922—can be recommended. So can the more difficult *Philosophy of Space and Time,* by Hans Reichenbach (New York, 1958), and *Philosophical Problems of Space and Time,* by Adolf Grünbaum (New York: Knopf, 1963). See also Grünbaum's paper "Geometry, Chronometry and Empiricism," in *Minnesota Studies in the Philosophy of Science,* edited by Herbert Feigl and Grover Maxwell, Vol. III (Minneapolis: University of Minnesota Press, 1962). A criticism of Grünbaum's views is given by Hilary Putnam in his paper "An Examination of Grünbaum's Philosophy of Geometry," in *Philosophy of Science, the Delaware Seminar,* edited by Bernard Baumrin, Vol. II (1962–1963; published New York: Interscience, 1963). Chs. 8 and 9 of Ernest Nagel's *Structure of Science* (New York: Harcourt Brace, 1961) are very useful. An interesting dialogue by A. S. Eddington, "What Is Geometry?," is a prologue to his *Space, Time and Gravitation* (Cambridge, U.K.: Cambridge University Press, 1920). On relativity, see Hermann Minkowski, "Space and Time," in *The Principle of Relativity,* by Albert Einstein, et al. (London: Methuen, 1923); Hans Reichenbach, "The Philosophical Significance of Relativity," and H. P. Robertson, "Geometry as a Branch of Physics," both in *Albert Einstein: Philosopher-Scientist,* edited by P. A. Schilpp, 2nd ed. (New York: Tudor, 1951); and Adolf Grünbaum, "The Philosophical Retention of Absolute Space in Einstein's General Theory of Relativity," in *Philosophical Review* 66 (1957): 525–534 (a revised version appears in *Problems of Space and Time,* edited by J. J. C. Smart, New York: Macmillan, 1963). See also J. A. Wheeler, "Curved Empty Space-Time as the Building Material of the Physical World: An Assessment," in *Logic, Methodology, and Philosophy of Science* edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski (Stanford, CA: Stanford University Press, 1962), pp. 361–374.

For a discussion of the asymmetry between clockwise and counterclockwise rotations in relation to the nonconservation of parity, which has some relevance to Kant's problem of the left and right hands, see the brilliant popular exposition by O. R. Frisch, "Parity Is Not Conserved, a New Twist to Physics?," in *Universities Quarterly* 11 (1957): 235–244, and the article by Philip Morrison, "Overthrow of Parity," in *Scientific American* 196 (April 1957). For Poincaré's views, see *Science and Hypothesis* (New York: Dover, 1952), especially pp. 72–73. In connection with the sharpening of the issue between absolute and relational theories of space and time into an issue of ontology, see W. V. Quine, *Word and Object* (Cambridge, MA: Technology Press of the Massachusetts Institute of Technology, 1960), especially Sec. 52, "Geometrical Objects." A book of readings on space and time is *Problems of Space and Time,* edited by J. J. C. Smart (New York: Macmillan, 1964).

I should like to thank Professor B. C. Rennie, who read an earlier draft of this entry and made helpful comments.

J. J. C. Smart (1967)

SPACE IN PHYSICAL THEORIES

Space here means the space of the science of mechanics, which encompasses planetary and celestial (i.e., "outer") space, but is presupposed by the motion—spatial change—of any bodies whatsoever, from the tiniest particles through human-sized bodies to the whole universe. The investigation of space has been perhaps the most fruitful interaction between physics and philosophy. Physics endows space with specific properties playing a crucial role in determining the motions of bodies, but, despite being omnipresent, space (prerelativistically) is frustratingly inert—not having even the indirect causal effects of subatomic particles, say. Thus physics ascribes substantive properties to space on the basis of indirect evidence, allowing metaphysical bias to influence understanding, and calling (in part) for philosophical clarification.

One of the main strands of this clarification involves the "absolute-relative" debate. In fact a number of (interconnected) debates go under this title, of which two are focused on in the historical development of mechanics: Of all the motions a body has (relative to different frames of reference), which if any are privileged or "absolute"? Are such absolute motions determined by the motions of bodies relative to one another, or by motions with respect to space itself: is space a real, substantial entity in addition to bodies? (A third important strand: Are all spatial properties extrinsic—that is, "relative"—or intrinsic?)

SPACE IN ARISTOTELIAN PHYSICS

In the European tradition, Eudoxus's (408–355 BCE) account of the motions of the heavens—which was later significantly extended by Ptolemy (c. 85–165 CE)—is probably the first "physical theory" (in anything like a modern sense) in which space plays a significant role. According to this theory the Earth is at rest at the center of the universe, surrounded by a series of concentric spheres, interconnected along their axes. The moon, sun, planets, and totality of fixed stars are each located on their own sphere, with the stars farthest out. The daily apparent motions of the heavens are explained by a daily rotation of the stellar sphere, which carries all the other spheres with it; the "wanderings" of the other bodies through the fixed stars are explained by the additional, slower rotations of the other spheres about their axes. Aristotle (384–322 BCE) provided the philosophical interpretation of this system: a finite, spherical universe with an absolute center (which Aristotle suggests is determined by its position relative to the circumference). Thus bodies do have an absolute motion, namely relative to the center, which is essential for Aristotle's mechanics: heavy bodies move naturally toward the center, light bodies away, and the heavenly element, ether, around circularly. (Note that Aristotle denied the existence of space separate from body: no vacuum and no pure extension.)

Although astronomers often took this model instrumentally, Aristotle's account was the context of debate over the nature of space until the eighteenth century, even after Nicolaus Copernicus (1473–1543) proposed that the earth moved around the sun. Questions seen as important during this period that had bearing on later developments include the possibility of the vacuum and whether God could move the entire universe.

SPACE IN CARTESIAN PHYSICS

In the Early Modern period, René Descartes (1596–1650) is a logical place to start despite numerous important predecessors, especially Galileo Galilei (1564–1642), because of his influence on both physics and its philosophy. Notable contributions include the development of mechanical explanation, conservation laws, and, with Pierre Gassendi (1592–1655), the correct "law of inertia": Bodies experiencing no net forces move at constant speeds along straight paths. According to Descartes, because matter and space have the same essence—"extension"—they are one and the same (Plato's *Timaeus* describes a similar view). This identification poses a problem: As a body moves, so does the matter that composes it and hence the space it occupies, but if it does not change with respect to space then it does not move! In his *Principles of Philosophy* (1644), Descartes's first solution is to relativize to reference bodies (selected arbitrarily): In thought people identify a relatively moving piece of matter=space as the same body, while they identify as the same spatial region those different pieces of matter=space that bear some fixed relations to the reference bodies. However, in addition to this "ordinary" concept of motion, Descartes defines motion "properly speaking" as displacement of a body from the bodies in contact with it (in accord with Aristotle's *Physics*, Book IV Chapter 4). Why there are dual accounts is a subject of dispute.

According to one interpretation, Descartes took relative motion to be fundamental, but sought to avoid the heretical denial of the earth's rest; because he wrote only a decade after Galileo was condemned (1633), such concern was real. Descartes claimed that the universe was a plenum in constant agitation, and explained the motions of the planets (including the earth) by postulating a giant

vortex of fine matter carrying them around the sun, like leaves in a whirlpool. Hence the earth is in relative motion around the sun and roughly at rest with its surroundings, and so both Copernicus and Aristotle were correct—in the "ordinary" and "proper" senses, respectively. The second interpretation claims that Descartes took proper motion more seriously, as the correct, "true" sense of motion in physics; in particular, his laws of collision are blatantly contradictory if taken to concern relative motion, but not if they concern proper motion, because it is "absolute" in the sense of being privileged over all other relative motions. (As Christiaan Huygens [1629–1695] realized, Descartes should have changed the laws to make them consistently describe relative motions, not relied on his absolute notion.)

SPACE IN NEWTONIAN PHYSICS

Although Descartes's views were influential, Isaac Newton's (1643–1727) physics and philosophy (arguably his epistemology as well as his metaphysics) were infinitely more successful. In his *Principia* (1687) and in an unpublished essay, *De Gravitatione* (undated), he attacks Descartes's views concerning space and motion and lays out his own. Newton claims that space is three-dimensional and Euclidean, persists through time, and is neither a substance such as mind or matter (because it has no causal powers—the law of inertia holds because space does not act on bodies) nor a property of substances (because in a vacuum there is space but no substance): Space is outside of the categories of traditional metaphysics. He takes it to be a pseudosubstance, causally inert, but metaphysically necessary for the existence of anything, including God, because everything exists somewhere. Commentators often stretch metaphysical categories, and count Newton's "absolute space" as a nonmaterial, nonmental substance, regions of which may be occupied by other substances: they rather inaccurately ascribe "substantivalism" to Newton.

Newton famously argues against the Cartesian view of space using the example of a bucket of water, though it is only one of a series of arguments he gives. If bucket and water, initially at rest, are set spinning about their axis, initially the water will remain at rest, and hence be in motion relative to its contiguous surroundings (the side of the bucket); the water will be rotating properly speaking. Later, friction with the sides of the bucket will have set the water rotating at the same rate as the bucket, and so it will be at "proper" rest, according to Descartes. In the first instance, because it is not yet rotating, the surface of the water will be flat, whereas in the second it will be concave (just like tea stirred in a cup). By Descartes's and Newton's (and most of their contemporaries') explicit principles, it follows that only in the second case is the water "truly," physically rotating. And so in the experiment the water has physical motion if and only if it has no motion properly speaking. Cartesian "ordinary" motion fares no better: The water spins at a unique height in the bucket, indicating a unique rate of rotation, while it moves at different rates relative to different reference bodies. Newton concludes that because true motion is neither kind of Cartesian motion, it must be the only other option on the table: motion relative to absolute space (which he calls "absolute motion," though it was seen above that proper motion too is "absolute" in the sense of being privileged).

LEIBNIZ'S RELATIONIST RESPONSE

Gottfried Leibniz's (1646–1716) position is complex: He argued persuasively against substantivalism, but was motivated by idiosyncratic metaphysics; and he gave a sophisticated account of "relationism"—space is not a substance, and all spatial properties and motions are determined by relations—but it conflicted with his theory of collisions (the so-called "Newtonian" or "classical" theory of elastic collisions). At the end of his life Leibniz arguably held: (1) that every body possesses a unique quantity of "living force" or "*vis viva*", measured by *mass* × *speed*^{2} (basically kinetic energy), and hence a unique speed; (2) that living force and pure Cartesian extension are "form" and "substance" in an updated Aristotelian metaphysics; (3) that force entails the laws of mechanics (living force and momentum are conserved in elastic collisions); (4) to avoid being an occult power, the actual force must have no detectable effects, so the laws must satisfy the "equivalence of hypotheses" and hold in all frames (Leibniz was mistaken to think this was true of his laws); and (5) that space is not only merely relative, so bodies and their relations exhaust all spatial facts, but also ideal (not a "well-founded phenomenon" in his terms), arguing in part that because no two things can literally stand in the same relation to a third, only a mental identification allows two things to stand in the same relative place one after the other. Thus Leibniz opposes both Descartes and Newton: Against Descartes he rejects the claims that space is matter (space is ideal, whereas matter is well-founded) and that "proper" motion is privileged (Descartes also held that *vis viva* was *mass* × *speed*); against Newton he rejects the view that space is absolute.

In his *Correspondence* (1715–1716) with Samuel Clarke (1675–1729), Leibniz gives relativity arguments

against Newton. He argues that because two systems differing only in their absolute positions or velocities cannot be told apart, they must not differ at all: that absolute locations and velocities, and absolute space itself, are unreal. While this argument impressed later empiricists, Leibniz himself argued from the theological "principle of sufficient reason." Leibniz claims that his relationism avoids Newton's arguments, but this is highly doubtful: He only hints at (in his *Specimen of Dynamics*, 1695) a relational account of rotating bodies (such as the bucket), and fails to see that Newton's arguments disprove the relativity of his own mechanics.

MODERN ARGUMENTS

In his *Science of Mechanics* (1893) Ernst Mach (1836–1916) criticized Newton for making a non sequitur: Rotation relative to the bucket fails to explain the curvature of the water, but it does not follow that the water must be rotating relative to absolute space—could not the curvature show motion relative to some other body? Mach's reading fails to understand how Newton refuted Descartes, and ignores his attack on relative ("ordinary") motion, but asks a reasonable question as a non-Cartesian relationist. Mach proposed that sufficiently massive bodies act to cause distant bodies to move in constant, linear relative motion unless acted on by forces: in particular, that the fixed stars determine which motions are inertial, not absolute space. (Newton considered this idea, but dismissed it because it involved action at a distance—a questionable argument given his theory of gravity.) Mach's arguments were influential on contemporary physicists who were developing the idea of an "inertial frame": a frame in which Newton's laws hold, and in particular in which bodies experiencing no net forces move inertially. In practice, physicists have since taken inertial frames to be sufficient, and viewed absolute space (if at all) as an early formulation of that idea, though whether this approach amounts to relationism is debatable.

Mach was also a hero of empiricist philosophers, however; beginning in the 1960s, a reappraisal of Newton led to a defense of absolute space (simultaneous with a general philosophical turn from strict empiricism toward realism). In the late twentieth and early twenty-first centuries, Newton is often taken to argue abductively that his theory gives the best explanation of the bucket; better than Descartes's, and by extension, because he offered no real theory, better than Mach's. It is (arguably) no non sequitur to infer that motion is absolute because Newtonian mechanics in absolute space explains better than any relational theory. Note that although Newton might have endorsed this argument as a response to Mach, it is weaker than Newton's demonstration of the inconsistency of Descartes.

A major innovation has been to transfer the arguments into the context of (nonrelativistic) spacetime. One can then distinguish (a) "Newtonian spacetime" with a preferred standard of rest—geometrically speaking, a "rigging" that picks out stationary trajectories—from (b) "Galilean spacetime" with only a preferred standard of constant motion—no rigging but an "affine connection" that picks out nonaccelerated, inertial trajectories. In (a) both velocity and acceleration are well defined and hence "absolute," but in (b) only acceleration is, avoiding (part of) Leibniz's relativity argument. Thus, plausibly, Newtonian mechanics—which distinguishes different states of absolute acceleration but not velocity—in Galilean spacetime offers the best mechanical explanations.

Modern substantivalists infer first from the need for well-defined accelerations in Newtonian mechanics, to spacetimes with "absolute structures" such as a connection, and then further to the substantiality of those spacetimes, particularly of Galilean spacetime. Several relationists have responded by arguing that acceleration can be understood without substantial spacetime: that Newtonian mechanics has a relational interpretation. Other relationists have attempted to construct a theory that does explain this as well as Newton: Most attempts rely on the fact that if it is postulated that the total angular momentum of a system (such as the whole universe) is zero, then Newtonian mechanics determines a well-defined evolution for the relative state of the system.

RELATIVITY

How does the absolute-relative debate change in relativity theory? Consider the special theory of relativity (henceforth "STR"). First distinguish "relativity" from relationism. Broadly speaking, a theory is relativistic if it admits no unmeasureable quantities. Then Newtonian mechanics in absolute space or Newtonian spacetime is not relativistic, because it admits absolute velocity, whereas Newtonian mechanics in Galilean spacetime and electromagnetism in Minkowski spacetime are relativistic. The relativity of STR is thus of a specific kind: So that no body can be said to be at rest, all must agree on the speed of light. Thus the difference between the relativity of STR and Newtonian mechanics lies in whether one takes account of electromagnetic phenomena, a difference that has no immediate bearing on whether space is absolute or relative. Indeed, Minkowski spacetime has an affine con-

nection, so acceleration is as absolute as in Galilean spacetime, thus the same question of whether a connection provides evidence for the substantiality of spacetime arises.

In the general theory of relativity ("GTR") spacetime is not a fixed background but is acted on by matter and has more robust causal powers. For example, rapidly rotating bodies (e.g., a black hole) can produce gravitational waves with the power to stretch and squeeze bodies as they pass through them. The causal powers of spacetime are a serious problem for relational interpretations of GTR: If a gravitational wave knocks a person down it would be odd to say that person's body merely moved relative to the ground "as if" a wave were present. Thus only a strictly relational theory in agreement with the evidence for GTR will suffice for the relationist. However, the causal nature of the spacetime of GTR makes it metaphysically different from Newton's, and so hardly vindicates prerelativistic substantivalism either.

GTR is sometimes mistakenly claimed to be relational. First, the action of matter on space means that the affine connection, and hence inertial motion, is dependent on the distribution of matter in distant regions (in the causal past), as Mach claimed. However, the distribution does not determine inertial motion, because the connection also depends on the geometry of spacetime in the causal past: for instance, even if there is no matter at all, the connection in a region is not fixed. Thus Mach's relationism is not vindicated by GTR. Second, the theory is "generally covariant": Its equations take the same form in every frame. Thus, unlike Newtonian mechanics, one cannot define absolute acceleration as acceleration in some privileged class of frames, and any relative frame will do for formulating the theory. But these points do not settle the absolute-relative debate: GTR has an affine connection, and every body still has an absolute acceleration. Further, Newtonian mechanics can be formulated generally covariantly too, so arguably general covariance shows nothing.

It is true, however, that, unlike Newtonian mechanics and STR, the dynamic nature of spacetime in GTR makes general covariance necessary: Intuitively, how can spacetime have privileged frames if spacetime is not independently given? More or less equivalently, GTR is "diffeomorphism invariant": If all the dynamical quantities in a model (the distribution of matter and the geometry of spacetime) are continuously rearranged over the points of spacetime then the result is still a model. (Spacetime theories with static geometries are also "diffeomorphism invariant" in the weaker sense that the diffeomorphism of a model is also a model if the dynamical quantities and the nondynamical geometry are permuted.) Diffeomorphism invariance drives the "hole argument" against substantivalism, because it entails a kind of indeterminism if the spacetime of GTR has a substantival interpretation. That antisubstantivalists and some substantivalists avoid such indeterminism by claiming that distinct diffeomorphic models represent the same physical world, which is to say that the physical content of the models is captured by the relation between matter and the geometry of space, because this is what the models have in common. If so, GTR is a theory of the relations between dynamical quantities, which is what the prerelativistic relationists sought, though in terms of different dynamic quantities, namely relative distances. Thus it can be argued that GTR is as sympathetic to prerelativistic relationism as to prerelativistic substantivalism (note that physicists tend to emphasize the relational nature of GTR far more than philosophers).

QUANTUM GRAVITY

What of the absolute-relative debate in a sought-after quantum theory of gravity, such as string theory? First, the interpretation of diffeomorphism invariance is important for certain approaches to quantizing general relativity, which some argue gives physical import to the philosophical debate concerning the hole argument. Second, space is likely to be an "effective" notion, which does not appear as a fundamental element of the theory, but only phenomenologically in particular circumstances. If so, then neither relationism nor substantivalism will be correct interpretations of quantum gravity, and the debate may seem doomed. However, quantum gravity could shed new light on the matter in the following sense. One could ask what quantities count as observables in the effective context: If the theory can be given completely in terms of observables relating bodies then effective space could be said to be relational, whereas if the theory contains observables concerning points of space, then it seems that effective space is substantial.

See also Hole Argument, The; Philosophy of Physics; Relativity Theory; Time in Physics.

Bibliography

Barbour, Julian B. *Absolute or Relative Motion?* Vol. 1: *The Discovery of Dynamics*. Oxford: Oxford University Press, 1989. Republished as *The Discovery of Dynamics*. Oxford: Oxford University Press, 2003.

Dainton, Barry. *Time and Space*. Montreal: McGill-Queen's University Press, 2001.

Earman, John. *World Enough and Spacetime: Absolute versus Relational Theories of Space and Time*. Cambridge, MA: MIT Press, 1989.

Friedman, Michael. *Foundations of Space-Time Theories*. Princeton, NJ: Princeton University Press, 1983.

Garber, Daniel. "Leibniz: Physics and Philosophy." In *The Cambridge Companion to Leibniz*, edited by Nicholas Jolley. Cambridge, U.K.: Cambridge University Press, 1995.

Huggett, Nicholas. *Space from Zeno to Einstein: Classic Readings with a Contemporary Commentary*. Cambridge, MA: MIT Press, 1999. This book contains relevant extracts from most of the historical works mentioned here.

Rovelli, Carlo. "Quantum Spacetime: What Do We Know?" In *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, edited by Craig Callender and Nicholas Huggett. Cambridge, U.K.: Cambridge University Press, 2001.

Rynasiewicz, Robert. "By Their Properties, Causes, and Effects: Newton's Scholium on Time, Space, Place, and Motion–I. The Text" and "By Their Properties, Causes and Effects: Newton's Scholium on Time, Space, Place, and Motion–II. The Context." *Studies in History and Philosophy of Science* 26 (1995): 133–153 and 295–321.

Sklar, Lawrence. *Space, Time and Spacetime*. Berkeley: University of California Press, 1977.

Sorabji, Richard. *Matter, Space and Motion: Theories in Antiquity and Their Sequel*. Ithaca, NY: Cornell University Press, 1988.

Stein, Howard. "Newtonian Spacetime." *Texas Quarterly* 10 (1967): 174–200.

Nick Huggett (2005)

SPANN, OTHMAR

(1878–1950)

Othmar Spann, the Austrian philosopher and sociologist, was born in Vienna and educated at the universities of Vienna, Zürich, and Tübingen. He was a professor at Brünn from 1909 to 1919, when he was appointed to a chair of economics and sociology at Vienna.

Spann contrasted his "neoromantic universalism"—called neoromantic by Spann to indicate his debt to Adam Müller—with "individualism," that is, with the doctrine that society derives its character from the independently existing qualities of the individual men composing it. He classified as individualist such allegedly erroneous doctrines as the economic liberalism of Adam Smith and David Ricardo, utilitarianism, the various "social contract" theories, "natural law" theories of social life, egalitarianism, anarchism, Machiavellianism, and Marxism. As this heterogeneous grouping suggests, Spann was less interested in discussing the individual merits and faults of these doctrines than in placing them with respect to his total intellectual system. Such an aim was entirely consistent with his universalistic tenet that wholes are logically prior to and more real than their parts. Particular intellectual doctrines, on this view, can be understood only in relation to the total worldview to which they belong.

Spann's main application of universalism was in his theory of society, widely acclaimed by fascists. What is spiritual (*das Geistige*) in an individual is never due to himself alone but is always "an echo of what another spirit excites in him." The development and persistence of spirituality must be understood in the context of personal relations falling under the heading of what Spann called *Gezweiung.* Individuals so related form a genuine whole, the reality of which is presupposed by, rather than a result of, the spiritual characteristics of the related individuals. Examples of *Gezweiung* are the relations between artist and public, mother and child, teacher and pupil. Spann was not merely making the formal logical point that if, for instance, one calls a man "a teacher," one implies that he has a pupil, and vice versa. He was saying something about the quality of the teacher's and the pupil's experiences; the teacher "learns by teaching," and the pupil incorporates some of the teacher's spiritual qualities into his own soul.

Spann held that it is the prior existence of such institutions as art, the family, and education that makes possible relations of *Gezweiung.* These institutions have both a higher degree of reality and a higher value than do individuals. One does not understand what education is unless one understands that there can be more and less satisfactory instances of the teacher-pupil relationship and that there could be no actual instance beyond conceivable improvement. Therefore, a knowledge of the ideal must precede understanding of particular cases, and the study of social institutions must be normative.

An institution is itself only a partial whole (*Teilganz*) belonging to a higher reality, society. Society, too, has a normative aspect; it involves a hierarchy of values in terms of which the *Teilgänze* are mutually related. There must be a corresponding hierarchy among the social sciences; particular social institutions and aspects can be studied only in the context of a general theory of society.

Spann's emphasis on hierarchy was reinforced by his insistence that all *Gezweiung* involves a relation between a leader and one who is led. It belongs to the nature of society that there should be "obedience of those low in the spiritual scale toward those more highly developed." In Spann's theory distributive justice, based on the idea of inequality of function, replaces liberty as the fundamental social value.

Spann's stress on inequality is reflected in his political program. His doctrine of estates (*Stände*) was intended to combine decentralization with a strengthening of authority in order to check socially deleterious individualist tendencies. Each industry would be directed by the "mentally most highly developed individuals" from labor unions and employers' unions, which would send representatives to a central representative *Ständehaus.* Property would be owned communally by the various estates, and each industry's legal problems would be handled by its own special courts.

See also Equality, Moral and Social; Holism and Individualism in History and Social Science; Smith, Adam.

Bibliography

WORKS BY SPANN

Die Haupttheorien der Volkswirtschaftslehre auf lehrgeschichtlicher Grundlage. 1910. Translated from the 19th German edition by Eden Paul and Cedar Paul as *Types of Economic Theory.* London, 1930. This translation also appeared under the title *History of Economics.* New York: Norton, 1930.

Kurzgefasstes System der Gesellschaftslehre. Berlin, 1914.

Fundament der Volkswirtschaftslehre. Jena, Germany, 1918.

Der wahre Staat. Leipzig, 1921.

Gesellschaftsphilosophie. Jena, Germany, 1932.

Kämpfende Wissenschaft. Jena, Germany: Fischer, 1934. Collected papers.

Erkenne Dich selbst! Eine Geistesphilosophie als Lehre vom Menschen und seiner Weltstellung. Jena, Germany, 1935.

Naturphilosophie. Jena, Germany, 1937.

Kategorienlehre, 2nd ed. Jena, Germany, 1939.

WORKS ON SPANN

Gerber, Carl. *Der Universalismus bei Othmar Spann in Hinblick auf seine Religionsphilosophie.* Berlin, 1934.

Haag, John. "Othmar Spann and the Quest for a 'True State.'" *Austrian History Yearbook* 12–13 (2) (1976–1977): 227–250.

König, Albert. *Emil Brunners Staatsauffassung und der Universalismus Othmar Spanns.* Bleicherode, 1938.

Räber, Hans. *Othmar Spanns Philosophie des Universalismus. Darstellung und Kritik.* Jena, Germany: Fischer, 1937.

Riha, Thomas J. F. "Spann's Universalism: The Foundation of the Neoromantic Theory of Corporative State." *Australian Journal of Politics and History* 31 (2) (1985): 255–268.

Vikor, Desider. *Economic Romanticism in the Twentieth Century: Spann's Attempt to Revolutionize Economic Theory.* New Delhi: New Book Society of India, 1964.

Wagner, H. G. *Essai sur l'universalisme économique. Othmar Spann.* Paris, 1931.

Wrangel, Georg. *Das universalistische System von Othmar Spann.* Jena, Germany, 1929.

Peter Winch (1967)
Bibliography updated by Philip Reed (2005)

SPAVENTA, BERTRANDO

(1817–1883)

Bertrando Spaventa, the Italian Hegelian philosopher, was born at Bomba in Abruzzo, educated in the seminary at Chieti, and taught for a time in the seminary at Monte Cassino before moving to Naples in 1840. There he became one of a small circle of liberal students associated with Ottavio Colecchi (1773–1847), who taught privately in opposition to the "official" philosophy of Pasquale Galluppi. Colecchi was himself a devotee of Immanuel Kant, but he read all the German idealists carefully and in the original. Spaventa, like the other young men in Colecchi's circle, was convinced that the real meaning of Kant's work was to be found in the later idealists, especially in G. W. F. Hegel, and the Hegelian interpretation of the *Critique of Pure Reason* always remained the nodal point of his own speculations.

Spaventa's younger brother, Silvio, was imprisoned at Naples for his part in the revolution of 1848, and Bertrando was forced to take refuge at Turin for ten years. This was the period during which most of his ideas took shape. By 1850 he had renounced the priestly office to which he had, with great reluctance, been ordained some years earlier in the hope that by preferment he could relieve the poverty of his family. In Turin he turned his hand to political journalism, writing philosophical and historical polemics against the church and particularly against the Jesuits. He was already an enthusiastic student of Giordano Bruno and Tommaso Campanella.

THE "CIRCULATION OF ITALIAN PHILOSOPHY"

The first fruits of Spaventa's labors were his "Studi sopra la filosofia di Hegel" (in *Rivista italiana,* n.s., [November 1850]: 1–30, and [December 1850]: 31–78) and his "I principî della filosofia pratica di Giordano Bruno" (in *Saggi di filosofica civile,* Genoa, 1851). His studies of Hegel were specifically concerned with the *Phenomenology,* but they contained the germ of Spaventa's most original and fruitful conception, which he termed "circulation of Italian philosophy." This germ was the claim, first voiced by Silvio Spaventa about 1844, that the real tradition of Italian philosophy had been cut off and driven into exile by the Counter-Reformation, so that "Not our own philosophers of the last two centuries, but Spinoza, Kant, Fichte, Schelling, and Hegel, have been the real disciples of Bruno, Vanini, Campanella, Vico and other great thinkers." In this view of the history of philosophy Spaventa's patriotism was neatly reconciled with his

political and intellectual liberalism. He could use it both against the defenders of the status quo and against the patriotic chauvinism of Antonio Rosmini-Serbati and Vincenzo Gioberti, who believed that their native tradition enshrined a truth that had become corrupted in the rest of Europe. Spaventa himself held at this time that, on the contrary, nothing of value had survived in contemporary Italian philosophy.

He began to shift from this position toward his doctrine of a completed circle when he studied Rosmini's work in connection with an article on Kant that he wrote in 1855. He decided then that everything good in Rosmini's theory of knowledge had been stolen from Kant. This unjust judgment at least involved the admission that there were valuable elements in Rosmini's thought. When Spaventa began, in 1857, to work on a critical survey of Galluppi and Gioberti in connection with a projected study of Hegel's *Phenomenology,* his attitude changed dramatically, and he ended by writing in 1858 one massive volume of a planned two-volume work, *La filosofia di Gioberti* (Naples, 1863). The view that he now took was that all the fruits of European speculation from René Descartes to Kant were to be found in the work of Galluppi and Rosmini when it was rightly understood, and that Gioberti was even moving at the end of his life toward a critical reconstruction of his system that would have made it clearly the culmination of post-Kantian speculation.

Thus, in its fully developed form, the thesis that Spaventa proclaimed to the new nation when he returned as professor at Bologna in 1860, and at Naples from 1861 onward, was that the metaphysics of modern idealism was born in Bruno, that Campanella's theory of knowledge foreshadowed all the problems of rationalism and empiricism which were finally resolved by Kant, and that the achievement of the Germans had been anticipated by Giambattista Vico and had at last returned to be integrated with its sources in Galluppi, Rosmini, and Gioberti. As history, this thesis becomes more dubious with every succeeding clause. It must be taken rather as an account of the historical genesis of Spaventa's own idealism and as a model of how an idealist of the Hegelian type must strive, in studying the history of philosophy, to integrate different aspects of the truth as they appear. From this standpoint we can see how the emphasis on concrete experience that Spaventa found in Bruno and Campanella led him to feel that the rather abstract formalism of Kant's transcendental unity of apperception must be integrated with Rosmini's theory of the self as rooted in a "fundamental feeling"; once this was done, the Rosminian-Giobertian doctrine of knowledge as the intellectual intuition of Being could be jettisoned. Spaventa's most fundamental philosophical insight is to be found in his critical analysis of the difficulties that arise from an intuitive theory of knowledge.

LATER STUDIES

The "circulation of Italian philosophy" and the critical reconstruction of Gioberti is, properly speaking, a sort of Italian version of the coming to consciousness of the Absolute in Hegel's *Phenomenology*; Spaventa is remarkable among the Hegelians of his generation in that he regarded the *Phenomenology* as being of equal importance with the *Logic* in Hegel's system and as the key to a right interpretation of the system. He always rejected the religious interpretation of Hegel given by the "Right" and defended at Naples by his better-known colleague Augusto Vera. To admit that the Idea was really superior to and independent of the laborious progress of the Spirit in history would have entailed falling back into just the sort of Platonic intuitionism that Spaventa had so trenchantly criticized. The Being from which Hegel's *Logic* begins must therefore be taken as the thinking being of the Absolute Spirit itself that emerges at the end of the *Phenomenology.* Thus a completely human or immanent interpretation of the *Logic* as an actual process of thinking, rather than as an ideal pattern of thought, can be given.

Just how the *Philosophy of Nature* fits into Hegel's system thereby becomes even more obscure; Spaventa did not concern himself with this problem as such, but his ready acceptance of the Darwinian theory forced it on him in another way when the positivists began to produce evolutionary explanations of the Kantian a priori. Pointing to the vicious circle involved in a causal explanation of our belief in causes, Spaventa began in his last years to work out a phenomenalist account of experience that would do justice to the positivist claims while remaining firmly founded on Kant's first *Critique.* He died, however, before his work was finished. *Esperienza e metafisica* was published at Turin in 1888.

Spaventa was never widely understood or appreciated in his own lifetime. His most sympathetic follower was Donato Jaja (1839–1914), who inspired Giovanni Gentile to collect and republish Spaventa's scattered essays, along with some unpublished manuscripts. As a result of Gentile's work, Spaventa's true stature and importance have been recognized; and in Gentile's own "actual idealism" the three distinct strands of Spaventa's thought—the Italian tradition, the Hegelian dialectic,

and critical phenomenalism—are woven into a single synthesis.

See also Bruno, Giordano; Campanella, Tommaso; Darwinism; Descartes, René; Fichte, Johann Gottlieb; Gentile, Giovanni; Gioberti, Vincenzo; Hegel, Georg Wilhelm Friedrich; Hegelianism; Kant, Immanuel; Phenomenalism; Rosmini-Serbati, Antonio; Schelling, Friedrich Wilhelm Joseph von; Spinoza, Benedict (Baruch) de; Vanini, Giulio Cesare; Vico, Giambattista.

Bibliography

ADDITIONAL WORKS BY SPAVENTA

Saggi di critica filosofica. Naples, 1867.

Scritti filosofici. Edited by Giovanni Gentile. Naples: Morano, 1900. Contains biography and full bibliography to 1900.

Principi di etica. Naples, 1904. A study of Hegel's ethics.

La filosofia italiana nelle sue relazione colla filosofia europea. Bari: Laterza, 1908.

Logica e metafisica. Bari: Laterza, 1911.

WORKS ON SPAVENTA

Cubeddu, I. *Bertrando Spaventa.* Florence: Sansoni, 1964. The most comprehensive monograph on Spaventa.

Grilli, M. "The Nationality of Philosophy and Bertrando Spaventa." *Journal of the History of Ideas* 2 (1941): 339–371.

H. S. Harris (1967)

SPECIAL SCIENCES

The special sciences are generally taken to include all the sciences above physics, including biochemistry, genetics and the various biological sciences, the brain sciences, cognitive science, psychology, and economics, amongst many others. Because of their growing success over the last century, the special sciences, and their results, play an increasingly central role in philosophy. This is true of issues in the philosophy of mind and psychology, such as the mind-body problem or the nature of emotion, but also in central debates in ethics concerning a person's moral psychology and its implications, in metaphysics, for instance in discussions of personal identity and the possibility of freewill, and in epistemology, through the manifold issues affected by the nature of human cognitive capabilities. Consequently, debates over the nature, and status, of special sciences are understandably vigorous, though unfortunately they are also especially challenging because of the wide range of issues they incorporate, the often technical formulations of positions, and the implicit nature of many of their commitments. Given these difficulties, one must, first, illuminate the key questions about special sciences and then, second, provide a road map to the major positions and ongoing areas of dispute.

There have arguably been two primary, and hard-fought, foundational issues about the nature of special sciences, though for historical reasons only one of these questions has received widespread explicit discussion. First, there is the issue of the dispensability of the various special sciences themselves. That is, whether humans will be able to completely replace the special sciences, their theories, explanations, laws, and, ultimately, predicates (i.e., words or terms), with the predicates, laws, explanations, and theories of physics. Because there is clearly no practical opportunity of actually dispensing with the special sciences in the foreseeable future, the contested question is whether *in principle* special sciences, at some future point, could be dispensed with in favor of a more fully developed physics. Can humans, in principle, dispense with the special sciences, their predicates, such as *neuron* or *diabetic*, and the explanations couched in terms of them? (In discussing this issue, the in principle dispensability of special science predicates will be referred to for simplicity). On one side, the inter-theoretic reductionist argues that in order to fully explain and understand the natural world one ultimately only needs physics and its predicates, whereas, on the other side, the inter-theoretic anti-reductionist argues that humans cannot do without the special sciences and their proprietary vocabulary.

In contrast, rather than focusing upon words or explanations, the second foundational topic asks which entities, for example properties and individuals such as neurons and being diabetic, should be accepted as the truth makers of the best scientific explanations and theories. In this debate, in one corner is the ontological reductionist who argues that, when properly understood, the sole truth makers for scientific theories and explanations are the entities of physics—thus, really, only individuals like quarks and their properties of spin, charm, and charge should be taken to exist. In the other corner is the ontological non-reductivist who argues that, in addition to the entities of physics, it must also be accepted that the world contains the properties and individuals apparently posited by the special sciences—for example, individuals such as neurons and properties such as being diabetic.

Given the thoroughly ontological nature of this second question, it must be carefully noted that some philosophers accept the existence of only one genuine issue—the first. For example, many scientific anti-realists, such as the positivists and their intellectual descendents, take broad ontological questions to be, in some sense, ille-

gitimate and argue that one must rest simply with the first kind of question about theories, explanations, and predicates. However, since the mid-1970s, scientific realism has reemerged and argues that the best scientific theories allow humans to know about entities in the world. As a consequence, many scientific realist philosophers now accept the legitimacy of both questions about what scientific predicates are in principle indispensable and the consequent issue of which worldly entities should be taken to be the truth makers of the true, and in principle indispensable, scientific theories and explanations using such predicates.

Bearing these two issues in mind, modern discussions of special sciences can be examined and arguably start, in the 1950s, with the positivists' account of special sciences, which grows from Ernest Nagel's (1961) model of inter-theoretic reduction. In its most plausible version, Nagel provided machinery that putatively allowed the laws of special sciences to be explained by using identity statements relating special science predicates and predicates of lower level sciences, in combination with the law statements of lower level sciences, to derive the law statements of special sciences. For this entry's purposes, what is important is that it was claimed that, in principle, one could consequently derive, and explain, all the laws of the special sciences from the laws of physics. Thus it was concluded that special sciences and their predicates are in principle dispensable. As a result, the Nagelian picture of special sciences takes them to be analogous to the line chefs who are needed in restaurants to speedily prep difficult and complex subject matter, but where ultimately the master chef, in physics and its predicates, would, in principle, suffice to get the job done (i.e., to explain and understand all phenomena).

As befits positivism's suspicion of ontology, the Nagelian picture is focused upon the relations of predicates, law statements, theories, and other semantic entities. However, obvious ontological conclusions flow from the Nagelian account, though, for the ideological reasons noted earlier, these implications were rarely made explicit. When one establishes identity statements, then one shows that there is only one entity referred to by two predicates, rather than two entities as was previously supposed. Through such identity statements one thus plausibly reduces one's ontology. Furthermore, if as a result of such identity statements one only needs physics and its predicates in order to account, in principle, for everything about the natural world, then, at least intuitively, parsimony considerations suggest that the entities of physics are the only entities that actually exist. In this manner, the Nagelian view of special sciences provides the background to recent debates with a trenchant defense of inter-theoretic reductionism, and the in principle dispensability of special sciences and their predicates, implicitly combined with a thorough ontological reductionism that merely accepts the existence of the entities of physics, such as quarks and their properties.

During the 1960s, 1970s, and 1980s, philosophers of science more closely examined the actual nature of particular special sciences, primarily psychology and biology, to show that, contrary to the Nagelian claims, such disciplines and their predicates are in principle indispensable. Though a range of evidence was used to defend this conclusion, Jerry Fodor's (1974) so-called Multiple Realization Argument was the most prominent of these defenses. The latter argument's crucial premise is the observation that the predicates of the special sciences refer to properties that are composed, or multiply realized, by heterogeneous combinations of the properties studied by physics. For example, the economic predicate "has monetary value" refers to the properties composed by the physically heterogeneous combinations of properties found in paper, metal, plastic, and even shells. Such multiple realization means that there is a failure in getting the identity claims necessary to drive the Nagelian program—having monetary value, and other special science properties, simply are not identical to any particular combination of physical properties.

As well as undermining the Nagelian's key argument for the dispensability of special sciences and their predicates, multiple realization was also used to provide positive arguments for the in principle indispensability of such predicates by Fodor, William Wimsatt (1976), Philip Kitcher (1984), and others. Though differing in their details, these positive arguments putatively show that—given the physical heterogeneity of the combinations of physical properties that realize special science properties—the predicates of physics will fail to articulate the commonalities between the multiply realized properties, like having monetary value, studied by the special sciences. For the predicates of physics simply frame the physical differences amongst the heterogeneous realizers of special science properties. Consequently, it is argued in various ways that the proprietary predicates of the special sciences are also necessary, in principle, to fully account for the multiply realized properties these disciplines study.

Though many philosophers of science have worked to articulate this position, for simplicity the latter account of special sciences will be referred to as the Fodorian

view; and, as well as its claims about predicates, this position again lends itself to further ontological conclusions. In fact, the realization of special sciences properties by the properties of physics is explicitly combined by Fodor, and others, with the idea of the implementation of special science mechanisms by mechanisms of physics (and implicitly with the constitution of individuals of the special sciences by the individuals of physics).

As a result, the Fodorian view is apparently a version of ontological non-reductivism, for it assumes that the world is a compositional hierarchy containing many levels of distinct properties, individuals, and mechanisms bearing complex compositional relations to other levels of entities until one bottoms out (so far as we now know) with the entities studied by physics. The Fodorian picture thus takes a diametrically opposed view of the special sciences than the Nagelian account, arguing that, rather than leaving too many cooks preparing the broth, the special sciences and their proprietary predicates are, in principle, necessary in order to fully understand and explain the variegated levels of multiply realized properties, multiply constituted individuals, and multiply implemented mechanisms that the special sciences take as their objects of inquiry.

As the dominant position, the Fodorian view has received sustained critical attention and two tendencies are worth noting here. First, there is a significant, and continuing body of work that follows various strands of the Nagelian view, either by seeking to provide technical machinery that establishes the in principle dispensability of special science predicates, or by looking at a wider range of scientific cases to drive such machinery, or both (Hooker 1981, Bickle 1998). However, Jaegwon Kim has recently pioneered a second approach that diverges radically from the Nagelian framework's semantic focus. As a response to the Fodorian view, Kim (1998) instead champions what might be dubbed the "metaphysics of science"—the careful examination of ontological issues as they arise in sciences. The resulting strategy proceeds, first, by more carefully examining the nature of an ontological claim about the special sciences central to the Fodorian picture, and then, second, by seeking to show that when the metaphysics of this notion is properly understood it fails to support the conclusions claimed by the Fodorians.

Perhaps the most important of these critical arguments focuses on the realization relation itself. Crudely put, certain property instances, the realizers, realize another property instance only if the causal powers contributed by the realizers non-causally suffice for the powers individuative of the realized property, but not vice versa. Kim (1998) has consequently argued that, given this core feature of the realization relation, considerations of ontological parsimony make it prima facie plausible that it should only be accepted that there are realizer property instances. This grounds a new form of ontological reductionism, what Kim terms the functionalization model, which uses the Fodorian view's own commitment to realization relations in the special sciences to reduce the ontology of the sciences simply to the ultimate realizers—the properties of physics. Other important examples of such arguments driven by work in the metaphysics of science are found in Kim (1992) and Lawrence Shapiro (2000), which each use more precise metaphysical examinations of multiple realization to attack the scientific legitimacy of multiply realized properties, as well as explanations using predicates referring to them and any science that seeks to study them—again turning the Fodorian account of special sciences against itself.

Naturally, there have been responses to such critical arguments focused on the metaphysics of science (see, for instance, Fodor [1997] and Gillett [2003]) and, as yet, it is far from clear where this renewed ontological focus will finally lead. However, ongoing debates over special sciences have arguably changed in a fundamental way, not least by the range of new questions faced. Can the ontological non-reductivism, the levels position, which many assume is the backbone of the Fodorian view, be sustained or does it collapse upon itself as Kim's functionalizing reductionism seeks to show? And is there space for a third, previously unappreciated, type of view about special sciences that combines ontological reductionism, driven by the metaphysical argument underpinning Kim's functionalizing reduction, and a commitment to the in principle indispensability of special sciences, founded upon the Fodorian's reasoning that complex aggregates can only be fully understood using special science predicates? As well as these concerns about global views of special sciences, one also now confronts a prior set of more specific issues about the foundations of the special sciences. For example, what is the nature of composition generally in the special sciences, as well as particular compositional relations such as the realization relations between properties in the sciences?

The answers to the more particular questions in the metaphysics of science are important because, as has been seen, they underpin many of the ongoing disputes between proponents of competing global accounts of the special science themselves. Moreover, all of these questions about the foundations of the special sciences will

only become more pressing. For as humans increasingly look to the sciences to understand their own nature, then what is said about special sciences, like genetics, neurophysiology, and psychology, will also have more and more obvious implications for what they must consequently conclude about themselves.

See also Emergence; Philosophy of Biology; Reduction.

Bibliography

Bickle, J. *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press, 1998.

Fodor, J. "Special Sciences: Or, the Disunity of Science as a Working Hypothesis." *Synthese* 28 (1974): 97–115.

Fodor, J. "Special Sciences: Still Autonomous after all these Years." *Philosophical Perspectives* 11 (1997): 149–163.

Gillett, C. "The Metaphysics of Realization, Multiple Realizability and The Special Sciences." *Journal of Philosophy* 22 (2003): 591–603.

Hooker, C.A. "Towards A General Theory of Reduction. Part I: Historical and Scientific Settings. Part II: Identity. Part III: Cross Categorial Reduction." *Dialogue* 20 (1981): 38–59, 201–36, 496–529.

Kim, J. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

Kim, J. "Multiple Realization and the Metaphysics of Reduction." *Philosophy and Phenomenological Research* 52 (1992):1–26.

Kitcher, P. "1953 and All That: A Tale of Two Sciences." *Philosophical Review* 93 (1984): 335–73.

Nagel, E. *The Structure of Science*. New York: Harcourt Brace, 1961.

Shapiro, L. "Multiple Realizations." *Journal of Philosophy* 19 (2000): 635–654.

Wimsatt, W. "Reductionism, Levels of Organization, and the Mind-Body Problem." In *Consciousness and The Brain*, edited by G. Globus, I. Savodnik, and G. Maxwell. New York: Plenum, 1976.

Carl Gillett (2005)

SPECIESISM

"Speciesism" is the name of a form of bias or discrimination that is much discussed in the contemporary debates over the moral status of animals. It amounts to discriminating on the basis of species; that is, it takes the fact that, say, baboons and humans belong to different species as a reason in itself to draw moral differences between them and on several counts.

First, speciesism sometimes manifests itself in consideration of who or what may be members of the moral community, of who or what is morally considerable (see Clark, Frey, Regan, Singer). For example, it is sometimes said that creatures who have experiences or are sentient count morally; to go on to affirm that (some) animals have experiences and are sentient but to deny that they count morally solely because they are not of the right species is a form of speciesism. If it really is the fact that creatures have experiences and are sentient that matters, then animals count; what has to be shown is why the fact that it is a baboon and not a human who has these characteristics matters morally.

Second, speciesism sometimes manifests itself in claims about pain and suffering. For instance, we usually take pain and suffering to be evils, to be things that blight a life and lower its quality, and animals can feel pain and suffer. Thus, suppose one pours scalding water on a child and on a cat: It seems odd to say that it would be wrong to scald the child but not wrong to scald the cat, since both feel pain and suffer, both have the quality of their lives diminished, and both instinctively reveal pain-avoidance behavior. To claim that scalding the child is wrong, but that scalding the cat is not wrong solely on the basis of the species to which each belongs is not in itself to give a reason why or how species-membership is morally relevant, let alone morally decisive (see Rachels 1990, Sapontzis 1987).

Third, speciesism sometimes manifests itself in claims about the value of life. Most of us think human life is more valuable than animal life; yet to think this solely on the basis of species exposes one to an obvious problem. If it is true that normal adult human life is more valuable than animal life, it by no means follows that all human life is more valuable than animal life, since it is by no means the case that all human lives are even remotely approximate in their quality. Thus, some human lives have a quality so low that those who are currently living those lives seek to end them; this, of course, is what the contemporary concern with euthanasia and physician-assisted suicide is all about. Indeed, some humans live in permanently vegetative states, where, as best we can judge, all talk of the quality of life seems beside the point. Are even these human lives more valuable than the lives of perfectly healthy baboons? To say that they are solely because they are human lives, lives lived by members of the species *Homo sapiens,* even though it is true that healthy baboons can do all manner of things, can have all manner of experiences, is in effect to say that species-membership makes the crucial difference in value. It is not apparent exactly how it does this (see Frey). Of course, certain religions and cultural traditions may hold that humans have greater value than do animals, no matter what the quality or kind of lives lived: But these very

same religions have put forward moral views that many today do not endorse, and these very same cultural traditions have held that, for example, whites are superior to blacks.

See also Animal Mind; Animal Rights and Welfare; Euthanasia; Racism; Singer, Peter.

Bibliography

Clark, S. R. L. *The Moral Status of Animals.* Oxford: Clarendon Press, 1977.

Clark, S. R. L. *The Nature of the Beast.* Oxford: Oxford University Press, 1982.

Frey, R. G. *Interests and Rights: The Case against Animals.* Oxford: Clarendon Press, 1980.

Frey, R. G. "Moral Standing, the Value of Lives, and Speciesism." *Between the Species* 4 (1988): 191–201.

Frey, R. G. *Rights, Killing, and Suffering: Moral Vegetarianism and Applied Ethics.* Oxford: Blackwell, 1983.

Rachels, J. *Created from Animals: The Moral Implications of Darwinism.* New York: Oxford University Press, 1990.

Regan, T. *The Case for Animal Rights.* Berkeley: University of California Press, 1983.

Regan, T., and P. Singer, eds. *Animal Rights and Human Obligations,* 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1989.

Sapontzis, S. F. *Morals, Reason, and Animals.* Philadelphia: Temple University Press, 1987.

Singer, P. *Animal Liberation,* 2nd ed. New York: New York Review of Books, 1990.

Singer, P. *Practical Ethics,* 2nd ed. Cambridge, U.K.: Cambridge University Press, 1993.

R. G. Frey (1996)

SPECIOUS PRESENT

See *Time, Consciousness of*

SPENGLER, OSWALD

(1880–1936)

The German writer Oswald Spengler was born at Blankenburg, Germany. Spengler is known almost entirely for his contribution to philosophy of history. After studying at the universities of Munich, Berlin, and Halle—chiefly natural science and mathematics, although he also read widely in history, literature, and philosophy—Spengler obtained a doctorate in 1904, with a thesis on Heraclitus, and embarked upon a career as a high school teacher. In 1911 he abandoned teaching to take up the penurious life of a private scholar in Munich, where the first volume of his only considerable work, *Der Untergang des Abendlandes* (*The Decline of the West*), gradually took shape. This volume was published in 1918 at the moment of his country's defeat in World War I. Its pessimistic conclusions so exactly suited the prevailing mood that its author rocketed to instant but short-lived fame.

An ardent nationalist, Spengler has sometimes been accused, especially because of his reactionary and quite undistinguished political writings after 1923, of having helped to prepare the way intellectually for fascism. He actually opposed Adolf Hitler's rise, but chiefly on the ground (as he put it) that what Germany needed was a hero, not a heroic tenor. He died in Munich in 1936, bitterly resentful of the drastic decline his reputation had suffered. It is doubtful that Spengler would have been greatly mollified by the revival of interest in his work that followed World War II, for this was due as much to the general stimulus given to speculation about history by Arnold Toynbee's popular *A Study of History* as to any belated recognition of the independent merits of Spengler's views.

HISTORY AS COMPARATIVE MORPHOLOGY

The Decline of the West, although fascinating in stretches, is an unsystematic, repetitive, obscurely written book. Its style is oracular rather than analytical; it offers more "insights" than arguments. Yet its major claims are reasonably clear. From the outset it calls for a "Copernican revolution" in our way of viewing human history that will at once undermine both the traditional ancient–medieval–modern framework generally employed by empirical historians (a framework that Spengler finds provincial) and the prevailing linear interpretation of most Western philosophers of history, whether progressive or regressive (which he finds naive). According to Spengler, history, steadily and objectively regarded, will be seen to be without center or ultimate point of reference. It is the story of an indefinite number of cultural configurations, of which western Europe is only one, that "grow with the same superb aimlessness as the flowers of the field." The careers of such cultures, he contends, constitute the only meaning to be found in the course of history as a whole; they are pockets of unconnected significance in a wilderness of human life, most of which is "historyless." All that philosophical study of history can attempt is a "comparative morphology of cultures"—an inquiry into the typical form of their life, their rhythms, and possibly their laws—aimed at giving cate-

gories and an interpretative framework to empirical historiography. In outline at least, this is the aim of Spengler's two massive volumes.

But what exactly are the cultures that provide subject matter for the morphological approach to history? In view of the common complaint that Spengler "biologizes" history, it should be noted that he represents cultures as spiritual phenomena, although rooted in a definite "natural landscape." A culture is the spiritual orientation of a group of people who have achieved some unitary conception of their world that informs all their activities—their art, religion, and philosophy, their politics and economics, even their warfare—and which is expressible in a distinctive concept of the space in which they are to live and act. This concept of space functions as the culture's "prime symbol" and is the key to the understanding of its history.

Thus, classical man, to whom Spengler applies Friedrich Nietzsche's term *Apollinian,* is said to have conceived of himself as living in a local, finite space, a visible, tangible here-and-now, of which the life-sized nude statue and the small columned temple are eminent expressions. The concept shows itself equally in such things as the circumscribed political life of the city-state and the practice of burning rather than burying the dead, as if the idea of eternity could not be squarely faced. By contrast, modern Western man conceives of himself as living in a space of boundless extent, his whole culture expressing a Faustian urge to reach out and fill it with his activity. Thus, the spires of Gothic cathedrals soar skyward, Western painting develops distant perspectives, music produces the expansive form of the fugue. Also typically Faustian are long-distance sailing and long-range weapons, the conquest of space by telephone, and the insatiable ambitions of Western statesmen (for whom, like Cecil Rhodes, "expansion is everything").

Other cultures each have their characteristic space concept. The ancient Egyptians saw their world in one dimension, and their architecture, which assumed the basic form of a corridor enclosed in masonry, expressed the notion of "moving down a narrow and inexorably prescribed life-path." The Russians, whom Spengler classifies as non-Western, have a "flat plane" culture, which, when free to do so, expresses itself in low-lying buildings and an ethics of undiscriminating brotherhood. The Arabian culture of the Middle East, which Spengler calls Magian, views the world mysteriously, as a cavern in which "light … battles against the darkness." Its architecture is consequently interior-oriented; its religion, magical and dualistic. Altogether, Spengler claims to identify nine (possibly ten) such cultures, which have emerged at various times from "the proto-spirituality of ever-childish humanity." But he does not rule out the possibility of others being discovered.

Spengler's concept of human cultures has some affinity with G. W. F. Hegel's concept of the state. Both envisage an organic unity of human attitudes and activities that express a definite form of the human spirit. Spengler never wrote the promised metaphysical work that might have made clearer the general status of "spirit" in his philosophy of history. But his concept of it certainly differs from Hegel's, for he denies that the spirituality of successive historical units taken together reveals the developing nature of spirit itself. The units have no rational connection with one another, Spengler maintains, denying categorically that one culture can ever really understand, learn from, or (strictly speaking) be influenced by another. The divergence of his approach from Hegel's is even greater in his account of the typical career of a culture. Whereas Hegel attempted to represent not only the succession of historical units, but also the succession of stages within each unit, as a rationally (that is, dialectically) ordered sequence, Spengler finds, instead, a pattern analogous to the life cycle of a plant or animal. Like biological organisms, cultures grow old. The qualitative changes that accompany the "aging" will be as apparent, to a historian possessing "physiognomic tact," as is a culture's original orientation.

CULTURAL CYCLE

Spengler often speaks of the aging of cultures in terms of the succession of the four seasons. They have their spring in an early heroic period when life is rural, agricultural, and feudal. In the Apollinian culture this was the Homeric period; in the Faustian it was the high Middle Ages. This is a time of seminal myths, of inspiring epic and saga, and of powerful mystical religion. With summer comes the rise of towns not yet alienated from the countryside, an aristocracy of manners growing up beside an older, lustier leadership, and great individual artists succeeding their anonymous predecessors. In the Apollinian culture this was the period of the early city-states; in the Faustian it was the time of the Renaissance, of William Shakespeare and Michelangelo, and of the Galilean triumphs of the uncorrupted intellect.

Autumn witnesses the full ripening of the culture's spiritual resources and the first hints of possible exhaustion; it is a time of growing cities, spreading commerce, and centralizing monarchies, with religion being challenged by philosophy and tradition undermined by

"enlightenment." In the classical world this was the age of the Sophists, of Socrates and Plato; in the West it was the eighteenth century, which reached the apogee of creative maturity in the music of Mozart, the poetry of Johann Wolfgang von Goethe, and the philosophy of Immanuel Kant. Transition to winter is characterized by the appearance of the megalopolis, the world city, with its rootless proletariat, plutocracy, esoteric art, and growing skepticism and materialism. It is an age, furthermore, of imperialism, of increasing political tyranny, and of almost constant warfare, as political adventurers skirmish for world empire. In general, culture loses its soul and hardens into mere "civilization," the highest works of which are feats of administration and the application of science to industry.

Faustian culture is, according to Spengler, currently well into its autumn period, at a point roughly equivalent to 200 BCE in the Apollinian culture. An early sign of our advanced cultural age is the career of Napoleon Bonaparte, who is morphologically contemporary with Alexander the Great; our Julius Caesar is yet to come. The moral is plain.

> We are civilized, not Gothic or Rococo, people; we have to reckon with the hard cold facts of *late* life, to which the parallel is to be found not in Pericles' Athens but in Caesar's Rome. Of great painting or great music there can no longer be, for Western people, any question.… Only *extensive* possibilities are open to them.

Young Faustians who wish to play a significant role in the gathering winter should, in other words, either join the army or enroll in a technological institute. Spengler hopes that enough of his countrymen will heed his advice to ensure that the Faustian equivalent of the Roman Empire will be German.

CULTURAL CYCLE AND DETERMINISM. Clearly, Spengler regards comparative morphology as a basis for predicting the future of a culture, given the stage it has reached. Spengler, in fact, represents his study as the first serious attempt to "predetermine history," and he offers comparative charts in support of his claim that the life cycle of a culture takes about one thousand years to work itself through.

It is nevertheless misleading to call Spengler's account of history deterministic without qualification. Unlike Toynbee's, for example, it offers no explanation of the origin of cultures; the sudden rise of a new "world experience" is left a cosmic mystery. Nor do Spengler's cultures disappear on schedule after reaching the stage of civilization; civilizations may last indefinitely, as the examples of India and China show. Even while alive, the working out of a culture's "destiny" leaves open many alternative possibilities; the themes, Spengler says, are given, but not the modulations, which "depend on the character and capacities of individual players." Thus, Germany was bound to be united in the nineteenth century; how it would be united depended on what Frederick William IV would do in 1848 and Otto von Bismarck in 1870. Spengler's historical "laws" are thus not envisaged as determining, but only as limiting, the actions of individuals. This is part of the rationale of his political activism.

The notion, furthermore, of a developing culture's being a *self*-determining system is qualified by Spengler's recognition of two ways in which its normal development may be frustrated. Thus, he claims that the Mexican culture had perished through external assault, "like a sunflower whose head is struck off by one passing." Spengler also concedes that a culture can sustain spiritual damage from too close proximity to a stronger one, resulting in what he calls pseudomorphosis. What originally led him to elaborate this idea was the confused development of the Magian culture, which came to life on the ground of the Apollinian before the older culture had passed away. In such cases, Spengler observes, the younger culture "cannot get its breath, and fails not only to achieve pure and specific expression-forms, but even to develop fully its own self-consciousness." The Russian culture—which, according to Spengler, was "prematurely born"—has similarly been deformed by intrusions of the Faustian culture, first in the "reforms" of Peter the Great and again in the Bolshevik Revolution. Since weaker cultures take on only certain outer forms of dominant ones, however, Spengler would deny that the doctrine of pseudomorphosis contradicts his claim that one culture never really influences another.

DIFFICULTIES IN SPENGLER'S THEORY

Like all large synoptic systems, Spengler's theory of history has been criticized for rearing its speculative superstructure on too shaky an empirical foundation. Even Toynbee has not escaped this charge, and in breadth of historical knowledge (if not always in perceptiveness) Spengler is vastly the inferior of the two. His knowledge of his cultures is much more uneven; all he really knows well is the Apollinian and Faustian. More important, what he does say at the detailed level all too often gives the appearance of special pleading. In some cases his morphological judgments are just a bit too ingenious to

be convincing, as when he declares that Rembrandt's brown is the color of Ludwig van Beethoven's string quartets. In other cases dubious value judgments seem to be traceable chiefly to the requirements of the overarching thesis, as when the Roman Empire, being a winter phenomenon, is represented as culturally sterile, in spite of Vergil, Horace, and Ovid. In still other cases critics have suspected Spengler, if not of falsifying, then at least of suppressing, known historical facts, as when he claims that classical man, by contrast with Magian man, was polytheistic, ignoring the almost uniform monotheism of the great Greek philosophers. Highhanded treatment of the details is made easier by the fact that what passes for empirical verification in Spengler's work is really only casual exemplification of his general ideas; he makes no attempt to test systematically, and possibly to falsify, a precisely articulated hypothesis about cultural development. And when the details become intransigent, much can be explained away as pseudomorphosis. Thus, high-rise buildings in Russia are called Western-inspired, and Hadrian's Pantheon (the "first mosque") is labeled an irruption of the Magian.

Even if Spengler's actual procedure were scientifically more acceptable, there would remain the basic weakness of any attempt to generalize about the whole of history from a mere eight to ten instances of cultural development, two of which are conceded in any case to be abnormal. Spengler's defenders, of course, have often denied the relevance of this sort of criticism. What he attempted, they claim, was not social science, not even philosophy of history in the sense of arguing to general conclusions from philosophical premises in the manner of Kant and Hegel. It was, rather, a vision of events, whose truth is the truth of poetry. From this standpoint Spengler's charts and tables are an unfortunate lapse that should not be taken too seriously; part of the value of his work lies in its imaginative imprecision. Certainly, Spengler himself declared that whereas nature should be studied scientifically, history should be studied poetically. As a defense against the empirical objection, however, this will not do. For poetry is not predictive. Spengler's theory is distinctive in insisting that the significant features of history are those that are focused by the historian's aesthetic judgment. But classification and simple induction of the sort characteristic of the underdeveloped sciences is as essential to his final conclusions as is aesthetic insight.

The weakness of Spengler's inductions might not have been so serious had he not been an uncompromising holist as well. He offers no explanation of the changes his cultures undergo; he makes no attempt to isolate the factors that might throw light on their "mechanism" and that might have afforded reasons for expecting such developments to continue. In fact, part of the function of the puzzling contrast he draws between the "causality" of nature and the "destiny" of history is to persuade us not to look for this sort of thing. Spengler seems to think of causality rather narrowly as a matter of physical interaction. His own model for historical development is the biological destiny of a seed, its tendency to grow into a plant of a definite kind, barring accidents and in spite of deformations—it being assumed that this is not explicable mechanistically. It is ironical that although Spengler himself, in elaborating this concept of explanation, claimed to be resisting inappropriate scientific approaches to history, it is precisely because of this approach that some critics have charged him with scientism. Idealist philosophers of history, for example, have regarded Spengler as a cryptopositivist because, in searching out the life cycle of cultures without trying to understand in detail and from the inside why the human participants acted as they did, he treats what he originally defined spiritualistically as if it were part of nature. The causation of action by human reason, these critics would say, is central to all explanations of historical change. By ignoring this, Spengler's theory falls into incompatible parts.

Many critics have held that an even more obvious contradiction vitiates much of what Spengler had to say about specifically historical understanding. According to him, the reason cultures never really influence one another is that they are never able to grasp one another's prime symbol—a doctrine of cultural isolation that Spengler extended even to such apparently recalcitrant subjects as mathematics (to Apollinians and Faustians, he says, number means entirely different things). But the notion that we can never understand what is culturally alien to us surely raises barriers to the sort of understanding claimed by Spengler himself; comparative morphology presupposes a correct grasp of what is being compared. Spengler tries to meet this difficulty with the ad hoc claim that a few intuitive geniuses may rise above the barrier of cultural relativism. Yet the fact that he offered his book to the general public surely betrays confidence in a rather wider distribution of transcultural insight than is strictly compatible with the impossibility of cultures' learning from one another. Nor is it helpful to suggest that cultures may learn without being influenced, for the reason for denying influence was the impossibility of understanding. The difficulty is compounded by Spengler's sometimes also denying that we can understand what is culturally "out of phase" with us, even though it

belongs to the past of our own culture. Thus, we are told that although Tacitus knew of the revolution of Tiberius Gracchus two and a half centuries earlier, he no longer found it meaningful. Together, Spengler's two limitations on the understanding lead to the conclusion that we can understand only ourselves. This is scarcely a promising position from which to develop a theory of historical inquiry.

See also Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Heraclitus of Ephesus; Kant, Immanuel; Nietzsche, Friedrich; Philosophy of History; Plato; Socrates; Sophists; Toynbee, Arnold Joseph.

Bibliography

The first volume of Spengler's *Der Untergang des Abendlandes,* subtitled *Gestalt und Wirklichkeit* (Munich: Beck, 1918), was revised in the definitive Munich edition of 1923. The second volume, *Welthistorische Perspektiven,* had been published there the previous year. A good English translation by C. F. Atkinson is available under the title *The Decline of the West* (London: Allen and Unwin, 1932), and there is an abridged edition of this prepared by Helmut Werner and Arthur Helps (New York: Knopf, 1962).

For an introductory discussion of the background and influence of Spengler's work, as well as a select bibliography, see H. S. Hughes, *Oswald Spengler: A Critical Estimate* (New York: Scribners, 1952). For further references see Patrick Gardiner, ed., *Theories of History* (Glencoe, IL: Free Press, 1959), p. 524. For criticism from the standpoint of both idealist philosophy of history and empirical historiography, see R. G. Collingwood, "Oswald Spengler and the Theory of Historical Cycles," in *Antiquity* 1 (1927): 311–325, 435–446.

W. H. Dray (1967)

SPINOZA, BENEDICT (BARUCH) DE

(1632–1677)

Dutch Jewish philosopher Benedict de Spinoza was best known for his *Ethics* (1677), which laid out in geometric form arguments for the existence of an impersonal God, the identity of mind and body, determinism, and a way of overcoming the dominance of the passions and achieving freedom and blessedness. His *Theological-Political Treatise* (1670) was a landmark in the history of biblical criticism. He was also, in that work, the first major philosopher in the Western tradition to argue for democracy and for freedom of thought and expression.

IN THE PORT OF AMSTERDAM (1632–1656)

Spinoza was born into the Portuguese Jewish community in Amsterdam in the same year Galileo published his *Dialogue Concerning the Two Chief World Systems*. His father, Michael, was an immigrant who had fled Portugal, with other members of his family, to escape the persecution of the Inquisition. At that time the Dutch Republic was one of the few places in Europe where Jews could worship freely. In Amsterdam Michael became a fairly prosperous merchant in the import-export business and a prominent member of the Portuguese synagogue.

But Baruch, as Benedict was first called, encountered his own problems with religious intolerance. In 1656, when he was twenty-three, the synagogue expelled him for what the sentence of excommunication described as "abominable heresies" and "monstrous deeds." Although Spinoza had received an orthodox religious education in his congregation's school, he rebelled early on against central tenets of Judaism and began to take an interest in the new philosophy of Descartes, Hobbes, and Galileo. After his excommunication he was known by the Latin version of his name, Benedict (which means "blessed" in Latin, as Baruch does in Hebrew).

Excommunication was a common form of discipline in the Amsterdam synagogue, often imposed for minor offenses and for short periods, with a provision that the sentence could be lifted if the offender performed some penance. Spinoza's excommunication was unconditional and quite harsh. The elders cursed him with exceptional severity; no one in the Jewish community (including members of his own family) could associate with him. For a long time historians did not know exactly what heresies he was accused of. But in the mid-twentieth century, research in the archives of the Inquisition disclosed a report from a Spanish priest who had spent several months in Amsterdam. His report revealed that the main doctrinal charges against Spinoza were: (1) that he held that God exists "only philosophically"; (2) that he maintained that the soul dies with the body; and (3) that he denied that the law of Moses was a true law. The "monstrous deeds" probably included his unrepentant resistance to authority when threatened with excommunication.

BECOMING A PHILOSOPHER (1656–1661)

Michael de Spinoza died two years before the excommunication. At that time Baruch took over the family busi-

ness in partnership with his younger brother, Gabriel. But the punishment prescribed for his heresy made it impossible for Benedict to continue running his father's firm (which was, in any case, in financial trouble as a result of the first Anglo-Dutch war). There is little definite information about Spinoza's life during the years immediately after his excommunication. Probably he remained in Amsterdam for most of this period, and began working as a lens grinder, a craft in which he earned a reputation for excellence. Perhaps he lodged at first with Francis van den Enden, a former Jesuit at whose school he had been learning Latin. Van den Enden may also have helped to shape his inclinations toward the new philosophy, religious heterodoxy, and democratic politics. Perhaps Spinoza earned room and board by assisting Van den Enden in teaching Latin. Very probably he played parts in the comedies of Terence, which Van den Enden had his students perform in 1657 and 1658. Possibly he assisted the Quakers in their attempts to convert the Jews by translating some of their literature into Hebrew.

Sometime between 1656 and 1661 it appears that Spinoza did some formal study of philosophy at the University of Leiden. The Dutch Republic was the first place where Cartesianism took hold, having been introduced in 1640 by Regius, a professor of medicine at the University of Utrecht. Cartesianism was highly controversial. Voetius, a professor of theology at Utrecht, challenged Regius's doctrine that the union of soul and body is one of two separate substances, defending the scholastic-Aristotelian doctrine that the soul is the substantial form of the body. In 1642 the university forbade the teaching of Cartesianism. Later in the 1640s there were similar controversies at the University of Leiden. In 1646 Heereboord, a professor of logic at that university, defended the Cartesian method of doubt as a way of achieving certainty. Revius, a professor of theology at Leiden, replied that the method of doubt would lead to atheism and accused Descartes of Pelagianism. In 1647 their controversy led the university to ban the discussion of Descartes' philosophy, pro or con. Nevertheless, in the late 1650s Leiden was a place where one could study Cartesian philosophy.

By the end of the 1650s, Spinoza had established a circle of friends, the most notable of whom were Jan Rieuwertsz, a bookseller and publisher of Dutch translations of Descartes' works, who was later to become Spinoza's publisher; Jan Glazemaker, translator into Dutch of Descartes'works, who was later to translate most of Spinoza's works into Dutch; Peter Balling, the Amsterdam agent of various Spanish merchants, who was to translate Spinoza's first published work, an exposition of Descartes, into Dutch; the brothers Jan and Adriaan Koerbagh, the latter of whom died in prison for publishing Spinozistic views; and Lodewijk Meyer, a prominent member of Amsterdam literary circles, who wrote, in 1666, a work entitled *Philosophy, Interpreter of Holy Scripture*.

Meyer's work anticipates some of the themes of Spinoza's *Theological-Political Treatise* (TPT), though it differs from Spinoza in the solution it proposes. Meyer complains that theologians try to settle their controversies by appeals to scripture but that their interpretations of scripture are so insecurely based that the controversies never end. Meyer thinks Descartes' work holds the key to ending these debates. He proposes to doubt everything alleged to be the teaching of scripture if it is not based on a solid foundation. Accepting the Cartesian doctrine that God is not a deceiver, and assuming that the books of the Old and New Testaments are the word of God, Meyer concludes that if a proposed interpretation of scripture conflicts with what philosophy shows to be the truth, we can reject that interpretation as false. This is a modernized version of the Maimonidean approach to scripture that Spinoza rejected in the TPT.

Spinoza's friends in Amsterdam shared an interest in Cartesian philosophy and in a religion which involves minimal theological doctrine, emphasizing the love of God and neighbor. Many were affiliated with the Collegiants, a liberal protestant group which had broken away from the Reformed Church after the Synod of Dort in 1618, and which had neither a clergy nor a creed. Many of Spinoza's friends also had a connection with the University of Leiden.

Evidently Spinoza began writing his earliest philosophical works during this period: almost certainly the never-finished *Treatise on the Emendation of the Intellect*; probably his *Short Treatise on God, Man, and his Well-Being*, a systematic presentation of his philosophy, foreshadowing his *Ethics*, but never put into final form; and an early version of the *Theological-Political Treatise*, which may have developed out of a defense of his religious opinions he wrote in Spanish, addressed to the synagogue. The *Treatise on the Intellect* was first published in his *Opera posthuma*; the *Short Treatise* was not discovered until the nineteenth century, in two manuscripts which apparently stem from a Dutch translation of a lost Latin original. The defense to the synagogue has never been found, though it seems possible to infer some of its likely content from the version of the *Theological-Political Treatise* published in 1670.

THE TREATISE ON THE EMENDATION OF THE INTELLECT

The order of composition of Spinoza's earliest works has been debated, but there now seems to be a consensus that the *Treatise on the Emendation of the Intellect* (TEI) is the earliest of his surviving works. It is a good place to start the exposition of Spinoza's philosophy, since it explains his motivation for becoming a philosopher. Spinoza begins the TEI by writing that experience had taught him that all the things men commonly pursue—wealth, honor and sensual pleasure—are empty and futile. The pursuit of these supposed goods does not lead to true peace of mind. Sensual pleasure is transitory and, when past, is followed by great sadness. The desires for honor and wealth are never satisfied; when we achieve some measure of them, our success leads only to a never-ending quest for more of the same. When we are unsuccessful, we experience great sadness. The pursuit of honor has the special disadvantage that it puts us at the mercy of others' opinion.

The pursuit of wealth is subject to the uncertainties of fortune, as Spinoza might have learned from his experience as a merchant during the first Anglo-Dutch war. So Spinoza says he finally resolved to seek a good which would give him a joy unalloyed with sadness and which he thought could be found in love for something eternal and infinite. Achieving that highest good, he concluded, would involve perfecting his own nature by acquiring knowledge of the union the mind has with the whole of nature. This decision evidently came only after the excommunication, though it probably culminated a period of reflection which began several years earlier.

Spinoza's primary purpose in this work is to develop a theory of knowledge which will enable him—"with others if possible"—to attain the knowledge which is the highest good. He conceives that project as requiring a healing and purification of the intellect. To this end he offers a classification of the different ways we can 'perceive' things so that he can choose the best. He enumerates four ways by which he has been lead to affirm something without doubt: (1) because someone has told him so; (2) because he has come to believe it by random experience; (3) because he has inferred the essence of a thing from something else (but not adequately); and (4) because he has come to perceive the thing through its essence alone or through knowledge of its proximate cause.

Of the numerous examples Spinoza gives of things he has come to believe in these ways, one must suffice here. Suppose we are given three numbers, *a*, *b*, and *c*, and wish to find a fourth number, *d*, which is to *c* as *b* is to *a*. (1) Some will be able to find *d* because they have been taught a rule which tells them to multiply *b* and *c*, and divide the product by *a*. (2) Others will construct that rule for themselves by generalizing from simple cases where the answer is obvious. (3) Still others will have learned the rule by working through its demonstration in Euclid's *Elements*. And finally, (4), some will simply see, intuitively, the answer to the problem, without going through any inferential process. Surprisingly, given his fondness for demonstration in the *Ethics*, Spinoza rejects *all* of the first three paths to knowledge, and he claims that only the fourth way of affirming things will lead us to the perfection we seek. But, he says ruefully, the things he has so far been able to understand by this kind of knowledge are very few.

The middle portion of the TEI is a search for a method of acquiring knowledge in this fourth way. The reasoning here is obscure and seems to present difficulties which may explain why Spinoza never finished this work. For example, he claims that truth needs no sign and that having a true idea is sufficient to remove all doubt. But the method is supposed to teach us what a true idea is and how to distinguish it from other perceptions. That quest seems to assume that we do need a sign to recognize a true idea.

The concluding sections of the work, however, contain suggestive hints about Spinoza's metaphysical views during this period. A proper application of the method, it seems, will require us to order our ideas in a way which reflects the order of things in nature, reflects, that is, the causal structure of nature. This in turn requires that we begin by understanding what he calls "the source and origin of Nature," which he identifies with "the first elements of the whole of nature." He then makes a distinction between 'uncreated' things that "require nothing but their own being for their explanation" and 'created' things, which depend on a cause (other than their own nature) for their existence. The first elements of the whole of nature would evidently be uncreated things which exist in themselves, independently of anything else. Spinoza explains that if something exists in itself, it is its own cause. Everything else in nature presumably would depend in some way on the first elements. But how do 'created' things depend on 'uncreated' things? And how can something be its own cause?

Toward the end of the TEI Spinoza makes another distinction, which may help to answer these questions. He distinguishes between what he calls the series of fixed and eternal things and the series of singular, changeable

things. The singular changeable things are apparently the particular, finite things we encounter in our daily experience. The fixed and eternal things are said to be present everywhere, to be the causes of all things, and to have laws "inscribed in them," according to which the singular, changeable things come to be and are ordered. There are, it appears, two causal orders, one of which relates singular, changeable things to other singular, changeable things, the other of which relates them to fixed and eternal things. The true progress of the intellect requires understanding how singular, changeable things are related to the series of fixed and eternal things. To trace their connection with the series of other singular, changeable things would be impossible, because of the infinity of that series. But it would also not give us insight into the essences of the singular changeable things.

What does this mean? In particular, what are these fixed and eternal things? One plausible conjecture is this: central to Descartes' philosophy is the claim that philosophy is like a tree whose roots are metaphysics, whose trunk is physics, and whose branches are all the other sciences. What underlies this metaphor is Descartes' idea—present both in his cosmological treatise, *The World*, and in his *Principles of Philosophy*—that the fundamental laws of physics—such as the principle of inertia and the principle of conservation of motion—can be deduced from the attributes of God (in particular, from his immutability). From these fundamental laws of physics, which apply to all bodies, we can deduce other, more specific laws which apply to particular kinds of bodies (such as magnets) and which are the subjects of the special sciences (such as medicine and mechanics). In principle it should be possible to deduce all the laws governing the operations of physical objects from the fundamental laws of physics. And everything which happens in the physical world (except insofar as it involves the intervention of mental acts, which are outside the causal network) is governed by scientific laws.

Suppose Spinoza accepted the broad outlines of this Cartesian vision of a unified science. He would not have accepted the idea that minds can operate as uncaused causes, interfering with what would otherwise be the course of physical nature. And he would not have accepted the idea that the will of a personal God is the ultimate cause of the fundamental laws of physics. But he does seem to have accepted the idea that there are fundamental laws of physics, from which all the other laws of physical nature can in principle be deduced, and that all the operations of physical objects can be understood in terms of these laws. On this hypothesis, the first elements of the whole of nature, which are among the fixed and eternal things, would be those general features of extended nature which the fundamental laws of physics describe. The other fixed and eternal things, which are connected in a finite series running between the first elements and the singular changeable things, would be the general features of nature which the derivative laws of physics describe. And the singular, changeable things would be the particular physical objects whose operations are explained by these laws. The order of ideal science reflects the causal structure of nature.

This account may give the impression that Spinoza thought of science as a wholly a priori enterprise which proceeds by the intuition of first principles and deduction of theorems from those first principles. But the final sections of the TEI make it clear that Spinoza recognized that achieving knowledge of singular, changeable things would require some appeal to experience. The laws of nature describe general, unchanging facts, which hold at all times and places. They are not sufficient by themselves to explain why events in the physical world happen at the particular times and places they do. To understand that, Spinoza thinks, we must appeal to "other aids," to experiments which will enable us to determine by what laws of eternal things the particular event occurred. But before we can conduct fruitful experiments, we must first come to understand the nature of our senses so that we will know how to use them. Since that would appear to require knowledge of singular things, there seems to be a problem of circularity here, which may be one reason why Spinoza never succeeded in finishing this treatise.

One puzzle about the TEI, not resolved by the above interpretation, is what the relation is between the "first elements of the whole of nature" and Spinoza's later metaphysical categories. In the TEI Spinoza never uses the terms "substance," "attribute," and "mode," which are fundamental to the metaphysics of the *Ethics*. If the first elements are the uncreated things Spinoza mentions in the TEI's theory of definition, then we might be inclined to identify them with the one substance, God. The uncreated things exist in themselves, or are their own cause, and the concept of existing in itself is one Spinoza later used to define substance. Moreover, the first elements are supposed to be "the source and origin of Nature." Although Spinoza does not refer to them as God, it is natural to think that "the source and origin of Nature" must be God in any philosophy which acknowledges the existence of God. The problem is that there is, evidently, a plurality of first elements, and only one substance, only

one God. The next work we consider may provide a solution to this puzzle.

THE SHORT TREATISE ON GOD, MAN, AND HIS WELL-BEING

It is clear that Spinoza intended the *Treatise on the Emendation of the Intellect* as a prelude to a systematic exposition of his philosophy; from the correspondence it seems almost certain that some version of *The Short Treatise on God, Man, and His Well-being* (ST) was the systematic exposition the TEI was intended to introduce. Spinoza probably began writing it while he was still living in Amsterdam, but he must have finished it after he moved to Rijnsburg in the summer of 1661, when he apparently sent a copy of the Latin manuscript back to his friends in Amsterdam. This manuscript would then have been translated into Dutch for the members of his circle who could not read Latin. It is that Dutch manuscript, or manuscripts descended from it, which provides the basis for our knowledge of the ST.

Spinoza was still uncertain about publishing the ST as late as April 1662, when he had already made a start on expounding his philosophy in the geometric style of his *Ethics*. He had initially written the ST at the request of his friends, but only for private circulation, not publication. It appears that he sent them the manuscript some time after he moved to Rijnsburg. He hesitated to publish this work because he knew it was theologically unorthodox and he was reluctant to invite the attacks he knew would come from the conservative Calvinist clergy.

The surviving manuscripts present many textual difficulties. Frequently we do not know whether what we are reading is originally from Spinoza's hand, an addition by an early reader, a mistranslation of the Latin original, or a copyist's error. It appears that even in those portions of the manuscripts we can confidently ascribe to Spinoza, the views he holds, or the ways he expresses or argues for those views, reflect an early, formative stage of his thought. There also seem to be different strata in the manuscripts themselves, reflecting different stages in his thought. Often the argument is quite obscure.

In spite of these difficulties, the ST can be very instructive. Many of the central theses of the *Ethics* are already present in this work; it is interesting to see the form they take here. Like Descartes, Spinoza holds that God exists necessarily. He accepts versions of the ontological and causal arguments Descartes had used to prove this in the *Meditations*. The work does not yet have the distinctively Spinozistic arguments used in the *Ethics*. He defines God as a being consisting of infinite attributes, each perfect in its kind. This is not a definition Descartes had explicitly given, though it is one he might have accepted. From the correspondence we know Spinoza thought it followed from the definition Descartes did give, that God is by definition a supremely perfect being.

Unlike Descartes, and anticipating the *Ethics* (though often with different arguments), Spinoza contends that no substance can be finite; that there are no two substances of the same kind; that one substance cannot produce another; that God is an immanent cause; that both thought *and* extension are attributes of God; that man is not a substance, but a mode of substance; that the human soul (or mind) is a mode of thought, the idea of its body, which is, a mode of extension. Spinoza also argues in this work for theses which appear in the *Ethics* without argument, such as the identification of God with Nature. Early in the ST he contends that, because no attributes can exist in the divine intellect which do not exist in Nature, Nature must be a being which consists of infinite attributes, each perfect in its kind. So Nature satisfies the definition of God.

The identification of God with Nature and the claim that God is an extended substance are only two of several claims Spinoza makes in this work which he might have expected to arouse theological opposition. Also provocative are his contentions that because God is supremely perfect, he could not omit doing what he does; and that the properties of God commonly included in lists of his attributes—omnipotence, omniscience, eternity, simplicity, and so on—are not, strictly speaking, divine *attributes*, which tell us what God is in himself, but only modes, which can be attributed to him in virtue of some or all of his attributes. Omniscience, for example, presupposes thought; so it must be a mode, not an attribute; but it applies to God only in virtue of the attribute of thought, not in virtue of the attribute of extension. Eternity, on the other hand, would apply to God in virtue of all of his attributes. But it is not an attribute, because it does not tell us *what* God is. It only tells us something about the manner of God's existence, that he exists timelessly and immutably. Spinoza also argues that, because God is omnipotent, he does not give laws to men which they are capable of breaking (who could disobey the will of an omnipotent being?); that he does not love or hate his creatures; and that he does not make himself known to man through words, miracles, or any other finite things.

The God of the ST, like the God of the *Ethics*, is a philosopher's God, an eternal first cause of all things, quite remote from the God who revealed himself to the Jews through his prophets, chose them as his people, per-

formed miracles on their behalf, rewarded them when they obeyed his laws, and punished them when they disobeyed. Presumably something like this is what Spinoza meant when he said to the elders of the synagogue that God exists "only philosophically" and that the law of Moses is not a true law, that it does not, as Judaism supposes, represent a divine command which people may either obey or disobey at their peril.

If there is no divine law which is binding on us, how, then, should we conduct ourselves? Here Spinoza develops at considerable length a theme he only hinted at it in the TEI: that we must set aside worldly goods to seek a good which can give us joy unmixed with sadness, transferring our love for finite, transitory things to something eternal and infinite, perfecting our nature by acquiring knowledge of "the union the mind has with the whole of nature." Progressing towards this perfection requires us to rid ourselves of irrational passions, which depend on the lowest form of cognition, opinion.

Like the *Ethics*, the ST (normally) counts three forms of cognition, not the four counted in the TEI. The first, opinion, combines the first two forms of perception enumerated in the TEI: beliefs we form on the basis of what others have told us and beliefs based on what the TEI called "random experience." As an example of an irrational passion based on opinion, Spinoza offers the hatred which Jews, Christians, and Muslims often have for one another, based on unreliable reports about the others' religions and customs, and/or hasty generalizations from an inadequate acquaintance with members of the other religion. 'Opinion' in the ST corresponds to what Spinoza calls 'imagination' in the *Ethics*.

We can make progress towards overcoming these irrational passions if we pass from opinion to what the ST sometimes calls 'belief' and sometimes calls 'true belief.' However designated, this stage of cognition involves more than what the phrases suggest: in Spinoza's usage 'true belief' implies not only that the belief is true but that the believer has a firm rational basis for it. True belief, the second of three modes of cognition in the ST, is equivalent to the third of the four modes of cognition in the TEI (and to what Spinoza calls 'reason' in the *Ethics*). So it would involve rational demonstration from certain premises.

How does true belief enable us to overcome our irrational passions? Partly, it seems, by eliminating beliefs formed through unreliable ways of perceiving things, but partly also by enabling us to recognize that man is a part of nature (where this implies that man must follow the laws of nature, that his actions are as necessary as those of any other thing in nature) and partly by teaching us that good and evil are not something inherent in the things we judge to be good and evil, but that they are related to human nature. The good is what helps us to attain what our intellect conceives to be perfection for a human being; evil is what hinders our attaining it (or does not assist it).

But as in the TEI, Spinoza does not think this form of cognition can take us all the way to our goal. That requires the highest form, which this work usually calls 'clear knowledge,' or 'science,' which we achieve when we are not merely convinced by reasons but are aware of and enjoy the thing itself. If we achieve this kind of knowledge of God, we will come to love Him and be united with Him, as we now love and are united with the body. In our union with Him, we will be released from the body and achieve an eternal and immutable constancy.

This affirmation that we can achieve immortality looks like a startling departure from one of the views for which Spinoza was condemned by the synagogue—that the soul dies with the body. In other respects the ST seems to remain committed to the early heresies and to enable us to understand Spinoza's reasons for holding them. In this instance, it looks as though Spinoza has reverted to what his community regarded as orthodox belief. But as we will see when we come to the *Ethics*, it does not appear that the 'immortality' Spinoza allows is a personal immortality.

In the preceding section we noted a puzzle about Spinoza's early metaphysics: How are the "first elements of the whole of nature," which the TEI said were the "source and origin of nature," related to the categories of Spinoza's later metaphysics? If the first elements are "uncreated things," then Spinoza's theory of definition in the TEI implies that they exist in themselves, which would mean that they are substances. But the first elements are evidently many; and there is supposed to be only one substance.

In the ST the answer appears to be that the first elements of nature are the attributes, which Spinoza defines as existing through themselves and known through themselves, in contrast with the modes, which exist through and are understood through the attributes of which they are modes. So the attributes taken individually satisfy the definition of substance that Spinoza will give in the *Ethics*. The reason there is nevertheless only one substance is that the many attributes are attributes of one being, God or Nature.

The ST also tells us what the other "fixed and eternal" things of the TEI might be. Here for the first time Spinoza makes his distinction between *natura naturans*, defined as a being we conceive clearly and distinctly through itself (all the attributes, or God), and *natura naturata*, the modes which depend on and are understood through God. He divides *natura naturata* into universal and particular modes, identifying only one universal mode in each attribute: motion in extension and intellect in thought. These he describes as infinite, eternal, and immutable, proceeding immediately from God, and in turn the cause of the particular modes, which are 'corruptible': they are changeable, have a beginning, and will have an end. The idea underlying the identification of motion as a "universal" mode of extension is that, in accordance with the mechanistic program of the new philosophy, the particular properties of individual extended objects are a function of the different degrees of motion of their component parts.

RIJNSBURG YEARS (1661–1663)

By mid-summer of 1661 Spinoza had moved to Rijnsburg, a quiet village near Leiden, which had been the center of the Collegiant sect. The extant correspondence begins during this period, so we are much better informed about these years in Spinoza's life. Much of the correspondence is with his Amsterdam friends, but his correspondents also include Henry Oldenburg, who became the first secretary of the nascent Royal Society, and Robert Boyle, the British chemist and advocate of the mechanical philosophy. By the fall Spinoza had begun to put his philosophy into geometric form. An early experiment with a geometric presentation appears as an appendix to the ST; another version can be reconstructed from the correspondence with Oldenburg, whom Spinoza had sent a draft which improved on the draft in the appendix of the ST.

In the following year, Spinoza undertook to teach Cartesian philosophy to a student named Casearius. He prepared for Casearius a geometric presentation of Part II of Descartes' *Principles of Philosophy*, which deals with the foundations of Cartesian physics, along with some thoughts on topics in metaphysics. When his friends learned of this work, they urged him to add to it a geometric presentation of Part I of Descartes' *Principles*; Lodewijk Meyer offered to write a preface for the work and help him polish it for publication. Spinoza agreed, hoping that by establishing himself as an expert in Cartesian philosophy, he would ease the way toward the publication of his own ideas.

PARTS I AND II OF DESCARTES' *PRINCIPLES OF PHILOSOPHY* DEMONSTRATED GEOMETRICALLY (1663)

Although the preface Meyer wrote for this work proclaimed that Spinoza's work was no more than an exposition of Descartes' *Principles*—and that this was true even for the appendix, which Spinoza called *Metaphysical Thoughts*—in fact his work is more than that. For one thing, Spinoza also draws on other Cartesian works in constructing his account of Descartes' philosophy. Sometimes his reconstruction implies a criticism of the way Descartes himself argued for his positions. Sometimes he is openly critical of Descartes' assumptions. And sometimes (particularly in the appendix) he uses this venue to develop his own ideas, independently of Descartes. An interesting example involves the question of miracles. He offers a reason for doubting them along the lines he subsequently published in the TPT. But in this work he does not endorse the argument; he merely leaves it as a problem for the theologians.

Perhaps his most important differences with Descartes in this mainly expository work are those he asked Meyer to call attention to in his preface: that he does not think the will is distinct from the intellect, or endowed with the freedom Descartes attributed to it; and that he does not think the human mind is a substance, any more than the human body is a substance. Just as the human body is "extension determined in a certain way, according to the laws of extended nature, by motion and rest, so also the human mind, or soul, is … thought determined in a certain way, according to the laws of thinking nature, by ideas" (Gebhardt I, 132). He also disassociates himself from the Cartesian claim that some things—such as the nature of the infinite—surpass human understanding. He claims that these and many other things can be conceived clearly and distinctly, provided the intellect is guided in the search for truth along a different path from the one Descartes followed. He does not say precisely how that path would have to differ, but he does say that the foundations of the sciences Descartes laid are not sufficient to solve all the problems arise in metaphysics. We need to find different foundations for the sciences.

VOORBURG (1663–C.1670)

In April 1663, shortly before the publication of his exposition of Descartes, Spinoza moved from Rijnsburg to Voorburg, a village outside the Hague. During his first two years in Voorburg, Spinoza must have worked intensively on his *Ethics*, for by the summer of 1665 he had a draft far

enough advanced that he was thinking about finding someone to translate it into Dutch. Having grown up in a community whose main languages were Spanish, Portuguese, and Hebrew, Spinoza did not feel entirely comfortable writing philosophy in Dutch. In 1665 he seems to have conceived the *Ethics* as being divided into three parts, the last of which would probably have corresponded roughly to the last three parts of the final version.

During this period he also entered into a correspondence with a Dutch merchant and would-be philosopher, Willem van Blijenbergh, who had read his exposition of Descartes and had many questions for the author. Van Blijenbergh wondered about the existence of evil, and about how, if evil existed, this fact could be reconciled with the creation of the world by God—and indeed, its continuous creation, from one moment to the next. He wanted to know what it meant to say that evil is only a negation in relation to God, and how he could distinguish which portions of *Descartes' Principles* merely articulated Descartes' views and which ones expressed Spinoza's views. He wondered what Spinoza's view of the relation between mind and body implied about the immortality of the soul.

Van Blijenbergh was a committed Christian who believed that scripture was the ultimate authority on any philosophical question it addressed. His approach to scripture was the opposite of Meyer's: If his reason persuaded him of something contrary to what scripture taught, he would mistrust his reason rather than scripture. This was not a promising basis for a dialogue with Spinoza. Spinoza found the exchange of letters an unproductive use of his time and broke it off as soon as he could. But the correspondence with Van Blijenbergh seems to have persuaded him that he must diminish the authority of scripture before he could get a fair hearing for his own philosophy. By the fall of that year he had set the *Ethics* aside to return to work on his *Theological-Political Treatise*, which he intended to "expose the prejudices of the theologians," clear himself of the charge of atheism, and argue for freedom of thought and expression, which he saw as threatened by the authority of the preachers.

Another stimulus for this shift in his writing may have been an incident involving his landlord, Daniel Tydeman, a painter and member of the Reformed Church. The minister of the local church had died, and Tydeman was on the committee appointed to select his successor. Tydeman seems to have been a theological liberal, perhaps with Collegiant inclinations. The committee nominated a man they found sympathetic theologically but encountered opposition from conservatives in the congregation, who sought to discredit the committee's candidate by claiming, among other things, that Tydeman had living in his house a former Jew, now turned atheist, who "mocked all religions" and was "a disgraceful element in the republic." The committee's candidate was rejected.

These were difficult years for the Dutch Republic. The plague had returned to Europe in 1663 and had been so virulent that Spinoza felt it necessary to leave Voorburg to spend several months of the winter of 1664 at the country house of relatives of a friend. Competition between the Dutch and the English for control of maritime trade led to war between the two countries from 1664 to 1667, the second such war in a little over a decade. No sooner had that war ended than there were threats of a new war with France, whose king, Louis XIV, had expansionist ambitions. And there was tension between the leaders of the Republic and the princes of the house of Orange.

This tension went back to the early days of the Republic. In the mid-sixteenth century the area now occupied by the independent nations of the Netherlands, Belgium, and Luxemburg was a unit within the Holy Roman Empire, ruled by the King of Spain. Toward the end of the century, the seven northern provinces (the modern Netherlands) succeeded in breaking away from Spanish rule, largely under the leadership of William I, Prince of Orange and Stadholder of the provinces of Holland, Utrecht, and Zeeland, though his son, Maurice of Nassau, also played a key role. The Stadholders were originally governors of the provinces, representing the Spanish crown and charged with the administration of justice. During the revolt against Spain, the Stadholders of the house of Orange sided with the rebels and provided the military leadership the provinces needed. Sometimes they worked in collaboration with the States-General, an assembly representing all the provinces. Sometimes they competed with the leadership of the States-General for power. Later princes of Orange developed monarchic ambitions.

In the late 1640s the Prince of Orange was William II, who unsuccessfully opposed the Treaty of Westphalia (1648), which ended the eighty-year war for Dutch independence from Spain (as well as the Thirty Years War, which had embroiled most of Europe since before Spinoza was born). The States-General, dominated by the province of Holland and Dutch mercantile interests, favored the treaty. When William died unexpectedly in 1650, the position of the Orange party was weakened. His son, William III, was not born until just after his father's death. For

many years the minority of the young prince provided the leaders of the States-General with an excuse to leave the office of Stadholder vacant. The functions the Stadholder had performed fell to the States-General, under the leadership of Jan de Witt, who generally had great success in defending his country against many challenges. But as William III neared adulthood, the tensions between the De Witt party and the Orange party increased, particularly when the affairs of the Republic were not going well, as was the case at the end of the 1660s.

Spinoza was sympathetic to the De Witt regime, strongly preferring it to the Orangist alternative. But some historians have exaggerated his closeness to De Witt, trusting too much to contemporary accounts. De Witt's political enemies, bent on discrediting him, sometimes claimed a close association between him and Spinoza—suggesting, for example, that De Witt had assisted in the editing and publishing of the TPT. And Spinoza's friends sometimes told similar stories—for example, that De Witt had often visited Spinoza to discuss affairs of state—apparently with the intention of magnifying Spinoza's reputation by associating him with a political leader whom many regarded as a hero. Though De Witt and Spinoza would have agreed in opposing the monarchic ambitions of the Prince of Orange, Spinoza was a democrat, whereas De Witt favored an oligarchic republic. They would have agreed in opposing the desire of the more conservative members of the clergy, in alliance with the princes of Orange, to enforce a strict Calvinist orthodoxy. But Spinoza favored a very expansive freedom of thought, whereas De Witt recognized the necessity, if only as a matter of practical politics, of making accommodations to the Reformed Church.

In the *Theological-Political Treatise* (TPT) Spinoza speaks in glowing terms about the freedom of the Dutch Republic:

> Since we happen to have that rare good fortune, that we live in a Republic in which everyone is granted complete freedom of judgment, and is permitted to worship God according to his understanding, and in which nothing is thought to be dearer or sweeter than freedom, I believed I would be doing something neither unwelcome, nor useless, if I showed not only that this freedom can be granted without harm to piety and the peace of the Republic, but also that it cannot be abolished unless piety and the Peace of the Republic are abolished.
>
> (GEBHART III, 7)

But Spinoza knew all too well that the Republic was not as free as he claimed.

In 1668 his friend Adriaan Koerbagh had published *A Flower Garden of All Kinds of Loveliness*, ostensibly a treatise explaining the meanings of foreign words which had become part of Dutch but in fact a critique of all the organized religions known in the Dutch Republic. In this acerbically written book, Koerbagh anticipated a number of the claims Spinoza made two years later in the TPT: He denied that the books of the Bible were written by the men to whom they were traditionally ascribed; he proposed that Ezra, the postexilic priest and scribe who wrote the book of Ezra, was responsible for the existing form of the Hebrew Bible, having compiled and attempted to reconcile the inconsistent manuscripts which had come down to him; and he argued that a proper interpretation of the Bible would require a thorough knowledge of the languages it was written in and the historical contexts its authors wrote in. Like Spinoza, he did not deny that there was something solid and consistent with reason in scripture; but that solid element in scripture was not its theology.

Koerbagh was arrested—along with his brother, Jan, who was suspected of complicity in the work—and, with the encouragement of the Reformed clergy, tried for blasphemy by the civil authorities in Amsterdam. Jan was released after a few weeks, but Adriaan was found guilty after a lengthy inquest, during which he was questioned about his association with Spinoza and Van den Enden. Sentenced to a fine of 4,000 guilders and ten years in prison, to be followed by ten years' exile, he died a little more than a year after his imprisonment from the harsh conditions in the prison.

The influence of the Reformed clergy on Dutch politics perhaps explains why Spinoza and the other members of his circle showed the interest they did in the work of Hobbes. Probably Spinoza had known some of Hobbes' work for years, since Hobbes' first published work of political philosophy, *De cive* (On the Citizen), had been available in a language he could read since 1642. It is likely that this would have been one of the works Van den Enden called to his attention when he was encouraging his interest in the new philosophy. But before 1667, Spinoza's inability to read English would have prevented him from gaining first-hand knowledge of *Leviathan*, which developed Hobbes' religious views more fully than *De cive* had. Two events in the late 1660s changed that: in 1667 Abraham van Berckel, a friend of Spinoza's (and of the Koerbagh brothers), translated *Leviathan* into Dutch; and in 1668 an edition of Hobbes' complete Latin works

(including a Latin translation of *Leviathan*) was published in the Netherlands. Although it may seem paradoxical to Anglophone readers of Hobbes, who think of him primarily as a defender of absolute monarchy, Hobbes' theory was attractive to republicans in the United Provinces because of his advocacy of state control over religion. In Holland in the 1660s conservative Christianity was a problem for them, much as it had been for the royalists in England in the 1640s.

THE THEOLOGICAL-POLITICAL TREATISE (1670)

It is no accident that Spinoza treats religion and politics in one work. The preface to the TPT illustrates one way in which these subjects are linked. Spinoza begins with reflections on the psychological origin of superstition, which he attributes to the uncertainty of our lives and the role fortune plays in them. Much of what happens to us depends on circumstances over which we have no control. We do not know whether things will go well or badly for us, and we fear what may happen if they go badly. So we would like to believe in some story which offers us the hope of gaining control over our lives. In this mood we may believe that the future can be predicted from the entrails of birds or affected by prayer and the performance of rituals. That belief puts us at the mercy of unscrupulous priests and the politicians who use them. "The greatest secret of monarchic rule," Spinoza writes in the preface, "is to keep men deceived, and to cloak in the specious name of religion the fear by which they must be checked, so that they will fight for slavery as they would for salvation, and will think it not shameful, but a most honorable achievement, to give their life and blood that one man may have a ground for boasting."

If the politicians use the priests to provide divine authority for their rule, the priests also use the politicians, trading their support for the enactment of laws condemning opinions contrary to those they endorse. These condemnations enhance their authority, giving official sanction to the idea that the priests have a special expertise in matters of religion. Spinoza speaks with respect of Christianity, which he sees as a religion whose true spirit calls for love, peace, restraint, and honesty toward all. But he deplores the fact that the Christians of his day are no more prone to display these virtues than the members of any other religion, a fact he attributes to the wealth, honor, and power accorded to its clergy. These incentives attract the worst kind of men to the ministry, men who for their personal ends are willing to exploit the credulity of the people for personal gain, to teach them contempt for reason, and to stir up hatred of those who disagree with them.

Spinoza proposes to remedy this evil by challenging the assumptions with which the priests approach scripture. They assume as a principle of interpretation that scripture is, in every passage, true and divine. Since scripture often appears to be inconsistent, they invent forced, reconciling interpretations whose only value is their apparent smoothing over of contradictions. And because scripture often appears to be contrary to reason in other ways, they are prone to invent metaphorical readings of scripture to make it conform to their beliefs. This procedure reverses the proper order of things. We should seek first to determine the meaning of scripture and only after that should we make a judgment about its truth and divinity.

But how should we determine the meaning of scripture? Spinoza's fundamental rule is that we should attribute to scripture as its teaching nothing we have not clearly understood from its history. By a "history of scripture" Spinoza understands, first, an account of the vocabulary and grammar of the language in which its books were written and which its authors spoke. This will tell us what meanings its words can have in ordinary usage and what ways of combining those words are legitimate. Second, a history of scripture must organize what scripture says topically, so that we can easily find all the passages bearing on the same subject; it must also note any passages which seem ambiguous or obscure or inconsistent with one another. Next, it must describe the circumstances under which the book was written, who its author was, what his character was, when he wrote and for what reason, for what audience, and in what language. And, finally, it must tell us how the book was first received, into whose hands it fell, how many different readings there are of various passages, and how it came to be accepted as sacred. What Spinoza is proposing here is that we apply to the interpretation of scripture the scholarly criteria Renaissance humanists had applied to the classics of pagan antiquity (with the exception that for the pagan works the question of their acceptance as sacred does not arise).

The result of applying these rules does not inspire confidence in the historical accuracy of scripture: the historical books were not written by the authors to whom tradition ascribed them—Moses, Joshua, Samuel, and so on—but were compiled by a much later editor, whose knowledge of the events these books described was based on manuscripts which had come into his possession but are now lost. Spinoza conjectures that this editor was

Ezra. Moreover, not only was Ezra's knowledge of the early history of the Jews second-hand knowledge of long-ago events, but he also reworked the texts to smooth out inconsistencies and make them tell the story he wanted to tell: that when the people of Israel obeyed God's laws, they prospered, whereas evil befell them when they disobeyed.

Not all of Spinoza's conclusions about the Bible were radically new. In the twelfth century Abraham ibn Ezra had hinted in his commentary on the Torah that the first five books of the Bible, in the form in which we have them, were written much later than the events they described. In the 1650s Isaac de la Peyrère and Thomas Hobbes had drawn similar conclusions more openly. But Spinoza was more systematic, thorough, and blunt than any of these predecessors. Unlike La Peyrère and Hobbes, he had the advantage of knowing the texts well in the original Hebrew, of knowing the medieval Jewish interpretive tradition, and of having a well-developed theory of interpretation, a theory which set a new standard for Biblical scholarship. Unlike Ibn Ezra, he did not pull his punches:

> Those who consider the Bible, as it is, as a letter God has sent men from heaven, will doubtless cry out that I have committed a sin against the Holy Ghost, because I have maintained that the word of God is faulty, mutilated, corrupted and inconsistent, which we have only fragments of it, and finally, which the original text of the covenant God made with the Jews has been lost.
>
> (GEBHARDT III, 138)

It's hardly surprising that when Hobbes read the TPT, he commented that he had not dared to write so boldly.

Spinoza did not object only that our knowledge of biblical history was based on unreliable texts, he also criticized biblical theology as embodying the opinions of men whose conception of God was based on the imagination rather than the intellect. The prophets, he argued, were outstanding for their personalities, their moral qualities, and their knack for expressing themselves in powerful language. But they were not philosophers. They thought of God as the maker of all things, existing at all times, who surpassed all other beings in power; but they did not understand that God was omniscient and omnipresent, or that He directed all human actions by his decree. They imagined that He had a body, which was visible (though you would die if you looked upon it), and that He had emotions, like compassion, kindness, and jealousy. Moreover, they were not strict monotheists. They believed that there were other Gods who were subordinate to the God of Israel and that He had entrusted the care of other nations to these lesser Gods. So their conceptions of God were very inadequate. And they often accommodated their theology to the even more primitive capacities of their audience.

In his rejection of Biblical theology, Spinoza even goes so far as to suggest that it is anthropomorphism to think of God as having a mind. What, then, can God be? Spinoza never answers that question directly, but he does say that God's guidance is "the fixed and immutable order of nature." When we say that all things are ordered according to the decree and guidance of God, this is the same as saying that all things happen according to the laws of nature. It is a natural consequence of this view that there can be no miracles, no divine interventions in the order of nature. If there were an event contrary to the laws of nature, that would be an event contrary to divine decree. If God is omnipotent, this is impossible.

God's omnipotence also makes it irrational to conceive of God as a lawgiver of the kind portrayed in the Bible. The biblical God is conceived as being like a king who issues commands which his subjects have the power to obey or disobey. They will prosper if they obey and suffer if they disobey. But the laws which are truly divine are principles of natural necessity—like the laws according to which motion is transferred from one body to another in a collision. No one has any choice but to "obey" these laws; it is not a contingent matter whether someone acts in accordance with them. (Nevertheless, even after stating this conclusion quite clearly early in the TPT, Spinoza regularly adopts some of this anthropomorphic language himself, later in his work, when he argues that the primary purpose of scripture is to encourage obedience to God, not to inculcate correct beliefs about God.)

Although Spinoza questions much of the history and theology of the Hebrew Bible—and delicately avoids any extended discussion of the Christian New Testament—he denies that he has spoken unworthily of scripture. Scripture is divine and sacred when it moves men toward devotion toward God, as it can do and often does. But it is not inherently sacred. If men neglect it, or interpret it superstitiously, as they can and often do, it is no more sacred than any other writing. There is a core ethical teaching in scripture which is so pervasive that it cannot have been corrupted by any misinterpretation: that we should love God above all else, and love our neighbors as ourselves; that we should practice justice, aid the poor, kill no one, covet no one's possessions, and so on. These prescriptions deserve our utmost respect. If we seek to

follow them wholeheartedly, we will be treating scripture as sacred, whether we think of those prescriptions as the commands of a heavenly king or regard them (in the manner of Hobbes) as theorems about what is conducive to our self-preservation and to living in the best way possible.

Spinoza does not endorse only the ethical teachings of scripture. He also thinks there are core theological teachings which are central to scripture and which are in some sense true: for example, that God exists; that he provides for all; that he is omnipotent; that things go well for those who observe their religious duties but badly for the unprincipled; that our salvation depends only on God's grace; and so on. In his way, he does endorse these teachings. But his approval of them is hedged. There is a popular way of understanding them which assumes that the God of whom they speak is a changeable personal agent who acts from freedom of the will, who prescribes laws as a prince does, and who has desires which humans will frustrate if they disobey his commands. And there is a philosophical way of understanding them, according to which God is the fixed and immutable order of nature who acts from the necessity of his own nature and whose "laws" are eternal truths, the violation of which is followed only by natural punishments, not supernatural ones. Presumably the philosophical way of understanding these doctrines is the right way to understand them from the standpoint of truth. But the popular way of understanding them is not to be despised *if* it produces conduct in accordance with the ethical teachings of scripture. If it does, it is to be respected, honored, and encouraged.

Insofar as Spinoza endorses a minimalist theology, which avoids most controversial doctrines, concentrating on those which elicit broad agreement and which emphasizes the importance of works as the path to salvation, the TPT is in the tradition of Erasmian liberalism. This outlook provides him with a religious argument for tolerating diversity of opinion in the realm of religion. Philosophy and theology are separate areas, neither of which should be the handmaiden of the other. Theology is concerned with revelation, which in turn is concerned with obedience, not with speculative truth. In judging whether or not a person's faith is pious, we must look only to his works. If they are good, his faith is as it should be.

In the political portions of the TPT, Spinoza supplements this religious argument for freedom of thought and expression with a political argument. He seeks to show, from fundamental political principles, that allowing this freedom is compatible, not only with religion, but also with the well-being of the state. Indeed, he will go further and argue that the well-being of the state *requires* freedom of thought and expression.

The foundations of his political thought look very Hobbesian; the liberal conclusions he draws from them seem rather un-Hobbesian. Like Hobbes, Spinoza believes that the condition of man in the state of nature—that is, in any state where there is no effective government—is wretched and insecure. Human beings are very egoistic. Everyone seeks what considerations he would develop to be to his own advantage, with little concern for the well-being of others or the long-term consequences of his actions or the moral repercussions for civil society. Moreover, humans generally have an impoverished understanding of what is in their interest, valuing such goods as wealth, honor, and sensual pleasure more than they should, and knowledge and the control of their passions less than they should. If they did not have laws to restrain them, laws which alter their calculations of self-interest, they would not practice justice and loving-kindness; their lives would be full of conflict, hatred, anger, deception, and misery. In the state of nature there is, by definition, no human law to restrain them. And Spinoza takes himself to have shown that God cannot be conceived as a lawgiver. It follows that in the state of nature, though each person is permitted to do whatever he has the power to do, he has no joy from this freedom.

But, like Hobbes, Spinoza also assumes that people are smart enough to see that their condition in the state of nature is wretched and to see what they must do to escape it: create a civil society by agreeing with other people to transfer their power to defend themselves to society, creating a collective entity which will have sufficient power to make and enforce laws for the common protection and advantage. Not only will this arrangement provide them with security, but it will also make possible cooperative enterprises which improve the lives of everyone in the state, enabling them to seek the highest good: the knowledge of things through their first causes, that is, the knowledge of God, which leads to the love of God. (Positing this—or anything else—as our highest good is very un-Hobbesian.)

In some respects, Spinoza goes further than Hobbes in his conception of what the creation of the state involves. He thinks that when individuals agree to form a civil society, they must surrender to it whatever rights they possessed in the state of nature. If they wanted to reserve certain rights to themselves, they would have to establish some means of protecting those rights; establishing these means would divide and consequently

destroy the sovereignty of the state. (Although Hobbes favored absolute sovereignty, he argued that some rights, like the right to defend oneself against attack, were inalienable.) Just as Spinoza thinks that the right of individuals in the state of nature is limited only by their power, so the right of the state is limited only by its power. Since it is not, and cannot be, bound by any laws, what it can do, it may do.

Is the formation of the state, then, really as rational an act as Spinoza presents it as being? The state, which can call upon the collective might of all (or at least, most) of its members, seems potentially much more dangerous to each of its members than any individual in the state of nature. As Locke wrote in response to the similar views of Hobbes, "this is to think that men are so foolish that they take care to avoid what mischiefs may be done them by polecats, or foxes, but are content, nay think it safety, to be devoured by lions." But Spinoza thinks people can rationally run this risk because he thinks that even in a monarchy or aristocracy the state will normally avoid commanding things contrary to the interests of the people. If it did, it would risk losing its power and hence its right to command.

Moreover, in the TPT Spinoza is mainly thinking of the state which emerges from this process as a democratic one, that is, one in which decisions of the state are to be made by a general assembly of all the people. He acknowledges that in certain circumstances other forms of political organization may be desirable. In his posthumously published *Political Treatise* (PT) he recommended ways of structuring monarchies and aristocracies which provide the citizens with protection from their rulers. But in the TPT he focuses most of his attention on democracy, which he regards as the most natural form of government.

In the state of nature all men were equal; they retain that equality in civil society when the state is a democracy because no one in a democracy is subject to his equals. In the state of nature, all people are free because they are subject to no laws; they retain their freedom in civil society insofar as they are subject only to laws in whose formation they have participated—laws, moreover, guided by the principle that the well-being of the people is the supreme law, not the well-being of the ruler. A man can be free even when he is acting according to a command, if the command is rationally aimed at his advantage. Indeed, he is truly free only when he is acting wholeheartedly according to the guidance of reason. (Unlike Hobbes, Spinoza favors a positive conception of liberty, not a negative one which regards it merely as the absence of impediments to the agent's preferred actions.) Rule by one man, or by a few men, might be justifiable if that man (or those men) had some ability which went beyond ordinary human nature. But Spinoza seems to think that this is not normally the case. And, like Machiavelli, Spinoza thinks that the people are less prone to unwise actions than are autocratic rulers.

To those of us who are accustomed to a system in which the actions of government are constrained by a written constitution which provides protection for individual liberties, it may seem that a political theory that calls for men to give up all their rights to the state is an unpromising basis for a defense of freedom of thought and expression—even if the state is a democratic one. Spinoza may have thought, as Rousseau did, that if the legislators are making laws which bind themselves as much as they do others, that fact will provide a sufficient incentive for them not to impose undue burdens. But this thought seems to ignore the possibility that a majority will make decisions that it believes to be for the common good, even if the minority regards them as tyrannical.

In the TPT Spinoza's primary remedy for this problem is not an institutional one. He relies on the facts that in his theory the right of the state is limited by its power and that its power is inevitably limited by the recalcitrance of human nature. Some of the things a state might wish to command are things its citizens cannot change at will, such as their beliefs and their emotions. The threat of punishment for believing or loving as a person does cannot cause that person to believe or love otherwise. But if the state lacks the power to control its citizens' beliefs and actions, then it also lacks the right to control these things.

The fact that the state lacks the right to control what it lacks the power to control, in itself, is no protection. But Spinoza emphasizes that it is impossible for people to surrender their right (or transfer their power) to the state in such a way that they are not feared by the people to whom they have surrendered their right. Any government is in greater danger from its own citizens than it is from any external enemy, for its control over its citizens and its ability to respond to enemies both depend ultimately on the voluntary obedience of a substantial number of its own citizens. Hobbes put the point well in *Behemoth*, his history of the English Civil War: "The power of the mighty hath no foundation but in the opinion and belief of the people.… If men know not their duty, what is there that can force them to obey the laws? An army, you will say? But what shall force the army?" Spinoza would almost certainly not have known *Behemoth*—which was finished in 1668 but first published in a pirated edition in

1679, and then only in English—but he might have come to appreciate the basic point by reading and reflection on Hobbesian works he did know, or by reflection on the works of classical historians like Tacitus and Quintus Curtius, whose writings may also have helped Hobbes see this point.

If all governments are vulnerable to destruction from within, those which seek to rule by violence are the most vulnerable. And no rulers are more violent than those which make it a crime to hold controversial opinions, since they criminalize behavior the citizen cannot change at will. When the government seeks to do what it cannot do, not only does it exceed its right, it also creates resentment among those citizens who feel they are being treated unjustly. It cannot do this without harm to its own power to maintain itself. The most the government can accomplish is to suppress the expression of opinion, not the opinions themselves. But to the extent that it succeeds in suppressing expression, it creates a culture in which people think one thing and say another. It destroys the honesty necessary to the well-being of the state, encouraging deception, flattery and treachery, all of which are destructive of the social order.

What is particularly pernicious about this result is that it makes enemies of just those citizens whose education, integrity of character, and virtue would make them most useful to the state. Spinoza is sometimes portrayed as the epitome of cool rationality, but on this subject he is passionate:

> What greater evil can be imagined for the State than that honorable men should be exiled as unprincipled because they hold different opinions and do not know how to pretend to be what they are not? What, I ask, can be more fatal than that men should be considered enemies and condemned to death, not because of any wickedness or crime, but because they have a mind worthy of a free man? Or that the gallows, the scourge of the evil, should become the noblest stage for displaying the utmost endurance and a model of virtue, to the conspicuous shame of the authorities?
>
> (GEBHARDT III, 245)

Spinoza may be thinking here of cases like that of Judah the Faithful, whom he refers to in his correspondence. Judah was a Spanish *converso* (that is, a Jew forcibly converted to Christianity) who reverted to Judaism. Burned at the stake by the Inquisition when Spinoza was twelve, his case was well-known in the Amsterdam Jewish community. As the flames roared up around him, he sang a hymn which begins "I offer up my soul to you, Oh Lord." He died still singing this hymn. Spinoza cites this case in response to a Christian correspondent who tried to persuade him of the truth of Christianity by citing the many martyrs who had died for their faith. Spinoza's reply was that Judaism claimed, with justice, to count many more martyrs to its faith.

THE HAGUE (C. 1670–1677)

Sometime during the winter of 1669–1670, Spinoza moved to the Hague, first renting a room from a widow and, after about a year, relocating to the home of the painter Hendrik van der Spyck, where he was to live for the rest of his life. In early 1670 the TPT was published in Amsterdam by Jan Rieuwertsz, but with a title page claiming publication in Hamburg, by a fictitious publisher named Heinrich Künraht. Reaction was immediate and vehement. In June the ecclesiastic court of the Reformed Church in Amsterdam condemned the work as "blasphemous and dangerous." Similar denunciations followed from church groups in The Hague, Leiden, and Utrecht. Nor was it only conservative Calvinists who were shocked by his work. Theological liberals, including those sympathetic to the new philosophy, such as Frans Burman and Philip van Limborch, also opposed it. Burman called it an "utterly pestilential book" which must be attacked and destroyed. Between 1670 and 1672 the church authorities repeatedly called for the suppression of the TPT, along with Meyer's *Philosophy, the Interpreter of Holy Scripture* and Hobbes' *Leviathan*.

Nevertheless, there was no formal prohibition of the TPT until 1674, and it did in fact circulate widely among the learned audience to whom it was addressed. This does not mean that the civil authorities tolerated it. De Witt's position seems to have been that the city governments had ample authority, under anti-Socinian legislation passed in 1653, to confiscate copies of Spinoza's book. There was no need to increase the notoriety of this book, and its sales, by calling special attention to it. In many parts of the Republic the civil authorities did make efforts to suppress it, as they did in the other countries to which it spread. That these efforts did not prevent the work from being widely read was due to the ingenuity and dedication of Spinoza's publisher. However, when Spinoza learned that Rieuwertsz had commissioned a Dutch translation of the TPT which would have made it available to a wider audience, he asked that it be withheld, as it was until sixteen years after his death.

1672 has been called a "year of disaster" in the history of the Dutch Republic. In March, England resumed its

naval war with the Republic, attacking a Dutch convoy. In April France declared war. In May the French army began its invasion, followed quickly by two German armies, under the Prince-Bishop of Münster and the Elector of Cologne. The overwhelming forces of the invaders quickly conquered most of the Dutch provinces. Only by opening the dykes to flood a large swath of land, from the Zuider Zee in the north to the river Waal in the south, was the government able to prevent the invaders from occupying the province of Holland.

These were extreme and unpopular measures. The people were deeply divided between those who wanted to surrender and those who wanted to resist. In June, Jan de Witt, wounded in an assassination attempt, resigned his position as Grand Pensionary, leader of the States of Holland. William III had been appointed captain-general of the army in February; in July he became Stadholder of the provinces of Zeeland and Holland, and the dominant political power in the Republic. In August De Witt's brother, Cornelis, who had been imprisoned on a charge of plotting against the Stadholder's life, was acquitted. When an angry mob gathered outside the prison where he was being held, Jan went to the prison to escort his brother to safety. The mob murdered both brothers, dismembering their bodies, roasting them and eating them. When Spinoza learned of this, he tried to rush into the street, carrying a sign reading *ultimi barbarorum*, "the worst of barbarians." Fortunately, his landlord prevented him from carrying out this act of protest.

In 1673 Spinoza had an opportunity to leave the Netherlands when the University of Heidelberg offered him a professorship. It appears that the Elector Palatine, who was responsible for the offer, knew Spinoza as the author of a highly regarded exposition of Descartes but not as the author of the TPT. He charged a professor at the university, Louis Fabritius, with the task of making the offer. Fabritius knew that Spinoza was the author of the TPT. He couched his offer in terms which he probably knew Spinoza would refuse, assuring him that he would have "complete freedom to philosophize" but noting that the Elector assumed Spinoza would not "misuse use that freedom to disturb the publicly established religion." In declining, Spinoza gave two reasons: first, he feared that teaching would interfere with his research, and second, he did not know what limits he would have to impose on himself to avoid appearing to disturb the established religion.

By 1675 Spinoza was satisfied enough with his revisions of the *Ethics* that he visited Amsterdam to give the manuscript to Rieuwertsz for publication. But the theologians learned of his plans and complained to the civil authorities. So Spinoza gave up on this attempt to publish his masterwork, leaving it to appear in his *Opera posthuma*.

In his last years Spinoza began two additional works which he did not live to finish: his *Compendium of Hebrew Grammar* and his *Political Treatise*. Both these works are in some sense byproducts of the TPT. The biblical criticism of the TPT had emphasized that to understand scripture it was essential to understand the language in which it was originally written. But Spinoza believed no existing grammar explained it adequately. The *Hebrew Grammar* was intended to fill that gap. And although the TPT had provided foundations for political philosophy, it had not dealt with practical questions about the merits of the different forms of government and the best ways of organizing them. The *Political Treatise* aimed to remedy that lack.

In February 1677 Spinoza died of a debilitating lung disease, probably aggravated by inhaling the glass dust produced by grinding lenses. By December his posthumous works were published in nearly simultaneous Latin and Dutch editions, the *Opera posthuma* and the *Nagelate schriften*. Because the Dutch translations must have been done from manuscripts rather than from the printed text of the Latin edition, the Dutch translations provide a check on the proofreading of the editors of the *Opera posthuma*, a fact which has aided recent critical editions of Spinoza's works. The Latin edition included the *Ethics*, the correspondence (originally seventy-five letters to and from Spinoza), and three unfinished works, the *Treatise on the Intellect*, the *Political Treatise*, and the *Hebrew Grammar*. Neither edition included the *Short Treatise*, manuscripts of which were not discovered until the nineteenth century. Subsequent scholarship has also added twelve letters to the correspondence. We'll conclude with an account of the works which first appeared posthumously, beginning with the *Political Treatise*.

THE POLITICAL TREATISE (1677)

Though Spinoza expressed a strong preference for democracy in the TPT, he also recognized that it might not be the most suitable form of government for all situations. Like Machiavelli, whose work he studied closely, he thought it was not an easy matter to impose a new form of government on people who had become accustomed to a different form. So part of what he seeks to do in the *Political Treatise* (PT) is to work out principles for organizing the alternatives he regards as inherently less desirable. He offers detailed proposals for the best way to

organize a monarchy or an aristocracy so that it can be stable and serve the interests of its citizens as well as possible.

The sensible design of any form of government must take into account the known features of human nature. For example, because no one has so powerful a mind that he always sees the good and never yields to his passions, and because "kings are not Gods, but men, who are often captivated by the Sirens' song," even in a monarchy it is unwise to put all decision-making power in the hands of one man. If it is necessary to have a monarchy, the king should be guided in his decisions by a large, broadly based council of advisors. Indeed, Spinoza proposes that the king be required to choose from among the proposals recommended by his council. He does not explain how this requirement is to be enforced.

Similarly, he thinks an aristocracy will work best if the power to make and repeal laws, and to appoint ministers of state, is granted to a large council drawn from the patrician class. He regards the size of that council as critical to its proper functioning, on the theory that the larger the deliberative body, the more apt it is to have in it some men outstanding for their wisdom, and the less apt it is to favor irrational policies. But he would provide a smaller council of syndics, also drawn from the patrician class, to insure that the legislative council follows the prescribed procedures and that the ministers faithfully execute the laws.

Spinoza intended to add a discussion of democracy to this work but lived to complete only a few paragraphs on that topic. What he does say about democracy has embarrassed many of his modern admirers because he excludes women from the political process on the ground that they are naturally unequal to men (and because men are apt to overrate the intelligence of beautiful women). We can only speculate about what else he might have said, but it seems likely that he would have acknowledged that even democracy—understood as a form of government in which all adult males who are neither servants nor criminals nor men of ill repute are entitled to vote in the legislative assembly and to hold political offices—has inherent problems that require some form of constitutional protections.

THE COMPENDIUM OF HEBREW GRAMMAR (1677)

As indicated above, Spinoza undertook this work because he believed that a thorough understanding of biblical Hebrew was essential for interpreting scripture, that no existing Hebrew grammar provided an adequate understanding of the language, and that he could succeed where his predecessors had failed. The first of these reasons would generally be acknowledged as valid. The second may have been true in Spinoza's day but is probably not true now. To what extent Spinoza's grammar has contributed to our improved understanding of the Hebrew language and the Bible is a matter for historians of Hebrew linguistics and biblical scholarship to judge. The primary question here is whether this work contributes anything to our understanding of Spinoza's philosophy. Regrettably the answer to that question seems to be "no."

ETHICS (1677)

The most important work included in the *Opera posthuma* is the *Ethics*, a systematic account of Spinoza's philosophy written in a style modeled on Euclidean geometry, beginning with a set of axioms and definitions, and attempting to show, by formal demonstrations, what conclusions these assumptions lead to. From time to time Spinoza interrupts the construction of proofs to elaborate on particularly important topics, in prefaces, scholia, and appendices. These tend to contain his most accessible and memorable passages. But the bulk of the work is written in a format which increases its difficulty for many readers, however much they may admire the commitment to rigor. The formal definitions Spinoza gives of his key terms sometimes raise more questions than they answer. The axioms are not always intuitively obvious. And the demonstrations are not always perspicuous. The forbidding style of the work may explain why, for the first hundred years after Spinoza's death, the TPT was the most influential of his main works. It was only toward the end of the eighteenth century that the *Ethics* began to find an appreciative audience.

Some of the difficulty of the work may be alleviated by recognizing that Spinoza does not expect his readers to find all the axioms obvious or all the demonstrations compelling. He arrived at his final set of axioms only by trying out different axiomatizations on his correspondents and modifying them in response to criticism, supplying arguments for assumptions the correspondents questioned. Often he provides more than one demonstration of a proposition, recognizing that his readers may not be convinced by the first demonstration. And at one point, having come to a conclusion he expects his readers to find particularly surprising, he implores them to refrain from judgment until they have followed the argument carefully to its conclusion. The implication seems to be that the system is to be judged partly by its ability to

explain, comprehensively and consistently, a wide range of data.

The work is divided into five parts. The first attempts to demonstrate the existence of God and determine his properties; the second explores the nature of the mind, with particular attention to the human mind; the third gives an account of man's emotional nature, systematizing what Spinoza takes to be the laws of human psychology; the fourth seeks to explain why we are so often the victims of self-destructive passions and propounds an ideal of human nature we can and should strive to attain; the fifth part tries to show how we can control our passions and achieve blessedness.

In the Appendix to Part I, Spinoza provides a useful summary of its main conclusions: that God, defined as a substance consisting of infinite attributes, exists necessarily; that God is the only substance, everything else being a mode of God; that God is the free cause of all things; that everything else is so dependent on God that it cannot be or be conceived without him; and that God has predetermined all things, not from freedom of the will, but from the necessity of his nature.

To this we might add that Spinoza also claims to show in Part I that infinitely many modes follow from the necessity of the divine nature. Some of these modes follow from God's *absolute* nature—that is, follow from God's nature unconditionally—and hence are themselves infinite and eternal. Other things—particular, finite things—express God's attributes in a determinate way, and do not follow from God's absolute nature, but from one of God's attributes insofar as it is modified by another modification which is also finite. So each finite mode has as part of its causal history an infinite series of other prior, particular, finite things.

Spinoza is often referred to as a pantheist, a term usually taken to mean that God is identical with nature, understood as the totality of things. But Spinoza identifies God with nature only in the sense that he identifies God with His attributes, those eternal elements in nature which exist in themselves and are conceived through themselves. When Spinoza identifies God with Nature, it is with what he calls *Natura naturans* (active or productive nature). The modes which follow from and express God's attributes he calls *Natura naturata* (passive or produced nature) (*Ethics* I, Prop. 29, Schol.). They are not God. Their defining properties are logically opposed to God's: they exist in another, through which they are conceived. Nor are they a part of God, since it is incompatible with God's nature to have parts (*Ethics* I, Prop. 29, Schol.).

Because everything which exists is either an attribute, whose existence is absolutely necessary, or a mode, and because all modes either follow from God unconditionally or else are necessary in relation to other modes of God, Spinoza concludes that there is nothing contingent in nature. All things are determined by the necessity of the divine nature to exist and act as they do. God could not have produced them in any other way than He did.

This is what Spinoza says. What does it mean? From the seventeenth century to the twenty-first many interpreters have understood the doctrine that there is only one substance, of which everything else is a mode—Spinoza's monism, in effect—as implying that there is only one ultimate subject of predication and that everything else is in some way a predicate of that one subject. This is a prima facie plausible way to understand his monism, given the close historical connection between the idea of substance and the idea of an ultimate subject of predication. But it is not obviously an attractive way of understanding Spinoza's monism on reflection. In what sense might a particular thing, like a human being, for example, be predicated of God?

When Pierre Bayle advanced this line of interpretation in the seventeenth century, he took it to imply that the properties of finite things must really be properties of God. And he understandably thought Spinoza's monism, so interpreted, was absurd. God would be constantly changing his properties as the properties of finite things changed (though Spinoza insists that God is immutable). He would have unseemly human properties, insofar as people behaved improperly or criminally (though Spinoza is resolutely opposed to anthropomorphism). And he would have contradictory properties at the same time, as one finite thing had one property and another had its opposite.

In the late twentieth century Jonathan Bennett (1985) advanced a variation on Bayle's interpretation which avoids some but not all of these unhappy consequences. He suggested that when we say of a finite thing that it has a certain property, what we are really saying is that the universe, conceived under one of God's attributes, has some property at a certain location. That property is not necessarily the one we ascribe to the finite thing. For example, when we attribute a property to a physical object, we are saying that the universe, conceived under the attribute of extension—that is, space itself—has some property at that particular point. If I say that the peach I am about to peel is ripe, I am saying that space has, in that region, some quality I conceptualize as

ripeness; I am not attributing ripeness either to that region of space or to space as a whole.

This interpretation avoids the problem of ascribing contradictory properties to God by understanding apparently contradictory predications as applying to the universe *at different locations*: Space is qualified here by whatever property I conceptualize as ripeness; it is qualified there by whatever property I conceptualize as unripeness. (How this works for modes of attributes other than extension is unclear.) It avoids the problem of ascribing human properties to God by remaining agnostic about the properties of space which underlie the properties we ascribe to humans and other finite things. (We do not know what properties of the universe underlie the fact that I love someone whom you do not love, but they are evidently not human properties.) But it does not avoid the problem that, on this view, God is constantly changing. Whenever some finite thing changes, God is changing at that location.

The main alternative interpretation (Curley 1969, 1988) emphasizes the equally strong traditional connection between the idea of substance and the idea of independent existence. When the TEI first introduced the contrast between things which exist in themselves and things which exist in something else, Spinoza glossed that contrast as one between things which are their own cause and things which are caused by something other than themselves. He did not explain it in terms of predication. The things which exist in themselves—the first elements of the whole of nature—were supposed to be fixed and eternal and to have laws "inscribed in them." If we identify the first elements of the whole of nature with the attributes, then we can infer that Spinoza conceived attributes like thought and extension as eternal entities involving laws of nature so fundamental that they do not admit of explanation in terms of anything more basic. On this reading Spinoza dreamed of a final scientific theory whose most basic principles would be, and could be seen to be, absolutely necessary. That is why the attributes exist in themselves and are conceived through themselves.

According to this interpretation, some things follow from the fundamental laws without the aid of any other propositions. These are the eternal, immutable things which follow from God's *absolute* nature, the infinite modes of the *Ethics* (or universal modes in the ST). They are those general features of reality corresponding to the derived laws of nature, like motion and rest, which involve laws pertaining to anything possessing motion or rest. They follow from the attributes because the lower level laws can be deduced from and hence explained by the most fundamental laws. (Spinoza provides us with a sketch of such a deduction in Part II of the *Ethics*.) Although these modes themselves are infinite (in the sense that the laws they involve apply throughout nature) and eternal (in the sense that the laws are immutable), the series of causes which produces them is finite. Explanation of one law by another deduces the less general law (say, a law governing the transfer of motion in a particular kind of impact) from more general laws (say, the law that motion is conserved in all causal interactions). The series of general causes must come to an end because there is a logical limit to the generality of laws. Once you have formulated a law so general that it applies to everything which possesses a certain attribute, no more general law is possible. It is thus in the nature of the attributes that they cannot be explained through anything else.

Other things—the finite modes of the *Ethics*, the singular changeable things of the TEI—do not follow from the *absolute* nature of God's attributes but do follow from God's attributes as modified by the infinite modes *and* other finite modes. This is a reflection of the fact that particular events cannot be explained by laws alone but require information about other particular events for their explanation. Their necessity is not absolute but relative to the existence of the other events essential to their explanation.

Something of this sort must be true if Spinoza's system is to allow for the reality of change. Spinoza insists that things follow from God's nature with the same necessity with which the properties of a triangle follow from its nature. This is why he is often criticized for assimilating the causal relation to that of entailment. If *everything* followed logically from the *absolute* nature of God, which is eternal and immutable, nothing could fail to share in that eternity and immutability. Because the infinite modes *do* follow from God's absolute nature, they share the eternity and infinitude of their cause. But not everything follows from God's absolute nature. Specifically, the particular finite things do not follow unconditionally from the infinite and eternal things. So its members are not infinite and eternal. This is why change is possible.

This dependence of the finite on other finite things also explains why the world must have no beginning. It contains particular things whose behavior can only be explained if we add information about antecedent conditions to the general facts we appeal to in our explanation. Those particular things constitute a series which cannot have an end, because each member of the series must have an explanation and can only be explained by the

existence of some particular thing(s) prior to it (plus the laws of nature).

This reading of Spinoza's metaphysics has the advantage of identifying something in nature—the first elements of nature, or the attributes—which can plausibly be thought to be eternal, immutable, ultimate principles of explanation for everything else in the universe. Because Spinoza's system requires something eternal and infinite as an object of the love which is supposed to provide us with pure joy, this seems an important consideration. This reading also has the advantage of identifying something in nature which can plausibly be thought to follow logically from the first elements alone and to function as an intermediate between the ultimate principles of explanation and the finite things whose behavior is to be explained. The idea that there is a series of infinite and eternal things intermediate between God and finite things is one of the most distinctive features of Spinoza's metaphysics in contrast to Cartesianism.

This reading also has what may be thought to be a disadvantage: it implies that not everything in nature is absolutely necessary. The finite modes are portrayed here as not following unconditionally from the fixed and eternal things but as requiring other finite modes for their explanation and as being necessary only in relation to those other finite modes and the infinite modes. But this feature of the interpretation may not really be a disadvantage; Spinoza's discussion of necessity suggests that he thought things are necessary in two very different ways (*Ethics* I, Prop. 33, Schol. 1). Some are necessary in virtue of their own nature; others are necessary in virtue of their cause. Particular finite things, such as this or that human being, do not involve any inherent necessity (*Ethics* II, Ax. 1). They are necessary just insofar as the order of nature (the series of prior finite causes) makes them necessary.

The theory of mind-body identity in Part II of the *Ethics* is best approached by viewing it as a subversion of Cartesian dualism. Descartes sought to make belief in personal immortality rational by showing that the mind and the body are really distinct from one another. His strategy was to set up a thought experiment in which we clearly and distinctly conceive the possibility of the mind's existing without the body. We can, he claimed, find reasonable grounds for doubting the existence of the whole physical world by reflecting on the powers of God. An omnipotent being could, if he chose, create in us representations of physical objects without creating any physical objects. But we cannot find reasonable grounds for doubting our own existence as thinking things. Any hypothesis we entertain to cast doubt on our existence, such as deception by God, will entail that we think, and hence, that we exist. So we are compelled to affirm our existence as thinking things but not compelled to affirm the existence of our body (or any other extended object).

If we can clearly and distinctly conceive of the mind as existing without the body, then it is logically possible for it to exist without the body. If it is logically possible for it exist without the body, then it *could* exist without the body. (If it is logically possible for two things to exist separately, then an omnipotent being could cause them to exist separately. And Descartes thinks he has shown that there is an omnipotent being. So the possibility of their existing separately is not merely a logical one. There is a being which has the power to bring this about, if he wishes.) But if two things are such that one can exist without the other, they are really distinct. This entails that the mind is not necessarily destroyed when the body is destroyed, and that establishes the *possibility* of immortality. Whether that possibility is realized depends on the inscrutable will of God. So Descartes makes no serious attempt to prove actual immortality.

Descartes did, however, modify the strictness of this dualism when he added that the mind is not present in the body "as a sailor is present in his ship," that it is, instead, closely united to it, so that mind and body together constitute one thing and are "substantially united." What seems to have motivated this doctrine of substantial union—which is not obviously consistent with the dualism—was Descartes' recognition that there is a particularly intimate connection between the human mind and the human body. When something happens in my body, normally I am not aware of it in the external way in which I am aware of things which happen in bodies not mine. I feel my body's need for food as hunger, its need for drink as thirst, damage done to it as pain, and so on. These interested, action-motivating bodily sensations are what make this particular body peculiarly mine.

Spinoza, too, seems to have been deeply impressed by the intimacy of the relationship Descartes described, and particularly by the facts that the mind's capacities are a function of those of the body and that changes in the mind strictly parallel those in the body. For example, my mind's capacity for higher-level thought seems to be a function of my brain's complexity; its ability to think clearly and its mood are both closely correlated with my body's blood alcohol level. A Cartesian might dismiss some of these phenomena as mere coincidences. Others he might regard as examples of the body acting on the mind. But Spinoza thinks that because mind and body belong to such fundamentally disparate conceptual cate-

gories, we cannot posit a causal relationship between them. And he would not dismiss any such regularity in nature as a coincidence. What we should say instead is that the mind and the body are one and the same thing conceived under different attributes.

Spinoza has a metaphysical argument for supposing that this identity of modes of thought with modes of extension exists not only in human beings but also runs throughout the whole of nature. Suppose that God is an infinite, perfect substance who possesses the attributes of thought and of extension. As an infinite and perfect thinking thing, he must have in his intellect an idea of every existing mode of extension. If he did not, there would be gaps in his knowledge. Equally, as an infinite and perfect thinking thing, he cannot have in his intellect an idea of a mode of extension as existing if no such mode exists. If he did, he would be in error. So in God there must be a one-to-one correspondence between the modes of extension which exist and their representations in God. Moreover, since this correspondence is necessary, it is not possible for the modes of thought to exist without their corresponding modes of extension. The converse is also impossible. This entails that the modes of thought and the modes of extension are not, in Cartesian terms, really distinct from one another. They are conceptually distinct, insofar as they are conceived under different attributes. This is why there can be no causal relation between them. But they are not capable of existing apart from one another.

This argument leads to some surprising conclusions from which Spinoza does not shrink. For example, it entails that every extended thing in nature corresponds to a mode of thought which is, in some sense, its "mind." This doctrine is known as panpsychism. Spinoza clearly does think that all finite physical things other than humans have something *like* the minds humans have. Insofar as he affirms a continuity between humans and other animals, his panpsychism seems quite reasonable, much more reasonable than the Cartesian view that nonhuman animals are merely machines without any sensations. Moreover, other philosophers before Spinoza—like Montaigne—had argued that animals were capable of displaying intelligence and emotions. What is puzzling about Spinoza's panpsychism is its apparent implication that even the simplest material objects have something like a mind. We can diminish the shock of this claim to some degree by recollecting that Spinoza would probably not think that the minds of the simplest material objects are *very much like* human minds. If our capacity for higher-order thinking depends on our having a very complex brain, then presumably a carbon atom does not have the capacity to solve quadratic equations. But it is still unclear what the ascription of mentality to very simple physical objects comes to.

One unsurprising consequence of this view of the relation between mind and body is that Spinoza denies that the mind is capable of acting freely in the way Descartes tended to understand freedom. Descartes was quite ambiguous about the kind of freedom he wanted to claim for us. In the Fourth Meditation he seemed, initially, to interpret freedom of the will indeterministically, as a power to either do something or not do it, independently of any external causes. Then he reflected that there were two cases where he might not, in fact, be able to act otherwise, though he did not want to deny that he was free in those cases: one is the case where he sees something so clearly that he cannot help but assent to it; the other is the case where God, in an act of grace, disposes his inmost thoughts in a certain way. So he revised his initial definition, adding a clause which would make freedom compatible with certain kinds of determinism: we can be free if our intellect presents something so clearly to the will that it cannot judge otherwise; and we can be free even if God is determining our actions, so long as we are not aware of that determination, so long as we seem to ourselves to be the initiators of our actions. But this was another area where he was unable to maintain consistency. In the *Principles of Philosophy* he reverted to an indeterminist conception of freedom and pronounced the problem of reconciling human freedom with God's preordination of all things insoluble.

Spinoza rejects any indeterminist conception of freedom. This was evident already in Part I of the *Ethics*, where he held that all finite things are determined to exist and act the way they do by an infinite series of prior finite things. But his acceptance of mind-body identity provides an additional reason for denying indeterminism in humans. Descartes would have allowed that determinism reigned in the physical world except insofar as minds were capable of intervening in it to cause events which would have gone differently but for that intervention. If the mind and the body are one and the same thing, conceived in different ways, then the mind will not be able to intervene in the physical world as an uncaused cause. The decisions of the mind are just the appetites of the body, conceived under a different attribute. When they are conceived under the attribute of extension, they are conceived as part of a causal network which determines their character. Since the order and connection of ideas mirrors the order and connection of extended things, modes

of thought must also be part of a causal network that determines their character, a network whose members are conceptually distinct from, but really identical with, the corresponding modes of extension. Spinoza concedes that it often seems to us that our acts of will have no antecedent causes; but he thinks all this shows is the inadequacy of our self-knowledge.

Consistently with this deterministic picture of things, Spinoza turns in Part III of the *Ethics* to an attempt to provide a systematic human psychology, explaining the laws according to which the human mind operates. He writes in the Preface to Part III,

> Nothing happens in nature which can be attributed to any defect in it, for nature [read *Natura naturans*] is always the same, and its virtue and power of acting are everywhere one and the same, i.e., the laws and rules of nature, according to which all things happen, and change from one form to another, are always and everywhere the same.
>
> (GEBHARDT II, 138)

So, if we are to understand anything, we must understand it in terms of the universal laws of nature. When we understand human actions and emotions in this way, we will no longer be disposed to curse them or find them ridiculous. We will see them as an inevitable result of the circumstances under which they occurred.

Like Hobbes, Spinoza makes the striving to persevere in existence the fundamental law of human behavior. He sees an analogy between that striving and the principle of inertia which was fundamental in the new physics and treats it as constituting the essence of each individual. His conviction that there is this analogy leads him to a revised understanding of what constitutes human activity: We should think of ourselves as active just to the extent that our actions can be adequately understood in terms of our striving to persevere in being. But he also thinks of the striving as encompassing more than just continuation in existence. In addition, it seeks to increase our perfection, or power of action. When we succeed in doing that, we experience the increase as joy; when our power of acting is diminished, we experience the decrease as sadness. In a way, Spinoza is a hedonist. We seek to maximize our joy and minimize our sadness. But the underlying changes in perfection, or power of action, are really at the core of these strivings.

Spinoza's psychology is generally egoistic in the sense that he thinks what we basically seek, insofar as we are active or self-determined (that is, insofar as what we do is determined by our own nature) is something we imagine to be good for ourselves (that is, to involve or lead to our joy). But his egoism does not exclude our taking an interest in the interests of others. If we conceive an external object—a person, or an institution, say—as a cause of joy in us, we will love that object and seek our own good by seeking its good. Similarly, if something in itself neutral is associated in our experience with something either positive or negative, we will come to have positive or negative feelings toward the inherently neutral thing. And to the extent that a thing is like us in some degree, we will tend to share its feelings: to feel sadness when it is sad, and joy when it is joyful. This is the psychological basis for pity and benevolence. We can minimize our own sadness and maximize our own joy by seeking to minimize the sadness of others like us and maximizing their joy.

These are fairly simple and benign cases. But the same psychological laws which explain pity and benevolence also explain, less happily, racial and religious hatred. We are less apt to feel sympathy for those we think of as unlike us. And we are apt to generalize to a whole group the negative emotions we have experienced toward some members of that group. What interests Spinoza most in human psychology is the complexity of our emotions and the psychological conflicts we regularly experience. If something affects us with both joy and sadness, we will feel conflicting emotions of love and hatred; a similar process will unfold if we imagine that something which usually affects us with sadness is like another thing that usually affects us with joy. The uncertainty of our knowledge of human affairs makes us prey to both hope and fear, which are inseparable from each other. But we are subject to wishful thinking, which inclines us to believe the things which give us hope. That is the root cause of superstition. And acting on irrational beliefs is a recipe for disappointment and despondency. Hatred, envy, and jealousy are as natural to us as love, benevolence, and friendship. These conflicting emotions are constantly fluctuating as external circumstances change, with the result that "we toss about, like waves on the sea, driven by contrary winds." For the most part we are not the masters of our fate.

Because Spinoza is a determinist who takes his doctrine to imply that we should bear calmly both good fortune and bad and condemn no one for his behavior, and because he frequently embraces subjectivist-sounding theories of ethical language—as when he writes that good and evil are nothing positive in things, considered in themselves, but just modes of thinking—it has often been thought that he has no ethical theory—or at least that he cannot consistently have one. But Spinoza called his mas-

terwork *Ethics*, and Part IV of that work is full of what look like ethical judgments. He tells us that the knowledge of God is the mind's greatest good, that joy in itself is good and sadness evil, that pleasure can be excessive and evil, that pain can be good, that love can be excessive, that hatred can never be good, and so on. How can these judgments be true if good and evil are only "modes of thinking"?

The answer seems to be that Spinoza makes a distinction between the ordinary, nonphilosophical use of ethical terms, which is highly subjective and undisciplined, and the philosophical use of the same language. If we reflect on the use of terms like *good* and *evil* in connection with members of a natural kind, like man, we will recognize that they signify varying degrees of approximation to an ideal of perfection or completeness. Unaided by philosophy, we are apt to have varying conceptions of that ideal. But there is a way of conceiving the ideal human being which will necessarily attract us as soon as we form a clear idea of it. Spinoza uses the term "free man" as a label for that ideal and the term "good" as a label for those things we know will help to achieve our goal.

The free person is defined as one who is led by reason alone and characterized by his disregard of death and concentration on life; by his willingness to accept risks, when that is called for, and his wisdom in determining when it is not called for; by his determination to avoid the favors of the ignorant, when accepting them might compromise his integrity; by his gratitude to other free men for their acts of genuine love and friendship; by his honesty; and by his obedience to the laws of the state, not from fear of punishment but from his commitment to the common good.

The psychology of Part III holds that all men, to the extent that they determine their own actions and are not the slaves of fortune, pursue what they take to be their own good. The ethical theory of Part IV holds out the ideal of the free man as an enlightened egoist. Freedom is not mere self-determination but informed self-determination. The free man recognizes that, left to himself, he would lead a miserable life, that achieving his optimal state requires the cooperation of other men, that nothing is more useful to him than his fellow men, and that they are the more useful the more they share his dedication to the pursuit of knowledge, a noncompetitive good which is only increased, not diminished, by being shared. He is not an ascetic. He knows that his body requires the moderate use of pleasant food and drink, and that beautiful natural objects and works of art, music, theater, and other such things are goods anyone can enjoy without detriment to others. He understands that the greater the joy with which we are affected, the greater the perfection to which we gravitate, and the more we participate in the divine nature. Spinoza is apprehensive about human sexuality, knowing how easily sexual desire can become obsessive and self-destructive.

The central problem of ethics for Spinoza is not that of knowing what is good but that of pursuing it single-mindedly. "I see and approve the better," he writes, quoting Ovid, "but I follow the worse." Parts III and IV are concerned with explaining why we are often unable to pursue the good we clearly see. Part V tries to help us overcome the unhealthy dominance of the passions which underlies this weakness of the will. Descartes, whose moral philosophy was heavily influenced by the Renaissance revival of stoicism, thought that the mind could exercise an absolute control over the passions. Spinoza is not so optimistic. But he does think that we can increase our power over them and make them less harmful to us.

One promising remedy for our harmful passions is to correct the false beliefs they often involve. Most of the emotions Spinoza analyzes in Part III incorporate some cognitive element. He defines hatred, for example, as sadness accompanied by the idea of something external to us as the cause of our sadness. Indignation is hatred toward someone whom we imagine as having done evil to someone (or something) else. If we come to understand that the person we hate or toward whom we feel indignation is at most a partial cause of those negative consequences, that his actions are no more than the most recent link in a chain of causes which extends into the infinite past, this will diminish our negative emotions toward that individual, redirecting them toward the prior causes and diffusing them over those causes. This process may not immediately diminish our overall level of negative emotions. But if it diminishes the negative feelings we have toward the proximate cause of our sadness, it may make it easier for us to behave well toward that person and break the vicious circle of harm and retaliation which is the cause of so much human misery.

Part V of the *Ethics* concludes with a puzzling series of propositions dealing with the eternity of the mind. Astonishingly, given his earlier doctrine that the mind and body are one and the same thing, conceived under different attributes, Spinoza now maintains that the human mind is not entirely destroyed with the body but that something of it remains which is eternal. The eternal portion of the mind is apparently the part which under-

stands things "under a species of eternity," that is, that sees them as necessary by understanding them under the second or third of the three kinds of cognition which the *Ethics* assumes, reason or intuitive science. Because Spinoza assumes that it is possible to increase our understanding of things by the second and third kinds of cognition—understanding more things in those ways at one time than we do at another—this implies that we can increase the portion of our mind which is eternal, even though eternity is supposed to entail that whatever is eternal has no relation to time. We can make sense of much of Spinoza's philosophy, but so far this part of the *Ethics* has resisted the best efforts of sympathetic interpreters. It is clear that it is not a doctrine of personal immortality, for Spinoza regards memory of the individual's past as essential to personal identity, and he is quite emphatic that the portion of the mind which is eternal has no memory of any past. Perhaps the best thing we can say is that Spinoza thought that there was some truth, badly articulated, in the traditional doctrine of personal immortality and thought (wrongly) that his philosophy could give a coherent explanation of that truth.

In another way, however, Spinoza may achieve some reconciliation with traditional religion in these final portions of the *Ethics*. Because he identifies God with nature (*natura naturans*), he can claim that the more we understand Nature, the more we understand God. When we understand nature by the third kind of cognition, intuitive science, we not only have the highest form of cognition we can have, but we also experience the greatest possible satisfaction. We then experience joy accompanied by the idea of God as the cause of our joy. This means that we love God. Together the knowledge of God and the love which is inseparable from that knowledge constitute our highest good, not because God is a king who will reward us with a happiness extrinsic to our love for him but because the knowledge and love of God inherently involve the highest happiness we can know.

This attempt at an accommodation with traditional religion may not succeed. It is true that Spinoza's "God" has many of the properties of God, as the concept of God came to be developed by philosophically minded theologians in Judaism and Christianity: He is a perfect being, infinite, eternal, the first cause of all things, himself neither needing nor being susceptible of any explanation. Because, in Spinoza's view, knowledge of God can be the cause of the greatest joy we can experience, he can be the object of a love which surpasses any love we can have for finite things. But because, according to Spinoza, God is supremely perfect, he is as incapable of joy (passage to a greater perfection) as he is of sadness (passage to a lesser perfection). So he is also incapable of love or hate, which are species of joy and sadness. We cannot rationally expect Spinoza's God to return our love. Nor can we expect him to watch over us like a loving father. Spinoza's God, being perfect, has no goals, no states he desires to reach (or maintain). To ascribe desire to Spinoza's God would be to conceive him as imperfect, a contradiction in terms. A fortiori, he is not seeking our welfare and cannot provide a refuge from the uncertainty of fortune. He cannot be affected by prayer or ritual. He does not issue laws accompanied by promises of reward for obedience and threats of punishment for disobedience. His laws are ones we cannot break.

Because Spinoza's God differs in so many respects from the God of traditional religion, even in its most philosophical forms, it is understandable that many religious-minded critics have regarded his philosophy as a form of atheism. But from Spinoza's point of view these criticisms only show a misunderstanding of the nature of God. The founders of the traditional religions, he thinks, were in a position like that of the first students of geometry, when geometry was still an empirical science. Relying on what Spinoza would call imagination, the early geometers had only very crude ideas of the objects they were studying. They could not have given a properly scientific definition of a triangle or a circle from which they could demonstrate precise theorems about the nature of these objects. So they made mistakes about them, thinking, for example, that the ratio of the circumference of a circle, to its diameter is 3:1.

But though they may not have had the same definitions of these objects as later geometers, they were still attempting to develop a theory of the same objects. They were just handicapped by the inadequate ideas they had about those things. Similarly handicapped by their reliance on imagination—on the dreams of prophets and reports of revelation passed down through tradition—the philosophers and theologians of the organized religions got some things right and many things wrong. They saw the truth, not clearly, but as if through a cloud. Spinoza's claim not to be an atheist depends on whether he was, as he believed, the Euclid of theology. Spinoza's admirers have inclined to the view that he was.

On the two hundredth anniversary of his death a collection was taken to erect a statue to Spinoza in the Hague. When the statue was unveiled in 1882, Ernest Renan concluded his address with words which sum up the feelings of those admirers: "Woe to him who in passing should hurl an insult at this gentle and pensive

head… This man, from his granite pedestal, will point out to all men the way of blessedness which he found; and ages hence, the cultivated traveler, passing by this spot, will say in his heart, 'The truest vision ever had of God came, perhaps, here.'"

See also Bayle, Pierre; Bennett, Jonathan; Boyle, Robert; Cartesianism; Democracy; Descartes, René; Determinism and Freedom; Essence and Existence; Ethics, History of; Galileo Galilei; Hobbes, Thomas; Human Nature; Jewish Philosophy; La Peyrère, Isaac; Laws, Scientific; Machiavelli, Niccolò; Mind-Body Problem; Panpsychism; Philosophy of Mind; Regius, Henricus (Henry de Roy); Rousseau, Jean-Jacques; Spinozism.

Bibliography

WORKS BY SPINOZA

The Collected Works of Spinoza. Vol. 1. Edited and Translated by Edwin Curley. Princeton, NJ: Princeton University Press, 1985. The second and final volume is in progress.

Spinoza Opera. 4 vols. Edited by Carl Gebhardt. Heidelberg: Carl Winters Universitäts-buchhandlung, 1925.This is the current standard original-language edition, which will eventually be superseded by Moreau's edition.

Spinoza Oeuvres. Edited by Pierre-François Moreau. Paris: Presses Universitaires de France. As of June 2005 only one volume had appeared (in 1999), containing the *Theological-Political Treatise*, with a critical edition of the text by Fokke Akkerman and a translation by Moreau and Jacqueline Lagrée. A comprehensive edition is planned, with critically edited versions of the original language texts and French translations on the facing pages.

Spinoza, Complete Works. Edited by Michael L. Morgan and translated by Samuel Shirley. Indianapolis: Hackett, 2002. This is the only complete English-language edition.

WORKS ABOUT SPINOZA

Allison, Henry. *Benedict de Spinoza: An Introduction*. New Haven, CT: Yale University Press, 1987.

Bennett, Jonathan. *A Study of Spinoza's Ethics*. Indianapolis, IN: Hackett, 1984.

Curley, Edwin. *Behind the Geometrical Method*. Princeton, NJ: Princeton University Press, 1988.

Curley, Edwin. "*Homo Audax*: Leibniz, Oldenburg and the *Theological-Political Treatise*." *Studia Leibnitiana*, supp. 1991.

Curley, Edwin. "Maimonides, Spinoza and the Book of Job." In *Jewish Themes in Spinoza's Philosophy*, edited by Heidi Ravven and Lenn Goodman. Albany: SUNY Press, 2002.

Curley, Edwin. "Notes on a Neglected Masterpiece, I: Spinoza and the Science of Hermeneutics." In *Spinoza: The Enduring Questions*, edited by Graeme Hunter. Toronto: University of Toronto Press, 1994.

Curley, Edwin. "Notes on a Neglected Masterpiece, II: The *Theological-Political Treatise* as a Prolegomenon to the *Ethics*." In *Central Themes in Early Modern Philosophy*, edited by J. A. Cover and Mark Kulstad, Indianapolis, IN: Hackett, 1990.

Curley, Edwin. *Spinoza's Metaphysics*. Cambridge, MA: Harvard University Press, 1969.

Curley, Edwin. "The State of Nature and its Law in Hobbes and Spinoza." *Philosophical Topics* 19 (1991): 97–117.

Curley, Edwin, and Pierre-François Moreau. *Spinoza: Issues and Directions*. Leiden: Brill, 1990.

Curley, Edwin, and Greg Walski. "Spinoza's Necessitarianism Reconsidered." In *New Essays on the Rationalists*, edited by Rocco Gennaro and Charles Huenemann. Oxford: Oxford University Press, 1999.

Damasio, Antonio. *Looking for Spinoza: Joy, Sorrow and the Feeling Brain*. Orlando, FL: Harcourt, 2003.

Della Rocca, Michael. *Representation and the Mind-Body Problem in Spinoza*. Oxford: Oxford University Press, 1996.

Donagan, Alan. *Spinoza*. Chicago: University of Chicago Press, 1988.

Garrett, Don, ed. *The Cambridge Companion to Spinoza*. Cambridge, U.K.: Cambridge University Press, 1996.

Garrett, Don. "*Ethics* IP5: Shared Attributes and the Basis of Spinoza's Monism." In *Central Themes in Early Modern Philosophy*, edited by J. A. Cover and Mark Kulstad, Indianapolis: Hackett, 1990.

Garrett, Don, "Teleological Explanation in Spinoza and Early Modern Rationalism." In *New Essays on the Rationalists,* edited by Rocco Gennaro and Charles Huenemann. Oxford: Oxford University Press, 1999.

Grene, Marjorie, ed. *Spinoza, a Collection of Critical Essays*. Garden City, NY: Anchor, 1973.

Israel, Jonathan. *The Radical Enlightenment, Philosophy and the Making of Modernity, 1650–1750*. Oxford: Oxford University Press, 2001.

Koistinen, Olli, and John Biro, eds. *Spinoza: Metaphysical Themes*. Oxford: Oxford University Press, 2002.

Lloyd, Genevieve. *Part of Nature, Self-knowledge in Spinoza's Ethics*. Ithaca, NY: Cornell University Press, 1994.

Mandelbaum, Maurice, and Eugene Freeman, eds. *Spinoza: Essays in Interpretation*. LaSalle, IL: Open Court, 1975.

Nadler, Steven. *Spinoza, A Life*. Cambridge, U.K.: Cambridge University Press, 1999.

Nadler, Steven. *Spinoza's Heresy*. Oxford: Oxford University Press, 2001.

Smith, Steven, *Spinoza, Liberalism and the Question of Jewish Identity*. New Haven, CT: Yale University Press, 1997.

Wolfson, Harry Austryn. *The Philosophy of Spinoza*. 2 vols. Cambridge, MA: Harvard University Press, 1934.

Yovel, Yirmiyahu, ed. *Desire and Affect: Spinoza as Psychologist*. New York: Little Room Press, 1999.

Yovel, Yirmiyahu, ed. *God and Nature*. Leiden: E. J. Brill, 1991.

Yovel, Yirmiyahu. *Spinoza and Other Heretics.* Vol. 1: *The Marrano of Reason*. Vol. 2: *The Adventures of Immanenece*. Princeton, NJ: Princeton University Press, 1989.

Yovel, Yirmiyahu, and Gideon Segal, eds. *Spinoza*. Aldershot, U.K.: Ashgate, 2002.

Yovel, Yirmiyahu, and Gideon Segal, eds. *Spinoza on Knowledge and the Human Mind*. Leiden: E. J. Brill, 1994.

Yovel, Yirmiyahu, and Gideon Segal, eds. *Spinoza on Reason and the Free Man*. New York: Little Room Press, 2004.

Edwin Curley (2005)

SPINOZISM

The term *Spinozism* has almost invariably been used, by both defenders and detractors, to refer to doctrines held or allegedly held by Benedict de Spinoza. Unlike "Platonism," for example, it has not generally been used to refer to a developing doctrine arising out of Spinoza's philosophy. In the seventeenth and eighteenth centuries the term was frequently used to disparage various types of atheistic doctrines that were held to be attributable to Spinoza. For almost a century after his death, his work was neglected by philosophers, execrated by orthodox theologians of diverse denominations, and slighted even by freethinkers. It is not always possible, however, to distinguish between those genuinely opposed to Spinoza's alleged atheism and those who really espoused atheism while pretending to disparage it.

BAYLE AND THE "PHILOSOPHES"

Spinoza's early reputation rested almost entirely on the long article in Pierre Bayle's *Dictionnaire philosophique* (1697), for some time the only readily accessible account of Spinoza's system. Bayle, like many others, admired Spinoza's life but abhorred his doctrine. In Spinoza he saw an application of his own thesis that atheism may coexist with the highest moral excellence. All agree, he wrote, that Spinoza was a "sociable, affable, friendly, and thoroughly good man. This may be strange, but no stranger than to see a man lead an evil life even though he is fully persuaded of the truth of the Gospel." But Bayle described Spinoza's philosophy as "the most absurd and monstrous hypothesis that can be envisaged, contrary to the most evident notions of our mind." Bayle's antagonism to Spinoza's philosophy arose primarily from his dissatisfaction with monism as a solution to the problem of evil. That such an extreme evil as war could exist among men who are but modes of one and the same infinite, eternal, and self-sufficient substance seemed particularly outrageous to him.

Voltaire, like Bayle, expressed esteem for Spinoza's life but had misgivings about his philosophy, although he did accord a measure of praise to the *Tractatus Theologico-Politicus.* Voltaire's understanding of Spinoza's *Ethics,* however, may be questionable, for he quoted from the inaccurate, popularized version by the Count de Boulainvilliers, published under the title *Réfutation de Spinoza* (Brussels, 1731). According to Voltaire, Spinoza's system was built on complete ignorance of physics and was the most monstrous abuse of metaphysics. In regarding the universe as a single substance Spinoza was, as he put it in his *Le philosophe ignorant* (Geneva, 1766), "the dupe of his geometrical spirit."

Denis Diderot, in the *Encyclopédie,* also closely followed Bayle's article in his criticism of Spinoza's philosophy, yet his own views unmistakably reveal Spinozist elements in denying the existence of a being outside, or separate from, the material universe. "There is," he wrote in *Entretiens entre d'Alembert et Diderot,* "no more than one substance in the universe, in man or in animal." Diderot's monism was not quite the same as Spinoza's metaphysical monism, for it was more pragmatic in nature. His "one substance" was merely material substance, not substance in Spinoza's sense of "that which is in itself, and conceived through itself … (and) of which a conception can be formed independently of any other conception (*Ethics,* Part I, Definition 3). The universe, for Diderot, was monistic in its material unity. Nonetheless, Spinoza's metaphysical monism could be considered as the logical basis for Diderot's materialist monism.

GERMANY

While Voltaire's and the Encyclopedists' interpretation of Spinoza was gaining currency in France, attempts were being made in Germany to reappraise his philosophy. This reexamination was an integral part of the German Enlightenment that, while sharing with its French and English counterparts the affirmation of the individual's right to question established truths, also sought to link this affirmation with religious faith rather than with skeptical disbelief. In the course of this quest Spinoza's image underwent a distinct change. From David Hume's ironically labeled "universally infamous" atheist, Spinoza became Novalis's *gottbetrunkener Mensch.* A number of leading German thinkers came increasingly to see in Spinoza's pantheism a profoundly religious conception and interpretation of the cosmos.

To some extent, the reversal in Spinoza's fortunes was also a corollary of the developments in science. Few of Spinoza's contemporaries who accepted the new scientific theories realized their theological implications. The intellectual reorientation in eighteenth-century Germany, on the other hand, was accompanied by a corresponding change in theological thinking. In the light of these changes Spinoza's philosophy appeared much less inimical to the essential truths of religion.

PANTHEISMUSSTREIT. Probably the strongest factor contributing toward the revival of interest in Spinoza's thought was the controversy that raged over Gotthold Lessing's alleged Spinozism. This dispute, sparked by the

disagreement between Moses Mendelssohn and F. H. Jacobi, came to involve almost every notable figure in the German literary world. Jacobi, in his account of a conversation with Lessing, claimed that the latter had been a Spinozist. According to this account Lessing said that the orthodox conceptions of deity were no longer satisfactory for him and that, if he were to call himself after any master, he knew of no other than Spinoza. Although Jacobi conceded that Spinoza's philosophy was logically unanswerable, he found it unacceptable on religious grounds; in religion, he felt, he had to take refuge in an act of faith, a "salto mortale" as he called it. Lessing sardonically replied that he was unable to trust his old limbs and heavy head for such a leap.

It should not, however, be inferred that Lessing's philosophical outlook was in every detail or even in essentials merely a reflection of Spinozist ideas. Lessing was far too independent a thinker to be subject to any single pervasive influence. He was also far less metaphysically oriented than Spinoza, and his faith in man's perfectibility was tempered by a shrewder realization of man's limitations than that of his world-shunning precursor. Nor must it be assumed that Lessing's exchanges with Jacobi can be taken at their face value. Lessing was fully aware of Jacobi's misconceptions in his approach to Spinoza and hardly took him seriously. He may have been speaking with tongue in cheek, and it would therefore be unwise to attach too great an importance to the views he espoused.

Lessing did succeed in eliciting Jacobi's admission that Spinoza's philosophy was the most rigorous and consistent intellectual enterprise ever attempted and in inducing him to study it more deeply. Although Jacobi's further studies did little to alter his conviction that Spinoza was an atheist and that final truths were to be found in the philosophy of the heart rather than in that of the understanding, they nonetheless helped to focus attention on Spinoza to an unprecedented degree. Two men in particular, Johann Gottfried Herder and Johann Wolfgang von Goethe, who were both on intimate terms with Jacobi, were the most directly affected. Herder openly called himself a Spinozist, although his ontology and cosmology had much more in common with the Earl of Shaftesbury's and Gottfried Wilhelm Leibniz's than with Spinoza's. Yet he insisted that by substituting his concept of *Kraft* for Spinoza's substance he was not fundamentally departing from Spinozist premises. Herder clearly did not realize how very different were his metaphysical presuppositions in postulating an ever-changing *Kraft* in place of Spinoza's unchanging substance and hence how profoundly at variance was his brand of monism with that of his great precursor, despite superficial similarities. Goethe, too, in his autobiography and in his correspondence with Jacobi, acknowledged a far greater debt to Spinoza than he really owed. In Book XIV of his *Dichtung und Wahrheit* he paid his eloquent tribute to Spinoza's influence:

> After I had looked around the whole world in vain for a means of developing my strange nature, I finally hit upon the *Ethics* of this man.… Here I found the serenity to calm my passions; a wide and free view over the material and moral world seemed to open before me. Above all, I was fascinated by the boundless disinterestedness that emanated from him. That wonderful sentence "he who truly loves God must not desire God to love him in return" with all the propositions on which it rests, with all the consequences that spring from it, filled my whole subsequent thought.

Yet Goethe's pantheism had far greater affinity with Herder's—and thus with Shaftesbury's and Leibniz'—than with Spinoza's. Like Herder's confessed Spinozism, Goethe's was much more the result of a poetical imagination and of an emotional craving than of logical analysis and philosophical understanding. Indeed, although G. W. F. Hegel regarded Spinoza's philosophy as philosophy par excellence and although Johann Gottlieb Fichte and Friedrich von Schelling took it as their starting points, the general nature of the Spinozist revival in Germany was literary rather than philosophical.

ENGLAND

Much the same was true of the Spinozist renaissance in England and to a lesser extent in France during the nineteenth century. Admittedly, deism in England had already displayed marked Spinozist characteristics, even if one cannot agree with Leslie Stephen that the "whole essence of the deist position may be found in Spinoza's *Tractatus.*" Few deists were consciously aware of the Spinozist heritage, and it was not until German thought had begun to make itself felt in the English literary world that Spinozism acquired significance as a subject of intellectual discourse.

Samuel Taylor Coleridge was undoubtedly the chief link in this transmission. To judge from Henry Crabb Robinson's account, Coleridge, when receiving from him Spinoza's *Ethics,* kissed Spinoza's face on the title page, said the book was his gospel, but—almost in the same

breath—proclaimed his philosophy false and hence incapable of affecting in the slightest his faith "in all the doctrines of Christianity, even of the Trinity." The ambivalence in Coleridge's attitude toward Spinoza, whom he praised as the "Hercules' pillar of human reason" and simultaneously assailed for his moral and religious views, followed a pattern characteristic of many Spinozists before him, most notably Jacobi. Like Jacobi, Coleridge paid tribute to the rigor of Spinoza's logic and commended his works as "medicinal" reading, while deploring their inadequacy as a philosophical basis of religious belief. Spinoza's *unica substantia,* Coleridge maintained, was not an object at all but a mere notion, a subject, of the mind. Spinoza committed the "most grievous error" of seeing God "in his *Might* alone … and not likewise in his moral, intellectual, existential and personal Godhead." In the *Biographia Literaria* Coleridge related that he had talked much to William Wordsworth about Spinoza, which would help to account for the undeniably Spinozist elements in Wordsworth's poetry. But like Coleridge and other English writers of this period, Wordsworth added nothing new to the conception of Spinozism.

NINETEENTH-CENTURY FRANCE

The reception of Spinoza in nineteenth-century France also witnessed no startling reinterpretations except that, as in Germany, the charge of atheism appeared to many to be quite unfounded. Like Lessing, Herder, and Goethe, Victor Cousin and his followers decisively dismissed the accusations to which Spinoza's *Ethics* had been subjected by orthodox Christians. Nonetheless, Théodore Jouffroy and Émile Saisset, both disciples of Cousin, had serious misgivings about Spinoza's pantheism, for it seemed to absorb the individual in too determinate a manner in the cosmic forces of the whole and thus to threaten the very possibility of human freedom. Paul Janet echoed these misgivings and declared that "the genius of Spinoza was therefore not well adapted to the French mind." But Jouffroy's detailed attention in his lectures at the Sorbonne to Spinoza's thought, and Saisset's publication of a French translation of Spinoza's works, helped to create an intellectual climate in which Spinoza's philosophy could no longer be ignored or lightly dismissed. Thenceforth very many French writers of note, from Edgar Quinet, Alphonse-Marie-Louis de Prat de Lamartine, and Jules Michelet to Georges Sand, Ernest Renan, and the Saint-Simonians felt impelled to grapple with Spinozist ideas.

RUSSIA

The spread and proliferation of interest in Spinozism could not help making its imprint on Russia, a country whose thinkers had for some time been increasingly fascinated by Western philosophical thought. Even more remarkable is the extent to which Russia maintained its preoccupation with Spinoza despite—or perhaps because of—the Bolshevik Revolution. No other pre-Marxian philosopher, with the possible exception of Hegel, has received as much attention in the Soviet Union. From 1917 to 1938, 55,200 copies of Spinoza's works were published in the Soviet Union, compared to 8,000 in the period from 1897 to 1916. Prerevolutionary literature on Spinoza had for the most part been critical and negative, but what non-Marxists considered Spinoza's chief philosophical defects later appeared to many Soviet writers as his strong points. Spinoza's political doctrines particularly appealed to the Marxists. Georgi Plekhanov came to see in Spinozism, when freed from its theological wrappings, a historical forebear of dialectical materialism, and he spoke of Marxism as a "variety of Spinozism." Following Marx and Engels, many Soviet writers credited Spinoza with having correctly solved the fundamental ontological problem concerning the relation of consciousness to being, and of thought to things. Indeed, admiration for Spinoza prompted some to call him "Marx without a beard." Spinoza's rejection of an act of creation, his denial of a continuing intervention in the governance of the world by a supernatural being, his acceptance of nature as something ultimate, self-caused, and "given," without limits of time or space, were all features not lost upon dialectical materialists. No less congenial was the determinism and naturalism of Spinoza's ethical and social philosophy that, while insisting on the possibility of arriving at objective and absolute truth, had analyzed the moral concepts of good and evil in terms of human desire and judgment. Finally, and most important, the allegedly passive role of thought in Spinoza's system, which several prerevolutionary writers had critically commented upon, was regarded in the Soviet Union as the most convincing proof of Spinoza's profound understanding of the historical process. Even if it is conceded that the Marxists revealed as many differences of emphasis in their positive appraisal of Spinoza's thought as did the non-Marxists in their negative approaches, the essentials of Spinoza's doctrines substantially engaged Russian philosophical thinking since the nineteenth century.

Spinozism, then, embodies no single consistent school of thought. Many who professed to admire and accept Spinoza's philosophical premises were as apt to

misunderstand and misinterpret them as those who despised them. Yet despite the diversity of meaning that the term underwent in different intellectual contexts and periods, its catalytic significance cannot be gainsaid.

See also Bayle, Pierre; Coleridge; Samuel Taylor; Lessing, Gotthold Ephraim; Pantheismusstreit; Spinoza, Benedict (Baruch) de.

Bibliography

Bayle, Pierre. *Dictionnaire historique et critique.* 2nd ed. Rotterdam, 1702.

Coleridge, Samuel Taylor. *Biographia Literaria.* London, 1817.

Colie, R. L. "Spinoza and the Early English Deists." *Journal of the History of Ideas* 20 (1959): 23–46.

Diderot, Denis. *Oeuvres complètes,* edited by J. Assezat and M. Tourneux. Vol. II. Paris, 1875–1877.

Dilthey, Wilhelm. "Aus der Zeit der Spinoza-Studien Goethes." *Archiv für Geschichte der Philosophie* 7 (1894): 317–341.

Goethe, Johann Wolfgang von. *Aus meinem Leben: Dichtung und Wahrheit.* 4 vols. Tübingen and Stuttgart, 1811–1833.

Herder, Johann Gottfried von. *Gott, einige Gespräche.* Gotha, 1787. Translated with an introduction by F. H. Burkhardt as *God, Some Conversations.* New York: Veritas Press, 1940.

Jacobi, Friedrich Heinrich. *Ueber die Lehre des Spinoza.* Breslau, 1785.

Jakobi, Max, ed. *Briefwechsel zwischen Goethe und F. H. Jakobi.* Leipzig, 1846.

Janet, Paul. "Spinoza et le Spinozisme." *Revue des deux mondes* 70 (1867): 470–498.

Janet, Paul. "Le Spinozisme en France." *Revue philosophique* 13 (1882): 109–132.

Kline, G. L. *Spinoza in Soviet Philosophy.* London: Routledge and Paul, 1952.

Lévy-Bruhl, Lucien. "Jacobi et le Spinozisme." *Revue philosophique* 27 (1894): 46–72.

Metzger, Lore, ed. "Coleridge's Vindication of Spinoza: An Unpublished Note." *Journal of the History of Ideas* 21 (1960): 279–293.

Morley, Edith J., ed. *Henry Crabb Robinson on Books and Their Writers.* 2 vols. London: Dent, 1938. These volumes contain the references to contemporary English books and their writers in Crabb Robinson's diary, travel journals, and reminiscences.

Rehorn, Karl. *G. E. Lessings Stellung zur Philosophie des Spinoza.* Frankfurt, 1877.

Stephen, Leslie. *History of English Thought in the Eighteenth Century.* London, 1876. Ch. 1, Par. 33.

Suphan, Bernhard. *Goethe und Spinoza, 1783–86.* Berlin, 1882.

Voltaire. *Le philosophe ignorant.* Geneva, 1766. Cited from Voltaire, *Mélanges.* Paris, 1961.

Frederick M. Barnard (1967)

SPIR, AFRIKAN ALEXANDROVICH
(1837–1890)

Afrikan Alexandrovich Spir, the Russian metaphysician was born in Elizavetgrad (present-day Kirovohrad) in the Ukraine, the son of a Russian doctor and a mother of Greek descent. Spir became interested in philosophy when, at the age of sixteen, he read Immanuel Kant's *Critique of Pure Reason,* a work that was to have a profound influence on him. He received no formal education in philosophy, however, and consequently never gained entry into philosophical circles, either in his native country or in Germany, where he settled in 1867. Spir attended a naval cadet school. He received both the Order of St. George and the Order of St. Andrew for his services as a naval officer. Before leaving Russia, he freed all his serfs and gave them land and lodging. He also gave away most of his money and lived on the income from the remainder. In 1869 Spir wrote that only two human activities have real worth—socially useful work and intimate discourse among people who think alike, yet in his lifetime Spir was denied both of these; indeed, few philosophers have been so isolated or ignored.

During the fifteen years Spir lived in Germany he published many articles and several books, including his major philosophical work, *Denken und Wirklichkeit* (Thought and reality; Leipzig, 1873), but notices and reviews were few. Bad health cut him off even further from the world. Hoping for a more receptive audience among French-speaking readers, Spir moved to Switzerland in 1882, but his work remained unknown and his views not understood. He died in Geneva, a Swiss citizen, just as his writing was beginning to attract attention.

Spir's later writings are on the whole restatements and clarifications of the metaphysical views presented in *Denken und Wirklichkeit,* which he felt might have been neglected because of its difficulty. In *Denken und Wirklichkeit* Spir argued that the task of philosophy is to seek absolutely true knowledge. In order to carry out this task, two immediately certain facts must be recognized: consciousness and the supreme law of thought, the principle of identity. This principle is the expression of a norm, of the a priori concept of the unconditioned, that is, of an object that is its own essence and is self-identical. To deny this concept is to deny that it can be conceived and, hence, that it can be denied. The principle of identity is seen to be the one synthetic a priori principle.

To the subjective necessity of this norm is added an objective proof: All our experience disagrees with it and,

therefore, it cannot be a mere generalization from experience. Finally, the principle of identity adds something to experience: All phenomena are organized as if they were self-identical; therefore the principle of identity is the condition of all the regularity of experience.

The unconditioned is, then, the norm, true essence, or God. The unconditioned, however, is not the source or ground of the conditioned: The norm cannot be the source of the abnormal, which contains elements of falsity foreign to the absolute. The relation of the absolute to the phenomenal can best be described analogously, as the relation of an object to its false idea. Having no relation to true being, the phenomenal world simply cannot be explained, its principle can only be thought of as its very abnormality, as its nonself-identity, as becoming. Hence the phenomenal world has no beginning and no end. At the same time, since it is conditioned by becoming, it strives for and evolves to what it is not, the normal. In man, empirical nature has evolved to consciousness, to the awareness of its abnormality. In this awareness man recognizes a norm. Thus he rises above empirical nature and sees the law of his true being as the law not of nature but of the norm, as the laws of morality and logic. Thus morality rises above natural science and, since the moral law is the norm, morality becomes religion.

See also Kant, Immanuel; Metaphysics, History of; Russian Philosophy.

Bibliography

WORKS BY SPIR

Recht und Unrecht. Leipzig: J.G. Findel, 1879. Translated into English by A. F. Falconer as *Right and Wrong.* Edinburgh: University Court of the University of St. Andrews, Oliver and Boyd, 1954.

Gesammelte Werke von A. Spir. 2 vols., 4th ed. Leipzig, 1908–1909.

WORKS ON SPIR

Jodl, Friedrich. "African Spir." *Zeitschrift für Philosophie* 98 (1891).

Lapshin, I. I. "A. Spir, sa vie, sa doctrine." *Bulletin de l'Association russe pour les recherches scientifiques à Prague* 7 (42) (1938).

Lessing, Theodor. *African Spirs Erkenntnislehre.* Erlangen, Germany, 1899. Dissertation.

Zacharoff, Andreas. *Spirs theoretische Philosophie.* Jena, Germany, 1910. Dissertation.

Mary-Barbara Zeldin (1967)

SPIRITO, UGO

(1896–1979)

Ugo Spirito, the Italian idealist philosopher, was born in Arezzo. He began his academic career as assistant to Giovanni Gentile at Rome and first established his reputation as an acute interpreter and trenchant defender of "actual idealism." He was also one of the founders of "corporative" economic studies in fascist Italy and always maintained an active interest in economics and in political and social science.

Spirito held that Gentile's "pure act" was not merely a philosophical concept but was also necessarily a concept of philosophy itself as an activity. This belief led Spirito in 1929 to proclaim the identity of philosophy and science, because all actual knowledge must be the solution of a determinate historical problem and neither philosophy nor science as they occur in actual experience can claim an absolute status independent of the history of their genesis and of the progress of further research. According to Spirito, the actual unity of philosophy and science is what is realized in the process of scientific research; his claim that the "pure act" is the conscious achievement of this unity led to the conception of life as research, set forth in his best-known book *La vita come ricerca*. In this work the absolute philosophical knowledge of traditional metaphysics was presented as the ideal limit toward which scientific research must forever tend but which it can never attain.

In later works, Spirito was led to an ever more strictly negative or critical conception of the task of philosophy because of the difficulty of defining this ideal goal and the paradox involved in discussing it without knowledge of it (which could only come from the secure possession of an eternal standpoint). The philosopher must confine himself to the task of identifying and exposing all claims to absolute knowledge and all forms of antihistorical dogmatism or superhistorical metaphysics wherever they occur. Such claims will otherwise impede the free advance of positive research, which includes all types of inquiry leading to the acquisition of knowledge, whether theoretical or practical. In aesthetics, for example, the philosopher must concentrate on removing prejudices created by definitions and philosophies of art; he must leave to artists, critics, and competent students the construction of the positive science of aesthetics.

This negative conception of the philosopher's task necessarily presupposes a positive philosophy of scientific research itself as a cooperative and progressive solution of problems that organized social groups of researchers

define for themselves. Theoretical problems are solved when science replaces personal opinion. Similarly, practical disagreements will be properly resolved only when scientific planning replaces the selfish initiatives of private individuals. The ideal of social competence must replace the ideal of personal culture in ethics and education, for only through commitment to membership in the community of positive research can an objective criterion of moral and practical values be found without recourse to any metaphysical or religious absolutes. Thus, Spirito inverted the conception of the relation between philosophy and science and between technical competence and general culture, which he found in Benedetto Croce and Gentile. He became one of the leaders of a new Hegelian left in Italy.

Bibliography

WORKS BY SPIRITO

Scienza e filosofia. Florence, 1933.

La vita come ricerca. Florence, 1937.

La vita come arte. Florence, 1941.

Opere complete, 12 vols. (as of 1965). Florence, 1950–.

La vita come amore. Florence: Sansoni, 1953.

Critica della democrazia. Florence: Sansoni, 1963.

Storia antologica dei problemi filosofici. Firenze: Sansoni, 1965–.

Il comunismo. Firenze: Sansoni, 1965.

Dal mito alla scienza. Firenze: Sansoni, 1966.

Machiavelli e Guicciardini. Firenze: Sansoni, 1968.

Giovanni Gentile. Firenze: G.C. Sansoni, 1969.

La vita come amore; Il tramonto della civilt‡ cristiana. Firenze: G.C. Sansoni, 1970.

L'idealismo italiano e i suoi critici. 2nd ed. Roma: Bulzoni, 1974.

La fine del comunismo. Roma: G. Volpe, 1978.

Tamassia, Franco. *L'opera di Ugo Spirito: Bibliografia.* Roma: Fondazione U. Spirito,1986.

Il Pensiero di Ugo Spirito. Roma: Istituto della Enciclopedia italiana, 1988–1990.

Memoirs of the Twentieth Century. Amsterdam; Atlanta, GA: Rodopi, 2000.

H. S. Harris (1967)
Bibliography updated by Michael J. Farmer (2005)

SPRANGER, (FRANZ ERNST) EDUARD

(1882–1963)

Eduard Spranger, the German philosopher and educator, was born in Grosslichterfelde, Berlin. He studied both mathematics and science at a *Realschule* and the humanities at a classical Gymnasium. At the University of Berlin he studied under Wilhelm Dilthey and Friedrich Paulsen and earned his right to lecture with *Wilhelm von Humboldt und die Humanitätsidee* (Berlin, 1909), a classic in the history of German humanism. He was called to the University of Leipzig as professor of philosophy in 1911 and to Berlin as professor of philosophy and pedagogy in 1920. He spent the most creative years of his career and exercised his greatest influence on the *Geisteswissenschaften* and on all levels of German education while at Berlin. In 1933 he submitted his resignation in protest against interference with university freedom by the new National Socialist government but was persuaded by many followers to retain his influential university position. In 1937/1938 he lectured in Japan. He was arrested and imprisoned in 1944 but was released upon the intercession of the Japanese ambassador. Appointed rector of the University of Berlin by the Allied military government in 1945, he found it impossible to accept interference by the East Berlin authorities and in 1946 accepted a professorship in philosophy at Tübingen, where he lectured until his retirement.

Spranger sought to further two projects begun by his teacher, Dilthey. One was an "understanding" (*verstehende*) psychology that would approach human life not with scientific abstractions but perceptively and with an appreciation of cultural values; the other was an attempt to provide a normative interpretation of the *Geisteswissenschaften.* The interdependence of these two problems led Spranger to a Hegelian position (toward which Dilthey himself had begun to turn before his death), and he became a leading figure of the German neo-Hegelian revival of the 1920s.

In his chief work, *Die Lebensformen* (Halle, 1914; translated by J. W. Pigors as *Types of Men,* Halle, 1928), Spranger undertook a typological analysis of personality through the use of the method of *Verstehen.* He held this method to be empirical in that it results in "an at least minimally categorialized after-experience." It is essentially an aesthetic perception of cultural forms in individual life and is motivated by a Platonic *eros*—a love for the personal values involved; this, Spranger insisted, does not interfere with its objectivity. Six forms of value—all of which are objectively rooted in the historical and cultural order, and each of which may dominate a person's life and evoke a reordering of the others in subordination to itself—determine six types of personality in modern culture—the theoretical, economic, aesthetic, social, political, and religious—which center, respectively, in the values of truth, utility, beauty, love; power, and, in religion, in the devotion to a vital totality of value. The moral

is not a distinct type of value but enters into all valuations. Spranger schematized these types into an ideal order without denying individual freedom in value selection.

Spranger's *Psychologie des Jugendalters* (Leipzig, 1924; 8th ed., 1926) applied his method and conclusions to the problems of youth. Four important attainments mark the sound growth of the adolescent: the discovery of self, the development of a life plan, the ordering of the self into the different spheres of human relations, and the awakening of the sexual life and *eros.* The six personality types developed in the *Lebensformen* can serve as a schema for comprehending the individual person in exploring these critical developments.

Spranger's analysis of the *Geisteswissenschaften* found application in his discussions of the ethical bases of modern culture and education. It combined criticism of the historical philosophies of society and culture with the development of a modified Hegelian theory of objective spirit. Subjective and objective spirit are in close interaction within every historically relative situation. To them Spranger added a third dimension of spirit, the normative. This, the relativized absolute spirit of G. W. F. Hegel, comprised the factors that serve a regulative role in history through art, religion, and philosophy. Responsibility for the actualization of the normative, however, lies in the individual; no cultural content becomes meaningful except "insofar as it is again and again created out of the attitude and the conscience of the individual soul."

After World War II Spranger turned to religious themes, particularly in *Die Magie der Seele* (Tübingen, 1947). This "magic of the soul," which is essential to the life of a culture, is constituted by the religious consciousness and serves not to meet immediate external goals but to augment the powers of the person himself. Faith is a "withdrawal into inwardness."

Spranger's work in the philosophy of education kept the classical humanistic ideal alive and exercised a liberating effect on all levels and dimensions of education. It found notable expression in classic studies of great figures in education—Wilhelm von Humboldt, Jean-Jacques Rousseau, Friedrich Froebel, Johann Heinrich Pestalozzi, and Johann Wolfgang von Goethe. Spranger was also involved in most of the ethical and cultural problems of German life, addressing himself to such challenges as labor education, vocational education, personal and vocational guidance, and juvenile delinquency. The eloquence of Spranger's lectures and writings, his personal warmth, felt by a wide circle of friends of all ages, and his combination of keen perception with deep moral concern made him one of the most admired and influential of German thinkers. His deep sense of the German tragedy, and his long preoccupation with its moral and historical causes and the moral cost of redemption, won for him, before he died, the most distinguished honors that his country could bestow.

See also Dilthey, Wilhelm; Froebel, Friedrich; Geisteswissenschaften; Goethe, Johann Wolfgang von; Hegel, Georg Wilhelm Friedrich; Hegelianism; Humanism; Humboldt, Wilhelm von; Paulsen, Friedrich; Pestalozzi, Johann Heinrich; Rousseau, Jean-Jacques.

Bibliography

Among Spranger's other works are *Wilhelm von Humboldt und die Reform des Bildungswesen* (Berlin, 1910); *Goethes Weltanschauung. Reden und Aufsätze* (Leipzig and Wiesbaden, 1942); *Gibt es ein Kulturpathologie?* (Tübingen, 1947); and *Aus Friedrich Froebels Gedankenwelt* (Heidelberg, 1951).

For a complete bibliography of Spranger's works, see Theodore Neu, *Bibliographie Eduard Spranger* (Tübingen: Niemeyer, 1958).

For discussions of Spranger's thought and influence, see *Erziehung zur Menschlichkeit–Festschrift für Eduard Spranger zum 75. Geburtstag,* edited by H. W. Bähr (Tübingen: Niemeyer, 1957); and *Eduard Spranger, Bildnis eines geistigen Menschen unserer Zeit,* edited by Hans Wenke (Heidelberg: Quelle and Meyer, 1957).

L. E. Loemker (1967)

STACE, WALTER TERENCE

(1886–1967)

Walter Terence Stace, the Anglo American empiricist philosopher, was born in London. He was graduated from Trinity College, Dublin, in 1908 and from 1910 to 1932 served in the civil service in Ceylon. During this period he published *A Critical History of Greek Philosophy* (London, 1920) and *The Philosophy of Hegel* (London, 1924). In 1932 he retired from the civil service to teach philosophy at Princeton University, where he remained until his academic retirement in 1955. He was president of the American Philosophical Association in 1949.

Stace's *The Theory of Knowledge and Existence* (Oxford, 1932) is the definitive statement of his general position on philosophical method. His argument rests on the claim that on strict empirical grounds the solipsist position is logically unassailable. Whereas philosophers such as George Santayana, starting with the same claim, appealed to a doctrine of "animal faith" and emphasized

the irrational element in belief in an external world, Stace carefully and in detail offered an analysis of the steps whereby we construct our conception of an external physical world out of the available data. He often spoke of his doctrine as a theory of *fictions*, but in print he preferred the word *constructions*. The point is that the construction of the fiction of an external world is neither irrational nor animal. It is a step-by-step inference that, although it fails to provide a logical answer to solipsism's claims, does satisfy human demands for reasons for belief. Ultimately our reasons for belief rest, according to Stace, upon two general claims that can be empirically supported—the claims that human minds are similar and that they labor together in common. These two empirical facts, and not logical proofs, support our commonsense beliefs. This thesis lies at the heart of most of Stace's later work.

Stace in this earlier period was an advocate of the sense-datum theory. In spite of continued association of his name with G. W. F. Hegel, he was chiefly indebted to David Hume, G. E. Moore, and Bertrand Russell. His main object of attack was Russell's *Our Knowledge of the External World*, which, according to Stace, constantly violates the principle of empiricism. In 1934 he published one of his best-known articles, "The Refutation of Realism" (*Mind* 43 [1934]: 145–155), in response to Moore's influential "The Refutation of Idealism." Moore's argument was based upon a distinction between sense data and our awareness of them. Stace replied that one can grant the distinction and still deny any force to the claim that sense data exist when not being perceived. He generalized the claim that there can be no good reason for believing any version of the proposition that entities exist unperceived. They may so exist, but it is absurd to claim that this can be empirically proved. It follows that where "such proof is impossible, the belief ought not to be entertained."

This argument seems, on the face of it, to contradict the thesis of *The Theory of Knowledge and Existence*. Stace always subsequently maintained, however, that his article had been misunderstood because it was not recognized as irony. He also insisted that Moore's article had been intended as humorous. The irony of his own consisted in showing that the simplest natural belief cannot be supported by strict logical proofs.

Stace's next major work was *The Concept of Morals* (New York, 1937). In one sense the main argument of the book might be, and has been, characterized as a version of subjectivism because it associates a general theory of the meaning of moral judgments with a general theory of man's wants and approvals. Perhaps the most permanently valuable aspect of the argument, however, is the attempt to disassociate the view he is defending from the label "subjectivist." Stace held that the proper contrast between subjectivism and objectivism is between views which make reasoned adjudication of ethical disputes impossible, and views which provide rational grounds for holding that one moral claim can be correct and its rivals mistaken. According to Stace, what makes his view objectivist in this significant sense is the connection between it and a general theory of man's nature, including his desires, wants, and approvals. The result is a modified version of utilitarianism based upon the same two principles emphasized in the theory of knowledge, the similarity of men's minds and the fact that they labor together in common.

In two articles ("Positivism," *Mind* 53 [1944]: 215–237; and "Some Misinterpretations of Empiricism," *Mind* 67 [1958]: 465–484) Stace distinguished empiricism from recent positivistic tendencies. The intention of both is to attack the attempt on the part of more recent logical empiricists, who, Stace claimed, associate empiricism with the demand for strict logical proofs.

In September 1948 Stace published in the *Atlantic Monthly* (pp. 53–58) an article titled "Man against Darkness." The thesis of the article, which Stace considered neither very original nor very shocking, was that the worldview endorsed by the physical sciences since the time of Galileo Galilei is incompatible with Christianity's traditional worldview. The violent reaction to this article stunned him. There followed *The Gate of Silence* (Boston, 1952), a book-length poem; *Philosophy and the Modern Mind* (New York, 1952), a careful historical study of the thesis that had been popularly stated in "Man against Darkness"; and *Time and Eternity* (Princeton, NJ, 1952), an essay in the philosophy of religion which many consider his most profound work.

No doubt partially because of the years he had lived in Ceylon, Stace was attracted to Hinayana Buddhism, and both *The Gate of Silence* and *Time and Eternity* reveal the extent of that influence on his later metaphysical thought. The theme of paradox runs throughout these works: "Men have always found that, in their search for the Ultimate, contradiction and paradox lie all around them.… Either God is a Mystery or He is nothing at all" (*Time and Eternity*, p. 8).

Thus, Stace now held that belief must transcend the confines of strict logic, and the rigorous empiricist ended by courting mysticism. Fully aware of this fact, Stace set himself to what he conceived to be his final philosophical

task—the reconciliation of empiricism and mysticism. The result was *Mysticism and Philosophy* (New York, 1960). He claimed (1) that the mystical experience is a fact, is unique, and is the same in all cultures; (2) that the interpretations of the mystical experience vary widely from culture to culture; and (3) that a genuine empiricism cannot ignore the mystical experience simply because it is logically paradoxical.

Throughout the somewhat otherworldly philosophical reflection of his later life, Stace retained an interest in practical problems. His *The Destiny of Western Man* (New York, 1942) was an expression of horror against the irrational totalitarianism that swept Europe in the 1930s. In February 1947 he published an article in the *Atlantic Monthly*, vigorously attacking the legal basis of Zionist arguments. In early 1960s he was concerned with the universal condemnation of colonialism, insisting that high generalizations be checked against the evidence. In a letter to the *New York Times* (February 4, 1964), he wrote that colonialism "civilized half the world at the cost of the loss of some *amour propre*, of some snobbishness, of some arrogance, of some hard feeling, but—in the case of the Romans and British, at any rate—of very little real cruelty, injustice or tyranny."

See also Buddhism; Empiricism; Mysticism, Nature and Assessment of; Solipsism.

Bibliography

ADDITIONAL WORKS BY STACE

"The Metaphysics of Meaning." *Mind* 44 (1935): 417–438.

A Critical History of Greek Philosophy. London, Macmillan; New York, St. Martin's Press, 1965.

Man against Darkness, and Other Essays. Pittsburgh: University of Pittsburgh Press, 1967.

The Nature of the World: An Essay in Phenomenalist Metaphysics. New York, Greenwood Press, 1969.

Time and Eternity: An Essay in the Philosophy of Religion. New York, Greenwood Press, 1969.

The Destiny of Western Man. Westport, CT, Greenwood Press, 1970.

The Theory of Knowledge and Existence. Westport, CT, Greenwood Press, 1970.

Mysticism and Philosophy. New York: Macmillan, 1972.

The Concept of Morals. Gloucester, Mass., P. Smith, 1975.

Religion and the Modern Mind. Westport, CT: Greenwood Press, 1980.

Works on Stace are A. J. Ayer, "The Principle of Verifiability," in *Mind* 45 (1936): 199–203, a reply to Stace's "The Metaphysics of Meaning"; C. J. Ducasse, a review of *Mysticism and Philosophy*, in *Journal of Philosophy* 59 (1962): 323–335; H. H. Price, "Mr. Stace on the Construction of the External World," in *Mind* 42 (1933): 273–298; and Ralph E. Stedman and H. B. Acton, "Mr. Stace's Refutation of Realism," in *Mind* 43 (1934): 349–353. On *The Theory of Knowledge and Existence*, see the critical notice by F. C. S. Schiller in *Mind* 42 (1933): 94–100, and the review by L. Susan Stebbing in *Philosophy* 8 (1933): 354–357.

James Ward Smith (1967)
Bibliography updated by Michael J. Farmer (2005)

STAËL-HOLSTEIN, ANNE LOUISE GERMAINE NECKER, BARONNE DE
(1766–1817)

Anne Louise Germaine Necker Baronne de Staël-Holstein, the French novelist and essayist, was born in Paris, the daughter of Suzanne Curchot and Jacques Necker, finance minister to Louis XVI. In 1786 she married Eric Magnus, baron of Stäel-Holstein, the Swedish ambassador to France, from whom she separated in 1797. In the year of her marriage she published her first novel, *Sophie,* and four years later a tragedy, *Jeanne Grey.*

Her interest in philosophy began with a study of Jean-Jacques Rousseau, whose fervent admirer she remained throughout her life. She incurred the hostility of Napoleon Bonaparte both by her frank criticism and by her liberalism, and her advocacy of a constitutional monarchy led to her being exiled in 1802. She made her first trip to Germany at this time, a trip that was the occasion of her book *De l'Allemagne.* This work was sent to the printer in 1810, but it was condemned by the censor and did not appear until 1813. After years of traveling, Mme. de Staël returned to Paris, where she remained until her death.

The philosophical ideas of Mme. de Staël are to be found mainly in two books, *De la littérature considérée dans ses rapports avec les institutions sociales* (1800) and *De l'Allemagne.* In the former she attempted to show the influence of religion, morals, and laws on literature and that of literature upon religion, morals, and laws. This book presupposed the perfectibility of man, as Mme. de Staël admitted, but human progress was not automatic; to come into being it required the constant and deliberate aid of education (*les lumières*), which could be provided only through literature. A second premise was that of national characters, the Greek being given to art, emulation, and amusement; the Roman, to dignity, gravity of speech, and rational deliberation. Later she contrasted the Northerner and the Southerner, in *De l'Allemagne* exemplified respectively by the German and the Frenchman.

Nevertheless, there is nowhere in Mme. de Staël's writings the notion of national souls or collective spirits (*Geister*). People to her were individuals, and whatever community of interests and talents they showed was to be attributed to the influence of other individuals.

Mme. de Staël never questioned the absolute value of personal liberty. This belief she attributed to Protestantism, her family religion. To her, Protestantism rested on the principle of personal interpretation, and the source of one's convictions was to be looked for in the heart, just as it was in the teachings of Rousseau's Savoyard vicar. She held that individual differences in temperament were irreconcilable, and believed that only statistics could help a statesman solve his people's ethical problems. It may have been this firmly rooted idea that made her fear the natural scientist as the tool of despots. The scientist, who rejects everything that cannot be reduced to mathematics, is always willing to pursue his own ends, regardless of the vital interests of his fellow men.

The chief contribution of *De l'Allemagne* to philosophy was that it acquainted Mme. de Staël's countrymen with the works of Immanuel Kant, Johann Gottlieb Fichte, Friedrich von Schelling, and Friedrich Schlegel. She presented their ideas simply and sketchily but on the whole correctly. In this way she helped break the hold that the sensationalism of the school of Étienne Bonnot de Condillac had had upon the French. Mme. de Staël wrote no book that can be considered as technical philosophy, but she represents the mind that has absorbed a philosophy as a technique of thinking and as a corrective to authoritarianism.

Bibliography

WORKS BY MME. DE STAËL

Lettres sur les ouvrages et le charactère de J.-J. Rousseau. Paris, 1788.

De l'influence des passions sur le bonheur. 2 vols. Lausanne, 1796. A defense of reason as a critical agent.

De la littérature considérée dans ses rapports avec les institutions sociales. Paris, 1800.

De l'Allemagne. 3 vols. London, 1813.

Oeuvres complètes, 17 vols. Paris, 1820–1821.

WORKS ON MME. DE STAËL

Blennerhasset, Charlotte. *Frau von Staël, ihre Freunde und ihre Bedeutung in Politik und Literatur.* 3 vols. Berlin: Gebrüder Paetel, 1887–1889. Translated by J. E. Gordon Cumming as *Mme. de Staël, Her Friends and Her Influence on Politics and Literature.* London: Chapman and Hall, 1889.

Gautier, Paul. *Mme. de Staël et Napoléon.* Paris: Plon-Nourrit, 1903.

Herold, J. Christopher. *Mme. de Staël, Mistress to an Age.* Indianapolis, 1958. A popular but well-documented biography.

Ollion, E. *Les idées philosophiques morales et pedagogiques de Mme. de Staël.* Mâcon, 1910. Thesis.

George Boas (1967)

STAHL, GEORG ERNST
(1660–1734)

Georg Ernst Stahl was a leading German medical scientist and chemist of his day. Stahl was appointed professor of medicine at the University of Halle in 1694, and from 1716 until his death he served as personal physician to Frederick William I of Prussia. His numerous medical writings had a strongly doctrinal tendency, which made them the source of lively, often bitter, controversy. His famous phlogiston theory, an erroneous explanation of the nature of combustion and calcination, was nonetheless, before Antoine Lavoisier's discoveries, instrumental in placing chemistry on a scientific basis. The same may be said of his studies concerning the properties and composition of acids, alkalis, and salts.

Led by his medical, rather than chemical, interests to philosophy, Stahl elaborated (particularly in his *Theoria Medica Vera*, 1707) a rigorous position of animism, affirming that the animal organism was formed, governed, and preserved by an immaterial principle, or soul. If Stahlian thought was indebted to the *archei* of J. B. van Helmont's occultist biology, and more broadly to both neo-Aristotelian and Neoplatonic versions of animism in the late Renaissance, his notion of soul, reflecting the impact of post-Cartesian dualism, was typical of his own period. He conceived of it as essentially a rational and spiritual substance distinct from matter, but simultaneously he assigned to it the ability to control the organism by an "unconscious" mode of activity. Thus, the soul not only thinks and wills but, having constructed its body, also excites, regulates, and sustains all involuntary and vital processes. It does so by the intermediary of movement, which Stahl regarded as an immaterial entity, for matter itself is held to be essentially passive and inert. The soul, by a specific energy, is supposed to communicate the "spiritual act" of movement to the organism in pursuance of its own aims.

This rather obscure view of things (which Gottfried Wilhelm Leibniz, among others, criticized) was not improved by Stahl's manner of expression, a mixture of dogmatic haughtiness and repetitious turgidity. If he

failed, moreover, to consider properly the various contradictions and difficulties peculiar to his position, this was due largely to his lack of interest in metaphysics as such. His animism was intended less as a philosophical contribution than as a theoretical standpoint from which to perceive and evaluate the phenomena of disease and health in accordance with an expectative approach to therapeutics. Even more significantly, it represented a protest against the dominant iatromechanist and iatrochemical schools, which at the time tended to see animate beings too naively and rigidly in terms of facile mechanical analogies and unexplained chemical reactions. But although Stahl's animism had the merit of emphasizing the presence of an irreducible "life force" having no equivalent in the machine, the omnipresent role allowed to this life force at the expense of a purely organic dynamism proved untenable.

The influence of Stahlianism was checked during the first half of the eighteenth century by the success of the mechanistic and empirical doctrines of Hermann Boerhaave and Friedrich Hoffmann. Subsequently, Stahl's medical philosophy was reinterpreted at the important Faculty of Montpellier, with the general result that its spiritualist aspect was abandoned as unscientific while its insistence on a metamechanical "vital principle" in the organism was adopted as profoundly valid. Stahl thereby came to be recognized as the founder of the vitalistic school of modern biology.

See also Cartesianism; Lavoisier, Antoine; Leibniz, Gottfried Wilhelm; Macrocosm and Microcosm; Panpsychism; Philosophy of Biology; Renaissance; Vitalism.

Bibliography

WORKS BY STAHL

Oeuvres médico-philosophiques et pratiques. 5 vols. Paris, 1859–1863.

WORKS ON STAHL

Duchesneau, Francois. "Stahl, Leibniz, and the Territories of Soul and Body." In *Psyche and Soma: Physicians and Metaphysicians on the Mind-Body Problem from Antiquity to Enlightenment*, edited by John Wright. Oxford: Clarendon Oxford Press, 2000.

Gottlieb, B. J. *Bedeutung und Auswirkungen Stahls auf den Vitalismus*. Halle, 1943.

Lemoine, Albert. *Le vitalisme et l'animisme de Stahl.* Paris: Baillière, 1864.

Metzger, Hélène. *Newton, Stahl, Boerhaave et la doctrine chimique*. Paris: Alcan, 1930.

Aram Vartanian (1967)
Bibliography updated by Tamra Frei (2005)

STAMMLER, RUDOLF

(1856–1938)

Rudolf Stammler was a German neo-Kantian legal philosopher. His first major work, *Die Lehre vom richtigen Recht,* outlined his philosophy of law, which was elaborated in subsequent works. Stammler sought to apply Immanuel Kant's distinction between pure and practical reason to the law. The embodiment of pure reason in legal theory is the concept of law, which Stammler defined as "combining sovereign and inviolable volition." The counterpart of practical reason is the idea of law, that is, the realm of purposes realized by volition. But whereas for Kant practical reason was not, like pure reason, a matter of intellectual perception, but of morality, Stammler sought to formulate a theoretically valid idea of justice. He based it on the community of purposes and the fact that man is a reasonable being, an end in himself. From this he derived two "principles of respect" and two "maxims of participation." The former are that no one's volition must be subject to the arbitrary desire of another and that any legal demand must be of such a nature that the addressee could be his own neighbor. The latter are that no member of a legal community must be arbitrarily excluded from the community and that a legal power may be exclusive only insofar as the excluded person can still be his own neighbor.

For Stammler these were not merely formal principles; they could be used to solve actual legal problems. He attempted, for example, to apply them to the legality of cartels and to the solution of disputes between upper and lower riparian owners over the use of water. His solutions were generally those of a moderate liberal.

Max Weber has shown in "Rudolf Stammlers Überwindung des materialistischen Geschichtsauffassung" (*Gesammelte Aufsätze zur Wissenschaftslehre,* Tübingen, 1922, pp. 291–359) that Stammler's alleged formal categories are in fact categories of progressive generalizations, the more general being relatively more formal than the less general. Stammler's main error was his attempt to make the idea of justice a matter of theoretical knowledge; it was therefore inevitable that he should confuse principles generally acceptable to a moderate liberal with universally valid principles of justice. His idea of justice is therefore a cross between a formal proposition and a definite social ideal, kept abstract and rather vague by the desire to remain formal. Stammler's chief merit remains his reintroduction of legal philosophy as a vital aspect of the study of law.

See also Continental Philosophy; Justice; Kant, Immanuel; Neo-Kantianism; Philosophy of Law, History of; Weber, Max.

Bibliography

WORKS BY STAMMLER

Wirtschaft und Recht. Halle, 1896. Law and economics are related as form and matter.

Die Lehre vom richtigen Recht. Halle, 1902. Translated by I. Husik as *The Theory of Justice.* New York: Macmillan, 1925.

Theorie der Rechtswissenschaft. Halle: Buchhandlung des Waisenhauses, 1911.

Lehrbuch der Rechtsphilosophie. Berlin and Leipzig: de Gruyter, 1922.

WORKS ABOUT STAMMLER

Ginsberg, Morris. "Stammler's Philosophy of Law." In *Reason and Unreason in Society: Essays in Sociology and Social Philosophy.* Cambridge, MA: Harvard University Press, 1948.

Stone, Julius. *Human Law and Human Justice.* London: Stevens, 1965.

Wolfgang Friedmann (1967)
Bibliography updated by Philip Reed (2005)

STATE

Before the sixteenth century the word *state* was used to refer to the *estates* of the realm or to kingly office or dignity, but not to an independent political community. Niccolò Machiavelli was largely responsible for establishing this modern usage. The change, however, was not in words only but also in ways of thinking about political organization and political relations. In feudal society a man figured in a network of quasi-contractual relations in which his political rights and duties were closely linked to land tenure and fealty. He was his lord's man and his king's man. The powers of kingship were only with difficulty distinguished from property rights. From the twelfth century on, the conceptions of Roman law began once more to influence political thought. Public authority was more sharply distinguished from private rights; the peculiar position of the king among his barons, which feudal writers recognized but found difficult to conceptualize, came to be expressed in Roman terms—the *princeps* was said to speak on behalf of the whole people and to exercise *imperium,* as distinct from a feudal privilege, because his care was for the whole *respublica.*

However, so long as barons could still simultaneously hold fiefs from different kings in different lands, the notion could not develop of the territorially defined state, making an exclusive claim to the allegiance of all who resided within its borders. The idea that men could be not only subjects of their king but also citizens of their state became possible with the consolidation of national monarchies in England, France, and Spain. Its development was assisted in the thirteenth century by the quickening of interest in Aristotle's ideas about the city-state and, in the early sixteenth century, by the Renaissance interest in the ancient Roman republic. Classical elements, then, were grafted onto the late medieval stock to produce the Renaissance state.

With the declining influence of such customary forms of regulation as feudal and manorial ties, the guild, and the family, the state became an indispensable category for any kind of speculative thought about society. Moreover, as the grip of custom slackened, men came to think that law might be made by an authoritative will rather than discovered by the understanding or known by tradition. The political order, as the authority structure through which law was created and which therefore conferred legal status and rights on all other forms of association, gained a corresponding preeminence. Out of the split in the universal church and the consequent alliance for mutual survival between protestant princes and religious reformers, there emerged the idea of a national church closely related to the state, further stressing that the state was a community or polity and not simply an aggregation of men who happened to owe allegiance to a common overlord. The consolidation of national states created a new state of nature—a world peopled by sovereign states recognizing no overriding authority and only tenuously subject, if at all, to a common law. Francisco Suárez, Francisco de Vitoria, Hugo Grotius, and Samuel von Pufendorf, the pioneers of international law, explored the relations between states in such a world; what was implied for the internal structure of a state was worked out by Jean Bodin and Thomas Hobbes.

IDENTITY OF THE STATE

Since the seventeenth century, political philosophers have been largely preoccupied with the relations of the state and the individual, with the citizen's rights, if any, against the state, with the right of the state to punish, to promote morality, or to regulate the affairs of other associations such as families, trade unions, and churches. These matters have been all the more troublesome because there is disagreement about the proper analysis of propositions about the state. For instance, what does it mean to say that a state has acted in a certain way, made a decision, adopted a policy, assumed responsibility, and so on? These are not statements about every one of its citizens,

nor are they simply statements about the acts of certain individuals who govern the state; for not all the actions of the person who for the time being is president are acts of the United States, nor is an act of the state always attributable to one person in particular. Hobbes was certainly mistaken when he argued that what made an aggregate of many men into one corporate person was that one man acted for the rest: "The *unity* of the representer, not the *unity* of the represented … maketh the person *one*" (*Leviathan* I, 16).

Again, what kind of sustained identity has the state, that one can speak of its enduring through many generations of natural lives? It is tempting to meet such a question with an organic analogy: Although the cells die and are replaced, the organism survives; although an action of an organism requires nothing more than the coordinated operations of its organs, it is not identical with the actions of any one or of all of them (unless their functions as elements in an organism are presupposed in the descriptions of their actions). The organism, it is often said, is a form of life transcending its parts; purposes are attributed to it that are not the purposes of any one of its parts or of all of them taken severally. Many writers, notably the Hegelians, have described the state in this way, exalting the interests of the state at the expense of the interests of its members considered as individuals.

A quite different account of the state has been given by writers who have employed atomic or contractual models, with explanatory analogies drawn from joint-stock corporations, clubs, or perhaps from mechanical contrivances. Thus, Hobbes talks of the state as an *artificial* man, contrived by an agreement of self-determining individuals. It can have no purposes not ultimately reducible to the purposes of individuals; its acts are those of a sovereign authorized to act on their behalf. The contractual analogy in Hobbes and John Locke is a device for explaining how and under what conditions the acts of one or a few ruling individuals could be attributed to a body composed of a multitude of free and autonomous persons, all with their own separate interests, yet each committed by his own consent to a public interest in which he has a personal stake.

The problem of meaning, however, must be distinguished from the moral problem of obligation. The notion of corporate action does not necessarily entail consent or authorization on the part of individual members, although it could be argued that without consent the individual could have no moral commitment or responsibility. Acts of the state are acts of persons in an official capacity, acting according to procedures and within the competence prescribed by the rules of its constitution. A president's actions are those of the United States only when they form part of a particular procedural routine; they then indicate appropriate responses by other officials. When the president acts in nonofficial roles—as father or as member of his golf club—his actions are incidents in what a Wittgensteinian would call different "games" and therefore have appropriately different implications. The enduring identity of a state can be correspondingly analyzed in terms of the endurance of its procedural order. The Constitution of the United States has had an unbroken history since its adoption in 1788; the changes it has suffered have all been valid according to the criteria it prescribes for itself.

This sort of analysis explains the personality and life of a state without resorting to organic analogies or to metaphysical notions of an order of being where a whole is greater than the sum of its parts. However, it does not deal with all the problems. Despite several revolutions since 1789, there is a sense in which the French state has a continuous history, unlike the Austro-Hungarian state that was destroyed after World War I and replaced by a number of successor states. If the population of an area continues to be governed undivided, as an independent political unit, there seem to be grounds for saying that it remains the same state, despite changes in regime. In the case of France, although formal continuity of legitimization broke down between, for instance, the Second Empire and the Third Republic, there is a continuity of tradition and, despite deep cleavages, a sense that however bitterly rival groups contend, they are nevertheless committed by their awareness of history and common culture to remaining in political association. A struggle to control or reconstruct the machinery of government is not necessarily, then, an attempt to break up the political association, as it was in the Austro-Hungarian Empire.

THE STATE AS AN ASSOCIATION

To call the state an association is to put it on the same footing as clubs, churches, and trade unions. There are features of the state, however, which, although no one of them is peculiar to the state alone, together make it a rather special case. For instance, because people do not usually become or remain members of a state by choice, and because a state exercises exclusive authority over everyone in a given territory, the concept of membership is hazier than in the case of voluntary associations. The state insists that not only its citizens but also everyone else in its territorial jurisdiction shall conform to its rules. Indeed, the notion of a citizen suggests a certain mini-

mum degree of active participation. This may be restricted, as it was in Athens, to a relatively small number of the resident native population. In that case, would the association include only the citizens? Are the rest outsiders on whom the state imposes its will, much as a trade union might insist that nonunionists shall not work for lower wages than its members? Or are citizens and noncitizens merely two classes of members, one with rights of participation, such as the right to vote, the others with private rights only?

Unlike trade unions, literary societies, joint-stock corporations, and guilds, the state's range of interests is very wide and, in principle, unlimited. This, too, is connected with its nonvoluntary character. Even allowing for migration and naturalization, people do not easily join or leave a state, and when they do, it is usually only with its permission. And whether they join it or not, they are subject to it if they reside in its territory. Consequently, the state does not need to define the terms and aims of their membership. Neither is there any higher authority which can rule, as the state's judicial authorities may do in relation to other associations, that a proposed act falls outside its terms of association and therefore infringes its members' rights. This indeterminancy of scope is a characteristic that the state shares with the family and even with some churches. Such associations have no defined set of aims: The behavior norms they sustain may govern a very wide, if fluctuating, segment of the social life of their members. And since the mid-1800s the effective sphere of the state has encroached increasingly on the spheres of other associations.

THE STATE AND CONFLICTS OF INTEREST

The state's territorial inclusiveness and the uncertain limits to its concern have led many political philosophers to assign to it a unique role among the forms of human association. Plato's *Republic* sketched an ideal state in which men's conflicting interests and energies were harnessed and reconciled by philosopher-rulers who would integrate them into a single-minded unity, the principles of which could be discovered by a philosophical insight. Aristotle claimed that, at its best, the Greek *polis* was the most perfect association because, while including lesser associations like the family and the village, it was large enough to provide within itself everything necessary for the good life. For Aristotle, citizenship was a matter not of passively enjoying rights but of participating energetically in the many-sided life of the *polis.* The Greek writers had in mind a small state, a face-to-face community capable of satisfying emotional needs that the impersonal mass state of the twenty-first century cannot. Nevertheless, the same completeness that Aristotle found in the *polis* has often been attributed to the modern state.

Jean-Jacques Rousseau, though tempted to identify the modern state with the *polis,* hesitated to do so unconditionally. He believed that the state was sufficient for the expression of all human excellencies. The vocation of the citizen was the highest to which a man could aspire. Participating in the expression of the general will for the common good of the whole association, the citizen rose above private interest and became a moral person, "substituting justice for instinct in his conduct.… man, who so far had considered only himself, finds that he is forced … to consult his reason before listening to his inclination" (*Social Contract* I, 8). Membership of the state was for Rousseau, as for Plato and Aristotle, a moral education; bad laws corrupted nature, good laws provided conditions for moral development and nobility of soul. Not only was nothing needed beyond the state but also, Rousseau suspected, lesser associations, by setting up partial or sectional interests as objects of loyalty, frustrated the public interest and corrupted the state. Nevertheless, the ideal state of Rousseau's *Social Contract* remained a city-state, small enough for everyone to know everyone else. The attempt by others to extend the conception to the nation-state led to confusion in theory and, in practice, to Jacobin totalitarianism.

G. W. F. Hegel transformed Rousseau's doctrine by substituting for personal, face-to-face relations a metaphysical dependence of parts on the whole. The state was the concrete universal, the individual a mere partial expression of it. Sectional associations had a function in organizing human interests. They operated, however, on a lower plane of reality than the state, a plane that Hegel termed "civil society." This was not a different order from the state but the same social organization viewed from the standpoint of the subjective ends that individuals set themselves. It was the plane of the free market economy motivated by the pursuit of profit and sectional advantage, where competitive conflicts are checked, ordered, and adjusted by the police. Nevertheless, unknowingly and despite themselves, individuals promoted ideal ends. Interests that from the subjective point of view of civil society were sectional and egoistic appeared objectively in the state as moments or partial expressions or functions of the greater whole. The state would then rightly regulate although not supplant such interests. For Plato and Rousseau the conflict of interests was a pathological symptom in a state; for Hegel it was an unreality masking

a fundamental unity that the state would safeguard if necessary. For all three there was a transcendent public interest in which the apparent interests of individuals are dissolved and fused.

There is, however, another view that takes the conflict of interests as a fundamental fact of nature; it can be controlled but never finally superseded. Machiavelli, Hobbes, and Jeremy Bentham were in this tradition. The state existed to regulate competition, since without it individual objectives would be mutually frustrating. The harmony it achieved, however, was artificial; the state remedied a desperate situation by altering the conditions under which men sought their own interests, deflecting them from antisocial ends by fear of punishment. Karl Marx and Friedrich Engels, agreeing that the state suppressed conflict, saw it as a strictly coercive instrument maintained by the dominant economic class to safeguard its privileges. But they believed that with the advent of a classless society, scarcity would give way to abundance, and conflict to harmony. The state would then wither away, to be replaced by a new administrative order without organized violence. The state, then, was a response to a pathological although historically necessary condition. Ultimately, however, the evolution of society would bring about the changes that would make Rousseau's vision possible. For Augustine the earthly state was the palliative for sin; for Marx it was the palliative for class conflict. But for both there was a condition of ultimate redemption, where the coercive state would have no place.

For John Locke civil society (equivalent in Locke's terms to the state) existed to safeguard the natural rights of individuals, which they could not successfully preserve in the state of nature. Nevertheless, because Locke considered people rational by nature and therefore ideally capable of living in peace according to the law of nature, the condition of conflict was pathological, not natural. However, the norm was not participation in a transcendent good but a condition in which everyone enjoyed their own area of legitimate privacy, troubled by neither private nor public intrusions. For Locke, as for Hobbes, the state's ends were reducible to those of individuals. Bentham put this quite unequivocally: "This public interest … is only an abstract term; it represents only the mass of the interests of individuals" (*Principles of the Civil Code, Works,* Vol. I, p. 321). The state had and could have no moral function except to arrange that as many people as possible should obtain as much as possible of whatever it was that they wanted. For some purposes all that was needed was for the state to uphold property and the sanctity of contract; economic motives in a free market would do the rest. But Benthamite utilitarianism was committed to active state policies wherever, as in public health, laissez-faire would not work. The Benthamite state was readily convertible to a Fabian policy of social engineering. But the objective would still be, in Roscoe Pound's phrase, "such an adjustment of relations and ordering of conduct as will make the goods of existence … go round as far as possible with the least friction and waste" (*Social Control through Law,* New Haven, CT, 1942, p. 65).

The view that politics is a matter of who gets what is substantially that of the group theorists in political science, such as A. E. Bentley and, more recently, Harold Lasswell, David Truman, and Robert Dahl. In their accounts, the state is dissolved into a "political process" which can be analyzed without residue in terms of the competitive pressures of interests. Whereas Locke and Rousseau would have agreed that the public interest was the proper end of state action (although possibly disagreeing in their accounts of it), many modern political scientists, Glendon Schubert, for instance, have rejected the concept of public interest as being so vague as to be useless or as being a device of politicians for advocating policies actually pursued for quite other reasons. Policy decisions, they argue, are the resultants of competing interests—there is no single interest that everyone would acknowledge, nor one that would be to everyone's advantage. Thus, there can be no public interest that the state ought to pursue.

An analysis like Schubert's depends, on the one hand, on the identification of interest and desire and, on the other hand, on interpreting "public" to mean "enjoyed by everyone." This was clearly not Rousseau's meaning. A citizen's interest was in being a person of a certain kind with characteristic excellences, attainable only in a healthy state. One might misguidedly desire what was not in his interest; so might all the citizens, for the will of all was not necessarily the same as the general will. But as long as their vision was clear, conflict was impossible because the public interest was whatever would be to *anyone's* advantage, insofar as he was capable of human excellence.

Political scientists mistrust such a theory, partly because it tends to describe the actual state as if it were the ideal and partly because it is evaluative, whereas they want theories to be descriptive and explanatory. What is in a man's interest, they say, is simply what he strives to get, irrespective of why he does so or with what wisdom. However, treating the state as simply an arena for sectional pressures has the drawback of disregarding or misconstruing the widespread opinion that to act in the

public interest is to be impartial between competing groups—that the state (or its rulers) is therefore in a special position as arbiter between group interests. This frequently gives state decisions a moral authority that a mere political barometer, responding to the greatest pressures, could never enjoy, and it provides politicians and public servants, potentially at least, with a range of motives that are quite unlike interests as usually understood.

Sheldon S. Wolin, in *Politics and Vision,* advanced the somewhat paradoxical thesis that despite the vast extension of governmental activity, there has been a steady depreciation of politics and the political order since the seventeenth century. This has been matched, he asserts, by a corresponding heightening of regard for nonpolitical institutions and associations—for society as distinct from the state. This "groupism" is regrettable, in Wolin's view, because the specialized roles adopted by the individual are no substitute for citizenship. Citizenship, as the individual's most general role, calls on him to choose regardless of special interests. As a member of a society bounded for most purposes by the state's frontiers, he is confronted with this demand only as a member of the state. As a trade unionist, for instance, he shares sectional loyalties with coworkers and is led to strive for advantages at the expense of other groups. To be conscious of oneself as a citizen, however, is to enjoy an integrative experience, which "demands that the separate roles be surveyed from a more general point of view." The political art, in Wolin's opinion, is that "which strives for an integrative form of direction, one that is broader than that supplied by any group or organization." Wolin comes close indeed to the view of Rousseau and Hegel that there is a concrete morality in the state. As a citizen one is asked to judge what would be to the advantage of *anyone,* their special circumstances aside. In this manner one approaches a moral judgment, an impartial assessment of claims in matters of general concern.

A further disadvantage of a fragmented vision of the political process is its tendency to miss the influence of the state, both as an idea and as a tradition, on the life of the society. As a trade union or a church is not simply an arena for its own sectional interests, so each state embodies a set of values and objects of loyalty which may greatly influence what its members consider their interests to be. Its manners and traditions leave their mark on them. Associations that participate in its political processes reflect its style, its modes of organization, and its procedures. Moreover, the state lays down terms on which its members deal with one another and with foreigners, establishing an area within its borders in which trade, communications, and movement are free, and regulating traffic that crosses them. Because of its regulative power, the texture of social relations is far closer within its boundaries than across them. It thus supplies not only a legal but also a general conceptual framework for much of our social thought and action. Thus, where we speak of Australian primary producers' associations, Australian football teams, and the Australian Political Studies Association, we speak not of the Australian state but of Australia.

This seems to support the Hegelian view of the state as a national community within which certain particular functions are promoted by sectional associations operating within it. But then one must distinguish the state in this sense from its governmental authority structure, which would be but one of its organs alongside trade unions, graziers' associations, and the like. For voluntary and sectional associations are not, like departments of state, of the navy, or of the post office, subordinate parts of the governmental structure, nor are their actions the acts of the state. This distinction would be quite consistent with a generalized although conditional duty on the part of sectional associations to submit to governmental authority. However, it would not be a duty owed by subordinate agencies to a superior but rather one owed by members of a society in which an authority is recognized as arbiter and coordinator of interests and as initiator of policies of general concern. This would also be consistent with the moral right of associations to defy the government should these functions be abused. The fact that the government is the executive agent of the politically organized state does not mean that its own views of the public interest or of a just settlement of conflicting claims must always and necessarily prevail.

The word *nation* is often used to refer to the state-community; so, in slightly different contexts, is the word *country.* Both words, however, have other meanings and overtones, *nation* being used of cultural groups which can transcend state frontiers or which may be minorities within a state, *country* referring more particularly to the state's territory or to the state as an international personality.

LIMITS OF STATE ACTION

Liberal political philosophers have tried to define necessary limits beyond which the activities of the state must not extend. Some, like Locke, account for the existence of the state in terms of some specific function, such as the safeguarding of natural rights. They then infer, by

analogy with the statement of aims in the articles of association of a club or joint-stock company, that the state would be exceeding its competence if it did more than that. Others have tried to define an area of private action that the state ought not invade. According to J. S. Mill, for instance, the state is never justified in restraining the action of a normal adult solely on the grounds that it is in his interests that it should. Some, like T. H. Green and Ernest Barker and, in a more sophisticated form, F. A. Hayek, have claimed that the state as a coercive organization has intrinsic limitations. Although it can hinder hindrances to the good life, it cannot force people to live that life; any form of activity, such as religion, art, or science, whose value lies in spontaneity or freedom of belief must therefore fall outside its scope. Barker argued that because the state's essential mode of action was through general rules, it was not apt for any field that, like industry, required ad hoc discretionary decisions. Such an argument depends, however, on a very doubtful kind of essentialism. The state has no one modus operandi. For the varied range of activities that states have undertaken since the mid-1800s, they have devised an equally varied range of techniques. They encourage the arts as well as censoring them. Nearly all modern states have very extensive responsibilities in education, industrial management, health insurance, and medical services, all of which have at one time been private undertakings and none of which involves coercion except in very remote or indirect ways. It does not follow from the state's monopoly of legitimate coercion that it can do nothing for which coercion is inappropriate. Nor need we suppose that, if there are indeed forms of social activity that the state has at present no satisfactory means of regulating, encouraging, or promoting, it may not yet invent them. Therefore, one cannot say in advance whether a given task would be more properly left to individual initiative or organized by governmental agencies. That depends on what can be done with the techniques available.

See also Aristotle; Augustine, St.; Bentham, Jeremy; Bodin, Jean; Engels, Friedrich; General Will, The; Green, Thomas Hill; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Hobbes, Thomas; Locke, John; Machiavelli, Niccolò; Marx, Karl; Mill, John Stuart; Nationalism; Plato; Political Philosophy, History of; Pufendorf, Samuel von; Punishment; Renaissance; Rousseau, Jean-Jacques; Social Contract; Society; Sovereignty; Suárez, Francisco; Vitoria, Francisco de.

Bibliography

Virtually all the works listed in the bibliography to the Sovereignty entry are also relevant to this topic. The works listed here are therefore additional references.

HISTORICAL TREATMENTS

Cassirer, Ernst. *The Myth of the State.* New Haven, CT: Yale University Press, 1946.

Jouvenel, Bertrand de. *Du pouvoir. Histoire naturelle de sa croissance.* Geneva, 1945. Translated by J. F. Huntington as *On Power: Its Nature and the History of Its Growth.* New York: Viking, 1949.

Nisbet, Robert. *The Quest for Community.* New York: Oxford University Press, 1953. Republished as *Community and Power.* New York: Oxford University Press, 1967.

Plamenatz, John. *Man and Society.* 2 vols. London: Longman, 1963.

Wolin, Sheldon S. *Politics and Vision.* Boston: Little, Brown, 1960.

ANALYTICAL WORKS

Benn, S. I., and R. S. Peters. *Social Principles and the Democratic State.* London: Allen and Unwin, 1959. Republished as *Principles of Political Thought.* New York: Collier, 1964.

Jouvenel, Bertrand de. *De la souveraineté.* Paris: M. T. Génin, 1955. Translated by J. F. Huntington as *Sovereignty: An Inquiry into the Political Good.* Chicago: University of Chicago Press, 1957.

Jouvenel, Bertrand de. *The Pure Theory of Politics.* New Haven, CT: Yale University Press, 1963.

Mabbott, J. D. *The State and the Citizen.* London: Hutchinson's University Library, 1948.

THE APPROACH OF MODERN POLITICAL SCIENTISTS

Bentley, A. F. *The Process of Government.* Chicago: University of Chicago Press, 1908.

Dahl, Robert A. *Modern Political Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1963.

Friedrich, Carl J. *Man and His Government.* New York: McGraw-Hill, 1963. Includes extensive bibliography.

Friedrich, Carl J., ed. *The Public Interest.* New York: Atherton Press, 1962. This is Vol. V of American Society of Political and Legal Philosophy, ed., *Nomos.*

Lasswell, Harold D., and Abraham Kaplan. *Power and Society: A Framework for Political Inquiry.* New Haven, CT: Yale University Press, 1950.

Schubert, Glendon. *The Public Interest.* Glencoe, IL: Free Press, 1960.

WORKS IN THE CLASSICAL LIBERAL TRADITION

Barker, Ernest. *Principles of Social and Political Theory.* Oxford, 1951.

Bentham, Jeremy. *A Fragment on Government* (London, 1776). Edited by W. Harrison. Oxford: Blackwell, 1948. This volume also contains *Introduction to the Principles of Morals and Legislation.*

Bentham, Jeremy. *Traités de législation, civile et pénale,* edited by E. Dumont. Paris, 1802. Translated by R. Hildreth as *The Theory of Legislation.* Boston: Weeks, Jordan, 1840; London: Trubner, 1864.

Bentham, Jeremy. *Works,* edited by John Bowring, 11 vols. London, 1843.

Hayek, F. A. *The Constitution of Liberty.* London: Routledge and K. Paul, 1960.

Locke, John. *Two Treatises of Government* (1690), edited by Peter Laslett. Cambridge, U.K.: Cambridge University Press, 1960.

Mill, J. S. *On Liberty* (1859) and *Considerations on Representative Government* (1861). Edited by R. B. McCallum. Oxford, 1946; London, 1947.

Popper, Karl. *The Open Society and Its Enemies.* 2 vols., 4th rev. ed. London: Routledge and K. Paul, 1962.

MARXIST VIEWS

Bottomore, T. B., and M. Rubel, eds. and trs. *Karl Marx: Selected Writings in Sociology and Social Philosophy.* 2nd ed. London: Penguin, 1961. Includes selected bibliography of Marx's writings.

Engels, Friedrich. *Der Ursprung der Familie, des Privateigenthums und des Staats.* Zürich, 1884. Translated by Lewis H. Morgan as *Origin of the Family, Private Property and the State.* 4th ed. London, 1946.

Lenin, V. I. *Gosudarstvo i Revoliutsiia.* Petrograd, 1918. Translated as *State and Revolution* in *Selected Works,* edited by J. Fineberg. Vol. VII. London, 1946.

Marx, Karl, and Friedrich Engels. *Die deutsche Ideologie* (1845–1846). In *Marx–Engels Gesamtausgabe.* Berlin, 1932. Part I, Vol. V. Parts I and III were edited by R. Pascal and translated by William Lough and Charles P. Magill as *The German Ideology.* London: Lawrence and Wishart, 1938.

Marx, Karl, and Friedrich Engels. *Manifest der kommunistischen Partei.* London, 1848. Edited with introduction by Harold J. Laski as *Communist Manifesto: Socialist Landmark.* London: Allen and Unwin, 1948.

CATHOLIC THEORIES OF NATURAL LAW

Entrèves, A. P. d'. *La dottrina dello stato; elementi di analisi e di interpretazione.* Turin: Giappichelli, 1962.

Maritain, Jacques. *Man and the State.* Chicago: University of Chicago Press, 1951.

FASCIST VIEWS

Gentile, Giovanni. *Genesi e struttura della società.* Florence: Sansoni, 1946. Translated and edited by H. S. Harris as *Genesis and Structure of Society.* Urbana: University of Illinois Press, 1960. The views of the official philosopher of Italian fascism from the standpoint of "actual idealism."

Stanley I. Benn (1967)

STATE [ADDENDUM]

In the past three centuries, states have replaced empires and tribes as the dominant form of political organization. But one clear lesson of the twentieth century is that the vast powers of states can be put to disastrous as well as beneficent ends.

Philosophical reflection about states often begins with Thomas Hobbes and the rational justification of social order as mutually advantageous. Many more contemporary philosophers have ignored the state, however, focusing instead on justice and the rights and liberties that states should respect. Indeed, the most important work in political philosophy in the twentieth century (Rawls 1972) does not discuss the state—it lacks even one entry for "the state" in its index.

In recent years there has, however, been a renewed interest in the state that has developed along several lines. Some have used modern game theory to pursue Hobbes's question of the possibility of a rational justification of the state. Others have studied the nature of the state itself and its relationship with other forms of social control, while some have questioned both the authority and the legitimacy of states. Another topic is the impact on states of global economic, social, and legal transformations.

Questions about the nature of states can be addressed either by considering the similarities and differences among states, nations and governments or by comparing states with other ways of maintaining social order. Nations and peoples are distinct from states, as evidenced by the fact that we often speak of "stateless" peoples such as Kurds and Palestinians. Nations and peoples are marked by common cultures and histories that provide the basis of a shared identity. Governments are also distinct from states: the head of the government in the United Kingdom (the prime minister) is not the head of state (the monarch is), just as the U.S. president is the head of state but not of the government as a whole. What, then, are states?

Unlike both ancient empires and the overlapping allegiances of feudal Europe, states claim sovereignty, and of a specific sort. Empires lacked clear territorial boundaries and often shared sovereignty with local rulers. In feudal Europe political power was fragmented among different and often overlapping jurisdictions that encompassed kings, lords, local rulers, bishops, and popes who demanded allegiance or taxes or both. Sovereign states differ from these forms of political control because they have a centralized and hierarchical organization ruling over a defined territory with established boundaries. A state also claims to be the ultimate source of legal authority and demands loyalty from all permanent inhabitants within its territory.

Although many assume that states' claims to authority and legitimacy could be vindicated—that states could be made just—anarchists have questioned both claims. Robert Paul Wolff (1976) attacked the state's authority by attacking authority in general. He argued that because people are responsible for their own decisions based on

reasons that they understand to be relevant, the claimed authority of states is illegitimate. One cannot both accept responsibility for one's own actions and submit to the authority of the state, said Wolff. This claim has spurred Joseph Raz (1979) and others to look more closely at authority. Raz agrees that authority involves a type of surrender or acquiescence of judgment, though he denies that this is always contrary to reason. He explains by distinguishing first-order reasons (where we weigh competing reasons and act accordingly) from second-order reasons that "preempt" first-order reasons. The eclipsing of first-order reasons by the authority's judgment suggests that Wolf is right in casting doubt on the state's claim that it is always an authority, although it also implies that it is sometimes not a violation of autonomy to decide to act for second-order reasons.

Robert Nozick (1974) raised questions not only about the state's authority but also about the widely presumed legitimacy of the state's use of coercive power. The only legitimate exercise of coercive power, he argued, would be vastly different from powers states commonly claim. A legitimate state's power is limited, for example, by people's rights to refuse to join the state or to join only on terms that are voluntarily. While Nozick defends the state's use of coercion to protect rights to property and life, he questions whether the many other, familiar coercive measures are legitimate—measures ranging from paternalistic efforts to protect people against themselves to laws preventing self-regarding but immoral acts to taxes aimed at redistributing wealth and providing social services. In painting an attractive and purportedly workable picture of an anarchist society, both Wolff and Nozick have encouraged a fresh look at states' claims to authority and legitimacy as well as at alternative methods of maintaining social order.

Economic, legal, and social forces are also affecting states. States traditionally claim both internal sovereign control over populations and immunity from external power, yet both ideas have come under increasing pressure from many different angles. As the world has become smaller and more integrated and corporations do business in different states, it is often important for states to harmonize laws governing commerce and immigration. Adding to these pressures for more cooperation has come a need to meet growing international problems such as environmental degradation and terrorism—neither of which can be effectively addressed without the cooperation of other states. This greater interdependence of states, and their mutual vulnerability, has even sparked renewed interest in possible preemptive actions against states as a form of self-defense.

Alongside these challenges to the external sovereignty of states has come greater emphasis on human rights, further weakening states' claims of internal sovereignty over their own populations. International tribunals, nongovernment aid organizations, and sometimes unilateral military action in the name of helping citizens or protecting them from their own states have all challenged the supremacy of state power. Yet despite all these forces working to limit states' sovereignty, terrorism has also brought home the importance of avoiding "failed states" in which terrorists can train and plan. So although states are losing authority and sovereignty because of globalization, mutual interdependence, and growing legal limits on their power, the prospect of failed states breeding terrorists abroad and anarchy at home has strengthened the case of defenders of the state power.

See also Anarchism; Authority; First-Order Logic; Hobbes, Thomas; Justice; Liberty; Nozick, Robert; Political Philosophy, History of; Rights; Sovereignty; Terrorism.

Bibliography

Fukuyama, Francis. *State-Building: Governance and World Order in the 21st Century*. Ithaca: Cornell University Press, 2004.

Green, Leslie. *The Authority of the State*. Oxford: Clarendon Press, 1990.

Morris, Christopher. *An Essay on the Modern State*. Cambridge, U.K.: Cambridge University Press, 1998.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1972.

Raz, Joseph. *The Authority of Law.* Oxford: Oxford University Press, 1979.

Wolff, Robert Paul. *In Defense of Anarchism.* 2nd ed. New York: Harper and Row, 1976.

John Arthur (2005)

STATEMENTS

See *Propositions*

STATISTICAL MECHANICS, PHILOSOPHY OF

See *Philosophy of Statistical Mechanics*

STATISTICS, FOUNDATIONS OF

Thorny conceptual issues arise at every turn in the ongoing debate between the three major schools of statistical theory: the Bayesian (B), likelihood (L), and frequentist (F). (F) rather uneasily combines the Neyman-Pearson-Wald conception of statistics as "the science of decision making under uncertainty" with Ronald A. Fisher's theories of estimation and significance testing, viewed by him as inferential. However, in keeping with his frequentist conception of probability, Fisher viewed the inferential theory of Thomas Bayes and Pierre Simon de Laplace as applicable only where the needed prior probability inputs are grounded in observed relative frequencies. Maximum likelihood estimates and significance tests were intended as substitutes for Bayesian inference in all other cases. (F), (B) and (L) all provide a framework for comparatively appraising statistical hypotheses, but Fisher questioned whether one can fruitfully assimilate the weighing of evidence to decision making.

Given the response probabilities for a diagnostic test shown in Table 1:

TABLE 1

	Positive	Negative
Infected (h)	0.95	0.05
Uninfected (k)	0.02	0.98

one may, following Richard M. Royall (1997, p. 2), usefully distinguish three questions of evidence, belief, and decision when a subject (S) tests positive:

Q1. Is this result evidence that S has the disease?

Q2. What degree of belief that S has the disease is warranted?

Q3. Should S be treated for the disease?

(L) addresses only Q1 and does so by what Ian Hacking (1965) dubs the law of likelihood (LL):

> evidence e supports hypothesis h over k if and only if $(Pe|h) > P(e|k)$; moreover, *the likelihood ratio* (LR), $P(e|h) : P(e|k)$, measures the strength of the support e accords h over k.

The LL follows from Bayes's fundamental rule for revising a probability assignment given new data. Indeed, Laplace arrived (independently) at this rule by appeal to the intuition that the updated odds in favor of h against k in light of e should be the product of the initial odds by the LR (Hald 1998, p. 158):

(1)
$$P(h \mid e) : P(k \mid e) = \frac{P(e \mid h)}{P(e \mid k)} \times \frac{P(h)}{P(k)}$$

If the rival (mutually exclusive) hypotheses h and k are treated as exhaustive, so that their probabilities sum to one, then (1) yields the usual form of Bayes's rule:

(2)
$$P(h \mid e) = \frac{P(e \mid h)P(h)}{P(e)}$$

with $P(e)$ usually given in the general case by the partitioning formula:

(3) $$P(e) = P(e|h_1)P(h_1) + \ldots + P(e|h_n)P(h_n)$$

with the (mutually exclusive) considered hypotheses h_1, . . . , h_n treated as exhaustive.

One also sees how (B) answers Q2 by multiplying the initial odds, based on what is known about the incidence of the disease, by the LR of 95/2 provided by a positive reaction. If the incidence of the disease is even as low as 1 per 1,000, the posttest (or "posterior") probability of infection may still lie well below 50 percent. Notice, too, that knowledge of the infection rate may rest on the same sort of empirical frequency data that underwrites the conditional probabilities of Table 1. When this is true, (L) and (F) have no qualms about applying (2) to answer Q2. They do not question the validity of (2), only whether the initial probabilities needed to apply it can be freed of the taint of subjectivism.

THE LIKELIHOOD PRINCIPLE

Statistical hypotheses typically assign values to one or more parameters of an assumed probability model of the experiment, for example, to the mean of a normal distribution or the probability of success in a sequence of Bernoulli trials. If θ is such a parameter and X the experimental random variable then

$$P(x|\theta)$$

is called the sampling distribution when considered as a function of the observation x and the likelihood function qua function of θ.

The case of randomly sampling an urn with replacement, with p the population proportion of white balls, affords a simple illustration. Then the probability of x white and n-x black in a sample of n is given by the binomial (sampling) distribution:

$$P(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

For comparing two hypotheses about p by the LR, the binomial coefficients cancel and so one may ignore them and define the likelihood function for this experiment by:

$$L(p) = p^x(1-p)^{n-x}$$

The value of p, which maximizes $L(p)$, is called the maximum likelihood (ML) estimate of p and is easily found, by calculus, to be x/n, the observed sample proportion (of white balls) or successes.

Consider, next, a second experiment in which one samples until the first success is observed. This happens on trial n with probability, $p(1-p)^{n-1}$, since $n-1$ failures must precede the first success. More generally, if one samples until the *rth* success is observed, this happens on trial n with probability:

$$P(r \mid n, p) = \binom{n-1}{r-1} p^r (1-p)^{n-r}$$

which reduces to $p(1-p)^{n-1}$ when $r = 1$. This sampling distribution is called the negative binomial (or waiting time) distribution; it gives rise to the same likelihood function as the first experiment.

Now suppose Jay elects to observe $n = 30$ trials and finds $x = 12$ successes, while May elects to sample until she finds $r = 12$ successes but that happens to occur on the thirtieth trial. In a literal sense, both experimenters have observed the same thing: twelve successes in thirty Bernoulli trials. One would think they would then draw the same conclusions. (F) violates this prescription, called the likelihood principle (LP). In so doing (F) allows the experimenter's intentions when to stop sampling to influence the evidential import of what is observed. It also makes the import of the outcome observed dependent on the entire sample space, hence, on outcomes that might have been but were not observed (see de Groot 1986, p. 417). By the same token, the unbiased estimators favored by (F), those centered on the true value of the parameter, violate the LP (p. 417), since this concept depends on all possible values of the estimator. Thus, the unbiased estimates of p are, respectively, k/n and $(k-1)/(n-1)$ for the two previous experiments. The LP virtually defines the difference between (B) and (L), on the one hand, and (F), on the other.

In effect, (B) and (L) charge (F) with inconsistency, with basing different assessments of the evidence (or different decisions to accept or reject hypotheses) on equivalent outcomes, for two outcomes are accounted equivalent by the LP if they define the same likelihood function. This charge of inconsistency can be carried to a higher metalevel since (F) accepts Bayes's rule (2), and with it the LP, when the prior probabilities are known from past frequency data. Hence (F) finds itself in the odd position of accepting or rejecting the LP according as the prior probabilities are "known" or "unknown." Charges of inconsistency are the weapon of choice in the ongoing battles between the three schools, beginning with the charge that Bayes's postulate for assigning a uniform distribution to a parameter about which nothing is known leads to inconsistent assignments. In the sequel, one will explore how consistency may be used instead to forge agreement.

FISHERIAN SIGNIFICANCE TESTS

Fisher (1935, chapter 2, the *locus classicus*) presented significance tests as analogues of the logicians' *modus tollens*: if A then B, not-B/$\therefore$not-A. When the probability, $P(e|h_0)$, falls below α, one counts e as evidence against h_0, the smaller α, the stronger the evidence. As Fisher describes it, the logic is "that of a simple disjunction: Either an exceedingly rare chance has occurred, or the theory is not true." Using (2), the probabilistic analogue of *modus tollens* is:

$$P(A \mid \bar{B}) = P(A)\left[\frac{P(\bar{B} \mid A)}{P(\bar{B})}\right]$$

which shows that for not-B to seriously infirm A requires, not merely that $P(\bar{B}|A)$ be small, but small relative to $P(\bar{B})$, so that some alternative to A must accord not-B a higher probability.

Much of Fisher's practice conforms to this precept. In his famous example of the tea-tasting lady (1935), the lady claims that she can tell whether tea or milk was infused first in a mixture of the two. To test her claim she is asked to classify eight cups of which four are tea-first and the other four milk-first, but, of course, she does not know which four. The relevant statistic is the number R of correct classifications and its sampling distribution on the null hypothesis that she lacks such ability is:

$$P(R = r \mid h_0) = \frac{\binom{4}{r}\binom{4}{4-r}}{\binom{8}{4}}$$

Notice, the probability that $R = r$ on the alternative hypothesis of skill cannot be computed so that likelihood

ratios do not exist. All that one has to work with is an intuitive rank ordering of the outcomes with larger values of R more indicative of skill. What $P(R \geq r^*|h_0)$ measures may be verbalized as "the probability of obtaining, by chance, agreement with the hypothesis of skill as good as that observed" (Fisher 1935, p.13). Although Fisher rejected the implication that by "disproving" the null hypothesis one "demonstrates" the alternative (p. 16), he also says that "we should admit that the lady had made good her claim" (p. 14) if she classified all eight cups correctly. He argues that one can (effectively) disprove the null hypothesis because it is "exact," while the alternative of skill is vague. However, this does not preclude one from adopting the natural view of most researchers that a significant result is evidence in favor of the alternative hypothesis. The null hypothesis is then cast in the subtly different role of a fixed point of comparison that permits computation of the relevant chance probability (Rosenkrantz 1977, chapter 9).

This is, in fact, the logic of most nonparametric tests, the Wilcoxon rank sum test for comparing two treatments being paradigmatic (see Hodges and Lehmann 1970, §§12.3–12.4, especially p. 333). Table 2 compares the survival times (in years) following a heart attack of $t = 6$ patients receiving a new treatment and $s = 4$ controls receiving the standard treatment, with their ranks in parentheses.

TABLE 2

Treated	7.3 (4)	17.2 (1)	6.1 (6)	11.4 (3)	15.8 (2)	5.2 (7)
Controls	1.4 (9)	0.6 (10)	5.0 (8)	6.7 (5)		

The sum, W_t of the ranks of the t-treated patients is a suitable test statistic, and under the null hypothesis that the new treatment is no better than the old, all $\binom{10}{6} = 210$ assignments of ranks 1 through 10 to the six treated patients are equiprobable. Hence, the paucity of possible rank sums as small as the observed value, $W_t = 1 + 2 + 3 + 4 + 6 + 7 = 23$, measures the strength of the evidence, the smaller this proportion the stronger the evidence of improved efficacy. Since only three other possible rank sums are as small as the observed value of W_t, the relevant proportion is $4/210 = .019$, or about 2 percent.

This same form of argument also enjoys widespread currency in the sciences, as when an anthropologist maintains that certain cultural commonalities are too numerous and striking to be ascribed to parallel development and point instead to contact between two civilizations, or when an evolutionist argues that the structural similarities between two organs that do not even perform the same function in two species are homologous and not merely analogous, hence indicative of common ancestry. Indeed, the rationale behind the principle of parsimony—that a phylogeny is more plausible if it requires fewer evolutionary changes—is this same piling up of otherwise improbable coincidences. And how improbable that various methods of reconstructing a phylogeny—for example, the ordering of fish, amphibians, reptiles, and mammals—based on the fossil record, homologies, serology, or DNA and protein sequencing should all agree if the phyla in question were separately created?

Fisher's foremost contribution to the design of experiments, randomization, also fits this logic (Fisher 1935, pp. 17–21, 41–44, 62–66). If, for example, the treated subjects of Table 2 were all younger than the controls, they might be expected to live longer in any case. However if, after controlling for such plainly relevant differences, the patients were assigned at random to the two groups, the chances are just one in $\binom{s+t}{t}$ that all treated subjects will share some hidden trait conducive to longevity that is lacking in the controls, thus removing any suspicion of selection bias. In addition, randomization underwrites the probability model of the experiment from which the sampling distribution of the chosen test statistic, W_T, is deduced (for a more leisurely discussion of randomization, see Hodges and Lehmann 1970, §12.1).

Since significance tests apply, on this reading, only when the likelihood function does not exist, they can be viewed as complements rather than alternatives to the methods of (B) or (L). Seen in this positive light, significance tests have a deeper Bayesian rationale. For the paucity of possible outcomes a model with zero or more adjustable parameters accommodates measures the support in its favor when the observed outcome belongs to this set (Rosenkrantz 1977, chapter 5). Echoing I. J. Good (who echoed Fisher), to garner support requires not just accuracy but improbable accuracy.

Moreover, the present formulation resolves many of the controversies that have swirled about significance testing (see Morrison and Henkel 1970), above all, the question whether a significant outcome with a small sample constitutes stronger evidence against null than one with a large sample (see Royall 1997, pp. 70–71). If, in

fact, the chance probability of agreement with the causal hypothesis of interest is the same in both cases, the evidence in favor of that causal hypothesis is also equally strong.

All these advantages notwithstanding, significant test results are still most widely viewed as evidence against the null hypothesis and, indeed, without reference to alternative hypotheses (see Fisher 1935, pp. 15–16; 1956, pp. 40–42; and for a critique of this viewpoint, Royall 1997, chapter 3). Thus, one classifies the observed outcome as evidence for or against h_0 not by comparing its probability on h_0 to its probability on alternative hypotheses but by comparing its probability on h_0 with that of other possible outcomes.

NEYMAN-PEARSON THEORY

In the late 1920s Jerzy Neyman and Egon S. Pearson (henceforth, NP) set forth a new approach to the testing of statistical hypotheses. Although initially presented as a refinement of Fisherian significance testing, NP actually addressed the different problem of testing one hypothesis against one or more alternatives in situations where the likelihoods do exist. In such cases, Fisher's practice, in accord with (L), was to compare the relevant hypotheses by their likelihoods. NP proposed, instead, to lay down in advance a rule of rejection, that is, a critical region *R* of the space of outcomes such that the tested hypothesis is rejected just in case the outcome actually observed falls in *R*.

In the simplest case of testing one point hypothesis, $h_0 : \theta = \theta_0$ against another, $h_1 : \theta = \theta_1$, called simple dichotomy, one can err not only by rejecting h_0 when it is true but also by accepting h_0 when the alternative hypothesis, h_1, is true. Plainly, one cannot reduce both these error probabilities,

$$\alpha = P(X \in R|h_0)$$

and

$$\beta = P(X \notin R|h_1)$$

without increasing the sample size. NP's recommended procedure was to so label the hypotheses that rejecting h_0 is the more serious error, fix α at a tolerable level, α_0, called the size or significance level of the test, and then among all tests of this size, $\alpha \leq \alpha_0$, choose the one that minimizes β, or, equivalently, maximizes the power $1 - \beta$. The test is thus chosen as the solution of a well-defined optimization problem, a feature modeled on Fisher's approach to estimation. The fundamental lemma of NP theory then affirms the existence of a unique solution, that is, the existence of a most powerful test of a given size. Finally, test statistics could then be compared in terms of their power. The overall effect was to unify point estimation, interval estimation (confidence intervals), and testing under the broader rubric of "decision making under uncertainty," a viewpoint made explicit in the later work of Abraham Wald. In this scheme of things, estimates, confidence intervals, and tests are to be judged solely in terms of such performance characteristics as their mean squared error or their error probabilities. That is, arguably, the feature of the approach that continues to exercise the most powerful influence on the orthodox (i.e., frequentist) school (see Hodges and Lehmann 1970, chapters 11–13; de Groot 1986, chapter 7).

These developments occurred in such rapid succession that they have yet to be fully digested. NP had uppermost in mind massed tests like screening a population for a disease, testing a new drug, or industrial sampling inspection where the same practical decision, such as classifying a patient as infected or uninfected, must be faced repeatedly. For such situations, a reliable rule that controls for the probability of error seemed preferable to an explicitly (Bayesian) decision theoretic treatment that would require prior probabilities that the statistician could not base on any objective rule, as well as on loss or utility functions that would vary even more from one policy maker to another. To be sure, one might know the distribution of the proportion of defectives from past experience with a manufacturing process and be able to supply objective cost functions, but such cases would be uncommon.

But even in cases where an assembly line approach seems appropriate, NP's recommended procedure is open to question. If the more serious type 1 error is deemed, say, a hundred times more serious than the less serious type 2 error, should one not prefer a test whose probability of committing the more serious error is correspondingly less than its probability of committing the less serious error? In short, why not minimize the weighted sum, $100\alpha + \beta$? After all, the result of fixing α at some tolerable level, then minimizing β, might be to drive β much lower than α, which is wasteful, or else to drive β so high as to render the test powerless. This point is not merely academic, for a random sample of some seventy-one clinical trials revealed that overemphasis on controlling type 1 error probability led to a 10 percent risk of missing a 50 percent therapeutic improvement (Good 1983, p. 144).

To minimize the total risk, $a\alpha + b\beta$, one finds, writing $f_i(x) = P(X = x_i|h_i)$, $i = 1, 2$, that

$$a\alpha + b\beta = aP(X \in R \mid h_0) + bP(X \notin R \mid h_1)$$
$$= aP(X \in R \mid h_0) + b[1 - P(X \in R \mid h_1)]$$
$$= b + \sum_{x \in R}[af_0(x) - bf_1(x)]$$

Hence, the total risk is minimized by making $af_0(x) - bf_1(x) < 0$ for all $x \in R$. Then h_0 is rejected when

$$f_1(x){:}f_0(x) > a{:}b$$

which says: Reject h_0 (in favor of h_1) when the LR in favor of h_1 exceeds the relative seriousness, $a{:}b$, of the two kinds of error. More advanced readers will recognize this as a Bayesian decision rule for the special case of constant regret functions, appropriate in situations where "a miss is as good as a mile," and equal prior probabilities. In the general case, one may interpret $a{:}b$: as the product of the prior odds by the ratio of the regrets. The fundamental lemma then drops out as an easy corollary (de Groot 1986, p. 444), where the most powerful test of size α has critical region, $R = \{x{:}f_1(x){:}f_0(x) > k\}$, with k the least number for which $P(X \in R|h_0) \leq \alpha$. The main virtue of this approach, however, is that it allows one to adjust the sample size so as to achieve a tolerable level of overall risk. Roughly speaking, one goes on sampling until the marginal cost of one more item exceeds the marginal risk reduction.

NP's decision theoretic formulation notwithstanding, users of statistical tests have continued to interpret them as evidence and to view NP tests as a refinement of Fisher's significance tests. One reason for this is that NP continued to use the language of hypothesis testing, of accepting or rejecting hypotheses. A more important reason is that in many, if not most, scientific inquiries, practical decisions are nowhere in view. Even where questions of public policy impinge, as in the smoking-cancer or charter school controversies, it is deemed necessary to first weigh the evidence before deciding what policy or legislation to adopt. The tendency of NP is to subsume the individual test under a rule of specifiable reliability. Rejection of h_0 at a 5 percent level does not mean that the probability is 0.05 that a type 1 error was committed in this case, much less that h_0 has probability 0.05 given the outcome. The error probability refers to the procedure, not the result. However, this raises new concerns.

Consider a test of normal means of common (unknown) variance, σ^2, $h_0{:}\mu = \mu_0$ *versus* $h_1{:}\mu = \mu_1$. The optimal 5 percent test rejects h_0 when $\bar{x} \geq \mu_0 + 1.64\sigma/\sqrt{n}$, where n is the sample size and $\bar{x} = (x_1 + \ldots + x_n)/n$ is the sample mean. For as Carl Friedrich Gauss first showed, $\bar{x} \sim N(\mu, \sigma^2/n)$, that is, the sample mean for independent and identically distributed normal variates, $X_i \sim N(\mu, \sigma^2/n)$, is normally distributed about their common mean, μ, with variance, σ^2/n, or precision, n/σ^2, n times that of a single measurement. For example, if $\mu_0 = 0$, $\mu_1 = \sigma^2 = 1$, and $n = 30$ so that $\sigma^2/n = 0.18$, then h_0 is rejected when $\bar{x} \geq .30$. However, $\bar{x} = .30$ is $.70/.18 = 3.89$ standard deviation units below the mean of $\mu = 1$ posited by h_1, and thus much closer to $\mu_0 = 0$. It is strange that such an observation should be interpreted as strong evidence against h_0. Indeed, the LR given a random sample of n measurements is:

$$\frac{(2\pi\sigma^2)^{n/2}\prod \exp(-\frac{1}{2\sigma^2}(x_i-\mu_0)^2)}{(2\pi\sigma^2)^{n/2}\prod \exp(-\frac{1}{2\sigma^2}(x_i-\mu_1)^2)} = \exp\{-\frac{1}{2\sigma^2}\sum[(x_i-\mu_0)^2-(x_i-\mu_1)^2]\}$$

which, using $\sum x_i = n\bar{x}$, simplifies further to:

(4)
$$f_0 / f_1 = \exp\left\{\frac{n(\mu_1-\mu_0)}{\sigma^2}\left(\bar{x} - \frac{\mu_0+\mu_1}{2}\right)\right\}$$

And with the values chosen, this specializes at the boundary point, $\bar{x} = 1.645\sigma/\sqrt{n}$, to

$$f_0/f_0 = \exp(1.645\sqrt{n} - 0.5n)$$

which tends to zero as $n \to \infty$. Even at a modest $n = 30$ one finds:

$$f_0/f_1 = \exp(1.645(\sqrt{30}) - 15) = 0.0025 = 1/400$$

or an LR in favor of the rejected h_0 of roughly 400:1.

Thus, one has a recognizable subset of the critical region, namely outcomes at or near the boundary, which more and more strongly favor the rejected hypothesis. The 5 percent significance level is achieved by a surreptitious averaging, for the critical region is built up by incorporating outcomes that give LR's greater than a critical value, starting with the largest LR and continuing until the size of the test is .05. Those first included give evidence against h_0 stronger than the significance level indicates, but the last few included often favor h_0. Better disguised examples of this phenomenon drawn from actual frequentist practice are given in chapter 9 of Jaynes (1983, especially pp. 182f), a critical comparison of orthodox and Bayesian methods that focuses on actual performance. For other criticisms of NP along these lines, see Fisher (1959, chapter 4), and John Kalbfleisch and D.A. Sprott, both of which repay careful study.

It is clear as well that NP violates the LP. In the example of binomial versus negative binomial given earlier,

Jay's most powerful 5 percent test of $h_0{:}p = \frac{1}{4}$ against $h_1{:}p = \frac{3}{4}$ rejects h_0 when $X \geq 12$ successes occur in the $n = 30$ trials, while May's best 5 percent test rejects h_0 when $n_0 \leq 29$, that is, when the twelfth success occurs on or before the twenty-ninth trial. Hence, they reach opposite conclusions when Jay records twelve successes and May obtains the twelfth success on the thirtieth trial. Notice, too, the outcomes 12 and 13 of Jay's experiment both favor h_0, even though the error probabilities of Jay's test are eminently satisfactory, with $\alpha \leq .05$ and $\beta = .0001$.

In keeping with the LP, it seems perfectly permissible to stop sampling as soon as the accumulated data are deemed sufficiently strong evidence for or against the tested hypothesis. This is, after all, the idea behind Wald's extension of NP theory to sequential tests (see Hodges and Lehmann 1970, §6.10). Could it really make a difference whether one had planned beforehand to stop when the sample proportion of defectives exceeds B or falls below A or decided this on the spur of the moment? To continue sampling till the bitter end in keeping with a preset sample size may place experimental subjects in needless jeopardy or even cause their death (for a chilling real-life example, see Royall 1997, §4.6). Thus, the ongoing debate over optional stopping raises serious ethical, as well as methodological, concerns.

(B) and (L) also permit enlarging a promising study to solidify the evidence, but because this can only increase the type 1 error probabilities, NP disallows it. This further points to the need to separate the presampling design of an experiment from the postsampling analysis of the resulting data.

But what about the fraud who resolves to go on sampling until some targeted null hypothesis is rejected? The reply to this objection to optional stopping is that while such deception is, indeed, possible using standard NP tests, for the power of such a test, as illustrated earlier, approaches one as the sample size increases, the chances of such deception using a likelihood criterion are remote. Using the familiar mathematics of gambler's ruin (de Groot 1986, §2.4), one can show, for example, that the probability of achieving an LR of 32 in favor of a cure rate of 75 percent for a new drug against the 25 percent rate of the drug currently in use, which requires an excess of $s - t \geq 4$ cures over noncures, is given by:

$$\frac{(q/p)^m - 1}{(q/p)^{m+4} - 1} = \frac{3^m - 1}{3^{m+4} - 1}$$

with $q = 1 - p$, which increases rapidly to its limit of 1/81 as $m \to \infty$.

In espousing an evidential interpretation of NP, Egon S. Pearson speaks of "a class of results which makes us more and more inclined . . . to reject the hypothesis tested in favor of alternatives which differ from it by increasing amounts" (1966, p. 173). Deborah G. Mayo, who defends an evidential version of NP, remarks that "one plausible measure of this inclination is the likelihood" (1996, p. 389), but Pearson rejects this on the grounds that "if we accept the criterion suggested by the method of likelihood it is still necessary to determine its sampling distribution in order to control the error involved in rejecting a true hypothesis" (quoted by Mayo 1996, p. 393). What Pearson, Mayo, and others fail to appreciate, however, is the possibility of retaining the law of likelihood while still assessing and controlling beforehand the probability of obtaining misleading results.

If a LR, $L = f_1/f_0$ greater than L^* is accounted strong evidence in favor of h_1 against h_0, then one may compute $P(f_1/f_0 \geq L^*|h_0)$ as readily as one computes $\alpha = P(X \in R|h_0)$, and in place of $\beta = P(X \notin R|\, h_1)$ one may compute $P(f_1/f_0 < L^*|h_1)$, which is the probability of misleading evidence against h_1. (It should be emphasized that it is the evidence itself that is misleading, not one's interpretation of it.)

An important general result, noted independently by C. A. B. Smith and Alan Birnbaum, affirms that the probability of obtaining an LR of at least k in favor of h_0 when h_1 holds is at most $1/k$:

(5) $$P(f_1/f_0 \geq k|h_0) \leq k^{-1}$$

For if S is the subset of outcomes for which the LR is at least k, then

$$P(L \geq k \mid h_0) = \sum_{x \in S} P(x \mid h_0) \leq k^{-1} \sum_{x \in S} P(x \mid h_1) \leq k^{-1}$$

Naturally, this universal bound can be considerably sharpened in special cases, as in the example of a would-be fraud. A specially important case is that of testing hypotheses about a normal mean of known variance with LR given by (4). If the distance $\Delta = |\mu_1 - \mu_0|$ is measured in units of the standard deviation of $\bar{x}$, $\Delta = c\sigma/\sqrt{n}$, one finds:

$$f_1 / f_0 \geq k \Leftrightarrow n\frac{\Delta}{\sigma^2}\left(\bar{x} - \frac{\mu_0 + \mu_1}{2}\right) \geq \ln k$$

$$\Leftrightarrow \bar{x} \geq \frac{\sigma^2}{n\Delta}\ln k + \frac{\mu_0 + \mu_1}{2}$$

$$\Leftrightarrow \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq \frac{\ln k}{c} + \frac{c}{2}$$

whence

$$P(f_1/f_0 \geq k \mid h_0) = 1 - \Phi\left(\frac{\ln k}{c} + \frac{c}{2}\right) = \Phi\left(-\frac{c}{2} - \frac{\ln k}{c}\right)$$

with $\Phi(x)$ the (cumulative) normal distribution. Hence, the probability of misleading evidence in this case is a maximum when $c/2 + ln\ k/c$ is a minimum. By calculus this happens when $c = \sqrt{2lnk}$, in which case $c = c/2 + lnk/c$. Thus,

(6) $$\max P(f_1/f_0 \geq k|h_0) = \Phi(-\sqrt{2lnk})$$

For example, for $k = 8$, $\Phi(-\sqrt{2ln8}) = .021$, while for $k = 32$, $\Phi(-\sqrt{2ln32}) = .0043$, which improve considerably on the universal bounds of 1/8 and 1/32. In fact, the ratio, $\Phi(-\sqrt{2lnk})/k^{-1}$ is easily seen to be decreasing, so that the relative improvement over the universal bound is greater for larger *k*. Royall (1997) greatly extends the reach of (6) by invoking the fact that the log-likelihood is asymptotically normal about its maximum (the ML estimate of the parameter) with precision given by the Fisher information, with an analogous result for the multiparameter case (Lindley 1965, §7.1; Hald 1998, p. 694).

The upshot is that one can retain the law of likelihood and the likelihood principle and still control for the probability of misleading evidence, the feature that lent NP so much of its initial appeal. This "Royall road" opens the way to further reconciliation of (F) with (B) and (L) and to the removal of many perplexing features of NP significance tests (Royall 1997, chapter 5). In retrospect, one sees that the significance level was made to play a dual role in NP theory as both an index of the evidence against null (Fisher's interpretation) and the relative frequency of erroneous rejections of the tested hypothesis. Fisher vigorously rejected the latter interpretation of significance levels and offered a pertinent counterexample (1956, pp. 93–96). He even says, "[T]he infrequency with which, in particular circumstances, decisive evidence is obtained, should not be confused with the force, or cogency, of such evidence" (p. 96).

NP's ban on optional stopping as well as on what Pearson brands "the dangerous practice of basing the choice of test . . . on inspection of the observations" (1966, p. 127) is rooted in a conception of testing as subsumption under a reliable rule. One's particular experiment is viewed as one trial of a repeatable sequence of identical experiments in which the considered hypotheses and a division of the outcomes into those supporting and those not supporting the tested hypothesis are specified in advance (compare Fisher 1956, pp. 81–82, who rejects this formulation in no uncertain terms). Thus, it is considered cheating to publish the error probabilities computed for a *post facto* test as if that test had been predesignated. See Mayo (1996, chapter 9) for numerous statements and illustrations of this stance, especially when she maintains, "Using the computed significance level in post-designated cases . . . conflicts with the intended interpretation and use of significance levels (as error probabilities)" (p. 317). Most textbooks are curiously silent on this issue (see Hodges and Lehmann 1970, chapters 11, 13; de Groot 1986, chapter 8), but Mayo's strictures seem to be widely shared by users of statistical tests. The question is whether a statistician, even an orthodox statistician, can function within the confines of such a strict predesignationism.

From Fisher on, modern statisticians have emphasized the importance of checking the assumptions of one's model, and, of course, these are not the object of one's test. Moreover, the most sensitive test of such common assumptions as independence, normality, or equality of variances, is often suggested by the deviations observed in one's data, thus violating Pearson's proscription. But, ironically, the most telling counterexamples come from the bible of NP theory, Erich Lehmann's classic, *Testing Statistical Hypotheses* (1959, p. 7). In testing a hypothesis about a normal mean of unknown variance, one cannot tell how large a sample is needed for a sharp result until one has estimated the variance. Or, again, if X is uniformly distributed in a unit interval of unknown location, one can stop sampling if the first two observations are (very nearly) a unit distance apart, but if the first n observations all lie within a tiny distance of each other, no more has been learned than the first two observations convey and one must go on sampling. In these workaday examples of Lehmann's, optional stopping is not optional; it is the only option.

Obviously, the issue just raised has strong links to the philosophy of science that holds that "evidence predicted by a hypothesis counts more in its support than evidence that accords with a hypothesis constructed after the fact" (Mayo 1996, p. 251). It would be digressive to enter into this issue here, so one must refer to Mayo (chapter 8) for further discussion and references, and to Stephen G. Brush (1994).

GOODNESS-OF-FIT TESTS

Karl Pearson's goodness-of-fit test (de Groot 1986, §§9.1–4; Hodges and Lehmann 1970, §11.3) rejects a multinomial model h_0 of categorical data when the deviation between observed (n_i) and predicted category counts (np_i) is improbably large conditional on h_0. The measure of deviation employed by Pearson is the chi-squared measure:

(7)
$$X^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = n \sum_{i=1}^{k} \frac{(f_i - p_i)^2}{p_i}$$

with $f_i = n_i/n$. Pearson showed that if h_0 is true, X^2 has, asymptotically, a chi-squared distribution with $v = k - 1$ degrees of freedom. The mean and variance are v and $2v$ and a rule of thumb is that roughly 90 to 95 percent of the probability mass of the chi-squared distribution lies to the left of the mean plus two standard deviations. These and other mathematically convenient features are, essentially, the only thing that recommends this particular measure of deviation (see the two texts just cited and Jaynes 2003, p. 299).

On the surface, Pearson's chi-squared test appears to test the goodness-of-fit of a model without reference to alternatives. (B) offers a less well known test whose rationale is best brought out by considering Jaynes's example of a thick coin (2003, p. 300) that may land on its edge with a probability of .002 and is otherwise balanced (h_0). In $n = 29$ tosses, $D = (n_1, n_2, n_3) = (14, 14, 1)$ is observed, that is, the coin lands on its edge once and lands heads and tails equally often, in an almost "best possible" agreement with h_0. However, $X^2 = 15.33$, which is more than seven standard deviations beyond the mean of 2. Defenders of the test will be quick to point out that the chi-square approximation to the distribution of X^2 breaks down when one or more of the expected counts is less than 5, but that is not the problem here. For one can use brute force to compute $P(X^2 \geq 15.33|h_0)$ exactly, since the only outcomes that give a smaller value of X^2 are $(l, 29 - l, 0)$ and $(29 - l, l, 0)$ with $4 \leq l \leq 14$. The sum of their probabilities on h_0 is 0.9435, whence $P(X^2 \geq 15.33|h_0) = 0.0565$. Hence, Pearson's test just fails by a whisker to reject h_0 at the 5 percent significance level conventionally associated with strong evidence against h_0. The source of the trouble is that X^2 wrongly orders the possible outcomes; some accounted less deviant than (14, 14, 1) are actually less probable on h_0. Ideally, outcomes less probable on h_0 should be accounted more deviant.

Given data $D = (n_1, \ldots, n_k)$, one might ask a somewhat different question than the one Pearson asked, namely: How much support is apt to be gained in passing to some alternative hypothesis? For as Fisher and others emphasize, before rejecting a model as ill fitting one should attempt to find a plausible alternative that fits the data better. Plausibility aside, there is always one alternative hypothesis—call it the tailored hypothesis—that fits D better than h_0 by positing the observed relative frequencies, $f_i = n_i/n$, as its category probabilities. In effect, one wants to test the given model against the ideally best-fitting alternative, and this prompts one to look at the LR in favor of $F = (f_1, \ldots, f_2)$ against the probability distribution $P = (p_1, \ldots, p_2)$ of h_0, namely, $\prod_{i=1}^{k} \frac{f_i^{n_i}}{p_i^{n_i}}$, or, better, at its logarithm, $\sum_{i=1}^{k} n_i \ln(f_i / p_i)$, which is additive in independent samples. This proves to be n times

(8)
$$H(F, P) = \sum_{i=1}^{k} f_i \ln(f_i / p_i)$$

which may be viewed as a measure of the nearness of F to P. Though (8) was used by Alan Turing and his chief statistical assistant, I. J. Good during World War II, Solomon Kullback, another wartime code breaker, was the first to publish a systematic treatment of its properties and applications to statistics, dubbing it discrimination information (see the entry on information theory). Since F is tailored to achieve perfect fit, $H(F, P)$ sets an upper limit to how much one can improve the fit to the data by scrapping h_0 in favor of a simple or composite alternative hypothesis (Jaynes 2003, pp. 293–297).

Happily, $\psi = 2nH(F, P)$ is also asymptotically distributed as χ^2_{k-1}, the chi-square variate with $k - 1 d.f.$ (degrees of freedom). This hints that Pearson's X^2 approximates ψ (Jaynes 1983, pp. 262–263). For example, Mendel's predicted phenotypic ratios of $AB{:}Ab{:}aB{:}ab = 9{:}3{:}3{:}1$ for a hybrid cross, $AaBb \times AaBb$, gave rise to counts of 315, 101, 108, and 32 among $n = 556$ offspring. This gives $X^2 = .4700$ and $\psi = .4754$. But when the expected category counts include a small value or the deviations are large, the approximation degrades, and with it the performance of Pearson's test. Thus, in Jaynes's (2003) thick coin example, X^2 rates the outcomes $(l, 29 - l, 0)$ and $(29 - l, l, 0)$ for $4 \leq l \leq 8$ as less deviant than (14, 14, 1) even though they are also less probable on h_0; by contrast, ψ errs only in failing to count (9, 20, 0) and (20, 9, 0) as less deviant than (14, 14, 1). Hence, the exact probability that ψ is less than its value of 3.84 at (14, 14, 1) is twice the sum of the probabilities (on h_0) of the outcomes $(l, 29 - l, 0)$ for $10 \leq l \leq 14$, or 0.7640, whence $P(\psi \geq 3.84|h_0) = .2360$. Clearly,

the ψ-test gives no reason to believe support can be much increased by passing to an alternative hypothesis, but it will be instructive to carry the analysis a step further.

The only plausible alternative that presents itself is the composite hypothesis, $H{:}p_1 = p_2 = \frac{1}{2}(1 - \theta), p_3 = \theta$ $(0 < \theta < 1)$, which includes h_0 as the special case $\theta = .002$. Since one d.f. is lost for each parameter estimated from the data in using Pearson's test (de Groot 1986, §9.2), this is one way of trading off the improved accuracy that results when a parameter is added against the loss of simplicity. It is insensitive, however, to whatever constraints may govern the parameters. A Bayesian treatment tests h_0 against the composite alternative $H - h_0$ (i.e., H exclusive of the value $\theta = .002$) and goes by averaging the likelihoods of the special cases of $H - h_0$ against a uniform prior of θ over its allowed range—unless more specific knowledge of θ is available. (The affect is to exact a maximum penalty for the given complication of h_0.) On canceling the multinomial coefficient and using the beta integral (*v.s.*), the ratio of the likelihoods reduces to:

$$P(D \mid h_0) : P(D \mid H - h_0) = .499^{28}(.002) : \int_0^1 [\tfrac{1}{2}(1-\theta)]^{28}\theta d\theta$$

$$= 30 \cdot 29 \cdot 2^{28}(.499^{28})(.002) = 1.645$$

Thus, the data $D = (14, 14, 1)$ favors h_0 over the composite alternative, and this remains true, albeit less strongly, if one integrates, say, from 0 to 0.1. By contrast, the chi-square test favors $H - h_0$ over h_0 by mere dint of the fact that the composite hypothesis includes the tailored hypothesis as a special case, namely, $\theta = 1/29$, for then the value of X^2 is zero. Thus, any complication of an original model that happens to include the tailored hypothesis will be preferred to the original model.

Notice, the parameter distribution must reflect only what is known before sampling. Unfortunately, more cannot be said about the different ways (F) and (L) handle the problem of trading off the improved accuracy gained in complicating a model, retaining the original model as a special case, against the loss of simplicity as compared to the Bayesian method just illustrated of averaging the likelihoods. For more on this, see Roger D. Rosenkrantz (1977, chapters 5, 7, and 11) and Arnold Zellner, Hugo A. Keuzenkamp, and Michael McAleer (2001) for other approaches.

PROBABILITY AS LOGIC

Bayesians view probability as the primary (or primitive) concept and induction or inference as derived (see Finetti 1938/1980, p. 194). They emphasize that their methods, properly applied, have never been rejected on the basis of their actual performance (Jaynes 1983, chapter 9; 2003, p. 143). As a corollary, they maintain that the canons of scientific method and inductive reasoning have a Bayesian rationale, while this is vigorously contested by frequentists (e.g., Mayo 1996, chapters 3 and 11). In particular, Bayesians evolved a mathematical analysis of inductive reasoning with its source in the original memoir of Thomas Bayes that includes purported solutions of the notorious problem of induction by Laplace (see Hald 1998, chapter 15) and de Finetti (1937/1981), as well as the equally notorious paradoxes of confirmation (see Good 1983, chapter 11; Rosenkrantz 1977, chapter 2).

Plainly, one's view of statistics is highly colored by one's interpretation of probability. The approaches of Fisher, Neyman, and Pearson, as well as that of most (L) proponents, like Royall, are grounded in a frequency interpretation that equates probabilities with asymptotically stable relative frequencies. The criticisms of the frequency theory, nicely summed up by L. J. Savage (1954, pp. 61–62), are, first, that it is limited (and limiting) in refusing to treat as meaningful the probabilities of singular or historical events, or (in most cases) scientific theories or hypotheses, like the hypothesis that smoking causes lung cancer, and, second, that it is circular. The model of random independent (Bernoulli) trials considered earlier is often held to justify the definition of probability as a limiting relative frequency, but all that theorem does is assign a high probability to the proposition that the observed relative frequency will lie within any preassigned error of the true probability of success in a sufficiently long sequence of such trials.

Savage's criticism along these lines is more subtle. Bayes saw that a distinctly inverse or inductive inference is needed to infer probabilities from observed frequency behavior. Thus, even Bayesians, like Good or Rudolf Carnap, who admit physical probabilities, insist that epistemic probabilities are needed to measure or infer the values of physical probabilities. A more sophisticated view is that physical probabilities arise from the absence of microscopic control over the outcome of one's experiment (see the final section).

Modern Bayesians have sought deeper foundations for probability qua degree of belief and the rules governing it in the bedrock of consistency. It is not merely "common sense reduced to a calculus" (Laplace) but a "logic of consistency" (F. P. Ramsey). Needed, in particular, is a warrant for (2), for it is in Bayesian eyes the basic (not to say the "bayesic") mode of learning from experience. Epistemologists of the naturalist school seriously ques-

tion this, as when Ronald N. Giere contends that "there are many different logically possible ways of 'conditionalizing' on the evidence, and no *a priori* way of singling out one way as uniquely rational" (1985, p. 336). Rather than multiply one's initial odds by the LR, why not by some positive power of the LR? At any rate, this marks a major parting of the ways in contemporary epistemology.

One Bayesian response has been to argue that alternatives to the usual rules of probability open one to sure loss in a betting context, to a so-called Dutch book. However, this justification imports strategic or game theoretic considerations of doubtful relevance, which is why Bruno de Finetti (1972), an early sponsor of the argument, turns, instead, to the concept of a proper scoring rule, a means of evaluating the accuracy of a probabilistic forecast that offers forecasters no incentive to announce degrees of prediction different from their actual degrees of belief. (It is rumored that some weather forecasters overstate the probability of a storm, for example, to guard against blame for leaving the citizenry unwarned and unprepared.) This move to scoring rules opens the way to a means-end justification of (2) as the rule that leaves one, on average, closest to the truth after sampling.

By far the most direct way of sustaining Ramsey's declaration that "the laws of probability are laws of consistency" is that developed by the physicist Richard T. Cox (1946). Besides a minimal requirement of agreement with common sense, his main appeal is to a requirement of consistency (CON), that two ways of doing a calculation permitted by the rules must yield the same result. In particular, one must assign a given proposition the same probability in two equivalent versions of a problem.

In a nutshell, Cox's argument for the product rule, $P(AB|C) = P(A|BC)P(B|C)$, from which (2) is immediate, exploits the associativity of conjunction.

First phase: Letting $AB|C$ denote the plausibility of the conjunction AB supposing that C, show that $AB|C$ depends on (and only on) $A|BC$ and $B|C$, so that

(i) $$AB|C = F(A|BC, B|C)$$

Moreover, by the requirement of agreement with qualitative common sense, the function $F(x, y)$ must be continuous and monotonically increasing in both arguments, x and y.

Second phase: Using first one side then the other of the equivalence of $(AB)D$ and $A(BD)$:

$$ABD|C = F(AB|DC, D|C) = F(F(A|BDC, B|DC), D|C)$$

$$ABD|C = F(A|BDC, BD|C) = F(F(A|BDC, F(B|DC, D|C))$$

leading by (CON) to the associativity functional equation first studied by Niels Henrik Abel in 1826:

(ii) $$F(F(x, y), z) = F(x, F(y, z))$$

Cox solved (ii) by assuming that, in addition, $F(x, y)$ is differentiable. An elementary approach sketched by C. Ray Smith and Gary J. Erickson (1990) based on functional iteration, due to J. Aczel, dispenses with this assumption and leads to the solution: $w(F(x, y)) = w(x)w(y)$, with w continuous and monotonic, hence to

(iii) $$w(AB|C) = w(A|BC)w(B|C)$$

Third phase: Specializing (iii) to the cases where A is certain or impossible given C, one deduces that $w(A|A) = 1$ and $w(A|\bar{A}) = 0$ or ∞. But these two choices lead to equivalent theories, so one may as well assume that $w(A|\bar{A}) = 0$ in line with the usual convention.

Cox (1946) gives a similar derivation of the negation rule. $P(A) + P(\bar{A}) = 1$, and in conjunction with the product rule just derived, this yields the sum rule as follows:

$$\begin{aligned} P(A \vee B \mid C) &= 1 - P(\bar{A}\bar{B} \mid C) \\ &= 1 - P(\bar{A} \mid \bar{B}C)P(\bar{B} \mid C) \\ &= 1 - [1 - P(A \mid \bar{B}C)]P(\bar{B} \mid C) \\ &= P(B \mid C) + P(A \mid \bar{B}C)P(\bar{B} \mid C) \\ &= P(B \mid C) + P(A\bar{B} \mid C) \\ &= P(B \mid C) + P(A \mid C)[1 - P(B \mid AC)] \\ &= P(A \mid C) + P(B \mid C) - P(AB \mid C) \end{aligned}$$

Notice, Cox's derivation is restricted to finite algebras of sets, though not to finite sample spaces.

Non-Bayesian methods (or surrogates) of inference, which *ipso facto* violate one or more of Cox's desiderata, tend to break down in extreme cases. For example, unbiased estimates can yield values of the estimated parameter that are deductively excluded and frequentist confidence intervals can include impossible values of the parameter. A weaker but more general result to account for this affirms that one maximizes one's expected score after sampling (under any proper scoring rule) with (2) in preference to any other inductive rule (Rosenkrantz 1992, p. 535). This optimality theorem, which seems to have many discoverers, affords a purely cognitive justification of (2) as the optimally efficient means to one's cognitive end of making inferences that leave one as close to the truth as possible. This rationale has been extended by inductive logicians to the justification of more specialized

predictive rules that are seen as optimal for universes or populations of specifiable orderliness (see Festa 1993).

An interesting implication of the optimality theorem is that it pays to sample, or that informed forecasts are better than those that lack or waste given information. To see this, compare (2) to the impervious rule that fixes updated probabilities at their initial values. Moreover, since the utility scoring rule, $S(R, h_i) = U(a_R, h_i)$, is proper, where a_R maximizes expected utility against the probability distribution, $R = (r_1, \ldots, r_n)$, over states of nature, one can expect higher utility after sampling as well, a result first given by Good (1983, chapter 17). Thus, both cognitive and utilitarian ends are encompassed.

The optimality theorem presents Bayesian conditioning as the solution of a well-defined optimization problem, thus connecting it to related results on optimal searching and sorting and continuing the tradition of Fisher, Neyman, Pearson, and Wald of viewing rules of estimation, statistical tests, and decision functions (strategies) as solutions of well-posed optimization problems.

THE CONTROVERSIAL STATUS OF PRIOR PROBABILITIES

Objections to (B) center on the alleged impossibility of objectively representing complete ignorance by a uniform probability distribution (Fisher 1956, chapter 2; Mayo 1996, pp. 72ff; Royall 1997, chapter 8). For if one is ignorant of V (volume), then, equally, one is ignorant of $D = 1/V$ (density), but a uniform distribution of V entails a nonuniform distribution of D and vice versa, since equal intervals of V correspond to unequal intervals of D, so it appears one is landed in a contradiction (for some of the tangled history of this charge of noninvariance, see Hald 1998, §15.6; Zabell 1988).

Bayesian subjectivists also deny that any precise meaning can be attached to ignorance (Savage 1954, pp. 64–66), but often avail themselves of uniform priors when the prior information is diffuse (e.g., Lindley 1965, p. 18). This affords a reasonably good approximation to any prior that is relatively flat in the region of high likelihood and not too large outside that region, provided there is such a region (or, in other words, that the evidence is not equally diffuse). For a precise statement, proof, and discussion of this so-called principle of stable estimation, see Ward Edwards, Harold Lindman and Leonard J. Savage (1965, pp. 527–534), as well as Dennis V. Lindley (1965, §5.2) for the important special case of sampling a normal population.

Bayesians have also used Harold Jeffreys's log-uniform prior with density

(9)
$$p(\theta|I_0) \propto \theta^{-1}$$

for a positive variate or parameter, $\theta > 0$, where I_0 represents a diffuse state of prior knowledge. (9) is equivalent to assigning $\ln\theta$ a uniform distribution, whence the name *log-uniform*. If θ is known to lie within finite bounds, $a \le \theta \le b$, the density (9) becomes

(9a)
$$p(\theta)d\theta = \frac{1}{\theta \ln R_\theta}$$

where $R_0 = b/a$, hence, the probability that θ lies in a subinterval, $[c, d]$ of $[a, b]$ is given by:

(9b)
$$P(c \le \theta \le d \mid I_0) = \frac{\ln d - \ln c}{\ln b - \ln a}$$

It follows that θ is log-uniformly distributed in $[a, b]$ if and only if, for any integer k, θ^k is log-uniformly distributed in $[a^k, b^k]$, since

$$\begin{aligned} P(c \le \theta^k \le d \mid I_0) &= P(c^{1/k} \le \theta \le d^{1/k} \mid I_0) \\ &= \frac{\ln d^{1/k} - \ln c^{1/k}}{\ln b - \ln a} \\ &= \frac{\ln d - \ln c}{k[\ln b - \ln a]} \\ &= \frac{\ln d - \ln c}{\ln b^k - \ln a^k} \end{aligned}$$

This at once resolves the objection from the (alleged) arbitrariness of parameterization mentioned at the outset. For V (volume) is a positive quantity, hence, the appropriate prior is, not uniform, but log-uniform, and it satisfies the required invariance: all (positive or negative) powers of V, including V^{-1}, have the same (log-uniform) distribution.

Its invariance would be enough to recommend (9), but Jeffreys provided further justifications (for his interesting derivation of 1932, see Jaynes 2003, p. 498). He did not, however, derive (9) from a basic principle clearly capable of broad generalization (Kendall and Stuart 1967, p. 152). Nevertheless, his insistence that parameters with the same formal properties be assigned the same prior distribution hinted at a *Tieferlegung*. And while the leaders of the Bayesian revival of the 1950s, Savage, Good, and Lindley, did not find in Jeffreys's assorted derivations of (9) a principle definite enough to qualify as a postulate of

rationality, they did clearly believe that given states of partial knowledge are better represented by some priors than by others they denigrated as pig-headed (Lindley 1965, p. 18) or highly opinionated (e.g., Zabell 1988, p. 157). Such out-of-court priors might be highly concentrated in the face of meager information or import a dependence between two parameters (de Groot 1986, p. 405). There matters stood when Jaynes published his fundamental paper, "Prior Probabilities" in 1968 (chapter 7 of Jaynes 1983).

Bayesian subjectivists are as committed to consistency as Bayesian objectivists, and to assign different probabilities to equivalent propositions or to the same proposition in two equivalent formulations of a problem is to commit the most obvious inconsistency. Savage (1954, p. 57), for one, viewed it as unreasonable to not remove an inconsistency, once detected.

Consider a horse race about which one knows only the numbers—better, the labels—of the entries. Since the labels convey no information (or so one is assuming), any relabeling of the horses leads to an equivalent problem, and the only distribution invariant under all permutations of the labels is, of course, the uniform distribution. Thus reinvented as an equivalence principle, Laplace's hoary principle of indifference is given a new lease on life: The vague notion of indifference between events or possibilities gives way to the relatively precise notion of indifference between problems (Jaynes 1983, p. 144). Two versions of a problem that differ only in details left unspecified in the statement of the problem are *ipso facto* equivalent (p. 144). In this restricted form Laplace's principle can be applied to the data or sampling distributions to which (F) and (L) are confined as well as to the prior distributions on which (B) relies. Indeed, from this point of view, "exactly the same principles are needed to assign either sampling distributions or prior probabilities, and one man's sampling probability is another man's prior probability" (Jaynes 2003, p. 89).

Invariance also plays a leading role in frequentist accounts of estimation and testing (Lehmann 1959, chapter 6). In testing a bivariate distribution of shots at a target for central symmetry, Lehmann notes, the test itself should exhibit such symmetry, for if not, "acceptance or rejection will depend on the choice of [one's coordinate] system, which under the assumptions made is quite arbitrary and has no bearing on the problem" (p. 213).

To see how the principle can be used to arrive at a sampling distribution, consider, again, Frank Wilcoxon's statistic, W_t, for the sum of the ranks of the t treated subjects, with W_c the corresponding statistic for the c controls, where $t + c = N$. Clearly, it is a matter of arbitrary convention whether subjects who show a greater response are assigned a higher or lower number as rank. In Table 2, the inverse ranks of the $t = 8$ treated subjects are, respectively, 10, 13, 8, 11, 12, and 7, where each rank and its inverse sum to $N + 1 = 14$. This inversion of the ranks leaves the problem unchanged. On the null hypothesis, h_0, that the treatment is without affect, both W_t and the corresponding statistic, W_t', for the sum of the inverse ranks, are sums of t numbers picked at random from the numbers *1* through N. Hence, W_t and W_t' have the same distribution, which we write as:

$$W_t \simeq W_t'$$

This is the invariance step where the Jaynesian principle of indifference is applied. Furthermore, since $W_t + W_t' = t(N+1)$, it follows that

$$E(W_t) = E(W_t') = t\frac{N+1}{2}$$

whence

$$W_t - t\frac{N+1}{2} = -\left[W_t' - t\frac{N+1}{2}\right] \simeq -\left[W_t - t\frac{N+1}{2}\right]$$

which implies that W_t is symmetrically distributed about its mean. Next, recenter the distribution by subtracting the minimum rank sum of $1 + 2 + \ldots + t = t(t + 1)/2$ from W_t, that is, define:

$$U_t = W_t - t\frac{t+1}{2}$$

and, similarly,

$$U_c = W_c - c\frac{c+1}{2}$$

for the controls. Then both U_t and U_c range from 0 to tc, have mean ½tc, and inherit the symmetry of W_t and W_t' about their mean, which suggests, but does not prove, that $U_t \simeq U_c$. This follows from

$$W_c - c\frac{N+1}{2} \simeq W_t - t\frac{N+1}{2}$$

using

$$W_c - c\frac{N+1}{2} = -\left[W_t - t\frac{N+1}{2}\right]$$

and the symmetry of W_t, while at the same time,

$$W_t - t\frac{N+1}{2} = U_t - \frac{tc}{2}$$

and

$$W_c - c\frac{N+1}{2} = U_c - \frac{tc}{2}$$

so that $U_t - ½tc \simeq U_c - ½tc$, or $U_t \simeq U_c$. Finally, from the common distribution of U_t and U_c, which is easily tabulated for small values of t and c using an obvious recurrence, and for large values using a normal approximation (Hodges and Lehmann 1970, chapter 12, especially p. 349), the distributions of W_t and W_c, with either convention governing the ranks, can be obtained.

Consider, next, Jaynes's (1983, p.126) derivation of the distribution of the rate parameter, λ, of the Poisson distribution (POIS):

$$p(n \mid \lambda, t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

which gives the probability that n events (e.g., accidents, cell divisions, or arrivals of customers) occur in an interval of time of length t. Nothing being said about the time scale, two versions of the problem that differ in their units of time are equivalent. Then the times t and t' in the two versions are related by

(i) $t = qt'$

so that corresponding pairs (λ, t)and (λ', t') satisfy $\lambda t = \lambda' t'$, or

(ii) $\lambda' = q\lambda$

Indeed, (ii) is what defines λ as a scale parameter. Then $d\lambda' = qd\lambda$, that is, corresponding intervals of time also differ by the scale conversion factor. Hence, if $f(\lambda)d\lambda$ and $g(\lambda')d\lambda'$ are the probabilities of lying in corresponding small intervals, $d\lambda$ and $d\lambda'$, then (step 1):

(iii) $f(\lambda)d\lambda = g(\lambda')d\lambda'$

since one is observing the same process in the two time frames, or, using (ii),

(iv) $f(\lambda) = qg(q\lambda)$

Now (step 2) invoke the consistency requirement to affirm that $f = g$, leading to the functional equation

$$f(\lambda) = qf(q\lambda)$$

whose (unique) solution (step 3) is readily seen to be $f(\lambda) \propto 1/\lambda$, the log-uniform distribution of Jeffreys. Thus, if all one knows about a parameter is that it is a scale parameter, then consistency demands that one assigns it a scale-invariant distribution. Following Jaynes (2003, §17.3) , it is instructive to compare the estimates of λ and powers thereof to which the log-uniform distribution leads with the unbiased estimates favored by frequentist theory.

Using the gamma integral,

$$\int_0^\infty \lambda^{k-1}e^{-\lambda}d\lambda = \Gamma(k) = (k-1)!$$

for integers k = 1, 2, 3, . . . , one sees that the rate parameter, λ, is also the mean and variance of POIS. Hence, the mean of any (integer) power of λ after observing n incidents in a chosen unit interval of time (used now in place of an interval of length t) is given by:

$$E(\lambda^k) = \frac{(n+k)!}{n!}$$

In particular, the posterior mean of λ is $n + 1$, that of λ^{-1} is n^{-1}, and that of λ^2 is $(n + 2)(n + 1)$, so that the variance of the posterior distribution is equal to $n + 1$, the same as that of λ, itself a kind of invariance. (F) favors using unbiased statistics (estimators) to estimate a parameter and then among them, choosing the one of minimum variance. That is, on the analogy to target shooting, one uses statistics centered on the bull's eye and most tightly concentrated there (Hodges and Lehmann 1970, chapter 8; de Groot 1986, §7.7). However, as Jaynes shows (2003, §17.3) this "nice property" is not so nice. For while the unbiased estimator of λ is n, which is reasonable and close to its (B) counterpart, the only unbiased estimator $f(n)$ of λ^2 when n is the number of incidents recorded in the unit of time, is

$$f(n) = n(n-1)$$

and $f(n) = 0$ otherwise. Thus, when $n = 1$ incident is observed, the unbiased estimate of λ^2 is zero, which entails that $\lambda = 0$. That is, one is led to an estimate of λ^2 that is deductively excluded by the observation. (It only gets better—or worse!—when one looks at higher powers of λ.) Moreover, no unbiased estimator of λ^{-1} exists. In essence, unbiased estimators are seen to be strongly dependent on which power of the unknown parameter one chooses to estimate, Bayes estimators (equating these with the mean of the posterior distribution) only weakly so.

It is also well known that, for any distribution, the sample variance, $n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$, is a biased estimator of

the population variance, σ^2, while $(n-1)^{-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ is unbiased. If, however, one's goal is to minimize the mean-squared error, $E_\theta[(\hat{\theta}-\theta)^2]$, of one's estimate $\hat{\theta}$ of θ (de Groot 1986, p. 412), the avowed goal of (F), then it can be shown that the biased estimator, $(n+1)^{-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$, of a normal population variance has, for every value of σ^2, a smaller MSE than either of the two cases of the class $c\sum_{i=1}^{n}(x_i-\bar{x})^2$ given earlier (de Groot 1986, pp. 414–415). Hence, the unbiased sample variance is dominated by a biased one; it is, in this precise sense of decision theory, inadmissible. Thus, the two leading (F) criteria of unbiasedness and admissibility are seen to conflict. This insight of Charles Stein's shows, too, that an unbiased estimator is by no means certain to have lower MSE than a biased one, for the MSE is a sum of two terms, the bias and the variance, and in the case at hand, the biased sample variance more than makes up in its smaller variance what it gives up in bias (for more on this, including the waste of information that often accompanies unbiased estimation, see Jaynes 2003, pp. 511ff).

If the density of a variate, X, can be written:

(10)
$$f(x\mid\mu,\sigma)=g\left(\frac{x-\mu}{\sigma}\right)$$

then μ is called a location parameter and σ a scale parameter of the distribution. For changes in μ translate the density curve along the x-axis without changing its shape, while changes in σ alter the shape (or spread) without changing the location. The exemplars are, of course, the mean and standard deviation of a normal distribution. Pretty clearly, Jaynes's derivation of the log-uniform distribution of a Poisson rate applies to any scale parameter (1983, pp. 125–127). That is the justification Jeffreys lacked, though anticipating it in his requirement that formally identical parameters should have the same distribution. The essential point is that not every transformation of a parameter leads to an equivalent problem. Even a subjectivist with no prior information about the population proportion p of some trait would balk at having his or her beliefs represented by a uniform prior of some high power of p.

Notice, the range of $\ln\sigma$ for $0 < \sigma < \infty$ is the whole real line, as is that of a uniform prior of a variate that can assume any real value. Such functions are, of course, nonintegrable (nonnormalizable) and are termed *improper*. They cause no trouble—lead to a normalizable posterior density—when the likelihood function tails off sufficiently fast, as it will when the sample information is non-negligible. In sampling a normal population of known precision, $h = \sigma^{-2}$, a normal prior, $N(\mu_0, h_0)$, of the unknown mean, μ, combines with the normal likelihood based on a random sample of size n to yield a normal posterior density, $N(\mu_1, h_1)$ with precision given by $h_1 = h_0 + nh$, the sum of the prior and the sample precision, and mean:

(11)
$$\mu_1=\frac{h_0\mu_0+nh\bar{x}}{h_0+nh}$$

a precision-weighted average of the prior mean and the sample mean (Lindley 1965, §5.1; Edwards, Lindman, and Savage 1965, pp. 535–538). Small prior precision, h_0, represents a poverty of prior information about the mean, and letting it approach zero yields a uniform prior as a limiting case. Then the posterior mean, μ_1, becomes the sample mean. This is a way of realizing Fisher's ideal of "allowing the data to speak for themselves" and can be applied in the spirit of the "jury principle" when the experimenter is privy to prior information not widely shared by the relevant research community. Priors that achieve this neutrality are termed *uninformative* or *reference priors* (see Loredo 1990, p. 119).

This example of closure—a normal prior combining with the (normal) likelihood to yield a normal postsampling distribution—is prototypic and one speaks of the relevant distribution as conjugate to the given likelihood function or data distribution. Other examples (de Groot 1986, pp. 321–327) include the beta:

$$f_\beta(p|a,b)dp = B(a,b)^{-1}p^{a-1}(1-p)^{b-1}$$

with $B(a,b)=\int_0^1 p^{a-1}(1-p)^{b-1}dp=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and $\Gamma(n) = (n-1)!$ when n is an integer, which combines with a binomial likelihood, $L(p) = p^r(1-p)^s$, to yield a beta posterior density, $f_\beta(p|a+r, b+s)$; or, again, the gamma distribution with density

$$f_\gamma(\lambda\mid a,b)d\lambda=\frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda}$$

which combines with a Poisson likelihood to yield a gamma posterior density (de Groot 1986, p. 323). In general, any (one-parameter) data distribution of the form:

(12) $$f(x|\theta) = F(x)G(\theta)\exp[u(x)\phi(\theta)]$$

will combine with a prior of the form, $p(\theta|I)d\theta \propto G(\theta)^a \exp(b\phi(\theta))$, to yield a density of the same so-called Koop-

man-Darmois form (Lindley 1965, p. 55). These are precisely the data distributions that admit a fixed set of sufficient statistics, namely, estimators of the unknown parameter(s) that yield the same posterior distribution as the raw data (Lindley 1965, §5.5; de Groot 1986, §6.7; or, for more advanced readers, Jaynes 2003, chapter 8).

The parameters of a conjugate prior represent a quantity of information. For example, for the beta prior, $a + b$ may be the size of a pilot sample or a virtual sample. By letting these parameters approach zero, one obtains an uninformed prior in the limit that represents, so to speak, the empty state of prior knowledge. The log-uniform prior (9) of a normal variance can be obtained in this way from the conjugate chi-squared prior (Lindley 1965, §5.3, p. 32), thus complementing its derivation as the distribution of a scale parameter about which nothing else is assumed.

In all the cases considered, the improper prior arises as a well-defined limit of proper priors. When this finite sets policy, which Jaynes traces to Gauss, is violated, paradoxes result, that is, in Jaynesian parlance, "errors so pervasive as to become institutionalized" (2003, p. 485). Such paradoxes can be manufactured at will in accordance with the following prescription:

(1) Start with a mathematically well-defined problem involving a finite set, a discrete or a normalizable distribution, where the correct solution is not in doubt;

(2) Pass to a limit without specifying how the limit is approached;

(3) Ask a question whose answer depends on how that limit is approached.

Jaynes adds that "as long as we look only at the limit, and not the limiting process, the source of the error is concealed from view" (p. 485).

Jaynes launches his deep-probing analysis of these paradoxes with the following exemplar, a proof that an infinite series, $S = \sum a_n$, converges to any real number x one cares to name. Denoting the partial sums, $s_n = a_1 + a_2 + \ldots + a_n$ with $s_0 = 0$, one has for $n \geq 1$:

$$a_n = (s_n - x) - (s_{n-1} - x)$$

and so the series becomes

$$S = (s_1 - x) + (s_2 - x) + (s_3 - x) + \ldots$$

$$-(s_0 - s) - (s_1 - x) - (s_2 - x) - \ldots$$

Since the terms $s_1 - x$, $s_2 - x$, . . all cancel out, one arrives at $S = -(s_0 - x) = x$.

Apart from assuming convergence, the fallacy here lies in treating the series as if it were a finite sum. The nonconglomerability paradox, which purports to show that the average, $P(A|I)$, of a bounded infinite set of conditional probabilities, $P(A|C_jI)$, can lie outside those bounds, also turns on the misguided attempt to assign these probabilities directly on an infinite matrix rather than approaching them as well-defined limits of the same probabilities on finite submatrices (Jaynes 2003, §15.3). Jaynes goes on to consider countable additivity, the Borel-Kolmogorov paradox, which involves conditioning on a set of measure zero, and the marginalization paradoxes aimed at discrediting improper priors. These paradoxes have little to do with prior probabilities *per se* and everything to do with ambiguities in the foundations of continuous probability theory.

Leaving these subtle fallacies to one side, one can apply Jaynes's policy of starting with finite sets and then passing to well-defined limits to another old chestnut, the water-and-wine paradox in which one is told only that the ratio of water (H) to wine (W) in a mixture lies between 1 and 2. Then the inverse ratio of wine to water lies between ½ and 1, and, in the usual way, a uniform distribution of one ratio induces a nonuniform distribution of the other. One can eliminate ambiguity, however, by quantizing the problem. There are, after all, just a finite number N of molecules of liquid, of which N_H are water molecules and N_W are wine molecules. Then the inequality, $1 \leq N_H{:}N_W \leq 2$, is equivalent to $N_W \leq N_H \leq 2N_W$, and so the admissible pairs $(N_H, N_W)$are:

$$\{(N_H, N - N_H){:}\tfrac{1}{2}N \leq N_H \leq \tfrac{2}{3} N\}$$

Moreover, this remains true when one starts with the other (equivalent) version of the problem in which the given is the inequality, $\tfrac{1}{2} \leq N_W{:}N_H \leq 1$, governing the inverse ratio. One then assigns equal probabilities to these $(\tfrac{2}{3} - \tfrac{1}{2})N = \tfrac{1}{6}N$ allowed pairs. Then to find, for example, the probability that $\tfrac{1}{2} \leq N_W{:}N_H \leq \tfrac{3}{4}$, one takes the ratio of the allowed pairs meeting this condition, which is equivalent to

$$\tfrac{4}{7}N \leq N_H \leq \tfrac{2}{3}N$$

to the total number, $N/6$, of allowed pairs to find, not 1/2, but

$$\left(\frac{2}{3} - \frac{4}{7}\right) / \frac{1}{6} = \frac{2/21}{1/6} = \frac{4}{7}$$

which is surprisingly close to:

$$\frac{\ln\frac{3}{4}-\ln\frac{1}{2}}{\ln 1-\ln\frac{1}{2}}=\frac{\ln 3-\ln 2}{\ln 2}=.58496$$

Or, again, the probability that N_W:N_H lies between ⅝ and $^{11}/_{13}$ is found to be 23/52 = .442, which is close to ($\ln^{11}/_{13}$ – ln⅝)/ln2 = .437. Thus, by assigning equal probabilities in a discrete version of the problem—the only invariant assignment—one appears to be led once more to the log-uniform prior.

Another familiar puzzle of geometric probability is Joseph Bertrand's chord paradox, which asks for the probability that a chord of a circle of radius *R* drawn at random exceeds the side $s = \sqrt{3}R$ of the inscribed equilateral triangle. Depending on how one defines "drawn at random," different answers result, and Bertrand himself seems to have attached no deeper significance to the example than that "la question est mal posee."

Like the water-and-wine example, this puzzle is more redolent of the faculty lounge than the laboratory, so following Jaynes (1983, chapter 8; 2003, §12.4.4), one can connect it to the real world by giving it a physical embodiment in which broom straws are dropped onto a circular target from a great enough height to preclude skill. Nothing being said about the exact size or location of the target circle, the implied translation and scale invariance uniquely determine a density:

(13)
$$f(r,\theta)=\frac{1}{2\pi rR}$$

for the center (r, θ) of the chord in polar coordinates. And since

$$\int_{r_1}^{r_2} f(r,\theta)dr=(2\pi R)^{-1}\ln(r_2/r_1)$$

it follows that annuli whose inner and outer radii, r_1 and r_2, stand in the same ratio should experience the same frequency of hits by the center of a chord. With $L = 2\sqrt{R^2-r^2}$ the length of a chord whose center is at (r, θ), the relative length, $x = L/2R$, of a chord has the induced density:

(13a)
$$p(x)dx=\frac{x}{\sqrt{1-x^2}}$$

Finally, since $L = \sqrt{3}R$ is the side-length of the inscribed equilateral triangle, the probability sought is:

$$\int_{\sqrt{3}/2}^{1} p(x)dx=\tfrac{1}{2}\int_{0}^{1/4} u^{-1/2}du=\frac{1}{2}$$

with $u = 1 - x^2$.

All these predictions of Jaynes's solution can be put to the test (for one such test and its outcome, see Jaynes 1983, p. 143). In particular, (13) tells one to which hypothesis space a uniform distribution should be assigned to get an empirically correct result, namely, to the linear distance between the centers of the chord and circle. There is no claim, however, to be able to derive empirically correct distributions *a priori*, much less to conjure them out of ignorance. All that has been shown is that any distribution other than (13) must violate one or more of the posited invariances. If, for example, the target circle is slightly displaced in the grid of straight lines defined by a rain of straws, then the proportion of hits predicted by that other distribution will be different for the two circles. However if, Jaynes argues (p. 142), the straws are tossed in a manner that precludes even the skill needed to make them fall across the circle, then, surely, the thrower will lack the microscopic control needed to produce a different distribution on two circles that differ just slightly in size or location.

The broom straw experiment, which readers are urged to repeat for themselves, is highly typical of those to which one is tempted to ascribe physical probabilities or objective chances, for example, the chance of 1/2 that the chord fixed by a straw that falls across the circle exceeds the side of the inscribed triangle. However, as Zabell (1988, pp. 156–157) asks, if there is a "propensity" or "dispositional property" present, of what is it a property? Surely not of the straws, nor, he argues, of the manner in which they are tossed. A skilled practitioner of these arts can make a coin or a die show a predominance of heads or sixes (see Jaynes 2003, chapter 10). Nor is it at all helpful to speak of identical trials of the experiment, for if truly identical, they will yield the same result every time. Zabell concludes that "the suggested chance setup is in fact nothing other than a sequence of objectively differing trials which we are subjectively unable to distinguish between." However, one may well be able to distinguish between different throws of a dart in terms of how tightly one gripped it, for example, without being able to produce different distributions on slightly differing targets. It is the absence of such skill that seems to matter, and that feature of the chance setup is objective. On this basis, Jaynes is led to characterize the resulting invariant distribution as "by far the most likely to be observed experimentally in the sense that it requires by far the least skill" (1983, p. 133).

For a different example, consider the law of first digits. Naive application of the principle of indifference at the level of events leads to an assignment of equal proba-

bilities to the hypotheses, h_d, $d = 1, 2, \ldots, 9$, that d is the first significant digit of an entry, X, in a table of numerical data. Nothing being said about the scale units employed, the implied scale invariance implies a log-uniform distribution of X with normalization constant, 1/ln10, since $a = 10^k \leq X < 10^{k+1} = b$ forces $1 = \int_a^b \frac{dx}{x} = \ln(b/a) = \ln 10$ (which is independent of k). Hence, d is the first significant digit with probability:

(14)

$$p_d = \int_{da}^{(d+1)a} \frac{dx}{x \ln 10} = \frac{1}{\ln 10} \ln(1 + d^{-1}) = \log_{10}(1 + d^{-1})$$

so that $p_1 = log_{10}\ 2 = .301, \ldots, p_9 = 1 - log_{10}\ 9 = .046$. Known earlier to Simon Newcomb, (14) was rediscovered in 1938, though not explained, by Frank Benford, who tested it against twenty tables ranging from the surface areas of rivers and lakes to the specific heats of thousands of compounds. Surprisingly, Benford found that (14) even applies to populations of towns or to street addresses, which are certainly not ratio scaled. The explanation lies in the recent discovery of T. P. Hill (1995) that "base invariance implies Benford's law." That is, (14) is invariant under any change of the base $b > 1$ of the number system. Moreover, since scale invariance implies base invariance—but not conversely—the scale-invariant tables for which (14) holds are a proper subset of the base-invariant ones. Indeed, Hill derives a more general form of (14) that applies to initial blocks of $k \geq 1$ digits of real numbers expressed in any base, namely:

(14a)

$$P\left(\bigcap_{i=1}^{k}\{D_b^{(i)} = d_i\}\right) = \log_b\left[1 + \left(\sum_{i=1}^{k} b^{k-i} d_i\right)^{-1}\right]$$

where $D_b^{(i)}(x)$ is the ith significant digit of x in base b. For example, for base ten, and $k = 2$, the probability that the first two digits are 3 and 7 is $\log_{10}[1 + (37)^{-1}] = .01158$, while, as one may verify, the probability that the second digit is d is given by:

(14b)

$$\sum_{k=1}^{9} p_{kd} = \sum_{k=1}^{9} \log_{10}[1 + (10k + d)^{-1}]$$

Hill's derivation of (14) is a beautiful and instructive exercise in measure theoretic probability, but the main point to register here is that (14) is not the chance distribution of any readily conceivable physical process or random experiment. One can be just as certain, though, that any list or table of numbers that violates (14) must yield different frequencies of first (second, . . .) digits when the scale or number system is changed. More generally, the output of a deterministic process, like that which generates the digits of π or random numbers, for that matter, can be as random as one likes under the most stringent criterion or definition of randomness. These categories, so commonly contrasted, are not mutually exclusive. However, it is far from clear how to characterize an intrinsically random physical process in a way that is free of circularity and amenable to experimental confirmation (Jaynes 2003, §10.5). Jaynes views such random processes as mythic products of what he labels the "Mind Projection Fallacy."

BAYES EQUIVALENCE

Part of the motivation of the frequency theory was to develop objective means of assessing the evidence from an experiment, leaving readers of the report free to supply their own priors or utility functions. However, this ideal of separating evidence from opinion is unrealizable because, first, the support of a composite hypothesis or model with adjustable parameters depends on the weights assigned its various simple components, and, second, because of the presence of so-called nuisance parameters.

For an example of the former (Royall 1997, pp. 18–19), one can compare the hypothesis (H) that the proportion p of red balls in an urn is either ¼ or ¾ with the simple hypothesis (k) that $p = ½$, given that a ball drawn at random is red. The bearing of this outcome is wholly dependent on the relative weights assigned the two simple components of H, namely, $p = ¼$ and $p = ¾$. Or, again, how does drawing an ace of clubs bear on the hypothesis that the deck is a trick deck (fifty-two copies of the same card) versus the hypothesis that it is a normal deck? If one's intuition is that a single card can tell one nothing, then one is implicitly assigning equal probabilities to all fifty-two components of the trick deck hypothesis, but if, for example, one has information that most trick decks are composed of aces or picture cards, then drawing that ace of clubs will favor (for one) the trick deck hypothesis by a factor ranging from 1 to 52.

In practice, (F) resorts to comparing two models by the ratios of their maximum likelihoods, as in the orthodox t-test for comparing two normal means (de Groot 1986, §8.6). This is often a good approximation to the Bayes factor, the ratio of average likelihoods, when the two models are of roughly equal simplicity (Rosenkrantz 1977, p. 99), but this practice is otherwise highly biased (in the colloquial sense) in favor of the more complicated hypothesis, as in the trick deck example.

An equally formidable bar to the separation of sample information and prior information is the presence of parameters other than the one of interest. In testing the equality of two normal means, the difference, $\bar{x}-\bar{y}$, of the sample means *means* different things depending on one's beliefs about the variances of the two populations. An even simpler example is random sampling of an urn of size N without replacement. If interest centers on the number R of red balls in the urn and N is also unknown, then an outcome, $D = (n, r)$, of r red in a sample of n, will mean different things depending on one's prior beliefs about the relationship, if any, between R and N. If, for example, extensive previous experience renders it almost certain that the incidence of a certain birth defect lies well below one in a thousand, then a sample of modest size in which such a defect occurs, for example, $(n, r) = (500, 1)$, tells one not merely that $N \geq 500$, the sample size, but (almost surely) that $N \geq 1{,}000$. Even so simple a problem as this appears to lie entirely beyond the scope of (F) or (L), but as Jaynes amply demonstrates, this shopworn topic of introductory probability-statistics courses takes on a rich new life when the inverse problem of basing inferences about N and R on observed samples is considered and different kinds of prior information are incorporated in the resulting data analysis (2003, chapters 3 and 6).

In general, (B) handles nuisance parameters by marginalization, that is, by finding the joint posterior density, say, $p(\theta_1, \theta_2|DI)$ for the case of two parameters, and then integrating with respect to θ_2:

$$p(\theta_1 \mid DI) = \int p(\theta_1, \theta_2 \mid DI)\, d\theta_2$$

the discrete analogue being $P(A|DI) = \sum P(AB_i|DI)$ for mutually exclusive and exhaustive B_i's. Thus, intuition expects that a more focused belief state will result when there is prior knowledge of θ_2 than when its value is completely unknown before sampling.

Consider the case of sampling a normal population when nothing is known about $\theta_1 = \mu$ *and* $\theta_2 = \sigma^2$, so that their joint prior is the Jeffreys prior:

$$p(\theta_1, \theta_2|I) = p(\theta_1|I)p(\theta_2|I) = \theta_2^{-1}$$

while the (normal) likelihood is

$$L(\theta_1,\theta_2) = (2\pi\theta_2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\theta_2}\sum (x_i-\theta_1)^2\right]$$

$$\propto \theta_2^{-\frac{n}{2}} \exp[-\{\nu s^2 + n(\bar{x}-\theta_1)^2\}/2\theta_2]$$

using the obvious identity, $\sum(x_i - \theta_1)^2 = \sum(x_i - \bar{x} + \bar{x} - \theta_1)^2 = \sum(x_i - \bar{x})^2 + n(\bar{x} - \theta_1)^2$, with $s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/\nu$ and $\nu = n - 1$. Multiplying this by the prior yields the joint postsampling density

$$p(\theta_1,\theta_2 \mid DI) \propto \theta_2^{-\left(\frac{n-2}{2}\right)} \exp[-\{\nu s^2 + n(\bar{x}-\theta_1)^2\}/2\theta_2]$$

up to a normalization constant. Then using

(*)
$$\int_0^\infty e^{-A/\theta}\theta^{-u}\,d\theta = \frac{\Gamma(u-1)}{A^{u-1}}$$

obtained from $\Gamma(u) = \int_0^\infty x^{u-1}e^{-x}dx$ by the substitution, $x = A/\theta$, the marginal posterior density of θ_1 is:

$$p(\theta_1 \mid DI) \propto \int_0^\infty \theta_2^{-\left(\frac{n+2}{2}\right)} \exp[-\{\nu s^2 + n(\bar{x}-\theta_1)^2\}/2\theta_2]\,d\theta_2$$

$$\propto \{\nu s^2 + n(\bar{x}-\theta_1)^2\}^{-\nu/2}$$

using (*) with $A = \nu s^2 + n(\bar{x} - \theta_1)^2$ and $u = (n + 2)/2$, whence

$$p(\theta_1|DI) \propto (1 + t^2/\nu)^{\frac{1}{2}(\nu+1)}$$

with $t = n^{\frac{1}{2}}(x - \theta_1)/s$. To find the normalization constant, one integrates on the right using the substitution, $t = \nu^{1/2}\left[\frac{x}{1-x}\right]^{1/2}$, with $dt = \frac{1}{2}\nu^{\frac{1}{2}}x^{-\frac{1}{2}}(1 - x)^{-\frac{3}{2}}dx$ and $1+\frac{t^2}{\nu} = \frac{1}{1-x}$ to obtain:

$$2\int_0^\infty (1+t^2/\nu)^{-\frac{1}{2}(\nu+1)}dt = \nu^{1/2}\int_0^1 x^{-\frac{1}{2}}(1-x)^{\frac{\nu}{2}-1}dx$$

$$= \nu^{\frac{1}{2}}\Gamma(\tfrac{1}{2})\Gamma(\tfrac{\nu}{2})/\Gamma(\tfrac{\nu+1}{2})$$

using the beta integral,

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \Gamma(a)\Gamma(b)/\Gamma(a+b).$$

Hence, the posterior density of the mean, using $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, is given by:

(15)
$$p(\theta_1 \mid DI) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})}\frac{1}{\left(1+\frac{t^2}{\nu}\right)^{\frac{1}{2}(\nu+1)}}$$

which is the density of the t-distribution with $\nu = n - 1$ degrees of freedom.

Thus, one has arrived in a few lines of routine calculation at the posterior (marginal) density of the mean when the variance is (completely) unknown. Were the variance known, the uniform prior of the mean leads, as was seen earlier, to a normal posterior distribution about the sample mean, x, with variance σ^2/n, or in symbols:

$$n^{\frac{1}{2}}(\theta_1 - \bar{x})\sigma \sim N(0, 1)$$

while the result, $n^{\frac{1}{2}}(\theta 1 - \bar{x})/s$, of replacing the population s.d., s, by the sample s.d., s when the former is unknown, has the t-distribution with $\nu = n - 1$ d.f. The density (15) is, like the normal density, bell-shaped and symmetric, but has larger tails (i.e., does not approach its asymptote, the x-axis, as rapidly as the normal curve) and is thus less concentrated. For example, for $\nu = 10$ degrees of freedom (d.f.), the 95 percent central region of the t-distribution is (–2.228, 2.228) while that of the normal is (–1.96, 1.96) ≈ (–2, 2). Thus, Bayesian updating confirms one's intuition that the postsampling belief function should be less concentrated when the variance is (completely) unknown than when it is known. Moreover, the t-distribution approaches normality rather rapidly as $\nu \to \infty$, and so the difference in the two states of prior knowledge is quickly swamped by a large sample. Already at $\nu = 20$, the 95 percent central region of (15) is (–1.98, 1.98), which is almost indistinguishable from the normal.

Because of a mathematical quirk, (F) interval estimates (confidence intervals) for a normal mean with variance unknown are numerically indistinguishable from (B) interval estimates (credence intervals) obtained from the posterior density, although their interpretation is radically different. For a normal distribution, the sample mean, $\bar{x}$, and sample variance, $s_n^2 = n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$, are independent (for a proof, see de Groot 1986, §7.3). As the normal distribution is the only one for which this independence of sample mean and sample variance obtains, it may justly be called a quirk. One shows, next, that if $Y \sim N(0, 1)$ and $Z \sim \chi_n^2$, then

$$X = \frac{Y}{\left(\frac{Z}{n}\right)^{1/2}}$$

has the t-distribution (15) with n degrees of freedom (§7.4). Now $Y = n^{\frac{1}{2}}(\bar{x} - \mu)/\sigma \sim N(0, 1)$ is standard normal, and it can be shown (de Groot 1986, pp. 391–392) that $Z = s_n^2 / \sigma^2 \sim \chi_{(n-1)}^2$, hence

$$U = \frac{Y}{\left(\frac{Z}{n-1}\right)^{1/2}}$$

has the t-distribution with $n - 1$ d.f. The crucial point is that σ^2 cancels out when one divides Y by $Z^{\frac{1}{2}}$ and so the distribution of U does not depend on the unknown variance. The nuisance is literally eliminated. Finally, (F) estimates of μ can be obtained from the distribution of U since $-c \le U \le c$ just in case $\bar{x} - c\sigma'/\sqrt{n} \le \mu \le \bar{x} + c\sigma'/\sqrt{n}$, writing $\sigma' = \left(\frac{s_n^2}{n-1}\right)^{1/2}$. Notice, however, the different interpretation. One thinks of $(\bar{x} - c\sigma', \bar{x} + c\sigma')$ as a random interval that contains μ with the specified probability, or long-run relative frequency in an imagined sequence of repetitions of the experiment.

The first thing that strikes one is how much more complicated this derivation of the sampling distribution of the relevant statistic is than the (B) derivation of the postsampling distribution (15) of μ. Even the modern streamlined derivation given in de Groot's (1986) text occupies nearly ten pages. William Seeley Gossett guessed the distribution by an inspired piece of mathematical detective work (for some of the relevant history, see Hald 1998, §27.5). The first rigorous proof was given in 1912 by a bright Cambridge undergraduate named R. A. Fisher. Gossett began his 1908 paper by noting that earlier statisticians had simply assumed "a normal distribution about the mean of the sample with standard deviation equal to $s/\sqrt{n}$" but that for smaller and smaller samples "the value of the s.d. found from the sample . . . becomes itself subject to an increasing error, until judgments reached in this way may become altogether misleading" (Hald 1998, p. 665). Fisher never tired of extolling "Student" (Gossett's pen name) for his great discovery, as well he might, for it is safe to say that without it, the (F) approach to statistics would never have gotten off the ground. For (F) would not then have been able to address the inferential problems associated with sampling a normal population for the vital case of small samples of unknown precision.

In essence, the pre-Gossett practice of replacing σ in $n^{\frac{1}{2}}(\bar{x} - \mu)/\sigma$ by its ML estimate, s_n, would be about the only option open to (F) or (L) if this nuisance parameter could not be eliminated. However, that is to treat the unknown parameter as if it were known to be equal to its estimated value—precisely what Gossett's predecessors had done. Complete ignorance of σ should result in a fuzzier belief state than when it is known (compare Roy-

all 1997, p. 158). The only remedy (L) offers (p. 158) when nuisance parameters really are a nuisance (and cannot be eliminated) is to use the maximum of the likelihood function taken over all possible values of the relevant nuisance parameter(s). However, this is to equate a model with its best-fitting special case, which is to favor the more complicated of two models being compared. Moreover, in the real world outside of textbooks, normal samples do not come earmarked "variance known" or "variance unknown."

To borrow "Example 1" from Jaynes (1983, p. 157), the mean life of nine units supplied by manufacturer A is 42 hours with s.d. 7.48, while that of four units supplied by B is 50 hours with s.d. 6.48. (F) proceeds in such cases to test the null hypothesis that the two s.d.'s are equal using the F-test originated by Fisher (de Groot 1986, §8.8). When the null hypothesis is accepted, (F) then treats the two s.d.'s as equal and proceeds to a two sample t-test of the equality of the means (de Groot 1986, §8.9), which is predicated on the equality of the two (unknown) variances. In the present example, the hypothesis of equal s.d.'s is accepted at the 5 percent significance level, but then the two-sample t-test (unaccountably) accepts the hypothesis that the two means are equal at a 10 percent level. Jaynes calculates odds of 11.5 to 1 that B's components have a greater mean life, and without assuming equality of the variances. Then he asks, "Which statistician would you hire?"

The (F) solution extends to the case where independent samples are drawn from two normal populations of unknown variance, provided the variances are known to be equal or to stand in a given ratio. However, when the variances are known (or assumed) to be unequal, (F) fragments into a number of competing solutions with no general agreement as to which is best (Lindley 1965, pp. 94–95; Kendall and Stuart 1967, pp. 139ff). W.-U. Behrens proposed a solution in 1929 that Fisher rederived a few years later using his highly controversial fiducial argument (see the entry on R. A. Fisher). As Harold Jeffreys points out (1939, p. 115), the Behrens solution follows in a few lines from (2) using the Jeffreys prior for the unknown parameters (also see Lindley 1965, §6.3).

However, what of the intermediate cases where the variances are not known to be equal (the two-sample problem) and not known to be unequal (the Behrens-Fisher problem). In his definitive treatment, G. Larry Bretthorst (1993) takes up three problems: (1) determine if the two samples are from the same normal population, (2) if not, find how they differ, and (3) estimate the magnitude of the difference. Thus, if they differ, is it the mean or the variance (or both)?

Consider, once more, Jaynes's example. Bretthorst finds a probability of 0.58 that the s.d.'s, σ_1 and σ_2, are the same, given the sample s.d.'s of 7.48 and 6.48, the inconclusive verdict intuition expects. (F) is limited to an unnuanced approach where one or the other of these alternatives must be assumed. The posterior distribution Bretthorst computes is a weighted average of those premised on equal and unequal population variances and thus lies between them (Bretthorst 1993, p. 190, and figure 1). By marginalization, it yields a 72 percent probability that the parent means are different. The analysis is based on independent uniform and log-uniform priors for the means and variances truncated, respectively, at 34 = 46 – 12 and 58 = 46 + 12, and at $\sigma_L = 3$ and $\sigma_H = 10$. This is not just for the sake of greater realism but to ensure that the posterior density is normalizable (p. 191). Doubling the range of the means lowers the probability that the parent populations differ from 0.83 to 0.765, a change of roughly eight percent, while doubling the range of the s.d.'s makes about a 2 percent difference. Hence, the inference appears to be reasonably robust. Finally, the Bayesian solution smoothly extends the partial solutions (F) offers when the variances are unknown; the (F) solutions appear as the limiting cases of the (B) solution when the probability that the variances are equal is either zero or one. This makes it hard for an (F) theorist to reject the Bayesian solution.

The (F) solutions also correspond to (B) solutions based on an uninformative prior. This Bayes equivalence of (F) interval estimates or tests is more widespread than one might suppose (Jaynes 1983, pp. 168–171, 175), but is by no means universal. Generalizing from the case of known variances, it would seem to hold when sufficient estimators of the parameter(s) of interest exist, no nuisance parameters are present, and prior knowledge is vague or insubstantial.

Confidence intervals for a binomial success rate, θ, are harder to construct than the CI's for a normal mean because here the population variance, $n\theta(1-\theta)$, depends on the parameter being estimated. The solution is to find for each value of θ, values $p_L(\theta)$ and $p_H(\theta)$, such that

$$P(p \geq p_L|\theta) = ½(1-\alpha)$$

and

$$P(p \leq p_H|\theta) = ½(1-\alpha)$$

as nearly as possible, where p is the proportion of successes in n trials. In other words, one finds a direct> 100(1

– α)% confidence interval for p for each value of the unknown success rate θ. Then the corresponding CI for θ comprises all those values whose direct CI contains the observed proportion p. For an example ($n = 20$) and a chart, see Kendall and Stuart (1967, pp. 103–105), whose obscure exposition makes this rather convoluted method seem even more mysterious. Plainly, finding such CI's is an undertaking, involving round off errors and approximations. By contrast, the Bayesian posterior density, given in the original memoir of Bayes, based on the uniform prior of θ, for r successes in n trials is

$$p(\theta \mid n, r) = \frac{(n+1)!}{r!(n-r)!}\theta^r(1-\theta)^{n-r}$$

with mean $(r + 1)/(n + 2)$ and variance, $f(1 - f)/(n + 3)$, where $f = r/n$. Hence, the Bayesian credence intervals assume the simple form,

$$f \pm k\sqrt{\frac{f(1-f)}{n+3}}$$

where k is 1.645, 1.96, and 2.57 for the 90, 95, and 99 percent intervals (using the normal approximation to the beta distribution). Jaynes finds (1983, p. 171) these Bayesian intervals are numerically indistinguishable from the CI's of the same confidence coefficient, leading him to wryly observe that the Bayesian solution Fisher denigrated as "founded on an error" delivers exactly the same interval estimates as the (F) solution at a fraction of the computational and mathematical effort. The reason for the equivalence is that, despite its great difference in motivation and interpretation, the (F) method of confidence intervals is based in this case on a sufficient statistic, the observed relative frequency f of success.

As Jaynes notes, the official doctrine of (F) is that CI's need not be based on sufficient statistics (Kendall and Stuart 1967, p. 153), and, indeed, the advertised confidence coefficient is valid regardless. Bayesian credence intervals, being based on the likelihood function, automatically take into account all the relevant information contained in the data, whether or not sufficient statistics exist. Thus, (F) methods not based on a sufficient statistic must perforce be wasting information, and the result one expects, given the optimality theorem, is a degradation of performance. The point is that the data may contain additional information that leads one to recognize that the advertised confidence coefficient is invalid (Loredo 1990, p. 117). The next several examples illustrate this and related points in rather striking fashion.

For a simple example (de Groot 1986, p. 400), let independent observations X_1 and X_2 be taken from a uniform distribution on the interval, $(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, with θ unknown. Then if $Y_1 = min(X_1, X_2)$ and $Y_2 = max(X_1, X_2)$, we have:

$$P(Y_1 \leq \theta \leq Y_2) = P(X_1 \leq \theta)P(X_2 \geq \theta) + P(X_2 \leq \theta)P(X_1 \geq \theta)$$
$$= \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{2}$$

Thus, if $Y_1 = y_1$ and $Y_2 = y_2$ is observed, (y_1, y_2) is a 50 percent CI for θ. However, what if $y_2 - y_1 \approx 1$? Then (y_1, y_2) is virtually certain to contain θ; indeed, one easily checks that it is certain to contain θ when $y_2 - y_1 \geq 1/2$. Thus, one has a recognizable subset of the outcome space on which the 50 percent confidence coefficient is misleadingly conservative.

For an example of the opposite kind, where confidence is misplaced, one can turn to "Example 5" of Jaynes (1983, p. 172f). A chemical inhibitor that protects against failures wears off after an unknown time θ and decay is exponential (with mean one) beyond that point, so that a failure occurs at a time x with probability

$$f(x|\theta) = \exp(\theta - x)h(x, \theta)$$

where $h(x, \theta) = 1$ if $\theta < x$ and is otherwise zero. Since this data distribution for n failure times factors as

$$f_n(x_1, \ldots, x_n|\theta) = \exp[-\textstyle\sum x_i][e^{n\theta}h(y_1, \theta)$$

the factorization criterion (de Groot 1986, p. 358) shows that $Y_1 = \min(X_1, \ldots, X_n)$ is a sufficient statistic. (Intuitively, the least time to a failure contains all the information in the n recorded failure times relevant to the grace period of θ.) With a uniform distribution of θ (which enters here as a positive location parameter), the posterior density of θ is proportional to $exp[n(\theta - y_1)]$ and yields for three observations, $(X_1, X_2, X_3) = (12, 14, 16)$, a 90 percent credence interval of $11.23 > \theta > 12.0$, in good accord with qualitative intuition. However, (F) doctrine directs one to an unbiased estimator, and the point of the example is to show what can happen when a CI is not based on a sufficient statistic. Since

$$E(X) = \int_\theta^\infty x e^{\theta - x} dx = \theta + 1$$

an unbiased estimator of θ is given by $\theta^* = n^{-1}\sum_{i=1}^{n}(X_i - 1)$. Notice, however, that this can be negative for permitted (positive) failure times, even though θ is necessarily nonnegative. The shortest 90 percent CI based on this statistic's sampling distribution

(found by computer, using an approximation) is $\theta^*–0.8529 < \theta < \theta^* + 0.8264$, or, since $\theta^* = 13$ for the three observations,

$$12.1471 < \theta < 13.8264$$

This consists entirely of values deductively excluded by the data! By contrast, the CI based on the sufficient statistic, the least of the failure times, is indistinguishable from its (B) counterpart.

Thus, Fisher was right to insist that his fiducial intervals be based on sufficient statistics. But, unfortunately, sufficient statistics do not always exist. A famous example is provided by the Cauchy distribution (the special case, $v = 1$, of Gossett's t-distribution), with density:

$$f(x \mid \theta) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$$

with θ a location parameter to be estimated. The Cauchy distribution has the peculiarity that the mean of any finite number of observations has the same (Cauchy) distribution as a single observation. Given, say, two observations, X_1 and X_2, the sampling distributions of either one or their mean, $\theta^* = ½(X_1 + X_2)$, are all the same, and so, if one's choice of estimator is to be guided solely by the sampling distributions of the candidates, as (F) doctrine dictates, then any of these statistics is as good as another for the purpose of estimating θ. However, would anyone be willing to use just the first observation and throw away the second? Or doubt that their mean is a better estimator of θ than either observation taken alone? In fact, the mean is the optimal Bayes estimator for any loss function that is a monotonically increasing function of the absolute error, $\left|\hat{\theta} - \theta\right|$, in the sense that it minimizes one's expected loss after sampling. (Lacking a prior for θ, (F) lacks any such clear-cut criterion of optimality.) Now, besides their mean, the two observations provide further information in the form of their range or half-range, $Y = ½(X_1 - X_2)$. Jaynes then calculates the conditional distribution of θ^* given Y, from which he calculates the probability that the 90 percent CI contains The true value of θ given the value of the half-range Y (1983, p. 279). The calculations show that for samples of small range, the .90 confidence coefficient is conservative: The CI for $y \leq 4$ will cover the true θ more than 95 percent of the time. However, for samples of wide range, $y \geq 10$, which comprise about 6.5 percent of the total, the CI covers θ less than 12 percent of the time.

By abandoning the principle of being guided only by the sampling distribution, (F) can also avail itself of the conditional distribution and base different estimates of θ on different values of Y, choosing for each observed y the shortest CI that, within that y-subclass, covers the true θ 90 percent of the time. For samples of narrow range, this delivers much shorter intervals than the standard 90 percent CI, while for samples of wide range, it covers the true θ more often with a join of two separate intervals. The resulting rule is uniformly reliable in never under or overstating its probability of covering the true θ, but by now one will have guessed that the uniformly reliable rule is the Bayesian rule!

A recurring theme of Jaynes's writings is that the (F) devotees of error probabilities and performance characteristics have never bothered to investigate the performance of the Bayesian solutions they denigrate as "founded on an error" or to compare their performance with their own preferred solutions. (B) methods based on uninformed priors capture Fisher's desideratum of "allowing the data to speak for themselves" as evidenced by their agreement with (F) methods based on sufficient statistics. It is then rather an onerous thesis to maintain that they fail to do this in cases where (F) lacks a solution or where, as it has just been seen, the (F) solution not so based leads to palpably absurd results or misleading statements of confidence. One can also sometimes criticize a frequentist solution as equivalent to a Bayesian solution based on an absurdly opinionated prior (see Jaynes 1983, p. 103).

Jaynes explains why (F) methods inevitably waste information as follows, "Orthodoxy requires us to choose a single estimator, $b(D) \equiv b(X_1, \ldots, X_n)$, *before we have seen the data*, and then use only $b(D)$ for the estimation" (2003, p. 510). The observed value of this statistic then places one on a manifold (or subspace) of n-dimensional space of dimension $n - 1$. If position on this manifold is irrelevant for θ, then $b(D)$ is a sufficient statistic, but if not, then D contains additional information relevant to θ that is not conveyed by specifying $b(D)$. (B) is then able to choose the optimal estimator for the present data set. The sampling distribution of $b(D)$ is simply not relevant, since one is free to choose different estimators or different CI's for different data sets.

INFORMED PRIORS AND ENTROPY

Of the many approaches to constructing uninformed priors, group invariance has been stressed because of its intimate ties to consistency. The same rationale underwrites a powerful extension of (2) to a more general rule of minimal belief change that goes by minimizing the cross-entropy deviation from an initial (pre) distribution among all those satisfying empirically given distribu-

tional constraints (see the entry on information theory). Recall, the cross entropy or discrimination information of a distribution $P = (p_1, \dots, p_n)$ with respect to $Q = (q_1, \dots, q_n)$ is defined by

$$H(P,Q) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right)$$

And when $Q = (n^{-1}, \dots, n^{-1})$ is a uniform distribution, the rule (*MINXENT*) of minimizing cross entropy specializes to the rule (*MAXENT*) of maximizing the (Shannon) entropy,

$$H(P) = -\sum_{i=1}^{n} p_i \ln p_i$$

which is a measure of the uncertainty embodied in *P*. Entropy figures centrally in Claude Shannon's mathematical theory of communication (information theory), and looks to be a fundamental concept of probability theory as well. Thus, sufficient statistics, informally defined as "preserving all the information in the data relevant to inferences about θ," do actually preserve information in the sense of entropy (Jaynes 2003, §14.2). Also see Jaynes (§17.4) for further links between sufficiency, entropy, Fisher information, and the Cramer-Rao inequality.

When the psi-test discussed earlier leads one to expect a significant improvement in support by moving to an alternative (and possibly more complicated) model, MINXENT can lead one to it, as in the example of a biased die discussed in information theory entry. Thus, MINXENT literally enables one to carve a model out of empirically given measurements or mean values. Jaynes's original (1957) application to equilibrium thermodynamics (Jaynes 1983, chapters 1–6) with later extensions to nonequilibrium thermodynamics (chapter 10, §D) remains the exemplar, but a veritable floodtide of additional applications to all areas of scientific research have since followed, as recorded in the proceedings of workshops on Bayesian and maximum entropy methods held annually since 1981. The inferential problems this opens to attack lie even further beyond the range of (F) or (L).

Moreover, many classical models like the exponential or Gaussian arise most naturally as maxent distributions. Thus, the exponential, with density, $f(x|\theta) = \theta exp(-\theta x)$, is the maxent distribution of a positive continuous *X* of know mean; the normal (or Gaussian) that of a distribution whose first two moments are known. Jaynes (2003, p. 208) makes a serious case that this best accounts for the ubiquity of the Gaussian as a distribution of errors or noise, so that it is neither "an experimental fact" nor a "mathematical theorem," but simply the most honest representation of what is typically known about one's errors, namely, their "scale" and that positive and negative ones tend to cancel each other.

MAXENT functions primarily, though, as a means of arriving at informed priors. The superiority of a Bayes solution will be more manifest, in general, when substantial prior knowledge is formally incorporated in the analysis. Research might disclose, for example, that horse 1 finished ahead of horse 2 in two-thirds of the races both entered. If that is all that is known, then one's prior for tonight's race must satisfy $p_1 = 2p_2$. (How should this information affect the odds on the other horses?) In the inventory example of Jaynes (2003, §14.7), successive pieces of information, bearing on the decision which of three available colors to paint the day's run of 200 widgets so as to ensure twenty-four-hour delivery, are assimilated, starting with the current stocks of each color, the average number of each color sold per day, the average size of an individual order for each color, and so on. This is not just an amusing and instructive example of entropy maximization, but, evidently, one with serious practical applications.

At the other extreme of uninformativeness, the Bayesian econometrician Arnold Zellner has used entropy to define a maximal data informative prior (MDIP) as one that maximizes

$$\int I(\theta)p(\theta)d\theta - \int p(\theta)\ln p(\theta)d\theta$$

the difference between the average information in the data density and the (variable) prior density, where

$$I(\theta) = \int f(x\,|\,\theta)\ln f(x\,|\,\theta)d\theta$$

measures the information conveyed by the data density, $f(x|\theta)$. Such a prior also maximizes the expected log-ratio of the likelihood function to the prior density (for a number of examples and yet another derivation of the Jeffreys log-uniform prior, see Zellner and Min 1993).

To Lindley's oft-repeated question, "Why should one's knowledge or ignorance of a quantity depend on the experiment being used to determine it?" Jaynes answers that the prior should be based on all the available prior information and "the role a parameter plays in a sampling distribution is always a part of that information" (1983, p. 352). However, it should not depend on the size of the sample contemplated (pp. 379–382).

Apart from the satisfaction of seeing that variously motivated lines of attack all lead to the same priors in the best understood cases, like location or scale parameters or regression coefficients (for which see Jaynes 1983, pp. 195–196), different methods can be expected to generalize in different ways when harder problems are addressed. Obviously, there is room for much creative thought here in what might be described as the new epistemology, the endeavor to accurately represent whatever is known in probabilistic terms—what Jaynes calls "that great neglected half of probability theory." Such research can be expected to further the development of artificial intelligence and the formation of consensus priors in decision theoretic or policymaking contexts.

SUMMARY: A BAYESIAN REVOLUTION?

Is the much heralded Bayesian revolution a *fait accompli*? In his account of scientific revolutions, Thomas Kuhn may have erred in some of the details but certainly convinced his readers that there is a pattern here, something to construct a theory of. That applies, in particular, to revolutionary theory change, the overthrow of an old paradigm in favor of the new. Brush (1994, p. 137) touches on several of the reasons that usually enter in speaking of the acceptance of wave mechanics. Based on what has been surveyed in this entry, Bayesians would lodge the following parallel claims:

- (B) offers simpler solutions to the salient problems of the old (F) paradigm
- (B) offers a unified approach to all inferential problems—indeed, to all three problems of evidence, belief, and decision mentioned at the outset
- (B) is able to pose and solve problems of obvious importance that lie beyond the range of (F) or (L), among them problems involving nuisance parameters and those amenable to MINXENT.

(B) also lays claim to greater resolving power in the detection of periodicities in time series or in separating periodicities from trends (Jaynes 2003, p. 125 and chapter 17).

(B) views (2) as embodying the entire logic of science. It has demonstrated that (2), as well as its extension to MINXENT, is anchored in the bedrock of consistency and that the price of inconsistency is inefficiency, the waste of information present in the data. Finally, (B) claims to be able to ascertain the limits of validity of the methods of (F) by viewing them as approximations to Bayesian methods. That, too, is highly characteristic of the claims a new paradigm lodges against the old. In any case, many time-honored procedures of (F), like significance tests or chi-square tests, retain an honorable place in the Bayesian corpus as approximate Bayes procedures, and where the elements needed for a Bayesian solution are lacking, one may use Bayesian logic to find a useful surrogate.

Critics will allege that Bayesians have not solved the "problem of the hypothesis space," namely, to which hypotheses should one assign probabilities? Jaynesians admit they have not solved this problem, but neither has anyone else. Jaynes's point, rather, is that the only way to discover that we have not gone to a deep enough hypothesis space is to draw inferences from the one we have. We learn most when our predictions fail, but to be certain that failed predictions reflect inadequacies of our hypothesis space rather than poor reasoning, "those inferences [must be] our *best* inferences, which make full use of all the information we have" (Jaynes 2003, p. 326).

See also Experimentation and Instrumentation; Probability and Chance.

Bibliography

Bretthorst, G. Larry. "On the Difference in Means." In *Physics and Probability: Essays in Honor of Edwin T. Jaynes*, edited by W. T. Grandy Jr. and P. W. Milonni. New York: Cambridge University Press, 1993.

Brush, Stephen G. "Dynamics of Theory Change: The Role of Predictions." *PSA 1994*. Vol. 2, edited by David Hull, Micky Forbes, and Richard Burian. East Lansing, MI: Philosophy of Science Association, 1994.

Cox, Richard T. "Probability, Frequency, and Reasonable Expectation." *American Journal of Physics* 14 (1946): 1–13.

De Finetti, Bruno. "La Prevision: ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincare* 7 (1937), 1–68. Translated as "Prevision: Its Logical Laws, Its Subjective Sources." In *Studies in Subjective Probability*. 2nd ed., edited by Henry Kyburg and Howard Smokler. New York: Wiley, 1981.

De Finetti, Bruno. "Sur la condition d'equivalence partielle." *Actualites scientifiques et industrielles*, No. 739 (Colloque Geneve d'Octobre 1937 sur la Théorie des Probabilites, 6tieme partie.) Paris: Hermann, 1938. Translated as "On the Condition of Partial Exchangeability." In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey. Berkeley: University of California Press, 1980.

De Finetti, Bruno. *Probability, Induction, and Statistics*. New York: Wiley, 1972.

De Groot, Morris. *Probability and Statistics*. 2nd ed. Reading, MA: Addison-Wesley, 1986.

Edwards, Ward, Harold Lindman, and Leonard J. Savage. "Bayesian Statistical Inference for Psychological Research." In *Readings in Mathematical Psychology*. Vol. 2, edited by. R. Duncan Luce, R. Bush, and Eugene Galanter. New York: Wiley, 1965.

Festa, Roberto. *Optimum Inductive Methods*. Dordrecht, Netherlands: Kluwer Academic, 1993.

Fisher, Ronald A. *The Design of Experiments*. Edinburgh, Scotland: Oliver and Boyd, 1935.

Fisher, Ronald A. *Statistical Methods and Scientific Inference*. 2nd ed. Edinburgh, Scotland: Oliver and Boyd, 1959. All of Fisher's three main books on statistics have been reprinted by Oxford University Press (2003) under the title *Statistical Methods, Experimental Design and Statistical Inference*.

Giere, Ronald N. "Philosophy of Science Naturalized." *Philosophy of Science* 52 (1985): 331–356.

Good, I. J. *Good Thinking*. Minneapolis: University of Minnesota Press, 1983.

Hacking, Ian. *Logic of Statistical Inference*. New York: Cambridge University Press, 1965.

Hald, Anders. *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley, 1998.

Hill, T. P. "Base Invariance Implies Benford's Law." *Proceedings of the American Mathematical Society* 123 (1995): 887–895.

Hodges, J. L., and Erich Lehmann. *Basic Concepts of Probability and Statistics*. 2nd ed. San Francisco: Holden-Day, 1970.

Jaynes, E. T. *Papers on Probability, Statistics, and Statistical Physics*, edited by R. D. Rosenkrantz. Dordrecht, Netherlands: D. Reidel, 1983.

Jaynes, E. T. *Probability Theory: The Logic of Science*, edited by G. Larry Bretthorst. New York: Cambridge University Press, 2003.

Jeffreys, Harold. *Theory of Probability*. Oxford, U.K.: Clarendon Press, 1939.

Kalbfleish, John G., and D. A. Sprott. "On Tests of Significance." *In Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science.* Vol. 2, edited by William L. Harper and Clifford A. Hooker. Dordrecht, Netherlands: D. Reidel, 1976.

Kendall, Maurice G., and Alan Stuart. *Advanced Theory of Statistics.* 2nd ed. New York: Hafner, 1967.

Lehmann, Erich. *Testing Statistical Hypotheses*. New York: Wiley, 1959.

Lindley, Dennis V. *Introduction to Probability and Statistics. Vol. 2, Inference*. Cambridge, U.K.: Cambridge University Press, 1965.

Loredo, T. J. "From Laplace to Supernova SN 1987A: Bayesian Inference in Astrophysics." In *Maximum Entropy and Bayesian Methods*, edited by Paul Fougere. Dordrecht, Netherlands: Kluwer Academic, 1990.

Mayo, Deborah G. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press, 1996.

Morrison, D., and R. Henkel, eds. *The Significance Test Controversy*. Chicago: Aldine Press, 1970.

Pearson, Egon S. *The Selected Papers of E. S. Pearson*. Berkeley: University of California Press, 1966.

Rosenkrantz, R. D. *Inference, Method, and Decision*. Dordrecht, Netherlands: D. Reidel, 1977.

Rosenkrantz, R. D. "The Justification of Induction." *Philosophy of Science* 59 (1992): 527–539.

Royall, Richard M. *Statistical Evidence: A Likelihood Paradigm*. London: Chapman and Hall, 1997.

Savage, L. J. *The Foundations of Statistics*. New York: Wiley, 1954.

Smith, C. Ray, and Gary J. Erickson. "Probability Theory and the Associativity Equation." In *Maximum Entropy and Bayesian Methods*, edited by Paul Fougere. Dordrecht, Netherlands: Kluwer Academic, 1990.

Zabell, S. L. "Symmetry and Its Discontents." In *Causation, Chance, and Credence*, edited by Brian Skyrms and W. L. Harper, 155–190. Dordrecht, Netherlands: Kluwer Academic, 1988.

Zellner, Arnold, and Chung-ki Min. "Bayesian Analysis, Model Selection, and Prediction." In *Physics and Probability: Essays in Honor of Edwin T. Jaynes*, edited by W. T. Grandy Jr. and P. W. Milonni. New York: Cambridge University Press, 1993.

Zellner, Arnold, Hugo A. Keuzenkamp, and Michael McAleer, eds. *Simplicity, Inference, and Modeling: Keeping It Sophisticatedly Simple*. New York: Cambridge University Press, 2001.

Roger D. Rosenkrantz

STEBBING, LIZZIE SUSAN

(1885–1943)

Lizzie Susan Stebbing, the English logician and philosopher, was born in London. A very delicate child, she received a discontinuous education until she went to Girton College, Cambridge, in 1906. While at Cambridge she happened to read F. H. Bradley's *Appearance and Reality,* which led to her interest in philosophy. She became a pupil of the logician W. E. Johnson. From 1913 to 1915 she lectured in philosophy at King's College, London; and she became a lecturer at Bedford College, London, in 1915 and a professor in 1933.

In London Stebbing's philosophical development was stimulated by the meetings of the Aristotelian Society, which were often attended by Bertrand Russell, A. N. Whitehead, and G. E. Moore; and she always acknowledged the philosophical influence of Moore as particularly strong. In 1931 she published *A Modern Introduction to Logic* and in 1937 *Philosophy and the Physicists,* which were by a considerable degree the most substantial of her books. She wrote numerous papers, the best of which are to be found in *Mind* and the *Proceedings of the Aristotelian Society.*

In philosophy Stebbing's main interests lay in the metaphysical questions posed by logic and in the foundations of science. Much of her work in these topics is contained in *A Modern Introduction to Logic.* The book's merit does not lie in any originality in formal logic, or even in its method of presenting formal structures, but rather in its clear exposition of the logical theories of the early twentieth century, together with a stimulating, lucid, perceptive account of the metaphysical problems

the new logical techniques either dispersed or clarified, and of the metaphysics that lay behind these logical theories. It was the first book on modern logic that introduced together and comprehensively both the formalism and its related philosophical problems. It is probably still the best introduction for a reader prepared to give serious thought to such problems.

In the professional journals Stebbing published papers on a range of topics closely related to those of *A Modern Introduction to Logic,* but her interests were not confined to such purely academic, though deeply absorbing, matters. She wrote several books on what one might call logic in practice. (Her book *Thinking to Some Purpose* is a good example both in its title and in its content.) She was strongly convinced of the importance of rationality and clarity in the conduct of human affairs and of the immense importance of knowledge. She attempted, therefore to expose the artifices by which hard facts are obscured in soft language, either so that the unscrupulous may deceive us or so that we may hide from ourselves what we do not wish to see. Her books in this field are especially valuable for their actual examples of irrationality and emotional persuasion in high places and on vital matters.

This commitment to rational clarity was combined with her more purely professional interests and skills in *Philosophy and the Physicists.* In the course of writing books with the ostensible aim of popularizing contemporary science, Sir James Jeans and Sir Arthur Eddington had argued that modern physics shows the world to be quite other than the sort of place it seems to be, not merely physically but also metaphysically. Both argued for idealist views of physics and, consequently, for a comfortable if imperfectly clear form of theism. In much of her book Stebbing exposed the fallacies, needless obscurities and mystifications with which the pages of Jeans and Eddington abound. *Philosophy and the Physicists* is an excellent piece of rational cool criticism, but a significant characteristic of the book is its implicit faith that we need not seek protection behind intellectual smoke screens and, indeed, that this sort of evasion prevents any really dignified adjustment to the human situation based on knowledge and reason. Stebbing deeply believed that such an adjustment is possible.

See also Bradley, Francis Herbert; Eddington, Arthur Stanley; Jeans, James Hopwood; Logic, History of: Modern Logic; Moore, George Edward; Russell, Bertrand Arthur William; Whitehead, Alfred North; Women in the History of Philosophy.

Bibliography

Books by Susan Stebbing are *A Modern Introduction to Logic* (London: Methuen, 1931; 2nd ed., revised, London, 1933); *Philosophy and the Physicists* (London: Methuen, 1937); *Thinking to Some Purpose* (Harmondsworth, U.K.: Penguin, 1939); and *Ideals and Illusions* (London: Watts, 1941).

The Aristotelian Society collected the material for *Philosophical Studies: Essays in Memory of L. Susan Stebbing* (London: Allen and Unwin, 1948), a volume of essays that contains an appreciation by John Wisdom and a full bibliography of Stebbing's writings.

G. C. Nerlich (1967)

STEFANINI, LUIGI

(1891–1956)

Luigi Stefanini, the Italian personalist philosopher, taught at Messina and Padua. He was a founder of the Gallarate movement and the founder and first editor of the *Rivista di estetica.* Much of Stefanini's own philosophy is to be found in his work on the history of philosophy. He tried to demonstrate by careful historical analysis that authentic religious and metaphysical needs are adequately met by certain historical positions, especially those of St. Augustine and St. Bonaventure. His guiding principle, "paradigmatism," is of Platonic and Neoplatonic origin and may be stated thus: that which is created in the image of another (as is man) has as its constitutive imperative, or life vocation, the expression in itself of its transcendental model.

Stefanini professed in turn Christian idealism, spiritualism, and personalism. His Christian idealism was based on a critique of Giovanni Gentile's claim that the self generates the self and the world and hence is the paradigm of the world. Stefanini held that the self apprehends itself not as self-generating but as created and therefore has its paradigm in an other. Art is an immediate expression of that other and provides an approach to the Christian experience, in which the image of God in the human subject is remodeled on the higher paradigm of Christ.

Stefanini's spiritualism began in a critique of historicism, phenomenology, and existentialism. All of these, he claimed, divide the transcendental from the existential. He sought to heal this split by the analysis of the self. The self is not existence as given (*Dasein*) but existence that utters itself. The self is spirit, or word, and this word does not utter, but alludes to, the Absolute; in this way it reveals its dependence. The purest form of this allusion to the Absolute is the Word of God, Christ. The vocation of the Christian is to utter that Word in himself.

Stefanini called his most mature thought "personalism." The self is central to every form of participation and is the only ultimate point of reference. But the self cannot sustain itself; it rests upon the other, and the transcendent is therefore the principle of the self's being. The self realizes itself as a person by its relation to the transcendent. It seeks to realize the transcendent in itself according to the limits and form of its own being.

See also Absolute, The; Augustine, St.; Bonaventure, St.; Existentialism; Gentile, Giovanni; Historicism; Idealism; Personalism; Phenomenology; Self.

Bibliography

WORKS BY STEFANINI

Idealismo cristiano. Padua: Zannoni, 1931.

Platone. 2 vols. Padua: Milani, 1932–1935.

Spiritualismo cristiano. Brescia, 1942.

La metafisica della persona. Brescia: Morcelliana, 1949.

La metafisica dell'arte. Brescia: Morcelliana, 1949.

WORKS ON STEFANINI

Bortolaso, Giovanni. "Uno spiritualista cristiano: Luigi Stefanini." *Civiltà cattolica* (1956) (1): 295–304.

Carlini, Armando. "Incontri e scontri con Stefanini e con Sciacca." *Giornale critico della filosofia italiana,* series III, 4 (1950): 841–893.

Chaix-Ruy, Jules. "Les philosophes italiens d'aujourd'hui." *Revue thomiste* 55 (1947): 407ff.

De Ruggiero, Guido. "Stefanini." *La critica* (1934): 383ff.

Luigi Stefanini. Filosofi d'oggi series. Turin. Contains a complete bibliography.

A. Robert Caponigri (1967)

STEFFENS, HENRICH
(1773–1845)

Henrich Steffens, the philosopher, scientist, and novelist and short-story writer was of Danish and German descent. He was born in Stavanger, Norway, the son of a physician in the service of the Dano-Norwegian monarchy. From 1790 to 1794 Steffens studied natural science, especially mineralogy and geology, in Copenhagen. He next studied natural history in Kiel, where he became interested in philosophy. In 1798 he moved to Jena, drawn not least by the natural philosophy of Friedrich von Schelling, whose *Erster Entwurf eines Systems der Naturphilosophie* had appeared in 1797. In Jena, Steffens met Schelling, Johann Wolfgang von Goethe, and August Schlegel; and in Berlin in 1799 he met Friedrich von Schlegel and Friedrich Schleiermacher.

In 1802 Steffens returned to Copenhagen to lecture on natural philosophy. Through his large audience he influenced the development of the romantic movement in Denmark, but he failed to obtain the university position he had hoped for, and in 1804 he accepted a chair in natural philosophy and mineralogy at the University of Halle. In 1811 he was appointed professor of physics in Breslau, where he remained, except for a brief period of service as a volunteer in the war against Napoleon Bonaparte in 1813–1814, until 1832. In that year Steffens became professor at Berlin, where he lectured on natural philosophy, anthropology, and geology until his death.

Steffens's philosophy was markedly influenced by Benedict de Spinoza and by Spinozistic pantheism, as well as by Schelling. Schelling's *Von der Weltseele, eine Hypothese der höheren Physik zur Erklärung des allgemeinen Organismus* (On the world-soul, a hypothesis of higher physics in explanation of the general organism) appeared in 1798, and in Steffens's *Beiträge zur innern Naturgeschichte der Erde* (Contributions to the inner natural history of the earth; 1801) the influence of Schelling is readily discernible. The title of Schelling's work gives an indication of the substance and trend of Steffens's philosophical thinking; it is a blend of natural science and speculative philosophy imbued with the general spirit of the romantic movement, somewhat less speculative than that of Schelling.

Steffens viewed the history of nature as a development or evolution from inorganic stages to organic and animate forms, governed by a divine purpose. His pantheism found characteristic expression in the view that nature itself is creative, the acme of the natural creative process being the free individual human personality, or spirit. According to Steffens's *Anthropologie* (1822) man is a living unity of spirit and nature—a microcosm, in the sense that the history of humankind mirrors the development of nature itself. He found in myths and mythological traditions a true, though symbolically expressed, understanding and knowledge of nature; however, he believed that a proper scientific study of nature was a necessary prerequisite for a correct interpretation of the meaning of myths.

See also Goethe, Johann Wolfgang von; Pantheism; Philosophy of Physics; Schelling, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von; Schleiermacher, Friedrich Daniel Ernst; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY STEFFENS

Beiträge zur innern Naturgeschichte der Erde. Freiburg, 1801.

Grundzüge der philosophischen Naturwissenschaft. Berlin: Realschulbuchhandlung, 1806.

Anthropologie. Breslau: J. Max, 1822.

Was ich erlebte, 10 vols. Breslau, 1840–1844. Steffens's autobiography; an important source for the study of the romantic movement in Germany.

WORKS ON STEFFENS

Petersen, Richard. *Henrich Steffens.* Copenhagen, 1881.

Waschnitus, Victor. *Henrich Steffens.* Neumünster, Germany, 1939.

Knut Erik Tranöy (1967)

STEIN, EDITH

(1891–1942)

Edith Stein was born into a German Jewish family on October 12, 1891, on Yom Kippur, in the Silesian capital Breslau, Germany (after 1945, Wroclaw, Poland). She was the youngest of eleven children, four of whom died in early childhood. Her father, Siegfried Stein (1844–1893), had a small trade with coals and wood and died too early for his youngest child to have any memory of him. Her mother, Auguste Stein, née Courant (1849–1936), was a matriarchal, warm-hearted woman who tried to educate her children in the traditional Jewish faith and in the celebration of the rituals. Nonetheless, the industrious and highly intelligent girl became an agnostic from her puberty onward and already in school became a champion of women's liberation.

After a brilliant performance on school examinations, she studied psychology with William Stern, philosophy with Richard Hönigswald, along with German literature and history, at the Universität Breslau from 1911 to 1913. One can obtain a good sense of her feelings from that period, up to her doctorate in 1916 from the Albert-Ludwigs-Universität Freiburg, from her fragmentary autobiography *Life in a Jewish Family*, written in 1933 but first published in 1965. In 1913 Stein went to Göttingen to study under the famous founder of phenomenology, Edmund Husserl (1859–1938), and with his assistant Adolf Reinach (1883–1917), whose death in World War I affected her very deeply. In 1915 she worked as a Red-Cross nurse in an international soldiers' recovery hospital in Weißkirchen, Mähren (now located in the Czech Republic). After completing her state examinations, she followed Husserl to Universität Freiburg in 1916, where she completed her dissertation *On the Problem of Empathy* summa cum laude. From 1917 to 1918 she served as Husserl's private assistant, transcribing, ordering, and completing his manuscripts, preparing for publication his *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, books 2 and 3, along with his *On the Phenomenology of the Consciousness of Internal Time*, later published in 1927 by Martin Heidegger.

Between 1918 and 1932 Stein attempted four times to qualify for a habilitation (the highest qualification in the German university), at the universities in Göttingen, Freiburg, Breslau, and Kiel, but she failed partly because she was female and partly because she was a Jew. During a deepening personal as well as academic crisis as her relationships with the phenomenologists Roman Ingarden and Hans Lipps weakened, she started studying classical Christian literature, especially St. Teresa of Ávila, as well as Martin Luther, Søren Kierkegaard, and St. Augustine. Her Catholic baptism on January 1, 1922, separated her in a painful way from her family, especially from her mother, who received a second, almost unsustainable blow in October 1933, when Stein entered the Carmelite order in Cologne. From 1923 until 1931, she worked as a teacher of German and history at a girls' college, Mädchen-Lyzeum, in Speyer on the Rhine, and from 1932 until March 1933 she taught as a docent at the Deutsches Institut für wissenschaftliche Pädagogik (German Institute for Scientific Pedagogy) in Münster. From 1928 through 1933, her spiritual mentor was Raphael Walzer OSB, arch abbot of the Benedictine monastery at Beuron. During the same period she became well known in Catholic circles in Germany, Austria (Salzburg, Vienna), and Switzerland (Zurich) through her lectures on Christian anthropology and Christian feminism.

After the removal of non-Aryans from official positions in the spring of 1933, Stein left the institute to fulfill her wish for a Carmelite existence. In April 1933 she wrote a famous letter to Pope Pius XI asking him to protest against the humiliation of Jews and predicting a coming prosecution of the Catholic Church too. From 1933 through 1938 she stayed in the Carmelite cloisters in Cologne, using the name Sister Teresia Benedicta a cruce of the Cross OCD. In 1939 she moved to the Carmelite cloisters at Echt, Netherlands. After the protest of Dutch Catholic bishops against prosecution of Jews, she and her sister Rosa were arrested by the Gestapo on August 2, 1942, brought first to the Dutch camps of Amersfort and Westerbork, and taken from there by train to Auschwitz. The day of her arrival on August 9, 1942, is most probably the day she was killed. In 1987 she was beatified, in

1998 sanctified, and in 1999 named copatroness of Europe by Pope John Paul II.

PHILOSOPHICAL WORKS

In the first, strictly phenomenological period of her writing while she was one of Husserl's leading students (1916–1922), Stein employed Husserl's phenomenological method in fundamental analyses in anthropology, focusing on psychology, psychophysical interactions, intersubjectivity, and personhood. Her dissertation investigated empathy (a field neglected by Husserl) as the basis for intersubjectivity and the experience of the other's and one's own body, referring to the tradition of Theodor Lipps, Max Scheler, and Alexander Pfänder, and then developing independent conclusions. In 1919, in *Einführung in die Philosophie*,Stein critiqued Husserl's idealistic position on the ego, contrasting his view of the monadic ego with arguments for a real external world. Her habilitation *Philosophy of Psychology and the Humanities* (1922/2000) differentiated the psyche and the soul with reference to causality and motivation. Causality determines the bound psyche with the help of conditions and psychic laws, while motivation inspires the free, creative will of the personal soul. Respectively they constitute the sensual, receptive subject and the rational, active subject. She takes an analogous approach in her treatment of the community and its transindividual reality in *Individuum und Gemeinschaft* (Individuality and community; 1922). The essential difference between psychic bindings and rationally deciding leads to the difference between psychology and the humanities (*Geisteswissenschaften*). The voluminous study *Eine Untersuchung über den Staat* (A study of the state; 1925) illuminates the ontological basis of sociology by differentiating between community and society and showing the roots of society in community and the roots of community in the individual.

In her second period after her baptism (1922–1937), Stein, in analyzing important parts of the Christian tradition but still doing so in a phenomenological way, was drawn to classical ontology and metaphysics. Inspired by the Jesuit Erich Przywara, in the 1920s Stein translated John Henry Newman's *Letters from the Anglican Period* and *Idea of a University*, and Thomas Aquinas's *Quaestiones disputatae de veritate* (Disputations on truth; 1931–1934) and *De ente et essentia* (On being and essence; unpublished yet). Her studies in Christian feminism and female education, including essays on Elisabeth of Thüringen and Teresa of Avila, revealed a remarkable phenomenology of womanhood, especially in reference to the interrelation of body, soul, self-concept, and being divinely gifted. While teaching in Münster from 1932 to 1933, she wrote a philosophical anthropology and a fragmentary theological anthropology in *Der Aufbau der menschlichen Person* (The structure of human person) and *Was ist der Mensch?* (What is a human being?). The difference, but also the possible connection, between phenomenological method and scholastic ontology is shown in *Was ist Philosophie? Ein Gespräch zwischen Edmund Husserl und Thomas von Aquino* (1929) a Platonic dialogue between Husserl and Aquinas, with Aquinas as the leading speaker. In *Potenz und Akt* (Potentiality and act; 1931) and *Endliches und ewiges Sein* (Finite and eternal being; 1936/37), Stein tried to reconcile phenomenology and scholastic philosophy in a contemporary fashion. Referring to Aristotle, Aquinas, Augustine, pseudo-Dionysius, Heidegger, Jean Hering, and Hedwig Conrad-Martius (her godmother, famous for a philosophy of nature and of space and time), Stein tried to analyze different conceptions of being and to reconcile phenomenology and classical and medieval ontology into a philosophy for all time. Though she started with Aquinas, who maintained an Aristotelian ontology, she ultimately ended up closer to Augustine's personalism and his trinitarian view of creation. The aim of her philosophy was a theory of the person, not of ontological being.

In her third period (1940–1942), Stein composed two important studies on Christian spirituality and mystics. To prepare for a modern analysis of the great Spanish Carmelite reformer John of the Cross (1542–1591), she translated the complete works of pseudo-Dionysius (the Areopagite), the father of occidental mysticism, and dedicated to him the essay "Wege der Gotteserkenntnis" (Ways to recognize God; 1940/41). She reconstructed and commented on the three classical Areopagitic ways of pursuing theology: the positive, the negative, and the mystical. As an immediate fruit of rethinking the basics of mysticism, Stein provided an immanent interpretation of the theory and poetry of mystical ascent by John of the Cross, in her last, almost completed work *The Science of the Cross* (1950/2002). In his three-dark-nights theory of spiritual development, one must pass through the night of sentiment, the night of mind, and the night of faith before ascending to God. She also held that one must annihilate the self before reaching the glory of God—a theory that sheds light on Stein's own inner spiritual development. Her reflections retain language and methods close to phenomenological research.

LEGACY

Until 1930 the writings and translations of Stein were published for the most part during her lifetime. All of her other works, letters, and uncompleted projects began to be published from 1950 to 1998 in *Edith Steins Werke* in 18 volumes by Herder in Freiburg. A new critical edition of all her writings, based on the complete material in the Carmelite Archive in Cologne and including translations and scattered pieces, is being projected from 2000 to 2010 as *Edith Stein Gesamtausgabe* (Complete works of Edith Stein) in 25 volumes, also by Herder. The interest in her life initially led to many hagiographic studies. Meanwhile, since the 1990s her philosophical work on Husserl and Heidegger has met with strong interest and received an increasingly positive appraisal. Stein's importance and influence in the history of phenomenology has yet to be fully explored.

See also Phenomenology; Thomism.

Bibliography

WORKS BY STEIN

The Collected Works of Edith Stein. Washington, DC: ICS Publications, 1986–2002. Vol. 1: *Life in a Jewish Family, 1891–1916: An Autobiography*, translated by Josephine Koeppel (1986). Vol. 2: *Essays on Woman*, translated by Freda Mary Oben (1996). Vol. 3: *On the Problem of Empathy*, translated by Waltraud Stein, 3rd, rev. ed. (1989). Vol. 4: *The Hidden Life*, translated by Waltraud Stein (1992). Vol. 5: *Self Portrait in Letters, 1916–1942*, translated by Josephine Koeppel (1993). Vol. 6: *The Science of the Cross* (1950), translated by Josephine Koeppel (2002). Vol. 7: *Philosophy of Psychology and the Humanities*, translated by Mary Catharine Baseheart and Marianne Sawicki (2000). Vol. 8: *Knowledge and Faith*, translated by Walter Redmond (2000). Vol. 9: *Finite and Eternal Being: An Attempt at an Ascent to the Meaning of Being*, translated by Kurt F. Reinhardt (2002).

Edith Steins Werke, edited by Lucy Gelber and Michael Linssen. Freiburg, Germany: Herder, 1950–.

Einführung in die Philosophie, edited by Claudia Mariéle Wulf. Freiburg, Germany: Herder, 2004.

Gesamtausgabe. Freiburg, Germany: Herder, 2000–.

Individuum und Gemeinschaft. Halle, Germany: M. Niemeyer, 1922.

Eine Untersuchung über den Staat. Halle, Germany: M. Niemeyer, 1925.

WORKS ON STEIN

Baseheart, Mary Catharine. *Person in the World: Introduction to the Philosophy of Edith Stein*. Dordrecht: Kluwer, 1997.

Beckmann, Beate. *Phänomenologie des religiösen Erlebnisses: Religionsphilosophische Überlegungen im Anschluß an Adolf Reinach und Edith Stein*. Würzburg, Germany: Königshausen und Neumann, 2003.

Beckmann, Beate, and Hanna-Barbara Gerl-Falkovitz, eds. *Edith Stein: Themen, Bezüge, Dokumente*. Würzburg, Germany: Königshausen und Neumann, 2003.

Fetz, Reto Luzius, Matthias Rath, and Peter Schulz, eds. *Studien zur Philosophie von Edith Stein*. Freiburg, Germany: Alber, 1993.

Gerl-Falkovitz, Hanna-Barbara. *Unerbittliches Licht: Edith Stein; Philosophie, Mystik, Leben*. 3rd ed. Mainz, Germany: Grünewald, 1999.

Herbstrith, Waltraud, ed. *Never Forget: Christian and Jewish Perspectives on Edith Stein*. Translated by Susanne Batzdorff. Washington, DC: ICS Publications, 1998.

Neyer, Amata. *Edith Stein: Her Life in Photos and Documents*. Translated by Waltraud Stein. Washington: ICS Publications, 1999.

Sawicki, Marianne. *Body, Text, and Science: The Literacy of Investigative Practices and the Phenomenology of Edith Stein*. Dordrecht, Netherlands: Kluwer, 1997.

Hanna-Barbara Gerl-Falkovitz (2005)

STEINER, RUDOLF
(1861–1925)

Rudolf Steiner, the German philosopher and occultist, was born in Kraljevic, Hungary, of Catholic parents. His early education was obtained at technical secondary schools and the Polytechnic Institute of Vienna. Steiner's anthroposophical teaching, presented as "spiritual science," is an extraordinary synthesis of "organic" ideas in nineteenth-century German thought with theosophical material and fresh occult intuitions. In 1902 Steiner became a lecturer and general secretary of the Theosophical Society's German branch, but his earlier thought had been basically formed between 1890 and 1897, years devoted to the study and editing of Johann Wolfgang von Goethe's scientific writings at the Goethe-Archiv in Weimar. In this time, and during a period (1897–1900) as editor of the *Magazin: Monatschrift für Litteratur*, he developed his own views of evolution, natural organization, and science through confrontation with the ideas of Charles Darwin, Ernst Heinrich Haeckel, Friedrich Nietzsche, and contemporary German philosophies.

Steiner presented his synthesis as a modern scientific and monistic world conception, despite the range of esoteric content it eventually included. His early work, *Philosophie der Freiheit* (1896), contained no occult material, but it left room for inclusion of such material by the theories of knowledge and of spiritual freedom which it expounded: Mechanistic science gives only abstract knowledge of some uniform relations in nature. The model for fuller knowledge of individual beings is the organic idea of a self-evolving and self-directing organ-

ism, which Goethe saw in the "primal plant." The method for generalizing such knowledge is one of intuitive thinking. Steiner espoused a "monism of thought": A valid world image is ever building as individual spirits live in (*miterleben*) the organic world process.

Heralding Nietzsche's independence of thought, Steiner followed him in rejecting both natural teleology and objective moral laws. Yet he maintained that Nietzsche was always protestingly and tragically dashing his free spirit against an alien culture and a limited science of nature. Nietzsche's doctrine of "eternal recurrence," however, was a factor that led Steiner to give sympathetic attention to Indian thought. Nature is, after all, but one manifestation of spiritual reality, which reveals itself more directly in thought and in art. Among Indian ideas which Steiner adopted while a theosophist is the fourfold construction of man on Earth as having the physical, the ether, the astral bodies, and the "I," with their respective powers of development and transformation.

After 1907 conflict with Annie Besant's pro-Hindu policies led Steiner to withdraw from the Theosophical Society, but he continued on an independent line of esoteric thinking, to which in 1913 he gave the name "anthroposophy." Natural evolution, he then taught, has thus far been a progression of bodily organizations into which "pure spirit" descends through successive reincarnations with the aim of producing individual self-consciousness. Reaching its apogee in the Renaissance, this development showed its dangerous limitations in nineteenth-century individualism. The societal remedy, Steiner declared in 1919, was not the collectivism of a totalitarian state but a "three-fold social organism," in which the juridical, spiritual, and economic spheres of life are independently organized as three autonomous interacting systems. Equality is a concept applying particularly to the juridical sphere of rights (which includes just compensation for work), liberty to the spiritual domain, and fraternity or voluntary cooperation to the economic organization of production.

Steiner's own interest lay primarily in the liberty of the spiritual sphere, which included great reaches of "cosmic memory." In future stages of evolution, spirit, without loss of self-consciousness, must ascend again through knowledge of its cosmic relations to its universality and transcendence over matter. Special organs ("the lotuses") must be cultivated to apprehend the higher worlds of spirit and the traces left by their events in the cosmic ether. These include the anti-Lucifer impulsions given by Buddha, Zarathustra, Plato, and Christ and the regenerative solar influence of the blood shed in the mystery of Golgotha.

After World War I Steiner was able to establish a cultural center, the Goetheanum, in Switzerland at Dornach, near Basel. His movement spread from Germany to England, the United States, and other countries. Anthroposophy was practiced at various levels of initiation; those not ready for the higher insights could participate in the preliminary disciplines. These included eurythmic dance, mystery plays, organic agriculture and therapy, and distinctive educational measures in a number of notable elementary schools, beginning with the Waldorf School in Stuttgart. While the higher aim of Steiner's pedagogy was to develop special powers of spiritual insight, the cultivation of moral balance, a harmony of virtuous dispositions intermediate between excesses and defects, was considered a prerequisite.

Bibliography

The writings of Rudolf Steiner are extensive. His autobiography, *Mein Lebensgang* (Dornach: Philosophisch-Anthroposophischer, 1925), was translated by Olin D. Wannamaker as *The Course of My Life* (New York, 1928).

Of basic interest to students of general philosophy are *Philosophie der Freiheit* (Berlin, 1896), translated as *Philosophy of Spiritual Activity* (2nd ed., rev. and enl., London, 1916); *Goethes Weltanschauung* (Stuttgart, 1897); and *Friedrich Nietzsche: Ein Kämpfer gegen seine Zeit* (1895; expanded 2nd ed., Dornach: Philosophisch-Anthroposophischer, 1926).

For further developments of his thought in various directions, see *Knowledge of the Higher Worlds and Its Attainment* (London and New York: Putnam, 1932); *The New Art of Education* (London: Anthroposophical, 1928); *The Problems of Our Time* (London and New York, 1919); and *The Writings and Lectures of Rudolf Steiner,* compiled by P. M. Allen (New York: Whittier, 1956).

A collected edition of Steiner's major writings was begun in observation of the centennial of his birth. Vol. I in English has appeared as *Cosmic Memory: Prehistory of Earth and Man,* (West Nyack, NY: Rudolf Steiner Publications, 1961); it is a translation by Karl E. Zimmer of *Aus der Akasha-Chronik.* Four additional volumes have been published. For additional information, the reader may consult the *Bibliographie der Werke Rudolf Steiners,* prepared by Guenther Wachsmuth (Dornach, 1942).

A secondary work of interest is J. W. Hauer, *Werden und Wesen der Anthroposophie* (Stuttgart, 1922).

Horace L. Friess (1967)

STEPHEN, LESLIE

(1832–1904)

Leslie Stephen, an English man of letters, was the son of James and Jane Venn Stephen, both of whom came from families in the innermost group of the reforming Evangelicals who formed the so-called Clapham Sect. He attended Eton, briefly and unhappily, and then went to Trinity Hall, Cambridge, where he was made a fellow in 1854. Fellows had then to be ordained in the Church of England, and Stephen took holy orders and eventually became a priest, although he was not deeply religious. At the same time, religious doubt and disaffection began to trouble him. In 1862, as a result of these doubts, he resigned his fellowship, and in 1864 he left Cambridge for good. By 1865 he had completely lost all religious belief. He settled in London and began writing for various journals. Thereafter he wrote continually, copiously, and on a very wide range of topics.

In 1867 he married William Makepeace Thackeray's daughter Harriet Marian. She died in 1875, leaving him with one child. Three years later he married Julia Jackson Duckworth, a widow. They had four children, one of whom became the writer Virginia Woolf. Julia Stephen died in 1895.

Stephen was for many years editor of the *Cornhill Magazine.* In 1882 he accepted an invitation to edit the newly projected *Dictionary of National Biography.* The success of the project was largely due to his lengthy period of arduous service in this position (he wrote 387 of the biographies himself). Stephen was knighted in 1901.

Stephen was not a considerable innovator, in philosophy, in historical method, or in literary criticism. He had, however, very great gifts of rapid narration and clear and lively exposition. His work on the history of thought is based on massive reading and wide acquaintance with the social, political, and religious aspects of the periods of which he wrote. If it is neither original in its criticism nor profound in its understanding of positions, it is still useful and has not been entirely superseded because of its grasp of the broader contexts of thought and the skill with which it brings out the continuities from one period to another and from earlier formulations of problems to later ones.

It was Stephen who made Thomas Huxley's coinage *agnostic* an English word, and the problems and beliefs springing from his agnosticism underlay both his major historical works and his philosophical writings. He rejected theism of the sort he had originally been taught because he rejected the doctrine of original sin and because the problem of evil seemed to him insoluble. To evade this problem by confessing the transcendence and incomprehensibility of God was, he thought, to change from a believer into a skeptic, and in that case the part of honesty was simply to avow oneself an agnostic. But true Victorian that he was, he felt that morality, by this view, becomes gravely problematical. If there is no deity to sanction moral principles, why will—why should—men obey them?

To answer these questions was part of Stephen's aim in his investigations of eighteenth-century thought. He dealt more systematically with them, and with others, in his least successful and most tedious book, *The Science of Ethics.* The agnostic, he held, must place morality on a scientific basis, and this means that there must be nothing in his ethics that is outside the competence of scientific inquiry. Brought up on John Stuart Mill and profoundly influenced by Charles Darwin, Stephen attempted to cut through what he impatiently dismissed as academic debates about morality by showing that moral beliefs were the result neither of excessively rational utilitarian calculation nor of mysterious intuition but of the demands of the social organism in its struggle for survival. Since the healthy survival of the social organism must increasingly coincide with conditions that bring the greatest happiness to the greatest number of those individuals who are the "cells" in the "social tissue," utilitarianism is not entirely false. But its atomistic analysis of society is erroneous, and its criterion of rightness is neither adequate nor entirely accurate. The healthy survival of society, and of oneself as part of it, can alone serve as sanction for morality, and the rules for that health, which are mirrored in our instincts and our deepest habits and appear in consciousness as intuitively known moral rules, can be put on a scientific basis only when we come to possess, as we do not yet, a scientific sociology.

Bibliography

WORKS BY STEPHEN

Stephen's works are far too numerous to be listed completely here. *Essays on Freethinking and Plainspeaking* (London: Longmans, Green, 1873) and *An Agnostic's Apology and Other Essays* (London: Smith, Elder, 1893) contain most of his better-known popular essays. *The Science of Ethics* (London: Smith, Elder, 1882) is his only purely philosophical work. His important historical studies are *History of English Thought in the Eighteenth Century* (2 vols., New York: Putnam, 1876; 3rd ed., 1902); *The English Utilitarians* (3 vols., London: Duckworth, 1900); and *Hobbes* (London: Macmillan, 1904). To these the lectures in *English*

Thought and Society in the Eighteenth Century (London, 1904) provide a valuable supplement.

WORKS ON STEPHEN

The standard biography is F. W. Maitland's charming *Life and Letters of Leslie Stephen* (London: Duckworth, 1906), which contains an adequate bibliography of Stephen's work. Noel Annan, in *Leslie Stephen* (London: MacGibbon and Kee, 1951), studies Stephen as a representative Victorian thinker and as a link between the Clapham Sect and the Bloomsbury Group.

J. B. Schneewind (1967)

STERN, LOUIS WILLIAM

(1871–1938)

Louis William Stern, the German philosopher and psychologist, was born in Berlin and received his PhD under Hermann Ebbinghaus in Berlin in 1892. From 1897 to 1915 he taught philosophy and psychology at the University of Breslau, and in 1915 he moved to Hamburg, where, in 1919, he helped to found the University of Hamburg. He was forced into exile in 1933 by the Nazi government and became professor of psychology and philosophy at Duke University. He died in Durham, North Carolina.

As a psychologist Stern revolted against the elementarism (the belief in the adequacy of analysis of consciousness into its elementary parts) current in Germany before the general acceptance of Gestalt psychology. In his early studies of the perception of change and motion, he employed phenomenological methods and anticipated some later developments in Gestalt psychology. He soon gave up psychophysical experimentation, however, and pioneered in various fields of applied psychology, such as psychology of childhood, forensic psychology, intelligence testing (he introduced the concept of the intelligence quotient), and vocational psychology. Stern's work in psychology was always timely and often ahead of his times; he therefore earned a reputation as a psychologist that he never enjoyed as a philosopher, for most of his philosophizing was either opposed to, or out of touch with, contemporary movements. Some resemblance to *Lebens-philosophie* can be discerned, but he had little contact with Wilhelm Dilthey and his circle. Stern's philosophy must be understood in conjunction with his own psychological work, as providing the presuppositions for his lifelong scientific focus on the individual person—not on elements in his behavior and not on abstract universal laws relating them, but on the unique man. Even against Gestalt psychology, which likewise rejected elementarism, Stern's motto was: "No *Gestalt* without a *Gestalter.*" The *Gestalter* was the person.

Stern called his philosophy critical personalism to distinguish it from other personalistic theories, such as animism, vitalism, and Cartesianism, which were based upon the familiar dualism of mind and body. For him the person was an integral totality (*unitas multiplex*) whose defining property was purposive activity. What is not a person is a thing. A thing is not a whole but merely an aggregate; not autonomous but determined from without; not concretely individual but fragmentary or abstract. The person-thing distinction does not correspond to the mind-body distinction; rather, Stern held, the person is "psychophysically neutral," and both mind and body are thinglike abstractions from the original concreteness of a person sufficiently complex to be called an organism. Only some persons are conscious; indeed, only some of them are living. The person-thing distinction is repeated hierarchically, and the world is a system of persons included in and inclusive of others. A thing is a person seen from the standpoint of the supervenient person; that is, a person which includes other persons as parts.

With this conception, which suggests Aristotle, Gottfried Wilhelm Leibniz, and Gustav Fechner, Stern formulated his theory of teleomechanics as a way of avoiding an ontological dichotomy between teleology and mechanism. Mechanical uniformities, patterns of thing-behaviors, are derivative from teleological activities of supervenient personal beings in which the things are components. By this theory Stern attempted to derive the formal concepts and principles of the thing-world as we know it, such as magnitude, uniformity, class, causality, space, and time. By making these concepts and principles derivative, not fundamental, Stern's theory gave metaphysical priority to teleological and irreducibly individualistic notions.

Since the concrete substances of the world are teleological both as goal-setting and as goal-realizing, Stern identified the concept of intrinsic value with that of genuine, or personal, being. There are values corresponding to every level of person, indeed to every individual in the hierarchy of persons. But whereas in the theory of teleomechanism persons become things in the context of supervenient persons and thereby have at most extrinsic value, Stern later explored interpersonal relations in which the autonomy of each person is preserved and heightened through those relations which constitute a higher person. To the teleomechanical (cosmological) relation between persons Stern now added the introcep-

tive (axiological) relation, by which ends and intrinsic values of other persons as such are used by each person as factors in his own selfhood and autonomous self-determination and growth. In the formation of more inclusive and autonomous persons, the value of the whole suffuses the included persons with a radiative value (*Strahlwert*) instead of depersonalizing them as merely instrumentally valuable.

Stern's studies of love, religion, art, history, and ethics are deep and perceptive applications of his account of introception and radiative values. The theory of radiative value is especially fruitful in his accounts of symbolism and expression in many fields, and in his theory of introception he attempted to rationalize the value-oriented assessment of total personality characteristic of his psychology of individual differences.

Stern's personalism differs from that of personal idealism in that it is neither theistic nor idealistic, nor so radically pluralistic. It has closer resemblances to Jan Christiaan Smuts's holism and to some phases of Max Scheler's theory of value.

Bibliography

Works by Stern are *Person und Sache,* 3 vols. (Leipzig: Barth, 1906–1924); *Personalistik als Wissenschaft* (Leipzig: Barth, 1932); and *Allgemeine Psychologie auf personalistischer Grundlage* (The Hague: Nijhoff, 1936), translated by H. D. Spoerl as *General Psychology from the Personalistic Standpoint* (New York: Macmillan, 1938). An autobiographical essay may be found in Carl Murchison, *History of Psychology in Autobiography* (Worcester, MA, 1930), Vol. I, pp. 335–388.

Lewis White Beck (1967)

STEVENSON, CHARLES L.

(1908–1979)

Charles L. Stevenson authored the first thorough emotivist, or noncognitivist, account of ethical language. Traditionally the study of ethics had involved a quest for the truth about what is good and right, but Stevenson abandoned that search and set out to investigate the practical use of ethical language to shape attitudes. In a series of articles, and in his 1944 book *Ethics and Language*, he proposed answers to classical philosophical questions about meaning and justification that set the agenda for the next several generations of moral philosophers.

Stevenson earned degrees at Yale and Cambridge before receiving his doctorate from Harvard in 1935. He then taught at Harvard and Yale, where his original and challenging ideas about ethics were not popular. In 1946 he joined the philosophy department at the University of Michigan, where he remained till his retirement.

By the time *Ethics and Language* appeared, a form of emotivism had been sketched by A. J. Ayer, who claimed that ethical utterances are disguised commands and exclamations. Other students of ethics and language had introduced behavioral accounts of meaning, drawing attention to the actual use of moral language and questioning the place of reason in ethics. Stevenson's contribution was to integrate these ideas into a coherent theory and to emphasize the complexity and importance of the expressive function and the dynamic power of ethical language.

Disagreements in ethics, according to Stevenson, involve "an opposition of purposes, aspirations, wants, preferences, desires, and so on" (Stevenson 1944, p. 3). He called such disagreements "disagreements in attitude" and contrasted them with "disagreements in belief." Ethical disagreements can be resolved by rational argument when they can be traced to disagreements in belief, but when disagreements in attitude remain after agreement about the facts has been reached, rational means will be of no use. When rational means fail, Stevenson noted, and even when they do not, we resort to a variety of non-rational methods. Non-rational persuasion exploits language that carries what Stevenson called "emotive meaning." Emotive meaning "is the power that a word acquires, on account of its history in emotional situations, to evoke or directly express attitudes, as distinct from describing or designating them" (Stevenson, 1944, p. 33). Stevenson explored the many ways in which words with positive or negative emotive meaning can be used by speakers aiming to persuade others (or themselves) to alter (or preserve) some attitude.

Turning to the question of meaning, Stevenson argued that we can explain the meaning of an utterance such as *X is good* if we can find a relevant, similar expression that is free from ambiguity and confusion, and that allows us to do and say everything we can do and say with the original expression. By leaving out any mention of emotive meaning, a "subjectivist" definition such as *X is good = I approve of X* fails because it distorts the nature of ethical disagreement, which is fundamentally a clash of attitudes. Stevenson's suggestion, which he characterized as his "first pattern of analysis," was that any adequate analysis of *X is good* will satisfy the following pattern:

X is good = I approve of X, do so as well.

The first element (*I approve of X*) gives a subjectivist descriptive meaning and is but one example from a long list of candidates. The second (*Do so as well*) represents the emotive meaning and indicates that exposure to utterances like *X is good* tends to bring about approval for *X*.

According to a first-pattern analysis, one persuades by making a straightforward ethical judgment, counting on the emotive meaning of the key terms to influence the attitudes of the audience. A second method of persuasion is illustrated by a "second pattern of analysis." Many words carry strong emotive meaning, and just as we can influence attitudes by an explicit ethical judgment, so we can operate more subtly by exploiting what Stevenson called a "persuasive definition." When we give or use a persuasive definition, we attach a new descriptive meaning to a term like *courage* or *justice* while keeping the emotive meaning unchanged. The point of doing this is to change the direction of peoples' interests. As Stevenson says, "Words are prizes which each man seeks to bestow on the qualities of his own choice" (Stevenson 1944, p. 213) If we can redefine *courage* to cover our strategic retreat, then we too can be called courageous. "True courage," we might say, "is knowing when to run."

Stevenson observed that when our persuasion fits the first pattern, "attitudes are altered by ethical judgments," and when it fits the second pattern, attitudes "are altered not only by judgments but by definitions" (Stevenson 1944, p. 210). The two patterns turn out to be equivalent in the sense that "for every second pattern *definition* there is a first pattern *judgment*, the latter being the persuasive counterpart of the former" (Stevenson 1944, p. 229).

Stevenson's analysis of meaning had consequences for his view of another metaethical issue, the question of justification. When disagreement in attitude is not rooted in disagreement in belief, then the notion of a "reason" expands to include "any statement about any matter of fact which any speaker considers likely to alter attitudes" (Stevenson 1944, p. 114). This claim led some critics to accuse Stevenson of wanting to replace ethical reasoning with propaganda, but actually he claimed only that rational methods have limits and that persuasion is in play even when rational methods are used and even when we are trying to change or preserve our own attitudes. The choice of methods, he pointed out, is always a normative one, but he consistently identified his own study as a descriptive analytical one and refused to moralize about the ways of moralists.

In addition to his landmark works on metaethics, Stevenson wrote on aesthetics, music, and verse. He was a serious amateur musician, frequently performing chamber music with his friends and family.

See also Ayer, Alfred Jules; Emotive Theory of Ethics; Ethical Subjectivism; Metaethics; Noncognitivism.

Bibliography

WORKS BY STEVENSON

"The Emotive Meaning of Ethical Terms." *Mind*46 (1937): 14–31.

Ethics and Language. New Haven, CT: Yale University Press, 1944.

Facts and Values. New Haven, CT: Yale University Press, 1963.

"Persuasive Definitions." *Mind* 47 (1938): 331–350.

WORKS ON STEVENSON

Ayer, A. J. *Language, Truth and Logic*. London: Golancz, 1936.

Goldman, Alvin I. and Jaegwon Kim, eds. *Value and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt.* Dordrecht; Boston: D. Reidel, 1978.

Urmson, J. O. *The Emotive Theory of Ethics.* London: Hutchinson, 1968.

Richard T. Garner (1996, 2005)

STEWART, DUGALD
(1753–1828)

Dugald Stewart was an Edinburgh professor of moral philosophy who expounded the common sense theory of Thomas Reid and the libertarian political economy of Adam Smith. He taught from 1785 until illness forced his retirement in 1809. An eloquent spokesman for Reid and Smith rather than an original thinker, he left no legacy of his own but conveyed theirs. He provided his classes with a feast of psychology, ethics, and intellectual history and was the first professor in Britain to offer a course in political economy, which he began in 1800. A defender of academic freedom (see Brown [2004, 657] and Veitch [1858, lxxv–lxxix on the Leslie affair]), he both consoled and disturbed his audience by sustaining its metaphysical prejudices against Humean skepticism while revising its economic and political ones. He was no utilitarian yet advocated private liberty and the open market as the route to general happiness. His renown as a teacher was sustained by his books, which were translated into German, French, and Italian. He was honored by learned societies in Russia, Italy, and America, as well as by the Royal Societies of Edinburgh and London. Poet Robert Burns summed Stewart up as four parts Socrates, four parts Nathaniel, and two parts Brutus. He meant that

Stewart combined philosophical wisdom, a prophetic sense of morality, and a republican inclination.

Stewart's birth in Edinburgh on November 22, 1753 was in every sense an academic one. Not only was his father, Matthew, a college professor, but he was actually born in the college itself since their house was one of the college buildings. His father's family came from the southwest of Scotland where his grandfather was a minister. His mother was the daughter of an Edinburgh lawyer from whom she inherited the small Ayrshire estate of Catrine where the family spent the summer and where he befriended Burns whose home was at nearby Mossgiel.

Stewart attended the High School of Edinburgh where he learned Latin and Greek and the literature of both civilizations. He formed a lifelong attachment to the classics, a taste he shared with his revered friend Smith. In old age both philosophers turned to the early authors for pleasure and consolation, Smith to Sophocles and Euripides, Stewart to the Latin poets. He would later find this school education helpful in following the lectures of Adam Ferguson, whose class in moral philosophy he attended at the College in Edinburgh, which later became Edinburgh University. Ferguson was steeped in Roman history and literature, which formed the background to his lectures on moral and political philosophy and on civil society and its progress.

At the college, Stewart was introduced by John Stevenson, professor of logic and metaphysics, to the philosophy of John Locke, which was dominant at the time but which Stewart was to reject largely under the influence of Reid but also under that of Ferguson, who inspired his love of moral philosophy and whose chair he was to occupy. Before replacing Ferguson and after completing his college studies, Stewart had unexpectedly to take his father's place as professor of mathematics because illness forced his premature retirement. His father had achieved a minor international reputation as a Euclidean geometer although he was a reactionary who disdained algebraic geometry. He probably schooled his son informally in his own subject. Although Stewart was a good mathematician, he preferred philosophy, in which subject Ferguson discovered his talent.

Ferguson's philosophy was eclectic but principally Stoic. The classical moralists on whom he modeled himself advanced their own individual conceptions of virtue, of which they were taken to be exemplars. Assuming that moral philosophy is a kind of practical wisdom, their aim was to advise their students morally and lead them towards virtue. Stewart followed Ferguson's lead in adopting this ideal and in regarding right and wrong as like *primary* qualities, such as hardness, and not like the *secondary* qualities of colour and taste. With Ferguson and Reid, he criticized the school of moral sense led by the Lockean Francis Hutcheson, professor of moral philosophy in Glasgow and Smith's teacher. Hutcheson, followed by David Hume, said that virtue and vice are perceived through moral sensations of pleasure and pain or displeasure. Reason, they thought, is indifferent to virtue, which is only discovered by the responsive heart. Their critics—Ferguson, Reid, and Stewart—proposed, on the contrary, that humans use rational intuition to see which actions are morally right or wrong. These qualities exist independently of feeling and sensation. If the two sides did not agree about how virtue is perceived and why it is pursued, they did agree that the fundamental virtues are those of benevolence and justice.

Though no populist, Stewart managed to be more supportive of the idea of liberal reform than Ferguson. He agreed with Ferguson on the need for political leadership by wise philosophers, though he was quite clear about the citizen's right to political representation and clear that personal liberty is sacred. If the citizen is to be led, then it is to be out of servitude toward liberty. He was therefore deeply interested in the French liberal movement, which was headed intellectually by Anne-Robert-Jacques Turgot, François Quesnay, and Marie-Jean-Antoine-Nicolas Caritat, Marquis de Condorcet. They saw the nation's economy as the means of raising the standard of living of all its citizens. The movement was taken over by extremist deputies in the Assembly and culminated in violence against the throne. This was not the intention of the economists, who were not arguing for populist control but for rule by platonic philosophers guiding the monarch.

Stewart visited Paris in 1788 and 1789 and met some of the reformist thinkers, who encouraged his belief in the peaceful benefits of economic reform under wise government. He subsequently explained his innocuous views on political reform in *Elements of the Philosophy of the Human Mind* (1792; 1818, Vol. 1, 234–276). But this had an un-looked-for consequence because it led Scotland's judiciary to suppose that he actually supported violent revolution. Included among those were two judges known personally to Stewart who wished him to tone down his political writings. He declined to alter the second edition (1802) of the offending text, explaining his reason in a footnote. Although he sympathized with French liberalism and, unusually for someone of his position, with the American assertion of political and eco-

nomic independence, he rejected violence as an instrument of change.

Stewart went to Glasgow to hear Reid lecture in 1772 just before he took over as deputy for his father. As professor of moral philosophy, Reid was famous for his theory of common sense and his criticism of Hume's skepticism and the theory that ideas are copies of sensations. It was Reid's theory of belief, or laws of belief, as Stewart preferred to phrase it, that specially appealed, and he dedicated his first book, *Elements*, to Reid in 1792. Stewart felt that describing Reid's work as an inquiry into the principles of common sense suggested quite wrongly that it was not a philosophical theory about a philosophical matter: There is no room for theory if it is only common sense. According to Stewart—though Reid did much in showing that sensation cannot explain central beliefs in personal identity, the external world, the past and the future—Reid made no progress on René Descartes's position on proof of the existence of the external world: In other words, we can only trust to our beliefs, not prove them. To advance further, Stewart revives a suggestion he attributes to Father Ruggero Giussepe Boscovich the eighteenth-century Jesuit natural philosopher, that belief in external objects comes from the experience of their resistance. Stewart enlarges the suggestion with an idea from Turgot that, if experience suggests its cause, it is repetition of the experience that suggests the continuity of that cause (*Philosophical Essays*, chs. 1 and 2, 115–148). This account does not, he admits, completely prove that there are external objects but, rather, explains the belief as an expectation that what resists being touched or pushed will do so again because it continues to exist when it is not being felt.

As did the despised Lockeans, Stewart believed that the philosophy of mind is a science in which data are our sensations, our thoughts, and our volitions. It tries to analyze states of consciousness without either aspiring to understand the ultimate nature of mind or trying to explain all belief by sensation and feeling. We are not directly conscious of mind, nor are we of matter. Although we do not know what matter is, nor what mind is, we do know that there are two fundamentally different kinds of experience. One suggests matter, the other mind. To materialists who said that if we do not know what matter or mind are, they might be the same thing, he replied in a footnote in the first part of the introduction to *Elements*: if they were the same, "it would no more be proper to say of mind, that it is material, than to say of body, that it is spiritual" (p. 5). It did not occur to Stewart that, since it is improper to say of what is spiritual that it is material, if mind is matter, it would be improper to say that it is spiritual but not improper to say that it is material. It was inconceivable to him, though not to others such as David Hartley and Joseph Priestley, that mind might be located in the nervous system and the brain.

See also Condorcet, Marquis de; Descartes, René; Ethics; Ferguson, Adam; Hartley, David; Hume, David; Hutcheson, Francis; Locke, John; Philosophy of Mind; Priestley, Joseph; Reid, Thomas; Smith, Adam; Social and Political Philosophy; Socrates; Stoicism; Turgot, Anne Robert Jacques, Baron de L'Aulne.

Bibliography

WORKS

Collected Works of Dugald Stewart, edited by Sir William Hamilton. Edinburgh: Constable, 1858.

Philosophical Essays. 2nd ed. Edinburgh: Constable, 1816.

Elements of the Philosophy of the Human Mind. 1st ed., 1792. 6th ed. [London], 1818.

STUDIES

Brown, Michael P. "Stewart Dugald." In *Oxford Dictionary of National Biography*. Vol. 52., edited by H. C. G. Matthew and Brian Harrison. Oxford: Oxford University Press, 2004.

Veitch, John. "A Memoir of Dugald Stewart." In *The Collected Works of Dugald Stewart*. Vol. X., edited by Sir William Hamilton. Edinburgh: Constable, 1858

V. M. Hope (2005)

STILLINGFLEET, EDWARD

(1635–1699)

Edward Stillingfleet, an English Protestant theologian, was born in Cranborne, Dorset. He entered St. John's College, Cambridge, in 1649. On graduating in 1653 he was elected a college fellow, but after a year went into private employment. He was appointed rector of Sutton, Bedfordshire, in 1657. The Church of England was then under Presbyterian administration, but Stillingfleet received episcopal ordination in a clandestine ceremony and readily conformed after the restoration of the monarchy in 1660. A popular preacher in London legal circles, he became rector of St. Andrew's, Holborn, London, in 1665, and in 1678 rose to be dean of St. Paul's. On the accession of William III (1650–1702) in 1689 Stillingfleet was created bishop of Worcester. He was active in the politico-theological controversies of the time, most of which had a philosophical dimension. None of his writings was narrowly or exclusively philosophical.

His first work was *Irenicum* (1659). Though ostensibly an attempt to restore Protestant unity after several decades of sectarian divisions, it had a disguised episcopalian agenda. Stillingfleet resumed the debate with less disguise in the 1680s amid growing fears of a Catholic revival, publishing *The Mischief of Separation* (1680), *The Unreasonableness of Separation* (1681), and *Origines Britannicae* (1685). In *Irenicum* he allowed that episcopacy, presbytery, and independency could all point to precedents from the apostolic period; thus, all three could coexist compatibly. By 1685, however, he was arguing that the original English church had been an episcopal foundation, independent of Rome.

Stillingfleet's most consistent claim was that the primitive churches constituted a single society within each political state. Citing the authority of both natural and scriptural law, he portrayed the church of his own day as a subsociety operating within and compatibly with the laws of civil society, under which its members receive or lose privileges in proportion to their conformity. This was "latitudinarianism," a scheme that, by distinguishing essential from inessential matters, aimed to comprehend all believers in a national church and opposed the legal toleration of dissenting denominations. On matters not dictated by natural or revealed law—including the balance between episcopal and other forms—the overriding issue was one of civil peace, for which the civil administration was legislator. But many dissenters believed that there were theological issues here on which the civil power was incompetent to arbitrate. By the time of Stillingfleet's later writings against separation, there was a growing lobby in favor of the tolerationist alternative. John Locke prepared a critique of Stillingfleet in 1681 that survives in manuscript.

A second important early work, *Origines Sacrae* (1662), attempted to demonstrate the rational foundations of Judeo-Christian monotheism. Stillingfleet presented a detailed philosophy of history, exploring the nature of historical evidence and the grounds of assent to testimony. He claimed to establish the general superiority of written records over tradition and of the biblical record over ancient pagan history. On these principles he defended the authenticity of the biblical miracles, but not others, as confirming the authority of a revelation. Central to his argument was the concept of moral certainty. This was a genuine certainty attainable in matters beyond reasonable doubt by persons in possession of normal reason and of the evidence, where part of the function of reason is to judge the type of evidence appropriate to the context. By this means one can attain certainty in doctrinal matters that are above reason but not contrary to it. One's confidence is underwritten by the certainty one has of the existence of God.

This was a different kind of certainty based on clear and distinct ideas, yet compatible with the recognition that the object of certainty is largely incomprehensible. Part of the inspiration here was Cartesian, but Stillingfleet's enthusiasm for Cartesianism moderated in his last years after he absorbed Henry More's criticisms of René Descartes's cosmology and saw the direction taken by some post-Cartesian thinkers such as Benedict (Baruch) de Spinoza. In 1697 he was at work on a new *Origines Sacrae*, but only a fragment survives.

The epistemology developed in *Origines Sacrae* provided the basis for a relentless polemic against Catholic views of the rule of faith, from *A Rational Account of the Grounds of Protestant Religion* (1664) to *The Doctrine of the Trinity and Transubstantiation Compared* (1687), with many intervening titles. Stillingfleet appealed to weakly formulated principles of reason and common sense to reiterate his conviction that the doctrine of the trinity, being derived from a historically sound scripture, albeit above reason, was an assured certainty of faith; whereas that of transubstantiation, being contrary to reason and sense, was not. The Catholics argued for an exact parallelism and believed that the Protestants had no reliable arbiter in their disagreements about biblical interpretation.

By 1687 Stillingfleet had opened up the debate over the identification of substance and the distinction of persons. This was an opportunity for a growing Unitarian movement on the edge of Anglicanism to weigh in, seeking to demonstrate on clear and distinct principles that both the trinity and transubstantiation were equally indefensible and to promote a revisionist account of the atonement. Simultaneously with this, a rising tide of deism—religious belief based on natural reason alone without revelation—was beginning to pose awkward questions about the credibility of revelation.

Stillingfleet had already attacked Socinianism, a continental form of Unitarianism, in 1669 and deism in 1677, without obvious effect. Beset with opposition on so many fronts, he published *A Discourse in Vindication of the Doctrine of the Trinity* (1696). He incorporated an attack on John Toland's deistic *Christianity Not Mysterious* (1696), implicating Locke as the supposed inspiration for Toland's rejection of truths above reason. As a result, his final years were taken up with a highly public dispute with Locke, each side contributing three pieces. The dispute was over whether Locke's philosophy was capable of

supporting what Stillingfleet considered the basic propositions of the creed. Confused by Locke's Cartesian language about clear and distinct ideas, he challenged Locke to show how such ideas could come by sensation or reflection. Locke, he complained, had a "new way" of ideas, one that left him apparently ambivalent over mind-body dualism, agnostic about substance and essence, and unable to demonstrate immortality or to explicate the distinction of persons on his philosophy: in short, unable to bring any certainty to matters of faith. Locke gave no quarter to Stillingfleet in his replies, insisting on the coherence of his philosophy and its compatibility with biblical doctrine but refusing to be drawn into theological debate. Where, however, Stillingfleet had identified ill-chosen uses of the phrase "clear and distinct ideas" in Locke's *Essay concerning Human Understanding*, Locke silently amended them in the fourth edition (1700).

See also Cartesianism; Deism; Descartes, René; Locke, John; More, Henry; Revelation; Socinianism; Spinoza, Benedict (Baruch) de; Toland, John.

Bibliography

WORKS BY STILLINGFLEET

The Works of That Eminent and Most Learned Prelate, Dr. Edw. Stillingfleet, Late Lord Bishop of Worcester: Together with His Life and Character. 6 vols. London: Henry and George Mortlock, 1710. A consolidated reissue of separately titled works and collections published in a uniform format between 1707 and 1710.

Three Criticisms of Locke. Hildesheim, Germany: Georg Olms, 1987.

The Philosophy of Edward Stillingfleet. 6 vols. Bristol, U.K.: Thoemmes Press, 2000.

WORKS ABOUT STILLINGFLEET

Carroll, Robert Todd. *The Common-Sense Philosophy of Religion of Bishop Edward Stillingfleet, 1635–1699*. The Hague, Netherlands: Nijhoff, 1975.

Griffin, Martin I. J., Jr. *Latitudinarianism in the Seventeenth-Century Church of England*, edited by Lila Freedman. Leiden, Netherlands: E. J. Brill, 1992.

Hutton, Sarah. "Science, Philosophy, and Atheism: Edward Stillingfleet's Defence of Religion." In *Scepticism and Irreligion in the Seventeenth and Eighteenth Centuries*, edited by Richard H. Popkin and Arjo Vanderjagt. Leiden, Netherlands: E. J. Brill, 1993.

Popkin, Richard H. "The Philosophy of Bishop Stillingfleet." *Journal of the History of Philosophy* 9 (1971): 303–319.

Reedy, Gerard. *The Bible and Reason: Anglicans and Scripture in Late Seventeenth-Century England*. Philadelphia: University of Pennsylvania Press, 1985.

Schwitzgebel, Gottfried. *Edward Stillingfleet als Kritiker der Ideenlehre John Lockes*. Frankfurt am Main, Germany: Peter Lang, 2000.

Stewart, M. A. "Stillingfleet and the Way of Ideas." In *English Philosophy in the Age of Locke*, edited by M. A. Stewart. Oxford, U.K.: Clarendon Press, 2000.

M. A. Stewart (2005)

STIRNER, MAX

(1806–1856)

Max Stirner was the nom de plume of the German individualist philosopher Johann Kaspar Schmidt. Born in Bayreuth, Bavaria, Schmidt had a poor childhood. His academic career was long and fragmented. From 1826 to 1828 he studied philosophy at the University of Berlin, where he fell under the influence of G. W. F. Hegel. After brief periods at the universities of Erlangen and Königsberg, he returned to Berlin in 1832 and with some difficulty gained a certificate to teach in Prussian Gymnasiums. Several years of poverty and unemployment followed, until Schmidt found a position as teacher in a Berlin academy for young ladies run by a Madame Gropius. After this he lived something of a double life: The respectable teacher of young ladies had for another self the aspiring philosophical writer who assumed the name of Stirner.

The immediate stimulus that provoked Stirner to write his one important book, *Der Einzige und sein Eigentum* (Leipzig, 1845; translated by Steven T. Byington as *The Ego and His Own,* New York, 1907), was his association with the group of young Hegelians known as Die Freien (the "free ones"), who met under the leadership of the brothers Bruno and Edgar Bauer. In this company Stirner met Karl Marx, Friedrich Engels, Arnold Ruge, Georg Herwegh, and many other revolutionary intellectuals. In the same circle he also met Marie Dahnhardt, whom he married in 1843 and who left him in 1847. Before the publication of his book Stirner produced only a few brief periodical pieces, including an essay on educational methods printed by Marx in *Rheinische Zeitung.*

THOUGHT

Der Einzige und sein Eigentum, a treatise in defense of philosophic egoism, carried to its extreme the young Hegelian reaction against Hegel's teachings. In part it was a bitter attack on contemporary philosophers, particularly those with social inclinations. Stirner's associates among Die Freien were rejected as strongly as Hegel and Ludwig Feuerbach.

Stirner's approach was characterized by a passionate anti-intellectualism that led him to stress the will and the

instincts as opposed to the reason. He attacked systematic philosophies of every kind, denied all absolutes, and rejected abstract and generalized concepts of every kind. At the center of his vision he placed the human individual, of whom alone we can have certain knowledge; each individual, he contended, is unique, and this uniqueness is the very quality he must cultivate to give meaning to his life. Hence, he reached the conclusion that the ego is a law unto itself and that the individual owes no obligations outside himself. All creeds and philosophies based on the concept of a common humanity are, in Stirner's view, false and irrational; rights and duties do not exist; only the might of the ego justifies its actions.

There is much in common between Stirner's embattled ego and Friedrich Nietzsche's superman; indeed, Stirner was seen as a forerunner of Nietzsche during the 1890s.

Stirner has often been included with the anarchist philosophers, and he has much in common with them. However, he differs from writers like William Godwin, Pierre-Joseph Proudhon, and Pëtr Alekseevich Kropotkin in that the idea of a system of natural law, or immanent justice, which human law negates, is essential to their points of view. Stirner, however, rejected the idea of any such law, and in this respect he stands nearer to certain existentialists and the nihilists. Furthermore, while the anarchist seeks freedom as his ultimate goal, Stirner regarded such an aim as always being limited by external necessities; in its place he sought uniqueness or "ownness." "Every moment," he said, "the fetters of reality cut the sharpest welts in my flesh. But *my own* I remain."

Stirner agreed with the anarchists, however, in regarding the state as the great enemy of the individual who seeks to fulfill his "own will." The state and the self-conscious and willful ego cannot exist together; therefore the egoist must seek to destroy the state, but by rebellion rather than by revolution. This distinction is essential to Stirner's doctrine. Revolution, in overthrowing an established order, seeks to create another order; it implies a faith in institutions. Rebellion is the action of individuals seeking to rise above the condition they reject; it "demands that one rise, or exalt oneself." Revolution is a social or political act; rebellion is an individual act, and therefore appropriate to the egoist. If rebellion prospers, the state will collapse.

In rebellion the use of force is inevitable, and Stirner envisaged "the war of each against all," in which the egoist fights with all the means at his command. This viewpoint led Stirner to justify and even to exalt crime. Crime is the assertion of the ego, the rejection of the sacred. The aim of egoist rebellion is the free wielding of power by each individual.

In Stirner's view the end of this process is not conflict but a kind of dynamic balance of power between men aware of their own might, for the true egoist realizes that excessive possessions and power are merely limitations on his own uniqueness. His assertion is based on the absence of submissiveness in others; the withdrawal of each man into his uniqueness lessens rather than increases the chance of conflict, for "as unique you have nothing in common with the other any longer, and therefore nothing divisive or hostile either." Stirner argued that far from producing disunity among individuals, egoism allows the freest and most genuine of unions, the coming together without any set organization of the "Union of Egoists," which will replace not only the state with its political repression but also society with its less obvious claims.

LATER YEARS

Der Einzige und sein Eigentum is not just a most extreme expression of individualism, it is also the single manifestation of Stirner's own revolt against a frustrating life that finally submerged him. In his totally undistinguished later years he embarked on a series of unsuccessful commercial ventures and translated English and French economists. His remaining book, *Die Geschichte der Reaktion* (Berlin, 1852), lacked the fire of discontent that made his earlier work so provocative. Stirner's last years were shadowed by declining powers and haunted by creditors; he died poor and forgotten in 1856.

Bibliography

For further information on Stirner, see Victor Basche, *L'individualisme anarchiste: Max Stirner* (Paris: Alcan, 1904); James Gibbons Huneker, *Egoists* (New York, 1921); and John Henry Mackay, *Max Stirner, sein Leben und sein Werk* (Berlin: Schuster and Loeffler, 1898).

George Woodcock (1967)

STÖHR, ADOLF
(1855–1921)

Adolf Stöhr, the Austrian philosopher, psychologist, and linguist, was born at St. Pölten and studied law and philology, then botany, and finally philosophy, at the University of Vienna. In 1885 he was appointed *Privatdozent* in theoretical philosophy at the same university, rising to associate professor in 1901 and to full professor of the philosophy of the inductive sciences in 1911. He pub-

lished some thirty works in logic, natural philosophy, psychology, and philosophy.

LANGUAGE AND THOUGHT

Stöhr developed his system of logic in the closest connection with the psychology of thought processes and linguistics. His work deals in great detail with the dependence of thought upon language (what he calls the glossomorphy of thought), and he warned against the dangerous consequences that flow from confusing forms of speech with forms of thought. Not only do we make use of language to fix our thoughts and to communicate our knowledge; we also think in our language, so that the structure of our thought reflects the logical forms of our language. When the course of thought becomes automatic, the result may be that self-critical thought is replaced by an "idle flow of speech" ("glossurgy"), which is frequently even self-contradictory.

Through such reflections Stöhr began the "critique of language" pursued later with such success by other important thinkers. With the aid of this critique, he sought above all to oppose the misuse of language in philosophy and to unmask the muddled philosophical thinking that gives rise to the reification of concepts, metaphors, and allegories. Because "our language compels us to designate consciousness as if it were constructed of a subject, of mental acts and of physical objects" (as in the sentence "I see an object"), the illusion arises that "thoughts have the form (*morphe*) of the language (*glossa*)." The final outcome is that fictions are taken for facts; metaphors, for that which is actually meant. Thus the fact of the psychological "I" is confused with the fiction of the mental "subject," and the fact of phenomenal matter as a complex of visual and tactile sensations is confused with the materialistic fiction of a metaphysical matter (*Wege des Glaubens,* pp. 20ff.).

METAPHYSICS

Stöhr distinguished three roots of metaphysical thinking: wonder at the facts (the "theorogonous" metaphysics of the "constructing imagination"); pain (the "pathogonous" metaphysics of the "suffering heart"); and glossomorphic confusion (the "glossogonous" metaphysics of the "rolling word"). Metaphysics can supply no universally valid knowledge because the transcendental is in principle unknowable; one can only "have faith" in the existence of something beyond experience. This metaphysical faith is the expression of a subjective reaction of the heart and is "lived." Knowledge cannot engender faith, and faith cannot substitute for knowledge; for the two are of an entirely different nature" ("Ist Metaphysik möglich?," p. 30). "Everyone proceeds along that path of faith which his whole constitution obliges him to take. There is neither an inductive nor a deductive proof for or against a faith" (*Wege des Glaubens,* p. 36).

Stöhr rejected both "pathogonous" and "glossogonous" metaphysics, and thus the whole of metaphysics in the traditional sense, with its claim to knowledge of the transcendental. Anyone who pretends to provide such knowledge is philosophizing both "pathogonously" and "glossogonously." Anyone who is unable to find the meaning of life in life itself, in the work and tasks of life, and therefore suffers in being alive, seeks that meaning beyond the world and life. Since he would like to convince others of the truth of his outlook on life and the world, which is directed to the beyond, he intentionally or unintentionally misuses language in order to offer rhetorical pseudo solutions to metaphysical pseudo problems as if they were genuine solutions to real problems.

Stöhr himself professed "theorogonous" metaphysics. He defined it as "the satisfaction of an artistic propensity by means of the elegant construction of a world view"—which, of course, must not contradict the facts. "Thus metaphysics, in contrast to the empirical sciences, does not grow through apposition, but continuous building, rebuilding and building anew" (*Lehrbuch der Logik,* p. 304). Stöhr constructed his own view of nature in this manner, not dogmatically but as an exercise, assigning more importance to the creation than to the validity of a system. (He often said in discussion: "I am only playing with these ideas. I do not say that this is the way things are. I do not say even that this is the way they probably are. All that I say is that this is the way they may be.")

NATURAL PHILOSOPHY

Stöhr attempted to explain the structure of matter and the peculiarities of organic happenings in conformity with his undogmatic approach. Since for him mechanism was the sole intelligible conception of nature, he sought to understand both the organic world and the inorganic world with the help of mechanistic conceptual models. Stöhr proved to be as original a thinker in the philosophy of nature as in logic and psychology. That many of his ingenious solutions to problems have become outmoded by the progress of the sciences does not alter the epistemological excellence of his clear and exact style of thought.

Bibliography

Stöhr's major works include *Umriss einer Theorie der Namen* (Leipzig and Vienna: Deuticke, 1889); *Alegbra der Grammatik* (Leipzig: Deuticke, 1898); *Philosophie der unbelebten Materie* (Leipzig, 1907); *Der Begriff des Lebens* (Heidelberg: Winter, 1909); *Lehrbuch der Logik* (Leipzig: Deuticke, 1910); "Ist Metaphysik möglich?" in *Jahrbuch der philosophischen Gesellschaft an der Universität zu Wien 1914 und 1915* (Leipzig, 1916), pp. 25–36; *Psychologie. Tatsachen, Probleme und Hypothesen* (Vienna and Leipzig, 1917; 2nd ed., Vienna and Leipzig: W. Braumüller, 1922); and *Wege des Glaubens* (Vienna and Leipzig, 1921).

Works on Stöhr include Franz Ferdinand Worlitzky, *Über die Philosophie Adolf Stöhrs,* a dissertation (Vienna, 1925), and Franz Austeda, "Der Oesterreicher Adolf Stöhr—einer der bedeutendsten Denker unseres Jahrhunderts," in *Neue Wege* (Vienna) (104) (1955): 3–5.

Franz Austeda (1967)
Translated by Albert E. Blumberg

STOICISM

Stoicism was a philosophical movement founded in Athens in the late fourth century BCE by Zeno of Citium. Although Stoicism was shaped by many philosophical influences (including the thought of Heraclitus), it was throughout its history an essential part of the mainstream Socratic tradition of ancient philosophy. Inspired as well by the Cynics (Zeno was taught by Crates, a student of Diogenes of Sinope), Stoicism developed alongside and in competition with Platonism and Aristotelianism over the next 500 years. For centuries it was the main rival to Epicurean thought as well. Virtually no works survive from the early period of the school's history. Yet its doctrines have been reconstructed with a fair level of reliability on the basis of later accounts, critical discussions by non-Stoics, and the surviving works of later Stoic writers.

HISTORY

When Zeno arrived in Athens, attracted from his home on Cyprus by Socratic philosophy, Plato's Academy was led by Polemo and was soon to make its historic shift away from what we now recognize as Platonism toward a form of skepticism under the leadership of Arcesilaus. Aristotle's legacy was still in the hands of Theophrastus, head of the Lyceum, though in the third century BCE the school would decline in philosophical power as it concentrated on more narrowly scientific problems. Nevertheless, the Aristotelian drive for broad-based philosophical synthesis had an impact on the shape of Stoicism. A significant group of philosophers, forming no particular school but many coming from nearby Megara, concentrated on dialectic as their principal activity. These included Stilpo, also interested in ethics and metaphysics, and Diodorus Cronus, whose sharply formulated arguments provided powerful challenges in physics and metaphysics and challenged the Stoics to develop dialectic as a central part of their system. The Cynics in turn championed nature (as opposed to narrow polis-based social norms) as the foundation of ethics. All of this contributed to Zeno's formation of a powerful philosophical system whose internal articulation into three parts (logic, physics, ethics) was inspired by the Academic Xenocrates.

Stoicism was named for Zeno's favorite meeting place, the Painted Stoa in the Athenian marketplace. The movement was concentrated in a formal philosophical school in Athens for more than 200 years until political changes resulting from Rome's rise to power led prominent philosophers to spread out around the Mediterranean world, especially to Rhodes, Alexandria, and Rome itself. The climax of this process came when the Roman general Sulla sacked Athens in 86 BCE during the Mithridatic Wars. By the end of the first century BCE, Stoic activity was widely dispersed and had become a central part of intellectual culture in the Greco-Roman world. In the early second century CE, the emperor Hadrian founded a chair of Stoic philosophy in Rome (as well as chairs for the other major schools). With the rise of Neoplatonism, Stoicism gradually faded in prominence, though its influence persisted until the end of antiquity. Its impact on medieval philosophy was sporadic, but in the Renaissance it became an important part of the philosophical legacy of the ancient world to modern philosophy.

PRINCIPAL STOICS AND THEIR WORKS

The founder of the school, Zeno, was a prolific author whose best-known work was his utopian *Republic*, influenced by his Cynic teachers and by Plato's *Republic*. He wrote extensively on ethics and politics (e.g., *On the Life according to Nature*; *On Law*; *On Human Nature*; *On Passions*; *On Greek Education*), on cosmology (*On the Universe*), on poetry (*Homeric Problems*; *On Listening to Poetry*), and on dialectic (*On Signs*; *Refutations*,; *Solutions*). Of his many students, some (Persaeus and Sphaerus) also involved themselves in politics. Cleanthes was a highly prolific writer in the areas of cosmology, physics, ethics, and dialectic. He was also known for his poetry, especially the *Hymn to Zeus* (which has survived entire) and for his interest in Heraclitus. Cleanthes' contemporary Aristo of Chios favored the Cynic side of the

school's heritage and rejected physics and dialectic in favor of a teaching based solely on ethics. Though eclipsed by Cleanthes (who succeeded Zeno as head of the school) and Chrysippus (the third head of the school), Aristo's influence continued to be felt at least until the first century CE.

Chrysippus, the great systematizer of the Stoic tradition, put the school's doctrines on a solid footing after a long period of debate and criticism, especially by the Academic Arcesilaus. Respected as a second founder of the school, he and his students dominated its leadership for many decades. He argued that Zeno's philosophy (as he interpreted it) was essentially correct and thereby stabilized the essential doctrines of the school, which nevertheless continued to be open to internal debate. A highly prolific author (more than 700 books are attributed to him and a partial catalog survives in book 7 of Diogenes' *Lives*), Chrysippus revised and rounded out the areas of physics and ethics and put dialectic, especially the study of formal inference and the theory of language, on a new foundation. He wrote a work in defense of Zeno's *Republic*, evidently declining to abandon the school's Cynic roots, a large number of works on logic and dialectic (including *Logical Investigations*, of which a few fragments have survived among the Herculaneum papyri), and a nearly equal number on logic and physics. The best attested work is certainly his *On Passions*, from which Galen quotes many passages in the course of his criticism of Stoic views on psychology and ethics.

The next phase in the school's history came in the late second and early first centuries BCE, when Panaetius of Rhodes and subsequently Posidonius of Apamea adopted a more open stance toward Platonic and Aristotelian approaches than seems to have been characteristic of Chrysippus. There was, however, no dramatic departure from the earlier school. Prominent among later Stoics is Seneca the Younger, a Roman politician of the first century CE. Many of his works, including the *Moral Epistles to Lucilius*, were highly influential in the early modern period. Other works of Seneca's include *On Benefits* (which offers important arguments in ethics) and *Natural Questions* (on physics and meteorology). His works form the earliest corpus of Stoic writing that has survived to the modern era. Another Stoic was Epictetus, a prolific writer and teacher, mostly of ethics, in the late first century CE. He owed a great deal to Musonius Rufus, a Roman citizen from Etruria who wrote in Greek in the early first century CE. Epictetus's lectures were very influential in later antiquity and the early modern period; this is especially true of his *Handbook*, a compendium drawn from the *Discourses*, which in turn was compiled by his student Arrian from his lectures. The emperor Marcus Aurelius left a set of personal philosophical reflections, *To Himself*, more commonly titled *Meditations*. In no sense a professional philosopher, Marcus combines a profoundly Stoic point of view, deeply influenced by Epictetus, with a more generalized "philosophical" stance reflecting influences from many traditions.

CENTRAL IDEAS

The concept of nature played a central role in Stoicism. The key to human fulfillment or happiness (*eudaimonia*) is living according to nature, and Stoic philosophy was based on this conception of the goal of life. The study of the natural world, physics, was a major occupation of virtually all Stoics (Aristo of Chios being a notable early exception). Human nature for the Stoics is characterized by a rationality that, when fully developed, is divine in its perfection. A deep expression of our nature and of that of the cosmos is our capacity for logic. Nature was formally defined as "a craftsmanlike fire, proceeding methodically to creation (*genesis*)" (Diogenes 7.156). God, a fully rational and providential force causally responsible for the world and its orderliness, was equated with nature. Whereas the divine craftsman of Plato's *Timaeus* stood outside the physical cosmos, the rational creator god of Stoicism is completely immanent in the material world.

The Stoics, more than any other ancient school, emphasized the interdependence among the parts of philosophy. They used various similes to illustrate the point. Philosophy is like an animal—logic is the bones and sinews; ethics the flesh; physics the soul. Or it is like an egg—logic is the shell; ethics the white; physics the yolk. Or like agricultural land—logic is the wall around the field; ethics the fruit; physics the land or trees that bear the fruit. Ideas varied about the ordering and relative importance of the three parts and their subdivisions, but all agreed that philosophy, when properly taught, demanded an intimate blend of all three disciplines, regardless of the pedagogical order chosen (Diogenes 7.39–41).

The Stoics based all areas of their thought on a rigorous metaphysical principle that sharply distinguished the corporeal and the incorporeal. The key to this distinction is the argument that only bodies can interact causally, an argument that seems to have emerged from a critique of Plato's metaphysics. Hence god, the soul, nature, and the principles that organize raw matter into intelligible natural kinds are all forms of matter for the Stoics. Even cognitive states such as knowledge are treated

as corporeal dispositions of the material mind, since they have causal impact; so too for virtue and other dispositions. Their theory of perception similarly posits corporeal entities, lending weight to their essentially empiricist epistemology. The Stoics recognized only four incorporeal entities: void, space, time, and "sayables" (*lekta*, roughly, the meanings of thought and speech). Each of these incorporeal entities is parasitic on bodies, a necessary feature of the world but in itself causally inefficacious.

In ethics the central concept was virtue, understood in a distinctively Stoic manner. Human life has a single goal (*telos*): to live according to nature. Following Aristotle, the Stoics called achieving this goal "happiness" (*eudaimonia*). Perfection of our intrinsically rational nature is the only way to do this. This perfection, which they called "virtue" (*aretē*), is the necessary and sufficient condition for achieving our goal. This robust conception of virtue is at the center of Stoic thought and became the defining feature of the school.

LOGIC

Stoic logic has two parts: dialectic and rhetoric. Dialectic is broader in scope than logic in the modern sense. Yet the Stoics made crucial advances even in logic understood in the narrower modern sense.

Traditionally, rhetoric had been the art of persuasion through speech. As such it was either condemned, as by Plato, or reformed, as by Aristotle. The Stoics restricted rhetoric by insisting that it, like other crafts, must be conducted under norms of truth and virtue. Hence rhetoric became the art of persuading an audience of the truth through orderly discourse and argument, differing from dialectic only in form; rhetoric is merely a more expansive way of achieving such conviction. As Zeno said, rhetoric is an open hand, while dialectic is a closed fist (Sextus 1935, 2.7 [= *Adv. Mathematicos* 2.7]). Stoic ideas about rhetoric understandably had limited influence.

In contrast, their dialectic had considerable influence, since it aimed to be a comprehensive study of human discourse and its relation to truth about the world. It covered the content of discourse as well as the utterances that express that content, both what is signified and what does the signifying. The relationship between linguistic signifiers and their meaning lies at the heart of Stoic dialectic. Accordingly, dialectic covered much of what we classify as epistemology and philosophy of language (including semantics), as well as the study of propositions and their relations. But since what is signified by speech are incorporeal sayables, dialectic also included aspects of metaphysics and philosophy of mind. The broad Stoic conception of dialectic also covered what we would consider linguistics and grammar, the parts of speech and various forms of speech acts; their theories had great influence on the development of grammar as a discipline.

In perception, on the Stoic theory, we receive through the senses representations of objects and events. A rational animal becomes aware of this representational content by way of a sayable (usually a proposition [*axiōma*], defined as what admits of being true or false), which is dependent on the physical change in the mind. We either assent to this proposition, reject it as being unrepresentative of its alleged correlate in the world, or suspend judgment about its truth. This is the heart of Stoic epistemology. Academic critics of the Stoic theory argued that no sensory representation could be satisfactorily reliable. In defending their theory (in part by positing self-verifying cataleptic representations) and in elaborating how perceptual experience formed the basis for concepts, memories, and the like, the Stoics expanded on the foundations for empirical epistemology that Aristotle had laid.

The most important aspect of Stoic logic is its study of the forms of argument, inference, and validity. Stoics undertook this to defend the truth of their substantive doctrines and to demonstrate the pervasiveness of rational structures in the world. Chrysippus went beyond that goal and plunged into had been the starting point, and the subject had been advanced by the development of challenging paradoxes and puzzles by Megarian and other dialecticians. Chrysippus made the logic of propositions and arguments into a discipline.

Stoic logic takes the proposition (*axiōma*, often symbolized by an ordinal number) as its basic unit of analysis and works with a small set of operators used to connect them: "if," "and," "not," and exclusive "or." Five basic inference forms were recognized; all valid arguments were supposed to be derivable from these indemonstrable arguments by purely logical means. Stoics attempted to prove this completeness claim with the aid of higher-order logical principles. The five indemonstrables are the following:

If the first, the second.
But the first.
Therefore, the second.

If the first, the second.
But not the second.
Therefore, not the first.

Not both the first and the second.
But the first.
Therefore, not the second.

Either the first or the second.
But the first.
Therefore, not the second.

Either the first or the second.
But not the second.
Therefore, the first.

PHYSICS

Stoic physics was, in its day, the most up-to-date and influential version of the nonatomistic physics pioneered by Empedocles and developed by Aristotle. Stoics posited a geocentric cosmos made up of earth, air, fire, and water arranged in four roughly concentric spheres. Although the cosmos has no void within it, it is surrounded by an indefinitely large void, which provides room for expansion when the cosmos reaches the end of its finite life span. The Stoics held that the cosmos was generated by the creative intelligence of Zeus and eventually ends by returning to the fire from which it was born. This process repeats itself forever—a doctrine that responds in part to Aristotle's arguments for the eternity of the cosmos. Since things expand when heated, the conflagration that occurs at the end of each cycle requires that there be empty space outside the physical world.

Zeus is a craftsman-god modeled on the creator god of Plato's *Timaeus* and initially identified with a kind of fire. Cosmogony begins when this fire transforms itself in a quasi-biological process that generates the four elements that are the stuff of the world. Fire has a dual role, both as the original divine source and as one of the four elements. Each element is analyzable into two principles, the active and the passive, but these principles are themselves corporeal. The active principle (like Aristotelian form) is immanent everywhere and is responsible for the structure and comprehensibility of things; hence it is often identified as god and reason, a creative form of fire that embodies a divine plan for every aspect of the physical world. This emphasis on unified and immanent divine power made the Stoics pioneers for later forms of pantheism.

Later Stoics (including Chrysippus) revised the role of fire and claimed that the immanent shaping power was better understood as *pneuma*, a unique blend of fire and air with an optimal combination of fluidity and tensile strength. *Pneuma* gives order and shape to things in varying degrees. In lifeless things like rocks it is a disposition (*hexis*), giving them coherence and shape. In plants it is their "nature" (*phusis*) and accounts for their ability to grow and change. In animals it accounts for the full range of dynamic attributes, including perception and desire; hence it is there called "soul" (*psuchē*). In humans and gods this divine shaping power is labeled "reason" (*logos*). These various forms of a single power unite all entities into a single order, the cosmos. Since both the active shaping power and the passive component of a thing are corporeal, the Stoics had to give an account of how two such bodies could be fused into a perfect mixture. Their sophisticated theory of "total blending" was frequently criticized, but the concept of *pneuma* itself had considerable influence in later centuries.

The Stoics analyzed each individual entity by means of a complex theory that today would fall under the heading of metaphysics. They posit four "genera" or kinds (less helpfully, "categories"), all of which apply to every object. First, each object can be treated as a "substrate"; this merely asserts that it is a material object, a being, without specifying its attributes. Second, each object is "qualified," endowed (by the active principle or by *pneuma*) with structure sufficient to make it a definite thing. Qualities are either common (making the object a kind of thing, such as a human) or peculiar (making it a unique individual, such as Socrates). The third genus specifies dispositions or conditions of an entity (Socrates may be courageous or have frost-bitten feet), while the fourth is termed "relative disposition" and picks out relations such as being the father of someone or being on the right of someone. Though we cannot be certain of all its details, this theory clearly provided the analytical framework for Stoic corporealist physics.

Since the cosmos is a whole united by reason (i.e., the *pneuma* that pervades it), it can be regarded as a single living entity. In this perspective, everything else is a part of the whole, even humans, whose reason is the same in nature as that of Zeus. Hence humans are uniquely situated in the world, subordinate to it as parts but able to understand in principle the unified plan determining all that happens.

From a theological perspective, this plan appears as a providential divine arrangement, but in Stoic physics, it is actually a mere consequence of Stoic causal determinism. There are no uncaused events, so all that happens is determined by antecedent events and states of affairs in the world. The world, then, is a network of causal relationships capable in principle of being explained. If this were not the case, there would be uncaused events, which Stoics thought unacceptable; even the principle of biva-

lence (the claim that every proposition is either true or false) would be threatened, and Chrysippus (contrary to Aristotle and Epicurus) held that this logical principle obtains even for future-tense propositions.

Human thoughts, actions, and decisions are a part of this causally deterministic system, but moral responsibility is not threatened (according to the Stoics), since the decisive causal factor is the character and disposition of the agent as he or she reacts to the world. Critics in the ancient world argued that causal determinism jeopardized moral accountability, but Chrysippus stoutly maintained a distinction between being caused (as human actions are) and being necessitated by factors wholly external to the agent. Stoic compatibilism still seems reasonable to many philosophers, but it remained contentious in the ancient world.

ETHICS

It is tempting to suppose that for the Stoics ethics is the most important branch of philosophy, subserved by logic and physics. But of all the similes used to described the relationship among the parts of philosophy, only two support this claim: Posidonius's assertion that ethics is like the soul of an animal (Sextus 1935, 1..19 [= *Adv. Mathematicos* 7.19]) and the claim that ethics is like fruit on the trees (Diogenes 7.39–41). Other Stoics make physics the culmination of philosophical activity. Three factors incline us to regard ethics as the core of Stoic thought: the pattern of ancient philosophical controversy, the accidental bias of the surviving sources, and the fact that Stoic physics is today more obviously obsolete than Stoic ethics. To yield to this tendency is to take sides in a debate within the ancient school, to support the Socratic mission of Aristo of Chios against, for example, Chrysippus, who regarded theology (part of physics) as the culmination of philosophy (Plutarch 1035a).

Philosophy is a craft for living (*technē tou biou*). As a craft, it is based on a body of knowledge, consists in a stable disposition of a rational agent, and has a determinate function (*ergon*) and goal (*telos*). Stoicism is firmly embedded in the eudaimonistic tradition of ancient ethics, where the goal is *eudaimonia*, conventionally translated as "happiness." For Stoics, the goal is to live in accordance with nature, and their claim is that this consists in living in accordance with virtue, since human virtue is the excellence of our nature. But our nature is fundamentally rational. Hence perfection of human reason is another summary expression of the goal. This remains a merely formal account until substantive Stoic views about human nature are considered. In contrast to Plato and Aristotle, Stoics denied that the mature human soul contains essentially irrational components. In Stoic thought, there is no lower part of the soul to be tamed and managed by reason; rather, our rational faculties have an affective component, and so emotion and desire are features of some of our cognitive processes. Further, the Stoics held that our rational nature is qualitatively the same as the divine reason embedded in nature, so that our goal requires living in accordance with both human nature and cosmic nature (Diogenes 7.88).

Like all living things, humans are shaped by a fundamental drive to preserve and enhance their nature, a drive visible even in infants but taking on its characteristic form when they mature. This basic drive involves a commitment to pursue the good, understood as what is truly beneficial. Stoics accept the Socratic argument that only virtue is consistently and genuinely beneficial, since an excellence cannot be misused. Other advantageous things (health, pleasure, social standing, etc.) admit of misuse, so their value is merely provisional. They are preferred but not good. There is a similar account of vice (the only truly bad thing) and disadvantageous things like disease and poverty, which are dispreferred but not genuinely bad. This basic duality in Stoic value theory is a central feature of Stoic ethics. Though it is rational to avoid dispreferred things and embrace preferred things in the course of a well-planned human life, only genuine goods demand unconditional commitment.

This is the basis for the notorious Stoic rejection of passions, which are understood as unreasonable and excessive reactions to preferred and dispreferred things. If sickness and poverty are not bad but merely dispreferred, we should not grieve over them (but, of course, we should do our best to avoid them). If wealth is not a strict good, we should not be elated at achieving it (though there is nothing wrong with enjoying it). If a favorable reputation in our community is not an unconditional good, then we need not fear losing it. If romantic attachments are worth having but are not the sine qua non of human flourishing, then we should pursue potential partners without obsession. And so forth. Life according to our purely rational nature will be free of passions, but not devoid of affect. For in a life of virtuous choices and actions, there will be many things to want, to shun, and to rejoice over. Such positive affective states were called *eupatheiai*.

Most Stoics accepted the doctrine of the unity of virtues, though there was serious debate about the nature of that unity. But all Stoics held that virtuous action was limited to the sage—a normative ideal of perfected virtue used as a benchmark for good action. The Stoics distin-

guished between appropriate actions (*kathēkonta*), which can be determined by the proper application of moral guidelines and maxims, and genuinely good actions (*katorthōmata*), which are appropriate actions performed from the perfected disposition of a sage. Nonsages may have little real chance to attain wisdom, but their constant striving to determine the appropriate thing and to do is guided by the ideal of the sage. Stoic recommendations for appropriate actions (such as participation in civic life, unless it is hopelessly corrupt) are routinely presented as descriptions of what the sage will do, yet Stoicism does not categorically prescribe any particular actions. Only the commands to follow (or accommodate oneself to) nature and to act virtuously are unconditional.

Stoic ethics is often portrayed as mired in paradox, but we can make better sense of the persistent philosophical appeal of Stoicism if we focus instead on Stoics' stringent and carefully formulated theories in all branches of philosophy and their insistence that these parts should fit together into a coordinated whole, that they should combine the best understanding of the natural world available in their day with a deep commitment to the exercise of human reason as the key to human fulfillment.

See also Arcesilaus; Aristotelianism; Chrysippus; Cleanthes; Cynics; Diodorus Cronus; Epictetus; Epicureanism and the Epicurean School; Greek Academy; Heraclitus of Ephesus; Marcus Aurelius Antoninus; Musonius Rufus; Panaetius of Rhodes; Posidonius; Seneca, Lucius Annaeus; Zeno of Citium.

Bibliography

WORKS BY STOICS

Diogenes Laertius. *Lives of Eminent Philosophers*. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1972.

Epictetus. *The Discourses of Epictetus*. London: J. M. Dent, 1995.

Epictetus. *Handbook of Epictetus*. Translated by Nicholas P. White. Indianapolis, IN: Hackett, 1983.

Marcus Aurelius. *Meditations*. Translated by A. S. L. Farquharson. New York: Knopf, 1992.

Plutarch. "On Stoic Self-Contradictions." In his *Moralia*. Vol. 13, Pt. 2: *Stoic Essays*, translated by Harold Cherniss. Cambridge, MA: Harvard University Press, 1976.

Seneca, Lucius Annaeus. *Seneca's Letters to Lucilius*. Translated by E. Phillips Barker. Oxford, U.K.: Clarendon Press, 1932.

Sextus Empiricus. *Against the Logicians*. Translated by R. G. Bury. Cambridge, MA: Harvard University Press, 1935.

Sextus Empiricus. *Against the Physicists*. *Against the Ethicists*. Translated by R. G. Bury. Cambridge, MA: Harvard University Press, 1936.

Sextus Empiricus. *Against the Professors*. Translated by R. G. Bury. Cambridge, MA: Harvard University Press, 1949.

Sextus Empiricus. *The Skeptic Way: Sextus Empiricus's "Outlines of Pyrrhonism."* Translated by Benson Mates. New York: Oxford University Press, 1996.

WORKS ON STOIC PHILOSOPHY

Algra, Keimpe, et al. *Cambridge History of Hellenistic Philosophy*. Cambridge, U.K.: Cambridge University Press, 1999.

Becker, Lawrence C. *A New Stoicism*. Princeton, NJ: Princeton University Press, 1998.

Bobzien, Susanne. *Determinism and Freedom in Stoic Philosophy*. Oxford, U.K.: Clarendon Press, 1998.

Frede, Michael. *Die stoische Logik*. Göttingen, Germany: Vandenhoeck und Ruprecht, 1974.

Hahm, David E. *The Origins of Stoic Cosmology*. Columbus: Ohio State University Press, 1977.

Inwood, Brad, ed. *The Cambridge Companion to the Stoics*. Cambridge, U.K.: Cambridge University Press, 2003.

Inwood, Brad. *Ethics and Human Action in Early Stoicism*. Oxford: Clarendon Press, 1985.

Long, A. A., ed. *Problems in Stoicism*. London: Athlone Press, 1971.

Long, A. A., and D. N. Sedley. *The Hellenistic Philosophers*. Chaps. 26–67. Cambridge, U.K.: Cambridge University Press, 1987.

Mates, Benson. *Stoic Logic*. Berkeley: University of California Press, 1953.

Rist, John M., ed. *The Stoics*. Berkeley: University of California Press, 1978.

Sandbach, F. H. *The Stoics*. London: Chatto and Windus, 1975.

Striker, Gisela. "Following Nature: A Study in Stoic Ethics." *Oxford Studies in Ancient Philosophy* 10 (1996): 1–73.

Brad Inwood (2005)

STOUT, GEORGE FREDERICK
(1860–1944)

George Frederick Stout was an English philosopher and psychologist. Records of Stout's early life are scant. He was born in South Shields, Durham. A clever boy at school, he went in 1879 to St. John's College, Cambridge, where he obtained first-class honors in the classical tripos with distinction in ancient philosophy and followed this with first-class honors in the moral sciences tripos with distinction in metaphysics. In 1884 he was elected a fellow of his college, and in 1891 he succeeded George Croom Robertson as editor of *Mind.* He was appointed Anderson lecturer in comparative psychology at Aberdeen in 1896; Wilde reader in mental philosophy at Oxford in 1899; and professor of logic and metaphysics at the University of St. Andrews in 1903. He remained at St. Andrews,

where he was instrumental in establishing a laboratory of experimental psychology, until his retirement in 1936. In 1939 he went to Sydney, Australia, to live with his son Alan, who had been appointed to the chair of moral and political philosophy at the University of Sydney. He spent the remaining years of his life joining vigorously in the discussions of a lively circle of younger philosophers at that university.

Stout's position in the history of philosophy and psychology is at the end of the long line of philosophers who, by reflective analysis, introspection, and observation, established the conceptual framework of what became in his time the science of psychology. He was a pupil of James Ward but not a mere disciple. He assimilated the essentials of Ward's system into his own philosophy of mind, but in the assimilation he transformed and extended them so that he created an entirely original and distinctive philosophy. Although he was formidable in polemical discussion, his bent was to constructive thinking. He assimilated many systems, boasting in later years, "I have got them all in my system" (idealism, realism, rationalism, and empiricism). He acknowledged indebtedness to philosophers as diverse as Benedict de Spinoza and Thomas Hobbes and to the last was preoccupied with the ideas of his contemporaries Bertrand Russell, G. E. Moore, and Ludwig Wittgenstein, and he was far from being unsympathetic to the increasingly influential schools of psychology: behaviorism and the hormic and gestalt psychologies.

In Stout's work there is a progressive development of three main theses: the doctrine concerning thought and sentience; the concept of the embodied self; and a doctrine concerning "conative activity." These central theses entail in their elaboration the reinterpretation of many of the concepts important, historically and analytically, in the philosophy of mind. It is difficult to distinguish clearly, although the attempt is rewarding, between changes (or developments) in Stout's views and changes merely in his terminology. In his earlier writings, for example, he was content to describe the ultimate data of our knowledge of the external world as "sensations." Later he followed Ward in using the term *presentations,* and finally he accepted *sense data* and *sensa* to facilitate discussion with the exponents of the prevailing phenomenalism of the day. The readiness to change his terminology was most striking in his many attempts to convey his distinctive doctrine of thought reference.

THOUGHT AND SENTIENCE

Since the time of George Berkeley there has been a widely accepted doctrine that cognition begins with simple sensations which are mental states and "in the mind"; that these sensations and their corresponding images are associated in order to form complex ideas; that some of these sensations and images are projected so as to appear as phenomena of the external world; and that these sensations are the ultimate basis of our beliefs about and our knowledge of the external world. Against this Stout set up the proposition that sense experience involves "thought reference" to real objects. As René Descartes had held that "thought" (as he used the term) implies a thinker, so Stout held that "thought" (in the same sense) implies something real and objective which is thought about.

This thesis, prominent in his *Analytic Psychology,* was expressed in terms of the concept of "noetic synthesis." In his characteristic conciliatory way he conceded the abstract possibility of "anoetic sentience" (sense experience without thought reference), but in subsequent writings he was inclined to deny both the occurrence of anoetic sentience and (to coin a phrase for him) "nonsentient noesis" (imageless thought or any form of thought reference independent of sense experience). In the elaboration of this thesis he offered a paradoxical theory of error—one difficult to refute or prove—to the effect that there can be no complete error, no sheer illusion, no pure hallucination. All errors are misinterpretations of fact. This thesis was later expressed in terms of "original meaning," in saying that every sense experience is apprehended as "conditioned by something other than itself," or as an "inseparable phase of something other than itself." It was developed with subtlety and in detail in the genetic psychology of the *Manual of Psychology.*

Following Ward, Stout attempted to give a natural history of the development of human awareness of the world which also offered grounds for our knowledge of what the world is really like. The central thesis here is that we must accept as primary not only the particular sense data of experience but also the categories or ultimate principles of unity: space, time, thinghood, and causality. These are not so much a priori cognitions as dispositions to organize experience in certain ways. We do not, for instance, have a priori knowledge that every event has a cause, but we have a disposition to look for causes. So, *mutatis mutandis,* with the other categories.

THE EMBODIED SELF

Stout, like Ward, accepted a two-dimensional, tripartite division of mental functions into cognition, feeling, and

conation; and he distinguished self, attitude, and object in each function. However, in the analysis of every concept in this scheme Stout modified every idea he took from Ward. He was more thoroughgoing in his adoption of Franz Brentano's principle that the essential component that distinguishes a mental function from a nonmental one is the attitude or way in which the subject is concerned with its objects. His most fundamental divergence from Ward was in his account of the knowing, feeling, and willing subject (self or ego). His differences from Ward are set out in detail in his important article "Ward as a Psychologist" (*Monist,* January 1926). Here he opposed to Ward's account of the pure ego his own view that the self as first known in sensible experience is that thing whose boundary from other things is the skin.

The *Manual of Psychology* contains a puzzling and confusing chapter, "Body and Mind," that combines a critique of the classical theories of interactionism, epiphenomenalism, and parallelism, all of which presupposed Cartesian dualism, with a defense of a version of parallelism that did not. This chapter puzzled students until, many years later, Stout was able to set out more clearly (especially in the Gifford Lectures) his basic philosophical thesis. This was a rejection of a dualistic ontology (that there are two sorts of substance, material things and minds) and a defense of a dualism of attributes—physical and mental—combined in a single entity, the embodied mind, which has both physical and mental attributes united somewhat as the primary and secondary characteristics are united in a material object as it is apprehended in naive perceptual situations. This view of the self entailed a corresponding reanalysis of the mental attitudes of cognition, feeling, and conation.

Stout discarded the dualism of substances but retained the dualism of qualities in his account of mental dispositions. These came to be described as "psychophysical dispositions" in accounts of the instincts, sentiments, attitudes, and other proposed ultimate sources of behavior. In this he anticipated and inspired the hormic psychology of William McDougall and, less directly, the theory of personality elaborated by Gordon Allport. McDougall was to describe the ultimate springs of human conduct in terms of certain innate primary psychophysical dispositions to perceive and attend to certain objects, to feel emotional excitement in the presence of such objects, and to experience an impulse to act in certain ways in regard to those objects. Allport later defined these sources of behavior as mental and neural "states of readiness" for such experiences and activities. In Stout these concepts are embodied in a more radical account of conative activity and conative dispositions.

CONATION

Although he accepted the classical tripartite division of mental functions, Stout accorded a certain priority to conation, so much so that he encouraged what has been described as the "conative theory of cognition," such as that developed by his contemporary Samuel Alexander. (The last paper published by Stout was "A Criticism of Alexander's Theory of Mind and Knowledge," *Australian Journal of Psychology and Philosophy,* September 1944.) The term *conative activity* covers all psychophysical processes which are directed to a goal (whether anticipated or not). It includes such cognitive processes as observation, recollection, and imagination, which are directed to the attainment of clearer and fuller perception of things present, the reconstruction of the past, and the comprehension of future possibilities. Conation is divided into practical and theoretical conation. Practical conative activity is directed to producing actual changes in the objects and situations with which the subject has to deal in the real world. Theoretical conation is directed to the fuller and clearer apprehension of such objects and situations. Stout's account of theoretical conation was in effect his account of attention. Attention is theoretical conation, although it incorporates practical conation through determining sensory-motor adjustments and the manipulation of instruments that facilitate clarity of perception.

Traditional accounts of association and reproductive and productive thinking were similarly revised and restated in conative terms. The law of association by contiguity was reformulated as the law of association by continuity of interest. One basic idea in all later theories of productive or creative thinking derives from Stout's account of "relative suggestion," an expression introduced by Thomas Brown that led to confusion between Stout's usage and Brown's.

In his treatment of all these concepts, Stout advanced beyond Ward and contributed significantly to the transition of psychology from a branch of philosophy to a science of human experience and behavior. These contributions were largely ignored, however, because of the powerful movements in psychology that were adverse to what had come to be described as "armchair psychology," that is, the purely formal analysis of psychological concepts. Stout's influence on philosophical thought outside his own circle of associates was also limited because of the reaction against "speculative" philosophy and the

increasing restriction of philosophical discussion to analysis, more especially to the analysis of linguistic usage.

Stout's philosophy was, mistakenly, treated as being in the tradition of metaphysical speculation and the creation of systems in the grand manner. His final position is most fully set out in the two volumes of Gifford Lectures. These embody many clarifications of concepts in the philosophy of mind and some acute criticism of earlier expositions of materialism and of contemporary phenomenalism. They contain the only records of Stout's views on aesthetics and ethics and his more tentative speculations concerning God, teleology, and the nature of material things. There is probably no philosopher who in his own thinking so smoothly made the transition from the prevailing idealism of the late nineteenth century to the prevailing critical, nonspeculative philosophy of the mid-twentieth century. Something of the idealist tradition is preserved in his sophisticated defense of philosophical animism, but more important are his detailed contributions to the transition from the philosophy of mind of the nineteenth century to that of the twentieth.

Bibliography

WORKS BY STOUT

Analytic Psychology. 2 vols. London: Sonnenschein, 1896.

A Manual of Psychology. London: University Correspondence College Press, 1899; 4th ed., rev. by C. A. Mace, London, 1929; 5th (and last) ed., London: University Tutorial Press, 1938. The 5th edition contains an appendix on gestalt psychology by R. H. Thouless and a supplementary note by Stout.

Studies in Philosophy and Psychology. London: Macmillan, 1930.

Mind and Matter. London and New York: Cambridge University Press, 1931. Vol. I of the Gifford Lectures.

God and Nature. Edited by Alan Stout. London: Cambridge University Press, 1952. Vol. II of the Gifford Lectures with a memoir by J. A. Passmore and a full bibliography.

WORKS ON STOUT

Broad, C. D. "The Local Historical Background of Contemporary Cambridge Philosophy." In *British Philosophy in the Mid Century,* edited by C. A. Mace. London: Allen and Unwin, 1957; 2nd ed., 1966.

Hamlyn, D. W. "Bradley, Ward and Stout." In *Historical Roots of Contemporary Psychology,* edited by B. B. Wolman. New York: Harper and Row, 1968.

Mace, C. A. "George Frederick Stout." *Proceedings of the British Academy* 31.

Mace, C. A. "The Permanent Contributions to Psychology of George Frederick Stout." *British Journal of Educational Psychology* 24, Part 2 (June 1954).

Passmore, J. A. *A Hundred Years of Philosophy,* 192–202 and passim. London: Duckworth, 1957.

C. A. Mace (1967)

STRATO AND STRATONISM

Little is known for sure about the life and work of Strato of Lampsacus, third head of Aristotle's school. He lived in Alexandria as tutor to the young Ptolemy Philadelphus for some time before he took over the leadership of the Peripatos; during this time he likely came into contact with the doctors and scientists patronized by the Ptolemaic court. He was head of Aristotle's school in Athens from Theophrastus's death in 286 BCE until his death in 268 or 269 BCE. The school seems to have dwindled into obscurity after Strato's time: Explanations offered for this include a suspect story that the school lost its library after Theophrastus's death.

Strato was known in antiquity as "the natural philosopher," possibly because of his insistence on separating the study of the natural world from any dependence on the divine. He reportedly ascribed all natural events to forces of weight and motion. He rejected Aristotle's doctrine of the fifth element, and also the idea that air and fire have an independent tendency to move upward, claiming instead that they are squeezed out by the fall of heavy bodies. His physics seems to have been basically Aristotelian, because he stressed the role of hot and cold in effecting change; yet he seems to have made changes in the doctrine of the void, because he held that it is at least possible within the cosmos. One report claims that he held that matter has passageways to allow the passage of light and heat. Controversy surrounds the relationship between Strato's view of the void and that of later Hellenistic theories of pneumatic effects. His best-known contributions to natural philosophy include attempts to prove the downward acceleration of falling bodies.

Besides work on logic, metaphysics, and ethics, Strato wrote a number of works on medical topics. Perhaps following Hellenistic medical research, he seems to have offered a naturalistic account of the soul, ascribing its functions to a substance, *pneuma*, carried in passageways throughout the body. He located the center of the soul's activity between the eyebrows, rejecting Aristotle's view that the heart is the center. He regarded reasoning as a causal movement in the soul, and offered lists of objections to Plato's arguments for the immortality of the soul.

Strato may have had some impact amongst the scientific figures in Alexandria, but his greatest notoreity was acquired some two millenium later. Ralph Cudworth characterized Strato's approach—which he called "hylozoism," the idea that matter is inherently alive—as a par-

ticularly pernicious brand of atheism. Although there is little evidence that this is Strato's view, his name became identified in the Enlightenment with a kind of naturalistic atheism.

See also Aristotelianism; Aristotle.

Bibliography

Until the expected reedition of the fragments of Strato of Lampsacus by R. W. Sharples, Rutgers University Studies in the Classical Humanities, the standard edition is F. Wehrli, *Die Schule des Aristoteles*. Vol. 5: *Straton von Lampsakos*. 2nd ed. (Basel, Germany: Schwabe, 1969). Recent studies include David Furley, "Strato's Theory of Void," in his *Cosmic Problems: Essays on Greek and Roman Philosophy of Nature*, 149–160 (Cambridge, U.K.: Cambridge University Press, 1989), and H. B. Gottschalk, "Strato of Lampsacus: Some Texts," *Proceedings of the Leeds Philosophical and Literary Society* 9 (1965): 95–182.

Sylvia Berryman (2005)

STRAUSS, DAVID FRIEDRICH
(1808–1874)

David Friedrich Strauss, the German theologian, historian of religion, and moralist, was born at Ludwigsburg in Württemberg. He studied from 1821 to 1825 at Blaubeuren, where he fell under the influence of the Hegelian theologian F. C. Baur, and at the Tübingen Stift from 1825 to 1831. He next attended the University of Berlin, where he heard lectures by G. W. F. Hegel and Friedrich Schleiermacher. In 1832 he went to the University of Tübingen as lecturer, remaining there until 1835, the year of the publication of the first volume of his most important work, *Das Leben Jesu kritisch bearbeitet* (2 vols., Tübingen, 1835–1836; translated from the 4th German edition by George Eliot as *The Life of Jesus Critically Examined,* London, 1848). The universal storm of public indignation that this book occasioned resulted in his dismissal from the university and his permanent retirement from academic life. Master of a clear and forthright prose style, Strauss had no difficulty supporting himself as a journalist and popular exponent of the view that religion—Christianity in particular—is an expression of the human mind's capacity to generate myths and treat them as truths revealed by God to man.

When he began his study of the Gospels, Strauss was neither a liberal nor a materialist. His original interests had been those of a Hegelian idealist; he had meant to study the available records of Jesus' life in order to distinguish their historically valid content from the theological accretions that had become associated with them during the first two centuries of the Christian era. His investigations convinced him, however, that the principal importance of the Gospels was aesthetic and philosophical, not historical. On the one hand, the Gospels provided insight into the Messianic expectation of the Jewish people in the late Hellenistic period; on the other hand, they reflected a memory of the exceptional personality of a great man, Jesus. Thus envisaged, the Gospels were a synthesis of notions peculiar to the Jews regarding the nature of world history and of certain moral teachings associated with the name of a purely human, yet historically vague, personality, presented in an aesthetically pleasing form for members of a new religious community that was both Jewish and Greek in its composition. For Strauss, the Gospels were, in short, interesting primarily as evidence of the workings of consciousness in the sphere of religious experience: they showed how the mind could fabricate miracles and affirm them as true, contrary to the Hegelian dictum, then regarded as an established truth, that the real was rational and the rational was real.

Had Strauss halted at this point, his work might have been ignored as merely another vestige of the free thought of the Enlightenment. Instead, he went on to argue that even if the historicity of the account of Jesus's life in the Gospels were denied, it need not follow that the Gospels were a product of conscious invention or fraud. He held, rather, that they could be said to belong to a third order of mental activity, called by Hegel unconscious invention or myth and defined by him as an attempt to envision the Absolute in terms of images derived from sensible experiences. As unconscious invention, the Gospels were to be viewed as poetic renderings of man's desire to transcend the finitude of the historical moment, as evidence of the purely human desire to realize the immanent goal of Spirit in its journey toward the Hegelian Being-in-and-for-itself. Thus, although Strauss had denied that the Gospels were evidence of the direct intrusion of the divine into history or even of the true nature of Jesus' life, he had, in his own view, at least salvaged them as documents in the history of human expression. In doing so, of course, he had reduced them to the same status as the pagan myths, legends, and epics.

In a second work, *Die christliche Glaubenslehre* (2 vols., 1840–1841), Strauss tried to clarify the theoretical basis of his original historical inquiry. He argued that Christianity was a stage in the evolution of a true pantheism that had reached its culmination in Hegelian philos-

ophy. What the poet and mystic took for God was nothing but the world—specifically, man in the world—conceived in aesthetic terms. Science studied the same phenomena that are governed by physical laws, and philosophy was, as Hegel had taught, mind reflecting on these prior activities of thought and imagination.

Das Leben Jesu became a cause célèbre in a Germany growing increasingly reactionary both politically and intellectually. The attack launched against Strauss from all quarters soon made him a symbol to German liberals; he was regarded as a martyr of science and freedom of thought. Accordingly, Strauss was drawn into political as well as theological polemics. In 1848 he published at Halle a defense of bourgeois liberalism, *Der politische und der theologische Liberalismus.* He later turned to the study of philosophical materialism (that of Friedrich Albert Lange and of Charles Darwin) and to the production of a series of historical works on leading advocates of freedom of thought in European history (for example, a long biography of Ulrich von Hutten, 1858, and a study of Voltaire, 1870). As he progressed, he repudiated the Hegelianism of his first book. In a preface to a later edition of *Das Leben Jesu,* he stated that he had undertaken it to show "to those to whom the conceptions … as to the supernatural character … of the life of Jesus had become intolerable … [that] the best means of effectual release will be found in historical inquiry." Abandoning the last residues of his earlier idealism, he argued that "everything that happens, or ever happened, happened naturally." He still recognized the aesthetic value of the Gospel account, but he now saw it as providing the image of the good life that had finally become possible on this earth because of the triumphs of science and industrial technology and the advance of political liberalism. It was this position that won for him the enmity of both Karl Marx and Friedrich Nietzsche. To Marx, he was the bourgeois *idéologue* par excellence, who tried to combine Christian sentimental ethics and the practices of capitalism in a single package. For Nietzsche, Strauss represented the German *Bildungsphilister* who made a show of intellectual radicalism but always left the conventional morality intact.

Strauss remained to the end of his life the spokesman of popular religious criticism, materialistic in his intention but Hegelian in method, a combination which allowed him to accommodate almost any position that appealed to him. After 1850 his political and social criticism became increasingly conservative—aristocratic, monarchical, and nationalistic. In part this transformation was due to the suspicion that popular democracy would be in general as unable to recognize genius as it had been unable to recognize, in particular, the value of Strauss's own works; but this transformation was also a result of his attempt to move from Hegelianism to positivism. In the second half of the eighteenth century, positivist social thought had become—as, for example, in Hippolyte Taine—a kind of crude determinism, hostile to any revolutionary impulse.

To the young Hegelians, who were already becoming aware of the methodological limitations of Hegel's late thought, *Das Leben Jesu* provided an impulse to the critical, empirical study of the historical milieus within which *Geist* supposedly manifested itself, and it thus prepared them to accept Leopold von Ranke's historicism. To German liberals, Strauss remained a symbol of the risks that had to be run by any German who presumed to espouse radical causes. The later Marxists regarded Strauss as merely a confused bourgeois who had blundered onto forbidden ground. For them, the way to a true revision of Hegelianism was provided by Ludwig Feuerbach. Feuerbach saw that the true importance of Strauss's *Das Leben Jesu* lay in a problem that remained implicit in the work and was hardly touched upon by Strauss himself: the psychological problem about the nature of the mythmaking mechanism that distinguishes man from the rest of nature. It was Feuerbach, then, rather than Strauss, who posed the question with which German philosophy had to come to terms in the 1840s—the question of the relation between human consciousness and its material matrix.

See also Darwin, Charles Robert; Enlightenment; Feuerbach, Ludwig Andreas; Hegel, Georg Wilhelm Friedrich; Hegelianism; Idealism; Lange, Friedrich Albert; Marx, Karl; Materialism; Miracles; Nietzsche, Friedrich; Positivism; Schleiermacher, Friedrich Daniel Ernst; Taine, Hippolyte-Adolphe; Voltaire, François-Marie Arouet de.

Bibliography

WORKS BY STRAUSS

Ulrich von Hutten. 2 vols. Leipzig, 1858. Translated by Mrs. G. Sturge as *Ulrich von Hutten, His Life and Times.* London, 1874.

Hermann Samuel Reimarus. Leipzig: Brockhaus, 1862.

Kleine Schriften biographischen, literarischen, und kunstgeschichtlichen Inhalts. Leipzig, 1862.

Der Christus des Glaubens und der Jesus der Geschichte. Berlin, 1865.

Kleine Schriften, Neue Folge. Berlin, 1866.

Voltaire: sechs Vortrage. Leipzig, 1870.

Gesammelte Schriften, edited by Eduard Zeller. 12 vols. Bonn: E. Strauss, 1876–1878.

Ausgewählte Briefe. Bonn, 1895.

WORKS ON STRAUSS

Barth, Karl. *Die protestantische Theologie im 19. Jahrhundert.* Zürich: Evangelischer, 1947. Translated by Brian Cozens as *Protestant Thought from Rousseau to Ritschl.* New York: Harper, 1959.

Feuter, Eduard. *Geschichte der neueren Historiographie.* 3rd ed. Munich and Berlin, 1936.

Nietzsche, Friedrich. *Unzeitgemüsse Betrachtungen, Erstes Stück.* Leipzig, 1873. Translated by A. M. Ludovici as *Thoughts out of Season.* New York, 1924.

Schweitzer, Albert. *Von Reimarus zu Wrede.* Tübingen: Mohr, 1906. Translated by W. Montgomery as *The Quest for the Historical Jesus.* London: Black, 1910.

Zeller, Eduard. *David Friedrich Strauss in seinem Leben und seinen Schriften.* 2 vols. Heidelberg, 1876–1878.

Hayden V. White (1967)

STRAWSON, PETER FREDERICK

(1919–)

Peter Frederick Strawson, the British philosopher, was educated at Christ's College, Finchley, and St. John's College, Oxford. He holds the BA and MA degrees and is a fellow of University College, Oxford.

LANGUAGE AND LOGIC

Strawson is a leading member of the circle of philosophers whose work is sometimes described as "ordinary language philosophy" or as "Oxford philosophy." Of his early work, the most influential and most controversial is the famous article "On Referring" (*Mind*, 1950), a criticism of the philosophical aspects of Bertrand Russell's theory of definite descriptions. According to Russell's theory any sentence of the form "The *f* is *g*"—for example, "The king of France is bald"—is properly analyzed as follows (in terms of our example): "There is a king of France. There is not more than one king of France. There is nothing which is king of France and which is not bald."

Strawson argues that this analysis confuses referring to an entity with asserting the existence of that entity. In referring to an entity, a speaker presupposes that the entity exists, but he does not assert that it exists, nor does what he asserts entail that it exists. Presupposition is to be distinguished from entailment. In asserting something of the form "The *f* is *g*," a speaker refers or purports to refer to an entity with the subject noun phrase, and to do so involves presupposing that there is such an entity, but this is quite different from asserting that there is such an entity.

According to Strawson this confusion between referring and asserting is based on an antecedent confusion between a sentence and the statement made in a particular use of that sentence. Russell erroneously supposes that every sentence must be either true, false, or meaningless. But, Strawson argues, sentences can be meaningful or meaningless and yet cannot strictly be characterized as true or false. Statements, which are made using sentences, but which are distinct from sentences, are, or can be, either true or false. The sentence "The king of France is bald" is indeed meaningful, but a statement made at the present time using that sentence does not succeed in being either true or false because, as there is presently no king of France, the purported reference to a king of France fails. According to Russell the sentence is meaningful and false. According to Strawson the sentence is meaningful, but the corresponding statement is neither true nor false because one of its presuppositions—namely, that there is a king of France—is false.

In another well-known article of this early period, "Truth" (*Analysis*, 1949), Strawson criticizes the semantic theory of truth and proposes an alternative analysis to the effect that "true" does not describe any semantic properties or, indeed, any other properties at all, because its use is not to describe; rather, we use the word *true* to express agreement, to endorse, concede, grant, or otherwise accede to what has been or might be said. Strawson explicitly draws an analogy between the use of the word *true* and J. L. Austin's notion of performatives. Like performatives, *true* does not describe anything; rather, if we examine its use in ordinary language, we see that it is used to perform altogether different sorts of acts.

This article gave rise to a controversy with Austin, a defender of the correspondence theory. The gist of Strawson's argument against the correspondence theory is that the attempt to explicate truth in terms of correspondence between statements on the one hand and facts, states of affairs, and situations on the other must necessarily fail because such notions as "fact" already have the "word–world relationship" built into them. Facts are not something which statements name or refer to; rather, "facts are what statements (when true) state."

In his first book, *Introduction to Logical Theory* (New York and London, 1952), Strawson continued his investigation of the logical features of ordinary language by studying the relations between ordinary language and formal logic. The book, he says, has two complementary aims: first, to compare and contrast the behavior of ordinary words with the behavior of logical symbols, and, second, to make clear the nature of formal logic itself. It is in

the first of these two enterprises that he has shown the more originality and aroused the more controversy. The theme of this part of the book is that such logical systems as the propositional and predicate calculi do not completely represent the complex logical features of ordinary language and indeed represent them less accurately than has generally been supposed. He argues that the logical connectives, especially "∨," "⊃," and "≡," are much less like "or," "if," and "if and only if" than is often claimed. In his discussion of predicate logic (Chs. 5 and 6), he continues the themes of "On Referring," arguing that certain orthodox criticisms which are made of traditional Aristotelian syllogistic fail because of a failure to appreciate the fact that statements made in the use of a sentence of the form "All *f*'s are *g*" presuppose the existence of members of the subject class.

Thus, for example, the question whether it is true that all John's children are asleep does not even arise if John has no children. Once it is seen that statements of the form "All *f*'s are *g*" have existential presuppositions, it is possible to give a consistent interpretation of the traditional Aristotelian system. The failure to understand this and the misconception regarding the relation of the predicate calculus to ordinary language are in large part due to the same mistakes that underlie the theory of descriptions: the failure to see the distinction between sentence and statement; the "bogus trichotomy" of true, false, or meaningless; and the failure to see the distinction between presupposition and entailment.

The final chapter of the book contains a discussion of probability and induction in which Strawson argues that attempts to justify induction are necessarily misconceived, since there are no higher standards to which one can appeal in assessing inductive standards. The question whether inductive standards are justified is as senseless as the question whether a legal system is legal. Just as a legal system provides the standards of legality, so do inductive criteria provide standards of justification. Underlying this point is the fact that inductive standards form part of our concept of rationality. It is, he says, a necessary truth that the only ways of forming rational opinions concerning what happened or will happen in given circumstances are inductive.

METAPHYSICS

In the middle 1950s Strawson's concerns shifted from investigations of ordinary language to an enterprise he named descriptive metaphysics. This enterprise differs from "revisionary metaphysics" in that it is content to describe the actual structure of our thought about the world rather than attempting to produce a better structure, and it differs from ordinary conceptual analysis in its much greater scope and generality, since it attempts to "lay bare the most general features of our conceptual structure."

These investigations resulted in the publication of a second book, *Individuals* (London, 1959). The book is divided into two parts. Part One, titled "Particulars," deals with the nature of and preconditions for the identification of particular objects in speech; Part Two, "Logical Subjects," concentrates on the relations between particulars and universals and on the corresponding and related distinctions between reference and predication and subjects and predicates. The first important thesis of the book is that from the point of view of particular identification, material objects are the basic particulars. What this means is that the general conditions of particular identification require a unified system of publicly observable and enduring spatiotemporal entities. The material universe forms such a system. Material objects can therefore be identified independently of the identification of particulars in other categories, but particulars in other categories cannot be identified without reference to material objects. This provides us, then, with a sense in which material objects are the basic particulars as far as particular identification is concerned.

A second thesis, one of the most provocative of the book, concerns the traditional mindogon;body problem. In Chapter 3, titled "Persons," Strawson attacks both the Cartesian notion that states of consciousness are ascribed to mental substances, which are quite distinct from but nonetheless intimately connected to bodies, and the modern "no-ownership" theory, according to which states of consciousness are not, strictly speaking, ascribed to anything at all. Both views, he argues, are ultimately incoherent. The solution to the dilemma posed by these views is that the concept of a person is a primitive concept. It is a concept such that both states of consciousness and physical properties are ascribable to one and the same thing—namely, a person. The concept of a mind is derivative from the primitive concept of a person, and the concept of a person is not to be construed as a composite concept made up of the concept of a mind and the concept of a body. The recognition of the primitiveness of the concept of a person enables us to see both why states of consciousness are ascribed to anything at all and why they are ascribed to the very same thing to which certain physical states are ascribed.

Most of Part Two of *Individuals* is devoted to an investigation of the problems of the relations of subjects

and predicates. Strawson considers two traditional ways of making the distinction between subject and predicate: a grammatical criterion in terms of the different kinds of symbolism for subject and predicate expressions and a category criterion in terms of the distinction between particulars and universals. He investigates the "tensions and affinities" between these two criteria, and he concludes that the crucial distinction between the way a subject expression introduces a particular into a proposition and the way a predicate expression introduces a universal into a proposition is that the identification of a particular involves the presentation of some empirical fact which is sufficient to identify the particular (this harks back to the doctrine of what is presupposed by identifying references in "On Referring" and *Introduction to Logical Theory*), but the introduction of the universal term by the predicate term does not in general involve any empirical fact. The meaning of the predicate term suffices to identify the universal that the predicate introduces into the proposition. One might say that identifying reference to particulars involves the presentation of empirical facts; the predication of universals involves only the presentation of meanings. This enables us to give a deeper sense to Gottlob Frege's notion that objects are complete—in contrast to concepts, which are incomplete—and it enables us to account for the Aristotelian doctrine that only universals and not particulars are predicable.

In tone, method, and overall objectives, *Individuals* stands in sharp contrast to Strawson's earlier work. Piecemeal investigation of ordinary language occurs here only as an aid and adjunct to attacking large traditional metaphysical problems. One might say that *Individuals* employs essentially Kantian methods to arrive at Aristotelian conclusions. Yet much of the book is at least foreshadowed by Strawson's earlier work, particularly "On Referring" and certain portions of his first book. The notion of descriptive metaphysics itself has been as influential as the actual theses advanced in *Individuals*. More than any other single recent work, this book has resurrected metaphysics (albeit descriptive metaphysics) as a respectable philosophical enterprise.

See also Performative Theory of Truth.

Bibliography

ADDITIONAL WORKS BY STRAWSON

"Necessary Propositions and Entailment Statements." *Mind* (1948).

"Ethical Intuitionism." In *Readings in Ethical Theory*, edited by Wilfrid Sellars and John Hospers. New York: Appleton-Century-Crofts, 1952.

"Particular and General." *PAS* supp. 27 (1953–1954).

"Critical Notice of Wittgenstein." *Mind* (1954).

"Presupposing: A Reply to Mr. Sellars." *Philosophical Review* (1954).

"A Logician's Landscape." *Philosophy* (1955).

"Construction and Analysis." In *The Revolution in Philosophy*, by A. J. Ayer et al. London: Macmillan, 1956.

"In Defense of a Dogma." *Philosophical Review* (1956). Written with H. P. Grice.

"Singular Terms, Ontology and Identity." *Mind* (1956).

"Logical Subjects and Physical Objects." *Philosophy and Phenomenological Research* (1956–1957).

"*Logic and Knowledge* by Bertrand Russell." *Philosophical Quarterly* (1957).

"Metaphysics." In *The Nature of Metaphysics*, edited by D. F. Pears. London: Macmillan, 1957. Written with H. P. Grice and D. F. Pears.

"Professor Ayer's *The Problem of Knowledge*." *Philosophy* (1957).

"Proper Names." *PAS* supp. 31 (1957).

"Propositions, Concepts and Logical Truths." *Philosophical Quarterly* (1957).

"Singular Terms and Predication." *Journal of Philosophy* (1961).

"Social Morality and Individual Ideal." *Philosophy* (1961).

"Freedom and Resentment." *Proceedings of the British Academy* (1962). Annual philosophical lecture.

"Determinism." In *Freedom and the Will*, edited by D. F. Pears. New York: St. Martin's Press, 1963. A discussion with G. J. Warnock and J. F. Thomson.

"Identifying Reference and Truth Values." *Theoria* (1964).

"Truth: A Reconsideration of Austin's Views." *Philosophical Quarterly* (1965).

Meaning and Truth. Oxford: Clarendon, 1970.

Logico-Linguistic Papers. London: Methuen, 1971.

Freedom and Resentment. London: Methuen, 1974.

Subject and Predicate in Logic and Grammar. London: Methuen, 1974.

"Does Knowledge Have Foundations?" *Teorema* MONO1 (1974): 99–110.

"Universals." *Midwest Studies in Philosophy* 4 (1979): 3–10.

Skepticism and Naturalism: Some Varieties. New York: Columbia University Press, 1985.

"Ma Philosophie: Son Developpment, Son Theme Centrale et Nature Generale," *Revue de Theologie et de Philosophie* 120 (1988): 437–452.

"Two Conceptions of Philosophy." In *Perspectives on Quine*, edited by Robert Barrett and Roger Gibson. Cambridge, MA: Blackwell, 1990.

Analysis and Metaphysics: An Introduction to Philosophy. Oxford: Oxford University Press, 1992.

"Comments on Some Aspects of Peter Unger's 'Identity, Consciousness and Value,'" *Philosophy and Phenomenological Research* 52 (1) (March 1992): 145–148.

"The Incoherence of Empiricism." *Proceedings of the Aristotelian Society*, Supp. Vol. 66 (1992).

"Knowing from Words." In *Knowing from Words*, edited by B.K. Matilal and A. Chakrabarti. Dordrecht: Kluwer, 1992.

"Individuals." In *Philosophical Problems Today*, Vol. I, edited by G. Fløistad. Dordrecht: Kluwer, 1994.

"The Problem of Realism and the A Priori." In *Kant and Contemporary Epistemology*, edited by Paolo Parrini. Dordrecht: Kluwer, 1994.

"My Philosophy" and "Replies." In *The Philosophy of P. F. Stawson*, edited by Pranab Kumar Sen and Roop Rekha Verma, New Delhi: Indian Council of Philosophical Research, 1995.

Entity and Identity and Other Essays. Oxford: Clarendon, 1997.

"Intellectual Autobiography" and "Replies." *In The Philosophy of P. F. Strawson*, edited by Lewis Edwin Hahn, Library of Living Philosophers Vol. 26, Chicago and LaSalle, IL: Open Court, 1998.

"Entailment and its Paradoxes." In *Realism*: *Responses and Reactions (Essays in Honor of Pranab Kumar Sen)*. New Delhi: Indian Council of Philosophical Research, 2000.

"What Have We Learned from Philosophy in the Twentieth Century?" In *The Proceedings of the Twentieth World Congress of Philosophy, Vol. 8: Contemporary Philosophy*, edited by Daniel Dahlstrom. Charlottesville: Philosophy Documentation Center, 2000.

"A Bit of Intellectual Autobiography." In *Strawson and Kant*, edited by Hans-Johann Glock. Oxford: Clarendon, 2003.

Logico-Linguistic Papers. Rev. ed. Aldershot, Hampshire and Burlington, VT: Ashgate, 2004.

Subject and Predicate in Logic and Grammar. Rev. ed. Aldershot, Hampshire, and Burlington, VT: Ashgate, 2004.

Strawson, Peter Frederick, and Arindam Chakrabarti, eds. *Universals, Concepts and Qualities: New Essays on the Meaning of Predicates*. Aldershot, Hampshire, and Burlington, VT: Ashgate, 2005.

WORKS ON STRAWSON

Aune, Bruce. "Feelings, Moods and Introspection." *Mind* (1963).

Ayer, A. J. "The Concept of a Person." In *The Concept of a Person and Other Essays*. London: Macmillan, 1963.

Bradley, M. C. "Mr. Strawson and Skepticism." *Analysis* (1959).

Campbell, Scott. "Strawson, Parfit and Impersonality." *Canadian Journal of Philosophy* 30 (2) (June 2000): 207–224.

Caton, C. E. "Strawson on Referring." *Mind* (1959).

Coval, S. C. "Persons and Criteria in Strawson." *Philosophy and Phenomenological Research* (1964).

Geach, P. T. "Mr. Strawson on Symbolic and Traditional Logic." *Mind* (1963).

Hacker, Peter. "Strawson's Concept of a Person." *Proceedings of the Aristotelian Society* 102 (2002): 21–40.

Hahn, Lewis Edwin, ed. *The Philosophy of P. F. Strawson*. Library of Living Philosphers, Vol. 26. Chicago and La Salle, IL: Open Court, 1998.

Hutcheson, Peter. "Vindicating Strawson." *Philosophical Topics* 13 (spring 1985): 175–184.

Jarvis, Judith. "Notes on Strawson's *Logic*." *Mind* (1961).

Lewis, H. D. "Mind and Body." *PAS* (1962–1963).

Millican, Peter. "Statements and Modality: Strawson, Quine and Wolfram." *International Journal of Moral and Social Studies* 8 (3) (autumn 1993): 315–326.

Pears, D. F. Review of *Individuals. Philosophical Quarterly* (1961).

Quine, W. V. "Mr. Strawson on Logical Theory." *Mind* (1953).

Russell, Bertrand. "Mr. Strawson on Referring." *Mind* (1957).

Sellars, Wilfrid. "Presupposing." *Philosophical Review* (1954).

Sen, Pranab Kumar, and Roop Rekha Verma, eds. *The Philosophy of P. F. Strawson*, New Delhi: Indian Council of Philosophical Research, 1995.

Szubka, Tadeusz. "P. F. Strawson." In *British Philosophers, 1900–2000*, edited by Peter S. Fosi. Columbia, SC: Bruccoli Clark Layman/Gale, 2002.

Urmson, J. O. Review of *Individuals. Mind* (1961).

Van Straaten, Zak, ed. *Philosophical Subjects: Essays Presented to P. F. Strawson*. Oxford: Clarendon, 1980.

Williams, B. A. O. Review of *Individuals. Philosophy* (1961).

John R. Searle (1967)
Bibliography updated by Vicki Lynn Harper (2005)

STRING THEORY

Physicists believe there to be four fundamental forces. Three of these—the electromagnetic, the strong force, and the weak force—are amalgamated in the standard model of elementary particle physics, a family of quantum field theories that has enjoyed stupendous empirical success. Gravity, the fourth and feeblest fundamental force, is the subject of a stupendously successful *non-quantum* field theory, Einstein's general theory of relativity (GTR). Desiring to fit all of fundamental theoretical physics into a quantum mechanical framework, and suspecting that GTR would break down at tiny ("Planck scale," i.e., 10^{-33} cm) distances where quantum effects become significant, physicists have been searching for a quantum theory of gravity since the 1930s. In the last quarter of the twentieth century, string theory became the predominant approach to quantizing gravity, as well as to forging a unified picture of the four fundamental forces. A minority approach to quantizing gravity is the program of loop quantum gravity, which promises no grand unification. Both attempts to quantize gravity portend a science of nature radically different from the Newtonian one that frames much of classical philosophical discourse. They also present gratifying instances of working physicists actively concerned with recognizably philosophical questions about space, time, and theoretical virtue.

THE STANDARD MODEL

String theory would quantize gravity by treating the gravitational force as other forces are treated. In the standard model, pointlike elementary particles, quarks, and leptons constitute matter. Each particle is characterized by invariants, such as mass, spin, charge, and the like. The matter-constituting particles have half-integer multiples of spin, which makes them fermions. Beside fermions, the standard model posits gauge bosons, "messenger parti-

cles" or carriers of the interaction, for each force in its ambit. Bosons are distinguished from fermions by having whole-integer multiples of spin. As early as 1934, preliminary work on the sort of coupling with matter required by a quantum theory suggested that, if the gravitational force had a gauge boson, it must be a mass 0 spin 2 particle, dubbed the *graviton*. No such particle is predicted by the standard model.

According to string theory, the elementary particles of the standard model are not the ultimate constituents of nature. Filamentary objects—strings—are. Different vibrational modes of these strings correspond to the different masses (charges, spins) of elementary particles. The standard model is recovered, and fundamental physics unified, in a string theory incorporating vibrational modes corresponding to every species of particle in the standard particle zoo (and so incorporating the strong, weak, and electromagnetic forces), as well as to the graviton (and so incorporating gravity).

THE EARLY YEARS OF STRING THEORY

String theory evolved from attempts, undertaken within the standard model in the 1970s, to model the strong nuclear force in terms of a band between particles. As a theory of the strong nuclear force, these attempts suffered in comparison to quantum chromodynamics. They also predicted the existence of a particle that had never been detected: a mass 0 spin 2 particle. In 1974, John Schwartz and Joël Scherk proposed to promote this empirical embarrassment to a theoretical resource: The undetected particle, they suggested, was in fact the graviton! (Further evidence that string theory encompasses gravity comes in the form of a consistency constraint on the background spacetime in which string theoretic calculations are carried out, which consistency constraint resembles the equations of GTR.)

String theory evolved piecemeal in the 1970s and 1980s, roughly by adapting perturbative approximation techniques developed for the standard model's point particles to stringy entities. One benefit of the adaptation was the suppression of infinities that arise in perturbative calculations for point particles. In the standard model, these infinities call for the expedient of renormalization, the barelyprincipled subtraction of other infinities to yield finite outcomes. Perturbative string theories require no such expedient. Worries that they harbored inconsistencies all their own, called *anomalies*, were allayed by Schwartz's and Michael Green's 1984 argument that string theories were anomaly-free—a result that galvanized research in the field.

By the early 1990s there were five different consistent realizations of perturbative string theory. These realizations shared some noteworthy features. First, their equations were consistent only in ten space-time dimensions. To accord with the appearance that space is three-dimensional, the extra six dimensions are supposed to be Planck-scale and compactified ("rolled up"). (The usual analogy invokes the surface of a cylinder, which is a two-dimensional object: one dimension runs along the length of the cylinder; the other is "rolled up" around its circumference. Supposing the rolled-up dimension to be small enough, a cylinder looks like a one-dimensional object, a line.) Details of the geometries of these extra dimensions influence the physics string theory predicts. These details are adjustable; only with certain choices of the geometries can string theory mimic the standard model.

The initial string theories dealt only with bosons. So that they might incorporate fermions as well, *supersymmetry* was imposed. That is, the equations of string theory were required to be invariant under half-integer changes in spin. Thus the theory predicts for every particle in the standard zoo that it has a supersymmetric partner. For the (spin 1/2) electron, a spin 0 "selectron;" for the (spin-1) photon, a spin 1/2 "photino," and so on. Of these supersymmetric partners, none are observable using present technologies. But there is hope of detecting the lightest, the neutralino, with the Large Hadron Collider, slated to come on-line at CERN in 2007.

Parameters describing, for example, coupling strengths or the volume of the compactified extra dimensions appear in string theories. This means that each string theory can be thought of as a member of a family of related string theories, obtained from the first by varying the values of these parameters. A *duality* is said to obtain between theories so related. In the mid-1990s, Ed Witten and others uncovered evidence of dualities connecting pairs in the set of five consistent perturbative string theories. This embolded Witten to propose that the existing, approximate, string theories were all approximations to a single underlying exact theory he dubbed "M-theory." Although the equations of M-theory are unknown, it is believed that they hold in an eleven-dimensional spacetime, and have eleven-dimensional supergravity (ironically enough, a leading contender for the title "theory of everything" which string theory dislodged in the early 1980s) as their low-energy limit. In addition to strings, M-theory boasts higher-dimensional supersymmetric objects—membranes—some theorists have put to cosmological use, for example, by maintain-

ing that the three spatial dimensions of this world are a three-brane moving through an eleven-dimensional universe harboring other worlds such as this one.

Most predictions of fledgling programs in quantum gravity are experimentally inaccessible, and liable to stay that way. But a nonempirical circumstance is widely believed to confirm string theory. In black hole thermodynamics (developed by Stephen Hawking, Jacob Bekenstein, and others), black holes are attributed properties, such as temperature and entropy, that obey thermodynamic laws. (For instance, entropy, identified as the surface area of a black hole's future event horizon, never decreases.) For certain black holes known as extremal black holes, string theoretic calculations exactly reproduce the Bekenstein entropy formula. Although there has never been an observation confirming (or disconfirming) black hole thermodynamics, the recovery of the black hole entropy formula is widely held to be evidence that string theory is on the right track.

More empirical tests have been proposed, none strong. For example, if the extra dimensions posited by string theory are large enough, new mechanisms for the production of microscopic black holes could be unleashed at energies attainable in the Large Hadron Collider. But string theory is not required to posit large extra-dimensions. So the failure of microscopic black holes to appear would not force the abandonment of string theory.

Despite its successes, there are causes for complaint about string theory. It is not an exact theory yet. Its predictions might seem unduly sensitive to the discretionary matter of the geometry of the extra dimensions. In addition to predicting the existence of the standard particles and the graviton, it predicts the existence of infinitely many particles, including supersymmetric particles, humans have not seen (yet). It requires seven extra spatial dimensions humans have not seen (yet). And as formulated *at present*, it takes place in a fixed space-time background.

STRING THEORY AND LOOP QUANTUM GRAVITY

The game of background-independent M-theory is afoot; some (e.g., Smolin 2001) hope that its pursuit will reveal connections between string theory and its main rival, loop quantum gravity. Background-independence is the rallying cry of the (much less populated) loop quantum gravity camp. Largely trained as general relativists, adherents of this approach take the fundamental moral of GTR to be that space-time is not a setting in which physics happens but is itself a dynamical object, malleable in response to the matter and energy filling it. Whereas string theory seeks a quantum theory of gravity on the model of early twenty-first century quantum theory of other forces—a model that adds a graviton to a particle zoo revealed by approximations carried out in a fixed spacetime background—loop quantum gravity seeks a quantum theory of gravity by quantizing gravity: that is, by casting GTR as a classical theory in Hamiltonian form, and following a canonical procedure for quantizing such theories. Insofar as GTR's variables determine the geometry of space-time, should the quantization procedure succeed, space-time itself would be the commodity quantized.

The quantization procedure is complicated by the fact that GTR is a constrained Hamiltonian system: its canonical momenta are not independent. Instead, they satisfy constraint equations that must be reflected in the final quantum theory. The origin of these constraint equations is the diffeomorphism invariance of GTR, that is, if one starts with a solution to the equations of GTR and smoothly reassigns the dynamical fields comprising that solution to the manifold on which they are defined, one winds up with a solution to the equations of GTR. Adherents of loop quantum gravity take diffeomorphism invariance to express the background independence of GTR.

Loop quantum gravity exploits a Hamiltonian formulation of GTR due to Abhay Ashtekar, a physicist at Syracuse University. Its quantization is set in a Hilbert space spanned by spin-network states: graphs whose edges are labeled by integer multiples of 1/2. Not set in some background space, these spin-network states are supposed to be the constituents from which space is built. Defined on their Hilbert space are area and volume (but not length) operators that have discrete spectra. A free parameter in the theory can be adjusted so that this quantization occurs at the Planck scale. On these grounds, its adherents claim loop quantum gravity to be a background-independent exact theory that quantizes space. Like string theory, loop quantum gravity finds quasi-confirmation in its accord with black hole thermodynamics: for all black holes, loop quantum gravity reproduces the Bekenstein entropy within a factor of 4.

Despite its successes, there are causes for complaint about loop quantum gravity. It does not incorporate the predictions of the standard model. So whereas it may be a quantum theory of gravity, it is not a theory of everything. More telling, loop quantum gravity as yet fails to

reflect the full diffeomorphism invariance of GTR in a way that is both consistent and has GTR as its classical limit. The sticking point is the classical Hamiltonian constraint, related to diffeomorphisms that can be interpreted as time translations. Until this constraint is wrangled, loop quantum gravity lacks a dynamics: it consists of a space of possible instantaneous spacetime geometries, without an account of their time development. Given loop quantum gravity's ideology of background-independence, this is disappointing.

There is no established philosophy of quantum gravity. But there is much to provoke the philosopher. What, according to string theory or loop quantum gravity, is the nature of space(-time)? How many dimensions has it? (These questions are complicated by dualities between string theories revealed by varying the volumes of their compactified geometries, as well as by the holographic hypothesis, according to which physics in the interior of a region—an n-dimensional space—is dual to physics on that region's boundary—an (n-1)-dimensional space.) The search for quantum gravity was set off by no glaring empirical shortcoming in existing theories, and has reached theories for which no empirical evidence is readily forthcoming. In the absence of empirical adequacy, other theoretical virtues occupy center stage: the ideal of unification, the capacity to reproduce the results, or preserve the insights, of other theories (even unconfirmed ones); the susceptibility of puzzles posed in one theoretical framework to solution techniques available in another. The nature of these virtues, and how best to pursue them, are often live questions for quantum gravity researchers. Their work holds interest for the methodologist and the metaphysician alike.

See also Atomism; Philosophy of Physics; Quantum Mechanics; Relativity Theory.

Bibliography

Callendar, Craig, and Nicholas Huggett, eds. *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*. Cambridge, U.K.: Cambridge University Press: 2001.

Greene, Brian. *The Elegant Universe*. New York: Vintage, 2000.

Polchinski, Joseph. *String Theory*. Cambridge, U.K.: Cambridge University Press, 1999.

Rovelli, Carlo. "Quantum Spacetime: What Do We Know?" Cornell University on-line arXiv, 1999. Available from http://lanl.arxiv.org/PS_cache/gr-qc/pdf/9903/9903045.pdf.

Rovelli, Carlo. "Strings, Loops and Others: A Critical Survey of the Present Approaches to Quantum Gravity." Cornell University on-line arXiv, 1998. Available from http://lanl.arxiv.org/PS_cache/gr-qc/pdf/9803/9803024.pdf.

Smolin, Lee. *Three Roads to Quantum Gravity*. New York: Basic Books, 2001.

Stachel, John. "Early History of Quantum Gravity." In *Black Holes, Gravitational Radiation and the Universe*, edited by B. R. Iyer and B. Bhawal. Dordrecht, Netherlands: Kluwer, 1999.

Wald, Robert M. *Quantum Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago: University of Chicago Press, 1994.

Laura Ruetsche (2005)

STRUCTURALISM, MATHEMATICAL

Structuralism is a view about the subject matter of mathematics according to which what matters are structural relationships in abstraction from the intrinsic nature of the related objects. Mathematics is seen as the free exploration of structural possibilities, primarily through creative concept formation, postulation, and deduction. The items making up any particular system exemplifying the structure in question are of no importance; all that matters is that they satisfy certain general conditions—typically spelled out in axioms defining the structure or structures of interest—characteristic of the branch of mathematics in question. Thus, in the basic case of arithmetic, the famous "axioms" of Richard Dedekind (taken over by Giuseppe Peano, as he acknowledged) were conditions in a definition of a "simply infinite system," with an initial item, each item having a unique next one, no two with the same next one, and all items finitely many steps from the initial one. (The latter condition is guaranteed by the axiom of *mathematical induction*.) All such systems are structurally identical, and, in a sense to be made more precise, *the shared structure* is what mathematics investigates. (In other cases, multiple structures are allowed, as in abstract algebra with its many groups, rings, fields, and so forth.) This structuralist view of arithmetic thus contrasts with the absolutist view, associated with Gottlob Frege and Bertrand Russell, that natural numbers must in fact be certain definite objects, namely classes of equinumerous concepts or classes.

Historically, structuralism can be traced to nineteenth-century developments, including the rise of the axiomatic method and of non-Euclidean geometries leading to the recognition of multiple abstract spaces independent of *physical* space and of spatial intuition. David Hilbert, whose work in the foundations of geometry (1959 [1899]) was especially influential in this regard,

remarked that "points, lines, and planes" could be read as "tables, chairs, and beer mugs" (suitably interrelated; Shapiro, p. 157). In instructive correspondence with Frege, Hilbert championed the structuralist view of axioms in pure mathematics as defining structures of interest rather than as assertions whose terms must already be understood. In the twentieth century, the development of modern algebra and set theory informed the influential views of the Bourbaki, who explicitly espoused a set-theoretic version of structuralism. Virtually any mathematical structure (or "space," e.g. metric, topological, and so forth) can be conceived or modeled as a set of objects with certain distinguished relations and/or operations on the set, and set theory has the resources for describing a wealth of interrelationships among structures, vital to advanced mathematics. The branch of logic known as *model theory* develops these ideas systematically.

Despite the success of set-theoretic structuralism in providing a unified framework for all major branches of mathematics, as an articulation of structuralism, it confronts certain problems. Notable among these is that it makes a major exception in its own case: despite the multiplicity of set theories (differing over axioms such as well-foundedness, choice, large cardinals, constructibility, and others), the axioms are standardly read as assertions of truths about "the real world of sets" rather than receiving a structuralist treatment. Questions then arise about this "fixed universe as background": How does one know about this real-world structure, how rich it is at its various levels, and how far its levels extend? The (putative) set-theoretic universe cannot be a set; yet as a totality of a different order, is it not indefinitely extendable, contrary to its purported universality? These and related questions have led some philosophers, logicians, and mathematicians to develop alternative ways of articulating structuralism.

ALTERNATIVE ARTICULATIONS TO STRUCTURALISM

The main alternatives to set-theoretic structuralism to be described here are, first, the view of structures as *patterns* or *sui generis universals*, developed by Michael Resnik and Stewart Shapiro, respectively; second, an eliminative, nominalistic *modal structuralism*, traceable in part to Russell and Hilary Putnam and developed by Geoffrey Hellman; and, finally, a version based on *category theory*, as a universal framework for mathematics independent of set theory, suggested by Saunders Mac Lane and others.

THE VIEW OF STRUCTURES AS PATTERNS OR UNIVERSALS. On the view of structures as patterns or universals, apparent reference to special objects in mathematics is taken at face value. Moreover, the reason that such objects are typically identified only by reference to operations and relations within a structure is that in fact they are *inherently incomplete*. They are to be thought of as *positions* or *places* in a pattern, on analogy with, say, the vertices of a triangle. For Resnik, identity and difference among positions make sense only in the context of a structure given by a theory. The number 2, say, is identified as the successor of 1, the predecessor of 3, and so on, but not intrinsically. Indeed, whether the natural number 2 = the real number 2 is indeterminate, except relative to a subsuming structure specified by a broader theory; and then it would still be indeterminate whether the numbers of the new theory were the same as or different from the respective old ones. This theory-relativity of reference and identity—besides leading to complications in the account of the common mathematical practice of embedding structures of a prior theory in those of a later one, as well as in the account of applications of mathematics—reflects Resnik's reluctance to think of patterns as an ontological foundation for mathematics. Talk of patterns may be only analogical, helping free one from the grip of traditional Platonism. Thus, a mathematical theory of structures is not given, in part because *its* objects could not then be identified with those of existing mathematical theories, defeating its purpose.

In contrast, Shapiro takes ontology seriously and develops an axiom system governing the existence of *ante rem* structures, abstract archetypes with places as objects, answering to that which particular realizations have in common. The background logic is second-order and the axioms resemble those of Zermelo-Fraenkel set theory but with an added *Coherence Postulate* guaranteeing an existing structure modeling any *coherent* second-order axiom system, where this new primitive is understood as analogous to the logical notion of *satisfiability*. Knowledge of key instances of this postulate arises naturally, it is argued, from their learning how to use mathematical language together with certain axioms characterizing the structure of interest (e.g. the principle of continuity of the real number system).

Although this view circumvents some of the objections raised against the set-theoretic version, it confronts a number of objections of its own. One (due to Jukka Keränen and John Burgess) points out that, whereas objects in a structure should be distinguishable entirely in terms of internal structural relationships, this is possible

only in cases admitting no nontrivial automorphisms (1-1 structure preserving maps from the class of places to itself other than the identity map). The natural numbers and the reals are "rigid" in this sense, but many nonrigid structures arise in mathematics (e.g. the complex numbers, permuting i and $-i$, or homogeneous Euclidean spaces under isometries, and so forth). A further objection finds a circularity in the account of abstraction offered; the relevant structural relations can only be distinguished from others generated, say, from permutations of objects if those objects (the places) can be picked out independently, contrary to the idea of "structural objects." (This revives a well-known argument of Paul Benacerraf against numbers as objects generally.) Finally, although not committed to any maximal universe of sets, *ante rem* structuralism seems committed to a universe of all places in structures, contrary to the view that any such totality should be extendable.

MODAL-STRUCTURALISM. Turning to modal-structuralism, this view dispenses with special structural objects and indeed even with structures as objects, recognizing instead the possibility that enough objects—of whatever sort one likes—could be interrelated in the right ways as demanded by axioms or conditions appropriate to the mathematical investigation at hand. As suggested by Russell, the irrelevance of any intrinsic features of "mathematical objects" arises through *generalization*: statements "about numbers," for instance, are not about special objects but about whatever objects there might be, collectively standing in the right sort of ordering. By speaking of wholes and parts and utilizing a logic of plurals—reasoning about many things at once without having to talk of sets or classes of them—such generalizations, even over functions and relations, can be framed in nominalistic terms. The effect is to generalize over "structures there might be" without actually introducing structures as entities. Extendability is respected, as it makes no sense to collect "all structures, or items thereof, that there might be." Assuming the logical possibility of countably infinitely many objects, one can recover full classical analysis and, with coding devices, modern functional analysis and more. The main price paid for all of this is the adoption of a primitive notion of *possibility*, something set theory *explains* in terms of the existence of models. The gain is a circumvention of problems of reference to *abstracta* and a natural way of respecting indefinite extendability of mathematical domains.

CATEGORY-THEORETIC STRUCTURALISM. The final approach considered here is based on category theory. Having arisen in mathematics proper to help solve problems in algebraic topology and geometry, it can also serve as a general framework for mathematics. Its basic concepts are *mappings* (*morphisms* or *arrows*) between *objects*, and their *compositions*. The objects are typically what the other approaches call structures, described in relation to other such objects via morphisms ("arrows only"), not internally via set membership. Morphisms typically preserve relevant structure (algebraic, topological, differentiable, and so forth). *Toposes* are families of objects and morphisms with richness comparable to models of Zermelo set theory; they can serve as universes of discourse for mathematics. Generalizations of set-theoretic ideas are provided (such as Cartesian product, function classes, and logical operations, which generally obey intuitionistic laws, i.e. excluding the "law of excluded middle," p or not p, for arbitrary p). In contrast to set theory with its fixed universe, topos theory promotes a pluralistic conception of "many worlds," functionally interrelated (cf. Bell).

It is clear that there are some interesting similarities between category-theoretic structuralism and modal-structuralism, and indeed the latter can be adapted to accommodate the former. Whether category-theoretic structuralism can stand on its own, however, is an open question that turns on such issues as whether its basic concepts are really intelligible without set theory, just what its background logic presupposes, and whether a theory of category of categories can serve as an autonomous framework.

In sum, structuralism has become a major arena for exploring central questions of ontology and epistemology of mathematics.

See also Mathematics, Foundations of; Nominalism, Modern; Realism and Naturalism, Mathematical.

Bibliography

Awodey, S. "Structure in Mathematics and Logic: A Categorical Perspective." *Philosophia Mathematica* 3 (4) (1996): 209–237.

Bell, J. L. "From Absolute to Local Mathematics." *Synthese* 69 (1986): 409–426.

Benacerraf, Paul. "What Numbers Could Not Be." In *Philosophy of Mathematics*, 2nd edition, edited by P. Benacerraf and H. Putnam, 272–294. Cambridge, U.K.: Cambridge University Press, 1983.

Burgess, John P. Review of *Philosophy of Mathematics: Structure and Ontology*, by Stewart Shapiro. *Notre Dame J. Formal Logic* 40 (2) (1999): 283–291.

Dedekind, Richard. *Was sind und was sollen die Zahlen?* Brunswick, Germany: Friedrich Vieweg and Sons, 1888. Translated as *The Nature and Meaning of Numbers*. In *Essays on the Theory of Numbers*, edited by W. W. Beman, 31–115 (New York: Dover, 1963).

Frege, Gottlob. *Wissenschaftlicher Briefwechsel*. Edited by G. Gabriel, H. Hermes, F. Kambartel, and C. Thiel. Hamburg, Germany: Felix Meiner, 1976. Translated as *Philosophical and Mathematical Correspondence* (Oxford: Blackwell, 1980).

Hellman, Geoffrey. "Structuralism." In *The Oxford Handbook of Philosophy of Mathematics and Logic*. Edited by Stewart Shapiro, 536-562. Oxford: Oxford University Press, 2005.

Hilbert, David. *Grundlagen der Geometrie*. Leipzig: Teubner, 1899. Translated by E. Townsend as *Foundations of Geometry* (La Salle, IL: Open Court, 1959).

Keränen, Jukka. "The Identity Problem for Realist Structuralism." *Philosophia Mathematica* 3 (9) (2001): 308–330.

MacBride, Fraser. "Structuralism Reconsidered." In *The Oxford Handbook of Philosophy of Mathematics and Logic*. Edited by Stewart Shapiro, 563-589. Oxford: Oxford University Press, 2005.

Mac Lane, Saunders. *Mathematics: Form and Function*. New York: Springer, 1986.

Putnam, Hilary. "Mathematics without Foundations." In *Philosophy of Mathematics*. 2nd edition. Edited by P. Benacerraf and H. Putnam, 295–311. Cambridge, U.K.: Cambridge University Press, 1983.

Resnik, Michael D. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press, 1997.

Russell, Bertrand. *Introduction to Mathematical Philosophy*. New York: Dover, 1993. First published 1919.

Shapiro, Stewart. *Philosophy of Mathematics: Structure and Ontology*. New York: Oxford University Press, 1997.

Geoffrey Hellman (2005)

STRUCTURALISM AND POSTSTRUCTURALISM

Structuralism emerged as a dominant intellectual paradigm in France in the late 1950s in part in response to the existentialist emphasis on subjectivity and individual autonomy—personified in the work and person of Jean-Paul Sartre—and in part as a reflection of the rising influence of research in the human sciences. In fact, structuralism has its origins in the work of the Swiss linguist Ferdinand de Saussure (1857–1913), whose 1906–1911 lectures at the University of Geneva, published on the basis of student notes in 1916 as the *Cours de linguistique générale*, provide structuralism's basic methodological insights and terminology. While Saussure's *Cours* makes frequent reference to a science that will study language as a system, it was the Russian-born linguist Roman Jakobson (1896–1982) who first used the term *structuralism* in 1929, and it was Jakobson who introduced the basic principles of Saussurean linguistics to both the anthropologist Claude Lévi-Strauss (1908–) and the psychoanalyst Jacques Lacan. Lévi-Strauss and Lacan, along with the philosopher Louis Althusser (1918–1990) and the literary theorist Roland Barthes, together are viewed as the dominant figures in French structuralism whose work in the 1950s and 1960s revolutionized how one thought about the human sciences, psychoanalysis, literature, and Marxism.

What unites these structuralist theorists is less a shared set of philosophical theses than a shared set of methodological assumptions and a willingness to work with the concepts of Saussurean linguistics. Drawing on the four binary oppositions central to Saussurean linguistics—signifier (*signifiant*) and signified (*signifié*), langue and parole, synchronic and diachronic, and infrastructure and superstructure—and privileging in their analyses the former term in each binary pair, the structuralists were able to develop theories that diminished the role of the individual subject or agent while highlighting the underlying relations that govern social and psychic practices.

Saussure defined the linguistic sign as the unity of a sound-image (signifier) and a concept (signified). The signifier is that aspect of a sign that can become perceptible, the psychological imprint of the word-sound or the impression it makes on one's senses, while the signified is a set of psychological associations, the mental picture or description associated with a signifier. In general, then, the signifier is the material (auditory or visual) component of a sign, while the signified is the mental concept associated with that sign. By langue, Saussure meant the set of interpersonal rules and norms that speakers of a language must obey if they are to communicate; langue is the theoretical system or structure of a language like English, French, or Italian. By contrast, parole is the actual manifestation of the system in speech and writing, the speech act, language as used. The distinction between langue and parole is the distinction between structure and event, between a collective product passively assimilated by the individual and the individual act.

By synchronic Saussure named the structural properties of a system at a particular historical moment, while the diachronic referred to the historical dimension of a language, the historical evolution of its elements through various stages. Finally, infrastructure refers to the set of underlying relations that explain the superstructure or surface structure that is open to observation and description. For Saussure, langue functions as the infrastructure

to parole as superstructure, while, oversimplifying greatly, on Althusser's reading of Marx, the relations of means and modes of material production are the infrastructure, while ideology (family, religion, law, social organizations, etc.) is the resultant superstructure, or on Lacan's reconstruction of Sigmund Freud, the dynamic relations among the id, ego, and superego play themselves out at the infrastructural level, while the observable superstructural effects are displayed through behavior.

Their social scientific emphasis on structures also led the structuralists to downplay the role of consciousness, which figured so prominently in existentialism and phenomenology. This deflation of the importance of consciousness and subjectivity—the so-called "death of the subject"—can be seen in all the structuralists' work. Lévi-Strauss's structural analysis of myths, for example, suggests we interpret myths as parole or speech acts that are not the articulations of any particular conscious subject but are instead expressions and variations of a few basic structural relations that form a culture's langue, the set of interpersonal rules and norms that operate unconsciously and that actors in a culture must obey if they are to function. So, in *The Raw and the Crooked* (*Le cru et le cuit*) (1964), Lévi-Strauss analyzes 187 separate myths, showing them all to be variations, transformations, reversals, inversions, and so on of a deep structural opposition between the raw and the cooked, which is itself at the superstructural level of myth the expression of the underlying infrastructural opposition of nature and culture.

This methodological privileging of structure—the underlying rules or general laws—over event—the act of articulating the myth—leads structuralism to place emphasis on synchronic relations rather than on diachronic developments. The structuralists are thus concerned with studying particular systems or structures under somewhat artificial and ahistorical conditions in the hope of explaining their present functioning, as we see in Althusser's concentration on the various ideological state apparatuses at work at a given time in a society rather than the historical evolution of these various cultural formations, or in Barthes's emphasis on writing (*écriture*) as a function that exceeds the author's desire to express or communicate (which Barthes associates with style).

Poststructuralism is the name bestowed in the English-speaking philosophical and literary communities on the ideas of several French philosophers whose work arose as a distinctly philosophical response to the privileging of the human sciences that characterized structuralism. Under the name *poststructuralism* are brought together a number of theorists and theoretical positions that, in France, are often positioned far apart. The name is, however, preferable to either *deconstructionism* or *postmodernism*, which are frequently taken to be synonymous with *poststructuralism* as a rubric under which are grouped together the work of Jacques Derrida, Michel Foucault, Gilles Deleuze, and Jean-François Lyotard, as well as Luce Irigaray, Julia Kristeva, and Hélène Cixous. By contrast, in France only Derrida would be associated with deconstruction, and only Lyotard with postmodernism and, contrary to their English-speaking reception, each of these philosophers is considered to have a distinct project that results only rarely in any two of them being treated together by interpreters sympathetic to their work.

One can locate the emergence of poststructuralism in Paris in the late 1960s: Foucault published *Les mots et les choses: Une archéologie des sciences humaines* in 1966; Derrida published *De la grammatologie*, *L'écriture et la différence* and *La voix et le phénomène* in 1967; and Deleuze published *Différence et répétition* in 1968 and *Logique du sens* in 1969.

While not wanting to overlook the important differences between these thinkers, there are nevertheless certain themes and trends that do emerge in various ways in the work of many of the French philosophers and theorists who follow structuralism. In some cases these should be seen as correctives to the excesses of structuralism, in other cases as various ways in which thinkers coming into prominence in the late 1960s and early 1970s were to give expression to the Nietzschean-Freudian-Marxian spirit of the times, and in still other cases as a way of retrieving themes from some of the French traditions that had fallen out of favor during the scientistic orientation of the 1950s and early 1960s—the return of certain ethical, spiritual, and religious themes, along with some positions associated with phenomenology and existentialism. What cannot be denied, and should not be underestimated, is that the hegemony of structuralist social scientific thinking in the late 1950s and early 1960s was followed by the reemergence of the value of specifically philosophical thinking.

One way to understand their specifically philosophical orientation is to note that while the poststructuralists, like their structuralist predecessors, drew heavily on the ideas of Marx and Freud, unlike the structuralists, they drew at least as much from the third so-called master of suspicion—Friedrich Nietzsche. Nietzsche's critique of truth, his emphasis on interpretation and differential relations of power, and his attention to questions of style

in philosophical discourse became central motifs within the work of the poststructuralists as they turned their attention away from the human sciences and toward a philosophical-critical analysis of writing and textuality (Derrida); relations of power, discourse, and the construction of the subject (Foucault); desire and language (Deleuze); questions of aesthetic and political judgment (Lyotard); and questions of sexual difference and gender construction (Irigaray, Kristeva, and Cixous).

And so, while the structuralist theorists had turned away from philosophy, theorists following structuralism readily identify themselves as philosophers. This is not surprising when one remembers that most of the poststructuralist philosophers "came of age" in an intellectual environment dominated by Sartre's existentialism and they all studied and were profoundly influenced by Maurice Merleau-Ponty's thinking on language and corporeality as well as Martin Heidegger's critique of the history of metaphysics. But unlike most philosophical thinkers in France who preceded the rise of structuralism, French philosophers after structuralism engage in philosophical reflection and analysis while taking account of the institutional and structural forces that inform philosophical thinking itself.

Although it is impossible to locate any set of themes that unite all the poststructuralist philosophers, it would not be inaccurate to note certain motifs that appear frequently in their works: an attention to questions of language, power, and desire that emphasizes the context in which meaning is produced and makes problematic all universal truth and meaning claims; a suspicion toward binary, oppositional thinking, often opting to affirm that which occupies a position of subordination within a differential network; a suspicion toward the figure of the humanistic human subject, challenging the assumptions of autonomy and transparent self-consciousness while situating the subject as a complex intersection of discursive, libidinal, and social forces and practices; and a resistance to claims of universality and unity, preferring instead to acknowledge difference and fragmentation. Situating these philosophical thinkers after structuralism, then, three themes in particular can be highlighted: the return to thinking historically, the return of thinking about the subject, and the emphasis on difference.

THE RETURN TO THINKING HISTORICALLY

There are many ways in which philosophical thinking in France after structuralism can be viewed as a corrective to the overemphasis on synchrony that one finds in structuralist writing. There is no single reason behind this, nor a single form in which French philosophy after structuralism seeks to think time, temporality, or history. But where the structuralists sought to understand the extratemporal functioning of systems (whether social, psychic, economic, or literary), thinkers like Foucault, Derrida, Deleuze, or Lyotard attend to the historical unfolding of the phenomena they choose to examine. In part, the attention to time, temporality, and history can be viewed as a consequence of the intellectual resources to which these thinkers appeal, resources that were not necessarily central to the work of their structuralist predecessors. Foucault, for example, draws on the study of the history of science and scientific change in the work of Georges Canguilhem (1904–1995) and Gaston Bachelard, while Deleuze returns to Henri Bergson's theories of time and *durée* (duration) as well as Nietzsche's eternal return. For Derrida, it is primarily Heidegger's focus on Being and the history of philosophy as a history of the forgetting of the ontological difference (the difference between Being and beings) that leads him to think in terms of the history of metaphysics as a history of logocentrism and ontotheology.

THE RETURN OF THINKING ABOUT THE SUBJECT

Where the rhetoric of the "death of the subject" was characteristic of the structuralists, this was never really the case with most of the philosophers labeled *post*structuralist. To be sure, thinkers like Derrida, Foucault, or Deleuze were never comfortable with the subject-centered thinking of the existentialists or phenomenologists. But they were equally uncomfortable with the straightforwardly antihumanist rhetoric of structuralist thinkers like Althusser or Lévi-Strauss. Thus, Derrida could reply to a question concerning the "death of the subject" that the subject is "absolutely indispensable" and that he does not destroy the subject but situates it in terms of "where it comes from and how it functions."

Even Foucault, who can arguably be associated with the rhetoric of the "death of the subject" in his works of the early 1960s, can at the same time be shown to have been thinking about the question of the construction of the modern subject throughout his oeuvre. That is to say, a distinction can and should be drawn between the "end of man" and the "death of the subject." It may be the case that Foucault's early work engages in thinking the end of man, as we can see, for example, in the closing pages of *The Order of Things* (*Les mots et les choses*). But it would be a mistake to equate the referent of "man" in these early

contexts with what Foucault means by "subject." There is no question that the subject named "man" in philosophical discourse, from René Descartes's *cogito* to Immanuel Kant's autonomous rational moral agent, is a concept toward which Foucault has little sympathy. But even in a supposedly antihumanist work like the essay "What Is an Author?" (1969) Foucault's desire to deflate the subject as epistemically and discursively privileged is not conjoined with an attempt to eliminate the subject entirely. Instead, Foucault seeks to analyze the subject as a variable and complex function of discourse and power, which, he writes, means to ask not "How can a free subject penetrate the substance of things and give it meaning?" but "How, under what conditions and in what forms, can something like a subject appear in the order of discourse? What place can it occupy in each type of discourse, what functions can it assume, and by obeying what rules?"

What this means, and what has been largely misunderstood by many of Foucault's critics, is that his so-called antihumanism was not a rejection of the human *per se*; it was instead an assault on the philosophically modern idea that sought to remove man from the natural world and place him in a position of epistemic, metaphysical, and moral privilege that earlier thought had set aside for God. Foucault's work is less an antihumanism than an attempt to think humanism and the subject after the end of (modern) man. Far from being a thinker of the "death of the subject," Foucault simply refuses to accept the subject as given, as the foundation for ethical or rational thinking. The subject is, instead, something that has been historically created and Foucault's work, in its entirety, is engaged in analyzing the various ways that human beings are transformed into subjects, whether subjects of knowledge, of power, of sexuality, or of ethics.

For feminist thinkers writing after structuralism, the question of the subject was also central to their work as they sought to challenge both philosophical and psychoanalytic assumptions concerning the subject as sexed or gendered male or masculine. The feminists don't object to the subject simply being sexed or gendered; it is the subject's being sexed/gendered male that is the object of their criticisms. Although there are important differences between the theoretical positions of Cixous, Irigaray, or Kristeva, insofar as these "difference feminists" argue for sexual difference and the significant and important differences between male and female desire, they had to argue that there were important differences between male and female subjects. And to make this argument required that they refuse to follow the structuralist project of entirely eliminating the subject.

So, for example, while Irigaray acknowledges that insofar as the logic of subjectivity has relegated women to the position of object, one should not give up the possibility of occupying the position of the subject insofar as this is a position that women have heretofore never been able to occupy. In fact, she suggests that insofar as the circulation of women as objects of social-sexual exchange has been foundational to the Western patriarchal social order, one should not underestimate the possibilities for radical social transformation if women were to finally emerge as "speaking subjects."

The "speaking subject" is also a central focus of Kristeva's work, as she defines her project of analytical semiology or semanalysis, in part, as one of reinserting subjectivity into matters of language and meaning. Such a subject would not, of course, be a Cartesian or Husserlian subject, who could function as a pure source of meaning. Rather, following the discoveries of Freud, Lacan, and structural linguistics (Saussure and Émile Benveniste [1902–1976]), the "speaking subject" will always be a "split subject," split between conscious motivations and the unconscious, between structure and event, and between the subject of the utterance (*sujet d'énonciation*) and the subject of the statement (*sujet d'énoncé*). Elsewhere, in *Revolution and Poetic Language* (*La révolution du langage poétique: L'avant-garde à la fin du XIXe siècle, Lautréamont et Mallarmé*) (1974), this subject is developed as a subject-in-process/on-trial (*sujet-en-procès*), a dynamic subject at the intersection of the semiotic and the symbolic, making itself and being made, but a subject nonetheless.

THE EMPHASIS ON DIFFERENCE

One of the essential themes of Saussure's linguistics was that "in language there are only differences *without positive terms*" (Saussure 1959). By this, he meant that language functions as a system of interdependent units in which the value of each constituent unit results solely from the simultaneous presence of other units and the ways each unit differs from the others. This attention to difference led the structuralists to emphasize in their analyses relations rather than things and to focus on the differential relations between the objects they studied rather than the objects themselves. While the structuralists all took note of this theme, the emphasis on difference did not become truly dominant until after the hegemony of the structuralist paradigm began to wane. It has already been noted that sexual difference is a theme that almost all the feminist thinkers after structuralism have addressed. Indeed, Irigaray goes so far as to suggest that,

if Heidegger is right in thinking that each epoch has but a single issue to think through, then "sexual difference is … the issue of our age" (Irigary 1993, p. 5). Similarly, Cixous sees the rigid conceptualization of sexual difference as what supports the identification of the male/masculine with the Same, while the female/feminine is rendered Other. For Cixous, the way out of this patriarchal system is not through the elimination of difference but through escaping the dominant logic of difference as hierarchal opposition to a new logic of difference in which "difference would be a bunch of new differences" (Cixous and Clément 1986).

Sexual difference is only one form in which the poststructuralist attention to difference has appeared. Insofar as Derrida's philosophical project began as an attempt to deconstruct the logocentric history of metaphysics as a metaphysics of presence that invariably privileges the temporal present, his coining of the neologism *différance* sought to situate at the foundation of deconstructive analysis an attention to difference by highlighting both meanings of the French verb *différer*: to defer in terms of delay over time and to differ in terms of spatial nonidentity. Insofar as *différance* names the movement of both temporal deferring and spatial differing, it stands as the transcendental condition for the possibility of differentiation, that is, *différance* is what makes differences possible.

This attention to difference—rather than a focus on identity or the Same—is particularly central to the projects of Lyotard and Deleuze. For Deleuze, whose work often takes a form of presentation much more in the mold of traditional philosophical analysis than the other philosophers writing after structuralism, difference has been a central and constant focus of his thinking. His *Nietzsche et la philosophie* (1962), which was the first of the major French interpretations of Nietzsche to appear, appeals to the concept of difference to show how Nietzsche departs from the Hegelian tradition (where Hegel's dialectic supersedes difference, Nietzsche's philosophy affirms it), to explicate Nietzsche's will to power (as the differential element between active and reactive forces), and to interpret Nietzsche's thought of the eternal recurrence (not as the eternal return of the same but as the repetition of difference). Deleuze develops these themes much further in *Différence et répétition* as he attempts to think the concept of difference in itself while challenging the metaphysical tradition for associating difference with opposition and the negative and privileging identity and the Same as primary.

For Lyotard, whose work is more closely tied to postmodernism than the other French philosophers, what characterizes the postmodern, as he puts it in the introduction to *The Postmodern Condition* (*La condition postmoderne: Rapport sur le savoir*) (1979), is an "incredulity toward metanarratives." Rather than naming a specific epoch, the postmodern names, instead, an antifoundationalist attitude that exceeds the legitimating orthodoxy of the moment. Postmodernity, then, does not follow modernity but resides constantly at the heart of the modern, challenging those totalizing and comprehensive master narratives (like the Enlightenment narrative of the emancipation of the rational subject or the Marxist narrative of the emancipation of the working class) that serve to legitimate its practices. In place of these grand meta- and master narratives, Lyotard suggests one looks instead to less ambitious "little narratives" that refrain from totalizing claims in favor of recognizing the specificity and singularity of events. To refuse to sanction the move to a metanarrative in the ethical, political, aesthetic, and metaphysical domains commits one to a philosophy of difference in that it accepts that oppositions will not be resolved in some higher unity and concludes that multiple and discordant voices are not only inevitable but desirable.

Beyond his postmodernist polemic, reflecting on difference operates at the core of what Lyotard considered his most important work, *Le Différend* (1983), in which he attempts to account for radical and incommensurable differences in the discourses of ethics and politics, that is, those incommensurable differences that will not admit any shared standard to which one could appeal in making judgments concerning what is different. The *différend* is thus defined as "a case of conflict, between (at least) two parties, that cannot be equitably resolved for lack of a rule of judgment applicable to both arguments" (Lyotard 1988). For Lyotard, once one has given up on master narratives, one must also give up on a master narrative of justice or the good to which all parties will agree. While such a master narrative is presupposed for a democratic politics based on consensus and agreement, the political question for Lyotard is ultimately the question of how to make decisions in the case of a *différend* in which, by definition, no consensus is possible. The choice, it would seem, is either violence or a new kind of political thinking that can accommodate *différends* in a shared social space where norms work to minimize evil rather than maximize good and where evil is itself defined in terms of the continued interdiction of different possibilities.

INFLUENCE

The impact of poststructuralism on philosophy, aesthetics, literary studies, and social theory has been extensive. While Continental philosophy was, during the 1970s, dominated by issues related to phenomenology, existentialism, and the works of Edmund Husserl, Heidegger, and Sartre, in the early 2000s the scope of Continental philosophy is increasingly focused on issues that originate in the works of post-1960 French thinkers. Derrida, and deconstruction, has been a major force in literary theory and criticism since the early 1970s. Since the early 1980s, Derrida has become a major influence in philosophical studies and he and Foucault have had the widest influence on English-language writers. Since 1980 other poststructuralist texts have appeared in translation and, as a consequence, we now see the impact on philosophers of Deleuze's important and innovative readings of major philosophical figures (David Hume, Benedict [Baruch] de Spinoza, Gottfried Wilhelm Leibniz, Kant, Nietzsche, and Bergson) as well as his analyses, alone and in collaboration with Félix Guattari, of psychoanalysis, cinema, art, literature, and contemporary culture; Lyotard's essays on politics, aesthetics, and art history, plus his important reflections on Kant's *Critique of Judgment* and questions of modernity and postmodernity; Irigaray's critical rereadings of Freud, the philosophical canon (Plato, Descartes, Heidegger, Nietzsche, Merleau-Ponty, and Emmanuel Levinas), and her reflections on language and sexual difference; Cixous's engendering writing and reflecting on its relations to the body, particularly the feminine body; and Kristeva's thinking on semiotics, abjection, and desire in language.

See also Art, Interpretation of; Deconstruction; Derrida, Jacques; Foucault, Michel; Literature, Philosophy of.

Bibliography

PRIMARY WORKS

Althusser, Louis. *Pour Marx*. Paris: F. Maspero, 1965. English-language translation: *For Marx*. Translated by Ben Brewster. London: Verso, 1990.

Cixous, Hélène, and Catherine Clément. *La jeune née*. Paris: Union générale d'Éditions, 1975. English-language translation: *The Newly Born Woman*. Translated by Betsy Wing. Minneapolis: University of Minnesota Press, 1986

Deleuze, Gilles. *Différence et répétition*. Paris: Presses universitaires de France, 1968. English-language translation: *Difference and Repetition*. Translated by Paul Patton. New York: Columbia University Press, 1994.

Deleuze, Gilles. *Logique du sens*. Paris: Éditions de Minuit, 1969. English-language translation: *The Logic of Sense*, edited by Constantin V. Boundas. Translated by Mark Lester with Charles Stivale. New York: Columbia University Press, 1990.

Deleuze, Gilles. *Nietzsche et la philosophie*. Paris: Presses universitaires de France, 1962. English-language translation: *Nietzsche and Philosophy*. Translated by Hugh Tomlinson. New York: Columbia University Press, 1983.

Deleuze, Gilles, and Félix Guattari. *Capitalisme et schizophrénie. 1. L'anti-Oedipe*. Paris: Éditions de Minuit, 1972. English-language translation: *Anti-Oedipus*. Translated by Robert Hurley, Mark Seem, and Helen R. Lane. New York: Viking Press, 1977.

Deleuze, Gilles, and Félix Guattari. *Capitalisme et schizophrénie. 2. Mille plateaux*. Paris: Éditions de Minuit, 1980. English-language translation: *A Thousand Plateaus: Capitalism and Schizophrenia*. Translated Brian Massumi. Minneapolis: University of Minnesota Press, 1987.

Deleuze, Gilles, and Félix Guattari. *Qu'est-ce que la philosophie?* Paris: Éditions de Minuit, 1991. English-language translation: *What Is Philosophy?* Translated by Hugh Tomlinson and Graham Burchell. New York: Columbia University Press, 1994.

Derrida, Jacques. *De la grammatologie*. Paris: Éditions de Minuit, 1967. English-language translation: *Of Grammatology*. Translated by Gayatri Chakravorty Spivak. Baltimore, MD: Johns Hopkins University Press, 1976.

Derrida, Jacques. *L'écriture et la différence*. Paris: Éditions du Seuil, 1967. English-language translation: *Writing and Difference*. Translated by Alan Bass. Chicago: University of Chicago Press, 1978.

Derrida, Jacques. *Marges de la philosophie*. Paris: Éditions de Minuit, 1972. English-language translation: *Margins of Philosophy*. Translated by Alan Bass. Chicago: University of Chicago Press, 1982.

Foucault, Michel. *The Essential Works of Michel Foucault, 1954–1988*. 3 vols., edited by Paul Rabinow. *Vol. 1: Ethics: Subjectivity and Truth*, edited by Paul Rabinow; translated by Robert Hurley (New York: New Press, 1997). *Vol. 2: Aesthetics, Method, and Epistemology*, edited by James D. Faubion; translated by Robert Hurley (New York: New Press, 1998). *Vol. 3: Power*, edited by James D. Faubion; translated by Robert Hurley (New York: New Press, 2000).

Foucault, Michel. *Histoire de la sexualité. 1. La volonté de savoir*. Paris: Éditions Gallimard, 1976. English-language translation: *The History of Sexuality. Vol. One: An Introduction*. Translated by Robert Hurley. New York: Vintage Books, 1980.

Foucault, Michel. *Les mots et les choses: Une archéologie des sciences humaines*. Paris: Éditions Gallimard, 1966. English-language translation: *The Order of Things: An Archeology of the Human Sciences*. New York: Vintage Books, 1970.

Foucault, Michel. *Surveiller et punir: Naissance de la prison*. Paris: Éditions Gallimard, 1975. English-language translation: *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books, 1977.

Irigaray, Luce. *Ce sexe qui n'en est pas un*. Paris: Éditions de Minuit, 1977. English-language translation: *This Sex Which Is Not One*. Translated by Catherine Porter with Carolyn Burke. Ithaca: Cornell University Press, 1985.

Irigaray, Luce. *An Ethis of Sexual Difference*. Translated by Carolyn Burke and Gillian C. Gill. Ithaca, NY: Cornell University Press, 1993.

Irigaray, Luce. *Speculum de l'autre femme*. Paris: Éditions de Minuit, 1974. English-language translation: *Speculum of the Other Woman*. Translated by Gillian C. Gill. Ithaca: Cornell University Press, 1985.

Kristeva, Julia. *Desire in Language: A Semiotic Approach to Literature and Art*, edited by Leon S. Roudiez. Translated by Thomas Gora, Alice Jardine, and Leon S. Roudiez. New York: Columbia University Press, 1980.

Kristeva, Julia. *Pouvoirs de l'horreur: Essai sur l'abjection*. Paris: Éditions du Seuil, 1980. English-language translation: *Powers of Horror: An Essay on Abjection*. Translated by Leon S. Roudiez. New York: Columbia University Press, 1982

Kristeva, Julia. *La révolution du langage poétique: L'avant-garde à la fin du XIXe siècle, Lautréamont et Mallarmé*. Paris: Éditions du Seuil, 1974. English-language translation: *Revolution in Poetic Language*. Translated by Margaret Waller. New York: Columbia University Press, 1984.

Lacan, Jacques. *Écrits*. Paris: Éditions du Seuil, 1966. English-language translation: *Ecrits: A Selection*. Translated by Alan Sheridan. New York: W. W. Norton, 1977.

Lévi-Strauss, Claude. *Anthropologie structurale*. Paris: Plon, 1958. English-language translation: *Structural Anthropology*. Translated by Claire Jacobson and Brooke Grundfest Schoepf. New York: Basic Books, 1963.

Lévi-Strauss, Claude. *Mythologiques. 1. Le Cru et le cuit*. Paris: Plon, 1964. English-language translation: *The Raw and the Cooked*. Translated by John and Doreen Weightman. New York: Harper and Row, 1969

Lévi-Strauss, Claude. *Les structures élémentaires de la parenté*. Paris: Presses universitaires de France, 1949. English-language translation: *The Elementary Structures of Kinship*. Edited by Rodney Needham. Translated by James Harle Bell and John Richard von Sturmer. Boston: Beacon Press, 1969.

Lyotard, Jean-François. *La condition postmoderne: Rapport sur le savoir*. Paris: Éditions de Minuit, 1979. English-language translation: *The Postmodern Condition. A Report on Knowledge*. Translated by Geoff Bennington and Brian Massumi. Minneapolis: University of Minnesota Press, 1984.

Lyotard, Jean-François. *Le différend*. Paris: Éditions de Minuit, 1983. English-language translation: *The Differend: Phrases in Dispute*. Translated by Georges Van Den Abbeele. Minneapolis: University of Minnesota Press, 1988.

Lyotard, Jean-François. *Discours, figure*. Paris, 1971.

Lyotard, Jean-François. *Economie libidinale*. Paris: Éditions de Minuit, 1974. English-language translation: *Libidinal Economy*. Translated by Iain Hamilton Grant. Bloomington: Indiana University Press, 1993.

Saussure, Ferdinand de. *Cours de linguistique générale*. Paris: Payot, 1915. English-language translation: *Course in General Linguistics*, edited by Charles Bally and Albert Sechehaye with the collaboration of Albert Riedlinger. Translated by Wade Baskin. New York: Philosophical Library, 1959.

SECONDARY LITERATURE

Bennington, Geoff. *Lyotard: Writing the Event*. Manchester, 1988.

Bogue, Ronald. *Deleuze and Guattari*. New York: Routledge, 1989.

Bogue, Ronald. *Deleuze on Cinema*. New York: Routledge, 2003.

Bogue, Ronald. *Deleuze on Literature*. New York: Routledge, 2003.

Bogue, Ronald. *Deleuze on Music, Painting and the Arts*. New York: Routledge, 2003.

Caws, Peter. *Structuralism: The Art of the Intelligible*. Atlantic Highlands, NJ: Humanities Press International, 1988.

Critchley, Simon, and William Schroeder, eds. *A Companion to Continental Philosophy*. Oxford: Blackwell, 1999.

Descombes, Vincent. *Le même et l'autre: Quarante-cinq ans de philosophie française (1933–1978)*. Paris: Éditions du Minuit, 1979. English-language translation: *Modern French Philosophy*. Translated by L. Scott-Fox and J. M. Harding. Cambridge, U.K.: Cambridge University Press, 1980.

Dews, Peter. *Logics of Disintegration: Post-Structuralist Thought and the Claims of Critical Theory*. London: Verso, 1987.

Dosse, Francois. *History of Structuralism. Vol. 1. The Rising Sign, 1945–1966. Vol. 2. The Sign Sets, 1967–present*. Translated by Deborah Glassman. Minneapolis: University of Minnesota Press, 1997.

Frank, Manfred. *Was ist Neostrukturalismus?* Frankfurt: Suhrkamp, 1983. English-language translation: *What Is Neostructuralism?* Translated by Sabine Wilke and Richard. Minneapolis: University of Minnesota Press, 1989.

Gutting, Gary. *French Philosophy in the Twentieth Century*. Cambridge, U.K.: Cambridge University Press, 2001.

Hardt, Michael. *Gilles Deleuze: An Apprenticeship in Philosophy*. Minneapolis: University of Minnesota Press, 1993.

Hoy, David Couzens, ed. *Foucault: A Critical Reader*. Oxford: Blackwell, 1986.

Lawlor, Leonard. *Thinking through French Philosophy: The Being of the Question*. Bloomington: Indiana University Press, 2003.

Lechte, John. *Fifty Key Contemporary Thinkers: From Structuralism to Postmodernity*. New York: Routledge, 1994.

Macksey, Richard, and Eugenio Donato. *The Structuralist Controversy: The Languages of Criticism and the Sciences of Man*. Baltimore, MD: Johns Hopkins Press, 1972.

Massumi, Brian. *A User's Guide to Capitalism and Schizophrenia. Deviations from Deleuze and Guattari*. Cambridge, MA, 1992.

Patton, Paul, ed. *Deleuze: A Critical Reader*. Oxford: Blackwell, 1996.

Readings, Bill. *Introducing Lyotard: Art and Politics*. New York: Routledge, 1991.

Schrift, Alan D. *Twentieth Century French Philosophy: Key Themes and Thinkers*. Oxford: Blackwell, 2005.

Sturrock, John, ed. *Structuralism and Since: From Lévi-Strauss to Derrida*. Oxford: Oxford University Press, 1979.

Williams, James. *Lyotard: Towards a Postmodern Philosophy*. Cambridge: Polity Press, 1998.

Alan D. Schrift (1996, 2005)

STUMPF, KARL

(1848–1936)

Karl Stumpf, the German psychologist and philosopher, was born in Wiesentheid, Bavaria. He studied law at Würzburg, but under the influence of Franz Brentano his interests turned to philosophy and psychology. In 1868 he took a degree at Göttingen, under Rudolf Hermann Lotze, with a dissertation on the relation between Plato's God and the Idea of the Good. In 1869 he entered the Catholic seminary in Würzburg, where he studied St. Thomas Aquinas and the Scholastics. A year later, having lost his faith in orthodox Christianity and having abandoned the idea of becoming a priest, he left the seminary and became docent at Göttingen, where he taught for three years. His acquaintances included the philosopher and psychologist Gustav Fechner, who used Stumpf as a subject for his experiments in aesthetics.

Stumpf's passionate fondness for music motivated his pioneering research in the psychology of sound perception. In 1873 he became professor of philosophy at Würzburg and in 1879, at Prague. His associates included Ernst Mach and Anton Marty. In 1884 he moved to Halle, where Edmund Husserl (who later dedicated his *Logische Untersuchungen* to Stumpf) became his student. Stumpf moved to Munich in 1889, but his heretical religious views made him uncongenial to some of his orthodox colleagues and to the authorities. He therefore accepted a professorship in Berlin in 1894. There he founded the Phonogram Archive, devoted to collecting recordings of primitive music, and the Psychological Institute, and for a time he directed research in Immanuel Kant and Gottfried Wilhelm Leibniz at the Academy of Sciences. Besides Husserl, his most famous student was Wolfgang Köhler, the Gestalt psychologist. William James, who praised Stumpf's *Tonpsychologie* very highly, was a friend and correspondent.

Stumpf contributed greatly to the development of psychology from a branch of philosophy into an empirical science. His own experimental work was largely concerned with acoustical phenomena, but he also wrote on other topics in psychology, such as the theory of emotions. As a philosopher, Stumpf was an empiricist who preferred John Locke and George Berkeley to the tradition of German idealism. He praised Kant for emphasizing the concepts of necessity and duty but rejected the view that the categories are a priori (by which Stumpf meant innate) and not derived from perceptions. The category of substance, or "thing," he maintained, is a concept that can be traced back to such actual experiences as that of perceiving the close interpenetration of the parts of a whole. The constituent characteristics of a sensory feeling, such as quality and intensity, form a whole rather than a mere aggregate. Experience includes the perceiving of relations; it does not consist merely of individual sensations that need to be related by the understanding.

In the realm of mental functions, all simultaneous states of consciousness and intellectual and emotional activities are perceived as a unity. The concept of a substance, whether of a physical or a psychical substance, is not that of a bundle of qualities, as with David Hume, but is a unity of qualities and relations. As for the concept of cause, Stumpf believed that both Kant and Hume were wrong; we can sometimes actually perceive a causal nexus as opposed to a mere sequence, and this experience is the origin of the category of cause. For example, when our thought processes are governed by some interest or mood, we do not first experience the interest and only subsequently its effects; rather, we are aware of the interest and its effects all at once. Thus we directly experience causality in our own internal activity. Without this we would not be conscious of reality. We transfer this awareness of causality to natural phenomena, although this projection is superfluous for scientific purposes where only lawlike sequences of events are needed.

Stumpf accepted a dualism of mind and nature but regarded the task of philosophy as the investigation of what mind and nature have in common. Philosophy is the science that studies the most general laws of the psychical and of the real. To be real means to have effects. The reality of our own mental states is the first datum. We recognize the reality of external objects as they affect us, having first acquired the idea of causality internally.

From Brentano, Stumpf took the fundamental notion of self-evidence. We experience the self-evidence of such judgments as $2 \times 2 = 4$, and this self-evidence cannot be further reduced. It is the subjective aspect of truth. Truth itself is that property of contents of consciousness whereby they compel assent. Truth is a function of that which is thought, not a function of the thinker. Stumpf explicitly rejected the positivist and pragmatist theories of truth.

Knowledge is of two sorts, a priori and a posteriori. A priori knowledge consists of deductions from self-evident propositions and from bare concepts. It ought to be expressed in hypothetical propositions, since no determination of fact is here made. Mathematical knowledge is of this type. If there are more geometries than one, all are a priori; only their applicability to objective space is an empirical question. A priori knowledge may be secured

from any concept. The mere concept of three tones implies a definite order according to which a tone of one pitch must be located between the other two. The concept of a tone series contains the possibility of its continuation ad infinitum. These are propositions that we know but that neither have nor require proof. They are analytic, not only known by means of our concepts but known because they are about our concepts. A posteriori knowledge, on the other hand, is of facts and laws. Both sensory contents and mental activities or functions are experienced directly. Stumpf introduced the term *Sachverhalte* (state of affairs) into philosophy, although he claimed only to have replaced Brentano's notion of "content of judgment" with the term.

Stumpf rejected the idea of vitalism or of any sort of life force, although he did not oppose empirical psychovitalism, the view that feelings, thoughts, and volitions can be stimuli for physical nerve processes. He argued that evolution did not dispose of the problem of teleology, since life itself, whose origin from nonliving atoms is so mathematically improbable, requires an explanatory hypothesis.

Bibliography

WORKS BY STUMPF

Über den psychologischen Ursprung der Raumvorstellung. Leipzig: Hinzel, 1873.

Tonpsychologie. 2 vols. Leipzig: Hirzel, 1883 and 1890.

Leib und Seele. Leipzig: Barth, 1903.

Erscheinungen und psychische Funktionen. Berlin, 1907.

Philosophische Reden und Vorträge. Leipzig: Barth, 1910. Contains various papers on evolution, the aesthetics of tragedy, and child psychology.

Die Anfänge der Musik. Leipzig: Barth, 1911.

Empfindung und Vorstellung. Berlin, 1918.

Franz Brentano. Munich: Beck, 1919.

Spinozastudien. Berlin, 1919.

Autobiography in *Die Philosophie der Gegenwart in Selbstdarstellung,* edited by Raymund Schmidt, Vol. V, 205–265. Leipzig, 1924. Gives the best summary of his philosophical views. An English translation is available in Charles Murchison, *Psychology in Autobiography,* Vol. I, 389–441. Worcester, MA, 1930.

William James. Berlin: Pan-verlag Rolf Heise, 1927.

Erkenntnislehre. 2 vols. Leipzig: Barton, 1939–1940.

WORKS ON STUMPF

Boring, E. G. *History of Experimental Psychology.* New York and London: Century, 1929. Gives fullest account of Stumpf's psychology and his classification of experiences in relation to that of Husserl and Brentano.

Spiegelberg, Herbert. *The Phenomenological Movement.* 2 vols. The Hague: Nijhoff, 1960. Vol. I.

Titchener, E. B. *Experimental Psychology.* New York and London: Macmillan, 1905. Vol. II, 161–163. Stumpf's psychophysics.

Titchener, E. B. "Prof. Stumpf's Affective Psychology." *American Journal of Psychology* 28 (1917): 263–277. Stumpf's theory of emotions.

Arnulf Zweig (1967)

STURZO, LUIGI

(1871–1959)

Luigi Sturzo, the Italian political figure and philosopher who elaborated a systematic historical anthropology, was born in Caltagirone, Sicily. He was ordained a priest in 1894 and received a doctorate in philosophy from the Gregorian University in Rome in 1898. He taught philosophy at the seminary in Caltagirone from 1898 to 1903. Sturzo served as mayor of Caltagirone from 1905 to 1920. He founded the Italian Popular Party in 1919 and served as its political secretary from 1919 to 1923. As early as 1926, in *Italy and Fascism,* Sturzo exposed the total economic concentration of power in the ruling radical right and the method of violence by which the power elite governed. His major works were written in exile in the period from 1924 to 1946 in Paris, London, and New York and were first published in translations. In recognition of his historic role in the birth of the Italian Republic, Sturzo was named a senator for life in 1952.

In philosophy Sturzo elaborated a "dialectic of the concrete" based primarily on the thought of St. Augustine, Gottfried Wilhelm Leibniz, Giambattista Vico, and Maurice Blondel. He opposed this dialectic to both absolute idealism, which he regarded as a necessitarian monism, and scholastic realism, which he considered a spectatorlike abstractionism. At the basis of his thought is historical man projected into "the fourth dimension, that of time." Man is at one and the same time individual and social, free and conditioned, structural and in process; he is a singular history in process rather than a nature fixed in essence. Man is never pure becoming, however, but a radical tendency toward reason in action.

Organically, man is constitutionally relational in his total organic connections. Socially, he is a manifold and simultaneous projection of collective purposes that are made concrete in social structures that embody his many needs in a dynamic interplay of primary and subsidiary associations.

When collective purposes become institutionalized and each social form presses for exclusive domination,

conflicts are engendered. If one form gains such domination, forces of renewal and reform are unwittingly released. Thus, driven by precarious and incomplete achievements, man advances by conquering new dimensions of experience, both personal and collective.

The most radical novelty and the most powerful solvent of conflicting interests is the concrete ingression of the divine into the total human process. This "historicization of the divine" in its empirical reality is both singular and collective and constitutes the driving force of human progress.

Although he recognized the recurrence of regression, Sturzo professed an enlightened optimism, similar to that of Pierre Teilhard de Chardin, born out of his vision of one humankind moving toward ever greater socialization through the growth of international consciousness as revealed in the rationalization of force and the repudiation of war.

Bibliography

WORKS BY STURZO

Opera Omnia. Edited under the auspices of the Istituto Luigi Sturzo. Bologna, 1953–. Of a projected 30 vols., 12 have been published.

The following works were first published in Italian in *Opera Omnia,* First Series.

La comunità internazionale e il diritto di guerra (1928). Vol. II. 1954. Translated by Barbara B. Carter as *The International Community and the Right of War.* London: Allen and Unwin, 1929.

La società: Sociologica storicista (1935). Vol. III. 1960. Translated by Barbara B. Carter as *The Inner Laws of Society.* New York: P.J. Kenedy, 1944.

Chiesa e stato: Studio sociologico-storico (1937). Vols. V–VI. 1959. Translated by Barbara B. Carter as *Church and State.* New York, 1939.

La vera vita (1943). Vol. VII. 1960. Translated by Barbara B. Carter as *The True Life.* Washington, DC: Catholic University of America Press, 1943.

Del metodo sociologico. Bergamo, Italy, 1950.

WORKS ON STURZO

De Rosa, Gabriele. *Storia del partito popolare.* Bari: Laterza, 1958.

Di Lascia, Alfred. "Sturzo." *Cross Currents* 9 (1959): 400–410.

Pollock, Robert C., ed. "Luigi Sturzo: An Anthology." *Thought.* 28 (1953): 165–202. Pollock's commentary is invaluable.

Timasheff, Nicholas S. *The Sociology of Luigi Sturzo.* Baltimore: Helicon Press, 1962.

Walsh, John, and Joan Quick. "A Sturzo Bibliography." *Thought* 28 (1953): 202–208.

Alfred Di Lascia (1967)

SUÁREZ, FRANCISCO

(1548–1617)

Francisco Suárez, the Spanish scholastic philosopher and theologian, "Doctor Eximius," was born at Granada. His father was a wealthy lawyer and Francisco was the second of eight sons, six of whom entered the religious life. In 1564 he applied for admission to the Jesuit order. Perhaps because of ill health he showed little promise at first, and he failed to pass the examinations. Suárez appealed the verdict of his examiners, but his second examinations were not much better than the first. The provincial agreed, however, to admit Suárez at a lower rank. Shortly after his admission to the order, he began his study of philosophy. He showed little promise in the next few months and considered abandoning his studies for a lesser occupation in the order. However, he was persuaded by his superior to continue his studies, and within the next few years he became an outstanding student. Completing his course in philosophy with distinction, he transferred to the theology curriculum at the University of Salamanca and soon became an outstanding theologian.

In 1571 he was appointed professor of philosophy at the Jesuit college in Segovia and shortly thereafter was ordained to the priesthood. From 1576 to 1580 he served at the University of Valladolid and was then honored with an appointment to the chair of theology at the Jesuit college in Rome. Five years later he was transferred to a similar chair at the University of Alcalá. He had now achieved considerable reputation as a theologian and in 1593 was singled out by Philip II of Spain for appointment to the chair of theology at the University of Évora in Portugal. The years at Évora saw the publication of such major works as the *Disputationes Metaphysicae* (1597); the *De Legibus ac Deo Legislatore* (1612); the *Defensor Fidei* (1613), a refutation of the *Apologia* of King James I of England; and the *Varia Opuscula Theologica* (1599), which embodied Suárez's contributions to the congruist movement. In 1616 Suárez retired from active teaching; he died the following year.

At the time of his death, Suárez's reputation as a philosopher and theologian was extraordinary, and his metaphysics dominated thought at Catholic and many Protestant universities for the next two centuries. René Descartes is said to have carried a copy of the *Disputationes* with him during his travels. The *Ontologia* of Christian Wolff owed much to Suárez, and Gottfried Wilhelm Leibniz read him avidly. Arthur Schopenhauer declared that the *Disputationes* was an "authentic com-

pendium of the whole scholastic wisdom." After Thomas Aquinas, to whom he owed much, Suárez is generally recognized as the greatest of the Scholastics. His philosophy will be considered under two headings, the metaphysics (including epistemology) and the philosophy of law.

METAPHYSICS

The metaphysics of Suárez is basically Aristotelian and Thomistic yet also highly original. It reveals remarkable erudition and a profound knowledge of his medieval predecessors. Some of the outstanding features of Suárez's metaphysics may be shown in a brief exposition of his views on the nature of metaphysics, the theory of distinctions, the principle of individuation, the problem of universals, the knowledge of singulars, the doctrine of analogy, the existence of God, and the problem of freedom.

NATURE OF METAPHYSICS. Suárez defined metaphysics as the science of being qua being. Taken as a noun, being signifies a real essence; taken as a participle, being refers to the act of existing. A real essence is noncontradictory, and by real Suárez means that which can or actually does exist in reality. Being may also be distinguished as real being and conceptual being. Real being may be immaterial, material, substantial, or accidental. The concept of being is analogical, derived from knowledge of the various kinds of real being; it is not univocal. The metaphysician is concerned primarily with immaterial being, and metaphysics is necessary for an understanding of sacred theology.

THEORY OF DISTINCTIONS. Like his predecessors Suárez held that in God essence and existence are one. Aquinas held that in finite beings essence and existence are really distinct. Suárez, however, maintained that the distinction is solely one of reason, a mental or logical distinction, for to assert a real distinction presupposes a knowledge of existence, and this would entail an essence of existence. To the Thomist objection that the denial of real existence destroys the contingency of created beings, Suárez replied that it is unnecessary to add a real distinction to establish the contingency, for it is in the nature of created being to be contingent. The emphasis upon essence in contrast to existence led Étienne Gilson to refer to Suárez's metaphysics as "essentialistic" in contrast to the "existentialistic" metaphysics of Aquinas.

PRINCIPLE OF INDIVIDUATION. The principle of individuation is neither the *materia signata* of Aquinas nor the *haecceitas* of John Duns Scotus, although Suárez agreed with Scotus that "individuality adds to the common nature [essence] something which is mentally distinct from that nature … and which together with the nature constitutes the individual metaphysically." In composite substances both form and matter individuate, for the essence of the individual is made up of both matter and form, with form the principal determinant. Individuals may be distinguished on the basis of their matter—for example, quantity—but their individuation is determined by form and matter, not by our mode of cognition.

PROBLEM OF UNIVERSALS. Universals have no existence either in reality or in individuals. There are only individuals; universals do have a foundation in reality, however, for the mind abstracts them from the likenesses of individuals. Suárez criticizes the Ockhamists for insisting that universals are only words or mental constructs, but it is difficult to dissociate his position from theirs, for he strongly insists that there are as many essences as individuals and that each individual being is an individual essence.

KNOWLEDGE OF SINGULARS. With Scotus, Suárez maintained that the intellect has a direct knowledge of singulars. "Our intellect knows the individual material object by a proper species of it … our intellects know individual material objects without reflection." Suárez maintained that the active intellect can have this kind of knowledge, for there is nothing contradictory about such knowledge and it is in conformity with experience. Furthermore, it is the function of the active intellect to make the passive intellect as similar as possible to the representation of the phantasms. Unlike Aquinas, Suárez maintained that the passive intellect can abstract the universal and that the active intellect can know the individual material object.

DOCTRINE OF ANALOGY. Suárez rejected the Scotist doctrine of the univocity of being. Like Aquinas he accepts the analogicity of being, but he insists that there is only an analogy of attribution—not of proportionality—which possesses an element of metaphor. "Every creature is being in virtue of a relation to God, inasmuch as it participates in or in some way imitates the being of God."

EXISTENCE OF GOD. A metaphysical rather than a physical proof is needed to establish the existence of God. The major defect in the Aristotelian argument from motion is the principle that "everything which is moved

is moved by another." For this principle Suárez substituted the metaphysical principle that "everything which is produced is produced by another." From this principle he argued that there must be an unproduced or uncreated being, for an infinite regress either of a series or a circle of finite beings cannot be accounted for. And even if an infinite series were accepted, such a series would depend on a cause external to it. From the conclusion that there exists an uncreated being, Suárez proceeded to demonstrate that there is only one such being. Regarding the nature of such a being, its perfection, wisdom, infinitude, and so on, he followed Aquinas.

PROBLEM OF FREEDOM. Like Luis de Molina, Suárez was convinced that the Thomist doctrine that God physically predetermines the free act of the individual nullified man's freedom. Suárez maintained that through the *scientia media* God knows from all eternity what an individual will do if his grace is extended to him, and he consequently gives sufficient grace to effect the congruent action of the individual's will with his grace.

PHILOSOPHY OF LAW

Although Aquinas's influence on Suárez is apparent, Suárez was a highly original and influential thinker in the philosophy of law. He effected the transition from the medieval to the modern conception of natural law, and his influence is particularly noticeable in the work of Hugo Grotius.

NATURE OF LAW. Suárez maintained that Aquinas's definition of law as "an ordinance of reason directed to the common good" placed an inordinate emphasis on reason or intellect. Suárez did not deny that reason has a part in the law, but he did hold that obligation is the essence of law and that obligation is essentially an act of will. He defined law as "an act of a just and right will by which a superior wills to oblige his inferior to do this or that."

ETERNAL LAW. Like Aquinas, Suárez distinguished between eternal, divine, natural, and human law. However, the treatment of each is based on Suárez's contention that law is fundamentally an act of will. Eternal law is the divine providence that extends to all creatures and from which the other laws are derived. Defined as "a free decree of the will of God, who lays down the order to be observed," it is immutable and has always existed with God. It differs from the other laws, whose origins depend upon their promulgation; the eternal law receives its promulgation only through the other laws. Man's knowledge of such a law is limited and is reflected in his acceptance of the divine law, the discovery of the natural law, and his promulgation of the human law.

DIVINE LAW. Divine law is the direct revelation of God—the Mosaic law. The power and the will of God are the source of man's obligation to obey the divine law. In contrast, the power and the obligation of the human law are directly the will of the legislator, although indirectly the will of God.

NATURAL LAW. Natural law receives considerable attention from Suárez. This law is the participation of the moral nature of man in the eternal law. The natural law is based on the light of reason, but it is the work of the divine will and not the human will; its ultimate source is God, the supreme legislator. The natural law is not identified with man's nature; it transcends his will. The precepts of the natural law are the general and primary principles—to do good and avoid evil; the more definite and specific principles—that God must be worshiped; and certain moral precepts that may be deduced from the primary principles—that usury is unjust, adultery wrong, and so on. There is no dispensation from the natural law; its precepts are immutable. Thus, the introduction of private property did not reflect a change in the natural law, for although the natural law conferred all things upon men in common, it did not positively enjoin that only this form of ownership should endure.

HUMAN LAW. Human law must be based on either the divine law or the natural law and is best exemplified in political philosophy. Following Aristotle, Suárez held that man is a social animal. He rejected the view that political society is artificial, the result of a social contract or an enlightened egoism. The state is natural, and the legislative power is derived from the community and exists for the good of the community. The ultimate source of such power is God, who bestows it as a natural property upon the community. Such power is actualized only upon the formation of a political society. The form of government is essentially a matter of choice by the people. The modernity of Suárez is revealed in his rejection of the medieval ideal of the imperial power. He accepted the sovereignty of individual rulers and was skeptical of the feasibility of a world state. In discussing the rule of tyrants, he distinguished between a legitimate ruler who behaves tyrannically and a usurping tyrant. Revolt against the latter is self-defense; it is even legitimate to resort to tyrannicide provided that the injustice is extreme and the appeal to authority impossible. In the case of the legitimate ruler, the people have a right to rebel, for they bestowed the

power upon the ruler. Tyrannicide is rejected here, and the rules of a just war must be followed. Suárez maintained that war is not intrinsically evil; just and defensive wars are permissible, and considerable attention is given to the conditions for waging a just war. Suárez also rejected the extremist views of papal power over temporal rulers, but he argued for the spiritual supremacy and jurisdiction of the papacy. This implied that the papacy has an indirect power to direct secular rulers for spiritual ends.

See also Aristotelianism; Duns Scotus, John; Essence and Existence; Gilson, Étienne Henry; Leibniz, Gottfried Wilhelm; Molina, Luis de; Natural Law; Ockhamism; Peace, War, and Philosophy; Philosophy of Law, History of; Schopenhauer, Arthur; Scientia Media and Molinism; Thomas Aquinas, St.; Thomism; Universals, A Historical Survey; Wolff, Christian.

Bibliography

MODERN EDITIONS

Works mentioned in text and others are to be found in the *Opera Omnia*. 28 vols. (Paris, 1856–1878). The *Disputationes Metaphysicae* have appeared in a Latin and Spanish edition (Madrid, 1960–1964). For the philosophy of law see especially the *Selections from Three Works of Francisco Suárez, S.J.* (*De Legibus, Defensor Fidei*, and *De Triplici Virtute Theologica*). 2 vols. (Classics of International Law, No. 20, Oxford: Clarendon Press, 1944). There is an introduction by J. B. Scott.

BIBLIOGRAPHICAL WORKS

McCormick, J. J. *A Suarezian Bibliography*. Chicago, 1937.

Mugica, P. *Bibliografica suareciana*. Granada, 1948.

Scorraille, R. de. *François Suarez de la compagnie de Jésus, d'après ses lettres, ses autres écrits inédits et un grand nombre de documents nouveaux*. 2 vols. Paris: Lethielleux, 1912–1913. The most authoritative bibliographical study.

METAPHYSICAL AND EPISTEMOLOGICAL STUDIES

Alejandro, J. M. *La gnoseologica del Doctor Eximio y la accusación nominalista*. Comillas, Spain, 1948.

Breuer, A. *Der Gottesbeweiss bei Thomas und Suarez*. Fribourg, Switzerland, 1930.

Conza, E. *Der Begriff der Metaphysik bei Franz Suarez*. Leipzig, Germany, 1928.

Descoqs, P. "Thomisme et Suarezisme." *Archives de philosophie* 4 (1926): 434–544.

Dumont, P. *Liberté humaine et concours divin d'après Suarez*. Paris, 1960.

Grabmann, M. "Die Disputationes Metaphysicae des Franz Suarez." In *Mittelalterliches Geistesleben*. 3 vols., Vol. I, 525–560. Munich: Hueber, 1926–1956.

Hellin, J. *La analogía del ser y el conocimiento de Dios en Suárez*. Madrid, 1947.

Mahieu, L. *François Suarez, sa philosophie et les rapports qu'elle a avec la théologie*. Paris, 1921.

Mullaney, T. *Suárez on Human Freedom*. Baltimore: Carroll Press, 1950.

PHILOSOPHY OF LAW STUDIES

Perena Vicente, L. *Teoría de la guerra en Francisco Suárez*. 2 vols. Madrid: Consejo Superior de Investigaciones Cientificas, Instituto Francisco de Vitoria, 1954.

Recaséns Siches, L. *La filosofía del derecho en Francisco Suárez*. Madrid, 1927.

Rommen, Heinrich. *Die Staatslehre des Franz Suarez*. Munich and Gladbach: Volksvereins-Verlag, 1926.

COLLECTED ESSAYS

Estudios eclesiásticos 22 (85–86) (1948).

Gemelli, A., ed. *Scritti vari*. Milan: Vita e pensiero, 1918.

Pensamiento, special issue, 4 (1948).

John A. Mourant (1967)

SUBCONSCIOUS

See *Unconscious*

SUBJECT AND PREDICATE

The contrast between "subject and predicate" is a significant one in at least four different realms of discourse: grammar, epistemology, logic, and metaphysics. A large number of philosophical problems have to do with how the distinction on one level is related to that on some other level; whether there really are four such distinct realms and, if so, how they bear on one another are matters of controversy.

GRAMMAR

In the realm of grammar, subject and predicate are sentence parts; they are, therefore, words or groups of words, and their definition and identification is a matter of syntax. In the simplest case, where the sentence consists of just two words, such as

(1) Bats fly,

(2) Fraser swims,

the subject is the noun and the predicate is the verb. Very few sentences are so simple, but an indicative sentence with just one noun and one verb remains a good paradigm for the grammatical categories of subject and predicate because we can see in it the form of the sentence stripped down to its essentials: If either of the two words were omitted, we would no longer have an indicative sentence. Furthermore, very many sentences of English, as

well as of other familiar European languages, break neatly and obviously into two parts corresponding to the noun and the verb in the paradigm, and modern linguistic analysis of sentence syntax generally begins by viewing a sentence as a noun phrase plus a verb phrase:

$$S \rightarrow NP + VP$$

Although subject-predicate sentences are very common in English and in other languages, this form of sentence is not the only one, other forms being exemplified in English by normal idiomatic expressions for commands, requests, salutations, and so on. These other forms of sentence, however, have traditionally been assimilated to the subject-predicate form through the assumption of an "unexpressed subject" or some other missing element. It once seemed reasonable to try to save appearances in this way because subject and predicate seemed to be universal grammatical categories, found not only in the European languages but also, for example, in Sanskrit. Recent familiarity with a wider variety of languages has shown that these categories are by no means universal, and it is doubtful whether any grammatical categories or linguistic forms are universal. Some linguists have proposed that topic and comment are found universally, although subject and predicate are not. These categories, however, do not have to do just with the arrangement of words in sentences but rather with knowing what is being discussed and understanding what is said about it; hence topic and comment are not purely grammatical categories. The present situation in linguistics may therefore be summed up by saying that subject and predicate are useful grammatical concepts but do not represent universal grammatical categories.

In philosophy the grammatical distinction between subject and predicate has been prominent at least since Plato, who, in the *Sophist*, distinguished nouns and verbs as two classes of names. It is fair to say, however, that in that discussion, as well as in subsequent ones, philosophers have been interested in this grammatical distinction primarily because of the use they might make of it in treating problems of epistemology, logic, and metaphysics.

EPISTEMOLOGY

In epistemology the contrast between subject and predicate is a contrast between that part of a sentence which serves to identify or designate what is being discussed and that part which serves to describe or characterize the thing so identified. The categories of subject and predicate have more claim to universality at the level of epistemology (semantics) than at the level of grammar (syntax). It is here that the hypothesis about topic and comment, mentioned earlier, has its significance, for the fact that every language has some grammatical device or other for identifying a subject, or topic, and predicating something of it, or commenting on it, largely accounts for our remarkable ability to translate the content of any message from one language into another.

The epistemological sense of subject and predicate has much in common with the grammatical sense: Sentences (1) and (2) can be taken as paradigms for both senses, and the grammatical subject very frequently identifies the subject of discourse. Nevertheless, the two senses are not identical. They diverge, for example, in sentences with a dummy subject. In "It is raining" the expletive "it" is the grammatical subject of the sentence, but since it does not designate anything at all, it does not designate or identify what the sentence is about. Other instances are more relevant to philosophical issues and may be controversial. Consider

(3) What is not pink is not a flamingo.

(4) What is not just is not to be done.

There is no difficulty with (4), for it says something about unjust acts, and hence its grammatical and epistemological subjects coincide. But (3) seems to be about flamingos rather than about nonpink things, even though it has the same grammatical form as (4). Perhaps this is because we directly recognize and classify things as flamingos and as unjust acts, and even as pink, whereas in order to call something "not pink" one would normally first recognize it as gray or blue or some other color. If this is correct, the epistemological subject of (3) is mentioned in the grammatical predicate rather than in the grammatical subject.

Another instance of the divergence of the epistemological and grammatical senses is in relational sentences, such as

(5) Andrew was hit by Bernard.

(6) The cat is between the bird and the snake.

Sentences (5) and (6) may be taken to be about the two persons and the three animals, respectively, and what is said about their epistemological subjects is that a certain relation is true of them. Treating (5) and (6) as having multiple subjects in this manner is much more congenial than is a grammatical analysis to what Bertrand Russell, among others, said about the importance of relations.

It should be noted that what counts as the epistemological subject of a statement may be determined in part by the context in which it is made: If Bernard is the "topic" of conversation, (5) would naturally be construed as a "comment" about him, but other conversations in which (5) occurs will be focused differently. The importance of context in determining what counts as a subject differentiates the epistemological conception of subject from all the others.

Predicates as well as subjects have required special treatment in epistemology. Immanuel Kant distinguished real predicates from grammatical or logical predicates, a real predicate being one that says something about the subject—that is, one which attributes some property to the subject. Kant's contention that "exists" is not a real predicate but only a grammatical or logical one provides the basis for his refutation of the Ontological Argument. Statements of identity have also been held by Gottlob Frege, Russell, Ludwig Wittgenstein, and others not to be genuine predications—or at least not to be straightforward ones. Hence, in

(7) Tully is Cicero

the words "is Cicero" would not express an epistemological predicate, although they assuredly constitute the grammatical predicate. These are matters that are still not so clear as they might be.

Some very important topics in semantics and the philosophy of language are connected with the epistemological contrast between subject and predicate. In order to know what a person is talking about, I must know to what (or to whom) certain words in his utterances refer; the problem of how words can have such reference is an important one. In order to understand what is said about the subject under consideration, I must further know what is signified or entailed or meant by certain other words the person uses, whence arises another important problem, how words come to have sense or connotation. The distinction between two such modes of meaning, characteristic respectively of subjects and of predicates, has a long history and is still a live issue. Plato, in the *Theaetetus* and the *Sophist*, distinguished the mode of meaning of nouns from that of verbs. More recently J. S. Mill's distinction between connotation and denotation and Frege's distinction between sense and reference have taken up the same theme and made it central to the philosophy of language.

LOGIC

In formal logic there has been a distinction between subject and predicate ever since Aristotle's pioneering work in the field, but a dispute about the nature and scope of the distinction separates traditional from modern logicians. Aristotle would regard sentences (1) and (2) as both having subject-predicate form, but only (1) could serve as a paradigm for his formal logic. In traditional formal logic what is important about the subject term in the paradigm is, roughly, that it comes at the beginning of the sentence and indicates what (or who) is being discussed and that its quantity can be expressed by "some" or "all" preceding the noun. The pattern involved is

$$S \text{ is } P,$$

and since every proposition must have a topic about which something is asserted, this pattern is held to be manifested universally in categorical propositions. In modern logic, on the other hand, what is important about the subject term is that it is a proper name and stands for an individual, and so only sentence (2) can serve as a paradigm of the subject-predicate form. The pattern involved is

$$Fa$$

(where "*F*" stands for some attribute and "*a*" is a proper name); this pattern never applies to general propositions, since fully general propositions contain quantifiers, variables, and predicate terms but no proper names. According to this view general propositions pertain just to predicates and are not subject-predicate propositions at all. Russell's famous attack on "subject-predicate logic" was an attack on the view that every proposition must have a logical subject.

From a formal point of view the issue can be seen as a dispute about whether the principle of transposition (or contraposition) applies to subject-predicate propositions. In traditional logic it does, for the complement of a predicate can serve as a subject. This is not the case in modern logic, however, where only singular terms count as subjects and where transposition applies only to complex propositions compounded with the "if-then" sentence connective. There is a related divergence in the treatment of existence. Kant, a typical traditional logician in this respect, called existence a "logical" predicate, although not a "real" one; in effect, the grammatical analysis of assertions of existence into subject and predicate is carried over into logic. In modern logic, on the other hand, existence is generally represented through quantification, rather than through a predicate.

Epistemological and metaphysical considerations are involved in this dispute about how to represent subjects and predicates in formal logic. Roughly speaking, traditional logic seems to favor some sort of realistic view of universals, since terms representing universals can serve as both logical subjects and logical predicates. In the notation of modern logic, on the other hand, only singular expressions can serve as logical subjects, and this rule seems to give prominence to individuals rather than to universals. But a variety of epistemological and metaphysical views can consistently be advanced by both traditional and modern logicians, and the ascendancy of modern logic can be attributed to its greater flexibility, adaptability, and power as a calculus, rather than to epistemological and metaphysical views associated with it. It seems prudent, therefore, to keep matters of perspicuous symbolism and logical transformation separate from other considerations.

To illustrate the problems about the relation of logical structure to epistemological structure, one might consider

(8) All ravens are black.

The epistemological subject of (8) is ravens, and hence one would go about confirming the proposition by examining ravens and finding them black. If, using the rule of transposition, we derive from (8) the logically equivalent form

(9) All nonblack things are nonravens,

one is tempted to assume that the epistemological subject and predicate of (8) have been similarly transposed, so that nonblack things is the epistemological subject of (9). This assumption gives rise to the so-called paradox of confirmation, for it then appears as though we might confirm (8) and (9) by examining nonblack things and finding them not to be ravens, contrary to our normal procedure for confirming such simple generalizations. One solution is to hold that transposition does not apply to the epistemological structure of a proposition, that the epistemological structure of a proposition is therefore not always parallel to its logical structure, and that the epistemological subject of (9) is the same as that of (8)—that is, ravens. But the desire to have epistemological structure unambiguously represented in logical notation is a powerful consideration for some philosophers, and hence the matter is still controversial.

METAPHYSICS

The distinctions between subject and predicate in grammar, epistemology, and logic have given rise to a variety of metaphysical doctrines. These doctrines deserve separate consideration because although they are closely related to the distinctions already sketched and are suggested by them, none follows from them.

Plato noted that applying different predicates to a subject often entails a change in the subject, whereas applying a predicate to different subjects does not entail a change in the predicate. He took this changelessness to be a mark of reality (as well as epistemological priority), and hence his theory of Forms gives great ontological prominence to predicates (concepts, universals—i.e., that which a grammatical predicate stands for). This bold thesis opened a long and continuing dispute about the nature of universals, the problem being to determine what ontological commitments, if any, are entailed by our use of predicative expressions (in the epistemological sense).

Aristotle, in contrast to Plato, gave ontological standing to subjects as well as to predicates. Discussing substance in his *Categories*, he defined "first substances" as things satisfying two conditions: (*a*) being subjects but never predicates and (*b*) not being in or of something else (as a color or surface must be the color or surface of some other thing). He then defined "second substances" as things satisfying the second condition but not the first. First substances are individuals. Second substances are species or universals and hence incorporate an element of Plato's metaphysics (although not all universals are substances). An attractive feature of Aristotle's metaphysical treatment of subjects is that it fits his conception of subjects in epistemology and logic: What we talk about and investigate (especially in biology, Aristotle's scientific forte) are individuals and species, and his logic allows both individual names and universal terms, including species names, to occur as logical subjects. But, in spite of its merits, Aristotle's metaphysical conception of subjects is often regarded as unsatisfactory, largely because of qualms about putting individuals and species in one basket, about distinguishing predicates that stand for substances from those that do not, and about the usefulness of traditional logic.

Gottfried Wilhelm Leibniz's doctrine of monads builds on Aristotle's conception of individual substance. But Leibniz considered Aristotle's definition inadequate, and he defined a monad or individual substance as a subject that contains all its predicates—that is, as an individual from whose "notion" it is possible to deduce all that may ever be truly predicated of it. Few philosophers have

thought there were any such substances. One difficulty may be that Leibniz attributed to his monads, which are epistemological subjects, the sort of identity that characteristically belongs to a predicate—namely, a definite set of entailments that define it.

Whereas Leibniz had only one kind of substance, G. W. F. Hegel allowed only one individual substance, the Absolute. The Absolute is the ultimate subject of every statement and resembles Leibniz's monads in that it contains all its predicates in the same sense as the monads are supposed to. Other philosophers have not been convinced of the existence of such a universal subject; Russell, who acknowledges Hegelian idealism to be a plausible account of the metaphysical implications of traditional logic, regards the doctrine as a reductio ad absurdum argument against a logic that analyzes every proposition as having a subject and a predicate.

Another interesting element of idealism is the concept of the concrete universal. Like the idea of a monad, this concept is an attempt to overcome the subject-predicate dualism by amalgamating features of both subjects and predicates in a single sort of entity. Whereas a monad is a subject with characteristics of a predicate (in that its identity is determined by what is logically contained in it, or entailed by it), a concrete universal is a predicate treated as a concrete individual thing.

One philosopher who accepted the subject-predicate dualism as a basis for his metaphysics was Frege. There are, he maintained, two radically different sorts of things, objects and concepts. Objects are complete, or "saturated," and stand on their own, so to speak; we have names for them and talk about them, but the name of an object can never be a grammatical or logical predicate. Concepts, or, more generally, what Frege called "functions," are incomplete, or "unsaturated"; they require an object to complete them and hence cannot stand alone, and a concept term is always a predicate, never a subject. Frege's dualistic view has been very influential with other philosophical logicians, including Russell, Wittgenstein, Rudolf Carnap, and P. T. Geach, but difficulties in Frege's formulation of it have impeded its general acceptance.

One difficulty is that even Frege wished to talk about concepts, and hence he had to suppose that each concept has a special object associated with it that serves only as an object to talk about when we mean to discuss the concept. A more serious difficulty is that the object-concept dualism does not fit with Frege's semantic distinction between sense and reference, which also arises from a consideration of subjects and predicates. One might expect that reference would be the mode of meaning characteristic of names of objects, and sense the mode of meaning characteristic of concept terms; however, both names and concept terms have both sense and reference. Frege had powerful reasons for what he said, but the final impression is that his two distinctions are distressingly unrelated; hence, the philosophers most influenced by him have differed from him. Russell, for example, vigorously rejected Frege's distinction between sense and reference (in his essay "On Denoting"), and Wittgenstein in his *Tractatus Logico-Philosophicus*, although indebted to Frege when he characterized his metaphysical objects, left no room for any other entities corresponding to Fregean functions.

Many analytic philosophers (which included Carnap, Ernest Nagel, and Max Black) hold that neither grammatical nor logical categories have metaphysical implications. P. F. Strawson, however, revived the issue among them by considering the implications and presuppositions of grammatical, logical, and epistemological subjects in his metaphysical essay *Individuals*. On balance, metaphysical skepticism must probably be considered as controversial as any of the metaphysical doctrines proposed.

See also Existence; Logic, History of; Meaning; Proper Names and Descriptions; Relations, Internal and External; Substance and Attribute; Universals, A Historical Survey.

Bibliography

For the linguistic aspects consult a standard work, such as Leonard Bloomfield, *Language* (New York: Holt, 1933), or R. A. Hall Jr., *Introductory Linguistics* (Philadelphia: Chilton, 1964). Russell's strictures against subject-predicate logic can be found in his discussions of Aristotle and Hegel in his *History of Western Philosophy* (New York: Simon and Schuster, 1945), and his *An Inquiry into Meaning and Truth* (New York: Norton, 1940) contains discussions of both logical and epistemological aspects of the question, as well as comments on Wittgenstein. Russell's paper "On Denoting" can be found in *Logic and Knowledge* (New York: Macmillan, 1956), together with other relevant papers. Frege's views are readily accessible in *Translations from the Philosophical Writings of Gottlob Frege*, edited by P. T. Geach and Max Black (Oxford: Blackwell, 1952). Wittgenstein's *Tractatus Logico-Philosophicus* has had a more profound influence with respect to semantics and logical form than any other twentieth-century work. Mill's discussion of connotation and denotation can be found in Book I of his *System of Logic*.

The chief items by Aristotle are *On Interpretation, Categories*, the opening paragraph of *Prior Analytics*, and Book Zeta of the *Metaphysics*; by Hegel, *The Phenomenology of the Mind*, especially the preface; by Leibniz, *Monadology*; by Carnap, P. A. Schilpp, ed., *The Philosophy of Rudolf Carnap* (La Salle,

IL: Open Court, 1963), the autobiographical essay and the replies; by Black, *Models and Metaphors* (Ithaca, NY: Cornell University Press, 1962); by Geach, *Reference and Generality* (Ithaca, NY: Cornell University Press, 1962).

G. E. M. Anscombe and P. T. Geach, *Three Philosophers* (Ithaca, NY: Cornell University Press, 1961), contains useful comments on the relevant views of Aristotle and Frege. One of the best more recent discussions of the influence of grammar on metaphysics is Morris Lazerowitz, "Substratum," in *Philosophical Analysis*, edited by Max Black (Ithaca, NY: Cornell University Press, 1950). For further references, see the articles cited in text and entries on the philosophers mentioned.

Newton Garver (1967)

SUBJECTIVISM IN ETHICS

See *Ethical Subjectivism*

SUBJECTIVIST EPISTEMOLOGY

A "subjectivist epistemology" is one that implies the standards of rational belief are those of the individual believer or those of the believer's community. Thus, subjectivism can come in either an individualistic form or a social form. A key negative test of subjectivism is whether an account implies that by being rational one is assured of having beliefs that are more reliable than they would be otherwise—that is, more reliable than they would be if one were not rational. Thus, reliabilist accounts of rational beliefs are paradigmatically objective. So are traditional foundationalist accounts. By contrast, if an account implies that the standards one must meet if one's beliefs are to be rational are those that one would regard as intellectually defensible were one to be ideally reflective (Foley 1987, 1993), then the account is subjective. Similarly, an account is subjective if it implies that one's beliefs are rational if they meet the standards of one's community (Rorty 1979) or the standards of the recognized experts in one's community (Stich 1985). Likewise, an account is subjective if it implies that one's beliefs are rational if they meet the standards of the human community at large, provided nothing else in the account implies that adhering to such standards will reliably produce true beliefs.

One of the considerations favoring a subjectivist epistemology is that it provides an attractive way of describing what is going on in skeptical scenarios—for example, one in which everything appears normal from my subjective point of view even though my brain has been removed from my body and placed in a vat, where it is being fed sensory experiences by a deceiving scientist. In such a scenario, almost everything I believe about my immediate surroundings would be false. Hence, I would have little knowledge about these surroundings, but what I believe about them might nonetheless be rational. Indeed, my beliefs would be as rational as my current beliefs about my surroundings. The most plausible explanation as to why this is so is that there is at least one important sense of rational belief according to which having rational beliefs is essentially a matter of meeting subjectively generated standards. Thus, by being envatted I may be deprived of the opportunity of having knowledge about my surroundings, but I am not necessarily also deprived of an opportunity of having rational beliefs.

See also Classical Foundationalism; Epistemology; Reliabilism; Social Epistemology.

Bibliography

Foley, R. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.

Foley, Richard. *Working without a Net*. New York: Oxford University Press, 1993.

Rorty, R. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press, 1979.

Stich, S. "Could Man Be an Irrational Animal?" *Synthese* 64 (1985): 115–135.

Richard Foley (1996)

SUBJECTIVITY

Subjectivity is, primarily, an aspect of consciousness. In a sense, conscious experience may be described as the way the world appears from a particular mental subject's point of view. The idea that there is a distinction between appearance and reality seems to presuppose the distinction between subjective and objective points of view.

THE TWO CONTROVERSIES

There are two principal controversies surrounding subjectivity: first, whether subjectivity, as it is manifested in consciousness, is an essential component of mentality; and second, whether subjectivity presents an obstacle to naturalistic theories of the mind.

THE FIRST CONTROVERSY. Most philosophers agree that intentionality—the ability to represent—is characteristic of mentality. However, there is strong disagree-

ment over whether subjectivity is also necessary. Those philosophers who think it is (e.g., Searle 1992) argue that true—or what they call "original"—intentionality can only be attributed to a conscious subject. In this view, representational properties can only be ascribed to unconscious states and to unconscious machines—such as computers and robots—in a derivative sense. With respect to computers, the claim is that their internal states only have meaning to the extent that people (conscious subjects) interpret them to mean something. On their own, these states are merely meaningless formal symbols. When it comes to unconscious states—such as the unconscious beliefs and desires hypothesized in Freudian psychology—the claim is that only by virtue of their effects on one's conscious beliefs and intentions do they have content. The source of all genuine meaning resides in conscious, subjective mental activity.

The basic argument for this position is that for something to count as a representation—as meaning something—there must be a subject for whom its meaning is significant; a subject who is aware of and appreciates what it means. Otherwise, the argument goes, without a subject who understands, interprets, and makes use of the meaning, there is no basis for saying it means anything at all. In particular, given that most conditions stateable in objective terms for what a brain state or a computer state represents leave room for alternative interpretations, it is only by reference to the awareness of a conscious subject that a representation acquires determinate content.

Other philosophers reject this assimilation of intentionality and subjectivity, arguing that a theory of intentionality—one that applies equally to conscious and unconscious states—can be developed independently of a theory of subjectivity (e.g., Dretske 1981 and Fodor 1987). Some theorists see no need at all to appeal to the interpretive activity of a conscious subject to fix the content of a representational state. In this view, meaning ultimately comes down to information, a notion that may be treated in objective terms.

Others agree that some appeal to the purposes of the agent is necessary in order to ground an assignment of meaning to brain states. However, they claim that it is not necessary to invoke the subjective character of a subject's conscious states for this purpose. Rather, it suffices to show that by assigning the relevant interpretation to the subject's internal states one may provide appropriate psychological explanations of the subject's behavior and explain that subject's success in his or her interactions with the environment. What a subject's beliefs and desires are about is determined, in this view, by the nature of the subject's interactions with the environment and the role these states play in her or his internal psychological economy. These are facts clearly stateable from an objective point of view; no special appeal to the subjective experience of the agent is required.

Just how serious the first controversy is depends considerably on one's stand with respect to the second one. Suppose that one adopts the position that only creatures possessing subjective, conscious states are capable of any mentality at all. Still, if one also thinks that possession of subjective consciousness is a perfectly natural phenomenon—itself explicable in physical, or objective terms—the sting is largely removed from this position. There is now no reason to think properly programmed computers or robots couldn't possess the full range of mental states, so long as they satisfied the naturalistic conditions for conscious subjectivity.

THE SECOND CONTROVERSY. With respect to this second question—whether or not subjectivity presents a problem for a naturalistic framework—one may reason as follows. A complete inventory of the world should, if it is truly complete, capture everything there is and everything going on. It seems natural to suppose that such a complete description is in principle possible, and is in fact the ideal aim of natural science. But some argue that facts that are essentially accessible only from a particular subject's point of view cannot be included in this allegedly complete objective description (Nagel 1974, 1986). If they cannot, this would seem to undermine the idea that the natural world constitutes a coherent, lawful, and objective whole.

For example, take the fact of one's own existence. You could read through this hypothetical exhaustive description of the world, and it would include a description of a body at a particular spatio-temporal location, with particular physiological (or even nonphysical) processes going on inside it. However, what would be missing is that this is your body—this is you. No collection of facts stateable in objective terms seem to add up to this body being yours.

Or take the problem of personal identity. From a point of view outside the subject, what it is that makes one the same person across time—whether it be a matter of bodily or psychological continuity—seems to admit of borderline cases or matters of degree, or other sorts of indeterminacy. Thought experiments involving split brains, machines that take "memories" from one brain and implant them in another, and the like, reveal just how

difficult it is to pin down personal identity as a determinate matter of objective fact. Yet, from the point of view of the subject, what it is to be oneself seems to be a clear-cut, all-or-nothing matter. Either one continues to exist or one doesn't. It is hard to reconcile the objective and subjective perspectives on this question.

One particularly difficult manifestation of the problem of subjectivity is how to account for the fact that there is "something it is like" to be certain objects (say a human being), or occupy certain states (say, visual experiences), but not others (say, a rock, and its states). This is also known as the problem of "qualia." From an objective point of view, there would seem to be nothing special about the neurological activity responsible for conscious experience that would explain what it's like for the subject. Two influential thought experiments starkly illustrate the problem.

Nagel (1974) presents the problem this way. Bats navigate in the dark by emitting high-pitched sounds and detecting their echos—a sensory system known as "echolocation." From an objective, third-person point of view, there is nothing especially difficult about understanding how this system works. While there are of course difficult technical questions, the idea that the bat extracts information concerning the location and movement of its target from the returning sound waves bouncing off of it is fairly clear. The problem emerges when one considers what echolocation is like for the bat, from its point of view. People know that there is something particular it is like to see a sunset, smell a rose, or feel a pain. There is every reason to believe that there is also something particular it is like to sense by echolocation. Yet, when the question is posed this way, it doesn't seem as if any of the details learned about the information-processing capabilities of the bat are helpful in answering this simple question: What is it like for the bat? It seems as if only by adopting the bat's point of view, by humans' experiencing echolocation, could one obtain a clue concerning what it is like.

Jackson (1982) asks people to consider the following situation. Imagine Mary, a neuroscientist who learns everything there is to know about the physiology and information processing involved in color vision. However, she learns this while restricted to a completely black and white environment, so that she herself never experiences color sensations. In a sense, she would be in the same position vis-à-vis everyone else that everyone else is vis-à-vis bats. At some point Mary is released from her purely black and white environment and allowed to see color. Suppose she now sees a red rose for the first time. It seems undeniable that her reaction would be one of wonder and novelty. "So that's what red looks like!" she might say. But now, if the subjective experience were adequately captured by the objective, third-person descriptions presented in her science texts, why should she experience novelty and wonder? That she would have this experience seems to demonstrate that what is apprehended from the first-person, subjective point of view is distinct from what is describable in objective, third-person terms.

Many philosophers argue that subjectivity does not present a special puzzle. For some (e.g., Searle 1992), it is just a fact that the world contains both objective facts and irreducibly subjective facts; their relation requires no explanation and produces no mystery. For most, though, the demystification of the subjective is accomplished by some sort of reductionist strategy (e.g., Lycan 1987 and 1990, and Rosenthal 1986), one that shows how to incorporate so-called subjective facts into an all-embracing, naturalistic and objective scientific framework. One influential model of subjectivity is the internal monitoring, or higher-order thought model. In this view, which fits well with a functionalist approach to the mind-body problem in general, subjectivity is principally a matter of some mental states representing other mental states. That is, to be aware of, or to apprehend from the first-person point of view, that one is having a certain experience, is merely to occupy a mental state that represents one as having that experience. If this is what subjectivity amounts to, then any model of the mind that builds in the requisite architectural features will explain subjectivity. A model of this sort of internal scanning already exists with computers.

Advocates for the view that subjectivity presents no special mystery sometimes point to the perspectival character of indexical expressions such as "I" and "here" for support. The idea is that it is generally acknowledged that the meaning of such expressions cannot be captured in nonindexical terms (Perry 1979), yet this doesn't give rise to any special philosophical problem or mystery. Because one cannot derive a statement containing an indexical expression from statements free of indexicals, one need not conclude that there are any special indexical facts that are indescribable in indexical-free terms. There are theories that take into account the special behavior of such terms consistent with a general theory that applies to nonindexical terms as well.

In the same way, goes the argument, subjective mental phenomena can be incorporated into a more general theory of the world that applies to nonsubjective phe-

nomena as well. For instance, whereas it may be true that Mary, in the Jackson example described above, could not predict what it would be like to see red from her knowledge of the neurophysiology of color vision, this need not be taken to show that there are irreducibly subjective facts. It could be that human beings possess a distinct representational system that is employed only when information comes directly from the sensory systems (Rey 1993). It is no surprise that the same fact can be represented in distinct ways, and that being represented in distinct ways may obscure its identity from the subject.

ELIMINATING SUBJECTIVITY

Yet another approach to the problem of subjectivity is eliminativism (e.g., Churchland 1985, Dennett 1991). Proponents of this view will grant that none of the models proposed to account for subjectivity really explains it; but, they argue, that is due to the human intuitive conception of subjectivity—indeed of consciousness in general—being too confused, or incoherent, to be susceptible to scientific explanation. Subjectivity just isn't a real phenomenon, so there's nothing in the end to explain.

See also Knowledge Argument; Qualia; Self.

Bibliography

Churchland, P. "Reduction, Qualia, and the Direct Introspection of Brain States." *Journal of Philosophy* 82 (1985).

Dennett, D. C. *Consciousness Explained*. Boston: Little, Brown, 1991.

Dretske, F. *Knowledge and the Flow of Information*. Cambridge, MA: Bradford Books/The MIT Press, 1981.

Fodor, J.A. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: Bradford Books/The MIT Press, 1987.

Jackson, F. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Levin, J. "Could Love Be Like a Heatwave?: Physicalism and the Subjective Character of Experience." *Philosophical Studies* 49 (1986).

Lycan, W. G. *Consciousness*. Cambridge, MA: Bradford Books/The MIT Press, 1987.

Lycan, W. G. "What Is the 'Subjectivity' of the Mental?" *Philosophical Perspectives.* Vol. 4, edited by J. Tomberlin. Atascadero, CA: Ridgeview Publishing, 1990.

McGinn, C. *The Problem of Consciousness*. Oxford: Basil Blackwell, 1991.

McGinn, C. *The Subjective View*. New York: Oxford University Press, 1983.

Nagel, T. *The View from Nowhere*. New York: Oxford University Press, 1986.

Nagel, T. "What Is It Like to Be a Bat?" *The Philosophical Review* Vol. 82 (1974), pp. 435–450

Perry, J. "The Problem of the Essential Indexical." *Nous* 13 (1979).

Rey, G. "Sensational Sentences." In *Consciousness: Psychological and Philosophical Essays*, edited by M. Davies and G. Humphreys, 240–257. Oxford: Blackwell, 1993.

Rosenthal, D. "Two Concepts of Consciousness." *Philosophical Studies* 49 (1986): 329–359.

Searle, J. *The Rediscovery of the Mind*. Cambridge, MA: Bradford Books/The MIT Press, 1992.

van Gulick, R. "Physicalism and the Subjectivity of the Mental." *Philosophical Topics* 16 (1985).

Joseph Levine (1996, 2005)

SUBLIME, THE

This title already raises the conundrum that "the sublime" has regularly, although in different ways, posed. The substantivized form of the adjective suggests something one could point to where sublimity resides. The sublime might even be misconstrued (as it was by Edmund Burke) as a property of certain objects. But the sublime refers to no thing; it is instead an effect produced by the limits of our capacities for perception and representation. As such the sublime has played a vital role in the history of aesthetic theory as well as in postmodernist debates about representation and the limits of knowledge.

The sublime was first theorized by the pseudonymous Longinus in *On the Sublime*, written in the first century CE. Longinus conceives sublimity as a quality of elevated prose of great rhetorical power. Not until the seventeenth century does the sublime become associated with natural phenomena, and then with the incomprehensible excesses of natural force. In *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and the Beautiful* (1757), Burke provided an empiricist account of kinds of objects and situations that induce sublime perceptual experiences. Where beauty is found for Burke in things, the perception of which seems to harmonize with human sensory capacities, the sublime object of perception challenges our senses or exceeds our perceptual grasp. Burke equivocated on the implications of his empiricism, however, by conceiving sublimity as a property of these perceptually challenging objects or scenes, rather than understanding sublimity as a kind of secondary quality to be located in the relationship between perceiver and perceived.

Immanuel Kant provided in his *Critique of Judgment* (1790) the essential formulation of the sublime that has organized most subsequent discussion. Beauty, sublimity, and aesthetic qualities generally are for Kant no proper-

ties of objects; they are the felt effects of judgments that reflective perceivers make on the form and content of their experience. A judging subject finds something beautiful when its appearance or design, without or before applying conceptual rules to it, invigorates her cognitive capacities generally, and inspires an imaginative appreciation of the object. Judgments of beauty, then, reflect a harmony between feeling and cognition that is absent in the judgment of the sublime. People feel sublimity, to the contrary, in cases where their efforts to comprehend something are stymied by vastness, complexity, or by the natural might of that which threatens to overwhelm them.

These varieties of sublimity reflect Kant's germinal distinction between the mathematical and the dynamical sublime. The subject encounters mathematical sublimity when attempting to comprehend perceptually an object too vast (the starry heavens) or too grand (the great pyramids, from the correct distance) to take in all at once. The mathematical sublime exceeds one's conceptual grasp by inducing in the subject perceptual riches too extensive to subsume satisfactorily under available categories. It points up the limits of human capacity to perceive comprehensively and to represent to humans conceptually what is perceived. The frustration of this nevertheless gives rise to aesthetic pleasure for Kant, because the humbling of certain human cognitive capacities reminds people of the superiority of reason's capacity to think the infinite. For this reason, the sublime has regularly invited a theological interpretation throughout the European tradition.

The judging subject feels dynamical sublimity when threatened by the extraordinary forces of violent nature. This strain of Kant's theory of sublimity inspired the subsequent generation of Romantic poets, not to mention the later Nietzschean appreciation of Dionysian artistic impulses. Throughout the nineteenth century, the sublime is associated with excesses of natural force, tormented outpourings of emotion, and the transgression of norms of representation. Hence in the twentieth century the effects of sublime experiences were embraced by the sequence of artistic avant-garde movements that sought to induce ecstatic or liminal aesthetic responses designed to challenge conventional artistic or cultural norms. What a culture already possesses the conceptual apparatus to represent adequately cannot be sublime; the goal of the avant-garde was to allude to something that defies available means of representation.

Not surprisingly, then, the sublime was of great interest to postmodern theorists of the late twentieth century. Developments in multiple fields (the crisis of representation in anthropology, attacks on the representational theory of the mind in philosophy) encouraged postmodernists to embrace sublimity as the irrational and humiliating counterpoint to modernist categorizing zeal and its bureaucratic rationality. To embrace sublimity and to induce its manifestation in judging subjects is, as Jean-François Lyotard put it in *The Postmodern Condition* (1984), "To present the fact that the unpresentable exists" (p. 78). Rather than regard that humbled subject as the last word on the sublime, however, future theorists of this perennial notion may see the sublime, that which challenges human perceptual and conceptual reach, as a regular inducement to strive to extend that reach, rather than a reason to cease the attempt.

See also Aesthetics, History of; Beauty; Burke, Edmund; Kant, Immanuel.

Bibliography

Burke, Edmund. *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and the Beautiful*, edited by Alan Phillips. Oxford: Oxford University Press, 1990.

Kant, Immanuel. *Critique of Judgment*. Translated by Werner S. Pluhar. Indianapolis: Hackett, 1987.

Longinus. *On Great Writing (On the Sublime)*. Translated by G. M. A. Grube. Indianapolis: Hackett, 1991.

Lyotard, Jean-François. *The Postmodern Condition*. Translated by Geoff Bennington and Brian Massumi. Minneapolis: University of Minnesota Press, 1984.

Kirk E. Pillow (2005)

SUBSTANCE AND ATTRIBUTE

The concepts of "substance and attribute" are the focus of a group of philosophical problems that have their origins in Greek philosophy and in particular the philosophy of Aristotle. The concepts are, of course, familiar to prephilosophical common sense. Yet although we are acquainted with the distinction between things and their properties and are able to identify the same things among the changing appearances they manifest in time, these commonsense notions give rise to a group of philosophical problems when we come to scrutinize them. Thus we may wonder what it is that remains the same when, for example, we say that the car has new tires and lights and does not run as smoothly as it used to, but is still the same car; or when we say that although we could hardly recog-

nize him, this man is the same one we went to school with thirty years ago.

It is interesting to note that the principal term for substance in the writings of Aristotle is *ousia,* a word that in earlier Greek writers means "property" in the legal sense of the word, that which is owned. (This sense is familiar in English in the old-fashioned expression "a man of substance.") The word *ousia* also occurs in philosophical writings before Aristotle as a synonym for the Greek word *physis,* a term that can mean either the origin of a thing, its natural constitution or structure, the stuff of which things are made, or a natural kind or species. The Latin word *substantia,* from which the English term is derived, is a literal translation of the Greek word *hypostasis* ("standing under"). This term acquired its philosophical connotations in later Greek and occurs principally in controversies among early Christian theologians about the real nature of Christ. A third philosophical term, *hypokeimenon* ("that which underlies something"), is used by both Plato and Aristotle to refer to that which presupposes something else.

There is, however, little of philosophical importance to be learned from the etymology of the terms in which problems are formulated and discussed. We shall first consider the questions to which the concepts of substance and attribute give rise in some of the philosophers for whom they have been important. We may then ask which of these questions remain as live philosophical issues at the present time and what answers can be given to these surviving questions.

ARISTOTLE

Aristotle's account of substance has been the most influential in the history of philosophy. His account is, however, obscure and probably inconsistent. The difficulties of elucidating and reconciling the various parts of his doctrine have been part of the cause of its influence—it has offered a continuing challenge to commentators and critics from Aristotle's time to the present. "Substance in the truest and primary and most definite sense of the word is that which is neither predicable of a subject nor present in a subject; for instance, the individual man or horse" (*Categories* 2A11). The explanation is obscure, but the examples cited leave no doubt of what Aristotle means here: Substance in the most basic sense of the word is the concrete individual thing. However, he goes on at once to mention a second sense of the word: "Those things are called substances within which, as species, the primary substances are included; also those which, as genera, include the species. For instance the individual man is included in the species 'man' and the genus to which the species belongs is 'animal'; these, therefore,—the species 'man' and the genus 'animal'—are termed secondary substances." These secondary substances are predicable of a subject. "For instance, 'man' is predicated of the individual man" (*Categories* 2A21–22), as when we say "Socrates is a man." Aristotle seems to have the idea here that essences or natures are substances, and the more qualities they comprise, the more substantial they really are; he explains, "Of secondary substances, the species is more truly substance than the genus, being more nearly related to primary substance" (*Categories* 2B7). For example, the species *Canis domesticus* shares more qualities in common with the individual dog Tray than does the genus *Canis.*

This notion of essences as substances is treated at length by Aristotle in the *Metaphysics* and seems to be his preferred sense of the term. The intimation that the more qualities something has, the more substantial it is, has the advantage of suggesting that being a substance is a matter of degree and not an all-or-nothing matter. This hint, which Aristotle does not develop, contains an important idea, as will be seen later. But the doctrine of secondary substances has little else to recommend it and involves a serious logical confusion between the relations of class membership and class inclusion, as well as the notorious difficulties of the doctrine of essences.

Aristotle's main purpose in the *Categories* is to contrast the independent way of existing proper to substances with the parasitic mode of being of qualities and relations. Substances can exist on their own; qualities and relations, only as the qualities of or relations between substances. The key to this distinction is given by the phrase "present in a subject." (The Greek word for "subject" here is *hypokeimenon,* literally "underlay.") Substances are never "present in a subject." This does not mean, as Aristotle explains, that a substance is never "present in" something else as a part of a whole. On the contrary, he cites heads and hands (*Categories* 8B15) as substances although they are parts of bodies. Rather, *x* is present in *y* when it is "incapable of existence apart from" *y.* This notion introduces a third sense of substance as that which is capable of independent existence. This sense is of considerable importance in later philosophy, but Aristotle does not develop it. He uses it chiefly to emphasize the distinction between substances on the one hand and their qualities and relations on the other. A quality—"red," "sweet," or "virtuous"—cannot exist apart from an *x* that has the quality. Relations such as "larger than" or

"to the left of" cannot occur in the absence of the *x* and *y* that they relate.

It is true, of course, as Aristotle's critics have pointed out, that it is no more possible for a substance to exist without qualities than for qualities to exist without a substance. However, it is possible to point to prima facie examples of qualities existing without substances—the blue of the sky, for instance, or a red afterimage floating in my visual field. Surely the sky is not a substance, nor is my visual field. However, one cannot point to any instances of substances existing without qualities. Even if it makes sense to suppose that such a thing could occur, it is clearly incapable of being identified. Aristotle does not consider these problems. What he seems to have meant, although he does not express himself clearly, is that what is capable of independent existence is the concrete individual thing, a substance with its qualities and in its network of relations to other substances. But even here there is an obvious difficulty. Once we introduce the notion of relations involving other substances, we put a restriction on independent existence.

A fourth criterion of substance is that "while remaining numerically one and the same, it is capable of admitting contrary qualities" (*Categories* 4A10). This Aristotle calls "the most distinctive mark of substance." This notion is developed, more by later philosophers than by Aristotle himself, into the conception of a center of change and so of a substratum that underlies and supports its qualities. Finally, Aristotle emphasizes the notion of substance as a logical subject, "that which is not asserted of a subject, but of which everything else is asserted" (*Metaphysics* 1029A8), and he links this sense of the term with the concept of substratum. This logical criterion has been criticized as making the notion of substance dependent on the structure of Greek (and some other Indo-European languages), in which subject-predicate sentences are a standard mode of expression, and upon a restricted and now outmoded view of logic in which all statements canonically expressed are in a form in which a predicate is affirmed of a subject. It is not the case that sentences in all languages fall into a subject-predicate form or that this form of expression is adequate for a developed logic.

OTHER PHILOSOPHERS

The various notions of substance as (1) the concrete individual, (2) a core of essential properties, (3) what is capable of independent existence, (4) a center of change, (5) a substratum, and (6) a logical subject are never thoroughly worked out and reconciled in Aristotle. He appears to emphasize now one and now another mark of substance as of paramount importance. The quotations cited above have been chiefly from the *Categories*; the topic is taken up and discussed at length in the *Metaphysics.* The discussion is tentative and not finally conclusive, but Aristotle seems to favor alternative (2), substance as essence, as his preferred sense. But the whole treatment is important not for the answers that he gives but for the questions that he raises. Discussions of substance in later philosophers have tended, with few exceptions, to take over one or more of the six senses proposed by Aristotle as the clue to the problem.

ATOMISTS AND MEDIEVALS. Of the philosophical theories of antiquity, one other is of some consequence. Ancient atomism, founded by Leucippus and Democritus, developed by Epicurus, and expressed in its most attractive form in the *De Rerum Natura* of the Roman poet Lucretius, suggests that the truly real and substantial elements of nature are the atoms out of which everything is composed. It is these that are fundamental, unchangeable, and, in the last resort, capable of independent existence. The problem of substance and attribute was not much discussed by the ancient atomists, but their theories provide material for an answer to the question raised by Aristotle.

During the Middle Ages, discussion of this problem was very naturally centered upon the theological repercussions of rival theories. In particular, the doctrines of the Incarnation of Christ and of transubstantiation depended for their rational justification upon a plausible theory of substance. But these theological outworks produced no new basic insights that can be regarded as an improvement on the work of Aristotle. Indeed, they are just variations upon Aristotelian themes.

DESCARTES. The revival of philosophy in the seventeenth century in a form that was relatively independent of the religious framework of medieval philosophy produced several systems for which the notion of substance is fundamental. In the work of René Descartes the concepts of substance and attribute become associated naturally with those of the conscious self and its states, and the problem of substance becomes associated with the problem of personal identity. Descartes had been thoroughly trained in the form of Aristotelian scholasticism current in his day, and his notions of substance are in part derived from this and in part inconsistent with it. He gives a formal definition of substance as follows: "Everything in which there resides immediately, as in a subject, or by means of which there exists anything that we per-

ceive, i.e. any property, quality, or attribute of which we have a real idea is called a *Substance*; neither do we have any other idea of substance itself, precisely taken, than that it is a thing in which this something that we perceive or which is present objectively in some of our ideas, exists formally or eminently. For by means of our natural light we know that a real attribute cannot be an attribute of nothing" (*Philosophical Works,* translated by Haldane and Ross, 2nd ed., Cambridge, U.K., 1931, Vol. II, p. 53). In other words, what we are directly aware of are attributes of things and not the things themselves. But it is a logically self-evident principle (known by "the natural light" of reason) that an attribute must be an attribute of something, and the something is a substance—known by this inference and not directly. So far Descartes does not depart from scholastic doctrine, but he goes on to affirm that substances have essential attributes. For example, thought is the essential attribute of mind, and extension is the essential attribute of matter. But he does not explain what a substance is apart from its essential property. What is the mind apart from thinking or matter apart from extension? Unless this question is answered, how can Descartes answer the later empiricist criticism that the concept of substance is meaningless because empty of content?

In another context (ibid., p. 101) he gives an alternative definition of substance. "Really the notion of *substance* is just this—that which can exist by itself, without the aid of any other substance." This second definition is a bad one, being circular in expression; but clearly Descartes has in mind both here and in the quotation above simply the Aristotelian criteria (3) and (5). On the basis of these definitions, Descartes postulates three types of substance: material bodies, minds, and God. But the first two, being in a certain sense dependent on God for their existence, clearly have a lower grade of substantiality. Descartes's conception of substance and attribute is made impossible to understand by the vagueness of the notion "attribute" by which he seeks to clarify the idea of substance. If "attribute" means "property or relation," it simply is not true that all attributes are attributes of substances. For example, a color may have properties that are not properties of the colored thing. It is true of the color *red* that it is produced by light of wavelength about 7000 angstrom units, but this is not true of red objects. In any case, it seems that Descartes has simply defined *substance* and *attribute* relative to each other so that his explanation is circular and thus uninformative: Attributes are what qualify substances and substances are what have attributes.

SPINOZA. Descartes's second account of substance as that which is capable of independent existence was taken up and developed by Benedict de Spinoza in his *Ethics.* Spinoza was a student of Descartes and may be regarded as one who developed some of Descartes's ideas to consistent but surprising conclusions. Reflecting on Descartes's second account of substance, Spinoza showed that if by *substance* we mean, according to his definition, "that which is in itself and is conceived through itself," it is easy to show that there can be only one such being, the whole universe. Thus Spinoza equated substance with God and nature, the three terms being synonymous for him. This "hideous hypothesis," in David Hume's ironical phrase, has won for Spinoza the inconsistent titles of atheist and pantheist. What he did, in fact, was to demonstrate the alarming consequences for religious orthodoxy of Descartes's second definition and to indicate obliquely that substantiality in this sense is a matter of degree. Nothing in the universe is completely independent of its environment, although some things are more independent than others. A human being has a certain degree of independence of his environment but can exist only within a certain range of temperature, pressure, and humidity, and with access to air, food, and water. Other things may be more or less independent of their surroundings, and the extent of their freedom in each case is an empirical question. Spinoza did not draw this conclusion, but it is implicit in his development of Descartes.

LEIBNIZ. Another rationalist philosopher, Gottfried Wilhelm Leibniz, makes the concept of substance fundamental to his philosophical system. He uses two of the Aristotelian criteria of substance, substance as a center of change and substance as a logical subject, but adds the concept of simplicity. The basic elements of Leibniz's metaphysical system were what he called monads. In his *Monadology* he defines *monad* as "nothing but a simple substance.… By 'simple' is meant 'without parts.'" That there are such simple substances follows, for Leibniz, from the admitted fact that there are compound things, which can be nothing but collections of simple things. Leibniz seems here to have been influenced by the arguments of the ancient materialists for the existence of atoms. His monads, however, were supposed to be immaterial substances, centers of change and thus subjects of predicates. Unfortunately, by describing his substances in this way, he deprives the term of meaning just as Descartes had done. He does indeed affirm that his monads are centers of activity, but this activity is manifested only in their tendency to move from one state to another. But if the essence of something is to be the *x* that under-

goes changes and of which predicates can be affirmed, it can have no positive character of its own. In Bertrand Russell's words, "substance remains, apart from its predicates, wholly destitute of meaning" (*The Philosophy of Leibniz,* p. 50).

LOCKE, BERKELEY, HUME. Leibniz had criticized the British empiricist philosopher John Locke for professing to find substance an empty concept. The weakness of Locke's criticisms of the concept was that he concentrated his attack on the notion of a substratum of qualities. This is not the most important of the Aristotelian senses of the term. But if "substratum" can be shown to be an empty notion, it is easy to raise skeptical doubts about some of the associated senses, particularly those of substance as a center of change, as the concrete individual, and as a logical subject. Locke points out that we find in experience groups of qualities that occur together in time and place. We therefore presume these qualities to belong to one thing and come to use one word, "gold," "apple," or "water" (whatever it may be) to refer to the collection of properties "which indeed is a complication of many ideas together." Further, "not imagining how these simple ideas *can* subsist by themselves, we accustom ourselves to suppose some *substratum* wherein they do subsist, and from which they do result, which we therefore call *substance*" (*Essay concerning Human Understanding,* Book II, Ch. 23).

Substance, then, is not a positive concept but merely an "obscure and relative" notion of "the supposed but unknown support of those qualities we find existing, which we imagine cannot exist *sine re substante* without something to support them." Since Locke has already tried to show that all our meaningful concepts originate in experience, substance is an awkward counterexample to his theory of knowledge. Indeed, he would probably have rejected it altogether but for certain associated moral and theological doctrines that his cautious and conformist temperament made him forbear to reject outright. Moreover, he seems to have been unable to reject Descartes's principle that attributes must inhere in a substance, although he does not submit this supposed logical truth to any rigorous examination.

However, Locke's empiricist successors, George Berkeley and Hume, were fully aware of the importance of Locke's criticism and his reduction of the notion to "an uncertain supposition of we know not what." Berkeley's attack on the concept of material substance owes much to Locke, and Hume was content to write off the whole idea as an "unintelligible chimaera." Moreover, Hume extended the skepticism of Locke and Berkeley in respect of material substance to question, on analogous grounds, the existence of spiritual substances or selves. It is clear that a mind whose function is merely to be the bearer of states of consciousness is as vacuous a notion as Locke's material "we know not what."

KANT. Immanuel Kant's *Critique of Pure Reason* (1781) transformed the notion of substance, as it did so many other philosophical concepts. In Kant's view, "substance" does not refer to a feature of the objective world independent of human thinking. On the contrary, the unity and permanence of substances are features contributed by the human understanding to the world of phenomena. This represents a very radical revision of the concept of substance. Substance shrinks from being a fundamental feature of the objective world to an aspect under which men cannot help classifying their experience—and they cannot help themselves not because of the nature of external reality but because of the structure of their own cognitive apparatus.

MODERN CRITICISM

Since Kant's day the permanent and valuable features of philosophy have been those that have grown out of the immense development of the formal and natural sciences from the end of the eighteenth century to the present, a development that has shown the falsity of the scientific assumptions on which the Kantian revolution was built. For example, Kant believed that Newtonian physics, Euclidean geometry, and Aristotelian logic were finally and beyond all question true of the world, and some features of his system depend on these assumptions. This development has presented the problem of substance as a problem soluble, if at all, in the light of empirical evidence drawn from the relevant sciences. It has, moreover, made clear that there is no one problem of substance but a number of subproblems that can be treated independently.

These problems can still be stated in something like their original Aristotelian form, but we may find ourselves looking in different areas of knowledge for their answers. There is no one unitary science, such as metaphysics or ontology, that can be looked to for a solution. For example, the notion of substance as a logical subject of predicates (as when we say of a piece of gold, "*It* is heavy," "*It* is yellow," "*It* is malleable," "*It* melts at 1063° C," and so on) is now seen to be a problem of interest to formal logic and to linguistics. It is a technical question of logic whether all sentences about individual things can be

(or must be) expressed in subject-predicate form. And it is a technical question of linguistics whether all languages use such a form to express these notions, or indeed have a subject-predicate syntax at all. (The answer in both cases seems to be "No.")

INDEPENDENT EXISTENTS. The question "What, if anything, is capable of independent existence?" can be seen, insofar as it relates to material things, to be a question to which physics, chemistry, and biology give us the answers. (If the question is asked about the existence of nonmaterial things such as numbers or propositions, we have first to make clear what is meant by "existence" in such contexts.) We see that *independent* is not a term with a clear meaning but, rather, is an elliptical expression. "X is capable of independent existence" means "X is capable of existing without regard to features $y_1, y_2, \cdots, y_n$ of its environment." Since these conditions are so numerous, it is easier to express the concept negatively: "X is not independent" means "X is incapable of existing apart from conditions $z_1, z_2, \cdots, z_n$" or "$z_1, \cdots, z_n$ are necessary conditions for the existence of X." On this interpretation, a substance in the sense of something that is capable of completely independent existence is something for whose existence there are no necessary conditions. The specific values of the variable z will vary with the value of X. For example, if X is a piece of ice or a lump of metal, one of the z's will be temperature; if X is a green plant, the z's will include light and oxygen; and so on. It may well be that nothing in the universe is independent of all conditions, but whether this is so is an empirical question.

ESSENCES. Aristotle's favorite, but least satisfactory, account of substance was that of substance as essence, an essence being a set of qualities that conjointly embody the nature of the thing they qualify, are grasped by intellectual intuition, and are expressed in the definition of the thing. But developments in the sciences (especially in biology) and in the philosophy of science over the past century have shown that this notion is illusory. Definitions, in the contemporary view, are either descriptions of current linguistic usage or recommendations for linguistic conventions. They cannot seek to explicate the essential nature of the definiendum because naturally occurring objects have no such invariable natures. Definitions in formal sciences like mathematics and logic do delineate the invariant properties of the definienda precisely because they are proposals for conventions.

SUBSTRATUM. There remains for consideration substance in the senses of (*a*) a center of change, (*b*) a substratum of qualities, and (*c*) the concrete individual thing. Senses (*a*) and (*b*) are closely akin and are both vulnerable to the empiricist line of criticism made famous by Locke. We may regard a particular thing as qualified by different properties at different times (for example, when an insect changes from egg to caterpillar to pupa to moth), or as qualified by a group of qualities at the same time (for example, when we say that a lump of sugar is white and sweet and soluble). Both of these ways of looking at substance lead to the unanswerable question "What is it that is the bearer of the qualities in each case?" But the answer to this cannot even be as satisfactory as Locke's "something we know not what," for by thus separating the subject (or hypothetical bearer of the qualities) from its predicates, we effectively prevent ourselves from saying anything about it. For to say anything about it is merely to assign to it one more predicate. This way of explaining substance makes it an empty concept.

Yet the obvious alternative to this blind alley seems no more promising. Suppose that when we say "Some apples are red" we do not mean what contemporary logic teaches us to mean: There is an x that has both the property of being an apple and the property of being red. Suppose that instead we mean: That set of particular properties which we call "apple" includes the further property of being red. Then the relation "being predicated of" turns out to be nothing more than the familiar relation of being a member of a group. This conclusion looks innocuous until we realize that this interpretation would make all subject-predicate affirmations either necessarily true or logically false. For the proposition "The set of properties $Q_1, Q_2, \cdots, Q_n$ contains the property Q_n" is a logically true statement. And if we amend it to make it informative thus: "The set of properties $Q_1, \cdots, Q_n$ contains the property Q_{n+1}" we do not have an informative proposition but, rather, a logically false one.

The way out of this dilemma is not to ask such misleadingly general questions as "What is the locus of change?" or "What is the bearer of properties?" We can ask for the detailed history of a particular thing, an insect, a plant, a man or what not, and the answer will be given to us by the relevant sciences—embryology, anatomy, physiology. We can ask for the detailed structure of a particular thing, a piece of gold, a moth, a man, or what not; again the relevant science—physics, chemistry, anatomy—will give us the answer if the answer is known. But we cannot ask for the history or structure of things in general, for there is no science of things in general.

CONCRETE INDIVIDUALS. A similar criticism awaits the last of the Aristotelian answers to the question about substance: A substance is a concrete individual thing. We cannot sensibly ask what makes things-in-general concrete individuals. The notion of a concrete individual thing is clear in its standard cases, like men, tables, mice, or stones. But it is unclear in its nonstandard applications. Is a cloud a concrete individual or is it just the particles that make it up that can be so called? Is a rainbow? Or a dream table? Can electrons be called individual things when it is impossible in principle to identify them and trace their continuous histories? Examples such as these show the futility of trying to find a general formula that will clarify the notion of a concrete individual thing. We can, of course, ask the psychologists what perceptual characteristics of things lead us to class them as individuals. That a set of jointly occurring properties stands out in our perceptual field, that it moves as one, that it persists through time—all these and other characteristics will lead us to regard a thing as a thing. But there is no decisive test which will enable us to decide, if we are doubtful, whether a certain *x* is really a concrete individual or not. In borderline cases this must be a matter for decision, not diagnosis.

Bibliography

BOOKS

Aristotle. *Metaphysics* and *Categories* (Oxford translations). Oxford, 1908 and 1928.

Blanshard, Brand. *The Nature of Thought.* London: Allen and Unwin, 1939. Ch. 3.

Broad, C. D. *An Examination of McTaggart's Philosophy.* Cambridge, U.K.: Cambridge University Press, 1933. Vol. I, Ch. 7.

Leibniz, G. W. *Monadology,* edited by Robert Latta. Oxford, 1898.

Locke, John. *An Essay concerning Human Understanding,* edited by J. Yolton. London, 1962. Book II, Ch. 23.

Russell, Bertrand. *The Analysis of Matter.* London: Kegan Paul 1927. Ch. 23.

Russell, Bertrand. *Human Knowledge.* London: Allen and Unwin, 1948.

Russell, Bertrand. *An Inquiry into Meaning and Truth.* London: Allen and Unwin, 1940.

Russell, Bertrand. *The Philosophy of Leibniz.* Cambridge, U.K., 1900.

Shoemaker, Sydney. *Self-Knowledge and Self-Identity.* Ithaca, NY: Cornell University Press, 1963.

Strawson, P. F. *Individuals.* London: Methuen, 1959.

ARTICLES

Anscombe, G. E. M. "Aristotle. " In *Three Philosophers.* Oxford, 1963.

Anscombe, G. E. M., and S. Körner. "Substance." *PAS* supp. 38 (1964).

Basson, A. H. "The Problem of Substance." *PAS* (1948–1949).

Bennett, Jonathan. "Substance, Reality, and Primary Qualities." *American Philosophical Quarterly* (1965).

Cousin, D. R. "Aristotle's Doctrine of Substance." *Mind* (1933).

Kneale, William. "The Notion of a Substance." *PAS* (1939–1940).

Lazerowitz, Morris. "Substratum." In *Philosophical Analysis,* edited by Max Black. Ithaca, NY: Cornell University Press, 1950.

Mei, Tsu-Lin. "Chinese Grammar and the Linguistic Movement in Philosophy." *Review of Metaphysics* 14 (1961). An original and penetrating critique of the notions of subject and predicate.

Sellars, Wilfrid. "Substance and Form in Aristotle." *Journal of Philosophy* (1957).

Stebbing, L. Susan. "On Substance." *PAS* (1929–1930).

D. J. O'Connor (1967)

SUBSTANCE DUALISM

See *Dualism in the Philosophy of Mind*

SUFI PHILOSOPHY

See *Sufism*

SUFISM

The origins of Sufism (*taṣawwuf* in Arabic), or Islamic mysticism, appear clearly in the spiritual practise of the Prophet Muhammad in seventh-century Arabia (Massignon 1954, Lings 1993). Sufism's key contemplative discipline, remembrance of God (*dhikr*), was practiced continually by the Prophet and is alluded to in fifteen verses of the Qur'ān. From this practise the Sufis developed an entire science of invocations and supplications (*adhkār*) designed to cultivate the heart, refine the soul, and elevate ordinary human consciousness into awareness of the ever-immanent divinity (Chittick 1987). There are nonetheless a number of formative influences on early Sufism that are extraneous to early Qur'ānic spirituality. Michael Sells (1996) has demonstrated that the heritage of pre-Islamic poetry provided numerous subthemes (for example, drunkenness, love-madness, perpetual wandering, the secret shared between lover and beloved) for later Sufi literature and poetry. Scholars such as D. Miguel Asin Palacios, Tor Andrae, Duncan Macdonald, Louis Massignon, Henry Corbin and Luce López-Baralt have revealed how some of the ascetic and mystical tendencies in early Sufism bear close resemblances to

Christian mysticism, a thesis adumbrated by Tor Andrea's *In the Garden of Myrtles: Studies in Early Islamic Mysticism* (1987).

QUR'ĀNIC ORIGINS AND FORMATIVE INFLUENCES FROM THE SEVENTH TO TENTH CENTURY

The word *Ṣūfī* as a technical term does not itself come into use before the end of the eighth century CE. The last of the following four possible etymologies of the word (there is no consensus) reflects the relation of the movement with Greek philosophy: from *Ahl-i Ṣuffa*, "the People of the Veranda," the Prophet's most intimate companions in seventh-century Medina; from *ṣafā*, meaning purity; from *ṣūf*, meaning wool; and from the Greek *sofos*, that is, *sagesse*, a cognate of *sophia* ("wisdom"). In the context of the last-cited etymology, Sufism appears to be related to Islamic "philosophy" or *falsafa* in Arabic, *faylasūf* (philosopher) being the Arabic transcription of the Greek *philosophos*. Although the terms *Sufi* and *Sufism* are historically applicable only to the type of mystic and mysticism developed within Islam, based upon pursuit of the Prophet's exemplary practice (*sunna*), it is undeniable that many of the theosophical elements in Sufism, especially as the mystical tradition changed and developed over the course of later centuries, are largely derived from Greek thought.

Mystical teachings are usually ascribed to a number of the Companions (*al-aṣḥāb*) of the Prophet and their "followers" (*al-tabā'iyun*) (Ernst 1999), the first and foremost being the fourth Sunni Caliph ʿAlī ibn Abī Ṭālib (d. 661) whose sermons, letters, poems, and maxims were compiled by Sharīf al-Raḍī (d. 1015) in the *Nahj al-balāghah*. ʿAlī features as the starting-point of all the esoteric initiatic chains of Sufism, whether Sunni or Shīʿite, and is recognized as the founder of two fundamental Sufi doctrines: renunciation of the world (*zuhd*) and spiritual poverty (*faqr*). His possession of gnostic insight and esoteric knowledge (*ʿilm-i ladunī*) is acknowledged by all Muslim theologians, Sufi mystics, and philosophers.

Ḥasan al-Baṣrī (d. 728), the principal founder of the early ascetic movement of Islam that later became known as Sufism, is listed as Imam ʿAlī's succeeding link in most Sufi initiatic chains among the "followers" of the Prophet's "Companions."

The next most significant figure in Sufi thought is the sixth Shīʿite Imam Abū Jaʿfar al-Ṣādiq (d. 765), the author of the earliest mystical Qur'ān commentary, described as "the soundest of all the Shaykhs, upon whom all of them rely. … He is the path-master of the people of love (*pīshvā-yi ahl-i ʿishq*) (ʿAttār 1993, p. 12). In fact, the love mysticism of Sufism may be traced back to both al-Ṣādiq and to his contemporary, Rābiʿa al-Adawiyya (d. 788–792), the most famous female Sufi in all history, of whom Ibn ʿArabī commented, "She is the one who analyzes and classes the categories of love to the point of being the most famous interpreter of love."

It was in the ninth century, when Greek philosophy was being introduced into Islam and when all the technical vocabulary of philosophy and theology in the Arabic language was being fashioned, that most of the basic technical terms, concepts, and categories of Sufism were also elaborated. It was probably in response to the Neoplatonic philosophers of the "School of Baghdad" (revolving around Caliph al-Ma'mūn, who supported the translation of Greek works into Arabic and Syriac) that the Sufis of the ninth century first began to use the term *mystical knowledge* (or *maʿrifat*) instead of *rational knowledge* (or *ʿilm*) to refer to the type of experiential, gnostic knowledge they possessed, in order to distinguish it from the mental, purely theoretical knowledge of their contemporaries, the Neoplatonists. (Danner 1987, p. 254).

It is not mere historical coincidence that both of these celebrated Schools of Baghdad—that of the philosophers and that of the Sufis—evolved at exactly the same time and place. From the early ninth century, Muslim Peripatetic philosophy and Sufi mysticism shared a common psychological vocabulary simultaneously fed by the two streams of Qur'ānic spirituality and Greek philosophical writings, which had been translated into Arabic. Although the intellectual contexts and applications of this vocabulary differed greatly, the lexicon of both was often identical; a huge stream of common terms flowed through both systems from the two sources. For instance, in psychology, both Sufi mystics and Peripateric philosophers shared a common terminology: for soul, *nafs*; for spirit, *rūḥ;* for heart, *qalb*, for phantasy, *wahm*; for imagination, *khiyāl;* for reason, *ʿaql*. While all these terms also figure prominently in the Qur'ān, they were corralled and culled as suitable translations (as Harry Wolfson [1935] established in a seminal article) by Muslim thinkers such as al-Farabi, Avicenna, and Ghazālī.

In the ninth century three mystics were of primary importance for the development of Sufi esoteric and mystical terminology. The first two are vaunted for their role in the development of psychospiritual terminology of Sufism, whereas the third is famous for his unusual but highly influential mystical theology. All three affected the

formulation of Sufi philosophy, if *philosophia* is understood in its literal sense as love of divine wisdom.

The first figure was al-Ḥarīth al-Muḥāsibī (d. 857), who lived and taught in Baghdad. From the standpoint of formulation of mystical doctrine, psychological examination of the spiritual life, and authorship of definitive textbooks on both subjects, he was indubitably the most illustrious Sufi of the ninth century. As "the real master of primitive Islamic mysticism," as Margaret Smith put it, most later elaboration and exposition of Sufi technical terminology—such as self-examination (*muḥāsaba*), contemplation (*muraqaba*), fear (*khawf*), hope (*raja*'), patience (*ṣabr*), contentment (*riḍā'*)—can be traced back to terminology that first appeared in his works.

The second figure, Dhū'l-Nūn al-Miṣrī (d. 861), "the founder of theosophical Súfiism," as Nicholson (1906) rightly calls him, played a formative role in the evolution of Sufi doctrine. He had been the first to provide a systematic teaching about the mystical states and spiritual stations (*aḥwāl u maqāmāt*) of Sufism and was also the first to discourse on mystical knowledge, or *maʿrifat*, and to distinguish it from academic knowledge, or *ʿilm*. He was also founder of the practice and theory of the "art of audition to music" and the first to describe in poetic detail the types of "ecstatic rapture" (*samāʿ* and *wajd*), which ensued from this aesthetic tool of contemplative vision. He was the also the first mystic to use the imagery of the wine of love and cup of mystical of gnosis poured out for the lover (Smith 1991).

However, it was the third figure, Abū Yazīd (or Bāyazīd) Bisṭāmī (d. 848 or 875), who personified the Muslim mystic par excellence and who served as the real cornerstone of the free-spirited classical Sufism of later generations. He is the most frequently cited mystic in Sufi poetry. Bāyazīdian Sufism still represents the zenith of anticlerical thinking in Islam. His paradoxical utterances (he wrote nothing down), transmitted by word of mouth by disciples, soon became the subject of intricately argued prose commentaries and complicated Sufi metaphysical compositions in prose and verse. A century after his death, a separate Bāyazīdian school came into being; some two centuries later this school's contours became intellectually formalized in ʿAlī Hujwīrī's (d. 1071) *Kashf al-maḥjūb*, a Persian manual of Sufi teachings and doctrine, in which Bāyazīd's followers are classified as comprising a separate school of thought known as the Ṭayfūriyya and described as advocates of rapture (*ghalabat*) and intoxication (*sukr*) as opposed to Junayd's School of Sobriety (*saḥw*). Of particular importance in Sufi philosophy is Bāyazīd's doctrine of *fanā*', or annihilation, of the selfhood or individual ego identity in God's Self-identity, enabling the mystic to contemplate God directly through God's own eye (Rūzbihān 1966, p. 115).

Aside from these three key Sufis, there were a number of other significant mystics in the history of ninth-century Sufism, most notably Ḥakīm al-Tirmidhī (d. 908), from the Transoxanian town of Tirmirdh, one of the most interesting and prolific authors to write on themes such as sanctity and prophethood. His works became the subject of commentaries by later Sufis such as Ibn ʿArabī.

The main center for the development of Sufi doctrine in the ninth and tenth centuries was Khurāsān, in northern Iran, and the city of Nishāpūr, which, following the fall of Baghdad to the Buwayhids in 945, became the center of Sunni Islam for the next two centuries. Nishāpūr was the center of the antiascetic Sufi school of the *Malāmatiyya* (lovers of blame), whose masters enjoined their students to practice psychological introspection into the blemishes of the "lower soul" (*nafs*), or ego, and to expose their personal faults in public. Its central teacher, Abū Ḥafṣ Ḥaddād (d. 874–879), advocated opening oneself to public blame, concealing all one's own praiseworthy virtues from public scrutiny while accusing oneself of spiritual shortcomings. Its two other main representatives in Nishāpūr, Hamdūn al-Qaṣṣār (d. 884) and Abū 'Uthmān al-Ḥīrī (d. 910), were famous for nonconformist mysticism: Qaṣṣār criticized as egotistical those who overtly perform *dhikr*, and al-Ḥīrī reproached as hypocritically impious those who engaged in acts of devotion with any degree of awareness of self.

Three important developments in Sufism—institutional, aesthetic and pedagogical—took place in Nishāpūr at the end of the ninth century. Regarding the institutional developments, Margaret Malamud (1977) and Jacqueline Chabbi (1994) have shown that, in the early ninth century, some of the earliest Sufi *khānaqāh*s (meeting houses) were established in Nishāpūr. Abū Saʿīd Abī'l-Khayr (967–1049) was the first person to formalize a program for institutional and communal living of disciples, codifying rules for novices in his Sufi *khānaqāh*. In mystical aesthetics, Abū Saʿīd is significant for having definitively integrated the practice of "audition to poetry with music" (*al-samāʿ*) into the Sufi devotional life. He pioneered the expression of mystical ideas in Persian verse, using the quatrain form (*rubāʿī*), in which he was the chief forerunner of Sanāʾī, ʿAṭṭār, and Rūmī (Graham 1999).

Fritz Meier (1999) has shown how a radical transformation in Sufism took place in Nishāpūr regarding the

theory of pedagogy and practice of the master-disciple relationship from end of the late ninth century onwards. The spiritual master, who had formerly figured merely as an academic instructor of a group of students, now became the main fulcrum of the *via mystica*. He was transformed into a spiritual trainer of adepts, a saint in whom the student—now disciple—is obliged to confide with childlike trust his inmost thoughts and grant unquestioning obedience, considering him as the absolute authority and ultimate judge in all matters. By the eleventh century, this aristocratic Nishāpūrian model of the spirituality came to prevail throughout the Sufi tradition worldwide.

The leader and founder of the other important mystical school of Sufism, which was centred in Baghdad, was Abū'l-Qāsim Junayd (d. 910), who perfected Muḥāsibī's orthodox teachings and utilized his terminology. Junayd's translation of Bāyazīd's sayings from Persian into Arabic and commentary on them were preserved in Abū Naṣr Sarrāj al-Ṭusī's (d. 378/988) *Illumination of Sufism* (*Kitāb al-luma' fī'l-taṣawwuf*), "the oldest surviving general account of Sufism" (Arberry 1950). Junayd elaborated Bāyazīd's doctrine of *fanā'* in depth and detail, careful to guard against the negative consequences of the doctrine, which, superficially considered, might be interpreted by Sufism's enemies as either a kind of an ontological nihilism or else a subjective interiorised pantheism; he thus rejected both the doctrine of *ḥulūl* ("incarnationism," whereby God infuses himself in man as one substance into another) and *iṭṭiḥād* ("unitive absorption" of the individual's finite selfhood in God). Junayd's sober integration of the theosophical teachings of Sufism with Islamic legalism constitutes the basis for the orthodox understanding of Sufism down to the present day.

Because of the century and city (Baghdad) in which he flourished, Junayd was highly influenced by the school of Islamic Neoplatonism that had been established there. The theory of Al-Fārābī (d. 950), known as the "second teacher" (*al-mu'allim al-thānī*) after Aristotle, was that religions constitute elaborate symbol systems to be interpreted by an elite group of sages. This rationalist esotericism found a fit gnostic reprise in Junayd's use of mystical terminology that employed Sufi symbolic sayings couched in an enigmatic and hermetic writing style (*ishārāt*). A comparison of Junayd's basic concepts (as Ali Hassan Abdel-Kader [1976] has shown) with those of Plotinus—the stages of the mystical path, the doctrine of the preexistence and postexistence of the soul, the theory of contemplation (*mushāhada*), and the idea that mundane beauty stimulates the longing of the soul for its home Yonder—reveals Junayd's intellectual fraternity with the great pagan philosopher of late antiquity.

Junayd's school of sobriety stands in contrast to the boldly unconventional mystical theology of his most celebrated contemporary, the great martyr of Sufism Manṣūr al-Ḥallāj (d. 922), to whose life and thought Louis Massignon consecrated a huge four-volume monograph, *La Passion de Husayn Ibn Manṣūr Ḥallāj: martyr mystique de l'Islam* (1982). As Massignon (1986) pointed out, Ḥallāj figures as a precursor of Ghazālī in his endeavor to bring dogma into harmony with Greek philosophy on the basis of mystic experience. Ḥallāj was a disciple of Sahl ibn 'Abd Allāh Tustarī (d. 896), famed for his esoteric Qur'ānic exegesis. Tustarī identified "the search for knowledge" (*ṭalab al-'ilm*) as incumbent upon all Muslims with mystical feeling and spiritual consciousness (*'ilm al-ḥāl*). He defined this consciousness as the deep-felt realization that God is the witness (*shāhid*) of the devotee's thoughts, words, and deeds, which, with practice, can be transmuted into realized sapience or existential verification of knowledge (*taḥqīq al-'ilm*).

At least two key philosophical doctrines in Sufism are traceable to Ḥallāj: first, the idea of Love (*'ishq*) as "essential desire" (that is, human erotic aspiration as identical with the divine Essence), which Ḥallāj's follower Abū al-Ḥasan al-Daylamī (tenth century), was first to attribute to him in the *Kitāb 'aṭf al-alif al-ma'lūf 'alā'l-lām al-ma'ṭūf* (The book of the inclination of the familiar alif toward the inclined lam), the first book on mystical love in Islam which drew on Sufism, philosophy, and Arabic court culture (*adab*). Ḥallāj's controversial usage of the Arabic *'ishq* (passionate love) for the human-divine relationship has startling similarities to the objections raised by Christian theologians against the use of the Platonic *eros* and the Latin *amor* as equivalents to the Pauline *agape*. Ibn Sīnā's (Avicenna, d. 1037) philosophical conception of love (*'ishq*) as the universal principle of being, animate and inanimate; his view of God as the First Beloved (*Ma'shūq-i awwal*) who is simultaneously loved, lover, and love, is connected with Ḥallāj's theory (Anwar 2003, Ernst 1994). Second, Ḥallāj's conception of divine union as embodying realization of the essential oneness or unification of the human spirit with God (*'ayn al-jam'*) was expressed notably in his shocking theopathic locution *Anā al-Ḥaqq* ("I am God"), for utterance of which he was martyred.

During the tenth century Persian mystics continued to compose manuals and systematic treatises on Sufism in Arabic: Abū Bakr Muḥammad al-Kalābādhī (1989) (d. 990, a native of Bukhara) wrote his pioneering *Introduc-*

tion to the Creed of the Sufis (*Kitāb al-taʿarruf li-madhhab ahl al-taṣawwuf*), an important introduction to—and integration of Islamic exotericism with—Sufism. In this work he prudently avoided any mention of Ḥallāj, still considered a heretic by the jurists. Another Sufi scholar, Abū Naṣr Sarrāj (d. 988) from Khurasan, wrote "the oldest surviving general account of Sufism" (Arberry 1950, p. 67). *Illumination of Sufism* (*Kitāb al-luma' fi'l-taṣawwuf*). One of Ḥallāj's masters, Abū Ṭālib al-Makkī (d. 996), composed the most celebrated Sufi textbook of the Baghdad School entitled *The Food of Hearts* (*Qūt al-qulūb*), which anticipated the reconciliation of mystical and legalistic Islam that would later appear in Ghazālī's works.

ABŪ ḤĀMID AL-GHAZĀLĪ'S ATTACK ON PHILOSOPHY AND THE RENAISSANCE OF SUFISM IN THE TWELFTH CENTURY

The birth of Islam's greatest mystical theologian, Abū Ḥāmid al-Ghazālī (in Ṭūs in Khurasan in 1058) occurred at the peak of the arch of the development of Islamic mystical tradition in eleventh-century Khurasan, at the precise cusp when one half of the tangent of the Persian-Arabic mystical tradition, buttressed by the rise of Arabic mystical literature (mostly composed by Persian Sufis), faced the other half of the arch's tangent, the first beginnings of Sufi literature in Persian. The two pillars of this arch were, respectively, the *malāmātī* Sufism of Abū Saʿīd Abī'l-Khayr and the Hellenistic philosophy of Abū ʿAlī Sīnā (Avicenna)—who, being affected and profoundly influenced the Sufism of his day, wrote a number of visionary works in Arabic (and the earliest philosophical work in Persian) that provided the speculative premises for the development of the love mysticism espoused by the later Persian Sufi poets.

So it is on the foundation of the Persian Sufi tradition that Ghazālī's theological achievement rests. Nearly all the major founders of Khurasani Sufism flourished during Ghazālī's era, having been born either in decades immediately before or after his birth. These included the likes of Abū 'Abd al-Raḥmān al-Sulamī (d. 1021), one of the main chroniclers of early Sufism, best known for his Arabic tract *The Generations of Sufis* (*Ṭabaqāt as-ṣūfiyya*), a compendium of the biographies of Sufis of five earlier generations that is a fundamental source for early Sufi history. 'Abd Allāh Anṣārī (d. 1089) of Heart, the leading stylist of Persian rhyming prose, translated and adapted Sulamī's tract into a Khurasanian dialect of New Persian. Almost as important as Sulamī's *Ṭabaqāt* is the best compendium of early Sufi doctrine, namely the *Treatise* (*Risāla*) on Sufism in Arabic by Abū'l-Qāsim al-Qushayrī (d. 1072) of Nishāpūr. All of these sources Ghazālī read and knew and often reproduced them verbatim in his works.

In his autobiography, *Al-Munqidh min al-ḍalāl*, Ghazālī records how he investigated the truth claims and methods advanced by four different schools of thought: scholastic theology (*Kalām*), Isma'ili pedagogy (*taʿlīm*), philosophy (*falsafa*), and Sufism (*taṣawwuf*); he concluded that the Sufi way is the highest and most perfect of them. The distinguishing dimension of Sufi teaching, he asserted, was that "it was not apprehended by study, but only by immediate experience (*dhawq*, literally "tasting"), by ecstasy, and by a moral change. (*Mā lā yumkin al-wuṣūl ilayih bā'l-taʿallum bul bā'l-dhawq wa'l-ḥāl wa tabaddal al-ṣafāt.*) I apprehended that the Sufis were men who had real experiences, not men of words (*arbāb al-aḥwāl, lā aṣḥāb al-aqwāl*)." The unstated implication of the Sufi experience was that it allowed the adept, without recourse to either theology or philosophy, to personally verify and partially access the experience of prophecy (Hodgson 1977). Ghazālī's approach to prophecy accorded with Avicenna's view of the faculty of intuition and imagination possessed by certain adept Sufis that enabled them to have access to illumination of the active intelligence (Griffel 2002). He believed that only the science of disclosure (*ʿilm al-mukāshafa*) allowed one to "gain knowledge of the meaning of prophecy and the prophet, and of the meaning of revelation" (*al-waḥy*) (Heer 1999, p. 247 and Ghazālī 1962, p. 47), which led to the privileging of esoteric visionary thinking in later Islamic epistemology.

Ghazālī consecrated two works to the Neo-Platonic philosophers, al-Fārābī and Avicenna in particular. The first of these works was his *Objectives of the Philosophers* (*Maqāṣid al-falāsifa*); written in Arabic, it closely followed Avicenna's Persian work *Dānish-nāma 'Alālī*, providing an overall account of the history of Muslim philosophy and a lucid exposition of the philosophical doctrines that he later means to criticize. The second work, *The Incoherence of the Philosophers* (*Tahāfut al-falāsifa*), was a decisive attack on the emanative metaphysics, causal theory, and the psychology of the philosophers (especially Avicenna); in this work he sets out to prove that the philosophers are unable to prove religious truths from a theoretical point of view.

Modern scholars disagree about Ghazālī's contribution to the development of later Islamic philosophy. Lenn Goodman (1992), Ahmed El-Ehwany (1995), and Fazlur Rahman (2000) view his emphasis on Sufism as fettering

philosophic method and stifling the development of science in Islam, whereas M. Hodgson (1977), S. H. Nasr, and Henry Corbin (1996) perceive his contribution as having provided an excellent philosophical basis for the rise of later Islamic intellectual mysticism (*ḥikmat* and *ʿirfān*). Although it is true that Ghazālī's *Tahāfut* put later Islamic philosophy on the defensive, his reinterpretion of *falsafa* made philosophical ideas more accessible in the Islamic intellectual milieu than they had previously been and provided a necessary niche for philosophy to flourish in orthodox Islamic theological thought. Because Sufi theories of knowledge took center stage in his epistemological thinking, from the post-Ghazālī period in Islam down to early modern times, esoteric modes of expression invariably came to enjoy great popularity. Ghazālī believed the sapience of the heart (*dhawq*) to be superior to rational knowledge (*ʿilm*) and thought that gnosis (*maʿrifat*) could be obtained by means of the Sufi practices of remembrance of God and contemplation (*al-dhikr wa'l-fikr*), visionary unveiling (*kashf*) and abstaining from all but God Almighty. In this respect, his views are identical to those of Ibn ʿArabī a century later, whose writings on these subjects closely resemble Ghazālī's.

His most important composition was a monumental opus divided into forty books entitled *Iḥyā' ʿulūm al-dīn* (*The Revivification of the Sciences of Religion*), which, in its day, was unique in its cosmopolitan scope and integration of technical terminology, ideas, and writings derived from diverse sources. The *Iḥyā'*, a highly successful attempt to revive Islamic faith and piety on the basis of Sufism, had a profound impact on the later Islamic theological tradition. It began, in fact, what has been described as "the thirteenth-century revival of Sufism" (Danner 1988) and "the reorientation of the piety of Islam on the basis of Sufism." Because of men such as Ghazālī, Sufism became "acceptable to the '*ulama*' themselves," so that "gradually Sufism, from being one form of piety among others, and by no means the most accepted one either officially or popularly, came to dominate religious life not only within the Jama'i-Sunni fold, but to a lesser extent even among Shi'is" (Hodgson 1977, 2:203).

Mention here must be made of an equally important figure in the history of Sufism, namely Ghazālī's brother Aḥmad Ghazālī (d. 1126), who was the foremost metaphysician of love in the Sufi tradition (Lombard 2003). His impact on the later Persian Sufi tradition was even more profound than that of his brother the theologian. Aḥmad was the teacher of two important figures in particular: Abū'l-Najīb al-Suhrawardī (d. 1168) (Pourjavady 2001), who was in turn the master of his nephew Shihāb al-Dīn Abū Ḥafṣ 'Umar Suhrawardī (d. 1234), the founder of the Suhrawardī order (famed as the "Mother of Sufi Orders"), who also authored the *'Awārif al-maʿārif*, a manual of Sufism so fundamental and all-encompassing that it was translated and adapted into Persian several times and taught throughout madrasas and *khānaqāh*s in the Indian subcontinent for centuries afterward. Aḥmad Ghazālī was also the master of the enigmatic mystical theologian and founder of Sufi speculative metaphysics: 'Ayn al-Quḍāt al-Hamadhānī (executed in 1132 by fanatical Muslim clerics for his uncompromising Sufi beliefs).

ILLUMINATIONISM AND THE RISE OF THE SUFI ORDERS

In terms of Islamic *philosophia*, the most important figure following Ghazālī was Shihāb al-Dīn Yaḥyā Suhrawardī (born in Suhraward, in northwest Persia, in 1154 and died in Aleppo in 1191), renowned as *Shaykh al-Ishrāq*, the "master of illuminationist theosophy" or the "sage of the theosophy of oriental lights." He was the most significant Platonic philosopher in the Eastern lands of medieval Islam. Described by Henry Corbin (1971, p. 340) as "an irregular Sufi of no formal affiliation," Suhrawardī traced his thought back to various sources: Islamicized Peripatetic philosophers (he followed Avicenna's metaphysics in many respects), the Hermetic tradition of Egypt (Hermes, Asclepius), the pre-Islamic Persians of Mazdean Iran (Kayomarth, Kaykhusraw and Zoroaster), and Greek thought (Socrates, Plato, Aristotle). His theosophy anticipated in Islam the universalistic philosophy of fifteenth-century Renaissance Platonists such as Gemistos Pletho and Marsilio Ficino. In the world of Islam, his writings were highly influential on the intellectual development of the Neoplatonist thinkers of seventeenth-century the School of Isfahan. Despite his Peripatetic roots, Suhrawardī featured Sufis in his works, considering them to be the true philosophers of Islam. In this context, he related a dream he had had of Aristotle in which the latter identified Bāyazīd Bisṭāmi, Sahl Tustarī, and Ḥallāj as the highest Muslim thinkers (Walbridge 2000).

Suhrawardī's epistemology was based on Sufi visionary experience, and in his major work, the *Philosophy of Oriental Illumination* (*Ḥikmat al-ishrāq*), he goes to considerable lengths of philosophical argument to prove the verity of mystical intuition (*kashf*). He calls this intuition "knowledge by presence" (*ʿilm-i ḥuḍūrī*), according to which the self can know things directly by virtue of the very presence of itself (Yazdi 1992). The doctrine of

knowledge by presence is one of Suhrawardī's distinctive contributions to philosophy, and his *ishrāqī* theosophy generated a philosophical school that still dominates traditional schools of Iranian thought today. His influence "was greater than that of Averroes, for while the latter was largely forgotten in the Islamic world, Suhrawardī has continually attracted Islamic readers, followers, and opponents up to our own day" (Walbridge 2000, p. 5).

The twelfth century was also graced by the presence of the founders of two of the most influential Sufi orders in later medieval Islam: Abū Yaʿqūb al-Hamadhānī (d. 1140), founder of the Naqshbandī order, and 'Abd al-Qādir al-Jīlānī (d. 1166), founder of the Qādirī order. Two of the greatest poets of Persian literary history flourished in the same century. Ḥakīm ("the Sage") Sanā'ī of Ghazna (d. between 1131 and 1150) was a pioneer in the development of the gnostic ghazal and the first Persian Sufi poet to blend poetic imagery of the sacred and the profane into a refined philosophical lyricism. Sanā'ī's follower, Niẓāmī (d. 1202), wrote a series of unrivaled romantic epics and much mystical poetry. Another important figure is Rūzbihān Baqlī (d. 1210), whose writings constitute "a vast synthesis and rethinking of early Islamic religious thought from the perspective of pre-Mongol Sufism" (Ernst 1996, p. x), furnishing us with "a vital resource for understanding the experiential basis, not simply of Persian Sufi literature, but of Sufism and indeed mysticism in general" (Ernst 1996, p. 11). His monumental *Commentary on the Paradoxes of the Sufis* (*Sharḥ-i shaṭḥiyyāt*) is an indispensable source for the interpretation of the higher reaches of Sufi apophatic theology.

The most important Persian Sufi poet of the twelfth century was Farīd al-Dīn ʿAṭṭār (d. 1221), the prolific author of numerous epic Persian poetic works. His seminal masterpiece, *The Conference of the Birds* (*Manṭiq al-ṭayr*), has been translated into most European languages. ʿAṭṭār's major prose work was the monumental compendium, in Persian, of biographies of the famous Sufis, *Tadhkirat al-awliyā'* (*Memoirs of the Saints*).

ʿAṭṭār's contemporary was Najm al-Dīn Kubrā (d. 1221), another important figure in medieval Sufism. The founder of the Kubrawiyya, also known as the Central Asian school of Sufism, Kubrā was known for his theory of light apparitions that are beheld by the spiritual imagination in the imaginal realm (*ʿālam al-mithāl*). These theories were elaborated by later Sufis of this order, who included some of the most important names of the twelfth and thirteenth centuries. Their interpretation of these phenomena, especially when combined with their adherence to the theomonist doctrine and technical terminology of Ibn ʿArabī, constitute one of the most important chapters in the history of Islamic mysticism.

Perhaps the most famous Kubrawī mystic was Najm al-Dīn Rāzī (d. 1256), author of the *Devotees' Highroad* (*Mirṣād al-ʿibād*), an important manual of Sufi methodology in which he elaborated the peculiarly Kubrawī notion of a series of subtle centers of perception (*laṭā'if*) (Rāzī 1986, p. 299ff.). He also explained the varieties of visionary contemplation (*mushāhadāt-i anwār*) (Rāzī 1986) and continued an esoteric commentary on the Qur'ān that had been begun by Najm al-Dīn Kubrā and completed by another Kubrawī master, 'Alā' al-Dawla Simnānī (d. 1326), who elaborated his own theory of the scripture's seven esoteric levels of meaning, each of which, he said, corresponded to a subtle center of light (*laṭīfa*) (Waley 1991, Elias 1995) and expressed the inner reality (*ḥaqīqa*) of one of the prophets.

The Kubrawī school also featured a number of other notable Sufis who flourished in Iran and Central Asia: Saʿd al-Dīn Ḥammūya (d. 1253), author of the *Al-Miṣbāḥ fi'l-taṣawwuf*; Sayf al-Dīn Bākharzī (d. 1260), author of the *Waqā'iʿ al-khalwa*; Abū'l-Mafākhir Yaḥyā Bakhrazī (d. 1335–1336), the author of an important Sufi manual, *Fuṣūṣ al-ādāb*; and 'Azīz-i Nasafī (d. between 1282 and1300), a Sufi philosopher from Uzbekistan who wrote a number of profoundly original works in Persian that still remain popular. In India, the Kubrawiyya played an important role down to fourteenth century. A disciple of Simnānī named Sayyid 'Alī Hamadānī (d. 1385) was the last great thinker of the order in Central Asia; he founded the Hamadānī line, and, according to legend, was responsible for the Islamization of Kashmir.

This order was also influential in China, where Sufism first established a foothold in the early fifteenth century. The writings of two Kubrawī masters, Rāzī and Nasafī, were among the first Islamic works that were translated into Chinese in the seventeenth century, thus forming the intellectual bedrock of the Chinese Islamic tradition. The development of Islam in China is inextricably connected with the translation of Sufi texts into Chinese. Prior to the twentieth century, only four Islamic books had been translated into Chinese, all of them Persian Sufi classics belonging to the Kubrawī and Ibn ʿArabī schools (Murata 1999). Sufism in China today remains dominated by the Naqshbandī and Qādirī orders (Gladney 1999).

RŪMĪ AND IBN ʿARABĪ

The thirteenth century was the golden age of Sufism, when the most celebrated Persian poet in Islamic history,

Jalāl al-Dīn Rūmī (d. 1273), appeared. He was the author of the most extensive collection of mystical poetry, with the widest pattern of meters yet seen in Persian poetry. His collection of mystical-erotic lyrics, the *Dīvān-i Shams-i Tabrīz* (compiled under the name of Shams-i Tabrīzī because the signature verse of nearly each poem bore the name "Shams," symbolic of the poet's absorption in his spiritual teacher of this name) totals some 35,000 verses. Each of these *ghazal*s (Arabic for "love-lyric") is between five and sixty lines long and expresses the mystery of their relationship, as well as the paradoxes and subtleties of the mystical theology of Sufism. Each poem was the product of an ecstatic experience realized by the poet under the influence of the Sufi music-and-dance (*samā*ʿ) ceremony, which came to be the hallmark of his order, called the Mevlevi in Turkey and later known in the West as the Whirling Dervishes. In the eighteenth and nineteenth centuries, the Mevlevi Order's (from Rūmī's sobriquet *Mawlānā*, "our teacher") exotic flowing skirts and hypnotic revolving dance became the most popular European tourist attraction east of Athens, prompting Alexander Pope in his *Essay on Man* to observe that "Eastern priests in giddy circles run, / And turn their heads to imitate the Sun."

During the last decade and a half of his life, Rūmī began to compose the *Mathnawī-yi maʿnawī* (Rhyming spiritual couplets), dictated to his disciples under the sway of rapture. Eventually comprising more than 26,000 couplets of didactic poetry, this mystical epic became Rūmī's chief literary monument. "Judged by modern standards," wrote R. A. Nicholson in 1925 in his introduction to his critical edition and translation of the poem, "the *Mathnawī* is a very long poem: it contains almost as many verses as the *Iliad* and *Odyssey* together and about twice as many as the *Divina Commedia*."

Islam's greatest mystical thinker, known as the *Magister Magnus* or Shaykh al-Akbar, Muḥyī al-Dīn Ibn ʿArabī of Spain (d. 1240), generated a new era of writing in the field of Islamic gnosis with a string of Sufi commentators on his works and a whole school of theosophy still vital in Iran, India, Turkey, North Africa, Malaysia, and neighboring areas. Ibn ʿArabī was a very prolific author and, with the possible exception of Ghazālī, has been the most extensively studied thinker in the Islamic world (Morris 1986–1987). He composed some 850 works; 700 of these are extant, and at least 450 of them are genuine. His writings were responsible for formalizing and crystallizing the largely orally transmitted doctrines of the founders of the various Sufi Orders and thus fostered a common heritage for Sufism, which was then in the process of "creating new structures and attracting a wider flock of followers." (Chodkiewicz 1991, p. 51).

His major work, *The Meccan Revelations* (*al-Futūḥāt al-makkiyya*), covers 2,580 pages of small Arabic script (in its new critical edition the work is projected to cover thirty-seven volumes of about 500 pages each). His most famous work, however, is a short work entitled *Fuṣūṣ al-ḥikam*, made up of twenty-seven chapters, each of which is devoted to the divine wisdom revealed in a particular prophet and specific divine word. Each of these prophets represents a different mode of knowing. The title may be translated as "Bezels of Wisdom," implying that each prophet in his human setting is a kind of gemstone in which "each kind of wisdom is set, thus making of each prophet the signet or sign, by selection, of a particular aspect of God's wisdom" (Austin 1980, p. 16). The first chapter of the book concerns Adam and the last concerns Muḥammad, although the prophets discussed in between are not dealt with in chronological order. For nearly five hundred years it was the most frequently commented upon work in Sufi and theological circles in the Middle East, Central Asia, and India. In fact, the *Fuṣūṣ* was the chief intellectual preoccupation of the Sufis in India, where commentaries were written on the book by Sayyid ʿAlī Hamadānī in Kashmir, Shaykh ʿAlī Mahaymī in Gujerat, and Muhammad Gisūdarāz in the Deccan (Ahmad 1963).

Ibn ʿArabī's name is inextricably associated with the doctrine of the "Unity of Being," "Oneness of Existence," or "theomonism" (*waḥdat al-wujūd*), which should not be confused with pantheism. In this view, God is identical to created beings in His manifestation but completely separate and distinguished from them in their essences, so there is no substantial continuity between God and creation. All living beings participate with God through the theophany of His divine Names (the Living, the Speaking, the Hearing, the Omniscient, and so on), for we are all manifestations of one Light—the orifices of being through which His illumination is shone. Existence thus manifests itself by means of epiphany or theophany (*tajallī*), of which there are two types: intellectual theophany (*tajallī ʿilmī*), which is a manifestation of Being that is termed the "Most Holy Emanation" (*fayḍ al-aqdās*), and existential theophany (*tajallī wujūdī*), which is termed the "Sacred Emanation" (*fayḍ al-muqaddas*). The first type of theophany belongs to the Divine Essence, appertaining to the World of Unity (*ʿālam al-aḥadiyya*); the second type hails from the World of Unicity (*ʿālam al-waḥda*). Unlike the Peripatetic philosophers and most Sunni theologians, Ibn ʿArabī believed nothing

to be external to the divinity or outside the Absolute. Existential multiplicity is not a kind of divine action outside of Being in its Essence and Attributes. He considered "Being as an unconditional absolute (*mawjūd-la-bi-sharṭ*) beyond all duality or multiplicity. According to him, the multiplicity which we observe at the sensible or spiritual levels does not affect the Unity of Being in its creative act. It simply represents its various degrees and many states. The existential theophanies, therefore, only constitute a facet of the Absolute-God who is One in His existence and many in His manifestations" (Yahia 1991, p. 36).

Knowledge of both existence and God can only be grasped imaginatively, that is, by intuitive disclosure (*kashf*) and contemplative insight (*shuhūd*), not through reason (*ʿaql*), because a likeness of God can be gained only by recourse to imagination, not reason. Ibn ʿArabī's doctrine of the metaphysical, transpersonal imagination (*khiyāl munfaṣil*), which possesses its own distinct independent ontological level (comparable to Jung's collective unconscious) lead him to espouse an epistemology that harmonizes reason and mystical insight (Chittick 1996, p. 666). God's self-manifestation (*ẓuhūr*) can thus be intuited through the theophany of His divine names, which are manifest to the visionary imagination of the mystic, who can thereby experience a supersensory reality (Izutsu 1994).

Ibn ʿArabī's writings, employing "all the tools of the theologians, philosophers, grammarians, and other specialists" (Chittick 1989, p. 289), generated "by far the most elaborate Islamic 'philosophy of religion' and religious life, a comprehensive metaphysics which offered an all-encompassing justification and explanation for the observed diversity of religions, philosophic, and spiritual 'paths' to God—whether within the multiple sects and schools of later Islamic culture, or in the wider, even multi-confessional context of the Ottoman, Safavid and Mogul empires." (Morris 1998, p. 23) As. T. Izutsu (1995, p. 552) has pointed out, "Even today the metaphysics of Ibn ʿArabī together with—or mingled with—that of Suhrawardī, the Master of Illumination (*Shaykh al-Ishrāq*), form the basis of the philosophical-gnostic world-view of Iranian Muslim intellectuals. In fact, one of his surnames, *Muḥyī al-Dīn*, meaning literally 'revivifier of religion,' manifests its living force when it is seen in terms of the role his thought has played in the historical formation of Iranian Islam."

Many of the greatest names in the annals of Persian Islam have counted themselves as disciples or at least interpreters of his doctrines. These include the likes of Awḥād al-Dīn Kirmānī (d. 1238), Ṣadr al-Dīn Qūnawī (d. 1274), Fakhr al-Dīn ʿIrāqī (d. 1289), Saʿīd al-Dīn al-Farghānī (d. 1299), ʿAzīz al-Dīn Nasafī (d. circa 1300), Mu'ayyid al-Dīn Jandī (d. 1301), ʿAbd al-Razzāq al-Kashānī (d. 1339), ʿAlā al-Dawla Simnānī, Dāwūd Qayṣārī (d. 1350), Rukn al-Dīn (Bābā Ruknā) Masʿūd Shirāzī (d. 1367), Maḥmūd Shabistarī (d. after 1339), Muḥammad Shirin Maghribī (d. 1408), Khwāja Muḥammad Parsā (d. 1419), Ṣā'īn al-Dīn Turkah Iṣfahānī (d. 1427), Shāh Nimatu'llāh Walī (d. 1431), and Shāh Dāʿī Shirāzī (d. 1464).

SUFISM IN THE SCHOOL OF ISFAHAN

Prior to the advent of the modern age, the most significant development in Islamic thought occurred in the philosophical collegium of Isfahan in Safavid Iran (1501–1722), a unique amalgam of Sufism, Shīʿism, Platonist Ishrāqī theosophy, and Islamic rationalism that was heavily grounded in the theosophical theories of classical Sufism. Although all its members exhibited a profound respect for the ethical, intellectual, and spiritual ideals of classical Persian Sufism, few of them seem to have openly accepted the requirement of following the *ṭarīqa* discipline involving obedience to a living master (*pīr, murshid*). The writings of its members are permeated with Shīʿite piety, imamology, and theology, and were intellectually inspired by the Illuminationist (*Ishrāqī*) theosophy of Shaykh Yaḥyā Suhrawardī, which mixed Peripatetic rationalism with Islamic Platonism. Its main thinker, Mullā Ṣadrā (d. 1650), drew heavily on other renowned Sufi authors such as Abū Naṣr Sarrāj, ʿAyn al-Quḍāt Hamadhānī, Abū Ḥāmid al-Ghazālī, and Ibn ʿArabī (Pourjavady 1999). In fact, as S. H. Nasr has noted, if viewed correctly in historical context, the entire later school of Ṣadrā's Transcendental Theosophy, both in Iran and India, might be better classified as a sort of "speculative Sufism" (*taṣawwuf-i naẓārī*) (Nasr 1993, p. 124) rather than as simply a species of philosophical mysticism (*ḥikmat*).

Hodgson (1977, 3:52) has noted how the Platonists of Isfahan may be compared at points with their contemporaries, the Cambridge Platonists of England in their ecumenical interests. Mīr Findiriskī (d. 1640–1641) was one of the major philosophers of the School of Isfahan and was committed to the transmission and translation of the Hindu holy books and scriptures into Persian; he composed a commentary on the *Yoga-Vāshishtha* of Vālmiki. The Muslimization of Hindu mystical thought that resulted from the efforts of such philosophers and translators both in Iran and India can be compared to Marsilio Ficino's Christianization of the Greek Neoplatonic classics in his translations of Plato and Plotinus into Latin.

SUFISM IN EASTERN EUROPE, THE MIDDLE EAST, IRAN, AND INDIA FROM THE SEVENTEENTH CENTURY TO THE PRESENT

Since the late eighteenth century Muslim Sufi orders throughout the world have been in the throes of crisis and transformation because of the combined influences of modernism, Islamist reformism, nationalism, and European colonialism. A key to these upheavals has been the continuing impact of fundamentalist Islamism on Sufism throughout the Islamic world, a trend that began in the early twentieth century.

Throughout the Sunni world, *Salafī*s (puritans claiming to be followers of the "pious forbears" of the Prophet)—particularly in Egypt—have attacked Sufism as "inauthentic," a "Trojan horse for unwarranted innovations that owe their origins to non-Muslim civilizations such as Greece, Persia, and India" (Cornell 2004, p. 59). The same attacks have occurred in other Sunni-dominated countries of the Middle East. In Algeria and Syria, Sufis are beleaguered on the one hand by the all-encroaching influence of Western secularism, which endorses the Western modernist view of mysticism as an anachronistic superstition, and Wahhābī scriptural literalists on the other.

In Eastern Europe, Sufism has been a significant force since the early fifteenth century, especially in Bosnia, where a number of leading intellectuals, thinkers, and poets, mostly followers of the Mevlevīya and Naqshbandī Orders, penned influential mystical treatises and books and wrote glosses on classical tracts. After the collapse of the Ottoman Empire, Albania became an important center of Sufism, with the majority of its inhabitants belonging to one or another Sufi order (Clayer 2001).

Since the early sixteenth century Sufism has been firmly established in Turkey "as a fundamental element of Ottoman Islamic society, where in the urban context, the Mevlevīya played an important role in the education of Ottoman elites and in the cultivation of Sufi and Persian literatures" (Lapidus 1992, p. 29). In Ottoman society, Rūmī's Mevlevī order, to which most of the country's intellectual and artistic elite belonged, became the greatest preserver of musical creativity in a religious context. The Mevlevīya produced some of Turkey's finest musicians and calligraphers and the most sophisticated religious poet of early modern times, Ghālib Dede (1799), whose poem *Beauty and Love* is a supreme work of world literature (Holbrooke 1994, Winter 1994). Although, by the end of the nineteenth century, almost every city in the Ottoman Empire possessed its own Mevlevī center (Zarcone 2000), by the early twentieth century, because of the Kemalist laws against the Orders, many of the Sufi centers were closed down or destroyed (Raudvere 2002). The law of September 1925, which stated that "from this day forth, there are not *tarikat*s, or *dervishes*, and *murid*s belonging to them, within the boundaries of the Turkish Republic" (Algar 1994, p. 55) explicitly banned all dervish gatherings, practices, and teachings. The Naqshbandī Order was subject to particular governmental persecution and harassment. Since the 1950s there has been a relaxation of some of these restrictions because of the Turkish government's attempt to harness Sufism's spiritual potential to further its own secularist sociopolitical agenda. Because the agenda of the Kemalist secular state is to counter Islamist fundamentalism with Sufism's mystical and moral universalism (ignoring its institutional, contemplative, and practical aspects), there has been a consequent revival of Sufi activities such as Mevlevī dervish dancing, and renewed interest in the cultural heritage of Sufi architecture, poetry, literature and music.

In Egypt, hardline Islamist ideologues such as Muḥammad Rashīd Riḍā (d. 1935) and Ḥasan al-Banna (d. 1949), founder of the Muslim brotherhood, condemned Sufism wholesale as a repository of corrupting opinions and ideas in Islam. Another Egyptian fundamentalist thinker, Sayyid Quṭb (d. 1966), argued that Sufism represented a debilitating, antirational, antiprogressive force in the Islamic tradition (Abu-Rabi 1988). For more than a century, Sufism in Egypt has been controlled by an elaborate state apparatus. Since 1903 the leaders of the Orders have been governed and often appointed by a Supreme Council of the Sufi Orders. In the interests of religious and state conformism, most of the transcendentalist, illuminationist, ecstatic, and unitive aspects of the Sufi tradition are publicly denigrated and suppressed in favor of a sober, reformist mysticism focused on communal moral virtues and study of *ḥadīth* and the Qur'ān. The doctrines of rapture and intoxication maintained by the great founders of Sufi theosophy such as Ḥallāj and Bāyazīd are frowned upon by the Sufi Council (Hoffman 1995).

In Saudi Arabia, Sufism is banned today by the ḥadīth-driven scripturalism of the Wahhābī literalist theologians. The entire corpus of Sufi writings, philosophy, poetry, theosophy, and literature—whether these be the more orthodox works of Ghazālī or the visionary meditations of Ibn ʿArabī, which were once accepted as a mainstay of traditional Islamic theology by a broad spectrum of believers—have been anathematized by the Wahhābī

hierarchy that controls the mosques, schools, and universities (Cornell 2004). Even the writings of great Sufi masters such as Aḥmad Ibn Idrīs (d. 1837) (the renowned Sufi saint of Moroccan origin who lived in Arabia and defended Ibn ʿArabī's Sufi doctrines in face of Wahhābī persecution) remain anathema to the Saudi fundamentalist state (Radtke 2000).

In Algeria during the nineteenth century, the Sufi orders played a leading role, among other Muslim groups, in fighting French imperialism, and stood in the vanguard of opposition to France's cultural and political colonialism (Benaissa 1997). During the twentieth century all the Sufi orders suffered persecution by the *Salafī* reformists, who accused them of backwardness and deviance from orthodoxy (Andezian 1994). In recent decades terrorist organizations, inspired by these same Algerian *Salafī*s, have continued their attack on Sufism, whereas the modernist secularist elements equate Sufism with decadence and backwardness, so that today "for many if not most educated Algerians, Sufism is virtually synonymous with 'maraboutism'—saint-worshipping idolatry, superstitious donning of amulets, snake-charming, etc." (Shah-Kazemi 1994, p. 171)

In Iran, most of the main nineteenth-century political reformers, such as Akhundzāda (d. 1878), Mīrzā Malkum Khān (d. 1908), and Mirzā Āqā Khān Kirmāni (d. 1896) attacked Sufism, castigating its alleged passivity and religious conformism (Lewisohn 1998–1999). Radical Iranian secular intellectuals of the early twentieth century, such as Aḥmad Kasravī (d. 1946) widened this critique to sweepingly condemned Sufism as "one of the deep-rooted and greatly misguided beliefs to have appeared in Islam" (Kasravī 1990, p. 79). In the Islamic Republic in the early twenty-first century, mystical philosophy (*ḥikmat*) is encouraged, and there has been a renaissance in the publication of works on classical *taṣawwuf*, with Sufis abounding in all major urban centers, but their activities and gatherings are often closely monitored by the fundamentalist state. Since 1978 the theocratic regime has tried to write Sufism out of the textbooks of Iranian history and to destroy the mausoleums of the masters and living institutions of the Orders which dot the country; nevertheless, both above and below ground the Sufi orders have managed to survive.

In Pakistan, there has been a renaissance in the publication of Sufi literature, much of it patronized by the state and nationalist interests, which underwrite editions and Urdu translations of prominent Sufi poets who composed verse in regional vernaculars. Works by the famous masters of the Chishtī, Suhrawardī, and Naqshbandī Orders from the thirteenth to the nineteenth centuries "are widely available for popular use through modern Urdu translations in India and Pakistan, and occasionally in other languages as well (Ernst 2000, p. 335). Pakistani modernists such as Muhammad Iqbal (d. 1938) have made use of classical figures such as Rūmī, Ḥallāj, and Junayd in their own writings to further their own personal philosophical agenda but have denounced khānaqāh-based Sufism and the master-disciple relationship; some have attacked as decadent the Sufi love mysticism of Persian poets such as Ḥāfiẓ. Recently, Sufism has sometimes been press-ganged to support nationalism—as in Z. A. Bhutto's claim that Sufi saints were forerunners of the modern Islamic state of Pakistan (Ernst 1997, pp. 79, 209).

From the tenth century onward, the Islamization of India "was achieved largely by the preaching of the dervishes, not by the word" (Schimmel 1975, p. 346). The two main Indian orders that dominated the cultural and religious life of the land were the Chishtiyya and Suhrawardīyya, which had been introduced into India with the foundation of the Sultanate of Dehli; within a short time thousands of their khānaqāhs and zāwiyahs had woven themselves into the complex religious culture of India, smoothing and softening relations between opposing religious identities. The rise of the Indian Bhakhti movements in the fourteenth and fifteenth centuries took place in the background and under the direct influence of *khānaqāh*-based Sufism of the Suhrawardī and Chishtī Orders (Nizami 1957).

The school of Ibn ʿArabī in India was sustained by Sufis of all the major Orders. The renewer of the Chishtī Order in northern India, 'Abd al-Quddūs Gangūhī (d. 1437), who had mastered the famous Hatha Yoga treatise Amrit Kund and who wrote Hindi poetry influenced by Nathpanthi Yogic and Bhakti traditions, strongly defended the philosophy of the 'Unity of Being' in his treatises and correspondence (Farooqi 2004, pp. 4–6). Some of the great Chishtī Sufis were ardent supporters of Ibn ʿArabī's theomonism. Shaykh Muḥibb-Allāh Ilāhābādī (d. 1648), a vicar of the grandson of 'Abd al-Quddūs Gangūhī, was known as the "Supreme Master" (Shaykh-i Kabīr) for works that defended and commented on Ibn ʿArabī's *Fuṣūṣ al-ḥikam*. (Farooqi 2004).

The rulers of the Mughal Empire, from Akbar the Great (d. 1604) down to Shāh Jahān (d. 1658), patronized Sufis of the Chishtī, Qādarī, and Naqshbandī Orders, and utilized Sufi ecumenical "unity of religions" theory to unite their Hindu and Muslim subjects. Many Sufis in India tried to bridge the differences between Hindu and

Muslim mysticism; hence one important service that the Sufi Orders and Sufis in South Asia performed was the promotion of sectarian harmony and interfaith tolerance (Islam 2002, p. 447). Mystics such as Niẓām al-Dīn Awliyā' and Dārā Shikūh were known for their tolerance of religious diversity and their appreciation of Hindu spirituality. Dārā Shikūh (d. 1659), the eldest son of the Mughal Emperor Shāh Jahān, wrote a comparative study of Sufi and Vedantic technical terms (*Conjunction of the two Oceans* [*Majma' al-baḥrayn*]) and a Persian translation of fifty-two Upanishads (*The Supreme Arcanum* [*Sirr-i akbar*]). This work was later translated into Latin by Anquetil-Duperron, inspiring Schopenhauer and a whole string of European and American philosophers after him throughout the nineteenth century. Sufis of the Chishtī and Qādirī Orders rendered the Bhagavadgītā into Persian three times during the Mughal period in the seventeenth century, with Ibn ʿArabī's theory of an underlying mystical unity of religions used by its translators to interpret Hinduism in the context of Islamic theomonism (Vassie 1999).

SUFISM IN THE CONTEMPORARY WEST

Up until the late eighteenth century, the cultural and intellectual influence of the Sufi tradition upon Western Europe had been marginal (Chodkiewicz 1994), although certain Sufi thinkers such as Ghazālī did have a formative influence upon certain Christian philosophers such as Raymond Llull (Urvoy 2004). In the nineteenth century, Persian Sufi theosophy and poetry entered the course of Western European thought through key representatives of the German Idealist and American Transcendentalist movements, particularly in the figures of Goethe in Germany and Emerson in North America, both of whom were profoundly influenced by translations of Persian Sufi mystical literature (Jahanpour 1999). During the twentieth century, the traditionalist school founded by the French metaphysician René Guénon (d. 1951)—who converted to Islam and spent the last twenty years of his life in Cairo as a Sufi shaykh of the North African Shadhili Order—have constituted the avant-garde of Sufi teaching in the West. Sufi Muslims among Guénon's followers included Frithjof Schuon, Titus Burckhardt, Martin Lings, and S. H. Nasr, whose writings endeavor to revive Muslim orthodox traditional Sufi teachings in the light of the *Sophia perennis*, aiming to address both Islamic orthodoxy and the ecumenical concerns of comparative religion. Other advocates of the *Sophia perennis* and followers of the traditionalist school who were deeply influenced by Sufism are Ananda Coomaraswamy and Aldous Huxley.

The renowned Greek-Armenian spiritual teacher G. I. Gurdjieff (d. 1949), who was steeped in Sufi theosophy, spread his teachings during the 1930s and 1940s throughout Europe and the United States through a wide circle of followers, such as P. D. Oupensky (d. 1947), P. L. Travers, René Daumal, and Maurice Nicoll. Many of Gurdjieff's followers articulated his esoteric teachings as being a kind of Sufism divorced from traditional Islam. During the same period, the so-called "Sufi Order of the West," founded by Ināyat Khān (d. 1927), an Indian musician of the Chishtī Order, preached Sufism in Europe and North America as a sort of woolly universal mysticism that could be detached from its Islamic roots. Idries Shah (d. 1996), a prolific author of more than twenty-five books on Sufism, did much to introduce Sufism to the educated middle classes in the West, particularly artists and intellectuals, teaching that Sufism lies at the heart of all religion, although his interpretation of Sufism was primarily a *malāmatī* rather than an orthodox Muslim one.

Over the past few decades, under the leadership of Dr. Javad Nurbakhsh, the Iranian Niʿmatu'llāhī order has become a major publisher of Sufi works in English, French, German, Russian, Spanish, and Italian. This order lays little emphasis on the Islamic dimension of Sufism, stressing its universalism and ethnic Persian origins. Since the 1980s, there has also been a renaissance of scholarship on classical Sufi texts in French, English, and German, and the publication of critical studies and editions of the works of the great Sufi saints in all the major European languages has blossomed. Rūmī has become the best-selling poet in the history of American poetry publishing.

There are today at least fifty different Sufi movements in North America, the literary output of which, as Marcia Hermansen (2000, p. 158) observes, "is by now so vast that it would require a volume rather than an essay to adequately discuss the history and doctrines of each of the groups in detail." Sufism and its Orders are today found throughout all the major countries of Europe; in Britain alone, there are at least twenty-five active orders whose followers' ethnic origins can be traced back to Pakistan, India, the Middle East, Iran, and West Africa (Geaves 2000).

See also al-Fārābī; al-Ghazālī, Aḥmad; al-Ghazālī, Muḥammad; Aristotle; Averroes; Avicenna; Corbin, Henry; Ficino, Marsilio; Galen; Ibn al-ʿArabī; Islamic Philosophy; Lull, Ramón; Mullā Ṣadrā; Mysticism, History of; Nasr, Seyyed Hossein; Neoplatonism; Plato;

Pletho, Giorgius Gemistus; Plotinus; Pope, Alexander; Socrates; Suhrawardī, Shihāb al-Dīn Yaḥyā; Zoroastrianism.

Bibliography

Abdel Kader, A. H. *The Life, Personality and Writings of al-Junayd*. London: Luzac & Co., 1976.

Abu-Rabi, Ibrahim M. "Al-Azhar Sufism in Modern Egypt: The Sufi Thought." *Islamic Quarterly* 32 (4) (1988): 207–235.

Ahmad, Aziz. "The Sufi and the Sultan in Pre-Mughal India." *Der Islam* 38 (1963): 142–153.

Algar, Hamid. "The Naqshbandi Order in Republican Turkey." *Islam World Report: Turkey: The Pendulum Swings Back* 1 (3) (1994): 51–67.

Andezian, Sossie. "Sufi Brotherhoods in Contemporary Algeria." *Islam World Report: Algeria: A Revolution Revisited* 1 (1) (1997): 114–127.

Andrea, Tor. *In the Garden of Myrtles: Studies in Early Islamic Mysticism*. Translated by B. Sharpe. Albany: State University of New York Press, 1987.

Anwar, Etin. "Ibn Sina's Philosophical Theology of Love: A Study of *Risalah fi al-ʿishq*." *Islamic Studies* 62 (2) (2003): 331–345.

Arberry, A. J. *Sufism: An Account of the Mystics of Islam*. London: Allen & Unwin, 1950.

ʿAṭṭār, Farīd al-Dīn. *Tadhkirat al-awliyāʾ*, edited by Muḥammad Istiʿlāmī. Tehran: Zawwār, 1993.

Benaissa, Omar. "Algerian Sufism in the Colonial Period." In *Islam World Report: Algeria: A Revolution Revisited*, edited by Reza Shah-Kazemi 1(1)(1997): 47–68.

Chabbi, Jacqueline. "Sufi Organizations and Structures of Authority in Medieval Nishāpūr." *International Journal of Middle Eastern Studies* 26 (3) (1994): 427–442.

Chittick, W. C. "Dhikr." In *Encyclopedia of Religion*, edited by M. Eliade, Vol. 4, 341–342. New York: Macmillan, 1987.

Chittick, W. C. "Ebn al-ʿArabī." *Encyclopædia Iranica* 7 (1996): 664–670.

Chittick, W. C. *The Sufi Path of Knowledge: Ibn al-ʿArabī's Metaphysics of Imagination*. Albany: State University of New York Press, 1989.

Chodkiewicz, Michel. "The Diffusion of Ibn ʿArabī's Doctrine." *The Journal of the Muhyiddin Ibn ʿArabī Society (JMIAS)* 9 (1991): 36–57.

Chodkiewicz, Michel. "La réception du soufisme par l'occident: conjectures et certitudes." In *The Introduction of Arabic Philosophy into Europe*, edited by C. Butterworth and Blake Kessel, 136–149. Leiden: Brill, 1994.

Clayer, Nathalie. "Taṣawwuf, Music and Social Change in the Balkans since the Beginning of the Twentieth Century with Special Consideration of Albania." In *Sufism, Music, and Society in Turkey and the Middle East*, edited by Anderes Hammarlund, Tord Olsson, and Elisabeth Ozdalga, 137–146. Istanbul: Swedish Research Institute, 2001.

Corbin, Henry. *History of Islamic Philosophy*.Translated by L. Sherrard and P. Sherrard. London: KPI and the Institute of Ismaili Studies, 1996.

Cornell, Vincent. "Practical Sufism: An Akbarian Foundation for a Liberal Theology of Difference." *JMIAS* 36 (2004): 59–84.

Danner, Victor. "The Early Development of Sufism." In *Islamic Spirituality*, Vol. 1: *Foundations*, edited by S. H. Nasr, 239–264. New York: Crossroad, 1987.

Danner, Victor. *The Islamic Tradition*. New York: Amity House, 1988.

El-Ehwany, Ahmed. "Ibn Rushd. E. The Way to Science." In *A History of Muslim Philosophy*, edited by M. M. Sharif, 540–564. Dehli: D.K. Publications, 1995.

Elias, Jamal. *The Throne Carrier of God: The Life and Thought of ʿAlāʾ ad-Dawla as-Simnānī*. Albany: State University of New York Press, 1995.

Ernst, Carl. *Rūzbihān Baqlī: Mysticism and the Rhetoric of Sainthood in Persian Sufism*. London: Curzon Press, 1996.

Ernst, Carl. "Rūzbihān Baqlī on Love as 'Essential Desire.'" In *God Is Beautiful and He Loves Beauty: Festschrift in Honour of Annemarie Schimmel*, edited by A. Geise and J. C. Bürgel, 181–189. Berlin: Peter Lang, 1994.

Ernst, Carl. *The Shambhala Guide to Sufism*. Boston: Shambhala, 1997.

Ernst, Carl. "Taṣawwuf." Sec. 7 in "Muslim India." In *The Encyclopedia of Islam*. 2nd ed. Vol. 10, 334–337. Leiden: Brill, 2000.

Ernst, Carl, tr. *Teachings of Sufism*. Boston: Shambhala, 1999.

Farooqi, N. R. "The Resurgence of the Chishtīs: A Survey of the Expansion and Fulfillment of a Sufi Order in Mughal India." *Islamic Culture* 78 (1) (2004): 1–26.

Geaves, Ron. *The Sufis of Britain: An Exploration of Muslim Identity*. Cardiff: Cardiff Academic Press, 2000.

Ghazālī, Abū Ḥāmid. *The Book of Knowledge, Being a Translation with Notes of the Kitāb al-ʿIlm of al-Ghazzālī's Iḥyāʾ ʿUlūm al-Dīn*. Translated by N. A. Faris. Lahore: Sh. Muhammad Ashraf, 1962.

Gladney, Dru C. "Central Asia and China." In *The Oxford History of Islam*, edited by J. Esposito, 448–457. Oxford: Oxford University Press, 1999.

Goodman, Lenn. *Avicenna*. London: Routledge, 1992.

Graham, Terry. "Abū Saʿīd ibn Abīʾl-Khayr and the School of Khurāsān." In *The Heritage of Sufism*, edited by L. Lewisohn, Vol. 1, 116–126. Oxford: Oneworld, 1999.

Griffel, Frank. "The Introduction of Avicennean Psychology into Muslim Theological Discourse." *Transcendent Philosophy* 4 (4) (2002): 359–370.

Hammarlund, Anders, Tord Olsson, and Elisabeth Ozdalga, eds. *Sufism, Music and Society in Turkey and the Middle East*. Istanbul: Swedish Research Institute, 2001.

Heer, Nicholas. "Abū Ḥāmid al-Ghazālī's Esoteric Exegesis of the Koran." In *The Heritage of Sufism*. Vol. 1: *Classical Persian Sufism: From Its Origins to Rumi*, edited by Leonard Lewisohn, 235–257. Oxford: Oneworld, 1999.

Hermansen, Marcia. "Hybrid Identity Formations in Muslim America: The Case of American Sufi Movements." *Muslim World* 90 (2000): 158–178.

Hodgson, Marshall G. S. *The Venture of Islam*. 3 vols. Chicago: University of Chicago Press, 1977.

Hoffman, Valerie. *Sufism, Mystics and Saints in Modern Egypt*. Columbia: University of South Carolina Press, 1995.

Holbrooke, Victoria. *The Unreadable Shore of Love*. Austin: University of Texas, 1994.

Ibn al-ʿArabī. *The Bezels of Wisdom*. Translated by R. J. W. Austin. New York: Paulist Press, 1980.

Islam, Riazul. *Sufism in South Asia: Impact on Fourteenth Century Muslim Society*. Karachi: Oxford University Press, 2002.

Izutsu, Toshiko. "An Analysis of *Wahdat al-Wujúd*: Towards a Metaphilosophy of Oriental Philosophies." In his *Creation and the Timeless Order of Things: Essays in Islamic Mystical Philosophy*, 66–97. Ashland, OR: White Cloud Press, 1994.

Izutsu, Toshiko. "The Basic Structure of Metaphysical Thinking in Islam." In his *Creation and the Timeless Order of Things: Essays in Islamic Mystical Philosophy*, 1–37. Ashland, OR: White Cloud Press, 1994.

Izutsu, Toshiko. "Ibn al-ʿArabī." In *Encyclopædia of Religion*, edited by M. Eliade. Vol. 6, 552–557. New York: Macmillan, 1995.

Jahanpour, Farhang. "Western Encounters with Persian Sufi Literature." In *The Heritage of Sufism*. Vol. 3: *Late Classical Persianate Sufism (1501–1750)*, edited by L. Lewisohn and David Morgan, 28–59. Oxford: Oneworld, 1999.

Kalābādhī. *Al-Taʿarruf li-madhhab ahl al-taṣawwuf*. Translated by A. J. Arberry as *The Doctrine of the Sufis*. Cambridge, U.K.: Cambridge University Press, 1989.

Kasravi, Ahmad. *On Islam and Shiʿism*. Translated by M. R. Ghanoonparvar. Costa Mesa, CA: Mazda, 1990.

Lapidus, Ira. "Sufism and Ottoman Islamic Society." In *The Dervish Lodge: Architecture, Art, and Sufism in Ottoman Turkey*, edited by Raymond Lifchez, 15–32. Berkeley: University of California Press, 1992.

Lewisohn, Leonard, ed. *The Heritage of Sufism*. Vol. 1: *Classical Persian Sufism from its Origins to Rumi 700–1300*; Vol. 2: *The Legacy of Mediæval Persian Sufism 1150–1500*; Vol. 3: *Late Classical Persianate Sufism 1501–1750*. Edited by L. Lewisohn and David Morgan. Oxford: Oneworld, 1999.

Lewisohn, Leonard. "An Introduction to the History of Modern Persian Sufism." A 2-part study in *The Bulletin of the School of Oriental & African Studies* 61, pt. 3 (October 1998): 437–464; 62, pt. 1 (February 1999): 36–59.

Lewisohn, Leonard. "Sufism and the School of Isfahan: *Taṣawwuf* and *ʿIrfān* in Late Safavid Iran (ʿAbd al-Razzāq Lahījī and Fayḍ-i Kashānī on the Relation of *Taṣawwuf, Ḥikmat* and *ʿIrfān*." In *The Heritage of Sufism*, Vol. III: *Late Classical Persianate Sufism: the Safavid and Mughal Period*, edited by L. Lewisohn and David Morgan, 63–134. Oxford: Oneworld, 1999.

Lifchez, Raymond, ed. *The Dervish Lodge: Architecture, Art, and Sufism in Ottoman Turkey*. Berkeley: University of California Press, 1992.

Lings, Martin. "The Koranic Origins of Sufism." *Sufi* 18 (1993): 5–9.

Lombard, Joseph. "Aḥmad al-Ghazālī and the Metaphysics of Love." PhD diss. Yale University, 2003.

Malamud, Margaret. "Remarques sur le développement historique des mouvements ascétiques et mystiques au Khurasan." *Studia Islamica* 46 (1977): 5–72.

Massignon, Louis. "Al-Ḥallādj." In *Encyclopedia of Islam*. 2nd ed. Vol. 3, 99–104. Leiden: Brill, 1986.

Massignon, Louis. *Essai sur les origines du lexique technique de la mystique musulmane*, nouvelle édition. Paris: J. Vrin, 1954. Translated by Benjamin Clark as *Essay on the Origins of the Technical Language of Islamic Mysticism*. Notre Dame, IN: University of Notre Dame Press, 1997.

Massignon, Louis. *La passion de Ḥusayn Ibn Mansūr Ḥallāj: martyr mystique de l'Islam*. Paris: Gallimard, 1975. Translated by H. Mason as *The Passion of Ḥallāj: Mystic and Martyr of Islam*. 4 vols. Princeton, NJ: Princeton University Press, 1982.

Meier, Fritz. "Khurāsān and the End of Classical Sufism." In *Essays on Islamic Piety and Mysticism*, translated by John O'Kane, 189–219. Leiden: Brill, 1999.

Morris, James. "...Except His Face: The Political and Aesthetic Dimensions of Ibn ʿArabī's Legacy." *Journal of the Muhyiddin Ibn ʿArabī Society* 23 (1998): 19–31.

Morris, James. "Ibn ʿArabī and His Interpreters." *Journal of the American Oriental Society* 106–107 (1986): I: 539–551; II: 732–756; (1987) II conclusion: 101–119.

Murata, Sachiko. "Sufi Texts in Chinese." In *The Heritage of Sufism*, edited by L. Lewisohn. Vol. 3, 376–388. Oxford: Oneworld, 1999.

Nasr, S. H. "'Irfān-i naẓārī u sayr u sulūk dar taṣawwuf." *Iran Nameh* 11 (1) (1993): 121–128.

Nasr, S. H., ed. *Islamic Spirituality*. Vol. I: *Foundations*; *Islamic Spirituality*. Vol. II: *Manifestations*. New York: Crossroad, 1987 and 1991.

Nasr, S. H. "The Place of the School of Isfahan in Islamic Philosophy and Sufism." In *The Heritage of Sufism*. Vol. 3: *Late Classical Persianate Sufism (1501–1750)*, edited by L. Lewisohn and David Morgan, 3–15. Oxford: Oneworld, 1999.

Nicholson, R. A. "A Historical Enquiry concerning the Origin and Development of Sufism, with a List of Definitions of the Terms *Ṣūfī* and *Taṣawwuf*, Arranged Chronologically." *Journal of the Royal Asiatic Society* Part I (1906): 303–353.

Nizami, K. H. "Some Aspects of Khanqah Life in Medieval India." *Studia Islamica* 8 (1957): 52–69.

Pourjavady, Nasrollah. "Ḥallāj va Bāyazīd Bisṭāmī az naẓār-i Mullā Ṣadrā." *Nashr-i Danish* 16 (3) (1999): 14–24.

Radtke, Bernd, John O'Kane, Knut S. Vikor, and R. S. O'Fahey. *The Exoteric Ahmad Ibn Idris: A Sufi's Critique of the Madhāhib and the Wahhābīs*. Leiden: Brill, 2000.

Rahman, Fazlur. *Revival and Reform: A Study of Islamic Fundamentalism*. Oxford: Oneworld, 2000.

Raudvere, Catharina. *The Book and the Roses: Sufi Women, Visibility and Zakir in Contemporary Istanbul*. Sweden: Bjärnums Tryckeri AB, 2002.

Rāzī, Najm al-Dīn. *Mirṣād al-ʿibād min al-mabdā' ilā'l-maʿād*. 2nd ed., edited by Muḥammad Amīn Riyāḥi. Tehran: Intishārāt-i ʿIlmī u farhangī, 1986.

Rūzbihān Baqlī Shīrāzī. *Sharḥ-i shaṭḥiyyāt*, edited by H. Corbin. Bibliothéque Iranienne 12. Tehran: Departement d'iranologie de l'Institut Franco-iranien, 1966.

Abū Naṣr Sarrāj al-Ṭūsī. *Kitāb al-Lumaʿ*, edited by R. A. Nicholson. Translated into Persian by M. Maḥabbati. Tehran: Intishārāt-i Asāṭīr, 2003.

Schimmel, Annemarie. *Mystical Dimensions of Islam*. Chapel Hill: University of North Carolina Press, 1975.

Sells, Michael, ed. *Early Islamic Mysticism: Sufi, Qur'an, Mi'raj, Poetic and Theological Writings*. New York: Paulist Press, 1996.

Shah-Kazemi, Reza. "From Sufism to Terrorism: The Distortion of Islam in the Political Culture of Algeria." *Islam World Report: Algeria: A Revolution Revisited* #1 (1) (1997): 160–192.
Smith, Margaret. "Dhū'l-Nūn." *Encyclopedia of Islam*. 2nd ed. Vol. 2, 242. Leiden: Brill, 1991.
Trimingham. J. S. *The Sufi Orders in Islam*. Oxford: Oxford University Press, 1973.
Urvoy, Dominique. "Soufisme et Dialogue Islamo-Chrétien." *Islamochristiana* 30 (2004): 55–64.
Vassie, Roger. "ʿAbd al-Raḥmān Chishtī and the Bhagavadgita: 'Unity of Religion' Theory in Practice." In *The Heritage of Sufism*. Vol. 2: *The Legacy of Mediæval Persian Sufism (1150–1500)*, edited by L. Lewisohn, 367–377. Oxford: Oneworld, 1999.
Walbridge, John, *The Leaven of the Ancients: Suhrawardī and the Heritage of the Greeks*. Albany: State University of New York Press, 2000.
Waley, Muhammad Isa. "Najm al-Dīn Kubrā and the Central Asian School of Sufism the Kubrawiyyah." In *Islamic Spirituality*. Vol. 2: *Manifestations*, edited by S. H. Nasr, 80–104. New York: Crossroad, 1991.
Winter, T. J. "Spiritual Life in Ottoman Turkey." In *Islam World Report: Turkey: The Pendulum Swings Back*, edited by Reza Shah-Kazemi, 1 (1) (1994): 32–41.
Wolfson, Harry. "The Internal Senses in Latin, Arabic, and Hebrew Philosophic Texts." *Harvard Theological Review* 28 (2) (1935): 69–133.
Yahia, Osman. "Theophanies and Lights in the Thought of Ibn ʿArabī." *JMIAS* 10 (1991): 35–44.
Yazdi, Mahdi Ha'iri. *The Principles of Epistemology in Islamic Philosophy: Knowledge by Presence*. Albany: State University of New York Press, 1992.
Zarcone, Th. "Ṭarīḳa, 5. In the Turkish Lands, from Anatolia to Eastern Turkestan." *Encyclopedia of Islam*. 2nd ed. Vol. 10, 250–253. Leiden: Brill, 2000.

Leonard Lewisohn (2005)

SUHRAWARDĪ, SHIHĀB AL-DĪN YAḤYĀ

(c. 549 AH/1155 CE–587 AH/1191 CE)

Shihāb al-Dīn Yahyā Suhrawardī was born in Suhraward, ancient Media, in northwestern Iran. He died in Aleppo, in the full bloom of youth, a victim of the vindictiveness of the doctors of the Law and of the fanaticism of Ṣalāḥ al-Dīn (the "Saladin" of the Crusaders). It is important that this philosopher not be confused with two other Sufis with similar names (Shihāb al-Dīn ʿOmar and Abuʾl-Najīb Suhrawardī).

A guiding thought dominates Suhrawardī's work: to restore the philosophy and theosophy of the sages of ancient Persia. Three centuries before it was effected in the works of the great Byzantine philosopher Georgius Gemistus Pletho, the conjunction of the names of Plato and Zoroaster was realized in the works of this thinker of Islamic Persia. Broadly outlined, this work (where the influence of Hermeticism and late Neoplatonism was also joined) brought forth an interpretation of the theory of Platonic Ideas in terms of Zoroastrian angelology. If his design reconciled itself with difficulty to the spirit of legalistic Islam, of religion and the Law, it was not, on the other hand, contrary to a spiritual Islam, bringing into play all its resources and profoundly influencing it. This employment in effect imposed on philosophy an exigency that assured it thenceforth of a completely characteristic place in Islam. Suhrawardī did not separate philosophy and spirituality; a philosophy that does not terminate in or at least tend toward a mystical and spiritual experience is a vain undertaking. Seeking out a mystical and spiritual experience without a preliminary philosophical position puts one in great danger of losing one's way. The influence of this doctrine has been considerable, especially in Iran, and endures even to the present.

The key word in Suhrawardī's entire work is (in Arabic) *Ishrāq.* Literally, it means the illumination of the sun when it arises (*Aurora consurgens*). Transposed to the spiritual plane, it means a type of knowledge which is the very *Orient* of knowledge. Suhrawardī's principal work is titled *Ḥikmat al-Ishrāq,* "Oriental" philosophy or theosophy (the term *ḥikmat ilāhīya* being the exact equivalent of the Greek *theosophia*). It deals with a philosophy that is Oriental because it is illuminative and illuminative because it is Oriental. Between these two terms there is reciprocity rather than opposition (as C. Nallino believed). The disciples and perpetuators of Suhrawardī are known as the *Ishrāqīyūn* or *Mashriqīyūn,* the "Orientals." Suhrawardī himself is designated as preeminently the *shaikh al-Ishrāq.* Prior to Islam, these "Orientals" are to him essentially the sages of ancient Persia. Their "philosophy of enlightenment" originated with the concept of *Xwarnah* (Light-of-Glory in the Avesta and Mazdaistic cosmology; *Khorreh* in Persian). In its turn, this concept dominates the entire work of the *shaikh al-Ishrāq.* "Oriental" knowledge, which is its subject matter, is essentially a discovered "presential" knowledge (*ʿilm hoḍūrī*), and intuitive perception, such as knowledge of oneself, in opposition to a type of representative knowledge (*ʿilm ṣūrī*), through the intermediary of a Form or a *species.*

This is why an entire section of our shaikh's work (among approximately fifty titles, a trilogy, each of whose constituent elements is composed of a logic, a physics, and a metaphysics) is dedicated to freeing philosophy from all accumulated obstacles attributable to the abstractions of the Peripatetics and the scholastic scholars

of Islam (the *Mutukallimūn*). This preliminary study was crowned with the work cited above, where, from the analysis of the concept of being as Light, the theory of the procession of beings of Light is disengaged (complex angelic hierarchies, deduced somehow from the esoteric interpretation of the laws of optics). To the structure of these hierarchies correspond those of the plans of the universe, which are "symbolic of each other." Suhrawardī, more particularly, seemed to have been the first to found, systematically, an ontology of the *mundus imaginalis* (*ʿālam al-mithāl*), a world of the Image and a world of the Souls (the *malakūt*), acting as an intermediary between the world of pure Intelligences (the *jabarūt*) and the sensible world. This is a world without which the visionary experiences of the prophets and mystics, as well as the suprasensible events that the philosophy of the Resurrection treats, would remain unexplained. From this another complete section of Suhrawardī's works, deliberately written in Persian, was introduced, especially to this world, as the first phase of spiritual initiation. It forms a cycle of symbolic tales in which Suhrawardī consciously followed Avicenna (Ibn Sīnā). He knew very well what he owed to Avicenna and why he was able to go further than he: Avicenna also had formulated the project of an "Oriental" philosophy, but he could not realize it, not having known its true source.

Thus did the work of the *shaikh al-Ishrāq* give rise in Islam to a current of philosophy and spirituality distinct from the three currents that are usually considered, that of *Kalām* (the rational scholastic scholars), that of the *falāsifa* (philosophers known as the Hellenists), and Sufism. It is currently said that the *Ishrāq* is to the philosophy of the *falāsifa* what Sufism is to the theology of the *kalam.* By doing this, Suhrawardī defended the cause of philosophy against the pious agnosticism of the literalist theologians, as well as against that of certain Sufi pietists. It was only because his work was ignored for so long a time in the West (where one was accustomed to assessing Islamic philosophy from the viewpoint of what was known of it by Latin Scholastics) that an exaggerated importance was attached to Averroes, whose work was considered as having attained the self-proclaimed pinnacle and terminal point of philosophy in Islam. Neither the Peripateticism of Averroes (with which the ontology of *Malakūt* was lost) nor the critique of the philosophy of Muḥammad al-Ghāzalī has had any influence on Oriental Islam, notably on Iranian philosophy. Even there, what develops is a "Suhrawardian Avicennism" to which is joined the influence of Ibn al-ʿArabī (of Andalusia, died 1240 CE, one of the greatest mystical theosophists of all time), which spread forth into the "prophetic philosophy" of Shīʿism. The influence of Suhrawardī's doctrines was later dominant in the School of Ispahan, in the sixteenth and seventeenth centuries, in the Iran of the Safavids (with the great names of Mīr Dāmād, Mullā Ṣadrā Shīrāzī, Moḥsen Fayż, Qāẓī Saʿid Qommī, and so forth), as it was also later preponderant in India in those circles influenced by the generous religious reform of Shāh Akbar. It still makes itself felt in Iran at the present time.

Bibliography

For an edition of the works of Suhrawardī, see *Opera Metaphysica et Mystica,* Vol. I, edited by Henry Corbin (Istanbul: Maarif Matbaasi, 1945), and *Oeuvres philosophiques et mystiques* (which is *Opera Metaphysica et Mystica,* Vol. II), edited by Henry Corbin (Teheran: Institut franco-iranien, 1952). The two volumes contain a long introduction in French.

See also Henry Corbin, *Histoire de la philosophie islamique,* Vol. I (Paris: Gallimard, 1964), pp. 284–304 and the detailed bibliography on pp. 360–361, and *Terre céleste et corps de résurrection: de l'Iran mazdeen a l'Iran shiʿite* (Paris: Buchet/Chastel, 1960), which contains translations of several of Suhrawardī's works.

Henry Corbin (1967)

SUHRAWARDĪ, SHIHĀB AL-DĪN YAḤYĀ [ADDENDUM]

(1155 or 1156–1191)

Shihāb al-Dīn Suhrawardī is one of the most well known, innovative, yet controversial Persian philosophers of the history of philosophy in Iran. He was executed in 1191 at the age of thirty-six by the express command of King Saladin, most probably for his illuminationist political doctrine. This doctrine is Platonist in principle, and is based on Fārābī's structure of the ideal republic, commonly known as the "Virtuous City," in which justice is achieved based on the enlightened rule of the inspired philosopher-sage. Later Shī'a scholastic political thought draws heavily on Suhrawardī's illuminationist political doctrine.

Suhrawardī authored nearly fifty works, many of them devoted to the systematic refinement and reconstruction of philosophical arguments of the prevailing Avicennan peripatetic system of his time. Suhrawardī's stipulated aim was to refine the Greek-inspired Avicennan texts, and as such he is one of the first philosophers to challenge the unquestioned superiority of Aristotle.

Suhrawardī's philosophical aim was not to refute rational philosophy, nor to reduce it to ill-defined mysticism; rather, his creative thinking represented a positive philosophical approach aimed primarily at constructing a consistent system to prove the rational validity of revealed knowledge, as well as the intuitive and the inspired, non-predicative cognitive modes.

Medieval historians and scholastic commentators recognize Suhrawardī's innovative thinking and named him the founder of a new system, the "philosophy of Illumination." Recent analytical studies of Suhrawardī's Arabic and Persian works that together define the new system have led to the revision of earlier Orientalist misrepresentations of Suhrawardī as a mystic or a theosopher. Suhrawardī was above all a rationalist thinker whose ambition in philosophy was to construct a consistent holistic system to remove presumed logical gaps in the Aristotelian scientific system known to him in Avicenna's peripatetic philosophical corpus. The aim of Suhrawardī's reconstructed system was to define a new scientific method named the "Science of Lights" (*al-ʿilm al-anwār*) that then is employed in the construction of a unified epistemological theory, named Knowledge by Presence (*al-ʿilm a-ḥuḍūrī*), capable of scientifically explaining an inclusive range of phenomena that cover the domains of sensation, intellection, intuition, inspiration, and revelation.

The Knowledge by Presence theory has been widely acclaimed in all major philosophical works in Arabic and Persian—from Suhrawardī's own time to the present—as the crowning achievement of the philosophy of illumination, and was later employed by the major Persian thinkers in their probing of theories of knowledge. For example, the much acclaimed seventeenth-century Persian scholastic philosopher, Mullā Ṣadrā, uses the illuminationist theory of Knowledge by Presence to, among other things, explain God's knowledge of things as well as man's knowledge of God. This knowledge by presence is of essence, and its construction exemplifies Suhrawardī's aim to refine and reconstruct peripatetic arguments, not to refute them. Suhrawardī attempted to prove that the Avicennan Essentialist Definition (*al-ḥadd al-tāmm*, similar to Aristotle's *horos* and *horismos*) does not provide knowledge of essence of primary principles; and that Aristotelian theory of intellectual knowledge—which in its Avicennan peripatetic formulation is seen as conjunction with the Active Intellect (acting as *dator formarum*), does not bestow principles of science to the knower.

In his analysis Suhrawardī first examined the logical law of identity and criticized knowledge by predication; he then took up the notion of union and conjunction in physics, finally constructing a unified theory as metaphysical law. The unified theory of Knowledge by Presence, then, is stated as an identity-preserving relation (literally an "illuminationist relation," *al-iḍāfa al-ishrāqiyya*) between the domains "knower" and "known," or the intellect and the intellected—or simply knowing and being. This type of knowledge is the technical refinement of Plato's "intellectual vision" plus Aristotle's logical notion of "quick wit" (*agkhinoia*); it posits priority to the self-conscious subject's immediate grasp of the real, manifest essence of objects. Suhrawardī's epistemological theory may be compared with Kant's notion of "immediate relation to objects," but is not to be reduced to Bertrand Russel's "knowledge by acquaintance," and in general anticipates Descartes's views on knowledge.

Suhrawardī's legacy defines the height of Arabic and Persian philosophy's twelfth-century rational response to the Ashʿarite and other Ghazzālī-inspired theological antirational dogma. This philosophical legacy continues to this day, where the philosophy of illumination is an accepted school of Islamic philosophy and is taught in Shiʿite scholastic circles in Iran. While the most major innovation of Suhrawardī's technical philosophical work may be seen in his unified epistemological theory, and while it is his illuminationist political doctrine that has had the widest impact on Persian intellectual and religious traditions, still illuminationist philosophy includes many technical innovations. To name a few: the definition of an independent modal operator in the construction of a superiterated modal proposition as the single form to which all types of propositions are reduced; the proof of the impossibility of the necessary and always true validity of the universal, affirmative proposition; reduction of the Figures of Syllogism, as well as other technical innovations.

Some of his ideas in ontology and cosmology should also be mentioned: In his system, God and the intellects are types of lights; creation is the propagation of abstract, countless, continuous lights as self-conscious entities, extended in durationless time from the source, becoming less intense with distance, and the source, the Light of Lights, is the most essentially luminous, thus the most visible and self-cognizant of all. The process of becoming indicates continuum being, and is defined by rapidly increasing sequences of light-essences within a time-space continuum, where measured time and Euclidean space apply to the corporeal realm, and time without measure and non-Euclidean space define a separate realm Suhrawardī names "Mundus Imaginalis," which is an

"amazing" boundary "wonderland" realm joining the domain "intellect" with the domain "soul." This realm of being is named in many later works as the locus of experiential knowledge, and the idea also impacted textual traditions beyond the purely philosophical, notably wide-ranging Persian mystical poetry.

See also Illuminationism.

Bibliography

Two of Suhrawardī's philosophical texts are now available in bilingual editions. These texts are the best source for the study of his thinking:

The Book of Radiance: A Parallel English-Persian Text, edited and translated, with introduction, by Hossein Ziai. Costa Mesa, CA: Mazda Publishers, 1998.

The Philosophy of Illumination: A New Critical Edition of the Text of Ḥikmat al-ishrāq *with English Translation, Notes, Commentary, and Introduction by John Walbridge & Hossein Ziai*. Provo, UT: BYU Press, 1999.

A number of studies have been published, and the reader may consult them as further reading on Suhrawardī and his philosophy of illumination:

Walbridge, John. *The Science of Mystic Lights: Qu'b al-Dīn Shīrāzī and the Illuminationist Tradition in Islamic Philosophy*. Harvard Middle Eastern Monographs 26. Cambridge, MA: Harvard University Press, 1992.

Ziai, Hossein. "al-Suhrawardī, Shihāb al-Dīn." In *Encyclopedia of Islam*, 2nd ed. 781–784.

Ziai, Hossein. "The Illuminationist Tradition." In *The Routledge History of Islamic Philosophy*, edited by S. H. Nasr and Oliver Leaman, 465–496. London: Routledge, 1995.

Ziai, Hossein. *Shams al-Dīn Muḥammad Shahrazūrī's* Sharḥ Ḥikmat al-Ishrāq: *Commentary on the Philosophy of Illumination*. Critical edition, with English and Persian introductions, notes, and indexes. *Cultural Studies and Research Institute* 736 (1993): xxxix, 766.

Ziai, Hossein. "Shihāb al-Dīn Yaḥyā Suhrawardī." In *The Routledge History of Islamic Philosophy*, edited by S. H. Nasr and Oliver Leaman, 434–464. London: Routledge, 1995.

Ziai, Hossein. "Source and Nature of Authority: A Study of Suhrawardī's Illuminationist Political Doctrine." In *The Political Aspects of Islamic Philosophy*, edited by Charles Butterworth, 304–344. Cambridge, MA: Harvard University Press, 1992.

Hossein Ziai (2005)

SUICIDE

What role may a person play in the end of his or her own life? Is suicide wrong, always wrong, profoundly morally wrong? Or is it almost always wrong, but excusable in a few cases? Or is it sometimes morally permissible? Is it not intrinsically wrong at all though perhaps often imprudent? Is it sick? Is it a matter of mental illness? Is it a private or a social act? Is it something the family, community, or society could ever expect of a person? Or is it solely a personal matter, perhaps a matter of right, based in individual liberties, or even a fundamental human right?

What role a person may play in the end of his or her own life is the central ethical issue in suicide around which a set of related issues also form: What should the role of other persons be towards those intending suicide? What should the role of medical and psychiatric clinicians be toward a patient who intends suicide since it is they who are said to be charged with protecting human life? What intervention may the state make to interfere with a person's intention to end his or her own life? What responsibility do others—both immediate others such as family and friends or more distant or generalized others such as employers or institutions or society as a whole—bear when a person commits suicide?

This spectrum of views about the *ethics* of suicide—from the view that suicide is profoundly morally wrong to the view that it is a matter of basic human right, and from the view that it is primarily a private matter to the view that it is largely a social one—lies at the root of contemporary practical controversies over suicide. These practical controversies include at least three specific matters of high contemporary saliency:

- *Physician-assisted suicide in terminal illness*, the focus of intense debate in parts of the world with people who have long life expectancies and with high-tech medical systems, particularly the Netherlands, the United States, the United Kingdom, Canada, Switzerland, Belgium, Germany, and Australia;
- *Hunger strikes and suicides of social protest*, as in Turkey, Northern Ireland, and wartime Vietnam;
- *Suicide bombings* and related forms of self-destruction employed as military, guerilla, or terrorist tactics in ongoing political friction, including kamikaze attacks by wartime Japan; suicide missions by groups from Tamil separatists to al-Qaeda, and suicide bombings in the conflicts in Israel, Palestine, Iraq, and elsewhere.

Ethical issues have occupied the center of attention in the philosophical discussion of suicide, but conceptual and epistemological ones also play a role, as do a broad range of further issues raised within world historical, religious, and cultural traditions.

CONCEPTUAL AND LINGUISTIC ISSUES

The term *suicide* carries extremely negative connotations. However, there is little agreement on a formal definition. Some authors count all cases of voluntary, intentional self-killing as suicide; others include only cases in which the individual's primary intention is to end his or her life. Still others recognize that much of what is usually termed suicide neither is wholly voluntary nor involves a genuine intention to die, such as suicides associated with depression or other mental illness. Many writers exclude cases of self-inflicted death that, while voluntary and intentional, appear aimed to benefit others or to serve some purpose or principle—for instance, Socrates drinking the hemlock, Captain Lawrence Oates's (1890–1912) walking out into the Antarctic blizzard to allow his fellow explorers to continue without him, or the self-immolation of war protesters. These cases are usually not called suicide but *self-sacrifice* or *martyrdom*, terms with strongly positive connotations.

Attempts to differentiate these positive cases from negative ones often seem to reflect moral judgments, not genuine conceptual differences, and the linguistic framing of a practice plays a substantial role in social policies about suicide. For example, supporters of physician-assisted suicide often use the term *aid-in-dying* as well as earlier euphemisms such as *self-deliverance* to avoid the negative connotations of *suicide* while opponents insist on the more negative term *suicide*. Islamic militants attacking civilians are called *martyrs* by their supporters and those who recruit them but *suicide bombers* by their targets and by the Western press.

Differences among languages also play a role in the conceptualization of suicide. While for example English, French, Spanish, and many other languages have just a single, primary word for suicide, German has four: *Selbstmord* (self-murder), *Selbsttötung* (self-killing), *Suizid* (the Latinate term), and *Freitod* (free death). This latter German term has comparatively positive, even somewhat heroic, connotations, making it possible for German-speakers to think about the deliberate termination of their lives in a linguistic way not easily available to speakers of English or other languages that rely on a single, principal term with strongly negative connotations.

Linguistic issues also arise in attempts to refer to the performance of the act of suicide. The expression *to "commit" suicide* has been common, echoing the phrase *to commit a crime*; contemporary suicidologists typically use a variety of less-stigmatizing alternatives, including *suicided*, *completed suicide*, and *died by suicide*.

Some authors claim that it is not possible to reach a rigorous formal definition of suicide and prefer a *criterial* or operational approach to characterizing the term, noting its varied, shifting, and often inconsistent range of uses. Translation from one language to another may also prove difficult since there is sometimes little way to preserve comparatively positive connotations of some terms. Cases of death from self-caused accident, self-neglect, chronic self-destructive behavior, victim-precipitated homicide, high-risk adventure, refusal of life-saving medical treatment, and self-administered euthanasia—all of which share many features with suicide but are not usually termed such—cause still further conceptual difficulty.

Nevertheless, conceptual and linguistic issues concerning suicide are of considerable practical importance in policy formation, affecting, for instance, coroners' practices in identifying causes of death, insurance disclaimers, psychiatric protocols, religious prohibitions, codes of medical ethics, laws prohibiting or permitting assistance in suicide, social stigma and respect, and public response to international and political issues such as suicide bombing and protest suicide.

EPISTEMOLOGICAL ISSUES

Closely tied to conceptual issues, the central epistemological issues raised by suicide involve the kinds of knowledge available to those who contemplate killing themselves. The issue of what, if anything, can be known to occur after death has generally been regarded as a religious issue answerable only as a matter of faith; few philosophical writers have discussed it directly, despite its clear relation to theory of mind. Some writers have argued that since we cannot have antecedent knowledge of what death involves, we cannot knowingly and voluntarily choose our own deaths; suicide is therefore always irrational. Others, rejecting this argument, instead attempt to establish conditions for the rationality of suicide. Others consider whether death is always an evil for the person involved and whether death is appropriately conceptualized as the cessation of life. Still other writers examine psychological and situational constraints on decision making concerning suicide. For instance, the depressed, suicidal individual is described as seeing only a narrowed range of possible future outcomes in the current dilemma, the victim of a kind of *tunnel vision* constricted by depression. The possibility of preemptive suicide in the face of deteriorative mental conditions such as Alzheimer disease is characterized as a problem of having to use that very mind which may already be deterio-

rating to decide whether to bear deterioration or die to avoid it. Still others suggest that suicide would be the most straightforward expression of normative skepticism, expressing the view that life has no meaning and nothing is of value.

SUICIDE IN WORLD HISTORICAL TRADITIONS: THE WEST

Much of the extremely diverse discussion of suicide in the history of Western thought has been directed to ethical issues. Plato acknowledges Athenian burial restrictions—the suicide was to be buried apart from other citizens with the hand severed and buried separately—and in the *Phaedo*, he also reports the Pythagorean view that suicide is categorically wrong. But Plato also accepts suicide under various conditions, including shame, extreme distress, poverty, unavoidable misfortune, and *external compulsions* of the sort imposed on Socrates by the Athenian court: Socrates was condemned to drink the hemlock. In the *Republic* and the *Laws*, respectively, Plato obliquely insists that the person suffering from chronic, incapacitating illness or uncontrollable criminal impulses ought to allow his life to end or cause it to do so. Aristotle held more generally that suicide is wrong, claiming in the *Nichomachean Ethics* that it is *cowardly* and *treats the state unjustly*. The Greek and Roman Stoics, in contrast, recommended suicide as the responsible, appropriate act of the wise man, not to be undertaken in emotional distress but as an expression of principle, duty, or responsible control of the end of one's own life, as exemplified by Marcus Porcius Cato Uticencis (Cato the Younger) (95 BCE–46 BCE), Lucretia (sixth century BCE), and Lucius Annaeus Seneca.

Although Old Testament texts describe individual cases of suicide (Abimilech, Samson, Saul and his armor-bearer, Ahithophel, and Zimri), nowhere do they express general disapproval of suicide. However, the Greek-influenced Jewish soldier and historian Flavius Josephus (37 CE–100 CE) rejects it as an option for his defeated army, and clear prohibitions of suicide appear in Judaism by the time of the Talmud during the first several centuries CE, often appealing to the Biblical text Genesis 9:5: "For your lifeblood I will demand satisfaction." New Testament does not specifically condemn suicide, and mentions only one case: the self-hanging of Judas Iscariot after the betrayal of Jesus. There is evident disagreement among the early Church Fathers about the permissibility of suicide, especially in one specific circumstance: among others, Eusebius Pamphilus (c. 264–340), Ambrose (c. 340–397), and Jerome (c. 342–420) all considered whether a virgin may kill herself in order to avoid violation.

While Christian values clearly include patience, endurance, hope, and submission to the sovereignty of God, values that militate against suicide, they also stress willingness to sacrifice one's life, especially in martyrdom, and absence of the fear of death. Some early Christians (e.g., the Circumcellions, a subsect of the rigorist Donatists) apparently practiced suicide as an act of religious zeal. Suicide committed immediately after confession and absolution, they believed, permitted earlier entrance to heaven. Rejecting such reasoning, St. Augustine asserted that suicide violates the commandment *Thou shalt not kill* and is a greater sin than any that could be avoided by suicide. Whether he was simply clarifying earlier elements of Christian faith or articulating a new position remains a matter of contemporary dispute. In any case, it is clear that with this assertion, the Christian opposition to suicide became unanimous and absolute.

This view of suicide as morally and religiously wrong intensified during the Christian Middle Ages. St. Thomas Aquinas argued that suicide is contrary to the natural law of self-preservation, injures the community, and usurps God's judgment "over the passage from this life to a more blessed one" (*Summa theologiae* 2a 2ae q64 a5). By the High Middle Ages the suicide of Judas, often viewed earlier as appropriate atonement for the betrayal of Jesus, was seen as a sin worse than the betrayal itself. Enlightenment writers began to question these views. Thomas More incorporated euthanatic suicide in his *Utopia*. In *Biathanatos*, John Donne (c. 1572–1631) treated suicide as morally praiseworthy when done for the glory of God—as, he claimed, was the case for Christ; David Hume mocked the medieval arguments, justifying suicide on autonomist, consequentialist, and beneficent grounds.

Later thinkers such as Mme. de Staël (Anne Louise Germaine, née Necker, the baroness Staël-Holstein)—although she subsequently reversed her position—and Arthur Schopenhauer construed suicide as a matter of human right. Throughout this period, other thinkers insisted that suicide was morally, legally, and religiously wrong: among them, John Wesley (1703–1791) said that suicide attempters should be hanged, and Sir William Blackstone (1723–1780) described suicide as an offense against both God and the king. Immanuel Kant used the wrongness of suicide as a specimen of the moral conclusions the categorical imperative could demonstrate. In contrast, the Romantics tended to glorify suicide, and

Friedrich Nietzsche insisted that "suicide is man's right and privilege."

Although religious moralists have continued to assert that divine commandment categorically prohibits suicide, that suicide repudiates God's gift of life, that suicide ruptures covenantal relationships with other persons, and that suicide defeats the believer's obligation to endure suffering in the image of Christ, the volatile discussion of the moral issues in suicide among more secular thinkers ended fairly abruptly at the close of the nineteenth century. This was due in part to Émile Durkheim's insistence that suicide is a function of social organization, and also to the views of psychological and psychiatric theorists, developing from Jean Esquirol (1772–1840) to Sigmund Freud, that suicide is a product of mental illness. These new *scientific* views reinterpreted suicide as the product of involuntary conditions for which the individual could not be held morally responsible. The ethical issues, which presuppose choice, reemerged only in the later part of the twentieth century, stimulated primarily by discussions in bioethics of terminal illness and other dilemmas at the end of life.

SUICIDE AND MARTYRDOM IN MONOTHEIST RELIGIOUS TRADITIONS

The major monotheisms, Judaism, Christianity, and Islam, all repudiate suicide though in each, martyrdom is recognized and venerated. Judaism rejects suicide but venerates the suicides at Masada and accepts *Kiddush Hashem*, self-destruction to avoid spiritual defilement. At least since the time of Augustine, Christianity has clearly rejected suicide but accepts and venerates martyrdom to avoid apostasy and to testify to one's faith. Islam also categorically prohibits suicide but at the same time defends and expects martyrdom to defend the faith. Yet whether the distinction between suicide and martyrdom falls in the same place for Judaism, Christianity, and Islam is not clear. Judaism appears to accept self-killing to avoid defilement or apostasy; Christianity teaches submission to death where the faith is threatened but also celebrates the voluntary embrace of death in such circumstances; some Islamic fundamentalists support the political use of suicide bombing, viewing it as consistent with Islam and its teachings of *jihad*, or holy war to defend the faith, though others view this as a corruption of Islamic doctrine.

Thus, while all three traditions revere those who die for the faith as martyrs and all three traditions formally repudiate suicide, at least by that name, the practices they accept may be quite different: Christians would not accept the mass suicide at Masada; Jews do not use the suicide-bombing techniques of their Islamic neighbors in Palestine; and Muslims do not extol the passive submission to death of the Christian martyrs, appealing on Quranic grounds to a more active self-sacrificial defense of the faith.

OTHER RELIGIOUS AND CULTURAL VIEWS OF SUICIDE

Many other world religions hold the view that suicide is prima facie wrong but that there are certain exceptions. Still others encourage or require suicide in specific circumstances. Known as *institutionalized suicide*, such practices in the past have included the *sati* of a Hindu widow who was expected to immolate herself on her husband's funeral pyre; the *seppuku* or hara-kiri of traditional Japanese nobility out of loyalty to a leader or because of infractions of honor; and, in traditional cultures from South America to Africa to China, the apparently voluntary submission to sacrifice by a king's retainers at the time of his funeral in order to accompany him into the next world. Inuit, Native American, and some traditional Japanese cultures have practiced voluntary abandonment of the elderly, a practice closely related to suicide, in which the elderly are left to die, with their consent, on ice floes, on mountaintops, or beside trails.

In addition, some religious cultures have held comparatively positive views of suicide, at least in certain circumstances. The Vikings recognized violent death, including suicide, as guaranteeing entrance to Valhalla. Some Pacific Island cultures regarded suicide as favorably as death in battle and preferable to death by other means. The Jains, and perhaps other groups within traditional Hinduism, honored deliberate self-starvation as the ultimate asceticism and also recognized religiously motivated suicide by throwing oneself off a cliff. On Mangareva, members of a traditional Pacific Islands culture also practiced suicide by throwing themselves from a cliff, but in this culture not only was the practice largely restricted to women, but a special location on the cliff was reserved for noble women and a different location assigned to commoners. The Maya held that a special place in heaven was reserved for those who killed themselves by hanging (though other methods of suicide were considered disgraceful), and, though the claim is disputed, may have recognized a goddess of suicide, Ixtab. Many other pre-Columbian peoples in the Western hemisphere engaged in apparently voluntary or semi-voluntary ritual self-sacrifice, notably the Aztec practice of heart sacrifice, which

was generally characterized at least at some historical periods by enhanced status and social approval. The view that suicide is intrinsically and without exception wrong is associated most strongly with post-Augustinian Christianity of the medieval period, surviving into the present; this absolutist view is not by and large characteristic of other cultures.

ETHICAL ISSUES IN CONTEMPORARY APPLICATION: PHYSICIAN-ASSISTED SUICIDE

The *right to die* movement emerging in the 1970s, 1980s, and 1990s, counting among its achievements the passage of *natural death*, *living will*, and *durable power of attorney* statues that gave patients greater control in decision making about their end-of-life medical care, also raised the question of what role the dying person might play in shaping his or her own death and what role the physician might play in directly assisting the patient's dying. These notions have often appealed to the concept of *death with dignity*, though the coherence of that notion is sometimes challenged. Public rhetoric quickly labeled the practice at issue *physician-assisted suicide* although less negatively freighted labels such as *physician-aid-in-dying* or *physician-negotiated death* have also been advanced as more appropriate.

Proponents of legalizing the practice have argued in its favor on two principal grounds: (1) autonomy, the right of a dying person to make his or her own choices about matters of deepest personal importance, including how to face dying, and (2) the right of a person to avoid pain and suffering that cannot be adequately controlled. Opponents offer two principal competing claims: (1) that fundamental moral principle prohibits killing, including self-killing, and (2) that allowing even sympathetic cases of physician assistance in suicide would lead down the *slippery slope*, as overworked doctors, burdened or resentful family members, and callous institutions eager to save money would manipulate or force vulnerable patients into choices of suicide that were not really their own. Pressures would be particularly severe for patients with disabilities, even those who were not terminally ill, and the result would be widespread abuse.

Compromise efforts, launched by bioethicists, physicians, legal theorists, and others on both sides, have focused primarily on the mercy argument from avoiding pain: It is claimed that improving pain control in terminal illness, including accelerated research, broader education of physicians, rejection of outdated concerns about addiction associated with opioid drugs, and recourse to terminal sedation or induced permanent unconsciousness if all else fails will serve to decrease requests for physician assistance in suicide. These compromise views also hold that assistance in suicide should remain, if available at all, a last resort in only the most recalcitrant cases.

However, although proponents of physician-assisted suicide welcome advances in pain control, many reject this sort of compromise arguing that it restricts the freedom of a person who is dying to face death in the way he or she wants. They point out that other apparent compromises, such as the use of terminal sedation, are both repugnant and can be abused, since full, informed consent may not actually be sought. Proponents also object on grounds of equity: It is deeply unfair, they insist, that patients dependent on life-support technology such as dialysis or a respirator can achieve a comparatively easy death at a time of their own choosing by having these supports discontinued—an action fully legal—but patients not dependent on life supports cannot die as they wish but must wait until the inevitable end when the disease finally kills them.

Many opponents of physician-assisted suicide reject attempts at compromise as well, sometimes arguing on religious grounds that suffering is an aspect of dying that ought to be accepted, sometimes holding that patients' wishes for self-determination ought not override the scruples of the medical profession, and sometimes objecting to any resort at all to assisted dying, even in very rare, difficult cases. And some who accept the claim that death is sometimes a benefit to which a person can be morally entitled still object that placing this choice in the hands of patient would make him or her worse off by obliging him or her to choose at all, even if the choice is against. There is little resolution, however, of the competing claims of autonomist and mercy claims on the one hand and wrongness-of-killing and social-consequences views on the other. Like the social arguments over abortion, disagreement continues both at the level of public ferment and at the deeper level of philosophical principle although the raising of the issue itself has meant far greater attention to issues about death and dying.

ETHICAL ISSUES IN CONTEMPORARY APPLICATION: SUICIDE IN OLD AGE

While comparatively rarely discussed in contemporary moral theory, the more difficult applied question concerns suicide in old age for reasons of old age alone though this is said to be an issue that will increasingly confront an aging society. In both historical argumentation and the very small amount of contemporary theoriz-

ing, the fundamental issues of suicide in old age concern two distinct sets of reasons for suicide, in practice often intertwined: (1) *Reasons of self-interest*: suicide in order to avoid the sufferings, physical limitations, loss of social roles, and stigma of old age; (2) *Other-regarding reasons*: suicide in order to avoid becoming a burden to others, including family members, caretakers, immediate social networks, or society as a whole.

Contemporary reflection, at least explicitly, countenances neither of these as adequate reasons for suicide in old age. With regard to self-interested reasons, modern gerontology maintains a resolutely upbeat and optimistic view of old age, insisting that it is possible to ameliorate many of the traditional burdens of old age—chronic illness, isolation, poverty, depression, and chronic pain—by providing better medical care, better family and caregiver education, and more comprehensive social programs. With respect to other-regarding reasons, including altruistic reasons, contemporary views consider it unconscionable—especially in the wealthy societies of the developed world—to regard elderly persons as burdens to families or to social units or to the society; nor is it thought ethically permissible to allow or encourage elderly persons to see themselves this way. While the notion that the elderly are to be venerated is associated primarily with the traditional cultures of the Asia, especially China, Western societies also insist, though sometimes ineffectually in a youth-oriented culture, on respect for the aged and on enhancing long lives. Simply put, the prevalent assumption in the Western cultures in the twenty-first century is that there can be no good *reasons* for suicide in old age even though suicide is frequent, especially in men in old age. Daniel Callahan (1930–), although opposing suicide in old age, points to contemporary medicine's relentless drive for indefinite extension of life, arguing that the elderly should forgo heroic life-prolonging care and refocus their attention instead on turning matters over to the next generation. Carlos Prado (1937–), exploring issues of declining competence, raises the issue of *preemptive* suicide in advanced age. Colorado Governor Richard Lamm's widely (mis)quoted remark that the elderly have a "duty to die," unleashed a small storm of academic and public discussion concerning suicide in terminal illness and in old age (Hardwig 1997).

Hints of real social friction can be seen over both self-interested and other-regarding and altruistic reasons for suicide in old age. Having fully legalized physician-assisted suicide and voluntary active euthanasia, the Netherlands is now considering whether to honor advance directives such as living wills in which a now-competent person requests physician-aided death after the onset of Alzheimer disease, a condition particularly frequent among the elderly. Double-exit suicides, often of married partners in advanced age even though only one is ill, startle public awareness. Disputes over generational equity in the face of rising health care costs question whether life prolongation means merely the extension of morbidity and whether health care ought to be preferentially allocated to the young rather than the old. The issue of whether a person may ethically and reasonably refuse medical treatment in order to spare health care costs to preserve an inheritance for his or her family is already beginning to be discussed; the same issue also raises the question of suicide. And issues about suicide in old age are posed by far-reaching changes in population structure, the *graying* of societies in Europe and the developed world: As birthrates fall and the proportion of retirees threatens to overwhelm the number of still-working younger people, could there be any obligation, as Euripides (c. 480–406 BCE) put it in *The Suppliants* nearly 2,500 years ago, go "hence, and die, and make way for the young"?

No party now encourages suicide for the elderly, and, indeed, no party even raises the issue; but the issue of suicide as a response to self-interested avoidance of the conditions of old age and to other-interested questions about social burdens of old age cannot be very far away. Drawing as they might on both Stoic and Christian roots in the West and on non-Western practices now coming to light, the ethical disputes over suicide in old age, independent of illness, are likely to intensify the currently vigorous debate over suicide in terminal illness: Can suicide in old age represent, as one author puts it, the last rational act of autonomous elders, or does it represent the final defeated event in a series of little tragedies of all kinds?

See also Aristotle; Augustine, St.; Consequentialism; Durkheim, Émile; Epistemology; Freud, Sigmund; Hume, David; Kant, Immanuel; More, Thomas; Nietzsche, Friedrich; Plato; Pythagoras and Pythagoreanism; Romanticism; Schopenhauer, Arthur; Staël-Holstein, Anne Louise Germanie Necker, Baronne de; Socrates; Stoicism;Thomas Aquinas, St.

Bibliography

Alvarez, Albert. *The Savage God: A Study of Suicide*. London: Weidenfeld and Nicolson, 1971.

Améry, Jean. *On Suicide: A Discourse on Voluntary Death* (1976). Translated by John D. Barlow. Bloomingon: Indiana University Press, 1999.

Thomas Aquinas. *Summa Theologiae*. 2a2ae, question 64, article 5, translated by Michael Rudick.

Aristotle. *Ethica Nicomachea* Book III, 1115a-1116a; Book V, 1138a, edited and translated by W.D. Ross. Oxford: Clarendon Press, 1925.

Baechler, Jean. *Suicides*. Translated by Barry Cooper. Oxford: Blackwell, 1979.

Barraclough, Brian M. "The Bible Suicides." *Acta Psychiatrica Scandinavia* 86 (1) (1992): 64–69.

Battin, Margaret Pabst. *Ending Life: Ethics and the Way We Die*. New York: Oxford University Press, 2005.

Battin, Margaret Pabst. *Ethical Issues in Suicide*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

Battin, Margaret Pabst. *The Least Worst Death: Essays in Bioethics on the End of Life*. New York: Oxford University Press, 1994.

Battin, Margaret Pabst, and Ronald W. Maris, eds. "Suicide and Ethics." *Suicide and Life-Threatening Behavior*. Spec. issue 13 (4) (1983).

Battin, Margaret Pabst, and David J. Mayo, eds. *Suicide: The Philosophical Issues*. New York: St. Martin's Press, 1980.

Battin, Margaret P., Rosamond Rhodes, and Anita Silvers, eds. *Physician-Assisted Suicide: Expanding the Debate*. New York: Oxford University Press, 1998. Includes Oregon Death with Dignity Act.

Beauchamp, Tom L. *Intending Death: The Ethics of Assisted Suicide and Euthanasia*. Upper Saddle River, NJ: Prentice-Hall, 1996.

Brandt, Richard B. "The Morality and Rationality of Suicide." In *A Handbook for the Study of Suicide*, edited by Seymour Perlin. New York: Oxford University Press, 1975.

Brock, Dan W. *Life and Death: Philosophical Essays in Biomedical Ethics*. Cambridge, U.K.: Cambridge University Press, 1993.

Brock, Dan W. "Physician-Assisted Suicide as a Last-Resort Option at the End of Life." In *Physician-Assisted Dying. The Case for Palliative Care and Patient Choice*, edited by Timothy E. Quill and Margaret Pabst Battin, 130–149. Baltimore: Johns Hopkins University Press, 2004.

Brody, Baruch A., ed. *Suicide and Euthanasia: Historical and Contemporary Themes*. Dordrecht, Netherlands: Kluwer, 1989.

Callahan, Daniel. *False Hopes: Why America's Quest for Perfect Health Is a Recipe for Failure*. New York: Simon and Schuster, 1998.

Cholbi, Michael. "Suicide." *Stanford Encyclopedia of Philosophy*. (Summer 2004 Edition), edited by Edward N. Zalta. Available at http://plato.stanford.edu/.

Cicero, Marcus Tullius. *On Old Age*. Trans. Frank O. Copley. Ann Arbor, The University of Michigan Press, 1967.

Daube, David. "The Linguistics of Suicide." *Philosophy and Public Affairs* 1 (4) (1972): 387–437.

DeLeo, Diego, ed. *Suicide and Euthanasia in Older Adults: A Transcultural Journey*. Göttingen: Hogrefe and Huber, 2001.

Donne, John. *Biathanatos* (1647), edited by Michael Rudick and Margaret Pabst Battin. New York: Garland, 1982.

Droge, Arthur J., and James D. Tabor. *A Noble Death. Suicide and Martyrdom among Christians and Jews in Antiquity*. San Francisco: HarperCollins, 1992.

Durkheim, Emile. *Suicide: A Study in Sociology*. Translated by John A. Spaulding and edited by George Simpson. New York: Free Press, 1951.

Emanuel, Ezekiel J. "What is the Great Benefit of Legalizing Euthanasia or Physician-Assisted Suicide?" *Ethics* 109 (3) (1999): 629–642.

Emanuel, Linda L., ed. *Regulating How We Die: The Ethical, Medical, and Legal Issues Surrounding Physician-Assisted Suicide*. Cambridge, MA: Harvard University Press, 1998.

Euripides. *The Suppliants*, line 1113. Translated by Arthur S. Way. In *Euripides*. London: William Heinemann; New York: Macmillan, 1912, vol. 3, p. 589.

Fedden, Henry Romilly. *Suicide: A Social and Historical Study*. London: Peter Davies, 1938.

Fischer, John Martin, ed. *The Metaphysics of Death*. Stanford, CA: Stanford University Press, 1993.

Fletcher, David B. "Holy Dying, Assisted Dying? An Anglican Perspective on Physician-Assisted Suicide." *Ethics and Medicine* 1 (2004): 35–43.

Foley, Kathleen, and Herbert Hendin. *The Case against Assisted Suicide: For the Right to End-of-Life Care*. Baltimore: Johns Hopkins University Press, 2002.

Francis, Leslie Pickering. "Assisted Suicide: Are the Elderly a Special Case?" In *Physician Assisted Suicide: Expanding the Debate*, edited by Margaret P. Battin, Rosamond Rhodes, and Anita Silvers, 75–90. New York: Routledge, 1998.

Gentzler, Jyl. "What Is a Death with Dignity?" *Journal of Medicine and Philosophy* 28 (4) (2003): 461–487.

Griffiths, John, Alex Bood, and Heleen Weyers. *Euthanasia and the Law in the Netherlands*. Amsterdam: Amsterdam University Press, 1998.

Hardwig, John. "Is There a Duty to Die?" *Hastings Center Report* 27 (2) (1997): 34–42.

Hill, Thoms E. "Self-Regarding Suicide: A Modified Kantian View." In *Autonomy and Self-Respect*, 85–103. Cambridge, U.K.: Cambridge University Press, 1991.

Humber, James M., and Robert F. Almeder, eds. *Is There a Duty to Die?* Totowa, NJ: Humana Press, 2000.

Hume, David. "Of Suicide." (1777). In *Essays: Moral, Political, and Literary*, 586–596. London: Oxford University Press, 1963.

Humphry, Derek. *Final Exit: The Practicalities of Self-Deliverance and Assisted Suicide for the Dying*. Eugene, OR: National Hemlock Society, 1991.

Humphry, Derek, with Ann Wickett. *Jean's Way*. London: Quartet Books, 1978.

Kamisar, Yale. "Some Non-religious Views Against Proposed 'Mercy-Killing' Legislation." *Minnesota Law Review* 42 (1958): 969–1041.

Kamm, F. M. "Physician-Assisted Suicide, the Doctrine of Double Effect, and the Ground of Value." *Ethics* 109 (3) (1999): 586–605.

Kant, Immanuel. *Lectures on Ethics*. Translated by Louis Infield. Indianapolis, IN: Hackett, 1980.

Kant, Immanuel. *Ethical Philosophy: The Complete Texts of the Grounding for the Metaphysics of Morals, and Metaphysical Principles of Virtue, Part II of the Metaphysics of Morals*. Translated by James W. Ellington. Indianapolis, IN: Hackett, 1983.

Kant, Immanuel. *The Critique of Practical Reason*. Translated by Lewis White Beck. New York: Macmillan, 1993.

Kass, Leon. "Death with Dignity and the Sanctity of Life." *Commentary* (March 1990): 33–43.

Korsgaard, Christine M. *The Sources of Normativity*. Cambridge, U.K.: Cambridge University Press, 1996.

Landsberg, Paul-Louis. *The Experience of Death; The Moral Problem of Suicide*. Translated by Cynthia Rowland. New York: Philosophical Library, 1953.

Maris, Ronald W., Alan L. Berman, and Morton M. Silverman. *Comprehensive Textbook of Suicidology*. New York: Guilford Press, 2000.

McIntosh, John L. Annual updated U.S. suicide statistics. Available from www.iusb.edu/~jmcintos/. Official data also available in hardcopy from *National Vital Statistics Reports*, National Center for Health Statistics, Hyattsville, Maryland.

Minois, George. *History of Suicide. Voluntary Death in Western Culture*. Translated by Lydia G. Cochrane. Baltimore: Johns Hopkins University Press, 1999.

More, Thomas. *Utopia*. Book II. Translated by Ralph Robinson, in *Three Early Modern Utopias*, edited by Susan Bruce. Oxford University Press, 1999, pp. 89-90

Murray, Alexander. *Suicide in the Middle Ages: The Violent against Themselves*. Vol. 1. Oxford: Oxford University Press, 1998.

Narveson, Jan. "Self-Ownership and the Ethics of Suicide." *Suicide and Life-Threatening Behavior* 13 (4) (1983): 240–253.

Nietzsche, Friedrich. *The Day of Dawn (Morgenrote)* Original edition, 1881. Stuttgart: Kroner Verlag, 1953, p. 210.

Oregon Health Department. Annual data on the Death with Dignity Act. Available from http://www.oregon.gov/.

Perlin, Seymour, ed. *A Handbook for the Study of Suicide*. Oxford: Oxford University Press, 1975.

Plato. *The Dialogues of Plato.* Phaedo 61B-69E, 116A–118A; Republic III 405A–410A; Laws IX 853A-854D, 862D-863A, 872D–873E. Translated by. Benjamin Jowett. New York: Random House, 1892, 1920, vol. I, pp. 444–453 and 499–501; 669–674; vol. II, pp. 599–600, 608, 617–618.

Prado, Carlos G. *The Last Choice: Preemptive Suicide in Advanced Age*. Westport, CT: Greenwood Press, 1990.

Putnam, Constance E. *Hospice or Hemlock? Searching for Heroic Compassion*. Westport, CT: Praeger, 2002.

Quill, Timothy E. *A Midwife Through the Dying Process: Stories of Healing and Hard Choices at the End of Life*. Baltimore: Johns Hopkins University Press, 1996.

Quill, Timothy E., and Margaret P. Battin, eds. *Physician-Assisted Dying. The Case for Palliative Care and Patient Choice*. Baltimore: Johns Hopkins University Press,2004.

Raymond, Diane. "'Fatal Practices': A Feminist Analysis of Physician-Assisted Suicide and Euthanasia." *Hypatia: A Journal of Feminist Philosophy* 14 (2) (1999): 1–25.

Regan, Donald H. "Suicide and the Failure of Modern Moral Theory." *Suicide and Life-Threatening Behavior* 13 (4) (1983): 276–292.

Sartorius, Rolf. "Coercive Suicide Prevention: A Libertarian Perspective." *Suicide and Life-Threatening Behavior* 13 (4) (1983): 293–303.

Seneca, Lucius Annaeus. *Ad Lucilium Epistulae Morales* (Moral letters to Lucilius). Vol. 2. Letters 70 ("On the Proper Time to Slip the Cable"), 77 ("On Taking One's Own Life"), 78 ("On the Healing Power of the Mind"). Translated by Richard M. Gummere, 57–73, 169–199. New York: G. P. Putnam's Sons, 1920.

Shneidman, Edwin S., ed. *Suicidology: Contemporary Developments*. New York: Grune and Stratton, 1976.

Shneidman, Edwin S., and Norman I. Farberow, eds. *Clues to Suicide*. New York: McGraw-Hill, 1957.

Seneca, Lucius Annaeus. "Letter on Suicide" and "On the Proper Time to Slip the Cable." In *Ad Lucilium Epistulae Morales*. Translated by Richard M. Gummere. London: Heinemann, 1920.

Sprott, Samuel E. *The English Debate on Suicide: From Donne to Hume*. La Salle, IL.: Open Court, 1961.

Steinbock, B. "The Case for Physician-Assisted Suicide: Not (Yet) Proven." *Journal of Medical Ethics* 331 (2005): 235–241.

Thomasma, David C., Thomasine Kimbrough-Kushner, Gerrit K. Kimsma, and Chris Ciesielski-Carlucci, eds. *Asking to Die: Inside the Dutch Debate About Euthanasia*. Dordrecht, Netherlands: Kluwer, 1998.

van der Maas, Paul J., Gerrit van der Wal, et al. "Euthanasia, Physician-Assisted Suicide and Other Medical Practices Involving the End of Life in the Netherlands, 1990–1995." *New England Journal of Medicine* 335 (22) (1996): 1699–1705.

Velleman, J. David. "Against the Right to Die." *Journal of Medicine and Philosophy* 17 (1992): 665–81.

Velleman, J. David. "A Right of Self-Termination? *Ethics* 109 (1999): 606–628.

Wellman, Carl. "A Legal Right to Physician-Assisted Suicide Defended." *Social Theory and Practice* 29 (1) (2003): 19–38.

Werth, James L. Jr. *Contemporary Perspectives on Rational Suicide*. Philadelphia: Taylor and Francis, 1999.

Windt, Peter, "The Concept of Suicide." In *Suicide: The Philosophical Issues*, edited by M. Pabst Battin and David J. Mayo, 39–47. New York: St. Martin's Press, 1980.

Margaret Pabst Battin (2005)

SULZER, JOHANN GEORG

(1720–1779)

Johann Georg Sulzer, the Swiss aesthetician, was born in Winterthur. After studying in Zürich under J. J. Bodmer, he became a tutor in a private home in Magdeburg in 1743. He then went to Berlin, where he became acquainted with Pierre-Louis Moreau de Maupertuis and Leonhard Euler. In 1747 he was appointed professor of mathematics at the Joachimsthaler Gymnasium and in 1763 he moved to the new Ritterakademie. Illness forced him to resign in 1773, but in 1775 he was appointed director of the philosophical section of the Berlin Academy, to which he had been elected in 1750.

Sulzer's *Allgemeine Theorie der Schönen Künste* (General theory of the fine arts) was originally planned as a revision of Jacques Lacombe's *Dictionnaire portatif des*

beaux-arts (1752), but it developed into an original encyclopedia covering both general aesthetics and the theory and history of each of the arts and of literature. The edition of 1796–1798, completed with biographical supplements by Christian Friedrich von Blankenburg, is still the best *summa* of German Enlightenment aesthetics and theory of art, as well as being an original contribution to aesthetics.

Sulzer's style, his psychological interests, and his unsystematic method were typical of the "popular philosophers." Because of his lack of system, and because his ideas are spread through the various articles of his encyclopedia, it is difficult to reduce his views to an organic and systematic whole.

Sulzer's aesthetics was inspired by Gottfried Wilhelm Leibniz, A. G. Baumgarten, G. F. Meier, Moses Mendelssohn, Joseph Addison, Edwards Young, and others. But the psychological character of Sulzer's work is even stronger than that of Baumgarten, Meier, or Mendelssohn. He was the first to find the source of beauty in the perceiving subject only, abandoning every residue of French classicism still present in his German predecessors.

Following Leibniz, Sulzer held that the essence and perfection of the soul consists in its activity of representation. The soul is representing sensibly when it is representing a multiplicity of partial representations taken as a whole. If it is representing every part of a representation as a distinct unit, it is thinking. Sensible representation is more effective than thought, and leads more readily to action. Thus the "lower faculty" of representation of traditional German psychology became more important relative to intellect in Sulzer than in Baumgarten or Meier.

Aesthetics, for Sulzer as for Baumgarten and Meier, was the theory of sensible representation. It explained how to arouse the soul to greater activity. This activity would make sensible representations more lively, and because the activity of representation was intimately connected with the feeling of pleasure, more pleasurable and beautiful.

By studying the psychological constitution of the soul it would be possible to deduce the general rules of the different arts—the more special rules can neither be deduced nor taught. The most important rule concerns the harmony of unity and multiplicity in the beautiful object as it arises out of the representative action of the soul. The object must conform to a spontaneous (*ungezwungen*) order and it must be coherent (*zusammenhängend*).

Sulzer held that beauty is judged by a special feeling—taste—that he sometimes seems to have held to be a function of a faculty different from intellect and the faculty of moral feeling but closely connected with both, particularly with the latter through the moral value of beauty. Taste itself is a transition between thinking and feeling.

Beauty, according to Sulzer, is a product of genius which is the highest stage of the spontaneous representative state of the soul. Genius is a natural force within the soul, and it acts unconsciously in a rational way. It does not, contrary to Baumgarten and Meier, create a new world. Art is an imitation of nature not because it copies nature, but because the artist of genius imitates nature's creative process. He creates nothing outside of nature, but something new within the natural world. In general, art is the expression of a psychological state of man; it imitates human nature in that it expresses nature through the representation of an object.

Sulzer, influenced by Johann Joachim Winckelmann, held that some works of art represent an ideal—that is, they express sensibly a general concept not mixed with anything particular.

In the theory of the individual arts Sulzer's most important contributions were in the aesthetics of music. Music, according to Sulzer, was the expression of passion. Opera, which is a union of all the arts, is the highest form of drama. Besides influencing musical theoreticians, Sulzer's aesthetics influenced Immanuel Kant and Friedrich Schiller; and although Sulzer was attacked by Johann Wolfgang von Goethe in 1772, his work was the foundation of the aesthetics of the *Sturm und Drang.*

Bibliography

WORKS BY SULZER

Versuch einiger moralischer Betrachtungen über die Werke der Natur. Berlin, 1745.

Untersuchungen über die Schönheit der Natur. Berlin, 1750.

Allgemeine Theorie der Schönen Künste. 2 vols. Leipzig: Weidmannschen, 1771–1774; 4th enlarged edition with supplements by C. F. von Blankenburg, 4 vols., Leipzig, 1796–1798.

Vermischte philosophische Schriften. 2 vols. Leipzig, 1773–1781.

WORKS ON SULZER

Dahne, M. *J. G. Sulzer als Pädagoge.* Leipzig, 1902.

Gross, Karl Josef. *Sulzers Allgemeine Theorie der schönen Künste.* Berlin: E. Ebering, 1905.

Heym, L. M. *Darstellung und Kritik der ästhetischen Ansichten J. G. Sulzers.* Leipzig, 1894.

Leo, Johannes. *Zur Entstehungsgeschichte der "Allgemeinen Theorie der schönen Künste" J. G. Sulzers.* Heidelberg, 1906.

Palme, Anton. *J. G. Sulzers Psychologie und die Anfänge der Dreivermögenslehre.* Friedenau, Germany: L. Schultz, 1905.
Tumarkin, Anna. *Der Asthetiker J. G. Sulzers.* Frauenfeld and Leipzig: Huber, 1933.
Wolf, R. *J. G. Sulzer aus Winterthur.* Zürich, 1860.

Giorgio Tonelli (1967)

SUMNER, WILLIAM GRAHAM
(1840–1910)

The American social philosopher, economist, and cultural anthropologist William Graham Sumner was graduated from Yale in 1863 and continued his studies at Geneva, Göttingen, and Oxford, with the aim of entering the Episcopal ministry. He did so in 1867, having returned to America the preceding year. Increasingly, however, this calling conflicted with his wider interests, and when in 1872 he was offered the chair of political and social science at Yale University, he gladly accepted it. He soon gained a considerable reputation as a teacher, publicist, and local politician, but his chief claim to renown derived from his studies in social development, culminating in his masterpiece, *Folkways* (1907).

Two conflicting impulses—polemical and scientific—dominated Sumner's approach to the study of society. It was undoubtedly the polemical impulse that fed the scientific. Dissatisfaction with the reformist dogmatism of his age prompted his search for a scientific basis for his own no less dogmatic advocacy of laissez-faire. In place of "political engineering" based on a facile and sentimental philosophy, Sumner advocated "social evolutionism" free from moralizing preconceptions.

Sumner identified the basic social forces with certain group habits, or "folkways," which, he held, operate on a subconscious level and reflect the spontaneous and the primary needs and interests of a given society, such as hunger, sex, vanity, and fear. These needs and interests, rather than conceptually formed purposes, determine the course of social development. Once the folkways attain persistence and stability, they become reinforced by more conscious processes, such as religious sanctions. Through repeated transmission they assume the status of sociomoral traditions, or "mores." The mores, supported by group authority, then function as the chief agencies of "legitimation"; they determine what shall be deemed right or wrong, or socially acceptable or unacceptable. The mores form the matrix into which an individual is born, and they pervade and control his ways of thinking in all the exigencies of life. The individual becomes critically conscious of his mores only when he comes into contact with another society with different mores or, if he lives in a society at a higher level of civilization, through literature.

Attempts to change a particular set of mores meet with considerable resistance, for they present themselves "as final and unchangeable, because they present answers which are offered as 'the truth'" (*Folkways,* Ch. 2, Sec. 83). Hence, Sumner argued, it was not likely that they could be substantively affected by revolutions or other predetermined acts or changed "by any artifice or device, to a great extent, or suddenly, or in any essential element" (ibid., Sec. 91). Legislation by itself can do little to bring about a transformation of social and moral values. To be truly effective, legislation must grow out of a people's mores; only then is it in keeping with their basic "interests." Nonetheless, Sumner did not deny the significance of legislation, as some commentators have suggested. Indeed, he believed it had a highly educative role, even when it was ineffective in achieving its intended ends. For "it is only in so far as things have been transferred from the mores into laws and positive institutions that there is discussion about them or rationalizing upon them" (ibid., Sec. 80). These unintended consequences, far from being a threat to the established system of mores, constitute a vital component of that system, since it is through such a "rationalizing" process that the mores develop "their own philosophical and ethical generalizations, which are elevated into 'principles' of truth and right" (ibid., Sec. 83).

Although Sumner had little faith in the efficacy of social and economic change produced by state intervention, he was by no means a fatalist or a blind defender of the status quo. A relativist in the tradition of Baron de Montesquieu and Johann Gottfried Herder, a conservative in the tradition of Edmund Burke and Alexander Hamilton, an individualist in the tradition of Thomas Jefferson and Wilhelm von Humboldt, a historicist in the tradition of Friedrich Karl von Savigny and the romanticists, a Spencerian and Darwinist by confession, Sumner believed that man could mold his social life only by paying heed to the "organic" nature of social growth, that he could modify its operative values only "by slow and long continued effort" (ibid., Sec. 91).

Starting from premises not unlike those of Karl Marx, Sumner was, in a sense, a social determinist. However, he recognized the dynamic role of beliefs and the operative value of ideas and, like Marx, he denied their independence from or superiority to material interests.

Material interests constituted both the primary source and the ultimate sanction of social action. Although they drew opposite inferences from their shared premises, and although they were both mistaken in their several dogmatisms and prophecies, Sumner and Marx nevertheless laid bare in an equally fearless manner many features of social development that their generation ignored.

See also Burke, Edmund; Darwinism; Herder, Johann Gottfried; Humboldt, Wilhelm von; Jefferson, Thomas; Marx, Karl; Montesquieu, Baron de; Savigny, Friedrich Karl von; Sociology of Knowledge.

Bibliography

WORKS BY SUMNER

A History of American Currency. New York: Holt, 1874.

Andrew Jackson as a Public Man. Boston: Houghton Mifflin, 1882.

What Social Classes Owe to Each Other. New York: Harper, 1883.

Alexander Hamilton. New York: Dodd Mead, 1890.

The Financier and the Finances of the American Revolution, 2 vols. New York: Dodd Mead, 1891.

A History of Banking in the United States. New York, 1896.

Folkways. Boston: Ginn, 1907.

Earth-Hunger and Other Essays. Edited by A. G. Keller. New Haven, CT: Yale University Press, 1913.

The Forgotten Man, and Other Essays. Edited by A. G. Keller. New Haven, CT: Yale University Press, 1918.

On Liberty, Society, and Politics: The Essential Essays of William Graham Sumner. Edited by Robert C. Bannister. Indianapolis: Liberty Fund, 1992.

WORKS ON SUMNER

Ball, H. V., G. E. Simpson, and K. Ikeda. "Law and Social Change: Sumner Reconsidered." *American Journal of Sociology* 67 (5) (March 1962).

Barnes, H. E. *An Introduction to the History of Sociology.* Chicago: University of Chicago Press, 1948. Ch. 6, "William Graham Sumner; Spencerianism in American Dress."

Curtis, Bruce. *William Graham Sumner.* Boston: Twayne, 1981.

Ellwood, C. A. *A History of Social Philosophy.* New York: Prentice-Hall, 1938. Ch. 28, "William Graham Sumner and Laissez Faire."

Hofstadter, Richard. *Social Darwinism in American Thought.* Boston: Beacon Press, 1955. Ch. 3, "William Graham Sumner."

Keller, A. G. *Reminiscences (Mainly Personal) of William Graham Sumner.* New Haven, CT: Yale University Press, 1933.

Pickens, Donald K. "Scottish Common Sense Philosophy and 'Folkways.'" *Journal of Thought* 22 (1987): 39–44.

Starr, H. E. *William Graham Sumner.* New York: Holt, 1925.

Frederick M. Barnard (1967)
Bibliography updated by Philip Reed (2005)

SUPERVENIENCE

There is supervenience when and only when there cannot be a difference of some sort A (for example, mental) without a difference of some sort B (for example, physical). When there cannot be an A-difference without a B-difference, then but only then A-respects supervene on B-respects. Supervenience claims are thus modal claims. They are claims to the effect that *necessarily*, there is exact similarity in A-respects whenever there is exact similarity in B-respects. So if, for example, mental properties supervene on physical properties, then, necessarily, individuals that are physically indiscernible (exactly alike with respect to every physical property) are mentally indiscernible (exactly alike with respect to every mental property). Thus, A-properties supervene on B-properties just in case how something is with respect to A-properties is a function of how it is with respect to B-properties.

Supervenience has been invoked in nearly every area of analytical philosophy. In addition to its having been claimed that mental properties supervene on physical properties, it has also been claimed that normative properties—moral, aesthetic, epistemic, and so on—supervene on natural properties, that general truths supervene on particular truths, and that modal truths supervene on nonmodal truths. Supervenience, moreover, has been used to distinguish various kinds of internalism and externalism: epistemic, semantic, and mental. And it has been invoked to test claims of reducibility and claims of conceptual analysis, both of which entail supervenience claims. Much of the philosophical work on supervenience itself, as opposed to its philosophical applications, has focused on distinguishing various varieties of supervenience, and examining their pairwise logical relations. But, before turning to the main varieties of supervenience, we can make some central points working just with the idea that there cannot be an A-difference without a B-difference.

1. MODEL FORCE

The term *cannot* in a supervenience claim can express logical impossibility, nomological impossibility (impossibility by virtue of laws of nature), or some other kind of impossibility. If it is logically impossible for there to be an A-difference without a B-difference, then A-properties logically supervene on B-properties; if that is only nomologically impossible, then there is merely nomological supervenience. The property *being a bachelor* logically supervenes on the set of properties {*being unmarried*, *being a man*} because it is logically impossible for indi-

viduals to differ with respect to *being a bachelor* without differing with respect to some property in that set. According to the Wiedemann-Franz law, the electrical conductivity of metals covaries with their thermal conductivity; thus, in metals, electrical-conductivity properties nomologically supervene on thermal conductivity properties (and vice versa).

2. THE RELATA OF THE SUPERVENIENCE RELATION

A difference can be a difference in any respect in which there can be a difference: a difference with respect to what properties something has, in what truths hold, in what conditions obtain, in what events occur, in what laws of nature there are, and so on. The *relata* of the supervenience relation thus seem many and varied. Indeed in "There cannot be an A-difference without a B-difference," A and B may range over nearly all manner of entities. It is often claimed, however, that nonempty sets of properties are the *primary relata* of the supervenience relation: either A-respects and B-respects will be properties in some nonempty sets of properties A and B, or else A-respects will supervene on B-respects in virtue of A-properties supervening on B-properties (Kim 1984). This view requires an "abundant" (as opposed to a "sparse") conception of properties, according to which properties "may be as extrinsic, as gruesomely gerrymandered, as miscellaneously disjunctive, as you please. … [They] far outrun the predicates of any language we could possibly possess. … In fact, the properties are as abundant as the sets themselves, because for any set whatever, there is the property of belong to that set" (Lewis 1986, 59–60).

Indeed, on this conception, there are even necessarily uninstantiated properties such as *being an electron and not being an electron*, and so properties are not always ways things might be. In the literature on supervenience, an abundance of properties is often assumed, and such will be assumed in this essay. But whether there is supervenience does not turn on whether there are abundant properties, or, if nominalists are right, even on whether there are properties at all. A nominalist could maintain that what A-predicates are true of something supervenes on what B-predicates are true of it. Nor does it turn on whether there is some uniform category of being the members of which are the *primary relata* of the supervenience relation. It does not even turn on whether there is a relation of supervenience in anything other than a merely pleonastic sense: talk of A bearing the supervenience relation to B might be taken to be just a way of saying that there cannot be an A-difference without a B-difference. What matters is that there be true statements of the form, "There cannot be an A-difference without a B-difference." And such there are in abundance, including many of philosophical interest.

3. LOGICAL PROPERTIES OF THE SUPERVENIENCE RELATION

Supervenience is reflexive, transitive, and nonsymmetric. Trivially, it holds when A = B and so is reflexive. It is also transitive, because if there cannot be an A-difference without a B-difference, and cannot be a B-difference without a C-difference, then there cannot be an A-difference without a C-difference. However, it is neither symmetric nor asymmetric, and so is nonsymmetric. Every reflexive case of supervenience is trivially symmetric. But, for instance, *being a bachelor* asymmetrically supervenes on {*being unmarried*, *being a man*}. James is a man and Vanessa is not, and so they differ with respect to B-properties. But since James is married, they are exactly alike with respect to *being a bachelor*: neither of them has that property.

4. SUPERVENIENCE AND ENTAILMENT

A notion of property entailment can be defined as follows: property P entails property Q if and only if it is logically necessary that whatever has P has Q. Supervenience shares with entailment the properties of being reflexive, transitive, and nonsymmetric. Property supervenience, however, is neither necessary nor sufficient for property entailment. The property *being a brother* entails the property *being a sibling*. But *being a sibling* does not supervene on *being a brother*. Thus, suppose that Sarah has a sister and that Jack is an only child. Then Sarah is a sibling and Jack is not, though neither is a brother. Property entailment thus does not suffice for supervenience.

It is often claimed in the literature that *logical* supervenience suffices for entailment (see, for example, Chalmers 1996). But that is not in general true. If A = {P&Q} and B = {P, Q}, then the A-property logically supervenes on B-properties, but no B-property entails the A- property. Indeed, every property F will supervene with logical necessity on its complement, not-F: Two things cannot differ with respect to F without differing with respect to not-F (and vice versa). But of course *being F* does not entail *being not-F* (McLaughlin 1995).

There seem, moreover, to be philosophically interesting cases of logical supervenience without entailment. Particular truths do not entail general truths. But general truths (arguably) supervene on particular truths (Skyrms 1981, Lewis 1986a). Bertrand Russell correctly noted:

"you cannot ever arrive at a general fact by [deductive] inference from particular facts, however, numerous" (1918/1992, p. 235, quoted in Bricker 2005). He concluded from this that "you must admit general facts as distinct from and over and above particular facts" (1981/1992, p. 236). If, however, general facts logically supervene on particular facts, then there is a sense in which that is not so, for once all the particular facts of a world are fixed, the general facts are fixed as well. A compelling case has been made that general facts logically supervene on particular facts, despite not being entailed by them (Bricker 2005).

5. SUPERVENIENCE AND ONTOLOGICAL PRIORITY

Many of the most interesting cases of supervenience are ones in which the subvenient factors are *ontologically prior* to the supervenient factors. Supervenience itself, however, is not an ontological priority relation. Ontological priority is irreflexive and asymmetric: Nothing can be ontologically prior to itself or be ontologically prior to something that is ontologically prior to it. But supervenience is reflexive and not asymmetric. Supervenience claims do not, in general, entail "in virtue of" claims. Every property supervenes on its complement, but of course nothing has a property F in virtue of having its complement not-F because nothing has both F and not-F (at least at the same time). Further, properties that everything necessarily has, and ones that nothing could possibly have, supervene on any property whatsoever. The necessary property *being an electron or not an electron* trivially supervenes on the property *being an antique*; and the necessarily uninstantiated property *being an electron and not being an electron* does well. The reason is that no two things can differ with respect to either such noncontingent property; and so, trivially, for any property, no two things can differ with respect to them without differing with respect to it. But there is no ontological priority in such cases. (McLaughlin 1995)

6. SUPERDUPER VENIENCE

Supervenience is just the relation of functional dependence: A-properties supervene on B-properties just in case how something is with respect to A-properties is a function of how it is with respect to B-properties. Given that, when A-properties supervene on B-properties, we expect there to be some explanation of why that is so. In the case of logical supervenience, the explanation might be that A-properties are necessary properties or that they are properties nothing could have. Or the explanation might be that A-properties are identical with B properties. Or the explanation might be that A-properties are determinables of B-properties and B-properties are all the determinates of A-properties, as *being colored* is a determinable of all the shades of color (*being red*, and so on), and they are determinates of *being colored*. And in the case of merely nomological supervenience, the explanation will appeal to a law of nature. (This list of possible explanations is not intended to be exhaustive.) When a supervenience relation is explainable, there is "superdupervenience" (Horgan 1993). Appeals to in principle unexplainable supervenience—supervenience without the possibility of superdupervience—would arguably be mystery-mongering.

7. SUPERVIENCE, CONCEPTUAL ANALYSIS, AND REDUCTION

Although logical supervenience does not suffice for conceptual analysis, the latter requires the former: if A-factors can be conceptually analyzed in terms of B-factors, then A-factors logically supervene on B-factors. Supervenience is thus useful in testing claims that a certain a kind of conceptual analysis is possible. According to a simple causal theory of perceptual knowledge, a subject's perceptual knowledge that P can be analyzed as P's bearing an appropriate causal connection to the subject's perceptual belief that P. To test the claim, one need not await a specific proposal as to what kind of causal connection is appropriate. For such a conceptual analysis is possible only if two believers that P cannot differ with respect to perceptually knowing that P without differing with respect to how the fact that P is causally connected to their belief that P. This supervenience thesis is open to refutation by a single counterexample. The well-known "fake barn country" case (Goldman 1976) yields a putative counterexample to this thesis. Thus, the claim that a certain kind of conceptual analysis is possible can be refuted by appeal to a false implied supervenience thesis (or, FIST). Claims that certain kinds of reductions are possible can be similarly tested by their implied supervenience theses. (McLaughlin 1995)

8. INDIVIDUAL/GLOBAL SUPERVENIENCE

There is a distinction between *individual* supervenience and *global* supervenience. The former concerns differences in individuals; the latter concern differences in possible worlds. The claim that individuals cannot differ with respect to their moral properties without differing with respect to their natural properties (Hare 1952) is an indi-

vidual supervenience thesis. The claim that possible worlds cannot differ with respect to what general truth hold in them without differing with respect to what particular truths hold in them is a global supervenience thesis.

9. STRONG/WEAK INDIVIDUAL SUPERVENIENCE

Two nonequivalent kinds of individual supervenience have been formulated as follows (see Kim 1987):

POSSIBLE-WORLDS WEAK INDIVIDUAL SUPERVENIENCE. A-properties weakly supervene on B-properties if and only if in any possible world *w*, B-indiscernible individuals in *w* are A-indiscernible in *w*.

POSSIBLE-WORLD STRONG INDIVIDUAL SUPERVENIENCE. A-properties strongly supervene on B-properties if and only if for any possible worlds *w* and w^*, and any individuals *x* and *y*, if *x* in *w* is B-indiscernible from *y* in w^*, then *x* in *w* is A-indiscernible from *y* in w^*.

The possible worlds quantified over might be all logically possible worlds or only all nomologically possible worlds (and so on); thus, weak and strong supervenience relations can have different modal strengths. As the names suggest, strong supervenience is stronger than weak supervenience (*modulo* sameness of modality). When the range of worlds is the same, strong supervenience of A-properties on B-properties entails weak supervenience of A-properties on B-properties, but the latter does not in general entail the former. Notions of weak and strong individual supervenience have also been formulated as follows, using the modal operator *necessarily* rather than quantification over possible worlds (Kim 1984).

OPERATOR-WEAK INDIVIDUAL SUPERVENIENCE. A-properties weakly supervene on B-properties if and only if necessarily, for any A-property *F*, if something has *F*, then there is a B-property G such that it has G, and whatever has G has *F*.

OPERATOR-STRONG INDIVIDUAL SUPERVENIENCE. A-properties strongly supervene on B-properties if and only if necessarily, for any *A*-property *F*, if something has *F*, then there is a B-property G such that it has G, and necessarily whatever has G has *F*.

The strong version is formulated exactly like the weak version except that it contains one more necessity operator. The two modal operators in the strong case can be the same or different. When all of the modal operators are the same, strong supervenience entails weak supervenience, but the latter does not in general entail the former.

If necessity is understood as universal quantification over possible worlds, then operator-weak supervenience entails world-weak supervenience, and operator-strong supervenience entails world-strong supervenience. However, the converse entailments do not hold in general. The operator definitions go beyond the idea that B-indiscernible individuals must be A-indiscernible. Operator-strong supervenience with logical necessity guarantees that every A-property is entailed by a B-property. And both operator-weak supervenience and operator-strong supervenience entail that if something has an A-property, then it has some B-property. Neither world-weak supervenience nor world-strong supervenience has that entailment, and so world-strong supervenience fails even to entail operator-weak supervenience (McLaughlin 1995). The property *being a bachelor* fails to even operator-weakly supervene on {*being unmarried*, *being a man*}, even though the former world-strongly supervenes on the later. The weak and strong operator definitions are, however, equivalent to the corresponding world-definitions in the special case of nonempty sets of properties closed under the Boolean operations of complementation and conjunction and/or disjunction, and ones involving quantification (Kim 1987). (The qualifiers *world* and *operator* will now be dropped.)

10. SUPERVENIENCE AND INTERNALISM/EXTERNALISM DISTINCTIONS

Individual supervenience has proved useful for formulating various kinds of internalism/externalism distinctions. For example, according to internalists about mental content, what content a mental state has will strongly supervene on intrinsic properties of the subject of the mental state. Content externalists deny such supervenience, and indeed typically deny there is even weak supervenience: they typically hold that two subjects within a possible world can be intrinsic duplicates while being in mental states with different contents. (Twin-Earth cases [Putnam 1975] are invoked in would-be arguments by appeal to FISTs against internalist theories of content.) Similarly, an internalist about epistemic justification asserts that whether a belief is justified strongly supervenes on what mental states the subject is in. Epistemic externalists deny that, and indeed deny that whether a belief is epistemically justified even weakly supervenes on what mental

states the subject is in. Moreover, supervenience has been employed to capture the traditional distinction between internal and external relations (Lewis 1986a): internal relations (such as *being taller than*) strongly supervene on the intrinsic natures of its *relata*, whereas external relations (such as *being three kilometers from*) fail to even weakly supervene on the intrinsic natures of its *relata*.

11. WEAK SUPERVENIENCE WITHOUT STRONG SUPERVENIENCE

There can be weak supervenience without strong supervenience. But when this is the case, we expect an explanation of why weak supervenience holds that does not entail that strong supervenience holds as well. In any possible world, if two individuals assert exactly the same propositions, then they are exactly alike in having asserted a true proposition: The one will have asserted a true proposition if and only if the other did. The explanation is that any proposition will have a unique truth value relative to a world. But since contingent propositions are true in some worlds but not in others, strong supervenience fails in the case in question. It has been claimed that, although moral properties weakly supervene on natural properties, they do not strongly supervene on them (Hare 1952). And it has been claimed that, although mental properties weakly supervene on physical properties, they do not strong supervene on them (Davidson 1985). Defense of these claims requires an explanation of why weak supervenience holds despite the failure of strong supervenience. Although attempts have been made to provide such an explanation in the moral case (Blackburn 1993), there has been no attempt in the mental case. Many philosophers doubt such an explanation is possible in the mental case.

12. GLOBAL SUPERVENIENCE

Global supervenience has been invoked in the formulation of various philosophical doctrines (see, for example, Horgan 1982, 1984; Haugeland 1982; Post 1987). David Lewis's (1986a, x) doctrine of Humean Supervenience, according to which everything supervenes on the pattern of perfectly natural qualitative properties across space-times points, is a global supervenience thesis. Although Donald Davidson (1970) proposed a weak individual supervenience thesis to characterize the dependency of mental properties on physical properties, several attempts have been made to characterize physicalism as a global supervenience thesis (Lewis 1983, Chalmers 1996, Jackson 1996).

For example, Frank Jackson has proposed the following formulation: Any possible world that is a minimal physical duplicate of our world is a duplicate *simpliciter* of it (1998, p. 12). A physical duplicate of our world is any world exactly like it in every physical respect—with respect to its worldwide pattern of distribution of physical properties and relations, its physical laws, and so on. A minimal physical duplicate is any physical duplicate that contains nothing other than what is metaphysically necessary to be a physical duplicate. It is controversial whether this thesis suffices for physicalism; unlike physicalism, it seems compatible with the existence of a necessarily existing God. But even if it does not suffice, if physicalism requires it, then it earns its keep. A substantive condition of adequacy on physicalism would be that it explain why the supervenience thesis is true. And physicalism itself would rendered testable, even in the absence of a fully adequate formulation. Given that we are phenomenally conscious, if, as some philosophers (Chalmers 1996) maintain, a "zombie world" is possible—a world that is a minimal physical duplicate of our world but entirely devoid of phenomenal consciousness—then physicalism is false. Of course, the success of this would-be refutation by appeal to a FIST turns on the controversial issue of whether a zombie world is indeed possible.

Global property supervenience has often been formulated as follows:

GLOBAL SUPERVENIENCE. A globally supervenes on B if and only if, for any possible worlds w_1 and w_2, if w_1 and w_2, have exactly the same worldwide pattern of distribution of B-properties, then w_1 and w_2 have exactly the same worldwide pattern of distribution of A-properties.

It is now usually acknowledged that the notion of a worldwide pattern of distribution of properties should be understood in terms of a kind of property-preserving isomorphism between worlds as follows (McLaughlin 1996, 1997; Stalnaker 1996):

> An isomorphism I between the inhabitants of any worlds w_1 and w_2 preserves F-properties if and only if, for any x in w_1, x has an F-property in w_1 just in case the image of x under I (the individual to which I maps x) has P in w_2.

13. WEAK, INTERMEDIATE, AND STRONG GLOBAL SUPERVENIENCE

A variety of different kinds of global supervenience has been formulated:

A-properties weakly globally supervene on B-properties if and only if, for any worlds w_1 and w_2, if there is a B-preserving isomorphism between w_1 and w_2, then there

is an A-preserving isomorphism between them (McLaughlin 1996, 1997; Stalnaker 1996; Sider 1999).

A-properties intermediately globally supervene on B-properties if and only if, for any worlds w_1 and w_2, if there is a B-preserving isomorphism between w_1 and w_2, then there is at least one isomorphism between them that is both A-and-B-preserving (Shagrir 2002, Bennett 2004).

A-properties strongly globally supervene on B-properties if and only if, for any worlds w_1 and w_2, every B-preserving isomorphism between w_1 and w_2 is an A-preserving isomorphism between them. (McLaughlin 1996, 1997; Stalnaker 1996; Sider 1999). Strong global supervenience entails intermediate global supervenience, which entails weak global supervenience. But the converse entailments all fail to hold in general.

There seem to be no cases of philosophical interest in which weak global supervenience holds, but both strong and intermediate global supervenience fail to hold. In some cases of interest, however, intermediate global supervenience holds, even though strong global supervenience may fail to hold. Many philosophers maintain that two numerically distinct objects can have the same spatiotemporal location and so be spatiotemporally coincident. A frequently cited would-be example is a clay statue and the lump of clay that makes it up. Even if they are spatiotemporally coincident throughout their existence—created at the same time and destroyed at the same time—they nevertheless have different modal properties: for example, the lump could survive being squashed, while the statue could not. But they have exactly the same categorical properties (mass, size, shape, and so on). If the statue is indeed not the lump, then the statue's modal properties will neither individually strongly nor individually weakly supervene on its categorical properties. (Multiple-domain individual supervenience will hold, however [see Kim 1988 and Zimmerman 1995].) And modal properties will fail to strongly globally supervene on categorical properties. But weak global supervenience (Sider 1999) and intermediate global supervenience (Bennett 2004) will both hold. An appeal to intermediate global supervenience would not by itself, however, solve "the grounding problem," the problem of how individuals with exactly the same categorical properties can differ in their modal properties (Bennett 2004). A solution to the grounding problem would have to explain why intermediate global supervenience holds and do so in a way that does not entail that coincident objects are identical.

14. SOME EQUIVALANCIES

The plethora of technical definitions of kinds of supervenience gives the appearance of more variety than there is. Strong individual supervenience entails strong global supervenience (Kim 1984), but strong global supervenience does not in general entail strong individual supervenience (Paull and Sider 1992). Nevertheless, strong individual supervenience and strong global supervenience are equivalent in cases in which the base set of properties B is closed under Boolean operations and ones involving quantification and identity (Stalnaker 1996). Strong individual supervenience is also equivalent to strong global supervenience in cases in which A and B are sets of intrinsic properties (Shagrir 2002, Bennett 2004). It has, moreover, been compellingly argued that in cases in which A and B are sets of intrinsic properties, weak and strong individual supervenience are equivalent as well. Weak individual supervenience, strong individual supervenience, and strong global supervenience are equivalent for sets of intrinsic properties.

See also Davidson, Donald; Knowledge and Modality; Lewis, David; Modality, Philosophy and Metaphysics of; Physicalism; Reduction; Russell, Bertrand Arthur William.

Bibliography

Bennett, K. "Global Supervenience and Dependence." *Philosophy and Phenomenological Research* 68 (2004): 510–529.

Blackburn, S. "Supervenience Revisited." Reprinted in Blackburn ed. *Essays in Quasi–Realism*. Oxford: Oxford University Press, 1993.

Bricker, P. "The Relation Between General and Particular: Entailment vs. Supervenience." In *Oxford Studies in Metaphysics*. Vol. 2, edited by D. Zimmerman. Oxford: Oxford University Press, 2005.

Chalmers, D. *The Conscious Mind*. New York: Oxford University Press, 1996.

Davidson, D. "Mental Events." In *Experience and Theory*, edited by L. Foster and J. W. Swanson. Amhurst: University of Massachusetts, (1970): 79–101.

Davidson, D. "Replies to Essays X–XIII." In *Essays on Davidson: Actions and Events*, edited by B. Vermazen and M.B. Hintikka. New York: Oxford University Press, 1985.

Goldman, A. "Discrimination and Perceptual Knowledge." *Journal of Philosophy* 18 (1976): 771–91.

Hare, R. M. *The Language of Morals*. Oxford: Clarendon Press, 1952.

Haugeland, J. "Weak Supervenience." *American Philosophical Quarterly* 19 (1982): 93–103.

Horgan, T. "Supervenience and Microphysics." *Pacific Philosophical Quarterly* 63 (1982): 29–43.

Horgan, T. "Cosmic Hermenutics." *Southern Journal of Philosophy* 22: *The Spindel Conference on Supervenience Supplement* 22 (1984): 19–38.

Horgan, T. "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind* 102 (1993): 555–86.

Jackson, F. *From Ethics to Metaphysics*. Oxford: Oxford University Press, 1996.

Kim, J. "Concepts of Supervenience." *Philosophy and Phenomenological Research* 45 (1984): 153–76. Reprinted in *Supervenience and Mind*, edited by J. Kim, 33–52. Cambridge, U.K.: Cambridge University Press, 1993.

Kim, J. "Strong and 'Global' Supervenience Revisited." *Philosophy and Phenomenological Research* 48 (1987): 315–26. Reprinted in Kim ed.1993, 79–91.

Kim, J., ed. *Supervenience and Mind*. Cambridge, U.K.: Cambridge University Press, 1993.

Kim, J. "Supervenience for Multiple Domains." *Philosophical Topics* 16 (1988): 129–50. Reprinted in Kim ed. 1993, 79–91.

Kim, J. "Supervenience as a Philosophical Concepts." *Metaphilosophy* 21 (1990): 1–27. Reprinted in Kim 1993 ed., 109–130.

Lewis, D. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61 (1983): 343–77.

Lewis, D. *On the Plurality of Worlds*. Oxford: Oxford University Press, 1986b.

Lewis, D., ed. *Philosophical Papers*, vol. II. Oxford: Oxford University Press, 1986a.

Marsh, R. C. *Logic and Knowledge*. London: Routledge, 1992.

McLaughlin, B. P. "Supervenience." In *Macmillian Encyclopedia of Philosophy, Supplementary Volume* 1996, edited by D. Borchert, 558–60.

McLaughlin, B. P. "Supervenience, Vagueness, and Determination." *Philosophical Perspectives* 11 (1997): 209–230.

McLaughlin, B. P. "Varieties of Supervenience." In *Supervenience: New Essays*, edited by E. Savellos and U. Yalcin. Cambridge, U.K.: Cambridge University Press, 1995.

Moyer, M. "The Strengths and Weakness of Weak and Strong Supervenience." Forthcoming.

Paull, R. C., and T. R. Sider. "In Defense of Global Supervenience." *Philosophy and Phenomenological Research* 52 (1992): 833–54.

Post, J. *The Faces of Existence*. Ithaca, NY: Cornell University Press, 1987.

Putnam, H. "The Meaning of 'Meaning.'" *Minnesota Studies in the Philosophy of Science* 7 (1975): 215–71.

Russell, B. *The Philosophy of Logical Atomism* (1918). Reprinted in Marsh ed. (1992): 177–281.

Shagrir, O. "Global Supervenience, Coincident Entities and Anti–individualism." *Philosophical Studies* 109 (2002): 171–196.

Sider, T. R. "Global Supervenience and Identity Across Times and Worlds." *Philosophy and Phenomenological Research* 59 (1999): 913–37.

Skyrms, B. "Tractarian Nominalism." *Philosophical Studies* 40 (1981): 199–206.

Stalnaker, R. "Varieties of Supervenience." *Philosophical Perspectives* 10 (1996): 221–41.

Zimmerman, D. "Theories of Masses and the Problems of Constitution." *Philosophical Review* 104 (1995): 53–110.

Brian P. McLaughlin (1996, 2005)

SUPPES, PATRICK
(1922–)

Patrick Suppes is an American philosopher and scientist. Born in Tulsa, Oklahoma, Suppes was educated at Tulsa Central High School and pursued undergraduate studies at the University of Oklahoma, the University of Tulsa, and the University of Chicago, with particular emphasis on physics and meteorology. He graduated from Chicago in 1943, then spent 1944 to 1946 in the United States Army Air Force. During graduate studies in philosophy at Columbia University in the years 1947 through 1950 Suppes studied with Ernest Nagel, combining courses in philosophy with further work in physics and mathematics. Somewhat surprisingly in the light of his later research in psychology, he did not study that subject at either the undergraduate or the graduate level.

Suppes received his Ph.D. in philosophy from Columbia University in 1950. His entire academic career has been spent at Stanford University, where he began as an assistant professor of philosophy in 1950. He subsequently held concurrent positions in the departments of psychology, statistics, and education, and from 1959 until his retirement directed the Institute for Mathematical Studies in the Social Sciences at Stanford, a research center he co-founded with the economist Kenneth Arrow. He has been a pioneer in computer-assisted education and in 1967, with the psychologist Richard Atkinson, founded a successful company, Computer Curriculum Corporation. He has received numerous honors during his career, culminating with the award of the National Medal of Science in 1990. Suppes retired from Stanford in 1992, but he has continued an active research program, including work on robotics and experimental work on the neural bases of language processing.

WORK

Suppes's work is unusual in its combination of significant scientific research with rigorous philosophical analysis, in its scope, and in its constructive orientation. It spans philosophy, psychology, probability and statistics, education, and computer science. The focus here is on his contributions to the philosophy of science, although his positions in that area are always deeply rooted in his scientific work. Throughout his career, Suppes has emphasized the

pluralistic and complex nature of actual, rather than idealized, scientific methods. For example, as early as 1962 he argued that there was a hierarchy of models between data and theory, anticipating in certain ways the later philosophical literature emphasizing the importance of scientific models. Many of Suppes's principal philosophical contributions have been in the area of formal methods, both as a way of injecting precision into philosophical questions and as an effective set of tools for producing answers to those questions.

At the heart of his philosophical work lies the semantic account of theories, of which Suppes—building on joint work with J. C. C. McKinsey and employing Alfred Tarski's work on formal models—was the primary developer. The semantic account, which is the chief rival to the syntactic account of theories, also served as the foundation for the later structuralist approach to theories. In Suppes's version of the semantic account, a theory is identified with a class of set-theoretical structures—models in the sense of mathematical logic. Thus, rather than a theory being a set of sentences or propositions represented in first order logic—the identification made by the logical empiricists and their successors, particularly Quine—a theory in Suppes's sense abstracts from a particular linguistic representation and focuses instead on what makes that theory true, using the full apparatus of set theory. Thus, Newton's, Hamilton's, and Lagrange's versions of classical mechanics are simply different linguistic representations of the same underlying semantic theory. This powerful foundational apparatus allows for an easy representation of the kind of mathematics needed for scientific theories—in contrast to first order logic, which is an apparatus that is too weak to capture large parts of standard mathematics. The apparatus employed in the semantic approach is especially useful in such areas as measurement theory, a subject to which Suppes has made contributions of permanent value. The semantic approach also leads naturally to a focus on axiomatized theories because this allows the content of the theories to be fully captured in an explicit, and often recursive, set of constraints. This emphasis on formal methods follows naturally from Suppes's view that there are only practical, rather than theoretical, differences between representations of mathematical theories and representations of scientific theories.

A key concept in Suppes's work is that of a representation theorem. A representation theorem for a set of models M asserts that there exists a subset R of M such that for any model m in M there is a model r in R that is isomorphic to m. Such representation theorems play a central role in measurement theory when R is a class of numerical measurement structures and M is the class of empirical models upon which measurement procedures are to be placed. Philosophically, the emphasis on identity up to isomorphism (or, more generally, homomorphism) entails that the abstract structure of systems is captured, rather than any intrinsic features that are unique to the system.

Suppes's other important contributions include his monograph on probabilistic causality that, together with Reichenbach's earlier treatment, began this distinctive and widely discussed approach to causation; his pioneering work on the identification of aural and visual language recognition using electroencephalographic brain data; his work exploring variant probability spaces in quantum theory; an exploration of Bayesian inference; and the role of invariances in classical and relativistic physics. As the culmination of developing a number of stochastic models of learning, Suppes proved in 1969 that any finite automaton could be represented by a stimulus-response learning model, a result of importance to controversies about the nature of language learning. Together with the work on theory structure and measurement theory, these form an impressive and permanent set of contributions to the philosophy of science.

Suppes's publications are demanding but always lucid; they invariably repay careful study. Inevitably, they only partially convey his considerable influence as a teacher and professional colleague, an influence grounded in equal parts of rigor, style, humor, and clarity. A comprehensive and detailed presentation of his mature views is given in *Representation and Invariance of Scientific Structures* (2002).

See also Causation: Philosophy of Science; First–Order Logic; Mathematics, Foundations of; Semantics; Structuralism, Mathematical.

Bibliography

WORKS BY SUPPES

"Models of Data." In *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*. Edited by E. Nagel, P. Suppes, and A. Tarski, 252–261. Stanford, CA: Stanford University Press, 1962.

"Stimulus-Response Theory of Finite Automata." *Journal of Mathematical Psychology* 6 (1969): 327–355.

A Probabilistic Theory of Causality. Amsterdam: North-Holland, 1970.

Foundations of Measurement. vols. 1–3. With D. H. Krantz, R. D. Luce, and A. Tversky. New York: Academic Press, 1971, 1989, 1990.

Stanford Faculty Web Pages. A complete bibliography of Patrick Suppes's publications through 2000. Available from http://www.stanford.edu/~psuppes.

Representation and Invariance of Scientific Structures. Stanford: CSLI Publications, 2002.

WORKS ABOUT SUPPES

Bogdan, R., ed. *Patrick Suppes*. Dordrecht, Netherlands: D. Reidel, 1979. Contains a personal and intellectual autobiography up to 1978.

Humphreys, P., ed. *Patrick Suppes: Scientific Philosopher*. Vols. 1–3. Dordrecht, Netherlands: Kluwer, 1994. Contains articles on most aspects of Suppes's research and his replies to each article.

Paul Humphreys (2005)

SUSO, HEINRICH
(1295/1300–1366)

Heinrich Suso, the Rhineland mystic, was born at Constance and early entered the Dominican order. A mystical experience at the age of eighteen set him on the path of asceticism, but a later one, between 1335 and 1340, led him to abandon self-mortification and to embark on an active career as preacher and spiritual adviser. As a result of attacks on some of his teachings and on his personal character, he was transferred to Ulm in 1348.

During his period of studies in Cologne, Suso had come into contact with Johannes Tauler and also came under the influence of Meister Eckhart. Indeed, in *Das Büchlein der Wahrheit* (*The Little Book of Truth,* c. 1327) he was bold enough to defend Eckhart against the doctrinal charges leveled against him, setting Eckhart's disputed doctrines alongside other quite orthodox statements made by him and providing interpretations that did not entail pantheistic conclusions.

Although Suso made use of the Eckhartian-sounding distinction between the undifferentiated Godhead and God as manifested in the persons of the Trinity, he did not hold that there was an ontological distinction within the divine Being. Rather, he held that the distinction was an intellectual one, made from the human point of view and dependent on our mode of trying to understand God's nature. Although Suso also used extreme Neoplatonic language in speaking of God as Nothing, he made it clear that this was simply to say that, because of God's complete simplicity, we cannot ascribe predicates to him in the sense in which they are applied to creatures. Suso went on to try to explain the contrasting and paradoxical multiplicity of God's nature, as exhibited in the Trinity, by the usual concept of eternal procession. Like his doctrine that the distinction between the Godhead and God as the Trinity is not an ontological one, the notion of procession should be taken in a way that does not imply the priority of God considered as a simple Nothing over God considered as the Father, Son, and Holy Spirit. Thus, Suso drew a strong distinction between the procession occurring within the divine Being and the creation of the world. The latter is a free act of God, and creatures owe their being to him; thus God is ontologically prior to the world. On the other hand, the internal dynamics of the Trinity are a perfect and eternal feature of God's life.

The idea of God as Nothing reflected, as did similar doctrines held by other medieval mystics, not only a view about predication in theology but also about the mystical experience itself. Thus, Suso characteristically spoke of that state in which the contemplative is taken out of himself and is made calm in the ground of the eternal Nothing. The fact that the contemplative experience is free from images and discursive thought is a sufficient explanation of the negative language used. Suso generally avoided the suggestion that the soul is merged with the Godhead and described the union as one of wills in which, however, the soul retains its identity. Nevertheless, there were times when he, orthodox as he generally was and wished to be, spoke of a substantial identification with the Godhead. Some explanation of this apparent inconsistency is found in his assertion that in the mystical state the individual is no longer aware of his own identity. It is afterward, and through going beyond a merely phenomenological description of the experience, that the mystic is able to give what he considers to be the correct theological account of it.

Suso's chief works were the autobiographical *Das Buch von dem Diener* (*The Life of the Servant*); the *Horologium Sapientiae,* which also occurs in a somewhat different German version as *Das Büchlein der ewigen Weisheit* (*The Little Book of the Eternal Wisdom*); and *Das Büchlein der Wahrheit* (*The Little Book of Truth*). The second of these, which is a dialogue about and meditation on the sufferings of Christ, attained a wide circulation, almost rivaling that of Thomas à Kempis's *The Imitation of Christ.* Because of the degree of openness in the description of his inner life, Suso's writings constitute a valuable source for the study of Christian mysticism.

Bibliography

The German writings can be found in Karl Bihlmayer, *Heinrich Suso: deutsche Schriften* (Stuttgart, 1907). The *Horologium Sapientiae* was edited by J. Strange (Cologne, 1861). Useful translations are J. M. Clark, *Little Book of Eternal Wisdom and Little Book of Truth* (London: Faber and

Faber, 1953) and *The Life of the Servant* (London, 1952). For a general introduction, see J. M. Clark, *The Great German Mystics* (Oxford: Blackwell, 1949), Ch. 4.

Ninian Smart (1967)

SWEDENBORG, EMANUEL

(1688–1772)

Emanuel Swedenborg, the scientist, biblical scholar, and mystic, was a member of a famous Swedish family of clergymen and scholars; his father was a prominent bishop and a prolific writer. Swedenborg studied the classics and Cartesian philosophy at Uppsala and became interested in mathematics and natural science. In 1710 he went abroad, spending most of the next five years in England, where he learned the Newtonian theories and developed a modern scientific outlook. After his return to Sweden in 1715, Swedenborg was appointed an assessor in the College of Mines by Charles XII. He held this office until 1747, when he resigned in order to devote his time to the interpretation of the Scriptures.

PHILOSOPHY OF NATURE

Swedenborg's many writings are characterized by great scholarship and by a fervent search for a synthesis of ancient wisdom and modern experience, empirical science, rationalistic philosophy, and Christian revelation. After some minor treatises on geological and cosmological problems, he published his first important work in 1734, *Opera Philosophica et Mineralia* (3 vols., Dresden and Leipzig); the first part of this work, *Principia Rerum Naturalium*, contains his philosophy of nature. Here Swedenborg used the concept of the mathematical point, which he described as coming into existence by motion from the Infinite. This point forms a nexus, or connection, between the Infinite and the finite world, and by its motion it creates aggregates of elements that build up the Cartesian vortexes, which are interpreted as the fundamentals of nature. The original motion in the Infinite, however, is not a mechanical motion but a kind of Leibnizian conatus, a motive force in nature that corresponds to will in human minds. In the first point there is a corresponding tendency, which transmits itself to the subsequent aggregates in this great chain of being.

The outlines of Swedenborg's natural philosophy are derived from René Descartes, Gottfried Wilhelm Leibniz, and other rationalists, but in the *Principia* Swedenborg was also inspired by empirical philosophy, especially that of John Locke. A similar English influence can be observed in Swedenborg's cosmology, which is set forth in the *Principia* and in a short hexaemeron titled *De Cultu et Amore Dei* (London, 1745). In these works Swedenborg presents a nebular hypothesis according to which the planets are formed of solar matter. It has been maintained that the planet theory of Immanuel Kant and Pierre Simon de Laplace might have been derived from Swedenborg via the comte de Buffon, but most probably the similarities between Swedenborg and Buffon depend on their common source of inspiration, Thomas Burnet's *Telluris Theoria Sacra* (The sacred theory of the earth; 1681). This treatise was widely known (even Samuel Taylor Coleridge admired it), and there is no doubt that it guided Swedenborg in his cosmology. Swedenborg's cosmology was essentially mechanistic, but like the great speculative philosophers of the seventeenth century, he attempted very early to find a theory that could combine these scientific hypotheses with Christianity.

Together with this mechanistic outlook there are several elements in Swedenborg's philosophy of nature that anticipate the organic theories set forth in his anatomic and psychological works. These works include *Oeconomia Regni Animalis* (2 vols., London and Amsterdam, 1740–1741), *Regnum Animale* (3 vols., The Hague and London, 1744–1745), and many other posthumously published treatises on the animal kingdom. The main problem concerning Swedenborg here is the relationship between soul and body. Since he was not satisfied by any of the current philosophical hypotheses, he turned to the study of contemporary microanatomy and physiology. His own theory, which is sometimes called the *harmonia constabilita* (coestablished harmony), is similar to Leibniz's theory of preestablished harmony. The two models are not identical, however, since there is a component of successive growing in Swedenborg's notion that is missing in the preestablished harmony.

In his physiological research Swedenborg starts with the study of the blood, which in its relation to the organization of the human body corresponds in some important ways to the role of the mathematical point as a nexus between the spiritual and the physical worlds. Swedenborg distinguishes several degrees of purity in the blood, with the highest degree corresponding to the Cartesian spiritous fluid. This fluid functions both as a concrete communication line between soul and body and as an abstract principle, a formative force of the body (*vis formatrix*). Swedenborg combined this concept of life force with Aristotle's concept of form and developed a teleological system very much like Leibniz's monadology.

DOCTRINE OF SERIES AND DEGREES. Swedenborg's system may be called the doctrine of series and degrees. The degrees are distinct links in the universal chain and form connected series of several kinds. Three of these series—the mineral kingdom, the plant kingdom, and the animal kingdom—belong to the earth. In these great series there are also subordinate series, down to the lowest elements. Each series has its first substance, which is dependent on the first series of nature. The first series of nature is an organic development of the concept of the mathematical point. Here, Swedenborg comes very close to the Neoplatonic conception of a world soul, a creative intellect from which the material world is called forth by the process of emanation. It seems probable that Aristotle's notion of the hierarchy of organisms was a decisive influence in the structuring of this gigantic system, in which Swedenborg has tried to arrange all series and degrees in a fixed order that determines all their interrelations. Swedenborg refused to follow Leibniz and Christian Wolff in calling his first substances monads because he did not look upon them as absolutely simple. For him they are created not directly from the Infinite but via the first substance of nature, in the same way that, according to the *Principia*, all natural elements are produced indirectly via the mathematical point.

The first substance of the series, its *vis formatrix*, determines the development of the whole series. There exists nothing in nature that does not belong to such a series. In the *Oeconomia* the human series consists of four degrees, the soul (*anima*), the reason (*mens rationalis*), the vegetative soul (*animus*), and the corresponding sense organs of the body, but in the theosophic writings after 1745 the series is reduced to three degrees with the *animus* subordinated to the *mens rationalis*. Nor is there any first substance of nature in these later works. The chain of the series extends up to God, who himself becomes the highest series.

PSYCHOLOGY

The philosophy of the theosophic period thus presents a kind of Neoplatonic emanation system, although in his earlier works Swedenborg was more influenced by contemporary philosophy. In his psychology he also turned to Locke, and his epistemology coincides with Locke's tabula rasa theory. According to Swedenborg, there are no innate ideas in the *mens rationalis*. He also thought, however, that all a priori knowledge is in the *anima* but that after the Fall of humanity the soul (*anima*) was separated from the body; this synthetic source of knowledge—in some ways corresponding to Locke's notion of intuitive knowledge—was thereby closed for ordinary people. If we could return to Adam's integrity before the Fall, it would be opened up anew. This dream of regaining paradise haunted Swedenborg in the decade before 1745, and he attempted to devise several methods for discovering this lost knowledge.

DOCTRINE OF CORRESPONDENCE

One of the best-known elements in Swedenborg's philosophy is his doctrine of correspondence. This doctrine parallels the speculations about *harmonia constabilita*, but it also has other connections with contemporary thought. The meaning of the term *correspondence* is stated in a short manuscript written in 1741 and titled *Clavis Hieroglyphica* (A Hieroglyphic Key; London, 1784). This work is an attempt to illustrate how linguistic terms may be used with three different meanings—the natural, the spiritual, and the divine. Later, this doctrine becomes the fundamental exegetic principle of the theosophic works. Swedenborg's doctrine of correspondence is an attempt to describe and explain the relations between the spiritual world and our material universe by means of linguistic analogies, the construction of which may be illustrated by the following example from *Clavis Hieroglyphica*.

> (1) There is no motion without *conatus*, but there is *conatus* without motion. For if all *conatus* were to break out into open motion the world would perish, since there would be no equilibrium. (2) There is no action without will, but there is will without action. If all will were to break out into open action man would perish, since there would be no rational balance or moderating reason. (3) There is no divine operation without providence, but there is indeed a providence not operative or effective. If all providence were operative and effective, human society would not be able to subsist such as it now is, since there would be no true exercise of human liberty. (*Psychological Transactions by Emanuel Swedenborg*, pp. 162–163)

The notions conatus, will, and providence correspond; so do world, humankind, and human society. By such means, the principles of the philosophy of nature are given a wider field of application, so that they reveal heavenly and divine secrets. Fundamentally, this doctrine may be interpreted as a variation of the Platonic theory of the relations between the world of ideas and the world of senses, but it is important to stress that Swedenborg

looked upon his system primarily as a synthesis of ancient wisdom and contemporary thought.

The *Clavis Hieroglyphica* is related to the interpretations of hieroglyphics that were made during the Renaissance. This is apparent in Swedenborg's use of excerpts from Wolff's *Psychologia Empirica* (1732) where the famous German rationalist discusses the Egyptian hieroglyphs and their mystic signification and gives examples from John Amos Comenius and others. More important, Wolff inspired speculation about the universal philosophical language, *mathesis universalium* (Swedenborg) or *characteristica universalis* (Leibniz). In a posthumously published manuscript (Stockholm, 1869), Swedenborg tried to formulate his psychophysical conclusions in algebraic formulas of sorts, and he declared his conviction that such an attempt might eventually succeed. But in the meantime he introduced in the *Clavis Hieroglyphica* what he called a key to natural and spiritual arcana by way of correspondences and representations. Thus, there is no doubt that the doctrine of correspondence must be regarded as Swedenborg's contribution to the solution of the problem of the philosophical language. It should be noted, however, that he seems to have been influenced by Nicolas Malebranche in respect to the correspondent relations between the mind and the cerebral base. Swedenborg also follows another fundamental thought of Malebranche, according to which the omnipotence of God functions in conformity with an eternal order (*l'ordre immuable*); this idea becomes prominent in Swedenborg's theosophic writings.

THEOSOPHIC WORKS

Swedenborg's scientific and theosophic works are closely related. The decisive difference is that Swedenborg after a profound spiritual experience in 1745 directed his reasoning exclusively toward the interpretation of Scripture according to the doctrine of correspondence. His first exegetic work is *Arcana Coelestia quae in Genesi et Exodo Sunt Detecta* (8 vols., London, 1749–1756), and it was followed by many others. In all his exegetic treatises Swedenborg also gives vivid descriptions of his experiences in the spiritual world. Apart from these descriptions we meet with the same main theories, although they have been developed into an emanationist theology. Like most of his contemporaries, Swedenborg had always been certain of the existence of spirits and angels, and in the exegetic works he went so far as to describe a comprehensive spiritual system. The spirits live in cities where they have an active social life with social functions (even marriage) corresponding to earthly conditions. The relegation of spirits to heaven or hell from the intervening spiritual world depends on the spirits themselves, since their utmost desire (*amor regnans*) leads them into suitable company.

Christ and the doctrine of atonement play a very insignificant role in Swedenborg's theology, and he dismissed the Trinity dogma. Christ is the *Divinum Humanum*, a manifestation in time of God himself. Swedenborg's theology is extremely intellectual and totally dependent on the interpretation of the divine word as the mediating link between the Creator and humankind. In the course of time decadent churches have destroyed the original meaning of this word, and Swedenborg saw his mission as the restoration of its primary sense. He identified his own exegetic activity with the return of Messiah and the foundation of the New Jerusalem. However, Swedenborg did not aspire to effect conversions but confined himself to explaining the spiritual meaning of the Scriptures. He felt he had been commanded to do this in his decisive vision of 1745.

CONCLUSION

This is not the place to discuss the difficult problem of Swedenborg's mental status. For many modern observers it is only too easy to look upon his theosophy as the result of a pathological development of a pronouncedly schizoid personality whose intense desire for synthesis could not be satisfied within the boundaries of science and normal experience. But this must remain speculation. What is certain is that hundreds of thousands of followers have seen in him a prophet and visionary explorer of divine secrets. He has had a wide influence in several fields of thought and art, especially in romantic and symbolist literature; for poets like Charles-Pierre Baudelaire and August Strindberg he was a teacher and predecessor. Swedenborg is, of course, not a philosopher in the modern meaning of the word, but he is an interesting representative of the mystical trend in eighteenth-century thought.

See also Aristotle; Buffon, Georges-Louis Leclerc, Comte de; Coleridge, Samuel Taylor; Comenius, John Amos; Kant, Immanuel; Laplace, Pierre Simon de; Leibniz, Gottfried Wilhelm; Locke, John; Mysticism, History of; Nature, Philosophical Ideas of; Neoplatonism; Wolff, Christian.

Bibliography

Swedenborg's manuscripts are deposited in the library of the Royal Academy of Science in Stockholm. The greater part of

them have been published (as photolithographs of the original or in edited translations or both) by the New Church societies, especially the Swedenborg Scientific Association in the United States, which is a great aid to scholars. Swedenborg wrote in Latin, but almost all of his works are available in English translations; a detailed but unfortunately obsolete bibliography is J. Hyde, *A Bibliography of the Works of Emanuel Swedenborg Original and Translated* (London: Swedenborg Society, 1906).

The following English translations of his many philosophical and scientific works can be recommended: *The Principia; or, The First Principles of Natural Things, Being New Attempts Toward a Philosophical Explanation of the Elementary World*, translated by A. Clissold, 2 vols. (London, 1846); *The Infinite and the Final Cause of Creation, Also the Intercourse between the Soul and the Body*, translated by J. J. G. Wilkinson (London, 1908); *Psychologica, Being Notes and Observations on Christian Wolff's "Psychologia Empirica" by Emanuel Swedenborg*, translated and edited by A. Acton (Philadelphia, 1923); *The Economy of the Animal Kingdom, Considered Anatomically, Physically, and Philosophically by Emanuel Swedenborg*, translated by A. Clissold, 2 vols. (London, 1845–1846); *The Fibre*, Vol. III of *The Economy of the Animal Kingdom, Considered Anatomically, Physically, and Philosophically by Emanuel Swedenborg*, translated and edited by A. Acton (Philadelphia, 1918); *A Philosopher's Note Book. Excerpts from Philosophical Writers and from the Sacred Scriptures on a Variety of Philosophical Subjects; Together with Some Reflections, and Sundry Notes and Memoranda by Emanuel Swedenborg*, translated and edited by A. Acton (Philadelphia, 1931); *The Brain Considered Anatomically, Physiologically, and Philosophically by Emanuel Swedenborg*, translated and edited by R. L. Tafel, 2 vols. (London, 1882–1887); *Three Transactions on the Cerebrum. A Posthumous Work by Emanuel Swedenborg*, translated and edited by A. Acton, 2 vols. (Philadelphia, 1937–1940); *Psychological Transactions by Emanuel Swedenborg*, translated and edited by A. Acton, 2nd ed. (Philadelphia, 1955); *Rational Psychology. A Posthumous Work by Emanuel Swedenborg*, translated and edited by N. H. Rogers and A. Acton (Philadelphia, 1950); *The Animal Kingdom Considered Anatomically, Physically, and Philosophically by Emanuel Swedenborg*, translated by J. J. G. Wilkinson, 2 vols. (Boston, 1858); *The Animal Kingdom*, Parts 4 and 5, translated and edited by A. Acton (Bryn Athyn, PA, 1928); *The Five Senses*, translated and edited by E. S. Price (Philadelphia, 1914); and *The Worship and Love of God*, translated by F. Sewall and A. H. Stroh (Boston, 1925).

The vast literature about Swedenborg is of unequal quality. An excellent survey is given in M. Lamm, *Swedenborg* (Stockholm, 1915); it has been translated by Ilse Meyer-Lüne as *Swedenborg: Eine Studie über seine Entwicklung zum Mystiker und Geisterseher* (Leipzig, 1922), and into French by E. Söderlindh as *Swedenborg* (Paris, 1936). This is still the best work available. In Ernst Benz, *Emanuel Swedenborg: Naturforscher und Seher* (Munich, 1948), there is more stress on theology and church history, but in general the author follows Lamm. A popular biography is S. Toksvig, *Emanuel Swedenborg, Scientist and Mystic* (New Haven, CT: Yale University Press, 1948). A modern solid monograph, although inspired by New Church teachings, is C. O. Sigstedt, *The Swedenborg Epic* (New York: Bookman Associates, 1952). An analysis of *De Cultu et Amore Dei*, which also deals with many of the philosophical and scientific problems in the rest of Swedenborg's production up to 1745, is I. Jonsson, *Swedenborgs Skapelsedrama "De Cultu et Amore Dei"* (Stockholm, 1961), written in Swedish with a summary in English.

Swedenborg's correspondence has been published in translations and with very informative commentaries in A. Acton, *The Letters and Memorials of Emanuel Swedenborg* (Bryn Athyn, PA: Swedenborg Scientific Association, 1948).

The biographical sources are collected in R. L. Tafel, *Documents concerning the Life and Character of Emanuel Swedenborg*, 3 vols. (London, 1875–1890).

Among the many useful studies by A. H. Stroh may be mentioned "The Sources of Swedenborg's Early Philosophy of Nature," Vol. III of *Emanuel Swedenborg: Opera Quaedam aut Inedita aut Obsoleta de Rebus Naturalibus*, published by the Royal Swedish Academy of Science (Stockholm, 1911), and "Swedenborg's Contributions to Psychology," in *Transactions of the International Swedenborg Congress* (London, 1911).

See also Clarke Garrett, "Swedenborg and the Mystical Enlightenment in Eighteenth-Century England," *Journal of the History of Ideas* (45 [1984]: 67–82).

Inge Jonsson (1967)

Bibliography updated by Tamra Frei (2005)

SWIFT, JONATHAN

(1667–1745)

Jonathan Swift, the British clergyman, moralist, satirist, poet, and political journalist, was born in Dublin, a few months after his father's death. He was educated at Kilkenny Grammar School and received his MA *speciali gratiâ* from Trinity College, Dublin, in 1686 and MA from Hart Hall, Oxford, in 1692. Periodically, from 1689 to 1699, he acted as secretary to Sir William Temple at Moore Park, Surrey. Ordained deacon and priest in the established church of Ireland, he was left by Temple's death in 1699 to make a career for himself. As domestic chaplain to the earl of Berkeley, lord justice of Ireland, he returned to Dublin and was granted the DD degree in 1701 by Trinity College.

In 1704 there appeared anonymously (his customary mode of publishing) *A Tale of a Tub* and *The Battle of the Books*, brilliant satires upholding the ancients against the moderns; assaulting both Catholic and Puritan theologies while upholding the *via media* of the Anglican Church; and castigating the shallowness of contemporary scholarship and literature. Thereafter Swift associated with the Whiggish wits in the circle of Joseph Addison and Richard Steele, contributing to the *Tatler* and laughing the astrologer John Partridge out of business in the hilar-

ious *Bickerstaff Papers* (1708–1709). Gradually, however, when the Whig ministry displayed no interest either in the welfare of the Irish church or in Swift's own ecclesiastical preferment, he veered toward the Tories. His literary friends now included Alexander Pope, John Gay, William Congreve, Matthew Prior, and John Arbuthnot, many of whom later joined with him in the famous Scriblerus Club dedicated to eternal warfare against the dunces.

In 1710 Swift assumed the editorship of the *Examiner*, thus becoming party spokesman for the new Tory ministry of Robert Harley and Lord Bolingbroke. He shortly resigned this post to work on *The Conduct of the Allies* (1711), a pamphlet designed so to sway public opinion as to bring about the end of the "Whiggish" War of the Spanish Succession, an event that occurred in 1713 with the Treaty of Utrecht. Swift was unable, however, to reconcile the ever increasing animosities between Harley (now Lord Oxford) and Bolingbroke, each of whom was surreptitiously treating with both Jacobite and Hanoverian claimants to the British crown. The death of Queen Anne in 1714 and the accession of George I (of Hanover) led to the downfall and disgrace of the Tory Party. Swift, having been installed the previous year as dean of St. Patrick's Cathedral in Dublin, retired to Ireland, a country whose people he despised. A fascinating record of events and personalities of the turbulent years of ecclesiastical and political intrigues, 1710–1713, is preserved in his letters to Esther Johnson, known as the *Journal to Stella*.

During the long years of "exile," Swift, paradoxically, became the national hero of Ireland, rising to her defense against the ruthless exploitation by the English. Two works are especially notable in this campaign. First, there was *The Drapier's Letters to the People of Ireland* (1724), which caused the king of England, the prime minister, and the Parliament to back down from the insult to the people of Ireland in the proposed coining of William Wood's copper halfpence. And second, there was *A Modest Proposal For preventing the Children Of Poor People From Being a Burthen to Their Parents or Country, And For making them Beneficial to the Publick* (1729), which employed shock technique to apprise the Irish people of the fact that slaughtering and dressing infants for the dinner tables of English absentee landlords was really little different from prevailing conditions, which allowed them to die of starvation. In the *Proposal* and other politicoeconomic publications Swift advocated what was later to be called the boycott. In 1726 the immortal social and political satire *Gulliver's Travels* was published in London. Minor works—economic, political, and satirical—continued to appear until about 1739. In 1742 Swift's health had deteriorated to the extent that, for his own protection, he was declared of unsound mind and memory and incapable of caring for himself or his estate. Today it is recognized that Swift was suffering from labyrinthine vertigo (Ménière's disease), a purely physical disease, and that in modern terminology he was not insane. He lingered on until 1745, when he died in his seventy-eighth year and was buried in St. Patrick's Cathedral, ironically leaving most of his estate for the founding of a hospital for the insane. His last words were "I am a fool." He had prepared for himself an epitaph in Latin that is translated "When savage indignation can no longer torture the heart, proceed, traveller, and, if you can, imitate the strenuous avenger of noble liberty." "Savage indignation" and the fight for "noble liberty" are truly the prime characteristics of Jonathan Swift.

RELIGION AND MORALITY

Never professing to be a philosopher, Swift was nevertheless a serious thinker on the problems of religion and morality; however, because of his pervasive use of irony, his writings in this area have not infrequently been misunderstood and maligned. Swift always maintained, and quite properly, that he was not attacking religion but the corruptions and excesses of religion and the abuses of reason. As dean, he performed all the functions of that office and was in every respect a sincere Christian. In his surviving sermons, only eleven of which are unquestionably authentic, he takes a commonsense (derived from the funded experience of humankind) approach to theology. The lingering Trinitarian controversy, which caused such bitterness and name-calling among the "orthodox" that Parliament prohibited further publication on the subject, Swift found thoroughly repugnant. In *A Letter to a Young Gentleman, lately enter'd into Holy Orders* (1720), Swift advised that the Christian mysteries should not be explicated by divines but should remain incomprehensible, for otherwise they would not be "mysteries." Though God-given, human reason is not infallible, because of the interests, passions, and vices of the individual. Although there is clearly a skeptical bent in Swift, he is not to be regarded as a skeptic. Mysteries (for instance, the Trinity) are to be accepted on faith (which is above reason) and asserted on the authority of the Scriptures. As Swift stated in a private letter, "The grand points of Christianity ought to be taken as infallible revelations." It was this orthodox insistence on revelation that made Swift the intractable enemy of the English deists, who maintained that knowledge is prior to assent or faith.

Swift's religious antirationalism, anti-intellectualism, and fideism are well illustrated in his writings against the deists: John Toland, Matthew Tindal, and Anthony Collins were his chief butts. Collins who, in his *Discourse of Free thinking* (1713), had twice taunted Swift by name, is subjected to Swiftian irony in *Mr. C——n's Discourse of Freethinking; put into plain English by way of Abstract, for the Use of the Poor* (1713). Grossly unjust to Collins though it is deliberately intended to be, Swift's work is a witty exploitation of antirationalistic and anti-intellectualistic arguments. The optimistic apriorism inherent in deism was repugnant to Swift, who as an essentially Christian pessimist was always less concerned with philosophical and theological niceties than with the practical problems of morality.

Swift's vital interest in morality is observable in *An Argument against Abolishing Christianity* (1711). This masterpiece of irony attacks the rationalistic deistical concept of a self-sufficient religion of nature that needs no special revelation by assuming the position that "real" Christianity is no longer capable of justification to a sophisticated age. However, "nominal" Christianity is justifiable on grounds of expediency: It may help to preserve pride, wealth, and power and, possibly, to prevent a drop in the stock market of as much as 1 percent. *A Project for the Advancement of Religion and the Reformation of Manners* (1709) urges Queen Anne to lead a moral crusade against existing vices in the nation. That Swift was not ironic but completely earnest in this project is certain because of the abhorrence of human vices and the necessity for reformation he expressed in many other writings.

Believing that man is not *animal rationale* but merely *rationis capax*, Swift discerns a negative philosophy of history in the human tendency to degenerate after a certain degree of order and virtue has been achieved. In this restrictive sense only is he to be called a Christian misanthrope or simply a misanthrope. Swift devoted his life to exposing cruelty, inhumanity, inordinate love of power, pride, corrupt politics, and political oppression and to inculcating integrity and virtue in its major aspects of magnanimity and heroism—yet with no illusion that human nature is capable of reaching virtue in an eminent degree. This satiric-moralistic aim, enhanced by Swift's comic vision, finds its most brilliant literary achievement in *Gulliver's Travels*, a work that always has, and always will, vex, shock, divert, and entertain the world.

See also Addison, Joseph; Bolingbroke, Henry St. John; Collins, Anthony; Gay, John; Pope, Alexander; Religion and Morality; Tindal, Matthew; Toland, John.

Bibliography

MODERN EDITIONS OF SWIFT

The Letters of Jonathan Swift to Charles Ford, edited by D. Nichol Smith. Oxford: Clarendon Press, 1935.

The Prose Works of Jonathan Swift. 14 vols., edited by Herbert Davis. Oxford: Blackwell, 1939–1968.

Journal to Stella, edited by Harold Williams. Oxford: Clarendon Press, 1948.

The Poems of Jonathan Swift. 3 vols., edited by Harold Williams. Oxford, 1958.

The Correspondence of Jonathan Swift. 5 vols., edited by Harold Williams. Oxford, 1963–1965.

MODERN WORKS ON SWIFT

Davis, Herbert. *The Satire of Jonathan Swift*. New York: Macmillan, 1947.

Donoghue, Denis. *Jonathan Swift: A Critical Anthology*. Harmondsworth: Penguin, 1971.

Downie, J. A. *Jonathan Swift, Political Writer*. Boston: Routledge & K. Paul, 1984.

Ehrenpreis, Irwin. *The Personality of Jonathan Swift*. London: Methuen, 1958.

Ehrenpreis, Irwin. *Swift: The Man, His Works, and the Age*. Vol. 1: *Mr. Swift and His Contemporaries*. London: Methuen, 1962.

Fauske, Christopher J. *Jonathan Swift and the Church of Ireland, 1710–1724*. Dublin: Irish Academic Press, 2002.

Ferguson, Oliver W. *Jonathan Swift and Ireland*. Urbana: University of Illinois Press, 1962.

Fox, Christopher, ed. *The Cambridge Companion to Jonathan Swift*. New York: Cambridge University Press, 2003.

Gilbert, Jack G. *Knaves, Fools, and Heroes: Jonathan Swift's Ethics*. Unpublished PhD diss., University of Texas, 1962.

Landa, Louis. *Swift and the Church of Ireland*. Oxford: Clarendon Press, 1954.

Landa, Louis. "Swift, the Mysteries, and Deism." In *Studies in English*, 239–256. University of Texas, 1944.

Looten, C. *La pensée religieuse de Swift*. Paris, 1935.

Quintana, Ricardo. *The Mind and Art of Jonathan Swift*. 2nd ed. London: Methuen, 1953.

Quintana, Ricardo. *Swift: An Introduction*. London and New York: Oxford University Press, 1955.

Quintana, Ricardo. *Two Augustans: John Locke, Jonathan Swift*. London: University of Wisconsin Press, 1978.

Rawson, Claude, ed. *Jonathan Swift: A Collection of Critical Essays*. Englewood Cliffs, NJ: Prentice Hall, 1995.

Wedel, T. O. "On the Philosophical Background of Gulliver's Travels." *Studies in Philology* 23 (1926): 434–450.

Williams, Kathleen. *Jonathan Swift and the Age of Compromise*. Lawrence: University of Kansas Press, 1958.

Ernest Campbell Mossner (1967)

Bibliography updated by Tamra Frei (2005)

SWINESHEAD, RICHARD

(mid-1300s)

Richard Swineshead (Swyneshed; on the Continent, more commonly Suiseth) is the name now commonly ascribed to the author of the *Book of Calculations* (*Liber Calculationum*) although in various manuscripts and printed editions he is also given the first names John, Raymund, Roger, and William, among others. Based on the work of James A. Weisheipl, a different person with the name Roger Swyneshed, who was a Benedictine monk at Glastonbury, is now credited with writing a work that is in some ways similar, titled *Descriptions of Motions* or *On Natural Motions* (*Descriptiones motuum* or *De motibus naturalibus*) dated to the mid-1330s and found in Erfurt manuscript Amplonian F 135, ff. 25va–47rb. This same Roger Swyneshed is credited with logical works *On Insolubles* and *On Obligations* (*De insolubilibus* and *De obligationibus*) connected to standard academic exercises within medieval universities. If the same person wrote all of these works, then his views must have matured and changed considerably between the writing of the various works. The following entry will be limited to a discussion of the author of the *Book of Calculations*. Those interested in the history of logic should turn first to the articles, listed below, by Paul Spade on Roger Swyneshed's works.

Documentary evidence indicates that Swineshead was a fellow of Merton College, Oxford, probably in 1340—certainly in 1344—and again in 1355. Manuscript copies of the *Book of Calculations* are often incomplete and arranged differently from the printed editions. The work shows clear influence of Thomas Bradwardine's *On the Proportions of Velocities in Motions* (1961 [1328]) and of William Heytesbury's *Rules for Solving Sophisms* (1494 [1335]). Influence of the *Book of Calculations* begins to show up in Paris before 1350. Through the early sixteenth century, the work was widely studied on the Continent, in Italy and Spain as well as France, leading to various propedeutic works explaining its methods to potential readers. G. W. Leibniz several times recommended that the book be reprinted, both as a gem of the early history of printing and because the author was among the first to introduce mathematics into natural philosophy or metaphysics. To that end Leibniz went so far as hire someone to copy the Venice, 1520, printed edition by hand in preparation for the reprinting. Although Leibniz's project never came to fruition, the hand copy still exists in the Niedersächsische Landesbibliothek in Hannover, Germany.

In the printed versions of the *Book of Calculations* there are sixteen treatises, which cover:

I. Intension and remission of forms.

II. (Measures of) difform qualities.

III. Intensity of elemental bodies having two unequally intense qualities.

IV. Intensity of mixed bodies.

V. Rarity and density.

VI. Augmentation.

VII. Reaction.

VIII. Powers of things.

IX. Difficulty of action.

X. Maxima and minima.

XI. Place of an element.

XII. Light sources.

XIII. Action of light sources.

XIV. Local motion.

XV. Motion in nonresisting media (in media with varying resistances).

XVI. Induction of the maximum degree.

What these treatises have in common is an effort to attach quantitative measures to physical entities. Swineshead first tries to establish scales of measure for static magnitudes, such as intensities of heat and cold. He then attempts to measure speeds of change in the three categories in which medieval Aristotelians believed motion to occur, namely place, quality, and quantity. Treatise XIV, on dynamics, assumes the truth of Bradwardine's rule stating that the velocities in motions depend on the ratios of forces to resistances, using a special sense of the variation of ratios connected with the notion of *compounding* ratios used in Euclid's *Elements* (Book VI, proposition 23). The *Book of Calculations* represents a stage in medieval intellectual development in which logic (including the theory of supposition) together with mathematics begin to move physics from the matrix of natural philosophy to the status of an exact science.

Most of the treatises of the *Book of Calculations* follow the standard scholastic format in which arguments are given for and against competing opinions before Swineshead settles on and argues for the theory he believes to be more correct. Like Heytesbury's *Rules for Solving Sophisms*, the *Book of Calculations* seems to have

been composed to provide university undergraduates with the analytical tools they needed to participate in disputations. As such, it is a good text to use for learning about the concepts and tools of fourteenth-century natural philosophy, including mathematics. Although the book does not expound its natural philosophical, let alone its metaphysical, foundations in detail, Swineshead appears to have agreed with the other Oxford Calculators, who (with the exception of Walter Burley) adopted the Scotistic addition theory of qualitative change and favored the ontological parsimony usually associated with William of Ockham. For more detail on the logical tools assumed by Swineshead, one should look to the work of Heytesbury, and for the natural philosophical background, to John Dumbleton's *Summa logicae et philosophiae naturalis*, as described in the work of Edith Sylla (1991b). A final fourteenth-century Oxford scholar whose work is related to that of Swineshead is Richard Kilvington, on whom there is significant recent scholarly work.

See also Aristotle; Bradwardine, Thomas; Burley, Walter; Heytesbury, William; Kilvington, Richard; Leibniz, Gottfried Wilhelm; Philosophy of Science; William of Ockham.

Bibliography

WORKS BY RICHARD SWINESHEAD

Book of Calculations. Only Treatise XI has been published in full, in the article by Hoskin and Molland (listed below). Edith Sylla has published a detailed outline of the entire work in her dissertation (listed below), as well as a similar outline for the *Descriptiones motuum* ascribed to Roger Swyneshed.

ROGER SWYNESHED'S LOGICAL WORKS

Spade, Paul Vincent, ed. "Roger Swyneshed's *Insolubilia*: Edition and Comments." *Archives d'histoire doctrinale et littéraire du moyen âge* 46 (1979): 177–220.

Spade, Paul Vincent, ed. "Roger Swyneshed's *Obligationes*: Edition and Comments." *Archives d'histoire doctrinale et littéraire du moyen âge* 44 (1977): 243–85.

WORKS ON RICHARD SWINESHEAD

Clagett, Marshall. "Richard Swineshead and Late Medieval Physics." *Osiris* 9 (1950): 131–161.

Clagett, Marshall. *The Science of Mechanics in the Middle Ages*. Madison: University of Wisconsin Press, 1959.

Crosby, H. Lamar Jr., ed. *Thomas of Bradwardine. His* Tractatus de Proportionibus. *Its Significance for the Development of Mathematical Physics*. Madison: University of Wisconsin Press, 1961.

Heytesbury, William. *Regule solvendi sophismatum*. Venice, 1494; many manuscripts.

Hoskin, Michael, and A. George Molland. "Swineshead on Falling Bodies: An Example of Fourteenth-Century Physics." *British Journal for the History of Science* 3 (1966): 150–182.

Murdoch, John, and Edith Sylla. "Swineshead, Richard." In *Dictionary of Scientific Biography*, Vol. 13. New York: Charles Scribner's Sons, 1976, pp. 184–213.

Spade, Paul. "Insolubilia," and "Obligations: Developments in the fourteenth century." In *The Cambridge History of Later Medieval Philosophy*, edited by Norman Kretzmann, Anthony Kenny, and Jan Pinborg. Cambridge, U.K.: Cambridge University Press, 1982.

Sylla, Edith. "Mathematical Physics and Imagination in the Work of the Oxford Calculators: Roger Swineshead's *On Natural Motions*." In *Mathematics and Its Application to Science and Natural Philosophy in the Middle Ages*, edited by Edward Grant and John Murdoch. Cambridge, U.K.: Cambridge University Press, 1987, pp. 69–101

Sylla, Edith. "The Oxford Calculators." In *The Cambridge History of Later Medieval Philosophy*, edited by Norman Kretzmann, Anthony Kenny, and Jan Pinborg. Cambridge: Cambridge University Press, 1982.

Sylla, Edith. "The Oxford Calculators and Mathematical Physics: John Dumbleton's *Summa Logicae et Philosophiae Naturalis*, Parts II and III." In *Physics, Cosmology and Astronomy, 1300-1700: Tension and Accommodation*, edited by Sabetai Unguru. Boston Studies in the Philosophy of Science, Vol. 126, Dordrecht; Boston: Kluwer Academic Publishers, 1991b, pp. 129–161.

Sylla, Edith. *The Oxford Calculators and the Mathematics of Motion, 1320–1350. Physics and Measurement by Latitudes*. Harvard University Dissertations in History of Science. New York: Garland Press, 1991a.

Weisheipl, James. "Roger Swyneshed, O. S. B., Logician, Natural Philosopher, and Theologian." In *Oxford Studies Presented to Daniel Callus*. Oxford: Clarendon Press, 1964. pp. 231–52.

Edith D. Sylla (2005)

SYLLOGISM

See *Logic, Traditional; Logical Terms, Glossary of*

SYLVESTER OF FERRARA, FRANCIS

(c. 1474–1528)

Francis Sylvester of Ferrara, a leading Thomistic commentator, sometimes listed under Francis, sometimes under his family name Silvestri, and cited in the Latin literature as Ferrariensis, was born in Ferrara, Italy. He entered the Dominican order in 1488, and took his magistrate in theology at Bologna in 1507. He later taught philosophy and theology at Bologna and other cities in northern Italy. Sylvester's "Commentary on *Summa Contra Gentiles*" has been printed with the definitive edition

of that work of St. Thomas Aquinas in the Leonine edition of *Opera Omnia S. Thomae* (Vols. XIII–XV, Rome, 1918–1926). Among his other philosophical writings are two commentaries on Aristotle: *Annotationes in Libros Posteriorum* (Venice, 1535), and *Quaestionum Libri de Anima* (Venice, 1535).

A critic of Scotist and Ockhamist thought, Sylvester of Ferrara held some highly personal views, modifying Thomism in directions different from those of his contemporary Cajetan. In psychology and epistemology, Sylvester taught a theory of intellectual abstraction by compresence in which the actual object of understanding is quite different from the intelligible determinant that is impressed on the possible intellect (*species impressa* is not the *intelligibile*). The agent intellect performs two distinct actions, one on the phantasm and the other on the possible intellect. He modified Thomas's view that the proper object of the understanding is the universalized nature of sensible things, by teaching that the possible intellect forms a proper concept of the singular. In metaphysics, he also modified Thomism, saying that pure essences—for example, the natures of angels—may be multiplied numerically in existence, although how this is done is unknown. Concerning the individuation of bodies, Sylvester held that this is accomplished by matter as marked by definite dimensions (*materia signata quantitate determinata*).

Perhaps Sylvester is best known for his explanation of metaphysical analogy as that general characteristic of beings whereby they all somewhat resemble each other and yet are different. Contrary to the theory of Cajetan that all analogy reduces to that of proportionality, Sylvester argued that in every instance of analogy there is a first analogate which determines the meaning of the other analogates (*analogia unius ad alterum*). In endeavoring to harmonize various texts of Thomas, Sylvester may have minimized the essential character of analogy, moving in the direction of attribution and metaphor.

Among twentieth-century followers of Sylvester's theory of analogy are such important Thomists as F. A. Blanche, J. M. Ramirez, and N. Balthasar.

Bibliography

For a concordance to the Latin text of the Commentary on *Summa Contra Gentiles,* use *Indices … in Commentariis … Caietani et … Ferrariensis,* in Vol. XVI of the Leonine edition of *Opera S. Thomae* (Rome, 1948). See also G. P. Klubertanz, *St. Thomas Aquinas on Analogy* (Chicago: Loyola University Press, 1960), pp. 10ff., and F. A. Blanche, "Surle Sens de quelques locutions concernant l'analogie," in *Revue des sciences philosophiques et théologiques* 10 (1921): 52–59 and 169–193.

Vernon J. Bourke (1967)

SYMBOLIC LOGIC

See *Logic, History of: Modern Logic*

SYMPATHY AND EMPATHY

The notions of empathy and sympathy have a muddled history, and they are often used interchangeably. Recently, efforts at clarifying the difference have focused on empathy first and proceeded to characterize sympathy by contrast. The contemporary philosophical conception of empathy has three aspects. If Sam empathizes with Maria's anger, then: 1) Sam has a representation of Maria as angry; 2) Sam comes to have his empathic experience because of his representation of Maria as angry; 3) Sam's experience involves experiencing a state that is similar to anger.

On most accounts, sympathy differs from empathy by being triggered solely by emotions that are linked with pain and involves—either as consequence or through sharing the other person's pain—feeling sorry for the other person or wanting to alleviate the other person's suffering. The phrases *feeling with* and *feeling for*, respectively, are often used to capture the difference between the two notions.

Concerning number one above, the main point of contention is whether it is a requirement that the representation of Maria as angry be true, or whether Sam can empathize with Maria even if Maria is not angry now. Concerning number two, the main issue is how to describe the process of coming to feel empathic because of someone else's emotion. Does it require imagining the other person's emotion/situation or is it the case that a purely causal story not involving imagination sufficient for empathy? Concerning number three, the question is how to characterize the kind of affective experience empathy is. Is it an emotion of the same type as that of the person empathized with? Or are there rather natural empathic counterpart emotions corresponding to the emotions of the person empathized with? Or does empathic experience involve having some nonemotional feelings associated with the emotion empathized?

Although all these questions are still debated, there are two points of agreement: Empathy is not an emotion,

but a phenomenon concerning the way one comes to be in touch with other people's emotions; in contrast, sympathy is, on one common conception, an emotional experience and amounts to something close to compassion. This contemporary understanding of empathy and sympathy has had many historical precursors under various confusing names. Most of these have focused on number two (i.e. the special way in which empathic experience is caused). Benedict de Spinoza's theory of affect imitation and David Hume's principle of sympathy, both central to these authors' conceptions of moral agency, exemplify the view that a fundamental trait of humanity resides in its capacity to experience other people's affects simply through the process of imagining these people experiencing these affects. The Scottish philosopher and economist, Adam Smith, held a similar view although his focus was on imagining other people's situations rather than affects.

The concept of empathy became prominent at the turn of the nineteenth century in German psychology and philosophy. It played an important role in elucidating human creatures' emotional engagement with the arts and how they come to interpret and understand each other as psychological beings. It was in this context that the term *empathy* itself was coined to translate the German word *Einfühlung* (i.e., "to feel one's way into"). Edmund Husserl, his student Edith Stein, and later Max Scheler are three philosophers whose contributions have shaped our present understanding of empathy. In particular, they each offered a particular elucidation of number three, insisting, each in their own way, that empathic experience cannot be of the same sort as the feeling that is the object of the empathic experience. Empathizing with someone who is angry would thus not involve oneself being angry, although it might involve the feelings associated with anger.

Interest in empathy and sympathy—and the broader interest in psychological simulation—has recently been driven by the thought that these phenomena are keys to the understanding of the development of moral agents. The idea—associated with a Humean take on morality—is that empathy is the most important source of one's understanding of others as beings with joys and sufferings directly dependent on the way one treats them. Hence the thought that moral sentiments and moral agency stem from a capacity to empathize with others. Contemporary empirical research on empathy has reinforced this idea. So has the existence of people (psychopaths) lacking both empathy and moral concern. However, the existence of people suffering from the same deprivation (some high-functioning autistic people) but manifesting a clear concern with morality suggests that empathy might only be a significant aid to moral growth, but not a necessary component of it.

See also Altruism; Moral Psychology; Moral Sentiments.

Bibliography

Darwall, Stephen. "Empathy, Sympathy, Care." *Philosophical Studies* 89 (1998): 261–282.

Hume, David. *A Treatise of Human Nature*. 2nd ed. Oxford: Oxford University Press, 1978.

Kennett, Jeanette. "Autism, Empathy, and Moral Agency." *The Philosophical Quarterly* 52 (2002): 340–357.

Nilsson, Peter. *Empathy and Emotions: On The Notion of Empathy as Emotional Sharing*. Sweden: Umea University, 2003.

Scheler, Max. *The Nature of Sympathy*. Translated by Peter Heath. London: Routledge, 1948.

Smith, Adam. *The Theory of Moral Sentiments*. Indianapolis: Liberty Fund, 1984.

Snow, Nancy E. "Empathy." *American Philosophical Quarterly* 37 (2000): 65–78.

Sober, Eliot, and David S. Wilson. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press, 1998.

Stein, Edith. *On the Problem of Empathy. The Collected Works of Edith Stein*. 3rd ed. Translated by W. Stein. Washington, DC: ICS Publications, 1989.

Julien A. Deonna (2005)

SYNONYMITY

"Synonymity" has been a major topic in philosophy since the publication of Rudolf Carnap's *Meaning and Necessity* in 1947, though it was discussed earlier in the writings of W. V. Quine and C. I. Lewis. After Quine and Morton White launched their attacks on the tenability of the analytic-synthetic distinction, around 1950, the two topics became closely linked.

SYNONYMITY AND THE ANALYTIC-SYNTHETIC DISTINCTION

Analytic statements, in Quine's account, fall into two classes. Those of the first class, exemplified by (1), are logically true.

(1) No unmarried man is married.

Quine has no objection to the notion of analytic truth as used here, for he has what he regards as an acceptable account of the notion of logical truth in terms of which the notion of analytic truth is partially explicated. "The relevant feature of this example is that it not merely is

true as it stands, but remains true under any and all reinterpretations of 'man' and 'married.' If we suppose a prior inventory of *logical* particles, comprising 'no,' 'un-,' 'not,' 'if,' 'then,' 'and,' etc. then in general a logical truth is a statement which is true and remains true under all reinterpretations of its components other than the logical particles" (all quotations from Quine are from "Two Dogmas of Empiricism").

All logical truths are analytic. The problems that beset analyticity, however, concern those purported analytic truths which are not logical truths. These are typified by

(2) No bachelor is married.

This is not a logical truth, for it does not remain true under every reinterpretation of its nonlogical components, "bachelor" and "married." If (2) is nevertheless to be considered analytic, it is because we can turn it into the logical truth (1) by replacing synonyms with synonyms. Thus, since "bachelor" and "unmarried man" are synonyms, we may replace the former with the latter in (2) in order to arrive at (1), a truth of logic.

It might appear that a generalization of the above considerations would yield a satisfactory account of the notion of an analytic statement. The generalization would go as follows: a statement is analytic if and only if it either (1) is a logical truth or (2) is transformable into a logical truth by the substitution of synonyms for synonyms. This account is rejected by Quine and White on the ground that synonymity (or synonymy, as Quine prefers) is no clearer a notion than analyticity. In Quine's words, "We still lack a proper characterization of this second class of analytic statements, and therewith of analyticity generally, inasmuch as we have had in the above description to lean on a notion of 'synonymy' which is no less in need of clarification than analyticity itself."

INTERCHANGEABILITY CRITERION OF SYNONYMITY

A natural response to Quine is that we can give an acceptable account of synonymity in terms of interchangeability. The suggestion is that the synonymity of two linguistic forms consists simply in their interchangeability in all contexts without change of truth-value—interchangeability, in Gottfried Wilhelm Leibniz's phrase, *salva veritate.* Benson Mates has offered an argument to show that if two expressions are synonymous they are interchangeable everywhere *salva veritate.* Following Gottlob Frege, Mates assumes that the meaning of a declarative sentence is a function of the meanings of the words which compose the sentence. Furthermore, two declarative sentences having the same meaning will necessarily have the same truth-value. It follows from these two assumptions that the replacement of a word in a sentence by another word synonymous with it cannot change the meaning of that sentence and hence cannot change its truth-value. Thus, if two words are synonymous they are interchangeable everywhere *salva veritate.*

In spite of the reasonableness of the above argument, the proposed interchangeability criterion soon runs into difficulty. Consider the synonymous pair "bachelor" and "unmarried man." The following statement is true:

(3) "Bachelor" has fewer than ten letters.

But the result of replacing the word *bachelor* by its synonym *unmarried man* is the false statement

(4) "Unmarried man" has fewer than ten letters.

This case can presumably be set aside on the ground that quoted expressions should themselves be understood as words functioning as names for their quoted contents. The interchangeability test is then interpreted as not applying to words such as *bachelor* when they appear as fragments of other words, such as "*bachelor.*" This makes the account of synonymity rest on the notion of wordhood, but Quine does not object on this account.

Perhaps Quine does not take seriously enough the difficulties involved here. Consider the synonymous pair "brothers" and "male siblings." Replacement of the former by the latter in

(5) *The Brothers Karamazov* is Dostoevsky's greatest novel

turns a true statement into one which is not true,

(6) *The Male Siblings Karamazov* is Dostoevsky's greatest novel.

Quine cannot object to this replacement for the same reason he objects to substitution of synonyms for synonyms within the context of quotation marks, for he cannot reasonably claim that titles are all single words.

The most serious problem connected with the interchangeability criterion is that the requirement is, apparently, too strong. Problems about wordhood aside, it is doubtful that paradigmatic synonym pairs like "bachelor" and "unmarried man" can pass the test. Consider the statement

(7) Jones wants to know whether a bachelor is an unmarried man.

Suppose it true, as it may well be, of some man named "Jones." Replacement of synonym for synonym here yields a statement that is no doubt false,

(8) Jones wants to know whether a bachelor is a bachelor.

CARNAP'S "INTENSIONAL ISOMORPHISM"

Carnap intended the concepts of intensional isomorphism and intensional structure to be explications of the ordinary notion of synonymity. Intensional isomorphism is explained in terms of logical equivalence (L-equivalence) when the usual application of the latter notion is extended beyond full sentences to cover various sentence parts. For example, two names "*a*" and "*b*" are L-equivalent if and only if "$a = b$" is logically true (L-true). Two (one-place) predicate expressions "*P*" and "*Q*" are L-equivalent if and only if "$(x)(Px \equiv Qx)$" is L-true. (This means that it is L-true that whatever has the property *P* also has the property *Q*, and conversely.) An analogous definition extends the notion of L-equivalence to many-place predicates (expressions for relations). Expressions for which the relation of L-equivalence has been defined in this manner are called "designators." If two designators are L-equivalent they are said to have the same intension.

Intensional structure is explained thus: "If two sentences are built in the same way out of corresponding designators with the same intensions, then we shall say that they have the same intensional structure" (all quotations from Carnap are from *Meaning and Necessity*). For example, consider the expressions "2 + 5" and "II sum V." These occur in a language *S* in which "2," "5," "II," and "V" are designations for numbers and "+" and "sum" signs for arithmetical operations. We suppose that according to the semantical rules of *S*, "2" is L-equivalent to "II" (and thus the two have the same intension), "5" is L-equivalent to "V," and "+" is L-equivalent to "sum." With regard to this example Carnap says, "…we shall say that the two expressions are *intensionally isomorphic* or that they have *the same intensional structure,* because they are not only L-equivalent as a whole, both being L-equivalent to '7,' but consist of three parts in such a way that corresponding parts are L-equivalent to one another and hence have the same intension." In our example corresponding parts correspond spatially, but this is not a necessary condition. Thus, Carnap regards "5 > 3" as intensionally isomorphic to "Gr(V,III)" because the (two-place) predicates ">" and "Gr" are L-equivalent and so are "5" and "V" and "3" and "III." The (two-place) predicates "correspond," regardless of their positions in the sentences. The sentence "(2 + 5) > 3" is intensionally isomorphic to "Gr(Sum(II,V),III)" because "2 + 5" is intensionally isomorphic to "Sum(II,V)" and the predicate expressions are L-equivalent, as are "3" and "III." On the other hand "7 > 3" is not intensionally isomorphic to "Gr(Sum(II,V),III)" even though "Gr" is L-equivalent to ">," "3" to "III," and "Sum(II,V)" to "7." They are not intensionally isomorphic because "Sum(II,V)" is not intensionally isomorphic to "7," although these expressions have the same intension (are L-equivalent). Intensional isomorphism of two expressions requires the intensional isomorphism of all corresponding subdesignators.

OBJECTIONS. Consider Carnap's extension of the use of "≡" so as to hold between predicators. According to this extension, if A_i and A_j are two predicators of degree 1, the following abbreviation is allowable:

$$A_i \equiv A_j \qquad \text{for } (X)(A_iX \equiv A_jX).$$

Now let us assume as L-true a sentence of the following form:

(1) $A_i \equiv A_j$.

This sentence will be intensionally isomorphic to

(2) $A_i \equiv A_i$.

But (1) is not intensionally isomorphic to

(3) $(X)(A_iX \equiv A_iX)$,

which is the definitional expansion of (2). Sentence (1) will not be intensionally isomorphic to (3), because (3) contains a designator, "(X)," which cannot be matched to a designator in (1). The point of this criticism is that an expression can be intensionally isomorphic to another expression without being isomorphic to a third expression which has the same meaning as the second according to a definition. For this reason intensional isomorphism seems not to be an adequate explication of synonymity.

In "A Reply to Leonard Linsky," Carnap says that the ordinary notion of synonymity is imprecise. He concludes that more than one explicans must be considered. He proposes a series of seven possible explicata, at least some of which would not be affected by the above criticism.

The most serious argument against Carnap's program is that of Benson Mates: Let "*D*" and "*D'*" be abbreviations for two intensionally isomorphic sentences. Then the following are also intensionally isomorphic:

(1) Whoever believes that *D* believes that *D*.

(2) Whoever believes that *D* believes that *D*'.

Now the following sentence is true:

(3) Nobody doubts that whoever believes that *D* believes that *D*.

But (4), which is intensionally isomorphic to (3), is very likely false:

(4) Nobody doubts that whoever believes that *D* believes that *D*'.

If anybody even doubts that whoever believes that *D* believes that *D*', then (4) is false, and the consequence is that two intensionally isomorphic sentences will differ in truth-value. But since two synonymous sentences cannot differ in truth-value, it follows that intensional isomorphism is not adequate as an explication for synonymity.

According to Hilary Putnam, Carnap believes that his theory in its present form cannot refute Mates's criticism. However, other philosophers (notably Alonzo Church) disagree with Putnam and (apparently) Carnap over the soundness of Mates's argument.

GOODMAN'S THEORY

One of the most widely discussed contributions to the topic of synonymity is Nelson Goodman's "On Likeness of Meaning." His view is particularly attractive to nominalistic philosophers who would avoid "abstract" entities, such as thoughts, senses, and meanings, in their semantical theories. Goodman proposes to explicate the notion of synonymity solely in terms of words and their "extensions"—the objects to which they apply. His account is confined to predicate expressions.

Suppose we say that two predicate expressions have the same meaning if and only if they have the same extensions—are true of the same things. A fatal objection to this view is that there are clear cases where two words have the same extension but do not have the same meaning. *Centaur* and *unicorn,* for example, have the same (null) extension, yet they differ in meaning.

We thus see that any simple identification of sameness of meaning of two expressions with sameness of extension must fail. But Goodman argues that we can still give an extensional account of sameness of meaning; although two words may have the same extension, certain predicates composed by making identical additions to these two words may have different extensions. *Centaur* and *unicorn* have the same (null) extension, but there are centaur pictures that are not unicorn pictures. Thus, "centaur picture" and "unicorn picture" have different extensions. Goodman concludes that "difference of meaning among extensionally identical predicates can be explained as difference in the extensions of certain other predicates. Or, if we call the extension of a predicate by itself its *primary* extension, and the extension of any of its compounds a *secondary* extension, the thesis is formulated as follows: two terms have the same meaning if and only if they have the same primary and secondary extensions." Suppose that in accordance with our nominalistic inclinations we exclude thoughts, concepts, attributes, meanings from the extensions under consideration. This means that when considering the identity of meaning of, for example, *centaur* and *unicorn* we will ignore such secondary extensions as those of "thought of a unicorn" and "thought of a centaur" or "concept of a unicorn" and "concept of a centaur." "If the thesis is tenable, we have answered our question by stating, without reference to anything other than terms and the things to which they apply, the circumstances under which two terms have the same meaning" (all quotations from Goodman are from "On Likeness of Meaning").

Let us see how Goodman's solution works. The predicates "(is the) morning star" and "(is the) evening star" have the same (primary) extension but differ in meaning. This difference is explained by Goodman as being due to a difference in the secondary extensions of these predicates. There are morning-star pictures that are not evening-star pictures and vice versa.

Now consider any predicates "*P*" and "*Q*." Consider the actual ink marks which constitute any inscription of the phrase "a *P* that is not a *Q*." Such an inscription will itself be part of the (secondary) extension of the predicate "*P*," for it will be part of the extension of the expression "*P*-description." But no inscription of the phrase "a *P* that is not a *Q*" will be part of the extension of the expression "*Q*-description." It follows from this that "*P*" and "*Q*" have different (secondary) extensions and hence that they are not synonymous. Since "*P*" and "*Q*" are any predicate expressions, no two predicates are synonymous. For example, any inscription of the phrase "a centaur that is not a unicorn" will be part of the extension of the expression "centaur description," but it will not be part of the extension of the expression "unicorn description." Hence, "centaur" and "unicorn" have different secondary extensions (though they have the same primary extension), so they differ in meaning.

ORDINARY-LANGUAGE VIEW

The discussions of the interchangeability criterion of synonymity and of Goodman's extensional criterion lead to

the same radical conclusion. No two expressions are synonymous. Many philosophers regard this result as a reductio ad absurdum of the proposed criteria. Goodman seems to regard the result as a reductio ad absurdum of what is "commonly supposed" about synonymity. It is not clear whether he thinks that these views are commonly supposed only by the philosophers who discuss such questions or that they are held by those who in ordinary language sometimes declare two words to be synonymous. What is "commonly supposed," according to Goodman, is that (1) some predicates are synonymous with others and (2) synonymous expressions can replace each other "in all nonextensional contexts without change of truth-value."

Goodman holds that the two requirements are incompatible, and we can see why. "A *P* that is not a *Q*" is a *P*-description, not a *Q*-description; "a *Q* that is not a *P*" is a *Q*-description, not a *P*-description. On the supposition that "*P*" and "*Q*" are synonymous the following two statements have the same truth-value, if the interchangeability criterion is correct.

(1) "A *P* that is not a *Q*" is a *P*-description.

(2) "A *P* that is not a *Q*" is a *Q*-description.

However, the first statement is true and the second false. Thus, the predicates "*P*" and "*Q*" are not interchangeable everywhere, even in extensional contexts. But since "*P*" and "*Q*" are any predicates, no predicates are interchangeable everywhere. It follows from this that either no predicates are synonymous or synonymous predicates are not interchangeable everywhere.

In the face of this dilemma Goodman takes the alternative of declaring that "the relation of exact synonymy between diverse predicates is null." This is to say that no two predicates (or expressions of any kind, presumably) are "exactly synonymous." To many it has seemed more reasonable to abandon the interchangeability criterion. If no two expressions are synonymous or mean exactly the same thing, it is hard to see how the expressions "synonymous expressions" and "mean exactly the same thing" could have any currency in our language. Is it really credible that whenever we say two expressions are synonymous we are wrong? Is it not much more likely that the philosophers who discuss these issues have supposed that our concepts are governed by criteria which in fact do not apply? Consider a dictionary of synonyms. Is it credible that it is wrong in every entry because no two terms are synonymous? Surely not.

The above, or something like it, represents the response of the ordinary-language philosophers to the radical conclusions discussed in the earlier parts of this article. Such philosophers observe that a pair of terms may be regarded as synonymous "for certain purposes." This requires that they be interchangeable not everywhere but only in contexts relevant to the given discussion. It is wrong, these philosophers argue, to treat language as though it were a calculus governed by exact rules. But it is one thing to complain that the philosophers have distorted our actual use of the concept of synonymity and quite another to supply a careful and complete account of what that use is. Such an account remains to be given.

Bibliography

Several of the papers cited below are reprinted in *Semantics and the Philosophy of Language,* edited by Leonard Linsky (Urbana: University of Illinois Press, 1952).

ANALYTIC-SYNTHETIC DISTINCTION

Quine, W. V. "Two Dogmas of Empiricism." In *From a Logical Point of View.* Cambridge, MA: Harvard University Press, 1953. Ch. 2.

White, Morton. "The Analytic and the Synthetic." In Linsky, *Semantics* (see above). Ch. 14.

White, Morton. *Toward Reunion in Philosophy.* Cambridge, MA: Harvard University Press, 1956.

INTERCHANGEABILITY

Frege, Gottlob. "On Sense and Reference." In *Translations from the Philosophical Writings of Gottlob Frege,* edited by Max Black and P. T. Geach. New York: Philosophical Library, 1952.

Mates, Benson. "Synonymity." In Linsky, *Semantics* (see above). Ch. 7.

Quine, W. V. "Notes on Existence and Necessity." In Linsky, *Semantics* (see above). Ch. 5.

INTENSIONAL ISOMORPHISM

Carnap, Rudolf. *Meaning and Necessity.* Chicago: University of Chicago Press, 1947.

Carnap, Rudolf. "A Reply to Leonard Linsky." *Philosophy of Science* 16 (4) (1949): 347–350.

Church, Alonzo. "Intensional Isomorphism and Identity of Belief." *Philosophical Studies* 5 (5) (1954): 65–73.

Church, Alonzo. "On Carnap's Analysis of Statements of Assertion of Belief." *Analysis* 10 (1950): 97–99.

Lewis, C. I. *An Analysis of Knowledge and Valuation.* La Salle, IL: Open Court, 1946.

Lewis, C. I. "The Modes of Meaning." In Linsky, *Semantics* (see above). Ch. 3.

Linsky, Leonard. "Some Notes on Carnap's Concept of Intensional Isomorphism and the Paradox of Analysis." *Philosophy of Science* 16 (4) (1949): 343–347.

Putnam, Hilary. "Synonymity and the Analysis of Belief Sentences." *Analysis* (April 1954): 114–122.

Sellars, Wilfrid. "Putnam on Synonymity and Belief." *Analysis* (1955): 117–120.

GOODMAN'S THEORY

Goodman, Nelson. "On Likeness of Meaning." In Linsky, *Semantics* (see above). Ch. 4.

Goodman, Nelson. "On Some Differences about Meaning." *Analysis* (March 1953): 90–96.

Rudner, Richard. "A Note on Likeness of Meaning." *Analysis* 10 (1950): 115–118.

Thomson, James. "Some Remarks on Synonymy." *Analysis* 12 (1952): 73–76.

ORDINARY-LANGUAGE VIEW

Rollins, C. D. "The Philosophical Denial of Sameness of Meaning." *Analysis* 11 (1950): 38–45.

Shwayder, David. "Some Remarks on 'Synonymity' and the Language of Semanticists." *Philosophical Studies* 5 (1) (1954): 1–5.

Leonard Linsky (1967)

SYNONYMITY [ADDENDUM]

Intuitively, two expressions are synonymous if and only if (iff) they have the same meaning. Despite the apparently straightforward nature of this definition, the notion of synonymy has been hard pressed in contemporary philosophy of language. Difficulties arise from two directions: general skepticism about intensional semantics and specific concerns involving substitution into intensional contexts.

QUINE AGAINST INTENSIONAL SEMANTICS

In "Two Dogmas of Empiricism" (1951), W. V. O. Quine leveled an influential attack on intensional semantic concepts, concepts that express meaning relations (e.g., "analytic," "synonymous," and "antonymous"). While Quine acknowledged that such concepts are as a group interdefinable, he argued that no members of the class can be made philosophically respectable on empiricist principles.

His argument can be stated as follows. To be acceptable, semantic concepts must be definable in terms that are either formal (i.e., purely logical or linguistic) or observational. Quine first argued that there is no noncircular, nonarbitrary formal definition of the relevant semantic concepts. Thus, any definition will have to be stated in observational terms. But the semantic concepts under consideration presuppose that sentences have meanings individually. Yet, except at the theoretical periphery, where one finds observation sentences, observation does not bear on sentences one by one (holism). Consequently, there is no way to assign observational meanings to sentences individually. It follows that no acceptable definition of intensional concepts is possible.

MATES'S PUZZLE

In "Two Dogmas," Quine noted the possibility of defining an analytic statement as one that may be turned into a logical truth by replacing synonyms with synonyms. Quine then considered the possibility of defining synonymy in terms of substituting *salve veritate* (i.e., substituting without changing the truth value) in all nonquotational contexts. The suggestion is a natural one, since substitution of synonymous expressions ought to preserve compositional meaning, which in turn ought to preserve truth. Moreover, Quine's reasons for rejecting this proposal were not particularly forceful.

Benson Mates (1952) soon formulated a powerful and independently puzzling argument against substitution *salve veritate* as an adequate basis on which to define synonymy. Take any two purportedly synonymous expressions, say "chew" and "masticate." Now consider the following truism:

> (1) Nobody doubts that whoever believes that *x* chews, believes that *x* chews.

According to the proposed definition of synonyms, (1) implies (2):

> (2) Nobody doubts that whoever believes that *x* chews, believes that *x* masticates.

But clearly (2) may be false even though (1) is true, as when someone is unsure whether or not "chew" and "masticate" are synonymous. Consequently, "chews" and "masticates" are not synonymous. Moreover, this same argument will work for *any* pair of purportedly synonymous expressions.

Mates's puzzle is philosophically interesting apart from the question of whether or not substitution *salve veritate* underwrites an adequate definition of synonymy. For, in a compositional semantics for a language, substitution *salve veritate* appears to be at least a necessary condition for any pair of words to be synonymous.

RESPONSES

Attempts to revive a definition of synonymy have relied on hidden quotation, have sought to define synonymy in terms of responses to stimuli, and have involved rejecting the requirement of a reductive definition.

METALINGUISTIC RESPONSES. A natural first response to Mates's puzzle is to treat sentences like (1) and (2) as in some way covertly involving direct quotation. Wilfrid Sellars (1955), for instance, thought that sentences (1) and (2) should be reinterpreted as (1*) and (2*):

> (1*) Nobody doubts that whoever believes "*x* chews," believes "*x* chews."

> (2*) Nobody doubts that whoever believes "*x* chews," believes "*x* masticates."

Church (1954) offers a slightly more complex variant of this approach.

Despite their initial appeal, such metalinguistic responses do not appear to do justice the issue. For instance, while (1*) and (2*) explicitly involve English expressions, the original sentences do not; there is not even the presumption that people who are counterexamples to (2) must speak English. On this point, Tyler Burge (1978) seems quite right to note that while linguistic considerations may well be involved in such claims as (1) and (2), this does not show that these considerations enter into the content of the attitude report.

NEO-QUINEANISM. A second approach to synonymy derives from the work of Quine himself. Despite his attack on intensional semantics, Quine (1960) was able to preserve a vestigial concept of synonymy. He called two sentences *stimulus synonymous* for a speaker at a particular time iff the speaker would accept or reject them both under the same range of observational conditions. (A similar but more tenuous definition can be given for words.) Yet for many sentences, our assent or dissent does not depend on observation. Quine's concept of stimulus synonymy is far less widely applicable than the intuitive notion.

Peter Pagin (2001) attempted to extend this general sort of definitional strategy. According to Quine's holism, sentences may be partially ordered by how closely tied they are to observation. Observation sentences are either accepted or rejected on the basis of current observation. Most of the remaining sentences of the language, however, are assigned truth values more or less likely to be revised in light of further observation. On this basis, Pagin defined two statements *A* and *B* to be equally revisable ($=_r$), for a speaker at a particular time, as follows: $A =_r B$ iff for any statement C, $A <_r C$ iff $B <_r C$. Equivalently, $A =_r B$ iff for any statement C, $C <_r A$ iff $C <_r B$. Here $<_r$ is the relation of being less revisable than (for a speaker at a time). Pagin then offers the following definition (in which $A(\alpha/\beta)$ is any statement that results from substituting α for β in A, not necessarily uniformly):

> Expressions α and β are synonymous iff for any statement A, $A =_r A(\alpha/\beta)$.

Pagin's definition runs into problems at the level of statement synonymy. Let *A* and *B* be two distinct sentences that happen to be equally prone to revision. Then, setting $\alpha = B$ and $\beta = A$, we have it that any two statements that are equirevisable are synonymous. But surely it is possible to have two nonsynonymous statements that are equally prone to revision in light of recalcitrant data. In addition, Pagin's definition appears to flounder on a variant of Mates's argument. For let α = "masticates" and β = "chews" and let *A* = sentence (1) above. Then (2) will arise by substitution of α for β in *A*. But (1) and (2) are plainly not equally revisable, and so fail to qualify as synonyms on the proposed definition. And this result will clearly generalize to any pair of distinct expressions. This result is not surprising, for if substitution of synonyms cannot preserve truth, it can hardly be expected to preserve revisability.

NEOINTENSIONALISM. A final approach to restoring a definition of synonyms involves rejecting Quine's demand for a reductive definition altogether. Over the years, Jerrold Katz has developed a distinctive non-Fregean version of this approach. Katz's neointensionalism (2004) consists of two major theses: (i) Expressions of the language have a sense structure specified in terms of their parts. (ii) The sense structure of an expression is specified independently of its referential properties. Thus, senses are not modes of presentation. Rather, they constitute an *autonomous* semantic level posited (on the basis of the judgments of competent speakers of the language) to account for the sense properties of expressions (e.g., being meaningful, being synonymous). On such a view, it is straightforward to define synonymy in terms of having the same meaning (sense), since there is no further requirement to analyze meanings in terms of nonintensional concepts.

But clearly it will not do to allow the two semantic levels (sense and reference) to come apart completely. We cannot have an expression that, for instance, is synonymous with the definite description "the first celestial body visible in the evening" but that refers to, say, Margaret Thatcher. Consequently, Katz proposed that while sense does not *determine* reference, it does *mediate* it; that is, having a sense is necessary (though maybe not sufficient) for reference. The picture that Katz paints is one where we develop an autonomous theory of sense on the evidence

of competent speakers' use of the language and then use that theory to constrain our theory of reference.

Katz's semantic theory does not appear, however, to avoid many of the objections that led to the downfall of its Fregean predecessors. Consider, for example, the fact that ordinary competent speakers of a language are occasionally radically mistaken about the nature of the entities about which they are speaking. Jonathan Cohen (2000), for instance, notes that in the past the best evidence from competent native speakers of English would have supported the hypothesis that the kind term "whale" included as a component of its sense the semantic marker FISH. Intuitively, however, those speakers were still referring to the same natural kind (the whale) as we do. But this judgment is inconsistent with Katz's proposal.

See also Analytic and Synthetic Statements; Analyticity; Meaning.

Bibliography

Burge, Tyler. "Belief and Synonymy." *Journal of Philosophy* 75 (1978): 119–138.

Cohen, Jonathan. "Analyticity and Katz's New Intensionalism: or, If You Sever Sense from Reference, Analyticity Is Cheap but Useless." *Philosophy and Phenomenological Research* 61 (2000): 115–135.

Church, Alonzo. "Intensional Isomorphism and Identity of Belief." *Philosophical Studies* 5 (1954): 65–73.

Katz, Jerrold. *Sense, Reference, and Philosophy*. Oxford, U.K.: Oxford University Press, 2004.

Mates, Benson. "Synonymity." In *Semantics and the Philosophy of Language*, edited by Leonard Linsky. Urbana: University of Illinois Press, 1952.

Pagin, Peter. "A Quinean Definition of Synonymy." *Erkenntnis* 55 (2001): 7–32.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Sellars, Wilfrid. "Putnam on Synonymity and Belief." *Analysis* 15 (1955): 117–120.

Marc A. Moffett (2005)

SYNTACTICAL AND SEMANTICAL CATEGORIES

The basis for any theory of syntactical categories is the linguistic fact that in all natural languages there are strings of (one or more) words which are mutually interchangeable in all well-formed contexts *salva beneformatione*—that is, with well-formedness (grammaticality, syntactical correctness) being preserved in the interchange—and that there are innumerable other strings which do not stand in this relation. Any theory of semantical categories rests on a similar fact, with *well-formed* replaced by *meaningful* or *semantically correct*, and *beneformatione* by *significatione*.

The relation between *well formed* and *meaningful* is, in general, complex, and neither term is simply reducible to the other. The English expression "Colorful green ideas sleep furiously" (to use an example given by Noam Chomsky) is, at least prima facie, syntactically well formed. Yet it is semantically meaningless, even though certain meanings can be assigned to it by special conventions or in special contexts. In contrast, many everyday utterances are syntactically ill formed (because of false starts, repetitions, and the like) but semantically perfectly meaningful, again at least prima facie.

Chomsky and his followers have recently stressed that for natural languages well-formedness and meaningfulness are mutually irreducible, but this view has not gone unchallenged. For constructed language systems, particularly those meant to serve as languages of science, it has generally been assumed that the notions of well-formedness and meaningfulness coincide.

Since the time of Aristotle it has been customary among philosophers to explain the linguistic facts about interchangeability by resort to ontological assumptions. Certain strings of words, it is said, are not well formed (or meaningful) because the entities denoted by the substrings (the meanings, denotata, etc., of these substrings) do not fit together. Edmund Husserl, one of the authors who dealt most explicitly with interchangeability, coined the term *meaning categories* (*Bedeutungskategorien*). He maintained that we determine whether or not two expressions belong to the same meaning category, or whether or not two meanings fit together, by "apodictic evidence." But his examples and terminology—for instance, the use of the expression "adjectival matter" (*adjektivische Materie*)—indicate that his apodictic evidence was nothing more than a sort of unsophisticated grammatical intuition, which he hypostatized as insights into the realm of meanings.

Husserl certainly deserves great credit for distinguishing between nonsense (*Unsinn*) and "countersense" (*Widersinn*), or, in modern terms, between strings that violate rules of formation and strings that are refutable by the rules of deduction. But he is also responsible for the initiation of a fateful tradition in the treatment of semantical (and syntactical) categories. This tradition assumes—sometimes without even noticing the problematic status of the assumption, more often with only

the flimsiest justification—that if two strings are interchangeable in some one context *salva beneformatione*, they must be so in all contexts.

This entry will discuss the chief modern contributions to the theory of syntactical and semantical categories. It will first outline the achievements of the Polish logician Stanisław Leśniewski and his pupil Kazimierz Ajdukiewicz. It will then evaluate the contributions by Rudolf Carnap and, in particular, stress the added flexibility gained by his decision not to adhere to Leśniewski's "main principle." Finally, the synthesis by Yehoshua Bar-Hillel of the insights of Ajdukiewicz and Carnap into a theory of syntactical categories and the demonstration by Chomsky of the essential inadequacy of categorial grammars for a description of the syntactical structure of natural languages will be mentioned.

LEŚNIEWSKI

In 1921, Leśniewski made an attempt to simplify Bertrand Russell's ramified theory of types but was not satisfied with the outcome. A type theory, however simplified and otherwise improved, remained for him an "inadequate palliative." He therefore began, the following year, to develop a theory of semantical categories that had greater appeal to his intuitive insights into the syntactical and semantical structure of "proper" language. For this purpose he turned from Russell to Husserl, of whose teachings he had learned from his teacher and Husserl's pupil, Kazimierz Twardowski, and, in particular, to Husserl's conception of meaning categories. As a prototype of a proper language, to which his theory of semantical categories was to be applied, Leśniewski constructed the canonical language *L*. Husserl's tacit assumption that if two strings are interchangeable in some one context *salva beneformatione*, they must be so in all contexts was elevated to the rank of the "main principle of semantical categories." Today Leśniewski's term *semantical categories* must be regarded as a misnomer, since the categorization was based on purely syntactical considerations. At the time, however, Leśniewski, like many other authors, believed that well-formedness and meaningfulness are completely coextensive for any proper language.

According to Leśniewski, each string, whether a single word or a whole phrase, of a proper language, and hence of his canonical language *L*, belongs to at most one category out of an infinitely extensible complex hierarchy. Strings are understood as tokens rather than as types. Moreover, two equiform tokens may well belong to different categories. This homonymy, however, never leads to ambiguity, since in any well-formed formula the context always uniquely determines the category of the particular token. In fact, Leśniewski exploited this homonymy for systematic analogy, with an effect similar to that obtained by Russell's exploitation of the typical ambiguity of strings (qua types).

Leśniewski excluded from the hierarchy only strings outside a sentential context, terms inside quantifiers binding variables, and parentheses and other punctuation signs. Defined constants were automatically assigned to categories by means of "introductory theses," as Leśniewski called those object-language sentences which, in his view, served to introduce new terms into an existing language. He gave rigid directives for the formation of introductory theses, assignment to a category being valid only after these theses were specified. The constructive relativity thus introduced was intended to take the place of the order restrictions by which Russell had sought to avoid the semantical antinomies.

In his canonical language Leśniewski worked with two basic categories, "sentences" and "nominals," and a potential infinity of functor categories. He admitted only indicative sentences; interrogatives, imperatives, hortatives, and the like were excluded. He explicitly rejected any categorial distinction between proper names and common nouns or between empty, uniquely denoting, and multiply denoting nominal phrases, although he later drew these distinctions on another basis. In the notation subsequently devised by Ajdukiewicz the category, say, of the sentential negation sign (that is, of a functor which, from a sentence as argument, forms a complex expression itself belonging to the category of sentences) is denoted by its "index" "*s/s*." The denominator of this "fraction" indicates the category of the argument and the numerator that of the resulting string. The index of such binary connectives as the conjunction sign is "*s/ss*." With "*n*" as the category index of nominals, "*n/n*" is assigned to "attributive adjectives" (but also to "nominal negators" such as "non-____"), "*s/n*" to "predicative intransitive verbs," "*s/nn*" to "predicative transitive verbs," "*s/n//s/n*" to certain kinds of "verbal adverbs," etc.

AJDUKIEWICZ

With the help of this notation Ajdukiewicz was able to formulate, in 1935, an algorithm for the determination of the syntactical structure of any given string in certain languages and, in particular, of its "syntactical connexity"—that is, its well-formedness. These languages had to embody, among other conditions, the Polish notation, in which functors always precede their arguments (thereby

freeing parentheses from their customary duty as scope signals and making them available for other duties) and had to be "monotectonic," in H. B. Curry's later terminology—that is, to allow just one structure for each well-formed formula. These conditions of course excluded the natural languages from coming under Ajdukiewicz's algorithm.

To illustrate: Let

Afagbc

be a string in a given language fulfilling the above conditions. Let "*n*" be the index of "*a*," "*b*," and "*c*," let "*s/n*" be the index of "*f*," let "*s/nn*" be the index of "*g*," and let "*s/ss*" be the index of "*A*." The index string corresponding to the given string is, then,

A	*f*	*a*	*g*	*b*	*c*
s/ss	*s/n*	*n*	*s/nn*	*n*	*n.*

Let the only rule of operation be the following: replace $\alpha/\beta\beta$ (where α and β are any index or string of indexes) by α (always applying the rule as far "left" as possible). One then arrives in two steps at the "exponent" "*s*," thus verifying that the given string is a sentence with the "parsing" *(A(fa)(gbc))*. The whole derivation can be pictured as follows:

```
s/ss  └s/n   n┘  s/nn   n    n
          |
s/ss      s     └s/nn   n    n┘
                        |
└s/ss     s             s┘
          |
          s
```

In 1951, Bar-Hillel adapted Ajdukiewicz's notation to natural languages by taking into account the facts that in such languages arguments can stand on both sides of the functor, that each element, whether word, morpheme, or other appropriate atom in some linguistic scheme, can be assigned to more than one category, and that many well-formed expressions will turn out to be syntactically ambiguous or to have more than one structural description. These changes greatly increased the linguistic importance of the theory of syntactical categories and initiated the study of a new type of grammars, the so-called categorial grammars.

Ajdukiewicz never questioned the validity of Leśniewski's main principle. Neither did Alfred Tarski at first. It was taken for granted in the main body of Tarski's famous 1935 paper, "Der Wahrheitsbegriff in den formalisierten Sprachen." (The concept of truth in formalized languages; whose Polish original dates from 1931.) The appendix to this paper voiced some doubts as to its intuitive appeal, but these doubts probably derived more from a growing preference for set-theoretical logics over type-theoretical ones than from straight linguistic considerations.

CARNAP

Rudolf Carnap, in *Der logische Aufbau der Welt* (1928), had few misgivings about applying the simple theory of types to natural languages. Like Russell, he made a halfhearted attempt to provide a quasilinguistic justification for the type hierarchy, and his notion of "spheres" (*Sphären*) occupies a position approximately midway between Russell's types and Leśniewski's semantical categories. Carnap's explanation of certain philosophical pseudo problems as based on a "confusion of spheres" (*Sphärenvermengung*) antedates Gilbert Ryle's discussion of "category mistakes" in his *Concept of Mind* (London, 1949) by more than twenty years. Both explanations rest on an uncritical implicit adherence to the "main principle," even though Leśniewski's formulation was not known to Carnap at the time he wrote his book, probably because Leśniewski's publications prior to 1929 were all in Russian or Polish. At the same time, neither Leśniewski, Ajdukiewicz, nor Tarski quotes Carnap's book in their pertinent articles. Ryle, in his book, does not mention any of these publications.

Carnap was apparently the first logician to use the term *syntactical categories*, in 1932. At that time he believed that all logical problems could be treated adequately as syntactical problems, in the broad sense he gave the term.

He was also the first to free himself from the main principle. It eventually occurred to him that this principle embodied an arbitrary restriction on freedom of expression. Any attempt to impose this restriction on natural languages resulted in an intolerable and self-defeating proliferation of homonymies, similar to the outcome of the attempt by Russell and some of his followers to impose type-theoretical restrictions on natural languages, other than the tolerable "typical" ambiguities. In some cases it sounded rather natural to invoke equivocation (which is, of course, a "nontypical" ambiguity)—in the tradition of Aristotle, who used this notion to explain the deviancy of "The musical note and the knife are sharp." But in innumerable other cases there were no independent reasons for such invocation, and the induced artificialities exploded the whole structure. For instance, very strong reasons seem to be required if one

were to assign the string "I am thinking of" to a different type or syntactical category each time the string following it belonged to a different type or category. For one may have after "I am thinking of" such varied strings as "you," "freedom," "the theory of syntactical categories," and "the world going to pieces."

In 1934, in *Logische Syntax der Sprache*, Carnap took implicit account of the possibility that two strings might be interchangeable in some contexts but not in all. He coined the term *related* for this relation and used *isogenous* for the relation of total interchangeability. Languages in which all strings are either pairwise isogenous or unrelated have, in this respect, a particularly simple structure. But there is no reason to assume that natural languages will exhibit this particularly simple structure. In fact, observing the main principle becomes a nuisance even for rich constructed language systems; as Carnap showed, the principle is not observed in some of the better-known calculi (perhaps contrary to the intention of their creators) with no real harm done.

BAR-HILLEL AND CHOMSKY

The relation "related" is clearly reflexive and symmetrical; hence, it is a similarity relation. The relation "isogenous" is, in addition, transitive; hence, it is an equivalence relation. Starting from these two relations, Bar-Hillel, in 1947, developed a theory of syntactical categories, illustrated by a series of model languages, all of which were, in a certain natural sense, sublanguages of English. In 1954, Chomsky developed a more powerful theory by taking into account, in addition, relations between the linguistic environments of the strings compared.

Recently, primarily owing to the insights of Chomsky and coming as a surprise to most workers in the field, it has become clear that interchangeability in context cannot by itself serve as the basic relation of an adequate grammar for natural languages. It may play this role for a number of constructed languages, and it certainly does so, for example, in the case of the standard propositional calculi. More exactly, it provides a satisfactory basis for what have become known as "phrase-structure languages," or what Curry calls "concatenative systems."

A phrase-structure language is a language (a set of sentences) determined by a phrase-structure grammar, the grammar being regarded as a device for generating or recursively enumerating a subset of the set of all strings over a given vocabulary. A phrase-structure grammar, rigorously defined, is an ordered quadruple $\langle V,T,P,S\rangle$, where V is a finite vocabulary, T (the terminal vocabulary) is a subset of V, P is a finite set of productions of the form $X \rightarrow x$ (where X is a string over $V–T$, the auxiliary vocabulary, and x is a string over V consisting of at least one word), and S (the initial string) is a distinguished element of $V–T$. Any terminal string (string over T) that can be obtained from S by a finite number of applications of the productions is a sentence. When the X's in all the productions consist of only one word the grammar is called a context-free, or simple, phrase-structure grammar.

Interchangeability in context seems also to be adequate for describing the surface structure of all English sentences but not for describing their "deep structure." It is powerful enough to enable us to analyze correctly the sentence "John loves Mary" (S) as a concatenate of a noun phrase (NP), consisting in this particularly simple illustration of a single noun (N), and a verb phrase (VP), consisting of a transitive verb (Vt) and another noun phrase itself consisting of a noun. Two customary representations of this analysis are the "labeled bracketing,"

$$({}_S({}_{NP}({}_N\text{John})({}_{VP}({}_{Vt}\text{loves})({}_{NP}({}_N\text{Mary}))))),$$

and the "inverted tree,"

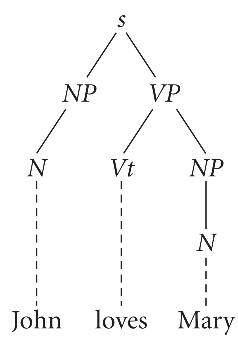

(both representations are simplified for present purposes). Interchangeability in context is likewise powerful enough to provide "Mary is loved by John" with the correct structure,

$$({}_S({}_{NP}({}_N\text{Mary})({}_{VP}({}_{\text{Pass}Vt}\text{is}({}_{Vt}\text{love})\text{-ed by})({}_{NP}({}_N\text{John}))))).$$

However, these analyses will not exhibit the syntactically (and semantically) decisive fact that "Mary is loved by John" stands in a very specific syntactical relation to "John loves Mary," namely that the former is the passive of the latter. No grammar can be regarded as adequate that does not, in one way or another, account for this fact. Transformational grammars, originated by Zellig Harris and considerably refined by Chomsky and his associates, appear to be in a better position to describe the deep structures of these sentences and of innumerable others.

Such grammars adequately account for the relation between the active and passive sentences and explain the fact that one intuitively feels "John" to be in some sense the subject of "Mary is loved by John," a feeling often expressed by saying that "John," though not the "grammatical" subject, is still the "logical" subject of the sentence. Transformational analysis shows that "John," though indeed not the subject in the surface structure of the given sentence, is the subject of another, underlying sentence of which the given sentence is a transform.

It has recently been proved that categorial grammars and context-free phrase-structure grammars are equivalent, at least in the weak sense of generating the same languages qua sets of sentences over a given vocabulary, though perhaps not always assigning the same structure(s) to each sentence. These sets can also be generated (or accepted) by certain kinds of automata, the so-called push-down store transducers. The connection that this and other results establish between algebraic linguistics and automata theory should be of considerable importance for any future philosophy of language.

DEVELOPMENTS IN THE 1960S

The early 1960s witnessed a revival of interest in the semantical categorization of expressions in natural languages, mostly under the impact of the fresh ideas of Chomsky and his associates. The whole field of theoretical semantics of natural languages is still very much in the dark, with innumerable methodological and substantive problems unsolved and sometimes hardly well enough formulated to allow for serious attempts at their solution. However, there is now a tendency to include indexes of semantical categories in the lexicon part of a complete description of such languages. These indexes, after application of appropriate rules, determine whether a given string is meaningful and, if it is, what its meaning is in some paraphrase of standardized form or, if it is not, how it deviates from perfect meaningfulness. In addition to semantical category indexes there are morphological, inflectional, and syntactical category indexes that determine whether the given string is morphologically and syntactically completely well formed, that present its syntactical structure in some standardized form, or that indicate the ways in which it deviates from full well-formedness.

Whether at least some semantical categories can, or perhaps must, be considered in some sense universal (language-independent) is a question that, like its syntactical counterpart, is now growing out of the speculative stage, with the first testable contributions beginning to appear. Investigations by Uriel Weinreich (1966) have cast serious doubts on the possibility of making a clear distinction between syntactical and semantical categories. Should these doubts be confirmed, the whole problem of the relation between these two types of categories will have to be reexamined.

See also Categories; Semantics, History of; Type Theory.

Bibliography

Ajdukiewicz, Kazimierz. "Die syntaktische Konnexität." *Studia Philosophica* 1 (1935): 1–27.

Bar-Hillel, Yehoshua. *Language and Information*. Reading, MA: Addison-Wesley, 1964.

Carnap, Rudolf. *Der logische Aufbau der Welt*. 2nd ed. Hamburg, 1961.

Carnap, Rudolf. *Logische Syntax der Sprache*. Vienna: Springer, 1934. Translated by Amethe Smeaton as *The Logical Syntax of Language*. New York: Harcourt Brace, 1937.

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.

Chomsky, Noam. "Formal Properties of Grammars." In *Handbook of Mathematical Psychology*, edited by R. Duncan Luce, R. R. Bush, and E. Galanter. Vol. II. New York: Wiley, 1963. Ch. 12.

Curry, H. B. *Foundations of Mathematical Logic*. New York: McGraw-Hill, 1963.

Husserl, Edmund. *Logische Untersuchungen*. 2nd ed., Vol. II. Halle: Niemeyer, 1913.

Luschei, E. C. *The Logical Systems of Leśniewski*. Amsterdam: North-Holland, 1962.

Suszko, Roman. "Syntactic Structure and Semantic Reference." *Studia Logica* 8 (1958): 213–244, and 9 (1960): 63–91.

Tarski, Alfred. *Logic, Semantics, Metamathematics*. Translated by J. H. Woodger. Oxford: Clarendon Press, 1956. Contains a translation of "Der Wahrheitsbegriff in den formalisierten Sprachen."

Weinreich, Uriel. "Explorations in Semantic Theory." In *Current Trends in Linguistics*, edited by Thomas A. Sebeok, Vol. III. The Hague: Mouton, 1966.

Yehoshua Bar-Hillel (1967)

SYNTACTICAL AND SEMANTICAL CATEGORIES [ADDENDUM]

Categorical distinctions in syntax and semantics are drawn on the basis of the distribution of linguistic expressions. According to the classical definitions, two expressions belong to the same syntactic category just in case they can be interchanged in every well-formed context *salva beneformatione* (without loss of well-formedness) and they belong to the same semantic category just

in case they can be interchanged in every meaningful context *salva significatione* (without loss of meaningfulness) (Bar-Hillel 1953). The question is what counts as interchange in a context. Interpreting this phrase naïvely will result in inadequate definitions: Intuitively, one cannot conclude that "You are bald" and "I am bald" belong to different syntactic categories because "Most people who like you are bald" is well-formed, whereas "Most people who like I am bald" is not. Likewise, it cannot be concluded that *cat* and *dog* belong to different semantic categories from the fact that *location* is meaningful, whereas *lodogion* is not. Interchange of non-constituents is irrelevant for syntactic or semantic categorization.

One might try to modify the classical definition minimally, by saying that two expressions belong to the same syntactic category just in case they occupy the same range of syntactic positions within well-formed complex expressions, and they belong to the same semantic category just in case they occupy the same range of syntactic positions within meaningful complex expressions. But this still leads to excessively fine-grained categories. If one insists that *book* and *books* must belong to different syntactic categories because the first but not the second can occur with the indefinite article, or that *year* and *century* belong to distinct semantic categories because the morpheme *-ly* can attach meaningfully to the first but (arguably) not the second, they will miss a number of crucial generalizations.

The most useful categories will group together linguistic expressions that share much, but perhaps not all their distribution. Consider, for example, the syntactic category of prepositions. A distributional pattern used in identifying these is that prepositions can be intensified by *right* or *straight*, whereas other kinds of words cannot. "He went straight down the ladder" and "He lives right in the center of town," for instance, are well-formed, whereas "He went straight crazy" and "He is right the center of attention" are ungrammatical. This is an important test even though it yields both false negatives and false positives. "He went straight home" and "He came right with a friend." The exceptions can be neglected because *home* is idiosyncratic (for some reason "He went to home" is ungrammatical, but "He went to his home" is fine), and the facts about *with* allow for a different explanation (*with* cannot be intensified with *straight* or *right* because, given its meaning, it cannot be intensified at all).

One common way to achieve flexibility in talking about the distribution of linguistic expressions is to introduce *features*. Features are properties of words and morphemes that are marked in the lexicon (according to a common conception of lexical items, they are nothing more than structured bundles of features). Complex expressions inherit some of the features of their constituents. Some features are both semantically and syntactically significant (e.g. [PAST] which is a property of 'loved Hugo' but not 'loves Hugo'), some are syntactically but not semantically significant (e.g. [ACCUSATIVE] which is a property of the first person pronoun in 'They want me to come' but not in 'They want that I come'), and some are semantically but not syntactically significant (e.g. [ADULT], which is a property of 'horse' but not of 'foal').

Classical definitions can be modified by saying that two expressions belong to the same syntactic (semantic) category whenever they share a syntactically (semantically) significant set of features. This allows someone to say, for example, that nouns form a single syntactic category exhibiting important subcategorial distinctions (e.g., between proper nouns and common nouns, or between count nouns and mass nouns) and also cross-categorical similarities (e.g., prepositions and nouns do not allow the prefix "un-).

CATEGORIAL GRAMMARS

Even if certain differences of distribution are allowed within categories, about a dozen of them are still needed, and if significant subcategories are taken into account, the tally will go well above 100. To systematize these, one needs to involve them in describing the syntax and the semantics of the language (or languages) to which they belong. One of the simplest conceivable ways this could be done is through a *categorial grammar* (Ajdukiewicz 1967, Bar-Hillel 1953, Lambek 1958)

Categorial Grammars make the following four fundamental assumptions. First, words and morphemes are assigned, *pace ambiguity*, a single syntactic category in the lexicon. Second, there are a few basic syntactic categories, and the rest are derived through a few schemata. Categorial grammars vary widely in what they allow as basic categories, but for the sake of illustration, let us take the base consisting of the categories *S*, *N*, and *NP*—the category of sentences, nouns, and noun phrases. All categorial grammars include derived types that can be generated by the schema: If *A* and *B* are categories, so is *A*/*B*. Third, there are a few syntactic operations, including the one for right-concatenation: If *e* is of syntactic category *A*/*B* and *e′* is of syntactic category *B*, then there is an expression *ee′* of syntactic category *A*. Fourth, every expression within a given syntactic category has the same type of semantic value, and its semantic category is determined by its type.

Again, categorial grammars differ in the system of semantic types they assume, but a fairly typical idea is to assign the type t to S, the type $\langle e,t\rangle$ to N, the type $\langle\langle e,t\rangle,t\rangle$ to NP, and to lay down the schema that if the type of A is α and the type of B is β then the type of A/B is (β,α). If the semantics is extensional, semantic values of type e are entities, semantic values of type t are truth values, and semantic values of type (β,α) are functions from semantic values of type β to semantic values of type α (note that although the syntactic categories N and NP are basic, their semantic types are not). This divergence could be avoided if we had a basic syntactic category P of proper names with the associated semantic type e—N and NP could then perhaps be identified with the derived syntactic categories S/P and $S/(S/P)$, respectively (Ajdukiewicz 1967). Assigning the semantic type e to proper names would seem to be desirable anyway—it seems plausible that if semantics is extensional, the semantic value of *John* should be John himself, not the (characteristic function of the) set of (characteristic functions of) sets containing John. But if proper names and noun phrases are allowed to belong to different syntactic categories, it must be explained why their distribution is so very similar, which is why Montague (1973) decided against this option. One could get around this difficulty by introducing type-lifting rules (Hendriks 1987, Partee 1987).

In a minimal categorial grammar, the only syntactic rule is right-concatenation (interpretated as functional application). This is clearly inadequate because it cannot capture syntactic generalizations about word order. A natural idea to rectify this shortcoming is to introduce the schemata that if A and B are categories, then so is $A\backslash B$, and that if e is of syntactic category $A\backslash B$ and e' is of syntactic category A then there is an expression $e'e$ of syntactic category B (left-concatenation is also interpreted as functional application). But the resulting framework is still much too restrictive.

Until the very end of the 1960s it was widely assumed that categorial grammars are inadequate as syntactic theories of natural languages. There have been three basic strategies to challenge this attitude. The first is to add a transformational component to categorial grammars, whereas the second involves adding free permutations and propose syntactic filters to eliminate the ungrammatical (but interpretable) expressions. (For the former strategy see Lewis [1970]; for the latter see Cresswell [1973]). Both of these lines concede that categorial grammar is incomplete and perhaps nonexplanatory as a syntactic theory, but they argue that it still is the best structure to base compositional semantics on. The third strategy is more ambitious: It extends the set of permissible syntactic operations beyond concatenation and thereby seeks to achieve descriptive adequacy and explanatory power. This is the avenue most categorial grammarians have followed since the early 1970s.

The most important extension of permissible syntactic operations is the introduction of a different sort of concatenation—one that is not interpreted as functional application, but as functional composition. The simplest one of these is: If e is of syntactic category A/B and e' is of syntactic category B/C, then ee' is of syntactic category A/C; if the semantic value of e is the function f and the semantic type of e' is the function g, then the semantic value of ee' is a function h such that for every x: $h(x)=f(g(x))$.

A categorial grammar can have mixed composition rules as well, allowing the composition of a left-slash category with a right-slash category. These rules allow for the construction of sentences containing more than one quantifier to have different derivational histories, which in turn can account for scope ambiguities. They also open up the possibility to construct and interpret non-constituents (such as "Ron loves" in "Ron loves spinach"), which in turn allows categorial grammar to deal with difficult coordination phenomena, such as "Ron loves and Mia hates spinach" (Dowty 1987). All this is done without the introduction of a separate level of logical form with phonologically empty elements. Obviously, to prevent overgeneration, the application of composition rules must be tightly constrained (Steedman 1987).

Although very much a minority view among syntacticians, categorial grammar can explain a good deal about the structure of natural languages; for a survey, see Jacobson (1996). The attempt to do away with any structure other than what is visible on the surface is philosophically intriguing, especially considering that it is often supposed to lead to the complete elimination of variables (Szabolcsi 1989, Jacobson 1999). That the elimination of variables from certain logical languages was possible without limiting their expressive power is well-known from Quine (1966), but the claim that all sentences lack these devices may have much more significant consequences for philosophy. For one thing, those who believe in being ontologically committed to the values of the variables that are quantified over must find where these variables are. Quine's answer was that they are within the formulae of a formal first-order language that are associated with sentences through regimentation—a process where the outcome depends on one's particular interests. The result was the doctrine of ontological relativity (Quine 1969).

STRUCTURAL VALIDITY

Suppose categorial grammarians are empirically adequate and there is no need for a separate level of logical form to account for syntactic generalizations. There still might be the need for logical forms to distinguish between lexical and structural entailments. Consider, for example, the contrast between "Lou is a bachelor; therefore Lou is unmarried" and "Martin walked quickly; therefore Martin walked." Intuitively, the first is valid in part because of what *bachelor* means, whereas the validity of the second is independent of the lexicon. The established account of structural validity rests on logical form: An entailment is structurally valid if, and only if, it is valid in virtue of logical form (i.e., if, and only if, the logical forms of the premise(s) logically entail(s) the logical form of the conclusion). For example, if the (simplified) logical forms within the second entailment are:

$$\frac{\exists e\ \ (\text{walk}\ (\text{Martin}, e) \wedge \text{quick}\ (e))}{\exists e\ \ \text{walk}\ (\text{Martin}, e),}$$

As Davidson (1967) has argued, the entailment is indeed structurally valid on the established account. However, if there is no separate logical form, structural validity must be understood in a different manner.

The obvious thing to say is that semantic categories can provide a definition of structural validity without taking a detour through logical forms. One can say that an entailment is structurally valid just in case any uniform substitution of expressions of the same semantic category within it results in a valid entailment (Evans 1976). If, as categorial grammar assumes, syntactic categories are associated with a unique semantic type, which in turn determines semantic category, one may replace *semantic* with *syntactic* in the above definition. An interesting consequence of this definition is that logical consequence expressed in natural language (setting aside cases like "Hugo walks; therefore Hugo walks") will not be structural. But this is arguably as it should be: the inference "Hugo walks and talks; therefore Hugo walks" is valid in part because of what *and* means—replace it with *or* and the resulting entailment is no longer valid. Although logical entailments are said to be valid in virtue of their form, except for the special case of concluding something from itself, their validity also rests upon the lexical meaning of logical constants.

See also Semantics; Syntax.

Bibliography

Ajdukiewicz, K. "Die syntaktische konnextität." In *Polish Logic, 1920–1939*. Translated by S. McCall. Oxford: Clarendon, 1967.

Bar-Hillel, Y. "A Quasi-Arithmetical Notation for Syntactic Description." *Language*, 29 (1953): 47–58.

Cresswell, M. *Logics and Languages*. London: Methuen, 1973.

Davidson, D. "The Logical Form of Action Sentences." In *The Logic of Decision and Action*, edited by N. Rescher. Pittsburgh, PA: University of Pittsburgh Press, 1967.

Dowty, D. "Type Raising, Functional Composition, and Non-constituent Conjunction." In *Categorial Grammars and Natural Language Structures*, edited by R. Oherle. Dordrecht, Netherlands: Reidel, 1987.

Evans, G. "Semantic Structure and Logical Form." In *Truth and Meaning: Essays in Semantics*, edited by G. Evans and J. McDowell. Oxford: Oxford University Press, 1976.

Hendriks, H. "Type Change in Semantics: the Scope of Quantification and Coordination." In *Categories, Polymorphism, and Unification*, edited by E. Klein and J. van Benthem. Amsterdam: University of Amsterdam ITLI, 1987.

Jacobson, P. "The Syntax-Semantics Interface in Categorial Grammar." In *The Handbook of Contemporary Semantic Theory*, edited by S. Lappin. Oxford U.K.: Blackwell, 1996.

Jacobson, P. "Toward A Variable-Free Semantics." *Linguistics and Philosophy* 22 (1999): 117–185.

Lambek, J. "The Mathematics of Sentence Structure." *American Mathematical Monthly* 65 (1958): 154–169.

Lewis, D. "General Semantics." In *Philosophical Papers*. Vol. 1. Oxford: Oxford University Press, 1970.

Montague, R. "The Proper Treatment of Quantification in Ordinary English." In *Formal Philosophy: Selected Papers of Richard Montague*, edited by R.H. Thomason. New Haven, CT: Yale University Press, 1973.

Partee, B. "Noun Phrase Interpretation and Type-Shifting Principles." In *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, edited by J. Groenendijk. Dordrecht, Netherlands: Foris, 1987.

Quine, W. V. O. "Variables Explained Away." *In Selected Logic Papers*. New York: Random House, 1966.

Quine, W. V. O. "Ontological Relativity." In *Ontological Relativity*. New York: Columbia University Press, 1969.

Steedman, M. "Combinatory Grammars and Parasitic Gaps." *Natural Language and Linguistic Theory*, 5 (1987): 403–440.

Szabolcsi, A. "Bound Variables in Syntax: Are There Any?" In *Semantics and Contextual Expression*, edited by R. Bartsch. Dordrecht, Netherlands: Foris, 1989.

Zoltán Gendler Szabó (2005)

SYNTAX

"Syntax" is the theory of the construction of sentences out of words. In linguistics, syntax is distinguished from morphology, or the theory of the construction of words out of minimal units of significance, only some of which are words. According to this division, it is a matter of morphology that the word *solubility* decomposes into

"dissolve" + "able" + "ity"; but it is a matter of syntax to analyze the construction of the sentence, "That substance is able to dissolve."

Although syntax is a traditional grammatical topic, it was only with the rise of formal methods growing out of the study of mathematical logic that the subject attained sufficient explicitness to be studied in depth, in works by Zelig Harris (1957) and Noam Chomsky (1957). Since then a flourishing field has been created; for it was rapidly discovered that the syntax of human languages was far more complex than at first appeared. In this respect, the development of syntax is comparable to other fields of cognitive science such as human vision, problem-solving capacities, and the organization of commonsense knowledge, all of which gave rise to difficult problems once the goal of fully explicit representation was put in place.

The dawn of syntax is marked by the realization that the structure of sentences is hierarchical; that is, that behind the linear order of words and morphemes that is visible in natural languages there is another organization in terms of larger or smaller constituents nested one within another. Description of sentences at this level is said to give their phrase structure. Moreover, phrases of a given kind can occur within others of the same kind: It is this recursive feature of language that enables sentences of arbitrary complexity to be constructed. The realization that phrase structure is recursive is very old. Assuming the categories of a complete noun phrase (NP) and sentence (S), Antoine Arnauld (1662) gives the examples (rendered here in English):

(1) ($_S$The divine law commands that [$_S$kings are to be honored])

(2) ($_S$[$_{NP}$Men [$_S$who are pious]] are charitable)

remarking that in (1) the embedded element "kings are to be honored" is a sentence occurring within a sentence, and that in (2) the relative clause has all the structure of a sentence, except that the relative pronoun "who" has replaced the subject.

In linguistic theory the recursive structure of syntax is expressed by principles of combination modeled after the clauses of an inductive definition. However, far more complex devices seem to be required for a compact description that helps to reveal the basis of the native speaker's ability. Chomsky's introduction of grammatical transformations opened the way to a variety of formalisms and developments (see Atkinson, Kilby, and Roca 1988 for a useful overview). Chomsky also initiated the conception of linguistic theory as a study of the acquisition of a system of linguistic knowledge, or competence. Any human language is acquirable under ordinary experiential conditions by any normal child. The space between empirical evidence and the resulting linguistic competence is sufficiently great that a kind of readiness for language, universal grammar in Chomsky's terminology, is presupposed. Contemporary theory seeks to probe the basis for this readiness in terms of innate rules and principles of grammar. For a more recent statement, see Chomsky and H. Lasnik (in Jacobs et al. 1993).

Within philosophy too the theory of syntax came to play an important role in the systematization of mathematics, and assumed central importance in Rudolf Carnap (1934). Carnap distinguished between grammatical syntax, of the sort that a linguist might give in a description of a language, and logical syntax, whose aim was not only to specify the class of sentences (or well-formed formulas of a calculus) but also to use formal methods in constructing a theory of logical consequence and logical truth. Carnap employed the distinction between grammatical form and logical form, which plays a crucial part in Ludwig Wittgenstein's views both in the *Tractatus* and in the *Philosophical Investigations,* and has become part of the lore of analytic philosophy. The scope of logical syntax in Carnap's terms took on much of the role of semantics in later philosophical discussion. Even with the later distinction between syntax and model-theoretic semantics, syntactic properties of formalized languages are still crucial for properties of systems of logic (soundness and completeness), and proof theory is established as a part of the syntax of mathematics.

In linguistic theory syntax and semantics have become increasingly intertwined disciplines, as it was realized that there are explanatory issues in relating linguistic forms to the specific meanings, or range of meanings, associated with them. S. Lappin (1995) contains a number of useful expositions on this theme; see also R. Larson and G. Segal (1995). The current research climate is in practice very different from conceptions associated with "ordinary language" philosophy: The contemporary view is not that ordinary speech lacks an exact logic, but rather that a diligent, collaborative effort is required to find out what the logic is. The concentration on logic implies that syntactic investigations have a metaphysical dimension. The patterns of inference of ordinary language call for formalization as part of a general account of the structure of individual human languages, or human language in general, and this formalization may in turn lead to proposals for reification, as in Donald Davidson's (1967) hypothesis that references to events are pervasive in ordinary action sentences.

On the side of linguistics proper, the problems of morphology have been treated in a progressively more syntactic manner as, for instance, our example *solubility* can be seen as built up by rules of a sort familiar from syntax. The result is the area now called morphosyntax, where the question whether morphology is a distinct level of linguistic organization is under active debate; see R. Hendrick (1995) for more recent discussion.

See also Arnauld, Antoine; Carnap, Rudolf; Chomsky, Noam; Davidson, Donald; Language; Logic, History of; Logical Form; Philosophy; Philosophy of Language; Proof Theory; Semantics; Wittgenstein, Ludwig Josef Johann.

Bibliography

Arnauld, A. *La logique, ou l'art de penser* (1662). Translated by J. Dickoff and P. James as *The Art of Thinking.* Indianapolis: Bobbs-Merrill, 1964.

Atkinson, M., D. Kilby, and I. Roca. *Foundations of General Linguistics.* 2nd ed. London: Unwin Hyman, 1988.

Carnap, R. *The Logical Syntax of Language.* London: K. Paul, Trench, Trubner, 1937.

Chomsky, N. *Syntactic Structures.* The Hague: Mouton, 1957.

Chomsky, N., and H. Lasnik. "The Theory of Principles and Parameters." In *Syntax: An International Handbook of Contemporary Research,* edited by J. Jacobs, A. von Stechow, W. Sternfeld, and T. Vennemann. Berlin and New York, 1993.

Davidson, D. "The Logical Form of Action Sentences." In *The Logic of Decision and Action,* edited by N. Rescher. Pittsburgh: University of Pittsburgh Press, 1967.

Harris, Z. "Co-Occurrence and Transformations in Linguistic Structure." *Language 33* 3 (1957).

Hendrick, R. "Morphosyntax." In *Government and Binding Theory and the Minimalist Program,* edited by G. Webelhuth. Oxford: Blackwell, 1995.

Lappin, S., ed. *The Handbook of Contemporary Semantic Theory.* Oxford: Blackwell, 1995.

Larson, R., and G. Segal. *Knowledge of Meaning.* Cambridge, MA: MIT Press, 1995.

James Higginbotham (1996)

SYNTAX AND SYNTACTICS

See *Semantics*

SYNTHETIC STATEMENTS

See *Analytic and Synthetic Statements*

T

TAGORE, RABINDRANATH
(1861–1941)

Rabindranath Tagore was an Indian writer and philosopher. Romain Rolland, referring to the Orient and the Occident, said that Tagore contributed more than anyone else toward "the union of these two hemispheres of spirit." Sarvepalli Radhakrishnan called Tagore "the greatest figure of the Indian renaissance."

Tagore was born in Calcutta, studied in London, returned to India, and was married in 1883. He founded Visvabharati, a university at Santiniketan (near Bolpur), became India's most popular poet, won the Nobel Prize for literature in 1913, and was knighted in 1915. He visited and lectured in Canada, the United States, South America, England and several countries of Europe, the Soviet Union, Turkey, Iran, Ceylon, China, and Japan. He was in personal contact with Henri Bergson, Benedetto Croce, Albert Einstein, Bertrand Russell, and other leading intellectual figures of his period.

Tagore wrote about fifteen books of philosophical lectures and essays, about one hundred books of verse (mostly in Bengali, and partly translated by himself from his own Bengali version into English), about fifty plays (in some of which he acted the main role), and about forty works of fiction. His main writings of philosophical interest are *Sadhana: The Realisation of Life* (1913), *Personality* (1917), *Creative Unity* (1922), *The Religion of Man* (1931), all published in London and New York, and *Man* (1937), published in Madras. His best-known poems appear in *Gitanjali* (Song offerings), translated by the author from the original Bengali, with an introduction by W. B. Yeats (1913); *The Crescent Moon,* likewise translated by the author from the original Bengali (1913); and *Fruit-Gathering* (1916), all published in London and New York. He produced some drawings and paintings, beginning about his seventieth year, and planned and produced ballets.

Tagore's basic philosophical position is one that recognizes the useful insights of the main opposing views on a given question. For example, concerning the transcendence or immanence of God, Tagore accepted, on the one hand, the value of the doctrine of Brahman as "the absolute Truth, the impersonal It, in which there can be no distinction of this and that, the good and the evil, the beautiful and its opposite, having no other quality except its ineffable blissfulness in the eternal solitude of its consciousness"; but he also felt, on the other hand, that "whatever name may have been given to the divine Real-

ity it has found its highest place in the history of our religion owing to its human character, giving meaning to the idea of sin and sanctity, and offering an eternal background to all the ideals of perfection which have their harmony with man's own nature" (*The Religion of Man*).

Similarly, he combined the best insights of humanists, who exalt man, and of otherworldly seekers of the World Force, who belittle man; of naturalists, who deny spirit, and of extreme partisans of spirit, who cut man off from nature; of individualists and universalists; of determinists and defenders of free will; of hedonists and ascetics; and of romanticists and realists.

In his social philosophy, as well as in his metaphysics, Tagore attempted to synthesize polar opposites. Neither wholly conservative nor wholly liberal, he favored gradual reform. This evolutionary note is reflected in his views on the economic order, public health, education, the social structure, national politics, and international affairs.

Tagore's emphasis on the mediating unity that embraces variety appears, for example, in *Sadhana,* where he wrote: "Facts are many, but the truth is one.… Man must clearly realise some central truth which will give him an outlook over the widest possible field. And that is the object which the Upanishad has in view when it says, Know thine own Soul. Or, in other words, realise the one great principle of unity that there is in every man."

In May 1930 Tagore delivered the Hibbert Lectures at Oxford. In the following year, the lectures were published in expanded form as a book, *The Religion of Man.* Tagore's mediationism appears in the book in such passages as the following: "The final freedom which India aspires after … is beyond all limits of personality, divested of all moral or aesthetic distinctions; it is the pure consciousness of Being, the ultimate reality." The yogi has claimed that through intensive concentration and quietude we do reach "that infinity where knowledge ceases to be knowledge, subject and object become one—a state of existence that cannot be defined.… India attunes man to the grand harmony of the universal, leaving no room for untrained desires of a rampant individualism to pursue their destructive career unchecked, but leading them on to their ultimate modulation in the Supreme."

See also Bergson, Henri; Brahman; Croce, Benedetto; Einstein, Albert; Humanism; Indian Philosophy; Russell, Bertrand Arthur William.

Bibliography

The philosophical works by Tagore that are mentioned above, as well as others, have been published in various editions, in English and in translation into European languages.

Works on Tagore's philosophy include the following (arranged in chronological order): Sarvepalli Radhakrishnan, *The Philosophy of Rabindranath Tagore* (London: Macmillan, 1918; reissued Baroda: Good Companions, 1961); Sachin Sen, *The Political Thought of Tagore* (Calcutta: General Printers and Publishers, 1947); B. G. Ray, *The Philosophy of Rabindranath Tagore* (Bombay: Hind Kitabs, 1949); V. S. Naravane, *Rabindranath Tagore, A Philosophical Study* (Allahabad: Central Book Depot, n.d.); and Sasadhar Sinha, *Social Thinking of Rabindranath Tagore* (New York: Asia Publishing House, 1962).

Of the many biographies of Tagore, mention may be made of Krishna Kripalani, *Rabindranath Tagore: A Biography* (New York: Oxford University Press, 1962).

William Gerber (1967)

TAINE, HIPPOLYTE-ADOLPHE

(1828–1893)

Hippolyte-Adolphe Taine was a philosopher, psychologist, historian, and critic. Taine and Ernest Renan were the leading French positivistic thinkers of the second half of the nineteenth century. As a result of Taine's great independence of mind, his life was not always comfortable. Discriminatory treatment from the authorities of the Second Empire led to his withdrawal from teaching from 1852 to 1863, when he was appointed an examiner at Saint-Cyr. The next year he became a lecturer at the École des Beaux Arts; from his lectures there came his famous *Philosophie de l'art,* At the intervention of the Catholic clergy, a French Academy award for his *Histoire de la littérature anglaise* was denied him, and he was elected to the academy only in 1878, after the fall of the Second Empire. By that time he had antagonized both liberals and Bonapartists by his ruthless destruction of the revolutionary and Napoleonic legends. Nevertheless, his influence was great and diversified. His positivistic and physiological approach to psychology was adopted by Théodule Ribot, Pierre Janet, and others, and his opposition to centralization and to revolutionary experiments attracted Catholic traditionalists such as Paul Bourget and Maurice Barrès, who, however, ignored his severe condemnation of the old regime and his outspoken sympathies for Protestant and parliamentary England.

Although Taine's philosophical views were formed early in life under the joint influence of Benedict de Spin-

oza, G. W. F. Hegel, and classical science, they were first systematically expounded in his *De l'intelligence.* The theory of mind presented in this book is based on Taine's general monism and determinism. Thus in the preface to the fourth edition (Paris, 1883), he stated his opposition to faculty psychology on the grounds that words such as *capacity, self, reason,* and *memory* suggest by their simplicity the existence of indivisible mental entities and thus prevent us from grasping the enormous complexities of the underlying psychological mechanisms. The self is nothing but a series of mental events. In his attack on the substantialization of the self and the reification of abstractions, Taine drew on psychopathology and neural physiology. Psychopathology shows how mental disease can dissociate the components of a complex phenomenon that appears subjectively as simple; neural physiology reveals the enormous complexity of the neural mechanism that underlies mental phenomena. Taine held a double-aspect theory of the relation between introspective data and public physical events; the mental and the physical are two sides of the same process, "two translations of the same text" (*De l'intelligence,* Book 4, Ch. 2). Taine's use of physiological analysis, his strictures on introspection, and his mechanistic determinism place him among the naturalists.

Like most of his contemporaries, Taine regarded classical science as complete, and its picture of nature as definitive. Like Herbert Spencer, Wilhelm Ostwald, and others, he regarded the law of conservation of energy as ultimate, as "the immutable ground of being," and the equivalence of cause and effect as a consequence of this law.

Taine applied his rigorous determinism to all phenomena—physical, mental, and social. There is little in his writings dealing directly with physical phenomena, but there is no question that the determinism of classical physics was for him an ideal model to which other sciences should conform. Thus in the introduction to his *Histoire de la littérature anglaise,* he proposed that every social phenomenon should be explained as the result of race, environment, and time—that is, of the particular psychosocial state of a society. Taine had already applied this method in previous essays, and he applied it in his *Philosophie de l'art* and later in his major historical work, *Les origines de la France contemporaine,* inspired by his reflections on the French defeat in 1870. The thesis of this monumental and controversial work is that there was one persistent theme—excessive centralization—underlying all the violent upheavals of modern France. Introduced by the Bourbons, it was strengthened by the French Revolution, which destroyed the natural provinces and replaced them by departments which were mere administrative appendixes of the central government; in the hands of Napoleon Bonaparte the centralized administrative structure was an efficient tool of internal control and external conquest, but it became an unwieldy bureaucratic machine as soon as it was deprived of Napoleon's ruthless energy.

Taine's detailed study of social conditions under the old regime, of revolutionary excesses, and of mob psychology after 1789 strengthened the inclination to pessimism present in his previous writings. This inclination found its most eloquent expression in the following passage: "Man is a nervous machine, governed by a mood, disposed to hallucinations, transported by unbridled passions, essentially unreasonable" (*History of English Literature,* Vol. II, p. 173). In *De l'intelligence* Taine had said that every image tends to acquire a hallucinatory intensity unless checked by the inhibiting influence of other images. Thus mental equilibrium and social stability are mere "happy accidents." Civilization is a mere surface beneath which lurk irrational drives always ready to break through.

See also Determinism in History; Hegel, Georg Wilhelm Friedrich; Ostwald, Wilhelm; Renan, Joseph Ernest; Ribot, Théodule Armand; Sociology of Knowledge; Spinoza, Benedict (Baruch) de.

Bibliography

WORKS BY TAINE

Essais de critique et d'histoire. Paris: Hachette, 1858.

Histoire de la littérature anglaise, 4 vols. Paris: Hachette, 1863–1864. Translated by H. van Laun as *History of English Literature,* 2 vols. Edinburgh, 1873.

Nouveaux essais de critique et d'histoire. Paris, 1865.

Philosophie de l'art. Paris: Baillière, 1865; 2nd ed., enlarged, Paris, 1880. First edition translated by J. Durand as *The Philosophy of Art.* New York, 1865.

De l'intelligence, 2 vols. Paris: Hachette, 1870. Translated by T. D. Hayes as *Intelligence.* London, 1871.

Les origines de la France contemporaine, 5 vols. Paris, 1876–1893.

Derniers essais de critique et d'histoire. Paris, 1894.

WORKS ON TAINE

Barzelotti, Giacomo. *Ippolito Taine.* Rome, 1896.

Bourget, Paul. *Essais de psychologie contemporaine.* Vol. I. Paris, 1883.

Chevrillon, André. *Taine: Formation de sa pensée.* Paris: Plon, 1932.

Faguet, Émile. *Politiques et moralistes du XIXe siècle.* 3rd series. Paris, 1900.

Giraud, Victor. *Essai sur Taine: Son oeuvre et son influence.* Paris: Hachette, 1901.
Kahn, Sholem J. *Science and Aesthetic Judgment: A Study in Taine's Critical Method.* New York: Columbia University Press, 1953.
Lacombe, Paul. *La psychologie des individus et des sociétés chez Taine.* Paris: Alcan, 1906.
Lacombe, Paul. *Taine, historien et sociologue.* Paris: V. Giard and E. Brière, 1909.

Milič Čapek (1967)

TAOISM

See *Chinese Philosophy: Daoism; Laozi; Mysticism, History of*

TARSKI, ALFRED

(1902–1983)

Alfred Tarski, the Polish-American mathematician and logician, was born in Warsaw, received his doctorate in mathematics from the University of Warsaw in 1924, and two years later was named docent. In 1939 he emigrated to the United States. Appointed lecturer in mathematics at the University of California (Berkeley) in 1942, he remained at that institution for the rest of his life, serving as professor of mathematics from 1946 and becoming professor emeritus in 1968.

MATHEMATICS

Tarski worked in both pure mathematics, especially set theory and algebra, and mathematical logic, especially metamathematics. This entry will not discuss his mathematical contributions, although some of them (in particular his famous theorem, established jointly with Stefan Banach, on the decomposition of the sphere, as well as his theory of inaccessible cardinals) have a definite bearing on the epistemology of mathematics. (See S. Banach and A. Tarski, "Sur la décomposition des ensembles des points en parties respectivement congruentes," *Fundamenta Mathematicae* 6 [1924]: 244–277.)

It should be noted that in these papers Tarski has not criticized the assumptions of set theory. Like most mathematicians he has simply accepted them as true. This attitude and a systematic use of set-theoretic concepts have profoundly influenced his work in logic and metamathematics. Unlike the followers of David Hilbert and of L. E. J. Brouwer, Tarski has not refrained from the use of infinitistic set-theoretical concepts. He finds a definition or a theorem to be acceptable if it is expressed or proved on the basis of set theory. This attitude, of course, is completely different from that of Hilbert's formalism or Brouwer's intuitionism.

As a consequence of this methodological attitude, Tarski has gained much freedom in introducing new notions and thus has put himself in a much more advantageous position than the adherents of Hilbert or Brouwer. Consider the following very simple but typical example. In *Logic, Semantics, Metamathematics* (p. 38) Tarski defines the set of consequences of a given set of axioms as the smallest set containing the axioms and closed with respect to the rules of proof, and on this definition he bases the whole theory of the consequence relation in the propositional calculus. A follower of Hilbert or Brouwer would never accept such a definition because he would regard the clarification of the notion of set (involved in this definition) as the ultimate aim of his activity.

The free use of set theory has enabled Tarski to extend the field of application of metamathematics (see, for instance, his investigations of "infinitary languages," discussed below) and has formed a natural basis for the development of his semantic method. This method can indeed be formulated only in a language that has considerable deductive strength and is provided with means to express definitions of a very complicated structure. The general theory of sets satisfies both these requirements.

Obviously Tarski's methodological attitude is rejected by the adherents of finitism and by all logicians who seek in metamathematics a justification or explanation of set theory.

METAMATHEMATICS

Metamathematics is a branch of mathematical logic that studies formal theories and solves problems pertaining to such theories. Tarski contributed so much to this field that he deserves to be regarded, with Hilbert, as its cofounder.

AXIOMATIC THEORY OF FORMAL SYSTEMS. In his early papers Tarski presented an axiomatic theory of arbitrary formal systems. A "theory" for him is a set (whose elements are called formulas) and a function (called the consequence function) that correlates a set of formulas with each such set; this new set is called the set of consequences of the first set. The consequence function is not wholly arbitrary; it must satisfy certain axioms that will not be reproduced here. Several metamathematical notions, such as consistency, completeness, and inde-

pendence, can be defined for theories in this abstract sense. All formal theories that were known in 1930 can be subsumed under this scheme. While this is no longer true today (see below), a relatively small rectification of Tarski's axioms would suffice to restore the universality of his scheme.

SYSTEMS BASED ON PROPOSITIONAL LOGIC. Besides discussing the most general scheme of formal theories, Tarski axiomatically described theories based on the classical propositional logic. Here the assumptions must, of course, be specialized. It is assumed, for example, that certain operations are defined on the set of formulas (the joining of two formulas by means of a connective). An example of an important property of consequence that Tarski took as an axiom is the deduction theorem. Its importance is that it provides the possibility of defining the consequence function in terms of one fixed set S_0 of sentences, specifically the set of consequences of the empty set. In concrete cases, S_0 consists of logical tautologies expressible in the given theory. In what follows, we shall speak of theories as being based on a logic L if S_0 is the set of tautologies of the logic L.

DESCRIPTION OF SYSTEMS. Tarski calls a set X of formulas a system if it is deductively closed, that is, if it is equal to its set of consequences. In "Grundzüge des Systemenkalküls" he formulated a general program aimed at describing all systems of a given theory. Tarski showed in this paper that in order to achieve this aim it is sufficient to describe all complete systems, and he illustrated his program in several simple but interesting cases of decidable theories. Many ideas developed in this paper were later incorporated by Tarski in the general theory of models.

SEMANTICS

In the early 1930s Tarski formulated the semantic method, which is his most important achievement in logic. The essence of the method consists in discussion of the relations between expressions and the objects they denote.

Tarski himself said that his semantics was a modest discipline. Yet the philosophical claims of semantics were ambitious from the start. Tarski's aim was "to construct—with reference to a given language—*a materially adequate and formally correct definition of the term 'true sentence,'*" a problem "which belongs to the classical questions of philosophy."

Almost from the beginning the methods of semantics exerted a profound influence on philosophers engaged in the construction and study of exact scientific languages. Semantics opened new possibilities in these studies, which formerly were limited to purely syntactic problems and thus were unable to express relations between languages and extralinguistic objects. Semantics offered a natural tool for the discussion of such relations. The price one had to pay was the use of a much stronger metalanguage than the one sufficient for syntax. At any rate, semantic methods became an accepted tool in the study of scientific languages: "Contemporary studies in the methodology of science are primarily concerned with the syntax and semantics of the language of science" (R. M. Martin, *Truth and Denotation,* Chicago, 1958, p. 16).

Tarski published little concerning the applicability of semantics to the study of empirical languages (see, however, his remarks in "The Semantic Conception of Truth and the Foundations of Semantics"). Rather, he limited himself to applications of his method to logic and mathematics. His most outstanding contributions in these areas will be described briefly.

INTERPRETATIONS OF PROPOSITIONAL CALCULI. The propositional calculus provides us with simple examples of semantic notions. Thus, the two-element Boolean algebra is an interpretation of the calculus; the propositional connectives are interpreted as functions whose arguments and values range over the algebra. We may accordingly conceive of the propositional calculus as a language that describes the two-element algebra. Instead of the two-element algebra we may take any other matrix for the propositional calculus. Thus, a formal calculus may have (and in general does have) many interpretations. Tarski early became acquainted with these notions through his collaboration with Jan Łukasiewicz, who in the 1920s initiated the metatheoretical investigation of propositional calculi. In a joint publication Tarski and Łukasiewicz gave a general set-theoretical definition of a matrix and showed its usefulness in various special problems.

MODELS. Models play the same role for theories based on (extensions of) the first-order functional calculus as that played by matrices for propositional calculi. If a theory T has as its primitive constants k predicates with $r_1, \cdot\cdot\cdot, r_k$ arguments, then a model for T is defined as an ordered $k+1$-tuple $\langle A, R_1, \cdot\cdot\cdot, R_k \rangle$, where R_i is a relation with r_i arguments ranging over A ($i = 1, \cdot\cdot\cdot, k$). A model determines a partition of sentences into two sets, one consisting of sentences that are true in the model and the

other of sentences that are false in the model. A formula that contains free variables is by itself neither true nor false in the model, but if arbitrary elements of A are correlated with the free variables of the formula, it becomes either true or false. In the first case we say that the elements of A correlated with the free variables satisfy the formula in the model. We have here an analogy with the situation in the propositional calculus: if a matrix is given and if its elements are correlated in an arbitrary way with the free variables of a formula, then the formula has a value that is an element of the matrix. This analogy between models and matrices was stressed in "The Concept of Truth in Formalized Languages," in *Logic, Semantics, Metamathematics,* pp. 152–278. (This is an English translation of an earlier paper.)

The notion of a model and some related semantic notions were known to mathematicians and logicians long before the work of Tarski. No one, however, was concerned to strive for such a degree of precision as Tarski maintained. The fruits of Tarski's approach are first, a precise set-theoretical description of the semantic notions, together with a meticulous discussion of the language in which these definitions are expressible; second, the discovery of general properties of these notions which sometimes are very startling; and third, the discovery of a broad field of applications.

The semantic notions, which before Tarski were used in solving relatively special problems concerning consistency and independence, now turned out to be powerful tools in dealing with many metamathematical investigations. For a philosopher the most important application of the semantic method is Tarski's theory of truth.

LOGICAL CONSEQUENCE. Logical consequence is defined as follows: a sentence F is a logical consequence of a set X of sentences if F is true in every model in which all sentences of X are true. For theories based on first-order logic this notion is coextensive with the syntactic notion of derivability (Gödel's completeness theorem). For theories based on the higher-order logics or on the various extensions of first-order logic, these notions are essentially different. Analyzing the intuitions underlying the notion of consequence, one arrives with Tarski at the conclusion that it is the semantic and not the syntactic notion that adequately describes the notion that is intuitively given. At the same time, many logics in which the consequence functions are defined semantically turn out to be free from defects resulting from the incompleteness phenomenon discovered by Kurt Gödel. This shows the essential gains brought by the acceptance of the semantically defined notion of logical consequence. What is lost is the finitary ("combinatorial") description of the consequence function.

DEFINABILITY. Like the notion of consequence, definability can be treated syntactically and semantically. Although investigations in both these directions were pursued in special cases before Tarski, it is only following Tarski's work that we can speak of a systematic theory of definability.

Syntactic theory of definability. Let T be a formal theory among whose constants there is a one-place predicate C. We say that C depends on other constants of T if there is a formula F free of C with exactly one free variable x such that the equivalence $C(x) \equiv F$ is provable in T. In special cases this notion was used long before Tarski; but Tarski was the first to formulate this notion precisely and in the general case, to discuss its properties, and to discover a far-reaching parallelism between the notions of consequence and definability. One of the most interesting results of his theory is a general formulation of a method (due in principle to A. Padoa) allowing one to establish the independence of a constant. Tarski also showed the universality of this method in cases in which the theory under consideration is based on second-order logic or its extensions; the case of theories based on first-order logic was decided much later by E. W. Beth.

Semantic notion of definability. Let M be a model as defined above. A subset S of A is called definable in M if there is a formula F with exactly one free variable such that an element a of A satisfies F in M if and only if a is an element of S. The formula F is called a definition of S in M.

The determination of the class of definable sets is an interesting problem that occupies a central place in investigations concerning the so-called hierarchies of sets. Without going into details, the aim of these investigations is to discuss sets obtainable from simple sets (which constitute the lowest level of the hierarchy) by means of fixed operations that lead to higher and higher levels. Hierarchies of this kind are discussed in mathematics (the Borel and the projective hierarchies) and in metamathematics (the arithmetical, the hyperarithmetical, and the analytic hierarchies). Tarski and Kazimierz Kuratowski in a joint paper described a method that in many cases allows one to infer directly, from the form of definition of a set, to which level of a given hierarchy this set belongs. Their method introduced essential simplifications into the theory of hierarchies.

The importance of these investigations for metamathematics will be clear if we reflect that, for example, Gödel's incompleteness theorem is an obvious corollary of the fact that the set of (numbers of) sentences derivable from the axioms of arithmetic does not belong to the lowest level of the arithmetical hierarchy. Tarski's work on definability is thus closely connected with problems of incompleteness. The most important result in this field is his theorem on truth, which says that under very general assumptions the set of (numbers of) sentences that are true in a model *M* is not definable in *M.* Gödel's incompleteness theorem for arithmetic and many related results are immediate corollaries of this theorem ("On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth," 1.939). Tarski's semantic theorem, however, requires for its formulation as well as for its proof a much stronger logical basis than the syntactic theorem of Gödel.

GENERAL THEORY OF MODELS. Notions closely related to models (as defined above) appeared in abstract mathematics independently of the logical investigations. Mathematicians were led to notions of this degree of generality by the development of abstract algebra. Tarski developed these algebraic investigations and tied them to metamathematics.

It is easy to explain the close connections between the general theory of models and the theory of systems. If we consider a theory whose consequence function is defined semantically, then every system is determined by the class of those models in which all sentences of the system are true. Conversely, every model determines a (complete) system consisting of sentences that are true in the model. However, different models may yield one and the same system.

Tarski and his students exploited these relationships especially for the case in which the theory under consideration is based on first-order logic. In this case it is irrelevant whether we accept the semantic or the syntactic notion of consequence, and we thus have the advantage of being able to use on the one hand the connection between systems and models and on the other the various properties of the consequence function that result from its syntactic definition. One of these properties is the so-called compactness of the consequence function, which states that if a set *X* of sentences is contradictory, then the same is true of a finite subset of *X.*

In his publications on the theory of models, which date as far back as 1949, Tarski sought to develop the theory in purely mathematical terms and avoided notions current in logic but less so in mathematics. Consequently his papers on the theory of models are more accessible to mathematicians than to logicians. The details of his highly technical works on the theory of models cannot be related here, and we must content ourselves with the brief indications given above.

GENERALIZATIONS OF FIRST-ORDER LOGIC. As was stated earlier, the general setting of model theory is meaningful for theories that are not necessarily based on first-order logic. Tarski suggested two important generalizations of first-order logic and showed that the model-theoretic approach to these logics leads to important discoveries.

The first of these logics is one with infinitely long formulas ("A Sentential Calculus with Infinitely Long Expressions"). Such formulas are, of course, abstract entities definable only in strong systems of set theory; nevertheless, Tarski showed that most of the questions formerly raised exclusively for theories based on ordinary logic are also meaningful for this abstractly described logic. The mathematically important work "Some Problems and Results Relating to the Foundations of Set Theory" resulted from a negative solution of the analogue of the compactness problem ("Some Model-Theoretical Results concerning Weak Second Order Logic," *Notices of the American Mathematical Society* 5, Abstract 550–6) for logics with infinitely long formulas.

Another important logic introduced by Tarski is weak second-order logic, that is, second-order logic in which the set variables are restricted to finite sets. For this logic as well, the semantic notion of consequence is definable only in a fairly strong system of set theory. Thus weak second-order logic, like the preceding one, is only an abstract construction. Tarski established various metamathematical properties of this logic (for instance, the analogue of the Skolem-Löwenheim theorem) and showed that they imply important mathematical consequences in algebra.

FURTHER CONTRIBUTIONS

DECISION PROBLEM AND UNDECIDABLE THEORIES. The decision problem for a theory *T* is the question whether there exists an algorithm allowing one to decide whether a sentence of *T* is or is not provable in *T.* Tarski discussed this problem for a large number of theories using the so-called method of the elimination of quantifiers, which originated with Thoralf Skolem ("The Concept of Truth in Formalized Languages," in *Logic, Language, Metamathematics,* p. 204). The most important

result in this direction was a positive solution of the problem in the case in which T is the first-order theory of the field of real numbers (*A Decision Method for Elementary Algebra and Geometry*). This result found numerous applications in algebra and geometry.

A theory for which the decision problem does not admit a positive solution is called undecidable. It was related above how Tarski deduced the incompleteness (and hence the undecidability) of arithmetic from his general theorem. His further efforts were directed toward establishing the undecidability of various very weak but mathematically interesting theories. To this end he introduced the important notion of essential undecidability. A theory is said to be essentially undecidable if all consistent extensions of it are undecidable. Tarski showed in *Undecidable Theories* (1953) that a theory that has a joint consistent extension with an essentially undecidable theory based on a finite number of axioms is itself undecidable, although in general not essentially undecidable. This theorem provided a basis for numerous undecidability results obtained partly by Tarski and partly by his collaborators.

INTUITIONIST AND MODAL LOGICS. Of the numerous papers that Tarski devoted to the propositional calculus, only those on the intuitionistic and modal propositional calculi can be mentioned here. In "Sentential Calculus and Topology" (*Logic, Semantics, Metamathematics,* pp. 421–454) he established a startling connection between intuitionistic logic and topology: he constructed matrices for the intuitionistic propositional calculus, using as elements closed subsets of a topological space. In his further work on this calculus, done jointly with J. C. C. McKinsey, he no longer used topological notions but worked instead with certain algebraic structures. The class of all subsets of a topological space and the class of all closed subsets of such a space are examples of such structures, which Tarski and McKinsey called closure algebras and Brouwerian algebras, respectively. Using them, they established several properties of the intuitionistic and modal propositional logics.

CYLINDRIC ALGEBRAS. The above papers give a good illustration of Tarski's growing tendency to deal with metamathematical problems by means of algebraic tools. Another example is his work on cylindric algebras. These algebraic structures are related to the predicate calculus with identity in the way Boolean algebras are related to the usual propositional calculus. Logics with infinitely long expressions can also be investigated by means of suitable cylindric algebras.

CALCULUS OF BINARY RELATIONS. The calculus of binary relations was created by Ernst Schröder but soon fell into oblivion. Tarski gave axioms for this calculus, investigated its relations to the predicate calculus, and initiated extensive work on the models of his axioms. Of the several applications of the calculus found by Tarski, the axiomatization of set theory without variables, the existence of undecidable subsystems of the two-valued propositional calculus, and a general method of reduction of the number of primitive terms of a theory should be mentioned.

PHILOSOPHY

In the rich bibliography of Tarski's publications there are almost no philosophical papers. The exceptions are "The Establishment of Scientific Semantics" and "The Semantic Conception of Truth and the Foundations of Semantics," which deal with the philosophical significance of semantics. A partial exception is Tarski's paper on the notion of truth (in *Logic, Semantics, Metamathematics,* pp. 153–278), although the bulk of it is devoted to a systematic exposition of semantics.

Tarski, in oral discussions, often indicated his sympathies with nominalism. While he never accepted the "reism" of Tadeusz Kotarbiński, he was certainly attracted to it in the early phase of his work. However, the set-theoretical methods that form the basis of his logical and mathematical studies compelled him constantly to use the abstract and general notions that a nominalist seeks to avoid. In the absence of more extensive publications by Tarski on philosophical subjects, this conflict appears to have remained unresolved.

See also Boole, George; Brouwer, Luitzen Egbertus Jan; Correspondence Theory of Truth; First-Order Logic; Gödel's Theorem; Hilbert, David; Kotarbiński, Tadeusz; Logic, History of; Łukasiewicz, Jan; Mathematics, Foundations of; Model Theory; Second-Order Logic; Semantics; Set Theory.

Bibliography

Tarski's scientific writings consist of more than one hundred articles and books, plus many abstracts and reviews. Among these the most important for logic and philosophy are the following:

"Sur les truth-functions au sens de MM. Russell et Whitehead." *Fundamenta Mathematicae* 5 (1924): 59–74.

"Grundzüge des Systemenkalküls." *Fundamenta Mathematicae* 25 (1935): 503.

"Der Wahrheitsbegriff in den formalisierten Sprachen." *Studio Philosophica* 1 (1935–1936): 261–405.
"Über unerreichbare Kardinalzahlen." *Fundamenta Mathematicae* 30 (1938): 68–89.
"On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth." *Journal of Symbolic Logic* 4 (1939): 105–112.
"On the Calculus of Relations." *Journal of Symbolic Logic* 6 (1941): 73–89.
"The Semantic Conception of Truth and the Foundations of Semantics." *Journal of Philosophy and Phenomenological Research* 4 (1944): 341–375. Reprinted in *Readings in Philosophical Analysis,* edited by H. Feigl and W. Sellars, 52–84. New York: Appleton-Century-Crofts, 1949.
"On Closed Elements in Closure Algebras." *Annals of Mathematics* 45 (1944): 141–191, and 47 (1946): 122–162. Written with J. C. C. McKinsey, with remarks by Tarski, 163–165.
"Some Theorems about the Sentential Calculi of Lewis and Heyting." *Journal of Symbolic Logic* 13 (1948): 1–15. Written with J. C. C. McKinsey.
"Some Notions and Methods on the Borderline of Algebra and Metamathematics." In *Proceedings of the International Congress of Mathematicians,* 705–720. Cambridge, MA, 1950.
A Decision Method for Elementary Algebra and Geometry. Santa Monica, CA, 1948; 2nd ed., Berkeley: University of California Press, 1951.
Undecidable Theories. Amsterdam: North-Holland, 1953. Written with A. Mostowski and R. M. Robinson.
Logic, Semantics, Metamathematics. Oxford: Clarendon Press, 1956. Tarski's papers on logic from 1923 to 1938, collected and translated by J. H. Woodger.
"A Sentential Calculus with Infinitely Long Expressions." *Colloquium Mathematicum* 6 (1958): 165–170. Written with Dana Scott. Remarks by Tarski, 171–176.
"Cylindric Algebras." In *Proceedings of Symposia in Pure Mathematics: II Lattice Theory,* 83–113. Providence, RI: American Mathematical Society, 1961. Written with Leon Henkin.
"Some Problems and Results Relating to the Foundations of Set Theory." In *Proceedings of the 1960 Congress on Logic, Methodology, and Philosophy of Science,*125–135. Palo Alto, CA, 1962.
"From Accessible to Inaccessible Cardinals." *Fundamenta Mathematicae* 53 (1964): 225–308. Written with H. J. Keisler.

Andrzej Mostowski (1967)

TARSKI, ALFRED [ADDENDUM]

Alfred Tarski was born in 1901 (not 1902, as stated in the original entry). The name on his birth certificate was Alfred Teitelbaum (variant: Tajtelbaum); he changed it to Alfred Tarski in 1924. That same year his dissertation, written under the direction of Stanisław Leśniewski, was published in two parts, the first under his birth name and the second under Alfred Tajtelbaum-Tarski; thereafter, all his articles and books were published under the name Alfred Tarski.

Tarski's immigration to the United States was somewhat accidental: He was attending a meeting of the Unity of Science at Harvard University in September 1939 when the Nazis invaded Poland and World War II began. Tarski was stranded and separated from his wife and two children, who were left behind in Warsaw (they were reunited after the war, but most of his family perished in the Holocaust). In 1942, after three years of casting about for a position, he received a one-year appointment as a lecturer at the University of California, Berkeley (UCB), from which he quickly rose to professor of mathematics in 1946. Working intensively with increasing success, he built a substantial graduate program in logic and, within a decade, Berkeley became a mecca for logicians worldwide. In 1957 Tarski was instrumental in creating the interdepartmental Program in Logic and Methodology of Science at UCB that mainly bridged the departments of mathematics and philosophy.

Tarski retired in 1968 but was recalled to teach for the next five years. He continued to do research and advise students until a year before his death in 1983. In the last decade of his life he received a number of honors, including honorary doctorates from the Universidad Católica de Chile, the Université d'Aix-Marseille II, and the University of Calgary; in addition, in 1981 he was awarded the Berkeley Citation, the highest honor that UCB can bestow. For a full biography see Anita Burdman Feferman and Solomon Feferman (2004).

From the 1960s to the end of his life, with the collaboration of colleagues and students, Tarski concentrated on the topics of axiomatic geometry and algebraic logic, while continuing to contribute to the areas of model theory, set theory, and universal algebra. His work on first-order systems of Euclidean geometry and the work that it led to in non-Euclidean geometry is described in a joint article with Steven R. Givant, "Tarski's System of Geometry" (1999). The research on relation algebra was capped by the joint monograph with Givant, *A Formalization of Set Theory without Variables* in 1987. In that it is shown how a wide variety of formal theories in the first-order predicate calculus, including set theory, can be axiomatized equivalently in purely quantifier-free relation-algebraic terms, even though those do not suffice in general to axiomatize first-order logic. The work on the algebraization of the full first-order logic with equality is exposited in the two substantial volumes of *Cylindric*

Algebras (1971–1985), written in collaboration with Leon Henkin and Donald Monk.

See also Model Theory.

Bibliography

WORKS BY TARSKI

Logic, Semantics, Metamathematics: Papers from 1923 to 1938. 2nd ed. Translated by J. H. Woodger, edited by John Corcoran. Indianapolis, IN: Hackett, 1983.

Collected Papers, Vols. 1–4, edited by Steven R. Givant and Ralph N. McKenzie. Basel: Birkhäuser, 1986.

WORKS BY TARSKI WITH OTHERS

Henkin, Leon, Donald Monk, and Alfred Tarski. *Cylindric Algebras.* 2 vols. Amsterdam: North-Holland, 1971–1985.

Tarski, Alfred, and Steven R. Givant. *A Formalization of Set Theory without Variables*. Providence, RI: American Mathematical Society, 1987.

Tarski, Alfred, and Steven R. Givant. "Tarski's System of Geometry." *Bulletin of Symbolic Logic* 5 (1999): 175–214.

WORKS ON TARSKI

Feferman, Anita Burdman, and Solomon Feferman. *Alfred Tarski: Life and Logic*. New York: Cambridge University Press, 2004.

Givant, Steven. "Bibliography of Alfred Tarski." *Journal of Symbolic Logic* 51 (4) (1986): 913–941.

Tarski's work is surveyed in a number of articles in the *Journal of Symbolic Logic* 51 (4) (1986) and 53 (1) (1988). There is a considerable secondary literature on Tarski's work, including the proceedings of the Tarski Centenary Conference held in Warsaw in 2001 that appeared in the *Annals of Pure and Applied Logic* 126 (1–3) (2004) and 127 (1–3) (2004).

Anita Burdman Feferman (2005)
Solomon Feferman (2005)

TAULER, JOHANNES
(c. 1300–1361)

The German mystic Johannes Tauler entered the Dominican order at Strasbourg about the age of fifteen and probably studied in the Dominican *studium generale* at Cologne, where he may have been taught by Meister Eckhart. He was certainly influenced by the latter and by the contemplative movement known as the *Gottesfreunde* (Friends of God). He was in Strasbourg at the time of Pope Innocent XXII's interdict on the city for taking the wrong side in the war between different sections of the Holy Roman Empire, but there is no good evidence for the story that during the Black Death he defied the interdict by administering sacraments to the dying. He remained a loyal and orthodox member of the church. Much legendary material surrounds his life, and various spurious works are attributed to him. It was on the basis of these sources that some earlier scholars mistakenly thought of Tauler as a precursor of the Reformation.

In his sermons, Tauler geared mystical teachings, which made use of Eckhartian and Neoplatonic concepts, to practical purposes. He was deeply committed to the view that mystical experiences are a nourishment to the soul in supporting the individual in a life of active love and that there are behavioral criteria for estimating their worth. He believed that in this active life we may possess God through a fusion of the divine and human wills. However, far from reducing contemplative religion to the exercise of good works, Tauler believed that the love of God and the love of men go together and that the former finds its consummation in the inner union of the soul with the Creator.

In principle, all men should be capable of this return of the soul to its Source (the notion of return was typical of the Neoplatonic tradition with which Tauler was acquainted). Two qualifications, however, must be made. First, the way of return, according to Tauler's account, involves great heroism and suffering. The creaturely side of man must be crucified. Self-mortification is a sign of burning love of God, and eventually the friend of God may acquire a real desire for, rather than an aversion to, suffering. In this emphasis on suffering, Tauler was strongly Christocentric in his preaching. But second, the fall of man has so tainted the human being that the divine light, which illuminates the contemplative and brings about the return to God, is something that man cannot achieve on his own. It is the gift of divine grace. Thus, the culmination of the mystic's quest is not a personal achievement of the mystic, but an enjoyment granted from beyond.

The importance of the need for grace gave Tauler's mysticism a firmly orthodox character. Nevertheless, he maintained that the operation of divine grace requires a right attitude on the part of men. Tauler speaks of God as a fisherman who lets down a baited hook into the ocean. Those fish who are not disposed toward the bait will not be hooked. This simile had its basis in Tauler's account of human psychology.

According to his psychology, three aspects of the soul can be distinguished. At the deepest level is the Ground of the soul—otherwise referred to as the Spark, the Apex (*Punkt*), and God in the soul—a concept deriving from Eckhart's teaching. However, Tauler is eager to assert that the Ground is God-given and is not an intrinsic, natural property of the individual. At another level, the soul possesses intellect, sense faculties, and will. Third, there is

what Tauler refers to as the heart (*das Gemüt*). The attitude of the individual toward the divine Being is determined by whether his heart is turned toward the Ground or away from it. If the former, God will descend, draw the spirit up to himself, and unite it with him. Man's choice is therefore essentially a choice of disposition. Once this choice has been made, God through his grace will conform the human will to his own. Thus, the end of the contemplative life is a state in which the mystic is, so to speak, "taken over" by God, so that all his actions express God's purposes rather than his own.

See also Eckhart, Meister; Mysticism, History of; Neoplatonism; Reformation.

Bibliography

Tauler's works include *Twenty-five Sermons,* translated by Susanna Winkworth, 2nd ed. (London, 1906); *The Sermons and Conferences of John Tauler,* edited and translated by Walter Elliott (Washington, DC: Apostolic Mission House, 1910); *Die Predigten Taulers,* edited by F. Vetter (Berlin: Weidmann, 1910); and *Johannes Tauler—Predigten,* edited by G. Hofmann (Freiburg, 1961).

See also J. M. Clark, *The Great German Mystics* (Oxford: Blackwell, 1949).

Ninian Smart (1967)

TAUTOLOGY

See *Logical Terms, Glossary of*

TAYLOR, ALFRED EDWARD

(1869–1945)

Alfred Edward Taylor, the British philosopher, was born at Oundle, Northamptonshire, and educated at New College, Oxford. His teaching experience was unusually varied: He was a fellow of Merton College, Oxford, 1891–1898; lecturer at Owens College, Manchester, 1898–1903; professor of logic and metaphysics at McGill University, Montreal, 1903–1908; professor of moral philosophy at St. Andrews University, 1908–1924; and professor of moral philosophy at Edinburgh, 1924–1941. His interests were also varied; not only was he an authority on Greek philosophy but he also made extensive contributions to current thinking on ethics, metaphysics, and the philosophy of religion. Taylor's thought was within the tradition of British neo-Hegelianism, but as his philosophy developed, other influences came in also, though he remained firmly attached to a theistic and spiritualist interpretation of reality.

In the field of Greek philosophy, Taylor is noted chiefly for his work on Plato. He gives a full-scale exposition of Plato's thought in *Plato: The Man and His Work* (London, 1926) and a detailed study of Plato's cosmology in *A Commentary on Plato's Timaeus* (Oxford, 1928). Even in these works Taylor's own philosophical interests assert themselves, notably in his attempt to minimize alleged differences between the Platonic and biblical ways of understanding creation and in his contention that the Demiurge of Plato is a creator in the full sense of the word.

Taylor's philosophy found early expression in *The Problem of Conduct* (London, 1901) and in *Elements of Metaphysics* (London, 1903). At this stage he was influenced primarily by F. H. Bradley and English idealism. Later, Platonism, Thomism, and even Bergsonism became important additional influences on his mature thought as expressed in *The Faith of a Moralist* (London, 1930), a work based on his Gifford Lectures of 1926–1928.

Here Taylor claims that if we take moral experience seriously, we must recognize that it points beyond itself to, and is completed in, religion and that we are thus led to theism. Moral experience does deserve to be taken seriously, for facts and values are given together and never occur in separation in our concrete experience of the world. A naturalistic philosophy that allows reality to fact but denies it to value is guilty of a false abstraction. This argument about the concreteness of experience is a necessary prolegomenon to Taylor's position as a whole, for if the values of the moral life were divorced from the facts of the world, then no argument from moral experience to the nature of reality could succeed.

Taylor's attempt to move from the facts of moral experience to a religious metaphysic turns on two main considerations. The first concerns the nature of the good at which the moral life aims. Is it a temporal good or is it an eternal good? Taylor contends that even to be able to ask this question and to be aware of the temporal dimension of our existence is to have begun to transcend the form of temporality. Further reflection shows that no merely temporal goods can satisfy the demands of man's nature. Such goods are defective in various ways; for instance, they can be attained only successively and cannot be enjoyed simultaneously. One might answer, of course, that this merely shows that human aspirations are doomed to frustration, but Taylor rejects this and claims that the facts of moral striving point to an eternal good.

The second consideration concerns the question of how such an eternal good is to be attained. Can man of himself attain to an eternal good? Taylor answers in the negative, for he sees sin and guilt as inhibiting the moral life and preventing man from reaching his goal. But again he does not accept this frustration as final. Man's unavailing endeavors to reach toward the eternal good are met by what Taylor calls the initiative of the eternal. This is the divine grace that reaches down to man and enables his moral fulfillment. Thus, the moral life finds its completion in religion; if we deny this, we are bound to say that the moral life is self-stultifying. To take its demands seriously is to believe that it makes sense, and according to Taylor, it makes sense only in the light of a theistic worldview.

The individual destined for an eternal good and enabled by divine grace to move toward that good is also assured of immortality. Hence, from consideration of the implications of the moral life alone we arrive at a kind of minimal theology, so to speak, of God, grace, and immortality. But Taylor, who was himself a devout churchman of the Anglican communion, asks whether this minimal theology does not, like morality, point beyond itself for completion. The concreteness that characterizes Taylor's starting point is apparent again in his conclusions, as he argues that a bare philosophical theism needs to be embodied in an actual historical religion. Although the philosopher does not appeal to revelation, his analysis can, Taylor believed, bring us to the point at which we see the need for a concrete revelation to complete the bare schema of philosophical theology. Philosophy makes it reasonable to expect that there would be such a revelation, and Taylor thinks that Christian revelation especially fulfills this expectation. He continued to wrestle with the problems of religion, which provide the themes for two of his last books, *The Christian Hope of Immortality* (London, 1938) and *Does God Exist?* (London, 1943).

See also Bradley, Francis Herbert; Cosmology; Ethics, History of; Good, The; Idealism; Moral Arguments for the Existence of God; Plato; Platonism and the Platonic Tradition; Thomism.

Bibliography

Additional works by Taylor are *Varia Socratica* (Oxford: J. Parker, 1911); *The Laws of Plato* (London: Dent, 1934); *Philosophical Studies* (London: Macmillan, 1934); and *Aristotle* (London: T. Nelson, 1943).

W. D. Ross, "Alfred Edward Taylor, 1869–1945," in *Proceedings of the British Academy* 31 (1945): 407–424, contains a bibliography of Taylor's writings.

John Macquarrie (1967)

TEILHARD DE CHARDIN, PIERRE

(1881–1955)

Pierre Teilhard de Chardin, the paleoanthropologist and Roman Catholic priest who advocated a doctrine of cosmic evolution, was born in Sarcenat, France. At the age of eighteen he entered the Jesuit order, and he remained a faithful member of it for the rest of his life. By the time he was ordained, his interest in science and the reading of Henri Bergson resulted in his becoming a fervent evolutionist. Association with the Bergsonian scholar Édouard Le Roy also deeply influenced his thought. It became one of Teilhard's aims to show that evolutionism does not entail a rejection of Christianity. He likewise sought to convince the church that it can and should accept the implications of the revolution begun by Charles Darwin, but he met with uniform opposition from ecclesiastical superiors.

In 1926 he was expelled from the Catholic Institute in Paris, at which he had taught after returning from service in World War I. Until 1946 he was "exiled" in China, where he participated in paleontological researches that led to the discovery of Beijing man. He also completed the manuscript of his major work, *Le phénomène humain (The Phenomenon of Man)*; but despite repeated applications to Rome he was refused permission to publish it. After his death the appearance of the work, along with his other essays, gave rise to controversies both inside and outside the church.

The evolutionism that Teilhard advocated is all-embracing and characterizes much more than living things. Teilhard contended that long before living things appeared on Earth, the basic stuff of the cosmos was undergoing irreversible changes in the direction of greater complexity of organization. Hence, nonliving nature is profoundly historical. It is not a system of stable elements in a closed equilibrium. On the contrary, it conforms at all stages to a "law of complexification," comparable in importance to the law of gravity and illustrated by the vast array of organic forms that have appeared in evolutionary history. The most recent of these forms is man.

When viewed "from without" by the physical sciences, man is a material system in the midst of other material systems. But each individual man experiences himself "from within" as a conscious being. Consciousness is thus directly identifiable as "spiritual energy." Teilhard maintained that *all* constituents of the cosmos, from elementary particles to human beings, have "a conscious

inner face that everywhere duplicates the material external face." Since this is so, the physical evolution of the cosmic stuff will at the same time be an evolution of consciousness. The more highly integrated a material system, the more developed its psychical interior will be. Thus, in the human brain an intense concentration, or "involution," of cells has led to the emergence of self-conscious thought, the most advanced stage reached by evolution thus far.

But greater developments are in store from the evolutionary convergence of disparate cultures and forms of consciousness. Man is now a single, interbreeding species expanding on the finite, spherical surface of the planet and still showing signs of biological immaturity. Furthermore, his capacity for self-conscious thought and the production of cultures has added a new "layer" to Earth's surface, which Teilhard calls the "noosphere," distinct from, yet superimposed on, the biosphere. The noosphere, or "thinking layer," forms the unique environment of man, marking him off from all other animals. The evolutionary convergence that it makes possible will be manifested externally in the unification of all human cultures into a single world culture. Paralleling this, a movement toward psychical concentration will occur, so that the noosphere will become involuted in a Hyperpersonal Consciousness "at a point which we might call *Omega*." Here evolution will reach the terminal phase of convergent integration.

Teilhard's concept of Point Omega is obscure, like other aspects of his evolutionism, because it is essentially the expression of a mystical vision. Omega is not identical with God but, rather, is God insofar as he determines the direction and goal of cosmic history. Hence, the evolutionary process is orthogenetic, although neither vitalistic nor wholly devoid of chance events. The integration of all personal consciousnesses at Omega will be achieved, Teilhard urged, through love, which forms *le milieu divin*, the spirit of Christ at work in nature.

Teilhard's doctrine tends to become pantheistic in certain of its formulations. On the whole, it is difficult to reconcile Teilhard's views either with orthodox Christian teaching or with a scientific theory of evolution. Yet the prose poetry of *The Phenomenon of Man* has stirred the imagination of theologians, philosophers, and scientists, even when it has not won their assent.

See also Bergson, Henri; Darwin, Charles Robert; Evolutionary Theory; Le Roy, Édouard; Pantheism.

Bibliography

WORKS BY TEILHARD DE CHARDIN

Oeuvres, 9 vols. Paris, 1955–. Ten volumes are projected.

Le phénomène humain. Paris: Seuil, 1955. Translated by Bernard Wall as *The Phenomenon of Man*. New York: Harper, 1959.

Lettres de voyage. Paris: Grasset, 1956. Translated by René Hague and others as *Letters from a Traveller*. London: Collins, 1962.

Le milieu divin. Paris: Seuil, 1957. Translated by Bernard Wall and others as *The Realm of the Divine*. New York: Harper, 1960.

L'avenir de l'homme. Paris: Seuil, 1959. Translated by Norman Denny as *The Future of Man*. New York: Harper, 1964.

Hymne de l'univers. Paris: Seuil, 1964. Translated by Simon Bartholemew as *Hymn of the Universe*. London: Collins, 1965.

WORKS ON TEILHARD DE CHARDIN

Cuénot, Claude. *Pierre Teilhard de Chardin, les grandes étapes de son évolution*. Paris: Plon, 1958. Translated by Vincent Colimore and edited by René Hague as *Teilhard de Chardin: A Biographical Study*. Baltimore: Helicon Press, 1965.

Francoeur, R. T., ed. *The World of Teilhard*. Baltimore: Helicon Press, 1961.

Raven, C. E. *Teilhard de Chardin: Scientist and Seer*. London: Collins, 1962.

Tresmontant, C. *Pierre Teilhard de Chardin, His Thought*. Baltimore: Helicon Press, 1959.

OTHER RECOMMENDED WORKS

Chauchard, Paul. *Man and Cosmos: Scientific Phenomenology in Teilhard de Chardin*. New York: Herder and Herder, 1965.

Delfgaauw, Bernard. *Evolution: The Theory of Teilhard de Chardin*. New York: Harper & Row, 1969.

Grau, Joseph A. *Morality and the Human Future in the Thought of Teilhard de Chardin: A Critical Study*. Rutherford: Fairleigh Dickinson University Press; Cranbury, NJ: distributed by Associated University Presses, 1976.

Lukas, Mary, and Ellen Lukas. *Teilhard*. Garden City, NY: Doubleday, 1977.

Rideau, Emile. *The Thought of Teilhard de Chardin*. New York: Harper & Row, 1967.

Speaight, Robert. *The Life of Teilhard de Chardin*. New York: Harper & Row, 1967.

Teilhard de Chardin, Pierre, and Henri de Lubac. *Letters from Paris, 1912–1914*. New York: Herder and Herder, 1967.

Teilhard de Chardin, Pierre, and Pierre Leroy. *Letters from My Friend, Teilhard de Chardin, 1948–1955: Including Letters Written during His Final Years in America*. New York: Paulist Press, 1980.

Teilhard de Chardin, Pierre. *Letters from Hastings, 1908–1912*. New York: Herder and Herder, 1968.

Teilhard de Chardin, Pierre. *Pierre Teilhard de Chardin: Writings*. Edited by Ursula King. Maryknoll, NY: Orbis Books, 1999.

Teilhard de Chardin, Pierre. *The Making of a Mind: Letters from a Soldier-priest, 1914–1919*. London, Collins, 1965.

Wildiers, N. M. *An Introduction to Teilhard de Chardin*. New York: Harper & Row, 1968.

T. A. Goudge (1967)
Bibliography updated by Michael Farmer (2005)

TELEOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD

The "Teleological Argument for the existence of God" is a member of the classic triad of arguments, which is completed by the Ontological Argument and the Cosmological Argument. Stated most succinctly, it runs:

> The world exhibits teleological order (design, adaptation).
>
> Therefore, it was produced by an intelligent designer.

To understand this argument, we must first understand what teleological order is.

TELEOLOGICAL ORDER

Generally speaking, to say that a group of elements is ordered in a certain way is to say that they are interrelated so as to form a definite pattern, but the notion of a definite pattern is vague. Any set of elements is interrelated in one way rather than another, and any complex of interrelations might be construed by someone as a definite pattern. Certain patterns are of special interest for one reason or another, and when one of these is exhibited, the complex would ordinarily be said to be ordered. Thus, when the elements form a pattern in whose perception we take intrinsic delight, we can speak of aesthetic order. When there are discernible regularities in the way, certain elements occur in spatiotemporal proximity, we can speak of causal order. The distinctive thing about teleological (Greek, *telos,* "end" or "goal") order is that it introduces the notion of processes and structures being fitted to bring about a certain result.

The usual illustrations of teleological order are from living organisms. It is a common observation that the anatomical structures and instinctive activities of animals are often nicely suited to the fulfillment of their needs. For example, the ears of pursuing, carnivorous animals, like the dog and the wolf, face forward so as to focus sounds from their quarry, while the ears of pursued, herbivorous animals, like the rabbit and the deer, face backward so as to focus sounds from their pursuers.

Examples of instinctive behavior are even more striking. The burying beetle deposits its eggs on the carcass of a small animal and then covers the whole "melange" with dirt to protect it until the young hatch out and find an ample supply of (hardly fresh) meat at hand.

If we are going to distinguish teleological order from causal order, we shall have to make explicit the tacit assumption that the result the structure or process in question is fitted to bring about is of value. Otherwise, *any* cause-effect relationship would be a case of teleological order. It is just as true to say that wind is fitted to produce the result of moving loose dirt into the air as it is to say that the mechanism of the eye is fitted to produce sight. The latter would be counted as an example of "design," whereas the former would not, because we regard sight as something worth having, whereas the movement of dirt through the air is not generally of any value. This has the important implication that insofar as it is impossible to give an objective criterion of value, it will not be an objective matter of fact that teleological order is or is not exhibited in a given state of affairs.

It is important to note that the term *design,* as used in this argument, does not by definition imply a designer. If it did, there could be no argument from design to the existence of God; we would have to know that the phenomena in question were the work of a designer before we could call them cases of design. We must define *design* in such a way as to leave open the question of its source. We have design in the required sense when things are so ordered that they tend to perform a valuable function. We might put this by saying that things are ordered as they would be if some conscious being had designed them, but in saying this we are not committing ourselves to the proposition that a mind has designed them. The equivalent terms *adaptation* and *teleological order* are not so liable to mislead in this way.

Arguments for the existence of God have been based on kinds of order other than the teleological. Exhortations to move from a consideration of the starry heavens to belief in God constitute an appeal to aesthetic order. It is sometimes claimed that we must postulate an intelligent creator to explain the regularity with which the solar system operates. Here it is causal order that is involved. Arguments like these are often not clearly distinguished from those based on teleological order, to which we shall confine our attention.

ARGUMENTS FROM PARTICULAR CASES OF DESIGN

The simplest form of the argument is that in which we begin with particular cases of design and argue that they can be adequately explained only by supposing that they were produced by an intelligent being. Thus William Paley, an eighteenth-century philosopher, in a classic formulation of the argument concentrated on the human eye as a case of design, stressing the ways in which various parts of the eye cooperate in a complex way to produce sight. He argued that we can explain this adaptation of means to end only if we postulate a supernatural designer. This is the heart of the teleological argument—the claim that adaptation can be explained only in terms of a designer. It always rests, more or less explicitly, on an analogy with human artifacts. Thus, Paley compared the eye to a watch and argued as follows: If one were to find a watch on a desert island, one would be justified in supposing that it was produced by an intelligent being. By the same token (the adjustment of means to ends) one is entitled, upon examination of the human eye, to conclude that it was produced by an intelligent being.

If it is asked why we should take artifacts as our model, the answer would seem to be this. Artifacts are certainly cases of design. In a watch, for example, the structure is well suited to the performance of a valuable function: showing the time. With artifacts, unlike natural examples of design, we have some insight into what is responsible for the adjustment of means to end. We can understand it because we can see how this adjustment springs from the creative activity of the maker, guided by his deliberate intention to make the object capable of performing this function. Hence, in natural cases of adaptation where the source of the adaptiveness is not obvious, we have no recourse but to employ the only way we know of rendering such phenomena intelligible—supposing them to stem from conscious planning. Since we do not observe any planner at work, we must postulate an invisible planner behind the scenes.

CRITICISMS. The comparison to artifacts was attacked by David Hume in his *Dialogues concerning Natural Religion,* in which he suggested that the production of artifacts by human planning is no more inherently intelligible than the production of organisms by biological generation. Why, asked Hume, should we take the former rather than the latter as the model for the creation of the world? Even if we admit that the world exhibits design, why are we not as justified in supposing that the world was generated from the sexual union of two parent worlds as in supposing that it was created by a mind in accordance with a plan? In answer to Hume it might be argued that creation gives a more satisfactory and a more complete explanation than generation because the generation consists of a reproduction of the same kind of thing and hence introduces another entity that raises exactly the same kind of question. If we are initially puzzled as to why a rabbit has organs that are so well adapted to the satisfaction of its needs, it does not help to be told that it is because the rabbit sprang from other rabbits with just the same adaptive features. If, on the other hand, we could see that the rabbit had been deliberately constructed in this way so that its needs would be satisfied, we would be making progress. To this Hume would reply that the mind of the designer also requires explanation. Why should the designer have a mind that is so well fitted for designing? Thus, this explanation also leaves problems dangling, but at least it is not just the same problem. If we were to reject every explanation that raised fresh problems, we would have to reject all of science.

DARWINIAN THEORY OF EVOLUTION. The development of the Darwinian theory of evolution opened up the possibility of a more serious alternative to the theistic explanation. According to this theory, the organic structures of today developed from much simpler organisms by purely natural processes. In this theory (as developed since Charles Darwin) two factors are considered to play the major role: mutations and overpopulation. (A mutation occurs when an offspring differs from its parents in such a way that it will pass this difference along to its offspring, and they will pass it along, and so on. It is a relatively permanent genetic change.)

The way these factors are thought to work can be illustrated by taking one of the cases of adaptation cited above. If we go back far enough in the ancestry of the dog, we will discover ancestors that did not have ears facing forward. Now let us suppose that a mutation occurred that consisted of an ear turned somewhat more forward than had been normal. Granting that organisms tend to reproduce in greater numbers than the environment can support, and hence that there is considerable competition for the available food supply, it follows that any feature of a given organism that gives it any advantage over its fellows in getting food or in avoiding becoming prey will make it more likely to survive and pass along its peculiarity to its offspring. Thus, within a number of generations we can expect the front-turned-ear proto-dogs to replace the others and be left in sole possession of the field. Since mutations do occur from time to time, and since some of them are favorable, we have a set of purely natural factors

by whose operation the organic world can be continuously transformed in the direction of greater and greater adaptation.

The Darwinian theory aspires to do no more than explain how more complex organisms develop from less complex organisms. It has nothing to say about the origins of the simplest organisms. However, no matter how simple the organism, its structure must be fitted to the satisfaction of its needs, or it will not survive. Therefore, Darwinian theory is not a complete explanation of the existence of teleological order in the world; it merely tells us how some cases develop from other cases. Hence, it alone is not an alternative to the theistic explanation, but in principle there is no reason why it should not be supplemented by a biochemical theory of the origin of life from lifeless matter. No such theory has yet been completely established, but progress is being made. When and if this is done, there will be an explanation of design in living organisms for which there is empirical support, and it can no longer be claimed that theism represents the only real explanation of such facts.

WHAT FOLLOWS FROM THE ARGUMENT. The other major deficiency in Paley's form of the argument is that, even if valid, it does not go very far toward proving the existence of a theistic God. The most we are warranted in concluding is that each case of design in the natural world is due to the activity of an intelligent designer. Nothing is done to show that all cases of design are due to one and the same designer; the argument is quite compatible with polytheism or polydaemonism, in which we would have one supernatural designer for flies, another for fish, and so on. Even if there is one, and only one, designer, nothing is done to show that this being is predominantly good rather than evil; neither is anything done to show that he is infinitely powerful or wise, rather than limited in these qualities. Of course the theist might seek to supplement this argument by others, but by itself it will not bear the weight.

ARGUMENT FROM THE UNIVERSE AS A WHOLE

No argument that, like the Teleological Argument, is designed to show that facts in nature require a certain explanation, can establish the existence of a deity absolutely unlimited in power, knowledge, or any other respect. By such reasoning we can infer no more in the cause than is required to produce the effect. This deficiency is irremediable. However, there is a simple way of eliminating competing scientific claims—by starting from the universe as a whole rather than from individual instances of design within the universe. There are different ways of doing this. We might think of the whole universe as instrumental to some supreme goal, or we might think of the universe as a unified system of mutually adjusted and mutually supporting adaptive structures.

Taking the whole universe as instrumental to some supreme goal would give us the strongest argument, for here the analogy with consciously designed artifacts is strongest. An artifact like a house, ship, or watch is designed for the realization of goals outside its internal functioning; it is intended to be used for something. Therefore, if the analogy with artifacts is the main support for the notion that the universe was the result of conscious planning, that support would be firmest if grounds were presented for thinking that the universe as a whole was well fitted to be used for something. And if this something were of maximum value, we would then have a basis for attributing supreme goodness to the designer.

However, this alternative is rarely taken, largely because it is difficult to decide on a suitable candidate for, in Alfred, Lord Tennyson's words, the "far-off divine event, toward which the whole creation moves." The most common suggestions are the greater glory of God and the development of moral personality. But in regard to the first, no one can really understand just what it would mean for a God who is eternally perfect to receive greater glory, and in regard to the second, even if we can overcome doubts that moral development is worth the entire cosmic process, it would seem impossible ever to get adequate grounds for the proposition that everything that takes place throughout all space and time contributes to this development.

The second interpretation, that the universe is a unified system of mutually adjusted and mutually supporting adaptive structures—has been tried more often. So conceived, the argument will run as follows.

(1) The world is a unified system of adaptations.

(2) We can give an intelligible explanation of this fact only by supposing that the world was created by an intelligent being according to some plan.

(3) Therefore, it is reasonable to suppose that the world was created by an intelligent being.

The famous formulation of the argument in Hume's *Dialogues* makes explicit the analogy on which, as we have seen, step two depends. Hume's formulation, which is substantially equivalent to the above, runs as follows.

(1) The world is like a machine.

(2) Machines are made by human beings, in accordance with plans.

(3) Like effects have like causes.

(4) Therefore, the world probably owes its existence to something like a human being, who operates in accordance with a plan.

TYPES OF ADAPTATION. If one is to think of the whole universe as a system of connected adaptations, he will consider kinds of adaptation other than that exemplified by the fitness of organisms to the conditions of life; this kind alone will not bear the whole weight. F. R. Tennant, who has developed the weightiest recent presentation of the teleological argument in his *Philosophical Theology,* discusses six kinds of adaptation:

(1) The intelligibility of the world. The world and the human mind are so related that we can learn more and more without limit.

(2) The adaptation of living organisms to their environments. This is the kind on which we have been concentrating.

(3) The ways in which the inorganic world is conducive to the emergence and maintenance of life. Life is possible only because temperatures do not exceed certain limits, certain kinds of chemical processes go on, and so on.

(4) The aesthetic value of nature. Nature is not only suited to penetration by the intellect; it is also constituted so as to awaken valuable aesthetic responses in man.

(5) The ways in which the world ministers to the moral life of men. For example, through being forced to learn something about the uniformities in natural operations, men are forced to develop their intelligence, a prerequisite to moral development. And moral virtues are acquired in the course of having to cope with the hardships of one's natural environment.

(6) The overall progressiveness of the evolutionary process.

Tennant admits that no one of these forms of adaptiveness is a sufficient ground for the theistic hypothesis, but he maintains that when we consider the ways in which they dovetail, we will see theism to be the most reasonable interpretation. Thus, the adjustment of lower organisms to the environment takes on added significance when it is seen as a stage in an evolutionary process culminating in man, which in turn is seen to be more striking when we realize the ways in which nature makes possible the further development of the moral, intellectual, and aesthetic life of man.

When the argument takes this form, it is no longer subject to competition from scientific explanations of the same facts. If our basic datum is a certain configuration of the universe as a whole, science can, by the nature of the case, offer no explanation. Science tries to find regularities in the association of different parts, stages, or aspects within the physical universe. On questions as to why the universe as a whole exists, or exists in one form rather than another, it is silent. Ultimately this is because science is committed to the consideration of questions that can be investigated empirically. One can use observation to determine whether two conditions within the universe are regularly associated (increase of temperature and boiling), but there is no way to observe connections between the physical universe as a whole and something outside it. Therefore, there is no scientific alternative to the theistic answer to the question "Why is the universe a unified system of adaptations?"

ALTERNATIVE EXPLANATIONS OF ADAPTATION. What alternatives to the theistic explanation of adaptation are there? In the literature on the subject one often encounters the suggestion that we have this kind of universe by chance. If we dismiss the animistic notion of chance as a mysterious agent, the suggestion that we have this kind of universe by chance boils down to a refusal to take the question seriously. It may be said that the fact that the universe as a whole exhibits teleological order is not the sort of thing that requires explanation. It is difficult to see what justification could be given for this statement other than an appeal to the principle that sense observation is the only source of knowledge and/or meaning.

One cannot perceive by the senses any relation between the physical universe as a whole, or any feature thereof, and something outside it on which it depends. Hence, an extreme form of empiricism would brand the question posed by the Teleological Argument as fruitless or even meaningless. If, on the other hand, the question is taken seriously, any answer will be as metaphysical as the theistic answer, for it is really a question as to what characteristics are to be attributed to the cause (or causes) of the universe. Do the relevant facts about the world most strongly support the theistic position that the cause is a perfectly good personal being who created the universe in the carrying out of a good purpose? Or is there some other view that is equally, or more strongly, supported by

the evidence? The Manichaeans held that the physical universe was the work of a malevolent deity and that man must separate himself from the body in order to escape this diabolical power and come into contact with the purely spiritual benevolent deity. It has also been held in many religions that the universe is the joint product of two or more deities who differ markedly in their characteristics. In Zoroastrianism it is held that the world is the battleground of a good deity and an evil deity, the actual state of affairs bearing traces of both. Indian religious philosophy typically regards the universe as resulting from a nonpurposive manifestation of, or emanation from, an absolute unity that is not personal in any strict sense.

EXTENT OF ADAPTIVENESS IN THE UNIVERSE. To evaluate the Teleological Argument in the light of competing explanations, we must ask whether the extent of adaptiveness in the universe is sufficient to warrant the theistic conclusion. As the problem is formulated in Hume's *Dialogues,* is there a close enough analogy between the universe and a machine? This requires judging the relative proportion of adaptive features to nonadaptive or maladaptive features. In addition to taking account of Tennant's enumeration of the ways in which the shape of things is instrumental to the realization of valuable ends, we must look at the other side of the picture and try to form an adequate impression of (1) the ways in which the shape of things is neutral, providing neither for good nor for evil, and (2) the ways in which the shape of things frustrates the search for value.

As for (1), as far as we can see, the distribution of matter and the variety of chemical elements in the world, to take two examples at random, could have been very different from what they are without reducing the chances of sentient beings leading satisfying lives.

As for (2), we begin to trespass onto the problem of evil, except that here we are interested in suffering and frustration not as possible disproofs of theism but as affecting the cogency of the Teleological Argument for the existence of God. There are many ways in which the organization of the world makes for disvalue rather than value in the lives of men and other sentient creatures. One need only mention the numerous sources of disease, the incidence of malformed offspring, the difficulty of attaining optimum conditions for the development of healthy personalities, and the importance of antisocial tendencies in human nature. It is quite possible, of course, that all the things that seem to be unfortunate features of the world as it exists are necessary elements in the best of all possible worlds. If we already believe that the world is the creation of a perfect deity, that carries with it the belief that these apparent evils are necessary even though we cannot see how they are. However, if we are trying to establish the existence of a perfect deity, we have to proceed on the basis of what we can see. And since, so far as we can see, the world would be better if the features listed above were altered, we cannot argue that the state of adaptiveness in the world requires explanation in terms of a perfectly good, omnipotent deity. But we have already seen, on other grounds, that the Teleological Argument cannot be used to establish the existence of a being unlimited in any respect.

The serious problem that remains is whether the total picture of adaptation and maladaptation, so far as we have it, gives sufficient support to the hypothesis that the world represents the at least partial implementation of a plan that is at least predominantly good. To resolve this problem we must weigh opposite factors and arrive at a final judgment of their relative importance. Unfortunately there are no real guidelines for this task. No one knows how much adaptation, relative to maladaptation, would warrant such a conclusion; and even if he did, he would not know what units to employ to perform the measurement. What is to count as one unit of adaptation? Do we count each individual separately, or is each species one unit? How can we compare the value of human knowledge with the disvalue of disease? It would seem that on this issue different positions will continue to be taken on the basis of factors outside the evidence itself.

See also Cosmological Argument for the Existence of God; Darwin, Charles Robert; Darwinism; Evil, The Problem of; God/Isvara in Indian Philosophy; Hume, David; Mani and Manichaeism; Ontological Argument for the Existence of God; Paley, William; Physicotheology; Popular Arguments for the Existence of God; Tennant, Frederick Robert; Theism, Arguments For and Against; Zoroastrianism.

Bibliography

In the Middle Ages there was general acceptance of an Aristotelian physics, according to which even purely physical processes were explained in terms of the natural tendency of a body toward an end. (Fire naturally tends to come to rest at the periphery of the universe.) Given this background, it was argued that the consideration of any natural processes led to the postulation of a designer. The argument in this form is found in Thomas Aquinas, *Summa Theologiae,* Part I, Question 2, Article 3. Contemporary Thomistic statements try to adjust this line of thought to modern physics. See G. H. Joyce, *The Principles of Natural Theology*

(New York: AMS Press, 1972); Réginald Garrigou-Lagrange, *God, His Existence and His Nature,* 2 vols. (St. Louis: Herder, 1934–1936); and D. J. B. Hawkins, *The Essentials of Theism* (New York: Sheed and Ward, 1949).

The influential presentation by the eighteenth-century thinker William Paley is to be found in his *Natural Theology: Or, Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature* (Indianapolis: Bobbs-Merrill, 1963). Important more recent formulations include F. R. Tennant, *Philosophical Theology,* 2 vols. (New York, 1928–1930), Vol. II, Ch. 4, and A. E. Taylor, *Does God Exist?* (New York: Macmillan, 1947).

Acute criticisms of the argument are to be found in David Hume, *Dialogues concerning Natural Religion*; Immanuel Kant, *Critique of Pure Reason,* Book II. Ch. 3; C. D. Broad, *Religion, Philosophy, and Psychical Research* (New York: Harcourt Brace, 1953); John Laird, *Theism and Cosmology* (New York, 1942); and J. J. C. Smart, "The Existence of God," in *New Essays in Philosophical Theology,* edited by Antony Flew and Alasdair MacIntyre (London: SCM Press, 1955).

William P. Alston (1967)

TELEOLOGICAL ARGUMENT FOR THE EXISTENCE OF GOD [ADDENDUM]

The argument from the causal order of the universe to the existence of an intelligent designer has experienced a revival in the work of Richard Swinburne and others (see also Collins 2002). Swinburne's cumulative case for God's existence is an argument to the best explanation, citing various pieces of data or evidence that are (a) relatively improbable on an assumption of naturalism but (b) relatively probable if theism is true (Swinburne 1979). One such datum is that the universe conforms to simple, mathematically formulable scientific laws—that is it exhibits *causal order*. (This differs from *spatial order*, an arrangement of parts that serves the purpose of a greater whole, as in an organism's suitability for its environment.) While theories of evolution partially undermine the argument from spatial order, they leave the following argument from causal order untouched:

(1) The universe conforms to formulas. ("The orderliness of nature is a matter of the vast uniformity in powers and liabilities of bodies throughout endless time and space, and also in the paucity of kinds of components of bodies" [Swinburne 1979, p. 140].)

(2) There are only two kinds of explanation for phenomena: scientific explanation and personal explanation (Swinburne 1979, pp. 140–141).

(3) No scientific explanation of (1) is possible. (The data in (1) concern the most basic or ultimate constituents of material bodies and the most fundamental physical laws; scientific explanation reaches no further.)

(4) Thus, either there is a personal explanation for (1) or it has no explanation (i.e., it occurs by chance).

(5) That there is a personal explanation for (1) is more probable than that it has no explanation.

(6) Hence, (1) confirms the hypothesis of a personal cause of the universe.

Naturalism offers no explanation for the causal order and fundamental intelligibility of the universe. Indeed, this type of order is surprising if the universe did not result from purpose or design. On the other hand, a personal being has reasons to produce causal order in the universe, due to aesthetic considerations—for example, order is more beautiful than chaos—and other value considerations (a universe with intelligent beings who can understand their world is preferable to a universe with no intelligent beings or with rational creatures whose attempts to "read the book of Nature" cannot succeed). Causal order combines with additional data that exhibit properties (a) and (b) above to support the further conclusion that *theism is more probable than naturalism*, even if the probability of theism is not greater than 0.5 or fifty percent.

Critics point to the difficulty of assigning a priori objective probabilities to large-scale metaphysical theories. Perhaps this can be blunted by appealing to epistemic probability—given what is known minus the assumption of intelligent design, it does not seem likely that the universe would exhibit such precise and ubiquitous causal order. Further scrutiny falls on the argument's conclusion, which posits only a personal cause, not a being with every perfection. Swinburne claims that considerations of simplicity lead to a positing of only one person—a person who has infinite knowledge and power, because any finite amount would require further explanation as to why the person has exactly this degree of knowledge or power.

Finally, the cumulative case argument draws upon further features of the universe that similarly confirm theism and disconfirm naturalism. Such features include the existence of a material universe, consciousness and

moral awareness, and evidence of providence, miracles, and religious experiences. (Note that in chapters ten and eleven, Swinburne argues that evil and suffering do not disconfirm theism. The claim that they do, he writes, "stems from a failure to appreciate the deepest needs of men … and the strength of the logical constraints on the kinds of world which God can make" (1979, p. 224).

See also Naturalism; Philosophy of Religion, History of; Physicotheology; Popular Arguments for the Existence of God; Religious Experience, Argument for the Existence of God; Theism, Arguments For and Against.

Bibliography

Collins, Robin. "Design and the Many-Worlds Hypothesis" In *Philosophy of Religion: A Reader and Guide*, edited by William Lane Craig. New Brunswick, NJ: Rutgers University Press, 2002.

Smart, John Jamieson Carswell, and John J. Haldane. *Atheism and Theism*. Oxford: Blackwell, 1996. Smart offers a critique of Haldane's theistic arguments that would apply equally to many of Swinburne's and Collins's arguments.

Swinburne, Richard. *The Existence of God*. Oxford: Clarendon Press, 1979.

Laura Garcia (2005)

TELEOLOGICAL ETHICS

Theories about what is right and wrong are standardly divided into two kinds: those that are teleological and those that are not. Teleological theories are ones that first identify what is good in states of affairs and then characterize right acts entirely in terms of that good. The paradigm case of a teleological theory is therefore an impartial consequentialist theory, such as hedonistic utilitarianism; defended by John Stuart Mill (1969) and Henry Sidgwick (1907), it says the right act is always the one whose consequences contain the greatest total pleasure possible. But the category of teleological ethics is normally thought to be broader than that of consequentialism, so there can be teleological theories that are not consequentialist. This can be so, however, in several different ways.

Hedonistic utilitarianism has three principal features: First, it identifies good states of affairs independently of claims about the right, so even pleasure in a wrong act, such as a sadist's pleasure in torturing, is intrinsically good; and these goods are always consequences in the ordinary sense of acts that produce them, that is, separate states that follow after the acts. Second, in evaluating consequences, utilitarianism weighs all people's pleasures impartially, so for any person, a stranger's pleasure counts just as much as his child's or even his own. Finally, utilitarianism characterizes right acts in terms only of promoting the good and, more specifically, of maximizing it, so the right act is always the one that produces the most good possible.

Although teleological theories must identify the good independently of the right, they can recognize many goods other than pleasure. Some possible goods, such as knowledge and artistic creativity, are, like pleasure, states of individual persons. Others involve patterns of distribution across persons, such as that they enjoy equal pleasures or, on a different view, pleasures proportioned to their merit. Yet others, such as the existence of beauty or of complex ecosystems, are independent of persons. (Goods of all three types are affirmed in the *ideal consequentialisms* of G. E. Moore (1903) and Hastings Rashdall (1907). These initial goods are all, like pleasure, consequences in the ordinary sense of acts that produce them, but other possible goods are not. Imagine that a theory values difficult activities because they are difficult. Then engaging in a difficult activity, such as playing chess, will promote value not just by producing it as an external consequence but also by instantiating it, or by having difficulty as an intrinsic feature. The same holds if a theory values action from a virtuous motive, such as a benevolent desire for another's pleasure. Then a benevolent act will contribute to value in part through an intrinsic feature—its being benevolent. This is a first way in which a theory can be teleological but not consequentialist: If consequentialism can value only the external consequences of acts, as some definitions assume, then a theory fits the broader but not the narrower concept if it values some intrinsic properties of acts. It can still evaluate acts by the total state of the world that will obtain if they are performed, but some relevant features of that state are now internal to them.

A teleological theory can also abandon the second feature of utilitarianism—its impartiality about the good. Thus, a teleological theory can be egoistic, telling individual agents to promote only their own pleasure, knowledge, or other goods, or, conversely, can say that they should promote only others' good and not their own. It can also embrace what C. D. Broad (1971) called "self-referential altruism," which says that while people should give some weight to everyone's good, they should care more about that of those who are close to them, such as their family and friends. These theories can still identify the good independently of the right and say right acts

maximize the good, but if it is essential to consequentialism to be impartial, as again some assume, they are teleological but not consequentialist.

These first two possibilities come together in a group of theories often categorized as teleological but not consequentialist—the eudaimonist theories of Aristotle and other ancient philosophers. They derive all moral requirements from a final end or good they call a person's *eudaimonia*, translated either as happiness or as flourishing. They are therefore formally egoistic since each person's final end is just that person's own eudaimonia. But they hold that a principal component of eudaimonia is moral virtue, which will express itself in virtuous acts such as helping others from benevolent motives. Eudaimonist theories can in principle yield the same substantive duties as utilitarianism, telling each person to maximize pleasure impartially. But their explanatory claims do not use the causal relation central to utilitarianism, saying, instead, that acts of helping others are required because they can instantiate moral virtue, which in turn instantiates part of eudaimonia.

Finally, a theory can abandon the third, maximizing feature of utilitarianism. This feature is extremely demanding since it implies that any time we do not do everything we can to benefit other people, which includes any time we relax or amuse ourselves, we act wrongly. One possibility, proposed by Michael Slote (1985), is to replace the maximizing principle with a *satisficing* one that says an act is right so long as its consequences are good enough, either in absolute terms or because they make some reasonable proportion of the greatest improvement the agent can make in the circumstances. Many writers see satisficing as consistent with consequentialism, but if it is essential to the latter to be maximizing, as some definitions imply, a satisficing principle again generates a nonconsequentialist teleology. A related possibility, proposed by Samuel Scheffler (1982), is to retain a maximizing principle but simultaneously grant agents an option to give somewhat more weight to their own good. Then, if they prefer a smaller benefit for themselves to a somewhat greater one for other people, they do not act wrongly, though if they preferred the greater good, they also would not act wrongly. The resulting view is probably not consequentialist since it does not contain only principles about promoting the good; but it arguably is teleological since its principles all do concern the good in some way.

More radical departures from maximizing may be possible. Teleological theories are commonly contrasted with deontological ones, which say an act can be wrong even if it has the best consequences. Thus, a deontological theory can say it is wrong to kill an innocent person even if that will prevent five other innocent people from being killed because doing so violates a moral constraint against killing; it can likewise contain constraints against lying, promise- breaking, and so on. A deontological theory is clearly nonconsequentialist, and it is also nonteleological if its constraints are independent of the good, say, if it contains independent, underived prohibitions of killing and lying. But some deontologists, who call their view Thomist, do connect constraints to the good. They start by identifying certain states of affairs as intrinsically good, say, pleasure, knowledge, and freedom. But they then claim that alongside a duty to promote these goods is a separate and stronger duty to respect them, which means not choosing against or intentionally destroying them. This second duty grounds constraints against killing, which destroys good human life; lying, which aims at the opposite of knowledge; and more.

But Thomists such as John Finnis (1980) call their view teleological since it is centered on goods that can and should be promoted. The same could not be said of Kantian deontologies, which ground constraints in respect for a value that is located in persons rather than in states of affairs and is not to be promoted since there is no duty to increase the number of valuable persons. But Thomist deontology shares enough assumptions with paradigmatically teleological theories that it arguably, if not uncontroversially, belongs in the category. (If so, deontological ethics contrasts with consequentialism but not necessarily with teleology.)

Teleological moral theories relate all moral duties to the goodness of states of affairs. They will therefore be rejected by those who think claims about intrinsic goodness are unintelligible or who hold, with Kant (1998), that the fundamental value is that of persons. These are minority views, however. Most philosophers accept as underived such claims as that pain is evil and knowledge good, so there is at least some moral duty to prevent the one and promote the other. The key issue about teleological ethics, then, is whether *all* duties can be related to the good. In addressing this issue, the many forms teleological ethics can take should be remembered. It can value not just pleasure but also, say, equal distribution and virtuous action; it can allow or even require agents to give more weight to some people's good; and it need not demand maximization of the good. But the question remains whether teleological ethics can recognize moral constraints, which can make it wrong to do what has the best effects. Strict consequentialists reject such con-

straints or claim that belief in them is justified only insofar as it has good consequences. But those who find constraints independently compelling will ask whether teleological ethics can accommodate constraints, as Thomist theories try to do, and, if so, whether it gives them the best explanation. If the answer to both questions is yes, then the teleological approach to ethics can capture a wide range of moral phenomena. If not, it will be unacceptable to those who think it sometimes wrong to do what will promote the most good.

See also Aristotle; Consequentialism; Deontological Ethics; Ethics, History of; Kant, Immanuel; Utilitarianism.

Bibliography

Aristotle, *The Nicomachean Ethics.* Translated by David Ross. Oxford: Oxford University Press, 1980.

Broad, C. D. "Self and Others." In *Broad's Critical Essays in Moral Philosophy*, edited by David R. Cheney. London: Allen & Unwin, 1971, pp. 262–282.

Finnis, John. *Natural Law and Natural Rights*. Oxford: Clarendon Press, 1980.

Frankena, William K. *Ethics*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1973.

Kant, Immanuel. *Foundations of the Metaphysics of Morals* Translated by Mary Gregor. Cambridge, U.K.: Cambridge University Press, 1998.

Mill, John Stuart. *Utilitarianism* (1861). In *Essays on Ethics, Religion, and Society*, edited by J. M. Robson. Vol. 10 of *Collected Works of John Stuart Mill.* Toronto: University of Toronto Press, 1969, pp. 203–259.

Moore, G. E. *Principia Ethica*. Cambridge, U.K.: Cambridge University Press, 1903.

Rashdall, Hastings. *The Theory of Good and Evil, A Treatise on Moral Philosophy*. 2 vols. London: Oxford University Press, 1907.

Scheffler, Samuel. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Clarendon Press, 1982.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1907.

Slote, Michael. *Common-Sense Morality and Consequentialism.* London: Routledge & Kegan Paul, 1985.

Thomas Hurka (2005)

TELEOLOGY

The term *teleology* locates a series of connected philosophical questions. If we grant that there is such a thing as purposive or goal-directed activity (as we must, since, for example, a political campaign aimed at victory represents a clear, uncontroversial case), we may ask the following questions: (1) By what criteria do we identify purposive activity? (2) What is the nature of the systems that exhibit purposive activity? (3) Does the nature of purposive activity require us to employ special concepts or special patterns of description and explanation that are not needed in an account of nonpurposive activity? And if we grant that there are objects and processes which perform functions (again, as we must, since no one would deny, for instance, that the human kidney performs the function of excretion), we may ask: (4) By what criteria do we identify functions? (5) What is the nature of the systems that exhibit functional activity? (6) Does the description of functions require special concepts or special patterns of analysis?

These six questions have been formulated with the help of a distinction between purposive and functional activity. Although the distinction is not always drawn in discussions of teleology, it is desirable for a number of reasons. It seems, at least prima facie, that the criteria of functional activity are quite distinct from the criteria of purposive activity: urine excretion, for example, seems to be a function by virtue of its role in the economy of a living organism, whereas activity seems to be purposive in virtue of the manner in which it is controlled. Thus, it seems at least logically possible that a purposive activity could perform no function, and that a function could be performed without purposive activity. Moreover, in view of this fundamental conceptual difference between purpose and function, we should expect the analysis of purposive and functional activity to show differences in logical pattern. On the other hand, it also seems clear that there are close connections between function and purpose; thus the final question: (7) What is the relation between ascriptions of function and ascriptions of purpose?

PURPOSE ACTIVITY

CRITERIA. A number of writers have proposed definitions of "goal-directed" or "purposive" action that leave open the question whether the action is intentional or in any way involves consciousness. R. B. Braithwaite suggests, as a behavioral criterion of goal-directed activity that either may or may not be goal-intended, "persistence toward the goal under varying conditions." This is a condensed version of very similar criteria offered by R. B. Perry, E. S. Russell, and A. Hofstadter. All presuppose that a goal may be identified and that both persistence and sensitivity to varying conditions may be located by reference to the goal. E. C. Tolman adds the requirement that purposive activity show "docility," that is, some improve-

ment in reaching the goal in the course of successive trials. But docility, however important it may be in the total picture of biological purposiveness, is surely not part of the criterion of purposiveness. Any abilities that are in fact learned could, in logical principle, be innate.

This criterion, in Braithwaite's form, is of course susceptible of considerable refinement; Braithwaite himself (in *Scientific Explanation*), for example, proposes a way of identifying variations in conditions as relevant variations for applying the criterion. Further possible refinements will be discussed in the next section.

The apparent circularity in the criterion—defining "goal-directed" in terms of a "goal"—is not serious. The location of persistence, sensitivity, and a goal may proceed together by a method of successive approximations. For example, a pattern of animal behavior may appear persistent and lead to a tentative identification of a goal, and the identification may be checked by looking for sensitivity to conditions or further evidence of persistence. A hypothesis about any one of the three—goal, persistence, sensitivity—can be confirmed by investigating either of the other two.

It seems clear that there are behavioral criteria for identifying purposive action, not only of human beings but also of other animals and of artifacts such as self-guided missiles. A pilot who watches a rocket approach in spite of his evasive maneuvers would rightly have no doubts about either the goal-directedness of the rocket's movements or the identity of its goal. No doubt the actual criteria of purposiveness that have been proposed suffer various shortcomings. In particular, they seem to lay down a necessary but not a sufficient condition. However, most philosophers would regard the program of seeking behavioral criteria as sound.

NATURE OF SYSTEMS SHOWING PURPOSIVE ACTIVITY. Is it possible for the philosopher, as distinct from the biologist, psychologist, or communications engineer, to say anything illuminating about the nature of the systems—men, mice, and missiles—that engage in purposive activity? He can at least examine more closely the behavioral criteria of purposiveness, in order to see whether there might be covert reference to the nature of the system in the criterias' actual application. A critic of the behavioral criteria might remark that a river is persistent in reaching the sea and is sensitive to the conditions necessary for reaching the sea—it detours all obstacles—but we would not call the flowing of a river purposive, nor would we call the sea or reaching the sea its goal. In short, the critic might say, a river is not the sort of thing to which we ever ascribe purposiveness.

Directive correlation. A number of philosophers, including Braithwaite, Ernest Nagel, George Sommerhoff, and Morton Beckner, have proposed ways of avoiding the difficulty about rivers and the like. Although there are differences in their accounts, they all adopt the strategy of regarding an activity as purposive only when its goal-seeking character is the outcome of relatively independent but dovetailing processes. Sommerhoff, for example, defines "purposive behavior" with the help of a concept he terms "directive correlation." Two variables, such as the position of a moving target and the direction in which an automatic target-tracking mechanism points, are said to be directively correlated with respect to a goal state (in this case, the state in which the mechanism points at the target) whenever: (1) The two variables are independent in the sense that any value of one is compatible with any value of the other; (2) The actual value of both, at a given time, is at least in part causally determined by the prior value of a "coenetic" (steering) variable (in the example, the coenetic variable is the same as one of the directively correlated variables, namely, the position of the moving target); and (3) the causal determination is such that the actual values of the directively correlated variables are sufficient for the realization of the goal state. Sommerhoff then defines "purposive behavior" as directively correlated behavior in which the coenetic variable is identical with one of the directively correlated variables.

Stipulations (2) and (3) make the notion of two processes dovetailing so as to achieve a goal as precise as the notion of causal determination; and stipulation (1) specifies that the processes must be independent. The requirement of independence rules out such cases as the river, for the direction in which a river flows is not independent of the lay of the land.

Sommerhoff's analysis is not without difficulties (see Nagel and Beckner), but it is undoubtedly correct in general approach. A system *S* that could exhibit directive correlation would satisfy a number of prior conceptions about purposive behavior; for instance, that *S* would employ information about its environment, particularly about an aspect of the environment associated with the goal, and that the behavior of *S* would be dependent upon a specialized physical hookup, such as some sort of circuitry.

It is now possible to suggest a schema for constructing a criterion of purposive activity that includes both a necessary and a sufficient condition and that incorporates some reference both to the empirical character of

the activity and to the nature of the system that engages in it. Activity is purposive if and only if it exhibits sensitivity and persistence toward a goal as a result of directive correlation.

NEED FOR SPECIAL CONCEPTS OR PATTERNS OF DESCRIPTION AND EXPLANATION. Purposive activity, in the analyses of Braithwaite and Sommerhoff described above, does not involve a special kind of causality but only a special organization of ordinary causal processes. If these analyses are correct, both living organisms and artificial machines are capable of purposive activity. If, therefore, special concepts or patterns of description and explanation are not needed in the case of purposive machines, it would appear that they are equally unnecessary in the case of organisms. Many philosophers have drawn this conclusion, and it must be admitted that accounts like Braithwaite's and Sommerhoff's constitute powerful arguments in its support.

There is room for some doubt, however. Even if we grant that purposive activity can be defined in terms that are equally applicable to organic and inorganic systems, it does not follow that all purposive activity can be explained on the model of inanimate activity. The most serious doubt concerns those purposive activities that may be described as the acts of agents, such as acts deliberately undertaken for the sake of a consciously envisaged end. Suppose, for example, that some or all of these acts of agents are in principle unpredictable—a view accepted by some philosophers. Then, if they can be explained at all, their explanation is essentially post hoc. The pattern of such explanation is not yet properly understood; nevertheless, there is at least some doubt that it can dispense with the conception of following a rule. But these considerations raise questions that cannot be pursued here.

FUNCTIONS

CRITERIA. When we assert truly—for example, that a function of the kidney is the excretion of urine—precisely what relations must hold between the kidney and excretion? It has been proposed, for example by Nagel, that such teleological terms as *purpose* and *function* can be eliminated in the following way: An expression such as "A function of the kidney is the excretion of urine" is translated into the nonteleological expression "The kidney is a necessary (or necessary and sufficient) condition of urine excretion." In general we may interpret Nagel as proposing a translation schema—For "*F* is the function of *A*," write "*A* is a necessary (or necessary and sufficient) condition of *F*"—that dispenses with teleological language and that also provides part of a criterion (a necessary condition) for identifying functions.

At best, however, Nagel's schema must be modified, for the possession of kidneys is neither a necessary nor a sufficient condition of urine excretion. It is obviously not sufficient; but it is also not necessary, since urine can also be excreted by various artificial devices. (If it is objected that these devices are themselves a sort of kidney, then the statement that a kidney is necessary for excretion reduces to a tautology.) Moreover, the translation schema is much less plausible when applied to organic functions that are ordinarily accomplished in distinct ways. Temperature regulation, for example, is a function of man's body hair; but hair is not necessary for heat regulation, since the function may be performed by other physical and physiological mechanisms. When we ascribe a function to the kidney or to body hair, we seem to be saying no more than that these structures contribute to certain processes; we leave open the question whether they are necessary or sufficient for the processes. The relation "contributing to" may be defined without employing teleological language. Let *F* be a process, some or all of which takes place in system *S*; and let *A* be a part of, or a process in, *S*. Finally, let the terms "*S*-like," "*F*-like," and "*A*-like" refer, respectively, to all those entities that answer to the definition of the terms employed in specifying *S, F,* and *A*. (In the example "A function of the kidney in vertebrates is the excretion of urine," all vertebrates are *S*-like, all cases of urine excretion are *F*-like, and all kidneys are *A*-like.) Then "*A* of *S* contributes to *F*" if and only if there exist *S*-like systems and states or environments of these *S*-like systems in which *F*-like processes occur and the possession of *A*-like parts or processes is necessary for the occurrence of *F*-like processes.

On this definition, we may say that in general a man's kidney contributes to the excretion of urine and that body hair contributes to heat regulation. And if we adopt the translation schema "For '*F* is the function of *A* in *S*,' write '*A* contributes to *F* in *S*,'" we may say, even in the case of a man whose bad kidneys have been bypassed to an artificial kidney, that the function of his flesh-and-blood kidneys is still the excretion of urine; they merely fail to perform it.

NATURE OF SYSTEMS SHOWING FUNCTIONAL ACTIVITY. Nagel's translation schema and the above modification of it provide a way of translating a teleological statement T_1 into a statement T_2 that does not employ explicitly teleological terms. Therefore, the satisfaction of T_2 by a given *A, F,* and *S* is a necessary condition of *F*'s

being a function of *A*. It is, however, not a sufficient condition; we may not in general translate T_2 into T_1. We would not say, for example, that the function of the ground is to hold up the rocks even though, in our technical sense, the ground contributes to the holding up of rocks. It would seem that out of the whole set of "contributing" cases, only a very restricted subset could be regarded as functions.

How may this subset be specified? We ordinarily attribute functions to two sorts of systems, artifacts and living things. We may consider first a simple artifact such as a cooking pan. We ascribe a function to the whole pan: cooking. Moreover, we also ascribe functions to parts and properties of the pan insofar as they contribute to its usefulness in cooking. For example, it is natural to think of the handle as providing a grip, of the rivets as fastening on the handle, and so on. In short, whenever we are prepared to acknowledge a single function *F,* we are also prepared to acknowledge a hierarchy of functions, with *F* at the top and the functions at each lower level contributing to all those above them.

The assignment of functions to living organisms proceeds on the same principle. There are two organic processes that are regarded as fundamental, the maintenance of life and reproduction. Alternatively, these two processes may be thought of as contributing to a single process, the maintenance of a species, which stands at the top of all functional hierarchies. The fundamental processes thus play a defining role in the identification of functions. The following schema lays down a necessary and sufficient condition of functional activity: F_1 is a function of *A* in *S* if and only if *A* contributes to F_1 in *S*; and F_1 is identical with or contributes to F_2 in *S,* where F_2 is either a purpose for which the artifact *S* is designed or the process of maintenance of the species of which *S* is a member.

The concept of an artifact may be interpreted quite broadly in order to include not only things like cooking pans but also all cultural products, such as works of art, language, and legal institutions. It makes sense, for example, on the above analysis and on this interpretation of *artifact,* to ask "What is the function of Ophelia in *Hamlet?*" and "What is the function of verb inflections in Japanese?" The justification for regarding maintenance of the species as a fundamental function, serving a logical role in functional analysis, is examined below.

NEED FOR SPECIAL CONCEPTS OR PATTERNS OF ANALYSIS. The definition of functional activity offered above provides a way of interpreting ascriptions of functions without using explicitly teleological expressions. However, there is a sense in which many of the concepts that are employed in the ascription of functions are implicitly teleological. Consider, for example, the concept of an "escape reaction." It is applied to a great variety of animal movements, such as flying up, forming dense schools, withdrawing into burrows, jumping into water, and gathering under the mother. These diverse reactions probably have no relevant feature in common other than a functional one; they all, in the technical sense, contribute to the avoidance of death by predation. Such functional concepts are common in the theory of animal behavior, in all branches of natural history, in physiology, and indeed in everyday language. The terms that we most commonly use, for example, in describing machines are defined functionally.

The view that teleological language can be eliminated from the language of science may be true; again, the most difficult cases concern human agency. But the program of eliminating teleological expressions even from biological theory must involve more than the elimination of such terms as *function, purpose, goal,* and *in order to.* If there is any point in eliminating these terms, there is just as much point in eliminating all concepts that are defined functionally, for "The function of this movement is to escape from a predator" is equivalent in asserted content to "This movement is an escape reaction." It is obviously true that the movement in question can be described, without employing the term *escape reaction,* as a movement that contributes to the avoidance of a predator. But if we eliminate the term *escape reaction,* we have excised from the language the term that applies not only to this movement but to all the diverse movements, in a variety of taxonomic groups, that serve this function.

The ascription of functions, therefore, does not require either an explicit or an implicit teleological vocabulary. It should be recognized, however, that the elimination of implicitly teleological expressions (concepts that are defined functionally) would result in a language for biological theory that would bear very little resemblance to the existing language.

Moreover, the difference would not be superficial; the rejection of functional concepts would amount to the rejection of a powerful and fruitful conceptual scheme. Our picture of living organisms as organized functional hierarchies is an essential part of the theory of natural selection; it is the foundation of physiology and morphology; and it is the basis of the medical view of disease as derangement of function. It is the fruitfulness of this conceptual scheme, embodied in a network of connected

functional concepts, that constitutes the justification for assigning to maintenance of the species its central logical role in the ascription of functions.

RELATION BETWEEN ASCRIPTIONS OF FUNCTION AND OF PURPOSE

We have drawn a sharp distinction between functional activities, which contribute to a "fundamental" process, and purposive activities, which are persistent, flexible patterns of directively correlated behavior. It is clear, however, that function and purpose are closely connected—so closely, indeed, that many writers have failed to see the distinction. These connections may be described as follows:

(*a*) Whenever we construct an artifact as an aid to our own purposive activities, we are willing to ascribe functions to the artifact and to its parts and properties.

(*b*) Many but by no means all organic functions are served by purposive activities. For example, temperature regulation in the mammals involves directive correlation, whereas the excretion of urine does not.

(*c*) Conversely, every organic mechanism that provides an organism with the means of purposive activity serves the function of maintenance of the species. This is an empirical fact. It does not mean, however, that each case of purposive activity, when it occurs, performs a function. A purposive activity that is ordinarily adaptive (functional under normal circumstances) can lead to disaster when the circumstances are abnormal. For example, the homing of a male moth on a female, directed by the attractant secreted by the female, is ordinarily both purposive and functional. But it can lead the moth to his death when the attractant is placed on a surface covered with an insecticide.

See also Braithwaite, Richard Bevan; Functionalism; Functionalism in Sociology; Nagel, Ernest; Organismic Biology; Perry, Ralph Barton; Speculative Systems of History; Teleological Argument for the Existence of God; Teleological Ethics.

Bibliography

Beckner, Morton. *The Biological Way of Thought.* New York: Columbia University Press, 1959.

Braithwaite, R. B. *Scientific Explanation.* Cambridge, U.K.: Cambridge University Press, 1953.

Braithwaite, R. B. "Teleological Explanations." *PAS,* n.s., 47 (1947): i–xx.

Hofstadter, A. "Objective Teleology." *Journal of Philosophy* 38 (2) (January 1941): 29–39.

Nagel, Ernest. "Teleological Explanation and Teleological Systems." In *Readings in the Philosophy of Science,* edited by H. Feigl and M. Brodbeck, 537–558. New York: Appleton-Century-Crofts, 1953. A revised and enlarged version of this article appears in Nagel's *The Structure of Science,* 401–428. New York: Harcourt Brace, 1961.

Perry, R. B. "A Behavioristic View of Purpose." *Journal of Philosophy* 18 (4) (February 1921): 85–105.

Russell, E. S. *The Directiveness of Organic Activities.* Cambridge, U.K.: Cambridge University Press, 1945.

Sommerhoff, George. *Analytical Biology.* London, 1950.

Tolman, E. C. *Purposive Behavior in Animals and Men.* New York: Century, 1932.

Morton Beckner (1967)

TELEOLOGY [ADDENDUM]

Teleological explanations are said to be forward looking. We ask why Lauren is walking and are told her purpose, which is to buy ice cream when she gets to the shop. Or we ask why vertebrates have kidneys and are told their function, which is filtering blood. In both cases, the end explains the means; something at a time is explained by something else at a later time. This inverts the usual order of causal explanations: If Johnny's throwing the ball explains the window breaking, his throwing preceded the breaking.

PURPOSIVE EXPLANATIONS

How does Lauren's purpose explain her walking? Many philosophers would now say that the relation between her purpose and her walking is a special instance of ordinary physical causation. On a standard version of physicalism, an agent's purpose consists of beliefs and desires, which involve brain states that represent what is believed and desired. If Lauren is walking to the shop to buy ice cream, she has both a desire to buy ice cream and a belief that walking to the shop will let her do so. It is not her buying ice cream but her intention to do so that causes her walking, and since her intention precedes her walking, the usual explanatory order is preserved.

Some physicalists question the causal power and explanatory relevance of beliefs and desires. For example, Jaegwon Kim (1998) argues that, given that mental properties cannot be strictly identified with basic physical properties (a thesis of functionalism), they are causally redundant, since basic physical properties suffice to cause behavior. And Jerry Fodor (1991) argues that, given that the contents of beliefs and desires depend on the relations of an agent to his or her environment (the thesis of content externalism), contents do not explain behavior, since

an agent's behavior is caused by his or her intrinsic properties. Similar doubts can be raised with regard to the causal power and explanatory relevance of functions. However, by no means is everyone persuaded by these arguments, and their conclusions are anyway consistent with (what Sydney Shoemaker calls) the core realizers of beliefs and desires being the causes of behavior.

FUNCTIONAL EXPLANATIONS

When functions are attributed to artifacts and components of organic systems, we seem to use a teleological notion of what something is for. "The switch has the function of turning on the light" seems equivalent to "The switch is for turning on the light" (that is why it is there). "Pineal glands have the function of secreting melatonin" seems equivalent to "Pineal glands are for secreting melatonin" (that is why they are there). Not all locutions involving the word *function* have this teleological flavor. "X *performs* the function of Z-ing" does not entail "X *has* the function of Z-ing" or "X is there *in order to* Z." So only function ascriptions of the latter kind are relevant here.

Artifact functions depend on the purposes of the people who design, make, or use the artifacts: The switch has the function of turning on the light because someone put it there (or later adapted it) for that purpose. Organic function ascriptions in biology seemed more puzzling once the bearers of the functions were no longer seen as God's artifacts.

However, many philosophers of biology now believe that natural selection can replace God in function ascriptions. A popular view, developed and defended by, among others, Larry Wright (1976), Ruth Millikan (1989), Karen Neander (1991a, 1991b), and Peter Schwartz (2002), is that the biological function of a trait is what that type of trait was selected for. According to this etiological theory of function, the pineal gland has the function of secreting melatonin because that is what pineal glands did that caused them to be preserved and/or proliferated in the population. This gives functional explanations of the teleological variety a parallel form to purposive explanations: They both explicitly refer to an effect of the item being explained, but in doing so they implicitly refer to a past event to explain it (intentional selection for the effect, or natural selection for the effect). Numerous objections to the etiological theory have been made, but while it has not gone entirely unscathed, in the view of most philosophers of biology it remains the theory to beat (although see, e.g., Christopher Boorse [2002], who strongly disagrees).

As with purposes, an important issue is the explanatory role of functions. According to Wright (1976), a trait's function explains why it is there. Robert Cummins (1975) argues against this, that functions explain how systems operate. An overall capacity of a complex system is explained by a functional analysis, which describes the contributing capacities of the parts of the system, and the contributing capacities of each of their parts, in turn. According to Cummins, a function of a component part is its contribution to a capacity under analysis.

A problem with Cummin's account is that it does not account for the normativity of function ascriptions. Function ascriptions are normative (although not prescriptive) in the sense that they permit the possibility of malfunction: For example, my pineal gland could have the function to secrete melatonin and at the same time it could lack the capacity to secrete melatonin because it is malfunctioning. His account also leaves a lot to be determined by the interests of the researcher. Which overall capacity is to be analyzed and in which environment its exercise is to be analyzed is settled by the interests of the researcher. Thus the account is not naturalistic (it makes use of intentional terms). It is also inaccurate. For example, those interested in explaining death by cancer can give a functional analysis of the kind that Cummins describes. But contributions to death by cancer are not normal (proper) functions by virtue of their role in producing death by cancer. These problems suggest that the analysis is at best incomplete as it stands.

While Cummins's (1975) analysis of functions is problematic, he is right about the importance of functional analysis. This has led some to suggest that biology employs two notions of function, with distinct explanatory roles: a teleological notion for teleological explanations and a notion of a contributing capacity for functional analysis. However, this cannot be the right way to understand their respective explanatory roles if the etiological analysis is the correct analysis of functional norms, since physiological biology, which provides functional analyses of living systems, makes important use of the distinction between normal and abnormal functioning in doing so. Neander (1991b) suggests that the teleological/etiological notion of a function permits an idealized functional analysis, the idea being that we describe the functional organization of a normal system (as opposed to the malfunctioning of an abnormal system) by describing the capacities for which each of its parts was selected.

Bibliography

Boorse, Christopher. "A Rebuttal on Functions." In *Functions: New Essays in the Philosophy of Psychology and Biology*, edited by André Ariew, Robert Cummins, and Mark Perlman. New York: Oxford University Press, 2002.

Cummins, Robert. "Functional Analysis." *Journal of Philosophy* 72 (1975): 741–765.

Fodor, Jerry. "A Modal Argument for Narrow Content." *Journal of Philosophy* 88 (1991): 5–26.

Kim, Jaegwon. *Mind in a Physical World*. Cambridge, MA: MIT Press, 1998.

Millikan, Ruth. "In Defense of Proper Functions." *Philosophy of Science* 56 (1989): 288–303.

Neander, Karen. "Functions as Selected Effects." *Philosophy of Science* 58 (1991a): 168–184.

Neander, Karen. "The Teleological Notion of Function." *Australasian Journal of Philosophy* 69 (1991b): 454–468.

Schwartz, Peter. "The Continuing Usefulness Account of Proper Function." In *Functions: New Essays in the Philosophy of Psychology and Biology*, edited by André Ariew, Robert Cummins, and Mark Perlman. New York: Oxford University Press, 2002.

Wright, Larry. *Teleological Explanation*. Berkeley: University of California Press, 1976.

Karen Neander (2005)

TELESIO, BERNARDINO

(1509–1588)

Bernardino Telesio, the Renaissance philosopher, was born at Cosenza, in Calabria, Italy. He studied philosophy, physics, and mathematics at the University of Padua, and received his doctorate in 1535. In Padua he became acquainted with the teaching of Aristotle and the two main Aristotelian schools, the Averroistic and the Alexandrist. Following the trend of the time, he devoted himself especially to the study of nature; but far from accepting the Aristotelian doctrine, he reacted vigorously against it. Telesio pursued his literary activity mostly at Naples, where he was a guest of the Carafa family, and at Cosenza. He enjoyed the friendship of several popes, and Gregory XIII invited him to Rome to expound his doctrine. He never engaged in any formal teaching, for he preferred to discuss his ideas in private conversations with friends.

Telesio is the author of the nine-book *De Rerum Natura luxta Propria Principia* (On the Nature of Things According to Their Principles; 1586) and of several philosophical opuscules. He proposed to interpret nature by following the testimony of the senses, rather than to attempt an explanation through the "abstract and preconceived ideas" of the Aristotelians. Nature must be studied in itself and in its own principles, which are matter and the two active forces of heat and cold. Matter is the passive, inert substratum of all physical change and is substantially the same everywhere. Unlike Aristotelian prime matter, which is pure potency, it is concrete and actual, and hence it can be directly perceived by the senses. Heat and cold are the two opposing forces responsible for all natural events; the first is represented by sky and the second by earth. Heat is also the source of life in plants and animals, as well as the cause of biological operations and some of the lower psychological functions in man. The whole of nature is animated and endowed with sensation in varying degrees (panpsychism). In addition to the vital principle there is present in man and animals "spirit," a very subtle material substance that emanates from the warm element and is generated with the body. Spirit is properly located in the brain and has the function of anticipating and receiving sense impressions. It has both an appetitive power and an intellective power of its own that correspond to the sensitive appetite and the cogitative power (*vis cogitativa*) of the Aristotelians.

Besides body and spirit, man has a *mens,* or *anima superaddita,* which is created by God and informs both body and spirit. This is roughly equivalent to the spiritual soul of Platonic -Augustinian tradition, whose operations transcend those of spirit and reach up to the divine. Apart from the natural drive or instinct of self-preservation, which Telesio attributed to all beings—including inorganic matter—man can also strive after union with God and contemplate the divine. This inner tendency of the *mens,* along with the need for proper sanctions in a future life in order to correct injustices, was one of the arguments used by Telesio to prove the immortality of the soul, which is known by revelation but can also be demonstrated by reason.

For Telesio self-preservation was man's supreme good. Just as in man there is a twofold intellect, one pertaining to the spirit and the other to the soul, so also there is in him a twofold appetitive power. The sensitive appetite tends toward temporal goods and its own preservation in this life; rational appetite or will tends toward immortal goods and its own preservation in a future, eternal life. Virtues are powers or faculties that enable man to achieve self-preservation; they are not merely habits, as Aristotle taught. There are virtues of the spirit and virtues of the soul. Among the virtues, sublimity and wisdom occupy a high place. Sublimity is not merely a particular virtue but virtue as a whole. It stands at the summit of all virtues and somehow includes all of them, for it directs all man's operations toward his supreme good. Wisdom helps man to attain to the knowledge of

God as creator of the universe and can reach out to the knowledge of the divine substance itself.

Although Telesio did not specifically treat the problem of God's existence (it was beyond the scope of his study), he touched incidentally upon Aristotle's argument from motion and criticized it on the ground that movement is an intrinsic property of heat, the first active principle of material beings. Accordingly, there is no need for an extrinsic agent to set the bodies in motion. Besides, an immovable mover that sets the heavens in motion, as conceived by Aristotle, is a contradiction. The existence of God is better proved from the wonderful order of the universe, which can only be the work of a divine mind.

As evidenced by this summary exposition of Telesio's thought, it would be wrong to call him a naturalistic philosopher, if the term *naturalism* is taken to mean a purely materialistic approach to reality. In his *De Rerum Natura* Telesio claimed to investigate the nature of things according to their intrinsic principles, and only incidentally spoke of their extrinsic causes. He gave us a philosophy of nature along the general lines of Aristotle's *Physics,* although from a different point of view and following a more scientific method; he did not intend to present a philosophy of reality as a whole. Briefly, he discussed nature or the world as it is in its concrete reality, not as it came about or in reference to the end for which it was made. His approach to man, knowledge, and morality was on the same plane. One should not be surprised, then, to find in his *De Rerum Natura* no special treatment of God, the spiritual soul, man's ultimate end, and other doctrines commonly held by Christian philosophers. His pertinent statements were nevertheless more than sufficient to show the personal convictions of their author. Thus, in his dedicatory letter to Ferdinand Carafa, duke of Nocera, he wrote: "Our doctrine, far from contradicting the senses and Holy Scripture … so agrees with them that it seems to stem directly from these two sources."

Telesio was called "the first of the moderns" by Francis Bacon, who claimed that Telesio was the first to raise the banner against Aristotle. This same phrase has been used in connection with Telesio by some modern historians of philosophy to indicate his revolt against the traditional teaching of the Catholic Church. The truth is that Telesio was neither a mere critic of Aristotle nor an antagonist of the church, to which he always professed loyalty. His modernity consists, rather, in the emphasis he placed on sense experience in the study of nature, thus paving the way for the scientific method of Galileo Galilei and his followers and opening a path in philosophy that was soon to be followed by Tommaso Campanella, Bacon himself, and Thomas Hobbes. It must be admitted that Telesio often discussed scientific problems with a philosophical method. The result was that his *De Rerum Natura,* a pioneering work of unquestionable value, was neither a scientific study nor a philosophical treatise, but a hybrid combination of science and philosophy not quite in agreement with the rigorous empirical method he professed to follow. This weakness in Telesio's system was pointed out by his contemporary Francesco Patrizi, the Neoplatonist.

See also Alexandrian School; Aristotelianism; Aristotle; Averroism; Bacon, Francis; Campanella, Tommaso; Neoplatonism; Patrizi, Francesco; Renaissance.

Bibliography

WORKS BY TELESIO

De Rerum Natura. Naples, 1586. New edition, edited by Vincenzo Spampanato. 3 vols. Vol. I, Modena: Formíggini, 1910; Vol. II, Genoa, 1913; Vol. III, Rome, 1923.

Varii de Naturalibus Rebus Libelli, edited by Antonio Persio. Venice, 1590. Various philosophical opuscules.

Solutiones Thylesii (and two other opuscules). In Francesco Fiorentino, *Bernardina Telesio.* 2 vols. Florence, 1872–1874. Works contained in an appendix.

WORKS ON TELESIO

Abbagnano, Nicola. *Bernardino Telesio.* Milan: Fratelli Bocca, 1941.

Di Napoli, Giovanni. "Fisica e metafisica in Bernardino Telesio." *Rassegna di scienze filosofiche* 6 (1953): 22–69.

Gentile, Giovanni. *Bernardino Telesio.* Bari: Laterza, 1911.

Soleri, Giacomo. *Telesio.* Brescia: Scuola, 1944.

Troilo, Erminio. *Bernardino Telesio.* 2nd ed. Modena, 1924.

Van Deusen, Neil. "The Place of Telesio in the History of Philosophy." *Philosophical Review* 44 (1935): 417–434.

Van Deusen, Neil. *Telesio, the First of the Moderns.* New York, 1932.

Bernardine M. Bonansea, O.F.M. (1967)

TEMPORAL OR TENSE LOGIC

See *Modal Logic*

TENNANT, FREDERICK ROBERT
(1866–1957)

Frederick Robert Tennant, the philosopher of religion and theologian, spent most of his life in Cambridge, England, and was educated at Cambridge University. He was a fellow of Trinity College and university lecturer in the philosophy of religion. His writings are in two main areas. In the strictly theological field he produced several influential studies of the concepts of sin and the fall of man, in which he diverged widely from the traditional Augustinian doctrines. In the philosophy of religion and the philosophy of science (in both of which his thought shows the influence of his Cambridge contemporary James Ward) Tennant's magnum opus is the two-volume *Philosophical Theology,* which develops, from foundations in the sciences, the thesis that there is "a theistic world-view commending itself as more reasonable than other interpretations or than the refusal to interpret, and congruent with the knowledge—i.e. the probability—which is the guide of life and science" (Vol. II, p. 245).

Tennant described his method as empirical rather than a priori. He meant (1) that his epistemology was based on a psychological examination of the cognitive capacities of the human mind, and (2) that his theistic argument was inductive, treating the existence of God as a hypothesis that goes beyond but builds upon the hypotheses of the special sciences.

Tennant argued in *Philosophy of the Sciences* that all knowledge, other than that in logic and mathematics, consists in probable interpretative judgments whose verification to the human mind is ultimately pragmatic. Thus, science and natural theology share a common method and status: "inductive science has its interpretative explanation-principles, … and its faith elements with which the faith of natural theology is, in essence, continuous" (p. 185). So Tennant can speak of theology as "the final link in a continuous chain of interpretative belief" (p. 184) and can say that "theistic belief is but a continuation, by extrapolation, or through points representing further observations, of the curve of 'knowledge' which natural science has constructed" (pp. 185–186). (For Tennant's conception of faith as the volitional element in the acquisition of all knowledge, scientific no less than religious, see the entry FAITH).

Tennant rejected religious experience—both the special experiences of the mystic and the less special religious experience of the ordinary believer—as a valid ground for belief in God, and he rested his entire case upon what he called the wider, or cosmic, teleology.

The version of the Argument to Design in Volume II of Tennant's *Philosophical Theology*—taking account as it does of David Hume's critique of the much simpler arguments of the eighteenth-century teleologists culminating in William Paley's *Natural Theology,* and taking account also of relevant developments in nineteenth-century and early twentieth-century science including the work of Charles Darwin—is probably the strongest presentation that has been written of this type of theistic reasoning. Serious discussions of the Teleological Argument should deal with it in the form provided by Tennant rather than in the relatively cruder versions of earlier centuries or of contemporary popular apologetics.

Tennant begins by making it clear, in accordance with his general theory of knowledge, that the argument is to provide "grounds for reasonable belief rather than rational and coercive demonstration." It employs a concept of probability that is not that of mathematics or logic but "the alogical probability which is the guide of life" and which, Tennant had already claimed in Volume I, is the ultimate basis of all scientific induction.

The argument itself does not rely (as did Paley's) on particular instances of apparent design in nature or on the arithmetical accumulation of these. Tennant allowed that each separate case of adaptation may be adequately explicable in purely naturalistic as well as in teleological terms. But he held that "the multitude of interwoven adaptations by which the world is constituted a theatre of life, intelligence, and morality, cannot reasonably be regarded as an outcome of mechanism, or of blind formative power, or of aught but purposive intelligence." (*Philosophical Theology,* Vol. II, p. 121).

His detailed argument contains the following strands:

(1) The basic instance of order is that the world stands in relation to human thought as something "more or less intelligible, in that it happens to be more or less a cosmos, when conceivably it might have been a self-subsistent and determinate 'chaos' in which similar events never occurred, none recurred, universals had no place, relations no fixity, things no nexus of determination, and 'real' categories no foothold" (p. 82).

(2) The internal and external adaptation of animal organisms can be accounted for in terms of an evolutionary process operating by means of natural selection; but how, other than by a cosmic pur-

pose, is that process itself to be accounted for? Here "The discovery of organic evolution has caused the ideologist to shift his ground from special design in the products to directivity in the process, and plan in the primary collocations" (p. 85).

(3) The emergence of organic life presupposes complex and specific preparatory processes at the inorganic level. Why has a universe of matter produced life and intelligence? If there were millions of universes, we might expect this to happen in a few of them. But there is only one universe. "Presumably the world is comparable with a single throw of dice. And common sense is not foolish in suspecting the dice to have been loaded" (p. 87).

(4) Nature produces in great abundance beauty that seems to exist only for the enjoyment of man. "Theistically regarded, Nature's beauty is of a piece with the world's intelligibility and with its being a theatre for moral life; and thus far the case for theism is strengthened by aesthetic considerations" (p. 93).

(5) Nature has produced man, with his ethical sense. If we judge the evolutionary process not by its roots in the primeval slime but by its fruits in human moral and spiritual experience, we note that "The whole process of Nature is capable of being regarded as instrumental to the development of intelligent and moral creatures" (p. 103).

(6) These five aspects of nature can individually be understood naturalistically. Nevertheless, taken as a whole they suggest a cosmic purpose that has used nature for the production of man. The more we learn of the complex conditions that had to come about before man could exist, "the less reasonable or credible becomes the alternative theory of cumulative groundless coincidence" (p. 106).

Having thus sought to establish theism as the most reasonable explanation of the world as a whole, Tennant discussed the problem of evil considered as challenging the theistic hypothesis, and he offered a theodicy that is typical of the thought of many British theologians on this subject in the twentieth century. This type of theodicy has an ancestry going back through Friedrich Schleiermacher to the early Hellenistic thinkers of the Christian church, especially Irenaeus, and it stands in contrast to the Augustinian and Latin tradition. For Tennant the possibility of the moral evil of sin was involved in the creation of free and responsible personal beings and was justified by the fact that only free persons can be the bearers of moral and spiritual values. Tennant saw the natural evil of pain in its many forms as a necessary concomitant of man's existence in a world that has its own stable structure and laws of operation; and it is justified by the fact that only in such an environment can the higher values of the human personality develop.

The same aspects of Tennant's thought constitute its strength from one philosophical point of view and its weakness from another point of view. He presented theology as an extension of science and theism as a hypothesis that is arguable in essentially the same sort of way as, for example, organic evolution. To some it will seem that by thus assimilating religious to scientific theorizing, Tennant made theology intellectually respectable; and this was his own view of the matter. To others, however, it will seem that Tennant was presenting religious belief in false colors. From their point of view, having excluded the true basis of religious faith in religious experience, Tennant attempted in vain to infer religious conclusions from nonreligious data, and by thus setting theistic belief upon a wrong and inadequate foundation, he has weakened rather than strengthened it.

See also Darwin, Charles Robert; Evil, The Problem of; Faith; Hume, David; Moral Arguments for the Existence of God; Paley, William; Schleiermacher, Friedrich Daniel Ernst; Teleological Argument for the Existence of God.

Bibliography

WORKS BY TENNANT

The Origin and Propagation of Sin. Cambridge, U.K.: Cambridge University Press, 1902.

The Sources of the Doctrines of the Fall and Original Sin. Cambridge, U.K.: Cambridge University Press, 1903.

The Concept of Sin. Cambridge, 1912.

Miracle and Its Philosophical Presuppositions. Cambridge, U.K.: Cambridge University Press, 1925.

Philosophical Theology, 2 vols. Cambridge, U.K.: Cambridge University Press, 1928 and 1930.

Philosophy of the Sciences. Cambridge, U.K.: Cambridge University Press, 1932.

The Nature of Belief. London: Centenary Press, 1943.

Philosophical Theology, Vol. 3. New York: Cambridge University Press, 1968.

WORKS ON TENNANT

Attfield, David. "The Morality of Sins." *Religious Studies* 20 (1984): 227–238.

Broad, C. D. "Frederick Robert Tennant, 1866–1957." *Proceedings of the British Academy* 44 (1960).

Buswell, J. Oliver. *The Philosophies of F. R. Tennant and John Dewey.* New York: Philosophical Library, 1950.

Mellor, D. H. "God and Probability." *Religious Studies* 5 (1968): 223–234.
Scudder, D. L. *Tennant's Philosophical Theology.* New Haven, CT: Yale University Press, 1940.
Smart, Ninian. "F. R. Tennant and the Problem of Evil." In his *Philosophers and Religious Truth.* London: S.C.M. Press, 1964.
Smart, Ninian. *Philosophers and Religious Truth.* New York: Collier, 1969.
Wynn, Mark. *God and Goodness: A Natural Theological Perspective.* London: Routledge, 1999.

John Hick (1967)
Bibliography updated by Christian B. Miller (2005)

TENSE

See *Appendix, Vol. 10*

TERESA OF ÁVILA, ST.

(1515–1582)

St. Teresa of Ávila, the Spanish mystic, was born of an aristocratic family in Ávila. In 1535 she entered a Carmelite convent there and four years later was prostrated by a long illness, probably of psychological origin. However, she had already felt the call to contemplation, and at about the age of forty, after a long struggle, she received a second "conversion," which turned her toward an intense practice of contemplation. Her order was relatively lax in its rules, and she felt impelled to begin a reform. In 1562 a reformed convent was established in Ávila under her direction. After five years, despite ill health and official opposition, she began energetically to spread the reform to other parts of Spain. She died in 1582, after a three-year illness. Her main works were her *Life* (1562–1565), *The Way of Perfection* (1565), and *The Interior Castle* (1577). The first is a full account of her inner experiences, and the last gives a more systematic description of the contemplative life.

Her account of the stages of mysticism, in the *Life*, uses the analogy of watering a garden by various means. Once the weeds have been uprooted, irrigation is needed. Those who bring the water from a well are compared to beginners in prayer and meditation. It is a laborious activity, involving the taming of the senses so that they are no longer distracting. The second stage of meditation is reached with the prayer of quiet. This is compared to irrigating the garden by a waterwheel. The third mode of watering is by a running brook: This corresponds to a state of contemplation in which effort is no longer needed, as if the work were done by the Lord. It is, according to St. Teresa, "a celestial frenzy," in which the faculties of sense perception no longer function. The soul no longer wishes to live in the world but solely in union with God. The intellect is worth nothing, for ordinary modes of understanding are considered irrelevant or nonsensical. In the fourth stage, which is compared to a shower falling on the garden, the soul is totally passive and receptive, all its faculties somehow united with God. The soul cannot properly understand what is occurring, but afterward it is certain that there has been a union with God.

In *The Interior Castle* St. Teresa supplements her earlier account, comparing the contemplative life to entering a castle or palace in which there are many rooms. These are arranged concentrically in six rings of rooms, or "mansions," round an inner chamber where the king lives. To enter this castle, prayer is needed. Ordinary Christians can enter the first three mansions through humility, meditation, and exemplary conduct; and the attainment of the third mansion represents the life achievement of many worthy Christians. But more remains in the spiritual life than such a virtuous existence. The fourth mansion corresponds to the "second water" of St. Teresa's earlier simile. In the fifth the soul seems to be asleep and unconscious both of the external world and of itself (although such language is analogical; the contemplative is not literally asleep). The soul is illuminated in this state by God. The sixth mansion is like a couple's first sight of one another at a betrothal. Finally, the soul enters the holy of holies. It seems as if this place is dark, because of the overpowering strength of the divine light. Here the soul has a direct vision of God, like the beatific vision to be enjoyed hereafter in heaven. Throughout these descriptions St. Teresa makes frequent use of the imagery of love and of marriage. The distinction between the "betrothal" and the "marriage" is found also in the writings of St. John of the Cross, a friend and follower of St. Teresa.

The detail and sensitivity of St. Teresa's autobiographical reports have given her a special importance in the history of mysticism.

See also John of the Cross, St.; Mysticism, History of; Mysticism, Nature and Assessment of; Women in the History of Philosophy.

Bibliography

St. Teresa's works are collected in Spanish as *Obras de Santa Teresa de Jesús*, edited by P. Silverio, 9 vols. (Burgos, 1915–1926); they appear in English as *Complete Works*, translated and edited by E. Allison Peers, 3 vols. (London:

Sheed and Ward, 1946). Also see E. Allison Peers, *Studies of the Spanish Mystics*, Vol. I (London: Sheldon Press, 1927).

OTHER RECOMMENDED WORKS

De Groot, Jean. "Teresa of Avila and the Meaning of Mystical Theology." In *Hispanic Philosophy in the Age of Discovery: Studies in Philosophy and the History of Philosophy,* Vol. 29, edited by Kevin White. Washington DC: Catholic University American Press, 1997.

Fales, Evan. "Scientific Explanations of Mystical Experiences, Part II." *Religious Studies* 32 (3) (1996): 297–313.

Waithe, Mary E. "Roswitha of Gandersheim, Christine Pisan, Margaret More Roper and Teresa of Avila" In *A History of Women Philosophers. Vol. 2: Medieval, Renaissance*. Norwell: Kluwer, 1989.

Ninian Smart (1967)
Bibliography updated by Tamra Frei (2005)

TERRORISM

Terrorism, whether practiced by states, substate groups, or individuals, is found throughout human history. Most historical accounts, however, focus on what they take to be forms of terrorism that are practiced by substate groups and individuals.

During Biblical times, Jewish Sicarii, known for their use of a short sword (sica), struck down rich Jewish collaborators who were opposed to violent resistance against their Roman conquerors. Later, in the eleventh and twelfth centuries, a group of Shiite Moslems, called the Assassins, opposed efforts to suppress their religious beliefs in Sunni-dominated Persia. Using daggers, the Assassins killed prefects, governors, and caliphs in front of many witnesses, thus ensuring their capture and execution because they believed that by their actions they would gain entry into paradise. Eventually, the group was suppressed by the Mongols in the thirteenth century.

In India, from the eleventh century on, a group called the Thugs was active until it was destroyed by the British in the nineteenth century. The Thugs ritually strangled their victims with a silk tie. They claimed allegiance to the goddess Kali, who it is said required them to kill in order to supply her with blood for nourishment.

Following the French Revolution, the Jacobins under Robespierre gave us the very term *terror*, unleashing a Reign of Terror between 1793 and 1794 upon all levels of French society. During this period, those executed included not only those accused of some offense or disloyalty, but sometimes their children, parents, or even grandparents as well.

Yet, it is not clear that all of these historical examples should be regarded, as they usually are, as acts of terrorism. Without a doubt, they are all cases in which terror (intense fear or fright or intimidation) is induced in large groups of people, but terrorism, as many have come to understand it, involves more than just this. First of all, many think that terrorism must have a political purpose—that it must aim to achieve some change in a government or governmental institution or policy. Now, this is true of most of the historical examples just cited, but it is not true of the Thugs of India whose goals were personal and religious rather than political. Second, many also think that terrorism must directly target innocents, a requirement that does not really hold of any of these historical examples except that of the Jacobins. The Sicarii targeted Jewish collaborators who in virtue of their collaboration were clearly not innocent. The Assassins attacked people in positions of political leadership who were responsible for the religious persecution against Shiite Moslems and so were not innocent. So the only really clear example we have here of terrorism is that of Robespierre's Reign of Terror, directed as it was at innocents as well as at those who were considered to be guilty of some offense. However, in the case of Robespierre's Reign of Terror, what we have is an example of state terrorism, not terrorism as practiced by substate groups or individuals.

Since 1983, the U.S. State Department has defined terrorism as follows: "Terrorism is premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents, usually intended to influence an audience." In a U.S. State Department document in which this definition is endorsed, there is also a section that discusses state-sponsored terrorism (Office of the Coordinator for Counterterrorism 2001). It is clear, then, that the U.S. State Department does not hold that only subnational groups or individuals can commit terrorist acts; it further recognizes that states can commit terrorist acts as well. So let us offer the following definition of terrorism, which is essentially the same as the U.S. State Department's definition once it is allowed that states, too, can commit terrorist acts and once it is recognized that it is through attempting to elicit terror (that is, intense fear, fright, or intimidation) that terrorists try to achieve their goals. The definition is: "Terrorism is the use or threat of violence against innocent people to elicit terror in them, or in some other group of people, in order to further a political objective."

Using this definition, there is no problem seeing the attacks on New York City and Washington, D.C., particu-

larly the attacks on the World Trade Center, as terrorist acts. Likewise, the bombing of the U.S. embassies in Kenya and Tanzania in 1998 as well as the suicide bombings directed at Israeli civilians are terrorist acts.

But what about the U.S. bombing of a pharmaceutical plant in Sudan with respect to which the United States blocked a United Nation's (U.N.) inquiry and later compensated the owner but not the thousands of victims who were deprived of drugs? Or what about the United States' $4 billion-a-year support for Israel's occupation of Palestinian lands, which began in 1969 and which is illegal, that is, in violation of U.N. resolutions that specifically forbid "the acquisition of territory by force" and which has resulted in many thousands of deaths? Or to go back further: What about U.S. support for the Contras in Nicaragua, and of death squads in El Salvador during the Reagan years, and the use of terrorist counter-city threats of nuclear retaliation during the Cold War and the actual use of nuclear weapons on Hiroshima and Nagasaki at the end of World War II resulting in over 100,000 deaths? Surely, all of these actions also seem to be either terrorist acts or support for terrorist acts according to our definition. How can we tell then, which, if any, of these terrorist acts or support for terrorist acts are morally justified?

Let us address this question from the perspective of the just war theory. In traditional just war theory, two requirements must be met in order to justify going to war. First, there must be a just cause. Second, just means must be used to fight the war. In order for there to be a just cause (1) There must be substantial aggression. (2) Nonbelligerent correctives must be either hopeless or too costly. (3) Belligerent correctives must be neither hopeless nor too costly.

Needless to say, the notion of substantial aggression is a bit fuzzy, but it is generally understood to be the type of aggression that violates people's most fundamental rights. To suggest some specific examples of what is and is not substantial aggression, usually the taking of hostages is regarded as substantial aggression while the nationalization of particular firms owned by foreigners is not so regarded. But even when substantial aggression occurs, frequently nonbelligerent correctives are neither hopeless nor too costly to pursue. And even when nonbelligerent correctives are either hopeless or too costly, in order for there to be a just cause, belligerent correctives must be neither hopeless nor too costly.

Traditional just war theory assumes, however, that there are just causes and goes on to specify just means as imposing two requirements: (1) Harm to innocents should not be directly intended as an end or a means. (2) The harm resulting from the belligerent means should not be disproportionate to the particular defensive objective to be attained. While the just means conditions apply to each defensive action, the just cause conditions must be met by the conflict as a whole.

Given the constraints imposed on just means, one might think that from the perspective of just war theory, acts of terrorism could never be morally justified. But this would require an absolute prohibition on intentionally harming innocents, and such a prohibition would not seem to be justified, even from the perspective of the just war theory. Specifically, it would seem that harm to innocents can be justified for the sake of achieving a greater good when the harm is: (1) trivial (e.g., as in the case of stepping on someone's foot to get out of a crowded subway), (2) easily reparable (e.g., as in the case of lying to a temporarily depressed friend to keep that person from committing suicide), or (3) nonreparable but greatly outweighed by the consequences of the action. Obviously, it is this third category of harm that is relevant to the possible justification of terrorism. But when is intentional harm to innocents nonreparable yet greatly outweighed by the consequences?

Consider the following example often discussed by moral philosophers: A large person who is leading a party of spelunkers gets stuck in the mouth of a cave in which flood waters are rising. The trapped party of spelunkers just happens to have a stick of dynamite with which they can blast the large person out of the mouth of the cave; either they use the dynamite or they all drown, the large person with them. Now, it is usually assumed in this case that it is morally permissible to dynamite the large person out of the mouth of the cave. After all, if that is not done, the whole party of spelunkers will die, the large person with them. So the sacrifice imposed on the large person in this case would not be that great.

But what if the large person's head is outside rather than inside the cave, as it must have been in the previous interpretation of the case. Under those circumstances, the large person would not die when the other spelunkers drowned. Presumably after slimming down a bit, the large person would eventually just squeeze out of the mouth of the cave. In this case, could the party of spelunkers trapped in the cave still legitimately use the stick of dynamite to save themselves rather than the large person?

Suppose there were ten, twenty, 100, or an even a larger number of spelunkers trapped in the cave. At some point, would not the number be sufficiently great that it would be morally acceptable for those in the cave to use

the stick of dynamite to save themselves rather than the large person, even if this meant that the large person would be morally required to sacrifice his life? The answer has to be yes, even if you think it has to be a very unusual case when we can reasonably demand that people thus sacrifice their lives in this way.

Is it possible that some acts of terrorism are morally justified in this way? It is often argued that the dropping of atomic bombs on Hiroshima and Nagasaki was so justified. President Truman, who ordered the bombing, justified it on the grounds that it was used to shorten the war. In 1945, the United States demanded the unconditional surrender of Japan. The Japanese had by that time lost the war, but the leaders of their armed forces were by no means ready to accept unconditional surrender. While the Japanese leaders expected an invasion of their mainland islands, they believed that they could make that invasion so costly that the United States would accept a conditional surrender.

Truman's military advisers also believed the costs would be high. The capture of Okinawa had cost almost 80,000 American casualties while almost the entire Japanese garrison of 120,000 men died in battle. If the mainland islands were defended in a similar manner, hundreds of thousands of Japanese would surely have died. During that time, the bombing of Japan would continue, and perhaps intensify, resulting in casualty rates that were no different from those that were expected from the atomic attack. A massive incendiary raid on Tokyo early in March 1945 had set off a firestorm and killed an estimated 100,000 people. Accordingly, Truman's Secretary of State James Byrnes admitted that the two atomic bombs did cause "many casualties, but not nearly so many as there would have been had our air force continued to drop incendiary bombs on Japan's cities" (Byrnes 1947, p. 264). Similarly, Winston Churchill wrote in support of Truman's decision: "To avert a vast, indefinite butchery … at the cost of a few explosions seemed, after all our toils and perils, a miracle of deliverance" (Churchill 1962, p. 634).

Yet the "vast, indefinite butchery" that the United States sought to avert by dropping atomic bombs on Hiroshima and Nagasaki was one that the United States itself was threatening, and had already started to carry out, with its incendiary attack on Tokyo. And the United States itself could have arguably avoided this butchery by dropping its demand for unconditional Japanese surrender. Moreover, a demand of unconditional surrender can almost never be morally justified since defeated aggressors almost always have certain rights that they should never be required to surrender. Hence, the United States' terrorist acts of dropping atomic bombs on Hiroshima and Nagasaki cannot be justified on the grounds of shortening the war and avoiding a vast, indefinite butchery if the United States could have secured those results simply by giving up its unreasonable demand for unconditional surrender. So, it is difficult to see how the dropping of atomic bombs on Hiroshima and Nagasaki could be justified acts of terrorism.

A more promising case for justified terrorism is the counter-city bombing of the British during the early stages of World War II. Early in the war, it became clear that British bombers could fly effectively only at night because too many of them were being shot down during day raids by German antiaircraft fire. In addition, a study done in 1941 showed that of those planes flying at night that were recorded as having actually succeeded in attacking their targets, only one-third managed to drop their bombs within five miles of what they were aiming at. This meant that British bombers flying at night could reasonably aim at no target smaller than a fairly large city. Michael Walzer (1992) argues that under these conditions, the British terror bombing was morally justified because at this early stage of the war, it was the only way the British had left to them to try to avert a Nazi victory. Walzer further argues that the time period when such terror bombing was justified was relatively brief. Once the Russians began to inflict enormous casualties on the German army and the United States made available its manpower and resources, other alternatives opened up. The British, however, continued to rely heavily on terror bombing right up until the end of the war, culminating in the fire-bombing of Dresden in which something like 100,000 people were killed. Nevertheless, for that relatively brief period when Britain had no other way to avert a Nazi victory, Walzer argues, its reliance on terror bombing was morally justified.

Suppose we agree with Walzer that British terror-bombing during the earlier stages of World War II was morally justified. Could there be a comparable moral justification for Palestinian suicide bombings against Israeli civilians? Israel has been illegally occupying Palestinian land since 1969 in violation of U.N. resolutions following the 1967 Arab–Israeli war. Even a return to those 1967 borders, which the U.N. resolutions require, still permits a considerable expansion of Israel's original borders as specified in the mandate of 1947. Moreover, since the Oslo Peace Accords in 1993 until 2001, Israeli settlements doubled in the occupied territories. Under Israel's prime minister Ariel Sharon, some thirty-five new settlements have been established in the occupied territories. In Gaza

in 2001, there were 1.2 million Palestinians and 4,000 Israelis, but the Israelis control 40% of the land and 70% of the water. In the West Bank, there were 1.9 million Palestinians and 280,000 Israelis, but the Israelis controlled 37% of the water.

In addition, Israel failed to abide by its commitments under the Oslo Peace Accords to release prisoners, to complete a third redeployment of its military forces, and to transfer three Jerusalem villages to Palestinian control. Moreover, at the Camp David Meeting in 2000, Israel's proposals did not provide for Palestinian control over East Jerusalem upon which 40% of the Palestinian economy depends. Nor did Israel's proposals provide for a right of return or compensation for the half of the Palestinian population that lives in exile, most of them having been driven off their land by Israeli expansion. So the Palestinian cause is arguably a just one, and clearly the Palestinians lack the military resources to effectively resist Israeli occupation and aggression by simply directly attacking Israeli military forces. The Israelis have access to the most advanced U.S. weapons and $4 billion-a-year from the United States to buy whatever weapons they want. The Palestinians have no comparable external support. Under these conditions, is there a moral justification for Palestinian suicide bombers against Israeli civilians? Assuming that the Palestinians lack any effective means to try to end the Israeli occupation or to stop Israel's further expansion into Palestinian territories other than by using suicide bombers against Israeli civilians, why would this use of suicide bombers not be justified in much the same way that Walzer justifies the British terror bombing in the early stages of World War II?

Much depends on what Israel's intentions are. If the Israelis have the ultimate goal of confining most Palestinians to a number of economically nonviable and disconnected reservations, similar to those on which the United States confines American Indian nations, would not the Palestinians have a right to resist that conquest as best they can, even if this involves the use of suicide bombers? Of course, everything here turns on a correct assessment of Israeli intentions and on whether Palestinians (and Israelis) have sufficiently exhausted the use of nonbelligerent correctives. The 2005 political overtures from Sharon might also indicate a new beginning. Only time will tell.

Starting with the just war theory, we have seen that there are morally defensible exceptions to the just means prohibition against directly killing innocents. The cave analogy argument aims to establish that conclusion. British terror bombing at the beginning of World War II, but not the American dropping of atomic bombs on Hiroshima and Nagasaki at the end of that war, seems to provide a real life instantiation of that argument. The Palestinian use of suicide bombers against Israeli civilians may or may not be a contemporary instantiation of that very same argument.

Yet, even if some acts of terrorism can be justified in this manner, clearly, most acts of terrorism cannot be so justified, and clearly, there was no moral justification for the terrorist attacks on New York City and Washington, D.C., particularly the attacks in the World Trade Center. For Americans, no act of terrorism compares with the September 11, 2001 (9/11), morning attacks on the World Trade Center and the Pentagon. Initial estimates put the number of dead from this terrorist attack at more than 5,000, but later the death toll was reduced to around 3,000. Comparisons were made to the Japanese attack on Pearl Harbor in 1941 where 2,403 sailors, soldiers, and civilians died. But the attack on a military outpost far removed from the American heartland is hardly comparable to an attack against targets in its largest city and in its capital. Nor was 9/11 carried out with the weapons of previous adversaries but by commandeering commercial aircraft with knives and box cutters and using them in murderous suicidal missions. So, this terrorism now faced is something new, something different, and, as a consequence, many people around the world feel vulnerable in a way they would have never thought possible before.

Even so, the question remains as to what is the appropriate response to unjustified terrorist acts. According to the just war theory, before using belligerent correctives, one must be sure that nonbelligerent correctives are neither hopeless nor too costly. The three weeks of diplomatic activity that the United States engaged in with the Taliban government of Afghanistan does not appear to have been sufficient to determine whether it was hopeless or too costly to continue to attempt to bring Osama bin Laden before a U.S court, or better, before an international court of law, prior to going to war against Afghanistan. The United States demanded that the Taliban government immediately hand over bin Laden and "all the leaders of Al Qaida who hide in your land" (Bush 2001). But was it reasonable to expect compliance from the Taliban, given that even after the overthrow of the Taliban government and the installation of a more friendly regime, the United States and its allies were still unable several years later to apprehend bin Laden and reduce the frequency of terrorist attacks sponsored by Al Qaida around the world? Was it reasonable for the United States to have expected the Taliban government, with its limited

resources and loose control over the country, to have done in three weeks what it was not able to accomplish after several years? Similar and even more telling questions can be raised about the decision to go to war against Iraq as a response to the threat of terrorism.

Terrorism, whether practiced by states, substate groups, or individuals, has a long and varied history. Whereas the practice can be generally condemned, many who condemn it most strongly are themselves engaged in terrorism or support for terrorism. More significantly, in order for responses to terrorism or the threat of terrorism to be morally justified, they must meet the requirements of the just war theory by first exhausting nonbelligerent correctives, and frequently, this is not done.

See also Just War Theory.

Bibliography

Bell, J. Bowyer. *Terror Out of Zion*. New York: St, Martin's Press, 1977.

Bickerton, Ian, and Carla Klausner. *A Concise History of the Arab-Israeli Conflict*. Upper Saddle River, NJ: Prentice Hall, 2002.

Byrnes, James F. *Speaking Frankly*. New York: Harper, 1947.

Crenshaw, Martha, "The Logic of Terrorism." In *Origins of Terrorism: Psychologies, Ideologies, Theologies, States of Mind*. Edited by Walter Reich. New York: Cambridge University Press, 1990.

Henderson, Harry. *Terrorism*. New York: Facts on File, 2001.

Kameel Nasr. *Arab and Israeli Terrorism*. London: McFarland, 1997.

Laquer, Walter. *The Age of Terrorism*. Boston: Little, Brown, 1987.

Malley, Robert, and Hussein Agha, "Camp David: The Tragedy of Errors." *New York Review of Books* 48 (13), 2001.

Nash, Jay Robert. *Terrorism in the 20th Century*. New York: M. Evans, 1998.

Office of the Coordinator for Counterterrorism. *Patterns of Global Terrorism-2000*, 2001.

Rappoport, David. "Religion and Terror: Thugs, Assassins, and Zealots." In *International Terrorism: Characteristics, Causes, Controls*. Edited by Charles Kegley, 147–149. New York: St. Martin's Press, 1990.

Simon, Jeffrey. *The Terrorist Trap*. Bloomington: Indiana University Press, 2001.

Walzer, Michael. *Just and Unjust Wars*, 2nd ed. New York: Basic Books, 1992

James P. Sterba (2005)

TERTULLIAN, QUINTUS SEPTIMIUS FLORENS

(c. 160–c. 220)

Quintus Septimius Florens Tertullian, the African Church Father, was born in Carthage and was converted to Christianity about 193. He made early use of his training in rhetoric and Roman law in two apologetic works, *Ad Nationes* and *Apologeticum*, written in 197. These owe much to earlier Greek Christian apologies and to the writings of Varro, an Augustan polymath who analyzed religion along Stoic lines; *Ad Nationes* seems to have been a first draft of the *Apologeticum*. Tertullian was the first Christian theologian to write in Latin, and most of his works deal with moral and theological issues; all contain elements of polemic either against various aspects of Greco-Roman culture or against Christian heresies. Tertullian's works can be dated by cross-references, allusions to current events, and by his gradual movement toward the ascetic-apocalyptic sect of the Montanists, advocates of the "new prophecy"; he became a Montanist about 206 and later became the leader of a Montanist group in Carthage. Nothing is known of his life after the time of his last literary work, written about 220.

His writings are vigorously, even violently, individualistic in style and often in content; he loved paradox and contradiction, going so far as to claim in *De Carne Christi* (Ch. 5) that the incarnation of Christ "*certum est quia impossibile*" ("is certain because impossible"). This claim seems to be based on a line of argument found in Aristotle's *Rhetoric* (Book 2, Ch. 23, Sec. 22): It is likely that unlikely things should happen. Tertullian's philosophical theology is derived largely from his Greek Christian predecessors (St. Justin Martyr, Tatian, St. Theophilus, Irenaeus); his own contributions are chiefly Stoic in origin. For him philosophy is partly, or sometimes, an enemy of religion ("What does Jerusalem have to do with Athens?"), sometimes an ally ("Seneca is often one of us").

Only two of Tertullian's nonapologetic works are primarily concerned with philosophical themes. One is the early treatise *Adversus Hermogenes*, in which he attacks the doctrine that matter is eternal and claims that Hermogenes derived this belief from Platonic and Stoic sources. His own arguments against the eternity of matter are partly a revision of a lost book by Theophilus, as the common Genesis text indicates. Hermogenes argued that the immutable God cannot have created the world from himself or have begun to create it ex nihilo; therefore he must have made it from matter, to which its

imperfections are to be ascribed. God continually "creates," influencing matter as a magnet influences iron. In reply, Tertullian insisted primarily on God's freedom from "necessity." God created by his free will and therefore was not limited by matter.

His other work of philosophical interest is the Montanist treatise *De Anima* (c. 210–213), which is intended to prove that Platonic teaching is false. The soul is actually corporeal and originates from a "soul-producing seed" at the moment of conception. It is not preexistent and does not transmigrate—an argument directed not only against Platonists but also against Christian heretics, chiefly Gnostic. Tertullian also discusses the human embryo and other related topics. His work is largely based on a treatise on the soul by the Greek physician Soranus, who wrote at Rome early in the second century. From Soranus, Tertullian derives most of his discussions of Plato, the Stoics, Aristotle, Heraclitus, and Democritus. Tertullian's importance thus lies in his mediation of earlier conceptions, Christian and pagan alike, and for his translation of Greek ideas into Latin.

See also Apologists; Aristotle; Heraclitus of Ephesus; Leucippus and Democritus; Plato; Platonism and the Platonic Tradition; Stoicism.

Bibliography

TEXTS AND TRANSLATIONS

First modern edition by F. Oehler, 3 vols., Leipzig, 1853–1854; A Reifferscheid et al. (*Corpus Scriptorum Ecclesiasticorum Latinorum* XX, XLVII, LXX. Vienna, 1890–); and E. Dekkers et al. (*Corpus Christianorum*. Series Latina, I–II, Turnhout: Brepols, 1954–).

English translations in *Ante-Nicene Christian Library* VII, XI, XV, XVIII. Edinburgh: T. and T. Clark, 1868–1870.

Castorina, E., ed. *De spectaculis*. Florence, 1961.

Evans, E., ed. *Adversus Praxeam*. London, 1948.

Evans, E., ed. *De carne Christi*. London 1956.

Evans, E., ed. *De resurretione carnis*. London 1960.

Pollmann, K. *Das Carmen adversus Marcionitas: Einleitung, Text, Übersetzung und Kommentar*. Göttingen: Vandenhoeck & Ruprecht, 1991.

Refoulé, R. F. *De praescriptione haereticorum. Sources Chrétiennes* XLVI. Tübingen, 1910; Frankfurt: Minerva, 1968.

Schneider, A., ed. *Ad nations*. Rome, 1968.

Waszink, J. P., ed. *De anima*. Amsterdam, 1947.

STUDIES

Aziza, Claude. *Tertullien et le judaïsme*. Paris: Belles Lettres, 1977.

Barnes, T. D. *Tertullian: A Historical and Literary Study*. Oxford: Clarendon, 1971; rev. ed., 1985.

Braun, R. et al., ed. *Chronica Tertullianea et Cyprianea, 1975–1994: Bibliographie critique de la première littérature latine chrétienne*. Paris, 1999.

Bray, G. *Holiness and the Will of God: Perspectives on the Theology of Tertullian*. London; Atlanta: John Knox Press, 1979.

Dekkers, E. *Tertullianus en de Geschiednis der Liturgie*. Münster, 1947.

Dunn, Geoffrey. *Tertullian*. London: Routledge, 2004.

Fredouille, J. C. *Tertullien et la conversion de la culture antique*. Paris: Études augustiniennes, 1972.

Moingt, J. *Théologie trinitaire de Tertullien*. Théologie 68–70. Paris: Aubier, 1966.

Nisters, B. *Tertullian: Seine Persönlichkeit un sein Schicksal*. Münsterische Beiträge zur Theologie 25 (1950).

O'Malley, T. P. *Tertullian and the Bible*. Latinitas Christianorum Primaeva 21. Utrecht: Dekker & Van de Vegt, 1967.

Osborne, E. F. *Tertullian, First Theologian of the West*. Cambridge, U.K.: Cambridge University Press, 1997.

Questen, J. *Patrology*, II. pp. 246–340. Utrecht: Spectrum Publishers, 1953.

Rankin, D. I. *Tertullian and the Church*. Cambridge, U.K.: Cambridge University Press, U.K., 1995.

Säflund, G. *De Pallio und stilistische Entwicklung Tertullians*. Skrifter utgivna av Svenska Institutet i Rom VIII, Lund: C. W. K. Gleerup, 1955.

Sider, D. *Ancient Rhetoric and the Art of Tertullian*. Oxford: Oxford University Press, 1971.

Vecchiotti, I. *La filosofia di Tertulliano. Un colpo di sonda nella storia del cristianesimo primitivo*. Urbino: Argalìa, 1970.

Robert M. Grant (1967)

Bibliography updated by Scott Carson (2005)

TESTIMONY

The term *testimony* in contemporary analytic philosophy is used as label for the spoken or written word, when this purports to pass on the speaker's or writer's knowledge, conveying factual information or other truth. Testifying, or giving testimony, is a linguistic action, and testimony is its result, an audible speech act of telling or more extended discourse (perhaps recorded), or a legible written text. Interest in the topic has grown rapidly since the publication of C. A. J. Coady's *Testimony: A Philosophical Study* (1992). Testimony in this broad sense includes the central case of one person telling something to another in face-to-face communication, as well as a range of other cases, from public lectures, television and radio broadcasts, and newspapers to personal letters and e-mails, all kinds of purportedly factual books and other publications, and the information recorded in train timetables, birth registers, and official records of many kinds.

PHILOSOPHICAL ISSUES ABOUT TESTIMONY

The key interest of testimony is as a source for individual human knowledge, alongside perception, memory, inference, and intuition. Thus attaining a correct account of its epistemology is the core organizing issue for explanatory philosophical theorizing about testimony. This interlocks with several other issues.

First, there is no believing what one is told, without first understanding it—grasping both content and force of the speech act. And knowledge of what one was told surely rests on knowledge that one was told. Thus an account of testimony needs to be supplemented with an account of linguistic understanding—both its psychology and its epistemology. Understanding in turn cannot be fully explained except as part of the large project of explaining linguistic meaning, the significance of words, which is grasped when a speech act is understood. Second, telling is just one of the many diverse activities that make up the human social institution of language. Why and how it is epistemically justified to believe the purport of a linguistic act of telling turns on the nature of that act. Appreciation of the interpersonal relations involved in linguistic exchange, especially the commitments and norms involved in the making and reception of the speech act of assertion, must inform our account of testimony.

Third, an account of what makes belief acquired from testimony become knowledge will be persuasive only if it instances a convincing general conception of knowledge; and similarly for justified testimonial belief. Fourth, how is testimony best individuated as an epistemic kind? It is clear that the following very broad category is not one about which any interesting generalizations may be made: whatever may, on occasion, be justifiedly inferred by an audience from observing someone assert that *P*. But exactly how narrow the kind is that we should discern as the core case—what we may call knowledge (or justified belief) from testimony—is debatable. In general knowledge from testimony that *P*, there will be knowledge with that same content *P*; but knowledge of an intended message can also be acquired through sarcasm and metaphor, and despite minor linguistic infelicity by the speaker. One may come to know that *P*, where one's knowledge rests essentially on the fact that *S* told one that *P*, but where one's reason for forming belief in what she said is not that one trusts her to know whereof she speaks, but that one has circumstantial evidence that her utterance, though not from knowledge, is nonetheless sure to be true. A speaker, for instance one whose job it is to instruct, may convey empirically well-established facts that she for perverse reasons does not believe. Can others acquire knowledge from her instruction?

These and other problem cases render the precise individuation of our epistemic kind a subtle and debatable matter. Some argue that the core case is confined to when the testifier speaks from her own knowledge, and her audience trusts her to do so, accepting her word for what she tells on that basis. This is argued to be the core case, because in it alone the audience accepts the teller's linguistic act of assertion at face value as what it purports to be, an expression of knowledge. She accepts the warrant to believe on her say-so offered by the teller. But others, considering cases such as those mentioned above, argue for a broader conception, on which it is not necessary that the testifier speak from knowledge in order for one to acquire knowledge from testimony.

THE IDEAL OF EPISTEMIC AUTONOMY VERSUS MODERN RELIANCE ON KNOWLEDGE AT SECONDHAND, FROM TESTIMONY

An individualist strand in Western philosophy castigates belief derived from testimony as epistemically inferior. Plato (in the *Theatetus*) and Augustine (in *De Magistro*) despised its secondhand character and denied that knowledge, as opposed to mere belief, can ever be acquired from it. Rene Descartes (in his *Meditations on First Philosophy*) insisted on building his knowledge afresh from individualist foundations, and John Locke (in his *Essay on the Human Understanding*) rejected "other men's opinions floating in one's brain" as never amounting to knowledge. They were correct that belief derived from testimony is epistemically problematic and arguably inferior in two related respects, entailed by its being knowledge at secondhand.

First, one who forms belief that *P* on trust in another's testimony does not herself possess the evidence for *P*, but instead a second-order warrant. Her own immediate basis for believing *P* is that she trusts her teller to knows whereof she speaks. This entails that the teller, or some other person or group of people upstream of her in a chain of testimony, possesses nontestimonial evidence establishing the truth of *P*. The trusting recipient of testimony is committed to belief in the existence of this evidence, of which she is personally ignorant, and that her informants have evaluated it correctly. Insistence that, for a first-class warrant amounting to knowledge one must possess the evidence for *P* oneself, would rule out all knowledge thus based on trust in the word of others—and hence, in others' honesty and epistemic good judgment.

Second, such trust is epistemically risky. One who testifies that *P* in an act of assertion purports to speak from knowledge. But her own belief may be false: she may have failed to form belief in an epistemically responsible way, or may have been the subject of bad epistemic luck, and may have fallen into honest error. Or she may be insincere, intent on deception. There are many entirely understandable and common human motives for this. Circumstances are many and frequent in which personal advantage may be gained by lying, and it can require altruism or courage to tell the truth in difficult circumstances. These risks incurred in believing what others tell us mean that we should place our trust in the word of others discriminatingly and circumspectly. The epistemically responsible recipient of testimony will be aware of the need for both sincerity and competence about her topic in her source, and her response will be mediated by this.

But the price of maintaining Descartes's ideal of epistemic self-reliance would be infeasibly high, in the condition of extensive division of epistemic labor that characterizes our modern, highly socialized existence. Topics that we know of, for the most part, only from testimony include: all of history, including our own early personal and family history; much of the geography and politics of the contemporary world; nearly all of knowledge in the various specialized domains of human inquiry—the natural and social sciences, humanities, and so forth. In addition, we rely heavily in our daily lives on the fruits of advanced technology, from plumbing and motor mechanics to information technology and dentistry, about which most of us know little. Each one of us would be unimaginably epistemically and practically impoverished without knowledge learned from trust in the testimony of others.

THE TASKS FOR A POSITIVE EPISTEMOLOGY OF TESTIMONY

A more constructive theoretical approach takes the primary task for epistemology to be the following: to explain precisely how and in what circumstances testimony can yield knowledge and justified belief. This task may be subdivided into micro and macro issues. The central case of testimony occurs when one person tells something to another, thereby expressing her knowledge, and the other understands and believes her, taking her word for it. When all goes as it should, knowledge is thereby shared, and by recursion of this mechanism it may be diffused through a community of speakers of a shared language. Our micro question is: How precisely is knowledge spread from teller to audience in this core process? What are the conditions for belief formed in what one is told to be justified, and knowledgeable?

The macro issues are: How pervasive is epistemic dependence on testimony, in the system of empirical belief of each of us? Can this epistemic dependence be eliminated, in principle or in practice? How much of one's belief system would be left, after such pruning? We have already seen that a very great deal of what an individual believes, in our modern society, is learned initially from testimony. This does not entail that these beliefs are still epistemically dependent on testimony, since the believer may later acquire other, independent evidence—for instance, when one sees for oneself a place of which one has previously only read. Support from coherence and inference to the best explanation may sustain a system of belief initially acquired from trust in testimony. But testimony plays a key role in putting in place the framework—of land masses and seas, cities and nations, natural and social history, and so forth, in terms of which we theorize our experiences. Thus the idea of eliminating dependence on testimony is problematic, and it is not clear that we have any beliefs that are entirely free of epistemic dependence on testimony—hence the unlivability of the supposed ideal of epistemic autonomy.

Hume (1777) thought that knowledge could be gained from testimony, but the warrant to believe it came only with empirical evidence of the reliability of testimony as a source. Reid (1764), in contrast, argued that human nature includes two complementary dispositions, to truthfulness and trustfulness, and that this engenders a defeasible a priori warrant to trust others' testimony. Their two views instance what may be called the reductionist versus the anti-reductionist stance regarding our micro question: What is the basis of a hearer's epistemic entitlement to trust what someone tells her? Coady argues against reductionism, in favor of the view that our knowledge from testimony can only be explained by positing an epistemic principle special to testimony. There is an a priori, albeit defeasible, epistemic entitlement to trust any giver of testimony: One may presume true whatever one is told, so long as one is not aware of evidence that defeats one's presupposition of the sincerity and competence of one's informant. Coady advances several arguments for this view. His first main argument is transcendental: We do gain justified belief, and knowledge, from testimony. But it is impossible noncircularly to establish that testimony is generally reliable; therefore (on pain of denying that testimony can yield knowledge) a hearer must be entitled in effect to presume this on no evidence.

His second argument invokes considerations about the interpretation of the language of a community, to argue that the supposition that all reports made in that community are false is incoherent. He suggests that this fact underwrites an epistemic right to trust on no evidence, in the absence of defeaters. Burge (1993) gives another argument for antireductionism: Testimony is presumed to come from a rational source, and in the absence of counterevidence, such a source is presumed true. Fricker (in Chakrabarti and Matilal 1994) argues against Coady's transcendental argument, and presses the presumptive case for reductionism, from the epistemic riskiness of trusting others. She argues that epistemic responsibility requires monitoring others for sincerity and competence, and believing what they tell only if there is empirical basis for trusting them.

Further questions include: What is the range of subject matters on which a person may properly defer to the word of another, so that testimony on it may properly be given and accepted? For instance, can one properly accept, even defer to, another's word on moral, or aesthetic matters? Extensive division of epistemic labor characterizes the sciences, and all academic disciplines in which there is a domain of specialized knowledge and inquiry. There are many issues about the nature of trust and epistemic dependence in these specialized epistemic domains. In the sciences, many results depend on collaborative research from large numbers of individuals, members of collaborating research teams. In history, the judicious evaluation of oral and written testimonial sources is methodologically crucial. The status of testimony in formal settings such as legal ones is another area of interest.

See also Augustine, St.; Descartes, René; Epistemology; Hume, David; Inference to the Best Explanation; Intuition; Knowledge and Truth, The Value of; Locke, John; Memory; Perception; Plato; Reid, Thomas.

Bibliography

Adler, Jonathan. *Belief's Own Ethics*. Cambridge, MA: MIT Press, 2002.

Audi, R. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. London and New York: Routledge, 1998.

Burge, T. "Content Preservation." *Philosophical Review* 102 (1993): 457–488.

Chakrabarti, A., and B. Matilal, eds. *Knowing from Words*. Dordrecht: Kluwer Academic Publishers, 1994.

Coady, C. A. J. *Testimony: A Philosophical Study*. Oxford: Clarendon Press, 1992.

Elgin, Catherine Z. "Take It from Me: The Epistemological Status of Testimony." *Philosophy and Phenomenological Research* 65 (2) (2002): 291–308.

Fricker, E. "Trusting Others in the Sciences: A Priori or Empirical Warrant?" *Studies in History and Philosophy of Science* 33 (2002): 373–383.

Goldman, A. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63 (1) (2001): 85–110.

Graham, Peter J. "The Reliability of Testimony." *Philosophy and Phenomenological Research* 61 (3) (2000): 695–709.

Hardwig, J. "The Role of Trust in Knowledge." *Journal of Philosophy* 88 (1991): 693–708

Hume, David. *Enquiry concerning Human Understanding* (1777), edited by P. H. Nidditch. Oxford: Clarendon Press, 1975.

Jones, Karen "Second-Hand Moral Knowledge." *Journal of Philosophy* 96 (1999): 55–78.

Lackey, J. "Testimonial Knowledge and Transmission." *Philosophical Quarterly* 49 (1999): 471–490.

Lackey, J., and E. Sosa, eds. *The Epistemology of Testimony*. Oxford, forthcoming.

Reid, Thomas. *An Inquiry into the Human Mind on the Principles of Common Sense* (1764), edited by Derek Brookes. University Park: Pennsylvania State University Press, 1997.

Elizabeth Fricker (1996, 2005)

TETENS, JOHANN NICOLAUS

(1736 or 1738–1807)

Johann Nicolaus Tetens, the German philosopher and psychologist, was born in Tetenbüll, Schleswig, in 1736 or in Tönnig, Schleswig, in 1738, and died in 1807. He studied at the universities of Rostock and Copenhagen and became a *Magister* at Rostock University in 1759. From 1760 until 1765, when he became director of the local Gymnasium, he taught physics at Bützow Academy. He was full professor of philosophy at the University of Kiel from 1776 to 1789, during which period he also carried out an official study of the local hydraulic installations on the North Sea coast. From 1789 until his death he had a brilliant career as a high financial official in Copenhagen.

Tetens was strongly influenced by J. C. Eschenbach, his teacher of philosophy at Rostock. Eschenbach was an eclectic who accepted some Leibnizian and Wolffian tenets but sided with the Pietists against Christian Wolff; nevertheless, he seems to have been influenced more by the Berlin Academy and by John Locke's empiricism than by C. A. Crusius. Tetens likewise was influenced by Locke and, after their publication, by Gottfried Wilhelm Leibniz's *Nouveaux Essais*. Among his contemporaries he was influenced by David Hartley, Abraham Tucker, J. G. Sulzer, Claude-Adrien Helvétius, and Charles Bonnet.

Tetens was one of the first in Germany to discuss David Hume at length. J. H. Lambert's *Architektonik* and Immanuel Kant's *Inaugural Dissertation* later played important roles in the development of Tetens's own views.

Tetens hoped to reform German metaphysics by using the critical approach of the new empirical psychology. He wished to restore metaphysics in a new form that would meet the criticisms based on the skeptical and psychological orientations of the English and French schools, then widely influential in Germany. On the other hand he defended phenomenalism against the adherents of the schools of common sense and of "popular philosophy."

In his first significant work, *Ueber die allgemeine spekulativische Philosophie* (On general speculative philosophy; Bützow and Wismar, 1775; reprinted Berlin, 1913), Tetens discussed the weaknesses of traditional metaphysics and proposed some remedies. He held that to reform metaphysics, the sources and development of metaphysical concepts must be investigated. The means of inquiry was "inner sense," or introspection. He tried to give purely psychological answers to psychophysiological problems on the one hand and to metaphysical problems on the other.

In this spirit, Tetens's major work, *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* (Philosophical essays on human nature and Its development; 2 vols., Leipzig, 1777; reprinted Berlin, 1913), was an extended inquiry into the origin and structure of knowledge. He distinguished three faculties of the human mind: understanding, will, and feeling of pleasures and pains. He stressed the independence of the third faculty from the first two. The three may be reducible to one, but if so, according to Tetens, we cannot know it.

The mind is essentially active. Even sensation implies a reaction of the subject to the thing sensed. There are three fundamental activities of representation: perception, reflection (or abstraction from perceptions), and fiction (or the construction of new ideas out of perceived and abstracted representations).

Relations are established among perceived things by means of "primary original notions of relationships," or "forms"; one such form is causal connection. The three activities of representation together with the forms bring about the "concept of an object." Tetens proposed a rule for deciding whether something exists subjectively or objectively—we attribute a sensation to a thing if the sensation is contained as a part in the entire sensation of the thing.

Tetens distinguished rational knowledge from sensible knowledge by its being general and necessary. Metaphysical first principles are undeniable because they are rooted in the essence of the ego. They are like natural laws to which the intellect is subjected. The intellect—or common sense—and reason are governed by different kinds of laws, and the confusion between the two kinds of laws brings them into conflict.

Tetens discussed with great insight many other extremely complicated problems in metaphysics, ethics, the philosophy of education, and the philosophy of language. His *Philosophische Versuche* exerted a tremendous influence on Kant while he was writing the *Critique of Pure Reason*, and the many similarities between their doctrines are evident. Tetens's doctrines may be compared to Kant's even in their speculative power and importance.

See also Bonnet, Charles; Crusius, Christian August; Empiricism; Hartley, David; Helvétius, Claude-Adrien; Hume, David; Kant, Immanuel; Lambert, Johann Heinrich; Leibniz, Gottfried Wilhelm; Locke, John; Sulzer, Johann Georg; Wolff, Christian.

Bibliography

Brenke, M. *J. N. Tetens' Erkenntnistheorie vom Standpunkte des Kritizismus*. Rostock: Boldt'sche Hofbuchdruckerei, 1901.

Fuchs, Arnold. *J. N. Tetens' pädagogische Anschauungen*. Langensalza, 1918.

Golembski, W. "Die deutsche Aufklärungsphilosophie als Quelle des Transzendentalismus, I, Die Ontologie des J. N. Tetens." *Bulletin International de l'Academie Polonaise des Sciences et des Lettres* (1934): 167–173.

Kuehn, Manfred. "Hume and Tetens." *Hume Studies* 15(2)(1989): 365–375.

Lorsch, J. *Die Lehre vom Gefühl bei Tetens*. Giessen, 1906.

Schinz, M. *Die Moralphilosophie von Tetens*. Leipzig, 1906.

Seidel, Arthur. *Tetens Einfluss auf die kritische Philosophie Kants*. Leipzig, 1932.

Sommer, R. *Grundzüge einer Geschichte des deutschen Psychologie und Aesthetik von Wolff–Baumgarten bis Kant–Schiller*, 260–302. Würzburg, 1892.

Störring, G. *Die Erkenntnistheorie von Tetens*. Leipzig, 1901.

Uebele, W. "Herder und Tetens." *Archiv für Geschichte der Philosophie* 18 (1905): 216–249.

Uebele, W. *J. N. Tetens nach seiner Gesamtentwicklung betrachtet mit besonderer Berücksichtigung des Verhältnisses zu Kant*. Berlin: Reuther and Reichard, 1911.

Giorgio Tonelli (1967)

Bibliography updated by Tamra Frei (2005)

THALES OF MILETUS

(sixth century BCE)

Thales of Miletus is widely depicted in ancient sources as a pioneering rationalist and the founding father of Greek philosophy, science, and mathematics. Famous for ingenuity in many areas, he was also numbered among the seven sages (*Sophoi* or wise men). Evidence for his life and thought is meager and often questionable. Although written work is attested, nothing survives and he probably wrote nothing (Greek script still had limited uses). The earliest extant reports come from the historian Herodotus (c. 484–between 430 and 420 BCE); other evidence derives largely from Aristotle and his younger colleagues, Theophrastus and Eudemus (fourth century BCE). Hence, the reliability of the evidence depends heavily on the accuracy of the information available to them. Their testimony has been challenged by many scholars. But recent studies afford grounds for confidence, in part by tracing how Thales' ideas were transmitted by his intellectual heirs, including his younger compatriots Anaximander and Anaximenes.

Thales is a pivotal figure not unlike Galileo Galilei. Before him come cosmogonic verse (influenced by Near Eastern and Egyptian traditions) and a century of rapid advances in Greek culture, most notably in civic institutions and technology (e.g., building, coinage, and writing). In his wake, empirical inquiry, abstract speculation, and critical debate flower. Although his role in those developments cannot be assessed precisely, it was probably seminal. Early sources tell of travel to Egypt (where Miletus had a major trading depot), regional diplomacy (advocating a federation of Ionian cities to counter aggressive foreign neighbors), and diverse feats of engineering (diverting the course of the Halys River), economics (monopolizing olive presses), and surveying (calculating the height of pyramids and the distance of ships from shore).

Thales' significance for the history of philosophy stems mainly from his insights in three areas: cosmology, astronomy, and geometry. He is best known today for the bold but obscure claim that water is the *archē* (source or basic causal factor) of everything, ostensibly on the grounds that moisture (not water narrowly defined but fluid generally) is both the "seed" (originating source) and "food" (source of growth and sustenance) of all things. What exactly Thales said or meant is unrecoverable. Aristotle, the primary source for these claims, calls him the founder of material explanation: specifying the material constituents responsible for persistence and change. Thales also proposed that the earth floats on water "like wood"; and he attributed earthquakes to the earth's occasional rocking. Related considerations probably included the mobility of water, its exceptional mutability (readily solidifying and vaporizing), and its ubiquity (falling from the sky, emerging from springs, and both surrounding the land and filling its depressions).

Antiquity admired Thales most for his astronomy. Most famous was his alleged prediction of a solar eclipse (securely dated to May 28, 585 BCE) that halted a major foreign battle. The story, which many scholars doubt, appears first in Herodotus, who says only that he forecast the year. But a newly recovered text on papyrus cites the astronomer Aristarchus of Samos (flourished c. 270 BCE) crediting Thales with discovering the cause of solar eclipses by first determining that they occur only at a new moon. Other reports of his stargazing are more credible: charting the periodic rising and setting of prominent stars and star clusters (as in Hesiod's verse, over a century earlier); introduction of a circumpolar constellation (Ursa Minor); and a rough determination of the solstices and equinoxes, which enabled him to correlate the annual cycles of the sun and stars more reliably, thereby improving Greek calendrical schemes. Methodical observation of the horizon was the basis for most of these discoveries, but study of the lunar cycle is also reported.

Several new insights in geometry are ascribed to Thales: the equality of the opposite angles formed by intersecting lines; the equality of the base angles in isosceles triangles; the bisection of circles by their diameters; the congruence of triangles having a side and two angles equal; and the proportionality of similar triangles. The latter two are cited in connection with practical procedures: the former to calculate the distance of ships, the second to calculate the height of pyramids in Egypt. The novelty of his ideas probably lay not in simply enunciating these elementary propositions, nor in their formal proof, but in asserting their universal scope on the basis of ad hoc reasoning or evidence. Other innovative ideas attributed to Thales include the earliest recorded explanation for the Nile's annual flooding (seasonal winds obstruct its flow), a claim that amber and magnets are animate (because they cause motion, though curiously not self-motion), and a claim that all things are full of gods (perhaps because full of water, which exhibits two standard attributes of divinity: it is both deathless and life-giving). Implicit in many of the views attributed to Thales are basic principles of rational inquiry and naturalistic explanation: observation, analysis, abstraction,

generalization, and regularity. Provided that some of this evidence is accurate, Thales may reasonably be counted as the first philosopher—well before the word was coined.

See also Anaximander; Anaximenes; Archē; Aristotle; Nomos and Phusis; Pre-Socratic Philosophy; Theophrastus.

Bibliography

TEXTS AND TRANSLATIONS

Diels, Hermann, and Walther Kranz, eds. *Die Fragmente der Vorsokratiker.* 3 vols. 6th ed. Berlin: Weidmann, 1951–1952. The standard (but incomplete) collection of ancient testimony (widely cited as DK or *VS*).

Kirk, Geoffrey S., J. E. Raven, and Malcolm Schofield. *The Presocratic Philosophers*. 2nd ed. New York: Cambridge University Press, 1983. A useful selection of texts with translation and judicious analysis.

STUDIES

Mansfeld, Jaap. "Aristotle and Others on Thales, or The Beginnings of Natural Philosophy." *Mnemosyne* 40 (1987): 109–129. Analyzes Aristotle's testimony.

O'Grady, Patricia F. *Thales of Miletus: The Beginnings of Western Science and Philosophy*. Aldershot, U.K.: Ashgate, 2002. Suggestive but often uncritical discussion.

Panchenko, Dmitri. "Thales and the Origin of Theoretical Reasoning." *Configurations* 3 (1993): 387–414. An incisive case for his philosophical significance.

White, Stephen A. "Thales and the Stars." In *Presocratic Philosophy*, edited by Victor Caston and Daniel W. Graham. Aldershot, U.K.: Ashgate, 2002. A favorable reassessment of evidence for his astronomical study.

Stephen A. White (2005)

THEISM

The central claim of theism is that God exists. According to a standard version of this doctrine, God is omniscient, omnipotent, perfectly good, and the creator of all contingent things. According to more developed versions, God intervenes in the created world in order to answer prayers and perform miracles. Developed versions of theism are often contrasted with deism because deists hold that God created the contingent world but does not subsequently intervene in it.

Various aspects of theism are discussed in the following articles in the Encyclopedia:

> *Agnosticism; Analogy in Theology; Atheism; Common Consent Arguments for the Existence of God; Cosmological Arguments for the Existence of God; Creation and Conservation, Religious Doctrine of; Degrees of Perfection, Argument for the Existence of God; Deism; Epistemology, Religious; Evil, The Problem of; Faith; Fideism; Foreknowledge and Freedom, Theological Problem of; God, Concepts of; Hiddenness of God; Infinity in Theology and Metaphysics; Miracles; Moral Arguments for the Existence of God; Mysticism, Nature and Assessment of; Ontological Argument for the Existence of God; Pantheism; Perfection; Philosophy of Religion; Philosophy of Religion, Problems of; Physicotheology; Popular Arguments for the Existence of God; Providence; Religious Experience; Religious Experience, Argument for the Existence of God; Revelation; Teleological Argument for the Existence of God; Theism, Arguments For and Against.*

THEISM, ARGUMENTS FOR AND AGAINST

Philosophy of religion enjoyed a renaissance in the final third of the twentieth century. Its fruits include important contributions to both natural theology, the enterprise of arguing for theism, and natural atheology, the enterprise of arguing against it. In natural theology philosophers produced new versions of ontological, cosmological, and teleological arguments for the existence of God. In natural atheology problems of evil, which have always been the chief arguments against theism, were much discussed, and philosophers debated proposed solutions to both the logical problem of evil and the evidential problem of evil.

NATURAL THEOLOGY

Building on work by Charles Hartshorne and Norman Malcolm, Alvin Plantinga (1974) formulated a model ontological argument for the existence of God that employs the metaphysics of possible worlds. Let it be stipulated that being unsurpassably great is logically equivalent to being maximally excellent in every possible world and that being maximally excellent entails being omnipotent, omniscient, and morally perfect. The main premise of Plantinga's argument is that there is a possible world in which unsurpassable greatness is exemplified. From these stipulations and this premise he concludes, first, that unsurpassable greatness is exemplified in every possible world and hence in the actual world and, second, that there actually exists a being who is omnipotent, omniscient, and morally perfect and who exists and has these properties in every possible world. The argument is valid

in a system of modal logic that can plausibly be claimed to apply correctly to possible worlds. Plantinga reports that he thinks its main premise is true and so considers it a sound argument.

However, he acknowledges that it is not a successful proof of the existence of God. A successful proof would have to draw all its premises from the stock of propositions accepted by almost all sane or rational persons. The main premise of this argument is not of that sort; a rational person could understand it and yet not accept it. In other words, not accepting the argument's main premise is rationally permissible. But Plantinga maintains that accepting that premise is also rationally permissible. Since he regards it as rational to accept the argument's main premise, he holds that the argument shows it to be rational to accept its conclusion. As he sees it, even though his ontological argument does not establish the truth of theism, it does establish the rational permissibility of theistic belief.

According to William L. Rowe (1975), Samuel Clarke has given us the most cogent presentation of the cosmological argument we possess. It has two parts. The first argues for the existence of a necessary being, and the second argues that this being has other divine attributes such as omniscience, omnipotence, and infinite goodness. As Rowe reconstructs it in contemporary terms, the first part of the argument has as its main premise a version of the principle of sufficient reason, according to which every existing thing has a reason for its existence either in the necessity of its own nature or in the causal efficacy of some other beings. It is then argued that not every existing thing has a reason for its existence in the causal efficacy of some other beings. It follows that there exists a being that has a reason for its existence in the necessity of its own nature. Next it is argued that a being that has a reason for its existence in the necessity of its own nature is a logically necessary being. It may then be concluded that there exists a necessary being.

Rowe takes care to ensure that his version of Clarke's argument is deductively valid. What is more, he maintains that the principle of sufficient reason that is its main premise is not known to be false because no one has set forth any convincing argument for its falsity. However, he claims that the argument is not a proof of the existence of a necessary being. As Rowe sees it, an argument is a proof of its conclusion only if its premises are known to be true, and no human knows that the principle of sufficient reason is true. Hence, even if the argument is sound, it is not a proof of its conclusion. Rowe leaves open the possibility that it is reasonable for some people to believe that the argument's premises are true, in which case the argument would show the reasonableness of believing that a necessary being exists. If the second part of the argument made it reasonable to believe that such a necessary being has other divine attributes, then the theist might be entitled to claim that the argument shows the reasonableness of theistic belief. So Rowe invites the theist to explore the possibility that his cosmological argument shows that it is reasonable to believe in God, even though it perhaps fails to show that theism is true.

Richard Swinburne's teleological argument is part of a cumulative case he builds for theism (Swinburne, 1979). Other parts of the case involve arguments from consciousness and morality, from providence, from history and miracles, and from religious experience. Each part of the case is supposed to increase the probability of theism; the case as a whole is supposed to yield the conclusion that, on our total evidence, theism is more probable than not. The existence of order in the universe is supposed to increase significantly the probability of theism, even if it does not by itself render theism more probable than not.

In constructing his teleological argument, Swinburne appeals to general physical considerations rather than specifically biological order. There is a vast uniformity in the powers and liabilities of material objects that underlies the regularities of temporal succession described by the laws of nature. In addition, material objects are made of components of very few fundamental kinds. Either this order is an inexplicable brute fact or it has some explanation. Explanatory alternatives to theism such as the committee of minor deities suggested by David Hume seem to Swinburne less probable than theism, because theism leads us to expect one pattern of order throughout nature, while we would expect different patterns in different parts of the universe if its order were the product of a committee. So the alternatives are that the temporal order of the world has no explanation and that it is produced by God.

It is a consequence of Bayes's theorem that this order increases the probability of theism if and only if it is more probable if God exists than that God does not exist. Swinburne offers two reasons for thinking that the order of the universe is more probable on theism than on its negation. The first is that the order seems improbable in the absence of an explanation and so cries out for explanation in terms of a common source. The second is that there are reasons for God to make an orderly universe: One is that order is a necessary condition of beauty, and there is good reason for God to prefer beauty to ugliness

in creating; another is that order is a necessary condition of finite rational agents growing in knowledge and power, and there is some reason for God to make finite creatures with the opportunity to grow in knowledge and power.

The teleological argument plays a limited role in Swinburne's natural theology. Since it is an inductive argument, it does not prove the existence of God. Swinburne does not claim that by itself it shows that theism is more probable than not; nor does he claim that by itself it establishes the rational permissibility of belief in God.

Hence, only modest claims should be made on behalf of these three arguments for theism. Their authors are well aware that they do not prove the existence of God. However, they may show that belief in God is reasonable or contributes to a cumulative case for the rationality of theistic belief.

PROBLEMS OF EVIL

According to J. L. Mackie (1955), the existence of a God who is omniscient, omnipotent, and perfectly good is inconsistent with the existence of evil. If this is correct, we may infer that God does not exist from our knowledge that evil does exist. A solution to this logical problem of evil would be a proof that the existence of God is, after all, consistent with the existence of evil. One way to prove consistency would be to find a proposition that is consistent with the proposition that God exists and that, when conjoined with the proposition that God exists, entails that evil exists. This is the strategy employed in Plantinga's free-will defense against the logical problem of evil (Plantinga, 1974).

The intuitive idea on which the free-will defense rests is simple. Only genuinely free creatures are capable of producing moral good and moral evil. Of course, God could create a world without free creatures in it, but such a world would lack both moral good and moral evil. If God does create a world with free creatures in it, then it is partly up to them and not wholly up to God what balance of moral good and evil the world contains. The gift of creaturely freedom limits the power of an omnipotent God. According to Plantinga, it is possible that every free creature God could have created would produce at least some moral evil. Hence, it is possible that God could not have created a world containing moral good but no moral evil.

Consider the proposition that God could not have created a world containing moral good but no moral evil and yet creates a world containing moral good. The free-will defense claims that this proposition is consistent with the proposition that God is omniscient, omnipotent, and perfectly good. But these two propositions entail that moral evil exists and thus that evil exists. Hence, if the defense's consistency claim is true, the existence of a God who is omniscient, omnipotent, and perfectly good is consistent with the existence of evil. Therefore, the free-will defense is a successful solution of the logical problem of evil if its consistency claim is true. That claim certainly appears to be plausible.

Most philosophers who have studied the matter are prepared to grant that the existence of God is consistent with the existence of evil. The focus of discussion has shifted from the logical to the evidential problem of evil. The evils within our ken are evidence against the existence of God. The question is whether they make theism improbable or render theistic belief unwarranted or irrational.

William L. Rowe (1988) presents the evidential problem of evil in terms of two vivid examples of evil. Bambi is a fawn who is trapped in a forest fire and horribly burned; Bambi dies after several days of intense agony. Sue is a young girl who is raped and beaten by her mother's boyfriend; he then strangles her to death. According to Rowe, no good state of affairs we know of is such that an omnipotent, omniscient being's obtaining it would morally justify that being's permitting the suffering and death of Bambi or Sue. From this premise he infers that no good state of affairs is such that an omnipotent, omniscient being's obtaining it would morally justify that being in permitting the suffering and death of Bambi or Sue. If there were an omnipotent, omniscient, and morally perfect being, there would be some good state of affairs such that the being's obtaining it would morally justify the being's permitting the suffering and death of Bambi or Sue. Hence, it may be concluded that no omnipotent, omniscient, and morally perfect being exists.

The first step in this argument is an inductive inference from a sample, good states of affairs known to us, to a larger population, good states of affairs without qualification. So it is possible that no good state of affairs known to us morally justifies such evils but some good state of affairs unknown to us morally justifies them. But Rowe argues that the inference's premise gives him a reason to accept its conclusion. We are often justified in inferring from the known to the unknown. If I have encountered many pit bulls and all of them are vicious, I have a reason to believe all pit bulls are vicious.

William P. Alston (1991) challenges Rowe's inference. As he sees it, when we justifiably infer from the known to

the unknown, we typically have background knowledge to assure us that the known sample is likely to be representative of the wider population. We know, for example, that character traits are often breed-specific in dogs. According to Alston, we have no such knowledge of the population of good states of affairs because we have no way of anticipating what is in the class of good states of affairs unknown to us. He likens Rowe's reasoning to inferring, in 1850, from the fact that no one has yet voyaged to the moon that no one will ever do so.

The disagreement between Rowe and Alston illustrates the lack of a philosophical consensus on a solution to the evidential problem of evil. It is safe to predict continued debate about whether horrible evils such as the suffering and death of Bambi or Sue provide sufficient evidence to show that theistic belief is unjustified or unreasonable.

See also Alston, William P.; Bayes, Bayes' Theorem, Bayesian Approach to Philosophy of Science; Clarke, Samuel; Cosmological Argument for the Existence of God; Evil, The Problem of; Hume, David; Mackie, John Leslie; Malcolm, Norman; Modality, Philosophy and Metaphysics of; Modal Logic; Ontological Argument for the Existence of God; Philosophy of Religion; Plantinga, Alvin; Religious Experience; Teleological Argument for the Existence of God.

Bibliography

Alston, W. P. "The Inductive Argument from Evil and the Human Cognitive Condition." In *Philosophical Perspectives,* Vol. 5: *Philosophy of Religion,* edited by James E. Tomberlin. Atascadero, CA: Ridgeview, 1991.

Mackie, J. L. "Evil and Omnipotence." *Mind* 64 (1955): 200–212.

Plantinga, A. *The Nature of Necessity.* Oxford: Clarendon Press, 1974.

Rowe, W. L. *The Cosmological Argument.* Princeton, NJ: Princeton University Press, 1975.

Rowe, W. L. "Evil and Theodicy." *Philosophical Topics* 16 (2) (1988): 119–132.

Swinburne, R. *The Existence of God.* Oxford: Clarendon Press, 1979.

Phillip L. Quinn (1996)

THEMISTIUS
(c. 317–c. 385 CE)

Themistius is one of the principal Greek commentators on Aristotle. He was born at Byzantium, the son of a philosopher (Eugenius), and received a traditional education in Greek culture at various locations. In his twenties Themistius established a philosophical school at Constantinople (as Byzantium had by then become), and prepared the paraphrases on several Aristotelian works that represent his main contribution to the ancient philosophical tradition. After about 350 CE he became involved in the political life of the eastern Empire, and served several emperors as an ambassador, administrator, and adviser. This phase of his career is richly documented in his orations, some of which reflect his philosophical interests.

Themistius cannot be easily labeled by his philosophical affiliation. His extant paraphrases of Aristotle's *De anima*, *De caelo*, *Metaphysics* Book 12, *Physics*, and *Posterior Analytics* follow the Aristotelian text closely and are designed to facilitate study. He was clearly influenced by the work of the great Peripatetic commentator Alexander of Aphrodisias. However, at times Themistius reveals some knowledge of the Platonic tradition, notably in his response to Aristotle's account of the intellect in *De anima* Book 3, chapter 5. He is most safely described as a philosophical scholar who absorbed the Platonic tradition without allowing it to dominate his interpretations, as it did in the case of later commentators, notably Simplicius and Philoponus.

Themistius was respected by Aristotelian commentators in later antiquity, in the Arabic, Hebrew, and western medieval Latin tradition, as well as during the Renaissance. Some of his texts are in fact extant only in Arabic. His interpretation of the active intellect was suggestive enough to allow for the notion of the immortality of the individual soul, and, as such, was welcome within the Christian tradition.

See also Aristotle.

Bibliography

Todd, Robert B. "Themistius." *Catalogus Translationum et Commentariorum* 8 (2003): 57–102 (bibliography at 68–72).

Todd, Robert B., trans. *Themistius on Aristotle on the Soul.* Ithaca, NY: Cornell University Press, 1996.

Todd, Robert B., trans. *Themistius on Aristotle Physics 4*. Ithaca, NY: Cornell University Press, 2003.

Vanderspoel, J. *Themistius and the Imperial Court: Oratory, Civic Duty and Paideia from Constantine to Theodosius*. Ann Arbor: University of Michigan Press, 1995.

Robert B. Todd (2005)

THEODICY

See *Evil, The Problem of; Leibniz, Gottfried Wilhelm*

THEODORIC OF CHARTRES

Theodoric of Chartres (or Thierry of Chartres) was a twelfth-century philosopher and younger brother of Bernard of Chartres. He appears first as a master in 1121, when he spoke in support of Peter Abelard at the latter's trial for heresy at Soissons. In the 1130s he was teaching the arts in Paris, and in 1142 he became chancellor at Chartres. He attended the trial of Gilbert of Poitiers at Rheims in 1148 and shortly afterward became a monk. The date of his death is unknown.

Theodoric's rhetorical teaching survives in a commentary on Cicero's *De Inventione.* Three versions of his exposition of Boethius's *De Trinitate* and a fragmentary exposition of Boethius's *De Hebdomadibus* are also extant, as is a commentary on the beginning of the book of Genesis (the *De Sex Dierum Operibus*). In the last-named work Theodoric's Platonizing cosmology and his mathematical bent found their expression. In his *Heptateuch,* a bulky collection of the sources for each of the seven liberal arts, Theodoric revealed his fidelity to the ancients. Grammar was represented by the works of Donatus and Priscian, rhetoric by Cicero, astronomy by Ptolemy; but the place of honor went to Boethius for his writings on music, arithmetic, geometry, and, especially, dialectic. Theodoric reproduced Boethius's translations and commentaries on the whole of Aristotle's *Organon,* with the exception of the *Posterior Analytics.*

Theodoric regarded the arts as the indispensable instrument of philosophy, which consisted of physics, mathematics, and theology. He based his Trinitarian speculation upon arithmetic, applying the Pythagorean-Platonic dialectic of unity-multiplicity to St. Augustine's dictum that the Father is unity, the Son equality, and the Spirit the agreement of unity and equality. Unity can only engender its equality; both are one substance but have different properties and are called persons by the theologians. Theodoric's argument emphasized the unity of the Trinity but made difficult a numerical distinction between the divine persons. The dialectic of unity-multiplicity was perhaps more appropriately used to explain the relationship of the Creator to creation. Unity is God and is immutable and eternal; the principle of multiplicity is the domain of creation. Unity is the *forma essendi* of creatures, their unique and entire being, totally and essentially omnipresent. Things are not pantheistically identified with the One; multiplicity is distinct from, and subordinate to, unity. The divine unity in an ineffable way absorbs the forms of all beings in itself, but only images of these forms are joined to matter. Theodoric's thought here moves close to his brother's theory of native forms.

Although Theodoric stressed the universal causality and omnipresence of the Creator, he presented creation as an ordered system of secondary causes. Matter was created by God from nothing, but the fashioning of the world out of the four elements occurred by the action of the circular motion of heaven and of the diffusion of heat in the underlying elements. The four elements of matter (which Genesis collectively designates by the names of heaven and earth) arranged themselves into four concentric spheres. The heaven of air and fire enveloped the water and Earth and, being supremely light, tended to move by turning about. Fire became ardent and illumined the air and heated the water, vaporizing it to reveal islands on Earth and to incubate life in the water and on land. The mechanistic character of this explanation is supplemented by a recognition of the role of spirit, which fills and animates the world. Through the "seminal reasons" introduced by God into creation, nature is capable of its own continuation after the completion of the work of six days. Theodoric's doctrine of creation represents an adventurous application of the teachings of the Platonic *Timaeus* to the biblical account.

Theodoric was a bold speculator, molded by and helping to mold the Platonic tradition of Latin Christendom. He seems also to have been the first medieval schoolman to have commented on the recently rediscovered *Prior Analytics* and *Sophistic Refutations* of Aristotle. Moreover, it was to him that Hermann of Carinthia sent his translation of Ptolemy's *Planisphere,* just as Bernard of Tours dedicated his *De Mundi Universitate* to Theodoric. Other disciples and admirers included Clarembald of Arras and John of Salisbury and, in the fifteenth century, Nicholas of Cusa.

See also Abelard, Peter; Aristotle; Augustine, St.; Bernard of Chartres; Bernard of Tours; Boethius, Anicius Manlius Severinus; Cicero, Marcus Tullius; Gilbert of Poitiers; John of Salisbury; Matter; Medieval Philosophy; Nicholas of Cusa; Platonism and the Platonic Tradition.

Bibliography

Commentary and text of works by Theodoric are *De Septem Diebus et Sex Operum Distinctionibus,* edited by N. M. Haring, in "The Creation and Creator of the World According to Thierry of Chartres and Clarenbaldus of Arras," in *Archives d'histoire doctrinale et littéraire du moyen âge* 22 (1955): 137–216; and the following additional articles by Haring: "The Lectures of Thierry of Chartres on Boethius' *De Trinitate,*" in *Archives …* 25 (1958): 113–226; "A Commentary on Boethius' *De Trinitate* by Thierry of Chartres," in *Archives …* 23 (1956): 257–325; and "Two Commentaries on Boethius (*De Trinitate* and *De Hebdomadibus*) by Thierry of Chartres," in *Archives …* 27 (1960): 65–136.

Extracts of Theodoric's commentary on Cicero's *De Inventione* were edited by P. Thomas, "Un commentaire du moyen âge sur la rhétorique de Cicéron," in *Mélanges Charles Graux* (Paris, 1884), pp. 41–45; by W. H. D. Suringar, *Historia Critica Scholiastarum Latinorum,* Vol. I (Leiden, 1834), pp. 213–252; and by R. Ellis, "Petronianum," in *Journal of Philology* 9 (1880): 61.

On the above extracts see P. Delhaye, "L'enseignement de la philosophie morale au XII^e siècle," in *Mediaeval Studies* 11 (1949): 77–99, Appendix C, 97–99; and F. Masai, "Extraits du commentaire de Thierry de Chartres au *De Inventione* de Cicéron," in *Scriptorium* 5 (1951): 117–120, 308–309. There is an edition of the prologue to the *Heptateuch* by E. Jeauneau in *Mediaeval Studies* 16 (1954): 171–175; there is a summary of the contents of the prologue in A. Clerval, *Les écoles de Chartres* (Paris, 1895), pp. 220–240.

A biography of Theodoric is A. Vernet, "Une épitaphe inédite de Thierry de Chartres," in *Receuil de travaux offerts à M. Clovis Brunel,* 2 vols. (Paris, 1955), Vol. II, pp. 660–667.

Theodoric's doctrines are discussed in Étienne Gilson, *History of Christian Philosophy in the Middle Ages* (London, 1955), pp. 145–148; J. M. Parent, *La doctrine de la création dans l'école de Chartres* (Paris and Ottawa, 1958), passim; and W. Jansen, *Der Kommentar des Clarenbaldus von Arras zu Boethius "De Trinitate"* (Breslau, 1926), passim.

David Luscombe (1967)

THEOLOGY

See *God, Concepts of; Philosophy of Religion; Philosophy of Religion, History of; Philosophy of Religion, Problems of; Physicotheology; Religion.*

THEOPHRASTUS

(372/1–282/1 BCE)

Born in Eresus on the Aegean island of Lesbos, Theophrastus moved to Athens, studying under Plato briefly and then Aristotle, soon becoming the latter's colleague. In 322/1 BCE he succeeded Aristotle as head of the Lyceum. The picture arising from his extant works is that of a conscientious scholar and researcher, with a marked emphasis on natural philosophy. His place as Aristotle's first successor has for a long time created the impression of a dogmatic and docile pupil, but a comparison with his master is invidious. A more acceptable perspective, established in antiquity (e.g., frag. 72A), is to view his work as trading on the presence of the Aristotelian corpus, while expanding and adjusting even fundamental aspects of the system where required. Exciting recent finds in Arabic and Syriac sources and the new 1992 edition of fragments (edited by Fortenbaugh et al.) have given us a better idea of his learning, independence of thought, and influence (all references to fragments are to 1992). Diogenes Laertius lists some two hundred titles in the Theophrastan corpus (D.L. 5.42–50), and only a fraction of these works survives. Yet what survives is sufficient to reveal him as a clever and productive philosopher and scientist with wide-ranging interests.

LANGUAGE AND LOGIC

Theophrastus made contributions to the theory of the syllogism (e.g., on the relation between the second and third figures), and he revised Aristotle's modal logic, suggesting that the conclusion has the same modality as the weaker premise, not the major premise (a weakest-link principle). He also proposed revising the system of dialectical predication, subsuming the four predicables under definition, perhaps to create "a single universal method" (frag. 124A–B), and he provided us with a definition of the dialectical *topos*" (not found in Aristotle) as an argumentative strategy or principle (frag. 122B). He is said to have introduced a doctrine of hypothetical syllogisms, possibly in collaboration with Eudemus of Rhodes (350–290 BCE). True to his reputation as a good speaker, his comments on language advance grammar and style, and he makes a notable effort to use appropriate language in each field.

PHYSICS AND SCIENCE

Of Theophrastus's work in the sciences, we still have two major works on plant taxonomy (*Enquiry into Plants*) and explanations for plants (*De causis plantarum* [Causes of plants]), famously influential on Carl Linnaeus (1707–1778); nine short tracts on the inanimate (e.g., winds, stones) and physiology (e.g., sweat, dizziness, fatigue); and fragments pertaining to meteorology, biology, epistemology, and psychology.

While maintaining an empiricist outlook, Theophrastus consistently dealt with issues of a fundamental nature (frags. 142–143). He added significantly to

the scientific methodology developed by Aristotle. The latter sought to describe a system of argumentation, providing the first attempt at a second-order language of research. This early scientific methodology was a mix of logical principles and rhetorical habits, combining forms of presentation and manipulation with rules of consistency and rigor. Theophrastus also believed in an appropriate method (*oikeios tropos*) for each field of research (*Metaphysics* 9a11).

In line with Peripatetic doctrine, Theophrastus attributed teleological order to nature, "which does nothing in vain" (e.g., *De causis plantarum* 1.1.1, 2.1.1, 4.4.2), though he allowed for exceptions to this general rule (see the next section). Another feature of Theophrastus's approach is his readiness to allow for multiple explanations for physical phenomena (found again in Epicurus in different form), which may signal a growing awareness that a universal theory is unattainable. A correct explanation should give a reason for puzzling facts (*De ventis* [On winds] 59), be coherent (*De odoribus* [On smells] 64), and harmonize with descriptions of the facts (*De causis plantarum* 1.1, 1.21.4). His views thus adumbrate a principle of falsification.

We can reconstruct significant aspects of Theophrastus's epistemology and psychology on the basis of mostly late sources, some going back to his own work *On the Soul*. His empiricist approach is evident in his claims that perception is crucial for knowledge (frags. 301B and 143), and that exceptional clarity (*to enarges*) is a criterion of truth shared by sensation and intellect (frag. 301A). Regarding Aristotle's *On the Soul*, he asked pertinent questions about the process of sensation (e.g., How does the sense organ become like the object? The answer is that the organ receives a universal form). His concerns over Aristotle's notoriously difficult account of intellect (*nous*) (*On the Soul* 3.5) are paraphrased in Themistius (frag. 307A) and the neo-Platonist Priscian (frag. 307B–D). He asked after the nature of intellect in relation to matter (both seem to be "nothing, but potentially all things"), and puzzled over how intellect and object might affect each other.

METAPHYSICS

Theophrastus's extant short tract on metaphysics, now considered to be a complete work, can be seen as a critical evaluation of Aristotle (and others), in particular, on first principles and the unmoved mover. He presented a range of connected problems that he did not always clearly resolve. (This is typical of his *aporetic* [doubt-prone] style, in this case perhaps because *Metaphysics* is an early work or because it is didactic or both.) He also showed himself to be preoccupied with the boundaries of explanation. For instance, he raised questions about what we can assume as fundamental principles and how many there are, and he looked at possible options (one, more than one) and their problems: A universe with one principle cannot be diverse, but a universe with two or more principles might lack coherence. His discussion of what kind of principles he envisages presents two options: Principles are either the ultimate sources of things (a foundationalist position) or else general laws governing everything (in which case, principles are rules of practice). He restricted the number of principles, and the scope of their influence in the physical realm. This allowed him to keep certain accidental occurrences (e.g., thunder, but also evil) outside the range of events with a final cause. Theophrastus's idea of limited teleology and purposiveness (Theophrastus, *Metaphysics* 7a19–b9, 10a21–23) is confirmed in Arabic sources. In his botanical works, however, he tried to accommodate anomalies within the Aristotelian framework (*De causis plantarum* 5). Obviously, Theophrastus's position complicated the Aristotelian position that "nature does nothing in vain."

ETHICS

Our material for Theophrastus's ethical views is rather uneven, ranging from comments on virtue to friendship and natural kinship between animals and humans. Of interest are the excerpts in Porphyry (c. 300 CE), which discuss forms of sacrifice and reasons for vegetarianism (frags. 531, 584). A lost work on friendship was quite influential, and he seems to have come up with new ideas on emotions (frags. 438–448). His collection of character sketches (*Characters*), hugely popular in the eighteenth century, presents psychological profiles in the style of contemporary comedy depicting men with serious character flaws. These profiles perhaps fit into the general framework of Aristotle's ethics. Aristotle's analysis of types (*Nicomachean Ethics* 2) and his doctrine of virtue as a mean or middle between vices help to understand these flaws as concrete examples of Aristotle's more abstract model. Some fragments support such a connection (*Characters*, p. 19). Theophrastus differs from Aristotle at least in focusing on faults and in adopting an anecdotal style of moral instruction.

HISTORY OF PHILOSOPHY (DOXOGRAPHY)

Theophrastus's critical evaluations of earlier philosophers (pre-Socratics, Platonists) are extant in short passages and the treatise *De sensibus* (On sensation), which is believed to be part of a larger work, perhaps his *Physical Opinions* (D.L. 5.48) or his *Reply to Physical Philosophers* (D.L. 5.46, frag. 241A). These comments represent important aspects of his methodology and his influence on the early history of philosophy. Not only do they show a greater awareness of the philosophical enterprise as a continuous discourse by their methodical preservation and assessment of past achievements, but they also illustrate, through his criticisms and convenient organization of materials, the reason for his impact on following ages. Theophrastus's work in this area contributed to the consolidation and preservation of philosophical debates in the Hellenistic schools.

See also Aristotle; Diogenes Laertius; Ethics, History of; Hellenistic Thought; History and Historiography of Philosophy; Logic, History of Metaphysics, History of; Neoplatonism; Peripatetics; Plato; Platonism and the Platonic Tradition; Porphyry; Pre-Socratic Philosophy; Teleology; Themistius; Virtue and Vice.

Bibliography

WORKS BY THEOPHRASTUS

Metaphysics. Translated by W. D. Ross and F. H. Fobes. Oxford, U.K.: Clarendon Press, 1929.

De causis plantarum. 3 vols. Translated by Benedict Einarson and George K. K. Link. Cambridge, MA: Harvard University Press, 1976–1990.

Theophrastus of Eresus: Sources for His Life, Writings, Thought, and Influence. 2 vols, edited and translated by William W. Fortenbaugh, P. M. Huby, R. W. Sharples, and D. Gutas. Leiden, Netherlands: Brill, 1992. Text and translation of sources up to the fourteenth century. References to fragments refer to this work.

Theophrastus of Eresus: Sources on Biology (Texts 328–435). Vol. 5. Commentary by R. W. Sharples. Leiden, Netherlands: Brill, 1995.

Theophrastus of Eresus: Sources on Physics (Texts 137–223). Vol. 3.1. Commentary by R. W. Sharples. Leiden, Netherlands: Brill, 1998.

Theophrastus of Eresus: Sources on Psychology (Texts 265–327). Vol. 4. Commentary by Pamela M. Huby. Leiden, Netherlands: Brill, 1999.

Characters. Translated by Jeffrey Rusten. Together with mimes by Herodas and Sophron and other mime fragments. 3rd ed. Cambridge, MA: Harvard University Press, 2002.

On Sweat; On Dizziness; and On Fatigue, edited by William W. Fortenbaugh, Robert W. Sharples, and Michael G. Sollenberger. Leiden, Netherlands: Brill, 2003.

WORKS ON THEOPHRASTUS

Baltussen, H. *Theophrastus against the Presocratics and Plato: Peripatetic Dialectic in the "De sensibus."* Leiden, Netherlands: Brill, 2000.

Diogenes Laertius. *Lives of Eminent Philosophers*. Bk. 5. Translated by R. D. Hicks. Cambridge, MA: Loeb Classical Library, 1925.

Diogenes Laertius. *The Vita Theophrasti*. Bk 5.36–57, edited by M.G. Sollenberger in *Rutgers Studies in the Classical Humanities* vol. 2, 1985, pp. 1–62.

Fortenbaugh, William W., and Georg Wöhrle, eds. *On the Opuscula of Theophrastus*. Stuttgart, Germany: F. Steiner Verlag, 2002.

Han Baltussen (2005)

THEORIES AND THEORETICAL TERMS

In mathematical logic, a theory is the deductive closure of a set of axioms (that is, the set of all propositions deducible from a set of axioms). In the early- and mid-twentieth century, philosophers of science, under the influence of Bertrand Russell's work in philosophy of language and philosophy of mathematics, attempted rationally to reconstruct scientific knowledge by representing scientific theories with the powerful conceptual tools provided by the theory of formal languages.

THE SYNTACTIC VIEW OF THEORIES

The *syntactic* view of theories (also called the received view) was developed by Rudolf Carnap, Ernest Nagel, Hans Reichenbach, and other logical empiricists. Like David Hume, these philosophers thought that insofar as scientific theories accurately describe the world, they cannot be known a priori, but they also recognized that some elements of our theoretical knowledge seem to be independent of the empirical facts. For example, Isaac Newton's second law states that the force on a body is proportional to the rate of change of its momentum, where the constant of proportionality is the inertial mass. This law cannot be tested in an experiment, because it is part of what gives meaning to the concepts employed to describe the phenomena. Hence, the logical empiricists argued, physical theories can be split into a part that expresses definitions of basic concepts and relations among them, and a part that relates to the world. The former part also includes the purely mathematical axioms of the theory and, trivially, all the logical truths expressible in the language of the theory. This part of the theory is a priori knowledge and concerns matters purely of convention. The factual content of the theory is confined to the

latter part, and hence the fundamental empiricist principle that the physical world cannot be known by pure reason is satisfied.

Empiricists argue that meaning must originate in experience, and the logical empiricists used this criterion to criticize speculative metaphysics and to place limits on legitimate scientific theorizing. However, we can have no direct experience of theoretical entities such as neutrinos or theoretical properties such as spin. How can theoretical terms be meaningful? The logical empiricists tried to use logic to show how the theoretical language of science is related to the everyday language used to describe the observable world. They were motivated by the verification principle, according to which a (nontautological) statement is meaningful if and only if it can be verified in the immediacy of experience, and the verifiability theory of meaning, according to which the meaning of particular terms (other than logical constants) is either directly given in experience or consists in how those terms relate to what is directly given in experience.

The idea is that a physical theory will have a canonical formulation satisfying the following conditions:

1. L is a first-order language with identity, and K is a calculus defined for L.

2. The nonlogical terms of L can be partitioned into two disjoint sets, one of which contains the *observation* terms, V_O, and the other of which contains the *theoretical* terms, V_T.

3. There are two sublanguages of L, and corresponding restrictions of K, such that one (L_O) contains no V_T terms and the other (L_T) no V_O terms. These sublanguages together do not exhaust L, of course, since L also contains *mixed sentences*.

4. The observational language L_O is given an *interpretation* in the domain of concrete observable entities, processes, events, and their properties. An interpretation of language L (in the model-theoretic sense used here) attributes a reference to each of the nonlogical terms in L at the metalinguistic level. If the axioms of a theory are true under some interpretation, then that interpretation is a *model* for the theory.

5. The theoretical terms of L are given a *partial interpretation* by means of two kinds of postulates: *theoretical postulates*, which define internal relations among the V_T terms and do not feature V_O terms, and *correspondence rules* or *bridge principles*, which feature mixed sentences and relate the V_T and V_O terms. (These correspondence rules are also known as "dictionaries," "operational definitions," and "coordinative definitions," depending on the author. All these terms designate a set of rules connecting theoretical terms to observable states of affairs.)

The theoretical postulates are the axioms of the theory, and the purely theoretical part of the theory is the deductive closure of these axioms under calculus K. The theory as a whole, TC, is the conjunction of T and C, where T is the conjunction of the theoretical postulates and C is the conjunction of the correspondence rules.

The logical empiricists soon abandoned the attempt to give language L_O an interpretation in terms of immediate experience. It was decided instead that it is just as good to opt for a physicalist language, that is, one that refers only to physical objects, properties, and events (Friedman 1999). Initially, it was required that the theoretical terms of L be given *explicit definitions* (this was Carnap's original goal, but he had abandoned it by the time of his 1936–1937 paper). An example of such a definition of a theoretical term V_T is the following:

$$\forall x(V_T(x) \leftrightarrow [Px \rightarrow Qx]),$$

where P is some preparation of an apparatus (known as a test condition) and Q is some observable response of the apparatus (so P and Q are describable in V_O terms alone). For example, an explicit definition of temperature can be given as follows: Any object x has temperature t if and only if when x is put in contact with a thermometer, it gives a reading of t. If theoretical terms could be so defined, this would show that they are convenient devices, can in principle be eliminated, and need not be regarded as referring to anything in the world (this view is called *semantic instrumentalism*).

It was soon realized that explicit definition of theoretical terms is highly problematic. Perhaps the most serious difficulty is that, according to this definition, if we interpret the conditional in the square brackets as material implication, theoretical terms are trivially applicable when the test conditions do not obtain (because if the antecedent is false, the material conditional is always true). If, in contrast, we interpret the conditional as strict implication, then the theoretical term is applicable only when the test conditions obtain. In other words, either everything never put in contact with a thermometer has temperature t (under material implication), or only those things put in contact with a thermometer are candidates for having temperature t (under strict implication). This is clearly inadequate, since scientists use the language of temperature as if things have a temperature whether anybody chooses to measure it or not.

The natural way to solve this problem is to allow subjunctive assertion in explicit definitions. That is, we define the temperature of object *x* in terms of what would happen if *x* were put in contact with a thermometer. Here temperature is understood as a dispositional property. Unfortunately, this raises further problems. First, unactualized dispositions, such as the fragility of a glass that is never damaged, seem to be unobservable properties, and they give rise to statements whose truth conditions are problematic for empiricists, namely counterfactual conditionals, such as "If the glass had been dropped, it would have broken," where the antecedent is false. Dispositions are also modal, that is, they involve possibility and necessity, and empiricists since Hume have disavowed objective modality. Like laws of nature and causation, dispositions are problematic for empiricists. Second, no one has ever provided explicit definitions for terms like "space-time curvature," "spin," and "electron," whether dispositional or not, and there are no grounds for thinking that they could be.

However, advocates of the syntactic view did not abandon the attempt to anchor theoretical terms to the observable world. This is the point of the correspondence rules that connect the theoretical terms with the observational ones and so ensure their cognitive meaningfulness. They do not define the former in terms of the latter; rather, together with the theoretical postulates, they offer a partial interpretation for them. The correspondence rules are also intended to specify procedures for applying the theory to the phenomena. Theoretical concepts such as those of vital forces and entelechies were criticized by the logical empiricists because their advocates failed to express them in terms of precise, testable laws.

According to the view developed so far, *TC* is fully interpreted only with respect to its V_O terms, which refer to ordinary physical objects (such as ammeters, thermometers, and the like) and their states; the V_T terms are only partially interpreted. The models of *TC* comprise all the possible interpretations of *TC* in which the V_O terms have their normal meanings and under which *TC* is true. The problem for the advocate of the syntactic approach is that there will be many models in general, so there is no unique interpretation for the theory as a whole. Hence, it would seem to make no sense to talk of *TC* being true or false of the world. Hempel (1963) and Carnap (1939) solved this problem by stipulating that *TC* is to be given an intended interpretation; theoretical terms are interpreted as (putatively) referring to the entities, processes, events, and properties appropriate to their normal meanings in scientific (and everyday) use.

Thus, if the meaning of the term "electron," say, derives from the picture of electrons as tiny billiard balls or classical point particles, this picture is important in determining what the theory of electrons refers to. Once the explicit-definition project is abandoned, one must accept that the meanings of theoretical statements lacking testable consequences are nonetheless important in determining the referents of the V_T terms. As Suppe put it, "When I give a semantic interpretation to *TC*, I am doing so relative to the meanings I already attach to the terms in the scientific metalanguage. In asserting *TC* so interpreted, I am committing myself to the meaning of 'electron' and so on, being such that electrons have those observable manifestations specified by *TC*" (1977, p. 92).

This version of the syntactic view is committed to the idea that theoretical terms have excess or surplus meaning over and above the meaning given by the partial interpretation in terms of what can be observed. Herbert Feigl explicitly recognized this in 1950 and was thus led to argue for the view that theoretical terms genuinely refer to unobservable entities (*scientific realism*).

Perhaps the most widespread criticism of the syntactic view is that it relies on the distinction between observational terms and theoretical terms. This distinction is supposed to correspond to a difference in how language works. Observational terms are more or less ostensibly defined and directly refer to observable features of the world, while theoretical terms are indirectly defined and refer to unobservable features of the world. Examples of the former presumably include "red," "pointer," "heavier than"; examples of the latter would include "electron," "charge density," "atom." Putnam (1962/1975) and many others have argued that there is no objective line to be drawn between observational and theoretical language, and that all language depends on theory to a degree. Moreover, eliminating theoretical terms, even if it were possible, would not eliminate talk of the unobservable, because it is possible to talk about the unobservable using V_O terms only, for example, by saying that there are particles that are too small to see. (William Demopoulos has argued that this criticism is irrelevant to the project of offering a rational reconstruction of theories.)

Whether or not the distinction between observational and theoretical terms can be drawn in a nonarbitrary way, the syntactic view also faces criticism concerning the correspondence rules. These rules were supposed to have three functions: (a) to generate (together with the theoretical postulates) a partial interpretation of theoretical terms, (b) to give the theoretical terms cognitive significance by connecting them with

what can be observed, (c) to specify how the theory is related to the phenomena. There are several problems concerning (c). First, if the correspondence rules are part of the theory, then whenever a new experimental technique is developed in the domain of the theory and the correspondence rules change to incorporate the new connections between theoretical terms and reality, the theory will change. This is counterintuitive. Another problem, raised by Suppe (1977), is that there are probably an indefinite number of ways of applying a theory, and so there ought to be an indefinite number of correspondence rules, but the formulation of the syntactic view requires that there be only finitely many. Furthermore, theories are often applied to phenomena by means of other theories used to establish a causal connection between the states of affairs described by the theory and the behavior of some measuring apparatus. For example, theories of optics are needed to link the occurrences of line spectra with changes in the energy states of electrons. The correspondence rules in this case will incorporate principles of optics to offer mechanisms and explanations for the behavior of measuring devices. Suppe concludes that correspondence rules are not an integral part of the theory as such but rather are auxiliary assumptions about how the theory is to be applied.

Nancy Cartwright (1983, 1989) and many others have argued that the syntactic view is misleading about how scientific theories are applied, because auxiliary assumptions about background conditions are rarely, if ever, sufficient for deriving concrete experimental predictions from a theory. Rather, these authors argue, the connections between abstract theory and concrete experiment are complex, nondeductive, and involve the use of many theories, models, and assumptions that are not yet part of the original theory.

THE SEMANTIC APPROACH TO SCIENTIFIC THEORIES

According to the semantic or model-theoretic view of theories, theories are better thought of as families of models rather than as partially interpreted axiomatic systems. Theories are "extralinguistic entities which may be described or characterized by a number of different linguistic formulations" (Suppe, p. 221).

To understand the semantic approach, first consider a modification of the syntactic view due to Ernest Nagel (1961) and Mary Hesse (1966). These authors insist that there are always models for a theory, whether true of the world or not. According to Nagel, "An interpretation or model for the abstract calculus … supplies some flesh for the skeletal structure in terms of more or less familiar conceptual or visualizable materials" (p. 90). He is here thinking of models like the billiard-ball model of a gas. This model supplies an iconic representation for the theory of gases (we interpret "gas molecule" as referring to a billiard ball and then picture the gas accordingly). This concrete picture allows the physicist to visualize the system and may also provide heuristic guidance for the future development of the theory. Hesse does not restrict models of theories to those that feature "familiar conceptual or visualizable materials," like the billiard-ball model. She regards mathematical structures specified by the formalism of a theory as a paradigm type of model. Indeed, she goes so far as to say that a model can be "any system, whether buildable, picturable, imaginable, or none of these, which has the characteristic of making a theory predictive" (1966, p. 19). In this she seems right in that many theories of contemporary physics, such as quantum mechanics, do not admit of models consisting of familiar or visualizable materials.

The origins of the semantic approach can be traced to Evert Beth and Patrick Suppes. The latter coined the slogan "[T]he correct tool for philosophy of science is mathematics, not meta-mathematics" (see for example, 1961/1969) and thought of theories as set-theoretic structures. Bas van Fraassen (1980, 1989) further elaborated and generalized Beth's approach: Theories are presented by specifying a class of state spaces with laws of coexistence (synchronic constraints) and laws of succession (diachronic constraints), which together specify the allowable trajectories for systems whose states are represented by parameters located in the state space. Examples of laws of coexistence are Boyle's gas law and the Pauli exclusion principle for energy states of electrons and other fermions; examples of laws of succession include the Schrödinger wave equation in quantum mechanics and Hamilton's equations of motion in classical mechanics.

An advantage claimed for the semantic approach is that it is closer to the practice of science, since scientists do not deduce empirical results directly from theories, but rather use theories in conjunction with models that apply to the system in question. Much of the practice of science concerns the development of new models to extend the domain of application of well-known theories. According to Ron Giere (1988) and Bas van Fraassen (1980, 1989), theories are partly linguistic entities insofar as they include various theoretical hypotheses linking models with systems in the real world, but are nonlinguistic insofar as they essentially involve populations of

models. Such models "are the means by which scientists represent the world" (Giere, p. 80). Properly speaking, then, a theory comprises the models it uses and hypotheses that assert a similarity between a real system and some aspects of a model (other aspects are left out because of idealization and approximation).

Giere leaves this relation of similarity unanalyzed. For van Fraassen, the relation between theories and the world is one of isomorphism: "To present a theory is to specify a family of structures, its *models*; and secondly, to specify certain parts of those models (the *empirical substructures*) as candidates for the direct representation of observable phenomena. The structures which can be described in experimental and measurement reports we can call *appearances*: the theory is empirically adequate if it has some model such that all the appearances are isomorphic to empirical substructures of that model" (1980, p. 64). The appearances are the representations of the phenomena, in other words, mathematical models of the data (Suppes 1962).

THE REFERENCE OF THEORETICAL TERMS

Theoretical terms that allegedly refer to unobservable entities cannot be defined ostensively. If the reference of theoretical terms, such as "electron," is fixed by the relevant scientific theory, the sense of such a term fixes its reference (this is called a *descriptivist theory of reference*). Thomas Kuhn (1962) argued that the sense of many scientific terms—terms such as "atom," "electron," "species," and "mass"—has changed considerably during the course of scientific revolutions. If the references of theoretical terms are fixed by the whole of the theories in which they feature, then any change in the latter will result in a change in the former.

In response, Hilary Putnam, in "Explanation and Reference" (1975), advocated a radically different account of the meaning of theoretical terms. He pointed out that most people have no idea how to link many terms with their references but nonetheless successfully refer to particular kinds of things using them. They do so by deferring to experts. For example, most people successfully use the word "platinum" even though lack an explicit definition and have no way to distinguish samples. Only a few experts have detailed criteria.

Putnam advocates a *causal theory of reference* for natural-kind terms. According to this theory, the referent of "water," for example, is whatever causes the experiences that give rise to talk of water. Reference is fixed not by the description associated with a term, but by the cause of the use of the term. This allows for continuity of reference across theory changes. Even though theories about electrons have changed, and hence the meaning of the term "electron" has changed, the term, Putnam argues, has always referred to whatever causes the phenomena that prompted its introduction, such as the conduction of electricity by metals.

THE RAMSEY-SENTENCE APPROACH TO THEORIES

Frank Ramsey argued that the content of a physical theory is captured in its *Ramsey sentence*, the result of taking an axiomatization of the form described above and replacing all the theoretical terms with variables and existentially quantifying over the latter. For example, $\emptyset(O_1, \ldots, O_n; T_1, \ldots, T_m)$ has the Ramsey sentence $\exists t_1, \ldots, \exists t_m \emptyset(O_1, \ldots, O_n; t_1, \ldots, t_m)$. In effect, the Ramsey sentence of a theory is a statement in higher-order logic that says that the theory has a model consistent with a fixed interpretation of the observational terms. Ramsey thus treated theoretical terms as disguised definite descriptions. The Ramsey sentence and the original theory both imply the same observational sentences involving *O*-terms, and hence the factual content of the latter is captured by the former. David Lewis (1970) used Ramsey's method to show how new theoretical terms could be defined in terms of antecedently understood theoretical terms, rather than observational terms.

The Ramsey-sentence approach to theories has been thought to show that scientific knowledge of the unobservable theoretical world is purely structural (Worrall 1989). This raises technical problems discussed in Demopoulos (forthcoming), Demopoulos and Friedman (1985), and Psillos (2000).

See also Philosophy of Science, Problems of; Scientific Realism.

Bibliography

Beth, Evert. "Towards an Up-to-Date Philosophy of the Natural Sciences." *Methodos* 1 (1949): 178–185.

Carnap, Rudolf. *Foundations of Logic and Mathematics*. Chicago; University of Chicago Press, 1939.

Carnap, Rudolf. "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–471; 4 (1937): 1–40.

Cartwright, Nancy. *How the Laws of Physics Lie*. Oxford, U.K.: Oxford University Press, 1983.

Cartwright, Nancy. *Nature's Capacities and Their Measurement*. Oxford, U.K.: Oxford University Press, 1989.

Demopoulos, William. "Carnap's Philosophy of Science." In *The Cambridge Companion to Carnap*, edited by R. Creath

and Michael Friedman. Cambridge, U.K.: Cambridge University Press, forthcoming.
Demopoulos, William, and Michael Friedman. "Critical Notice: Bertrand Russell's *The Analysis of Matter: Its Historical Context and Contemporary Interest*." *Philosophy of Science* 52 (1985): 621–639.
Feigl, Herbert. "Existential Hypotheses: Realistic versus Phenomenalistic Interpretations." *Philosophy of Science* 17 (1950): 35–62.
Friedman, Michael. *Reconsidering Logical Positivism*. Cambridge, U.K.: Cambridge University Press, 1999.
Giere, Ronald N. *Explaining Science*. Chicago: University of Chicago Press, 1988.
Hempel, Carl. "Implications of Carnap's Work for Philosophy of Science." In *The Philosophy of Rudolf Carnap*, edited by Paul Schilpp. LaSalle, IL: Open Court, 1963.
Hesse, Mary. *Models and Analogies in Science*. Oxford, U.K.: Oxford University Press, 1966.
Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
Lewis, David. "How to Define Theoretical Terms." *Journal of Philosophy* 67 (1970): 427–446.
Nagel, Ernest. *The Structure of Science*. New York: Harcourt, Brace, and World, 1961.
Psillos, Stathis. "Carnap, the Ramsey-Sentence, and Realistic Empiricism." *Erkenntnis* 52 (2000): 253–279.
Putnam, Hilary. "Explanation and Reference." In his *Mind, Language, and Reality*. Vol. 2 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1975.
Putnam, Hilary. "What Theories Are Not" (1962). In his *Mathematics, Matter, and Method*. Vol. 1 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1975.
Ramsey, Frank Plumpton. "Theories" (1929). In his *Foundations of Mathematics and Other Logical Essays*, edited by R. B. Braithwaite, 212–236. Paterson, NJ: Littlefield and Adams, 1960.
Suppe, Frederick, ed. *The Structure of Scientific Theories*. 2nd ed. Urbana: University of Illinois Press, 1977.
Suppes, Patrick. "A Comparison of the Meaning and Use of Models in Mathematics and the Empirical Sciences" (1961). In his *Studies in the Methodology and Foundations of Science*, 10–23. Dordrecht, Netherlands: Reidel, 1969.
Suppes, Patrick. "Models of Data." In *Logic, Methodology, and the Philosophy of Science*, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski, 252–267. Stanford, CA: Stanford University Press, 1962.
Van Fraassen, Bas C. *Laws and Symmetry*. Oxford, U.K.: Oxford University Press, 1989.
Van Fraassen, Bas C. *The Scientific Image*. Oxford, U.K.: Oxford University Press, 1980.
Worrall, John. "Structural Realism: The Best of Both Worlds?" *Dialectica* 43 (1989): 99–124.

James Ladyman (2005)

THEORIES OF TYPES

See *Russell, Bertrand Arthur William (section on logic and mathematics); Type Theory*

THEOSOPHY

See *Steiner, Rudolf; Swedenborg, Emanuel*

THINKING

"Thinking" is an essentially human activity occurring in two basic forms. We may think in order to attain knowledge of what is, must, or may be the case; we also may think with a view to making up our mind about what we will or will not do. Following Aristotle, these two forms of thought may be called, respectively, contemplation and deliberation. Both forms may be carried on well or badly, successfully or unsuccessfully, intelligently or stupidly. When contemplation is successful, it terminates in a conclusion; successful deliberation terminates in a decision or resolution. Again following Aristotle, the form of reasoning involved in contemplation may be called theoretical, and the form involved in deliberation may be called practical. Obviously, our day-by-day reasoning in ordinary life is an untidy mixture of both these basic forms.

Less generally, thinking is commonly understood as a largely covert activity, something done mainly *in foro interno*. This activity is also conceived of as intentional in Franz Brentano's sense of "being directed towards an object." For whether we are trying to solve a logical puzzle or are in the process of making up our minds about what to say to a noisy, officious neighbor, we are thinking about something or other. This object (or subject) of our thinking may be either abstract or concrete. We may think about courage, justice, or humanity just as easily as we think about our neighbors and friends, our flowers and the evening sunset. In thinking about these various objects, whether abstract or concrete, we are also necessarily thinking something about them. We think of them as having various features, as doing something or other, or as being related in this or that way to other things of various sorts. For convenience, we may express the last fact about thinking by saying that our specific thoughts have contents as well as objects. We may think that the rain is welcome, that Mary is enchanting, that debts ought to be paid, or that triangularity entails trilaterality.

Another distinctive feature of particular thoughts is that the language used to describe them is nonextensional in a rich sense that is commonly called intentional. As Roderick Chisholm has pointed out, this type of discourse has three distinguishing marks. For one thing, some sentences used to describe thoughts or to ascribe them to thinkers may contain a substantive expression (a name or description) in such a way that neither the sen-

tence nor its negation implies either the existence or the nonexistence of that thing to which the substantive expression truly applies. An example of such a sentence, which illustrates that one may think about nonexisting objects, is "Tommy is thinking about Santa Claus."

Second, a noncompound sentence about thinking may contain a prepositional clause in such a way that neither the sentence nor its negation implies either the truth or the falsity of the propositional clause. An example of such a sentence, emphasizing that one may think what is false, is "It occurred to Jones that demons cause schizophrenia." Finally, a sentence like "Mary thought that the author of *Waverley* wrote *Ivanhoe*" has the peculiarity that although Walter Scott is the author of *Waverley*, one cannot infer that Mary thought that Scott wrote *Ivanhoe.* This last mark of intentionality implies that although things or events have many names and may be described in many different ways, the fact that a person thinks of them in connection with one name or description does not imply that he thinks of them in connection with some other name or description.

From these few remarks about the nonextensional character of discourse about thoughts, several important conclusions about the nature of thinking may immediately be drawn. First, of all the logically equivalent linguistic forms that may be used to describe either the object or the content of a person's thought, only one such form is in most cases strictly applicable. This suggests that thinking something about a particular subject generally involves conceiving of the subject under a certain name or description and attributing something to the subject according to a fairly specific form of attribution. To the extent that the name or description and the attribution are expressible in certain specific words, it will not, in general, be true that an expression or description of the thought in some other words will be equally accurate. The force of this point may be put by saying that at least some thoughts are essentially conceptual, tied to a particular mode of conceiving of a thing or attribute, and felicitously expressed only in specific verbal forms.

Another consequence of these considerations is that certain thoughts have a particular logical form. This emerges not only from the fact that in most thoughts a subject (or object) is in some way characterized, so that the thinking may involve the idea of, schematically, *S*'s being *M*, but also from the possibility that certain logical forms may be involved in a thought while equivalent forms are not. Thus, from "Jones thought that it will rain or snow," it does not follow that Jones thought that it will not both not rain and not snow, even though what is thought in these two cases is logically equivalent by virtue of De Morgan's laws. (One reason that this implication does not hold is that Jones may never have heard of these laws.)

Taking all of what has been said about particular thoughts into account, it appears that as ordinarily conceived, the thoughts involved in both contemplation and deliberation have the following basic features. First, they are characteristically, but perhaps not necessarily, carried on *in foro interno*. Second, they are directed toward an object or a number of objects, and they either attribute something to, or deny something about, this object or objects. Third, the language used to describe them is nonextensional in the sense of possessing at least one of the three intentional marks mentioned above. Fourth, thoughts are often conceived in relation to, and are felicitously expressible by, specific verbal forms; that is, they are often essentially linguistic or conceptual. Finally, particular thoughts have some kind of logical form; they may be categorical, hypothetical, disjunctive, universal, particular, and the like. In general, it may be said that the philosophical task of analyzing the concept of thinking must yield an explanation of exactly what sort of activity thinking is and of how and to what extent it can possess the features just mentioned.

TRADITIONAL THEORIES

A survey of the full range of views on thinking that have been influential in the history of philosophy would reveal, roughly speaking, that most important theories of thinking have been variants of one or more of the following basic views: Platonism, Aristotelianism, conceptualism, imagism, psychological nominalism, and behaviorism. A brief description and criticism of these may thus serve as a useful introduction to the philosophical theory of thinking.

According to the Platonist, thinking is either a dialogue in the soul involving mental words that refer to Forms (such as Redness, Triangularity, Flying) and, possibly, to individuals (such as Socrates) or a spiritual activity of inspecting or recollecting Forms and discerning their natures and interrelations. According to Aristotelianism, thinking is an act of the intellect in which a thing's essence, or intelligible form, actually qualifies the intellect; to think about humanity is for one's intellect to be informed by—literally, to share—the essence humanity. To the extent that one thinks something about humanity—for instance, that it involves animality—one's intellect is also informed by this other essence, the latter being perhaps part of the former.

For conceptualists (the rationalists, for example, and Immanuel Kant) thinking is an activity of bringing concepts or ideas before the mind, these being either innate and applicable to the world in virtue of God's grace (René Descartes, Gottfried Wilhelm Leibniz) or else formed by abstraction from sense experiences and thus actually sharing the abstract features of those experiences (John Locke and, for empirical concepts only, Kant). For imagists (George Berkeley, David Hume) thinking is basically a sequence of episodes involving images; these images are tied to certain "habits," which are the inveterate tendencies of the mind to move from one image to another. To think about triangularity, according to this view, is to imagine some particular triangle while disposed to pass on to other images "of the same sort."

According to the psychological nominalist (such as Thomas Hobbes when he speaks of reasoning) thinking is literally a dialogue in the soul (or, better, in the head) involving the use of verbal images, or mental words, which denote things or classes of things. In this view a complete thought is a mental utterance of a sentence, such as "Tom is tall." Finally, according to behaviorism, thinking is either thoughtful overt speech—thoughtful in the sense that it is in accordance with various principles of relevance, evidence, or inference that the agent is prepared to cite in explanation of his behavior—or a changing series of dispositions to behave intelligently that the agent can at any time avow.

SOME BASIC DIFFICULTIES. One perennial problem peculiar to the Platonic approach is that of accounting for one's ability to learn about the Forms and thus of learning to think. The trouble is that Forms are conceived of as independent of the changing world in which we live, and Plato's suggestion (in the *Phaedo*) that man was born with an ability to "recollect" the Forms experienced in another life is scarcely acceptable to a contemporary thinker. Also, since Forms are conceived of as distinct from the common domain of sense experience, there is a profound difficulty about how to justify knowledge of the Forms. Plato had argued in the *Theaetetus* that true knowledge "can give an account of itself," but it seems that a satisfactory answer has not been given to the question of how agreement in argument or a man's ability to answer objections brought against his view shows knowledge of an independent world of Forms. This problem has been posed more recently, for instance by W. V. Quine, as a demand that the Platonist provide clear, objective criteria for the identity of such strange otherworldly entities as propositions and attributes.

A basic problem for the Aristotelian is to account for the logical form of a thought—that is, for the fact that one may think "If *p* were the case, *q* would be the case" or even "It will either rain or snow." The reason for difficulty here is that there are no intelligible forms corresponding to subjunctive conditionality, to disjunction, or, indeed, to any other logical relation, and it is by no means clear how the intelligible essences that do inform the intellect can be joined to constitute a thought about something conditional or disjunctive. Also, since all general ideas are presumably to be extracted from the sensible forms of experienced objects, thought about what is unobservable, like electrons and negative charges, seems to be impossible as well.

Apart from their highly questionable theories of intelligible essences, one basic drawback common to the Platonic and the Aristotelian views of thinking is their difficulty in accounting for a man's ability to think about particular, nonabstract objects. In the *Sophist*, Plato does, it is true, suggest that some of the mental words of a soul's dialogue may refer to particulars such as Socrates, but his general position is that the objects of thought must be unchanging, intelligible objects, which are universal rather than particular. In arguing that the individuality of a thing is determined by its matter, which is essentially a potentiality rather than an actuality, Aristotle was committed to a similar view, although his medieval heirs argued that particulars could be thoroughly conceived of if, like angels and gods, they constituted the only possible members of a species.

John Duns Scotus, philosophizing as a modified Aristotelian, attempted to get around this difficulty by arguing that particulars are merely congeries of universals. This view, although common in the objective idealism of the nineteenth century, faces a serious problem of distinguishing actual from merely possible particulars or, as Leibniz would have expressed it, of distinguishing a world containing a certain actual particular from a merely possible world containing a "compossible" particular. This Leibnizian type of objection tends to be expressed today by saying that the language used to characterize actual, as opposed to merely possible or fictional, particulars is essentially token reflexive, involving an implicit reference to the speaker: adequate identification of a particular concrete thing cannot be given wholly in context-independent general terms (see Stuart Hampshire and P. F. Strawson).

A difficulty common to conceptualism and Aristotelianism is that in most of their forms they involve an untenable theory of concept formation—namely,

abstractionism. As Peter Geach pointed out, this theory fails even for the favorite examples of the abstractionist since one cannot abstract the concept of color from an experience of scarlet, the latter not being redness plus a differentia. Conceptualists also share with Aristotelians the difficulty already noted of giving an adequate account of the logical form of various thoughts. Kant, a conceptualist, went further than most in the attempt, but he was forced to bring in a priori categories and to insist that men are born with an innate ability to think according to such patterns as "All … are …" and "Either … or …." His approach in this regard was unsatisfactory not only because it is out of line with the well-attested fact that one must *learn* to think according to certain patterns but also because there are no special patterns in accordance with which all men *must* think. (On the last point see B. J. Whorf.)

Imagism shares with Aristotelianism and conceptualism the difficulty of accounting for the logical forms of thought, but it faces the added difficulty of explaining our ability to think of things never perceived, like infinity and million-sided polygons. Although psychological nominalism escapes these difficulties with ease, it runs headlong into the objection that we do not constantly mutter words to ourselves throughout every thinking moment. This objection is not meant to imply that we never think in words; its point is, rather, that we do not always do so and that it is not essential to our thinking one thing rather than another that we experience some verbal imagery. The final alternative, behaviorism, is simply Procrustean as a theory of thinking, for it ignores the plain fact that we do commonly think to ourselves *in foro interno*. As a result of this failure, the behaviorist is unable to account satisfactorily for the changes in behavior and behavioral dispositions that are frequently brought about by our silent deliberation and contemplation.

MERITS OF TRADITIONAL THEORIES. Although each theory just discussed has serious drawbacks and can therefore be said to fail in some measure or other, each nevertheless has some hold on the truth. Thus, the Platonist's idea that thinking is a kind of dialogue in the soul is not entirely empty, for while all thinking is not inner speech pure and simple, it is still true that it is generally like inner speech in crucial respects and that it is felicitously expressed in verbal discourse. The implication that thinking may be carried out *in foro interno* and yet not be mere inner speech is also shared by conceptualism and imagism. The latter has the added advantage of accounting for the occasional utility of imagistic thinking, as in pondering the location of a town on a map, the kind of angle formed by certain intersecting lines, and so on (see H. H. Price). Psychological nominalism actually accounts for most features of conceptual thinking except for the possibility of its occurring without verbal imagery. The forms of thought are explained by reference to the forms of the sentences used in inner speech, the object and content of a thought are explained with reference to the words used, and so on.

Behaviorism, finally, although not without its shortcomings, does have the advantage of accounting for the important fact that some episodes of thinking, such as resolves and decisions, essentially involve behavioral dispositions: If a man is not moved, or disposed, to do *A* when he believes he is in circumstances *C*, he is not, *ceteris paribus*, resolved or decided to do *A* in *C*. The crucial importance of this tie-up between certain forms of thought and behavioral dispositions is that it shows how an explanation of behavior in terms of reasons (rather than causes) can be acceptable. Without this tie-up we would have to say that a man's reasons for acting are strictly irrelevant to the question of why he so acted, for the intellect could not then "move a man to act."

TOWARD AN ADEQUATE ACCOUNT

A useful way of working out an account of thinking free from the drawbacks of traditional theories is to examine Gilbert Ryle's influential critique of all theories that insist that thinking must be done *in foro interno*. According to his argument in *The Concept of Mind*, all such theories are based on the mistaken idea that nonhabitual, intelligent human behavior is always guided by silent thought, whose presence explains why the behavior occurs and why it is intelligent. In Ryle's opinion this persistent idea is plainly untenable and leads to a vicious regress. This regress occurs because thinking is itself an activity that is admittedly done well or badly, intelligently or stupidly. This being so, the idea in point would imply that the intelligent character of thinking requires explanation by further thinking, which in turn guides the first thinking and explains why it occurs, why it is intelligent, and the like. Since this further thinking will itself be done well or badly, intelligently or stupidly, it will also require explanation by a third line of thinking and so on without end.

In rejecting this traditional idea, Ryle argues that reference to interior and anterior acts of thinking is not in any way needed for the explanation of most intelligent behavior. In his view a form of behavior, especially verbal behavior, may be regarded as intelligent, thoughtful, or even rational if it is done in accordance with certain principles of inference, evidence, relevance, and so on. That

the behavior is in accordance with these principles does not mean that they are rehearsed in thought while the behavior is being carried out. On the contrary, it means only that the behavior conforms to, or is in line with, these principles and that the agent is disposed to cite or at least to allude to them if called upon to explain his behavior. Thus, if a man calculates out loud, then—assuming that this calculation is done in accordance with principles in the above sense—there is no need to introduce any further thought episodes to account for the fact that he arrives at a certain conclusion or resolution; the steps that led him to the conclusion or resolution are already laid bare. If the calculation shows intelligence or ingenuity, it does so by virtue of the relations between the overt steps; going from a premise to a conclusion is not proved reasonable or unreasonable, rational or irrational, by reference to something other than the premise and the conclusion. When we have the premise and the conclusion, we have all we need to decide whether the inference was reasonable. Even if we were to allude to interior steps of reasoning in order to explain a man's actions, we would have to appraise those steps in light of the same principles. Therefore, it may, in fact, be said that purely overt calculation or deliberation is itself a process of thinking and that thinking is not something that is necessarily done silently in the soul. In other words, overt thinking is just as useful a mode of thinking as any other, and there is no need, even no point, in always hunting for hidden acts of thought.

CRITICISM OF RYLE'S APPROACH. Although there is considerable plausibility to Ryle's approach, it must be granted that not all the calculation or deliberation that accounts for a man's actions is done out loud or on paper. In fact, nothing is more obvious than the fact that a good share of one's calculation is not done overtly and that reference to silent thought is constantly and legitimately made in order to account for activities that would otherwise remain inexplicable. Thus, a man may make a move in chess after sitting in silent anguish for long minutes at the board; and the intelligence of this move will remain a stubborn question mark until, perhaps after the game, he outlines the strategy behind it. The same is true in countless other cases. On being asked a question, the mathematics student may close his eyes for a minute before giving the answer, and when the answer is given, he can usually follow it with a proof, a line of reasoning he will claim to recall having thought out *in foro interno*.

Ryle was, of course, aware of these cases in *The Concept of Mind*, and he attempted to account for them by arguing that a man can learn to mutter to himself as well as mutter out loud. Thus, when pressed, Ryle could not entirely dispense with the traditional conception of covert thinking; in regarding it as "inner speech" he was, in fact, squarely in the tradition of Hobbes, and his view is thus subject to the same fundamental difficulty—namely, that to most it seems plainly false that inner speech occurs whenever one can correctly be said to think *in foro interno*.

THE ANALOGY THEORY

Although Ryle's view of thinking does not, as a whole, succeed, in the opinion of the present writer it does come close to the truth. For while silent thought need not be inner speech, it may still be an activity that is at least formally analogous to speech. In what sense "formally analogous"? In the sense in which chess played with pennies and nickels is formally analogous to chess played with standard pieces or in which the Frenchman's "Il pleut" is formally analogous to the Englishman's "It is raining": the same basic moves are made, but the empirical features of the activities are different. Thus, while the thought p is empirically different from the act of saying that p (in that the former need not even involve verbal imagery), it may still be regarded as formally the same: Both are activities that conform to the same principles and have many of the same implications. This sort of formal identity among empirically different activities is, of course, hard to state clearly, but at least an intuitive sense of what is meant by speaking of such an identity can be conveyed by the following analogy. Saying that p is a formal analogue of thinking that p in the way that playing "Texas chess" (with automobiles on certain counties) is a formal analogue of playing ordinary chess (with ivory pieces on checkered boards). What is essential in both cases is that formally analogous activities are carried on in accordance with the same basic principles—the principles or rules of chess, on one hand, and various principles of inference and relevance, on the other.

This theory of thinking, which may be called the analogy theory, does more than merely correct the shortcomings of Ryle's view. It also seems to account for all of the distinctive features of conceptual thinking that were mentioned earlier. Since it also appears to possess none of the drawbacks of traditional theories, it is perhaps the most satisfactory account of thinking yet developed by philosophers.

See also Being; Empiricism; Intention; Universals, A Historical Survey.

Bibliography

CONTEMPORARY DISCUSSIONS

Alston, William P. "The Role of Reason in the Regulation of Belief." In *Rationality in the Calvinian Tradition*, edited by N. Wolterstorff et al. Lanham, MD: University Press of America, 1983.

Bealer, George. "A Theory of Concepts and Concept Possession." *Proceedings of the Tenth Annual SOFIA Conference*, edited by Enrique Villanueva. Atascadero, CA: Ridgeview Press.

Bergmann, Gustav. "Intentionality." In *Meaning and Existence*. Madison: University of Wisconsin Press, 1960.

Chisholm, Roderick. "Sentences about Believing." *PAS* 56 (1955–1956): 125–148.

Chisholm, Roderick, and Wilfrid Sellars. "Intentionality and the Mental." In *Minnesota Studies in the Philosophy of Science*, Vol. II, edited by Herbert Feigl and Grover Maxwell. Minneapolis: University of Minnesota Press, 1958.

Davidson, Donald. "What Thought Requires." In *The Foundations of Cognitive Science*, edited by Joao Branquinho. Oxford: Clarendon Press, 2001.

Gibbard, Allan. "Thoughts and Norms." *Nous Supplement* 13 (2003): 83–98.

Ginnane, W. J. "Thoughts." *Mind* 49 (1960): 372–390.

Hampshire, Stuart. *Thought and Action*. London: Chatto and Windus, 1959.

Harman, Gilbert. *Change in View: Principle of Reasoning*. Cambridge, MA: MIT Press, 1986.

Harman, Gilbert. *Thought*. Princeton, NJ: Princeton University Press, 1973.

Jeffrey, Richard. *The Logic of Decision.* 2nd ed. Chicago: University of Chicago Press, 1983.

Kenny, Anthony. *Action, Emotion, and Will*. London: Routledge and Kegan Paul, 1963.

Lehrer, Keith. *Metamind*. Oxford: Clarendon Press, 1990.

Price, H. H. *Thinking and Experience.* 2nd ed. London: Hutchinson, 1969. An extremely valuable discussion of traditional theories of thinking.

Quine, W. V. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949.

Skyrms, Brian. *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard University Press, 1990.

Stalnaker, Robert C. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Stein, Edward. *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Oxford: Oxford University Press, 1996.

Strawson, P. F. *Individuals*. London: Methuen, 1959.

Whorf, B. J. *Language, Thought, and Reality*. Cambridge, MA: Technology Press of Massachusetts Institute of Technology, 1956.

TRADITIONAL THEORIES

Detailed information on traditional theories may be found in F. C. Copleston, *A History of Philosophy.* 7 vols. (New York, 1959). Specific reference may be made, however, to the following classics, which are published in numerous editions: Aristotle, *De Anima* and *Nicomachean Ethics*, Book 5; George Berkeley, *Principles of Human Knowledge*, introduction; René Descartes, *Meditations* and *Principles of Philosophy*; Thomas Hobbes, *Leviathan*, Part I, Ch. 5, and *Elements of Philosophy*, Chs. 2–3; David Hume, *Enquiry concerning Human Understanding*, Secs. 2–4, and *Treatise of Human Nature*, Secs. 1–4; Immanuel Kant, *Critique of Pure Reason* and *Critique of Practical Reason*; G. W. Leibniz, *Leibniz Selections*, edited by P. P. Wiener (New York, 1951); John Locke, *Essay concerning Human Understanding*, Book II, Chs. 1, 2, 7, 8; Plato, *Phaedo, Republic, Sophist*, and *Theaetetus*.

PSYCHOLOGICAL STUDIES

Bartlett, F. C. *Thinking*. London: Allen and Unwin, 1953.

Bruner, J. S., J. J. Goodnow, and G. A. Austin. *A Study of Thinking*. New York: Wiley, 1956.

THE ANALOGY THEORY

Aune, Bruce. *Knowledge, Mind, and Nature: An Introduction to Theory of Knowledge and the Philosophy of Mind*. New York: Random House, 1967. Ch. 8.

Aune, Bruce. "On Thought and Feeling." *Philosophical Quarterly* 13 (1963): 1–12.

Geach, Peter. *Mental Acts*. London, 1957.

Sellars, Wilfrid. "Physical Realism." *Philosophy and Phenomenological Research*15 (1954): 13–32. An illuminating discussion of conceptualism and recent Platonism.

Sellars, Wilfrid. *Science, Perception, and Reality*. London, 1964.

Bruce Aune (1967)

Bibliography updated by Benjamin Fiedor (2005)

THOMAS À KEMPIS

(1379/1380–1471)

Thomas à Kempis, the writer on asceticism and probable author of *The Imitation of Christ,* was born in Kempen, near Düsseldorf, Germany. He belonged to the Brethren of the Common Life, a group that was much influenced by Jan van Ruysbroeck and whose organization centered on the Windesheim community. The major part of Thomas's life was spent at the Augustinian monastery of St. Agnes, near Zwolle.

Thomas's writings on the interior life and ways of practicing virtue are not philosophical or theoretical but are purely practical in intent. This is true also of *The Imitation of Christ,* about whose authorship there has been much dispute. It is not altogether certain that the work, really a set of four treatises, should be attributed to Thomas. The oldest manuscripts date from about 1422 and contain only the first book, and the first complete edition goes back to 1427. Since the work is not quoted earlier than the fifteenth century, it seems likely that it originated during Thomas's lifetime. Moreover, the style is remarkably like that of writings that can certainly be ascribed to him (a statistical investigation has also supported this). For these reasons we can rule out certain

speculative attributions (to Jean Gerson and to John Gersen, in the thirteenth century). On the other hand, the first attribution of the book to him occurred rather late, in the second edition of an account of the Windesheim community written in the latter part of the fifteenth century. The fact that Thomas signed a manuscript of the *Imitation* is not conclusive, for he was, like his fellow monks, a copyist and also signed a Bible. But the balance of probability is that Thomas himself compiled the work anonymously, and he certainly incorporated into it materials not original to himself, especially in the first book.

The wide circulation of the book was partly due to the efforts of the copyists at Windesheim, but it was also due to the kind of piety it recommended. The second part of the full title (*Of the Imitation of Christ and of Contempt for All Worldly Vanities*) indicates that its teachings were adapted to the monastic life—and indeed it was primarily intended as a handbook for monks. But its tender concentration on the figure of Jesus made attractive its doctrine of resignation—the surrendering of all worldly concerns to the service of, and imitation of, Christ. Moreover, it gave very concrete guidance on many problems—for example, how to distinguish the results of grace from natural acts and propensities. The most notable feature of the book, however, is its uncompromising and uncomfortable insistence on self-mortification as preparation for grace and the presence of the true Lover of the soul, Christ. The "imitation" of Christ that Thomas recommends is not a simple copying of Jesus but acting by analogy with Jesus, whose life was mainly characterized, according to Thomas, by suffering and self-sacrifice.

The first book has mainly to do with the moral reform of the individual. The second concerns the preparation for the interior or illuminative life. The third consists in a dialogue between Christ and the soul that gives a further exposition of ascetic practices, and one or two passages give a hint of the kind of mystical experience awaiting those who truly love Christ. The fourth book is a manual for those who receive Holy Communion.

There is very little theology in the *Imitation.* Thomas seems to have been reacting against the speculations of academic theology, for he wrote: "Of what use is your highly subtle talk about the blessed Trinity, if you are not humble?" and "I would rather feel compunction than be able to produce the most precise definition of it." The strongly practical bent of the work, in any event, gave it a continuing relevance to the Christian life and enabled it to achieve the status of a classic ranking, in Christian piety, with *Pilgrim's Progress.*

See also Asceticism; Gerson, Jean de; Ruysbroeck, Jan van; Virtue and Vice.

Bibliography

WORKS BY THOMAS

Opera et Libri Vite Fratris Thome a Kempis. Edited by P. Danhausser. Nuremberg: Per Caspar Hochfeder, 1494. Critical edition by M. J. Pohl, 7 vols. Freiburg im Breisgau, 1902–1922.

Prayers and Meditations on the Life of Christ. Translated by William Duthoit London, 1904.

The Founders of the New Devotion. Translated by J. P. Arthur. London, 1905.

The Chronicle of the Canons Regular of Mount St. Agnes. Translated by J. P. Arthur. London: K. Paul, Trench, Trubner, 1906.

Meditations and Sermons on the Incarnation, Life and Passion of Our Lord. Translated by Dom Vincent Scully. London, 1907.

Sermons to the Novices Regular. Translated by Dom Vincent Scully. London: K. Paul, Trench, Trubner, 1907.

The Imitation of Christ. Translated by Ronald Knox and Michael Oakley. New York: Sheed and Ward, 1960. The freshest and most direct translation.

WORKS ON THOMAS

Huijben, Jacques, and Pierre Debougnie. *L'auteur ou les auteurs de L'Imitation.* Louvain, 1957. On the question of authorship of the *Imitation.*

Scully, Dom Vincent. *Life of the Venerable Thomas à Kempis.* London: R. and T. Washbourne, 1901.

Yule, George Udney. *The Statistical Study of Literary Vocabulary.* Cambridge, U.K.: Cambridge University Press, 1944. On the question of the authorship of the *Imitation.*

Ninian Smart (1967)

THOMAS AQUINAS, ST.
(c. 1224–1274)

St. Thomas Aquinas, the Catholic theologian and philosopher, was born at Roccasecca, Italy, the youngest son of Landolfo and Teodora of Aquino. At about the age of five he began his elementary studies under the Benedictine monks at nearby Montecassino. He went on to study liberal arts at the University of Naples. It is probable that Thomas became a master in arts at Naples before entering the Order of Preachers (Dominicans) in 1244. He studied in the Dominican courses in philosophy and theology, first at Paris and, from 1248 on, under Albert the Great at Cologne. In 1252 he was sent to the University of Paris for advanced study in theology; he lectured there as a bachelor in theology until 1256, when he was awarded the magistrate (doctorate) in theology. Accepted after some opposition from other professors as a fully

accredited member of the theology faculty in 1257, Thomas continued to teach at Paris until 1259.

Thomas Aquinas then spent almost ten years at various Dominican monasteries in the vicinity of Rome, lecturing on theology and philosophy (including an extensive study of the major works of Aristotle) and performing various consultative and administrative functions in his order. In the fall of 1268 Aquinas returned for his second professorate in theology at the University of Paris. He engaged in three distinct controversies: against a group of conservative theologians who were critical of his philosophic innovations; against certain radical advocates of Aristotelianism or Latin Averroism; and against some critics of the Dominicans and Franciscans and their right to teach at the university. Many of Aquinas's literary works were in process or completed at this time. It is thought that he was provided with secretarial help in this task, partly in view of the fact that his own handwriting was practically illegible. Called back to Italy in 1272, Aquinas taught for a little more than a year at the University of Naples and preached a notable series of vernacular sermons there. Illness forced him to discontinue his teaching and writing toward the end of 1273. Early in 1274 he set out for Lyons, France, to attend a church council. His failing health interrupted the trip at a point not far from his birthplace, and he died at Fossanova in March of that year.

The writings of Thomas Aquinas were produced during his twenty years (1252–1273) as an active teacher. All in Latin, they consist of several large theological treatises, plus recorded disputations on theological and philosophical problems (the "Disputed Questions" and "Quodlibetal Questions"), commentaries on several books of the Bible, commentaries on twelve treatises of Aristotle, and commentaries on Boethius, the pseudo-Dionysius, and the anonymous *Liber de Causis.* There are also about forty miscellaneous notes, letters, sermons, and short treatises on philosophical and religious subjects. Although Aquinas's philosophic views may be found in almost all his writings (thus the "Exposition of the Book of Job" reads like a discussion among philosophers), certain treatises are of more obvious interest to philosophers. These are listed in detail at the end of this entry.

GENERAL PHILOSOPHICAL POSITION

In the main, Aquinas's philosophy is a rethinking of Aristotelianism, with significant influences from Stoicism, Neoplatonism, Augustinism, and Boethianism. It also reflects some of the thinking of the Greek commentators on Aristotle and of Cicero, Avicenna, Averroes, Solomon ben Judah ibn Gabirol, and Maimonides. This may suggest that we are dealing with an eclectic philosophy, but actually Aquinas reworked the speculative and practical philosophies of his predecessors into a coherent view of the subject that shows the stamp of his own intelligence and, of course, the influence of his religious commitment.

One of the broad characteristics of Aquinas's work in philosophy is a temperamental tendency to seek a middle way on questions that have been given a wide range of answers. This spirit of moderation is nowhere better illustrated than in his solution to the problem of universals. For centuries philosophers had debated whether genera and species are realities in themselves (Plato, Boethius, William of Champeaux) or mere mental constructs (Roscelin, Peter Abelard). What made this odd discussion important was the conviction (certainly shared by Aquinas) that these universals (such as humanity, justice, whiteness, dogness) are the primary objects of human understanding. Most thinkers in the Middle Ages felt that if something is to be explained, it must be treated in universal terms. Therefore, the problem of universals was not simply an academic question.

Aquinas's position on this problem is now called moderate realism. He denied that universals are existing realities (and frequently criticized Plato for having suggested that there is a world of intelligible Forms), but he also insisted that men's universal concepts and judgments have some sort of foundation in extramental things. This basis for the universality, say of humanity, would consist in the real similarity found among all individual men. It was not that Aquinas attributed an actual, existent universal nature to all individual men: that would be an extreme realism. Rather, only individuals exist; but the individuals of a given species or class resemble each other, and that is the basis for thinking of them as universally representative of a common nature.

Thomas's spirit of compromise as a philosopher was balanced by another tendency, that toward innovation. His original Latin biographers all stress this feature of his work. Thomas introduced new ways of reasoning about problems and new sources of information, and he handled his teaching in a new way. In this sense Thomas Aquinas was not typical of the thirteenth century and was perhaps in advance of his contemporaries.

FAITH AND RATIONAL KNOWLEDGE

As Aquinas saw it, faith (*fides*) falls midway between opinion and scientific knowledge (*scientia*); it is more than opinion because it involves a firm assent to its

object; and it is less than knowledge because it lacks vision. Both are intellectual acts and habits of assent: in the case of faith a person is not sufficiently moved by the object to accept it as true, so, by an act of will, he inclines himself to believe. Knowledge implies assent motivated by a personal seeing of the object without any direct influence from will. Where objects of belief have to do with divine matters that exceed man's natural cognitive capacity, the disposition to believe such articles of religious faith is regarded as a special gift from God. Reason (*ratio*) is another type of intellectual activity: Simple understanding and reasoning differ only in the manner in which the intellect works. Through intellection (understanding) one knows simply by seeing what something means, while through reason one moves discursively from one item of knowledge to another. (These functions of believing and knowing are treated in many places by Aquinas: *Summa Contra Gentiles* III, 147; *In Boethii de Trinitate,* Ques. II and III; *Summa Theologiae* I, Ques. 79–84.)

Aquinas thought that philosophy entailed reasoning from prior knowledge, or present experience, to new knowledge (the way of discovery) and the rational verification of judgments by tracing them back to more simply known principles (the way of reduction). Where the basic principles are grasped by man's natural understanding of his sensory experiences, the reasoning processes are those of natural science and philosophy. If one starts to reason from judgments accepted on religious faith, then one is thinking as a theologian. Questions V and VI of *In Boethii de Trinitate* develop Aquinas's methodology of the philosophical sciences: philosophy of nature, mathematics, and metaphysics. He distinguished speculative or theoretical reasoning from the practical: The purpose of speculation is simply to know; the end of practical reasoning is to know how to act. He described two kinds of theology: The philosophical "theology," metaphysics, which treats divine matters as principles for the explanation of all things, and the theology taught in Scripture, which "studies divine things for their own sakes" (*In Boethii de Trinitate* V, 4 c).

Thus philosophy, for Aquinas, was a natural type of knowledge open to all men who wish to understand the meaning of their ordinary experiences. The "philosophers" whom he habitually cited were the classic Greek, Latin, Islamic, and Jewish sages. Christian teachers mentioned by Aquinas were the "saints" (Augustine, John of Damascus, Gregory, Ambrose, Dionysius, Isidore, and Benedict); they were never called Christian philosophers. The word *theology* was rarely used by Aquinas. In the first question of his *Summa Theologiae* he formally calls his subject sacred doctrine (*sacra doctrina*) and says that its principles, unlike those of philosophy, are various items of religious faith.

Thus, Thomas Aquinas was by profession a theologian, or better, a teacher of sacred doctrine who also studied and wrote about philosophy. He obviously used a good deal of pagan and non-Christian philosophy in all his writings. His own understanding of these philosophies was influenced by his personal faith—as almost any man's judgment is influenced by his stand for or against the claim of religious faith—in this sense Thomism is a "Christian philosophy." Aquinas did not ground his philosophical thinking on principles of religious belief, however, for this would have destroyed his distinction between philosophy and sacred doctrine, as presented in the opening chapters of the first book of *Summa Contra Gentiles.* One of the clearest efforts to maintain the autonomy of philosophy is found in Aquinas's *De Aeternitate Mundi* (about 1270), in which he insists that, as far as philosophical considerations go, the universe might be eternal. As a Christian, he believed that it is not eternal.

Among interpreters of Aquinas there has been much debate whether his commentaries on Aristotle deal with his personal thinking. It is generally agreed even by non-Thomists (W. D. Ross, A. E. Taylor) that these expositions are helpful to the reader who wishes to understand Aristotle. It is not so clear whether the mind of Aquinas is easily discernible in them. One group of Thomists (Étienne Gilson, Joseph Owens, A. C. Pegis) stresses the more obviously personal writings (such as the two *Summa*'s) as bases for the interpretation of his thought; another school of interpretation (J. M. Ramírez, Charles De Koninck, J. A. Oesterle) uses the Aristotelian commentaries as the main sources for Aquinas's philosophic thought.

THEORY OF KNOWLEDGE. The Thomistic theory of knowledge is realistic. (This theory is presented in *Summa Theologiae* I, 79–85; *Quaestiones Disputatae de Veritate* I, II; *In Libros Posteriorum Analyticorum* I, 5; II, 20.) Men obtain their knowledge of reality from the initial data of sense experience. Apart from supernatural experiences that some mystics may have, Thomas limited human cognition to sense perception and the intellectual understanding of it. Sense organs are stimulated by the colored, audible, odorous, gustatory, and tactile qualities of extramental bodies; and sensation is the vital response through man's five external sense powers to such stimulation. Aquinas assumed that one is cognitively aware of red flowers, noisy animals, cold air, and so on. Internal sensa-

tion (common, imaginative, memorative, and cogitative functions) works to perceive, retain, associate, and judge the various impressions (phantasms) through which things are directly known. Man's higher cognitive functions, those of understanding, judging, and reasoning, have as their objects the universal meanings that arise out of sense experience. Thus, one sees and remembers an individual apple on the level of sensation—but he judges it to be healthful because it contains vitamins, or for any other general reason, on the level of intellectual knowledge. Universals (health, humanity, redness) are not taken as existing realities but are viewed as intelligibilities (*rationes*) with a basis in what is common to existents. As a moderate realist, Aquinas would resent being classified as a Platonist; yet he would defend the importance of our knowledge of the general and common characteristics of things.

Although human cognition begins with the knowing of bodily things, man can form some intellectual notions and judgments concerning immaterial beings: souls, angels, and God. Aquinas taught that man does this by negating certain aspects of bodies (for instance, a spirit does not occupy space) and by using analogy. When the notion of power is attributed to God, its meaning is transferred from an initially physical concept to the analogous perfection of that which can accomplish results in the immaterial order. Thomas did not think that men, during earthly life, can know the nature of God in any adequate, positive way.

Discursive reasoning was taken as an intellectual process moving from or toward first principles in logical processes of demonstration (the ways of discovery and reduction, described above). In one way, sense experience is the first principle (starting point) for all of man's natural knowledge. This is one aspect of Aquinas's empiricism. Following Aristotle's *Posterior Analytics,* Thomas taught that many sensations combine to form a unified memory, and many memories constitute sense experience. From this manifold of experience, by a sort of sensory induction, there arises within human awareness a beginning (*principium*) of understanding. Such first principles are not demonstrated (they naturally emerge from sense cognition), but they become the roots for consequent intellectual reasoning. A doctor who tries a variety of remedies to treat headaches eventually notices that one drug works well in almost all cases—at some point he grasps the universal "Drug *A* is a general remedy for headache." From this principle he proceeds rationally to order his practice. If he becomes a teacher of medicine, he uses such a theoretical principle to instruct others. This is the basis of the life of reason.

PHILOSOPHY AND THE PHYSICAL WORLD

In his exposition of the *Liber de Causis* (Lect. 1), Aquinas described a sequence of philosophic studies: logic, mathematics, natural philosophy (physics), moral philosophy, and, finally, metaphysics. The first kind of reality examined in this course would be that of the physical world. (At the start of the next century, John Duns Scotus criticized Thomas for attempting to base his metaphysics and his approaches to God on physics.) Interpreters still debate whether Aquinas himself felt that this was the order to be followed in learning philosophy, or whether he was merely reporting one way that the "philosophers" had taught it. In any case, the philosophical study of bodies, of mobile being in the Aristotelian sense, was important to Aquinas. One group of his writings (*De Principiis Naturae,* parts of Book II of the *Summa Contra Gentiles,* the treatise *De Aeternitate Mundi*) offers a quite personal treatment of this world of bodies. Another set of writings (the commentaries on Aristotle's *Physics* and *De Generatione et Corruptione*) shows how indebted Aquinas was to Aristotle in his theory of physical reality.

MATTER AND FORM. The philosophy of nature (*phusis*) was understood as the study of a special kind of beings, those subject to several kinds of change. Physical beings have primary matter as one component and, depending on their species or kind, substantial form as their other integral principle. Neither matter nor form is a thing by itself; matter and form are simply the determinable and determining factors within any existing physical substance. Like Aristotle, Aquinas took it that there are many species of bodily substances: all the different kinds of inanimate material (wood, gold, water, etc.) and all the species of plants and animals. Within each such species there is one specifying principle (the substantial form of wood, potato plant, or dog), and the many individual members of each species are differentiated by the fact that the matter constituting dog *A* could not also constitute dog *B* (so viewed, matter is said to be quantified, or marked by quantity).

CHANGE. Being mobile, physical beings are subject to four kinds of change (*motus*): of place (locomotion), of size (quantitative change), of color, shape, and so on (qualitative change), and of species of substance (generation and corruption, substantial change). Basically, prime matter is that which remains constant and provides con-

tinuity during a change from one substance to another. When a pig eats an apple, that part of the apple really assimilated by the pig becomes the very substance of the pig; some factor in the apple, the prime matter, must continue on into the pig. All four types of change are explained in terms of the classic theory of four causes. The final cause is the answer to the question "why" something exists or occurs; the agent or efficient cause is the maker or producer of the change; the material cause is that out of which the change comes; and the formal cause is the specifying factor in any event or existent. So used, "cause" has the broad meaning of raison d'être.

SPACE AND TIME. Certain other points in Aquinas's philosophy of nature further illustrate the influence of Aristotle. Place, for instance, is defined as the "immobile limit of the containing body" (*In IV Physicorum* 6). Moreover, each primary type of body (the four elements still are earth, air, fire, and water) is thought to have its own "proper" place. Thus, the place for fire is "up" and that for earth is "down." Some sort of absolute, or box, theory of space may be presupposed; yet in the same passage Aquinas's discussion of the place of a boat in a flowing river indicates a more sophisticated understanding of spatial relativity. Time is defined, as in Aristotle, as the measure of motion in regard to "before" and "after." Eternity is a type of duration differing from time in two ways: The eternal has neither beginning nor termination, and the eternal has no succession of instants but exists entirely at once (*tota simul*).

ENCOURAGEMENT OF SCIENCE. Doubtless Aquinas's philosophy of the physical world was limited and even distorted by certain views and factual errors derived from Aristotle and from thirteenth-century science. Apart from the mistaken hypothesis that each element has its proper place in the universe, Thomas also used the Eudoxian astronomy, which placed the earth at the center of a system of from 49 to 53 concentric spheres. (Besides the Commentary on *De Caelo* II, 10, and the Commentary on *Meteorologia* II, 10; see *Summa Contra Gentiles* I, 20, and *Summa Theologiae* I, 68, 4 c.) At times Thomas showed an open mind on such questions and an ability to rise above the limitations of his period. His Commentary on Aristotle's *Metaphysics* (Lect. 1 on Book III and Lect. 9 on Book XII) provides a key instance. Pointing out that astronomers differ widely on the number and motions of the planets, Aquinas recommended that one study all the reports and theories of such scientists, even though these scientific explanations are not the last word on the matter and are obviously open to future revision. He further compared the study of physical science to the work of a judge in a court of law. One should listen to, and try to evaluate, all important testimony before attempting to formulate one's own judgment on the problems of contemporary science. This is Aquinas at his best, hardly a philosophical dogmatist.

HUMAN FUNCTIONS AND MAN'S NATURE

Anthropology, or psychology, in the classical sense of the study of man's psyche, forms an important part of Aquinas's philosophy. His view of man owed much to the Aristotelian treatise *On the Soul,* to the Christian Platonism of Augustine and John of Damascus, and to the Bible. This part of Aquinas's thought will be found in *Scriptum in IV Libros Sententiarum* (Commentary on the *Sentences*) I, Dists. 16–27; *Summa Contra Gentiles* II, 58–90; *Quaestio Disputata de Anima*; the *Libros de Anima*; and *Summa Theologiae,* I, 75–90.

Aquinas's usual way of working out his theory of human nature was first to examine certain activities in which man engages, then to reason to the kinds of operative powers needed to explain such actions, and finally to conclude to the sort of substantial nature that could be the subject of such powers. He described the biological activities of man as those of growth, assimilation of food, and sexual reproduction. A higher set of activities included sensory perception, emotive responses to what is perceived, and locomotion: These activities man shares with brute animals. A third group of activities comprises the cognitive functions of understanding, judging, and reasoning, as well as the corresponding appetitive functions of affective inclination toward or away from the objects of understanding. To these various functions Aquinas assigned generic powers (operative potencies) of growth, reproduction, sensory cognition and appetition, physical locomotion, and intellectual cognition and appetition (will).

Reexamining these functional powers in detail, Aquinas distinguished five special sense powers for the cognition of physical individuals: sight, hearing, smell, taste, and touch. These functions and powers are called external because their proper objects are outside the mental awareness of the perceiver: This is essential to epistemological realism. Following these are four kinds of internal sensory activities: the perceptual grasping of a whole object (*sensus communis*), the simple retention of sensed images (imagination), the association of retained images with past time (sense memory), and concrete discrimination or judgment concerning individual things

(cogitative sense, particular reason). Still on the level of sensory experience, Aquinas (here influenced by John of Damascus) described two kinds of appetition (emotion): A simple tendency toward or away from what is sensed as good or evil (this affective power is called the concupiscible appetite), and a more complicated sensory inclination to meet bodily threats, obstacles, and dangers by attacking or avoiding them or by putting up with them (this affective power is called irascible appetite). Eleven distinct kinds of sensory passions (emotions) are attributed to these two sensory appetites: love, desire, delight, hate, aversion, and sorrow to the concupiscible; fear, daring, hope, despair, and anger to the irascible. Much of this psychological analysis is quite sophisticated, employing data from Greek, Roman, and early Christian thought and also using the physiological and psychological treatises of Islamic and Jewish scholars. It also forms the basis of the analysis of human conduct in Thomistic ethics.

On the higher level of distinctively human experience, Aquinas found various other activities and powers. These are described in his commentary on Book III of Aristotle's *De Anima,* in the *Summa Contra Gentiles* (II, 59–78), and in Questions 84–85 of the *Summa Theologiae.* The general capacity to understand (*intellectus*) covers simple apprehension, judging, and reasoning. The objects of intellection are universal aspects (*rationes*) of reality. Since universal objects do not exist in nature, Aquinas described one intellectual action as the abstraction of universal meanings (*intentiones*) from the individual presentations of sense experience. This abstractive power is called agent intellect (*intellectus agens*). A second cognitive function on this level is the grasping (*comprehensio*) of these abstracted meanings in the very act of cognition; this activity is assigned to a different power, the possible intellect (*intellectus possibilis*). Thus, there are two quite different "intellects" in Thomistic psychology: One abstracts, the other knows. No special power is required for intellectual memory; the retention of understandings is explained by habit formation in the possible intellect.

WILL. Affective responses to the universal objects of understanding are functions of intellectual appetition. Considered quite different from sensory appetition, this is the area of volition, and the special power involved is the will (*voluntas*). Aquinas distinguished two kinds of volitional functions. First, there are those basic and natural tendencies of approval and affective approach to an object that is judged good or desirable without qualification. In regard to justice, peace, or a perfectly good being, for instance, Aquinas felt that a person's will would be naturally and necessarily attracted to such objects. This natural movement of the will is not free. Second, there are volitional movements toward or away from intellectually known objects that are judged as partly desirable or as partly undesirable. Such movements of will are directed by intellectual judgments evaluating the objects. In this case volition is said to be "deliberated" (specified by intellectual considerations) and free. It is in the act of decision (*arbitrium*) that man is free. Aquinas did not talk about "free will"; the term *libera voluntas* is found only twice in all his works, and then in a nontechnical usage; rather, he spoke of free choice or decision (*liberum arbitrium*). Man, by virtue of his intellectual powers, is free in some of his actions.

SOUL. Although Aquinas sometimes spoke as if these various "powers" of man were agents, he formally stressed the view that it is the whole man who is the human agent. A human being is an animated body in which the psychic principle (*anima*) is distinctive of the species and determines that the material is human. In other words, man's soul is his substantial form. Some of man's activities are obviously very like those of brutes, but the intellectual and volitional functions transcend materiality by virtue of their universal and abstracted character. Aquinas took as an indication of the immateriality of the human soul the fact that it can understand universal meanings and make free decisions. The soul is a real part of man and, being both immaterial and real, it is spiritual. From certain other features of man's higher activities, especially from the unity of conscious experience, Aquinas concluded to the simplicity and integration of man's soul: It is not divisible into parts. This, in turn, led him to the conclusion that the soul is incapable of corruption (disintegration into parts) and thus is immortal.

Since Thomas thought the soul incapable of being partitioned, he could not explain the coming into being of new human souls by biological process. He was thus forced to the view that each rational soul is originated by divine creation from nothing. Human parents are not the total cause of their offspring; they share the work of procreation with God. This view explains why Aquinas put so much stress on the dignity and sanctity of human reproduction, which he regarded as more than a biological function. When he claimed, in his ethics, that the begetting and raising of children is the primary purpose of married life, he was not thinking of simple sexual activity but of a human participation in God's creative function. This does not mean that man is the highest of God's creatures; Aquinas speculated that there are other kinds of purely intellectual beings with activities, powers, and

natures superior to those of men. These are angels. Thomas Aquinas is called the Angelic Doctor in Catholic tradition because of his great interest in these purely spiritual but finite beings. They would constitute the highest realm of the universe.

METAPHYSICS AND REAL BEING

Aquinas devoted much thought to the question "What does it mean to be?" Many Thomists think that his greatest philosophical ability was shown in the area of metaphysics. His general theory of reality incorporates much of the metaphysics of Aristotle, and some interpreters have seen Thomistic metaphysics as but a baptized Aristotelianism. Recent Thomistic scholarship has selected two non-Aristotelian metaphysical teachings for new emphasis: the theory of participation and the general influence of Platonic metaphysics (L. B. Geiger, Cornelio Fabro, R. J. Henle), and the primacy of *esse,* the fundamental act of being (Gilson, Jacques Maritain, G. P. Klubertanz). Because *esse,* which simply means "to be," is sometimes translated as "existence," this second point of emphasis is called by some writers the existentialism of Thomistic metaphysics. It has little, however, to do with present-day existentialism. A major treatment of metaphysical problems is to be found in Aquinas's long Commentary on Aristotle's *Metaphysics,* but here again the problem is to decide how much is Thomistic. Some very competent scholars (Pegis, Gilson) regard this work as a restatement of Aristotelianism; others (De Koninck, Herman Reith) consider the Commentary to be a key exposition of Aquinas's own metaphysics. It is admitted by all that there are some explanations in it that are not found in Aristotle.

Metaphysics, for Aquinas, was the effort to understand reality in general, to find an ultimate explanation of the manifold of experience in terms of the highest causes. His predecessors had variously described the subject matter of this study as existing immaterial substances, as the most universal and common aspects of being, as the first causes of all things, and as the divine being in itself. Commenting on these opinions in the prologue to his Commentary on the *Metaphysics,* Aquinas remarked: "Although this science considers these items, it does not think of each of them as its subject; its subject is simply being in general." In this sense, he called the study of being "first philosophy."

ANALOGY. It is distinctive of Aquinas's thought to maintain that all existing realities, from God down to the least perfect thing, are beings—and that "being" has in this usage an analogical and not a univocal meaning. In a famous passage (*In I Sententiarum* 19, 5, 2, ad 1) Aquinas describes three sorts of analogy: one in which a given perfection is present in one item but only attributed to another; one in which one perfection exists in a somewhat different way in two or more items; and one in which some sort of remote resemblance or community is implied between two items which have no identity either in existence or in signification. "In this last way," Aquinas adds "truth and goodness, and all things of this kind, are predicated analogously of God and creatures." In later works the notion of proportionality is introduced to develop the concept of the analogy of being. Vision in the eye is a good of the body in somewhat the same way that vision in the intellect is a good of the soul. Similarly, the act of being in a stone is proportional to the act of being in a man, as the nature of a stone is proportional to the nature of man. Whereas some interpreters feel that the analogy of proportionality is the central type of analogy of being, others insist that Aquinas used several kinds of analogy in his metaphysics.

BEING AND ESSENCE. One early but certainly personal presentation of the metaphysics of Aquinas is to be found in the brief treatise *De Ente et Essentia,* which was strongly influenced by Avicenna. His usage of basic terms of analysis, such as being (*ens*), essence (*essentia*), nature, quiddity, substance, accident, form, matter, genus, species, difference, immaterial substance (*substantia separata*), potency, and act, is clearly but rather statically defined in this *opusculum.* Additional precisions, particularly on the meaning of element, principle, cause, and *esse,* are to be found in the companion treatise, *De Principiis Naturae.* A more dynamic approach to being and its operations is offered in the *Quaestiones Disputatae de Potentia Dei* and in Part I of the *Summa Theologiae.*

Fundamental in the metaphysical thinking of Aquinas is the difference between *what* a being is and the fact *that* it is. The first is a question of essence; the second is the act of being, *esse.* Essences are many (various kinds of things—stones, cows, air, men) and are known through simple understanding, without any necessity of adverting to their existence or nonexistence. For a thing *to be* is entirely another matter; the fact that something exists is noted in human experience by an act of judgment. Many essences of things are material, but there is nothing about *esse* that requires it to be limited to materiality. This proposition (to be is not necessarily to be material) is the "judgment of separation" (*In Boethii de Trinitate* V, 3). Many Thomists now regard it as a funda-

mental point of departure for Aquinas's metaphysical thinking.

There are also certain most general features of real beings that transcend all division into genera and species; these are convertible with metaphysical being. In other words, they are coextensive and really identical with being. Such transcendentals are thing (*res*), something (*aliquid*), one, true, good, and (according to some interpreters) beautiful. The more important of these transcendentals suggest that every being is internally undivided but externally distinct from all else (*unum*), that every being has some intelligible meaning (*verum*), and that every being is in some way desirable (*bonum*). The theory of transcendentals is much more expanded and stressed in later scholasticism than in Aquinas's own writings. He barely touches upon it in Questions I and XXI of *De Veritate* and in the discussion of God's attributes in *Summa Theologiae* (I, Ques. 6, 11, 16).

POTENCY AND ACT. Potency and act are important principles in Aquinas's metaphysical explanation of the existence and operation of things. In *De Potentia Dei* (I, 1) Aquinas pointed out that the name "act" first designated any activity or operation that occurs. Corresponding to this sort of operational act is a dual meaning of potency (or power). Consider the activity of sawing wood: The passive potency of wood to be cut is required (water, for instance, cannot be sawed); also required is the active potency of the sawyer to do the cutting. In addition, in the same text, Aquinas says that the notion of "act" is transferred to cover the existence of a being. Essential potency, the metaphysical capacity to exist, would correspond to this act of being (*esse*). In this way the theory of act and potency was applied to all levels of being. At the highest level, God was described as Pure Act in the existential order, but this did not prevent Aquinas from attributing to God an active potency for operating.

FINALITY. Still another dimension of metaphysical reality, for Aquinas, was that of finality. He thought of all activities as directed toward some end or purpose, a basic assumption in Aristotle. But Aquinas developed this tendential, vector characteristic of being and applied it to the inclination of possible beings to become actual. The finality of being, in Thomism, is that dynamic and ongoing inclination to be realized in their appropriate perfections that is characteristic of all realities and capacities for action. In this sense the finality of being is an intrinsic perfectionism in the development of all beings. Aquinas also held that all finite beings and events are tending toward God as Final Cause. This is metaphysical finality in the sense of order to an external end. This theme runs through Book III of *Summa Contra Gentiles.*

PHILOSOPHY AND GOD

The consideration of the existence and nature of God was approached by Aquinas both from the starting point of supernatural revelation (the Scriptures), which is the way of the theologian, and from the starting point of man's ordinary experience of finite beings and their operations, which is the way of the philosopher: "The philosophers, who follow the order of natural cognition, place the knowledge of creatures before the divine science; that is, the philosophy of nature comes before metaphysics. On the other hand, the contrary procedure is followed among the theologians, so that the consideration of the Creator precedes the consideration of creatures" (*In Boethii de Trinitate,* Prologue). In the same work (II, 3 c) we are told that the first use of philosophy in sacred doctrine is "to demonstrate items that are preambles to faith, such as those things that are proved about God by natural processes of reasoning: that God exists, that God is one," and so on.

Aquinas recognized two types of demonstration, one moving from cause to effects and the other from effects back to their cause. The arguments that he selected to establish that God exists use the second procedure and are technically called *quia* arguments. In other words, these proofs start with some observed facts of experience (all Aquinas's arguments to God's existence are a posteriori) and conclude to the ultimate cause of these facts. Well aware of his debt to his predecessors, Aquinas outlined three arguments for the existence of God in *De Potentia Dei* (III, 5 c). The first shows that, since the act of being is common to many existents, there must be one universal cause of all (Plato's argument, Aquinas noted); the second argument starts from the fact that all beings in our experience are imperfect, not self-moved, and not the source of their actual being, and the reasoning concludes to the existence of a "mover completely immobile and most perfect" (Aristotle's argument); the third argument simply reasons from the composite nature of finite beings to the necessary existence of a primary being in which essence and the act of existing are identical (Avicenna's proof). Aquinas felt that these two pagan philosophers and an Islamic thinker had successfully established the conclusion "that there is a universal cause of real beings by which all other things are brought forth into actual being."

THE "FIVE WAYS." The most famous of the arguments are the "Five Ways" (*Quinque Viae*) of reasoning to the conclusion that God exists (*Summa Theologiae* I, 2, 3, c). All these ways employ the principle of causality and start from empirical knowledge of the physical world. They are not entirely original with Aquinas, depending not only on Plato, Aristotle, and Avicenna but also on Augustine and especially on Moses Maimonides. The First Way begins with the point that things in the world are always changing or moving and concludes to the existence of one, first, moving Cause. The Second Way argues from the observation of efficient production of things in the universe to the need of an existing, first, efficient Cause. The Third Way reasons from the contingent character of things in the world (none of them has to be) to the existence of a totally different kind of being, a necessary one (which has to be). The Fourth Way argues from the gradations of goodness, truth, and nobility in the things of man's experience to the existence of a being that is most true, most good, and most noble. The Fifth Way starts from the orderly character of mundane events, argues that all things are directed toward one end (the principle of finality), and concludes that this universal order points to the existence of an intelligent Orderer of all things. At the end of his statement of each "way," Thomas simply said, "and this is what all men call God," or words to that effect. Obviously, he presupposed a common meaning of the word *God* in the dictionary or nominal sense. There is disagreement among interpreters as to whether the "ways" are five distinct proofs or merely five formulations of one basic argument. Most Thomists now favor the second view.

Aquinas favored the argument from physical motion (*prima autem et manifestior via est*). The *Summa Contra Gentiles* (I, 13) offers an extended version of this first argument and frankly indicates its relation to the ideas in the last books of Aristotle's *Physics.* The other four ways are but briefly suggested in the *Summa Contra Gentiles.* In another, much neglected, work (*Compendium Theologiae* I, 3) the first way is stated clearly and concisely. Before attempting to establish in detail the various attributes of God, such as divine unity, one should consider whether he exists. Now, all things that are moved must be moved by other things; furthermore, things of an inferior nature are moved by superior beings. (Aquinas's examples are chosen from thirteenth-century physics and astronomy, in which the four basic elements were thought to be under dynamic influence of the stars, and lower celestial bodies were considered to be moved about by those at a greater distance from Earth. How much of the force of this argument may depend on outmoded science is a matter of debate in present-day Thomism.)

Aquinas next argues that the process in which *A* moves *B, B* moves *C,* and so on cannot be self-explanatory. His way of saying this is "This process cannot go on to infinity." He concludes that the only possible explanation of the series of physical motions observed in the universe requires the acceptance of the existence of a different sort of "mover"—a being that is not moved by another, in other words, a first mover. This would have to be a real being, of course, and of a quite different nature from bodily things. He eventually suggests that this "first mover existing above all else" is what Christians call God.

In the same passage from the *Compendium,* two other facets of the argument from motion are introduced. First, Aquinas claims that all causes observed as acting in the physical universe are instrumental in character and must be used, as it were, by a primary agent. This primary agent is again another name for God. To suppose that the universe is self-explanatory is, to Aquinas, like thinking that a bed could be constructed by putting the tools and material together, "without any carpenter to use them." This is an important case of the conception of God as a divine craftsman. In the second place, this text suggests briefly that an infinite series of moved movers is an impossibility; the length of the series has nothing to do with its explanatory function, if all its members be finite. Finally, any such series requires a first mover (primary in the sense of causality, not necessarily of chronological priority). This first mover would be a Supreme Being. It is obvious that many of the attributes of God are already implied in the argument for divine existence.

KNOWLEDGE OF GOD. Regarding the nature and attributes of God, Aquinas's greatest emphasis fell on how little we really know about the Supreme Being. In a series of articles (*Summa Theologiae* I, 86–88) on the objects of human knowledge, he reiterated his position that man is naturally equipped to understand directly the natures of material things; further, that man is aware of his own psychic functions as they occur but that all man's understanding of the nature of his own soul, of immaterial substances such as angels, and of infinite immaterial being (God) is achieved by dint of discursive and indirect reasoning. There is, of course, a wide gap between material and immaterial substances. Yet both these types of finite beings fall within the same logical genus, as substances, and thus bodies and created spirits have some aspects in common. On the other hand, God is an immaterial being of an entirely different nature from that of

bodies or even of created spirits. Between God and creatures there is no univocal community: That is to say, God does not fall within the same genus, either real or logical, as any other being. Hence, God's nature transcends all species and genera. Man's natural knowledge of God's nature is therefore very imperfect, achieved by negating various imperfections found in finite beings: Thus, God is not in time, not in place, not subject to change, and so on. Furthermore, man may reach some semipositive knowledge of God by way of analogy: Thus, God is powerful but not in the finite manner of other beings; he is knowing, willing, and so on.

PROVIDENCE. Divine providence is that attribute of God whereby he intelligently orders all things and events in the universe. As Aquinas explained it in the *Summa Contra Gentiles* (III), God both establishes the plan (*ratio*) in accord with which all creatures are kept in order and executes this plan through continued governance of the world. Literally, providence means "foresight," and this required Aquinas to face certain problems traditionally associated with any theory of divine foreknowledge. First of all, he insisted that such a view of divine providence does not exclude chance events from the universe. In one sense, a chance event occurs apart from the intention of the agent. However, what is intended by one agent may involve another agent who is unaware of the intention of the first. Hence, a plurality of real but imperfect agents sets the stage for chance: God knows this and permits it to occur.

EVIL. In the *Quaestiones Disputatae de Malo* and elsewhere Aquinas agreed with Augustine that evil (both physical and moral) is a privation of goodness, of perfection, in being or in action. This does not deny the fact that evil really occurs but asserts that it is like a wound in being (the phrase is Maritain's); and, like any defect, evil is important by virtue of what is lacking. As to why a perfectly good God will allow evil to occur, Thomas argued that the possibility of evil is necessary so that many goods may be possible. "If there were no death of other animals, there would not be life for the lion; if there were no persecution from tyrants, there would be no occasion for the heroic suffering of the martyrs" (*Summa Theologiae* I, 22, 2, ad 2).

FREEDOM. Aquinas also did not admit that divine foreknowledge is opposed to the exercise of human freedom. His explanation of this point (in *Summa Theologiae* I, 103, 7 and 8) is complicated and not easy to state briefly. In effect, human freedom does not imply absolute indeterminism (action that is uncaused). What a man does freely is caused by himself, as a knowing and willing agent. God makes man capable of choosing well or ill, permits man to do so freely, and knows what man will accomplish. What appears to be necessitated from one point of view may be quite contingent and free from another viewpoint. From God's vantage point in eternity, human actions are not affairs of past or future but are events within the all-inclusive present of a divine observer who witnesses these events but does not determine them.

ETHICS AND POLITICAL PHILOSOPHY

The foregoing problems and considerations fall within Aquinas's speculative philosophy. His practical philosophy, aimed at the intelligent performance of actions, is divided into ethics, economics (treating problems of domestic life), and politics. In all three areas the thinking is teleological; finality, purposiveness, and the means-end relation all are aspects of Thomistic teleology. Rationally controlled activities must be directed to some goal; they are judged good or bad in terms of their attainment of that goal and in terms of the means by which they attain (or fail to attain) that end.

Aquinas dealt with the theoretical analysis of ethical activities in a long series of works: the *Scriptum in IV Libros Sententiarum,* Book III; *Summa Contra Gentiles* III, 114–138; the *In X Libros Ethicorum; Quaestiones Disputatae de Malo*; and the *Summa Theologiae,* Part II. Most of these works take the approach of moral theology, viewing moral good and evil in terms of accord or discord with divine law, which is revealed in Scripture and developed and interpreted in Christian tradition. Thomas himself did not consider moral theology to be a part of philosophy, and it will not be further considered here, except as throwing incidental light on his ethical position.

VOLUNTARY ACTION. Aquinas's ethics consists of a study of good and evil in human conduct, from the point of view of man's achievement of ultimate happiness. Not all the actions in which man is involved are truly human but only those accomplished under control of man's intellect and will. The primary characteristic of human conduct, according to Aquinas, is not so much freedom as voluntariness. His description of voluntary activity is a development of the teaching of Aristotle. Several factors are required for a voluntary action. There must be sufficient knowledge on the part of a moral agent that a given action is within his power; he cannot be entirely ignorant of the kind of action that he is performing or of the

means, circumstances, and end of his action. Violence, under certain conditions, modifies the voluntariness of one's actions—as do certain kinds of uncontrollable feelings. Furthermore, as Aquinas saw it there are two opposites to what is voluntary. The "involuntary" is a contrary: It represents a diminution of voluntariness. Thus, an action that is partly involuntary is also partly voluntary and is, to a greater or lesser extent, imputable to the agent. On the other hand, the "not-voluntary" is the contradictory of what is voluntary, and an agent who is not voluntary is not morally responsible for his action.

NATURAL LAW. Most surveys of ethical theories classify Aquinas's ethics as a natural law theory. He described natural law as a rational participation in the eternal law of God and suggested that all men have a sufficient knowledge of what is morally right (the *justum*) to be able to regulate their own actions. In a famous passage (*Summa Theologiae* I–II, 94, 2) Aquinas explained the way in which he thought that rules of natural law are known. The judgment of *synderesis* (an intellectual quality enabling any man to intuit the first principle of practical reasoning) is simply the proposition "Good should be done and sought after; evil is to be avoided." (Most modern Thomists take this rule as a formal principle in the Kantian sense, requiring further knowledge to fill in the content of specific moral rules.) Aquinas then proceeded to describe three kinds of inclinations natural to man: that of man's substantial nature toward the conservation of its own existence and physical well-being, that of man's animal nature to seek such biological goods as sexual reproduction and the care of offspring, and that of man's reason whereby he tends toward universal goods, such as consideration of the interests of other persons and the avoidance of ignorance. All three kinds of inclinations are presented as natural and good, provided they are reasonably pursued. They form the bases from which one may conclude to a number of rules of natural moral law. Aquinas never attempted to make an exhaustive listing of the precepts of such a law; nor did he consider such a codification advisable.

In point of fact, the natural law approach to moral theory is not the only, and not the best, classification of Aquinas's ethics. Particularly in view of various shifts in the meaning of "law" since the time of Aquinas (notably a growing stress on law as a fiat of legislative will), it can be positively misleading to limit Aquinas's ethics to a natural law position. He defines law in general as "any ordinance of reason that is promulgated for the common good by one who has charge of a community" (*Summa Theologiae* I–II, 90, 4 c). "Reason" is the key word in this definition. Right reason (*recta ratio*) is the justification of ethical judgment in Aquinas's thought. "In the case of volitional activities, the proximate standard is human reason (*regula proxima est ratio humana*) but the supreme standard is eternal law. Therefore, whenever a man's action proceeds to its end in accord with the order of reason and of eternal law, then the act is right; but when it is twisted away from this rightness, then it is called a sin" (21, 1 c).

REASON, GOODNESS, AND JUSTICE. Thomistic ethics requires a person to govern his actions as reasonably as he can, keeping in mind the kind of agent that he is and the position that he occupies in the total scheme of reality. Man's own good is achieved by the governance of his actions and feelings under rational reflection—and God does not require anything else. "For we do not offend God, except by doing something contrary to our own good" (*Summa Contra Gentiles* III, 121–122). It is a part of being reasonable to respect the good of others. The moral good, then, is not so much what men are obligated to do by an all-powerful legislator; rather, it is that which is in accord with the reasonable perfecting of man. In becoming a better agent within himself, man is making himself more fit for ultimate happiness and for the vision of God. This kind of ethics resembles a self-perfectionist theory, without idealist overtones.

Aquinas based much of his teaching on ethical rules on the theory of natural justice found in Book V of the *Nicomachean Ethics.* All things have specific natures that do not change: Dogs are dogs and stones are stones. Certain functions are taken as natural and appropriate to given natures: Eating is an act expected of a dog but not of a stone. Human nature shares certain functions with the higher brutes but is distinguished by the performance of rational activities. Some of these typical functions are always the same in relation to man's nature and ethical rules pertaining to these do not change. Aquinas's example of such an immutable rule of justice is simply "Theft is unjust." Other ethical judgments, however, are not essential to justice (for example, detailed ordinances that contain many variable factors); these secondary rules are by no means absolute and immutable. Examples would be rules concerned with taxation, buying and selling, and other such circumstantially variable regulations. Moral law is composed of both types of rules and is neither absolute nor immutable in all its requirements.

CONSCIENCE. In *De Veritate* (XVII) Aquinas referred to moral conscience as a concrete intellectual judgment whereby the individual agent decides for himself that a

given action or feeling is good or bad, right or wrong, to be done or not to be done. Conscience was not considered a special power or moral sense, nor was it viewed as the source of universal moral convictions. For Aquinas it was simply a man's best practical judgment concerning a concrete moral problem. As such, moral conscience is a person's internal guide to good action; one acts immorally in going against his conscience, for it is his best judgment on a matter. If it is not his best judgment, then the person is clearly required to make a better effort to reach a conscientious decision. Reasonable consideration of a proposed action includes thinking of the kind of action that it is (the formal object), the purpose to which it is directed (the end), and the pertinent circumstances under which it is to be performed. These three moral determinants were used by Aquinas to complete the theory of right reasoning in *De Malo* (II, 4 c, ad 2, ad 5).

FAMILY. Aquinas also considered man in his social relations. In the *Summa Contra Gentiles* (III, cc. 122–126) the family is regarded as a natural and reasonable type of small society, designed to provide for the procreation and raising of children and for the mutual good of husband and wife. (The material on matrimony in the so-called *Supplement* to the *Summa Theologiae* was excerpted from Book IV of the *Scriptum in IV Libros Sententiarum* and does not represent Aquinas's mature thought.) The main reason why people get married, Aquinas thought, is to raise children, so his approach to the family was child-oriented. There should be but one husband and wife in a family; they should stay together until the children are fully grown and educated; they should deal honestly and charitably with each other as marriage partners. Many of Aquinas's arguments for monogamy and the indissolubility of the marriage bond are but restatements of similar reasonings in Aristotle's *Politics.*

POLITICAL THEORY. Aquinas's family, living in southern Italy, had been closely allied with the imperial government: His father and at least two of his brothers were in the service of Emperor Frederick II. Aquinas thus grew up with monarchic loyalties. However, early in life he joined the Dominicans, a religious community remarkable for its democratic and liberal practices. As a result Aquinas's political philosophy (in *De Regno,* in *In Libros Politicorum,* and in *Summa Theologiae,* I–II, passim) stressed the ideal of the limited monarchy, or that kind of state which Aristotle had called the *politeia.* The purpose of the state is described as to provide for temporal peace and welfare. Political society is quite different from ecclesiastical society (the church), whose end is otherworldly. Here again Aquinas always stressed the central role of reason: "Divine justice (*ius divinum*) which stems from grace does not cancel human justice which comes from natural reason." There is no detailed theory of government in Aquinas's writings.

ART AND AESTHETICS

In his theory of art Aquinas was quite abstract and intellectualistic, taking Aristotle's *Rhetoric, Poetics,* and *Nicomachean Ethics* (Book VI) as his major sources. He used a new awareness of the spiritual and moral dimensions of the beautiful, found seminally in the mystical Neoplatonism of Dionysius the Pseudo-Areopagite, to develop the fragmentary aesthetics of Aristotelianism. Most of these precisions are found in Aquinas's commentary on the fourth chapter of Dionysius's *De Divinis Nominibus.*

Art is understood to be a special habit, or acquired skill, of the practical intellect, which is simply man's possible intellect applied to problems of action. Prudence, the key practical habit in moral discourse, is defined as right reason in doing things (*recta ratio agibilium*). Similarly, art is defined as right reason in making things (*recta ratio factibilium*). These two practical habits are not confused. Elsewhere it is explained: "The principle of artifacts is the human intellect which is derived by some sort of similitude from the divine intellect, and the latter is the principle of all things in nature. Hence, not only must artistic operations imitate nature but even art products must imitate the things that exist in nature" (*In I Politicorum* 1). Some artifacts are merely useful; others may be beautiful; and still others may exist only in the order of thought (Aquinas took seriously the dictum that logic is an art).

He regarded the beautiful and the good as really identical but insisted that they differ in their formal meanings (*rationes*). Where the good is simply that which all desire, the beautiful is that which gives pleasure when perceived (*quod visum placet*). Three aspects of the beautiful are distinguished: integrity (*integritas sive perfectio*), due proportion (*debita proportio sive consonantia*), and brilliance (*claritas*). Each of these aesthetic factors is taken as capable of variation in degree and appeal.

These notions on the general meaning of Beauty were used not to describe the attraction of a life of sacrifice but of spiritual perfection as a member of a religious community, such as the Dominicans. "In fact," Aquinas wrote, "there are two kinds of beauty. One is spiritual and it consists in a due ordering and overflowing of spiritual goods. Hence, everything that proceeds from a lack of spiritual good, or that manifests intrinsic disorder, is ugly.

Another kind is external beauty which consists in a due ordering of the body" (*Contra Impugnantes Dei Cultum et Religionem* 7, ad 9). He was actually defending the practice of begging, as used in the mendicant orders. Aquinas agreed that there is something distasteful about begging but argued that it is an admirable exercise of humility, when religiously motivated. Here again the concept of purpose, teleological order, is central.

Metaphysical participation recurs as a key theme in Aquinas's discussion of the manner in which the manifold of creation shares in the transcendent beauty of God. All lower beauties are but imperfect manifestations of one highest *pulchritudo.* This is Dionysian mystical aesthetics and is presented in *In Dionysii de Divinis Nominibus* (IV, 5–6).

AUTHORITY AND INFLUENCE

Aquinas has been given a special position of respect in the field of Catholic scholarship, but this does not mean that all Catholic thinkers agree with him on all points. Within three years of his death a number of propositions closely resembling his philosophic views were condemned as errors by Bishop Tempier of Paris. This episcopal condemnation was formally revoked in 1325. Thomistic thought met much criticism in the later Middle Ages. Since the Renaissance nearly all the popes have praised Aquinas's teaching; the one who provided for the first collected edition of his works (St. Pius V) also did the same for St. Bonaventure, a Franciscan, and proclaimed both Doctors of the Church. In the ecclesiastical law of the Catholic Church, revised in 1918, canon 589:1 states that students for the priesthood are required to study at least two years of philosophy and four of theology, "following the teaching of St. Thomas." Further, canon 1366:2 directs professors in seminaries to organize their teaching "according to the method, teaching and principles of the Angelic Doctor."

Actually, Thomism has never been the only kind of philosophy cultivated by Catholics, and from the fourteenth century to the Enlightenment, Thomism was rivaled and sometimes obscured by Scotism and Ockhamism.

In 1879, with the publication of the Encyclical *Aeterni Patris* by Pope Leo XIII, the modern revival of Thomism started. While this document praised Thomism throughout, Pope Leo added this noteworthy qualification: "If there be anything that ill agrees with the discoveries of a later age, or, in a word, improbable in whatever way—it does not enter Our mind to propose that for imitation to our age" (Étienne Gilson, ed., *The Church Speaks to the Modern World,* New York, 1954, p. 50.)

In 1914 a group of Catholic teachers drew up a set of twenty-four propositions that, they felt, embodied the essential points in the philosophy of Aquinas. The Sacred Congregation of Studies, with the approval of Pope Pius X, published these "Twenty-four Theses" as clear expressions of the thought of the holy Doctor. (Original Latin text in *Acta Apostolicae Sedis* 6 [1914]: 384–386; partial English version in Charles Hart, *Thomistic Metaphysics,* Englewood Cliffs, NJ, 1959, passim.)

The first six theses attempt a formulation of the general metaphysical position of Aquinas. All beings are composed of potential and actual principles, with the exception of God, who is pure act. The divine *esse* (act of being) is utterly simple (that is, without parts or constituents) and infinite in every way. Other beings are composite; their acts of existing are limited in character and merely participated. In general, metaphysical being may be understood in terms of analogy: God's being and that of created things do not belong within the same genus, but there is some remote resemblance between divine and nondivine beings. To satisfy competing theories of analogy that developed in Renaissance Thomism, the theses describe this metaphysical analogy in terms of both attribution (following Francisco Suárez) and proportionality (following Cardinal Cajetan). The real distinction between essence and *esse* is stressed in the fifth thesis, while the difference between substance and accidents is stated in the sixth (accidents *exist in* some substance but never, in the natural course of things, exist by themselves). Marking a transition to special metaphysics (cosmology and philosophical psychology), the seventh proposition treats a spiritual creature as composed of essence and *esse,* and also of substance and accidents, but denies that there is any composition of matter and form in spirits.

A series of theses (VIII to XIII) describe bodily beings as constituted of prime matter and substantial form, neither of which may exist by itself. As material, bodies are extended in space and subject to quantification. Matter as quantified is proposed as the principle that individuates bodies. The location of a body in place is also attributed to quantity. Thesis XIII distinguishes nonliving from living bodies and makes the transition to a group of propositions concerned with human nature and its activities. The life principle in any plant or animal is called a soul, but, in the case of the human animal, the soul is found to be a principle of a very special kind. Theses XIV to XXI focus on the vital nature and functions of

man. His soul is capable of existing apart from the human body; it is brought into existence directly by God's creative action; it is without constituent parts and so cannot be disintegrated, that is to say, the human soul is immortal. Moreover, man's soul is the immediate source of life, existence, and all perfection in the human body. Subsequent propositions emphasize the higher human functions of cognition and volition, and they distinguish sensitive knowledge of individual bodies and their qualities from intellectual understanding of the universal features of reality. Willing is subsequent to intellectual cognition, and the free character of volitional acts of choice is strongly asserted.

The last three theses offer a summary of Aquinas's philosophic approach to God. The divine existence is neither directly intuited by the ordinary man nor demonstrable on an a priori basis. It is capable of a posteriori demonstration using any of the famous arguments of the Five Ways; these arguments are briefly summarized. Thesis XXIII reaffirms the simplicity of God's being and maintains the complete identity between the divine essence and *esse.* The final thesis asserts the creation by God of all things in the universe and stresses the point that the coming into existence and the motion of all creatures are to be attributed ultimately to God as First Cause.

These twenty-four theses represent a rigid and conservative type of Thomism. Many modern Catholic philosophers, while recognizing that these propositions do express some of the basic themes in the speculative thought of Aquinas, doubt that it is possible to put the wisdom of any great philosopher into a few propositions and prefer to emphasize the open-minded spirit with which Aquinas searched for information among his predecessors and approached the problems of his own day. After all, it was Aquinas who remarked that arguments from authority are appropriate in sacred teaching but are the weakest sort of evidence in philosophic reasoning.

See also Abelard, Peter; Aesthetics, History of; Albert the Great; Aristotelianism; Aristotle; Augustine, St.; Averroes; Avicenna; Being; Boethius, Anicius Manlius Severinus; Bonaventure, St.; Cajetan, Cardinal; Cicero, Marcus Tullius; Duns Scotus, John; Empiricism; Enlightenment; Essence and Existence; Eternal Return; Ethics, History of; Faith; Gilson, Étienne Henry; Ibn Gabirol, Solomon ben Judah; John of Damascus; Liber de Causis; Maimonides; Maritain, Jacques; Metaphysics, History of; Neoplatonism; Ockhamism; Plato; Pseudo-Dionysius; Roscelin; Ross, William David; Scientia Media and Molinism; Scotism; Stoicism; Suárez, Francisco; Taylor, Alfred Edward; Thomism; Universals, A Historical Survey; William of Champeaux.

Bibliography

WORKS BY THOMAS AQUINAS

Collected Editions

Opera Omnia, 25 vols. Parma, 1852–1873. Reprinted, New York, 1948–1950. A noncritical but almost complete edition.

Opera Omnia, 34 vols. Paris, 1871–1882. A noncritical but almost complete edition.

S. Thomae Aquinatis, Opera Omnia. Rome, 1882–. The Leonine Edition is recognized by scholars as the standard edition of the works of Aquinas.

Individual Works

Scriptum in IV Libros Sententiarum (1252–1257). 4 vols, edited by Pierre Mandonnet and M. F. Moos. Paris, 1929–1947. No English version.

De Ente et Essentia (1253). Vol. XLIII of Leonine edition. Translated by Armand Maurer as *On Being and Essence.* Toronto, 1949.

De Principiis Naturae (1253). Vol. XLIII of Leonine edition. Translated by V. J. Bourke in *The Pocket Aquinas,* edited by V. J. Bourke, 61–77. New York: Washington Square Press, 1960.

Contra Impugnantes Dei Cultum et Religionem (1256). Vol. XLI of Leonine edition. Translated by J. Proctor as *An Apology for the Religious Orders.* Westminster, MD, 1950.

Quaestiones Disputatae de Veritate (1256–1259). Vol. XXII of Leonine edition. Translated by R. W, Mulligan et al. as *Truth,* 3 vols. Chicago, 1952–1954.

Quaestiones Quodlibetales (1256–1272). Vol. XXV of Leonine edition. Partially translated by S. Edwards as *Quodlibetal Questions 1 and 2. Medieval Sources in Translation,* Vol. 27. Toronto, 1983.

In Librum Boethii de Trinitate Expositio (1257–1258). Vol. L of Leonine edition. Translated by R. E. Brennan as *The Trinity and the Unicity of the Intellect.* St. Louis, 1946.

In Librum Dionysii de Divinis Nominibus (1258–1265), edited by Ceslas Pera. Turin, 1950. Partially translated by V. J. Bourke in *The Pocket Aquinas,* 269–278. New York: Washington Square Press, 1960.

Summa de Veritate Catholicae Fidei Contra Gentiles (1259–1264), Vols. XIII–XV of the Leonine edition. Rome, 1918–1930. Also published in a one-volume "manual" edition. Turin and Rome, 1934. Translated by A. C. Pegis, J. F. Anderson, V. J. Bourke, and C. J. O'Neil as *On the Truth of the Catholic Faith,* 5 vols. New York, 1955–1957.

De Emptione et Venditione (1263). Vol. XLII of Leonine edition. Translated and edited by Alfred O'Rahilly as "Notes on St. Thomas on Credit." *Irish Ecclesiastical Record* 31 (1928): 164–165. Translation reprinted in *The Pocket Aquinas,* 223–225. New York: Washington Square Press, 1960.

Quaestiones Disputatae de Potentia Dei (1265). Edited by R. M. Spiazzi. Turin, 1949. Translated by Lawrence Shapcote as *On the Power of God.* Westminster, MD: Newman Press, 1952.

Commentaries on Aristotle (1265–1273), listed in next section.

De Regno (1265–1266). Vol. XLII of Leonine edition. Translated by G. B. Phelan and I. T. Eschmann as *On*

Kingship. Toronto: Pontifical Institute of Mediaeval Studies, 1949.

Compendium Theologiae (1265–1269). Vol. XLII of Leonine edition. Translated by Cyril Vollert as *Compendium of Theology.* St. Louis, 1957.

Summa Theologiae (1265–1273). Vols. IV–XII of the Leonine edition. Rome, 1918–1930. Reprinted, Turin, 1934. Translated by the English Dominican Fathers as *The Summa Theologica,* 22 vols. London, 1912–1936. Revision of part of the English Dominican Fathers' translation appears in *Basic Writings of Saint Thomas Aquinas,* edited by A. C. Pegis. New York: Random House, 1945.

Quaestiones Disputatae de Spiritualibus Creaturis (1267). Vol. XXIV of Leonine edition. Translated by John Wellmuth and Mary Fitzpatrick as *On Spiritual Creatures.* Milwaukee, 1949.

Quaestio Disputata de Anima (1269). Vol. XXIV of Leonine edition. Translated by J. P. Rowan as *The Soul.* St. Louis, 1949.

Quaestiones Disputatae de Malo (1269–1272). Vol. XXIII of Leonine edition. Translated by J. and J. Oesterle as *Disputed Questions on Evil.* Notre Dame, IN: University of Notre Dame Press, 1983.

Quaestiones Disputatae de Virtutibus (1269–1272). Edi Vol. XXIV of Leonine edition. Partially translated by J. P. Reid as *The Virtues in General.* Providence, RI, 1951.

In Job Expositio (1269–1272). In Vol. XVII of Leonine edition. Rome, 1962. Translated by A. Damico as *The Literal Exposition on Job: A Scriptural Commentary concerning Providence.* Classics in Religious Studies 7. Atlanta, 1989.

In Evangelium Joannis Expositio (1269–1272), edited by Raphael Cai. Turin, 1952. Partially translated by J. A. Weisheipl and F. R. Larcher as *Commentary on the Gospel of Saint John,* Part I. Albany, NY, 1980.

De Unitate Intellectus (1270). Edited by Vol. XLIII of Leonine edition. Translated by Sister Rose E. Brennan as *The Unicity of the Intellect.* St. Louis: Herder, 1946.

De Substantiis Separatis (1271). Vol. XL of Leonine edition. Translated by F. J. Lescoe as *Treatise on Separate Substances.* West Hartford, CT: Saint Joseph College, 1963.

De Aeternitate Mundi (1271). Vol. XLIII of Leonine edition. Translated by Cyril Vollert as *On the Eternity of the World.* Milwaukee: Marquette University Press, 1964.

In Librum de Causis (1271), edited by H. D. Saffrey. Fribourg, 1954. Translated by V. A. Guagliardo et al. as *Commentary on the Book of Causes.* Thomas Aquinas in Translation Series. Washington, DC, 1996.

Commentaries on Aristotle

In Libros Posteriorum Analyticorum. In Vol. I of the Leonine edition. Rome, 1882. Translated by Pierre Conway. Quebec, 1956.

In Libros de Anima. In Vol. XLV of the Leonine edition. Also edited by R. M. Spiazzi. Turin, 1955. Translated by Kenelm Foster and Silvester Humphries as *Aristotle's De Anima with the Commentary of St. Thomas.* London and New Haven, CT, 1951.

In Libros de Caelo et Mundo. In Vol. III of the Leonine edition. Rome, 1886. Also edited by R. M. Spiazzi. Turin, 1952. No English version.

In X Libros Ethicorum. In Vol. XLVII of the Leonine edition. Also edited by R. M. Spiazzi. Turin, 1949. Translated by C. I. Litzinger as *Commentary on the Nicomachean Ethics,* 2 vols. Chicago: Regnery, 1964.

In Libros de Generatione et Corruptione. In Vol. III of the Leonine edition. Rome, 1886. Also edited by R. M. Spiazzi. Turin, 1952. No English version.

In Libros Peri Hermeneias. In Vol. I of the Leonine edition. Rome, 1882. Also edited by R. M. Spiazzi. Turin, 1955. Translated by Jean Oesterle as *Aristotle on Interpretation—Commentary by St. Thomas and Cajetan.* Milwaukee: Marquette University Press, 1962.

In Libros de Memoria et Reminiscentia, et de Sensu et Sensato. In Vol. XLV of the Leonine edition. Also edited by R. M. Spiazzi. Turin, 1949. No English version.

In XII Libros Metaphysicorum. In Vol. XLVI of the Leonine edition. Also edited by R. M. Spiazzi. Turin, 1950. Translated by J. P. Rowan as *Commentary on the Metaphysics of Aristotle,* 2 vols. Chicago: Regnery, 1961.

In Libros Meteorologicorum. in Vol. III of the Leonine edition. Rome, 1886. Also edited by R. M. Spiazzi. Turin, 1952. No English version.

In VIII Libros Physicorum. In Vol. II of the Leonine edition. Rome, 1884. Also edited by P. M. Maggiolo. Turin, 1954. Translated by R. J. Blackwell and others as *Commentary on Aristotle's Physics.* New Haven, CT: Yale University Press, 1963.

In Libros Politicorum. In Vol. XLVIII of the Leonine edition. Also edited by R. M. Spiazzi. Turin, 1951. Translation of Book III, Lectures 1–6 by E. L. Fortin and Peter D. O'Neill in *Medieval Political Philosophy,* edited by Ralph Lerner and Muhsin Mahdi, 297–334. New York: Free Press of Glencoe, 1963.

WORKS ON THOMAS AQUINAS

Guides

Bergomo, Petri de. *Tabula Aurea in Omnia Opera S. Thomae Aquinatis.* Bologna, 1485. Reprinted in Vol. XXV of *S. Thomae, Opera Omnia.* Parma, 1873.

Busa, Roberto. *Index Thomisticus.* Stuttgart-Bad Cannstatt: Fromann-Holzboog, 1974–1980.

Deferrari, Roy J., Sister M. Inviolata Barry, and Ignatius McGuiness. *A Lexicon of St. Thomas Aquinas, Based on the Summa Theologica and Selected Passages of His Other Works.* Washington, DC: Catholic University of America Press, 1949.

Indices … in Summam Theologiae et Summam Contra Gentiles. In Vol. XVI of the Leonine edition. Reprinted with some omissions in *Editio Leonina Manualis.* Rome, 1948.

Schütz, Ludwig. *Thomas-Lexikon.* Paderborn, 1895. Reprinted New York, 1949.

Studies

Aertsen, Jan. *Nature and Creature: Thomas Aquinas' Way of Thought.* Leiden: Brill, 1988.

Anderson, J. F. *An Introduction to the Metaphysics of St. Thomas Aquinas.* Chicago: Regnery, 1953. Collection of basic texts with clear explanations; useful for the theory of analogy.

Bourke, V. J. *Aquinas' Search for Wisdom.* Milwaukee: Bruce, 1965. A factual biography excluding pious legends and situating Thomas's thought in the context of his life.

Burrell, David. *Knowing the Unknowable God: Ibn Sina, Maimonides, Aquinas.* Notre Dame: University of Notre Dame Press, 1986.

Callahan, Leonard. *A Theory of Esthetic according to St. Thomas.* Washington, DC: Catholic University of America Press, 1927. Rethinks the original theory but also interprets the fragmentary maxims of Aquinas on beauty and art.

Chenu, M. D. *Toward Understanding Saint Thomas.* Chicago: Regnery, 1964. Masterful introduction to the advanced study of the works of Thomas Aquinas.

Copleston, F. C. *Aquinas.* Harmondsworth, U.K.: Penguin, 1955. This is a well-written exposition of the philosophy of Thomas, with some effort to meet the criticisms of British analysts.

Davies, Brian. *The Thought of Thomas Aquinas.* Oxford: Clarendon Press, 1992.

De Koninck, Charles. "Introduction à l'étude de l'âme." *Laval théologique et philosophique* 3 (1947): 9–65. Reprinted in Stanislas Cantin, *Précis de psychologie thomiste.* Quebec, 1948. LXXXIII, 173.

Descoqs, Pedro. *Thomisme et scolastique.* Paris: G. Beauchesne, 1927. Representative Suarezian criticism of the metaphysics of Aquinas; insists that there are other valid positions in philosophy.

Duns Scotus, John. *De Primo Principio.* Translated and edited by Evan Roche. St. Bonaventure, NY: Franciscan Institute, 1949. Important for fourteenth-century criticism by Duns Scotus of the general metaphysics and natural theology of Thomas.

Fabro, Cornelio. *La nozione metafisica di partecipazione secondo S. Tommaso.* Milan, 1939; 2nd ed., Turin, 1950. Outstanding study of the metaphysics of Thomism, stressing the Platonic elements.

Finnis, John. *Aquinas: Moral, Political, and Legal Theory.* New York: Oxford University Press, 1998.

Flannery, Kevin. *Acts Among Precepts: The Aristotelian Structure of Thomas Aquinas's Moral Theory.* Washington DC: Catholic University of America Press, 2001.

Garrigou-Lagrange, Réginald. *Reality: A Synthesis of Thomistic Thought.* St. Louis: Herder, 1950. Adapted from a long article in the *Dictionnaire de théologie catholique,* this work represents a most conservative type of Thomism.

Geiger, L. B. *La participation dans la philosophie de saint Thomas.* Paris, 1942. Excellent discussion of the Platonic themes in Thomas's metaphysics; with Fabro and Geiger it has become clear that Thomism is not a "baptized" Aristotelianism.

Gilby, Thomas. *The Political Thought of Thomas Aquinas.* Chicago: University of Chicago Press, 1958. Work of popularization that offers, however, a rather balanced view of the monarchic and democratic elements in Thomas's ideas on government.

Gilson, Étienne. *The Christian Philosophy of St. Thomas Aquinas.* New York: Random House, 1956. With "Catalogue of St. Thomas's Works" by I. T. Eschmann. A highly esteemed exposition; the French original followed the order of the *Summa Contra Gentiles*; this English revision stresses the idea that Thomism is a Christian philosophy.

Grabmann, Martin. *Die Werke des hl. Thomas von Aquin.* 3rd ed. Münster, 1949. The most complete study of the chronology and authenticity of the writings of Aquinas.

Henle, R. J. *Saint Thomas and Platonism.* The Hague: Nijhoff, 1956. Good example of a careful textual study using the method of parallel passages taken in chronological sequence.

Hughes, Christopher. *On a Complex Theory of a Simple God: An Investigation in Aquinas' Philosophical Theology.* Ithaca: Cornell University Press, 1989.

Jaffa, H. V. *Thomism and Aristotelianism. A Study of the Commentary by Thomas Aquinas on the Nicomachean Ethics.* Chicago: University of Chicago Press, 1952. A sharp criticism of Aquinas as an interpreter of Aristotle's practical philosophy.

Kenny, Anthony, ed. *Aquinas: A Collection of Critical Essays.* Garden City: Doubleday, 1969.

Kenny, Anthony. *Aquinas on Mind.* London: Routledge, 1993.

Klubertanz, G. P. *St. Thomas Aquinas on Analogy.* Chicago: Loyola University Press, 1960. Advanced and difficult study of all that Aquinas has written on metaphysical analogy; stresses the pluralism inherent in the theory: several types of analogy are at work.

Kluxen, Wolfgang. *Philosophische Ethik bei Thomas von Aquin.* Mainz, 1964. Important discussion of the relation between moral theology and ethics.

Kretzmann, Norman. *The Metaphysics of Creation: Aquinas's Natural Theology in Summa Contra Gentiles II.* Oxford: Clarendon Press, 1999.

Kretzmann, Norman. *The Metaphysics of Theism: Aquinas's Natural Theology in Summa Contra Gentiles I.* Oxford: Clarendon Press, 1997.

Kretzmann, Norman, and Stump, Eleonore, eds. *The Cambridge Companion to Aquinas.* Cambridge, U.K.: Cambridge University Press, 1993.

Lonergan, Bernard. *Verbum: Word and Idea in Aquinas.* Edited by David Burrell. Notre Dame, IN: University of Notre Dame Press, 1967.

Maritain, Jacques. *The Angelic Doctor: The Life and Thought of St. Thomas Aquinas,* rev. ed. New York, 1958. Popular work stressing papal approval of Thomistic philosophy and the special status of Aquinas in Catholic educational programs.

Maritain, Jacques. *Art and Scholasticism.* New York: Scribners, 1930. An almost classic account of the aesthetic theory of Thomas, with some adaptation to artistic problems in the twentieth century.

Maurer, Armand, ed. *St. Thomas Aquinas, 1274–1974: Commemorative Studies,* 2 vols. Toronto: Pontifical Institute of Medieaval Studies, 1974.

McInerny, Ralph. *Aquinas and Analogy.* Washington, DC: Catholic University of America Press, 1996.

Meyer, Hans. *The Philosophy of St. Thomas Aquinas.* St. Louis: Herder, 1944. In the guise of an exposition of Aquinas's thought, this is actually a trenchant critique of Thomism.

Oesterle, John A. *Ethics: The Introduction to Moral Science.* Englewood Cliffs, NJ: Prentice-Hall, 1957. Representative textbook emphasizing the autonomy and strictly philosophical character of Thomistic ethics.

Owens, Joseph. *An Elementary Christian Metaphysics.* Milwaukee: Bruce, 1963. Advanced textbook in the tradition of the Christian philosophy interpretation of Aquinas.

Owens, Joseph. *St. Thomas Aquinas on the Existence of God: Collected Papers of Joseph Owens,* edited by John Catan. Albany: State University of New York Press, 1980.

Pasnau, Robert. *Thomas Aquinas on Human Nature: A Philosophical Study of Summa Theologiae Ia 75–89.* Cambridge, U.K.: Cambridge University Press, 2001.

Pieper, Josef. *Guide to Thomas Aquinas.* Translated by Richard and Clara Winston. New York: Pantheon, 1962. A readable and elementary introduction by a noted German Thomist.

Rahner, Karl. *Geist im Welt. Zur Metaphysik der endliche Erkenntnis bei Thomas von Aquin.* Innsbruck and Leipzig, 1939. Outstanding German study of the theory of knowledge.

Ramírez, J. M. *De Auctoritate Doctrinae Sancti Thomae Aquinatis.* Salamanca, 1952. Representative of the well-informed and broadminded views of some Spanish Dominicans on the status of Thomism in Catholic philosophy.

Rommen, Heinrich. *The State in Catholic Thought.* St. Louis: Herder, 1945. Situates Thomistic political philosophy in the broader context of Catholic thinking as a whole.

Reith, Herman. *The Metaphysics of St. Thomas Aquinas.* Milwaukee: Bruce, 1958. Standard exposition, not difficult to read. Includes 200 pages of quotations from many works of Thomas.

Ryan, J. K. "St. Thomas and English Protestant Thinkers." *New Scholasticism* 22 (1) (1948): 1–33; (2) (1948): 126–208. Excellent survey of the reactions to, and criticisms of, Aquinas in post-Reformation British thought.

Smith. Gerard. *Natural Theology: Metaphysics II.* New York: Macmillan, 1951. On Thomas's philosophical approach to God; Ch. 16 discusses divine foreknowledge and human freedom.

Stump, Eleonore. *Aquinas.* London: Routledge, 2003.

Torrell, Jean-Pierre. *Saint Thomas Aquinas,* Vol. 1: *The Person and His Work.* Translated by Robert Royal. Washington, DC: Catholic University of America Press, 1996.

Ude, Johannes. *Die Autorität des hi. Thomas von Aquin als Kirchenlehrer.* Salzburg, 1932. Basic study of the question "Must Catholics agree in philosophy and theology with Thomas Aquinas?" The answer is "No."

Walz, Angelus, *Saint Thomas d'Aquin.* French adaptation by Paul Novarina. Louvain and Paris, 1962. One of the best biographies; this French printing has much more information than the English version, *Saint Thomas Aquinas,* translated by S. T. Bullough. Westminster, MD: Newman Press, 1951.

Westberg, Daniel. *Right Practical Reason: Aristotle, Action, and Prudence in Aquinas.* Oxford: Clarendon Press, 1994.

Wippel, John. *The Metaphysical Thought of Thomas Aquinas: From Finite Being to Uncreated Being.* Washington, DC: Catholic University of America Press, 2000.

Wittmann, Michael. *Die Ethik des hl. Thomas von Aquin.* Munich, 1933. Good historical study of Thomistic ethics in relation to Aristotle, the Stoics, and the Fathers of the Church.

WORKS ON THOMISM

Bourke, V. J. *Thomistic Bibliography, 1920–1940.* St. Louis: Modern Schoolman, 1945.

Bulletin thomiste (since 1921).

Ingardia, Richard. *Thomas Aquinas: International Bibliography 1977–1990.* Bowling Green, OH: Philosophy Documentation Center, 1993.

Mandonnet, Pierre, and Jean Destrez. *Bibliographie thomiste.* Le Saulchoir, Belgium: Revue des Sciences Philosophiques et Thélogiques, 1921; rev. ed., edited by M. D. Chenu. Paris: Vrin, 1960.

Répertoire bibliographique. Annual supplement to *Revue philosophique de Louvain* (Belgium).

Wyser, Paul. *Thomas von Aquin* and *Der Thomismus,* fascicles 13–14 and 15–16 of *Bibliographische Einführungen in das Studium der Philosophie.* Fribourg, Switzerland, 1948ff.

Vernon J. Bourke (1967)
Bibliography updated by Christian B. Miller (2005)

THOMASIUS, CHRISTIAN
(1655–1728)

Christian Thomasius was a philosopher and jurist and the first important thinker of the German Enlightenment. He was born in Leipzig, the son of the Aristotelian philosopher Jakob Thomasius, who had been a teacher of Gottfried Wilhelm Leibniz. Christian, after studying philosophy and law at the universities of Leipzig and Frankfurt an der Oder, began lecturing at Leipzig in 1682. His theological enemies forced him to move in 1690 to the Ritterakademie in Halle. He helped to found the University of Halle, became professor of law there in 1694, and later was Geheimrat (privy counselor) and rector of the university.

LAW AND THEOLOGY

Thomasius followed his father, as well as Hugo Grotius and Samuel von Pufendorf, in the study of natural law. He sought a foundation for law, independent of theology, in man's natural reason. Like Pufendorf he opposed the orthodox Lutheran view that revelation is the source of law and that jurisprudence is subordinate to theology. He held that law is based on common sense and on truths common to all religions. On the other hand, many precepts traditionally held to be absolute were only the result of the historical development of a given nation, subject to change and justifiable only in terms of the characteristics of that nation. Thomasius asserted the right of free and impartial interpretation of the Bible and of God's laws, reacting against orthodox Lutheran exegesis and the intricacies and dogmatism of scholastic theology. He condemned fanaticism and the persecution of heretics and preached toleration of differing religious beliefs.

Thomasius opposed the episcopal system of church government, which asserted the rights of consistories and of theological faculties in church affairs, and supported a territorial system of church government, in which the

government would have control of church administration but not of dogma. In dogma neither state nor consistories and faculties should have power; the latter should make decisions concerning dogma, but individual churches and Christians should be free to accept or reject them. Thomasius thus sought to break the power of the governing bodies of the church, which were dominated by intolerant orthodox Lutherans, and to subordinate the church to the government, which by natural law should be supreme within the state. It was these doctrines that forced Thomasius's expulsion from Leipzig and led to his reception at Halle by the Prussian government, which was more liberal in religious matters.

EDUCATION AND THE NATURE OF MAN

Thomasius held that philosophy should be practical and should concentrate on man, his nature, and his needs. He opposed the Aristotelian scholasticism of orthodox Lutheranism because its abstractions and speculative complexities were useless in life. His *Introductio ad Philosophiam Aulicam* (An Introduction to Philosophy for the Courtier; Leipzig, 1688) was in the tradition of Renaissance humanistic pedagogy. It advocated a worldly education intended to produce "courtiers" (politicians, diplomats, and bureaucrats) rather than the "pedantic" scholastic education of the universities. The German states established after the Thirty Years' War were organizing centralized governments and modern administrations on the French model, and they needed officials with the practical education Thomasius advocated. Thomasius's model was the education given in the German *Ritterakademien* (schools for the nobility), and he himself introduced this practical, worldly education into the teaching of the Halle faculty of law.

The *Introductio* was intended as the first of a series of texts furthering Thomasius's educational goals. In it Thomasius advocated eclecticism and disapproved of sectarianism and quarrels between schools of thought. He held that philosophy should be independent of revealed theology and founded on the observation of reality. Metaphysics was harmful and should be confined to a short terminological excursus. For Thomasius theoretical philosophy comprised natural theology, physics, and mathematics. The *Introductio* presented his theory of man and covered psychology and theory of knowledge, knowledge being obtained through the senses only. Thomasius was a nominalist, and he was skeptical about rationally proving God's existence. He closed with a summary of logic, both practical and theoretical. Thomasius continued the educational program of the *Introductio* in his *Einleitung zu der Vernunfft-Lehre* (Introduction to logic; Halle, 1691), *Einleitung zur Sitten-Lehre* (Introduction to ethics; Halle, 1692), *Ausübung der Vernunfft-Lehre* (Practical logic; Halle, 1693), and *Ausübung der Sitten-Lehre* (Practical ethics; Halle, 1696), all of which introduced the use of German into university teaching.

In the *Introductio* and other works Thomasius's eclecticism and opposition to dogmatism, his empiricism, his concentration on description of human nature and the giving of advice for practical behavior, are evident. His eclecticism and opposition to dogmatism was connected with the tradition of Peter Ramus that survived in the school of John Amos Comenius and with Thomasius's philosophical individualism. He often presented his doctrines as only hypothetical and spoke of "my own" philosophy, renouncing absolute truth. Thomasius's concentration on the practical was influenced by such writers as Pierre Charron and Baltasar Gracián. Besides his texts he wrote special works on "prudence" (*Klugheit, prudentia*), giving advice for persons in different situations and positions.

Thomasius held that logic should be simple, should avoid the scholastic syllogistic treatment, and should be based on personal experience. Its goal should be not only the demonstration but also the discovery of truth. In line with his empiricism and opposition to dogmatism, Thomasius wrote much on probability and combined his discussion of logic with psychology and sociology.

Thomasius believed that Christian ethics must be based on rational love. Love, in its different forms, is the basic impulse in man. The will is independent of reason and is the origin of evil.

PIETISM

About 1694 Thomasius underwent a personal religious and philosophical crisis. Influenced by certain Pietist thinkers, he lost faith in the natural goodness and intellectual power of man and held that virtue and truth could be reached only through God's grace, man being otherwise vicious and blind. He solemnly disavowed his former errors in a public confession. By 1705 Thomasius showed a renewed faith in human freedom and goodness and in the natural light. The period from 1694 to 1705 is known as Thomasius's Pietist period, but his acceptance of Pietism was eased by substantial similarities between his own views and those of the Pietists. Both opposed "pedantry," Aristotelianism, Lutheran orthodoxy, the episcopal system of church government, and intolerance;

both were also eclectic and empirical and avoided scholastic abstractions and theological subtleties. A personal acquaintance with the Pietist A. H. Francke played an important part in Thomasius's temporary conversion to other Pietist views.

METAPHYSICS

Thomasius's two works on metaphysics were published at Halle during his Pietist period, the *Confessio Doctrinae Suae* in 1695 and the *Versuch vom Wesen des Geistes* (An Essay on the Essence of Spirit) in 1699. Like Paracelsus, Valentin Weigel, Jakob Boehme, and others before him, Thomasius presented a mystical or theosophical variety of animism or vitalism. The world, both spiritual and material, is animated by a spirit created by God. Truth can be found only in the Bible as made clear by divine illumination. Although such views were held by some Pietists, they were not confined to them, and Thomasius continued to hold them after his Pietist period. Perhaps Thomasius's metaphysics was influenced not only by Pietism but also by the school of Comenius, who influenced Thomasius in other ways, and by the Hermetic school of medicine and chemistry, which had a mystically based experimental attitude. The latter possibility especially would explain Thomasius's combination of empiricism and a mystical metaphysics advanced only as a hypothesis.

INFLUENCE

Thomasius's most important followers were either Pietists or their sympathizers, and his views soon became the official Pietist philosophy. The theologian Joachim Lange in particular stressed Thomasius's Pietism and held that divine illumination was the only source of truth. By 1710 Thomasius's followers had displaced the Aristotelians in nearly all the German universities. Lange led the first attacks against the new doctrines of Christian Wolff, but Thomasius, true to his spirit of toleration, did not participate in the attack. Wolffianism became dominant after 1730, but a few Pietist centers remained. Later, the work of the Pietists A. F. Hoffmann and Christian August Crusius helped to bring about the renewal of German philosophy after 1760, which culminated in the critical philosophy of Immanuel Kant.

See also Aristotelianism; Boehme, Jakob; Charron, Pierre; Comenius, John Amos; Crusius, Christian August; Empiricism; Enlightenment; Gracián y Morales, Baltasar; Grotius, Hugo; Hermeticism; Holism and Individualism in History and Social Science; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Paracelsus; Pietism; Philosophy of Law, History of; Pufendorf, Samuel von; Ramus, Peter; Toleration; Wolff, Christian.

Bibliography

ADDITIONAL WORKS BY THOMASIUS

Institutiones Iurisprudentiae Divinae. Frankfurt and Leipzig, 1688.

Die neue Entfindung einer wohlgegründeten und für das gemeine Wesen höchstnötige Wissenschaft. Das Verborgene des Hertzens anderer Menschen auch wider ihren Willen aus der täglichen Conversation zu erkennen. Halle, 1692.

Kleine deutsche Schriften. Halle, 1701.

Fundamenta Iuris Naturae et Gentium. Halle, 1705.

Kurtzer Entwurff der politischen Klugheit. Halle, 1705.

WORKS ON THOMASIUS

Barnard, F. M. "The Practical Philosophy of Christian Thomasius." *Journal of the History of Ideas* 32 (1971): 221–246.

Battaglia, Felice. *Cristiano Thomasius*. Rome, 1935.

Bienert, Walther. *Der Anbruch der Christlichen deutschen Neuzeit dargestellt an Wissenschaft und Glauben des Christian Thomasius*. Halle: Akademischer, 1934.

Bloch, Ernst. *Christian Thomasius*. Berlin: Aufbau-Verlag, 1953.

Fleischmann, Max, ed. *Christian Thomasius, Leben und Lebenswerk*. Halle: Niemeyer, 1931.

Hunter, Ian. "Christian Thomasius and the Desacralization of Philosophy." *Journal of the History of Ideas* 61(4) (2000): 595–616.

Jaitner, W. R. *Thomasius, Rüdiger, Hoffmann und Crusius*. Bleicherode, 1939.

Kayser, R. *Christian Thomasius und der Pietismus*. Hamburg, 1900.

Link Salinger, R. "Christian Thomasius: Of Moral Philosophy and Natural Law." In *Philosophy and Culture.* Vol 2, edited by Venant Cauchy. Montreal: Montmorency, 1988.

Neisser, Liselotte. *Christian Thomasius und seine Beziehung zum Pietismus*. Munich, 1928.

Schneiders, Werner. *Recht, Moral und Liebe. Untersuchungen zur Entwicklung der Moralphilosophie und Naturrechtslehre des 17 Jahrhunderts bei Christian Thomasius*. Münster, 1961.

Schroder, Peter. "Thomas Hobbes, Christian Thomasius and the Seventeenth-Century Debate on the Church and State." *History of European Ideas* 23(2-4) (1997): 59–79.

Schubert-Fikentscher, Gertrud. *Unbekannter Thomasius*. Weimar: Böhlaus Nachfolger, 1954.

Wolf, Erik. *Grotius, Pufendorf, Thomasius*. Tübingen, 1927.

Giorgio Tonelli (1967)

Bibliography updated by Tamra Frei (2005)

THOMAS OF YORK
(1220/1225–1260/1269)

Thomas of York, the English metaphysician and theologian, joined the Franciscan order by 1245, and he became doctor of theology at Oxford in 1253. He was fifth lecturer to the Oxford Franciscans (1253/1254) and sixth lecturer at the Cambridge convent (1256/1257). Thomas was the protégé of both Adam Marsh and Robert Grosseteste, whose tradition he followed. He wrote a treatise, *Manus Quae contra Omnipotentem* (The hand which is raised against the almighty), supporting St. Bonaventure in the battle between seculars and mendicants at Paris.

His major work, *Sapientiale,* written between 1250 and 1260 and never finished, is the earliest known metaphysical *summa* of the thirteenth century. It makes use of all the major writers of antiquity, as well as the Muslim and Jewish philosophers (particularly Avicebron and Maimonides), the Church Fathers, and his immediate predecessors at Paris and Oxford. Although he presents all the important opinions on each point, he is not a mere compiler but an original and profound philosopher who had mastered the entire corpus of knowledge available.

In the *Sapientiale* he treats all the standard metaphysical problems, both general and specific (a distinction he seems to have been the first to make), from an essentially Augustinian standpoint. His theory of matter is eclectic: There is a universal matter that is pure potentiality, and matter understood simply as privation. Heavenly bodies, for example, lack the second kind. Because in act they are already everything they are capable of becoming, they are free of any privation. He subscribes to a modified form of Grosseteste's light metaphysics, including a form of corporeity that is present in every body. Since form is the principle of individuation, however, there must be a plurality of forms in any given body. (Thomas does not explicitly raise this question, but it is implicit in much that he says.) He is very clear, though, that the soul cannot be a form perfecting that of the body. It is itself composite and is related to the body "as a pilot is to a ship." The soul is able to gain knowledge by abstracting universals from singulars through sense (the complete universal can be known from one singular), but it gains more certain knowledge from above, receiving ideas from Ideas through interior illumination.

Thomas maintained the distinction in creatures between essence and existence, the latter characterized by composition from matter and form, and the mark of a creature's contingency. His emphasis on the contingency of creation prevented his arriving at a clear-cut assertion of the efficacy of natural causes, although he usually seems to favor this position.

Finally, Thomas was a vigorous proponent of what had become the typical Franciscan position since Grosseteste, denying the eternity of the world, of time, of matter, and of motion, and refusing any accommodation to the Aristotelian or Averroistic schools.

See also Augustinianism; Averroism; Bonaventure, St.; British Philosophy; Essence and Existence; Grosseteste, Robert; Ibn Gabirol, Solomon ben Judah; Maimonides; Metaphysics, History of; Patristic Philosophy.

Bibliography

An edition of the *Sapientiale* is being prepared by R. J. O'Donnell, C.S.B. Pending its appearance, see D. E. Sharp, *Franciscan Philosophy at Oxford in the Thirteenth Century* (Oxford: Oxford University Press, 1930), pp. 49–114; Felix Treserra, "De Doctrinis Metaphysicis Fratris Thomae de Eboraco," in *Analecta Sacra Tarraconensis* 5 (1929): 33–102; and A. B. Emden, *A Biographical Register of the University of Oxford to A.D. 1500* (Oxford: Clarendon Press, 1957–1959), Vol. III, pp. 2139–2140.

Richard C. Dales (1967)

THOMISM

The epithet "Thomist" has been applied since the fourteenth century to followers of St. Thomas Aquinas; the earlier "Thomatist," occasionally used, was dropped toward the end of the fifteenth century. The term has a different implication according to the three main historical periods that can be distinguished. First, until the beginning of the 1500s, during a period of vigorous Scholasticism and competition among several schools, Thomism stood in metaphysics for the doctrine of a composition of essence and existence in all created beings; and in noetics it opposed both nominalism and the Neoplatonic concept of illumination by the Ideas. Second, from the sixteenth until the eighteenth century Thomism flourished in the golden age of Spanish Scholasticism. (At this time Thomists unreservedly applied to theology the metaphysical concept of the premotion of all secondary causes by the first cause.) Third, beginning about the middle of the nineteenth century there was a revival of Thomism that was authoritatively endorsed by the Catholic Church. Since then it has been claimed for Thomism that it represents the *philosophia perennis* of the West; Thomists have engaged in many-sided dialogue with thinkers from other traditions and disciplines and

have been constructive in applying Thomistic principles to modern social and political problems.

We shall take these periods in order, noting beforehand that a unified philosophy, inspired by the writings of Thomas, persists throughout. In the philosophy of Thomas phenomenology is not divided from ontology; the world is real and composed of many real and distinct things, all deriving from one fount and all related by the analogy of being. Man is a single substance composed of body and soul; his knowledge begins from experience of the material world, and his understanding is developed through reason; his free activity determines his personal and eternal destiny.

THIRTEENTH TO SIXTEENTH CENTURY

When Thomas died in 1274, much of his teaching was still regarded as startling. Despite the affection in which he had been held (this was greater in the faculties of arts than in those of divinity) and despite his writings against the Latin Averroists, there developed a bitter opposition expressed in criticism and censure. It came from the representatives of the traditional Augustinian theology and was reinforced by the Franciscan masters. Conservative, yet by no means obscurantist, they included Thomas in their suspicions of what can be simplified as the "this-worldliness" of the new Aristotelianism. Étienne Tempier, bishop of Paris, was commissioned by Pope John XXII (Peter of Spain, the famous logician, who was an able natural philosopher) to investigate the charges against the new philosophy; he exceeded his instructions and in 1277, in a scissors-and-paste syllabus, he condemned 219 propositions, about a dozen of which can be traced to Thomas. In the same year Robert Kilwardby, the ex-provincial of the English Dominicans and now the archbishop of Canterbury, forbade the teaching of Thomas at Oxford, and his successor, John Peckham, acridly continued the same policy; they led the group called the *Cantuarienses.* As is evidenced in William de La Mare's list of correctives (*correctoria*) issued to be appended to Thomas's writings, many of the points at issue were highly technical, and some of them may now seem even trivial; the debate, much of which Thomas himself anticipated in his *Quaestiones Quodlibetales,* revolves round what to him were contrasts—but to his critics were conflicts—between nature and grace, reason and faith, determinism and freedom, the existence of the universe from eternity and its beginning in time, the soul as biological form and as spirit, and the role of the senses and of divine enlightenment in the acquisition of knowledge.

Although the censures had no force outside Paris and Oxford and the criticisms were more moderate in substance than they were in tone (they judged Thomas to be dangerous rather than heretical), his fellow Dominicans were quick to rally to his defense, to get the condemnations reversed and to correct the corrections, which they called corruptions. Thomas's old master, Albert the Great, so much the leader of the new movement that it has been called Albertino-Thomism, interposed at Paris; Pierre of Conflans, archbishop of Corinth, and Giles of Lessines remonstrated with Kilwardby; and Richard Clapwell, prior of Blackfriars, Oxford, progressively adopted Thomas's positions and stoutly maintained them against Peckham. The school was strengthened by a brilliant group of English and French Dominicans, and it was adopted by the Dominican order at successive general chapters. It could always count on support from the Roman Curia, which was favorably inclined toward Greek philosophy. The Ecumenical Council of Vienne (1311–1312) endorsed man's psychophysical unity, and in 1323 John XXII canonized Thomas and solemnly commended his doctrine. Henceforth he was a received authority.

Among the Thomists of these first fifty years John of Paris and Thomas Sutton were outstanding; other noteworthy teachers were Raymond Martin, a contemporary of Thomas who worked on the frontiers of Arabic science, William of Macclesfield, William of Hothun (archbishop of Dublin), Thomas Joyce (Jorz), Robert of Orford, Rambert of Bologna, Bernard de la Treille (Bernard of Trilia), Hervé de Nedellec, Nicholas Trivet, James of Lausanne, Ptolemy of Lucca, Peter de la Palu, James of Metz (uneasily attached to the school), and Remigio de Girolami, the master of Dante Alighieri. In their hands the distinctions between essence and existence, matter and form, and substance and accident became sharper, although some of these scholars were reluctant to go beyond Aristotle to support, as Thomas did, the concept of an act of a form. Of particular interest is a German group deriving more directly from Albert than from Thomas and imbued with strains of Neoplatonism from Proclus and Avicenna; within this group were Ulrich of Strasbourg, Dietrich of Vrieberg (Freiburg), Berchtold of Mosburg, and, most famous of all, Meister Eckhart, whose Thomism is not generally considered to have been unequivocal. All these men were Dominicans; the secular master Peter of Auvergne and the Augustinian friars Giles of Rome and James of Viterbo can also be ranged with them.

As the later Middle Ages drew on, the enterprise of integrating a wide-ranging philosophy in theology was succeeded by more piecemeal investigations, and the schools settled down to their own party lines with a sharpened logic but some loss of originality. In the rivalry between the Dominicans and the Franciscans, Thomism was matched against Scotism, and this set the tone of its development: In fact, however, as Dominic de Soto later acknowledged, the agreements between the two were more important than their differences. Moderate realism was represented at all the universities and adhered to at Louvain, at Cologne, and later at Heidelberg. Thomism itself must be reckoned a minority movement, and some prominent Dominicans did not belong to the school. Durandus of Saint-Pourçain steadily ran counter to Thomas's teaching, and the Cambridge Dominican Robert Holkot did not fall in with it. A central figure is John Capreolus, called the *Princeps Thomistarum,* whose writings are a mine of information on the disputes with Scotists and Ockhamists. Although Capreolus chose Thomas's "Commentary on the *Sentences*" for his expositions rather than the better organized *Summa Theologiae,* he, together with Serafino Capponi de Porrecta, bequeathed to their order the habit of systematically articulating the whole corpus of Thomas's teaching. Less confined to the classroom and closer to life and the historical movement of ideas was St. Antoninus, archbishop of Florence, the moralist who is a major authority for medieval economics.

The influence of the Renaissance was already beginning to make itself felt, and the first period of Thomism closed nobly in north Italy with Bartholomew of Spina, Crisostomo Javelli, Francis Sylvester (or Ferrariensis), and Thomas de Vio (or Cajetan). The last two, the classical commentators on the *Summa contra Gentiles* and the *Summa Theologiae,* respectively, were friends and opponents, particularly on the metaphysics of analogy. Both were responsive to the renewed vitality of Latin Averroism, and for them the unity of their school lay more in an inner consistency of approach than in a common subscription to a list of propositions, such as marked later Scholasticism when it had retreated or been banished from the profane world into the ecclesiastical academies. Cajetan, the master of a nervous style that fitted the subtle analysis at which he excelled, was a good scholar and a man of affairs. His standing in the school is second only to that of Thomas himself, although there is some question whether he was not a better Aristotelian than a Thomist. It is alleged that his emphasis on existence as the act of substance rather than on *esse* as the act of being may have encouraged the habit of discussing essences apart from existence, which was treated as a predicate.

SIXTEENTH TO NINETEENTH CENTURY

The second period, coterminous with the golden age of Spain, also had its origins in Burgundy and also declined through an inability to adjust to an expanding world outside its frontiers. In the fifteenth century Dominic of Flanders developed Thomas's exposition of the *Metaphysics,* and Peter Crockaert of Brussels, the master of Francisco de Vitoria (the father of international law), was the first of a great line of masters associated with the University of Salamanca. It was the faculty of this university that intervened with the Spanish government to humanize colonial policy. They forsook the crabbed angularities of fifteenth-century Scholasticism for a more flowing baroque style; at the same time, however, they found what they regarded as the formal logic of Aristotle to be a sufficient instrument for their debates, and the advances made on it (the *subtilitates anglicanae*) were neglected. Although they are chiefly famous as Tridentine divines, the theological questions that they considered—the relations of efficacious grace and free will, of authority and conscience—occasioned sustained philosophical discussion.

Among these sixteenth-century authors, the following are well worth study: Melchior Cano for scientific method and Bartholomew de Medina, Dominic de Soto, and Martin de Ledesma for moral theory. Dominic Báñez is much admired for his high Thomism in metaphysics and natural theology. These were Dominicans, but the best-known writer of the group is the Jesuit Francisco Suárez, who is impressive by virtue of the breadth of his interests and the organization of his voluminous writings, although strict Thomists would reckon him an eclectic and would think that he achieved his clarity by too concrete a habit of thought. The Jesuits were at this time taking the lead in higher education, and of all the orders they were the most aware of contemporary scientific research. Courses of philosophy began to be given apart from theology, and the teamwork of the Jesuits at Coimbra produced the volumes titled *Conimbricenses* (1592), and of the Carmelites at Alcalá de Henares those titled *Complutenses* (1624). In twentieth-century Thomistic studies John of St. Thomas perhaps became more influential than Cajetan, and his *Cursus Philosophicus,* digested in Josef Gredt's *Elementa Philosophica,* may be recommended as of lasting value.

Yet by the end of the seventeenth century Thomism was important only in the centers of ecclesiastical learning; it was part of the establishment, more honored, perhaps, than listened to. Its monument is the Casanata Library in Rome, founded with two chairs of Thomist exegesis. Its philosophy served mainly as a prolegomenon to theological studies and was conducted in the "essentialist" temper of Gottfried Wilhelm Leibniz and Christian Wolff. In this spirit Antoine Goudin wrote his significantly titled *Philosophia Juxta D. Thomae Dogmata* (Milan, 1676), which by 1744 had gone through fourteen editions. Salvatore Roselli's six-volume *Summa Philosophiae* (Rome, 1777) was written in response to the reiteration of the Dominican commitment to Thomas's doctrine made by the master general, John Thomas Boxadors. Both works influenced the revival of Thomism in the next century. But few Thomists took part in the dialogue of philosophers from René Descartes to G. W. F. Hegel, and the writings of the school were studied only by those with antiquarian tastes or a special interest in the history of philosophy.

NINETEENTH AND TWENTIETH CENTURIES

The situation began to change about the middle of the nineteenth century. A circle of teachers at Piacenza, Naples, and Rome who were dissatisfied with the eclectic doctrines that then served for clerical studies and were critical of the developed Kantianism of Georg Hermes, the accommodated Hegelianism of Anton Günther, the antirationalism of traditionalism, and the ontologism of Antonio Rosmini began to look to the synthesis of Thomas. The Dominicans themselves had remained faithful to Thomas, but their temper was somewhat rabbinical and concentrated on the letter of the text; and except in Spain and southern Poland they had been scattered in the troubled times after the French Revolution. At the beginning of the nineteenth century a secular canon, Vincenzo Buzzetti, inspired two brothers, Serafino and Domenico Sordi, who later became Jesuits, and Giuseppe Pecci, the brother of the future Leo XIII, to the work of the restoration of Thomism. They were joined by Gaetano Sanseverino, who contributed the five-volume *Philosophia Christiana* (Naples, 1853), and were supported by the influential Jesuit periodical *Civiltá cattolica.* The movement gathered strength with the affirmation of the rights of reason at the First Vatican Council (1869–1870) and with the teaching of two great professors at the Gregorian University, Matteo Liberatore and Josef Kleutgen, and of two Dominican cardinals, the Corsican Thomas Zigliara and the Spaniard Zefirín Gonzales. Finally, Leo Kill's encyclical *Aeterni Patris* (1879) sounded the recall to Thomas's basic doctrines in order to meet modern needs. Succeeding popes have reinforced this recommendation, not without embarrassment to those not wedded to Thomas's system, and even to those Thomists who would not have philosophy inculcated according to administrative needs. In practice, however, and despite the scares of the Modernist movement and the antimetaphysical temper since the 1940s, the injunctions have not proved irksome; and many forward-looking thinkers have discovered that Thomas was a benign and generous patron of their studies.

A history of neo-Thomism—the title is not relished by many in the school who do not see themselves committed to an absolute system—remains to be written. One characteristic of neo-Thomism has been its willingness to assimilate influence from outside its own tradition, which is a tribute to the depth and versatility of its principles. Another is that it has not been preoccupied with ecclesiastical matters; it inspired the social teaching of Leo XIII, with the result that many laypeople and statesmen have consulted it in developing the ideals and practice of Christian democracy. Nor has the conduct of speculation been reserved to clerics, and in the mid-twentieth century Thomism had no names more eminent than those of Jacques Maritain and Étienne Gilson. Although it appeals primarily to Catholics, its adherents are not necessarily Catholics, or even Christians. It presents no fixed image of conformity.

The Spanish works of high Thomism (the names of Norberto del Prado and Jaime Ramírez may be mentioned) have seemed to stand apart from the streams of contemporary thought, and the chief agencies that have taken Thomism into the world debate have been the University of Louvain and the French Dominicans. The Institut Supérieur at Louvain was founded in 1889 by Désiré Mercier, later cardinal, to bridge the gap between modern science and philosophy, particularly with respect to the problem of knowledge. In connection with this effort, the work of Joseph Maréchal was noteworthy. The French Dominicans have made contributions important both in critical research and in the popularization of Thomistic philosophy, and they have been alert to consider the most seemingly disparate interests; their periodicals, the *Revue des sciences philosophiques et théologiques* and the *Revue thomiste,* provide probably the best index to the activities of the school. From the universities of Munich and Münster has come important work, and the names of Martin Grabmann and Otto Geyer are illustrious. Other out-

standing figures are Réginald Garrigou-Lagrange of the University of St. Thomas in Rome and R. Welty and I. M. Bocheński of the University of Fribourg. A strong stream of Thomism is evident in the work of A. E. Taylor at Edinburgh, Kenneth Kirke at Oxford, E. L. Mascall at London, and Mortimer Adler at Chicago. Distinguished work comes from the Medieval Institute in Toronto, and there are flourishing centers of Thomistic study in Washington, D.C.; River Forest, Illinois; St. Louis; Montreal; and Sydney. The enumeration, however, is incomplete and perhaps invidious. The bibliographies of the *Bulletin thomiste* bear witness to a worldwide interest in Thomistic thought on the part of both philosophers and theologians.

See also Albert the Great; Aristotle; Augustinianism; Averroism; Avicenna; Báñez, Dominic; Capreolus, John; Cajetan, Cardinal; Dante Alighieri; Descartes, René; Eckhart, Meister; Essence and Existence; Garrigou-Lagrange, Réginald Marie; Giles of Rome; Gilson, Étienne; Hegel, Georg Wilhelm Friedrich; Holkot, Robert; John of Paris; John of St. Thomas; Kilwardby, Robert; Leibniz, Gottfried Wilhelm; Maréchal, Joseph; Maritain, Jacques; Medieval Philosophy; Mercier, Désiré Joseph; Neoplatonism; Ockhamism; Peckham, John; Proclus; Renaissance; Rosmini-Serbati, Antonio; Scientia Media and Molinism; Scotism; Soto, Dominic de; Suárez, Francisco; Taylor, Alfred Edward; Thomas Aquinas, St.; Ulrich (Engelbert) of Strasbourg; Vitoria, Francisco de; Wolff, Christian.

Bibliography

Bourke, V. J. *Thomistic Bibliography, 1920–1940.* St. Louis, 1945.

Dezza, Paolo. *Alle origini del Neotomismo.* Milan: Fratelli, 1940.

Gilson, Étienne. *History of Christian Philosophy in the Middle Ages,* 277–294, 361–436, 471–485. New York: Random House, 1955. This work contains excellent bibliographies.

Mandonnet, Pierre, and Jean Destrez. *Bibliographie thomiste.* Paris, 1921.

Roensch, F. J. *The Early Thomistic School.* Dubuque, IA: Priory Press, 1964.

Wulf, Maurice de. *Histoire de la philosophie médiévale,* 5th ed., Vol. II, 33–51, 112–151, 197–203, 272–277. Louvain: Institut Supérieure de Philosophie, 1925.

Thomas Gilby, O.P. (1967)

THOMISM [ADDENDUM]

The most important development in Thomism since the original entry has been increased interest in St. Thomas Aquinas among philosophers trained in the analytic tradition. The pioneer was Peter Geach, whose essay on "Aquinas" in *Three Philosophers* (1961) has proved to be seminal. Although often critical of what he takes to be Aquinas's positions, Anthony Kenny's numerous publications—covering such diverse philosophical topics as God, mind, and metaphysics—have been influential in making Aquinas more accessible.

The most comprehensive attempt to argue for the contemporary relevance of Aquinas to analytic philosophers is Eleonore Stump's wide-ranging *Aquinas* (2003). The emergence of philosophy of religion as a recognized discipline within analytical philosophy departments has generated greater interest in Aquinas among a wide variety of theists. Norman Kretzmann, in *The Metaphysics of Theism* (1997) and *The Metaphysics of Creation* (1999), has argued that Aquinas's natural theology as developed in the first three books of the *Summa contra gentiles* is the richest and most impressive resource for the development of a contemporary theistic metaphysics. David Burrell has repeatedly argued, especially in *Freedom and Creation in Three Traditions*(1993), that Aquinas is an important resource for philosophy of religion in an ecumenical spirit as modeled on Aquinas's own dialogue with Muslim and Jewish interlocutors.

Interest in Aquinas has also flourished in ethics. Alasdair MacIntyre, in *Three Rival Versions of Moral Inquiry* (1990), argues for the rational superiority of the Thomistic moral tradition to the failed legacy of the Enlightenment project and the incoherence of Friedrich Nietzsche's genealogy of morals, provoking a large body of secondary literature.

Thomists have traditionally sought to extract from Aquinas a natural-law ethic that could provide the foundation for arguments with those who do not share similar theological commitments. John Finnis's work, especially in *Natural Law and Natural Rights* (1980), is the most influential attempt to articulate a Thomistic theory of natural law that is more attractive to those who accept the modern starting point of individual natural rights. Finnis's argument that the first principles of practical reason indicate a number of irreducible and incommensurable goods as integral to human fulfillment has been criticized by other Thomists (for example, Russell Hittinger) on the grounds that it is incompatible with Aquinas's claim that the contemplation of God is constitutive of human flourishing.

It should be noted that Thomists trained in a more classically historical approach to Aquinas have made notable recent contributions. The works of John F. Wip-

pel and W. Norris Clarke in metaphysics are especially important. In noting this other strain within Thomism, we come to the abiding tension between traditional fidelity to the central commitments of Aquinas and the development of insights that can engage contemporary problems and modes of discourse. In the previous generation of Thomists, the battle was over whether Aquinas could be brought into dialogue with post-Kantian German philosophy; now the focus has shifted to analytic philosophy. Traditional Thomists worry that analytic readings of Aquinas distort his thought, through both the failure to understand it in its original context and the imposition of foreign metaphysical and epistemological dogmas. More analytically-minded Thomists worry that traditional approaches to Aquinas render his thought irrelevant.

See also Enlightenment; MacIntyre, Alasdair; Natural Law; Neo-Kantianism; Nietzsche, Friedrich; Thomas Aquinas, St.

Bibliography

Davies, Brian. *The Thought of Thomas Aquinas*. London: Oxford University Press, 1992.

Kerr, Fergus. *After Aquinas: Versions of Thomism*. London: Blackwell, 2002.

Kretzmann, Norman, and Eleonore Stump, eds. *The Cambridge Companion to Aquinas*. Cambridge, U.K.: Cambridge University Press 1993.

Shanley, Brian J. *The Thomist Tradition*. Dordrecht, Netherlands: Kluwer, 2002.

Brian Shanley, O.P. (2005)

THOMSON, JUDITH JARVIS
(1929—)

Judith Jarvis Thomson has made major contributions to moral theory and metaphysics. In addition to several books in these areas, she has written more than seventy articles on a range of topics, including action theory, philosophy of mind, and philosophy of science. She was educated at Barnard College, Cambridge University, and Columbia University, the last awarding her a doctoral degree in 1959. Since 1962, Thomson has taught at the Massachusetts Institute of Technology, where she became a full professor in 1969.

In moral theory, much of Thomson's work concerns what it is to have a moral right. Thomson's 1971 article "A Defense of Abortion"—an important contribution not only to ethics but also to feminist philosophy— revolutionized the abortion debate, which had previously focused largely on the question of whether the fetus has a right to life. Thomson grants, for the sake of argument, that the fetus has a right to life, but argues that it does not follow that abortion is impermissible. She asks you to imagine waking up in the hospital with your kidneys connected to the circulatory system of a famous violinist with a fatal kidney ailment; the violinist will die without the continued use of your body (no one else with the requisite blood type can be found). It is not obvious that you must continue to lend the violinist the support of your body; thus the fact that something has the right to life, together with the fact that it will die without the continued use of your body, does not obviously show that you must continue to lend it that support. Thus, in Thomson's words, "the right to life will not serve the opponents of abortion in the very simple and clear way" they thought it would.

Thomson's views about rights are further developed in her 1976 and 1985 essays "Killing, Letting Die, and the Trolley Problem" and "The Trolley Problem" (among other essays collected in *Rights, Restitution, and Risk* [1986]). These two essays focus on issues surrounding the problem, due to Philippa Foot (1967), of explaining why it would be impermissible for a surgeon to cut up one patient to save five who need organs, but permissible for a trolley driver to divert a runaway trolley onto a track where it will kill one person from a track where it would kill five. Foot's suggestion is that the duty not to kill is more stringent than the duty to save: Whereas the surgeon chooses between killing one and letting five die, and so should let five die, the trolley driver chooses between killing five and killing one, and so should kill one. Thomson objects that Foot's solution cannot account for the fact that it would be permissible for a *bystander* to flip the switch that diverts the trolley from killing the five, even though the bystander, like the surgeon, chooses between killing one and letting five die. Solving this problem—the Trolley Problem—requires a more subtle understanding of what rights are and which we have. Thomson's *The Realm of Rights* (1990) addresses these issues in detail.

Even if we grant that the distinction between killing and letting die does not solve the trolley problem, we may still think that the distinction is morally important. Many philosophers have thought that whether it is depends on what it consists in, metaphysically. Thomson's "Critical Study of Jonathan Bennett's *The Act Itself*" (1996) suggests that the metaphysical distinction is, roughly, that "there is a method in the making," whereas allowing

something to happen does not involve bringing it about by any method—"there is no *how* about it." "Physician Assisted Suicide: Two Moral Arguments" (1999) poses a serious challenge to those who think that while it is morally permissible for a doctor to accede to a patient's request to "let nature take its course"—either by not supplying, or by disconnecting life support—it is impermissible for a doctor to supply or administer a lethal drug at the patient's request. Along the way, Thomson makes the point that the killing/letting die distinction might itself be a moral distinction rather than a metaphysical distinction that makes a moral difference. In particular, it might be a necessary condition on an agent's letting someone die that she "have a liberty-right to act as she does."

The second major theme of Thomson's work in moral theory is her anticonsequentialism. One source of support for this comes from what she takes to be the moral theorist's data: our settled moral judgments about particular examples (for example, that the surgeon may not cut up the healthy patient to save five). Another, developed in "The Right and the Good" (1997) and *Goodness and Advice* (2001), is that the consequentialist's basic idea—that morality requires one to act in such a way as to make the world better than it otherwise would have been—is meaningless: there is no such relation as "better than." If there were such a relation, Thomson argues, then we could make sense of the question: Which is better, St. Francis or chocolate? But the question doesn't make sense: The goodness of a saint is an entirely different property from the goodness of chocolate. If all goodness is, as Thomson puts it, "goodness in a way," then the consequentialist owes us an account of what he or she means when he or she says that we ought to act so as to make the world better than it otherwise would have been. Thomson argues that no such account is available.

A third theme in Thomson's work in moral theory is her opposition to expressivist and relativist views about the content of moral claims. In *Moral Relativism and Moral Objectivity* (1996), coauthored with Gilbert Harman, Thomson defends moral objectivism, Harman defends moral relativism, and each replies to the other's arguments. One exchange concerns Harman's influential argument that moral theory cannot be justified in the same way that scientific theory can: our evidence that scientific hypotheses are true is that the truth of those hypotheses would explain what scientists observe, whereas moral hypotheses are explanatorily inert (1977). Thomson replies that our evidence that moral hypotheses are true is that they would be explained *by* observation: the data explain the hypotheses rather than the other way around.

In metaphysics, one strand of Thomson's work concerns questions about the persistence of material objects through change. "Parthood and Identity Across Time" argues against the thesis that objects have, in addition to spatial parts, temporal parts. According to Thomson, that thesis is absurd, because it implies that "[a]s I hold the bit of chalk in my hand, new stuff, new chalk keeps constantly coming into existence *ex nihilo*." "The Statue and the Clay" (1998) concerns the related issue of how artifacts are related to the material of which they are composed. Thomson argues that artifacts are not identical to but rather constituted by quantities of matter, and she provides a much-needed definition of the constitution relation, which previous writers on the topic had left unexplained.

The killing/letting die distinction is at the intersection of metaphysics and moral theory, along with many of Thomson's other interests—including causation, action, and agency. *Acts and Other Events* (1977) concerns events, their causes, and parts, and presents important challenges to rival theories of events and action. "The Time of a Killing" (1971) and "Causation: Omissions" (2001) also address metaphysical issues that bear on moral problems. Indeed, contemporary philosophy is indebted to Thomson for showing that metaphysics and ethics are often so intimately connected.

See also Abortion; Consequentialism; Euthanasia; Feminist Philosophy; Foot, Philippa; Harman, Gilbert; Objectivity in Ethics; Suicide.

Bibliography

WORKS BY JUDITH JARVIS THOMSON

"The Time of a Killing." *Journal of Philosophy* 68 (1971): 115–132.

"Killing, Letting Die, and the Trolley Problem." *The Monist* 59 (1976): 204–217.

Acts and Other Events. Ithaca: Cornell University Press, 1977.

"Parthood and Identity Across Time." *Journal of Philosophy* 80 (1983): 201–219.

"The Trolley Problem." *Yale Law Journal* 94 (1985): 1395–1415.

Rights, Restitution, and Risk. Cambridge, MA: Harvard University Press, 1986.

The Realm of Rights, edited by William Parent. Cambridge, MA: Harvard University Press, 1990.

"Critical Study of Jonathan Bennett's *The Act Itself*." *Nous* 30 (1996): 545–557.

"The Right and the Good." *Journal of Philosophy* 94 (1997): 273–298.

"The Statue and the Clay." *Noûs* 32 (1998): 149–173.

"Physician Assisted Suicide: Two Moral Arguments." *Ethics* 109 (1999): 497–518.

Goodness and Advice. Princeton, NJ: Princeton University Press, 2001.

"Causation: Omissions." *Philosophy and Phenomenological Research*, 2003.

WORKS ABOUT JUDITH JARVIS THOMSON

Bennett, Jonathan. *The Act Itself*. Oxford: Oxford University Press, 1998.

Foot, Philippa. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* (1967): 5–15.

Harman, Gilbert. *The Nature of Morality*. Oxford: Oxford University Press, 1977.

Harman, Gilbert, and Judith Jarvis Thomson. *Moral Relativism and Moral Objectivity*. Oxford: Blackwell, 1996.

Sider, Theodore. *Four-Dimensionalism*. Oxford: Oxford University Press, 2001.

"A Defense of Abortion." *Philosophy and Public Affairs* 1 (1971): 47–66.

Sarah McGrath (2005)

THOREAU, HENRY DAVID

(1817–1862)

Henry David Thoreau once described himself as "a mystic, a transcendentalist, and a natural philosopher." If this description does some justice to the extent of Thoreau's eclecticism, it nevertheless obscures those characteristics that made him important during his lifetime and still remain significant today, for Thoreau was an anarchist and revolutionary who created a highly articulate literature of revolt. Born at Concord, Massachusetts, the son of a pencil maker, Thoreau emerged from Harvard in 1837 with testimonials signed by Dr. George Ripley, Ralph Waldo Emerson, and the president of the university, all of whom attested, in glowing terms, to his moral and intellectual integrity. After a brief skirmish with school teaching, Thoreau became infected with the ideas of the New England transcendentalists, gave up all plans of a regular profession, and devoted himself to literature and the study of nature. His remarkable practical skills and intimate knowledge of the Concord countryside enabled him to earn his living independently, largely through pencil making and surveying, for the rest of his life.

From 1841 to 1843 Thoreau resided with Emerson. This brought his intellectual development roughly into line with the ideas of transcendentalists such as Amos Bronson Alcott, Margaret Fuller, and Ellery Channing, all of whom he came to know well. Thus, philosophically, Thoreau's reaction against the still fashionable sensationalism of John Locke and the theistic utilitarianism of William Paley was aided by ideas derived from the Scottish philosophers of common sense, who, in turn, formed a bridge to the idealism of Samuel Taylor Coleridge, Thomas Carlyle, and the Germans. Emerson also directed Thoreau to the English metaphysical poets and to Johann Wolfgang von Goethe. But despite this deep and undeniable cultural *rapprochement* it would be a misunderstanding to see Thoreau merely as Emerson's most eccentric disciple. Thoreau's individuality was maintained even at the intellectual level. He also studied New England history and legend, the life of the Indian, and early accounts of American travel and exploration; he probably had a better knowledge of the Greek and Latin classics than Emerson and certainly knew more about Oriental scriptures, of which he possessed an excellent collection. Above all, Thoreau's knowledge of natural history, motivated not so much by a desire for scientific understanding as by a need for concrete communion with nature, marks him off from the rest of Emerson's circle.

NATURE AND SOCIETY

Thoreau's writings everywhere bear the stamp of aboriginal practicality that also made him unique as a person. Society and nature were not for Thoreau, as they were for so many romantic thinkers, dialectical opposites whose inner identity was simply in need of philosophical explication. For him they involved a genuine contrast that he had personally experienced as a professional "saunterer" in and around Concord. Nature represented for Thoreau an "absolute freedom and wildness," whereas society provided "a freedom and culture merely civil." In his writing, as in his life, he attempted to implement the view that man should be regarded "as an inhabitant, or a part and parcel of Nature, rather than a member of society." It is only through a sustained involvement with the vast "personality" of nature that man can simplify his existence, clarify his senses, drive life into a corner, and reduce it to its lowest terms, thus achieving in practice a purer and tougher form of that self-reliance extolled, somewhat abstractly, by Emerson.

With these objects in mind, in the spring of 1845 Thoreau began building himself a hut on the shore of Walden Pond, a small lake then about a mile and a half south of Concord village. There he lived alone, with occasional visits to the village and from friends, until September 1847. His mode of life at the pond is described in *Walden, or Life in the Woods* (1854). For Thoreau Walden was an experiment in individualistic anarchism, just as Fruitlands and the Brook Farm community were for other transcendentalists attempts to revert to more "nat-

ural" modes of communal existence. But Thoreau had little confidence in collective protests against the existing social order, inspired by the doctrines of François Marie Charles Fourier. For him individual communion with nature was more fundamental than relationships with other men, even in societies where the worse forms of economic alienation have been overcome. For, unlike any social experience, the experience of nature becomes as much a discipline for the moral will as a stimulant to creative imagination. But essentially it is the spontaneity of wildness or nature that is to be favorably contrasted with the politico-economic organization of advanced European and New England societies. For, wrote Thoreau, "all good things are wild and free." The creative spontaneity of nature that is so crucial for man's spiritual well-being is embodied in all enduring products of culture—in the *Iliad* and *Hamlet,* in religious scriptures, in music, and especially in mythologies of all kinds. Commerce—"that incessant business"—and its political manifestations are indeed "vital functions of human society," yet a bare minimum of time should be consciously spent on them. They are "*infra*-human, a kind of vegetation," whose operations, like those of the human body, should be performed for the most part automatically, unconsciously. Far from viewing economic success alone as the sign of achievement or virtue, Thoreau believed that "to have done anything by which you earned money merely is to have been truly idle, or worse."

Despite the acquisitive basis of New England society, Thoreau saw a vision of true freedom in the expansion of the western frontier. For him the West was identical with the wild, and "wildness is the preservation of the world." These ideas, which constitute Thoreau's most persuasive expressions of revolt against bourgeois society, are best seen in his essays "Walking" (1862) and "Life without Principle" (1863).

REVOLUTION AND REFORM

Thoreau's essay "Civil Disobedience" (1849) has been the most influential of his works because of its overt political implications. It was, for example, a reading of this essay in 1907 that helped Mohandas Gandhi develop his own doctrine of passive resistance. Here Thoreau advocates active rebellion against the state. This involves what he calls "action from principle" on the basis of an intuitive perception of what is right, which is roughly equivalent to acting on the dictates of one's own conscience. He boldly asserts that "the only obligation which I have a right to assume is to do at any time what I think right." Action thus motivated "changes things and relations" and is therefore "essentially revolutionary." Radical social reforms, such as the abolition of slavery (for which Thoreau agitated throughout his life), can be effected not by petitions to elected representatives of government or by other indirect democratic means but only when each right-minded individual takes direct action on his own part. This would consist in withdrawing his allegiance "in person and property" from the government that supports or permits the abuse in question. Such is the form of "peaceful revolution" Thoreau himself attempted to put into practice by refusing to pay taxes. Despite its localized New England context and its relative lack of theoretical sophistication, it is possible to see Thoreau's doctrine of civil disobedience as historically linked, through the revolutionary element in European idealism, with the larger protest against the established order represented more notably by Søren Kierkegaard's *The Present Age* (1846) and the *Communist Manifesto* (1847). Like Karl Marx, Thoreau sought the dismantling of existing institutions in an attempt to discover an economy that would provide full human satisfaction. Yet like Kierkegaard he insisted on maintaining the uniqueness of the individual as the ultimate source of value; he attempted, however, to overcome the isolation his radical views forced upon him by means of a dialogue not with God but with nature.

See also Anarchism; Carlyle, Thomas; Channing, William Ellery; Coleridge, Samuel Taylor; Common Sense; Emerson, Ralph Waldo; Fourier, François Marie Charles; Goethe, Johann Wolfgang von; Kierkegaard, Søren Aabye; Locke, John; Marx, Karl; New England Transcendentalism; Paley, William; Sensationalism; Utilitarianism.

Bibliography

PRIMARY WORKS

The most complete edition of Thoreau's works is that published by Houghton Mifflin in 20 volumes: *The Writings of Henry David Thoreau* (Boston, 1906). See also *The Writings of Henry David Thoreau,* edited by Elizabeth Hall Witherall (Princeton, NJ: Princeton University Press, 1981–).

SECONDARY WORKS

H. S. Salt, *Life of Henry David Thoreau* (London: Bentley, 1890; reissued in an abridged version, 1896), is indispensable for an understanding of Thoreau's personality. For other details see Stanley Cavell, *The Senses of Walden,* expanded edition (Chicago: University of Chicago Press, 1980); Robert D. Richardson Jr., *Henry Thoreau: A Life of the Mind* (Berkeley: University of California Press, 1986); Robert Sattelmeyer, *Thoreau's Reading: A Study in Intellectual History with Bibliographical Catalogue* (Princeton, NJ: Princeton University Press, 1988); Leo Stoller, *After Walden: Thoreau's*

Changing Views on Economic Man (Stanford, CA: Stanford University Press, 1957); Floyd Stovall, ed., *Eight American Authors: A Review of Research and Criticism* (New York: Modern Language Association of America, 1956; reprinted, New York: Norton 1963).

Michael Moran (1967)
Bibliography updated by Philip Reed (2005)

THOUGHT, LAWS OF

See *Laws of Thought*

THOUGHT EXPERIMENTS IN SCIENCE

Thought experiments in science are generally characterized by contrast to actual experiments: The former are conducted by engaging in an imaginative act, the latter by manipulating features of the observed world. So if to perform an (actual) scientific experiment is to conduct an empirical test under controlled conditions with the aim of illustrating, supporting, or refuting some scientific hypothesis or theory, then to perform a scientific thought experiment is to reason about an imaginary scenario with a similar aim. In the case of actual experiments, the theory-relevant evidence generally takes the form of data concerning the behavior of the physical world under specific conditions; in the case of thought experiments, the theory-relevant evidence generally takes the form of intuitions (or predictions) concerning such behavior. In both instances, imagining or performing the experiment ostensibly results in new knowledge about contingent features of the natural world. The primary philosophical puzzle concerning scientific thought experiment is how (if at all) contemplation of a merely imaginary scenario can provide this. (Cf. Kuhn 1964/1977.)

TERMINOLOGICAL ISSUES

The earliest uses of the expressions *Gedankenexperiment* and *mit Gedanken experimentieren* seem to be in the writings of the Danish Kantian Hans Christian Örsted (1811) and the German polymath Georg Christoph Lichtenberg (1793) respectively. However, contemporary use of the term stems from its apparently independent coinage by Ernst Mach, who introduced the expression *Gedankenexperiment* in an 1897 essay of the same name, and discussed a number of examples that have remained central to present-day discussions. Though the historical record is a bit unclear on this point (because later editions of works often insert the word where it was not originally used), it seems to have taken roughly four decades following the publication of Mach's essay for the term *thought experiment* to become widespread in scientific circles. In particular, despite his thorough knowledge of Mach's corpus, Einstein seems not to have used the term to describe his own thinking, at least not in his written works. (Cf. Lichtenberg 1793/1983; Mach 1897; Mach 1905/1976; Schildknecht 1990, 147ff; Witt-Hansen 1976).

Despite the absence of a specific term for the technique, the method was widely employed long before it was labeled. Contemplating imaginary cases in order to develop scientific theory was central to the practice of ancient and medieval natural philosophy, despite the apparent absence of any articulated experimental methodology. And it played a crucial role in the development of early modern natural science. Indeed, some have argued that thought experiment was the predominant mode of scientific investigation prior to the scientific revolution (cf. King 1991, Rescher 1991).

This points to a certain ambiguity in the term's application. Given the characterization offered above, it is a bit challenging to distinguish scientific *thought experiment* as such from scientific *thought* in general, because the latter largely consists in reasoning about (less or more detailed) imaginary scenarios as a way of testing or illustrating (more or less tentative) hypotheses. Indeed, nearly every exercise in a standard physics textbook would, by these criteria, count as a thought experiment. As a matter of sociological fact, however, the expression tends to be reserved for cases where a fairly detailed scenario is contemplated in order to invoke intuitions that help to illustrate or support a specific and novel scientific hypothesis, or to refute a specific and otherwise plausible scientific hypothesis. (A parallel set of definitional and historical issues confronts the analogous term in philosophy, where the term "thought experiment" is generally used to refer to the consideration of fairly detailed, often physically unrealized, scenarios in order to invoke intuitions concerning the proper application of some concept.) Perhaps because of these definitional difficulties, philosophical discussions of scientific thought experiment have focused primarily on a small stable of canonical examples. (For a comprehensive bibliography, see Gendler 2000.)

EXAMPLES

Among the three most widely discussed scientific thought experiments in the philosophical literature are Galileo's refutation of the Aristotelian view that heavy bodies fall faster than light ones, Stevin's determination of the

FIGURE 1

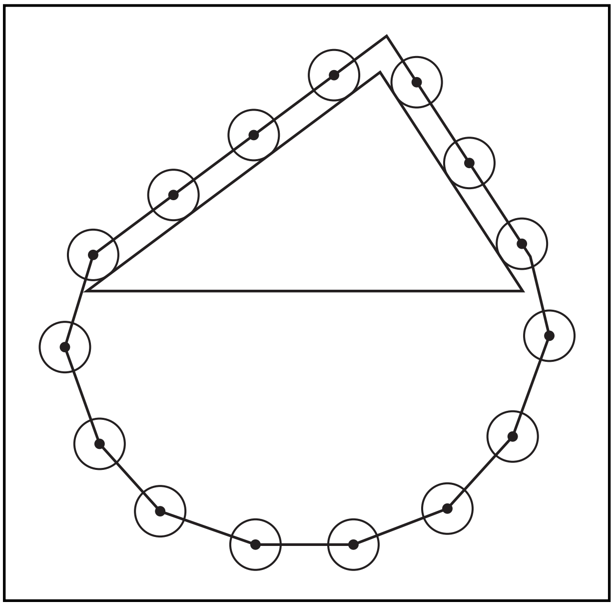

amount of force required to prevent an object from sliding down a frictionless inclined plane, and Einstein's demonstration of the relativity of simultaneity by consideration of the moving train. These exemplify respectively the role of scientific thought experiments in refuting, supporting, and illustrating scientific theories.

In Galileo's *falling body* thought experiment, by which Galileo is said to have refuted the Aristotelian theory that heavier bodies fall faster than lighter ones, Galileo imagines two otherwise similar bodies of differing weights that are strapped together and dropped from a significant height. If one accepts the Aristotelian assumption that natural speed is proportional to weight, and accepts that there is no fact of the matter about whether the strapped body is one entity or two (that is, if one accepts that entification is not physically determined), then it seems that two outcomes are predicted: on the one hand, the lighter body should slow down the heavier whereas the heavier speeds up the lighter, so the combined object should fall with a speed that lies between the natural speeds of its components; on the other hand, because the weight of the two bodies combined is greater than the weight of the heavy body alone, their combination should fall with a natural speed greater than that of the heavy body. Galileo's suggested resolution to the paradox is to assume that the natural speed with which a body falls is independent of its weight, that is, that "both great and small bodies … are moved with like speeds" (Galileo 1638/1989, pp. 107–109; cf. Gendler 1998, 2000).

FIGURE 2

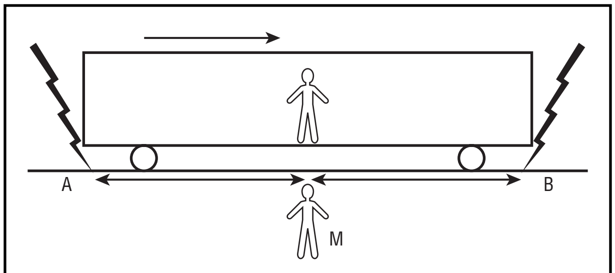

From perspective of observer on embankment

In Stevin's *inclined plane* thought experiment, which served as Mach's original example of the term, Stevinus establishes the amount of force required to prevent an object from sliding down a frictionless inclined plane by imagining a connected string of beads hung across a triangular prism with a horizontal base (as illustrated in figure 1). Consideration of this imaginary setup convinces him that the balls are in a state of equilibrium—that is, that the chain moves neither to the left nor to the right (else, it seems, the system would be in a state of perpetual motion, for because the beads are of equal weight and hung equally along the string, if the current state is one of disequilibrium, so too would be the state into which the system moved as the result of the string sliding.) He next imagines cutting the string at the two lower corners, so that only the beads along the two diagonal planes remain. Given that beads were in equilibrium prior to the cutting, and that the lower part of the loop exerts equal force on both sides of the string, the balls can be expected to remain in equilibrium afterwards. Because the number of beads along each side is proportional to the length of the plane, and because the beads are of equal weight spaced equidistantly, it follows that two bodies on two different, inclined planes are in balance if their weights are proportional to the lengths of the two planes. (Stevin 1955 [1586], pp. 175–179)

In Einstein's *moving train* thought experiment, Einstein illustrates the relativity of simultaneity by imagining a situation in which there are two people, one standing at a point, call it M, along the embankment of a railroad track, the other riding on a train that is moving with respect to the embankment. He then supposes that lightning strikes the embankment at two points, A and B, which are a significant distance from one another, but equidistant from M.

From the perspective of the person standing on the bank, the two flashes occur simultaneously: that is, the ray of light that is emitted from point A reaches M at

exactly the same moment as the ray of light that is emitted from point B (see figure 2).

But from the perspective of the person on the moving train, the two flashes are not simultaneous, because (considered with reference to the embankment) she is rushing toward the beam emitted from B, and away from the beam emitted from A. (Note that from her perspective, it is the person on the embankment who is in motion in the direction of A. Note further that neither frame of reference is privileged in any way.) Because the speed of light is constant, the B-light will reach the passenger earlier than the A-light, so from her perspective, the two flashes are not simultaneous: the B-flash occurs first. Einstein concludes: "We thus arrive at the important result: Events which are simultaneous with reference to the embankment are not simultaneous with respect to the train, and *vice versa* … unless we are told the reference-body to which the statement of time refers, there is no meaning in a statement of the time of an event" (Einstein 1961, p. 26).

PHILOSOPHICAL ISSUES

Philosophical discussions of scientific thought experiment have primarily focused on two related questions. The first, which may be called the "what" question, concerns what sort of knowledge one gains from the contemplation of imaginary cases: do they provide one with new knowledge about contingent features of the natural world, or do they instead provide knowledge of some other sort? The second, which might be called the "how" question, concerns the process by which such knowledge is obtained: what, if anything, is epistemically distinctive about the process of thought-experimental reasoning?

THE "WHAT" QUESTION

A strong case can be made for the view that scientific thought experiments do not, in themselves, provide new knowledge about contingent features of the natural world: to the extent that they provide new knowledge, that knowledge concerns necessary truths. So, for example, the reader who works through Einstein's moving train thought experiment does not thereby gain novel knowledge of the (apparently) contingent truth that simultaneity is relative. What one gains instead is new knowledge of the (apparently) necessary truth that, if the speed of light is constant, then simultaneity is relative, which can then be combined with one's antecedent knowledge that the speed of light is constant in order to gain knowledge of the consequent. Likewise in the case of Stevin: What the thought experiment reveals is not the (apparently) contingent fact that the force required to hold a ball in place along an inclined plane is inversely proportional to the length of the plane, but rather to the (apparently) necessary truth that if certain sorts of states are equilibrium states, then the force required is inversely proportional to length. A person combines independent knowledge of this conditional with prior (empirically obtained) knowledge of statics and dynamics, and thereby gains knowledge of the consequent. So too in the Galileo case: What the reader gains is not new knowledge of the (apparently) contingent truth that the speed at which a body falls is independent of its weight, but rather the (apparently) necessary truth that, if entification is not a physically determined matter, then natural speed is independent of weight. And one can combine this conditional knowledge with one's empirically obtained knowledge of the antecedent to derive the conclusion.

Those who wish to challenge this position must argue that it is by engaging in this particular instance of thought-experimental reasoning that knowledge of the relevant contingent antecedent is gained. This is least plausible in the case of illustrative thought experiments that evoke intuitions about highly theoretical properties (e.g., the Einstein case), and most plausible in the case of supportive or refutory thought experiments that evoke physical intuitions (e.g., Galileo, Stevin). So it might be argued that it is precisely by contemplating the imaginary scenario in question that a person might come to know (the contingent fact) that the balls do not move in the Stevin example, or in the Galileo example (the contingent fact) that it is not a physically determined matter whether the strapped objects form one entity or two: though the intuitions evoked by the cases have their ultimate basis in experience (or the accumulated experience encapsulated by evolution [cf. Shepard undated]), the general information they encapsulate was too unsystematized to count as knowledge prior to engaging in the act of directing imagining. Something like this view appears to have been held by Mach, who writes:

> Unquestionably in the assumption from which Stevinus starts, that the endless chain does not move, there is contained primarily only a *purely instinctive* cognition. He feels at once, and we with him, that we have never observed anything like a motion of the kind referred to, that a thing of such a character does not exist. This conviction has so much logical cogency that we accept the conclusion drawn from it respecting the law of equilibrium on the inclined plane without the thought of an objection, although the law, if pre-

sented as the simple result of an experiment, otherwise put, would appear dubious. (Mach 1976 [1905], p. 34)

THE "HOW" QUESTION

A number of recent discussions of thought-experimental cognition have focused on whether the structured contemplation of imaginary examples produces distinctive sorts of cognitive access to the knowledge they do give (whether or not that knowledge concerns contingent features of the natural world). In a series of widely discussed articles, John Norton (1991, 2002, 2004, and references contained therein) has defended a view that he calls "empiricism" according to which "thought experiments are just ordinary argumentation, disguised in some vivid picturesque or narrative form. As a result," he contends, "they can do nothing more epistemically that can ordinary argumentation" (2002, p. 1). On this view, knowledge obtained through scientific thought experiment is the result of inference from known premises to inductively or deductively implied conclusions: "the actual conduct of a thought experiment consists of the execution of an argument" (Norton 2002, p. 4).

Norton's view has been widely discussed and criticized by those who hold that contemplation of well-articulated specific imaginary cases can give access to inchoate information about patterns of experience to which people lack independent propositional or conceptual access. Some have suggested that thought experiment does this by exploiting the same cognitive mechanisms that mental models do (cf. Nersessian 1993; Miscevic 1992); others have suggested that certain thought experiments work by evoking quasi-sensory intuitions, resulting in new beliefs about contingent features of the natural world that are produced not inferentially, but quasi-observationally (Gendler 2004). Yet others have stressed other aspects of the similarities between thought experiments and actual experiments (for example, their indifference to certain sorts of changes of content but not others), contending that insofar as the latter are not arguments, neither are the former (cf. Arthur 1999, Bishop 1999, Gooding 1992, Sorensen 1992.)

A final contrasting view, advanced in a series of papers and books by James Robert Brown (e.g., Brown 1991, 2002, 2003, 2004, and references contained therein) is that in certain instances (the Galileo case being one) engaging in thought-experimental reasoning provides "*a priori* (though still fallible) knowledge of nature" derived through a process of what Brown terms "platonic insight" (2002, p. 2). "Thought experiments," he writes, "are our telescopes to see into the abstract realm"; by making use of "the mind's eye," they allow us to perceive the laws of nature "*a priori*" (2004, p. 113). The laws in question are necessary rather than contingent, involving "relations between objectively existing abstract entities" (2002, p. 2). Such a view will be appealing only to those who accept Brown's platonist metaphysics along with its corresponding epistemology.

See also Experimentation and Instrumentalism; Scientific Method.

Bibliography

Arthur, Richard. "On Thought Experiments as *A Priori* Science." *International Studies in the Philosophy of Science* 13 (3) (1999): 215–229.

Bishop, Michael. "Why Thought Experiments Are Not Arguments." *Philosophy of Science* 66 (1999): 534–541.

Brown, James Robert. *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. New York: Routledge, 1991.

Brown, James R. "Thought Experiments." In *The Stanford Encyclopedia of Philosophy (Summer 2002 Edition)*, edited by Edward N. Zalta. Available from http://plato.stanford.edu/archives/sum2002/entries/thought-experiment/.

Brown, J. R. "Why Thought Experiments Transcend Experience." In *Contemporary Debates in Philosophy of Science*, edited by Christopher Hitchcock. Oxford: Blackwell, 2004.

Brown, James Robert. "Peeking into Plato's Heaven." *Philosophy of Science* 71 (2004): 1126–1138.

Einstein, Albert. *Relativity: The Special and General Theory*. New York: Crown Publishers, 1961.

Galilei, Galileo. *Two New Sciences, Including Centers of Gravity and Force of Gravity and Force of Percussion* [1638]. Translated by Stillman Drake. Toronto: Wall & Thompson, 1974. Revised edition, 1989.

Gendler, Tamar Szabó. "Galileo and the Indispensability of Scientific Thought Experiment." *British Journal for the Philosophy of Science* 49 (3) (1998): 397–424.

Gendler, Tamar Szabó. *Thought Experiment: On the Powers and Limits of Imaginary Cases*. New York: Garland Press, 2000.

Gendler, Tamar Szabó. "Thought Experiments Rethought—and Reperceived." *Philosophy of Science* 71 (2004): 1152–1163.

Gooding, David C. "The Cognitive Turn, or, Why Do Thought Experiments Work?" In *Cognitive Models of Science*, edited by Ronald Giere, 45–76. Minneapolis: University of Minnesota Press, 1992.

Hitchcock, Christopher, ed. *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell, 2004.

Horowitz, Tamara, and Gerald Massey, eds. *Thought Experiments in Science and Philosophy*. Savage, MD: Rowman and Littlefield, 1991.

King, Peter. "Mediaeval Thought-Experiments: The Metamethodology of Mediaeval Science." In *Thought Experiments in Science and Philosophy*, edited by Tamara

Horowitz and Gerald Massey, 43–64. Savage, MD: Rowman and Littlefield, 1991.

Kuhn, Thomas. "A Function for Thought Experiments" [1964]. Reprinted in *The Essential Tension*. Chicago: University of Chicago Press, 1977.

Lichtenberg, Georg Christoph. *Schriften und Briefe: Sudelbücher, Fragmente, Fabeln, Verse* (Erster Band), edited by Franz H. Mautner. Frankfurt: Insel Verlag, 1983.

Mach, Ernst. "Über Gedankenexperimente." *Poskes Zeitschrift für den physikalischen und chemischen Unterricht* (January 1897): 1–5.

Mach, Ernst. "Über Gedankenexperimente." *Erkenntnis und Irrtum*, 183–199. Leipzig: Johann Ambrosius Barth,1905. Translation of 1926 edition of *Erkenntinis and Irrtum* by Thomas J. McCormack and Paul Foulkes as "On Thought Experiment." *Knowledge and Error*, 134–147 (Dordrecht, Netherlands: Reidel, 1976).

Miscevic, Nenad. "Mental Models and Thought Experiments." *International Studies in the Philosophy of Science* 6 (3) (1992): 215–226.

Nersessian, Nancy. "In the Theoretician's Laboratory: Thought Experiment as Mental Modeling." *Proceedings of the Philosophy of Science Association* 2 (1993): 291–301.

Norton, John. "Thought Experiments in Einstein's Work." In *Thought Experiments in Science and Philosophy*, edited by Tamara Horowitz and Gerald Massey, 129–148. Savage, MD: Rowman and Littlefield, 1991.

Norton, John. "Why Thought Experiments Do Not Transcend Empiricism." In *Contemporary Debates in Philosophy of Science*, edited by Christopher Hitchcock. Oxford: Blackwell, 2004.

Norton, John. "On Thought Experiments: Is There More to the Argument?" *Philosophy of Science* 71 (2002): 1139–1151.

Rescher, Nicholas. "Thought Experiments in Pre-Socratic Philosophy." In *Thought Experiments in Science and Philosophy*, edited by Tamara Horowitz and Gerald Massey, 31–41. Savage, MD: Rowman and Littlefield, 1991.

Schildknecht, Christiane. *Philosophische Masken: Literarische Formen der Philosophie bei Platon, Descartes, Wolff und Lichtenberg*. Stuttgart, Germany: Metzler, 1990.

Shepard, Roger. "Thought Experiments on Falling Bodies." Available from http://www.cs.northwestern.edu/~paritosh/links/shepard.html

Sorensen, Roy. *Thought Experiments*. New York: Oxford University Press, 1992.

Stevin, Simon. *The Principal Works of Simon Stevin*. Vol. 1: *General Introduction and Mechanics*, [1586], edited by E. J. Dijksterhuis. Amsterdam: C. W. Swets and Zeitlinger, 1955.

Witt-Hansen, Johannes. "H. C. Örsted, Immanuel Kant, and the Thought Experiment." *Danish Yearbook of Philosophy* 13 (1976): 48–65.

Tamar Szabó Gendler (2005)

THUCYDIDES

(460–399 BCE)

Thucydides wrote a history of the epic struggle between Athens and Sparta. His work has proved to be—as he hoped—a "possession for all time," though perhaps not in quite the way he intended. Virtually every age, every occasion, every interpreter, has appropriated a different Thucydides and a different masterpiece. Both the author and the work remain enigmatic.

The reliable biographical details are few, and all derive from his own account. Thucydides son of Olorus was an Athenian, born around 460 BCE. In his analysis of the causes, symptoms, and consequences of the plague that devastated Athens a few years after the outbreak of hostilities with Sparta, Thucydides drew on his own experience of the illness. He was for a time prominent in Athenian public life. During the war, he attained the office of general, one of the very few elected positions in the Athenian democracy (most offices were allocated by lot), and was sent to Thrace, perhaps because of his connections and influence there. In 423 BCE, his fellow citizens banished him for failing to reach the Athenian colony of Amphipolis in time to rescue it from the Spartans. Athens' loss was posterity's gain: Thucydides proceeded to travel the Greek world and gather information for his history from a variety of sources including, as he noted, the Spartans and their allies. He lived to see the end of the war he chronicled, though his narrative breaks off seven years earlier, in 411 BCE.

The history is no less difficult to pin down than the historian, in part because it gives eloquent voice to the various protagonists in the conflict. The history therefore provides ample fodder for a variety of interpretations. Thucydides has been dubbed a scientific historian by some, a dramatist by others. His history is said by some to argue for a realist view of human affairs and international relations, by others to demonstrate the fallacy of such a view.

Thucydides' history is more and other than the sum of its parts. The complexity of his account cannot be reduced; but it can be understood, by taking seriously several considerations. First, Thucydides chose to write history, not tragic poetry, philosophical dialogues, or medical treatises. He explicitly commits himself to giving an accurate account, based on firsthand knowledge or scrupulous inquiry. In the case of the speeches, he states that since it was not possible to "carry them word for word in one's memory," he makes the speakers say what in his judgment is "demanded of them by the various occa-

sions, while adhering as closely as possible to the general sense of what they really said." (1.22.1) In so doing, Thucydides does not abandon history for drama or dogma, but rather insists on the need for and the possibility of rigorously truthful historical interpretation.

Second, certain aspects of human nature (including judgment, passion, chance, the need for security, and the desire for power and gain) form the backbone of Thucydides' attempt to explain and interpret—not merely recount—the events of his time. Different speakers appeal to these concepts in different ways under different circumstances, and so does Thucydides himself when he characterizes the sources and trajectory of Athenian imperial power and the polarization of the Greek world. These building blocks of an intelligible history are therefore not to be seen as static truths, but construed instead in terms of the relationship between actions and contexts over time. It is, for example, not true that the will to power is the fundamental and inexorable force in human affairs, but rather that the will to power leads to greatness for some and security for others until the Greek world is fully polarized, at which point it is essential and possible to exert self-control.

Third, Thucydides was not writing in a vacuum. His decision to write an interpreted history is a response to challenges raised by the experience of democracy. Thucydides' history is intended as a political argument and a political education, and effective as such only to the extent that it is an accurate and intelligible history. Throughout the fifth century the Athenians wrestled with the question of how a polity that gives equal access to decision-making power to all citizens, including those without breeding, education, or property, can possibly achieve order, freedom, or the collective good. Protagoras of Abdera, one of the Sophists, or teachers of the art of politics, argued that participation in democratic practices facilitated self-expression while promoting self-restraint.

By the time of the war, continued reflection on the democratic experience had spawned the view that *Nomos* (law or custom), self-imposed as it was by the people, or by a majority, was in fact an artificial constraint, unrelated to the well-being of any particular citizen. Political deliberation was characterized as a manipulative process designed to advance the interests of some at the expense of others. In response to these challenges to the belief that man's good could be secured through democratic political interaction, some thinkers (Socrates among them) appealed to the force of reason, detached from the realm of politics and persuasion, as the fundamental criterion of the good for man; others (Callicles, in Plato's *Gorgias*) appealed to the force of desire and ambition, likewise detached from social convention. Neither view could accommodate the complexities of the human condition: the real constraints on any person or polity's will to power, and the no less authentic claims of personal needs and passions against the single-minded cultivation of the rational soul.

By the time Thucydides came to write his history, and in part because of the process he charts, the significance of these various aspects of the human experience had become all too evident. He portrays the social and ethical corrosion caused by the polarization of the Greek world, both within and among states, and by war, which he calls a "harsh schoolmaster." Thucydides offers history as a way for people to think and act prudently under such conditions. An interpreted history—which engages the reader's emotions as well as their reason—extends the range of man's experience and cultivates their capacity for judgment under trying circumstances, an appreciation of the need for self-control, and an ability to exercise it.

Historical analysis is most effective when it informs political leadership, as occurred in Athens under the guidance of Pericles (495–429 BCE). As Thucydides portrays him, Pericles sought to educate the Athenians about their real condition, its sources and implications, in such a way as to enable them to anticipate and reconsider their responses. Thucydides acknowledges that this kind of historical leadership did not always work—even when Pericles was alive—and gave way to demagoguery and distortions of the truth after he died. Thucydides himself has acquired a reputation for hostility to democracy because he inclines at times toward institutional substitutes for the dynamic cultivation of judgment through democratic interaction. But his characterization of the respective strengths and weaknesses of the Athenians and the Spartans points to Thucydides' belief that the most admirable polity—the one capable of understanding and responding to the world as it really is—is a democratic polity, like Athens, that cultivates initiative, flexibility, passion, freedom, and is guided by prudent leadership—and by history.

See also Nomos and Phusis; Socrates; Sophists.

Bibliography

Cochrane, C. N. *Thucydides and the Science of History*. New York: Russell & Russell, 1965.

Connor, W. Robert. *Thucydides*. Princeton, NJ: Princeton University Press, 1984.

Cornford, F.M. *Thucydides Mythistoricus*. London: E. Arnold, 1907.

Crane, Gregory. *Thucydides and the Ancient Simplicity: The Limits of Political Realism*. Berkeley: University of California Press, 1998.

Farrar, Cynthia. *Origins of Democratic Thinking: The Invention of Politics in Classical Athens*. Cambridge, U.K., and New York; Cambridge University Press, 1988.

Romilly, Jaqueline de. *Thucydides and Athenian Imperialism*. Translated by Philip Thody. New York: Barnes and Noble, 1963.

Strassler, Robert B., ed. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War*. New York: Free Press, 1996.

Williams, Bernard. *Truth and Truthfulness: An Essay in Genealogy*. Princeton, NJ: Princeton University Press, 2002.

Cynthia Farrar (2005)

THÜMMIG, LUDWIG PHILIPP

(1697–1728)

Ludwig Philipp Thümmig, the German Wolffian philosopher, was professor of philosophy at Halle from 1717 until 1723 when he was expelled with Christian Wolff. On Wolff's recommendation he was appointed professor of philosophy at the Collegium Carolinum in Kassel, but he ended his career as an instructor of pages. His early death prevented him from regaining a decent position when Wolff's fortunes improved.

Thümmig was one of Wolff's earliest pupils, and his *Institutions Philosophiae Wolffianae* (2 vols., Frankfurt and Leipzig, 1725–1726) was intended as a short and more readily understandable presentation, closer to the doctrines of traditional philosophy, of the doctrines presented in Wolff's German works. The work was written in Latin to prevent misunderstandings arising out of Wolff's new German terminology. The order of presentation of the main subjects covered, and the sharp separation between the topics treated in the discussions of the main branches of philosophy, were probably suggested by Wolff and were later adopted by him in his own Latin works. Unlike Wolff in his German works, Thümmig discussed cosmology before psychology, and divided psychology into empirical and rational branches. This order became traditional in the Wolffian school and was adopted by Wolff himself in his Latin works.

Thümmig used the traditional language and manner of exposition to make Wolff's doctrines more acceptable. He introduced non-Wolffian elements into his solution to the problem of preestablished harmony. He also differed from Wolff in regarding the study of natural law as a theoretical science (*scientia legum naturalia*) but ethics and politics as practical sciences whose purpose was to reach an agreement between man's real condition and the natural law.

See also Cosmology; Natural Law; Psychology; Wolff, Christian.

Bibliography

Thümmig's works include *Demonstratio Immortalitas Animae ex Intima Eius Natura Deducta* (Halle, 1721), and *Meletemata Varii et Rarioris Argumenti* (Braunschweig and Leipzig, 1727).

For a discussion of Thümmig, see Max Wundt, *Die deutsche Schulphilosophie im Zeitalter der Aufklärung,* 212–214. (Tübingen: Mohr, 1945).

Giorgio Tonelli (1967)

TILLICH, PAUL

(1886–1965)

Paul Tillich, the German American theologian, was born in Starzeddel in eastern Germany, the son of a Lutheran pastor. He received a theological and philosophical education and was ordained in the Evangelical Lutheran Church in 1912. He served as an army chaplain during World War I and then taught theology and philosophy at Berlin, Marburg, Dresden, and Frankfurt. On Adolf Hitler's advent to power in 1933, Tillich immigrated to the United States, serving as professor of systematic theology and philosophy of religion at Union Theological Seminary from 1933 to 1956. From 1956 until his death he held chairs at Harvard and at the University of Chicago.

ANXIETY

Tillich's religious thought has been enormously influential, particularly in English-speaking countries. He was strongly influenced by existentialism, and he held, as did Søren Kierkegaard, that religious questions are appropriately raised only in relation to problems that are inherent in the "human situation" and that theological claims are not mere responses to theoretical puzzles. Thus, Tillich presents Christian doctrines as resolutions of practical problems. His discussion of anxiety in *The Courage to Be* is a good example of his method. He first analyzes thoroughly and with great sensitivity what he considers the three great anxieties of modern man—the anxiety of death, that of meaninglessness, and that of guilt. These three forms of anxiety are three modes of response to var-

ious kinds of threats from nonbeing, threats to which existence as such is subject. As a practical solution to this practical problem, theology presents God. By participating in God, who is the infinite power to resist the threat of nonbeing, man acquires the courage to exist fully, even in the face of such anxiety. Similarly, when a person becomes deeply aware of historical existence as full of ambiguities, he becomes filled with perplexities and despair. The Christian answer is the notion of the Kingdom of God, which is the meaning, fulfillment, and unity of history.

KNOWLEDGE OF REALITY

Tillich's concern was with the religious significance of the "human situation," and he held that religious questions arise out of human problems. In a similar vein, the only basis for an understanding of the ontological structure of reality is the analysis of human existence, of man's encounter with his environment. We can grasp the being of other things only by analogy with man. Tillich, in the first volume of his *Systematic Theology,* sees man as "that being in whom all levels of being are united and approachable." But man is not merely "an outstanding object among other objects." He is the "being who asks the ontological question and in whose self-awareness the ontological answer can be found." Man can proceed in this way "because he experiences directly and immediately the structure of being and its elements"—because "the interdependence of ego-self and world is the basic ontological structure and implies all the others." Man is a self; "therefore selfhood and self-centeredness must be attributed … to all living beings and, in terms of analogy, to all individual *Gestalten* even in the inorganic realm." In accordance with this view, Tillich takes concepts that he supposes to have their primary application to human existence—individualization and participation, dynamics and form, freedom and destiny—and designates them as the elements constituting ontological structure, applying them to being as such.

FAITH

Tillich conceives of faith or, as he calls it, "ultimate concern" as a way of organizing human experience and activity. In his view, faith is an unconditional surrender to something and the willingness to recognize it as an absolute authority; an expectation that one will in some way receive a supreme fulfillment through encounter and commerce with it; a discovery that everything in one's life and one's world is significant only insofar as it is in some way related to it; and experiencing it as holy—that is, reacting to it with an intimate blend of a sense of awe, mystery, and fascination.

Every human being, Tillich believed, has such an ultimate concern, but the objects of the concern vary enormously. Supernatural beings, historical persons whether religious or secular, nations, social classes, political movements, cultural forms like painting and science, material goods, social status—any of these may be the object of an ultimate concern. But despite what Tillich said, it would seem that such orientation around a single object is a rare achievement. Most people, it would seem, have several major interests. Moreover, there is a crucial difference between concern with an object, whether existent or thought to exist, and concern for the realization of some end. The significance of taking an end, like social status, as having authority is not clear. Nevertheless, Tillich's analysis of religiosity is a penetrating one, and it reveals the important resemblances between religiosity and nonreligious modes of personal organization.

GOD

Tillich tried to show that the religious life is more than an organization of human feelings and attitudes and that it involves a reference to a reality outside itself, a reference that can be validated. Although Tillich did not, like Kierkegaard, deny the religious relevance of rational investigation, and although he did think that ontology gives some support to religion, he did not believe in the validity of traditional metaphysical proofs of specifically religious doctrines and in particular of the existence of a personal God. Tillich did not, in fact, accept the notion of a personal deity. For him the doctrine of a supernatural person, like all religious doctrines, is to be conceived as an attempt to symbolize an ultimate reality, "being-itself," which is so ultimate that all that can literally be said about it is that it is ultimate. If the God of theism is a person, the often repeated charge that Tillich is really an atheist thus seems justified; yet Tillich can point out that in the past Christian theology has repeatedly found difficulty in the notion that God is a person in any straightforward or literal sense.

THE ULTIMATE

Tillich defended his view that religious faith is objectively valid by claiming that an ultimate concern must necessarily have what is metaphysically Ultimate as its object. It is not clear, however, that if a concern is ultimate (in the sense of being the dominant interest of a person), the object of the concern is necessarily Ultimate in the relevant sense; that is, that the object of the concern is that on

which all else depends for its being. Tillich has argued elsewhere that one can be ultimately concerned only with what is metaphysically Ultimate. Nothing can properly be of ultimate concern unless it is the ultimate determiner of the reality and meaning of our existence, and only being-itself occupies this position. From this conclusion it is only a short step to say that in ultimate concern one is always really concerned with being-itself, whether one realizes it or not.

RELIGIOUS SYMBOLS

But if being-itself is always the object of ultimate concern, what is the status of the various nonultimate entities on which ultimate concern seems to be focused? According to Tillich, as we have seen, the object of an ultimate concern is generally something relatively concrete, such as a person or a social group, and not, at least not consciously, some ineffable metaphysical Ultimate. Tillich claims that these concrete objects function as symbols of the Ultimate. They manifest the Ultimate to those who experience them as holy, and for those persons they *point to* the Ultimate; through them the individual participates in the Ultimate. Thus, ultimate concern has in a sense a double object. Unfortunately, Tillich never gave an intelligible account of these closely interrelated concepts of symbolizing and pointing to, which are so crucial for his position. Pointing to the Ultimate cannot consist in calling the Ultimate to mind, for admittedly most people have no such concept. The main difficulty is that being-itself is given such a fundamental position in Tillich's metaphysical scheme that one necessarily is related to being-itself at every moment in any way in which anyone could conceivably be related to it. Thus, if it is possible to speak of beings participating in being-itself, then each being necessarily so participates at every moment of its existence. There seems to be no room for any special contact with being-itself that could be generated by religious symbols when they are "pointing to it."

DEFENSE OF CHRISTIANITY

As a Christian theologian, Tillich wanted to demonstrate that among ultimate concerns the Christian concern is the most adequate. He sometimes said that some ultimate concerns are "idolatrous" because they are directed at finite objects rather than at the Ultimate. But by his own principles Tillich could not say this, because every case of ultimate concern involves a concrete object that manifests or points to the Ultimate. If it did not so function, it would not be a case of ultimate concern. The only possible way of showing that one ultimate concern is more adequate than another would be to show that it served better as a symbol of being-itself. But since nothing can be said literally about being-itself except that it is Ultimate, a feature that nothing else can share, it is not clear how this could be done. Tillich's own argument for the superiority of Christianity seems itself to be in symbolic terms. He said that by dying on the cross, Jesus Christ, who is the basic symbol of being-itself in Christianity, underlined the fact that symbols have their significance not in themselves but as manifesting the Ultimate.

See also Atheism; Existentialism; Kierkegaard, Søren Aabye; Ontological Argument for the Existence of God; Philosophy of Religion, History of; Philosophy of Religion, Problems of.

Bibliography

WORKS BY TILLICH

Translations of important works include *The Religious Situation,* translated by H. R. Niebuhr (New York: Holt, 1956), and *The Interpretation of History,* translated by N. A. Razetski and E. L. Talmey (New York: Scribners, 1936).

Tillich's magnum opus is *Systematic Theology.* 3 vols. (Chicago: University of Chicago Press, 1951–1963). Vol. I contains most of the philosophically interesting sections. *The Courage to Be* (New Haven, CT: Yale University Press, 1952) is an existentially toned account of the spiritual situation of contemporary man. *The Dynamics of Faith* (New York: Harper, 1957) is the best popular account of Tillich's position. *Biblical Religion and the Search for Ultimate Reality* (Chicago: University of Chicago Press, 1955) defends the thesis that ontological speculation has relevance to biblical religion. *Theology of Culture,* edited by R. C. Kimball (New York: Oxford University Press, 1959), is a collection of Tillich's writings on the mutual relevance of theology and various aspects of culture.

On the notion of a religious symbol, see "The Religious Symbol," in *Journal of Liberal Religion* 2 (1) (Summer 1940): 13–33, and "Theology and Symbolism," in *Religious Symbolism,* edited by F. E. Johnson (New York, 1955), pp. 107–116.

WORKS ON TILLICH

For discussions of Tillich's works, see C. W. Kegley and R. W. Bretall, eds., *The Theology of Paul Tillich* (New York: Macmillan, 1952); W. P. Alston, "Tillich on Idolatry," in *Journal of Religion* 38 (1958): 263–267; Paul Edwards, "Professor Tillich's Confusions," in *Mind* 74 (1965): 192–214; and W. P. Alston, "Tillich's Conception of a Religious Symbol," in *Religious Experience and Truth,* edited by Sidney Hook (New York: New York University Press, 1961), pp. 12–26.

OTHER RECOMMENDED TITLES

Adams, James et al., eds. *The Thought of Paul Tillich.* San Francisco: Harper and Row, 1985.

Crossman, Richard. *Paul Tillich: A Comprehensive Bibliography and Keyword Index of Primary and Secondary Writings in English.* Metuchen, NJ: Scarecrow Press, 1983.

Pauck, Wilhelm, and Marion Pauck. *Paul Tillich: His Life and Thought.* New York: Harper and Row, 1976.

Rowe, William. *Religious Symbols and God: A Philosophical Study of Tillich's Theology.* Chicago: University of Chicago Press, 1968.

Scharlemann, Robert. *Reflection and Doubt in the Thought of Paul Tillich.* New Haven, CT: Yale University Press, 1969.

Thatcher, Adrian. *The Ontology of Paul Tillich.* Oxford: Oxford University Press, 1978.

William P. Alston (1967)
Bibliography updated by Christian B. Miller (2005)

TIME

Time has frequently struck philosophers as mysterious. Some have even felt that it was incapable of rational discursive treatment and that it was able to be grasped only by intuition. This defeatist attitude probably arises because time always seems to be mysteriously slipping away from us; no sooner do we grasp a bit of it in our consciousness than it has slipped away into the past. This entry will argue, however, that this notion of time as something that continually passes is based on a confusion.

ST. AUGUSTINE'S PUZZLES

The apparent mysteriousness of time can make puzzles about time seem more baffling than they are, even though similar ones arise in the case of nontemporal concepts. St. Augustine, in his *Confessions*, asks, "What is time?" When no one asks him, he knows; when someone asks him, however, he does not know. He knows how to use the word "time" and cognate temporal words, such as "before," "after," "past," and "future," but he can give no clear account of this use. Trouble arises particularly from the form in which he puts his question: "What is time?" This looks like a request for a definition, and yet no definition is forthcoming. However, most interesting concepts cannot be elucidated by explicit definitions. Thus, to explain the meaning of the word "length," we cannot give an explicit definition, but we can do things that explain how to tell that one thing is longer than another and how to measure length. In the same way, it is possible to give an account of the use of the word "time" even though it is not possible to do so by giving an explicit definition. In short, this puzzle of St. Augustine's is not of a sort that arises peculiarly in the case of time. Beyond pointing this out, therefore, it is not appropriate here to go further into the matter.

Augustine was also puzzled by how we could measure time. He seems to have been impressed by the lack of analogy between spatial and temporal measurement. For example, one can put a ruler alongside a tabletop, and the ruler and the tabletop are all there at once. However, if one were to measure a temporal process, it would be done by comparing it with some other process, such as the movement of the hand of a watch. At any moment of the comparison, part of the process to be measured has passed away, and part of it is yet to be. It is not possible to get the thing to be measured in front of a person all at once, as one could with the tabletop. Moreover, if two temporal processes are compared—say, a twenty-mile walk last week with a twenty-mile walk today—they are compared with two different movements of a watch hand, whereas two different tabletops are compared with the same ruler. Augustine is led to see a puzzle here because he demands, in effect, that non-analogous things should be talked about as though they were analogous.

In any case, the two things are not, in fact, as non-analogous as they appear to be at first sight. If we pass to a tenseless idiom in which material things are thought of as four-dimensional space-time solids, the difference becomes less apparent. For in the case of the tables we compare two different spatial cross sections of the four-dimensional object that is the ruler with spatial cross sections of the two tables. Augustine seems to have been influenced by the thought that the present is real, although the past and future are not (the past has ceased to exist, and the future has not yet come to be); consequently, the measurement of time is puzzling in a way in which the measurement of space need not be (where the whole spatial object can be present now). This thought—that the present is real in a way in which past and future are not real—is part of the confusion of the flow or passage of time. This is not to say that presentism has not recently been intelligently defended, however implausibly, as by John Bigelow (1996). Apodeictic proof has rarely been possible in metaphysics, and we fall back eventually on trading plausibilities. One of the central objections to presentism is the difficulty it has in analyzing cross-temporal statements such as "Smith will have come before you have finished breakfast." Perhaps the most important objection relates to the explanatory value of four-dimensional space-time in relativity theory to be discussed below.

THE MYTH OF PASSAGE

We commonly think of time as a stream that flows or as a sea over which we advance. The two metaphors come to much the same thing, forming part of a whole way of thinking about time that D. C. Williams has called "the myth of passage"(Williams 1951). If time flows past us or if we advance through time, this would be a motion with respect to a hypertime. For motion in space is motion with respect to time, and motion of time or in time could hardly be a motion in time with respect to time. Ascription of a metric to time is not necessary for the argument, but supposing that time can be measured in seconds, the difficulty comes out clearly. If motion in space is feet per second, at what speed is the flow of time? Seconds per what? Moreover, if passage is of the essence of time, it is presumably the essence of hypertime, too, which would lead one to postulate a hyper-hypertime and so on ad infinitum.

The idea of time as passing is connected with the idea of events changing from future to past. We think of events as approaching us from the future, whereupon they are momentarily caught in the spotlight of the present and then recede into the past. Yet in normal contexts it does not make sense to talk of events changing or staying the same. Roughly speaking, events are happenings to continuants—that is, to things that change or stay the same. Thus, we can speak of a table, a star, or a political constitution as changing or staying the same. But can we intelligibly talk of a change itself as changing or not changing?

It is true that in the differential calculus we talk of rates of change changing, but a rate of change is not the same thing as a change. Again, we can talk of continuants as coming into existence or ceasing to exist, but we cannot similarly talk of a "coming-into-existence" itself as coming into existence or ceasing to exist. It is nevertheless true that there is a special class of predicates, such as "being past," "being present," "being future," together with some epistemological predicates such as "being probable" or "being foreseen," with respect to which we can talk of events as changing. Significantly enough, these predicates do not apply to continuants. We do not, for example, naturally talk of a table or a star as "becoming past" but of its "ceasing to exist." There is something odd about the putative properties of pastness, presentness, futurity, and the like, whereby events are supposed to change. One might conjecture that the illusion of the passage of time arises from confusing the flow of information through our short-term memories with a flow of time itself.

TOKEN-REFLEXIVE EXPRESSIONS. Leaving aside the epistemological predicates, we may suspect that the oddness arises because the words "past," "present," and "future," together with "now" and with tenses, are token-reflexive, or indexical, expressions. That is, these words refer to their own utterance. If italics are allowed to indicate tenselessness in a verb, then if one says, "Caesar *crosses* the Rubicon," the speaker does not indicate whether the crossing is something before, simultaneous with, or after the assertion. Tenseless verbs occur in mathematics where temporal position relative to a person's utterance is not even in question. Thus, we can say, "2 + 2 *is* equal to 4" not because we wish to be noncommittal about the temporal position of 2 + 2 as being 4 but because it has no temporal position at all.

The token-reflexiveness (or more generally the indexicality) of the word "past" can be seen, for example, if a person who said that a certain event *E* is past could equally well have said, "*E is* earlier than this utterance." Similarly, instead of saying, "*E* is present," he could say, "*E is* simultaneous with this utterance," and instead of "*E* is future," he could say, "*E is* later than this utterance." The phrase "*E* was future" is more complicated. It means that if someone had said, "*E* is future" or "*E is* later than this utterance," at some appropriate time earlier than the present utterance (the utterance which we now refer to as "this utterance"), he would have spoken truly. Thus, if we say that in 1939 the battle of Britain was in the future, we are putting ourselves into the shoes of ourselves as we were in 1939, when, given a certain amount of prescience, we might have said truly, "The battle of Britain *is* later than this utterance." Apart from this imaginative projection, we are saying no more than that the battle of Britain *is* later than 1939. Another way of dealing with this problem, one that is preferred by Michael Tooley (1997) would be to interpret the token reflexive expressions as referring not to utterances but to times of utterance.

It follows that there is a confusion in talking of events as changing in respect of pastness, presentness, and futurity. These are not genuine properties, which can be seen if the token-reflexiveness is made explicit. "*E* was future, is present, and will become past" goes over into "*E is* later than some utterance earlier than this utterance, *is* simultaneous with this utterance, and *is* earlier than some utterance later than this utterance." Here the reference is to three different utterances. However, if we allow simultaneity, being later, and being earlier as relations to times as well as events we could render the tensed sentence above by saying, "*E is* later than some time earlier than this utterance, *is* simultaneous with this utterance, and *is*

earlier than some time later than this utterance." Also, the troubling sentence "Once there were no utterances" could go over to "There *are* times earlier than this utterance when there were no utterances." A failure to recognize the direct or indirect indexicality of words such as "past," "present," and "future" can lead us to think wrongly of the change from future to past as a genuine change, such as the change in position of a boat that floats down a river.

Nevertheless, there is probably a deeper source of the illusion of time flow. This is that our stock of memories is constantly increasing, and memories are of earlier, not of later, events. It is difficult to state this matter properly because we forget things as well as acquire new memories. With a very old man there may well be a net diminishing of his stock of memories, and yet he does not feel as if time were running the other way. This suggestion is therefore tentative and incompletely worked out. Possibly we confuse a flow of information through our short-term memories with a flow of time itself (Smart 1987). The subordinate question of why our memories are of the past, not of the future, is an extremely interesting question in its own right and will be answered in a later section.

TENSES. Not only words such as "past" and "future" but also tenses can be replaced by the use of tenseless verbs together with the phrase "this utterance." Thus, instead of saying, "Caesar crossed the Rubicon," we could have said, "Caesar *crosses* the Rubicon earlier than this utterance." For the present and future tenses we use "simultaneous with this utterance" and "later than this utterance." Of course, this is not a strict translation. If one person says, "Caesar *crosses* the Rubicon earlier than this utterance," that person refers to his utterance, whereas if another person says, "Caesar crossed the Rubicon," she is implicitly referring to her utterance. Nevertheless, a tensed language is translatable into a tenseless language in the sense that the purposes subserved by the one, in which utterances covertly refer to themselves, can be subserved by the other in which utterances explicitly refer to themselves.

A second qualification must be made. In the case of spoken language the token or "utterance" can be taken to be the actual sounds. In a written language the "token," the configuration of ink marks, is something that persists through time. By "this utterance" we must therefore, in the case of written language, understand the coming-into-existence of the token or perhaps the act of writing it. It has sometimes been objected that this account will not stand because "this utterance" means "the utterance which is *now*," which reintroduces the notion of tense. There does not seem to be any reason, however, why we should accept this charge of circularity. We have as good a right to say that "now" means "simultaneous with this utterance" as our opponent has to say that "this utterance" means "the utterance which is now." The notion of an utterance directly referring to itself does not seem to be a difficult one.

Tenses and their cognates may be seen to be indexical expressions. The truth conditions of sentences containing them cannot be given by translation into a nonindexical language. Nevertheless they can be given in a nonindexical *metalanguage*. The idea derives from Donald Davidson and is advantageous because there is a recursively specifiable infinity of sentences in a language but not of utterances or inscriptions. Equally with the token reflexive account it removes the mystery that one might feel about tenses and cognate expressions.

Tensers, such as Quentin Smith (1993), argue that the words "past," "present" and "future" refer to intrinsic properties of events, though Smith defines "past" and "future" in terms of "present." This makes him in a sense a presentist, though only a mild one as he does not deny the reality of the past and future. Davidson's suggestion for the semantics of tenses is to say that (say) "I will come" is true as (potentially) spoken by person *P* at time *t* if and only if *P comes* later than *t*. As Heather Dyke, in her doughty defense of the token-reflexive approach (Dyke 2002, 2003), has remarked, without the "potentially" (of which critics of modal logic may be suspicious) the Davidsonian schema comes out trivially true in cases where (say) "I will come" is not uttered by *P* at *t*. Perhaps one might reply that trivial truth is still truth and so harmless, or one might treat the Davidsonian schema as an idealization. Dyke has urged that one should abandon aspirations of the old token reflexive theory for a translation of tensed sentences into tenseless ones but argue that a tensed sentence states the same fact about the world as can be stated by a tenseless one. Thus she wants a semantics based on tokens of sentences, not sentences, and so abandons recursiveness. A similar appeal to the notion of "fact" is made by D. H. Mellor in his influential *Real Time II* (1998), where he says that ontology can be separated from considerations of semantics. Of course this metaphysical notion of "fact" has been thought problematic, as by Davidson himself. Nevertheless, the difference between the token reflexive account and the metalinguistic one is not of great ontological significance. Dyke contests arguments by Quentin Smith (1993), who has been an immensely prolific defender of the tensed notion of time.

DURATION. The philosophical notion of duration seems to be heavily infected with the myth of passage. Thus John Locke in his *Essay concerning Human Understanding* (1690) says that "duration is fleeting extension" (bk II, ch. 14, paragraph 1). In the early nineteenth century, Henri Bergson (1910, 1911, 1913) made the notion of duration (*durée*) central in his philosophy. According to him, physical time is something spatialized and intellectualized, whereas the real thing, with which we are acquainted in intuition (inner experience), is duration. Unlike physical time, which is always measured by comparing discrete spatial positions—for example, of clock hands—duration is the experienced change itself, the directly intuited nonspatial stream of consciousness in which past, present, and future flow into one another. Bergson's meaning is unclear, partly because he thinks that duration is something to be intuitively—not intellectually—grasped. Duration is closely connected in his thought with memory, for in memory, Bergson says, the past survives in the present. Here he would seem to be open to the objection, urged against him by Bertrand Russell in his *History of Western Philosophy* (1945), that he confuses the memory of the past event with the past event itself, or the thought with that which is thought about.

Even though the Bergsonian notion of duration may be rejected because of its subjectivism and because of its close connection with the notion of time flow or passage, there is nevertheless a clear use of the word "duration" in science and ordinary life. Thus, in talking about the duration of a war, we talk simply about the temporal distance between its beginning and its end.

MCTAGGART ON TIME'S UNREALITY. The considerations thus far adduced may well be illustrated by considering how they bear on John McTaggart Ellis McTaggart's well-known argument for the unreality of time, which was put forward in an article in *Mind* (1908) and in his posthumous *Nature of Existence* (1927). For McTaggart, events are capable of being ordered in two ways. First, they can be ordered in respect to past, present, and future. He calls this ordering of events "the *A* series." Second, events can be ordered in respect to the relations "earlier than" and "later than." He calls this "the *B* series." McTaggart then argues that the *B* series does not by itself give all that is essential to time and that the *A* series is contradictory. Neither leg of his argument can stand criticism. His reason for saying that the *B* series misses the essence of time is that time involves change and yet it always is, was, and will be the case that the Battle of Hastings, say, is earlier than the Battle of Waterloo. It has already been shown, however, that it is not just false but also absurd to talk of events' changing. The Battle of Hastings is not *sempiternally* earlier than the Battle of Waterloo; it simply *is* (tenselessly) earlier than it. The notion of change is perfectly capable of being expressed in the language of the *B* series by saying that events in the *B* series *differ* from one another in various ways. Similarly, the proposition that a thing changes can be expressed in the language of the *B* series by the statement that one spatial cross section of it *is* different from an earlier one, and the proposition that it does not change can be expressed by saying that earlier and later cross sections *are* similar to one another. To express the notion of change, we are therefore *not* forced to say that events change. Nor, therefore, are we forced into referring to the *A* series, into saying that events change (in the only way in which we can plausibly say this) in respect to pastness, presentness, and futurity.

Nevertheless, if we do retreat to the language of the *A* series, we can perfectly well do so without contradiction. Just as McTaggart erred by using tensed verbs when talking of the *B* series, he in effect made the correlative error of forgetting tenses (or equivalent devices) when talking of the *A* series. For the contradiction that he claimed to find in the *A* series is that because any event is in turn future, present, and past, we must ascribe these three incompatible characteristics to it; but an event cannot be future, present, or past *simpliciter* but only with reference to a particular time—for example, one at which it was future, is present, and will be past. If we restore the tenses, the trouble with the *A* series disappears. Unsuccessful though McTaggart's argument is, it provides an excellent case study with which to elucidate the relations between tensed and tenseless language.

SPACE-TIME

The theory of relativity illustrates the advantages of replacing the separate notions of space and time by a unified notion of space-time. In particular, Minkowski showed that the Lorentz transformations of special relativity correspond to a rotation of axes in space-time. He showed how natural the kinematics of special relativity can seem, as opposed to Newtonian kinematics, in which, in effect, we should rotate the time axis without correspondingly rotating the space axes. Since the theory of relativity it has become a commonplace to regard the world as a four-dimensional space-time manifold. Nevertheless, even in the days of Newtonian dynamics, there was nothing to prevent taking this view of the world, even though it would not have been as neat as it is in relativity theory. If we pass to the four-dimensional way of looking at things, it is important not to be confused about certain

conceptual matters. Confusion will arise if the tenseless way of talking, appropriate to the four-dimensional picture, is mixed with our ordinary way of talking of things as enduring substances, "the permanent in change."

In ordinary language the word "space" itself is used as the name of a continuant. We can say, for example, that a part of space has become, or has continued to be, occupied. Space-time, however, is a "space" in a tenseless sense of this word, and because time is already in the representation, it is wrong to talk of space-time as itself changing. Thus, in some expositions of relativity it is said that a certain "world line" is a track along which a material body moves or a light signal is propagated. The body or light signal, however, cannot correctly be said to move through space-time. What should be said is that the body or the light signal *lies* (tenselessly) along the world line. To talk of anything's moving through space-time is to bring time into the story twice over and in an illegitimate manner. When we are talking about motion in terms of the space-time picture, we must do so in terms of the relative orientations of world lines. Thus, to say that two particles move with a uniform nonzero relative velocity is expressed by saying that they *lie* (tenselessly) along straight world lines that are at an angle to one another. Similarly, the recent conception of the positron as an electron moving backward in time is misleading because nothing can move, forward or backward, in time. What is meant is that the world lines of a positron and electron, which are produced together or which annihilate one another, can be regarded as a single bent world line, and this may indeed be a fruitful way of looking at the matter.

In popular expositions of relativity we also read of such things as "consciousness crawling up the world line of one's body." This is once more the confusion of the myth of passage and, hence, of the illegitimate notion of movement through space-time. It is instructive to consider how H. G. Wells's time machine could be represented in the space-time picture. A moment's thought should suffice to indicate that it cannot be represented at all. For if a line is drawn extending into the past, this will simply be the representation of a particle that has existed for a long time. It is not surprising that we cannot represent a time machine because the notion of such a machine is an incoherent one. How fast would such a machine flash over a given ten-second stretch? In ten seconds or minus ten seconds? Or what? No sensible answer can be given, for the question is itself absurd. The notion also involves the contradiction, pointed out by D. C. Williams in his article "The Myth of Passage" (1951) that if a person gets into a time machine at noon today, then at 3 a.m., say, that person shall be *both* at 3 p.m. today *and* at, say, a million years ago. There is nevertheless a more consistent notion of time travel though misleadingly so called. A person as a space-time entity might lie along a bent-back world line. It might curve back and then would go back to your great grandmother's time and then a bit forward while you saw your great grandmother. Paradox lurks because if the great grandmother had been shot you would not have existed. David Lewis has proposed a banana skin solution. Since you could not have shot your great grandmother some accident, such as your slipping on a banana skin or your pistol jamming, must have prevented you from harming her. One would wish, however, for a solution of the paradox by reference to the laws of nature.

Though D. H. Mellor ably defends the four-dimensional ontology in his *Real Time II*, he nevertheless says something that may puzzle four-dimensionalists—for example, that a person from birth to death, or a stone over a long period of time, is said to have a certain property at time *t*, but not that a mere time slice or temporal stage of the person or stone has the property. The puzzle is perhaps resolved if we note that Mellor thinks of the thing S as reidentifiable or a sortal as discussed by Peter Strawson. This is understandable because a child could hardly—and an adult could not easily—reidentify the mereological fusion of a bird, a bishop, and Mount Everest. Even so, the four-dimensionalist need not discern a difference between "S is A at *t*" and "S at *t* is A." The time slice may be referred to by reference to the salient four-dimensional object of which it is a slice. Mellor rightly stresses the importance for agency and practical matters of notions of reidentifiable sortals and for the determination of the strengths of beliefs and desires by a method originally due to F. P. Ramsey.

ABSOLUTE AND RELATIONAL THEORIES

Isaac Newton held to an absolute theory of space and time, whereas his contemporary Gottfried Wilhelm Leibniz argued that space and time are merely sets of relations between things that are in space and time. Newton misleadingly and unnecessarily expressed his absolute theory of time in terms of the myth of passage, as when he confusingly said, "Absolute, true and mathematical time, of itself and from its own nature, flows equably without relation to anything external" (*Principia,* in the Scholium to the Definitions of *Mathematical Principles of Natural Philosophy*). The special theory of relativity has made it impossible to consider time as something absolute;

rather, it stands neutrally between absolute and relational theories of space-time. The question as between absolute and relational theories of space-time becomes especially interesting when we pass to the general theory of relativity. According to this theory, the structure of space-time is dependent on the distribution of the matter in the universe. In most forms of the theory there is nevertheless a residual space-time structure that cannot be thus accounted for. A curvature is usually attributed to space-time even in the complete absence of matter, and the inertia of a body, according to this theory, depends in part on this cosmological contribution to the local metrical field and hence not solely on the total mass of the universe, as a purely relational theory would require.

Research on this question is still going on, and until it has been decided, Mach's principle (as Einstein called it), according to which the spatiotemporal structure of the universe depends entirely on the distribution of its matter, will remain controversial. But even if Mach's principle were upheld, it might still be possible to interpret matter, in a metaphysical way, as regions of special curvature of space-time. Graham Nerlich (1994) has given a striking and simple argument against those who, like Leibniz, defend relational theories by asking how one could tell whether everything had not doubled in size. He pointed out that this depends on the assumption that space is Euclidean. Relational theorists usually make the relevant relation that of cause and effect. If this is defined by the use of counterfactual propositions one may object that the murkiness or contextual nature of these contrasts with the absolute theory's reliance on the limpid clarity of geometry. Here I use "absolute" to contrast with 'relational' not as contrasted with "relativistic." An objection to a causal theory of time is that there could be uncaused events and that there are uncountably more space-time points than there are events. Michael Tooley separately assumes an ontology and topology of instants of time, but uses a causal theory to define temporal direction.

TIME AND THE CONTINUUM

An absolute theory of space-time, as envisaged above, need not imply that there is anything absolute about distance (space-time interval). Because of the continuity of space-time, any space-time interval contains as many space-time points as any other (that is, a high infinity of them); space and time do not possess an intrinsic metric, and there must always be an element of convention in definitions of congruence in geometry and chronology, as Adolf Grünbaum has pointed out (Grünbaum1973). This means that the same cosmological facts can be expressed by means of a variety of space-time geometries, provided that they have the same topological structure. (Topology is that part of geometry which treats only of those properties of a figure which remain the same however that figure is transformed into a new one, with the sole restriction that a point transforms into one and only one point and neighboring points transform into neighboring ones. Thus, the surface of a sphere and that of a cube have the same topology, but that of a sphere and that of an infinite plane do not.)

ZENO AND CANTOR. The continuity of space and time can be properly understood only in terms of the modern mathematical theory of infinity and dimensionality. Given the concepts available to him, Zeno rightly rejected the view that an extended line or time interval could be composed of unextended points or instants. (See Aristotle, *Physics* 231a20–231bl8 and *De Generatione et Corruptione*, 316al5–317al7.)

In modern terms it may be said that not even a denumerable infinity of points can make up a nonzero interval. Cantor has shown, however, that there are higher types of infinity than that which belongs to denumerable sets, such as the set of all natural numbers. Cantor showed that the set of real numbers on a line, or segment of a line, is of a higher type of infinity than is the set of natural numbers. Perhaps the right cardinality of "dimensionless points" *can* add up to a nonzero length. This answer is on the right track. Nevertheless, the cardinality of a set of points does not by itself determine dimensionality.

For example, Cantor showed that there is a one-to-one mapping between the points of a plane and the points of a line. However, a mathematical theory of dimension has been developed that accords with our intuitions in assigning 0, 1, 2, 3, and so on, dimensions respectively to points, lines, planes, volumes, and so on, and which also assigns dimensions to other sorts of sets of points. For example, the set of all rational points on a line has dimension 0. So does the set of all irrational points. In these cases an infinity of "unextended points" does indeed form a set of dimension 0. Because these two sets of points together make up the set of points on a line, it follows that two sets of dimension 0 can be united to form a set of dimension 1. Strictly speaking, it is even inaccurate to talk of "unextended points." It is sets of points that have dimension. A line is a set of points, and the points are not parts of the line but members of it. The modern theory of dimension shows that there is no inconsistency in supposing that an appropriate nondenu-

merable infinity of points makes up a set of greater dimensionality than any finite or denumerable set of points could.

The theory of the continuum implies that if we take away the lower end of a closed interval, what is left is an open interval, an interval without a first point. In fact, Zeno's premises in his paradox of the dichotomy do not lead to paradox at all but are a consistent consequence of the theory of the continuum. Motion is impossible, according to the paradox of the dichotomy, because before one can go from *A* to *B*, one must first get to the halfway mark *C*, but before one can get to *C*, one must get to the halfway mark *D* between *A* and *C*, and so on indefinitely. It is concluded that the motion can never even get started. A similar argument, applied to time intervals, might seem to show that a thing cannot even endure through time. The fallacy in both cases comes from thinking of the continuum as a set of points or instants arranged in succession. For if a continuous interval had to consist of a first, second, third, and so on point or instant, then the dichotomy would provide a fatal objection. However, points or instants do not occur in succession, because to any point or instant there is no *next* point or instant. Such considerations enable us to deal with Zeno's paradox of Achilles and the tortoise, in which similar difficulties are supposed to arise at the *latter* end of an open interval.

KANT'S ANTINOMIES. A related paradox is Kant's first antinomy, in his *Critique of Pure Reason* (1929 [1781]). As was shown by Edward Caird (1889) in his commentary on Kant's *Critique*, the antinomies (or paradoxes which Kant had constructed about space, time, and causality) were as important as Hume's skeptical philosophy in arousing Kant from his "dogmatic slumbers." Kant's first antinomy relates to both space and time; the concentration here is on *Critique* as it relates to time. There are two antithetical arguments. The first states that the world had a beginning in time, whereas the second, with equal plausibility, seems to show that the world had no beginning in time. The first argument begins with the premise that if the world had no beginning in time, then up to a given moment an infinite series of successive events must have passed. But, says Kant, the infinity of a series consists in the fact that it can never be completed. Hence, it is impossible for an infinite series of events to have passed away.

It can be seen that Kant's argument here rests partly on the myth of passage. Kant thinks of the world as having come to its present state through a series of past events, so that an infinite succession would therefore have had to be completed. Otherwise, he would have been just as puzzled about the possibility of an infinite future as about an infinite past, and this does not seem to have been the case. Just as the sequence 0, 1, 2 … can never be completed in the sense that it has no last member, the sequence ——, –2, –1, 0 cannot be completed in the sense that it has no first member. This is not to say, of course, that an infinite set need have either a first or last member. Thus, the set of temporal instants up to, but not including, a given instant, has neither a first nor last member. However, Kant is clearly thinking not of the set of instants but of a sequence of events, each taking up a finite time. The set of instants does not form a sequence because there are no instants that are next to one another. Kant's definition of infinity, besides being objectionably psychologistic, is clearly inapplicable to infinite sets of entities which do not form a sequence, such as the points on a line or a segment of a line. Concerning an infinite set of events which form a sequence, however, Kant is not justified in supposing that its having a last member is any more objectionable than its having a first member. There is a perfect symmetry between the two cases once we rid ourselves of the notion of passage—that is, of the one-way flow of time.

In Kant's antithetical argument, he argues that the world cannot have had a beginning in time, so that, contrary to the thesis of the antinomy, there must have been an infinity of past events. His reason is that if the world had begun at a certain time, all previous time would have been a blank and there would be no reason that the world should have begun at the time it did rather than at some other time. Previously, Leibniz had used the same argument to support a relational theory of time. If time is constituted solely by the relations between events, then it becomes meaningless to ask questions about the temporal position of the universe as a whole or about when it began. In an absolute theory of time (or of space-time) Kant's problem remains, but further discussion of it cannot be pursued here because it would involve a metaphysical discussion of causality and the principle of sufficient reason.

TEMPORAL ASYMMETRY

We have just seen that Kant was puzzled about the infinity of the past in a way in which he was not puzzled about the infinity of the future. Further, it has been suggested that the myth of passage had something to do with this inconsistency. If we reject the notion of passage, we find ourselves with a new, though soluble, problem. This is the apparent temporal asymmetry of the universe, which

contrasts sharply with its large-scale spatial symmetry. For example, if we look out at the galaxies, they appear to be distributed evenly in all directions, and yet a time direction seems to be specified by the fact that they are all receding from one another, not approaching one another. On a more mundane level, the temporal asymmetry of the universe is forcibly striking in many ways. For example, there is nothing in our experience analogous to memory but with respect to the future. Nor is there anything like a tape recording or a footprint of the future—that is, there are no *traces* of the future. A memory is indeed a special case of a trace. This asymmetry about traces explains how we can be so confident about the past history of the human race and about the past evolution of living creatures, whereas it would be a bold person who would try to guess the political history of even the next hundred years or the organic evolution of the next few millions. The question "Why are there traces only of the past, not of the future?" is thus a fundamental one.

We must first rule out a purely verbalistic answer to this question. Someone might say that traces are always of the past, never of the future, because it is part of the meaning of the word "trace" that traces are of earlier, not of later, events. This would be to suppose that the earlier question is as stupid as the question "Why are bachelors always male, never female?" This account of the matter is not good enough. Admittedly, in the English language as it is, the expression "female bachelor" is a self-contradictory one. Nevertheless, it is easy to imagine a variant of English in which "bachelor" simply meant "not yet married person" and according to which spinsters could therefore be called "bachelors." For example, if one were to call a spinster a "female analogue" of a bachelor, then it is possible to silence the verbalistic objection to the question about why traces are always of the past, never of the future, by recasting it in the form "Why are there no future analogues of traces?"

TEMPORAL ASYMMETRY AND PHYSICAL LAWS. The temporal directionality of the universe or, at the very least, of the present cosmic era of the universe would therefore appear to be a deep-lying cosmological fact, which is not to be glossed over by verbalistic explanations. How is it to be explained? We must first dismiss the suggestion that the asymmetry lies in the laws of physics. The laws of classical dynamics and electromagnetism, as well as of quantum mechanics, are all expressed by time-symmetrical differential equations. In other words, if $f(t)$ is a solution to these equations, so is $f(-t)$. (Actually to take care of recondite matters, twenty-first century physicists believe not in T symmetry but in CPT symmetry, reversal of time, reversal of charge, and reversal of parity. P symmetry can be thought of as reversal in a space mirror just as C symmetry is a matter of thinking of an antiparticle as a backwards-in-time particle. So CPT symmetry can be thought of as a deeper form of space-time symmetry.)

It follows that if a cinematographic film were taken of any process describable by means of these laws and then run backward, it would still portray a physically possible process. It is true that phenomenological thermodynamics would provide a contrary case, because its second law does contain time explicitly. Thus, if someone put a kettle full of ice on a hot brick, that person finds that the system turns into one in which a kettle full of water sits on a cool brick. A film of this process cannot be reversed to show a process which is possible in phenomenological thermodynamics; we cannot have a system of a kettle filled with water on a cool brick turning into one in which the water has frozen and the brick has become hot. In spite of all this it must still be asserted that the laws of nature are time symmetrical. This is because phenomenological thermodynamics provides only an approximation of the truth (it is refuted by the phenomenon of Brownian motion, for example) and, more importantly, because the detailed explanation of the facts of which phenomenological thermodynamics treats at the surface level is to be found in statistical thermodynamics. Statistical thermodynamics bases itself on the laws of mechanics, which are time symmetrical.

According to statistical thermodynamics, the situation in which the water in the kettle freezes while the brick gets hotter is indeed a physically possible one, though it is an almost infinitely unlikely one. Why it is unlikely has to do not with the laws of nature themselves but with their boundary conditions. There is indeed a puzzle here, because if all the velocities of a closed system are reversed, what results is a configuration that, according to statistical mechanics, is as likely as the original one. Therefore, the process seen on the reversed cinematographic film should be as likely as the original one. The answer to this objection (the reversibility objection) lies in the fact that corresponding to a given macroscopic description (cold kettle on hot brick, say), there is a whole ensemble of possible microstates. It follows that though any microstate is as probable as any other, this is not so with macrostates, and given the information that a body is in a macrostate *A*, it is highly probable that it will turn into a macrostate *B* rather than vice versa if *B* corresponds to an ensemble of microstates which is vastly

more numerous than the ensemble of microstates corresponding to *A*.

An analogy with a pack of cards will help to make this clear. Consider a well-shuffled pack of cards. Any order of the cards is as probable as any other provided that the order is precisely described. Given any one such order *P*, it is, of course, just as probable that in shuffling, *P* will turn into the order (call it *Q*) in which the pack is arranged in suits as that *Q* would turn into *P*. But if *P* is described simply as haphazard, there is a vast number of states other than *P* which are also haphazard. Thus, although a shuffling which turns *Q* into *P* is no more probable than one which turns *P* into *Q*, there are far more shufflings which turn *Q* into a state abstractly described as haphazard than there are shufflings which turn a particular haphazard state—say, *P*—into *Q*.

Suppose we started with our cards arranged in suits, the state *Q*. If we shuffled them, they would soon get into what we should call a well-shuffled state. Nevertheless, if we went on shuffling long enough, we should eventually get back to the unshuffled state *Q*. This illustrates the following interesting point. Let us for the moment toy with the almost certainly false cosmological hypothesis that the universe is a finite nonexpanding collection of particles without spontaneous creation or annihilation. Then, just as with our pack of cards, such a universe will eventually return to any given state. The universe will get more and more shuffled until we get the so-called heat death, in which everything is a featureless uniformity and will then become less and less disordered. In the era in which, as we should put it, the universe was getting less disordered, time would seem to run in the opposite direction to that in which it seems to run to us. (Thus, denizens of this era would still say that the universe was getting more disordered.) Indeed, there would be an infinite sequence of cosmic eras, much as is supposed in some Buddhist cosmologies, except that time would seem to run in opposite ways in alternate eras. In a sufficiently large view there would be temporal symmetry in this universe, though not on the scale of any single cosmic era. This is what makes the hypothesis of a finite nonexpanding universe philosophically instructive, even though it is probably contrary to fact.

TRACE FORMATION AND ENTROPY. It is now possible to deal with the formation of traces. Although a wide, relatively isolated part of the universe is increasing in its state of being shuffled, or, to use the more precise notion developed by physicists, in its entropy, subsystems of the wider system may temporally decrease in shuffling, or entropy. Thus, an isolated system, such as that consisting of a cube of ice in a beaker of water, may well have lower entropy than its surroundings. This reduction of entropy is bought at the expense of a more than compensating increase of entropy in the surroundings. There will, for example, be an increase of disorderliness in the system containing the coal and air that react chemically and drive the generators that provide the electric power that drives the refrigerator that makes the ice cube. (The system consisting of coal and oxygen is a more highly ordered one than is that which consists of the ashes and used up air.) Eventually the ice cube melts and becomes indistinguishable from the water in which it floated.

BRANCH SYSTEMS. The formation of a trace is the formation of a subsystem of temporarily lower entropy than that of its surroundings, and the trace is blotted out when the entropy curve of the subsystem rejoins that of the larger system. A footprint in sand is a temporarily highly ordered state of the sand; this orderliness is bought at the expense of an increased disorderliness (metabolic depletion) of the pedestrian who made it, and this extra orderliness eventually disappears as a result of wind and weather. Hans Reichenbach (1956) calls such systems of temporarily lower entropy "branch structures." It is an observable fact, and one to be expected from considerations of statistical thermodynamics, that these branch structures nearly all (in practice, quite all) go in the same direction. This direction defines a temporal direction for the universe or at least for our cosmic era of it.

On investigation it will be seen that all sorts of traces, whether footprints on sand, photographs, fossil bones, or the like, can be understood as traces in this sense. Indeed, so are written records. The close connection between information and entropy is brought out in modern information theory, the mathematics of which is much the same as that of statistical thermodynamics. A coherent piece of prose is an ordered part of the universe, unlike a completely random sequence of symbols.

It is possible that the formation of branch systems may be linked to deeper cosmological facts. Thomas Gold (1958, 1962) has argued persuasively that the formation of such a system is possible only because the universe provides a sink for radiation, and this is possible, again, only because of the mutual recession of the galaxies. It may therefore ultimately be the expansion of the universe that accounts for the direction of time. Beyond noting this interesting suggestion of a link between the small-scale and large-scale structure of the cosmos, we can for our present purposes take the formation of branch sys-

tems for granted without linking it to uncertain cosmological speculations.

POPPER'S ACCOUNT. The theory of branch systems outlined above has been developed rigorously by Reichenbach and Grünbaum, whose work partly goes back to that of Ludwig Boltzmann (1895). (A rather similar account of temporal direction has been independently given by O. Costa de Beauregard [1963].) We must now consider a different account of the direction of time, one that was conceived by Karl Raimund Popper.

Slightly changing Popper's example, consider a spherical light wave emitted from a source, as when a small electric bulb is turned on. Consider how this process would look in reverse. We should have a large spherical wave contracting to a point. This would be causally inexplicable. In order to get a spherical light wave coming in from the depths of an infinite space, we should have to suppose a coordinated set of disturbances at every point of a vast sphere, and this would require a deus ex machina. Moreover, this would still not provide the reverse of an outgoing wave expanding indefinitely. Thus, although the contracting wave is as much in accordance with the laws of optics as is the expanding one, it still is not compatible with any physically realizable set of initial conditions. Once more, as with the Reichenbach-Grünbaum solution, it can be seen that temporal asymmetry arises from initial, or boundary, conditions, not from the laws of nature themselves.

Popper's criterion of temporal direction does not shed light on the concept of trace, as does the criterion of branch systems. And traces, particularly memory traces, give us our vivid sense of temporal asymmetry in the world. It is also interesting that if we consider a finite but unbounded nonexpanding universe, a contracting spherical wave would be physically realizable. Just as an expanding series of concentric circles on the earth's surface which have their original center at the North Pole would become a series of circles contracting to the South Pole, so in a symmetrical, finite, but unbounded universe a spherical wave expanding from a center would eventually become a contracting wave, shrinking to the antipodal point of the point of emission. If we included the facts of radiation in our finite nonexpanding universe, we should have to suppose a finite but unbounded space, and Popper's criterion of temporal direction would become inapplicable. Including such facts would therefore also not conflict with our supposition of alternate cosmic eras in such a universe. In such a universe the Reichenbach-Grünbaum account of temporal direction for particular cosmic eras would still be applicable. There are still anthropocentricities to be brought to light, a task which has been impressively achieved by Huw Price in his book *Times Arrow and Archimedes' Point* (1996). He has clearly discussed the time symmetry (or one might say CPT symmetry) of microphysics. On the macro level, causation is at least in our cosmic era asymmetrical because the concept of it is closely related to that of agency and so to the temporal asymmetry of memory traces.

What is presented here is not an analysis of the ordinary language concept of earlier and later. This is learned to some extent ostensively, and we may perfectly well know how to use words such as "earlier" and "later" without knowing anything about entropy or branch systems. As Wittgenstein might have said, "We know the language game." Here the concern is with a deeper problem: what are the general features of the universe which enable us to play the language game? Indeed, if the universe did not contain traces, it would be impossible for there to be any thought at all. It should be noted that Mellor in his aforementioned book rejects the relevance of considerations of entropy and the like and relies on the notion of probability: the cause is an event that raises the objective chance of the event that is the effect. As mentioned above, Tooley also has a causal account. Even so, considerations of entropy could be needed to explain the asymmetry of causation on the macro level. On the micro level, causation is time symmetric and Price has neatly suggested defending locality, and perhaps hidden variables, in quantum mechanics and in the face of John Bell's well-known inequality, by means of backward causation. Curiously, according to Price, Bell had once considered such a solution but had rejected it for dubious philosophical reasons connected with the notion of free will.

COMPROMISE THEORIES

Storrs McCall and Michael Tooley have proposed theories that contain elements of both tensed and tenseless theories. Tooley, in his *Time, Tense, and Causation* (1997), worked out a sophisticated theory that is partly similar to one that C. D. Broad proposed in his *Scientific Thought* (1923). According to this view, only past and future are real and the universe is continually getting bigger as more and more of the future becomes present and past. Tenseless theorists will still see this as open to the objections to notions of time flow and of absolute becoming that were canvassed above. So also will they see McCall's theory according to which reality keeps getting smaller. McCall is inspired by the Everett-Wheeler interpretation of quantum mechanics. Space-time reality is like a giant poplar

tree with branches corresponding to possible futures, with trunk, branches of branches, and so on, all pointing up in timelike directions. At every interaction between particles, branches (real possibilities) get lopped off. According to the tenseless theorist, reality must be like a stack of poplar trees, ordered according to the inclusiveness of the sets of branches. The mind boggles. Tooley's (though not McCall's) theory requires an absolute present and Tooley is bold enough to consider modifying special relativity. However, a reconciliation with special relativity could have been acquired at less cost as follows. The equality in all directions of the cosmic background radiation may give an approximation to a preferred frame of reference at each point of space. This will, because of the expansion of the universe, yield a curved hypersurface of cosmic simultaneity. Tooley defends his view of the increase of reality against the objection that it requires a hypertime. However, time travel is not like space travel because we may travel to a place, say the Taj Mahal, where we have not been before. The four-dimensional equivalent of a place is a timelike world-line, which in the example may intersect the world line of the Taj Mahal. The space of commonsense talk and of Newton's *Principia* is a continuant, not like the atemporal space of Euclid. Tooley's cutting off of the future may put in question the explanatory (as opposed to instrumental) value of full Minkowski space, though perhaps less so than presentism.

The tenseless four-dimensional account sits well with mereology, the theory of part and whole. Indeed some philosophical problems come out as easily as shelling peas when one goes four-dimensional. Consider Robert Louis Stevenson's story of Dr. Jekyll and Mr. Hyde, in which the personalities of the virtuous Jekyll and the criminal Hyde alternate in the one body. Mereology distinguishes three objects, the spatiotemporally scattered objects Jekyll and Hyde and the continuous fusion of these two. The problem is not one about identity, which is a clear notion in logic, but about "person" and the problems about these are more legal and psychiatric than philosophical.

CAUSAL THEORIES OF TIME

There are theories of the structure of time, or of space-time, that are based on the notion of causality. Objections to such theories have been made as follows (Smart 1987). How do we deal with points of space-time that are not occupied with events that are neither causes nor effects? Perhaps we could rely on causal connectibility and not on connectedness. Connectibility is a modal notion and so will not be liked by philosophers such as those influenced by W. V. Quine, who are suspicious of modality. In special relativity the notion of connectibility can be defined directly in terms of the geometry of Minkowski space by that of belonging in the same double light cone and then properties of space-time defined by axioms. Still, in face of the beautiful clarity of geometry we may prefer to characterize space-time directly, without trying to define the geometry by reference to causality. Tooley avoids these objections because he has an absolute theory of space-time and uses causality simply to define temporal direction. Possibly some of these objections make difficulty for Mellor who has a relational theory. However his notion of probability is that of objective chance and may depend on a theoretical posit and avoid modality. Tooley also needs a realistic theory of causality which some philosophers will find problematic.

TIME AND FREE WILL: THE SEA FIGHT TOMORROW

It is sometimes thought that the picture of the world as a space-time manifold is incompatible with free will. It is thought that if a single action of one's future actions exists (tenselessly) in the space-time manifold, then it is fated that the person will do this action; one cannot be free not to do it. To evade this conclusion, philosophers have sometimes been inclined to reject the theory of the manifold and also to deny that propositions about the future have to be either true or false. This view can be contested at several levels. First, the fact that this singular future action exists in the space-time manifold does not mean that the person is fated to do it, in the sense that the person comes to do it independently of what it was he or she does in the meantime. It will still be that person's choice. Second, the doctrine of the space-time manifold does not even imply determinism. Determinism asserts that the laws of nature connect earlier and later spatial cross sections of the manifold in a determinate way, whereas indeterminism denies this. Indeterminism is compatible with the theory of the manifold as such but is no friend to free will. Acting by pure chance is not being free. Third, it could be argued that free will is perfectly compatible with determinism anyway. On three counts, therefore, we may assert that the theory of space-time has, in fact, nothing at all to do with the question of free will.

Aristotle canvassed some of these matters in his well-known passage about the sea battle (*De Interpretatione*, ch. 9). Aristotle held that it is necessary that either there will be a sea battle tomorrow or there will not be, but that

it is not necessary that there will be a sea battle tomorrow, nor is it necessary that there will not be a sea battle tomorrow. He held, however, that all present and past events are necessary, as are some future ones, such as an eclipse of the moon. It is clear, therefore, that Aristotle's notion of necessity here is not the modern notion of logical necessity. Nor by "necessary" can he even mean "predictable" or "retrodictable." Because past events, though not all retrodictable, may have at least left traces, perhaps Aristotle may have meant by "necessary" something like "knowable in principle." But how about past events whose traces have been blotted out? It is hard to give a coherent interpretation of Aristotle here, and certainly to try to give one would be to go into metaphysical subtleties not especially connected with time. Some commentators have interpreted Aristotle as saying that the proposition "There will be a sea battle tomorrow" is neither true nor false. It would seem, however, that this was not Aristotle's view.

Finally, it must be pointed out that the difference between past and future is misleadingly expressed by the common remark that we can change the future but not the past. It is true that we can affect the future and we cannot affect the past. We cannot, however, *change* the future, for the future is what it will be. If a person decides to take the left-hand fork in a road instead of the right-hand one, that person has not changed the future, for in this case the future *is* that person's going left. To talk of changing the future is indeed to relapse into talking of events changing and of the notion of passage.

See also Causal Approaches to the Direction of Time; Physics and the Direction of Time; Time, Being and Becoming.

Bibliography

GENERAL WORKS

An important book is Hans Reichenbach, *The Philosophy of Space and Time* (New York: Dover, 1958), which is a translation of a book published in German in 1928. Reichenbach partially modified his ideas in this book in his later work *The Direction of Time* (Berkeley: University of California Press, 1957). A wide-ranging book on the philosophy of time is *The Natural Philosophy of Time* by the cosmologist G. J. Whitrow (London: Nelson, 1961). Other important books are Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Knopf, 1963), and O. Costa de Beauregard, *La notion du temps* (Paris: Hermann, 1963) and *Le second principe de la science du temps* (Paris: Éditions du Seuil, 1963). See also Adolf Grünbaum's long essay "The Nature of Time" in *Frontiers of Science and Philosophy*, edited by R. G. Colodny (Pittsburgh: University of Pittsburgh Press, 1962).

A valuable essay is Wilfrid Sellars's "Time and the World Order," in *Minnesota Studies in the Philosophy of Science*, vol. 3 (Minneapolis: University of Minnesota Press, 1962). An earlier essay of considerable value is the article "Time" by C. D. Broad in *Encyclopedia of Religion and Ethics*, vol. 12 (Edinburgh and New York: 1921). Several essays of interest are also in *The Problem of Time*, University of California Publications in Philosophy, vol. 18 (Berkeley: University of California, 1935).

A book of readings on space and time is *Problems of Space and Time*, edited and with an introduction by J. J. C. Smart (New York: Macmillan, 1964). The history of philosophical thought about time is presented in M. F. Cleugh, *Time and Its Importance for Modern Thought* (London: Methuen, 1937).

ST. AUGUSTINE'S PUZZLES

Augustine's reflections on time are to be found in his *Confessions*, bk. 11, chaps. 14–28; a good translation is A. C. Outler, *St. Augustine: Confessions and Enchiridion*, vol. 7 of the Library of Christian Classics (Philadelphia: Westminster, 1955). Augustine much influenced Wittgenstein. For evidence of this see Ludwig Wittgenstein, *The Blue and Brown Books*, p. 26 (Oxford: Harper, 1958). W. H. Watson, a physicist who had attended lectures on philosophy by Wittgenstein, quotes a passage from Augustine on time (in order to show what a typical philosophical problem is like) in the first chapter of his *On Understanding Physics* (Cambridge: The University Press, 1938). Part of the article by J. N. Findlay, "Time: A Treatment of Some Puzzles," in *Logic and Language*, first series, edited by A. G. N. Flew (Oxford: Blackwell, 1951), discusses Augustine's puzzles. See also Ronald Suter, "Augustine on Time, With Some Criticisms from Wittgenstein," in *Revue internationale de philosophie* 16 (1962): 319–332.

THE MYTH OF PASSAGE

On the topic of the myth of passage see especially D. C. Williams's brilliant criticism in "The Myth of Passage" in *Journal of Philosophy* 48 (1951): 457–472. In chap. 35 of C. D. Broad, *Examination of McTaggart's Philosophy*, vol. 2, pt. 1 (Cambridge: The University Press, 1938), are relevant arguments against the notion of passage, even though in the end Broad does not free himself from it. An earlier view of Broad's is given in his *Scientific Thought* (London: K. Paul, Trench, Trubne, 1923). On Broad's changing views about time see C. W. K. Mundle, "Broad's Views About Time," in *The Philosophy of C. D. Broad*, edited by P. A. Schilpp (La Salle, IL: Tudor, 1959). A criticism of the notion of passage is in J. J. C. Smart, "The River of Time," in *Essays in Conceptual Analysis*, edited by A. G. N. Flew (London: Macmillan, 1956), and "Spatialising Time" in *Mind* 64 (1955): 239–241. A contrary point of view is defended by A. N. Prior, "Changes in Events and Changes in Things" (The Lindley Lecture, University of Kansas, 1962); "Time After Time," in *Mind* 67 (1958): 244–246; and "Thank Goodness That's Over," in *Philosophy* 34 (1959): 12–17. The last mentioned article defends tensed theories as explaining the difference between our attitudes to past and future pains respectively. A reply by Jonathan Cohen to the last article is to be found in the same volume, and a recent explanation of a biological point of his is in "'Thank Goodness That's Over': The Evolutionary Story," by J. Maclaurin and H. Dyke

in *Ratio* 15 (2002): 276–292. Ned Markosian has defended passage in his "How Fast Does Time Pass?" in *Philosophy and Phenomenological Research* 53 (1993): 829–844.

J. McT. E. McTaggart's argument for the unreality of time is to be found in his *Philosophical Studies*, chap. 5 (London: E. Arnold, 1934; originally published as an article in *Mind*, 1908), and in his *Nature of Existence*, vol. 2, chap. 33 (Cambridge: The University Press, 1927). For criticisms of this see C. D. Broad's *Examination of McTaggart's Philosophy*; Paul Marhenke's article in the book *The Problem of Time*; D. W. Gotshalk's "McTaggart on Time" in *Mind* 39 (1930): 26–42; and part of D. F. Pears's article "Time, Truth and Inference," in *Essays in Conceptual Analysis*, edited by A. G. N. Flew. On the other side see Michael Dummett, "A Defense of McTaggart's Proof of the Unreality of Time," in *Philosophical Review* 69 (1960): 497–504; and L. O. Mink, "Time, McTaggart and Pickwickian Language," in *Philosophical Quarterly* 10 (1960): 252–263. A sympathetic and scholarly work on McTaggart is P. T. Geach, *Truth, Love, and Immortality* (London and Berkeley: Hutchinson, 1979).

On tenses and similar token-reflexive expressions see Hans Reichenbach, *Elements of Symbolic Logic*, secs. 50–51 (New York: Macmillan, 1947); Nelson Goodman, *The Structure of Appearance*, chap. 11 (Cambridge, MA: Harvard University Press, 1951); and Bertrand Russell, *An Inquiry into Meaning and Truth*, chap. 7 (New York, 1940). Also see Yehoshua Bar-Hillel, "Indexical Expressions," in *Mind* 63 (1954): 359–379; Jonathan Cohen, "Tense Usage and Propositions," in *Analysis* 11 (1950–1951): 80–87; and R. M. Gale, "Tensed Statements," in *Philosophical Quarterly* 12 (1962): 53–59, together with ensuing discussion notes on this. The article by Sellars, "Time," has much on tenses. Zeno Vendler, "Verbs and Times," in *Philosophical Review* 66 (1957): 143–160, shows that tenses have more functions than one might first suppose. A tense logic is worked out by A. N. Prior in his *Time and Modality* (Oxford: Clarendon, 1957); and R. M. Martin, in his review of this book in *Mind* 68 (1959): 271–275, questions whether this is legitimately part of logic. See also Jonathan Cohen's critical notice of the same book in *Philosophical Quarterly* 8 (1958): 266–271. A tenseless language is advocated by W. V. Quine, *Word and Object*, sec. 36 (Cambridge, MA: Technology Press of the Massachusetts Institute of Technology, 1960). Tensed language is advocated by J. N. Findlay in his article, "An Examination of Tenses," in *Contemporary British Philosophy*, edited by H. D. Lewis (New York: Macmillan, 1956). The token reflexive approach has been well defended by Heather Dyke in several articles, especially "Tokens, Dates and Tenseless Truth Conditions," in *Synthese* (2002): 329–351, and "Tensed Meaning: A Tenseless Account," in the *Journal of Philosophical Research* 28 (2003): 65–81. The tenseless metalinguistic account of the semantics of indexicals is both defended and attacked in articles in *The New Theory of Time*, by L. N. Oaklander and Quentin Smith (New Haven, CT: Yale University Press, 1994). See also J. J. C. Smart, "Time and Becoming," reprinted in his *Essays Metaphysical and Moral* (Oxford: Blackwell, 1987), which contains the conjecture mentioned in the text of a possible source of the illusion of the flow of time. Jeremy Butterfield, in his "Seeing the Present," in *Mind* 93 (1984): 161–176, relates the different ways that common sense sees space and time respectively to the difference between the high velocity of light compared with the timescale of our physiological and electrochemical processes. Also Smart's criticism of causal theories of time and his defense of the reality of the future are reprinted in the same volume. Quentin Smith defends his own tensed theory in his *Language and Time* (New York: Oxford University Press, 1993).

Other articles are R. G. Collingwood, "Some Perplexities About Time," in *PAS* 26 (1925–1926): 135–150; and the symposium "Time and Change" by J. Macmurray, R. G. Braithwaite, and C. D. Broad in *PAS*, Supp., Vol. 8 (1928): 143–188. On the status of the past see A. J. Ayer, "Statements About the Past," in his *Philosophical Essays* (London: Macmillan, 1954).

See also Richard Taylor, "Spatial and Temporal Analogies and the Concept of Identity," in *Journal of Philosophy* 52 (1955): 599–612; and "Moving About in Time" in *Philosophical Quarterly* 9 (1959): 289–301; as well as Bernard Mayo, "Objects, Events, and Complementarity," in *Philosophical Review* 70 (1961): 340–361.

F. H. Bradley's argument for the unreality of space and time is given in his *Appearance and Reality*, 2nd ed., chap. 2 (Oxford: Clarendon, 1930). Henri Bergson's accounts of time and duration are given in his *Time and Free Will* (New York: Macmillan, 1910), *Matter and Memory* (New York: Macmillan, 1911), and *Introduction to Metaphysics* (London: Putnam, 1913). Bertrand Russell in his *History of Western Philosophy* (London: Allen and Unwin, 1945) gives a succinct criticism of Bergson. Like Bergson's, A. N. Whitehead's metaphysics took for granted a form of the myth of passage. His views are to be found especially in *An Enquiry concerning the Principles of Natural Knowledge* (Cambridge: The University Press, 1920), chaps. 3–6, and parts of *Process and Reality* (Cambridge: The University Press, 1929). See also V. C. Chappell, "Whitehead's Theory of Becoming," in *Journal of Philosophy* 58 (1961): 516–528.

SPACE-TIME

Hermann Minkowski's classic paper "Space and Time" can be found in *The Principle of Relativity*, a collection of papers by Einstein and others, translated by W. Perret and G. B. Jeffery, with notes by Arnold Sommerfeld (London: Methuen, 1923). Popular accounts can be found in A. S. Eddington, *Space, Time and Gravitation* (Cambridge: The University Press, 1920), and Moritz Schlick, *Philosophy of Nature*, chap. 7 (New York: Philosophical Library, 1949). Milič Čapek, in his *The Philosophical Impact of Contemporary Physics* (Princeton, NJ: Van Nostrand, 1961), criticizes the theory of the space-time manifold and defends the concept of becoming.

ABSOLUTE AND RELATIONAL THEORIES

A relational theory of space and time is defended by Leibniz. See especially his third and fifth papers in *The Leibniz-Clarke Correspondence*, edited by H. G. Alexander (Manchester: Manchester University Press, 1956) A brilliant argument against Leibniz is in Graham Nerlich, *What Spacetime Explains* (Cambridge, U.K.: Cambridge University Press, 1994). On space-time in the general theory of relativity see Adolf Grünbaum's paper "The Philosophical Retention of Absolute Space in Einstein's General Theory of Relativity" in *Problems of Space and Time*, edited by J. J. C. Smart (New York: Macmillan, 1964), and references given therein. Also see Graham Nerlich, *The Shape of Space*, 2nd

ed. (Cambridge, U.K.: Cambridge University Press, 1994). The issue between three and four dimensionalism is thoroughly discussed in Theodore Sider, *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Clarendon, 2001).

TIME AND THE CONTINUUM

A good discussion of the paradoxes of Zeno will be found in Adolf Grünbaum, *Modern Science and Zzeno's Paradoxes* (London: Allen and Unwin, 1968). Since 1951 many articles on Zeno's paradox of Achilles and the tortoise have appeared in *Analysis*. See also V. C. Chappell, "Time and Zeno's Arrow," in *Journal of Philosophy* 59 (1962): 197–213; and Harold N. Lee, "Are Zeno's Paradoxes Based on a Mistake?" in *Mind* 74 (1965): 563–570. Also of interest is Paul Benacerraf, "Tasks, Super-Tasks and the Modern Eleatics," in *Journal of Philosophy* 59 (1962): 765–784. A useful account of Zeno's paradoxes is to be found in Kathleen Freeman, *Pre-Socratic Philosophers: A Companion to Diels, Fragmente der Vorsokratiker*, 3rd ed. (Oxford: Basil Blackwell, 1946).

Kant's antinomies about space and time occur in *The Critique of Pure Reason*. There is a translation of this book by Norman Kemp Smith (London: Macmillan, 1929). Zeno's and Kant's antinomies are discussed by Bertrand Russell in lectures 6 and 7 of *Our Knowledge of the External World* (London: W. W. Norton, 1922). See also C. D. Broad, "Kant's Mathematical Antinomies," in *PAS* 55 (1954–1955): 1–22. The commentary by Edward Caird, mentioned in the present article, is *The Critical Philosophy of Immanuel Kant* (Glasgow: J. Maclehose, 1889).

THE DIRECTION OF TIME

Besides Reichenbach's book *The Direction of Time* and the book by Grünbaum, *Philosophical Problems*, see especially Adolf Grünbaum's paper "Carnap's Views on the Foundations of Geometry" in *The Philosophy of Rudolf Carnap*, edited by P. A. Schilpp (La Salle, IL: Open Court, 1962), which, despite its title, contains a thorough discussion of the present problem, and Grünbaum's essay "The Nature of Time." See also Erwin Schrödinger's fine paper "Irreversibility" in *Proceedings of the Royal Irish Academy* 51 (1950): 189–195; and Norbert Wiener, "Newtonian and Bergsonian Time," which is chap. 1 of *Cybernetics*, 2nd ed. (New York: M.I.T. Press, 1961). Also see Ludwig Boltzmann, "On Certain Questions of the Theory of Gases," in *Nature* 51 (1895): 413–415. Reichenbach's book depends to a great extent on Boltzmann's ideas. There is a readable treatment of some of these issues in the final appendix of Schlick's *Philosophy of Nature*. A different solution to the problem is to be found in notes by K. R. Popper in *Nature* 177 (1956): 538; also vol. 178 (1956): 382; vol. 179 (1957): 1,297; and vol. 181 (1958): 402–403, in connection with which see the note by E. L. Hill and Adolf Grünbaum, in *Nature* 179 (1957): 1,296–1,297. See also O. Costa de Beauregard, "L'Irreversibilité quantique, phénomène macroscopique," in *Louis de Broglie*, edited by A. George (Paris, 1953). Grünbaum has examined Popper's view in his essay "Popper on Irreversibility" in *The Critical Approach to Science and Philosophy: Essays in Honor of Karl Popper*, edited by Mario A. Bunge (New York: Free Press of Glencoe, 1964). There are two beautiful articles titled "The Arrow of Time" by the cosmologist Thomas Gold in *La Structure et l'évolution de l'univers*, proceedings of the eleventh Solvay Conference, pp. 81–91 (Brussels: R. Stoops, 1958), and in *The American Journal of Physics* 30 (1962): 403–410. "The Direction of Time" by Max Black in his *Models and Metaphors* (Ithaca, NY: Cornell University Press, 1962), is written from the point of view that scientific considerations are irrelevant to the problem of the direction of time. D. H. Mellor rejects the relevance of considerations of entropy and statistical mechanics in his *Real Time II* (Cambridge, U.K.: Routledge, 1998). An absolutely outstanding discussion of temporal symmetry and asymmetry in which he identifies unrecognized anthropocentric confusions is Huw Price, *Time's Arrow and Archimedes' Point* (New York: Oxford University Press, 1996).

A readable discussion of the experiment by James H. Christenson, James W. Cronin, Val L. Fitch, and René Turlay, which suggests a possible violation of time symmetry in the laws of nature themselves, can be found in Eugene P. Wigner's article "Violations of Symmetry in Physics" in *Scientific American* 213 (December 1965): 28–42.

TIME AND FREE WILL: THE SEA FIGHT TOMORROW

On fatalism see R. D. Bradley, "Must the Future Be What It Is Going To Be?" in *Mind* 68 (1959): 193–208; Richard Taylor, "Fatalism," in *Philosophical Review* 71 (1962): 56–66, with the discussion on this by Bruce Aune in the same volume, pp. 512–519; and A. J. Ayer, "Fatalism," in his *The Concept of a Person and Other Essays* (London: Macmillan, 1963). On the sea battle see Aristotle, *De Interpretatione*, chap. 9. Extensive notes and a translation can be found in J. L. Ackrill's *Aristotle's Categories and De Interpretation,* vol. 1 of the complete works of Aristotle edited by Jonathan Barnes (Princeton; NJ: Princeton University Press, 1984). This passage has also been translated and discussed by G. E. M. Anscombe in "Aristotle and the Sea-Battle" in *Problems of Space and Time*. See also Colin Strang, "Aristotle and the Sea Battle," in *Mind* 69 (1960): 447–465. Many journal articles on the subject, following on D. C. Williams's interesting "The Sea-Fight Tomorrow," appear in *Structure, Method, and Meaning*, edited by Paul Henle, Horace M. Kallen, and Susanne K. Langer (New York: Liberal Arts Press, 1951). See especially the discussion note "Professor Donald Williams on Aristotle" by Leonard Linsky and the rejoinder by Williams in *Philosophical Review* 63 (1954): 250–255, and Richard Taylor, "The Problem of Future Contingents," and Rogers Albritton's reply in *Philosophical Review* 66 (1957): 1–46. The seventeenth-century English philosopher Thomas Hobbes also wrote on the sea-fight; see his *Works*, edited by William Molesworth, vol. 4, p. 277 (London: J. Bohn, 1839), and discussion by A. G. N. Flew, "Hobbes and the Seafight," *Graduate Review of Philosophy* 2 (1959): 1–5.

Other references are to Storrs McCall, "Objective Time Flow," in *Philosophy of Science* 43 (1976): 337–362; and his "A Dynamic Model of Temporal Becoming," *Analysis* 44 (1984): 172–176; and to McCall's book, *A Model of the Universe: Space-Time, Probability and Decision* (Oxford: Clarendon, 1994). Mellor's *Real Time II* was published in London in 1998. Michael Tooley's *Time, Tense and Causation* was published in Oxford in 1997. See also Michael Tooley, "The Metaphysics of Time" in *The Argument of Time*, edited by Jeremy Butterfield, pp. 21–42 (London: Oxford University

Press, 1999), and "Basic Tensed Sentences and their Analysis" in *Time, Tense, and Reference*, edited by Aleksander Jokic and Quentin Smith, pp. 409–447 (Cambridge, MA: MIT Press, 2003). John Bigelow defends presentism in his "Presentism and Properties," in *Philosophical Perspectives* 10, *Metaphysics* (1996): 35–52.

J. J. C. Smart (1967, 2005)

TIME, BEING, AND BECOMING

The major debate in the philosophy of time, being, and becoming is between defenders of the tenseless theory of time and defenders of the tensed theory of time. During the late twentieth century into the early twenty-first century, the tenseless theory of time was defended by such philosophers as D.H. Mellor, Graham Nerlich, and L. Nathan Oaklander. The tenseless theory implies that temporal features of events consist only of relations of simultaneity, earlier, and later than, and that all events are ontologically equal, regardless of when they occur. The tensed theory, which has many versions, is advocated by such philosophers as William Lane Craig, Quentin Smith, and Michael Tooley. The tensed theory of time implies that some or all of the words *past*, *present*, and *future* are needed to describe time, although what is understood by the words *future*, *present*, and *past*, or by their usage as parts of phrases or sentences (e.g., whether or not they express analyzable or unanalyzable concepts) is a matter that varies among tensed theorists.

THE OLD AND NEW TENSELESS AND TENSED THEORIES OF TIME

For most of the twentieth century, the debate was between defenders of the old tenseless theory of time and defenders of the old tensed theory of time, concerning whether or not tensed sentence tokens are translatable by tenseless sentences. If a tensed sentence token, call it S, such as the sentence token "John was running" can be translated by a tenseless token, such as "John is (tenseless) running earlier than S," then the tensed token S conveys no more temporal information than the tenseless token. Consequently, the defender of the old tenseless theory of time maintained that temporal properties and relations can consist only of the relations of earlier than, later than, and simultaneous with. Some of the main developers of the old tenseless theory are Bertrand Russell (1903, 1906, 1915)—Russell is the first twentieth century defender of the tenseless theory against the tensed theory of time—Hans Reichenbach (1947), J.J.C. Smart (1963, 1966), and Adolf Grünbaum (1973). Smart (1980) was also one of the main founders of the new tenseless theory of time.

Proponents of the old tensed theory of time argued that these sentence tokens cannot be translated. For example, "John (is) running earlier than S" does not convey the temporal information of whether John's running is past, present, or future. Because "John was running" conveys that it is past, this sentence token cannot have the same semantic content (or the same meaning, or express the same proposition) as the tenseless token, and therefore cannot be translated by the tenseless token. Some of the most influential defenders of the old tensed theory of time are C.D. Broad (1923)—who is the first twentieth century defender of the tensed theory and critic of the tenseless theory—A. N. Prior (1967, 1968, 1979), Richard Gale (1962, 1968), and George Schlesinger (1981).

In response to criticisms advanced by the old tensed theory of time, defenders of the tenseless theory largely accepted the argument of Gale and others that tensed sentence tokens cannot be translated by tenseless ones; however, the tenseless theorists now argued that the truth conditions of tensed sentence tokens are tenseless. For example, Mellor (1981) argued that the token S of "John was running" is not translatable by a token "John is (tenseless) running earlier than S", but is true if, and only if, John is (tenseless) running earlier than S. The new tenseless theory of time was in place by 1981, due primarily to the independent work of Mellor (1981) and Smart (1980) (see also Anderson and Faye [1980], Faye [1981], and Oaklander [1984]). The main developments and defenses of various versions of the new tenseless theory from the mid-1980s to the early twenty-first century were made for the most part by L. Nathan Oaklander, but also by Heather Dyke (2002a, 2002b, 2003), Robin Le Poidevin (1992, 2003), Graham Nerlich (1998), L.A. Paul (1997), J. M. Mosersky (2000), and others.

The emergence of the new tenseless theory in the 1980s inspired the new tensed theory of time, whose unifying theme was a criticism of the new tenseless theory and the development of ontologies for a tensed theory that were able to overcome the hurdles set by the new tenseless theorists. Criticisms of one of the two main versions of the new tenseless theory, Mellor's token-reflexive theory, appeared in Graham Priest's (1986, 1987) work, and criticisms of the two main versions of the new tenseless theory (Smart's and Mellor's) appeared in Smith's (1987, 1993) work.

The classification of the new tenseless theories of time into two versions, the token-reflexive version and

the date-involving version, was made in the course of Smith's (1987, 1993) criticisms of these theories. One criticism of the former is that the tenseless token-reflexive theory of tensed sentence tokens in natural language is mistaken because (among other reasons) the truth conditions of a tensed sentence token S cannot be about S itself, as well as what S is about. Suppose there are two simultaneous utterances, the utterance U of "The talk will begin in an hour" and the utterance S of "The talk will begin in sixty minutes." These two utterances, given that they occur at the same time, are logically equivalent. It is impossible for the talk to begin in an hour unless it begins in sixty minutes and vice versa. But the token-reflexive truth conditions of S and U are not logically equivalent. U is true if, and only if, the talk begins one hour later than U and S is true if, and only if, the talk begins sixty minutes later than S, whereas because "the talk begins in an hour" and "the talk will begin in sixty minutes" are logically equivalent, it is neither necessary nor sufficient for S's truth that the talk begin one hour later than U. It is not necessary because there is a possible world in which S is true, but in which U is not uttered.

Further, it is sufficient for S's truth that the talk begins one hour later than the time at which U, as a matter of fact, occurs, regardless of whether or not U occurs; if U had not occurred, S would still be true. We have two logically equivalent, simultaneous, tensed sentence tokens that have logically inequivalent truth conditions—which not only fails to explain the logical equivalence of the tensed sentence tokens, but leads to an implicit contradiction. If S and U entail each other, and S and U are each logically equivalent to their respective truth conditions clauses SC and UC, then it follows by the transitivity of logical equivalence that SC and UC are logically equivalent. Because SC and UC are not logically equivalent, SC is not a truth conditions clause for S and UC is not a truth conditions clause for U.

This and other criticisms appear to have motivated an abandonment of the new token-reflexive tenseless theory of time by its originators and developers—Mellor, Oaklander, Paul (1997), Le Poidevin (2003), and so on, as well as by critics who are tensed theorists—Craig (1996, 2000a), Peter Ludlow (1999), and so on. However, Oaklander (2003, 2004), as well as Dyke (2000a, 2002b, 2003), have spent much time developing versions of what Oaklander calls the newer token-reflexive tenseless theory, which they argue are immune to Smith's criticisms. Because Dyke's and Oaklander's theories have not yet been critically evaluated, it must be said that the token-reflexive theory, in its newer version, remains an obstacle in the tensed theorist's path.

The other version of the new tenseless theory of time is the date-theory. This may be criticized by arguing that the new tenseless date-involving truth conditions are neither necessary nor sufficient for the truth of tensed sentence tokens. It appears to be false, for example, that "Jane is running" as uttered at noon on July 1, 1994, is true if, and only if, Jane runs at noon on July 1, 1994. There are possible worlds in which the mentioned sentence utterance, call it U, is true and yet it is false that Jane is running at noon on July 1, 1994.

Suppose, for instance, that times are sets of simultaneous events and that noon on July 1, 1994, refers to the set of simultaneous events that is actually 1,993 years, six months, and twelve hours after the conventionally assigned birth date of Jesus. There is a possible world exactly similar to the actual world except for the fact that the utterance U belongs to a different set of simultaneous events, a set that includes every event included on July 1, 1994, at noon (which means it includes Jane's running), except for some minor difference; say, the set does not include the decision actually made by David to have lunch. Because U occurs simultaneously with Jane's running in this world, U is true; nonetheless, it does not occur at noon on July 1, 1994. Thus date-involving truth conditions do not appear to be necessary for the truth of tensed sentence tokens.

Suppose, in contrast, that one does not reduce times to sets of events, adopting instead a substantival theory that regards times as particulars in their own right, particulars identified by their position in a time sequence, essentially dated (and metricated) in relation to earlier and later times; times may be occupied by events or sets of events, but the times are neither identical with nor necessarily contain their occupants.

The same time (e.g., May 1, 2005, at noon) may have different occupants in different possible worlds. One of the arguments against a substantival version of the tenseless date-theory concerns the date-theory that a sentence token S of "Jane is running" that is uttered at noon on May 1, 2005, is true if, and only if, Jane is (tenseless) running on May 1, 2005, at noon. Suppose Jane is running at this time. Because we are assuming a substantival version of the date-theory, the mentioned time has the essential date property of being May 1, 2005.

In other words, the time is metricated (identified as a part of a sequence of equal-lengthed intervals and assigned a specific ordinal in this sequence, convention-

ally abbreviated as 5/1/2005) and this metricated time remains identical across possible worlds even if it has different occupants in these worlds. There is a possible world similar to the actual world except that Jane is not running at noon on May 1, 2005, and S does not occupy the time on May 1, 2005, at noon. Instead, S occupies a later time, on May 2, 2005, at noon and Jane is running at noon on May 2, 2005, in this world.

The token S of "Jane is running" on noon, May 2, in this second world is true because S occurs simultaneously with Jane's running. And yet the purported date truth conditions it is supposed to have would imply S is false because it cannot be true unless the date is May 1. But how could the token S of "Jane is running" be false if Jane is running simultaneously with the token S of "Jane is running"? This indicates that the truth condition sentence: "A token S of 'Jane is running' that is uttered at noon on May 1, 2005, is true if, and only if, Jane is (tenseless) running on May 1, 2005 at noon" is false. It is false because the token S is true in the second world even though Jane is not running on May 1 in that world (note that S is here being used in the actual world as a modally stable tag [Marcus 1961] that serves to refer directly to S in both worlds). Thus, the alleged date-involving truth condition sentence does not give us a correct necessary condition ("only if") of S's truth.

Oaklander (1994) responds to these arguments of Smith (1987, 1993) by changing the new date-theory to a still newer date-theory and thus avoids the problem Smith mentions. The newer date-theory, Oaklander says, is that the correct truth condition sentence is that the token S of "Jane is running" uttered at noon on May 1, 2005, in world W, is true at noon on May 1, 2005, in W if, and only if, Jane is (tenseless) running at noon on May 1, 2005, in W." Because the possible world W is mentioned in the truth-condition sentence, the objection based on what occurs in a different possible world is avoided.

This newer theory may seem prima facie plausible. But a closer look shows that, by virtue of being world-indexed, it is irrelevant to the semantic content, truth value, and truth conditions of the token S. If we take any true extensional sentence, such as "The sun is shining on Mount Everest at noon on May 1, 2005," substitute it for the extensional clause after the biconditional, namely, "Jane is (tenseless) running at noon on May 1, 2005," retain the world-index "in W," then we also have a true truth condition sentence for the token S-in-W. If we take any true, contingent, extensional, sentence token T, operate on it to produce the world-indexed operand T-in-W, then T-in-W is necessarily true and fulfills the criteria of being both sufficient and necessary for the truth of S-in-T. But whether or not the sun is shining on Mount Everest has no bearing on the truth or falsity of the sentence token S, which is the sentence token whose truth conditions are being discussed by the tensed theorist and the tenseless theorist. Accordingly, world-indexing the clauses before and after the biconditional does not solve the problem of the truth conditions of a token of "Jane is running" that is uttered at noon on May 1, 2005. We can see that a problem with Oaklander's newer date-theory is that it has, in effect, changed the subject.

The subject is the truth conditions of the non-world-indexed, tensed sentence token, the May 1, 2005, at noon token S, "Jane is running." Oaklander changed the subject to world-indexed sentence tokens, such as the truth conditions of S-in-W, and whether or not the tenseless date-theorist can provide tenseless truth conditions—for S-in-W does nothing to answer Smith's argument that the new tenseless date-theory of time cannot provide satisfactory truth conditions for the tensed sentence token S.

But Oaklander's modal argument is not the only objection that can be brought against Smith's arguments against the new tenseless date-theory of time. Oaklander has advanced further arguments challenging Smith's arguments against the new tenseless date-theory, as have Le Poidevin (2003), Mosersky (2000), L.A. Paul (1997), and Nerlich (1998). Furthermore, arguments in favor of a tensed date-theory have been made by Tooley (1997, 2001, 2003) who also presents arguments against Smith's criticisms of the new date-theory. Whether or not a date-theory of time is viable remains an issue upon which there is as of yet no common consensus.

An equally crucial issue concerns the relation of the new tensed theory of time to the sciences. Smith emphasized (1985, 1993) that the new tensed theorist must show that the crucial sort of scientific theses, the theses predominately found in the central observational part of the sciences, include tensed sentence tokens. These tensed tokens are used to confirm the theoretical claims of the sciences (keeping in mind, of course, the context relativity of the theoretical/observational distinction) and Smith argues that these tensed sentence tokens are logically incoherent if they lack tensed truth conditions.

A long-standing mistake, championed most influentially by Grünbaum, is that tensed statements, if they belong to the sciences, must belong to the theoretical part of physics (specifically, to the basic equations, and the semantic content of the constants and parameters in these equations). This is wrong because the semantic content of the tenses of verbs, and the semantic content of

temporal pronouns (*now*, *yesterday*) are essentially observational and by definition belong to the observational, not theoretical, statements in the sciences.

For example, Alexander Friedman's solution to the Einstein equation belongs to the theoretical part of big bang cosmology, but the observational information that the big bang occurred fifteen billion years ago essentially belongs to the observational part of big bang cosmology (see Smith 1985, 1993; Smith and Oaklander 1994). This shows that some of the more superficial evaluations of Smith's *Language and Time* (1993) are mistaken (e.g. the evaluation that it is not based on science but ordinary language analysis of the sort done in the 1950s in England). For it is based, not on ordinary language, but, instead, on the observational part of science, on confirmation theory, logic, and on the deep structure of natural languages (1993, Ch. 6.6) studied in linguistics. However, one of the most conceptually precise and accurate explorations of this notion, Nerlich's *Time and Spacetime* (1998), takes the ingenious route of eliminating the presentness part of the deep structure of a sentence, while still retaining the propositional relation.

Nerlich predicts that Smith will answer his critique by appealing to ordinary language, rather than to science (to which Nerlich appeals). But section 1.5 of *Language and Time* suggests otherwise. Smith would say that he appeals to the conditions in the universe that make true the tensed observation sentence tokens in the observational part of science. What is reported in these observation sentence tokens is the condition that the empirical datum observed is past to some degree or is or will be present in a certain amount of time. Nerlich appeals to the theoretical parts of the special and general theories of relativity. Smith appeals to the tensed observation sentence tokens that confirm the theoretical parts of special and general relativity.

Dennis Sciama (1973, pp. 24–25), for example, made the observation (relative to the observational/theoretical distinction in big bang cosmology): "in its *present* state the universe is far too dilute to be able to thermalize radiation in the time available (10^{10} years) … we conclude that at sometime *in the past* the universe must *have been* sufficiently dense to thermalize radiation.… According to the standard cosmological models the universe thus would require a universal density of at least 10^{-14} gm cm^{-3} (that is about 10^{15} times larger than the *present* mean density. [my italics]" P.A.M. Dirac (1983, p. 47) observes that "the *present* velocity of recession is 10^{-3} [my italics]" I. D. Novikov (1974, p. 273) observes that "the Universe expands isotropically with a high degree of accuracy at *the present time* …This is valid for at least some period in *the past* too." A philosopher of the observational part of science will find that the tense in the verb phrases of the observation sentence tokens are surface manifestations of the deep structure of language, a structure that includes only propositions that have presentness as a part. This deep structure, like Ludlow's (1999) deep structure, is a structure of mind-independent reality. This investigation of the deep structure of scientific observation sentence tokens is a primary task of Smith's *Language and Time*.

The misunderstanding of Smith's work as being ordinary language analysis rather than scientific analysis may be because the tenseless theory is often associated with more scientifically inclined philosophers and the tensed theory with more ordinary language inclined philosophers. This association is largely a myth. Not only Smith but also Storrs McCall (1994), Tooley (1997), Craig (2000b, 2003), Mauro Durato (1995), and many others have developed tensed theories in terms of or in relation to the physical sciences. Many tenseless theorists, such as Mellor (1981), Oaklander (1994, 2003), Dyke (2002), Le Poidevin (1992, 2003), Paul (1997), and others have based their theory in large part on analysis of ordinary language.

DISTINCTIONS BETWEEN TENSED AND TENSELESS EXISTENCE

One of the oldest and most important ontological distinctions in the philosophy of time concerns the "full/empty" versions of the tensed theory of time. Broad's theory (1923) and Tooley's theory (1997, 2001, 2003) imply an empty future and full present and past; that is, the future is nonexistent (nothing exists later than the present time) and the present and past are full (existent). Schlesinger's (1981) theory implies a full future, present, and past and, likewise, McCall's (1994) theory implies a full future, containing real possibilities, and a full present and past; in McCall's theory, the present and past are both real and actualized possibilities, whereas the future consists of real but unactualized possibilities. Bell (1987) articulates a theory with an empty past and Others, such as Prior (1967, 1968), Craig (2000a), John Bigelow (1996), Mark Hinchliff (1996), and Ludlow (1999), hold an empty past, full present, and empty future theory.

Smith and Tooley introduced new but different ways to understand the empty/full ontology. But many philosophers have misunderstood both of their (very different) ontologies to be full, tenseless ontologies. A clari-

fication of their ontologies will be helpful, starting with Smith's ontology.

Most tensed theorists, from 1996 to 2005, (with exceptions, such as Tooley and McCall) call themselves *presentists*. Many of these tensed theorists believe Prior coined this neologism as a name for his theory of temporal solipsism (only what is present is real and possesses properties) and they see themselves as developers of the Priorian tradition. But this widespread belief is because of a misunderstanding of the use of *presentism*. Prior did not coin the neologism *presentist* and never used this word even once in his entire corpus. Nor did Prior's early disciples, such as Genevieve Lloyd (1977; 1978), Ferrel Christensen (1974), and others, use the words *presentism* and *presentist*. Contrary to widespread belief, there was no standard use of this term prior to *Language and Time*, which was published in 1993.

The words *presentism* and *presentist* appear nowhere in philosophy journals and books in the 1950s, 1960s, and 1970s. In the 1980s there were two articles in which *presentism* appears; one by Robert Adams (1986), where he rejects presentism, and in a reply to Adams, where Jonathan Kvanvig (1989) defines *presentism* in a way that contemporary philosophers would call a non-presentist theory. The philosopher who first called himself a presentist and who first called the theory he was advocating presentism was Smith (1993). Far from it being the case that *presentism* was regularly used since Prior's 1950s and 1960s publications, the use of *presentism* did not become widespread until readers of Smith (1993) had time to read the book, write an article, and have it published, that is, with the first post-Smith publications beginning in 1996 (Bigelow 1996, Hintchliff 1996 and others).

By 1997 and 1998 presentism had become the most widely used name of a theory of time (replacing, for example, the names *A-theory*, *tensed theory*, *theory of temporal becoming*, and so on). The false belief that Prior and his 1970s disciples used *presentism* to name Prior's theory partly explains the false belief that Smith misused this word since he had a different theory than Prior. Thus, Smith is typically classified with Tooley and McCall as a contemporary non-presentist who takes tense seriously. The truth is the reverse. Smith correctly used *presentism* and the hundreds of contemporary philosophers who discuss presentism are misusing this word, because of their mistaken belief that it was in wide use prior to Smith (1993) to denote a Priorian version of the tensed theory of time.

The important point is not the mere terminological one that if *presentism* is used accurately (on the causal chain theory of reference), Smith is a presentist, Prior is not a presentist, and the post-1993 philosophers who call themselves presentists are not, in fact, presentists because they do not hold a version of Smith's presentism. The ontologically important issue concerns the presuppositions about the empty/full distinction that led philosophers of time to believe that Smith's presentism was a full tenseless existence theory. Philosophers interpreted him as maintaining that all times exist equally, in an irreducible, tenseless sense of *exists*. But Smith maintained exactly the opposite theory. He held that no times, events, or anything else exist tenselessly; that only one time exists in the present tensed sense; and that past and future times either no longer exist or do not yet exist.

Smith writes: "'x exists'" in the tenseless sense means 'x existed, exists, or will exist' where the middle 'exists' is present tensed … and 'x exists' in the present tensed sense means, or is logically equivalent to, 'x is present' (Smith 1993, p. 165). In fact, Smith argues that there is no tenseless semantic content of *is* or *exists* so that *tenselessly exists* is merely a syntactical string whose semantic content is *existed*, *exists* (present tense), or *will exist*.

This seems to be what post-1993 philosophers meant by their use of *presentism*, so, despite their false beliefs about the correct use of the word *presentism*, it may seem that Smith is a presentist in the same sense in which later philosophers used or misused this word. But there is one main difference: Smith (1993, 2003) argued that past and future tensed sentence tokens can be true in the sense of correspondence only if past and future events presently possess properties of pastness or futurity. Although these past and future events do not exist in the present tense sense of this word, their exemplification of pastness or futurity exists or presently obtains. For the sake of brevity, Smith says that what is past or future may be said to exist in an artifical present tense sense, namely, to presently possess pastness or futurity. If commentators on Smith's ontology distinguished this artificial present tense sense of "exists" from both the natural, genuine present tensed sense of "exists" (is present) and the reductive tenseless use of "exists," all explained in (1993, p. 165), and *if* the philosophers commented that only in the artificial present tense sense of "exists" do all times exist equally, then this would be a correct attribution (even if the artificial sense appears in only three sentences in the book [1993, p.165]. But their criticism is instead based on mistakenly attributing to Smith's times an equal, primitive, tenseless existence and a "spotlight" version of the tensed theory of time, such as Schlesinger (1981) held.

It is noteworthy that both Smith and Prior recognized that pastness and futurity have presentness as part of their meaning, which must be reflected in one's ontology. Both agree with the statement that pastness is (identically) *present* pastness. Using the example of Whitrow's lecture, Prior notes (1979, p. 258): "its pastness is its present pastness, so that although Whitrow's lecture isn't now present and so isn't real, isn't a fact, nevertheless its pastness, its *having* taken place, is a present fact, is a reality, and will be one as long as time lasts." This is also Smith's position, except Smith proceeds to develop an ontological analysis of these statements and Prior does not. Prior merely gives syntactic rules for translating tensed sentences into the syntactically regimented sentences of tense logic (which have operators such as "It was the case that," "It will be the case that," and so on). Peter Ludlow notes (1999, p. 100): "Prior never actually gave a semantics for his tense logic." In addition, Tooley (1997, p. 164) points out some problems with Prior's syntactics for his tense logic. "But, while treating tensed terms as operators on sentences may be convenient for the formulation of a logic of tense, is it also metaphysically perspicuous? I do not believe that it is. In order for a given regimentation of tensed sentences to be metaphysically perspicuous, the syntax needs to reflect the structure that would need to be present in states of affairs to render tensed sentences true." Tooley shows it does not and concludes that the tense-logical reformulation of a natural sentence "does not get one back to the state of affairs in the world that makes the original sentence true. The tense-logical formulation appears, therefore, to leave it completely obscure what sorts of states of affairs are truth-makers for tensed sentences." (Tooley, 1997, p. 166).

More recently, Smith (2002) has developed a different ontology than his (1993), a theory he calls Degree Presentism. This theory implies there are no properties of pastness, presentness, or futurity. Each entity tenselessly stands in a relation to the present of being earlier than it by a certain amount of time, being later than it by an amount of time, or being simultaneous with the present. Only the present exists to the maximal degree. What is earlier or later than the present lacks the amount of existence that is measured by its temporal distance from the present. Something one second earlier than the present is not maximally existent but rather exists to the lower degree of being one second distant from the present.

A recent, non-presentist, tensed account is Tooley's (1997) theory. Here the central ontological claim is that the past and the present are real, but the future is not, while the main semantical claims are, first, that when the terms *past*, *present*, and *future* are used in ordinary sentences, they involve an indexical element that refers directly to the time that the utterance is made; secondly, that there are non-indexical, tensed concepts that are more basic, such as the concepts of being past at time t, or future at time t, or present at time t; and, thirdly, that those more basic tensed concepts can in turn be analyzed. Thus it was claimed, for example, that the sentence "E is (tenseless) present at time t" could be analyzed, using a temporally-indexed notion of actuality, as "E is actual as of time t and nothing later than t is actual as of time t" (Tooley 2003).

The idea that the terms *past*, *present*, and *future*, as used in ordinary sentences, involve an indexical element, and that it is expressions such as *present at time t* that are more basic, suggested to some philosophers that the theory advanced by Tooley was in fact a full tenseless existence theory. For it is often held, by advocates of tensed views, as well as by defenders of tenseless approaches, that the sentence "E lies (tenseless) in the present at time t" is logically equivalent to "E is (tenseless) simultaneous with time t". But these two sentences are, Tooley argues, not equivalent. The reason is that the former, in view of the term *present*, entails the fundamental idea of the tensed theory of time, that time is dynamic, but the latter, which contains instead the word *simultaneous*, does not entail this. For because the sentence "E lies (tenseless) in the present at time t" means the same as "E is actual as of time t, and only times earlier than t are also actual as of t," the truth of this sentence entails an empty future, because it entails that no future state of affairs is actual as of time t (Tooley 2003).

Thus Tooley writes: "The analysis needed here rests upon the claim that the present is the point at which events and states of affairs come into existence, and the basic idea is that, since this view of the present entails that future events and states of affairs are not yet real, an event is present at a given time if and only if the totality of what is actual as of that time does not contain an event or state of affairs that is later than the event in question" (Tooley 2003, p. 438).

But what account can be given of the core notion on which this approach rests—that is, the concept of being actual as of a time? Is it a tensed notion, or a tenseless notion? The most natural view would seem to be that it is a tensed notion. It is true that tensed concepts are typically defined in terms of the concepts of past, present, and future, and such an account entails that the concept of being actual as of a time is not a tensed notion, because it can be argued that it is not analyzable in terms of the

concepts of past, present, and future. However, the temporally-relativized concepts of a proposition's being true at a time, and of a state of affairs being actual as of a time are integral to dynamic conceptions of time, and have no place in tenseless approaches. Accordingly, it seems natural to conclude that tensed temporal concepts are best viewed as including both tensed concepts in the narrow sense of concepts involving ideas such as past, present, and future, and also the temporally-indexed concepts of truth and actuality that are crucial for tensed conceptions of time.

Advocates of tenseless approaches to time have argued (Smart 1981, Mellor 1998), however, that the only way one can make sense of such a temporally-indexed notion of actuality is by saying that E is actual as of time t only if E occurs at or earlier than t. If this view is right, then Tooley's approach collapses into a tenseless account. But this criticism would in fact be very wide-ranging indeed, because arguably what is central to any tensed approach to time is the idea that at least some propositions can have different truth values at different times. If this is right, any tensed approach to time requires a temporally-indexed conception of truth, and this combined with a correspondence theory of truth, means that tensed approaches to time need a temporally-indexed conception of actuality. So if the latter can only be understood tenselessly, no tensed theory of time can be correct.

These explanations of Tooley's and others' theories gives a substantive presentation of the novel ideas that are currently under discussion as of 2005. The tensed/tenseless theories and debates are attracting an increasing number of philosophers. The creativity, the new and more complex arguments, and the increasingly precise conceptual distinctions exhibit the advancement or progress of philosophy in a very clear and positive light.

See also Being; Ontology, History of; Prior, Arthur Norman; Reichenbach, Hans; Russell, Bertrand Arthur William; Smart, John Jamieson Carswell; Time.

Bibliography

Adams, Robert. "Time and Thisness." *Midwest Studies in Philosophy* 11 (1986): 315–329.

Anderson, H. B. and J. Faye. "On Future Sea-Battles." In *Studien antik.* Copenhagen: Museum Forlag, 1980.

Bigelow, John. "World Enough for Times." *Nous* 25 (1991): 1–19.

Bigelow, John. "Presentism and Properties." In *Philosophical Perspectives*. Cambridge, MA: Blackwell, 1996.

Broad, C.D. *Examination of McTaggart's Philosophy*. Cambridge, U.K.: Cambridge University Press, 1938.

Broad. C.D. *Scientific Thought*. New York: Harcourt, Brace, and Co., 1923.

Christensen, Ferrel. "McTaggart's Paradox and the Nature of Time." *Philosophical Quarterly* 24 (1974): 289–299.

Craig, William Lane "Presentism: A Defense." In *Time, Tense, and Reference*, edited by A. Jokic and Q. Smith. Cambridge, MA: MIT Press, 2003.

Craig, William Lane. "Tense and the New B-theory of Language." *Philosophy* 71 (1996):5–26.

Craig, William Lane. *The Tensed Theory of Time*. Dordrecht; Boston: Kluwer, 2000a.

Craig, William Lane. *The Tenseless Theory of Time*. Dordrecht; Boston: Kluwer, 2000b.

Dirac, P.A.M. "Fundamental Constansts and Their Development in Time." In *The Physcist's Conception of Nature*, edited by J. Mehra. Dordrecht: Reidel, 1973.

Dyke, Heather. "McTaggart and the Truth About Time." In *Time, Reality, and Experience*, edited by Craig Callender. Cambridge, U.K.: Cambridge University Press, 2002b.

Dyke, Heather. "Tensed Meaning: A Tenseless Account." *Journal of Philosophical Research* 27 (2003): 67–83.

Dyke, Heather. "Tokens, Dates, and Tenseless Truth Conditions." *Synthese*: 131 (2002a): 329–351.

Durato, Mauruce. *Time, Tense and Reality*. Berlin: Spring Verlag, 1995.

Faye, Jan. *An Essay in Natural Philosophy of Time and Causation*. Copenhagen: Forlag, 1981.

Gale, Richard. *The Language of Time*. London: Routledge, 1968.

Gale, Richard. "Tensed Statements." *Philosophical Quarterly* 12 (1962): 53–59.

Geach, Peter."The Perils of Pauline." *The Review of Metaphysics* (1969).

Grünbaum, Adolf. *Philosophical Problems of Space and Time*. Dordrecht; Boston: Reidel, 1973.

Hinchliff, Mark. "McTaggart, Change, and Real Tense." In *The Importance of Time*, edited by L. Nathan Oaklander. Dordrecht; Boston: Kluwer, 2001.

Hinchliff, Mark. "The Puzzle of Change." *Metaphysics: Philosophical Perspectives* (1996): 119–136.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Kvanvig, Jonathan. "Adams on Actualism and Presentism." *Philosophy and Phenomenological Research* 5 (1989): 289–298.

Le Poidevin, R. *Change, Cause, and Contradiction: A Defense of the Tenseless Theory of Time*. Cambridge, MA: MIT Press, 1992.

Le Poidevin, R. "Tenses Need Real Times." In *Time, Tense, and Reference*, edited by Quentin Smith and A. Jokic. Cambridge, MA: MIT Press, 2003.

Levison, Arnold. "Events and Time's Flow." *Mind* 96 (1987): 341–352.

Lloyd, Genevieve. "Tense and Predication." *Mind* 86 (1977): 433–438.

Lloyd, Genevieve. "Time and Existence." *Philosophy* 53 (1978): 215–228.

Ludlow, Peter. *Semantics, Tense, and Time*. Cambridge, MA: MIT Press, 1999.
Ludlow, Peter "Presentism, Triviality, and Varities of Tensions." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman. Oxford: Clarendon Press, 2004.
Markosian, Ned. "A Defense of Presentism." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman. Oxford: Oxford University Press, 2004.
Markosian, Ned. "How Fast Does Time Fly?" *Philosophy and Phenomenological Research* 53 (1993): 829–844.
Markosian, Ned. "On Language and the Passage of Time." *Philosophical Studies* 66 (1992): 1–26.
Marcus, Ruth Barcan. "Intensionality in Modal Languages." *Synthese* (1961).
McCall, Storrs. *A Model of the Universe*. Oxford: Oxford University Press, 1994.
Mellor, D. H. *Real Time*. Cambridge, U.K.: Cambridge University Press, 1981.
Mellor, D, H. *Real Time II*. London; New York: Routledge, 1998.
Mosersky, J.M. "Tense and Temporal Semantics." *Synthese* 124 (2000): 257–279.
Nerlich, Graham. "Time as Spacetime." In *Questions of Time and Tense*, edited by Robin Le Poidevin. Oxford: Oxford University Press, 1998.
Novikov, I. D. "Isotropization of of Homogeneous Cosmological Models." In *Confirmation of Cosmological Theories with Observational Data*, edited by M. S. Longair. Dordrect: Reidel, 1974.
Oaklander. L. N. *The Importance of Time*. Dordrecht; Boston: Kluwer, 2001.
Oaklander, L. N. *The Ontology of Time*. Buffalo, NY: Prometheus Books, 2004.
Oaklander, L. N. *Temporal Relations and Temporal Becoming*. Lanham, MD, 1984.
Oaklander, L. N. "Two Versions of the New B-Theory of Language." In *Time, Tense, and Reference*, edited by Quentin Smith and A. Jokic. Cambridge, MA: MIT Press, 2003.
Oaklander, L. N., and Q. Smith, eds. *The New Theory of Time*. New Haven, CT: Yale University Press, 1994.
Paul, L. A. "Truth Conditions of Tensed Sentence Types." *Synthese* 111 (1997): 53–71.
Percival, Phillip. "Presentism." In *Time, Reality, and Experience*, edited by Craig Callender. Cambridge, U.K.: Cambridge University Press, 2002.
Priest, G. "Mellor on Tense." *Analysis* 44 (1986).
Priest, G. "Tense, *Tense*, and TENSE." *Analysis* 46 (1987): 184–187.
Prior, A. N. *Time and Modality*. Oxford. Oxford University Press, 1957.
Prior, A.N. "The Notion of the Present." *Studium Generale* 23 (1979): 245–258.
Prior, A. N. *Papers on Time and Tense*. Oxford: Oxford University Press, 1968.
Prior, A. N. *Past, Present, and Future*. Oxford: Oxford University Press, 1967.
Reichenbach, Hans. *Elements of Symbolic Logic*. New York: Oxford University Press, 1947.
Russell, Bertrand. "Review of Hugh MaColl's Symbolic Logic and its Application." *Mind* 15 (1906): 256–257.
Schlesinger, George. *Aspects of Time*. Indianapolis, IN: Hackett, 1981.
Sciama, Dennis. "The Universe as a Whole." In *The Physcist's Conception of Nature*, edited by J. Mehra. Dordrecht: Reidel, 1973.
Smart, J. J. C. (1971). "Causal Theories of Time." In *Basic Issues in the Philosophy of Time*, edited by E. Freeman and W. Sellars. La Salle: Open Court, 1971.
Smart, J. J. C. *Philosophy and Scientific Realism*. London: Routledge and Kegan Paul, 1963.
Smart, J. J. C. "The Reality of the Future." *Philosophia* 10 (1981) 141–150.
Smart, J. J. C. "The River of Time." In *Essays in Conceptual Analysis*, edited by A. Flew. London: Routledge and Kegan Paul, 1966.
Smart, J. J. C. "Time and Becoming." In *Time and Cause*, edited by P. van Inwagen. Dordrecht, Netherlands: Reidel, 1980.
Sider, Ted. *Four Dimensionalism*. New York: Oxford University Press, 2001.
Smith, Q. "Actuality and Actuality as of a Time." In *The Importance of Time*, edited by L.N. Oaklander. Dordrecht; Boston: Kluwer, 2001.
Smith, Q. *Language and Time*. New York: Oxford University Press, 1993.
Smith, Q. "The Mind-Independence of Temporal Becoming." *Philosophical Studies* (1985).
Smith, Q. "The Multiple Uses of Indexicals." *Synthese* (1989).
Smith, Q. "Problems with the New Tenseless Theory of Time." *Philosophical Studies* (1987).
Smith, Q. "Reference to the Past and Future" In *Time, Tense, and Reference*, edited by Q. Smith and A. Jokic. Cambridge, MA: MIT Press, 2003.
Smith, Q. "Time and Degrees of Existence." In *Time, Reality, and Experience*, edited by Craig Callender. Cambridge, U.K.: Cambridge University Press, 2002.
Smith, Q., and L. N. Oaklander. *Time, Change, and Freedom*. New York: Routledge, 1994.
Tooley, Michael. *Time, Tense, and Causation*. Oxford: Oxford University Press, 1997.
Tooley, Michael. "Response to the Comments on *Time, Tense, and Causation* by Storrs McCall, Nathan Oaklander, and Quentin Smith." In *The Importance of Time*, edited by L.N. Oaklander. Dordrecht; Boston: Kluwer, 2001.
Tooley, Michael. "Basic Tensed Sentences." In *Time, Tense, and Reference*. New York: Routledge, 2003.
Williams, Clifford. "The Date-Analysis of Tensed Sentences." *Australian Journal of Philosophy* 70 (1992): 198–203.
Zimmerman, Dean. "Temporary Intrinsics and Presentism." In *Metaphysics: The Big Questions*, edited by Peter van Inwagen. Malden, MA: Blackwell, 1998.

Quentin Smith (2005)

TIME, CONSCIOUSNESS OF

William James's discussion of the perception of time in *Principles of Psychology* (Vol. I, Ch. 15) provides a con-

venient starting point for a discussion of the "consciousness of time." James's main concern was to give an empiricist account of our temporal concepts. This is clear from the Lockean question with which he started: "What is the *original* of our experience of pastness, from whence we get the meaning of the term?" (p. 605) and from his answer that the "prototype of all conceived times is the specious present, the short duration of which we are immediately and incessantly sensible" (p. 631). A contemporary empiricist might formulate James's thesis thus: that all other temporal concepts can be defined in terms of the relation "earlier than" and that this relation is sense given or can be ostensively defined so that even if a person does not use the term *specious present*, he is obliged to say that some earlier events are still, in some sense, present to us when we are sensing a later event.

Consider why James used the term *specious present* in describing such facts. He quoted with approval a passage by E. R. Clay, who invented this term; the quotation shows that they both assumed that the philosophically correct use of "present" is to refer to the boundary, conceived of as a durationless instant, between past and future. They pictured time as a line of which the specious present is a segment whose later boundary is the real present and hence concluded that the specious present and its contents are really past. James used two phrases that suggest that the specious present also includes a bit of the future; one, when he said that it has "a vaguely vanishing backward and forward fringe" (p. 613) and, two, when he said that it is "a saddle-back from which we look in two directions into time" (p. 609). This view is implied by nothing else he said, so we shall ignore the paradoxes it would needlessly generate and concentrate on what James said frequently: that we are continuously directly perceiving or intuiting a past duration and its contents.

James illustrated the concept of the specious present by citing experiments carried out by Wilhelm Wundt and his pupil Dietze designed to measure the duration of the longest group of sounds that a person can correctly identify without counting its members. According to Wundt, this duration is 6 seconds; according to Dietze, it is 12 seconds. James equated this period (6 to 12 seconds) with the duration of the specious present (and failed to add the qualification "for hearing"). The ability that Wundt and Dietze were investigating is a familiar one. Hearing a series of sounds as a melody or as a sentence involves recognizing them as forming a temporal pattern, or Gestalt. Another familiar experience is sometimes cited in this context: The chiming of a clock may not be noticed until it has stopped, yet we can still attend to the sounds and, one is inclined to say, inspect them; we can notice facts about them—for example, that there are five or ten chimes. Since James applied the concept of the specious present by reference to such auditory experiences, he was committed to saying that a sound that audibly terminated 5 or 10 seconds ago is still being directly perceived. Now, this seems inconsistent. "I am now directly perceiving (or sensing) *X*" seems to imply "*X* is now present and exists simultaneously with my perceiving (sensing) it."

This criticism was made by H. J. Paton (*In Defence of Reason*, pp. 105–107) against the account of the specious present given by Bertrand Russell and C. D. Broad. Russell and Broad had, however, applied the concept of the specious present differently from James. They appealed to the fact that we see things moving, that we see the second hand of a watch moving in a way that we cannot see the hour hand moving. They took this to imply that we simultaneously sense the second hand (or, rather, the corresponding sensa) occupying a series of adjacent positions. To this Paton replied, "If in a moment I can sense several different positions of the second-hand, then these different positions would be sensed as being all at the same moment.… What I should sense would be not a movement, but a stationary fan covering a certain area and perhaps getting gradually brighter towards one end.… You can't see a sensum that isn't there. If you see it, it is there at the time you see it." Paton concluded that awareness of the positions of the second hand prior to the present instant must be ascribed to memory. Paton, however, overlooked a fact about vision. What he failed to find when he looked at the second hand is found when we look at things that move (traverse a given optical angle) more quickly. If, in the dark, you watch someone rotating a lamp at the appropriate speed, you see a moving ring of light or if, in daylight, you hold a bright object—for instance, a watch—and move it fairly quickly across your visual field while gazing at a point in the middle of its path (place 1), you can still, momentarily, see a streak in place 1 when the watch is seen, out of the corner of your eye, to have halted at place 2. Such facts provide a second way of applying the concept of the specious present.

Our philosophical problem is to analyze and describe the experiences in question in a way that avoids contradictions and which, if we are empiricists, is consistent with saying that temporal relations are given in experience. We shall examine several alternative accounts of the relevant facts but first note that the account one finds appropriate will depend on one's philosophical standpoint, especially concerning the nature of the mind and of perception. Obviously, it makes a difference whether

one conceives of the self as, for example, an immaterial substance that transcends time or as a physical organism, whether one holds a realist or a representative theory of perception. Paton assumed, as did Russell and Broad, that what we see are sensa, conceived of as entities numerically distinct from physical objects, and Paton asserted that sensa can exist only at the moment at which they are sensed. Whether this dictum need be accepted will be discussed later.

Our problem is also phenomenological. The specious present doctrine dissolves into a platitude unless we draw a distinction between what is "sensed" (or "immediately experienced" or "directly perceived") and what is "perceived" (or "perceptually accepted, recognized, or judged"). No one doubts that we perceive things changing, that it is correct to speak of "seeing" a thing move, and so on. The phenomenological question is whether, in such cases, the very recent positions or states of things are still being sensed. In posing the problem in this way, we are not committed to a representative theory of perception or to a sensum terminology. As we are using "to sense" and kindred verbs to say that we perceive more than we sense—that we see an orange as juicy and solid when all that we sense is its front surface—does not entail that the things we sense are numerically distinct from the things we perceive—the orange.

ATTEMPTED SOLUTIONS

TIME AS THE FOURTH DIMENSION. A simple solution seems to be open to anyone who accepts the thesis that the physical world is a four-dimensional manifold. If, accordingly, we (learn to) think of physical objects as four-dimensional solids in describing which tenseless verbs must be used, it is a corollary that what is visually sensed is not an instantaneous cross section of the four-dimensional manifold, but a short slice thereof, about one-tenth of a second long in the time dimension. Suppose you see a meteor flash across the sky. If you hold a realist theory of perception, you would say that what you sense is a short slice of the history of the four-dimensional meteor. If you identify conscious states with brain processes, you would say that what you sense is a short slice of certain of your four-dimensional brain cells. And in these sentences "short slice of the history of" would be used literally, since you are presumably following mathematicians such as Hermann Minkowski in treating time as if it were another spatial dimension, which is "at right angles to each of the other three" (whatever this may mean apart from indicating what sort of diagrams to draw).

This account would satisfy the empiricist insofar as it implies that temporal intervals and relations are sense given in the same sense as that in which spatial intervals and relations are sense given. This account, however, does not seem viable. If the physical world were a four-dimensional manifold, it would be logically impossible for its contents—four-dimensional solids—to move or otherwise change unless they did so in a time that is distinct from the one which has been spatialized (and such motion would not concern us since we do not observe motions of four-dimensional solids). The four-dimensional conceptual scheme would permit no use for the basic concepts in terms of which we do (and must?) interpret our experience—notably, our concept of a physical thing as a three-dimensional entity that can move and change, our concept of a physical event as a change in one or more such physical things, and our concept of physical causation as a relation between such physical events.

Now, it is a ground-floor empirical fact that we observe things moving and changing. Anyone who adopts the four-dimensional world theory is therefore obliged to tell us what it is that moves or changes. Since he is treating the physical world as changeless, the only answer he can give is that it is our states of consciousness that change as we become successively aware of adjacent cross sections of the four-dimensional world. But this makes sense only if we, the observers, are not in space-time (and one would still have to acknowledge a [real] time dimension other than the one that has been spatialized, in which our states of consciousness are successive). Our first account of the specious present could be accepted by a dualist if he could show that it is possible to dispense with our concepts of physical things, events, and causes. We may well doubt whether he can do this, for even the physicists cannot formulate many of their questions without using our conceptual scheme.

AUGUSTINE AND BROAD. James followed Clay in assuming that the philosophically correct use of *present* is to refer to a durationless instant. We christen this "the punctiform present (PP) assumption." Anyone who makes this assumption is committed to saying that apart from its later boundary the specious present is really past, and he is thereby disposed to say (1) that the contents of the specious present consist of images or "representations" of what has just been sensed and (2) that what these images represent is known only by memory. Here we have a second way of describing the relevant experiences.

This way of thinking is found in Augustine's classical discussion of time (*Confessions*, Book XI, Secs. 10–28). Augustine claimed that no one would deny that the present has no duration, and surprisingly, until recently no one has. Augustine combined the PP assumption with another that he deemed self-evident—that everything which is past or future does not (now) exist. He proceeded logically to the conclusions that when a person perceives or measures time, what he is attending to is "something which remains fixed in his memory" and therefore that time is not "something objective" (Sec. 27). He ended by, in effect, defining "past" in terms of human memories and "future" in terms of human expectations (Sec. 28). (These conclusions suited Augustine, for his purpose in discussing time was to show that it is meaningless to ask what God was doing before he made heaven and earth; see Secs. 10–13, 30.)

Idealists may be happy to accept Augustine's conclusion that time is unreal (subjective), but many philosophers and psychologists who do not accept this conclusion have found themselves in a quandary as a result of taking for granted Augustine's premises. Their quandary is that however one applies the concept of the specious present, if its contents are described as sensa or images, the sensa or images which a person has at any durationless instant are present at (that is, simultaneous with) that instant, but then whatever relations may hold between such sensa or images, temporal precedence cannot be among them, for this relation holds between things that are not simultaneous. One is then driven to say that awareness of the nontemporal features of one's sensa or images somehow stimulates one to construct ideas of temporal relations that are not sense given. James quoted several psychologists who got into this quandary, but he showed no sign of recognizing its (for him) unacceptable implications—that it obliges one either to deny the objective reality of time or to appeal to an intuition or a priori knowledge of time.

The paradoxical implications of Augustine's premises are clearly exhibited in Broad's account of time in his *Examination of McTaggart's Philosophy*. Broad here abandoned the account of the specious present he had given in *Scientific Thought*, where he had spoken of an event's being present throughout a finite process of sensing. He now asserted that it is only "instantaneous event-particles" which are "present in the strict sense," and he spoke of events (event-particles) becoming (coming into existence) and passing away (ceasing to exist). He was thus committed to the strange metaphysical theory according to which each event-particle is created and annihilated at "successive" instants, and the answer to the question "What exists at present?" would have to be "A set of simultaneous event-particles," though during the time it takes you to utter this phrase, an infinite number of such sets would have been born and died.

Why has the PP assumption been treated as self-evident by so many eminent thinkers? No one has claimed that the correct (strict) use of "here" is to refer to a Euclidean point; why have so many philosophers assumed that the correct (strict) use of "now" or "present" is to refer to a durationless instant? That it rejects, by implication, the PP assumption is a merit of the now popular token-reflexive analysis of sentences containing "now" or "present" or a verb in the present tense. In this analysis "now" is rendered "simultaneous with this utterance," and uttering a sentence takes a second or two. But this analysis is open to two objections: (1) that when one says "It is (now) raining," one is not referring to one's own utterance and (2) that when one refers to "the present war," the duration of the war does not coincide with one's utterance.

To remedy these objections, we need to jettison the traditional oversimplified assumption that the only temporal relations are earlier than, simultaneous with, and later than (the only relations that could hold between durationless instants); we need to recognize the numerous perceptible temporal relations between durations or processes (for example, sounds), the relations that are formally analogous to those that can hold between two segments of varying lengths belonging to the same straight line (coincidence, adjacence, partial and complete overlapping). We may then say "It is (now) raining" equals "The falling of rain (here) overlaps temporally with this" where "this" refers to the duration of the speaker's so-called specious present.

AN EMPIRICIST SOLUTION. The first solution we considered could be accepted only by a dualist who holds that minds are not in space-time (and René Descartes's problems concerning the connection between mind and body would become much more acute, since one's body is being conceived of as a four-dimensional solid). The second solution we considered is consistent only with either a form of idealism that denies the objective reality of time or a form of rationalism which treats our knowledge of time as a priori. If we reject the premises used by Augustine and many others, we can find a solution that is consistent with empiricism and with the views that time order is an objective feature of the world and that we, whatever else we may be, are physical creatures. Consider

first the proposition that what is past or future cannot (now) exist. We may reply that "existence" should be predicated, in any tense, only of things (continuants), not of events, which happen or occur, and not of processes, which go on. Admittedly, past or future events are not now happening, and past or future processes are not now going on, but, of course, many of the things, including people, which existed at past times and which will exist at future times exist now.

We must also reject the PP assumption and may define "present" as the duration of the speaker's specious present. But can we, for this purpose, employ either or both of the methods of interpreting "the specious present"? James's method would make the specious present 6 to 12 seconds long; Russell's would make it about one-tenth of a second, so we can scarcely combine these interpretations. In Wundt's experiments, cited by James, the subjects were attending to sounds that had audibly terminated, though they were still presented in the sense that the subject could still "hear" them. If we say that a sound that has audibly terminated is still present, this would be inconsistent, for "it *has* audibly terminated" implies "it is past." We ought surely to describe the duration of the specious present, as interpreted by James, as "the span of immediate memory for hearing," and to call this a *specious* present is appropriate.

Does a similar objection arise if we define "present" as the duration of what is visibly sensed, when, for example, we see a meteor? Can we describe this experience by saying that we simultaneously sense the meteor occupying a series of different places throughout a fraction of a second? Those who accept the PP assumption will say, "No. When the meteor has visibly reached place 2, it is no longer in place 1, where it was one-tenth of a second earlier, and we cannot sense a thing occupying a place in which it no longer is; thus, the fading sensation of the meteor must be ascribed to (immediate) memory." But why the "must"? In discussing such phenomenological problems, for which ordinary language was not designed, it is not decisive to appeal to the "correct" (normal) use of language, but note that "remember" is not used in the way prescribed by our critic. In our earlier example, moving a watch across one's field of vision, we should say that the streak at place 1 is seen, not that it is merely remembered.

The experiences we have in seeing such movements can be described by saying that visual sensations linger and very rapidly fade. (This fact rarely obtrudes on us because we follow a moving object in which we are interested by head or eye movements and do not attend to the resultant blurring of background objects.) But are we obliged to describe the facts by saying that a moving object can be simultaneously seen (sensed) in a series of different positions? We are obliged to do this if we adopt a realist theory of perception. Consider the case of the moving watch. The realist holds that what is sensed is a surface of the watch, and as we conceive such a physical object, it cannot occupy different regions of space at the same time; thus, the realist must describe this experience by saying that, for a very short time, a person still senses (very indistinctly) the watch at place 1 when it has visibly reached place 2. But this argument is not sufficient if one adopts a representative theory of perception, or phenomenalism. For then one may, apparently, say that what one senses is a contemporary instantaneous streaky sensum at place 1.

But can one consistently say this? To say this involves conceiving a sensum as an entity that exists only at a durationless instant. This generates paradox since one will have to say that we falsely believe that we see something moving and that this belief is somehow generated by our sensing a compact series of instantaneous and stationary sensa the later members of which differ in their spatial relations from the earlier; one will also be unable to give an empiricist account of how we come by the notions earlier and later. To try to get out of this quandary, the user of the sensum language may amend his account and say that what we sense is the contemporary instantaneous state of a sensum; then he is conceiving of a sensum as a continuant (albeit a short-lived one)—that is, as something which endures and can change. Those who use sensum language usually do talk of sensa moving and changing.

Since sensa may be and often are conceived of as short-lived continuants, the user of the sensum language is free to drop the PP assumption. The latter implies that the phenomenological objects (images or sensa) which a person has or is aware of at any durationless instant, must be present at—that is, simultaneous with—that instant, and this implies that temporal precedence cannot be sense given. If, however, a sensum is conceived of as a continuant, we may say that the same sensum is present throughout a short period, that successive states or positions of the sensum are present at a given instant, and that a person can still sense a visual sensum where it was one-tenth of a second ago. Paton's statement "You can't sense a sensum that isn't there. If you see it, it is there at the time you see it" was intended to refute the possibility that one can simultaneously sense a sensum occupying a series of adjacent positions, but such dicta cannot be treated as synthetic a priori propositions. Philosophers

make the rules of the sensum language as they go along, and there seem to be no clear and accepted rules for translating "visual sensations linger and fade" into this language. If we use this language, we are free to adopt rules that allow empiricists to say what they need to say—that is, that temporal relations between different sensa and different states of the same sensum are sense given.

Few philosophers would now accept Immanuel Kant's view that time (conceived of as an infinite continuum) is an intuited datum or his view that our knowledge of time is a priori (*Critique of Pure Reason*, "Transcendental Aesthetic," II, Sec. 4). Most modern philosophers would agree with James that time is a notion that we construct from temporal relations which are sense given. Such philosophers must surely accept the thesis that temporal relations are sense given within the present and that this duration of which we are in James's words "incessantly sensible" ought to be called "the *conscious* present." Clay and James called this duration "the specious [that is, pseudo] present" because they assumed that only its later boundary should be called "the real present."

FINAL CONSIDERATIONS. The besetting sin of philosophers, scientists, and, indeed, all who reflect about time is describing it as if it were a dimension of space. It is difficult to resist the temptation to do this because our temporal language is riddled with spatial metaphors. This is because temporal relations are formally analogous to spatial relations—for example, the formal resemblance between the overlapping of two sticks and the overlapping of two sounds disposes us to forget that in the latter case "overlapping" is used metaphorically. If we picture the passing of time in terms of movement along a line, we are led to ask "What moves?" and are disposed to answer, like Edmund Husserl, "Events keep moving into the past" and to forget that "move" is now being used metaphorically, that events cannot literally move or change. As J. J. C. Smart asserted, things change, events happen ("The River of Time," *Mind* 58 [1949]: 483–494). Those who spatialize time, conceiving of it as an order in which events occupy different places, are hypostatizing events. The temptation to hypostatize events is presumably the result, at least in part, of the linguistic fact that the terms, which can be said to stand in temporal relations like simultaneous with and earlier than, are event expressions. Those who ponder about time are forever using event expressions as their main nouns, and they frequently seem to forget what events are—changes in three-dimensional things. What we perceive and sense are things changing. Time is a nonspatial order in which things change.

This conclusion is deflationary. Poets, mystics, and metaphysicians naturally prefer more exciting ways of talking about time. It is ironical that although Henri Bergson forcibly criticized the spatialization of time, he based his metaphysical theories largely upon describing time in spatial images and metaphors. Bergson argued that our spatialized concept of time is an intellectual construct which misleadingly represents real concrete time (*durée*), which is grasped by, and belongs only to, inner consciousness (*Time and Free Will*). In describing *durée*, however, he said things that are difficult to reconcile and, in some cases, to interpret at all. *Durée* is said to *flow* (p. 221), yet its different moments are said to *permeate* one another (pp. 110 and 133) and to be *inside* one another (p. 232). Bergson did not recognize that these are as much spatial metaphors as is describing time as linear. It was his own metaphors and his implicit use of the PP assumption that led Bergson to his paradoxical conclusions—for example, that "duration and succession belong not to the external world, but [only] to the conscious mind" (p. 120). We cannot prevent metaphysicians who are so inclined from trying to reduce things to events or processes or to expand things into four-dimensional solids, but such intellectual acrobatics are unnecessary, apart from the paradoxes that they generate. Our consciousness of time's "flow" is our consciousness of things changing.

See also Augustine, St.; Bergson, Henri; Broad, Charlie Dunbar; Consciousness; Consciousness in Phenomenology; Gestalt Theory; Husserl, Edmund; James, William; Kant, Immanuel; McTaggart, John McTaggart Ellis; Russell, Bertrand Arthur William; Smart, John Jamieson Carswell; Space; Wundt, Wilhelm.

Bibliography

The interest of twentieth-century philosophers in time stemmed largely from the writings of Henri Bergson, who held that understanding the nature of time is the key to the main problems of philosophy. His first important book, *Essai sur les données immédiates de la conscience* (Paris: Alcan, 1889), was translated by F. L. Pogson as *Time and Free Will* (New York: Macmillan, 1910). This contains what purports to be a phenomenological description of time consciousness, but from the start Bergson's language is permeated with idealist metaphysics. Edmund Husserl discussed problems concerning awareness of time in his *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins* (Halle, 1928), which has been translated by J. S. Churchill as *The Phenomenology of Internal Time-Consciousness* (Bloomington: Indiana University Press, 1964). In *An*

Outline of Philosophy (London: Allen and Unwin, 1927), pp. 204–205, and *The Analysis of Mind* (London: Macmillan, 1921), pp. 174–175, Bertrand Russell presented, very briefly, the kind of solution argued for above, but he did not acknowledge any of the difficulties that others have found in this concept. C. D. Broad has made two detailed attempts to analyze the concept of the specious present, in *Scientific Thought* (London: Kegan Paul, 1923), pp. 346–358, and *Examination of McTaggart's Philosophy*, Vol. II (Cambridge, U.K.: Cambridge University Press, 1938), Ch. 35. He used similar diagrams in each book, but what these are said to symbolize differs greatly in each. His earlier account can be criticized for its use of the concept of momentary acts of sensing, but this could have been remedied. In his later account he ended by describing the specious present doctrine as a verbal trick for trying to reconcile contradictory propositions. It looks as if Broad was converted by the sort of criticism made by H. J. Paton in his paper "Self-Identity," *Mind* 38 (1929): 312–329, later reprinted in his *In Defence of Reason* (London and New York: Hutchinson, 1951). J. D. Mabbott criticized his own odd interpretation of the specious present doctrine in "Our Direct Experience of Time," *Mind* 60 (1951): 153–167. C. W. K. Mundle challenged Mabbott's interpretation and discussed several alternatives in "How Specious Is the 'Specious Present'?," *Mind* 63 (1954): 26–48, and later critically examined three different accounts of time contained in Broad's writings in "Broad's Views about Time," in *The Philosophy of C. D. Broad*, edited by P. A. Schilpp (New York: Tudor, 1959). The thesis criticized above, that the physical world should be conceived as a four-dimensional manifold, is argued in J. J. C. Smart's *Philosophy and Scientific Realism* (New York: Humanities Press, 1963).

OTHER RECOMMENDED TITLES

Butterfield, Jeremy. "Seeing the Present." *Mind* 93 (1984): 161–176.

Campbell, John. *Past, Space, and Self*. Cambridge, MA: MIT Press, 1994.

Hestevold, H. Scott. "Passage and the Presence of Experience." *Philosophy and Phenomenological Research* 50 (1990): 537–552.

Hirsh, I. J., and J. E. Sherrick. "Perceived Order in Different Sense Modalities." *Journal of Experimental Psychology* 62 (1961): 423–432.

Le Poidevin, Robin. "The Experience and Perception of Time." In *The Stanford Encyclopedia of Philosophy*. Available from http://www.plato.stanford.edu.

Le Poidevin, Robin, and M. MacBeath. *The Philosophy of Time*. Oxford: Oxford University Press, 1993.

McInerney, P. *Time and Experience*. Philadelphia: Temple University Press, 1991.

Mellor, H. *Real Time II*. London: Routledge, 1998.

Newton-Smith, W. *The Structure of Time*. London: Routledge and Kegan Paul, 1980.

Oaklander, L. Nathan. "On the Experience of Tenseless Time." *Journal of Philosophical Research* 18 (1993): 159–166.

Oaklander, L. Nathan, and Quentin Smith, eds. *The New Theory of Time*. New Haven, CT: Yale University Press, 1994.

Prior, A. N. *Past, Present and Future*. Oxford: Oxford University Press, 1967.

Reichenbach, H. *The Philosophy of Space and Time*. Translated by M. Reichenbach. New York: Dover, 1957.

Russell, Bertrand. "On the Experience of Time." *Monist* 25 (1915): 212–233.

Schlesinger, G. *Aspects of Time*. Indianapolis: Hackett, 1980.

Shoemaker, S. "Time without Change." *Journal of Philosophy* 66 (1969): 363–381.

C. W. K. Mundle (1967)
Bibliography updated by Benjamin Fiedor (2005)

TIME, DIRECTION OF

See *Causal Approaches to the Direction of Time; Physics and the Direction of Time*

TIME IN CONTINENTAL PHILOSOPHY

The attempt to conceive time, time's relation to human experience, and the makeup of the universe is perhaps the central problem of twentieth-century Continental philosophy. Time emerged as a central problem in late nineteenth century German philosophy where temporality became increasingly identified with consciousness and mind. Franz Brentano's work provided an impetus for Edmund Husserl's analyses of *internal time-consciousness*, and Wilhelm Dilthey and Husserl were both influential for Martin Heidegger's fundamental ontology. In France, before these phenomenological approaches had been worked out, Henri Bergson reconceived time in a way that anticipated them and profoundly influenced later French thought.

In general, Bergson calls on metaphysics (that is, Platonism and its latest version in Kant) to embrace the reality of movement, change, becoming, and time. The originality of this thinking consists in differentiating between abstract representations of time and the immediate givenness of *pure duration* in consciousness. In *Time and Free Will* (1910), he distinguishes duration from time understood as a homogeneous medium in which moments are represented as juxtaposed to one another like points on a line. His concern is that this representation of time confuses duration with spatial extension, generating metaphysical problems involving motion (see Zeno's paradoxes) and free will. In duration, Bergson says, moments are not mutually external but interpenetrating (multiplicity); states of consciousness are not separate and distinct but combined and continuous (unity); and actions are not the realization of preexisting possibil-

ities but the fruit of the self's organic evolution through time. Later in the century, Gilles Deleuze will appropriate the Bergsonian concept of heterogeneous and yet continuous multiplicity in his own considerations of time (see below).

In *Matter and Memory* (1991), Bergson's greatest book, he defines duration as the unconscious conservation of memories, which progressively insert themselves into hesitations in the stimulus–response circuits of living bodies. Bergson thus conceives the past as surviving independent of perceived or recollected images, that is, independent of presence. The connection of duration to the past and to anticipated actions transforms duration into the *vital impetus* (*élan vital*), which Bergson presents in *Creative Evolution* (1998). Here he offers an alternative to views of evolution that reduce time to the mechanical realization of preexisting possibilities. Such views treat life as a closed system in which "all is given" (p. 37). The notion that all possibilities are already given renders time meaningless.

After psychology and evolutionary biology, Bergson brings his conception of time to bear on physics. In *Duration and Simultaneity* (1999), he aims to show how duration can resolve the paradoxes surrounding Einstein's special theory of relativity. The concepts of simultaneity and succession presuppose a consciousness in which events are contemporaneous or follow one another. Bergson argues that physicists are incorrect to conclude that a plurality of times exists. Different times assigned to different systems of reference are indeed measurable, but they have no duration other than that of the physicist performing the calculations and therefore no reality. Not surprisingly, Bergson's views have been the center of controversy, and they remain indicative of profound differences between philosophical and scientific ways of conceiving time.

Like Bergson, Husserl originally devoted his attention to describing time as it is given to consciousness, investigating how things and events are represented as continuing over time. How, for instance, is a melody given as a unified object even though its beginning *runs off* into the past before its end arrives? Husserl's response to this question can be found in his lectures *Concerning the Phenomenology of Internal Time Consciousness* (1905). During the period of these lectures, Husserl was developing his phenomenological method of reduction. The objective time of things or events in the world must be suspended or reduced, that is, made relative to consciousness, which, for Husserl, is defined by intentionality. Intentionality turns out to be fundamentally time-consciousness. The appearance of temporal objects (i.e., things identical over time) is analyzed into the contents and the acts of consciousness (the subjective correlates of the contents).

Husserl adopts Brentano's idea that an objective unity in time requires acts of presentation that join its preceding phases with its current phase, for example, the notes of a melody that are sinking away into the past with the note that is heard now. Past notes must be not only retained but also modified so that they are connected to those that follow without being jumbled together. However, Husserl rejects Brentano's claim that the contents of perception, which represents only what is given in the present, are supplemented by imagination, which reproduces those contents with the stamp of having passed. He contends that the consciousness of a note as having just passed is essentially different from recollection or memory, which would rely on an image. In other words, Husserl distinguishes between *retention*, an impressional consciousness that holds on to what was given in perception as it sinks away into the past, and *secondary memory*, a representational consciousness that makes present again what had already run off into the past. He argues that perception of a temporal object, whether enduring unchanged or changing successively, implies different modes of apprehension of the same contents, and retention accounts for the interplay of sameness and difference.

Later, Jacques Derrida will argue that this interplay of sameness and difference blurs Husserl's essential distinction between retention and representation (see below). For Husserl, however, the interplay of sameness and difference also occurs in relation to the future. Like retention, *protention*, the anticipation of what is immediately to come, is a form of impressional (or nonrepresentational) consciousness. Retention and protention constitute the temporal horizon of what is no longer present and what is not yet present for any *primal impression*. These modes of impressional consciousness constitute the temporality of immanent temporal objects. Consciousness of these objects is oriented by a *now-point*, but Husserl maintains that this point is an ideal limit and that the phases of time-consciousness comprise a *living present*.

What Husserl calls the living present implies another and more fundamental level of consciousness: the *absolute flow of time-constituting consciousness*. With regard to a unity constituted in time, we are aware of the threefold temporal intentional dimensions of the object in retention, primal impression, and protention. There is

not only the unity of an object through its appearances across time as one and the same object—vertical intentionality—but also the unity of consciousness across the differences in objects that appear for consciousness—transverse intentionality. According to Husserl, the ultimate constituting flow, in which these unities are constituted at once in a double-intentionality, is not itself constituted in time. For this reason, it is difficult to speak of the ultimate ground of temporality as either in time or outside of it, and Husserl refers to it as *quasi*-temporal.

Heidegger's standard criticism is that Husserl, despite the radicality of his descriptions of time-consciousness, never posed the question of the being of consciousness. Therefore, in *Being and Time* (1962), Heidegger reopens "the question about the meaning of being" (p. 2), which has been forgotten since the time of Plato and Aristotle, and approaches time as "the horizon for the understanding of being" (p. 39). To gain access to this horizon, following Husserl, Heidegger engages in a phenomenological analysis of the modes of temporality underlying *existence* (*Dasein*, a term that indicates not only human existence but also being itself). He shows in the first division of *Being and Time* that *Dasein* consists in a structure of care, which intertwines being *ahead of itself*, being *already in the world*, and being *alongside things*. Although anticipated by others, Heidegger's innovation is to show how the past and the future, not the present, define time.

Heidegger begins the second division of *Being and Time* with an analysis of death and finitude and attempts to show how temporality is *the ontological meaning of care*. Because death is *my* death, it makes me break free of inauthentic (group) existence where I do not take responsibility for my possibilities of existence. In contrast, authentic *being-toward-death* is a mode of existence called *anticipatory resoluteness* in which I freely take up my possibilities, opening the horizon of authentic temporality. By *repeating* the existential analysis, Heidegger grounds *Dasein's* ontological structure in temporality. He shows how the originary unity of the structure of care is grounded in the temporal *ecstases* of the future, having-been, and the present. He then distinguishes between the authentic and inauthentic modes of these *ecstases*, contrasting the everyday phenomena of awaiting, *making-present*, and forgetfulness, with the authentic modes of anticipation, the moment (*Augenblick*), and repetition. He also gives a temporal interpretation of structures introduced in the first division—understanding, affectedness (*Befindlichkeit*), falling, and discourse—explicating the temporal conditions for the disclosedness of *Dasein* as being-in-the-world. The temporal interpretation opens the way for a consideration of *Dasein's* historical character.

By means of determining the existential foundation of historical research and historical truth—appropriating Dilthey's idea of hermeneutics—Heidegger shows that our *reckoning* of historical or natural events that occur *in time* is derived from primordial temporality. This derivative character of something being in time leads him to account for phenomena of intratemporality through the temporal structures of *Dasein's* concern with the world, always directed toward a for-the-sake-of-which, that makes measuring time possible. The ordinary understanding of time as an infinite, irreversible sequence of nows originates, Heidegger says, from the *ecstatic-horizonal unity of temporality*.

Heidegger continues the project of *Being and Time* in subsequent lecture courses, including *The Basic Problems of Phenomenology*, which includes a deconstruction of Aristotle's theory of time and an account of how time as it is ordinarily understood presupposes originary temporality. While in *Being and Time* he focuses on the ecstatic character of temporality, the basis of *Dasein's* existence as a *thrown projection*, in *Basic Problems*, he turns his attention to its horizonal schema, or the enclosure of the ecstatic opening. Heidegger focuses especially on the present and its horizon, which he calls *praesens*, to show that Kant understands being on the basis of presence. (A deconstruction of Kant's ontology appears in Heidegger's second book, *Kant and the Problem of Metaphysics* [1929].) For Heidegger, since the ancient Greeks, being has been defined as *ousia*, which he interprets as *constant presence*. Consequently, the relation between being and time has traditionally been understood on the basis of one ecstasis: the present. For Heidegger, a *temporal ontology* is the necessary corrective for this privilege of the present.

In the early 1930s, Heidegger appeals to a notion of the event (*Ereignis*) as a new way to conceive how being comes into presence without recourse to the self-projection of *Dasein*. In this period, Heidegger begins thinking of time in terms of the play of space–time (*Zeitraum*). Much later, he reformulates his approach to temporality in the lecture *On Time and Being* (2002) in which he considers time as the unity of three dimensions of givenness, whose interplay constitutes yet a fourth dimension, which he calls *nearness*. Although Heidegger's thought turns away from *Dasein*, from the human being, toward *Ereignis*, the event of appropriation, the *inner co-belonging* of being and time, remains a fundamental question

for him. Indeed, the event of appropriation, for Heidegger, is the event of thinking, which is a kind of memory.

Both Jean-Paul Sartre and Maurice Merleau-Ponty carry Heidegger's project of a phenomenological ontology forward, making temporality integral to their major works. Sartre's *Being and Nothingness* (1993) revolves around the fundamental ontological difference between being-for-itself (Sartre speaks of both consciousness and *Dasein*) and being-in-itself (brute objects). For Sartre, all other accounts of subjectivity (for instance, that of Bergson) have confused the for-itself with the in-itself. Human beings have *no* determining essence; they are *nothing* and therefore they are radically free. Temporality comes into play in this dialectic of being and nothingness because freedom is future oriented. Beginning with the concrete phenomena of *my* particular past, present, and future, Sartre works toward an account of their general form and their unity. He argues that temporality is a structure of being-for-itself that implies separation and synthesis, multiplicity and unity, of the different temporal phases. He dubs this "profound cohesion and dispersion" (p. 195) of temporality a *diasporatic* mode of being-for-itself.

Nevertheless, for Merleau-Ponty, Sartre's idea of a radical voluntarism requires the emphasis of dispersion and separation over cohesion and synthesis. So, in the *Phenomenology of Perception* (1962), Merleau-Ponty develops a phenomenological ontology of time without recourse to Sartre's categories of being-for-itself and being-in-itself. Merleau-Ponty rejects both the early Bergsonian characterization of time as immediately given to consciousness and the Husserlian view that consciousness constitutes time. In order to show how time originates in a synthesis without ever being completely deployed, he directs attention to the "*field of presence* as the primary experience in which time and its dimensions make their appearance" (p. 416). In the primordial field of presence, he says, time is a single thrust, a "bursting forth or dehiscence," and, in Heidegger's words, an *ek-stase*. For Merleau-Ponty, time has a *sense*, which gives it an abiding character (without sense ever being eternal like a Platonic idea).

Merleau-Ponty's concept of sense negotiates the transition from passivity to spontaneity. In opposition to Sartre, therefore, Merleau-Ponty maintains that temporality does not confirm absolute freedom (pure spontaneity) but only the possibilities of commitment and refusal afforded by the historical and corporeal situation. Later, in a sometime bitter debate with Sartre, Merleau-Ponty argues in *Adventures of the Dialectic* (1973) that politics and temporal ontology are interwoven in a way that Sartre misses. He worries that Sartre's early ontology implies that a choice takes place in the instant by fiat, or else it has always already taken place. For Merleau-Ponty, choices, and especially political choices, must repeat a sense given in the past and open a sense continuing into the future.

Despite the dominance of Sartre and Merleau-Ponty's existentialism, Emmanuel Levinas's thought eventually comes to be recognized as providing an important approach to time. Against Bergson's duration, Levinas stresses the instant, an event that comes from the future and is always other than what I have experienced. In *Time and the Other* (1987), he describes this alterity with regard to death, also challenging Heidegger's existential analysis. For Levinas, death is defined not by nothingness but by mystery since it cannot be grasped. Whereas Heidegger allows for a mastery of death and the future in anticipation, Levinas thinks that they are absolutely other. Unlike Heidegger's *Augenblick*, the instant disrupts the solitude and virility of the subject for Levinas, so that time is a relationship with the radically other. In this way, Levinas's discourse of the other moves from ontology to ethics, and in later works, especially *Totality and Infinity* (1969), he continues to consider the ethical significance of time. Like Levinas, Derrida is inspired by the phenomenological approach to time. In *Speech and Phenomena* (1973), Derrida *deconstructs* Husserl's phenomenology of language in the *Logical Investigations* (1901) by means of Husserl's own descriptions of internal time-consciousness. What is at issue is the momentary (and therefore temporal) self-understanding of meaning in an internal dialogue. According to Derrida, with the distinction between expression and indication, Husserl maintains that in an internal dialogue, I understand the meaning of my own expression in the very moment when I speak; there is no mediation of the linguistic phoneme, and no difference between me as speaker and me as hearer, only immediate presence to myself.

Yet, in his early lectures on time-consciousness, Husserl speaks of retention being a *nonperception*. If it is nonperception (without which there could be no living present), retention could not be a pure presence and would have to involve some sort of absence, difference, and mediation. Retention is thus, as Derrida says, *a trace*. This is not a return Brentano's view that imagination lends the experience of time to perception. Rather, Derrida means that the genetic source of the difference between imagination and perception lies in the difference

between retentional trace (repetition in the most general sense) and primal impression. The trace implies a kind of spatial distance within my internal dialogue, as if I were speaking not to the one who is closest to me (myself) but to someone else, someone past, someone distant, someone other. Derrida elaborates on the relationship between time and language in "*Ousia* and *Grammē*" (1982), challenging Heidegger's distinction between primordial and derivative temporality and showing how Heidegger's own thought remains oriented by the value of presence.

Finally, Deleuze offers a variety of approaches to time, also influenced by Husserl and Heidegger but especially by Bergson. In *Bergsonism* (1991), he focuses especially on Bergson's concept of duration, defining it as a *qualitative multiplicity* in which there is continuity and heterogeneity. For Deleuze, continuity does not eliminate difference but, rather, makes it be internal (in contrast to Levinas's and Derrida's emphasis on exteriority). Deleuze pushes Bergson's thought further in *Difference and Repetition* (1994), where he discusses three syntheses of time: habit, memory, and the empty form of time. Here he provides his account of the living present, the past in general, and the future as absolutely new (with regard to Friedrich Nietzsche's eternal return). In *The Logic of Sense* (1990), Deleuze opposes thinking of time in terms of the present through the distinction between Chronos and Aion. While Chronos signifies the time of a present that comprehends or mixes together the past and the future, Aion divides the present into the past and the future. As an *instant without thickness*, dividing time in two directions at once, Aion signifies a continuous and heterogeneous multiplicity. Deleuze identifies Aion with the pure, empty form of time that has "unwound its own circle, stretching itself out in a straight line" (p. 165). Later, Deleuze offers commentaries on Bergsonian duration in *Cinema 1: The Movement-Image* (1983) and *Cinema 2: The Time-Image* (1985), and he describes how modern directors achieve a *direct presentation of time*.

In twentieth century Continental philosophy, there have been several major shifts. Bergson challenges thinking of time in terms of space, Husserl describes the quasi-temporal origin of time, and Heidegger calls into question the privilege of presence. Subsequently, Sartre and Merleau-Ponty recognize the need to come to terms with the relation between temporality and sense. The nonpresence of the instant and the trace orient Levinas's and Derrida's thinking, and Deleuze also displaces the time of the present. On the horizon of these philosophies of difference emerging in the 1960s, we find in Michel Foucault and others a renewed concern with *place* that rivals an alleged *temporocentrism* of mainstream Continental philosophy, and it remains to be seen whether time will continue to be a central problem.

See also Aristotle; Bergson, Henri; Brentano, Franz; Deleuze, Gilles; Derrida, Jacques; Dilthey, Wilhelm; Einstein, Albert; Ethics; Foucault, Michel; Heidegger, Martin; Husserl, Edmund; Infinity in Mathematics and Logic; Kant, Immanuel; Levinas, Emmanuel; Merleau-Ponty, Maurice; Nietzsche, Friedrich; Platonism and the Platonic Tradition; Sartre, Jean-Paul.

Bibliography

Ansell Pearson, Keith. *Philosophy and the Adventure of the Virtual: Bergson and the Time of Life*. New York: Routledge, 2002.

Bergson, Henri. *Creative Evolution*. Translated by Arthur Mitchell. Mineola, NY: Dover Publications, 1998 [1907].

Bergson, Henri. *Duration and Simultaneity: Bergson and the Einsteinian Universe*. Translated by Mark Lewis and Robin Durie; edited by Robin Durie. Manchester, U.K.: Clinamen Press, 1999 [1922].

Bergson, Henri. *Matter and Memory*. Translated by N. M. Paul and W. S. Palmer. New York: Zone Books, 1991 [1896].

Bergson, Henri. *Time and Free Will: An Essay on the Immediate Data of Consciousness*. Translated by F. L. Pogson. London: Dover Publications, 1910 [1889].

Brough, John B., and Lester Embree, eds. *The Many Faces of Time*. Dordrecht, MA: Kluwer Academic, 2000.

Carr, David. *Time, Narrative, and History*. Bloomington: Indiana University Press, 1986.

Casey, Edward. *Remembering: A Phenomenological Study*. Bloomington: Indiana University Press, 1987.

Chanter, Tina. *Time, Death, and the Feminine: Reading Levinas with Heidegger*. Stanford, CA: Stanford University Press, 2001.

Dastur, Françoisie. *Heidegger and the Question of Time*. Translated by François Raffoul and David Pettigrew. Atlantic Highlands, NJ: Humanities Press, 1998 [1990].

De Boer, Karin. *Thinking in the Light of Time: Heidegger's Encounter with Hegel*. Albany: State University of New York Press, 2000.

Deleuze, Gilles. *Bergsonism*. Translated by Hugh Tomlinson and Barbara Habberjam. New York: Zone Books, 1991 [1966].

Deleuze, Gilles. *Cinema 1: The Movement-Image*. Translated by H. Tomlinson and B. Habberjam. Minneapolis, MN: Zone Books, 1986 [1983].

Deleuze, Gilles. *Cinema 2: The Time-Image*. Translated by H. Tomlinson and R. Galeta. Minneapolis, MN: Zone Books, 1989 [1985].

Deleuze, Gilles. *Difference and Repetition*. Translated by Paul Patton. New York: Columbia University Press, 1994 [1968].

Deleuze, Gilles. *Logic of Sense*. Translated by Mark Lester and Charles Stivale. New York: Colombia University Press, 1990 [1969].

Derrida, Jacques. "*Ousia* and *Grammē*: Note on a Note from Being and Time." In *Margins of Philosophy*. Translated by

Alan Bass. Chicago: University of Chicago Press, 1982 [1972].

Derrida, Jacques. *Speech and Phenomena, and Other Essays on Husserl's Theory of Signs*. Translated by David Allison. Evanston, IL: Northwestern University Press, 1973 [1967].

Durie, Robin, ed. *Time and the Instant: Essays in the Physics and Philosophy of Time*. Manchester, U.K.: Clinamen Press, 2000.

Gallagher, Shaun. *The Inordinance of Time*. Evanston, IL: Northwestern University Press, 1998.

Heidegger, Martin. *Being and Time*. Translated by J. Macquarrie and E. Robinson. (Especially sections 65, 69, 78–81), New York: Harper, 1962 [1927].

Heidegger, Martin. *Contributions to Philosophy: From Enowning*. Translated by Parvis Emad and Kenneth Maly. Bloomington: Indiana University Press, 1999 [1989].

Heidegger, Martin. *History of the Concept of Time: Prolegomena*. Translated by Theodore Kisiel. (Sections 1–3, 18, 32–36.), Bloomington: Indiana University Press, 1985 [1979].

Heidegger, Martin. *On Time and Being*. Translated by Joan Stambaugh. Chicago: University of Chicago Press, 2002 [1969].

Heidegger, Martin. *The Basic Problems of Phenomenology*. Translated by Albert Hofstadter. Bloomington: Indiana University Press, 1982 [1975]. (Especially sections 19–22).

Heidegger, Martin. *The Concept of Time*. Translated by W. McNeill. Oxford: Blackwell Publishers, 1992 [1989].

Heidegger, Martin. "The Concept of Time in the Science of History" (1915). Translated by H. S. Taylor, H. W. Uffelmann, and J. Van Buren. In *Supplements*, edited by John Van Buren, 49–60. Albany: State University of New York Press, 2002.

Heidegger, Martin. *Kant and the Problem of Metaphysics*. Translated by Richard Taft, Bloomington: Indiana University Press, 1997 [1929].

Husserl, Edmund. *Concerning the Phenomenology of Internal Time Consciousness, Husserliana*. Vol. 10. Translated by John B. Brough; edited by Rudolf Boehm. Dordrecht: Kluwer Academic, 1990 [1928].

Husserl, Edmund. *Die 'Bernauer Manuscripte' über das Zeitbewußtseins (1917/1918), Husserliana*. Vol. 33, edited by R. Bernet and D. Lohmar. Dordrecht: Kluwer Academic, 2001.

Husserl, Edmund. *Logical Investigations*. Vols. 1 and 2. Translated by J. N. Findlay; edited by Dermot Moran. London: Routledge, 2001 [1900/01].

Kisiel, Theodore *The Genesis of Heidegger's Being and Time*. Part III. Berkeley: University of California Press, 1993.

Levinas, Emmanuel. *Time and the Other*. Translated by Richard Cohen. Pittsburgh: Duquesne University Press, 1987 [1947].

Levinas, Emmanuel. *Totality and Infinity*. Translated by Alphonso Lingis. Pittsburgh: Duquesne University Press, 1969 [1961].

McCumber, John *Time in the Ditch: American Philosophy and the McCarthy Era*. Evanston, IL: Northwestern University Press, 2001.

Merleau-Ponty, Maurice. *Adventures of the Dialectic*. Translated by Joseph Bien. Evanston, IL: Northwestern University Press, 1973 [1955].

Merleau-Ponty, Maurice. *Phenomenology of Perception*. Translated by Colin Smith. London: Routledge, 1962 [1945].

Negri, Antonio. *Time for Revolution*. Translated by Matteo Mandarini. London: Continuum International, 2004.

Pöggeler, Otto. "Pathways of 'Time.'" *The Paths of Heidegger's Life and Thought*. Translated by J. Bailiff. Amherst, NY: Humanities Press, 1998.

Protevi, John. *Time and Exteriority*. Lewisburg, PA: Bucknell University Press, 1995.

Ricoeur, Paul. *Time and Narrative*. Vols. 1–3. Translated by Kathleen McLaughlin and David Pellauer. Chicago: University of Chicago Press, 1990 [1983].

Sallis, John. "Time Out…" In *Echoes: After Heidegger*, 44–75. Bloomington: Indiana University Press, 1990.

Sartre, Jean-Paul. *Being and Nothingness*. Translated by Hazel E. Barnes. New York: Washington Square Press, 1993 [1943].

Von Herrmann, Friedrich-Wilhelm. "*Being and Time* and *The Basic Problems of Phenomenology*". Translated by Parvis Emad and Kenneth Maly. In *Reading Heidegger: Commemorations*, edited by John Sallis, 118–135. Bloomington: Indiana University Press, 1993.

Wood, David. "Reiterating the Temporal: Toward a Rethinking of Heidegger on Time." In *Reading Heidegger: Commemorations*, edited by J. Sallis, 136–159. Bloomington: Indiana University Press, 1993.

Wood, David. *The Deconstruction of Time*. Evanston, IL: Northwestern University Press, 2001 [1989].

Heath Massey and Leonard Lawlor (2005)

TIME IN PHYSICS

No one conception of time emerges from a study of physics. One's understanding of physical time changes as science itself changes, either through the development of new theories or through new interpretations of a theory. Each of these changes and resulting theories of time has been the subject of philosophical scrutiny, so there are many philosophical controversies internal to particular physical theories. For instance, the move to special relativity gave rise to debates about the nature of simultaneity within the theory itself, such as whether simultaneity is conventional. Nevertheless, there are some philosophical puzzles that appear at every stage of the development of physics. Perhaps most generally, there is the perennial question, Is there a "gap" between the conception of time as found in physics and the conception of time as found in philosophy?

One can understand all of these changes and controversies as debates over what properties should be attributed to time. The history of the concept of time in physics can then be understood as the history of addition and subtraction of these properties, and the philosophical controversies thus understood as debates about particular

additions and subtractions. Just as one may take a set of numbers and impose structure on this set to form the real number line, one may also take the set of moments or events (which will be used interchangeably) and impose various types of structure on this set. Each property attributed to time corresponds to the imposition of a kind of structure upon this set of events, making sense of different claims about time. Let us begin with a bare set of events and successively add structure to this set. In particular, it helps to differentiate *ordering* properties, *topological* properties, and *metrical* properties of time.

ORDER

It seems clear that different times are ordered to some extent. Intuitively, one can give a set an order by making sense of what times are between what other times. The time the cake baked is between the time of mixing the ingredients and the time of eating the cake; eating the cake is between the baking and the feeling full, and so on. One can therefore impose an ordering on this set of events by adding a ternary "between-ness" relation of the form: "x is between y and z" defined for some or all moments in the set. If betweenness is defined for some but not all distinct triples of moments, then it can be said that one has a *partially ordered* set; if betweenness is defined for every triple of the set, then it can be said that one has a *totally ordered* set. Newtonian physics, as will be shown, totally orders classes of simultaneous events. Relativistic physics, by contrast, will only partially order the set of all events.

Between-ness as defined above is not always sufficiently powerful to order topologically nontrivial sets. To see this, consider a circle with four members of the set on it: "1" at twelve o'clock, "2" at three o'clock, "3" at six o'clock, and "4" at nine o'clock. Because the set is closed, 2 is between 1 and 3, between 3 and 4, and between 1 and 4. Consequently, the between-ness relation is blind to the difference between this layout and the same but with "3" at three o'clock and "2" at six o'clock. For such sets more machinery is needed to order the set.

An ordering does not disclose much about the set of moments, $\{t_1, t_2, t_3\ldots\}$. It does not imply whether t_2 is as far from t_1 as from t_3. Nor does it imply a direction, whether times goes from t_1 to t_3 or t_3 to t_1. Although the baking example suggests a natural direction to the set of times, an ordering is strictly independent of a direction. Nor does the ordering specify the dimensionality of the set or most other properties one normally attributes to time. The next level of structure, topology, will help make sense of some of these attributions to time.

TOPOLOGY

Topological properties are those that are invariant under "smooth" transformations. Technically, these transformations are one-to-one and bicontinuous; and what they leave invariant is the so-called neighborhood structure that is given by picking out a family of open subsets closed under the operations of union and finite intersection. Intuitively, the transformations that leave this structure unchanged correspond to operations such as stretching or shrinking, as opposed to operations such as ripping and gluing. A coffee cup and a doughnut are, topologically speaking, the same shape; if made out of an infinitely pliable rubber, one could be smoothly transformed into the other. Being closed like a circle, having an edge, and being one-dimensional are examples of topological properties. No amount of stretching and shrinking can (for instance) make the circle into a line, make an edge disappear, or make a one-dimensional set two-dimensional.

Many issues in the philosophy of time are in fact questions about the topology of time: is time closed or open? discrete? branching? two-dimensional? oriented (directed)? Formally, the answers to these questions are determined by the topological structure of time.

METRIC

Once topological structure is added to the set of times, most temporal properties are determined. However, there is still a major one remaining: duration. Of the set $\{t_1, t_2, t_3\ldots\}$ it is still not known whether t_2 is as far from t_1 as it is from t_3—even after all topological properties are specified. The temporal distance between two moments is not a topological invariant, for it can be smoothly stretched or shrunk. To capture the idea of temporal distance, a *metric* must be put on the topological structure. The temporal metric is a function that gives one a number, the temporal distance or duration, between any pair of times. (In relativity what is imposed instead is a *spacetime* metric; see below.)

In principle, an infinite number of possible metrics are mathematically possible. One might choose a metric that makes the duration between 1980 and 1990 twice the duration between 1990 and 2000. However, such a choice would make a mess of almost all of science. It would entail, for instance, that the earth went twice as fast around the sun in the 1990s as it did in the 1980s. One would then have to adjust the rest of physics so as to be compatible with this result. As Hans Reichenbach stresses, there are simpler and more complex choices of temporal metric.

TIME IN CLASSICAL PHYSICS

Time in classical physics is normally assumed to have the ordering, topological, and metrical structure of the real number line. That is, it is one-dimensional, continuous, infinite in both directions, and so on. The temporal metric is just the one used for the real line: between any two times, a and b, the duration is b–a. Time in classical physics does have a number of remarkable properties, of which three will be mentioned here. The first two concern the metrical properties of time, whereas the third is more a property of the dynamics than of time itself.

First, the metric of time is independent of the metric of space. This feature implies that the amount of time between any two events is path-independent: if persons A and B leave an event e_1 and then meet at a later event e_2, the amount of time that has elapsed for A is equal to the amount of time that has elapsed for B. The distinct spatial distances traveled by A and B are irrelevant to how much time has passed between e_1 and e_2.

Second, simultaneity is absolute. Before explaining "absolute," consider the "simultaneous with" relation. For any event e, there is a whole class of events that are simultaneous with e. Indeed, the "simultaneous with" relation is an equivalence relation in classical physics. Equivalence relations are reflexive, symmetric, and transitive; for this example, what is important is that they partition a set into disjoint subsets. Hence the "simultaneous with" relation partitions the set of all events into proper subsets, all of whose members are simultaneous with one another. It is these classes of simultaneous events, rather than the events themselves, that are totally ordered. What is interesting about this partition in classical physics is that it is unique. Classical physics states that every observer, no matter their state of motion, in principle agrees on whether any two events are simultaneous. This observation translates into only one partition (or foliation) being the right one. In this sense simultaneity is absolute—it does not depend on one's frame of reference but is an observer-independent fact of the Newtonian world.

Third, classical physics is time reversal invariant. Consider a sequence of particle positions over time, (x_1,t_1), (x_2,t_2), $(x_3,t_3)\ldots(x_n,t_n)$. The fundamental classical laws of evolution are such that if this sequence is a solution of the laws, then so is the time-reversed sequence $(x_n,t_n)\ldots(x_3,t_3)$, (x_2,t_2), (x_1,t_1). The classical laws are invariant under the transformation of –t for t. This is true also of arbitrarily large multi-particle systems and even of classical fields. If a bull entering a china shop and subsequently breaking vases is a lawful history, then so is a bunch of scattered vase shards spontaneously jumping from the ground and forming perfect vases while a bull backs out of a china shop.

TIME IN SPECIAL RELATIVITY

In classical physics, material processes take place on a background arena of space and time, described above. The move from classical physics to special relativity is usually taken as a change in the background arena from classical space and time to the "spacetime" of Hermann Minkowski. This new entity, spacetime, is fundamental, and space and time only exist in a derivative fashion. On this conception, there is not one metric for time and another for space; rather, there is one spacetime metric supplying spatiotemporal distances between four-dimensional events. These spacetime distances are invariant properties of the spacetime. Time can be decoupled from space only in an observer-dependent way; each distinct possible inertial observer (one who feels no forces) carves up spacetime into space and time in a different way. In a sense, there is no such thing as time in Minkowski spacetime, if by "time" one conceives of something fundamental.

There are, however, two "times" in Minkowski spacetime that correspond to different aspects of classical time, namely, "coordinate" time and "proper" time. Let us take coordinate time first. Think of an arrow in three-dimensional Euclidean space. One can decompose this arrow relative to an arbitrary basis {x,y,z} by measuring how far the arrow extends in the x-direction, how far in the y-direction, and how far in the z-direction, where x, y, and z are perpendicular, and the arrow's base lies at the origin. The same arrow would decompose differently in a different basis {x',y',z'}. As one can decompose a vector in Euclidean space along indefinitely many different bases, so too can one decompose a four-dimensional spacetime vector along many different bases in Minkowski spacetime. Mathematically, coordinate time in special relativity is just one component of an invariant spacetime four-vector, just as y' is one component of a Euclidean spatial vector. In the Euclidean case, the value of the arrow along the first component of the decomposition varies with basis; so too in spacetime, the value of the first component—here, coordinate time—varies with frame of reference.

The second bit of residue of the classical time is the so-called proper time. The proper time is a kind of parameter associated with individual trajectories in spacetime. It is often thought of as a kind of clock tied to an object through its motion. This time is a scalar—that is, just a number—and as such is an invariant of the

spacetime. All observers will agree on the value of proper time for A as he travels from e_1 to e_2; all will agree on the value of proper time for B as she travels from e_1 to e_2; and all will agree that these values will not be the same if they take different paths. Unlike with classical time, the temporal distance in Minkowski space is not independent of spatial distance. The amount of time between any two events is path-dependent: if persons A and B leave an event e_1 and then meet at a later event e_2, the amount of time that has elapsed for A is in general not equal to the amount of time that has elapsed for B. Spatial distances can only be completely disentangled from temporal distance in a given inertial frame of reference.

Time in classical physics plays the role of coordinate time and the role of proper time. A little reflection reveals that it can accomplish this task because in classical physics the amount of time between any two events is path-independent.

Three consequences of the shift to special relativity ought to be highlighted. First, simultaneity is not absolute in Minkowski spacetime. Simultaneity is a temporal feature, yet the temporal does not disentangle from the spatial except within an inertial reference frame. What events are simultaneous with one another is observer-dependent. Given spacelike-related events e_1 and e_2, inertial observer A may (rightly) say they are simultaneous whereas inertial observer B, traveling at a constant velocity with respect to A, may (rightly) say e_1 is earlier than e_2. In Minkowski spacetime, they do not disagree over any observer-independent fact of the matter. In terms of the earlier discussion, it can then be said that the "simultaneous with" relation partitions Minkowski spacetime, but only within a frame of reference.

Second, the temporal ordering in Minkowski spacetime is partial, not total. The only temporal ordering that all observers agree on is the ordering among "timelike" events. Timelike related events are those that are in principle connectible by any particle going slower than the speed of light in a vacuum. Think of all the events that can be reached from any given event that way. Consider the event of your elementary school graduation (e_1) and the event of your high school graduation (e_2). Obviously sub-luminal particles could make it from one to the other; for instance, you are a set of such particles. Due to the finite speed of light, however, there are many events that such particles could not reach—for example, whatever was going on at Alpha Centuri simultaneous with (in your reference frame) e_2. What happened on Alpha Centuri simultaneous with e_2 is not an observer-independent fact. But that e_2 follows e_1 is an observer-independent fact. Only the timelike related events are invariantly ordered.

Third, and perhaps most famously, in a sense time passes more slowly for a moving observer than for one at rest. Consider two inertial observers, A and B, traveling at a constant velocity relative to one another, and let a clock be at rest in A's frame. Looking at the ticks of the clock, the special relativistic metric entails that B will conclude that the clock in A's frame is running slow. This effect, known as time dilation, is entirely symmetrical: A would find a clock at rest in B's frame to be running slow, too. Time dilation has many experimentally confirmed predictions, such as that atomic clocks on planes tick slowly relative to clocks on land and that mesons have longer lifetimes than they should from the earth's frame of reference.

TIME IN GENERAL RELATIVITY

General relativity, unlike special relativity, treats the phenomenon of gravitation. It famously does away with Newton's gravitational force, understanding gravitational phenomena as instead a manifestation of spacetime curvature. Loosely put, the idea is that matter curves spacetime and spacetime curvature explains the gravitational aspects of matter in motion. Hence the largest conceptual difference between special and general relativity is that Minkowski spacetime is flat whereas general relativistic spacetimes may be curved in an indefinite number of ways. Otherwise, as regards time, again there is a division between coordinate time and proper time, no privileged foliation of spacetime, only a partial temporal ordering, and the possibility of time dilation.

In terms of the previous division, curvature is a metrical property, so the primary difference between special and general relativity is that the former's metric is merely one of the many possible metrics allowed by the latter. General relativity places various constraints between the spacetime metric, or geometry, and the distribution of matter-energy. Thinking of these constraints as the laws of general relativity, general relativity claims a variety of spacetime geometries are physically possible. Because these different metrics allow and sometimes demand different topologies and even orderings, time may have dramatically different ordering, topological, and metrical properties depending on the spacetime model. Some consequences of this fact are especially worthy of note.

First, there are spacetimes without a single global moment. In special relativity, simultaneity was observer-dependent. Minkowski spacetime could be carved up, or foliated, into a succession of three-dimensional spaces

evolving along a one-dimensional time an indefinite number of ways—a distinct foliation for every possible inertial observer. Though this may also be the case in general relativity, there are spacetime models that prohibit even one foliation of spacetime into space and time. The famous Gödel spacetime, named after the great logician Kurt Gödel, is an example of such a spacetime. Due to the effects of curvature, in such spacetimes it is impossible to find even a single global always-spatial three-dimensional surface. There is no global moment of time in such spacetimes. There is no way to conceive of world history, in such a spacetime, as the successive marching of three-dimensional surfaces through time.

Second, perhaps most famously, general relativity has models that permit interesting time travel. In these models a traveler can start off at event e, and by traveling always to the local future (that is, into e's future lightcone), eventually come back to events that are to e's past (that is, in e's past lightcone). Indeed, these models will allow one to travel back to an earlier event: an observer's worldline may intersect e, and then after some proper time has elapsed, intersect e again. These "causal loops" are called closed timelike curves. Of the many models that allow time travel, the Gödel model is again remarkable for it allows the time traveler the fullest menu of possibilities: in the model, it is possible (given enough time and energy) to get from any event e_1 to any other event e_2 on the entire spacetime, including the case where $e_1=e_2$.

Third, whether time is infinite or finite can be an observer-dependent fact. When discussing Minkowski spacetime it was noted that there are different ways to decompose spacetime into space and time; alternatively, there are generally many ways to foliate a spacetime. When nontrivial topologies are considered, there are spacetimes consistent with general relativity that make whether time is infinite or finite a foliation-dependent matter. That is, there are foliations of one and the same spacetime that make time finite and foliations that make time infinite. In spacetimes admitting two such foliations, the age-old question of whether time is finite or infinite would be answered with a convention. In such a world there is no coordinate-independent fact of the matter regarding how long time persists. The universe might last an infinite amount of time according to one coordinization, or language, and a finite amount of time according to another coordinization, or language.

TIME IN FUTURE PHYSICAL THEORIES

As mentioned, because physical theories are always changing, there is no one conception of time emerging from a study of physics. On the horizon of research are the various programs of "quantum gravity," the would-be theory that unifies or at least makes consistent our best theory of matter, quantum field theory, and the best theory of spacetime, general relativity. Though speculative, virtually all of these programs are entertaining dramatic changes for the conception of spacetime, ranging from the idea that spacetime is discrete to the idea that time is an emergent property arising from some more fundamental stuff.

PHILOSOPHICAL CONTROVERSIES

There are many philosophical problems concerning time in physics. Philosophers have discussed the physical possibility of time travel in general relativity, the possibility of discrete time, the nature of time reversal invariance, the possibility of backward causation in physics, such as in the Wheeler-Feynman time-symmetric version of electromagnetism, the possibility of time emerging from something more fundamental in quantum gravity, and more. In addition, it will not be surprising that many topics typically dealt with in the context of space also have temporal counterparts. The absolute-versus-relational debate, famously discussed by Gottfried Leibniz and Samuel Clarke and more than a hundred authors thereafter, is often discussed in the classical context of space; but those arguments apply equally well to the case of time, and in the modern version of the debate, to spacetime. And the many deliberations surrounding the conventionality of the metric apply just as well to the temporal metric as the spatial metric (and of course the spacetime metric). Here the discussion focuses on whether physical time captures all the fundamental properties of time and the so-called problem of the direction of time.

TENSE

In the famous terminology of J. E. McTaggart, the temporal relations of earlier than, later than, and simultaneous with are called "B-properties" and the monadic properties of past-ness, present-ness, and futurity are called "A-properties." Those who argue that the B-properties are the fundamental features of time are dubbed advocates of the "tenseless" theory of time; those who argue that instead the A-properties are fundamental are dubbed advocates of the "tensed" theory of time. Much of the work in philosophy of time, especially throughout the twentieth century, can be described as a debate between tensers and detensers.

Because the categories "tensed" and "tenseless" are broad umbrellas covering many different doctrines, it is probably best not to think of this as one debate. A better way to frame the debate is to conceive it on the model of the debate between mind-body dualists and materialists. Dualists find the description of the mind by the natural sciences to be either incomplete or simply wrong. Various features of mental states—for example, consciousness—are said to be either left out or indescribable by these natural sciences. Materialists counter either by denying the reality of these features or by explaining why the natural sciences do manage to explain such features.

One can conceive the debate regarding time in the same mold. Though the features attributed to time vary with physical theory, some philosophers feel that physical theory has consistently missed out on one or more essential properties of time. Physical theory orders some or all of the events in time, just as the relations of right and left order events in space. In classical (relativistic) physics, for any (some) pair of events, e_1, e_2, physical theory states whether e_1 is earlier, later, or simultaneous with e_2. The theories use relational temporal properties and not monadic ones. One can of course say e_1 is to the past of e_2, but that is just to say that e_1 is earlier than e_2. Physical theory seems to require only tenseless temporal relations. Broadly speaking, the debate is between those who would add some metaphysical feature to time as it is found in science and those who would not. Various arguments are adduced to show that such features are needed or not needed, compatible with science or incompatible, and so on. Consider now three features often felt to be left out by physical time.

THE PRESENT. Physical theory does not identify which time is Now. That is, it judges which events are earlier, later, and simultaneous with which other events, but it fails to mention which among all sets of events are the present ones. Some philosophers argue, based on experience, analysis of ordinary language, or study of puzzles surrounding change, that physical theory misses out on a genuine property of time, Now-ness. Others reply that the idea of a metaphysically special present is wrong-headed. Linguistic features of the now are explained via the properties of indexicals in general. Because one would not reify the here, one should not reify the now. Attempts are then made to show that the language, thought, and behavior attributing objectivity to the present can be explained by facts about human beings and their typical physical environments.

FLOW OR BECOMING. Physical theory also does not describe a property corresponding to the flow of time or to a process of becoming. Again, the different events are ordered, have a certain distance from one another, and so on, but there does not seem to be anything that flows (such as the Now). Nor is there a distinction made among events, such that it makes sense to talk about the Now turning an unreal future real. Again, some philosophers argue, based on experience or the study of various puzzles, that there is genuine becoming in the world. C. D. Broad, for example, proposed a model wherein the past and present are real and the future successively becomes present and hence real.

TIME'S ARROW. If physical time is time-reversal invariant, then nowhere does it distinguish one direction of time. But there are many asymmetric processes: physical ones, such as the radiation and thermodynamic asymmetries; metaphysical ones, such as the asymmetry of causation and of counterfactual dependence; epistemological ones, such as that one typically knows more about the past than the future; and emotional ones, such as that people usually care more about the future than the past. To explain one or more of these asymmetries, some philosophers have posited a directionality to physical time. Others answer that that the physical asymmetries do not themselves need explanation and that they in turn can explain the other asymmetries. To mention one possible sequence of moves, one might try to show that the thermodynamic and radiative temporal asymmetries explain the memory asymmetry (people have memories of the past, not the future), the memory asymmetry explains the knowledge asymmetry, and the knowledge asymmetry explains the psychological asymmetry.

There are also two famous conceptual arguments against the idea that time itself flows (and depending on the model of becoming, against becoming). One, McTaggart's Paradox, claims that the idea of time flowing leads to a logical contradiction. Essential to the idea that time flows, says McTaggart, is the idea that events change their A-properties: for instance, the event of Socrates's death was future, then present, and then past. So every event has all three monadic properties. But this is in straightforward conflict with the claim if an event is future it is not past. McTaggart and his supporters claim that any way of discharging the contradiction by insisting that events are not at the same time past, present, and future leads to infinite regress.

Another argument, by the philosophers C. D. Broad and J. J. C. Smart, begins by noting that change is always

the change of some property with respect to time. Movement, for example, is having different locations at different times. So if time flows—if, say, the Present moves—then Broad and Smart suggest that it must be that the Present moves with respect to time. But this time, Smart claims, must be a hyper-time; and if this hyper-time is a kind of time, it must flow with respect to a hyper-hyper-time, and so on. There are too many responses to this argument to consider them all here.

It should not be surprising that considerations from physics enter these debates.

SPECIAL RELATIVITY AND TENSE

Some also argue that a metaphysically distinguished present is inconsistent with special relativity. The reason is obvious: since simultaneity is relative, how can a monadic feature of events such as presentness be frame-dependent? In Minkowski spacetime, there will be cases where for observer O_1, e_1 is present and e_2 is later, whereas for observer O_2, e_2 is present and e_1 is later. Assuming presentness is not frame-dependent, there appears to be a contradiction. This argument, originally made by Hilary Putnam and C. W. Rietdijk, also would affect positions claiming time flows, if the flowing is done by a unique present. Even if correct, by itself this argument does not tell how to arrange the conflict into premises and conclusion. Does relativity disprove the present or does the present disprove relativity? Naturalistically inclined philosophers are loath to consider the latter reading; but strictly speaking, if there were enough prior reason to believe in a privileged present, then alternatives to Minkowski spacetime would need to be considered—such as embedding relativistic phenomena in classical space and time in the manner H. A. Lorentz favored.

GENERAL RELATIVITY AND TENSE

From the perspective of general relativity, the attack on tenses from special relativity seems rather limited. Minkowski spacetime may locally be a good approximation to whatever the true global spacetime is, but strictly speaking special relativity is only valid on planes that are tangent to mere points of the general relativistic geometry. There appears no particular reason to think that general relativity's impact on the tenses debate will mirror special relativity's impact.

As mentioned, general relativity takes from special relativity a division between coordinate time and proper time and only a partial temporal ordering. The question is whether it banishes a privileged foliation of spacetime into space and time. The answer depends on the particular spacetime model and what one means by "privileged." In some models, ones with realistic distributions of matter and energy, one can define a global cosmic time. Cosmic time is defined with respect to the mean motion of matter. The possibility exists of a tenser using cosmic time, which mimics some features of classical time, as the time of becoming, passage, and so on. Challenges to this use include the fact that cosmic time can only be defined in some subset of the solutions to Einstein's field equations, and questions of arbitrariness in the choice of a cosmic time function.

With the possibility of cosmic time in mind, Kurt Gödel argued that general relativity, far from rescuing tenses, in fact showed that time is "ideal," or not fundamental. Reflecting on the odd eponymous spacetime mentioned above, Gödel states that it is obvious that time does not flow in the spacetime he discovered. But that means, Gödel says, that time does not flow in the spacetime of the actual world either. Why? In brief, his idea is that time flow should not be contingent, yet because Gödel spacetime enjoys the same laws of nature as does the actual world, it differs from this world only in the contingent distribution of matter and energy. Indeed, Gödel goes so far as to presume time's flow is essential to time, and hence concludes that Gödel spacetime shows that there is no such thing as time in this world.

THE PROBLEM OF THE DIRECTION OF TIME

So far this entry has described issues concerning time in fundamental or near-fundamental physics. There also exists a philosophical problem arising from an apparent conflict between the way microphysics seems to treat time and the way macroscopic physics treats time. While microphysics may be time reversal invariant, the physics describing macroscopic behavior such as the warming or cooling of bodies to room temperature, the expansion of gases, and so on, is not time reversal invariant. Consider the volume of an initially localized sample of a light gas released in the corner of a room. As time goes on, it will spread through its available volume: (v_1, t_1) (v_2, t_2) (v_3, t_3)…, where $v_3>v_2v_1$ and $t_3>t_2>t_1$, and so on. While classical mechanics implies that the opposite shrinking process from v_3 to v_1 is lawful, thermodynamics states that it is not.

The science of statistical mechanics seems to reconcile the two by introducing probabilistic considerations: the process from v_3 to v_1 is possible, says statistical mechanics, but highly unlikely, whereas the process from v_1 to v_3 is highly likely. However, statistical mechanics

itself is time reversal invariant. It manages to state that evolution from v_3 to v_1 is unlikely and v_1 to v_3 likely. Looked at more closely, however, it implies that *given* v_1, v_3 is more likely in either time direction. In other words, it rightly states that v_3 is a likely state to evolve to, but it also implies that it is a likely state to have evolved from. The second implication is obviously wrong. This problem and related ones occupied many of the founders of statistical physics, including Ludwig Stephan Boltzmann, J. C. Maxwell, Joseph Loschmidt, and Ernest Zermelo. Solutions to the problem seem to require inserting a temporal asymmetry somewhere in the physics, either by assuming temporally asymmetric boundary conditions or by introducing new laws of nature.

See also Philosophy of Physics; Relativity Theory.

Bibliography

Albert, David. *Time and Chance*. Cambridge, MA: Harvard University Press, 2000.

Arntzenius, Frank, and Tim Maudlin. "Time Travel and Modern Physics." In *The Stanford Encyclopedia of Philosophy (Spring 2002 Edition)*, edited by Edward N. Zalta. Available from http://plato.stanford.edu/. Reprinted in *Time, Reality and Experience*, edited by Craig Callender. Cambridge, U.K.: Cambridge University Press, 2002.

Barbour, Julian. *The End of Time*. London: Weidenfeld & Nicholson, 1999.

Broad, C. D. *Examination of McTaggart's Philosophy*, vol. 2, part 1. Cambridge, U.K.: Cambridge University Press, 1983.

Butterfield, Jeremy, ed. *The Arguments of Time*. Oxford: The British Academy, 1999.

Callender, Craig, ed. *Time, Reality and Experience*. Cambridge, U.K.: Cambridge University Press, 2002.

Callender, Craig "Thermodynamic Asymmetry in Time." In *The Stanford Encyclopedia of Philosophy (Spring 2002 Edition)*, edited by Edward N. Zalta. Available from http://plato.stanford.edu/.

Craig, William L. *Time and the Metaphysics of Relativity*. Dordrecht, Netherlands: Kluwer, 2001.

Dainton, Barry. *Time and Space*. Ithaca, NY: McGill-Queen's University Press, 2001.

Dorato, Mauro. *Time and Reality: Spacetime Physics and the Objectivity of Temporal Becoming*. Bologna, Italy: CLUEB, 1995.

Earman, John. "An Attempt to Add a Little Direction to 'The Problem of the Direction of Time.'" *Philosophy of Science* 41 (1974): 15–47.

Earman, John. *Bangs, Crunches, Whimpers and Shrieks*. Oxford: Oxford University Press, 1995.

Einstein, Albert. "On the Electrodynamics of Moving Bodies," as reprinted and translated in *The Principle of Relativity*, 35–65. New York: Dover, 1952.

Friedman, Michael. *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton, NJ: Princeton University Press, 1983.

Gödel, Kurt. "A Remark about the Relationship Between Relativity and Idealistic Philosophy." In *Albert Einstein: Philosopher-Scientist*, edited by P. Schilpp, 557–562. La Salle, IL: Open Court, 1949.

Grünbaum, Adolf. "The Meaning of Time." In *Basic Issues in the Philosophy of Time*, edited by E. Freeman and W. Sellars, 195–228. La Salle, IL: Open Court, 1971.

Grünbaum, Adolf. *Philosophical Problems of Space and Time*, (2nd, enlarged edition). Dordrecht, Netherlands: D. Reidel, 1973.

Horwich, P. *Asymmetries in Time: Problems in the Philosophy of Science*. Cambridge, MA: MIT Press, 1987.

Kroes, Peter. *Time: Its Structure and Role in Physical Theories*. Synthese Library, 179. Dordrecht, Netherlands: Kluwer, 1985.

Le Poidevin, Robin, ed. *Questions of Time and Tense*. Oxford: Oxford University Press, 1998.

Le Poidevin, Robin, and Murray McBeath, eds. *The Philosophy of Time*. Oxford: Oxford University Press, 1993.

Lewis, David, "The Paradoxes of Time Travel." In *Philosophical Papers*, vol. 2, edited by David Lewis. Oxford: Oxford University Press, 1986.

McCall, Storrs. *A Model of the Universe*. Oxford: Clarendon Press, 1994.

McTaggart, J. M. E. "The Unreality of Time." *Mind* New Series 68 (1908): 457–484.

Mellor, D. H. *Real Time II*. London: Routledge, 1998.

Minkowski, H. "Space and Time," as reprinted and translated in *The Principle of Relativity*, 73–91. New York: Dover, 1952.

Nerlich, Graham. *What Spacetime Explains*. Cambridge, U.K.: Cambridge University Press, 1994.

Newton-Smith, W. H. *The Structure of Time*. London: Routledge, 1980.

Poincaré, Henri. *Science and Hypothesis*. New York: Dover, 1952.

Price, Huw. *Time's Arrow & Archimedes' Point: New Directions for the Physics of Time*. New York: Oxford University Press, 1996.

Putnam, Hilary. "Time and Physical Geometry." *Journal of Philosophy* 64 (1967): 240–247. Reprinted in Putnam's *Collected Papers*, vol. 1. Cambridge, U.K.: Cambridge University Press, 1975.

Reichenbach, Hans. *The Philosophy of Space and Time*. New York: Dover, 1957.

Reichenbach, Hans. *The Direction of Time*. Berkeley: University of California Press, 1958.

Rietdijk, C. "A Rigorous Proof of Determinism Derived from the Special Theory of Relativity." *Philosophy of Science* 33 (1966): 341–344.

Savitt, Steven, ed. *Time's Arrows Today: Recent Physical and Philosophical Work on the Direction of Time*. Cambridge, U.K.: Cambridge University Press, 1995.

Savitt, Steven. "Being and Becoming in Modern Physics." In *The Stanford Encyclopedia of Philosophy (Spring 2002 Edition)*, edited by Edward N. Zalta. Available from http://plato.stanford.edu/.

Sklar, Lawrence. *Space, Time, and Spacetime*. Berkeley: University of California Press, 1974.

Sklar, Lawrence. "Up, Down, Left and Right, Past and Future." *Nous* 15 (1981): 111–129. Reprinted in *The Philosophy of*

Time, edited by Robin Le Poidevin and Murray McBeath. Oxford: Oxford University Press, 1993.
Smart, J. J. C. *Problems of Space and Time*. London: Macmillan, 1964.
Stein, H. "On Relativity Theory and Openness of the Future." *Philosophy of Science* 58 (1991): 147–167.
Swinburne, Richard. *Space and Time*. London: Macmillan, 1968.
Whitrow, G. *The Natural Philosophy of Time*. Oxford: Oxford University Press, 1961. (2nd edition, 1980.)
Williams, Donald C. "The Myth of Passage." *Journal of Philosophy* 48 (1951): 457–472.
Yourgrau, Palle. *Gödel Meets Einstein: Time Travel in the Gödel Universe*. La Salle, IL: Open Court, 1999.

Craig Callender (2005)

TIMON OF PHLIUS

(320–230 BCE)

Most of Timon's importance rests upon his reputation as a reporter, but he was also responsible for one or two original twists to the philosophy of his master—Pyrrho. He was a literary virtuoso, composing in a variety of verse forms. Seventy-one fragments of his poetry survive in quotations by later writers, sixty-five of them deriving from one work, the *Silloi*, a mock-epic series of lampoons in verse. The majority of them deal with philosophers other than Pyrrho, whom Timon attacks with wit and verve, frequently in pointed parody of Homeric verse; but Timon's purpose is to exalt Pyrrho at their expense: "Truly, no other mortal could rival Pyrrho; such was the man I saw, unproud, and unsubdued by everything which has subdued known and unknown alike, volatile crowds of people, weighed down in all directions by passions, opinion, and vain legislation" (Diels 1901, pgs. 8 and 9).

Timon portrays his hero as a superman: "Old man, how and whence did you find escape from the bondage of opinions and the empty wisdom of the sophists? How did you break the chains of all deception and persuasion? You did not concern yourself with what winds pass over Greece, and from what and into what each thing passes" (Diels 1901, p.48).

This philosophical hagiography deliberately recalls that of Socrates (note the rejection of natural science in the last fragment); Pyrrho is presented as a man apart from and immune to the seductive claims of pseudo-knowledge. But in the verse little of genuine philosophical substance is found, apart from the rejection of anything that smacks of dogmatic opinion: dogma unsupportable by persuasive argument, and the implication that such a rejection brings with it tranquillity.

But Timon also wrote prose works and a crucial report of one of them, *Pytho*, survives in a fragment of the Peripatetic Aristocles (around the first century CE), itself preserved in a text of Eusebius. Timon is reported as saying that anyone seeking happiness should consider these three questions: How are things by nature? What attitude should we adopt toward them? What will be the outcome for those who have this attitude? And he goes on to report (controversially) Pyrrho's answer: Things are indifferent, unmeasurable, and undecidable; neither sensation nor judgment is determinably true or false; and so one should not be opinionated, but be uncommitted and unwavering, saying about everything that it no more is than is not, or that it both is and is not, or that it neither is nor is not. Once accepted, the result is tranquility. In other words, we do not know how things really are; and once we accept that inability, it does not matter. However, Timon's Pyrrho, in contrast with later Pyrrhonians, claimed to be purveying a practical truth, albeit a skeptical one; in his other philosophical poem, *Images*, Timon writes: "The story of the truth has a correct rule, namely the nature of the divine and the good, from which derives the most equable life for man" (Diels 1901, p. 68).

The same poem contained the line: "the appearance prevails everywhere, wherever it comes from" (Diels 1901, p. 70). Here Timon encapsulates the central tenet of later skeptical philosophy, that one can neither question, nor go beyond, the content of appearances. Again anticipating a skeptical topos, in a work *On Sensations*, he wrote "that honey is sweet I do not affirm, but I accept that it appears so" (Diels 1901, p. 74).

In these passages, we may perhaps discern Timon's independent philosophizing; and reports in Sextus attribute views to Timon himself rather than via him to his master. In *Against the Geometers*, Sextus Empiricus attacks geometers on the ultimately Platonic grounds that they assume as firm principles what are in fact mere hypotheses, alluding to Timon's *Against the Physicists* as saying that one should investigate whether anything should be accepted on the basis of a hypothesis. Sextus gives no context; but the title of Timon's volume suggests that he would not have had the geometrical notion specifically in mind, but rather have been more generally concerned with the epistemic status of allegedly explanatory postulates. In this, too, he anticipates characteristic moves of later Pyrrhonism, in particular that encapsulated in the fourth mode of Agrippa.

Timon also dealt with time. Sextus reports that he argued against the indivisibility of the momentary present on the grounds that "no divisible thing such as

becoming or perishing, can come to be in an indivisible time" (Diels 1901, p. 76). Change involves a complex of distinct states: They cannot be squeezed into a partless present. That the present was a punctual; *now* was a tenet of Aristotelianism; the idea that no change can occur in a punctual present being a feature of Zeno's arrow paradox. Timon's argument was not, probably, very original in content. But it does show him adopting material supplied by the philosophical tradition and turning it to distinctively skeptical ends, something that was itself distinctive of the later skeptical tradition, and apparently unanticipated by anything we know of in Pyrrho. Thus if Timon's argument was unoriginal, the use to which it was put may well not have been. And herein lies his personal contribution to the development of Greek skepticism.

See also Agrippa; Ancient Skepticism; Pyrrho.

Bibliography

Bett, R. *Pyrrho, His Antecedents, and His Legacy*. Oxford; New York: Oxford University Press, 2000.

Diels, H. *Poetarum Philosophorum Fragmenta*. Berlin: Weidmann, 1901.

Hankinson, R. J.*The Sceptics*. 2nd ed. London; New York: Routledge, 1998.

R. J. Hankinson (2005)

TINDAL, MATTHEW

(1657?–1733)

Matthew Tindal, the English jurist, Whig propagandist, and deist, was born at Beer Ferris, Devonshire, the son of John Tindal, a minister. After an early education in the country, he proceeded to study law at Oxford, first at Lincoln College and later at Exeter College. In 1678 he was elected to a law fellowship at All Souls' College. In 1679 he received the BA and the BCL degrees and in 1685 the DCL. In 1685 he was also admitted as an advocate at Doctors' Commons, a society of ecclesiastical lawyers, with a pension of £200 a year for the remainder of his life. While at Oxford and under the influence of the high churchman George Hickes, he defected from the Church of England and became a Roman Catholic for a brief period, but he recanted in 1688. Soon thereafter, he began to publish a long series of tracts and books, culminating in 1730, when he was over seventy years old, with *Christianity as Old as the Creation.* Frequently called "the deist's Bible," this work elicited more than 150 replies, including Bishop Butler's famous *Analogy of Religion* (1736).

At Oxford, Tindal's enemies accused him of gluttony but granted that he was so abstemious in the drinking of wine that he frequently outsmarted them in argument. Dr. Edmund Gipson, bishop of London, however, won a posthumous "victory" over Tindal when he managed to acquire the manuscript of a second volume of *Christianity as Old as the Creation* and deliberately burned it. The same forged will (probably by Eustace Budgell) that made this action possible also deprived Tindal's nephew of his property.

Tindal died stoically in 1733 and was buried in Clerkenwell Church, London. Without question the most learned of the English deists, Tindal consistently referred to himself as a "Christian deist."

EARLY POLITICAL PUBLICATIONS

Tindal did not begin to publish until he was middle-aged. A first series of tracts, *Essay of Obedience to the Supreme Powers* (1694), *Essay on the Power of the Magistrate and the Rights of Mankind in Matters of Religion* (1697), *The Liberty of the Press* (1698), and *Reasons against restraining the Press* (1704), all showed low church and Miltonic influences. Tindal first gained notoriety with *The Rights of the Christian Church Asserted, against the Romish, and all other Priests who claim an Independent Power over it* (1706), which brought over twenty answers. A sequel, *A Defence of the Rights of the Christian Church* (1709), was condemned by the House of Commons and burned in 1710 by the common hangman. These early works are strongly Whiggish, anti-authoritarian, and anticlerical in tone; they argue for freedom of the press and for general toleration (except for atheists)—principles that were to be even more forcefully urged in *Christianity as Old as the Creation.* For his radical political view that although the magistrate has power to legislate in the area of religion, he has no authority to compel conformity and that persecution of nonconformity not only violates natural law but is also futile, Tindal, like many other deists, was branded by the orthodox as "Spinozan."

"THE DEIST'S BIBLE"

Christianity as Old as the Creation: Or, The Gospel A Republication of the Religion of Nature appeared in 1730 with subsequent editions in 1731, 1732, and 1733; in 1741 it was translated into German by Johann Lorenz Schmidt, a writer in the Leibniz-Wolff tradition. Although the work makes frequent mention of John Locke, it is fundamentally rationalistic, and it is the rationalistic side of Locke that is emphasized—that morality is capable of demonstration and is therefore true, that whatever is

known to be true on the basis of reason cannot be falsified by revelation, that the Bible must be read like any other book, that without reason any religion can be held to be true because of the power of tradition.

As is implied by the subtitle, Tindal's thesis is an elaboration of the proposition from Dr. Thomas Sherlock, bishop of Bangor and later of London, quoted on the title page: "The Religion of the Gospel is the true original Religion of Reason and Nature.… And its Precepts declarative of that original Religion, which was as old as the Creation." Citation from the rationalistic orthodoxy of such latitudinarians as Archbishop Tillotson, Samuel Clarke, and Thomas Sherlock, a deceptive device frequently employed by the deists, provides some indication of how close in thought rationalistic orthodoxy and rationalistic deism actually were.

Tindal's use of Sherlock's thesis, developed in a dialogue between *A* (Tindal) and *B* (an objector to, and a questioner of, *A*), is entirely negative. The Scriptures, with all the ambiguities that have confused the Church Fathers, the Schoolmen, and modern theologians, are really a work of supererogation. Although never stated in so many words, it is clear that Tindal's radical anticlericism challenged the validity of all historical religions and established churches.

On the critical and historical side, the Scriptures are examined and attacked by Tindal in great detail to expose the imperfect morality of certain Old Testament heroes and, to some extent, of certain parables of the New Testament. Even worse, according to Tindal, priestcraft and tradition, working together, have corrupted the texts and confused the people. Churches have used the teachings of the New Testament to acquire new members and have then used the teachings of the Old Testament to keep members in line. Tindal was incensed that priests first tempt men to examine their faith and then punish them for so doing if, perchance, their interpretations differ from those established by tradition and authority. This side of Tindal's work greatly influenced Voltaire.

On the philosophical side it is Tindal the rationalist, rather than the critic and moralist, who was the "Christian deist," for Tindal, like Lord Herbert of Cherbury before him, took what Alexander Pope was to call "the high Priori Road." God is conceived of as the God of reason, and because human nature is inalterable, man's reason has known His being and attributes from the beginning of time. Rational man, then, reasons downward from the divine perfections to morality and religion. All men, whether of the highest intellect or the meanest capacity, declares Tindal, are equally capable of knowing the immutable law of nature or reason and the religion of nature. In this respect Tindal is close to the more "orthodox" theologians of the waning rationalist or latitudinarian school in Britain represented by Archbishop Tillotson, Samuel Clarke, and Thomas Sherlock. The book concludes with Tindal's statement of his three basic notions about natural religion. First, there are things that show, by their inner nature, that they are the will of an infinitely wise and good God (for example, the relations between God and man, the immutability of morality). Second, there are things that have no worth in themselves, which are to be considered solely as means (forms of worship, positive regulations and precepts); these are to be used as men see fit in their quest for happiness. Third, there are things (the vested interests of priestcraft, miracles, "enthusiasm") so indifferent that they cannot be considered as either means or ends, and if emphasis is placed on them in religious matters, the worst sort of superstition ensues—and superstition is the enemy of true religion.

Tindal does not consider the fact that many people are totally incapable of right reason, a point that was dutifully reported by many of his opponents. The philosophical argument of Bishop Butler repudiated rationalism as the chimerical building of the world upon hypothesis in the manner of René Descartes. The paradoxical and abusive Bishop Warburton was content to dismiss Tindal's apriorism as "the silliest, and most wretched Error, in an age of Paradoxes." Tindal is the last and most influential of the British deists who sought to keep the movement on a high intellectual level.

See also Butler, Joseph; Clarke, Samuel; Deism; Descartes, René; Locke, John; Patristic Philosophy; Pope, Alexander; Voltaire, François-Marie Arouet de.

Bibliography

There is no collected edition of the works of Tindal. Early documents and "lives" include *A Copy of the will of Dr. Matthew Tindal, with an account of what pass'd concerning the same, between Mrs. Lucy Price, Eustace Budgell esq.; and Mr. Nicholas Tindal* (London, 1733); *Memoirs Of The Life and Writings of Matthew Tindall, LL.D. With A History Of The Controversies Wherein he was Engaged* (London, 1733); and *The Religious, Rational and Moral Conduct of Matthew Tindal, LL.D., late fellow of All Souls', by a member of the same college* (London, 1733).

There are no full-scale modern treatments of Tindal, but the following works are of some importance: Alfred O. Aldridge, "Polygamy and Deism," *Journal of English and Germanic Philology* 48 (1949): 343–360; Rosalie L. Colie, "Spinoza and the Early English Deists," *Journal of the History of Ideas* 20 (1959): 23–46; Arthur Friedman, "Pope and Deism," in *Pope*

and His Contemporaries: Essays Presented to George Sherburn, edited by James L. Clifford and Louis A. Landa (Oxford: Clarendon Press, 1949), pp. 89–95; Isabel Rivers, *Reason, Grace, and Sentiment: A Study of the Language of Religion and Ethics in England, 1660–1780,* 2 vols. (Cambridge, U.K.: Cambridge University Press, 1991–2000); Ernest Gordon Rupp, *Religion in England 1688–1791* (Oxford: Clarendon Press, 1986); Ernest Sirluck, "*Areopagetica* and a Forgotten Licensing Controversy," *Review of English Studies* 11 (1960): 260–274; Leslie Stephen, *History of English Thought in the Eighteenth Century,* 2 vols. (Bristol, U.K.: Thoemmes, 1991); Norman L. Torrey, *Voltaire and the English Deists* (New Haven, CT: Yale University Press, 1930), Ch. 5.

See also the general bibliography under the Deism entry.

Ernest Campbell Mossner (1967)
Bibliography updated by Philip Reed (2005)

TOLAND, JOHN

(1670–1722)

John Toland was an English deist, philosopher, diplomat, political controversialist, secular and biblical scholar, and linguist. Christened "Janus Junius" in the Roman Catholic Church, Toland later took the name of John. He was born near Londonderry, Ireland, possibly of partial French extraction. At the age of sixteen he ran away from school to become a Protestant Whig. In 1687 he turned up at Glasgow University and in 1690 was awarded an MA at Edinburgh University. For two years he studied at the University of Leiden under Friedrich Spanheim the younger, and in 1694 he settled at Oxford for some time to carry on research in the Bodleian Library. "The Character you bear in Oxford," he was informed by a correspondent, "is this; that you are a man of fine parts, great learning, and little religion."

The stream of books and pamphlets, mostly anonymous or pseudonymous, that followed has been estimated by various authorities to range from thirty to one hundred. His most famous work, *Christianity not Mysterious: Or, A Treatise Shewing That there is nothing in the Gospel Contrary to Reason, Nor above it: And that no Christian Doctrine can be properly call'd A Mystery,* appeared in 1696, when he was but twenty-five years old, elicited some fifty refutations and prosecution in both England and Ireland. In Ireland it was condemned by Parliament and ordered to be burned by the common hangman; an order was issued for the author's arrest. In England it was presented as a nuisance by the grand jury of Middlesex and roundly denounced in Parliament and in pulpit. In 1697, Toland replied to the Irish condemnation with the *Apology for Mr. Toland* and in 1702 to the English with *Vindicius Liberius: Or, Mr. Toland's Defence of himself.*

POLITICS

Toland's political publications are numerous. He was always the defender of toleration and the opponent of superstition and enthusiasm, a consistent Whig and a Commonwealth man. Outspoken and not very politic, he dedicated several of his tracts to the Whig deist Anthony Collins, who held similar convictions. Among Toland's more important political publications are the *Life of John Milton* (1698) and *Amyntor: Or, a Defence of Milton's Life* (1699), both of which have religious as well as political overtones. In 1701 the *Art of Governing by Parties* and *Anglia Libera: Or, the Limitation and Succession of the Crown of England explain'd and asserted* were published; the latter, supporting the Act of Settlement, was well received by Sophia, electress of Hanover. As a result Toland became secretary to the embassy to Hanover under Lord Macclesfield and presented a copy of the act and the book to Sophia. She was not, however, entirely pleased with his *Reasons for addressing his Majesty to invite into England their Highnesses, the Electress Dowager and the Electoral Prince of Hanover* (1702). Nevertheless, the electress was instrumental in introducing Toland to the court of Berlin and to her daughter Sophia Charlotte, wife of Frederick, the first king of Prussia. For the queen he composed *Letters to Serena* (1704) and *An Account of the Courts of Prussia and Hanover* (1705). At the invitation of the electress, Toland met Gottfried Wilhelm Leibniz and held numerous discussions with him in the presence of the queen. The two philosophers, though disagreeing on certain fundamentals, respected each other, kept up a correspondence for years, and to some extent were mutually influenced.

CAREER

Toland's chaotic career worsened throughout his life. He had early been under the political patronage of the third earl of Shaftesbury and later under that of Robert Harley, Lord Oxford. For the earl of Shaftesbury he had written political tracts, but Toland lost his friendship by publishing one of the earl's works, *An Inquiry concerning Virtue,* without authorization. For Harley he wrote political tracts and brought out an edition of James Harrington's *The Commonwealth of Oceana* with a biography but lost his friendship in 1714 with the *Art of Restoring* and *The Grand Mystery Laid Open,* wherein he implied distrust of his patron's loyalty to the Hanoverian succession. Of necessity, he became a Grub Streeter and lost everything

in the South Sea Bubble of 1720. As a result he either wrote or revised someone else's text of *The Secret History of the South-Sea Scheme.* The following year his health went into a rapid decline, abetted by the inept treatment of a physician, which inspired the indomitable Toland, ill as he was, to write a tract titled *Physic without Physicians* ("They learn their Art at the hazard of our lives, and make experiments by our deaths"). In 1722 he died in extreme poverty.

CHRISTIANITY NOT MYSTERIOUS

Like David Hume in "Of Miracles" (1748), Toland found an appropriate quotation for his title page from Archbishop Tillotson: "We need not desire a better Evidence that any Man is in the wrong, than to hear him declare against Reason, and thereby acknowledge that Reason is against him." The first edition appeared anonymously, but the second edition of the same year (1696) bore Toland's name.

Always professing some form of theism here and in subsequent writings, Toland, in his work, has affinities with the rationalistic religious common notions of Lord Herbert of Cherbury and with the empiricism and commonsense approach of John Locke in *An Essay concerning Human Understanding* (1690) and *Reasonableness of Christianity* (1695). He remained, however, fundamentally a rationalist in the line of Giordano Bruno, René Descartes, Benedict de Spinoza, and Leibniz.

Drawing freely upon Lord Herbert, the Cambridge Platonists, and Locke, though without naming names, Toland set out to prove that no Christian doctrine is mysterious—that is, above reason: "Could that Person justly value himself upon his being wiser than his Neighbors, who having infallible Assurance that something call'd a Blictri had a Being in Nature, in the mean time knew not what this Blictri was?" Faith and revelation involve both knowledge and assent, but revelation must rely upon the evidence of faith. In the Gospels, Toland correctly points out, "mystery" does not designate what cannot be known by man but, rather, what is revealed only to the chosen few. Faith, the hallmark of Puritanism, is consequently of no avail without the confirmation of reason.

Like many of the deists Toland argued that priestcraft introduced mysteries and then fostered them by ceremonies and discipline. Unlike Bishop Warburton, that eighteenth-century colossus of controversy who is alleged to have said, "Orthodoxy is my doxy; heterodoxy, another man's doxy," Toland ends *Christianity not Mysterious* with "I acknowledge no Orthodoxy but the Truth."

It was widely believed that Toland was a disciple of Locke, and he had been described to Locke by William Molyneux in 1697 as "a candid Free-Thinker, and a good Scholar." However, when *Christianity not Mysterious* aroused such a stir, Locke, who seems hardly to have realized the logical consequences of his own Arminianism (witness his prolonged controversy with Bishop Stillingfleet), repudiated any approval of his so-called disciple.

BIBLICAL CRITICISM

Oddly enough, Toland's biblical criticism first appears in the seemingly innocuous *Life of John Milton,* wherein, suggesting that the *Eikon Basilike* was not written by Charles I but was a priestly forgery, he proceeds to remark that many supposititious pieces under the name of Christ and his apostles had been accepted in the period of primitive Christianity. Divines rushed in where scholars feared to tread, charging Toland with attacking the authenticity of the Gospels. Toland speedily responded with *Amyntor,* which contains a catalog of apocryphal pieces twenty-two pages in length and is one of the earliest examinations of scriptural canon by an Englishman. Though in no sense definitive, Toland's catalog forced the issues of the canon and of early church history upon the scholars. Christ did not, he declares, institute one religion for the learned and another for the vulgar.

Toland's exploration of early Jewish religion and of the Druids' religion—he was an adept in the Celtic language—led him to the conviction that the simplicity of reason has been corrupted by the machinations of priestcraft. *Letters to Serena* explores somewhat unsystematically the beginnings of religion, examining the origin and force of prejudices, the history of the immortality of the soul among the heathens, the origin of idolatry, and motivations of heathenism. These and other explorations embryonically anticipate Hume in the *Natural History of Religion* (1757) and the *Dialogues concerning Natural Religion* (1779).

Toland argued that belief, prejudice, and superstition are ingrown from infancy. "You may reason yourself into what religion you please; but, pray, what religion will permit you to reason yourself out of it?"

He found a perfect example of surviving simple intuitive religion in a French letter written in 1688 from Carolina: "We know our Saviour's precepts without observing them, and they [the Indians] observe them without knowing him." As Toland put it elsewhere, "Those who live according to Reason … are Christians, tho' they be reputed Atheists." In "Hodegus," an essay of 1720, he inter-

prets Old Testament miracles by a naturalistic method, thereby anticipating Hermann Samuel Reimarus and the German rationalistic school of biblical exegesis.

PHILOSOPHICAL DEVELOPMENT

Toland's rationalism led him to translate and to defend Bruno's Latin treatise of 1514 on the infinite universe and innumerable worlds. In turn, he proceeded into a variety of naturalistic monism, which eventuated in pantheism. In the *Letters to Serena* he attacked Spinoza for his disavowal of the necessity of motion to matter, but in later works he had lavish praise for much of Spinozism. *Socinianism truly stated: being an example of fairdealing in theological Controversy,* a work of 1705 in which is found the first use of the word *pantheist,* is essentially pantheistic.

Toland's final statement, however, if it is to be taken seriously, was published in 1720 in Holland; termed "Cosmopoli," it was issued under the pseudonym Janus Junius Edganesius (indicating Inis-Eogan or Eogani Insuli, the northernmost peninsula of Ireland and the place of Toland's birth). *Pantheisticon: sive Formula celebrandae Sodalitatis,* the work referred to and translated into English in 1751, has been variously interpreted as a serious exposition of the philosophy of pantheism, a literary hoax, a sort of litany in derision of Christian liturgies, a mask to disguise atheism, a modernized version of the secret doctrines of Freemasonry, and a device to stimulate new thinking. The work consists of a dialogue between the president of a pantheistical society which acknowledges no other God than the universe and its members, who respond to his endeavors to inspire them with the love of truth, liberty, and health, cheerfulness, sobriety, temperateness, and freedom from superstition.

It is sufficiently evident that Toland was not a really original thinker but one who reflected many influences. Born Roman Catholic, he became Protestant. He was a latitudinarian, a freethinker, a deist, a materialist, and a pantheist. In a Latin epitaph that he composed for himself, he laid claim to the knowledge often languages. He was a prolific writer on many subjects, sometimes confused and contradictory, sometimes foreshadowing aspects of modern thought. In his life of fifty-two years his restless, inquiring mind was ever active, his accomplishments were manifold, and he was an internationalist of consequence in the Age of Enlightenment.

See also Bruno, Giordano; Cambridge Platonists; Collins, Anthony; Deism; Descartes, René; Enlightenment; Herbert of Cherbury; Hume, David; Leibniz, Gottfried Wilhelm; Locke, John; Materialism; Milton, John; Pantheism; Rationalism; Reimarus, Hermann Samuel; Shaftesbury, Third Earl of (Anthony Ashley Cooper); Spinoza, Benedict (Baruch) de; Stillingfleet, Edward; Toleration.

Bibliography

ADDITIONAL WORKS BY TOLAND

Other works of Toland not named above but worth mentioning include *Adeisidaemon, sive Titus Livius a Superstitione Vindicatus* (The Hague: Thomam Johnson, 1709); *Nazarenus, or Jewish, Gentile, and Mahometan Christianity* (London: J. Brown, 1718); *Tetradymus,* containing "Hodegus," "Clidophorus," "Hypatia," and "Mangoneutes" (London, 1720); and *The Miscellaneous Works of Mr. John Toland, Now first published from his Original Manuscripts,* with a life by Pierre Des Maizeaux, ed., 2 vols. (London, 1747), a reprint, with some additions, of *A Collection of Several Pieces by Mr. John Toland,* 2 vols. (London, 1726).

WORKS ON TOLAND

Champion, Justin. *Republican Learning: John Toland and the Crisis of Christian Culture, 1696–1722.* Manchester, U.K.: Manchester University Press, 2003.

Cragg, G. R. *From Puritanism to the Age of Reason.* Cambridge, U.K.: Cambridge University Press, 1950. Ch. 7.

Daniel, Stephen H. *John Toland: His Method, Manners, and Mind.* Kingston, ON: McGill-Queen's University Press, 1984.

Dyche, Eugene I. "The Life and Works and Philosophical Relations of John (Janus Junius) Toland, 1670–1722." *Abstracts of Dissertations, the University of Southern California,* 64–69. Los Angeles, 1944.

El Nouty, Hassam. "Le Panthéisme dans les lettres françaises au XVII[e] siècle: Apercus sur la fortune du mot et de la notion." *Revue des sciences humaines* 97 (1960): 435–457.

Heinemann, F. H. "John Toland and the Age of Enlightenment." *Review of English Studies* 20 (1944): 125–146.

Heinemann, F. H. "John Toland and the Age of Reason." *Archiv für Philosophie* 4 (1950): 35–66.

Heinemann, F. H. "John Toland, France, Holland, and Dr. Williams." *Review of English Studies* 25 (1949): 346–349.

Heinemann, F. H. "Prolegomena to a Toland Bibliography." *Notes and Queries* 185 (1943): 182–186.

Heinemann, F. H. "Toland and Leibniz." *Philosophical Review* 54 (1945): 437–457.

An Historical Account of the Life and Writings of the late eminently famous Mr. John Toland. London, 1722. "By one of his Intimate Friends."

Jacob, Margaret C. *The Radical Enlightenment.* London: Allen and Unwin, 1981.

Lange, Friedrich A. *History of Materialism.* New York: Humanities Press, 1950.

Lantoine, Albert. *Un précurseur de la franc-maçonnerie, J. Toland.* Paris, 1927.

McGuinness, Philip, Alan Harrison, and Richard Kearney, eds. *John Toland's Christianity not Mysterious.* Dublin: Lilliput Press, 1997.

Nourrison, Jean Felix. *Philosophies de la nature. Bacon-Bayle-Toland-Buffon.* Paris, 1887.

Simms, J. G. "John Toland (1670–1722), a Donegal Heretic." *Irish Historical Studies* 16 (1969): 304–320.

Sullivan, Robert E. *John Toland and the Deist Controversy.* Cambridge, MA: Harvard University Press, 1982.

See also the general bibliography under Deism.

Ernest Campbell Mossner (1967)
Bibliography updated by Philip Reed (2005)

TOLERATION

"Toleration" is a policy of patient forbearance in the presence of something that is disliked or disapproved of. Toleration must thus be distinguished from freedom or liberty precisely because it implies the existence of something believed to be disagreeable or evil. When freedom or liberty is said to prevail, no criticism, moral or otherwise, is entailed of the people who are said to be free or of the use to which such people put their freedom. Indeed, there are some writers who would reserve the words *liberty* and *freedom* for the rightful exercise of human choice, thinking, with the poet John Milton, that "only the good man can be free." Toleration, on the other hand, has an element of condemnation built into its meaning. We do not tolerate what we enjoy or what is generally liked or approved of. We speak of freedom of speech, of worship, and of movement—speech, worship, and movement being good or ethically neutral things. But when we speak of toleration, we speak of the toleration of heretics, dissenters, or atheists, all of whom were once thought to be wrongdoers, or we speak of the toleration of prostitution, gambling, or the drug traffic, all of which are still generally regarded as evils. To tolerate is first to condemn and then to put up with or, more simply, to put up with is itself to condemn.

T. S. Eliot once surprised his readers by saying, "The Christian does not wish to be tolerated." He did not mean, as some supposed, that the Christian yearned for martyrdom. He meant that the Christian did not wish to be put up with. The Christian wanted something better—to be respected, honored, loved. And what Eliot said in the name of Christians would doubtless also be said by Jews, Muslims, Mormons, African Americans, or any other minority group that finds itself tolerated by a larger society. Toleration is always *mere* toleration. It is less than equality just as it is distinct from liberty, and it is sharply at variance with fraternity. For these reasons toleration is far from an ideal policy; it is contaminated, so to speak, by that very implication of evil which its meaning contains.

Toleration must also be distinguished from indifference. A man who has no feelings about something is indifferent to it, not tolerant, for if he has no feelings, he cannot be said to dislike or disapprove of it. He cannot claim to put up with what troubles him in no way. It has sometimes been said by critics of religious toleration that such toleration is evidence of indifference to religion and that indifference to religion is bad. Here one must distinguish a logical connection from a historical one. It may well be a historical fact that the growth of religious toleration as a government policy in France and England during the eighteenth century was due to a diminution of religious fervor, to an increase in worldliness, and in a word, to indifference. Even so, however, the toleration must be distinguished from the indifference, for the words have significantly different meanings. There have been many men, like Thomas Hobbes, who were personally indifferent to religion but opposed to religious toleration, and many, like John Locke, who had strong religious beliefs but who favored religious toleration.

ALTERNATIVES TO TOLERATION

The alternative to toleration is often said to be persecution. This is a misleading dichotomy. Persecution is by definition always wrong. Moral condemnation is part of the meaning of the word. Yet who is to say that the alternative to toleration is always a wrong policy? Is the suppression of the drug traffic, for example, wrong? Is it persecution? It would be perverse to say that everything that is not tolerated is persecuted. Persecution is one alternative to toleration. However, there is another alternative which must be expressed in more neutral language, though, of course, it is one of the central difficulties of all social theory that neutral language is not always at one's disposal. Almost all the words we use in discussing social and political problems have a normative element in them. We might be wise, for lack of a better term, to rely on the word *suppression* as the alternative to *toleration.* To ask whether the persecution of religious dissenters was justifiable in thirteenth-century Europe is to prejudice the issue from the outset by speaking of persecution. But one might have an impartial discussion about whether the suppression of religious dissent was justifiable at that time and place, for even those who practiced it would agree to calling it suppression.

Many writers have opposed policies of toleration, but few have ventured to defend intolerance. This is clearly because intolerance in private life is considered a moral

defect or weakness, a defect allied to arrogance, narrow-mindedness, and impatience. Hence, intolerance has an unpleasant ring. James Fitzjames Stephen frankly advocated intolerance in opposition to John Stuart Mill's policy of toleration, but though Stephen's arguments were of a kind more likely to appeal to the majority, his success with the public was conspicuously less than Mill's; manifestly, Stephen had made an infelicitous choice of language. Most supporters of what Stephen called intolerance have preferred to speak of order, discipline, authority, or control in putting forward a case for suppression against one of toleration.

PAGAN AND CHRISTIAN ATTITUDES

The central problem of toleration in Western history was for centuries the problem of religious toleration. This is one of the consequences the West has faced because its religion is Christianity. Polytheistic religions are by nature more tolerant. The Greeks, for example, were conservative in the matter of religious ceremonies and institutions, but they admitted a great variety of theological beliefs. Where there were many gods, there could be many dogmas. And although Socrates and the Pythagoreans were persecuted, it was not on religious grounds but because they were accused of threatening the morality and political security of the community. The Romans were less steady in their policy, alternating between policies of general permissiveness and repression of particular sects—notably, but not exclusively, the Christians. Roman toleration was limited by at least one specifically religious notion, namely, the belief that the traditional deities would punish a whole people for the offense of those who failed to worship them.

The early Fathers of the Christian church, having themselves been cruelly persecuted by the Romans, were in favor of religious toleration as a principle. But as soon as Constantine made Christianity a state religion, the pagans, who had once been the persecutors, became the persecuted. Nevertheless, it may be recorded that the Christian repression of paganism never went to the cruel lengths to which Roman repression had gone. St. Augustine, an early advocate of suppressing heretics, went out of his way to say that the death penalty for heresy was wrong. The comparatively few pagans who were put to death by the Christian emperors were usually executed on charges of sorcery rather than of worshiping false gods.

This policy of moderate repression continued throughout the early Middle Ages. In the late Middle Ages, the Renaissance, the Reformation, and the Counter Reformation, toleration was virtually repudiated on principle by European Christians. The few Christians who continued to favor religious toleration are conspicuous for that very reason. They include the Anabaptists in Germany, the Arminians in Holland, Huldrych Zwingli in Switzerland, Sebastian Castellio in France, Socinus in Poland. But the main Protestant churches, whether Lutheran, Calvinist, or Anglican, were not conspicuously more tolerant than the Catholic Church. The Catholic Church's chief instrument of religious discipline was the Inquisition, which freely employed torture as well as the death penalty in its endeavors to recover erring souls for God.

Christian arguments in defense of repression are several. Some writers repeat the old pagan argument that God is offended by heretical practices and is likely to inflict disasters on the whole community as a punishment. Other writers stress the point that heresy is a crime, a form of revolt against lawful authority, a culpable betrayal of promises made (even if only by proxy) at baptism. Crime, it is argued, cannot be tolerated. A more sophisticated argument maintains that the authority of the church is as essential to the continued existence of civil society as is that of the state; hence, those men who defy the church are akin to those who repudiate their duty to the king. Thus, members of such religious sects as the Cathari, Waldenses, and Albigenses are regarded by certain Catholic theorists as seditious rebels who have put themselves in a state of war with the sovereign power. The true religion seals men together in the safety of the commonwealth; dissent and heresy are therefore likely to open the way to anarchy. Furthermore, it is held by all these Christian writers that to tolerate heresy is to do no service to the man concerned, for to leave him alone in his error is to leave him in a state of sin, faced with the prospect of eternal damnation in the life to come. It is thus thought to be no real cruelty to inflict painful penalties, even death itself, on an erring man if by so doing one is sparing him the far greater torments of hell.

PHILOSOPHICAL ARGUMENTS FOR TOLERATION

The philosopher who is best known for having addressed himself to the Christian arguments for suppression was the Englishman John Locke. In the seventeenth century Christians were generally beginning to lose confidence in the old policy of repression, although it was still being practiced. The unity of Christendom was plainly ended and not likely to be recovered. Protestantism in its various forms had come to be almost as great a power in the world as Catholicism. The old notion of one true faith

against heresy had lost its meaning. Besides, although Protestantism in its leading forms did not preach toleration, it preached a gospel that led inexorably to the demand for toleration; the Protestant doctrine that every man must be a priest unto himself gave the dissenter just as good grounds as the orthodox believer for claiming that his faith was true. Confidence in the utility—and justice—of suppressing unorthodox opinions was shaken by such writers as Pierre Bayle (1647–1706), who in his *Pensées sur la comète* (1682), argued that morality is independent of religion.

Locke's plea for toleration, set forth in his *Epistola de Tolerantia,* published in 1688, was not the first such plea, but it was the earliest systematic argument in its favor. Locke's first point is that repression is not an effective policy. Force can be used to make a man go through the motions of a given form of Christian worship, but force cannot make a man entertain any faith or belief in the privacy of his soul. What force can do is make a man pretend to be an orthodox believer. And such a policy, says Locke, is not only useless but also morally harmful since it is bound to breed hypocrisy. Locke thus totally rejects the Catholic argument that force—let alone torture and death—can bring any man to salvation.

Second, Locke rejects the traditional argument that a man's obligation to the church is equal to his obligation to the state and that civil society will lapse into anarchy if religious dissent is tolerated. Locke describes the church as a "voluntary society" which has a mission in the world quite independent of the functions of the state. The church exists to save men's souls, and it can fulfill this mission only by persuasion, by essentially nonviolent means. The state, on the other hand, exists to protect men's rights—their lives, liberties, and estates—so that the use of force as an ultimate sanction is a necessary part of the state's function. The state has no concern with the salvation of men's souls, just as the church has no concern with the use of force. Nor has the state any knowledge of what the true religion is. The Persian ruler believes it is Islam; the Spanish ruler believes it is Catholicism; the English king believes it is Anglicanism. They cannot all be right. Therefore, that a religion is established is no evidence that it is the true religion. Each man has his own faith, and every person's conscience is entitled to the same respect.

Locke's theory of toleration was intimately connected with his theory of freedom. Since he held that one of the most fundamental reasons for the existence of the state was the preservation of man's natural right to liberty, he argued that the government was entitled to use force against an individual only when it was necessary to protect the rights of others. Certain things, Locke agreed, could not be tolerated: (1) the propagation of "opinions contrary to human society, or to those moral rules which are necessary to the preservation of civil society"; (2) any claim "to special prerogative opposite to the civil right of the community"; (3) the activity of "persons who are ready on any occasion to seize the government, and possess themselves of the estates and fortunes of their fellow subjects"; (4) transferring allegiance to a foreign prince; and (5) denying the existence of God.

Locke's reason for withholding toleration from atheists was the rather quaint one that a man who did not believe in God could not take a valid oath and that oaths and covenants were "the bonds of human society." Locke was unwilling to extend toleration to Roman Catholics, not on religious grounds but because he held, with some reason, that Roman Catholics were not loyal subjects of the English crown, since they owed their first allegiance to a foreign prince, the pope.

Locke's argument for toleration, which seemed distinctly avant-garde when it was first published, eventually came to be regarded as common sense. Indeed, even Catholic teaching on the subject of toleration moved toward Locke's position. Later Catholic apologists distinguished between (1) theological dogmatic toleration, (2) practical civil toleration, and (3) public political toleration. The first, theological dogmatic toleration, was resisted as firmly as ever. The teaching of the Catholic Church was held to be the absolute and certain truth; thus, to tolerate any opinion at variance with it would be to tolerate falsehood, and the clear duty of the rational mind to uphold truth and deny falsehood imposed an equally categorical duty to deny any religious or moral teaching at variance with the teaching of Rome, which is infallible. However, what is called practical civil toleration was gradually accepted by Catholics. First, it was said to be the Christian's duty to distinguish between the error and the man who erred. Error was always to be opposed, but the man who erred was to be regarded, in full Christian charity, as a fellow man and, therefore, not to be persecuted. On public political toleration, later Catholic theory was somewhat ambiguous. This was because of the need to claim for Catholic minorities in Protestant states the utmost possible toleration without equally committing Catholic governments to tolerating Protestant minorities. Thus, the principle of public political toleration was admitted to vary between its application in a secular state and in a "truly Christian state."

The outstanding exponent of the case for greater toleration in the nineteenth century was John Stuart Mill. In many ways his argument followed the lines laid down by Locke, but Mill put fewer limitations on toleration than did Locke. He was more insistent that the only justification for interfering with any man's liberty was a reasonable assurance that some danger or threat to the liberty of another was involved. Again, where Locke was exclusively concerned with the protection of individual liberty from the interference of state and church, Mill was increasingly concerned with the limitations on human freedom that stemmed from unwritten law—the pressure of convention and public opinion. Mill wanted to see toleration extended from the realm of politics to that of morals and manners, to all self-regarding actions, as he called them. Mill, as a Victorian, lived, of course, in a society that not only frowned on things like free love, adultery, and Sabbath-breaking but also vigorously applied the social sanction of ostracism to any who committed these sins. Mill felt that people were more oppressed and hemmed in by the unwritten laws than they were by laws enforced by the state and that human freedom and variety could not flourish in a repressive atmosphere. Mill demanded toleration because he held that liberty, individuality, and variety were of the highest ethical value; they were what made man "nobler to contemplate."

Mill's ablest critic, James Fitzjames Stephen (in his book *Liberty, Equality, Fraternity,* written in reply to Mill's essay *On Liberty*), argued that intolerance was a necessary preservative of society. The modern liberal state was possible precisely because society was able to discipline itself through unwritten laws. It was a good thing for men to be compelled by social intolerance to keep laws of conduct that the wisdom of the ages had shown to be good. Mill's claim that there was a class of self-regarding actions that had a right to be tolerated because they did not affect others was, in Stephen's view, unfounded; almost everything a man did affected someone else. Suicide, intemperance, debauchery, and so forth were not things that injured the agent alone. The class of self-regarding actions was virtually an empty one. And since almost all conduct was other-regarding, society had a right to interfere as widely as it did. Stephen argued that the general run of men did not have the wit to think out moral codes of their own or the strength of character to obey such codes if they established them. Hence, some form of external sanction was needed if morality was to be upheld. Stephen also rejected Mill's view that variety was a good thing in itself. Goodness, he agreed, was varied, but that did not mean that variety itself was good; a nation in which half the population was criminal would be more diversified than a wholly honest one, but it would not be a better nation. Dissent for its own sake Stephen condemned as frivolous and sentimental Bohemianism. Eccentricity was a mark of weakness rather than of strength; and constraint, far from being an evil, was a great stimulus to exertion. Stephen even held that the intolerance that went with the Puritan spirit had been one of the chief factors enabling England to surge ahead of other nations in making industrial and social progress.

POLITICAL TOLERATION

With the rise of totalitarian governments in the twentieth century, the problem of toleration took on a new aspect. For democratic and freedom-loving governments the toleration of intolerance became an acute problem. In 1936 the British government introduced a ban on political uniforms because of the disturbances caused by Oswald Mosley's fascist movement and its black-shirted adherents; an attempt was made under Harold Wilson's Labour government in 1965 to proscribe acts of racial discrimination. After World War II the United States was troubled by the difficulty of deciding how much toleration could be safely extended to communists when several communists proved to be Russian or Cuban agents and when all communists seemed to have a more pronounced loyalty to the Soviet Union than to the United States. The position of the communists in twentieth-century America was thought to resemble that of the Catholics in seventeenth-century England, and many Americans recalled Locke's view that such persons had forfeited their right to toleration. Other Americans argued that repression was futile; the interdiction of open communist organizations would do little to protect the state from secret and more sinister communist activities. Hence, an abridgment of political toleration would do no good to anyone, for it would simply create martyrs without eliminating spies. Thus, the argument both for and against political toleration in the twentieth century cannot be said to have differed greatly from the debate concerning religious toleration that exercised the minds of earlier generations.

See also Augustine, St.; Bayle, Pierre; Eliot, Thomas Stearns; Freedom; Hobbes, Thomas; Liberty; Locke, John; Mill, John Stuart; Milton, John; Socinianism.

Bibliography

Bonet-Maury, G. *Histoire de la liberté de conscience en France.* Paris: Alcan, 1909.

Budziszewski, J. *True Tolerance: Liberalism and the Necessity of Judgment.* New Brunswick, NJ: Transaction, 1992.

Creighton, M. *Persecution and Tolerance.* London: Longmans Green, 1895.
Devlin, P. A. *The Enforcement of Morals.* London: Oxford University Press, 1965.
Hart, H. L. A. *Law, Liberty and Morality.* London: Oxford University Press, 1963.
Horton, John, and Peter P. Nicholson, eds. *Toleration: Philosophy and Practice.* Aldershot, U.K.: Avebury, 1992.
Jordan, W. K. *The Development of Religious Toleration in England.* London: Allen and Unwin, 1932–1940.
Kamen, Henry Arthur Francis. *The Rise of Toleration.* London: Weidenfeld and Nicolson, 1967.
Kilcullen, John. *Sincerity and Truth: Essays on Arnauld, Bayle, and Toleration.* Oxford: Oxford University Press, 1988.
King, Preston. *Toleration.* London: Allen and Unwin, 1976.
Lecky, W. E. *History of the Rise of Rationalism in Europe.* London: Longmans, 1877.
Lecler, J. *Histoire de la tolérance au siècle de la Réforme.* Paris, 1955.
Locke, John. *Epistola de Tolerantia.* Gouda, 1688.
Luzzatti, L. *La libertà di conscienza.* Milan, 1909.
Mendus, Susan, ed. *Justifying Toleration. Conceptual and Historical Perspectives.* Cambridge, U.K.: Cambridge University Press, 1988.
Mendus, Susan. *Toleration and the Limits of Liberalism.* Atlantic Highlands, NJ: Humanities Press, 1989.
Mendus, Susan, and David Edwards, eds. *On Toleration.* Oxford: Clarendon Press, 1987.
Mill, John Stuart. *On Liberty.* London, 1859.
Newey, Glen. *Virtue, Reason, and Toleration: The Place of Toleration in Ethical and Political Philosophy.* Edinburgh: Edinburgh University Press, 1999.
Rawls, John. *Political Liberalism.* New York: Columbia University Press, 1993.
Ruffini, F. *La libertà religiosa.* Florence, 1901.
Seaton, A. A. *The Theory of Toleration under the Later Stuarts.* Oxford, 1911.
Stephen, J. F. *Liberty, Equality, Fraternity.* London, 1861.
Tinder, Glenn E. *Tolerance and Community.* Columbia: University of Missouri Press, 1995.
Voltaire, F.-M. de. *Traité sur la tolérance.* Paris, 1763.
Wolff, Robert Paul, Herbert Marcuse, and Barrington Moore Jr. *A Critique of Pure Tolerance.* London: Jonathan Cape, 1969.

Maurice Cranston (1967)
Bibliography updated by Philip Reed (2005)

TOLETUS, FRANCIS
(1532–1596)

Francis Toletus, the first important Jesuit philosopher, was born in Córdoba, Spain. He studied philosophy at the University of Valencia and theology at the University of Salamanca under Dominic de Soto. While a professor of philosophy at Salamanca, Toletus entered the Jesuit order (1558). He taught philosophy at the order's Roman College from 1559 to 1563 and theology from 1563 to 1569. In 1593 Toletus became the first Jesuit cardinal. He died in Rome.

Toletus's Latin philosophical works include commentaries on the logic, physics, and psychology of Aristotle; Toletus's commentary on Thomas Aquinas's *Summa* (*Enarratio in Summam Theologiae Divi Thomae*) also contains philosophical material. In all these works his views are Thomistic with many personal modifications. In the theory of knowledge, Toletus taught that individual things are directly apprehended by the intellect, that the primary object of knowledge is a sort of particularized form (*species specialissima*) and not being in general (*Physica,* Venice, 1600, p. 12), that intellectual abstraction is simply a precision from accidents and a consideration of the substance of anything (*De Anima,* Venice, 1575, p. 170), that the agent intellect may be fundamentally the same power as the possible intellect (*De Anima,* Venice, 1586, pp. 144–146). His metaphysics is distinguished by a theory of triple acts in the same being: formal, entitative, and existential (*Physica,* p. 33). The existential act is limited in two ways: by the receptive potency and by its efficient cause (*Enarratio,* Vol. I, p. 118). He denied that essence and existence are really distinct principles (*Physica,* p. 34; *Enarratio,* Vol. I, p. 79), and that matter is pure potency; it has its own actuality (*Physica,* pp. 32–36), but form is the principle of individuation (*De Anima,* p. 163). The number of the categories (ten) in Aristotle's logic is merely probable. It is possible rationally to demonstrate the existence of God but the famous "five ways" of Thomas are incomplete; they do not establish the key attributes of God (*Enarratio,* Vol. I, 69).

See also Aristotle; Epistemology; Epistemology, History of; Soto, Dominic de; Thomas Aquinas, St.; Thomism.

Bibliography

The Latin philosophical works are collected in *Omnia quae Hucusque Extant Opera* (Lyons, 1586). See also *Enarratio in Summam Theologiae Divi Thomae.* 4 vols. (Rome, 1869–1870).
Works on Toletus include L. Morati, "Toledo, Francisco de," *Enciclopedia filosofica* (Venice and Rome, 1957), Vol. IV, cols. 1216–1217 (an excellent entry to which the present account is indebted); C. Giacon, *La seconda scolastica* (Milan: Fratelli Bocca, 1944), pp. 25–44, 51–65.

Vernon J. Bourke (1967)

TOLSTOY, LEV (LEO) NIKOLAEVICH
(1828–1910)

Lev (Leo) Nikolaevich Tolstoy, the renowned Russian novelist, won worldwide fame as a moralist and sage for his antiecclesiastical interpretation of Christianity and fervent preaching of nonviolence. A well-read amateur in philosophy from the age of fifteen, Tolstoy displayed serious philosophical interests in his greatest novel, *War and Peace* (1865–1869), and in 1874 he began an increasingly anguished philosophical and religious quest, seeking a reason for living. His spiritual crisis, dramatically described in *My Confession* (1879), was resolved by a return to the Christian faith of his youth, but in a radically different form based on his reading of selected New Testament texts. The new creed, further elaborated in such works as *What People Live By* (1881) and *What I Believe* (1883), was the foundation for the philosophical and hortatory works on morality, society, and culture that dominated his writing during the last three decades of his life.

PHILOSOPHY OF HISTORY

Tolstoy conceived *War and Peace* as a grand historical narrative embodying conclusions he had reached, partly under the influence of Schopenhauer, concerning causality in history and especially the interplay of freedom, chance, and necessity; the novel's two epilogues address these themes explicitly. It is in the nature of human consciousness, Tolstoy argued, to conceive of oneself and others as free agents whose actions may have a significant impact on the world—in the case of so-called great figures like Napoleon, a determining impact. Yet no individual is more than one node in a vast and unpredictable web of interacting forces, conscious and unconscious, contingent and necessary. Hence individuals cannot with any assurance foresee the effects of their own or others' actions (a point to which Tolstoy returned in his case against violence), and great men do not make history. He delights in describing, for example, how the tide of a decisive battle can be turned by the behavior of a single rank-and-file soldier—although this example undercuts his own arguments against attributing a determining influence to any one person. Tolstoy's philosophy of history is analyzed insightfully in Isaiah Berlin's classic study, *The Hedgehog and the Fox* (1957).

METAPHYSICS AND EPISTEMOLOGY

In *My Confession* Tolstoy expressed his disillusionment with all attempts by human reason, whether philosophical or scientific, to explain how life can have meaning when it inevitably ends in death. Meaning, he decided, can be imparted to a finite life only by linking it with an eternal, infinite reality—by which he meant the spiritual reality of the Christian God—and such union with an infinite deity is achievable only through an act of faith. Though itself "unreasonable," the primitive act of faith answers the ultimate question posed by reason without disqualifying reason from serving as the standard of truth on other questions. Tolstoy accordingly sought to develop something he had dreamed of as early as 1855: a rational religion, one stripped of everything unreasonable, including miracles, sacraments, mysticism, clergy, rituals, special buildings, and dietary rules. Tolstoy's standard of reasonableness proved to be highly fluid and subjective, however. In a Rousseauian spirit he rejected much of modern science and technology as products of false reason, and the mysticism he condemns in some contexts appears to be embraced in others.

Tolstoy's metaphysical views are a form of Christian idealism based on a dualism of matter and spirit. Reality is bifurcated into an infinite, eternal divine world and a finite, temporal material world, with human beings mirroring this division in their possession of a body and a soul. The universal divine reality is manifested in the human soul in the form of love, so that only when people are vehicles of universal love are they living a "true" life, "a life divine and free" (Edie 1976, p. 218). In several respects, however, Tolstoy departed from the commonly accepted Christian versions of this picture, prompting the Russian Orthodox church to excommunicate him in 1901. He opposed Trinitarianism and denied the special divinity of the man Jesus, contending that he was no different in nature from any other son of God. Further, despite frequent references to God as a "Father," Tolstoy did not subscribe to a personal conception of God. His conception, rather, as Richard F. Gustafson has argued in *Leo Tolstoy, Resident and Stranger* (1986), is panentheistic: God is both transcendent and immanent; He is "beyond the world of space and time but includes within Him all the world of space and time" (Gustafson 1986, p. 101). Tolstoy also rejects personal immortality in the sense of an individual life after death, holding rather that individuals attain immortality by merging with the infinite. Gustafson sees the influence of Eastern Christianity in Tolstoy's theology, whereas David Kvitko, in *A Philosophic Study of Tolstoy* (1927), argues that Tolstoy's metaphysical

views in general were indebted more to Buddhism than to Christianity. Tolstoy's interest in and extensive knowledge of Chinese philosophy has been well documented by the sinologist Derk Bodde in *Tolstoy and China* (1950).

ETHICS

Tolstoy states that he found the true meaning of Christ's teaching in the Sermon on the Mount as reported in the gospel of Saint Matthew, the text that became the focal point of his thinking about personal and social morality. From the sermon he distilled a moral code consisting of five commandments: first, do not be angry; second, do not lust; third, do not take oaths; fourth, do not resist evil by force; and fifth, love all people, including your enemies. The first, fourth, and fifth commandments are expressions of what, to Tolstoy, was the unique Christian understanding of the universally recognized law of love (the Old Testament's injunction to love one's neighbor as oneself). All the great religions of antiquity, as he explained later in *The Law of Violence and the Law of Love* (1908), considered love a virtue, but only Christianity acknowledged it as a categorical demand, as "the supreme law of human life—i.e., in such a way as not to admit of exceptions in any case" (Edie 1976, p. 217). Christ, in other words, recognized the law as prohibiting *all* use of violence.

Tolstoy was called upon repeatedly to justify his absolutist interpretation of the law, and he did so consistently and with great vigor, not hesitating to condemn violence even when used in self-defense against a mad dog or against a savage who is preparing to slaughter one's children. To support his position he relies not simply on his religious faith but on two philosophical objections to violence that undeniably carry some weight, though perhaps not enough to justify his extreme stance. The first, echoing his skepticism about predictability in *War and Peace*, is that arguments for the use of violence to stop evil rest on the dubious assumption that we can reliably foresee and control the future. The second is that the use of force generates more force in return, making it counterproductive. As the acknowledged prophet of nonviolent resistance, Tolstoy found a devoted disciple in Mohandas Gandhi (with whom he corresponded) and a host of admirers among figures as diverse as Clarence Darrow and Ludwig Wittgenstein.

Tolstoy's second commandment—do not lust—although logically unrelated to the law of love, was advanced with equal maximalism. He treated it as not only a condemnation of extramarital relations but also as a call for celibacy even in marriage. In defending the ideal of universal celibacy he was unmoved by the argument (offered before the development of artificial insemination) that if his ideal were realized, it would mean the end of the human race. His response was, first, that humanly irresistible lapses would more than suffice for the continuation of the species; and second, that in any event, physical extinction would eliminate only the troublesome animal dimension of humanity and thus would be no great loss. Tolstoy's interest in the themes of sexuality and sexual misconduct (to which he himself confessed) gave him literary subjects—especially in later works such as *The Kreutzer Sonata* (1889) and *Resurrection* (1895–1899)—and some awareness of feminist issues.

SOCIAL AND POLITICAL THOUGHT

As the institution that claims a monopoly on the use of violence in society, the state was an obvious target for Tolstoy's moral indignation, and his antistate position ranks as one of the most sweeping in the annals of nonviolent anarchism. He opposed not only serving in the military or the police but also all activity that promotes or supports state force indirectly, such as paying taxes, serving on juries, and holding public office. Moreover, he condemned private ownership and other institutions that are sustained by the threat of state force. Tolstoy saw the gospel injunction against oath-taking (the third of his five commandments) as a recognition of the evils of acknowledging state authority; it confirmed his conviction that there was divine sanction for civil disobedience.

Although Tolstoy himself held a minor position as a justice of the peace in the early 1860s, his other civic activities after his army service (which ended in 1856) were outside any official sphere. In 1859 he founded a school for peasant children on his estate at Iasnaia Poliana and for the next few years devoted much attention to pedagogical theory and practice, producing essays (discussed in Charles Baudouin's *Tolstoi: The Teacher* [1923]) of interest to historians and theorists of education. During the famines of 1873 and 1891–1892, he worked tirelessly in the Russian countryside to organize relief efforts, publicly castigating the tsarist government for its incompetent handling of the crises. Later in the 1890s he provided moral and material support to the *Dukhobors* (literally, "spirit wrestlers"), a Russian sect that attempted to practice Christian anarchism on principles paralleling his own, and he spearheaded the successful drive to arrange for their mass relocation to Canada to escape tsarist persecution. Tolstoy's criticisms and civic initiatives angered the authorities, but he was protected

from serious reprisals (other than excommunication) by the enormous popular respect he enjoyed.

AESTHETICS

The most professional and enduring of Tolstoy's philosophical writings, despite its eccentric conclusions, is his book *What Is Art?*, originally published serially in 1897–1898 in the leading Russian journal of philosophy. The work is valued for its systematic approach to aesthetic philosophy, beginning with a critical survey of earlier attempts to define art and ending with a clear and forceful presentation of an expressionist theory centering on the notion of the communication of emotion from artist to audience.

"Art begins," Tolstoy wrote, "when one person, with the object of joining another or others to himself in one and the same feeling, expresses that feeling by certain external indications." The aim is achieved when the feeling is successfully transmitted or, as Tolstoy puts it, when "the spectators or auditors are infected by the feeling which the author has felt." The feeling transmitted, he adds, may be "very strong or very weak, very important or very insignificant, very bad or very good"; any feeling will do as far as art per se is concerned (pp. 121–123). From a strictly aesthetic point of view, then, the worth of art depends simply on its emotional infectiousness, which Tolstoy traced to the individuality, clarity, and sincerity of the feeling conveyed.

Tolstoy is by no means satisfied with a merely aesthetic approach to art, however, and the center of gravity of his treatise soon shifts to the moral demands that art, like every other aspect of culture, must satisfy. Art, according to Tolstoy, must reflect the loftiest religious perception of its time, which means in the modern day that the artist is called upon to communicate feelings flowing from "a perception of our sonship to God and of the brotherhood of man" (p. 240). This does not imply, as some of Tolstoy's critics have charged, that art can be of value only if it transmits specifically religious emotions. Tolstoy indeed esteems religious art as the highest form, but he also strongly commends the whole range of what he calls "universal" art, or art that simply promotes "the loving union of man with man" by transmitting "even the most trifling and simple feelings if only they are accessible to all men without exception, and therefore unite them" (pp. 240–241). What even the most generous critic finds hard to accept, however, is that on Tolstoy's criteria (and by his own admission) simple folk songs are greater music than Beethoven symphonies, and *Uncle Tom's Cabin* is a greater novel than *War and Peace*.

See also Aesthetics, History of; Anarchism; Art, Expression in; Art, Value in; Life, Meaning and Value of; Mysticism, Nature and Assessment of; Schopenhauer, Arthur; Violence; Wittgenstein, Ludwig Josef Johann.

Bibliography

Baudouin, Charles. *Tolstoi: The Teacher*. Translated by Fred Rothwell. New York: Dutton, 1923.

Berlin, Isaiah. *The Hedgehog and the Fox: An Essay on Tolstoy's View of History*. New York: Mentor Books, 1957.

Bodde, Derk. *Tolstoy and China*. Princeton: Princeton University Press, 1950.

Edie, James M., James P. Scanlan, Mary-Barbara Zeldin, and George L. Kline, eds. *Russian Philosophy*. Vol. 2, *The Nihilists; the Populists, Critics of Religion and Culture*. Knoxville, TN: University of Tennessee Press, 1976.

Egan, David R., and Melinda A. Egan, eds. *Leo Tolstoy: An Annotated Bibliography of English-Language Sources to 1978*. Metuchen, NJ: Scarecrow Press, 1979.

Gustafson, Richard F. *Leo Tolstoy, Resident and Stranger: A Study in Fiction and Theology*. Princeton, NJ: Princeton University Press, 1986.

Hopper, Stanley R., ed. *Lift up Your Eyes: The Religious Writings of Leo Tolstoy*. New York: Julian Press, 1960.

Kvitko, David Iur'evich. *A Philosophic Study of Tolstoy*. New York: Author, 1927.

Orwin, Donna Tussing. *Tolstoy's Art and Thought, 1847–1880*. Princeton: Princeton University Press, 1993.

Silbajoris, Rimvydas. *Tolstoy's Aesthetics and His Art*. Columbus, OH: Slavica, 1991.

Tolstoy, Leo. *The Complete Works of Count Tolstoy.* 24 vols, edited and translated by Leo Wiener. Boston: Dana Estes, 1904.

Tolstoy, Leo. *The Kingdom of God Is Within You: Christianity Not as a Mystic Religion but as a New Theory of Life*. Translated by Constance Garnett. Lincoln, Nebraska: University of Nebraska Press, 1984.

Tolstoy, Leo. *My Confession, My Religion*. Midland, MI: Avensblume Press, 1993.

Tolstoy, Leo. *Polnoe sobranie sochinenii* [Complete collected works]. 90 vols. Moscow: Gos. izd. khudozhestvennoi literatury, 1928–1958.

Tolstoy, Leo. *War and Peace*. Translated by Constance Garnett. New York: Modern Library, n.d.

Tolstoy, Leo. *What Is Art?* Translated by Alymer Maude. London: Oxford University Press, n.d.

James P. Scanlan (2005)

TOTALITARIANISM

See *Fascisim*

TOUCH

Two bodies are said to be touching if there is no spatial gap between some point on the surface of one and some point on the surface of the other. If one of the touching bodies is that of a sentient being, it may be aware of certain properties of the other body: for instance, that it is hot or cold, rough or smooth, wet or dry, hard or soft, sweet or sour. The sentient being is said to be aware of an object's sweetness or sourness by taste. (Aristotle attributes our distinguishing taste from touch to the fact that only a part of our flesh is sensitive to flavor.) The remaining properties the sentient being is said, in common speech, to be aware of by touch. Accordingly, touch appears in the traditional list of senses, with sight, hearing, and so on.

ARISTOTLE

Aristotle remarks that in the case of touch the contraries hot-cold, dry-moist, and hard-soft do not seem to have a single subject in the way in which the single subject of the properties acute-grave and loud-soft is sound, which is perceived by hearing. This may lead one to say that there are really a number of different senses that are mistakenly referred to as one sense, touch, perhaps because the body of a sentient being must touch an object in order for it to be aware by any of them of that object's properties. Or one may say that there is a single subject of the different contraries, namely, a material thing, and that there is only one sense, touch, whereby we are aware of the different properties of which the material thing is a subject. If one takes the latter course, it may appear that touch is the only sense whose proper object is the material world.

LOCKE, BERKELEY, AND CONDILLAC

To John Locke, it seemed that "the idea most intimately connected with and essential to body, so as nowhere else to be found or imagined, but only in matter" was the idea of solidity. This idea is received by touch and "arises from the resistance which we find in body to the entrance of any other body into the place it possesses."

As Locke held it to be by touch that we receive the idea of solidity, the idea essential to body, so George Berkeley, in his *Essay towards a New Theory of Vision*, held it to be touch alone that directly acquaints us with the external world. He abandoned this view in *The Principles of Human Knowledge*, maintaining that the objects of touch are as much sensations as are the objects of sight.

Locke regarded solidity as a "simple idea": "If anyone asks me what this solidity is, I send him to his senses to inform him." Later philosophers have tried to explain what is involved in the sensation of solidity. Étienne Bonnot de Condillac distinguished it from the sensations of sound, color, and smell, since a person knows his own body by it. If a person presses his hand against his chest, his hand and chest "will be distinguished from one another by the sensation of solidities which they mutually give each other." Thus, involved in the notion of a sensation of solidity is the notion of the recognition as such of a feeling given to a part of the body. If organic sensations were not localized in the body, a person could never know his own or any other body by touching it, for "it is only with extension that we can construct extension, just as it is only with objects that we can construct objects."

H. H. PRICE

H. H. Price carried the analysis a step further. He divided touch "into three distinct types of sensation: contact sensation proper, muscular sensation, and the sensation of temperature." The perception of solidity involves both contact sensation proper and muscular sensation. The latter is "essentially a modification of the voluminous life-feeling [that] might also be described as our sense of embodiment." Muscular strain is felt at a place in the body and as having vectorial character, that is, originating from or tending toward a certain direction. A person experiences the solidity of something when the resistance he feels on pressing it "is actually felt as coming from within the closed boundary which contact-sensation reveals.… Thus the tactual conception of Matter is strictly speaking tactuo-muscular or contactuo-muscular."

LOCAL SIGN THEORY

The analyses of both Condillac and Price specify organic sensations as being localized. As Condillac expressed it, to know its body the child must "perceive its sensations, not as modifications of its soul, but as modifications of the organs which are their occasional causes." Condillac cannot explain "how the self which is only in the soul appears to be found in the body … it is enough that we observe this fact." The alternatives are either that a person is born with the capacity to locate organic sensations or that he acquires this capacity. Most philosophers hold the capacity to be acquired, although they differ widely in the accounts they give of how it is acquired; whether by the person's learning to interpret some feature of the sensa-

tion as a sign of its location (the so-called local sign) or in some other way.

MOVEMENT AND TOUCH

Perhaps the most important recent contribution to the problem of how touch mediates awareness of its objects was made by David Katz in "Der Aufbau der Tastwelt." Summarizing Katz's conclusions, Maurice Merleau-Ponty expresses the crux of the matter as being that "the movement of one's body is to touch what lighting is to vision.… When one of my hands touches the other, the hand that moves functions as subject and the other as object. There are tactile phenomena, alleged tactile qualities, like roughness and smoothness, which disappear completely if the exploratory movement is eliminated. Movement and time are not only an objective condition of knowing touch, but a phenomenal component of tactile data. They bring about the patterning of tactile phenomena, just as light shows up the configuration of a visible surface."

BODY-OBJECT RELATION

With the view that the objects of touch are physical objects may be contrasted the view that we are not aware of the object we touch but of a relation holding between our body and that object. It is a fact that how warm an object feels to an observer depends causally on the warmth of the part of the observer's body with which he is touching it. We notice the temperature of a hand that is colder or warmer than our own. Aristotle explains this in terms of his theory of sensation as the assimilation in form of the organ to the object. D. M. Armstrong mentions it, together with the fact that a person can say immediately with what portion of his body he is in contact with an object perceived by touch, in support of his theory that all immediate tactual perception involves perception of a relation holding between the observer's body and the object he is touching. As evidence for his theory, Armstrong holds that "hardness and softness as immediately perceived by touch, are obviously relative to the hardness or softness of our flesh." It is unclear from this evidence whether Armstrong is justified in claiming more than that how things feel to us depends on the condition of the part of the body with which we feel them.

See also Aristotle; Armstrong, David M.; Berkeley, George; Colors; Condillac, Étienne Bonnot de; Locke, John; Merleau-Ponty, Maurice; Sensa; Sound.

Bibliography

Aristotle. *De Anima*. Book II, Ch. 11.

Armstrong, D. M. *Bodily Sensations*. London and New York: Routledge and Paul, 1962. Chs. 3–5.

Berkeley, George. *An Essay towards a New Theory of Vision*. Dublin, 1709. Secs. 45–49.

Condillac, E. B. de. *Treatise on the Sensations*. Paris and London, 1754. Part II, Chs. 1–5.

Grice, H. P. "Some Remarks about the Senses." In *Analytic Philosophy.* 1st series. Oxford: Blackwell, 1962.

Katz, David. "Der Aufbau der Tastwelt." *Zeitschrift für Psychologie* 11 (1925).

Locke, John. *An Essay concerning Human Understanding*. London, 1690. Book II, Ch. 4.

Merleau-Ponty, Maurice. *Phenomenology of Perception*, 313–317. New York: Humanities Press, 1962.

O'Shaughnessy, B. "The Sense of Touch." *Australasian Journal of Philosophy* 67 (1989): 37–58.

Perkins, M. *Sensing the World*. Indianapolis: Hackett, 1983.

Price, H. H. "Touch and Organic Sensation." *PAS* 44 (1943–1944).

Urmson, J. O. "The Objects of the Five Senses." *Proceedings of the British Academy* 54 (1968): 117–131.

Vesey, G. N. A. *The Embodied Mind*. London: Allen and Unwin, 1965. Ch. 4. Contains an exposition and criticism of the local sign theory and references to other work on the subject.

Warnock, G. J. *Berkeley*. Harmondsworth, U.K., 1953. Ch. 3.

G. N. A. Vesey (1967)
Bibliography updated by Benjamin Fiedor (2005)

TOYNBEE, ARNOLD JOSEPH
(1889–1975)

Arnold Joseph Toynbee was in the twentieth century the foremost contemporary representative of what is sometimes termed "speculative philosophy of history." In some respects he occupied a position analogous to that of Henry Thomas Buckle in the nineteenth century. Like Buckle, he sought to discover laws determining the growth and evolution of civilization and to do so within the context of a wide comparative survey of different historical societies; like Buckle again, the results of his investigation became a storm center of controversy and criticism. To support his hypotheses, Toynbee, however, was able to draw on a vast fund of material of a kind unavailable to his Victorian predecessor, and the imposing examples and illustrations in which his work abounds make Buckle's much-vaunted erudition look strangely threadbare. As a consequence, Toynbee's historical theory is worked out in far greater detail; in fact, it represents a highly articulated and complex structure with many ram-

ifications and appendages. Moreover, the materialist optimism underlying Buckle's linear conception of history as a continuous progressive development is wholly absent from Toynbee's analysis of the rise and decay of different cultures, while, in place of Buckle's positivistic rationalism, there runs through all Toynbee's work, especially his later books, a strain of mysticism and religious idealism.

Toynbee was educated at Balliol College, Oxford, and was a tutor there from 1912 to 1915. Subsequently, he became professor of Byzantine and modern Greek language, literature, and history at London University (1919–1924) and then for thirty years held the post of director of studies in the Royal Institute of International Affairs. He wrote on a wide variety of topics concerning Greek history, international politics, and contemporary affairs, but his main work was his *A Study of History*, the first ten volumes of which were published between 1934 and 1954. As of 1967, two other volumes appeared, the last, titled *Reconsiderations*, being largely an attempt to meet points raised by his numerous critics and, where he has thought it necessary, to qualify previous claims in the light of their objections. Toynbee always listened carefully to those who have disagreed with him, although he has apparently never felt that their observations justified any major revision of his views.

A STUDY OF HISTORY

Toynbee claimed that his project was first suggested to him when, at the beginning of World War I, he became aware of certain striking affinities between the courses taken by the Greco-Roman and modern European civilizations. It occurred to him that similar parallels might be discernible elsewhere, that there is, as he puts it, "a species of human society that we label 'civilisations'" and that the representatives of this species which have thus far appeared on this planet may exemplify in their various histories a common pattern of development. With this idea forming in his mind, Toynbee came across Oswald Spengler's *Decline of the West*, in which he found many of his own intimations affirmed and corroborated. Nevertheless, it seemed to Toynbee that Spengler's account was defective in important ways. The number of civilizations examined (eight) was too small to serve as a basis for safe generalization; little attempt was made to explain why cultures rise and decline in the manner described; and, in general, Spengler's procedure was marred by certain a priori dogmas that distorted his thinking, leading him to display at times a cavalier disregard for the facts. What was required was a more empirical approach, one in which it was clearly recognized that a problem of explanation existed and that the solution of this problem must be in terms of verifiable hypotheses that can stand the test of historical experience.

THE PATTERN OF HISTORY. Toynbee repeatedly referred to his own method as essentially "inductive." His aim (initially, at least) was to "try out the scientific approach to human affairs and to test how far it will carry us." In undertaking this program, he was insistent upon the need to treat as the fundamental units of study "whole societies," as opposed to "arbitrarily insulated fragments of them like the nation-states of the modern West." In contrast with Spengler, he claimed to have identified twenty-one examples (past and present) of the species "civilization," though he admitted that even this number is inconveniently small for his purpose—"the elucidation and formulation of laws." He argued, however, that a significant degree of similarity is discernible between the careers of the societies he examined and compared; certain stages in their respective histories can be seen to conform to a recognizable pattern too striking to be ignored, a pattern of growth, breakdown, and eventual decay and dissolution. Within this pattern certain recurrent "rhythms" may be detected.

When a society is in a period of growth, it offers effective and fruitful responses to the challenges that present themselves; when in decline, on the other hand, it proves incapable of exploiting the opportunities and of withstanding or overcoming the difficulties with which it is confronted. Neither growth nor disintegration, Toynbee holds, is necessarily continuous or uninterrupted. In disintegration, for instance, a phase of rout is frequently succeeded by a temporary rally, followed in turn by a new, more serious relapse. As an example he cited the establishment of a universal state under the Augustan Pax Romana as a period of rally in the career of the Hellenic civilization, coming between a time of troubles which, in the form of revolutions and internecine wars, preceded it and the first stages of the Roman Empire's final collapse, which followed in the third century. Toynbee contended that clearly comparable rout-rally rhythms have manifested themselves in the disintegration of many other civilizations, such as the Chinese, the Sumerian, and the Hindu. In these, too, we encounter the phenomena of increasing standardization and loss of creativity that were apparent when the Greco-Roman society was in decline.

HISTORICAL MODELS. Toynbee's tendency to interpret the history of other civilizations in terms suggested by that of the Hellenic culture is marked, and many of his opponents have claimed that it has led him into imposing

artificial schemes upon the past and into postulating parallels by no means borne out by the historical material. In his most recent work Toynbee has shown himself to be sensitive to criticism of this kind. He has maintained, however, that for an investigation of the kind he envisaged it was at least essential to start with a model of some sort, his chief doubts being whether the model he chose was ideally suited to his purpose and whether a future student of the comparative history of civilizations would not be better advised to employ a diversity of specimens, rather than a single example, to guide his inquiries.

However, it is not clear that in proposing this amendment to his original procedure, Toynbee has fully appreciated the principal points at issue. He still seems to be searching for some single pattern of interpretation to which the histories of particular societies can be seen to stand as specimen cases, and in so doing, he overlooks two considerations, both of which have been stressed by various critics.

First, he continues to leave obscure the question of how the identity of a given civilization is to be determined. This is by no means a trivial matter, since in his practice Toynbee has often given the impression of identifying civilizations by reference to the very principles of development that in other places he has claimed to have elicited purely through an empirical survey of their actual careers. He thereby exposes himself to the charge of treating as factual discoveries what are no more than disguised tautologies.

Second, it has been argued that insofar as the term suggests an explanatory device capable of rendering intelligible a certain range of phenomena, Toynbee's references to models in the context cited are misleading. To maintain that a number of other societies have tended to follow a path significantly similar to the course taken by a selected specimen is by itself to explain nothing; at best, it is to point out that there is something *requiring* explanation—namely, the existence of the similarities in question. But although such an objection has force, Toynbee has, in fact, attempted to account for the correlations he believes himself to have discovered. He is not, as some have alleged, content simply to enumerate like instances and has always taken the problem of seeking explanations seriously. Thus, when trying to account for the disintegration of civilizations, he has invoked such notions as the "intractability of institutions" and the "nemesis of creativity," as well as pointing to the development of "internal" and "external" proletariats and of "dominant," as opposed to "creative," minorities.

Whether the explanations he has sought to provide are plausible or convincing is, of course, another matter. Frequently, they seem to involve an appeal to laws too vague to afford adequate support, and at other times Toynbee enlists the services of highly dubious or irrelevant analogies. He also tends to treat literary or folk myths as if they in some way gave evidential backing to his generalizations.

ORDER OR CHAOS. In defending his position, Toynbee has frequently attacked what he calls "antinomian historians," upholders of "the dogma that in history no pattern of any kind is to be found." He has argued that to deny the existence of patterns is implicitly to deny the possibility of writing history, for patterns are presupposed by the whole system of concepts and categories a historian must use if he is to talk meaningfully about the past.

But patterns of what sort? Toynbee sometimes implies that it is essential to choose between two fundamentally opposed views. Either history as a whole conforms to or manifests some unitary order and design, or else it is a "chaotic, disorderly, fortuitous flux" which defies intelligible interpretation. As examples of the first he cites the "Indo-Hellenic" conception of history as "a cyclic movement governed by an Impersonal Law" and the "Judeo-Zoroastrian" conception of it as governed by a supernatural intellect and will. A combination of these ideas appears to underlie Toynbee's own picture of the human past as it finally emerges in *A Study of History*, particularly in the later volumes, where the suggestion that the rise and fall of civilizations may be susceptible to a teleological interpretation is explicitly put forward.

It would seem, however, that Toynbee has posed his dilemma in altogether too simple terms. There are a number of familiar ways in which historians may be said to reduce the material of history to order and coherence, none of which involves the acceptance of all-embracing beliefs regarding the historical process as a whole of the type he instances. Of course, if the notion of the intelligibility of the past is initially defined in a manner that presupposes the validity of such beliefs, it is possible to accuse historians who deny that it is necessary or even legitimate to adopt them of making nonsense of their subject. But why, it may be asked, should such a stipulation be accepted?

REPUDIATION OF OLDER SCHEMES. In fact, Toynbee does not really intend to advance so exclusive a claim. He does not deny that historians may be able to make sense

of particular segments of human history without being committed to universalistic positions of the sort mentioned, imperfect and incomplete though such explanations must ultimately be judged to be. He does, however, strongly suggest that the piecemeal approaches and categories of traditional history leave much to be desired, applying to them such terms as *archaic, infantile*, and *crude*. Here, possibly, lies the true source of his objections to "antinomianism." He wishes to condemn the old structures and clichés, the worn axioms unconsciously assumed in conventional historical thought. In particular, he is critical of the lines along which historians have been prone to cut up the past, both geographically and temporally. He distrusts the artificial cohesion they have projected into certain periods through the use of comprehensive simplifying labels like "the Renaissance" and "the Middle Ages," and he questions the unity and self-sufficiency implicit in their conception of "European history."

It is, of course, perfectly acceptable to appraise and seek to revise the conceptual schemes of previous historians in the light of fresh empirical knowledge and discoveries, but it is quite another thing to propound a general theory of historical development which appears in its final form to rely heavily upon extrahistorical considerations and preconceptions of a metaphysical or religious kind. Toynbee has perhaps never sufficiently appreciated the force of this distinction; even so, it would be churlish not to recognize the imaginative fertility, the sheer inventiveness, which is so marked a feature of his system, whatever its shortcomings in other respects. *A Study of History* is rich in methodological suggestions and contains a profusion of original interpretative concepts and frameworks. Whether any of these will be found of value by future historians or social scientists remains to be seen.

FREEDOM AND LAW IN HISTORY

A word may be said about Toynbee's views regarding the future of Western civilization and their relation to his general theory. He frequently speaks as if Western society were in an advanced state of breakdown; at the same time he repeatedly shows himself unwilling to draw the conclusion that it is in fact doomed to final disintegration, and he speaks of the possibility of a "reprieve" granted by God. The "determinism" implicit in his thought when he is seeking to apply "the scientific approach to human affairs" tends thus to conflict with the "libertarian" principles to which he claims to subscribe when discussing the nature of human actions and which are connected with his own metaphysical and religious beliefs. The later volumes of the *Study* display a persistent uneasiness over this apparent contradiction, yet it cannot be said that the efforts he has made in these volumes to reconcile the roles of law and freedom in history have proved satisfactory. Rather, they serve to highlight the logical difficulties that had already revealed themselves at earlier stages in Toynbee's work.

See also Determinism and Freedom; Libertarianism; Philosophy of History; Spengler, Oswald.

Bibliography

WORKS BY TOYNBEE

A Study of History. 12 vols. London: Oxford University Press, 1934–1961.

A Study of History: Abridgment of Volumes I–X, edited by D. C. Somervell. Oxford, Vols. I–VI, 1946; Vols. VII–X, 1957.

Civilisation on Trial. London: Oxford University Press, 1948.

An Historian's Approach to Religion. London: Oxford University Press, 1956.

Experiences. London, New York: Oxford University Press, 1969.

With G. R. Urban. *Toynbee on Toynbee: A Conversation between Arnold J. Toynbee and G.R. Urban*. New York: Oxford University Press, 1974.

Arnold Toynbee, a Selection from His Works. Edited by E. W. F. Tomlin. Oxford, U.K.; New York: Oxford University Press, 1978.

WORKS ON TOYNBEE

Dray, W. H. "Toynbee's Search for Historical Laws." *History and Theory* 1 (1960): 32–54.

Geyl, Pieter. *Debates with Historians*. London, 1955.

McNeill, William Hardy. *Arnold J. Toynbee, a Life*. New York: Oxford University Press, 1989.

Montagu, Ashley, ed. *Toynbee and History: Critical Essays and Reviews*. Boston: Sargent, 1956.

Morton, S. Fiona. *A Bibliography of Arnold J. Toynbee*. Oxford; New York: Oxford University Press, 1980.

Ortega y Gasset, José. *An Interpretation of Universal History*. New York: Norton, 1973.

Samuel, Maurice. *The Professor and the Fossil: Some Observations on Arnold J. Toynbee's "A Study of History."* New York: Knopf, 1956. A criticism of Toynbee's views on the Jews.

Trevor-Roper, H. R. "Arnold Toynbee's Millenium [*sic*]." In his *Men and Events*. New York: Harper, 1957.

Walsh, W. H. "Toynbee Reconsidered." *Philosophy* 38 (1963).

Patrick Gardiner (1967)

Bibliography updated by Michael J. Farmer (2005)

TRACY, DESTUTT DE

See *Destutt de Tracy, Antoine Louis Claude, Comte*

TRADITIONALISM

"Traditionalism" was a philosophy of history and a political program developed by the Counterrevolutionists in France. It was ultramontane in politics and anti-individualistic in epistemology and ethics.

It was the common belief of both those who favored the French Revolution and those who opposed it that the revolution was prepared by the *philosophes.* Voltaire and Jean-Jacques Rousseau were invoked by both parties as having been either the initiators of much-needed reforms or the corrupters of youth. The intellectual differences among the *philosophes* were minimized. The Revolutionary Party believed that Voltaire and Rousseau were the leaders of two schools of thought, both of which removed the seat of authority from the group—society or the nation or the church—to the individual, and that the two schools disagreed only on the question of whether authority was vested in the reason or in feeling (*sentiment*). The Voltairians were said to be individualistic rationalists; the Rousseauists individualistic sentimentalists. In short, the Voltairians were supposed to believe that any individual, by the use of reason alone, could reach all attainable truth in any field; the Rousseauists, that one had only to look into his "heart" to achieve the same result. Oversimplified as this was as history, it was common belief.

The philosophy of Comte Joseph de Maistre (1753–1821) and of his alter ego, Vicomte de Bonald (1754–1840), was developed in conscious reaction to individualism. De Maistre and Bonald were rationalistic, but they maintained that the reason to be trusted was that of the group, not that of the individual. The common reason, like the common sense, was lodged in a superindividual being, manifested in tradition and expressed in language. The superindividual being was the Roman Catholic Church, the authority of which was binding not only on its avowed members, but on all people. The church alone had direct access to the source of truth (God) and for 1,800 years had remained steadfast and unshaken in its dogmas. Since truth must be one and everlasting, the traditionalists were persuaded by a simple conversion of the proposition that where there was a single and everlasting set of ideas, it must be true. "No human institution has lasted eighteen centuries," de Maistre wrote in *Du Pape* (3 vols., Lyons, 1821). Therefore, he inferred, the church must be superhuman or divine.

Human nature can be understood only by seeing humanity as an integral part of the church. The human individual is but a fragment of a whole. He is completely dependent on society for his bodily welfare and even for his thoughts, for his thoughts are internal speech, and no language is either that of a single individual or created by an individual. Combating the theory that language was invented, de Maistre argued, as Rousseau and Thomas Reid had done before him, that thought is required for invention and language must therefore have existed before it could be invented. Language is the thought of the race expressing itself. It is also rational—we cannot express emotions and sensations linguistically. We speak our thoughts; we speak *of* our feelings and emotions. Since the traditionalists were French, they turned to the history of France for their evidence and found it in the antiquity of the Capetian dynasty, founded, in their view, by Louis the Pious in the ninth century, if not by Charlemagne; in the genesis of French from Latin; and in the primacy of Catholicism in France, which was converted from paganism by Dionysius the Areopagite, the first pagan to be converted by St. Paul.

The supremacy of the pope in both religious and secular affairs was emphasized by de Maistre. Although there might be two swords, the spiritual and the temporal, the latter was wielded, in the language of Boniface VIII, at the pleasure and sufferance of the priest (*ad nutum et patientiam sacerdotis*). This factor of the traditionalists' teachings led to ultramontanism, which, when vigorously preached by Hugues Félicité Robert de Lamennais (1782–1854) in the nineteenth century, was condemned by the pope.

Another type of traditionalism was espoused by Pierre-Simon Ballanche (1776–1847). In his major work, *Palingénésie sociale* (1827–1829), Ballanche developed a philosophy of history based on man's fall from primordial innocence. However, he maintained that there could be steady progress toward universal rehabilitation. In upholding the possibility of human progress, Ballanche differed from Bonald and de Maistre, for whom time and change, variety and multiplicity, were inherently evil. To Ballanche they were the only condition of redemption. He was convinced of the ultimate perfection of humankind, at which time all that is potential in the human essence would be realized. All men were to be rehabilitated, regardless of their present merits. There was no eternal hell. Even religion would progress, in that God would reveal its truths bit by bit as humankind became worthy of receiving them. Each man would have to make himself worthy by listening to his heart, an appeal to personal interpretation that was considered heretical.

Although Ballanche agreed with Bonald and de Maistre that the understanding of history could come only from seeing the designs of God in every historical event, he did not believe that government should be theocratic. On the contrary, the two swords must be wielded by two separate powers. The secular power, however, should not be in the hands of the people; they should be permitted to voice their aspirations only so that the sovereign might accept them.

It remained for Lamennais to carry traditionalism to its logical conclusion. Beginning with the strictest form of ultramontanism, he developed into a heresiarch, never realizing that he was moving away from the course of reason. If the pope was the head of the church and the church was superior to the state, then the pope should be recognized as the one sovereign and autonomous being on Earth. The sole test of certitude, Lamennais maintained, lay in the racial reason, and this collective reason was tradition. Tradition gives society its unity, and its unity fosters civilization. However, society to Lamennais was not France; it was humanity. And since civilization was Catholicism, national boundaries were artificial and should be eliminated except for practical purposes. The common sense of humankind, in which he believed as did the Stoics, was nothing that could be substantiated by the reason. It was the reason. One must submit to tradition in order to avoid the divisive effects of sectarianism. When the state put obstacles in the way of such submission, then rebellion was legitimate. However, this involved freedom of conscience, of the press, and of education, if it was to be practiced. It was at this point that Gregory XVI in his encyclical *Mirari Vos* intervened to silence Lamennais.

Traditionalism as a body of doctrine was condemned in 1855 in a decretal against Augustine Bonnetty (1798–1879), a priest. The theory directly condemned was the *fidéisme* of the Abbé Bautain (1796–1867), which Bautain had retracted in 1840. Since the identity of reason, common sense, and tradition demanded prerational assertions, faith seemed to be the only thing left to which the traditionalist might appeal. However, this raised faith to a position above that of reason, contrary to the doctrine of the church. The rationalistic position of the church was confirmed at the third session of the Vatican Council in 1870.

See also Bonald, Louis Gabriel Ambroise, Vicomte De; Conservatism; Lamennais, Hugues Félicité Robert de; Maistre, Comte Joseph de; Philosophy of History; Reid, Thomas; Rousseau, Jean-Jacques; Voltaire, François-Marie Arouet de.

Bibliography

Ballanche, Pierre Simon. *Oeuvres complètes.* 4 vols. Paris, 1830. See especially *Palingénésie sociale* (first published in 2 vols., Paris, 1827–1829); *Le vieillard et le jeune homme* (first published Paris, 1819); *Vision d'Hébal* (first published Paris, 1831).

Berlin, Isaiah. "Joseph de Maistre and the Origins of Fascism." In his *The Crooked Timber of Humanity,* edited by Henry Hardy. New York: Knopf, 1991.

Boas, George. *French Philosophies of the Romantic Period.* Baltimore: Johns Hopkins Press, 1925. Ch. 3.

Denziger, Heinrich. *Enchiridion Symbolorum et Definitionum,* 13th ed. Freiburg, 1921. No. 1613 et seq.; No. 1649 et seq.

Ferraz, Marin. *Histoire de la philosophie en France au XIX^{e} siècle: Traditionalisme et ultramontanisme.* Paris: Didier, 1880.

Gunn, J. Alexander. *Modern French Philosophy.* London: Unwin, 1922. Ch. 7.

Klinck, David. *The French Counterrevolutionary Theorist Louis de Bonald (1754–1840).* New York: Peter Lang, 1996.

Laski, Harold. *Authority in the Modern State.* New Haven, CT: Yale University Press, 1919. On the condemnation of traditionalism.

Lebrun, Richard. *Joseph de Maistre: An Intellectual Militant.* Kingston: McGill-Queen's University Press, 1988.

Reedy, W. Jay. "The Traditionalist Critique of Individualism in Post-Revolutionary France: The Case of Louis de Bonald." *History of Political Thought* 16 (1) (1995): 49–75.

Stearns, Peter N. *Priest and Revolutionary: Lamennais and the Dilemma of French Catholicism.* New York: Harper and Row, 1967.

George Boas (1967)
Bibliography updated by Philip Reed (2005)

TRAGEDY

The two main strands in the history of philosophical reflection on tragedy, as a genre of art, can both be seen as having their origins in Plato's critique of tragic poetry in the *Republic* and other dialogues. It is there that we find their first sustained philosophical treatment; and with respect to this small part of it, at least, Alfred North Whitehead's characterization of the history of philosophy as a series of footnotes to Plato is not too fanciful.

TRAGEDY AND EMOTION

One strand of thought focuses on the character and value of our experience of tragedy, and can be seen in Plato's charge that tragedy (and indeed mimetic poetry in general) "gratifies and indulges the instinctive desires … with its hunger for tears and for an uninhibited indulgence in grief"; that "it waters [passions] when they ought to be allowed to wither, and makes them control us when we ought, in the interests of our own greater welfare and

happiness, to control them" (1987, 606a). Plato's thought that the emotional dimension of our experience of tragedy is particularly significant has been taken up in a variety of directions by other philosophers.

In the *Poetics*, Aristotle argued that tragedy's capacity to arouse the emotions of pity and fear in its audience, so far from rendering it intellectually and morally damaging, is in fact a source of its value: Tragedy aims at emotional effect not for its own sake, or for the sake of gratifying or indulging its audience, he argued, but rather in such a way as to bring about a *catharsis* of the tragic emotions. Precisely what Aristotle meant by *catharsis* is far from clear, and has been the topic of much scholarly debate: The notion has been understood in terms of purgation (of excessive or pathological emotion), of purification, and of intellectual clarification, to mention only some of the most influential of the interpretations that have been offered. Whatever its precise meaning may be, however, it is clear that Aristotle took *catharsis* to be a process or experience that in one way or another is conducive to emotional health or balance, such that our emotional experience of (well-written) tragedy is not indulgently sentimental and opposed to "our better nature," as Plato argued, but is rather an essential element in a fully comprehending attitude to what a work depicts.

Aristotle linked *catharsis* with the pleasure that we take in tragedy: The fact that mention of the former comes at the end of his definition of tragedy suggests that he takes it to be in some sense the goal of works of this sort, and (an appropriate form of) the latter is said to be "what the poet should seek to produce." His defense of the value of our emotional experience of tragedy in terms of *catharsis* is thus at least implicitly a defense of it in terms of tragic pleasure; and a debate related to, and at least as extensive as that concerning the meaning of "*catharsis*," has its origins in his characterization of tragic pleasure as "the pleasure derived from pity and fear by means of imitation [mimesis]" (1967, 1453b). For how is it that one can derive pleasure from what Aristotle himself describes elsewhere (notably in the *Rhetoric*) as painful feelings? This question is a more difficult relative of one prompted by Plato's reference to the fact that "when we hear Homer or one of the tragic poets representing the sufferings of a hero and making him bewail them at length … even the best of us enjoy it" (1987, 605c-d): How is it that in engaging with a work of tragedy one is able, or is enabled by the work, to enjoy the depiction of human suffering?

Debate surrounding these and related questions was particularly prevalent in eighteenth-century British philosophy and criticism, attracting contributions from such figures as Lord Kames, James Beattie, and Joseph Priestley, as well as, more influentially, David Hume, Adam Smith, and Edmund Burke. Some contributors to the debate focus on the question of how one can respond with pleasure to what tragedy depicts: Edmund Burke, for example, in his *A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and the Beautiful*, took the problem to lie in the "common observation" "that objects which in the reality would shock, are in tragical, and such like representations, the source of a very high species of pleasure" (1990, p. 41), and thus in effect construed the problem as one concerning the consistency of one's patterns of response. (As, in a sense, did Plato, though he took the inconsistency between our responses to depictions of suffering in tragedy and our responses to suffering "in reality" to lie not in the fact that the former involve pleasure and the latter "shock" or horror, but rather in that in the former we give vent to our emotions whereas in the latter we strive "to bear them in silence like men.")

Discussions that remain exclusively occupied with the pleasure that Plato holds that one takes in what tragedy depicts often proceed by attempting to resolve the apparent inconsistency in one's patterns of response by pointing to relevant differences between the contexts in question: for example, one's awareness of the fictional status of tragedy, the contribution of artistry, and "aesthetic distance" have all been cited as aspects of our experience of tragedy that are not involved in our experience of actual suffering, the functioning of which explains why pleasure is a characteristic element of the former while typically absent from the latter. However, such discussions risk missing the more difficult issue that arises from Aristotle's characterization of tragic pleasure. For if that characterization is right, the peculiarity of the latter is not simply that it occurs in response to the depiction of things that in other contexts do not give one pleasure, but rather that it is a variety of pleasure that is intimately bound up with painful feeling; as he put it, it is the pleasure "of," or "derived from," such feeling.

The more sophisticated treatments of our emotional experience of tragedy have attempted to address this. Burke, for example, suggested that the apparent inconsistency between one's responses to tragedy and one's responses to actual suffering is illusory; in fact, he held, we are just as disposed to take pleasure in actual sufferings as we are in depictions of suffering, and in both cases our response is based on sympathy, a psychological mechanism that involves pain at the distress of its objects,

but also (in order to foster its occurrence) pleasure: "as our Creator has designed we should be united by the bond of sympathy, he has strengthened that bond by a proportional delight" (1990, p. 42). Adam Smith made a similar point when he argued that it is because of its social utility that the experience of sympathy, even when the emotions communicated sympathetically are painful, is naturally pleasurable to human beings.

This account of the matter, though clearly based on a Humean theory of the passions, was rejected by Hume himself, on the grounds that the operation of sympathy is not always pleasurable: If it were, he suggested in a letter to Smith, "an hospital would be a more entertaining place than a ball." (A point anticipated in its spirit if not its tone, by Burke, who suggested that people do indeed find public executions more compelling than "the most sublime and affecting tragedy we have.") Hume's own account of what he described as the seemingly "unaccountable pleasure which the spectators of a well-written tragedy receive from sorrow, terror, anxiety, and other passions that are in themselves disagreeable and uneasy" 1987, p. 216) is by far the most discussed by contemporary contributors to the debate, although it is more interesting as an application of his theory of the passions than it is as an account of our experience of tragedy.

Hume suggested that the spectators' pleasure and their "disagreeable and uneasy" emotions are initially responses to different aspects of a work of tragedy: their distress is a response to what the work depicts, their pleasure a response to the "eloquence" and "genius" with which it depicts it. To leave the matter at that would clearly miss the problem posed by Aristotle's characterization of tragic pleasure. But Hume went on to argue that these responses merge, as the pleasure, which is dominant, overpowers, and somehow "converts" the distress in such as way as to reinforce the former: "The impulse or vehemence, arising from sorrow, compassion, indignation, receives a new direction from the sentiments of beauty. The latter, being the predominant emotion, seize the whole mind, and convert the former into themselves, at least tincture them so strongly as totally to alter their nature" (1987, p. 220). Contemporary discussions of Hume's account have focused on just what this "conversion" of emotion is supposed to involve, for Hume himself was less than clear on the matter. Whatever it does amount to, however, it is clearly dependent on Hume's associationist psychology, and is unlikely to survive the rejection of this.

Philosophical discussion of tragic pleasure, or what scholars often refer to as "the paradox of tragedy," has continued on very much the lines established by eighteenth-century thinkers, though a new slant on the matter (and indeed on the nature of catharsis) has been introduced by philosophers and others influenced by the methods and findings of psychoanalytic theory. It remains a recurring theme in contemporary philosophy of art.

THE PROFUNDITY OF TRAGEDY The second major strand in the history of the philosophy of tragedy is represented in Plato's discussion of the epistemic credentials of tragic poetry, so to speak, where he argued that the tragedian has neither knowledge nor true belief concerning that of which he writes, and (hence) that tragedy cannot be a source of knowledge. Plato's target here is the view that "the tragedians … are masters of all forms of skill, and know all about human excellence and defect and about religion" (1987, 598d-e), or more broadly the thought that tragedy's distinctiveness has to do with its capacity to prompt, and to suggest authoritative answers to, questions of a distinctively ethical sort. Despite Plato's efforts, the appeal of this line of thought survived his critique, not least due to the support that some found for it in Aristotle's claim that "poetry is a more philosophical and more serious business than history" (1987, 1451b), a claim made in the context of his attempt to show that the tragedian's art is, despite Plato's arguments to the contrary, a *technê*, a productive activity that employs rational means or principles in the pursuit of a predetermined practical end. The thought that tragedy is an especially philosophical form of art received its most sustained treatment in nineteenth-century German philosophy and criticism, where versions of it were expounded by Gotthold Lessing, Friedrich Schiller, Friedrich Schlegel, August Wilhelm Schlegel, and Johann Goethe, as well as, and from a philosophical point of view more notably, by Georg Hegel, Arthur Schopenhauer, and Friedrich Nietzsche.

Hegel argued that the business of Classical tragedy—its "essential basis"—is to demonstrate "the validity of the substance and necessity of ethical life" (1975, Vol. 2, p. 1222). It achieves this first by showing the "collision" between different aspects of the ethical that occurs when the latter is fragmented and particularised in human social life: thus he claimed that Sophocles' *Antigone* dramatizes the collision between the authority of the state (represented by Creon) and family love (represented by Antigone). These aspects of ethical life collide because "each of the opposed sides, if taken by itself, has justification; while each can establish the true and positive content of its own aim and character only by denying and

infringing the equally justified power of the other" (1975, Vol. 2, p. 1196). The task of tragedy is then to show the "resolution" of conflict of this sort, which it can do in a variety of ways. The most satisfying form of resolution, Hegel claimed, involves the destruction of the characters who embody "false one-sidedness," as happens in *Antigone*, but "the unity and harmony of the entire ethical order" may also be effected and exemplified by the surrender of the hero (as in *Oedipus the King*), the reconciliation of opposing interests (as in the *Eumenides*), or "an inner reconciliation" in the tragic hero himself (as in *Oedipus at Colonus*).

Although he held that tragedy was at its most beautiful in the classical period, Hegel argued that it is in what he called Romantic tragedy that art is at its most philosophical, or, in his terms, comes closest to "bringing to our minds and expressing the Divine, the deepest interests of mankind, and the most comprehensive truths of spirit" (1975, Vol. 1, p. 7). The subject matter of tragedy by this stage of its development is "the subjective inner life of the character," and at its best, which Hegel thought was in Shakespeare's hands, these characters are "concretely human individuals," "free artists of their own selves" (Vol. 2, pp. 1227–1228). Tragedy at this stage represents not collision between particularised ethical powers, as did classical tragedy, but either (and, Hegel claimed, unsatisfactorily) collision between different aspects of a character's personality, or (in what he held are the finest examples of Romantic tragedy) between the character and external circumstances. Tragedy of the latter sort presents the "progress and history of a great soul, its inner development, the picture of its self-destructive struggle against circumstances, events, and their consequences" (Vol. 2, p. 1230).

Hegel's claim that the importance of tragedy lies in its capacity to reveal important truths about the human condition is echoed by Schopenhauer. Indeed, like Hegel, Schopenhauer saw the arts in general as engaged fundamentally in the same task as philosophy; both, as he said, "work at bottom towards the solution of the problem of existence" (1969, Vol. 2, p. 406). Tragedy, Schopenhauer held, is "the summit of poetic art," for in dramatising "the terrible side of life … the unspeakable pain, the wretchedness and misery of mankind, the triumph of wickedness, the scornful mastery of chance, and the irretrievable fall of the just and the innocent," tragedy reveals to us more clearly than anything else the most important feature of reality: "the antagonism of the will with itself" and the fact that "chance and error" are "the rulers of the world" (1969, Vol. 1, pp. 252–253). However, in Schopenhauer's view tragedy is significant not merely because of the importance of what it reveals to us concerning the nature of reality, but also because in the experience of tragedy one may come to recognize the only appropriate response to the terrible truth it presents. This is to adopt an attitude of "resignation": as Schopenhauer put it, "The horrors on the stage hold up to [the spectator] the bitterness and worthlessness of life, and so the vanity of all its efforts and endeavours. The effect of this impression must be that he becomes aware … that it is better to tear his heart away from life, to turn his willing away from it, not to love the world and life" (Vol. 2, p. 435) The greatest tragedies, Schopenhauer said, are those in which this attitude of resignation is not only suggested by a work but also demonstrated by its characters.

If Schopenhauer was less concerned with particular works of tragedy than Hegel, Nietzsche was still less so. In *The Birth of Tragedy*, his infrequent references to particular works of Greek tragedy betray very little of the knowledge of this part of literary history that he surely had; and the Aeschylus, Sophocles, and Euripides whom he discussed in that work figure not as artists in a history of a genre of art, but rather as symbols or personifications of different cultural points or tendencies in Nietzsche's working out of a genealogy of the tragic spirit. The main symbols in this genealogy are those of Dionysus and Apollo, Greek deities whom Nietzsche used creatively to stand for both metaphysical and artistic categories. The Apollonian spirit is that which is concerned with appearances, with the world as composed of individuals; what it offers us is "beautiful illusion" (1993, p.15). The Dionysian spirit is that through which this illusion is shattered, and what is revealed to us reality as it truly is: an endless and pointless struggle of things in flux. As its objects are illusory, the Apollonian vision is too fragile to sustain human beings indefinitely. But with its object of what Nietzsche described as a "witch's brew" of "lust and cruelty" (p. 19) the Dionysian vision is too terrible for human beings to survive. The "supreme goal" of art, Nietzsche claimed, is to allow us to escape this dichotomy.

Art, at its highest, does not attempt to evade the Dionysian truth but rather, by somehow (and in a way that Nietzsche is never very clear about) mediating it through the Apollonian, renders it bearable and even something to be exulted in. Nietzsche suggested that the tragedies of Aeschylus and Sophocles, in which, as he put it, "Dionysus speaks the language of Apollo, but Apollo finally speaks the language of Dionysus" (p. 104) are instances of such art. But he also held that the tragic spirit was almost immediately extinguished in tragedy (in the

literary-historical sense), snuffed out by Euripides' rejection of Dionysiac wisdom in favor of Socratic rationality. Nor, he held, is the tragic spirit to be found in post-Renaissance tragedy, in which music, through which the Dionysian wisdom is expressed, plays no substantial role. In fact, Nietzsche believed, at least at the time when he wrote *The Birth of Tragedy*, if not for long afterward, the only art capable of rediscovering the spirit of tragedy is the music-drama of Richard Wagner, the dedicatee of *The Birth of Tragedy*.

The concern with tragedy as a source of insight into problems that are in the broadest sense problems of ethics, which is exhibited in different ways by Hegel, Schopenhauer, and Nietzsche, has been taken up distinctively in contemporary Anglo-American philosophy by Stanley Cavell, who has shown how Shakespearean tragedy can be read as working out problems of skepticism, and as occupied with "how to live at all in a groundless world"; by Martha Nussbaum (1986), who has taken up Hegel's concern with the ethical dilemmas posed in classical tragedy; and by Bernard Williams (1993), who finds in classical tragedy an exploration of the nature of necessity which challenges Kantian conceptions of the voluntary, of obligation, and of responsibility. Here, as in contemporary discussion of the so-called "paradox of tragedy," Plato's fascination with tragedy, though not his condemnation of the art form, lives on.

See also Aristotle; Beattie, James; Burke, Edmund; Cavell, Stanley; Emotion; Goethe, Johann Wolfgang von; Greek Drama; Hegel, Georg Wilhelm Friedrich; Home, Henry; Hume, David; Katharsis; Lessing, Gotthold Ephraim; Nietzsche, Friedrich; Nussbaum, Martha; Plato; Priestley, Joseph; Schiller, Friedrich Wilhelm Joseph von; Schlegel, Friedrich von; Schopenhauer, Arthur; Smith, Adam; Whitehead, Alfred North; Williams, Bernard.

Bibliography

Aristotle. *Poetics*. Translated by Gerald F. Else. Ann Arbor: University of Michigan Press, 1967.

Beistegui, M., and S. Sparks. *Philosophy and Tragedy*. London: Routledge, 2000.

Budd, M. *Values of Art*. London: Allen Lane, 1995.

Bungay, Stephen. *Beauty and Truth: A Study of Hegel's Aesthetics*. New York: Oxford University Press, 1984.

Burke, Edmund. *A Philosophical Enquiry*, edited by Adam Phillips. New York: Oxford University Press, 1990.

Carroll, N. *The Philosophy of Horror, or Paradoxes of the Heart*. New York: Routledge, 1990.

Cavell, Stanley. *Disowning Knowledge in Six Plays of Shakespeare*. Cambridge, U.K.: Cambridge University Press, 1987.

Eaton, Marcia. "A Strange Kind of Sadness." *Journal of Aesthetics and Art Criticism* 41 (1) (1982): 51–63.

Feagin, Susan. "The Pleasures of Tragedy." *American Philosophical Quarterly* 20 (1) (1983): 95–104.

Hegel, G. W. F. *Aesthetics: Lectures on Fine Art*. 2 vols. Translated by T. M. Knox. Oxford: Clarendon Press, 1975.

Hume, David. "Of Tragedy." In *Essays Moral, Literary and Political*, edited by Eugene F. Miller. Indianapolis, IN: Liberty Classics, 1987.

Kuhns, R. *Tragedy: Contradiction and Repression*. Chicago: University of Chicago Press, 1991.

Lamarque, P. "Tragedy and Moral Value." *Australasian Journal of Philosophy* 73 (1995): 239–249.

Levinson, J. "Horrible Fictions." *Journal of Aesthetics and Art Criticism* 49 (1991): 253–258.

Levinson, J. "Music and Negative Emotion." *Pacific Philosophical Quarterly* 63 (1982): 327–346.

Morreall, John. "Enjoying Negative Emotions in Fiction." *Philosophy and Literature* 9 (1) (1985): 95–102.

Neill, A. "Hume's 'Singular Phænomenon.'" *British Journal of Aesthetics* 39 (1999): 112–125.

Neill, A. "'An Unaccountable Pleasure': Hume on Tragedy and the Passions." *Hume Studies* 24 (1998): 335–354.

Nietzsche, Friedrich. *The Birth of Tragedy*. Translated by Shaun Whiteside and edited by Michael Tanner. London: Penguin Books, 1993.

Nussbaum, Martha. *The Fragility of Goodness*. Cambridge, U.K., Cambridge University Press, 1986.

Packer, M. "Dissolving the Paradox of Tragedy." *Journal of Aesthetics and Art Criticism* 47 (1989): 211–219.

Peckham, M. *Beyond the Tragic Vision*. New York: George Braziller, 1962.

Plato. *The Republic*. Translated by Desmond Lee. London: Penguin Books, 1987.

Ridley, A. "Tragedy and the Tender-Hearted." *Philosophy and Literature* 17 (1993): 234–245.

Schier, Flint. "The Claims of Tragedy: An Essay in Moral Psychology and Aesthetic Theory." *Philosophical Papers* 18 (1) (1989): 7–26.

Schier, Flint. "Tragedy and the Community of Sentiment." In *Philosophy and Fiction: Essays in Literary Aesthetics*, edited by Peter Lamarque. Aberdeen: Aberdeen University Press, 1983.

Schopenhauer, Arthur. *The World as Will and Representation*. 2 vols. Translated by E. F. Payne. New York: Dover, 1969.

Williams, Bernard. *Shame and Necessity*. Berkeley: University of California Press, 1993.

Alex Neill (2005)

TRANSCENDENTALISM

See *Kant, Immanuel; Neo-Kantianism; New England Transcendentalism*

TRESCHOW, NIELS
(1751–1833)

Niels Treschow, the Norwegian philosopher, defended a monism strongly influenced by Benedict de Spinoza and Gottfried Wilhelm Leibniz. Treschow was born at Drammen, Norway. He studied at the University of Copenhagen, where he became a professor in 1803. In 1813 he left Denmark to become the first professor of philosophy at the University of Oslo, but he held the post for only one year before entering government service.

Treschow's philosophical views are based on an idea of the unity of all things and on a concept of God similar to that of Spinoza. However, Treschow wanted to combine the idea of God's immanence, the idea that God is in all things, with the idea of God's transcendence, the idea that God is above all things. God is not the unity of all things but rather that which makes all things into a unity; as such, God is not an abstraction but a real individual, "unchangeable, eternal, and independent" (*Om Gud, Idee- og Sandseverdenen*, Vol. I, p. 81). The nature of God is manifest in our consciousness. God, or the One, "stands in the same relation to the manifold produced by it as does our mind to its thoughts, feelings, and decisions" (p. 115). Our consciousness "pictures the Absolute One."

In his psychology also, Treschow tried to uphold a Spinozistic view, opposing the Cartesian dualism of soul and body. "Man may indeed be considered composite," Treschow said, but not a composite of soul and body, for these are both different aspects of the same thing as it is a possible object of the inner and outer sense (see *Om den Menneskelige Natur*, p. 11).

Treschow also commented on the problem of universals and individuals. He criticized the tendency of abstract philosophers to give priority to universals and to regard individual things and events as instances and exemplifications of universals. The concrete individual, he held, is prior in existence and in knowledge. Only individuals exist, and universals are merely means toward the recognition and description of individual things. An individual thing cannot be fully grasped, however, since this would involve recognizing what is at the basis of all its various states, the idea that expresses all these states.

Since only individuals are real, universal concepts, or concepts of species of things, are "artificial," and so also is any classification of things into more or less fixed kinds. The "specific nature of man" is in a way a fiction, but man has developed gradually from some animal in which the specifically human dispositions potentially inhered, and the natural history of man is part of the history of the whole of nature. In his philosophy of history Treschow tried to substantiate his claim that man descended from some species of animal. Humankind's gradual development is due to the interaction of external and internal conditions. The fact that the individual physically and mentally goes through the various phases of the historical development of the species was to Treschow another proof of the primacy of the individual.

See also Cartesianism; Holism and Individualism in History and Social Science; Leibniz, Gottfried Wilhelm; Spinoza, Benedict (Baruch) de; Universals, A Historical Survey.

Bibliography

WORKS BY TRESCHOW

Gives der Noget Begreb Eller Nogen Idee om Enslige Ting? (Are there concepts or ideas about particular things?). Copenhagen, 1804.

Elementer til Historiens Philosophie (Elements of the philosophy of history). Copenhagen, 1811.

Om den Menneskelige Natur, Især fra Dens Aandelige Side (Human nature, especially its mental aspects). Copenhagen, 1812. The first Scandinavian work on empirical psychology.

Om Gud, Idee- og Sandseverdenen (On God and the worlds of ideas and sensations), 3 vols. Christiania, 1831–1833.

WORKS ON TRESCHOW

Høffding, Harald. *Danske Filosofer* (Danish philosophers). Copenhagen, 1909.

Schmidt-Phiseldech, K. *Niels Treschows Historiefilosofi*. Copenhagen, 1933.

Stybe, Svend Erik. "Niels Treschow (1751–1833), A Danish Neoplatonist." *Danish Yearbook of Philosophy* 13 (1976): 29–47.

Svendsen, P. *Gullalderdrøm og Utviklingstro* (The dream of the golden age and the belief in evolution). Oslo, 1940.

Winsnes, A. H. *Niels Treschow. En Opdrager til Menneskelighet* (Niels Treschow. An educator to humanity). Oslo, 1927.

Anfinn Stigen (1967)
Bibliography updated by Tamra Frei (2005)

TROELTSCH, ERNST
(1865–1923)

Ernst Troeltsch, the German theologian and social scientist, was born near Augsburg in Bavaria. He studied Protestant theology at the universities of Erlangen, Göttingen, and Berlin, and after three years as a Lutheran curate in Munich, he returned to the University of Göttingen as a lecturer in theology. He became extraordinary professor at Bonn in 1892, and in 1894 ordinary professor of systematic theology at Heidelberg, a position that

he held for twenty-one years. He also served as a member of the Bavarian upper legislative house. In 1915 he moved to a chair of philosophy in the University of Berlin, serving concurrently as a member of the Prussian Landtag and as undersecretary of state for religious affairs.

Troeltsch contributed to the philosophy and sociology of religion and also to cultural and social history, ethics, and jurisprudence. His work raised in many related fields the much-debated questions of the extent and limitations of the historicosociological method. He played a leading role in the clarification of the conception of historicism and made important contributions to the study of methodology in the historical sciences. By recognizing the impact of sociological and historical thinking on the shaping of modern mentality, Troeltsch became involved in the intractable problems of the relation between absolute ethical and religious values and historical relativity. He remained uncompromisingly sincere in revealing the difficulties of this approach and admitted to not being able to surmount them or to reconcile conflicting results in an all-embracing theory.

Troeltsch's intellectual development was bound up with his recognition of the importance of historical change. He chose theology as the field in which, in his own words, "one had access to both metaphysics and the extraordinarily exciting historical problems." The historical theology devoid of metaphysics of his teacher Albrecht Ritschl stimulated him to radical doubt of the validity of Ritschl's own procedure, although with Ritschl Troeltsch accepted the Kantian primacy and underivative character of the basic structure of human morality. He argued that moral awareness was basic to the human constitution and that it was only during the course of historical development that morality and religion became connected and interdependent. To understand Christian ethics as the supreme manifestation of such historical combination was nevertheless his aim in *Grundprobleme der Ethik* (written 1902; in *Gesammelte Schriften*, Vol. II).

Troeltsch was aware of the problems arising from two basic assumptions: (1) the Kantian thesis that the formal necessities and laws of morality are irreducible and (2) the equally basic assumption of materialist ethics that what we study are the manifestations of a grown and growing morality in religious, social, and political consciousness. Thus Immanuel Kant's formalism changed in Troeltsch's hands from a means of critical analysis to an attempt to provide an ontology of personality. The point of reference for an understanding of the moral person is no longer the will as such, but morality as realizing itself through persons in history.

Troeltsch's major work is *Die Soziallehren der christlichen Kirchen und Gruppen* (Tübingen, 1912, translated by Olive Wyon as *The Social Teaching of the Christian Churches*, London and New York, 1931). It is a collection of many detailed studies in Christian social ethics published earlier in the *Archiv für Sozialwissenschaft und Sozialpolitik*, with new chapters on Calvinism, the sects, and mysticism. The work is unified by the sociological formulation of the entire history of the Christian churches.

It is easy to see how Troeltsch maneuvered himself into what has been described as the "crisis of historicism." For despite his insistence on the formal a priori of morality and the necessity of thinking of some values and norms as transcending historical change and accident, Troeltsch could not avoid the suggestion that the explanation of a given phenomenon can be adequately provided only by an account of its genesis.

Troeltsch faced the problems his position posed for Christian ethics and theology, with their claims to historically unique or historically transcendent values. In *Die Trennung von Staat und Kirche* (Separation of state and church) he spoke of the polymorphous truth of the churches. This conception was still present in his later attempts to reconcile the absolutist claims of Christian revelation—which as monomorphous truth belongs strictly to the early church—with the later developments of the three great Christian forms of social expression: the church, the sects, and mysticism.

Troeltsch made reliable and learned contributions to the history of ideas, notably his analysis of the role of Protestantism in the formation of the modern world and his searching studies of the differentiation of Protestantism into Calvinism and Lutheranism with their important differences in ethos. He was in basic agreement with his friend Max Weber, whose theses he summarized and elaborated. His important contributions to the conception of group personalities are generally recognized in sociology, philosophy, and jurisprudence. His work on the great social groups—family, guild, state, and church—owed much to Otto von Gierke's *Genossenschaftsrecht*, but Troeltsch went beyond Gierke's emphasis on corporative formations to a study of their personal aspect.

Troeltsch's political thought emerged from his wide learning in the history of ideas. After World War I he was among those German thinkers who realized that Germany's disastrous estrangement from the West was based on a divergence in political philosophy. He urged a return of German political thinking to the position of the eigh-

teenth-century Enlightenment, before the romantic glorification of the state. He thought that this position was compatible with Western thought, as rooted in Stoic and Christian ethics with their essential respect for the individual person that grew into the modern democratic idea of the rights of man. Troeltsch made the point that German political thinking had yet to learn from the West not to despise arrogantly the serious possibilities of compromise.

In 1922 Troeltsch collected his writings on the philosophy of history under the title *Der Historismus und seine Probleme* (Historicism and its problems). Material toward a projected second volume is contained in *Christian Thought, Its History and Application* (London, 1923, edited by Friedrich von Hügel; published in German under the title *Der Historismus und seine Überwindung*, Berlin, 1924).

See also Enlightenment; Historicism; Kant, Immanuel; Philosophy of Religion; Religion and Morality; Ritschl, Albrecht Benjamin; Weber, Max.

Bibliography

WORKS BY TROELTSCH

Vernunft und Offenbarung bei J. Gerhard und Melanchthon. Göttingen: Vandenhoeck and Ruprecht, 1891.

Die Absolutheit des Christentums und die Religionsgeschichte. Tübingen: Mohr, 1902.

Psychologie und Erkenntnistheorie in der Religionswissenschaft. Tübingen: Mohr, 1905.

Protestantisches Christentum und Kirche in der Neuzeit. Leipzig and Berlin, 1906.

Die Trennung von Staat und Kirche. Tübingen, 1907.

Die Bedeutung des Protestantismus für die Entstehung der modernen Welt. Munich: Oldenbourg, 1911. Translated by W. Montgomery as *Protestantism and Progress*. London: Williams and Norgate, 1912.

Die Bedeutung der Geschichtlichkeit Jesu für den Glauben. Tübingen: Mohr, 1911.

Gesammelte Schriften. 4 vols. Tübingen: Mohr, 1912–1925. Does not contain all the works.

Augustin, die christliche Antike und das Mittelalter. Munich: Oldenbourg, 1915.

Spektator-Briefe, Aufsätze über die deutsche Revolution und die Weltpolitik 1918/1922. Tübingen: Mohr, 1924.

Deutscher Geist und Westeuropa. Tübingen: Mohr, 1925.

Vorlesungen über "Glaubenslehre." Munich: Duncker and Homblot, 1925.

Troeltsch, Ernst. *Writings on Theology and Religion*. Atlanta: John Knox Press, 1977.

Troeltsch, Ernst. *Kritische Gesamtausgabe*. Edited by Friedrich Wilhelm Graf. Berlin; New York: W. de Gruyter, 1998.

WORKS ON TROELTSCH

Alberca, Ignacio Escribano. *Die Gewinnung theologischer Normen aus der Geschichte der Religion bei E. Troeltsch*. Munich: Hueber, 1961.

Bodenstein, Walter. *Neige des Historismus, Ernst Troeltschs Entwicklungsgang*. Gütersloh: Gütersloher Verlagshuas, Gerd Mohn, 1959.

Chapman, Mark D. *Ernst Troeltsch and Liberal Theology: Religion and Cultural Synthesis in Wilhelmine, Germany*. Oxford; New York: Oxford University Press, 2001.

Clayton, John Powell. *Ernst Troeltsch and the Future of Theology*. Cambridge; New York: Cambridge University Press, 1976.

Drescher, Hans-Georg. *Ernst Troeltsch: His Life and Work*. Minneapolis: Fortress Press, 1993.

Gayhart, Bryce A. *The Ethics of Ernst Troeltsch: A Commitment to Relevancy*. Lewiston, NY: E. Mellen Press, 1990.

Graf, Friedrich Wilhelm, and Hartmut Ruddies. *Ernst Troeltsch Bibliographie*. Tübingen: J.C.B. Mohr, 1982.

Kaftan, Theodor. *Ernst Troeltsch*. Schleswig, 1912.

Kasch, Wilhelm F. *Die Socialphilosophie von Ernst Troeltsch*. 1963.

Köhler, Walther. *Ernst Troeltsch*. Tübingen: Mohr, 1941.

Myers, Max A., and Michael R. LaChat. *Studies in the Theological Ethics of Ernst Troeltsch*. Lewiston: E. Mellen Press, 1991.

Ogletree, Thomas W. *Christian Faith and History: A Critical Comparison of Ernst Troeltsch and Karl Barth*. New York: Abingdon Press, 1965.

Pauck, Wilhelm. *Harnack and Troeltsch: Two Historical Theologians*. New York: Oxford University Press, 1968.

Renz, Horst, Friedrich Wilhelm Graf, and Trutz Rendtorff. *Troeltsch-Studien*. Gütersloh: Mohn, 1982.

Rintelen, Fritz-Joachim von. "Der Versuch einer Überwindung des Historismus bei Ernst Troeltsch." *Deutsche Vierteljahrsschrift* 8 (1930): 324–372.

Scheler, Max. "Ernst Troeltsch als Soziologe." *Kölner Vierteljahrshefte für Soziologie* (1923–1924).

Tillich, Paul. "Ernst Troeltsch." *Kant-Studien* 29 (1924): 351–358.

Vermeil, Edmond. *La pensée religieuse d'Ernst Troeltsch*. Paris, 1922.

Wichelhaus, Manfred. *Kirchengeschichtsschreibung und Soziologie im neunzehnten Jahrhundert und bei Ernst Troeltsch*. Heidelberg: Winter, 1965.

Eva Schaper (1967)
Bibliography updated by Michael J. Farmer (2005)

TRUBETSKOI, EVGENII NIKOLAEVICH
(1863–1920)

A Russian philosopher, law specialist, religious and political figure, Evgenii Trubetskoi was a member of one of the oldest aristocratic families of Russia. He received an

excellent education, graduating from the Department of Law of Moscow University (1885) and earning a master's degree in philosophy for his work on St. Augustine (1892) and a doctorate for his work on Pope Gregory VII (1897). He taught law and philosophy in Iaroslavl' (1886–1897), Kiev (1897–1905), and Moscow (1905–1917), where he was elected chair of philosophy after the sudden death of its former head, his brother Sergei Trubetskoi (1862–1905). Parallel to his teaching career, he was active in Russian cultural, academic, and political circles. Trubetskoi was one of founders of several philosophical associations (Psychological Society at Moscow University, Vladimir Solov'ev Religious-Philosophical Society, and others); he was a leading figure of the publishing house Put (The Way) and of the group of religious thinkers affiliated with it, who represented the so-called "neo-Slavophile" current in Russian culture. He was one of the founders and leaders of the Constitutional Democratic (Kadet) Party; he was editor in chief (1906–1910) of the liberal-conservative magazine *Moskovsky Ezhenedel'nik* (Moscow weekly); a member of the State Council in 1916–1917; and a participant in the Council of the Russian Orthodox Church in 1917–1918. After the Revolution of 1917 he adopted a sharply anti-Bolshevik stance and joined the White Army. Trubetskoi died in Novorossiysk at the Black Sea, where the defeated army was preparing to leave Russia.

Trubetskoi was a prolific author, whose writings embrace many fields: religion, philosophy, law, and politics. In the last years of his life he wrote valuable studies on Russian icon painting, as well as fairy tales and his memoirs. His main works, in which he presents an original philosophical system, are *Mirosozertsanie V. S. Solov'eva* (V. S. Solov'ëv's world view, 2 vols., Moscow, 1913), *Metafizicheskie predposylki poznaniya* (Metaphysical premises of knowledge, Moscow, 1917), and *Smysl zhizni* (The meaning of life, Moscow, 1918).

His system belongs to the school of Russian religious philosophy founded by Vladimir Solov'ëv and often referred to as "metaphysics of All-Unity." Trubetskoi's place in this school, which includes Pavel Florenskii, Sergei Bulgakov, Lev Karsavin, Nikolai Losskii, and other principal Russian religious thinkers of the twentieth century, is determined by a special attachment of his philosophy to the thought of the founder of the school (this attachment was enhanced by the fact that Trubetskoi and his brother Sergei were close personal friends of Solov'ëv). Other thinkers in the school are more independent of Solov'ëv, adopting from him just a few key ideas, such as "All-Unity," "Sophia the Wisdom of God," or "Godmanhood," and often criticizing him. In the case of Trubetskoi, however, the entire body of his philosophy emerges out of the critical analysis of Solov'ëv's metaphysics.

Trubetskoi defines the message of Solov'ëv's oeuvre as the teaching on "Godmanhood," and reviews all of this vast and heterogeneous work, selecting a certain core that conveys the message rightly and truly. (He leaves out of the core mainly what he calls Solov'ëv's "Utopias": ideas of theocracy, androgynous love, or the absolute nature of the Roman pope's authority). Then he sets the task of developing this core into a systematic philosophy, complementing it with new ideas and concepts. Due to such a method of "immanent critique," his study of Solov'ëv becomes the basis of his own philosophy.

As for new concepts introduced by Trubetskoi, the most important is "Absolute Consciousness," which is his version of Solov'ëv's All-Unity. Each thing or phenomenon is endowed, for Trubetskoi, with its "meaning" or "truth," conceived epistemologically, as a content of a certain consciousness or, in the tradition of Christian Platonism, as "God's idea" of the thing in question; Absolute Consciousness is defined as the set of all such truths. It is structured into the "exoteric" sphere (God's ideas pertaining to the things of the world) and "esoteric" sphere (God's ideas about Himself).

Taking this concept as his point of departure, Trubetskoi develops, first of all, a detailed theory of cognition. In putting the emphasis on cognition, he was influenced by the Western philosophy of his time, dominated as it was by Neo-Kantianism; but at the same time, following the traditional line of Solov'ëv and much of Russian thought in general, he adopts a critical attitude toward both Kant and Neo-Kantianism. Thus the main part of his theory of cognition takes the form of a critical analysis of Kantian epistemology, aiming to disclose implicit "metaphysical (i.e., ontological) premises" in the latter, and to subordinate epistemology to ontology. Attempts of this kind, often described as "the overcoming of Kant," were typical of Russian philosophy of that period and were dubbed "ontological epistemology" by Nikolai Berdiaev. Trubetskoi's theory of cognition is not the most successful of such attempts, since his treatment of such basic concepts as truth and consciousness is clearly in the Kantian line, and his critical attitude is in fact rather superficial.

A devoted Orthodox Christian of traditionalist views, Trubetskoi believed that in trying to describe the inner dynamics of the Absolute, philosophy risks falling into "Gnosticism" and "Schellingianism." Thus his ontol-

ogy, presented chiefly in his last work *Smysl zhizni* (The meaning of life), is a traditional Christian philosophy of God and world, or theodicy, developed with the aid of Solov'ëvian concepts of Godmanhood and Sophia (the latter is identified by Trubetskoi with the exoteric sphere of Absolute Consciousness). The final goal of the course of the world is the "conversion of everything human and, even more, everything terrestrial, into Godmanhood" (*Smysl zhizni*, p. 225). The attainment of this goal is not, however, guaranteed; Trubetskoi resorts to his sophiology to describe the path toward it, which he calls the "process of Godmanhood."

Because of the existence of evil and the freedom of the will, each creature may or may not approach its ideal image in Sophia; in Trubetskoi's terms, it possesses both "sophianic and antisophianic potentials." Thus he considers various spheres of reality, presenting a detailed classification of sophianic and antisophianic elements in each sphere: For example, light is regarded as sophianic and darkness as antisophianic. While it may be questionable as an ontology, this approach becomes fruitful when applied to phenomena of Russian art and culture; in particular, it serves as the underpinnings for Trubetskoi's interpretation of the Russian icon as "contemplation in colors," which won wide recognition.

While hardly the best-known or most profound example of Russian thought, Trubetskoi's philosophy nonetheless demonstrates typical features of the Russian religious-philosophical renaissance: its origins in Solov'ëv's thought; its leanings toward religious and mystical experience, resulting in a mixture of theological and philosophical discourse; and its striving to combine this discourse with the "last word" in Western philosophy.

See also Berdyaev, Nikolai Aleksandrovich; Bulgakov, Sergei Nikolaevich; Florenskii, Pavel Aleksandrovich; Kant, Immanuel; Karsavin, Lev Platonovich; Losskii, Nikolai Onufrievich; Neo-Kantianism; Russian Philosophy; Schelling, Friedrich Wilhelm Joseph von; Solov'ëv (Solovyov), Vladimir Sergeevich; Trubetskoi, Sergei Nikolaevich.

Bibliography

WORKS BY TRUBETSKOI

Entsiklopediia prava (The encyclopedia of law). St. Petersburg: LAN, 1999.

Icons: Theology in Color. Translated by Gertrude Vakar. New York: St. Vladimir Seminary Press, 1971.

Iz proshlogo: vospominaniia; iz putevykh zametok bezhentsa (From the past: Memoirs; Travel notes of a refugee). Tomsk: Vodolei, 2000.

Mirosozertsanie V. S. Solov'eva (V. S. Solov'ëv's world view). Moscow: Medium, 1995.

Smysl zhizni (The meaning of life). Moscow: Respublika, 1994.

Voina i mirovaia zadacha Rossii (War and Russia's global task). Moscow: I. D. Sytin, 1915.

WORKS ABOUT TRUBETSKOI

Zenkovsky, V. V. *A History of Russian Philosophy*. Translated by George L. Kline. London: Routledge, 2003.

Segey Horujy (2005)

TRUBETSKOI, NIKOLAI SERGEEVICH

(1890–1938)

A Russian linguist, ethnologist, and student of culture, Nicolai Trubetskoi was one of the founders of Eurasianism. His father, Sergei Trubetskoi, was a well-known philosopher and the first elected rector of Moscow University. Although a descendant of an old aristocratic family, he played an outstanding role in the democratization of Russian life. Unfortunately, his life was cut short: He died less than a month after his election at the age of forty-three; the same fatal ailment (heart disease) killed his son, who lived to be only forty-eight.

It is hard to determine to what extent Trubetskoi's family was responsible for his future scholarly and political views, but certain influences are apparent. He grew up in a devout Orthodox family and owed a great deal to his religious upbringing. The history and meaning of Christianity interested both father and son. The same holds for the relations between Christianity and other religions. The least one can say about Trubetskoi's worldview is that it was formed in a highly cultured religious family with a strong interest in Russia's history and destiny.

In 1905, when Trubetskoi was fifteen years old, he published his first article, but his scholarly interests date back to 1903. He was a typical child prodigy and in this respect he continued the tradition of his incredibly gifted family. Trubetskoi's article was published in the prestigious *Etnograficheskoe obozrenie* (Ethnographic review). It treats the Finnish song "Kulto neito" in light of the theory of survivals. His contributions to the same journal appeared regularly until World War I.

Like many of his peers, Trubetskoi did not go to school: His teachers were private tutors. In 1908 he entered Moscow University and declared his major in the philosophical-psychological department. Disappointed with its curriculum, after two semesters he transferred to the Department of Linguistics but never lost interest in

philosophy. His indebtedness to Georg Hegel is unmistakable, and in matters of history he was an extreme determinist. However, his Orthodoxy can explain his teleological position as well as his affinity with Hegel. As a prospective philologist Trubetskoi studied old languages and the comparative method. He also continued his studies of non-Indo-European languages and folklore (especially Finno-Ugric and Caucasian).

In 1913 Trubetskoi graduated with a work on the expression of the future in Indo-European and stayed at the university to prepare for advanced exams and eventually to join the faculty. He spent the next year in Leipzig, where he heard the lectures of the greatest comparative scholars of that time. On his return to Moscow, he married Vera Petrovna Bazilevskaia (1892–1965). In 1915 he passed his master's exams and in 1916 received the rank of adjunct professor. The 1917 Bolshevik Revolution found him in the Caucasus, and he never saw Moscow again. He migrated south with the White Army and eventually came to Constantinople. There he received an offer from Sofia University and spent two years as a docent in Bulgaria. In Sofia in 1920, Trubetskoi published his book *Rossiia i chelovechestvo* (Russia and mankind), which inaugurated Eurasianism, a trend that later enjoyed great popularity among the Russian émigrés between two world wars.

The main idea of Eurasianism is that Russia belongs to the East rather than to the West and has little to do with "the Romano-Germanic" world. Trubetskoi's diatribe against the West is oddly at variance with his upbringing, for he was a classic product of European culture, but it accords well with his lifelong interest in non-Indo-European languages and oral tradition and his glorification of the morals of nomadic peoples. It therefore comes as no surprise that his next book bears the title (in translation) *The Legacy of Genghis Khan: A Perspective on Russian History Not from the West but from the East* (1925). Trubetskoi's attack on European ethnocentricity found many supporters and many opponents among his contemporaries, but after World War II his theories merged with those of the anticolonial movement, which explains a renewed interest in them. His Eurasianist works and the trend he initiated have been studied extensively in many countries, and the foundational texts have been translated into several "Romano-Germanic" languages. After the collapse of the Soviet Union, they were also published in Russia.

Trubetskoi's position in Bulgaria was precarious, but an offer from Vienna University to become a professor of Slavic secured his future, and in the autumn of 1922 the most productive period of Trubetskoi's life began. In Vienna he taught all the Slavic languages and literatures, and his lectures, published posthumously, provide a good idea of his activities. Eurasianism too remained at the center of his interests. However innovative his ideas on Russian history and its future and however original his contributions to the study of Russian literature, especially medieval, may be, it is his linguistic work that made him world famous. Trubetzkoi is the founder of a branch of linguistics known as phonology. His main ally in that endeavor was Roman Jakobson, another expatriate from Moscow, who lived in Czechoslovakia. He and Trubetskoi became the main inspiration of a group of linguists known as the Prague Circle.

The focus of phonology is not on the production of the sounds of speech but on their ability to distinguish meaning, form oppositions, and change as elements of a system and a self-regulating code. Sounds viewed from this perspective are called phonemes. Phonology (that is, functional phonetics) served as the basis of what came to be known as structuralism. The conceptual apparatus of phonology was later extended to the other areas of linguistics, mythology, folklore, literary studies, anthropology, psychology, and even geography, with varying success. Although the Prague version of structuralism is not the only one, it is arguably the most influential. Trubetskoi developed his ideas in numerous publications, but his main book appeared posthumously.

On March 13, 1938, German troops occupied Austria. All his life Trubetskoi suffered from various illnesses; the spring of 1938 was an especially hard period for him. The Gestapo subjected him to a long interrogation, and his papers were impounded. The search and the interrogation had a devastating effect on Trubetskoi. Dangerous symptoms developed in his lungs, and on June 25 he died.

See also Eurasianism; Hegel, Georg Wilhelm Friedrich; Phonology; Trubetskoi, Sergei Nikolaevich.

Bibliography

WORKS BY NIKOLAI TRUBETSKOI

Principles of Phonology. Translated by Christiane A. M. Baltaxe. Berkeley: University of California Press, 1969. Originally published as *Grundzüge der Phonologie* (Travaux du cercle linguistique de Prague 7. Prague: Cercle linguistique de Copenhague et Ministè, publ. de la Rèpublique Tcéco-Slovaque, 1939). The book contains a bibliography of Trubetskoi's works through 1968.

Theory and History of Literature. Vol. 72: *Writings on Literature*, edited, translated, and introduced by Anatoly Liberman.

Minneapolis: University of Minnesota Press, 1990. Includes, among others, Trubetzkoy's lectures on Dostoevsky.

WORKS ABOUT NIKOLAI TRUBETSKOI

Liberman, Anatoly, ed. *The Legacy of Genghis Khan and Other Essays on Russia's Identity*. Ann Arbor: Michigan Slavic Publications, 1991. Liberman's postscript contains a detailed biography of Trubetzkoy and an analysis of his Eurasianist works.

Liberman, Anatoly, ed. *Studies in General Linguistics and Language Structure*. Translated by Marvin Taylor and Anatoly Liberman. Durham, NC: Duke University Press, 2001.

Nos. 2–4 contain numerous references to the literature on Trubetzkoi.

Anatoly Liberman (2005)

TRUBETSKOI, SERGEI NIKOLAEVICH

(1862–1905)

Sergei Nikolaevich Trubetskoi was a Russian philosopher, socially conscious essayist, and man of public affairs. After graduating from the historico-philological department of Moscow University in 1885, he remained at the university. In 1890 he defended his master's dissertation, "Metafizika v Drevnei Gretsii" (Metaphysics in ancient Greece), and in 1900 he defended his doctoral dissertation, "Uchenie o Logose v ego istorii" (The doctrine of the logos in its history). From 1900 to 1905 he served as one the editors of the journal *Voprosy filosofii i psikhologii* (Questions of philosophy and psychology). He actively participated in the Zemstvo movement, becoming one of its spiritual leaders. Starting in 1901, at the beginning of the student disturbances, he came out for the institution of university autonomy. After Moscow University was granted autonomy in 1905, he was chosen as its head. However, the wave of disturbances at that time had swamped the university, putting liberal defenders of academic freedoms in a difficult position and leading to Trubetskoi's untimely death.

In his philosophical views Trubetskoi is close to Vladimir Sergeevich Solov'ëv. Like Solov'ëv, Trubetskoi experienced the influence of the Slavophiles, German idealism, and ancient Platonism, uniting Christianity and Platonism in his doctrine. However, Trubetskoi did not share Solov'ëv's mysticism: If for Solov'ëv the central theme was the doctrine of Sophia, Trubetskoi's main work was devoted to the theme of the Logos. And it is not by chance that an early work of Trubetskoi's that was devoted to sophiology remained unfinished; in his works this theme is represented by a theory of the world soul, where Platonism is united with a Kantian doctrine of *a priori* forms of sensation.

TRUBETSKOI'S CONCRETE IDEALISM

Trubetskoi's conceptions received their most complete exposition in his works *O prirode chelovecheskogo soznaniia* (On the nature of human consciousness; 1889–1891) and *Osnovaniia idealizma* (The foundations of idealism; 1896). He called his doctrine *concrete idealism*, in contrast to the abstract idealism of classical German philosophy.

As his starting point, Trubetskoi takes not abstract concept (of the type of Georg Wilhelm Friedrich Hegel's *pure being*) but concrete being, real entity as the subject of all definitions, which reveals thought in this subject. Being necessarily precedes thought; if the contrary is assumed, one arrives at panlogism, that is, at the production from abstract thought of all the abundance of its definitions. According to Trubetskoi the eternal actual consciousness (God) precedes every finite (becoming) consciousness; he thus rejects the pantheistic doctrine of Johann Gottlieb Fichte, Friedrich Wilhelm Joseph von Schelling, and Hegel concerning humanity as the "becoming God" and defends the positions of theism. Attempting to prove that being cannot be reduced to a logical idea and that general concepts are only relations of thought to its object, Trubetskoi at the same time recognizes the spiritual nature of reality, the rational laws of the cosmic Logos according to which both natural life and human life are ordered.

In attempting to remain on the foundation of rationalism, the philosopher, however, does not consider reason to be the sole source of knowledge. Just as in man it is possible to identify three faculties—sense perception, thought, and will—so knowledge, too, is realized with the aid of experience, conditioned by the *a priori* laws of perception (universal sensationalism), with the aid of reason, which reveals the lawful connection of phenomena, the universal correlatedness of that which exists, and finally with the aid of faith, which establishes the reality of the entities one thinks and perceives. The object of faith is an autonomous living power, defined as spirit; faith, according to Trubetskoi, is the recognition of "real entities or subjects independent of us" (1994, p. 671). With this, faith "convinces us of the reality of the external world, of the reality of objects of sense perception and reason" (p. 665).

In contrast to Solov'ëv, Trubetskoi does not identify faith with intellectual intuition or with inspiration: True

to Orthodox tradition, he is careful to separate faith from imagination and places the moral or ethical sphere above the aesthetic sphere. Both in God and in humans the foundation of the personality is will; and therefore being is revealed to faith as a faculty of will. However, Trubetskoi does not oppose faith and reason, revelation and speculation, but points to their unity, emphasizing that "the concept of the Logos is connected with Greek philosophy, in which it arose, and with Christian theology, in which it took firm root" (1994, p. 44).

In accordance with this conviction, Trubetskoi devotes his chief historico-philosophical work, *Metaphysics in Ancient Greece* (1890), to Greek philosophy, where the concept of the Logos was formed, and his chief historico-theological work, *The Doctrine of the Logos in Its History*, to the Christian understanding of the Logos, which was developed in the struggle with Judaism and Gnosticism. Greek philosophy, according to Trubetskoi, is one of the spiritual sources of Christianity. It is not antagonistic to Christianity, not the cause of the distortion of the original Evangelical faith, as many Protestant theologians have asserted. Nevertheless, Trubetskoi recognizes the achievements of Protestant scholarship, in particular that of the historical criticism of Adolf Harnack, to whose discoveries he attempts to give his own interpretation, on the basis of the Orthodox patristic tradition. With his thoughts developing in the spirit of this tradition, Trubetskoi displays a critical attitude toward Solov'ëv's theocratic utopia and toward his interpretation of the Bible through the prism of mystical symbolism and Catholic orthodoxy.

THE SOBORNOST (CONCILIAR NATURE) OF CONSCIOUSNESS

In analyzing the nature of human consciousness, Trubetskoi poses the complex philosophical question about the interrelation of the individual and the universal. According to Trubetskoi this question has not been resolved in European philosophy: Neither empiricism nor idealism have been able to explain the nature of consciousness, and therefore the nature of personality has not been understood. The empiricists identified personality with individual internal states of consciousness, with a set of psychical associations (psychologism) that do not have objective logical significance. By contrast, German idealism dissolved personality in a universal principle, making it a disappearing "moment" in the development of the absolute spirit.

According to Trubetskoi the common root of modern European philosophy in its two variants is subjectivism, originating in Protestantism. Having shown that it is impossible to explain consciousness either as a property of the separate empirical individual or as a product of a universal generic principle, Trubetskoi, following the Slavophiles, arrives at the conclusion that the personal, finite consciousness can be understood only if one admits the sobornost (conciliar nature—from "church council") of consciousness, the common or communal nature of the latter. He considers that this is the only way one can explain man's ability to gain universal and necessary knowledge of reality and to gain an understanding of other people and of the surrounding world. Sobornost as the essence of consciousness is conceived by Trubetskoi as guaranteeing the objectivity of knowledge. For him, the premise of this objectivity and therefore of the possibility of communal consciousness (consciousness rooted in sobornost) is the existence of the eternally actual consciousness, that is, the consciousness of the divine person of the Creator.

Sobornost is a kind of perfect society or a "metaphysical socialism." "Individualistic psychology and subjective idealism both lead to the rejection of the individual soul, but metaphysical socialism, the recognition of the sobornost of consciousness, grounds our faith in this soul. If it is grounded abstractly, isolated individuality tends to become a zero, nothing; individuality is preserved and actualized only in society, and in fact only in the perfect society" (1994, p. 577). The perfect society is an ideal toward which humankind strives. This society must be ruled by the law of love, and love is "the unity of all in one, the consciousness of all in oneself and of oneself in all" (p. 592). But such love, according to Trubetskoi, is unrealizable in natural human union. It presupposes the divine-human union, or the Church.

Just as reason is a property of the universal subject, sense perception, too, according to Trubetskoi, should not be considered to belong only to the individual consciousness. There exists a certain universal sense perception whose bearer is the world soul as its subject, distinct from God. Trubetskoi conceives this bearer as a cosmic entity, or as the world in its psychical foundation, thanks to which the world appears as a living and animate organism. Remaining an adherent of the Logos complemented by faith, Trubetskoi is convinced that at the foundation of the world there lies a rational and loving principle, and for this reason the world is essentially good. This is the source of Trubetskoi's optimism, of his energy, and of his indefatigable academic and public activity.

See also Fichte, Johann Gottlieb; Harnack, Carl Gustav Adolf von; Hegel, Georg Wilhelm Friedrich; Idealism;

Platonism and the Platonic Tradition; Schelling, Friedrich Wilhelm Joseph von; Solov'ëv (Solovyov), Vladimir Sergeevich; Trubetskoi, Evgenii Nikolaevich.

Bibliography

WORKS BY TRUBETSKOI

Sobranie sochinenii (Collected works). Vols. 1–6. Moscow, 1906–1912.

Sochineniia (Works). Moscow, 1994.

WORKS ON TRUBETSKOI

Bohachevsky-Chomiak, Martha. *Sergei N. Trubetskoi: An Intellectual among the Intelligentsia in Prerevolutionary Russia*. Belmont, MA: Nordland, 1976.

Elansky, P. P. "Kn. S. N. Trubetskoi i filosofiia" (Prince S. N. Trubetskoi and philosophy). *Mysl' i slovo* (Moscow) (1917).

Gaidenko, P. P. "'Konkretnyi idealizm' S. N. Trubetskogo" (The "Concrete Idealism" of S. N. Trubetskoi). *Voprosy literatury* 9 (1990).

Lopatin, L. M. *Kniaz' S. N. Trubetsk oi i ego obshchee filosofskoe mirosozertsanie* (Prince S. N. Trubetskoi and his general philosophical worldview). Moscow, 1906.

Kotliarevsky, S. A. "Mirosozertsanie kniazia S. N. Trubetskogo" (The worldview of Prince S. N. Trubetskoi). *Voprosy filosofii i psikhologii* 131 (1) (1916).

Rachinsky, G. A. "Religiozno-filosofskie vozzreniia kn. S. N. Trubetskogo" (The religious-philosophical views of Prince of S. N. Trubetskoi). *Voprosy filosofii i psikhologii* 131 (1) (1916).

Sbornik rechei, posviashchennykh pamiati S. N. Trubetskogo (Collection of speeches dedicated to the memory of S. N. Trubetskoi). Moscow, 1909.

Zen'kovskii, V. V. *A History of Russian Philosophy*. 2 vols. Translated by George L. Kline. New York: Columbia University Press, 1953. Originally published in Russian under the title *Istoriia russkoi filosofii* in 1948 and 1950.

P. Gaidenko (2005)
Translated by Boris Jakim

TRUTH

Theories of truth investigate truth as a property of one's thoughts and speech. We attribute truth and falsity to a wide variety of so-called *truth-bearers*: linguistic items (sentences, utterances, statements, and assertions), abstract items (propositions), and mental items (judgments and beliefs). What is the property we are attributing when we call a truth-bearer true? The question is crucial because of truth's involvement in central philosophical claims: For example, it is often said that truth is the aim of science, that the meaning of a sentence is given by the conditions under which it is true, that logical validity is the preservation of truth, or that ethical statements are neither true nor false. A proper understanding of truth promises to illuminate fundamental issues in metaphysics, the philosophy of language, logic, and ethics.

The two traditional theories of truth are the correspondence theory and the coherence theory. Further theories of truth have emerged since the last part of the nineteenth century, most notably the pragmatic theory, the identity theory, and the semantic theory. There has also been a reaction against the idea that truth has a substantive nature to uncover, which has led to markedly increased support for so-called deflationary theories of truth.

A different motivation for theorizing about truth is the challenge posed by the semantic paradoxes, especially the Liar paradox. Theories of truth prompted by the Liar tend to be concerned less with the nature of truth, and more with the logic and semantics of the predicate *true*. There has been surprisingly little contact between these two groups of theories (though see Priest, Beall, and Armour-Garb 2005).

THE CORRESPONDENCE THEORY OF TRUTH

According to the correspondence theory truth consists in correspondence to the facts. A truth-bearer (say, the proposition that snow is white) is true if and only if it corresponds to a fact (that snow is white). Broadly speaking, truth is a relational property between truth-bearers on the one side and the world on the other.

There is the suggestion of the correspondence account in Plato's *Sophist* (263b), where in Theaetetus's presence the Stranger contrasts the true statement "Theaetetus sits" with the false statement "Theaetetus flies": "The true one states about you the things that are as they are … [w]hereas the false statement states about you things different from the things that are." In *Categories* Aristotle writes, "The fact of the being of a man carries with it the truth of the proposition that he is … the truth or falsity of the proposition depends on the fact of the man's being or not being" (14b14–22; see also 4b8). The correspondence idea may also be present in Aristotle's famous definition of truth, "To say of what is that it is, and of what is not that it is not, is true" (*Metaphysics* Γ, 1011b25). Echoes of the Platonic-Aristotelian account are present in the Stoics and medieval philosophers (e.g., St. Thomas Aquinas, William of Ockham, and Jean Buridan), and many modern philosophers from René Descartes onward endorse the correspondence idea, though with little or no discussion.

A classic statement of the correspondence theory is given by G. E. Moore: To say of a given belief that it is true "is to say that there is in the Universe *a* fact to which it corresponds" (1953, p. 302). Moore takes it that we are all perfectly familiar with the relation of correspondence, "That there *is* such a relation, seems to me clear; all that is new about my definitions is that they concentrate attention upon just *that* relation, and make it the essential point in the definitions of truth and falsehood" (p. 304). Moore's remarks bring out both a strength and a weakness of the correspondence theory. The correspondence theory is the most natural account of truth—it seems that no one need deny that a true belief corresponds to how things are. But this raises the suspicion that the correspondence theory is platitudinous—to say that a truth-bearer corresponds to the facts is just an elaborate way of saying that it is true. There is no distinctive theory of truth unless more can be said about the correspondence relation. And Moore admits that he can offer no analysis of it; the best he can do, he says, is to "define it in the sense of pointing out what relation it is, by simply pointing out that it is *the* relation which does hold between this belief, if true, and this fact, and does not hold between this belief and any other fact" (p. 301).

Bertrand Russell (1906–1907, 1912/1959) attempts to shed light on the correspondence relation by arguing for a structural isomorphism or congruence between beliefs and facts. Beliefs and facts are structured complexes, and when a belief-complex is suitably congruent with a fact-complex, the belief is true. Consider Othello's belief that Desdemona loves Cassio. According to Russell, believing is a four-place relation; in the present case it is the cement that unites Othello, Desdemona, the loving relation, and Cassio into one complex whole. The last three items are what Russell calls the *objects* in the belief, and these objects are ordered in a certain way by the believing relation (Othello believes that Desdemona loves Cassio, not that Cassio loves Desdemona). Now consider another complex unity, Desdemona's love for Cassio, composed of the objects in Othello's belief. Here, the loving relation is the cement that binds together Desdemona and Cassio in the same order that they have in Othello's belief. If this complex unity exists, then it "is called the fact corresponding to the belief. Thus a belief is true when there is a corresponding fact, and is false when there is no corresponding fact" (p. 129).

OBJECTIONS TO THE CORRESPONDENCE THEORY

It is central to Russell's elucidation that there is a structural congruence between the content of a true belief and the corresponding fact—for example, between the proposition expressed by the sentence "Desdemona loves Cassio" and the fact that Desdemona loves Cassio. But sentences and the propositions they express come in a variety of logical structures—negations, conditionals, universal generalizations, and so on. Are there, then, "funny facts": negative facts, hypothetical facts, universal facts, and other logically complex facts? It might seem that the real world—the world of dated, particular events and things in specific spatial and temporal orderings—just does not seem able to contain anything of this kind of complexity: negative, universal, or hypothetical situations, for example. We seem to be presented with a dilemma: either facts are too "linguistic," too closely tied to the logical structures of our language, or facts are worldly items that are not structurally congruent with the propositions we express.

Russell (1956) and Ludwig Wittgenstein (1922) go on to develop their philosophy of logical atomism, according to which there are no logically complex facts, only atomic facts. True propositions that are logically simple or atomic correspond to atomic facts, but logically complex true propositions no longer correspond to logically complex facts. Rather, complex propositions are recursively broken down into the simple propositions that compose them, and the truth of complex propositions is ultimately explained via the atomic facts to which true atomic propositions correspond. Difficulties remain, however: certain complex propositions, for example, "because" statements and subjunctives, are resistant to a recursive breakdown into simple components; and we can still ask whether universal facts are required for true universal generalizations, and negative facts for true negations. Despite these well-known problems, versions of logical atomism are not without their supporters (e.g., see Armstrong 1997). In a different vein J. L. Austin avoids "funny facts" by denying that correspondence is a matter of structural congruence, "There is no need whatsoever for the words used in making a true statement to 'mirror' in any way, however indirect, any feature whatsoever of the situation or event" (1999, p. 155)—even a single word or simple phrase can correspond to a complex situation. Rather, correspondence is a correlation that is determined by our linguistic conventions: it is "*absolutely and purely* conventional" (p. 154).

A far-reaching and influential family of objections to the correspondence theory takes issue with a certain distinction of standpoints that the theory seems to imply. There is the standpoint we occupy when we judge, say, that there are cows in the garden, and then there is the standpoint we occupy when we determine whether our judgment is true. When we occupy this latter standpoint, the correspondence theory seems to require us to judge whether our judgment is appropriately related by correspondence to the facts. Gottlob Frege (1999) objects that there really is no further standpoint to take up, and no further judgment to make—rather we should simply verify whether there are cows in the garden. This line of thought leads Frege to the conclusion that truth is undefinable; it also tends toward deflationism, since it may seem that truth drops out of the picture.

According to another line of objection, it is an illusion that we can have access to an unvarnished realm of facts with which to compare our judgment. Our knowledge of the world is mediated by our descriptions, interpretations, and judgments; we cannot step outside our own system of beliefs and compare those beliefs with "bare reality." Since the correspondence theory says that truth consists in correspondence to the facts, and those facts are inaccessible to us, we can never know that a judgment is true, and we are led to skepticism. Those who endorse this line of criticism typically associate the correspondence theory with metaphysical realism and advocate instead some form of antirealism and an "epistemic" account of truth, say, in terms of verification (like the logical positivists) or assertibility (see Dewey 1938, Dummett 1978).

THE COHERENCE THEORY OF TRUTH

If we cannot judge a belief against the facts, perhaps we should judge it against our other beliefs: does it "hang together" with the rest of our beliefs? The coherence theorist says that the truth of a belief consists in its coherence with other beliefs. Given some favored coherent set of beliefs, the truth of any of its members consists in its membership in that set—in this way the skeptic is disarmed, since truth no longer requires access to an independent realm of facts. Versions of the coherence theory have been attributed to Benedict (Baruch) de Spinoza, Immanuel Kant, Johann Gottlieb Fichte, and Georg Wilhelm Friedrich Hegel (see by way of comparison Walker 1989), and the theory was championed by idealists, including Harold H. Joachim (1906) and Brand Blanshard (1939), at the end of the nineteenth century and the beginning of the twentieth. Joachim rejects Descartes' idea that we can know truths individually, "The ideal of knowledge for me is a system, not of *truths*, but of truth"; knowledge of an individual truth "is the smallest and most abstracted fragment of knowledge, a mere mutilated shred torn from the living whole in which alone it possessed its significance" (1906, p. 48). So Joachim advocates a thoroughgoing holistic view of knowledge and of truth, "Truth in its essential nature is that systematic coherence which is the character of a significant whole" (p. 50). The coherence theory was subsequently adopted by some logical positivists, notably Otto Neurath (1959), who, like Joachim, endorsed a holistic view of knowledge and truth, and combined it with the positivists' verificationist doctrine that no sense can be attached to a reality that goes beyond what can be verified or falsified by the empirical methods of science.

There are attractive features of the coherence theory. In favor of holism, we can say that statements like "The Enlightenment brought about the French Revolution" and "Neutrinos lack mass" cannot be understood in isolation from a good deal of history and science; and we do often test the truth of a statement against a large body of background statements. But the coherence theory is a theory of the nature of truth, not a theory of how we test for truth, and as such it has been the target of a number of objections. Russell (1906–1907), Moritz Schlick (1959), and others have argued that an arbitrary set of propositions, say, those of a fairy tale or a good novel, would count as a set of truths as long as the propositions cohere with one another—where coherence is taken in the sense of consistency or compatibility. An appeal to comprehensiveness seems not to help the coherence theorist here: Given a coherent set of propositions however large, there will always be equally large coherent sets incompatible with it (and with each other). And placing restrictions on membership in the favored set—for example, admitting only our actual beliefs, or ideal beliefs held at the end of inquiry—seems to tie truth less to coherence and more to the successful tracking of the facts. A further objection derives from Russell: Suppose we have a large, coherent set of propositions about, say, the nineteenth century, and suppose that we can coherently add the proposition that Bishop Stubbs wore episcopal gaiters. According to the coherence theory this proposition is true, in virtue of its membership in a coherent set. If we protest that we cannot be committed to its truth because we do not know whether it is true or false, then we are using true and false in a way that the coherence theorist does not recognize. The difficulty is compounded if we now run the argument with the proposition that Bishop Stubbs *did not* wear episcopal gaiters (further discussion

of the coherence theory can be found in Putnam [1981], Blackburn [1984], Davidson [1984], and Walker [1989]).

THE PRAGMATIC THEORY OF TRUTH

The pragmatic theory of truth is associated primarily with the American pragmatists Charles S. Peirce and William James, and their influence can still be felt in the work of, for example, Richard Rorty (1982) and Robert B. Brandom (1994). According to Peirce we are to understand any idea or object through its practical effects, "Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object" (1955b, p. 31).

Peirce applies his rule to the idea of *reality*: the practical effect that real things have on us "is to cause belief" (1955b, p. 36), and so the question is how to distinguish true belief from false belief. Peirce's answer is that the true beliefs are the ones to which we will all agree, and only the methods of science can realize the hope of reaching this consensus. Peirce writes, "This great hope is embodied in the conception of truth and reality. The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real. That is the way I would explain reality" (p. 38). This is not the independently existing reality associated with the correspondence theory: For Peirce, what is special about science is its ability to settle opinion, and reality is whatever settled opinion says it is.

James applies Peirce's rule directly to truth. The practical effects of true beliefs are successful actions, beneficial dealings with the world; truths are "invaluable instruments of action" (1907, p. 97), truths "*pay*" (p. 104). And so, in accordance with Peirce's rule, truth *is* what is useful, what "works." James places less emphasis than Peirce on consensus and scientific method (indeed, Peirce renamed his theory "pragmaticism" to distance it from James's version). James applies his theory to individuals' beliefs as well as collective beliefs, and religious and metaphysical beliefs as well as empirical ones (e.g., "On pragmatist principles, if the hypothesis of God works satisfactorily in the widest sense of the word, it is true" [p. 143]).

It is standardly objected that we can have beneficial false beliefs and detrimental true beliefs. My false belief that I play the violin beautifully may in fact improve my performance; my true belief that I do not may worsen it. James has the resources for a response. While "the true is only the expedient in our way of thinking," truth is the expedient in a strong sense, "expedient in the long run and on the whole of course" (1907, p. 106). We have to take the long view: I may perform well this time, but overall I will be better served by an accurate assessment of my talents. The long view must be taken not only of individuals' beliefs, but of whole theories—Ptolemaic astronomy was expedient for centuries (p. 107). "The 'absolutely' true, meaning what no farther experience will ever alter, is that ideal vanishing-point towards which we imagine that all our temporary truths will some day converge. … Meanwhile we have to live today by what truth we can get today, and be ready tomorrow to call it falsehood" (pp. 106–107).

Along with a controversial commitment to relativism, James presents here a holistic theme that may suit his pragmatism: It is perhaps more plausible that the truth of an entire system of belief, as opposed the truth of our beliefs taken individually, is a matter of its working for us. Taken this way, pragmatism may be seen as a version of the coherence theory. Still, a basic objection remains: It is plausible that a body of truths should be useful or coherent, but it does not follow that truth *consists in* utility or coherence—a correspondence theorist will say that truths are useful and mutually coherent just because they correspond to the world.

THE IDENTITY THEORY OF TRUTH

Despairing of the correspondence theory, F. H. Bradley wrote, "if we are to advance, we must accept once for all the identification of truth with reality" (1999, pp. 35–36). Here, Bradley seems to embrace the identity theory of truth: a truth does not correspond to a fact, but is identical to a fact (Bradley's view is discussed in Candlish 1995). Another influence is Frege's remark, "A fact is a true thought" (1999, p. 101), though Frege himself did not endorse the identity theory. Versions of the theory are defended by Jennifer Hornsby (1997) and Julian Dodd (2000). The theory may appear counterintuitive: If true mental items—true judgments or true beliefs—are facts, then it seems that the mind contains facts, that mind and world are literally the same. It may also be argued that the theory is unstable, collapsing into deflationism or leading to the elimination of true judgments altogether—"straight to thought's suicide," as Bradley puts it (1893, p. 150).

THE SEMANTIC THEORY OF TRUTH

The semantic theory of truth originates with the mathematician and logician Alfred Tarski (1930–1931/1983, 1999). Tarski sought a *definition* of truth that was formally correct and met the following constraint: It must imply all sentences of the form exemplified by

> "Aardvarks amble" is true if and only aardvarks amble,

that is, all sentences of the form: p is true if and only if p. These so-called T-sentences are so basic to truth, Tarski thought, that they must follow logically from any adequate definition—in this way, he said, we do justice to Aristotle's definition (see the previous discussion). Indeed, Tarski regarded each T-sentence as a "partial definition" of truth, and if we were dealing with a finite language (in the sense that it contains only finitely many sentences), we need only list all the associated T-sentences for a complete definition of truth for that language (see 1930–1931/1983, pp. 251–253). But since Tarski was after a definition of truth for formal languages that were infinitary, such a list is not feasible. So instead Tarski provided a recursive definition—not of truth, though, but of the more basic notion of *satisfaction*. In the simplest kind of case, satisfaction is a relation between an object and a predicate—for example, a London bus satisfies the predicate *is red*. Satisfaction is defined recursively, first for predicates (of a given language) that exhibit no logical complexity, and second for those that do. Tarski then defined truth in terms of satisfaction. The result was a definition of truth for formal languages that was formally precise and implied the T-sentences.

It is remarkable that both correspondence theorists and deflationists have found Tarski's account congenial. Correspondence theorists are drawn to satisfaction as a word-world relation and to the possibility that the correspondence relation between a sentence and a fact can be broken down into relations between parts of sentences (predicates and names) and the things they refer to (e.g., Devitt 1991). This raises the hope that correspondence is no more mysterious than the semantic relations between predicates and names and their referents. Deflationists, in particular disquotationalists, are drawn to the idea that the T-sentences say all there is to say about truth, as will be seen later on. Tarski himself emphasized the neutrality of his theory: "We may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naïve realists, critical realists or idealists, empiricists or metaphysicians—whatever we were before. The semantic conception is completely neutral toward all these issues" (1999, p. 140).

Tarski's aim was not to uncover the nature of truth, but to place the concept of truth beyond suspicion. On the one hand, he thought, truth is fundamental to science, logic, and metamathematics; on the other hand, truth has an "evil reputation" because of its involvement with the Liar paradox. Tarski's aim was to find a way of defining truth in terms that no one could question:

> The definition of truth, or of any other semantic concept, will fulfil what we intuitively expect from every definition; that is, it will explain the meaning of the term being defined in terms whose meaning appears to be completely clear and unequivocal. And, moreover, we have then a kind of guarantee that the use of semantic concepts will not involve us in any contradictions. (1999, p. 127)

Anyone wishing to turn Tarski's definition into a fully general account of truth faces a number of obstacles. Tarski defined truth only for regimented, formal languages, not for natural languages like English; the definition is a definition of truth for a given language, not for truth *simpliciter*; and the definition, according to Hartry Field (1972), fails to explain truth since it merely reduces truth to further semantic notions that are not themselves adequately explained.

DEFLATIONARY THEORIES OF TRUTH

Deflationists say that "substantive" theories of truth—such as the correspondence and coherence theories—are radically misguided: there is no substantive property of truth to theorize about. According to Frank Ramsey truth is *redundant*, "It is evident that 'It is true that Caesar was murdered' means no more than that Caesar was murdered" (1999, p. 106). Truth is less easily eliminated from generalizations like "Everything Socrates says is true," but Ramsey argues that it can be done (p. 106). The word *true* disappears, and any reason to investigate the nature of truth disappears along with it. According to a more sophisticated version of the redundancy theory, the prosentential theory of truth (Grover, Camp, and Belnap 1975), the word *true* is not even a genuine predicate, but a mere component of prosentences. If I say "That is true" in response to a claim of yours, I have produced not a sentence but a prosentence, referring back to your sentence just as the pronoun *he* may refer back to the name *John*. We might think of "That is true" as hyphenated, with no more internal structure than the pronoun *he*. On the prosentential view, *true* does not survive as a discrete property-denoting predicate. P. F. Strawson's (1949) variant of the redundancy theory attributes to *true* a performative role: we use *true* not to pick out a property, but to perform speech-acts such as endorsing, agreeing, and conceding.

Disquotationalists also ascribe to *true* a role different from that of ordinary predicates. According to the dis-

quotational theory of truth—championed by W. V. Quine (1970) and further developed and defended by Field (e.g., see 1994)—to say that a sentence is true is really just an indirect way of saying the sentence itself. There really is no more to the truth of the sentence "Penguins waddle" than is given by the Tarskian T-sentence

> "Penguins waddle" is true if and only if penguins waddle,

and the totality of T-sentences tells the whole story about truth. This prompts the question: Why not dispense with the truth predicate in favor of direct talk about the world? The disquotationalist will respond by pointing to generalizations such as "Every sentence of the form 'p or not p' is true" (see Quine 1970, pp. 10–13). In such a case we could dispense with the truth predicate only if we could produce an infinite conjunction of sentences of the form "p or not p": "Aardvarks amble or aardvarks do not amble, and bison bathe or bison do not bathe, and … ." But we cannot produce infinitely long sentences. So to achieve the desired effect, we generalize over sentences, and then, via the truth-predicate, bring them back down to earth by disquoting them. The truth-predicate is a device for disquotation. Despite surface appearances, *true* does not denote a property or relation—it is a logical device. So there is no property of truth to explore and no work for truth to do beyond its logical role.

The disquotational theory takes the truth-bearers to be sentences, and this raises a concern about the scope of the theory (for further concerns, see David 1994). Suppose that on the authority of others I believe that Dmitri is always right, though I speak no Russian. I say, with apparent understanding, "What Dmitri says is true." But according to disquotationalism understanding what I have said is just a matter of understanding what Dmitri said; and since I cannot understand what Dmitri said, I cannot understand what I have said. Disquotationalists typically relativize their theory to the sentences of a given natural language such as English. And since an English speaker will not understand every sentence of English, some disquotationalists recognize the need to go further and restrict the theory to the sentences of a given speaker's idiolect (those sentences that the speaker understands). This seems to lead us away from the commonsensical notion of truth—ordinarily, it seems, we can apply the notion of truth to foreign sentences, and to sentences of English that we do not yet understand. In short, the concept of truth seems not to depend on the sentences that a speaker happens to understand at a given time. The challenge to the disquotationalist (taken up by Field and others) is to ease the counterintuitive restrictions on disquotational truth in ways that do not compromise the theory.

These difficulties for disquotationalism might motivate a different choice of truth-bearer—propositions instead of sentences. Paul Horwich (1998) presents a *minimal theory of truth*, according to which a complete account of truth is given by the propositional analogues of Tarski's T-sentences:

> The proposition that aardvarks amble is true if and only aardvarks amble; The proposition that bison bathe is true if and only if bison bathe,

and so on, ad infinitum. Far from being restricted to speakers' idiolects, *true* applies to all propositions, including those expressed by sentences we do not understand. But now there is a new set of concerns. First, since we do not understand every proposition, we will understand only a fraction of the axioms that compose the minimal theory—and so our grasp of truth must always remain partial. Second, since the minimal theory describes truth in a piecemeal way, for each proposition individually, it does not include any generalizations about truth. So it may be objected that the theory cannot explain generalizations such as "Only propositions are true"—the theory does not tell us what is not true, so it does not rule out, for example, the absurdity that the Moon is true. (For more on this objection, see Anil Gupta [1993]; Christopher S. Hill [2002] offers a version of minimalism that is responsive to it.) Third, consider the form shared by Horwich's axioms: the proposition that p is true if and only p. To obtain an axiom, we must be careful to replace each occurrence of p by English tokens of the same sentence-type, with the same meaning. But now sentences appear to be back in the picture—together with the substantive semantic notion of meaning, which may not be as free of involvement with truth as minimalism requires.

This last remark relates to a general challenge faced by all forms of deflationism. Deflationists typically focus on uses of *true* such as "'Aardvarks amble' is true," or "Most of what Socrates says is true"—what we may call *first-order* uses, where *true* applies to a particular truth-bearer or a set of truth-bearers. But *true* is also used in other ways: for example, consider the claim that the meaning of a sentence is given by its truth-conditions or the claim that to assert is to put forward as true. These uses of *true*, call them *second order*, purport to explain meaning and assertion. Unlike first-order uses, they do not apply to any particular truth-bearers, and so it is not easy to see how they might be treated as redundant and eliminable or given a merely disquotational role. These second-order uses must be explained. Moreover, the

deflationist must show that it is possible to explain meaning and assertion (and many other concepts apparently related to truth, such as validity, belief, verification, explanation, and practical success) in terms that assign to truth a limited logical role or no role at all.

THEORIES OF TRUTH AND THE LIAR

One version of the Liar paradox is generated by the self-referential sentence:

(1) (1) is false.

Suppose that (1) is true, then what it says is the case, and so (1) is false. On the contrary, suppose that (1) is false—then since that is what (1) says it is, (1) is true. A contradiction is reached either way and we are landed in paradox.

Hierarchical theories of truth have perhaps been the orthodox response to the Liar. Let L_0 be a fragment of English that does not contain the predicate *true*. Let true-in-L_0 be the truth predicate for L_0, holding of exactly the true sentences of L_0. If true-in-L_0 is itself a predicate of L_0, then we can construct the Liar paradox in L_0 via the sentence "This sentence is not true-in-L_0." Accordingly, the predicate *true-in-L_0* is confined to a richer *metalanguage* for the *object language* L_0. But on pain of the Liar, this metalanguage cannot contain its own truth predicate; for that a further metalanguage is needed. In this way a hierarchy of languages is generated, each language beyond L_0 containing the truth predicate for the preceding language. By a celebrated theorem of Tarski's (1930–1931/1983), no classical formal language can contain its own truth predicate, and we are led to a hierarchy of formal languages. Some have carried over this result to natural languages as a way of dealing with the Liar, though Tarski did not endorse this move. Russell's hierarchical approach was embodied in his theory of types and orders (1967). It is often complained that hierarchical approaches force an unnatural regimentation on a natural language like English; Russell himself at one time called the approach "harsh and highly artificial."

Another kind of approach abandons classical semantics—usually it is the principle of bivalence ("Every sentence is true or false") that is rejected. If we can motivate the existence of truth-value gaps, then we can say that (1) is neither true nor false and avoid the contradiction. Saul Kripke's (1975) influential theory of truth takes Liar sentences to be "gappy" because they are *ungrounded*: any attempt to evaluate a Liar sentence leads only to sentences involving *true* or *false*—in the case of (1), we are repeatedly led back to (1) itself. Kripke constructs a language that, remarkably, contains its own truth and falsity predicates. It cannot, however, accommodate the predicates "is false or gappy" or "not true"—and so ultimately we cannot dispense with a hierarchy.

The revision theory of truth (Gupta and Belnap 1993) is formally a variant of Kripke's theory, but provides a distinctive way of explaining the meaning of truth. Truth is taken to be a circular concept, and the revision theory describes how its meaning is given by the Tarskian T-sentences via a dynamic process that, through systematic revisions, provides better and better approximations of the extension of *true*.

Contextual theories of truth are motivated by so-called strengthened reasoning about the Liar. Start with a Liar sentence, say,

(2) (2) is not true.

Reasoning in the usual way, we will find that (2) is pathological. But then we may infer

(3) (2) is not true.

Now (2) and (3) are composed of the same words with the same meanings, and yet one is pathological and the other is true. Contextual theorists claim that this change in truth status without a change in meaning is best explained by a contextual shift (compare "I'm hungry" said before dinner and "I'm hungry" said after dessert). Most contextual theories are hierarchical (e.g., Burge 1979, Barwise and Etchemendy 1987), though Keith Simmons (1993) develops a suggestion of Kurt Gödel's, according to which an unstratified concept of truth applies everywhere except for certain singularities.

Any purported solution to the Liar faces the so-called Revenge Liar—a version of the Liar couched in the terms of the solution. Truth-value gap approaches must deal with the Liar sentence "This sentence is false or gappy," hierarchical approaches with "This sentence is not true at any level," and contextual theories with "This sentence is not true in any context." With no agreed-on solution in sight, and with the constant threat of Revenge Liars, some have concluded that we must cut the Gordian knot and embrace the contradictions associated with the Liar. According to dialetheists such as Graham Priest (1987) there are sentences that are both true and false, and among them are the Liar sentences (for critical discussions of dialetheism, see Priest, Beall, and Armour-Garb 2004). Besides meeting the obvious charge of counterintuitiveness, dialetheists must underwrite their theory with a plausible paraconsistent logic (a logic that challenges the principle that everything follows from a con-

tradiction) and ensure that dialetheism is not itself vulnerable to a Revenge Liar.

See also Meaning; Semantics.

Bibliography

Alston, William P. *A Realist Conception of Truth*. Ithaca, NY: Cornell University Press, 1996.

Aristotle. *Categories* and *Metaphysics*. In *The Basic Works of Aristotle*, edited by Richard McKeon. New York: Random House, 1941.

Armstrong, David M. *A World of States of Affairs*. New York: Cambridge University Press, 1997.

Austin, J. L. "Truth." In *Truth*, edited by Simon Blackburn and Keith Simmons, 149–161. New York: Oxford University Press, 1999. Austin's essay originally appeared in *Proceedings of the Aristotelian Society* in 1950.

Barwise, Jon, and John Etchemendy. *The Liar: An Essay on Truth and Circularity*. New York: Oxford University Press, 1987.

Blackburn, Simon. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford, U.K.: Clarendon Press, 1984.

Blackburn, Simon, and Keith Simmons, eds. *Truth*. New York: Oxford University Press, 1999.

Blanshard, Brand. *The Nature of Thought*. London: Allen and Unwin, 1939.

Bradley, F. H. *Appearance and Reality: A Metaphysical Essay*. New York: Macmillan, 1893.

Bradley, F. H. "On Truth and Copying." In *Truth*, edited by Simon Blackburn and Keith Simmons, 31–45. New York: Oxford University Press, 1999. Bradley's essay originally appeared in *Mind* in 1907 and also appeared in *Essays on Truth and Reality* in 1914.

Brandom, Robert B. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press, 1994.

Burge, Tyler. "Semantical Paradox." *Journal of Philosophy* 76 (1979): 169–198.

Candlish, Stewart. "The Truth about F. H. Bradley." *Mind* 98 (1989): 331–348.

Candlish, Stewart. "Resurrecting the Identity Theory of Truth." *Bradley Studies* 1 (1995): 116–124.

David, Marian. *Correspondence and Disquotation: An Essay on the Nature of Truth*. New York: Oxford University Press, 1994.

Davidson, Donald. "A Coherence Theory of Truth and Knowledge." In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, edited by Ernest LePore, 307–319. Oxford, U.K.: Basil Blackwell, 1984.

Davidson, Donald. "The Folly of Trying to Define Truth." *Journal of Philosophy* 93 (1996): 263–279.

Devitt, Michael. *Realism and Truth*. 2nd ed. Oxford, U.K.: Basil Blackwell, 1991.

Dewey, John. *Essays in Experimental Logic*. Chicago: University of Chicago Press, 1916.

Dewey, John. *Logic: The Theory of Inquiry*. New York: Henry Holt, 1938.

Dodd, Julian. *An Identity Theory of Truth*. New York: St. Martin's Press, 2000.

Dummett, Michael. *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, 1978.

Etchemendy, Jon. "Tarski on Truth and Logical Consequence." *Journal of Symbolic Logic* 53 (1988): 51–79.

Field, Hartry. "Deflationist Views of Meaning and Content." *Mind* 103 (411) (1994): 249–284.

Field, Hartry. "Tarski's Theory of Truth." *Journal of Philosophy* 69 (13) (1972): 347–375.

Frege, Gottlob. "The Thought: A Logical Inquiry." In *Truth*, edited by Simon Blackburn and Keith Simmons, 85–105. New York: Oxford University Press, 1999. Frege's essay was originally published in German in 1918 under the title "Der Gedanke: eine logische Untersuchung." It was also reprinted in English under its current title in the journal *Mind* in 1956.

Grover, D., J. Camp, and N. Belnap. "A Prosentential Theory of Truth." *Philosophical Studies* 27 (1975): 73–125.

Gupta, Anil. "Minimalism." In *Philosophical Perspectives*. Vol. 7: *Language and Logic*, edited by James E. Tomberlin, 359–369. Atascadero, CA: Ridgeview Press, 1993.

Gupta, Anil, and Nuel Belnap. *The Revision Theory of Truth*. Cambridge, MA: MIT Press, 1993.

Hill, Christopher S. *Thought and World*. New York: Cambridge University Press, 2002.

Hornsby, Jennifer. "Truth: The Identity Theory." *Proceedings of the Aristotelian Society* 97 (1997): 1–24.

Horwich, Paul. *Truth*. 2nd ed. New York: Oxford University Press, 1998.

James, William. *The Meaning of Truth: A Sequel to "Pragmatism."* New York: Longmans, Green, 1909.

James, William. *Pragmatism, a New Name for Some Old Ways of Thinking: Popular Lectures on Philosophy*. New York: Longmans, Green, 1907.

Joachim, Harold H. *The Nature of Truth: An Essay*. Oxford, U.K.: Clarendon Press, 1906.

Kirkham, Richard L. *Theories of Truth: A Critical Introduction*. Cambridge, MA: MIT Press, 1992.

Kripke, Saul. "Outline of a Theory of Truth." *Journal of Philosophy* 72 (1975): 690–716.

Martin, Robert L. *Recent Essays on Truth and the Liar Paradox*. New York: Oxford University Press, 1984.

Moore, G. E. *Some Main Problems of Philosophy*. New York: Humanities Press, 1953.

Neurath, Otto. "Protocol Sentences." In *Logical Positivism*, edited by A. J. Ayer, 199–208. Glencoe, IL: Free Press, 1959. Neurath's essay was published in 1932–1933 under the title "Protokollsätze" in the journal *Erkenntnis*.

Peirce, Charles S. "The Fixation of Belief." In *Philosophical Writings of Peirce*, edited by Justus Buchler, 5–22. New York: Dover Publications, 1955a. Peirce's essay originally appeared in *Popular Science Monthly* in 1877.

Peirce, Charles. "How to Make Our Ideas Clear." In *Philosophical Writings of Peirce*, edited by Justus Buchler, 23–41. New York: Dover Publications, 1955b. Peirce's essay originally appeared in *Popular Science Monthly* in 1878.

Plato. *Sophist*. In *The Collected Dialogues of Plato, Including the Letters*, edited by Edith Hamilton and Huntington Cairns. New York: Pantheon Books, 1961.

Priest, Graham. *In Contradiction: A Study of the Transconsistent*. Dordrecht, Netherlands: Nijhoff, 1987.

Priest, Graham, J. C. Beall, and Brad Armour-Garb. *The Law of Non-Contradiction: New Philosophical Essays*. Oxford, U.K.: Clarendon Press, 2004.

Putnam, Hilary. *Reason, Truth, and History*. New York: Cambridge University Press, 1981.

Quine, W. V. *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice-Hall, 1970.

Ramsey, Frank. "Facts and Propositions." In *Truth*, edited by Simon Blackburn and Keith Simmons, 106–107. New York: Oxford University Press, 1999. Ramsey's essay originally published in *Proceedings of the Aristotelian Society* in 1927.

Rorty, Richard. *Consequences of Pragmatism: Essays, 1972–1980*. Minneapolis: University of Minnesota Press, 1982.

Russell, Bertrand. "Mathematical Logic as Based on the Theory of Types." In *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by Jean van Heijenoort, 150–182. Cambridge, MA: Harvard University Press, 1967. Russell's essay originally published in *American Journal of Mathematics* in 1908.

Russell, Bertrand. "On the Nature of Truth." *Proceedings of the Aristotelian Society* 7 (1906–1907): 28–49.

Russell, Bertrand. "The Philosophy of Logical Atomism," 1918 lectures. In *Logic and Knowledge: Essays, 1901–1950*, edited by Robert Charles Marsh. London: Allen and Unwin, 1956.

Russell, Bertrand. *The Problems of Philosophy* (1912). New York: Oxford University Press, 1959.

Schlick, Moritz. "The Foundation of Knowledge." In *Logical Positivism*, edited by A. J. Ayer, 209–227. Glencoe, IL: Free Press, 1959. Schlick's essay was originally published in German in 1934 under the title "Ueber das Fundament der Erkenntnis" in the journal *Erkenntnis*.

Simmons, Keith. *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. New York: Cambridge University Press, 1993.

Strawson, P. F. "Truth." *Analysis* 9 (6) (1949): 83–97.

Tarski, Alfred. "The Concept of Truth in Formalized Languages" (1930–1931). In *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. 2nd ed. Translated by J. H. Woodger; edited by John Corcoran. Indianapolis, IN: Hackett, 1983.

Tarski, Alfred. "The Semantic Conception of Truth and the Foundations of Semantics." In *Truth*, edited by Simon Blackburn and Keith Simmons, 115–143. New York: Oxford University Press, 1999. Tarski's essay originally appeared in *Philosophical and Phenomenological Research* in 1944.

Walker, Ralph C. S. *The Coherence Theory of Truth: Realism, Anti-Realism, Idealism*. London: Routledge, 1989.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. 1921. New York: Harcourt, Brace, 1922.

Keith Simmons (2005)

TRUTH AND FALSITY IN INDIAN PHILOSOPHY

By and large, classical Indian philosophy treats truth within an epistemological context, and different theories of truth are connected with different theories of knowledge. Truth is regarded as a property of cognitions, not of sentences or propositions, although it is presupposed that a true cognition, if appropriately verbalized, would be expressed by a true statement. Cognitions form dispositions or beliefs, but the concept of a belief is also not in the forefront in classical Indian analyses. Modern interpreters tend to use the term *veridicality*, rather than *truth*, because of this focus. Cognitions are episodic psychological events divided into types according to epistemic and other criteria, and perceptual, inferential, testimonial, and hypothetical veridical (true) cognitions are not only the results of processes that are veritable "knowledge sources" (*pramāṇa*) but are also causes of effort and action, including speech. A cognition has objecthood, its indication or intentionality, which is a feature it can share with other cognitions: two people can have the same cognition in this sense. Against such a background, contested issues include, most notably, the nature of veridicality as a cognitive property and the nature of justification, that is, how veridicality is known.

PRECLASSICAL AND EARLY CLASSICAL METAPHYSICS

Classical Indian philosophy proper stretches from about 100 BCE to the modern period (1800s and beyond). Earlier Vedic and Upanishadic thought, along with rejection of a Brahmanic worldview by Buddhists, Jainas, and materialists, sets the stage for the professional reasoners. According to yogis and mystics of an early age (recorded in Upanishads, "mystic treatises," from about 800 BCE) consciousness has lost its native state of bliss and self-awareness. It can be recovered through meditation and various practices of yoga and religious discipline. Buddhist literature develops the theme: The world is a dream from which one needs to awaken to an emptiness brimming with delight and compassion, or, in still later Hindu literature, awaken to one's true self as one with the Absolute Brahman. Nonveridical perception is held up as an analogue to one's everyday lack of awareness of Brahman (*nirvāṇa*).

Brahman is the real, the "truth" in a metaphysical sense, and spiritual knowledge, which is compared to veridical perception and is true in some higher sense of the word. Mystical sublation shows Brahman to be the

real (*sat*, being), as a sublating perception shows a rope formerly misperceived as a snake to be the rope that it is. Such reasoning becomes crystallized as the doctrine of two truths common in much Buddhism and Vedānta (i.e., Hindu schools of Upanishadic philosophy). Indeed, Advaita (Nondualist) Vedānta develops a theory of three truths: the true (cognition, or consciousness, of Brahman), the indeterminable (cognition that is true of the world but not of Brahman, for example, a veridical cognition of water), and the false (not true of the world, for example, a dream or mirage). In Buddhism, a four-cornered negation is said to characterize nirvāṇa or speech about nirvāṇa: not F, not not-F, not (F and not-F), and not not (F and not-F). The logic and language of everyday life do not apply.

Metaphysical controversy marks the beginning of classical Indian philosophy, which is defined by texts devoted to systematic presentation of worldviews complete with supporting arguments and attacks on rival theories. Jaina logicians developed a theory of seven-truth perspectives to support their nonabsolutism (*anekāntavāda*) in metaphysics or perspectivalism, the view that truth is relative to a perspective. (Some have seen in this way *ahiṃsā* [non-injury]—the core teaching of Jainism—applied to the life of the mind.) Every philosophy has something to be said for it. Every judgment has a grain of truth, as tied to a particular take on things, but, likewise, the negation of every judgment, and their combination. A fourth *naya* (perspective) is inexpressibility: every cognition has something about it that is paradoxical or ineffable in another fashion. Further combinations result in seven modes.

Jainas aside, disputes between idealists and realists dominate the earlier centuries of classical philosophy. A school of direct realists, Nyāya (Logic), argues that the intentionality of even a nonveridical cognition hits a feature of the world, albeit misplaced. When one misperceives mother-of-pearl as silver, the silver-hood of which one is aware exists elsewhere. Had one not experienced it previously, one would not misperceive in this way ("It's silver"). The mother-of-pearl misperceived as silver is real, and so, too, the silver-ness wrongly indicated. Buddhists and other classical idealists argue, in sharp contrast, that one's desires and interests shape one's perceptions and all determinate cognition. Illusion shows that there are no objects independent of consciousness, since the false is seen to appear as the true.

Regarding the nature of veridicality, realists tend to embrace varieties of a correspondence theory. A cognition is veridical just in case the object cognized is cognized as being some way it is in fact. Whether there need be congruence between the object as qualified (thing-ontological relation-property) and the cognition as structured (qualificandum-qualificative relation-qualifier) was debated for several centuries. Realist camps explain illusion in different ways. Prābhākara Mīmāṃsakas deny that the intentionality of cognitions ever in itself misfires. The problem lies in confusing a perceiving and a remembering occurring at the same time. Nyāya philosophers hold that a nonveridical cognition presents something in some way that it is not, analyzing the error, "That is silver," as perceptual. That is, according to them silver-hood is projected into the sensory flow by a dispositional misfiring, the thing being in fact shell. They say that the view that there are two cognitions occurring simultaneously, a perceiving and a remembering (along with a failure to notice the difference), is wrong for several reasons. A single cognition stream defines a person's mental life. The nonveridical cognition of shell presents the thing perceptually as silver such that one says of the thing in front, "That is a piece of silver," and reaches out to pick it up. The thing perceived as silver motivates one's effort and action (including speech).

Prābhākara Mīmāṃsakas nevertheless join with Nyāya in seeing cognitive objects both as out there in the world and as structured: Property-bearers, which are enduring entities, are qualified by properties, some of which change (e.g., color) and some of which are essential to the thing qualified (e.g., cow-hood or being earthen). Cognition is similarly structured on the Nyāya theory, presenting qualificandum as qualifier. Thus, when there is a match between how an object is presented cognitionwise with the thing as it is in the world, the cognition is true.

Buddhists and other idealists tend to adopt a pragmatic theory. A cognition is veridical just in case it proves workable in helping one get what one wants and avoid what one wants to avoid. Realists agree that cognition is in this way useful and that sometimes one knows that a cognition is true by inferring its truth from the success of the action it guided. But realists see the nature of truth as correspondence. The Buddhists see workability not just as a mark of the truth but as truth itself. One calls cognitions true that make one successful, and false those that lead to frustrated efforts instead. Insofar as cognitive contents or indications are verbalizable, they are useful fictions, since the real is unverbalizable, knowable only through direct perception. Direct perception has unique particulars as object, not the general concepts contemplated by the mind. Concepts are mental constructions,

and what one says depends on mental projections on things that are ungeneralizable as things in themselves, as self-characterized particulars (*svalakṣaṇa*).

Later Buddhist logicians use an exclusion theory of concepts (*apoha*) in working out principles of logic and epistemology. The *apoha* theory seems motivated by Buddhist nominalism. A causally ordered series of particulars is conveniently designated a cow, though, strictly speaking, the series is a mental projection on fleeting particulars, none of which is either a cow or a non-cow. Designations exclude the least adequate concepts ("not a non-cow and so a cow"), according to one's desires and purposes; they do not apply directly to things in themselves. This view does not result in skepticism, since from one's everyday perspective truth is unproblematic. One distinguishes the veridical and the nonveridical by their perceived effects, satisfactions, or frustrations of desire through action undertaken on the basis of a belief (or mental construction, *kalpanā*).

GENUINE SOURCES OF KNOWLEDGE AND THEIR IMITATORS

Normative epistemology centers on the distinction between the veritably true cognition and its veritable knowledge source in distinction to the cognition that might seem to be veridical with the right pedigree but is in fact false and unreliable. Some kind of foul-up or deviation is to be suspected in a process resulting in the nonveridical. Though the evaluative paradigm is psychological and causal, inferential fallacies are discovered along with other epistemic faults. Indeed, long lists of fallacies appear in logic textbooks of both Buddhists and Hindus, including a majority of those known to the Aristotelian tradition and modern textbooks of critical reasoning. Veridicality is the ultimate touchstone, and disputants, given their differences on the nature of truth, rather surprisingly agree on fallacies and other concrete patterns of epistemic deficiency. Fallacies include nongenuine provers (*hetvābhāsa*), that is, evidence that seems to indicate a probandum in question but fails to secure the truth.

The distinction between the apparent (but false) and the genuine is made early in a metaphysical context, in the *Nyāyasūtra*, where it is used to refute the illusionist who would deny the reality of everyday objects. Things could be unreal or nonexistent, like dream objects. The epistemologist's knowledge source may itself be an illusion. Vātsyāyana (c. 400) points out in his *Nyāyasūtra* commentary (4.2.34) that the concept of the apparent whatever (as an apparent person that is really a post misperceived in the distance) presupposes the concept of the genuine variety (formed from previous experiences of persons). The apparently F could not be recognized without knowledge of things that are F genuinely. Thus, the concept of the illusory is parasitic on that of the veridical. If all cognitions were false, the cognition of the falsity would also be false. This is nonsense. Falsity requires an appreciation of truth. Thus, there is no reason to think that all objects and knowledge sources could be pretenders.

Despite such metaphysical argument, it is in epistemology where the distinction is most exploited. What is a genuine knowledge source (*pramāṇa*) as distinct from the imitator or pseudo (*ābhāsa*, thus *pramāṇābhāsa*)? People are subject to cognitive error of several types including logical error (*anumānābhāsa*), of which the *hetvābhāsas* (apparent [but false] reasons or provers) are the most discussed. Illusion is apparent (but false) perception (*pratyakṣābhāsa*). Understanding a false statement and being misled by the testimony of the deluded or of a deceiver, which is a form of *śābdābhāsa* (apparent [but false] testimony), will be treated separately later on. In general, if a cognition that appears to be, for example, perceptual from a first-person point of view is nonveridical (however defined), it is no result of perception as a genuine knowledge source, but of a cousin process, a close cousin, perhaps, indistinguishable from the real McCoy by the cognizer at the time. Much effort, under different flags, goes into trying to specify the features of cognitive processes that are marks of the one or the other, the genuine truth-generator or the imitator. The issues are complex, as can be guessed simply from the fact that at least thirty distinct definitions of truth and falsity are examined by late classical philosophers.

FALSE STATEMENTS AS NONGENUINE TESTIMONY

Classical Indian theories of meaning are mainly referentialist, and it is interesting to see how a false statement is analyzed by the classical epistemologists. Such enquiry also connects with questions about the lack, in Indian ontologies, of an exact equivalent of Western philosophy's "proposition." What is said about false claims, statements that seem meaningful but fail to hit the facts? Only the Nyāya view will be laid out; other schools present variations.

A case of *śābdābhāsa* (pseudoknowledge from testimony) may be taken to originate in a false statement of a speaker that a hearer understands and accepts, having no reason not to. As with perceptual cognition where there is

no block, testimonial uptake and acceptance are normally fused. A blocker (*pratibandhaka*) would be, for example, the hearer knowing in advance the opposite or knowing the speaker is a liar or deluded, the statement not being syntactically well formed or meeting certain conditions called semantic expectation (one cannot understand the statement, for example, "He wets with fire," since wetting is done only with water). Given no blockage, the false statement has a role in the generation of the hearer's comprehending and accepting cognition, which is false.

Taking the objecthood of that cognition to be the target of inquiry (a homonym misunderstood as well as a lie could constitute the deviant source), the Nyāya philosopher analyzes it in much the same manner as with apparent perception. The way (*prakāra*) that an object, a qualificandum, is being cognized would indicate a qualifier that exists elsewhere than in the thing. The standard realist story about how qualifiers, which are real-world realities, form dispositions (*saṃskāra*), which are inappropriately aroused, is available here as with other forms of cognitive error. The peculiarity of testimonial pseudo-knowledge concerns the speaker's statement being a causal factor in the generation of the hearer's nonveridical testimonial cognition. Nevertheless, it is the result—how the hearer understands the statement—that is targeted in the standard account of apparent (but false) testimony.

HOW IS VERIDICALITY KNOWN?

Prominent in classical debates about veridical cognitions and their sources is the issue of how veridicality is known. Prābhākara Mīmāṃsakas and Vedāntins say there is a kind of self-certification (*svataḥprāmāṇya*) at least with respect to certain contents or a cognition's own occurrence. Nyāya philosophers and others say that certification requires apperception, a second-order awareness, and certification by inferential means. The nature of the justificational inferences becomes central. Bhāṭṭa Mīmāṃsakas propose that while every cognition wears veridicality on its face—at least one assumes veridicality as a default—decertification is possible. Vedāntins tend to insist that there is a self that is essentially self-aware and the precondition of all cognition and experience. They view the other-certificationalists (*parataḥprāmāṇyavādin*) as confused about self-knowledge, though they may get the story right about knowledge of the external world, at least provisionally right, until the dawning of spiritual knowledge (*vidyā*).

On all views, confidence in a cognition's truth prompts effort and action; there are differences about whether the confidence has to be in some sense self-conscious. Realists of the two-cognition persuasion on illusion support a self-certificationism by taking a noncongruent correspondence view of the nature of truth. Idealists, too, often attack the qualificandum-qualifier structure supposed by Nyāya.

In Nyāya certification is said to proceed in three ways. First, a knowledge source can be identified by intrinsic features and in relation to a cognition in question as its result. Second, a cognition's veridicality can be certified with respect to its fruit, success of effort and action—a way that is also tied to causal relations and that is accepted by practically all disputants. The third procedure involves typifying. As mentioned, a cognition belongs to a type in virtue of its objecthood, its indicating, for instance, "*a* is F." Such objecthood can be shared with other cognitions, belonging to other people and to the cognizing subject at other times. So once a cognition as specified by its objecthood has been certified, a later cognition known to be a token of that type would also be certified.

Self-certificationists say that certification rides piggyback on apperception or whatever the way it is that a particular cognition is itself cognized. It appears that in this way ethical prescriptions of scripture can be upheld. They require no external justification. Certain Buddhists admit a form of certificational inference that looks like a kind of *a priori* knowledge, whereas Nyāya philosophers view all inference as depending crucially on prior perceptions.

Against the other-certificationists it is argued that, given that veridicality is in question, no certification would be possible, since only a cognition known to be veridical could possibly provide certification. If a certificational inference is required to show that a cognition is veridical, then there would have to be another inference to certify it, and one lands in an impossible regress and skepticism. Without the possibility of knowing that a cognition is veridical, trust in cognition would fly away. However, normally one does trust one's cognition, as is proved by one's behavior. Thus, however a veridical cognition is itself known or cognized, in that way its veridicality is also known, argue the self-certificationists. Other-certificationists respond by agreeing that an assumption of veridicality is a cognitive default, such that a cognition normally would not require certification to spark unhesitating effort and action. A cognition may nevertheless be called into doubt by good reasons, reasons that make one desist and reconsider.

Pseudocertification, on the Nyāya view, is possible but the presumption is also against it. Pseudocertification is certification that seems right from a first-person perspective but is misleading in fact. Apparent certification can be defeated (*bādhita*) by one's coming to learn something that undermines or rebuts a putatively certificational pseudoinference, whereas genuine certification requires that there be no ultimate defeater (*bādhaka*) in fact, that is, that one's evidence for regarding a cognition as veridical would hold no matter what else one comes to know. Established positions (*siddhānta*) serve as winnowing devices, and what one already knows can prevent wrong cognitions from arising. But one is not infallible. Just about any cognition, including an apparent certification, can prove to be wrong. But cognition of a cognition's veridicality, as distinct from a first-order assumption of truth, presents a higher barrier to doubt. Not only would there have to be good reasons to doubt the original cognition but also further reasons to question its certification.

The realist admission of a fallibilism that has few exceptions leaves the door wide open for the Advaitin nonrealist. Late Advaita Vedānta develops its two- or three-truth theory in sophisticated polemics where the Advaitin takes a minimalist position about the Upanishadic truth that Brahman is everything. World description may be left to the realists (science). The way that Brahman is the world is not statable (cognizable) in language that conflicts with statements (cognitions) about everyday things. Realism holds only provisionally.

See also Knowledge in Indian Philosophy; Logic, History of: Logic and Inference in Indian Philosophy; Meditation in Indian Philosophy; Mind and Mental States in Indian Philosophy; Negation in Indian Philosophy; Philosophy of Language in India; Universal Properties in Indian Philosophical Traditions.

Bibliography

Chakrabarti, Arindam. "Telling as Letting Know." In *Knowing from Words: Western and Indian Philosophical Analysis of Understanding and Testimony*, edited by Bimal Krishna Matilal and Arindam Chakrabarti. Dordrecht, Netherlands: Kluwer Academic, 1994.

Chakrabarti, Kisor Kumar. "Some Remarks on Indian Theories of Truth." *Journal of Indian Philosophy* 12 (4) (1984): 339–356.

Ganeri, Jonardon. *Semantic Powers: Meaning and the Means of Knowing in Classical Indian Philosophy*. Oxford, U.K.: Clarendon Press, 1999.

Hayes, Richard P. *Dignāga on the Interpretation of Signs*. Dordrecht, Netherlands: Kluwer Academic, 1987.

Krishna, Daya, et al., eds. *Samvāda: A Dialogue between Two Philosophical Traditions*. New Delhi, India: Indian Council of Philosophical Research, 1991.

Matilal, Bimal Krishna. *The Character of Logic in India*, edited by Jonardon Ganeri and Heeraman Tiwari. Albany: SUNY Press, 1998.

Matilal, Bimal Krishna. "A Realist View of Perception." In *The Philosophy of P. F. Strawson*, edited by Pranab Kumar Sen and Roop Rekha Verma. New Delhi, India: Indian Council of Philosophical Research, 1995.

Mohanty, Jitendra Nath. *Reason and Tradition in Indian Thought: An Essay on the Nature of Indian Philosophical Thinking*. Oxford, U.K.: Clarendon Press, 1992.

Nyāyasūtra (Nyāyadarśanam). With four commentaries, the *Nyāyasūtra-bhāṣya* of Vātsyāyana, the *Nyāyasūtrāvārttika* of Uddyotakara, the *Nyāyasūtravārttikatātpāryaṭīkā* of Vācaspati Miśra, and the *Vṛtti* of Viśvanātha, edited by A. M. Tarkatirtha, Taranatha Nyayatarkatirtha, and H. K. Tarkatirtha. New Delhi, India: Munshiram Manoharlal, 1985.

Phillips, Stephen H. *Classical Indian Metaphysics: Refutations of Realism and the Emergence of "New Logic."* Chicago: Open Court, 1995.

Phillips, Stephen H., and N.S. Ramanuja Tatacharya. *Epistemology of Perception: Gageśa's "Tattvacintāmaṇi."* Vol. I, *pratyakṣa-khaṇḍa, Introduction, Translation, and Commentary*. New York: American Institute of Buddhist Studies, 2004.

Potter, Karl H. "Does Indian Epistemology Concern Justified True Belief?" *Journal of Indian Philosophy* 12 (4) (1984): 307–328.

Ram-Prasad, Chakravarthi. *Advaita Epistemology and Metaphysics: An Outline of Indian Non-realism*. London: RoutledgeCurzon, 2002.

Rao, Srinivasa. *Perceptual Error: The Indian Theories*. Honolulu: University of Hawaii Press, 1998.

Ruegg, D. Seyfort. "Does the Mādhyamika Have a Thesis and Philosophical Position." In *Buddhist Logic and Epistemology: Studies in the Buddhist Analysis of Inference and Language*, edited by Bimal Krishna Matilal and Robert D. Evans. Dordrecht, Netherlands: D. Reidel, 1986.

Stephen H. Phillips (2005)

TRUTHLIKENESS

Truth is the aim of inquiry. Despite this, progress in an inquiry does not always consist in supplanting falsehoods with truths. The history of science is replete with cases of falsehoods supplanting other falsehoods. If such transitions are to constitute epistemic progress, then it must be possible for one falsehood better to realize the aim of inquiry—be more truthlike, be closer to the truth, or have more verisimilitude—than another. The notion of "truthlikeness" is thus fundamental for any theory of knowledge that endeavors to take our epistemic limitations seriously without embracing epistemic pessimism.

Given that truthlikeness is not only a much-needed notion but rich and interesting, it is surprising that it has attracted less attention than the simpler notion of truth. The explanation is twofold. First, if knowledge requires truth, then falsehoods cannot constitute knowledge. The high value of knowledge has obscured other epistemic values such as the comparative value of acquiring more truthlike theories. Second, if knowledge requires justification, then the notion of probability often takes center stage. There has been a long and deep confusion between the notions of subjective probability (seemingly true) and the notion of truthlikeness (similarity to the truth; Popper, 1972). This, together with the high degree of development of the theory of probability, obscured the necessity for a theory of truthlikeness.

Sir Karl Popper was the first to notice the importance of the notion (1972, chap. 10 and addenda). Popper was long a lonely advocate of both scientific realism and fallibilism: that, although science aims at the truth, most theories have turned out to be false and current theories are also likely to be false. This seems a bleak vision indeed and fails to do justice to the evident progress in science. Popper realized that the picture would be less bleak if a succession of false (and falsified) theories could nevertheless constitute steady progress toward the truth. Further, even if actually refuted by some of the data, the general observational accuracy of a false theory might be good evidence for the theory's approximate truth, or high degree of truthlikeness. That our theories, even if not true, are close to the truth, may be the best explanation available for the accuracy of their observable consequences (Boyd, 1983; Putnam, 1978, chap. 2).

Note that truthlikeness is no more an epistemic notion than is truth. How truthlike a theory is depends only on the theory's content and the world, not on our knowledge. The problem of our epistemic access to the truthlikeness of theories is quite different from the logically prior problem of what truthlikeness consists in.

Popper proposed a bold and simple account of truthlikeness: that theory B is more truthlike than theory A if B entails all the truths that A entails, A entails all the falsehoods that B entails, and either B entails at least one more truth than A or A entails at least one more falsehood than B (Popper, 1972).

This simple idea undoubtedly has virtues. Let the *Truth* be that theory that entails all and only truths (relative to some subject matter). On Popper's account the Truth is more truthlike than any other theory, and that is as it should be. The aim of an inquiry is not just some truth or other. Rather, it is the truth, the whole truth, and nothing but the truth about some matter—in short, the Truth—and the Truth realizes that aim better than any other theory. The account also clearly separates truthlikeness and probability. The Truth generally has a very low degree of (subjective) probability, but it definitely has maximal truthlikeness. Furthermore, the account yields an interesting ranking of truths—the more a truth entails, the closer it is to the Truth.

Popper's account also has some defects. For example, it does not permit any falsehood to be closer to the Truth than any truth. (Compare Newton's theory of motion with denial of Aristotle's theory.) But its most serious defect is that it precludes any false theory being more truthlike than any other (Miller, 1974; Tichý, 1974). The flaw is simply demonstrated. Suppose theory A entails a falsehood, say *f*, and we attempt to improve on A by adding a new truth, say *t*. Then the extended theory entails both *t* and *f* and hence entails their conjunction: *t*&*f*. But *t*&*f* is a falsehood not entailed by A. Similarly, suppose A is false and we attempt to improve it by removing one of its falsehoods, say *f*. Let *g* be any falsehood entailed by the reduced theory B. Then $g \supset f$ is a truth entailed by A but not B. (If B entailed both *g* and $g \supset f$, it would entail *f*.) So truths cannot be added without adding falsehoods, nor falsehoods subtracted without subtracting truths.

Maybe this lack of commensurability could be overcome by switching to quantitative measures of true and false logical content. Indeed, Popper proposed such accounts, but the problem they face is characteristic of the content approach, the central idea of which is that truthlikeness is a simple function of two factors—truth-value and logical content/strength (Kuipers, 1982; Miller, 1978). If truthlikeness were such a function, then among false theories truthlikeness would vary with logical strength alone. There are only two well-behaved options here: Truthlikeness either increases monotonically with logical strength, or else it decreases. But strengthening a false theory does not itself guarantee either an increase or a decrease in truthlikeness. If it is hot, rainy, and windy (h&r&w), then both of the following are logical strengthenings of the false claim that it is cold (~h): It is cold, rainy, and windy (~h&r&w); it is cold, dry, and still (~h&~r&~w). The former involves an increase, and the latter a decrease, in truthlikeness.

A quite different approach takes the likeness in truthlikeness seriously (Hilpinen, 1976; Niiniluoto, 1987; Oddie, 1981; Tichý, 1974, 1976). An inquiry involves a collection of possibilities, or possible worlds, one of which is actual. Each theory selects a range of possibilities

from this collection—that theory's candidates for actuality. A proposition is true if it includes the actual world in its range. Each complete proposition includes just one such candidate. The Truth, the target of the inquiry, is the complete true proposition—that proposition that selects the actual world alone. If worlds vary in their degree of likeness to each other, then a complete proposition is the more truthlike the more like actuality is the world it selects. This is a promising start, but we need to extend it to incomplete propositions. The worlds in the range of an incomplete proposition typically vary in their degree of likeness to actuality, and the degree of truthlikeness of the proposition should be some kind of function thereof: average likeness is a simple suggestion that yields intuitively pleasing results. (For a survey, see Niiniluoto, 1987, chap. 6.) The framework can also be used in the analysis of related notions such as approximate truth or closeness to being true (Hilpinen, 1976; Weston, 1992).

There are two related problems with this program. The first concerns the measure of likeness between worlds. It would be a pity if this simply had to be postulated. The second concerns the size and complexity of worlds and the number of worlds that propositions typically select. Fortunately, there is available a handy logical tool for cutting the complexity down to a finite, manageable size (Niiniluoto, 1977; Tichý, 1974, 1976). We can work with kinds of worlds rather than whole words. The kinds at issue are specified by the constituents of first-order logic (Hintikka, 1965), a special case of which are the maximal conjunctions of propositional logic (like h&r&w, ~h&r&w, ~h&~r&~w). Constituents have two nice features. First, each depicts in its surface structure the underlying structure of a kind of world. And, second, like the propositional constituents, they are highly regular in their surface structure, enabling degree of likeness between constituents to be extracted. (The world in which it is cold, rainy, and windy [~h&r&w] is more like the world in which it is hot, rainy, and windy [h&r&w] than it is like the world in which it is cold, dry, and still [~h&~r&~w]. In the propositional case, just add up the surface differences.) Since every statement is logically equivalent to a disjunction of constituents, we have here the elements of a quite general account of truthlikeness, one that can be extended well beyond standard first-order logic (Oddie, 1986, chap. 5).

Not just any features count in a judgment of overall likeness. Such judgments clearly presuppose a class of respects of comparison. The possibilities specified by h&r&w and ~h&r&w differ in one weather respect and agree on two, whereas those specified by h&r&w and ~h&~r&~w differ in all three. But now consider the following two states (where ≡ is the material biconditional): hot ≡ rainy, and hot ≡ windy. The possibility specified by h&r&w can equally be specified by h&(h ≡ r)&(h ≡ w); ~h&r&w by ~h&~(h ≡ r)&~(h ≡ w); and ~h&~r&~w by ~h&(h ≡ r)&(h ≡ w). Counting differences in terms of these new features does not line up with our intuitive judgments of likeness. Unless there is some objective reason for counting the hot-rainy-windy respects rather than the hot-(hot ≡ rainy)-(hot ≡ windy) respects, truthlikeness (unlike truth) seems robbed of objectivity.

This is the main objection to the likeness program (Miller, 1974). If sound, however, it would reach far indeed, for perfectly analogous arguments would establish a similar shortcoming in a host of important notions—similarity in general, structure, confirmation, disconfirmation, fit of theory to data, accuracy, and change (Oddie, 1986, chap. 6). The advocate of the objectivity of such notions simply has to grasp the nettle and maintain that some properties, relations, and magnitudes are more basic or fundamental than others. Realists, of course, should not find the sting too sharp to bear.

See also Aristotle; Confirmation Theory; Newton, Isaac; Philosophy of Science, History of; Philosophy of Science, Problems of; Popper, Karl Raimund; Propositions; Putnam, Hilary; Realism; Truth.

Bibliography

Boyd, R. "On the Current Status of the Issue of Scientific Realism." *Erkenntnis* 19 (1983): 45–90.

Hilpinen, R. "Approximate Truth and Truthlikeness." In *Formal Methods in the Methodology of Empirical Sciences,* edited by M. Przelecki, K. Szaniawski, and R. Wojcicki. Dordrecht: Reidel, 1976.

Hintikka, J. "Distributive Normal Forms in the Calculus of Predicates." In *Formal Systems and Recursive Functions,* edited by J. N. Crossley and M. A. Dummett. Amsterdam: North-Holland, 1965.

Kuipers, T. A. F. "Approaching Descriptive and Theoretical Truth." *Erkenntnis* 18 (1982): 343–387.

Kuipers, T. A. F., ed. *What Is Closer-to-the-Truth?* Amsterdam: Rodopi, 1987.

Miller, D. "Distance from the Truth as a True Distance." In *Essays on Mathematical and Philosophical Logic,* edited by J. Hintikka, I. Niiniluoto, and E. Saarinen. Dordrecht: Reidel, 1978.

Miller, D. "Popper's Qualitative Theory of Verisimilitude." *British Journal for the Philosophy of Science* 25 (1974): 166–177.

Niiniluoto, I. "On the Truthlikeness of Generalizations." In *Basic Problems in Methodology and Linguistics,* edited by R. E. Butts and J. Hintikka. Dordrecht: Reidel, 1977.

Niiniluoto, I. *Truthlikeness.* Dordrecht: Reidel, 1987.

Oddie, G. *Likeness to Truth.* Dordrecht: Reidel, 1986.
Oddie, G. "Verisimilitude Reviewed." *British Journal for the Philosophy of Science* 32 (1981): 237–265.
Popper, K. R. *Conjectures and Refutations,* 4th ed. London, 1972.
Putnam, H. *Meaning and the Moral Sciences.* London: Routledge and Kegan Paul, 1978.
Tichý, P. "On Popper's Definitions of Verisimilitude." *British Journal for the Philosophy of Science* 25 (1974): 155–160.
Tichý, P. "Verisimilitude Redefined." *British Journal for the Philosophy of Science* 27 (1976): 25–42.
Weston, T. "Approximate Truth and Scientific Realism." *Philosophy of Science* 59 (1992): 53–74.

Graham Oddie (1996)

TRUTH-TABLES

See *Logical Terms, Glossary of*

TSCHIRNHAUS, EHRENFRIED WALTER VON

(1651–1708)

Ehrenfried Walter von Tschirnhaus (or Tschirnhausen), the German mathematician and physicist, was born in Kieslingswalde, near Görlitz, and became count of Kieslingswalde and Stolzenberg. He studied mathematics at Görlitz and at the University of Leiden, where the Cartesian philosophers Adriaan Heereboord and Arnold Geulincx were teaching. After serving with the Dutch in 1672 during a war with France, Tschirnhaus studied further in Leiden and in Germany, and in 1674 he traveled to London, Paris, Rome, Sicily, and Malta. He met Benedict de Spinoza in Holland, English scientists in London, and he undoubtedly met Cartesian philosophers and scientists such as Jacques Rohault and Pierre-Sylvain Régis in Paris. Tschirnhaus finally settled down in Kieslingswalde. He established several factories for manufacturing glass and for grinding magnifying glasses, and was associated with J. F. Böttger in the development of Meissen porcelain. Tschirnhaus published various essays on mathematics and optics in the *Acta Eruditorum* from 1682 to 1698, and a philosophical treatise, *Medicina Mentis* (Amsterdam, 1687; 2nd ed. revised, Leipzig, 1695; reprinted with introduction by W. Risse, Hildesheim, 1964), on methodology, logic, and theory of knowledge, which also explained some of his geometrical discoveries.

Medicina Mentis followed Tschirnhaus's scientific interests; but some general features of the treatise were derived from Cartesianism, Spinoza, English empiricism, and, in some respects, from Gottfried Wilhelm Leibniz. Tschirnhaus's "mental medicine" was intended as a method of discovering rational truth as a basis of a happy life. Only true knowledge can tame the passions, which are the source of error and therefore of unhappiness.

Knowledge comes only from the senses, but purely sensible knowledge—which Tschirnhaus called imagination—is passive, approximate, and relative, and must be governed by rigid precepts. Reason abstracts from imagination, producing universal and strict concepts. The intellect considers things "as they exist in themselves"; that is, it penetrates their "real nature" and connects in one whole the real thing and its sensible and abstract representations. Reason operates by analysis, intellect by synthesis.

Only intellectual knowledge can reach truth and be communicated. Falsehood arises when intellect works like imagination. The criterion of truth is "what can be conceived"—that is, ideas insofar as they may be connected or not connected with one another. This criterion does not rest simply on an abstract rule to be applied in each case, but on the possibility of connecting ideas in a comprehensive system. But for Tschirnhaus this system was not, as for the rationalists, a closed and independent cognitive order. He considered the intellectual faculty to be the source of logical truth. But metaphysical truth comes from experience, and it is truth insofar as it has been deduced from experience by reasoning conforming to logical standards, and insofar as it is confirmed "through evident experiments."

Intellectual knowledge operates by elaborating simple concepts, or "definitions"; by deducing simple properties, or "axioms," from them; and by connecting the definitions in all possible ways to produce "theorems." Tschirnhaus held that definition is real. It is a knowledge of causes that enables us to reproduce the object. In its highest stages intellectual knowledge is knowledge of the natural world. Science is a whole, and should conform to the methodological ideal of mathematical clarity. Physics is the foundation of the other sciences. By demonstrating the rationality and necessity of all events, physics leads us to recognize divine providence. Human freedom arises from the command of God.

Although Tschirnhaus's *Medicina Mentis* was quite famous in its own day and its methodology was an important source of Christian Wolff's ideas, it exerted no direct influence on the German Enlightenment.

See also Cartesianism; Empiricism; Geulincx, Arnold; Leibniz, Gottfried Wilhelm; Régis, Pierre-Sylvain; Rohault, Jacques; Spinoza, Benedict (Baruch) de; Wolff, Christian.

Bibliography

Campo, Mariano. *Cristiano Wolff e il razionalismo precritico*. Milan: Vite e pensiero, 1939. Vol. I, Part 1, Ch. 2.

Klüger, Richard. *Die pädagogischen Ansichten des Philosophen Tschirnhaus*. Leipzig, 1913.

Kulstad, Mark A. "Leibniz, Spinoza, and Tschirnhaus: Metaphysics a Trois, 1675–1676." In *Spinoza: Metaphysical Themes*, edited by Olli Koistinen. New York: Oxford University Press, 2002.

Kunze, A. "Lebensbeschreibung des E. W. von Tschirnhaus und Wurdigung seiner Verdienste." *Neues Lausitzer Magazin* 43 (1866).

Radetti, G. "Cartesianesimo e spinozismo nel pensiero di E. W. von Tschirnhaus." *Rendiconti della Classe di Scienze morali, storiche e filologiche della Accademia Nazionale dei Lincei*, Sec. 6, 14 (5 and 6) (1938): 566–601.

Reinhardt, K. *Beiträge zur Lebensgeschichte von E. W. von Tschirnhausen*. Meissen, Germany, 1903.

Schönfeld, Martin. "Dogmatic Metaphysics and Tschirnhaus's Methodology." *Journal of the History of Philosophy* 36 (1) (1998): 57–76.

Van Peursen, C. A. "E. W. Von Tschirnhaus and the "Ars Inveniendi." *Journal of the History of Philosophy* 54 (3) (1993): 395–410.

Verweyen, Johannes. *E. W. von Tschirnhaus als Philosoph*. Bonn, 1906.

Weissenborn, H. *Lebensbeschreibung des E. W. von Tschirnhaus*. Eisenach, Germany, 1866.

Giorgio Tonelli (1967)
Bibliography updated by Tamra Frei (2005)

TUNG CHUNG-SHU

See *Dong Zhongshu*

TURGOT, ANNE ROBERT JACQUES, BARON DE L'AULNE

(1727–1781)

The French statesman, economist, and philosopher of history Anne Robert Jacques Turgot, Baron de l'Aulne, was born in Paris. He began formal theological training in 1743, anticipating a career in the church. As a young scholar at the Sorbonne (1749–1751) he showed brilliant promise in several writings on the philosophy of history. In 1752 he left the service of religion to become a magistrate, and from 1753 to 1761 he fulfilled the legal and administrative duties of a master of requests. His writings in this period included contributions to the *Encyclopédie* in metaphysics, linguistics, science, economics, and political theory, as well as short writings over a similarly broad range of fields, but his contemplated major work on the history of human progress never materialized. From 1761 to 1774 he served as the enlightened intendant (royal administrator) of Limoges; in this period and later, economic subjects predominated in his writings. Appointed minister of marine by Louis XVI in 1774, he was very shortly afterward transferred to the crucial position of comptroller general of finance. In this post Turgot instituted economies, corrected abuses in the taxation system, established free grain trade within France, and suppressed the guilds and the labor services. Opposition at court and in the Parlement of Paris, and the withdrawal of royal support, led to his resignation after twenty months (1776), thus ending the last attempt at thoroughgoing reform of the ancien régime in France before the Revolution.

ECONOMIC AND SOCIAL THEORIES

Turgot's economic theories are expressed most fully in his *Réflexions sur la formation et la distribution des richesses* (1766, published serially 1769–1770; translated as *Reflections on the Formation and the Distribution of Riches,* New York, 1898). In this and other works his basic principles are essentially physiocratic: The sole ultimate source of wealth is land, and only the growth and the unhindered flow of capital can create prosperity. Assuming that the French economy would continue to be largely agrarian, Turgot advocated a gradual simplification and moderation of taxation, looking toward the day when only landowners would be taxed, on the basis of a careful assessment of their profits, and when restrictions and impositions upon commerce and industry might be altogether abolished.

Turgot's general political thought, based on a belief in paternalistic, enlightened monarchy, is of less interest than his two *Lettres à un grand vicaire sur la tolérance* (Letters to a grand vicar on toleration, 1753, 1754; in *Oeuvres,* Vol. I) concerning governmental toleration of religion. In these letters he defended a broad toleration of different faiths but maintained that the state may offer special protection to the "dominant" or most numerous religion, as a useful guide to men in their uncertainties. He nevertheless held that some sects—those too rigid, irrational, morally or socially burdensome, or politically subversive—are not worthy of such protection, but

should simply be tolerated; Roman Catholicism, he noted, might be considered by some to be such a sect. The dogma of infallibility is dangerous if it is false, and "it is certainly false or inapplicable when the exercise of infallibility is confided to those who are not infallible, that is to princes and governments" (*Oeuvres,* Vol. I, p. 425). Intolerance, unworthy of a gentle and charitable Christianity, must in any case be eradicated, for the rights of society are not greater than those of individuals, and individual conscience is no proper concern of government.

PHILOSOPHY OF HISTORY

To the philosopher, Turgot's importance may well derive from his early writings on the theory of history, notably his *Tableau philosophique des progrès successifs de l'esprit humain* (Philosophic panorama of the progress of the human mind, 1750; in *Oeuvres,* Vol. I), and his "Plan de deux discours sur l'histoire universelle" (Plan of two discourses on universal history; c. 1750, in *Oeuvres,* Vol. I). Upon the basis of contemporary psychological sensationalism, and with a nod to Providence, Turgot constructed a broad theory of human progress reflecting past theories and foreshadowing later ones.

In contrast to the phenomena of the world of nature, trapped in unprogressive cycles of birth and death, Turgot postulated the infinite variability and indeed the perfectibility of humankind. In the past and in the future, as knowledge and experience accumulate, man's reason, passions, and freedom permit him to escape from the repetitive cycles of external nature. Movement and change give rise to new relationships, and thus all experience is instructive; even passion and error, calamity and evil providentially contribute to humankind's advance. Indeed, the ambitions and the vices of men and the barbarities of warfare, however morally reprehensible, may often rescue humankind from stagnation or mediocrity.

The vital medium of progress, wrote Turgot, is the process of human communication. Ideas deriving from sensations are developed through the use of signs, pictures, and especially language, by which knowledge and experience are transmitted and augmented from generation to generation. Since above all it is the man of genius who can grasp the implications and make articulate the lessons of experience, it is society's duty to encourage natural genius and to heed its advice. "Moral" circumstances, such as the cultivation of genius, are more important in determining the extent and nature of progress than are such physical circumstances as climate.

Progress is uneven throughout man's history. Moreover, it varies necessarily in the different areas and aspects of human activity, such as science, technology, morality, and the arts. Progress in the arts, for example, is always radically limited by the nature of man himself, since the goal of the arts is pleasure alone, whereas speculative scientific knowledge can be as infinite as the natural universe. And each area of activity has its own rules of progress. In his discussion of scientific progress, Turgot suggested three historical stages of development (anticipating Auguste Comte's system): the anthropomorphic or supernatural, the abstract or speculative, and the empirical-mathematical.

For Turgot the broad tempo of progress was increasing in the mid-eighteenth century; indeed, despite instances of momentary or partial decadence, any wholesale retrogression of humankind was now impossible. Surely, he wrote, the general momentum of science, buttressed by mathematics, was irreversible. Yet Turgot, especially in his later years, had frequent doubts, and he was well aware of the forces of error and evil in the world, both in the past and in the happier future. The historical continuity so much stressed in his writings in fact ruled out any immediate, thorough renovation of humankind. Certainly the future would not bring the radical break with a deplorable past that was intimated in the thought of many another writer of the Enlightenment. Because the element of empiricism was seldom wholly absent in Turgot, his historical thought, although undoubtedly optimistic, was never unreservedly utopian.

See also Encyclopédie; Enlightenment; Philosophy of History; Progress, The Idea of.

Bibliography

Turgot's works were published as *Oeuvres de Turgot et documents le concernant,* edited by Gustave Schelle, 5 vols. (Paris: Alcan, 1913–1923). See also: *Turgot on Progress, Sociology, and Economics: A Philosophical Review of the Successive Advances of the Human Mind,* translated and edited by Ronald L. Meek (Cambridge, U.K.: Cambridge University Press, 1973).

For literature on Turgot, see Douglas Dakin, *Turgot and the Ancien Régime in France* (London: Methuen, 1939), which has an extensive bibliography; Edgar Faure, *La disgrâce de Turgot* (Paris: Gallimard, 1961), which also has an extensive bibliography; Philippe Fontaine, "Turgot's 'Institutional Individualism,'" *History of Political Economy* 29 (1) (1997): 1–20; Malcolm Hill, *Statesman of the Enlightenment: The Life of Anne-Robert Turgot* (London: Othila Press, 1999); Frank E. Manuel, *The Prophets of Paris* (Cambridge, MA: Harvard University Press, 1962), Ch. I, pp. 11–51; and Robert Nisbet, "Turgot and the Contexts of Progress," *Proceedings of the American Philosophical Society* 119 (3) (1975): 214–222.

Henry Vyverberg (1967)
Bibliography updated by Philip Reed (2005)

TURING, ALAN M.
(1912–1954)

Alan Mathison Turing was born June 23, 1912, in London and died June 7, 1954, at his home near Manchester. He suffered the conventional schooling of the English upper-middle class, but defeated convention by becoming a shy, eccentric but athletic Cambridge mathematician. The Second World War transformed Turing's life by giving him a crucial role in breaking German ciphers, with particular responsibility for the Atlantic war. Thereafter Turing led the design of electronic computers and the program of artificial intelligence. In 1950 he began another career as a mathematical biologist, but was assailed by prosecution for homosexuality. His last two years, though overshadowed by punishment and security risk status, saw vigorous and defiant work until his death by cyanide poisoning.

Turing's paper *Computing Machinery and Intelligence* appeared in 1950. This, his only contribution to a philosophical journal, was to become one of the most cited. He considered the question "can a machine think" and gave an argument that broke with all previous speculation about homunculi and robots, and from all earlier discussion of mind, matter, freewill, and determinism. It was based on his own elucidation of mathematical computability, as achieved in 1936. It also reflected his unique experience with practical computation.

Turing's computability arose from the long search for a logical basis to mathematics, in which Bertrand Russell had played a prominent part. In 1931 Gödel showed that no formal proof scheme such as Russell had envisaged could encapsulate mathematics. In 1935 Turing seized on the further outstanding question, of whether there could be a definite method for deciding whether a given proposition was susceptible to formal proof. The question turned on finding a definition of "method," and this Turing supplied with his "Turing machine" construction. This was mathematically equivalent to the definition of "effectively calculable" offered by the logician Alonzo Church a little earlier, but Church accepted that Turing's argument gave it a natural and compelling rationale. Their assertions, taken together, are referred to as the Church-Turing thesis. On accepting this thesis, it follows that there is no effective method for deciding provability. Many other mathematical questions of decidability have likewise been resolved.

Turing's thesis was based on analyzing the actions of a human mind when following a rule, and translating it into formal actions of reading and writing. More generally, Turing's formalism was intended to capture what could be carried out by a "purely mechanical process," interpreting this as one that "could be carried out by a machine." Thus Turing found a new connection between the mind and the material world. On the one hand, he gave a new logical analysis of mental operations, but on the other hand, the criterion of "effectiveness" implied something that could be implemented physically.

As mathematics, Turing's argument meant encoding operations *on* symbols *by* symbols, rather as Gödel encoded theorems *about* numbers *by* numbers. Turing exploited this by describing a "universal" machine, which could do the work of any Turing machine. This concept led directly to the modern computer in which program and data are stored and manipulated alike as symbols. In 1936 Turing had no technology for implementing this idea. He did further important work exploring the mathematics of uncomputability, which touched on the role of human intuition in mathematics. He also discussed the foundations of mathematics with Wittgenstein. But then six years of war work brought him back to the "universal machine." He had gained the experience of advanced electronics and hence the means of putting his idea into practice.

Turing's central interest in computing lay in its role in investigating the nature of the mind. He described his post-war computer plan as "building a brain," and asserted with increasing confidence that *any* action of the mind, including creative acts, could be described as computable operations. Turing's sophisticated cryptanalytic work had impressed him with the apparently limitless scope of the computable. He now discounted arguments derived from Gödel's theorem suggesting a noncomputable aspect to the human mind. He emphasized that any computable operation could be implemented on a single universal machine: the computer. Hence, the computer could rival human intelligence.

Turing's 1950 paper summarized these arguments for a wide readership. His underlying view assumed a physical basis for Mind, but rather than argue for this he appealed to an argument from external observation. He held that a computer exhibiting the appearance of intelligence should be credited with intelligence. He thus avoided discussing the reality of consciousness, and sought to sidestep its traditional philosophical primacy. Instead, he illustrated his "imitation game" with a provocatively wide view of "intelligence," and took pleasure in playing the role of a new Galileo, defying orthodox belief in the uniquely human nature of mind.

This "imitation game," the so-called "Turing Test" for intelligence, was not the only content of this paper. He also sketched a constructive program for Artificial Intelligence research, which he saw as a combination of "top-down" methods by programming and "bottom-up" methods using networks capable of developing functions through training. Turing saw self-modification in machines as a key analogy with human mental development. His doubts and reservations centered on the question of defining a valid line separating the mind from the external world with which it interacts.

Turing made a prophecy of progress within fifty years, which though cautiously expressed, still proved over-optimistic. Some artificial intelligence protagonists have come to see Turing's ambitious goal as a distraction from systematic research. But many thinkers have found it vital to continue Turing's arguments. Lucas revived the objection from Gödel's theorem, which Turing had dismissed. Hofstadter and Dennett then vigorously defended Turing's view. A new argument was made by Penrose. This shares with Turing a wholly materialist viewpoint, but holds that there must be uncomputable elements in the physics of the brain, arising from the reduction process of quantum mechanics. A late talk given by Turing indicates that he, too, considered this question, but death cut off the physical investigations he undertook in 1953 and 1954. The relationship of computability to physics, in particular to the material basis of mind, is the central question left by Turing's work.

As a human being Alan Turing was highly willful and far from soulless, yet he sought to mechanize will and mocked the concept of soul. He was highly original and resisted social conformity, yet attempted to explain creativity as a process of learning. Truthfulness was paramount to him, yet he committed himself to state secrets and defined intelligence by imitation. The paradoxical life and death of Alan Turing continue to fascinate.

See also Artificial Intelligence; Church, Alonzo; Computability Theory; Computing Machines; Gödel, Kurt; Logic Machines; Machine Intelligence; Russell, Bertrand Arthur William.

Bibliography

PRIMARY WORKS

The Essential Turing, edited by B. J. Copeland. Oxford: Oxford University Press, 2004.

Collected Works of A. M. Turing, edited by R. O. Gandy et al. Amsterdam: Elsevier, 1992–2001.

SECONDARY WORKS

Diamond, C., ed. *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge 1939*. Hassocks, U.K.: Harvester, 1976.

Herken, R., ed. *The Universal Turing Machine*. Oxford: Oxford University Press, 1988.

Hofstadter, D. R., and D. C. Dennett. *The Mind's I*. New York: Basic, 1981.

Hodges, A. *Alan Turing: The Enigma*. New York: Simon & Schuster, 1983.

Hodges, A. *Turing: A Natural Philosopher*. London: Routledge, 1997.

Penrose, R. *The Emperor's New Mind*. Oxford: Oxford University Press, 1989.

Teuscher, C., ed. *Alan Turing: Life and Legacy of a Great Thinker*. Berlin: Springer, 2004.

Andrew Hodges (2005)

TURING AND COMPUTABILITY

See *Logic, History of: Modern Logic: Since Gödel: Turing and Computability Theory*

TURING TEST

See *Machine Intelligence; Turing, Alan M.*

TWARDOWSKI, KAZIMIERZ

(1866–1938)

Kazimierz Twardowski had a twofold role in the recent history of philosophy. He had a decisive influence on Polish philosophy in the twentieth century; and at the turn of the twentieth century he contributed to the transformation of European philosophy in its search for new, intellectually responsible methods of philosophical inquiry. His conception of philosophy and his specific contributions to epistemology, philosophical psychology, and theory of science helped to pave the way for the emergence of phenomenology and of some forms of analytic philosophy.

Twardowski was born in Vienna. He studied philosophy at the University of Vienna, where he came under the influence of Franz Brentano. In 1892 he received a PhD degree from the university, and he became a lecturer there in 1894. In 1895 he was appointed to a chair of philosophy at the University of Lwów, where he taught until 1930.

Like Brentano, he wanted philosophy to be scientific, which to him meant a rejection of grandiose but nebulous speculation, an unrelenting war on conceptual confusion and linguistic obscurity, and a painstaking analysis of clearly defined problems, which through elimination of conceptual sloppiness, leads to empirically verifiable conclusions. No wholesale condemnation of metaphysics was intended by these methodological injunctions. Nevertheless, Twardowski was increasingly aware of the boundary beyond which the method of philosophy, as conceived by him, could not reach and beyond which a philosopher qua philosopher should remain silent.

More specifically, the basic philosophical science, avoiding both irresponsible speculation and skepticism, was to be the Brentanist "descriptive psychology," understood as a sort of empirical inquiry, but distinct from experimental psychology. Twardowski, however, went well beyond Brentano and contributed to the demise of psychologistic accounts of meaning and of psychologism in general. In an early and influential book, *Zur Lehre vom Inhalt und Gegenstand der Vorstellungen,* Twardowski introduced a sharp distinction between the mental act, its content, and its object. The distinction between content, which is mental and a part of a person's biography, and object, which is not, was overlooked by Brentano and the early Alexius Meinong but became crucial for Twardowski and led him to a general theory of objects of thought. These ideas influenced Meinong, Edmund Husserl, and to some extent Moritz Schlick, and through them much of early-twentieth-century philosophy. The difficulties of Twardowski's theory of objects, with its attending danger of overpopulating the Platonic heaven, led later to Stanisław Leśniewski's "ontology" and Tadeusz Kotarbiński's "reism." Twardowski's conclusions, far from supporting psychologism, implied a sharp separation of logic and philosophy from psychology. Moreover, the actual procedure of this "psychological investigation" did not look much like psychology either. Phenomenologists have seen in it the germ of the ideas that reappeared both in the later Husserl and in the realist branch of phenomenology. Up to a certain point, it is equally plausible to construe Twardowski's contributions as an early attempt to develop a philosophical psychology, in the sense of an examination of the logical geography of mental concepts.

Twardowski's later work included a further analysis of mental concepts; the formulation of a nonpsychologistic and non-Platonizing account of logic, based on the distinction between acts and their products; the extension of a similar line of reasoning to a general theory and classification of the sciences; and an examination, on several occasions, of various methodological issues of psychology. This included a critique of reductive materialism and a defense of introspection as a source of knowledge. One of his most influential works, "O tak zwanych prawdach względnych," was a lucid critique of relativism.

A strong sense of the scholar's social responsibilities, heightened by the special circumstances of Polish history, led Twardowski to devote more and more time to educational activities, to the detriment of his own work, but to the lasting benefit of Polish philosophy.

As a teacher, Twardowski transformed Polish philosophy and endowed it with a distinct style. He did not preach any particular weltanschauung, and his influence—not unlike that of G. E. Moore—was due less to his specific doctrines than to his way of doing philosophy, his qualities of character, his intellectual integrity, and the impact of his personal example. The school that he created was not linked by a common allegiance to any philosophical creed, but rather by a common acceptance of rigorous standards of professional excellence. Most of his pupils went their own independent ways, representing a wide spectrum of philosophical opinion, but they never ceased to express their gratitude to him. The best-known among them, Jan Łukasiewicz, Leśniewski, Kazimierz Ajdukiewicz, and Kotarbiński, differed from Twardowski methodologically in their emphasis on the philosophical relevance of symbolic logic. Twardowski's influence, transmitted by his numerous students—philosophers and nonphilosophers—went far beyond academic circles and, fostering the ethos of free and responsible inquiry in all areas of intellectual life, became a significant factor in the history of Polish culture.

Twardowski organized the teaching of philosophy in Poland, initiated regular meetings of philosophers, founded the first Polish psychological laboratory (1901), the Polish Philosophical Society (1904), and in 1911 the quarterly journal *Ruch Filozoficzny,* which he edited until his death. In 1935 he became the chief editor of *Studia Philosophica,* a periodical publishing works of Polish philosophers in foreign languages. He was also active as the editor of several different series of original works and translations, many of them inspired by him, such as Władysław Witwicki's masterful translations of Plato.

See also Brentano, Franz; History and Historiography of Philosophy; Husserl, Edmund; Kotarbiński, Tadeusz; Leśniewski, Stanisław; Łukasiewicz, Jan; Meinong, Alexius; Moore, George Edward; Phenomenology; Plato; Schlick, Moritz.

Bibliography

WORKS BY TWARDOWSKI

Idee und Perzeption. Vienna, 1892.

Zur Lehre vom Inhalt und Gegenstand der Vorstellungen. Vienna, 1894.

Wyobrażenia i pojęcia. Lwów, 1898.

"O tak zwanych prawdach względnych," *Księga Pamiątkowa Uniwersytetu Lwowskiego.* Lwów, 1900. Translated into German by M. Wartenberg as "Über sogenannte relative Wahrheiten." *Archiv für systematische Philosophie* 8 (4) (1902): 415–447.

Zasadnicze pojęcia dydaktyki i logiki. Lwów, 1901.

"Über begriffliche Vorstellungen." In *Beilage zum XVI Jahresbericht der Philosophischen Gesellschaft a. d. Universität zu Wien.* Leipzig, 1903.

O filzofji średniowiecznej. Lwów, 1910.

O metodzie psychologji. Warsaw, 1910.

O czynnościach i wytworach. Kraków, 1911.

O psychologji, jej przedmiocie, zadaniach i metodzie. Lwów, 1913.

Rozprawy i artykuły filozoficzne. Lwów, 1927.

Wybrane pisma filozoficzne. Edited by T. Czeżowski. Warsaw: Panstwowe Wydawn, Naukowe, 1965.

WORKS ON TWARDOWSKI

Discussions of Twardowski in English include T. Czeżowski, "Tribute to Kazimierz Twardowski on the 10th Anniversary of His Death," in *Journal of Philosophy* 57 (1960): 209–215; T. Czeżowski, "Kazimierz Twardowski as Teacher," in *Studia Philosophica* 3 (1948): 13–17; J. N. Findlay, *Meinong's Theory of Objects* (London: Oxford University Press, H. Milford, 1933); Roman Ingarden, "The Scientific Activity of Kazimierz Twardowski," in *Studia Philosophica* 3 (1948): 17–30; and Z. A. Jordan, *Philosophy and Ideology* (Dordrecht, Netherlands: Reidel, 1963).

A bibliography of writings by and on Twardowski until 1938, compiled by D. Gromska, can be found in *Ruch Filozoficzny* 14 (1938): 14–39. Additional bibliography can be found in Z. A. Jordan's book.

George Krzywicki-Herburt (1967)

TYCHE

See *Moira/Tychē/Anakē*

TYPE THEORY

Type theory, in one sense, is the view that some category of abstract entities—sets, in the simplest example, but there are analogous views of properties, relations, concepts, and functions—come in a hierarchy of levels, with an entity of one level applying to (having as members, or having as instances, or…) entities only of a lower level. Such a view gives an intuitively comprehensible picture of the universe of abstracta and provides a principled way of avoiding Bertrand Arthur William Russell's Paradox and its analogues. In a second sense, the term refers to any of a wide range of formal axiomatic systems embodying some form of the view. The present entry gives a short history of the view and a brief survey of the systems.

The systems are generally formulated in *many-sorted* quantificational logic, with a separate alphabet of quantified variables ranging over each type of entity. Axiomatically, they incorporate the rules of propositional logic (usually though not always classical) and of quantifier logic, the latter reduplicated for each alphabet of variables. Beyond this the most important axioms postulate the existence entities of the various types. For versions of Simple Type Theory, these are typically unrestricted comprehension principles: For any type, *t*, there is a set (or property, or …) of a higher type containing as members (or having as instances, or …) all the entities of type *t* satisfying an arbitrarily chosen formula of the language. For versions of Ramified Type Theory, this is restricted: Only such entities are postulated as can have their membership (or …) specified by formulas in which certain sorts of variables do not occur. References are given below to works in which precise formulations can be found; Alonzo Church is particularly helpful in this matter.

Type theory as a way of avoiding the set-theoretic paradoxes is one of Russell's great contributions to the study of the foundations of mathematics, but the idea of a hierarchy with sets (or set-like entities …) coming in levels is a bit older. Schroeder had presented a version of it, and Gödel Frege had based his foundational system on it. For Frege, the hierarchy of entities reflected the hierarchy of grammatical categories in an (idealized) language. At the bottom there were *objects*, the referents of singular terms. Predicates (either simple or complex) then stood for *concepts*, which he conceived of as so different from objects that it was an abuse of language to try to say anything of the two together (hence his avowedly nonsensical dictum about the concept horse not being a concept). Linguistic constructions with *blanks* that can be filled by simple or complex predicates—his prime examples were the first-order quantifiers "it holds of every object that it ___" and "it holds of at least one object that it ___"—he construed as a sort of second-level predicate and took to denote second-level concepts: entities as different from (first-level) concepts as they are from objects. And so (in principle; in practice he made little use of higher levels) on up. The grammar of his formal system reflected this: No term or variable for an item of one level was allowed to stand in the positions filled by terms or variables for

items of other levels. Had he rested content with this machinery, his formal system would have been a version of (what later came to be called) the Simple Theory of Types and demonstrably consistent.

To carry out his project of giving a logicist foundation for arithmetic, however, he had to prove that there were infinitely many objects, and to do this, he postulated that every concept had an object—its *Werthverlauf*—as a kind of shadow. These objects functioned essentially as sets: Frege was able to define membership by saying that one object was a member of a second iff the second was the *Werthverlauf* of a concept holding of the first. Since a *Werthverlauf* could have any objects whatever—including *Werthverlaufs!*—as members, the derived set theory was untyped. Frege was able to prove in it the existence of an infinite set, and to interpret (a variant of) the Peano axioms for Arithmetic. Russell was able to derive in it his contradiction. (There is a readable account of the derivation of Arithmetic in the untyped set theory in Hatcher 1982.)

Russell was not initially attracted to Frege's linguistic hierarchy. He wanted to formulate a general metaphysical theory and to describe the differences between horses and concepts by denying of the one the very same predicates he affirmed of the other. During his period of experimentation after the discovery of the paradoxes, he toyed with and rejected versions of type theory, finally coming to it by an indirect route.

Sets (Russell said *classes*) themselves—entities satisfying an axiom of extensionality—he was willing to give up as excess ontological baggage. A set is typically defined by giving an open formula that specifies its membership, and Russell preferred to think in terms of nonextensional entities designated by the formulas. He thought of sentences as standing for *propositions*, which he took to be complex entities built up out of the items designated by the words in a sentence in a way that paralleled the syntactic construction of the sentence. The items expressed by open formulas he called *propositional functions*: things which, when given some entity as argument, would yield the proposition that would be expressed by inserting a name of the argument in place of the free variable. Russell's Paradox, however, does not depend on the assumption of extensionality: A *naïve* theory of propositional functions is inconsistent in the same way as naïve set theory. If a consistent theory of propositional functions could be found, however, a theory of sets could be interpreted in it by contextual definition: Statements about sets would be interpreted as statements about propositional functions to which the differences between extensionally equivalent functions were irrelevant. Apparently almost immediately after seeing how to dispense with the strange entities he called *denoting concepts*—"On Denoting" eliminates them by giving a new analysis of the propositions he had thought of as containing them—he thought of what might be called a theory of *virtual* propositional functions, a theory in which, though neither classes nor propositional functions were postulated as entities, statements *apparently* referring to them could be formulated.

On this theory reference to a propositional function (say, *X is a horse*) would be replaced by reference to a pair of entities: one of the propositions that might have been taken as a value of the function (for example, *Bucephalus is a horse*) along with one of the component entities of that proposition (Bucephalus, in this case). The key notion was one of the *substitution* of an entity for one of the constituents of a proposition: Rather than saying that Traveler, for example, satisfies the propositional function *X is a horse*, on the new theory, we will say that the proposition obtained from *Bucephalus is a horse* by substituting Traveler for Bucephalus in it is a true one. (Note that this notion of substitution is not a syntactic one: We are substituting one flesh-and-blood horse for another in a proposition, construed as a complex but nonlinguistic entity. The developed formalization of the theory, of course, has provisions for substitution of names in the sentences expressing propositions!)

Since the place of a variable for propositional functions is taken by two variables for entities (one for a proposition, one for a *designated argument*), and a variable for a higher-level propositional function taking first-level propositional functions will similarly be replaced by three variables and so on, this theory gives the effect of a typed theory of propositional functions: When references to and quantifications over propositional functions are replaced with terms and variables for propositions and other entities recognized by the theory in the way sketched, it will be impossible to say that a higher-level item serves as an argument for a lower! (Russell described the theory in his 1906 essay, which he withdrew before publication. Remarkably, essentially the same system was developed again, apparently independently, several decades later. For discussions of Russell's theory and his reasons for abandoning it, see Peter Hylton [1980] and Gregory Landini [1998].)

Russell took propositions and propositional functions to be the objects of cognitive attitudes and the meanings of linguistic expressions as well as the fundamental objects of mathematics. In trying to formulate a

general theory of entities that could serve all these functions, he confronted not only the set-theoretic paradoxes but also those now classed as *semantic* or *intentional*, and they drove him to an even more restrictive form of type theory. This was the *Ramified Theory of Types*, first presented in Russell (1908) (a paper largely recycled in the introduction to *Principia Mathematica* two years later). On this theory there is a kind of double hierarchy. Propositional functions are classified not only by the arguments they take but also by the conceptual resources that go into their definitions: A propositional function can only have arguments of certain lower levels, but two propositional functions taking exactly the same arguments may be of different types if in formulating them (it is best, here, to think of the functions as the meanings of open formulas) one quantifies over entities of a different level.

Start with a domain of nonabstract or nonconceptual entities as a bottom level. The level of a propositional function will be at least one higher than that of its argument (or, in the case of a relational function, the highest level of its arguments). It will only have this minimum level, however, if no quantified variables are used in its formulation that range over entities of a higher level than its arguments, and the general rule is that the level of a propositional function is one greater than the highest level of its arguments or of the entities quantified over in its formulation, whichever is higher. (Propositions, since they do not take arguments, formed a single type on his earlier approaches, but the Ramified Theory divides them into a hierarchy based on the quantifications involved in them. This makes possible a quick dissolution of many semantic paradoxes: When Epimenides says that every proposition asserted by a Cretan is false, he asserts a proposition of a higher level than those he quantifies over, and so his assertion does not cover the proposition he himself has asserted.)

The Ramified Theory, though notationally complicated, has a perspicuous semantic interpretation which make its ontological commitments seem fairly innocent: Kurt Gödel, in a note added to reprintings of his 1944 essay, speaks of it as embodying a strictly nominalistic (or strictly antirealistic) kind of constructivism about abstract entities. From a mathematical point of view, it is a very weak theory: When it is supplemented by an Axiom of Infinity (stating that there are infinitely many objects of the lowest level), it suffices to derive a certain portion of elementary number theory, but only a restricted portion. In order to provide a foundation for classical mathematics, Russell added the Axioms of Reducibility. These maintained the type distinctions of the Ramified Theory (allowing Russell to appeal to them in dealing with the semantic and intentional paradoxes) but postulated the existence of enough *predicative* propositional functions (functions, that is, of the lowest possible level for their arguments) to provide a model of the mathematically stronger Simple Theory of Types, and the mathematical work of *Principia Mathematica* is then essentially conducted in the Simple Theory. (The clearest account of Ramified Type Theory and its use in analyzing the paradoxes is in Church [1976], to which sections 58 and 59 of Church [1956] can serve as an introduction.)

In the early 1920s two alternatives to the Ramified Theory were proposed. One was described in the introduction and appendices Russell wrote for the second edition of *Principia Mathematica* (1925). It was noted by Gödel in 1944, but otherwise seems to have been ignored until the 1990s. On this theory the two factors in the Ramified Theory's classification are separated. Each function has a simple type depending only on the arguments it takes, and also a ramification level determined by what entities are quantified over in its formulation. A function of higher simple type (one, that is, that can take functions as arguments) can be affirmed of any argument of the appropriate simple type, even an argument whose ramification level is higher than its own. Each quantified variable for propositional functions, however, is restricted to range only over functions of a certain ramification level. Gödel (1944) noted that this system was acceptable to the same nominalistic constructivism as Ramified Type Theory. One way of making this precise is that, as shown in A. P. Hazen and J. M. Davoren (2001), the 1925 system, like the Ramified system, can be given a semantics on which quantification over objects other than the basic, bottom-level, ones is interpreted substitutionally.

In Appendix B to the second edition of *Principia Mathematica*, Russell gave what he claimed was a derivation of the principle of mathematical induction in his new system, but the proof contains an essential error. Landini (1996) gives a correct proof of induction but uses an additional extensionality axiom that is not valid on the nominalistic semantics. The exact mathematical strength of the 1925 system, supplemented by an Axiom of Infinity, is not clear: It will not suffice for the full strength of (first-order) Peano Arithmetic, but it may yield a richer fragment of it than the Ramified system.

At about the same time, F. P. Ramsey (1925) proposed abandoning ramification altogether, giving a formulation of the Simple Theory of Types. On this view, the basic objects, or *individuals*, form one type and the types of other entities are defined exclusively by the types of

arguments they can take: Properties of individuals will form one type, properties of properties of individuals another, relations between individuals and properties of individuals a third, and so on. The theory need not be extensional: It can allow distinctions between properties holding of exactly the same objects, and both Aldo Bressan (1972) and Montague (cf. Daniel Gallin 1976) developed versions based on modal logic, the first seeking applications in the formalization of physical theory and the second a variety of semantic and conceptual analyses. For mathematical purposes Ramsey assumed extensionality; on this assumption, propositional functions of a single argument amount to sets, those of more than one to *relations-in-extension*. The resulting system is described (and compared with the Ramified Theory) in sections 34–36 of W. V. Quine (1969).

Obviously, Ramsey and those who have followed him have abandoned Russell's attempt to deal with the semantic and intentional paradoxes through type distinctions. Their view was that the set-theoretic paradoxes were adequately handled by the Simple Theory of Types, and the others essentially involved other concepts—semantic or cognitive/epistemological—and were so properly dealt with by separate theories. Alfred Tarski showed that the semantic paradoxes could be avoided by invoking a doctrine of levels of language (clearly foreshadowed by Russell at the end of his introduction to Ludwig Wittgenstein's essay [1921].) Russell seems to have thought the intentional paradoxes were best handled by assimilating them to semantic paradoxes through a kind of *language of thought* idea, which he discussed in Appendix C to the second edition of *Principia Mathematica*.

The extensional Simple Theory of Types, without an axiom of infinity, was proven consistent by Gerhard Gentzen (1936) (one of the successes of Hilbert's Program!). With an axiom asserting the existence of infinitely many individuals, it becomes a usable system of set theory, strong enough to derive most of the mathematics actually used in the natural sciences. As such it was taken as *standard* by many researchers between the publication of *Principia Mathematica* and the 1930s: Gödel (1931) and Tarski (1935) both assume it as their background system. Subsequent set theorists have preferred other axiomatizations, such as Zermelo-Fraenkel set theory, but (as described in sections 37–38 of Quine [1969]), they can be seen as natural generalizations of Simple Type Theory. To get from a system like Ramsey's to one like Zermelo's, one makes five changes:

(1) abandoning relational types (reducing relations to sets by using the Wiener-Kuratowski analysis of ordered pairs),

(2) abandoning the many-sorted formal language, with its separate alphabets of variables ranging over different types of entity, in favor of a description of the whole hierarchy in a first-order language with a single sort of variable,

(3) making the hierarchy cumulative, so a set can have members of any lower level rather than being restricted to members of the immediately preceding level,

(4) allowing sets of infinitely high level, which can have members of all finite levels,

(5) reformulate the axioms to give an elegant systematization of the new framework.

The first is just a simplification, adding nothing to the system. The second would be perverse if we still, like Russell in the first decade of the twentieth century, thought of the entities in the hierarchy as the meanings of expressions in our language, but is natural if we think of them in a Platonistic way as entities independent of our thought or language. The third can be shown to be harmless, and the fourth, though a significant enrichment of the system, is natural after the third. The fifth is not trivial: The resulting systems are stronger than the Type Theory we started with, and would be even if we left out the infinite types. Conceptually, however, the generalized type-theoretic way of thinking about set theory is very satisfying. The stages of George Boolos (1971) are very reminiscent of Russellian types.

Church (1940) makes a different generalization. As Russell's *propositional function* suggests, a property can be thought of as a function mapping arguments (of appropriate type) to propositions (or, assuming extensionality, to truth values). Church assumes two basic types, of individuals and truth values, and represents properties as functions from entities of some type to truth values, and then adds types for other kinds of function: Thus, there is a type of functions from individuals to individuals, a type of functions from individuals to (functions from individuals to individuals), and so on. The formal language embodying this conception is based on a typed version of Church's *lambda calculus*, an elegant notation for the representation of recursive functions. (Montague's intensional logic, mentioned above, is essentially a modal version of Church's system.)

All the type theories mentioned so far have been based on classical logic (or, in modal extensions thereof). The considerations that motivate *intuitionistic logic* are independent of those leading to type theories (recall that Russell's Paradox works in essentially the same way in intuitionistic as in classical logic!), and variants of all of these systems based on intuitionistic logic are possible. They have received some study under the name *theory of species*. Joachim Lambek and P. J. Scott (1986) present what is essentially an intuitionistic variant of the system of Church (1940), showing that it has natural connections with mathematical *category theory*. Per Martin-Löf (1984), with greater philosophical attention to intuitionistic concerns about the meaningfulness of mathematical assertions, has developed fragments of intuitionistic type theory in a series of publications, with Martin-Löf serving as a summary of his work to that point. Since intuitionistic proofs often provide information that can be used to define algorithms, there has been considerable interest in Martin-Löf's and similar systems among computer scientists; Thompson (1991) provides an introduction to the systems and their applications. The area is one of active research by logicians, and efforts to develop more powerful and general theories have encountered difficulties, as witnessed by Thierry Coquand (1994), of a kind that would have been familiar to early twentieth-century researchers in the foundations of mathematics.

See also Epistemology; Frege, Gottlob; Gödel, Kurt; Intuitionism and Intuitionistic Logic; Mathematics, Foundations of; Russell, Bertrand Arthur William.

Bibliography

Boolos, George. "The Iterative Conception of Set." *Journal of Philosophy* 68 (1971): 215–231.

Bressan, Aldo. *A General Interpreted Modal Calculus*. New Haven: Yale University Press, 1972.

Burgess, John, and A. P. Hazen. "Predicative logic and Formal Arithmetic." *Notre Dame Journal of Formal Logic* 39 (1998): 1–17.

Church, Alonzo. "A Formulation of the Simple Theory of Types." *Journal of Symbolic Logic* 5 (1940): 65–68.

Church, Alonzo. "Axioms for Functional Calculi of Higher Order." In *Logic and Art*. Edited by Richard Rudner and Israel Scheffler. Indianapolis, IN: Bobbs Merrill, 1972.

Church, Alonzo. "Comparison of Russell's Resolution of the Semantical Antinomies with That of Tarski." *Journal of Symbolic Logic* 41 (1976): 747–760. Reproduced with revisions in *Recent Essays on Truth and the Liar Paradox*. Edited by R. L. Martin. Oxford: Clarendon Press, 1984.

Church, Alonzo. *Introduction to Mathematical Logic*. Princeton, NJ: Princeton University Press, 1956.

Church, Alonzo. "Russellian Simple Type Theory." *Proceedings and Addresses of the American Philosophical Association* 47 (1974): 21–33.

Church, Alonzo. "Schroeder's Anticipation of the Simple Theory of Types." *Erkenntnis* 10 (1976): 407–411.

Coquand, Thierry. "A New Paradox in Type Theory." In *Logic, Methodology and Philosophy of Science IX*. Edited by Dag Prawitz, Brian Skyrms, and Dag Westerstahl. Amsterdam: North-Holland, 1994.

Fitch, Frederic B. "Propositions as the Only Realities." *American Philosophical Quarterly* 8 (1971): 99–103.

Gallin, Daniel. *Intensional and Higher-order Modal Logic*. Amsterdam: North-Holland, 1976.

Gentzen, Gerhard. "Widerspruchsfreiheit d. Stufenlogik." *Mathematische Zeitschrifft* 41 (1936): 357–366. English translation in *Collected Papers of Gerhard Gentzen*. Edited by M. E. Szabo. Amsterdam: North-Holland, 1969.

Gödel, Kurt. "Russell's Mathematical Logic." In *The Philosophy of Bertrand Russel*. Edited by P. A. Schilpp. Evanston: Northwestern University, 1944.

Gödel, Kurt. "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I." *Monatshefte für Mathematik und Physik* 38 (1931): 173–198.

Hatcher, William S. *The logical Foundations of Mathematics*. New York: Pergamon Press, 1982.

Hazen, A. P., and J. M. Davoren. "Russell's 1925 Logic." *Australasian Journal of Philosophy* 78 (2001): 534–556.

Hylton, Peter. "Russell's Substitutional Theory." *Synthese* 45 (1980): 1–31.

Lambek, Joachim, and P. J. Scott. *Introduction to Higher Order Categorical Logic*. Cambridge, U.K.: Cambridge University Press, 1986.

Landini, Gregory. "The Definability of the Set of Natural Numbers in the 1925 *Principia Mathematica*." *Journal of Philosophical Logic* 25 (1996): 597–615.

Landini, Gregory. *Russell's Hidden Substitutional Theory*. New York: Oxford University Press, 1998.

Linsky, Bernard. *Russell's Metaphysical Logic*. Stanford, CT: CSLI, 1999.

Martin-Löf, Per. *Intuitionistic Type Theory*. Naples: Bibliopolis, 1984.

Pelham, Judy, and Alasdair Urquhart. "Russellian Propositions." In *Logic, Methodology and Philosophy of Science IX*. Edited by Dag Prawitz, Brian Skyrms, and Dag Westerstahl. Amsterdam: North-Holland, 1994.

Quine, W. V. *Set Theory and its Logic*. Cambridge, MA: Harvard University Press, 1969.

Ramsey, F. P. "The Foundations of Mathematics." *Proceedings of the London Mathematical Society* 25 (1925): 338–384.

Russell, B. A. W. "Mathematical Logic as Based on the Theory of Types." *American Journal of Mathematics* 30 (1908): 222–262.

Russell, B. A. W. "On the Substitutional Theory of Classes and Relations."(Originally written 1906 but withdrawn from publication.) In *Essays in Analysis*. Edited by D. Lackey. London: Allen & Unwin, 1973.

Sambin, Giovanni, and Jan Smith. *Twenty-Five Years of Constructive Type Theory*. Oxford: Oxford University Press, 1998.

Tarski, Alfred. "Der Wahrheitsbegriff in den formalisierten Sprachen." *Studia Philosophica* 1 (1935): 261–405.

Thompson, Simon. *Type Theory and Functional Programming*. Wokingham, U.K.: Addison-Wesley, 1991.

Whitehead, Alfred North and Bertrand Russell. *Principia Mathematica*. Cambridge, U.K.: Cambridge University Press: 1910. Revised edition 1925.

Wittgenstein, Ludwig. "Logisch-Philosophische Abhandlung." *Annalen der Naturphilosophie* 14 (1921): 185–262. Reprinted in book form with facing English translation as *Tractatus-Logico Philosophicus*.

Allen P. Hazen (2005)

UGLINESS

Aesthetics has often been described as the philosophical study of beauty and "ugliness." It is important at the outset to see what is involved in this familiar definition, for it embodies a view of ugliness and of its role within aesthetic theory that has been the major source of contention in historical debates on the concept. The first thing to note about this view is that it takes ugliness to be a category that properly falls within aesthetic theory. Ugliness designates aesthetic disvalue as beauty designates positive aesthetic value. The two therefore constitute a value polarity analogous to right and wrong in ethics or to truth and falsehood in epistemology. Just as the field of ethics comprises responsible human actions of which some are evil and blameworthy, so, among perceptual objects, there are some that have negative aesthetic value. This does not mean that such objects simply lack the characteristics by virtue of which things are beautiful; it means, rather, that they possess recognizable properties that are the opposites of those found in beautiful objects.

The relation between beauty and ugliness has commonly been conceived in hedonistic terms, that is, whereas a beautiful object is a source of pleasure in the spectator, an ugly object arouses its opposite, pain. Plato, in numerous instances, takes beauty to be characteristically pleasurable (*Hippias Major* 297–299, *Philebus* 50–52, *Laws* II). Aristotle perpetuates this view, and in his study of specific art forms (notably tragedy) he holds that it is the proper function of these forms to create pleasure. Yet it is clear in his classic *Poetics* that he is troubled by the seeming conflict between this view of art and the empirical fact that works of art often represent objects and events that are ugly. Aristotle raises the question first in regard to the type of visual art that depicts things "which in themselves we view with pain" (IV). He does not doubt, however, that the painting itself arouses pleasure, a phenomenon that is explained by our intellectual interest in recognizing the object. Comedy, moreover, "imitates" men who are ignoble and therefore ludicrous; and though this is a kind of ugliness, the comedy is, for reasons that Aristotle does not specify, kept from being painful (V). Finally, though the protagonist is a good man who suffers adversity, tragedy is not merely shocking (XIII).

Thus Aristotle initiated the controversy over the "paradox of tragedy" that has survived to the present day. As has been shown, this paradox is not the sole instance of the problem of ugliness in art, but it states the problem

most acutely, both because tragedy is almost the only artistic genre whose subject matter is necessarily sorrowful or pathetic and because of the preeminent value that has traditionally been claimed for works in this genre. Why do we esteem narratives of evil and suffering? The poetic values of tragic literature, the ennobling courage of the hero, the insight and wisdom gained by the spectator—these are among the usual solutions of the paradox. All of them consider the ugly as only a single aspect of the work of art, for they all undertake to show that within the work as a whole the ugliness is somehow transcended. Hence they presuppose that some objects, such as the preartistic model of tragic plot, are "painful in themselves," and therefore ugly.

Throughout aesthetic theory, ugliness is discussed mainly by those philosophers who deny precisely this assumption. They wish to hold that ugliness does not exist, and since their thesis runs counter to ordinary belief, they are constrained to justify it. In Augustine, the unreality of ugliness is enjoined by his most fundamental philosophical doctrines. Stated theologically, the world and everything in it have been created by an infinitely good God, as an expression of his goodness; stated metaphysically, existence is not neutral with respect to value and disvalue, but is rather an embodiment, through and through, of positive value. In such a worldview, the apparent presence of evil of any kind poses a problem, and Augustine considers sin and blindness just such problems. But aesthetic disvalue is a particular issue for him because his conception of reality is conspicuously aesthetic. All things are images of the ideas of form and harmony that exist in the mind of God, and together they make up an internally ordered unity. The categories of Greek aesthetic theory are thus writ large in his metaphysics.

To say that a thing can exist at all only if it possesses form, and that, indeed, its existence cannot be conceived of apart from form, implies the solution of Augustine's problem. Objects are beautiful by virtue of their form, but if this is so, then ugliness does not exist, since sheer formlessness cannot exist. The opposite of beauty is not anything real, but merely the absence or "privation" of positive value. But now the argument seems to prove almost too much, for it appears to deny the possibility of the very facts—that is, apparently ugly objects—which gave rise to it in the first place. Augustine therefore employs the notion of "degrees" of value characteristic of metaphysical optimism and idealism. An object may not have the form appropriate to things of its kind, but this lack constitutes a relative deficiency of beauty, not sheer ugliness. Moreover, such objects must be seen not in isolation but as parts of the universe as a whole. Seeming ugliness sets off, and thereby enhances, the beauty of the world. Augustine uses the same argument in the case of objects, such as dangerous animals, which are not in any clear way lacking in form, but are considered ugly because they are displeasing or offensive to the sight.

However, when "form" has been construed less broadly than it was by Augustine, it has been used to differentiate beauty from genuine ugliness. During the sixteenth and seventeenth centuries, numerous treatises were devoted to particular arts, on the model of the *Poetics*. The properties of form that a work must possess in order to achieve beauty are specified precisely and narrowly. These include the "unities" in drama (Pierre Corneille) and the "correct" anatomical proportions in the visual arts (Albrecht Dürer). A work of art that lacks these properties is still recognizably a drama or a sculpture and therefore has some organization or structure. Yet it is not only deficient in beauty but really ugly.

This assured and unequivocal way of distinguishing ugliness was called into question, however, by the rebellion against the "rules" of form that was carried on throughout the eighteenth century. The rules were found to be too parochial and constricting. Yet the distinction between beauty and ugliness might still have been drawn, by reference to felt experience rather than to the object, if the hedonistic theory of value had been consistently preserved. But examination of aesthetic experience (of the sublime) reveals that it engenders feelings that are akin to pain. Sublime objects are overwhelming, menacing, intractable to understanding and control. And yet such experiences, because they are intensely moving, are of great value. Thus, both formalism and hedonism, which had traditionally sustained the duality of beauty-ugliness, are impugned. More fundamentally still, the eighteenth century first established aesthetics as an autonomous and systematic discipline. The question "What counts as a properly aesthetic phenomenon?" was then raised explicitly for the first time. The answer to this question, as we shall see, ultimately determines whether ugliness is a category of aesthetic disvalue. In all these ways, the eighteenth century provided impulse and direction to the vigorous prosecution in recent thought of what was first called, at the close of that century, "the theory of ugliness" (Friedrich von Schlegel, 1797).

According to two of the most influential answers to the question raised above, the aesthetic is to be found either (1) wherever some conceptual theme is embodied in an object that can be grasped by sense and imagination

or (2) wherever some sensory structure expresses to the observer its distinctive feeling-quality. Any object of either kind is said to possess beauty. Ugliness, traditionally, is the "opposite" of aesthetic value. But what would be the opposites of these two conceptions of the aesthetic? In the first case, the opposite would be found in some sensory presentation devoid of intellectual significance or, alternatively, in pure concepts, such as certain of those of science and philosophy, which are beyond imagination. Such objects, however, do not exemplify aesthetic disvalue; rather they fall wholly outside of the realm of the aesthetic as it is defined according to this theory. In the second case, similarly, a thing completely lacking any emotional tone—if any such thing exists—is simply nonaesthetic.

This conclusion, however, fails to take into account ugliness in the usual sense—that is, what we perceive as being displeasing or revolting. W. T. Stace, a recent exponent of the first theory mentioned above, which he took over from G. W. F. Hegel, suggests that what is thus excluded from the aesthetic should be called "the unbeautiful"— "the mere negative absence of beauty"—rather than the ugly. Ugliness itself is a "species" of beauty that is present whenever such concepts as evil and disaster enter into the aesthetic object. The pain that such concepts arouse in us is moral, not aesthetic, and it is usually overcome by the aesthetic pleasure we gain from the total object. Bernard Bosanquet develops the second theory, derived from Benedetto Croce, by arguing that most of what is usually found to be ugly is deemed so because of "the weakness of the spectator." Either the work of art makes very great demands on his emotional capacities or, as in satiric comedy, it offends his moral beliefs; the "weakness," however, is remediable. Such a work of art is therefore more properly considered an instance of "difficult beauty" than of ugliness. Are there any objects at all that come within the realm of the aesthetic and are genuinely (or, as Bosanquet says, "invincibly") ugly? Bosanquet is "much inclined" to think that there are none. Given his view that the expressive is the aesthetic, and that "every form expresses" and is therefore beautiful, it is difficult to see how there could be any such object. He holds, however, that ugliness is to be located in what is only incipiently and partially expressive, that is, in a work of art that suggests some feeling but does not coherently elaborate and fulfill the suggestion, as in sentimental or "affected" art.

The traditional polarity of beauty-ugliness marks the distinction between aesthetic value and disvalue. Both the above theories conceive of the aesthetic in such a way that they leave little or no room for disvalue. Yet both Stace and Bosanquet regard the aesthetic experience as pleasurable. At the same time, they want to make room for art that is tragic, demonic, "difficult" (Stace, for example, cites the sculpture of Jacob Epstein). Therefore, as has just been shown, they seek to reconcile the painfulness of such art with the positive value that it necessarily possesses as an aesthetic object. In the case of Bosanquet, however, the question should be raised whether the expression of feeling is universally accompanied by pleasure. Historically, the concept of "expression" has tended to accommodate emotions of every kind within art, even those, as in an art of violence or outrage, which are "darkest." Successful artistic expression can render such emotions more, rather than less, concentrated and painful, and if it be urged that pleasure is taken in the unity and power of the artist's conception, there are, according to Bosanquet's theory, many nonartistic aesthetic objects that are intensely expressive and for which this explanation will not hold. Since there is no necessary or logical connection between "expression" and "pleasingness," it must be decided empirically whether, even when "the weakness of the spectator" is overcome, his experience of the expressive object has a positive hedonic tone. Stace's view that the painfulness of the theme of a work of art is moral, not aesthetic, seems more like definitional legislation than an insight into aesthetic experience. Moral perplexity and frustration are integral to such art as tragedy, and their painfulness enters into our perception of the total work of art. Stace's view, too, is a defense of hedonism. Yet there is no reason a priori to hold to hedonism in aesthetics, and indeed these difficulties cast doubt on such a theory. The term *ugliness*, in the sense of what is preponderantly painful, may still be used to designate one kind of aesthetic object without any implications of disvalue. So considered, "*X* is ugly but aesthetically good" is not self-contradictory and may indeed be something that we want to and have to say. Those modern artists who have vigorously repudiated the pleasingness of beauty as the goal of their creative efforts have made this way of speaking sound less implausible than it once did.

The graver and more basic question is whether ugliness, in the broader sense of negative aesthetic value, is, for aesthetic theory, otiose. Doubtless, we also want to say sometimes that the work of art is bad. Bosanquet, however, takes genuine ugliness to be at least partially expressive, and if we follow this lead, badness must be construed as a deficiency or relatively slight degree of aesthetic goodness. The work achieves less than it promises, the nonartistic object is lacking in vitality or charm. According to this view, then, there is no opposite to aesthetic

value, but only, as Augustine said, a "privation" of it. On the other hand, this may be thought to be a gratuitous misreading of those properties that are commonly held to constitute ugliness or that are adduced as reasons for judging a thing ugly. Muddy orchestration or incoherent plot structure are, significantly, opposites to orchestral clarity or unity of plot, and they are equally real and present to awareness. In the absence of compensating virtues, objects that possess them are "positively bad."

No matter whether the denial of negative value should, finally, be tolerated or rejected, it is fair to say that this denial is less vexing in aesthetics than in ethics or epistemology. The explanation lies, in large part, in Bosanquet's notion of "the weakness of the spectator." The determination of beauty and ugliness is much more closely tied to the perceptual and emotional capacities of the spectator and to the attitudes that affect them than it is to moral and cognitive values. This leads us to think that the experience of negative value (though not that of positive value) results from a failure to see what is yet there to be seen. Thus, the transvaluation of what had previously been accounted ugly, which is endemic to the history of art and taste, is characteristically credited with being an enlargement of sympathy and a refinement of discrimination. The more obdurate cognitive and moral judgments of falsehood and evil, however, are not characteristically altered in this way. Can any limits, therefore, be set to what sensibility finds to be aesthetically good? To define the field of the aesthetic in such a way that all things are seen to possess positive value formalizes the endless catholicity of aesthetic interest. Freed from the exigencies of morality and the biases of perceptual habit, the aesthetic approach to the world, at the hypothetical limit, fixes upon any tone or shade the quality of any ambience. In John Keats's words, it "has as much delight in … an Iago as an Imogen." But if everything engages and rewards aesthetic perception, then either "aesthetic disvalue" is a self-contradiction or else it denotes nothing.

See also Aesthetic Qualities; Aesthetics, History of; Aesthetics, Problems of; Aristotle; Augustine, St.; Beauty; Bosanquet, Bernard; Croce, Benedetto; Hegel, Georg Wilhelm Friedrich; Humor; Plato; Pleasure; Schlegel, Friedrich von; Stace, Walter Terence; Tragedy; Visual Arts, Theory of the.

Bibliography

Bosanquet, Bernard. *Three Lectures on Aesthetic*. London, 1923. Lecture III is the starting point for any serious discussion of the problem of ugliness.

Chapman, Emmanuel. *Saint Augustine's Philosophy of Beauty*. New York: Sheed and Ward, 1939. Those of Augustine's writings on aesthetics that have survived are scattered among several different treatises. This convenient volume locates these passages.

Pepper, Stephen C. *The Basis of Criticism in the Arts*. Cambridge, MA: Harvard University Press, 1946. Contrasts the meaning and status of ugliness in some of the major aesthetic theories. In the author's own view, ugliness is a moral rather than an aesthetic category.

Rosenkranz, Karl. *Ästhetik des Hässlichen*. Köenigsberg, 1853. The classic post-Hegelian study of ugliness. Addicted somewhat to multiplying conceptual distinctions but a thoughtful and suggestive analysis. Argues that art would be unduly narrow if it did not include the ugly and that ugliness is not simply a device to enhance beauty.

Santayana, George. *The Sense of Beauty*. New York: Scribners, 1936. An influential contemporary statement of hedonist aesthetics.

Stace, W. T. *The Meaning of Beauty*. London: G. Richards and H. Toulmin, 1929.

Véron, Eugène. *Esthétique*. Paris: C. Reinwald, 1878. Translated by W. H. Armstrong as *Aesthetics*. London: Chapman and Hall, 1879. An early and forceful statement of the theory that the artist seeks emotional self-expression rather than beauty and will therefore exploit rather than mitigate the ugliness of his subject.

OTHER RECOMMENDED WORKS

Coleman, Earle J. "The Beautiful, the Ugly and the Tao." *Journal of Chinese Philosophy* (1991): 213–226.

Devereaux, Mary. "The Ugly." *The American Society for Aesthetics Newsletter*, vol. 24, no. 3 (Winter 2005): 1–2.

Guyer, Paul. "Thomson's Problems with Kant: A Comment on Kant's Problems with Ugliness." *Journal of Aesthetics and Art Criticism*, vol. 50, no. 4 (Fall 1992): 317–319.

Higgins, Lesley Hall. *The Modernist Cult of Ugliness: Aesthetic and Gender Politics*. New York: Palgrave Macmillan, 2002.

Kieran, Matthew. "Aesthetic Value: Beauty, Ugliness and Incoherence." *Philosophy*, vol. 72, no. 281 (1997): 383–399.

McCall, Robert E. "The Metaphysical Analysis of the Beautiful and the Ugly" (with comment by James P. Reilly). *Proceedings and Addresses of the American Philosophical Association*. Vol. 30, 1956, pp. 137–153.

Mesa, James, P. "The Good, the Bad, and the Ugly: The Aesthetic in Moral Imagination." In *Beauty, Art and the Polis*, edited by Alice Ramos, Catholic University of America Press, 2000, pp. 237–244.

Miller, William Ian. *The Anatomy of Disgust*. Harvard University Press, 1998.

Pickford, R. W. "The Psychology of Ugliness." *British Journal of Aesthetics* 9 (1969): 258–270.

Raters, Marie-Luise. "Unbeautiful Beauty in Hegel and Bosanquet." *Bradley Studies*, vol. 7, no. 2 (Autumn 2001): 162–176.

Shier, David. "Why Kant Finds Nothing Ugly." *British Journal of Aesthetics*, vol. 38, no. 4 (1998): 412–418.

Thomson, Garrett. " Kant's Problems with Ugliness." *Journal of Aesthetics and Art Criticism*, vol. 50, no. 2 (Spring 1992): 107–115.

Jerome Stolnitz (1967)

Bibliography updated by Mary Devereaux (2005)

ULRICH (ENGELBERT) OF STRASBOURG

(fl. 1248–1277)

Ulrich (Engelbert) of Strasbourg was a scholastic philosopher and theologian, priest, and author. A member of the Dominican priory at Strasbourg in the German province, Ulrich studied under Albert the Great at Cologne, together with Thomas Aquinas and Hugh of Strasbourg, between 1248 and 1254. During those years Ulrich heard Albert expound the Dionysian corpus and the *Ethics* of Aristotle. As a lecturer in theology at Strasbourg, Ulrich acquired considerable fame for his learning; among his illustrious disciples was Lector John of Fribourg.

The ancient catalogs attribute to Ulrich commentaries on Aristotle's *Metheora* and *De Anima,* Peter Lombard's *Sentences,* and the book of Ecclesiastes. His only extant work, however, is a remarkable compendium of theology titled *De Summo Bono,* planned and probably written in eight books. Only the first book and fragments of others have been published, and the known manuscripts end with Book VI, tr. 5. This compendium was composed between 1262 and 1272 and marks a notable advance over the earlier *summas* of William of Auxerre, Alexander of Hales, and Albert the Great. It is divided into (1) introduction to theology, (2) essence of the supreme Good, (3) Trinity in general, (4) the Father and creation, (5) the Son and incarnation, (6) the Holy Spirit and sanctification, (7) sacraments, and (8) ultimate beatitude.

The doctrinal framework of Ulrich's thought is predominantly Augustinian and Neoplatonic, depending largely on Pseudo-Dionysius, Avicenna, *Liber de Causis,* and Albert. For Ulrich man has a rational predisposition for knowing the existence of God as the supreme cause. This knowledge is rendered more precise, although not comprehensive, by the traditional three ways: (1) negating imperfections found in creatures (for example, as creatures are finite, God is infinite); (2) seeing God as the ultimate cause of all perfections; and (3) recognizing the transcendence of those perfections in God. God created the universe in a hierarchical order ranging from the first luminous intelligence through lesser intelligences, man, animals, plants, elements, and material principles. In all creatures there is a real distinction between essence and existence, and in all material substances there is only one substantial form. Created intellectual substances, seeing the eternal Ideas in God, illuminate lesser intelligences to know truth. The human mind has four immediately evident (*per se nota*) rules by which it can investigate theology, the science of the faith: God is the supreme Truth and cause of all truth; primary Truth can neither deceive nor be deceived, therefore his Word should be believed; we should believe everything clearly revealed by God through his spokesmen; Scripture is true precisely because God gave it to us in that way. Unlike these rules, the articles of faith are not immediately evident, but in the light of faith and these rules, the articles of faith become objects of scientific study.

For five years (1272–1277) Ulrich was provincial of the German province before the General Chapter of Bordeaux assigned him to Paris to lecture on the *Sentences* and to obtain his degree in theology. He died, probably in 1278, before becoming a master; in the manuscripts he is designated a bachelor in theology.

See also Albert the Great; Alexander of Hales; Aristotle; Augustinianism; Avicenna; Medieval Philosophy; Neoplatonism; Peter Lombard; Pseudo-Dionysius; Thomas Aquinas, St.

Bibliography

Daguillon, Jeanne. *Ulrich de Strasbourg, O.P., la "Summa de Bono,"* Lib. I. Paris, 1930. Bibliothèque Thomiste, Vol. XII.

Glorieux, Palémon. *Ulrich de Strasbourg au XIII[e] siècle,* 2 vols. Paris, 1933–1934. Vol. I, pp. 145–151.

Grabmann, Martin. "Studien über Ulrich von Strassburg." In *Mittelalterliches Geistesleben; Abhandlungen zur Geschichte der Scholastik und Mystik,* 3 vols. Munich: Hueber, 1926–1956. Vol. I, pp. 147–221.

Quétif, Jacques, and Jacques Échard. *Scriptores Ordinis Praedicatorum Recensiti,* 2 vols. Paris: Lutetiæ Parisiorum, apud J.-B.-C. Ballard, and N. Simart, 1719–1721. Vol. I, p. 256.

Théry, Gabriel. "Originalité du plan de la Summa de Bono d'Ulrich de Strasbourg." *Revue Thomiste* 27 (1922): 276–297.

James A. Weisheipl, O.P. (1967)

ULTIMATE MORAL PRINCIPLES: THEIR JUSTIFICATION

See *Moral Principles: Their Justification*

UNAMUNO Y JUGO, MIGUEL DE

(1864–1936)

The Spanish philosopher of life Miguel de Unamuno y Jugo's concern was neither with the problems of linguis-

tic clarification and conceptual analysis nor with speculative metaphysical constructions but, rather, with coming to terms with life both intellectually and emotionally. The symbols Unamuno used are related to Spanish life and destiny and his way of thinking was Spanish, but his message is universal. He expressed himself symbolically, through poetry, religious writings, and the novel, and through the general evocative and emotive character of his prose. However, his efforts to give literal articulation to the mystery and anguish of his existence make him a philosopher rather than exclusively a novelist or poet. The style of philosophy that Unamuno represents must at all times emanate from the world situation and the life situation of the individual philosopher. It follows that Unamuno's philosophy is to be found not only in his writings but also in his general mode of life, particularly in his conspicuous political actions at a time of serious turmoil in Spain.

In view of this it is quite proper to call Unamuno an existentialist. First, his philosophy clearly wells up from his own human situation in space and time. Second, his writings tend to be emotive rather than intellectual. He wished to express not exact ideas but feelings; and feelings are often more accurately expressed in the turgid and quasi-sentimental language of Unamuno than in logical exegesis. Third, his subject matter was existential—death and anxiety, doubt and faith, guilt and immortality. Fourth, he traced the sources of his thought to such existentialist precursors as Blaise Pascal and Søren Kierkegaard and found kinship with anyone who stressed intuition and subjectivity in the life of man and in the construction of worldviews—with men like Arthur Schopenhauer, Friedrich Nietzsche, and William James. Finally, Unamuno's philosophy, like Kierkegaard's, was deliberately unsystematic, an expression of his wrestling with existence, and any systematic account of that expression must falsify or at least distort the facts of experience.

LIFE

Don Miguel de Unamuno y Jugo was born in the Basque city of Bilbao. He studied philosophy and classics at the University of Madrid and moved to Salamanca in 1891 as professor of Greek at the university there. He was associated with the university for most of the rest of his life, being appointed rector in 1901 and named rector for life in 1934. Unamuno's first published work, *En torno al casticismo* (On purism; 1895), was a historical and political work that questioned and examined the place of Spain in the modern world. His first novel, *Paz en la guerra* (Peace in war; 1897), sometimes called the first existentialist novel, was based on his early memories of the siege of Bilbao in 1873. In the novel *Amor y pedagogía* (Love and pedagogy; 1902) Unamuno tried to show the basic failure of science in dealing with human and humanistic problems. *Amor y pedagogía* describes a man's attempt to educate his family scientifically and the dismal failure of this attempt. *Vida de Don Quijote y Sancho* (*Life of Don Quixote and Sancho*, 1905) foreshadowed many of the themes of Unamuno's masterpiece, *Del sentimiento trágico de la vida en los hombres y en los pueblos*. The *Vida de Don Quijote* is a plea for salvation through the anguish and passion experienced by the man of flesh and blood. *Del sentimiento trágico de la vida* (*The Tragic Sense of Life*), which appeared in 1913, expresses Unamuno's intemperate longing for eternal life and his desperate search for some solace in the exploration of the tension and conflict that exists between faith and reason. The novel *Niebla* (*Mist*) was published in 1914, and in 1917 Unamuno's modern version of the problem of Cain, *Abel Sánchez*, appeared. In 1924 Unamuno was deported to Fuerteventura in the Canary Islands for his unrelenting attack on the totalitarianism of General Primo de Rivera. He managed to escape to France and remained in exile until 1930, when Rivera's dictatorship fell. Unamuno was reinstated as rector of the University of Salamanca the next year.

From 1931 to 1933 Unamuno served in the Cortes, the constituent assembly of the Spanish republic, as an independent Republican deputy. His last and greatest novel appeared in 1933. *San Manuel bueno, mártir* (Saint Emanuel the Good, martyr) describes the agony of a priest who finds it impossible to believe. Unamuno's independence, individualism, and patriotism led to his being dismissed from his rectorship in 1936. He at first favored the nationalists in the Spanish Civil War, but he came to feel that neither side was working for the best interest of either Spain or humanity. During the last year of his life he was under house arrest in Salamanca.

CENTRAL THEMES

To characterize Unamuno's basically unsystematic philosophical position is difficult. A few themes can be isolated from his philosophy, however, and may be generalized as follows:

(1) Unamuno's interest was primarily in the individual rather than in social reality, and thus his philosophy extols the agony and the importance of the individual. In this context Unamuno's Spanishness becomes not a social ideal but the expression of his individuality.

(2) He emphasized the importance of personal integrity. Truthfulness to oneself and total honesty in ideals are the hallmarks of the philosophical man.

(3) He saw his function—and that of philosophers generally—as that of a Socratic gadfly to the community. The philosopher is needed to reawaken us to our genuine nature, to our authentic problems, and to the honest attempts to resolve them.

(4) Much of Unamuno's life was spent in agony over the conflict between faith and reason. Reason alone—which Unamuno invariably associated with skepticism—cannot lead to any kind of fundamentally hopeful knowledge. Faith can do so, but faith exists only in the shadow of the despair that is reason; it has no independent and positive existence. Faith can never totally dispel reason, and reason always leads to despair. The logic of the heart is hopeful and gives meaning to life, but it is never strong enough to fully set aside the darkness of the logic of the head.

(5) Unamuno's general conception of religion was related to the tension between faith and reason. Although Catholicism did not fully satisfy either his emotions or his reason, Unamuno felt that religion is a necessity of life. We must risk faith in the way that Pascal wagered, James willed, and Kierkegaard leaped. We must, for profoundly pragmatic reasons, live as if God does in fact exist.

(6) The above views led to the doctrine that commitment is one of the central features of the authentic life. An authentic life is dedicated to and identified with an ideal, an ideal that genuinely emanates from the depths of each man. The truth of such a commitment can be vindicated and confirmed only by the heart; but since reason casts permanent doubt on that commitment, a blind, courageous leap of faith is needed for authentic human existence.

(7) Life thus becomes a vague, brittle, and tenuous cluster of experiences between two awesome, incomprehensible, and impenetrable barriers of nothingness: birth and death. Only through a foundationless but fervid commitment can man escape, at least temporarily, the despair of meaninglessness.

(8) Unamuno loved Spain and was an impartial observer and recorder of the Spanish temperament. According to Unamuno, the Spaniard—like his paradigm Don Quixote—wants adventures, willingly risks revolution for the establishment of utopian societies, and is impractical. But there is also a practical side to the Spaniard, symbolized by Sancho Panza, which often degenerates into blind formalism, intolerance, religious bigotry, and unprincipled commercialism.

Unamuno's commitment to Spain embraced his commitment to the Catholic Church. However, it was only his heart that pulled him toward the church; his reason pulled him away from it. This excruciating tension between his fervent emotional need and hope for the presence of an enveloping and supporting God and for certainty with respect to the immortality of the soul on the one hand, and the fact that he found this world picture rationally untenable on the other hand, was central to Unamuno's philosophy.

GOD AND EXISTENCE

The problem of human existence, in Unamuno's famous formulation, is *el sentimiento trágico de la vida* (the tragic sense of life); it is the fact that there is sorrow that has no resolution and evil that has no redemption. We should weep, not because it helps but precisely because it avails us nothing. If we recognize the pervasiveness of hopelessness and despair, we can at least experience the brotherhood of man. Without disease or defect (be it sin in paradise, a weak species of apelike man, or immunization—the momentary creation of an illness for the sake of health) there can be no progress. Philosophy in this sense is eminently practical: *Primum vivere, deinde philosophari*—"man philosophizes in order to live." "He philosophizes either in order to resign himself to life or to seek some finality in it, or to distract himself and forget his griefs, or for pastime and amusement" (*The Tragic Sense of Life*, p. 29).

The most attractive solution to the problems of human existence, to "the tragic sense of life," is the hope for eternal life expressed in man's perennial hunger for immortality. This hunger has two dimensions—it refers either to the nondestruction of the soul or to the merger of the soul with the universe or the totality of being. In connection with the first of these dimensions, Unamuno seems to have held that the destruction of a man's consciousness is an a priori impossibility: We cannot even conceive of the nonexistence of consciousness, since that conception is itself an act of consciousness. In connection with the second, he concluded that man is nothing if he is not everything—to exist is yearning to reach all space, all time, all being. To be a man is to seek to become God.

Unless man is God, he is not even man: "Either all or nothing!" was Unamuno's motto.

Catholicism promises immortality, but modern rationalism denies it. As a consequence, fundamental doubt sets in, doubt that is both passionate and rational. Such tense but mature insight, however, does lead to some solace: "But here, in the depths of the abyss, the despair of the heart and of the will and the skepticism of reason meet face to face and embrace like brothers" (ibid., p. 106). Man must reach the depths of despair, doubt, and agony in order to arrive at the solid "foundation upon which the heart's despair must build up its hope." Furthermore, the agony that arises out of the tensions of passionate doubt and total rational skepticism when both are focused on the problem of eternal life may also form "a basis for action and morals."

Tension is the essence of life, and the tension that leads to agony is also the tension that allows man to feel his existence; pure consciousness deserves only suicide. Life, to be felt as real, *as there*, as existing, must be a life of passion. This truth is well illustrated by love, which for Unamuno is basically sexual love. In the tensions and paradoxes of love—as well as in compassion and pity—man experiences the richness, concreteness, and fullness of his existence. Consciousness, in this sense, is knowledge through participation; it is "co-feeling."

The hope for immortality is supported by the notion of God. The traditional arguments for the existence of God prove nothing other than that we have the idea of God. The God who is the idea of excellence and the first mover is a fleshless and passionless abstraction and cannot soothe the anguish of man's existence. This abstraction is not what the heart craves. The strongest conclusion of reason is that we "cannot prove the impossibility of His existence." Belief in God is an expression solely of man's longing for the rich and concrete experience of his existence and of his determination to live by this longing and make it a basis for action. Man's agonizing hunger for the divine—even though it cannot be satisfied directly—leads to hope, faith, and charity, and eventually to his sense of beauty and of goodness.

There are other typically existentialist themes in Unamuno's philosophy:

(1) Man is painfully aware of his contingency. That he exists or that he is the particular person he happens to be is neither necessary nor permanent.

(2) To assuage his anguish, man must feel his existence, even if he is led to suffering. He must learn to experience his uniqueness by expanding the range and the self-consciousness of his perceptions of the world.

(3) All existence is a mystery: Consciousness is a mystery, contingency is a mystery, absurdity is a mystery, and anguish is a mystery.

(4) Love is the basic force of human existence. It encompasses all the conative relations of man to being and enables him to overcome the anguish of his contingency by giving him the rich feeling of his own existence.

(5) The central temporal dimension of human existence is the future, which leads to a desire for immortality and to a concern with death. This focus on the future is expressed in Unamuno's use of *esperar*: It means both the joys of hope and the anguish of eternal waiting. The structure of the future expresses both man's determination to continue to live and his permanent dissatisfaction and despair concerning existence.

(6) Goals are self-created and are permanent commitments.

(7) Finally, Unamuno's views on the nature of language foreshadow those of Maurice Merleau-Ponty and Martin Heidegger. Language is a mode of being. Living, not only knowing, is expressed in certain basic forms, one of which is language. Language thus is not symbolic but the actual embodiment of an idea. Without language an idea could not exist.

EPISTEMOLOGY AND METAPHYSICS

Truth, according to Unamuno, is subjective; it exists only as it is manifested in authentic belief. Belief, in turn, is an expression of man's total being and consequently is realized in action. Objective truth is, strictly speaking, a meaningless conception. Through its identity with belief and action, truth is ultimately an act of will. It is a will to create; and the will as creator wants and loves at the same time. Because of this personal and volitional factor in truth, the opposite of truth is not error but the lie. This subjective view of truth gives a distinct idealistic, even mystical, cast to Unamuno's thought. All knowledge about man and the world is subjective in the sense that it begins with first-person experience. To think of truth as transcending first-person experiences is, strictly speaking, a contradiction, because the very program of transcending first-person experiences is a first-person project and concept and a construction. There is, however, another kind of truth, illustrated by mathematics, which is the

function of reason alone, whereas true belief is a function of man's whole being.

Unamuno followed Heraclitus in holding that reality is a state of permanent flux, so that no two experiences are ever the same. There are two metaphysical alternatives. Reality may be a vast sea of consciousness with my subjectivity at the center. There is no easy way to distinguish this consciousness from a mere dream. Its sole foundation is the fact that I experience it and that I will it to be real. Unamuno ultimately rejected this view. The other view is that the focus of our being may be outside ourselves. We may identify ourselves with the realities of other people, with trees, flowers, and mountains. This orientation, to which Unamuno did not accede fully but which he preferred, is close to objective idealism and to naturalism. In either view, man and world are intimately meshed.

See also Common Consent Arguments for the Existence of God; Existentialism; Faith; Heidegger, Martin; Heraclitus of Ephesus; Immortality; James, William; Kierkegaard, Søren Aabye; Life, Meaning and Value of; Love; Merleau-Ponty, Maurice; Nietzsche, Friedrich; Pascal, Blaise; Reason; Schopenhauer, Arthur.

Bibliography

WORKS BY UNAMUNO

En torno al casticismo. Madrid, 1895.

Paz en la guerra. Madrid: F. Fe, 1897; 3rd ed., Madrid, 1931.

Amor y pedagogía. Barcelona, 1902; 2nd ed., Madrid, 1934.

Vida de Don Quijote y Sancho, según Miguel de Cervantes Saavedra, explicada y comentada. Madrid, 1905; 2nd enlarged ed., Madrid, 1914. Translated by H. P. Earle as *Life of Don Quixote and Sancho according to Miguel de Cervantes Saavedra Expounded with Comment*. New York: Knopf, 1927.

Poesías. Bilbao, 1907.

Por tierras de Portugal y España. Madrid, 1911.

Rosario de sonetos líricos. Madrid: Imprenia española, 1911.

Contra esto y aquello. Madrid: Renacimiento, 1912; 4th ed., Buenos Aires: Espasa-Calpe Argentina, 1941.

El espejo de la muerte. Madrid, 1913.

Del sentimiento trágico de la vida en los hombres y en los pueblos. Madrid, 1913. Translated by J. E. Crawford Flitch as *The Tragic Sense of Life in Men and in Peoples*. London, 1921.

Niebla. Madrid, 1914. Translated by Warren Fite as *Mist. A Tragi-comic Novel*. New York, 1928.

Ensayos, 7 vols. Madrid, 1916–1918.

Abel Sánchez. Madrid: Renacimiento, 1917. Translated by Anthony Kerrigan in *Abel Sanchez and Other Stories*. Chicago: Gateway, 1956.

Andanzas y visiones españolas. Madrid: Renacimiento, 1922; 2nd ed., Buenos Aires: Espasa-Calpe Argentina, 1941.

De Fuerteventura á París. Paris: Editorial Excelsior, 1925.

La agonía del cristianismo. Madrid: Renacimiento, 1931. Translated by K. F. Reinhardt as *The Agony of Christianity*. New York: Ungar, 1960.

San Manuel bueno, mártir y tres historias más. Madrid: Espasa-Calpe, 1933.

Epistolario á Clarín. Madrid, 1938.

La ciudad de Henoc. Comentario 1933. Mexico City: Editorial Séneca, 1941.

Obras completas, 7 vols. Edited by Manuel García Blanco. Madrid: Aguado, 1950–1959.

Poems. Translated by Eleanor Turnbull. Baltimore, 1952.

Cancionero, diario poético. Buenos Aires: Losada, 1953.

Teatro. Barcelona, 1954.

España y los españoles. 2 vols. Madrid: A. Aguado, 1955.

Inquietudes y meditaciones. Madrid: A. Aguado, 1957.

Cincuenta poesías inéditas (1899–1927). Madrid: Ediciones de los Papeles de Son Armadans, 1958.

Our Lord Don Quixote: The life of Don Quixote and Sancho, with Related Essays. Princeton, NJ: Princeton University Press, 1967.

Desde el miradol de la guerra. Paris: Centre de recherches hispaniques, 1970.

The Last Poems of Miguel de Unamuno. Princeton, NJ: Princeton University Press, 1974.

Ficciones: Four Stories and a Play. Princeton, NJ: Princeton University Press, 1976.

Escritos socialistas, 1894–1922. Madrid: Ayuso, 1976.

Artículos olvidados sobre españa y la primera guerra mundial. London: Tamesis Book, 1976.

El torno a las artes. Madrid: Espasa-Calpe, 1976.

Crónica política española (1915–1923). Salamanca: Grupo Editorial Ambos Mundos, 1977.

República española y españa republicana (1931–1936). Salamanca: Almar, 1979.

La vida literaria. Madrid: Espasa-Calpe, 1981.

Ensueño de una patria. Valencia: Pre-Textos, 1984.

The Private World: Selections from the Diario Intima and Selected Letters. Princeton, NJ: Princeton University Press, 1985.

Dos articulos y dos discursos. Madrid: Editorial Fundamentos 1986.

El resentimiento trágico de la vida. Madrid: Alianza Editorial, 1991.

Artículos en 'la nación' de buenos aires, 1919–1924. Salamanca: Ediciones Universidad de Salamanca, 1994.

WORKS ON UNAMUNO

Ferrater Mora, José. *Unamuno*. Translated by Philip Silver. Berkeley: University of California Press, 1962.

Meyer, François. *L'ontologie de Miguel de Unamuno*. Paris: Presses Universitaires de France, 1955.

Oromi, Miguel. *Pensamiento filosófico de M. de Unamuno*. Madrid, 1943.

Rudd, Margaret Thomas. *The Lone Heretic*. Austin: University of Texas Press, 1963.

Peter Koestenbaum (1967)
Bibliography updated by Thomas Nenon (2005)

UNCERTAINTY PRINCIPLE

See *Quantum Mechanics*

UNCONSCIOUS

Under the impact of new developments in science, ideas in all fields are undergoing rapid change. This is especially true of the twentieth-century conception of the *unconscious,* the term being used here in a general sense for all those mental processes of which the individual is not aware while they occur in him.

The present interest in the unconscious is a result of the advance of science and psychology since the mid-1800s, and to understand this interest requires some knowledge of the history of ideas. But the timing of this outburst of interest, its intensity (which is greatest in the English-speaking countries and least in Russia and China), and the particular conception of the unconscious that is now dominant are mainly due to one man, Sigmund Freud. His high degree of success in creating widespread appreciation of the power of the unconscious makes the improvement of his conception of it a matter of great importance. Fortunately, a historical survey can not only put recent sectarian conflicts in perspective but can also throw light on aspects of the unconscious that have long been recognized by philosophers and humanists but that receive inadequate emphasis in Freudian theory.

There have been few peoples since, say, 3000 BCE who have not possessed myths expressing a sense of the power of divine or natural agencies to influence the individual without his being aware of that influence. Before the emergence of clear conceptions regarding nature and man there prevailed a sense of the continuity of phenomena, and it was taken for granted that man was part of a totality in which anything might influence anything else. This assumption of continuity is evident in much Eastern thought. Western recognition, from around 1600 CE, of unconscious mental processes, at first philosophical but gradually becoming more scientific, may be superficially regarded as the rediscovery of something that had long been taken for granted in certain Eastern traditions and also in some Greek and Christian writings. Plotinus held that "the absence of a conscious perception is no proof of the absence of mental activity," Augustine was interested in memory as a faculty extending beyond the grasp of the conscious mind, Thomas Aquinas developed a theory of the mind covering "processes in the soul of which we are not immediately aware," and most mystics assumed that insights might be gained by a process of inner reception in which the conscious mind is passive.

But these early ideas lack an essential feature of the modern concept of the unconscious that became possible only after Western thought had set out on the search for precision and scientific validity and, in doing so, had separated the conscious mind from material processes; that is, this became possible only from about 1600 on, or after René Descartes. For the ultimate purpose of the concept of unconscious mental processes is to link conscious awareness and behavior with its background—a system of processes of which one is not immediately aware—and to establish this connection without losing the benefits of scientific precision. Here lies the weakness of the concept of the unconscious: It cannot be made fully acceptable to the scientific age until some science or union of sciences has provided an adequate conception of the unity and continuity of conscious thought, unconscious cerebral processes, physiological changes, and the processes of growth. In fact, the idea of the unconscious (or some equivalent) can acquire scientific status only after a unified picture of the human organism has repaired the intellectual lesions created by Cartesian and other dualistic or specialized methods.

DESCARTES TO FREUD

It is useful, if oversimplified, to consider that Descartes, by his definition of mind as awareness, provoked as a reaction the Western "rediscovery" of unconscious mental processes. During the two and a half centuries between Descartes's *Discourse on Method* (1637) and Freud's first interest in the unconscious, many philosophers, psychologists, biologists, novelists, and poets recognized that mental activity of various kinds occurs without awareness. This view was reached through introspection, through observation, or through attempts to create a theory of the working of the mind. By the last decades of the nineteenth century it was so widespread in Germany and Britain, and to a lesser extent in France, that one can say that by then the existence of the unconscious mind had become a common assumption of educated and psychological discussions; however, its structure, mode of operation, and role in illness were left for the twentieth century to explore.

Here we can consider only a few names out of many, selected either because they were influential or because their ideas represent an advancing understanding.

Our survey opens at the moment when Cartesian thought was acquiring influence. Ralph Cudworth, English divine and philosopher, wrote in 1678:

> There may be some vital energy without clear consciousness or express attention—Our human souls are not always conscious of whatever they have in them—that vital sympathy, by which our soul is united and tied fast to the body, is a thing that we have no direct consciousness of, but only in its effects—There is also a more interior kind of plastic power in the soul … whereby it is formative of its own cogitations, which it itself is not always conscious of. (*True Intellectual System of the Universe,* Book I, Ch. 3)

Many other thinkers of the seventeenth and eighteenth centuries expressed similar ideas, at first mainly in relation to the cognitive aspects, such as perception and memory. Gottfried Wilhelm Leibniz introduced the notion of a quantitative threshold. For him ordinary perceptions were the summation of countless small ones, each of which we are not aware of, because they lie below this threshold.

Two eighteenth-century figures were among the first to direct attention to the emotional aspects of the unconscious mind. Jean-Jacques Rousseau tried to explore the unconscious background of his own temperament and to discover the reason for his fluctuating moods ("It is thus certain that neither my own judgment nor my will dictated my answer, and that it was the automatic consequence of my embarrassment"), and J. G. Hamann, a German religious philosopher, studied the deeper levels of his own mind as evidenced in his experience of conversion, in the emotional life, and in imaginative thinking ("How much more the formation of our own ideas remains secret!").

Between 1750 and 1830 a number of German philosophers and poets increasingly emphasized the emotional and dynamic aspects of the unconscious. Johann Gottfried Herder stressed the role of unconscious mental processes in relation to the imagination, dreams, passion, and illness. Johann Wolfgang von Goethe expressed in poems and *aperçus* his sense of the fertile interplay of conscious and unconscious in the creative imagination, "where consciousness and unconsciousness are like warp and weft." Johann Gottlieb Fichte treated the unconscious as a dynamic principle underlying conscious reason. G. W. F. Hegel based his philosophy on the conception of an unconscious historical process becoming in the individual a partly conscious will. For Friedrich von Schelling unconscious nature becomes conscious in the ego.

Many of the romantic writers and poets, particularly in Germany and England, echoed what was in the air: a vivid sense of the powerful, dark, yet creative aspects of the unconscious mind. Thus, J. P. F. Richter wrote: "The unconscious is really the largest realm in our minds, and just on account of this unconsciousness the inner Africa, whose unknown boundaries may extend far away."

Another sequence of German thinkers made the idea of the unconscious a commonplace of European educated circles by about 1880: Arthur Schopenhauer, C. G. Carus, Gustav Fechner, Eduard von Hartmann, and Friedrich Nietzsche. Schopenhauer took the idea of a mainly unconscious will in nature and in man as his central theme. Carus, physician and friend of Goethe, opened his *Psyche* (1846) with the words: "The key to the understanding of the character of the conscious lies in the region of the unconscious" and presented Goethe's favorable view of the unconscious. Fechner, like Freud (who expressed a debt to him), regarded the mind as an iceberg largely below the surface and moved by hidden currents. He used the concept of mental energy, a topography of the mind, an unpleasure-pleasure principle, and a universal tendency toward stability. Von Hartmann's *Philosophy of the Unconscious* (1869) gave a survey of a vast field of unconscious mental activities, and this book enjoyed a great success in Germany, France, and England. He discussed twenty-six aspects of the unconscious and converted the Goethean ideas of Carus's *Psyche* into a grandiose metaphysical system. Nietzsche, in his penetrating insights into the unconscious, reflected what was already widespread but gave it a new intensity. "The absurd overvaluation of consciousness …. Consciousness only touches the surface …. The great basic activity is unconscious …. Every sequence in consciousness is completely atomistic …. The real continuous process takes place below our consciousness; the series and sequence of feelings, thoughts, and so on, are symptoms of this underlying process …. All our conscious motives are superficial phenomena; behind them stands the conflict of our instincts and conditions."

Nietzsche had cried, "Where are the new doctors of the soul?" Soon after, Freud started on his task: to begin afresh, unprejudiced by all this speculation, and to try to identify the precise structure of unconscious processes and their role in particular mental disturbances, so that lesions of the mind might be repaired by systematic techniques. We are not here concerned with his methods of therapy or with their degree of efficacy but with his steadily developing and often modified theory of the unconscious mind.

Freud was not the first to develop a systematic theory of conflicts in the unconscious. J. F. Herbart had put forward a theory of the operation of unconscious inhibited ideas and their pressure on consciousness, and of the resulting conflict between conscious and unconscious ideas at the threshold of consciousness. But he had little immediate influence. Meanwhile, a school of medical thought was developing in England that treated the patient as a unity, took for granted the interplay of unconscious and conscious, and sought to use this way of thinking in its approach to mental illness. William Hamilton, student of medicine and metaphysics, lectured on the role of the unconscious, particularly in relation to emotions and action, thus providing the background for the psychiatrist H. Maudsley and the naturalist W. B. Carpenter. Maudsley's *The Pathology of Mind* (1879) expresses this English school of thought about the unconscious and is included in the references given by Freud in his *Interpretation of Dreams* (1900), while Carpenter's *Principles of Mental Physiology* (1876) discusses "unconscious cerebration." A group of physicians in Germany were pursuing similar lines of thought, but for these figures and for the French interest in hypnotism, which exerted a strong influence on depth psychology, the reader must turn to histories of psychiatry.

During the 1870s several theories of unconscious organic memory were developed, and between 1880 and 1910 physicians and philosophers in many countries were concerned with various aspects of the unconscious (see references given in the surveys cited below).

FREUD

Sigmund Freud, even late in life, had no idea how extensive attention to the unconscious had been. Today we need to see him in perspective in order to strengthen what was weak in his ideas and so to advance toward a complete theory of the unconscious mind in health as well as in sickness. A more detailed survey of Freudian theory and method is given elsewhere; here we can treat only those aspects of his ideas that are directly relevant to the theory of the unconscious.

For Freud all mental processes are determined by natural laws, ultimately by those governing chemical and physical phenomena; they are associated with quantities of psychic energy that strive toward release and equilibrium; the primary driving force is instinctual energy (libido, a concept that was at first narrowly, then more widely interpreted) expressing an often unconscious wish, and moving from unpleasure to physical pleasure (pleasure principle); the predominant energy is sexual, but other forms are present, and Freud later assumed two basic instincts, sexuality in a broad sense and aggression (Eros and Thanatos). The establishment of civilized life involves restraints on sexual activity, and the unconscious proper (in Freudian theory the accessible unconscious being called the preconscious) consists of instinctual energies, either archaic or repressed during the life of the individual, particularly in childhood (universal incestuous desires of the earliest years, adolescent frustrated dreaming, aggressive impulses, etc.); these are available only through the use of special techniques. A genetic or developmental approach to mental illness is therefore essential. Forgetting is an active process in which painful memories are repressed.

The Freudian unconscious is a pool of mainly repressed energies, distorted by frustration and exerting a stress on conscious reason and its shaping of the patterns of daily life. The strain produced by this stress, present in some degree in all civilized men and women, is seen in neurosis. It is only by exceptional luck in heredity or experience that civilized man can avoid this tragic and potentially universal feature of modern life, the major influence of the unconscious being antagonistic to reason. This doom and neurosis he can escape (wholly, Freud thought at first; later he had doubts) by becoming aware of his situation and gaining insight into the particular traumatic experiences that created his neurosis. Freud began with an unquestioning conviction that insight brought recovery. The interpretation of dreams (which are symptoms and express wish fulfillment) and the process of free association can render accessible the regions of the unconscious producing the neurosis and can make possible a cure. Myths express for communities what dreams do for the individual. Later, Freud developed his ego theory, dividing the mind into three areas: the id, or basic instincts; the ego, or rational part of the mind that deals with reality; and the superego, a differentiated part of the ego that results mainly from the child's self-identification with his parents. This triple division overlaps awkwardly with the unconscious-conscious dichotomy, and here the theory becomes obscure. It left Freud unsatisfied—indeed, late in his life he stated that understanding of the deepest levels of the mind was not yet possible.

These are, in condensed form, the main ideas that make up the core of the Freudian theory of the unconscious, leaving aside his many applications of it. The theory, in its most characteristic form, is a description of the pathology of civilized man, although for Freud this

implied little restriction, since all suffer in some degree from the neurosis of civilization.

When this theory is reviewed today, most agree that Freud's general conception of a repressed unconscious, and its relation to child sexuality, aggression, defense mechanisms, sublimation, and so forth, is a permanent contribution of the highest importance. On the other hand, his sharp categories (conscious-unconscious, wishful-realistic, stages of sexual development, etc.) are merely, as he himself recognized, provisional steps toward the truth. But his theory suffers from a more radical weakness than these.

Freud's attitude toward the unconscious has been regarded as biological. But it was not so in a genuine sense, for all viable organisms display an organizing principle, not yet understood, which ensures that everything occurs in support of the continuation of life. This coordinating and formative principle underlies all organic properties, including the processes of the human unconscious, such as the imaginative and inventive faculties without which civilization could not have developed. It has been widely recognized that this factor—although it had been emphasized in earlier views of the unconscious, for example, by Cudworth, Goethe, Fichte, Schelling, Samuel Taylor Coleridge, and Carus—is not adequately represented in the Freudian theory, perhaps because it was neglected by the physicochemical approach to organisms dominant when Freud was shaping his ideas. His theory of the mind is overly analytic or atomistic and must be complemented by a general and powerful principle of coordination.

ADLER, JUNG, AND RANK

The lack of a general principle of coordination was recognized by three of Freud's colleagues, Alfred Adler, Carl Gustav Jung, and Otto Rank, who, from different points of view, stressed the potential integration and self-organizing power either of the unconscious or of the mind as a whole. Adler treated the person as a unity; he did not regard the unconscious-conscious division as basic and held that the inaccessible unconscious contains elements that have never been repressed but are simply not yet understood and are unconsciously assumed in the endeavor to adapt socially and to overcome supposed or real weaknesses.

The individual's aspiration or unconscious need to realize a potential unity was more deeply appreciated by Jung. He created the concept of the collective unconscious, which is not a "group mind" but the deepest level in the individual mind, consisting of potentialities for ways of thinking shared by all men because their genetic constitutions are closely similar and their family and social experiences share certain universal features. In a given society the collective unconscious contains particular traditional symbols or archetypes that organize thought and action. This sociological concept of the deeper mental levels involves a historical background in which ritual, myth, symbol, and religious attitude play organizing and integrating roles that contribute to the strength and stability of the psyche and that are subject to an underlying tendency developing a differentiated unity in the person (individuation). The tension of superficially opposed aspects in the unconscious mind produces autonomous foci of energy, acting as complexes. The ultimate aim for Jung was not discovery of truth but acceptance of the role of deep psychology in the present historical situation: assistance in the search for life-enhancing significance in the fate of living in a scientific age at a time when traditional sources of strength have been weakened but a fully comprehensive scientific truth is not yet in sight. In this search, psychology enters realms that previously belonged to history, philosophy, and religion. Jung's ideas form part of a discursive communication of attitudes, rather than being steps toward an ultimately confirmable theory of unconscious mental processes.

Rank stressed the role of religious and aesthetic traditions in shaping the unconscious, and he saw in the life will a factor making for integration. The writings of these three display agreement that Freud, particularly in his early work, overemphasized the role of genital sexuality, unduly neglected the historical background of the individual unconscious, and failed to allow for the role of factors making for coordination both within each Freudian level of the mind and between the various levels.

THE FUTURE OF THE CONCEPT

It has been observed (by Ira Progoff and others) that, mainly in their later years, Freud, Adler, Jung, and Rank all looked toward a future theory of the mind based on what perhaps can best be called the organic core of the mind (similar to Jung's objective psyche and psychoid) and capable of covering all human mental faculties, man's cultural history, his imagination, his mental illnesses and health. This still lies ahead. It seems that no important basic advance has been made in the theoretical understanding of the unconscious mind since then; certainly no one has yet made a satisfactory synthesis of the reliable features of their views. Thus, there has been a pause in the advance of the theory of subjective deep psychology.

Freud hoped for assistance from the neurophysiology of the brain, but this has not yet come.

We should now consider what the unconscious has stood for in the minds of different groups. The mystics saw it as the link with God; the Christian Platonists as a divine creative principle; the romantics as the connection between the individual and universal powers; the early rationalists as a factor operating in memory, perception and ideas; the postromantics as organic vitality expressed in will, imagination, and creation; dissociated Western man as a realm of violence threatening his stability; physical scientists as the expression of physiological processes in the brain that are not yet understood; monistic thinkers as the prime mover and source of all order and novelty in thought and action; Freud (in his earlier years) as a melee of inhibited memories and desires the main influence of which is damaging; and Jung as a prerational realm of instincts, myths, and symbols often making for stability. It is natural to seek a common principle underlying these partial truths, but we do not possess the unified language in which to express it scientifically.

The formulation of a valid theory of the integrated human mind and of its various pathologies would imply the possibility of a transformation in man and his unconscious toward a more harmonious condition accompanied by the development of a social order that does not bring with it inescapable neurosis. This may seem a distant hope. But recent advances in biology and medicine have opened new vistas of improvement, and no survey of the idea of the unconscious would be complete without a glance into this possible future for theory and practice, for therein may lie the deepest reason for the fascination that the idea has for so wide a public.

This sketch of the idea of the unconscious has neglected its recent applications to religion, art, the history of science, philosophy, literature (Marcel Proust believed that the reality of experience lies in the unconscious), ethics, and justice. In all these realms the main effect has been to broaden, deepen, and loosen traditional conceptions. But the unification of scientific principles, so badly needed today, still lies ahead. In this an improved conception of the unconscious must play a crucial role.

See also Adler, Alfred; Augustine, St.; Carus, Carl Gustav; Cudworth, Ralph; Descartes, René; Fechner, Gustav Theodor; Fichte, Johann Gottlieb; Freud, Sigmund; Goethe, Johann Wolfgang von; Hamann, Johann Georg; Hamilton, William; Hartmann, Eduard von; Hegel, Georg Wilhelm Friedrich; Herbart, Johann Friedrich; Herder, Johann Gottfried; Jung, Carl Gustav; Leibniz, Gottfried Wilhelm; Nietzsche, Friedrich; Proust, Marcel; Psychoanalytic Theories, Logical Status of; Rousseau, Jean-Jacques; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Thomas Aquinas, St.

Bibliography

HISTORICAL SURVEYS

It is remarkable that no authoritative critical study has yet been made of all ideas of the unconscious from earliest times to the present; the following works are useful historical surveys:

Ellenberger, H. "The Unconscious before Freud." *Bulletin of the Menninger Clinic* 21 (1957): 3.

Margetts, E. L. "Concept of Unconscious in History of Medical Psychology." *Psychiatric Quarterly* 27 (1953): 115.

Whyte, L. L. *The Unconscious before Freud.* New York: Basic, 1960; paperback ed., 1962.

Zilboorg, Gregory. *History of Medical Psychology.* New York: Norton, 1941.

COMMENTARIES

Drews, A. C. H. *Psychologie des Unbewussten.* Berlin, 1924.

Geiger, Moritz. "Fragment über den Begriff des Unbewussten und die psychische Realität." *Jahrbuch für Philosophie und phänomenologische Forschung* 4 (1921): 1–137.

Jones, Ernest. *The Life and Work of Sigmund Freud,* 3 vols. New York: Basic, 1953–1957.

MacIntyre, Alasdair C. *The Unconscious: A Conceptual Analysis.* New York, 1958.

Miller, J. G. *Unconsciousness.* New York, 1942.

Northridge, William L. *Modern Theories of the Unconscious.* London, 1924.

Progoff, Ira. *Death and Rebirth of Psychology.* New York, 1942.

Taylor, W. S. "Psycho-analysis Revised or Psychodynamics Developed?" *American Psychologist* (November 1962): 784.

ORIGINAL WORKS

Carus, C. G. *Psyche.* Pforzheim, 1846; 3rd ed., Stuttgart, 1860.

Cudworth, Ralph. *True Intellectual System of the Universe.* London, 1678.

Freud, Sigmund. *Die Traumdeutung.* Vienna, 1900. Translated by James Strachey as *The Interpretation of Dreams,* in *Standard Edition of the Complete Psychological Works,* James Strachey, general editor, 24 vols. New York, 1953–1964. Vols. IV and V.

Freud, Sigmund. "The Unconscious." In *Collected Papers,* edited by James Strachey, 5 vols. London: Hogarth Press, 1924–1950. Vol. IV, *Papers on Metapsychology and Papers on Applied Psychoanalysis,* p. 98.

Hartmann, Eduard von. *Die Philosophie des Unbewussten,* 3 vols. Berlin, 1869. Translated by W. C. Coupland as *The Philosophy of the Unconscious.* London: K. Paul, Trench, Trubner, 1883.

Jung, C. G. *Über die Psychologie des Unbewussten.* Zürich: Rascher, 1943. Translated by R. F. C. Hull as *Two Essays on Analytical Psychology,* Vol. VII of *Collected Works.* New York, 1953; paperback ed., New York: Meridian, 1956.

Jung, C. G. *Wandlung und Symbole der Libido.* Leipzig, 1912. Translated by Beatrice M. Hinkle as *Psychology of the Unconscious.* London, 1916. Rev. ed. and new translation by R. F. C. Hull from the 4th ed. (1952), *Symbole der Wandlung,* as *Symbols of Transformation,* published as Vol. V of *Collected Works,* Herbert Read et al., eds. London and New York, 1956.

Leibniz, G. W. *Die philosophischen Schriften.* Edited by C. J. Gerhardt, 7 vols. Berlin, 1875–1890. Vol. V, p. 48, Vol. VI, p. 600.

Schopenhauer, Arthur. *Die Welt als Wille und Vorstellung.* Leipzig, 1819. Translated by R. B. Haldane and J. Kemp as *The World as Will and Idea,* 3 vols. London: Trubner, 1883–1886.

Lancelot Law Whyte (1967)

UNDERDETERMINATION THESIS, DUHEM-QUINE THESIS

Underdetermination is a relation between evidence and theory. More accurately, it is a relation between the propositions that express the (relevant) evidence and the propositions that constitute the theory. The claim that evidence underdetermines theory may mean two things: first, that the evidence cannot prove the truth of the theory, and second, that the evidence cannot render the theory probable. Let us call the first deductive underdetermination and the second inductive (or ampliative) underdetermination. Both kinds of claims are supposed to have a certain epistemic implication, namely that belief in theory is never warranted by the evidence. This is the underdetermination thesis.

DEDUCTIVE UNDERDETERMINATION

Deductive underdetermination is pervasive in all interesting cases of scientific theory. If the theory is not just a summary of the evidence, the evidence cannot determine, in the sense of proving, the theory. For instance, no finite amount of evidence of the form Aa_i & Ba_i can entail an unrestricted universal generalization of the form *All A's are B*. Deductive underdetermination rests on the claim that the link between evidence and (interesting) theory is *not* deductive. What is the epistemic problem it is supposed to create? Given that the link is not deductive, it is claimed that we can never justifiably believe in the truth of a theory, no matter what the evidence is. However, it would be folly to think that deductive underdetermination creates a genuine epistemic problem. There are enough reasons available for the claim that belief in theory can be justified even if the theory is not proven by the evidence: Warrant-conferring methods need not be deductive.

Deductive underdetermination speaks against simplistic accounts of the hypothetico-deductive method, which presuppose that the epistemic warrant for a theory is solely a matter of entailing correct observational consequences. Two or more rival theories (together with suitable initial conditions) may entail exactly the same observational consequences. Given the above presupposition, it follows that the observational consequences cannot warrant belief in one theory over its rivals. Though simplistic accounts of the hypothetico-deductive method need to be jettisoned, there are ways to meet the challenge of deductive underdetermination, even if we stay close to hypothetico-deductivism. Since theories entail observational consequences only with the aid of auxiliary assumptions, and since the available auxiliary assumptions may change over time, the set of observational consequences of a theory is not circumscribed once and for all. Hence, even if, for the time being, two (or more) theories entail the same observational consequences, there may be future auxiliary assumptions such that, when conjoined with one of them, they yield fresh observational consequences that can shift the evidential balance in favor of it over its rivals. Besides, a more radical (though plausible) thought is that theories may get (indirect) support from pieces of evidence that do not belong to their observational consequences.

INDUCTIVE UNDERDETERMINATION

Inductive underdetermination takes for granted that any attempt to *prove* a theory on the basis of evidence is futile. Still, it is argued, no evidence can confirm a theory or make it probable, or no evidence can confirm a theory more than its rivals. This claim is rather odd. In all its generality, it is a recapitulation of inductive skepticism. If induction lacks justification, then no inductively established theory is warranted by the evidence. Yet induction does not lack justification. In any case, according to recent externalist-reliabilist theories of justification, belief in theory is justified if induction is reliable; and there is no argument that it is *not*. If inductive scepticism is set aside, inductive underdetermination must relate to problems with the theory of confirmation. For on *any* theory of confirmation, the evidence (even if it is restricted to observational consequences) can render a theory probable or more probable than its rivals. That is, the evidence can raise the probability of a theory. So inductive underdetermination must rest on some argu-

ments that question the confirmatory role of the evidence vis-à-vis the theory. There is a battery of such arguments, but they may be classified under two types.

The first capitalizes on the fact that no evidence can affect the probability of the theory unless the theory is assigned some nonzero initial probability. In fact, given the fact that two or more rival theories are assigned different prior probabilities, the evidence can confirm one more than the others, or even make one highly probable. The challenge, then, is this: Where do these prior probabilities come from? A total denial of the legitimacy of any prior probabilities would amount to inductive skepticism. Inductive underdetermination would be inductive skepticism. The more interesting version of inductive underdetermination does not challenge the need to employ prior probabilities, but rather their epistemic credentials. If, it is argued, prior probabilities have epistemic force, then the evidence can warrant a high degree of belief in a theory (or greater degree of belief in a theory than its rivals). But, it is added, how can prior probabilities have any epistemic force?

The subjective Bayesians' appeal to subjective prior probabilities (degrees of belief) accentuates rather than meets this challenge. Bayesians typically argue that, in the long run, the prior probabilities wash out: even widely different prior probabilities will converge, in the limit, to the same posterior probability, if agents conditionalize on the same evidence. But this is scant consolation because, apart from the fact that in the long-run we are all dead, the convergence-of-opinion theorem holds only under limited and very well-defined circumstances that can hardly be met in ordinary scientific cases. The alternative is to claim that prior probabilities have epistemic force because they express rational degrees of belief, based, for instance, on plausibility or explanatory judgements. This claim faces many challenges, but its defense might well be necessary for blocking the epistemic implications of inductive underdetermination. In its favor, it can be said that rational belief in theory is not solely a matter of looking for strict observational evidence.

The second type of argument rests on the claim that theories that purport to refer to unobservable entities are, somehow, unconfirmable. The problem is supposed to be that since there cannot be direct observational access to unobservable entities, no observational evidence can support the truth of a theory that posits them, and no evidence can support a theory more than others that posit different unobservable entities. The distinctive element of the second type of argument is that the resulting inductive underdetermination is selective. It does *not* deny that observational generalisations can be confirmed. Hence, it does not deny that the evidence can confirm or render probable observational theories. It denies that the same can be the case for theories that refer to unobservable entities.

Even if a sharp distinction between observable and unobservable entities were granted (though it is by no means obvious that it should), this selective inductive underdetermination has a bite only if the methods that lead to, and warrant, belief in observable entities and observational generalizations are different from the methods that lead to, and warrant, belief in theories that posit unobservable entities. Yet the methods are the same. In particular, explanatory considerations play an indispensable role in both cases. In the end, this kind of selective inductive underdetermination undermines itself: it either collapses into inductive skepticism or has no force at all.

EMPIRICAL EQUIVALENCE

It is commonly argued that there can be totally empirically equivalent theories— that is, theories that entail exactly the same observational consequences under any circumstances. In its strong form, this claim (let's call it the Empirical Equivalence Thesis, *EET*) asserts that *any* theory has empirically equivalent rivals (some of which might be hitherto unconceived). *EET* is an entry point for the epistemic thesis of total underdetermination: that there can be no evidential reason to believe in the truth of *any* theory. But there is no formal proof of *EET*, though a number of cases have been suggested ranging from Descartes' "evil demon" hypothesis to the hypothesis that for every theory T there is an empirically equivalent rival asserting that T is empirically adequate yet false, or that the world is *as if* T were true. One can, of course, argue that these rival hypotheses have only philosophical value and drive only an abstract philosophical doubt. In science, it is often hard to come by just one totally empirically adequate theory, much less a bunch of them.

Yet it seems that there is a genuine case of empirical equivalence of theories of quantum mechanics. Alternative interpretations of the quantum-mechanical formalism constitute empirically equivalent but different theories that explain the world according to different principles and mechanisms. The most typical rivalry is between the orthodox understanding of quantum theory—the "Copenhagen interpretation," according to which a particle cannot have a precise position and momentum at the same time—and the Bohmian understanding of quantum theory—the hidden-variables inter-

pretation, according to which particles always have a definite position and velocity, and hence momentum. On Bohm's theory, particles have two kinds of energy: the usual (classical) energy and a "quantum potential" energy. More recently, there have been three particularly well-developed theories (the Bohmian quantum mechanics, the many-worlds interpretation, and the spontaneous-collapse approach) such that there is no observational way to tell them apart. And it seems that there *cannot* be an observational way to tell them apart. This situation is particularly unfortunate, but one may respond that the ensued underdetermination is local rather than global; hence the possible skepticism that follows is local.

The Duhem-Quine thesis has been suggested as an algorithm for generating empirically equivalent theories. Briefly put, this thesis starts with the undeniable premise that all theories entail observational consequences only with the help of auxiliary assumptions and concludes that it is always possible that a theory, together with suitable auxiliaries, can accommodate *any* recalcitrant evidence. A corollary, then, is that for any evidence and any two rival theories T and T', there are suitable auxiliaries A such that T' and A will be empirically equivalent to T (together with its own auxiliaries). Hence, it is argued, no evidence can tell two theories apart. It is questionable that the Duhem-Quine thesis is true. There is no proof that *nontrivial* auxiliary assumptions can always be found.

But let us assume, for the sake of the argument, that it is true. What does it show? Since the Duhem-Quine thesis implies that any theory can be saved from refutation, it does create some genuine problems to a falsificationist (Popperian) account of theory testing— that is, the view that theories are tested by attempting to refute them. If attempted refutations are the sole test for theories, two incompatible theories that are not refuted by the evidence are equally well tested by it. But the Duhem-Quine thesis does not create a similar problem to an inductivist. From the fact that any theory can be suitably adjusted so that it resists refutation it does not follow that all theories are equally well confirmed by the evidence. An inductivist can argue that the empirical evidence does not lend equal inductive support to two empirically congruent theories. It is not necessarily the case that the auxiliary assumptions that are needed to save a theory from refutation will themselves be well supported by the evidence. Since it is reasonable to think that the degree of support of the auxiliary assumptions associated with a theory is reflected in the degree of support of the theory, it follows that not all theories that entail the same evidence are equally well supported by it.

EET has generated much philosophical discussion. An argument favored by the logical positivists is that such cases of total underdetermination are illusions: the rival theories are simply notational variants. This move presupposes that theories are not taken at face value. For anyone who does not subscribe to a verificationist criterion of meaning, this move is moot. It does make sense to say that there *can* be distinct but totally empirically equivalent theories. The hard issue is not to exclude their possibility on a priori grounds but to find ways to distinguish their epistemic worth, should we find ourselves in such a predicament.

Another move, favored by Quine, is to go for pragmatism: The balance is shifted to the theory *we* (our community) favor, simply because it is *our* theory. This raises the spectre of epistemic relativism. Yet another move is to go for skepticism: among rival totally empirically equivalent theories one is true, but we cannot possibly come to know or justifiably believe which this is. This skeptical answer might be supplemented with some differential stance towards the rival theories, but this differential treatment will not be based on epistemic reasons but rather on pragmatic considerations. Indeed, social constructivists have seized upon this in order to claim that social, political, and ideological factors break observational ties among theories: hence, they argue, belief in theory is socially determined.

The general problem with the skeptical move is that it rests on a restricted account of what counts as evidence (or reason) for justified belief; it counts only observations as possible epistemic reason for belief. But rational belief may well be a function of other epistemic reasons—for instance, the theoretical virtues that a theory possesses. This last thought ushers in yet another possibility: that empirically equivalent theories may well differ in their explanatory power. Insofar as explanatory power can offer epistemic credentials to a theory, it can break supposed epistemic ties among totally empirically equivalent rivals. This move makes rational belief a more complex affair and tallies with the intuitions of scientific and common sense. Yet it faces the problem of justifying the claim that theoretical virtues are epistemic reasons— that is, that a virtuous theory (a theory with great explanatory power) is more likely to be true than a less virtuous one.

This is not an unsolvable problem. There are, broadly, two ways to tackle it. One is to argue (rather implausibly) that some theoretical virtues are constitutive marks of truth. The other is to argue for a broad concep-

tion of evidence that takes the theoretical virtues to be empirical and contingent marks of truth. A central element in this latter argument is that theories can get extra credence by entailing novel predictions—that is, predictions such that information about the predicted phenomenon was not previously known and not used in the construction of the theory. In the end, the epistemic relations between evidence and theory cannot be exhausted by their logico-semantic relations.

See also Confirmation Theory; Scientific Realism.

Bibliography

Cushing, J. T. *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*. Chicago: The University of Chicago Press, 1994.

Devitt, Michael. "Underdetermination and Realism." *Philosophical Issues* 12 (2002): 26–50.

Earman, John. "Underdetermination, Realism, and Reason." *Midwest Studies in Philosophy* 18 (1993): 19–38.

Laudan, Larry. "Demystifying Underdetermination." *Minnesota Studies in the Philosophy of Science* 14 (1990): 267–297.

Newton-Smith, William. "The Underdetermination of Theory by Data." *Aristotelian Society* Suppl. 52 (1978): 71–91.

Psillos, Stathis. *Scientific Realism: How Science Tracks Truth*. London and New York: Routledge, 1999.

Quine, W. V. "On Empirically Equivalent Systems of the World." *Erkenntnis* 9 (1975): 313–328.

Stathis Psillos (2005)

UNITY AND DISUNITY IN SCIENCE

Unity covers a wide range of loosely connected ideas in science, differently analyzed by different interpreters. Generally, they are expressions, or echoes, of the idea that science can succeed in providing one consistent, integrated, simple, and comprehensive description of the world. This entry will provide a historical perspective on such ways of thinking about unity in science. (Readers should bear in mind that the real history is much more complex and interesting than the following microsketch, which is intended only to introduce the leading ideas.)

MECHANISMS AND LAWS

The scientific revolution of the seventeenth century involved consolidation of the "mechanical (or corpuscularian) philosophy" according to which natural phenomena are to be understood in terms of shaped matter in motion, with the natural world likened to a giant mechanism. Natural philosophy could look for unity in this regard by thinking of the parts of the world machine as all governed by the same simple set of rules or laws. Isaac Newton's mechanics could be seen in this regard as a paradigm of unification, showing how the same laws covered motion in both the heavens and on Earth.

But there was a monkey wrench in this mechanist paradigm: Newton's law of gravity involved "action at a distance," inadmissible by most seventeenth-century interpreters as a legitimate mechanical principle. Mechanism required contact action. Newton's official response was that "I make no hypotheses," that is, no hypotheses or speculations about what the underlying real mechanism of gravity might be. Instead, he presented his mechanics as "mathematical only," that is, mathematical principles by which motions can be reliably and accurately described but with no pretense to describing what makes things move as they do. Accordingly, some of Newton's successors thought of unity in theory and in science in terms of a simple set of general, mathematical laws that integrate, by covering, a wide range of phenomena that otherwise might seem independent, and all this without any thought of underlying mechanisms. This will be referred to as the "nomological attitude."

These two ideas, seeing disparate phenomena as manifestation of one underlying mechanism or covered by one set of simple laws, interacted and intertwined during the eighteenth and nineteenth centuries. For example, James Clerk Maxwell worked to treat first electric and magnetic effects and then discovered he could also cover optical phenomena, thinking of all of these first as manifestations of one underlying mechanism, developing the laws that might govern such a mechanism, and then letting go of the postulated underlying mechanism as unverifiable speculation in favor of the general laws that had emerged. Heinrich Rudolf Hertz maintained that Maxwell's theory is Maxwell's equations, and eventually Albert Einstein's special relativity did in the speculated stuff of electromagnetic mechanisms, the luminiferous aether.

The opposition of mechanisms versus laws also played out, with the opposite result, during the second half of the nineteenth century over the issue of atoms. The predictive and explanatory success of chemistry, as well as the nascent kinetic theory (statistical mechanics), emboldened some to see atoms and molecules as real cogs in the cosmic machine. Others scoffed at postulation of things too small to see or individually detect as "metaphysics," not science. Continuum mechanics and even contact action presented severe problems for an atomistic

theory. The speculated indivisibility of atoms, though mentioned by some, was not really the issue. Rather, it was whether one could correctly think of the underlying order in terms of discrete parts interacting in something like the mechanist tradition or whether this should be seen, at best, as a kind of pretty imaginative picture, while scientific truth was exhausted by mathematical laws in the nomological tradition.

The issue of atoms came to a head in the first decade of the twentieth century in the work augmented and integrated by Jean-Baptiste Perrin. Perrin catalogued the astonishingly numerous and diverse facts that could be encompassed by postulating atoms: constant ratios in chemistry, relative atomic weighs, diffusion and other fluctuation phenomena, osmotic pressure, behavior of electrolytes, specific heat, behavior of thin materials, even why the sky is blue. Perrin tabled sixteen independent ways of reaching the same estimate of Avogadro's number. Einstein's theory of Brownian motion proved especially effective—in a sense one could "see" the causal effects of individual molecular collisions. A vast range of otherwise diverse observable phenomena were unified in the sense of interpreting them as the manifestation of the properties and behavior of atoms. By 1913 most of the physics community accepted atoms as real.

Electric, magnetic, and optical phenomena unified by Maxwell's laws. Perrin's diverse phenomena unified by postulation of atoms. Though they are in some ways polar attitudes, mechanistic and nomological thinking really cannot operate without one another. To provide unifying explanations, mechanisms need to be governed by laws, and laws, if they are to do more than exhaustively list superficially observable phenomena, must at least have the form of describing some conceptually more economic structure.

REDUCTIONISM

The nineteenth century saw explosive development of the natural sciences, emboldening some toward the end of the century to speculate that physics was almost completed with little left to do but to work out the applications to other natural phenomena. Contrary to what one might have imagined, the shocks of relativity and quantum mechanics in the first quarter of the twentieth century initially encouraged rather than tempered such scientific utopian attitudes. Some strands of positivism in the second quarter of the century described unity of science in terms of unity of language and methods; others took the spirit of unification to its logical extreme, emphasizing axiomatic formulation and developing the idea of reduction of all natural phenomena to "fundamental physics" in the spirit of the logicists' hope of reducing all of mathematics to logic. By the 1950s and 1960s reductionistic thinking had taken a deep hold on much thinking in both philosophy and science, no doubt encouraged by advances within science in subjects such as quantum chemistry and microbiology. Unity now took the form of (expected) chains of reductive definitions, identifying not just complex physical, but biological, psychological, and social phenomena with the behavior of physical parts, everything ultimately to be described in terms of the laws of fundamental physics.

Again a monkey wrench, or this time two, brought the reductionist juggernaut to a halt. In the 1970s and 1980s philosophy of science became acutely aware of difficulties with the whole reductionist program. The reversal began with the collapse of the two show cases: claimed deductive reduction of thermodynamics to statistical mechanics and of Mendelian to molecular genetics. Temperature is in fact realized by mechanisms in addition to mean kinetic energy, and in principle could be realized in indefinitely many ways. There is no neat one trait–one gene correlation and the developmental effects of any one bit of DNA depend, not just on its genetic, but on its overall environmental context. If temperature and genes are multiply realizable by disparate physical constructs, then surely also, for example, are mental states. Higher level objects and phenomena may still all be physically realized, but in such diverse ways that the program of reduction by definitions and deduction loses plausibility. *Unity* no longer seems such an apt term.

This first basis for some kind of disunity was followed in the 1980s and 1990s by a second. Nancy Cartwright, Ronald N. Giere, and others have pointed out that, whatever the ultimate aims of science or of some scientists might be, the science we actually have, now or any time in the foreseeable future, hardly follows the pattern of calculation of phenomena from universally applicable, exact, true laws or of description in terms of mechanisms known or even believed to operate exactly as described. Rather, science uses laws in the construction of idealized models, always limited in scope, and even where they apply never exactly correct. Rather than providing descriptions that set out exactly what the phenomena are, the laws of science are only true, or at least only exactly true, of the idealized models that in turn enable us to understand phenomena and their hidden sources in terms of the idealizations to which the phenomena are similar. For the puny minds of even the best physicists, to understand the fluid properties of water we need to resort

to continuum hydrodynamical models, while to understand dispersive phenomena we turn to the discontinuous models of statistical mechanics. "Foundational" theories fare no better. Quantum field theory and general relativity each idealize away from the phenomena of the other, are mutually inconsistent, and have no humanly accessible direct application to most phenomena of human interest. The science we have displays disunity on a grand scale.

UNITY AND DISUNITY IN SCIENCE

Or does it? Few would dispute the claims just listed about science and idealized models. But many challenge the interpretation of these facts as constituting disunity in any weighty sense. Since *unity* and *disunity* have no well-established univocal usage and are susceptible to expropriation as rhetorical weapons by advocates of one or another larger position, we have difficulty in saying just what the issue really is, let alone in resolving it. Yet there are interesting and important issues here, ones that it is suggested we do not understand at all well. For elaboration let us, with hindsight, revisit the unification afforded by the postulation of atoms.

Descriptions of none of the phenomena described as manifestations of the existence and behavior of atoms follow from the bare postulation of atoms alone. We require assumptions, not only about the properties and behavior of atoms but also—for many of the phenomena—about a great deal else. The accounts based on the postulation of atoms hardly constitute the deductions imagined by the reductionists. Rather, they work, often fortuitously, by appealing to a helter-skelter of plausible assumptions, phenomenological observations, disconnected results from other accounts, and a wide range of approximative mathematical methods and experimental techniques from independently practiced fields. Nonetheless, all these accounts have at their core the assumption that material is composed of relatively stable and discrete parts with properties that admit of systematic investigation. In all the admittedly disunified messy process of science the postulation of atoms is doing real and systematic work—we would not have this body of accounts without the postulation of atoms.

This kind of intertheoretic asymmetry occurs broadly. Quantum theory plays a role in understanding chemistry that chemistry does not play in understanding quantum theory, and similarly for chemistry and biology, biology and psychology, and many other pairs of theories and theoretical domains. Clearly, such asymmetry has to do with the circumstance that parts of an object or process play crucial roles in the behavior of the containing whole. But one does not yet understand at all clearly the nature of such intertheoretic relations—reductionism was a vast oversimplification. The mirage of some kind of simple unity was the artifact of imagining that the human mind could get its head around all of the natural world, exactly and, at least potentially, in all its detail. This will not happen, at least not until long after this encyclopedia has become hopelessly out of date. In the mean time we face the complex and interesting challenge of charting the complex interplay of elements of unity and disunity in the science we know.

See also Philosophy of Physics; Reduction; Special Sciences.

Bibliography

Darden, Lindley, and Nancy Maull. "Interfield Theories." *Philosophy of Science* 44 (1977): 43–64

Maudlin, Tim. "On the Unification of Physics." *Journal of Philosophy* 93 (1996): 129–144.

McRae, Robert F. *The Problem of the Unity of the Sciences: Bacon to Kant*. Toronto: University of Toronto Press, 1961.

Morrison, Margaret. *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. New York: Cambridge University Press, 2000.

Nagel, Ernest. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace, 1961.

Oppenheim, Paul, and Hilary Putnam. "The Unity of Science as a Working Hypothesis." In *Minnesota Studies in the Philosophy of Science*. Vol. 2, edited by Herbert Feigl, Michael Scriven, and Grover Maxwell. Minneapolis: Minnesota University Press, 1958.

Wilson, Mark. "Mechanics, Classical." In *The Routledge Encyclopedia of Philosophy*. Vol. 6, edited by Edward Craig, 251–259. London: Routledge, 1998.

Paul Teller (2005)

UNIVERSAL PROPERTIES IN INDIAN PHILOSOPHICAL TRADITIONS

EARLY GRAMMARIANS ON UNIVERSALS OF WORDS AND MEANINGS

In ancient India systematic metaphysics started with a linguistic turn. Ontological concepts and controversies arose in the context of musings on meanings of words and debates on declensions, unlike in ancient Greece,

where metaphysics arose out of wondering about numbers, figures, and nature. In Pāṇini's grammar and his early commentaries (between the fourth and second centuries BCE) the three crucial technical terms for a universal—*sāmānya*, *jāti*, and *ākṛti*—were already explicitly in use. Philosophers of language dabbled in metaphysics since Patañjali's "Great Commentary" to Pāṇini's grammar. The device of adding a *tva* or *tā* (roughly equivalent to the English "ness") to any nominal root x, yields, as meaning, the property of being x. From substance (*dravya*) one can thus mechanically abstract substanceness (*dravya-tva*), from real (*sat*) and reality (*sattā*). With this device in place it was natural to make the distinction between an individual substance and the property that makes it what it is, its abstract essence. But even to parse this talk of concrete cows rather than of the bovine essence, the grammarians drew the distinction between talking about one particular cow and talking about any cow or a cow in general (VMB on Pāṇini sutra 1.2.58 and 1.2.64). The distinction between the general and the particular also came up for discussion in the context of the logic of pluralization. What allowed one to say "trees" or "men" instead of using the word for a tree or man as many times as the number of trees one referred to? It must be because the direct meaning of a common noun is the shared universal property of the referents that one could eliminate all but one remaining occurrence of that word, when speaking generally. One could also issue universalizable moral imperatives such as, "A cow ought not to be killed," which, Patañjali jokes, is not obeyed by simply sparing the life of one single cow.

Jāti (a word that, in modern Indian vernaculars, has come to mean "a class," "a caste," or even "a nation")—the Sanskrit counterpart of the Latin "genera"—is used by Pāṇini for a shared property of all the particulars of one natural kind, which serves also to distinguish any one of them from things of other kinds. The particulars are called *vyakti*—a word that etymologically suggests a distinct concrete manifestation of common and uncommon properties. The problem with this universalist theory of meaning—defended by Vajapyāyana—was that when, in a descriptive or prescriptive sentence, the action denoted by the verb has to hook up with what the noun means, it has to be a particular. For, after all, no one can bring cowness, cut the tree-essence, or meet humanity on the street.

Thus, in Indian philosophical semantics the dispute between those who insisted that a word primarily means a universal and their rivals who held that it must be particular substances that are the first meanings of words is at least twenty-two centuries old. The word often used for *universal* by Patañjali was *ākṛti* (literally "shape"), which is more reminiscent of form than a property. In answer to the basic question "What is a word?" Patañjali considers the option, "Is it that which remains non-distinct among distinct individuals, un-torn when individuals are torn down?" and answers, "No, that is not the word, that is only the universal (*ākṛti*).

The need to switch to imperishable universals as meanings was felt both by the grammarians and the Mīmāṃsā school of Vedic hermeneutics for whom the authority of authorless sentences of the Vedas rested on their eternity. The relation between words and objects was said to be entrenched and eternal. If perishable particular horses, cows, humans, and plants were the meanings of words, how could they be the eternally connected meanings of these beginningless Vedic words? The word *gauḥ* (cow) is therefore best taken to be eternally connected to the timeless bovine essence.

The first clear recognition of the need to postulate universals might have come, not so much from the theory of meaning but from reflecting on the generality or repeatability of the audible words themselves. That there could be many pronunciations or distinguishable phonations of the same word was seen to be an unquestionable example of the one-in-many. That naturally went hand in hand with the idea of the real word-type existing timelessly there independently of its temporal perishable token-utterances. Later, in the philosophy of Bhartrhari, sometimes called a linguistic nondualist, word-universals and meaning-universals and one's natural tendency to superimpose the former on the latter were elaborately discussed, because it was easy to confuse them with Bhartrhari's single most important metaphysical concept of a speech-bud or linguistic-potentiality (*sphoṭa*) in all consciousness, where signifier and signified exist undivided, waiting to blossom into articulated structures of sentences.

In the context of interpreting Kātyāyana's aphorism, "the word-meaning-relation being fixed," Patañjali mentions two alternative ways of taking the concepts of form (*ākṛti*) and content or substance (*dravyam*). In the first sense forms are universal properties that remain unchanged while individual material substances come and go, hence the forms must be those fixed meanings. In the second sense, somewhat like René Descartes's lump of wax, the substance continues to exist, retaining its sameness while the structures vary or perish, hence the substance or content must be that fixed meaning. If one defines the universal as the invariant across variations of individuals, then that definition fits both the form (under

the first interpretation) as well as the indestructible content (under the second interpretation). One encounters a similar clash of intuitions in Ludwig Josef Johann Wittgenstein's *Tractatus Logico-Philosophicus*, where, about the ultimate constituents of all atomic facts, one finds the remark, "Objects are form and content." This idea of the enduring stuff of changing entities as a ground of sameness, found in early grammarians' and Advaita Vedāta thought, was later on picked up by the Jaina notion of a vertical universal (*ūrdhvata-sāmānya)*, as against the more common property-universal that was termed *horizontal universal* (*tiryak-sāmānya)*.

THE HOT TOPICS FOR DEBATE

Between fifth and fifteenth centuries the debate between mainstream Nyāya-Vaiśeika and Mīmāṃsā realists and Buddhist nominalists raged around the existence of eternal essences. The major points of disputation were:

(1) Must one explain the use of a common noun or the experience of community across a plurality of particulars by postulating a single real property inherent in each of those particulars? (Vaiśeika and Mīmāṃsā said yes with some caveats, and Buddhists said no.)

(2) Is this property totally distinct from the individuals that exemplify it? (Vaiśeika said yes, and Bhāṭṭa Mīmāṃsā said yes and no.)

(3) Does a universal exist only in all its own instances or are universals omnipresent ? (This is a trick question set up by the Buddhist nominalist, answered cautiously by Vaiśeika.)

(4) Do universals have any role in causation? (Vaiśeika said that they can cause one's awareness of them. For Buddhists anything that is eternal must be causally barren, hence nonexistent. For Udayanācārya [tenth-century Nyāya-Vaiśeika] nomic relations of necessary concomitance are ontologically founded on the universals inherent in the causes and effects.)

(5) Can the work that is done by universals be done by relations of resemblance between particulars? (Vaiśeika said no, Jainism and Madhva Vedāta said yes).

THE CLASSICAL NYĀYA-VAIŚEIKA REALISM ABOUT UNIVERSALS

Universals come to occupy a crucial role as the fourth type of real, in the scheme of six basic categories of reals or "things-meant-by-words" (*padārthas*)—notice again the semantic orientation—listed in the Vaiśeika sutras of Kaṇāda. In that canonical scheme, after the three types of unrepeatables—substances, particular qualities, and motions—come common properties. Although substances, qualities, and motions are entities of different types, they share one common property: They are all real. What is this realness that is common to all substances, qualities, and motions? Realness is a generic essence present in many substances, qualities, and motions. It is a universal, the highest one. Then there are less general features as well, the substance-hood shared by all substances, the quality-hood common to all qualities, and the motion-hood inherent in all motions. These second-tier universals are called common-uncommon since they function as defining properties belonging to all the members of the class to be defined, and lacked by all else.

The Vaiśeika sutra's word for universal is *sāmānya*" (the phonetic resemblance with "sameness" may not be entirely accidental), meaning "what is common." The word for an individuator or particularity is *víseṣa*, which means "uncommon feature" or "specialty," the difference-maker. Flower-ness could be a common property, shared by roses, jasmines, and sunflowers. But the same property would be a difference-maker when you compare a rose with fruits, seeds, stones, and animals, since none of these except the rose has flower-ness. Hence, Kaṇāda's aphorism, "Universal and particularity depend upon understanding" (VS:1/2/3).

Commentators hasten to point out that this formulation does not mean that universals are subjective or invented by one's ways of understanding the world. All it means is that one finds out by the verdict of one's understanding whether some property is a pure universal or also a demarcator, as shown earlier.

Four broad arguments are generally proposed by these staunch realists for proving the existence of universal properties:

(1) The evidence of sense-perception is the strongest of all. Unless one is threatened by a logical inconsistency, one must admit some common recurrent entity in each of those many things that sense-perception shows one to be of the same kind. This class-character, the basis for one's sense of sameness (*anugata-pratīti*), is a universal.

(2) The argument from the meaning of general words runs as follows. A learnable common noun such as *bird* can denote an unlimited number of particulars of enormous variety. How the same word with the same meaning can correctly apply to so

many diverse particulars calls for an explanation. The explanation must lie in a distinction between reference (*śakya*) and sense (*śakyatāvacchedaka*). Thanks to the existence of an objective universal, for example, bird-ness, which serves as the same sense, the same word can distributively refer to all birds or any bird. This does not boil down to one of the early extreme views that the bare particular or the pure universal is the primary meaning of a word. It is the balanced view that the meaning of a word is a particular possessing a general property that serves as the common mode of presentation of its unlimited number of referents.

(3) Then one has the argument from lawlike causal connections. Fire is a substance, but when it causes burning, its causal efficacy is not determined by its simply being a substance, for, then any substance would burn. To explain what makes fire—and not any other substance—the cause of burning, one needs to postulate fire-ness as the property that limits the causality of fire toward this effect. With the advent of extremely technical New Nyāya (around the thirteenth century) the need to have limiters (*avacchedaka*) of cause-hood and effect-hood became the standard ground for ontological commitment to universals.

(4) Admission of universals also helped Nyāya solve the problem of justifying the inductive leap from observation of a few cases to a universal generalization covering all cases of a concomitance. The common property observed in a few instances can, as it were, put one in direct perceptual touch with all the other instances where also it inheres, not in their individual details but in a generic way. Here, the universal itself is supposed to play the role of the operative connection between the sense-organ and the apparently unobserved instances of that universal.

With all these supporting arguments for its existence, the precise definition offered by Nyāya-Vaiśeika settled down to this, "A universal is that which, being eternal, is inherent in many." Not any quality inhering in a substance is a universal. A wish inheres in a soul, but it is a short-lived episode, not a universal. Colors are not universals in this system because they are unrepeatable qualities clinging to the particular surfaces. All colors share the universal color-hood. But two red apples have two distinct red colors in them, just as each of them would have a distinct falling-motion when they both fall. A universal must subsist wholly in each of its instances by the special relation of inherence. A universal must be wholly inherent in each of its instances. The word *inherent* must be taken seriously. A single string may be running through many flowers, but it is only in contact with them, the whole string is not inherent in any one of them.

What is inherence? It is a kind of being-in, the converse of which is an intimate "having." Humanity inheres in me, just in case I have humanity. Now, having can be of many kinds. Things have qualities and motions. Wholes have parts. I have a pen in my hand. A rich man has a big house. The logical structure of each of these relations of characterization, constitution, contact, and ownership, however, is utterly different. All four are more or less aptly reportable by the use of the preposition *in* or *of*: the taste is in the apple, the room is or consists in the walls, roof, and floor, the pen is in between the fingers, and the house is of the rich merchant. Still, one initial grouping could be made to clarify their distinct structures. The taste and the room cannot exist without the apple or the room-parts. The taste cannot float about on its own, minus the apple. The room cannot stand independently of the walls. But that pen can easily exist untouched by the hand, and that house can change hands.

So, the first two relations hold between pairs that are "incapable of standing apart from one another" (*ayutasiddha)*, whereas the other two relations hold between pairs that are "capable of standing apart from one another" (*yutasiddha)*. However tightly my ring is stuck to my finger, it is not inherent in it as inseparably as finger-ness is inherent in my fingers. It is no physical glue but a metaphysical inseparability that joins the goat-ness to the goat, ties up the running and black color of the goat to the goat, as well as binds the goat to its body-parts. The kind of being-inseparably-in that connects the universal to its instances has to be distinguished from the way a berry lies in a bowl. For the sake of economy—the principle of not multiplying entities beyond necessity—the mainstream Nyāya-Vaiśeika metaphysicians posit only one single such relation as enough to link innumerable pairs of universals and particulars, qualities and substances, and wholes and parts. For systemic reasons, this relation is supposed to be eternal as well. And this is inherence (*samavāya*). Even other universal-friendly realists, such as the Bhāṭṭa Mīmāṃsaka, give Vaiśeika a lot of grief over this peculiar theory of the exemplification. The Bhāṭṭas themselves take the relation between a universal and its own exemplifier to be identity-in-difference. The Buddhist logician finds both inherence and identity-in-difference equally unpalatable.

Though one cannot experience Vaiśeika universals by themselves, they are ontologically independent of the

particular instances. Even when all cows are destroyed in the world, cow-ness will still be around, for otherwise the possibility of a fresh cow coming to be remains inexplicable.

REAL UNIVERSALS AND TITULAR PROPERTIES: ON BEING A COOK

Though all universals are common features, not all common features corresponding to multiply applicable descriptions are, strictly speaking, universals. Being a Brahman (a member of the highest priestly intellectual class) is taken to be a natural kind by Nyāya-Vaiśeika in the face of vehement opposition by anticaste Buddhists and Jainas. But being a cook is the standard example of a common feature that is not a real universal. The Nyāya-Vaiśeika philosophers suggest six tests that an alleged (semantically suggested) property must pass to count as a genuine universal. These tests or hurdles are called universal-blockers:

(1) If a property has only a single exemplifier, then it is not a universal. "Being the Statue of Liberty" is not a universal, neither is time-hood, because there is no more than one Statue of Liberty, one time.

(2) If two properties have exactly the same extension, for example, the property of being a Homo sapiens and the property humanity, they cannot be two distinct universals.

(3) The domains of two universals can be either completely disjoint or one of them completely included in the other. They cannot be partially intersecting and partially excluding each other. Thus, being material and having a limited size cannot both be universals in Vaiśeika ontology, because while lots of things have both the properties, open space is supposed to be material yet not limited in size, while the internal sense-organ is supposed to be limited in size but immaterial. Whether crosscutting disqualifies both the properties or only one of them, and whether the neat ontological hierarchy that is presupposed by this universal-blocker is integral to a realist metaphysics have been the subject of much contemporary debate (see Shastri 1964, Mukhopadhyaya 1984).

(4) A regress-generating property is not a universal. Universal-hood is not a universal, although all universals seem to have that property in common. Because then one could multiply levels of universals endlessly. Universals do not have further universals in them.

(5) When the nature of a characteristic is to merely distinguish its bearer, for example, one earth-atom, from another particular of that kind, such ultimate individuators should not be brought under a general category of individuator-hood, for that militates against their necessarily unique nature. Failing this test, the alleged generality individuator-ness (*visesatva*) fails to qualify as a universal within Vaiśeika atomism.

(6) The feature must bear inherence and no other relation to its bearer. Inherence-hood is not a universal because, had it been one, it would have to be related by inherence to inherence, which would be absurd. An absence cannot be a universal. Nor could the negativity common to all absences be a universal. Even though every rabbit is hornless, neither the absence of horn itself nor the absence-ness of the absence resides in rabbits or absences by inherence. Besides these, compound properties such as being a sturdy black cow or being either a cow or a buffalo are ruled out because universals are supposed to be simple.

What happens to the properties that, thus, get disqualified by a universal-blocker? They are thrown into the mixed pile of titular, surplus, or imposed properties (*upādhi*). They could still be of much theoretical and practical use. Not only nonnatural generalities like being a New Yorker, but even is-ness, knowability, and positive presence (shared by items of all the six categories—substance, quality, motion, universal, inherence, and final individuator—but not found in absences) are merely titular properties. Knowability and existence (is-ness) are (intensionally) distinct properties, in spite of being equiextensive, because they are not universals.

HOW ARE UNIVERSALS KNOWN?

One needs philosophical reasoning to grasp such deep universals as substance-hood, because many instances of substance-hood, such as time, atoms, other people's souls, are not objects of perception. If the instances are perceptible, the universals must be directly perceptible as well. One sees flower-ness in a flower, just as one sees its hue and smells its fragrance. According to Nyāya epistemology, to see Black Beauty as a horse one must first see its horse-ness (which is a perceived universal, though it is not perceived to be a universal).

But many strong arguments could be given against the perceptibility of universals (NM, ch VII). The following are a couple of examples:

If properties were perceived, one would perceive them even at the time of encountering the first exemplifier, but one does not. Hence properties are abstracted, not seen. Both the premises of this argument, of course, could be questioned. For the empirical knowledge of a common property to dawn gradually, a recognition must take place in the second, third, and subsequent sightings of the instances. To be faithful to the form of that recognition, "I have seen this sort of animal before," is to admit that even in the first instance that sortal property was seen.

Here is another antiperception argument. If properties were objects of perception, they would be causes of perception, but they are not. Therefore, they are not perceived. Again, both the premises are rejected by the Nyāya realists. Pot-ness need not itself reflect light back into the retina for it to be causally relevant to the visual perception of pot-ness. As long as the pot in which it inheres is in contact with the seeing eyes, it has a causally operative connection with the appropriate sense-organ. If, of course, perception is defined as prelinguistic and nonconceptual (as some Buddhists have done) and universals are taken to be word-generated concepts, then to use that definition as an argument for imperceptibility of universals would be crudely question-begging.

With Fregean sensibilities one could propose another quick argument against the perceptibility of universals. Universals are not objects but functions. Therefore, they are not objects of perception. Still, there is a clear shift in the meaning of "object" between the premise and conclusion of this argument. There is a basic (rationalist?) resistance even among realists in the West to admit sense-perception of universals, because universals are supposed to belong to the intelligible realm. In *The Problems of Philosophy* Bertrand Russell claims that one has direct acquaintance with universals, but that acquaintance is not meant to be sensory. It is only David M. Armstrong, whose view about universals comes close to Nyāya-Vaiśeika realism, who seems to have warmed up to the idea of perceiving universals.

ATTACKS FROM THE BUDDHIST NOMINALIST

Vaiśeika's first argument for the existence of universals depends on the generalization, "In every case, the sense of commonness or similarity felt by word-users must be spawned by an objective universal." Surely, this generalization is riddled with counterexamples. One has just seen earlier how people feel a sense of similarity across many cooks, yet the Nyāya-Vaiśeika realists refuse to admit cook-ness as a universal. There is no good reason to posit these weird entities, and every reason to eliminate them. So claimed the Sautrāntika-Yogacara Buddhists, "It does not come there (from another place), it was not there already, nor is it produced afresh, and it has no parts, and even when it is elsewhere it does not leave the previous locus. Amazing indeed is this volley of follies!" (PV 1.152–153).

With this oft-quoted remark Dharmakirti (1994) summarizes his battery of objections against the Nyāya-Vaiśeika theory of universals. How can a universal remain the same while existing in distinct things and places? Does it scatter itself into parts or does it live in its entirety in each instance? When the locus moves, does it move? If cow-ness is everywhere, why is it absent in a horse? If it is only where its instances are now, then how does it travel to a new place when a cow is born there? It does not pervade the place where an individual is located, for then the place itself would be its instance, yet how can it manage to inhere in the individual that occupies that place? If the particular instance is needed as a revealer of the ubiquitous universal, how come one cannot perceive the cow—its revealer—independently of noticing the universal cow-ness? A lamp reveals the preexistent pot in a room, but one does not need to see the pot first before one notices the lamp (PV 156).

Most of these difficulties, the realists retorted, suffer from a category-mistake. They assume that a universal is just another kind of super-particular. But a universal is not a spatiotemporal thing, and that is why multiple-location without divisibility is not a problem for it. In spite of such robust responses Buddhist antirealism about universals became more trenchant in the second millennium until such caustic were remarks directed at the Vaiśeika realists, "One can clearly see five fingers in one's own hand. One who commits oneself to a sixth general entity finger-hood, side by side with the five fingers, might as well postulate a horn on top of one's head."

APOHA SEMANTICS: THE BUDDHIST EXCLUSIONIST ACCOUNT OF CONCEPT-FORMATION

Buddhist logicians have an error-theory about universals and permanent substances. There are nothing but momentary quality-particulars in the world. But the human mind, afflicted by recurrence-wishes and language-generated conventional myths, has a tendency to cluster some of them together first in the fictional form of enduring substantial things and then further classify these "things" into types. This illusion of generality, of

course, has some pragmatic value, because, except in contemplative experience, most of one's working cognitions of the world take the form of predictive or explanatory inferences on the basis of these apparently general features and their mutual connections.

When a particular cow (which is a fictional cow-shape superimposed on certain packets of quality-tokens) is seen to be other than all other animals, the original indeterminate (concept-free) perceptual content somehow causally triggers off this difference-obliterating tendency. The particular cow-image is made to "fit" this linguistic and imaginative exclusion from the complementary class of horses, rabbits, pillars, and such things. The specificity of the particular cow—its numerical detailed differences from other cows—is ignored; instead, this mere exclusion from noncows is foisted on to the perceptual content as a predicate. This exclusion masquerades as the universal cow-ness. To take Dharmakirti's (1994) example, the universal antipyretic-ness is a useful figment of imagination. In the external world there is no single shared intrinsic property of different medicinal plants all of which work as fever-reducers, except that they are other than those things that fail to relieve fever. Antipyretic-ness is an erroneous reification of this mere exclusion (*apoha*). This, in a nutshell, is the apoha nominalism of the Yogacara Buddhist logicians.

MILDER NOMINALISMS: RESEMBLANCE THEORIES

In the middle of this great battle between the realists and nominalists, the Jaina syncretists step in with the reconciliatory message that every object of knowledge has an alternatively more-than-one (*anekānta*) nature—particularity and generality are just two of them. One cannot doubt that things do objectively resemble each other. These resemblances are real relations. But both the things and their mutual resemblances are particulars. Nothing has the burden of being repeatable.

The Jainas reject the Buddhist version of nominalism, more or less on the same grounds as Kumārila Bhāṭṭa, the great Mīmāṃsaka, rejected it. Positive predicates, Kumārila had objected, cannot all be given a negative meaning. Since these exclusions are nonentities invented by erroneous imagination, to say that all one's words mean them is to turn all words into empty terms. Indeed, since all exclusions are equally hollow in content, distinguishing one from another would be like trying to distinguish one imaginary nonexistent from another. Only those denials make sense that have something positive to deny. Since all descriptions capture only negations, this theory, ironically, strips one's negations of all meaning, since there is nothing left to deny.

Jaina thinkers reject the exclusionism of the Buddhist but use the Buddhist criticisms to reject the Vaiśeika realism. In its place they propose this resemblance theory. Prabhācandra anticipates the Russellian objection that at least all these resemblance-relations would ultimately need a shared resemblance-universal. His answer to it is that, just as a Vaiśeika final individuator (*víseṣa*) does not need another distinguisher, one resemblance does not need a higher level resemblance or universal to explain why all those resemblances are similar. While accounting for the similarity between ground-level particulars, they also account for their own similarity to each other. Versions of this theory were adopted by followers of Rāmānuja (qualified monist Vedāta) as well as by Madhva (dualist Vedātin) logicians. Vyāsatīrtha of the latter school clarified how a single resemblance can reside, as it were, with one leg in the resembler and with another leg simultaneously in many other similar particulars.

The category of resemblance admitted by these philosophers is different from the resemblance admitted by Prābhākara Mīmāṃsakas, for the latter were realists about universals, while the Jainas and the Madhvas rejected, as logically redundant, both universals and inherence. The only difference between Prābhākara and Vaiśeika as regards universals centers on their conceptions of inherence.

CONTRASTS WITH WESTERN METAPHYSICS OF FORMS AND PROPERTIES

It should be clear by now that there is no core theory of universals shared by all the Indian philosophers. But one can discern five broad features that distinguish the Indian theories of universals from their Western counterparts:

(1) Even the strongest realist position of the Nyāya-Vaiśeika never took the form of the realism of Plato's theory of ideas. Indian realists about universals were equally realists about the perceptible particulars of the external world. Earthly particulars were never thought to be less real copies of thinkable universals, even by those who believed in universals.

(2) Even if one concedes that the Nyāya universals were closer to Aristotle's universal properties, which are immanent in the worldly particulars, Aristotle could never agree that universals are themselves directly perceived, which is the standard Nyāya position.

(3) The peculiar form that nominalism took in the Indian Buddhist theory of word-meanings as exclusions does not have any parallel in the West. One finds an interestingly different counterpart of the Jaina and Madhva theories of resemblance in Nelson Goodman, but exclusion-nominalism remains a unique contribution of Indian Buddhism.

(4) Most Western realist accounts of universals take colors and such qualities, as well as relations such as "being larger than," as paradigm examples of universal properties. In Indian realist thought the distinction between such particular qualities (*guṇa*) and universal properties (*jāti*) has been sacrosanct. It is only recently that the idea of particular qualities is gaining ground in Western analytic metaphysics of tropes. Even relations are not treated as genuine universals by any classical Indian realist.

(5) The controversial and complex theory of inherence as a single concrete connector joining not only universals and their instances but also particular qualities to substances and, most puzzlingly, wholes to their parts is totally foreign to the Western realists.

See also Atomic Theory in Indian Philosophy; Knowledge in Indian Philosophy.

Bibliography

IN SANSKRIT

Udayana. *Kiranavali*. Calcutta, India: Bibliotheca Indica, 1956.

Jayanta Bhāṭṭa. *Nyayamanjari*. Part 2, Chapter 5. Mysore, India: Oriental Research Institute, 1983.

Dharmakirti. *Pramana Varttikam*, edited by Swami Dwarikadas Shastri. Varanasi, India: Bauddhabharati, 1994.

Asoka Pandita. "Samanyadusanadikprasarita." In *Six Buddhist Nyāya Tracts*, edited by Haraprasad Shastri. Calcutta, India: The Asiatic Society, 1989.

Patañjali. *Vyakarana-Mahabhasyam*. Vols. 1–2, edited by Bala Shastri. Delhi, India: Pratibha Prakashan, 2001.

IN ENGLISH

Bhattacharya, Gopinath. "Some Aspects of the 'Universal.'" In *Essays in Analytical Philosophy*. Calcutta, India: Sanskrit Pustak Bhandar, 1989.

Chakrabarti, Kisor. "The Nyaya-Vaisesika Theory of Universals." *Journal of Indian Philosophy* 3 (1975): 363–382.

Dasgupta, Surendranath. *The Mahābhāsya of Patañjali: With Annotations*, edited by Sibajiban Bhattacharyya. New Delhi, India: Indian Council of Philosophical Research, 1991.

Dravid, Raja Ram. *The Problem of Universals in Indian Philosophy*. Delhi, India: Motilal Banarsidass, 1972.

Halbfass, Wilhelm. *On Being and What There Is: Classical Vaiśesika and the History of Indian Ontology*. Albany: SUNY Press, 1992.

Hayes, Richard P. *Dignāga on the Interpretation of Signs*. Dordrecht, Netherlands: D. Reidel, 1988.

Herzberger, Radhika. *Bhartrhari and the Buddhists: An Essay in the Development of Fifth and Sixth Century Indian Thought*. Dordrecht, Netherlands: D. Reidel, 1986.

Mukhopadhyaya, P. K. *Indian Realism*. Calcutta, India: Firma K. L., 1984.

Shastri, Dharmendra Nath. *Critique of Indian Realism*. Agra, India: Agra University, 1964.

Siderits, Mark. "More Things in Heaven and Earth." *Journal of Indian Philosophy* 10 (1982): 187–208.

Arindam Chakrabarti (2005)

UNIVERSALS

See *Universals, A Historical Survey*

UNIVERSALS, A HISTORICAL SURVEY

The word *universal,* used as a noun, has belonged to the vocabulary of English-writing philosophers since the sixteenth century, but the concept of universals, and the problems raised by it, has a far longer history. It goes back through the *universalia* of medieval philosophy to Aristotle's *τὰ καθόλου* and Plato's *ἐίδη* and *ἰδέαι*. Indeed, Plato may be taken to be the father of this perennial topic of philosophy, for it is in his dialogues that we find the first arguments for universals and the first discussion of the difficulties they raise. Plato believed that the existence of universals was required not only ontologically, to explain the nature of the world that as sentient and reflective beings we experience, but also epistemologically, to explain the nature of our experience of it. He proposed a solution to his problem, but he also recognized the objections to his particular solution. Ever since, except for intervals of neglect, philosophers have been worrying about the nature and status of universals. No account has yet been propounded that has come near to receiving universal acceptance; this reflects not merely disagreement on the answers to be offered but also, and perhaps more importantly, disagreement on exactly what the questions are that we are, or should be, trying to answer.

That in some sense or other there are universals, and that in some sense or other they are abstract objects—that is, objects of thought rather than of sense perception—no philosopher would wish to dispute; the difficulties begin when we try to be more precise. They

may be indicated (although not defined) by the abstract nouns that we use when we think about, for example, beauty, justice, courage, and goodness and, again, by the adjectives, verbs, adverbs, and prepositions that we use in talking of individual objects, to refer to their qualities and to the relations between them. In saying of two or more objects that each is a table, or square, or brown, or made of wood, we are saying that there is something common to the objects, which may be shared by many others and in virtue of which the objects may be classified into kinds. Not merely is such classification possible, for scientific and other purposes; it is unavoidable: All experience is of things as belonging to kinds, however vague and inarticulate the classification may be. Whatever we see (to take sight as an example) we see as *a something*—that is, as an object of a certain kind, as having certain qualities, and as standing in certain relations to other objects—and although every individual object is unique, in that it is numerically distinct from all others, its features are general, in that they are (or might be) repeated in other objects. Even if there were only one red object in the world, we would know what it would be like for there to be others, and we would be able to recognize another if we were to meet with it.

Generality is an essential feature of the objects of experience, recognition of generality is an essential feature of experience itself, and reflection of this generality is shown in the vocabulary of any language, all the words of which (with the exception of proper names) are general. Universals are, by tradition, contrasted with particulars, the general contrasted with the numerically unique, and differing theories of universals are differing accounts of what is involved in this generality and in our experience of it. The leading theories of universals—realism, conceptualism, nominalism, and resemblance theories—can best be explained by an examination of the doctrines of the main exponents. In following that sequence we shall be adhering approximately (although not precisely) to the chronological order in which the rival theories developed, and we shall be historically selective, in that we say almost nothing of the periods in the history of philosophy during which the controversies continued (for example, medieval philosophy) but of which a detailed knowledge is not necessary to a general understanding of the issues involved. The aim here is to present the different views that have been held, not to trace the fortunes of each view throughout the history of the subject.

REALISM

Realist and conceptualist theories of universals are, by long tradition, regarded as opposed because according to realism universals are nonmental, or mind-independent, whereas according to conceptualism they are mental, or mind-dependent. For the realist, universals exist in themselves and would exist even if there were no minds to be aware of them; if the world were exactly what it is now, with the one difference that it contained no minds at all, no consciousness of any kind, the existence of universals would be unaffected. They are public somethings with which we are somehow or other acquainted, and a mindless world would lack not universals but only the awareness of them: They would be available for discovery, even if there were nobody to discover them. For the conceptualist, on the other hand, universals are in the mind in a private sense, such that if there were no minds, there could be no universals, in the same way as there could be no thoughts or imagery or memories or dreams. As will be seen, whatever may be said for or against realism, pure conceptualism cannot be a satisfactory theory, for it is essentially incomplete; it says something about our consciousness of universals but nothing at all about any basis for this consciousness. Consequently, philosophers who have been conceptualists either have been so because they have been interested only in the epistemological question, in the conceptual structure of human thought and experience, or have combined their conceptualism with another theory designed to answer the ontological question—that is, the question what there is in the world corresponding to our mental concepts or ideas, what our concepts are concepts of. The antithesis between the two theories of realism and conceptualism is not, therefore, as clear-cut as it has often been presented to be.

The two main versions of realism are those of Plato and Aristotle. Plato's came first, and the difficulties it raised, some raised by Plato himself, others added by Aristotle, were what led Aristotle to devise his own quite different, but still realist, account. Plato and Aristotle were both realists in that they accorded to universals an existence independent of minds; where they differed was on the nature of the existence and the status that they believed universals to possess.

PLATO. Although it is possible to give, in some detail, a statement of what may be called Plato's theory of universals, and to give it full documentary support by quotations from his writings, we would be mistaken to regard it as a final and fully worked out theory. It was a theory toward which Plato can be seen working his way through-

out his philosophical career, not so much by independent arguments as by intertwining strands of thought, all leading in the same general direction. There were a number of facts about the world and our experience of it by which he was impressed and puzzled. His theory evolved as an explanation of them, but he was never satisfied that he had solved his problem. He was his own first critic, and a penetrating one, and to the end of his life he was torn, as is brought out in his dialogue *Parmenides,* between the conviction that his theory was fundamentally correct and the recognition that it posed problems that he found himself unable to solve. It should not be thought, therefore, that he ever produced a final account that he was prepared to rest content with and that needed an Aristotle to find fault with it.

Plato's interest in questions about universals was first aroused by Socrates, by whom he was greatly influenced, whom he introduced as one of the speakers in all his dialogues (with the single exception of *The Laws*), and who in all but the later dialogues appears as the central character actually directing the conversation. Unfortunately, we are presented with difficulties of interpretation, the details of which we shall not enter into here, because our knowledge of Socrates is derived entirely from descriptions given by other writers, one of whom was Plato. Hence arises the problem of deciding which of the doctrines ascribed to "Socrates" in the Platonic dialogues are those of the actual Socrates and which of them are extensions or even entirely new doctrines developed by Plato himself. In general, it is accepted that the "Socrates" of the early dialogues does represent the views, and even more the methods of philosophical inquiry, of Socrates himself but that as time went on Plato more and more used him as the spokesman of Plato's own views, the transitional stage being marked by such dialogues as *Phaedo* and the *Republic.* We may conclude that while Socrates did not explicitly hold a theory of universals (and we have Aristotle's word for it, in *Metaphysics* 1078b, that Socrates did not hold the view Plato put forward), his philosophical questions were such that Plato held they could not be answered except by such a theory; in other words, Plato, in putting a theory of universals into Socrates' mouth, was not attributing it to Socrates as what he had actually expounded but was maintaining it as the logical consequence of Socrates' own arguments: Socrates stopped short of propounding such a theory himself but was logically committed to it.

Socrates' main interest was in the human virtues, and his aim was to secure a satisfactory definition of the virtue under discussion. His questions were all of the form "What is *X?,*" where "*X*" stood for beauty, courage, piety, justice, and so on, in one case (*Meno*) even virtue itself. The answers that he received he rejected because they were too narrow or too wide, but more commonly because instead of giving the essential definition of the virtue they gave instances of it or mentioned kinds of it. Thus, it was no answer to the question "What is piety?" to reply that a man is acting piously if he prosecutes a murderer; again, it was no answer to the question "What is virtue?" to reply that the virtue of a man consists in managing a city's affairs capably, that a woman's virtue consists in managing her domestic affairs capably, that there are different virtues for an old man and a young man, for a free man and a slave, and so on. Granted that there are many virtues, what is wanted is the one and the same form that they all have and by which they are virtues. The search, then, is for the single and essential form common to all things of the same kind, by virtue of which they are things of the same kind.

The "things" about which Socrates in fact asked his questions were limited because his philosophical interest was limited, but even he did not confine himself to human conduct. He acknowledged, for instance, that health or size or strength must be the same in all its instances, with the consequence that we answer the question "What is health?" only when we have given the essence of health—that is, what is common and peculiar to all instances of health. Plato took this further and maintained (although not without hesitation) that there must be an essence common to all things of a given kind, whatever that kind was. It would apply not only to abstract virtues, such as justice and courage, but also to natural objects, such as trees, and to artifacts, such as tables. An object would not be a table unless it had the same essence (of tablehood) as all other tables; despite the different shapes and sizes that individual tables may possess, there must be a single form or essence, common to them all, which constitutes their being tables and distinguishes them from other objects, such as chairs or beds. Plato summarized his position in the statement "We are in the habit of postulating one single form for each class of particulars to which we give the same name" (*Republic* 596A). And he held it to be true not only of objects designated by nouns (such as "bed" and "table") but also of attributes or qualities indicated by predicates (such as "beautiful" and "greater than"). As there must be a form or essence of bedhood somehow common to all beds, so there must be a form or essence of beauty (or the beautiful) common to all things that are beautiful.

So far Plato had done nothing more than take over the Socratic contrast between the single general, essential form common to a class of particulars and the particulars themselves and extend it more widely than Socrates had done: He found the same contrast not only in the realms of ethics, aesthetics, and mathematics but also in the everyday world of sense experience. But he went on to ask the questions that Socrates had never asked, namely what are we to say about the relationship between the universal form and its particular manifestations, and what are we to say about the nature and existence of the universal itself? His answer was to develop the theory known as the theory of Forms, according to which each universal is a single substance or Form, existing timelessly and independently of any of its particular manifestations and apprehended not by sense but by intellect. His arguments can be distinguished, although not entirely separated, into two general kinds, metaphysical or ontological and epistemological. If knowledge is to be possible at all (and Plato did not doubt either that it was possible or that in certain spheres it was actual), it must be of what is stable and unchanging.

However, the familiar world of ordinary experience does not meet this requirement, for the one constant and striking feature of all objects (and their qualities) in this world is that they are subject to change and decay: Both natural objects and artifacts come, or are brought, into being, undergo changes throughout their existence, and sooner or later die or disintegrate and disappear. This is the Heraclitean doctrine of flux, which Plato accepted and which he believed required as its counterpart a nonsensible realm of unchanging stability, without which there could be no knowledge. What can be known must be real, unitary, and unchanging: These are the Forms. Particulars are only semireal, real to the extent that in some way or other, or to some degree or other, they manifest the Forms, unreal to the extent that being material, they lack the perfection of pure Forms and are subject to the laws of material change and decay. Thus, Forms are required, to confer on particulars such reality as they do have, to constitute their being what they are and of what kinds they are. A bed is a bed rather than a table because it somehow manifests the Form Bed. A Form is required not only to explain a particular object's being what it is but also to cause its being what it is; the doctrine is thus not merely a logical but a metaphysical doctrine. Plato emphasized this in the analogy of the sun (*Republic* VI), where he compared the chief Form of all, the Form of the Good, with the sun, which as the light-giving and life-giving agent in the physical world is the prime material cause of natural life as well as of our awareness, through our senses, of the material world.

Another consideration that led Plato to suppose the Forms as transcendent substances was the presence of what he thought to be contradictions in the material world: What is real cannot contain contradictions; therefore the material world cannot be more than an appearance of reality. That a single object should be both beautiful (in one respect) and ugly (in another), or large (in comparison with a second object) and at the same time small (in comparison with a third), was enough, in his view, to show that the Forms were more than immanent. Therefore, not only must there be Forms in order to cause particulars to be what they are, but the Forms must be separate from the particulars because they must be free of the imperfection and defectiveness with which particulars are inevitably infected. The Forms are thus not only independent substances but perfect and ideal patterns, which particulars must fall short of.

This comes out especially in the consideration of mathematical (primarily geometrical) and value concepts, namely those of ethics and aesthetics. For a line to be straight or a figure to be circular, there must be the Forms of Straightness and Circularity. But it is well known that no actual line is ever perfectly straight and no figure is ever perfectly circular; however carefully and precisely drawn, it possesses some curves or kinks that more minute scrutiny could disclose. And what we are thinking about when we study or discuss a geometrical theorem is not the diagram of the circle drawn, freehand or mechanically, on the blackboard but the circle represented by the diagram. We thus have both the diagram of the circle, adequate as a diagram but imperfect as a circle, and the perfect Form of Circularity of which it is a diagram. While this gives rise to the question, which cannot be pursued here, whether Plato distinguished between the Form of Circularity (of which there could not be more than one) and a Perfect Circle (of which, if there could be one, there could be more than one—as required by, for example, a theorem involving two intersecting circles), there is no doubt that he did think a Form not only was the perfect pattern, of which a particular was an imperfect manifestation, but also was what the particular would be if, *per impossibile,* it could be perfect. Thus, to take an aesthetic example, Beauty (or the Beautiful) not only is the pattern that beautiful particulars inadequately manifest but also is itself perfectly beautiful; it is a substance possessing in perfection the essence that its derivative particulars possess only partially or in some degree. As Plato came to realize later (*Parmenides* 131ff.), and as

Aristotle repeated, if a Form stands to its particulars as "one over many," and if the Form is an ideal pattern of which the particulars are imperfect copies, then an infinite regress argument (known as the third-man argument) is generated: For the Form to be predicable of itself as well as of its particulars, it must share a character with them; but then there will be a Form of this character; this second Form will be predicable of itself, requiring a third Form of it, a fourth, and so on ad infinitum.

As was indicated above by the geometrical example, Plato believed that his theory of Forms accounted for the possibility of knowledge of universal truths, which was the only kind of knowledge strictly meriting the name. When, by working out or following the proof, we learn that a square constructed on the diagonal of a given square has an area equal to double the area of the given square, we have learned a truth that is necessary and universal. It is not something that happens, as a matter of fact, to be true of the squares in our diagram but might turn out not to be true of some other squares; that is, it is not an empirical generalization that subsequent experience might show to be false as a generalization. We have a piece of a priori knowledge, which no possible experience could affect, namely that if a square has a given area, and if a second square has its sides equal in length to a diagonal of the original square, then the area of the second square must be double the area of the first. Our knowledge is not knowledge of our diagram squares, or any others that we care to draw, for, as we have seen, they are not in fact squares. But it is knowledge, and the only thing, therefore, that it can be knowledge of is the Form Square (or the Square).

What defeated Plato in any attempt to give a complete account of his theory was the problem of describing the relation of Forms to particulars. In different places he spoke of the Forms "being in" their particulars, of particulars "participating in" their forms, and of particulars "copying" their forms. Literal interpretation of any of these phrases gives rise to logical difficulties, and to take them metaphorically is to leave the statement of the theory imprecise and the problem unanswered. In Plato's final writings (*Epistle* VII) on the subject there are signs that he was inclined to think that the fault lay with the inadequacy of language to describe what he wanted to describe, but the trouble is deeper than mere paucity of vocabulary. We can form some kind of a picture of his two worlds if we think of the world of Forms as actually existing somewhere, populated by objects like the Standard Meter and the Standard Pound, and we can then think of actual particulars as being imperfect copies of the originals. But that picture, taken literally, is false, because Plato's Forms do not exist in a place or at a time. The mystery of their "existence" becomes impenetrable when we are asked to use the word *exist* in a way that we are incapable of conceiving. In his theory of Forms, with the Forms not immanent but transcendent, the problem of their relation to particulars becomes not almost impossibly difficult to solve but in principle insoluble.

ARISTOTLE. Aristotle, Plato's pupil and successor, is often regarded as the careful scientific-minded thinker, anxious to restrain philosophy within the range of the observable and to avoid the imaginative speculations of Plato. While this picture is in general correct and in particular fits Aristotle's criticisms of Plato's theory of Forms regarded as universals, his own theory of a Form as the object of a definition that describes a thing's essential nature becomes in the end as obscure as Plato's. His criticism that Plato's theory does nothing to provide a scientific explanation of the nature of things applies equally forcibly to his own theory of essences, and natural science, as we know it, began to progress only when, many centuries later, it liberated itself from this aspect of Aristotelianism.

But Aristotle's theory of universals, which is nowhere fully elaborated and has to be pieced together from different passages, is important, both because it offered an alternative to Plato's and because it is more obviously attractive to common sense. His objections to Plato are numerous and detailed but are not all of equal weight. Basically, apart from the infinite regress argument, which he took over from Plato, they come to two: First, that Plato, by making the Forms perfect, separate substances, introduced an unnecessary and unhelpful duplication, and second, that Plato confused the categories of substance and property. Nothing is accounted for by making the Forms perfect patterns of particulars. To attempt to explain the nature of one set of entities by postulating a second and better set does not solve a problem but merely repeats it at a different level: Whatever the question was that needed to be answered about particulars, it will need to be answered again about the Forms; mere multiplication answers nothing. Second, Plato was guilty of a logical mistake in treating a Form both as an individual substance (which the "separation" thesis requires) and as a property (which it would have to be to be a universal). Substances are individuals and have properties, but they cannot be properties, yet Plato's theory treats them as both.

For Aristotle the only true substances were single individual objects, such as Socrates or this table. (It is true that Aristotle introduced a difficulty by treating genus and species also as substances, for they are what it is the aim of science to know, but they are secondary substances, and the knowledge we may gain of them is knowledge about primary substances—that is, the individual objects met with in experience.) Universals, therefore, are not substances existing independently of particulars. They exist only as common elements in particulars: The universal *X* is whatever is common to, or shared by, all *x*'s; it is what is predicated of the individual. Individual objects are to be classified into kinds according as they share the same property, and the kinds are to be subdivided into genus and species by the differences between more determinate properties. Thus, all colored objects belong to the genus "color" because they all alike have the property of being colored, whereas red objects and green objects belong to different species of the genus, because the first have the property of being colored red and the second have the property of being colored green. One of the primary tasks of natural science is to divide and classify natural objects by genus and species into the real kinds to which, by nature, they belong.

Aristotle's theory is more economical than Plato's, requiring only one world of being instead of two, the contrast between the two theories being indicated by the labels that they later acquired in medieval scholastic philosophy: Plato's was a theory of *universalia ante rem* (universals independent of particulars), and Aristotle's of *universalia in rebus* (universals in things). And with the possible exception of ideal concepts, such as those of geometry, which Plato had argued had no actual instances, Aristotle's account seems better to fit a fact, or what we take to be a fact, of human experience, namely that a particular really is an instance of its universal. Not only should we say that we get our idea of red, for example, from seeing red objects, such as fire engines or ripe tomatoes, but we should also say (except for philosophical theories of perception) that the object really was red, not that (as with Plato) the tomato tried unsuccessfully to be red but that (with Aristotle) it actually was red. The properties that an object has, and that together constitute its nature, its being an object of that kind, whatever that kind may be (for example, whether it is a horse or a table), are really in the object, in some sense of "in." If objects do not and cannot possess any of the characteristics that according to experience and the scrutiny of observation they appear to have, then scientific knowledge becomes either altogether impossible or unrelated to the natural world. Aristotle's view avoids the Platonic paradox that nothing in the observable world can ever be what it seems to be.

The contrast between the two views comes out again in their accounts of how we apprehend universals. They are agreed both that awareness of universals is implicit in ordinary sense experience (for it is this awareness that conditions our experience as being what it is) and that we are aware of universals not by sense itself but by intellect. Plato could not say that we become aware of them by abstraction from particular instances, because they have separate existence and never are more than defectively instantiated: If our concept of *X* were only what we could abstract from imperfect instances, we never could apprehend *X* itself. Therefore there must be some other mode of apprehension, which Plato called *ἀνάμνησις* (usually translated as "recollection" but less misleadingly interpreted in this context as "recovery"). The human soul has prenatal knowledge of universals and of their mutual relations, and postnatal experience of the ordinary world serves, or may serve, to revive this knowledge in suitable circumstances. Thus, experience does not directly provide us with new apprehensions (of universals) or with new knowledge of necessary truths (connections between universals) but acts as a stimulus to remind us of what we already know but have hitherto in this life forgotten. Plato's argument here, if it is to be regarded as an argument, is a transcendental one (in Immanuel Kant's sense of the word): Our knowledge is a priori, that is, of such a kind that we cannot get it from experience, although we do get it in experience; therefore it must be innate, that is, knowledge of what we originally knew prior to any experience. As a transcendental argument it could be effective only if it could be shown that there was no other possible way of accounting for our apprehension of universals and our knowledge of universal truths. And Aristotle thought that there was another, less fanciful and less speculative way, derived from actual experiences and memories of previous experiences.

Apprehension of a universal, or formation of a concept, is not a sudden once-and-for-all business, given in a single experience, but a gradual process. Sense perception gives rise to memory, and memory conditions subsequent perceptions, so that they are not merely perceptions but recognitions of what is in some degree or other familiar from previous perceptions. Awareness of characteristics thus becomes clearer and more explicit with the growth and variety of experience. By a process of induction, namely intuitive induction, the first primitive awareness of a universal (necessary to any perception) becomes stabilized in the mind, leading ultimately to a clear and

articulate concept of it. Thus, for Aristotle, as for Plato, grasp of universals is by the intellect, but it is by the intellect gradually working on what it is at first dimly and indeterminately conscious of in the data of sense perception. A simple example from arithmetic will illustrate his point. As children we learn to count. We get the idea of 2 from being faced with pairs of objects, and we learn that 2 + 2 = 4 from coming to "see," for instance, that two apples plus two other apples are equal in number to four other apples. But we also come, sooner or later, to "see" that the number 2 characterizes any pair of objects, and that 2 + 2 = 4 is a necessary truth, applicable to any two pairs compared with a quartet. We have the power, which becomes actualized in experience, of intuiting clearly the universal in the particular and of intuiting the necessary in the matter of fact; this, for Aristotle, is the beginning of scientific knowledge.

AUGUSTINE. Medieval philosophy was not primarily interested in questions about the nature of human knowledge. But its concern with metaphysics, especially in those aspects that carried theological implications, led to a continuation of the dispute between the two versions of realism and later to a nominalist rejection of both. Platonic realism was championed by St. Augustine, for whom divine illumination performed much the same function as Plato's Form of the Good, rendering intelligible by its light the necessity of eternal truths that the human intellect could grasp. Man is above the beasts, not only because he can acquire, by the mind alone, knowledge of eternal truths, but also because even in sensation he judges of material objects by incorporeal standards: In judging a physical object to be beautiful he implies the objective existence of Beauty, both as a universal and as a standard. Again, the intelligible structure of the temporal world, which the reason of man (but not the senses of the beasts) can grasp, is itself nontemporal; for example, the concepts and truths of mathematics, although empirically applicable, are timeless necessities. Ideas as objective essences are exemplars contained "in the divine intelligence." Thus, Plato's theory of Forms enters theology, and the question arises whether Augustine in his theory of ideas supposed that men were in direct contact with the mind of God. It is fairly clear that he did not but much less clear how he could avoid it.

THOMAS AQUINAS. The leading exponent of Aristotelian realism was Thomas Aquinas, who, although professing the greatest reverence for Augustine, departed widely from Augustine's views. Thomas's metaphysics is, like Aristotle's, teleological, maintaining that the nature of things and events is to be explained in terms of the ends that they serve, and he extended Aristotle's contrasts between potentiality and act, between form and matter, and between essence and existence. Essences are universals, which have no being apart from existence but which are intelligible without the supposition of existence. The existence of things does not follow from their essence—otherwise existence could not be, as it clearly is, contingent. Universals are apprehended directly by the mind, but only in the material things the nature of which they comprise; they are not to be found in themselves, although by the processes of abstraction and comparison the mind can approximate to thinking of them in themselves. The chief follower in the Thomist tradition was John Duns Scotus, who nevertheless rejected much in Thomas, such as the distinction between essence and existence, and followed Avicenna in differentiating between the "thisness" of an individual object (which distinguishes it from other objects of the same kind) and the nature of an individual object (which distinguishes it from objects of other kinds).

CRITICISM OF REALISM. Although each of the two versions of realism received vigorous support in the long disputes of medieval philosophy, and although Augustinianism for a time prevailed, Aristotle's version has had the longer-lasting influence, especially on philosophers brought up in the British tradition of empiricism. That things do have common characters and that the characters are objectively real seems hardly deniable, and this is part of what Aristotle's theory asserts. But although it is more hardheaded than Plato's, it does raise its own difficulties, two of which may be mentioned. First, how much does it in fact explain of what it purports to explain? We do not account for two tables' being tables better by saying that they have a single characteristic (or set of characteristics) in common than by saying that they are both imitations of a single Form. And if what is to be accounted for is rather our ground for saying that they are tables, which is a question not about their being tables but about our justification for believing or claiming to know that they are, then admittedly we are perceptually aware of the characteristics of each, and of their similarity. But is saying that some (or all) of the characteristics of the one table are like (even exactly like) the characteristics of the other what the Aristotelian means to do when he maintains that there is a universal common to them (and any other tables)? This may be doubted, for the Aristotelian asserts that a single universal is present in each of the objects, or that each is an instance of it, all the objects of a given kind sharing in the universal of that kind.

But this is metaphorical talk, and to explain by metaphor is not to explain at all. As a descriptive statement "These two tables are the same shape" is unobjectionable; as an explanatory statement it is less obviously illuminating. Second, Aristotle's supposition that objects belong to real kinds, which are there for us to discover, ignores the fact that distinctions between kinds or classes are not found but made by us, as was later emphasized by John Locke. This difficulty is not fatal to the Aristotelian theory, which could accommodate it by emphasizing different levels of determinacy in a universal or class characteristic, but it leads to the question, pursued by Ludwig Wittgenstein in the twentieth century, whether it is necessary that any single characteristic at all be common to all members of a single class. If it is not necessary, our recognition of objects as belonging to a certain class does not have to depend on the apprehension of a universal shared by all its members, for it may be that nothing, even in the metaphorical sense, is shared. Aristotle's theory, which prima facie has the merits of being simple and realistic, is perhaps both too simple and not realistic enough.

CONCEPTUALISM

As has already been indicated, conceptualism should not be regarded strictly as a rival theory to realism, even if some of its exponents have mistakenly so regarded it. Starting from an extreme Aristotelian position, that everything which exists is particular, conceptualism concentrates on the fact that generality is an essential feature of both experience and language, and it seeks to answer the question how mental concepts are formed, how they can be general if the data of experience from which they are formed are particular, and how words are general in their significance. Nominalism carries the process further by maintaining that only words are general. Both theories, even if they answered their own question satisfactorily, would have to face the question what basis in reality there is for the generalization inherent in experience, thought, and language. Some versions ignore this question altogether; others answer it in terms of the similarities and differences to be found between particulars. The essential difference between the theories of conceptualism and nominalism is that while both profess to answer a question about language—how words are general, or how words have meaning—nominalism does it more economically, without interposing concepts between words and what words stand for. The conceptualist says that a word is general or meaningful because in the mind there is a corresponding general concept; he then has to explain what a general concept is. The nominalist thinks that the meaningfulness of a word can be accounted for without postulating a separate mental entity called a concept.

Conceptualism is primarily associated with the three classical British empiricists, Locke, George Berkeley, and David Hume, all of whom propounded views about what, in the terminology of the time, were called general ideas. They were all empiricists in that they agreed that all ideas, or the elements that ideas are composed of, come from, and can come only from, experience: The mind can work on what is given to it by sense experience but can neither have ideas prior to any experience (a denial of the doctrine of innate ideas and, by implication, of Plato's suggestion of prenatal acquaintance with the Forms) nor create ideas de novo. Thus, the essence of empiricism is the Epicurean doctrine, given fresh impetus in the seventeenth century by Pierre Gassendi, that *nihil est in intellectu nisi prius fuerit in sensu* ("Nothing is in the mind which is not first in the senses").

LOCKE. John Locke was first in the field, with his *Essay on Human Understanding* (1690), a long, rambling, and discursive work composed and revised over many years. Unfortunately, the passages in the *Essay* in which he discussed general ideas, or, as he more commonly and perhaps misleadingly called them, "abstract ideas," are neither so clearly thought out and expressed nor perhaps even so consistent as to save him from varying interpretations. The initial difficulty concerns the word *idea* itself, which is the key word of his philosophy, but which he neither defined nor used so as to escape ambiguity. Sometimes when he spoke of ideas in the mind he appears to have meant mental images such as occur in remembering, imagining, and dreaming; in this view thinking is done in images, which are particular in their occurrence and existence but somehow become general in their use. At other times he meant, or at least has been taken to have meant, that abstract ideas are mental entities different from images. At still other times he showed signs of using the word *idea* not as the name for any mental occurrence at all but as shorthand for the meaning of a word. Thus, the idea of red would be not an image of something red but what we mean by the word *red* or what we think an object to be when we think it is red; to have the idea of red is to be able to use the word *red* correctly and to be able to discriminate correctly between those objects that are red and those that are not. Attention here will be paid mainly to the first view, of ideas as images, for it is a conceptualist view; so would be the second, that general ideas are mental occurrences different from images, but this

appears to be a view that Berkeley fathered on Locke rather than one Locke actually held.

According to Locke we form general ideas by a process of abstraction from particular ideas. In two different places he gave what appear to be two different accounts of abstraction. In the *Essay on Human Understanding* (Book III, Ch. 3) he said that a general idea—for example, of man—is formed by leaving out of the particular ideas of various individual men all features that are not common to them all and retaining only what is common to them all. The general idea of animal is arrived at by still further leaving out, "retaining only a body, with life, sense and spontaneous motion, comprehended under the name 'animal.'" If this passage were taken in isolation, regardless of what else Locke said on the matter, there would be something to be said for the Berkeleian interpretation. For Locke appears to have been saying that we start with a number of particular images, each, for example, of a different individual man of our acquaintance, and end with something that is still an image but is now a ghostly general image, characterized not by any of the features that are peculiar to any of the individual men but only by all those that all men share. It was not difficult for Berkeley to ridicule as logically absurd the suggestion of a mental image, all the features of which are (as, in this view, they would be) determinables. In his polemic Berkeley did not consider the possibility that Locke might have been getting at something different, namely that mental images may be indeterminate, so that the logical laws of contradiction and excluded middle do not apply to them; for instance, a mental image of a cloudless night sky is an image of a number of stars but of no precise number.

Locke's other account of abstraction, however, which occurs earlier in the *Essay,* seems to be the one he seriously intended. For he came back to it again later in the work than the passage just discussed, and it may even be that in that passage he thought he was still giving the same view as before. In Book II (Ch. 11, Sec. 9) he thus described abstraction:

> The mind makes the particular ideas, received from particular objects, to become general; which is done by considering them as they are in the mind such appearances—separate from all other existences, and the circumstances of real existence.… This is called *abstraction,* whereby ideas taken from particular things become general representatives of all of the same kind.… Thus, the same colour being observed today in chalk or snow, which the mind yesterday received from milk, it considers that appearance alone, makes it a representative of all of that kind; and having given it the name "whiteness," it by that sound signifies the same quality wheresoever to be imagined or met with; and thus universals, whether ideas or terms are made.

It should be noted, from the last phrase, that Locke was using the word *universal* in the subjective conceptualist way, to indicate a concept or idea, not that of which it is the idea. If there is a problem of objective universals raised by a number of things being "all of the same kind" or "the same quality wheresoever met with," Locke showed no sign here of being troubled by it. He was interested only in the question how we form the general ideas that undoubtedly we do have (for without them thought, language, and even experience as we know it would be impossible) when every idea or image that occurs in our consciousness is a particular occurrent. I cannot form an image of whiteness or of white, only of a white something, such as a piece of white chalk or a white snowball. The general idea is not a different idea from the particular idea, somehow extracted from it. It is the particular idea regarded in a special way. First, the mind attends only to a certain aspect of the idea and ignores the rest; second, it treats the idea in that aspect as representative of everything that is similar in that aspect. If *abstraction* is perhaps not the most happily chosen term here, at least Locke's meaning is clear, and he repeated it several times later. A general idea is not one that has a different kind of existence from particulars; all ideas, he said, are particular in their existence. A general idea is a particular idea, used in respect to some aspect as representative of a class, namely the class of things determined by the aspect attended to; in thinking or talking about whiteness the ideas of the piece of white chalk, the snowball, and the glass of milk will all do equally well.

Just how far Locke regarded himself as committed to ideas as images and how far he would have regarded his account as being philosophical rather than psychological (if he could have been induced to accept the distinction) is hard to say. But it is fairly clear that his account is not philosophically satisfactory. He showed himself to be well aware that the real problem is one concerning the applicability and use of general words or terms. But as must have been obvious to him, significant use of words in speech or writing is not in fact paralleled by a corresponding string of introspectable images. Therefore, at best, his claim that a general word is meaningful because it stands for a general idea would have to involve "stand for"

in a dispositional sense; that is, a word is meaningful if a corresponding idea can be found for it. Even then he would be open to the nominalist criticism that nothing is explained simply by duplicating a general word with a general idea. Furthermore, he stressed that almost all thought is verbal: The use of nonverbal imagery in thinking is restricted to a very narrow and primitive level. And, in fact, in the latter part of the *Essay* he showed signs of interpreting ideas not as pictures corresponding to words but as meanings of words, particularly when he was discussing modes—that is, concepts not necessarily used with existential reference. To have an idea, for example, of murder or of gratitude is to understand and use the words *murder* and *gratitude* in a certain way, and to have a correct idea is to understand and use the words in the same way others do. The question whether *A* has shown gratitude in his conduct to *B* is a question not only what *A*'s conduct has been but also whether it sufficiently fits the accepted sense of *gratitude.*

Finally, Locke extended this to all general ideas and rejected the Aristotelian thesis that apprehending universals is apprehending real kinds, or real principles of classification. In maintaining this he was making a move toward a kind of nominalism, for he was emphasizing the fact that concepts, other than those determined by technical or arbitrary definition, are open-ended. We do not find objects and their features divided by nature or God into real and objectively delimited classes; we observe objects and their features, but the distinction between one class and another is something we ourselves make by criteria of convenience and utility. Similarities and differences are there for us to observe; whether the similarities are sufficiently close so that we can place the objects in the same or in different classes is for us to decide. A modern example would be the question whether a machine can think, or whether a computer can remember. Such a question, Locke would insist, is to be answered only by seeing what operations the machine performs and then deciding whether they are sufficiently close to what we mean by *thinking* or *remembering* when we talk of our own activities to make it reasonable, rather than misleading, to describe them in these terms.

A consequence of this kind of conceptualism will be that concepts are not permanently fixed, as on a simple realist theory they would be; a concept is liable to development and change, as fresh experience or changes of view show the need or utility of it. For example, a central question of twentieth- and twenty-first-century sociology, which concerns not only moral outlooks but also legal decisions and the development of law and penal policy, is the question under what conditions a man is to be held not responsible for his physical actions. But the answer to the question is not to be reached simply by determining whether the physical, psychological, and medical facts of a particular case place it inside or outside the accepted scope of responsibility; it also leads to examining the notion of responsibility itself, which in the slow process of time undergoes modification. Experience being ineluctably conceptual, not only are concepts derived from experience, but concepts shape experience itself, as indeed Aristotle had hinted. If there were nothing else valuable in conceptualism, it would be of importance as a corrective to the naïveté of extreme realism, which suggests that all the material of human experience falls into a scheme of pigeonholes or a fixed mold and that the task of inquiry is simply to find out what the scheme or mold is.

BERKELEY. George Berkeley, Locke's immediate successor and fiercest critic, devoted the whole introduction of his main philosophical work, *The Principles of Human Knowledge* (1710), to a violent attack on Locke's theory of abstract ideas, for reasons perhaps not primarily concerned with universals at all. However, it is extremely doubtful whether he had, in fact, either studied Locke carefully enough or interpreted him correctly. Berkeley's own theory of general ideas as particular ideas that become "general by being made to represent or stand for all other particular ideas of the same sort" is expressed in a way that might be a verbatim quotation from Locke himself (cf. Locke, *Essay,* Book III, Ch. 3, Sec. 13: "Ideas are general when they are set up as representatives of many particular things.… [They] are all of them particular in their existence … their general nature being nothing but the capacity they are put into, by the understanding, of signifying or representing many particulars"). And Hume's enthusiastic comment that Berkeley's view of general ideas as particular ideas used generally is "one of the greatest and most valuable discoveries that has been made of late years in the republic of letters" does Hume little credit; his examination of Locke was clearly no more thorough than Berkeley's had been.

If Berkeley had done nothing but propound his account of general ideas, his contribution would have been nil. But, in fact, he did much more. Aware that a central strand in the supposed problem of universals was the fact of language and appreciating the question how sounds made by the human larynx or marks made on paper could be used to convey a meaning (this too had been stressed by Locke), he protested against the simple view of *unum nomen unum nominatum,* that every time

the same word is used it is accompanied in the mind by the same idea. First, this is empirically false, as anybody could find out by noticing the many different ideas (images) he might have on the different occasions he used the word; for example, *red* might be accompanied sometimes by an image of a red dress, sometimes by an image of a red apple, a red flower, and so on, which might in any case all be different shades of red. Furthermore, it is not even true that every time a man uses a word that can be accompanied by an image, it is accompanied by one. The actual occurrence of an image, if not necessary, could not help to explain the meaningfulness of a word. Sometimes Berkeley wrote as if an image were necessary in a dispositional sense; a word is significant if a suitable image can be had or produced to correspond to it. Thus, he compared a use of language—for instance, in conversation—to the use of algebraic symbols in a calculation: We can represent a given quantity by the symbol *x,* and we carry out the calculation without all the time thinking of the quantity represented by *x*; what matters is that we can, at any time we want to, replace *x* by the quantity. Similarly, words for the most part, as actually employed, function as cashable counters.

But Berkeley went on to emancipate himself even from this tenuous servitude to ideas as images. He hinted at it when he said that the important thing is the definition of a word, not the occurrence or recurrence of an idea: "It is one thing for to keep a name constantly to the same definition, and another to make it stand every where for the same idea: The one is necessary, the other useless and impracticable." But later he went even further and suggested what can be described as an operational theory of meaning. This is nowhere fully developed, chiefly because he abandoned serious philosophical inquiry while still a young man, but unmistakable indications of it persist throughout his writings.

In the *Principles* they appear in two ways: (*a*) the reminder of the diversity of function of language; and (*b*) the doctrine of "notions." The tendency among philosophers to try to explain the significance of words in terms of corresponding ideas was due to a simple and entirely false view of language, namely that its sole function was informing, or "the communication of ideas"; this made it easier to think of ideas as pictures translated into words by the speaker and retranslated into pictures by the hearer. (The modern television analogy of visual pictures translated into radio signals by the transmitter and retranslated into visual pictures by the receiving set would not be entirely inapt.) But as Berkeley rightly emphasized, to inform is not the function of language, only one of its functions. It has others, "the raising of some passion, the exciting to or deterring from an action, the putting the mind in some particular disposition"—to which we could add still others, such as asking questions, praying, vowing, swearing, making promises, declaring intentions, and expressing wishes or fears.

It is not entirely clear exactly what Berkeley intended the doctrine of "notions" to be. He acknowledged that his own principles did not allow him to say that we have (or can have) ideas of everything we may significantly talk of, because they did not allow him to say that we have ideas of mind or spirit (ideas being passive and mind or spirit being active); yet a man who uses the words *mind* and *spirit* (to which Berkeley added all words denoting relations) is not uttering meaningless gibberish. Therefore, it must be true of at least some words that we "know or understand what is meant" by them although we can have no corresponding ideas. In these cases we have notions. Notions, as they appear in the *Principles,* do not solve any problem (if one exists) regarding how words that cannot be paralleled by ideas can be significant—they merely occur as a label for the fact that there are such words. They are not the answer but appear to be Berkeley's name for the question. If by "having a notion of *x*" he meant "knowing or understanding the meaning of the word *x,* although not being able to have an idea of *x,*" then the question how one can know or understand the meaning of an idealess word is not answered by saying that he has a notion, and there is no reason to think that Berkeley deluded himself into supposing that his doctrine of notions actually gave an answer to anything. The *Principles* takes the matter no further than the negative conclusion not only that a word need not be accompanied by an idea but also that some words cannot be. This is the beginning of an admission that the intelligibility of language neither requires nor is illuminated by suppositions about mental imagery.

In a much later work, *Alciphron* (1732), Berkeley returned to the topic and showed how (with the examples of force from physics and grace from theology) although frontal questions such as "What is force?" and "What is grace?" could produce no answer, yet these were genuine concepts, because it was true that the use of them (or of the words *force* and *grace*) could lead to fruitful results. Or again, "the algebraic mark, which denotes the root of a negative square, hath its use in logistic operations, although it be impossible to form an idea of any such quantity." In allowing that a concept could be fertile even though it could not be cashed, Berkeley was at once breaching the walls of strict empiricism and anticipating

the theory construction of modern science, particularly of modern physics.

HUME. Immediately after Berkeley came David Hume, the third of the great British empiricists and the one who has had the most lasting influence on subsequent developments in the philosophy of that school. He devoted an early section of his *Treatise of Human Nature* (1739) to the subject of abstract ideas (Book I, Part i, Sec. 7), professing to accept Berkeley's doctrine of general ideas and producing arguments to confirm it. But in fact he was not merely repeating Berkeley's views. He took one step backward in maintaining that the use of every general word must be accompanied by a particular mental idea: "'Tis certain that we form the idea of individuals, whenever we use any general term." But he took several steps forward in suggesting how a given idea can represent others of the same kind—that is, how the idea can become general.

Hume's emphasis on the role of the word was even stronger than Berkeley's had been. Whereas Berkeley had supposed that a word becomes general by its relation to a particular but representative idea, Hume put it the other way round, that a particular idea becomes general by being "annexed to a certain term." "All abstract ideas are really nothing but particular ones … but, being annexed to general terms, they are able to represent a vast variety." Where Berkeley had contented himself with maintaining that an idea became general by representing all ideas of that kind, Hume offered an account of how a particular idea could represent others that were not at the time present to the mind. It did this through custom or habit, by the association of ideas and the association of words. At any given moment a man has only one individual idea before his mind, but because of the resemblances that he has found in his experience, the one individual idea is associated with others of the same kind, which are not actually present to the mind at the time but which would be called up by the stimulus of a suitable experience or a suitable word. Thus, the possession of a general idea or a concept becomes a mental disposition, the readiness, engendered by custom, to have some idea belonging to a given kind, when the appropriate stimulus occurs, and the acquisition of a concept will be the gradual process of (1) learning by experience and habituation to recognize instances and to discriminate between them and instances of a different concept; and (2) having the appropriate associations and dispositions set up in one's mind. To have a concept actually in mind at any given time is to have in mind an individual idea plus the appropriate associative dispositions.

Hume assigned words a key role in his doctrine of association of ideas, supposing that particular ideas, which resemble one another somewhat but not exactly or in all respects, tend to be associated with one another because each is associated with the same general word. The differences between a ripe tomato and a scarlet-painted automobile are more numerous and conspicuous than their similarities, but the idea of the one can readily be associated with that of the other by the fact that the word *red* is used of each, and thus the idea of either could serve as representative of the class of red objects, whatever the variety of objects and the differences between the many shades of red displayed. "A particular idea becomes general by being annex'd to a general term; that is, to a term, which from a customary conjunction has a relation to many other particular ideas, and readily recalls them in the imagination." One could say that according to Hume we learn to think by learning to talk, not the other way round, and that in learning to talk the chief influence is that of custom and association. Here Hume failed, as nominalism also failed, to see that the attempt to account for the generality of an idea in terms of the generality of a word will not do, if taken only as far as he took it. In the sense in which he insisted that every idea is particular, so is every word. Whatever reasons there are for denying the existence of general ideas as distinct from particular ideas will also be reasons for denying the existence of general words as distinct from particular words. Paradoxical though it may seem, the sense in which the word *red* may be said to be general is such that the word *red* cannot occur in any sentences at all, for what occurs in a particular sentence is a particular word *red.* The fourth word in the sentence "Some automobiles are red" may be very like the first word in the sentence "Red tomatoes are ripe," but they are different individual words, occupying different positions in space (as printed). Even in this case they are not exactly alike (for the first does not, and the second does, start with a capital letter), and other "reds" could be even more unlike—for instance, if they were printed in different fonts of type or were written down by different people.

Consideration of this point would have required Hume to say about a word's being general what he (like Locke and Berkeley) said about an idea's being general, namely that it was based on (or constituted by) the resemblance between particulars. (Difficulties in making out somebody's handwriting stem precisely from its deviating more than usual from the familiar resemblances.) Conceptualism therefore comes down, in the persons of these three authors, on the side of resemblance as being the ontological basis of general ideas. All that actually

exists is individual; generalization, or concept formation, is possible only to the extent that individual objects and occurrences, their features, and the relations between them display perceptible resemblances to a greater or lesser extent. But Hume offered, or at least hinted at, a more sophisticated version of resemblance. According to Locke, two objects would resemble each other if they possessed certain features in common—that is, if certain features of the one were identical (in an Aristotelian sense) with certain features of the other. Thus, one object possessing features *abcd* would resemble another possessing features *adef,* but less closely than it resembled one possessing features *acdf.* But Hume saw that this raised difficulties for simple (or unanalyzable) ideas or qualities—for example, that "*blue* and *green* are different simple ideas, but are more resembling than *blue* and *scarlet*; tho' their perfect simplicity excludes all possibility of separation or distinction." They may resemble each other "without having any common circumstance the same." The notion of resemblance as an ultimate relation, without requiring that the respect in which two objects resemble each other should be a quality identical in each, propounded here by Hume, has been taken further in later developments of his theory.

NOMINALISM AND RESEMBLANCE

NOMINALIST THEORIES. The nominalist view, that only names (or, more generally, words) are universal, "for the things named are every one of them singular and individual" (Hobbes, *Leviathan,* Ch. 4), has had a very long history. It was the subject of much controversy in medieval philosophy, more for the theological heresies it was believed to engender than on grounds of logic, and it was advanced again in the seventeenth century by Thomas Hobbes.

Of the medievalists mention need be made only of two, one early and the other late. Peter Abelard, although fiercely critical of the extreme nominalism of Roscelin de Compiègne, was strongly influenced by it. For Abelard a universal was not a sound (*vox*), as it was for Roscelin, but a word (*sermo*)—that is, a meaningful sound—and it acquired its meaning from its referential use, the reference being mediated by a general idea that is a composite image. Thus, although Abelard was described by his successors as a nominalist, he was only partly and confusedly so; he could as well be called a conceptualist, or even a moderate realist.

William of Ockham, a polemical figure who was pronounced a heretic and excommunicated, produced a number of logical works in which he developed a battery of arguments against realism and supported a form of nominalism. According to him, universals are terms or signs standing for or referring to individual objects and sets of objects, but they cannot themselves exist. For what exists must be individual, and a universal cannot be that; the mistake of supposing that it could was the fatal contradiction of Platonic realism. And Aristotelian realism was no better, for it involved its own contradiction, that the identical universal should be present in a number of particulars. Real universals are neither possible nor needed. Rather, universals are predicates or meanings, possessing logical status only, required for thought and communication, not naming anything that could possibly exist.

In its extreme form, that there is nothing common to a class of particulars called by the same name other than that they are called by the same name, nominalism is so clearly untenable that it may be doubted whether anybody has actually tried to hold it. If all the individuals (objects, qualities, or whatever they were) called by the same name—for example, "table"—had nothing in common but being called by the same name, no reason could be given why just they and no others had that name, and no reason could be given for deciding whether to include an object in or to exclude it from the class. On a realist view certain objects are called "tables" because they are tables (that is, they partially embody a Platonic Form of tablehood or possess a common Aristotelian feature of tablehood). On an extreme nominalist view they are tables only because they are called "tables," and no answer at all can be given to the question why certain objects are (or are to be) called "tables" and others not. Perhaps the only extreme nominalist has been Humpty Dumpty. ("'When *I* use a word, it means just what I choose it to mean—neither more nor less.' 'The question is,' said Alice, 'whether you *can* make words mean different things.' 'The question is,' said Humpty Dumpty, 'which is to be master—that's all.'") Moderate nominalism, while retaining the view that only words are universals, saves itself from total subjectivity by basing the use of words on the resemblances between things. Hobbes, for example, in the *Leviathan* (Ch. 4) said: "One universal name is imposed on many things, for their similitude in some quality or other accident." So *table* is a universal word, applicable to any individual objects between which a certain resemblance holds. Objects, their qualities, and their relations are all individual, the only thing that is general being the word that is applicable to objects (or qualities, or relations) of a given class in virtue of the resemblances between them.

Nominalism and a conceptualism such as Hume's here converge, differences being in approach and emphasis rather than in substance. And nominalism must in the end reduce itself to a resemblance theory that, if acceptable, finally renders nominalism unnecessary. Nominalism's only reason for insisting on the universality of the word is its denial of the universality of the thing: Things are individuals, and the properties of a thing are individual to it. But the universality of the word depends on resemblances between things; thus, nominalism requires a resemblance theory. However, as was already mentioned in reference to Hume, the nominalist must, to be consistent, go further and recognize that what he says of things, if true of them, must be true of words also, which requires him to make what logicians have called the "type-token" distinction. Any occurrence of the word *red* is individual ("red" as a token), and two occurrences of what would be called the same word ("Red" as a type) are occurrences of the same word only in that they resemble each other in the relevant ways. Thus, the universal word *Red* becomes the class of the resembling individual words "red," "***red***," "RED," and so on, and once the universality of a word has been analyzed along these lines, the reason for saying that only words are universal is gone, for exactly the same account can be given by the resemblance theory of universality in things. Nominalism was able to present the appearance of being a distinct theory of universals only as long as its exponents and critics alike failed to apply to it the type-token distinction. Once that is applied, words are seen to be on all fours with things, and the question becomes, for words as for things, whether generality can be analyzed simply in terms of resemblances between individuals, as Hume suggested. Nominalism not only requires the support of a resemblance theory to explain how a word can have a general use but also, in its only consistent form, is a resemblance theory.

RESEMBLANCE THEORIES. Whether or not Hume actually held what might be described as the pure-resemblance theory, that is the only form of resemblance theory that is distinctive. The version advanced by Locke, and possibly by Berkeley, too, according to which the degree of resemblance between two objects depends on the extent of qualitative identity between them, collapses into a modified Aristotelian realism. Pure resemblance, although allowing that if two objects resemble each other there must be some respect in which they are similar, would deny that this respect is to be regarded as an identical something common to both; not to deny this would be to reintroduce the Aristotelian universal. Red objects are to be called red simply because they resemble each other in a way in which they do not resemble blue objects, or hard objects, or smooth objects, or spherical objects. Nothing is described by saying that the universal red is what is common to any pair of red objects that is not more accurately and less misleadingly described by saying that both are red—that is, resemble each other in respect of each being red. There is a similarity between the red of the one and the red of the other, and the similarity might be anything from being virtually exact (as in two new red postage stamps of the same denomination) to being only approximate and generic (as in two flags of widely different shades of red, one flag, in addition, being bright and new, the other old and faded). The world is made up of individual things and events, with their individual qualities and relations, and with resemblances in different respects and of differing degrees. Were it not for such resemblances (and contrasting differences), concept formation and language would be impossible; indeed, biological survival would be impossible, too. The resemblance theory is metaphysically the most economical, but it has objections to face, notably two: (1) It does not succeed in dispensing with universals in a traditional sense, such as the Aristotelian, because resemblance itself will have to be such a universal, and if it is, there is no ground for denying others. (2) As two objects that resemble each other must be similar in some respect, the respect must be something common to both.

Although these two objections are frequently reiterated, it is not clear that either has great force, as is shown by H. H. Price's detailed discussion in *Thinking and Experience* (1953). The argument that the resemblance theory requires resemblance itself to be a universal in a sense in which the theory denies that there are any universals has been the more persistent; it is particularly associated with Bertrand Russell (although he was not the first to propound it). But although he advanced it in two books widely separated in time, *Problems of Philosophy* (1912) and *An Inquiry into Meaning and Truth* (1940), his confidence seems to have diminished. Originally he maintained a realist theory of universals, of a Platonic kind, and held that it could be proved, at least in the case of relations, that there must be such universals.

> If we wish to avoid the universals *whiteness* and *triangularity,* we shall choose some particular patch of white or some particular triangle, and say that anything is white or a triangle if it has the right sort of resemblance to our chosen particular. But then the resemblance required will have to be a universal.

That is, we could theoretically dispense with universals of quality by analyzing them in terms of relation, and ultimately in terms of the relation of resemblance. The latter we cannot dispense with, for if we say that the resemblance between a pair of similar particulars is itself a particular relation, we shall then have to admit a resemblance between that resemblance relation and the resemblance relation holding between another pair of similar particulars; the only way to save ourselves from an infinite regress (of resemblances between resemblances between resemblances … between resemblance relations) is to admit that "the relation of resemblance must be a true universal. And having been forced to admit this universal, we find that it is no longer worth while to invent difficult and unplausible theories to avoid the admission of such universals as whiteness and triangularity." In this respect, Russell held, the rationalists were right, as against the empiricists like Hume: The existence of real universals has been proved, at least in the case of the relation of resemblance, and no good reason is left for denying it in the case of other relations and of qualities.

Some years later, in *The Analysis of Mind* (1921), Russell showed more hesitation, when he wrote, "I *think* a logical argument could be produced to show that universals are part of the structure of the world." Finally, in the *Inquiry,* after repeating his original argument, he said, "I conclude, therefore, though with hesitation, that there are universals, and not merely general words."

Price seems to have lost confidence in the validity of Russell's proof even more thoroughly, and far more rapidly. In *Thinking and Representation* (1946) he accepted that resemblance has to be a universal and repeated that the most the resemblance theory would have achieved "would be to reduce all other universals to this one relational universal." He went on: "This is a very notorious difficulty, and perhaps by much repetition it has become a bore. Yet I do not think it has ever been answered." But in *Thinking and Experience* (1953) he thought the difficulty could be answered, and he spent several pages answering it. Admittedly, his first argument is hardly convincing, namely that the opponents of the resemblance theory (such as Russell) are begging the question by assuming the very thing that they have to prove, that there are universals: From the fact that the theory analyzes all other alleged universals in terms of resemblance, and that it is ultimate, it does not follow that resemblance is a universal. We cannot answer the question whether there are any universals by replying that even if there are no other universals, resemblance must be one. Against Price here, it may be doubted whether Russell's objection is of this question-begging form. The objection, rather, is that the only way of avoiding the admission of resemblance as a universal leads to a vicious infinite regress. Nevertheless, Russell's objection is invalid, as the next stage of Price's answer shows. It is true that the resemblance theory would have to admit different orders or levels of resemblance, resemblances between pairs of particulars, resemblances between these resemblances, and so on ad infinitum. But there is nothing logically vicious or unintelligible about that. The resemblance that we notice between any pair of similar individuals is as individual as they and as the qualities of each; the resemblance we notice or can find between such a resemblance relation and another resemblance relation holding between another pair of similar individuals is itself individual; the process can be continued as long as patience and imagination hold out. We do not need a real universal of resemblance to stop the regress, simply because the regress does not need to be stopped. The fallacious assumption at the root of this objection to the resemblance theory is not the question-begging assumption that there are universals but the assumption that unless there are, a vicious regress is generated.

The merit of the resemblance theory is that it does not confuse, as the realist theories arguably did, the roles of explanation and description. Why or how tables are tables rather than chairs, and elephants are elephants rather than tigers, is not answered by saying that each is what it is because it instantiates the appropriate universal. The only explaining that has to be done on why a given object is a table is to be done in causal terms. What does have to be explained is something about ourselves, namely how it is that we can (indeed, must) experience, in terms of kinds and generality, that we form concepts, and that we develop language for communication. That experience, thought, and language depend on the use of universals, in some sense, is undeniable, and the explanation of this is to be given by a suitably illuminating description of the world we experience. About ourselves the question of universals is a question of explanation. About our world the question of universals is a question of description, and this the resemblance theory seems adequately, and nontendentiously, to provide.

In the twentieth century, philosophers paid far more attention to actual language and, largely under the influence of Ludwig Wittgenstein, came to appreciate that even if the notion of there being (in some sense) something common to all instances covered by a single general word is true of some words, it is not true of all, and that even the resemblances within a group of things all called

by the same general name may be what Wittgenstein called "family resemblances"—the vague and overlapping likenesses that one sees between the different members of a family. His own example is what "we call 'games.'" He meant "board-games, card-games, ball-games, Olympic games, and so on. What is common to them all? Don't say: 'There *must* be something common, or they would not be called "games"'—but *look and see* whether there is anything common to all" (*Philosophical Investigations,* I, Sec. 66). There is nothing common to all games, only "similarities, relationships, and a whole series of them at that." The concept of causality, too, has stubbornly resisted the attempts of philosophers to analyze it, as though there were only one *it* to analyze—although the hint that it really requires the Wittgenstein treatment first came from Aristotle himself.

The history of the subject of universals has come a long way from looking for a general entity for which a general word is to be the name (Plato), via looking for recurring identities (Aristotle), selected identities (Locke), and resemblances (Hume), to looking for varying and overlapping resemblances and recognizing that only vain servitude to a theory insists on trying to find what is common to a whole range of overlaps (Wittgenstein). Furthermore, with the development of semantics it has come to be appreciated that not all general words are, even in a stretched sense, "names" at all. They can be significant for their syntactical function, indicating, for instance, condition or conjunction or contrast ("if," "and," "although") or, again, attitudes, outlooks, or degrees of confidence ("perhaps," "probably," "certainly"). The philosophical history of universals has been plagued by the persistent treatment of words as names, which has been made easier by philosophers' taking as their examples only objects and their qualities. But questions about universals are questions about generality, and generality is the essential feature of all words, not just of those that might plausibly be called names.

See also Abelard, Peter; Aristotle; Augustine, St.; Avicenna; Berkeley, George; Empiricism; Epistemology, History of; Gassendi, Pierre; Hobbes, Thomas; Hume, David; Illumination; Kant, Immanuel; Laws of Nature; Locke, John; Medieval Philosophy; Plato; Properties; Realism; Relations, Internal and External; Roscelin; Russell, Bertrand Arthur William; Semantics; Socrates; Subject and Predicate; Thomas Aquinas, St.; William of Ockham; Wittgenstein, Ludwig Josef Johann.

Bibliography

PLATO AND ARISTOTLE

Plato introduced his theory of Forms into many different dialogues, in particular *Phaedo, Republic,* and *Parmenides*; in the last of these he summarized the trend of his thought in earlier dialogues and subjected it to criticism, which was further developed by Aristotle, as in *Metaphysics* M. Aristotle's own views are briefly indicated in *Posterior Analytics* II, 19. Sir David Ross, in *Plato's Theory of Ideas* (Oxford: Clarendon Press, 1951), provides a useful account and discussion of the development of Plato's views and Aristotle's criticisms.

MEDIEVAL PHILOSOPHY

Some account of medieval philosophy's treatment of the theme of universals is given in Father Frederick Copleston's *A History of Philosophy,* Vols. II and III (London, 1950). A more detailed discussion of the four key figures in the dispute between realism and nominalism—Augustine, Abelard, Thomas, and William of Ockham—is to be found in M. H. Carré's *Realists and Nominalists* (London: Oxford University Press, 1946). *Selections from Mediaeval Philosophers,* edited by Richard McKeon (London, 1928), contains a few relevant passages. Copleston, in his bibliographies, provides references to editions of the full texts, where available, and to the appropriate volumes of J. P. Migne's *Patrologia Latina.*

SEVENTEENTH- AND EIGHTEENTH-CENTURY PHILOSOPHY

Hobbes's few remarks on universals are to be found in his *Elements of Philosophy,* I, 2, and in *Leviathan,* Ch. 4. Locke scattered comments all over his diffuse and repetitious *Essay on Human Understanding,* but the main entries are II, xi, and III, iii. Berkeley devoted the whole of the introduction to his *Principles of Human Knowledge* to the subject and returned to it, in a rather more sophisticated way, in *Alciphron,* 7.4. Hume dispatched it briskly in his *Treatise of Human Nature,* I, i, 17. Thomas Reid, in his *Essays on the Intellectual Powers of Man,* V, 6, subjected the other philosophers to telling criticism and foreshadowed modern tendencies.

RECENT PHILOSOPHY

In *Studies in Philosophy and Psychology,* Vols. XV–XVII (London: Macmillan, 1930), G. F. Stout reprinted three relevant papers, the last criticizing the resemblance theory and advocating the view of a universal as a "distributive unity" of a class. Bertrand Russell followed his paper "On the Relation of Universals and Particulars," in *PAS* (1911–1912), with *Problems of Philosophy* (London: Williams and Norgate, 1912), which contains two chapters on the subject; it is taken up again in *Analysis of Mind* (London: Macmillan, 1921) and *Inquiry into Meaning and Truth* (London: Allen and Unwin, 1940). Russell's views are the subject of an article by O. K. Bouwsma in *Philosophical Review* (1943). Other relevant articles are F. P. Ramsey, "Universals," in *Mind* (1925); A. J. Ayer, "On Particulars and Universals," in *PAS* (1933–1934); R. I. Aaron, "Two Senses of the Word Universal," in *Mind* (1939), and "Our Knowledge of Universals," in *Proceedings of the British Academy* (1944); Morris Lazerowitz, "The Existence of Universals," in *Mind*

(1946); Nelson Goodman and W. V. Quine, "Steps towards a Constructive Nominalism," in *Journal of Symbolic Logic* 12 (1947); W. V. Quine, "On What There Is," in *Review of Metaphysics* (1948–1949); A. N. Prior, in *Mind* (1949); A. C. Lloyd, "On Arguments for Real Universals," in *Analysis* (1951); D. F. Pears, "Universals," in *Philosophical Quarterly* (1950–1951); R. B. Brandt, "The Languages of Realism and Nominalism," in *Philosophy and Phenomenological Research* (1956–1957); Arthur Pap, in *Philosophical Quarterly* (1959–1960); and Renford Bambrough, "Universals and Family Resemblances," in *PAS* 61 (1960–1961). The last paper takes as its point of departure the "family resemblance" account of the use of general words given by Ludwig Wittgenstein in *The Blue and Brown Books* (Oxford: Blackwell, 1958), pp. 17–27, and *Philosophical Investigations* (Oxford: Blackwell, 1953), Secs. 65–77. A general survey of the problems connected with universals is undertaken, at a level of no great philosophical difficulty, by R. I. Aaron in *The Theory of Universals* (Oxford: Clarendon Press, 1952) and, more briefly, by A. D. Woozley in *Theory of Knowledge* (London: Hutchinson's University Library, 1949). Other books, each containing several chapters on the subject, are Nelson Goodman's *Structure of Appearance* (Cambridge, MA: Harvard University Press, 1951), John Holloway's *Language and Intelligence* (London: Macmillan, 1951), and, most detailed of all, H. H. Price's *Thinking and Experience* (Cambridge, MA: Harvard University Press, 1953). Papers by I. M. Bocheński, Alonzo Church, and Nelson Goodman are included in the symposium *The Problem of Universals* (Notre Dame, IN: University of Notre Dame Press, 1956).

A. D. Woozley (1967)

UTILITARIANISM

"Utilitarianism" can most generally be described as the doctrine that states that the rightness or wrongness of actions is determined by the goodness and badness of their consequences. This general definition can be made more precise in various ways, according to which we get various species of utilitarianism.

ACT AND RULE UTILITARIANISM

The first important division is between "act" utilitarianism and "rule" utilitarianism. If, in the above definition, we understand *actions* to mean "particular actions," then we are dealing with the form of utilitarianism called act utilitarianism, according to which we assess the rightness or wrongness of each individual action directly by its consequences. If, on the other hand, we understand *actions* in the above definition to mean "sorts of actions," then we get some sort of rule utilitarianism. The rule utilitarian does not consider the consequences of each particular action but considers the consequences of adopting some general rule, such as "Keep promises." He adopts the rule if the consequences of its general adoption are better than those of the adoption of some alternative rule.

Since, in this context, the word *rule* can be interpreted in two ways, to mean either "possible rule" or "rule actually operating in society," there are actually two species of rule utilitarianism. If we interpret *rule* simply as "possible rule," we get an ethical doctrine strongly resembling that of Immanuel Kant. It is true that Kant is not normally regarded as a utilitarian, but nevertheless a utilitarian strain can be detected in his thought. If we interpret his categorical imperative, "Act only on that maxim through which you can at the same time will that it should become a universal law," as meaning "Act only on that maxim which you would like to see established as a universal law," and if liking here is determined by the individual's feelings as a benevolent man, then we get a version of utilitarianism which may usefully be called Kantianism. It is true that Kant would object to this appeal to feelings of benevolence and would wish to distinguish sharply between *willing* and "*wanting* or *liking.* Nevertheless, it is far from clear how Kant's distinction can be defended; and when he elucidates his general principle by means of examples, he does indeed tend to think in terms of the consequences that we should like to see brought about. However, the word *Kantianism* is used here merely as a useful and perhaps not inappropriate label; whether Kant himself would approve of its present application is not an important issue in the present discussion.

If, in our definition of *utilitarianism,* we interpret the word *rule* as "actual rule," or "rule conventionally operative in society," we get a form of rule utilitarianism that has been propounded in recent times by Stephen Toulmin, who seems mainly concerned with the justification, and in some cases the reform, of rules of conduct that are actually operative in society.

When we think of the writers with whom the term *utilitarianism* is most naturally associated, namely, Jeremy Bentham, J. S. Mill, and Henry Sidgwick, we must think of utilitarianism primarily as act utilitarianism. However, controversy has developed over whether Mill should not rather be interpreted as a rule utilitarian, and there has also been much discussion of the rival claims of act and rule utilitarianism to be viable ethical theories.

R. M. Hare, in his book *Freedom and Reason* (Oxford, 1963), has recently argued that there is no clear distinction between act and rule utilitarianism, since if a certain action is right, it must be the case that any action just like it in relevant respects will also be right. If these respects are then specified in detail, we get a rule of the form "Do

actions of this sort." A defender of the distinction between act and rule utilitarianism could reply that since the situations in which actions occur are infinitely variable, and since no two actions have quite the same sorts of consequences, the act utilitarian may not be able to describe the "relevant respects" mentioned above in any less general form than "The action is of the sort that has the best consequences." But if this is so, Hare's principle that if an action is right then any action which is like it in the relevant respects is also right does not yield a sufficiently particular form of rule to justify the assimilation of act and rule utilitarianism.

EGOISTIC AND UNIVERSALISTIC UTILITARIANISM

Act utilitarianism, unlike rule utilitarianism, lends itself to being interpreted either in an egoistic or in a nonegoistic way. Are the good consequences that must be considered by an agent the consequences to the agent himself (his own happiness, for example), or are they the consequences to all humankind or even to all sentient beings? If we adopt the former alternative, we get egoistic utilitarianism; and if we adopt the latter alternative, we get universalistic utilitarianism. Since what is best for me is unlikely to be what is best for everyone, it is clear that there is not only a theoretical but also a practical incompatibility between egoistic and universalistic utilitarianism. This was not always seen by the early utilitarians, who sometimes seem to have confused the two doctrines. There is, in fact, even a pragmatic inconsistency in egoistic utilitarianism, since an egoist, on his own principles, would be unlikely to wish to be seen in his true colors, and so would have no motive for expressing his ethical doctrine. In this entry we shall be concerned with utilitarianism in the universalistic sense.

HEDONISTIC AND IDEAL UTILITARIANISM

Another distinction, which cuts across that between act and rule utilitarianism, is the distinction between hedonistic and ideal utilitarianism. Utilitarianism has been defined above as the view that the rightness or wrongness of an action depends on the total goodness or badness of its consequences. A hedonistic utilitarian will hold that the goodness or badness of a consequence depends only on its pleasantness or unpleasantness. As Bentham put it, quantity of pleasure being equal, pushpin is as good as poetry. An ideal utilitarian, such as G. E. Moore, will hold that the goodness or badness of a state of consciousness can depend on things other than its pleasantness. According to him, the goodness or badness of a state of consciousness can depend, for example, on various intellectual and aesthetic qualities. In his calculations, the ideal utilitarian will be concerned not only with pleasantness and unpleasantness, but also with such things as knowledge and the contemplation of beautiful objects. He may even hold that some pleasant states of mind can be intrinsically bad, and some unpleasant ones intrinsically good. J. S. Mill took up an intermediate position. He held that although pleasantness was a necessary condition for goodness, the intrinsic goodness of a state of mind could depend on things other than its pleasantness, or, as he put it, there are higher and lower pleasures.

It should be noted that we have assumed that the only things that can be intrinsically good or bad are states of consciousness. Other things can of course be extrinsically good or bad. For example, an earthquake is normally extrinsically bad, that is, it causes a state of affairs that is on the whole intrinsically bad. Moreover, a utilitarian can hold that something that is intrinsically bad, such as the annoyance of remembering that we have forgotten to do something, is extrinsically good, for it is a means to a set of consequences that are on balance intrinsically good. G. E. Moore held that states of affairs other than states of consciousness could be intrinsically good or bad. For an ideal utilitarian, this is a theoretically possible contention, but nevertheless, few ideal utilitarians would find the contention a plausible one, and we shall therefore ignore it in this article.

NORMATIVE AND DESCRIPTIVE UTILITARIANISM

Utilitarianism may be put forward either as a system of normative ethics, that is, as a proposal about how we ought to think about conduct, or it may be put forward as a system of descriptive ethics, that is, an analysis of how we do think about conduct. The distinction between normative and descriptive utilitarianism has not always been observed. It is important to bear carefully in mind the distinction between normative and descriptive utilitarianism and to note that objections to descriptive utilitarianism do not necessarily constitute objections to normative utilitarianism.

HISTORICAL REMARKS

Properly speaking, utilitarianism began with Jeremy Bentham (1748–1832), who was a universalistic hedonistic act utilitarian. He put forward his view essentially as normative ethics, but he was unclear about the distinction between normative and factual utterances and may justly

be accused of committing what Moore later called the naturalistic fallacy—the fallacy of claiming to deduce ethical principles solely from matters *of* fact. (David Hume had in effect pointed out this fallacy before Bentham's time.)

PRECURSORS OF UTILITARIANISM. Anticipations of Bentham are to be found in the history of ethics. In ancient times Aristippus of Cyrene and Epicurus propounded hedonistic theories. However, their doctrines approximate egoistic rather than universalistic utilitarianism, despite the fact that they were unclear about the difficulty of reconciling the two doctrines and hence tried to have it both ways. The same might be said of Abraham Tucker and William Paley, the more immediate precursors of Bentham, who also injected certain theological conceptions into their systems. The tension between egoistic and universalistic hedonism can also be detected in the eighteenth-century French writer Claude-Adrien Helvétius, who appears to have influenced Bentham; also, the political philosopher William Godwin should be mentioned. David Hume is often classified as a utilitarian, but he used utility not as a normative or even as a descriptive principle, but as an explanatory one: When asked why we approve of certain traits of character, he would point out that they are traits which either are useful or are immediately agreeable. Both because he used the principle of utility in an explanatory way and because he was primarily concerned with the evaluation of traits of character (virtues and vices and the like) rather than with the question of what actions ought to be done, it is not advisable to regard Hume as a utilitarian.

J. S. MILL. As was mentioned above, there has been some controversy over whether J. S. Mill (1806–1873) ought to be regarded as an act utilitarian or as a rule utilitarian. Mill does not make his position on this issue very clear. Probably he was not very well aware of the distinction, and in any case he would probably have thought it a fairly unimportant one, since he was mainly concerned with the opposition between utilitarianism in general and other systems of ethics that were quite nonutilitarian. Although Bentham had on at least one occasion used the word *utilitarian,* it was Mill who introduced it into philosophy. He appropriated it, with some change of meaning, from a passage in the Scottish novelist John Galt's *Annals of the Parish* (Edinburgh, 1821).

SIDGWICK. We can with some confidence classify Mill as a normative utilitarian rather than a descriptive one, but the first utilitarian philosopher who was very explicit on this issue was Henry Sidgwick (1838–1900). Sidgwick understood that there is a distinction between normative and factual sentences, although, like G. E. Moore (1873–1958), he thought that ethical principles could be the objects of intellectual intuition. Sidgwick was a universalistic hedonistic utilitarian, but he was also strongly attracted by the claims of egoism. He saw more clearly than earlier writers that there was a theoretical inconsistency in being both an egoistic and a universalistic utilitarian, and he considered the possibility that there might be theological sanctions that would reconcile the two views, if not in theory, then at least in practice.

LATER UTILITARIANS. Moore and Hastings Rashdall were ideal universalistic utilitarians, although Moore, with his principle of organic unities, and Rashdall, with his importation into the utilitarian calculations of the moral worth of the actions themselves, introduced considerations which, if taken seriously, would seem to vitiate the truly utilitarian character of their theories.

A subtle form of rule utilitarianism of the sort we have called Kantianism was propounded in 1936 by R. F. Harrod. Contemporary writers such as Stephen Toulmin, P. H. Nowell-Smith, John Rawls, K. E. M. Baier, and M. G. Singer have propounded views that either are or approximate rule utilitarianism. R. B. Brandt has been sympathetic to rule utilitarianism and has recently defended a rather subtle and complex version of it.

ANALYSIS AND CRITIQUE

UTILITARIANISM AS A DESCRIPTIVE ETHICS. It is fairly easy to show that both act utilitarianism and rule utilitarianism are inconsistent with usual ideas about ethics, or what can be called the common moral consciousness. For the principles of both systems will in some cases lead us to advocate courses of action that the plain man would regard as wrong. Consider, for example, the case of a secret promise to a dying man. To ease his dying moments, I promise him that I will deliver a hoard of money, which he entrusts to me, to a rich and profligate relative of his. No one else knows either about the promise or the hoard. On utilitarian principles, it would appear that I should not carry out my promise. I can surely put the money to much better use by giving it, say, to a needy hospital. In this way I would do a lot of good and no harm. I do not disappoint the man to whom I made the promise, because he is dead. Nor, by breaking the promise, do I do indirect harm by weakening men's faith in the socially useful institution of promise making and promise keeping, for on this occasion no one knows

about the promise. Normally, of course, an act utilitarian will keep a promise even when the direct results are not beneficial, because the indirect effects of sowing mistrust are so harmful. This consideration clearly does not apply in the present instance. The plain man, however, would be quite sure that the promise to the dying man should be kept. In this instance, therefore, we have a clear case in which utilitarianism is inconsistent with the way in which, for the most part, people in fact think about morality.

The rule utilitarian, on the other hand, would probably agree with the plain man in the above case, because he would appeal to the utility of the rule of promise keeping in general, not to the utility of the particular act of promise keeping. Nevertheless, cases can be brought up that will show the incompatibility of even rule utilitarianism with the common moral consciousness. For example, a riot involving hundreds of deaths may be averted only by punishing some innocent scapegoat and calling it punishment. Given certain empirical assumptions, which may perhaps not in fact be true, but which in a certain sort of society might be true, it is hard to see how a rule utilitarian could object to such a practice of punishing the innocent in these circumstances, and yet most people would regard such a practice as unjust. They would hold that a practice of sometimes punishing the innocent would be wrong, despite the fact that in certain circumstances its consequences would be good or that the consequences of any alternative practice would be bad. In this instance, then, there is a conflict between even the rule utilitarian and the plain man. (This is not, of course, to say that in fact, in the world as it is, the rule utilitarian will be in favor of a practice of punishing the innocent, but it can be shown that in a certain sort of world he would have to be.)

ACT UTILITARIANISM AS A SYSTEM OF NORMATIVE ETHICS. Both act and rule utilitarianism fail, then, as systems of descriptive ethics. But act utilitarianism as a system of normative ethics would seem to have certain advantages over both rule utilitarianism and nonutilitarian, or deontological, systems of ethics (a deontological system of ethics is one that holds that an action can be right or wrong in itself, quite apart from consequences). Moreover, the failure of act utilitarianism as a descriptive system is the source of its interest as a possible normative system: If it had been correct as a descriptive system, then the acceptance of it as a normative system would have left most men's conduct unchanged.

No proof of utilitarianism. A system of normative ethics cannot be proved intellectually. Any such "proof" of utilitarianism as was attempted by Bentham or Mill can be shown to be fallacious. (Mill disclaimed the possibility of proof and spoke more vaguely of "considerations capable of determining the intellect," but he presented an attempted proof nonetheless.) Sidgwick and Moore were clearer on this point and saw that ethical principles cannot be deduced from anything else. They appealed instead to intellectual intuition, but recent developments in epistemology and other fields of philosophy have made the notion of intellectual intuition a disreputable one. The tendency among some more recent writers, such as C. L. Stevenson, R. M. Hare, and P. H. Nowell-Smith, has been to regard assertions of ultimate ethical principles and valuations as expressions of feeling or attitude, or as akin to imperatives rather than to statements of fact. In this respect, they develop further the position held much earlier by Hume. Now if we abandon a cognitivist theory about the nature of moral judgments, such as was held by Sidgwick or Moore, and adopt the view that ultimate ethical principles depend only on our attitudes, that is, on what we like or dislike, we must give up the attempt to prove any ethical system, including the act-utilitarian system. We may nevertheless recommend such a system. We may also try to show inconsistencies or emotionally unattractive features of various possible alternative systems.

Appeal to generalized benevolence. In putting forward act utilitarianism as a normative system, we express an attitude of generalized benevolence and appeal to a similar attitude in our audience. (The attitude of generalized benevolence is not the same as altruism. Generalized benevolence is self-regarding, and other-regarding too—I count my happiness neither more nor less than yours.) Of course, we all have in addition other attitudes, self-love, and particular likes and dislikes. As far as self-love is concerned, either this will be compatible with generalized benevolence or it will not. If the former, then self-love does not conflict with act utilitarianism, and if the latter, nevertheless self-love then will be largely canceled out, as among a number of people engaged in discussion.

Arguments against deontological systems. As to particular likes and dislikes, an important case concerns our liking for obeying the rules of some deontological ethics in which we have been raised. However, the following persuasive considerations can be brought up as arguments against the adherent of a deontological system of ethics. It can be urged that although the dictates of a generalized benevolence might quite often coincide with those of an act-utilitarian ethics, there must be cases in

which the two would conflict with one another. Would the benevolent and sympathetic persons to whom we conceive ourselves to be appealing be happy about preferring abstract conformity with an ethical rule, such as "Keep promises," to preventing avoidable misery of his fellow creatures?

It will be noticed that the above defense of utilitarianism against deontology is purely persuasive, an appeal to the heart and not to the intellect. It is based on the metaethical view that ultimate ethical principles are expressions of our attitudes and not the findings of some sort of intuition of ethical fact. An intellectualist in metaethics, such as W. D. Ross, could well resist our appeal to feeling by saying that it is possible to see that his deontological principles are correct, and that whether we like them or not is beside the point.

Weakness of rule utilitarianism. In defending act utilitarianism, then, we appeal to feelings, namely, those of generalized benevolence. Since people possess other attitudes too, such as loyalty to a code of morals in which they have brought up, the possession of feelings of generalized benevolence is not a sufficient condition of agreement with the act utilitarian. But it is a necessary condition. Now the rule utilitarian also appeals ultimately to feelings of generalized benevolence. Like the deontologist, however, he is open to the charge of preferring conformity with a rule to the prevention of unhappiness. He is indeed more obviously open to such a charge, since he presumably advocates his rule utilitarian principle because he thinks that these rules conduce to human happiness. He is then inconsistent if he prescribes that we should obey a rule (even a generally beneficial rule) in those cases in which he knows that it will not be most beneficial to obey it. It will not do to reply that in most cases it is most beneficial to obey the rule. It is still true that in some cases it is not most beneficial to obey the rule, and if we are solely concerned with beneficence, in these cases we ought not to obey the rule. Nor is it relevant that it may be better that everybody should obey the rule than that nobody should. That the rule should always be obeyed and that it should never be obeyed are not the only two possibilities. There is the third possibility that sometimes it should be obeyed and sometimes it should not be obeyed.

Hedonistic act utilitarianism. We shall therefore neglect rule utilitarianism as a system of normative ethics, and consider only act utilitarianism, which will be conveniently put forward in a hedonistic form. The reader will easily be able to adapt most of what is said to cover the case of ideal utilitarianism. Indeed, in many cases the differences between hedonistic and ideal utilitarianism are not usually of much practical importance, since the hedonist will usually agree that the states of mind the ideal utilitarian regards as intrinsically good, but which he does not, are nevertheless extrinsically good. Bentham would say that Mill's higher pleasures, if not intrinsically better than the lower ones, are usually more "fecund" of further pleasures. This is not to say, however, that there are no cases in which there would not be a significant difference between hedonistic and ideal utilitarianism.

The act-utilitarian principle can now be put in the following form: "The only reason for performing some action *A*, rather than various alternative actions, is that *A* results in more happiness (or more generally, in better consequences) for all humankind (or perhaps all sentient beings) than will any of these alternative actions." Since this principle expresses an attitude of generalized benevolence, we can expect to find a good deal of sympathy for it among the sort of people with whom it would be profitable to carry on a discussion about ethics. It may therefore be possible to obtain wide assent to the principle, provided that we can develop its implications in a clear and consistent manner and that we can show that certain common objections to utilitarianism are not as valid as they are supposed to be. We have already seen that certain objections, based on "the common moral consciousness," fail because they are valid only against descriptive utilitarianism and not against normative utilitarianism.

Determining consequences. Utilitarianism would be an easier doctrine to state if we could assume that we could always tell with certainty what all the consequences of various possible actions would be, and if we could assume that very remote consequences need not be taken into account. In applying the utilitarian principle, we would simply have to envisage two or more sets of consequences extending into the future, and ask ourselves, as sympathetic and benevolent men, which of these we would prefer. There would be no need for any calculation or for any summation of pleasures. We would simply have to compare two or more possible total situations. Sometimes, indeed, the postulate that we need not consider very remote situations will not be necessary. For example, if it be admitted that, on the whole, people are more happy than not, a man and woman who are left alive as sole representatives of the human race after some atomic holocaust could, as utilitarians, decide to have children in the hope that the world would once more be populated indefinitely far into the future. This is because although the generations will extend indefinitely far into the future, there is reason to believe that each generation will

be happy rather than unhappy, while if no children are had, there will be no succeeding generations at all, and so no possibility of happiness accruing in the future. In normal cases, however, we do need to assume that remote consequences can be left out of account. Surely, however, this is a plausible assumption, for on the whole, the goodness and badness of very remote consequences are likely to cancel out. In any case, if this assumption cannot be made, also difficulties will arise for many deontological systems (for example, the system of W. D. Ross), which allow beneficence as one principle among others.

Unfortunately, however, we do not know with certainty what the various possible consequences of our actions will be. This uncertainty would not be so bad provided we could assign numerical probabilities to the various consequences. We could then still employ a method similar to that of envisaging total consequences. A very simplified example may make this clear. Suppose that the only relevant consequences are, on the one hand, a 3/5 probability of Smith's being in some state *S*, and on the other hand, if we do an alternative action, a 2/7 probability of Jones's being in some state *T*. We simply envisage 21 people just like Smith in state *S* as against 10 people just like Jones in state *T*. It should be evident how, in theory at least, this method could be extended to more complex cases. However, numerical probabilities can rarely be assigned to possible future events, and the utilitarian is reduced to an intuitive weighting of various consequences with their probabilities. It is impossible to justify such intuitions rationally, and we have here a serious weakness in utilitarianism. It is true that this weakness also extends to prudential decisions, and most people think that they can make prudential decisions with some rationality. But this is not of much help, since in propounding a normative system we are concerned with what we ought to think, not with how we do think. Utilitarianism is therefore badly in need of support from a theory whereby, at least roughly or in principle, numerical probabilities could be assigned to all types of events.

THE PLACE OF RULES IN ACT UTILITARIANISM. Even the act utilitarian cannot always be weighing up consequences. He must often act habitually or in accordance with rough rules of thumb. However, this does not affect the value of the act-utilitarian principle, which is put forward as a criterion of rational choice. When we act habitually we do not exercise a rational choice, and the utilitarian criterion is not operative. It is, of course, operative when we are deciding, on act-utilitarian principles, the habits or rules of thumb to which we should or should not school ourselves. The act utilitarian knows that he would go mad if he deliberated on every trivial issue, and that if he did not go mad he would at least slow up his responses so much that he would miss many opportunities for probably doing good. He may also school himself to act habitually because he may think that if he deliberated in various concrete situations, his reasoning would be distorted by a selfish bias.

APPLICATIONS

UTILITARIANISM AND GAME THEORY. The act utilitarian will of course use as some of his premises propositions about how other members of the community are likely to act. For example, if certain individuals are adherents of a deontological morality, their actions will tend to be made predictable and their behavior will constitute valuable information for the act utilitarian when he is planning his own actions. Thus, an act utilitarian who has something important to do with his time may be wise to abstain from voting in an election (assuming that there is no legal compulsion to vote), for he will reflect that most people will in fact go to vote and that elections are very rarely decided by a single vote.

But how should the act utilitarian reason if he lives in a society in which everyone else is an act utilitarian? He needs information about what other people will do, but since they reason as he does, what they will do depends on what they think he will do. There is a circularity in the situation that can be resolved only by the technique of the theory of games.

Moral philosophers have commonly failed to give the correct solution to this sort of question. In the case in which the act utilitarian is asking whether he should do an action *A* or not do it, moral philosophers have commonly envisaged only two possibilities: Either everyone does *A* or no one does *A*. They have failed to notice the possibility of what, in the theory of games, is called a mixed strategy. Each act utilitarian can give himself a probability *p* of doing *A*. Thus, in the case of the voting, each act utilitarian might toss pennies or dice in such a way as to give himself a certain probability *p* of voting, so that the best possible proportion of people will turn up to vote and a small proportion will be free to do other things. The calculation of *p* is a simple maximization problem, provided that we know numerical values of the probabilities and numerical values of the various consequences of alternative actions. Of course, this is unlikely to be the case, and the question of a mixed strategy is usually more of theoretical than of practical importance. Moreover, in very many important cases the effect of even a few people acting in a certain way is, in practice, so dis-

astrous that the probability we should give ourselves of acting in this way may be so small that we may as well say, like the rule utilitarians, that we would never do it.

UTILITARIANISM AND PRAISE AND BLAME. Not only do we use moral language to deliberate about what we should do, but we also use moral language to praise people and blame them. Suppose that we use the words "good action" and "bad action" to convey praise and blame, and "right action" and "wrong action" to evaluate what ought to be done. On act-utilitarian principles, then, a right action is one that produces the best consequences. A good action is one that should be praised. Normally we will wish to praise right actions and blame wrong ones, but this is not invariably the case. As Sidgwick has pointed out very clearly, when, as utilitarians, we assess agents and motives as good or bad, the question at issue is not the utility of the actions but the utility of praise or blame of them. Suppose that the only way in which a soldier can save the lives of half a dozen companions is by throwing himself upon a grenade that is about to explode, thus taking upon himself the full impact of the blast and inevitably being killed. The act utilitarian would have to say that the soldier ought to sacrifice himself in this way. Nevertheless, he would not censure the soldier or say that he had acted from a bad motive if he had refrained from this heroic act and his companions had been killed. There is nothing to be gained by censuring someone for lack of extraordinary heroism, and probably much harm in doing so. The act utilitarian should say that the soldier's motive was not a bad one, although his action was as a matter of fact a wrong one.

Consider a case in which an action, normally of trivial import, happens to have very unfortunate consequences. A man with a head cold goes to the office, instead of nursing his illness at home. He is visited by an eminent statesman, who catches the cold and, in consequence, is not quite at his best in carrying out some delicate negotiations. These negotiations fail just by a hairsbreadth, whereas if the statesman had been fully fit they would have succeeded. In consequence, thousands of people die from starvation, a misfortune that would have been avoided if the negotiations had succeeded. These deaths from starvation would therefore not have occurred if the man with a head cold had not gone to his office in an infectious state. Someone may be tempted to argue as follows: "Surely it is not a very wrong action to go to the office suffering from a head cold. In some cases, where important work has to be done, it may even be praiseworthy. But in this case the action had very bad consequences, and so the utilitarian must say that it is very wrong. There must therefore be something wrong with utilitarianism." The utilitarian must reply that the objector is confusing two things, the rightness or wrongness of an action and the praiseworthiness or blameworthiness of it. The action, he can consistently say, was very wrong, but it was not very bad: That is, it ought not to be blamed very much, if at all. If we blame it, we are concerned with the utility of discouraging similar actions on the part of other people, and since going to the office with a head cold is not normally productive of very bad consequences, this action, although in fact very wrong, was not a very bad or blameworthy one.

Another reason why utility (or rightness) of an action does not always coincide with utility of praise or blame of it, and hence with its goodness or badness, is that, as Sidgwick pointed out, although universal benevolence, from the act-utilitarian view, is the ultimate standard of right and wrong, it is not necessarily the best or most useful motive for action. For example, although family affection may not always act in the same direction as generalized benevolence, it very frequently does so, and is a much more powerful motive than the latter. The act utilitarian may well think it useful to praise an action done from family affection in order to strengthen and encourage this motive, even when in fact the action was not generally beneficial.

Similarly, members of a community may act according to some traditional code of rules and may be likely to become simply amoral if a premature attempt is made to convert them to utilitarianism. A utilitarian may well think, therefore, that he ought to support this traditional nonutilitarian code of morals, if its general tendency is at all beneficent. He may therefore apportion praise and blame among members of this community according to whether their actions are in conformity with this code, and not according to whether they are right or wrong from the utilitarian standpoint. The relations between act utilitarianism and the traditional morality of a community in which an act utilitarian may find himself are very complex, and have been quite thoroughly investigated by Sidgwick.

See also Aristippus of Cyrene; Baier, Kurt; Bentham, Jeremy; Brandt, R. B.; Consequentialism; Deontological Ethics; Egoism and Altruism; Epicurus; Game Theory; Godwin, William; Good, The; Happiness; Hare, Richard M.; Hedonism; Helvétius, Claude-Adrien; Hume, David; Kant, Immanuel; Metaethics; Mill, John Stuart; Moore, George Edward; Paley, William; Pleasure; Punishment; Rashdall, Hastings; Rawls, John;

Ross, William David; Sidgwick, Henry; Stevenson, Charles L.

Bibliography

HISTORY OF UTILITARIANISM

For the history of utilitarianism, see especially H. Sidgwick, *Outlines of the History of Ethics* (London, 1946), with an additional chapter by Alban G. Widgery; Ernest Albee, *A History of English Utilitarianism* (New York: Macmillan, 1902); and Leslie Stephen, *The English Utilitarians* (London: Duckworth, 1900), which contains references to the works of Tucker and Paley. The works of Helvétius are not readily accessible, but a list of various editions can be found in the bibliography of Ian Cumming, *Helvétius, His Life and Place in the History of Educational Thought* (London: Routledge and Kegan Paul, 1955).

MAJOR UTILITARIANS

For works by Bentham, see especially his *Fragment on Government and Introduction to the Principles of Morals and Legislation,* edited by Wilfred Harrison (Oxford, 1948), and his *Deontology,* edited by John Bowring (London and Edinburgh, 1843). A very scholarly modern work on Bentham is David Baumgardt, *Bentham and the Ethics of Today, with Bentham Manuscripts—Hitherto Unpublished* (Princeton, NJ: Princeton University Press, 1952). For works by J. S. Mill, see *Utilitarianism, On Liberty, Essay on Bentham, Together with Selected Writings of Jeremy Bentham and John Austin,* edited with an introduction by Mary Warnock (Malden, MA: Blackwell, 2003). M. St. J. Packe, *The Life of John Stuart Mill* (London: Secker and Warburg, 1954), gives evidence on p. 53 (footnote) that Bentham used the word *utilitarian* before Mill. An interesting discussion of Mill's ethical principles is given in Karl Britton, "Utilitarianism: the Appeal to a First Principle," in *PAS* 60 (1959–1960): 141–154. Britton is also the author of an account of Mill's philosophy, *John Stuart Mill* (London: Penguin, 1953). A nineteenth-century criticism of utilitarianism is J. Grote, *An Examination of the Utilitarian Philosophy* (Cambridge, U.K., 1870). For Sidgwick, see his *Methods of Ethics,* 7th ed. (London, 1907), which is discussed by C. D. Broad in *Five Types of Ethical Theory* (London: Kegan Paul, 1930). Godwin's moral philosophy is critically expounded by D. H. Monro in his *Godwin's Moral Philosophy* (London: Oxford University Press, 1953). Ideal utilitarianism may be studied in G. E. Moore, *Principia Ethica* (Cambridge, U.K.: Cambridge University Press, 1903) and *Ethics* (London: Williams and Norgate, 1912), and in H. Rashdall, *Theory of Good and Evil* (Oxford: Clarendon Press, 1907). A sympathetic discussion of utilitarianism will be found in A. J. Ayer's essay "The Principle of Utility," in his *Philosophical Essays* (London: Macmillan, 1954).

RULE AND ACT UTILITARIANISM

Many modern writers have espoused views that can be described as rule utilitarianism. See especially Stephen Toulmin, *The Place of Reason in Ethics* (Cambridge, U.K.: Cambridge University Press, 1951); P. H. Nowell-Smith, *Ethics* (London: Penguin, 1954); J. Rawls, "Two Concepts of Rules," in *Philosophical Review* 64 (1955): 3–32; K. E. M. Baier, *The Moral Point of View* (Ithaca, NY: Cornell University Press, 1958); M. G. Singer, *Generalization in Ethics* (New York: Knopf, 1961). J. O. Urmson, "The Interpretation of the Philosophy of J. S. Mill," in *Philosophical Quarterly* 3 (1953): 33–39, interprets Mill as a rule utilitarian. His view is contested by J. D. Mabbott, "Interpretation of Mill's *Utilitarianism,*" in *Philosophical Quarterly* 6 (1956): 115–120. The issue between act and rule utilitarianism is in effect discussed by A. C. Ewing, "What Would Happen If Everyone Acted Like Me?" in *Philosophy* 28 (1953): 16–29, and by A. K. Stout, "But Suppose Everybody Did the Same?" in *Australasian Journal of Philosophy* 32 (1954): 1–29. J. J. C. Smart, "Extreme and Restricted Utilitarianism," in *Philosophical Quarterly* 6 (1956): 344–354, is a defense of act as against rule utilitarianism. The terms *extreme* and *restricted* are used here instead of the more appropriate words *act* and *rule.* These last were introduced by R. B. Brandt in his *Ethical Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1959), which contains good discussions of act and rule utilitarianism. Brandt has also developed a complex and subtle form of rule utilitarianism in his paper "In Search of a Credible Form of Rule Utilitarianism," in *Morality and the Language of Conduct,* edited by George Nakhnikian and Héctor-Neri Castañeda (Detroit: Wayne State University Press, 1953). See also H. D. Aiken, "The Levels of Moral Discourse," in *Ethics* 62 (1952): 235–248. Rule utilitarianism is criticized by H. J. McCloskey, "An Examination of Restricted Utilitarianism," in *Philosophical Review* 66 (1957): 466–485. A Kantian type of rule utilitarianism is presented by R. F. Harrod, "Utilitarianism Revised," in *Mind* 45 (1936): 137–156, and J. C. Harsanyi, "Ethics in Terms of Hypothetical Imperatives," in *Mind* 67 (1958): 305–316. Jonathan Harrison's article "Utilitarianism, Universalisation and Our Duty to Be Just," in *PAS* 53 (1952–1953): 105–134, discusses important issues and includes a criticism of Harrod.

NEGATIVE UTILITARIANISM

K. R. Popper, *The Open Society and Its Enemies,* 3rd ed., Vol. 1 (London, 1957), Ch. 5, note 6, has put forward some considerations that suggest the possibility of expressing utilitarianism in terms of the prevention of misery rather than in terms of the promotion of happiness, although Popper himself does not seem to be a utilitarian. Such a "negative utilitarianism" has been criticized by R. N. Smart, "Negative Utilitarianism," in *Mind* 67 (1958): 542–543.

EXPOSITIONS AND CRITICISMS

An exposition of act utilitarianism as a normative system is given by J. J. C. Smart, *An Outline of a System of Utilitarian Ethics* (Melbourne: Melbourne University Press and University of Adelaide, 1961). An introductory textbook from an act-utilitarian point of view is C. A. Baylis, *Ethics, the Principles of Wise Choice* (New York: Holt, 1958). There are useful chapters on utilitarianism in John Hospers, *Human Conduct* (New York: Harcourt Brace, 1961). The inconsistency of an egoistic utilitarianism is pointed out by B. H. Medlin, "Ultimate Principles and Ethical Egoism," in *Australasian Journal of Philosophy* 35 (1957): 111–118. J. Rawls, in his article "Justice as Fairness," in *Philosophical Review* 67 (1958): 164–194, holds that one must never act solely to increase the general happiness if in so doing one makes any particular person unhappy. I. M. Crombie, in his article "Social Clockwork and Utilitarian Morality" in

Christian Faith and Communist Faith, edited by D. M. Mackinnon (London: Macmillan, 1953), suggests that a utilitarian could accuse the deontologist of a sort of idolatrous attitude toward rules. Another valuable point made in this article is that utilitarianism is in a certain way a self-correcting doctrine. See also A. I. Melden, "Two Comments on Utilitarianism," in *Philosophical Review* 60 (1951): 508–524; H. W. Schneider, "Obligations and the Pursuit of Happiness," in *Philosophical Review* 61 (1952): 312–319; and S. M. Brown Jr., "Utilitarianism and Moral Obligation," ibid., 299–311, which led to comments by C. A. Baylis and John Ladd, ibid., 320–330. J. O. Urmson, "Saints and Heroes," in *Essays in Moral Philosophy,* edited by A. I. Melden (Seattle: University of Washington Press, 1958), makes some distinctions that he tentatively suggests may be accommodated more easily by a utilitarian than by a nonutilitarian ethics. A pioneering application of the theory of games to problems of moral philosophy is to be found in R. B. Braithwaite, *Theory of Games as a Tool for the Moral Philosopher* (Cambridge, U.K.: Cambridge University Press, 1955).

METAETHICAL THEORIES

For the noncognitivist theories of metaethics of C. L. Stevenson, R. M. Hare, and P. H. Nowell-Smith, see C. L. Stevenson, *Ethics and Language* (New Haven, CT: Yale University Press, 1944); R. M. Hare, *The Language of Morals* (Oxford: Clarendon Press, 1952); and P. H. Nowell-Smith, *Ethics* (London: Penguin, 1954). Hare's sequel, *Freedom and Reason* (Oxford: Clarendon Press, 1963), contains an interesting chapter on utilitarianism.

J. J. C. Smart (1967)

UTILITARIANISM [ADDENDUM]

J. J. C. Smart's advocacy of utilitarianism has been perhaps the most influential since Henry Sidgwick's nearly a century earlier. Nevertheless, there have been some significant developments since Smart's work, outlined here.

Fundamental to Smart's approach is his thesis that there can be no proof of ultimate normative moral principles. In this respect, ultimate normative principles, Smart thinks, are unlike many other kinds of claims. For example, some claims are true because of the definitions of the terms in them ("Bachelors are unmarried"). And, setting aside worries about induction, we observe that some claims are proven false by empirical investigation ("Drinking caffeinated coffee makes you sleepy"), and that other claims are confirmable by empirical investigation ("sugar dissolves in boiling water"). Ultimate normative principles, however, are different. They are not true by definition. They are neither refutable nor confirmable by purely empirical investigation. And ultimate normative principles are *basic*, that is, are not derivable from something deeper. So they cannot be proved, Smart thinks.

Indeed, in Smart's view, to endorse some ultimate moral principle is not to express a cognition, that is, a belief. Smart held that moral judgments essentially express something noncognitive, such as a sentiment, an attitude, or a commitment. So Smart was a noncognitivist in ethics. Yet Smart did think that there is a way of supporting ultimate normative principles—by appeal to generalized benevolence. He meant that we might find certain proposed moral principles attractive from the point of view of impartial concern for all. This point of view accords any benefit or harm to any individual the same weight as it accords to the same size benefit or harm to any other individual. Smart's idea is that, from the point of view of impartial benevolence, utilitarianism is virtually irresistible.

Four years after the publication of Smart's entry in the first edition of this encyclopedia, John Rawls published *A Theory of Justice* (1971). Many of the theses in Rawls's book had been put forward by him or others before, but his book solidified support for many of these theses. The net effect was that Rawls's book changed the landscape in moral and political philosophy.

Rawls's most pervasive influence was in what we might call moral methodology. He championed the search for "reflective equilibrium" between the specific moral judgments that we make after careful reflection and the general moral principles that we affirm after careful reflection. We seek general principles that are consistent with and provide some justification for our more specific judgments. At least to some extent, we are willing to adjust our specific judgments to fit with the best general principles we can find. When specific moral judgments and general principles cohere, we have reflective equilibrium.

This picture of theorizing in normative ethics has been widely accepted by moral philosophers, though the picture has been interpreted in a variety of different ways. One point about it that seems incontrovertible is that achieving reflective equilibrium between one's specific moral judgments and one's more general moral principles hardly proves that the judgments and principles are true. A coherent set of beliefs or commitments can be deeply misguided rather than true. Coherence is not a sufficient condition of truth.

But the consistency of any two beliefs with one another is a necessary condition of their both being true.

So if moral commitments are beliefs, they had better be consistent with one another. Even if moral commitments are not really beliefs but instead are noncognitive states, they are faulty if inconsistent.

As Smart noted, act utilitarianism conflicts with many popular moral commitments. Suppose that we would produce slightly greater net aggregate welfare by, for example, breaking a promise or stealing or framing an innocent person than we would by not doing such a thing. Act utilitarians must favor breaking the promise or stealing or framing an innocent person in such circumstances. But here act utilitarianism seems, to most people, deeply mistaken.

In his entry on this topic in the first edition, Smart replied with a rhetorical question: "Would the benevolent and sympathetic persons to whom we conceive ourselves to be appealing be happy about preferring abstract conformity with an ethical rule, such as 'Keep promises,' to preventing avoidable misery of his fellow creatures?" Part of the rhetorical force of this question comes from the implication that we are here choosing between conforming to a rule when this would benefit no one and breaking a rule when this would prevent avoidable misery.

If such are indeed the circumstances, many people would agree with Smart that it would be right to break the promise, because of the following argument. An action is wrong only if it harms someone. Breaking a promise would, in some cases, harm no one. So, in these cases, breaking a promise could not be wrong. If one could prevent avoidable misery in a way that would not be wrong, then one would be morally required to do so. Thus, in the case posed by Smart, one would be morally required to break the promise.

There are various problems with this argument. For example, there are problems with its first premise, that an action is morally wrong only if it harms someone. Suppose that I broke a promise to you, or stole from you, in a way that harmed no one. Could not such an act be wrong despite the fact that no one was harmed?

Rather than pursue that question, let us turn to the more important question: What about cases where breaking a rule would benefit some people but harm someone else? For instance, we have a rule against framing people, particularly innocent people. Breaking this rule would harm the person framed, but others might benefit. Consider a (very unlikely) situation where for some reason the harm caused to the innocent person is less than the aggregate benefit to others. Now suppose that framing the innocent person would produce at least a little greater net welfare than any alternative possible act. In any such case, act utilitarianism licenses framing the innocent person under such circumstances. Some philosophers try to defend act utilitarianism here by arguing that framing an innocent person could be the lesser of two evils in catastrophic cases. Suppose that the only way of preventing hundreds or thousands or even millions of innocent people from dying is to frame some innocent person for some crime. In such catastrophic cases, many people would admit that morality would reluctantly allow, or even require, framing the innocent person.

But those who admit this need hardly go as far as act utilitarianism does. Act utilitarianism holds that breaking a promise or stealing or injuring or even framing an innocent person is morally right not only when such an act would prevent a catastrophe but also when it would produce only a little greater net aggregate welfare than not performing the act. This act-utilitarian claim is terribly counterintuitive.

This is not the only place where act utilitarianism departs dramatically from our intuitive views. Consider the act-utilitarian view of praise and blame. As Smart explained, act utilitarians since Sidgwick have held that an action is to be praised if and only if praising it maximizes utility, and blamed if and only if blaming it maximizes utility. Act utilitarianism also holds that an action, which might be praised or blamed, is morally right if and only if the action maximizes utility. So what about cases where utility would be maximized if blame were directed at an action that itself maximized utility? Act utilitarianism holds that in such cases the right action should be blamed. Likewise, act utilitarianism can hold that a wrong action (one that failed to maximize utility) should be praised if praising the wrong action would for some reason maximize utility.

Again, these are counterintuitive claims. Common moral awareness sees a much tighter connection between an action's being morally right and its being praiseworthy than act utilitarianism allows. Equally, common moral awareness sees a much tighter connection between an action's being morally wrong and its being blameworthy than act utilitarianism allows.

Another way in which act utilitarianism is counterintuitive has come to light as a result of an article published by Peter Singer in 1971, the same year in which Rawls's book was published. On highly plausible assumptions about the disparity in wealth in the world, the needs of the billion worst off, the diminishing marginal utility of wealth, and the unwillingness of most others to contribute significantly to reducing world poverty, act utili-

tarianism seems to demand nothing less than huge sacrifices from the average individual in a relatively wealthy country for the sake of helping the poorest in the world. An average individual in a relatively wealthy country could save someone's life by making a small contribution to one of the best relief organizations. That is true of the first contribution an individual might make, but also of each of very many further contributions. True, if most average individuals in wealthy countries made personal contributions to the most efficient poverty-relief organizations, each of them might not need to contribute more than a few percent of annual income. But in fact the average person in the relatively wealthy country knows that most others will not give anything at all. In this context, act utilitarianism makes extreme demands on the average person in relatively wealthy countries.

How far do these demands go? Act utilitarianism requires you to keep making contributions until you reach a point where further contributions on your part fail to maximize net aggregate utility. The most obvious way further contributions could fail to maximize net aggregate utility would be for the harm to you and your dependents to be at least as great as the benefits produced for the recipients of aid. Another way in which further contributions from you could fail to maximize net aggregate utility would be for those contributions to undermine your capacity to make more contributions later. In any case, act utilitarianism seems to require enormous sacrifices from the average person in a relatively wealthy country to rescue the needy of the world. A very high level of personal sacrifice for such a worthy cause is obviously admirable. And *some* level of personal sacrifice for such a worthy cause does seem morally *required*. But the level of sacrifice that is required by morality seems, intuitively, nowhere near as high as act utilitarianism claims it is.

Much of the work on utilitarianism since the publication of Rawls's book has focused on whether any version of the theory has intuitively acceptable implications. For example, Derek Parfit (1984) has sought a utilitarian principle with intuitively acceptable implications about how large the population should be. Again, Fred Feldman (1997) and Shelly Kagan (1999) have suggested supplementing act utilitarianism with a principle of desert: It matters not just how much net benefit is produced but also that benefits go to the deserving rather than to the undeserving. A difficulty with the latter approach, however, is that it appears to presuppose and leave unexplained principles about desert that a different utilitarian approach could easily explain. Why should only those who do certain kinds of acts be punished? Why should those who do certain other kinds of things be praised or otherwise rewarded? Because significant benefits will result from social practices of punishing those who do certain kinds of act and from social practices of praising and rewarding those who do certain other kinds of act. This is a rule-utilitarian explanation.

Indeed, if we are looking for a version of utilitarianism to be in reflective equilibrium with the moral judgments we make after careful reflection, act utilitarianism seems quite inferior to rule utilitarianism. Rule utilitarianism claims that an act is wrong if it is forbidden by rules whose internalization would produce the greatest (expected) utility. Rules that forbid promise breaking, stealing, lying (including framing the innocent), physical attack, and so on, produce greater utility than rules that allow these kinds of acts. So rule utilitarianism has no difficulty explaining why these kinds of acts are wrong.

Smart did not agree. He suggested that most people believe that a practice of framing the innocent would be wrong even in a possible world in which such a practice, as a rule, would maximize utility. But is there an empirically possible world in which not just one instance, but a general *practice*, of framing the innocent would maximize utility? Surely there is not *if* any such practice would have to be publicly known. For if a practice of framing the innocent became publicly known, public confidence in the police and courts would quickly dissipate, with terrible consequences for social order.

This point brings out an important difference between act utilitarianism and rule utilitarianism. Again as Sidgwick (1907) noticed, act utilitarianism might endorse what he called an "esoteric morality," that is, a principle determining right and wrong whose correctness should be known about by less than everyone, perhaps even by only a few. In contrast, rule utilitarians are hostile to the idea of secret rules determining what people are or are not morally allowed to do. As John Harsanyi (1982, 1993) has stressed, rule utilitarianism, in evaluating any proposed code of rules, attaches great importance to the expectations and incentives that would follow from public knowledge of the social acceptance of the rules.

Following Richard Brandt (1979), rule utilitarians have also stressed that the costs of getting a code of rules internalized by new generations of agents must be counted as part of the cost/benefit assessment of that code. The focus here is on new generations so as not to let the cost/benefit assessment of a code be influenced by which rules a society happens to accept already.

As rules become more numerous and complicated and as they demand more self-sacrifice, the costs of getting new generations to internalize them increase. At some point, the costs of yet more rules, or of greater complication, outweigh the benefits. Likewise, at some point the costs of getting new generations to internalize yet more demanding rules about helping the world's needy will outweigh the benefits of having agents willing to make the sacrifices necessary to help. So there are compelling rule-utilitarian reasons to restrict the number, complexity, and demandingness of rules. These restrictions help rule utilitarianism generate rules that accord with our intuitive views and conflict with act-utilitarian demands.

For about thirty years after the publication of Smart's entry in the first edition of this encyclopedia, most philosophers were persuaded by his objection that rule utilitarianism is fatally flawed. Smart wrote, "The rule utilitarian also appeals ultimately to feelings of generalized benevolence. … He is then inconsistent if he prescribes that we should obey a rule (even a generally beneficial rule) in those cases in which he knows that it will *not* be most beneficial to obey it." If what ultimately matters is how well individuals' lives go, why follow a rule when breaking it would maximize how well individuals' lives go? This objection to rule utilitarianism can be formulated as follows:

Premise 1. Rule utilitarians' overarching aim is to maximize utility.

Premise 2. Rule utilitarians endorse what conflicts with that aim, since their theory requires us to follow certain rules even when following those rules would not maximize utility.

Premise 3. It is inconsistent to maintain an overarching aim and then to endorse what conflicts with that aim.

Conclusion. Rule utilitarians are inconsistent.

One rule-utilitarian response to this objection is to reject its second premise. In other words, this response admits that in cases where following some generally beneficial rule would not maximize utility, the rule should not be followed. The suggestion might be that rule utilitarianism itself has a rule for abnormal cases where following the normal rules would not maximize utility. This rule might be, "In such cases, do whatever will maximize utility." But this defense of rule utilitarianism threatens to collapse rule utilitarianism into act utilitarianism. Such a collapse would be fatal to rule utilitarianism. For if rule utilitarianism ends up endorsing the very same acts that act utilitarianism endorses, why bother with rule utilitarianism, since it is the more complicated of these two theories?

A better way to defend rule utilitarianism is to attack the first premise of the objection, by denying that rule utilitarians must have maximizing utility as their overarching aim. Consider moral agents of which the following statements are true:

- Their fundamental moral motivation is to act in ways that are impartially justifiable.
- They believe that acting on impartially justifiable rules is impartially justifiable.
- They believe that rule utilitarianism is the best theory of impartially justifiable rules.

Agents with this psychological profile are rule utilitarians, but these agents do not have maximizing utility as their overarching aim. So rule-utilitarian agents need not be inconsistent.

Even if rule-utilitarian agents need not be inconsistent, is their theory itself nevertheless inconsistent? Rule utilitarianism consists of two principles: the principle that rules should be selected in terms of their expected utility, and the principle that the rules thus selected determine what kinds of acts are morally wrong. These two principles do not conflict with one another. And neither of them expresses an overarching aim to maximize utility. The theory simply does not contain that overarching aim. Thus, rule utilitarianism can consistently require us to follow certain rules even on occasions when following these rules would not maximize utility.

The ultimate justification for rule utilitarianism may come from its ability to provide general principles that accord with our more specific moral judgments. Admittedly, achieving reflective equilibrium between our principles and our more specific moral judgments cannot establish that the principles or judgments are true. Nevertheless, if rule utilitarianism is attractive in its own right, and if it underwrites and ties together all our more specific moral judgments, and it if does this more securely than any rival general principle does, we have good grounds for accepting rule utilitarianism.

But does rule utilitarianism succeed in providing a general principle that underwrites and ties together all our more specific moral judgments? The answer is uncertain. With Smart's objections to rule utilitarianism now answered, the theory is again under development. However, the possibility remains that the theory will be refuted by counterexample. This will happen if it is dis-

covered that the implications of rule utilitarianism for some kind of case are just too counterintuitive, i.e., if they conflict sharply with our very confident convictions about what is morally required in that kind of case.

One area of persistent controversy is over what constitutes the good that the rules should maximize. Many prominent utilitarians have held that utility is a matter exclusively of welfare and that welfare consists exclusively of net pleasure. But many philosophers have held that there is more to welfare than net pleasure. Some, for example, have held that making significant achievements, obtaining important knowledge, and having deep friendships constitute benefits to the individual, that is, constitute additions to the individual's welfare, beyond whatever pleasure the individual directly or indirectly gets from these things.

There is also controversy about whether rule utilitarianism is right to evaluate rules purely in terms of how much aggregate welfare would result. Many philosophers hold that not only the aggregate amount but also its distribution matters. So, for example, many philosophers hold that an outcome containing greater aggregate welfare might be less good than an alternative containing less aggregate welfare if the worst-off individuals in the outcome with less aggregate welfare are less badly off than are the worst-off in the outcome with greater aggregate welfare. For the sake of illustration, consider an artificially simple example in which the world consists of only two groups of people and only two codes of rules to compare:

First code of rules

	Units of welfare	
	Per person	Per group
10,000 people in group A	1	10,000
100,000 people in group B	10	1,000,000
Total welfare for both groups		1,010,000

Second code of rules

	Units of welfare	
	Per person	Per group
10,000 people in group A	8	80,000
100,000 people in group B	9	900,000
Total welfare for both groups		980,000

Such examples exert strong pressure on us to accept that benefits to worse-off individuals should be accorded more importance than the same-size benefits to the better-off. Philosophers who accept this often call themselves rule consequentialists instead of rule utilitarians.

An as yet unresolved difficulty is whether rule utilitarianism retains its fundamental impartiality if, in the cost/benefit assessment of rules, benefits to the worst-off are accorded more importance than the same-size benefits to the better-off. Certainly, one of the chief attractions of utilitarianism is its fundamental impartiality. This is not something to be jettisoned lightly.

See also Bentham, Jeremy; Brandt, R. B.; Consequentialism; Mill, John Stuart; Sidgwick, Henry; Smart, John Jamieson Carswell; Teleological Ethics.

Bibliography

Adams, R. M. "Motive Utilitarianism." *Journal of Philosophy* 73 (1976): 467–481.

Bailey, James Wood. *Utilitarianism, Institutions, and Justice*. New York: Oxford University Press, 1997.

Bales, R. E. "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?" *American Philosophical Quarterly* 8 (1971): 257–265.

Brandt, Richard B. "Some Merits of One Form of Rule-Utilitarianism" (1967). In his *Morality, Utilitarianism, and Rights*, 111–136. New York: Cambridge University Press, 1992.

Brandt, Richard B. *A Theory of the Good and the Right*. Oxford, U.K.: Oxford University Press, 1979.

Brink, David O. *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press, 1989.

Broome, John. *Weighing Goods*. Oxford, U.K.: Blackwell Publishers, 1991.

Crisp, Roger. *Mill on Utilitarianism*. London: Routledge, 1997.

Cullity, Garrett. *The Demands of Affluence*. Oxford, U.K.: Oxford University Press, 1994.

Feldman, Fred. *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*. New York: Cambridge University Press, 1997.

Foot, Philippa. "Utilitarianism and the Virtues." *Mind* 94 (1985): 196–209.

Griffin, James. *Well-Being: Its Meaning, Method, and Moral Importance*. Oxford: Oxford University Press, 1987.

Hardin, Russell. *Morality within the Limits of Reason*. Chicago: University of Chicago Press, 1988.

Hare, R. M. *Moral Thinking*. Oxford, U.K.: Oxford University Press, 1981.

Harsanyi, John. *Essays on Ethics, Social Behaviour, and Scientific Explanation*. Dordrecht, Netherlands: Reidel, 1976.

Harsanyi, John. "Expectation Effects, Individual Utilities, and Rational Desires." In *Rationality, Rules, and Utility: New Essays on the Moral Philosophy of Richard Brandt*, edited by Brad Hooker, 115–126. Boulder, CO: Westview Press, 1993.

Harsanyi, John. "Morality and the Theory of Rational Behaviour." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams, 39–62. Cambridge, U.K.: Cambridge University Press, 1982.

Harsanyi, John. "Rule Utilitarianism and Decision Theory." *Erkenntnis* 11 (1977): 25–53.

Hooker, Brad. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford, U.K.: Oxford University Press, 2000.

Hooker, Brad, Elinor Mason, and Dale E. Miller, eds. *Morality, Rules, and Consequences*. Lanham, MD: Rowman and Littlefield, 2000.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101 (1991): 461–482.

Jamieson, Dale, ed. *Singer and His Critics*. Oxford, U.K.: Blackwell Publishers, 1999.

Johnson, Conrad. *Moral Legislation*. New York: Cambridge University Press, 1991.

Kagan, Shelly. *The Limits of Morality*. Oxford, U.K.: Oxford University Press, 1989.

Kagan, Shelly. "Equality and Desert." In *What Do We Deserve?* edited by Louis Pojman and Owen McLeod, 298–314. New York: Oxford University Press, 1999.

Kelly, Paul J. *Utilitarianism and Distributive Justice: Bentham and the Civil Law*. New York: Oxford University Press, 1990.

Mackie, J. L. "The Disutility of Act-Utilitarianism." *Philosophical Quarterly* 23 (1973): 289–300.

Mackie, J. L. "Morality and the Retributive Emotions" (1982). In his *Persons and Values*, 206–219. New York: Oxford University Press, 1985.

Mason, Elinor. "Can an Indirect Consequentialist Be a Real Friend?" *Ethics* 108 (1998): 386–393.

McNaughton, David, and Piers Rawling. "On Defending Deontology." *Ratio* 11 (1998): 37–54.

Mulgan, Tim. *The Demands of Consequentialism*. Oxford, U.K.: Oxford University Press, 2001.

Murphy, Liam. *Moral Demands in Non-ideal Theory*. New York: Oxford University Press, 2000.

Parfit, Derek. "Equality and Priority." *Ratio* 10 (1997): 202–221.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Oxford University Press, 1984.

Pettit, Philip. "Consequentialism and Moral Psychology." *International Journal of Philosophical Studies* 2 (1994): 1–17.

Pettit, Philip. "The Consequentialist Perspective." In *Three Methods of Ethics*, edited by Marcia Baron, Philip Pettit, and Michael Slote, 92–174. Malden, MA: Blackwell Publishers, 1997.

Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13 (1984): 134–171.

Rawls, John. "Outline for a Decision Procedure in Ethics." *Philosophical Review* 60 (1951): 177–197.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Regan, Donald. *Utilitarianism and Co-operation*. Oxford, U.K.: Oxford University Press, 1980.

Scanlon, T. M. "Contractualism and Utilitarianism." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams. Cambridge, U.K.: Cambridge University Press, 1982.

Scheffler, Samuel. *The Rejection of Consequentialism*. Oxford, U.K.: Oxford University Press, 1982. Rev. and exp. ed., 1994.

Sen, Amartya, and Bernard Williams, eds. *Utilitarianism and Beyond*. Cambridge, U.K.: Cambridge University Press, 1982.

Shaw, William. *Contemporary Ethics: Taking Account of Utilitarianism*. Cambridge, MA: Blackwell Publishers, 1999.

Sidgwick, Henry. *Methods of Ethics*. 7th ed. London: Macmillan, 1907.

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1971): 229–243.

Smart, J. J. C. "Outline of a System of Utilitarian Ethics." In *Utilitarianism: For and Against*, by J. J. C. Smart and Bernard Williams, 3–74. Cambridge, U.K.: Cambridge University Press, 1973.

Streumer, Bart. "Can Consequentialism Cover Everything?" *Utilitas* 15 (2003): 237–247.

Sumner, L. W. *The Moral Foundation of Rights*. New York: Oxford University Press, 1987.

Sumner, L. W. *Welfare, Happiness, and Ethics*. Oxford, U.K.: Oxford University Press, 1996.

Unger, Peter. *Living High and Letting Die: Our Illusion of Innocence*. New York: Oxford University Press, 1996.

Williams, Bernard. "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, by J. J. C. Smart and B. Williams, 77–150. Cambridge, U.K.: Cambridge University Press, 1973.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Williams, Bernard. *Making Sense of Humanity and Other Philosophical Papers*. New York: Cambridge University Press, 1995.

Williams, Bernard. *Morality: An Introduction to Ethics*. New York: Harper and Row, 1972. See especially the final chapter.

Williams, Bernard. *Moral Luck*. Cambridge, U.K.: Cambridge University Press, 1981.

Wolf, Susan. "Moral Saints." *Journal of Philosophy* 79 (1982): 419–39.

Brad Hooker (2005)

UTOPIAS AND UTOPIANISM

The word *utopia* was invented by Thomas More, who published his famous *Utopia* (in Latin) in 1516. More coupled the Greek words *ou* (no, or not) and *topos* (place) to invent a name that has since passed into nearly universal currency. Further verbal play shows the close relation between utopia and *eutopia*, which means "the good [or happy] place." Through the succeeding centuries this double aspect has marked the core of utopian literature, which has employed the imaginary to project the ideal. (This is not to deny that More's own attitude towards the ideal society he imagined may well have been ambivalent.)

The words utopia and utopian, however, have been put to many uses besides the one suggested by More's book. Common to all uses is reference to either the imaginary or the ideal, or to both. But sometimes the words

are used as terms of derision and sometimes with a vagueness that robs them of any genuine usefulness. For example, a proposal that is farfetched or implausible is often condemned as utopian, whether or not the proposal has any idealistic content. In another, closely related pejorative use, utopian designates that which is unacceptably different from the customary or is radical in its demands. The connotation of impossibility or complete impracticality serves to discredit a threatening idealism. Similarly, daydreams and fantasies—psychologically driven and frequently bizarre expressions of private ideals—are called utopian, as if utopia were synonymous with deviant or deranged thinking. Even when the word is used without hostility, its coverage is enormously wide. Almost any expression of idealism—a view of a better life, a statement of basic political commitments, a plea for major reform in one or another sector of social life—can earn for itself the title utopian. Furthermore, all literary depictions of imaginary societies are called utopian, even if they are actually dystopias (bad places) that represent some totalitarian or fiendish horror, or are primarily futuristic speculations about technical and scientific possibilities that have no important connection to any idealism.

Much historical experience is reflected in this variety of usage. Indeed, the ways in which utopia (and utopian) are used can be symptomatic of prevailing attitudes towards social change in general. Nevertheless, clarity could be served if we see the core of utopianism as speculation, in whatever literary form, about ideal societies and ideal ways of life for whole populations, in which perfection, defined in accordance with common prepossessions and not merely personal predilections, is aimed at. Perfection is conceived of as harmony, the harmony of each person with himself or herself and with the rest of society. (If there must always be war, then utopian war is waged only against outsiders.) The tradition of utopian thought, in this core sense, is thus made up of elaborated ideas, images, and visions of social harmony.

Not discussed in this entry is dystopian speculation in many genres about the near or distant future, in which the condition of human life is degraded or deformed. In many cases, dystopia shares with utopia a total vision of an imaginary society; but a deliberate hell, not a planned heaven. What brings such a condition into being is zeal to maintain the power of the ruling group, not the project of human well-being. An oppressive and tenacious dictatorship holds sway. The most famous example is the sadistic dystopia of George Orwell's *Nineteen Eighty-Four* (1949).

INSPIRATION OF UTOPIANISM

The forerunners of the utopian tradition are the fables and myths of the golden age, the Garden of Eden, or some benign state of nature. These inherited stories, although of considerable antiquity, look back to some even more remote time in the misty past when harmony was allegedly the normal condition of life. Remorse or nostalgia is the usual accompaniment of these stories. Reality is not what it was, and worldly good sense holds that it is not likely that life will ever be again what it was—except perhaps through some divine intercession.

An uncontrived harmony characterized the primal felicity. Simple people led lives as simple as themselves; because human nature was undeveloped, they were easily made content. If the glories and pleasures of civilization were missing, so were its artificialities, corruptions, and physical and psychological sufferings. Whenever disgust or disenchantment with civilization has become acute, these old stories are retold in order to expose the faults of civilization. But apart from their role in this fundamentally self-conscious method of striking at an existing order, these stories are primarily interesting as repositories of the age-old longings of ordinary humanity. All that the world is not is summed up in short and supposedly seductive descriptions. Sometime long ago, when people were still in touch with their uncontaminated nature, they lived without domination, irrational inequality, scarcity, brutalizing labor, warfare, and the tortures of conscience; they lived without disharmony in any form. The good life is, in the first instance, defined by the absence of these things. Although fondness for an early simplicity may seem regressive—an ignoble attachment to a primitive and subhuman harmony—a principal impetus for utopianism is undeniably to be found here.

The later tradition not only fills out the picture that is only a sketch in the old myths, but more important, transcends the old myths. Whatever wistfulness for the golden age may be present, there is general agreement that primal harmony cannot be regained. The condition of harmony, which defines the good life, must be civilized. It may be more or less complex, more or less scientific, more or less abundant, more or less hierarchical, more or less free, but it must be organized and institutionally articulated (and almost always governed). Throughout the utopian tradition, reality is not defied to the extent of wishing away the idea of a settled society. In Plato's *Republic*, Socrates can dwell only briefly on the excellence of an amiably anarchic rusticity (no war, no class-strife, no politics, no meat-eating, no philosophy or sciences or high art) before his admirer Glaucon, with the

stinging phrase "city of pigs," forces him to turn his thoughts to the ideal city (the city of justice, which is founded on the initial unjust act of taking land from others). This transition can be taken as typical of utopianism as a whole.

VARIETIES OF UTOPIANISM

Even with a scrupulous adherence to the definition of utopianism as the succession of ideas, images, and visions of social harmony, the relevant texts are extremely numerous. The main types of utopias include, first, and most properly utopian, descriptions of imaginary societies held to be perfect or much closer to perfection than any society in the real world. They are located in the past, present, or future and are contained in treatise, novel, story, or poem with varying degrees of detailed specification and imaginative inventiveness.

The second type of utopia—closely allied to the first—is found in those works of political theory in which reflection on the fundamental questions of politics leads the theorist beyond politics to consider the social and cultural presuppositions of the ideal political order and the ends of life which that political order (placed in a certain social and cultural setting) can and should facilitate. Whereas the political theorist comes to the forms and purposes of all institutional life by way of political concerns and, as it were, incidentally, the intentionally utopian writer, with Thomas More as the model, works out from the start a comprehensive view of the ideal society and its way of life, a view in which political forms need not be of central importance. Some works of political theory—Plato's *Republic*, for example—so capably discuss nonpolitical matters that they fit into either category.

Those philosophies of history that culminate in a vision of achieved perfection are a third cluster of writings that are not imaginary projections of the ideal but display instead metaphysical optimism of a total kind. These are the theories of inevitable progress created by such thinkers as the Marquis de Condorcet, Herbert Spencer, and Karl Marx. Marx, for one, indignantly fought against inclusion in the utopian tradition because he presented himself as an antiutopian realist blessed with unique insight into the nature of the historical process and its necessary workings carried even to the future, not as an idealist preaching to the world an ahistorical conception of the ideal. For all that, others have taken his writings as belonging in the utopian tradition. Roughly, the same holds for Spencer and some other philosophers of history. No list of the major sources of utopian literature would be acceptable without theorists of inevitable progress.

Fourth are those works—sometimes called philosophical anthropologies—in which the writer attempts not only to isolate the instincts, traits, and capacities that are peculiar to humanity among all species in nature, but also to specify what is genuinely human rather than merely conventional, and what human growth and fuller realization would be. These discourses are not always consciously utopian; they may be directed to individual reformation or to preparation for the afterlife. Furthermore, the discussion may be carried on without reference to concrete social practices and institutions. That is, philosophical anthropologies aim to assess the various kinds of human activity, the various pleasures open to human beings, or the various styles of life made possible by advancing civilization or cumulatively progressive science. A few examples are Schiller's *On the Aesthetic Education of Man* (1795), Ernst Bloch's *The Principle of Hope* (1955–1959), Herbert Marcuse's *Eros and Civilization* (1955), and Norman O. Brown's *Love's Body* (1966). But despite the abstract quality of philosophical anthropology, and whatever the intentions of a given writer, it would be unduly constrictive to omit altogether this literature from an account of utopianism. When its idealism is manifest, philosophical anthropology is thus highly relevant to or allied with utopianism.

In the fifth group are prophecies of profound alteration for the better in human existence made by religious groups, statements of purpose made by revolutionary groups, and blueprints offered by individuals, sects, and secular associations. Obviously, not all activist and reformist political and religious groups have sought to remake society completely, in conformity with the utopian aim of harmony. Nevertheless, many groups have not been satisfied merely to speculate about the ideal society but have sought to realize it, either by persuasion or violence. Examples are the sixteenth-century Anabaptist millenarian, or chiliastic, movements in Europe, radical Protestant groups in the English civil war in the middle of the seventeenth century, and some of the marginal radical figures in the French Revolution, such as Gracchus Babeuf. And in the nineteenth century, especially in the United States, small bands of eager people, religious or simply high-minded, formed utopian communities on unoccupied land, enclaves in isolation from the larger society. Some residues continued to exist after the nineteenth century. In the second half of the twentieth century, for example, a few communal utopian experiments in the United States were inspired by *Walden Two*

(1948), the utopian novel by B. F. Skinner, a behavioral psychologist.

CAUSES OF UTOPIANISM

The literature of social harmony is thus extensive and diverse. Some periods and some cultures have been richer in utopianism than others. The question therefore arises as to why some persons become utopian in their thought or, more rarely, in their action. What causes the desire for change to be absolute, the character of idealism to be extreme and uncompromising, the passion for harmony so averse to the normal condition of dispute and dissonance? Several answers are found scattered in the history of utopianism; some indicate urgency, others do not.

First, some intellectuals simply need to invent worlds. The construction of a utopia, even if only on paper, is a godlike act and resembles the creation of a fictional world by the nonutopian novelist. A utopia can thus be an effort at mastering the complexity of social phenomena; part of the effort consists of rearranging social phenomena to form a more rational or beautiful pattern. In short, one impulse that sustains utopianism, from Plato to the latest science fiction, is to give imagination free rein. This is serious intellectual playfulness. (The same could be said about philosophers of inevitable progress, howsoever they present their optimism.)

Another cause is the desire for moral clarity. In the course of carrying one's demands on social reality as far as possible, one may achieve a fixed—potentially rigid—position in relation to that reality. As a consequence, reality can be constantly put to the test. To the utopian writer, improvisation that allows purposes to emerge from the onrush of experience or waits for new means to suggest or impose new ends is nothing more than a passive or complacent or naive immersion in reality or a confused and unprepared reception of it. Although utopian writers may do nothing to improve society, they may still deem it worthwhile to preserve the concept of the ideal. This may be thought desirable even in comparatively decent societies; to insist on the distinction between the acceptable and the ideal can have a chastening influence on those who govern as well as on those who happily go along. The utopian writer in all varieties of utopianism promotes dissatisfaction and self-criticism, with the risk, of course, of simultaneously provoking a reinvigorated defense of the status quo.

A further cause of utopian thought—and one that lacks the quality of comparative detachment present in the two preceding ones—is the wish to subject society to a total indictment. What is involved here is not a sense that things could be, or may always be, much better than they are but that everything, or nearly everything, is intolerable—inhumanly oppressive—and deserves to go under. There is the direct, unappeasable indictment of established institutions, the way of Jean-Jacques Rousseau in his discourses, William Blake in some of his long poems, Marx and Engels in *The Communist Manifesto*, or D. H. Lawrence in his two books on the unconscious. In works of this sort hatred of social reality may be stronger than love of any alternative; the positive utopianism may be only implicit.

Other works propose, in contrast, that existing social conditions are a spurious utopia: the mass pleasures, whether technological or licentious, provided by affluent society block the way to a genuine transformation of the human condition into a genuine utopia. Such was the theme of the Frankfurt School of social critique in the middle third of the twentieth century. The indictment of society is indirect when the utopianism is explicit and the practices of the ideal society are sketched. And because the main aim is to indict, the practices of the ideal society are, at least in large part, the contradiction of those in existence. The utopian imagination in these instances is hemmed in by the grave defects of the real world; the urge is strong to replace them by conditions that in no way resemble them or to discredit them intellectually. Utopian writing so motivated may blend into radical satire aimed at the status quo and produce a work as great as Jonathan Swift's *Gulliver's Travels* (1726). Or it may produce works such as William Morris's *News from Nowhere* (1891) that are plainly archaistic and may expose themselves to the charge of immaturity or irrelevance. Almost all utopian works contain curiosities and excesses, which may often be explained as compensatory responses to especially terrible features of the real world.

A similar cause of utopian thought is tactical. There are times when it may appear to those bent on reforming society that overstatement is necessary for some degree of success. That is, utopian works need not harbor utopian intentions or even an abstract utopian commitment. Although writers may lavish great energy on making their utopias plausible and attractive, they may aspire only to contribute to the gradual and partial amelioration of their societies. By painting fair pictures of felicity and suggesting that the world is, as presently made up, remote from that felicity, they may encourage an innovating spirit. At the same time, these utopias will give at least guidelines for reform. There may be no real expectation that the utopia will ever fully materialize or, indeed, that pure felicity can be had on any terms. Nevertheless, with-

out that exaggeration, less-than-utopian reform would perhaps be too modest or too slow. Much depends on the persuasiveness of the writer's scheme. For that reason the utopias of reform tend to be less free in their speculation and are content to suggest the completion of certain good tendencies in the real world rather than trying to overturn it theoretically. Edward Bellamy's *Looking Backward, 2000–1887* (1888) is an example of this tactic.

The last cause of utopian thought is the most obvious—the conviction that the whole truth about human well-being in a setting of social harmony is known, can be imparted, and should be acted on. There is, of course, a wide variety in the historical situations that call forth such an overweening attitude. But if some radical Protestant groups (such as the German Anabaptists of the sixteenth century), some utopian movements of the nineteenth century (such as those inspired by the Comte de Saint-Simon, Charles Fourier, and Robert Owen), and those Marxists who are quasi-utopian are exemplary, there must be a sense of deep, intolerable wrong. There must also be a sense of enormous possibility, of not only righting the wrong but also going beyond to perfection itself, and either an overpowering group- or self-confidence or the conviction that the utopian leaders and their following are the instruments of some higher will or the culmination of some impersonal process. The word *messianism* perhaps best summarizes some manifestations of this utopian spirit.

USES OF UTOPIANISM

Apart from their place in history, of what use are the works of utopianism? When utopian writings are deliberate constructions of whole societies, readers may think that utopianism is simply a scattering of uninhabited palaces—grand imaginary structures that may amuse realists if not filling them with contempt. But utopianism is more than its core, the deliberate constructions made by the imagination. The utopian aspiration is found in various modes of writing, and is sometimes oblique or even hidden. Is there, however, something of enduring value, in all these modes, even the deliberate constructions, apart from any question of application? There are, in fact, several benefits conferred by utopianism.

As already noted, a cause of utopian writing is playful delight in the act of imagining new kinds of social reality. This delight can be answered by the pleasure the reader takes in the results of that playfulness. The standards for judging utopianism (in any of its modes) from this point of view are primarily aesthetic—plausible novelties in the projected way of life, clever and ingenious details, daring departures from customary practices. The inner coherence of the utopian ideal matters more than any closeness to probability, although naturally too much strain on belief weakens the pleasure. Admiration for the skill of the utopian writer may be mixed with appreciation for being allowed to contemplate what it would mean to live other lives. No stimulus to make one's own better need be felt. This may make the utopian enterprise somewhat precious, but it can be a source of guiltless satisfaction even to the most conservative temperaments. The utopian works of H. G. Wells are famous for their power to gratify the taste for sampling different worlds, however else they may instruct.

A second use of utopianism is as a record of human aspiration. For the record to be complete, many other kinds of utterance must be consulted, but the various modes of utopianism supply a valuable indication. They are peculiarly vivid forms taken by changeable human longings underlain by permanent human wants. Read with due allowance for their often lopsided or eccentric quality, they will shed vivid light on their times. The desperation of a given historical period, together with the limits of its hopefulness, may emerge from a study of its utopian writings. The abundance or paucity of utopian writing is itself an aid to understanding a period.

Third is the contribution of several modes of utopian literature to general sociology. The great constructed utopias—Plato's *Republic* and *Laws*, the relevant parts of Aristotle's *Politics*, More's *Utopia*, Tommaso Campanella's *The City of the Sun* (1623), Morelly's *Code de la nature* (1755), the writings of Saint-Simon and Fourier (early nineteenth century), H. G. Wells's *A Modern Utopia* (1905)—incorporate a great deal of sociological wisdom. Common to these and other utopias is the idea of the integration of social institutions in its most intense version, utopian harmony. To utopian writers no habit or practice seems innocent of significance for the proper maintenance of the utopian society. Utopian writers are therefore constantly pointing out connections between things that appear unrelated. Part of utopian analysis consists in the attempt to identify the major elements of society and to demonstrate how they act on one another and how each must be adjusted to the others if the best possible world is to be attained. For all their care, utopian writers commit a radical abstraction when they create their images of perfection, but this is the price paid by all general sociology, including that which is wholly neutral and descriptive.

The last use of utopianism is moral. Utopian literature (including the literature relevant or allied to it) is a

repository of reflection on human nature. Although not directly concerned to expose frailty, to scrutinize motives, and to astonish with cynical revelations, utopian literature has in it much hard psychological intelligence. Utopian writers disagree among themselves on the extent to which human nature is reformable, but rarely is this problem treated lightly. Indeed, it is usually acknowledged as the problem requiring the deepest study; it is also the source of the greatest hesitation. The principal mission of utopianism is to encourage the hope that human nature is changeable for the better beyond the limits assigned by worldly pessimism or theological despair. That the real world, despite its amazing pluralist variety, still does not exhaust the possibilities of human nature is the heart of utopianism. The long series of utopian texts enlarge the world by suggesting new character types and new social milieus in which these types could emerge. They also enlarge the world by their claim that the societies of the world ignore, repress, distort, or destroy human potentialities that have not yet been fulfilled.

It is true that the concept of harmony rules out some segment of the spectrum of human nature. The essence of antiutopianism is the charge that any imaginable utopia, like any generous philosophical anthropology, actually impoverishes human nature by not allowing scope to those traits—wildness, excess, discontent, perversity, risk-taking, heroism—that threaten harmony. If therefore the precondition of a harmonious life is the thorough manageability of people, allegedly for their own good, human nature must suffer a terrible diminution. Such diminution is the awful hidden human sacrifice that utopianism exacts with a good conscience. What intensity of experience, what craving for more than satisfaction, what passion for the unknown and the unlimited, would be left? Humanity should always face difficulties that are impossible or nearly impossible to overcome.

For many people, perhaps most, utopia can and does already appear in experiences and temporary conditions, in moments and episodes, in the world as it is. Each person's utopia is different from everyone else's. Utopia cannot be an uninterrupted and common way of life for a whole society. The only genuine utopia is actual life, and every proposed utopia is a dystopia. The critique of utopianism is without doubt a rich field and numbers Friedrich Nietzsche and Fyodor Dostoyevsky among its luminaries. A shrewd and witty antiutopian satire is Aldous Huxley's novel *Brave New World* (1932). A related antiutopian theme is that utopias are often driven by a strong passion for equality that threatens to efface all that is fine or rare in life or that can be created by or appeal to only a few. Utopias level society and thus work to make people more or less uniform and interchangeable; preaching individual expressive growth, utopias often destroy the social and psychological conditions of such growth. Utopian harmony is only monotonous.

In rebuttal, utopian writers and their sympathizers are proud to confine their imagination to the realm of the largest happiness. Within that realm, utopians say, much more human excellence is possible than many people commonly think. That would be proven, if only the world, or a part of it, could be transformed or would become more permissive. Without subscribing to any set of specific utopian ideas, one can appreciate—at least to a moderate extent—the efforts of utopian writers to rescue this sentiment from the disparagement of those who believe, explicitly or not, that pain not only will but should remain, if not definitive of the human condition, then its substratum.

See also Civil Disobedience; Cosmopolitanism; Multiculturalism; Postcolonialism; Republicanism.

Bibliography

Buber, Martin. *Paths in Utopia*. 1949. Translated by R. F. C. Hull. Boston: Beacon, 1970.

Bury, J. B. *The Idea of Progress*. 1932. New York: Dover, 1955.

Cioran, E. M. *History and Utopia*. 1960. Translated by Richard Howard. New York: Seaver Books, 1987.

Claeys, Gregory, and Lyman Tower Sargent, eds. *The Utopia Reader*. New York: New York University Press, 1999.

Cohn, Norman. *The Pursuit of the Millennium.* 2nd edition. 1957. New York: Oxford University Press, 1970.

Gray, Alexander. *The Socialist Tradition, Moses to Lenin*. London: Longman's Green, 1946.

Horsburgh, H. J. N. "The Relevance of the Utopian." *Ethics* 67 (1957): 127–138.

Kateb, George, ed. *Utopia*. New York: Atherton, 1971.

Kateb, George. *Utopia and Its Enemies.* 2nd edition. New York: Schocken, 1972.

Lovejoy, Arthur O., et al., eds. *A Documentary History of Primitivism*. 1935. New York: Octagon Books, 1965.

Mannheim, Karl. *Ideology and Utopia*. 1929, 1936. Translated by Louis Wirth and Edward Shils. New York: Harcourt, Brace, 1968.

Manuel, Frank E., ed. *Utopias and Utopian Thought*. Boston: Beacon, 1966.

Manuel, Frank. E., and Fritzie P. Manuel, eds. *French Utopias*. New York: Free Press, 1966.

Manuel, Frank E., and Fritzie P. Manuel. *Utopian Thought in the Western World*. Cambridge, MA: Harvard University Press, 1979.

More, Thomas. *Utopia,* edited and translated with an introduction by David Wootton. Indianapolis: Hackett, 1999.

Mumford, Lewis. *The Transformations of Man*. New York: Harper, 1956.

Negley, Glen. *Utopian Literature: A Bibliography*. Lawrence, KS: Regents Press of Kansas, 1978.

Negley, Glenn, and J. Max Patrick, eds. *The Quest for Utopia*. New York: H. Schuman, 1952.

Noyes, John Humphrey. *History of American Socialisms*. 1870. New York: Dover, 1961.

Polak, Fred L. *The Image of the Future*. 1961. 2 vols. Translated by Elsie Boulding. New York: Elsevier, 1973.

Popper, Karl. *The Open Society and Its Enemies.* 5th edition. 2 vols. Princeton, NJ: Princeton University Press, 1966.

Sargent, Lyman Tower. *British and American Utopian Literature 1516–1975: An Annotated Chronological Bibliography*. New York: Garland, 1988.

Thrupp, Sylvia, ed. *Millennial Dreams in Action*. New York: Schocken, 1970.

George Kateb (1967, 2005)

V

VACUUM AND VOID

See *Quantum Mechanics*

VAGUENESS

A term is vague if, and only if, it is capable of having borderline cases. All borderline cases are inquiry-resistant: Senator Hillary Clinton is a borderline case of "chubby" because, given her constitution, no amount of conceptual or empirical investigation can settle the question of whether or not she is chubby. Notice that this is not vagueness in the sense of being underspecific. If her spokesperson states that the senator weighs between 100 and 200 pounds, reporters will complain that the assertion is too obvious to be informative—not that the matter is indeterminate.

Typically, borderline cases lie between clear negative cases and clear positives. Moreover, the transition from clear to borderline cases will itself be unclear. If one thousand women queue in order of weight, there is no definite point at which the definitely non-chubby end and the borderline chubby begin. In addition to this second order vagueness: There is third order vagueness: There is no definite point at which the definitely definite cases end and the indefinitely definite ones begin.

Vagueness is responsible for Eubulides' 2,400-year-old sorites paradox. This conceptual slippery slope argument can be compactly formulated with the help of mathematical induction:

> Base step: A collection of 1 million grains of sand is a heap.
>
> Induction step: If a collection of n grains of sand is a heap, then so is a collection of $n - 1$ grains.
>
> Conclusion: One grain of sand is a heap.

Long dismissed as a sophism, the sorites began to acquire respect in the 1970s. By 1990, its status was comparable to Eubulides' other underestimated paradox, the liar.

Eubulides may have intended the sorites to support Parmenides' conclusion that all is one. For one solution is to deny the base step on the grounds that there really are no heaps. Since a sorites paradox can be formulated for any vague predicate for ordinary items (cloud, chair), the solution only generalizes by a rejection of common sense. In any case, a few contemporary metaphysicians have championed this radical position. A less strident group

hopes that the sorites will be rendered obsolete by science's tendency to replace vague predicates by precise ones.

VIEWS ON VAGUENESS

C. S. Pierce was the first philosopher to propose that logic be revised to fit vagueness. Pierce developed a form of many-valued logic. "Hillary Clinton is chubby" is assigned a degree of truth between 1 (full truth) and 0 (full falsehood), say .5. Truth-values of compound statements are then calculated on the basis of rules. Disjunctions are assigned the same truth value as their highest disjunct. Conditionals count as fully true only when the antecedent has a truth-value at least as high as the consequent. This "fuzzy logic" undermines the induction step of the sorites. As the progression heads into the borderline zone, the consequent has a value a bit lower than the antecedent. Although a small departure from full truth is normally insignificant, the sorites accumulates marginal differences into a significant difference.

Supervaluationists deny that borderline statements have any truth-value at all. Words mean what we intend them to mean. Since there has been no practical need to decide every case, our words are only partially meaningful. We are free to fill in the gaps as we go along. If a statement would come out true regardless of how the gaps were filled, then we are entitled to deem the statement as actually true. This modest departure from truth-functionality lets the supervaluationists count "Clinton is chubby" or "Clinton is not chubby" as true even though neither disjunct has a truth-value. Indeed, all the tautologies of classical logic will be endorsed by this principle. All the contradictions will be likewise rejected. This suggests a solution to the sorites paradox. For every precisification of "heap" makes the induction step come out false.

Supervaluationism resonates with the use theory of meaning. If a term gets its meanings from linguistic practices, then the incompleteness of those practices will generate semantic gaps. In his work, Derek Parfit (1984) provides the example of a club that stops meeting. After a while, some of the members of the club start meeting again. Is this a new club or has the old club been revived? Parfit maintains this question is empty; there is no true answer or false answer. There might have been a correct answer if the founders had written a constitution that specified the conditions under which the club persists. But the club was an informal institution. Parfit believes our concept of personhood has a similar level of informality. There is vagueness as to when a fetus develops into a person, vagueness as to when brain damage suffices to end a person, and vagueness as to whether a person survives various hypothetical processes such as teletransportation.

Vagueness raises a methodological issue in philosophical analysis. What should be done with borderline cases? In his work, Nelson Goodman (1951) states a good theory is entitled to decide these "don't care" cases. To the victor go the spoils! Others are more sympathetic to the principle of coordinated indeterminacy; we should prefer theories that preserve gaps.

Aristotle postulated we should not demand more precision than the subject matter allows. But Goodman's argument is suspicious of any *a priori* assessment of how much precision is permitted. Just as we may be surprised to find that an apparently determinate question lacks a determinate answer (such as "What time is at the North Pole?"), we may be surprised that an apparently indeterminate question has a determinate answer. For instance, Ernst Mach dismissed the question "Is heat the absence of coldness or is coldness the absence of heat?" as a scholastic quibble. Atomists later showed that coldness is the absence of heat.

Israel Scheffler (2001) traced the belief that there are empty questions to the analytic-synthetic distinction. After all, a borderline case is supposed to be semantically indeterminate. We are supposedly unable to conceive of how the addition of a single grain could turn a non-heap into a heap. Scheffler believes that rejection of analytic-synthetic distinction would prevent intellectualism defeatism. In his work, he urges philosophers to stick to classical logic and persist with inquiry.

Epistemicists embraced Scheffler's logical conservatism but offered a new foundation for defeatism. They said vagueness is ignorance. "Clinton is chubby" has an *unknowable* truth-value. Consequently, the induction step of the sorites is plain false; there is an n such that n grains of sand make a heap but $n - 1$ does not. So there is no need to change logic. Instead we should change our beliefs about language.

The basic objection to epistemicism is that it requires a linguistic miracle. How could our rough and ready practices ensure a threshold for "heap" and "chubby"? Given that the threshold for "heap" exists, what explains our ignorance of it?

Timothy Williamson (1994) answers that knowledge requires a margin for safety. Suppose case n is an F and case $n + 1$ is a non-F that is indistinguishable from case n. The correctness of your belief that n is an F would then be a matter of luck. Since knowledge is incompatible with

luck, you would not really know that *n* is an F. So given that there is a threshold for F-ness, you cannot know it. In his work, Williamson reconciles ignorance with the use theory of meaning by emphasizing the chaotic complexity of linguistic practice. Our computational resources are not sufficient to settle all cases.

Is Williamson's ignorance too relativistic? Parfit's intuition is that no amount of investigation can settle the question of whether the club is old or new—not merely that no amount of *human* investigation is enough. If Williamson were right, then extraterrestrial anthropologists could figure out whether Parfit's club was new by applying their superior intellects. Indeed, since there is variation in human cognition, Williamson's account seems to permit borderline status to vary a bit from speaker to speaker. Supervaluationists and fuzzy logicians claim an advantage because their borderline cases are absolute.

Roy Sorensen (2001) suggests that the epistemicist can model absolute borderline cases with truth-maker gaps. A truth-maker is a state of affairs that makes a proposition true. All contingent propositions that are definitely true have truth-makers. But some truths lack truth-makers. Applying a predicate to a borderline case yields a proposition with a free-floating truth-value. Since we can learn the truth-values of contingent propositions only through connections with their truth-makers, indefinite truths are absolutely unknowable. Since there are borderline cases of "has a truth-maker" there will also be absolute higher order vagueness.

See also Fuzzy Logic; Goodman, Nelson; Many-Valued Logics; Parfit, Derek; Peirce, Charles Sanders.

Bibliography

Evans, Gareth, "Can There Be Vague Objects?" *Analysis* 38 (1978): 208.

Fine, Kit, "Vagueness, Truth, and Logic." *Synthese* 30 (1975): 265–300.

Goodman, Nelson. *The Structure of Appearance*. New York: Bobbs-Merrill, 1951.

Parfit, Derek.*Reasons and Persons*. Oxford: Clarendon, 1984.

Sheffler, Israel. *Beyond the Letter*. London: Routledge, 1979.

Sorensen, Roy. *Vagueness and Contradiction*. New York: Oxford University Press, 2001.

Wheeler, Samuel. "Megarian Paradoxes as Eleactic Arguments." *American Philosophical Quarterly* 20 (1983): 287–295.

Williamson, Timothy. *Vagueness*. London: Routledge 1994.

Roy Sorensen (1996, 2005)

VAIHINGER, HANS

(1852–1933)

Hans Vaihinger, the German philosopher of the "as if," was born in a devout home near Tübingen. Although he developed unorthodox religious views at an early age, he attended the Theological College of the University of Tübingen. Vaihinger wanted to be a man of action, but his extreme nearsightedness forced him into scholarly pursuits. He regarded the contrast between his physical constitution and the way he would like to live as irrational, and his defective vision made him sensitive to other frustrating aspects of existence.

Vaihinger eventually became a professor of philosophy at Halle, but failing vision necessitated his giving up his duties in 1906. He then turned to completing his most important work, *Die Philosophie des Als-Ob* (Berlin, 1911; translated by C. K. Ogden as *The Philosophy of "As If,"* New York, 1924), which had been started in 1876. The volume went through many editions and made the philosophy of fictions well known. Vaihinger also achieved renown as an Immanuel Kant scholar and founded the journal *Kant-Studien.* He also founded (with Raymund Schmidt) the *Annalen der Philosophie,* a yearbook concerned with the "as if" approach. He was much interested in the theory of evolution and emphasized the biological function of thought. On occasion he expressed himself sharply. For example, when quite young he defined humankind as "a species of monkey suffering from megalomania." This resulted in considerable controversy, and Vaihinger later seemed to regret this definition, although he still found some merit in it.

GENERAL POINT OF VIEW

In many ways Vaihinger was attracted to apparent inconsistencies. Although he held theological doctrines to be false in any literal or factual sense, Vaihinger, somewhat like George Santayana, found considerable aesthetic and ethical merit in Christian doctrines. Both idealism and materialism interested him, but he found either alone to be unsatisfactory. Indeed, he regarded the problem of the relation of matter to mind as logically insoluble. He was much influenced by Kant and emphasized the importance of categories supplied by the mind in the perception of objects; yet he wanted to modify Kant in a more materialistic and empirical direction.

Vaihinger's urge to absorb elements of apparently conflicting approaches is illustrated by the label he chose for his philosophy: idealistic positivism or positivist idealism. He was impressed by F. A. Lange's *History of Mate-*

rialism and respected both Lange's Kantian views and his great knowledge of the natural sciences. But even Lange's neo-Kantianism needed to be made more empirical and positivistic, in Vaihinger's view. This was to be achieved by recognizing the necessity and utility of acting on the basis of fictions that are known to be false.

Vaihinger praised Arthur Schopenhauer's pessimism and irrationalism. Too many philosophers (especially G. W. F. Hegel) had believed that the ideal of philosophy was to furnish a rational explanation for everything. But for Vaihinger both nature and history contain many irrational elements, and he regarded Schopenhauer as one of the few philosophers sincere enough to emphasize that irrationality.

Vaihinger maintained that pessimism gives moral strength, enables one to endure life, and helps to develop a more objective view of the world. He emphasized that in his opinion the difficulties of Germany, and especially its defeat in World War I, were largely attributable to the prevailing optimism of German idealism. He saw a close relation between philosophy and practical politics, arguing that a "rational pessimism" might have prevented the war.

FICTIONS

The Platonic myths were the first stimuli to Vaihinger's eventual theory of fictions. Later, Kant's antinomies also were influential. Lange had said, "Man *needs* to supplement reality by an ideal world of his own creation"; Vaihinger expanded this view and applied it to science, metaphysics, theology, social ideals, and morality. Fictions are not to be mistaken for true propositions, for fictions are known to be false. They contradict observed reality or are self-contradictory, and so they falsify experience. Something can work *as if* true, even though false and recognized as false.

Vaihinger distinguished his philosophy from any pragmatism that holds that a statement is true if it is useful in practice. In contrast, he argued: "An idea whose theoretical untruth or incorrectness, and therewith its falsity, is admitted, is not for that reason practically valueless and useless; for such an idea, in spite of its theoretical nullity may have great practical importance" (*The Philosophy of "As If,"* p. viii). Nevertheless, he admitted that in practice pragmatism and fictionalism had much in common, especially in their acknowledgment of the significance of heuristic ideals.

Nor can fictionalism be identified with any variety of skepticism. Vaihinger interpreted skepticism as the doubting of some view. Fictionalism does not doubt the correctness of its fictions; it knows them to be wrong. Vaihinger thought that the label "skepticism" was applied to his philosophy because of its views on God and immortality. He suggested that the label "relativism" (in the sense of opposition to absolutism) better fitted his views.

FICTIONS AND HYPOTHESES. Vaihinger distinguished between hypotheses and fictions. Methodologically they are very different, but they are similar in form and hard to separate in practice. According to Vaihinger, a hypothesis is "directed toward reality" and is subject to verification, but fictions are never verifiable, for they are known to be false. In the case of a number of competing hypotheses, the most probable is selected, but in the case of a number of competing fictions, the most expedient is chosen. Vaihinger held that to treat "Man is descended from the lower mammals" as a hypothesis is to say that we believe that if we had lived at the appropriate time, we would have perceived the ancestors of man, that we may still find the remains of those ancestors, and so on. In contrast, Johann Wolfgang van Goethe's notion of an animal archetype of which all known animal species are modifications was a fiction. Goethe did not believe the archetype had ever existed; he was saying that all animals could be regarded as if they were modifications of the single type.

Goethe's fiction was of considerable value despite its falsity, since it suggested a new classificatory system and had heuristic value for Darwin's later theory. Hypotheses, then, are constructed with the hope of verification, but "the fiction is a mere auxiliary construct, a circuitous approach, a scaffolding afterwards to be demolished." Thus, what is untenable as a hypothesis, especially if exceptions to it are discovered, may be useful as a fiction. Hypotheses are verified by experience, but fictions are justified by the services they render, by their utility.

CHARACTERISTICS OF FICTIONS. Fictions have four general characteristics: (1) They either deviate from reality or are self-contradictory. (2) They disappear either in the course of history or through logical operations and are used only provisionally. (3) The users of a fiction normally are consciously aware that the fiction lays no claim to being true; frequently in the history of thought, however, the first users of a fiction mistake it for a hypothesis. (4) Fictions are the means to some definite end; fictions lacking that expediency are mere subjective fancies.

THE UTILITY OF FICTIONS. Vaihinger adopted a basically biological account of the utility of fictions and made

lengthy comparisons of psychical and physical processes, holding that the same general notion of utility applies in both cases. He specifically mentioned "ready adaptation to circumstances and environment," the maintenance of a "successful reaction" to external impulses and influences, and "the adoption and acceptance or the repulsion of new elements." A Kantian emphasis also appears in this context. The psyche is not a receptacle into which sense impressions are poured but is, rather, a "*formative force,* which independently changes what has been appropriated." It is also assimilative and constructive. Logical thought, using fictions, "is an active appropriation of the outer world."

EXAMPLES OF FICTIONS. Vaihinger discussed in great detail specific fictions used in diverse realms of discourse. God and immortality have already been mentioned. It may be a great convenience to act as if the cosmos were orderly and created by an all-powerful and all-good God and as if man were immortal. The virgin birth is another "beautiful, suggestive and useful myth." Vaihinger agreed with Kant that despite the scientific difficulties of the notion, it has practical utility as an excellent symbol of humankind triumphantly resisting evil and raising itself above temptation. In science the atom is a fiction. Both those who defended the literal reality of the atom and the early positivists who rejected its reality on the grounds that atomic theory was internally contradictory were mistaken. The atom is, rather, "a group of contradictory concepts which are necessary in order to deal with reality."

A materialistic notion of the world is false if taken as a hypothesis but is a necessary and useful fiction. Materialism, Vaihinger held, simplifies our notion of the external world and helps to bolster a scientific outlook. Natural scientists carry on their work as if an external material world existed independently of perceiving subjects, and thus science can "proceed on the basis of relations far simpler than those actually presented to a careful observation of reality itself" (ibid., p. 200). The notion of a vital force in biology, while full of difficulties, may have some use as a fiction. Vaihinger regarded such a fiction as "an abbreviation for the sum of all the causes that determine the phenomena of life" (ibid., p. 212). It enables us to express some matters in a simpler way than we otherwise could. To cite one final example, doctrines in social theory, such as the notion of an original social contract, may be helpful. An extremely complicated situation can be grasped by adopting a fiction that deliberately substitutes for "the complete range of causes and facts" a part of that range.

Vaihinger's theory of fictions can be regarded as a denial of the view of W. K. Clifford and others that belief should always be proportionate to the evidence. Intellectually, practically, and morally we need false but expedient fictions to cope with the world. Many traditional philosophic views are mistaken in that they confuse the human need for certain doctrines with the truth of those doctrines; but various forms of skepticism, positivism, and materialism are wrong in assuming that because certain doctrines are false, they should be eliminated.

THEORY OF MIND

According to Vaihinger, all knowledge "is a reduction of the unknown to the known, that is to say a comparison." He held that there are limitations to all thought, although he did not wish to lament them; we cannot leap out of our skins and somehow attain what we cannot attain. These limitations apply not only to man but also to "the highest Mind of all," and they come about because thought originated as a means to an end. The end is to serve the will to live.

THE PURPOSE OF THOUGHT. Vaihinger held that "the test of the correctness of a logical result lies in *practice,* and the purpose of thought must be sought not in the reflection of a so-called objective world, but in rendering possible the calculation of events and of operations upon them" (ibid., p. 5). The purpose of thought is not correspondence with an assumed objective reality; nor is it the theoretical reconstruction of an outer world within consciousness; nor is it the comparison of things and logical constructs. It is pragmatic in the sense that successful logical products enable us to "*calculate events that occur without our intervention.*"

Vaihinger maintained that nature proceeds entirely according to "hard and unalterable laws … but thought is an adaptable, pliant, and adjustable organic function." Very probably the most elementary physical processes contain certain strivings. In organic beings, those strivings develop into impulses. Man, in his evolutionary development from the animals, has had those impulses transformed into will and action. Thus ideas, judgments, and conclusions act as means of survival.

SENSELESS PROBLEMS. Vaihinger put great stress on what he termed the "Law of the Preponderance of the Means over the End." According to this law, the well-adapted means to a specific end everywhere have a tendency to become independent and ends in themselves. Thus the mind sets itself impossible problems that cannot

be solved, even by "the highest Mind of all," just because no mind was developed for those purposes. Eventually "emancipated thought" sets for itself senseless problems, among which Vaihinger listed questions about the origin of the world, the formation of matter, the origin of motion, the meaning of the world, and the purpose of life. He gave particular attention to the relation of mind and matter. His philosophy was admittedly inconsistently dualistic; on the one hand it reduced all reality to sensations, and on the other it reduced all reality to matter. But Vaihinger insisted that no logical, rational unification is possible through any philosophy and that the question of the relation of mind to matter is as senseless as that of the purpose of existence.

However, a nonrational solution is possible to the various world-riddles: "in intuition and in experience all this contradiction and distress fades into nothingness." Experience and intuition, Vaihinger said, are "higher than all human reason," and we do not "understand the world when we are pondering over its problems, but when we are doing the world's work." Experience and intuition give us the harmonious unity that reason cannot supply. Philosophers are especially prone to torture themselves with unanswerable questions; the wise man is content if life is successful on the level of practice. Shifts, probably unwarranted, in the meaning of such terms as *understand* occur here, but Vaihinger's main point seems to be that there are nonrational solutions to questions which have no rational answers.

THOUGHT AND REALITY. Subjective events alter reality either by adding to it or by subtracting from it. Yet correct practical results are frequently obtained, and in that sense "thought tallies with reality." Hence, both what Vaihinger called logical optimism, the assumption that thought mirrors reality, and logical pessimism, the assumption that thought is always deceptive, need to be avoided. Senseless questions will not be answered in the future by some new philosophic synthesis but, rather, are explained by "looking backwards," by discovering their psychological origin.

RELIGION

Vaihinger's views on religion illustrate his general reluctance to accept either alternative of some of the traditional philosophic polarities. His early rationalistic, ethical theism later developed into a variety of pantheism. His pantheism then became, during his stay at Tübingen, a kind of Kantian agnosticism and then something close to Schopenhauerian atheism. Vaihinger saw no need to adopt a negative view toward the historical forms of the church and its various dogmas. But even though he regarded many Christian doctrines as fictions of considerable ethical and aesthetic value, doubt entered. For example, although he thought it was a fiction satisfying to many to take the world as if created, or at least regulated, by "a more perfect Higher Spirit," he further insisted that a supplementary fiction was necessary, holding that the "order created by the Higher Divine Spirit had been destroyed by some hostile force."

Vaihinger believed Friedrich Carl Forberg's views on religion were overly neglected. He agreed with Forberg that "theoretical atheism" was harmless and that everyone should have "an attack" of such atheism at least once, in order to find out whether he desired the good for its own sake or merely for some advantage either in this world or in a future world. On the other hand, Vaihinger deplored "practical atheism," understood as the failure to act so as to make the world better. Religion became a mode of behavior rather than the acceptance of certain theoretical views.

Vaihinger held, in agreement with Forberg, that the striving toward the kingdom of God is what matters, not the achieving of it. In fact, it is very likely that the kingdom of God is an actual impossibility. The man who neglects none of his duties to his fellows and helps to further the common good, even though convinced that the world is filled with wickedness and stupidity, practices true religion. Religion is not the belief in the kingdom of God but the attempt to make it come about while recognizing its impossibility. Vaihinger argued that this was the general view of Kant. He believed that this religion not only had warmth and poetry but also "represents in its radical form the highest point to which the human mind, or rather the human heart, is capable of raising itself."

See also Clifford, William Kingdon; Fictionalism; Goethe, Johann Wolfgang van; Hegel, Georg Wilhelm Friedrich; Idealism; Kant, Immanuel; Lange, Friedrich Albert; Materialism; Neo-Kantianism; Pantheism; Pessimism and Optimism; Pragmatism; Santayana, George; Schopenhauer, Arthur.

Bibliography

Additional works by Vaihinger are *Hartmann, Dühring und Lange* (Iserlohn: J. Baedecker, 1876); *Kommentar zu Kants Kritik der reinen Vernunft,* 2 vols. (Leipzig, 1881–1892); *Nietzsche als Philosoph* (Berlin: Reuther and Reichard, 1902); and *Die Philosophie in der Staatsprüfung* (Berlin: Reuther and Reichard, 1906).

The English translation by C. K. Ogden of *The Philosophy of "As If"* was made from the sixth German edition, specially revised by Vaihinger for the English-speaking philosophical world; it also contains a lengthy and helpful autobiography of Vaihinger that emphasizes the intellectual origins of his views.

See also W. Del Negro, "Hans Vaihinger's philosophisches Werk mit besonderer Berücksichtigung seiner Kantforschung," in *Kant-Studien* (1934): 316–327.

Rollo Handy (1967)

VAILATI, GIOVANNI

(1863–1909)

Giovanni Vailati, the Italian analytical philosopher and historian of science, was born at Crema, Lombardy. He studied engineering and mathematics at the University of Turin, where he later became an assistant to Giuseppe Peano (1892) and Vito Volterra (1895) and lectured on the history of mechanics (1896–1899). In 1899 he resigned his university post to be free for independent work, earning his living by teaching mathematics in high schools. By the end of his life Vailati's ideas were internationally recognized; some of his writings had been translated into English, French, and Polish, and he was personally acquainted with many of the important scholars of his time. He was forgotten after his death, however, and only since the late 1950s has he received renewed attention.

The main feature of Vailati's thought is his methodological and linguistic approach to philosophical problems. Rather than propounding anything resembling a doctrine, Vailati presented concrete examples of how to apply his new methods. He left no complete book, but only some two hundred essays and reviews on a great number of problems in several academic disciplines. The best way to indicate the range of his philosophical interests is, therefore, to report the titles of his most important essays in philosophy. In chronological order, they are "The Importance of Investigating the History of the Sciences" (its bearing on the understanding of scientific method); "Deductive Method as a Tool for Inquiry"; "Questions of Words in the History of Science and Culture" (on semantical problems); "The Difficulties that Impair Any Attempt Rationally to Classify the Sciences"; "The Logical Bearing of Brentano's Classification of Mental Facts"; "The Applicability of the Concepts of Cause and Effect in Historical Sciences"; "The Most Modern Definition of Mathematics" (Bertrand Russell's); "The Role of Paradoxes in Philosophy"; "The Tropes of Logic" (in which the important point is made that induction cannot be grounded, because if it were grounded, it would become deduction); "The Hunt for Antitheses" (an attack on the philosophical tendency toward unification and a defense of analysis); "The Distinction between Knowing and Willing"; "The Search for the Impossible" (which contains an assessment of G. E. Moore's *Principia Ethica* and an acceptance of his method); "Pragmatism and Mathematical Logic"; "Toward a Pragmatic Analysis of Philosophical Terminology"; "A Handbook for Liars" (a review of Giuseppe Prezzolini's *The Art of Persuading*); and "The Grammar of Algebra" (containing a comparison of the syntax of ordinary language with that of algebra).

Vailati's next important work, "Language as an Obstacle to the Elimination of Illusory Contrasts," is possibly his most concentrated inquiry into the relation between speech and thought and into the influence of speech on thought. Finally should be mentioned the papers Vailati wrote with his pupil, Mario Calderoni—"The Origins and Fundamental Idea of Pragmatism," "Pragmatism and Various Ways to Say Nothing," and "The Arbitrary in the Operation of the Mental Life." To all these articles Vailati brought a sense of humor; independence of judgment; a mind as cautious, matter-of-fact, and candid as one could wish for in a philosopher; complete control of mathematics, symbolic logic, and the history of the subject being examined; and an extremely concentrated style.

PHILOSOPHY

For Vailati, philosophy is no superscience that can teach scientists what they should do. It cannot make discoveries; it can only prepare the intellectual climate and furnish some of the necessary tools. It is a neutral enterprise that can receive contributions from people holding different personal beliefs and conceptions. It should avoid the struggle between systems which, "let us hope, will some day end like the reported fight between the two lions who ate one another up leaving only their tails on the ground" (*Scritti,* p. 652). As it has no special field of its own, philosophy should not construct any special language or resort to any jargon but should take into account what is already present in language. When a philosopher wants to ban a problematic term to avoid a related problem, he deludes himself; and when he substitutes for an ordinary-language term a technical term of his own or one drawn from a special science, his policy reminds one of "the advice given to children in jest that one can catch a bird by putting salt on its tail" (ibid., p. 315). The right policy consists in correcting the use of the ordinary term—in using it "technically," if you like, but in a tech-

nical use as near as possible to its ordinary use. On the other hand, Vailati denounced as misleading similarity in verbal form or in grammar as contrasted with similarity in thought. He defended the independence of the philosopher with respect to usage as such.

Vailati wrote his most rewarding pages on such subjects as definitions, the difference between statements and other types of sentences, the logic of dispositional expressions versus categorical ones, axioms and postulates, deduction and induction, and the use of experiments. Also of importance are several papers on analytical ethics.

Vailati held that "opinions, whether true or false, are always *facts,* and as such they deserve and require to be made the object of research and verification" (ibid., p. 65). Semantically, this is possible because we can understand and talk about sentences of which it cannot be said that they are either true or false. Indeed, "the question of determining *what we mean* when we propound a given proposition is entirely different from the question of deciding *whether it is true or false*" (ibid., p. 923). On the other hand, mere understanding should not be confounded with scientific method, nor does the study of all that can be significantly said supply us with criteria for assessing truth and falsity. One cannot even begin to deal with the question whether a sentence is true or false before settling the question of what is meant by it. But to decide truth or falsity one must connect present and future experiences in terms of prevision, and propositions and facts in terms of intersubjective verification, both in science and in philosophy. In both "it must be demanded of anybody who advances a thesis that he be capable of indicating the facts which according to him should obtain (or have obtained) if his thesis were true, and also their difference from other facts which according to him would obtain (or have obtained) if it were not true" (ibid., p. 790).

VAILATI'S "PRAGMATISM"

Vailati was a liberal analytical philosopher of the kind that has flourished in England and the United States since World War II. However, he is usually referred to as the chief Italian "Peircean," or "logical," pragmatist. He was indeed one of the first to read Charles Sanders Peirce correctly and to carefully distinguish his thought from William James's. But Vailati's thought was too complex and his acquaintance with the history of ideas too thorough, and the concept of pragmatism is itself too manifold, to call him only a pragmatist. Although he stressed the importance of Peirce, he traced Peirce's ideas back to George Berkeley and even to Plato's *Theaetetus,* claiming that Socrates was presented in that work as "defending against Protagoras the thesis now supported by Peirce under the name of 'pragmatism'" (ibid., p. 921). If Vailati was impressed by Peirce's criteria for meaning and truth, he was equally impressed by Peano's work in mathematical logic, Ernst Mach's principle of the economy of thought, Moore's approach to ethics and Russell's to mathematics, Franz Brentano's classification of mental phenomena, the Gottfried Wilhelm Leibniz revival (to which Vailati contributed), and James's conception of consciousness.

Vailati did not possess Peirce's speculative power and overwhelming originality, but neither did he share the American's ontological troubles and commitments, and he gave his own researches a more empirical and methodological bent. By "pragmatism" Vailati meant mainly a new freedom of thought, a refusal to subscribe to any given doctrine, a willingness to use new intellectual techniques, and a cooperative attitude toward philosophical problems. He possessed new methods and new ways of thought which were neither positivistic nor idealistic; and he needed a new banner under which to fight his intellectual battle within Italian philosophy, which was then in the process of passing over from nineteenth-century positivism to the neoidealism of Benedetto Croce and Giovanni Gentile. Vailati's very individual position within that process helps to account for the long silence about his work, some other reasons being the scattered nature of his publications, the fact that he was in advance of his time, and the intervention of World War I and Italian fascism.

HISTORICAL WORK

As a historian Vailati dealt chiefly with mechanics, logic, and geometry. He made important contributions to the study of post-Aristotelian Greek mechanics, of Galileo Galilei's forerunners, of definition in Plato and Euclid, of the influence of mathematics on logic and epistemology, and of Gerolamo Saccheri's work in logic and in non-Euclidean geometry. He gave a remarkable representation (much more than a translation) of Book A of Aristotle's *Metaphysics.* He was particularly interested in the dialectic of continuity and change, in how "the same" problems are faced and solved in different ways in different periods; which, owing to his constant interest in language, meant that he traced the history of the relations between concepts and terms.

Vailati's work as a historian and as an analytical philosopher were closely interwoven; they are two applications of the same attitudes and methods. He saw the

difference between theoretical and historical research not so much in their subject matters as in their approach to their subject matters. Philosophers and scientists, he held, should cooperate in historical research and remember that no history is complete unless the social background of ideas is taken into account. In science, past results are not "destroyed" by new ones, for new results make old ones even more important in the very process of superseding them. "Every error shows us a rock to be avoided, while not every discovery shows us a path to be followed" (ibid., p. 65). By his awareness of the importance and his command of the methodology of historical research, Vailati avoided the abstract ahistorical atmosphere and the scientifically biased attitude of many logical positivists.

LOGIC

Vailati wrote some early papers in symbolic logic, but he was chiefly interested in the function of logic within philosophy. He attacked confusions between logic and psychology and between logic and epistemology.

CORRESPONDENCE

Vailati's thought cannot be completely evaluated until the hundreds of letters he wrote to Mach, Brentano, Peano, Croce, Volterra, Giovanni Papini, Prezzolini, Giovanni Vacca, and many others, are published. Many concern topics not dealt with in the *Scritti.* These letters constitute one of the last large scientific correspondences of the eighteenth-century kind. They will throw new light on the intellectual history of Europe around 1900 and possibly establish connections hitherto unnoticed or only suspected.

See also Berkeley, George; Brentano, Franz; Croce, Benedetto; Galileo Galilei; Gentile, Giovanni; Geometry; Induction; James, William; Language and Thought; Leibniz, Gottfried Wilhelm; Logic, History of; Mach, Ernst; Moore, George Edward; Papini, Giovanni; Peano, Giuseppe; Peirce, Charles Sanders; Plato; Pragmatism; Propositions; Scientific Method; Semantics; Socrates.

Bibliography

Vailati's manuscripts (some still unpublished) and many of the letters he received are in the Institute for the History of Philosophy of the State University of Milan. The only almost-complete edition of his papers is the *Scritti* (Florence, 1911), which was followed by two anthologies: *Gli strumenti della conoscenza,* edited by Mario Calderoni (Lanciano, 1911), and *Il pragmatismo,* edited by Giovanni Papini (Lanciano, 1911). *Il pragmatismo* includes a completion of Vailati's notes for a book on pragmatism. Some of his best essays were first reprinted in *Il metodo della filosofia,* edited by Ferruccio Rossi-Landi (Bari, 1957), and in *Scritti di metodologia scientifica e di analisi del linguaggio,* edited by M. F. Sciacca (Milan, 1959). Complete collections of the philosophical papers and of the correspondence are in preparation.

The first contemporary scholar to point out Vailati's importance was Eugenio Garin, in 1946; see his *Cronache di filosofia italiana* (Bari: Laterza, 1955), Ch. 5, Sec. 5. See also Ferruccio Rossi-Landi's introduction to *Il metodo della filosofia*; Rossi-Landi's "Materiale per lo studio di Vailati, in *Rivista critica di storia della filosofia* 12 (1957): 468–485 and 13 (1958): 82–108, with extensive bibliographies and an attempt to classify all of Vailati's papers; and Rossi-Landi's "Some Modern Italian Philosophers," in *Listener* 17 (1450 and 1451) (1957): 59–61 and 97–98. The most complete study is a special issue of the *Rivista critica di storia della filosofia* 18 (1963): 273–523, which contains essays by twenty authors.

Ferruccio Rossi-Landi (1967)

VALENTINUS AND VALENTINIANISM

Valentinus (mid-2nd century CE) was the founder of what came to be one of the most influential Gnostic sects of heretical Christianity. Little can be known with certainty about either his life or his teachings, apart from what has been preserved for us in the writings of the church fathers, much of which is reported only very sketchily, with a view toward refutation. The discovery, in 1945, of important Coptic texts at Nag Hammadi has improved our understanding of his thought, but the texts discovered there (principally the so-called *Evangelium Veritatis* [Gospel of truth]) represent the thought of the various schools drawing inspiration from his teachings and cannot reasonably be attributed to Valentinus himself. St. Irenaeus (*Adversus Haereseis* I) and others assert that he was a native of Egypt, where he is said to have studied under Theodas, alleged to have been a pupil of St. Paul, but reports of both the connection to Egypt and to St. Paul may be motivated by a desire to put him into a certain tradition, whether mystical or theological. St. Irenaeus also reports that he lived in Rome during three pontificates (Hyginus, 136–140; Pius, 140–155; Anicetus, 155–166), and Tertullian (*Adversus Valentinianos*) says that he was in communion until he was passed over for the episcopacy (possibly in favor of Pius, though this is not clear), whereupon he left the church. Tertullian also mentions large numbers of followers (*frequentissimum*

plane collegium inter haereticos, *Adv. Val.* I), some of whom appear to have founded movements of their own, for example, Theodotus, Heracleon, Florinus, Ptolemaeus, and Marcus—these last two serving as particular targets for St. Irenaeus.

The philosophical and theological system of Valentinus bears some similarities to Platonism, though it has also been suggested, with much less plausibility, that his system was founded upon principles drawn from the Ophites, a Gnostic sect particularly devoted to the role of the serpent as metaphor and, in some cases, object of worship. If there were Pythagorean elements, as has also been suggested, they have been very cleverly disguised. Like Pythagoreanism, however, we may say that Valentinianism as we know it comes primarily from the writings of his disciples (and from his critics among the fathers) rather than from any writings of his own that have come down to us. It is possible to divide his followers into two "schools," one in the East (the "Anatolian" or "Oriental") and one in the West (the "Italian"). It has been alleged by some scholars that the Eastern school better preserved the teachings of Valentinus himself, but of course in the absence of empirical data it is impossible to make such a judgment without begging the question. More is said about the connection to Platonism below.

The Valentinians posit a primal being, Bythos (from the Greek *buthos*, "the depth," or "abyss"), who existed before all else, though in some sources he is portrayed as eternally coexisting with the Silence or Contemplation that is his thought. From this primordial pair arose, by emanation, three "syzygies" (Greek *suzugia*, "pair"), pairs of beings known as "aeons" (Greek *aiôn*, literally "age" or "generation" but also personified as a title for a divine being), which may have been conceived as aspects of divinity, though this interpretation possibly reflects a Trinitarian influence that may have been alien to Valentinus. (Some evidence suggests that Valentinus tried to remain in communion with the church, in which case he may have tried to formulate his ideas in a manner conducive to orthodoxy; on the other hand, the refutations of his followers would have been put into the terms and relations most natural to the orthodox writers of the refutations.)

The syzygies themselves represent cosmological opposites such as male and female, and it may be this aspect of the system that has suggested to some a Pythagorean influence. From this first triad of syzygies emanate other aeons, until there are thirty in all. These fifteen syzygies of thirty aeons make up the so-called *pleroma* (Greek *plêrôma*, "fullness," or "satiety"), a realm of immaterial, spiritual being. The last aeon to arise by emanation from the original triad is Sophia who, being farthest from the source of Being, managed through weakness to fall into sin and produce an offspring, Achamoth. If we care to take the comparisons with Platonic metaphysics seriously, we may note that Achamoth appears to represent a metaphysical principle of mimesis, for it creates a rival world, the *kenoma* (Greek *kenôma* "emptiness," or "vacuum"), in imitation of the *pleroma*, and a rival being, the Demiurge, in imitation of Bythos.

The Demiurge is clearly intended to be the God of the Old Testament, since he sets about creating the heavens and the earth of Genesis and everything in them. In particular, he creates humankind out of matter (Greek *hulê*) by imparting into it something of his own psychic substance (Greek *psukhê*). In addition to these two aspects of humankind, the "psychic" and the "hylic," a third, spiritual element, the "pneumatic" (Greek *pneuma*), was incorporated into our nature, apparently without the Demiurge's knowledge.

As in other Gnostic systems, humankind falls into classes that depend upon the degree to which members of the class have access to the saving knowledge (Greek *gnôsis*) that will enable them to escape the temporally finite material existence of the *kenoma* and enter into the eternal bliss of the *pleroma*. In the Valentinian system there are three classes: the *pneumatikoi* (that is, the Valentinians themselves) represent the spiritual, or highest, class, to whom full *gnôsis* has been given; the lowest class, the *hulikoi*, are those whose material aspect dominates and who are thus doomed never to escape from the *kenoma* and who will be destroyed along with it at the end of time; somewhere between lie the *psukhikoi*, or "psychics," the non-Valentinian Christians who can attain a kind of pseudo-salvation by means of faith and good works that will enable them to enter into the same plane of existence as the Demiurge. Christ is an aeon among the original thirty who unites himself (either at conception or at baptism) with the human Jesus of Nazareth (who is present only in a docetic sense), who is then the first to bring *gnôsis* to the rest of humankind.

Apart from the role of the Christ aeon and Jesus of Nazareth, there is little here to suggest Christian origins, in spite of Valentinus's reported desire to remain in communion with the orthodox church, and this fact has prompted some scholars to suggest that the Valentinians were, in fact, merely borrowing from pagan versions of Gnosticism. However, as with the connections to the Ophites, the Platonists, and the Pythagoreans, this is mere speculation. The evidence regarding Valentinus himself is

so thin, and that regarding the Valentinian schools so varied and contradictory, that it is quite difficult, if not impossible, to make any clear and non-circular case for the influences and origins of any aspect of the system as a whole. At best, similarities to other philosophical systems can be noted, but it is difficult to draw any secure conclusions about influences. The putative connection to Platonism, for example, clearly lies in the positing of two "realms," one ideal and the other material, with different sorts of beings inhabiting each and the material representing a kind of "falling away" from the ideal; but this kind of metaphysical system can be found in Jewish thought that either predates or is fully independent of Platonism. Of greater significance would seem to be Achamoth as a principle of mimesis, but that construal of his role in the system is already an interpretation beyond what can be found in the actual Valentinian texts, and it cannot serve to establish a definite link with Platonist thought. Similarly, it is perhaps tempting to see Pythagorean "dyads" in the Valentinian syzygies, but mere parallelism is insufficient to establish genuine borrowing.

See also Gnosticism; Platonism and the Platonic Tradition; Pythagoras and Pythagoreanism.

Bibliography

TEXTS AND TRANSLATIONS

Clement of Alexandria. *Stromateis*, edited by J. Ferguson. Washington, DC: Catholic University Press of America, 1991.

The Facsimile Edition of the Nag Hammadi Codices. Department of Antiquities, Arab Republic of Egypt in Conjunction with UNESCO. Leiden: E. J. Brill, 1972–1984.

Foester, W., and R. McL. Wilson, eds. *Gnosis: A Selection of Gnostic Texts*. Oxford: Clarendon Press, 1972–1974.

Hippolytus. *Refutatio omnium Haeresium*, edited by M. Marcovich.Berlin: De Gruyter, 1986.

Irenaeus. *Adversus omnes Haereses* (*Elegkos kai Anatropê tês Pseudônumou Gnôseôs*), editions by A. Stieren (Leipzig: T. O. Weigel, 1848–1853) and W. Harvey (Cambridge: Cambridge University Press, 1857); critical text by A. Rousseau, L. Doutreleau, et al. (*Sources Chrétiennes*, 100, 152 f., 210 f., 263 f., and 293 f. Paris: Éditions du Cerf, 1965–1982).

Layton, Bentley. *The Gnostic Scriptures: A New Translation with Annotations and Introductions*. London: SCM, 1987.

Nag Hammadi Studies. Leiden: E. J. Brill, 1975– (for main editions of texts).

Robinson, James M., ed. *The Nag Hammadi Library in English*. Leiden and New York: E. J. Brill, 1977, 1996.

Tertullian. *Adversus Valentinianos*, edited by J.-C. Fredouille, *Sources Chrétiennes* 280, 281. Paris: Éditions du Cerf, 1980–1981.

Völker, W. *Quellen zur Geschichte der christlichen Gnosis*. Tübingen: Mohr, 1932, esp. pp. 57–141.

STUDIES

Barc, Bernard. *Colloque international sur les textes de Nag Hammadi*. Québec: Presses de l'Université Laval, 1981.

Bermejo Rubio, Fernando. *La escisión imposible: Lectura del gnosticismo valentiniano*. Salamanca: Universidad Pontificia, 1998.

Cross, F. L., ed. *The Jung Codex*. London: Mowbray, 1955.

Davison, J. E. "Structural Similarities and Dissimilarities in the Thought of Clement of Alexandria and the Valentinians." *Second Century* 3 (1983): 201–217.

Dawson, David. *Allegorical Readers and Cultural Revision in Ancient Alexandria*. Berkeley: University of California Press, 1992.

DeConick, April D. "The Great Mystery of Marriage: Sex and Conception in Ancient Valentinian Traditions." *Vigiliae Christianae* 57 (2003): 307–342.

Desjardins, Michel. "Reconstructing Valentinian Families." In *Religionswissenschaft in Konsequenz: Beiträge im Anschluss an Impulse von Kurt Rudolph*, edited by Rainer Flasche et al. Münster: LIT Verlag, 2000.

Desjardins, Michel. *Sin in Valentinianism*. Atlanta, GA: Scholars Press, 1990.

Desjardins, Michel. "The Sources for Valentinian Gnosticism. A Question of Methodology." *Vigiliae Christianae* 40 (1986): 342–347.

Edwards, Mark J. "Gnostics and Valentinians in the Church Fathers." *Journal of Theological Studies* 40 (1989): 26–47.

Good, D. J. "Sophia in Valentinianism." *Second Century* 4 (1984): 193–201.

Holzhausen, Jens. "Irenäus und die valentinianische Gnosis." *Vigiliae Christianae* 55 (2001): 341–355.

Layton, Bentley. *The Rediscovery of Gnosticism*. Proceedings of the International Conference on Gnosticism at Yale, New Haven, CT, March 28–31, 1978. Leiden: E. J. Brill, 1980.

Markschies, Christoph Johannes. *Valentinus Gnosticus?: Untersuchungen zur valentinianischen Gnosis: mit einem Kommentar zu den Fragmenten Valentins*. Tübingen: Mohr, 1991.

McGuire, Anne M. "Valentinus and the *Gnostike Hairesis*: Irenaeus, *Haer*. I.XI.l and the Evidence of Nag Hammadi, XVIII,1." *Studia Patristica* 18, Papers of the Ninth International Conference on Patristic Studies, Oxford, 1983, edited by Elizabeth A. Livingstone, 247–252. Kalamazoo, MI: Cistercian Publications, 1985.

Pagels, Elaine. "Irenaeus, the 'Canon of Truth', and the 'Gospel of John'": 'Making a Difference'; Through Hermeneutics and Ritual." *Vigiliae Christianae* 56 (2002): 339–371.

Pagels, Elaine. "A Valentinian Interpretation of Baptism and Eucharist, and its Critique of Orthodox Sacramental Theology and Practice." *Harvard Theological Review* 65 (1972): 153–170.

Poirier, Paul-Hubert. "Pour une histoire de la lecture pneumatologique de Gn 2, 7." *Revue des Études Augustiniennes* 40 (1994): 1–22.

Quispel, Gilles. "The Original Doctrine of Valentinus the Gnostic." *Vigiliae Christianae* 50 (1996): 327–352.

Quispel, Gilles. "Valentinus and the *Gnostikoi*." *Vigiliae Christianae* 50 (1996): 1–4.

Sagnard, Francois. *La Gnose valentinienne et le témoignage de Saint Irénée*. Paris: J. Vrin, 1947.

Simonetti, Manlio. "Valentinus gnosticus." *Cassiodorus* 1 (1995): 197–205.

Stead, G. C. "The Valentinian Myth of Sophia." *Journal of Theological Studies* 20 (1969): 75–104.

Townsley, A. L. "St. Irenaeus' Knowledge of Presocratic Philosophy." *Rivista di Storia e Letteratura Religiosa* 12 (1976): 374–379.

Scott Carson (2005)

VALÉRY, PAUL

(1871–1945)

As a law student in Montpellier, Valéry published poems and befriended such influential authors as André Gide and Stéphane Mallarmé. As a result of a personal crisis in 1892, he resolved to abandon literature and devote himself to his autodidactical pursuit of knowledge. While serving in the Ministry of War, and then as private secretary to a powerful businessman, Valéry found time to read and write. In 1894 he began the first of some 261 notebooks in which he developed his matinal reflections for over fifty years. At Gide's instigation Valéry began to prepare a volume of poems, and ended up writing *La jeune parque* (The young fate) (1917), a hermetic allegory of consciousness that established him as an eminent French poet. In 1927 Valéry was elected as a member of the French Academy. He went on to lecture and write about an astounding array of topics, including science, history, architecture, dance, the visual arts, literature, politics, globalization, modern warfare, psychology, and moral philosophy. His achievement includes volumes of poetry, melodramas written to the music of Arthur Honnegger, philosophical dialogues, and numerous collections of essays and aphorisms. A chair in poetics was created for Valéry at the Collège de France in 1936.

Valéry's relation to philosophy was ambivalent. The philosopher, he ironizes, is a "specialist of the universal" (*Oeuvres*, vol. 1, p. 1235). And the universal is only what is "grossier" (coarse or crude) enough to be so (*Oeuvres*, vol. 2, p. 881). The philosopher is an artist who does not admit it. Every abstract theory is at bottom a fragment of an autobiography. Words that serve people perfectly well in ordinary transactions become the object of infernal, Sisyphean labors when philosophers wrongheadedly take words as ends instead of means and look for their ultimate meanings. Words are like a board thrown across an abyss; we can cross over if we move quickly, but not if we linger and test the board's strength. As the past no longer exists, the idea of historical truth is problematic. Origins are elusive, and "everything begins as an interruption" (*Oeuvres*, vol. 2, p. 881).

In spite of his misgivings about philosophical generalizations, Valéry did elaborate various philosophical theses, especially in aesthetics. He critiques inspirationist models of artistic creation; moments of inspiration can only produce fragments. The making of artworks is always a combination of deliberate and spontaneous processes, only their proportion varies. Appreciating a work requires the imaginative reconstruction of the creative process. Yet the creator's thoughts about a work's meaning have no special privilege. In literature, language is an end in itself. Poetry is to prose as dancing is to walking. To describe or sum up a work—in five hundred words or more—is necessarily to fail to convey what is most essential to it.

See also Aesthetics, History of; Aesthetics, Problems of; Philosophy of Language.

Bilbiography

PRIMARY WORKS

The Collected Works of Paul Valéry, edited by Jackson Mathews. 15 vols. Princeton, NJ: Princeton University Press, 1956–1975.

Oeuvres, edited by Jean Hytier. 2 vols. Paris: Gallimard, 1957–1960.

Cahiers. 29 vols. Paris: Centre National de la Recherche Scientifique, 1957–1961.

SECONDARY WORKS

"Selected Bibliography." Valéry Studies, University of Newcastle upon Tyne. Available from http://www.paulvalery.org.

Paisley Livingston (2005)

VALLA, LORENZO

(1407–1457)

Lorenzo Valla, the Italian humanist, is best known as the man who exposed the Donation of Constantine and thus undermined a leading argument for papal sovereignty in the secular realm. This fact and the reputation for hedonism derived from his youthful work *De Voluptate* (On pleasure) have conspired to invest Valla with an air of disrepute that he probably does not deserve. In particular, this reputation does not do justice to Valla's efforts on behalf of a return to the spirit of the Gospel or to his respect for Paul and the early Greek and Latin Church Fathers, in which he clearly anticipates later developments. Nor does it recognize his passion for historical

truth and for the defense of plain speaking against what he regarded as metaphysical obscurity and verbalizing. Valla was perhaps the most versatile of the humanists; he initiated a series of attacks upon Scholastic logic, theology, and law, in addition to his contributions to historical and textual criticism.

Valla was above all a brilliant philologian and a staunch champion of the new humanities; most of his writing is best understood from this point of view. Valla was born in Rome. He learned Latin and Greek there and perhaps in Florence, and he spent three formative years, from 1431 to 1433, teaching rhetoric at the University of Pavia. Pavia was a lively center of humanists, and it may have been here that Valla heard the discussions of ancient ethics that prompted him to write the earliest of his extant works, the dialogue generally known under the title "On Pleasure" (Valla actually called it "On the True Good"). Several versions of this dialogue appeared, with the speeches variously assigned to different contemporaries of Valla. Contrary to a widespread impression, Valla does not directly endorse Epicurean ethics in the work; he permits speakers to present Stoic and Epicurean ethics and then, in the person of a third speaker, criticizes their views from a Christian standpoint. This third speaker clearly represents the convictions of Valla himself. The Stoic spokesman presents a defense of Stoic *honestas* or virtue, together with a quite un-Stoic complaint against nature, "which has made men so prone to vice." An Epicurean replies, at much greater length, in defense of nature and "utility." Utility is equated with pleasure and described as a mistress among her handmaidens, the virtues, rather than as a harlot among honest matrons. The third speaker criticizes both of his predecessors and argues that the true Christian should disregard the goals of this life and concentrate on the joys that await him in Heaven. However, this speaker accepts without challenge the equating of "the useful" with pleasure; he insists only that the pleasures a Christian should pursue are not those of this world. Thus, despite his rejection of Epicurean morality, Valla's description of heavenly pleasures is more graphic than we are accustomed to expect from a Christian writer. Renaissance joie de vivre is allowed to assert itself only in a future life. Does Valla depart radically from earlier Christian doctrine, or does he simply make explicit what would constitute the traditional Christian hope if it were spelled out? Obviously there is room for disagreement here, but there can be no disagreeing with the view of the eminent historian Eugenio Garin that Valla's work on pleasure represents a major Renaissance document.

After sojourns in various Italian cities, Valla entered the service of King Alfonso of Aragon, with whom he remained from 1435 to 1448. During this time in Naples, and probably in connection with Alfonso's quarrels with the pope, Valla wrote his most renowned work—his exposure as a forgery of the supposed Donation of the Emperor Constantine of the Western Empire to Pope Sylvester. Although he was anticipated in this by several earlier writers, among them Nicholas of Cusa, Valla's treatise stands out as a very effective piece of historical criticism and, incidentally, a strong plea for the spiritual purity of the Holy See. In view of the latter it should not appear surprising that Valla was later accepted into the pontifical secretariat and spent the remaining years of his life in Rome. The genuineness of Valla's respect for historical truth and his scorn for superstition is shown in such statements as this in the treatise on the Donation: "A Christian man who calls himself the son of light and truth ought to be ashamed to utter things that not only are not true but are not even likely."

While with King Alfonso, Valla also wrote a work on free will, *De Libero Arbitrio*, in which he takes issue with Boethius's treatment of free will in the *Consolation of Philosophy*. In his dialogue Valla distinguishes God's foreknowledge, which cannot be said to be the cause of our volitions, from his will. God's accurate prediction that Judas will become a traitor does not excuse Judas. But Valla refuses to deal with the further question of whether God's will, which cannot be denied, takes away human choice. The divine will, he argues, is known neither to men nor to angels; we stand by faith, not by the probability of reasons.

A similar reluctance to engage in argumentative philosophizing appears in the treatise *Dialectic*, an attack upon conventional Aristotelian logic, printed a half-century after Valla's death. Valla here pleads for the elimination of empty subtleties and vain word-juggling. "Let us conduct ourselves more simply and more in line with natural sense and common usage," he says. "Philosophy and dialectic … ought not to depart from the most customary manner of speaking." Valla's treatment of the Aristotelian categories is not without interest. The Latin word for entity (*entitas*), for example, is simply a coinage of a participle from the verb "to be" that does not occur in standard Latin and hence ought to be regarded with suspicion. To say that a stone is an entity (*lapis est ens*) amounts to no more than saying that it is a thing (*res*), which is perfectly satisfactory and more clear. Therefore, Aristotle's metaphysics, which deals with "being qua being," is meaningless, suggesting as it does that what "is"

is "able not to be." Having protested the positing of mysterious entities, quiddities, and essences and having equated substances with bodies or things, Valla then reduces the remaining nine categories of Aristotle to two: quality and action. Definitions, according to Valla, are explications of all the qualities and actions that are present in a thing. In the course of his exposition, Valla has occasion to challenge the validity of many scholastic distinctions: for example, those between the concrete and the abstract, between matter and form, and so on. Unsatisfactory as Valla's own offerings may be (they are not clearly dedicated to the solution of any specific philosophical problems), nevertheless it must be admitted that a fresh consideration of technical terms was certainly called for at the time and was eventually carried through by later critics.

Valla displays great sensitivity to nuances of meaning in his *Elegantiae Linguae Latinae* (Elegancies of the Latin language), in which he makes careful analyses of the usage of many Latin terms. Critics have observed that Valla's own style was not as elegant as it could have been, but his advice was widely consulted.

Valla was often accused of bad form in his attacks on people and schools of thought, but one must recall that invectives and ad hominem attacks were the order of the day. In the Renaissance professional rivalry did not bother to conceal itself under polite or semipolite discussions of issues. Valla defended himself against the charge of malevolence and vindictiveness in a letter to Giovanni Serra, in which he concludes: "I do not censure all authors, but only a few, … not all philosophers but some from all sects, not the best but the worst, not impudently but calmly, ready to accept correction should it prove valid."

See also Aristotelianism; Aristotle; Epicureanism and the Epicurean School; Hedonism; Humanism; Italian Philosophy; Nicholas of Cusa; Renaissance; Stoicism.

Bibliography

Valla's works were collected in an edition that was published in Basel (1540). It has been reproduced, along with some treatises and letters missing from the earlier edition, in an edition with an introduction by Eugenio Garin (Turin, 1962). The standard biography is still Girolamo Mancini's *Vita di Lorenzo Valla* (Florence, 1891). Valla's treatise on free will is available in English in *The Renaissance Philosophy of Man*, edited by Ernst Cassirer, Paul Oskar Kristeller, and John H. Randall Jr. (Chicago: University of Chicago Press, 1948). *The Treatise of Lorenzo Valla on the Donation of Constantine* has been translated by C. B. Coleman (New Haven, CT: Yale University Press, 1922). The treatise on pleasure was published under the title *De verro falsoque bono* by Maristella De Panizza Lorch (Bari, Italy: Adriatica, 1970).

OTHER RECOMMENDED WORKS

Camporeale, Salvatore I. "Lorenzo Valla's 'Oratio' on the Pseudo-Donation of Constantine: Dissent and Innovation in Early Renaissance Humanism." *Journal of the History of Ideas* 57 (1) (1996): 9–26.

Connell, William J. "Lorenzo Valla: A Symposium." *Journal of the History of Ideas* 57 (1) (1996): 1–7.

Jardine, Lisa. "Lorenzo Valla: Academic Skepticism and the New Humanist Dialectic." In *The Skeptical Tradition*, edited by Myles Burnyeat. Berkeley: University of California Press, 1983.

Jardine, Lisa. "Lorenzo Valla and the Intellectual Origins of Humanist Dialectic." *Journal of the History of Philosophy* 15 (1977): 143–164.

Johnson, Lawrence J. "The Linguistic Imperialism of Lorenzo Valla and the Renaissance Humanists." *Interpretation* 7 (1978): 29–49.

Monfasani, John. "Lorenzo Valla and Rudolph Agricola." *Journal of the History of Philosophy* 28 (2) (1990): 181–200.

Monfasani, John. "Was Lorenzo Valla an Ordinary Language Philosopher?" *Journal of the History of Ideas* 50 (1989): 309–323.

Paganini, Gianni. "Hobbes, Valla and the Trinity." *British Journal for the History of Philosophy* 11 (2) (2003): 183–218.

Synan, Edward A. "Boethius, Valla, and Gibbon." *Modern Schoolman* 69 (3–4) (1992): 475–491.

Trinkaus, Charles. "Lorenzo Valla on the Problem of Speaking about the Trinity." *Journal of the History of Ideas* 57 (1) (1996): 27–53.

Neal W. Gilbert (1967)
Bibliography updated by Tamra Frei (2005)

VALUE AND VALUATION

The terms *value* and *valuation* and their cognates and compounds are used in a confused and confusing but widespread way in our contemporary culture, not only in economics and philosophy but also and especially in other social sciences and humanities. Their meaning was once relatively clear and their use limited. *Value* meant the worth of a thing, and *valuation* meant an estimate of its worth. The worth in question was mainly economic or quasi economic, but even when it was not, it was still worth of some sort—not beauty, truth, rightness, or even goodness. The extension of the meaning and use of the terms began in economics, or political economy, as it was then called. *Value* and *valuation* became technical terms central to that branch of economics which was labeled the theory of value. Then German philosophers, especially Rudolf Hermann Lotze, Albrecht Ritschl, and Friedrich Nietzsche, began to take the notion of value and

values in a much broader sense and to give it primary importance in their thinking.

Philosophers from the time of Plato had discussed a variety of questions under such headings as the good, the end, the right, obligation, virtue, moral judgment, aesthetic judgment, the beautiful, truth, and validity. In the nineteenth century the conception was born—or reborn, because it is essentially to be found in Plato—that all these questions belong to the same family, since they are all concerned with value or what ought to be, not with fact or what is, was, or will be. All these questions, it was believed, may not only be grouped under the general headings of value and valuation but are better dealt with and find a more systematic solution if they are thought of as parts of a general theory of value and valuation that includes economics, ethics, aesthetics, jurisprudence, education, and perhaps even logic and epistemology. This conception matured in the 1890s in the writings of Alexius Meinong and Christian von Ehrenfels, two Austrian followers of Franz Brentano. Through them and through others like Max Scheler and Nicolai Hartmann, two twentieth-century German followers of Edmund Husserl (himself influenced by Brentano), the idea of a general theory of value became popular on the Continent and in Latin America. It had some influence in Great Britain, in the works of Bernard Bosanquet, W. R. Sorley, J. M. Mackenzie, John Laird, and J. N. Findlay, but rather less than elsewhere, for, on the whole, British philosophers have held to more traditional terms such as *good* and *right.* But it received an excited welcome in the United States just before and after World War I. The idea was introduced by Hugo Münsterberg and W. M. Urban, taken up by Ralph Barton Perry, John Dewey, D. H. Parker, D. W. Prall, E. W. Hall, and others, and later refurbished by S. C. Pepper and Paul W. Taylor. This wide-ranging discussion in terms of *value, values,* and *valuation* subsequently spread to psychology, the social sciences, the humanities, and even to ordinary discourse.

PHILOSOPHICAL USAGES

The uses of *value* and *valuation* are various and conflicting even among philosophers, but they may perhaps be sorted out as follows. (1) *Value* (in the singular) is sometimes used as an abstract noun (*a*) in a narrower sense to cover only that to which such terms as *good, desirable,* or *worthwhile* are properly applied and (*b*) in a wider sense to cover, in addition, all kinds of rightness, obligation, virtue, beauty, truth, and holiness. The term can be limited to what might be said to be on the plus side of the zero line; then what is on the minus side (bad, wrong, and so forth) is called disvalue. *Value* is also used like *temperature* to cover the whole range of a scale—plus, minus, or indifferent; what is on the plus side is then called positive value and what is on the minus side, negative value.

In its widest use *value* is the generic noun for all kinds of critical or pro and con predicates, as opposed to descriptive ones, and is contrasted with existence or fact. The theory of value, or axiology, is the general theory of all such predicates, including all the disciplines mentioned above. The classic example in English of this approach is the work of R. B. Perry. In its narrower use, *value* covers only certain kinds of critical predicates and is contrasted with descriptive predicates and even with other critical ones like rightness and obligation. In this case the theory of value, or axiology, is a part of ethics, rather than the other way around. The work of C. I. Lewis is the best example of the narrower approach.

Those who take the wider approach sometimes distinguish "realms of value"; Perry and Taylor, for example, list eight of these: morality, the arts, science, religion, economics, politics, law, and custom or etiquette. Even when *value* is used in the narrower sense, several meanings of the term, or kinds of value, are sometimes distinguished. (The narrower distinctions may also be recognized by those who use value in the wider sense.) These meanings correspond to the senses or uses of *good,* which G. H. von Wright prefers to call "forms" or "varieties of goodness." Many classifications of kinds of value, or forms of goodness, have been proposed. Lewis distinguishes (*a*) utility or usefulness for some purpose; (*b*) extrinsic or instrumental value, or being good as a means to something desirable or good; (*c*) inherent value or goodness, such as the aesthetic value of a work of art in producing good experiences by being contemplated or heard; (*d*) intrinsic value, or being good or desirable either as an end or in itself, which is presupposed by both (*b*) and (*c*); (*e*) contributory value, or the value that an experience or part of an experience contributes to a whole of which it is a part (not a means or an object). A stick of wood may be useful in making a violin, a violin may be extrinsically good by being a means to good music, the music may be inherently good if hearing it is enjoyable, the experience of hearing it may be intrinsically good or valuable if it is enjoyable for its own sake, and it may also be contributively good if it is part of a good evening or weekend.

Dewey, however, attacks the distinction between means and ends while stressing the notion of total value or goodness on the whole—goodness when all things are considered. To Lewis's list of kinds of value, some writers, W. D. Ross for instance, would add moral value, the kind

of value or goodness that belongs to a virtuous man, to good motives, or to morally approved traits of character. Von Wright distinguishes instrumental goodness (a good knife), technical goodness (a good driver), utilitarian goodness (good advice), hedonic goodness or pleasantness (a good dinner), and welfare (the good of man). He also mentions moral goodness but argues that it is a subform of utilitarian goodness; Ross would deny this.

(2) *Value* as a more concrete noun—for example, when we speak of "a value" or of "values"—is often used (*a*) to refer to what is valued, judged to have value, thought to be good, or desired. The expressions "his values," "her value system," and "American values" refer to what a man, a woman, and Americans value or think to be good. Such phrases are also used to refer to what people think is right or obligatory and even to whatever they believe to be true. Behind this widespread usage lies the covert assumption that nothing really has objective value, that *value* means being valued and *good* means being thought good. But the term *value* is also used to mean (*b*) what has value or is valuable, or good, as opposed to what is regarded as good or valuable. Then *values* means "things that have value," "things that are good," or "goods" and, for some users, also things that are right, obligatory, beautiful, or even true.

In both usage (*a*) and usage (*b*) it is possible to distinguish different kinds of values, corresponding to the different kinds of value or forms of goodness mentioned above. It is also common to distinguish more or less clearly between material and spiritual values or among economic, moral, aesthetic, cognitive, and religious values.

Some philosophers, especially those influenced by Scheler and Hartmann, think of *value* as a general predicate like "color," which subsumes more specific value predicates analogous to "red" or "yellow." They call these more specific value predicates "values" (*Werte, valeurs*). Just as "a color" does not mean "a thing that has color" but a particular color like red, so "a value" does not mean "a thing that has value" but a particular kind of value, like pleasure value or courage value. These philosophers call a thing that is good "a good" or "a value carrier," not "a value." Since the adjective *valuable* simply means "having value" or "being good" in some sense (or, perhaps better, "having a considerable amount of value"), much of the above will apply to it, mutatis mutandis.

(3) *Value* is also used as a verb in such expressions as "to value," "valuating," and "valued." *Valuing* is generally synonymous with *valuation* or *evaluation* when these are used actively to mean the act of evaluating and not passively to mean the result of such an act. But sometimes *valuation* and *evaluation* are used to designate only a certain kind of valuing, namely, one that includes reflection and comparison. In either case *valuation* may be, and is, used in wider or narrower senses corresponding to the wider and narrower uses of *value.* For Dewey and Richard M. Hare it covers judgments about what is right, wrong, obligatory, or just, as well as judgments about what is good, bad, desirable, or worthwhile. For Lewis *valuation* covers only the latter use. The expression "value judgment" is also used in both of these ways. Among the writers who distinguish two main kinds of normative discourse, evaluating and prescribing, some, like Taylor, classify judgments of right and wrong as well as judgments of good and bad under evaluations and judgments, using *ought* under prescriptions; others put judgments of right and wrong under prescriptions.

Dewey always distinguishes two senses of "to value." It means either (*a*) to prize, like, esteem, cherish, or hold dear, or (*b*) to apprize, appraise, estimate, evaluate, or valuate. In the second sense reflection and comparison are involved; in the first sense they are not. In the first sense, he seems to regard mere desiring or liking as a form of valuing. Others often follow him in this, but some writers limit valuing to acts in which something is not merely desired or liked but judged to be good or to have value. Even Perry, who holds that the statement "*X* is good" = "*X* has positive value" = "*X* is an object of favorable interest," insists that we must distinguish between desiring *X* and judging *X* to have value, which would be judging *X* to be desired.

Thus, words such as *value* and *valuation* may be, and are, used in a variety of ways, even when they are used with some care—which is, unfortunately, not often the case both in and out of philosophy. In using the terms, one should choose a clear and systematic scheme and use it consistently. Because of the ambiguity and looseness that the terms often engender, it would seem advisable to use them in their narrower senses or not at all, keeping to more traditional terms such as *good* and *right,* which are better English, whenever possible.

PHILOSOPHICAL THEORIES

Philosophical theories of value and valuation, whether conceived in the wider or in the narrower manner and whether formulated in the traditional or in the newer "value" vocabulary, have been of two sorts. Normative theories make value judgments or valuations; they tell us what is good or what has value, what is bad, and so on. Metanormative theories analyze value, valuation, and

good; they neither make value judgments in this way nor tell us what is good or has value. Instead, they define what goodness and value are and what it means to say that something is good or has value. Sometimes philosophers also offer descriptive generalizations about what is valued or regarded as good in some culture or group of cultures, and explanatory theories about why this is so valued or regarded (David Hume, Moritz Schlick, F. C. Sharp, John Ladd). However, this is usually ancillary to their discussions of normative or metanormative questions. In themselves such descriptive and explanatory theories belong to anthropology, psychology, and sociology, not to philosophy. Recently, many analytical philosophers have been maintaining that even normative theories, however important they may be, have no place in philosophy proper, where theories of value and valuation should be limited to metanormative questions.

NORMATIVE THEORIES. In the broader conception, a normative theory of value must show, at least in general outline, what is good, bad, better, and best, and also what is right, obligatory, virtuous, and beautiful. In the narrower conception, normative theories of value have usually addressed themselves primarily to the question of what is good in itself or as an end or what has intrinsic value, an approach that Dewey has persistently attacked. They ask not what goodness and intrinsic value are but what the good is, what has value for its own sake, what is to be taken as the end of our pursuit or as the criterion of intrinsic worth.

Some theories have answered that the end or the good is pleasure or enjoyment or, alternatively, that the criterion of intrinsic value is pleasantness or enjoyableness. More accurately, they say that only experiences are intrinsically good, that all experiences that are intrinsically good are pleasant and vice versa, and that they are intrinsically good because and only because they are pleasant. These are the hedonistic theories of value, held by such thinkers as Epicurus, Hume, Jeremy Bentham, J. S. Mill, Henry Sidgwick, von Ehrenfels, Meinong (at first), and Sharp. There are also quasi-hedonistic theories in which the end or the good is said to be not pleasure but something very similar, such as happiness, satisfaction, or felt "satisfactoriness," to use Lewis's term. Examples are to be found in the writings of Dewey, Lewis, Parker, P. B. Rice, and perhaps Brand Blanshard.

Antihedonistic theories are of two kinds. Some agree that there is, in the final analysis, only one thing that is good or good-making but deny that it is pleasure or any other kind of feeling. Aristotle says it is eudaemonia (excellent activity); Augustine and Thomas Aquinas, communion with God; Benedict de Spinoza, knowledge; F. H. Bradley, self-realization; Nietzsche, power. Others, such as Plato, G. E. Moore, W. D. Ross, Laird, Scheler, Hartmann, and Perry, are more "pluralistic," holding that there are a number of things that are good or good-making in themselves. They differ in their lists but all include two or more of the following: pleasure, knowledge, aesthetic experience, beauty, truth, virtue, harmony, love, friendship, justice, freedom, self-expression. Of course, hedonists and other "monistic" thinkers may also regard such things as intrinsically good, but only if and because they are pleasant, self-realizing, or excellent.

METANORMATIVE THEORIES. The scope of metanormative theories may also be inclusive or limited, but both kinds will pose similar questions and offer similar answers. Their questions and answers have been variously stated in the formal or material mode, or the linguistic or nonlinguistic, but they will not be classified here.

One question or group of questions posed by metanormative theories concerns the nature of value and valuation: what is goodness or value? what is the meaning or use of *good?* what is valuing? what are we doing or saying when we make a value judgment? A subquestion here is what moral value and evaluation are, and how they are distinct from nonmoral value and valuation, if at all. Another question or set of questions has to do with the justification or validity of value judgments and normative theories: can they be justified or established with any certainty by some kind of rational or scientific inquiry? can they be shown to have objective validity in any way? if so, how? what is the logic of reasoning in these matters, if there is one? Here a subquestion is what is the logic of moral justification or reasoning, if there is one, and is it in any way distinctive. Beyond this there is an even more "meta" level of questioning: what is the nature of a metanormative theory, and how can it be defended? This last problem, as well as the subquestions just mentioned, has frequently been discussed in the twentieth century and earlier but will not be considered here.

In reply to the first question or group of questions, some philosophers have held that terms like *value* and *good* stand for properties; that in value judgments we are ascribing these properties to objects or kinds of objects (including activities and experiences), although we may also be taking pro or con attitudes toward them; and that, therefore, value judgments are descriptive or factual in the sense of truly or falsely ascribing properties to things. They are therefore cognitivists or descriptivists in value

theory. Of these the naturalists add that the property involved is a natural or empirical one, which can be defined. Aristotle, von Ehrenfels, and Perry claim that value is the relational property of being an object of desire or interest (an interest theory of value); Parker, that it is the satisfaction of desire (another interest theory of value); Lewis and Rice (as well as the early Meinong), that it is the quality of being, enjoyed or enjoyable in some way (the affective theory of value). George Santayana seems sometimes to hold one of these views, sometimes another, and sometimes to regard value as an indefinable natural quality ascribed to what we desire or enjoy.

Other cognitivists add that value or goodness is a metaphysical property that can neither be observed by or in ordinary experience nor made an object of empirical science. Examples of metaphysical definitions are being truly real (Neoplatonists), being ontologically perfect (Hegelian idealists), or being willed by God (theologians). Still others assert that intrinsic goodness or value is an indefinable nonnatural or nonempirical quality or property different from all other descriptive or factual ones (they even describe it as being nondescriptive or nonfactual). These philosophers are called intuitionists or nonnaturalists (Plato, Sidgwick, Moore, Ross, Laird, Scheler, Hartmann, and perhaps the later Meinong). They all hold that value belongs to objects independently of whether we desire, enjoy, or value them, and even independently of God's attitude toward them—as some metaphysical theorists and naturalists also do. Meinong, Scheler, Hartmann, and Hall contend that value is intuited through the emotions even though it is objective; Sidgwick, Ross, Laird, and others, that it is an object of intellectual intuition.

In the mid-twentieth century many writers, both analytical philosophers and existentialists, have taken the position that value terms do not stand for properties, natural or nonnatural, and that value judgments are not property-ascribing statements but have some other kind of meaning or function. These writers have therefore been called noncognitivists or antidescriptivists. Their positive theories are varied. Some argue that value judgments are wholly or primarily embodiments or expressions of attitude, emotion, or desire, and/or instruments for evoking similar reactions in others (A. J. Ayer, Bertrand Russell, Charles L. Stevenson). Others maintain that this account of value terms and judgments is inadequate and that value judgments are to be thought of as prescriptions, recommendations, acts of grading, or simply as valuations, not something else (Hare, Taylor, Stephen E. Toulmin, Patrick H. Nowell-Smith, R. W. Sellars, and J. O. Urmson).

Whether value judgments are susceptible to being justified or proved, and, if so, how, depends very considerably on the position taken in answer to the questions regarding the meaning of *good.* Some value judgments are derivative—for instance, the conclusion of the following inference:

What is pleasant is good.
Knowledge is pleasant.
Therefore, knowledge is good.

The real question is about the justification of basic or nonderivative value judgments. According to the intuitionist, such judgments cannot be justified by argument, but they do not need to be, since they are intuitively known or self-evident. According to the naturalist, they can be established either by empirical evidence (in Perry's view, by empirical evidence about what is desired) or by the very meaning of the terms involved (analytically or by definition). According to the metaphysical and theological axiologist, they can be established either by metaphysical argument, or by divine revelation, or by definition. Noncognitivists, being of many persuasions, have various views about justification. Some extreme emotivists and existentialists assert or imply that basic value judgments are arbitrary, irrational, and incapable of any justification (Ayer and Jean-Paul Sartre). Others believe that there are intersubjectively valid conventions, like "What is pleasant is good," which warrant our arguing from certain considerations to conclusions about what is good (Toulmin). Still others contend, in different ways, that attitudes, recommendations, commitments, conventions, and, hence, value judgments may be rational or justified, even if they cannot be proved inductively or deductively (Hare, Taylor, J. N. Findlay, and, up to a point, Stevenson).

See also Aesthetic Experience; Aristotle; Augustine, St.; Ayer, Alfred Jules; Beauty; Bentham, Jeremy; Blanshard, Brand; Bosanquet, Bernard; Bradley, Francis Herbert; Brentano, Franz; Dewey, John; Ehrenfels, Christian Freiherr von; Epicurus; Freedom; Good, The; Hare, Richard M.; Hartmann, Nicolai; Hume, David; Husserl, Edmund; Justice; Lewis, Clarence Irving; Lotze, Rudolf Hermann; Love; Meinong, Alexius; Mill, John Stuart; Moore, George Edward; Nietzsche, Friedrich; Perry, Ralph Barton; Plato; Pleasure; Ritschl, Albrecht Benjamin; Ross, William David; Russell, Bertrand Arthur William; Santayana, George; Sartre, Jean-Paul; Scheler, Max; Schlick, Moritz; Sellars, Roy Wood; Sidgwick, Henry; Spinoza, Benedict (Baruch) de; Stevenson,

Charles L.; Thomas Aquinas, St.; Truth; Virtue and Vice; Wright, Georg Henrik von.

Bibliography

INTRODUCTIONS

Frankena, William K. *Ethics.* Englewood Cliffs, NJ: Prentice-Hall, 1963. Elementary and systematic.

Frondizi, Risieri. *What Is Value?* Translated by Solomon Lipp. La Salle, IL, 1963. A useful elementary historical and critical work.

Nowell-Smith, Patrick H. *Ethics.* London: Penguin, 1954. A British systematic and analytical approach.

SYSTEMATIC DISCUSSIONS AND HISTORY

Blanshard, Brand. *Reason and Goodness.* New York: Macmillan, 1961. Scholarly and clear.

Laird, John. *The Idea of Value.* Cambridge, U.K.: Cambridge University Press, 1929. Scholarly and complex.

USAGES OF TERMS

Garnett, A. Campbell. *The Moral Nature of Man.* New York: Ronald Press, 1952. Ch. 4 is a useful, clear critique of prevailing semantic confusion.

Lepley, Ray, ed. *The Language of Value.* New York: Columbia University Press, 1957.

Lepley, Ray, ed. *Value: A Cooperative Inquiry.* New York: Columbia University Press, 1949. Both this and the following work illustrate usages, and both contain some good essays by Americans.

CONTINENTAL EUROPEAN WORKS

Ehrenfels, Christian von. *System der Werttheorie,* 2 vols. Leipzig: Reisland, 1897–1898. An early theory of value.

Hartmann, Nicolai. *Ethics,* 3 vols. Translated by Stanton Coit. London, 1932. An elaborate example of German intuitionism.

Meinong, Alexius. *Psychologisch-ethische Untersuchungen zur Werttheorie.* Graz, 1894. Another early theory of value and valuation.

Von Wright, Georg H. *The Varieties of Goodness.* London: Routledge and Kegan Paul, 1963. One of the latest theories of value.

BRITISH WORKS

Findlay, John N. *Values and Intentions.* New York: Macmillan, 1961. A value theory influenced by both British and Continental writers.

Hare, Richard M. *The Language of Morals.* Oxford: Clarendon Press, 1952. An imperativist analysis of value judgments.

Moore, G. E. *Principia Ethica.* Cambridge, U.K.: Cambridge University Press, 1903. A famous early analytical approach.

Urmson, J. O. "On Grading." *Mind* 59 (1950). A later analytical approach.

AMERICAN WORKS

Dewey, John. *Theory of Valuation.* Chicago: University of Chicago Press, 1939. A brief statement of the instrumentalist theory of values.

Lewis, C. I. *An Analysis of Knowledge and Valuation.* La Salle, IL: Open Court, 1946. A variation of a pragmatic theory.

Perry, Ralph B. *General Theory of Value.* New York: Longmans, Green, 1926. A classic work.

Perry, Ralph B. *Realms of Value.* Cambridge, MA: Harvard University Press, 1954. Restatement of Perry's value theory; applications to various fields.

Stevenson, Charles L. *Ethics and Language.* New Haven, CT: Yale University Press, 1944. An emotive theory.

Taylor, Paul W. *Normative Discourse.* Englewood Cliffs, NJ: Prentice-Hall, 1961. "The first fullscale attempt to use the 'informal logic' approach in general theory of value" (preface, p. xi).

William K. Frankena (1967)

VALUE AND VALUATION [ADDENDUM]

The ambiguities in the use of *value* and related words that William Frankena acutely summarized persist. But there has been some further work on value, especially in the narrow ethical sense of what is desirable or worth pursuing, that deserves comment. Some of the most interesting recent research has been by psychologists.

In relation to aesthetic value, psychologists have investigated how the kinds of experiences that we think point toward aesthetic value in their objects are produced. The psychology of identification with characters in fiction has been a fertile subject. So has the role of the unexpected (or not entirely expected) in appreciation of music. In all of this, there has been a tendency to connect aesthetic value with the quality of experiences that works of art (or beauties in nature) provide to those who are prepared to respond to them.

In relation to value in the narrow ethical sense of what is desirable or worth pursuing, much psychological research has investigated what people find satisfying in the present, or what they can be expected to find satisfying later (when they experience it), or what they prefer for the future. The evidence can seem telling, and yet many philosophers would want to distinguish sharply between what people like or prefer on one hand, and what has or would have value in their lives on the other.

There is the further complication that what has value in one life might have less or more within the context of a different kind of life. A kind of experience could be wonderful in one life and routine in another. C. I. Lewis (1883–1964) spoke of contributory value within a context. This suggests a contrast between instrumental value—the value that something causally has as a means to something else—and two kinds of noninstrumental value. One, which might be termed intrinsic, is a fixed

value (regardless of context) that something has because of what it is. The other is a noninstrumental value that depends on context.

How relevant is psychological research to judgments of what has value in a life? This is a vexed issue, one that connects with the long-standing philosophical problem of the relation between facts and ethical values (in various senses of value). This is sometimes spoken of as the problem of the *is* and the *ought*.

PSYCHOLOGICAL RESEARCH ON SUBJECTIVE WELL-BEING. The term *well-being* is sometimes used as a translation of Aristotle's *eudaemonia*, his term for the consortium of values in a desirable kind of life. Your subjective well-being is what you would estimate as the degree of desirability in the life you have. Many people tend to assume that their subjective well-being would go up sharply if they got a great deal more money, or if they were much luckier in getting what they wanted (and as a result had much more pleasure). Recent psychological research has tended to undermine these assumptions.

Australian work on the *hedonic treadmill* has shown that, while the subjective well-being of lottery winners can be expected to go up for a short period, it then tends to return to roughly the pre-lottery levels. This is because of adaptation: After a while it simply takes more to satisfy the newly wealthy person. Conversely, people who have been rendered paraplegic in their youth, tend after a while to return to prior levels of satisfaction in life. It takes less to give them pleasure.

There are exceptions to this. People who become paraplegic at an advanced age are much less likely to bounce back. Also there is evidence to suggest that newly raised levels of satisfaction that are linked to sense of self are more likely to persist. Mihaly Csikszentmihalyi has outlined a broad class of exceptions, having to do with experiences of losing oneself in a sequence of skilled activity. His subjects report these as continuing to be peak experiences. This lends psychological support to arguments like that of Plato's *Philebus*, that not all pleasures are alike and that some should be assigned much higher value than others.

Happiness is sometimes regarded as an index of the desirability of a life. Michael Argyle's data show that factors important to happiness cannot easily be reducible to pleasure. The element of one's attitude toward oneself, and toward the life they lead, is prominent in this. Someone who is very lucky in getting pleasures might all the same dislike themselves and not be happy.

Much of this psychological evidence can seem telling, especially in undermining simple views of what might make a life desirable. But it still can seem an open question whether a life that is happy and involves high subjective well-being is really a desirable one. What of someone who has an accident, as a result undergoes a right-side frontal lobotomy (becoming an idiot), and then is happy as can be? We tend to pity (rather than envy) such a person. Many of us also would not envy the sadist who has a very lucky run of victims. Conversely, could the life of someone who does not have an especially great amount of pleasure and is not unusually happy in any normal sense (e.g., Ludwig Wittgenstein) be unusually desirable? Norman Malcolm gives as words of the dying Wittgenstein, "Tell them I've had a wonderful life."

Even if satisfaction is not an index of the desirability of a life, it could be a factor. Many philosophers from Confucius and Aristotle on have taken it as obvious that a desirable kind of life must have at least a moderate degree of inner satisfaction. Even if psychological data do not entail judgments of value, one could hold that they can count in favor of certain judgments. Many philosophers, following Stephen Toulmin (b. 1922), have insisted that there can be reasons in support of ethical conclusions. If so, it is plausible to hold that psychological data often do provide reasons.

CAN A JUDGMENT OF VALUE BE CORRECT? One powerful reply to any line of thought that holds that psychological data provide evidence of what is desirable in life is this: Ethical judgments, including those of value in the narrow sense, it will be said, merely express the attitude of the person who accepts them. There is no truth here about something that is objectively the case.

One way of considering this issue is by examining degrees and kinds of objectivity. This is central to David Hume's essay *The Sceptic*. Richard W. Miller (b. 1945), like Hume, has examined differences between ethical and aesthetic objectivity.

There also has been a frontal assault on the notion that ethical judgments (including judgments of value) can have opinion-independent correctness. This was developed by A. J. Ayer and Charles Stevenson, and subsequently has been refined by Bernard Williams and Gilbert Harman. One argument for it is this. We can know that something is the case only if that it is the case plays a causal role in our coming to believe that it is. Scientific knowledge meets this requirement. But our ethical judgments can be causally explained without bringing in any alleged fact that they are correct. They can be

explained in terms of a collection of personal and social factors, including temperament, upbringing, acculturation, and so on. Hence there are no ethical facts, and certainly no opinion-independent facts about what really is a desirable kind of life.

A variety of issues are relevant. There is the nature (and perhaps the legitimacy) of the fact-value contrast. A naturalist in ethics need not take it as expressing any deep truth. There also are philosophers (e.g., P. F. Strawson and Bede Rundle) who have insisted on the interpretative elements in anything that we would term a fact, so that a fact is not anything in the world. It might generally be the case that what is judged to be a value also has an interpretative element, and, if so, an ethical judgment that encapsulated a very widely shared interpretation might look like a fact. G. E. M. Anscombe maintained that it was a *brute fact* that she owed her grocer money for potatoes that he had delivered at her request.

There also is the matter of the causal analysis of how people come to have the ethical beliefs they have. Plainly, factors such as upbringing and acculturation normally have a very large role, and it may be that often they are the whole story. But there are occasions when someone who has been brought up with a certain ethical view of *X* actually experiences *X*, and feels forced to change her or his mind. *X* might be a social practice that one had been taught was perfectly acceptable, but, looked at closely, seemed disgusting. Or *X* might be a highly recommended way of life, but after you have entered onto it seems somehow lacking. How thorough a knowledge do we have of the causation (all the causal factors) in such cases?

Sometimes people do have a sense that the cause of their rejecting, say, slavery was that it simply turned out to be disgusting—or that the recommended way of life just did not seem all that good, at least for them. They think, in short, that their changed opinion was caused by a sense of what ethically was the case. Can such a judgment about a causal relation have any validity? It can be plausible to hold that sometimes people do have an immediate awareness of a causal relation that does not require derivation from a covering generalization. Anscombe gives the example in *Intention* of knowing that one's fright was caused by the horrid face at the window.

DETERMINING WHAT IS DESIRABLE IN LIFE. If it is the case that some judgments of value in the narrow sense are better than others, then it is natural to ask how these can be arrived at or grounded. Many philosophers, as different from one another as Aristotle, Friedrich Nietzsche, and G. E. Moore have offered answers. Aristotle provides general considerations determining what can count as eudaemonia. The nature of humans as rational is given weight, as is the desirability of a life not far from what one imagines as the life of the gods.

Moore's nomination of intuition as the source of judgments of goodness is a way of saying that there is no strictly rational procedure. Nevertheless the final chapter of *Principia Ethica* contains a list of what seemed to him to be factors that would have a high degree of value. These include consciousness of Beauty and personal affection for someone worthy of it. James Griffin (b. 1933) also has provided a provisional list of major *prudential* values, approximating major noninstrumental values that can be attained. These include accomplishment, autonomy and liberty, understanding, enjoyment, and deep personal relations.

Both Moore's and Griffin's lists emerge as results of general reflection, doubtless with personal experience as part of its base. It is possible though to be highly skeptical of generalization about value. One can hold that the nuances of individual cases that fit under a general heading can make a major difference to the values. This is a point made in the "On the Three Evils" section of Nietzsche's *Thus Spake Zarathustra*, and echoed in Albert Camus's (1913–1960) *The Myth of Sisyphus*.

Along these lines, you could think it possible that something might contribute more (or less) of noninstrumental value in your life than something of the same general description would to someone else's life. Besides this, you can doubt that some familiar claims about what contributes noninstrumental value to lives have much validity for the general run of cases. Is there any way to arrive at a well-based answer to such questions?

One reply is this: A person can be in a good position to make a judgment about the noninstrumental value of *X* in a particular life, or the noninstrumental value *X* tends to have in lives in general, if you have a very good idea of what it is like to have *X* in a life. This could be the result of close observation of someone in whose life *X* is a part. Also, biographies and literary works, if they can be trusted, might sometimes provide such an idea.

The most common route though is to have experienced *X* in one's own life. This can be compared to being an eyewitness to an event. Some eyewitnesses are more reliable than others, and no eyewitness is guaranteed never to make mistakes. But to be an eyewitness is to be in a better position to know what happened than would otherwise be the case. Similarly, to have experienced *X* in your life is generally to be in a better position to judge the

noninstrumental value of *X*—at least in the context of the one life you know best. Mistakes are possible; but so also (if this line of thought is sound) is knowledge, and the knowledge may well be particular rather than general.

See also Good, The; Intrinsic Value; Value and Valuation.

Bibliography

DETERMINING WHAT IS DESIRABLE IN LIFE

Griffin, James. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford, U.K.: Clarendon Press, 1986. A nuanced and subtle study of value and issues related to value which repays repeated readings.

Kupperman, Joel. "The Epistemology of Noninstrumental Value." In *Philosophy and Phenomenological Research*, 66 (2005). A defense of claims that there can be knowledge of non-instrumental value.

FACT AND VALUE

Anscombe, G. E. M. "On Brute Facts." In *Analysis*, 18 (1958): 69–72. A strong attempt to undermine the widely assumed opposition between facts and values.

Rundle, Bede. *Facts*. London, 1993. Am examination of facts that is subversive of some widely accepted philosophical ideas.

OBJECTIVITY AND CORRECTNESS IN JUDGMENTS OF VALUE

Harman, Gilbert. *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press, 1977. This contains an argument for relativism, presented with some panache.

Miller, Richard W. "Three Versions of Objectivity: Aesthetic, Moral, and Scientific." In *Aesthetics and Ethics*, edited by Jerrold Levinson. Cambridge, U.K.: Cambridge University Press, 1998: 26–58. An exceptionally careful and nuanced discussion of varieties of objectivity.

Williams, Bernard. "The Truth in Relativism." In *Moral Luck: Philosophical Papers, 1973–1980*. Cambridge, U.K.; New York: Cambridge University Press, 1981. A subtle and provocative attack on some ideas of objective truth in ethics.

PSYCHOLOGICAL INVESTIGATIONS

Argyle, Michael. *The Psychology of Happiness* London; New York: Methuen, 1987.A clear account of factors relevant to happiness.

Brickman, P., D. Coates, and R. Janoff-Bulman, "Lottery Winners and Accident Victims: Is Happiness Relative?" In *Journal of Personality and Social Psychology*, 37 (1978): 287–302. The classic study of the hedonic treadmill.

Csikszentmihalyi, Mihaly. *Flow: the Psychology of Optimal Experience*. New York: Harper & Row, 1990. A study of experiences that are prized in people's lives.

Kahneman, Daniel. "Objective Happiness." In *Well-Being: the Foundations of Hedonic Psychology*, edited by D. Kahneman, E. Diener, and N. Schwartz. New York: Russell Sage Foundation, 1999. A strong general account of recent research.

Oatley, Keith, and Mitra Gholamain. "Emotions and Identification: Connectionsbetween Readers and Fiction." In *Emotion and the Arts*, edited by Mette Hjort and Sue Laver. New York, 1997. A fine example of psychological investigaton that sheds light on the production of aesthetic value.

Joel J. Kupperman (2005)

VALUE JUDGMENTS

See *Value and Valuation*

VALUE OF KNOWLEDGE AND TRUTH, THE

See *Knowledge and Truth, The Value of*

VAN FRAASSEN, BAS

(1941–)

Bas van Fraassen was born in Goes, in the Netherlands, on April 5. He lived in Holland until he was fifteen years old, when he moved with his family to Canada. After finishing his undergraduate studies in philosophy (with honors) at the University of Alberta in 1963, he went to the University of Pittsburgh for his Ph.D., which he completed in 1966 with a dissertation on the causal theory of time that was supervised by Adolf Grünbaum. He taught at Yale University, the University of Toronto, and the University of Southern California before moving to Princeton University, where he has been a Professor of Philosophy since 1982.

Van Fraassen has made seminal contributions to several areas of philosophy, and his work can be roughly divided into three major "periods": (i) the philosophical logic phase (1966–1979); (ii) the constructive empiricist period (1980–1993); and (iii) the empirical stance phase (1994 to the present). But throughout these periods, there has been a unified vision underlying his approach, with two crucial features: (a) the search for an empiricist (antirealist and, in a sense, antimetaphysical) approach to science and philosophy more generally; and (b) an attempt to preserve through this empiricism "classical" features of the domain under consideration—by taking scientific theories literally, retaining classical logic whenever possible, and resisting the need for introducing causally irrelevant items (such as possible worlds).

In the philosophical logic phase, this vision is articulated through the development of several proposals

guided by techniques from philosophical logic. For instance, van Fraassen's method of supervaluations provides a way of retaining classical logic (or, at least, classical logic's theorems), even in the presence of truth-value gaps. This method can then be used to accommodate logical paradoxes, such as the Liar ("This sentence is not true"). Van Fraassen's early work on space-time theories also illustrates the empiricist component of the vision, with the development of interpretations of space-time theories that do not presuppose the existence of absolute space (1970). Moreover, in his development of a semantics for free logic, van Fraassen assumed only existing individuals in the domain, thereby avoiding a commitment to nonexistent objects that early work in the area had presupposed. Finally, van Fraassen's early theory of meaning relations among predicates and modality does not involve any commitment to real modalities in nature.

Several of these problems can be approached from a unified perspective with the development of constructive empiricism (van Fraassen 1980). This is a view about the aim of science: the search for empirically adequate theories. The constructive empiricist articulates something novel: an empiricist alternative to scientific realism that avoids the early pitfalls of logical positivism. As opposed to logical positivism, the constructive empiricist takes scientific theories literally; there's no attempt to reformulate such theories in some formal language. And as opposed to scientific realism, the constructive empiricist puts forward an interpretation of science in which scientific theories need not be true to be good, as long as they are empirically adequate (and informative). To flesh out the proposal, van Fraassen argues that it is possible to make sense of scientific methodology from this viewpoint, and highlights, in particular, the crucial role played by models in scientific theorizing. He develops a new version of the semantic approach to scientific theories, insisting that to present a theory is to specify a class of models rather than to provide a list of axioms in a formalized language. As opposed to earlier positivist proposals, van Fraassen's work articulates a theory of the pragmatics of explanation that does not require scientific theories to be true for them to be explanatory. He also advances a new interpretation of probability that is compatible with the rejection of real modalities in nature.

Constructive empiricism's lack of commitment to metaphysically dubious notions (at least from an empiricist perspective)—such as laws of nature, possible worlds, and real modalities in nature—is developed further in van Fraassen's book *Laws and Symmetry* (1989). The book argues that attempts to characterize the notion of law of nature are doomed to failure because either they are unable to justify the inference from *It is a law that P* to *P*, or they fail to identify the features that make *P* a law in the first place. As an alternative, van Fraassen suggests that many roles that traditional philosophical proposals have assigned to laws of nature can be accommodated without commitment to the latter—provided we examine the role played by symmetry (roughly, transformations that leave certain structures invariant). A detailed case for this proposal in the context of quantum mechanics and a thorough development of an empiricist view of quantum theory is then articulated in *Quantum Mechanics: An Empiricist View* (1991).

After the development of the details of constructive empiricism, a more general question arises: How is it possible to be an empiricist instead of just developing an empiricist approach to science? To elaborate a broader perspective on empiricism that includes constructive empiricism as a particular case is a major goal of van Fraasen's empirical stance (2002). Instead of articulating empiricism as a doctrine (a set of beliefs), van Fraassen insists that empiricism should be conceptualized as a stance: an attitude, an epistemic policy. This move has several advantages. First, it avoids the incoherence of certain earlier empiricist proposals that failed to meet their own empiricist standards and ended up being meaningless or lacking any content. Second, the move also provides a novel way of understanding our practice, in particular the role of experience in our epistemic life, and how to make sense of scientific revolutions as a decision problem. The crucial features of van Fraassen's earlier works are also found here, notably in the development of an empiricist perspective that preserves the "classical" features of the phenomena under consideration.

See also Empiricism; Laws of Nature; Liar Paradox, The; Logical Paradoxes; Philosophy of Science, Problems of; Pragmatics; Presupposition; Realism.

Bibliography

Churchland, P. M., and C. A. Hooker, eds. *Images of Science: Essays on Realism and Empiricism, with a Reply by Bas C. van Fraassen*. Chicago: The University of Chicago Press, 1985.

van Fraassen, B. C. *The Empirical Stance*. New Haven, CT: Yale University Press, 1991.

van Fraassen, B. C. *An Introduction to the Philosophy of Time and Space*. New York: Random House, 1970.

van Fraassen, B. C. *Laws and Symmetry*. Oxford: Clarendon Pres, 1989.

van Fraassen, B. C. *Quantum Mechanics: An Empiricist View*. Oxford: Clarendon Press, 1991.

van Fraassen, B. C. *The Scientific Image*. Oxford: Clarendon Press, 1980.

Otávio Bueno (2005)

VANINI, GIULIO CESARE

(1584 or 1585–1619)

Giulio Cesare Vanini was born in Taurisano, in the province of Lecce, Italy, in 1584 or early in 1585. After completing a course of study in law in Naples, he proceeded to Padua to study theology. He entered the order of the Carmelites, and he visited various Italian cities—Venice, Genoa, and perhaps Bologna—and traveled in Germany, England, and France. In 1612, in England, he abjured, but, having aroused suspicion because of his ideas, he moved on again. In 1615, in Lyon, he published his *Amphitheatrum Aeternae Providentiae* (published by the widow of Antoine De Harsy), and in 1616, in Paris, the dialogues, in four books, *De Admirandis Naturae Reginae Deaeque Mortialium Arcanis* (published by Adrian Périer). Both works were given the regular permission of the ecclesiastical authorities but nevertheless aroused suspicions. Vanini then went to Toulouse, where he taught and practiced medicine. In August 1618 he was arrested by the Inquisition. He was condemned, and then in February 1619 burned to death after horrible torture.

Vanini's work, which shows repeatedly a kinship with that of Averroes, reflects above all the influence of the writers of the fifteenth and sixteenth centuries, among whom he had a particular predilection for Pietro Pomponazzi, whom he called his master, the prince of the philosophers of his century, and a second Averroes ("in his body Pythagoras would have placed the spirit of Averroes"). Next to Pomponazzi he placed Girolamo Cardano, Julius Caesar Scaliger, and numerous others, whom he drew from freely. His liberal use of other sources, long passages of which he inserted, even verbatim, into his own works, has caused several recent historians to speak of plagiarism and of writings that are "devoid of originality and scientific integrity." In reality, his attitude toward using the writings of others was common in his time; the present-day preoccupation with the citation of sources did not exist (certain Latin writings of Giordano Bruno are a case in point). Furthermore, the writings from which Vanini borrowed generally underwent a marked transformation in his pages.

Intensely critical of all revealed religions (his "atheism" stemmed from this), Vanini believed strongly in the divinity of nature and in the immanence of God in nature, which is eternal and eternally regulated by strict laws ("Natura Dei facultas, imo Deus ipse"). He held that the world is without origin, at least so far as could be established by natural religion. The human spirit is material, the soul mortal. Using arguments and themes taken from Cardano, Vanini stated that there is a natural explanation for all supposedly exceptional and miraculous phenomena in universal determinism; and thus, going back to Pomponazzi, he interpreted rationally all the aspects and forms of religious life.

Despite his frequent declaration that, as a Christian, he would continue to accept on faith even that which reason had disproved, the radical bent of Vanini's criticism escaped no one, and, as the seventeenth century progressed, he became almost a symbol of "atheistic and libertine" thought.

See also Atheism; Averroes; Bruno, Giordano; Laws of Nature; Pomponazzi, Pietro.

Bibliography

WORKS BY VANINI

Luigi Corvaglia, ed., *Le opere di Giulio Cesare Vanini e le loro fonti,* 2 vols. (Milan: Società Anonima Editrice Dante Alighieri, 1933–1934), is a reprint of the original editions with the texts of the "sources" printed alongside to show the "plagiarism." See also Guido Porzio, *Le opere di Giulio Cesare Vanini tradotte per la prima volta in italiano con prefazione del traduttore,* 2 vols. (Lecce, Italy, no date; published 1913), which includes biography, documents, complete bibliography.

WORKS ON VANINI

See F. Fiorentino, "Giulio Cesare Vanini e i suoi biografi," in *Studi e ritratti della rinascenza* (Bari, Italy: Laterza, 1911); E. Namer, "Nuovi documenti su Vanini," in *Giornale critico della filosofia italiana* 13 (1932): 161–198; E. Namer, *Documents sur la vie de Jules-César de Taurisano* (Bari, Italy, no date; published 1965); John Owen, *The Skeptics of the Italian Renaissance* (London: Sonnenschein, 1893), pp. 345–419; G. Spini, "Vaniniana," in *Rinascimento* 1 (1950): 71–90; G. Spini, *Ricerca dei libertini* (Rome: Editrice Universale de Roma, 1950), pp. 117–135.

Eugenio Garin (1967)
Translated by Robert M. Connolly

VARISCO, BERNARDINO

(1850–1933)

Bernardino Varisco, the Italian metaphysician, was born at Chiari (Brescia). It was only in the later part of his long life that he developed his philosophy, for he began as a teacher of science and his early outlook was characterized

by empiricism and positivism. These views found expression in *Scienza e opinioni* (1901). Thereafter he became interested in the problem of reconciling the scientific and religious ways of understanding the world and moved into metaphysics. In 1906 he was appointed professor of theoretical philosophy at the University of Rome, where he remained until his retirement in 1925. His metaphysic was a philosophy of spirit in the manner of Gottfried Wilhelm Leibniz and Rudolf Hermann Lotze and won him a considerable reputation in Italy and elsewhere.

The empiricism of Varisco's earlier phase was still apparent in the approach that he employed in constructing his distinctive philosophy. His starting point is the given fact of a plurality of conscious subjects. Each of these has its own private perspective upon the world, and each is also a spontaneous center of activity. In the personal subject, a high level of rationality and self-consciousness has been reached, but this is surrounded by an extensive penumbra of subconsciousness. Varisco thinks of conscious life as shading off imperceptibly into lower levels. Below the level of man's personal existence there is animal life, and it is argued that this in turn shades off into so-called inanimate existence. Thus, Varisco arrives at a kind of monadology, or panpsychism. Reality is made up of an infinite number of subjects, although at the level of inanimate nature these subjects are very primitive and have nothing like the self-consciousness of the personal human subject.

Varisco's metaphysic has a dynamic aspect, for these subjects are in constant action and interaction. The variations set up are of two kinds. Some arise from spontaneous activity in the subjects themselves, and in this way Varisco provides for freedom and for what he calls an "alogical" factor in reality. The other kind of variations arises from the mutual interaction of the subjects, and this happens in regular ways, so that the universe has also an ordered, logical character.

The most obscure and presumably the weakest part of Varisco's philosophy is his attempt to move from the plurality of subjects to a unitary reality. His appeal is to the notion of "being," which, implicitly or explicitly, is present in every act of thought whereby a subject grasps an object. Being is identified with the universal subject, with thinking itself in all particular subjects and in the world. In *I massimi problemi,* Varisco says explicitly that the universal subject is a logical conception that falls short of the notion of a personal God, although he believed that teleology and the conservation of value point toward theism. However, in his posthumous *Dall'uomo a Dio* (1939) he completes his pilgrimage from positivism to theism, arguing for a God who limits himself by his creation so that men can cooperate with him in creative activity. Such a view, he believed, supports a religious attitude to life and is especially compatible with Christianity.

See also Empiricism; Leibniz, Gottfried Wilhelm; Lotze, Rudolf Hermann; Metaphysics; Panpsychism; Positivism.

Bibliography

WORKS BY VARISCO

Scienza e opinioni. Rome, 1901.

La conoscenza. Pavia, 1905.

I massimi problemi. Milan, 1910. Translated by R. C. Lodge as *The Great Problems.* London: Allen, 1914.

Conosci te stesso. Milan, 1912. Translated by Guglielmo Salvadori as *Know Thyself.* London: Allen and Unwin, 1915.

Sommario di filosofia. Rome, 1928.

Dall'uomo a Dio. Padua: CEDAM, 1939.

WORKS ON VARISCO

Chiapetta, L. *La teodicea di Bernardino Varisco.* Naples, 1938.

De Negri, E. *La metafisica di Bernardino Varisco.* Florence, 1929.

Drago, P. C. *La filosofia di Bernardino Varisco.* Florence: Monnier, 1944.

Librizzi, C. *Il pensiero di Bernardino Varisco.* Padua: CEDAM, 1942; rev. ed., 1953.

John Macquarrie (1967)

VARONA Y PERA, ENRIQUE JOSÉ
(1849–1933)

Enrique José Varona y Pera was a Cuban philosopher, statesman, and man of letters. Beginning in the mid-1870s, Varona dominated Cuban intellectual life for fifty years. He was a professor of philosophy at the University of Havana, was founding editor of *Revista cubana,* and took an active part in education and politics. A former member of the Spanish Cortes, he became a revolutionary colleague of José Martí, was appointed secretary of public instruction and fine arts after the 1898 revolution, and served as vice-president of Cuba from 1913 to 1917.

Varona, one of the leading Latin American positivists, adapted French positivism and British empiricism to the contemporary sociopolitical and cultural situation of Cuba. Logic, psychology, and ethics were his primary philosophic concerns.

J. S. Mill's analysis of induction served as the basis of Varona's work in logic. As a scientific study of the ways in which man thinks and learns, logic assists in providing methodologies for the particular sciences as well as for the educational process. There are three stages in any mental act: The first and third are directed toward the object of experience, the second consists exclusively of mental activity. Unrelated data are obtained from nature; they are then related significantly in terms of ideal constructs, and the resultant schema is again compared with experience through controlled experimentation.

In psychology the root problem is that of human freedom. Varona subordinated the study of psychology to that of physiology and accepted a strictly deterministic position. However, his concern for the political and cultural independence of Cuba demanded an interpretation of man that provided room for freedom. Although man is not free, the development of intelligence provides him with the ability to avoid being an automaton, to understand the nature of causal determination, and thereby to "train and direct it, which is tantamount to overcoming it."

The proper approach to the study of ethics is genetic. Morality is based on the social nature of man, which, in turn, has its roots in the evolutionary biological process. "Man is not sociable because he is moral.… Man becomes moral by virtue of being sociable" (*Conferencias filosóficas, tercera serie: Moral* [Havana, 1888], p. 10). Just as the biological organism is dependent upon its natural environment, so the human organism is dependent upon its social environment. Such social dependence constitutes social solidarity. Awareness of this dependence and conscious accommodation of the individual to the social milieu constitutes moral behavior.

Throughout Varona's work and especially in a final book of aphorisms, *Con el eslabón* (Manzanillo, 1927), a subtle, penetrating irony concerning the foibles of human thought and existence was evident.

See also Empiricism; Ethics; Latin American Philosophy; Logic, History of; Mill, John Stuart; Positivism; Psychology.

Bibliography

Varona's collected works are *Obras,* 4 vols. (Havana, 1936–1938). Works on Varona include John H. Hershey, "Enrique José Varona, Cuban Positivist," in *Humanist* 3 (January 1944): 164ff.; Medardo Vitier, *La filosofia en Cuba* (Mexico City: Fondo de Cultura Económica, 1948), Ch. 11; and Humberto Piñera Llera, *Panorama de la filosofia cubana* (Washington, DC: Unión Panamericana, 1960), Ch. 5.

Fred Gillette Sturm (1967)

VASCONCELOS, JOSÉ
(1882–1959)

José Vasconcelos, the Mexican politician and philosopher, was born in Oaxaca. Vasconcelos was active in the Mexican revolution, directed the reform of Mexican education as secretary of education in the early 1920s, ran unsuccessfully for the presidency in 1929, and subsequently was exiled for a time. He was rector of the National University of Mexico, visiting professor at the University of Chicago, and director of the Biblioteca Nacional de México. The sources of his philosophy were Pythagoras, Plotinus, Arthur Schopenhauer, Friedrich Nietzsche, A. N. Whitehead, and especially Henri Bergson. Of Latin American philosophers, Vasconcelos is the most original, venturesome, and impassioned.

He called his philosophy aesthetic monism, scientific realism, and organic logic. The system he developed stressed intuition in addition to scientific experience; the particular, concrete, and heterogeneous; organic wholes; the fluid, living, and psychical; and the methods of art rather than mathematics. The true method of philosophy, Vasconcelos claimed, is to understand the particular phenomenon, not by reducing it to the universal but by relating it to other particulars in an organic whole in which unity is achieved without sacrifice of individuality.

The pervasive term in Vasconcelos's theory of reality is energy, which is unformed in its primordial condition but takes on determinate structures in the three phenomenal orders of the atomic, cellular, and spiritual. The transformation in recent physics of the elementary particle from a rigid body to an "individualized dynamic frequency," Vasconcelos held, emphasizes activity and novelty in the atom, which are reminders of spirit. In the cellular order, internal purposes are introduced. Spirit is eminently creative, but its action follows structures, or a priori methods, of logical inference for intellect, of values or norms for will, and of aesthetic unities for feeling. The early thought of Vasconcelos was pantheistic, finding the creative principle in the self-sufficient pervasive energy of the world. His later thought, after he had returned to the Roman Catholic Church, was theistic. It appears that in both periods "spirit," rudimentary or refined, was basic to his view of reality.

In Vasconcelos's aesthetics may be found implications for both reality and the life of spirit. The work of art, an emotionally intuited image, observes principles which, although more lucid in the work itself, have general application in reality. A musical scale is constructed by the musician out of the continuum of natural pitches; its

members are discrete tones separated by intervals or jumps. The activity of constructing this scale is analogous to that of intelligence in separating and ordering the objects of sensation; the discontinuity of the tones is similar to that of quantum phenomena in physics. Musical compositions observe three modes of aesthetic unity—melody, harmony, and rhythm—in which the heterogeneous or discontinuous is unified without loss of diversity. A true metaphysics, fortified by modern science, finds the same types of unity in reality, unlike mathematics, which unifies by reduction to homogeneous quantities.

Art, according to Vasconcelos, expresses the transformations of the spirit in the pursuit of value. He distinguished three kinds of art. Apollonian art is formal and intellectual. It can be saved from decay in giganticism or sensuality only by a shift to the Dionysian mode of passionate affirmation of the human will. Dionysian art does not decline; passion either destroys the spirit or saves it by a change to religious ardor. In mystical art, passion is directed from a temporal and human object to an eternal and divine object. Passion need not retreat from fate, as the Greeks thought; as Christianity discovered, it can be fully satisfied in the divine.

A similar conclusion occurs in the ethics of Vasconcelos. A terrestrial ethics, exemplified diversely in empiricism, hedonism, Confucianism, humanism, and socialism, does not take man beyond his animal and human condition. (Apart from this deficiency, a limited socialism stripped of Marxist theory has merit; Vasconcelos was critical of capitalism.) Metaphysical ethics attempts to go further in the name of reason; but the rational universal law of Immanuel Kant is a discipline appropriate for things and not for spirits. The highest ethics is revelatory; it combines transcendence, emotional illumination, and infinite love. Vasconcelos highly praised the wisdom of Buddhism and of Christianity, but he preferred Christianity because of its affirmation of life.

See also Aesthetics, History of; Bergson, Henri; Intuition; Kant, Immanuel; Latin American Philosophy; Nietzsche, Friedrich; Plotinus; Pythagoras and Pythagoreanism; Schopenhauer, Arthur; Scientific Realism; Whitehead, Alfred North.

Bibliography

Principal works by Vasconcelos include *Pitágoras: Una teoría del ritmo* (Havana: El Siglo XX, 1916); *El monismo estético* (Mexico City: Murguia, 1918); *Tratado de metafísica* (Mexico City: México joven, 1929); *Ética* (Madrid: Aguilar, 1932); *Estética* (Mexico City: Botas, 1936); *El realismo científico* (Mexico City: Centro de Estudios Filosóficos de la Facultad de Filosofía y Letras, 1943); and *Lógica orgánica* (Mexico City: Edición de el Colegio Nacional, 1945).

Also see Patrick Romanell, *Making of the Mexican Mind* (Lincoln: University of Nebraska Press, 1952), Ch. 4.

Arthur Berndtson (1967)

VASQUEZ, GABRIEL
(1549–1604)

Gabriel Vasquez, the neo-Scholastic theologian, was born at Villascuela del Haro, Spain, and died at Alcalá. Educated in the Jesuit houses of study in Spain, he taught moral philosophy at Ocaña from 1575 to 1577 and theology at Madrid and Alcalá. Eventually he succeeded Francisco Suárez in the chair of theology at Rome, where he taught from 1585 to 1592. His *Commentaria ac Disputationes in Primam Pattern S. Thomae* (8 vols., Alcalá, 1598–1615), a lengthy commentary on Part I of Thomas Aquinas's *Summa Theologiae*, contains much philosophical speculation. A posthumously published summary of this work, *Disputationes Metaphysicae* (Madrid, 1617), helped to popularize his philosophy.

Vasquez's most influential contribution lies in his distinction between the formal concept in the understanding (a mental entity, or "idea," constituting knowledge, *qualitas ipsa cognitionis*) and the objective concept that is the reality that is known (*res cognita*) through the formal concept (*Commentaria* I, 76, nn. 2–5). Since, in the view of Vasquez, the actual being (*esse*) of the thing that is known is identified with the act whereby it is known (*cognosci*), we may have here one of the sources of idealism in modern philosophy. There is little doubt that René Descartes's Jesuit teachers knew the thought of Vasquez, and hence the Cartesian teaching that ideas are direct objects of knowledge may owe a good deal to Vasquez (see the study by R. Dalbiez). Like Suárez, Vasquez introduced many changes into Thomistic metaphysics. He rejected the view that essence and existence are really distinct, opposed the theory that act is limited by the potency in which it is received, and argued that matter as marked by quantity (*materia signata quantitate*) cannot be the principle that individuates bodily things.

In psychology Vasquez also had teachings that are highly personal. He saw no reason for postulating two intellectual powers in man (agent and possible intellects, in Thomas) and implied that the one understanding can do the work of both. He regarded man as a composite of soul and body, but he treated these two "parts" almost as if they were two different substances joined together by a

peculiar sort of metaphysical semireality that he called a "mode." Here again, we may have a source of Descartes's mind-body problem and of the psychophysical parallelism of post-Cartesianism.

In his long discussion of St. Thomas's proofs for the existence of God, Vasquez again showed a critical attitude toward the thought of Thomas. In place of the traditional Five Ways of demonstration (which require the acceptance of a metaphysics of causality), Vasquez described a whole new series of arguments of his own. God's existence is demonstrated from the claim that morality requires it (an argument that reappears in Immanuel Kant) and from various types of "spontaneous assents" based on what one learns from parents, on a survey of the whole of reality (*ex rerum universitate*), and on our knowledge of the divine conservation and governance of the world (*Commentaria* I, 19, nn. 9–12). It is evident that Vasquez's work is one of the reasons that Thomism came to be misunderstood in modern philosophy.

See also Descartes, René; Kant, Immanuel; Scotism; Suárez, Francisco; Thomas Aquinas, St.; Thomism.

Bibliography

Dalbiez, R. "Les sources scolastiques de la théorie cartésienne de l'être objectif." *Revue d'histoire de la philosophie* 3 (1939): 464–472.

Gilson, Étienne. *Études sur le rôle de la pensée médiévale dans la formation du système cartésien*, 203–207. Paris: Vrin, 1930.

Solana, M. *Los grandes escolásticos españoles*, 109–128. Madrid: Ratés, 1928.

Wells, Norman J. "John Poinsot on Created Eternal Truths vs Vasquez, Suarez and Descartes." *American Catholic Philosophical Quarterly* 3 (1968): 425–446.

Vernon J. Bourke (1967)
Bibliography updated by Tamra Frei (2005)

VASUBANDHU

(fl. fourth or fifth century CE)

Vasubandhu was an Indian Buddhist philosopher who made significant contributions to the clarification and development of the Indian Buddhist schools of philosophy traditionally classified as the Vaibhāṣika (or Sarvāstivāda), the Sautrāntika, and the Yogācāra (or Cittamātra). Erich Frauwallner argued (1951), on the basis of a study of Vasubandhu's biographers, Paramārtha (499–569), Bus-ton (1290–1364) and Tāranātha (1575–1634), that there were two Vasubandhus, one who composed Yogācāra works and lived in the fourth century CE, and another who lived in the fifth century CE and composed treatises from the Vaibhāṣika and Sautrāntika points of view. But later studies (Jaini 1959, Anacker 1998) disputed Frauwallner's argument and advanced the hypothesis that there was only one author of these works and that he lived in the fourth century CE According to Buddhist tradition, Vasubandhu was at first an orthodox follower of the Vaibhāṣika school, and, after having allied himself with the Sautrāntika school, was convinced by his half-brother, Asaṇa, to accept the Mahāyāna scriptures (which were not accepted by the Vaibhāṣikas or Sautrāntikas) and to adopt the theses of the Yogācāra school.

VASUBANDHU'S CONTRIBUTIONS TO THE VAIBHĀṢIKA AND SAUTRĀNTIKA PHILOSOPHIES

Vasubandhu's contribution to the Vaibhāṣika philosophy is his masterly treatise the *Abhidharmakośa* (Treasury of knowledge). In this work he sets out in verse theses held in most of the Vaibhāṣika schools. One of the most fundamental of these theses is that what truly exists (that is, what exists apart from being conceived) is a substantially real permanent or impermanent phenomenon (*dharma*) or a collection of substantially real impermanent phenomena that is by convention conceived as a single entity of a certain kind. The treatise as a whole explains the world of conventional phenomena in terms of how its underlying substantially real phenomena are caused to combine and separate to perpetuate our rebirth and suffering and how, by eliminating their causes, our rebirth and suffering can be eliminated.

To this work Vasubandhu added a prose treatise, the *Ātmavādapratiṣedha* (Refutation of the theory of a self). In it he defends the theory of persons of the Vaibhāṣikas, who believe that we, as persons conceived from the first-person singular perspective, suffer and are reborn because we misapprehend ourselves as selves in the sense of being substantially real phenomena. We can become free from rebirth and suffering by realizing that we are not substantially real phenomena. Nonetheless, he believes, we ultimately exist insofar as we are the collections of substantially real impermanent aggregates (*skandhas*) of which our bodies and mental states are composed; only these aggregates are found, by direct perception and correct inference, to be the phenomena on the basis upon which we conceive ourselves as persons. Vasubandhu then presents objections to the theories of persons held in the unorthodox Vaibhāṣika school called the Vātsīputrīya and in the Hindu school called the Vaiśeṣika. According to the Vātsīputrīyas, we ultimately

exist without being collections of such phenomena, and according to the Vaiśeṣikas, we ultimately exist as permanent and partless substantially real phenomena. Vasubandhu claims that the Vātsīputrīyas' arguments for their theory are inconsistent with other theses they should, as Vaibhāṣikas, accept, and argues that their theory, like that of the Vaiśeṣikas, has the absurd consequence that we are completely different from, and so causally unrelated to, our aggregates.

In reply to the objections of the Vātsīputrīyas and Vaiśeṣikas—that the Vaibhāṣika theory of persons implies that we are not the same over time, do not possess mental states, and so on— Vasubandhu explains how, in spite of our reducibility to collections of impermanent aggregates, we are said to be the same over time, to possess mental states, and so on. Vasubandhu also briefly rejects the thesis of Nāgārjuna, the founder of the Mādhyamika school of Indian Buddhist philosophy, that nothing is substantially real, which, he believes, implies that we do not ultimately exist at all, since we could not in that case be reducible in existence to collections of substantially real phenomena.

Vasubandhu's most important contribution to the development of the Sautrāntika school was the *Abhidharmakośabhāṣya* (Commentary on the treasury of knowledge), a prose commentary on the verses in the *Abhidharmakośa*. In this work he adopts the Sautrāntika project of correcting the ontological excesses of the Vaibhāṣika school by showing that they are not supported by Buddhist scriptures. Although Vasubandhu accepts the Vaibhāṣika thesis that what exists is either a substantially real phenomenon or a collection of substantially real impermanent phenomena, he argues that the Vaibhāṣikas introduce more substantially real phenomena than are needed in order to explain how suffering and rebirth arise and are eliminated. For instance, he rejects the Vaibhāṣika explanation of how substantially real phenomena that have occurred in the past or will occur in the future can be apprehended if they do not ultimately exist at the time they are being apprehended. Their explanation is that substantially real phenomena ultimately exist in the past, present, and future insofar as they possess a real nature (*svabhāva*) by virtue of which they can be identified by themselves; they are said to be past phenomena when they have already exercised their characteristic causal power, to be present phenomena when they are exercising it, and to be future phenomena when they have not yet exercised it. Vasubandhu's basic objection to this explanation is that it unnecessarily introduces into their basic ontology past and future substantially real phenomena, because it is possible to apprehend substantially real phenomena that have ceased to exist and have not yet come to exist.

Among the many other theses of the Vaibhāṣikas he rejects are the theses: (i) that there can be a cause that is simultaneous with or can follow its effect (he claims that a cause must always precede its effect), (ii) that a future result of an action must occur in the same person who performed the action because there is present in the continuum of the person's aggregates of body and mind a separate substantially real phenomenon that causes the retention of the seed the action produces in the same causal continuum (he claims that the retention of the seed is due to the causal relationship between the phenomena in the continuum), and (iii) that an impermanent phenomenon can exist for more than an instant (he believes that an impermanent phenomenon by nature ceases to exist as soon as it arises).

It may have been during his Sautrāntika period (though some scholars think it was when he had already become a follower of the Yogācāra movement) that Vasubandhu wrote a number of treatises on logic in which he presents revisions and clarifications of forms of argument used by Indian philosophers in debate. In the *Vādavidhi* (The way of argument), part of which has survived, Vasubandhu anticipates some of the views of the Buddhist logician Dignāga, a circumstance that perhaps explains why he is sometimes said to be one of Dignāga's teachers.

VASUBANDHU'S CONTRIBUTIONS TO YOGĀCĀRA PHILOSOPHY

One of Vasubandhu's earliest contributions to the clarification and development of Yogācāra thought may be the *Pañcaskandhakaprakaraṇa* (A treatise on the five aggregates), which is an attempt to improve upon Asaṇa's account of the five aggregates in the *Abhidharmasammucaya* (Compendium of knowledge). In the *Karmasiddhaprakaraṇa* (A treatise on the establishment of Karma), Vasubandhu argues that the workings of the law of actions and their results are not correctly explained by the orthodox Vaibhāṣikas or by the Vātsīputrīyas and that the law's explanation requires reference to the Yogācāra theory that there is, apart from the six types of consciousnesses that are associated with the six types of organs of cognition, a storehouse consciousness (*ālayavijñāna*) that carries the seeds of all experiences and that this consciousness is not the substantially real self that we misapprehend it to be.

In the *Trimśikākārikāvṛtti* (Thirty verse treatise) and the *Trisvabhāvanirdeśa* (Teaching on the three natures), Vasubandhu explains how consciousness functions in terms of its three natures. He argues that persons and other phenomena are just ever-changing manifestations of consciousness, which is itself a beginningless sequence of momentary mental states that takes three different forms. Its most basic form is that of the storehouse consciousness, which is a beginningless sequence of mental states in which is stored the seeds that are produced by actions and give rise to their results. In dependence upon this sequence as an underlying support, it takes the forms of the afflicted mind (*kliṣṭamanas*), which is a sequence of minds that misapprehend the first sequence as a substantially real self, and of a sequence of six organ-dependent cognitions of objects. All three of these ever-changing forms of consciousness, Vasubandhu adds, are mental constructions and are to be eliminated on the path to Buddhahood.

The conceptual framework Vasubandhu uses to explain how mental constructions can cease to exist is that consciousness possesses three natures (*svabhāva*s). They are its nature of being dependent upon causes and conditions (*paratantra*), its nature of falsely appearing to be divided into a mind that grasps an object and an object that is grasped by it (*parikalpita*), and its thoroughly established nature (*pariniṣpannasvabhāva*) of not in fact being divisible into a mind that grasps an object and an object that is grasped by it. To become free from mental constructions and the rebirth and suffering they occasion, we need to realize in what way consciousness, in relation to its possession of these three natures, is without a nature (*niḥsvabhāvatā*). In relation to consciousness possessing the nature of appearing to be divided into a mind that grasps an object and an object that is grasped by it, consciousness is by its own nature without such a nature. In relation to consciousness possessing the nature of being dependent upon causes and conditions, consciousness is without a nature by virtue of which it could come to be by itself. In relation to consciousness possessing a thoroughly established nature, consciousness is without a nature by virtue of which it is divisible into a mind that grasps an object and an object that is grasped by it. To become free of rebirth and suffering and become a Buddha, Vasubandhu explains, we need to enter into a state of consciousness that is free from all mental constructions.

In Vasubandhu's *Viṃśatikākārikāvṛtti* (Twenty verse treatise) and his own commentary on it, he answers objections to the central theses of the Yogācāra philosophy. He says that the things we believe to exist apart from mind (that is, the things we believe to be external objects) are mere mental constructions (*vijñaptimātra*), because what does not exist apart from mind appears, because of the constructive activity of mind, to exist apart from mind, just as what does not exist apart from sight appears, because of an eye disorder, to exist apart from sight. In reply to the objection that if there are no external objects, perceptions cannot be distinguished from one another and the same objects cannot be perceived by different persons, he argues that perceptions in dreams differ from one another in spite of lacking external objects as causes and that many different persons perceive the same objects as a result of similar actions performed in the past. He also argues that the suffering that is experienced by beings in the hell realms is not produced by external objects, because otherwise the hell-guardians, who are said in scripture not to suffer in these realms, would suffer along with those reborn in those realms. He adds that there can be no atoms of which external objects are composed, since they could not possess different sides as parts and so could not occupy space and be combined to compose external objects, which are said to occupy space.

Vasubandhu also composed many commentaries on Yogācāra treatises and Mahāyāna scriptures. Important among those that have survived (either in Sanskrit or in Tibetan or Chinese) are the *Madhyāntavibhāgabhāṣya* (Commentary on the separation of the middle from the extremes), the *Mahāyānasūtrālaṃkārabhāṣya* (Commentary on the ornament of the Mahāyāna Sūtras), the *Mahāyānasaṃgrahabhāṣya* (Commentary on the compendium of Mahāyāna), and the *Dharmadharmatāvibhāgavṛtti* (Commentary on the distinction between phenomena and their true nature).

See also Buddhism; Buddhism—Schools; Causation in Indian Philosophy; Indian Philosophy; Logic, History of: Logic and Inference in Indian Philosophy; Nāgārjuna; Self in Indian Philosophy.

Bibliography

Anacker, Stefan. *Seven Works of Vasubandhu, The Buddhist Psychological Doctor*. Delhi: Motilal Banarsidass, 1998.

Duerlinger, James. *Indian Buddhist Theories of Persons: Vasubandhu's "Refutation of the Theory of a Self."* Curzon Critical Studies in Buddhism Series. London and New York: RoutledgeCurzon, 2003.

Frauwallner, Erich. *On the Date of the Buddhist Master of the Law, Vasubandhu*. Rome: Serie Orientale Roma, III, 1951.

Jaini, Padmanabh S. "On the Theory of the Two Vasubandhus." *Bulletin of the London School of Oriental and African Studies* 22 (1959): 48–53.

Kochumuttom, Thomas A. *A Buddhist Doctrine of Experience: A New Translation and Interpretation of the Works of Vasubandhu the Yogācārin*. Delhi: Motilal Banarsidass, 1982.

Lusthaus, Dan. *Buddhist Phenomenology: A Philosophical Investigation of Yogācāra Buddhism and the Ch'eng Wei-shih Lun*. Curzon Critical Studies in Buddhism Series. London: RoutledgeCurzon, 2002.

Potter, Karl H. *Encyclopedia of Indian Philosophies*. Vol. 3, *Buddhist Philosophy from 100 to 350 A.D.* Delhi: Motilal Banarsidass, 1999.

Pruden, Leo. M. *Abhidharmakośabhāṣyam by Louis de La Vallée Poussin*. 4 vols. Berkeley: Asian Humanities Press, 1990.

Stcherbatsky, Theodore. *The Central Conception of Buddhism and the Meaning of the Word "Dharma."* Delhi: Motilal Banarsidass, 1970.

Wood, Thomas E. *Mind Only: A Philosophical and Doctrinal Analysis of the Vijñānavāda*. Delhi: Motilal Banarsidass, 1994.

James Duerlinger (2005)

VAUVENARGUES, LUC DE CLAPIERS, MARQUIS DE

(1715–1747)

The French moralist and epigrammatist Luc de Clapiers, marquis de Vauvenargues, was born at Aix-en-Provence. He early revealed a lofty character that despised egotism and pettiness. Ambitious for glory, he became an army officer at the age of seventeen, despite a weak physique. He served throughout the Italian campaign of 1734. The later German campaign of 1741, especially the harsh retreat from Prague, ruined his health, forcing him to retire at the age of twenty-six. His hope of a career in diplomacy was dashed by lack of fortune and protection. While vainly waiting at Aix for replies to his petitions for appointment to a post, he contracted a severe case of smallpox that left him disfigured and sickly. His last years were spent in Paris, in unhappy poverty and solitude (despite Voltaire's admiration), but he endured the injustice of men and events with stoic resignation rather than with bitterness. During this period he wrote his *Introduction à la connaissance de l'esprit humain* (Paris, 1746; augmented edition, 1747), which included the supplement "Réflexions et maximes." He also wrote character sketches in the fashion of Jean de La Bruyère, although less brilliantly, and *Réflexions sur divers auteurs*, a work of generally sound and objective criticism. He is particularly known for his maxims.

Vauvenargues's life and writings are characterized by their contradictions rather than by their consistency. Weak in health, he had a proud, heroic soul; poverty-stricken, he refused to consider gainful work out of aristocratic prejudice and a dislike for restraint. A lover of peace, he praised war and the martial virtues; opposed to ethical absolutes, he considered greatness of soul and action to be absolute virtues. Extremely unhappy and frustrated in life, his writings are resolutely optimistic; almost without friends, his correspondence reveals a noble ideal of friendship. Inclined to sentiment, he was from youth enamored of Plutarch, Seneca, and the Stoic attitudes.

Vauvenargues was a vigorous but not a profound or systematic thinker. He is notable for his incisive insights and formulations, principally in regard to character and moral ideals. He was a deist and not a Christian; but, believing religion necessary to social order, he opposed the propaganda of the *philosophes*. His philosophy, however, was secular in spirit, concerned with the problem of human nature and of what men should be and how they should live. He defended the worth of human nature both against the pessimism of the Christian doctrine of original sin and the corrosive cynicism of Duc François de La Rochefoucauld. Like other thinkers of his time, he justified the passions. Following Benedict de Spinoza, he divided the passions into two kinds, according to their motivation: "They have their principle in the love of being [and desire for its] perfection, or in the feeling of its imperfection or withering." However, he warned against submitting to a single dominating passion. In a phrase that calls to mind both Blaise Pascal and Reinhold Niebuhr, Vauvenargues said of man, "The feeling of his imperfection makes his eternal torture." Although he believed that man's need for greatness and importance is laudable, he also maintained that men should respond with charity to the needs of others. Vauvenargues's moments of humanitarianism, however, were devoid of sentimentalism.

Vauvenargues wished to defend the value of self-interest, which is naturally a good, and also to preserve the ethical character of acts. He adopted two main approaches. Before Jean-Jacques Rousseau did, Vauvenargues distinguished between *amour propre* and *amour de nous-mêmes. Amour de nous-mêmes* allows us to seek happiness outside ourselves: "One is not his own unique object." There is, then, a difference between the satisfaction of *amour propre* and its sacrifice. Against those who held that all acts are motivated by self-interest Vauvenargues maintained that it is absurd to call sacrifice of life,

for example, an act of self-interest, for in such an act we consider ourselves as the least part of the whole and lose everything. Still combating La Rochefoucauld, Vauvenargues also argued that the criterion of acts is their effect on others; acts are virtuous if they tend to the good of all, even if they also satisfy self-interest. This definition opened a line of argument that had dangerous consequences in the hands of the materialists: (1) If each man must satisfy his self-interest where he can, men may be considered "fortunately born" or "unfortunately born" but not responsible for their acts. (2) Ethical and political considerations became fused, and eventually, with Rousseau, Johann Gottlieb Fichte, and G. W. F. Hegel, this led to the concept of the "ethical state." How should acts be judged? "Reason deceives us more often than the heart," declared Vauvenargues; like Rousseau, he trusted the "first impulse."

Vauvenargues believed that in regard to happiness, too, each man must follow his fated way; no philosophical formula can guide him. But he did offer one principle: "There is no enjoyment except in proportion as one acts, and our soul possesses itself truly only when it exerts itself completely." To give up action is to fall into nothingness. Existence is a function of becoming. Vauvenargues satirized pitilessly both the indolent and those who engage in aimless agitation. Activity, courage, glory, and ambition summarize his ideal of life and his concept of virtue. Greatness of soul is consistent with evil, as in Catiline; all depends on character and education. The great soul does not care about public esteem; true glory is an intimate feeling, self-satisfying to the point where it may paradoxically disdain action.

Although Vauvenargues was not interested in political philosophy, he did argue against the notion that men are, or may be naturally, politically or socially equal: "Law cannot make men equal in spite of nature." Hierarchy, in all respects, is inevitable.

Vauvenargues frequently espoused contradictory views. Although he developed no important theoretical positions, he occupies a leading rank in the long line of what the French term "moralists," excelling in psychological portraits and the striking but abstract formula of the maxim.

See also Ethics, History of; Fichte, Johann Gottlieb; Hegel, Georg Wilhelm Friedrich; La Bruyère, Jean de; La Rochefoucauld, Duc François de; Niebuhr, Reinhold; Pascal, Blaise; Plutarch of Chaeronea; Rousseau, Jean-Jacques; Seneca, Lucius Annaeus; Spinoza, Benedict (Baruch) de; Stoicism; Voltaire, François-Marie Arouet de.

Bibliography

Vauvenargues's works were published in a collected edition as *Oeuvres completes*, in 4 vols. (Paris, 1821); his moral works were published as *Oeuvres morales*, in 3 vols. (Paris, 1874).

For literature on Vauvenargues, see F. Vial, *Une philosophie et une morale du sentiment: Luc de Clapiers, marquis de Vauvenargues* (Paris: Droz, 1938), and G. Lanson, *Le Marquis de Vauvenargues* (Paris: Hachette, 1930).

OTHER RECOMMENDED WORKS

Fine, Peter M. *Vauvenargues and La Rochefoucauld*. Manchester, U.K.: Manchester University Press, 1974.

Meagher, Paul. "Vauvenargues' Quest for a New Moral Philosophy." *Lumen* 13 (1994): 125–136.

L. G. Crocker (1967)

Bibliography updated by Tamra Frei (2005)

VAZ FERREIRA, CARLOS

(1872–1958)

Carlos Vaz Ferreira, the Uruguayan educator and philosopher, was born in Montevideo. He became a professor of philosophy and rector at the University of Montevideo and played a prominent part in the theory and administration of primary and secondary education in Uruguay. He wrote voluminously and was a popular lecturer. As a result, he was for several decades a major intellectual force in his country. At various times and in various respects, he was influenced by Herbert Spencer, J. S. Mill, William James, and Henri Bergson, without full commitment to any of them.

Vaz Ferreira was impressed by the fluid complexity of experience, thought, and reality. Words and logical forms impose false precision and system on the contents of thought. The remedy is not a flight from reason but the development of a plastic reason close to experience, life, and instinct, alert to degrees of probability and unwilling to assent beyond the warrant of the question and evidence. The formulation and disposition of metaphysical questions requires the highest degree of caution, but metaphysics is both legitimate and necessary. It is impossible to move far in science without running into metaphysical questions, and it is necessary to cultivate metaphysics in order to understand the symbolic and limited nature of science and to counteract the bad metaphysics that comes into being when metaphysics is neglected. Vaz Ferreira was critical of positive religion but

was sympathetic to religion as the emotional apprehension of a possible transcendent being.

The ethics of Vaz Ferreira showed the same skepticism fused with marked human warmth and moral insight. Ethical principles cannot be stated without exceptions or descent into casuistry. Ideals clash and choices are usually between alternatives that contain some evil. An ethically sensitive person therefore is more subject than others to doubt, crisis, and remorse: satisfied conscience is more readily found in those who have a narrow awareness and ready formulas. But an ethically sensitive person may exemplify the perfection of individual morality, in which are combined a feeling for each individual act and a care for all possible results. Vaz Ferreira held that there has been moral progress in the course of history: Ideals have been added from time to time, more persons now share to some degree in all ideals, and there is greater resistance to evil.

See also Appearance and Reality; Bergson, Henri; Experience; James, William; Latin American Philosophy; Metaphysics; Mill, John Stuart.

Bibliography

Vaz Ferreira's principal works include *Los problemas de la libertad* (Problems of liberty; Montevideo, 1907); *Conocimiento y acción* (Knowledge and action; Montevideo: Mariño y Caballero, 1908); *Moral para intelectuales* (Ethics for intellectuals; Montevideo, 1909); *El pragmatismo* (Montevideo, 1909); *Lógica viva* (Living logic; Montevideo, 1910); *Sobre los problemas sociales* (On social problems; Montevideo, 1922); *Fermentario* (Montevideo, 1938).

See also Arturo Ardao, *Introducción a Vaz Ferreira* (Montevideo: Barreiro y Ramos, 1961).

Arthur Berndtson (1967)

VEBLEN, THORSTEIN BUNDE

(1857–1929)

Thorstein Bunde Veblen, the American economist and social theorist, is perhaps best known for his ironic style, a style that was at one with his life. Although he is still thought of abroad as the most influential American social scientist, among social scientists in America his influence has almost vanished. He is virtually unknown to college students, even if a scattered lot of Veblen's concepts—most obviously, "conspicuous consumption"—are unwittingly part of their speech and analyses.

Born on a Wisconsin farm, Veblen developed the most comprehensive and penetrating analysis of American industrial society in the early twentieth century. He emphasized qualitative relationships in the historical process, and his aim was an inclusive theory of social change. However, the largest number of those who have walked in Veblen's footsteps are known for quantitative, essentially unhistorical, often antitheoretical investigations. Where his followers have not deviated from his work in these ways, they have in another: Veblen called for, if he did not usually practice, dispassionate social analysis; many of his most fervent disciples are also quite fervent in their social analyses.

Like his contemporary, Charles S. Peirce, Veblen was a scholar of great intellectual achievement whose academic career was, at best, undistinguished. He took his doctorate in philosophy at Yale, whence he moved to Cornell to study economics. In a year he moved to the new University of Chicago, where he taught, and he also edited the *Journal of Political Economy.* Before long acrimony between Veblen and the administration over his academic and social nonconformity developed to a point where the happiest step for all concerned was for Veblen to leave Chicago. That experience, added to by similar ones at his next teaching post at Stanford, prompted Veblen to write one of his most scathing, if also very useful and sound, books: *The Higher Learning in America: A Memorandum on the Conduct of Universities by Businessmen* (New York, 1918). The original subtitle, abandoned for one reason or another, was "A Study in Total Depravity."

Stanford and Veblen failed to cement relations, and Veblen drifted to the University of Missouri, where he was sheltered by the eminent economist Herbert Davenport. Lectures at the New School for Social Research in New York City, and a brief interlude with the federal government, for which he wrote memoranda connected with World War I, ended Veblen's professional career. The department of economics at Cornell chose to add him to its faculty but that wish was denied by the university administration. Veblen spent his last few years unproductively, in a cabin in the Stanford hills, where he died, embittered against society.

The prime influences on Veblen appear to have been David Hume, Charles Darwin, and Karl Marx—although the influence of each was much transmuted by the mind and the circumstances of Veblen. The skepticism of Hume and the evolutionary approach of Darwin combined with the American scene to impel Veblen to launch a barrage of telling criticism (in essays in *The Place of Sci-*

ence in Modern Civilization, New York, 1919) at what he took to be the metaphysical, teleological, and optimistic qualities of Marxian analysis. But Veblen was not so much a critic as an adaptor of Marx, and his own works may be looked at most usefully in that light.

Darwinian concepts aside, the starting point of Veblen's analysis of society and of social change was fundamentally Marxian. The relationship of tension and change that Marx attributed to the conflict between "the forces of production" and "the mode of production" are present in Veblen's close equivalents, technology and institutions. For both men this relationship deserves and requires investigation within a framework of history (for Marx) or the genetic process (for Veblen).

But if the starting point for Veblen was the same as that of Marx, it was also there that basic similarities ended. For Marx the nineteenth-century assumptions of rationality went unquestioned, but for Veblen those assumptions were high on the list of matters to be investigated. As a consequence Veblen believed that a theory of social change required the integration of social psychology (and the psychology of related matters, such as nationalism and patriotism) with economics, politics, and history. Stemming from this is another difference: For Marx there were "general laws of motion of capitalist society" discoverable by the investigator; for Veblen those general laws had to be so qualified by national and cultural differences that it was not only plausible but also probable that capitalism would work out differently in different nations. Thus the very general quality of the conclusions to be found in *Capital,* when compared with Veblen's differing expectations for capitalism in Great Britain and Germany (in *Imperial Germany and the Industrial Revolution,* New York, 1915) and in the United States (in *The Theory of Business Enterprise,* New York, 1904, and in *Absentee Ownership,* New York, 1923). The point is illustrated by Veblen's findings about Japan and Germany, which (with much prescience) he saw as facing very much the same future despite their very different economic histories. For Veblen the decisive factors for the two nations were those making for extreme nationalism and social irrationality, moving them in much the same direction at much the same speed.

There is a final and striking difference between Marx and Veblen. In addition to his role as a social scientist, Marx was a political activist and propagandist, and his scientific writings were integrally connected with his political aims, concerning which Marx was optimistic. Veblen was politically aloof, except for a few periods such as his wartime propagandistic activity, and his role was that of Cassandra. Marx saw the class struggle as the means by which the contradictions between the forces and the mode of production would one day necessarily bring about the desired socialist society. Although Veblen would have found that socialist society less repulsive than the capitalist society he analyzed, his mood was gloomy and his vision apocalyptic, as suggested in one of his better-known but by no means unrepresentative observations in *The Instinct of Workmanship* (New York, 1914, p. 25): "history records more frequent and more spectacular instances of the triumph of imbecile institutions over life and culture than of peoples who have saved themselves alive out of a desperately precarious institutional situation, such, for instance, as now faces the people of Christendom."

Veblen's critical energies were spent most persistently in attacking the business system and nationalism, in that order. But he reserved his most savage wit for organized religion, which he considered a special—and the most successful—form of salesmanship (see the appendix to Ch. 11 of *Absentee Ownership*), manned by mental defectives whose business it is "to promise everything and deliver nothing."

See also Darwin, Charles Robert; Hume, David; Marx, Karl; Nationalism; Peirce, Charles Sanders; Philosophy of Social Sciences.

Bibliography

Of Veblen's 11 books, his first, *The Theory of the Leisure Class* (New York: Macmillan, 1899), was the most influential and most fundamental. See also *The Engineers and the Price System* (1919; New York: Harcourt Brace, 1963).

For the definitive biography of Veblen, see Joseph Dorfman, *Thorstein Veblen and His America* (New York: Viking Press, 1934).

For a more recent account of Veblen as an economist, see Douglas Dowd, *Thorstein Veblen* (New York: Washington Square Press, 1964). See also David Riesman, *Thorstein Veblen, a Critical Interpretation* (New York: Scribners, 1953; 2nd ed., 1960). Riesman is one of Veblen's severer critics.

ADDITIONAL SOURCES

Diggins, John P. *The Bard of Savagery: Thorstein Veblen and Modern Social Theory.* New York: Seabury Press, 1978.

Qualey, Carlton C., ed. *Thorstein Veblen; the Carleton College Veblen Seminar Essays.* New York: Columbia University Press, 1968.

Tilman, Rick. *The Intellectual Legacy of Thorstein Veblen.* Westport, CT: Greenwood Press, 1996.

Douglas F. Dowd (1967)
Bibliography updated by Philip Reed (2005)

VECCHIO, GIORGIO DEL

See *Del Vecchio, Giorgio*

VEDA

See *Indian Philosophy*

VEDANTA

See *Indian Philosophy*

VENN, JOHN

(1834–1923)

The British logician John Venn was born at Drypool, Hull, the elder son of the Reverend Henry Venn, a prominent evangelical divine. After early education at Highgate and Islington proprietary schools, he entered Gonville and Caius College, Cambridge, in 1853. On graduating Sixth Wrangler in 1857, he became a fellow and remained on the foundation for sixty-six years, until his death. During the last twenty years of his residence he was also president of the college. Venn took orders in 1858 and served as a curate in parishes near London before returning to Cambridge as college lecturer in moral sciences in 1862. He married in 1867. In 1869 he was Hulsean lecturer and published thereafter a work titled *On Some Characteristics of Belief* (London, 1870), but contact with Henry Sidgwick and other Cambridge agnostics, plus the reading of Augustus De Morgan, George Boole, J. Austin, and J. S. Mill had the effect of transferring his interests from theology almost wholly to logic, and in 1883 he gave up his orders without altogether withdrawing from the church. In the same year he became a fellow of the Royal Society and took the degree of doctor of science.

Venn was among those responsible for the development of the moral sciences tripos at Cambridge and in the course of his teaching published successively the three works by which he is now remembered: *The Logic of Chance* (London, 1866; 3rd ed., 1888); *Symbolic Logic* (London, 1881; 2nd ed., 1894); and *The Principles of Empirical or Inductive Logic* (London, 1889; 2nd ed., 1907). In 1888 he presented his extensive collection of books on logic to the university library, and he turned in later years to antiquarian pursuits, writing the history of his college and his family and collaborating with his son, J. A. Venn, in the preparation of Part I of *Alumni Cantabrigienses* (4 vols., London, 1922). Venn was an accomplished linguist and throughout most of his long life an active botanist and mountaineer. In addition to designing a simple mechanical contrivance to illustrate his well-known logical diagrams, he is said to have invented a very successful machine for bowling at cricket.

Venn has no strong claim to be regarded as an original thinker. His general position in philosophy was that of an orthodox, though unusually cautious and skeptical, empiricist. Outside the fields of logic and methodology he contributed little of importance, and even within them his role was essentially that of a critic and expositor of ideas first mooted by other men. In that capacity, however, his writings are marked by an acumen, learning, and lucidity that rank them among the best productions of their day. Within its limits, therefore, his reputation is still a high one.

LOGIC

Venn was a follower of Boole and to a lesser extent of Mill and a defender of both against the criticisms of William Stanley Jevons on the one hand and of the idealist logicians on the other. His *Symbolic Logic* is an attempt to show not merely that the Boolean algebra "works" but also that it is in the main line of historical tradition and that its supposedly mathematical obscurities are in fact intelligible from a purely logical point of view. Like De Morgan, he is aware of the element of convention in the choice of a logical standpoint and hence of the possibility of alternative versions of the basic propositional forms. He thus contrasts the four Aristotelian (or "predicative") types of proposition with the eight forms of Sir William Hamilton (which reduce on analysis to the five possible relations of inclusion and exclusion between pairs of classes), and compares them both with the fifteen possibilities that arise on his own "existential" view, based on the emptiness or occupancy of the four "compartments" marked out by a pair of terms and their negatives. Unlike some of his predecessors, he sees the difference as one of convenience rather than correctness, and so finds it unnecessary to dispute the merits of the older logic in order to vindicate the claims of the new. A similar tolerance is apparent in his treatment of the vexed issue concerning the "existential import" of propositions, where, after careful discussion, he opts for the presumption that universal propositions do not imply the existence of members in the subject class—a view that the great majority of writers from J. M. Keynes onward have since found reason to accept. Less open-minded, perhaps, is his attitude to Jevons's reforms of the Boolean calculus; but

he made several improvements of his own, notably in the writing of particular propositions as inequations, and, by the introduction of his diagrammatic methods, he did more than anyone else to render the workings of that calculus intelligible to the nonmathematical mind.

PROBABILITY

The Logic of Chance is also a work of much value to those embroiled in the mathematical complications of the theory of probability. The rationalistic handling of this subject by earlier writers was not to Venn's taste, and he recognized more clearly than they did the difficulties of relating their a priori computations to the realities of uncertain reasoning in everyday life. Following the suggestions of Leslie Ellis, he therefore identifies the probability of events not with the amount of belief it is rational to have in them but with their statistical frequency of occurrence in the generic class of events to which they belong. He assumes, that is, that the world contains series of resembling events in which individual irregularity in the possession of properties is combined with aggregate regularity "in the long run." The assignment of probability to a type of event is thus a mere matter of ascertaining the relative frequency with which it tends, increasingly, to occur as the series is extended to large numbers; and this is, in principle, not a subjective affair but a perfectly empirical and objective type of inquiry into the properties of a certain kind of group. To define probability in this way is, as Venn realized, to restrict it more narrowly than is usually done. No meaning can properly be attached to the probability of a single event, and the notion becomes equally inapplicable to the large range of judgments expressing partial belief (in theories and the like) that had hitherto been dealt with under this head. There are difficulties, moreover (as he also recognized), in assuming that observed frequencies are a reliable clue to "long-run" or "limiting" frequencies—that it is possible, in effect, on inductive grounds to arrive at such long-run frequencies by means of sample observations, however extended. For such a conclusion can itself be only probable, and that in a sense which Venn does not offer to define. Thus a knowledge of statistical frequency, even if obtainable, would be no sufficient ground for preferring one expectation to another. Probability, as Venn conceives it, is clearly not the guide of life.

SCIENTIFIC METHOD

The frequency theory of probability has had able defenders since Venn's time and is now less vulnerable to criticism. His version of it remains, however, the classical one, and the majority of later exponents acknowledge their debt to him. By comparison, the scientific methodology set forth in *Empirical Logic* has suffered somewhat from its association with that of Mill, on which it is largely modeled and whose conclusions it largely accepts. Venn differs from Mill chiefly in setting greater store by laws of coexistence than by laws of causal succession. The idea of causation he considers too crude and popular in conception to be of much use in science, and he is accordingly skeptical as to the value of the inductive methods. So far from being a reliable instrument for the discovery of causes, Mill's canons of induction are effective, he thinks, only where the conditions of the problem and its possible solutions have been narrowly circumscribed in advance, and under ordinary circumstances this can seldom be done. Inductive procedures are thus by no means so conclusive as Mill supposed, though we are not therefore justified in assuming, with Jevons, that they can be rationalized by appeal to the calculus of probability. Judgments of probability themselves make use of induction, and the two must therefore be kept, so far as possible, distinct. More generally, the use of formal methods in the classification, ordering, and prediction of natural phenomena can never be more than approximate, owing to the number of simplifying assumptions necessary before it can get under way. Venn's subsidiary discussions of definition, division, hypothesis, measurement, and so on, are similarly concerned to stress the difficulties of applying principles to cases and the amount that is taken for granted in doing so. Though less closely acquainted than some other writers with the details of scientific practice, he is also less liable than most to mistake the logic of science for a description of its technique.

See also Austin, John; Boole, George; British Philosophy; De Morgan, Augustus; Hamilton, William; Jevons, William Stanley; Logic Diagrams; Logic, History of; Mill, John Stuart; Mill's Methods of Induction; Probability and Chance; Sidgwick, Henry.

Bibliography

Venn has been somewhat neglected by historians of philosophy and no comprehensive study of him exists. For a serviceable brief account, see J. A. Passmore, *A Hundred Years of Philosophy* (London: Duckworth, 1957), pp. 134–136. His views on probability are most fully criticized in J. M. Keynes, *Treatise on Probability* (London: Macmillan, 1921).

P. L. Heath (1967)

VERIFIABILITY PRINCIPLE

The most distinctive doctrine of the logical positivists was that for any sentence to be cognitively meaningful it must express a statement that is either analytic or empirically verifiable. It was allowed that sentences may have "emotive," "imperative," and other kinds of meaning (for example, "What a lovely present!" or "Bring me a glass of water!") even when they have no cognitive meaning, that is, when they do not express anything that could be true or false, or a possible subject of knowledge. But—leaving aside sentences expressing analytic statements—for a sentence to have "cognitive," "factual," "descriptive," or "literal" meaning (for example, "The sun is 93 million miles from the earth") it was held that it must express a statement that could, at least in principle, be shown to be true or false, or to some degree probable, by reference to empirical observations. The iconoclasm of the logical positivists was based on this criterion of meaning, for according to the verifiability principle a great many of the sentences of traditional philosophy (for example, "Reality is spiritual," "The moral rightness of an action is a nonempirical property," "Beauty is significant form," "God created the world for the fulfillment of his purpose") must be cognitively meaningless. Hence, like Ludwig Wittgenstein in the *Tractatus Logico-Philosophicus*, they held that most of the statements to be found in traditional philosophy are not false but nonsensical. The verifiability principle, it was maintained, demonstrates the impossibility of metaphysics, and from this it was concluded that empirical science is the only method by which we can have knowledge concerning the world.

The verifiability principle stands historically in a line of direct descent from the empiricism of David Hume, J. S. Mill, and Ernst Mach. It has some affinities with pragmatism and operationalism, but it differs from them in some important respects. Pragmatism, as presented by C. S. Peirce, William James, and John Dewey, is the view that the "intellectual purport" of any symbol consists entirely in the practical effects, both on our conduct and on our experiences, that would follow from "acceptance of the symbol." This view, unlike the verifiability principle, makes the meaning of a sentence relative to certain human interests and purposes and to the behavior adopted for the realization of these purposes. Operationalism, as held by P. W. Bridgman and others, is the view that the meaning of a term is simply the set of operations that must be performed in order to apply the term in a given instance. Thus, according to this view, the meaning, or rather *a* meaning, of the term *length* is given by specifying a set of operations to be carried out with a measuring rod. Moritz Schlick and other logical positivists sometimes said that the meaning of a sentence is the method of its verification. But, unlike the advocates of operationalism, they meant by "the method of verification" not an actual procedure but the logical possibility of verification. The verifiability principle had among its immediate antecedents Schlick's *Allgemeine Erkenntnislehre* (Berlin, 1918) and Rudolf Carnap's *Der logische Aufbau der Welt* (Berlin, 1928). It was first formulated explicitly by Friedrich Waismann in his "Logische Analyse des Wahrscheinlichkeitsbegriffs" (1930) and subsequently by Schlick, Carnap, Otto Neurath, Hans Reichenbach, Carl Hempel, A. J. Ayer, and other logical positivists in numerous publications.

PROBLEMS RAISED BY THE PRINCIPLE

The controversial questions concerning the principle are: (1) What is it to be applied to—propositions, statements or sentences? (2) Is it a criterion for determining what the meaning of any particular sentence is, or is it simply a criterion of whether a sentence is meaningful? (3) What is meant by saying that a statement is verifi*able*, or falsifi*able*, even if in practice it has not been, and perhaps cannot be, verified, or falsified? (4) What type of statement directly reports an empirical observation, and how do we ascertain the truth-value of such a statement? (5) Is the principle itself either analytic or empirically verifiable, and if not, in what sense is it meaningful? (6) Is the question that the principle is intended to answer (that is, the question "By what general criterion can the meaning or the meaningfulness of a sentence be determined?") a logically legitimate question?

WHAT IS THE PRINCIPLE TO BE APPLIED TO? In some of the earlier formulations of the verifiability principle it is presented as a criterion for distinguishing between meaningful and meaningless propositions. However, in an accepted philosophical usage, every proposition is either true or false, and hence a fortiori a proposition cannot be meaningless. To meet this point some of the later exponents of the principle say that a grammatically well-formed indicative sentence, whether it is cognitively meaningful or not, expresses a "statement"; the term *proposition* is retained for what is expressed by a cognitively meaningful sentence—that is, propositions are treated as a subclass of statements. The verifiability principle is then presented as a criterion for distinguishing between meaningful and meaningless statements. This procedure, however, presupposes a usage for "cognitively meaningful sentence," and indeed it is sentences that are normally said to be meaningful or not.

Consequently, in still other formulations the principle is presented as applying directly to sentences; the objection to this is that sentences are not normally said to be true or false, and hence they are not said to be verifiable or falsifiable.

In order to meet these difficulties, sentences, statements, and propositions may be distinguished in the following way: A sentence, as we shall understand it, belongs to a particular language, it is meaningful or not, but it is not properly said to be true or false, or to stand in logical relations to other sentences, or to be verifiable or falsifiable. A statement is what is expressed in certain circumstances by an indicative sentence, and the same statement may be expressed by different sentences in the same or in different languages; a statement is properly said to be true or false, it does stand in logical relations to other statements, and it is verifiable or falsifiable. What can or cannot be said of statements applies equally to propositions, except that a proposition cannot be meaningless, that is, it cannot be expressed by a meaningless sentence.

For convenience we shall sometimes speak of sentences as being verifiable or not, and of statements as being meaningful or not. But, more strictly, we shall understand the verifiability principle as claiming that the cognitive meaning or meaningfulness of a *sentence* is to be determined by reference to the verifiability (or falsifiability) of the *statement* expressed by the sentence.

A CRITERION OF MEANING OR MEANINGFULNESS? The earliest presentations of the verifiability principle identified the meaning of a sentence with the logical possibility of verifying the corresponding statement, and apparently, in the last analysis, with the occurrence of certain experiences. This has some initial plausibility in the case of "empirical sentences," that is, sentences containing, apart from nondescriptive expressions, only empirical predicates (for example, "red," "round," "middle C"). An empirical predicate is, by definition, one that stands for a property that can be observed or experienced. Consequently, in the case of such a sentence as "This is red," there is a natural tendency to say that the meaning of the sentence is given by the experience that would verify it. The meaning is understood by anyone who can use the sentence for the purpose of identifying red objects when he sees them and cannot be understood by anyone who cannot identify red objects. It might be argued that a congenitally blind person could be said to understand the sentence "This is red" if he were able to identify red objects in some other way, by touch, for example. But in that case, an early adherent of the verifiability principle might reply, the predicate "red" has, for the person in question, not a visual but a tactual meaning. Our ability to understand empirical predicates, he might say, is plainly restricted by our capacity for sensory discrimination. For example, a person may be able to give a verbal definition of "C♭" as "the note midway between the notes designated by 'C' and 'C♯'"; but there is an important sense in which he does not know what "C♭" means if he is not able to discriminate quarter tones. It may be fairly objected, however, that this argument rests on the ambiguities of the words *meaning, stands for*, and *designates*; for example, the sense in which a term may be said to have a "tactual meaning" if it designates something tactual is not the sense in which a sentence may have a "cognitive or factual meaning." Moreover, it cannot be correct to identify the meaning of a sentence with the experiences that would verify it, for the characteristics that can be appropriately attributed to an experience cannot be appropriately attributed to the meaning of a sentence, nor conversely—for example, the meaning of a sentence does not occur at a particular time or with a certain intensity, as does an experience. And finally, if the meaning of a sentence were identified with the experiences of a particular person, the verifiability principle would result in a radical form of solipsism.

To meet these objections some other early formulations of the principle identified the meaning of a statement with that of some finite conjunction of statements directly reporting empirical observations. As will appear in more detail later, there are two main replies to this: (1) there are many types of statement whose meaning is not equivalent to that of any finite conjunction of observation statements, and (2) to identify the meaning of one statement with that of another is simply to say that the two statements have the same meaning, and this is not to explain or to give the meaning of the original statement.

For the foregoing reasons, it cannot be held that the verifiability principle is a criterion for determining the *meaning* of any particular sentence. In its later formulations it is presented simply as a criterion for determining whether a sentence is cognitively or factually *meaningful*.

STRONG VERIFIABILITY. In their early formulations Waismann, Schlick, and others held that the cognitive meaning of a sentence is determined completely by the experiences that would verify it conclusively. According to Waismann, for example, in "Logische Analyse des Wahrscheinlichkeitsbegriffs," "Anyone uttering a sentence must know in which conditions he calls the statement true or false; if he is unable to state this, then he does not

know what he has said. A statement which cannot be verified conclusively is not verifiable at all; it is just devoid of any meaning." This was sometimes called the requirement of "strong verifiability." It says, in effect, that for any statement *S* to be cognitively meaningful there must be some finite consistent set of basic observation statements $O_1 \cdots O_n$, such that *S* entails and is entailed by the conjunction of $O_1 \cdots O_n$. The principal objections to this requirement are: (1) a strictly universal statement, that is, a statement covering an unlimited number of instances (for example, any statement of scientific law), is not logically equivalent to a conjunction of any finite number of observation statements and hence is not conclusively verifiable; (2) any singular statement about a physical object can in principle be the basis of an unlimited number of predictions and hence is not conclusively verifiable; (3) statements about past and future events, and statements about the experiences of other people, are not conclusively verifiable; (4) even if an existential statement (for example, "Red things exist" or "At least one thing is red") is verifiable in the required sense, its denial cannot be verifiable in this sense, for its denial (for example, "Red things do not exist" or "Everything is nonred") is a strictly universal statement. Hence, the requirement of strong verifiability would have the strange consequence that the denial of an existential statement would never be meaningful, and this would involve the rejection of the fundamental logical principle that if a statement *S* is true, then not-*S* is false, and that if *S* is false, then not-*S* is true; (5) if a statement *S* is meaningful by the present requirement and *N* is any meaningless statement, then the molecular statement *S or N* must be meaningful; (6) the present requirement presupposes that observation statements are conclusively verifiable, for unless this is so, no statement at all, not even a statement that is logically equivalent to a finite conjunction of observation statements, will be conclusively verifiable—or cognitively meaningful.

FALSIFIABILITY. It was sometimes suggested that conclusive falsifiability rather than conclusive verifiability should be the criterion of a cognitively meaningful statement. The criterion of conclusive falsifiability says, in effect, that a statement *S* is meaningful if and only if not-*S* is conclusively verifiable. Consequently, objections analogous to those already considered still apply: (1) existential statements are not conclusively falsifiable, for if *S* is an existential statement, not-*S* is a strictly universal statement; (2) even if a universal statement is conclusively falsifiable, its denial is not conclusively falsifiable, since its denial is an existential statement. Hence, the present criterion would have the consequence that the denial of a universal statement would never be meaningful, and again this would involve the rejection of the fundamental principle of logic mentioned before; (3) the present criterion is open to the special objection that a universal statement (for example, "Whatever is pure water boils at 100° C.") would be meaningful, that is, conclusively falsifiable, only if the corresponding negative existential statement (for example, "There is an instance of pure water that does not boil at 100° C.") were assertable, and a fortiori meaningful; but this negative existential statement would be meaningful, that is, conclusively falsifiable, only if the corresponding universal statement were assertable, and a fortiori meaningful. To escape from this circle it would be necessary to have a different and independent criterion of significance for either universal or existential statements; (4) if *S* is meaningful by the present requirement and *N* is any meaningless statement, then *S and N* must be meaningful; (5) again, the present requirement presupposes that basic observation statements are conclusively verifiable.

CONFIRMABILITY. To meet the preceding difficulties the later formulations of the verifiability principle require of a meaningful statement that it should be related to a set of observation statements in such a way that they provide not conclusive verifiability but simply some degree of evidential support for the original statement. This was sometimes called the requirement of "weak verifiability." It says that for any statement *S* to be cognitively meaningful there must be some set of basic observation statements $O_1 \cdots O_n$ such that *S* entails $O_1 \cdots O_n$ and that $O_1 \cdots O_n$ confirms, or gives some degree of probability to, S. A formulation of this kind was given by Ayer in the first edition of *Language, Truth and Logic* (1936). He held that a statement is verifiable, and hence meaningful, if one or more observation statements can be deduced from it, perhaps in conjunction with certain additional premises, without being deducible from these other premises alone. The qualification concerning additional premises is introduced to allow, among other things, theoretical statements in science to be verifiable.

But this formulation, as Ayer recognizes in the second edition of his book, permits any meaningless statement to be verifiable. For if *N* is any meaningless statement and *O* some observation statement, then from *N* together with the additional premise *if N then O* the observation statement *O* can be deduced, although *O* cannot be deduced from the additional premise alone. To meet objections of this kind Ayer introduces a number of conditions; he says (1) "a statement is directly verifiable if it is either itself an observation-statement, or is such that

in conjunction with one or more observation-statements it entails at least one observation-statement which is not deducible from these other premises alone," and (2) "a statement is indirectly verifiable if it satisfies the following conditions: First, that in conjunction with certain other premises it entails one or more directly verifiable statements that are not deducible from these other premises alone; and secondly, that these other premises do not include any statement that is not either analytic, or directly verifiable, or capable of being independently established as indirectly verifiable."

These conditions are designed inter alia to prevent obviously meaningless statements from being verifiable simply by occurring as components of verifiable molecular statements as in the objection to the requirement of strong verifiability (see above), and the objection to the requirement of conclusive falsifiability. The conditions are, however, insufficient for this purpose. As Hempel remarks, according to the present formulation if *S* is meaningful, then *S and N* will be meaningful, whatever statement *N* may be. And Alonzo Church has shown that given any three observation statements O_1, O_2, and O_3, no one of which entails either of the others, and any statement *N*, it is possible to construct a molecular statement from which it follows that either *N* or not-*N* is verifiable. Such a molecular statement is one of the form $(\sim O_1 \cdot O_2) \vee (O_3 \cdot \sim N)$. For $(\sim O_1 \cdot \sim O_2) \vee (O_3 \cdot \sim N)$ together with O_1 entails O_3, and so the molecular statement is directly verifiable; but *N* together with $(\sim O_1 \cdot O_2) \vee (O_3 \cdot \sim N)$ entails O_2, and therefore *N* is indirectly verifiable. Alternatively, $(\sim O_1 \cdot O_2) \vee (O_3 \cdot \sim N)$ may by itself entail O_2, and in that case $\sim N$ and O_3 also entail O_2, and therefore $\sim N$ is directly verifiable.

Difficulties of the kind raised by Hempel and Church obtain when a component of a molecular statement is superfluous as far as the verifiability of the molecular statement is concerned, that is, when the inclusion or exclusion of the component makes no difference to the verifiable entailments of the molecular statement. To eliminate components of this kind, R. Brown and J. Watling have proposed that for a molecular statement to be verifiable, either directly or indirectly, it must contain "only components whose deletion leaves a statement which entails verifiable statements not entailed by the original statement, or does not entail verifiable statements entailed by the original statement." This stipulation is designed to ensure that every component of a verifiable molecular statement either is independently verifiable (that is, "entails verifiable statements not entailed by the original statement") or else contributes to the meaning of the molecular statement in such a way that the molecular statement entails verifiable statements not entailed by any of its components (that is, any of the components alone "does not entail verifiable statements entailed by the original statement"). The intention of these stipulations is to ensure that a meaningless statement cannot occur as a component of a verifiable molecular statement and derive verifiability from the statement in which it occurs.

In two important articles titled "Testability and Meaning" (1936–1937), Carnap distinguished the testing of a sentence from its confirmation; a sentence is "testable" if we know of a particular procedure (for example, the carrying out of certain experiments) that would confirm to some degree either the sentence or its negation. A sentence is "confirmable" if we know what kind of evidence would confirm it, even though we do not know of a particular procedure for obtaining that evidence. Carnap considers four different criteria of significance—complete testability, complete confirmability, degree of testability, and degree of confirmability. All of these exclude metaphysical statements as being meaningless. The fourth criterion is the most liberal and admits into the class of meaningful statements empirical statements of the various kinds that were excluded by the requirement of conclusive verifiability or the requirement of conclusive falsifiability.

Each of Carnap's criteria determines a more or less restrictive form of empiricist language, and this, according to his view, is the same thing as a more or less restrictive form of empiricism. Carnap is largely concerned in these articles with giving a technical account of the formal features of such languages. One of the most serious difficulties he encounters is that of giving a satisfactory account of confirmability. His procedure is, in effect, to regard as cognitively meaningful all and only those statements that can be expressed in a formalized empiricist language.

Similarly, Hempel, in his article "Problems and Changes in the Empiricist Criterion of Meaning" (1950), discussed the proposal that a sentence has cognitive meaning if and only if it is translatable into an empiricist language. A formalized language is characterized by enumerating the formation and transformation rules of its syntax and the designation rules for the terms of its basic vocabulary. An empiricist language is one in which the basic vocabulary consists exclusively of empirical terms. As Hempel explains, dispositional terms may be introduced by means of "reduction sentences," and the theoretical constructs of the more advanced sciences (for

example, "electrical field," "absolute temperature," "gravitational potential") can be accommodated by allowing the language to include interpreted deductive systems.

Hempel claims for his criterion that it avoids many of the difficulties of the earlier formulations of the verifiability principle. The logic of a formalized language may ensure that no universal or existential statement is excluded from significance merely on account of its universal or existential form and also that for every significant statement its denial is also significant. The vocabulary and syntax of a formalized empiricist language ensures that no meaningless statement will be admitted as significant, even by occurring as a component of a verifiable molecular statement.

Nevertheless, leaving purely formal objections aside, the main difficulty of both Carnap's and Hempel's treatment of the verifiability principle is that of giving an adequate characterization of an empiricist language. An "empirical term" or an "observation predicate" is one that designates a property that is in principle observable, even though in fact it is never observed by anyone. But if the property has never in fact been observed, how are we to know that it is observable?

It may be said that a basic observation statement "*Pa*," asserting that an object *a* has the observable property *P*, is meaningful only if the experiences that would verify the statement could occur. But "could" here cannot mean "factually could," since we can speak meaningfully of occurrences that are factually impossible. Apparently what is meant is that the experiences in question must be logically possible. But then it seems that the only sense that can be given to saying that the *experiences* are logically possible is that the statement "*Pa*" is contingent. However, in "*Pa*" the object *a* is simply named or referred to, and the property *P* ascribed to it—and it seems that every statement of this form must be contingent. Thus, unless a further explanation of the expression "observation predicate" is forthcoming, we have no way of distinguishing between those basic observation statements that are meaningful and those that are not.

OBSERVATION STATEMENTS. Schlick, in an early article titled "A New Philosophy of Experience," claimed that to understand a proposition we must be able to indicate exactly the particular circumstances that would make it true and those that would make it false. "Circumstances" he defined as facts of experience; and thus it is experience that verifies or falsifies propositions. An obvious objection to this view is that sense experience is essentially private, and hence apparently the cognitive meaning of every statement must be essentially private. Schlick attempted to avoid this objection by distinguishing between the content and form of experience. The content, he said, is private and incommunicable—it can only be lived through. But the form of our experiences, he claimed, is expressible and communicable, and this is all that is required for scientific knowledge. However, Schlick's distinction between content and form cannot save his view from the objection of solipsism; for if the meaning of every descriptive expression is to be found, in the last analysis, in private experience, then this is so not only for qualitative words but also for the relational words that are supposed to describe the form of experience.

Thus, the first problem concerning statements reporting empirical observations is that they should be expressible in such a way that their meaning is not private to any one observer. The logical possibility of verifying a given statement can then be explained without mentioning the experiences of any particular person or indeed the experiences of anyone at all. If basic observation statements can be formulated in the required way, they express logically possible evidence, and hence any statement suitably related to a set of observation statements is verifiable in principle, even though no one is ever in a position to have the relevant experiences, that is, to verify the statement in question.

In order to achieve this result some adherents of the verifiability principle regard certain statements describing physical objects as basic (for example, "This is a black telephone"); others attempt to achieve the same result while still regarding sense-datum or phenomenal statements as basic (for example, "Here now a black patch" or "This seems to be a telephone"). In either case, there is the difficulty of explaining how these statements are related to the experiences that would verify them.

The question whether a statement reporting an empirical observation is conclusively verifiable is, as we have seen, of importance for the criterion of conclusive verifiability and for that of conclusive falsifiability. It has also been thought to be of importance for the criterion of weak verifiability or confirmability, for, it has been said, unless basic statements are certain, or in some sense incorrigible, no other statement can be even probable or confirmable. Finally, as we noted before, there is also the problem of explaining what is meant by saying that a basic observation statement is verifiable in principle, that is, that certain experiences are logically possible, if in fact the experiences in question never occur.

IS THE PRINCIPLE ITSELF MEANINGFUL? It is sometimes objected that the verifiability principle itself, according to the criterion it lays down, must be either analytic or empirically verifiable if it is to be cognitively meaningful. But if it is analytic, then it is tautological and uninformative; at best it only exemplifies a proposed use of the terms "cognitive meaning" and "understanding." And if it is empirically verifiable, then it is a contingent statement about the ordinary use or some technical use of these terms and at best is only confirmable to some degree by the relevant evidence. In either case, it is objected, the principle cannot be the decisive criterion of cognitive meaning that its adherents suppose it to be.

One reply to this objection is that a criterion that determines a certain class of statements cannot have the same logical status as the statements in question. For example, the statement that expresses the principle of causality in effect determines a class of statements, namely, the class of causal statements, but obviously it is not itself a causal statement. Similarly, the verifiability principle, which claims to delimit the class of cognitively meaningful statements, cannot be expected to have the same logical status as the statements it delimits.

In order to understand the status of the verifiability principle, in the form in which it was held by the logical positivists, the following considerations are relevant: (1) They claimed that an essential difference between their empiricism and the earlier empiricism of Hume, Mill, and Mach was that it was based not on any particular psychological assumptions but only on considerations of logic. They may have believed that it is factually impossible for us to have experiences radically different in kind from those that we now have, but they did not present the verifiability principle as stating or implying this. But then, if the possibility of mystical or religious experiences is allowed, it seems that at least some metaphysical statements are verifiable and therefore meaningful. This conclusion has been accepted by some later adherents of the verifiability principle, but it is evident that the logical positivists wished to present their criterion of meaning in such a way that it would exclude all metaphysical statements from the class of meaningful statements.

(2) It might be argued, as Ayer once did, that it is meaningful to say that mystics have unusual experiences, but that nevertheless we can have no grounds for supposing that their experiences are relevant to the truth or falsity of any statement of fact, since we have no grounds for thinking that the "object" of such experiences could be described in ordinary empirical terms. The statement "Mystics have experiences that they report by the sentence 'Reality is One'" is empirically verifiable in the ordinary way. But the statement "Reality is One" is not empirically verifiable in the ordinary way. To this, however, the mystic may reply that he can describe in ordinary empirical terms the kind of preparation or discipline he recommends, and if we are not willing to carry out the appropriate procedure we are simply refusing to consider the possibility of verifying mystical statements. The antimetaphysical import of the verifiability principle, he may say, is apparently based on the assumption that we cannot have experiences radically different in kind from those that we now have.

(3) Some of the logical positivists (Schlick, the early Ayer) claimed that the verifiability principle is in effect a statement of the sense of "cognitive or factual meaning" and "understanding" that is actually accepted in everyday life. Schlick, for example, said that the verifiability principle is "nothing but a simple statement of the way in which meaning is actually assigned to propositions, both in everyday life and in science. There never has been any other way, and it would be a grave error to suppose that we believe we have discovered a new conception of meaning that is contrary to common opinion and which we want to introduce into philosophy" ("Meaning and Verification"). But, as we have seen, if the verifiability principle is simply a contingent statement about a certain linguistic usage, its logical status cannot justify the degree of confidence that its adherents place in it.

(4) Finally, the principle has been regarded as a recommendation or a decision concerning the use of the expression "factually meaningful statement." It has been claimed that this decision prevents radical intellectual confusion and that it promotes clarity in the discussion of many philosophical questions. Carnap and Ayer, among others, have taken this view of the status of the verifiability principle. It should be noted that this does not imply that the principle is regarded as an analytic or necessarily true statement. A principle that expresses a linguistic recommendation is no doubt closely related to a corresponding analytic statement, but the recommendation itself is not tautological and uninformative. A recommendation or a decision has a different logical status; it is not successful by being true or unsuccessful by being false.

MORE RECENT CRITICISMS. Following the later work of Wittgenstein it is now widely held among philosophers that to ask whether a sentence is meaningful is simply to ask whether the words that compose the sentence are used according to the rules or practice of a language. Understanding a word, it is said, does not involve "know-

ing what the word stands for" or "being able to recognize what the word designates"; it involves only the ability to use the word in accordance with certain linguistic rules. Furthermore, the rules governing the correct use of different kinds of words differ enormously, and hence there is not just one way of misusing the words that occur in a sentence and thereby rendering the sentence meaningless. Each of the sentences "I do not exist," "The round square feels depressed," "Nonbeing is infinitely perfect," and "The Absolute enters into but transcends all change" involves a violation of one or more linguistic rules, but of quite different rules. Consequently, it is said, it is not possible to give a general criterion of the meaningfulness of a sentence. The verifiability principle is an attempt to answer the question "Under what conditions is a sentence cognitively or factually meaningful?," but this question, according to the view now widely held, is not one to which it is possible to give an answer that is both general and informative. Two further criticisms are made of the verifiability principle: (1) the principle, it is said, is not at all a criterion of the meaningfulness of a sentence but simply a characterization of an "empirical sentence," (2) the principle confuses the question of whether a sentence is meaningful with the different question of whether the statement it expresses can be known to be true or false. These more recent objections to the verifiability principle occur in most post-Wittgensteinian discussions of the topic of meaning. A useful summary of the arguments is given by J. L. Evans in "On Meaning and Verification."

Truth theory of meaning. It is convenient to begin by examining the second of these two further criticisms. It is concerned with the fact that one component of the verifiability principle is the thesis that the meaning of a statement is given by its truth conditions. This idea, which may be called "the truth theory of meaning," had been employed and stated by philosophers before the discussions of the Vienna circle. It is assumed, for example, by Bertrand Russell in his theory of descriptions. And Wittgenstein, in the *Tractatus*, said explicitly, "To understand a proposition means to know what is the case if it is true."

The formal correctness of this view can be seen from the following definition of the meaning of a statement in terms of its truth conditions. "*Die Sonne scheint* means *that the sun is shining* $=_{\mathrm{Df}}$ *Die Sonne scheint* is true if, and only if, the sun is shining"; in general, "*S* means *that p* $=_{\mathrm{Df}}$ *S* is true if, and only if, *p*." Nevertheless, it has to be admitted that the truth theory provides no effective clarification of the notion of cognitive or factual meaning. For even if the truth conditions of a statement *S* can be enumerated exhaustively in terms of a finite conjunction of observation statements $O_1 \ldots O_n$ (and, as we have seen, in very many cases this cannot be done) this entitles us to assert only that *S* and $O_1 \ldots O_n$ have the same meaning. But this does not clarify what the meaning of *S* is, or what it is for *S* to be meaningful. To say simply that two statements have the same meaning is not to say what either statement means or what it is for either statement to be meaningful.

For the kind of clarification that is being sought we now need a different and independent explanation of the meaning of an observation statement. Furthermore, the definition of the meaning of a statement in terms of its truth conditions provides no clarification unless the notion of truth is further explained. The truth of a statement can be defined in terms of its meaning in the following way. "*Die Sonne scheint* is true $=_{\mathrm{Df}}$ *Die Sonne scheint* means *that the sun is shining*, and the sun is shining"; in general "*S* is true $=_{\mathrm{Df}}$ *S* means *that p*, and *p*." But obviously it would be circular to employ this definition of truth in an attempt to clarify the notion of cognitive meaning. The two preceding definitions show, however, that there is a close connection between the notion of cognitive or factual meaning and the notion of truth. And hence, in reply to the second of the two further criticisms of the verifiability principle mentioned above, it may be argued that there must be a close connection between understanding a sentence as expressing a statement of fact and its being possible for one to know whether the statement is true or false.

Meaning and experience. The first of the two further criticisms of the verifiability principle is concerned with the fact that another component of the principle is the thesis that the truth conditions of a statement can be known only by reference to experience. This is the traditional doctrine of empiricism or positivism. The logical positivists (with the exception of Neurath, Carnap, and others, who at one time adopted a "coherence theory" of truth) held this view on the grounds that there are only two ways in which the truth-value of a statement can be ascertained, either a priori or a posteriori. According to their doctrine, if a statement can be known to be true a priori, then it is analytic and tautological and hence not a statement of fact. Therefore, if a statement is a statement of fact, it cannot be known a priori—its truth-value can be ascertained only by reference to experience. The simple dichotomy (either a priori or a posteriori) on which this argument is based has been criticized in more recent philosophy. W. V. Quine, for example, maintains that for the most part the statements that compose the corpus of

knowledge have their truth-values determined by linguistic and pragmatic considerations, as well as by the occurrence of certain sensory experiences. He allows, however, that statements "on the periphery" have their truth-values determined by experience. Thus, even in a more qualified version of empiricism the difficulty still remains of making clear what it is to know that a statement is true "by reference to experience."

Nevertheless, the criticism of the verifiability principle now being considered admits that for a sentence to be an "empirical sentence" it must express a statement that is in some sense verifiable, that is, the truth conditions of which can be known by reference to experience. And it may be argued that the grounds on which this is admitted are such that they compel a similar admission for every sentence that can be understood as expressing a statement of fact. It is evident that if a form of language can be used to describe the world—that is, to make statements—its rules cannot be wholly syntactical, that is, of the kind that govern simply the formation and transformation of sentences in the language. For the language to be descriptive it must also have semantic rules, for example, rules that relate the use of its basic predicates to certain states of affairs in the world. Semantic rules may be said to govern directly the use of basic predicates and to govern indirectly, via definitions and other syntactical means, the use of nonbasic predicates. The more detailed analysis of a semantic rule—that is, an account of how such rules function in a language—is a difficult matter that we need not attempt here. For our present purpose it is sufficient to note that it would be a contradiction to say that a language was descriptive but had no semantic rules; similarly, it would be a contradiction for someone to say that he could understand a sentence as expressing a statement although he had not been able to ascertain the semantic rules of the language in which the sentence was expressed.

We can now see why many present-day philosophers say that the verifiability principle is simply a characterization of an empirical sentence. If a sentence is used to describe an experienceable state of the world, then the semantic rules governing its predicates relate those predicates, directly or indirectly, to that state of the world. It follows that the sentence expresses a statement that is in principle verifiable. But consider the position of a philosopher who maintains that he uses certain sentences to make statements about the world, although these statements are not verifiable in any sense at all. This position seems to be simply incoherent. If the sentences in question express statements, the use of the predicates that occur in them must be governed by semantic rules; how can these rules be known or explained to anyone else if the states of affairs which the sentences are supposed to describe are not experienceable in any way at all? The philosopher in question may eventually admit that the relevant states of the world are, after all, experienceable—but intuitively or by some other special kind of experience. This, apparently, would be a psychological claim, to the effect that we are capable of types of experience other than those we usually associate with the normal functioning of our sense organs. The onus of proof to show that such experiences are possible plainly rests upon the philosopher in question. But even if such experiences do occur, and are of such a kind that they can be associated, via semantic rules, with the descriptive expressions of a language, this will not provide an exception to the requirement laid down by the verifiability principle—it will, in fact, be simply an extension of that requirement to types of sentences that formerly could not be understood as expressing statements of fact.

For a further examination of this question, it would seem that the correct approach would be to give a completely general analysis of "knowing the use of a predicate." Such an analysis cannot be given here, but the following outline may be suggested. In the case of a basic predicate it may be held that (1) an essential part of the use of the predicate is to identify a property, (2) an ability to use the predicate to identify the relevant property does not constitute knowing its use, unless the user also knows what the ability consists in, and (3) the user cannot be said to know this if it is impossible for him to have any kind of experience of the property in question.

Thus, to revert to the first and main criticism of the verifiability principle, it may be admitted that to ask whether a sentence is meaningful is to ask whether the constituent words are used according to the rules of a language. And it may be admitted that the rules governing the use of different kinds of words differ immensely and that there is not just one way in which a sentence can be meaningless. Nevertheless, if the foregoing remarks are correct, a sentence cannot be understood as expressing a statement unless the use of the descriptive expressions that occur in it are governed by semantic rules; and these rules cannot be known or explained to anyone else unless it is possible for the users of the language to have some kind of experience of the states of the world to which the descriptive expressions in question are related. These requirements are, perhaps, all that is essential in the claim made by the verifiability principle in its later formulations.

See also Basic Statements; Logical Positivism.

Bibliography

FORMULATIONS AND FAVORABLE DISCUSSIONS

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1936; 2nd ed., 1946.

Ayer, A. J. "Logical Positivism—A Debate" (with F. C. Copleston). In *A Modern Introduction to Philosophy*, edited by Paul Edwards and Arthur Pap. Glencoe, IL: Free Press, 1957; 2nd ed., rev., New York, 1965.

Ayer, A. J. "The Principle of Verifiability." *Mind* 45 (1936).

Ayer, A. J. "Verification and Experience." *PAS* 37 (1936–1937). Reprinted in *Logical Positivism*, edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

Brown, R., and J. Watling. "Amending the Verification Principle." *Analysis* 11 (1950–1951).

Carnap, Rudolf. "Testability and Meaning." *Philosophy of Science* 3 (1936): 419–471; 4 (1937): 1–40. Reprinted in *Readings in the Philosophy of Science*, edited by Herbert Feigl and May Brodbeck, 47–92. New York: Appleton-Century-Crofts, 1953.

Hempel, C. G. "The Concept of Cognitive Significance: A Reconsideration." *Proceedings of the American Academy of Arts and Sciences* 80 (1951): 61–77.

Hempel, C. G. "Problems and Changes in the Empiricist Criterion of Meaning." *Revue internationale de philosophie* 4 (1950): 41–63. Reprinted in *Semantics and the Philosophy of Language*, edited by Leonard Linsky. Urbana: University of Illinois Press, 1952. Also reprinted in *Logical Positivism*, edited by Ayer (see above).

Reichenbach, Hans. "The Verifiability Theory of Meaning." *Proceedings of the American Academy of Arts and Sciences* 80 (1951). Reprinted in *Readings in the Philosophy of Science*, edited by Herbert Feigl and May Brodbeck. New York: Appleton-Century-Crofts, 1953.

Schlick, Moritz. "Form and Content." Three lectures given in London in 1932. Reprinted in *Gesammelte Aufsätze*.

Schlick, Moritz. "Meaning and Verification." *Philosophical Review* 45 (1936): 339–368. Reprinted in *Gesammelte Aufsätze* and in *Readings in Philosophical Analysis*, edited by Herbert Feigl and Wilfrid Sellars. New York: Appleton-Century-Crofts, 1949.

Schlick, Moritz. "A New Philosophy of Experience." *Publications in Philosophy of the College of the Pacific*, No. 1. Stockton, CA, 1932. Reprinted in *Gesammelte Aufsätze 1926–1936*. Vienna: Gerold, 1938.

Schlick, Moritz. "Positivismus und Realismus." *Erkenntnis* 3 (1932–1933): 1–31. Reprinted in *Gesammelte Aufsätze* and translated in Ayer, ed., *Logical Positivism* (see above).

Waismann, Friedrich. "Logische Analyse der Wahrscheinlichkeitsbegriffs." *Erkenntnis* 1 (1930–1931).

Whiteley, C. H. "On Meaning and Verifiability." *Analysis* (1938–1939).

Whiteley, C. H. "On Understanding." *Mind* 58 (1949).

CRITICISMS

Barnes, W. H. F. "Meaning and Verifiability." *Philosophy* 14 (1939).

Berlin, Isaiah, "Verification." *PAS* 39 (1938–1939).

Church, Alonzo. Review of Ayer's *Language, Truth and Logic*, 2nd ed. *Journal of Symbolic Logic* 14 (1949).

Copleston. F. C. "Logical Positivism—A Debate" (with A. J. Ayer). In *A Modern Introduction to Philosophy*, edited by Edwards and Pap (see above).

Copleston, F. C. "A Note on Verification." *Mind* 59 (1950). Reprinted, with "A Further Note on Verification," in *Contemporary Philosophy*. London: Burns and Oates, 1956.

Ducasse, C. J. "Verification, Verifiability, and Meaningfulness." *Journal of Philosophy* 33 (1936): 230–236.

Evans, J. L. "On Meaning and Verification." *Mind* 62 (1953): 1–19.

Ewing, A. C. "Meaninglessness." *Mind* 46 (1937): 347–364.

Kneale, W. C. "Verifiability." *PAS*, Supp. 19 (1945).

Lazerowitz, Morris. "The Positivistic Use of 'Nonsense.'" *Mind* 57 (1946). Reprinted in *The Structure of Metaphysics*.

Lazerowitz, Morris. "The Principle of Verifiability." *Mind* 46 (1937): 372–378.

Lazerowitz, Morris. "Strong and Weak Verification." *Mind* 48 (1939) and 59 (1950). Reprinted in *The Structure of Metaphysics*. London: Routledge and Paul, 1955.

Lewis, C. I. "Experience and Meaning." *Philosophical Review* 43 (1934). Reprinted in *Readings in Philosophical Analysis*, edited by Herbert Feigl and May Sellars. New York: Appleton-Century-Crofts, 1953.

O'Connor, D. J. "Some Consequences of Professor Ayer's Verification Principle." *Analysis* 10 (1949–1950).

Russell, Bertrand. *An Inquiry into Meaning and Truth*. New York: Norton, 1940.

Russell, Bertrand. "On Verification." *PAS* 38 (1937–1938): 1–15.

Stace, W. T. "Metaphysics and Meaning." *Mind* 44 (1935): 417–438. Reprinted in *A Modern Introduction to Philosophy*, edited by Paul Edwards and Arthur Pap. Glencoe, IL: Free Press, 1957; 2nd ed., New York, 1965.

Stebbing, L. Susan. "Communication and Verification." *PAS*, Supp. 13 (1934).

Waismann, Friedrich. "Language Strata." In *Logic and Language*, edited by A. G. N. Flew. 2nd series. Oxford: Blackwell, 1953.

Waismann, Friedrich. "Verifiability." *PAS*, Supp. 19 (1945): 119–150. Reprinted in *Logic and Language*, edited by A. G. N. Flew. 1st series. Oxford, 1951.

Warnock, G. J. "Verification and the Use of Language." *Revue internationale de philosophie* 5 (1951): 307–322. Reprinted in *A Modern Introduction to Philosophy*, edited by Paul Edwards and Arthur Pap. Glencoe, IL: Free Press, 1957; 2nd ed., New York, 1965.

Wisdom, John. "Metaphysics and Verification." *Mind* 47 (1938): 452–498. Reprinted in *Philosophy and Psycho-analysis*. Oxford: Blackwell, 1953.

Wisdom, John. "Note on the New Edition of Professor Ayer's *Language, Truth and Logic*." *Mind* 57 (1948). Reprinted in *Philosophy and Psycho-analysis*.

RELATED DISCUSSIONS

Alston, W. P. "Pragmatism and the Verifiability Theory of Meaning." *Philosophical Studies* 6 (1955).

Bridgman, P. W. "Operational Analysis." *Philosophy of Science* 5 (1938).

Grice, H. P. "Meaning." *Philosophical Review* 46 (1957): 377–388.

Heath, A. E. "Communication and Verification." *PAS*, Supp. 13 (1934).

MacKinnon, D. M. "Verifiability." *PAS*, Supp. 19 (1945).

Morris, C. W. "The Concept of Meaning in Pragmatism and Logical Positivism," in *Proceedings of the Eighth International Congress of Philosophy* (1934). Prague, 1936.

Nagel, Ernest. "Verifiability, Truth, and Verification." *Journal of Philosophy* 31 (1934): 141–148. Reprinted in *Logic without Metaphysics*. Glencoe, IL: Free Press, 1956.

Nelson, E. J. "The Verification Theory of Meaning." *Philosophical Review* 43 (1954).

OTHER RECOMMENDED TITLES

Bealer, George. "The Incoherence of Empiricism." *The Aristotelian Society*, Supp. 66 (1992): 99–138. Reprinted in *Rationality and Naturalism*, edited by S. Wagner and R. Warner. Notre Dame, IN: University of Notre Dame Press, 1993.

Coffa, Alberto. *The Semantic Tradition from Kant to Carnap*. Cambridge, U.K.: Cambridge University Press, 1991.

Cohen, Jonathan. "Is a Criterion of Verifiability Possible?" *Midwest Studies in Philosophy* 5 (1980): 347–352.

Dummett, Michael. "The Metaphysics of Verificationism." In *The Philosophy of A. J. Ayer*, edited by Lewis Hahn. La Salle, IL: Open Court, 1992.

Edington, D. "The Paradox of Knowability." *Mind* 94 (1985): 557–568.

Feigl, H. "The Weiner Kreis in America." In *The Intellectual Migration*, edited by D. Fleming and B. Bailyn. Cambridge, MA: Harvard University Press, 1969.

Hempel, Carl. *Aspects of Scientific Explanation*. New York: Free Press, 1965.

Johnston, Mark. "Verificationism as Philosophical Narcissism." In *Philosophical Perspectives, 7: Language and Logic*, edited by James Tomberlin. Atascadero, CA: Ridgeview Press, 1993.

Okasha, Samir. "Verificationism, Realism, and Scepticism." *Erkenntnis* 55 (2001): 371–385.

Quine, W. V. O. "Two Dogmas of Empiricism." In *From a Logical Point of View: Nine Logico-Philosophical Essays*. Cambridge, MA: Harvard University Press, 1980.

Smart, J. J. C. "Verificationism." In *Cause, Mind, and Reality: Essays Honoring C. B. Martin*. Norwell, MA: Kluwer, 1989.

R. W. Ashby (1967)
Bibliography updated by Benjamin Fiedor (2005)

VERIFIABILITY PRINCIPLE [ADDENDUM]

The doctrines associated with the slogan that meaning is the mode of verification continued to develop in the last four decades of the twentieth century. While the exact formulation of the principle was itself controversial, the essential idea was to link semantic and epistemic concerns by letting the meaning of an expression be its role within an empirical epistemology. At the same time the fortunes of logical empiricism, the movement associated with verificationism, changed substantially as well. First, as philosophers who conspicuously did not identify themselves with logical empiricism moved to center stage, the movement as a separately identifiable phenomenon virtually ceased to exist. This did not dispose of verificationism, however, for often the later philosophers' views were strikingly similar to the logical empiricism that they supposedly replaced, just as the criticisms of logical empiricism were often pioneered by the logical empiricists themselves. The second major change in the fortunes of this view was the renewal of interest in the history of philosophy of science, especially in the histories of the logical empiricists themselves. Now freed from the myopia that comes from being part of the fray, philosophers were able to explore the roots of logical empiricism, what held it together as a movement, which of its doctrines were central or peripheral, and even which views look more plausible in hindsight than they did before their systematic interconnection could be appreciated.

One root of verificationism lies in the increasing professionalization of both the sciences and philosophy around the turn of the twentieth century. The sciences tended to emphasize the importance of empirical investigation, to explore its scope and limits, and to deplore as metaphysical any claims not based on evidence. Correspondingly, many philosophers claimed for themselves a nonempirical source of knowledge concerning things higher or deeper than mere observation could reveal, that is, concerning metaphysics. Logical empiricism grew out of methodological discussions within science rather than philosophy, and many of its central proponents were trained in the sciences. True, logical empiricism made special accommodation for the a priori domains of mathematics and logic. But these were technical subjects of use within the sciences and for which there were increasingly well-developed modes of conflict resolution. Moreover, the way in which the accommodation was reached, namely through the logical analysis of language, especially the language of science, comported well with a basic empiricism and provided no comfort to traditional philosophy.

A second root of verificationism lies in Bertrand Russell's reaction to the paradox that bears his name (viz., a contradiction that arises when sets can contain themselves) and in Ludwig Wittgenstein's further elaboration of a related idea. In order to avoid the paradox, Russell had restricted the grammar so that apparent assertions of sets containing themselves were no longer well formed. Similarly, Wittgenstein emphasized that some combinations of words were neither true nor false but just non-

sensical; they were, he said, metaphysical. This seemed to offer the ideal diagnosis of the sought-after distinction: Scientifically respectable claims were either empirically meaningful in virtue of having some appropriate relation to the observations that would be the source of their justification, or else they were true in virtue of the language itself; traditional metaphysics, by contrast, was simply unintelligible. Phrased in this way, the verifiability principle leaves as a separate question the issue of what the appropriate relation to observation would be.

It has also become clearer what the logical status of the principle itself is. Initially, these philosophers could imagine that they were saying something about language in general or about the language of science. But as it became apparent that there were alternative languages to be considered, it became obvious that the principle could be put as a proposal for a language or as an analytic or empirical claim either about a particular language or about a range of languages. Perhaps the dominant form of the principle is as a proposal for a language to explicate the linguistic practices that are already largely in place in the sciences. As a proposal, it is not a claim, and hence neither true nor false, but not thereby unintelligible. If the proposal is adopted, the corresponding claim about the language that has those rules would be analytic. There would also be the empirical claim that we had adopted such a language and even empirical claims about that language if it were specified as, say, the language that is now used in contemporary physics.

So construed, many of the objections that were first made to the principle (and which continued to be made through the period in question) can be seen to be wrongheaded. The most persistent of these criticisms is that the principle renders itself an unintelligible claim. Whether construed as a proposal, as an analytic claim, or as an empirical one, this is just a (willful) misunderstanding. The same can be said for the criticism that it renders all philosophy meaningless. Equally misguided is the repeated objection that the principle cannot be right because we can understand a sentence without knowing whether it is true. Obviously, the principle in no way denies this truism.

Potentially more serious is the idea that all attempts to specify the principle have failed and are thus likely to continue to do so. Reinforcing this idea are papers by Carl Gustav Hempel (1950, 1965) that, while they are not really histories, strike many readers as signed confessions of complicity in a series of disasters. In defense of the principle it must be said that, except for those immediately around Wittgenstein, complete verifiability was virtually never at issue. Even in the *Aufbau*, where the general question is raised many times, all but one formulation are much more liberal. Similarly, strict falsifiability was never proposed as a criterion of meaningfulness. Concerning the more fertile ground of confirmation and disconfirmation, the difficulties seem to have arisen because the formulations tried both to link semantic and epistemic concerns *and* to specify a complete theory of confirmation. This latter task is so difficult that we should not expect early success nor conclude from failure that the enterprise is misguided—any more than we give up physics simply because we still lack the final theory.

There were, of course, other sources of difficulty. Many attempts, such as A. J. Ayer's, tried to apply a criterion of meaningfulness at the level of whole sentences even though those sentences could contain meaningless parts. More successful in this regard was Rudolf Carnap's "Methodological Character of Theoretical Concepts" (1956), which applied the criterion at the level of primitive terms. In a paper that was famous despite being unpublished for many years, David Kaplan (1975) provided two counterexamples to Carnap's criterion. These examples were widely regarded as decisive, but Richard Creath (1976) showed that one of the examples missed its mark and the criterion could be patched in a natural way so as to avoid the other. Less easily dismissed is W. Rozeboom's (1960) criticism that Carnap's criterion ties meaningfulness to a particular theory when it should apply only to the language. Finally, Carnap's criterion, like many others, seems to presuppose that the theory/observation distinction can be drawn at the level of vocabulary. There came to be general agreement that this presupposition is mistaken and distorts any criterion based on it. In fairness, it must be admitted that some theory/observation distinction is essential to a healthy empiricism and that Carnap was from the very beginning fully aware of the limitations of formulating the distinction in this way. Finding a satisfactory way is still an unsolved problem.

W. V. O. Quine is often associated with the demise of logical empiricism, and his "Two Dogmas of Empiricism" (1951) is often thought to have rejected verificationism decisively. It would be more accurate to say that he rejected the idea that individual sentences could be separately confirmed, but he did not resist linking meaningfulness with confirmation holistically construed. Indeed, his demand that behavioral criteria be provided for analyticity to render it intelligible is exactly parallel to Carnap's demand for correspondence rules to render theoretical terms meaningful. Moreover, Quine's argument from the indeterminacy of translation to the unin-

telligibility of interlinguistic synonymy makes sense only if meaning and confirmation are somehow linked as in the verifiability principle.

So what then of this link between semantic and epistemic issues? At least there is much to be said for it. A theory of meaning should give accounts of meaningfulness (having a meaning), of synonymy (having the same meaning), and of understanding (knowing the meaning). The verifiability principle provides *a* way of doing these things not provided by simply identifying various entities as "the meanings" of expressions. Moreover, it provides *a* defense against wholesale skepticism by tying what we know to how we know. And finally, it provides *a* way of dealing with the so-called a priori by making those claims knowable in virtue of knowing the meanings of the expressions involved. No doubt there are others ways, perhaps even equally systematic ways, of accomplishing these ends, and no doubt these other paths should be investigated as well. But the basic idea behind the verifiability principle, namely that semantical and epistemic questions should be linked, is far from refuted, and its promise is far from exhausted.

See also Analyticity; Ayer, Alfred Jules; Carnap, Rudolf; Empiricism; Epistemology; Hempel, Carl Gustav; Language; Meaning; Philosophy; Philosophy of Science, History of; Philosophy of Science, Problems of; Quine, Willard Van Orman; Russell, Bertrand Arthur William; Semantics; Skepticism, History of; Verifiability Principle; Wittgenstein, Ludwig Josef Johann.

Bibliography

Antony, Louise. "Can Verificationists Make Mistakes?" *American Philosophical Quarterly* 24 (1987): 225–236.

Ayer, A. J. *Language, Truth and Logic*. London: Gollancz, 1936.

Ayer, A. J. "The Principle of Verifiability." *Mind* 45 (1936): 199–203.

Bercic, Boran. "On the Logical Status of the Principle of Verifiability." *Synthesis Philosophica* 15 (2000): 9–26.

Carnap, R. "The Methodological Character of Philosophic Problems." In *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, edited by H. Feigl and M. Scriven. Minneapolis, 1956.

Carnap, R. "On the Character of Philosophic Problems." Translated by W. M. Malisoff. *Philosophy of Science* 1 (1934): 5–19.

Carnap, R. "On the Use of Hilbert's ϵ-Operator in Scientific Theories." In *Essays on the Foundations of Mathematics*, edited by Y. Bar-Hillel et al. Jerusalem: Magnes Press, Hebrew University, 1961.

Coffa, Alberto. *The Semantic Tradition from Kant to Carnap*. Cambridge, U.K.: Cambridge University Press, 1991.

Cohen, Jonathan. "Is a Criterion of Verifiability Possible?" *Midwest Studies in Philosophy* 5 (1980): 347–352.

Creath, R. "On Kaplan on Carnap on Significance." *Philosophical Studies*30 (1976): 393–400.

Creath, R. "Was Carnap a Complete Verificationist in the *Aufbau?*" *PSA* 1 (1982): 384–393.

Dummett, Michael. "The Metaphysics of Verificationism." In *The Philosophy of A. J. Ayer*, edited by Lewis Hahn. La Salle, IL: Open Court, 1992.

Glymour, C. *Theory and Evidence*. Princeton, NJ: Princeton University Press, 1980.

Hempel, Carl. *Aspects of Scientific Explanation*. New York: Free Press, 1965.

Hempel, C. G. "Problems and Changes in the Empiricist Criterion of Meaning," *Revue internationale de philosophie* 4 (1950): 41–63.

Johnston, Mark. "Verificationism as Philosophical Narcissism." In *Philosophical Perspectives, 7: Language and Logic*, edited by James Tomberlin. Atascadero, CA: Ridgeview Press, 1993.

Kaplan, D. "Significance and Analyticity: A Comment on Some Recent Proposals of Carnap." In *Rudolf Carnap, Logical Empiricist: Materials and Perspectives*, edited by J. Hintikka. Dordrecht: Holland, 1975.

Moser, Paul K. "Epistemology (1900–Present)." In *Routledge History of Philosophy*, Vol. 10: *Philosophy of the English Speaking World in the 20th Century*, edited by John Canfield. London: Routledge, 1996.

Okasha, Samir. "Verificationism, Realism, and Scepticism." *Erkenntnis* 55 (2001): 371–385.

Quine, W. V. O. "Things and Their Place in Theories." In his *Theories and Things*. Cambridge, MA: Harvard University Press, 1981.

Quine, W. V. O. "Two Dogmas of Empiricism." *Philosophical Review* 60 (1951): 20–43. Reprinted in his *From a Logical Point of View: Nine Logico-Philosophical Essays*. Cambridge, MA: Harvard University Press, 1980.

Quine, W. V. O. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Rozeboom, W. "A Note on Carnap's Meaning Criterion." *Philosophical Studies* 11 (1960): 33–38.

Richard Creath (1996)

Bibliography updated by Benjamin Fiedor (2005)

VICO, GIAMBATTISTA

(1668–1744)

Born in Naples, Italy, in 1668, Giambattista Vico is best known for his critique of the Cartesian method and his philosophy of history. Beyond these areas, he is also known for contributions to linguistic theory, legal history, and cultural anthropology. Many have construed Vico as an eighteenth-century thinker who expressed the germ of ideas more fully developed in the nineteenth century. Thus, for example, Karl Löwith understands Vico's master work *The New Science* to anticipate "not only fundamental ideas of Herder and Hegel, Dilthey and Spengler, but also the more particular discoveries of Roman

history by Niebuhr and Mommsen, the theory of Homer by Wolf, the interpretation of mythology by Bachofen, the reconstruction of ancient life through etymology by Grimm, the historical understanding of laws by Savigny, of the ancient city and of feudalism by Fustel de Coulanges, and of the class struggles by Marx and Sorel" (1949, p. 115).

The familiar picture of Vico as the "great anticipator" contains some truth. More recent scholarship, in contrast, has tried to understand Vico as a thinker in his own right. The result has been a proliferation of different and often incompatible interpretations. These include views of Vico as a pioneer of contemporary hermeneutics; a creator of the modern social sciences; an architect of a uniquely Christian synthesis of philosophy and poetry; an advocate of a naturalistic Epicureanism thinly disguised as orthodox piety; a proponent of a Counter-Enlightenment approach to politics; and an author of a "genealogy of morals" that exposes the roots of modern secularism in pagan idolatry, divination, and sacrifice.

Rather than comment on rival interpretations of Vico, I here invite the reader to consider some aspects of what Vico himself regards as a continuous project of thought. This project begins with the works he published in 1709 and 1710 (*On the Study Methods of Our Time* and *On the Most Ancient Wisdom of the Italians*), runs through his jurisprudential writings from 1720 to 1722 (*Universal Right*), and concludes with the three major versions of *The New Science* (1725, 1730, 1744).

ANTI-CARTESIAN WRITINGS

In 1709 Vico published a version of the inaugural oration he delivered at the University of Naples in the preceding year, under the title *De nostri temporis studiorum ratione* (*On the Study Methods of Our Time*). In that work, which does not mention Descartes by name, Vico considered the art of "criticism" (*critica*), juxtaposing it with the art of "topics" (*topica*). Characteristic of criticism, in Vico's sense of the term, is a "dry and attenuated method of argumentation" that he associated with the Stoics and their then contemporary counterparts. Vico chided critics for wanting to purify, from even the suspicion of falsehood, their first truths, which they took to exist "above, outside of and beyond all images of bodies" (1990, Vol. 1, p. 104). His argument against criticism involves two main claims. The first claim is that to prioritize criticism in the education of children is unwise. Youths taught not to accept anything unless it can be certified by a rationalistic standard will have bad memories, impoverished imaginations, and a knack for rashly entering into "astonishing and unaccustomed ventures" (1990, Vol. 1, p. 104). The second claim is that criticism is poorly suited to discover truth. Because "the invention of arguments is prior by nature in the judgment of truth" (1990, Vol. 1, p. 106), criticism has no work to do unless the mind has investigated and brought to light the full range of relevant possibilities. The success of this prior investigation, Vico thought, depends upon the exercise of memory and imagination, especially in assisting the mind as it runs through the commonplaces. These mental capacities, Vico argued, are smothered by premature indoctrination in criticism, but can be developed through an immersion in topics.

In *On the Study Methods of Our Time* (1709/1988), Vico protests against what he regards as the domination of Cartesian criticism, but he does not oppose it as such. In *On the Most Ancient Wisdom of the Italians*, a work published a year later, 1710, he became more explicit in his opposition to Descartes. In that work, Vico charged Descartes with dogmatism, attributing to him the desire to consider all truths doubtful until metaphysically established by the principle "Cogito, ergo sum" ("I think; therefore I am"). Vico argued that, contrary to how he presented himself, Descartes is far from original. He noted that the use of the evil genius was anticipated by the Stoic in Cicero's *Academia* (45 BCE), and that the *cogito* principle was already enunciated by the slave Sosia in Plautus's *Amphitryo* (186 BCE). Vico does not claim that the *cogito* principle is false; he merely holds, "It is an ordinary cognition that happens to any unlearned person such as Sosia, not a rare and exquisite truth that requires such deep meditation by the greatest of philosophers to discover it" (1971, p. 73). The *cogito* principle is not only hackneyed, according to Vico; it is also unable to meet the skeptic's argument. For the *cogito* principle to provide knowledge of the nature of the mind, it would have to grasp the causes of thought (for Vico, as for Aristotle, knowledge is knowledge of causes). According to Vico, the *cogito* principle furnishes only consciousness (*conscientia*) of thinking, without illuminating its causes, and thereby fails to provide knowledge (*scientia*).

Like Francis Bacon before him and Immanuel Kant after him, Vico sought a middle path between dogmatism and skepticism. Against the skeptics, whom he represented as tracing absence of knowledge to a universal ignorance of causes, Vico pointed to domains in which we possess knowledge of the causes of things, because we originate them ourselves. His examples were synthetic geometry, painting, sculpture, ceramics, architecture—crafts in which skepticism has no application, unlike

those of rhetoric, politics, and medicine, which are "conjectural" arts in that they do not teach the forms by which their subject matter is created. Vico formulated a second argument, more theological in character, against the skeptics. Although skeptics properly observe that we do not know the causes of things that we are merely acquainted with (here we have consciousness or awareness, but not knowledge), it does not follow that these things lack causes. The pertinent question, according to Vico, is not "Do they have causes?" but "Where are the causes located?" If the causes are truly unknown, as the skeptic argues, they cannot be within us. But they must exist somewhere, in some locus or receptacle outside the self. This locus Vico named the "comprehension of causes, in which is contained all genera, or all forms, through which all effects are given" (1971, p. 75). Since this "comprehension" is infinite and necessarily prior to finite body, it is nothing other than God, "and indeed the God whom we Christians profess" (1971, p. 75).

In place of the *cogito* principle, Vico proposed his own version of a first truth, crystallized in his principle "Verum et factum convertuntur" ("The true and the made are convertible") (1971, p. 63). Although Vico claimed to derive the *verum-factum* principle philologically, he also understood it to be the core of a new anti-Cartesian epistemology and metaphysics. The core of the new metaphysics was that to know something is to make it, where making is collecting or gathering elements into a whole. Strictly speaking, only God conforms to the *verum-factum* principle, because he uniquely contains "the elements of things, extrinsic and intrinsic alike" (1971, p. 63). Because God makes elements and contains them within himself, he can arrange them perfectly, with utter precision and control. God's understanding of the elements of things is self-knowledge. Human beings, by contrast, do not possess such understanding of the elements. Since the human mind does not contain the elements of things within itself, it thinks about them through representations, at one remove, as it were. "Thought [*cogitatio*] is therefore proper to the human mind, but understanding [*intelligentia*] proper to the divine mind" (1971, p. 63). Human thinking, Vico concluded, should be understood as "participation in reason" (1971, p. 63). Thus, in contrast to the dogmatists, who exalt human truth, Vico downgraded it. Unlike the skeptics, however, he did not intend to deny its claims altogether: "Humanity is neither nothing, nor everything" (1971, p. 81).

A final dimension of Vico's early polemic is what might be called his "genealogical" critique of Descartes. In the second of two responses to Cartesian critics, Vico suggested that Descartes maliciously neglected ancient philosophers to promote his own doctrines. He was even so bold as to suggest that Descartes was an intellectual tyrant: "Descartes has done what those who have become tyrants have always been wont to do. They came to power proclaiming the cause of freedom. But once they are assured of power, they become worse tyrants than their original oppressors" (1971, p. 167). Vico unmasked Descartes's appeal to the natural light of reason as an excuse to avoid the labor of erudition and to avoid reading texts in the original languages. Vico also indicted Descartes for concealing the nature of his sources. In wanting his readers to believe that he had no significant predecessors or important teachers, Descartes "gathers the fruit of that plan of wicked politics, to destroy completely those men through whom one has reached the peak of power" (1971, p. 167). Descartes's Machiavellian cunning inspired him to lie about his origins: "Although he can dissimulate the fact with the greatest art in what he says, he was versatile in every sort of philosophy" (1971, p. 167). As an alternative to what he regarded as the uncandid fable of Descartes's *Discourse on the Method*, Vico proposed his own *Autobiography* where he sought to "narrate plainly and step by step the entire series of Vico's studies with the candor proper to a historian" (1990, Vol. 1, p. 7).

THE TURN TO HISTORY

In 1716, Vico began producing philosophical history, composing (though hampered by a severe cramp in his left arm) *The Life of Antonio Carafa* (which only appears in the eight-volume collection of Vico's work published by Laterza called *Opere di G. B. Vico*). At that time he discovered *On the Law of War and Peace*, by the Dutch jurist Hugo Grotius (1583–1645). Impressed with Grotius's work, Vico made him the last of his "four authors." The first three authors whom Vico privileged in his *Autobiography* were Plato, Cornelius Tacitus (c. 56–c. 120), and Francis Bacon. Vico associates Plato with "universal knowledge" that contemplates "man as he ought to be" (1990, Vol. 1, p. 29). The Roman historian Tacitus, by contrast, offered "counsels of utility" pertaining to "man as he is" (1990, Vol. 1, p. 29). Uniting Platonic "esoteric wisdom" and Tacitean "vulgar wisdom" is Bacon, "at one and the same time a universal man in theory and in practice" (1990, Vol. 1, p. 30). Despite his ambition, Bacon failed intellectually to encompass "the universe of cities and the course of all times, or the extent of all nations" (1990, Vol. 1, p. 44). Grotius, however, "embraces in a system of uni-

versal law the whole of philosophy and philology" (1990, Vol. 1, p. 44). Vico described his own ambition in similar terms. He sought to reconcile "the best philosophy, that of Plato made subordinate to the Christian religion," with a type of philology that "contains within itself the history of languages and the history of things " (1990, Vol. 1, p. 45).

To bring this reconciliation about, Vico began researching the history of Roman law after reading and annotating Grotius. The first fruit of this inquiry was several volumes collected under the title of *Diritto Universale* (*Universal Right*; 1720–1722/2000). Vico's occasion for writing this work was his desire to demonstrate his qualifications for a chair in law at the University of Naples paying six times as much as his position in rhetoric, which he would hold for most of his life. The intellectual wellspring for the work was Vico's desire to address the question whether justice is natural or merely conventional. Vico reduced contemporary answers to this question to two positions. First, there was the stance that he associated with "the skeptics," a category that included Epicurus, Niccolò Machiavelli, Thomas Hobbes, Benedict de Spinoza, and Pierre Bayle. Their common argument is that justice is not natural, but rooted in fear, chance, or necessity. Second, Vico considered the possibility that justice is grounded in the social nature of humans as a necessary condition for maintaining social order. This was the strategy of Grotius, who claimed to treat the rational basis of law in a quasi-mathematical manner, abstracting from particulars. Vico faults Grotius for excessive abstraction. Rather than bring his profound philological learning to bear in his attempt to counter the reduction of justice to expediency, Grotius depended on abstract and rationalistic arguments that are not persuasive against the skeptics. The positive aim of the *Universal Right* is to replace Grotius's system with a new conception that places particular facts and universal truths in a more illuminating relationship.

This attempt required Vico to turn his attention to the history of legal concepts, particularly the law of nations. Against Grotius's tendency to treat the law of nations (*ius gentium*) and natural law (*ius naturale*) as if they were not only distinct but also separate and autonomous, Vico attempted to exhibit natural law as present within the law of nations, which in time becomes civil law (*ius civile*). This attempt required Vico to argue that natural law has a dual origin: a metaphysical origin in eternal truth and a historical origin in the customs of human society. These dual sources can ultimately be traced to a single origin, God, whom the work identified as the "one principle and one end of universal law" (1974, p. 341). Vico ordered the volumes of the *Universal Right* according to a tripartite scheme intended to reflect the "origin" of divine and human things, their "cycle" (progress and return), and their "constancy."

Vico began the *Universal Right* with a brief consideration of trinitarian theology, followed by an exploration of the virtue possible for fallen humanity. In terms reminiscent of Augustine, Vico made the following identifications: "The force of truth [*vis veri*], or human reason is virtue insofar as it fights self-love [*cupiditas*]; the same virtue is justice insofar as it directs and equalizes utilities" (1974, p. 57). To support his antiskeptical contention that "right is in nature," Vico argued that humans are naturally social, despite their love of self. Although humanity is fallen, it possesses certain "affections" that manifest themselves in facial expressions, which are the beginnings of "expressive language" (1974, p. 59). To recognize distress in the face of another and to acknowledge this pain are natural to humans: "Man differs from animate brutes not only by reason and language, but also by his countenance" (1974, p. 59). From such commiseration in humankind, Vico infers that prior to any calculation of self-interest, "man will bring help to men" (1974, p. 59). Hence, society is natural to human beings and is made possible by sharing advantages.

Here one can perceive how historical consciousness enters into Vico's thinking about justice. The question "Does right exist in nature?" becomes a question about the social nature of humankind, which in turn Vico resolves into a historical inquiry about human nature in the primal state. To anchor in history his conviction that justice is natural, and thereby remedy what he regards as the chief failing of Grotius's natural law, Vico is driven to a philosophical and philological investigation of human origins.

How can Vico reconcile the claim that our concept of justice is, in some sense, subject to historical development, with an affirmation of its eternity and immutability? Vico addresses this question in the chapter of the *Universal Right* with the long title "Utility [*utilitas*] Is the Occasion, Nobility [*honestas*] Is the Cause, of Right [*ius*] and Human Society" (1974, p. 61). Historical occasions are not the cause or sufficient reason of the idea of justice, because "flux cannot generate the eternal, as bodies cannot generate anything above body" (1974, p. 61). Hence justice cannot be reduced to what promotes the advantage or interest of particular individuals; neither the first nor final cause of justice is utility. Yet occasions when issues of advantage and interest arise arouse the "will to

justice." Through the pursuit of their own advantage, "men, naturally social and divided, weak and needy from original sin, are brought to cultivate society, that is, to celebrate their social nature" (1974, p. 61). Vico concludes, "As the body is not the cause but the occasion by which the idea of truth is aroused in the mind of men, so utility of the body is not the cause but the occasion by which the will to justice is aroused in the soul" (1974, p. 61).

Vico's use of Nicolas de Malebranche's distinction between cause and occasion protects him from reducing justice to the merely conventional. It does so, however, by elevating instances that would strike some as mere historical accident to the rank of the philosophically significant "occasions" on which human knowledge of justice depends. If Vico is to make this high valuation of occasion and custom plausible, he must construct a historical narrative that depicts how equity (*aequum bonum*) expanded over time, and yet maintain the eternity of the concept. Vico attempted this task in the long section of the first part of the *Universal Right*, which purports to describe the cycle of universal right. To provide additional confirmation of his findings, both philosophical and philological, he added a second volume to the work, titled *De constantia jurisprudentis* (On the Constancy of the Jurisprudent). The first chapter of this work begins with the declaration "a new science is attempted" (*nova scientia tentatur*), and marks the transition to the final phase of his thought, contained in *The New Science*.

VICO'S NEW SCIENCE

The composition of the *Universal Right* established Vico as an erudite scholar, but it did not win him the law chair that he sought. Deciding to compose in the language of his countrymen, rather than that of the university, Vico wrote, in 1725, the first part of his autobiography and a first draft of *The New Science*. Now lost, this draft assumed the form of a negative critique of the "improbabilities, absurdities, and impossibilities that his predecessors had rather imagined than thought out" (1990, Vol. 1, p. 54). Because Vico could not afford to print the work as it stood, he decided to rewrite it using a "positive method that would be more concise and thus more efficacious" (1990, Vol. 1, p. 54). The result of this effort is the first version of *The New Science* (1725/1984). Its full title indicates the continuity with his previous work: *Principles of a New Science of the Nature of Nations, from Which Are Derived New Principles of the Natural Law of Peoples*.

In the subsequent versions of *The New Science* (1730, 1744), Vico placed less emphasis on the specifically political problematic. His larger aim was to achieve a new understanding of the origins of human culture. Vico thought that prior attempts to achieve this goal were vitiated by methodological errors characteristic of both philosophers and philologists. Philosophers, Vico argued, confuse their own refined natures with that of the first humans, who were necessarily simple and crude. They project their own "esoteric wisdom" and mental habits onto the primitive mind, which is not capable of advanced conceptual thinking. This projection is rooted in the "conceit of scholars," the habit of supposing that what contemporary thinkers know "is as old as the world" (*The New Science*, para. 127). Yet philologists (poets, historians, orators, grammarians) are no more helpful for understanding human origins, according to Vico. This is not only because they lack access to relevant data, but also because they are susceptible to the "conceit of nations"—the prejudice that "before all other nations, [one's own nation] invented the comforts of human life and that its remembered history goes back to the very beginning of the world" (*The New Science*, para. 125). Against the background of this twin failure, Vico concluded, "We must reckon as if there were no books in the world" (*The New Science*, para. 330).

Vico's attempt to transcend philosophy and philology assumed the form of a system that aspires to contain the virtues and avoid the vices of each. In its final exposition in 1744, the system began with a chronological table that outlines "the world of the ancient nations," followed by an enumeration of 114 "axioms" that purport to organize the material of the chronological table into a coherent whole. Against the inclination to despair that any recovery of remote human origins is possible, Vico proposed "the eternal and never failing light of a truth beyond all question: that the world of civil society has certainly been made by men, and that its principles are therefore to be found within the modifications of our own human mind" (*The New Science*, para. 331). Vico was pessimistic about the ultimate intelligibility of the world of nature, "which since God made it, He alone knows" (*The New Science*, para. 330). The civil world, however, is eminently knowable: "Since men made it, men could come to know it" (*The New Science*, para. 331). Here Vico reformulated the *verum-factum* principle that he articulated in the *Ancient Wisdom* of 1710. From the *verum-factum* principle, Vico went on to identify three "universal and eternal principles (such as every science must have) on which all nations were founded and still preserve themselves" (*The New Science*, para. 332). These are religion, marriage, and burial. The core of *The New Science* is the attempt to read human culture as the exhibition of these principles in a variety of guises, mutually ordered

by what Vico called a "divine legislative mind" and, more simply, "Providence" (*The New Science*, para. 133).

Vico's emphasis on Providence is appropriate, because it is the first and principal "aspect" of the final version of *The New Science*. Vico lists seven aspects of his total conception: (1) "a rational civil theology of divine providence," (2) a "philosophy of authority," (3) a "history of human ideas," (4) "a philosophical criticism that grows out of the history of ideas," (5) "an ideal eternal history traversed in time by the histories of all nations," (6) "a system of the natural law of the peoples," (7) "principles of universal history" (*The New Science*, paras. 385–399).

The New Science is known both for its method of investigation and its substantive conclusions. Regarding method, Vico proclaimed his desire to begin where his subject matter begins, with the assumption that the nature (*natura*) of civil phenomena is intelligible only through their birth (*nascimento*). If there are several possible ways of conceiving the history of an idea or institution, Vico argued that we should focus on the possibility whose manner is most orderly and conducive to the preservation of the human race. Such an "order of things cannot be approached directly, but must be sought through the "order of ideas" and "order of language." As a preliminary to accomplishing the goal of the new science, to disclose the necessary substructure of the civil world, Vico asked the reader whether he can imagine more, fewer, or different causes than the ones he finds. Near the end of the section "Method" of Book 1, Vico declared that his aim was to clean, piece together, and restore "the great fragments of antiquity, hitherto useless to science because they lay begrimed, broken, and scattered" (*The New Science*, para. 357). The light shed by excavation and reconstruction would enable him, Vico thought, to trace "all the effects narrated by certain history" to their originating institutions, "as to their necessary causes" (*The New Science*, para. 358). Not all readers have found persuasive Vico's claim to strict logical necessity. Rather than defend the claim, many contemporary interpreters have advanced the weaker argument that a Viconian perspective is able to render intelligible aspects of the civil world (especially myth, custom, law, poetry) that would otherwise remain obscure.

The content of Vico's new science resists summary description. Its basic scheme is the division of human history into three periods: the age of gods, the age of heroes, and the age of humankind. In the age of gods, "every gentile nation had its Jove" (*The New Science*, para. 193). In every pagan culture, the sky came to be identified as a god who speaks in the language of lightning and thunder. "Jove" was the work of the "theological poets," who created the "first divine fable" and believed it themselves. The practical effect of Jove was to settle the wandering first humans and to set up a system of primitive religion based on divination and sacrifice. Vico's attitude toward primitive religion was complex. The fables created (or "feigned") by the theological poets were based on a "credible impossibility: it is impossible that bodies should be minds, yet it was believed that the thundering sky was Jove" (*The New Science*, para. 383). Yet Vico's attitude toward pagan religion is not one of enlightened condescension. "Through the thick clouds of those first tempests, intermittently lit by those flashes, they made out this great truth: that divine providence watches over the welfare of all mankind" (*The New Science*, para. 385). Thus ran Vico's partial defense of the primitive mind: It apprehended a truth, even if in distorted fashion, that later philosophers (especially the Epicureans and their then contemporary counterparts) altogether missed.

In the age of gods, primitive humans are incapable of proper political organization. There were no cities, only families governed by the "cyclopean paternal authority" of the fathers. The heroic age began with the founding of the cities, prompted by the need of family fathers to unite for the sake of self-defense against their increasingly resentful slaves (the "*famuli*"). Nominating one of their number as king, the fathers generated "severe aristocratic commonwealths" (*The New Science*, para. 663). Vico's narrative of the genesis of heroic commonwealths from the "state of the families" was a polemic directed against Hobbes and "the three princes of natural law," whom he identified as Grotius, the English jurist John Selden (1584–1654), and the German natural-law philosopher Samuel von Pufendorf. Based on neither contract nor self-interest, heroic commonwealths were essentially religious in character. Viewing themselves as descendants of the gods, the heroes secure their dominance through myths that define the plebeians as less than fully human (because they were not of divine descent), and thereby exclude them from citizenship. Toward heroic civil institutions as well, Vico's attitude was complex. On the one hand, he appreciated the gravity and reverence characteristic of aristocratic virtue, especially as expressed in Roman jurisprudence. On the other hand, he sympathized with the plebeians and their struggle for liberty and equality. As with the age of gods, determining Vico's judgment about the merits of the heroic age is a difficult matter of interpretation.

What prompted the transition from the heroic to the human age was the increase in self-knowledge on the part of the plebeians, as encoded in the poetic character of the Athenian lawgiver Solon (c. 630–c. 560 BCE). Once they came to fully recognize their equal humanity, the plebeians began to demand participation in civil society. At this point human nature became "benign," as exemplified by the Roman general Scipio Africanus (236–184 or 183 BCE), the Athenian statesman Aristides the Just (c. 530–c. 468 BCE), and Socrates. The form of government changed from aristocratic to democratic, issuing in "free popular commonwealths." Initially, this appeared to be progress. Philosophy (enabled by the trope of irony) came onto the scene, leading to a purification of the "vulgar wisdom" that developed in the divine and heroic ages. But the "political philosophy" of Plato and Aristotle, of which Vico approved, gave way to "monastic or solitary philosophy," as represented by the Stoics and the Epicureans. "As the popular states become corrupt, so also did the philosophies. They descended to skepticism. Learned fools fell to calumniating the truth" (*The New Science*, para. 1102). In the first phase of the human age, humans were "benign," but their quest for pleasure and luxury led them to become "delicate" and finally "dissolute" (*The New Science*, para. 242). Under the influence of radically antitraditional philosophy that sets itself against "common sense," the citizens, growing ever more atomistic, eventually become "aliens in their own nations" (*The New Science*, para. 1008). Vico indicated three remedies to the problem of social fragmentation: monarchy, conquest by more unified nations, and destruction followed by a return to the age of gods.

Vico's philosophy of decline appears inextricably linked to the decline of philosophy. According to one twentieth-century student of Vico, the last phase of the age of men is a condition where "thought still rules, but a thought which has exhausted its creative power and only constructs meaningless networks of artificial and pedantic distinctions" (Collingwood 1946, p. 67). This is the condition of "beasts made more inhuman by the barbarism of reflection than the first men had been made by the barbarism of sense" (*The New Science*, para. 1006). Yet along with the fatalistic strain of Vico's view of history, one must consider his evident belief in the power of his new science to inspire a rapprochement between philology and philosophy, tradition and reason, politicians and academics. Is such an equilibrium possible? If so, what form would it take? For both students of Vico and social philosophers, these questions remain.

See also Aristotle; Bachofen, Johann Jakob; Bacon, Francis; Bayle, Pierre; Cartesianism; Cicero, Marcus Tullius; Dilthey, Wilhelm; Epicureanism and the Epicurean School; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Hobbes, Thomas; Homer; Kant, Immanuel; Machiavelli, Niccolò; Malebranche, Nicolas; Marx, Karl; Myth; Niebuhr, Reinhold; Philosophy of History; Philosophy of Language; Plato; Pufendorf, Samuel von; Savigny, Friedrich Karl von; Sociology of Knowledge; Socrates; Sorel, Georges; Spengler, Oswald; Stoicism.

Bibliography

WORKS BY VICO

On the Study Methods of Our Time (1709). Translated by Elio Gianturco. Ithaca, NY: Cornell University Press, 1990.

On the Most Ancient Wisdom of the Italians (1710). Translated by L. M. Palmer. Ithaca, NY: Cornell University Press, 1988.

Universal Right (1720–1722). Translated by Giorgio Pinton and Margaret Diehl. Amsterdam: Rodopi, 2000.

The New Science of Giambattista Vico (1725–1744). Translated by Thomas G. Bergin and Max H. Fisch. Ithaca, NY: Cornell University Press, 1984.

The Autobiography of Giambattista Vico (1725–1731) Translated by Max Harold Fisch and Thomas Goddard Bergin. Ithaca, NY: Cornell University Press, 1944.

Opere di G. B. Vico. 8 vols. Edited by Fausto Nicolini. Bari, Italy: Laterza, 1911–1941.

Opere filosofiche. Edited by Paolo Cristofolini. Florence, Italy: Sansoni, 1971.

Opere giuridiche. Edited by Paolo Cristofolini. Florence, Italy: Sansoni, 1974.

Opere. 2 vols. Edited by Andrea Battistini. Milan, Italy: Mondadori, 1990.

WORKS ON VICO

Adams, H. P. *The Life and Writings of Giambattista Vico*. London: George Allen and Unwin, 1935.

Amerio, Franco. *Introduzione allo studio di G. B. Vico*. Turin, Italy: Societa Editrice Internazionale, 1947.

Bedani, Gino. *Vico Revisted: Orthodoxy, Naturalism, and Science in the "Scienza Nuova."* New York: St. Martin's Press, 1989.

Berlin, Isaiah. *Vico and Herder*. London: Hogarth Press, 1976.

Botturi, Francesco. *La sapienza della storia: Giambattista Vico e la filosofia pratica*. Milan, Italy: Università Cattolica del Sacro Cuore, 1991.

Collingwood, Robin. *The Idea of History*. Oxford: Clarendon Press, 1946.

Croce, Benedetto. *The Philosophy of Giambattista Vico*. Translated by R. G. Collingwood. New York: Russell and Russell, 1964.

Hösle, Vittorio. *Introduzione a Vico: La scienza del mondo intersoggettivo*. Milan, Italy: Guerni e Associati, 1997.

Lachterman, David. "Mathematics and Nominalism in Vico's *Liber Metaphysicus.*" In *Sachkommentar xu G. Vico's Liber*

Metaphysicus, edited by Stephan Otto and Helmut Viechtbauer, 47–85. Munich: Wilhelm Fink Verlag, 1985.

Lilla, Mark. *G. B. Vico: The Making of an Anti-modern*. Cambridge, MA: Harvard University Press, 1993.

Löwith, Karl. *Meaning in History*. Chicago: University of Chicago Press, 1949.

Mazzotta, Giuseppe. *The New Map of the World: The Poetic Philosophy of Giambattista Vico*. Princeton, NJ: Princeton University Press, 1999.

Milbank, John. *The Religious Dimension in the Thought of Giambattista Vico, 1668–1744*. Lewiston, NY: Edwin Mellen Press, 1991.

Miner, Robert. *Vico, Genealogist of Modernity*. Notre Dame, IN: University of Notre Dame Press, 2002.

Mooney, Michael. *Vico in the Tradition of Rhetoric*. Princeton, NJ: Princeton University Press, 1985.

Pompa, Leon. *Vico: A Study of the "New Science."* 2nd. ed. Cambridge, U.K.: Cambridge University Press, 1985.

Schaeffer, John. *Sensus Communis: Vico, Rhetoric, and the Limits of Relativism*. Durham, NC: Duke University Press, 1990.

Verene, Donald Philip. *Vico's Science of Imagination*. Ithaca, NY: Cornell University Press, 1981.

Robert C. Miner (2005)

VIENNA CIRCLE, THE

See *Logical Positivism*

VIOLENCE

"Violence" is derived from the Latin *violentia,* "vehemence," which itself comes from *vis* (force) + *latus* (to carry) and means, literally, intense force. Violence shares its etymology with violate, "injure." *Violence* is used to refer to swift, extreme force (e.g., a violent storm) and to forceful injurious violation (e.g., rape, terrorism, war).

Violence has received some philosophical consideration since ancient times, but only since the twentieth century has the concept of violence itself been of particular concern to philosophers. Perhaps this is due to the exponential growth in the efficiency of and access to the means of violence in the modern era, to the unprecedented carnage the twentieth century saw, or to the emergence of champions of nonviolence such as Mohandas Gandhi and Martin Luther King Jr. Beyond clarifying the concept of violence, philosophical argument has turned to the moral and cultural justifiability of violence to achieve personal, social, or political ends.

Philosophers do not achieve consensus about the concept. Often, violence is taken to consist in overt physical manifestations of force. These may be on the scale of individuals (e.g., mugging) or of nations (e.g., war). In its primary use *violence* refers to swift, extreme physical force typically involving injury and violation to persons or property. There is increasing philosophical interest in a wider use of the term extending beyond the overtly physical to covert, psychological, and institutional violence. In this broader sense racism, sexism, economic exploitation, and ethnic and religious persecution all are possible examples of violence; that is, all involve constraints that injure and violate persons, even if not always physically.

Concerning the moral and political justifiability of using violence to achieve personal or social ends, again philosophers disagree. Some have taken violence to be inherently wrong (e.g., murder), while most have taken it to be an open question whether violence is normatively justifiable. Terrorism presents a special case. It is aimed at randomly selected innocent victims in an effort to create general fear, thus sharpening focus on the terrorists' cause or demands. This random targeting of innocents accounts for the near universal moral condemnation of terrorism, despite the dominant view that violence in general is not inherently wrong.

Arguments purporting to justify violence do not value it in itself but as a means to an end sufficiently good to outweigh the evils of the injury or violation involved. Often, such justifiable violence is seen as a necessary means to important ends; that is, the good achieved by justifiable violence could not be achieved without it. Arguments challenging the justifiability of violence tend to reject the claim to necessity, arguing for nonviolent means, or to deny the claim that violation and injury are outweighed by the ends achieved. Such arguments may be against violence per se or merely against particular violent acts.

Georges Sorel's *Reflections on Violence* (1908) is the earliest extensive philosophical work devoted to the subject. While Karl Marx saw a role for violence in history, it was secondary to the contradictions inherent in collapsing systems. Sorel synthesizes Marx's proletarianism, Pierre-Joseph Proudhon's anarchism and Henri Bergson's voluntarism, defending revolutionary trade unionism in its efforts to destroy the existing institutional order. Sorel advocates the violent general strike as the means of class warfare against the state and owners of industry.

In *On Violence* (1970) Hannah Arendt reviews the twentieth-century apologists for violence in an effort to explain the increasing advocacy of violence, especially by the new left. She questions Mao Zedong's "Power grows out of the barrel of a gun" and articulates the position that

power and violence are opposites. For Arendt the extreme of violence is one against all while the extreme of power is all against one. Power is acting in concert with others while violence is acting with implements against others. Loss of power leads some to try to replace it with violence. But violence is the opposite of power and cannot stand in its stead. Arendt concedes that violence can be justified but insists that it is only in defense against clear, present, immediate threats to life where the violence does not exceed necessity and its good ends are likely and near.

Newton Garver's "What Violence Is" (1975) extends the discussion to covert, psychological, and institutional violence. According to Garver, "Any institution which systematically robs certain people of rightful options generally available to others does violence to those people" (p. 420). Despite his sympathy with nonviolence, Garver claims that it is not a viable social goal. Violence between nations may be reduced but not eliminated.

See also Anarchism; Arendt, Hannah; Bergson, Henri; King, Martin Luther; Marx, Karl; Pacifism; Proudhon, Pierre-Joseph; Racism; Sexism; Social and Political Philosophy; Sorel, Georges; Voluntarism.

Bibliography

Arendt, H. *On Violence.* New York: Harcourt Brace, 1970.

Corlett, J. Angelo. *Terrorism: A Philosophical Analysis.* Dordrecht: Kluwer Academic, 2003.

Cotta, S. *Why Violence? A Philosophical Interpretation.* Gainesville, FL: University Presses of Florida, 1985.

Garver, N. "What Violence Is." In *Today's Moral Problems,* edited by R. Wasserstrom. New York: Macmillan, 1975.

Gray, G. *On Understanding Violence Philosophically.* New York: Harper and Row, 1970.

Holmes, R. L. "Violence and the Perspective of Morality." In *On War and Morality.* Princeton, NJ: Princeton University Press, 1989.

Honderich, Ted. *Violence for Equality: Inquiries in Political Philosophy.* Harmondsworth, U.K.: Penguin, 1980.

Reitan, Eric. "The Moral Justification of Violence: Epistemic Considerations." *Social Theory and Practice* 28 (3) (2002): 445–464.

Schaffer, J. A., ed. *Violence: Award Winning Essays in the Council for Philosophical Studies Competition.* New York: McKay, 1971.

Sorel, G. *Reflections on Violence* (1908). Translated by T. E. Hulme and J. Roth. Glencoe, IL: Free Press, 1950.

Weber, Samuel, and Hent de Vries, eds. *Violence, Identity, and Self-Determination.* Stanford, CA: Stanford University Press, 1997.

Wolff, R. P. "On Violence." *Journal of Philosophy* 66 (1969): 601–616.

Duane L. Cady (1996)
Bibliography updated by Philip Reed (2005)

VIRTUE AND VICE

Assuming that human agents possess settled dispositions or character traits, some of which are especially deemed worthy of praise while others deserve blame or reproach, moral philosophers have long treated the first sort under the category "virtue" and their opposites under the general term "vice." The *fin-de-siecle* revival of the virtue tradition in normative ethics as a third force, alongside Kantianism and consequentialism, has resulted in focused attention by theorists of all persuasions on the nature and proper role of virtues and vices in any comprehensive treatment of morality. Thus, two consequentialists (Driver 2001, Hurka 2001) have produced full-length treatments of the virtues, and there has been a growing appreciation of the key role of virtue in Immanuel Kant's ethics (Herman 1993, O'Neill 1996, Wood 1999). While the attention to virtue among Kantians and neo-Kantians is not too surprising, since much of Kant's later work was devoted to working out the important role that virtue and character play in morality (the weighty concluding section of the 1797 *Metaphysics of Morals* is rightly titled "The Doctrine of Virtue"), the consequentialist turn to virtue is, perhaps, more surprising. Jeremy Bentham, for example, gave a rather rude treatment of virtue in his *Deontology*, as recently described by Julia Annas (2002).

AN EMPIRICAL CHALLENGE TO TRAITS OF CHARACTER

This recent consequentialist vindication of virtue can involve a considerable departure from the paradigmatic picture of virtues and vices as traits of character, however. Tom Hurka (2001), for example, defines moral virtues and vices as responsive *attitudes* taken up toward intrinsic goods and evils, in explicit opposition to the view going back to Aristotle that treats them as stable dispositions or persisting states of persons. In this identification Hurka is acknowledging a controversy stemming from certain results in social psychology that some philosophers have taken to rule out on empirical grounds any robust conception of personality traits. Extreme situationists argue on the basis of considerable experimental evidence that the layperson's readiness to attribute to themselves and others robust character traits that are stable across situations, both over time and in various circumstances, and that can be used to predict behavior, is undermined by what has been termed "the power of the situation."

In experiments no longer permitted by twenty-first-century ethical guidelines, subjects were duped into administering what they were led to believe were severe electric shocks to their "victims" or invited to "role-play" as prison guards to such an extent that the subsequent sadistic behavior caused the researchers to abort the exercise. In addition, we have increasing evidence from developments at prisons in Iraq and other places around the world that average American young people, in stressful environments, can engage in dehumanizing practices that shock almost all of us. Gilbert Harman, considering both experimental and real-life examples of such catastrophic character failure, has forcefully pressed the negative implications he sees for the very foundations of virtue theory: "I myself think it is better to abandon all thought and talk of character and virtue. I believe that ordinary thinking in terms of character traits has had disastrous effects on people's understanding of each other. … I think we need to get people to stop doing this. We need to convince people to look at situational factors and to stop explaining things in terms of character traits. We need to abandon all talk of virtue and character, not find a way to save it by reinterpreting it" (1999/2000, p. 224).

Such a sweeping dismissal of all talk of character traits is, arguably, an overly simplified reading of the relevant personality studies (see Matthews, Deary, and Whiteman 2003 for a synthesis of the empirical evidence favoring interactionism, the view that behavior is a function of both personality differences and situational influences). Yet even the more balanced presentation of a similar skepticism in John Doris's 2002 study surely calls for critical appraisal by virtue theorists of any normative persuasion. Annas (2002), Swanton (2003), and other virtue ethicists have responded to the challenge. There is also room for more detailed treatments integrating social psychology, personality theory, and ethical theory, preferably by collaborating researchers with relevantly different research interests and, perhaps, in newly designed psychological experiments designed to test for cross-situational attribution of virtues and vices (see Cawley, Martin, and Johnson 2000).

The exploration of this basic challenge to virtue theory promises to carry on the pioneering work of Owen Flanagan, who first brought philosophers' attention to the situationist challenge and who championed what he labeled the "Principle of Minimal Psychological Realism": "Make sure when constructing a moral theory or projecting a moral ideal that the character, decision processing, and behavior prescribed are possible, or are perceived to be possible, for creatures like us" (1991, p. 32). This call for ethicists to take note of social-scientific findings dovetails nicely with recent philosophical calls for naturalist or science-friendly approaches to the philosophy of mind, epistemology, and metaphysics. The principle is best thought of as giving contemporary substance to the familiar principle that "ought" implies "can."

VIRTUE THEORY AS DISTINCT FROM VIRTUE ETHICS

A distinction should be drawn, then, between virtue theory taken quite generally and virtue ethics proper, where virtue theory covers any theoretical treatment of the nature of virtue and vice, even if their role in the theory is not central, and virtue ethics privileges them in some way or other. In Christine Swanton's self-consciously pluralistic conception (2003), virtue ethics, like consequentialism, should be seen as a broad genus encompassing various species. Thus, alongside the familiar neo-Aristotelian varieties of virtue ethics (Foot 2001, Hursthouse 1999), there is room for Michael Slote's "agent-based" account (1992), which opposes the neo-Aristotelian emphasis on the agent's happiness and well-being (*eudaimonia*) as grounding the goodness of virtue insofar as its presence helps the agent to flourish in a social context, in favor of the view that various inner traits and motives are admirable on their own. James Martineau thus joins Friedrich Nietzsche in the pluralist pantheon of virtue ethicists, alongside Thomas Aquinas and David Hume and their Greek and Roman forebears.

Any version of virtue ethics gives primacy of place to moral character over action, to the aretaic over the deontic, and sees the individual's development of virtues and elimination of vices as the best assurance that good deeds (right actions) will be forthcoming. Thus, for the virtue ethicist, the familiar bumper sticker's call for "*random* acts of kindness" seems incoherent as well as quixotic. If people cultivate the virtue of kindness, they can be reliably counted on to perform kind actions in a variety of circumstances, to adjust their reactions to others' needs consistently and appropriately, by expressing a suitable interpersonal sensitivity, rather than by following formulaic prescriptions or rules for conduct. An honest person, for example, will not only tell the truth when called upon to do so but will also not shade it or allow others to dissemble. The honest person will not resent just criticism, abide flattery, envy rogues and rascals alike, or engage in any number of sharp practices in business dealings.

Dishonest people, in contrast, will predictably exhibit the opposite sorts of behavioral tendencies. They will lie when convenient, cheat on their taxes, allow oth-

ers to think them more deserving than they truly are, overlook mistakes on restaurant checks that are in their favor, and so on. For both the virtuous and the vicious, then, character structures will be expressed in a variety of ways and across a variety of circumstances, although some core traits will remain at the center of the individual's personality.

COMPARING VIRTUE AND VICE

It may be thought that a certain asymmetry will be found when comparing virtue and vice, with the former, perhaps, more predictable in its natural expression than the latter. A coward, it may be thought, might not run from some dangers and might not fear a wide range of things. Perhaps the Falstaffian figure that comes to mind is just a stereotype, and real cowards are much more selective in avoiding danger, rhetorical war hawks avoiding the draft by enrolling in college, perhaps, but not avoiding the most intimidating teachers or toughest courses.

This impression might simply reflect the fact that virtue theorists say much more about positive traits and much less about negative ones. It is the virtues, after all, that the theorist is trying to inculcate; detailed descriptions of the vices are often left out or given short shrift. The theorist accentuates the positive, perhaps. Aristotle, in his general theory of the virtues as the means between vices on both sides, one of excess and the other of deficiency, had a great deal to say about the vices and saw them as having the same psychological structures in the soul as the virtues. For him, vices were equally "settled dispositions" (*hexeis*), results of the wrong sort of habituation as opposed to the right kind. In departing from Aristotle in this regard, owing to our relative disenchantment with his general theory of excellence (*aretē*) as a mean, we moderns may well have tended to downplay the phenomenology of vice.

Tom Hurka's categorization of the range of vices (2001), from the *pure* ones (e.g., malice, *Schadenfreude*, sadism) at one end of the spectrum, through those of *indifference* (e.g., callousness, sloth, smugness), to the mildest forms at the other end, which he calls vices of *disproportion* (e.g., foolhardiness, avarice, intemperance), is a welcome reminder of the richness of our moral vocabulary and of the basic symmetry to be found when comparing virtue and vice. They both come in various forms and degrees, and can be similarly graphed by intensity and the relative value of their respective objects and fields. One important vice, hypocrisy in all of its manifestations, is the subject of the 2004 book by Bela Szabados and Eldon Soifer, who treat it from Kantian, consequentialist, and virtue ethicist perspectives. The philosophical fortunes of vice are thus on the rise.

THE PROBLEM OF VAGUENESS IN APPEALS TO VIRTUE

Critics of virtue ethics as a serious competitor in normative ethical theory have found it wanting in its vague decision procedure for deciding difficult cases. Moreover, by comparison with consequentialism and deontology, virtue ethics has made few contributions to the field of applied ethics. As for the last charge, the scene is shifting a great deal, since it is common these days to have virtue ethics treated alongside its more familiar predecessors with equal billing, as it were, in textbooks. In the subfield of professional ethics, Justin Oakely and Dean Cocking (2001) have deployed the resources of virtue ethics, comparing them favorably with Kantian and utilitarian approaches. The idea of a good general practitioner, whether in law, medicine, or business, is ripe for development along the lines of virtue ethics. Oakley and Cocking address a number of difficult issues from this angle in the course of their book.

One chief worry is the seeming vagueness of the advice to follow the example of the ideally virtuous person, especially in displaying the exquisite sensitivity to concrete detail supposedly exhibited by the practically wise (*phronimos*), which moral particularists and antitheorists tend to highlight. John McDowell (1998) and Martha Nussbaum (1986), among others working within the Aristotelian framework, have stressed the advantages of thinking of moral choice as uncodifiable, as the product of particular judgments made on the spot by individuals who embody the relevant virtues and are thereby in a better position than others to rightly perceive and assess the immediate needs of the situation. A virtuous friend, for example, is in the best position to give painful yet necessary advice to an individual, at the right time, with the right affect, neither too forcefully nor unclearly phrased, with due allowance for the receptivity and ability of the other to listen and take it in at that time. Similarly, the temperate person hits the right target in choosing bodily pleasures, adjusting intake by giving due attention to the situation (e.g., a party or a wake) and its demands (e.g., the need to stay alert and focused versus an opportunity to relax).

Christine Swanton (2003) has developed this ancient target analogy so favored by Aristotle and the Stoics in compelling fashion. She defines a virtuous act as one that hits the target of the relevant virtue, and she stresses the vicissitudes and complexities of "moral archery." Imagine

that you are at a conference where you spot a stranger with some command of English who cannot (as you can) fully appreciate the sophisticated and scintillating philosophical discussion going on. You decide to devote your energies to the apparent needs of the stranger, leave the meeting room and make conversation, only to discover that this is more difficult than you imagined, definitely not enjoyable, and, the truth be told, perhaps not as helpful to the other as you had hoped. He could just as easily have spent time at the book exhibit while you stayed in the session, and you could have met him there in due time. The point is that while a kind person might have impressions calling for an expression of virtue, the exact specification of what is kind in the precise circumstances is not at all clear in advance or even *in situ*. Even the ideal moral archer may miss the target for reasons extremely hard to calculate in advance. Nonetheless, sensitivity to the particular environment is the distinct strength of the ideally virtuous agent.

Against this sort of appeal James Griffin has forcefully replied, citing the implausibility of "an ideally virtuous person, whose dispositions are in perfect balance and who therefore is better able to perceive situations correctly, including features that general principles often fail to capture. This is another piece of over-ambition in ethical theory" (1996, p. 115). While Griffin's complaint stems from his general pessimism about the ambitions of a normative theory to take us deeply into the solution of practical moral problems, virtue ethicists do have a special responsibility to be more precise than they have been.

Rosalind Hursthouse (1999) has been quite sensitive to this particular charge and has emphasized that the alleged imprecision of virtue ethics is in part an artifact of the fact that most ethicists are so familiar with, and not explicit about, the basic principles of the main normative theories on offer. Consider the following principles (one for virtue ethics, one for consequentialism, and one for deontology):

> (VEP) An action is right if and only if (iff) it is what a virtuous agent, acting in character, would do in the circumstances.
>
> (CP) An action is right iff it promotes the best consequences.
>
> (DP) An action is right iff it accords with a correct moral rule or principle.

Since ways of filling out the consequentialist and deontological proposals come so readily to mind, we can immediately think of various ways to give more substance and specificity to (CP) and (DP). For example, in the consequentialist case we envisage utilitarian attention to quantity and quality of pleasure, satisfaction of preferences, or maximization of happiness These criteria are applied to acts themselves or to rules for choosing acts as in versions of rule utilitarianism. In the deontological case, we think of moral rules and principles, such as being commanded by God or in accord with natural law, licensed by the categorical imperative, responsive to the formula of humanity, chosen by free agents in an ideal initial bargaining position, etc.

Because ethicists since the enlightenment have been unaccustomed to filling in the details of any virtue theory, (VEP) can seem hopelessly vague to those whose historical perspective begins more or less with Kant. Hursthouse argues that when the most basic principles are staked out as starkly and simply as above, (VEP) has as much clearly marked precision as (DP) and (CP). As we become accustomed to the workings of the moral imagination of those at home with the virtues, we will find it easier to fill in (VEP) with alternative specifications, compare the advantages of each, and weigh and balance the strengths and weaknesses of a variety of historical and contemporary proposals of virtue theorists. Perhaps it will also be easier to see how society at large harbors and encourages various vices and character defects in our social, political, and personal lives. Surely, greed and ruthlessness in business and carelessness of citizens in rich nations lead people to ignore the needs of the planet and its less fortunate inhabitants, and hence lead to poverty and environmental degradation.

One attractive feature of a virtue-theoretical approach to morality is the fact that most communities around the world, however different they are in culture and religion and a myriad other ways, tend to organize their early moral education of children around the promotion of virtue and the avoidance of vice. It may well be that, in trying to reach across cultural divides to find a common moral vocabulary with which to address the pressing moral issues of global reach, we would do well to supplement the categories so familiar since the Enlightenment in the West (e.g., duty, utility, costs versus benefits) with the highly nuanced and richly textured vocabulary of virtue and vice.

See also Evil; Moral Psychology; Virtue Ethics.

Bibliography

Annas, Julia. "Should Virtue Make You Happy?" *Apeiron* 35 (4) (2002): 1–19.

Cawley, Michael J., James E. Martin, and John A. Johnson. "A Virtues Approach to Personality." *Personality and Individual Differences* 28 (2000): 997–1013.

Crisp, Roger, ed. *How Should One Live? Essays on the Virtues*. Oxford, U.K.: Clarendon Press, 1996.

Doris, John. *Lack of Character: Personality and Moral Behavior*. New York: Cambridge University Press, 2002.

Driver, Julia. *Uneasy Virtue*. New York: Cambridge University Press, 2001.

Flanagan, Owen. *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press, 1991.

Foot, Philippa. *Natural Goodness*. New York: Oxford University Press, 2001.

Griffin, James. *Value Judgement: Improving Our Ethical Beliefs*. Oxford, U.K.: Clarendon Press, 1996.

Harman, Gilbert. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99 (1999): 315–331.

Harman, Gilbert. "The Nonexistence of Character Traits." *Proceedings of the Aristotelian Society* 100 (1999/2000): 223–226.

Herman, Barbara. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Hooker, Brad, and Margaret Little, eds. *Moral Particularism*. Oxford, U.K.: Clarendon Press, 2000.

Hurka, Thomas. *Virtue, Vice, and Value*. New York: Oxford University Press, 2001.

Hursthouse, Rosalind. *On Virtue Ethics*. New York: Oxford University Press, 1999.

McDowell, John. *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, 1998

Matthews, Gerald, Ian J. Deary, and Martha Whiteman. *Personality Traits*. 2nd ed. New York: Cambridge University Press, 2003.

Nussbaum, Martha. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. New York: Cambridge University Press, 1986.

Oakley, Justin, and Dean Cocking. *Virtue Ethics and Professional Roles*. New York: Cambridge University Press, 2001.

O'Neill, Onora. "Kant's Virtues." In *How Should One Live? Essays on the Virtues*, edited by Roger Crisp. Oxford, U.K.: Clarendon Press, 1996.

Sherman, Nancy. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. New York: Cambridge University Press, 1997.

Slote, Michael. *From Morality to Virtue*. New York: Oxford University Press, 1992.

Swanton, Christine. *Virtue Ethics: A Pluralistic View*. New York: Oxford University Press, 2003.

Szabados, Bela, and Eldon Soifer. *Hypocrisy: Ethical Investigations*. Peterborough, ON: Broadview Press, 2004.

Wood, Allen W. *Kant's Ethical Thought*. New York: Cambridge University Press, 1999.

Lawrence J. Jost (2005)

VIRTUE EPISTEMOLOGY

"Virtue epistemology" has a narrow and a broad sense. In the narrow sense, the central claim of virtue epistemology is that, perhaps with some minor qualifications aside, knowledge is true belief resulting from intellectual virtue. On this view, the intellectual virtues are stable dispositions for arriving at true beliefs and avoiding false beliefs. Put another way, the intellectual virtues are reliable dispositions: either reliable powers, such as accurate perception and sound reasoning, or reliable character traits, such as intellectual honesty and intellectual carefulness.

In the broad sense, virtue epistemology is the position that the intellectual virtues are the appropriate focus of epistemological inquiry, whether or not knowledge can be defined in terms of such virtues, and whether or not such virtues can be understood as dispositions toward true belief. In this broad sense, the intellectual virtues continue to be understood as excellences of cognitive agents, but it is left open whether such excellences make the agent reliable, and whether the agent's being reliable is even relevant in the most important kinds of epistemic evaluations.

A number of claims have been made on behalf of virtue epistemology. As noted, virtue epistemologists claim that the resources of virtue theory can help to explicate a range of important kinds of epistemic evaluation. They have also claimed that virtue epistemology can provide an adequate response to skepticism, that it can solve Gettier problems, that it can contribute to a unified theory of value across epistemology and ethics, and that it can overcome the debates between internalism and externalism and between foundationalism and coherentism.

One issue that has been much discussed in the literature concerns the nature of the intellectual virtues. More specifically, it concerns the relationship between the intellectual virtues and the moral virtues. On one side of this debate are those who think that the intellectual virtues are much like the moral virtues. On this view, the intellectual virtues are such character traits as intellectual courage, intellectual honesty, and intellectual carefulness. For example, Linda Zagzebski (1996) takes Aristotle's account of the moral virtues as her model for the intellectual virtues, arguing that Aristotle was mistaken to insist on a strong distinction here. Other virtue epistemologists, such as Ernest Sosa, follow Aristotle in thinking of the intellectual virtues as reliable powers or abilities. Thus Aristotle took intuition into first principles and demonstrative reason to be paradigmatic intellectual

virtues. Updating Aristotle's list of the virtues, Sosa considers reliable perception and various sorts of sound inductive reasoning too to be paradigmatic epistemic virtues.

Despite these differences among virtue epistemologists, there are points in common as well. For one, all virtue epistemologists begin with the assumption that epistemology is a normative discipline. The main idea of virtue epistemology is to understand the kind of normativity involved in a virtue-theoretic model of knowledge. This idea is best understood in terms of a thesis about the direction of analysis. Just as virtue theories in ethics try to understand the normative properties of actions in terms of the normative properties of moral agents, so virtue epistemology tries to understand the normative properties of beliefs in terms of the normative properties of cognitive agents. Hence virtue theories in epistemology have been described as person-based rather than belief-based, just as virtue theories in ethics have been described as person-based rather than act-based.

VIRTUE AND KNOWLEDGE

A major motivation for applying virtue theory to the theory of knowledge is that the position explains a wide range of our pretheoretical intuitions about who knows and who does not. Thus suppose we think of intellectual virtues as reliable powers, and we think of knowledge as true belief grounded in such powers. This would explain why beliefs caused by clear vision, mathematical intuition, and reliable inductive reasoning typically have positive epistemic value, and why beliefs caused by wishful thinking, superstition, and hasty generalization do not. Namely, the former beliefs are grounded in intellectual virtues, whereas the latter beliefs are not. Another advantage of a virtue approach is that it seems to provide the theoretical resources for answering important kinds of skepticism. For example, by making epistemic evaluation depend on instancing the intellectual virtues, the approach potentially explains how justified belief and knowledge are possible for beings like us, and even if we cannot rule out skeptical possibilities involving evil demons or brains in vats. The idea is that actually instancing the virtues is what gives rise to knowledge, even if we would not have the virtues, or they would not have their reliability, in certain nonactual situations.

THE ANALYSIS OF KNOWLEDGE. In 1963, Edmund Gettier wrote a short paper purporting to show that knowledge is not true justified belief. His argument proceeded by way of counterexamples, each of which seemed to show that a belief can be both true and justified and yet not amount to knowledge. Here are two examples in the spirit of Gettier's originals:

Case 1. On the basis of excellent reasons, *S* believes that her coworker Mr. Nogot owns a Ford: Nogot testifies that he owns a Ford, and this is confirmed by *S*'s own relevant observations. From this *S* infers that someone in her office owns a Ford. As it turns out, *S*'s evidence is misleading, and Nogot does not in fact own a Ford. However, another person in *S*'s office, Mr. Havit, does own a Ford, although *S* has no reason for believing this (Lehrer 1965).

Case 2. Walking down the road, *S* seems to see a sheep in the field and on this basis believes that there is a sheep in the field. However, owing to an unusual trick of the light, *S* has mistaken a dog for a sheep, and so what she sees is not a sheep at all. Nevertheless, unsuspected by *S*, there *is* a sheep in another part of the field (Chisholm 1977).

In both cases the relevant belief seems justified, at least in senses of justification that emphasize the internal or the subjective, and in both cases the relevant belief is true. Yet in neither case would we be inclined to judge that *S* has knowledge. From the perspective of virtue theory, there is a natural way to think about the two cases. It is natural to distinguish between achieving some end by luck or accident, and achieving the end through the exercise of one's abilities (or virtues). This suggests the following difference between Gettier cases and cases of knowledge. In Gettier cases, *S* believes the truth, but only by accident. In cases of knowledge, however, it is no accident that *S* believes the truth. Rather, in cases of knowledge, *S*'s believing the truth is the result of *S*'s own cognitive abilities—believing the truth can be credited to *S*. To put this another way, in cases of knowledge, *S* believes the truth because *S* is intellectually virtuous. Below are four formulations of this idea:

> We have reached the view that knowledge is true belief out of intellectual virtue, belief that turns out right by reason of the virtue and not just by coincidence. (Sosa 1991)

> Knowledge is a state of true belief arising out of acts of intellectual virtue. (Zagzebski 1996)

> When a true belief is achieved non-accidentally, the person derives epistemic credit for this that she would not be due had she only accidentally happened upon a true belief.… The difference that makes a *value* difference here is the variation in the degree to which a person's abilities, powers, and skills are causally responsible for the outcome, believing truly that *p*. (Riggs 2002)

> When we say that *S* knows *p*, we imply that it is not just an accident that *S* believes the truth with respect to *p*. On the contrary, we mean to say that *S* gets things right with respect to *p* because *S* has reasoned in an appropriate way, or perceived things accurately, or remembered things well, etc. We mean to say that getting it right can be put down to *S*'s own abilities, rather than to dumb luck, or blind chance, or something else. (Greco 2004)

More needs to be said here. In particular, virtue theorists must provide an account of the difference between getting things right by accident and getting things right because one believes out of epistemic virtue. The four quotations above imply that the distinction involves the notions of cause and causal explanation: in cases of knowledge, *S*'s believing the truth is caused by (or explained by) the fact that *S* believes out of epistemic virtue. But these key notions are difficult, and there is no agreement among virtue theorists about how they should be understood.

SKEPTICISM. The problem of skepticism has received sustained attention in the theory of knowledge. Skepticism is best thought of as a theoretical problem, rather than as a practical problem or an existential problem. The problem is not that we might not know what we think we know. Neither is it that we cannot act until skeptical doubts have been adequately laid to rest. Rather, skeptical arguments constitute theoretical problems in the following sense: they begin from premises that seem eminently plausible, and proceed by seemingly valid reasoning to conclusions that are outrageously implausible. The task for a theory of knowledge is to identify some mistake in the skeptical argument and to replace it with something that is theoretically more adequate. It has been argued that a virtue-theoretic approach promises resources for doing just this. To see how, it will be helpful to consider two skeptical arguments.

The first belongs to a family of skeptical arguments, all of which claim that our knowledge of the world depends on how things appear through the senses, and that there is no good inference from how things appear to how things actually are. Here is the argument put formally:

> 1. All of our beliefs about the world depend, at least in part, on how things appear to us via the senses.
>
> 2. The nature of this dependency is broadly evidential: the fact that things in the world appear in a certain way is often our reason for thinking that they are that way.
>
> 3. Therefore, if I am to know how things in the world actually are, it must be via some good inference from how things appear to me. (By 1, 2)
>
> 4. But there is no good inference from how things appear to how things are.
>
> 5. Therefore, I cannot know how things in the world actually are. (By 3, 4)

The argument is a powerful one. Premises (1) and (2) say only that our beliefs about the world depend for their evidence on how things appear to us. That seems undeniable. Premise (4) is the only remaining independent premise, but there are good reasons for accepting it. One reason is that there seems to be no noncircular argument from appearance to reality. This is because any such argument would have to include a premise about the reliability of sensory appearances, but it is hard to see how that such a premise could be justified without relying on sensory appearances to make the case. Second, even if we could formulate a noncircular argument from appearances to reality, no such inference would be psychologically plausible, since we do not make inferences when we form beliefs about objects on the basis of sensory appearances. This is because an inference takes us from belief to belief, but we typically do not have beliefs about appearances. In the typical case, we form our beliefs about objects in the world without forming beliefs about appearances at all, much less by inferring beliefs about the world from beliefs about appearances.

Something in the skeptical argument is not innocent, of course. Here is a suggestion on what it is. The skeptical argument begins with the claim that beliefs about the world depend for their evidence on how things appear, and it concludes from this that knowledge of the world requires a good inference from appearances to reality. But this line of reasoning depends on an implicit assumption: that sensory appearances ground beliefs about the world by means of an inference. It is perhaps at this point that the skeptical reasoning is mistaken, and virtue theory gives us resources for saying why.

Let us define an inference as a movement from premise beliefs to a conclusion belief on the basis of their contents and according to a general rule. According to virtue theory, this is one way that knowledge can be grounded, since making a reliable inference (one in which the general rules used are good ones) is one way of virtuously forming a belief. But it is not the only way. For example, perceptual beliefs are reliably, and therefore vir-

tuously, formed, but not by means of a general rule taking one from belief to belief. When one forms a perceptual belief about the world, one does not begin with a belief about how things appear and then infer a belief about objects in the world. Rather, the process is more direct than that. In a typical case, one reliably moves from appearances to reality without so much as a thought about the appearances themselves, and without doing anything like following a rule of inference. Put simply, our perceptual powers are not reasoning powers. Rather, they are intellectual virtues in their own right, and therefore capable of grounding knowledge directly.

Consider now a different line of skeptical reasoning. René Descartes believes that he is sitting by the fire in a dressing gown. Presumably, he has this belief because this is how things are presented to him by his senses. However, Descartes reasons, things could appear to him just as they do even if he were in fact not sitting by the fire, but were instead sleeping or mad or the victim of a deceiving demon. Again, the point is not that these other possibilities are practical possibilities, or that they are in some sense causes for concern. Rather, the possibilities point to a theoretical problem: On the one hand, it seems that good evidence must rule out alternative possibilities. On the other hand, it seems that Descartes's evidence does not rule out the alternative possibilities in question. But then how can Descartes know that he is sitting by the fire?

Once more it has been argued that a virtue approach has the resources for solving the problem. As stated above, intellectual virtues, including our perceptual powers and our reasoning abilities, may be thought of as intellectual powers or abilities. Yet in general, abilities and powers can achieve success only in relevantly close possible worlds. In other words, to say that someone has an ability to achieve *X* (hitting baseballs, for example) is to say that he would be successful in achieving *X* in a range of situations relevantly similar to those in which he typically finds himself. But then possibilities that do not occur in relevantly similar situations, like the extreme possibilities of skeptical arguments, do not count in determining whether a person has some ability in question. For example, it does not count against Babe Ruth's ability to hit baseballs that he cannot hit them in the dark. Likewise, it does not count against our perceptual powers that we cannot discriminate real fires from demon-induced hallucinations. Accordingly, virtue theory explains why our inability to rule out Descartes's possibility of a demon is irrelevant to whether we have knowledge. Namely, knowledge is true belief grounded in intellectual virtue. The fact that our intellectual faculties would be unreliable in worlds where demons induce perceptions is irrelevant to whether they count as epistemically virtuous in the actual world.

NONTRADITIONAL PROBLEMS

As noted above, a number of virtue epistemologists are interested in traditional problems of epistemology, such as the analysis of knowledge, the nature of epistemic justification, and the problem of skepticism. These philosophers argue that a virtue approach in epistemology provides new insights into old problems. A second camp explicitly advocates a shift away from the traditional problems of epistemology and argues that a virtue approach is the best vehicle for achieving the new focus. These theorists agree that the intellectual virtues should play a central role in epistemology, but they prefer to ask different questions and engage in different projects.

Lorraine Code (1984, 1987) argues for the importance of epistemic responsibility, or the responsibility to know well. Code thinks that such responsibility is related, but not reducible, to our moral responsibility to live well. Redirecting epistemology in this way, she argues, constitutes a more adequate development of the initial insights of virtue epistemology. This is because, in part, the notion of responsibility emphasizes the active nature of the knower, as well as the element of choice involved in the knower's activity. Only an active, creative agent can be assessed as responsible or irresponsible, as having fulfilled obligations to fellow inquirers, and so on. Moreover, placing emphasis on virtue and responsibility has consequences for both how epistemology should be conducted and the kind of epistemological insights to be expected. Echoing a point by Alasdair MacIntyre, Code argues that an adequate understanding of what it is to be virtuous requires placing virtuous selves in the unity of thick narratives. A consequence of this is that we should not expect to describe tidy conditions for justification and knowledge. The relevant criteria for epistemic evaluation are too varied and complex for that, and so any simple theory of knowledge will distort rather than adequately capture those criteria. This does not mean, however, that insight into the nature and conditions of justification and knowledge is impossible. Rather, such insight is to be gained by narrative history rather than theory construction of the traditional sort.

James Montmarquet (1987, 1993) investigates the topic of doxastic responsibility, or the kind of responsibility for beliefs that can ground moral responsibility for actions. Often enough, the morally outrageous actions of tyrants, racists, and terrorists seem perfectly reasonable,

even necessary, in the context of their distorted belief systems. To find their actions blameworthy, we have to find their beliefs blameworthy as well, it would seem. A virtue account, Montmarquet argues, provides what we are looking for. Precisely because it understands justification in terms of epistemically virtuous behavior, in such an account, justified (and unjustified) beliefs can be under a person's control. And this allows relevant beliefs to become appropriate objects of blame and praise.

A common objection to this sort of view, and to virtue accounts in general, is that judgments of responsibility are inappropriate in the cognitive domain. The idea is that judgments of praise and blame presuppose voluntary control, and that we lack such control over our beliefs. Montmarquet responds to this objection by distinguishing between a weak and a strong sense of voluntary control. Roughly, a belief is voluntary in the weak sense if it is formed in circumstances that allow virtuous belief formation. This kind of voluntariness amounts to freedom from interference or coercion. A belief is voluntary in the strong sense (again roughly) if it is fully subject to one's will. Montmarquet concedes that responsibility requires weak voluntary control, but argues that we often have this kind of control over our beliefs. On the other hand, we do not typically have strong voluntary control over our beliefs, but responsibility does not require it.

Finally, Jonathan Kvanvig (1992) has argued for a more radical departure from traditional epistemological concerns. According to Kvanvig, traditional epistemology is dominated by an "individualistic" and "synchronic" conception of knowledge. From the traditional perspective, an important task is to specify the conditions under which individual *S* knows proposition *p* at time *t*. Kvanvig argues that this perspective should be abandoned in favor of a new social and genetic approach. Whereas the traditional perspective focuses on questions about justified belief and knowledge of individuals at particular times, a new genetic epistemology would focus on the cognitive life of the mind as it develops within a social context. In the new perspective, questions concerning individuals are replaced with questions concerning the group, and questions concerning knowledge at a particular time are abandoned for questions about cognitive development and learning. Kvanvig argues that virtues are central within the new perspective in at least two ways. First, epistemic virtues are essential to understanding the cognitive life of the mind, particularly the development and learning that takes place over time through mimicking and imitating virtuous agents. Second, in a social and genetic approach, epistemic virtues play a central role in characterizing cognitive ideals. For example, a certain structuring of information is superior, Kvanvig argues, if an epistemically virtuous person would come to possess such a structure in appropriate circumstances.

See also Aristotle; Code, Lorraine; Descartes, René; MacIntyre, Alasdair; Skepticism, Contemporary; Sosa, Ernest.

Bibliography

Axtell, Guy. "Epistemic-Virtue Talk: The Reemergence of American Axiology?" *Journal of Speculative Philosophy* 10 (3) (1996): 172–198.

Axtell, Guy, ed. *Knowledge, Belief, and Character*. Lanham, MD: Rowman and Littlefield, 2000.

Axtell, Guy. "Recent Work in Virtue Epistemology." *American Philosophical Quarterly* 34 (1) (1997): 410–430.

Axtell, Guy. "The Role of the Intellectual Virtues in the Reunification of Epistemology." *Monist* 81 (3) (1998): 488–508.

Chisholm, Roderick. *Theory of Knowledge*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1977.

Code, Lorraine. *Epistemic Responsibility*. Hanover, NH: Published for Brown University Press by University Press of New England, 1987.

Code, Lorraine. "Toward a 'Responsibilist' Epistemology." *Philosophy and Phenomenological Research* 45 (1) (1984): 29–50.

DePaul, Michael, and Linda Zagzebski, eds. *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford, U.K.: Oxford University Press, 2004.

Fairweather, Abrol, and Linda Zagzebski, eds. *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford, U.K.: Oxford University Press, 2001.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121–123.

Goldman, Alvin. "Epistemic Folkways and Scientific Epistemology." In his *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge, MA: MIT Press, 1992.

Greco, John. "Knowledge as Credit for True Belief." In *Intellectual Virtue: Perspectives from Ethics and Epistemology*, edited by Michael DePaul and Linda Zagzebski. Oxford, U.K.: Oxford University Press, 2004.

Greco, John. *Putting Skeptics in Their Place*. New York: Cambridge University Press, 2000.

Greco, John, ed. *Sosa and His Critics*. Oxford, U.K.: Blackwell, 2004.

Greco, John. "Virtue Epistemology." In *The Stanford Encyclopedia of Philosophy*, Fall 1999, edited by Edward N. Zalta. Available from http://plato.stanford.edu/.

Greco, John. "Virtues in Epistemology." In *The Oxford Handbook of Epistemology*, edited by Paul Moser. Oxford, U.K.: Oxford University Press, 2002.

Hookway, Christopher. "Cognitive Virtues and Epistemic Evaluations." *International Journal of Philosophical Studies* 2 (2) (1994): 211–227.

Kvanvig, Jonathan. *The Intellectual Virtues and the Life of the Mind*. Savage, MD: Rowman and Littlefield, 1992.
Lehrer, Keith. "Knowledge, Truth, and Evidence." *Analysis* 25 (1965): 168–175.
Lehrer, Keith. *Theory of Knowledge*, 2nd ed. Boulder, CO: Westview Press, 2000.
Montmarquet, James. "Epistemic Virtue." *Mind* 96 (1987): 482–497.
Montmarquet, James. *Epistemic Virtue and Doxastic Responsibility*. Lanham, MD: Rowman and Littlefield, 1993.
Plantinga, Alvin. *Warrant and Proper Function*. Oxford, U.K.: Oxford University Press, 1993.
Riggs, Wayne. "Reliability and the Value of Knowledge." *Philosophy and Phenomenological Research* 64 (1) (2002): 79–96.
Sosa, Ernest. "Beyond Internal Foundations to External Virtues." In *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*, by Laurence BonJour and Ernest Sosa. Oxford, U.K.: Blackwell, 2003.
Sosa, Ernest, "Intellectual Virtue in Perspective." In his *Knowledge in Perspective*. Cambridge, U.K.: Cambridge University Press, 1991.
Sosa, Ernest. *Knowledge in Perspective*. Cambridge, U.K.: Cambridge University Press, 1991.
Sosa, Ernest. "The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge." *Midwest Studies in Philosophy* 5 (1980): 3–25.
Zagzebski, Linda. *Virtues of the Mind*. Cambridge, U.K.: Cambridge University Press, 1996.
Zagzebski, Linda. "What Is Knowledge?" In *The Blackwell Guide to Epistemology*, edited by John Greco and Ernest Sosa. Oxford, U.K.: Blackwell, 1999.

John Greco (1996, 2005)

VIRTUE ETHICS

In 1930 C. D. Broad first proposed to divide ethical theories into two classes, teleological and deontological, thereby introducing a dichotomy that quickly became standard in ethics. Teleological theories were defined as ones that hold that the moral rightness of an action is always determined by its tendency to promote certain consequences deemed intrinsically good; deontological theories, as ones that deny this claim. Broad's dichotomy was widely accepted as being exhaustive, but in fact there are two fundamental classes of normative moral judgments that do not fit easily into it. First, it focuses on rightness or obligation, excluding moral judgments concerning what is admirable, good, excellent, or ideal. Second, it concerns only actions and their consequences, saying nothing about moral judgments concerning persons, character, and character traits.

The contemporary movement known as virtue ethics is usually said to have begun in 1958 with Elizabeth Anscombe's advice to do ethics without the notion of a "moral ought." Although her own critique of moral-obligation concepts (viz., that they have meaning only within religious frameworks that include the notion of a divine lawgiver) did not gain widespread acceptance among secular ethicists, her constructive proposal to look for moral norms not in duty concepts but within the virtues or traits of character that one needs to flourish as a human being quickly caught on. Soon thereafter philosophers such as Alasdair MacIntyre, Philippa Foot, Edmund Pincoffs, and many others began to articulate and defend a third option in normative ethics: one whose chief concern was not a theory of morally right action but rather those traits of character that define the morally good or admirable person.

Phrases such as "revival of" or "return to" often precede mention of virtue ethics in contemporary discussions, and it is generally true that questions about the virtues occupy a much more prominent place in ancient and medieval moral philosophy than in moral theories developed since the Enlightenment. But it is important to note that the conscious awareness of virtue ethics as a distinct way of theorizing about ethics arose from within contemporary Anglo American ethical theory. Virtue ethics took root as a reaction against the underlying common assumptions of both teleological and deontological ethical theories and has achieved its greatest critical success as a protest against these accepted ways of doing normative ethics. Accordingly, one can view virtue ethics as having two complementary aspects: a critical program that presents a critique of the prevailing assumptions, methods, and aspirations of normative teleological and deontological moral theories; and a constructive program, in which an alternative virtue-oriented normative moral conception is developed and defended.

THE CRITICAL PROGRAM

At this first level virtue theorists are not necessarily committed to defending a full-scale alternative to existing ethical theory programs but rather to showing why such approaches are systematically unable to account satisfactorily for moral experience. Major criticisms made by virtue theorists against their opponents include the following.

OVERRELIANCE ON RULE MODELS OF MORAL CHOICE. Utilitarians and Kantians, it is held, both mistakenly view universal and invariable principles and laws as being exhaustive of ethics. But real-life moral exemplars do not simply deduce what to do from a hierarchy

of timeless, universal principles and rules. They possess sound judgment skills that enable them to respond appropriately to the nuances of each particular situation in ways that go beyond mere mechanical application of rules.

OVERLY RATIONALISTIC ACCOUNTS OF MORAL AGENCY. Traditional moral theorists, it is held, too often assign a merely negative role in the moral life for desires and emotions. However, morally admirable people are not simply people who do their duty, but people who do so with the right kinds of emotions. Additionally, though many teleologists and deontologists do acknowledge the importance of motives in ethics, they typically mislocate them in abstractions such as "the greatest happiness principle" or "the moral law" rather than in particular persons and our relationships to them.

FORMALISM. Mainstream teleological and deontological theorists tend to focus exclusively on conceptual analyses of their favored duty-concepts and then on logical arguments based on such analyses. Additionally, they tend to view moral questions as arising only when an individual agent is trying to decide what to do in certain problematic situations. These methodological commitments result in a view of morality that is impoverished and overly restrictive. Virtue theorists, on the other hand, are much more open to drawing connections between morality and other areas of life such as psychology, anthropology, history, art, and culture. Their long-term agent-perspective also enables them to correctly view moral deliberation and choice as involving much more than snapshot decisions.

THE CONSTRUCTIVE PROGRAM

In offering their alternative, virtue theorists face the fundamental task of showing how and why a virtue-oriented conception of ethics is superior to its act- and duty-based competitors. In what ways is moral experience better understood once virtue-concepts become the primary tools of analysis? Here one may distinguish two general tendencies: Radical virtue ethics attempts to interpret moral experience and judgment without employing duty-concepts at all (or at least by claiming that such concepts are always derivable from more fundamental ones concerning good people—for example, "morally right" acts might be defined simply as those acts performed by moral exemplars); moderate virtue ethics seeks to supplement standard act approaches with an account of the virtues. The former approach tends to view teleological and deontological ethical theories as totally misguided; the latter sees them merely as incomplete. Major issues confronting constructive virtue ethics programs include the following.

DEFINING MORAL VIRTUE. What counts as a moral virtue and why? Is there any plausible way to distinguish between moral and nonmoral virtues? How exactly do virtues relate to actions, reasons, principles, rules, desires, emotions? Are virtues beneficial to their possessors, and, if so, are they too self-centered to count as moral traits?

JUSTIFYING THE VIRTUES. How can we establish the validity of those character traits defined as moral virtues, once the option of appealing to the value of the acts that the virtues tend to encourage is ruled out? Traditionally, moral virtues have been defined as traits that human beings need in order to live well or flourish. But does the idea of flourishing provide solid enough ground on which to base the moral virtues? Is it still possible to speak accurately of *a* single human function, or is human life more variously textured than the classical picture allows? How and why is evidence of flourishing necessarily evidence of moral virtuousness? On the other hand, if one declines to issue pronouncements about "the human *telos*" and instead opts for a softer, more pluralistic functionalism that seeks to define virtues in terms of different kinds of human purposes or practices, can one still arrive at a substantive notion of the virtues that holds that they are more than local cultural products?

APPLYING THE VIRTUES. How do the virtues relate to one another in real life? Is there anything to the ancient "unity of virtues" thesis (which, on the Aristotelian model, views *phronesis* or practical wisdom as generating and uniting all of the moral virtues), or does it make sense to hold that a person might possess one moral virtue such as courage and nevertheless lack others? How many different moral virtues are there? Are some more fundamental than others? Can they be ranked in order of importance? Do virtues ever conflict with one another? What kinds of specific practical guidance do we get from the virtues, especially in cases where they appear to conflict with one another (e.g., honesty vs. kindness, love vs. fidelity)?

It should come as no surprise that radical virtue-ethics approaches have attracted far fewer followers than more moderate versions and that the critical program has had a much stronger influence on contemporary ethical theory than has the constructive program. Those who turn to late-twentieth-century work in virtue ethics in hopes of finding greater consensus on either theoretical

or normative issues than exists among ethical theorists elsewhere are bound to be disappointed. Still, it is no small sign of virtue ethics's success that contemporary ethical theorists of all persuasions are addressing questions of character, agency, and motivation as never before—and that there now exist greater realism and humility among contemporary philosophers concerning how ethical theory should proceed and what it might reasonably accomplish.

See also Anscombe, Gertrude Elizabeth Margaret; Broad, Charlie Dunbar; Consequentialism; Deontological Ethics; Kant, Immanuel; Metaethics; Utilitarianism.

Bibliography

Annas, Julia. "Virtue Ethics." In *The Oxford Companion to Ethical Theory*, edited by David Copp. Oxford: Oxford University Press, 2004.

Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33 (1958): 1–19. Reprinted in her *Collected Philosophical Papers*. Vol. 3. Minneapolis, 1981.

Aristotle. *Nicomachean Ethics*.

Broad, C. D. *Five Types of Ethical Theory*. London: Kegan Paul, 1930. See pp. 206–207 for Broad's division of ethical theories into deontological and teleological.

Flanagan, O., and A. O. Rorty, eds. *Identity, Character, and Morality: Essays in Moral Psychology*. Cambridge, MA: MIT Press, 1990. Nineteen commissioned essays; see esp. part 5.

Foot, Philippa. *Natural Goodness*. Oxford: Oxford University Press, 2001.

Foot, Philippa. *Virtues and Vices*. Berkeley: University of California Press, 1978. Reprinted in *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford: Oxford University Press, 2002.

French, P. A., T. E. Uehling, and H. K. Wettstein, eds. *Ethical Theory: Character and Virtue*. Midwest Studies in Philosophy. Vol. 13. Notre Dame, IN: University of Notre Dame Press, 1988. Twenty-nine commissioned essays.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford: Oxford University Press, 1999.

Johnson, Robert. "Virtue and Right." *Ethics* 115 (2003): 810–834.

Kruschwitz, R. B., and R. C. Roberts, eds. *The Virtues: Contemporary Essays in Moral Character*. Belmont, CA: Wadsworth, 1987. Seventeen essays. The first anthology on the topic. Includes an extensive bibliography of relevant works published up to 1985.

MacIntyre, A. *After Virtue*. 2nd ed. Notre Dame, IN: University of Notre Dame Press, 1984.

MacIntyre, A. *Dependent Rational Animals*. Chicago and La Salle, IL: Open Court, 1999.

McDowell, John. *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, 1998.

Philosophia 20 (1990). Double issue on virtue, with special reference to Philippa Foot's work. Thirteen commissioned essays.

Pincoffs, E. L. *Quandaries and Virtues*. Lawrence: University Press of Kansas, 1986.

Slote, M. *From Morality to Virtue*. New York: Oxford University Press, 1992.

Slote, M. *Morals from Motives*. Oxford: Oxford University Press, 2001.

Sreenivasan, Gopal. "Errors about Errors: Virtue Theory and Trait Attribution." *Mind* 111 (2002): 47–68.

Statman, D., ed. *Virtue Ethics*. Edinburgh: Edinburgh University Press, 1996.

Swanton, Christine. *Virtue Ethics: A Pluralistic View*. Oxford: Oxford University Press, 2003.

Wallace, J. *Virtues and Vices*. Ithaca, NY: Cornell University Press, 1978.

Robert B. Louden (1996)
Bibliography updated by Rosalind Hursthouse (2005)

VISUAL ARTS, THEORY OF THE

There are competing views on what qualifies photographs, paintings, sculpture, and architecture as visual arts. This entry focuses on theories of vision and their implications for claims about each of these four art forms. There is also debate over whether it is desirable to identify these major categories of art in terms of particular sense modalities. What is partly at issue is whether vision and visual experience are isolated from other sense modalities. The status of photography, painting, sculpture, and architecture as major art forms is by no means beyond challenge; they, along with their paradigm cases, exhibit considerable variation within and across cultures, and through time.

PHOTOGRAPHY

Photography, like vision, seems to have an especially intimate connection with the world by virtue of a causal or "mechanical" process that is describable in purely physical terms. Interestingly, this alleged mechanical connection has also been responsible for the lion's share of skepticism about whether photography is indeed an art. The basic idea is that the appearance of a photograph is, like visual experience itself, dependent in a special way on the presence of the targeted object or scene. The claim is not (necessarily) that a photograph looks like the object or scene in the world, but rather that the way the photograph looks is, in an important way, independent of intentions or other mental states of the photographer, even if the photographed scenes are staged or an object's appearance is manipulated or disguised. The possibilities for manipulation and disguise, in fact, motivate a distinc-

tion, fundamental to most theories of photography as an art, between what is *photographed* and what the photograph is a *picture* of, that is, what it pictorially represents. For example, a photograph of the forequarters of a jackrabbit, suitably enlarged, cropped, and merged with a photograph of the hindquarters of an antelope may end up as a picture of a fictional animal, a "jackelope." What is photographed is due to a process independent of a photographer's mental states, while what is pictured is conceived as dependent at least in part on the artist's intentions or cultural context.

Kendall Walton has argued that the viewer of a photograph literally sees the object that has been photographed (commonly known as the transparency thesis), that a photograph's transparency constitutes one type of photographic realism, and that this realism accounts for a significant part of a photograph's power. Walton claims that, just as telescopes enable us to see things far away, photographs enable us to see things in the past. However, one may accept the relevance of what is photographed to the work's content but reject the claim that photographs are transparent, that is, that one *sees* the object or scene. For example, in one account of what it is to see something, one's visual experience provides information about the spatial location of the viewer, so-called egocentric spatial information, in relation to what is seen. Since neither photographs nor paintings provide such information, it is concluded that neither is transparent. Some allege that photographs may provide such information, such as information that the viewer of the photograph is at that time standing where the photographer was when he or she took the photograph. Others object that, even though seeing generally carries egocentric spatial information, it does not always do so, such as when one sees something in a series of mirrors.

Options multiply. Some allege that one does not see actual objects or scenes in mirrors, but only their reflections. A variant view takes the relevant concept to be what can generally be expected from a given type of perceptual process, rather than what it always provides (Cohen and Meskin 2004). The crucial point for this view is that it is reasonable to expect egocentric spatial information from vision but not from photographs or paintings. Further, it needs to be acknowledged that what can reasonably be expected may vary in relation to context and an individual's powers or background of experience. The increasing ease with which digital images can be manipulated in fact makes it reasonable to be skeptical about many of their alleged information-bearing properties.

Some art photographers, not surprisingly, have made the alleged realism and associated power of photography part of the subject matter of their work. Jerry Uelsmann's combinations of photographs within the same image give them a surreal and sometimes mystical character. Zeke Berman constructs and photographs little stage sets that create visual ambiguities in the photograph's pictorial space. Manipulated photographs of fantastical animals are part of installations designed by Joan Fontcuberta and Pere Formiguera as a send-up of the supposed objectivity of photographic documentation in ethnological and anthropological studies. Artists may also use photographs of some objects—qua photographs of those objects—as materials for making pictures having a different content, connecting with a general question in the visual arts over whether and how the character of the materials artists use affect the content or significance of the work.

Suppose, for the sake of argument, that the way the objects in a photograph look is dependent on the presence of the objects in front of the camera at a given time. Nigel Warburton (1988) criticizes this "snapshot" account of seeing on the grounds that, in ordinary seeing, visual experiences of an object change as the object and viewer move in relation to each other. Warburton concludes that, because photographs—like paintings—do not have this property, viewers of the photograph do not literally see the objects photographed. The relevant visual concern then becomes how one looks at something rather than what one sees, which in turn raises questions about relationships among vision, space, and time that are relevant to all of the visual arts.

PAINTING

Painting is sometimes thought to be the visual art par excellence. Confusingly, however, the term *painting* is frequently used to indicate drawings, prints, collage, and almost any other method or materials used to create something that, crudely put, can be hung on a wall. Literally construed, paintings are composed of paint; how artists work with different physical materials, such as paint, to make art would seem to be relevant, even central, to appreciating them. Paintings, broadly construed, may also pictorially represent things, arguably in virtue of the two-dimensional array of line, shape, and color, abstracted from whatever medium is used. The development of various technologies to mass-produce two-dimensional arrays raises the question whether merely being a two-dimensional array is enough to warrant status as art.

Paintings and drawings are plausibly thought of as physical objects; prints, such as Rembrandt's *Medea* (1648), are not. Jerrold Levinson (1996, p. 131) calls prints (and other types of art) that have many impressions, such that no individual impression is identical with the work itself, "multiples." But not all prints are multiples. Monoprints, like paintings and drawings, are singular because they are, by definition, produced by a process in which only one physical object can count as a genuine exemplar (to use Levinson's terminology) of the work. Photographs, depending on what photographic process is used, may also be multiple or singular.

These ontological differences have implications for whether one can see the (allegedly) visual work of art. Impressions of a print can be seen, but the print itself, as opposed to its exemplars, cannot be seen (or, less precisely, it can be seen only "in" or "through" impressions of it). Prints may also have multiple states—stages in the printmaking process—some of which may be considered to be works in themselves and each of which may have multiple impressions. Prints may be grouped together as a suite, such as the four plates of Hogarth's *The Analysis of Beauty*, raising the possibility that the set constitutes the work of art. Even if one accepts impressions of prints as works of art in their own right, they are still impressions of a print, which is not itself a physical object.

Other media, such as mosaics, introduce further complications, and may undermine the precision of the singular/multiple distinction. Tesserae can be mass-produced and combined formulaically to cover a surface with a pattern or image, which would seem to make them multiples, though mosaics of this type are rarely considered works of art. Highly sophisticated forms, such as those that evolved under the rule of Justinian, by contrast, have greater claim to be singular works of art. They are products of a workshop tradition very similar to that which persisted for centuries in Europe for painting. The master was responsible for the overall design and implementation of its most important components, such as the figures, especially faces and hands; assistants provided backgrounds and possibly drapery. Rubens's assistants painted large portions of works that we identify as singular works by Rubens; Constable's *Salisbury Cathedral* was so popular that he painted seven of them. Are they copies of a single work—a sort of prototype that cannot itself be seen—or seven different paintings that are visually virtually identical?

Titles are linguistic entities that may be given by painters themselves, making them clearly part of the artwork, though only debatably part of the painting. An inscription of the painting's title—or of other words, for that matter—in the painting itself may be a visually significant property of the painting, but as a linguistic entity or property, it is generally nonvisual. The caveat is necessary since some types of linguistic inscriptions—calligraphy, for example—have visually significant properties. But a painting may have a title that is not inscribed and hence not at all visible, yet is still part of the work.

Singular works of visual art, such as paintings and drawings, are physical objects, but this does not preclude them from having representational and expressive qualities, or from playing a role in a culture and in history, including the history of art. Many artifacts—furniture, tools, and televisions—have functions. Paintings are artifacts with the function of providing certain kinds of visual experiences. Nelson Goodman trenchantly criticizes the idea that pictures show us "the way things look." Richard Wollheim argues that the crucial visual experience is what one can *see in* a painting. As he puts it, "The marked surface must be the conduit along which the mental state of the artist makes itself felt within the mind of the spectator if the result is to be that the spectator grasps the meaning of the picture" (1987, p. 22). The artist's hand is to guide the perceiver's eye; how an artist works with the physical materials, as a medium, is essential to the painting's meaning.

Making a work of art, however, may require more than what a lone painter can do or what any given viewer can see in what is created. Arthur Danto proposes that even ordinary objects can be "transfigured" into art by the existence of a theory and history of art, which is something "the eye cannot descry" (1960, p. 580). The art world of the mid-twentieth century subsumed not only painted surfaces but also commonplace objects, visually indistinguishable from ordinary, everyday objects or real things, into the category of art. Everyday artifacts can constitute a medium, not merely materials, for making art, standing alone or as part of a construction or installation.

A contrasting view proposes that visual works of art are pure appearances and denies that any of them, not merely multiples, are physical objects. These virtual objects, as Susanne Langer characterizes them, are "created solely for the eye" (1953, p. 10), and as such have no practical purpose or function dependent on the physical characteristics of their constitutive materials. Clive Bell's (1958) concept of significant form as line, color, their relationship, and a sense of space also depends on the purely visual, independent of both concepts and use. Both take the value of the visual arts to be in the visual

experience one has in the presence of a work. Neither view accommodates the attribution of different contents to objects that are visually indistinguishable, something that Wollheim and Danto, in their different ways, are at pains to allow.

Visual experience is not merely ocular. Paintings, drawings, prints, and even photographs may be created with the expectation that perceivers' visual experiences will not be isolated from other sense modalities any more than from their concepts and beliefs. At a minimum, as Wollheim points out, we see with an *embodied* eye: at a distance and from a given (literal) point of view. The interplay of figure and ground is not merely visual, but relates to experiences of physical proximity and distance. Volumes and shapes are apprehensible by both vision and touch. Studies show that congenitally blind subjects can identify the content of raised line drawings of outline shape, contour, and even vanishing-point perspective. Paul Crowther (2002) goes further, taking the relevant connections to be not only between vision and touch, but between visual and motor *exploration* of the world and of the work. Past actions, including working with the relevant types of materials, affect visual experience. Chinese calligraphy, for example, is a semi-pictorial, linguistic inscription, and part of the tradition of appreciating it is to mimic making the brushstrokes in the air, using memory and imagination.

SCULPTURE

Sculptures and paintings relate differently to the space around them, depending on what one takes to be the paradigm cases of each. Suppose we take paintings that employ vanishing-point perspective, where there is a point internal to the space represented in the picture from which things are shown (its so-called "internal depiction point"), as its paradigm cases (Hopkins 2004). Such painting, like vision, is perspectival in that it organizes what it represents from a particular point of view. Sculpture in the round, by contrast, occupies a space that is continuous with that of the perceiver. It organizes the space around it, drawing on the perceiver's and the represented object's potential for movement and action. It has no internal depiction point, and hence is not perspectival. Tactile and somatosensory phenomena have a more obvious role to play in the appreciation of sculpture in the round, something that may be seen as a resource that enhances its power, or as an appeal to the "lower senses" as compared with the cognitively more esteemed sense of sight.

Sculpture might seem to be by definition three-dimensional, and visible as such, though the existence of multiples confounds this simple requirement, as they do with paintings and prints. Cast sculptures are multiples; one sees the exemplars, such as Rodin's *The Thinker*, but not the work itself. Cast and molded pieces are routinely hand-worked in various ways—painted, appliquéd, carved, and so on—motivating the acceptance of such pieces as works themselves rather than merely as exemplars of a type that cannot itself be seen. Installation art that is to be installed differently in different sites also challenges the idea that one sees the work rather than a particular installation of it.

An alternative strategy for distinguishing sculpture from pictures takes the sculptural to be a property of a work rather than sculpture as a category of art, so that a single work may have both sculptural and pictorial properties (Koed 2005). The basic idea is that materials—which can include paint—are treated as a sculptural medium when their three-dimensionality is used for representational purposes. For example, Paleolithic cave paintings are sculptural in exploiting protrusions from the cave wall to emphasize the swell of a bison's forequarters, and pictorial in exploiting line and color, applied to the surface of the cave wall, to represent a particular, or a particular type of, animal.

Bas-relief, including painted bas-relief, as a mode of representation, may actually have a closer connection with ordinary vision than either vanishing-point perspective painting or sculpture in the round, and hence is arguably a better candidate for being a paradigm case of visual art than painting or sculpture. Sculpture in the round exploits a crucial feature of ordinary vision that pictures do not, that is, one's visual experience changes as one moves around the object (except *per accidens*, as when looking at a sphere in a cylindrical room). However, sculpture accomplishes this by replicating, or at least approximating, the three-dimensional shape of an object rather than by representing it. One could say it presents, rather than represents, the shape of the object. In contrast, when pictures represent the shape of an object, they do not generate the changes in visual experience that ordinary seeing does as the perceiver moves about (except again, *per accidens*, as with pictures of relatively flat objects, such as pieces of paper). Visual experiences of bas-relief, however, change in relevant ways as one changes physical position in relation to the relief, though the relief does not replicate the shape that the represented object has in the round (again, except accidentally). Indeed, a low relief representation of a relatively flat

object, such as a piece of paper, may have greater physical depth than both what it represents and three-dimensional representations of it. Bas-relief, within the requisite distance, also enhances without replicating the visual experiences of three-dimensional form and depth that are due to binocular disparity.

Installation art can be treated as a form of sculpture in an "expanded field " (Krauss 1983), though the relationship of sculpture to its surrounding spaces is better illuminated by contrasting it with installation art. Artists have control over the entire designated space for their installations, rather than merely over the construction of individual objects. A good case can be made for installation art as a distinct category of art, one where the artist has control over the entire space that the installation occupies, in contrast to a type of art such as sculpture, where the work may be seen as controlling the space that surrounds it. Video installations employing speech and music are certainly not merely visual and are probably best grouped with other video art and film. Further, the status of the objects within installations—and some installations contain no objects at all—is different from sculpture because the space of the installation is exploited in a variety of ways and is often treated as a gray area between life and art. Perceivers may be required to "complete" the work, for example, by stepping on a switch that turns on a light, as one engages in the semi-voyeuristic activity of viewing Marcel Duchamp's *Étant donnés*. Installations often create environments that prompt self-consciousness or reflection on one's habitual actions or role as viewer.

The continuity of the installation space with lived space can be facilitated by the use of everyday objects and by invitations to treat the material components of the piece in ways ordinarily forbidden with art, such as when museum-goers are invited to walk on Carl Andre's 144 zinc squares. The use of ephemeral materials, such as banana peels, critiques the timelessness associated with traditional sculpture in the round by creating objects that one is not merely to see but to see deteriorate over time. Museums routinely display, for our visual delectation, objects that were created to be used and not merely looked at, such as illuminated manuscripts, ceramics, and furniture, which may already show signs of deterioration and wear, itself a candidate for visual appreciation, as with the Japanese *sabi* aesthetic developed by the sixteenth-century tea master Rikyu.

ARCHITECTURE

One could develop a theory of architecture as a purely visual art by separating its form from its function; alternatively, one might posit the appearance of functionality, apprehended by the imagination, as the object of experience when treating architecture as a purely visual art. A more promising approach takes functionality as integrating one's experience of form, just as pictorial content informs one's appreciation of a painting. Deep traditions in architectural theory see it as structuring, in a positive way, how one lives and works. Architecture, as an art—though not as a purely visual art—is thus conceived and evaluated by its contribution to, or inhibition of, domestic life or commercial work. If architecture is the attempt to build well, it will accommodate and fulfill purposes that are partly informed by individual needs and desires, and partly by cultural and social realities.

Herbert Read takes the monument—a solid, sculptured edifice—as a paradigm category of visual art, attenuating toward sculpture in one direction and architecture in another. Monuments, like architecture, have functions, so his typology of visual art does not provide any assistance on how to reconcile a work's function with its status as a visual art. But it does provide a way of thinking of something built, as opposed to something sculpted, as an organic whole. A striking example is the Rajrani Temple in Bhubaneswar, a solid temple with no interior. Its inverse is exemplified by cave temples of Ajanta and the Kailasanatha at Ellora, which are not so much built as carved out of "living" rather than dismembered chunks of rock. Its sculpture is subservient to, and inseparable from, the temple's overall form. As one conceptually pulls sculpture away from the monument, however, it goes through a phase that makes it vulnerable to the charge that it is mere decoration, as in Robert Venturi's characterization of architecture as a "decorated shed," which simultaneously ridicules both the structure and its ornament. Thinking of architecture in this way invites the question "What needs to be added to a building to make it architecture?", which has as little promise as the fundamental question in philosophy of architecture as does "What needs to be added to a set of lines, shapes, and colors to make it a picture?" in philosophy of painting.

A recurrent issue in the philosophy of the visual arts is whether visually indistinguishable replicas have any status as art, let alone the same status, or meanings, as what they replicate. In architecture, the replica takes the form of the historic reconstruction of a building or a set of buildings. As with painting, one must ask from what vantage point, and to whom, the buildings are supposedly

visually indistinguishable. Supposing that the relevant view is from the street, some cities have ordinances that allow the gutting of a building's interior, provided that the building's facade is preserved. Some designated historic districts in cities in the United States prohibit indistinguishable replicas or restorations, to ensure that anyone with even a minimally informed eye will know that whatever looks like an original will in fact be one. In contrast, the inauthenticity of the prettified "reconstruction" of a colonial village at Williamsburg, Virginia, has been criticized, even ridiculed, by comparing it with the deliberate artifice of the "leisure entertainment" of Disneyland. But Disneyland's "Main Street," it should be noted, is sometimes identified as installation art. It is a site for various ordinary activities, but different enough not to be confused with everyday life. Modified reconstructions that are altered to be economic, practical, and family-friendly are commonly criticized for the same reasons as copies or forgeries of paintings: they are not visually the same.

See also Art, Definitions of; Art, Ontology of; Art, Performance in; Art, Representation in.

Bibliography

Bell, Clive. *Art* (1913). New York: Putnam, 1958.

Berger, John. *Ways of Seeing*. London: British Broadcasting Company and Penguin Books, 1972.

Cohen, Jonathan, and Aaron Meskin. "On the Epistemic Value of Photographs." *Journal of Aesthetics and Art Criticism* 62 (2) (2004): 197–210.

Crowther, Paul. *The Transhistorical Image: Philosophizing Art and Its History*. Cambridge, U.K.: Cambridge University Press, 2002.

Currie, Gregory. "Photography, Painting and Perception." *Journal of Aesthetics and Art Criticism* 49 (1) (1991): 23–29.

Danto, Arthur, and Sidney Morgenbesser, eds. *Philosophy of Science.* New York: Meridian Books, 1960.

Graham, Gordon. "Architecture." In *Oxford Handbook of Aesthetics*, edited by Jerrold Levinson, 555–571. New York: Oxford University Press, 2003.

Hopkins, Robert. "Painting, Sculpture, Sight, and Touch." *British Journal of Aesthetics* 44 (2) (2004): 149–166.

Hopkins, Robert. *Picture, Image, and Experience*. Cambridge, U.K.: Cambridge University Press, 1998.

Koed, Eric. "Sculpture and the Sculptural." *Journal of Aesthetics and Art Criticism* 63 (2) (2005): 147–154.

Krauss, Rosalind. "Sculpture in the Expanded Field." In *The Anti-Aesthetic: Post Modern Culture*, edited by Hal Foster, 31–42. London and Sydney: Pluto Press, 1983.

Langer, Susanne. *Feeling and Form*. New York: Scribners, 1953.

Levinson, Jerrold. "Titles." *Journal of Aesthetics and Art Criticism* 44 (1) (1995): 29–39. Reprinted in *Music, Art, and Metaphysics*, 159–178. Ithaca, NY: Cornell University Press, 1990.

Levinson, Jerrold. "The Visual Work of Art." In *The Dictionary of Art*. London: Macmillan, 1996. Reprinted in *The Pleasures of Aesthetics*, 129–137. Ithaca, NY: Cornell University Press, 1996.

Lopes, Dominic McIver. "Art Media and the Sense Modalities: Tactile Pictures." *Philosophical Quarterly* 47 (189) (1997): 425–440.

Margolis, Joseph. "Painting." In *The Blackwell Guide to Aesthetics*, edited by Peter Kivy, 215–229. Malden, MA: Blackwell Publishing, 2004.

Maynard, Patrick. "Drawing and Shooting: Causality in Depiction." *Journal of Aesthetics and Art Criticism* 44 (1) (1985): 115–129.

Maynard, Patrick. *The Engine of Visualization: Thinking Through Photography*. Ithaca, NY: Cornell University Press, 1997.

Mey, Kerstin. "Sculpture." In *Encyclopedia of Aesthetics*. Vol. 4, edited by Michael Kelly. New York: Oxford University Press, 1998.

Read, Herbert. *The Art of Sculpture*. Princeton, NJ: Princeton University Press, 1961.

Rosenthal, Mark. *Understanding Installation Art: From Duchamp to Holzer*. Munich, Berlin, London, New York: Prestel, 2003.

Savedoff, Barbara E. *Transforming Images: How Photography Complicates the Picture*. Ithaca, NY: Cornell University Press, 2000.

Scruton, Roger. *The Aesthetics of Architecture*. Princeton, NJ: Princeton University Press, 1979.

Venturi, Robert. *Learning from Las Vegas*. Cambridge, MA: MIT Press, 1977.

Walton, Kendall. "Transparent Pictures: On the Nature of Photographic Realism." *Critical Inquiry* 11 (1984): 246–276.

Warburton, Nigel. "Seeing through 'Seeing Through Photographs.'" *Ratio* n.s. 1 (1988): 64–74.

Wollheim, Richard. *Art and Its Objects*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1980.

Wollheim, Richard. *Painting as an Art*. Princeton, NJ: Princeton University Press, 1987.

Susan L. Feagin (2005)

VITALISM

"Vitalism" is primarily a metaphysical doctrine concerning the nature of living organisms, although it has been generalized, by Henri Bergson for example, into a comprehensive metaphysics applicable to all phenomena. We shall examine vitalism only as a theory of life.

There have been three general answers to the question "What distinguishes living from nonliving things?" The first, and currently most fashionable, answer is "A complex pattern of organization in which each element of the pattern is itself a nonliving entity." In this view, a living organism, and each of its living parts, is exhaustively composed of inanimate parts; and these parts have

no relations except those that are also exhibited in inanimate systems. The second answer is "The presence in living systems of emergent properties, contingent upon the organization of inanimate parts but not reducible to them." This answer resembles the first in acknowledging that a living system is exhaustively composed of nonliving parts; it holds, however, that the parts have relations in the living system that are never exhibited in inanimate systems. The third, and least fashionable, answer is "The presence in living systems of a substantial entity that imparts to the system powers possessed by no inanimate body." This is the position of vitalism. It holds, first, that in every living organism there is an entity that is not exhaustively composed of inanimate parts and, second, that the activities characteristic of living organisms are due, in some sense, to the activities of this entity.

THE VITAL ENTITY

The vital entity that animates an organism may, for brevity, be termed its "Life"—a usage that is in fact supported by vitalistic writings. The first thesis of vitalism may be stated as: The Life of an organism is substantial, but it is not—or at least not totally—made up of nonliving substance.

To say that the Life is substantial is to indicate that it has always been conceived more or less closely in accordance with an available doctrine concerning the nature of substance. All vitalists have, for example, held that the Life of an organism is a particular, not a universal; that it is the subject of predicates and not only a predicate; and that it is an agent possessing some degree of autonomy with respect to the body it animates. Most, but not all, vitalists have also maintained that Life, or at least an aspect of it, is capable of existence apart from its organism.

NAÏVE VITALISM

In addition to regarding Life as a substance, all vitalists have adopted a model that helps to specify the sort of substance it is. It may be helpful at this point to distinguish between naive and critical vitalism. Naive vitalism is embedded in common sense in much the same way as a version of mind-body dualism: everyday speech, common maxims, and habitual metaphors all suggest and support it. This type of vitalism, for example, is simply the most direct and literal interpretation of such expressions as "He lost his life," "a lifeless corpse," "A cat has nine lives," and "Scientists will someday create life in the test tube." When the average man thinks about the nature of life at all, he is likely to be guided by these and similar expressions. Naive vitalism has been and indeed still is the popular doctrine. The model of Life adopted by the naive vitalist is the most familiar one available; Life is regarded as a material substance, usually as a fluid body.

In the most primitive forms of vitalism, the Life is flatly identified with a material fluid, the breath, or the blood. This view just misses materialism; it is vitalistic only because the fluid is assigned properties unlike those of any other material body, for example, the power of sensation. Slightly less primitive is the view that Life is a fluid like the blood, only invisible and rather more fiery. The doctrine of the spirits as it occurs in Galen and his successors is an example of this sort of vitalism. The process of etherealizing the Life culminates in the view that it is a fluid but one that is assigned no properties other than its power of animating an organism. This is still a prevalent view and was present, for example, in Mary Shelley's *Frankenstein.*

CRITICAL VITALISM

Although it has conceptual and historical roots in the material substance models, critical vitalism is far more sophisticated. Its various versions have been elaborated by professional philosophers and biologists; indeed, its two outstanding exponents, Aristotle and the twentieth-century biologist and philosopher Hans Driesch, were professionals in both fields. Aristotle's writings, especially his treatises *On the Soul* and *On the Generation of Animals,* are the standard works of vitalistic doctrine. In them Aristotle established four traditions that, it can be said, virtually determined the course of subsequent critical vitalism: he identifies what has been called here the Life of an organism with its psyche; he locates purposive activity, organic unity, and embryological development as the phenomena that vitalism must take most seriously; he argues that the activities of the part must be understood by reference to the form of the whole and that morphogenesis must be understood by reference to the form of the adult; and finally, he describes the manner of the psyche's influence on its organism as formal, not efficient, causation. In short, critical vitalism after Aristotle takes the soul as the model of the Life and attributes to Life the power of achieving and maintaining organic form.

NATURE AND HISTORY. Vitalism was defined above as a metaphysical doctrine in the sense that it is formulated with a degree of vagueness sufficient to exempt it from empirical refutation. However, this is not to say that vitalism has no implications concerning matters of fact. By means of very plausible arguments, vitalists have derived

empirical consequences, some of which have been falsified and some verified. For example, it was argued that since the Life is the blood, a transfusion of blood into a corpse would bring it to life. This experiment failed, but the failure obviously did not refute every version of vitalism or even the doctrine that the Life is the blood. More seriously, Driesch argued that if vitalism is true, then a bit of embryonic tissue that ordinarily develops into a particular organ ought to be capable of developing into other organs. It does happen that some embryonic tissue has this capability. But although Driesch cites such an experiment, he did not actually predict its results. Had they been unfavorable, Driesch would still have had a way to save vitalism. For although he is willing to set limits to the regulative powers of the Life, he gives no antecedent specification of these limits.

In short, vitalism is irrefutable. When this is coupled with the tendency to describe the Life in terms that are among the most problematical in philosophy, it is easy to see that vitalism is subject to the worst aspects of intellectual obscurantism. Its leading exponents, for instance, William Harvey, Georg Stahl, G. L. L. Buffon, Caspar Wolff, J. F. Blumenbach, Lorenz Oken, and K. E. von Baer, represent no improvement upon Aristotle either in the philosophical elaboration of vitalism or in its application to biological phenomena. The long period from Aristotle to Driesch, on the contrary, was characterized by confused invasions of naive vitalism; by the proliferation of such ad hoc entities as life forces, formative impulses, generative fluids, animal heat, and animal electricity; and by the merging of vitalistic thought with other fragments of biological metaphysics, such as the doctrine that living things are arranged along a linear scale corresponding to degrees of perfection (the *scala naturae*), and the archetypal conceptions of organic form. Moreover, vitalism showed a curious tendency to come out on the losing side of biological controversy: After Charles Darwin, it was anti-Darwinian; and it supported the view that organic syntheses could be effected only in a living organism. It also supported the useless and misleading conception of a primordial living substance, the protoplasm, a term and idea that unfortunately still survive.

HANS DRIESCH. After Bergson, Hans Driesch is the best-known twentieth-century vitalist. (Bergson will not be considered here since his biological views are intelligible only as an application of his more general metaphysics.) Driesch's position may be described as Aristotelianism painstakingly applied to modern findings—some of them the result of his own laboratory researches—in physiology and embryology. He also provides three empirical proofs of vitalism.

Driesch defines vitalism as "the theory of the autonomy of the processes of life." It is doubtful that this rules out any biological theories at all, but it does locate Driesch's major concern. He explicitly distinguishes between vitalism and animism, but he does not define *animism.* The term seems to be roughly equivalent to naive vitalism. He also considers vitalistic the view that the parts of an organic system can be understood only by reference to the form of the whole—a view that might preferably be classified as "organismic." But the latter distinction had not been clearly drawn in Driesch's time; he is quite correct in assuming that organismic biology is closer to the vitalistic tradition than, for example, Cartesian mechanism is.

According to Driesch, the Life of an organism is a substantial entity, an entelechy. Driesch employs this term as a mark of respect for Aristotle, although he does not use it with Aristotle's meaning. For Driesch, the entelechy is an autonomous, mindlike, nonspatial entity that exercises control over the course of organic processes; it is not actuality or activity in Aristotle's sense.

Driesch admits that the laws of physics and chemistry apply to organic changes. There is even a sense in which everything that happens in the organism is subject to physicochemical explanation. We may consider, for example, the first division of a fertilized ovum into two blastomeres (daughter cells). Even this relatively simple event can be analyzed as a complex sequence of cooperating chemical syntheses and mechanical movements resulting in, among other things, the duplication of the nucleus, the migration of the daughter nuclei into the opposite sides of the egg, and the formation of a cell membrane between them. Each step in each sequence is a physicochemical event and could be, at least in principle, described and explained as such. But chemistry and physics cannot explain why the steps occur when and where they do. Thus—and on this point some interpretation is necessary—although each event that constitutes first cleavage is physicochemical, it is subject only to post hoc explanation in physicochemical terms. The state of the egg and its environment at time t does not determine what events will begin at later time $t + dt$. But the latter events, after they have occurred, can be exhibited as consequences of events that ended at t. The state of the egg at t determines a range of possibilities; the entelechy influences the course of cleavage, in Driesch's terms, selectively "suspending" and "relaxing the suspension" of these possibilities.

An analogy may shed some light on this doctrine. Suppose that a person's voluntary acts are undetermined, at least at the physicochemical level; that for example, whether or not I clench my fist is not decided by the laws of physics and chemistry. Then the constitution of my body at a given time presents two possibilities, both within my organic capacity: to clench my fist or not. My choice to clench it is analogous to the action of an entelechy. The clenching could not by hypothesis have been predicted on physicochemical grounds, but after its occurrence it can be explained as the outcome of a sequence of physical and chemical events.

Driesch conceives of the laws of nature as placing constraints on the possible activities of a system. For example, the first principle of energetics (thermodynamics) states simply that whatever happens, energy is conserved, but conservation of energy is compatible with any number of actual changes in the system. The entelechy operates in the region of possibilities left open by the operation of laws. Driesch favors a particular metaphor: the entelechy is like an artist who gives form to a material medium, the medium itself both providing possibilities and presenting limitations.

There are, according to Driesch, three "empirical proofs" of vitalism.

(1) In 1888 the German biologist Wilhelm Roux performed the following experiment. Just after the first cleavage of a frog's egg he killed one blastomere with a hot needle. He allowed the other to develop, and it formed a half embryo, resembling a normal embryo that had been cut in two. Roux concluded that the egg is essentially a machine; after cleavage half its parts are in each blastomere.

Driesch performed a similar experiment in 1891 with the eggs of a sea urchin. He separated the blastomeres after first cleavage but found that instead of forming a half embryo, each blastomere developed into a perfect but half-sized larva. This result, Driesch argued, is incompatible with Roux's theory of the successive subdivision of the germ machinery. No machine that could build an organism could possibly build the same organism after it was chopped in two.

Subsequent embryologists have multiplied cases similar to that of Driesch's urchin eggs. Parts of embryos often can generate other than their normal parts. Driesch assigns the term *harmonious equipotential system* to wholes whose parts cooperate in the formation of an organic unity, if the parts themselves also have the potentiality of forming other parts of the unity. The existence of harmonious equipotential systems constitutes the first proof of vitalism.

(2) The formation of a whole sea urchin larva from a single blastomere—one that under ordinary circumstances would form one half of the larva—also provides an illustration of what Driesch calls a "complex equipotential system," that is, a system in which a part, the blastomere, forms a whole, the larva, when it would ordinarily form only a part. The existence of complex equipotential systems provides the second proof.

(3) The third proof is the existence of agency; its paradigm is deliberate human action. The action of an entelechy has been compared to conscious choice, and, indeed, Driesch regards human agency as a special mode of the entelechy's regulation of living processes. But agency characterizes other vital processes as well, especially embryological development. Unfortunately, his definition of agency as "an individual 'answer' to an individual stimulus—founded upon an historical basis" is not made clear.

Vitalism is not a popular theory among biologists, for many reasons apart from its affinity with various lost causes. The successful elucidation of various pieces of biological machinery (for example, the rather successful models of cleavage that at least outline a possible chemical explanation of equipotentiality) have rendered Driesch's first and second proofs rather suspect and, in general, have fostered confidence in the future of nonvitalist theory. There have been numerous philosophical criticisms of vitalism, most of them centering on the rather obvious point that vitalism provides nothing more than pseudoexplanation. The strongest case for vitalism can be summarized as follows: With respect to invulnerability to criticism, vitalism and its most plausible alternatives are in exactly the same position. The various lines of contemporary argument against the possibility of accounting for human agency on an inorganic model lend some support to the vitalist contention that physics and chemistry extend over only some aspects of organic activity.

See also Aristotelianism; Aristotle; Bergson, Henri; Buffon, Georges-Louis Leclerc, Comte de; Darwin, Charles Robert; Driesch, Hans Adolf Eduard; Harvey, William; Materialism; Oken, Lorenz; Organismic Biology; Philosophy of Biology; Stahl, Georg Ernst.

Bibliography

Agar, W. E. *A Contribution to the Theory of Living Organisms,* 2nd ed. Melbourne, 1951.

Aristotle. *The Works of Aristotle Translated into English,* 11 vols. Edited by J. A. Smith and W. D. Ross. Oxford: Clarendon Press, 1908–1952. See *De Generatione Animalium* (Vol. II) and *De Anima* (Vol. III).

Bergson, Henri. *Creative Evolution.* Translated by Arthur Mitchell. New York: Holt, 1911.

Driesch, Hans. *The History and Theory of Vitalism.* London: Macmillian, 1914.

Driesch, Hans. *The Science and Philosophy of the Organism.* London: A. and C. Black, 1908.

Schlick, Moritz. *Philosophy of Nature.* New York: Philosophical Library, 1949. A classic critique of vitalism.

Schubert-Soldern, Rainer. *Mechanism and Vitalism: Philosophical Aspects of Biology.* Translated by C. E. Robin. Notre Dame, IN: University of Notre Dame Press, 1962.

Morton O. Beckner (1967)

VITORIA, FRANCISCO DE

(1492/1493–1546)

Francisco de Vitoria, the political and legal philosopher and theologian, was born in Vitoria, capital of the Basque province of Álava, Spain. While still a boy, he joined the Dominican order in Burgos, and in 1509 or 1510 he was sent to the Collège Saint-Jacques in Paris, where he finished his courses in the humanities and went on to study philosophy and theology. While a student of theology, he directed an edition of the *Secunda Secundae* ("Second Part of the Second Part" of the *Summa*) of St. Thomas Aquinas. The date of his ordination is unknown. From 1516 to 1522 or 1523 he taught theology in the *écoles majeures* of the Collège Saint-Jacques and edited the *Sermones Dominicales* of Peter of Covarrubias, the *Summa Aurea* of St. Antoninus of Florence, and the *Diccionario moral* of Peter Bercherio. He obtained the licentiate and doctorate in theology in 1522. After teaching theology at St. Gregory's monastery in Valladolid from 1523 to 1526, he won by competition the "chair of prime," the most important chair of theology, at the University of Salamanca and held it until his death. Melchior Cano, Mancio, Ledesma, Tudela, Orellana, and Barron, among others, were his disciples. Vitoria helped to formulate the imperial legislation regarding the newly discovered American territories.

With the exception of the prologues to his editions of the works mentioned, Vitoria published nothing during his lifetime. His works include *lecturas* (his class lectures as preserved in the notes taken by his disciples), many of which have been published recently; *relectiones* (extraordinary lectures, which are summaries or popularizations of his ordinary lectures), published for the first time in 1557; and several writings on different topics. Vitoria is famous chiefly for his *relectiones*, the most important of which are *De Potestate Civili, De Potestate Ecclesiae Prior, De Potestate Ecclesiae Posterior, De Potestate Papae et Concilii*, and, particularly, *De Indis* and *De Iure Belli*.

According to Vitoria, political society (*respublica*) is a perfect, self-sufficient society, a moral and juridical person. It is a natural, not a conventional, society. In other words, it is required by nature and has its end set by nature. Actual states are the result of positive human acts, but men are obliged by natural law to live in some form of political society, outside of which no good or full human life is possible. The end of society is twofold: to promote the common good and virtuous life of its citizens and to protect their rights. The proximate origin of political society is the will of families. Authority is an essential property of the state, for without it the organic unity of the citizens and their activity, necessary for the attainment of the common temporal good, would be impossible. Like every natural right, authority derives ultimately from nature's author and resides originally in the body politic. However, since political society is incapable of exercising public authority directly, it must transfer it to one or several rulers. Particular forms of government depend on the will of the citizens. The absolutely best form is monarchy, "for the whole world is most wisely ruled by one Prince and Lord." The reason behind this claim is that monarchy, better than any other form, creates and preserves the necessary unity of social action without unduly curtailing the citizen's freedom; "freedom in monarchy," Vitoria remarked, "is no less than in democracy, wherein discussions and seditions, inimical to liberty, are the unavoidable result of the participation of many in government."

Beyond individual states there is a larger society, the international society constituted by the whole human family. It, too, is natural and necessary, although less strictly so, for the satisfaction of man's needs and the development and perfection of his faculties. International society possesses its own authority, which is immanent in the whole of humankind. From this universal authority derive the laws that establish the rights and correlative duties of the different states. The sum of these laws forms the *ius gentium*, which is partly made up of conclusions drawn from the principles of natural law by natural reason and partly of positive customs and treaties among nations. Vitoria established the chief rights of every nation, whether great or small, as the right to existence; the right to juridical equality; the right to independence (except where a nation is juridically and politically so

immature as to be incapable of self-rule, in which case a more civilized nation may temporarily administer it under mandate or keep it in trusteeship); the right to free communication and trade, denial of which by another nation could justify war; and the right—and the duty—of every state to intervene in defense of nations victimized by domestic tyrants or threatened or attacked by stronger nations.

War is licit as a last resort, according to Vitoria, when all other means of persuasion have failed. The cause that justifies a war, whether defensive or offensive, is the violation of a right. An essential condition for the licitness of a war is that the evils resulting from it will not be greater than the good intended. Defensive war can be justly undertaken by any person; offensive war can be launched only by public authority. The ruler waging a just war is invested with power by human society. Just as the state has the power to punish criminals among its citizens, so humankind has the power to punish a nation guilty of injustice. All means necessary for the attainment of victory are permissible in a just war. Once victory is achieved, the conquering nation should exercise its rights over the conquered with moderation and Christian charity.

The thesis that Vitoria was the founder of modern international law has been definitively established by numerous scholars. It was officially acknowledged in 1926, when the Dutch Association of Grotius gave the University of Salamanca a gold medal coined to honor Vitoria as the founder of international law. Also in 1926 the Asociación Francisco de Vitoria was founded in Spain for the purpose of studying and spreading Vitoria's ideas through publications, conferences, and special courses at the University of Salamanca.

See also Authority; Peace, War, and Philosophy; Philosophy of Law, History of; Thomas Aquinas, St.

Bibliography

WORKS BY VITORIA

Commentarios á la "Secunda Secundae" de Santo Tomás, 5 vols. Edited by Vicente Beltrán de Heredia. Salamanca, 1932–1935.

Relectiones Theologicae, 3 vols. Edited and translated into Spanish by Luis G. Alonso Getino. Madrid, 1933–1936. There is an English translation and photographic reproduction of the 1696 edition of *Relectiones* in the series Classics of International Law. Washington, DC, 1917.

WORKS ON VITORIA

Beltrán de Heredia, Vicente. *Francisco de Vitoria*. Barcelona: Editorial Labor, 1939.

Beltrán de Heredia, Vicente. "Vitoria (François de)." In *Dictionnaire de théologie catholique*. Vol. XV, cols. 3117–3133.

Brady, Bernard V. "An Analysis of the Use of Rights Language in Pre-Modern Catholic Social Thought." *Thomist* 57 (1) (1993): 97–121.

Davis, G. Scott. "Conscience and Conquest: Francisco De Vitoria on Justice in the New World." *Modern Theology* 13 (4) (1997): 475–500.

Doyle, John P. "Vitoria on Choosing to Replace a King." In *Hispanic Philosophy in the Age of Discovery: Studies in Philosophy and the History of Philosophy,* Vol. 29, edited by Kevin White. Washington DC: Catholic University American Press, 1997.

Getino, Luis G. Alonso. *El maestro fray Francisco de Vitoria: su vida, su doctrina y su influencia*. Madrid, 1930.

Giacon, Carlo. *La seconda scolastica*, Vol. I: *I grandi commentatori di San Tommaso*, 163–213. Milan: Fratelli Bocca, 1944.

Hartigan, Richard S. "Francesco De Vitoria and Civilian Immunity." *Political Theory* 1 (1973): 79–91.

Morris, John F. "The Contribution of Francisco De Vitoria to the Scholastic Understanding of the Principle of the Common Good." *Modern Schoolman* 78 (1) (2000): 9–33.

Naszályi, Aemilius. *Doctrina Francisci de Vitoria de Statu*. Rome, 1937.

Reichberg, Gregory M. "Francisco De Vitoria, De Indis and De Iure Belli Relectiones (1557): Philosophy Meets War." In *The Classics of Western Philosophy: A Reader's Guide*, edited by Jorge J. E. Gracia. Malden, MA: Blackwell Publishing, 2003.

Sanchez Sorondo, Marcelo. "Vitoria: The Original Philosopher of Rights." In *Hispanic Philosophy in the Age of Discovery: Studies in Philosophy and the History of Philosophy,* Vol. 29, edited by Kevin White. Washington DC: Catholic University American Press, 1997.

Scott, James Brown. *The Spanish Origin of International Law*, Vol. I: *Francisco de Vitoria and His Law of Nations*. Oxford: Clarendon Press, 1934.

Soder, Josef. *Die Idee der Völkergemeinschaft. F. de Vitoria und die philosophischen Grundlagen des Völkerrechts*. Frankfurt am Main: Metzner, 1955.

Solana, Marcial. *Historia de la filosofía española*, 3 vols., Vol. III, 43–91. Madrid: Real Academia de Ciencias Exactas, Físicas y Naturales, 1941.

Truyol y Serra, Antonio. *Los principios del derecho público en Francisco de Vitoria*. Madrid, 1946.

Felix Alluntis, O.F.M. (1967)
Bibliography updated by Tamra Frei (2005)

VIVES, JUAN LUIS
(1492–1540)

Juan Luis Vives, the Spanish humanist, was born in Valencia and died in Bruges. Considerably younger than such scholars as Desiderius Erasmus, Guillaume Budé, and John Colet, Vives deserves an honorable place among them for his moral seriousness, sincerity of religious

belief, promotion of education, and social concern, as manifested in projects for the promotion of peace and the relief of the poor. In many of these respects Vives is approached only by his nearer contemporary, Thomas More; his character emerges very favorably from any comparison with the earlier group. His efforts to secure patronage from the nobility did not blind him to the plight of those more needy than he, nor did he engage in the acrimonious personal quarrels that marred the character of some humanists.

Vives was a fine scholar and an excellent writer. After initial schooling in Spain he went to Paris to attend the university. Here he found still active a school of terminist logicians and physicists whose influence extended, so Vives tells us, to all the higher faculties. The earlier Oxford and Paris developments in logic and physics were being studied by teachers under the influence of the Scottish philosopher and theologian John Major. But the new learning was gaining favor, and there were signs among both students and teachers of dissatisfaction with the nominalist approach. Two of Vives's own teachers, Gaspar Lax and John Dullaert, told him that they were sorry that they had wasted so much time on "useless little questions." The "little questions" concerned such issues as the logical analysis of signification and of inference, as well as the quantification of physical phenomena. The complaint voiced by Vives and by many other humanists concerned not so much the intrinsic value of these discussions as the fact that they were permitted to invade all other fields of learning, often to the exclusion of the proper subject matter. Vives particularly disliked the petty vindictiveness and personal egoism displayed by younger men who delighted in scoring points over older opponents. When Vives returned to the University of Paris after his sojourn at Louvain, he expected to meet with a cool reception because of his book *Adversus Pseudodialecticos* (Against the pseudo dialecticians; 1520), in which he sharply criticized the academic climate at the university. To his surprise, he was warmly received, as he told Erasmus in a letter of 1520, and was assured that terminist quibbling was no longer tolerated in nonlogical discussions.

Vives's criticism of school philosophy was one of the more moderate and informed humanist attacks. He held Aristotle and the other ancients in high regard but deplored the failure of their followers to observe nature afresh. Vives condemned the undue humility of those who claimed to be only "dwarfs, standing on the shoulders of giants": If we cannot see farther than our predecessors, he insists, it is not because we are dwarfs and they giants but because we are lying prostrate on the ground, having given up the search for the truth. Vives insists as strongly as did Lorenzo Valla that philosophical terminology should not be artificial; the usage of such ancient writers as Cicero and Seneca should be taken as models. Philosophers should not depart too far from the speech of the people. Vives admitted, however, that it may occasionally be necessary for philosophers to coin terms of their own as well as to clarify those in ordinary usage.

Vives's own philosophy may be characterized as Augustinian in its general outlines, with eternal salvation and the vision of God overriding lesser concerns. It is in the light of this general orientation that his much discussed "empiricism" must be evaluated. Of all things on Earth, it is man's own soul that it most behooves him to know, by means of direct observation. But undue curiosity concerning other things, especially concerning their "inner natures and causes," is out of place and, indeed, impious. To inquire too curiously into the elements, the forms of living beings, or the number, magnitude, disposition, and powers of natural objects is to "tear the seventh veil." Such an attitude is certainly not favorable to purely theoretical scientific inquiry. But Vives's central concern is with man's felicity, and only to the extent that inquiry into nature serves to promote man's felicity is it admissible as part of the curriculum of studies. This curriculum would stress the useful arts, to the analysis of which Vives devoted great attention. In common with humanists in general, Vives stressed the utility of the arts and insisted that they must be systematized or brought into rules and precepts so as to be applicable to the purposes of ordinary life. Inordinate attention to their logical analysis must be curtailed; instead, students are to be constantly reminded of the empirical origins of useful knowledge. In his discussion of method in the arts, Vives explicitly drew on Galen as well as on suggestions in Aristotle.

Neither history nor theology is an art from this standpoint, since neither subject has been reduced to rules. Vives was impatient with the school theology of his time; he found little of value in the controversies between Scotists and Thomists and disliked their fanaticism: "They would accuse each other of heresy if it were not for the mellowing effect of the customs of the school." It has been aptly remarked that Vives's religious thought has close affinities with northern Pietism as exemplified by the Brethren of the Common Life, the movement that left such an impression on Erasmus. In keeping with this is Vives's obvious sympathy for the common people, a note conspicuously absent from the writings of many other humanists.

On a few points Vives specifically rejected Platonism—for example, in maintaining that God does not require divine Ideas and that we do not have reminiscences of Ideas from our past lives. Vives prefers to explain the insights of Plato's doctrine of reminiscence by means of certain natural relationships between the human mind and "those first true seeds of knowledge whence all the rest of our knowledge springs," called anticipations by the Stoics. This Stoic doctrine merges easily in Vives's thought, as in that of many of his contemporaries, with an appeal to common sense (*sensus communis*), which here takes on its modern flavor. Common sense furnishes us with an argument for God's existence, there being no people so benighted as to be completely destitute of some knowledge, however dim, of God. Human minds, furthermore, are all informed with the need to worship God, but what form this worship takes is a matter of human persuasion. Here we may trace the influence of Florentine Platonism, with which Vives was quite familiar. Perhaps from the same source is Vives's often repeated assertion that nothing would be more wretched than man if his actions aimed only at earthly ends. He condemns the vices of pleasure (*voluptas*) and pride (*superbia*) as roundly as any other medieval writer. Pride is responsible for the "frenzied craving for knowledge" shown by some men who are anxious to appear distinguished among their fellow men. Only piety, however, can permanently satisfy man and give him rest.

See also Aristotle; Cicero, Marcus Tullius; Colet, John; Erasmus, Desiderius; Galen; Humanism; Major, John; More, Thomas; Pietism; Platonism and the Platonic Tradition; Scotism; Seneca, Lucius Annaeus; Stoicism; Thomism.

Bibliography

Vives's writings, with the exception of his commentaries on Augustine's *City of God*, were collected by Mayáns y Ciscar, 8 vols. (Valencia, 1782). The biography by Adolfo Bonilla y San Martín, *Luis Vives y la filosofía del renacimiento* (Madrid, 1903), is reliable but not too detailed. See also *Vives: On Education, a translation of the De Tradendis Disciplinis*, translated by Foster Watson (Cambridge, U.K., 1913), and Vives's *De Anima et Vita*, edited by M. Sancipriano (Florence, 1954).

OTHER RECOMMENDED WORKS

Hidalgo Serna, Emilio. "'Ingenium' and Rhetoric in the Work of Vives." *Philosophy and Rhetoric* 16 (1983): 228–241.

Noreña, Carlos G. *Juan Luis Vives and the Emotions*. Carbondale: Southern Illinois University Press, 1989.

Noreña, Carlos G. *Studies in Spanish Renaissance Thought*. The Hague: Martinus Nijhoff, 1975.

Noreña, Carlos. "Was Juan Luis Vives a Disciple of Erasmus?" *Journal of the History of Philosophy* 7 (1969): 263–272.

Neal W. Gilbert (1967)

Bibliography updated by Tamra Frei (2005)

VLASTOS, GREGORY

(1907–1991)

Gregory Vlastos led a revival of interest in ancient philosophy and was the first American scholar to deploy the methods of analytic philosophy in this area. Best known for his work on the philosophy of Socrates, he also published widely on Plato and on topics in pre-Socratic philosophy. Before turning to ancient philosophy, he published works in social and political theory, and his writings on justice continue to be influential.

He was born in the Greek community of Istanbul, raised as a Protestant, and educated at Roberts College (an American-sponsored institution of secondary and higher education in Istanbul). He took a bachelor of divinity degree in 1929 from the Chicago Theological Seminary and proceeded to Harvard University, where, after studying philosophy under Raphael Demos and Alfred North Whitehead, he was awarded his PhD in 1931. In that year he took a position at the Queen's University in Kingston, Ontario. He served in the Canadian Air Force during World War II. In 1948, he joined the Sage School of Philosophy at Cornell. In 1954–1955, he was a member at the Institute for Advanced Study in Princeton and, in 1955, joined the Department of Philosophy at Princeton University, which he served for many years as Stuart Professor and then chairman. He was president of the Eastern Division of the American Philosophical Association in 1965–1966. In 1976, he moved to the University of California at Berkeley, where he remained until his death.

Vlastos had a huge influence on the next generation of scholars of ancient philosophy, which has been led in the United States largely by his students, proteges, and members of the seminars he conducted for young college teachers. Many of these became highly distinguished: Richard Kraut, Terence Irwin, A. P. D. Mourelatos, Alexander Nehamas, Gerasimos Santas, and Nicholas Smith, to name a few.

Vlastos began the revolution in Platonic studies with his article, "The Third Man Argument in the *Parmenides*" (1954), which rendered the argument in formal terms and ignited a debate (joined by such notable philosophers as Peter Geach and Wilfrid Sellars) over both the sound-

ness of the argument and its purpose. Vlastos concluded that the argument revealed Plato's "honest perplexity" about the theory of forms. Vlastos held a developmental view of Plato: Early dialogues (those with affinities to the *Apology*) were mainly innocent of metaphysics, middle dialogues (such as the *Republic*) were committed to a theory of the degrees of reality, and later dialogues showed Plato to be critical toward his former metaphysical theories.

On the theory of forms in Plato, Vlastos wrote a number of important papers, of which "Degrees of Reality in Plato" (1965) is the most famous. He explained, in a way that has been the basis for most subsequent work in this area, what Plato could mean by saying that a form was more real than its sensible instances: The form is cognitively more dependable.

Vlastos brought attention to Plato's writings about love and friendship, raising the question whether an individual person could be an object of love on the Platonic theory, which seems to place the Form of Beauty itself at the apex of love. Vlastos saw that Plato represented Socrates as a teacher who failed more often than he succeeded, and, in a famous essay, he attributed Socrates' failure to an inability to respond to his students with love ("The Paradox of Socrates," 1971).

Drawn early to Socrates' single-minded devotion to the care of the soul, Vlastos brought out the problem in Socrates' doctrine of the unity of virtue: Why, if they are one, do they have different definitions? His solution was that the virtues are not strictly identical, but biconditionally related in such a way that having any virtue implies having the others.

During his Berkeley period he generated his most influential work—a set of articles and a book about the Platonic Socrates that defined the subject for the next generation of scholars. He established a method for identifying the philosophy of Socrates, taking Plato's works to reflect the philosophy of Socrates insofar as they are compatible with Plato's *Apology of Socrates*, which he supposed to be an adequate historical guide on philosophical points.

In one of his most influential pieces, "The Socratic Elenchus" (1983), Vlastos identified the method Socrates uses in certain early dialogues as elenchus (a kind of cross-examination), about which Vlastos asked the question that has been fundamental to subsequent research. Socrates, he pointed out, depends on the elenchus for both negative conclusions, refuting the bogus knowledge-claims of others, and, for positive results, supporting his own ethical views. Yet the method seems to have no foundation aside from the assent of Socrates' interlocutors. Vlastos suggested that the method winnows out the interlocutors' false views, leaving ones that are likely to be true, thus providing credibility for those views that fall short of certainty, but nevertheless provide practical grounds for Socrates' moral teaching. Socrates' disclaimer of knowledge was not a lie, as many believed in antiquity, but a case of what Vlastos called "complex irony": the complex truth behind it is that Socrates lacks certainty, while maintaining what Vlastos called "elenctic knowledge," knowledge supported by the elenchus. In this way Vlastos introduced a new understanding of Socratic irony, which was to give a title to his last book.

Just before his death, in *Socrates, Ironist and Moral Philosopher* (1991), Vlastos brought together his conclusions about Socrates, of which the most important was that Socrates was a trend-setting innovator in moral theory, as "the first to establish the eudaimonist foundation of ethical theory," and, moreover, "the founder of the non-instrumentalist form of eudaimonism held in common by … all Greek moral philosophers except the Epicureans" (1991, p. 10). Even more revolutionary, according to Vlastos, Socrates rejected the traditional morality of retaliation, the idea that justice requires people to harm their enemies.

Vlastos had a gift for identifying questions of interpretation that drew other philosophers into discussion, both of his proposed answers and of the questions themselves. He never ceased to express a love for his subject that was infectious and has been passed down to subsequent generations of scholars. Whether or not the answers he gave will survive the test of scholarly debate, his questions will continue to define that debate.

See also Eudaimonia; Justice; Plato; Socrates.

Bibliography

WORKS BY GREGORY VLASTOS

"The Third Man Argument in the *Parmenides*." *Philosophical Review* 63 (1954): 319–349.

"The Paradox of Socrates." In *The Philosophy of Socrates: A Collection of Critical Essays*. Garden City: Anchor Books, 1971, pp 1–21.

Platonic Studies. 2nd ed. Princeton, NJ: Princeton University Press, 1981. Contains, among other essays, these that are cited in this article: "Degrees of Reality in Plato" (1965), "The Individual as Object of Love in Plato" (1969), and "The Unity of the Virtues in the *Protagoras*" (1971).

"The Socratic Elenchus." *Oxford Studies in Ancient Philosophy* 1 (1983): 27–58.

"Socrates' Disavowal of Knowledge." *Philosophical Quarterly* 35 (1985): 1–31.

Socrates: Ironist and Moral Philosopher. Ithaca, NY: Cornell University Press, 1991.

Socratic Studies, edited by Myles Burnyeat. Cambridge, U.K.: Cambridge University Press, 1994. Contains, among other essays, these that are cited in this article: "The Socratic Elenchus: Method is All" (revised from the 1983 paper), and "Socrates Disavowal of Knowledge" (revised from the 1985 paper).

Studies in Greek Philosophy. 2 vols, edited by Daniel W. Graham. Princeton, NJ: Princeton University Press, 1995.

WORKS ABOUT GREGORY VLASTOS

Annas, Julia. *The Morality of Happiness*. New York: Oxford University Press, 1993.

Irwin, Terence. *Plato's Ethics*. New York: Oxford University Press, 1995.

Paul Woodruff (2005)

VOID

See *Quantum Mechanics*

VOLITION

The action of opening a door by pushing on it is composed of the agent's action of voluntarily exerting force with his or her arm and hand plus that action's causing the door to open. Is the voluntary exertion of arm and hand similarly composed of an action producing a result? There is a clear candidate here for the role of result—namely, the limb's exerting force. It could have exerted exactly the same force, by means of just the same muscle contractions, without the agent's voluntarily exerting the force with it. So the exerting of force by the limb is only a part of the whole action. But does the remainder consist of this part's being caused by action of the agent? Philosophers disagree on the answer to this question. Section I below offers one way of spelling out an affirmative answer (which is developed more fully in Ginet [1990, ch. 2]). Section II briefly sketches some alternative views.

SECTION I

When one voluntarily exerts force with a limb, the action that causes the limb to exert force is a mental action, which, following an old tradition in philosophy and psychology, is called *volition*. We view such exertions as voluntary because we experience them as directly under our control. This is most clear in those cases of voluntary exertion where we have to concentrate on what we are doing with the body—such as my experience of trying an unfamiliar dance movement with my left leg. Here my attention is focused on my exertion with the leg. I note just how I am trying to exert it and just how the exertion feels. This contrasts with my moving my left leg in the course of walking along enjoying a fine day, where I do not attend at all to my exertion with the leg. I do it, as we say, "automatically," perhaps without even noting that I am now exerting that leg. But the difference is between these cases should not be exaggerated. It is not at all like the difference between one of the foregoing experiences and an exertion of my body that I experience as purely involuntary, such as the movement of my lower leg in response to a sharp tap just below the kneecap. In this last case, though I experience the leg's exertion, I do not experience it as something that I voluntarily determine. But my experience of voluntary exertions, even when it is most nonattentive, is colored with the sense of my making them happen.

I experience my voluntary acts as the specific exertions they are—at least in those respects that I voluntarily determine. If in walking I had made an appreciably different movement with my leg at one point than the one I actually made—taking a much longer step, say, than the one I actually took—my experience of making the movement would have been correspondingly different, whether or not I was attending to the experience.

The normal subjective experience of voluntarily exerting the body in a certain way is a compound of two significantly different parts. There is, first, a perceptual aspect. One perceives the exertion in a certain direct way, not visually or by feeling it with some other part of one's body. But the experience of voluntary exertion is more than the direct perception of the exertion. I could feel my arm exerting force in just the same way it does when I thrust it upward without experiencing this exertion as something I make happen. I could experience it as something that just happens to me, unconnected with my will, while at the same time perceiving the exertion of the arm as just like one I might have produced voluntarily. The voluntariness of the experience of voluntary exerting is a further part of it, distinct from the perceptual part, an aspect that would be more conspicuous by its absence than it is by its presence.

It is this nonperceptual part of the experience that is volition. This part could occur all by itself, unaccompanied by perception of exertion. It could seem to me that I voluntarily exert a force upward with my arm without at the same time having the sense that I feel the exertion happening. The arm feels paralyzed and anaesthetized.

Neither sort of impoverished experience—seeming to feel an exertion without seeming voluntarily to make it or seeming voluntarily to make it without seeming to feel it—happens very often. But both do in fact occasionally occur. And we know enough about how our experience depends on what happens in our neural system to know how it is possible in principle to produce either sort. Seeming to make an exertion without seeming to feel it could be produced by depriving a subject of the input neural capacity to perceive the exertion while leaving unimpaired their output neural capacity to make the exertion. And we could produce the experience of seeming to feel a given sort of exertion while lacking the sense that it is voluntary by giving to the perceptual system of a subject who is not trying to make any exertion the same neural input that causes a subject to feel that sort of exertion when he or she makes it voluntarily.

The mental action of volition is not an antecedent of the experience of voluntarily exertion, not a prior mental occurrence that triggers the whole package of the exertion and the experience of it. Rather it is that *part* of the experience whose presence is what makes the exertion seem voluntary and whose absence would make it seem involuntary.

Volition is the means by which I cause my body's exertion when I voluntarily exert it. For my volition counts as my trying to exert it—that is, as my trying to cause it to exert. So when I succeed, it is by this trying, this volition, that I cause it.

Volition resembles certain other mental actions (such as deciding) in having intentional content. The volition involved in my voluntarily exerting a certain force with my arm is volition to exert that sort of force. Its being a volition to exert a certain force with my arm is not a matter of what it causes but an intrinsic property of the mental act itself, in the same way that it is an intrinsic property of a certain act of deciding that it is a deciding to raise my arm.

Volition is an intentional mental occurrence whose content (or object) does not go beyond exerting force with one's body in the immediate present. Occurrent intention and occurrent desire are other sorts of intentional mental occurrences whose contents are not so restricted. Volition to exert in a certain way is not a kind of occurrent desire to exert in that way. For one thing, volition is action and not desire; not even occurrent desire is action. For another thing, it is possible to have volition to exert a certain way without at the time in any way desiring or intending to exert in that way. This would happen, for example, if I were sure that my arm is paralyzed and tried to exert it just to see what it is like to experience inefficacious volition. If I were mistaken about my arm's being paralyzed, I would exert it voluntarily but not intentionally. This shows also that volition to exert in a certain way is not a kind of decision or intention to exert in that way.

Volition differs from deciding also in not being a single-shot mental act with a static content. Volition is a fluid mental activity whose content is continually changing. At each moment, it is concerned only with bodily exertion in the immediate present. I can all at one time decide to swim another length of the pool, but I cannot all at one time have the volition to make the whole sequence of bodily exertions involved in turning a doorknob and pulling the door open, any more than I can perform that sequence of exertions all at one time. Volition is part of the experience of voluntary exertion and its content, unlike the content of a decision or intention, is as much tied to the immediate present as is voluntary exertion itself.

As we approach an instant, the content of volitional activity approaches an unchanging, frozen proposition about the immediate present. What I will at a particular moment is to exert at that moment a determinate degree of force in a determinate direction with one or more parts of my body. I do not will to *move* my body. The content of volition at a moment is not concerned with movement, which takes time, but only with exertion of directed force at that moment. Temporally extended movements are the objects of intentions rather than volitions. Volitions do not plan ahead, not even a little bit. Volitions do not *plan* at all. They *execute* (or try to execute). I have an intention as to what course of movement my body is to take over the next few moments, and in light of that intention I go through a certain course of volitional activity over the period of the movement, willing at each point, in light of my perceptions, the directed force needed at that point to keep the movement on the path prescribed by my intention. Volition is analogous to steering with a steering wheel rather than to steering with buttons that trigger preset patterns of movement. If there are mental triggers of sequences of voluntary exertion (as there may be in familiar, practiced movements), the volitional activity is not the trigger but rather part of what is triggered.

When I exert voluntarily, my volition is not just that my body exert but that I exert with my body. I will not just exerting but exerting caused by me. I will that my willing—this very volition of whose content we speak—cause the exertion. The content must refer to the volition

of which it is the content and say that this volition is to cause the body to exert in a certain way.

The content of my volition at an instant could be expressed by me in a proposition of the following form: "I will that this willing cause my bodily part B to exert force of degree F in direction D." Here F is a certain range of degrees of force. and D is a certain range of directions. What I will is never absolutely precise with respect to the degree or direction of the force. When I begin to move a lever, the degree and direction of the force exerted by my arm, as measurable by a precision instrument, could vary within certain limits and still fit the content of my volition. Gaining more finely tuned control of one's body is at least partly a matter of becoming able to will contents that are more determinate.

SECTION II

Several philosophers have put forward accounts of voluntary bodily action that incorporate something like volition but differ from the foregoing account of it in one way or another. Hugh McCann (1972, 1974, 1976 [all reprinted in McCann 1998]) presents an account that is nearly the same as the foregoing one. One minor difference is that on McCann's view, volition (willing) to exert entails intending to exert. John Searle (1983) gives to something he calls *intention in action* a role similar to the one given volition in the foregoing account in that it is the initial part, rather than a cause or accompaniment, of an action. But it differs in that an intention is not an action, whereas a volition is. Alvin Goldman (1976) gives the name "volition" to a certain kind of *occurrent desire*, but an occurrent desire is also not an action. Wilfrid Sellars (1976) gives the name to an *occurrent intention* or *decision* to act in a certain way; a decision is, like volition, a mental action, but a decision is intrinsically an intention to exert the body in a certain way, whereas a volition is not.

Larry Davis (1979) uses "volition" to name not a conscious mental activity of which we are directly aware but a functionally defined subconscious mental process that is not part of our experience but is posited by theory as that which causes the bodily exertion and the agent's belief that he or she is acting. Frederick Adams and Alfred Mele hold that "the major functional roles ascribed to volition are nicely filled by a triad composed of intention, trying, and information feedback" (1992, p. 323). Trying to A, on their account, "is an event or process that has A-ing as a goal and is initiated and (normally) sustained by a pertinent intention. Successful tryings to A, rather than causing A-ings, are A-ings." So, on their view, in one's voluntary exertion with one's limb, the trying to exert that is involved is to be identified not with a mental action that causes the exertion but with the whole voluntary exertion. There is no mental part of the action that causes the rest. Mele does hold (2002) that any action must have a proximal mental cause—namely, an intention to act straightaway.

According to Timothy O'Connor (2000), an action of a person involves *agent causation*. The mark of an action is that the agent, the enduring entity that is the person, and not any mental or other event causes the event parts of action. The initial event the agent causes in voluntarily exerting in a certain way could, on this view, be volition as characterized in section I above, but O'Connor himself takes it to be an "executive state of intention" to act in that way (p. 72).

The *tryings* of Jennifer Hornsby (1980) are mental actions and, in her account of action, play a role in causing bodily events analogous to that played by volition in the foregoing account. But on her account the momentary content of a trying can specify a temporally extended sequence of bodily exertion and even external consequences of these (for example, the content can be to open a door). This and the fact that for her a trying implies intending or desiring the content of the trying make her tryings significantly different from the volitions described section I.

See also Determinism, A Historical Survey; Goldman, Alvin; Searle, John; Sellars, Wilfrid.

Bibliography

Adams, Frederick, and Alfred R. Mele. "The Intention/Volition Debate." *Canadian Journal of Philosophy* 22 (1992): 323–338.

Davis, Lawrence. *Theory of Action*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

Ginet, Carl. *On Action*. New York: Cambridge University Press, 1990.

Goldman, Alvin. "The Volitional Theory Revisited." In *Action Theory*, edited by Myles Brand and Douglas Walton. Dordrecht: D. Reidel, 1976.

Hornsby, Jennifer. *Actions*. London: Routledge & Kegan Paul, 1980.

James, William. "The Will." Chap. 27 in *Principles of Psychology*. 2 vols. Mineola, NY: Dover Publications, 1950.

Libet, Benjamin, Anthony Freeman, and Keith Sutherland, eds. *The Volitional Brain: Towards a Neuroscience of Free Will.* Imprint Academic, 1999.

McCann, Hugh. "Is Raising One's Arm a Basic Action?" *Journal of Philosophy* 68 (1972): 235–250.

McCann, Hugh. "Trying, Paralysis, and Volition." *Review of Metaphysics* 28 (1976): 423–442.

McCann, Hugh. "Volition and Basic Action." *Philosophical Review* 83 (1974): 451–473.

McCann, Hugh. *The Works of Agency*. Ithaca, NY: Cornell University Press, 1998.
Mele, Alfred. *Motivation and Agency*. New York: Oxford University Press, 2003.
O'Connor, Timothy. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press, 2000.
Searle, John. *Intentionality*. Cambridge, U.K.: Cambridge University Press, 1983.
Sellars, Wilfrid. "Volitions Re-Affirmed." In *Action Theory*, edited by Myles Brand and Douglas Walton. Dordrecht: D. Reidel, 1976.

Carl Ginet (2005)

VOLNEY, CONSTANTIN-FRANÇOIS DE CHASSEBOEUF, COMTE DE *(1757–1820)*

Constantin-François de Chasseboeuf, comte de Volney, the French *philosophe* and historian, was born in Anjou. He early showed a scholarly disposition, and at fifteen he asked for Hebrew lessons in order to verify translations of the Bible. Inheriting independent wealth, he left for Paris at seventeen, turned down his father's plea to study law, and, interested in the relation between the moral and the physical aspects of man, chose medicine instead. He also pursued his study of history and languages, and he became involved in the polemics and ideological struggles of the time. In 1783 he gave himself the name Volney and left for Egypt and Syria "to acquire new knowledge and embellish the rest of my life by an aura of respect and esteem." After eight months in a Coptic monastery, devoted to mastering Arabic, he spent three and a half years traveling on foot throughout Egypt and Syria. The resulting *Voyage en Égypte et Syrie* (1787) is his most enduring production. A remarkable travel book, it differs from those of the romantic travelers (such as François René de Chateaubriand) by its impersonality and its careful, objective account of physical, political, and moral conditions. It was used as a guide by Napoleon Bonaparte's armies.

After his return to France, his prestige assured, he was placed in charge of commercial relations with Corsica and, on the outbreak of the Revolution, was elected a representative of the third estate. His revolutionary career was quite distinguished; he defended civil rights and freedoms, attacked the church strongly, and later opposed the excesses of the Jacobins. In 1792 he bought land in Corsica and showed how products of the New World could be successfully transplanted. There he met and became friendly with Napoleon, whose greatness he foresaw. Forced to leave because of unrest in Corsica, he subsequently spent ten months in prison, falsely accused of being a royalist, until he was released after the ninth of Thermidor. Appointed professor of history in the new École Normale, he developed a critical methodology for historical investigation. When that institution was suppressed in 1795, he went to the United States. Well received by George Washington, he was happy at first. John Adams, however, was unforgiving of Volney's severe criticisms of his political writings, and he felt an animosity toward the French as a result of the XYZ Affair. In addition, a theological quarrel with Joseph Priestley, who was then in America, did not dispose Adams favorably toward visiting philosophers. Accused of being a secret agent, Volney was forced to leave America in 1798, but by then he had traveled all over the country. In 1803 he published *Tableau du climat et du sol des États-Unis d'Amérique*, an objective description famous for its picture of Niagara Falls; in the preface he told of his persecutions.

Back in France, Volney cooperated in Napoleon's coup of the 18th Brumaire and was named senator. However, he frequently opposed Napoleon's dictatorial tendencies, and he also opposed the Concordat of 1801. Napoleon ridiculed him along with his whole group of *idéologues* (including Pierre Cabanis and Comte Antoine Destutt de Tracy), but he later made Volney a count. Volney, however, supported the Restoration and was rewarded with a peerage. Volney was known for his independence and for his ill-tempered, overbearing character.

WORKS

Volney's most famous work is *Les ruines, ou Méditations sur les révolutions des empires* (1791), a work conceived in Benjamin Franklin's study in Paris. Widely read and admired during his lifetime and later, it now seems a shallow piece of rhetoric. It was much read in English, under the title *The Ruins of Empires* (1792). The author contemplates the ruins of Palmyra and wonders how powerful empires, seemingly destined to last forever, succumbed to the universal law of change and destruction. A belated example of "philosophic" polemics, *Les ruines* promoted deism by a comparative study of religious doctrines and practices, preached tolerance and free inquiry, the unalienable rights of men and peoples, and the right of self-government. Some ethical ideas were sketched, which Volney developed in *La loi naturelle*. Thus, man in the state of nature "did not see at his side beings descended from the heavens to inform him of his needs which he

owes only to his senses, to instruct him of duties which are born solely of his needs."

Even more interesting as a reflection of moderate views held by *philosophes* at the end of the century is Volney's *La loi naturelle, ou Catéchisme du citoyen français* (1792). In this work he affirmed a natural law given by God, but this natural law is essentially physical ("the regular and constant order by which God rules the universe"). The moral aspect of natural law is only an extension of the biological requirement for self-preservation and "perfection" on the part of the individual and the species. Consequently, morals could become an exact science. In this work, as in *Les ruines*, Volney praised the harmony and order of relationships in the universe, declaring that man is no exception to their rule; yet within this impersonal natural law he discerned purpose and final causes, namely, the happiness and perfection of the individual. Physical suffering has a useful natural function, and the advantage of greater sensitivity in man is compensated by the disadvantage of greater suffering. Law is a command (or prohibition) followed by reward or punishment. Moral law depends on general and constant rules of conduct that inhere in the order of things. Moral law is not obvious; rather, it forms "in its developments and consequences, a complex ensemble that requires the knowledge of many facts and all the sagacity of reasoning." The basic principle of natural law is self-preservation, not happiness, which is "an article of luxury." Pleasure and pain are the mechanisms by which natural law works. Men are aware of these laws only in society. Life in society is man's true natural state, since it is necessary for his self-preservation; in what is called the state of nature, man was only a miserable brute. Volney's formulations reveal the infiltration of naturalistic viewpoints into natural law theory. The whole moral dimension of human life is reduced to a basic biological law, and all of morality is based on narrow utilitarian values.

Volney was also the author of works on biblical chronology (hostile to orthodox interpretations) and on ancient history. He proposed a universal alphabet and the study of culture through language.

See also Cabanis, Pierre-Jean Georges; Chateaubriand, François René de; Deism; Destutt de Tracy, Antoine Louis Claude, Comte; Franklin, Benjamin; Laws of Nature; Priestley, Joseph.

Bibliography

Volney's complete works were published as *Oeuvres complètes* in 8 volumes (Paris, 1821) and in 11 volumes (Brussels, 1822).

For literature on Volney, see J. Barni, *Les moralistes français au XVIII[e] siècle* (Paris, 1873); J. Gaulmier, *Volney* (Paris, 1959); and A. Picavet, *Les idéologues* (Paris, 1891).

See also Counihan, Roberat D., "The Political Philosophy of Volney: Case History of French Revolutionary Intellectualism" (MA thesis; University of North Carolina, Chapel Hill, 1969).

L. G. Crocker (1967)

VOLSKI, STANISLAV
(1880–1936?)

Stanislav Volski was the assumed name of Andrei Vladimirovich Sokolov, the Russian Marxist journalist and philosopher. Volski studied at Moscow University but was expelled in 1899. He was active in the Bolshevik faction until March 1917, when he broke with V. I. Lenin. In 1909 Volski published the only pre-Soviet book-length treatise on Marxist ethical theory, but its "Nietzschean" individualism had little impact on the development of Marxism-Leninism. In the 1920s and 1930s Volski was reduced to the status of literary popularizer and translator. The date and circumstances of his death are still unknown.

According to Volski, class solidarity and discipline are tactically essential to victory in the class struggle, but all binding norms will vanish with the defeat of capitalism. Under socialism individuals will be "freed from the numbing pattern of coercive norms" and from the "idea of duty," the "inevitable companion of bourgeois society" (*Filosofiya Borby,* p. 272).

Volski saw societies as weapons that individuals use in their struggle with nature. Typically, in bourgeois societies (based on fixed division of labor), individuals are free to develop only within the narrow confines of their occupational specialties. As a result they are self-alienated, conformist, and myopic. But in socialist society (based on variable division of labor), harmoniously self-determining individuals will grow into unique selfhood as ends in themselves. Their absolute value as persons will not be a formal postulate or imperative, as was claimed by the Russian Kantian Marxists, but rather a goal to be achieved by free struggle and social creativity. In this process "the socialization of methods is accompanied by an individualization of goals" (ibid., p. 300). "Struggle," Volski declared, "is the joy of being," and "socialism is

freedom of struggle; everything that increases struggle is good, everything that diminishes it is bad" (ibid., pp. 306, 302).

Assimilating Friedrich Nietzsche's insight that "enemy" means not "villain," but "opponent," Volski claimed that I should grant full freedom to the individual whose ideal is inimical to mine and that I should strive to make him an "integral personality," working with him to remove external obstacles to our sharp and clear collision. In struggling with me, he enriches me, enlivening my highest values. "Of all those who surround me, … the most precious, most essential is he with whom I struggle for life and death." He is both friend and enemy, and we share the "morality of 'friend-enemies'—the morality of the future" (ibid., pp. 310, 311).

See also Lenin, Vladimir Il'ich; Marxist Philosophy; Nietzsche, Friedrich; Russian Philosophy; Socialism.

Bibliography

Filosofiya Borby: Opyt Postroyeniya Etiki Marksizma (The philosophy of struggle: an essay in Marxist ethics). Moscow, 1909.

Sotsialnaya Revolyutsiya na Zapade i v Rossi (The social revolution in the West and in Russia). Moscow, 1917.

"Volski." In *Bolshaya Sovetskaya Entsiklopediya* (Great Soviet encyclopedia), 1st ed. Moscow, 1929. Vol. XIII, Cols. 66–67.

George L. Kline (1967)

VOLTAIRE, FRANÇOIS-MARIE AROUET DE

(1694–1778)

François-Marie Arouet de Voltaire encapsulates the spirit of the French Enlightenment in both his refusal to develop a philosophical system and his clear concern for social and political issues. But he is also representative of the eighteenth century in his deep attachment to John Locke's epistemological thought, his emphasis on the limited nature of human understanding, and his commitment to popularizing philosophy, especially by handling it through the medium of novels and tales in which irony often functions as an ad hominem argument. It is thus that he fulfilled the role of philosopher and that his philosophy met the needs of his times, times characterized by a break with seventeenth-century dogmatism and an intensification of the critique of the political and religious spheres aiming to bring forth a morality on the human scale, centered on the values of tolerance and respect for others. Those values were soon to bear fruit in the doctrine of the Rights of Man.

Born in Paris to an established bourgeois Parisian family, François-Marie Arouet, who took the name Voltaire in 1718, received a sound education from his Jesuit teachers at the Collège Louis-le-Grand and soon managed to make his way into the most brilliant Parisian intellectual milieu of his time. There, he gave evidence of his poetic talent and satiric verve—the latter cost him a brief exile to the Netherlands in 1713 and periods of imprisonment in the Bastille in 1717–1718. In the years that followed, he issued an epic poem, *Henriade* (1723), celebrating the tolerance of King Henry IV of France and entrenching his literary prestige on the Parisian intellectual scene. A romantic quarrel with the chevalier de Rohan in 1726 resulted in Voltaire's being exiled to England, where he lived until 1728, taking advantage of the circumstances to improve his English and absorb English culture, especially in the field of philosophy. During this period, he read William Shakespeare, deepened his knowledge of Locke and Isaac Newton, became familiar with Deism, and made the acquaintance of Jonathan Swift, Alexander Pope, John Gay, and doubtless George Berkeley. This sojourn also enabled him to take a detached perspective on French intellectual, political, and religious life.

On his return to France, he published *Temple du goût* (1733), which anticipates his praise for French classicism in 1751 in *Siècle de Louis XIV*; *Épître à Uranie* (1732), an early challenge to the notion of divine goodness; and the famous *Lettres philosophiques* (1734), which contain the essentials of the philosophical plan he subsequently sought to carry out. These were followed by *Remarques sur les pensées de Pascal* (1734). The publication of *Lettres philosophiques*, which discredited the regime under which France was governed by contrasting it to the more liberal English model, resulted in exile once again, this time to the home of Madame du Châtelet in Lorraine. Voltaire took advantage of this extended retreat (1734–1749)—which was broken up by excursions to Paris and Sceaux to advance his candidacy for official positions (historiographer royal in 1745 and election to the Académie française in 1746)—to produce the some fifty tragedies and comedies that won him literary renown; gather together documents on history; work on philosophy (*Traité de métaphysique* dates from 1734); and publish his *Éléments de la philosophie de Newton* (1738), on the thinker with whose approach to physics Voltaire's metaphysical theism was in sympathy.

After Madame du Châtelet's death in 1749 and a brief stay in Paris, Voltaire went into voluntary exile at the court of Frederick II of Prussia, with whom he had been corresponding for years. It was during his Prussian period, in 1751, that he published *Siècle de Louis XIV*. A quarrel with Frederick about a diatribe against Pierre-Louis Moreau de Maupertuis published by Voltaire led to his departure from Berlin in 1753. He went to Paris and from there to Geneva, Switzerland, where he settled in 1755. His Geneva period saw the start of his collaboration on the *Encyclopédie*, the publication of his *Essai sur les moeurs* (1756), and the production of works, like the celebrated *Candide*, that were increasingly critical of established religion. To protect himself against possible reprisal, Voltaire decided in 1760 to permanently settle in Ferney, France, which sits near the French-Swiss border. It was here that he became truly celebrated and his home took its place among the most fertile centers of intellectual activity of the time, thanks to his sustained correspondence with the elite of Europe, including Catherine II of Russia. Here, too, he wrote many novels and tales that enhanced his fame and he took up his role as the opponent of injustice, defending victims of intolerance and fanaticism. A case in point is his well-known struggle on behalf of the Protestant merchant Jean Calas, who was unjustly condemned, tortured, and executed.

Voltaire's struggles to promote religious tolerance cannot be viewed separately from his all-out attack on Catholicism in many vigorously worded pamphlets such as *Sermon des Cinquante* (1762), *Questions sur les miracles* (1765), and *Examen important de Milord Bolingbroke, ou le tombeau du fanatisme* (1767). However, it was his battles in defense of justice that won him a special place in the hearts of his contemporaries, who gave him a triumphant welcome on his return to Paris in 1778 to present the last of his tragedies, *Irène*. Voltaire died in Paris on May 30, 1778, aged eighty-four. The clergy of that city refused to give him a Christian burial, so his body was transported to the Abbey of Scellières, near Troyes. Subsequently, during the Revolutionary period, his remains were returned to Paris and buried in the Pantheon.

PHILOSOPHY

Although he was fully familiar with the French tradition, especially Michel Eyquem de Montaigne, Pierre Gassendi, René Descartes, and Pierre Bayle, thinkers with a common interest in skepticism, following his stay in London Voltaire drew the essentials of his philosophical position from the English tradition. From Locke's thought, he adopted the critique of the notion of innate ideas; the role assigned to philosophical inquiry as the means for best determining the faculties and limits of human understanding; and the acceptance of the unknowable nature of the essence of things. These precepts set him on the road to ontological skepticism. Doubt regarding external things was mirrored by doubt about human interiority, concerning that it is possible to believe that its distinguishing constituent, thought, is nothing more than a product of matter. Locke had indicated the possibility of "thinking matter" and Voltaire gives him a degree of credit for this but does not attempt to decide the question, because, as he says in *Le philosophe ignorant* (1767), one's knowledge of substance, whether material or spiritual, is not a given:

> Once again, what I am saying is not that it is matter that thinks in us; I am saying, with [Locke], that it does not behoove us to state it is impossible for God to cause matter to think, that it is absurd to state this, and that it is not up to earthworms to limit the power of the Supreme Being. (Art. 29; in *Oeuvres complètes*, ed. Moland, vol. 26).

Is Voltaire duping his readers here to lead them toward atheism? Not at all. His invocation of the divine is sincere and flows from his engagement with English thought. For it is from Newton that he drew the notion that the universe is a manifestation of the existence of God and that gravitational physics appears to prove that matter submits to the laws decreed by its creator. In response to criticism of Newton that characterized gravitational attraction as an occult quality of a kind equivalent to the notorious Cartesian vortices, Voltaire bent to the task of showing that an unknown cause can be proven to exist from its effects. Thus, even if attraction is not a perceivable thing, it is nevertheless the case that its existence is a true fact, because it is possible to prove its effects and calculate its proportions, even while acknowledging that this phenomenon's ends are hidden from one and known to God alone. Along the road to probabilistic knowledge of the natural order, Newton had opened up a way by proposing a procedure featuring the integrity and prudence implied by the watchword *hypothesis non fingo* (I feign no hypotheses). Allying Locke with Newton thus led Voltaire to a theistic vision consisting, on the one hand, of admitting the existence of God, conceived as the sole necessary being—but without saying anything about God's attributes nor the ends of God's creation—while on the other hand admitting the existence of a finite and contingent matter that requires divine aid to be set in motion.

METAPHYSICS

THE ONTOLOGICAL STATUS OF REALITY. Anticipating the definition of metaphysics proposed by Étienne Bonnot de Condillac in his *Essai sur l'origine des connaissances humaines* (1746), in which "ambitious metaphysics," which presumes to discover all and know all, is distinguished from "restrained metaphysics," which contains its inquiries within the limits of the weaknesses of the human mind, Voltaire, following in Locke's footsteps, conceives of metaphysics as a naturally limited science whose methods can only be founded on empiricism. As he wrote to Frederick II, "Metaphysics, in my opinion, is made up of two things, the first what all men of good sense know, the second what they will never know." In this light, Voltaire's skepticism can be termed *Zetetic* (to make use of an ancient term): that is, it is perennially in search of truth, even though truth is by nature destined to escape it, and it perennially revisits its own assumptions, accepting that over time some of its initial convictions will be subjected to critique or abandoned.

If there is one point on which Voltaire's position was to remain unchanged, it is surely the existence of two opposed substances: God and matter. His conviction on this score led him to oppose both the materialists and Berkeley's immaterialism. Still, Voltaire's conception of the relationship between these two substances underwent continuous change. The existence of matter appears obvious, at least in its phenomenal manifestation: it is sufficient to allow objects to take their effect on the senses to be persuaded of their presence. Belief in the existence of God rests on two banal proofs, recalled in *Traité de métaphysique*: the proof from ultimate causation (God is the architect of a world that acknowledges its Demiurge) and the proof *a contingentia mundi*, according to which the ultimate reason for things can only be found in a necessary Being who constitutes the ultimate explanation for them. (Voltaire subsequently abandoned the latter proof, retaining only the teleological one.) There flows from this the existence of this necessary Being, conceived as infinite, whose infinity is expressed through its eternity, immensity, and omnipotence. One can see why Voltaire opposed materialism all his life: it appeared to him to be an untenable form of reductionism, as well as to confuse two distinct levels by ascribing the quality of necessity to necessarily contingent matter.

Having acknowledged the existence of two substances, it is necessary to consider their relationship and in particular the two delicate matters of creation and of the existence of evil. The problem of creation is presented as early as *Traité de métaphysique* in the form of a set of alternatives: Either God drew the world out of nothingness or else he drew the world out of himself. The first alternative is doubtful: How can something be drawn from nothing? The second is equally so: It comes down to conceiving the world as a part of the divine essence. Logically, then, one must conclude that the world has eternal existence, but that would presuppose an eternity other than divine eternity.

The hypothesis of God's freedom makes it possible to settle this question: It is because God is free that he created the world at the moment he wished to. However, this brings one back to the first difficulty, that of creation ex nihilo, which was deemed untenable from the outset. As early as the *Éléments* in 1738, Voltaire had turned to the concept of divine decree to reconsider the idea of the existence of necessary and eternal matter. In *Tout en Dieu*, he explains the eternity of matter with a simple argument: Since God is the first cause and every cause has effects, one can conclude that God has been acting for all eternity and therefore that the material world is eternal. In 1768, in *Philosophe ignorant*, Voltaire was to reach the inevitable conclusion implied by this argument when he reasoned that the world is a form of eternal emanation from God, while guarding against pantheistic slippage and definitively rejecting the Christian concept of creation ex nihilo.

THE PHENOMENAL STATUS OF REALITY So much for relations between God and the material world. What of the more specific relationship between the soul and the body? First, it is necessary to be able to be sure of the existence of the soul. Now, if God has the power to give to matter the possibility of thought, why would he burden himself with useless substance? Called on to choose between pure idealism and strict materialism, Voltaire preferred to invoke his ignorance of this subject and to maintain doubt, "because it is just as presumptuous to say that a body organized by God Himself cannot receive the thought of God Himself as it is ridiculous to say that spirit cannot think" (*Philosophe ignorant*, art. 29; in *Oeuvres complètes*, ed. Moland, vol. 26). It is easy to foresee that doubt would also prevail on the question of the form taken by human freedom, which may in reality consist of pure material determinism or be a reflection within one of divine freedom.

In fact, over time, Voltaire did come close to a deterministic position that led him, in the name of the principle of parsimony (which makes it superfluous to hypothesize a soul acting on the body), to explain the process of cognition wholly in materialist terms and to

deny the Cartesian concepts of liberty of indifference and free will. Thus, in the entry on freedom in the *Dictionnaire philosophique*, *freedom* is defined strictly in negative terms, as the ability to do what one wishes, or rather as will that is determined by the set of causes that constitute the world —causes that ultimately refer to a prime mover that is their reason. The materialism that makes it possible to describe the order of the world and the laws of that order, and thus human actions as a part of it, must always be framed as being dependent on a spiritual principle that is alone capable of explaining its proper functioning. This accounts for Voltaire's glowing praise of Nicolas Malebranche in *Tout en Dieu*, since occasionalism is the system that provides the most correct explanation for the interactions that occur in the world, which at bottom have only one true cause: God.

PHILOSOPHICAL OPTIMISM

Whereas Voltaire's position on the question of creation and divine and human freedom evolved only somewhat, there is one problem in connection with which his intellectual evolution was radical, that of the existence of evil. In his early writings, he seems not to grasp the real difficulty posed by the existence of physical and moral suffering (and in this he is close to Pope and Gottfried Wilhelm Leibniz), making it vanish by adopting the perspective of the whole: If, since God himself is good, the organization of the universe as a whole is good, then the evil that one sees appearing here and there is justified at the holistic level. Indeed, it may not even be evil, since the notion of evil is always relative and its existence undoubtedly has a function, that of revealing the beauty of the whole, just as shadows are necessary to accentuate the effects of light in a picture.

But the 1755 Lisbon earthquake played for Voltaire the role that Auschwitz and Dachau played for philosophers in the second half of the twentieth century: it was a revelation of evil that is absolute because wholly gratuitous. *Poème sur le désastre de Lisbonne* (1756) and *Candide* (1759) show Voltaire attaining awareness of the positive existence of evil, evil that appears to have no possible justification. And yet God exists and, as a free being, he must be responsible for the disasters caused by the natural laws that he has willed. Must one therefore assign the fault to God, which would constitute true blasphemy? Voltaire is unafraid to affirm precisely that: since evil exists, it must be necessary that this be so, with evil being a necessary condition of divine action. In contrast to Leibniz, who claims to justify the existence of evil and thus rescue the principle of God's goodness, Voltaire seeks to excuse God by showing that undoubtedly he did his best but did not create the best of all possible worlds, and by acknowledging that the ultimate explanation for the reality of evil exceeds the bounds of one's understanding.

RELIGION AND ETHICS

If one restricts oneself to the etymological significance of the word *religion*, which evokes the linking of individuals to one another, Voltaire must be said not to have had a religion, because for him the relationship with the divine is strictly personal and requires no collective rite. But if one agrees to conceive of religion as a specific relationship linking the human to the divine, Voltaire was a fully religious person. To be religious is, for Voltaire to worship God as the reasonable cause of everything that happens; to thank him for having allowed one to benefit from it and marvel at it; and not to seek to adopt the divine perspective and claim thereby to understand its decrees, but to wish humbly to understand why something that happens in one way does not happen in another. It is thus up to reason to lead one to the Supreme Being, which is itself universal reason, and not up to faith, which wraps things up in mystery and relies on miracles to better subordinate weak minds and enable priests to exercise power over them. Voltaire's theism is in no sense a natural theology; but it aims to be a purified form of natural religion, along the lines set down by Herbert of Cherbury, and is wholly opposed to both positive religion and atheism.

Voltaire's opposition to atheism is categorical and rests on a simple argument: The laws of the physical world are so reasonable that they necessarily presuppose an intelligent artisan. His opposition to established religion is equally categorical. His celebrated watchword, "*Écrasez l'infâme* (Erase the infamy)," is a reminder of how violently he struggled against Christianity, especially toward the end of his life, when fear of political power, the enforcer of religious power, had diminished in him. His exasperation was directed less against the message of Christ, which he incorporated into a universalist conception of human values, than against what the church as an institution had done with that message and against the methods it had used to disseminate it (e.g., superstition, the worship of relics, faith in miracles, the establishment of the Inquisition, and incitement to fanaticism).

In his struggle against "*l'infâme*," he used every available weapon and did not hesitate to borrow alike from Christians and atheists, skeptics and deists,—those of their arguments that seemed to him the strongest. Over the course of this long struggle, Voltaire's immense erudi-

tion stood him in good stead, and he was effective at searching out the most convincing reasoning wherever necessary, turning to the European scholarship of previous centuries as well as to his contemporaries. He invoked Italian (Giordano Bruno and Giulio Cesare Vanini), English (John Toland, Anthony Collins, Matthew Tindal, Thomas Woolston, Henry St. John Bolingbroke, and Thomas Chubb), German (Henricus Cornelius Agrippa von Nettesheim, and Desiderius Erasmus), and French writers (Théophile de Viau, Jacques Vallée des Barreaux, François de La Mothe Le Vayer, Charles de Marguetel de Saint-Denis [Seigneur de Saint-Évremond], Pierre Bayle, and Julien Offray de La Mettrie).

In 1762, Voltaire went so far as to publish a long extract from the *Testament* of Jean Meslier, a text that was extremely hard on Christianity, written by one who knew it well because he had served it for many years as the curé of Étrépigny, France. Voltaire took care to touch up the text perceptibly, with a view to preserving natural religion and keeping only those criticisms that targeted revealed religion. Why preserve natural religion and not be satisfied with an internal religion that would amount at bottom to a system of morality? This is accounted for by Voltaire's anthropological pessimism. Human beings would not respect the rules of morality if there was no religion to bring those rules before their consciences. In truth, religion and morality are one and the same, as is to be inferred from a formulation found in chapter 4 of the *Oreilles du comte de Chesterfield et le chapelain Goudman* (1775): "Let us do our duty to God, let us worship Him, let us be just: that is what our true praise and true prayers consist of." (*Oeuvres complètes*, ed. Moland, vol. 39). In other words, religion is the morality of the weak and morality the religion of the strong. It would be possible to do without religion if everyone was wise and respected the moral law engraved in every heart. But that is not the case, and that is why religion retains its usefulness, as does the notion of punishment and reward following death, which alone can serve to temper bad inclinations and make social life not only possible but indeed agreeable.

But what morality is one speaking of, and how does Voltaire picture it? On this score, it is possible to draw an analogy between the natural world and the moral world. Just as the laws of the natural world can be uncovered by one who applies one's intelligence to the matter, those of the moral world are unveiled if one takes the trouble to reflect on them; and in light of such reflection, they lead one to distinguish right from wrong. What makes it possible to differentiate morality from particular systems of ethics specific to a given people is its universality, that it transcends not just borders but centuries. The beauty of a moral act does not change with time; the truth of moral values is not subject to relativism. Thus, it will always be right to defend the poor and the oppressed and always wrong to condemn without proof. That is how setting an example of virtue by practicing it confers a kind of immortality. In the West, Socrates exemplifies this truth; in the East, Confucius. At bottom, in the eyes of Voltaire (who on this score is heir to the Greeks), a philosopher's value resides more in the way he or she has lived life than in the system he or she has sought to build.

JUSTICE AND TOLERATION

Voltaire's involvement in social issues can be explained on the basis of his philosophical convictions. Since moral law exists, it must operate to the benefit of others and rest on the justice one owes to other natural beings, human beings in particular. In fact, virtue is nothing more than beneficence directed toward one's neighbor. The inverse is also true: Vice is malice directed toward that same neighbor. In this connection, nothing aroused greater indignation in Voltaire than the excesses of religious fanaticism. Under the Ancien Régime, these excesses were tolerated politically, the government often serving as an accomplice to them and never as the detached judge of collective passions or of the crying injustices to which such passions gave rise. In the manner of an anthropologist, Voltaire ascribed the weakness for fanaticism to Westerners only, ever concerned to seek the welfare of others even at their own expense and seeing in Easterners a willingness to be satisfied with complete indifference to their neighbors.

In light of these views, it is possible to understand Voltaire's militant stance in favor of enlightened despotism: It is the corollary of his anthropological pessimism, requiring a strong but just prince to ensure that the diverse factions that constitute the state do not destroy each other. For Voltaire was not just a philosopher; he was also a historian, and he knew that, because human beings prey on each other, barbarity is always at the gates, bringing the possibility of massacres in its train. Voltaire sought to serve as the unquiet watchman of the Enlightenment, to ensure that the light shed by his times should not be swallowed up in total darkness.

Voltaire played this role of watchman by defending unjustly accused contemporaries, as witness his efforts on behalf of Pierre-Paul Sirven, Thomas Arthur Lally, baron De Tollendal, the Chevalier Jean-François de La Barre, and especially Calas *père* and *fils*. With the Calas affair, the most celebrated cause defended by Voltaire, tolerance

became his primary concern and, little by little, he let go of reflection in favor of action, conscious that only involvement by philosophers makes the exercise of justice possible and that, without such involvement, justice would remain an abstract notion reigning over a heaven of Platonic ideas.

In *Traité sur la tolérance à l'occasion de la mort de Jean Calas* (1763), one can discern Voltaire's method for bringing about the triumph of a cause that he deems just. This work is a treatise in name only: It brings together an account of the Calas affair with past examples of fanaticism, general historical reflections on tolerance, a dialogue between a dying man and a well man, and a letter to the Jesuit Father Le Tellier, all designed to reveal the possible breakdown of tolerance, before concluding with an account of the most recent decree regarding the Calas family. Making use of all the stylistic resources Voltaire had at his disposal, this work seeks to convince by playing on readers' emotions. Taking readers from laughter to tears, it designedly forces them to pity the Calas family, a technique calculated to bring about awareness of the Calases' true misfortune.

Voltaire undoubtedly realized early on that his struggle would not suffice if it were not backed up by a complete recasting of legislation with a view to limiting injustice. This is what lies behind his strong interest in Cesare Bonesana Beccaria's masterwork, *Essay on Crimes and Punishment*, which he read and commented on with minute attention. His reading of Beccaria led him to believe that only judicial reform would make possible the real-life implementation of Enlightenment ideals. An echo of this concept of judicial reform is found in his *Prix de la justice et de l'humanité* (1777), composed one year before his death. Here, Voltaire advances his vision of a society built on just laws, one that prefers prevention to punishment, tolerance to fanaticism. He lauds the principle that the punishment should fit the crime and criticizes capital punishment and recourse to torture; and he insists the law must have a public nature and must not be obeyed unless it is known to all (as Thomas Hobbes had already stipulated in *Leviathan*). Furthermore, the law must be applied by judges of integrity, chosen on the basis of merit and not by reason of their social origins. In this regard, Voltaire is one of the main sources of inspiration for the ideals of the French Revolution.

HISTORICAL PHILOSOPHY

Voltaire's historical project cannot be dissociated from his philosophical and moral concerns. Once again, an analogy helps clarify the point: Since both the natural world and the moral world are governed by laws, it must also be possible to identify those of the historical world. To do so, a rigorous method is necessary, one that admits only acknowledged facts and repudiates mythical discourse, just as Voltaire undertook to do in his *Histoire de Charles XII* (1739). More than a methodology, historical work must have its own proper end, that of extracting coherent meaning from the mass of historical data. It is for this reason that, in *Siècle de Louis XIV*, Voltaire abandons narrative history (the approach he had taken with *Charles XII*, for example) in favor of a more general historiography—philosophical this time—that seeks to present the state of mind of a century and not to analyze the personal strengths and shortcomings of an individual. In thus depicting a vast panorama of human history, in which individual actions are brought into relation with an organized whole, Voltaire anticipates the Hegelian concept of the spirit of a people (*Volksgeist*).

It is with *Essai sur les moeurs et l'esprit des nations* (1756), however, that Voltaire let go the approach of a history limited to an individual or a century, to seek to extract from a mass of historical data a vision of human becoming made possible by an analysis of the mores and spirit of nations. Thus, rather than perceiving in the long view of history a movement toward salvation, as had Jacques Bénigne Bossuet in his *Discours sur l'histoire universelle*, Voltaire sees in it the immanent progress of civilization founded ultimately on universal morality and rationality. This movement of universal reason, however, does not have the character of necessity, since breaches of universal moral obligation are always possible. The concept of a universal history is merely a way of expressing a finding that one reports on in one's capacity as a historian reflecting on human history as a whole. This finding comes down to the view that it is reasonable to believe that the essence of reason consists of a permanent striving toward the good. As to knowing whether this is really so, and especially whether it will always be so in the future, Voltaire refrains from judgment: here as elsewhere, he adopts the role of skeptic rather than that of dogmatist.

See also Agrippa von Nettesheim, Henricus Cornelius; Atheism; Bayle, Pierre; Berkeley, George; Bolingbroke, Henry St. John; Bossuet, Jacques Bénigne; Bruno, Giordano; Chubb, Thomas; Clandestine Philosophical Literature in France; Collins, Anthony; Condillac, Étienne Bonnot de; Deism; Descartes, René; Enlightenment; Erasmus, Desiderius; Ethics, History of; Gassendi, Pierre; Gay, John; Innate Ideas; La Mettrie, Julien Offray de; La Mothe Le Vayer, François de; Leibniz,

Gottfried Wilhelm; Locke, John; Meslier, Jean; Montaigne, Michel Eyquem de; Newton, Isaac; Pessimism and Optimisim; Philosophy of History; Pope, Alexander; Socrates; Swift, Jonathan; Tindal, Matthew; Toland, John; Vanini, Giulio Cesare; Woolston, Thomas.

Bibliography

WORKS BY VOLTAIRE

Oeuvres complètes. 53 vols., edited by L. Moland. Paris: Garnier frères, 1877–1885.

Voltaire's Correspondence. 135 vols., edited by Theodore Besterman. Geneva, Switzerland: Institut et Musée Voltaire, 1953–1977.

Complete Works of Voltaire. 85 vols., edited by Theodore Besterman, W. H. Barber, and N. Cronk, Geneva, Switzerland: Institut et Musée Voltaire, 1968–2001.

Corpus des notes marginales de Voltaire. 5 vols. Berlin: Akademie-Verlag, 1979–1994.

WORKS ABOUT VOLTAIRE

Besterman, Theodore. *Studies on Voltaire and the Eighteenth Century*. 380 vols. Geneva, Switzerland: Insitut et Musée Voltaire, 1956–2000.

Brooks, Richard A. *Voltaire and Leibniz*. Geneva, Switzerland: Librairie Droz, 1964.

Brumfitt, John H. *Voltaire, Historian*. London: Oxford University Press, 1958.

Carré, Jean Raoul. *Consistance de Voltaire le philosophe*. Paris: Boivin et cie, 1938.

Dédéyan, Charles. *Voltaire et la pensée anglaise*. Paris: Centre de documentation universitaire, 1956.

Gay, Peter. *Voltaire's Politics: The Poet as Realist*. Princeton, NJ: Princeton University Press, 1959.

Goulemot, Jean, André Magnan, and Didier Masseau, eds. *Inventaire Voltaire*. Paris: Gallimard, 1995.

Lanson, Gustave. *Voltaire*. Paris: Hachette et cie, 1906.

Lauer, Rosemary Z. *The Mind of Voltaire: A Study in His "Constructive Deism."* Westminster, MD: Newman Press, 1961.

Lepape, Pierre. *Voltaire le conquérant: Naissance des intellectuels au siècle des Lumières*. Paris: Editions du Seuil, 1994.

Maestro, Marcello T. *Voltaire and Beccaria as Reformers of Criminal Law*. New York: Columbia University Press, 1942.

Martin-Haag, Éliane. *Voltaire: du cartésianisme aux Lumières*. Paris: Vrin, 2002.

McKenna, Antony. *De Pascal à Voltaire: Le rôle des Pensées de Pascal dans l'histoire des idées entre 1670 et 1734*. 2 vols. Oxford, U.K.: The Voltaire Foundation at the Taylor Institution, 1990.

Pellissier, Georges. *Voltaire philosophe*. Paris: A. Colin, 1908.

Pomeau, René. *La religion de Voltaire*. Paris: Librairie Nizet, 1956.

Pomeau, René. *Politique de Voltaire*. Paris: A. Colin, 1963.

Porset, Charles. *Voltaire humaniste*. Paris: Editions maçonniques de France, 2003.

Ridgeway, Ronald S. *La propagande philosophique dans les tragédies de Voltaire*. Geneva, Switzerland: Institut et Musée Voltaire, 1961.

Schwarszbach, Bertram Eugene. *Voltaire's Old Testament Criticism*. Geneva, Switzerland: Librairie Droz, 1971.

Torrey, Norman Lewis. *Voltaire and the English Deists*. New Haven, CT: Yale University Press, 1930.

Trousson, Raymond, Jeroom Vercruysse, and Jacques Lemaire, eds. *Dictionnaire Voltaire*. Paris: Hachette, 1994.

Trousson, Raymond, and Jeroom Vercruysse, eds. *Dictionnaire général de Voltaire*. Paris: Honoré Champion, 2003.

Wade, Ira O. *The Intellectual Development of Voltaire*. Princeton, NJ: Princeton University Press, 1969.

Waterman, Mina. *Voltaire, Pascal, and Human Destiny*. New York: King's Crown Press, 1942.

Sébastien Charles (2005)

VOLUNTARISM

The term *voluntarism* (from the Latin *voluntas,* "will") applies to any philosophical theory according to which the will is prior to or superior to the intellect or reason. More generally, voluntaristic theories interpret various aspects of experience and nature in the light of the concept of the will, or as it is called in certain older philosophies, passion, appetite, desire, or *conatus.* Such theories may be psychological, ethical, theological, or metaphysical.

PSYCHOLOGICAL VOLUNTARISM

Voluntaristic theories of psychology represent men primarily as beings who will certain ends and whose reason and intelligence are subordinate to will. The outstanding classical representatives are Thomas Hobbes, David Hume, and Arthur Schopenhauer. Hobbes, for example, thought that all voluntary human behavior is response to desire or aversion, which he brought together under the name "endeavor"; he based his ethical and political theories chiefly on this claim. Hume maintained that reason has no role whatever in the promptings of the will; that "reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them." Schopenhauer, the outstanding voluntarist of them all, believed that the will is the very nature or essence of man and indeed of everything, identifying it with the "thing-in-itself" that underlies all phenomena.

The point of all such theories can best be appreciated by contrasting them with the more familiar theories of rationalism found, for example, in Plato's dialogues or René Descartes's *Meditations.* Plato thought that men ideally perceive certain ends or goals by their reason and

then direct their wills to the attainment of these ends or goals. This is why he thought no man could knowingly will evil. Thus in the *Symposium* he traced the ascent of the soul toward higher and higher ends, the supposition being that these ends are apprehended first by the senses and then ultimately by the pure or unfettered intelligence, which enlists the will or desire for their pursuit. The corruption of a man was for Plato precisely the dominance of the will, that is, of a man's appetites or desires, this being a deviation from what human nature ideally should be. Descartes, similarly, supposed that the understanding first grasps certain ideas or presents certain ends to the mind and that the will then either assents or withholds its assent, thus following rather than directing the understanding.

Voluntarist theories reject this general picture as the reversal of the truth. Ends and goals, according to these theories, become such only because they are willed; they are not first perceived as ends and then willed. Hume in particular maintained that no sense can be made of the idea, so central to Plato's philosophy, of reason directing the passions, or even of its ever conflicting with them. Reason, he argued, is concerned entirely with demonstrations (deduction) or with the relations of cause and effect (induction). In neither case can it give us ends or goals. Mathematics is used in mechanical arts and the like, but always as a means of attaining something that has nothing to do with reason. The computations of a merchant, for example, can be fallacious, but the ends for which they are undertaken can in no sense be fallacious or irrational. They can only be wise or foolish, that is, such as to promote or to frustrate other ends that are again products of the will. Similarly, Hume thought that no discovery of causal connections in nature can by itself have the least influence on the will. Such discoveries can only be useful or useless in enabling men to choose appropriate means to certain ends, which are in no way derived from reason. "It can never in the least concern us to know," Hume said, "that such objects are causes, and such others effects, if both causes and effects be indifferent to us." Reason therefore can never produce actions or impulses, nor can it oppose them. An impulse to act can be opposed only by a contrary impulse, not by reason. There can, accordingly, be no such thing as a conflict between reason and passion, and the only way in which willed behavior can be "irrational" is for it to be based upon some misconception—for instance, on some erroneous conception of what is a fit means to the attainment of an end that is entirely the product of the will.

The theories of other voluntarists do not differ essentially from Hume's theory, although there are differences of emphasis. All agree that men are moved by their impulses, appetites, passions, or wills and that these are incapable of fallacy or error. There is thus no such thing as a rational or irrational will, although one may will imprudently in relation to other things that one wills. J. G. Fichte expressed this idea when he said that a free being "wills because it wills, and the willing of an object is itself the last ground of such willing."

ETHICAL VOLUNTARISM

It is obvious that the voluntarist conception of human nature contains implications of the highest importance for ethics. If ends or goals are entirely products of the will and the will is neither rational nor irrational, then ends themselves cannot be termed either rational or irrational and it becomes meaningless to ask whether this or that end is really good or bad independently of its being willed. Hobbes drew precisely this conclusion. To say that something is good, he said, is to say nothing more than that it is an object of one's appetite, and to say that something is bad is only to say that one has an aversion to it. Good and bad are thus purely relative to desires and aversions, which are, of course, sometimes quite different in different men. Wise behavior, on this conception, can be nothing other than prudence, that is, the selection of appropriate means to the attainment of whatever goals one happens to have. Hobbes thought that there is one goal, however, that is fairly common to all men: the goal of self-preservation. His political philosophy thus consisted essentially of formulas by means of which men can preserve themselves in safety and security within a commonwealth.

Essentially the same ideas were defended by Socrates' contemporary, Protagoras, and are reflected in his maxim that "man is the measure of all things." They also find expression in the philosophy of William James and are, in fact, an important aspect of pragmatism in general. James thought that things are good solely by virtue of the fact that they are "demanded," that is, that someone wants them or lays claim to them, and he noted that such a demand might be for "anything under the sun." Considered apart from the demands of sentient beings, nothing in the universe has any worth whatsoever. Hence James concluded that the only proper ethical maxim is to satisfy as many demands as possible, no matter what these happen to be, but at the "least cost," that is, with the minimum of frustration to other demands. It is clear that within the framework of voluntaristic theories like this,

no meaning can be attached to asking what is truly worthy of one's desires, unless this question is interpreted to mean "What is in fact satisfying of one's desires?"; nor does it make sense to seek, as did Immanuel Kant, any metaphysical principles of morals. Truth and falsity in ethics are exhausted in questions as to the truth or falsity of various opinions concerning the utility of proposed means to the achievement of ends, that is, to the satisfaction of appetite, desire, and demand. They have no relevance to any questions concerning ends themselves.

THEOLOGICAL VOLUNTARISM

Just as the theories thus far described give prominence to the human will over human reason, so certain theological conceptions give prominence to the divine will. Perhaps the most extreme form of theological voluntarism is exemplified in the thinking of St. Peter Damian (1007–1072). He maintained that human reason or "dialectic" is worthless in theological matters, for the simple reason that the very laws of logic are valid only by the concurrence of God's will. God is omnipotent, he said, and can therefore render true even those things reason declares to be absurd or contradictory. It is thus idle for philosophers to speculate upon what must be true with respect to divine matters, since these depend only on God's will.

A very similar idea has found expression in many and various forms of fideism, according to which the justification of religious faith is found in the very act of faith itself, which is an act of the will, rather than in rational proof. Thus Søren Kierkegaard described purity of heart as the willing of a single thing and emphatically denied that such notions as reason and evidence have any place in the religious life. William James, following suggestions put forth by Blaise Pascal, similarly justified the will to believe, defending the absolute innocence, under certain circumstances, of religious belief entirely in the absence of evidence. Many contemporary religious leaders, pressing the same notion, give prominence to the idea of religious commitment, suggesting that religion is primarily a matter of the will rather than of reason. This is, in fact, traditional in Christian thought, for even the most philosophical and rationalistic theologians, such as St. Anselm of Canterbury, have almost without exception given priority to the act of faith, maintaining that religious belief should precede rather than follow rational understanding. This idea is expressed in the familiar dictum *credo ut intelligam,* which means "I believe, in order that I may understand."

Perhaps no religious thinker has stressed the primacy of God's will in questions of morality more than Kierkegaard, who seems to have held that the divine will is the only and the ultimate moral justification for any act. Strictly understood, this means that an action that might otherwise be deemed heinous is not so, provided it is commanded by God. In the fourteenth century this was quite explicitly maintained by William of Ockham. William said that the divine will, and not human or divine reason, is the ultimate standard of morality, that certain acts are sins solely because they have been forbidden by God, and other acts are meritorious only because they have been commanded by God. He denied that God forbids certain things because they are sins or commands certain things because they are virtues, for it seemed to him that this would be a limitation upon God's will. There can be, he thought, no higher justification for any act than that God wills it, nor any more final condemnation of an act than that God forbids it. The moral law, accordingly, was for William simply a matter of God's free choice, for God's choice cannot be constrained by any moral law, being itself the sole source of that law. This view is frequently echoed in religious literature but usually only rhetorically.

METAPHYSICAL VOLUNTARISM

A number of thinkers have believed that the concept of the will is crucial to the understanding of law, ethics, and human behavior generally; a few have suggested that it is crucial to the understanding of reality itself. Such suggestions are found in the philosophies of Fichte, Henri Bergson, and others, but in no philosophy does it have such central importance as in that of Arthur Schopenhauer. Schopenhauer thought that will is the underlying and ultimate reality and that the whole phenomenal world is only the expression of will. He described living things as the objectifications of their wills and sought to explain not only the behavior but also the very anatomical structures of plants, animals, and men in terms of this hypothesis. The will was described by Schopenhauer as a blind and all-powerful force that is literally the inexhaustible creator of every visible thing. The sexual appetite, which he considered to be fundamentally the same in all living things, was described by him as a blind urge to live and to perpetuate existence without any goal beyond that, and he denied that it had anything whatever to do with reason or intelligence, being in fact more often than not opposed to them. The religious impulse found in all cultures at all times was similarly explained as the response to a blind and irrational will to possess endless existence. In the

growth and development of all living things Schopenhauer discerned the unfolding of the will in nature, wherein certain things appear and transform themselves in accordance with a fairly unvarying pattern and in the face of obstacles and impediments, solely in accordance with what is willed in a metaphysical sense but entirely without any rational purpose or goal. On the basis of this voluntarism, he explained ethics in terms of the feelings of self-love, malice, and compassion, all of which are expressions of the will, and he denied—in sharp contrast to Kant—that morality has anything to do with reason or intelligence. He argued that men have free will only in the sense that every man is the free or unfettered expression of a will and that men are therefore not the authors of their own destinies, characters, or behavior. Like other voluntarists, Schopenhauer thus emphasized the irrational factors in human behavior and, in doing so, anticipated much that is now taken for granted in those sophisticated circles that have come under the influence of modern psychological theories.

See also Anselm, St.; Bergson, Henri; Descartes, René; Determinism, A Historical Survey; Dialectic; Ethics, History of; Fichte, Johann Gottlieb; Fideism; Hobbes, Thomas; Hume, David; James, William; Kant, Immanuel; Kierkegaard, Søren Aabye; Pascal, Blaise; Peter Damian; Plato; Protagoras of Abdera; Schopenhauer, Arthur; Socrates; Volition; William of Ockham.

Bibliography

A good general work on voluntarism that is both historical and critical is Vernon J. Bourke's *Will in Western Thought* (New York: Sheed and Ward, 1964).

Thomas Hobbes's ethical and political theories are developed in his *Leviathan,* of which there are many editions. See also *Body, Man, and Citizen,* a selection from Hobbes's writings edited by Richard S. Peters (New York: Collier, 1962).

David Hume's defense of psychological voluntarism is best expressed in his *Treatise of Human Nature,* Book II, Part 3, especially Section 3; the quotations in this article are taken from L. A. Selby-Bigge's edition of that work (Oxford, 1888 and 1955).

For a fairly concise expression of Schopenhauer's voluntarism, see *The Will to Live,* a collection of his essays edited by R. Taylor (New York, 1962).

J. G. Fichte develops the idea of an active power in nature in the first book of his *The Vocation of Man,* edited by R. M. Chisholm (New York: Liberal Arts Press, 1956). The quotation from Fichte comes from his *The Science of Rights,* translated by A. E. Kroeger (Philadelphia: Lippincott, 1869), p. 193.

One of the clearest ancient defenses of a pragmatic basis of laws and institutions is given in Plato's *Protagoras,* where it is ascribed to Protagoras and criticized by Socrates. William James's ethical voluntarism is developed in his essay "The Moral Philosopher and the Moral Life," and the application of his principles to religious belief is given in his "The Will to Believe," both of which are found in nearly all editions of his popular essays.

The theological voluntarism of St. Peter Damian, as well as William of Ockham's ethical theories, are very well summarized in Frederick Copleston's *History of Philosophy,* Vols. II and III (London, 1950 and 1953). See also Étienne Gilson's *Reason and Revelation in the Middle Ages* (New York: Scribners, 1938) for a clear account of the opposition between rationalism and fideism. Kierkegaard has eloquently expressed the opposition between reason and religion in many writings, but see particularly his *Purity of Heart,* translated by Douglas V. Steere (New York: Harper, 1938).

Richard Taylor (1967)

VYSHESLAVTSEV, BORIS PETROVICH

(1877–1954)

Boris Petrovich Vysheslavtsev, the Russian philosopher and religious thinker, was born in Moscow. He studied at the University of Moscow under the Russian jurist and philosopher P. I. Novgorodtsev and later at the University of Marburg under the neo-Kantians Hermann Cohen and Paul Natorp. Upon the publication in 1914 of his dissertation, *Etika Fikhte* (Fichte's ethics), he received a doctorate from the University of Moscow and in 1917 was made professor of philosophy at that institution. Expelled from the Soviet Union in 1922, he emigrated first to Berlin, then in 1924 to Paris, where he became a professor at the Orthodox Theological Institute and was associated with Nikolai Berdyaev in affairs of the Russian *émigré* press. Prior to World War II Vysheslavtsev was active in the ecumenical movement. From the time of the German occupation of France until his death he lived in Switzerland.

Vysheslavtsev's lifelong concern with the themes of irrationality and the absolute was already evident in his work on Johann Gottlieb Fichte. He there asserted that beyond the sphere of rationality or "system" lies the irrational sphere, infinite and incapable of being systematized. Through the antinomy of these spheres philosophy arrives at recognition of the Absolute as the infinity that transcends the universe and all oppositions, even the opposition between Georg Cantor's "actual" and "potential" infinities. Because the Absolute underlies every rational construction, it is irrational. It cannot be exhausted by any concept but is "the mysterious limitlessness which is revealed to intuition."

According to Vysheslavtsev, the essence of man's ethical and religious life consists in his relation to the Absolute. He explored this relation in subsequent works, principally *Etika preobrazhennogo erosa* (The ethics of transfigured Eros), emphasizing the irrational forces in man and interpreting Christian doctrine in the light of the depth psychology of Carl Jung and the French psychoanalyst Charles Baudouin. Vysheslavtsev argued that moral laws cannot guide human conduct successfully, because they are rational rules directed to the conscious will and are defeated by the "irrational antagonism" that stems from man's subconscious. For moral ideals to be significant and effective they must take possession of the subconscious, which they can do only if they are reached through the sublimation of subconscious impulses. Sublimation, operating through the imagination, transforms man's lower impulses into higher ones and turns his inherent, arbitrary freedom into moral freedom that seeks the good. Such sublimation is aided by divine grace and is possible only where the soul turns freely toward the Absolute. Christian ethics is not an ethics of law but "the ethics of sublimation."

In his later years Vysheslavtsev increasingly concerned himself with social problems and wrote a major work on modern industrial culture, *Krizis industrial'noi Kul'tury* (The crisis of industrial culture), and a trenchant philosophical critique of Soviet Marxism, *Filosofskaia nishcheta marksizma* (The philosophical poverty of Marxism).

See also Absolute, The; Berdyaev, Nikolai Aleksandrovich; Cantor, Georg; Cohen, Hermann; Fichte, Johann Gottlieb; Jung, Carl Gustav; Natorp, Paul; Rationality; Russian Philosophy.

Bibliography

WORKS BY VYSHESLAVTSEV

The Eternal in Russian Philosophy. Translated by Penelope V. Burt. Grand Rapids, MI, and Cambridge, U.K.: Eerdmans, 2002.

Etika Fikhte (Fichte's ethics). Moscow, 1914.

Etika preobrazhennogo erosa (The ethics of transfigured eros). Moscow: Respublika, 1994.

Khristianstvo i sotsial'nyi vopros (Christianity and the social question). Paris: YMCA Press, 1929.

Krizis industrial'noi kul'tury: marksizm, neosotsializm, neoliberalizm (The crisis of industrial culture: Marxism, neosocialism, neoliberalism). New York: Chalidze Publications, 1982.

Serdtse v Khristianskoy i Indiyskoy Mistike (The heart in Christian and Indian mysticism). Paris: YMCA Press, 1929.

Filosofskaia nishcheta marksizma (The philosophical poverty of Marxism). Frankfurt am Main: Posev, 1957.

Zwei Wege der Erlosung: Erlosung als Losung des tragischen Widerspruchs (Two paths of salvation: Salvation as resolution of the tragic contradiction). Zurich: Rhein-Verlag, 1937.

WORKS ON VYSHESLAVTSEV

Beliaev, M. M., et al. *O Rossii i russkoi filosofskoi kul'ture: filosofy russkogo posleoktiabr'skogo zarubezh'ia* (Russia and Russian philosophical culture: The philosophers of the Russian post-October emigration). Moscow: Nauka, 1990.

Kline, G. L. "A Philosophical Critique of Soviet Marxizm." *Review of Metaphysics*, 9 (1955) 1: 90–105.

Redlikh, Roman, ed. *Dialektika Vysheslavtseva* (Vysheslavtsev's dialectic). Frankfurt am Main: Posev, 1973.

V. V. Zenkovsky. "B. P. Vysheslavtsev, kak Filosof" (B. P. Vysheslavtsev as philosopher). *Novy Zhurnal*, 15 (1955): 249–261.

V. V. Zenkovsky. *Istoriya Russkoy Filosofii*. 2 vols. Paris, 1948–1950. Translated by G. L. Kline as *A History of Russian Philosophy*, 2 vols., New York and London, 1953.

James P. Scanlan (1967)
Bibliography updated by Vladimir Marchenkov (2005)

W

WAHLE, RICHARD

(1857–1935)

Richard Wahle, the Austrian philosopher and psychologist, was born in Vienna. He was appointed Privatdozent in philosophy at the University of Vienna in 1885. A decade later he was called to a professorship in philosophy at the University of Czernowitz, where he taught until 1917. From 1919 to 1933 he again lectured at the University of Vienna. Possessed of originality and an unusually lively style, he published a number of books in the fields of psychology, general philosophy, and ethics.

Wahle is known especially for his relentlessly sharp critique of traditional philosophy, particularly of metaphysics, which he regarded as "one of the most dangerous breeding-places of empty phrases." An absolute, true knowledge, of the sort to which metaphysics aspires, cannot exist. For all knowledge consists in nothing more than that "an image (or idea) is given in dependence on the self"; a reality existing in itself can never be known. Against the traditional philosophical and metaphysical "delusion of knowledge," Wahle set his own positivistic "philosophy of occurrences," according to which the "given" constitutes the sole admissible point of departure for philosophical thought. What are empirically given to us, however, are only freely suspended, surfacelike, passive, powerless "occurrences" (the contents of perception and imagination) that are the effects of unknown "really operative, powerful substantial primitive factors," which remain forever hidden and are in principle unknowable. Wahle's epistemological standpoint, described also as "antisubjectivist product-objectivism" or "agnostic product-realism," lies beyond the antitheses of materialism and spiritualism, realism and idealism (or phenomenalism), objectivism and subjectivism. He regarded all of these positions as false because things are neither essence nor appearance but simply complexes of "occurrences," and the subjective and the objective are identical inasmuch as only neutral "occurrences" are given to us. Thus Wahle's antimetaphysical and skeptical agnosticism leads from illusory knowledge to genuine ignorance, which is the only attainable goal for philosophy.

As a psychologist, Wahle firmly rejected any kind of metaphysics of the soul, as well as faculty psychology and the depth psychology of the unconscious (psychoanalysis). A satisfactory explanation of mental processes, he held, can result only from connecting them with the corresponding physiological prerequisites. There are no independent psychical unities (like the ego), forces, acts, or powers; they appear to exist only because of an inexact

style of expression. For example, the ego is neither substance nor force; it is not an independent, simple, active thing at all but only a designation for a certain sphere of occurrences. Similarly, the will is said to be "the reflex action become stable under the accompaniment of images following a concurrence of reflex movements" (*Über den Mechanismus des geistigen Lebens,* p. 371).

Wahle attached special value to obtaining as penetrating an analysis as possible of those mental happenings that proceed essentially in "additive series." In such happenings, besides association, the "constellation" (the state of excitation of the brain at the given moment) is particularly significant. Organic sensations and bodily determinations, as well as the motor system, also play an important part in the processes of thinking, feeling, and willing. Wahle saw in the operations of the brain the antecedents or representatives of conscious processes; to the momentary molecular change of an entire specific brain region corresponds a concrete peculiarity of the given image. The brain, however, is not the "cause" of the mental occurrences or experiences but only the "necessary co-occurrence" of any such occurrence. Both psychopathological phenomena and the origin and formation of character can be understood only physiologically, more particularly from the more or less disturbed (in the case of psychopathology) or undisturbed (in the case of character formation) combined action of a very few elementary brain functions.

Wahle's reflections on the philosophy of culture and history were tinged with skepticism and pessimism, as was his conception of the intellectual capacity and ethical worth of man. Whatever meaning there is in life derives from the existence of love, joy, and pain. Life's highest wisdom is embodied in fulfilling the challenge to be happy with a modesty that is noble, free of illusion, and resigned.

See also Agnosticism; Ethical Subjectivism; Idealism; Metaphysics; Objectivity in Ethics; Pessimism and Optimism; Phenomenalism; Psychoanalysis; Psychology; Realism; Skepticism, History of; Unconscious.

Bibliography

WORKS BY WAHLE

Das Ganze der Philosophic und ihr Ende. Vienna and Leipzig, 1894.

Über den Mechanismus des geistigen Lebens. Vienna and Leipzig, 1906.

Josua. Munich, 1912. A second edition was published in 1928.

Die Tragikomödie der Weisheit. Vienna and Leipzig, 1915; 2nd ed., 1925.

Entstehung der Charaktere. Munich: Drei Masken, 1928.

Grundlagen einer neuen Psychiatrie. Vienna, 1931.

Fröhliches Register der paar philosophischen Wahrheiten (Cheerful catalog of the few philosophical truths). Vienna and Leipzig, 1934.

WORKS ON WAHLE

Flinker, Friedrich. *Die Zerstörung des Ich. Eine kritische Darlegung der Lehre Richard Wahles.* Vienna and Leipzig, 1927.

Hochfeld, Sophus. *Die Philosophie Richard Wahles und Johannes Rehmkes Grundwissenschaft.* Potsdam, 1926.

Franz Austeda (1967)
Translated by Albert E. Blumberg

WALLACE, ALFRED RUSSEL
(1823–1913)

Alfred Russel Wallace, the English naturalist and coformulator with Charles Darwin of the theory of natural selection, was born at Usk, Monmouthshire. He was largely self-educated, having left school at fourteen to serve as a surveyor's assistant with his brother. Like many of his contemporaries he acquired an early taste for the study of nature. But he also read widely and was influenced by the works of Alexander von Humboldt, Thomas Malthus, and Charles Lyell, as Darwin was. In 1844, while teaching school at Leicester, he met the naturalist H. W. Bates (1825–1892), who introduced him to scientific entomology. The two men later embarked on a collecting trip to the Amazon, where Wallace remained for four years examining the tropical flora and fauna.

In 1854, after a brief visit to England, Wallace set out by himself for the Malay Archipelago. He subsequently wrote an account of this trip, *The Malay Archipelago* (London, 1869), which is a fascinating narrative. When he returned in 1862, he had become a convinced evolutionist and was known in scientific circles for his formulation of the theory of natural selection. Another of his scientific contributions was "Wallace's line," a zoogeographical boundary he drew in 1863 to separate Indian and Australian faunal regions, and which was assumed to pass through the middle of the archipelago.

The rest of Wallace's long life was spent in England, except for a lecture tour of the United States in 1887 and short visits to the Continent. Darwin, Lyell, Thomas Henry Huxley, John Tyndall, and Herbert Spencer were among his most intimate friends. He wrote extensively on a wide variety of subjects, but biological interests remained central to his outlook and are reflected in such books as *The Geographical Distribution of Animals* (Lon-

don and New York, 1876), *Darwinism* (London and New York, 1889), *Man's Place in the Universe* (London and New York, 1903), and *The World of Life* (London and New York, 1910).

Wallace first thought of the theory of natural selection in February 1858, when he was ill with a fever at Ternate in the Moluccas. The occasion gave him time to reflect on the mechanism by which species might be altered. He outlined the theory rapidly in a paper, "On the Tendency of Varieties to Depart Indefinitely from the Original Type," and sent it to Darwin, who saw that Wallace had hit upon exactly the theory that he himself had formed and privately written down in 1842. With characteristic generosity he proposed that Wallace's outline should be published immediately. Lyell, however, urged a compromise that resulted in a joint communication from Darwin and Wallace that was read at the Linnaean Society on July 1, 1858. The two men thus received equal credit for the new doctrine, although Darwin was actually the pioneer. The joint communication created no stir at the meeting. However, it was later clearly recognized as a revolutionary document that demolished forever the ancient idea of the fixity of species by formulating a scientific theory of how species change and how their adaptations are secured at each stage of the process.

When Darwin published his famous books, the accord between him and Wallace began to disappear. The view expressed in *The Origin of Species* that evolution required the operation of factors of a Lamarckian as well as of a selective sort was unacceptable to Wallace. For him "natural selection is supreme" and is the sole means of modification, except in the case of man. Hence he became, like August Weissmann, an apostle of neo-Darwinism. This led him to hold that every phenotypic character of an organism must be useful to that organism in the struggle for life; the principle of utility is of universal application.

With regard to human evolution Wallace differed from Darwin in affirming that man's mental powers, especially "the mathematical, musical and artistic faculties," have not been developed under the law of natural selection. These faculties point to the existence in man of something that he has not derived from his animal progenitors, "something which we may best refer to as being of a spiritual essence." It came into action when man appeared on the evolutionary stage. As he grew older, Wallace put more and more emphasis on the spiritual agency, so that in *The World of Life* it is described as "a Mind not only adequate to direct and regulate all the forces at work in living organisms, but also the more fundamental forces of the whole material universe." For many years Wallace was interested in spiritualism and psychical research. A pamphlet that he published in 1866, *The Scientific Aspect of the Supernatural,* discussed such matters as clairvoyance, apparitions, animal magnetism, and the problem of miracles. It was clear that he took them seriously, and they influenced his general outlook. All this was far removed from anything Darwin was prepared to countenance.

Apart from the theory of natural selection, Wallace's most enduring work was his *Geographical Distribution of Animals.* He also made acute judgments on anthropological matters, such as the evolutionary significance of the human brain and human intelligence. Thus he contended that the brain is a specialized organ that has freed man from the dangers of specialization by vastly increasing his adaptability and that man's intelligence has allowed him to evolve without undergoing major somatic changes. Yet despite Wallace's fertility in producing ideas and his command of a wide array of facts, he never quite succeeded in relating the two. His ideas were not carefully analyzed or tested. At bottom he was a naturalist, with a deep love of nature and an inexhaustible passion for collecting.

See also Darwin, Charles Robert; Darwinism; Evolutionary Theory; Huxley, Thomas Henry; Malthus, Thomas Robert; Naturalism; Philosophy of Biology.

Bibliography

ADDITIONAL WORKS BY WALLACE

Contributions to the Theory of Natural Selection. New York: Macmillan, 1870.

My Life: A Record of Events and Opinions. New York: Dodd Mead, 1905.

WORKS ON WALLACE

Eiseley, Loren. *Darwin's Century.* Garden City, NY: Doubleday, 1958.

George, Wilma Beryl. *Biologist Philosopher: A Study of the Life and Writings of Alfred Russel Wallace.* New York: Abelard-Schuman, 1964.

Hogben, L. T. *A. R. Wallace.* London: Society for Promoting Christian Knowledge, 1918.

Marchant, J. *A. R. Wallace: Letters and Reminiscences.* London: Cassell, 1916.

T. A. Goudge (1967)

WANG BI
(226–249 CE)

Third-century Chinese philosopher Wang Bi (226–249 CE) achieved fame as an interpreter of the *Laozi* and the *Yijing* (Classic of changes), whose radical reformulation of the concept of Dao as nonbeing (*wu*) helped spark a new current of thought called *Xuanxue* (Learning of the mysterious), sometimes translated as "neo-Daoism." To Wang, Confucius, Laozi, and the other sages of old had discerned the true meaning of Dao as the root of all beings. This was misunderstood, which necessitated a reinterpretation of the classical heritage.

Wang probed the basis of interpretation and argued that words do not fully express meaning. This was a major debate in early medieval Chinese philosophy. Against earlier commentators who reduced meaning to reference, Wang believed that words are necessary but insufficient for understanding and sought to uncover the fundamental ideas that unite the classics. Famously, Wang declared that words must be forgotten before meaning can be understood.

From this hermeneutical perspective, Wang approaches the meaning of Dao, bringing into view both its transcendence and creative power. According to the *Laozi* (also known as *Daodejing*, the "Classic of the Way and Virtue"), Dao is nameless and formless; yet, it is also the beginning of all things. To Wang, this encapsulates the mystery (*xuan*) of Dao and discloses the central insight that "all beings originate from nonbeing" (*Laozi* commentary, chs. 1, 40).

The *Laozi* states, "Dao gives birth to one," which produces "two" and the rest of creation (ch. 42). Whereas commentators before Wang generally took this to mean that the Dao produced the original "vital energy" (*qi*), which in turn generated the yin and yang energies, Wang focused on the logic of creation. The many can be traced to "one" in the sense of a necessary ontological foundation, but "one" does not refer to any agent or substance. The ground of beings cannot be itself a being; otherwise, infinite regress cannot be overcome. "Beginning" is not a temporal reference but indicates logical priority. "One" is but another term for Dao and should be understood metaphysically as "nonbeing"; "it is not a number," as Wang asserts in his commentary to the *Yijing*, but that which makes possible all numbers and functions. Nonbeing—literally "not having" any property of being—is not a "something" of which nothing can be said; rather, it is a negative concept that sets the Dao categorically apart from the domain of beings and in so doing preserves the transcendence of Dao without compromising its creative power.

The Daoist world reflects a pristine order. This is to be understood in terms of constant principles (*li*) that govern the universe. They do not derive from an external source, but in the light of nonbeing can only be said to be "naturally so" (*ziran*), which Wang describes as "an expression for the ultimate" (*Laozi* commentary, ch. 25). Similarly, human nature should be viewed as "one," understood as what is true (*zhen*) in human beings.

The concept of *ziran* also sets the direction of Daoist ethics and politics. Effortlessly and spontaneously, nature accomplishes its myriad tasks and provides for all beings. In principle, the human world should also be naturally simple, noncontentioius, and self-sufficient. If present realities deviate from this order, it is imperative to recover what is true, to reorient human thinking and action by realizing *ziran*, and in this sense to return to Dao. This is how Wang interprets the key Daoist concept of nonaction (*wuwei*).

Nonaction does not mean total inaction or any esoteric technique to get things done; instead it is a mode of being characterized by the absence of desires, which corrupt one's nature. This, too, follows from the analysis of nonbeing. Genuine well-being can only be measured by the extent to which one is not being fettered by desires, or not having the kind of interest-seeking thought/action that invariably precipitates disorder. Nonaction acts constantly to diminish desires—and to diminish any false sense of self that engenders desires—until one reaches the tranquil depth of emptiness and quiescence. This defines not only the goal of self-cultivation but also that of government.

The order of nature encompasses the family and the state. Their hierarchical structure is rooted in the principles governing the Daoist world. The key to Daoist government lies in "honoring the root and putting to rest the branches." At the policy level, this means not burdening the people with excessive taxation, heavy punishment, and war, which Wang considered the bane of Chinese politics. Following nonaction, the ruler needs only to ensure that obstructions to human flourishing are removed. At a deeper level, desires must be put to rest so that the root may grow; that is, the ruler must embrace emptiness and enable those under the spell of desires to reclaim their true nature.

To many of Wang's contemporaries, the ideal reign of *ziran* can only be realized by a sage, who is utterly different from ordinary human beings in that he is endowed

with an extraordinarily pure *qi*-constitution and is inherently without desires and emotions. Wang Bi, however, argued that the sage is different from ordinary human beings only in terms of his profound "spirituality and enlightenment." In his humanity, the sage "cannot be without sorrow and pleasure to respond to things," but he is not burdened by them. Sage nature signifies complete self-realization.

While standing under tradition—whether in hermeneutics, metaphysics, or concerning the nature of the sage—Wang came to understand it anew. The philosophy of nonbeing made a strong impact on the development of Buddhist philosophy. The concept of *li* (principle) played a pivotal role in later neo-Confucian philosophy. In both instances, Wang's contribution is substantial.

See also Chinese Philosophy: Daoism; Guo Xiang.

Bibliography

Chan, Alan K. L. *Two Visions of the Way*. Albany: State University of New York Press, 1991.

Lynn, Richard J., tr. *The Classic of Changes: A New Translation of the* I Ching *as Interpreted by Wang Bi*. New York: Columbia University Press, 1994.

Lynn, Richard J., tr. *The Classic of the Way and Virtue: A New Translation of the* Tao-te ching *of Laozi as Interpreted by Wang Bi*. New York: Columbia University Press, 1999.

Wagner, Rudolf G. *A Chinese Reading of the Daodejing: Wang Bi's Commentary on the Laozi with Critical Text and Translation*. Albany: State University of New York Press, 2003.

Wagner, Rudolf G. *The Craft of a Chinese Commentator: Wang Bi on the Laozi*. Albany: State University of New York Press, 2000.

Wagner, Rudolf G. *Language, Ontology, and Political Philosophy in China: Wang Bi's Scholarly Exploration of the Dark (Xuanxue)*. Albany: State University of New York Press, 2003.

Alan K. L. Chan (2005)

WANG CHONG

(c. 27–100)

Historically speaking, Wang Chong is one of the best-known thinkers of Han China (221 BCE–220 CE), but the significance of his ideas is far less certain. Wang's native province of Guiji stood on the southeast margins of the Han Empire. Although once studying in the capital Luoyang, he remained basically an obscure local figure. He wrote several books and the most important and only surviving one is the *Lunheng*. This book was not known to the national elite community until the late second century, since then being recognized as a major intellectual work.

Modern opinions split on the nature of the *Lunheng*. Many believe the book reveals Wang as an iconoclast and skeptic who courageously denounced the Confucian orthodoxy and prevalent superstitions. Some, in contrast, consider him a mere rhetorician whose inconsistent arguments seek to justify the existence of people like himself, namely, conceited scholar–officials suffering world failures. The truth probably lies somewhere in between.

Consisting of eighty-five chapters and covering many subjects, the *Lunheng* is not easy to characterize. "*Lun*" means discourse while "*heng*" signifies to weigh or to measure. Wang Chong took the title to mean *discourses as measurements*. This book was thus purported to be a critique of common beliefs. Wang's most obvious target is the so-called theory of "interaction between Heaven and Man." This theory maintains that Heaven regulates, and acts in response to, human behavior. Early Han proponents of Confucianism relied heavily on this theory in their attempt to construct a doctrine as the orthodox ideology for both the state and society. They depicted Heaven as the guardian of Confucian values. It, for instance, punishes human misconduct, particularly that of rulers, by either generating anomalous natural phenomena or bringing down disasters. Wang denied categorically that Heaven was possessed of a will or that the world had any purpose. His critique went beyond a particular theory of heaven. He was deeply opposed to magic itself, especially the kind we now call *sympathetic magic*. This is by no means trivial considering the fact that magic and magical thinking dominated Han life. Wang also found fault with sagely figures, such as Confucius and Mencius. All these critiques earned him the reputation as a great rationalist. There may be some truth to this seemingly anachronistic representation. Wang actually described his project as one to make distinctions between the real and the fanciful although his basis for making such distinctions is sometimes alien to us today.

The *Lunheng* contains evident contradictions in its arguments. The most controversial part of this book is its discussion concerning fate. Whereas denying the existence of a heavenly will, Wang insisted upon predetermined fate. He contended that all human conditions were unavoidable and that the events of an individual's life were in no way related to that person's quality or conduct. He developed complex theories of fate, not unlike a modern economist trying to decipher the invisible hand working in the financial market. Wang's ideas on this sub-

ject were unconvincing to many and opened the door to the charge that his philosophical contentions were largely self-serving.

In terms of writing style in the *Lunheng*, Wang has been accused of being unstructured and redundant. But Wang can be very witty. To give just one example, a famous moral tale relates that upon hearing her husband was killed in war, a woman wailed with such a grief that a city wall collapsed. To this Wang asks: If one cries at water and fire in a state of true grief, can the water be roused to extinguish the fire? In this regard, Wang may be considered a minor Voltaire of early China.

It is easily noticeable that Wang attacked fiercely certain ideas and sayings associated with Confucianism not long after it emerged as the state orthodoxy for the first time in Chinese history. Yet that impression can be misleading. Wang's true target was what he saw as the fanciful thoughts of his time, some of which were used to establish the authority of Confucianism. He had no quarrel with core Confucian values, and indeed promoted the position of Confucian scholars in his book. Despite his rather modest agenda, *xuanxue*—antitraditionalists who arose a century after his death—drew on the *Lunheng* for inspiration. In this peculiar way, Wang helped to bring about a major change in the history of Chinese philosophy.

See also Chinese Philosophy; Confucius; Determinism, A Historical Survey; Mencius; Voltaire, François-Marie-Arouet de.

Bibliography

Fung, Yu-lan. *A History of Chinese Philosophy*. Vol. 2 Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1953.

Satō, Kyōgen. *Ronkō no kenkyū*. Tokyo: Sōbunsha, 1981.

Wang, Ch'ung. *Lun-heng*. 2 vols. Translated by Alfred Forke. Leipzig: 1907–1911. Reprinted, New York: Paragon Book Gallery, 1962.

Jo-shui Chen (2005)

WANG FUZHI

(1619–1692)

Wang Fuzhi was a Chinese philosopher in the late neo-Confucian School. After his initial attempt to resist the Manchu invasion of China had failed, he devoted the rest of his life to the reinterpretation of Chinese philosophical classics and the development of his own philosophical view. The last seventeen years of his life were spent as a hermit at the foot of a barren mountain which he named "the boat mountain" (*chuanshan*); hence his well-known alias: Wang Chuanshan. His copious works were first published posthumously by his son. Most notable among his works are: *Du Sishu Daquan Shuo* (Discourse on reading the great collection of commentaries on the four books), *Zhouyi Waizhuan* (External commentary on the book of changes), *Zhouyi Neizhuan* (Internal commentary on the book of changes), *Du Tongjian Lun* (A treatise on reading Tongjian), and *Zhuangzi Zhengmeng Zhu* (Commentary on Zhang Zai's zhengmeng).

Wang Fuzhi's metaphysics places the cosmic principle (*li*) in the midst of cosmic energy (force; *qi*), thereby denying any transcendent status of the cosmic principle. The universe is constituted by *qi*, which develops in accordance with a certain order. According to Wang Fuzhi, this order does not exist prior to the development of *qi*; it is simply "the way things are" as well as "the way things ought to be" for cosmic energy. *Qi* is self-regulating in virtue of this internal cosmic principle; therefore, *qi* is not a blind force. Wang Fuzhi not only acknowledges the orderliness of *qi*, but also recognizes the all-encompassing nature of *qi*. The universe is filled with *qi* from time immemorial; cosmic states are simply the different developmental stages of *qi*. When *qi* condenses, it composes myriad things; when material objects disintegrate, everything returns to the rarified form of *qi*. In this respect, his metaphysics follows directly from that of Zhang Zai.

In addition to advocating the unity between principle and *qi*, Wang Fuzhi also espouses the unity between Dao and concrete things (*qi*—a different word from the cosmic energy *qi*). Dao is the way particular things are and the way they ought to be. According to Wang Fuzhi, Dao does not have any a priori status; it does not exist independently of concrete things. In other words, Dao is postdevelopmental in the production of concrete things, just as cosmic principle (*li*) is postdevelopmental in the activities of *qi*. To Wang Fuzhi, only the concrete cosmic energy (*qi*), and the concrete objects composed of *qi*, are real. His metaphysics has often been interpreted as a form of materialism and realism.

Because *qi* constantly evolves and transforms itself, the universe perpetually generates and renews itself. When applied to the human world, this cosmology entails that human history is not predetermined. Wang Fuzhi's philosophy of history is modernistic in spirit, for he holds that the modern is more advanced than the ancient; ancient laws and morals do not necessarily apply to the

contemporary world. To find the best way to govern, people need to deal with the present context and understand the present societal needs. A good ruler is one who understands and aims to meet his or her people's wants and desires. Following Mencius, Wang Fuzhi argues that people's common desire is nothing but the satisfaction of their basic needs in life. These desires are natural to human beings; they are thus not morally blameworthy.

Wang Fuzhi rejects Buddhists' renouncement of human desires; he also criticizes the Cheng-Zhu School's doctrine that one needs to extinguish human desires in order to exemplify the Heavenly principle. He advocates the unity of the Heavenly principle and human desires: the principle of heaven lies in nothing but what the people desire in common. An ideal state of the world is reached when all people can have their basic desires satisfied. To Wang Fuzhi, human history is simply a reflection of human nature; human politics is solely determined by what the people want in common. This view reaffirms the Confucian humanism underlined in classic Confucianism.

See also Chinese Philosophy: Confucianism.

Bibliography

MAJOR WORKS BY WANG FUZHI

Zhouyi waizhuan (External commentary on the book of changes), 1655.

Laozi yan (Extended interpretation on Laozi), 1655.

Shangshu yinyi (The extended meaning of the book of history), 1663.

Du Sishu daquan shuo (Discourse on reading the great collection of commentaries on the four books), 1665.

Chunqiu shilun (A general treatise on Chun Qiu), 1668.

Liji zhangju (A textual annotation on the book of rites), 1673–1677.

Zhouyi daxiang jie (Interpretation on the images of the book of changes), 1676.

Siwen lu (Record of thoughts and questions), post 1677.

Zhangzi Zhengmeng zhu (Commentary on Zhang Zai's zhengmeng), post 1677.

Sishu xunyi (A contemporary interpretation of the meaning of the four books), 1679.

Zhuangzi jie (Interpretation on Zhuangzi), 1681.

Zhouyi neizhuan (Internal commentary on the book of changes), 1685.

Du Tongjian lun (A treatise on reading Tongjian), 1687–1691.

Song lun (A treatise on the Song dynasty), 1691.

WORKS ON WANG FUZHI

Black, Alison Harley. *Man and Nature in the Philosophical Thought of Wang Fu-Chih*. Publications on Asia of the Henry M. Jackson School of International studies, No 41. Seattle, WA: University of Washington Press. 1989.

Hou, Wai-lu, and Chi-chih Chang. "Philosophical thought of Wang Fu-chih." *Chinese Studies in History and Philosophy* 1 (Spring 1968): 12–28

Liu, JeeLoo. "Is Human History Predestined in Wang Fuzhi's Cosmology?" *Journal of Chinese Philosophy*28(3), 2001, 321–337.

Liu, JeeLoo. "Wang Fuzhi." In *The Encyclopedia of Chinese Philosophy*, edited by Antonio S. Cua. New York: Routledge, 2003, 748–755.

McMorran, Ian. "Late Ming Criticism of Wang Yang-Ming: The Case Of Wang Fu-Chih." *Philosophy East & West* 23, 1973, 91–102.

McMorran, Ian. *The Passionate Realist: An Introduction to the Life and Political Thought of Wang Fuzhi, (1619–1692)*. Hong Kong: Sunshine, 1992.

Teng, S.Y. "Wang Fu-chih's Views on History and Historical Writing." *Journal of Asian Studies* 28, 111–123.

Wright, Kathleen. "The Fusion of Horizons: Hans-Georg Gadamer and Wang Fu-Chih." *Continental Philosophy Review* 33 (3), 2000: 345–358.

Zhang, Jiemo. "On the Aesthetic Significance of Wang Fuzhi's Theory of the Unity of Poetry and Music, with Criticisms of Certain Biases." *Chinese Studies in Philosophy* 21(3), Spring 1990, 26–53.

Jeeloo Liu (2005)

WANG SHOU-JEN

See *Wang Yang-ming*

WANG YANG-MING

(1472–1529)

Wing-tsit Chan reminds the reader that "the philosophy of Wang Yang-ming is a vigorous philosophy born of serious searching and bitter experience" (1963, Chan's introduction, p. ix). Wang's doctrine of the unity of knowledge and action, for example, may be regarded as a forceful and concise way of stating the unity of his life and teaching during his formative years. For Wang, learning to become a sage involved a serious and resolute commitment to Dao or *ren* (humanity)—the ideal of "forming one body" with all things in the universe. Says Wang: "The great man regards Heaven, Earth, and the myriad things as one body (*yiti*). Moreover, the *ren*-person also forms one body with plants, stones, tiles, mountains, and rivers" (1963, p. 272).

Alternatively, one may characterize Wang's vision of the highest good as an ideal of the universe as a harmonious moral community. A commitment to the vision of *ren* is a commitment to the task of clarifying the concrete significance of the vision—an ideal theme rather than an

ideal norm as a basis for deriving precepts. An ideal theme is a unifying perspective, a point of orientation, not a fixed principle of conduct. For expressing his vision Wang sometimes used the term Dao (way) instead of *ren*. Dao and *ren* differ in the direction of stress. On the one hand, *ren* stresses the significance of Wang's moral vision as residing in affectionate human relationships, a habitat that is capable of indefinite expansion and ultimately embraces the whole universe. Dao, on the other hand, stresses the ongoing course of changing circumstances that calls for an exercise of the agent's sense of rightness (*yi*). The unlimited possibilities of the concrete significance of Dao cannot be exhausted with any claim to finality (*dao wu zhongqiong*). Notably Wang sometimes uses the term *tianli* (heavenly principle, pattern, rationale) to express his vision of the highest good. *Tianli* is inherent in *xin* (heart/mind); often it is obscured by the presence of selfish desires.

Except for its ethical significance, Wang shows little interest in the pursuit of factual knowledge. Unlike Zhu Xi (1130–1200), who emphasizes the significance of *li* (principle, pattern, or rationale) in the investigation of things (*gewu*) in the Great Learning, Wang focuses instead on the rectification of the mind (*zhengxin*) that deviates from his moral vision. Rectification of the mind involves, in particular, an acknowledgment of the unity of moral knowledge and action (*zhixing heyi*), an enlargement of the scope of moral concern in the light of the vision of *ren*, rather than extensive acquisition of factual knowledge.

Wang's doctrine of the unity of knowledge and action is sometimes stated as the unity of moral learning and action (*xuexing heyi*). Wang's discussion involves two different senses of *zhi*, corresponding to two senses of knowledge. For convenience, this entry will use the distinction between prospective and retrospective moral knowledge—that is, knowledge acquired anterior or prior to action and knowledge posterior to action.

Prospective moral knowledge, for the most part, is a product of learning, an acknowledgment of the projective significance of the standards embedded in the various notions of Confucian virtues. Prospective moral knowledge is implicit in Wang's compendious remark that "knowledge is the direction of action and action is the effort of knowledge" (1963, p. 11) As prospective knowledge, and by virtue of its cognitive content, it provides a direction or a leading idea (*zhuyi*) for actual conduct. Another compendious remark appears to make use of both prospective and retrospective senses of moral knowledge: "knowledge is the beginning of action and action is the completion of knowledge" (1963, p. 11). Wang's emphasis on personal realization of his moral vision is an emphasis on retrospective moral knowledge. For Wang, the transition from prospective to retrospective knowledge involves a variety of intellectual acts (inquiry, understanding, sifting, or discrimination) and volitional acts (involving resolution, intention, moral desire, and the purity of moral motives in the endeavor to achieve the ideal of *ren*). More especially, in his mature thought, Wang constantly focused on extending *liangzhi*, commonly rendered as "innate or intuitive knowledge of the good."

Liangzhi, in the sense of the ability of moral discrimination, while basic, cannot capture the depth of Wang's concern in his teaching of extending *liangzhi*. While the human mind is in the rudimentary sense consciousness, without a commitment to the vision of *ren*—alternatively to Dao or *tianli*—it would be indifferent to moral concern. Possessed of *liangzhi*, the human mind as informed by the vision will be distinctively marked as moral consciousness. As Wang was wont to say, it is *liangzhi* that manifests *tianli* or *liangzhi* that manifests Dao. As the intrinsic quality (*benti*) of the moral mind, *liangzhi* is "naturally intelligent, clear and unbeclouded" (1963, p. 274). This notion of *liangzhi* as the seat of moral consciousness does involve *liangzhi* in the sense of moral discrimination, and significantly stresses the exercise of clear intelligence in discerning the moral import of particular situations. As embodying the concern for *tianli*, *liangzhi* is properly considered a personal standard; that is, a standard for making autonomous judgment of the moral quality of thought and actions, as well as feelings. Thus Wang's notion of *liangzhi* cannot be understood apart from his vision and confidence in the mind as possessing its own capability of realizing the vision.

Liangzhi, being an active concern of the moral mind with *tianli*, clearly involves the determination to its actualization. As embodying this active concern with *tianli*, *liangzhi* cannot be rendered as *intuition*, as this term is used in Ethical Intuitionism. Genuine perplexity arises in changing or exigent circumstances, where established standards do not provide clear guidance (Cua 1982, ch. 3). While *liangzhi* is inherent in all minds, the distinguishing characteristic of the sage lies in his or her attitude toward study and reflection. As invested with *tianli*, *liangzhi* is indeed a standard, but it does not issue recipes for coping with changing circumstances. Wang believed that *liangzhi* can provide unerring guidance, but it is unclear how he could account for failure in extending *liangzhi* and the relation between moral and factual

knowledge. Focus on the nature of retrospective moral knowledge and experience may provide a critical point of departure for developing the notion of *liangzhi* in Confucian ethics.

See also Confucianism; Zhu Xi.

Bibliography

Chang, Carsun. *The Development of Neo-Confucian Thought*. Vol. 2. New York: Bookman Associates, 1962.

Ching, Julia. *To Acquire Wisdom: The Way of Wang Yang-ming.* New York: Columbia University Press, 1976.

Cua, A. S. "Between Commitment and Realization: Wang Yang-ming's Vision of the Universe as a Moral Community." *Philosophy East and West* 43 (4) (1993): 611–49.

Cua, A. S. *Moral Vision and Tradition: Essays in Chinese Ethics*. Washington, DC: The Catholic University of America Press, 1998.

Cua, A. S. *The Unity of Knowledge and Action: A Study in Wang Yang-ming's Moral Psychology*. Honolulu: University Press of Hawaii, 1982.

Deustch, Elliot, et al. "Proceedings of East-West Philosophers' Conference on Wang Yang-ming: A Comparative Study." *Philosophy East and West* 23 (1, 2) (1973).

Tu, Wei-ming. *Neo-Confucian Thought in Action: Wang Yang-ming's Youth (1472–1509)*. Berkeley: University of California Press, 1976.

Wang Yang-ming. *Instructions for Practical Living and Other Neo-Confucian Writings*. Translated by Wing-tsit Chan. New York: Columbia University Press, 1963.

Wang Yang-ming. *The Philosophical Letters of Wang Yang-ming*. Translated by Julia Ching. Columbia: University of South Carolina Press, 1972.

Antonio S. Cua (2005)

WAR

See *Peace, War, and Philosophy*

WATSUJI TETSURŌ

(1889–1960)

Watsuji Tetsurō, the best philosopher of ethics of modern Japan, was known also for his studies of cultural history. He was born in Himeji and died in Tokyo. Watsuji's work can be divided into three categories: his early literary efforts, his philological and historical studies, and his works on an ethical system. Gifted with literary talent, he wrote some short novels and a play while still studying philosophy, but these had no great success. Among his early philosophical essays are those on Friedrich Nietzsche (Tokyo, 1913) and on Søren Kierkegaard (Tokyo, 1915). His cult of ancient Greece, manifested in *Gūzō saikō* (The revivals of the idols; Tokyo, 1918), developed into an interest in the cultural history of his own country. His first work on this subject was *Nihon kodai bunka* (Ancient Japanese culture; Tokyo, 1920). Japanese culture and character were to be the subject of his constant study, as was attested by his *Nihon seishin-shi* (The history of Japanese spirit; 2 vols., Tokyo, 1926, 1934). Meanwhile, his other studies, based on philological research, covered the textual questions about Homer, primitive Christianity, early Buddhism, and Confucius. While these works differ in scientific value, they contain many insights and reveal him as more a litterateur than an expert philologist and historian. This is obvious in his well-known *Fūdo* (Tokyo, 1934; translated as *A Climate,* 1961), a work of psychological intuition and deep sensibility rather than a scientific or philosophical study of the conditioning effect of climate on culture.

A turning point in his career was his appointment as assistant professor of ethics at Kyoto University (1925). Out of his lectures at Kyoto grew his *Ningengaku toshite no rinrigaku* (Ethics as anthropology), a treatise of systematic ethics, initiated in 1931. Watsuji's ethic was designed as a Japanese system based upon the essential relationships of man to man, man to family, and man to society. In contrast with the private, individual ethics of the West, his ethic sees man as involved in community and society. *Rinri* (ethics) in Sino-Japanese characters meant for him the principle (*ri*—or *li* in Chinese) of companionship (*rin*). Furthermore, he introduced the Buddhist dialectic elements (negation of negation) to show how the individual is absorbed into the whole. It is true that in postwar years he rewrote the parts of his ethics concerning the state and the emperor. Yet his achievement was that he systematized—although in Western categories—a traditional ethics that is a substantial part of the ethos of Japan and also of China. His attitude toward East-West contacts may be surmised from his *Sakoku Nihon no higeki* (National seclusion, Japan's tragedy; Tokyo, 1951). His two-volume *Nihon rinri shisō-shi* (History of Japanese ethical thought; Tokyo, 1952) is a major contribution to the subject. Western philosophers who had a great influence upon Watsuji were Edmund Husserl and Martin Heidegger.

See also Buddhism; Confucius; Ethics, History of; Heidegger, Martin; Homer; Husserl, Edmund; Japanese Philosophy; Kierkegaard, Søren Aabye; Nietzsche, Friedrich.

Bibliography

Japanese primary sources include *Watsuji Tetsurō zenshū* (The complete works of Watsuji Tetsurō), 27 vols. (Tokyo: Iwanami Shoten, 1961–1992), and *Jijoden no kokoromi* (An attempt in autobiography; Tokyo, 1961).

Watsuji's *Fūdo* has been translated into English as *A Climate* by G. Bownas (Tokyo, 1961). A discussion of his work may be found in G. K. Piovesana, *Recent Japanese Philosophical Thought, 1862–1962* (Tokyo: Enderle Bookstore, 1963), pp. 131–145.

Gino K. Piovesana, S.J. (1967)

WAYLAND, FRANCIS

(1796–1865)

Francis Wayland, the American Baptist clergyman, educator, and moral philosopher, was one of the central figures in the modification of American collegiate education. As president of Brown University (1827–1855), he introduced proposals to ease the rigidity of the classical curriculum by an approximation of the later elective system. With his mentor, Eliphalet Nott of Union College, Schenectady, New York, Wayland approved of the substitution of modern language study for at least some of the required Greek and Latin, encouraged training in science and its practical application, and advocated a more professional faculty employed for longer terms. To some degree his interest in these reforms was the result of his Jeffersonian philosophy of democracy. He was completely in accord with Thomas Jefferson's insistence that a republican government can flourish only if the voters are well educated. He argued, too, that native talent was widely diffused and should be given the opportunity to develop through education.

Philosophically, Wayland was a naive realist of the Scottish school of philosophy. His theory of knowledge was basically Lockean sensationalism supported by a faculty psychology. Knowledge is gained by a combination of experience and intuition, leading to inductive generalizations whose certainty he did not question. Ultimately Wayland's epistemology rests upon a theistic assumption, that there is a correspondence between what man finds in the universe and what God put there for man to find. However, Wayland's most important contribution to American philosophic development was moral rather than epistemological. His textbook, *The Elements of Moral Science,* first published in 1835, was very widely used and served as a model for many imitators. In this book Wayland departed from the William Paley form of utilitarian ethics that had been taught in the colleges and introduced an ethical position more dependent upon the deontological position characteristic of Bishop Butler. The Enlightenment emphasis on the rights of man was subordinated to a philosophicoreligious stress upon ethics as a system of duties. The moral quality of an action is declared to reside in its intention rather than in its consequences.

Wayland's moral theory led him to an increasing rejection of the institution of slavery. At first he found intolerable only the thought of being himself a slave owner; later he came to feel that all property in human beings was intolerable. From a mildly antislavery position in 1835, he moved to vigorous abolitionism and support of the Union cause in the Civil War. To at least some of the Southern defenders of slavery, Wayland became the archenemy, particularly because of his insistence that the Scriptures cannot be used to support the institution of slavery. Wayland's exchange of letters with Richard Fuller, a Southern clergyman, published as *Domestic Slavery Considered as a Scriptural Institution* (New York and Boston, 1845), presents the arguments on both sides most effectively.

See also Butler, Joseph; Enlightenment; Jefferson, Thomas; Paley, William; Philosophy of Education, History of; Realism.

Bibliography

Wayland's *The Elements of Moral Science* has appeared in the John Harvard Library (Cambridge, MA: Belknap Press of Harvard University Press, 1963) with an extended introduction by the editor, J. L. Blau. Wayland's other major works are *The Elements of Political Economy* (New York: Leavitt Lord, 1837), *Thoughts on the Present Collegiate System in the United States* (Boston: Gould, Kendall and Lincoln, 1842), and *The Elements of Intellectual Philosophy* (Boston and New York, 1854). For discussions of Wayland see J. L. Blau, *Men and Movements in American Philosophy* (Englewood Cliffs, NJ: Prentice-Hall, 1952) and Wilson Smith, *Professors and Public Ethics* (Ithaca, NY: American Historical Association and Cornell University Press, 1956).

J. L. Blau (1967)

WEAKNESS OF THE WILL

The primary philosophical topic explored under the rubric "weakness of the will" is roughly what Aristotle called *akrasia*. This classical Greek term is formed from the alpha privative (basically, a negation sign) and *kratos*, meaning "strength" or "power." The power at issue is the power to control oneself in the face of actual or antici-

pated temptation. So *akrasia* is deficient self-control. Self-control, in this sense, may be understood as constituted primarily by a robust capacity to see to it that one does what one believes to be best on the whole when tempted to do otherwise. The self-controlled person, Aristotle writes, "is in such a state as … to master even those [temptations of a certain kind] by which most people are defeated," and the akratic person "is in such a state as to be defeated even by those … which most people master" (*Nicomachean Ethics* 1150a11–13).

In Plato's *Protagoras*, Socrates says that the common view about akratic action is that "many people who know what it is best to do are not willing to do it, though it is in their power, but do something else" (352d). Here he raises (among other issues) the central question in subsequent philosophical discussion of *akrasia*: Is strict akratic action possible? Strict akratic action may be defined as free, intentional action that is contrary to a conscious belief that the agent has at the time to the effect that it would be best to *A* (or best not to *A*)—best from the perspective of his own values, desires, beliefs, and the like, as opposed, for example, to a common evaluative perspective that he does not endorse. In this entry, I call beliefs with all the properties just mentioned *P beliefs*.

A feature of paradigmatic strict akratic actions that is typically taken for granted and rarely made explicit is that the *P* beliefs with which they conflict are rationally acquired. In virtue of clashing with the agent's rationally acquired *P* beliefs, akratic actions are subjectively irrational (to some degree, if not without qualification). There is a failure of coherence in the agent of a kind directly relevant to assessments of the agent's rationality. This kind of failure would be exhibited, for example, by a student who freely goes to a party tonight even though he or she has a *P* belief that it would be best not to go and to study instead.

To some theorists (e.g., R. M. Hare, Socrates, and Gary Watson), the threat that strict akratic action poses to our ability to make sense of human action seems so severe that they deem strict akratic action conceptually or psychologically impossible. Many others, including Donald Davidson, Alfred Mele, David Pears, and Amelie Rorty, try to accommodate strict akratic action in a general theory of human action.

SKEPTICISM ABOUT STRICT AKRATIC ACTION

For the purposes of this entry, it may be assumed (*P1*) that people sometimes act freely and (*P2*) that people sometimes perform intentional actions that are contrary to their *P beliefs*. Some compulsive hand-washers or crack cocaine addicts may occasionally confirm *P2*. But acting contrary to one's *P* belief is not sufficient for acting akratically; one's action must also be free. Some philosophers argue that strict akratic action is impossible because actions contrary to the agent's *P* beliefs are necessarily unfree.

Assumptions *P1* and *P2* and the following assertion form a consistent triad: (*UF*) All actions contrary to the agent's *P* belief are unfree. How might a philosopher try to defend *UF* while granting *P1* and *P2*? Here is a sketch of one such defense (Harepresents a similar argument in chapter 5):

Argument A

A1. Having a *P* belief that it is best to *A* now is conceptually sufficient for having an intention to *A* now.

A2. Any agent who intends to *A* now but does not *A* now is unable to *A* now.

A3. Such an agent, being unable to *A* now, is compelled to perform—and therefore unfreely performs—whatever pertinent intentional action he now performs.

Premise *A2* is falsified by simple counterexamples. A professional pitcher who intends to throw a pitch in the strike zone may accidentally miss even though he was able to do what he intended. Of course, the failures in alleged strict akratic actions may be different in important ways, and it may be claimed that *A2* simply needs to be revised to capture the difference. One likely suggestion is that in alleged strict akratic actions, the failure involves a change of intention—for example, a change from intending to study to intending to attend a party—whereas the pitcher's failure does not. Now, either the change of intention is paired with a corresponding change of belief or it is not. If there is a change of belief that matches the change of intention—for example, a change to believing that it would be better to attend the party—then the agent does not act contrary to his current *P* belief in executing that intention. But it is assumed that some actions are contrary to their agents' current *P* beliefs, and the skeptic is supposed to be arguing that all such actions are unfree. So suppose that the change of intention is not paired with a corresponding change of belief and that the agent's *P* belief persists. Then *A1* is false. It is falsified by an agent who had intended in accordance with a *P* belief but no longer so intends even though the belief persists.

A1 is in dire straits anyway, given *P2*. Consider compulsive hand-washers or crack cocaine addicts who believe that it is best not to wash their hands now or not to use crack now, but who do so anyway—intentionally and unfreely. If *A1* is true, they are intentionally washing their hands or using crack while intending not to do so. Although this may be conceptually possible—for example, perhaps an agent with a split brain may intend not to *A* while also intending to *A* and acting on the latter intention—it is a highly implausible hypothesis about representative cases of the kind at issue. A much more plausible hypothesis is that although the troubled agents believe that it would be best not to wash their hands now or not to use crack now, they lack a corresponding intention and instead intend to do what they are doing.

A3 also is problematic. Bob has been dieting and believes it best to order a low-calorie salad for lunch today. Unfortunately, he is tempted by several other items on the menu, including a hamburger, a steak, and a pork sandwich. He orders the steak. Even if Bob was unable to order the salad, we would need an argument that he was compelled to order the steak—that, for example, ordering the burger was not a live option.

Gary Watson offers the following argument for *UF*:

Argument B

B1. An agent's succumbing to a desire contrary to his *P* belief cannot be explained by his choosing not to resist nor by his making a culpably insufficient effort to resist.

B2. Only one explanation remains: The agent was unable to resist.

So *UF*. All actions contrary to the agent's *P* belief are unfree.

Watson argues that an agent's choosing not to resist cannot explain strict akratic action, for to make such a choice "would be to change" one's *P* belief (p. 337). For example, "The weak drinker's failure to resist her desire to drink is a failure to implement her choice not to drink. To choose not to implement this choice would be to change her original judgment, and the case would no longer be a case of failure to implement a judgment" (pp. 336–337). Watson also contends that an insufficient effort cannot be due to a belief that the effort is not worth the trouble, since the belief that it is worth the trouble is implicit in the violated *P* belief (p. 338). Nor, he argues, can the insufficient effort be explained by a misjudgment of "the amount of effort required," for misjudgment is "a different fault from weakness of will" (p. 338).

In some alleged instances of strict akratic action, agents believe that it would be best to *A*, choose accordingly, and then backslide while retaining that belief. In others, agents with the same *P* belief do not choose accordingly; they do not make the transition from belief to intention. Although Watson has the former kind of case in mind, it is useful to attend to a case of the latter kind. Imagine, if you can, that a drinker, Drew, who has had one shot of bourbon and needs to drive home soon, believes that it would be best to switch now to coffee but neither chooses nor intends to do so and intentionally drinks another bourbon. The reader is not asked to imagine that Drew akratically drinks the second bourbon; it is left open that she drinks it unfreely. If Drew can believe that it would be best not to drink a second bourbon without choosing accordingly, then she can fail "to resist her desire to drink" without there being any failure on her part "to implement her choice not to drink." If she makes no such choice, she does not fail to implement it. And if there is no such failure of implementation, then the reason Watson offers for maintaining that the agent "change[d] her original judgment" is undercut.

A scenario in which a belief-matching choice is made will be discussed shortly. The plausibility of scenarios of the present sort deserves a bit more attention now. Consider the following story. On New Year's Eve, Joe, a smoker, is contemplating kicking the habit. Faced with the practical question of what to do about his smoking, he is deliberating about what it would be best to do about it. He is convinced that it would be best to quit smoking sometime, but he is unsure whether it would be best to quit soon. Joe is under a lot of stress, and he worries that quitting smoking now might drive him over the edge. Eventually, he judges that it would be best to quit by midnight. But he is not yet settled on quitting. Joe tells his partner, Jill, that he has decided that it would be best to stop smoking, beginning tonight. Jill asks, "So is that your New Year's resolution?" Joe sincerely replies, "Not yet; the next hurdle is to decide to quit. If I can do that, I'll have a decent chance of kicking the habit."

This story at least has the appearance of coherence. Seemingly, although Joe decides that it would be best to quit smoking, he may or may not choose (i.e., form the intention) to quit. Watson offers no argument for the incoherence of stories of this kind. (It has not been claimed that Joe is a free agent.)

If Drew can fail to resist her desire for a second bourbon without changing her belief about what it is best to do, what about Lucy, who, like Drew, takes another bourbon despite believing that it would be best to switch now

to coffee, but, unlike Drew, chooses to switch now to coffee when she makes her judgment? Watson would say (*W1*) that Lucy's "failure to resist her desire to drink [a second bourbon] is a failure to implement her choice not to drink," (*W2a*) that "to choose not to implement this choice [is] to change her original judgment," (*W2b*) that to choose not to resist her desire to drink a second bourbon is to change that judgment, and (*W3*) that Lucy's drinking the second bourbon is therefore not a strict akratic action, since it is not contrary to her *P* belief (pp. 336–337). Is *W2a* or *W2b* true? Watson offers no argument for either, and some stories in which analogues of both are false certainly seem coherent.

Here is one such story. Alex's friend, Bob, has proposed that they affirm their friendship by becoming blood brothers, since Alex is about to go away to prep school. The ceremony involves the boys' cutting their own right palms with a pocket knife and then shaking hands so that their blood will mingle. Alex is averse to cutting himself, but he carefully weighs his reasons for accepting the proposal against his competing reasons (including his aversion), and he judges that it would be best to accept the proposal and to perform the ceremony at once. He chooses, accordingly, to cut his hand with the knife straightaway. Without considering that he may find the task difficult, he grasps the knife and moves it toward his right palm with the intention of drawing blood. However, as he sees the knife come very close to his skin, he intentionally stops because of his aversion. He chooses not to implement his original choice just now, and he chooses not to resist his aversion further just now. Alex abandons his original choice. But he has not changed his mind about what it is best to do, and he is upset with himself for chickening out. (Soon, Alex resolves to try again, this time without looking. The second attempt succeeds.)

If this story is incoherent, Watson should explain why. If he were to assent to *A1*, he could appeal to it here: since Alex no longer intends to cut his hand straightaway, it would follow that he no longer believes that it would be best to cut it straightaway. But Watson rejects *A1* to accommodate compulsives who act contrary to a *P* belief.

EXPLAINING STRICT AKRATIC ACTION

Imagine that although Jack believes that it would be better to study tonight for tomorrow's test than to attend a friend's party, he goes to the party and does not study. To the extent that his belief is sensitive to his motivational states (e.g., his desire to get a decent grade on the test), it has a motivational dimension. That helps explain why strict akratic action is regarded as theoretically perplexing. How, some philosophers wonder, can the motivation that is directly associated with a belief of this kind—in this case, Jack's motivation to study—be outstripped by competing motivation, especially when the competing motivation (a desire to have fun tonight) has been taken into account in arriving at the belief?

One answer (defended in Mele 1987) rests partly on the following two theses and on various arguments for those theses.

> *P* beliefs normally are formed at least partly on the basis of our evaluation of the objects of our desires (i.e., the desired items).
>
> The motivational force of our desires does not always match our evaluation of the objects of our desires.

If both theses are true, it should be unsurprising that sometimes, although we believe it better to *A* than to *B*, we are more strongly motivated to *B* than to *A*. Given how our motivation stacks up, it should also be unsurprising that we *B* rather than *A*.

Thesis 1 is a major plank in a standard conception of practical reasoning. In general, when we reason about what to do, we inquire about what it would be best, or better, or good enough, to do, not about what we are most strongly motivated to do. When we ask such questions while having conflicting desires, our answers typically rest significantly on our assessments of the objects of our desires—which may be out of line with the motivational force of those desires, if thesis 2 is true.

Thesis 2 is confirmed by common experience and thought experiments and has a foundation in empirical studies. Desire-strength is influenced not only by our evaluation of the objects of desires, but also by such factors as the perceived proximity of prospects for desire-satisfaction, the salience of desired objects in perception or in imagination, and the way we attend to desired objects (as Ainslie, Metcalfe and Mischel, and others have observed). Factors such as these need not have a matching effect on assessment of desired objects.

Empirical studies of the role of representations of desired objects in impulsive behavior and delay of gratification (reviewed in Mele 1995) provide ample evidence that our representations of desired objects have two important dimensions, a motivational and an informational one. Our *P* beliefs may be more sensitive to the informational dimension of our representations than to the motivational dimension, with the result that such beliefs sometimes recommend actions that are out of line

with what we are most strongly motivated to do at the time. If so, strict akratic action is a real possibility—provided that at least some intentional actions that conflict with agents' *P* beliefs at the time of action are freely performed. To be sure, it has been argued that no such actions can be free, but, as the preceding section indicates, representative arguments for that thesis are unpersuasive.

Unless a desire of ours is irresistible, it is up to us, in some sense, whether we act on it, and it is widely thought that relatively few desires are irresistible. Arguably, in many situations in which we act against our *P* beliefs, we could have used our resources for self-control in effectively resisting temptation. Normal agents can influence the strength of their desires in a wide variety of ways. For example, they can refuse to focus their attention on the attractive aspects of a tempting course of action and concentrate instead on what is to be accomplished by acting as they judge best. They can attempt to augment their motivation for performing the action judged best by promising themselves rewards for doing so. They can picture a desired item as something unattractive—for example, a wedge of chocolate pie as a wedge of chewing tobacco—or as something that simply is not arousing. Desires normally do not have immutable strengths, and the plasticity of motivational strength is presupposed by standard conceptions of self-control. Occasionally, we act contrary to our *P* beliefs, and it is implausible that, in all such cases, we are unable to act in accordance with those beliefs.

The key to understanding strict akratic action is a proper appreciation of the point that the motivational force or causal strength of a motivational attitude need not be in line with the agent's evaluation of the object of that attitude. Our *P* beliefs are based, in significant part, on our assessments of the objects of our desires; and when assessment and motivational force are not aligned, we may believe it better to *A* than to *B* while being more strongly motivated to *B* than to *A*. If while continuing to have that belief, we freely do *B*, our action is strictly akratic.

See also Aristotle; Davidson, Donald; Hare, Richard M.; Plato; Power; Socrates.

Bibliography

Ainslie, George. *Breakdown of Will*. Cambridge, U.K.: Cambridge University Press, 2001.

Aristotle. *Nicomachean Ethics*. In *The Complete Works of Aristotle: The Revised Oxford Translation,* edited by Jonathan Barnes. Princeton, NJ: Princeton University Press, 1984.

Charlton, William. *Weakness of Will*. Oxford: Basil Blackwell, 1988.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Dunn, Robert. *The Possibility of Weakness of Will*. Indianapolis, IN: Hackett, 1987.

Hare, R. M. *Freedom and Reason*. Oxford: Oxford University Press, 1963.

Mele, Alfred R. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press, 1995.

Mele, Alfred R. *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. New York: Oxford University Press, 1987.

Metcalfe, Janet, and Walter Mischel. "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower." *Psychological Review* 106 (1999): 3–19.

Pears, David. *Motivated Irrationality*. Oxford: Oxford University Press, 1984.

Plato. *Protagoras*. Translated by C. C. W. Taylor. Oxford: Clarendon Press, 1976.

Rorty, Amelie. "Where Does the Akratic Break Take Place?" *Australasian Journal of Philosophy* 58 (1980): 333–346.

Watson, Gary. "Skepticism about Weakness of the Will." *Philosophical Review* 86: 316–339.

Alfred R. Mele (2005)

WEBER, ALFRED

(1868–1958)

The German sociologist and philosopher of history Alfred Weber, like his older brother Max, studied law and political economy in preparation for a legal career and later changed to sociology and university teaching. Alfred Weber's academic career began in 1899 at the University of Berlin and continued at the University of Prague (1904), where he came into contact with Tomáš Masaryk, then professor of sociology. From 1907 to 1933, Weber held a professorship at Heidelberg; in 1933 he resigned at the rise of the Hitler regime. It was due largely to him that the Heidelberg Institute of Social Sciences became one of the chief centers of sociopolitical research during the Weimar Republic, and under his direction it regained its renown after World War II.

Having established his reputation as an economic sociologist by the publication in 1909 of his work on the location of industry (*Über den Standort der Industrien*), Weber turned to historical and cultural-sociological studies, culminating in his main work, *Kulturgeschichte als Kultursoziologie* (1935). In this work he attempted to discover by sociological analysis the chief structural constituents of the historical process. These constituents he distinguished as the social process, the civilization

process, and the culture process; although he distinguished between them, he emphasized their relatedness within the diverse constellations of a given historical continuum. By "social process" Weber understood the reoccurrence of certain societal sequences that, notwithstanding individual variations, reveal sufficient uniformity to provide the basis for a comparative study of different peoples. As an example of such a social process, Weber cited the succession from kinship organization to territorial groupings in diverse sociohistorical entities. The "civilization process" was for him essentially the growth of knowledge concerning the techniques of controlling natural and material forces. Weber regarded the discovery of these techniques as a continuous and cumulative progress permitting, by virtue of the transferability of such knowledge, an element of homogeneity amid the otherwise heterogeneous sociohistorical circumstances.

Weber's main attention was focused on the "culture process," which he did not regard as transferable. Culture can be understood only by recognizing the historical uniqueness of each case, since culture derives from the creative spontaneity of man, which in turn is the expression of an "immanent transcendence" that is not susceptible to the generalizing methods of science. There can therefore be no causal laws in the domain of culture. To assert their existence seemed to Weber no less mistaken than Herbert Spencer's "wrong-headed social evolutionism" (*Farewell to European History,* p. 49). Like Johann Gottfried Herder, for whom he had a profound admiration, Weber deplored what he called the Enlightenment's "dogmatic progressivism" as a "dangerous sort of optimism" (loc. cit.). The progressivist, evolutionary thesis stemmed, in Weber's opinion, from confusing the culture process with the civilization process, thus misconceiving the nature of culture, for culture does not follow any definite or lineal order of development but occurs sporadically, defying the causal determinism that operates in the realms of science and technology.

Weber's theory of immanent transcendentalism also colored his political views. In place of state socialism (whether of the Bismarckian or the Marxist-Leninist kind), he advocated a "debureaucratized" form of "free socialism," under which man's functional role within the social system would never be that of a mere functionary whose inner sense of right and wrong could be made subservient to reasons of state.

Weber's insistence on viewing the historical world of man as a realm where transcendental but (in contrast to G. W. F. Hegel) immanent determinants are at least as decisive as empirical or material factors reveals not only his fundamental disagreement with the Marxist school of historical determinism but also his most significant point of departure from the sociological methodology of his older brother. Unlike Max Weber, Alfred Weber could not conceive of a meaningful sociological interpretation or explanation of human thought or action that aimed to dispense with a value-oriented perspective.

Alfred Weber may possibly have exaggerated the difference between his methodological approach and that of his brother; it may well be true to say with Arnold Brecht that it is a difference of degree rather than of kind, that Alfred Weber was a latent and partisan relativist and Max Weber an overt and neutral one (*Political Theory,* Princeton, NJ, 1959, p. 278). Be that as it may, Alfred Weber's stress on a specifically historicocultural approach to sociology, no less than his denial of the validity of the naturalistic method in the sphere of human affairs, contributed to the relative lack of understanding of his theories by many contemporary sociologists.

Whatever the ultimate assessment of Alfred Weber as a sociologist, his penetrating insight into the forces that shape human history and his uncompromising adherence to the principle of individual social responsibility place him high in the tradition of thinkers of integrity in scholarship and in action.

See also Enlightenment; Functionalism in Sociology; Hegel, Georg Wilhelm Friedrich; Herder, Johann Gottfried; Masaryk, Tomáš Garrigue; Philosophy of History; Weber, Max.

Bibliography

WORKS BY WEBER

Über den Standort der Industrien. Tübingen, 1909. Translated by C. J. Friedrich as *Theory of the Location of Industries.* Chicago: University of Chicago Press, 1929.

Gedanken zur deutschen Sendung. Berlin: Fischer, 1915.

"Prinzipielles zur Kultursoziologie." *Archiv für Sozialwissenschaft und Sozialpolitik* 47 (1920–1921): 1–49. Translated by G. H. Weltner and C. F. Hirschman as *Fundamentals of Culture-Sociology.* New York, 1939.

Die Not der geistigen Arbeiter. Munich, 1923.

Deutschland und die europäische Kulturkrise. Berlin: Fischer, 1924.

Die Krise des modernen Staatsgedankens in Europa. Stuttgart: Deutsche Verlags-Anstalt, 1925.

Ideen zur Staats- und Kultursoziologie. Karlsruhe: Braun, 1927.

Kulturgeschichte als Kultursoziologie. Leiden: Sijthoff, 1935.

Das Tragische und die Geschichte. Hamburg: Govert, 1943.

Abschied von der bisherigen Geschichte. Hamburg: Claassen and Goverts, 1946. Translated by R. F. C. Hull as *Farewell to*

European History. New Haven, CT: Yale University Press, 1948.

Der dritte oder der vierte Mensch. Munich: Piper, 1953.

Einführung in die Soziologie. Munich: Piper, 1955.

WORKS ON WEBER

Barnes, H. E., and Howard Becker. *Social Thought from Lore to Science,* 2 vols. New York: Heath, 1938. Vol. II, Ch. 20.

Colvin, Milton. "Alfred Weber—The Sociologist as a Humanist." *American Journal of Sociology* 65 (1959): 166–168.

Neumann, Sigmund. "Alfred Weber's Conception of Historico-Cultural Sociology." In *An Introduction to the History of Sociology,* edited by H. E. Barnes, 353–361. Chicago: University of Chicago Press, 1948.

Salomon, Albert. "The Place of Alfred Weber's *Kultursoziologie* in Social Thought." *Social Research* 3 (1936): 494–500.

Frederick M. Barnard (1967)

WEBER, MAX

(1864–1920)

Max Weber, the German sociologist, historian, and philosopher, was raised in Berlin. His father was a lawyer and National Liberal parliamentary deputy, his mother a woman of deep humanitarian and religious convictions. The Weber household was a meeting place for academics and liberal politicians. From 1882 to 1886 Weber studied law at the universities of Heidelberg, Berlin, and Göttingen, except for a year of military training. His doctoral dissertation (1889) was on medieval commercial law, and he continued his researches into legal history with a study of Roman agrarian law. In 1890 he was commissioned by the Verein für Sozialpolitik to investigate the social and economic plight of the east German agricultural worker. Between 1894 and 1897 he was professor of economics, first at Freiburg, then at Heidelberg. During the next four years, however, a severe nervous illness forced him into academic retirement and kept him from productive work. His health never recovered sufficiently for him to resume an academic career, and he spent the years preceding World War I mainly at Heidelberg as a private scholar, although he became associate editor of the *Archiv für Sozialwissenschaft und Sozialpolitik* in 1903. During the war he was director of army hospitals at Heidelberg. As a consultant to the German armistice commission at Versailles he helped to draw up the memorandum on German war guilt; he also advised the commission that prepared the first draft of the Weimar constitution. Late in the war, Weber had accepted a temporary teaching post at the University of Vienna, and in 1919 he became professor of economics at Munich. He died shortly thereafter.

SOCIOLOGY, POLITICS, ETHICS, AND ECONOMICS

Weber was attracted to practical politics as well as to scholarship, and he had a vivid sense of the political and cultural significance of historical and sociological investigations. Nevertheless, he insisted that these two "callings" must be kept apart, for both political and academic reasons. His east German agrarian studies had convinced him that the decline of the *Junkers* as a positive political force made it necessary to foster a professional class of politicians who could direct the German administrative machine. He condemned Otto von Bismarck for having failed to cultivate such a class and for thus paving the way for the political dilettantism to which Weber attributed most of the weaknesses of German diplomacy. He also argued that scientific and philosophical inquiries into social phenomena were not capable of settling disputes about ethical and cultural values, commitment to which was a sine qua non of worthwhile political activity. Empirical scientific investigation could lead to the discovery of the ultimate motives of human behavior, which would serve as a preliminary to an adequate causal explanation of historical events; it could demonstrate the means necessary to given ends; and it could show otherwise unsuspected by-products of alternative policies. Philosophical analysis could lay bare the conceptual structure of various evaluative systems, place them with respect to other possible ultimate values, and delimit their respective spheres of validity. But such studies could not show that any particular answers to evaluative questions were correct. Weber pointed out that an evaluative choice does not depend merely on technical considerations applied to given ends; it is inherent in the very nature of the criteria used to discuss such questions that dispute about those criteria is both possible and necessary. There would be something incoherent in the idea that such disputes could ever be definitively settled.

Weber argued that the blurring by academic writers of the distinction between fact and value characteristically led to two unwarranted prejudices. First, because of the academic's duty to examine all sides of any question, he was likely to develop a predilection for the middle course, although a compromise "*is not by a hairbreadth more scientifically true* than the most extreme ideals of the parties of the left or right." Second, because the scientific investigator's methods were peculiarly well adapted to discovering the probable results of policies, he was likely

to think that a policy's value must also be settled by reference to results. But, Weber argued, policies could be rational, not merely in the sense of adapting means to ends (*zweckrational*), but also in the sense that they consistently and genuinely express the attachment to certain values of an agent who is indifferent to the achievement or nonachievement of further ends (*wertrational*).

Weber denied that any form of social activity could be purely economic. All activities have an economic aspect insofar as they face scarcity of resources and thus involve planning, cooperation, and competition. But economic considerations alone cannot explain the particular direction taken by any social activity or movement; for this, other values have to be taken into consideration. Further, the sociologist's own culturally conditioned values are already involved in the way in which he has isolated an intelligible field of study from the infinite complexity of social life. Hence, there is a certain subjectivity of value at the very foundations of social scientific inquiry, but this need not damage the objectivity of the results of such inquiry.

VERSTEHEN AND CAUSAL EXPLANATION

Social phenomena involve the actions of agents who themselves attach a sense (*Sinn*) to what they are doing. Correspondingly, sociology requires an understanding (*Verstehen*) of the sense of what is being studied. Without it, Weber argued, the sociologist would not even be in a position to describe the events he wants to explain. In this respect Weber was squarely in the tradition of G. W. F. Hegel, Wilhelm Dilthey, and Heinrich Rickert, but he developed these philosophical ideas into a methodology and applied it to a vast spectrum of empirical data.

Verstehen is particularly susceptible to the investigator's subjective bias, and the sense of unfamiliar forms of activity is likely to be interpreted by reference to what is familiar, but perhaps only superficially similar. Weber therefore thought that *Verstehen* must be supplemented by what he sometimes seemed to regard as a distinct method of inquiry, causal explanation. He argued that causal explanations in sociology are, as such, completely naturalistic and that the social sciences are distinguished by the addition of *Verstehen.* He did not always see clearly that a method which is to serve as a check on rashly subjective misinterpretations of the sense of an activity must itself be capable of producing more correct interpretations. Nor did he always understand that what he called causal explanation, therefore, must itself already involve the concept of *Verstehen.*

This point can be illustrated by Weber's treatment of authority (*Herrschaft*). As a prelude to a causal treatment, he tried to define authority naturalistically in terms of statistical laws expressing "the probability that a command with a given specific content will be obeyed by a given group of persons." The presence of expressions such as "command" and "obeyed" in this definition shows that it already presupposes *Verstehen.* This continues to hold for Weber's further treatment of the various types of legitimation in terms of which he classified authority: the traditional, the rational (bureaucratic), and the charismatic (involving attachment to the person of a powerful individual leader—Weber regarded charismatic authority as a principal source of social change). Here, as elsewhere in his work, the appeal to statistical laws must be understood as ancillary to the process of arriving at an adequate *Verstehen* and not as belonging to a distinct method of causal inquiry.

THE "IDEAL TYPE"

Both *Verstehen* and causal explanation are again involved in Weber's account of the use of "ideal types" in historical and sociological inquiries. Whereas a purely classificatory concept is reached by abstraction from a wide range of phenomena with differing individual characteristics, an ideal type is intended to illuminate what is peculiar to a given cultural phenomenon. Its most characteristic use is in connection with types of rational behavior. The ideal type is a model of what an agent would do if he were to act completely rationally according to the criteria of rationality involved in his behavior's sense. On the one hand, the ideal type facilitates *Verstehen* in that, although not itself a description of reality, it provides a vocabulary and grammar for clear descriptions of reality. On the other hand, although the ideal type is not itself a causal hypothesis, it is an aid to the construction of such hypotheses for the explanation of behavior that deviates from the ideal-typical norm. Weber regarded the three forms of authority (traditional, rational, and charismatic) as well as the theory of the market in economics as ideal types. The most succinct and celebrated application of the concept, as well as of most of his other methodological ideas, is to be found in *The Protestant Ethic and the Spirit of Capitalism.* In this work Weber argued that the development of European capitalism could not be accounted for in purely economic or technological terms but was in large part the result of the ascetic secular morality associated with the twin emphases in Calvinistic theology on predestination and salvation.

See also Authority; Determinism, A Historical Survey; Dilthey, Wilhelm; Hegel, Georg Wilhelm Friedrich; Philosophy of Social Sciences; Rickert, Heinrich.

Bibliography

WORKS BY WEBER

Gesammelte Aufsätze zur Religionssoziologie. 3 vols. Tübingen: Mohr, 1920–1921. Vol. I, Part 1, translated by Talcott Parsons as *The Protestant Ethic and the Spirit of Capitalism.* London: Allen and Unwin, 1930.

Gesammelte politische Schriften. Munich: Drei Masken, 1921.

Gesammelte Aufsätze zur Wissenschaftslehre. Tübingen: Mohr, 1922.

Wirtschaft und Gesellschaft. Tübingen: Mohr, 1922. 2nd ed., 2 vols. Tübingen, 1925. Part 1 translated by A. M. Henderson and Talcott Parsons as *The Theory of Social and Economic Organization.* Glencoe, IL, 1947. Vol. II, pp. 735ff., translated by Don Martindale and Gertrude Neuwirth as *The City.* New York: Free Press, 1958. Selections (chiefly from Ch. 7) in *Max Weber on Law in Economy and Society,* translated by E. A. Shils and Max Rheinstein. Cambridge, MA: Harvard University Press, 1954.

Wirtschaftsgeschichte. Munich: Duncker and Humblot, 1923. Translated by F. H. Knight as *General Economic History.* London: Allen and Unwin, 1927.

Gesammelte Aufsätze zur Soziologie und Sozialpolitik. Tübingen: Mohr, 1924.

From Max Weber: Essays in Sociology. Translated by H. H. Gerth and C. Wright Mills. New York: Oxford University Press, 1946.

Schriften zur theoretischen Soziologie, zur Soziologie der Politik und Verfassung. Frankfurt: Schauer, 1947.

On the Methodology of the Social Sciences. Translated and edited by E. A. Shils and H. A. Finch. Glencoe, IL: Free Press, 1949.

WORKS ON WEBER

Antoni, Carlo. *Dallo storicismo alla sociologia.* Florence, 1940. Translated by H. V. White as *From History to Sociology.* Detroit, MI: Wayne State University Press, 1959.

Aron, Raymond. *Essai sur la théorie de l'histoire dans l'Allemagne contemporaine.* Paris: Vrin, 1938.

Aron, Raymond. *La sociologie allemande contemporaine.* Paris: Alcan, 1935. Translated by M. Bottomore and T. Bottomore as *Modern German Sociology.* London, 1957.

Beetham, David. *Max Weber and the Theory of Modern Politics,* 2nd ed. Cambridge, U.K.: Polity Press, 1985.

Behnegar, Nasser. *Leo Strauss, Max Weber, and the Scientific Study of Politics.* Chicago: University of Chicago Press, 2003.

Bendix, Reinhard. *Max Weber. An Intellectual Portrait.* London: Heinemann, 1960.

Diggins, John P. *Max Weber: Politics and the Spirit of Tragedy.* New York: Basic, 1996.

Henrich, Dieter. *Die Einheit der Wissenschaftslehre Max Webers.* Tübingen, 1952.

Käsler, Dirk. *Max Weber: An Introduction to His Life and Work.* Translated by Philippa Hurd. Chicago: University of Chicago Press, 1988.

Lennert, Rudolf. *Die Religionstheorie Max Webers.* Stuttgart, 1935.

Mayer, J. P. *Max Weber and German Politics.* London: Faber and Faber, 1944.

Parsons, Talcott. *The Structure of Social Action.* New York: McGraw-Hill, 1937.

Scaff, Lawrence. *Fleeing the Iron Cage: Culture, Politics, and Modernity in the Thought of Max Weber.* Berkeley: University of California Press, 1989.

Swedberg, Richard. *Max Weber and the Idea of Economic Sociology.* Princeton, NJ: Princeton University Press, 2000.

Weber, Marianne. *Max Weber. Ein Lebensbild.* Tübingen: Mohr, 1926.

Peter Winch (1967)
Bibliography updated by Philip Reed (2005)

WEIL, SIMONE

(1909–1943)

The French author and mystic Simone Weil was born in Paris into a well-to-do family of distinguished intellectuals. During her lifetime she published only articles, dealing mainly with political and social issues, in obscure syndicalist sheets. Her uncompromising dedication to the search for truth and social justice as a way of life made her a significant though much debated personality. She lived a life of stringent deprivation. In spite of ill health she worked in factories, joined the anti-Franco volunteers in Spain, and worked as a farm laborer in the south of France after the 1940 defeat. After 1942 she lived in exile in New York and then in England. Jewish by birth, she wished to partake fully in the suffering of the victims of Nazism, and she allowed herself to die of hunger.

While in her twenties she was trained by Alain (Émile Auguste Chartier) in philosophy and logic. She had a voracious, relentless mind, and her studies included Greek, Latin, Sanskrit, several modern languages, philosophy, Western and Oriental religions, science, mathematics, and literature. Her writings are primarily based on textual comment and syncretic, ahistoric, and controversial interpretations. Her thought is rooted in Platonic and Stoic philosophy reinterpreted in terms of an apparently genuine mystical experience—in 1938 Weil experienced a moment of supernatural revelation and union with Christ. It gave her a mystical sense of vocation as possessor of a truth that she was delegated to transmit.

The bulk of her work, touching on the social, moral, aesthetic, and religious facets of life, was posthumously published. The published works combine fragments, more or less consistently developed and sometimes rather speciously selected, from her notebooks, letters, articles,

and memoranda. The three-volume *Cahiers* (two volumes in the English translation) gives the integral but still fragmentary manuscript text from which the first published volumes were drawn.

A systematic interpretation of her work is problematical and, besides, could do her sometimes brilliant, sometimes obscure, paradoxical writing scant justice. Her thought is concentrated in two areas, the social and metaphysical, linked by her special concept of the human person. In a universe ruled by an iron, impersonal necessity, the human being shows an ineradicable expectation of goodness that is the sacred part of the human person. Society, the collective in whatever form, is the "large animal" offering the individual a false transcendency. Modern industrial society uproots but offers no values corresponding to the sacred aspirations of the individual. Not until labor and thought coincide and work is reintegrated into the spiritual edifice of society will the individual regain a sense of freedom, dignity, and community.

Central to Weil's thought is the fundamental human frustration caused by the inherent contradiction between two forces—the rigorous mechanical necessity at work in the universe and the inner expectation of good. Weil developed her metaphysics from this central conflict. She presents a dialectic of divine creation and voluntary personal "decreation" or disindividualization whereby the creature relinquishes the particular and becomes annihilated in divine love through methodical destruction of the self. The destruction of the self is to be attained first by rigorous use of discursive reason pushed to its ultimate limits, at which point there will remain only a wall of unpassable contradictions representing the absurdities of the human condition. The second step is the way of the mystics and involves nondiscursive disciplines—attention, waiting, "transparency," an inner void, and silence followed by certainty. Both methods of approach are apparent in her writing. Her God is impersonal and passive because all-loving. Only through a voluntary withdrawal of God could the act of creation take place. Evil, felt by man as suffering and apprehended by the understanding as the incomprehensible, is the paradoxical lot of the creature because of the nature of the initial act of finite creation by the infinite being.

See also Mysticism, History of; Mysticism, Nature and Assessment of; Platonism and the Platonic Tradition; Stoicism; Women in the History of Philosophy.

Bibliography

WORKS BY WEIL

For bibliographies see *Selected Essays, 1934–1943*, translated by Richard Rees (London: Oxford University Press, 1962), pp. 228–229, and Michel Throut, *Essai de bibliographie des écrits de Simone Weil*, Vol. 26, *Archives des lettres modernes* (October 1959, Paris).

The notebooks have been published in French as *Cahiers*, 3 vols. (Paris: Plon, 1951, 1953, 1956), and in English translation by Arthur F. Wills as *The Notebooks of Simone Weil*, 2 vols. (New York: Putnam, 1956).

The published selections of her writings include *La pesanteur et la gràce* (Paris: Plon, 1946), translated by Emma Craufurd as *Gravity and Grace* (London: Routledge, 1952); *L'enracinement* (Paris: Gallimard, 1949), translated by Arthur F. Wills as *The Need for Roots* (New York: Putnam, 1952); *Attente de Dieu* (Paris: La Colombe, 1950), translated by Emma Craufurd as *Waiting for God* (New York: Putnam, 1951); *La connaissance surnaturelle* (Paris: Gallimard, 1950); *La condition ouvrière* (Paris: Gallimard, 1951); *Lettre à un religieux* (Paris: Gallimard, 1951), translated by Arthur F. Wills as *Letter to a Priest* (London: Routledge, 1953); *Intuitions pré-chrétiennes* (Paris: La Colombe, 1951), parts of which have been selected, edited, and translated by Elisabeth Geissbuhler as *Intimations of Christianity among the Ancient Greeks* (London: Routledge, 1957); *La source grecque* (Paris: Gallimard, 1953); and *Écrits de Londres* (Paris: Gallimard, 1957).

Additional Works

Oppression et liberté. Paris: Gallimard, 1955.

Écrits de Londres et dernières lettres. Paris: Gallimard, 1957.

Leçons de philosophies (Roanne 1933–1934). Edited by Anne Reynaud. Paris: Plon, 1959.

Essai biographique et critique. Paris: Presses Universitaires de France, 1960.

Écrits historiques et politiques. Paris: Gallimard, 1960.

Pensées sans ordre concernant l'amour de Dieu. Paris: Gallimard, 1962.

Sur la science. Paris: Gallimard, 1966.

Réflexions sur les causes de la liberté et de l'oppression sociale. Paris: Gallimard, 1980.

Œuvres completes. 6 vols. Paris: Gallimard, 1989–1994.

Additional Translations

First and Last Notebooks. Translated by Richard Rees. London: Oxford University Press, 1970.

Formative Writings: 1929–1941. Edited and translated by Dorothy Tuck McFarland and Wilhelmina Van Ness. Amherst: University of Massachusetts Press, 1987.

Gateway to God. Edited by David Raper. Glasgow: Collins, Fontana, 1974.

The Iliad, or, The Poem of Force. Translated by Mary McCarthy. Wallingford, PA: Pendle Hill, 1956.

Lectures on Philosophy. Translated by H. Price. Cambridge: Cambridge University Press, 1978.

On Science, Necessity, and the Love of God. Collected, translated, and edited by Richard Rees. London: Oxford University Press, 1968.

Oppression and Liberty. Translated by Arthur Wills and John Petrie. Amherst: University of Massachusetts Press, 1973.

Seventy Letters. Translated by Richard Rees. London: Oxford University Press, 1965.
Simone Weil, an Anthology. Wallingford, PA: Pendle Hill, 1956.
The Simone Weil Reader. Edited by George A. Panichas. New York: David McKay, 1977.

WORKS ON WEIL

Allen, Diogenes, and Eric O. Springsted. *Spirit, Nature and Community: Issues in the Thought of Simone Weil*. Albany: State University New York Press, 1994.
Anderson, David. *Simone Weil*. London: SCM Press, 1971.
Bell, Richard H., ed. *Simone Weil's Philosophy of Culture*. Cambridge, U.K.: Cambridge University Press, 1993.
Blum, Lawrence, and Victor Seidler. *A Truer Liberty: Simone Weil and Marxism*. London: Routledge and Kegan Paul, 1989.
Cabaud, Jacques. *Simone Weil: A Fellowship in Love*. New York: Channel Press, 1964.
Coles, Robert. *Simone Weil: A Modern Pilgrimage*. Reading, MA: Addison-Wesley, 1987.
Davy, Marie-Magdeleine. *The Mysticism of Simone Weil*. Translated by Cynthia Rowland. Boston: Beacon Press, 1951.
Dietz, Mary G. *Between the Human and the Divine: The Political Thought of Simone Weil*. Otowa, NJ: Rowman & Littlefield, 1988.
Dunaway, John M., and Eric O. Springsted. *The Beauty That Saves: Essays on Aesthetics and Language in Simone Weil*. Macon, GA: Mercer University Press, 1966.
Fiori, Gabriella. *Simone Weil: An Intellectual Biography*. Translated by Joseph R. Berrigan. Athens: University of Georgia Press, 1989.
Gray, Francine du Plessix. *Simone Weil*. New York: Viking, 2001.
Kempfer, G. *La philosophie mystique de Simone Weil*. Paris: La Colombe, 1960.
Little, J. P. *Simone Weil: Waiting on Truth*. Oxford: St. Martin's Press, 1988.
Malan, I. R. *L'enracinement de Simone Weil*. Paris: Didier, 1961.
McClellan, David. *Utopian Pessimist—The Life and Thought of Simone Weil*. New York: Poseidon Press, 1990.
McFarland, Dorothy Tuck. *Simone Weil*. New York: Unger, 1983.
Moulakis, Anthanasios. *Simone Weil and the Politics of Self-Denial*. Translated by Ruth Hein. Columbia: University of Missouri Press, 1998.
Nevin, Thomas. *Simone Weil, Portrait of a Self-Exiled Jew*. Chapel Hill: University of North Carolina Press, 1991.
Perrin, J. M., and Gustave Thibon. *Simone Weil as We Knew Her*. Translated by Emma Craufurd. London: Routledge and Kegan Paul, 1953.
Pétrement, Simone. *Simone Weil: A Life*. New York: Pantheon, 1976.
Rees, Richard. *Simone Weil: A Sketch for a Portrait*. Carbondale: Southern Illinois University Press, 1966.
Rhees, Rush. *Discussions of Simone Weil*. Edited by D. Z. Phillips. Albany: State University of New York Press, 1998.
Springsted, Eric O. *Simone Weil and the Suffering of Love*. Cambridge, U.K.: Cowley, 1986.
Veto, Miklos. *The Religious Metaphysics of Simone Weil*. Translated by Joan Dargan. Albany: State University of New York Press, 1994.
Winch, Peter. *Simone Weil: "The Just Balance."* Cambridge, U.K.: Cambridge University Press, 1980.

Germaine Brée (1967)
Bibliography updated by Thomas Nenon (2005)

WELL-BEING

See *Eudaimonia; Happiness; Self-Interest.*

WESTERMARCK, EDWARD ALEXANDER
(1862–1939)

Edward Alexander Westermarck is best known as an anthropologist and sociologist; he is important in philosophy, however, as an exponent of a subjectivist theory of ethics, which he illustrated and supported by a survey of the actual variations in moral ideas. He himself made it clear in *Memories of My Life* that his interest in the sociology of morals arose from a concern with the philosophical question of the status of moral judgments and not vice versa.

Westermarck was born in Helsinki, Finland, of Swedish ancestry and was educated at the University of Helsinki. After 1887 he lived partly in England and partly in Finland, but he also made lengthy visits to Morocco from 1897 on. He was lecturer in sociology at the University of London from 1903 and professor of sociology there from 1907 to 1930; professor of practical philosophy at the University of Helsinki from 1906 to 1918; and professor of philosophy at the Academy of Abo from 1918. Westermarck did not marry, and his life was spent mainly in research, writing, and university teaching. On occasion, however, he joined other Finnish intellectuals in defense of their country's national interests, and he took a leading part in the founding of people's high schools for the Swedish-speaking population of Finland and of the Swedish university at Abo in Finland, of which he became the first rector in 1918.

As an undergraduate Westermarck became (and thereafter remained) an agnostic. The theme of his last book, *Christianity and Morals,* is that the moral influence of Christianity has been, on the whole, bad rather than good. He found German metaphysics distasteful but was attracted by English empiricism, especially that of J. S. Mill and Herbert Spencer. This interest, together with the aim of using the library of the British Museum, attracted Westermarck to England. Through an interest in evolu-

tion he was led to the investigation of the history of marriage, which was to be the subject of his first book. Though much of his later work was based on his own observations and personal knowledge of Morocco, all Westermarck's early anthropological research was carried out in the reading room of the British Museum. On each topic that he studied, he painstakingly collected an enormous volume of data from a wide range of sources. His aim was never merely to amass evidence, however, but to draw general conclusions from it. In *The History of Human Marriage,* for example, he rejected the widely accepted theory of primitive promiscuity or communal marriage, severely criticizing the use of supposed "survivals" as evidence for it and showing that the actual evidence pointed to the extreme antiquity of individual marriage. And throughout this work evolution by natural selection is used as a guiding principle in forming theories and explanations.

Westermarck's second and longest work, *The Origin and Development of the Moral Ideas,* written from 1891 to 1908, is partly philosophical and partly sociological. He began by propounding the subjectivist view of ethics presupposed in the whole plan of the investigation. No ethical principles are objectively valid; moral judgments are based not on the intellect but on emotions; there can be no moral truths. "Consequently the object of scientific ethics cannot be to fix rules for human conduct … its task can be none other than to investigate the moral consciousness as a fact." Thus, he discussed the nature and origin of the specifically moral emotions and the analysis of moral concepts, and he carefully examined and attempted to explain the conflicting tendencies to pass moral judgments on overt acts or exclusively on the will.

The bulk of this work treats the moral ideas comparatively and historically in order to confirm this account of the moral consciousness. Westermarck surveyed the varying attitudes and practices of many human societies on such topics as homicide, blood revenge, charity, slavery, truthfulness, altruism, asceticism, regard for the dead, and regard for supernatural beings. This detailed survey showed the continuity between moral and nonmoral retributive emotions and traced the variations in moral ideas to a number of causes.

General conclusions do not readily emerge from this mass of information, but some widely held views are conclusively proved to be false. There is no simple path of moral advance through history; many of the sentiments and rules that we associate with moral refinement are found in primitive peoples, while more barbarous views and practices have sometimes accompanied the advance of civilization. Nevertheless, Westermarck did indicate a few main trends that he expected to continue—the expansion of the altruistic sentiment, the increasing influence on moral judgments of reflection as opposed to sentimental likes and dislikes, and the restricting of religion to the function of supporting ordinary moral rules as opposed to special religious duties.

Ethical Relativity is Westermarck's most exclusively philosophical work. It repeated much from the early chapters of *The Origin and Development of the Moral Ideas,* but it argued more directly for the subjectivist view of ethics and replied to such critics of the earlier work as G. E. Moore, Hastings Rashdall, and William McDougall. Westermarck began by saying that if moral judgments state objective truths, there must be considerations by which their truth can be established, but he showed that typical ethical theories, including hedonism, utilitarianism, evolutionary ethics, rationalism, and the various accounts of a special "moral faculty," are quite unable to defend their basic principles. He recognized that the variability of moral judgments did not in itself disprove objectivism, but he argued that the persistent disagreement even on fundamental principles among the most thoughtful of moral specialists tells strongly against every form of intuitionism. He admitted that our ordinary moral judgments make a claim to objectivity, but he rightly insisted that this does not show that any judgments have objective validity. Our moral judgments result from the "objectivizing" of moral emotions, this being just one example of "a very general tendency to assign objectivity to our subjective experience." This point is of radical importance, for it undermines all attempts to support ethical objectivism by appealing to the meaning of moral terms and incidentally reveals Westermarck's firm grasp of essentials that are often obscured by the current preoccupation with the use of ethical language.

To the argument that the subjectivist theory is fatal to our spiritual convictions and aspirations, Westermarck replied that a scientific theory would not be invalidated even if it were shown to be harmful and that in any case subjectivism, by making people more tolerant and more critically reflective, is likely to do more good than harm. In reply to McDougall he defended his view that there are distinguishable moral emotions, marked off by apparent impartiality.

An important part of *Ethical Relativity* and the earlier work is the analysis of particular moral concepts to show exactly how they are related to emotions. Among other things Westermarck insisted that although the con-

cept of "moral goodness" is based on approval, those of "right," "ought," and "duty" rest not on approval but on disapproval, of what ought not to be done or ought not to be omitted.

Westermarck admitted that the variability of moral judgments is due largely to differences in knowledge and beliefs, especially religious beliefs, and that insofar as variability can be thus explained, it is not evidence against the objective validity of ethics. However, some variations—in particular, in the breadth of the altruistic sentiment—are due to emotional differences. The gradual extension of morality until it enjoins respect for all humankind and even for animals is due to the expansion of this altruistic sentiment, not to reason or religion. Not only particular moral judgments, but also the broader features of normative theories, are explained by the emotional basis of ethics. This applies not only to various hedonistic views, which are obviously linked to the source of the moral emotions in pleasure and pain, but also to the ethics of Immanuel Kant, which Westermarck criticized very thoroughly, concluding that "in his alleged dictates of reason the emotional background is transparent throughout" (p. 289).

Westermarck's ethical subjectivism belongs to a persistent, though often unpopular, tradition in philosophy. He himself particularly commended Adam Smith's *Theory of Moral Sentiments.* Westermarck's own chief contributions are his stress on "objectivization," his careful analysis of moral concepts in relation to the emotions, and his moderate and cautious use of the argument from the variability of moral judgments, backed by immense evidence of this variability. His criticism of many contrary views and his defense of his own theory against contemporary critics are also effective, though he did not develop very far the logical and epistemological considerations that tell against the objectivist view of ethics. He formulated his account with considerable care. By making it clear that moral judgments do not report the feelings of the speaker or of anyone else and that moral terms are not necessarily simply expressive of the immediate feelings of the speaker, he protected his view against the stock objections to cruder versions of subjectivism, and he left room for the part played by social demand and custom in the genesis of morality. His formulations are, perhaps, still open to more refined objections, for to give any adequate account of moral concepts is a difficult task. There are also difficulties in his theory of the moral emotions. Nevertheless, some contemporary moral philosophers believe that Westermarck's views on ethics are substantially correct and that he made an important contribution to the development and defense of views of this kind.

See also Ethical Subjectivism; Ethics and Morality; Ethics, History of; Kant, Immanuel; McDougall, William; Mill, John Stuart; Moore, George Edward; Rashdall, Hastings; Smith, Adam.

Bibliography

WORKS BY WESTERMARCK

The Origin and Development of the Moral Ideas, 2 vols. London: Macmillan, 1906–1908.

The History of Human Marriage, 5th ed., 3 vols. London: Macmillan, 1921.

The Goodness of Gods. London: Watts, 1926.

A Short History of Marriage. New York: Macmillan, 1926.

Memories of My Life. London: Allen and Unwin, 1929.

Early Beliefs and Their Social Influence. London: Macmillan, 1932.

Ethical Relativity. London: Kegan Paul, 1932.

Three Essays on Sex and Marriage. London: Macmillan, 1934.

The Future of Marriage. London: Macmillan, 1936.

Christianity and Morals. London: Kegan Paul, 1939.

WORKS ON WESTERMARCK

Discussions of Westermarck's philosophical views can be found in Paul Edwards's *The Logic of Moral Discourse* (Glencoe, IL: Free Press, 1955), pp. 46–50, 61–64; G. E. Moore's *Philosophical Studies* (London: Routledge, 1922), pp. 332–336; and L. A. Reid's review of Westermarck's *Ethical Relativity,* in *Mind* 42 (1933): 85–94.

J. L. Mackie (1967)

WEYL, (CLAUS HUGO) HERMANN

(1885–1955)

(Claus Hugo) Hermann Weyl, the German-American mathematician, physicist, and philosopher of science, was born in Elmshorn, Germany, and died in Zürich. He studied at Munich and received his Ph.D. in 1908 from Göttingen, where he was Privatdozent from 1910 to 1913. He taught at the Eidgenossische Technische Hochschule in Zürich from 1913 to 1930, lecturing at Princeton in 1928-1929. He taught at Göttingen again from 1930 to 1933 and then returned to Princeton, remaining at the Institute for Advanced Study until 1953, when he became emeritus. He became a naturalized citizen in 1939. In 1925 he received the Lobachevski Prize for his research in geometrical theory. Weyl received many honorary degrees and was a member of numerous scientific societies and a civilian member of the Office of Scientific Research and Development in 1944.

Weyl's *Raum, Zeit, Materie* (Berlin, 1918; translated by H. L. Brose from the 4th German edition as *Space-Time-Matter,* London, 1922) is a classic in relativity theory. Weyl also made significant contributions to the formalization of quantum theory (*Gruppentheorie und Quantenmechanik,* Leipzig, 1928; translated by H. P. Robertson as *Theory of Groups and Quantum Mechanics,* London, 1931). Perhaps his most important contribution of philosophical interest in this book was his attempted solution to the problem of a unified field theory in relativity. Such a theory would ultimately express in one general invariant mathematical tensor equation or law the characteristics of gravitational, electric, and magnetic fields, and show the so-called elementary particles (such as electrons or protons) as derivative from that equation. That is, the discontinuous "particles" would be generated and controlled by the continuous unified field. In 1950, in a new preface to *Space-Time-Matter,* Weyl wrote that after his own first attempt at formulating such a theory, "Quite a number of unified field theories have sprung up in the meantime. They are all based on mathematical speculation and, as far as I can see, none has had a conspicuous success." He explained that "a unitary field theory … should encompass at least three fields: electromagnetic, gravitational, and electronic. Ultimately the wave fields of other elementary particles will have to be included too, unless quantum physics succeeds in interpreting them all as different quantum states of one particle." (In quantum theory all particles have associated wave fields.) No such theory has as yet been successfully formulated, despite even Albert Einstein's final heroic and desperate attempts along this line.

Weyl also showed the validity in general relativity of a variational principle of least action. He dealt in some detail with the problem of action at a distance by examining and defining more precisely the notion of gravitational waves propagated at a finite speed (the speed of light), as is held in general relativity, in contrast to the older Newtonian theory of an infinite or indefinitely high speed for all gravitational influences. Weyl also espoused a cosmological model in which all observers located on different galaxies anywhere would have equivalent overall views of the universe.

Weyl's *Das Kontinuum* (Leipzig, 1918) consists, first, of a logical and mathematical analysis of groups and functions and deals with such questions as the axiomatic method (in the manner of David Hilbert), the natural numbers (including Richard's antinomy), and the iteration and substitution principles of formal mathematical systems. Second, Weyl analyzed the concept of number in general, in conjunction with the notion of the continuum: the logical foundations of the infinitesimal calculus, with applications to spatial and temporal continua, magnitudes and measures, curves and surfaces. In all of this he explicitly used the ideas of Georg Cantor, Bertrand Russell, A. N. Whitehead, Jules Henri Poincaré, Augustin-Louis Cauchy, Richard Dedekind, Gottlob Frege, Ernst Zermelo, and Henri Bergson. Throughout, he attempted to distinguish the abstract, idealized, schematized ("objective") mathematical continua of space and time from the intuitive, phenomenal ("subjective") space and time personally and immediately experienced by each individual. Weyl acknowledged a debt to the ideas of Bergson concerning "duration" as given in phenomenal or intuitive time.

Weyl's definitive work in the philosophy of science, *Philosophie der Mathematik und Naturwissenschaft* (Munich, 1927; translated by O. Helmer, revised and augmented, as *Philosophy of Mathematics and Natural Science,* Princeton, NJ, 1949), dealt with pure and applied mathematics. In pure mathematics, he discussed mathematical logic and axiomatics, number theory and the continuum, the infinite, and geometry. In the natural sciences, he explained basic questions concerning space, time, and the transcendental world, with special concern for the epistemological problem of subject and object. The transcendental world is, of course, the Kantian idea with Weyl's added notion that this world might be knowable by the physicist. But the question of knowing was precisely the epistemological problem that troubled Weyl, as will be seen below.

In this work Weyl also discussed methodological problems in the theory of measurement and in the formation of scientific concepts and theories. Finally, he attempted to offer a general "physical picture of the world" in the course of analyzing the ideas of matter and causality.

The first German edition of *Philosophy of Mathematics and Natural Science* was written just before the broader philosophical implications of quantum theory had been recognized; hence Weyl added several appendices to the English edition in which he coped with the newer problems. In Appendix C he declared that "whatever the future may bring, the road will not lead back to the old classical scheme." Thus, Weyl had no real hope that a classical mechanical model would ever again be established as the basis of objective reality, and he explicitly emphasized that in quantum theory the relations between subject and object "are more closely tied together than classical physics had recognized." Weyl's notion of the vagueness

of the distinction between subject and object in quantum theory has deeper metaphysical implications, of which fact he was clearly aware. How could we know the real world apart from our interactions with it and apart from the consequent indeterminacy in such "knowledge"? What, then, is the physical "object" apart from our subjective knowledge of it?

Weyl's final work was *Symmetry* (Princeton, NJ, 1952), published on the eve of his retirement from the institute. In it Weyl related the precise geometrical concept of symmetry to the vaguer artistic ideas of proportion, harmony, and beauty. In this account he was sensitive to the ideas of Plato and other great Greek classical aestheticians. His illustrated survey ranged from Sumerian art forms through the ancient Greeks and the medievals, and down to contemporary physicists, crystallographers, and biologists, briefly mentioning modern women's fashions.

See also Bergson, Henri; Cantor, Georg; Confirmation Theory; Frege, Gottlob; Hilbert, David; Mathematics, Foundations of; Philosophy of Science, History of; Philosophy of Science, Problems of; Plato; Poincaré, Jules Henri; Relativity Theory; Russell, Bertrand Arthur William; Whitehead, Alfred North.

Bibliography

Other works by Weyl of interest to philosophers of science are: *Die Idee der Riemannschen Fläche* (Leipzig, 1913); *The Classical Groups* (Princeton, NJ: Princeton University Press, 1939); *Algebraic Theory of Numbers* (Princeton, NJ: Princeton University Press, 1940); *Metamorphic Functions and Analytic Curves* (Princeton: Princeton University Press, 1943); and *The Structure and Representation of Continuous Groups* (Princeton: Princeton University Press, 1955).

For further works by Weyl and for works on him, see *Biographical Memoirs of the Fellows of the Royal Society* 3 (1957): 305–328.

Carlton W. Berenda (1967)

WHATELY, RICHARD

(1787–1863)

Richard Whately, the English logician, was a fellow of Oriel College and archbishop of Dublin. In 1860 Augustus De Morgan said of Whately that "to him is due the title of the restorer of logical study in England." Between 1826, the year Whately's *Elements of Logic* was published, and 1860, George Boole, De Morgan, and John Stuart Mill were writing. It is therefore natural to expect to find adumbrations of their work in Whately, but in his systematic and formal treatment of logic there are remarkably few. Mill did mention that Whately revived the discussion of connotative terms (called attributive by Whately). Whately's section on "the drift of propositions," which is original and perceptive, was ignored until the twentieth century. Yet this is all that was original, and it is to be found only in later editions.

This systematic section was based on Henry Aldrich's cram book, *Artis Logicae Compendium,* published in 1691 and still used at Oxford in Whately's day. The section was conservative. All propositions were considered to be subject–copula–predicate in form. All arguments were held to be reducible to syllogisms and syllogisms to be based on the *dictum de omni et nullo,* for this is the dictum of the first figure, and the other figures reduce to the first. Modal and hypothetical propositions were squeezed into subject–copula–predicate form. Disjunctives were reduced to hypotheticals and then treated as such.

Why, then, did De Morgan regard Whately as the "restorer of logical study in England"? The book was something of a best seller and the style, roughly Gilbert Ryle vintage 1826, is excellent. But this was not enough.

Whately's achievement was not so much in logic as in moral metalogic; he explained what logicians should have been doing. When he wrote, nearly 250 years after Francis Bacon, no British philosopher had made a convincing reply to the charges leveled against logic from the time of the Renaissance. The case was lost by default, and the status of logic sank so far that it ceased to be something a philosophical system must make room for, as geometry was, and became something that must accommodate itself to the convenience of the system. Therefore, logic had been continually rewritten to suit current philosophical speculation. The status of logic could not be restored until the subject matter was defined, the rewriting ended, and the charges against it answered.

Logic, said Whately, is "entirely conversant about language," and it is only as reasoning is expressed in language that logic can study it. He was not concerned with whether reasoning can be carried out some other way—by, say, "abstract ideas." This delimitation of the subject for investigation was neutral and did not necessitate subscribing to the nominalism Whately took over from Thomas Hobbes.

Once the subject was delimited, the charges against logic could be more effectively answered. Whately granted the common objection, voiced by John Locke, that man argued correctly before syllogism was heard of;

nevertheless, putting arguments in logical form provides a test of validity. This test applies in all fields. There is no logic peculiar to science or religion. Induction is not a new method of reasoning, as Bacon claimed. Induction means, first, a form of argument; but inductions of this sort are syllogistic. Induction also means generalizing from instances. This is not the province of logic, and logic cannot guarantee the truth of premises so reached. While it is true that in syllogism the conclusion contains nothing that is not in the premises, this does not render it futile, as George Campbell and others had held. "It is peculiarly creditable to Adam Smith and Malthus, that the data from which they drew such important conclusions had been in everyone's hands for centuries" (Whately, *Elements of Logic,* Book IV, Ch. 2, Sec. 4).

By example as well as by argument Whately combated the view that "logic is the Art of bewildering the learned by frivolous subtleties." He illustrated points and drew exercises from discussions in science, sociology, and religion, and thus exhibited logic in use.

Whately's *Elements of Rhetoric* (London, 1828) dealt with the effectiveness of arguments, but it also contains interesting material on such subjects as plausibility and argument from analogy. *Historic Doubts Relative to Napoleon Buonaparte* (London, 1819) is a witty and attractive reductio ad absurdum of David Hume's short way with miracles. Whately edited and annotated works of William Paley and Bacon, noting the naturalistic fallacy in Paley. He also wrote much on questions of the day relating to Ireland and on religion and economics.

See also Bacon, Francis; Boole, George; De Morgan, Augustus; Fallacies; Hobbes, Thomas; Hume, David; Induction; Logic, History of; Logic, Traditional; Malthus, Thomas Robert; Mill, John Stuart; Paley, William; Smith, Adam.

Bibliography

In addition to the works cited in the text, see E. J. Whately, *Life and Correspondence of Richard Whately, D.D.,* 2 vols. (London: Longmans, Green, 1866).

Mary Prior (1967)

WHEWELL, WILLIAM
(1794–1866)

William Whewell, the British philosopher and historian of science, was born in Lancaster. He spent the greater part of his life at Trinity College, Cambridge, as an undergraduate, fellow, and tutor, and finally as master of Trinity from 1841 until his death. He twice served as vice chancellor of Cambridge University, and he also taught mineralogy and later (1838–1855) moral philosophy.

Whewell's output was exceptional both in its abundance and in its diversity. Save for a dozen papers on the tides (1833–1850), however, his scientific works were devoted not so much to research as to teaching (*Mechanical Euclid,* Cambridge, U.K., 1837) or popularization and to apologetics (*Astronomy and General Physics,* London, 1833; *Plurality of Worlds,* London, 1853). In addition to his scientific writings he published a number of works in moral philosophy (*Elements of Morality, Including Polity,* 2 vols., London, 1845; *Lectures on Systematic Morality,* London, 1846; *Lectures on the History of Moral Philosophy in England,* London, 1852) and pedagogy (*Principles of English University Education,* London, 1837; *Of a Liberal Education,* London, 1845). He also produced editions, with prefaces, notes, and in some instances translations, of works by Isaac Newton, Joseph Butler, Hugo Grotius, Plato, and others, as well as sermons, poetry, and occasional or polemical essays.

However, his principal work—in length, scope, and the central position it occupies in his thought—is constituted by the *History of the Inductive Sciences, from the Earliest to the Present Time* (3 vols., London, 1837) and the *Philosophy of the Inductive Sciences, Founded upon Their History* (London, 1840). The former, one of the first general histories of natural science, is erudite yet perfectly readable. The latter, revised and enlarged for its third edition, was published in three parts under separate titles: *History of Scientific Ideas* (2 vols., London, 1858); *Novum Organon Renovatum* (London, 1858); and *On the Philosophy of Discovery* (London, 1860).

According to Whewell, the theory of induction, which had been examined to the point of exhaustion after Francis Bacon formulated it as a program for future science, should be taken up again in view of the fact that the sciences called inductive have been actually established. Notwithstanding the opinions of the "writers of authority" invoked by J. S. Mill, the word *induction* can now validly signify only one thing: the method of construction employed in those sciences that all modern thinkers agree to call inductive. And the only means of becoming acquainted with this method is to see it at work in history. (This is the source of the close connection between the two works, the *History* and the *Philosophy,* which matured simultaneously over a period of many years.)

INDUCTION AND HISTORY

The study of history reveals an inductive process that does not resemble the generalizing argument of the logicians. In the first place, the induction practiced by the scientist is not reasoning that is valid *vi formae* (by virtue of its form). It is quite another way of arriving at truth: a venturesome course taken by the mind, which, as if deciphering a cryptogram, tests or tries out various hypotheses in turn, until by a "happy guess" it hits upon the relevant idea. The question therefore is not under what conditions this procedure is logically correct—it never is—but simply whether its result is sound. Care and rigor assert themselves in the experimental control of the inductive proposition, and not in its elaboration, which allows great freedom to the imagination. It is fruitless to try to set up an "inductive logic" that is symmetrical with deductive logic and that formulates canons analogous to those of the syllogism.

In the second place, scientific induction consists not in generalizing the observed facts but in colligating them, in binding them together by the intelligible unity of a new conception. Finding this conception requires the initiative of genius. Generalization comes afterward; the decisive discovery is the forging of the idea. Once this idea has taught us how to read experience, it becomes incorporated into experience; and it seems to us that we see it there. Thus, the contribution of the mind to knowledge is ignored: this is the source of the empiricist error. One forgets that the facts have little by little been given form by ideas and that the facts of today (such as the fact that the earth revolves) are the hypotheses of yesterday; our facts are realized theories.

INDUCTION AND IDEAS

Whewell's epistemological analyses have a general philosophical import; indeed, they furnish an indispensable basis for the theory of knowledge. Whewell was one of the first to whom the thought occurred that such a theory could rely validly only on the history of the sciences, examining how this exemplary form of knowledge had developed. Such an examination seemed to him to justify what one might call an inductive rationalism. All knowledge requires an ideal element just as much as an empirical one. By reason of this "fundamental antithesis" Whewell's philosophy at one and the same time is, in contrast with that of the apriorists, a philosophy of induction, and in contrast with that of the empiricists, a philosophy of the idea. Even the experimental sciences rest on certain axioms whose character as necessary truths—acknowledged to the point that one cannot distinctly conceive their negation—can be explained only by the presence in our mind of certain "fundamental ideas." Number, space, time, cause, medium, polarity, affinity, symmetry, resemblance, final cause—new ideas are added to those that precede as one descends the ladder of the sciences. It was this notion that largely inspired Antoine Cournot.

But such a rationalism, stamped with the influence of Immanuel Kant, is by no means bound up with a deductive idealism. The fundamental ideas are illuminated for us only progressively, in the course of our effort to interpret experience. They become elements of the structure of reason; and the principles that they govern pass little by little, as they are better understood, from the status of happy guesses to that of necessary truths that education then makes permanent in the public mind. Through this bold conception of how self-evidence develops, the theory of fundamental ideas is joined with the theory of induction, the idea as category with the idea as hypothesis. Here there would have been a prefiguring of modern theories of the self-construction of the reason had not theological preoccupations led Whewell to locate these "fundamental ideas," from all eternity, in the divine understanding. As a result the apparent invention of these ideas by man is ultimately reduced to a simple discovery.

Although Whewell's authority was recognized, his philosophy was received only with reservation. His theory of fundamental ideas ran counter to the empiricist tradition, and freethinkers regarded the theological setting of the theory as an anachronism. The logicians, for their part, complained that Whewell's theory of induction had altered the sense of the word by wrongly assimilating inductive method to the method of hypothesis and that it had neglected the question of proof. In all these respects Mill was his typical opponent. It is worth remarking, however, that neither he nor the other critics attacked Whewell's most daring and most novel notions, the interesting nature of which seems to have escaped them: the incorporation of ideas into the facts and the development of self-evidence.

See also Bacon, Francis; Butler, Joseph; Cournot, Antoine Augustin; Epistemology; Epistemology, History of; Grotius, Hugo; Induction; Kant, Immanuel; Mill, John Stuart; Newton, Isaac; Philosophy of Science, History of Plato.

Bibliography

In addition to Whewell's works cited in the text, see I. Todhunter, *William Whewell, an Account of His Writings with Selections from His Literary and Scientific Correspondence,* 2 vols. (London: Macmillan, 1876), and Mrs. Stair Douglas, *The Life and Selections from the Correspondence of William Whewell* (London, 1879).

For further information on Whewell's philosophy, see M. R. Stoll, *Whewell's Philosophy of Induction* (Lancaster, 1929); Robert Blanché, *Le rationalisme de Whewell* (Paris: Alcan, 1935); C. J. Ducasse, "Whewell's Philosophy of Scientific Discovery," in *Philosophical Review* 60 (1951): 56–69 and 213–234; and Silvestro Marcucci, *L'"idealismo" scientifico di William Whewell* (Pisa: Istituto di Filosofia, 1963).

Robert Blanché (1967)
Translated by Albert E. Blumberg

WHICHCOTE, BENJAMIN
(1609–1683)

Benjamin Whichcote, the guiding spirit of the Cambridge Platonists, was born at Whichcote Hall, Stoke, Shropshire, of "an ancient and honourable family." He was admitted to Emmanuel College, Cambridge, in 1626 and in 1633 was elected a fellow of Emmanuel. Whichcote was renowned as a college tutor for the number and the character of his pupils, who included John Smith and John Worthington, and for the personal attention he paid to them. Ordained deacon and priest in 1636, he was in the same year appointed Sunday afternoon lecturer at Trinity Church in Cambridge, a post he held for nearly twenty years and by virtue of which he exerted considerable influence on the moral and religious life of Cambridge. At a time of violent, dogmatic theological controversy, his sermons were a fervent plea for liberality and toleration. It was his habit to speak from notes; he introduced into pulpit oratory a new, vigorous, colloquial, epigrammatic style in contrast to the traditional formal discourse. Various versions of his Sunday lectures, reconstructed from notes, were published after his death in 1683 and constitute his most substantial work.

In 1643 he temporarily left Cambridge to become rector of North Cadbury in Somerset, where he married. The following year he was invited back to Cambridge to become provost of King's College, the former provost having been ejected by the Puritan Parliament. He accepted only after great hesitation and secured special provision for the support of the former provost. Alone among the newly appointed heads of colleges, he refused to subscribe to the National Covenant, by which he would have sworn to support Calvinist forms of church government and doctrine. He secured a similar exemption for the fellows of his college. In 1650 he was elected vice-chancellor of the university.

His influence at Cambridge was now at its height and aroused considerable alarm among his more orthodox Calvinist colleagues. Especially alarmed was his former tutor at Emmanuel, Anthony Tuckney. In July 1651 Whichcote preached a commencement sermon as vice-chancellor that provoked a lively controversy between Whichcote and Tuckney in the form of letters. Tuckney accused Whichcote of laying too much stress on reason and too little on faith, of being unduly influenced by pagan ideas and by the Dutch Arminians, of being too tolerant of unorthodoxy. In reply Whichcote denied that it is possible to emphasize reason unduly, reason being "the candle of the Lord." Faith not founded on reason was mere superstition. His own ideas, he maintained, derived from meditation rather than from reading; he knew little or nothing, he said, of the Arminians (this is scarcely credible) but was not ashamed of having learned from Plato. As for tolerance, the Christian's duty is to regard with charity the views of other Christians, however mistaken he takes them to be, and to minimize rather than to exaggerate differences. Reason, tolerance, the minimizing of differences—these qualities were characteristic of Whichcote personally and were central to his moral and religious outlook.

With the restoration of Charles II, Whichcote was dismissed as provost of King's College. He complied with the Act of Uniformity and was permitted to preach, finally becoming vicar of St. Lawrence Jewry, London, where he is buried. In London as in Cambridge his sermons, especially those he delivered regularly in the City at the Guildhall, attracted congregations considerable in both quality and numbers. He died as a result of a cold contracted while visiting Ralph Cudworth at Cambridge.

Whichcote wrote nothing. He was essentially a teacher who needed the inspiration of an audience that was physically present. His views have to be extracted from his correspondence, his sermons, and the aphorisms set down in his manuscripts. His leading ethical principle was that actions are good and bad, right and wrong, in their own nature, not because they are commanded or forbidden; the goodness of an action derives from its conformity with the nature of things as apprehended by reason. In his own teaching this principle is invoked against the Calvinist doctrine that moral laws are simply expressions of God's will, but his pupils were able to turn these principles against Thomas Hobbes's doctrine that moral laws are expressions of the will of the sovereign. Which-

cote initiated the rationalistic tendency in British ethics, which runs through Cudworth, Samuel Clarke, and Richard Price to our own times. But there is nothing dry or formalistic in his rationalism; his emphasis is not on obedience to rules of conduct but on affection and spontaneity. He thought of religion and morality as liberating rather than as imposing rules.

In theology his influence encouraged the development of the characteristically "liberal" point of view, with its emphasis on goodness rather than on creeds. He thought that the Calvinists, in treating as of central importance questions of creeds, government, and ritual, made the same mistake as the high church Anglicans to whom they were so bitterly opposed. These were matters about which men should be left free to differ, choosing whatever forms and formulations help them to live better lives. This was the side of Whichcote's teaching that caught the attention of the third earl of Shaftesbury, who edited a volume of Whichcote's sermons in 1698; historically, it issues in eighteenth-century deism and nineteenth-century liberal theology, as represented, for example, in the work of Matthew Arnold, a great admirer of the Cambridge school.

See also Cambridge Platonists.

Bibliography

In 1685 there appeared in London *Select Notions of that Learned and Reverend Divine of the Church of England, Dr. Whichcote*, described as being "faithfully collected from him by a pupil and particular friend of his"; the *Select Sermons* were edited with a preface by the third earl of Shaftesbury in 1698. *Several Discourses*, edited by John Jeffrey, was published in 1701 (London); Jeffrey also edited the first edition of *Moral and Religious Aphorisms*, published in 1703 (Norwich), and a sermon, *On the True Nature of Peace in the Kingdom or Church of Christ* (1717). The most useful edition of the discourses is *The Works of the Learned Benjamin Whichcote, D.D.* (Aberdeen, 1751); for the aphorisms see *Moral and Religious Aphorisms*, edited by Samuel Salter (London, 1753), which also includes the correspondence with Tuckney. There is a modern edition of the *Aphorisms* with an introduction by Dean Inge (London, 1930). Ernest Trafford Campagnac, *The Cambridge Platonists* (Oxford: Clarendon Press, 1901), contains considerable selections from Whichcote.

OTHER RECOMMENDED WORKS

Davenport, Paul M. *Moral Divinity with a Tincture of Christ: An Interpretation of the Theology of Benjamin Whichcote, Founder of Cambridge Platonism*. The Hague: Martinus Nijhoff, 1972.

Gill, Michael B. "The Religious Rationalism of Benjamin Whichcote." *Journal of the History of Philosophy* 37 (2) (1999): 271–300.

Greene, Robert A. "Whichcote, the Candle of the Lord, and Synderesis." *Journal of the History of Ideas* 52 (4) (1991): 617–644.

Greene, Robert A. "Whichcote, Wilkins, 'Ingenuity,' and the Reasonableness of Christianity." *Journal of the History of Ideas* 42 (1981): 227–252.

Patrides, C. A., ed. *The Cambridge Platonists*. Cambridge, U.K.: Cambridge University Press, 1980.

Roberts, James D. *From Puritanism to Platonism in Seventeenth Century England*. The Hague: Martinus Nijhoff, 1968.

John Passmore (1967)
Bibliography updated by Tamra Frei (2005)

WHITEHEAD, ALFRED NORTH

(1861–1947)

Alfred North Whitehead, the philosopher and mathematician, made one of the outstanding attempts in his generation to produce a comprehensive metaphysical system that would take account of scientific cosmology.

Whitehead was born at Ramsgate on the Isle of Thanet and wrote of his boyhood in a country vicarage on the East Kent coast in the "Autobiographical Notes" (*The Philosophy of Alfred North Whitehead*, pp. 3–14) and, more vividly, in some of the essays in *Essays in Science and Philosophy* (pp. 3–52). The religious (Anglican) background of his home and the experience of companionship with strong characters in a close-knit community made impressions that left their mark on his later philosophy. With these went a Wordsworthian sense of man's continuity with nature. In his education at Sherborne, an ancient public school in Dorset, he was taught the classics and history, less in a detached spirit of scholarship than as exercises in the study of what Michael Oakeshott has called "the practical past"—a living tradition illustrating general ideas and pointing to analogies in contemporary life. This approach to history remained with him and is apparent in his philosophical books, especially *Science and the Modern World* and *Adventures of Ideas*. It is a use of history in the spirit of what Edmund Burke called "philosophic analogy."

Whitehead also learned a good deal of mathematics at Sherborne, and in 1880 he went to Trinity College, Cambridge, with a scholarship in mathematics. In 1884 he was elected to a fellowship at Trinity. Bertrand Russell was his most distinguished pupil, and from 1900 to 1911 they collaborated on the *Principia Mathematica*, which attempted to prove that mathematics could be deduced from premises of formal logic. In his obituary note on

Whitehead, Russell wrote that although one or the other would take primary responsibility for writing some parts, every part was always discussed by both of them, the whole work being a complete collaboration. W. V. Quine, in his essay "Whitehead and the Rise of Modern Logic," called *Principia Mathematica* "one of the great intellectual monuments of all time." (The fourth volume, which Whitehead was to have written on the logical foundations of geometry, never appeared.)

Whitehead resigned his lectureship from Cambridge in 1910 and moved to London. He taught at the University of London until 1914, when he became professor of applied mathematics at the Imperial College of Science and Technology. During this period Whitehead did his most intensive work in the philosophy of science.

In 1924, Whitehead accepted an invitation to a chair in philosophy at Harvard University. He was then sixty-three; the transfer gave him the opportunity to develop his philosophy of science into a full-scale metaphysical philosophy.

Whitehead's work is commonly described as falling into the three periods indicated above: the early years in Cambridge up to 1910, when he was collaborating with Russell on the logical foundations of mathematics; the middle years in London up to 1924, when he was writing on the philosophy of science; and the last years in America, when he wrote first and foremost as a metaphysician. This division can, however, be overstressed. The philosophical interests explicit in his later work can be found implicitly in the earlier work, and some of the general assumptions of Whitehead's logical and mathematical work influence the later philosophy. Rather than as a succession of interests, his thought can best be interpreted as a developing unity. This is the approach of Victor Lowe in the essay "The Development of Whitehead's Philosophy" and in his book *Understanding Whitehead*. Wolfe Mays has remarked that the progression of Whitehead's thought can be looked on as a spiral, returning to certain general notions from different standpoints, rather than as a succession of stages.

LOGICAL FOUNDATIONS OF MATHEMATICS

Whitehead and Russell had been working independently on the logic of mathematics. Russell had become acquainted with the work of Giuseppe Peano in 1900 (Gottlob Frege's work came to their attention shortly after) and was working on *Principles of Mathematics* (Cambridge, U.K., 1903). Since 1891, Whitehead had been working on *A Treatise on Universal Algebra*, for which he was made a fellow of the Royal Society in 1903. In the *Treatise* he developed some ideas of Hermann Grassmann's *Ausdehnungslehre* (theory of extension) of 1844 and 1862, attempting to give a general formal description of addition and multiplication that would hold for all algebras. The *Treatise* was little noticed at the time; it is discussed by Quine in the essay "Whitehead and the Rise of Modern Logic."

In 1906 the Royal Society published Whitehead's memoir *On Mathematical Concepts of the Material World*, in which he put forth an interpretation of concepts formalized in a logico-mathematical scheme as basic notions describing the material world. Whitehead sought to define the concepts of a geometry from which, as a formal system, the theorems of Euclidean geometry can be derived and which can be interpreted by notions of space, time, and matter. At this early stage he was already dissatisfied with the Newtonian scheme of the material world as composed of atoms each occupying a position in absolute space at an absolute time. In *On Mathematical Concepts of the Material World* the ultimate entities that compose the universe are said to be lines of force. A particle is the field of a line of force at a point; particles are thus defined as elements in a field, and a point as not just having simple location in space but as an element in a linear polyadic relation *R*, so that *R* (*a, b, c*) means the points *a, b, c* are in linear order. This makes the notion of both a point and a particle a vector and not a scalar one.

Whitehead had been impressed as an undergraduate by J. J. Thomson's lecture "The Poynting Flux of Energy in Electrodynamics," describing the transmission of energy with quantitative flow and definite direction (see *Adventures of Ideas*, p. 238); in *The Philosophy of Whitehead* (pp. 235–260) Mays comments on the significance of this notion of the flux of energy for Whitehead's later work, leading to a view of nature as routes of events or occasions inheriting from each other. Lowe says that the developments in physics that interested Whitehead when he wrote the memoir were vector physics, the theories of molecular and submolecular energetic vibration, and the rise of "field" as a basic concept. The influence of all these ideas, generalized in different terminologies, can be seen throughout his work.

PHILOSOPHY OF SCIENCE

The twofold interest in logico-deductive schemes and in empirical interpretations can also be traced throughout Whitehead's work. Indeed, he saw the connection between such schemes and the vague world of our experience as the central problem of philosophy. He sought

the connection by describing a logical scheme as a systematic and generalized formulation of relationships crudely observable in experience.

The next link in this line of thought is the development of his method of extensive abstraction. There is an exposition of this in "The Anatomy of Some Scientific Ideas" (*The Organization of Thought*, Ch. 7); it is also discussed in *An Enquiry concerning the Principles of Natural Knowledge* (Part III). The method of extensive abstraction is a topological device by which such geometrical elements as points are defined, through concepts of "whole and part" and "overlapping," as relations between volumes of a certain shape extending over others of like shape—for example, rectangles, circles, or ellipses—so that a pattern like a nest of Chinese boxes is produced:

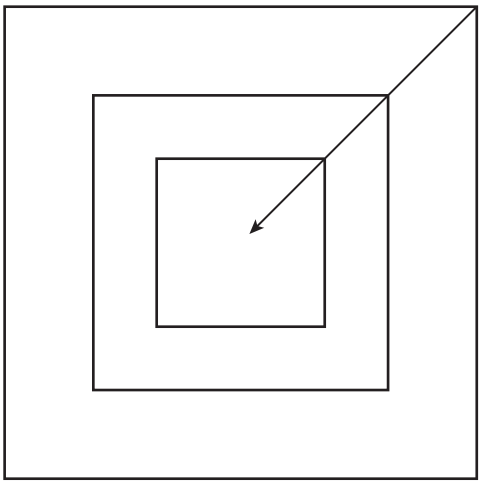

A "point" is not an ideal entity at the center or even an ideal limit of this route of approximation. It is defined as the whole convergent set. Similarly, a straight line can be defined as the direction of a route of overlapping ellipses or oblong rectangles, for example:

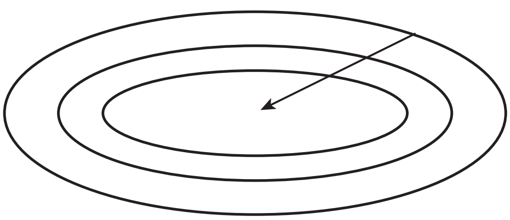

Whitehead looks on this type of definition as having an analogy in a perceived relation. No one can perceive Euclidean points with position and no magnitude or lines with length and no breadth, but volumes extending over other volumes can be perceived. The relations of "extending over" as formulated in the method of extensive abstraction are topological constructs, making precise relations that are also perceptible. This attempt to combine a view of logical schemes as reached from perceived relations with a view of them as theoretical constructs for which interpretations may be sought in experience underlies much of Whitehead's work.

OBJECTS AND EVENTS. A combination of theoretical construction and alleged derivation from experience also appears in Whitehead's analysis of nature in terms of "events" and "objects" given in the books of his middle period, *The Principles of Natural Knowledge* and *The Concept of Nature*. He claimed continually that the starting point is empirical. Just as in his earlier *On Mathematical Concepts of the Material World* he had attacked the notion of atoms externally related to one another in absolute space and time, so in his later analysis of nature (which he defined as "disclosed in sense experience"), he attacked the ultimacy of the Humean analysis of our experience into distinct impressions of sensation, such as visual sensations of colored patches. He believed that our more deep-seated experience was of something going on with spatiotemporal spread. This "passage" of nature could be divided into "events," so that its constituents are thought of not as enduring atoms but as happenings that can be described as events extending over other events. The writing of this article is a slice of the passage of nature, an event extending over the writing of this sentence, which is an event extending over the writing of this word. Thus, we converge by a route of approximation to what is happening here and now (again, an application of the basic notion of a pattern of volumes and durations extending over one another).

Events display recurrent patterns, the forms and properties of which Whitehead called "objects" and, in the later books, "eternal objects." This is his version of the problem of universals as abstract forms of recurrent recognizable characteristics in the passage of nature. The phrase "eternal objects," along with the interest in Plato shown in his later work, particularly in *Process and Reality*, might suggest that Whitehead took a Platonic realist view of a realm of such abstract entities. This is not so; his view was nearer to the Aristotelian one of *universalia in rebus* or, in his own phrase, "seeking the forms in the facts." His "objects" are "ingredients" in the process of events; they are "pure potentials" actualizable in an indefinite number of instances. At the same time he was no nominalist; the objects are more than names for observed resemblances. They are properties and relations that are exemplified in recurrences in patterns that can be precisely formulated.

Different types of objects can be distinguished. First, there is a "sense object"; for example, a color like Cambridge blue is perceived as situated in an event. A sense object requires a relation between a "percipient event," the "situation" to which it is referred, and active and passive conditioning events relating the percipient event to the situation. Second, there is the "perceptual object," a determinate association of sense objects in a series of situations strung together in a continuity and perceived as one prolonged event—for instance, that red and black coat. Perceptual objects can be delusive, as in reflections in mirrors or diffractions in water. Third, "physical objects" are those objects whose relations to events condition the appearance of the perceptual objects, as, for instance, the straight stick that appears bent in water. Fourth, "scientific objects" are inferred, nonperceived objects, such as "electrons," that account for the general properties and relations within events that constitute the situations in which physical objects are ingredients. At the stage of science in which Whitehead was writing he instanced electrons as the ultimate scientific objects. He would no doubt have welcomed the further refinements that have occurred since in discoveries of fundamental particles.

Whitehead would also have seen these developments as supporting his distinction between "uniform" and "nonuniform" objects. A uniform object is located in an event throughout a duration and also characterizes any slice of that duration. Perceptual objects are normally uniform; a bar of iron as perceived in any duration however small is still a bar of iron. A nonuniform object needs a minimum time span in order to be expressed at all; he thought a molecule, for instance, cannot exist in a lesser time than that required by the periodicity of its atomic constituents. Whitehead was impressed by the possibility suggested by the physics of his time that the ultimate scientific objects might be nonuniform rather than uniform. The development of quantum theory reinforced this idea. The notion of atomic events, or "occasions," displaying nonuniform objects and forming continuities through their overlapping so that they produce physical and perceptual objects, becomes a crucial one in Whitehead's later work. The distinctions and relations between different levels of objects are discussed in *The Principles of Natural Knowledge* (Ch. 7) and, more briefly, in the papers "Time, Space and Material" and "Uniformity and Contingency."

RELATIVITY PRINCIPLE. Objects situated in events form patterns among themselves that are constituents in wider patterns, finally dependent on a uniform pervasive pattern that expresses the uniformity of nature as an ongoing passage of related events with spatiotemporal spread. The attempt to unify notions of space, time, and matter, along with his attempt, stemming from *On Mathematical Concepts of the Material World*, to relate these to a set of formal notions underlying a geometry, led Whitehead to have a particular interest in Albert Einstein's general and special theories of relativity. Whitehead published his own alternative in *The Principle of Relativity* (1922). He refused to give a crucial role to special facts, notably the velocity of light, and, unlike Einstein, insisted that space must be "homaloidal" (that is, of uniform spread). His reason for this seems to follow from his view of abstraction, which led him to think that a logico-mathematical scheme of notions must be precisely realized in the physical world. Whitehead also believed that the possibility of measurement depended on exact congruence between one region of space and another, independently of physical bodies. Thus, though there are analogies in their conception of relativity, Whitehead's view depends on there being a noncontingent uniformity in spatial relations and is less open to experimental applications.

Whitehead's theory is set forth in his book *The Principle of Relativity* and in his article "Einstein's Theory: An Alternative Suggestion," contributed to *The Times* in 1920 and reprinted in *The Interpretation of Science*. Whitehead's views on relativity have not, however, been taken up by physicists.

METAPHYSICS

Science and the Modern World (given as Lowell Lectures at Harvard in 1925) is perhaps the most inspired expression of Whitehead's metaphysical philosophy. It is a book in which lucid and illuminating reflections on the history of science in relation to philosophy are interspersed with technically difficult passages; the book might have been written, as one reviewer remarked, by Dr. Jekyll and Mr. Hyde. But the technical passages are less overlaid with idiosyncratic terminology and a labored attempt at producing a system than is *Process and Reality* (1929). Those who find *Process and Reality* excessively forbidding can gain a very fair impression of the best of the later Whitehead by going from *Science and the Modern World* to his last books, *Adventures of Ideas* and *Modes of Thought*.

In *Modes of Thought* the analysis of nature into events and objects becomes an analysis of nature into "actual occasions," understood as unities that synthesize their relations to other occasions in their own "processes of becoming." Such a unity is called a "concrescence of prehensions," from *concrescere*, "to grow together," the

end product being something *concretum*, "concrete," and from *prehendere* "to grasp," suggesting an active relationship but not necessarily a conscious awareness (as is suggested by the word *apprehension*).

Instead of events extending over other events Whitehead now spoke of "societies" (also called *nexūs*, the plural of the Latin *nexus*) of actual occasions, which can be structured by subsocieties and which can inherit characteristics from one another in serial order, in which case they are called "enduring objects." "The real actual things that endure [such as stones or animal organisms] are all societies. They are not actual occasions" (*Adventures of Ideas*, p. 262).

This general view of larger units in nature as systems of smaller units with their own inner structure is called "the philosophy of organism." The notion of organism had already been defined in *The Principles of Natural Knowledge* (p. 3) as "the concept of unities functioning and with spatio-temporal extensions," a notion that it is said cannot be expressed in terms of a material distribution at an instant. (The definition of nonuniform objects as needing a time span for their expression may be recalled.) It is suggested that the notion of organism, thus interpreted, could be a unifying one between the physical and biological sciences, physics becoming the study of the smaller and biology of the larger organisms.

PERCEPTION. In the earlier books Whitehead had attacked the "bifurcation of nature" as the kind of view of appearance and reality that assigns secondary qualities such as colors to subjective experience and primary qualities to the physical sphere. Instead of this division he wrote about perception as nature ordered in a perspective from the standpoint of an event within nature itself called the percipient event, all perceived qualities being qualities of nature in that perspective.

In *Science and the Modern World* and in *Symbolism* (1927) the view of perception is developed in terms of what it is to be a percipient event. We start from the notion of an actual occasion as a "prehending" entity in active interaction with its whole environment. The primitive mode of perception is not, Whitehead insisted, an apprehension of clear-cut sense data or Humean "impressions of sensation." Rather, it is a vaguer sense of environing realities pressing in upon us. Whitehead called this "perception in the mode of causal efficacy" and thought that it is mediated primarily through kinesthetic organic sensation. "Philosophers," he said, "have disdained the information about the universe obtained through their visceral feelings, and have concentrated on visual feelings" (*Process and Reality*, p. 169 [184]; references to *Process and Reality* give the page of the Cambridge edition, 1929, followed by the page of the New York edition, 1929). This is a causal, not a phenomenalist, view of perception, in which the functioning of the physiological organism (disregarded by David Hume) is crucial. Environing events are mediated through the organism, becoming finally transmuted into conscious sensations, which are then projected as sensa qualifying regions of the contemporary world (this is called "symbolic reference" and "perception in the mode of presentational immediacy"). Since there is a time lag between the transmission of influences from the environment and the projection of sensa onto the contemporary world (events that are strictly contemporaneous must in Whitehead's view be causally independent), there is always a chance that perception in the mode of presentational immediacy will not give veridical information about the state of the environment, as when we perceive a yellow patch in the sky that we take to be a star, though the star has long since gone out of existence.

In "the mode of causal efficacy" the qualities of environing events are mediated through organic experiences of the percipient's body. The most difficult aspect in Whitehead's theory is the transmutation of an emotional organic experience into a sensum. He found a link in our use of color words such as *red* and *green* to describe certain affective states.

> This notion of the sensa as qualifications of affective tone is a paradox for philosophy, though it is fairly obvious to common sense. A red-irritation is prevalent among nerve-racked people and among bulls. The affective tone of perception in a green woodland in spring can only be defined by the delicate shades of the green. (*Adventures of Ideas*, p. 315)

But can an irritation be "red" except by metaphor (waiving the question of whether bulls do have color vision), and does Andrew Marvell's "green thought in a green shade" mean that "green" characterizes the thought or, rather, that there is an overwhelming awareness of green in the environment?

PROCESS AND REALITY. Whitehead's comprehensive metaphysical philosophy was presented in "An Essay on Cosmology," in *Process and Reality*, based on the Gifford Lectures given at the University of Edinburgh during the 1927–1928 session. Whitehead distinguished cosmology from metaphysics (which he held dealt with the formal character of all facts), maintaining that cosmology

described the general characteristics of our "cosmic epoch." That is, it took account of the empirical character of a particular type of world order—in the case of our world order, one characterized by electromagnetic events, dimensions, shapes, and measurability. Laws of nature, Whitehead held, were not part of the ultimate metaphysics of the universe; they could change their character with the rise and fall of different cosmic epochs dominated by different kinds of facts.

Process and Reality is a very difficult book, partly because of its vocabulary and not least when words of ordinary speech, such as *feelings*, are used with special meaning. Its manner of presentation is also difficult; the reader is confronted in the second chapter with the "categoreal [*sic*] scheme," comprising a category of the ultimate, 8 categories of existence, and 27 categories of explanation. He may find it advisable to read on and turn back to the scheme in the hope that what is there set out in summary form may become clearer in the light of the further discussions.

Lowe, in *Understanding Whitehead*, gives what is probably the most balanced presentation of Whitehead's work as a whole. Some of its notions are interpreted by analogy with more traditional metaphysical ones in Ivor Leclerc's *Whitehead's Metaphysics*, where comparison starts from the Aristotelian discussion of what it is to be a complete fact. Some aspects of the notions of "actual entities," "eternal objects," and their relations are considered in detail by William A. Christian in *An Interpretation of Whitehead's Metaphysics*; he has a particular interest in Whitehead's doctrine of God and its resemblance to and difference from more traditional views. The main drawback of these otherwise able books is that they seek to elucidate Whitehead's system in its own terms. It is likely that the contribution of *Process and Reality* can be estimated only if philosophers working independently of direct exegesis find that some of its ideas can be developed, perhaps in different terminology, and put to use in particular philosophical problems. It is likely, too, that these will be ways of thinking that take more account of the philosophy of science and vary more from the main tradition of European metaphysics than do these authors. It is a merit in Mays's book *The Philosophy of Whitehead* that it points out that behind *Process and Reality* lies the influence of Whitehead's early interest in axiomatic systems, as well as in electromagnetic field theories, especially the notion of the flow of energy. The book, however, criticizes Whitehead's realist metaphysical cosmology from the standpoint of a different philosophy of science.

It would be impossible to epitomize *Process and Reality* even in a longer treatment than can be given here. Attention can, however, be called to certain features. There is continuity with lines of thought in the earlier books, but the language becomes more naturally applicable to sentient experience. This is partly due to Whitehead's reading of Henri Bergson, F. H. Bradley, and William James, all of whom influenced him in shaping his own particular form of organic pluralism. It is also, however, due to a deliberate onslaught on the notion of "vacuous actuality," existence entirely devoid of subjective experience. Thus, Whitehead's "actual entities," while still linear events, are presented as processes of self-formation with "subjective aim." Actual entities are "epochal" happenings that take a minimal time span to become and which then perish; they are succeeded by others that conform to them and thus secure the continuity which Whitehead held was necessary if we are to have recognition of enduring objects and the expectation of continuing regularities which he believed to be necessary if induction is to be justified. The overlapping of events by other events in a field becomes the "objectification" of an actual entity in other actual entities, whereby the "feelings" and qualities of one entity are transmitted to others.

The notion of objectification is one of the most difficult of all Whitehead's views, and it is doubtful whether any satisfactory elucidation of it has yet been made. He envisaged objectification as more than a response to a stimulus and more than a causal interaction; in some sense it is a genuine reenactment of the feelings of one actual entity in another, and he maintained that we can experience this transition of feeling. The use of the term *feeling* presents great difficulty. Whitehead used it as a technical term for "the basic generic operation of passing from the objectivity of the data to the subjectivity of the actual entity in question" (*Process and Reality*, p. 55 [65]). This is to maintain that every entity, however lowly, appropriates its responses to the rest of its world in some form of sentient experience, but this does not necessarily involve consciousness. Consciousness he saw as a rare kind of sentience arising within experience; experience does not, as idealists have held, arise within consciousness.

The difficulties in this theory stem partly from Whitehead's insistence that there should not be basically different kinds of entities in the world—organic and inorganic, for instance, or minds and bodies. All entities should display the same general character. He then took certain psychological notions and generalized them (by

claiming that consciousness is incidental, not essential) to cover biological and even physical processes.

> I find myself as essentially a unity of emotions, enjoyments, hopes, fears, regrets, valuations of alternatives, decisions—all of them subjective reactions to the environment as active in my nature. My unity—which is Descartes' "*I am*"—is my process of shaping this welter of material into a consistent pattern of feelings. The individual enjoyment is what I am in my role of a natural activity, as I shape the activities of the environment into a new creation, which is myself at this moment; and yet, as being myself, it is a continuation of the antecedent world. (*Modes of Thought*, p. 228)

As a description of the kind of concrescence of prehensions I find myself to be, this is persuasive. Extended downward to describe the inner life of molecules, it strains the imagination. The possibility of making this generalization depends, Whitehead said, on our holding that "the energetic activity considered in physics is the emotional intensity entertained in life" (ibid., p. 232). Thus, Whitehead did not concern himself with the issue of freedom versus determinism as a special problem in human action. Insofar as actual entities conform to their environment and immediate past, there is determinism; insofar as any entity modifies its response through its unique subjective element of feeling, there is freedom. So freedom is a "clutch at novelty" that can appear at any point in nature.

Is it, in fact, possible to make the same general categories cover every kind of existent? Whitehead rejected "emergence" views, according to which different levels of existents may display special irreducible properties. (This view also has its difficulties.) Moreover, when Whitehead made the same "categoreal" characteristics apply to all actualities, it is possible that some of the notions he thus generalized may be of a more abstract type than others with which he connected them; one may suspect, for example, that this is so in the case of energy and emotion. Also, he held that all forms of experience—physiological and psychological and the distinctive kinds of the latter, such as moral, aesthetic, and religious—must be particular exemplifications of the same basic principles. It is by no means evident that a coherent theory of experience must imply this; there may be reasons why the principles of aesthetics, for example, might differ from those of morality or religion.

NATURAL THEOLOGY

Whitehead's interest in religion runs throughout his philosophy and is by no means confined to its later phase, though it is there that he sought to express it in a natural theology. He saw religion as sustaining a sense of the importance of an individual's experience within the social relationships and experience of his life. Beyond this broadly sociological interest, he held that religion was also concerned with permanence amid change. He connected the idea of permanence with the conception of a general ordering of the process of the world that could provide the ground first of "extensive connection," then of all more specific orderings. The ordering of the world, called "the primordial nature of God," has been compared by Mays to a sort of cosmic propositional function, a "form of definiteness" that can then be instantiated by "values," which are actual processes of events. But though Whitehead did indeed speak of the primordial nature of God as a "conceptual prehension" and, as such, "deficient in actuality," the interpretation of it as simply a formal schema omits the point that to Whitehead the notion of "conceptual prehension" includes "appetition," an urge toward the realization of the forms (or eternal objects) so prehended. This drive to realization is said to supply all particular actual entities with their "subjective forms," and God is thus represented as "the principle of concretion" whereby actual processes take their rise. God does not create other actual entities; he provides them with an initial impetus to self-creation. Each actual entity, including God, is a particular outcome of "creativity," which is said to stand for the continual process by which the many elements in the world are synthesized into new unities, each being called a "concrescence," described as a "production of novel togetherness." It is the creative advance into novelty of a pluralistic process. In response to the processes of becoming of the other actual entities of the world, God acquires a "consequent nature," in which they are "objectified" (again this difficult notion of reenactment) in his own self-formation, which appears to be coterminous with the process of nature.

The difficulties in Whitehead's natural theology are great, not least because he used traditional religious language in ways that may suggest misleading analogies. The most perceptive development of his natural theology is that of Charles Hartshorne, especially in *Philosophers Speak of God* (with William L. Reese, Chicago, 1953) and *The Logic of Perfection* (La Salle, IL, 1962). Hartshorne states, however, that his own views in natural theology were taking shape before he came in contact with Whitehead's work, which acted as a reinforcement.

WHITEHEAD'S INFLUENCE

It was suggested above that Whitehead's contribution may best appear if other philosophers find seminal ideas in it that they can develop independently. Hartshorne's work in natural theology may be one example; others would be work on concepts on the border between the physical and biological sciences, such as W. E. Agar's *A Contribution to the Theory of the Living Organism* (Melbourne, 1943), J. H. Woodger's *Biological Principles* (London, 1929), and R. S. Lillie's *General Biology and Philosophy of Organism* (Chicago, 1945). Some sociologists have also found support in Whitehead for views of societies as ongoing processes composed of subsocieties with ramified interrelations. H. H. Price has shown interest in the phenomenology of organic rather than visual sensations (see his paper "Touch and Organic Sensation," *PAS* 44 [1943–1944]: 1–30, especially his treatment of what he calls "bilateral dynamic transactions"). The main influence on contemporary philosophy is no doubt the pioneering logical work of *Principia Mathematica*.

Whitehead received the rare distinction of being awarded the Order of Merit. He had a gift for writing that showed itself at its best in the striking phrase and the vivid metaphor or analogy (some of these have been collected by A. H. Johnson in *The Wit and Wisdom of Alfred North Whitehead*, Boston, 1947). His style is less happy when this very gift of fine writing tempted him to be vaguely grandiose. Hence, rigorous critical interpretation is needed, which is more likely to be rewarding insofar as it leads to more than pure commentary.

See also Logic, History of.

Bibliography

WORKS BY WHITEHEAD

A bibliography of Whitehead's works up to 1941, including articles in journals and references to selected reviews of books, was compiled by Victor Lowe and R. C. Baldwin and is in *The Philosophy of Alfred North Whitehead*, edited by Paul A. Schilpp (Evanston, IL: Northwestern University Press, 1941, 2nd ed., New York: Tudor, 1951). George L. Kline lists Whitehead's works that have been translated into other languages in *Process and Divinity; Philosophical Essays Presented to Charles Hartshorne*, edited by William L. Reese and Eugene Freeman (La Salle, IL: Open Court, 1964), pp. 235–268. A number of essays in this book deal with aspects of Whitehead's work.

Lucien Price's *Dialogues of Alfred North Whitehead* (London, 1961) contains transcripts from memory of some of Whitehead's conversations in his last years. However, the dialogues show the side of Whitehead that came out in conversation with a classical humanist and give little impression of him as a philosopher of science.

Only principal logical and philosophical books and articles are listed below.

A Treatise on Universal Algebra. Cambridge, U.K.: Cambridge University Press, 1898.

On Mathematical Concepts of the Material World. Philosophical Transactions of the Royal Society of London, Series A (1906). Reprinted in *Alfred North Whitehead: An Anthology*, edited by F. S. C. Northrop and Mason W. Gross. Cambridge, U.K.: Cambridge University Press, 1953, pp. 11–82. This book also has extracts from Whitehead's main works and a note on terminology by Gross.

Principia Mathematica, 3 vols. Cambridge, U.K.: Cambridge University Press, 1910–1913. Written with Bertrand Russell.

An Introduction to Mathematics. London: Williams and Norgate, 1911.

The Organization of Thought. London: Williams and Norgate, 1917. This work, with others presented to the Aristotelian Society between 1916 and 1923, has been reprinted in *The Interpretation of Science*, edited by A. H. Johnson. New York: Bobbs-Merrill, 1961. The chief essays are "Space, Time and Relativity" (1915), "Time. Space and Material" (1919), and "Uniformity and Contingency" (1922); the book also contains other occasional papers on the philosophy of science and some occasional addresses given in Whitehead's last years.

An Enquiry concerning the Principles of Natural Knowledge. Cambridge, U.K.: Cambridge University Press, 1919.

The Concept of Nature. Cambridge, U.K.: Cambridge University Press, 1920.

The Principle of Relativity. Cambridge, U.K.: Cambridge University Press, 1922.

Science and the Modern World. New York: Macmillan, 1925.

Religion in the Making. New York: Macmillan, and Cambridge, U.K.: Cambridge University Press, 1926.

Symbolism, Its Meaning and Effect. New York; Macmillan, 1927: Cambridge, U.K.: Cambridge University Press, 1928.

Process and Reality. New York: Macmillan, and Cambridge, U.K.: Cambridge University Press, 1929. Two separate editions.

The Function of Reason. Princeton, NJ: Princeton University Press, 1929.

The Aims of Education and Other Essays. New York: Macmillan, and London: Williams and Norgate, 1929. This includes shorter versions of *The Organization of Thought* and "Space, Time and Relativity," as well as "The Anatomy of Some Scientific Ideas."

Adventures of Ideas. New York: Macmillan, and Cambridge, U.K.: Cambridge University Press, 1933.

Nature and Life. Chicago: University of Chicago Press, and Cambridge, U.K.: Cambridge University Press, 1934. The best general presentation of Whitehead's dissatisfaction with the Newtonian scheme of the material world.

Modes of Thought. New York: Macmillan, and Cambridge, U.K.: Cambridge University Press, 1938.

Essays in Science and Philosophy. New York: Philosophical Library, 1947.

"Immortality" and "Mathematics and the Good," in Schilpp, op. cit. Whitehead's last two lectures.

WORKS ON WHITEHEAD

Haack, S. "Descriptive and Revisionary Metaphysics." *Philosophical Studies* 35 (1979): 361–371.

Hartshorne, Charles. *Whitehead's Philosophy: Selected Essays, 1935–1970*. Lincoln: University of Nebraska Press, 1972.

Leclerc, I., ed. *The Relevence of Whitehead*. New York: Humanities Press, 1961.

Leclerc, I. *Whitehead's Metaphysics: An Introductory Exposition*. 2nd ed. London: Allen & Unwin, 1965.

Lowe, V. *Alfred North Whitehead: The Man and His Work*. 2 vols. Baltimore, MD: The Johns Hopkins University Press, 1985, 1990.

Lowe, V. *Understanding Whitehead*. 2nd ed. Baltimore, MD: The Johns Hopkins University Press, 1966.

Nobo, J. L. *Whitehead's Metaphysics of Extension and Solidarity*. Albany: State University of New York Press, 1986.

Palter, R. M. *Whitehead's Philosophy of Science*. Chicago: University of Chicago Press, 1960.

Quine, W. V. "Whitehead and the Rise of Modern Logic." In *The Philosophy of A. N. Whitehead*, edited by P. A. Schilpp. La Salle, IL: Open Court, 1941.

Ross, Stephen David. *Perspectives in Whitehead's Metaphysics*. Albany: State University of New York Press, 1983.

Russell, Bertrand. "Whitehead and Principia Mathematica." *Mind* 57 (1948): 137–138.

Dorothy M. Emmet (1967)
Bibliography updated by Benjamin Fiedor (2005)

WHY

Lack of clarity about the uses of the word *why* is responsible for confusion on a number of philosophical fronts. In this entry we shall confine ourselves to two groups of topics where greater attention to the proper and improper behavior of this word might well have avoided the adoption of misguided theories. There is, first, the contrast, or the alleged contrast, between the "how" and the "why" and the view, shared by writers of very different backgrounds, that science can deal only with how-questions. Second, there are certain "ultimate" or "cosmic" questions, such as "Why do we exist?" or, more radically. "Why does the world exist?" or "Why is there something rather than nothing?" Some, like Arthur Schopenhauer and Julian Huxley, regard these questions as unanswerable; others, like Étienne Gilson and F. C. Copleston, believe that they can be answered; but whether these questions can be answered or not, it seems to be widely agreed that they are very "deep." These questions, in the words of the British astrophysicist A. C. B. Lovell, raise problems "which can tear the individual's mind asunder" (*The Individual and the Universe,* New York, 1961, p. 125). Speaking of the question "Why is there something rather than nothing?," Martin Heidegger first remarks that it is "the fundamental question of metaphysics" and later adds that "with this question philosophy began and with this question it will end, provided that it ends in greatness and not in an impotent decline" (*An Introduction to Metaphysics,* p. 20).

HOW AND WHY

The contrast between the how and the why has been insisted on for two rather different reasons. Some writers have done so in the interest of religion or metaphysics. Their position seems to be that while science and empirical research generally are competent to deal with how-questions, the very different and much deeper why-questions are properly the concern of religion or metaphysics or both. Thus, in a widely read book the British psychiatrist David Stafford-Clark insists that the confusion between the how and the why is the "fundamental fallacy" behind "the whole idea that science and religion are really in conflict at all" (*Psychiatry Today,* Harmondsworth, U.K., 1952, p. 282). Sigmund Freud in particular is accused of committing this fallacy in his antireligious writings. Stafford-Clark is not at all opposed to Freudian theory so long as it confines itself to the how of psychological phenomena. Psychoanalysis cannot, however, "begin by itself to answer a single question as to why man is so constructed that they should happen in this way" (p. 287). Although he repeatedly expresses his own fervent belief in God, Stafford-Clark unfortunately does not tell us how religion answers the question why man is "constructed" the way he is. Perhaps he would answer it along the lines in which Isaac Newton answered a similar question about the sun. "Why is there one body in our system qualified to give light and heat to all the rest," Newton wrote in his first letter to Richard Bentley, "I know no reason, but because the author of the system thought it convenient" (*Opera,* London, 1779–1785, Vol. IV, pp. 429ff.).

Similar views are found in the writings of many professional philosophers. Thus, writing of Newton's work on gravitation, A. N. Whitehead observes that "he [Newton] made a magnificent beginning by isolating the stresses indicated by his law of gravitation." But Newton "left no hint, why in the nature of things there should be any stresses at all" (*Modes of Thought,* New York and Cambridge, U.K., 1938, pp. 183–184). Similarly, discussing the limitations of science, Gilson declares that "scientists never ask themselves *why* things happen, but *how* they happen.… Why anything at all is, or exists, science knows not, precisely because it cannot even ask the question" (*God and Philosophy,* New Haven, CT, 1959, p.

140). For Gilson the two topics mentioned at the beginning of this entry appear to merge into one. The why of particular phenomena, he seems to argue, cannot be determined unless we answer the question "why this world, taken together with its laws … is or exists" (p. 72).

Among those who have asserted that science can only deal with how-questions there are some who are not at all friendly to metaphysics or religion. These writers usually add to their remarks that science cannot handle why-questions the comment that no other enterprise fares any better. This "agnostic positivism," as we may call it, goes at least as far back as David Hume. We know, he writes, that milk and bread are proper nourishment for men and not for lions or tigers, but we cannot "give the ultimate reason why" this should be so (*An Inquiry concerning Human Understanding,* Sec. IV, Part I). Hume seems to imply that this unhappy state can never be remedied, regardless of the advances of physiology or any other science. Several writers in the second half of the nineteenth century advanced this position under the slogan "The task of science is to describe phenomena, not to explain them." Ernst Mach, Gustav Kirchhoff, and Joseph Petzoldt were among the best-known figures in central Europe who advocated this view. In England, Karl Pearson, its most influential exponent, conceded that there was no harm in speaking of "scientific explanations" so long as *explanation* is used "in the sense of the descriptive-*how*" (*The Grammar of Science,* Everyman edition, 1937, p. 97). We can indeed "describe how a stone falls to the earth, but not why it does" (p. 103). "No one knows why two ultimate particles influence each other's motion. Even if gravitation be analyzed and described by the motion of some simpler particle or ether-element, the whole will still be a description, and not an explanation, of motion. Science would still have to content itself with recording the *how.*" No matter how far physics may progress, the why will "remain a mystery" (p. 105).

It is important to disentangle purely verbal from substantive issues in all of this. Insofar as the various writers we have quoted merely wish to assert that causal statements and scientific laws in general are contingent and not logically necessary propositions, little exception could be taken to their remarks. However, they are, or at least they appear to be, saying a great deal more. They all seem to agree that there is a class of meaningful questions, naturally and properly introduced by the word *why* in one of its senses, which cannot be answered by the use of empirical methods. Writers belonging to the first group claim that the answers can be obtained elsewhere. The agnostic positivists maintain that human beings cannot obtain the answers at all.

It is this substantive issue which we shall discuss here, and it is necessary to point out that there are numerous confusions in all views of this kind. To begin with, although this is the least important observation, *how* and *why* do not always have contrasting functions but are in certain situations used to ask the very same questions. Thus, when we know or believe that a phenomenon, *A,* is the cause of another phenomenon, *X,* but at the same time are ignorant of the "mechanics" of *A*'s causation of *X,* we indifferently use *how* and *why.* We know, for example, that certain drugs cure certain diseases, but our knowledge is in a medical sense "purely empirical." Here we would be equally prepared to say that we do not know "why" the drug produces the cure and that we do not know "how" it does this. Or, to take a somewhat different case, it is widely known that cigarette smoking is causally connected with lung cancer. It is also known that sometimes two people smoke the same amount and yet one of them develops lung cancer while the other one does not. In such a case the question naturally arises why cigarette smoking, if it is indeed the cause at all, leads to cancer in one case but not in the other. And we would be just as ready to express our ignorance or puzzlement by saying that we do not know how it is as by saying that we do not know why it is that smoking produced cancer in the first man but not in the second. In all such cases it is clear that science is in principle competent to deal with the "why" no less than with the "how," if only because they are used to ask the very same questions.

It is undeniable, however, that in certain contexts *how* and *why* are used to ask different questions. This contrast is most obvious when we deal with intentional, or more generally with "meaningful," human actions. What seems far from obvious, what in fact seems plainly false, is that empirical methods are not in principle adequate to determine the answers to why-questions in these contexts. Let us take as our example the theft of the Star of India sapphire and other gems from the Museum of Natural History in New York. We can here certainly distinguish the question why the burglary was committed from the question how it was carried out. The latter question would concern itself with the details of the act—how the thieves got into the building, how they immobilized the alarm system, how they avoided the guards, and so on. The why-question, by contrast, would inquire into the aim or purpose of the theft—were the thieves just out to make a vast amount of money, or were there perhaps some other aims involved, such as proving to rival gangs

how skillful they were or showing the incompetence of the police force?

Now, the aim or purpose of a human being is surely not in principle undiscoverable, and frequently we know quite well what it is. The person himself usually, though not always, simply knows what his aim is. An orator, for example, who is advocating a certain policy, ostensibly because it is "for the good of the country," may at the same time know perfectly well that his real aim is personal advancement. It used to be said that in such situations a human being knows his own purpose by means of "introspection," where introspection was conceived of as a kind of "inner sense." This way of talking is not inappropriate to situations in which somebody is confused about his own motives, for then special attention to his own feelings, resembling in some ways the effort to discriminate the detailed features of a physical scene, may well be necessary in order to ascertain his "true" aims.

Much more commonly, however, a human being simply knows what his aims are, and it would be much better to say that he knows this "without observation" than that he knows it by introspection. In order to find out the purpose of somebody else's action, it is in countless instances sufficient to ask the person a direct question about his aim. Where the agent's veracity is suspect or where a person is the victim of self-deception, it is necessary to resort to more elaborate investigations. In the former type of case one might ask the agent all kinds of other questions (that is, questions not directly about the purpose of his action), one might interview his friends and acquaintances and other witnesses of his conduct, one might tap his telephone and employ assorted bugging devices, and one might perhaps go so far as to question him after the administration of "truth" drugs. In the latter type of case it may not be possible to ascertain the real purpose unless the person undertakes psychiatric treatment. While the practical difficulties in the way of discovering the purpose of an action are no doubt insurmountable in many cases of both these types, empirical procedures are clearly in principle adequate to this task.

We also contrast how- and why-questions when the latter are not inquiries into the purpose of any agent. Here, however, *how* has a different meaning from any previously discussed. In all examples so far considered, how-questions were in one way or another causal questions—"How did the thieves carry out their plan of stealing the Star of India?" is a question about the means of achieving a certain goal, and "How is it that smoking produces cancer in one man but not in another?," although not a question about means, is nevertheless about the processes leading to a certain result. These causal "hows" should be distinguished from what one may call the "how" of "state" or "condition." "How cold does it get in New York in the winter?" "How does the decline in his powers manifest itself?" "How is his pain now—is it any better?" are examples of the "how" of state or condition, and it is how-questions of this kind which we contrast with nonteleological why-questions—"Why does it get so cold in New York in the winter?" "Why did his powers decline so early in life?" "Why is his pain not subsiding?"

It is sometimes maintained or implied, as in the remarks of Stafford-Clark quoted earlier, that why-questions are invariably inquiries about somebody's purpose or end—if not the purpose of a human being, then perhaps that of some supernatural intelligence. This is clearly not the case. There can be no doubt that *why* is often employed simply to ask questions about the cause of a phenomenon. Thus the question "Why are the winters in New York so much colder than in Genoa, although the two places are on the same geographical latitude?" would naturally be understood as a request for information about the cause of this climatic difference, and it is not necessary for the questioner to suppose that there is some kind of plan or purpose behind the climatic difference in order to be using the word *why* properly. In saying this, one is not begging any questions against the theory that natural phenomena like the cold of the winter in New York are the work of a supernatural being: One is merely calling attention to what is and what is not implied in the ordinary employment of *why* in these contexts.

Let us briefly summarize the results obtained so far: In some situations *how* and *why* are naturally employed to ask the very same questions; when we deal with intentional human actions, we naturally use *why* to inquire about the purpose or goal of the agent and *how* to learn about the means used to achieve that goal; finally, how-questions are frequently used to inquire about the state or condition of somebody or something, while why-questions inquire about the cause of that state or condition without necessarily implying that any purpose or plans are involved. In all these cases it appears to be in principle possible to answer why-questions no less than how-questions, and this without the aid of religion or metaphysics.

THE THEOLOGICAL "WHY"

Let us turn now to what we earlier called "cosmic" why-questions. Two such cosmic "whys" need to be distin-

guished, the first of which, for rather obvious reasons, will be referred to as the theological "why." Here the questioner would be satisfied with a theological answer if he found such an answer convincing in its own right. He may or may not accept it as true, but he would not regard it as irrelevant.

Gilson, whose remarks on the limitations of science were quoted earlier, immediately supplies the answer to the "supreme question" which science "cannot even ask." Why anything at all exists must be answered by saying:

> [Each] and every particular existential energy, and each and every particular existing thing depends for its existence upon a pure Act of existence. In order to be the ultimate answer to all existential problems, this supreme cause has to be absolute existence. Being absolute, such a cause is self-sufficient; if it creates, its creative act must be free. Since it creates not only being but order, it must be something which at least eminently contains the only principle of order known to us in experience, namely, thought. (*God and Philosophy,* p. 140)

There is no doubt that many people who ask such questions as "Why does the universe exist?" or "Why are we here?" would also, at least in certain moods, be satisfied with a theological answer, though they would not necessarily accept all the details of Gilson's Thomistic theology. It should be emphasized that one does not have to be a believer in God to be using *why* in this way. The American playwright Edward Albee, for example, once remarked, "Why we are here is an impenetrable question." Everyone in the world, he went on, "hopes there is a God," and he later added, "I am neither pro-God nor anti-God" (*New York Times,* January 21, 1965). Albee's question "Why are we here?" evidently amounts to asking whether there is a God and, if so, what divine purposes human beings are supposed to serve. He does not definitely accept the theological answer, presumably because he feels unsure of its truth, but he does regard it as very much to the point.

It should be observed in passing that people frequently use the word *why* to express a kind of cosmic complaint or bewilderment. In such cases they are not really asking for an answer, theological or otherwise. This use of *why* is in some respects similar to the theological "why" and may not inappropriately be referred to as the quasi-theological "why." A person who is and regards himself as a decent human being, but who is suffering a great deal, might easily exclaim "Why do I have to suffer so much, when so many scoundrels in the world, who never worked half as hard as I, are having such a lot of fun?" Such a question may well be asked by an unbeliever who is presumably expressing his regret that the workings of the universe are not in harmony with the moral demands of human beings. Even when believers ask questions of this kind, it may be doubted that they are invariably requesting information about the detailed workings of the Divine Mind. In the deeply moving first-act monologue of *Der Rosenkavalier,* the Marschallin reflects on the inevitability of aging and death:

> I well remember a girl
> Who came fresh from the convent to be forced
> into holy matrimony.
> Where is she now?
> . . . .
> How can it really be,
> That I was once the little Resi
> And that I will one day become the old
> woman?

How, she exclaims, can something like this be? She is far from doubting the existence of God and proceeds to ask:

> Why does the dear Lord do it?

And worse, if he has to do it in this way:

> Why does He let me watch it happen
> With such clear senses? Why doesn't He hide it
> from me?

The Marschallin obviously does not expect an answer to this question, not, or not merely, because she thinks that the world's metaphysicians and theologians are not quite up to it. She is not, strictly speaking, asking a question but expressing her regret and her feeling of complete helplessness.

However, let us return from the quasi-theological to the theological "why." The difficulties besetting an answer like Gilson's are notorious and need not be reviewed here at length. There are the difficulties, much stressed by recent writers, of saying anything intelligible about a disembodied mind, finite or infinite, and there are further difficulties of talking meaningfully about the creation of the universe. There are the rather different difficulties connected not with the intelligibility of the theological assertions but with the reasoning used to justify them. Schopenhauer referred to all such attempts to reach a final resting place in the series of causes as treating the causal principle like a "hired cab" which one dismisses when one has reached one's destination. Bertrand Russell objects that such writers work with an obscure and objectionable notion of explanation: to explain something, we are not at all required to introduce a "self-sufficient"

entity, whatever that may be. Writing specifically in reply to Gilson, Ernest Nagel insists that it is perfectly legitimate to inquire into the reasons for the existence of the alleged absolute Being, the pure Act of existence. Those who reject such a question as illegitimate, he writes, are "dogmatically cutting short a discussion when the intellectual current runs against them" (*Sovereign Reason,* Glencoe, IL, 1954, p. 30). Without wishing to minimize these difficulties, it is important to insist that there is a sense in which the theological why-questions are intelligible. The question can be answered for such a person if it can be shown that there is a God. If not, it cannot be answered. Albee and Gilson, for example, do not agree about the truth, or at any rate the logical standing, of the theological assertion, but they agree that it is relevant to their cosmic why-question. There is thus a sense in which the questioner here knows what he is looking for.

THE SUPERULTIMATE "WHY"

The theological "why" must be distinguished from what we are here going to call the superultimate "why." A person who is using *why* in the latter way would regard the theological answer as quite unsatisfactory, not (or not just) because it is meaningless or false but because it does not answer his question. It does not go far enough. For granting that there is a God and that human beings were created by God to serve certain of his purposes, our questioner would now ask "Why is there a God of this kind with these purposes and not another God with other purposes?" or, more radically, he would ask "Why was there at some time God rather than nothing?" The biblical statement "In the beginning God created heaven and earth," Heidegger explicitly remarks, "is not an answer to … and cannot even be brought into relation with our question." The believer who stops with God is not pushing his questioning "to the very end" (*An Introduction to Metaphysics,* pp. 6–7). (It is not certain how somebody pressing the superultimate why-question would react to the rejoinder of those theologians who maintain that God exists necessarily and that hence the question "Why was there at some time God rather than nothing?" is illegitimate. In all likelihood he would support the view, accepted by the majority of Western philosophers since Hume and Immanuel Kant, that it makes no sense to talk about anything, natural or supernatural, as existing necessarily.)

There are times when most people would regard these superultimate why-questions as just absurd. Stafford-Clark himself speaks with impatience of the "rumination" and the tedious and interminable speculations of obsessional patients. "'Why is the world?' was a question to which one patient could find no answer but from which he could find no relief" (*Psychiatry Today,* p. 112). Yet, at other times, most of us are ready to treat these why-questions as supremely profound, as riddles to which it would be wonderful to have the answer but which, because of our finite intellects, must forever remain unsolved. It is true that certain philosophers, like Friedrich von Schelling and Heidegger, who have frequently been denounced as obscurantists, have laid special emphasis on superultimate why-questions; but it would be a total misunderstanding of the situation to suppose that more empirical philosophers, or indeed ordinary people, are not given to asking them or to treating them with great seriousness. It is almost unavoidable that any reasonably intelligent and reflective person who starts wondering about the origin of the human race, or animal life, or the solar system, or our galaxy and other galaxies, or about the lack of justice in the world, the brevity of life, and seeming absolute finality of death, should sooner or later ask "Why this world and not another—why any world?"

The scientist Julian Huxley is as far removed in temperament and philosophy from Heidegger as anybody could be. Yet he also speaks of the "basic and universal mystery—the mystery of existence in general … why does the world exist?" For Huxley it is science that "confronts us" with this mystery, but science cannot remove it. The only comment we can make is that "we do not know." We must accept the existence of the universe "and our own existence as the one basic mystery" (*Essays of a Humanist,* London, 1964, pp. 107–108). Ludwig Büchner was a materialist and an atheist, and yet he repeatedly spoke of the "inexplicability of the last ground of things." Nor are superultimate why-questions confined to those who do not believe in God or who have no metaphysical system. Schopenhauer was supremely confident that his was the true metaphysic, but he nevertheless remarks in the concluding chapter of his main work that his "philosophy does not pretend to explain the existence of the world in its ultimate grounds.… After all my explanations," he adds, "one may still ask, for example, whence has sprung this will, the manifestation of which is the world.… A perfect understanding of the existence, nature, and origin of the world, extending to its ultimate ground and satisfying all demands, is impossible. So much as to the limits of my philosophy, and indeed of all philosophy" (*The World as Will and Idea,* 3 vols., translated by R. B. Haldane and J. Kemp, London, 1883, Ch. 50)

Similarly, Voltaire, who was a firm and sincere believer in God and who never tired of denouncing atheists as blind and foolish, nevertheless asked, at the end of the article "Why?" in his *Philosophical Dictionary,* "Why is there anything?," without for a moment suggesting that an appeal to God's creation would be a solution. William James, too, although he repeatedly defended supernaturalism, never claimed that it provided an answer to the question "How comes the world to be here at all instead of the non-entity which might be imagined in its place?" Philosophy, in James's opinion, whether it be naturalistic or supernaturalistic, "brings no reasoned solution" to this question, "for from nothing to being there is no logical bridge" (*Some Problems of Philosophy,* New York, 1911, pp. 38–40). "The question of being," he observes later in the same discussion, is "the darkest in all philosophy. All of us are beggars here, and no school can speak disdainfully of another or give itself superior airs" (ibid., p. 46).

Having pointed out how widespread is this tendency to ask and take seriously the superultimate why-question, it is necessary to explain why, in the opinion of a number of contemporary philosophers, it must nevertheless be condemned as meaningless. It is the mark of a meaningful question, it would be urged, that not all answers can be ruled out a priori; but because of the way in which the superultimate why-question has been set up, it is logically impossible to obtain an answer. It is quite clear that the questioner will automatically reject any proposed answer as "not going back far enough"—as not answering his why. "All explanation," in the words of Peter Koestenbaum, an American disciple and expositor of Heidegger, "occurs within that which is to be explained … so the question applies to any possible answer as well" ("The Sense of Subjectivity," p. 54), that is, there cannot be an answer. If, however, a question can be put at all, to quote Wittgenstein,

> then it *can* also be answered … doubt can only exist where there is a question; a question only where there is an answer, and this only where something *can* be *said.* (*Tractatus Logico-Philosophicus,* 6.5 and 6.51)

It must be emphasized that the superultimate "why" does not express ignorance about the "early" history of the universe. Büchner, for example, had no doubt that matter was eternal and that nothing which could be called "creation" had ever occurred; Voltaire similarly had no doubt that the physical universe was created by God and that God had always existed—yet both of them asked the superultimate "why" and regarded it as unanswerable. No doubt, some who have asked superultimate why-questions would, unlike Büchner and Voltaire, declare themselves ignorant of the remote history of the universe, but it is not this ignorance that they are expressing by means of the superultimate "why."

Those who insist that the superultimate why-question is meaningful do not usually deny that it very radically differs from all other meaningful why-questions. To mark the difference they occasionally refer to it by such labels as "mystery" or "miracle." Thus Koestenbaum remarks that "questions of this sort do not lead to answers but to a state of mind that appreciates the miracle of existence," they call attention to "the greatest of all mysteries" (op. cit., pp. 54–55). Heidegger writes that the question "is incommensurable with any other" (*An Introduction to Metaphysics,* p. 4) and subsequently observes that "not only what is asked after but also the asking itself is extraordinary" (ibid., p. 10).

Calling the superultimate why-question a "mystery" or a "miracle" or "incommensurable" or "extraordinary" does not in any way remove the difficulty: It is just one way of acknowledging that there is one. If it is granted that in all other situations a question makes sense only if an answer to it is logically possible, one wonders why this principle or criterion is not to be applied in the present case. If the defender of the meaningfulness of the superultimate why-question admits that in the "ordinary" sense the question is meaningless but that in some other and perhaps deeper sense it is meaningful, one would like to be told what this other and deeper sense is.

The point of the preceding paragraphs is sometimes expressed in a way that is not totally satisfactory. It is maintained that a question does not make sense unless the questioner knows what kind of answer he is looking for. However, while the fact that the questioner knows the "outline" of the answer may be a strong or even conclusive reason for supposing that the question is meaningful, the converse does not hold. One can think of examples in which a question is meaningful although the person asking it did not know what a possible answer would look like. Thus somebody might ask "What is the meaning of life?" without being able to tell us what kind of answer would be relevant and at a later time, after falling in love for the first time, he might exclaim that he now had the answer to his question—that love was the meaning of life. It would be much better to say in such a case that the question, as originally asked, was not clear than to say that it was meaningless. It is not objectionable to condemn a question as meaningless on the ground that the questioner does not know what he is looking for if in the context this is a way of saying that he has ruled out all

answers a priori; and very probably those who express themselves in this way do not mean to point to some contingent incapacity on the part of the questioner but, rather, to a disability consequent upon the logical impossibility of obtaining an answer to the question. It is similar to saying that it is inconceivable that 3 plus 2 should equal 6 when we do not mean to assert a contingent fact about a certain incapacity on the part of human beings but, rather, that "3 plus 2 equals 6" is a self-contradiction.

The conclusion that the superultimate why-question is meaningless can also be reached by attending to what has here happened to the word *why.* A little reflection shows that in the superultimate question "why" has lost any of its ordinary meanings without having been given a new one. Let us see how this works when the question is put in the form "Why does the universe exist?" and when the "universe" is taken to include everything that in fact exists. In any of its familiar senses, when we ask of anything, *x,* why it happened or why it is what it is—whether *x* is the collapse of an army, a case of lung cancer, the theft of a jewel, or the stalling of a car—we assume that there is something or some set of conditions, other than *x,* in terms of which it can be explained. We do not know what this other thing is that is suitably related to *x,* but unless it is in principle possible to go beyond *x* and find such another thing, the question does not make any sense. (This has to be slightly modified to be accurate. If we are interested in the "why" of a state of *x* at a certain time, then the answer can certainly refer to an earlier state of *x.* This does not affect the issue here discussed since, in the sense with which we are concerned, reference to an earlier state of *x* is going beyond *x.*) Now, if by "the universe" we mean the totality of things, then our *x* in "Why does the universe exist?" is so all-inclusive that it is logically impossible to find anything which could be suitably related to that whose explanation we appear to be seeking. "The sense of the world," wrote Wittgenstein, "must lie outside the world" (*Tractatus Logico-Philosophicus,* 6.41), but by definition nothing can be outside the world. Heidegger, who avoids the formulation "Why does the universe exist?" and who instead inquires into the why of *das seiende* (the official translation of this term is "the essent," but Koestenbaum and others quite properly translate it as "things"), nevertheless makes it clear that *das seiende* here "takes in everything, and this means not only everything that is present in the broadest sense but also everything that ever was or will be." "Our question," he writes a little later, presumably without seeing the implications of this admission, "reaches out so far that we can never go further" (*An Introduction to Metaphysics,* p. 2).

For anybody who is not clearly aware of what we may call the logical grammar of *why* it is very easy to move from meaningful why-questions about particular things to the meaningless why-question about the universe. This tendency is aided by the picture that many people have of "the universe" as a kind of huge box that contains all the things "inside it." Voltaire's article "Why?," from which we quoted earlier, is a good example of such an illegitimate transition. Voltaire first asks a number of why-questions about specific phenomena, such as

> Why does one hardly ever do the tenth part good one might do? Why in half Europe do girls pray to God in Latin, which they do not understand? Why in antiquity was there never a theological quarrel, and why were no people ever distinguished by the name of a sect?

He then gets more and more philosophical:

> Why, as we are so miserable, have we imagined that not to be is a great ill, when it is clear that it was not an ill not to be before we were born?

A little later we have what may well be a theological "why":

> Why do we exist?

Finally, as if there had been no shift in the meaning of *why* Voltaire asks:

> Why is there anything?

It should be noted that the argument we have just presented is not in any way based on an empiricist meaning criterion or on any question-begging assumptions in favor of naturalism. Anybody who uses the word *universe* in a more restricted sense, so that it is not antecedently impossible to get to an entity that might be the explanation of the universe, may be asking a meaningful question when he asks "Why does the universe exist?" Furthermore, even if *universe* is used in the all-inclusive sense, what we have said does not rule out the possibility that God or various divine beings are part of the universe in this sense. The point has simply been that the word *why* loses its meaning when it becomes logically impossible to go beyond what one is trying to explain. This is a matter on which there need not be any disagreement between atheists and theists or between rationalists and empiricists.

It will be well to bring together the main conclusions of this entry:

(1) There is a sense in which *how* and *why* have roughly the same meaning. In this sense science is perfectly competent to deal with the *why.*

(2) There are certain senses in which *how* and *why* serve to ask distinct questions, but here too both types of questions can in principle be answered by empirical procedures.

(3) One of the cosmic "whys"—what we have called the theological "why"—is used to ask meaningful questions, at least if certain semantic problems about theological utterances are disregarded. It was pointed out, however, that this does not imply that the theological answers are true or well supported.

(4) Some apparent questions introduced by "why" are really complaints and not questions, and for this reason unanswerable.

(5) What we have called the superultimate "why" introduces questions that are devoid of sense, whether they are asked by ordinary people in their reflective moments or by philosophers.

See also Explanation; Gilson, Étienne Henry; Heidegger, Martin; Hume, David; Mach, Ernst; Newton, Isaac; Pearson, Karl; Petzoldt, Joseph; Schelling, Friedrich Wilhelm Joseph von; Schopenhauer, Arthur; Voltaire, François-Marie Arouet de; Whitehead, Alfred North; Wittgenstein, Ludwig Josef Johann.

Bibliography

Wittgenstein returned to a discussion of cosmic why-questions in a lecture given in 1930 which was published for the first time under the title "A Lecture on Ethics," in *Philosophical Review* (1965). He makes it clear that although he regards the questions as nonsensical, he "deeply respects" the tendency to ask such questions. The complete text of Voltaire's article "Why?," sometimes called "The Whys," is available in the six-volume edition of the *Philosophical Dictionary* published in London by J. Hunt and H. L. Hunt in 1824. Views similar to those expressed in the last section of the present article are defended in John Passmore, "Fact and Meaning," in *Thinking and Meaning* (Louvain and Paris, 1963). Jean-Paul Sartre appears to reach similar conclusions in the final section of *Being and Nothingness,* translated by H. E. Barnes (New York: Philosophical Library, 1956).

Heidegger's fullest discussion of the superultimate why-question occurs in Ch. 1 of *Einführung in die Metaphysik* (Tübingen: Niemeyer, 1953), translated by Ralph Manheim as *An Introduction to Metaphysics* (New Haven, CT: Yale University Press, 1959). Koestenbaum's treatment is contained in his "The Sense of Subjectivity," in *Review of Existential Psychology and Psychiatry* 2 (1962): 47–64. Max Scheler discusses the superultimate why-question in his essay "Vom Wesen der Philosophic und der moralischen Bedingung des philosophischen Erkennens," in *Gesammelte Werke,* edited by Maria Scheler, Vol. V (Bern: Francke, 1954). His position seems to be very similar to that of Heidegger and other existentialists. Scheler concludes that "he who has not, as it were, looked into the abyss of the absolute Nothing will completely overlook the eminently positive content of the realization that there is something rather than nothing" (pp. 93–94).

The only detailed attempt to reply to arguments such as those urged in the present entry and to show that the superultimate why-question is meaningful, although it is in principle unanswerable, is found in M. K. Munitz, *The Mystery of Existence* (New York: Appleton-Century-Crofts, 1965). Clearly theological uses of "why" occur in Ch. 7 of Richard Taylor, *Metaphysics* (Englewood Cliffs, NJ: Prentice-Hall, 1963) and in F. C. Copleston's remarks in his debate with A. J. Ayer, "Logical Positivism," in *A Modern Introduction to Philosophy,* edited by Paul Edwards and Arthur Pap, 2nd ed. (New York: Free Press, 1965). There are some interesting remarks on what we have here been calling the quasi-theological "why" in Ch. 14 of S. E. Toulmin, *The Place of Reason in Ethics* (Cambridge, U.K.: Cambridge University Press, 1950).

The general topic of what makes a question meaningful has only very rarely been discussed by philosophers. Rudolf Carnap, in *Der logische Aufbau der Welt,* Part V, Sec. E (Berlin: Weltkreis, 1928; 2nd ed., Hamburg, 1961), and Moritz Schlick, in "Unanswerable Questions?," in *Philosopher* (1935), reprinted in his *Gesammelte Aufsätze* (Vienna: Gerold, 1938), propose empiricistic meaning criteria and conclude that questions that cannot even in principle be answered must be condemned as meaningless. However, as was pointed out in the text, this conclusion does not depend on the adoption of an empiricistic meaning criterion. Thus the phenomenologist Oskar Becker writes that "according to the principle of transcendental idealism a question which is in principle undecidable has no sense—to it there corresponds no possible state of affairs which could supply an answer" ("Beiträge zur phänomenologischen Begründung der Geometrie und ihrer physikalischen Anwendungen," in *Jahrbuch für Philosophie und phänomenologische Forschung* 6 (1923): 412. There are numerous suggestive remarks in Ch. 20 of Friedrich Waismann's posthumously published *The Principles of Linguistic Philosophy* (New York: St. Martin's Press, 1965).

On *how* and *why,* in addition to the works quoted in the text, mention should be made of James Martineau, *Modern Materialism* (New York: Putnam, 1877), where the view is defended that science cannot deal with the "why." Agnostic positivism is defended in E. W. Hobson, *The Domain of Natural Science* (Cambridge, U.K.: Cambridge University Press, 1923). A J. Ayer in the debate with Copleston supports the position that science can handle why-questions so long as they are intelligible.

When we ask why a person acted in a certain way or why he holds a certain belief, we frequently ask for an explanation in terms of reasons. It has been argued by a number of recent writers that such explanations cannot be regarded as a species of causal explanation—at any rate in the sense in which we habitually search for causal explanations in the natural sciences. This topic has not been discussed in the present entry since it is treated at some length elsewhere in this encyclopedia (see the entry Philosophy of History).

Paul Edwards (1967)

WIGGINS, DAVID
(1933–)

David Wiggins was professor of philosophy at Bedford College, London; professor of philosophy at Birkbeck College, London; Wykeham Professor of Logic at Oxford University; and a fellow of New College, Oxford. He has published in metaphysics, philosophy of language, moral and political philosophy, and the history of philosophy. His major works are *Identity and Spatio-temporal Continuity*; *Sameness and Substance*; *Needs, Values, and Truth: Essays in the Philosophy of Value*; and *Sameness and Substance Renewed*.

The most influential part of Wiggins's work has been in metaphysics, where he has developed a fundamentally Aristotelian conception of substance, enriched by insights drawn from Putnam (1975) and Kripke (1980). His works also contain influential discussions of the problem of personal identity, which Wiggins elucidates via a conception that he calls the "Animal Attribute View."

Wiggins's metaphysic of substance embodies several contentions. The first is that a distinction can be drawn between sortal and nonsortal concepts, the former providing answers to the question "What is it?" asked of a substance. If *a* and *b* are the same, there must be an answer to the question "The same what?" This answer can be provided by a sortal concept satisfied by both *a* and *b*. This thesis implies that any substance satisfies at any time some sortal or other.

Wiggins also maintains that any substance must satisfy the same substance sortal throughout its existence, though it will also satisfy various phase sortals that apply to it only at certain stages of its career. For example, "child" is a phase sortal, while "man," Wiggins says, is a substance sortal. Protean change is not possible. Following Quine (1960), some opponents of this view hold that substances are not to be distinguished from events or processes, and can be thought of as having temporal parts. These proponents of "four-dimensionalism," as the doctrine of temporal parts is commonly called, also typically hold that any temporal part of one object and any temporal part of the same or another object can be thought of as constituting a third object (Quine 1960, Lewis 1986). There is, for example, the object consisting of the first decade of Aristotle and the third decade of the Eiffel Tower. This thesis is sometimes referred to as mereological universalism, or unrestricted composition. Wiggins's thesis that any substance must satisfy some one substance sortal throughout its existence is intended to be inconsistent with mereological universalism. More fundamentally, Wiggins argues against four-dimensionalism.

Another significant component of Wiggins's metaphysics is his denial of relative identity. Wiggins maintains that identity is not relative to different sortals, in the sense that *a* and *b* may be the same *f* but different *g*'s. The relative-identity thesis was introduced into modern debate by Peter Geach (1972) and appears to be illustrated by familiar kinds of change. For example, an old general is the same person or human being as the young boy he was, but he is not the same child, since the old general is not a child. Again, if a piece of clay is reshaped to make different statues, it is the same piece of clay throughout, but not the same statue. To deal with such examples, Wiggins appeals to (1) the distinction between phase sortals and substance sortals and (2) the distinction between constitution and identity. The first type of example, he suggests, can be dealt with merely by paying proper attention to tense: The general was the same child as the boy, and the boy will be the same man as the general. In the second type of case he suggests that we must recognize that the piece of clay is distinct from all the statues it successively constitutes. We can correctly say that the clay is at one time a statue of Goliath, say. But this is because one of the meanings of "is" is "constitutes"—a meaning that must be recognized in addition to the "is" of predication and the "is" of identity.

Wiggins opposes relative identity because he sees it as incompatible with Leibniz's Law, the principle that if *a* is identical with *b*, *a* and *b* must share all their properties. Some opponents of Wiggins have criticized his distinction between constitution and identity, which allows the possibility of two things in the same place at the same time (Lewis 1986). Others have questioned his positive argument that Leibniz's Law and relative identity are incompatible. Debate about these matters continues.

One sortal concept to which Wiggins has given special attention is that of a person. In *Identity and Spatio-temporal Continuity* (1967) and its successors, he developed his response to the problem of personal identity originating in the writings of John Locke, with particular reference to the writings of Bernard Williams (1973), Derek Parfit (1984), and Sydney Shoemaker (1963). In response to the famous Reduplication Argument against Lockean accounts of personal identity in terms of consciousness, put forward by Williams, Wiggins insists that the concept of a person, as a genuine sortal concept, must satisfy "the *a* and *b* rule," that whether later *a* is identical with earlier *b* can depend only on facts about *a* and *b* and the relations between them. This entails a

rejection of the modified Lockean "best candidate" type of account of personal identity developed by Shoemaker and endorsed by Parfit.

Wiggins also rejects Parfit's thesis that identity is not what matters in survival. Finally, he rejects Locke's distinction between man and person, and endorses the thesis that persons just are animals (more specifically, human beings). Many philosophers have accepted the distinction on the basis of thought experiments in which, for example, brains are transplanted from one skull into another, with consequent transference of memory and character traits. Wiggins suggests that in such cases the same human being (not merely the same person) has different bodies successively. More fundamentally, he denies the real possibility of such cases. In the last position, he is influenced by the work of Kripke and Putnam. In this area too, Wiggins's position remains one of the options subject to current debate and development. The "animalist" position is developed in different ways by van Inwagen (1990) and Olson (1997), and is opposed by Shoemaker (1963).

See also Aristotle; Identity; Kripke, Saul; Leibniz, Gottfried Wilhelm; Lewis, David; Locke, John; Meaning; Parfit, Derek; Personal Identity; Putnam, Hilary; Quine, Willard Van Orman; Shoemaker, Sydney; Williams, Bernard.

Bibliography

Geach, Peter. *Logic Matters*. Oxford, U.K.: Basil Blackwell, 1972.

Kripke, Saul. *Naming and Necessity*. Oxford, U.K.: Basil Blackwell, 1980.

Lewis, David. *On the Plurality of Worlds*. Oxford, U.K.: Basil Blackwell, 1986.

Olson, Eric. *The Human Animal*. Oxford, U.K.: Oxford University Press, 1997.

Parfit, Derek. *Reasons and Persons*. Oxford, U.K.: Clarendon Press, 1984.

Putnam, Hilary. "The Meaning of 'Meaning.' " In his *Mind, Language, and Reality*. Vol. 2 of *Philosophical Papers*. Cambridge, U.K.: Cambridge University Press, 1975.

Quine, Willard Van Orman. *Word and Object*. Cambridge, MA: MIT Press, 1960.

Shoemaker, Sydney. *Self-Knowledge and Self-Identity*. Ithaca, NY: Cornell University Press, 1963.

Van Inwagen, Peter. *Material Beings*. Ithaca, NY: Cornell University Press, 1990.

Wiggins, David. *Identity and Spatio-temporal Continuity*. Oxford, U.K.: Basil Blackwell, 1967.

Wiggins, David. *Needs, Values, and Truth: Essays in the Philosophy of Value*. 2nd ed. Oxford: Blackwell, 1991.

Wiggins, David. *Sameness and Substance*. Oxford, U.K.: Basil Blackwell, 1980.

Wiggins, David. *Sameness and Substance Renewed*. Cambridge, U.K.: Cambridge University Press, 2000.

Williams, Bernard. *Problems of the Self*. Cambridge, U.K.: Cambridge University Press, 1973.

Harold W. Noonan (2005)

WILDE, OSCAR FINGAL O'FLAHERTIE WILLS

(1854–1900)

Born in Dublin to artistically minded parents, Wilde studied for three years at Trinity College in Dublin, and then at Magdalen College in Oxford, where his tutors included the English art critic John Ruskin and the English essayist Walter Pater. At the age of twenty-four he moved to London, where he very quickly became a conspicuous figure on the social scene, celebrated for his wit, personality, and self-consciously foppish dress sense. He married in 1884, had two children, and then, within a couple of years, noticed that he was homosexual. He fell in love with Lord Alfred Douglas in the early 1890s, was repeatedly and publicly denounced by Douglas's father, the Marquess of Queensberry, until eventually Wilde sued for libel, and lost. This led to his trial and conviction for sodomy, and to a sentence of two years' hard labor, which he served first in Wandsworth prison and then in Reading gaol. He was released in 1897, and spent the remaining years of his life as a social outcast in France, cash-strapped and increasingly ill. When he died, he was just forty-six.

Although Wilde is chiefly remembered for his one-liners—not unreasonably, given how good so many of them are—he was a more versatile writer than this fact might suggest. He published prose fiction, including a collection of fairy stories, *The Happy Prince and Other Tales* (1888), and a novel, *The Picture of Dorian Gray* (1891); he published verse, most notably "The Ballad of Reading Gaol" (1898); he dabbled in social commentary of a utopian bent, as seen in "The Soul of Man under Socialism" (1891); and he was a highly successful dramatist, with the best of his plays, *The Importance of Being Earnest* (1894), still being performed regularly in the twenty-first century. He also wrote essays and dialogues on art and art criticism, the most important of which, "The Decay of Lying" and "The Critic as Artist," were among the pieces that he published in 1891, under the title *Intentions*.

Wilde was not a philosopher, and it is an interesting question whether, or to what extent, he can be taken to

have contributed to philosophy. His most obvious connection to the subject, after all, is the rather unusual one of being, not the originator of a philosophical position, but the emblem or embodiment of one: Wilde stands for aestheticism in much the way that Lord Byron, for instance, stands for Romanticism. And this is a role that Wilde cultivated assiduously.

The term "aestheticism" refers to a cluster of more or less closely related views (often glossed as "art for art's sake"), rather than to a single theory or system; and many of these views enjoyed wide currency in the second half of the nineteenth century, not least through the writings of Ruskin and Pater. Perhaps the most characteristic tenet of aestheticism is the claim that aesthetic value is independent of and/or superior to other kinds of value. From this standpoint, the preeminently Victorian habit of bringing moral values to bear on the assessment of art—of asking, if not first then certainly foremost, whether such-and-such a work is edifying, say, or is likely to deprave—was point missing and philistine. Instead, the aestheticists insisted, the question should be whether a given work is beautiful. As Wilde put it in the preface to *Dorian Gray*: "There is no such thing as a moral or an immoral book. Books are well written, or badly written. That is all." And this statement means that the artist's task cannot be didactic: "An ethical sympathy in an artist is an unpardonable mannerism of style" (Wilde 1949, p. 5).

Taken in its stronger form—that aesthetic value is both independent of and superior to other kinds of value—the aestheticist tenet prompts a view not merely about art, but also about life. It encourages the thought that one should try to turn oneself into a work of art, to understand oneself in aesthetic terms rather than moral ones, say, and this is a project to which Wilde devoted considerable effort, claiming (to André Gide) that he had put his genius into his life, and only his talent into his work. His dress sense, his manner, and above all his style, were carefully calculated for aesthetic effect: "To me," as a character in *Dorian Gray* says, "Beauty is the wonder of wonders. It is only shallow people who do not judge by appearances. The true mystery of the world is the visible, not the invisible …" (Wilde 1949, p. 29). And so successful was Wilde in cultivating his public persona that when Gilbert and Sullivan's operetta *Patience* was first performed in 1881, no one doubted after whom the dandified aesthete, Reginald Bunthorne, had been modeled.

"A critic cannot be fair in the ordinary sense of the word," Wilde wrote. "It is only about things that do not interest one that one can give a really unbiassed opinion, which is no doubt the reason why an unbiassed opinion is always absolutely valueless" (Wilde 1907, p. 153). This thought—an outright rejection of the value of disinterestedness in the experience of art—perhaps has a claim to be regarded as Wilde's most original contribution to the philosophy of art, shades of Ruskin notwithstanding. Wilde insisted that "it is only by intensifying his own personality that the critic can interpret the personality and work of others," and even went so far as to accord a higher value to the critic's work than to the artist's (Wilde 1907, p. 127), a relative estimation, incidentally, that proved to be prophetic of much that passed for literary studies in the later twentieth century. Wilde developed some of these thoughts, and they are interesting. But it is hard not to feel that they are, in the end, really only a side product of the much more pressing business of turning his life into art, of striking a stylish pose that should, above all, be effective, even if, as he himself averred, "All art is quite useless" (Wilde 1949, p. 6).

See also Aesthetic Qualities; Aesthetics, History of; Art, Value in; Beauty; Humor; Pater, Walter Horatio; Romanticism; Ruskin, John; Value and Valuation.

Bibliography

WORKS BY OSCAR WILDE

The Artist as Critic: Critical Writings of Oscar Wilde, edited by Richard Ellmann. Englewood Cliffs, NJ: Prentice Hall, 1969.

Intentions. London: The English Library, 1907.

The Annotated Oscar Wilde: Poems, Fiction, Plays, Lectures, Essays, and Letters, edited by H. Montgomery Hyde. New York: Potter, 1982.

The Complete Letters of Oscar Wilde, edited by Merlin Holland and Rupert Hart-Davis. New York: Holt, 2000.

The Picture of Dorian Gray. London: Penguin Books, 1949.

The Picture of Dorian Gray: Authoritative Texts, Backgrounds, Reviews and Reactions, Criticism, edited by Donald L. Lawler. New York: Norton, 1988.

WORKS ABOUT OSCAR WILDE

Bloom, Harold, ed. *Oscar Wilde*. New York: Chelsea, 1985.

Brown, Julia Prewitt. *Cosmopolitan Criticism: Oscar Wilde's Philosophy of Art*. Charlottesville: University of Virginia Press, 1997.

Chai, Leon. *Aestheticism: The Religion of Art in Post-Romantic Literature*. New York: Columbia University Press, 1990.

Cohen, Philip K. *The Moral Vision of Oscar Wilde*. Rutherford, NJ: Fairleigh Dickinson University Press, 1978.

Ellmann, Richard. *Oscar Wilde*. New York: Knopf, 1987.

Eltis, Sos. *Revising Wilde: Society and Subversion in the Plays of Oscar Wilde*. Oxford: Clarendon, 1997.

Freedman, Jonathan, ed. *Oscar Wilde: A Collection of Critical Essays*. Upper Saddle River, NJ: Prentice Hall, 1996.
McCormack, Jerusha, ed. *Wilde the Irishman*. New Haven, CT: Yale University Press, 1998.
Price, Jody. *A Map with Utopia: Oscar Wilde's Theory for Social Transformation*. New York: Peter Lang, 1996.
Raby, Peter, ed. *The Cambridge Companion to Oscar Wilde*. Cambridge, U.K.: Cambridge University Press, 1997.
Schmidgall, Gary. *The Stranger Wilde: Interpreting Oscar*. New York: William Abrahams, 1994.
Shewan, Rodney. *Oscar Wilde: Art and Egoism*. New York: Barnes, 1977.

Aaron Ridley (2005)

WILL

See *Determinism, A Historical Survey*; *Volition*

WILLIAM HEYTESBURY

See *Heytesbury, William*

WILLIAM OF AUVERGNE

(c. 1180–1249)

William of Auvergne (or Paris) was born in Aurillac in the province of Auvergne. He was a master of theology at Paris by 1223 and was consecrated bishop of Paris in 1228. His chief philosophical works are *De trinitate, seu De primo principio* (c. 1223; translated as *The Trinity, or The First Principle*), which presents his metaphysics; *De universo* (c. 1231–1240; translated as *The Universe of Creatures*); and *De anima* (c. 1240; translated as *The Soul*); all parts of his seven-part *Magisterium divinale et sapientiale*. These works were written in a literary and highly personal style influenced by Latin translations of Avicenna.

Reacting to the teaching of many then newly circulating translations of Greek and Arabic texts of metaphysics and natural philosophy, and writing under early-thirteenth-century prohibitions at Paris, William attempted to identify and refute the errors of these works. But he was also greatly influenced by their teachings when they accorded with Christian faith, and incorporated them into an outlook influenced by St. Augustine.

Especially influenced by Avicenna, William was the first Latin thinker to base his metaphysics on Avicenna's distinction between being and essence. According to William, everything that exists is a possible being, whose essence is distinct from its being, or a necessary being, whose essence and being are identical. There must be a single necessary being, God or the first being, from whom existing possible beings receive their being. William described existing possible beings as composed of being and essence, raising the question of whether he, like Aquinas, posited a real distinction of being and essence in creatures. From Boethius, William took a related distinction between being (*esse*) and what a thing is (*quod est*). Identifying what a thing is with its essence, he distinguished beings by participation, whose essence is distinct from their being, from beings by essence, whose being and essence are identical. Beings by participation, he argued, must partake of their being from a unique being by essence, God.

Despite care to avoid the errors of non-Christian thinkers, William himself sometimes treads on dangerous ground. At one point he describes God as the being of everything, suggesting pantheism. At other times, emphasizing God's power in opposition to the necessitarian tendencies of Arabic thought, he writes as though creatures are not genuine causal agents but merely conduits of God's causal power. Such statements, however, probably do not reflect his considered views.

A key error that William identified in Avicenna (misidentified as Aristotle) was his doctrine of creation. According to this doctrine, God does not, as Christians think, create all things freely and contingently from nothing, but necessarily emanates a single intelligence or spiritual being. From this being necessarily emanate in turn further intelligences and the heavenly spheres, a process ending with the emanation of human souls and things of the sublunary world from the tenth intelligence. William took this doctrine to result from an incorrect application of the principle that from what is one, insofar as it is one, comes only one. Drawing on the doctrine of the divine will of the Jewish thinker Avicebron (1021–1058), William argued instead that God created the world not insofar as he is one, but insofar as he is free.

William also attacked Avicenna's and Aristotle's non-Christian doctrine that the world exists without beginning. The first Latin thinker to treat the issue in depth, he refuted a battery of arguments for an eternal world and presented lengthy arguments for its beginning. Several of these arguments, some drawn from the sixth-century Alexandrian thinker John Philoponus, allege that a world without beginning involves paradoxes of infinity, and would be popular with later Franciscan thinkers, including Bonaventure.

William's *The Soul* is the most substantial early-thirteenth-century treatment of the soul. Despite using Aristotle's definition of the soul as the perfection of an organic body potentially having life, William in fact adopted a non-Aristotelian conception of the soul as an incorporeal, indivisible, simple substance, identifying it with the whole human being and treating the body as its prison or cloak. To show the distinctness of soul and body, William used Avicenna's "floating man' argument that someone floating in the air without use of the senses would know the existence of his soul, but not of his body. William rejected a plurality of distinct souls in a human being corresponding to the vegetative, sensitive, and rational vital functions, attributing these functions to a single rational soul. Perhaps the first Latin thinker to hold that souls and angels are wholly immaterial without any kind of matter, William argued at length that the soul survives destruction of the body and is immortal.

In epistemology William was concerned to attack the doctrines of an agent intelligence and an agent intellect. The former doctrine, found in Avicenna, posits that intelligible forms are impressed on the human intellect by the tenth intelligence. William objected that this is incompatible with our need to study to acquire knowledge. The doctrine of an agent intellect, according to William, posits within the human soul two intellects, a receptive or material intellect and an active or agent intellect, which impresses intelligible forms on the material intellect. Noting the popularity of this doctrine in his day, William objected that it is incompatible with the simplicity of the soul and would mean that we know everything that can naturally be known. His positive account of knowledge is unclear, however, being expressed in imprecise and metaphorical terms. It has been suggested that he treated God as an agent intellect. But in fact he held only that God impresses on the human intellect the principles of truth and morality; once these principles are known, the whole soul can acquire scientific knowledge directly without the mediation of any agent intellect within or outside it.

Early to advocate a voluntarist conception of free will, William held that the will is king and noblest power in the soul, with command over its other powers, and is counseled by the intellect. The will itself must be capable of apprehension and cognition if it is not to be blind, and the intellect likewise has a kind of appetite. The will cannot be forced, prevented, or necessitated. William wrote that he was puzzled that Aristotle had not considered the will.

Eminent in his day, William influenced Aquinas's metaphysics of being and essence and Franciscan thinkers' arguments for the beginning of the world. His works survive in many manuscripts, suggesting an influence whose full extent remains to be studied.

See also Agent Intellect; Avicenna; Thomas Aquinas, St.

Bibliography

WORKS BY WILLIAM OF AUVERGNE

De immortalitate animae. In *Des Dominicus Gundissalinus Schrift von der Unsterblichkeit der Seele*, edited by Georg Bülow. Münster, Germany: Aschendorff, 1897.

"Tractatus Magistri Guillelmi Alvernensis *De bono et malo*," edited by J. Reginald O'Donnell. *Mediaeval Studies* 8 (1946): 245–299.

"Tractatus secundus Guillelmi Alvernensis *De bono et malo*," edited by J. Reginald O'Donnell. *Mediaeval Studies* 16 (1954): 219–271.

Opera omnia (1674). 2 vols. Frankfurt am Main: Minerva, 1963.

Il "Tractatus de gratia" di Guglielmo d'Auvergne, edited by Guglielmo Corti. Rome: Lateran University, 1966.

De trinitate, edited by Bruno Switalski. Toronto: Pontifical Institute of Mediaeval Studies, 1976.

The Trinity, or The First Principle. Translated by Roland J. Teske and Francis C. Wade. Milwaukee, WI: Marquette University Press, 1989.

The Immortality of the Soul. Translated by Roland J. Teske. Milwaukee, WI: Marquette University Press, 1991.

The Universe of Creatures. Translated by Roland J. Teske. Milwaukee, WI: Marquette University Press, 1998. A partial translation.

The Soul. Translated by Roland J. Teske. Milwaukee, WI: Marquette University Press, 2000. Contains a full bibliography.

WORKS ON WILLIAM OF AUVERGNE

Caster, Kevin J. "The Distinction between Being and Essence according to Boethius, Avicenna, and William of Auvergne." *Modern Schoolman* 73 (1996): 309–332.

Gilson, Etienne. "La notion d'existence chez Guillaume d'Auvergne." *Archives d'histoire doctrinale et littéraire du moyen âge* 21 (1946): 55–91.

Gilson, Etienne. "Pourquoi saint Thomas a critiqué saint Augustin." *Archives d'histoire doctrinale et littéraire du moyen âge* 1 (1926): 5–127.

Kramp, Josef. "Des Wilhelm von Auvergne 'Magisterium Divinale.' " *Gregorianum* 1 (1920): 538–613; 2 (1921): 42–103, 174–195.

Marrone, Steven P. *William of Auvergne and Robert Grosseteste: New Ideas of Truth in the Early Thirteenth Century*. Princeton, NJ: Princeton University Press, 1983.

Masnovo, Amato. *Da Guglielmo d'Auvergne a S. Tommaso d'Aquino*. 3 vols. 2nd ed. Milan, Italy: Vita et Pensiero, 1946.

Moody, Ernest A. "William of Auvergne and His Treatise *De anima*." In his *Studies in Medieval Philosophy, Science, and Logic*, 1–109. Berkeley: University of California Press, 1975.

Teske, Roland J. "William of Auvergne on the 'Newness' of the Word." *Mediaevalia: Textos e Estudios* 7–8 (1995): 287–302.

Teske, Roland J. "William of Auvergne's Rejection of the Agent Intellect." In *Greek and Medieval Studies in Honor of Leo Sweeney, S.J.*, edited by William J. Carroll and John J. Furlong, 211–235. New York: Peter Lang, 1995.

Valois, Noël. *Guillaume d'Auvergne, évêque de Paris (1228–1249): Sa vie et ses ouvrages*. Paris: Picard, 1880.

Neil Lewis (2005)

WILLIAM OF CHAMPEAUX

(c. 1070–1121)

William of Champeaux, born at Champeaux near Melun, was perhaps a student of Anselm of Laon. William was held in high esteem by his contemporaries for his mastery of grammar, logic, and rhetoric, as well as for theological speculations. By 1100 he was the Master of the Cathedral School associated with Notre Dame in Paris, the most prestigious position available for a philosopher; he held the rank of archdeacon, and was a confidante of Philip I. In this period Peter Abelard was first William's student and then his rival in public debates over philosophy. In 1108, William entered the Abbey of St. Victor newly established outside the walls on Paris, on the south bank of the Seine, and apparently continued to teach while there. In 1114 William was made bishop of Châlons-sur-Marne, a position he held until his death there in 1121.

William's theological views are presented in a compendium of short discussions, each addressed to a particular question: his *Sententiae* (partially printed by Lefèvre). His views on logic, language, metaphysics, and rhetoric are preserved in many manuscripts and by later authors, most notably by Abelard; little of this material has yet been edited or sorted out, and there is no scholarly consensus about which views can reliably be attributed to William, although it seems clear that William lectured and perhaps wrote extensively on the liberal arts. Abelard mentions in passing William's claim that every sentence has both a grammatical and a logical sense (*Logica ingredientibus* 7, Glosses on the "Topics" 271–273); that present-tense sentences about nonexistents should be interpreted figuratively (*Dialectica* 135–136); and that differentiae are only accidentally related to the genera they differentiate (*Dialectica* 541). But the best-known and most widely attested philosophical views of William of Champeaux have to do with the problem of universals.

According to Abelard, William initially held a position known as "material essence realism": One and the same material essence is found in distinct individuals of the same species, which are distinguished from one another by the addition of further forms to the material essence. When challenged by Abelard, William modified his position to hold that the same thing (the material essence) is not literally present in different things; distinct things are called the same "indifferently." This latter position seems to be endorsed in William's discussion of the Trinity in his *Sententiae*. Abelard presents William's positions briefly in his *Historia calamitatum* 65–66, and William's positions along with his criticisms at length in his *Logica ingredientibus* 1, Glosses on the "Isagoge" 11–17 and *Logica nostrorum petitioni sociorum* 512–517. William's replies are not known independently.

See also Abelard, Peter; Propositions; Saint Victor, School of; Universals, A Historical Survey.

Bibliography

SELECTED WORKS

Fredborg, Karin M. "The Commentaries on Cicero's *De inventione and Rhetorica ad Herennium* by William of Champeaux." *Cahiers de l'institut du moyen-âge grec et latin* 17 (1976): 1–39.

Iwakuma, Yukio. "The *Introductiones dialecticae secundum Wilgelmum and secundum magistrum G. Paganellum*." *Cahiers de l'institut du moyen-âge grec et latin* 63 (1993): 45–114.

Iwakuma, Yukio. "William of Champeaux and the *Introductiones*." In *Acts of the 10th European Symposium on Medieval Logic and Semantics*, edited by H. A. G. Braakhuis and C. H. Kneepkens. Leiden: Brill 1996.

Lefèvre, Georges. *Les variations de Guilliame de Champeaux et la question des Universaux*. Lille, France: Université de Lille 1898. Includes William's *Sententiae*.

Marenbon, John. *The Philosophy of Peter Abelard*. Cambridge, U.K.: Cambridge University Press, 1997. Discusses reports of William's teaching in unedited manuscripts.

Mews, Constant. "Logica in the Service of Philosophy: William of Champeaux and His Influence on the Study of Language and Theology in the Twelfth Century." In *Studien zur Abtei Sankt Viktor zu Paris und zu den Viktorinern*, edited by Rainer Berndt (Corpus victorinum. Instrumenta 1), Berlin 2004.

Rosier-Catach, Irène. "Abélard et les grammairiens: sur le verbe substantif et la prédication." *Vivarium* 41 (2003): 175–248.

Peter King (2005)

WILLIAM OF CONCHES

William of Conches, the twelfth-century Chartrain philosopher, was born at Conches in Normandy at the end of the eleventh century. He probably studied under Bernard of Chartres, learning at least grammar from him, and began teaching in the early 1120s. About 1140 William, who was perhaps now in Paris, had John of Salisbury as one of his pupils; John found him perpetuating the spirit of Bernard's own teaching. However, opposition from less lettered philosophers led William to return to his native Normandy under the protection of Duke Geoffrey Plantagenet, whose son, the future Henry II of England, he taught. He died sometime after 1154.

William left glosses on Priscian in both an early and a later version, and recent evidence suggests that he may have written glosses on Juvenal. However, his other surviving writings testify above all to a considerable achievement in philosophy and in scientific thought. They include a commentary on the *Consolation of Philosophy* by Boethius that is dependent on older glosses but is animated by an ampler philosophical and physical interest; glosses upon Macrobius; a first version of a commentary upon the Chalcidian version of Plato's *Timaeus*; and a systematic work, the *Philosophia Mundi,* which ranges widely over the topics of God, the universe, and man. William considers the nature of God and his relationship to creation; he also considers the structure and composition of the universe, the elements, the heavens, motion, and geography. Finally, he examines the biology and psychology of man.

These were all youthful writings, completed by the early 1120s. In a second version of his commentary on the *Timaeus,* William abandoned his former assimilation of the Platonic world soul with the Holy Spirit of Christian doctrine. In the later 1140s he continued to modify youthful theses and produced a masterpiece, the *Dragmaticon Philosophiae,* cast in the form of a dialogue with Duke Geoffrey. In this work, which built upon the earlier *Philosophia Mundi,* William developed his physical and astronomical interests and produced the most up-to-date scientific encyclopedia of the mid-twelfth century. Like the *Philosophia Mundi,* it was widely circulated. Some historians consider William to be the author of the *Moralium Dogma Philosophorum,* an influential collection of moralist citations from Scripture, the Church Fathers, and ancient pagan writers.

Much of William's philosophical effort was directed toward ensuring that Christian theology embraced the study of the universe and of man. He saw in Plato's *Timaeus* a doctrine of creation that helped to explain the account given in the book of Genesis. He identified the Platonic archetypal world with the wisdom of God, the Logos of Christian belief. He firmly underlined St. Paul's teaching on the intelligibility of this world (Romans 1:20). The created universe bears the imprint of its creator, and its harmony reveals the fundamental attributes of God—power, wisdom, and goodness. These aspects of God are commonly signified by the names of three divine persons, but William was preoccupied with the creative activity of the Trinity rather than with the intimate relationships of the divine life. Stressing the cosmological function of the Holy Spirit, William presented the third person of the Trinity as the principle of life that animates the world and, in his earlier writings, as identical with the *anima mundi,* or world soul, of Platonic doctrine. Conservative theological opinion was thereby antagonized.

After 1140 William of St.-Thierry, the Cistercian friend of Bernard of Clairvaux, launched an attack against the grammarian of Conches, as he had earlier against Peter Abelard. He criticized William for following Abelard and for transgressing the limits of theological inquiry set by the fathers of the church. He accused the Chartrain of Sabellianism and of subordinationism in his cosmological interpretation of the Trinity, and of materialism in making God an immanent regulatory principle of the universe. In the *Dragmaticon* William yielded somewhat to these criticisms, but he was also influenced by new translations of Greek and Arabic medical writings. His animistic vision of the universe was now tempered by an increased insistence on the power of secondary causes, of nature itself to sustain the universe in cooperation with God. William arrived at a new sense of the autonomous value of nature, and he offered many new perspectives. On the individual human soul and its faculties he joined the medical theories of the newly translated *Pantegni* of ʿAlī ibn al-ʾAbbas and of the *Isagoge* of Johannitius to the traditional Boethian doctrine. Stimulated by the *Pantegni* as well as by Vergil and Lucretius, he criticized the traditional theory of the four elements as the first principles of things. The Ptolemaic theory of planetary motion appeared in William's *Dragmaticon,* which became a striking witness to the broadening of the contemporary scientific horizon.

See also Abelard, Peter; Bernard of Chartres; Bernard of Clairvaux; Boethius, Anicius Manlius Severinus; Chartres, School of; Creation and Conservation, Religious Doctrine of; God, Concepts of; John of Salisbury; Lucretius; Medieval Philosophy; Plato.

Bibliography

Extracts from the *Commentary on Boethius' Consolation of Philosophy* are in J. M. Parent, *La doctrine de la création dans l'école de Chartres* (Paris: Vrin, 1938), pp. 122–136, and, edited by C. Jourdain, in *Notices et extraits des manuscrits de la Bibliothèque Impériale,* Vol. 20, Part II (Paris, 1862), pp. 40–82.

Glosses on Macrobius' Commentary on the Dream of Scipio is unedited, but see E. Jeauneau in *Archives d'histoire doctrinale et littéraire du moyen âge* 27 (1960): 17–28.

Extracts from *Glosses on Plato's Timaeus* have been edited by V. Cousin, in *Ouvrages inédits d'Abélard* (Paris, 1836), pp. 646–657; by J. M. Parent, op. cit., pp. 137–177; and by T. Schmid, in *Classica et Mediaevalia* 10 (1949): 220–266; there is also an edition of Schmid's version by E. Jeauneau, *Glosae super Platonem* (Paris: J. Vrin, 1965).

Extracts from *Glosses on Priscian,* edited by E. Jeauneau, are in *Recherches de théologie ancienne et médiévale* 27 (1960): 212–247.

Philosophia Mundi may be found in *Patrologia Latina,* edited by J. P. Migne, Vol. 172, Cols. 39–102 (under Honorius Augustodunensis), and in Vol. 90, Cols. 1127–1178 (under Bede).

Dragmaticon Philosophiae, edited by G. Gratarolus, was published under the title of *Dialogus de Substantiis Physicis* (Strasbourg, 1567).

See also *Moralium Dogma Philosophorum,* edited by J. Holmberg (Uppsala: Almqvist and Wiksells, 1929.)

A study of William is T. Gregory, *Anima mundi. La filosofia di Guglielmo di Conches e la scuola di Chartres* (Florence: Sansoni, 1955).

David Luscombe (1967)

WILLIAM OF MOERBEKE

(c. 1215–c. 1286)

William of Moerbeke, one of the most competent and influential translators of Greek philosophical texts in the Middle Ages, was born at Moerbeke, near Ghent. He spent a number of years at the papal court in various Italian cities and also lived for some time in Greece and Asia Minor. His translations of Aristotle and other Greek authors began to appear about 1260. At the court of Pope Urban IV (1261–1264) in Orvieto, he made the acquaintance of his fellow Dominican, Thomas Aquinas, then beginning his series of Aristotelian commentaries, who encouraged him in his project of translating Aristotle. For several years before his death William was archbishop of Corinth.

Despite the claims that have sometimes been made about him, William of Moerbeke was not the first to translate the bulk of the Aristotelian corpus directly from Greek into Latin. It is true that in the twelfth century Western scholars had necessarily depended on translations from the Arabic, made in Spain or Sicily, for their knowledge of Aristotle. In the thirteenth century, however, at least partly as a result of the Fourth Crusade, a wider dissemination of Greek scholarship and easier access to Greek manuscripts encouraged Western translators to work directly from Greek originals, and many new translations came into use in the first half of the century. Thus, William's translation of Aristotle's *Metaphysics,* for example, while it may have been the first complete version, was apparently the third Latin translation to be made from the original text. A translation from Greek into Latin (the so-called *Metaphysica Vetus*) was in use at Paris as early as 1210, some time before the appearance of the *Metaphysica Nova,* based on the Arabic version, and a second translation from the Greek (the *Translatio Media*) seems to have been used by Albert the Great as the basis of his commentary. Many other works of Aristotle were similarly available by the middle of the thirteenth century in translations from the Greek as well as from the Arabic. While the extent of his indebtedness to earlier translators has not yet been precisely determined, William is known to have used some of the existing translations from the Greek in his own work.

Considered in themselves, then, William of Moerbeke's translations of Aristotle must be reckoned a less than revolutionary contribution to Aristotelian studies in the medieval West. It is not even known with certainty how far Thomas Aquinas, the outstanding interpreter of Aristotle in the thirteenth century, made use of his colleague's work. Nevertheless, William's translations of Aristotle and of other Greek philosophers, taken as a whole, can be said to have inaugurated a new phase of Aristotelian scholarship in Latin Christendom.

To begin with, William's new translations and revised versions of Aristotle's works gave the West a much more accurate text of "the Philosopher" than it had hitherto possessed. As a translator he was unquestionably superior in most respects to his predecessors. His strict adherence to the letter of the original text has been stigmatized as slavish, but it made his translations an unrivaled instrument of exact philosophical scholarship in his day.

Furthermore, William's translations of various post-Aristotelian authors helped Western scholars to form a clearer picture of the history of Greek philosophy and of the distinctive traits of Aristotle's doctrine. The Arabic versions of Aristotle's works had reached the West in the company of Neoplatonizing commentaries and Neoplatonic writings falsely attributed to Aristotle. Thanks to William's translations of important commentaries by Alexander of Aphrodisias, Simplicius, Themistius, and

John Philoponus, and of the *Elementatio Theologica* and other works of the Neoplatonist Proclus, the figure of the historical Aristotle stood out much more clearly than before, and Western thinkers were enabled to distinguish more precisely between the Platonic and Aristotelian approaches to philosophy. William's translation of Proclus was especially important in this connection, showing as it did that the influential *Liber de Causis,* far from being a genuine work of Aristotle, was in fact derived from Proclus's *Elementatio Theologica.*

Through his translation of Proclus William also influenced the development of medieval Neoplatonism. The works that he translated gave a fresh stimulus to the Neoplatonic school formed by Ulrich of Strasbourg and other disciples of Albert the Great and through that school helped to shape the mystical doctrine of Meister Eckhart.

See also Albert the Great; Alexander of Aphrodisias; Aristotelianism; Aristotle; Eckhart, Meister; Liber de Causis; Medieval Philosophy; Neoplatonism; Philoponus, John; Proclus; Simplicius; Themistius; Thomas Aquinas, St.; Ulrich (Engelbert) of Strasbourg.

Bibliography

See Martin Grabmann, *Guglielmo di Moerbeke, O.P., il traduttore delle opere di Aristotele* (Rome: Gregorianum, 1946).

Eugene R. Fairweather (1967)

WILLIAM OF OCKHAM

(c. 1285–1349)

William of Ockham, the most influential philosopher of the fourteenth century, apparently was born sometime between 1280 and 1290 at the village of Ockham, in Surrey, near London. Entering the Franciscan order at an early age, he commenced his course of theological study at Oxford in 1309 or 1310, and completed the requirements for the degree of master of theology with the delivery of his lectures on Peter Lombard's *Book of Sentences* in 1318–1319, or, at the latest, 1319–1320. Although an old tradition indicated that he studied under John Duns Scotus, it seems unlikely that he did so, since Duns Scotus left Oxford at the beginning of the century and died in 1308. Ockham's writings show intimate familiarity with the teachings of Duns Scotus, but this is explained by the dominant position Duns Scotus had acquired at Oxford, particularly within the Franciscan order.

Ockham's lectures on the *Sentences* made a profound impression on the students of theology at Oxford, but his new way of treating philosophical and theological questions aroused strong opposition by many members of the theological faculty. Normally the completion of his lectures on the *Sentences,* which gave Ockham the status of a *baccalaureus formatus* or *inceptor,* would have been followed by award to him of a teaching chair in theology. The granting of his teaching license was prevented by the chancellor of the university, John Lutterell, who in 1323 went to the papal court at Avignon to present charges against Ockham of having upheld dangerous and heretical doctrines. Because Ockham's academic career was thus interrupted while he was an *inceptor* awaiting award of the teaching license, he came to be known as "the venerable inceptor"—a title later misconstrued as meaning "founder of nominalism" (*inceptor scholae nominalium*).

Ockham was summoned to Avignon in 1324 to answer the charges against him, and he remained there four years, awaiting the outcome. A commission of theologians appointed by Pope John XXII to examine Ockham's writings submitted two lists of suspect doctrines in 1326, but there is no evidence of any final action having been taken on the charges that, in any case, were relatively mild. Despite the lack of a teaching chair, Ockham was extremely active during these years in developing his theological and philosophical positions, writing treatises and commentaries on logic and physics, a variety of treatises on theological questions, and an important series of quodlibetal questions that, presumably, he debated orally at Oxford or at Avignon.

In 1327, while at Avignon, Ockham became involved in the dispute then raging over the question of apostolic poverty, in which the general of the Franciscan order, Michael of Cesena, took a position opposed by the pope. Asked to study the question, Ockham found that a previous pope, Nicholas III, had made a pronouncement that fully supported the position of Cesena and of the majority of the Franciscans. When this controversy reached a critical stage in 1328, and it became evident that John XXII was about to issue an official condemnation of the position held by the Franciscans, Cesena and Ockham, along with two other leaders of the Franciscan opposition, fled from Avignon and sought the protection of Emperor Louis of Bavaria, who had repudiated the authority of the Avignon papacy in connection with the issue of succession to the imperial crown. Immediately after their flight from Avignon, Ockham and his companions were excommunicated by the pope for their refusal to submit to his authority.

Under the emperor's protection Ockham took up residence in Munich and devoted his full energies to writing a series of treatises on the issue of papal power and civil sovereignty, in which he held that John XXII had forfeited his right to the papal office by reason of heresy. When John XXII died in 1334, Ockham continued his polemic against the succeeding Avignon popes until 1347, when Louis of Bavaria died and the antipapal position became a lost cause. There is evidence that Ockham at that time sought reconciliation with the papal authority and with the rest of his own order, but the outcome is unknown. It is believed that he died in 1349, a victim of the Black Plague that, in the middle of the fourteenth century, took the lives of most of the intellectual leaders of northern Europe and played a major part in bringing about the cultural decline that lasted for more than a century.

WRITINGS

Ockham's writings fall into two distinct groups associated with the two different periods of his career. All of the political and polemical treatises directed against the Avignon papacy were written during his residence in Munich, between 1333 and 1347. Of these treatises many are solely of historical interest; but the lengthy *Dialogus Inter Magistrum et Discipulum,* written between 1334 and 1338, the *Octo Quaestiones Super Potestate ac Dignitate Papali,* written in 1340, and the *Tractatus de Imperatorum et Pontificum Potestate,* composed around 1347, present Ockham's philosophy of church and state and convey his deep-rooted convictions concerning the religious mission of the church.

The nonpolitical writings that embody Ockham's distinctive contributions to philosophy and theology were probably all written while he was at Oxford and at Avignon, between 1317 and 1328. The earliest of these include the lectures on the *Sentences,* a lengthy exposition of Aristotle's *Physics* extant only in manuscript form, and literal commentaries on Porphyry's *Isagoge* and on Aristotle's *Categoriae, De Interpretatione,* and *De Sophisticis Elenchis*; the first three of the commentaries were published at Bologna in 1496 under the title *Expositio Aurea … Super Artem Veterem* (Golden Exposition … of the Ancient Art). Ockham's most important work on logic, completed before he left Avignon, was a systematic treatise titled *Summa Logicae,* extant in several printed editions. An incomplete *Summulae in Libros Physicorum* (also given the title *Philosophia Naturalis*) contains an independent treatment of the subjects dealt with in the first four books of Aristotle's *Physics,* and was printed in several editions, beginning in 1495. In manuscript form only there is a work titled *Quaestiones Super Libros Physicorum,* which was probably one of his later writings; it covers, in the form of disputed questions, most of the topics treated in his earlier literal commentary on the *Physics* but reflects some changes in his views that occurred after the earlier work had been written. Two short compendia of logic, each extant only in a single manuscript version, are believed to be authentic works of Ockham, but they add nothing significant to the doctrines of his *Summa Logicae.*

Of Ockham's theological writings the lectures on the first book of the *Sentences,* known as the *Ordinatio* because Ockham revised and edited them for circulation, are of primary importance. Printed at Lyons in 1495, along with Ockham's lectures on the other three books of the *Sentences,* they are called the *Reportatio* because the text is derived from stenographic versions of the lectures as they were delivered. A modern critical edition of both parts of these lectures on the *Sentences* is very much needed. Of comparable importance for the understanding of Ockham's philosophical and theological doctrines are the quodlibetal questions, printed at Paris in 1487 and again at Strasbourg in 1491 under the title *Quodlibeta Septem.* Three other certainly authentic theological treatises, composed during the Oxford-Avignon period, are the *Tractatus de Corpore Christi* and *Tractatus de Sacramento Altaris,* which have been regularly printed together under the second of these titles, and the *Tractatus de Praedestinatione et de Praescientia Dei et de Futuris Contingentibus,* of which a modern edition, edited by Philotheus Boehner, was published in 1945. The 1495 Lyons edition of Ockham's theological works includes *Centiloquium Theologicum,* whose authenticity has been questioned by many scholars but without decisive evidence. In describing the philosophical doctrines of Ockham, use will be made chiefly of the *Commentary on the Sentences,* the *Summa Logicae,* and the *Quodlibeta Septem.*

CHARACTER OF OCKHAM'S PHILOSOPHY

Ockham's major contributions to the development of late medieval and early modern philosophy were in the areas of epistemology, logic, and metaphysics. His approach to these problems and his concern with them were those of a scholastic theologian, as had been the case with Thomas Aquinas, Duns Scotus, and other leading scholastic thinkers of the thirteenth century.

The basic problem of scholastic theology since the beginning of the thirteenth century had been that of finding a means of accommodating the philosophical system of Aristotle within the dogmatic framework of Christian doctrine. To achieve such an accommodation was a philosophical task because no alteration in the articles of the faith could be allowed, and consequently all elimination of contradictions had to be achieved by internal criticism or reinterpretation of the philosophical assumptions and arguments of Aristotle. Aquinas had sought to achieve an essentially external accord between natural philosophy and Christian theology, such as would leave the Aristotelian system internally intact. The Franciscan theologians, from St. Bonaventure to Duns Scotus, had considered this inadequate and had sought to achieve the required integration of philosophy and theology by exploiting the more Platonic elements of the Aristotelian system, much as the Greek Neoplatonists and the Muslim philosopher Avicenna had done. All of the thirteenth-century syntheses of philosophy and theology involved, in one form or another, the metaphysical and epistemological doctrine of realism—the doctrine that the human intellect discovers in the particulars apprehended by sense experience an intelligible order of abstract essences and necessary relations ontologically prior to particular things and contingent events and that from this order the intellect can demonstrate necessary truths concerning first causes and the being and attributes of God.

EMPIRICISM AND NOMINALISM. Ockham's significance, both as a theologian and as a philosopher, lay in his rejection of the metaphysical and epistemological assumptions of medieval realism, and in his reconstruction of the whole fabric of philosophy on the basis of a radical empiricism in which the evidential base of all knowledge is direct experience of individual things and particular events. The counterpart of this epistemological empiricism was the nominalistic analysis of the semantical structure and ontological commitment of cognitive language that Ockham developed in his logical writings. Ockham's empiricism was not phenomenalistic or subjectivistic, and it could be called a realistic empiricism according to a modern usage of "realism"; it presupposed and was based on the principle that the human mind can directly apprehend existent individuals and their sensible qualities, and that it can also directly apprehend its own acts. Insofar as Ockham is called a nominalist, his doctrine is not to be construed as a rejection of any ontological determination of meaning and truth, but rather as an extreme economy of ontological commitment in which abstract or intensional extralinguistic entities are systematically eliminated by a logical analysis of language.

OCKHAM'S RAZOR. The principle of parsimony, whose frequent use by Ockham gained it the name of "Ockham's razor," was employed as a methodological principle of economy in explanation. He invoked it most frequently under such forms as "Plurality is not to be assumed without necessity" and "What can be done with fewer [assumptions] is done in vain with more"; he seems not to have used the formulation "Entities are not to be multiplied without necessity." The principal use made by Ockham of the principle of parsimony was in the elimination of pseudo-explanatory entities, according to a criterion he expresses in the statement that nothing is to be assumed as necessary, in accounting for any fact, unless it is established by evident experience or evident reasoning, or is required by the articles of faith.

POSITIVE THEOLOGY. As applied by Ockham, the principle of parsimony resulted in an empiricist criterion of evidence that left little room for a natural theology. But since it also reduced physics and cosmology to the status of positive sciences without metaphysical necessity, it left room for a positive theology based on revelation and faith that could no more be refuted than it could be demonstrated by any necessary reasons or observational evidence. Moreover, this positive theology, in which God is conceived as the omnipotent creator of all finite things whose creative and causal action is wholly free and unnecessitated, provided an indirect justification of Ockham's philosophical empiricism, since it demanded a conception of the world of created things as radically contingent in both their existence and their interaction. Ockham made full use of the doctrine of divine omnipotence as an ad hominem argument against those who sought to discredit his philosophical doctrine on theological grounds; philosophically, however, the doctrine was equivalent to the principle that whatever is not self-contradictory is possible, and that what is actual, within the range of the logically possible, cannot be established by reason alone but only by experience.

CRITIQUE OF REALISM

Ockham's epistemology and metaphysics were designed to resolve a basic problem that the Scholastics had inherited from the Greek philosophical tradition and that may be summed up in the paradoxical thesis that the objects of thought are universal, whereas everything that exists is singular and individual. Seeking to overcome this gap between the intelligible and the existent, the earlier

Scholastics had elaborated various forms of the doctrine called moderate realism, according to which there are common natures in individual existing things, distinct from their individuating principles although not separable except in thought. On the psychological side, these doctrines held that the human intellect abstracts, from the particular presentations of sense experience, an intelligible species, or likeness, by means of which it apprehends the common nature apart from the individuating conditions. The varieties of this moderate realism turned on the answer to the question of whether, in an individual, the common nature is (1) really distinct from the individuating principle or (2) "formally distinct," as Duns Scotus proposed or (3) distinct only according to the mode of consideration although involving some "foundation in the thing" for such distinguishability, as Aquinas held.

Ockham considered all forms of this doctrine of common natures in individual things to be self-contradictory and irrational. If the human nature of Socrates is really distinct from Socrates, then it is not Socrates' nature or essence, for a thing cannot be said to be essentially something that it really is not. If the common nature is anything at all, it is either one thing or many things; if one and not many, it is not common but singular, and if not one but many, then each of the many is singular and there is still nothing common.

CRITICISM OF THE SCOTIST VIEW. The answer of Duns Scotus—that the common nature is really identical with, but formally distinct from, the *haecceitas* or individuating differentia that was said to contract the specific nature to singularity—was an attempt to find something intermediate between identity and nonidentity. Ockham argued, against the Scotist thesis, that if the specific nature and the individuating difference are really identical, they cannot be formally distinct; and if they are formally distinct, they cannot be really identical. Duns Scotus had claimed that they are both really identical and formally distinct. Let *a* and *b* represent the individual difference and the specific nature, respectively. Then, since *a* is not formally distinct from *a,* it follows that if *a* is identical with *b,* then *b* is not formally distinct from *a.* Similarly, since *a* is not formally distinct from *a,* then if *b* is formally distinct from *a, b* is not identical with *a.* In these arguments Ockham employs, with great effectiveness, the principle commonly ascribed to Gottfried Wilhelm Leibniz—that if two things are identical, whatever is true of one is true of the other; and if something is true of one that is not true of the other, they are not identical.

CRITICISM OF THE THOMIST VIEW. The third answer—that the same thing is singular and universal according to different ways of considering it—is ridiculed by Ockham on the ground that what a thing is in itself can in no way depend on how someone thinks of it. "For with the same ease I could say that a man considered in one way is an ass, considered in another way he is an ox, and considered in a third way he is a she-goat" (*Expositio Super VIII Libros Physicorum,* in *Ockham: Philosophical Writings,* edited by Philotheus Boehner, p. 14). Nor can it be said, as Aquinas appears to say in his *De Ente et Essentia,* that the nature or essence of a thing is in itself neither individual nor universal but is made singular by being received in individuating matter and is made universal by being received into the mind. Anything whatsoever, Ockham insists, is one thing and a singular thing by the very fact that it is a thing, and it is impossible that its unity or singularity is due to something added to it.

OCKHAM'S POSITION. It remains, then, that universality and community are properties only of signs—of language expressions and of the acts of thought expressed by them. The problem of universals therefore is not a metaphysical problem of explaining how abstract common natures are individuated to singular existence, nor is it a psychological problem of explaining how the intellect can abstract from the images of sense experience a common nature inherent in the individuals experienced; for there are no common natures to be individuated or to be abstracted. The problem of individuation is a logical problem of showing how general terms are used in propositions to refer to individuals signified by them; this problem is resolved in terms of the quantifying prefixes and other syncategorematic determinants of the referential use of terms in propositions. As an epistemological problem, the problem of universals is that of explaining how experience of individual existing things can give rise to concepts of universal character and to universally quantified propositions that hold for all objects signified by the subject term. The basis of Ockham's answer to these problems is given in his doctrine of intuitive and abstractive cognition.

INTUITIVE AND ABSTRACTIVE COGNITION

The doctrine of intuitive and abstractive cognition is formulated at the beginning of Ockham's *Commentary on the Sentences* in connection with the question of whether evident knowledge of theological truths can be acquired by man in this life. After distinguishing apprehension

from judgment as a distinct act of the intellect, and after showing that every act of judgment presupposes an act of apprehension of what is signified by the terms of the proposition expressing such a judgment, Ockham distinguishes two kinds of intellectual apprehension, intuitive cognition and abstractive cognition.

Intuitive cognition is defined as an act of apprehension in virtue of which the intellect can evidently judge that the apprehended object exists or does not exist, or that it has or does not have some particular quality or other contingent condition; in short, an intuitive cognition is an act of immediate awareness in virtue of which an evident judgment of contingent fact can be made.

Abstractive cognition is defined as any act of cognition in virtue of which it cannot be evidently known whether the apprehended object exists or does not exist, and in virtue of which an evident contingent judgment cannot be made. That these two ways of apprehending the same objects are possible is clear from experience; while I am observing Socrates sitting down, I can evidently judge that Socrates is seated, but if I leave the room and then form the judgment that Socrates is seated, it is not evident, and may indeed be false.

The important point in this distinction is that intuitive and abstractive cognition do not differ in the objects apprehended, but solely in the fact that intuitive cognition suffices for making an evident contingent judgment concerning the object apprehended, whereas an abstractive cognition does not. Nor is the distinction one between sensation and thought, for however much it may be true that affection of the senses by the external object is a necessary condition for an intuitive cognition of a sensible object, the intuitive cognition is an intellectual act that is presupposed by the act of judgment whose evidence is derived from it. Neither is the distinction one between direct awareness of the object and awareness of something representing the object in its absence; both kinds of apprehension are directly of the object. It is not even logically necessary that the object of an intuitive cognition be present or actually existent, although if, by the power of God, an intuitive cognition of an object were preserved after the object was removed or destroyed, it would then yield the evident judgment that the object was not present or that it did not exist; for it is self-contradictory, and hence not even within the power of God, for a cognition to yield an evident judgment that an object exists if the object does not exist.

INTUITIVE COGNITION OF NONEXISTENTS. Ockham must admit that an intuitive cognition of a nonexistent object is logically possible because an intuitive cognition, however much it may be caused by the presence of its object, is not identical with its object; hence it is not self-contradictory that it exists without the object's existing. And if we suppose that any effect that can be produced by a created cause can be produced by God without the created cause, this logical possibility could be realized by the power of God. In this way God could, and according to Christian belief did, produce intuitive cognitions of future things and events by which the prophets and saints had evident knowledge of what did not yet exist; and God himself, who apprehends all things intuitively and not abstractively, is aware not only of the things he has created but of all the things he does not choose to create. Thus, an intuitive cognition of a nonexistent object is logically possible, although it is realizable only by the power of God. Without such divine intervention, however, such cognitions can arise only if the object is present to the knower; and the judgments to which intuitive cognitions can give rise, in the natural course of events, are affirmative judgments of present existence and present fact.

INTUITIVE COGNITION OF MENTAL STATES. Ockham does not restrict the objects of intuitive cognition to objects perceptible to the external senses but includes nonsensible actualities that are apprehended introspectively, such as thoughts, volitions, and emotions. Thus the intellect, by reflecting on its own acts, can form evident judgments of the existence of those acts; for example, if I am intuitively aware of Socrates being seated, I can not only judge evidently that Socrates is seated, but I can also give evident assent to the second-order proposition "I evidently know that Socrates is seated." Although Ockham generally holds that the reflexive act is distinct from, and posterior to, the direct act, he speaks as if the evidence of the reflexive act can include that of the direct act.

DERIVATION OF ABSTRACTIVE COGNITIONS. Given an intuitive cognition of some object or event, the intellect thereby acquires an abstractive cognition of the same object or event, which it retains as a *habitus,* or acquired capacity, to conceive the object without any causal concurrence by the object itself; thus, objects that we have experienced intuitively can be apprehended abstractively, the only difference being that the abstractive cognition does not suffice to make evident a contingent judgment concerning the object thought of. If we leave out of account the logically possible case of God's producing an abstractive cognition without a preceding intuitive cogni-

tion, the principle holds, according to Ockham, that no abstractive cognition can be had that is not derived from an intuitive cognition of the object or objects conceived. This principle, which corresponds to David Hume's thesis that there is no idea which is not derived from one or more impressions, is basic to Ockham's theory of natural knowledge and its source of evidence.

UNIVERSALITY OF ABSTRACTIVE COGNITION. In his earlier formulation of the doctrine of intuitive and abstractive cognition, Ockham supposed that the abstractive cognition immediately derived from an intuitive cognition is a concept only of the singular object of the intuitive cognition. But in his *Quodlibeta* (Quod. I, q. 13) he states that a simple abstractive cognition cannot be a concept peculiar to one singular object to the exclusion of other objects that would, if apprehended intuitively, yield a wholly similar concept. Thus the universality of the concept, in this later theory, is immediately involved in the transition from intuitive to abstractive cognition. The operation is analogous to that of deriving, from a proposition of the form *Fa,* the open sentence *Fx,* which becomes a general proposition when the free variable *x* is bound by a quantifying prefix. In Ockham's terminology, the abstractive cognition has signification but acquires supposition only by formation of a judgment or proposition.

CONCEPTS. The concept, or universal in the mind, is a cognition of objects in virtue of which it cannot be evidently judged that they exist or do not exist. But what sort of reality is such a cognition or concept? One opinion is that the concept is a mental image or species which, because it is a resemblance of the external objects, causes the intellect to become aware of those objects. But Ockham points out, as Hume did later, that such a species could in no way represent to the intellect the objects of which it is a likeness, unless these objects were already known to it—no more, Ockham says, than a statue of Hercules could represent Hercules, or be recognized as his likeness, if the viewer had never seen Hercules.

In his *Commentary on the Sentences* Ockham mentions three theories of the concept as "probable" or tenable. According to the first theory, the concept is not a reality existing in the mind or outside the mind but is the being conceived of the external objects, the *esse obiectivum* of the objects—a view that was held by Peter Aureol and had adherents down to the time of René Descartes, who in the *Meditations* used this notion of the "objective being" of the concept in proving God's existence from his idea of God. Of the concept thus conceived, Ockham says that its being is its being understood—*eorum esse est eorum cognosci.* A second theory supposes that the concept is a real quality in the soul, used by the intellect for the individuals of which it is a concept, just as a general term in a proposition is used for the individuals of which it is a sign. A third theory, which Ockham finally adopted, is that the concept is merely the act of understanding the individual things of which it is said to be a concept. This theory is preferred on grounds of economy, for inasmuch as any of the theories requires that the intellect apprehend the extramental individuals, this function can be satisfied by the act of understanding without need of any other mental vehicle serving as surrogate for the objects.

Generality of concepts. The question may well be raised of how a concept derived from intuitive apprehension of a single object can constitute an act of understanding a definite set of objects—not any objects whatsoever but just those objects to which the concept is applicable or which, if directly experienced, would elicit that concept. Why should an intuitive cognition of Socrates yield a general concept applicable to just those individuals of which it is true to say "This is a man"? Ockham says that this is because the objects are similar, on which account the abstractive concept elicited by experience of one of the objects is ipso facto a concept of all similar objects. The realist might well insist that Ockham, in supposing this similarity in things, is covertly reintroducing the doctrine of common natures; but Ockham replies that similar individuals are similar by reason of what each individual is in itself, and not by reason of anything common. Two things are similar, for example, in being singular things, but this is not because there is one singularity common to the two things. Thus a concept can be a single act of understanding many individuals that are similar, without being an act of understanding anything other than just those individuals themselves. Again the analogy with the open sentence *Fx* is suggested, for if we should ask what things satisfy this function, the answer is that it is any of those things such that *Fx* holds for it. The obvious circularity of this question and answer indicates that any explanation that can be given of the fact that things are conceived in a universal manner by intelligent beings must itself use such universal concepts and thereby must presuppose the fact to be explained.

Concepts as natural signs. In this account Ockham describes concepts as natural signs whose relation to the things conceived is established not by human choice but by the fact that an act of understanding has no content other than the objects understood and arises in the first

instance only through direct experience of such objects. Ockham seems to recognize the futility of seeking to account for the possibility of knowledge as such by means of a particular branch of knowledge like physics or psychology; "*natura occulte operatur in universalibus* [nature works in a hidden manner in the case of universals]," he remarks, and is content to leave it at that.

LOGIC AND THEORY OF SCIENCE

Although the human intellect, according to Ockham, can directly apprehend and conceive the individual things that exist independently of our thought, the objects of knowledge (in the sense of *scire*) are propositions, formed within our minds by operations we freely perform by combining concepts derived from intuitive cognitions of things. Only propositions can be true or false, and since knowledge is of the true, its objects are propositions—complexes of signs put together by us. Logic is concerned with these ways of putting concepts together, insofar as these operations affect the truth or falsity of the resultant propositions.

Ockham was skilled in the formal logic developed in the arts faculties of the universities on foundations laid in the twelfth century by Peter Abelard, and represented in the thirteenth century by the treatises of the so-called terminist logicians William of Sherwood and Peter of Spain. The distinctive feature of this logic was its use of the concept of the supposition of terms in formulating the syntactical and semantical properties of cognitive language. In his *Summa Logicae* Ockham systematized the contributions of his predecessors in a reformulation of the whole content of Aristotelian logic on semantical foundations of a purely extensional character. These foundations, exhibited in his analysis of the signification of terms and of the truth conditions of propositions, reveal the ontological basis of his empiricist theory of knowledge and of scientific evidence. Some preliminary distinctions made at the beginning of Ockham's work on logic are important for understanding this analysis.

LOGIC AS A SCIENCE OF LANGUAGE. Logic, as a *scientia sermocinalis,* or science of language, deals with language as a system of signs that can be used in making true or false statements about things signified by those signs. The expressions of spoken and written language are instituted by convention to signify what is naturally signified (or intended) by acts of thought constituting the "inner discourse of the soul." Logic studies the properties of language expressions insofar as they embody the logically essential functions of mental discourse. Medieval logicians distinguished language signs into two basically different types: categorematic signs, which have independent meaning and can function as subjects and predicates of propositions, and syncategorematic signs, which have no independent meaning but exercise various logical functions with respect to the categorematic signs.

This important distinction corresponds to that made in modern logic between descriptive signs and logical signs. The categorematic signs, normally called terms, were divided into two distinct and nonoverlapping semantical types: terms of first intention, which signify things that are not language signs, and terms of second intention, which signify language signs or the concepts expressed by them, as signs. This distinction corresponds to that now made between the descriptive signs of the object language and the descriptive signs of the metalanguage. In Ockham's view, most of the metaphysical labyrinths in which the thirteenth-century Scholastics became entangled, such as the problem of universals in re, arose from the logical mistake of construing terms of second intention as terms of first intention; thus, because the term *man* is predicable of (or inheres in) the singular names "Socrates" and "Plato," they supposed that what is signified by the term *man* is some single reality that inheres in the individuals named by the names "Socrates" and "Plato."

SUPPOSITION. "Supposition" is defined by Ockham as the use of a categorematic term, in a proposition, for some thing or things—normally, for the thing or things it signifies. But terms can be used nonsignificatively as names of the concepts they express or as names of the spoken or written words of which they are instances. When used nonsignificatively as the name of the word, they were said to have material supposition; when used nonsignificatively as naming the concept expressed by the word, they were said to be used with simple supposition; but when used significatively for the things signified by them and understood by the concept or act of understanding expressed by them, they were said to be used in personal supposition. The earlier terminist logicians, who were metaphysical realists, had construed simple supposition as the use of a term for the universal nature that they supposed to exist in the individuals denoted by the term in its personal supposition—which is why they called this use simple (or absolute) supposition. But Ockham, who held that universality is a property only of concepts or language signs, rejected this interpretation and construed simple supposition as the use of a term for the concept or mental intention expressed by it.

The ontological foundations of Ockham's logic are exhibited in his analysis of the terms of first intention that Aristotle classified, in his *Categoriae,* as so many different ways of signifying "primary substances"—that is, concrete individuals. The terms Aristotle grouped under the category of substance, as signifying beings qua beings according to what they essentially are, were said by Ockham to be absolute terms, terms that signify nothing other than the individuals for which they can stand when used in propositions with personal supposition. The concrete terms of the so-called categories of accident, which are predicable of substance terms but signify them only as "of such quality," as "so big," or as "in such a place," were called by Ockham connotative terms—terms that refer obliquely to something other than the thing or things for which they can stand, and imply some contingent factual condition determining the range of objects for which the term can stand. The oblique reference may be to a part or parts of the object directly denotable by the term, to a quality of the object, or to some other thing or things with respect to which the denoted thing stands in some contingent relation—for instance, the term *father* stands for one thing by referring to another thing (a child) and implying that the child was generated by the person who is directly designated by the term *father.*

NOMINALISM. Ockham's nominalism consists in his refusal to construe abstract terms as names of entities distinct from the individual things signified by absolute terms. The realists, while conceding that the concrete forms of connotative terms stand for substances, held that their oblique reference is to entities distinct from these substances but inhering in them—these distinct entities are directly named by the abstract forms of such connotative terms. Thus the term *father,* in their view, connotes an entity called fatherhood and implies that it inheres in the thing denoted by the term *father.* Similarly the term *large,* although predicable of terms signifying substances, was said to connote an entity, distinct from such substances but inhering in them, called quantity or magnitude. Ockham was willing to grant that terms signifying sensible qualities, such as *white, hot,* and *sweet,* connote entities that are distinct from substances and are directly signified by the abstract terms *whiteness, heat,* and *sweetness*; hence he admitted as absolute terms the abstract forms of those qualitative predicates. But in all other cases he held that connotative terms, whether concrete or abstract, signify no entities other than those directly signifiable by substance terms or by these absolute quality terms. What the realists had done, in Ockham's view, was to treat facts about substances as entities distinct from those things, as if the fact that a man is six feet tall is an entity distinct from the man but inhering in him, or as if the fact that Socrates has fathered a son is an entity distinct from Socrates and from his son.

From a logical point of view, Ockham's analysis is a restriction of the domain of reference of terms, or of the domain of objects constituting possible values of the variable of quantification, to individual substances and singular (not common) sensible qualities. Ontologically, this means that the only things that there are, are individual substances and equally individual qualities. All terms that are not direct names (or absolute signs) of these objects are predicate terms which, although referring to no other objects than these, do so by indicating a contingent fact about such objects.

In thus impoverishing the domain of objects of reference, Ockham enriches the domain of truths to be known about these objects. The frequent charge that Ockham atomized the world by refusing to recognize relations as real entities distinct from substances and qualities fails to take account of the fact that the connotative terms relate the individuals by implying factual conditions by which the objects are tied together in an existential sense—something that cannot be done by treating relations as entities distinct from their relata and, in effect, as just another class of substances. From Ockham's point of view, it was the realists who atomized the world by treating all predicates as absolute names.

In rejecting the thesis that predicates designate entities distinct from the individuals denoted by absolute terms, Ockham rejects the interpretation of the affirmative copula as a sign of the inherence of an abstract entity in the individuals denoted by the subject term. The truth condition of an affirmative categorical proposition, in Ockham's interpretation, is that subject and predicate "stand for the same." Thus, in the proposition "Socrates is an animal," it is not indicated that Socrates has animality or that animality inheres in Socrates, but it is indicated that the individual denoted by the name "Socrates" is an individual for which the term *animal* stands and which it signifies. In universally quantified propositions, the affirmative copula indicates that every individual for which the subject term stands is something for which the predicate term stands; and in particular, or existentially quantified, propositions, the affirmative copula indicates that there is at least one individual signified by the subject term that is also signified by the predicate term.

This analysis of general propositions corresponds closely to the modern formulas $(x)Fx \supset Gx$ and $(\exists x)Fx \cdot Gx$, except that the medieval analysis requires existential

import as part of the truth condition of the universal affirmative and does not require existential import as a truth condition of the particular negative. In order for subject and predicate to stand for the same, there must be something they stand for; but it is not required that they stand for something in order that they not stand for the same thing. Ockham skillfully carried out the formal development of truth rules for propositions of more complex forms and for various modalities and used them in formulating inference rules both for syllogistic arguments and for arguments based on truth-functional relations between unanalyzed propositions.

SCIENTIFIC KNOWLEDGE. The Aristotelian dictum that science is of the universal was accepted by Ockham in the sense that scientific knowledge is of propositions composed of universal terms, quantified universally for all the individuals signified by the subject term and having the properties of necessity and evidence. Strictly speaking, scientific knowledge is only of demonstrable conclusions evident by reason of indemonstrable, necessary, and evident premises from which they are logically deducible. But Ockham extends the notion of *scientia,* defined as evident grasp of a proposition that is true, to include the indemonstrable premises of demonstrations and also to include evident knowledge of contingent propositions in virtue of intuitive cognition.

EVIDENCE AND SELF-EVIDENCE. Since, for Ockham, the universal propositions of scientific demonstrations are formed only from concepts by which things are apprehended abstractively and without evidence of their existence, the question of what kind of evidence such propositions can have is a crucial question for him. This problem reduces to that of the evidence of the indemonstrable premises of the sciences. Aristotle's characterization of such premises as necessary, self-evident (*per se nota*), and primary could not be accepted by Ockham without considerable qualification. First of all, he says that no such propositions are necessary as assertoric categorical propositions, but are necessary only if they are construed as conditionals or as propositions concerning the possible (*de eo quod potest esse*). Second, he distinguishes between two kinds of evidence that such propositions, construed as conditionals or as of the mode of possibility, may have: the proposition may be evident by the meaning of its terms (*per se nota*) or evident by experience (*nota per experientiam*). The first kind of evidence is obtained through the premises of mathematical demonstrations and by those premises of the natural sciences that are analytically evident by the definition of the terms. But in every natural or physical science there are premises that are not *per se nota* but are established by generalization from singular contingent propositions evident by intuitive cognition; such are the premises that state causal laws or correlate dispositional properties with their commensurately universal subject terms.

INDUCTION. What justifies the passage from singular propositions evident by direct experience to universal propositions affirmed for all possible cases? How does evident knowledge that this particular wood is combustible, acquired by direct observation of its burning, allow us to know that any piece of wood, if subjected to fire in the presence of air, will burn? Ockham invokes as justification for such generalized propositions a rule of induction, described as a *medium extrinsecum,* that corresponds to the principle of the uniformity of nature—that all individuals of specifically similar nature (*eiusdem rationis*) act or react in similar manner to similar conditions. He regards this principle as analytically evident from the meaning of "similar nature"; but since it is logically possible, and hence possible by the power of God, that an effect can be produced without its natural cause, the application of this rule of induction in establishing general premises or laws on the basis of experience of particular cases is valid only within the general hypothesis of the common course of nature (*ex suppositione communis cursus naturae*). Consequently, the evidence of such premises of the natural or positive sciences is not absolute but hypothetical. It should be further noted that Ockham, and his contemporaries as well, drew a sharp distinction between what comes to be by nature and what comes to be by the action of voluntary intelligent agents, both man and God. The principle that like causes produce like effects under like conditions is considered valid only on the supposition that no voluntary agencies are involved.

There is a marked analogy between Ockham's view of the evidential status of the premises of the empirical sciences and that of the premises of positive (or revealed) theology. In the one case their evidence is conditional on the hypothesis of a common course of nature, and in the other on the hypothesis of a revealed order of grace freely (and hence not necessarily) provided by God for the salvation of human souls. Neither hypothesis is logically or metaphysically necessary, and each is, in its own domain, used as a methodological principle pragmatically justified by its fruitfulness. What corresponds to Pelagianism in theology is dogmatic Aristotelianism in natural philosophy, and Ockham takes due precautions against both.

METAPHYSICS AND THEOLOGY

Ockham's metaphysics is primarily a critique of the traditional metaphysical doctrines of his scholastic predecessors. Most of these doctrines represent, in Ockham's view, confusions of logical and physical concepts or of ways of signifying things and the things signified. Such is the case with the supposed distinction, in things, between their essence and their existence, and with the distinction between potential and actual being; to say that something exists does not mean that there is something which is of itself nonexistent to which existence is added, and to say that something exists potentially does not mean that "something which is not in the universe, but can exist in the universe, is truly a being" (*Summa Logicae Pars Prima,* 1951, p. 99, ll. 55–58). These are distinctions between two modalities of statements, assertoric and *de possibili,* and not between things denoted by the terms of statements. The old issue of whether "being" is predicated univocally, equivocally, or analogically of substances and accidents, and of God and creatures, is resolved by saying that in the sense in which "being" is equivalent to "something," it is predicated in the same way of everything there is; but if "univocal" is taken as meaning that the term signifies everything according to a single determinate concept, the term *being* is equivocal and has as many meanings as there are kinds of things. The first sense is like saying $(x)(x = x)$; the second, or equivocal use, is indicated if we say "*to be a man* is not *to be white.*"

SUBSTANCE. The term *substance,* for Ockham, has the sense of Aristotle's primary substance, or *ὑποκείμενον,* rather than the sense of intelligible essence, or *τὸ τὶ ἦν εἶναι.* Basically, substance is conceived as the individual subject or substratum of qualities, and with regard to corporeal substances Ockham indicates that we are aware of substances only as the subject of sensible qualities. Thus he says that "no external corporeal substance can be naturally apprehended in itself, by us, however it may be with respect to the intellect itself or any substance which is of the essence of the knower" (*Commentary on the Sentences* I, d. 3, q. 2), and he adds that "substance is therefore understood in connotative and negative concepts, such as 'being which subsists by itself,' 'being which is not in something else,' 'being which is a subject of all accidents,' etc." (ibid.). These remarks suggest that the general terms of the category of substance are not as absolute as Ockham elsewhere supposes, and that the only nonconnotative concept is the transcendental concept "being" or "thing"; on this basis, general names are eliminated in favor of connotative predicates, proper names are eliminated in favor of descriptive phrases, and the whole category of substance is reduced to the referential function expressed in language by the phrase "thing such that … ," or by what is equivalent to the bound variable of quantification. Historically, Ockham's conception of substance as the posited (or "supposited") referent of the connotative predicates points toward John Locke's "something I know not what" characterization of substance; similarly, Ockham's treatment of sensible qualities as entities distinct from substances (and by the power of God separable, as in the Sacrament of the Altar), along with his contention that quantitative predicates signify nothing other than substances having parts outside of parts, pointed the way to the seventeenth-century treatment of qualities as secondary and quantitative attributes as primary.

MATTER AND FORM. With respect to the notion of cause, Ockham effected a considerable modification of the traditional Aristotelian doctrine. The intrinsic causes, matter and form, were construed physically rather than metaphysically; matter is not, for Ockham, a pure potentiality but is actual in its own right as body having spatially distinguishable parts, its extension being, in the scholastic terminology, the form of corporeity. The concept of form likewise is understood physically in the sense of *μορφή* rather than of *εἶδος,* and tends to be understood as shape and structure of the material parts. This is shown in Ockham's rejection of the notion of a form of the whole (*forma totius*) and in his thesis that a whole is its parts. Many pages of Ockham's works are devoted to the thesis, defended with an almost ferocious intensity, that quantity is not any entity other than substance (or quality), but is substance or sensible qualities as divisible into parts, or as numerable. This doctrine clearly suggests the later view that the primary qualities signified by quantity terms constitute the real essence of substances.

EFFICIENT CAUSES. The tendency toward a more mechanistic theory of natural substances and events is evident in Ockham's treatment of efficient causality. He says that one thing is said to be cause of another if, when it is present, the effect follows, and when it is not present, the effect does not occur. Such a causal relation can be known only by experience, and it is impossible to deduce a priori, from knowledge of one thing, that something else must result from it. This is so on the general epistemological principle that from the cognition of one thing we cannot acquire "first knowledge" of another thing which is really distinct from it but must have intuitive cognition of the latter in itself. Hence the knowledge that one thing is the cause of another, or that something is caused by

some other definite thing, is acquired only if we have intuitive cognition of each of the two things and repeated experience of their concomitance or sequence.

Like Hume, Ockham bases our knowledge of causal relations on experience alone and rejects the doctrine that the effect is virtually in its cause and deducible from the essential nature of the cause. But he is not skeptical with regard to the objectivity of causation; his point is that the only evidence we have of causal connections is experience of observed sequences. Although we cannot establish the causal relations between things a priori, and must accept the principle of the uniformity of nature as an act of faith, Ockham's faith in this principle appears to be as firm as his faith in the revealed doctrines of theology. In his *Summulae Physicorum* (II, c. 12) he says: "Leaving out of consideration all free and voluntary agencies, whatever happens by [natural] causes occurs of necessity and inevitably, and nothing of that sort occurs by chance" (1637 ed., p. 14).

FINAL CAUSES. The Aristotelian doctrine that nature acts for an end is interpreted by Ockham as a pure metaphor. In his *Quodlibeta* (Quod. IV, qq. 1 and 2) he states that it cannot be shown by any self-evident premises or by experience that any effect whatsoever has a final cause, whether distinct from the agent or not distinct from the agent; for that which acts by necessity of nature acts uniformly under like conditions, and it cannot be shown that it does so because of some end desired or aimed at. We speak of natural processes as having ends, not because the agents are really "moved by desire" but simply because natural bodies under similar conditions are observed to act in determinate ways, as if aiming at an end. But such language is purely metaphorical.

In applying his strict criteria of evidence to the doctrines of Aristotelian physics and cosmology, Ockham shows that many principles which Aristotle took to be necessary and self-evident are not. The arguments that celestial bodies have no matter and are ingenerable and incorruptible, that there cannot be a plurality of worlds, and that action at a distance is impossible were held by Ockham to be inconclusive and nonevident. Although Ockham was not concerned with establishing a new physics and cosmology to replace that of Aristotle, his critical treatment of Aristotle's arguments and his constant insistence on the possibility of different theories equally capable of accounting for the facts to be explained were influential in creating the intellectual environment in which later fourteenth-century philosophers explored new physical theories and laid some of the foundations for the scientific revolution of the seventeenth century.

THEOLOGICAL KNOWLEDGE. As a theologian, Ockham was concerned with the question of the cognitive status of theology. The thirteenth-century Scholastics had, for the most part, characterized theology as a science, on the ground that it contains truths which are necessary and "in themselves" evident, even though most of them are not evident to man in his present condition. The question of how we can know that a proposition is evident-in-itself, when it is not evident to us, was answered by saying that a person who does not know geometry may yet be fully assured that a theorem which is an object of belief to him is an object of scientific knowledge to the expert mathematician. Thus, Aquinas said that the articles of faith from which the theologian demonstrates his conclusions are accepted as evident in the light of a higher science (that of God), much as the astronomer accepts the theorems of geometry as premises for his astronomical reasonings but nevertheless demonstrates the conclusions of astronomy in a scientific manner.

Ockham, in a question of his *Commentary on the Sentences* (Prologue, q. 7), examines this and other similar arguments and rejects them as invalid. Every truth evidently known, he says, is either self-evident (*per se nota*), deduced from such, or is evident from intuitive cognition; but the articles of faith are not evidently knowable by man in any of these ways in his present life, for if they were, they would be evident to infidels and pagans, who are not less intelligent than Christians. But this is not the case. Furthermore, it cannot be maintained that theology is a science because it carries out valid processes of deduction of conclusions from the premises accepted on faith, for conclusions cannot be any more evident than the premises from which they are derived.

IMPOSSIBILITY OF NATURAL THEOLOGY. Ockham subjects the *prolegomena fidei,* or propositions about God held to be evidently knowable on natural grounds, to the criteria of evidence and proof that pertain to the natural or philosophical sciences. The issue of whether there is a natural theology as a part of philosophy reduces to the question of whether, from analytic premises evident from the meaning of the terms or from empirical evidence provided by direct experience of the object of theology, such a science is possible. It is conceded by all that man, in his present life, does not have intuitive cognition of God—not, certainly, by getting a degree in theology. But Ockham had argued, with respect to any naturally acquired

knowledge, that it is only by intuitive cognition of an object that we can evidently judge that it exists—and the only objects of which we can have simple abstractive concepts are those we have experienced intuitively or those specifically similar to them. From this it follows that we cannot have any simple and proper concept of God nor any direct evidence of his existence. Can we, then, from concepts derived from experience of other things, form a complex concept or description uniquely applicable to God and prove that an object satisfying this nominal definition exists?

CRITIQUE OF PROOFS FOR GOD'S EXISTENCE. Ockham admits that a descriptive concept of God can be formed from the concept of "being" or "thing" in its univocal (but empty) sense, along with such connotative or negative terms as "nonfinite," "uncaused," and "most perfect." But proving that there exists an object so describable is another matter. The arguments by which his predecessors had attempted to prove God's existence are examined by Ockham with great thoroughness in his *Commentary on the Sentences,* in the *Quodlibeta,* and in the possibly inauthentic *Centiloquium Theologicum.* St. Anselm's so-called Ontological Argument is analyzed (and shown to consist of two different arguments) but is rejected as invalid; and the old arguments from degrees of perfection are disposed of without difficulty.

It is chiefly the causal arguments, in the form used by Duns Scotus, that Ockham takes seriously; and these he examines with extraordinary care because of the way in which Duns Scotus used the concept of infinity in formulating them. Ockham's great logical skill is revealed at its best in his patient and remorseless untangling of the subtleties of the Scotist arguments. Those involving final causality are shown to have no force in themselves, so that the main issues are faced in the arguments from efficient causes. The thesis that there cannot be an infinite regress in the order of efficient causes is rejected as nonevident if the causes are successive in a temporal sense, but Ockham is willing to grant that there cannot be an infinite regress of "conserving causes," since these would have to exist simultaneously. Ockham does, therefore, allow that the existence of at least one conserving cause can be proved if it is granted that there are things whose existence is dependent on conservation by something else; but he immediately points out that we could not prove that there is only one such conserving cause, nor could we prove that the celestial spheres are not sufficient to account for the conservation of the things in the world. Thus the value of this argument for theological purposes is very slight indeed. It is also clear that a natural theology, in the sense involving strictly scientific or evident demonstrations, is completely ruled out by Ockham's basic epistemological principles.

He is willing to concede that it is "probable" that there is one supreme being, that this being is the cause of at least part of the movements and order of the world, and that this being is of an intellectual nature; but since Ockham defines "probable," following Aristotle's *Topics,* as an argument or premise that appears to be true to everyone, to the majority, or to the wisest, all this means is that most people, and the philosophers of old, have believed that there is a deity of this sort.

POSITIVE THEOLOGY. To conclude, from Ockham's merciless criticism of alleged proofs of theological beliefs, that he was an unbeliever and a religious skeptic would be a mistake—although some have drawn this conclusion. There is much evidence in Ockham's writings of an intense loyalty to the Christian faith and of full commitment to the articles of faith as divinely revealed. What Ockham appears to have found objectionable in the theological work of his contemporaries was their attempt to prove what cannot be proved and their loading of theology with pseudo explanations that merely blunted and obscured the tremendous implications of the fundamental articles of the Christian faith. The omnipotence of God and his absolute freedom are the two articles of Christian belief that Ockham never loses sight of; and in his internal treatment of the content of Christian doctrine, just as in his internal treatment of natural philosophy, Ockham invokes these articles of faith as justification for an empiricist or positivistic position. Just as the hypothesis of the common course of nature is a methodological postulate of physical explanation, so the order of grace as set up in the sacramental system and laws of the church is accepted as a postulate of the Christian life; but just as God is not bound or obligated by the order of nature he has established, so he is not bound or obligated by the order of grace he has established as the "common way" of salvation of souls. Neither order is necessary in itself or a necessary consequence of God's being or essence; the utter contingency of the created world, whose existence and order is a sheer fact without any metaphysical ground of necessity, is for Ockham a consequence of the omnipotence and absolute freedom of God that cannot, and should not, be softened or obscured by attempts to construe it in terms of the metaphysics of pagans and infidels.

ETHICAL AND POLITICAL DOCTRINES

In contrast with most of the thirteenth-century scholastic doctors, Ockham made little attempt to formulate a rational psychology or theory of the human soul. In his *Quodlibeta* (Quod. I, q. 10) he raises the question of whether it can be demonstrated that the intellective soul is a form of the body. Since the Council of Vienne had ruled a few years before that this Thomist doctrine was *de fide* (although the formulation was ambiguous enough to allow some latitude), Ockham was not as critical of it as he might otherwise have been. He points out that a person following natural reason would no doubt suppose that his own acts of understanding and of will, of which he has intuitive cognition, are acts of his substantial being or form; however, he would not suppose this to be an incorruptible form separable from his body but rather an extended and corruptible form like that of any other material body. If, however, we must understand by "intellective soul" an immaterial and incorruptible form that exists as a whole in the whole body and as a whole in each part, "it cannot be evidently known by reason or experience that such a form exists in us, nor that the understanding proper to such a substance exists in us, nor that such a soul is a form of the body. Whatever the Philosopher thought of this does not now concern me, because it seems that he remains doubtful about it wherever he speaks of it. These three things are only matters of belief" (Quod. I, q. 10).

Ockham thought that the Franciscan doctrine of a plurality of forms in the human being is more probable on natural grounds than the doctrine of a single form; indeed, if matter has its own corporeal form (*forma corporeitatis*) as extended substance, the sensitive soul would be a distinct form of organization of this matter; and the intellectual soul, if immortal and incorruptible, might well be in the organic body as a pilot is in his boat. But the only evident knowledge we have of ourselves as minds is the intuitive cognition of our acts of thinking and willing, and the subject of these acts is not apprehended directly as a substance or form. Nor is the faculty psychology elaborated by the earlier Scholastics, with its distinctions of active and passive intellect and of really distinct powers within the soul, evident or necessary. We are aware of the soul only as that which thinks and wills; and since the person who thinks is not other than the person who wills, the terms *intellect* and *will* refer to precisely the same subject, and not to distinct entities or faculties within that subject.

FREE WILL. If it is only by intuitive cognition of our own acts that we are aware of ourselves as intelligent beings, it is only in this way that we are aware of ourselves as voluntary agents free to choose between opposite actions. Ockham defines freedom (*libertas*) as "that power whereby I can do diverse things indifferently and contingently, such that I can cause, or not cause, the same effect, when all conditions other than this power are the same" (Quod. I, q. 16). That the will is free, he says, cannot be demonstratively proved by any reason, "because every reason proving this assumes something equally unknown as is the conclusion, or less known." Yet this freedom can be evidently known by experience, he says, because "a man experiences the fact that however much his reason dictates some action, his will can will, or not will, this act" (Quod. I, q. 16).

This liberty of will, for Ockham, is the basis of human dignity and of moral goodness and responsibility, more than the power of thinking—although the two are mutually involved. The seat of morality is in the will itself, Ockham says, "because every act other than the act of will, which is in the power of the will, is only good in such manner that it can be a bad act, because it can be done for an evil end and from an evil intention" (Quod. III, q. 13). Also, every action, other than the act of willing itself, can be performed by reason of natural causes and not freely, and every such action could be caused in us by God alone instead of by our will; consequently, the action in itself is neither virtuous nor vicious, except by denomination from the act of the will. Not even Immanuel Kant was more concerned to distinguish morality from legality, or the good will from the right action. Ockham had, in Peter Abelard, a medieval precedent for this emphasis.

FREE WILL AND GOD'S FOREKNOWLEDGE. Having thus affirmed the total freedom and integrity of the human will, Ockham was faced with the problem of reconciling this with the doctrine of divine foreknowledge of future contingent events, among which the decisions of the human will must be counted. The answer, apparently considered sufficient by Aquinas, that God sees, in one eternal glance, all the decisions of each soul, now and to come, is not sufficient for Ockham. God's intellect is not distinct from his will and his omnipotent causality of all things; hence, says Ockham, "either the determination or production of the created will follows the determination [of the divine will], or it does not. If it does, then the created will acts just as naturally as any natural cause … and thus, the divine will being determined, the created will acts accordingly and does not have the power of not acting accordingly, and consequently no act of the created will is

to be imputed to it" (*Commentary on the Sentences,* d. 38, q. 1). Ockham considers the problem of how God knows, with certainty and from all eternity, the contingent and free decisions of the human will, an insoluble problem; for both the freedom of the human will and the power of God to know all contingent acts of created beings must be conceded. "It is impossible," he says, "for any [created] intellect, in this life, to explain or evidently know how God knows all future contingent events" (d. 38, q. 1).

PROBLEM OF EVIL. While recognizing the Aristotelian conception of natural good and of virtuous choices in accordance with right reason, Ockham is primarily concerned with the theological norm of moral goodness, which is the will of God expressed in the commandments of both the Old Testament and the New Testament, whereby man is obligated (but not coerced) to love and obey God above all else. Thus, what God wills man to do of man's free will defines the right, and disobedience to God's will defines sin. This provides a solution of the old problem of evil, or of God as cause of the sinful acts of man; for since moral evil is the doing of the opposite of what one is obligated to do, and since God is not obligated to any act, it is impossible for God to sin by his causal concurrence in the production of an act sinfully willed by the creature. But Ockham raises an interesting paradox in this connection by supposing that God might command a man to hate him (or to disobey him). To obey God is to love God, and to love God is to do his will; but if it is God's will that I do not do his will, I do his will if I don't, and don't do it if I do. Hence, this command is impossible for a creature to fulfill; and although there would seem to be no patent self-contradiction in supposing that God could issue such a command, it would seem to be self-contradictory, and hence impossible, for God to will that this command be fulfilled.

GOD'S FREEDOM. Although Ockham recognizes that God has established laws binding the Christian to live in a certain way as a member of the church, participant in its sacraments, and believer in its articles of faith, this fact imposes no obligation on God either to bestow eternal life on the Christian who obeys God's precepts and loves him above all else, or to withhold eternal life from those who do not follow God's laws and love him above all else. "It is not impossible," Ockham says, "that God could ordain that a person who lives according to right reason, and does not believe anything except what is conclusive to him by natural reason, should be worthy of eternal life" (*Commentary on the Sentences* III, q. 8). Similarly, although according to the established order an infused grace is required for a man to be eligible for acceptance by God, Ockham insists that God is not necessitated, by reason of such a created grace given to a man, to confer eternal life on him—"always contingently and freely and mercifully and of his own graciousness he beatifies whomsoever he chooses … purely from his kindness he will freely give eternal life to whomsoever he will give it" (*Commentary on the Sentences* I, d. 17, q. 1).

What is distinctive of Ockham's theological point of view is its emphasis on the freedom and spontaneous liberality of God and on the "givenness" of the world that God creates. This stands in sharp contrast to the Muslim characterization of God as the necessary being whose act is equally necessary and therefore determinant of necessity in all that occurs in the created world. Ockham's doctrine of divine omnipotence is not to be understood, as some have done, on the analogy of an oriental potentate issuing arbitrary commands as a pure display of power; rather, it is grounded in the conception of a goodness that is purely spontaneous and unnecessitated, whose gift of existence to creatures and of freedom of choice to man is a perfectly free gift with no strings attached. Ockham's theology of divine liberty and liberality is the complement of his philosophy of radical contingency in the world of existing finite beings and of the underivability of matters of fact from any a priori necessity.

CHURCH AND STATE. Ockham's political and polemical writings on the issue of papal power eloquently convey the thesis that the law of God is the law of liberty and not one of oppression or coercion. The treatise *De Imperatorum et Pontificum Potestate* (On the Power of Emperors and Popes), dealing with the papal claim to plenitude of power, makes this very clear. Christ, in instituting the church, did not give Peter a plenitude of power that would give him the right to do everything not explicitly forbidden by divine or natural law; rather, Peter was given a limited and defined sphere of authority and power. Therefore, Ockham argues, the pope has no authority to deprive any human being of his natural rights or of the rights and liberties given to man by God. "As Christ did not come into the world in order to take away from men their goods and rights, so Christ's vicar, who is inferior and in no way equal to him in power, has no authority or power to deprive others of their goods and rights" (*De Imperatorum* …, p. 10, ll. 12–15). Ockham specifies three of these inalienable rights: first, all those rights that non-Christians justly and admittedly enjoyed before the coming of Christ—for any of these rights to be taken from Christians by papal authority would be to make the liberty of Christians less than that of pagans and infidels;

second, the disposition of temporal things belongs not to the papal authority but to the laity, according to the words of Christ that the things that are Caesar's should be rendered unto Caesar; third, although the pope is charged with the teaching of God's word, maintenance of divine worship, and provision of such things as are necessary for the Christian in his quest for eternal life, the pope has no power to command or requisition those things that are not necessary to this end, "lest he should turn the law of the Gospels into a law of slavery."

On the important question of who is to be the judge of what is necessary for the legitimate ends of the church, Ockham holds that this cannot be the prerogative of the pope, of those under his command, or of the civil rulers. The ultimate decision should be sought in the Gospel, interpreted not by the clergy alone but by "the discretion and counsel of the wisest men sincerely zealous for justice without respect to persons, if such can be found—whether they be poor or rich, subjects or rulers" (*De Imperatorum*…, p. 27, ll. 17–20). This not very practical proposal nevertheless suggests that the membership of the Christian community as private individuals, rather than as officeholders, constitutes the true church. Yet Ockham is not, like Marsilius of Padua, against the principle of the pope as head of the church and vicar of Christ; he only seeks safeguards against abuse of the papal office and illegitimate assumption of tyrannical powers by holders of that office. Legitimate sovereignty, whether papal or civil, is not despotism; the dominion a master has over a slave is not the kind of authority exercised legitimately by a king, pope, or bishop. A pope may turn out to be a heretic and may be deposed—not by the emperor but only by a general council of the church. The imperial power derives from God, not directly but by way of the people who confer upon the emperor his power to legislate; the imperial power is not, as the popes had claimed, derived from the papacy. Ockham's political theory, insofar as it was formulated at all in his polemical writings, was not secularist or anticlerical; it was against absolutism in either church or state and much concerned that the "law of force," which is characteristic of the civil state, should not be adopted by the papal authority, lest the law of God, which is a law of liberty, be corrupted and degraded by temporal ambitions and lust for power.

See also Abelard, Peter; Anselm, St.; Aristotle; Avicenna; Bonaventure, St.; Degrees of Perfection, Argument for the Existence of God; Descartes, René; Determinism, A Historical Survey; Duns Scotus, John; Empiricism; Evil, The Problem of; Hume, David; Induction; Intentionality; Leibniz, Gottfried Wilhelm; Logic, History of; Marsilius of Padua; Medieval Philosophy; Neoplatonism; Ockhamism; Ontological Argument for the Existence of God; Peter Aureol; Peter Lombard; Peter of Spain; Realism; Semantics, History of; Socrates; Thomas Aquinas, St.; Universals, A Historical Survey; William of Sherwood.

Bibliography

WORKS BY OCKHAM

The critical editions of Ockham's works are in three series:

Opera Theologica (*Op. Theol.*). 10 vols., edited by Gedeon Gál et al. St. Bonaventure, NY: Franciscan Institute Publications, 1967–1986.

Opera Philosophica (*Op. Phil.*). 7 vols., edited by Gedeon Gál et al. St. Bonaventure, NY: Franciscan Institute Publications, 1974–1988.

Opera Politica (*Op. Pol.*). 4 vols., edited by H. S. Offler. Vol. I–III. Manchester, U.K.: Manchester University Press; Vol. IV, Oxford: Oxford University Press, 1940–1997 (incomplete).

The main works are:

Scriptum in Librum Primum Sententiarum (*Ordinatio*), edited by Gedeon Gál, Stephen Brown, Jerry I. Etzkorn, and Francis E. Kelley. *Op. Theol* I–IV, 1967–1979.

Quaestiones in Libros Secundum, Tertium et Quartum Sententiarum (*Reportatio*), edited by Gedeon Gál, Rega Wood, Jerry I. Etzkorn, Francis E. Kelley, and Romuald Green. *Op. Theol.* V–VII, 1981–1984.

Quodlibeta Septem, edited by Joseph C. Wey. *Op. Theol.* IX, 1980. English translation by Alfred J. Freddoso and Francis E. Kelley, *Quodlibetal Questions.* 2 vols. New Haven, CT: Yale University Press, 1991.

Summa Logicae, edited by Philotheus Boehner, Gedeon Gál, and Stephen Brown. *Op. Phil.* I, 1974. English translation of Part I by Michael J. Loux, *Ockham's Theory of Terms.* Notre Dame, IN: University of Notre Dame Press, 1974. English translation of Part II by Alfred J. Freddoso and Henry Schuurman, *Ockham's Theory of Propositions,* Notre Dame, IN: University of Notre Dame Press, 1980.

Expositio in Librum Porphyrii De Praedicabilibus, edited by Ernest A. Moody. In *Op. Phil.* II, 1978: 8–131. English translation by Eike-Henner W. Kluge, " William of Ockham's Commentary on Porphyry. " *Franciscan Studies* 33 (1973): 171–254, and 34 (1974): 306–382.

Tractatus de Praedestinatione et de Praescientia Dei Respectu Futurorum Contingentium, edited by Philotheus Boehner and Stephen Brown. In *Op. Phil.* II, 1978: 505–539. English translation by Marilyn McCord Adams and Norman Kretzmann, *Predestination, God's Foreknowledge and Future Contingents.* New York: Appleton-Century-Crofts; 2nd ed., 1983.

Expositio in Libros Physicorum Aristotelis, edited by Vladimir Richter et al. *Op. Phil.* IV–V, 1985.

Brevis Summa Libri Physicorum, edited by Stephen Brown. In *Op. Phil.* VI, 1984: 1–134. English translation by Julian Davies, *Ockham on Aristotle's Physics.* St. Bonaventure, NY: Franciscan Institute, 1989.

Opus Nonaginta Dierum, edited by H. S. Offler and J. G. Sikes, in *Op. Pol.* I, 2nd ed., 1974: 287–368, and *Op. Pol.* II, 1963: 375–858. English translation by John Kilcullen and John Scott, *The Work of Ninety Days.* CD-ROM, Bowling Green, OH: Philosophy Documentation Center.

Breviloquium de Principatu Tyrannico, edited by H. S. Offler, in *Op. Pol.* IV, 1997: 79–277. English translation by John Kilcullen, *A Short Discourse on Tyrannical Government.* Cambridge, U.K.: Cambridge University Press, 1992.

Dialogus, edited by and English translation by John Kilcullen et al. in progress on the web. Available from http://www.britac.ca.uk/pubs/dialogus/ockdial.html.

A Letter to the Friars Minor and Other Writings. English translation by John Kilcullen of selected political writings from Ockham. Cambridge, U.K.: Cambridge University Press, 1995.

On the Power of Emperors ad Popes (*De Imperatorum et Pontificum Potestate*), edited by and English translation by Annabel S. Brett. Bristol, U.K.: Thoemmes Press for the University of Durham, 1998.

WORKS ON OCKHAM

Abbagnano, Nicola. *Guglielmo di Ockham.* Lanciano, Italy: Carabba, 1931.

Adams, Marilyn McCord. *William Ockham.* 2 vols. Notre Dame, IN: University of Notre Dame Press, 1987.

Baudry, L. *Guillaume d'Occam.* Vol. I: *L'homme et les oeuvres.* Paris, 1950. Contains extensive bibliography of works on Ockham; full information on Ockham's published and unpublished works will be found on pp. 273–294.

Baudry, L. *Lexique philosophique de Guillaume d'Ockham.* Paris, 1958.

Baudry, L. *Le Tractatus de Principiis Theologiae attribué à G. d'Occam.* Paris, 1936.

Beckmann, Jan P. *Ockham-Bibliographie 1900–1990.* Hamburg: Felix Meiner, 1992.

Biard, Joël. *Guillaume d'Ockham. Logique et philosophie.* Paris: P.U.F., 1997.

Biard, Joël. *Logique et théorie du signe au XIVe siècle.* Paris: Vrin, 1989.

Boehner, Philotheus. *Collected Articles on Ockham.* St. Bonaventure, NY, 1956.

Courtenay, William J. *Schools and Scholars in Fourteenth-Century England.* Princeton, NJ: Princeton University Press, 1987.

Federhofer, Franz. *Die Erkenntnislehre des Wilhelm von Ockham.* Munich, 1924.

Freppert, Lucan. *The Basis of Morality according to William Ockham.* Chicago: Franciscan Herald Press, 1988.

Goddu, André. *The Physics of William of Ockham.* Leiden: Brill, 1984.

Gottfried, Martin. *Wilhelm von Ockham.* Berlin: de Gruyter, 1949.

Guelluy, Robert. *Philosophie et théologie chez Guillaume d'Ockham.* Louvain and Paris, 1947.

Hochstetter, Erich. *Studien zur Metaphysik und Erkenntnislehre Wilhelms von Ockham.* Berlin, 1927.

Holopainen, Taina M. *William of Ockham's Theory of the Foundations of Ethics.* Helsinki: Luther-Agricola Society, 1991.

Lagarde, Georges de. *La naissance de l'esprit laïque au déclin du moyen âge,* Vols. IV–VI. Paris, 1942–1946.

Leff, Gordon. *William of Ockham. The Metamorphosis of Scholastic Discourse.* Manchester, U.K.: Manchester University Press, 1975.

Maurer, Armand. *The Philosophy of William of Ockham in the Light of Its Principles.* Toronto: Pontifical Institute of Mediaeval Studies, 1999.

McGrade, Arthur Stephen. *The Political Thought of William of Ockham.* Cambridge, U.K.: Cambridge University Press, 1974.

Michon, Cyrille. *Nominalisme. La théorie de la signification d'Occam.* Paris: Vrin, 1994.

Moody, E. A. *The Logic of William of Ockham.* New York and London: Sheed and Ward, 1935.

Moody, E. A. *Truth and Consequence in Medieval Logic.* Amsterdam, 1953.

Moser, Simon. *Grundbegriffe der Naturphilosophie bei Wilhelm von Ockham.* Innsbruck, 1932.

Panaccio, Claude. *Le discours intérieur. De Platon à Guillaume d'Ockham.* Paris: Éd. du Seuil, 1999.

Panaccio, Claude. *Les mots, les concepts et les choses. La sémantique de Guillaume d'Occam et le nominalisme d'aujourd'hui.* Montréal: Bellarmin, and Paris: Vrin, 1992.

Panaccio, Claude. *Ockham on Concepts.* Aldershot, U.K.: Ashgate, 2004.

Scholz, Richard. *Wilhelm von Ockham als politischer Denker und sein Breviloquium de Principatu Tyrannico.* Leipzig: Hiersemann, 1944.

Shapiro, Herman. *Motion, Time and Place according to William Ockham.* St. Bonaventure, NY, 1957.

Spade, Paul Vincent, ed. *The Cambridge Companion to Ockham.* Cambridge, U.K.: Cambridge University Press, 1999.

Spade, Paul Vincent. *Lies, Language and Logic in the Late Middle Ages.* London: Variorum Reprints, 1988.

Spade, Paul Vincent. *Thoughts, Words and Things. An Introduction to Late Mediaeval Logic and Semantic Theory.* Available from http://www.pvspade.com/Logic/index.html, 1996.

Tachau, Katherine H. *Vision and Certitude in the Age of Ockham.* Leiden: Brill, 1988.

Vasoli, Cesare. *Guglielmo d'Occam.* Florence, 1953. Contains an extensive bibliography of works on Ockham.

Vignaux, Paul. *Justification et prédestination au XIV[e] siècle.* Paris: Leroux, 1934.

Vignaux, Paul. "Nominalisme" and "Occam." In *Dictionnaire de théologie catholique.* 15 vols. Paris, 1903–1950. Vol. XI, cols. 733–789 and 864–904.

Vignaux, Paul. *Le nominalisme au XIV[e] siècle.* Montreal, 1948.

Webering, Damascene. *The Theory of Demonstration according to William Ockham.* St. Bonaventure, NY: Franciscan Institute, 1953.

Wolter, Allan B. et al. *William of Ockham (1285–1347). Commemorative Issue.* 3 vols. *Franciscan Studies* 44–46 (1984–1986).

Wood, Rega. *Ockham on the Virtues.* West Lafayette, IN: Purdue University Press, 1997.

Zuidema, Sytse. *De Philosophie van Occam in Zijn Commentaar op de Sententien,* 2 vols. Hilversum, Netherlands, 1936.

Ernest A. Moody (1967)
Bibliography updated by Claude Panaccio (2005)

WILLIAM OF SHERWOOD

(1200/1210–1266/1271)

William of Sherwood, or Shyreswood, was an English logician. All that is known for certain of William of Sherwood's life is that in 1252 he was a master at Oxford, that he became treasurer of the cathedral church of Lincoln soon after 1254, that he was rector of Aylesbury and of Attleborough, that he was still living in 1266, and that he was dead in 1271. From references in his works, however, and from the fact that his logic almost certainly had a direct influence on the logical writings of Peter of Spain, Lambert of Auxerre, Albert the Great, and Thomas Aquinas, all of whom were at Paris around the same time, it seems undeniable that he taught logic there from about 1235 to about 1250.

William's impact on his contemporaries went unacknowledged except by Roger Bacon, who, in his *Opus Tertium* (1267), described him as "much wiser than Albert [the Great]; for in *philosophia communis* no one is greater than he." Bacon's phrase *philosophia communis* must refer to logic; no other kind of work can be definitely attributed to William, and his logical works certainly were influential. They consist of an *Introductiones in Logicam,* a *Syncategoremata,* a *De Insolubilibus* (on paradoxes of self-reference), an *Obligationes* (on rules of argument for formal disputation), and a *Petitiones Contrariorum* (on logical puzzles arising from hidden contrariety in premises). Only the first two were ever published; they are longer and far more important than the last three. A commentary on the *Sentences,* a *Distinctiones Theologicae,* and a *Conciones* (a collection of sermons) have also been attributed to William, though their authenticity is seriously questioned.

The *Introductiones* consists of six treatises, the first four and the last one of which correspond (very broadly) to Aristotle's *De Interpretatione, Categories, Prior Analytics, Topics,* and *Sophistical Refutations,* in that order. The third treatise contains the earliest version of the mnemonic verses for the syllogism "Barbara, Celarent …," and there are other interesting minor innovations in those treatises. The most important novelties are concentrated in the fifth treatise, "Properties of Terms"; it contains the logico-semantical inquiries that gave the terminist logicians their name. William recognizes four properties of terms—*significatio, suppositio, copulatio,* and *appellatio.* The last three may be very broadly described as syntax-dependent semantical functions of a term's *significatio,* which is its meaning in the broadest sense.

In order to distinguish such medieval contributions from strictly Aristotelian logic, thirteenth-century philosophers spoke of them as *logica moderna.* When William wrote, *logica moderna* was thought of as having two branches, *proprietates terminorum* and *syncategoremata.* In his separate treatise on the latter, William investigates the semantical and logical properties of such syncategorematic words as *every, except, only, is, not, if, or, necessarily.* Both branches may be said to be concerned with the points of connection between syntax and semantics and with the effect those points have on the evaluation of inferences. William's treatment of both is marked by a concern with the philosophical problems to which they give rise.

The ingredients of the *logica moderna* certainly antedate William's writings, but his may very well be the earliest full-scale organization of those elements in the way that became characteristic of medieval logic after his time.

See also Albert the Great; Aristotle; Bacon, Roger; Logic, History of; Medieval Philosophy; Peter of Spain; Semantics, History of; Thomas Aquinas, St.

Bibliography

Martin Grabmann has edited *Die Introductiones in logicam des Wilhelm von Shyreswood, Literarhistorische Einleitung und Textausgabe* in *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Philosophische-historische Abteilung,* Jahrgang 1937, Vol. 10 (Munich, 1937). Norman Kretzmann has translated *William of Sherwood's Introduction to Logic* (Minneapolis: University of Minnesota Press, 1966), and J. Reginald O'Donnell has edited "The Syncategoremata of William of Sherwood" in *Medieval Studies* 3 (1941): 46–93.

Norman Kretzmann (1967)

WILLIAMS, BERNARD

(1929–2003)

Bernard Arthur Owen Williams, an English philosopher, was educated at Balliol College, Oxford, and received his BA in 1951. He was a Fellow of All Souls College, Oxford, and went on to teach at New College, Oxford, University College London, and Bedford College, London, before moving in 1967 to Cambridge as Knightbridge Professor and Fellow of King's College; he was Provost of King's from 1979 to 1987. In 1988 he became a professor at Berkeley, then in 1990 was appointed White's Professor of Moral Philosophy at Oxford. An English public figure as well as a distinguished thinker, he was chairperson of the

Committee on Obscenity and Film Censorship and served on the Royal Commission on Gambling, the Labour Party's Commission on Social Justice, and the Independent Inquiry into the Misuse of Drugs Act, as well as on the Board of the English National Opera. He was knighted in 1999.

Williams was a brilliant and versatile contributor to many branches of philosophy and its history. Trained in classics, he wrote about Plato and Aristotle, and also, in *Shame and Necessity*, about the ethical consciousness of classical Greece as revealed in its literature, law, and culture. He wrote an important book about René Descartes, and was profoundly drawn to the work of Friedrich Nietzsche. But his main contributions are his own ideas about knowledge, truth, reality, the self, ethics, and morality.

Williams did not offer a systematic philosophical theory and was distrustful of such theories; instead he tried to bring clarity and a recognition of complexity and historical contingency to a number of central philosophical problems. A theme throughout his work was how to combine the point of view of the individual with the conception of the world encouraged by the scientific ideal of objectivity and its kin. An early example is his paper, *The Self and the Future*, about the problem of personal identity over time, which showed that the first-person conception of the self is more favorable to a physical condition of personal identity than to a condition based on psychological similarity.

In his book on Descartes, he introduced the fruitful notion of the absolute conception of reality—a conception that would be free of every contingency of the human perspective and would therefore describe the world as it is in itself, not merely as it appears to us—, or the world that is there anyway, as he put it. This conception drives the pursuit of scientific objectivity, but also raises the question whether humans can reasonably hope to approach it. Williams thought the view *sub specie aeternitatis* was a reasonable goal for science, but rejected its authority for ethics.

He used the term *ethics* for the general topic of how to live, and *morality* for the special type of modern theory of right and wrong that is based on some form of impartiality or universalizability over all persons. Impartial morality, he argued in *Ethics and the Limits of Philosophy* and elsewhere, does not have an adequate basis in human motivation for the authority it claims over the individuals to whom it is addressed. The appropriate standpoint for assessing human conduct is *from here* not from an external vantage point assumed to be the same for everyone.

Williams held more generally that all reasons for action are internal reasons, by which he meant reasons derived from some desire or interest already present in the agent's subjective motivational set. External reasons, such as those Kant imagined the categorical imperative to provide, do not exist. It follows that moral requirements in particular must be rooted in already existing desires and commitments, and that they may be less than universal in their application. Williams also embraced a qualified relativism, whereby we can morally appraise only forms of life that constitute real options for us: It makes no sense for us to judge either right or wrong the moral beliefs of a medieval samurai, for example.

He had a large impact on moral philosophy through his claim that impersonal morality undermines the integrity of individual life by requiring us to detach from our most fundamental projects and personal commitments, the things that give life its substance and make it worth living. Utilitarianism does this by asking that we regard the attainment of our own aims simply as part of the general welfare, and ourselves as instruments of the universal satisfaction system. But Kantian universalisability, too, requires us to act on our deepest commitments only under the authorization of the higher-order principle that anyone in our situation may do the same—for example, rescue one's own child from drowning rather than a stranger. This, said Williams, is one thought too many. The core of personal life cannot survive subordination to the impersonal standpoint. The exploration of this critique and responses to it have become a focal point of moral theory.

Williams was skeptical about what he called the morality system, and of ethical theory, but he was not a moral skeptic: Morality, he thought, should seek confidence rather than theoretical foundations, and he himself held strong moral views. He believed that ethical judgments were often supported by less universal, more local grounds—particularly judgments involving thick moral concepts like cruelty, courage and chastity. But he drew the corollary that ethical knowledge expressed by those concepts can be lost if the practices and forms of life that underlie them disappear.

Williams formulated the important concept of moral luck, a term he invented for the phenomenon of our moral vulnerability to factors that are not under our control, so that what we are guilty of may depend partly on the actual, and not merely the foreseeable, results of our choices. This possibility was strenuously denied by Kant, but it is central to the moral content of tragedy, one of Williams's great subjects. He rejected the ideal of finding

principles of choice which would guarantee that if we follow them, we will have no reason to reproach ourselves later, whatever happens.

His final book, *Truth and Truthfulness*, pursued the reconciliation of his commitment to objectivity about factual, scientific, and historical truth with his resistance to the claims of objectivity in ethics. He attacked general postmodernist skepticism about truth, explained the vital moral importance of respect for factual truth, especially in politics, and analyzed the historical development of our ideas about truth, lying, and authenticity, starting with an imagined prehistory and then proceeding from the ancient world to the present.

See also Aristotle; Descartes, René; Ethics, History of; Kant, Immanuel; Nietzsche, Friedrich; Personal Identity; Plato; Truth; Utilitarianism.

Bibliography

PRIMARY WORKS

Morality: An Introduction to Ethics. New York: Harper & Row, 1972.

Problems of the Self; Philosophical Papers, 1956–1972. Cambridge U.K.: Cambridge University Press, 1973.

"A Critique of Utilitarianism." In *Utilitarianism; For and Against*, by J. J. C. Smart and Bernard Williams. Cambridge U.K.: Cambridge University Press, 1973.

Descartes: the Project of Pure Inquiry. Hassocks, U.K.: Harvester Press, 1978.

Moral Luck: Philosophical Papers, 1973–1980. Cambridge, U.K.; New York: Cambridge University Press, 1981.

Ethics and the Limits of Philosophy. Cambridge, MA: Harvard University Press, 1985.

Shame and Necessity. Berkeley: University of California Press, 1993.

Making Sense of Humanity and Other Philosophical Papers, 1982–1993. Cambridge, U.K.; New York: Cambridge University Press, 1995.

Truth and Truthfulness: An Essay in Genealogy. Princeton NJ: Princeton University Press, 2002.

Thomas Nagel (2005)

WILSON, EDWARD O.

(1929–)

Edward O. Wilson was born in Birmingham, Alabama, on June 10. His first degree was in biology from the University of Alabama. He moved north to Harvard as a graduate student, remaining there for the rest of his working life, first as a doctoral student, then as a junior fellow, and next as a member of the department of biology (later the department of organismic biology), retiring 2000 as a University Professor. Wilson is married with one child. He has received much acclaim, including the Pulitzer Prize for nonfiction (twice), the Craaford Prize of the Swedish Academy of Science, membership in the National Academy of Sciences, and fellowship in the Royal Society.

Wilson's abiding passion has been the world of ants. He has authored books on their nature, their behavior, and their classification. His *magnum opus* is *The Ants* (1990), jointly authored with Bert Holldöbler. This book won Wilson one of his Pulitzer Prizes. Another of Wilson's interests, arising from the ant studies, has been biogeography, the study of the distributions of organisms. With the late Robert MacArthur, in the 1960s, Wilson proposed an important theory of island flora and fauna, arguing that immigration and emigration and extinction eventually reach equilibrium. The ants also led naturally to an interest in chemical communication, with Wilson studying the use of pheromones for information transmission.

From here, Wilson was led into more general issues pertaining to social behavior, and a trilogy ensued. First there was *The Insect Societies* (1971), in which Wilson considered what we now know about the insects and their behaviors, paying special reference to the so-called social insects (especially the hymenoptera: the ants, the bees, and the wasps). Next came *Sociobiology: The New Synthesis* (1978), a book that popularized the term "sociobiology" (meaning the study of social behavior from an evolutionary perspective), in which Wilson extended and developed his thinking, covering the whole of the animal kingdom, including our own species. Finally there was *On Human Nature* (1978), written in a somewhat more popular fashion, and for which Wilson won the other of his Pulitzer Prizes. In this final book of the trilogy, Wilson turned exclusively to humankind, arguing that much that we know about the evolution of social behavior in other animals applies almost equally to humans.

Wilson's forays into human sociobiology were highly controversial. Some critics contended that in the guise of objective science, he simply defended conservative views of society, while social scientists argued that he had no feeling for the subtleties and ranges of human culture. Wilson defended and extended his thinking, pointing out that taking a biological perspective does not at once commit one to a hard-line deterministic position. It has never been his position that the genes are the sole causal factor behind human nature. It is just that biology must be

accorded equal causal weight in human affairs alongside the environment and culture.

More and more, through the 1980s, Wilson turned to philosophical questions. With respect to the theory of knowledge (epistemology), Wilson stresses the interconnected nature of our understanding. He wants to show that everything can be explained in just a few basic principles. The Victorian polymath William Whewell, in his *The Philosophy of the Inductive Sciences*, spoke of the highest kind of knowledge as being that which connects together the most disparate areas of science. Whewell spoke of such connection as a "consilience of inductions," and this phrase prompted Wilson to call one of his books *Consilience* (1998), referring to its plea that we bind together all aspects of human knowledge.

Along with epistemology, ethics has always been an interest of Wilson's. His hero in this field is Herbert Spencer, and although Wilson would not want to associate himself with the negative connotations of attempts to link evolution and morality—especially with so-called Social Darwinism—Wilson stands right in the tradition of those who argue that morality is and must be based in human nature as created and preserved by evolution. What is of great importance to Wilson is the need to be sensitive to the environment around us. He speaks of "biophilia," the human love of nature. He believes that we need nature not just to sustain us but also because, in a totally artificial world, we humans would wither and die. Our evolution has tied us to both physical and psychological needs of other organisms. This means that the Wilsonian categorical imperative focuses on biodiversity. In a world without many species, humans are condemned. Following his own prescriptions, for the past decade Wilson has been ardently committed to the preservation of the Brazilian rain forests.

Like Spencer and all other traditional thinkers of this ilk, Wilson turns to notions of progress to link evolution and ethics. Most particularly, he denies that the evolutionary process is one of aimless meandering. Rather, Wilson interprets it as showing an upward rise, from lesser to greater, with humans at the top. Wilson's thinking on this point is part and parcel of his feelings about ultimate questions. An intensely religious man who lost his faith in Christianity in his teens, Wilson was able to replace it with a new religion: Darwinism. He sees religion as an essential part of human culture, binding the tribe together, but he argues that this religious cohesion can endure in the modern age only with the propagation of new "myths" (his word). This is the essential message of Wilson's *On Human Nature* (1978). This is the story of evolution with the philosophical foundation of materialism. For Wilson, science, ethics, and religion are as one. They make for the ultimate consilience.

See also Darwinism; Evolutionary Ethics; Materialism; Organismic Biology; Philosophy of Biology; Whewell, William.

Bibliography

The key books by Wilson include the work on biogeography coauthored with Robert MacArthur, *The Theory of Island Biogeography* (Princeton, NJ: Princeton University Press, 1967); the trilogy *The Insect Societies* (Cambridge, MA: Harvard University Press, 1971); *Sociobiology: The New Synthesis* (Cambridge, MA: Harvard University Press, 1975); *On Human Nature* (Cambridge, MA: Harvard University Press, 1978); the coauthored work with Bert Holldöbler, *The Ants* (Cambridge, MA: Harvard University Press, 1990); *The Diversity of Life* (Cambridge, MA: Harvard University Press, 1992); and *Consilience* (New York: Knopf, 1998). The work by William Whewell that so influenced Wilson is *The Philosophy of the Inductive Sciences* (London, 1840; London: Cass, 1967). A more detailed overview of Wilson's work and thinking can be found in Michael Ruse, *Mystery of Mysteries: Is Evolution a Social Construction?* (Cambridge, MA: Harvard University Press, 1999). Also informative, not only about Wilson but about his various colleagues, friends, and enemies, is his autobiography *Naturalist* (Washington, DC: Island Press, 1994).

Michael Ruse (2005)

WINCKELMANN, JOHANN JOACHIM
(1717–1768)

Johann Joachim Winckelmann, the German art historian and founder of scientific archaeology, was born at Stendal in Prussia. After early schooling in Stendal and Berlin, he studied theology and classics at Halle and mathematics and medicine at Jena. He held a series of minor positions and then became a librarian at Nöthnitz, near Dresden, where he met many artists and critics who stimulated his interest in the fine arts. Influenced by the papal nuncio in Dresden, Winckelmann became a Catholic; and in 1755, after the publication of his first important work, *Gedanken über die Nachahmung der griechischen Werke in der Malerei und Bildhauerkunst* (Thoughts on the imitation of Greek works in painting and sculpture; Dresden and Leipzig, 1754), he went to Rome on a royal subsidy. In Rome he was supported by various high churchmen. In 1758 he visited Naples, Herculaneum, and Pompeii and spent a longer period in Florence. In 1760 he became

librarian and surveyor of antiquities to Cardinal Albani and wrote his *Anmerkungen über die Baukunst der Alten* (Remarks on the architecture of the ancients; Leipzig, 1762). In 1763 he was appointed general surveyor of antiquities for Rome and Latium. While general surveyor he published *Abhandlung über die Fähigkeit der Empfindung des Schönen in der Kunst und dem Unterricht in derselben* (Treatise on the power of feeling beauty and on teaching it; Dresden, 1764); *Geschichte der Kunst des Alterthums* (History of ancient art; Dresden, 1764); and *Versuch einer Allegorie, besonders für die Kunst* (An essay on allegory, especially for art; Dresden, 1766). In 1768 Winckelmann was murdered in an inn at Trieste.

Winckelmann was the founder of classical archaeology and of art history. He was the first person to consider a work of art not only as an item of contemplative pleasure and imitation or as an object of erudite commentary and psychological characterization, but as a creation of a particular nation and period with its own special geographical, social, and political conditions, which expresses the style of the spirit of the milieu as a whole.

Winckelmann's aesthetic theory is found mostly in scattered remarks in his works on ancient art, and his ideas were constantly evolving. They were methodological by-products of his work as a historian systematizing the history of ancient art. For these reasons any reconstruction of Winckelmann's aesthetic doctrines is controversial. These views were nevertheless systematized by his contemporaries, and extended from ancient art to literature both ancient and modern.

Winckelmann was dissatisfied with all received definitions of beauty, and he held that beauty is indefinable—that it is one of the greatest mysteries of nature, and beyond the limits of human understanding. (There is nevertheless an absolute standard of taste. But this cannot be deduced; it must be grasped through a deeper insight into actual works of art.) One general characteristic of beauty is proportion; but to dead proportion must be added living form.

Expression (*Ausdruck*) is a lower stage of beauty. It is a lively imitation of both the soul and the body as passive and active. Pure beauty is reached through the stillness of this feeling of life. The highest stage of beauty arises from the unification of expression and pure beauty in grace. By this unity beauty becomes an appearance of divinity in the representation of a sensible object. The unity of a work of art arises mainly from simplicity (*Einfalt*) and measure (*Mässigung*), or the harmony of opposing traits—for instance, understanding and passion. This process of unification corresponds to the rise from sensible to ideal beauty, or from the imitation of nature to the creation of a higher nature. The observation of nature gives us the means of overcoming spurious standards of beauty and a set of samples to be used by the intellect in creating the higher nature.

Beauty is felt by the senses, but it is understood and created by the intellect (*Verstand*)—which is the faculty of ideas as well as of distinct concepts. The "ideal" (*Das Ideale*), or "spirit" (*Geist*), is the most important and controversial notion in Winckelmann's aesthetics. One kind of ideal is created when an artist combines in one unique whole elements of beauty among different natural objects—for example, by constructing a perfect female figure from separate parts imitating parts of different real women, each of which is the most perfect of its kind. A superior kind of ideal arises when the choice of parts is directed not only by a feeling for proportion, but by a supernatural idea translated into matter—for example, the superhuman perfection of a particular human type or quality such as the combination of attractive manhood and pleasing youthfulness in the Apollo del Belvedere, or of enormous pain in a great soul in the Laocoön. The second kind of ideal is not abstracted from experience, but is derived from an intuition of the beauty of God himself. It is realized through a creative process like that of God creating his own image in man. Ideal beauty of the second kind must show "noble simplicity and quiet greatness" (*edle Einfalt und stille Grösse*). Immanuel Kant later systematized this double conception in his *Critique of Judgment.*

Because beauty in its highest form is spiritual, it must suggest a deeper ethical meaning. These ethical thoughts are the content of real art. Art makes them intuitively known through allegory. Nature also presents allegories to man; and man himself spoke through images before he spoke in rational language. Painting, sculpture, and poetry all express through allegory invisible things; and thus allegory is the foundation of the unity of the different fine arts.

Simplicity, or unity, gives distinctness (*Deutlichkeit*) to a work of art. Winckelmann held therefore that there is an intuitive, or sensible, distinctness, whereas the then current psychology admitted only intellectual distinctness and allowed only clarity to sensibility. Kant, later, was the first to introduce the concept of intuitive distinctness into the theory of knowledge.

Winckelmann saw in Greek art the standard of ideal beauty. The Greek man was the most spiritually and ethically balanced, and therefore the most physically perfect, because of various climatic, geographical, historical,

social, and political conditions. Greek artists could therefore use the most beautiful human specimens as models, and they should be imitated by modern artists. Imitation of nature and imitation of the Greeks is the same thing.

See also Aesthetic Judgment; Aesthetics, History of; Art, Value in; Beauty; Kant, Immanuel.

Bibliography

WORKS BY WINCKELMANN

Monumenti antichi inediti, 2 vols. Rome, 1767.

Werke, 12 vols. Edited by Joseph Eiselein. Donaueschingen, 1825–1829.

Werke. Edited by W. Rehm and H. Diepolter. Berlin, 1952–. Critical edition.

Kleine Schriften und Briefe. Edited by W. Senf. Weimar: H. Böhlaus Nachfolger, 1960.

WORKS ON WINCKELMANN

Aron, Erich. *Die deutsche Erweckung des Griechentums durch Winckelmann und Herder.* Heidelberg, 1929.

Baumecker, Gottfried. *Winckelmann in seinen Dresdner Schriften.* Berlin: Junker and Dünnhaupt, 1933.

Curtius, Ludwig. *Winckelmann und seine Nachfolge.* Vienna: Schroll, 1941.

Hatfield, H. C. *Winckelmann and His German Critics, 1755–1781.* New York: King's Crown Press, 1943.

Justi, Carl. *Winckelmann und seine Zeitgenossen,* 3 vols. Leipzig, 1866–1873; 4th ed. Leipzig: Koehler and Amelang, 1943.

Rehm, W. *Winckelmann und Lessing.* Berlin, 1941.

Vallentin, Berthold. *Winckelmann.* Berlin: Bondi, 1931.

Zbinden, W. *Winckelmann.* Bern, 1935.

Giorgio Tonelli (1967)

WINDELBAND, WILHELM
(1848–1915)

The German philosopher and historian of philosophy Wilhelm Windelband was born in Potsdam and educated at Jena, Berlin, and Göttingen. He taught philosophy at Zürich, Freiburg im Breisgau, Strasbourg, and Heidelberg. He was a disciple of Rudolf Hermann Lotze and Kuno Fischer and was the leader of the so-called southwestern German (or Baden) school of neo-Kantianism. He is best known for his work in history of philosophy, to which he brought a new mode of exposition—the organization of the subject by problems rather than by chronological sequence of individual thinkers. As a systematic philosopher he is remembered for his attempt to extend the principles of Kantian criticism to the historical sciences, his attempt to liberate philosophy from identification with any specific scientific discipline, and his sympathetic appreciation of late nineteenth-century philosophy of value.

Windelband believed that whereas the various sciences (mathematical, natural, and historical) have specific objects and limit their investigations to determined areas of the total reality, philosophy finds its unique object in the knowledge of reality provided by these various disciplines taken together as a whole. The task of philosophy, he held, was to explicate the a priori bases of science in general. The aim of philosophy was to show not how science is possible but why there are many different kinds of science; the relationships that obtain between these various sciences; and the nature of the relation between the critical intelligence—the knowing, willing, and feeling subject—and consciousness in general.

According to Windelband, both the triumphs and the limitations of contemporary philosophical thought had their origins in Immanuel Kant's thought. Kant had established the dogma that all knowledge must be of the type provided by the natural sciences. But, Windelband held, if knowledge is limited to only that which can be contained within the categories as set forth in the *Critique of Pure Reason,* then the kinds of activities associated with the will and the emotions—that is to say, the subjects of Kant's second and third critiques—are removed from the province of knowledge. The inadequacies of the Kantian identification of knowledge in general with natural scientific knowledge alone had been demonstrated by the post-Kantian idealists, who sought to construct a theory of knowledge capable of appreciating "the needs of modern culture, and … the historical material of ideas" (*History of Philosophy,* p. 569). Idealism failed, however, because it ended by hypostatizing a spiritual sphere that presumably was separate from the world of matter and that operated according to principles utterly different from those which science explicated in general causal laws. Thus, whereas Kantianism had failed to include ethics and aesthetics within the domain of scientific philosophy, idealism failed to provide a place for those aspects of the world revealed by the natural sciences and eternally established as causally determined. It thus appeared to late nineteenth-century thinkers that there were at least two levels of reality, one spiritual and historical, the other material and determined; and it seemed that knowledge itself, far from being one, was at least twofold. On the one hand, it was empirical and discovered laws; on the other hand, it was rational and revealed the essential freedom behind the laws. Such at least had been the contention of Wilhelm Dilthey and the neo-ide-

alists. As long as this division persisted, Windelband held, pessimism, the denial of philosophy, must flourish also.

The way out of the difficulty was to be provided by a fundamental reappraisal in philosophy, a reconsideration of modern thought ab initio. For Windelband this meant primarily an attempt to find a way to apply the technique of transcendental deduction to the historical as well as the physical sciences. It also meant liberation from the notion that natural science was the archetype of all knowledge.

In an early address, "Was ist Philosophie?" (1882), Windelband distinguished between theoretical judgments (*Urteile*) and critical judgments (*Beurteilungen*). The former expressed the "mutual implicativeness" (*Zusammengehörigkeit*) of two "representational contents" (*Vorstellungsinhalte*); the latter expressed the relation between the judging consciousness (*beurteilenden Bewusstsein*) and the object represented (see *Präludien,* Vol. I, p. 29). Theoretical judgments are judgments of fact and are always positive; their purpose is to extend the limits of knowledge in a given science. Critical judgments, however, can be either positive or negative, and they express the position assumed by the subject when a given theoretical judgment is endowed with a status as means to some end.

The individual sciences expand the series of theoretical judgments; philosophy examines the relations between the ability of individual consciousness to render judgments and that "consciousness in general" (*Bewusstsein überhaupt*) which is the intuited basis of every critical judgment. Philosophy, then, "has its own proper field and its own problem in those values of universal validity that are the organizing principles for the functions of culture and civilization and for all the particular values of life. But it will describe and explain those values only that it may give an account of their validity; it treats them not as facts but as norms" (*History of Philosophy,* pp. 680–681). The various sciences are concerned with facts, which they organize in different ways according to the ends for which those facts are "constructed." Philosophy, however, is concerned with the processes by which events attain the status of facts for particular sciences.

Critical judgments, then, are rendered in respect not of what is but of what ought to be; in accordance not with laws but with norms. There is a "normative consciousness" (*Normalbewusstsein*) presupposed by philosophy; this "normative consciousness" is *in abstracto* the same as that which, *in concreto,* underlies every scientific, moral, and aesthetic experience. It is not to be thought of as either a metaphysical or a psychological entity. It is, rather, merely the "sum-total of the inter-connections and relations between existents" (*Logic,* p. 59). These relations "are not themselves existents, either as things, as states, or as activities; they can only become 'actual' as the content of the psychical functions of knowing.… In itself the realm of the valid is nothing else than the form and order under which that which exists is determined" (ibid.). It follows, then, that "this whole is closed to our knowledge; we shall never know more than a few fragments of it, and there is no prospect of our ever being able to patch it together out of the scraps that we can gather" (ibid., p. 65). Therefore, philosophy cannot end in science or in any practical rule of life; it can only point the attention of humanity to the sensed "principles of absolute judgment" that are presupposed in every human confrontation of the world in scientific, moral, and aesthetic experience.

Windelband regarded as baseless every attempt to distinguish between the different disciplines that constitute science on the basis of a presumed essential difference between their objects. The disciplines are distinguished only by their methods, which are in turn functions of the ends or values informing them as instruments of culture. In the address "Geschichte und Naturwissenschaften" (1894), he distinguished between the natural sciences and the historical sciences, and he argued that the natural sciences aim at the construction of general laws and "explain" an event by identifying it as an instance of a general law. Historical sciences, on the other hand, are individualizing; they concentrate on specific events and attempt to determine their specific physiognomy or form. Natural science Windelband termed *nomothetic*; historical science, *idiographic.* But, he added, any given object could be studied by both kinds of science. A mental event, if viewed under the aspect of physical causality—as an instance of the working of some general law—was a natural event. That same mental event, described in its individuality and valued for its deviation from the class to which it belonged, became an object of the idiographic sciences. Positivists erred in holding that every event must be viewed nomothetically, just as idealists erred in thinking that certain kinds of events cannot be so viewed. The total picture of the world that consciousness is in principle able to construct can be constructed only through the use of both kinds of investigation. No single event can be deduced from general laws, and no law can be framed out of the contemplation of a single event. "Law and event remain together as the ultimate, incommensurable limits of our representation of the world" (*Präludien,* Vol. II, p. 160).

See also Consciousness; Dilthey, Wilhelm; Fischer, Kuno; History and Historiography of Philosophy; Idealism; Kant, Immanuel; Lotze, Rudolf Hermann; Neo-Kantianism.

Bibliography

WORKS BY WINDELBAND

Die Lehre vom Zufall. Berlin, 1870.

Präludien: Aufsätze und Reden zur Einführung in die Philosophie, 2 vols. Freiburg im Breisgau, 1884; 5th ed., Tübingen, 1914.

Lehrbuch der Geschichte der Philosophic. Tübingen, 1892; 14th ed., revised by Heinz Heimsoeth, Tübingen, 1948. Translated by J. H. Tufts as *History of Philosophy.* New York: Macmillan, 1893; 2nd ed., New York: Macmillan, 1901.

"Die Prinzipien der Logik." In *Enzyklopädie der philosophischen Wissenschaften,* by Wilhelm Windelband and Arnold Ruge. Tübingen, 1912. Translated by B. E. Meyer as *Logic.* London: Macmillan, 1913.

"Geschichtsphilosophie: Eine Kriegsvorlesung, Fragment aus dem Nachlass." Edited by Wolfgang Windelband and Bruno Bauch. *Kantstudien, Ergänzungshefte im Auftrag der Kantgesellschaft* (38) (1916): 5–68.

WORKS ON WINDELBAND

Collingwood, R. G. *The Idea of History,*165–168. Oxford: Clarendon Press, 1946.

Gronau, G. "Die Kultur und Wertphilosophie Wilhelm Windelbands." In *Die Philosophie der Gegenwart.* Langensalza, 1922.

Rickert, Heinrich. *Wilhelm Windelband.* Tübingen, 1915.

Rossi, Pietro. *Lo storicismo tedesco contemporaneo,*149–207. Turin: Einaudi, 1956.

Hayden V. White (1967)

WISDOM

"Wisdom" in its broadest and commonest sense denotes sound and serene judgment regarding the conduct of life. It may be accompanied by a broad range of knowledge, by intellectual acuteness, and by speculative depth, but it is not to be identified with any of these and may appear in their absence. It involves intellectual grasp or insight, but it is concerned not so much with the ascertainment of fact or the elaboration of theories as with the means and ends of practical life.

WISDOM LITERATURE

Concern with the art of living long preceded formal science or philosophy in human history. All ancient civilizations seem to have accumulated wisdom literatures, consisting largely of proverbs handed down from father to son as the crystallized results of experience. Perhaps the most ancient known collection of these sayings is the Egyptian "Wisdom of Ptah-hotep," which comes down from about 2500 BCE. The writings of Confucius (sixth century BCE) and Mencius (fourth century BCE), though more sophisticated, are still concerned chiefly with the Dao, the good or normal human life. The early writers of India held views at once more speculative and more disillusioned than those of China; both Buddhists and Hindus found the greatest happiness of man in deliverance from the grinding round of suffering and death and in absorption into ātman or nirvāṇa, where personality and struggle alike disappear. But large parts of the Bhagavad-Gita and the Dhammapada, two classics among the scriptures of India, are devoted to maxims and counsels for the conduct of life.

Of far greater influence in the West has been the wisdom literature of the Hebrew people, which consists of the more philosophical parts of the Old Testament and the Apocrypha. Perhaps the most important of these are the books of Job, Proverbs, and Psalms and the apocryphal book called The Wisdom of Solomon. There is no certain knowledge of who wrote any of them; they are probably the work of many men, extending over centuries. They differ strikingly from the writings of Greek and Chinese moralists in the closeness with which morality is identified with religion. The Hebrew sages were all monotheists who held that God fashioned the world but remained outside it; he had made his will known in the law delivered to Moses. This law set the standard and pattern of goodness for all time; the good man will make it his study and seek to conform his life to it. At the same time these sages reduced the miraculous element in Jewish history; they made no claim to being inspired themselves, and inclining, indeed, to assume that the sole motive of conduct was self-advantage, they offered their prudential maxims as not only conforming to the divine law but as also the product of good sense and sound reason. There is very little evidence that they were affected by Greek thought, though Greek influence must have flowed around them after the conquests of Alexander. It is possible that in their cool and reasonable note, contrasting so sharply with the visionary fervor of the prophets, there is an echo of the reflective thought of Greece.

The Greeks had a wisdom literature of their own that long preceded the appearance of their great philosophers. Hesiod (eighth century BCE) and Theognis (sixth century BCE) summed up in poetic form the maxims of traditional morality. Pythagoras (sixth century BCE), a curious combination of mathematician and religious seer, seems to have found in philosophy the guide of prac-

tical life. This view was further developed by the Sophists, who, at a time when libraries and universities were unknown, undertook to instruct young men in the arts, theoretical and practical, that were most likely to lead to success. In their emphasis on success, however, there was something skeptical and cynical; the art of life tended in their teaching to become the sort of craft that enabled one by clever strategy to achieve place and power.

THE GREEK CONCEPTION

The first full statement and embodiment of the classic Greek conception of wisdom came with Socrates (c. 470–399 BCE), who insisted that virtue and knowledge were one, that if men failed to live well, it was through ignorance of what virtue really was. He had no doubt that if men knew what virtue was, they would embody it in their conduct. Thus, he set himself to define the major virtues with precision. His method was to consider particular instances of them and bring to light the features they had in common; this would give the essence and true pattern of the virtue in question. He did not profess to be satisfied with the results of his inquiries, but his acuteness and thoroughness made him the first of the great theoretical moralists, and the courage with which he carried his principles into both life and death gave him a unique place in Western history.

The stress on wisdom was maintained by his disciple Plato. For Plato there are three departments of human nature, which may be described as the appetites, directed to such ends as food and drink; the distinctively human emotions, such as courage and honor; and reason. Of these reason is the most important, for only as impulse and feeling are governed by it will conduct be saved from chaos and excess; indeed, in such government practical wisdom consists. In one respect Aristotle carried the exaltation of reason further than Plato; in addition to this practical wisdom, he recognized another and purely intellectual virtue, the wisdom that pursues truth for its own sake and without reference to practice. In this pursuit, which can be followed effectively only by the philosopher, lay the highest and happiest life.

It was among the Stoics, however, that guidance by reason was most seriously and widely attempted. In the thought of the Roman emperor Marcus Aurelius (121–180 CE), both nature and human nature are determined by causal law, and the wrongs and insults that other men inflict on us are therefore as inevitable as the tides. The wise man will understand this inevitability and not waste his substance in futile indignation or fear. He will conform himself to nature's laws, recognize that passion is a symptom of ignorance, free himself from emotional attachments and resentments, and live as far as he can the life of a "passionless sage." The account given by Marcus Aurelius in his famous journal of his struggle to order his practice and temper by this ideal of austere rationality has made his little book a classic of pagan wisdom.

MODERN PHILOSOPHERS

The opinions of modern philosophers on the meaning of wisdom are too various for review here. But it can be noted of these thinkers, as it was of Marcus Aurelius, that their standing as purveyors or exemplars of wisdom bears no fixed relation to their eminence as philosophers. If their chief work lies, as Immanuel Kant's does, in the theory of knowledge, or as John McTaggart Ellis McTaggart's does, in technical metaphysics, it may have no obvious bearing on practical life. Furthermore, by reason of an unhappy temperament, some philosophers of name and influence, such as Jean-Jacques Rousseau, have been far from notable exemplars of wisdom in either controversy or conduct. On the other hand, there are thinkers who have shown in their writing, and sometimes also in their lives, so large a humanity and good sense that they have been held in especial esteem for their wisdom whether or not they have been of high philosophical rank. Michel Eyquem de Montaigne and Ralph Waldo Emerson are examples on one level; John Locke, Bishop Butler, John Stuart Mill, and Henry Sidgwick are examples from a more professional level. Among technical thinkers of the first rank, a figure who has left a deep impression for a wisdom serene and disinterested, though a little above the battle, is the famous philosopher of Amsterdam, Benedict de Spinoza (1632–1677).

COMPONENTS OF WISDOM

Are there any traits uniformly exhibited by the very diverse minds that by general agreement are wise? Two traits appear to stand out—reflectiveness and judgment.

REFLECTIVENESS. By reflectiveness is meant the habit of considering events and beliefs in the light of their grounds and consequences. Conduct prompted merely by impulse or desire is notoriously likely to be misguided, and this holds true of both intellectual and practical conduct. Whether a belief is warranted must be decided by the evidence it rests on and the implications to which it leads, and one can become aware of these only by reflection. Similarly, whether an action is right or wrong depends, at least in part, on the results that it produces in

the way of good and evil, and these results can be taken into account only by one who looks before he leaps. Common sense, with its rules and proverbs, no doubt helps, but it is too rough and general a guide to be relied on safely; and the reflective man will have at his command a broader view of grounds and consequences, causes and effects. He will more readily recognize the beliefs of superstition, charlatanism, and bigotry for what they are because he will question the evidence for them and note that when reflectively developed, they conflict with beliefs known to be true. In the same way he will be able to recognize some proposals for action as rash, partisan, or shortsighted because certain consequences have been ascribed to them falsely and others have been ignored. In some activities wisdom consists almost wholly of such foresight. A general, for example, is accounted wise if he can foresee in detail how each of the courses open to him will affect the prospects of victory.

JUDGMENT. There is a wisdom of ends as well as of means, which is here denoted by "judgment." The goal of the general—namely, victory—is laid down for him, but the ordinary man needs the sort of wisdom that can appraise and choose his own ends. The highest wisdom of all, Plato contended, is that required by the statesman, who is called upon to fix both the goals toward which society strives and the complex methods by which it may most effectively move toward them. Unfortunately, at this crucial point where the ends of life are at issue, the sages have differed profoundly. Some, like Epicurus and Mill, have argued for happiness; others, like the Christian saints, for self-sacrificing love; others, such as Friedrich Nietzsche, for power. Many philosophers of the twentieth century came to hold that this conflict is beyond settlement by reason, on the ground that judgments of good and bad are not expressions of knowledge at all but only of desire and emotion. For these thinkers there is properly no such thing as wisdom regarding intrinsic goods; knowledge is confined to means.

Whatever the future of this view, common opinion is still at one with the main tradition of philosophy; it regards the judgment of values as a field in which wisdom may be preeminently displayed. It must admit, however, that this judgment is of a peculiar kind; it seems to be intuitive in the sense that it is not arrived at by argument nor easily defended by it. One may be certain that pleasure is better than pain and yet be at a loss to prove it; the insight seems to be immediate. And where immediate insights differ, as they sometimes do, the difference appears to be ultimate and beyond remedy. Must such wisdom end in dogmatic contradiction and skepticism?

That it need not do so will perhaps be evident from a few further considerations. First, differences about intrinsic goods may be due to mere lack of knowledge on one side or the other. The Puritans who condemned music and drama as worthless could hardly have excluded them if they had known what they were excluding; in these matters wider experience brings an amended judgment. Second, what appears to be intuitive insight may express nothing more than a confirmed habit or prejudice. Where deep-seated feelings are involved, as in matters of sex, race, or religion, the certainty that belongs to clear insight may be confused with the wholly different certainty of mere confidence or emotional conviction. Fortunately, Sigmund Freud and others have shown that these irrational factors can be tracked down and largely neutralized. Third, man's major goods are rooted in his major needs, and since the basic needs of human nature are everywhere the same, the basic goods are also the same. No philosophy of life that denied value to the satisfactions of food or drink or sex or friendship or knowledge could hope to commend itself in the long run.

It should be pointed out, finally, that the judgment of the wise man may carry a weight out of all proportion to that of anything explicit in his thought or argument. The decisions of a wise judge may be implicitly freighted with experience and reflection, even though neither may be consciously employed in the case before him. Experience, even when forgotten beyond recall, leaves its deposit, and where this is the deposit of long trial and error, of much reflection, and of wide exposure in fact or imagination to the human lot, the judgment based on it may be more significant than any or all of the reasons that the judge could adduce for it. This is why age is credited with wisdom; years supply a means to it whether or not the means is consciously used. Again, the individual may similarly profit from the increasing age of the race; since knowledge is cumulative, he can stand on the shoulders of his predecessors. Whether individual wisdom is on the average increasing is debatable, but clearly the opportunity for it is. As Francis Bacon, a philosopher whose wisdom was of the highest repute, remarked, "We are the true ancients."

See also Bacon, Francis; Butler, Joseph; Confucius; Emerson, Ralph Waldo; Epicurus; Freud, Sigmund; Locke, John; Marcus Aurelius Antoninus; McTaggart, John McTaggart Ellis; Mencius; Mill, John Stuart; Montaigne, Michel Eyquem de; Nietzsche, Friedrich; Philosophy; Plato; Pythagoras and Pythagoreanism; Sidgwick, Henry; Spinoza, Benedict (Baruch) de; Socrates; Stoicism.

Bibliography

For proverbial wisdom see Archer Taylor, *The Proverb* (Cambridge, MA: Harvard University Press, 1931), and—old but suggestive—R. C. Trench, *Proverbs and Their Lessons* (London and New York, 1858).

For the problems of determining right and wrong, see any first-rate work on ethics, such as Henry Sidgwick, *The Methods of Ethics,* 7th ed. (Chicago: University of Chicago Press, 1962).

For an analysis of reflection, see, for example, John Dewey, *How We Think* (Boston: Heath, 1910).

For the place of reason in valuation, see L. T. Hobhouse, *The Rational Good* (New York: Holt, 1921), or Brand Blanshard, *Reason and Goodness* (New York: Macmillan, 1961).

For some useful popular works see T. E. Jessop, *Reasonable Living* (London, 1948); H. C. King, *Rational Living* (New York, 1912); and A. E. Murphy, *The Uses of Reason* (New York: Macmillan, 1943).

Brand Blanshard (1967)

WISDOM, (ARTHUR) JOHN TERENCE DIBBEN

(1904–1993)

(Arthur) John Terence Dibben Wisdom, the British analytic philosopher, was closely associated with Ludwig Wittgenstein, whose chair in philosophy at Cambridge he held. Wisdom became professor of philosophy there in 1952. He took his B.A. degree at Cambridge in 1924 and his M.A. there in 1934.

The philosophical problem on which Wisdom wrote the most is the question of what the nature of philosophy is, and his writings reflect his changing views concerning the proper answer to this question. His writings can be divided into two groups: those through 1934, putting forward one answer to the question, and those after 1936, consisting of successive attempts to make clear a quite different view of the nature of philosophy, along with applications of this new approach to a number of familiar first-level philosophical problems.

LOGICAL CONSTRUCTIONS

Wisdom's first book, *Interpretation and Analysis* (1931), compares Jeremy Bentham's notion of a "fiction" with Bertrand Russell's idea of a logical construction—a central notion of British philosophizing in the 1920s and 1930s. According to the theory of logical constructions, to say that a kind of entity *X* is a logical construction out of entities of kind *Y* is to say that statements about entities of kind *X* are translatable into statements about entities of kind *Y,* the *Y*'s being "more ultimate," "more fundamental," than the *X*'s. (It was often said to be less misleading to say, not "*X*'s are logical constructions," but "'*X*' is an incomplete symbol.") Thus, for example, it was said that nations, which are, after all, a kind of "abstraction," are logical constructions out of their nationals, and this meant that statements about, for example, England and France are translatable into statements about Englishmen and Frenchmen. The translation was to be performed not merely by replacement of the words—for "England is a monarchy" does not mean the same as "Englishmen are a monarchy"—but also by changing the predicates, and no doubt the new predicates would be more complicated. Nevertheless, a fact about England is not something "over and above" a fact or set of facts about Englishmen. And other things, too, were said to be logical constructions: propositions were said to be logical constructions out of sentences, people out of mental and bodily events, material objects (including human bodies) out of sense data, and so on. Indeed, Russell and others used the notion very widely; Ockham's razor (according to which "entities must not be multiplied beyond necessity") was given the modern form: supposedly transcendent or abstract entities are everywhere to be regarded as logical constructions out of the more concrete entities given in sense experience. This procedure has the advantage of explicitly blocking a mistaken inference that may arise, for example, from George Berkeley's analysis of a material object as a "congeries of ideas" (for "ideas," read "sense data"). Analyzing it in this way suggests, for example, that the apple I hold in my hand is made of sense data and that I would be eating sense data if I ate the apple. But to say that the apple is a *logical* construction out of sense data is only to say that statements about it are translatable into statements about sense data.

G. E. Moore had written (in "A Defense of Common Sense") that the work of the philosopher was not to find out whether this or that (supposed) matter of fact really was a fact but rather to find the analysis of what we know in knowing the things we do unquestionably know. Thus, I know for certain that I have two hands, but what is the analysis of what I know in knowing this? The followers of Russell and the early Wittgenstein ("logical atomists," as they have been called) saw their task as the analysis of such statements into "atomic statements," which are logically and epistemologically fundamental; they sought to provide translations of statements containing the expression "*X*" into statements that do not contain "*X,*" thus justifying the claim that *X*'s are logical constructions.

The first exhaustive treatment of this central notion is to be found in Wisdom's series of five articles titled

"Logical Constructions," which appeared in successive issues of *Mind* from 1931 to 1933. The first three of these essays discuss the relation between sentences in general and the facts expressed by them; the governing idea comes from Wittgenstein's *Tractatus,* where a sentence (on Wisdom's interpretation) is said to be a picture of the fact it expresses. Wisdom tries to bring out precisely what this comes to, in the case not only of such "simple" sentences as "Wisdom killed Al Capone" but also of negations, generalizations, and compound sentences. The fourth and fifth essays are concerned more specifically with logical constructions: How precisely is the analysandum (for example, a statement about sense data) related on the one hand to the fact it pictures and on the other hand to the analysans (a statement about an external object) and the fact it pictures?

In the last of the five essays philosophy is identified with analysis, which is said to provide the required translations. Philosophical propositions are thus verbal (that is, about words), differing only in aim or intention from those of writers of dictionaries: "The philosophical intention is clearer insight into the ultimate structure" of facts, and "philosophic progress does not consist in acquiring knowledge of new facts but in acquiring new knowledge of facts."

The essays "Ostentation" (1933) and "Is Analysis a Useful Method in Philosophy?" (1934) also deal with logical constructions.

THE NEW APPROACH TO PHILOSOPHY

Wittgenstein, who had been away from Cambridge since before World War I, returned there in 1929; his writings from then on show a gradual change in his conception of the nature of philosophy and of language. Wisdom himself returned to Cambridge in 1934 (he had for some years been teaching philosophy at St. Andrews University in Scotland), and his thinking was then strongly influenced by the new view of philosophy being worked out by Wittgenstein. Wisdom's essay "Philosophical Perplexity" (1936) shows that by 1936 a striking change had taken place,

No doubt many within the analytic movement had felt uneasiness about its program, and there had been criticism of the movement from its beginnings, but this was the first appearance in print of an alternative to the earlier reductive account of what philosophers are and ought to be doing. (Wittgenstein's writings of the period were not published until much later, after his death.)

According to the new conception of philosophy (set out briefly in "Philosophical Perplexity" and in greater detail in "Metaphysics and Verification," 1938), philosophical claims are answers to questions of the forms "What are *X*'s?," "What is it to know that here is an *X*?," "Are there any *X*'s?," "Is there any such thing as knowing that here there is an *X*?," where "*X*" is replaced by some very general term such as "material object," "soul," or "causal connection." Answers to the first pair of questions are of two and only two forms: the reductive (*X*'s are logical constructions out of *Y*'s; knowledge that here is an *X* is really knowledge about *Y*'s), and the transcendentalist (*X*'s are unanalyzable, are ultimate; knowledge that here is an *X* is unique, a special way of knowing appropriate only to *X*'s). A philosopher's answers to the second pair of questions will be connected with his answers to the first pair—for example, a reductionist is less likely to be a skeptic (although some have been both reductionists and skeptics with respect to, say, material objects), whereas a transcendentalist is more likely to fall into skepticism.

In view of their form, answers to the first pair of questions are apt to appear to be strictly definitional (as when one says "Fathers are male parents"), and answers to the second pair may appear to be making straightforward empirical points (as when one says what goes on inside Earth). But the philosopher does neither of these things. A philosophical question arises out of a dissatisfaction with the "categories of being" (in the formal mode, "kinds of statement") implicit in our ordinary way of talking. Reductive answers to the first pair of questions and skeptical answers to the second pair are disguised proposals of alternative categorizations; transcendentalist answers to the first pair of questions and nonskeptical answers to the second pair are disguised proposals that we retain the categorizations already marked in the language. The various answers all bring home to us the likenesses and differences between "categories of being" that are either concealed by or implicit in our ordinary way of talking.

Consider, for example, a certain kind of skepticism about material objects. The skeptic says, "We don't really know that there is cheese on the table" and "It would be well if we prefixed every remark about material things with 'probably.'" Such skepticism draws our attention to a likeness shared by all statements about material objects and to a difference between all such statements on the one hand and statements about sensations on the other. The skeptic forces us to see that if a man makes a statement about a material object—whatever the object, whatever the circumstances—then it always makes sense

for us to say "But perhaps he is mistaken"; whereas if he says he is having this or that sensation or sense experience, it would not make sense to say this of him. Ordinary language conceals this, for we ordinarily mark a difference among material-object statements; we say that some are at best probable (such as reports about what is going on inside Earth) and that others (such as reports about what is going on inside our fists) are as certain as any statement about a sensation or experience. Of course the job remains of showing why it strikes the skeptic—and us—as important to mark what is pointed to in his claim.

Consider the reductionist view of material objects (see "Metaphysics and Verification"). The reductionist says, "Material objects are logical constructions out of sense data." He draws our attention to a likeness between material-object statements and a certain kind of statement about sense data, a likeness in their mode of verification; if you have already found out that this has, does, and will continue to appear to be (say) a bit of cheese, then there is nothing further to do in the way of finding out whether or not it is a bit of cheese. Ordinary language conceals this likeness, for our ordinary use of the words is such that it is simply false to say that "This is a bit of cheese" means the same as "This has, does, and will appear to be a bit of cheese." Or, as it might be put, the reductionist draws our attention to a likeness between the statement "A material-object statement means the same as a certain complex sense-datum statement" and ordinary statements of the form "'*X*' means the same as '*Y*'" that we would unhesitatingly accept as true; and a difference between it and many ordinary statements of the form "'*X*' means the same as '*Y*'" that we would unhesitatingly reject as false.

Whether a philosophical claim is true is not the important question; what we should do with respect to a philosophical question about the nature of *X*'s and our knowledge of *X*'s is to bring out in full all the features of *X*'s that incline one to opt for this or that philosophical answer—thereby bringing out the relevant likenesses and differences between *X*'s (or statements about *X*'s) and other kinds of entities (or kinds of statements). In this way we obtain that illumination of the category of *X*'s which alone can answer the dissatisfaction that was expressed in our philosophical question.

Any account of the nature of a philosophical claim is itself a philosophical claim (for example, an answer to the question "What are philosophical claims?") and is itself to be dealt with in this way. In the essays already mentioned Wisdom also tries to bring out the likenesses and differences between philosophical claims and other kinds of claims that have been stressed by those who supposed that philosophical claims tell us facts about the world and by those who said that these claims are merely verbal.

"OTHER MINDS"

The papers mentioned so far are primarily concerned with expounding Wisdom's new view of the nature of philosophy, and the first-level philosophical claims considered there appear for the most part as examples; by contrast, his series of papers titled "Other Minds" (which appeared in successive issues of *Mind* between 1940 and 1943) is concerned mainly with the first-level questions relating to our knowledge of other minds, and the second-level question on the nature of philosophy is discussed largely in order to shed light on the first-level questions. His aim in these papers is to bring out all the problems that issue in the question "Do we ever know what anyone else is thinking, feeling, experiencing …?" and to give them the sort of treatment he has said a philosophical problem calls for. Roughly, papers I and II bring out the likenesses and differences between statements about other minds and statements about invisible currents flowing through wires; III compares the philosopher's and the plain man's use of "It's at best probable" and "We know by analogy"; IV and V deal with telepathy and extra or extended ways of knowing in general; VI and VII show what considerations rule out the possibility that one should have "direct" knowledge of the sensations of others—that is, knowledge of the kind one has of one's own sensations (this is done by showing what makes a statement be a statement that is not merely about one's own sensations); and VIII deals with the status of the statement "No one has any knowledge at all apart from knowledge as to his own sensations of the moment."

The difference in conception of the nature of philosophy between Wisdom's later work and, for example, the "Logical Constructions" papers has often been discussed. It is therefore worth mentioning that there is also considerable continuity. As previously noted, Wisdom had earlier thought of "the philosophical intention [as] clearer insight into the ultimate structure" of facts; in "Philosophical Perplexity" he still regarded it as a search for "illumination of the ultimate structure of facts." He did not, in this paper of 1936 or in any of his later works, regard philosophy as merely the study either of the workings of language for its own sake or of the confusions of ordinary language. The analogy he later drew between philosophy and psychoanalysis led many people to think he regarded philosophy as strictly a kind of therapy. But this was never his view, and indeed one may regard his

successive efforts to characterize the philosophical enterprise as attempts to bring out just what sort of insight and understanding the philosopher does provide (see, for example, "Gods" and "Philosophy, Metaphysics and Psychoanalysis").

WISDOM AND WITTGENSTEIN

It is dangerous to talk about the conception of philosophy held by the later Wittgenstein—there are very few remarks on the nature of philosophy in Wittgenstein's posthumously published *Philosophical Investigations,* and those he does make are obscure. Nevertheless, Wittgenstein's manner of dealing with philosophical problems there suggests that Wisdom differs from him at least in his attitude toward philosophy. While Wisdom always acknowledged his great debt to Wittgenstein, he says of him in "Philosophical Perplexity," "He too much represents [philosophical theories] as merely symptoms of linguistic confusion. I wish to represent them as also symptoms of linguistic penetration." And he reminds us repeatedly that we are not to take his work as representing Wittgenstein's own views.

In sum, Wisdom's view is that the goal of philosophy is an understanding of just what philosophers have at all times sought to understand—"time and space, good and evil, things and persons." In making their case, philosophers have always appealed to linguistic usage—in "The Metamorphosis of Metaphysics" (reprinted in *Paradox and Discovery*) Wisdom brings out the similarity between contemporary linguistic philosophy and older forms of speculative philosophy. But he also reminds us that good philosophy of any age gives us a clearer view not merely of how we may go wrong in our talking and thinking but of how we may go right.

See also Analysis, Philosophical; Bentham, Jeremy; Berkeley, George; Logic, History of; Moore, George Edward; Other Minds; Russell, Bertrand Arthur William; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY WISDOM

Interpretation and Analysis. London: K. Paul, Trench, Trubner, 1931.

"Logical Constructions." *Mind* 40–42 (1931–1933).

Problems of Mind and Matter. Cambridge, U.K.: Cambridge University Press, 1934. A work of the same period as the "Logical Constructions" papers; it is concerned with perception and with the relation between a man's body and mind in virtue of which it is true to say that that body and mind are his body and mind.

Other Minds. Oxford: Blackwell, 1952. Contains the eight papers that originally appeared in *Mind,* 1940–1943, as well as a symposium contribution titled "Other Minds" that was originally published in *PAS* supp. 20 (1946) and "The Concept of Mind" and "Metaphysics," the presidential address to the meeting of the Aristotelian Society, November 1950.

Philosophy and Psychoanalysis. Oxford: Blackwell, 1953. Contains "Ostentation," "Is Analysis a Useful Method in Philosophy?," "Philosophical Perplexity," "Metaphysics and Verification," "Gods," "Philosophy, Metaphysics and Psychoanalysis," and other articles and reviews.

Paradox and Discovery. Oxford: Blackwell, 1965. Contains "The Metamorphosis of Metaphysics," "A Feature of Wittgenstein's Technique," "The Logic of God," "Paradox and Discovery," and other short pieces.

WORKS ON WISDOM

Gasking, D. A. T. "The Philosophy of John Wisdom, I and II." *Australasian Journal of Philosophy* (1954).

Passmore, John. *A Hundred Years of Philosophy,* 367–368, 434–438. London: Duckworth, 1957.

Urmson, J. O. *Philosophical Analysis,* 76–85, 169–182. Oxford: Clarendon Press, 1956.

Judith Jarvis Thomson (1967)

WITTGENSTEIN, LUDWIG JOSEF JOHANN
(1889–1951)

Ludwig Josef Johann Wittgenstein, the Austrian-British philosopher, was born in Vienna, the youngest of eight children. Ludwig's paternal grandfather, a convert from Judaism to Protestantism, had been a wool merchant in Saxony before moving to Vienna. Ludwig's father, Karl Wittgenstein, had, as a strong-willed boy, rebelled against a classical education, running away to America when he was seventeen. After two years he returned to Vienna and underwent a brief training in engineering. He went to work as a draftsman, designed and largely directed the construction of a steel-rolling mill, became its manager, in ten years' time was the head of a large steel company, and subsequently organized the first cartel of the Austrian steel industry. Ludwig's mother was the daughter of a Viennese banker. She was a Roman Catholic, and Ludwig was baptized in the Catholic Church. Ludwig had four brothers and three sisters; all the children were generously endowed with artistic and intellectual talent. Their mother was devoted to music, and their home became a center of musical life. Johannes Brahms was a frequent

visitor and a close friend of the family. One of Ludwig's brothers, Paul, became a distinguished pianist.

Ludwig was educated at home until he was fourteen. He was an indifferent student, and apparently his greatest interest was in machinery; a sewing machine that he constructed was much admired. His parents decided to send him to a school at Linz, in Upper Austria, that provided preparation in mathematics and the physical sciences rather than a classical education. After three years at Linz, Wittgenstein studied mechanical engineering for two years at the Technische Hochschule at Charlottenburg, in Berlin. He left this school in the spring of 1908 and went to England. In the summer of 1908 he experimented with kites at a kite-flying station in Derbyshire. That fall he registered as a research student of engineering at the University of Manchester. He engaged in aeronautical research for three years and designed a jet-reaction engine and a propeller.

Wittgenstein's interest began to shift to pure mathematics and then to the philosophical foundations of mathematics. He chanced upon Bertrand Russell's *Principles of Mathematics* and was greatly excited by it. He decided to give up engineering and to study with Russell at Cambridge. At the beginning of 1912 he was admitted to Trinity College, where he remained for the three terms of 1912 and the first two terms of 1913. Under Russell's supervision he applied himself intensively to logical studies and made astonishing progress. Soon he was engaged in the research that culminated in the logical ideas of the *Tractatus.*

Wittgenstein's most intimate friend during those early years at Cambridge was David Pinsent, a fellow student, to whom he later dedicated the *Tractatus.* When they met in the spring of 1912, Wittgenstein, in addition to studying logic, was doing experiments in the psychological laboratory on rhythm in music. He and Pinsent were united by strong musical interests. They had a repertoire of forty of Franz Schubert's songs, whose melodies Wittgenstein would whistle while Pinsent accompanied him on the piano. Wittgenstein could play the clarinet and had an excellent memory for music and an unusual gift for sight-reading. He retained a deep interest in music throughout his life; in his philosophical writings there are many allusions to the nature of musical understanding.

In 1912, Wittgenstein was doing his first extensive reading in philosophy, and according to Pinsent he expressed "naive surprise" that the philosophers whom he had "worshipped in ignorance" were after all "stupid and dishonest and make disgusting mistakes!" He and Pinsent made holiday junkets to Iceland and Norway, Wittgenstein paying all expenses. Pinsent found Wittgenstein a difficult companion: irritable, nervously sensitive, often depressed. But when he was cheerful he was extremely charming. Sometimes he was depressed by the conviction that his death was near at hand and that he would not have time to perfect his new ideas in logic, sometimes by the thought that perhaps his logical work was of no real value. Even so, his general frame of mind was less morbid than before he had come to Cambridge. For a number of years previously there had hardly been a day, he told Pinsent, in which he had not thought of suicide "as a possibility." Coming to study philosophy with Russell had been his "salvation."

Wittgenstein worked with fierce energy at his logical ideas. In the spring of 1913 he submitted to hypnosis with the hope that in the hypnotic trance he could give clear answers to questions about difficulties in logic. He entertained a plan of going to live in seclusion in Norway for some years, devoting himself to logical problems. The reasons he gave to Pinsent were that he could do better work in the absence of all distractions, but he also said that "he had no right to live in a world" where he constantly felt contempt for other people and irritated them by his nervous temperament. Wittgenstein acted on his plan and lived in Norway from the latter part of 1913 until the outbreak of World War I. He stayed on a farm at Skjolden and later built a hut, where he lived in complete seclusion.

During this period Wittgenstein corresponded with Russell. His letters were warmly affectionate and were full of the excitement of his logical discoveries. However, he expressed the conviction that he and Russell had such different "ideals" that they were not suited for true friendship. Two people can be friends, he said, only if both of them are "pure," so that they can be completely open with one another without causing offense. A relationship founded on "hypocrisy" is intolerable. He and Russell should break off entirely or else limit their communications to their logical work. Both of them have weaknesses, but especially himself: "My life is *full* of the most hateful and petty thoughts and acts (this is *no* exaggeration)." "Perhaps you think it is a waste of time for me to think about myself; but how can I be a logician if I am not yet a man! *Before everything else* I must become pure."

When war broke out Wittgenstein entered the Austrian Army as a volunteer. He served in an artillery group on a vessel on the Vistula and later in an artillery workshop at Kraków. He was ordered to an officers' training school and subsequently served on the eastern front and later with mountain artillery in the southern Tyrol. Dur-

ing these years he continued to work at his book, writing down his philosophical thoughts in notebooks that he carried in his rucksack. He completed the book in August 1918; when he was taken prisoner by the Italians in November, he had the manuscript with him. From his prison camp near Monte Cassino he wrote to Russell, to whom the manuscript was subsequently delivered by diplomatic courier through the offices of a mutual friend, J. M. Keynes.

While serving on the eastern front Wittgenstein bought at a bookshop in Galicia a copy of one of Lev Tolstoy's works on the Gospels, which apparently made a deep impression on him. In the prison camp in Italy he read a standard version of the Gospels, possibly for the first time, and is reported to have been disturbed by much that he found in it and to have questioned its authenticity, perhaps because of the differences from Tolstoy's version.

Wittgenstein was anxious to have his book, *Logisch-Philosophische Abhandlung,* published immediately. Shortly after his release from imprisonment and his return to Vienna, in August 1919, he offered it to a publisher. He believed that his book finally solved the problems with which he and Russell had struggled. From Russell's letters, however, he concluded that Russell had not understood his main ideas, and he feared that no one would. He and Russell met in Holland in December 1919 to discuss the book. Russell undertook to write an introduction for it, but the following May, Wittgenstein wrote to Russell that the introduction contained much misunderstanding and he could not let it be printed with his book. Subsequently the publisher with whom he had been negotiating rejected the book. Wittgenstein wrote to Russell, in July 1920, that he would take no further steps to have it published and that Russell could do with it as he wished. The German text was published in 1921 in Wilhelm Ostwald's *Annalen der Naturphilosophie.* The following year it was published in London with a parallel English translation, under the title *Tractatus Logico-Philosophicus.* A new and improved English translation was published in 1961.

Most of the notebooks used in the preparation of the *Tractatus* were destroyed on Wittgenstein's order. Three of them, however, from the years 1914–1916, were accidentally preserved and were published in 1961 with a parallel English translation. The notebooks present a vivid picture of the intensity of Wittgenstein's struggles with the problems of the *Tractatus,* and they sometimes help to show what the problems were.

Soon after his return to civilian life Wittgenstein decided to become a schoolteacher. He attended a teacher-training course in order to receive a certificate, and in the fall of 1920 he began teaching classes of children aged nine and ten in the village of Trattenbach in Lower Austria. He was an exacting teacher. He did not get on with his colleagues and was often depressed. When he was transferred to another village he was somewhat happier, for one of the teachers, Rudolf Koder, was a talented pianist. The two of them devoted many afternoons to music, Wittgenstein playing the clarinet or whistling. He remained a schoolteacher until 1926. In 1924 he prepared a dictionary of six thousand to seven thousand words for the use of pupils in the elementary schools of the Austrian villages; this small book was published in 1926.

When his father died, in 1913, Wittgenstein inherited a large fortune. In the summer of the following year he wrote to Ludwig von Ficker, editor of the literary review *Der Brenner,* proposing to send a large sum of money to be distributed among needy Austrian poets and artists. The poets Rainer Maria Rilke and Georg Trakl received sizable gifts of money from this anonymous source. Upon his return to civilian life after the war, Wittgenstein gave his fortune to two of his sisters. Part of the reason for this action was that he did not want to have friends for the sake of his money, but undoubtedly it was largely due to his inclination toward a simple and frugal life.

During his years as a teacher, until Frank Ramsey visited him in 1923, Wittgenstein probably gave no thought to philosophy. Ramsey, a brilliant young mathematician and philosopher at Cambridge, had just completed a review of the *Tractatus* and was eager to discuss the book with its author. He found Wittgenstein living in extreme simplicity in a small village. In explaining his book, to which he was willing to devote several hours a day for a fortnight or more, Wittgenstein would become very excited. He told Ramsey, however, that he would do no further work in philosophy because his mind was "no longer flexible." He believed that no one would understand the *Tractatus* merely by reading it but that some day some person would, independently, think those same thoughts and would derive pleasure from finding their exact expression in Wittgenstein's book.

After his resignation as a schoolteacher in 1926, Wittgenstein inquired at a monastery about the possibility of entering upon monastic life, but he was discouraged by the father superior. In the summer of that year he worked as a gardener's assistant with the monks at Hütteldorf, near Vienna. Meanwhile, one of his sisters had commissioned the architect Paul Engelmann to build a

mansion for her in Vienna. Engelmann, a friend of Wittgenstein's, proposed to him that they undertake it jointly. Wittgenstein agreed and actually became the directing mind in the project, which occupied him for two years. The building has been described by G. H. von Wright as "characteristic of its creator. It is free from all decoration and marked by a severe exactitude in measure and proportion. Its beauty is of the same simple and static kind that belongs to the sentences of the *Tractatus.*" During the same period Wittgenstein did some work in sculpture.

Moritz Schlick, a professor in Vienna, had been deeply impressed by the *Tractatus.* He managed to establish contact with Wittgenstein and apparently prevailed upon him to attend one or two meetings of the group founded by Schlick, known as the Vienna circle. Subsequently Schlick and Friedrich Waismann paid visits to Wittgenstein, in which he expounded some ideas that were passed on to other members of the circle.

In January 1929 he returned to Cambridge to devote himself again to philosophy. What produced this renewal of interest is unknown, but it is said that it was provoked by a lecture he heard L. E. J. Brouwer give in Vienna in 1928 on the foundations of mathematics. Wittgenstein found he would be eligible to receive the Ph.D. degree from Cambridge if he submitted a dissertation, whereupon he submitted the *Tractatus.* Russell and G. E. Moore were appointed to give him an oral examination, which they did in June 1929. Moore found the occasion "both pleasant and amusing." Trinity College granted Wittgenstein a research fellowship. At this time he published a short paper, "Some Remarks on Logical Form," which he soon came to think was weak and confused. This paper and the *Tractatus* were the sole philosophical writings of his that were published in his lifetime.

Wittgenstein began to give lectures in January 1930. He remained at Cambridge until the summer of 1936, when he went to live for a year in his hut in Norway and to begin writing the *Philosophical Investigations.* In 1937 he returned to Cambridge and two years later succeeded Moore to the chair of philosophy.

Wittgenstein's lectures made a powerful impression on his auditors. They were given without notes or preparation. Each lecture was new philosophical work. Wittgenstein's ideas did not come easily. He carried on a visible struggle with his thoughts. At times there were long silences, during which his gaze was concentrated, his face intensely alive, and his expression stern, and his hands made arresting movements. His hearers knew that they were in the presence of extreme seriousness, absorption, and force of intellect. When he spoke his words did not come fluently, but they came with force and conviction. His face was remarkably mobile and expressive when he talked. His eyes were often fierce, and his whole personality was commanding. His lectures moved over a wide range of topics and were marked by great richness of illustration and comparison. Wittgenstein attacked philosophical problems energetically, even passionately. Unlike many other philosophers, who really want to retain the problems rather than to solve them, Wittgenstein's desire was to clear them up, to get rid of them. He exclaimed to a friend: "My father was a business man and I am a business man too!" He wanted his philosophical work to be businesslike, to settle things.

When he was not working at philosophy Wittgenstein could sometimes, with a friend, put on a charming mood of mock seriousness in which he said nonsensical things with utmost gravity. These lighthearted moments were, however, comparatively infrequent. Most commonly his thoughts were somber. He was dismayed by the insincerity, vanity, and coldness of the human heart. He was always troubled about his own life and was often close to despair. Human kindness and human concern were for him more important attributes in a person than intellectual power or cultivated taste. He had an acute need for friendship, and his generosity as a friend was striking. At the same time it was not easy to maintain a friendly relationship with him, for he was easily angered and inclined to be censorious, suspicious, and demanding.

In World War II Wittgenstein found it impossible to remain a spectator. He obtained a porter's job at Guy's Hospital in London and worked there from November 1941 to April 1943. He was then transferred to the Royal Victoria Infirmary in Newcastle, where he served as a "lab boy" in the Clinical Research Laboratory until the spring of 1944. He impressed the doctors for whom he worked by the prolonged and concentrated thought he gave to their medical problems. This hard thinking would often result in a new way of looking at the problems. At Newcastle, Wittgenstein devised a simple technique for estimating the area of war wounds that proved of value in determining their treatment.

In 1944 he resumed his lectures at Cambridge. But he became increasingly dissatisfied with his role as a teacher. He feared that his influence was positively harmful. He was disgusted by what he observed of the half understanding of his ideas. "The only seed I am likely to sow is a jargon," he said. He strongly disliked universities and academic life. He felt an increasing need to live alone, per-

haps occasionally seeing a friend, and to devote his remaining energies (for several years he had been repeatedly unwell) to finishing the *Investigations.*

In the fall of 1947 he finally resigned his chair. He sought a secluded life, first in the Irish countryside near Dublin, then in an isolated cottage on the west coast of Ireland. He worked hard when his health permitted it. In the summer of 1949 he went to spend three months with a friend in the United States. Upon his return to England, in the fall, he was discovered to have cancer. He wrote that he was not shocked by this news because he had no wish to continue living. During part of 1950 he visited his family in Vienna, then went to Oxford to live with a friend, and afterward made a trip to Norway. In 1951 he moved to the home of his physician in Cambridge. Wittgenstein had expressed an aversion to spending his last days in a hospital, and his doctor had invited him to come to his own home to die. Wittgenstein was deeply grateful for this offer. Knowing that death was imminent, he continued hard at work. The philosophical thoughts that he wrote in his notebooks at this time are of the highest quality.

On April 27 he was taken violently ill. When his doctor informed him that the end had come he said, "Good!" His last words, before he lost consciousness, were "Tell them I've had a wonderful life!" He died on April 29, 1951.

THE *TRACTATUS*

The *Tractatus* is a comprehensive work of extreme originality, yet it is less than eighty pages long. It is arranged as a series of remarks numbered in decimal notation. The following propositions are distinguished by their numbering as the primary theses of the book:

(1) The world is everything that is the case.

(2) What is the case, the fact, is the existence of states of affairs.

(3) A logical picture of facts is a thought.

(4) A thought is a sentence with a sense.

(5) A sentence is a truth-function of elementary sentences.

(6) The general form of a truth-function is $[\bar{p}, \bar{\xi}, N(\bar{\xi})]$.

(7) Whereof one cannot speak, thereof one must be silent.

Erik Stenius has perceptively remarked that the book has a "musical" structure and that the numbering brings out a "rhythm of emphasis": these seven main propositions are "forte" places in the rhythm.

THE PICTURE THEORY. In a notebook Wittgenstein wrote (*Notebooks,* p. 39): "My *whole* task consists in explaining the nature of sentences." (The German *Satz* will be translated sometimes as "sentence," sometimes as "proposition.") What makes it possible for a combination of words to represent a fact in the world? How is it that by producing a sentence I can say something—can tell someone that so-and-so is the case?

Wittgenstein's explanation consists in the striking idea that a sentence is a picture. He meant that it is literally a picture, not merely like a picture in certain respects. Apparently this thought first occurred to him during the war, when he saw in a magazine an account of how a motorcar accident was represented in a law court by means of small models (see *Notebooks,* p. 7). So he said: "A proposition is a picture of reality. A proposition is a model of reality as we think it to be" (*Tractatus,* 4.01). The dolls and toy cars could be manipulated so as to depict different ways in which the accident might have taken place. They could be used to construct different propositions about the accident—to put forward different accounts, different models of what took place. Wittgenstein's general conception was that when we put a sentence together we construct a model of reality. "In a proposition a situation is, as it were, put together experimentally" (4.031).

One would not normally think that a sentence printed on a page is a picture. According to the *Tractatus* it really is a picture, in the ordinary sense, of what it represents. Wittgenstein conceived the proof of this to be that although words we have not previously encountered have to be explained to us, when we meet for the first time a sentence that is composed of familiar words, we understand the sentence without further explanation. "I understand a sentence without having had its sense explained to me" (4.021). This can appear to one as a remarkable fact. If it is a fact, the only possible explanation would be that a sentence shows its sense. It shows how things are if it is true (4.022). This is exactly what a picture does. A sentence composed of old words is able to communicate a new state of affairs by virtue of being a picture of it.

In any picture, according to the *Tractatus,* there has to be a one-to-one correspondence between the elements of a picture and the things in the state of affairs its represents. If one element of a picture stands for a man and another for a cow, then the relationship between the picture elements might show that the man is milking the

cow. A picture is a fact, namely the fact that the picture elements are related to one another in a definite way. A picture fact shows that the things the picture elements stand for are related in the same way as are the picture elements.

Since a sentence is held to be a picture, there must be as many elements to be distinguished in it as in the state of affairs it portrays. The two must have the same logical or mathematical multiplicity. Again, this does not seem to be true of our ordinary sentences. For Wittgenstein this meant not that it is not true but that our sentences possess a concealed complexity that can be exhibited by analysis.

According to the *Tractatus* a picture must have something in common with what it pictures. This common thing is the picture's "form of representation." There are different kinds of pictures, different pictorial notations, different methods of projection. But all pictures must have in common with reality the same logical form in order to be able to picture reality at all, either truly or falsely. This logical form, also called "the form of reality," is defined as the possibility that things in the world are related as are the elements of the picture (2.18, 2.151). Sentences, since they are pictures, have the same form as the reality they depict.

WHAT CANNOT BE SAID. A picture can depict reality, but it cannot depict its own form of representation. It depicts (represents) its subject from "outside," but it cannot get outside itself to depict its own form of representation. A picture of another form might depict the representational form of a given picture; for instance, a picture in sound might depict the representational form of a picture in color. But in order for the one to represent the form of the other, there must be something that is the same in both. "There must be something identical in a picture and what it depicts, to enable the one to be a picture of the other at all" (2.161). Therefore, logical form, the form of reality, which all pictures must possess, cannot be depicted by any picture.

This consideration must apply to sentences, too. We make assertions by means of sentences. With a sentence we say something. We say how things are. Things in the world are related in a certain way, and we try to describe that. But we cannot describe how our sentences succeed in representing reality, truly or falsely. We cannot say what the form of representation is that is common to all sentences and that makes them pictures of reality. We cannot say how language represents the world. We cannot state in any sentence the pictorial form of all sentences. "What can be said can only be said by means of a sentence, and so nothing that is necessary for the understanding of *all* sentences can be said" (*Notebooks,* p. 25).

This doctrine implies that in a sense one cannot say what the meaning of a sentence is. With regard to the sentence "*a* is larger than *b,*" one can explain to a person what "*a*" and "*b*" each refer to and what "larger" means, but there is not a further explanation to give him, namely what "*a* is larger than *b*" means. We understand the elements of a sentence, and we see how they are combined. But we cannot say what this combination means. Yet we grasp its meaning. In some sense we know what it means, because the sentence shows its meaning. Anything that can be said can be said clearly, but not everything that is understood can be said. In a letter to Russell, Wittgenstein remarked that his "main contention" was this distinction between what can be said in propositions—that is, in language—and what cannot be said but can only be shown. This, he said, was "the cardinal problem of philosophy."

THE NATURE OF THOUGHT. The picture theory of propositions is at the same time an account of the nature of thought. Wittgenstein said: "A thought is a sentence with a sense" (*Tractatus,* 4). This implies that thinking is impossible without language. Since a thought is a sentence and a sentence is a picture, a thought is a picture. The totality of true thoughts would be a true picture of the world.

The view that a thought is a sentence seems to imply that the words of a sentence could be the constituents of a thought. But in a letter written to Russell shortly after the *Tractatus* was completed, Wittgenstein explicitly denied this. A thought consists not of words "but of psychical constituents that have the same sort of relation to reality as words. What those constituents are I don't know." "I don't know *what* the constituents of a thought are but I know *that* it must have such constituents which correspond to the words of Language" (*Notebooks,* pp. 130, 129). It would appear from these remarks that Wittgenstein's view was not that a thought and a sentence with a sense are one and the same thing but that they are two things with corresponding constituents of different natures. Each of these two things is a picture. "Thinking is a kind of language. For a thought too is, of course, a logical picture of a sentence, and therefore it just is a kind of sentence" (*Notebooks,* p. 82).

To say that a state of affairs is conceivable (thinkable) means that we can make a picture of it (*Tractatus,* 3.001). A thought "contains" the possibility of a state of affairs, for the logical form of the thought is the possibility that

things in the world are combined in the way the constituents of the thought are combined. Whatever is conceivable is possible. In a spoken or written sentence a thought is "made perceptible to the senses." All thoughts can be stated in sentences; what cannot be stated cannot be thought.

A consequence of these views is that the form of representation of propositions (the form of reality, logical form), which cannot be stated, also cannot be thought. Language shows us something we cannot think. A function of philosophy is to indicate (*bedeuten*) what cannot be said (or thought) by presenting clearly what can be said. According to the *Tractatus,* therefore, there is a realm of the unthinkable that, far from being a mere wind egg, is the foundation of all language and all thought. In some way we grasp this foundation of thought (what we do here cannot really be said); it is mirrored in our thoughts, but it cannot be an object of thought.

Obviously the *Tractatus* is a thoroughly metaphysical work; this is not a minor tendency of the book. Yet it was once widely regarded as being antimetaphysical in its outlook. There is some excuse for this interpretation, since at the end of the book Wittgenstein said that the correct philosophical method would be to prove to anyone who wants to say something metaphysical that he has failed to give a meaning to certain signs in his sentences (6.53). But Wittgenstein did not reject the metaphysical; rather, he rejected the possibility of stating the metaphysical.

NAMES AND OBJECTS. The conception of propositions, and therefore of language, in the *Tractatus* rests on the notion of a name. This is defined as a "simple sign" employed in a sentence. A simple sign is not composed of other signs, as, for example, the phrase "the king of Sweden" is. The word *John* would satisfy this requirement of a simple sign. But a further requirement of a name is that it should stand for a simple thing, which is called an "object." According to the *Tractatus* the object for which a name stands is the meaning of the name (3.203). It is easy to determine whether a sign is composed of other signs but not whether it stands for something simple.

Wittgenstein conceived of objects as absolutely simple and not merely as simple relative to some system of notation. "Objects make up the substance of the world. That is why they cannot be composite.… Substance is what exists independently of what is the case.… Objects are identical with the fixed, the existent.… The configuration of objects is the changing, the mutable" (2.021, 2.024, 2.027, 2.0271).

A name is not a picture of the object it stands for, and therefore a name does not say anything. A picture in language—that is, the sentence—can be formed only by a combination of names. This combination pictures a configuration of objects. The combination of names is like a *tableau vivant* (4.0311). (One might think here, for example, of a group of people posed to represent *The Last Supper*). A name is a substitute for an object, and a combination of names portrays a configuration of objects—that is, a state of affairs (*Sachverhalt*).

A reader of the *Tractatus* will be perplexed to know what examples of names and of objects would be. No examples are given. It is said that names occur only in "elementary" propositions, but there are no examples of the latter notion. Wittgenstein was not able to come to any conclusion about examples. The *Notebooks* show that he was very vexed by this problem. He struggled with the question of whether "points of the visual field" might be simples (see, for example, p. 45). Sometimes he wondered whether any ordinary name whatsoever might not be a "genuine" name. And he wondered whether his watch might not be a "simple object" (*Notebooks,* pp. 60–61). His final conviction that there are absolutely simple objects was purely a priori. He wrote in his notes:

> It seems that the idea of the *simple* is already to be found contained in that of the complex and in the idea of analysis, and in such a way that we come to this idea quite apart from any examples of simple objects, or of propositions which mention them, and we realize the existence of the simple object—*a priori*—as a logical necessity. (*Notebooks,* p. 60)

The "logical necessity" arises from the requirement that propositions have a definite sense. "The demand for simple things is the demand for definiteness of sense" (*Notebooks,* p. 63). As it is put in the *Tractatus,* "The requirement that simple signs be possible is the requirement that sense be definite" (3.23). An indefinite sense would be no sense at all. A proposition might be ambiguous, but the ambiguity would be between definite alternatives: either this or that.

The sentences of everyday language are in perfect logical order. This order rests on the simples—that which is fixed, unchangeable, hard (*das Harte: Notebooks,* p. 63). The simples and their configurations—that is what order is. Wittgenstein said: "Our problems are not abstract, but perhaps the most concrete that there are" (*Tractatus,* 5.5563).

ELEMENTARY PROPOSITIONS. A combination of genuine names is an elementary proposition. It is not analyzable into other propositions. "It is obvious that the analysis of propositions must bring us to elementary propositions which consist of names in immediate combination" (4.221). An elementary proposition shows (represents) a certain configuration of simple objects.

The picture theory is meant to hold for all genuine propositions, not merely for elementary propositions. Wittgenstein said without qualification: "A proposition is a picture of reality" (4.01, 4.021). Elementary and nonelementary propositions are equally pictures: the difference is that in an elementary proposition the pictorial nature is manifest. "It is *evident* that we perceive (*empfinden*) an elementary proposition as the picture of a state of affairs" (*Notebooks,* p. 25). But Wittgenstein admitted that most sentences do not seem to be pictures.

> At first sight a sentence—one set out on the printed page, for example—does not seem to be a picture of the reality with which it is concerned. But no more does musical notation at first sight seem to be a picture of music, nor our phonetic notation (letters) to be a picture of our speech. And yet these sign-languages prove to be pictures, even in the ordinary sense, of what they represent. (*Tractatus,* 4.011)

All genuine propositions, according to the *Tractatus,* are analyzable into elementary propositions. This analysis of our ordinary propositions, with their complicated modes of symbolizing—their various "methods of projection"—will make manifest their concealed pictorial nature. In his introduction to the *Tractatus,* written for the first English edition, Russell said:

> Mr. Wittgenstein is concerned with the conditions for a logically perfect language—not that any language is logically perfect, or that we believe ourselves capable, here and now, of constructing a logically perfect language, but that the whole function of language is to have meaning, and it only fulfils this function in proportion as it approaches to the ideal language which we postulate.

That this is an incorrect account of the *Tractatus* is sufficiently shown by Wittgenstein's remark "All the propositions of our everyday language are actually in perfect logical order, just as they are" (5.5563). The analysis achieved by the philosophical logician will not create order where previously there was no order; instead, it will make evident what is already there.

Every genuine proposition has one and only one complete analysis into elementary propositions (3.25). This is so even if every fact consists of infinitely many states of affairs and every state of affairs is composed of infinitely many simple objects (4.2211). The completely analyzed proposition will consist of simple names; the meaning of each simple name will be a simple object; the particular way in which the names are combined in the proposition will say that the simple objects in the world are related in the same way. To understand the completely analyzed proposition one need only understand the names—that is, know what objects they stand for. What their combination means will be immediately evident. Understanding a proposition requires merely understanding its constituents (4.024).

As Rush Rhees has remarked, the idea that there are elementary propositions is not an arbitrary assumption. Wittgenstein was trying to solve the question of how language and thought can be related to reality. His basic intuition was that language pictures reality. If this is so, then among the sentences of language there must be some that show their sense immediately, which, of course, does not mean that their truth is self-evident. Wittgenstein had no criteria for identifying elementary propositions and could give no general account of their subject matter. But if his intuition was right, then there must be elementary propositions—that is, propositions that show their sense immediately and of which all other propositions are "truth-functions." If this were not so, no sentence could say anything or be understood (Rush Rhees, "The *Tractatus*: Seeds of Some Misunderstandings," pp. 218–219).

THEORY OF TRUTH-FUNCTIONS. A truth-function of a single proposition *p* is a proposition whose truth or falsity is uniquely determined by the truth or falsity of *p*; for example, *not-p* (*p* is false) is a truth-function of *p.* A truth-function of two propositions *p, q* is a proposition whose truth or falsity is uniquely determined by the truth or falsity of *p, q*; for instance, "*p, q* are both true" is a truth-function of *p, q.* According to the *Tractatus* (5) every genuine proposition is a truth-function of elementary propositions. (It is an interesting and difficult question whether this doctrine follows from the picture theory or, on the other hand, is even compatible with it.) If two nonelementary propositions *r* and *s* are truth-functions of some of the same elementary propositions, then *r* and *s* will be internally related: For instance, one of them may logically follow from the other, or they may be contradictories or contraries of each other. If we see the internal structure of two propositions, we know what log-

ical relations hold between them. We do not need, in addition, a knowledge of logical principles. We can actually do without the formal principles of logic, "for in a suitable notation we can recognize the formal properties of propositions by mere inspection of the propositions themselves" (6.122).

Wittgenstein employed a technique (known as the method of truth tables) for making manifest the truth conditions of a proposition that is a truth-function of other propositions—that is, for exhibiting the relation between the truth or falsity of the latter and the truth or falsity of the former.

There are two limiting cases among the possible groupings of truth conditions of propositions. One case would be when a proposition was true for all truth possibilities of the elementary propositions; this proposition is called a tautology. The other would be when a proposition was false for all the truth possibilities; this proposition is called a contradiction. Although it is convenient to refer to tautologies and contradictions as "propositions," they are actually degenerate cases, not genuine propositions. They are not pictures of reality. They do not determine reality in any way. They have no truth conditions, since a tautology is *un*conditionally true and a contradiction *un*conditionally false. Wittgenstein compared a genuine proposition, a picture, to "a solid body that restricts the freedom of movement of others." In contrast a tautology (for example, "He is here, or he is not here") "leaves open to reality the whole of logical space." No restriction is imposed on anything. A contradiction (for example, "He is here, and he is not here") "fills the whole of logical space and leaves no point of it for reality" (4.461, 4.462, 4.463).

According to the *Tractatus* the so-called propositions of logic, logical truths, principles of logic are all tautologies. They express no thoughts. They say nothing. We could do without them. But they are not nonsense, for the fact that a certain combination of propositions yields a tautology reveals something about the structures of the constituent propositions. "That the propositions of logic are tautologies *shows* the formal—logical—properties of language, of the world" (6.12).

NECESSITY. Wittgenstein's picture theory and his explanation of logical truth lead to an interesting doctrine of necessity and also to a denial of any knowledge of the future. Genuine propositions say only how things are, not how things must be. The only necessity there can be is embodied in tautologies (and the equations of mathematics). Neither tautologies nor equations say anything about the world. Therefore, there is no necessity in the world. "Outside of logic everything is accidental" (6.3). One proposition can be inferred from another proposition only if there is an internal, structural connection between them. The existence of one state of affairs cannot be inferred from the existence of another, entirely different, state of affairs (5.135). But that is what an inference to a future state of affairs would have to be. Thus Wittgenstein declared that we do not know whether the sun will rise tomorrow (6.36311).

WILL AND ACTION. If we conceive of an act of will (a volition) as one occurrence and the transpiring of what is willed as an entirely different occurrence, it follows from the foregoing doctrines that there can be, at most, a merely accidental correlation between one's will and what happens in the world. I cannot make anything happen—not even a movement of my body. "The world is independent of my will" (6.373). In his notes Wittgenstein gave this idea dramatic expression: "I cannot bend the happenings of the world to my will: I am completely powerless" (*Notebooks,* p. 73).

ETHICS. According to the picture theory a proposition and its negation are both possible; which one is true is accidental. Wittgenstein drew the conclusion that there can be no propositions of ethics. His thought here was that if anything has value, this fact cannot be accidental: the thing must have that value. But everything in the world is accidental. Therefore there is no value in the world. "In the world everything is as it is, and everything happens as it does happen: *in* it no value exists—and if it did, it would have no value" (*Tractatus,* 6.41).

This view is an absolute denial not of the existence of value but of its existence in the world. Propositions can state only what is in the world. What belongs to ethics cannot be stated; it is "transcendental" (6.421). The world, and what is in the world, is neither good nor evil. Good and evil exist only in relation to the subject (the ego). But this "subject" to which Wittgenstein referred is also "transcendental." It is not in the world but is a "limit" of the world (5.5632).

THE MYSTICAL. In the view of the *Tractatus* there are a variety of things that cannot be stated: the form of representation of propositions, the existence of the simple objects that constitute the substance of the world, the existence of a metaphysical subject, of good and evil—these things are all unsayable. Wittgenstein seems to have believed that we have thoughts on these matters only

when we view the world as a limited whole. This latter experience is what he called "the mystical" (6.45).

Although one cannot say anything on these metaphysical topics included in the mystical, this is not because they are absurd but because they lie beyond the reach of language. "Unsayable things do indeed exist" (*Es gibt allerdings Unaussprechliches*: 6.522). This itself is something unsayable. It is one of those sentences of his own of which Wittgenstein declared that although they can produce philosophical insight, they are actually nonsensical and eventually must be "thrown away" (6.54). The final proposition of the book ("Whereof one cannot speak, thereof one must be silent") is not the truism one might take it to be, for it means that there is a realm about which one can say nothing.

THE *TRACTATUS* AND LOGICAL POSITIVISM. The *Tractatus* exerted a considerable influence on the so-called Vienna circle of logical positivism. Moritz Schlick, the leader of this movement, declared that the *Tractatus* had brought modern philosophy to a "decisive turning point." It is true that there is some agreement between the predominant views of the Vienna circle and the positions of the *Tractatus*—for example, that all genuine propositions are truth-functions of elementary propositions, that logical truths are tautologies and say nothing, and that philosophy can contain no body of doctrine but is an activity of clarifying thoughts.

But there are fundamental differences. The Vienna circle did not adopt the picture theory of propositions, which is the central idea of the *Tractatus.* A conspicuous doctrine of the circle was that all genuine propositions are reducible to propositions that report "direct perception" or what is "immediately given in experience." This doctrine is not found in the *Tractatus.* A corollary to it is the famous positivist thesis "The meaning of a statement is its method of verification." But the topic of verification is not even brought into the *Tractatus.* The only proposition there that seems to resemble this thesis is the following: "To understand a proposition means to know what is the case if it is true" (4.024). Even here nothing is explicitly said about verification, and a comment immediately following this remark shows that Wittgenstein was not thinking about verification. A proposition, he said, "is understood by anyone who understands its constituents." That is to say, if you understand the words in a sentence, you thereby understand the sentence. There is no mention of a requirement that you must know how to verify what it says.

As previously noted, Wittgenstein was tempted by the suggestion that "points in the visual field" are examples of the simples out of which all meaning is composed. But the final view of the *Tractatus* is that the simples are fixed, immutable things, which exist "independently of what is the case." If so, they cannot be described by propositions and cannot be given in experience. The *Tractatus* does not contain, therefore, an empiricist theory of meaning. What it holds is that to understand any sentence one must know the references of the names that compose it; that is all. When you understand a sentence you know how reality is constituted if the sentence is true, regardless of whether you know how to verify what it says. The picture theory is not a verification theory of meaning. It is ironical that the role of verification in meaning and understanding receives much attention in Wittgenstein's later philosophy, which obviously is not positivistic, but none at all in the reputedly positivistic *Tractatus.*

Logical positivism and the author of the *Tractatus* were both opposed to metaphysics, but in different ways. For positivism there is nothing at all behind metaphysical propositions except possibly their authors' emotions. "Metaphysicians are musicians without musical ability," said Rudolf Carnap. In the view of the *Tractatus* one may gain insights into the presuppositions and limits of language, thought, and reality. These metaphysical insights cannot be stated in language, but if they could be, they would be true insights and not mere muddles or expressions of feeling.

The foregoing sketch of the *Tractatus* has omitted many of its important topics. Wittgenstein wrote in his notes, "My work has extended from the foundations of logic to the nature of the world." In his preface to the *Tractatus* he expressed the opinion that he had obtained the final solution of the problems treated in the book, but he added that one value of his work is that "it shows how little is achieved when these problems are solved."

THE "NEW" PHILOSOPHY

In 1929, Wittgenstein returned to Cambridge, after an absence of more than fifteen years, to resume philosophical research and to lecture. From then until his death he did a huge amount of writing. Among the first works of this period were two large typescript volumes. One, which was composed in the period 1929–1930, has been published under the title *Philosophische Bemerkungen.* The other is a systematic work of nearly 800 typewritten pages written between 1930 and 1932. In both of these volumes Wittgenstein reexamined the problems of the

Tractatus and revised what he had written there. This led him to questions he had not previously considered. Perhaps it can be said that he found that the logical investigations of the *Tractatus* and its supreme problem of the relation of language to reality had drawn him more and more into questions in the philosophy of psychology. These volumes seem to show that the change from the *Tractatus* to the *Philosophical Investigations* was an intensive but continuous development rather than a sudden revolution.

In 1933–1934, Wittgenstein dictated to his students a set of notes that came to be called the *Blue Book,* and in 1934–1935 he dictated another set, later known as the *Brown Book.* (Although Wittgenstein always wrote in German, the *Blue Book* and the *Brown Book* were dictated in English.) Both circulated widely in typescript, and Wittgenstein's new ideas began to create a stir. The *Blue Book* is clear and lively and is perhaps the beginner's best introduction to Wittgenstein. Nevertheless, it is a comparatively superficial work; Wittgenstein never regarded it as more than a set of class notes. The *Brown Book,* on the other hand, he regarded for a short time as a draft of something that might be published. He worked at a revision but gave it up in 1936, when he began to write the *Philosophical Investigations.* Wittgenstein refrained from publishing the *Investigations* during his lifetime, but his explicit wish was that it be published posthumously, a wish that he probably did not have with respect to any of the rest of the voluminous work he produced between 1929 and 1951.

The *Philosophical Investigations* was published in 1953 in two parts. Part I was written in the period 1936–1945 and Part II between 1947 and 1949. Concurrently with the *Investigations,* Wittgenstein did other writing, which was closely related to the topics of the *Investigations* or even overlapped it. From the years 1937 to 1944 there are extensive manuscripts on the philosophy of logic and mathematics. *Remarks on the Foundations of Mathematics,* published in 1956, consists of selections, made by the editors, from this material. A quantity of writing in the form of loose notes, probably from the years 1947 to 1949, is of the same subject matter and quality as the latter part of Part I of the *Investigations.* Wittgenstein's last manuscript notebooks, from the years 1949 to 1951, treating questions about belief, doubt, knowledge, and certainty, also contain much material that should eventually be published.

PHILOSOPHICAL INVESTIGATIONS

Wittgenstein believed that the *Investigations* could be better understood if one saw it against the background of the *Tractatus.* A considerable part of the *Investigations* is an attack, either explicit or implicit, on the earlier work. This development is probably unique in the history of philosophy—a thinker producing, at different periods of his life, two highly original systems of thought, each system the result of many years of intensive labors, each expressed in an elegant and powerful style, each greatly influencing contemporary philosophy, and the second being a criticism and rejection of the first.

Apparently it is possible for a serious student of Wittgenstein to form the impression that "the *Investigations* basically contains an application of the main ideas of the *Tractatus* to several concrete problems, the only difference being the use of language-games instead of the language of the natural sciences which formed the theoretical background of the *Tractatus.*" This view is thoroughly mistaken, as will be seen.

THE WHOLE OF LANGUAGE. It is held in the *Tractatus* that any proposition presupposes the whole of language. "If objects are given, then at the same time we are given *all* objects. If elementary propositions are given, then at the same time *all* elementary propositions are given" (5.524). "If all objects are given, then at the same time all *possible* states of affairs are also given" (2.0124). An elementary proposition is a combination of names, and in order to understand the proposition one must in some sense "know" the objects for which the names stand. In understanding any proposition at all one must know some objects, and therefore, as stated, one must know all objects and all possibilities. Any proposition whatsoever carries with it the whole of "logical space." This view is connected with the idea that there is an essence of propositions. The essence of propositions is "the essence of all description, and thus the essence of the world" (5.4711). The essence of propositions is the same as "the universal form of proposition" (*Die allgemeine Satzform*). That there is a universal form of proposition is proved by the fact that all possibilities—i.e., all forms of proposition—"must be *foreseeable*" (*Notebooks,* p. 89; *Tractatus,* 4.5).

The *Investigations* emphatically rejects the idea that each proposition carries with it the whole of language. A sentence does presuppose a "language game," but a language game will be only a small segment of the whole of language. An example of a language game is the following, which appears at the beginning of the *Investigations* (Sec. 2): There are a builder and his helper. The building

materials are blocks, pillars, slabs, and beams. The two men have a language consisting of the words *block, pillar, slab, beam.* The builder calls out one of the words and the helper brings the building material that he has learned to bring at that call. Wittgenstein called the words and the actions with which they are joined a language game (*Sprachspiel*). He said that it is complete in itself and could even be conceived to be the entire language of a tribe. If we think it is incomplete we are only comparing it with our more complex language. In the *Brown Book* there is the analogy of someone's describing chess without mentioning pawns. As a description of chess it is incomplete, yet we can also say that it is a complete description of a simpler game (*Blue and Brown Books,* p. 77). This simpler game does not presuppose chess, nor does the part played, for example, by the word *block* in the game of Sec. 2 imply its use in descriptions or questions.

According to the *Tractatus* every form of proposition can be anticipated because a new form of proposition would represent a new combination of simple objects in logical space. It would be like grouping the pieces on a chessboard in a new way. It would be a different arrangement of what you already have. But in Wittgenstein's later philosophy a new language game would embody a new "form of life," and this would not merely be a rearrangement of what was there before. Suppose the people of a certain tribe use language to describe events that are occurring or have occurred (such as men walking, running, or fighting, or the weather), or that they believe have occurred, but they do not have any imaginative use of language. They do not lie, pretend, make supposals, or engage in any imaginative play. Nor does any behavior of pretending occur: the children do not ever, for example, walk on all fours and growl as if they were lions. These people would not understand kidding. If one of us said to them something obviously false and then laughed, they would not know how to take it. (We should remember that among ourselves we differ greatly in our responsiveness to joking and pretense.) What these people lack is not words but the behavior and reactions that enter into the language games of imagination. Are they capable of foreseeing a use of language to convey a play of imagination? They do not even understand it when they encounter it. A new use of language embedded in a new form of life could not be anticipated, any more than could the rise of nonobjective painting.

THE ESSENCE OF LANGUAGE. The *Tractatus* assumes that there is a universal form of language, just as it assumes (6.022) that there is a universal form of number—that which is common to all numbers. The *Investigations* rejects this assumption. There is nothing common to the various forms of language that makes them language. There is not something common to all language games, just as there is not something common to all games. We are asked to consider the various kinds of games there are (for example, board games, card games, ball games) and the variety within each kind. If we pick out a feature common to two games we shall find that it is absent from some other place in the spectrum of games. Not all games are amusing, not all involve winning or losing, not all require competition between players, and so on. What makes all of them games, what gives unity to those activities, is not some feature present in all games but a multitude of relationships "overlapping and criss-crossing." Wittgenstein employed the analogy of a family resemblance. One can often see a striking resemblance between several generations of the same family. Studying them at close hand one may find that there is no feature common to all of the family. The eyes or the build or the temperament are not always the same. The family resemblance is due to many features that "overlap and criss-cross." The unity of games is like a family resemblance. This is also the case with sentences, descriptions, and numbers.

> Why do we call something a "number": Well, perhaps because it has a—direct—relationship with several things that have hitherto been called number; and this can be said to give it an indirect relationship to other things we call the same name. And we extend our concept of number as in spinning a thread we twist fibre on fibre. And the strength of the thread does not reside in the fact that some one fibre runs through its whole length, but in the overlapping of the fibres. (Sec. 67)

One of the remarkable features of the *Investigations* is the detail and ingenuity of Wittgenstein's examination of some sample concepts (*reading, deriving, being guided*: Secs. 156–178) in order to bring out the variety of cases that fall under them and to prove that they are not united by an essence. If these concepts do not have an essential nature, then neither do the concepts of *description, proposition,* and *language.* The *Tractatus* was wrong in a most fundamental assumption.

ABSOLUTE SIMPLES. The *Tractatus* held that the ultimate elements of language are names that designate simple objects. In the *Investigations* it is argued that the words *simple* and *complex* have no absolute meaning. It has to be laid down, within a particular language game, what is to

be taken as simple and what composite. For example, is one's visual image of a tree simple or composite? The question makes no sense until we make some such stipulation as that if one sees merely the trunk, it is simple, but if one sees trunk and branches, it is composite.

> But isn't a chess board, for instance, obviously, and absolutely composite?——You are probably thinking of the composition out of thirty-two white and thirty-two black squares. But could we not also say, for instance, that it was composed of the colours black and white and the schema of squares? And if there are quite different ways of looking at it, do you still want to say that the chessboard is absolutely "composite"? … Is the colour of a square on a chessboard simple, or does it consist of pure white and pure yellow? And is white simple, or does it consist of the colours of the rainbow?——Is this length of 2 cm. simple, or does it consist of two parts, each 1 cm. long? But why not of one bit 3 cm. long, and one bit 1 cm. long measured in the opposite direction? (Sec. 47)

By such examples Wittgenstein tried to show that the ideas of "simple" and "complex" are necessarily relative to a language game. The notion of a simplicity that is not relative but absolute, because all of language is based on it, is a philosophical "super-concept." We have an image but we do not know how to apply it: we do not know what would be an example of an absolute simple.

In the *Tractatus* the existence of simple objects was conceived as following from the requirement that the sense of sentences be definite. In the *Investigations* this requirement is regarded as another philosophical illusion. We have imagined an "ideal" of language that will not satisfy actual needs. A sharp boundary has not been drawn between, for example, games and activities that are not games. But why should there be one in general? Precision and exactness are relative to some particular purpose. The guests are to arrive exactly at one o'clock, but this notion of exactness would not employ the instruments and measurements of an observatory. "No *single* ideal of exactness has been laid down; we do not know what we should be supposed to imagine under this head" (Sec. 88). Losing sight of the fact that there are different standards of exactness for different purposes, we have supposed that there is a certain state of complete exactness underneath the surface of our everyday speech and that logical analysis can bring it to light. We have supposed, therefore, that a proposition would have one and only one complete analysis.

In searching for the ideal of perfect exactness we become dissatisfied with ordinary words and sentences. We do not find in actual language the pure and clear-cut structure that we desire. The more closely we examine actual language, the sharper becomes the conflict between it and our philosophical ideal. The latter now begins to seem empty. We do not even understand how it could be realized in actual language. We have been bewitched by a picture. Instead of trying to perceive in our language a design too fine to grasp, we need to see more clearly what is really there. We should abandon preconceived ideas and hypotheses and turn to description, the purpose of which will be to remove our philosophical perplexities. The substitution of description for analysis, and the new conception that nothing is hidden, is a major change from the *Tractatus.*

MEANING AS USE. If the picture theory is the central feature of the *Tractatus,* it is important to see how Wittgenstein's new thinking judged that theory. Surprisingly, there is not much explicit discussion of it, and the remarks that do occur are usually enigmatic. But if we take a long view of the new philosophy, there can be no question that it rejects the picture theory. In the later work as well as the earlier, Wittgenstein was concerned with the question, How can a sentence say something; how can language represent reality? The first sentence of the *Blue Book* is "What is the meaning of a word?" and it might equally well have been "What is the meaning of a sentence?" Both philosophical systems are centered on the same question, but the answer given in the second is entirely different. Instead of holding that a sentence has meaning or sense because it is a picture, the *Investigations* says that the meaning of a sentence is its "use" (*Gebrauch*) or "employment" (*Verwendung*) or "application" (*Anwendung*).

Some readers of Wittgenstein have doubted that he spoke of the use of a sentence, and others have thought that in any case it is wrong to speak this way. There is no question on the first point. Wittgenstein spoke of the "use" of a sentence in many passages. For example: "But doesn't the fact that sentences have the same sense consist in their having the same *use*?" (*Investigations,* Sec. 20); there are "countless different kinds of use of what we call 'symbols,' 'words,' 'sentences'" (Sec. 23).

The other objection may be important. Some philosophers want to say that a sentence cannot have a use. Words have a use; we learn the use of words, not of sentences. We understand sentences without having their

sense explained to us, because we understand the use of the words that compose them.

What is espoused here is really the ground of the picture theory of the *Tractatus* (cf. *Tractatus,* 4.021, 4.026, 4.027). In the *Investigations* there is more than one objection to the above argument. Wittgenstein denied that we always understand a sentence, even if it is a grammatically correct sentence whose words we do understand. If someone says, for example, that the sentence "This is here" (saying which, he points to an object in front of him) makes sense to him, "then he should ask himself in what special circumstances this sentence is actually used. There it does make sense" (Sec. 117). "A philosopher says that he understands the sentence 'I am here,' that he means something by it, thinks something—even when he doesn't think at all how, on what occasions, this sentence is used" (Sec. 514). Wittgenstein was saying that these sentences have sense only in special circumstances; in other circumstances we do not understand them—that is, we do not know what to do with them.

The view of the *Tractatus* is entirely different. An elementary sentence is a combination of names, and if we know what the names refer to, then we understand the sentence, for it shows its sense. "Circumstances" have nothing to do with it. The *Investigations* regards this view as absurd. What does the sentence "I am here" show? Certainly it does not show its use. What can it mean to say that it shows its sense? A significant sentence is a tool with which a certain job is done. By looking at a sentence you cannot always tell whether it is a tool and, if it is, what job it is used for. The *Investigations* denies the claim that was the basis of the picture theory, namely that "we understand the sense of a propositional sign without its having been explained to us" (*Tractatus,* 4.02).

In holding that (in many cases) the meaning of an expression is its use, Wittgenstein was not declaring that the words *meaning* and *use* are general synonyms. By the "use" of an expression he meant the special circumstances, the "surroundings," in which it is spoken or written. The use of an expression is the language game in which it plays a part. Some readers have arrived at the mistaken idea that by the "use" of an expression Wittgenstein meant its ordinary or its correct use: They have thought that he was an "ordinary-language philosopher." But Wittgenstein studied any use of language, real or imaginary, that may illuminate a philosophical problem. Often he invented language games that corresponded to no actual use of language (see, for example, *Blue and Brown Books,* pp. 103–104, 110). The language games are "*objects of comparison* which are meant to throw light on the facts of our language by way not only of similarities, but also of dissimilarities" (*Investigations,* Sec. 130).

The *Tractatus* holds that language is ultimately composed of names, that the meaning of a name is a simple object, and that the sense of a sentence arises from the names that compose it. One name stands for one thing, another for another thing, and the combination pictures a state of affairs (4.0311). Thus, naming is prior to the sense of sentences (although it is also said that a name has meaning only in a sentence: 3.3). A sentence says something because it is composed of names that stand for things. In the *Investigations* two objections are made against this notion of the priority of names. First, the meaning of a word is never the thing, if there is one, that corresponds to the word (Sec. 40). Second, before one can find out what a name stands for one must already have mastered the language game to which the name belongs. In order to learn the name of a color, a direction, a sensation, one must have some grasp of the activities of placing colors in an order, of reading a map, of responding to the words, gestures, and behavior that are expressions of sensation. Merely pointing at something and saying a word achieves nothing. The kind of use the word will have, the special circumstances in which it will be said, must be understood before it can even be a name.

One could say that the *Tractatus* conceives of a significant sentence as having the nature of a mechanism. If the parts fit, then the whole thing works: you have a picture of reality. If the parts do not fit, they are like cogwheels that do not mesh. There is, as it were, a clash of meanings. But in the *Investigations* we read: "When a sentence is called senseless, it is not as it were its sense that is senseless" (Sec. 500). If someone said to us, for example, "My head is asleep," we should be perplexed. It would be no help if he said: "You know what it is for an arm or a leg to be asleep. I have the same thing, except that it is my head." Here we do not know what the "same" is. It is not that we see that the meaning of "head" is incompatible with the meaning of "asleep." We do not perceive a clash of meanings. But we do not know what behavior and circumstances go with this sentence. It is not that we see that it cannot have a use (because the words do not fit together). The fact is that it does not have a use: we do not know in what circumstances one should say it. "Look at the sentence as an instrument, and at its sense as its employment!" (Sec. 421). Instead of the fundamental notion being the right combination of words and the sense of the sentence being explained in terms of it, it is the other way around: whether the sentence has an "employment" (*Verwendung*) is what is fundamental.

This would be our only criterion for whether there is a sense-making combination of parts.

One additional criticism of the picture theory will be noted. Suppose that a sentence were a picture. There would still be a question of how we should apply the picture. If someone showed you a drawing of a cube and told you to bring him one of those things, you might in good faith bring him a triangular prism instead of a cube. More than one way of taking the drawing was possible. It suggests a cube, but it is possible to interpret the drawing differently. A picture represents an old man walking up a steep path leaning on a stick. But could it not also represent him as sliding down the hill in that position? For us it is more natural to take it in the first way, but the explanation of this does not lie in anything intrinsic to the picture. A picture of a green leaf might be understood to be a representation of the color green, or of a specific shade of green, or of leaf shape in general, or of a particular shape of leaf, or of foliage in general, and so on. How a picture is used will determine what it is a picture of. It cannot, therefore, be a fundamental explanation of the sense of sentences to say that they are pictures. Wittgenstein hinted that the picture theory is plausible because we tend to think of portraits that hang on our walls and are, as it were, "idle." If we consider instead an engineer's machine drawing or an elevation with measurements, then the activity of using the picture will be seen to be the important thing (Sec. 291).

LOGICAL COMPULSION. Our discussion may suggest the following view: How a word, sentence, or picture is interpreted determines what use is made of it. How a man responds to an order, for example, depends on how he understands it, and whether the one who gave the order will be satisfied with that response will depend on what he meant by it. If someone understands the algebraic formula determining a numerical series, then he will know what numbers should occur at various places in the expansion of the series. What a person deduces from a proposition will depend entirely on his understanding of the proposition. Wittgenstein once wrote (in a pre-*Tractatus* notebook): "What propositions follow from a proposition must be completely settled before that proposition can have a sense" (*Notebooks,* p. 64). By virtue of grasping the meaning or sense of an expression we know how to employ it: we know when to say it and what action it calls for. Instead of meaning being identical with use, it comes before use, and use is based on it. When you hear a sentence and understand it or give an order and mean it, the action required in responding to the sentence or obeying the order is already, in a queer sense, taken in your mind. In your act of meaning or understanding, "your mind as it were flew ahead and took all the steps" before they were taken physically (*Investigations,* Sec. 188). In taking, or accepting, those physical steps, you would be ratifying what has already transpired in your mind. To do differently would be inconsistent with the previous mental act. Consistency, rationality, requires you to take these steps or draw these conclusions. Understanding carries compulsion with it.

This idea of "logical compulsion" is vigorously attacked in the *Investigations* and in Wittgenstein's writings on the foundations of mathematics. Was Wittgenstein rejecting deductive reasoning and logical necessity? No. He was rejecting this picture of logical necessity, namely that when I have understood a proposition and there is a question of what follows from it, I have to deduce such-and-such consequences because it was already settled in my understanding of the proposition that it would have those consequences. Wittgenstein's criticism of this imagery creates a continuity between his philosophy of psychology and his philosophy of logic. A part of his criticism could be put as follows: Suppose that two people, *A* and *B,* have received the same instruction in elementary arithmetic. They have been given the same rules and illustrations and have worked through the same examples. Later, when they are required to perform some arithmetical operation, *A* does it right and *B* wrong, although *B* thinks he has done it correctly. We shall say that *A* understood the problem and *B* did not. What does this come to? It could have been that the sole difference between them was that *A* wrote down correct numbers and *B* incorrect ones. If this fact is our criterion of a difference of understanding, then it is wrongheaded to postulate a difference of understanding to explain the fact that *A* and *B* wrote down different answers.

The inclination to insert an act or state of understanding as an intermediary between, for example, hearing an order and executing it is an example of what is called in the *Brown Book* (*Blue and Brown Books,* p. 143) "a general disease of thinking." It consists in always looking for (and "finding") mental states and acts as the sources of our actions. Other examples of this inclination are thinking that one must know where one's pain is before one can point to the place, thinking that we call various shades of red by the name "red" because we see something in common in all of them, thinking that we speak of "looking in our memory for a word" and of "looking in the park for a friend" because we have noticed a similarity between the two cases.

The assumption of mental states to explain our actions comes from a "one-sided diet." If we let our view range over the family of cases of "differences of understanding," we shall discover some in which the only difference between two people who understood a certain proposition differently consists in their having drawn different conclusions from it.

Must we believe, then, that our understanding does not reach beyond the particular training we received and the examples we studied? No. There is a good sense in which it reaches beyond, for we do go on to apply rules in new cases in what we agree is the same way we were taught. Does this agreement have to be explained by the fact that our understanding has penetrated to the essence of the examples? No. This agreement is one of the "extremely general facts of nature" (*Investigations,* pp. 56, 230) that underlie our concepts. We do handle new cases in the same way. If this strikes us as mysterious, it is a symptom of our confusion. We are trying to imagine that the future steps are taken in the mind, "in a queer sense," before they are taken in reality—as if the mind were a machine that already contained its future movements (*Investigations,* Secs. 193–195).

Wittgenstein was saying that our understanding of a rule is not a state that forces us to apply the rule in a particular way. Someone who has received the ordinary instruction in arithmetic or chess and has applied it normally in the past could go on in the future in a different way but still be a rational person. Perhaps he could even give a reasonable defense of his divergence.

If this is true, it makes it seem that there are no rules, for a rule forbids some things and requires others. It appears that anything goes, anything can be justified. But then understanding, meaning, language itself all crumble away because they imply rules.

Wittgenstein was not denying, however, that there are rules and that we follow them. He held that the way a rule is applied in particular cases determines its meaning. A rule, as it is formulated in a sentence, "hangs in the air" (*Investigations,* Sec. 198). What puts it on the ground, gives it content, is what we say and do in actual cases. And on this there is overwhelming agreement: we nearly always say and do the same. It is this agreement that determines whether a particular action is in accordance with a rule. Rather than to say that we agree because we follow rules, it is more perceptive to say that our agreement fixes the meaning of the rules, defines their content. In a sense the content of the rules grows as our practice grows. Instead of thinking of humankind as coerced by the rules of logic and mathematics, we should consider that human practice establishes what the rules are.

PRIVATE RULES. The idea that the content of a rule can be fixed only by a practice provides a transition to one of the most subtle topics of the *Investigations,* namely the treatment of "private language." The conception that a significant sentence is a picture was replaced in Wittgenstein's thought by the conception that the sense of a sentence is determined by the circumstances in which it is uttered. Swinging a stick is a strike and pushing a piece of wood is a move—in the circumstances of games. Likewise, saying some words is making a decision—in certain circumstances. In one set of circumstances saying a particular sentence would be asserting something; in other circumstances saying those same words would be asking a question; in still others it would be repeating what someone had said.

This is a difficult conception to grasp. We feel a strong inclination to say that the only thing that determines the sense of what someone says is what goes on in his mind as he says it. As John Locke put it, "Words, in their primary or immediate signification, stand for nothing but *the ideas in the mind of him that uses them.*" Whether some words you uttered expressed a question or an assertion is solely a matter of whether there was a question or an assertion in your mind. What the occasion was, what happened before and after, what persons were present—those circumstances are irrelevant to the sense of your words. The only "circumstance" that matters is the mental occurrence at the time of utterance.

Wittgenstein fought hard and resourcefully against this objection. One technique he used was to describe different cases of deciding, asserting, intending, expecting, and so on. The purpose of this was to show that when one utters some words that express, for instance, a decision, one cannot pick out anything that occurred (for example, a thought, an image, some spoken words, a feeling) such that one wants to call that the act of deciding.

This technique, although powerful, may provoke the response that the only thing proved is the intangibility, the indescribability, of the mental phenomenon in question. William James remarked about the intention of saying a thing before one has said it: "It is an entirely definite intention, distinct from all other intentions, an absolutely distinct state of consciousness, therefore; and yet how much of it consists of definite sensorial images, either of words or of things? Hardly anything!" This intention has "a nature of its own of the most positive sort, and yet what can we say about it without using words that belong

to the later mental facts that replace it? The intention *to-say-so-and-so* is the only name it can receive" (*Principles of Psychology,* New York, 1890, Vol. I, p. 253). Likewise, the decision to stay an hour longer cannot be expressed in any other words than those, yet it is a quite definite mental occurrence; one knows it is there!

Wittgenstein opposed this conception not with further description but with an argument. It is the following: If a decision or expectation or sensation were a state or event that was logically independent of circumstances, then no one, not even the subject of the supposed event, could ever determine that it had occurred. First, how would one learn what, for example, deciding is? Since circumstances are supposed to be irrelevant, one could not learn it by observing other people. Apparently one would have to learn what deciding is from one's own case. But as Wittgenstein remarked: "If I know it only from my own case, then I know only what *I* call that, not what anyone else does" (*Investigations,* Sec. 347). Thus it would be unverifiable whether two people refer to the same phenomenon by the word *deciding.* But worse is to come. One could not even take comfort in the thought "At least I know what *I* call 'deciding.'" You might believe that you have always called the same thing by that name. Yet nothing could determine that this belief was right or wrong. Perhaps the private object constantly changes but you do not notice the change because your memory constantly deceives you (*Investigations,* p. 207)! The idea that you might have a language with logically private rules—that is, rules that only you could understand because only you could know to what the words refer—is a self-contradictory idea. Following a rule implies doing the same, and what "the same" is can only be defined by a practice in which more than one person participates.

Wittgenstein's rejection of the intrinsically private, inner object is a consequence of his new conception of meaning. Language requires rules, and following a rule implies a customary way of doing something. It could not be that only once in the history of humankind was a rule followed (Sec. 199). An expression has a meaning only if there is a regular, a uniform, connection between saying the expression and certain circumstances. When we call something measuring, for example, a part of the uniformity we require is a constancy in the results of measurement (Sec. 242). A person can be guided by a signpost only if there is a regular way of responding to signposts. The meaning of an expression is its use—that is to say, the language game in which it occurs—that is to say, the uniform relation of the expression to certain circumstances. Wittgenstein made explicit the connection between this view of the nature of meaning and his attack on "private" mental contents when he said that following a rule is a practice and therefore one cannot follow a rule "privately" (Sec. 202).

See also Brouwer, Luitzen Egbertus Jan; Existence; James, William; Keynes, John Maynard; Language; Logical Positivism; Logic, History of; Mathematics, Foundations of; Moore, George Edward; Number; Ostwald, Wilhelm; Proper Names and Descriptions; Propositions; Ramsey, Frank Plumpton; Rilke, Rainer Maria (René); Russell, Bertrand Arthur William; Schlick, Moritz; Thinking; Tolstoy, Lev (Leo) Nikolaevich; Volition; Wright, Georg Henrik von.

Bibliography

WORKS BY WITTGENSTEIN

Tractatus Logico-Philosophicus. London: Routledge, 1922. Contains the German text of *Logisch-Philosophische Abhandlung,* with English translation on facing pages, and an introduction by Bertrand Russell. Republished with a new translation by D. F. Pears and B. F. McGuinness. London: Routledge and Paul, 1961.

"Some Remarks on Logical Form." *PAS* supp. 9 (1929): 162–171.

Philosophical Investigations. Edited by G. E. M. Anscombe, Rush Rhees, and G. H. von Wright, translated by G. E. M. Anscombe. Oxford: Blackwell, 1953. Contains German text of *Philosophische Untersuchungen,* with English translation on facing pages.

Remarks on the Foundations of Mathematics. Edited by G. E. M. Anscombe, Rush Rhees, and G. H. von Wright, translated by G. E. M. Anscombe. Oxford: Blackwell, 1956. Parallel German and English texts.

The Blue and Brown Books: Preliminary Studies for the Philosophical Investigations. Oxford: Blackwell, 1958. With a preface by Rush Rhees.

Notebooks 1914–1916. Edited by G. E. M. Anscombe and G. H. von Wright, translated by G. E. M. Anscombe. Oxford: Blackwell, 1961. Parallel German and English texts.

Philosophische Bemerkungen. Frankfurt, 1964.

"A Lecture on Ethics." *Philosophical Review* 74 (1965). A paper delivered at Cambridge in 1929 or 1930.

WORKS ON WITTGENSTEIN

Biography

Gasking, D. A. T., and A. C. Jackson. "Ludwig Wittgenstein," memorial notice. *Australasian Journal of Philosophy* 29 (1951).

Malcolm, Norman. *Ludwig Wittgenstein: A Memoir.* London: Oxford University Press, 1958.

Russell, Bertrand. "Ludwig Wittgenstein," memorial notice. *Mind* 60 (1951).

Ryle, Gilbert. "Ludwig Wittgenstein," memorial notice. *Analysis* 12 (1951).

Von Wright, G. H. "Ludwig Wittgenstein: A Biographical Sketch," Malcolm, op. cit.

Wisdom, John. "Ludwig Wittgenstein, 1934–1937." *Mind* 61 (1952).

Tractatus

Anscombe, G. E. M. *An Introduction to Wittgenstein's Tractatus.* London: Hutchinson University Library, 1959.

Black, Max. *A Companion to Wittgenstein's Tractatus.* Ithaca, NY: Cornell University Press, 1964.

Colombo, G. C. M. Critical introduction and notes to the Italian translation of the *Tractatus.* Milan and Rome, 1954.

Daitz, E. "The Picture Theory of Meaning." In *Essays in Conceptual Analysis,* edited by Antony Flew. London: Macmillan, 1956.

Griffin, James. *Wittgenstein's Logical Atomism.* Oxford: Clarendon Press, 1964.

Hartnack, Justus. *Wittgenstein og den moderne Filosofi.* Copenhagen, 1960. Translated into German as *Wittgenstein und die moderne Philosophie.* Stuttgart, 1962.

Hintikka, Jaakko. "On Wittgenstein's Solipsism." *Mind* 67 (1958).

Keyt, David. "A New Interpretation of the *Tractatus* Examined." *Philosophical Review* 74 (1965).

Keyt, David. "Wittgenstein's Notion of an Object." *Philosophical Quarterly* 13 (1963).

Maslow, Alexander. *A Study in Wittgenstein's Tractatus.* Berkeley: University of California Press, 1961.

Pitcher, George. *The Philosophy of Wittgenstein.* Englewood Cliffs, NJ: Prentice-Hall, 1964.

Ramsey, F. P. Critical notice of the *Tractatus.* In *The Foundations of Mathematics,* edited by R. F. Braithwaite. London: K. Paul, Trench, Trubner, 1931.

Rhees, Rush. "Miss Anscombe on the *Tractatus.*" *Philosophical Quarterly* 10 (1960).

Rhees, Rush. "The *Tractatus*: Seeds of Some Misunderstandings." *Philosophical Review* 72 (1963).

Russell, Bertrand. "The Philosophy of Logical Atomism." In *Logic and Knowledge,* edited by R. C. Marsh, London: Allen and Unwin, 1956.

Schlick, Moritz. "The Turning Point in Philosophy." In *Logical Positivism,* edited by A. J. Ayer. Glencoe, IL: Free Press, 1959.

Stenius, Eric. *Wittgenstein's Tractatus.* Ithaca, NY: Cornell University Press, 1960.

Wienpahl, Paul D. "Wittgenstein and the Naming Relation." *Inquiry* 7 (1964).

Blue and Brown Books, Investigations, Remarks on the Foundations of Mathematics

Albritton, Rogers. "On Wittgenstein's Use of the Term 'Criterion.'" *Journal of Philosophy* 56 (22) (1959).

Ayer, A. J. "Privacy." *Proceedings of the British Academy* 45 (1959).

Ayer, A. J., and Rush Rhees. "Can There Be a Private Language?" PAS supp. 28 (1954). Symposium.

Bambrough, Renford. "Universals and Family Resemblances." *PAS* 61 (1960–1961).

Bouwsma, O. K. "The Blue Book." *Journal of Philosophy* 58 (6) (1961).

Buck, R. C. "Non-other Minds." In *Analytical Philosophy,* edited by R. J. Butler. Oxford: Blackwell, 1962.

Carney, J. D. "Private Language: The Logic of Wittgenstein's Argument." *Mind* 69 (1960).

Cavell, Stanley. "The Availability of Wittgenstein's Later Philosophy." *Philosophical Review* 71 (1962).

Chihara, C. S. "Mathematical Discovery and Concept Formation." *Philosophical Review* 72 (1963).

Chihara, C. S. "Wittgenstein and Logical Compulsion." *Analysis* 21 (1961).

Cook, J. W. "Wittgenstein on Privacy." *Philosophical Review* 74 (1965).

Cowan, J. L. "Wittgenstein's Philosophy of Logic." *Philosophical Review* 70 (1961).

Dummett, Michael. "Wittgenstein's Philosophy of Mathematics." *Philosophical Review* 68 (1959).

Feyerabend, Paul K. "Wittgenstein's *Philosophical Investigations.*" *Philosophical Review* 64 (1955).

Fodor, J. A., and J. J. Katz. "The Availability of What We Say." *Philosophical Review* 72 (1963).

Gasking, D. A. T. "Avowals." In *Analytical Philosophy,* edited by R. J. Butler. Oxford: Blackwell, 1962.

Hardin, C. L. "Wittgenstein on Private Languages." *Journal of Philosophy* 56 (1959).

Hartnack, Justus. *Wittgenstein og den moderne Filosofi,* see above.

Kreisel, Georg. "Wittgenstein's *Remarks on the Foundations of Mathematics.*" *British Journal for the Philosophy of Science* 9 (1958).

Kreisel, Georg. "Wittgenstein's Theory and Practice of Philosophy." *British Journal for the Philosophy of Science* 11 (1960).

Malcolm, Norman. "Behaviorism as a Philosophy of Psychology." In *Behaviorism and Phenomenology,* edited by T. W. Wann. Chicago: William Marsh Rice University by the University of Chicago Press, 1964.

Malcolm, Norman. *Dreaming.* New York: Humanities Press, 1959.

Malcolm, Norman. *Knowledge and Certainty.* Englewood Cliffs, NJ: Prentice-Hall, 1963. See "Wittgenstein's *Philosophical Investigations*" and "Knowledge of Other Minds."

Moore, G. E. "Wittgenstein's Lectures in 1930–1933." In *Philosophical Papers.* New York: Macmillan, 1959.

Nell, E. J. "The Hardness of the Logical 'Must.'" *Analysis* 21 (1961).

Pitcher, George. *The Philosophy of Wittgenstein.* Englewood Cliffs, NJ: Prentice-Hall, 1964.

Rhees, Rush. Preface to *The Blue and Brown Books.* Oxford: Blackwell, 1958.

Rhees, Rush. "Some Developments in Wittgenstein's View of Ethics." *Philosophical Review* 74 (1965).

Rhees, Rush. "Wittgenstein's Builders." *PAS* 60 (1959–1960).

Strawson, P. F. Critical notice of *Philosophical Investigations. Mind* 63 (1954).

Thomson, Judith J. "Private Languages." *American Philosophical Quarterly* 1 (1964).

Waismann, Friedrich. "Notes on Talks with Wittgenstein." *Philosophical Review* 74 (1965).

Wellman, Carl. "Our Criteria for Third-Person Psychological Sentences." *Journal of Philosophy* 58 (1961).

Wellman, Carl. "Wittgenstein and the Egocentric Predicament." *Mind* 68 (1959).

Wellman, Carl. "Wittgenstein's Conception of a Criterion." *Philosophical Review* 71 (1962).

Wisdom, John. "A Feature of Wittgenstein's Technique." *PAS* supp. 35 (1961).

Wisdom, John. *Philosophy and Psychoanalysis.* Oxford, 1953.

Norman Malcolm (1967)

WITTGENSTEIN, LUDWIG JOSEF JOHANN [ADDENDUM 1]

Of Ludwig Josef Johann Wittgenstein's philosophical writings available in print, by far the greater part was published after the 1967 *Encyclopedia of Philosophy.* The year 1967 also saw the publication on microfilm of Wittgenstein's *Nachlass.* In addition to the *Nachlass* itself and the posthumously published material from it, there has become available since 1967 a considerable body of Wittgenstein's letters, records of conversations with him, and notes taken by students at his lectures. Altogether, vastly more material is available to the student of Wittgenstein than there was in the mid-1960s. The *Tractatus* and the *Philosophical Investigations* remain, however, the central works for anyone trying to understand Wittgenstein's philosophy. The other writings do give a far fuller understanding of how Wittgenstein's later thought developed; they make clear important continuities between earlier and later work that had been difficult to see earlier. The recognition of these continuities can, for example, be seen in several of the essays in Peter Winch (1969), including Winch's own introductory essay on the unity of Wittgenstein's philosophy. Hidé Ishiguro (1969), in that volume, established that Wittgenstein's connection between meaning and use was not new in his later philosophy. He had always tied meaning to use; what was new in the later work, Ishiguro argued, was the willingness to consider a great variety of different kinds of use besides stating of facts; and Winch notes also the importance in Wittgenstein's later work of the idea that what we call "stating a fact" can itself be many different sorts of thing. A very important continuity noted by Anthony Kenny (1973) lies in Wittgenstein's conception of philosophy itself, including the contrast he made between philosophy and natural science, and the central role he gave to descriptions (rather than proofs) within philosophy.

The material written in the late 1940s and just before Wittgenstein's death shows how Wittgenstein's thought developed after the completion of what was published as Part I of *Philosophical Investigations.* He mentioned to friends his intention (never carried out) of replacing much of what is in the last thirty pages or so of Part I with what is in Part II, along with related material (subsequently published as *Remarks on Philosophical Psychology* and *Last Writings on Philosophical Psychology,* Vol. 1). His comment helps make clear how he saw the investigations of psychological concepts that occupy so much of Part II of the *Investigations* and of the related manuscripts. He is not turning away from the central questions about language in the *Investigations* to new and unrelated topics. Those questions themselves led him repeatedly into detailed examination of such matters as how what is going on in our minds bears on whether we speak with understanding or rather only as parrots might. The late writings show also his concern with the question, important to him from the 1930s onward, how what is given in experience is relevant to the concepts we grasp. These issues are closely related also to the investigations in *Remarks on Colour* (1977), drawn from manuscripts from the last eighteen months of Wittgenstein's life.

Wittgenstein was greatly stimulated by G. E. Moore's attempts to reply to skeptical arguments by asserting things he took it to be plain that he knew (for example, that Earth had existed for a long time) and by Moore's discussion of the paradoxical character of saying "I believe he has gone out, but he has not." Moore's paradox about belief provides a focus for some of Wittgenstein's discussions of psychological concepts in Part II of the *Investigations* and the related manuscripts. Moore's commonsense response to skepticism provided the impetus for Wittgenstein's treatment of skepticism and knowledge in *On Certainty.* He criticized Moore for having misunderstood the concept of knowledge on the model of that of belief and doubt; and indeed *On Certainty* is to some degree continuous with Wittgenstein's other discussions of psychological concepts. But it also stands on its own as an investigation of how certainty forms a part of our various language games and of the role played in those language games by empirical propositions that are not questioned. Wittgenstein's methods in *On Certainty* have been applied by other philosophers in discussions of religious and ethical claims, but he himself does not attempt to apply general principles about doubt, certainty, or knowledge to ethics or religion. (Some of his views about ethics and religion, as well as about art and other topics, have been gathered from various manuscripts and published in *Culture and Value.*)

There is a group of questions about how Wittgenstein saw the relation between facts and the language games in which we are engaged and about how far his

approach, in his later philosophy, involves some kind of idealism or relativism. Do facts exercise any sort of control on the character of our concepts? If there were people who engaged in language games very different from ours—if there were, for example, people who thought one could travel to the moon while in a dream—would we be in a position to criticize such people as fundamentally in error? Several of Wittgenstein's works published after 1967 are particularly relevant to these questions, including *On Certainty, Zettel* (a collection of remarks Wittgenstein had cut from various manuscripts, mostly from the late 1940s), and Wittgenstein's "Remarks on Frazer's *Golden Bough*" (included in Wittgenstein, 1993). Wittgenstein's discussions of mathematics also bear directly on the question how free we are in our development of concepts: What would we be getting wrong if our mathematics, or our logic, were very different? In these discussions Wittgenstein is frequently responding to Gottlob Frege's conception of objectivity in logic and mathematics.

RECEPTION OF WITTGENSTEIN'S PHILOSOPHY

Philosophers are far from agreement on how Wittgenstein's philosophical achievements can be assimilated or indeed whether they should be. There are many philosophers who regard Wittgenstein's influence as pernicious and who think that the best response to his philosophy is to ignore it. This view rests sometimes on the idea that his philosophy developed to meet his personal needs and is irrelevant to the genuine interests of contemporary philosophy. A second kind of response to Wittgenstein involves making a sharp distinction between, on the one hand, the important philosophical claims and arguments that are thought to be in his work or implied by it and, on the other, his own understanding of his philosophy as not involving disputable theses or explanations and as aiming to dissolve philosophical problems rather than to find the correct answers. If that distinction is made, it may then be held that we should simply ignore his views about philosophy (which it may also be held are inconsistent with his own practice) and should instead pay attention to the theses and arguments (on which, on this view, his reputation must properly rest). Philosophers who read Wittgenstein in this way do not agree among themselves whether the theses in question are true, the arguments sound; nor do they agree about what the extractible theses are supposed to be. Thus, for example, those who ascribe to him theses about the necessary conditions for a language disagree about whether these conditions include the necessity that a speaker of any language have been at least at some time a member of a community of speakers. A third distinct kind of response to Wittgenstein takes seriously his conception of philosophical problems as dependent upon our misunderstandings of the workings of our language; they arise when language is allowed to go "on holiday." And so any adequate approach to these problems depends on coming to see how we are led into them; it will not issue in solutions that leave unchanged our idea of the problems themselves. Finally, some elements of Wittgenstein's approach to philosophical problems, and his criticisms of standard philosophical moves in response to them, have also been treated as important and interesting by those who, like Richard Rorty, wish to see analytical philosophy replaced by some other kind of intellectual activity.

The philosophical disputes about Wittgenstein's work have been focused to a considerable degree on the issues discussed by Norman Malcolm in the original *Encyclopedia* piece, including the relation between meaning and use, the possibility of a private language, and the objectivity of rules. Much recent controversy has been inspired by the writings of Michael Dummett and Saul Kripke. Dummett reads Wittgenstein as putting forward an antirealist theory of meaning; Kripke has argued that Wittgenstein in the *Investigations* presents a new skeptical problem and a skeptical solution to it. Responses to Dummett and Kripke have made clear the importance of understanding Wittgenstein's aims, his desire to show how our misconceptions can make something perfectly ordinary appear problematic; thus, it is the step in our arguments at which the ordinary first appears problematic that we fail to note, and to which we need to attend.

See also Analysis, Philosophical; Dummett, Michael Anthony Eardley; Frege, Gottlob; Kripke, Saul; Malcolm, Norman; Meaning; Moore, George Edward; Philosophy; Rorty, Richard; Skepticism, History of.

Bibliography

WORKS BY WITTGENSTEIN

Philosophical Remarks. Edited by R. Rhees; translated by R. Hargreaves and R. White. Oxford, 1964; 2nd ed., Oxford: Blackwell, 1975.

Zettel. Edited by G. E. M. Anscombe and G. H. von Wright; translated by G. E. M. Anscombe. Oxford: Blackwell, 1967; 2nd ed., 1981.

On Certainty. Edited by G. E. M. Anscombe and G. H. von Wright; translated by G. E. M. Anscombe and D. Paul. Oxford: Blackwell, 1969.

Proto-tractatus. Edited by B. F. McGuinness, T. Nyberg, and G. H. von Wright; translated by D. F. Pears and B. F. McGuinness. Ithaca, NY: Cornell University Press, 1971.

Philosophical Grammar. Edited by R. Rhees; translated by A. Kenny. Oxford: Blackwell, 1974.

Remarks on Colour. Edited by G. E. M. Anscombe; translated by L. McAlister and M. Schättle. Oxford: Blackwell, 1977.

Remarks on the Foundations of Mathematics. Edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe; translated by G. E. M. Anscombe, 3rd ed. Cambridge, MA: MIT Press, 1978.

Culture and Value. Edited by G. H. von Wright; translated by P. Winch. Oxford: Blackwell, 1980.

Remarks on the Philosophy of Psychology. Vol. 1, edited by G. E. M. Anscombe and G. H. von Wright; translated by G. E. M. Anscombe. Vol. 2, edited by G. H. von Wright and H. Nyman; translated by C. G. Luckhardt and M. A. E. Aue. Chicago: University of Chicago Press, 1980.

Last Writings on the Philosophy of Psychology. Edited by G. H. von Wright and H. Nyman; translated by C. G. Luckhardt and M. A. E. Aue. 2 vols. Chicago: University of Chicago Press, 1982–1992.

Werkausgabe. 8 vols. Frankfurt, 1989.

Philosophical Occasions, 1912–1951. Edited by J. Klagge and A. Nordmann. Indianapolis: Hackett, 1993. Contains all Wittgenstein's shorter published writings; some letters and records of lectures; also a full account of the Wittgenstein *Nachlass.*

Wiener Ausgabe. Edited by M. Nedo. 22 vols. Vienna: Springer-Verlag, 1993–.

The Published Works of Ludwig Wittgenstein. Edited by H. Kaal and A. McKinnon. Electronic text database.

LECTURES AND CONVERSATIONS

Wittgenstein's Lectures, Cambridge, 1930–1932. Edited by D. Lee. Totowa, NJ: Rowman and Littlefield, 1980.

Wittgenstein's Lectures, Cambridge, 1932–1935. Edited by A. Ambrose. Totowa, NJ: Rowman and Littlefield, 1979.

Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge, 1939. Edited by C. Diamond. Ithaca, NY: Cornell University Press, 1976; Chicago: University of Chicago Press, 1989.

Wittgenstein's Lectures on Philosophical Psychology 1946–47. Edited by P. T. Geach. Chicago: University of Chicago Press, 1988.

Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann. Edited by B. F. McGuinness; translated by J. Schulte and B. F. McGuinness. New York: Barnes and Noble, 1979.

WORKS ON WITTGENSTEIN

Anscombe, G. E. M. "The Question of Linguistic Idealism." In *From Parmenides to Wittgenstein.* Minneapolis: University of Minnesota Press, 1981.

Canfield, J. V., ed. *The Philosophy of Wittgenstein.* 15 vols. New York: Garland, 1986. Comprehensive collection of over 250 articles.

Cavell, S. *The Claim of Reason.* Oxford: Clarendon Press, 1979.

Conant, J. "Kierkegaard, Wittgenstein, and Nonsense." In *Pursuits of Reason,* edited by T. Cohen et al. Lubbock: Texas Tech University Press, 1993.

Diamond, C. *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind.* Cambridge, MA: MIT Press, 1991.

Dummett, M. *Truth and Other Enigmas.* London: Duckworth, 1978.

Goldfarb, W. "I Want You to Bring Me a Slab: Remarks on the Opening Sections of the *Philosophical Investigations.*" *Synthese* 56 (1983).

Goldfarb, W. "Kripke on Wittgenstein on Rules." *Journal of Philosophy* 82 (1985).

Hacker, P. M. S. *Insight and Illusion: Themes in the Philosophy of Wittgenstein.* Oxford, 1972; rev. 2nd ed., Oxford: Clarendon Press, 1986.

Holtzman, S. H., and Leich, C. M., eds. *Wittgenstein: To Follow a Rule.* London: Routledge and Kegan Paul, 1981.

Ishiguro, H. "Use and Reference of Names." In *Studies in the Philosophy of Wittgenstein,* edited by P. Winch. London: Routledge and Kegan Paul, 1969.

Kenny, A. *Wittgenstein.* Cambridge, MA: Harvard University Press, 1973.

Kripke, S. *Wittgenstein on Rules and Private Language.* Cambridge, MA: Harvard University Press, 1982.

Malcolm, N. *Nothing Is Hidden: Wittgenstein's Criticism of His Early Thought.* Oxford: Blackwell, 1986.

McDowell, J. *Selected Papers,* Vol. 1. Cambridge, MA, 1996.

McGuinness, B. F. *Wittgenstein, a Life: Young Ludwig: 1889–1921.* Berkeley: University of California Press, 1988.

Monk, R. *Ludwig Wittgenstein: The Duty of Genius.* New York: Free Press, 1990.

Pears, D. F. *The False Prison.* 2 vols. Oxford: Clarendon Press, 1987–1988.

Rhees, R. *Discussions of Wittgenstein.* London: Routledge and Kegan Paul, 1970.

Shanker, S. G., ed. *Ludwig Wittgenstein: Critical Assessments.* 4 vols. London: Croom Helm, 1986. Comprehensive collection of 104 articles.

Shanker, V. A., and S. G. Shanker. *A Wittgenstein Bibliography.* London: Croom Helm, 1986. Covers primary sources and over 5,400 items on Wittgenstein.

Sluga, H., and D. Stern, eds. *Cambridge Companion to Wittgenstein.* Cambridge, U.K.: Cambridge University Press, 1996. Includes a full bibliography of Wittgenstein's writings and good selective bibliography of secondary literature.

Winch, P., ed. *Studies in the Philosophy of Wittgenstein.* London: Routledge and Kegan Paul, 1969.

Wright, C. *Wittgenstein on the Foundations of Mathematics.* Cambridge, MA: Harvard University Press, 1980.

Cora Diamond (1996)

WITTGENSTEIN, LUDWIG [ADDENDUM 2]

Although aesthetics was a subject of deep and lifelong importance to Ludwig Wittgenstein, he wrote very little directly on the topic. He did, however, write remarks on the visual arts, literature and poetry, architecture, and especially music throughout his multifarious writings on

the philosophies of language, mind, mathematics, psychology, and philosophical method. A number of these remarks, including some from his more personal notebooks, are collected in *Culture and Value*, and scholars have the collected notes from a course of lectures he gave in Cambridge in 1938. In those lectures Wittgenstein was quick to differentiate between types of questions, particularly between questions of empirical psychology and aesthetic questions (he said that, while he was interested in scientific issues, only conceptual and aesthetic issues could truly grip him).

He also looked, with at the time unprecedented detail, into the nuances of humankind's actual critically descriptive aesthetic language, showing how remote such context-specific articulations are from questions of the highest level of aesthetic generality, e.g. "What is Beauty?" He also showed how particularized aesthetic judgments can be supported by reasons as they emerge within a particularized context of aesthetic perception and evaluation, but *without* recourse to a more general theory that underwrites the judgment. Wittgenstein also investigated, and underscored the importance of, the contextual backdrop and the artistic tradition from which a work emerges; aesthetic reasoning, he suggested, very often proceeds by comparative juxtaposition, not by a form of deductive argumentation from general principles (and yet it is, in a full-blooded sense, reasoning nonetheless).

Scholars also have the record by G. E. Moore of Wittgenstein's lectures of 1930–1933, a document that has been of particular value to those working in the philosophy of criticism. In them, Wittgenstein made one link between the philosophies of language and of art explicit, developing a similarity between the meaning of the word "game" and the word "art." Like the class of all games, he suggested, art has no single essence, common property, or unitary feature present in all cases and by virtue of which the object in question is justifiably characterized as a work of art. This thought, along with the writings in his *Philosophical Investigations* concerning "family resemblance" concepts, i.e. concepts or classes whose members may exhibit some overlapping characteristics but no one defining feature in common, generated the view (articulated in the writings in the 1950s of Morris Weitz, William Kennick, and others) that art is itself an "open concept."

As such, it would prove intrinsically resistant to any traditional or essence-capturing definition; writers on aesthetics of the period frequently endorsed an "anti-essentialism" on these grounds. But this led, in turn, to the counter-argument (beginning with Maurice Mandelbaum) that the defining feature making essentialistic definition possible after all may not be an exhibited property, specifically that it may be relational in nature (just as it is a relational, ascertainable, and category-membership-determining fact about a person that she is or is not a grandmother, but this will not be a visually discernible or "exhibited" property). This was followed in turn by institutional theories of art (developed, in very different ways, by Arthur Danto and George Dickie, among others) designed to capture art's essence, the single property that at bottom makes it what it is. Debate about the viability, the general applicability, and the degree of illumination provided by such accounts, continues to the present.

Other strands of Wittgenstein's philosophy as they relate to aesthetic considerations have also been taken up since the 1950s and 1960s and continue into the early twenty-first century. These include studies in the 1970s and 1980s of the significance of Wittgenstein's remarks on aspect-perception and "seeing-as" in connection with problems of the visual discernment of representational content in a marked surface (by Richard Wollheim, who amended the concept to that of "seeing-in," and by others) and in connection with the perception of expressive properties and the use of expressive predicates (by Benjamin Tilghman and others). Others have continued to explore areas that extend well beyond the quite narrow issue of definition versus anti-essentialism (mistakenly, and ironically, regarded by many as the essence of the significance of Wittgenstein's later philosophical writings for aesthetic understanding). These include studies, in the 1990s to the 2000s, of the significance of Wittgenstein's remarks on "language-games" and a "form of life" in his philosophy of language for literary language as well as, conversely, the value of literary cases for work in the philosophy of language, studies of his remarks on music, studies of the complex interrelations between philosophical conceptions of linguistic meaning and aesthetic theory, studies of the relations between ethical and aesthetic values, studies of the legacy of romanticism in relation to Wittgenstein's later thought, studies of Wittgenstein's writings on self-reference and self-description for questions concerning autobiographical language and self-knowledge, and assessments of Wittgenstein's writings for literary aesthetics. Taken as a whole, late-twentieth-century and early-twenty-first-century work on Wittgenstein's aesthetics has shown that the focus on definition was only one aspect among many.

See also Aesthetics, History of; Art, Expression in; Art, Representation in; Danto, Arthur; Moore, George Edward; Visual Arts, Theory of the; Wollheim, Richard.

Bibliography

WORKS BY LUDWIG WITTGENSTEIN

Philosophical Investigations. 3rd ed. Translated by G. E. M. Anscombe. New York: Blackwell, 1968.

Lectures and Conversations on Aesthetics, Psychology, and Religious Belief. Edited by Cyril Barrett. Berkeley: University of California Press, 1972.

Culture and Value. Edited by G. H. von Wright. Translated by Peter Winch. Chicago: University of Chicago Press, 1980.

WORKS ABOUT LUDWIG WITTGENSTEIN

Allen, Richard, and Malcolm Turvey, eds. *Wittgenstein, Theory, and the Arts*. New York: Routledge, 2001.

Cavell, Stanley. *Must We Mean What We Say?: A Book of Essays*. New York: Cambridge University Press, 1976.

Dauber, Kenneth, and Walter Jost, eds. *Ordinary Language Criticism: Literary Thinking after Cavell after Wittgenstein*. Evanston, IL: Northwestern University Press, 2003.

Eldridge, Richard. *Leading a Human Life: Wittgenstein, Intentionality, and Romanticism*. Chicago: University of Chicago Press, 1997.

Elton, William, ed. *Aesthetics and Language*. Oxford: Blackwell, 1954.

Hagberg, Garry L. *Art as Language: Wittgenstein, Meaning, and Aesthetic Theory*. Ithaca, NY: Cornell University Press, 1995.

Hagberg, Garry L. *Meaning and Interpretation: Wittgenstein, Henry James, and Literary Knowledge*. Ithaca, NY: Cornell University Press, 1994.

Kennick, William E. "Does Traditional Aesthetics Rest on a Mistake?" *Mind* 67 (3) (July 1958): 317–334.

Lewis, Peter B., ed. *Wittgenstein, Aesthetics, and Philosophy*. Burlington, VT: Ashgate, 2004.

Mandelbaum, Maurice. "Family Resemblances and Generalization concerning the Arts." *American Philosophical Quarterly* 12 (3) (1965): 219–228.

Moore, George Edward. "Wittgenstein's Lectures, 1930–1933." In *Philosophical Papers*, 252–324. London: Allen & Uwin, 1959.

Tilghman, B. R. *But Is It Art? The Value of Art and the Temptation of Theory*. Oxford and New York: Basil Blackwell, 1984.

Tilghman, B. R. *Wittgenstein, Ethics, and Aesthetics: The View from Eternity*. Albany: State University of New York Press, 1991.

Weitz, Morris. "The Role of Theory in Aesthetics." *Journal of Aesthetics and Art Criticism* 15 (1) (September 1956): 27–35.

Wollheim, Richard. *Arts and Its Objects*. New York: Cambridge University Press, 1992.

Garry L. Hagberg (2005)

WODEHAM, ADAM

(c. 1298–1358)

Adam Wodeham studied theology with Walter Chatton. The man he held in high esteem, his friend and mentor, was, however, William Ockham. All three men were Englishmen and fellow Franciscans. But whereas Chatton systematically opposed Ockham's views, Wodeham rose to Ockham's defense. As a teacher of theology himself, Wodeham lectured on Peter Lombard's *Sentences*. He did so three times, in London, Norwich (c. 1330), and Oxford (1332). The text of only the last two lectures survive, and only the second has been printed in a modern critical edition. Wodeham developed his own philosophical and theological doctrines by rethinking those of Ockham, some of which he considerably altered. This entry mentions only his most original contributions to philosophy proper.

LANGUAGE AND THOUGHT

Wodeham agreed with Ockham that the languages humans speak derive their meaningfulness from an intrinsically significant mental language, common to all intellects. The terms of that language are concepts. Concepts are acts of apprehending individual things. Some are singular, by which a given individual thing is apprehended, as when we see a thing or remember one we have earlier seen. Others are general, as, for example, the concept corresponding to the word "rose," by which we apprehend all actual and possible roses indiscriminately. Mental sentences too are acts of apprehension. When we form a mental sentence, however, we apprehend a thing of a different sort, Wodeham thought, namely a state of affairs. For example, a rose being a flower is apprehended not by a concept, but by the mental correlate of "a rose is a flower." Concepts and mental sentences are to be regarded as signifying those very things we apprehend by them.

ONTOLOGY

Wodeham's ontology is thus twofold. It contains a restricted ontology of concrete individuals, a strictly nominalist ontology, but in its full extension it also includes states of affairs, and therefore abstract things. Accordingly, Wodeham regarded words such as "being," "thing" and "something" as having two senses. In one sense of "thing," only concrete individuals, actual or possible, are things. In another sense, states of affairs, though they are abstract entities, are things, whether they obtain or can obtain, or not. Wodeham recognized both affirmative and negative states of affairs. Discussing Augustine, he remarks that the person who prefers not to exist over existing in misery can be correctly described as preferring one thing over another, though both things are states of affairs, one negative, the other affirmative.

BELIEF AND KNOWLEDGE

Much of our intellectual activity consists in forming beliefs. We form a belief when we judge a state of affairs to obtain. We cannot form a belief, then, unless we first form a mental sentence by which we apprehend the relevant state of affairs. In some cases, it appears to us that the state of affairs we are considering obtains. The mental sentence by which we are apprehending it is then called "evident." Whenever we form an evident mental sentence, we tend to judge accordingly. There are, however, as Wodeham notes, degrees of evidence. At its lower degree, the evidence of a mental sentence is potentially outweighed by reasons we have or might have to dissent or doubt. We then judge accordingly only if we fail to bring these reasons to mind. At its higher degree, by contrast, the evidence of a mental sentence cannot be outweighed by any reasons to the contrary, and we are therefore compelled to judge accordingly. The sentence "If equals are subtracted from equals, the remainders are equal" has this degree of evidence, whereas the sentence "This boat is moving" has the lower degree of evidence. Wodeham assumed that if a mental sentence has the higher degree of evidence, its truth is guaranteed. On this assumption, he rules that only beliefs caused by mental sentences that have the higher degree of evidence (or that follow just as evidently from such sentences) are acts of knowledge. All other beliefs, whatever their cause, are matters of fallible opinion or perhaps of faith, but not of knowledge.

INFLUENCE

Wodeham's views, in particular on ontology, were extremely influential. In reaction to them, Parisian scholars of the mid-fourteenth century divided into two camps: those who recognized states of affairs and those who denied them. John Buridan was their most prominent opponent. He rejected, therefore, Wodeham's semantics of sentences, though not his semantics of terms. Authors who recognized the existence of states of affairs in addition to that of concrete individuals include Gregory of Rimini and Nicolas Oresme.

See also Chatton, Walter; Ockhamism; William of Ockham.

Bibliography

WORKS BY WODEHAM

Adam Goddam super quattuor libros sententiarum. Edited by John Major. Paris: P. le Preux, 1512. Contains Wodeham's Oxford lectures abbreviated (in c. 1375) by Henry Totting de Oyta, who inserted in the text comments of his own.

Lectura secunda in librum primum sententiarum. Edited by Rega Wood and Gedeon Gál. St. Bonaventure, NY: Franciscan Institute, 1990. Contains Wodeham's Norwich lectures.

Tractatus de indivisibilibus. Edited by Rega Wood. Dordrecht, Netherlands: Kluwer, 1998. Discusses alternative conceptions of the structure of the continuum as they connect with alternative attempts to solve the paradoxes of the infinite.

WORKS ON WODEHAM

Courtenay, William. *Adam Wodeham: An Introduction to His Life and Writings*. Leiden, Netherlands: Brill, 1978. Contains valuable historical and manuscript information.

Gál, Gedeon. "Adam Wodeham's Question on the 'Complexe Significabile' as the Immediate Object of Scientific Knowledge." *Franciscan Studies* 37 (1977): 66–102. This article revealed Wodeham as the source of Gregory of Rimini's theory of propositionally signifiable things—that is, of states of affairs.

Karger, Elizabeth. "Adam Wodeham on the Intentionality of Cognitions." In *Ancient and Medieval Theories of Intentionality*, edited by D. Perler. Leiden, Netherlands: Brill, 2001.

Karger, Elizabeth. "Ockham and Wodeham on Divine Deception as a Skeptical Hypothesis." *Vivarium* 42 (2) (2004): 225–236.

Karger, Elizabeth. "William of Ockham, Walter Chatton and Adam Wodeham on the Objects of Knowledge and Belief." *Vivarium* 33 (2) (1995): 171–196.

Sylla, Edith Dudley. "God, Indivisibles and Logic in the Later Middle Ages: Adam Wodeham's Response to Henry of Harclay." *Medieval Philosophy and Theology* 7 (1) (1998): 69–87.

Elizabeth Karger (2005)

WOLFF, CHRISTIAN
(1679–1754)

Christian Wolff was a rationalist polymath and an influential leader of the early German Enlightenment. He was born in Breslau into an impoverished family of leather workers. In his academic career, he gained renown by teaching mathematics and became famous for systematizing and updating the German philosopher and mathematician Gottfried Wilhelm Leibniz. Wolff pioneered socio-economics, framed the idea of subsidiarity (the EU welfare model), and made lasting contributions to international law. He developed German into a philosophical language (e.g., coining *Begriff*), created a terminology still in use in the twenty-first century (e.g., "monism" and "dualism"), and dominated continental thought before Immanuel Kant in Germany, Switzerland, Poland, Southeast Europe, and Russia. In his philosophical work, he revived ontology as a systematic framework for the

empirical sciences, and expanded the geometric method, a mathematical design for rational thought and conceptual reasoning. He advanced the first formal theory of evolution and defined the ecological and cosmological notion of a world as a network of worldlines (*nexus rerum*). Like Leibniz, he sided with the Jesuit accommodation in the Rites Controversy (1610–1724). Unlike Leibniz, he openly declared himself a neo-Confucian in the textual tradition of Zhu Xi (1130–1200).

This bold move resulted in his exile in 1723 and spawned the Pietism Controversy 1723–1740. His Christian critics denounced him as a pagan, "Spinozist," and atheist, while Thomasius attacked him as a "new, insolent Confucian" in 1726. His pupils lost teaching posts in Prussia and Swabia; his texts were outlawed at Halle in 1723 and in Prussia in 1729. His opponents were Christian fundamentalists influenced by Martin Luther, Philipp Jakob Spener, and John Calvin. They relented in the 1730s, when it became undeniable that Wolff accommodated mainstream opinions and retracted his provocative metaphysical claims. But he never retracted his arguments for academic freedom, especially as a freedom from religious dogma. He was celebrated as "the teacher of Germany" (*praeceptor Germaniae*) who yielded to his critics by choosing Sir Isaac Newton over Leibniz and Christ over Confucius, while preserving the unity of his system of ideas in a reformulated encyclopedic Latin oeuvre.

At Marburg, he served an enlightened Calvinist ruler, the Landgrave of Hesse-Cassel. He was invited to join Utrecht University and to lead Russia's and Prussia's academies. After the coronation of Frederick II, he left for a royal welcome in Berlin in 1740. His return to Halle, which was condoned by the king, was seen as a cultural feat for Prussia and was a legal victory for reason. The elector of the Holy Roman Empire and founder of the Bavarian Academy elevated him to nobility. Baron Wolff died on his estate (*Rittergut*) near Leipzig in 1754. He was the chief German thinker after Leibniz and before Kant.

CONTEXT, WORK, AND IMPACT

Wolff was born January 24, 1679, in the capital of Silesia (Breslau, present-day Wroclaw), in the Protestant northeast of Hapsburg, Austria (present-day Poland). He was the only survivor of six children by a tanner. Following his father's wishes, he attended Breslau's Lutheran School and majored in divinity at Jena in 1699. He changed his course of studies to mathematics and went to Leipzig to earn his magister degree in 1702. With a thesis on ethics according to the mathematical method, he won a magister legens in 1703, entitling him to teach. He taught mathematics as an adjunct professor at Leipzig and joined the staff of *Acta Eruditorum*, the first academic journal in Germany, published in Leipzig. For the *Acta*, he wrote as a specialist in mathematics but soon branched out to other fields, such as military architecture natural history, and natural philosophy. In 1706, for instance, Wolff reviewed the *Optics* (1704) by Newton (1642–1727) and the expanded *True Physics* (1705) by Newton's student John Keill (1671–1721). The Swedish invasion of Saxony in 1706 (Great Northern War 1700–1721) made Wolff leave Leipzig; Gottfried Wilhelm Baron von Leibniz (1646–1716) helped him to find employment at Halle University as a professor of mathematics. In 1709, he established himself as an expert in the quantitative dynamics of gases (with *Aerometry*).

With these credentials in natural philosophy, Wolff taught logic (1709), next ontology, and eventually ethics—in violation of administrative rules, because philosophy classes had been the exclusive turf of the theology faculty. Despite resistance by the Pietist mayor August Hermann Francke (1663–1727) and the evangelical theologian Joachim Lange (1670–1744), Wolff taught outside his area until 1723. In 1709, he was elected to the Royal Society, and in 1711 to the Berlin Academy.

With the four-volume *Foundations of All Exact Sciences* (1710), Wolff made a name for himself as the leading author of up-to-date German textbooks on the new quantitative sciences. In 1711 he wrote an anonymous review of a handbook (1710) by François Noël on China's geography and astronomy and on Chinese measurements for *Acta Eruditorum*. In 1712 he anonymously contributed to *Acta* a review on Alexandre [sic: François] Noël's translation of six Confucian classics. He wrote the four-volume *Elements of Universal Mathematics* (1713–1715), the so-called *German Logic* (*Rational Thoughts on the Forces of the Human Mind*, 1713), and a *Mathematical Dictionary* (1716). Staying in Halle, he declined calls to Marburg (1714), Wittenberg (1715), Jena (1716), and Leipzig (1716). In 1715 he became court councilor (*Hofrat*) and also professor of physics at Halle; Peter I (the Great, 1672–1725) asked him to serve as a tsarist advisor in St Petersburg. In 1718 he defended Confucian secular humanism and supported Chinese morals in *Reason of Wolff's Classes in Mathematics and Global Philosophy* (Ratio praelectionum Wolfianarum [in] mathesin & philosophiam universam).

In 1719 he published *German Metaphysics (Rational Thoughts on God, World, Human Soul, and All Things in General*), his best-known work. It was read as a revolu-

tionary and secular system; it was a best-seller and the program for a new philosophical network. His Swabian pupil Georg Bernhard Bilfinger (1693–1750) called the network the "Leibnizian-Wolffian School Philosophy."

Although this label irritated Wolff, Bilfinger was being honest. Leibniz was Wolff's most famous mentor, from whom he appropriated main ideas of the monadology and natural dynamics. He also followed Leibniz's rational theodicy. Later, however, Wolff's Leibnizian label turned into a misnomer. Spurred into action by the angry ideological critique of these subversive ideas, and their negative repercussions, Wolff spoke out against them and distanced himself from the deeper implications of ideas such as "monad" and "preestablished harmony." Most students who followed him in this moderation fared well nationally. Others, who resisted this about-face and insisted on the revolutionary significance of Leibniz's ideas in their Wolffian integration, found themselves marginalized (even by Wolff) or driven into exile. Bilfinger, exiled to Russia, was the most radical early interpreter who was not rejected by the later Wolff.

The Leibnizian-Wolffian School Philosophy grew to include female naturalists and free-thinkers, such as the karmic pantheist Johanna Charlotte Unzer (b. Ziegler 1725–1782); among its supporters abroad was the later Newtonian Gabrielle de Châtelet (1706–1749). Early continental feminists celebrated Wolff. Early (male, German) members were known as the textbook authors. The School Philosophy bred a new generation of Enlightenment thinkers, such as the poet and philosopher Johann Christoph Gottsched (1700–1766), and it culminated in the work of Alexander Gottlieb Baumgarten (1714–1762). Baumgarten's *Metaphysica* (1739) was the definitive textbook (used by Kant), and his *Aesthetica* (1750) was the historic Wolffian basis of modern aesthetics.

In 1719 Halle University elected Wolff to serve as its provost (*prorektor*). In 1721 he ended the two-year term with *Speech on the Morals of the Chinese*, a public address to an audience of more than a thousand. He refused to submit the text to the next provost (Lange) for religious scrutiny, which prompted the Pietists to conspire at the royal court. Around the same time, Wolff wrote German *Ethics* (1720), *Politics* (1721), *Physics* (1723), *Teleology* (1724), and *Physiology* (1725).

On November 8, 1723, King Frederick William I (r. 1713–1740) sentenced Wolff to death but granted his exile from Prussia if he left within two days. He fled to Marburg, called by Landgrave Charles I of Hesse (1654–1730). He took the mathematics and physics chair held by Denis Papin (1647–c. 1712), who had co-invented the steam engine with Leibniz (1690). Tsarina Catherine (1684–1727, Empress 1725) offered Wolff the vice presidency of the Russian Academy (in 1723 and 1725). By 1728 his fame had vastly increased the student numbers at Marburg, but he remained a target of Pietists and Calvinists.

Wolff qualified his early liberal challenges in detailed replies to critics (*Schutzschriften* to Lange and Johann Budde in 1724; *Notes to Tübingen Theology* in 1725). He moderated his secular ontology with *Comments to German Metaphysics* in 1724, published his own edition of the speech on Confucius (*Oratio de Sinarum philosophia practica* in 1726, with Bilfinger), and fought for academic freedom (*Preliminary Discourse on Philosophy* in 1728). In 1729 fundamentalists succeeded in having all his works declared illegal in Prussia.

While Wolff taught in Hesse, he was made honorary professor of the Russian Academy at St. Petersburg in 1725. Writing now for a wider European audience, he reformulated his views in a Latin series, with *Rational Philosophy or Logic* (1728; its preface is the *Preliminary Discourse on Philosophy as Such*, which he expanded into a separate work), followed by *First Philosophy or Ontology* in 1730, *General Cosmology* in 1731, *Empirical Psychology* in 1732, and *Rational Psychology* in 1734. *Natural Theology* (1736–1737) and *Global Practical Philosophy* (1738–1739) completed the group. The Latin series replaced the German textbooks, and the new set reveals his rejection of charges of paganism and "free-thinking." These works allowed Wolff's mainstream academic acceptance.

In 1733, the French Academy elected Wolff to one of its eight foreign members. Lobbied by a Wolffian (a warrior, Prince Leopold of Anhalt-Dessau 1676–1747), Frederick William I of Prussia certified Wolff at Marburg as a state counselor of Hesse, now ruled by Frederick I (1676–1751, king of Sweden since 1720; landgrave of Hesse since 1730). In 1734 Prussia rescinded the 1723 arrest warrant; Frankfurt at the Oder offered him a position; the Prussian Academy offered him the vice-presidency; and Halle University allowed his return. He stayed at Marburg until 1740, with students such as Mikhail Lomonossov (1711–1765), the founder of Moscow University (1755).

In 1740, Frederick II (the Great, 1712–1786) promoted Wolff to Prussian privy counselor, offered him the presidency of the Academy, and welcomed him back to Halle as an interdisciplinary professor of mathematics, law, and public policy. Meanwhile, the Leibnizian-

Wolffian School Philosophy had evolved to the leading cultural movement of the German Age of Reason. With the foundation of debate clubs such as the Society of the Friends of Truth (1736, which coined the slogan *sapere aude!*—dare to understand!) and the creation of a host of journals, the rational matrix of the early Enlightenment framed by Wolff had spread into the civil and public sphere of continental Europe. His students, driven from Prussia, taught in other parts of Germany, in Bavaria, Switzerland, Austria, Italy, and Russia.

The 1740 coronation of Frederick II was a pivotal event in Wolff's lifetime. Frederick was an avowed agnostic, who had been imprisoned by his Pietist father Frederick William I. The coronation of the jailed "atheist" was a triumph for the Enlightenment. Frederick's alliance with Wolff was a cultural feat for Prussia and signaled the better protection of academic freedom, the first political harbinger of Germany's later division of church and state.

Back in Halle, Wolff served as the university chancellor in 1743. There he developed a system of natural law (*Natural Law*, 8 vol., 1740–48) and outlined a theory of international law (*International Law*, 1749), which he grounded on natural law (*Principles of Natural and International Law*, 1750). In 1752 he was elected to the Italian Academy in Bologna. His final works were *Moral Philosophy* (1750–1753) and *Economics* (1754–1755). This late series repeats his early praise for the Mandarin-run welfare state of China as an exemplary administrative framework and informed Prussian political economy until 1786, when Frederick's successor returned to more parochial Lutheran values. Political economy had been taught since the creation of cameral chairs by Frederick William I, for training Prussia's tax revenue administrators (a century before the field was read at Oxford).

On September 10, 1745, Wolff was made imperial baron of the Holy Roman Empire (*Reichsfreiherr*) by his pupil Maximilian Joseph III (1727–1777), the enlightened Bavarian king (elector since 1745), who founded the Academy of Sciences at Munich, which later advanced stellar optics, helioscopy, and spectral analysis (e.g., Fraunhofer, 1814). Wolff acquired the feudal seat Klein-Dölzig in Saxony in 1748 and retired from teaching. He had single-handedly changed the German and East European landscape of legal, secular, and social thought—the thrust of his arguments had been so persuasive that they were seen as mainstream a mere generation after they had been first branded as extreme.

Baron Wolff died on his estate near Leipzig on April 9, 1754. His Leibnizian-Wolffian School, then the popular German philosophy, was already besieged by the critiques of the young Pietist theologian Christian August Crusius (1715–1775), whose philosophical tracts appeared in the 1740s. The Lisbon tsunami (November 1, 1755), the worst tectonic disaster in recorded European history, with 70,000 deaths, was internationally seen as a refutation of Leibniz's theodicy of the "best of all possible worlds" and turned Wolff's metaphysical framework, with its optimistic, anthropocentric outlook, into the butt of skeptical mockery.

Wolff advanced the continental Age of Reason and systematized early modern thought. Georg Friedrich Wilhelm Hegel (1770–1831), Karl Marx (1818–83), and Friedrich Nietzsche (1844–1900) dismissed him as an obsolete thinker. Kant (1724–1804), who called him the greatest of all German philosophers, joined Wolff's metaphysical viewpoint to its logical opposite, Humean skepticism, as the dialectic field for the collective "critical path of reason" (1781). Wolff created the grammar for the social sciences, integrated law and economics, and built the foundation (partly with his work on architecture and design, and partly via Gottsched and Baumgarten) for the later discipline of aesthetics.

INFLUENCES ON WOLFF

The earliest influences informing Wolff's intellectual development were Christian theology and the literary Baroque. His father, Christoph Wolff, had intellectual aspirations, and his family followed the Lutheran faith. His birth place Breslau was multidenominational, a regional result of the settlements after the Thirty Years War (1618–1648). In this Protestant city, which involved western Calvinist and eastern Jewish communities, he attended the Lutheran gymnasium (senior high school or community college) and distinguished himself in debates with students from the Roman Catholic school run by the Jesuit order. Wolff's rector was the poet Gryphius (1616–1684), a Baroque student of Martin Opitz's earlier *Book of German Poetry* (1624). Gryphius worked for a linguistic and cultural renewal of Germany, devastated by the genocide. His critique of protestant Aristotelianism, as a reactionary paradigm, exposed Wolff to problems of scholastic authority and to intolerant flaws in the campus doctrine.

In Jena and Leipzig, Wolff reacted to Gryphius' critique by turning to the so-called renegades of his day, René Descartes (1596–1650), Ehrenfried Walter v. Tschirnhaus (1651–1708), and Leibniz. Wolff proposed settling neo-scholastic issues by constructing a new design of conceptual analysis and logical deduction, which he applied to formal, natural, and moral philoso-

phy. In Jena, he studied the geometric method by Erhard Weigel (1625–1699) and a similar method proposed by Descartes. In Leipzig, he studied Tschirnhaus' art of invention (*ars invenienda*), a version of the geometric method influenced by optical ideas of Baruch Spinoza (1632–1677) and by catoptrics and dioptrics, the calculus of mirror reflection and lens refraction. Tschirnhaus used his art of invention for the reverse chemical engineering of Chinese porcelain (1708). Wolff applied the geometric method to conceptual reasoning, sharing Tschirnhaus' and Spinoza's hope that the free-spirited rational quests for scientific discovery would create civil happiness.

In Leipzig and Halle, Wolff interpreted Leibniz's monadology as a system of reflective substances. These ultimate and indivisible points are nature's energetic sources of material arrays; twenty-first-century scholars might call such monads powerpoints. Wolff shared Leibniz's interest in Chinese ontology and understood this model of reality as a rational matrix of interactive objects. Yet Wolff was not sure about the depth of physical interaction, repeatedly changing his mind over whether the energetic reciprocity of nature extends to the free powerpoints in the foundational Leibnizian monadology.

In Marburg, he rejected Leibniz's preestablished harmony and studied physical influx, a model of causation proposed by the Spanish scholastic Francisco Suárez (1548–1617). In 1724 he argued that influxionist causal processes govern the natural elements, only to change his mind again and to become ultimately noncommittal about any rational account of natural causes.

In 1726 he appropriated the principle of decorum from his ex-colleague Christian Thomasius (1655–1728). For Thomasius, the decorum was the rational ground of any good legislation. Thomasius defined it as the form of fair distribution and equated it with the Golden Rule (using it for legal briefs against witch trials and in defense of free sexual liaisons). Wolff read the principle of decorum in a wider sense, as the basic way of civil progress and as a human mirror of cosmic development. He identified it with the convergent arrows of civilization and evolution that are tipped toward perfection. This near-mystical reading of the decorum Wolff claimed as his own, but he acknowledged its previous account in the Book of Rites (*Li Ji*; especially *Da Xue* or "Great Learning" and *Zhong Yong* or "Doctrine of the Mean"). Wolff's principle of decorum (flat out rejected by Thomasius in 1726) was informed by Bilfinger and by the Jesuits Philippe Couplet, Athanasius Kircher, and Noël.

Wolff was also influenced by Lange, Hugo Grotius (1583–1645), and John Locke (1632–1704). Lange's attacks prompted him to retract some of his ontological claims for a metaphysical skepticism compatible with Lutheran doctrine. Wolff's caution was influenced by Newton's rules for philosophy (1687) and by Locke's empiricism. Locke was systematically used by the Pietists to shore up their fundamentalism against rationalist claims. Yielding to English and Saxon critics, Wolff rejected Leibniz's dynamics for Newton's mechanics, thus supporting the majority opinion of the day. But he did not entirely retract his earlier views. The theory of natural law, as developed above all by Grotius (see the subtitle of Wolff's *Reason* [1718; 2nd ed. 1735]), allowed him to make his rational point, while diplomatically avoiding farther and more controversial implications of the same ideas.

MATHEMATICS AND LOGIC

Wolff's initial series of mathematical works are systematic expositions of the scientific knowledge of the day, reflecting the state of the art in geometry, arithmetic, and algebra, as well as of the newly advanced calculus (following Leibniz, not Newton, as nineteenth-century mathematicians would do after Wolff as well). Wolff's mathematical works (1710–1716) do not give much space to statistics and stochastic. In part, this neglect had a historical reason. The revolutionary advances in the theory of probability (e.g., Jakob Bernoulli's *Ars coniectandi*, 1713) were made when part of this series was already in press. Moreover, the physical significance of probabilistic tools was shown later (e.g., by Daniel Bernoulli's *Hydrodynamica*, 1733), and only after Wolff had published his logics (1713 and 1728). While Johann Bernoulli (1667–1748) had written on waves, curves, and integrals earlier, Wolff apparently did not know what to make of it.

Wolff's methodological ideal is Euclidean geometry, an axiomatic and deductive system, which was to him the perfect science of nature. He trusted that all natural events, however vague, incoherent, or ambiguous they may seem, express invariant rational patterns, which one should be able to determine as clear and distinct truths. Probabilistic tools fail to reveal such geometric exactitude, and this is a sign of the limitation of the tools, and not the real limit of the events modeled by them. Wolff's nature is rationally ordered; its ways are logical; and science is "the art of demonstration" (*Logic* vii § 1).

Wolff's scientific works were without equal; they democratically addressed a general readership and popularized science in Germany. The *Foundations*, for instance, is a survey of mathematics, geography, mechanics, hydraulics, ballistics, war tactics, fortress design (*Fes-*

tungsbau), and civil architecture. These textbooks were used in Germany for decades; in the Balkans, such as Romania, and in Eastern Europe, such as the Ukraine, these texts were taught well into the nineteenth century. Wolff pioneered the distinction of pure and applied research; he stressed their equal significance, and he saw in mathematics the common denominator of all science.

Wolff regards logic as a system of universal relations, in contrast to Thomasius, Locke, and Lange (who looked at logic either with Christian disdain or as synonymous with natural sense). Against Arnauld's *Logic* (1662), Wolff argued that conceptual organization is not just a mnemonic tool or a *palais de mémoire* for arranging and retrieving stores of knowledge, but also the mirror of the order of nature (1713). The function of Wolff's logic is the theoretical clarification of natural data and the practical enlightenment of secular reason.

The early Wolff discussed logic together with psychology (1713); later, he joined logic to ontology (1728). Wolff's logic involves concepts, propositions, and the map of syllogistic arguments. The logic of scientific discovery works with definitions, laws, and experience. Since the truth-content of propositions and their relations reflect the cosmos, truth is inseparable from the order of events in physical and ultimate reality. As science is the art of demonstration, logic is the art of invention (*ars inveniendа*) in scientific work. Propositions can serve as hypotheses that support deductive networks of explanations, and they are also testable. The value of hypotheses depends on experiments. As positive results make hypotheses probable, negative results call them into question; further data will have to determine whether a hypothesis is to be revised or dismissed.

ONTOLOGY AND METAPHYSICS

Wolff described reality as the sum of observable things, whose actions and properties are ordered by small dynamic elements or substances (*Metaphysics*, 1719). The empirical structure is the world, defined as an interactive, developing web of things (*nexus rerum*), whose natural basis is the ontological system of rationally accessible simple elements. The substantial basis and the objective superstructure are a coherent whole, the order of nature.

The order of nature is ruled by the principle of (the impossibility of) contradiction—it is impossible for something to be and not to be at the same time; existential differences emerge only in time. The history of nature is the logical flow of its causal processes; their beginnings and ends differ, but transitions are lawfully harmonized. This causal logic obeys the principle of sufficient reason.

This order covers all reality. Its ontological basis is Leibniz's array of monads, organic, conscious, and indivisible force points, which function as Aristotelian entelechies, a primordial software of elementary action, material trade, and environmental fate. In the naturally evolving cosmos, all stuff, things, minds, and networks integrate in an ultimate harmonious and spiritual rule. Wolff's metaphysics combines ontology with a system of spirits (rational psychology or the "*pneumatic* of minds"), a system of nature (rational cosmology or the "world-doctrine"), and a system of divinity (rational theology or the "natural God-scholarship"). Being, minds, empirical reality, and supreme law are radically unified in an emphatically coherent, intelligible, and predictable order of nature.

Wolff framed this system as a rational reply to the scientific unifications by Nicolas Copernicus (1473–1543), Galileo (1564–1642), Johannes Kepler (1571–1630), and Descartes. This ontology is a conjectured "final theory" for all future research. Its problem is its unity—if the divine law integrates in natural order, then "God" is at risk of becoming Spinoza's *natura naturans* or turning into a cosmic energy flow.

As God is at risk of being merged with the cosmos in Wolff's system, freedom is at risk of being dissolved in a divinely deterministic blueprint of creative processes. For Wolff, any effect results from a prior sufficient reason according to lawful and rational patterns. But if all that happens is in principle predictable, where will this leave spontaneity, or the causation of willful and free actions?

The standard answer—freedom has its seat in the soul—does not quite map onto Wolff's system because of his Leibnizian leanings. Souls are simple substances, and all such monads strive and reflect in an interplay the steps of which are harmoniously preestablished. Christian critics objected that all humans are sinners; that "sinning" means the buck stops with the blameworthy person; and that God, who created persons, gave them free will. But if all personal actions resulted from a preestablished arrangement by God at creation, God would be guilty of human evil, and persons would be wheels in a world-machine (Lange, *Causa Dei*, 1723). This Pietist objection to Wolff's metaphysics was construed as a political charge that soldiers going AWOL cannot be blamed for desertion, which led the Prussian king to look at Wolff as a traitor to be fired, punished, and exiled.

Wolff's revised causal ontology drops the preestablished harmony of elementary souls for the addition of real interactions on the level of monads (henceforth called only "simple substances"; *Comments to German*

Metaphysics, 1724). Since substances are invested with a spontaneous power, they affect one another, and in this sense one soul can freely lead another soul into sin. Substances also affect things, like bodies, and hence souls can freely sin in their embodiments.

Paradoxically, the result of Wolff's revision is an even tighter rational order of the universe—as empirical structures form a *nexus rerum*, their basis is to be explicated as a network of elements or *nexus elementorum* (*Comments*, 1724). Now everything is purposeful. Nature's order has a supreme and final regularity. Apparent flaws, like evil, are transient and local phenomena but are not integral parts of the design; the general thrust of the natural network mirrors a pervasive goal-directedness. For Wolff, the whole creation reflects its first cause, whose effects are always good to its creatures, particular to humans (*German Teleology*, 1724).

But this revision does not let Wolff's metaphysics off the Spinozist hook. For as nature is a lawfully evolving framework, things are always getting better, and there is no need for a meddling celestial God to perform miracles on Earth. Since miracles break the natural flow, the logic of the cosmic order reveals miracles as making a causal mess—so requiring more miracles (*miraculum restitutionis*), ontological cleaning crews that restore the causal order broken by the initial miracle (*German Metaphysics*, 1719 and *Cosmologia Generalis*, 1731). The Christian notion of God, in its Catholic and Lutheran senses, does not "fit" the Wolffian reality of being, whereas a stipulated rational and dynamic wave-front, benevolently "naturing" nature, is its ontological consequence.

Wolff's identification of this dynamic ordering as the principle of decorum, which "waves" micro- and macroscopic worlds along their inexorable ways toward perfection (preface to *Speech on Chinese*, 1726), triggered another evangelical outcry and more charges of Spinozism, paganism, and atheism. Wolff replied by defining this power as "God" in the standard Lutheran sense (*Detailed News*, 1726).

Still, evangelicals objected to this metaphysics; they disliked Wolff's (qualified) embrace of Newton as early as 1719. (Pietists roundly rejected the content of *Principia* until midcentury.) That Wolff integrated the laws of motion, and included the technical concepts of mass and force (*Ontology*, 1729), made his world-idea seem all the more deterministic, material, and machinelike. As his critics reminded him (such as Lange in *Brief Sketch of the Axioms in Wolff's Philosophy Harmful to Natural and Revealed Religion*, 1736), the issue is over the elementary matrix of causal interplays. Just as Leibniz's preestablished harmony invites the problem of freedom vis-a-vis dogmas of sin, Wolff's interacting monads, the *nexus elementorum* of1724, draw this charge from another angle: If all was lawfully ordered, where would this leave room for surprises, or for human willfulness?

Wolff's final revisions amounted to a withdrawal from any causal claims and to a self-imposed silence on the issue of the behavior of elements. He vetoed identifying substances or souls with monads (*General Cosmology*, 1731 §182). The three possible metaphysical explanations of causal phenomena—physical influx, occasionalism, and preestablished harmony—all have their pros and cons, but which one would really be right no one can say (*Rational Psychology*, 1734). Wolff's order of nature, no matter which logical moves he made, kept provoking political and clerical critique. In 1734, he gave up on first causes and on mind-body interactions.

ETHICS AND AESTHETICS

Wolff's epistemological platform is the Cartesian cogito, the living being full of doubts, or the human power for reasoning things out. In reason, helped by experience and observation, one discovers the laws of nature in their present workings and in their evolutionary thrust toward a perfected state. In a historical sense, natural laws are the forms of progressive realization and organization, ultimately of nature itself. In a semantic sense, these laws, in their worked-out patterns, generate ever richer information or essential being, which is the best reality in perfection. In a practical sense, the laws of nature point to the final form of the natural good. Hence Wolff's practical law of nature is divinely inspired, aesthetically ideal, and morally binding.

If one wonders why beauty and the good should come about, Wolff argued, one will see that both are the clear and distinct ideas that prevail in the self-realization of nature's law. Why should a person be moral? By reason one knows "what the law of nature wants to get"; and "therefore a reasonable human being does not need any additional laws," for the progressively perfecting law of nature is humanity's law in light of reason. (*Ethics* § 23 1720; also *Global Practical Philosophy* § 268 1738).

Regardless of which metaphysical theory suits the causation of free actions best, the power of reason can shed light on the natural law and thus enlighten human choices. This law or decorum is the formal pattern of perfection. The idea of perfection is the declared source ("*fons … mea*") of Wolff's entire practical philosophy (as outlined in the preface of his *Moral Philosophy*, 1750).

Conceptually, perfection is the consensus of variety; Wolff defined consensus dynamically, as the interactive trend toward fair trade. Scientifically, in the twenty-first century, Wolff's idea of the naturally self-perfecting consensus is reflected in the ecological understanding of climax communities, environmental integrity, and biological diversity. Practically, for Wolff, perfection is the categorical duty and the moral imperative—do what makes the state of oneself and others more perfect; refrain from making it less perfect. Thus the natural law commands to work out the state of the art of the commerce of living forces, each of which freely wants to realize its material momentum in an ever more complex nature.

Accordingly, good and evil (just like beauty and ugliness) can be defined over their relative degree of systemic perfection—from the perspective of integrity and design, nasty and repugnant events are imperfect. The duty to realize well-ordered frames and a sustainable consensus, no matter its particular instantiation, has political and civil implications.

The enlightened sovereign regards the state like a house that needs to be built in the best way, through an efficient allocation of essential weights, for the sake of maximal strength of the whole. The ruler ought to order and maintain the best administrative design for the common good or the welfare of the people. The welfare state, whose revenues help weaker social groups for the sake of a tighter social contract, is Wolff's design (*Principles of Natural Law*, 1754 § 1022). It is inspired by the form of Mandarin administration under the neo-Confucian Qing rulers (since 1644). Wolff's take on the natural law is also shaped by Thomasius, Grotius, and Samuel Pufendorf (1632–94). The political task of the ruler is formally equivalent to the aesthetic task of a designer or architect. Architecture is Wolff's ideal art (his focus would provoke later aestheticians to criticize Wolff for roundly neglecting poetics). As architecture points to material blueprints of well-ordered frames that efficiently distribute mass in elegant designs, Wolff's intellectual concern is to advance the art's form and make it more of a science (*Universal Mathematics*, 1713).

Wolff is the father of German aesthetics, but he did not develop a specific theory of art. Instead, he laid its consistent foundation in philosophical terms. He argued for two aspects of the mind, cognitive powers and sentient will, and derived knowledge from sensation. The impressed data are ordered by the mind, and this order reveals a form—in the terms of Arthur Schopenhauer (1788–1860), the arena of appearances displays the handwriting of the natural force or will-to-life. The law of this form is the decorum; this law reveals geometrically and naturally elegant shapes. This design guides cosmic processes toward their historical unfolding into a final state of the art.

This metaphysical concept of perfection is a physically constant cosmological operation. In Wolff's reading, this operation is an evolutionary vector of material interplays toward complexity. Material interplays develop as progressive consensual grids, and the decorum is their entelechy: a rational, benevolent, substantial *conatus*. Wolff's principle (*prima principia decori*) is binding for ethics, politics, economics, and social order. As the decorum is evident to the unbiased observer, specific religions can illustrate it, but theology, whatever its type, is not a privileged perspective. Theology is an "art," but playful arts contain superior information only if they evolve into science. Architecture is about the design of material structures. The perception of good design elicits pleasure. In this Wolffian sense, the good and the beautiful do not depend on God's arbitrary will but instead on the rational order of nature. Monotheistic revelation is not needed; reason is enough.

INFLUENCE

The paradox of Wolff's influence is that he was the most successful early modern German thinker while suffering the same fate as Newton, the leading scientist of the era—his declared ideas were so persuasive that they were not just academically successful but also soon perceived as oddly trivial. Progress after Wolff was made by critique, by integrating Wolff's ideas in larger models. But while Newton remains admired, Wolff was forgotten after two generations. Later thinkers, from Kant to Marx, regarded him as part of the establishment that needed to be overcome. As a result of their intellectual impact, Wolff was not taught in the twentieth century.

In the eighteenth century, Wolff completed the step from the early Enlightenment to the apex of the Age of Reason, an age that culminated in the split of church and state (1740) and in the American (1776) and French (1789) revolutions. At the start of the era, witches were burnt; priests, preachers, and feudal lords reigned supreme; and the commoners had little to say. Wolff's political legacy was the influence on the academies of the day of his philosophical reflections on rational design, on logical reasons, and on the civil merit in questioning authority. For Voltaire (1694–1778), Wolff defined the Enlightenment—"*Federico regnante, Wolfio docente* (Frederick reigns; Wolff teaches)."

The integration of Wolff's liberal humanistic ideas in Prussian governance by Frederick the Great played no small part in Prussia's advancement to a world power. Wolff's system engaged Kant and Hegel, and thus ensured the continuity of continental thought from Spinoza to the present. During his lifetime, his followers were the Leibnizian-Wolffian school philosophers, who discussed *German Metaphysics* and organized an academic network. His system became the paradigm of German thought until the rise of Kant's star in the 1780s. Some students deserted to the Pietists and advanced in Halle. Daniel Strähler (1692–1750) criticized Wolff in his *Examination of Wolff's Rational Thoughts* (1723).

Other disciples, who stuck to their guns, were fired and driven out, such as Christian Gabriel Fischer (1686–1751) from Königsberg and all of Prussia (1725). Ludwig Philipp Thümmig (1697–1728) left with Wolff in 1723, went to Cassel (ruled by the Landgrave of Hesse), and published the first exegesis, *Principles of Wolffian Philosophy* (1725–1726). Wolffians gained nationwide appointments and ruled the intellectual field well into the 1770s. Bilfinger, the author of the *Elucidations* (1725), went to Tübingen. Johann Friedrich Stiebritz (1707–1772) taught at Gießen and Frankfurt, and wrote *Wolffian Thought Condensed* (1744–1745). Johann Franz Coing (1725–1792) went to Marburg in 1753 and wrote *System of God, Human Soul, World, and the First Principles of Human Cognition* (1765). The philologist, literary critic, and playwright Gottsched taught ontology at Leipzig and produced with *First Principles of Human Cognition* (1765), the most celebrated interpretation next to Baumgarten's. Johann Peter Reusch, who went to Jena in 1738, followed suit with *Metaphysical System (1734)*.

The works by Friedrich Christian Baumeister (1709–1795) at Wittenberg and Görlitz, *Elements of Rational Thought* (1735) and *Ontological Primer* (1738), gained wide circulation. Andreas Böhm (1720–1790) at Gießen contributed to the debate with *Metaphysics* (1753). Johann Nikolaus Frobesius (1701–1756) at Helmstedt (whose poet laureate was the female Wolffian Unzer) supplied with *Outline of Wolff's Metaphysics-System* (1730) yet another perspective. Israel Gottlieb Canz (1690–1753) at Tübingen (after Bilfinger was fired on behest of the theologians) contributed to the Jewish reception that influenced Moses Mendelssohn (1729–1786) with *The Use of Leibnizian-Wolffian Thought in Theology* (1728), *All Moral Disciplines* (1739), *Basics of Human Cognition* (1741), and *Elementary Philosophy* (1744). The Pietist Martin Knutzen (1713–1751) at Königsberg contributed *Elements of Rational and Logical Thought* (1744) before parting ways with Wolff over the theological ramifications of causal patterns.

Johann (Jean) Henri Samuel Formey (1711–1797), secretary of the Berlin Academy, thought that Enlightenment should not be a male affair and trained female intellectuals with the six-volume *La Belle Wolffienne* (1741–1753). One result of Formey's work was to create a social space for Unzer, the female thinker of the age. Unzer learnt from the Wolffian Georg Friedrich Meier (1718–1777) and from the psychologists in her family at Halle. She wrote a phenomenology of embodiment based on Wolff and Spinoza (*Outline of Philosophy for Females* (1751; 2nd ed. 1767).

Wolff's influence culminated in Kant. Kant arrived on the scene with a critique of Wolff's Newtonian departure from Leibniz (1749). Later, he integrated Wolff's and Euler's ideas into predictions of Earth's rotational and environmental fate, as well as into the discoveries of the daily rhythm of coastal winds, the coriolis turn of trade winds, and the seasonal cycle of the monsoon (1754–1757). In his critical phase, he denounced Wolff as a "dogmatic philosopher" and regarded him as the polar opposite to Hume; the *Critique* (1781) ends with a proposed middle way (a la Bilfinger) between the two heuristic extremes. Wolff's challenge is the natural law, the decorum, or rite of nature. The effect of Wolff's early *Aerometry* on Kant's rational apercus of climate patterns remains provocative to the twenty-first century, in light of current information on global warming. In modern times, Wolff's impact on the socioeconomic shape of the European Union (Maastricht treaties) is recognized, but his views on natural frames or "houses" (*oikos*), and on their internal dynamic interplays, are not topics of philosophical research.

See also Arnauld, Antoine; Baumgarten, Alexander Gottlieb; Bilfinger, Georg Bernhard; Calvin, John; Confucius; Copernicus, Nicolas; Cosmology; Crusius, Christian August; Descartes, René; Enlightenment; Galileo Galilei; Gottsched, Johann Christoph; Grotius, Hugo; Hegel, Georg Wilhelm Friedrich; Hume, David; Kant, Immanuel; Kepler, Johannes; Knutzen, Martin; Leibniz, Gottfried Wilhelm; Locke, John; Luther, Martin; Marx, Karl; Meier, Georg Friedrich; Mendelssohn, Moses; Monism and Pluralism; Newton, Isaac; Nietzsche, Friedrich; Ontology; Pietism; Schopenhauer, Arthur; Spinoza, Benedict (Baruch) de; Suárez, Francisco; Thomasius, Christian; Thümmig, Ludwig Philipp; Tschirnhaus, Ehrenfried Walter von; Voltaire, François-Marie Arouet de; Women in the History of Philosophy; Zhu Xi (Chu Hsi).

Bibliography

WORKS BY CHRISTIAN WOLFF

Gesammelte Werke [Collected works]. Edited by J. École, H. W. Arndt, C. A. Corr, J. E. Hofmann, and M. Thomann. Hildesheim: Georg Olms, 1965ff. [1965ff.]

All of the following works are contained in Wolff's Gesammelte Werke. *They were originally published at different times and in different cities throughout Germany, as reflected below, but have been gathered together in this modern reprint collection of Wolff's work.* (GW = *Gesammelte Werke*; Roman numerals refer to the German (I) or Latin (II) series in GW; Arabic numerals refer to the volumes in the individual series; a second Arabic numeral after volume number (e.g. 1.1.) refers to an individually released part of the volume.)

Aërometriae Elementa in quibus aliquot aëris vires ac proprietates iuxta methodum Geometrarum demonstrantur [*Aerometry*]. Leipzig, 1709.

Anfangsgründe aller mathematischen Wissenschaften [*Foundations of All Exact Sciences*]. 4 vols. Frankfurt/Leipzig, 1710. In GW I: 12–15.

Elementa matheseos Universae [*Universal Mathematics*]. 1st ed. 2 vol.; 2nd ed. 4 vol. Halle, 1713–1715. In GW II: 29–33.

Vernünftige Gedanken von den Kräften des menschlichen Verstandes und ihrem richtigen Gebrauche in der Erkenntnis der Wahrheit [*German Logic*]. Frankfurt/Leipzig, 1713. In GW I: 1.

Mathematisches Lexicon [*Mathematical Dictionary*]. Leipzig, 1716. In GW I: 11.

Ratio praelectionum Wolfianarum [in] mathesin et philosophiam universam et opus Hugonis Grotii de jure belle et pacis [*Reason of Wolff's Classes*]. Halle, 1718. In GW II: 36.

Vernünftige Gedanken von Gott, der Welt und der Seele des Menschen auch allen Dingen überhaupt [*German Metaphysics*]. Halle, 1719. In GW I: 2.

Vernünftige Gedanken von der Menschen Tun und Lassen zur Beförderung ihrer Glückseligkeit [*German Ethics*]. Frankfurt/Leipzig, 1720. In GW I: 4.

Vernünftige Gedanken von dem gesellschaftlichen Leben der Menschen und insbesondere dem Gemeinwesen zur Beförderung der Glückseligkeit des menschlichen Geschlechts [*German Politics*]. Frankfurt/Leipzig, 1721. In GW I: 5.

Vernünftige Gedanken von den Wirkungen der Natur [*German Physics*]. Halle, 1723. In GW I: 6.

Vernünftige Gedanken von den Absichten der natürlichen Dinge [*German Teleology*]. Frankfurt/Leipzig, 1724. In GW I: 7.

Anmerkungen über die vernünftigen Gedanken von Gott, der Welt und der Seele des Menschen auch allen Dingen überhaupt, zu besserem Verstande und bequemeren Gebrauche derselben [*Comments to German Metaphysics*]. Frankfurt, 1724. In GW I: 3.

Vernünftige Gedanken von dem Gebrauch der Teile in Menschen, Tieren und Pflanzen [*German Physiology*]. Leipzig, 1725. In GW I: 8.

Oratio de Sinarum philosophia practica / Rede über die praktische Philosophie der Chinesen. Frankfurt, 1726. Edited and translated by Michael Albrecht as *Speech on the Morals of the Chinese* (Hamburg: Felix Meiner, 1985).

Ausführliche Nachricht von seinen eigenen Schriften, die er in deutscher Sprache von den verschiedenen Teilen der Weltweisheit herausgegeben, auf Verlangen ans Licht gestellt [*Detailed News*]. Frankfurt, 1726. In GW I: 9.

Discursus praeliminaris de philosophia in genere [*Preliminary Discourse on Philosophy*]. Frankfurt, 1728. In GW II: 1.1.

Philosophia rationalis sive Logica, methodo scientifica pertractata et ad usum scientiarum atque vitae aptata [*Rational Philosophy or Logic*]. Frankfurt, 1728. In GW II: 1.2–3.

Philosophia prima, sive Ontologia, methodo scientifica pertractata qua omnis cognitiones humanae principia continentur [*First Philosophy or Ontology*]. Frankfurt/Leipzig, 1731. In GW II: 3.

Cosmologia generalis, methodo scientifica pertractata qua ad solidam, inprimis Dei atque naturae, cognitionem via sternitur [*General Cosmology*]. Frankfurt/Leipzig, 1731. In GW II: 4.

Psychologia empirica, methodo scientifica pertractata qua ea, quae de anima humana indubia experientiae fide constant, continentur et ad solidam universae philosophiae practicae ac theologiae naturalis tractationem via sternitur [*Empirical Psychology*]. Frankfurt/Leipzig, 1732. In GW II: 5.

Psychologia rationalis, methodo scientifica pertractata qua ea, quae de anima humana indubia experientiae fide innotescunt, per essentiam et naturam animae explicantur, et ad intimiorem naturae ejusque autoris cognitionem profutura proponuntur [*Rational Psychology*]. Frankfurt/Leipzig, 1734. In GW II: 6.

Theologia naturalis, methodo scientifica pertractata [*Natural Theology*]. 2 vols. Frankfurt/Leipzig, 1736–1737. In GW II: 7.1–7.2 and 8.

Philosophia practica universalis, methodo scientifica pertractata [*Global Practical Philosophy*]. 2 vols. Frankfurt/Leipzig, 1738–1739. In GW II: 10–11.

Jus Naturae, methodo scientifica pertractatum [*Natural Law*]. 8 vols. Frankfurt/Leipzig/Halle, 1740–1748. In GW II: 17–24.

Jus Gentium, methodo scientifica pertractatum in quo jus gentium naturale ab eo, quod voluntarii, pactitii et consuetudinarii est, accurate distinguitur [*International Law*]. Halle, 1749. In GW II: 25.

Institutiones juris naturae et gentium in quibus ex ipsa hominis natura continuo nexu omnes obligationes et jura omnia de deducuntur [*Principles of Natural and International Law*]. Halle, 1750. In GW II: 26.

Philosophia moralis sive Ethica, methodo scientifica pertractata [*Moral Philosophy*]. 4 vols. Halle, 1750–1753. In GW II: 12–16.

Oeconomica, methodo scientifica pertractata [*Economics*] 2 vols. Halle, 1754–1755. In GW II: 27–28.

WORKS ABOUT CHRISTIAN WOLFF

Albrecht, Michael. "Einleitung." In *Oratio de Sinarum philosophia practica*, edited by Michael Albrecht. Hamburg: Meiner, 1985.

Arndt, Hans Werner. "Rationalismus und Empirismus in der Erkenntnislehre Christian Wolffs." In *Christian Wolff 1679–1754: Interpretationen zu seiner Philosophie und deren Wirkung*, edited by Werner Schneiders, 31–47. Hamburg: Meiner, 1983.

Backhaus, Jürgen, ed. *Christian Wolff and Law & Economics: The Heilbronn Symposium*. Hildesheim: Olms, 1998.

Backhaus, Jürgen. "Christian Wolff on Subsidiarity, the Division of Labor, and Social Welfare." *European Journal of Law and Economics* 4 (1997): 129–146.

Beck, Lewis White. *Early German Philosophy: Kant and His Predecessors*. Cambridge, MA: Harvard University Press, 1969.

Biller, Gerhard. *Wolff nach Kant: eine Bibliographie*. Hildesheim: Olms, 2004.

Bissinger, Anton. "Zur metaphysischen Begründung der Wolffschen Ethik." In *Christian Wolff 1679–1754: Interpretationen zu seiner Philosophie und deren Wirkung*, edited by Werner Schneiders, 148–160. Hamburg: Meiner, 1983.

Campo, Mariano. *Cristiano Wolff e il razionalismo precritico* (1939). 2 vols. Hildesheim: Olms, 1980.

Corr, Charles A. "Christian Wolff and Leibniz." *Journal of the History of Ideas* 36 (1975): 241–262.

Corr, Charles A. "Introduction." In *Vernünftige Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt* [*German Metaphysics*]. Halle, 1719. In GW I: 2.1–47.

Drechsler, Wolfgang. "Christian Wolff (1679–1754): A Biographical Essay." *European Journal of Law and Economics* 4 (1997): 111–128.

École, Jean. "A propos du project de Wolff d'écrire une "Philosophie des Dames." *Studia Leibnitiana* 15 (1983): 46–57.

École, Jean. "La critique wolffienne du spinozisme." *Archives de Philosophie* 46 (1983): 553–567.

École, Jean. *La Métaphysique de Christian Wolff*. 2 vols. Hildesheim: Olms, 1990.

École, Jean. *Nouvelles etudes et nouveaux documents photographiques sur Wolff*. Hildesheim: Olms, 1997.

Heimsoeth, Heinz. "Christian Wolffs Ontologie und die Prinzipienforschung Immanuel Kants." In *Studien zur Philosophie Immanuel Kants: Metaphysische Ursprünge und ontologische Grundlagen*, edited by Heinz Heimsoeth. Kant-Studien Ergänzungsheft 71, Bonn: Bouvier, 1956.

Krause, Günter. "Christian Wolff and the Classics of Scientific Socialism." *European Journal of Law and Economics* 4 (1997): 285–297.

Ludovici, Carl Gustav. *Ausführlicher Entwurf einer vollständigen Historie der Wolffischen Philosophie*. 3 vols. Leipzig, 1736–1737. In *Werke* III: 1. Hildesheim: Olms, 2003.

Philipp, Wolfgang. *Das Werden der Aufklärung in theologiegeschichtlicher Sicht*. Göttingen: Vandenhoeck & Ruprecht, 1957.

Reinert, Erik S., and Arno M. Daastøl. "Exploring the Genesis of Economic Innovations." *European Journal of Law and Economics* 4 (1997): 233–283.

Rutherford, Donald. "Idealism Declined: Leibniz and Christian Wolff." In *Leibniz and His Correspondents*, edited by Paul Lodge, 214–237. Cambridge, U.K.: Cambridge University Press, 2004.

Rudolph, Oliver-Pierre. "Mémoire, réflexion et conscience chez Christian Wolff." *Revue philosophique* 3 (2003): 351–360.

Schneiders, Werner, ed. *Christian Wolff 1679–1754: Interpretationen zu seiner Philosophie und deren Wirkung*. Hamburg: Meiner, 1983.

Schönfeld, Martin. "Christian Wolff and Leibnizian Monads." *Leibniz Society Review* 11 (2002): 81–90.

Schönfeld, Martin. "German Philosophy after Leibniz." In *Companion to Early Modern Philosophy*, edited by Steven Nadler, 545–561 Oxford, U.K.: Blackwell, 2002.

Schönfeld, Martin. *The Philosophy of the Young Kant: The Precritical Project*. New York: Oxford University Press, 2000.

Schönfeld, Martin. "The Tao of Königsberg." *Florida Philosophical Review* 3 (2003): 5–32.

Stolzenberg, Jürgen, and Oliver-Pierre Rudolph, eds. *Christian Wolff und die Europäische Aufklärung. Akten des 1. Internationalen Christian-Wolff-Kongresses, Halle (Saale), 4.–8. April 2004*. Hildesheim: Olms, 2005.

Tonelli, Giorgio. "Der Streit über die Mathematische Methode in der Philosophie in der ersten Hälfte des 18. Jahrhunderts und die Entstehung von Kants Schrift über die 'Deutlichkeit.'" *Archiv für Philosophie* 9 (1959): 37–66.

Tonelli, Giorgio. *Elementi methodologici e metafisici in Kant dal 1747 al 1768*. Turin: Edizione di Filosofia, 1959.

Tonelli, Giorgio. "Wolff, Christian." In *Encyclopedia of Philosophy*, edited by Paul Edwards. 8 vols. New York: Macmillan, 1967.

Vleeschauwer, Herman Jean de. "La genèse de la méthode mathématique de Wolff." *Revue belge de philologie et d'histoire* 11 (1931): 651–677.

Wundt, Wilhelm. *Die deutsche Schulphilosophie im Zeitalter der Aufklärung*. Tübingen: Mohr, 1924.

Zeller, Eduard, "Wolffs Vertreibung aus Halle." In *Vorträge und Abhandlungen*, edited by Eduard Zeller. 2 vols. Leipzig: Fues, 1865.

Martin Schönfeld (2005)

WOLLASTON, WILLIAM

(1659–1724)

Born in 1660 at Coton-Clanford in Staffordshire, England, William Wollaston entered Sidney Sussex College, Cambridge, in 1674 as a pensioner. After receiving his MA in 1681, he took up the post of assistant master of Birmingham Grammar School. In his late twenties he unexpectedly came into a large inheritance and subsequently married a wealthy heiress with whom he had eleven children. Retiring to a life devoted to domestic matters, he began writing treatises on philosophical and ecclesiastical questions. In 1691 his *The Design of Part of the Book of Ecclesiastes* was published. His one important philosophical work, *The Religion of Nature Delineated*, was first published in 1724, with eight more editions following by 1759. Although he wrote many other treatises, he burned most of them toward the end of his life. He died in 1724, wealthy and esteemed. Queen Caroline had a bust of him placed along with those of Isaac Newton, John Locke, and Samuel Clarke in the royal garden at Richmond, England.

Wollaston is often grouped with Clarke as an unflinching defender of the kind of moral rationalism

that David Hume, among others, opposed. Clarke, along with many other philosophers of the period, was motivated to write on moral philosophy in reaction to Thomas Hobbes's work, which he regarded as both wrong and dangerous. Wollaston was one of the few who did not join the debate with Hobbes; as a result, his work is, for the period, unusually free of polemics.

WOLLASTON'S CRITERION OF IMMORALITY

Clarke argued that wrong actions are unfit or inappropriate to the real nature and relations of things. At one point he characterizes evildoers as attempting "to make things be what they are not, and cannot be," which he thought was as absurd as trying to change a mathematical truth. Wollaston constructs his entire moral theory around this idea. But unlike Clarke, for whom the basic moral notions are fitness and unfitness, Wollaston argues that moral goodness and evil can be reduced to truth and falsehood.

His argument has two stages. In the first, he argues that we are able to say things not only with words but also with actions. Beginning by defining true propositions as those that "express things as they are," he argues that actions may express, declare, or assert propositions, by which he means something more than that we understand gestures such as laughing, weeping, or shrugging. To use his example, if one group of soldiers fires on another, the first group's actions declare that the second is its enemy. If it turns out that the second group is not the first group's enemy, its declaration is false. Since we can understand actions, they—like sentences—have meaning, and whatever has a meaning is capable of truth and falsity.

Wollaston acknowledges that some actions have only conventional meaning—taking one's hat off when praying is a sign of reverence for Christian men but not for Jewish men. According to him, words always have a conventional meaning. He thinks, however, that many actions have a natural meaning that cannot be changed by agreement or force. For example, by using and disposing of something, I signify that it is mine. If it is not mine, my actions declare something false. When actions have natural meaning, Wollaston maintains that they express propositions more strongly than do mere words.

In the second stage of his argument, Wollaston proposes what he thinks is the basic criterion of immoral actions, "No act of any being, to whom moral good and evil are imputable, that interferes with any true proposition, or denies anything to be as it is, can be right" (1724, p. 13). Since immoral actions deny things to be what they are, they express false propositions. If I break a promise, I falsely declare that I never made one. If I am ungrateful, I falsely assert that I never received favors. To treat things as being what they are not is, for Wollaston, irrational in the sense that it is one of the greatest absurdities, "It is to put bitter for sweet, darkness for light, crooked for straight, etc." (p. 15).

TRUTH, HAPPINESS AND REASON

Wollaston goes on to try to show that "the way to happiness and the practice of reason" come to the same thing: they are both acting in conformity to truth (1724, p. 52). He thinks the nature of human beings is such that aim at their own happiness. Not only is happiness our natural good but we also have a duty to strive for our own happiness as well as the happiness of others. Anticipating Jeremy Bentham, Wollaston defines happiness as the "true quantity of pleasure": pleasures and pains may be measured in terms of their intensity and duration. We are happy when the sum total of pleasures exceeds the sum total of pains. Just as happiness cannot be achieved by anything that interferes with morality (truth), so the practice of truth (acting morally), Wollaston argues, cannot make a person unhappy. Morality and happiness are congruent, if not in this world, then in the afterlife.

Wollaston thinks that we are first and foremost rational creature. On his view, reason—or, more precisely, right reason—enables us to discover truth. When our actions are in accord with right reason, they express truths. To act according to right reason is thus the same as acting according to truth. It is reason's nature to command, he maintains, and as rational creatures, reason ought to govern us. Not only does reason enable us to discover which actions are morally good, but Wollaston also assumes that our motivation to act morally comes from reason. He argues that true happiness can be achieved only by pursuing means that are consistent with our rational nature, concluding that the "truest" definition of morality is "the pursuit of happiness by the practice of truth and reason" (1724, p. 52).

Belief in God underpins Wollaston's moral theory. God is the author of nature, including our nature as rational beings. The truths we should aim to mirror in our actions are God's truths. They are natural, however, because we are able to grasp them by reason unaided by divine revelation. Thus, there is, he claims, such a thing as natural religion.

CRITICISMS OF WOLLASTON

Wollaston is perhaps best known today not because of what he wrote, but because of the criticisms Hume and others brought against his theory. While his theory was popular during his lifetime, it was, and continues to be, subject to misinterpretations and parodies. Some of this was fostered by Wollaston's tendency to state his views in rhetorical or even paradoxical terms, for example, saying that an evildoer "lives a lie" or that "the *true quantity of pleasure* differs not from that *quantity of true pleasure*" (1724, pp. 11, 36). To the annoyance of some commentators, he included many footnotes in which he quotes in the original from Greek, Roman, Hebrew, and Arabic sources.

While many objections to Wollaston are based on misinterpretations of his view, some are so hilarious that they should be taken as parodies rather than as serious criticism. John Clarke (1725), offers the following quip. If expressing truth is our aim, a person should "spend his time in thrumming over such worthy and weighty propositions as these, 'a man's no horse, a horse, no cow, a cow no bull, nor a bull an ass" (p. 19). Hume (1978), following the eighteenth-century sentimentalist Francis Hutcheson (2002), often takes Wollaston's criterion of wrong actions to be the intention to cause false beliefs in others. He illustrates this reading with the absurd example of someone walking by an open window and seeing Hume cavorting with his neighbor's wife and being caused to falsely believe she is his wife. Hume responds that if that is the case, then the wrongdoing is unintentional since the adulterer's intention is to satisfy his lust and passion, not to cause false beliefs in others. Furthermore, if he had taken the precaution of shutting the window, his actions would not have been immoral, since they would not have caused false beliefs in others.

Some criticisms of Wollaston are directed to his view that wrong actions express falsehoods. The most telling is that his criterion is circular. It is wrong for me to take off with your property, Wollaston says, because I falsely declare it to be mine, not yours. But if we ask why this is what my action means, the answer is that the fact that it is yours means that I should not steal it. In every case the truth that is supposedly denied by a wrong action already has moral content. Clarke (1725) was the first to raise the problem of circularity, but the best-known formulation is Hume's (1978). Richard Price, the eighteenth-century rationalist, and J. L. Mackie (1980), the late twentieth-century sentimentalist, offer similar versions.

Both Hume (1978) and Mackie (1980) object to Wollaston's theory on motivational grounds, arguing that reason alone cannot move us. Both also argue that while people often refrain from performing an action because they see that it is unjust or immoral, no one refrains because he or she thinks it expresses a falsehood. Hutcheson (2002) and the twentieth-century philosopher Joel Feinberg (1977) worry that the fact that truth and falsehood do not come in degrees implies that on Wollaston's view "all crimes must be equal." Wollaston foresaw this criticism and argued that an offense increases with the importance of the truth denied. By introducing the idea of the importance of truth, however, Wollaston abandons his claim that conformity to truth is the only criterion of wrongness. Despite these criticisms, however, philosophers such as Feinberg (1977) and Mackie (1980) find Wollaston's idea that actions have meaning to be philosophically interesting.

See also Action; Bentham, Jeremy; Clarke, Samuel; Ethics, History of; Evil; Feinberg, Joel; Hobbes, Thomas; Hume, David; Hutcheson, Francis; Locke, John; Mackie, John Leslie; Newton, Isaac; Price, Richard; Rationalism.

Bibliography

WORKS BY WOLLASTON

The Design of Part of the Book of Ecclesiastes: Or, The Unreasonableness of Men's Restless Contention for the Present Enjoyments, Represented in an English Poem. London: James Knapton, 1691.

The Religion of Nature Delineated. London: n.p., 1724.

WORKS ABOUT WOLLASTON

Clarke, John. *Examination of the Notion of Moral Good and Evil Advanced in the Religion of Nature Delineated*. London: A. Bettesworth, 1725.

Feinberg, Joel. "Wollaston and His Critics." *Journal of the History of Ideas* 38 (1977): 345–352.

Hume, David. *A Treatise of Human Nature*. 2nd ed., edited by L. A. Selby-Bigge. Oxford, U.K.: Clarendon Press, 1978.

Hutcheson, Francis. *An Essay on the Nature and Conduct of the Passions and Affections: With Illustrations on the Moral Sense*, edited by Aaron Garrett. Indianapolis, IN: Liberty Fund, 2002.

Joynton, Olin. "The Problem of Circularity in Wollaston's Moral Philosophy." *Journal of the History of Philosophy* 22 (1984): 435–443.

Mackie, J. L. *Hume's Moral Theory*. London: Routledge and Kegan Paul, 1980.

Tweyman, Stanley. "Truth, Happiness and Obligation: The Moral Philosophy of William Wollaston." *Philosophy* 51 (1976): 35–46.

Charlotte R. Brown (2005)

WOLLHEIM, RICHARD
(1923–2003)

Richard Arthur Wollheim, an English philosopher, was born in London. After service in World War II, where he rose to captain, he returned to Balliol College, Oxford, first to continue the study of history (in which he received a bachelor of arts degree in 1946), then philosophy, politics, and economics (in which he received a bachelor of arts degree in 1948). He was Grote Professor of Philosophy of Mind and Logic at University College London, 1963–1982; professor of philosophy at Columbia University, 1982–1985; Mills Professor of Intellectual and Moral Philosophy at the University of California at Berkeley, 1985–2002; and professor of philosophy and the humanities at the University of California at Davis, 1989–1996. He was elected a fellow of the British Academy in 1972 and of the American Academy of Arts and Sciences in 1986; and was vice-president of the British Society of Aesthetics, 1968–1993, and president, 1993–2003. His writings focused principally on two subjects: art and human psychology. He made outstanding contributions not just to general but also to substantive aesthetics, above all the philosophy of painting. His unrivalled knowledge of psychoanalytic theory enabled him to write a masterly account of Sigmund Freud's thought and endowed his work in the philosophy of mind with its distinctive character. The strength of his contributions to the advancement of psychoanalytic theory were recognized in the profession by the honors accorded him by the British Psychoanalytical Society and the International Psychoanalytical Association, among others. He died in London.

AESTHETICS

Wollheim's aesthetics is marked by its psychological orientation, manifest in his account of the nature of art, artistic meaning, pictorial representation and artistic expression. In his works, Wollheim argued that art is a form of life (in Ludwig Wittgenstein's sense), artistic activity and appreciation requiring the existence of practices and institutions, art being an essentially historical phenomenon, the changes to which it is inevitably subject affecting the conceptual structure that surrounds it. The aim of artists is, he maintained, to endow their work with a meaning determined by the intentions that guide their activity; the distinctive function of the spectator is to grasp that meaning, to retrieve those intentions, which is achieved, if the artist fulfilled them, by engaging with the work and undergoing the experience the artist intended it to provide.

This psychological account of artistic meaning and understanding is applied to the art of painting in what is perhaps Wollheim's masterpiece, *Painting as an Art,* which maintains that great art is, as is the socialism he embraced throughout his life, rooted in the assumption of a common human nature. A painting's meaning (each painting having one and only one meaning), which is visual, is revealed in the experience induced in an adequately sensitive and informed spectator who looks at the surface of the painting as the fulfilled intentions of the artist led him or her to mark it. He distinguished five principal kinds of primary pictorial meaning achievable by a work: representational, expressive, textual, historical, and metaphorical; he identified what he characterized as secondary meaning, which is what the act of giving a picture its primary meaning meant to the artist; and he illustrated these categories with a remarkable series of challenging interpretations of works by some of the painters he most admired.

He elucidated two other central issues, the nature of pictorial representation and of artistic expression, in psychological terms, each exploiting a species of perception. Pictorial representation is a function of "seeing-in," a perceptual experience which consists of two aspects, the configurational being the seeing of a marked surface, the recognitional being the seeing in this surface of something—a plane of color, perhaps—in front of or behind something else. Artistic expression, at least that involved in the art of painting, is a function of "expressive perception," a perceptual experience with three aspects, the first representing the world as "corresponding" to an affective condition, the second being an affect in the viewer that is "of a piece" with the corresponding condition, and the third being a revelation or intimation of the origin, either of the experience itself or of the kind to which it belongs, in so-called "complex" projection.

Wollheim also advanced an account of the ontology of art. He argued that the fundamental distinction within works of art is between individuals and types, some works of art being individuals, the rest types. Furthermore, every work of art belonging to the same art belongs to the same category, type or individual as the case may be, and, for all works of art, the identity of a work of art is determined by the history of its production.

PSYCHOLOGY

His investigation of the question, What is it to lead the life of a person?, claims a fundamental status for the nature of the process that mediates between a person and the life he or she leads—the leading of a life. This process is consti-

tuted by interactions between a person's past, present, and future, and to elucidate this Wollheim presented a typology of the mind, distinguishing mental dispositions from mental states, and proceeds to examine their interactions as well as those among the various systems of the mind, the conscious, the preconscious, and the unconscious. The aim is to outline a philosophy of mind of a kind that psychoanalytic theory requires and it is studded with profound observations of human life that even those sceptical of psychoanalysis stand to benefit from. His study of the emotions, which he "repsychologized," attributing to them psychological reality, represents them as mental dispositions that cause their manifestations, assigning them a particular role within the psychology of the person— that of providing the person with an attitude to the world. He sketched and then developed in great detail a characteristic history, one the recognition of which is essential to understanding what an emotion is.

This proceeds from the "originating condition" of emotion—the satisfaction or frustration of a desire, actually or merely believed in or prospective— through the "precipitating factor," to the transformation of the "originating condition," the experience of satisfaction or frustration being "extroverted," the "precipitating factor" being perceived to correspond to the experience and becoming the object of an emotion, and then, finally, to internal and external manifestations of the emotion and other outcomes. Two of the so-called moral emotions, shame and guilt, which are given extended treatment, are represented as deviating from this characteristic history, incorporating the psychoanalytic notion of fantasy as an essential ingredient of their nature.

See also Art, Expression in; Art, Representation in.

Bibliography

WORKS BY RICHARD WOLLHEIM

Sigmund Freud. New York: Viking, 1971.

On Art and the Mind. Cambridge, MA: Harvard University Press, 1974.

Art and Its Objects: With Six Supplementary Essays. 2nd ed. New York: Cambridge University Press, 1980.

The Thread of Life. New York: Cambridge University Press, 1984.

Painting as an Art. London: Thames and Hudson, 1987.

The Mind and Its Depths. Cambridge, MA: Harvard University Press, 1993.

On the Emotions. New Haven, CT: Yale University Press: 1999.

WORKS ABOUT RICHARD WOLLHEIM

Hopkins, Jim, and Anthony Savile, eds. *Psychoanalysis, Mind, and Art: Perspectives on Richard Wollheim.* Cambridge, MA: Blackwell, 1992.

van Gerwen, Rob, ed. *Richard Wollheim on the Art of Painting: Art as Representation and Expression.* New York: Cambridge University Press, 2001.

Malcolm Budd (2005)

WOLLSTONECRAFT, MARY
(1757–1797)

Mary Wollstonecraft has long been recognized as one of the most influential feminist theorists in history, largely through her *Vindication of the Rights of Woman* (1792). Late-twentieth-century scholarship also began to explore her other texts and their significance.

Wollstonecraft's work is a product of the late Enlightenment, emphasizing the need to achieve virtue and progress through development of reason and sensibility. It also reflects ideas of the Dissenters and political radicals who stood among the relatively few English supporters of the French Revolution. Wollstonecraft's early mentors were Richard Price and Joseph Priestley. The circle with whom she continued to associate included writers and artists such as William Blake, Thomas Paine, Henry Fuseli, and William Godwin. Like them, she opposed slavery, standing armies, and many elements of political patriarchy such as primogeniture, aristocracy, and probably monarchy. She shared their critique of the corrupting influence of political and social institutions structured around "unnatural distinctions" based on rank, property, religion, or profession.

Wollstonecraft's most distinctive and well-known contribution was to extend this analysis to demand an end to unnatural distinctions based on sex and family relations. As she wrote in the *Rights of Woman,* if observation could not prove that men had more natural capability for reason than women, they could claim no superiority over women and certainly no right to rule them. In analysis shaped by John Locke and Jean-Jacques Rousseau (but one that attacked Rousseau for his views on women), she concluded that education, experience, and the "present constitution of society," and not nature, created most observed character differences between men and women.

She argued that unnatural distinctions between women and men tended toward the same effects as other unjust power relations: They corrupt the character of all parties to the relationship, rendering the dominant party dependent on its power and making the subordinate party resort to cunning and unvirtuous strategies of self-preservation. In the case of women she pointed to the use

of beauty as what might now be called a "weapon of the weak." Unlike better-known democratic theorists of her era, she applied an antipatriarchal analysis commonly used on institutions such as government to the family itself.

She advocated altering the social practices such as dress, courtship, employment, and family relations that had given men power over women and kept both from virtue. She sought expanded work opportunities for women. She proposed development of a public school system educating girls and boys and children of different classes similarly and together, at least for the early years of their schooling, and wanted girls to study subjects that had been forbidden to them. Her final, unfinished novel, *Maria, or the Wrongs of Woman,* underscored the necessity of women's ability to support themselves, divorce, and have rights over their children.

Although she is most famous for her arguments on women's rights, other contributions are worth noting. Her *Vindication of the Rights of Men* (1790) was one of the first attacks on Edmund Burke's *Reflections on the Revolution in France,* and it engaged his work on the sublime and the beautiful, thus integrating aesthetics and politics in a critique of Burke's defense of monarchy, aristocracy, and pomp. Her further exploration of the French Revolution in the *Historical and Moral View of the Origin and Progress of the French Revolution* (1794) contains an underrated inquiry into the nature of political history and the relationship between ideals and human action. Wollstonecraft's *Letters Written during a Short Residence in Sweden, Norway, and Denmark* influenced the early generation of English Romantics, including Samuel Taylor Coleridge, Robert Southey, William Wordsworth, and Percy Bysshe Shelley and his wife, Wollstonecraft's daughter, Mary Shelley.

See also Analytical Feminism; Beauty; Blake, William; Burke, Edmund; Coleridge, Samuel Taylor; Enlightenment; Feminist Ethics; Feminist Philosophy; Godwin, William; Locke, John; Paine, Thomas; Price, Richard; Priestley, Joseph; Rousseau, Jean-Jacques; Shelley, Percy Bysshe; Ugliness; Women in the History of Philosophy.

Bibliography

WORKS BY WOLLSTONECRAFT

The Works of Mary Wollstonecraft. Edited by J. Todd and M. Butler. New York: New York University Press, 1989. Includes all of Wollstonecraft's works (other than letters). Among the most important are:

Mary: A Fiction (1788).

A Vindication of the Rights of Men, in a Letter to the Right Honorable Edmund Burke (1790).

A Vindication of the Rights of Woman with Strictures on Moral and Political Subjects (1792).

An Historical and Moral View of the Origin and Progress of the French Revolution; and the Effect It Has Produced in Europe (1794).

Letters Written during a Short Residence in Sweden, Norway, and Denmark (1796).

Maria, or the Wrongs of Woman (post.).

WORKS ON WOLLSTONECRAFT

Poovey, M. *The Proper Lady and the Woman Writer: Ideology as Style in the Works of Mary Wollstonecraft, Mary Shelley, and Jane Austen.* Chicago: University of Chicago Press, 1984.

Sapiro, V. *A Vindication of Political Virtue: The Political Theory of Mary Wollstonecraft.* Chicago: University of Chicago Press, 1992.

Tomalin, C. *The Life and Death of Mary Wollstonecraft.* New York: Harcourt Brace Jovanovich, 1974.

Virginia Sapiro (1996)

WOMEN IN THE HISTORY OF PHILOSOPHY

The standard twentieth-century histories of European philosophy do not include women as important, original contributors to the discipline's past. Some relegate a few to footnotes; most omit women entirely. Recent research, inspired by the influence of feminist theory, and by a renewed interest in the historiography of philosophy, has uncovered numerous women who contributed to philosophy over the centuries.

Women's representation in philosophy's history was not always as marginal as it came to be by the opening of the twentieth century. For example, in the seventeenth century, Thomas Stanley's history mentioned twenty-four women philosophers of the ancient world, while Gilles Ménage discussed some seventy, including women Platonists, Academicians, Dialecticians, Cyrenaics, Megarians, Cynics, Peripatetics, Epicureans, Stoics, and Pythagoreans. With respect to the moderns, the seventeenth-century treatises of Jean de La Forge and Marguerite Buffet provided doxographies of women philosophers. Even in the nineteenth century, when women were virtually being erased from the standard histories, Lescure, Joël, Foucher de Careil, and Cousin wrote special studies on female philosophers.

Published 1987–1991, *A History of Women Philosophers,* volume 1, *600 BC–500 AD,* edited by Mary Ellen Waithe, has provided a detailed discussion of the following figures: Themistoclea, Theano I and II, Arignote,

Myia, Damo, Aesara of Lucania, Phintys of Sparta, Perictione I and II, Aspasia of Miletus, Julia Domna, Makrina, Hypatia of Alexandria, Arete of Cyrene, Asclepigenia of Athens, Axiothea of Philesia, Cleobulina of Rhodes. Hipparchia the Cynic, and Lasthenia of Mantinea. In addition to the medieval and Renaissance philosophers discussed in the second volume of Waithe's *History* (Hildegard of Bingen, Heloise, Herrad of Hohenbourg, Beatrice of Nazareth, Mechtild of Magdeburg, Hadewych of Antwerp, Birgitta of Sweden, Julian of Norwich, Catherine of Siena, Oliva Sabuco de Nantes Barrera, Roswitha of Gandersheim, Christine de Pisan, Margaret More Roper, and Teresa of Avila), scholars have recently begun to focus attention on such humanist and Reformation figures as Isotta Nogarola, Laura Cereta, Cassandra Fidele, Olimpia Morata, and Caritas Pickheimer.

THE SEVENTEENTH CENTURY

In the early modern period women's initial published philosophical endeavors inserted argumentation into the largely literary genre of the *querelle des femmes,* or woman question. Thus, Marie de Gournay, adopted daughter of Michel Eyquem de Montaigne, in *The Equality of Men and Women* (1622) replaced persuasive force based on example with skeptical and fideistic arguments; Anna Maria van Schurman's *Whether a Maid May Be a Scholar?* (1659) and Sor Juana Inés de la Cruz's "Response to Sor Filotea" (1700) used scholastic models of argumentation to discuss woman's nature and her relation to learning. By 1673, when Bathsua Makin published *An Essay to Revive the Ancient Education of Gentlewomen,* an unbroken, explicitly acknowledged line of influence ran from Gournay through van Schurman to Makin. In the second half of the century, partly in response to the writings of Desiderius Erasmus, Juan Luis Vives, and François de Salignac de La Mothe Fénelon, a number of treatises on the education of girls appeared, stressing its importance for religion and society. Authors included the Port Royal educator Sister Jacqueline Pascal and Madame de Maintenon.

In the second half of the Age of Reason women also produced numerous works on morals and the passions, including the maxims of Marguerite de La Sablière, Marquise de Sablé, and Queen Christina of Sweden. Perhaps the most well-known seventeenth-century woman writer of moral psychology is Madeline de Scudéry, of whom Gottfried Wilhelm Leibniz said that she had "clarified so well the temperaments and the passions in her … conversations on morals."

Another type of philosophical writing by women, the treatment of natural philosophy, begins to appear after 1660. In Paris Jeanne Dumée and, in England, Aphra Behn argued in defense of Nicolas Copernicus. But by far the most prolific female philosopher then was Margaret Cavendish, who published over a half dozen books on natural philosophy in which she advanced a unique combination of hard-nosed materialism together with an organic model of natural change and a denial of mechanism.

Of Anne Conway Leibniz said, "My philosophical views approach somewhat closely those of the late Countess of Conway." Her metaphysical treatise argued against René Descartes, Benedict de Spinoza, and Thomas Hobbes in favor of a monistic vitalism. On the Continent Princess Elisabeth of Bohemia, whose letters to Descartes had exposed the weakness of the latter's published views on mind-body interaction and free will, discussed Conway's philosophy with a Quaker correspondent. Seventeenth-century England also produced Mary Astell, who in the appendix to the *Letters concerning the Love of God* (1695) argued against occasionalism. In *A Serious Proposal to the Ladies, Part II* (1697), Astell offered women a manual for improving their powers of reasoning, a work that was influenced by Descartes and the Port Royal logicians. Damaris Cudworth Masham also argued against occasionalism in *Discourse concerning the Love of God* (1696). In *Occasional Thoughts* (1705) she defended a number of Lockean views on knowledge, education, and the relative merits of reason and revelation. Masham also corresponded with Leibniz on metaphysical issues, especially his views on substance; yet despite this scholarly career, she stood in need of defense against the charge that the arguments addressed to Leibniz could not have been written by a woman. It was Catherine Trotter Cockburn who came to her defense. Cockburn wrote a number of philosophical works, including *A Defence of Mr. Locke's Essay of Human Understanding* (1702) and a vindication of the views of Samuel Clarke.

In France in the final years of the seventeenth century, Gabrielle Suchon published, arguably, the most ambitious philosophical text that had yet been written by a woman on the Continent: *Treatise of Morals and of Politics* (1693), which included book-length treatments of liberty, science, and authority. Excerpts of her work were published in the scholarly journals of the time, but since the *Treatise* was published under a pseudonym, Suchon fell into oblivion by the late eighteenth century. (Anonymous authorship similarly led to Conway's erasure.)

THE EIGHTEENTH CENTURY

In England Catherine Macaulay published a critical treatment of Hobbes's political philosophy and her magnum opus, *Letters on Education* (1790), to which Mary Wollstonecraft explicitly acknowledges her debt in her own *Vindication of the Rights of Woman* (1792). By the end of the century Mary Hays's *Female Biography* (1803) demonstrated that English women were beginning to trace a history of feminist social and political philosophy that reached back about 100 years to Astell. At the turn of the century, with the growing professionalization of philosophy and placement of it over against the belles lettres and religion, women were producing philosophy stripped of its moorings within discussions of the woman question and theology, and written in journalistic style, as evidenced in Mary Shepherd's book-length treatments of causation, skepticism, and knowledge of the external world, with their attendant criticisms of such figures as David Hume and George Berkeley.

In Enlightenment France Anne Dacier published a translation and commentary for the writings of Marcus Aurelius and entered the debate about the ancients versus the moderns in her *The Causes of the Corruption of Taste* (1714). Dacier's salonist friend, the marquise de Lambert, published a number of works on morals, the passions, education, and woman's status, which continued to be published a century later. Sophie de Grouchy, Marquise de Condorcet, added to her translation of Adam Smith's. *Theory of the Moral Sentiments* her own blend of rationalist ethics and moral sentiment theory in her eight letters on sympathy.

Prior to the French Revolution philosophy of education, in particular, critical responses to Jean-Jacques Rousseau's *Émile,* occupied a prominent place in women's philosophical writings, as exemplified in Louise d'Epinay's *The Conversations of Emilie* (1774) and the works of Mme. de Genlis. In addition to her work on education Louise-Marie Dupin also left an extensive manuscript, *Observations on the Equality of the Sexes and of Their Difference,* which she dictated to her secretary, Rousseau. The French Revolution moved the issue of woman's education into the arena of the rights of a woman as a citizen. Perhaps the most famous of these treatises is Olympe de Gouge's *Declaration of the Rights of Woman* (1791).

In the area of natural philosophy there is no question but that Émilie du Châtelet deserves recognition as an important figure of the eighteenth century. Her *Principles of Physics* (1740) and her letters on the "active force" controversy (1742) attempt to reconcile what she takes to be most useful in Newtonian mechanics and Leibnizian philosophy. Du Châtelet also published a *Discourse on Happiness* (1779) and essays on the existence of God, the formation of color, and grammatical structure.

By the end of the century French women were producing broad critiques of culture and the arts, as evidenced in the mathematician Sophie Germain's *General Considerations on the State of the Sciences and Letters* (1833) and Madame de Staël's *On the Influence of the Passions on the Happiness of Individuals and Nations* (1796).

Germany spawned two critical treatments of Immanuel Kant's views on women: the first by an unidentified "Henriette" and the second by Amalia Hoist. In Switzerland Marie Huber's publications included three Enlightenment texts on the principles of natural religion: *The World Unmask'd* (English translation, 1736), *The State of Souls Separated from their Bodies* (English translation, 1736), and *Letters on the Religion Essential to Man* (English translation, 1738).

In Russia Catherine the Great's correspondence with Voltaire was published posthumously. Finally, in Italy Laura Bassi publicly disputed philosophical theses and published five lectures on natural philosophy; Maria Agnesi discussed logic, metaphysics, and Cartesian physics in *Philosophical Propositions* (1738); and Giuseppa Barbapiccola translated and wrote a critical introduction for Descartes's *Principles of Philosophy* (1731).

The information now available about women philosophers and ongoing research on this topic will provide us with a richer picture of philosophy's significant figures, topics, and styles of argumentation. It is to be hoped that future histories of philosophy will reflect this richer panorama of the past.

See also Berkeley, George; Conway, Anne; Copernicus, Nicolas; Descartes, René; Erasmus, Desiderius; Feminist Philosophy; Fénelon, François de Salignac de la Mothe; Gournay, Marie le Jars de; Hildegard of Bingen; Hobbes, Thomas; Hume, David; Hypatia; Kant, Immanuel; Leibniz, Gottfried Wilhelm; Locke, John; Marcus Aurelius Antoninus; Montaigne, Michel Eyquem de; Rousseau, Jean-Jacques; Spinoza, Benedict (Baruch) de; Vives, Juan Luis; Voltaire, François-Marie Arouet de; Wollstonecraft, Mary.

Bibliography

THE SEVENTEENTH CENTURY

Astell, M. *Letters concerning the Love of God between the Author of the Proposal to the Ladies and Mr. John Norris* (London, 1695); *A Serious Proposal to the Ladies Part II: Wherein a*

Method is offer'd for the Improvement of their Minds (London, 1697); *Some Reflections Upon Marriage* (London, 1700); *The Christian Religion as Profess'd by a Daughter of the Church of England* (London, 1705).

Behn, A. "The Translator's Preface." In B. le Bovier de Fontenelle, *A Discovery of New Worlds,* translated by A. Behn (London, 1688).

Cavendish, M. L., Duchess of Newcastle. *Philosophical and Physical Opinions* (London, 1655); *Orations of Divers Sorts* (London, 1662); *Philosophical Letters: or, Modest Reflections upon some Opinions in Natural Philosophy Maintained By Several Famous and Learned Authors of this Age* (London, 1664); *Observations upon Experimental Philosophy* (London, 1666); *Grounds of Natural Philosophy* (London, 1668).

Christina, Queen of Sweden. *L'ouvrage de loisir* (c. 1670–1680) and *Les sentiments héroïques* (c. 1670–1680), with *Réflexions diverses sur la Vie et sur les Actions du Grand Alexandre, Les Vertues et vices de Caesar,* and correspondence, in J. Arckenholtz, *Mémoires concernant Christine, reine de Suède,* 4 vols. (Leipzig, 1751–1760).

Conway, A., Viscountess. *The Principles of the Most Ancient and Modern Philosophy* (Latin translation: Amsterdam, 1690; English retranslation: London, 1692; both reprinted: The Hague, 1982); *The Conway Letters,* edited by M. H. Nicholson and S. Hutton (Oxford: Clarendon Press, 1992).

Dumée, J. *Entretien sur l'opinion de Copernic touchant la mobilité de la terre* (Paris, n.d.); ms. c. 1680, Bibliothèque Nationale Fonds français 1941.

Elisabeth, Princess of Bohemia. Her letters in: *Oeuvres de Descartes,* edited by C. Adam and P. Tannery (Paris, 1897–1913; rev. ed. 1964–1974); N. Malebranche, *Correspondance, actes et documents 1638–1689,* edited by A. Robinet, Vol. 18 (Paris, 1961); *Papers of William Penn,* Vol. 1, edited by M. Dunn and R. Dunn (Philadelphia, 1981).

Gournay, M. le Jars de. *L'egalité des hommes et des femmes* (Paris, 1622); *L'ombre de la Damoiselle de Gournay* (Paris, 1626); *Les advis ou Les presens de la Demoiselle de Gournay* (Paris, 1634).

Juana Inés de la Cruz, Sor. *Carta athenagórica de la madre Juana Inés de la Cruz* (Puebla de los Angeles, 1690); *Fama, y obras póstumas del fenix de Mexico, Decima Musa, Poetisa Americana* (Madrid, 1700); *Obras completas,* edited by A. Méndez Plancarte (A. Salceda), 4 vols. (Mexico City, 1951–1957).

Lettres, Opuscules et Mémoires de Mme. Périer et de Jacqueline, Soeurs de Pascal. Edited by P. Faugère. Paris, 1845.

Maintenon, F. d'Aubigné. *Lettres sur l'éducation des filles,* edited by Th. Lavallée (Paris, 1854); *Entretiens sur l'éducation des filles,* edited by Th. Lavallée (Paris, 1854); *Lettres historiques et édifiantes adressées aux dames de St.-Louis,* edited by Th. Lavallée, 2 vols. (Paris, 1856); *Conseils et instructions aux demoiselles pour leur conduite dans le monde,* 2 vols. (Paris, 1857).

Makin, B. *An Essay to Revive the Antient Education of Gentlewomen.* London, 1673.

Masham, D. C. A *Discourse concerning the Love of God* (London, 1696; French translation, 1705); *Occasional Thoughts in Reference to a Vertuous or Christian Life* (London, 1705); letters to Locke in *The Correspondence of John Locke,* edited by E. S. de Beer, 8 vols. (Oxford, 1976–85); letters to Leibniz in *Die Philosophischen Schriften von Leibniz,* edited by C. I. Gerhardt, 7 vols. (Berlin, 1875–1890).

*Réflexions ou Sentences et Maximes morales de Monsieur de la Rochefoucauld, Maximes de Madame la marquise de Sablé. Pensées diverses de M. L. D. et les Maximes chrétiennes de M***** (Mme. de La Sablière) (Amsterdam, 1705).

Schurman, A. M. van. *Amica dissertatio inter Annam Mariam Schurmanniam et Andr. Rivetum de capacitate ingenii muliebris ad scientias* (Paris, 1638); *Opuscula, hebraea, graeca, latina, gallica, prosaica et metrica* (Leiden, 1648).

Scudéry, M. de. *Discours sur la gloire* (Paris, 1671); *Conversations sur divers sujets,* 2 vols. (Paris, 1680); *Conversations nouvelles sur divers sujets,* 2 vols. (Paris, 1684); *Conversations morales,* 2 vols. (Paris, 1686); *Nouvelles Conversations de morale,* 2 vols. (Paris, 1688); *Entretiens de morale,* 2 vols. (Paris, 1692).

Suchon, G. *Traité de la morale et de la Politique* (Lyon, 1693); *[Traité] Du célibat Volontaire, ou la Vie sans engagement, par Demoiselle Suchon* (Paris, 1700).

THE EIGHTEENTH CENTURY

Agnesi, M. G. *Propositiones Philosophicae.* Milan, 1738.

Barbapiccola, G. E. *I Principii della Filosofia.* Turin, 1722.

Bassi, L. M. C. *Philosophica Studia* (49 theses disputed for the doctorate; Bologna, 1732); *De acqua corpore naturali elemento aliorum corporum parte universi* (theses for a disputation; Bologna, 1732); the following appear in *De Bononiensi Scientiarum et Artium Instituto atque Academia Commentarii: De aeris compressione* (1745); *De problemate quodam hydrometrico* (1757); *De problemate quodam mechanico* (1757); *De immixto fluidis aere* (1792).

Cockburn, C. Trotter. *The Works of Mrs. Catherine Cockburn, Theological, Moral, Dramatic, and Poetical,* edited by T. Birch. London, 1751.

Dacier, A. L. *Réflexions morales de l'empereur Marc Antonin* (Paris, 1690–1691; English translation: London, 1692); *Des causes de la corruption du goût* (Paris, 1714).

D'Épinay, L. *Les conversations d'Émilie.* Leipzig, 1774; Paris, 1781.

Documents of Catherine the Great: The Correspondence with Voltaire and the Instruction of 1767 in the English text of 1768, edited by W. F. Reddaway. Cambridge, U.K., 1931.

du Châtelet, G. É. Le Tonnelier De Breteuil, Marquise. *Institutions de Physique* (Paris, 1740); *Réponse de Madame**** [du Châtelet] *à la lettre que M. de Mairan … lui a écrite le 18 février sur la question des forces vives* (Brussels, 1741); *Dissertation sur la nature et la propagation du feu* (Paris, 1744); *Principes mathématiques de la philosophie naturelle,* 2 vols. (Paris, 1756); *Réflexions sur le bonheur* in *Opuscules philosophiques et littéraires, la plupart posthumes ou inédits* (Paris, 1796); essays in Ira O. Wade, *Studies on Voltaire with Some Unpublished Papers of Mme du Châtelet* (Princeton, NJ: Princeton University Press, 1947).

Dupin, L. *Portefeuille de Mme Dupin.* Paris, 1884.

Genlis, S. F. du Crest de Saint-Aubin, Comtesse de. *Adèle et Théodore ou lettres sur l'éducation* (Paris, 1782; English translation: London, 1783); *Discours sur la suppression des couvents de religieuses et l'éducation publique des femmes* (Paris, 1790).

Germain, S. *Oeuvres philosophiques de Sophie Germain, suivies de pensées et de lettres inédites.* Paris, 1879; 1896.

Gouges, O. de. *Les droits de la femme* (Paris, [1791]); *Oeuvres,* edited by Groult (Paris, 1986).

Grouchy, S. de, Marquise de Condorcet. *Lettres sur la Sympathie.* In *Théorie des Sentimens Moraux.* Vol. 2. Paris, 1798.

Hoist, A. *Über die Bestimmung des Weibes zur öhern Geistesbildung.* Berlin, 1802.

Huber, M. *Le monde fou préféré au monde sage* (Amsterdam, 1731); *Le systême des anciens et des modernes, … sur l'état des âmes séparées des corps* (London, 1731), both in English translation as *The World Unmask'd, or the Philosopher the greatest Cheat in Twenty Four Dialogues … To which is added, The State of Souls Separated from their Bodies … In Answer to a Treatise entitled, An Enquiry into Origenism* (London, 1736); *Lettres sur la religion essentielle à l'homme, distinguée de ce qui n'en est que l'accessoire* (Amsterdam, 1738; English translation, 1738).

Lambert, A. de. *Réflexions Nouvelles Sur Les Femmes par une Dame de la Cour de France* (Paris, 1727); *Lettres sur la véritable éducation* (Paris/Amsterdam, 1729); *Traité de l'Amitié, Traité de la Vieillesse, Réflexions sur les Femmes, sur le Goût, sur les Richesses* (Amsterdam, 1732); *Oeuvres complètes …* (Paris, 1808).

Macaulay, C. S. *Loose Remarks on Certain Positions to be found in Mr. Hobbes's Philosophical Rudiments of Government and Society* (London, 1767); *Letters on Education* (London, 1790).

Shepherd, M. *An Essay Upon the Relation of Cause and Effect, controverting the Doctrine of Mr. Hume …* (London, 1824); *Essays on the Perception of an External Universe* (London, 1827); "Lady Mary Shepherd's Metaphysics," *Fraser's Magazine for Town and Country* 5 (30) (July 1832).

Staël, G. de. *De l'influence des passions sur le bonheur des individus et de nations.* Paris, 1796.

Wollstonecraft, M. *Thoughts on the Education of Daughters* (London, 1787); *A Vindication of the Rights of Men* (London, 1790); *A Vindication of the Rights of Woman* (London, 1792).

WORKS ON WOMEN PHILOSOPHERS

Albistur, M., and D. Armogathe. *Histoire du féminisme français.* Vol. 1. Paris: Femmes, 1977.

Buffet, M. *Nouvelles observations sur la langue françoise … Avec les éloges d'illustres sçavantes tant anciennes que modernes.* Paris, 1668.

Cousin, V. *Jacqueline Pascal: Premières études sur les femmes illustres et la société du XVIIe siècle.* Paris, 1844. *Madame de Sablé: Nouvelles Etudes sur la société et les femmes illustres du dix-septième siècle.* Paris, 1854. *La société Française au XVIIe Siècle d'après Le Grand Cyrus de Mlle de Scudéry.* 2 vols. Paris, 1858.

Dronke, P. *Women Writers of the Middle Ages.* Cambridge, U.K.: Cambridge University Press, 1984.

Foucher de Careil, L. *Descartes et la Princesse Palatine, ou de l'influence du cartésianisme sur les femmes au XVIIe siècle.* Paris, 1862.

Foucher de Careil, L. *Descartes, la princesse Elisabeth et la reine Christine.* Paris, 1909.

Harth, E. *Cartesian Women.* Ithaca, NY: Cornell University Press, 1992.

Joël, K. *Die Frauen in der Philosophie.* Hamburg, 1896.

King, M. L. *Women of the Renaissance.* Chicago: University of Chicago Press, 1991.

Kristeller, P. O. "Learned Women of Early Modern Italy: Humanists and University Scholars." In *Beyond Their Sex: Learned Women of the European Past,* edited by P. Labalme. New York: New York University Press, 1980.

La Forge, J. de. *Le cercle des femmes sçavantes.* Paris, 1663.

Le Doeuff, M. "Long Hair, Short Ideas." In *The Philosophical Imaginary.* Stanford, CA: Stanford University Press, 1989.

Lescure, M. de. *Les femmes philosophes.* Paris, 1881.

Ménage, G. *Historia mulierum philosopharum.* Lyon, 1690; English translation, 1702; new English ed. by B. Zedler, Lanham, MD: University Press of America, 1984.

Merchant, C. *The Death of Nature: Women, Ecology, and the Scientific Revolution.* San Francisco: Harper and Row, 1980.

O'Neill, E. "Disappearing Ink: Early Modern Women Philosophers and Their Fate in History." In *Philosophy in a Feminist Voice,* edited by J. Kourany. Princeton, NJ: Princeton University Press, 1998.

Schiebinger, L. *The Mind Has No Sex? Women in the Origins of Modern Science.* Cambridge, MA: Harvard University Press, 1989.

Stanley, T. *A History of Philosophy,* 3 vols. London, 1687.

Waithe, M. E., ed. *A History of Women Philosophers.* 3 vols. Dordrecht: Nijhoff, 1987–1991.

Wilson, K., and F. Warnke, eds. *Women Writers of the Seventeenth Century.* Athens: University of Georgia Press, 1989.

Eileen O'Neill (1996)

WOODBRIDGE, FREDERICK JAMES EUGENE
(1867–1940)

Frederick James Eugene Woodbridge, the American educator, was born in Windsor, Ontario, and attended Amherst College, Union Theological Seminary, and the University of Berlin. He taught philosophy at the University of Minnesota (1894–1902) and Columbia University (1902–1937). At Columbia he also served as dean of the faculty of political science, philosophy, and pure science (1912–1929). Like his colleague John Dewey, he had great influence as a teacher. His influence was less widespread than was Dewey's and was more confined to professional philosophers, but it went deep and is clearly responsible for the revival in the United States of Aristotelian trends of thought. His successor at Columbia University as teacher of the history of philosophy, John H. Randall Jr., is a notable instance of his influence.

REALISM AND NATURALISM

In describing his own philosophical position Woodbridge used the terms *realism* and *naturalism.* By *realism* he

meant that life and mind are products that develop, here and there, in the course of the manifold developments in the natural world. Mind, life, consciousness, and soul are activities of certain types of bodies; they never appear apart from those bodies, although mind, once it has emerged in Nature, may come to guide and thus to master some of the occurrences in the world about it. Consciousness is an awareness of some of the things in the environment; it salutes, as it were, those things. Consciousness, far from being the source of the objective world, presupposes its existence. In all this realistic position Woodbridge regarded himself, quite correctly, as reaffirming in modern terms some basic themes of Aristotle's metaphysics.

By *naturalism* Woodbridge meant much the same thing as he meant by *realism.* Naturalism, he said, "is an attitude and not a doctrine." Some contemporary writers used the word *Nature* to indicate a norm of perfection that the historical processes in this world seldom bring to fulfillment. Others, especially theologians, used it to connote an inferior mode of being, contrasting it with an allegedly superior spirit or supernature. Woodbridge avoided such implied judgments. He wrote, in a hitherto unpublished letter of July 24, 1939:

> Let Nature be, as I love to put it, heaven and earth, the sea, and all that in them is, and I do not see how one can here complain of ambiguity; there is no mistaking what is named by the name Nature. So now I have adopted the practice of spelling Nature with a big N to indicate that it is a name given and not a predicate with implications. It is a name for the clearly identified subject-matter of all inquiry, so that now we can ask what Nature is and proceed at once to look for answers.

In other writers the word *naturalism* often introduced untested presuppositions and undetected prejudices. Woodbridge took Nature as anything and everything we encounter and want to investigate. He abjured "anticipations of nature" and made no commitments, in advance of careful study and research, as to the "interpretations of nature" that investigation would reveal to be proper and true. Nature is what we find around us, whether we are looking on a top closet shelf, or through telescopic instruments at stellar universes that are distant in both time and space, or at the evidences for ancient cities that long ago disappeared from view. Daily life, technical science, and history alike presuppose Nature; that is, all these kinds of quests for knowledge presuppose simply that there is much to investigate. *Naturalism,* in Woodbridge's sense of the term, is not a thesis about what kind of world we have; it is a summons to unbiased research.

Woodbridge's writings reflect, in their form as well as in their content, the attitude he called naturalism and realism. He had no interest in producing an intricate tome designed systematically to account for the existence of everything. Rather, he wrote outstanding essays, in each of which he pushed some one line of analysis as far as he then could. His interests are revealed by the titles of his essays: "Substance," "Teleology," "Creation," "Structure," "Evolution," "Behaviour," "Sensations," "Mind," and "Man." In these essays he examined the question of what thing or process or aspect of the world we isolate for inspection when we speak, for example, of "substance" or "teleology." The positions these essays expose are consistent enough, to be sure. But no one is a premise from which others are deduced; rather, each is a fresh inquiry into some facet of Nature. Moreover, Woodbridge maintained that all the possible investigations that might be undertaken still would not exhaust the intricacies of Nature. We may reach some profound conclusions, but we can never properly say concerning any or all of our conclusions that we have discovered the whole truth about Nature.

TIME AND CHANGE

The most influential of Woodbridge's writings are his discussions of time and change (see particularly Ch. 2 of *The Purpose of History*). Woodbridge argues that what happens at any time is not simply or wholly the effect of what has already happened; an event is dependent upon its past as the material upon which activity may be expended, but it is also a new and fresh expenditure of activity upon that material. What occurs is reconstruction, transformation, remaking. What was is thus pushed back into the past, and what becomes takes the place of what was. Time does not move from past through present to future; rather, it moves from the possible to the actual, that is, from one of the potentialities of what formerly was to a single actuality that is brought into existence by an action (whether that action be unconscious chance or conscious choice) upon what was. What comes to us from the past offers us opportunities and often imposes cruel limitations, but it does not make our choices for us. Rather, it allows us to realize our ends insofar as we have understanding of the potentialities it contains. History has no one end; it includes many processes with their many, often incompatible, ends. And human choices, insofar as they are intelligent, may well be effective to some degree. A natu-

ralistic theory of Nature thus issues in a humanistic theory of man.

See also Aristotelianism; Aristotle; Consciousness; Dewey, John; Metaphysics; Naturalism; Realism.

Bibliography

WORKS BY WOODBRIDGE

The Purpose of History. New York: Columbia University Press, 1916.

The Realm of Mind. New York: Columbia University Press, 1926.

The Son of Apollo: Themes of Plato. Boston: Houghton Mifflin, 1929.

Nature and Mind. New York: Columbia University Press, 1937. A volume of essays presented to Woodbridge on his seventieth birthday by students, colleagues, and friends; contains a bibliography of his writings.

An Essay on Nature. New York: Columbia University Press, 1940.

Aristotle's Vision of Nature. Edited with an introduction by J. H. Randall Jr. New York: Columbia University Press, 1965.

WORKS ON WOODBRIDGE

Cohen, Morris R. *American Thought,* 315–316. Glencoe, IL: Free Press, 1954.

Cohen, Morris R. Chapter 17, "Later Philosophy," of *Cambridge History of American Literature,* Vol. III, pp. 263–264. Cambridge, U.K.: Cambridge University Press, 1921.

Costello, Harry T. "The Naturalism of Frederick J. E. Woodbridge." In *Naturalism and the Human Spirit,* edited by Y. H. Krikorian, 295–318. New York: Columbia University Press, 1944.

Lamprecht, Sterling P. *Our Philosophical Traditions,* 486–497. New York: Appleton-Century-Crofts, 1955.

Randall, John H., Jr. "Dean Woodbridge." *Columbia University Quarterly* 32 (December 1940): 324–331.

Randall, John H., Jr. "Introduction" and "The Department of Philosophy." In *A History of the Faculty of Philosophy, Columbia University,* 3–57, 102–145. New York: Columbia University Press, 1957.

Sterling P. Lamprecht (1967)

WOODGER, JOSEPH HENRY
(1894–1981)

Joseph Henry Woodger, the British biologist, was born at Great Yarmouth, Norfolk. He was graduated from University College, London, where he studied zoology, and after war service returned there to teach. The rest of his academic career was associated with the University of London, as reader in biology from 1922 to 1947 and professor of biology from 1947 to 1959. In the term of 1949-1950 he was appointed Tarner lecturer at Trinity College, Cambridge, whose philosophers—C. D. Broad, Bertrand Russell, and Alfred North Whitehead—greatly influenced his early outlook. Later, the influence of the logicians Rudolf Carnap and Alfred Tarski can be seen in his writings, some of which are highly formal studies of the language and principles of biology. The chief work of his early period is *Biological Principles* (1929); the two best-known works of his later period are *The Axiomatic Method in Biology* (1937) and *Biology and Language* (1952).

Underlying the whole of Woodger's activities as a philosopher of science is his concern with a single problem generated by "the contrast between the brilliant skill, ingenuity and care bestowed upon observation and experiment in biology, and the almost complete neglect of caution in regard to the definition and use of the concepts in terms of which its results are expressed." The effect of this has been to arrest the development of the life sciences. Hence, in *Biological Principles* Woodger proposed to examine a number of key concepts that have entered into the chronic controversies and antitheses of biology, such as those between mechanism and vitalism, preformation and epigenesis, teleology and causation, structure and function, organism and environment, and body and mind. He employed the techniques of analysis made familiar by the Cambridge philosophers of the time. These techniques required clarity and precision in the use of ordinary English expressions, but no use of logical symbolism was introduced. Woodger showed that many of the traditional disputes arose either from failure to eliminate metaphysical elements from biological topics or from shortcomings in the biologists' language, which was often sloppy and imprecise. Trouble was also caused by the implicit adoption of theories of knowledge that were not critically evaluated. He objected to phenomenalism, for example, because the arguments used by phenomenalists presupposed the very knowledge that they declared unattainable—knowledge about brains and sense organs as physical objects in the world. In his own alternative to phenomenalism, Woodger contended that the existence of such objects is a hypothesis that "seems unavoidable for anyone who does not believe that when he uses language he is always talking to himself" (*Biology and Language,* p. 69).

In his subsequent work Woodger turned to mathematical logic as a means of reconstructing the language of biology. Here he made some pioneer contributions. *The Axiomatic Method in Biology* used the machinery of

Whitehead and Russell's *Principia Mathematica* to construct a logical calculus that could be applied to certain nonmetrical concepts of genetics, embryology, and taxonomy. The standard apparatus of logical constants, logical variables, postulates, and theorems was taken over, and to it was added a set of ten undefined "biological constants" together with postulates concerning them. The resulting axiom system permitted the deduction of a number of consequences in the form of precise specifications of such notions as "gametes," "zygotes," "cell hierarchies," "alleles," and so on. A simplified version of this calculus was given in *The Technique of Theory Construction* (Chicago, 1939), in which a specimen theory that is a fragment of the earlier system was neatly developed.

Biology and Language showed how these matters could be approached from the reverse direction. In a section devoted to the reconstruction of the language of genetics, Woodger began not by axiomatizing the set of genetical statements but by recasting observation records in symbolic form and then introducing piecemeal the technical vocabulary needed to move to successively higher levels of theory. This book went beyond classical symbolic logic in its discussion of the language of evolutionary studies, where Woodger developed a special branch of set theory in order to reconcile the gradualness in evolutionary changes with the demand that passage from one taxonomic category to another must take place in one generation.

Logicians have been more appreciative than biologists of Woodger's "experiments" in applied logistic. The abstract formalisms are clear, rigorous, and interesting as logical exercises. Yet although the claims made for them are modest, it might well be argued that it is premature to produce axiomatizations of existing biological knowledge or even that biology is not the sort of science that can be fully reconstructed in axiomatic terms.

See also Broad, Charlie Dunbar; Carnap, Rudolf; Organismic Biology; Philosophy of Biology; Philosophy of Science; Russell, Bertrand Arthur William; Tarski, Alfred; Whitehead, Alfred North.

Bibliography

WORKS BY WOODGER

Biological Principles. London: K. Paul, Trench, Trubner, 1929.

The Axiomatic Method in Biology. Cambridge, U.K.: Cambridge University Press, 1937.

Biology and Language. Cambridge, U.K.: Cambridge University Press, 1952.

WORKS ON WOODGER

Berkeley, E. C. "Conditions Affecting the Application of Symbolic Logic." *Journal of Symbolic Logic* 7 (1942): 160–168.

Carnap, Rudolf. *Introduction to Symbolic Logic and Its Application,* 213–225. New York: Dover, 1958.

Goudge, T. A. "Science and Symbolic Logic." *Scripta Mathematica* 9 (1943): 69–80.

Gregg, J. R., and F. T. C. Harris, eds. *Form and Strategy in Science: Studies Dedicated to Joseph Henry Woodger on the Occasion of His Seventieth Birthday.* Dordrecht, Netherlands: Reidel, 1964.

T. A. Goudge (1967)

WOOLSTON, THOMAS

(1670–1731)

Thomas Woolston, the English divine, religious controversialist, freethinker, and deist, was born in Northampton, the son of a successful tradesman. After schooling there and at Daventry, he entered Sidney Sussex College, Cambridge, in 1685, the same college from which the deist William Wollaston had graduated a few years earlier. Woolston received the BA in 1689 and the MA in 1692. In 1691 he was made fellow of the college and proceeded to take orders, achieving the BD in 1699. The study of Origen early led him to an allegorical interpretation of the Scriptures. He was subsequently accused of derangement of the mind and in 1720 was deprived of his fellowship. Two years later he retaliated by printing and dedicating to the master of the college *The Exact Fitness of the Time in Which Christ Was Manifested in the Flesh, Demonstrated by Reason, Against the Objections of the Old Gentiles, and of Modern Unbelievers,* a discourse that he had delivered twenty years earlier as a public exercise both in the chapel of the college and in St. Mary's Church. The theme of this work is expressed in the words "The first Reason, why *the then Greatness of the Roman Empire was a fit Circumstance of Time for the Mission of Christ,* is, that He might better manifest his Divine Authority and Commission to the civil Powers of the World."

A long series of heterodox religious pamphlets followed that led to unsuccessful prosecution by the government in 1725 and culminated in 1729 with conviction for blasphemy. Woolston was sentenced to a fine of £100, a year's imprisonment, and security for good behavior during life. Failure to meet the fine brought about confinement until his death in January 1731. Samuel Clarke, the rationalistic theologian, had made unsuccessful efforts to get Woolston released. A five-volume edition of Woolston's *Works* was published in 1733.

Woolston's first ironical application of Origen's allegorical method of scriptural interpretation appeared in 1705 under the title of *The Old Apology for the Truth of the Christian Religion Against the Jews and Gentiles Revived.* His anticlerical campaign, particularly directed at those who refused the allegorical way, inspired a number of tracts. *Four Free-Gifts to the Clergy* (1723–1724) accused the "ministers of the letter" of being worshipers of the apocalyptic beast and ministers of Antichrist. *The Moderator Between An Infidel and an Apostate* with its two supplements, all of 1725, continued the attack, the "infidel" being the greatly admired Anthony Collins and the "apostate" being a literal-minded divine. In reality the tracts are defenses of the freethinking Collins and attacks on the clergy who had abandoned the allegorical methods of the Church Fathers.

Another series of tracts from 1727 to 1729 began with *A Discourse On the Miracles of Our Saviour In View of the Present Controversy Between Infidels and Apostates.* Here again Woolston was the disciple of Collins, who had promised to write on the miracles but had never got around to it. In all events, however, Woolston is much more outspoken than Collins would possibly have been. Each of these six tracts, in which he frequently employs the device of an imaginary friend, a learned rabbi, as interlocutor, is ironically dedicated to a different bishop of the Church of England. It is argued that the only evidence for the messiahship of Jesus is found in the Old Testament prophecies, and both prophecy and fulfillment must be interpreted as parables. Many events of Jesus' life (especially the miracles) are patently absurd if given a literal interpretation. Jesus was a spiritual Messiah, healing distempers of the soul, not of the body. Hell, Satan, and the devils are in reality states of mind. Starting with the minor miracles, Woolston deals with fifteen in all, concluding with the Resurrection.

If all of Woolston's allegorizing be madness, there is yet method in it. A man of considerable learning, Woolston employs a racy, colloquial, and frequently witty style. For example, the rabbi comments, "I can't read the Story [of the apparitions of Jesus after his death] without smiling, and there are two or three Passages in it that put me in Mind of Robinson Cruso's filling his Pockets with Biskets, when he had neither Coat, Waste-coat, nor Breeches on."

Up to the last Woolston consistently denied that he was an infidel, avowing that he was a believer in the truth of Christianity. His faith in Christianity is perhaps still open to question, but it is certain that he was a deist, whether rationalistic or Christian. He was never a religious fanatic. Voltaire was much impressed by Woolston's attacks on the miracles and made much use of them.

On all occasions Woolston defended universal and unbounded religious toleration and freedom of thought and of publication. Conversely, he insisted that a hired and established priesthood is the root of all evil, and he vigorously defended such "freethinkers" as the Quakers. Ironically, he was the victim of the authoritarian principles he had dedicated his life to eradicate.

See also Deism.

Bibliography

See *Life* of Woolston, prefixed to Vol. I of his *Works,* 5 vols. (London, 1733); *The Life of the Reverend Mr. Thomas Woolston* (London, 1733); Norman L. Torrey, *Voltaire and the English Deists* (New Haven, CT: Yale University Press, 1930), Ch. 4. See also the general bibliography under the Deism entry.

Ernest Campbell Mossner (1967)

WORLD SOUL

See *Macrocosm and Microcosm*; *Panpsychism*; *Pantheism*

WRIGHT, CHAUNCEY

(1830–1875)

Chauncey Wright, the American philosopher and mathematician, was born in Northampton, Massachusetts. On the surface, his life was completely uneventful. From 1852 to 1870 he worked as a mathematician for the *Nautical Almanac*; he was twice a lecturer at Harvard College—in psychology in 1870 and in mathematical physics in 1874—and he occasionally tutored private pupils. In 1860 he was elected a fellow of the American Academy of Arts and Sciences, of which he was later secretary. He visited Charles Darwin in England in 1872—the major social event of his life. Between 1864 and 1875 he contributed numerous articles to the *North American Review* and the *Nation.* His longer articles were published posthumously in 1877 under the title *Philosophical Discussions*; his *Letters* appeared in 1878.

Wright was not successful as a lecturer, but he was a splendid tutor, and many interested individuals sought to converse with him. It was through this easy interchange of ideas that men such as Charles Sanders Peirce, William

James, and Oliver Wendell Holmes Jr. came to feel the influence of his philosophy. Wright was the mentor of the Metaphysical Club, which met in Cambridge in the early 1870s and included Peirce, James, and Holmes among its members.

ROLE OF SCIENTIFIC CONCEPTS

Wright was America's first technically proficient philosopher of science. He constantly criticized Herbert Spencer as being ignorant of the nature of scientific inference. Spencer tried to assemble all the results of scientific investigation and to fit them together into a total picture of the universe. However, Wright claimed, the theoretical concepts and principles of science are not simply summaries of events; rather, they are tools for extending our concrete knowledge of nature. Theoretical concepts, he said, are finders, not merely summaries, of truth.

Some commentators point out that this "working hypothesis" notion of scientific principles is similar to John Dewey's instrumentalism. According to Dewey, all ideas are working hypotheses and all thinking is experimental, scientific thinking being only a limiting case in the sense of having ideal controls. Wright, however, did not formulate an instrumental view of mind in anything like this general sense. All he did was to emphasize the "working hypothesis" nature of scientific concepts; he did not generalize this interpretation into an account of all thinking. To say that Wright "prefigured" Dewey's brand of pragmatism can mean no more than that he provided the logic of scientific inference that later philosophers generalized into a pragmatic view of mind.

SCIENTIFIC EXPLANATION

Wright distinguished two types of scientific explanation. First, an event can be explained by stating the cause of its occurrence even when it is not possible to show that the characteristics of the event are resultants of any combination of characteristics of the cause. Second, in cases like the parallelogram of forces, one can explain not only the occurrence of an event but also its characteristics as resultants of some combination of characteristics of its cause. Wright felt that some events could never be explained in this second sense, and hence he was advocating, in an embryonic way, a doctrine of emergence. He also believed that this distinction would allow a universal determinist, or necessitarian, to account for novelty and newness in the universe. Furthermore, he thought it provided the means for formulating an enlightened materialist doctrine—namely, that all mental events can be explained by physical events in the first sense but not in the second sense.

EVOLUTION

Wright analyzed the logical structure of evolutionary thought in his articles "The Limits of Natural Selection" (1870), "The Genesis of Species" (1871), and "Evolution by Natural Selection" (1872). He called these articles his definition and defense of Darwinism, and Darwin was sufficiently impressed to reprint "The Genesis of Species" and distribute it in England. Since Wright was answering specific questions, his essays have a piecemeal quality, but they are filled with enlightening points. Of particular interest are his comparison of explanation in biology with explanation in geophysics, his analyses of "accident" and "species," and his defense of "every event has a cause" as a presupposition of scientific investigation.

COSMOLOGY

In his cosmological essays Wright condemned the nebular hypothesis and criticized Spencer's defense of it. He referred to the production of systems of worlds as "cosmic weather." He believed that cosmic events, like ordinary weather, show on the whole no development or any discernible tendency whatever. In the stellar world there is a doing and undoing without end. Wright based his non-developmental view on what he called the principle of countermovements, "a principle in accordance with which there is no action in nature to which there is not some counter-action" (*Philosophical Discussions,* p. 9). He was, obviously, much impressed with the conservation principles of physics. Beginning with his concept of countermovements, and depending primarily upon the first law of thermodynamics and the conservation of angular momentum, he worked out a technical and elaborate hypothesis about the origin of the sun's heat and the positions and movements of planets.

OTHER DOCTRINES

Epistemologically, Wright was in the Humean tradition, but unlike many British empiricists he emphasized the empirical verification of beliefs and was indifferent to the origins of belief. Concerning religion, he was an agnostic. James observed that "never in a human head was contemplation more separated from desire." Wright simply had no desires about God one way or another. In moral philosophy he was a utilitarian, defending, in particular, J. S. Mill's views.

The metaphysical topics that most interested Wright were self-consciousness and a priori knowledge. In *Philosophical Discussions* (pp. 199–266), after sketching a naturalistic account of self-consciousness, Wright tried to show that the notion of substance was meaningless. He believed that ultimate reality consisted of "neutral phenomena" and that the distinction between subject and object is only a classification through observation. Wright's position was essentially a neutral monism and was a precursor of William James's notion of pure experience.

Unlike most nineteenth-century philosophers, Wright did not deny the existence of a priori knowledge. Quite to the contrary, he insisted that all knowledge, even the perception of qualities as well as relations and abstract concepts, has an a priori element, and this element can be explained experientially (*Letters*, pp. 123–135). This analysis is particularly interesting at the present, since we are currently offered various forms of "factual" or "pragmatic" concepts of a priori knowledge.

See also Cosmology; Darwin, Charles Robert; Darwinism; Dewey, John; Evolutionary Theory; Explanation; James, William; Knowledge, A Priori; Mill, John Stuart; Peirce, Charles Sanders; Pragmatism; Utilitarianism.

Bibliography

WORKS BY WRIGHT

Philosophical Discussions. Edited by C. E. Norton. New York: Holt, 1877.

Letters. Edited by J. B. Thayer. Cambridge, MA: Wilson, 1878.

Philosophical Writings. Edited by Edward H. Madden. New York: Liberal Arts Press, 1958.

WORKS ON WRIGHT

Madden, Edward H. *Chauncey Wright.* New York: Washington Square Press, 1964.

Madden, Edward H. *Chauncey Wright and the Foundations of Pragmatism.* Seattle: University of Washington Press, 1963. Contains complete bibliography of articles on Wright.

Edward H. Madden (1967)

WRIGHT, GEORG HENRIK VON

(1916–2003)

Georg Henrik von Wright held the Swedish language chair of philosophy at the University of Helsinki from 1946 through 1948 and from 1952 through 1961; in between he was professor at the University of Cambridge (1948–1951). From 1961 until his retirement he was a research professor in the Academy of Finland. A member of the Swedish-speaking minority in Finland, von Wright lived almost all of his life in Helsinki.According to von Wright, the major influences on his philosophy were Eino Kaila, an important and charismatic figure in Finnish philosophy; G. E. Moore; and Ludwig Wittgenstein.

Kaila sparked von Wright's interest in formal matters and his use of logical methods. Moore's writings may have inspired von Wright's unpretentiousness and unrelenting quest for clarity. Wittgenstein had a profound personal influence on von Wright—he was Wittgenstein's student, then his successor as professor in Cambridge, and finally, with G. E. M. Anscombe and Rush Rhees, one of his literary executors. Yet Wittgenstein's philosophical influences on von Wright's work are less apparent.

Throughout life von Wright combined, to an extent that is not common among today's academic philosophers, two rather different approaches to philosophy: one the passionate commitment of the humanist and the other the detached objectivity of the scholar. The former approach is exemplified by a number of books in Swedish such as *Tanke och förkunnelse* (Thought and prophecy) (1955), *Humanismen som livshållning* (Humanism as a way of life) (1978), and *Vetenskapen och förnuftet* (Science and reason) (1986). With his largely pessimistic views about the future of humankind, von Wright has won wide public acclaim in the Nordic countries, particularly in Sweden.

In the rest of the world, von Wright is best known for his academic work. He wrote on induction and probability (*The Logical Probability of Induction* [1941]; *A Treatise on Induction and Probability* [1951]) and on ethics (*The Varieties of Goodness* [1963]). But his main reputation lies in modal logic and in the theory of action. In *An Essay in Modal Logic* (1951), von Wright developed his method of distributive normal forms and analyzed a number of modal systems, one of which is nowadays usually referred to the Gödel/Feys/von Wright system T. In this work von Wright recognized the possibility of modal logics of knowledge and belief (that is, logics in which the modal box operator is interpreted as "the agent knows that" or "the agents believes that"); it was he who introduced the terms *epistemic logic* and *doxastic logic*, respectively, for these kinds of logic. This theme was later developed in great detail by von Wright's countryman and one-time student Jaakko Hintikka.

Von Wright's paper "Deontic Logic" in *Mind* (1951) opened up the new field of deontic logic and was the first in a long series of papers and books in which von Wright

elaborated and deepened his analysis. One important insight was that the fruitful study of deontic logic requires a logic of action as a basis, and in *Norm and Action* and many later works he tried to lay the foundations of such a logic. He is unique among early action theorists in letting his formal logic of action inform the philosophy of action and vice versa.

According to von Wright, to act is to interfere with the course of nature—to bring about a change, to bring about an event. This view led him to question the relationship between action and causality and eventually convinced him that an explanation of human action in purely causal terms will always leave out something important. In *Explanation and Understanding* (1973), he presented an influential examination of practical syllogisms: although they cannot possess logical validity in the ordinary sense, nevertheless they may be accepted as explanations *ex post actu*.

See also Anscombe, Gertrude Elizabeth Margaret; Ethics, History of; Hintikka, Jaakko; Humanism; Induction; Modal Logic; Moore, George Edward; Probability and Chance; Wittgenstein, Ludwig Josef Johann.

Bibliography

WORKS BY VON WRIGHT

The Logical Problem of Induction. Acta Philosophica Fennica, fasc. 3. Helsinki: Societas Philosophica, 1941.

"Deontic Logic." *Mind* 60 (1951): 1–15.

"An Essay in Modal Logic." *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland, 1951.

A Treatise on Induction and Probability. International Library of Psychology, Philosophy, and Scientific Method. London: Routledge& Kegan Paul, 1951.

Tanke och förkunnelse (Thought and prophecy). Helsinki: Söderström, 1955.

Norm and Action: A Logical Inquiry. London: Routledge & Kegan Paul, 1963.

The Varieties of Goodness. London: Routledge & Kegan Paul, 1963.

Explanation and Understanding. Ithaca, NY: Cornell University Press; London: Routledge & Kegan Paul, 1971.

Humanismen som livshållning och andra essayer (Humanism as a way of life and other essays). Helsinki: Söderström, 1978.

Vetenskapen och förnuftet: ett försök till orienteering (Science and reason: An attempt at orientation). Helsinki: Söderström, 1986.

BOOKS ABOUT VON WRIGHT

Schilpp, P. A., and L. E. Hahn. *The Philosophy of Georg Henrik von Wright*. Vol. 19 of *The Library of Living Philosophers*. La Salle, IL: Open Court, 1989.

Krister Segerberg (1996, 2005)

WUNDT, WILHELM

(1832–1920)

Wilhelm Wundt, the German philosopher and psychologist who founded the first psychological laboratory and won world fame as a teacher and scholar, was born in Neckarau, a suburb of Mannheim. After studying medicine at the universities of Tübingen, Heidelberg, and Berlin, he was a *Privatdozent* from 1857 to 1864 at the Physiological Institute founded by Hermann von Helmholtz in Heidelberg. At the age of twenty-four he became so severely ill that he was given up by his physicians and remained close to death for several weeks. In this time of crisis he developed his most essential religious and philosophical views, and also his ideas concerning the mental.

In a series of contributions to the theory of sense perception, published between 1858 and 1862, Wundt's interest in psychological problems, an interest derived from his physiological studies, becomes clear. He gave his first psychological lecture in 1862, and in 1863 his *Vorlesungen über die Menschen- und Tier-Seele* (2 vols., Leipzig, 1863, translated by J. G. Creighton and E. B. Titchener as *Lectures on Human and Animal Psychology,* London, 1896). A series of lectures given in 1864 on the fundamentals of physiological psychology was published at Leipzig in 1874 as *Grundzüge der physiologischen Psychologie* (translated by E. B. Titchener as *Principles of Physiological Psychology,* New York, 1904), his chief work. In the same year Wundt was called to the professorship in inductive philosophy at Zürich. In 1875 he accepted a call to Leipzig, where he founded the world's first experimental laboratory in psychology, the Institut für Experimentelle Psychologie, in 1879. Students from many countries throughout the world became devoted disciples and returned home to found similar institutions.

As a young man in Heidelberg, Wundt was a member of the Baden Stände assembly and the presiding officer of the Heidelberg Society for Workingmen's Education; he was in favor of a patriotic socialism. During the Franco-Prussian War of 1870–1871 he served as an army doctor. As an old man he was rector of Leipzig University (1900) and was overwhelmed with national and international honors and titles. Although in his last years he was practically blind, he did not retire from his teaching position until 1917. A philosophical autobiography was prepared for publication in the year of his death in Grossbothen, near Leipzig.

PHILOSOPHY

As a philosopher Wundt was self-taught. He published a system of logic (*Logik,* 2 vols., Stuttgart, 1880–1883; 4th and 5th eds., 3 vols., 1919–1924), a system of ethics (*Ethik,* Stuttgart, 1886; 5th ed., 3 vols., 1923–1924), and a system of philosophy (*System der Philosophie,* Leipzig, 1889; 4th ed., 2 vols., Leipzig, 1919) during the 1880s. He later wrote on historical subjects (*Die Nationen und ihre Philosophie,* 1915; *Leibniz,* 1916). Wundt was a voluntarist and a follower of the German school of idealism; as such he was indebted to Gottfried Wilhelm Leibniz in particular, and also to Arthur Schopenhauer and G. W. F. Hegel. He opposed sensationalism, materialism, and the relativity of values; nevertheless, he drew ideas from contemporary positivism, particularly in his eclectic historicism and his theoretical inclination to a sociological collectivism. This positivist tendency, noticeable until the middle of his career, especially as a kind of defense against metaphysics, was overcome late in his life. Wundt's main concern in logic was exactness in formal derivations; in ethics it was to secure the Leibnizian morality, based on duty, against contemporary utilitarianism and hedonism on the one hand and subjectivism and relativism on the other. Wundt also essentially followed Leibniz in his parallelist treatment of the mind-body problem.

GENERAL PSYCHOLOGY

If in his philosophy Wundt was primarily an eclectic and historical encyclopedist, he demonstrated his originality in psychology, where he achieved worldwide fame as the real founder of the science and its methodology. However, he was far from wanting to destroy the interconnection between psychology and philosophy. He regarded psychology as the common basis for all scientific and cultural knowledge and the bond uniting all the individual sciences, and therefore as the "science directly preparatory to philosophy."

Nevertheless, Wundt resisted "psychologism" as later formulated and criticized by Edmund Husserl—that is, the reduction of cultural organization and normative evaluations to mere mental processes and the relativization of the timelessly valid to the mere here and now in consciousness.

One of Wundt's main concerns was to investigate conscious processes in their own context by experiment and introspection. He regarded both of these as "exact methods," interrelated in that experimentation created optimal conditions for introspection. Where the experimental method failed, Wundt turned to other "objectively valuable aids," specifically to "those products of cultural communal life which lead one to infer particular mental motives. Outstanding among these are speech, myth, and social custom." Wundt's two main fields of investigation and his two main works, the *Physiologische Psychologie* and his *Völkerpsychologie* (Folk psychology, or Psychology of nations; 2 vols., Leipzig, 1904; 3rd ed., 10 vols., Leipzig, 1911–1920), correspond to this methodological division.

As a follower of Leibniz, Wundt maintained a strict psychophysical parallelism in his basic concepts and rejected any form of theory of reciprocal interaction (causation); however, he limited the mental to the realm of conscious events ("the actual"), in what F. A. Lange referred to as "psychology without soul." Experience should be investigated in its context, "as it is actually given to the subject." In contrast with the natural sciences, the subject matter of psychology is "the content of experience in its immediate nature, unmodified by abstraction and reflection." This claim, which in today's terminology is a strictly phenomenological one, was accompanied by a demand for explanations derived from strict necessity and based on as complete an analysis as possible of the direct, complex findings. Wundt modified the categories of explanation by assuming a unique "psychic causality," which he sought to distinguish from scientific or mechanical causality as including motivation. At this point in his thinking, again following Leibniz, he fought against British and French sensationalism and materialism.

Despite his stress on analytic observation, many notions of Wundt's psychology are transitional to the modern *Ganzheitspsychologie* (psychology of totalities, psychology of wholes) of Felix Krueger and others, among them the "principle of creative resultants or synthesis," which allows perception to transcend a mere addition of stimuli; the "unity of the frame of mind"; and the "value-grade of the total," or feeling and emotion. In his theory of the types of feelings Wundt went beyond the narrow dimensions of pleasure and displeasure, and developed the concept of "total feeling." Although Wundt sought to investigate the elements of conscious processes and their connecting forms, he cannot be counted among the classical sensationalist psychologists because his theory of actuality refers to constantly changing processes rather than to static elements.

Wundt designated the basic mental activity "apperception." Apperception is a unifying function that should be understood as an activity of the will. Feelings are attitudes adopted in apperception toward its individual contents. Thus apperception is simultaneously a descriptive

and an explanatory concept. It remained for Krueger, Wundt's pupil and his successor at Leipzig, to remove the limitation to the "pure mental actuality" (structural psychology) and thereby pave the way for the psychology of personality.

Many aspects of Wundt's empirical physiological psychology are still fruitful today. Among them are his principles of mutually enhanced contrasts and of assimilation and dissimilation, for instance, in color and form perception, and his advocacy of "objective" methods of expression and of recording results, especially in language. Another is the principle of heterogony of ends, which states that multiply motivated acts lead to unintended side effects, which in turn become motives for new actions.

SOCIAL PSYCHOLOGY

Wundt believed that his principles of physiological psychology were provable and confirmable in the nonexperimental realm of social, developmental, or cultural psychology, which he called *Völkerpsychologie.* In this field sociological considerations, and particularly the encyclopedic presentation of materials from history and from the other *Geisteswissenschaften* (roughly, "cultural and social sciences," or "humanities"), became Wundt's main concern, overshadowing actual psychological questions. The "objective products of the collective intellect" in nations—speech, myth (religion), and social custom (law)—that were the original subjects of *Völkerpsychologie* came in practice to include social structures and the arts. In Wundt's analysis, which he applied to an incredible amount of material and which was necessarily modified by later progress in the cultural and social sciences, the principle of the social, prehistoric, collective determination of intellectual development dominated. Concern with the individual and with individual development was neglected for this sociogenetic problem. There is, besides, a methodological gap between phenomenological and experimental psychology and cultural psychology, as was emphasized by Wilhelm Dilthey and Eduard Spranger, wide enough to endanger the unity of psychology.

Despite the outmoded material it contains, Wundt's gigantic lifework still offers a powerful inspiration that has never been totally exhausted, at least partly because, since his time, psychology and the *Geisteswissenschaften* have continued to move further apart. Felix Krueger said at Wundt's grave, "In him faithfulness to fact was raised to the level of genius." Thoroughness and methodical acuity, combined with universal versatility, created something unique in his work. Wundt has been extolled as the last "polyhistor." Education and aesthetics were the only fields to which he made no contribution. E. G. Boring computed his total published output at 53,000 pages—an entire library. The complete list of his works, published by his daughter Eleonore Wundt in 1926, is a hefty brochure. In both philosophy and psychology Wundt's oscillation between idealistic and positivistic tendencies kept him bound to his time and caused a notable lack of consistency. He was a major pioneer of both scientific and cultural psychology, even though he was unable to integrate them. The unity of all sciences through psychology and the development of philosophy out of psychology remain as transient theoretical postulates unrealizable and unrealized by developments since his death.

See also Apperception; Dilthey, Wilhelm; Geisteswissenschaften; Hegel, Georg Wilhelm Friedrich; Helmholtz, Hermann Ludwig von; Husserl, Edmund; Idealism; Introspection; Krueger, Felix; Lange, Friedrich Albert; Leibniz, Gottfried Wilhelm; Materialism; Mind-Body Problem; Phenomenology; Positivism; Psychology; Schopenhauer, Arthur; Sensationalism; Spranger, (Franz Ernst) Eduard; Voluntarism.

Bibliography

ADDITIONAL WORKS BY WUNDT

Erlebtes und Erkanntes, Selbstbiographie (Things experienced and perceived, autobiography). Stuttgart, 1920.

Wilhelm Wundts Werk, ein Verzeichnis seiner sämtlichen Schriften (Wilhelm Wundt's work, a list of his complete writings). Edited by Eleonore Wundt. Munich: Beck, 1927.

WORKS ON WUNDT

Boring, E. G. *A History of Experimental Psychology,* 2nd ed., 318–347. New York: Appleton-Century-Crofts, 1950.

Heussner, A. *Einführung in Wilhelm Wundts Philosophie und Psychologie.* Göttingen: Vandenhoeck and Ruprecht, 1920.

Hoffmann, Arthur. "Wilhelm Wundt, eine Würdigung." *Beiträge zur Philosophie des deutschen Idealismus* 2 (1922).

König, Edmund. *W. Wundt, Seine Philosophie und Psychologie.* Stuttgart: Frommann, 1901; 3rd ed., 1912.

Nef, Willi. *Die Philosophie Wilhelm Wundts.* Leipzig: Meiner, 1923.

Peters, R. S., ed. *Brett's History of Psychology,* 479–488. London, 1953.

Petersen, Peter. *Wilhelm Wundt und seine Zeit.* Stuttgart. 1925.

A Wilhelm Wundt Archive, established by his daughter in his house at Grossbothen, was transferred to the Psychological Institute of the University of Leipzig at her death and is administered by the institute.

Albert Wellek (1967)
Translated by Tessa Byck

WYCLYF, JOHN

(c. 1320–1384)

John Wyclyf, the scholastic philosopher and ecclesiastical reformer, was born in the north of England, near Richmond. He spent most of his adult life in and around Oxford; he served several parishes as priest and held a series of prebends that gave him a modest income. On several occasions he was asked his opinion in matters of government policy toward the papacy, and he appeared once before Parliament. In 1374 Wyclyf was a member of a royal commission of three that met with representatives of the papal Curia at Bruges to attempt to solve the impasse between England and the papacy over England's refusal to pay the Peter's pence. Later he became an adherent of and adviser to the duke of Lancaster, John of Gaunt, who protected Wyclyf when, under pressure from the English hierarchy, he was charged with heresy. Wyclyf retired, probably on Lancaster's advice, from active public life to his parish at Lutterworth in 1382. In that year he suffered a paralytic stroke but continued his prolific writing until his death, from a second stroke, two years later.

Wyclyf's literary life may be divided into three periods. During the first period, from about 1358 to 1372, he was primarily an academic philosopher, lecturing on logic and metaphysics in orthodox terms. During the second period, from 1372 to 1377 or 1378, he began to apply his realist philosophy to the problems of church and state, an application that resulted in his doctrine of dominion. In the last period, from 1377 or 1378 to 1384, he went much further in his investigation of the basis and structure of the Roman church and came to conclusions quite openly antipapal. During this period papal bulls were aimed against him (1377); he was twice haled before local bodies on orders from Rome; and many of his conclusions were specifically condemned, although he was not personally disciplined. These same conclusions, in addition to many more, were condemned by the Council of Constance in 1415.

Wyclyf's philosophical presuppositions colored all his thought. The transition from one period of his life to another was barely perceptible and he was able, late in his life, to refer to his earlier expressions with few apologies. In the atmosphere of mid-fourteenth-century Oxford, Wyclyf early had to take a position toward the *universalia post rem* of William of Ockham's nominalism, then popular and persuasive. He rejected its priority of the particulars over universals in favor of the older Augustinian tradition of *universalia ante rem.* Once he had accepted this position, he followed it to its logical conclusions and constructed a *summa de ente* in twelve books that, while not so systematic as most other *summae* of the thirteenth and fourteenth centuries, nevertheless dealt in great detail with the salient points of dispute between the nominalists, the *doctores moderni,* as he called them, and the protagonists of universal ideas.

THE *SUMMA DE ENTE*

Following his early works on logic, written probably between 1360 and 1365, Wyclyf's *Summa de Ente* occupied him until at least 1370, when his attention was diverted to theology. The *Summa* in its final form consists of two books of six treatises each. The first book treats being in general, the doctrine of universals, and the nature and function of time. These questions are approached from the point of view of man and his cosmos. The second book is pure theology: God's intellection, his knowledge, his will, the Trinity, his ideas and his power to create outside himself. In Wyclyf's grand design the first book is anthropology and the second book is theology. Universals thus may be considered the human parallel of God's ideas. Knowing only the *Timaeus* of Plato's works, Wyclyf adhered to Plato as he knew him from Augustine. His realism was uncompromising. Universals exist *ante rem,* temporally and logically prior to the particular. "The idea is therefore essentially the divine nature and formally the *ratio* according to which God intelligizes [*intelligit*] creatures." These ideas make up the creative mind of God. In a parallel fashion the universal (on man's level) is its singular. The singular participates in its universal, which is by nature a projection of an idea in the mind of God. As a creation of God's mind, the singular is incapable of annihilation. For God to allow a singular to be annihilated would be to permit the annihilation of a part of himself—an obvious impossibility.

As he articulated this line of thought, Wyclyf was led to examine the church's doctrine of transubstantiation. He reasoned that the church held that in the Eucharist the substance of bread and wine was annihilated. From about 1379 he attacked the doctrine vehemently on purely logical and philosophical grounds. This position in turn was bitterly attacked by orthodox theologians and later formally anathematized at the Council of Constance. In view of his basic realism Wyclyf could not have done otherwise than he did.

THE CHURCH

About 1374 Wyclyf had begun a spirited defense of the doctrine of dominion. This concept of the sanctions of power was rooted in Augustine and had recently been

propounded by Richard FitzRalph, archbishop of Armagh in Ireland. Dominion or lordship is founded in grace, and he who is without grace has no proper right to exercise dominion. Applied to the religious hierarchy, it would have deprived many of the higher clergy of their power and emoluments.

In 1378 Wyclyf was led, by an incident involving the theory and practice of sanctuary, to examine the nature of the church and the relations of the papacy with the English crown. In the course of the dispute arising from the publication of his views, he came to the clear conclusion that the pope and the cardinalate were unnecessary and that in England the king should control the church, allowing for counsel and advice of theologians in matters of theology.

Wyclyf was a stout defender of the Pauline-Augustinian doctrine of predestination, which he related to and strengthened with his doctrines of universals and necessity. The implications of predestination did not favor a highly organized ecclesiastical organization; if a believer is predestined by God to salvation from all eternity, the church would soon have no reason for existence. Individualism in religious matters could hardly be tolerated by the establishment.

In the last years of his life Wyclyf composed a second *summa,* a *Summa Theologica,* also in twelve books. Not a *summa* in the thirteenth-century style, it was a series of polemical treatises concerned with problems in church or national polity, in defense of his contested opinions. In presentation he remained a Schoolman to the end, but his ideas were disruptive of the establishment, and opposition, at Oxford and in London, was determined and ruthless. The opposition to his efforts at reform is somewhat surprising, in view of his highly pronounced English nationalism; but English clerics were his bitterest opponents. In Wyclyf's view, his thought and action were consistent and consistently rooted in the doctrine of divine ideas, the creative *rationes* by which the universals existed before the particular and were exhibited in the particular, *essentialiter, formaliter, et eternaliter.*

See also Augustine, St.; Augustinianism; Determinism, A Historical Survey; Medieval Philosophy; Plato; Realism; William of Ockham.

Bibliography

The Wyclif Society published 33 volumes of Wyclyf's works (London, 1883–1922) but omitted some important philosophical treatises. See also S. H. Thomson, *Joh. Wyclif Summe de Ente libb. I et II* (Oxford: Clarendon Press, 1930; reissued Boulder, CO, 1956), a Latin text with critical introduction; and A. D. Breck, ed., *Joh. Wyclyf De Trinitate* (Boulder: University of Colorado Press, 1962).

The standard life of Wyclyf is H. B. Workman, *John Wyclif,* 2 vols. (Oxford: Clarendon Press, 1926). See also K. B. McFarlane, *John Wycliffe and the Beginnings of English Nonconformity* (London: English Universities Press, 1952); and J. A. Robson, *Wyclif and the Oxford Schools* (Cambridge, U.K.: Cambridge University Press, 1961).

S. Harrison Thomson (1967)

X

XENOPHANES OF COLOPHON

(c. 570 BCE—c.475 BCE)

Like the other founders of Greek philosophy, Xenophanes lived in Ionia and investigated natural phenomena such as the basic substances, the history and structure of the cosmos, and weather phenomena. He is best known for his criticisms of religious beliefs and practices, for his own conception of the divine, and for being the earliest philosopher to discuss epistemological questions. A poet who traveled widely in Greek lands, he composed his philosophical work in verse, presumably for performance, which suggests that his radical theological views were not abhorrent to his audiences. Some forty fragments of his writings survive, more than one hundred lines, far more than what remains from any earlier philosopher.

His *theological* fragments consist in statements that seemingly criticize the anthropomorphic polytheism of Greek tradition and in pronouncements on the true nature of god. He claims that (just like the Greeks) Ethiopians and Thracians believe their gods look like themselves (frag. 16) and that if animals could draw, horses would depict their gods as horses, oxen as oxen, etc. (frag. 15). He reproaches the revered poets Homer and Hesiod for ascribing to the gods actions humans consider immoral (frag. 11). He does not argue that these diverse accounts of the divine are false or even contradictory, but the remark about animals seems intended to ridicule the differing human (including Greek) beliefs about the gods. Nor is the reproach about the gods' behavior an argument, but it further undermines tradition: Greeks not only think the gods are like humans, they think they are immoral too!

Abandoning the Olympian gods led Xenophanes not to atheism but to new opinions on the nature of the divine and a new way of apprehending it. God "always remains in the same place, moving not at all" (frag. 26); "not at all like mortals in body or thought" (frag. 23); "is one, greatest among gods and men, all of him sees, all of him thinks, all of him hears" (frag. 24); "without toil he shakes all things by the thought of his mind" (frag. 25). Fragments 24 and 25 probably assert omniscience and omnipotence. Xenophanes presents a nonanthropomorphic god possessing cognitive abilities corresponding to human ones but far exceeding humans in power. It is a theistic account since "shakes all things" seems to mean that god controls and causes all events in the cosmos. Xenophanes may also have been a monotheist. If so, he

was the first Greek to adopt this revolutionary view. The relevant text is fragment 23, whose opening words can be translated either "god is one" or "one god." The next phrase, "greatest among gods and men," suggests a plurality of gods, so the god Xenophanes describes would be the supreme god but not the only one. But it can be objected that his criticisms of the traditional anthropomorphic gods and his belief in a supreme god that governs everything tell against polytheism. This objection is reinforced by the report that he said it is unholy for any god to have a master and that no god is deficient in anything at all (Testimony 32), claims hard to square with a belief that combines polytheism with a single supreme deity. These are strong motives for taking "among gods and men" not to imply polytheism. One way is to take it as a *polar expression*, as if an atheist said that there is no god in heaven or earth, using "in heaven or earth" (ironically) to mean simply "anywhere." But many are dissatisfied by this solution, and there is no consensus on the question of Xenophanes's monotheism.

Xenophanes gives no argument for the existence or the nature of his supreme deity. He seems not to have questioned the existence of the divine. The only reason given for any of its attributes is that "it is not fitting for him to go to different places at different times" (frag. 26). Not tradition or other authority, but Xenophanes' sense of what befits the divine, is his criterion for determining god's nature. In this limited sense we find in Xenophanes the beginnings of rational theology.

Three fragments introduce important issues in epistemology although their meaning is disputed. "By no means did the gods intimate all things to mortals from the beginning, but in time, by searching, they discover better" (frag. 18) may refer specifically to the intellectual progress being made by Xenophanes and his fellow early philosophers and emphasize the importance of empirical work for making advances. Certainly, some of Xenophanes's new ideas on natural phenomena were based in observation and investigation, as opposed to mere theorizing. "No man has seen nor will anyone know the clear truth about the gods and all the things I speak of. For even if someone were to say exactly what has been brought to pass, he still does not know, but belief is fashioned over all things" (frag. 34) distinguishes truth, knowledge, and belief and denies that true beliefs and assertions amount to knowledge. It may indicate a skepticism about the possibility of acquiring knowledge of the subjects studied by the early philosophers. If so, the progress heralded in fragment 18 must fall short of certain knowledge. We must remain with beliefs, which may be better or worse: They may be better or worse supported by investigations, which themselves may be more or less thorough and careful. Fragment 35, which may be the conclusion of Xenophanes's discussion of these topics, advises, with modesty uncharacteristic of the Presocratics: "Let these things be believed as like the truth." Xenophanes's views remain on the level of beliefs; if he has *searched* well, his views will be *better*—possibly true or closer to the truth than conflicting views. But even if they are, they cannot be known to be more like the truth, only believed to be so.

See also Epistemology; Homer; Philosophy of Religion.

Bibliography

TEXTS

Diels, Hermann, and Walther Kranz, eds. *Die fragmente der Vorsokratiker, griechisch und deutsch*. Vol. 1, 6th ed. Berlin: Weidmann, 1951, p. 126–138.

Heitsch, E. *Xenophanes: Die fragmente*. Munich: Artemis, 1983.

Lesher, James H. *Xenophanes of Colophon: Fragments; A Text and Translation with Commentary*. Toronto: University of Toronto Press, 1992.

Studies

Barnes, Jonathan. *The Presocratic Philosophers*. Rev. ed. London: Routledge & Kegan Paul, 1982, p. 82–99, 137–143.

Fränkel, Hermann. "Xenophanesstudien," *Hermes* 60, 1925, p. 174–192. Reprinted in *Wege und Formen frühgriechischen Denkens*. 3rd ed. Munich: Beck, 1968. Translation of part of this paper by M. R. Cosgrove under the title "Xenophanes' Empiricism and his Critique of Knowledge" in *The Pre–Socratics: A Collection of Critical Essays*. Edited by Alexander P. D. Mourelatos. Garden City, NY: Doubleday, 1974, p. 118–131.

Guthrie, W. K. C. *A History of Greek Philosophy*. Vol. 1. Cambridge, U.K.: Cambridge University Press, 1962, p. 360–402.

Kirk, Geoffrey S., John E. Raven, and Malcolm Schofield, eds. *The Presocratic Philosophers: A Critical History with a Collection of Texts*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1983, p. 163–180.

Richard McKirahan (2005)

XENOPHON
(c. 430 BCE–c. 350 BCE)

Xenophon was an Athenian citizen, soldier, gentleman-farmer, historian, and author of many varied and often graceful prose works. When young he knew Socrates, whom he consulted before joining, in 401, the famous expedition to Persia narrated in his masterpiece, the

Anabasis. Xenophon played a part in leading the defeated remnant back to Greece. Meanwhile, in 399, Socrates had been executed on trumped-up charges. In the subsequent pamphleteering, Xenophon wrote in Socrates' defense. His so-called *Apology of Socrates* is an unconvincing footnote to Plato's; but later he compiled his extensive and valuable *Memorabilia* (Recollections of Socrates) the work that has given Xenophon, not himself a philosopher, considerable importance to all post-Socratic philosophers. In it Xenophon supplemented his defense of Socrates against specific charges (made in a pamphlet by Polycrates) with a more general description of his character as a man, a friend, and a teacher, strongly emphasizing his beneficial influence on all who knew him and, for illustration, recording many conversations in which Socrates' views or methods were displayed. Xenophon claimed to have heard many of these conversations himself; others were reported to him by friends among the original interlocutors. Some longer sequences of conversations follow up related topics, but individual conversations are never sustained as long as even a short Platonic dialogue.

Undeniably, Xenophon's Socrates is less lively in discussion than Plato's and far less impressive in defending his paradoxes. The difference reveals the gulf between Plato and his contemporaries in literary skill and in philosophical understanding. But there is no need to reject Xenophon's testimony, despite persistent attacks by scholars on his honesty. Xenophon's picture of Socrates is his own, drawn from his own and his friends' memories of Socrates, not plagiarized from other "Socratic" writers any more than from Plato; it is authenticated precisely by its failings. Xenophon saw Socrates as a man of enormously strong moral character and a teacher of moral principles revolutionary for their day in their demand for unselfishness and self-control. Xenophon only half understood the philosophical significance of Socrates's views, and for fuller understanding we must turn to Plato; but Xenophon occasionally added important details, and with allowance for his limitations an impression of Socrates can be obtained from him that helps us to discern very generally the area in which Plato was presenting his own arguments and no longer those of Socrates.

Xenophon's Socrates demonstrates repeatedly the practical importance of knowledge. He advises young men ambitious to be generals and politicians to acquire knowledge, and draws analogies to show that all skills must be learned; he discusses their respective skills with a painter, a sculptor, a breastplate maker, and even, humorously, with a courtesan. He does not try, as Plato's Socrates did, to question the significance of the craftsmen's knowledge, but only to show that their knowledge can be usefully increased by deeper understanding of the purposes of their various crafts. In turn, he is suspicious of the purely theoretical study of astronomy and geometry beyond their practical uses. Xenophon stresses, nevertheless, that Socrates himself was not ignorant of theoretical science.

Xenophon does not quote in so many words the Socratic paradox "no one errs voluntarily," but he does state that Socrates did not distinguish knowledge from self-control and identified justice and all other virtues with knowledge; knowledge of justice or piety is what produces the just or pious man. Characteristically, however, he repeatedly shows Socrates warning against "weakness of will," and forgets that in the Socratic view, strictly speaking, this could not occur; his admiration of Socrates' own self-control leads him to praise self-control as an independent virtue.

Xenophon occasionally reproduces a Socratic elenchus, or interrogation demonstrating an interlocutor's ignorance, and comments that Socrates used this method to stimulate moral improvement in his pupils by inducing them to acquire knowledge. Xenophon shows no grasp of elenchus as a philosophical weapon for testing arguments, nor indeed of the Platonic Socrates' insistence that consciousness of one's ignorance may be the best one can achieve. Xenophon's Socrates uses no "irony," but states positive views quite unreservedly. He is interested in definitions and unlike Plato's Socrates confidently provides them; rather surprisingly, he is willing to define *good* and *beautiful* as relative to utility. Perhaps out of many suggestions intended by Socrates to be tentative, or to show the difficulties of definition, Xenophon—in pursuit of certainty—isolated a few solutions as final.

Xenophon at one point describes Socrates' method as "leading the discussion back to its basic premise (*hypothesis*)" by establishing, for example, an agreed general definition of the good citizen before assessing a particular citizen's goodness; he tells us that Socrates regarded agreement in discussion as the best guarantee against error. This account of *hypothesis* is much simpler than Plato's in either *Meno* or *Phaedo,* but it is abundantly exemplified in Plato's early dialogues. Xenophon nowhere ascribes to Socrates any theory of Forms, but he quotes a suggestion of Socrates that etymologically "to perform dialectic" means "to arrange things in classes."

Xenophon's entertaining *Symposium* (Banquet) and *Oeconomicus* (Household management) display Socrates taking part in sustained discussions; but here this is a lit-

erary device with no biographical intention, and in any case little is attributed to Socrates. Xenophon's idealizing *Cyropaedia* (Education of Cyrus) shows very slight Socratic influence.

See also Medieval Philosophy; Plato; Socrates; Universals, A Historical Survey.

Bibliography

Xenophon's complete works may be found in both Greek and English in the seven-volume Loeb Classical Library edition (London and New York, 1914–1925). The *Memorabilia* in the Loeb edition, translated by E. C. Marchant, appeared in 1923.

For studies on Xenophon's Socrates, see A. Delatte, *Le troisième Livre des souvenirs socratiques de Xénophon* (Liège: Faculté de Philosophie et Lettres, 1933), and R. Simeterre, *La théorie socratique de la vertuscience selon les "Mémorables" de Xénophon* (Paris, 1938).

David B. Robinson (1967)

XENOPHON [ADDENDUM]

See Appendix, Vol. 10

XUNZI

(fl. 295–238 BCE)

Among the classical Confucian thinkers of the Warring States period (Zhanguo 475–221 BCE), Xunzi plays a commanding role in the systematic development and defense of Confucian Tradition. Xunzi's teachings are contained in the *Xunzi*, compiled by Liu Xiang of the Former Han (206 BCE–8 CE). Although some scholars have questioned the authenticity of some of the essays, this work shows remarkable coherent and reasoned statements of the central aspects of the Confucian ethical and political vision of a harmonious and well-ordered society. Moreover, especially impressive is Xunzi's wide-ranging interest in such timeless issues as the ideal of the good human life, relation between morality and human nature, the nature of deliberation, ethical discourse and argumentation, moral agency and moral knowledge, the ethical significance of honor and shame, ethical uses of historical knowledge, moral education, and personal cultivation. Because of the comprehensive and systematic character of his philosophical concerns, Xunzi is sometimes compared to Aristotle.

Whereas both Mencius and Xunzi are exponents and defenders of Confucius's ideal of well-ordered society, traditional Chinese scholars often distinguish their thought by the contrast between government by *ren* or benevolence and government by *li* (rites, rules of proper conduct). However, for both, the key concepts are *ren*, *yi* (righteousness, rightness, fittingness), and *li*. Xunzi writes:

> The *dao* (Way) of former kings consists of exaltation of *ren* and acting in accord with the Mean. What is meant by the Mean? I answer that: "*li* and *yi*." *Dao* is not the *dao* of Heaven, nor is it the *dao* of the Earth. It is the *dao* that guides humanity, the *dao* embodied in the lives of the paradigmatic individuals. (*ruxiao pian*, ch. 8)

Unlike Mencius, Xunzi was a forceful advocate of abolition of hereditary titles. Even more important, an enlightened ruler will enrich the state and its people with ample surplus to cope with untoward circumstances, protect the country with strong military defense measures in the spirit of *ren*, and promulgate and efficiently administer ethically legitimate laws and institutions. Thus an enlightened ruler is one who is good at organizing the people in society in accordance with the requirements expressed in *ren*, *yi*, and *li*. Some key aspects of Xunzi's philosophy are highlighted below.

Xunzi is best known for his thesis that human nature (or *xing*) is bad (*e*), and that any goodness man experiences is a direct result of activity that is constructive and productive (*wei*). Xunzi appeals to presumably established linguistic usages of *shan* and *e*: "All men in the world, past and present, agree in defining *shan* [goodness] as that which is upright, reasonable, and orderly, and *e* [badness] as that which is prejudiced, irresponsible, and chaotic." (*xing 'e pian*, ch. 23). Xunzi continued: "Now suppose man's nature was in fact intrinsically upright, reasonable, and orderly—then what need would there be for sage kings and *li* [rules of proper conduct] and *yi* [righteousness]?" (*ruxiao pian*, ch. 8).

In light of Xunzi's definitions of *shan* and *e*, it seems clear that these are evaluative terms based on his normative conception of moral and political order. The original human nature (*xing*) is normatively neutral. It consists of feelings (*qing*) such as "love, hate, joy, anger, sorrow, and pleasure," and desires (*yu*), which are responses to the arousal of feelings. What makes these feelings and desires problematic is that in the absence of the guidance of *li* and *yi*, humans tend to pursue their satisfaction without regard to other persons' needs and desires. And given

human partiality and scarcity of resources, conflict is inevitable.

Li and *yi* are the products of the constructive activity (*wei*) of the sages. Emphasis on *li* (ritual, rites, rules of proper conduct) is the hallmark of Xunzi's ethics. The *li* are formal prescriptions or rules of proper conduct. Although the *li* represent an established ethical tradition, they do not always provide adequate guidance in dealing with changing circumstances of human life. As markers of Dao (the Way), "the *li* provide models, but no explanations"; their primary function is regulation of conduct—defining the boundaries for the pursuit of desires. Notably, Xunzi also stresses the supportive and ennobling functions. Ultimately, the *li* promote the ennoblement of human characters by investing them with qualities of *ren* (benevolence) and *yi*. For Xunzi, the ultimate end of learning is to become a sage that embodies Dao—that is, *ren*, *yi*, and *li*. Ordinary humans are capable of becoming sages if they make efforts to understand the rationales and practice of these virtues.

Xunzi elaborates a complex theory concerning the capacity for knowing Dao and the significance of ethical commitment to the practice of Dao. Knowing Dao is the precondition to approving the Dao as the guide of human life. Xunzi is insistent that Dao is a whole consisting of many corners (*yu*) or aspects. All humans are liable to *bi* (obscuration, blindness), the beclouding of mind that leads to construing one aspect and ignoring an equally important aspect. Philosophers are especially prone to be victims of *bi*. For example, Mozi was beset by preoccupation with utility and failed to understand the importance of the beauty of form; Zhuangzi was beset by preoccupation with heaven and failed to understand the importance of humanity. Xunzi admits that Dao is a proper subject of discourse, but contentious reasoning must be avoided. The participants in reasoned discourse must be benevolent (*renxin*) and impartial (*gongxin*), and have a learning or receptive attitude (*xuexin*) toward competing views. For the telos in argumentation is to resolve problems of common concern in the light of Dao, not to win in disputation.

See also Confucius; Mencius.

Bibliography

Cua, Antonio S. *Ethical Argumentation: A Study in Hsun Tzu's Moral Epistemology*. Honolulu: University of Hawaii Press, 1985.

Cua, Antonio S. "Ethical Significance of Shame: Insights of Aristotle and Xunzi." *Philosophy East and West* 53 (2) (2003): 147–202.

Cua, Antonio S. "The Ethical and the Religious Dimensions *Li*." *Review of Metaphysics* 55 (3) (2002): 501–549.

Cua, Antonio S. *Human Nature, Ritual, and History: Studies in Xunzi and Chinese Philosophy*. Washington: The Catholic University of America Press, 2005.

Cua, Antonio S. "The Possibility of Ethical Knowledge: Reflections on a Theme in the *Hsün Tzu*." In *Epistemological Issues in Ancient Chinese Philosophy*, edited by Hans Lenk and Gregor Paul. Albany: State University of New York Press, 1993.

Dubs, H. H. *Hsuntze: The Moulder of Ancient Confucianism*. London: Arthur Probsthain, 1927.

Goldin, Paul R. *Rituals of the Way: The Philosophy of Xunzi*. La Salle, IL: Open Court, 1996.

Kline, T. C., and Philip J. Ivanhoe, eds. *Nature, and Moral Agency in the* Xunzi. Indianapolis: Hackett, 2000. Updated bibliography.

Knoblock, John. *Xunzi: A Translation and Study of the Complete Works*. 3 vols. Stanford, CA: Stanford University Press, 1988, 1990, and 1994. Vols. 1 and 3 contain excellent bibliographies.

Lee, Janghee. *Xunzi and Early Chinese Naturalism*. Albany: State University of New York Press, 2005.

Li, Disheng. *Xunzi jishi (Xunzi: An Exegesis)*. Taipei: Xuesheng, 1979.

Machle, Edward J. *Nature and Heaven in the Xunzi*. Albany: State University of New York Press, 1993.

Vittinghoff, Helmolt. "Recent Bibliography in Classical Chinese Philosophy." *Journal of Chinese Philosophy* 28 (1, 2) (2001): 1–208.

Y

YAMAGA SOKŌ

(1622–1685)

Yamaga Sokō was a Japanese Confucianist of the *kogakuha,* or "school of ancient learning," and codifier of the ethics of the military class, *Bushidō,* the "way of the warrior." He was born in Aizu, Fukushima prefecture. At nine he entered the school of Hayashi Razan in Edo (Tokyo), where he learned the official Zhu Xi doctrine. Interested in military science, he became a master of it. He taught it first at the castle of Lord Asano of Akō (Hyogo prefecture) and later in Edo, where the novelty of his advocating the use of firearms attracted many followers. In 1666 he wrote *Seikyō yōroku* (The essence of Confucianism), a blunt critique of Zhu Xi's ideas. For this and for his innovations in military science, he incurred the wrath of his two former teachers, Hayashi and the military expert Hōjō Ujinaga, and was exiled from Edo. For the rest of his life he lived under mild confinement at the castle of Lord Asano, instilling into the samurai of Akō the loyalty that was to make forty-seven of them famous for revenging their lord by slaying the man who had disgraced him and dutifully committing hara-kiri. Their deed and death was immortalized in the drama *Chūshingura.*

In the preface to *Seikyō yōroku,* Yamaga clearly states the program of the "school of ancient learning," adding that the doctrine of Confucius and the ancient sages had been obscured by interpreters and commentators. He dismisses Mencius, Zhu Xi, and Wang Yangming easily; he rejects the "great ultimate" (*taikyoku*) of Zhu Xi as a later Buddhist interpolation in Confucianism. The universe, he holds, is explained by the movement of yin and yang, the passive and active elements, and it has no beginning or end. Human nature is neither good nor bad, but ethically neutral. He stresses self-interest, but he urges that common utility take precedence over it.

The term *Bushidō* is a recent one, coined long after his death, but its meaning is clearly traceable to two of his books, *Shidō* and *Bukyō shōgaku.* His "way of the warrior" consists of ethical norms and practical means of fostering in oneself a sense of loyal duty (*gi*) toward one's lord. Mental training is paramount; serenity, sincerity, magnanimity, introspection, and self-restraint are the virtues to be cultivated. Yamaga praised the ancient Chinese sages but he was a strong nationalist who extolled Japan over China.

See also Chinese Philosophy; Confucius; Hayashi Razan; Human Nature; Japanese Philosophy; Mencius; Wang Yang-ming; Zhu Xi (Chu Hsi).

Bibliography

For Yamaga's works see *Yamaga Sokō zenshū* (Complete works of Yamaga Sokō), edited by Hirose Yutaka, 15 vols. (Tokyo: Iwanami Shoten, 1940). For discussion, see Hori Isao, *Yamaga Sokō* (Tokyo, 1963) and W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 394–410, which contains selections in translation.

Gino K. Piovesana, S.J. (1967)

YAMAZAKI ANSAI

(1618–1682)

Yamazaki Ansai, the Japanese Confucianist notable for his ethical bent and Confucian rationalization of Shintoism, was raised at Kyoto in a Buddhist monastery. He was so unruly that he was sent to Tosa (now the city of Kōchi) on Shikoku Island, where he came under the influence of Tani Jichu (1598–1649), the originator of the southern branch of the Zhu Xi school of Confucianism in Japan. Having discarded Buddhism, Yamazaki taught Zhu Xi Confucianism in Kyoto and Edo (Tokyo) from 1648. Uncompromising in character, he condescended in 1665 to become the official scholar of Hoshina Masayuki, lord of Aizu (in northeast Japan). At Hoshina's death in 1672 Yamazaki returned to Kyoto and developed his Confucian Shintoism.

Though a stern Confucianist teacher he gathered around him more than six thousand students; among the best were Asami Keisai (1652–1711), Satō Naokata (1650–1719), and Miyake Shōsai (1662–1741). They formed the Kimon or Ansai school. However, Yamazaki's Shintoism held the seed of disharmony; before his death this school split into four. He urged the ethical formula *keinai gigai,* that is, "Devotion within, righteousness without." By "devotion" he meant not simply Confucian self-cultivation but rather a religiously rectified mind related to cosmic reason. By "righteousness" he meant virtue toward others. His maxim, "Learning is knowing and practice," suggests a middle way between overemphasis on mastery of the mind and overemphasis on social virtues.

Yamazaki's Shintoism deserves attention because of its Confucian rationalism and the influence it had in the revival of Shintoist studies in Japan. It is called *Suika Shintō* and elaborates on Confucian cosmogony to explain Japan's mythological creation chronicles. Trying to see a rational core in these legends, he developed the Shinto creed, borrowing from neo-Confucianism. His best pupils, however, did not follow him in his Shintoist phase; and the *kokugakusha,* the "national learning scholars," did not become the purveyors of a rationalized Shintoism. His most lasting impact was made through his popularization of Confucian ethics and indirect fostering of loyalism toward the emperor. This last trend was exemplified in Asami Keisai, Yamagata Daini, and in the school of Mito historians. Yamazaki is, however, given credit for later loyalist and nationalist trends.

See also Buddhism; Chinese Philosophy; Confucius; Japanese Philosophy; Loyalty; Nationalism; Rationalism; Virtue and Vice; Zhu Xi (Chu Hsi).

Bibliography

For Japanese sources, see *Yamazaki Ansai zenshu* (Yamazaki Ansai: complete works), 5 vols. (Nagoya: Hatsubaijo Matsumoto Shoten, 1937), and Bitō Masahide, *Nihon hōken shisōshi kenkyu* (Studies on the history of feudal thought in Japan; Tokyo, 1961), pp. 40–99. An English source is W. T. de Bary, Ryusaku Tsunoda, and Donald Keene, eds., *Sources of Japanese Tradition* (New York: Columbia University Press, 1958), pp. 363–371.

Gino K. Piovesana, S.J. (1967)

YANG XIONG

(53 BCE–18 CE)

Having achieved his youthful ambition to become court poet, Yang Xiong spent his thirties and forties producing the occasional *fu* (rhapsodic poems) the throne required. Sometime around his fiftieth year, perhaps in reaction to the factionalized politics at the capital, Yang came to disparage his own poetic genius, equating the verbal pyrotechnics with childish games injurious to the moral process. In consequence, Yang turned to composing and then defending three works, the *Taixuan jing* (*Canon of Supreme Mystery*; c. 4 CE), the *Fayan* (*Model Sayings*; c. 12 CE), and the *Fangyan* (*Dialect Words*; unfinished?). Creating these new "classics" (*jing*) required greater ingenuity on Yang's part than writing *fu*, for Yang sought to capture both the inner message and the outer form of the canonical works: The *Mystery* was patterned after the *Yijing* (*Classic of Changes*); the *Model Sayings*, after the *Lunyu* (*Analects*); and the *Fangyan* claimed inspiration from the ancient Chou transcriptions of the *Odes* and possibly also the *Erya*, an early word list ascribed to Confucius. By such bold attempts at "renewing the old," Yang would restore the authentic teachings of the sages.

In imitation of the *Yijing*, an abstruse divination text turned philosophical work by the addition of "Ten

Wings," the *Taixuan jing* unfolds on two levels: For the ordinary reader, its divinatory formulae prescribe the virtues of humility, respect, and cautiousness that make for social order and personal safety. More sophisticated readers correlate a series of vignettes drawn from daily life and keyed to the calendar with graphic emblems, cryptic summaries, and Yang's own auto-commentaries to discover the complex relations binding human conduct and preordained fate. In Yang's view, four main factors determine the quality of life: Time, Tools, Position, and Virtue. Although the workings of fate (*ming*)—equated in Yang's work with Time—lie outside human control, time's depredations may be offset to some extent by other factors under better human control. Using the most advanced scientific theories of his time, Yang sketches the finely tuned cycles of yin/yang, and the Five Phases, relating them to decision-making and the hierarchical orders of civilization. In outlining these regularities, Yang touches upon the main topics of Han debate, including the existence of ghosts and providence, the role of divination and the divine, the origins and stages of the universe, and definitions of "good rule."

If the single most important theme of Yang Xiong's *Mystery* is the interaction between human will and divine fate, the *Fayan* sees single-minded devotion to the Good leading to an exquisite appreciation of the social and cosmic orders which itself constitutes the highest happiness of which humans are capable. In its brief dialogues, the *Fayan* constructs a compelling argument in favor of this inherently unprovable assertion by juxtaposing hypothetical cases with the examples of famous men and women, so as to assert three linked propositions: First, a crucial distinction exists between popular "heroes" and current officeholders and the "true" Ru faithful to Confucian ideals who neither pursue material success nor confuse the subtle Way with factual knowledge or rule-making. Second, the very process of learning to intuit the sages' intent so hones the learner's being that it gradually experiences the most exquisite pleasure known to humankind, a kind of moral connoisseurship called "the ultimate in discrimination" (*zhishi*) (chap. 6). Third, this therapeutic and pleasurable journey toward Goodness is the only sure reward for an expenditure of effort, as the pursuit of Goodness is "easy": it entails no trickery or treachery; it imparts mental equilibrium along with an ability to understand and predict human behavior (chaps. 2, 9); and it reveals an entire world marvelously balanced.

Given the broad strokes of the *Mystery* and the sweeping claims of the *Fayan*, some find it hard to place the *Fangyan*, a meticulous record of dialect expressions within the extended Chinese cultural sphere. The melodic patterns of human speech—as well as musical rhythms, the calligraphic forms of written characters, and the geographical configurations of the earth—intimate the divine order. Word patterns in particular fascinate Yang, for "words are the music of the heart-mind (*xin*); and writings, its painting" (*Fayan*, chap. 5). Yang's highest goal was to employ artistic forms to excite the sensibilities so that they might become more receptive to the serious business of moral edification. Therefore, Yang was the first to develop theories of aesthetic concepts and the hermeneutic enterprise, then to demonstrate the emotive power of language through his own rhetorical masterpieces.

During his lifetime, devoted disciples regarded Yang as Master, though some contemporaries mistrusted Yang's incredible versatility. Following his death, Yang was elevated to the pantheon by many. Han Yu (768–842), for example, named Yang Xiong as the single master qualified to "transmit the [Confucian] Way" after Mencius; Sima Guang (1018–1086) went further, insisting that neither Mencius nor Xunzi could compare with Yang. However, some Song thinkers, especially Zhu Xi (1130–1200), condemned Yang for his eclecticism, his arrogance in daring to create classics, his willingness to serve two dynastic courts, and his outright rejection of the Mencian theory of human nature. Only with the Qing Evidential Research movement did interest in Yang's work revive.

See also Aesthetics, History of; Chinese Philosophy; Confucius; Han Yu; Hermeneutics; Mencius; Time; Xunzi; Zhu Xi (Chu Hsi).

Bibliography

WORKS BY YANG XIONG

Fang-yen. Annotated and translated by Paul L. M. Serruys as *The Chinese Dialects of Han Time According to Fang Yen*. Berkeley: University of California Press, 1959.

Fayan yishu. Compiled by Wang Rongbao. Beijing: Zhonghua, 1987.

Taixuan jiaoyi. Compiled by Zheng Wangeng. Beijing: Shifan daxue, 1989. The best edition of the *Taixuan jing*.

Fayan zhu. Annotated by Han Jing. Beijing: Zhonghua, 1992.

The Canon of Supreme Mystery. Translated with commentary by Michael Nylan. Albany: State University of New York Press, 1993.

Fayan. Translated by Michael Nylan. Seattle: University of Washington Press, forthcoming.

WORKS ABOUT YANG XIONG

Knechtges, David R. *The Han Rhapsody: A Study of the fu of Yang Hsiung, 53 B.C.–A.D. 18*. Cambridge, U.K.: Cambridge University Press, 1976.

Pan Ku. *The Han shu Biography of Yang Hsiung (53 B.C.–A.D. 18)*. Translated and annotated by David R. Knechtges. Tempe: Arizona State University Press, 1982.

Michael Nylan (2005)

YANG ZHU

(c. 440–c. 380 BCE)

Not much has been discovered about Yang Zhu the person from the documents that still exist. However, the *Mencius*, the *Xunzi*, the *Hanfeizi*, the *Lushi Chunqiu*, the *Huainanzi*, and the *Lunheng* all confirm that Yang's school was one of the most influential in pre-Qin China. For Mencius, Yang and Mo Di were the most influential thinkers prior to Mencius's time, although he criticized Yang's emphasis on the individual and its anarchist consequence, as well as his selfishness and apathy to the public interest. These criticisms, however, are somewhat misleading for an understanding of the true nature of Yang's thought.

In the past, Chinese intellectuals were led to believe that "Yang Zhu chooses to exist only for his own self, and does nothing for the world, not even by drawing one hair of his" (*Mencius* 3B 9). Yet an unbiased understanding, based on existing texts, reveals that Yang cherished the value of life and the authenticity of self. For example, the Hanfeizi said that Yang was one who "despised things and values life" (*Hanfeizi Jijie*, p. 353). In the *Lushi Chunqiu*, it was said that "Scholar Yang elevates the self" (*Lüshi Chunqiu Jishi*, p. 803). And, according to the *Huainanzi*, "To keep the totality of one's natural life and conserve the authenticy of one's self, not to burden one's body with external things. This is that upon which Yangzi stands, yet it is criticized by Mencius" (Liu An 1985, p. 218).

These comments allow us to reread more coherently the *Yang Zhu* and other chapters of the *Liezi*, where many texts related to Yang were presented (even if these works are seen by many scholars as having been forged by later hands). In the *Liezi*, when Yang is asked by Qinzi whether he would agree to lose one hair to help out the whole world, he answers that the "human world is for certain not to be helped out by one hair" (*Liezi*, p. 218). Yang's emphasis is on "keeping the totality of one's natural life" and "conserving the authenticity of one's self," statements that can be understood in reference to his philosophy of body, in which he claims that the appropriate satisfaction of human desires and the economy of energy are essential in attaining the wholeness of one's own life.

Yang's emphasis on the authenticity of self is more understandable to twenty-first century readers: He is more like modern thinkers in that he underlines the autonomy of self in respect to all external determinations. *Autonomy* in this sense means the spontaneous unfolding of one's own nature—a nature not to be determined by external entities, either real or ideal, but to be determined internally by one's own self, which is different from Kant's idea of autonomy as positing norms by one's own free will. With his idea of autonomy, Yang made the distinction between "fled-away-persons" (*dunren*) and "conforming people" (*shunmin*). The *dunren* were escapists from their own natural self in living at the mercy of external factors. By contrast, the *shunmin* were those who did not run after external values and were free with the authenticity of their life, closely related to the self's autonomy.

Yang's philosophical anthropology is somewhat similar to St. Augustine's philosophy in *City of God* and Arnold Gehlen's in *Man, His Nature and Place in the World*. Yang believed that human intellect developed out of biological weakness, and from these weak biological conditions a person "should use things to nourish his own nature, let his own intellect develop without appealing to physical force" (*Liezi*, p. 224). The reason to use human intellect was for the purpose of conserving one's life by using natural resources without the necessity of appealing to physical force when competing with stronger animals. Based on this, Yang developed a philosophy of learning. Beginning with the tenet that life is a basic value, and avoiding losing oneself by embarking upon too many different courses of learning, Yang posited the authenticity of life as the final unity of all learning. The *Shuofu* chapter of the *Liezi* states that, "Because of too many deviations in roads, one can not find one's lost sheep; with too many deviations in learning, the learner would lose his own life" (*Liezi*, p. 254). Yang's pragmatist vision of learning meant to learn for the purpose of conserving life and its development according to self-authenticity, which for him were the ultimate values of human existence.

See also Augustine, St.; Chinese Philosophy; Confucius; Determinism and Freedom; Gehlen, Arnold; Han Fei; Kant, Immanuel; Mencius; Mozi; Xunzi.

Bibliography

Graham, Angus C., trans. *The Book of Lieh-tzǔ*. New York: Columbia University Press, 1990.

Hsü Wei-yü. *Lüshi Chunqiu Jishi* [Collected commentaries on spring and autumn of Mr. Lü]. Vol. 2. Taipei: Shijie Bookstore, 1988.

Liezi. In *Er Shi Er Zi* [Works of twenty-two masters]. Vol. 2. Taipei: Prophet Press, 1976.

Liu An. *Huainanzi* [The book of masters of Huainan]. Annotated by Gao Yu. Taipei: Shijie Bookstore, 1985.
Mencius. Translated by D. C. Lau. Rev. ed. Hong Kong: Chinese University Press, 2003.
Wang Xianshen. *Hanfeizi Jijie* [Collected explanations of Hanfeizi]. Taipei: Shijie Bookstore, 1988.

Vincent Shen (2005)

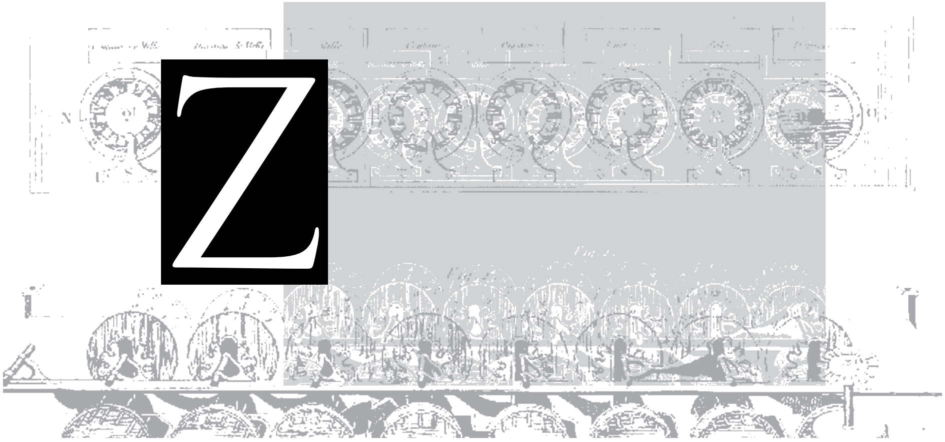

ZABARELLA, JACOPO

(1532–1589)

Jacopo Zabarella was one of the leading Aristotelians of the sixteenth century. He taught at the University of Padua for twenty-five years, from 1564 until his death. The fruit of these years of lecturing is contained in his printed works, which include treatises on Aristotelian logic and natural science. His writings in logic, and especially on scientific method, earned Zabarella a reputation as the most outstanding logician of his time; they continued to be read by school philosophers in Germany and Italy for several generations after his death and still command respect as interpretations of Aristotle.

Zabarella proceeds in characteristic scholastic fashion, examining and resolving, independently of each other, a sequence of issues. In the process he canvasses the views of an impressive number of predecessors among the Latins and seems fully conversant with Greek philosophy, including the Greek commentators on Aristotle. The doctrines discussed by Zabarella range, as is usual with scholastic writers, over an immense amount of material, basically that presented by Aristotle in his *Organon* and in the *Libri Naturales*. As a philosopher Zabarella is willing to leave certain arguments to the theologians—for example, whether God could have created prime matter without form. "My advice is to dispute in Aristotelian, not theological, fashion," he remarks. This does not mean, however, that Zabarella was not willing to consider and even to endorse arguments of a strictly philosophical nature presented by theologians; hence, the names of Thomas Aquinas, John Duns Scotus, Gregory of Rimini, and many others frequently occur in his works, along with the appeals to Averroes so frequent among Italian philosophers of his time. Analysis of the arguments advanced by predecessors constitutes one part of Zabarella's presentation (*ratio*); he also appeals to experience (*experientia*), his own or that of most people. Thus, he mentions having climbed the highest hill in the vicinity of Padua, seeing clouds below, and learning when he descended in the evening that it had rained in the valley during the day. But there is no reference to controlled experiment in his writings; in this respect he remained a bookish philosopher, like most university professors of his time.

No one has followed Zabarella carefully through the maze of his discussions in order to secure a clear view of his total thought. The studies we have are partial and will doubtless require revision in the light of increased knowl-

edge of the whole tradition he represents. Nevertheless, some of his conclusions can be definitely stated.

Zabarella regards Aristotle's science as perfect with respect to structure and form, imperfect only with regard to its subject matter. He compares Aristotle's writings on natural science with Euclid's *Elements* and suggests that the philosopher of nature can easily derive theorems of physics from the principles contained in them. Zabarella does not envisage the possibility that Aristotle's approach might be supplemented by mathematics. The fourteenth-century attempts at quantification in physics originating at Paris and Oxford had been transported to Italy by such teachers as Paul of Venice, but Zabarella does not seem aware of these developments. He did not welcome novel hypotheses, preferring, for example, to stand by Aristotle's explanation that the movement of projectiles can be attributed to pushing by the surrounding air (*antiperistasis*). Zabarella rejects the view that the "preceding motion is the cause of the greater velocity of the following motion."

In his discussions of the heavens, Zabarella betrays no concern with the Copernican theory published during his youth. He seems slightly dubious about the epicycles of the astronomers, but in this he was no doubt simply reflecting the doubts of Averroes. Zabarella endorses the view, also derived from Averroes, that the "confused" knowledge of the world supplied by the natural scientist must be made "distinct" by the metaphysician. For example, he concedes that the argument, "Since there is eternal movement, there must be an eternal mover," may be established by the natural scientist, whose bailiwick is the consideration and causal explanation of things in motion. But consideration of immaterial substances in themselves (the "eternal motors") must be left to the metaphysician.

Contemporaries had raised a difficulty in connection with certain mutually canceling actions in nature ("reactions"), which seemed to them to defy the Aristotelian dictum "Nature never does anything in vain." Zabarella points out that such mutual frustration nevertheless does not frustrate nature in general, since all things turn out according to the law of universal nature (*ex lege naturae universalis*).

Another question much discussed in scholastic physics concerned the elements in what we would call chemical compounds (called "mixtures" by the Schoolmen). Do they persist in existence after losing their sensible identity as elements and becoming part of the compound? Various solutions had been proposed to this problem; Zabarella accepts that of Averroes—the same "reality" of the elementary forms of matter is in the elements and in the mixture, but their "formality" is changed.

In Aristotelian metaphysics and philosophy the distinction between matter and form is crucial and difficult, especially in its application to human beings. School philosophers of Zabarella's time exercised a great deal of ingenuity in order to make sense of the Aristotelian doctrine that the soul is the form of the body. There were two main opinions: one, that the soul is a "form giving being" to man; the other, that the soul is merely a "form assisting" in man's operation, much as a sailor presides over the operation of an already formed ship. Zabarella chooses the former interpretation, although not without vacillation.

On another much disputed question, concerning the perception of sense qualities, Zabarella endorses the view of Albert the Great that there is no need to postulate an "active sense" (*sensus agens*); certain sensed qualities have it in themselves to multiply their "spiritual" species in the medium, in contrast to such other qualities as heat, which really produce their counterparts in the medium and in the sense of touch.

Zabarella decisively rejects the Averroist thesis of the unity of the intellect, insisting that the intellect is multiplied according to the number of individual men. The intellect is the form of man; since it is not itself "in act," it is able to receive all things spiritually and hence is capable of knowing all things.

LOGIC

Zabarella's most original contributions lie in his logical works. The nature of logic and its relation to other disciplines were controversial matters even in antiquity, and these controversies were renewed during the Renaissance. Zabarella sides with the Greek commentators on Aristotle in maintaining that logic is not strictly a part of philosophy but an instrumental discipline furnishing other arts and sciences with tools of inquiry. Two of these tools are order and method. Order is an intellectual habit that teaches us how to dispose suitably the parts of any given discipline so that we can learn it more easily. Method is also an intellectual instrument producing knowledge of the unknown from that which is known, but it permits us to draw syllogistic inferences. The nature of both order and method must be clarified by an analysis of their objectives: ease of learning in the case of order, perfect knowledge (*cognitio*) in the case of method.

These analyses are set forth in Zabarella's treatise *De Methodis* (On methods), in which he challenges two schools of thought prevalent in his time. One, drawn from Neoplatonic commentators on Aristotle, held that there are four methods employed in the arts and sciences: demonstrative, definitive, divisive, and resolutive. The other, advocated by medical men and drawn from Galen, held that there are three orders of teaching any discipline. Zabarella presents a simplified version, reducing the number of orders and methods to two. Contemplative disciplines are transmitted by the compositive order, practical or operative disciplines by the resolutive, which begins with the end to be achieved in any pursuit and reasons backward to an initial step in its direction.

This was traditional Aristotelian doctrine, but Zabarella's elaboration of compositive and resolutive methods was more original. In the natural sciences there are two things to be studied, substances and accidents. Substances can be investigated only by the resolutive method, which begins with sensible effects and "resolves" them into their causes. We know substances when we possess definitions of them, but these definitions, contrary to received opinion, are not "methods." Accidents, on the other hand, can be demonstrated by the demonstrative or compositive method once the principles discovered by the resolutive method are available.

In his work "On the Regress," Zabarella analyzes a special form of demonstration in which "the cause and the effect reciprocate, and the effect is more known to us than the cause." The best example of such a regress is to be found, Zabarella tells us, in Aristotle's *Physics*. We know in a confused way that where there is generation, there is matter, but only demonstration makes it clear to us why matter is the cause of generation. We must make use of a "mental examination," which tells us that matter is "that which is apt to receive all forms and privations."

Zabarella reaffirms man's central place in the universe; the operation of the most outstanding part of man is his highest perfection, and this is to be found in contemplation. Man is of a middle nature; he is the most noble animal, created in the image of God, but there is also a sense in which he is ignoble and imperfect, the sense in which we say, "To sin is human" or "After all, he is only a man." Such concern for placing man in nature probably echoes fifteenth-century humanism.

See also Albert the Great; Aristotelianism; Aristotle; Averroes; Duns Scotus, John; Galen; Gregory of Rimini; Humanism; Logic, History of; Paul of Venice; Scientific Method; Thomas Aquinas, St.

Bibliography

None of Zabarella's works has been translated into English, and this is unfortunate, since he ranks high as an expositor of Aristotle. Furthermore, copies of his *Opera Logica* (published first in Venice, 1578, but many times thereafter) are hard to obtain. The same may be said of his *De Rebus Naturalibus* (Venice, 1590) and his commentaries on the *Physics* and *De Anima*. A modern edition of the *De Methodis* and other logical works would be welcome and would furnish us with one of the most sophisticated expositions of school logic and thinking concerning scientific method to be given during the Renaissance.

For studies on Zabarella, see Ernst Cassirer, *Das Erkenntnis problem in der Philosophie und Wissenschaft der neueren Zeit* (Berlin: Cassirer, 1906), Vol. I, pp. 134–141; John Herman Randall Jr., *The School of Padua and the Emergence of Modern Science* (Padua, 1961), pp. 49–63 (gives ample quotations in Latin); N. W. Gilbert, *Renaissance Concepts of Method* (New York: Columbia University Press, 1960), Ch. 7; J. J. Glanville, "Zabarella and Poinsot on the Object and Nature of Logic," in *Readings in Logic*, edited by R. Houde (Dubuque, IA: Brown, 1958); Riccardo Pozzo, ed., *The Impact of Aristotelianism on Modern Philosophy (Studies in Philosophy and the History of Philosophy, Vol. 39)* (Washington DC: Catholic University of America Press, 2004); and William A. Wallace, "Circularity and the Paduan 'Regressus': From Pietro D'abano to Galileo Galilei." *Vivarium* 33(1)(1995): 76–97.

Neal W. Gilbert (1967)
Bibliography updated by Tamra Frei (2005)

ZARATHUSTRA

See *Zoroastrianism*

ZEN

See *Buddhism—Schools: Chan and Zen*

ZEN'KOVSKII, VASILII VASIL'EVICH

(1881–1962)

Vasilii Vasil'evich Zen'kovskii, a Russian philosopher and theologian, was born in Proskurov into the family of a teacher. Zen'kovskii studied natural sciences, history, and philology at Kiev University. In 1913–1914 he continued his education in Germany, Austria, and Italy. Following his return to Russia he was appointed a professor of psychology at Kiev University (1915–1919). In 1919 he immigrated to Yugoslavia, where he worked as a professor at the University of Belgrade (1920–1923). In 1923 he

moved to Czechoslovakia, where he became the director of the Academy of Education in Prague (1923–1926). In 1926 he settled in France, where he was a professor of the Theological Academy in Paris until his death. In 1944 he was elected as dean of the academy. Like many Russian intellectuals of the time, Zen'kovskii went through a spiritual crisis in his youth. He became an atheist when he was fifteen years old, but later returned to the church and dedicated all of his life to developing and promoting Christian philosophy and education. In 1942 he was ordained to Orthodox Christian priesthood.

PHILOSOPHY

Zen'kovskii belongs to a pleiad of prominent Russian thinkers who carried on Russia's intellectual tradition after the 1917 Communist Revolution and continued it outside the homeland despite the hardships of emigration. In the history of Russian thought Zen'kovskii is best known for his two-volume classic *Istoriia russkoi filosofii* (History of Russian philosophy; 1948–1950), which still remains an unsurpassed contribution to the field. He also authored many works in philosophy, theology, psychology, pedagogy, and literary history that left a notable mark on Russian culture. Overall, his philosophical system may be described as "Orthodox universalism" (Sapov 1995) or, in Zen'kovskii's own words, as an "experiment in Christian philosophy."

Zen'kovskii began his scholarly career with the study of psychic causality. He was interested in the phenomenon of religious consciousness, more particularly in the origin of the idea of God in the human mind. According to Zen'kovskii neither the social nor the subconscious sphere could produce in human consciousness such an idea that had its true roots in the mystical experience of the interconnection between the human being and the divine realm. He points out that some people apparently lack this inner vision, and as a result they advance theories that reduce religious experience to other forms of human activity, as was the case, for example, with Karl Marx, Émile Durkheim, or Sigmund Freud.

In his epistemological views Zen'kovskii rejects the autonomy and self-sufficiency of human reason. He develops a "Christocentric understanding of knowledge," which postulates that Christ as divine Logos (John 1:1) represents the ultimate generating and regulating power of human intellectual activities. More specifically, as Vadim Sapov notes, Zen'kovskii defends the "concept of 'ecclesial reason,' according to which one should search for the metaphysical basis of knowledge in the notion of the Church" (1995, p. 204) as the living body of Christ.

In his youth Zen'kovskii was to a considerable extent influenced by the nineteenth-century Russian philosophers Lev Mikhailovich Lopatin and Vladimir Sergeevich Solov'ëv (Solovyov), and his ontology also bears certain similarities to the Solov'ëvian tradition. Zen'kovskii combines here the elements of philosophy and theology by focusing on the concept of creation. He develops his own version of Sophiology that represents a variation of the Sophiological teachings of Solov'ëv and later of Sergei Nikolaevich Bulgakov and that centers around the notion of Sophia or God's Wisdom as the bridge between the creator and the creatures. In his Sophiological doctrine Zen'kovskii distinguishes between "ideas in God" and "ideas in the world" or between divine and created Sophia. Divine Sophia stands for God's plan of creation, while created Sophia represents the ideal foundation of the universe itself. Divine and created aspects of Sophia are connected with each other as the archetype and its image or Logos.

The concept of human personhood occupies the central place in Zen'kovskii's philosophical system. Every human being, in his view, is unique and experiences a different combination of genetic, social, and spiritual influences. Acts of freedom that are rooted in the metaphysical depth of one's self also constitute an inalienable part of the human person. Without divine grace such freedom, however, almost inevitably leads humanity to evil. The original sin that limits the creative potential of free will finds its manifestation in the "split between reason and heart." Hence, the purpose of human life consists in the restoration of lost spiritual wholeness through the church. Accordingly, the main task of any pedagogical efforts must be directed to helping the young generation in its efforts toward such a spiritual transformation.

THEOLOGICAL TEACHINGS

Zen'kovskii's theological teachings are collected in his *Apologetika* (Apologetics; 1957), which aims at defending Christian worldview against the challenges of modern culture and science. Here as elsewhere it is hard to dissociate Zen'kovskii's religious views from his philosophical argumentation. The work addresses a variety of issues from the dogmatic question of creation to the controversial problem of freedom. When facing the paradox of freedom versus evil, Zen'kovskii joins many other Russian thinkers, including Nikolay Aleksanrovich Berdyayev, in arguing that human freedom is totally unrestricted. In *Apologetics* he points out that "freedom is a true freedom only if it is unlimited—in it is God's likeness" (1997, p. 406). He adds, however, that, the "Lord can commit to

death, total destruction those individuals who resist a complete harmonization of being" (p. 229).

While Berdyayev in his philosophy questions divine omnipotence to proclaim the ultimate power of freedom, Zen'kovskii believes in the all-powerful God but seems to undermine God's all-goodness by forecasting a complete extermination of the wicked in the future. He refers to the authority of the Bible, according to which the "second death, i.e. annihilation awaits those who will not want to come back to God" (1997, p. 302). This interpretation reveals some of the aspects of Zen'kovskii's Orthodox Christian thought that today's readers may find rather conservative, if not fundamentalist.

See also Berdyaev, Nikolai Aleksandrovich; Bulgakov, Sergei Nikolaevich; Determinism and Freedom; Durkheim, Émile; Freedom; Freud, Sigmund; Lopatin, Lev Mikhailovich; Marx, Karl; Philosophy of Religion, History of; Russian Philosophy; Solov'ëv (Solovyov), Vladimir Sergeevich.

Bibliography

WORKS BY ZEN'KOVSKII

Problema psikhicheskoi prichinnosti (The problem of psychic causality). Kiev, Russia: 1914.

Russkie mysliteli i Evropa (Russian thinkers and Europe). Paris: YMCA Press, 1926.

Dar svobody (The gift of freedom). Paris: 1928.

Problema vospitaniia v svete khristianskoi antropologii (The problem of education in the light of Christian anthropology). Paris: YMCA Press, 1934.

Istoriia russkoi filosofii. 2 vols. Paris: YMCA Press, 1948–1950.

Osnovy khristianskoi filosofii (Principles of Christian philosophy). Moscow: Kanon, 1997.

WORKS ON ZEN'KOVSKII

Lossky, N. O. *History of Russian Philosophy*. New York: International Universities Press, 1951.

Sapov, Vadim. "Zen'kovskii, Vasilii Vasil'evich." In *Russkaia filosofiia. Malyi entsyklopedicheskii slovar'* (Russian philosophy: A small encyclopedic dictionary), edited by A. I. Aleshin, 202–205. Moscow: "Nauka," 1995.

Mikhail Sergeev (2005)

ZENO OF CITIUM

(334–262/1 BCE)

Zeno, creator of the philosophical system that became known as Stoicism, was born probably in 334 BCE in Citium, a coastal settlement in southeastern Cyprus, whih was largely Hellenized by that time. His family may well have been of Phoenician origin (as was a significant minority of the population). At the age of twenty-two, he left for Athens. There he spent the next decade or so studying philosophy with various teachers. In time a group formed round Zeno himself; and because these "Zenonians" met in a public colonnade named the Painted Stoa, they came to be called Stoics. Zeno evidently established a prominent position in Athenian society. In his later years Antigonus Gonatas, the Macedonian monarch, attempted without success to attract him to his court, while the Athenians themselves voted him public honors in both life and death, particularly because of the exemplary moral example he had set. "More self-controlled than Zeno" became the benchmark phrase. He died in 262/1 BCE.

Zeno's philosophical hero was Socrates. The Stoics, so Philodemus tells us, were prepared to be known as "Socratics"; and Stoicism is best understood as a theoretical articulation of Socrates' intellectualist ethics, buttressed by a monistic metaphysics that is at once materialist and pantheist. Zeno's early attraction to the Socrates portrayed in Xenophon's *Memorabilia* is attested to in an anecdote that associates it with the influence exercised over him by his first teacher, the Cynic philosopher Crates. He appears to have cultivated a Cynicizing image in his own lifestyle. Zeno was noted for frugality, stamina, unsociability—and a Laconic sharpness in repartee. His *Republic*, the first book he wrote, constituted a critique of Plato's great work so uncompromisingly Cynic that Stoics of Cicero's time tried either to disown or to bowdlerize it.

Here Zeno rejects the need for an elaborate educational system; he sweeps away institutions such as temples, law courts, gymnasia; he abolishes coinage. Women are to wear the same clothing as men. Any man may mate with any woman: Gone is all Plato's sexual regulation. Gone, too, is Plato's insistence on a rigidly stratified class structure. All that is required for true citizenship is virtue. Single-minded Cynic rejection of every conventional value is the short way to acquire that, and thus to help build a community of the virtuous in the here and now. But Zeno also invoked a more positive and distinctively Socratic idea in this context. Eros—the god of erotic love—was to be the deity presiding over Zeno's city, bringing it friendship, freedom, and concord. The wise and virtuous will, like Socrates, seek out young people whose physical attractions indicate a propensity to virtue. By such relationships the bonds of society are to be forged.

Like all Zeno's writings, the *Republic* is now lost. Quite a number of other book titles are preserved, indi-

cating a much wider range of philosophical preoccupations than are typical of the Cynics or of Socrates himself. Extended verbatim quotations are rare, but doctrines and especially definitions are cited in a variety of later classical authors. From these it is clear that the main structure of Stoic ethics was already articulated in Zeno's own pioneering work. Thus he endorses the Socratic idea that virtue is exclusively a matter of knowledge and wisdom, and that because it is, on its own, sufficient for happiness, the human goal consists in living in accordance with virtue. More innovative is Zeno's way of explaining what it means to be wise, and how in living wisely a person "follows nature." He took an expression in common moral discourse—*kathêkon*: what is incumbent upon me, my duty. Although (or perhaps in part because) it had never received any previous philosophical attention, he made it elemental within his own ethics. By a characteristic piece of etymologizing, *kathêkon* is explained as behavior that "comes in accordance with" the nature of a human being, or more generally an animal or plant of a particular kind. In a human it is what reason enjoins or forbids. Virtue or excellence in a person accordingly consists in "reason consistent and firm and unchangeable," and "living consistently" is by the same token the human goal: *eurhoia biou*, "success in life" (but etymologically its "life's smooth current").

Virtue is therefore not an ideal remote from everyday life but something focused on duties that are incumbent upon the ordinary person: honoring parents, serving country, spending time with friends, taking proper care of your health. An unqualified Cynic might have regarded most such things as indifferent to happiness. Zeno did not flatly disagree. But at this point he made another innovative move, decisive for the shape of Stoic ethics and for attacks upon it, ancient and modern. Some things indifferent for happiness (such as natural ability, beauty, health, wealth) are "preferred," like favorites at court, as according with nature; others (such as their opposites) not, as contrary to nature. Ordinarily reason will enjoin behavior designed to secure those that are preferred. But not always. Self-mutilation may be in order if the only alternative is military service with a tyrant in an unjust cause. What really matters for happiness is listening to right reason and acting accordingly, even if it is only the perfectly rational or wise person—the "sage"—who manages to do that with complete consistency. Consequently it is paradoxically the sage alone who is truly rich, strong, beautiful, and so on.

Knowledge, too, was, in Zeno's assessment, commonly accessible, not the preserve of philosophy or the sciences. Like the Socratic Stilpo, another of his teachers, Zeno rejected Platonic universals. As in ethics, so in epistemology he introduced fresh vocabulary to express the new idea he wanted to make fundamental: *katalêpsis*, "cognitive grasp." All of us—wise or wretched fools (which is what we are if we do not attain virtue and wisdom)—have a reliable basis for navigating the world we inhabit: sensory impressions conveying a grasp of reality that could not be wrong. Zeno used his hands to illustrate the point. An open palm represents what it is to receive an impression. Closing the fingers a little signifies assent. Clenching the fist is *katalêpsis*: Assent that is unquestionably right. The need for a concept of secure rational understanding—*epistêmê*—on the Platonic model is not denied. Zeno illustrated this by clasping his clenched right fist tightly and forcibly with his left hand. The point? Contra Plato, there can be no secure understanding without the kind of cognitive grasp of the sensible world that is made available to everyone by a providential Nature.

Belief in a providential nature—which the Stoics identified with God and Zeus and Fate—was something Zeno found Socrates arguing in Xenophon's *Memorabilia* (1.4, 4.3), most compellingly in the inference that, just as the physical stuff we are made of is supplied by the world about us, so our intelligence must derive from a cosmic intelligence. Zeno had studied logic with the dialectician Diodorus Cronus (author of the famous Master Argument) and formulated a pithy syllogism to express the point in causal and biological terms:

> What emits seed of something rational is itself rational.
> But the world emits seed of something rational.
> Therefore the world is rational.

Zeno seems to have considered the Socratic provenance of this line of reasoning particularly significant. But he was also anxious to claim the support of the entire philosophical tradition so far as he could. He exploited Plato's *Timaeus* (29B–30B) to argue:

> The rational is superior to the nonrational.
> But nothing is superior to the world.
> Therefore the world is rational.

The same was true, he argued, of "intelligent" and "ensouled."

From the Academic philosopher Polemo, yet another of his teachers, Zeno may have learned to find in the *Timaeus* something no less important: the duality of God and matter that he made fundamental to his own monistic metaphysics. But for theory about the cosmos, no pre-

vious philosopher was more important to him than Heraclitus. It must have been his reading of Heraclitus that convinced Zeno that nature was to be understood in terms of fire—its methodical crafting of the coming into being of things, its transformations, and the periodic cosmic holocausts it fuels. The richness of the Heraclitean resonances in Stoicism is now most apparent in the *Hymn to Zeus* of Zeno's pupil and successor as head of the school, Cleanthes.

See also Cicero, Marcus Tullius; Cleanthes; Cynics; Diodorus Cronus; Epistemology; Heraclitus of Ephesus; Logic, History of; Philodemus; Plato; Socrates; Stoicism; Xenophon.

Bibliography

BOOKS

Arnim, Hans von, ed. *Stoicorum Veterum Fragmenta. Vol. I: Zeno et Zenonis Discipuli*. Leipzig: Teubner, 1903.

Diogenes Laertius. *Lives of Eminent Philosophers*. 2 vols. Translated by R. D. Hicks. Cambridge, MA: Harvard University Press, 1925.

Pearson, A. C. *The Fragments of Zeno and Cleanthes*. Cambridge, U.K.: Cambridge University Press, 1891.

Scaltsas, Theodore, and Andrew S. Mason, eds. *The Philosophy of Zeno: Zeno of Citium and His Legacy*. Larnaka: The Municipality of Larnaka, 2002.

Schofield, Malcolm. *The Stoic Idea of the City*. 2nd edition. Chicago: The University of Chicago Press, 1999.

CHAPTER ARTICLES

Decleva Caizzi, F. "The Porch and the Garden: Early Hellenistic Images of the Philosophical Life." In *Images and Ideologies: Self-Definition in the Hellenistic World*, edited by Anthony W. Bulloch, Eric S. Gruen, Anthony A. Long, and Andrew Stewart. Berkeley: University of California Press, 1993.

Schofield, Malcolm. "Cicero, Zeno of Citium, and the Vocabulary of Philosophy." In *Le Style de la Pensée: Receuil de Textes en Hommage à Jacques Brunschwig*, edited by M. Canto-Sperber and P. Pellegrin. Paris: Les Belles Lettres, 2002.

Sedley, David N. "The Origins of Stoic God." In *Traditions of Theology: Studies in Hellenistic Theology*, edited by Dorothea Frede and André Laks. Leiden: Brill, 2002.

JOURNAL ARTICLES

Dorandi, Tiziano. "Filodemo, Gli Stoici (P.Herc. 155 e 339)." *Cronache Ercolanesi* 12 (1982): 91–133.

Long, Anthony A. "Socrates in Hellenistic Philosophy." *Classical Quarterly* 38 (1988): 150–71.

Mansfeld, Jaap. "Diogenes Laertius on Stoic Philosophy." *Elenchos* 7 (1986): 295–382.

Schofield, Malcolm. "The Syllogisms of Zeno of Citium." *Phronesis* 28 (1983): 31–58.

Malcolm Schofield (2005)

ZENO OF ELEA
c. 490–430 BCE

According to Plato (*Parmenides* (127A–C), Zeno was born around 490 BCE. He was a citizen of Elea, a Greek city in southern Italy with which Parmenides was also associated. Little is known about his life. The setting of Plato's *Parmenides* is a visit Zeno and Parmenides made to Athens in Socrates' youth (around 450 BCE), but since the conversation in that dialogue between Parmenides and Socrates certainly did not take place, there is no strong reason to believe that the visit did either. According to tradition, Zeno died heroically defying a tyrant in Elea. Philosophically he was a follower of Parmenides, whose doctrines he defended by arguing against opposing views; hence Aristotle called him the father of dialectic. Although Zeno wrote a book containing forty arguments against plurality. very little of his writing remains; approximately twenty lines of quotations, supplemented by relatively scanty testimonia. We have information about a dozen of his arguments. Under these circumstances, Zeno's immense influence on the history of philosophy is all the more remarkable.

Plato, our earliest witness, depicts Zeno as defending Parmenides' views against people who ridiculed Parmenides on the grounds that his views have absurd consequences. Zeno paid them back in their own coin, pursuing implications of the opposing views, which he showed have consequences even more absurd than those the opponents claimed to follow for Parmenides (*Parmenides* 128C–D). Zeno's book comprised a series of polemical arguments that employed the strategy *reductio ad absurdum* against the claim that there exists more than one thing (ibid. 128B–D).

Although Plato's account fits some of Zeno's arguments, it does not hold for them all. Several argue that motion cannot exist, another that the senses fail to discern the truth, another that things do not have locations. And so it is unclear how reliable Plato (whose reports of some other early philosophers are unreliable) is as a source on Zeno. Some scholars deny that Parmenides was a monist at all, or in the relevant sense, and some have held that some of Zeno's arguments tell as strongly against Parmenides' monism as they do against his opponents' pluralism. If this is correct, then Plato's account of Zeno's arguments is wholly misguided. Others have also argued that Zeno is better defined as a proto-sophist, a paradox-monger who constructed ingenious arguments with perverse conclusions, without any philosophical commitments at all.

Despite these concerns, the text of this encyclopedia follows the traditional view that Parmenides believed that there exists only one entity, which is motionless and changeless (it has other attributes as well); that the human senses are entirely deceptive as a source of knowledge of reality; and that Zeno defended this theory through arguments that derive absurd consequences not only from the assumption that there exists a plurality of entities but also from the assumptions that motion and change exist, and other assumptions that humans make about the world. Plato's account is taken to be essentially correct; when it states that Zeno defended Parmenides' view that there is just one thing, it is quoting this core Parmenidean thesis as a shorthand method of referring to the entire theory.

Scholars have disputed the identity of Parmenides' opponents against whom Zeno directed his arguments. Some held that they were the Pythagoreans, but the case collapsed for lack of evidence. Others have suggested that the opponents were not actual objectors but any possible objectors, that Zeno constructed a series of arguments that systematically refuted all possible alternative theories—for example, the theory that motion is continuous and also the theory that motion is discrete—but that interpretation failed for the same reason. What remains is the most natural interpretation, that Parmenides' opponents were people (ordinary folk and philosophers as well) who found Parmenides' views obviously, radically, and amusingly wrong because they conflict so strongly with humankind's most deeply held beliefs about the world.

The Zenonian legacy is a number of arguments known as paradoxes because of their implausible conclusions. Many of them have the form of an antinomy, which is a special kind of *reductio* argument. Zeno proves a thesis by demonstrating that its contradictory has incompatible consequences. Since the consequences cannot both be true, the contradictory of the original thesis is false, so the thesis itself is true. As a matter of fact, Zeno's arguments do not contain the final move, which is characteristic of reductio arguments: they stop when they have shown that the contradictory of the thesis is false and do not draw the inference that the thesis itself is true. It has therefore been claimed that the arguments are not reductio arguments at all. But this criticism affects only the form, not the intent of the argument; they are reductio arguments in spirit if not in letter.

ARGUMENTS AGAINST PLURALITY

Several of Zeno's arguments against plurality survive. They include the argument of both like and unlike; the argument of both large and small; and the argument of both limited and unlimited.

ARGUMENT OF BOTH LIKE AND UNLIKE. The first argument against plurality is as follows: (a) If things are many, they must be both like and unlike; but (b) what is like cannot be unlike and what is unlike cannot be like; therefore (c) there cannot be many things (*Parmenides* 127D). The meaning of (a) is unclear: In what way are many things both like and unlike? One attempt to explicate it as follows. If there are many things, each of them is like itself in that everything that is true of it is true of it. This is trivially true. But if one thing (A) is counted as like another (B) only if everything that is true of A is true of B and/or vice versa, and if A is unlike B if and only if A is not like B (i.e., "like" and "unlike" are contradictories, as (b) indicates), then *any* two things are unlike one another.

For example, even if A and B are as alike as two peas in a pod, A will be unlike B because it is true of A that it is A but it is not true of B that it is A. Following this interpretation, which places a very strong condition on things being "like," (a) is true but (b) is false, so it follows that the alleged impossibility is not impossible at all. For impossibility to occur, the things would have to be both like and unlike the same thing (whereas here A is like one thing (A) and unlike something else (B)). Further, they would have to be both like and unlike the same thing in the same respect (since A can be like B in color but unlike B in weight) and at the same time (since A can be like B in color at one time but not another). The paradox fails on the interpretation given. It also fails if one admits a weaker condition for one thing being like another. For example, if one counts A as like B if at least one thing true of A is also true of B, so that A will be unlike B only if nothing true of A is true of B, the alleged impossibility again proves perfectly possible, since the only way something can be both like and unlike is by being like one thing and unlike something else. Other attempts to reconstruct the argument have been proposed, but none has yet succeeded in making it plausible, so it seems likely that Zeno's first argument is fallacious.

ARGUMENT OF BOTH LARGE AND SMALL. Two of Zeno's surviving five fragments contain parts of a different and more complex argument against plurality. The argument claims that if things are many, they are both large and small: (a) so large that they are infinite and (b)

so small that they have no size. The argument consists of two separate parts, one showing that things are large and one that they are small. It is an antinomy, but in this case Zeno argues that each branch of the antinomy is subject in its own right to a serious objection.

The entire argument for (a) has survived, but only part of the argument for (b). The proof of (b) came first, and the part that is reported is as follows: "Nothing has size because each of the many things is the same as itself and one." Zeno then argues that anything without size, thickness, or bulk does not exist: "If it is added to something else that exists, it will not make it any larger. For if it were of no size and were added, what it is added to cannot increase in size. It follows immediately that what is added is nothing. But if when it is subtracted the other thing is no smaller, and it is not increased when it is added, clearly the thing added or subtracted is nothing." (DK 29B2) This argument holds for three-dimensional bodies (though not for other kinds of things: I do not become larger by becoming happier, though one might say that happiness is added to me), so it is reasonable to take Zeno as arguing against the kind of pluralism that supposes that there exists a plurality of bodies (physical pluralism). What is missing is a reason to hold that "nothing has size because each of the many things is the same as itself and one."

The argument for (a) states: "If it exists, each thing must have some size and thickness, and a part of it must be apart from the rest. And the same reasoning holds for the part that is in front: that too will have size and part of it will be in front. Now to say this once is the same as to keep saying it forever. No such part of it will be last, nor will there be one part unrelated to another. Therefore, if there are many things they must be both small and large; so small as not to have size, but so large as to be unlimited." (DK 29B1) The first claim follows from (b). Zeno proceeds on the assumption that size implies divisibility: any body can be divided into spatially distinct parts, each of which is itself a body. This in turn entails divisibility without limit: the process of subdividing never reaches an end, so the parts are so large as to be unlimited.

Most scholars believe that the argument claims to prove that the size of the totality of the parts is infinitely large. If so, it is fallacious. All it proves is that number of the parts is infinitely large, and as the series $1/2 + 1/4 + 1/8 + \dots$ (whose sum is 1) shows, the sum of an infinite series need not be infinite. In the present case, the size of the totality of the parts remains equal to the size of the original whole. But if we adopt another interpretation the argument is valid. Since the argument focuses not on the size of the parts, but on the process and the products of division, the problem it raises concerns not the size of the totality of the parts but the possibility of completing the division. According to this interpretation, Zeno is demonstrating a difficulty in ordinary notions of physical bodies and spatial extension. People think that bodies are divisible and Zeno points out there is no reason to postulate that divisibility is impossible beyond some minimum size. It follows that bodies are infinitely divisible: even a small body is large enough to have an infinite number of parts. This conclusion is surprising enough to be worthy of Zeno.

The account of division just given suggests a way to supply the missing step in the argument for (b). How does the innocuous fact that something is the same as itself and one imply that it has no size? Perhaps it is because being "one" entails having no parts—otherwise it would be many. Since (as the process of division shows) anything with size can be divided into parts, only something without size will have no parts and so be "one." And then the argument for (b) comes into play.

ARGUMENT OF BOTH LIMITED AND UNLIMITED. The argument is: "(a) If there are many things, they must be just as many as they are, neither more nor less. But if they are just as many as they are, they must be limited. (b) If there are many things, the things that exist are unlimited, since between things that exist there are always others, and still others between those. Therefore the things that exist are unlimited." (DK 29B3) Branch (a) of this antinomy amounts to the claims that any plurality of things consists of a definite number of things and that any definite number is limited. The latter of these is equivalent to the claim that there is no such thing as a definite unlimited number. It has been objected that this last claim is false, since some infinite collections are pluralities that in a relevant way are definite and yet not "just as many as they are." But Zeno did not have the modern understanding of the infinite available to him, and the notion of "unlimited" with which he was working (in which the word means "inexhaustible" or "endless") makes it reasonable, even truistic, to say that an unlimited collection of things has no definite number. The former claim, that every plurality contains a definite number of things, as at least superficially plausible, which is enough to launch the paradox. Whether or not it is true will depend on how an individual counts the things in question (and perhaps their parts as well—see the paradox of both large and small), which Zeno does not specify.

Branch (b) can be interpreted in several ways, some of them anachronistic (for example, that the plurality in question is not three-dimensional objects but mathematical points on a line) and some open to obvious objections (for example, if Zeno is talking about three-dimensional objects that can touch one another, it is just false that there are always other objects in between). The source for this argument suggests a more interesting approach, saying: "In this way he proved the quantity unlimited on the basis of bisection." (Simplicius, *In Physica* 140, 33). Simplicius need not have quoted all the Zenonian text he had access to, and since he quoted part of the argument he could very well have known the rest of it. In the kind of division referred to, an object is first cut in half, then one of the halves is cut in half, and so on ad infinitum. If there are two adjacent objects A and B, this argument can be used to prove not that A and B have other objects in between them, but that there is no part of A nearest to B. If A is adjacent to B on the left, then the right half of A (which is itself a part of A) is in some sense nearer to B than A is, and so is the right half of that half, and so on. The point is the same as that of the argument discussed above: when A is divided in this way it turns out to have an unlimited number of parts. And again, this conclusion follows validly if one assumes certain views about physical bodies and spatial extension.

ARGUMENTS AGAINST MOTION

Four of Zeno's arguments against motion were particularly difficult to refute, according to Aristotle, who summarized them and offered solutions. They are the Dichotomy (or the Stadium); the Achilles; the Flying Arrow; and the Moving Rows. The following exposition is based mainly on Aristotle's penetrating discussion.

THE DICHOTOMY (OR THE STADIUM). This paradox argues that motion does not exist because it requires something impossible to happen. In order to cross a stadium from the starting line (A) to the finish line (B), after setting out one must reach A_1, the midpoint of the interval AB, before reaching B, then A_2, the midpoint of the interval A_1B, and so on. Each time one reaches the midpoint of an interval one still has another interval to cross with a midpoint of its own. There is an infinite number of intervals to cross. But it is impossible to cross an infinite number of intervals. Therefore one cannot reach the finish line.

The backbone of the argument lies in the following claims. (a) To move any distance one must always cross half the distance; (b) there is an infinite number of half-distances; (c) it is impossible to get completely through an infinite number of things one by one in a finite time; therefore (d) it is impossible to move any distance.

Aristotle, the primary source for the paradox, discusses the paradox several times in *Physics* (233a21, 239b9, 263a4). He rejects the inference to (d) on the grounds that the time of the motion is not finite, but infinite. Not that he supposes that every motion takes an infinite length of time; rather, as he has argued elsewhere in the *Physics* (6, 1–2), time is divisible in the same way that the distance traversed is divisible. If it takes a minute to cross the whole distance, it takes half a minute to cross the first half-distance, a quarter of a minute to cross the second half-distance, and so on. As the distances become smaller so does the time required to cross them, and the time interval required for the whole movement can be divided into the same number of subintervals as the number of subintervals into which the distance of the whole movement can be divided. So the time (just like the distance) is infinite in one respect (Aristotle calls this "infinite by division") and finite in another ("in extent").

Aristotle, however, does not stop here. He observes, "This solution is sufficient to use against the person who raised the question … but insufficient for the facts of the matter and the truth" (*Physics* 263a15), and then proceeds to discuss a deeper issue that the paradox raises: whether it is possible at all to perform an infinite number of acts, even the acts of getting through the sequence of decreasing time intervals. Granted that if one can do it, it will take a finite time, but can we do it at all?

Aristotle's solution to this stronger version of the paradox relies on his distinction between the actual infinite and the potential infinite. It is impossible to complete an actually infinite number of tasks, but possible to complete tasks that are potentially infinite. A line or a time-interval contains a potentially infinite number of points or instants. A point is actualized by stopping there; an instant is actualized by stopping then. Crossing the distance by making a single continuous movement does not actualize any of the midpoints. Hence, according to Aristotle's analysis, motion is possible because it does not involve completing an infinite number of tasks. Aristotle's final position on the paradox is that (d) does not follow from (a) (b) and (c), and Zeno committed an elementary blunder in supposing that it does, and moreover that (b) is true only if taken to claim that there is a potentially infinite number of half-distances, whereas (c) is true only if taken to refer to an actually infinite number of things and additionally if the proviso "in a finite time" is deleted.

THE ACHILLES. This paradox too argues against the possibility of motion. The swiftest runner (Achilles) gives the slowest (traditionally a tortoise, although no mention of the reptile occurs in Aristotle's account) a head start. But then he cannot catch up. He must first reach the tortoise's starting point (A), by which time the tortoise will have moved ahead some distance, however small, to another point (A_1). Getting to A proves to be only the first stage of a longer race. In the second stage of the race Achilles must reach A_1, but by then, the tortoise will have gone ahead an even smaller distance, to A_2, and so on. Each time Achilles reaches the point from which the tortoise has started, the tortoise is no longer there, so Achilles never catches up.

Aristotle observes, "This is the same argument as the Dichotomy, but it differs in not dividing the magnitude in half": Achilles runs more than twice as fast as the tortoise. Therefore, on the basis of his analysis of the Dichotomy argument, Aristotle thinks that the Achilles goes as follows: (a) To catch up with the tortoise, Achilles must always reach the point from which the tortoise started; (b) There is an infinite number of such starting points; (c) It is impossible to get completely through an infinite number of things one by one in a finite time; Therefore (d) Achilles cannot catch up with the tortoise. Unlike the Dichotomy, this argument does not conclude with the statement that motion is impossible. However, since the nature of motion implies that a faster runner will eventually catch up with a slower one, Zeno's conclusion that this cannot happen *entails* that motion cannot exist. According to Aristotle's analysis, though (which remains the dominant interpretation), the Achilles is fallacious since it commits the same mistake as the Dichotomy.

However, Aristotle's own statement of the Achilles (*Physics* 239b14) suggests that this interpretation is mistaken. The passage reads: "The slower will never be caught by the swiftest. For the pursuer must first reach the point from which the pursued departed, so that the slower must always be some distance in front." This summary says nothing about there being an infinite number of starting points or about the impossibility of performing an infinite number of tasks, or performing them in a finite time. Rather, the paradox turns on the words "always" and "never," which points to a different interpretation of the argument: (a) Achilles catches up with the tortoise when he reaches the point where the tortoise then is; (b) each time, before catching the tortoise, Achilles must reach the point from which the tortoise started; (c) when Achilles reaches the point from which the tortoise started, the tortoise has moved ahead; therefore, (d) the tortoise is always some distance ahead of Achilles [from (b and c)]; therefore (e) Achilles never catches up [from (d)].

This argument is different from the Dichotomy argument and is not open to the same objection. Where the Dichotomy is based on the impossibility of performing an infinite number of tasks, the Achilles turns on the words "always" and "never." The Achilles challenges the existence of motion if (e) is taken to assert that there is no time at which is it true that Achilles reaches the point where the tortoise then is; and this is in fact is the natural way to understand (e). But if in (e) "never" means "there is no time at which is it true that… " then in order for the argument to go through, (d) "always" must correspondingly mean "at all times is it true that … ." So, (d) must be taken to claim that the tortoise is ahead of Achilles at all times. In faact, this is a valid inference: If the tortoise is always (in this sense of "always') ahead, then Achilles never (in the corresponding sense of "never") catches up. But (d) appears obviously false, since faster things do in fact catch up with slower things. In the argument, (d) follows from (b) and (c), but these premises do not entail that the tortoise is ahead of Achilles at all times (as is needed for the argument to go through to (e)), only that the tortoise is still ahead at every time during the race. For example, if the tortoise's head start is nine miles and its speed is 1 m.p.h. while Achilles' speed is 10 m.p.h, then Achilles catches up with the tortoise at the end of one hour. During the race—before the hour is over—Achilles is always catching up and the tortoise is always ahead. But the scope of "always" is restricted to the time during which Achilles has not yet caught up; it does not have unrestricted scope ("at all times") as is needed for (d) to entail (e). As was noted above, (e) will follow only if there is no time at which it is true that Achilles has caught up, and the argument—in particular (b) and (c)—has given no reason to believe this.

THE FLYING ARROW. Aristotle's summary is as follows: If everything is always at rest when it is in a space equal to itself, and what is moving is always at an instant, the moving arrow is motionless" (*Physics* 239b5). The argument is incomplete as it stands and has been completed in various ways, one of which is the following: (a) Whenever something is in a space equal to itself, it is at rest (from Aristotle's summary); (b) an arrow is in a space equal to itself at each instant of its flight (supplemented); therefore (c) an arrow is at rest at each instant of its flight [from (a) and (b)]; (d) what is moving is always at an instant (from Aristotle's summary); therefore (e) during the whole of its flight the arrow is at rest [from (c) and (d)].

Aristotle objects: the argument "follows from assuming that time is composed of instants; if this is not conceded, the deduction will not go through" (*Physics* 239b31). This fastens on the move from (d) to (e). Aristotle's view of time, that it is not composed of instants, defeats the paradox. It can also be objected (again on Aristotelian grounds, see *Physics* 6, 3) that rest and motion take place over time intervals, not at instants. Motion requires occupying different places at different times; it is measured by the distance covered in an interval of time; nothing can move in an instant or for an instant. Likewise, rest is properly understand as the absence of motion: something is at rest during a time interval when it is not in motion. It makes no more sense to speak of rest in an instant or for an instant than it does to say that it is moving in or for an instant. This constitutes an objection to (c).

Another objection concerns (a), which implies that something in motion is not in a space equal to itself. But what does this mean? When is the moving thing not occupying a space equal to itself, and in what way? Two possible answers to the first of these questions are that it does not occupy a space equal to itself over the entire duration of its motion and (ii) at an instant during its motion. On interpretation (i) the idea is that in its motion the arrow occupies different positions at different instants and the sum (in some sense of the word) of those positions is larger than any of the individual positions. If the arrow initially occupies position AB (extending from point A to point B) and ends up at position CD (where the distance from C to D is equal to the distance from A to B), then the distance from A to D is equal to the space the arrow is in during the whole of its flight, and the distance from A to D is larger than the distance from A to B. Conversely, during any period when the arrow is at rest, it will be in a space equal to itself. Interpretation (i) makes sense of (a), but if make the argument invalid. Because (a) concerns motion and rest over the duration of the motion, which is an interval of time, not at an instant, and it is in general illegitimate to infer a conclusion about the behavior of something at individual instants in an interval from its ehavior during the interval as a whole, or vice versa. Consquently the inferences to (c) and (e) are invalid. On interpretation (ii) the move to (c) is valid, but there is no obvious reason why Zeno should have thought or should have expected anyone to agree that things change size during their motion, so that at any instant of its flight an arrow is larger or smaller than when it is at rest. Thus the argument fails: On one interpretation (i) it is invalid and on another (ii), although valid, it contains an unacceptable premise.

THE MOVING ROWS. Aristotle reports this argument as follows: "The fourth argument concerns equal bodies moving in a stadium alongside equal bodies in the opposite direction, the one group moving from the end of the stadium, the other from the middle, at equal speed. [Zeno] claims in this argument that it follows that half the time is equal to the double. … Let A's represent the equal stationary bodies, B's the bodies beginning from the middle, equal in number and size to the A's, and C's the bodies beginning from the end, equal in number and size to these and having the same speed as the B's. It follows that the first B is at the end at the same time as the first C, as the B's and C's move alongside one another, and the first C has come alongside all the B's but the first B has come alongside half the A's. And so the time is half. For each of them is alongside each thing for an equal time. It follows simultaneously that the first B has moved alongside all the C's, for the first C and the first B will be at the opposite ends simultaneously, because both have been alongside the A's for an equal amount of time" (*Physics* 239b33).

In discussing this passage, Simplicius, in *Physics* 1016, 19, provides diagrams to illustrate the starting position and the finish:

DIAGRAM 1: Starting position:

A A A A

D B B B B →   E

← C C C C

DIAGRAM 2: Finishing position:

A A A A

D   B B B B   E

C C C C

The kernel of the argument is as follows: (a) The time it takes the first B to have come alongside four C's is equal to the time it takes the first B to have come alongside two A's; (b) the first B is alongside each A and also alongside each C for the same amount of time; (c) but during its motion B is alongside two A's and B is alongside four C's; therefore (d) the total time B is alongside the A's is half the total time B is alongside the C's [from (b) and (c)]; therefore (e) half the time is equal to the double [from (a) and (d)]. Here "the double" refers to the time taken in being alongside the four C's; it means "the double of half the time," not "the double of the whole time." (Another

iteration of the argument will yield the conclusion that half the whole time equals double the whole time.)

Aristotle claims the argument is based on an elementary mistake: "The mistake is in thinking that an equal magnitude moving with equal speed takes an equal time in moving alongside something in motion as it does in moving alongside something at rest" (*Physics* 240a1). Thus, (b) is false, and consequently so are (d) and (e). Aristotle's analysis is correct if Zeno is treating the motion of extended bodies over a continuous magnitude. But could Zeno have committed so gross a blunder?

An influential interpretation acquits Zeno of this charge. Zeno is arguing not against the ordinary view of time (and perhaps space and motion as well) as being continuous, but against another possible view, that they are discrete: there are "atoms" of time and space, and motion proceeds in atomic "jumps," going from one atomic location to the next from one atomic instant to the next. Either something is moving or it is not; if it is moving, it is in successive locations at successive instants, if it is not, it is in the same location at successive instants. By hypothesis the B's and the C's are moving. One instant after the instant they occupy the starting position (Diagram 1) they will occupy the position illustrated in Diagram 3:

DIAGRAM 3: Position after one step:

A A A A

D B B B B → E

← C C C C

One instant later, they will occupy the position illustrated in Diagram 2. And contrary to what happens if space and time are continuous, there is no instant at which the lead B is next to the lead C (as in Diagrams 4 and 5).

DIAGRAM 4: This position does not occur:

A A A A

D B B B B → E

← C C C C

DIAGRAM 2: Nor does this position:

A A A A

D B B B B → E

← C C C C

Those who hold this interpretation have claimed the Moving Rows argument to be Zeno's most sophisticated argument and one that tells decisively against the view that time and space are atomic. But there are two obstacles to it. First, it conflicts with Aristotle's statement of the argument, which states that "each of them is alongside each thing for an equal time"; as just noted the lead B is never alongside the lead C. Second, there is no evidence in favor of it. Our sources give no hint that the bodies are atomic bodies or the times are atomic instants and there is no reason to think that such a theory of space and time had been considered by anyone as early as Zeno. The only reason given to support this interpretation is that Zeno was too clever to make the mistake that Aristotle finds—an assessment that is refuted by the equally elementary mistake diagnosed in the paradox of like and unlike.

TWO MORE PARADOXES

Zeno did not limit himself to arguments against the existence of plurality and motion. Two other arguments—the Millet Seed and the Place of Place—survive that challenge other deeply held beliefs.

THE MILLET SEED. This argument apparently criticizes the senses, therefore supporting Parmenides' view that the senses are radically unreliable. It is preserved in the form of a dialogue between Zeno and Protagoras (Simplicius, *In Physica*, 1108.18). In essence it states: (a) One millet seed or one ten-thousandth of a millet seed does not make a sound when it falls; (b) a bushel of millet seeds makes a sound when it falls; (c) there is a ratio between the bushel of millet seeds and one millet seed or one ten-thousandth of a millet seed; (d) the sounds made by the bushel, the millet seed, and the ten-thousandth of a millet seed have the same ratios as the ratios identified in (c); therefore (e) a millet seed makes a sound when it falls, and so does one ten-thousandth of a millet seed [from (b) and (d)]. (e) contradicts (a), which depends on the evidence of hearing. Therefore, hearing is unreliable.

Aristotle rebuts the paradox by saying that a threshold of force is needed to produce sound, and that the force of one millet seed falling is below the threshold. Other solutions suggest themselves as well.

THE PLACE OF PLACE. This argument is reported in several sources, including Aristotle's *Physics* (209a23, 210b22) and Simplicius's *In Physica*. Its essence is as follows: (a) Everything that exists is in a place; therefore (b) place exists; therefore (c) place is in a place [from (a) and

(b)]; (d) but this goes ad infinitum. Therefore (e) place does not exist.

Aristotle and his followers rebutted the argument by denying (a): not everything that exists is in a place, "for no one would say that health or courage or ten thousand other things were in a place" (Eudemus, quoted in Simplicius *In Physica* 563.25); and "nothing prevents the first place from being in something else, but not in it as in a place" (Aristotle, *Physics* 210b24). One can grant that a three-dimensional object has a place without conceding that its place is the kind of thing that can have a place. Alternatively one might accept the reasoning through (d) but deny that (d) entails (e). Not all infinite regresses are vicious.

CONCLUSION

The present treatment has offered versions of the most important of Zeno's surviving arguments and has suggested ways to refute them. This follows the tradition in discussing Zeno and the other Eleatic philosophers that has been dominant since Plato (*Sophist* 258B–D). Aristotle employed this practice and not just as a matter of historical interest. His philosophical method required him to take his predecessors' views into account and find solutions for puzzles and problems they presented, and his views on place, time, motion, and the infinite were framed with Zeno's paradoxes in mind. Philosophical interest in Zeno was renewed (notably by Bertrand Russell) after the modern conception of the infinite had been elaborated; once again contemporary philosophical tenets were employed to refute the paradoxes (principally the Dichotomy, the Achilles, and the Flying Arrow) and the challenge they present to ordinary views of space, time, and motion, and once again the discussion went beyond what Zeno proposed and encompassed related puzzles that his paradoxes suggested.

This astonishing ability to invent exciting and fruitful paradoxes is not Zeno's only contribution to philosophy. If Parmenides was the first pre-Socratic philosopher to employ deductive arguments, Zeno was the first to do so in prose, and his fragments show that he made great advances over Parmenides in the clarity of his reasoning and the complexity of his arguments. Also noteworthy is his use of deductions to point out the danger of maintaining familiar beliefs without examining them. These contributions easily outweigh any errors one may (frequently by employing concepts, distinctions and proof techniques that were not developed for centuries or millennia after Zeno's time) detect in his arguments.

See also Aristotle; Dialectic; Infinity in Mathematics and Logic; Logic, History of; Logical Paradoxes; Melissus of Samos; Motion; Parmenides of Elea; Plato; Russell, Bertrand Arthur William; Set Theory; Simplicius; Socrates.

Bibliography

PRIMARY TEXTS

Caveing, Maurice. *Zénon d'Élée. Prolégomènes aux doctrines du continu*. Paris: Vrin, 1982.

Diels, Hermann, and Walther Kranz. *Die Fragmente der Vorsokratiker*. 5th ed. Dublin and Zurich: Weidmann, 1951.

Lee, H. D. P. *Zeno of Elea*. Cambridge, U.K.: Cambridge University Press, 1936.

Untersteiner, Mario. *Zenone. Testimonianze e Frammenti*. Florence: La Nuova Italia, 1963.

SECONDARY TEXTS

Barnes, Jonathan. *The Presocratic Philosophers*. 2nd ed. Boston: Routledge & Kegan Paul, 1982.

Black, Max. "Achilles and the Tortoise." *Analysis* 11 (1951): 91–101.

Dillon, John. "More Evidence on Zeno of Elea?" *Archiv für Geschichte der Philosophie* 58 (1976): 221–222.

Dillon, John. "New Evidence on Zeno of Elea?" *Archiv für Geschichte der Philosophie* 56 (1974): 127–131.

Fränkel, Hermann. "Zeno of Elea's Attacks on Plurality." *American Journal of Philology* 63 (1942): 1–25, 193–206.

Furley, David. *Two Studies in the Greek Atomists*. Princeton, NJ: Princeton University Press, 1967.

Grünbaum, Adolf. *Modern Science and Zeno's Paradoxes*. Middletown, CT: Wesleyan University Press, 1967.

Guthrie, W. K. C. *A History of Greek Philosophy*, Vol. 2: *The Presocratic Tradition from Parmenides to Democritus*. Cambridge, U.K.: Cambridge University Press, 1965.

Hinton, J. M., and C. B. Martin. "Achilles and the Tortoise." *Analysis* 14 (1954): 56–68.

Kirk, G. S., and J. E. Raven. *The Presocratic Philosophers.* Cambridge, U.K.: Cambridge University Press, 1957.

Kirk, G. S., J. E. Raven, and M. Schofield. *The Presocratic Philosophers*. 2nd ed. Cambridge, U.K.: Cambridge University Press, 1982.

Lear, Jonathan. "A Note on Zeno's Arrow." *Phronesis* 26 (1981): 91–104.

McKirahan, Richard D. "La dichotomie de Zénon chez Aristote." In *Qu'est-ce que la philosophie présocratique*, edited by André Laks and Claire Louguet, 465–496. Villeneuve d'Ascq (Nord): Presses Universitaires du Septentrion, 2002.

McKirahan, Richard D. "Zeno." In *The Cambridge Companion to Early Greek Philosophy*, edited by A. A. Long, 134–158. Cambridge, U.K.: Cambridge University Press, 1999.

Owen, G. E. L. "Zeno and the Mathematicians." *Proceedings of the Aristotelian Society* 58 (1957–58): 199–222.

Russell, Bertrand. *Our Knowledge of the External World*. Chicago: Open Court, 1914.

Ryle, Gilbert. *Dilemmas*. Cambridge, U.K.: Cambridge University Press.

Salmon, Wesley, ed. *Zeno's Paradoxes*. New York: Bobbs-Merrill, 1970.

Solmsen, Friedrich. "The Tradition about Zeno of Elea Re-Examined." *Pronesis* 16 (1971): 116–141.

Sorabji, Richard. *Time, Creation, and the Continuum: Theories in Antiquity and the Early Middle Ages*. Ithaca, NY: Cornell University Press, 1983.

Tannery, Paul. *Pour l'histoire de la science*. Hellène. Paris: Alcan, 1887.

Thomas, L. E. "Achilles and the Tortoise." *Analysis* 12 (1952): 92–94.

Vlastos, Gregory. "A Note on Zeno's Arrow." *Phronesis* 11 (1966): 3–18.

Vlastos, Gregory. "Plato's Testimony concerning Zeno of Elea." *Journal of Hellenic Studies* 95 (1975): 136–162.

Vlastos, Gregory. "A Zenonian Argument against Plurality." In *Essays in Ancient Greek Philosophy*, edited by John P. Anton and George L. Kustas, 119–144. Albany: State University of New York Press, 1971.

Vlastos, Gregory. "Zeno of Elea." In *Encyclopedia of Philosophy*, 1st ed., edited by Paul Edwards, 369–379. New York: Macmillan, 1967.

Vlastos, Gregory. "Zeno's Race Course: With an Appendix on the Achilles." *Journal of the History of Philosophy* 4 (1966): 95–108.

Richard McKirahan (2005)

ZENO'S PARADOXES

See *Infinity in Mathematics and Logic*; *Zeno of Elea*

ZHANG ZAI

(1020–1077)

Born into a family from Kaifeng in Henan Province, Zhang Zai, styled Zihou, lived in a small town called Hengqu of Mei County in modern Shaanxi Province for the major part of his life and hence was known as Hengqu. After a few years of strenuous study of Daoism and Buddhism, he was encouraged by Fan Zhongyan to study *Zhongyong* (The doctrine of the mean) when he was only 21. He thus left Daoism and Buddhism behind and returned to the Confucian classics in a quest for a philosophy of the Confucian Way (*dao*). Like Zhou Dunyi, Zhang Zai finally set his mind on the *Yijing* (Book of changes) and change (*yi*) as the very essence of the Way. Zhang Zai's main work was *Zheng meng* (Rectifying the obscure), in which he developed his metaphysics of vital energy (*qi*). In this treatise he became the first philosopher to expound on vital energy as the essence of the Way and thus provide a systematic foundation for understanding and developing the cosmology and ontology of change in the Confucian tradition. Included in *Zheng meng* is the noteworthy essay "Ximing" (Western inscription), which gives a deeply felt statement of his view on the cosmos, human life, and ideal Confucian practice.

In comparison with Zhou Dunyi, who developed a cosmology of change in terms of the abstract notions of the great ultimate (*taiji*) and rationality (*li*), Zhang Zai sought a more unified and yet more detailed description of the formation and transformation of all things in the world in terms of vital energy. Zhang Zai's metaphysics of the ubiquitous vital energy both inspires and justifies his theory of the human mind as endowed with both cognitive and ethical capacity. Like Zhou, Zhang Zai applied his cosmology to his life and strove to be a Confucian sage. In his mind, the ideal of a Confucian sage was to let morality guide one's heart and mind on earth (and to prepare for heaven) while following the teachings of past sages, all in hopes of improving the destiny of the living and establishing a peace that would last for generations.

In his metaphysics of vital energy and dialectics of the transformation of vital energy, Zhang Zai conceives of vital energy as primarily subsisting in the great void (*taixu*) and as the primordial source of the generation of things in the world. The great void gives rise to vital energy, which differentiates yin and yang and the five powers (*wuxing*), which then gives rise to all the things in the world. In this process of generation, rationality (*li*), as the order and form of things, arises naturally from the vital energy. Unlike Zhu Xi (1130–1200) after him, Zhang Zai never views rationality as an autonomous or independent category of reality. Instead, he regards rationality as always inherent in the vital energy, and he regards all things as transformations of the vital energy, which alone determines the formation and destruction of things.

In his reflections on human nature, Zhang Zai distinguishes between the nature of heaven and earth (*tiandi zhi xing*) and the nature of temperament and desires (*qizhi zhi xing*) of a person. The former is rooted in the primary unformed vital energy, and the latter arises from the formed body of a person. The moral virtue in a person consists in grasping one's primary nature and controlling one's secondary nature.

In connection with this distinction of two natures, Zhang Zai also makes a distinction between knowledge of virtues (*dexing zhi zhi*) and knowledge of seeing and hearing (*jianwen zhi zhi*). The first sort of knowledge comes not from seeing and hearing but from reflection on the nature of heaven and earth until one sees the functions and powers of the Way and understands how one

embodies these functions and powers and can channel them to transform oneself into a virtuous sage. For Zhang Zai, cultivating one's nature not only opens one's mind to understanding and knowledge of the ultimate reality but also leads to human goodness (*ren*). When the mind understands ultimate reality, it can unify and command one's nature and emotions, because it can relate to and embody ultimate reality as the ultimate ground of unity and integration.

In conclusion, Zhang Zai's philosophy, as presented in his essay "Ximing," embodies a deep cosmic piety of the Confucian tradition that is both ethical and religious in spirit. In lieu of an explicit organized religion, Confucianism reaches for a cosmic sentiment of piety rooted in self-cultivation of a human-cosmic bond that would transcend and dissolve the problems of life and death. Hence Zhang Zai's final statement in "Ximing": "In life I feel at ease; in death I will be at peace"

See also Cheng Hao; Cheng Yi; Chinese Philosophy: Confucius; Shao Yong; Zhou Dunyi.

Bibliography

Zhang Zai. *Zhang Zai ji*. Beijing: Zhonghua shuju, 1978. Contains *Zheng meng* (Rectifying the obscure), *Jingxue li ku* (Treasury of li in the Confucian classics), *Hengqu "Yi" shuo* (Zhang Zai's discourse on the *Yijing*). This is best edition of Zhang Zai's work.

Chung-ying Cheng (2005)

ZHOU DUNYI

(1017–1073)

Zhou Dunyi was the first eleventh-century Chinese thinker who argued for the inseparability of metaphysics and ethics. His two works—*Taiji tushuo* (An Explanation of the Diagram of the Great Ultimate) and *Tongshu* (Penetrating the Book of Changes)—were major neo-Confucian writings on the metaphysical nature of moral cultivation.

In the *Taiji tushuo*, Zhou Dunyi comments on the Diagram of the Great Ultimate (*Taiji tu*). The Diagram, created by the Daoist Chen Tuan (c. 906–989), consists of five circles. The top circle is an empty one, symbolizing the universe as a self-generative and self-reproducing entity. The second circle contains intermixing semi-circles of dark and light colors, with the dark color representing the *yin* (the yielding cosmic force) and the light color the *yang* (the active cosmic force). The third circle is a group of five small circles, each represents one of the Five Phases (*wu xing*)—water, fire, wood, metal, and earth. Describing biological reproduction, the fourth circle depicts how the *yin* moves the female, and the *yang* the male. Building on the fourth circle, the fifth circle likens the process by which the myriad beings are produced through the union of the two sexes.

For Zhou, the Diagram of the Great Ultimate is a graphic depiction of the two-way flow between the whole and the part, the one and the many. Reading from the top to the bottom, the Diagram shows how the one gives rise to the many. It explains the ways in which the intermixing of the *yin* and *yang* creates the Five Phases and the multitude of beings. However, reading from the bottom to the top, the Diagram describes how the many are in fact one. It traces the steps by which the myriad beings are derived from the Five Phases and the *yin* and *yang*. No matter whether it is from one to many or from many to one, the Diagram shows that the universe is an organic system wherein part and whole play equal role. On this basis, Zhou explains the metaphysical nature of moral cultivation. He suggests that human beings, given their sensibility and consciousness, are free to decide whether they are active participants or stubborn obstructers of the universe's self-renewal. Hence, daily moral practices are as much metaphysical as ethical, involving a conscious decision to render human activities to be a part of the universe's self-regeneration.

In the *Tongshu*, Zhou Dunyi further explains the metaphysical nature of moral cultivation. According to Zhou, there are two reasons why the innate human goodness is called sincerity (*cheng*). First, the innate human goodness, although available to every human being, is hidden. One has to uncover it by being honest and true to oneself. Second, because all beings in this universe are intricately connected as a family of beings, to be true to oneself requires being true to others. Thus sincerity has to be rooted in altruism. For Zhou, Yan Hui (Confucius's favorite student) is a prime example of the cultivation of sincerity. Materially, Yan Hui was in an uninviting situation—having only a single bamboo dish of rice, a single gourd dish of drink, and living in a mean narrow lane. But spiritually, Yan Hui was always upbeat because he had developed a noble state of mind that linked him to the universe.

In paying tribute to Yan Hui, Zhou Dunyi in effect redefines the Confucian learning. In earlier times, learning was understood by Confucian scholars as being a loyal government official. Hence, successful prime ministers (such as Yi Yin of the Shang Dynasty in the seven-

teenth century BCE) were considered to be exemplary students of Confucius. By promoting Yan Hui as the true student of Confucius, Zhou sees learning as an individual quest for broadening the mind. A learned person, then, is not just a person of action; he is also a person of the right mind, who recognizes the inherent connections among all beings in this universe. By focusing on the cultivation of the mind, Zhou helps to distinguish neo-Confucianism from Classical Confucianism.

See also Cheng Hao; Cheng Yi; Confucius; Shao Yong; Zhang Zai; Zhu Xi.

Bibliography

Chan, Wing-tsit, tr. *A Source Book in Chinese Philosophy*. Princeton, NJ: Princeton University Press, 1963.

Fung, Yu-lan. *A History of Chinese Philosophy.* Vol. 2. Translated by Derk Bodde. Princeton, NJ: Princeton University Press, 1953.

Graham, A. C. *Two Chinese Philosophers: Ch'eng Ming-tao and Ch'eng Yi-chuan*. La Salle, IL: Open Court, 1992.

Huang Zongxi. *Song Yuan xue'an* [Intellectual Biographies of Song and Yuan Dynasties]. Taibei: Taiwan Zhonghua shuju reprint.

Mou Zongsan. *Xinti yu xingti* [The Substance of the Mind and the Substance of Human Nature]. Vol. 1. Taibei: Zhengzhong shuju, 1969.

Tu, Wei-ming. *Confucian Thought: Selfhood as Creative Transformation*. Albany: State University of New York Press, 1985.

Zhou Dunyi. *Zhou Dunyi ji* [Collected Works of Zhou Dunyi]. Beijing: Zhonghua shuju, 1990.

Tze-ki Hon (2005)

ZHUANGZI

(b. 369 BCE)

Zhuangzi, the greatest Daoist next to Laozi, was also known by his private name, Zhou. Not much is known about his life except that he was a minor government official at one time and that he later declined a prime ministership in the state of Chu to retain his freedom. Although Zhuangzi and Mencius were contemporaries, they were not acquainted with each other's teachings. Zhuangzi advanced the concept of Dao and gave Daoism a dynamic character. To him, Dao as Nature is not only spontaneity but also a constant flux, for all things are in a state of perpetual "self-transformation," each according to its own nature and in its own way. If there is an agent directing this process, there is no evidence of it. Things seem to develop from simple to higher life and finally to man, but man will return to the simple stuff, thus completing a cycle of transformation.

In this unceasing transfiguration, things appear and disappear. In such a universe "time cannot be recalled" and things move like "a galloping horse." They seem to be different, some large and some small, some beautiful and some ugly, but Dao equalizes them as one. This is Zhuangzi's famous doctrine of the "equality of all things." According to it, reality and unreality, right and wrong, life and death, beauty and ugliness, and all conceivable opposites are reduced to an underlying unity. This is possible because all distinctions and oppositions are merely relative, because they are the result of a subjective point of view, because they mutually cause each other, and because opposites are resolved in Dao. By the doctrine of "mutual causation" Zhuangzi meant that a thing necessarily produces its opposite; for instance, "this" implies "that," life ends in death, construction requires destruction, and so forth. By the resolution of opposites Zhuangzi meant that a thing and its opposite, both being extremes, need to be synthesized. But the synthesis is itself an extreme that requires a synthesis. At the end Dao will synthesize all, in a dialectic manner not unlike that of G. W. F. Hegel.

In Zhuangzi's philosophy the pure man abides in the great One, wherein he finds purity and peace. He becomes a "companion of Nature" and does not substitute the way of man for the way of Nature. He rejects all distinctions and seeks no self, fame, or success. He seeks "great knowledge," which is all-embracing and extensive, and discards "small knowledge," which is partial and discriminative. He "fasts in his mind" and "sits down and forgets everything"—especially the so-called humanity and righteousness of hypocritical society; he "travels in the realm of infinity." In this way he cultivates "profound virtue," and achieves a "great concord" with Dao. Herein he finds spiritual peace and "emancipation."

Both the mystical and fatalistic elements are obvious, and in these Zhuangzi went beyond Laozi. He was also more transcendental, for while Laozi's chief concern was how to govern, Zhuangzi's primary interest was to "roam beyond the mundane world," in spite of the fact that his ideal being is "sagely within" and "kingly without," that is, both transcendental and mundane. Nevertheless, Zhuangzi stresses the individual more than does Laozi. To be in accord with Dao, everything must nourish its own nature and follow its own destiny. The eagle should rise to the clouds, but the dove should hop from treetop to treetop. If a man were to shorten the crane's neck because it is long or to lengthen the duck's leg because it is short,

that would be interfering with Nature. Spiritual freedom and peace can be achieved only through knowing one's own nature and capacity and being able to adapt oneself to the universal process of transformation. Although the ultimate goal is oneness with Dao, one's individuality is to be clearly recognized. Individual differences are not to be taken as basis for discrimination, but neither are they to be denied or ignored. This respect for individual nature and destiny eventually led to the emphasis on the particular nature in neo-Daoism.

See also Chinese Philosophy; Hegel, Georg Wilhelm Friedrich; Laozi; Mencius.

Bibliography

Available in English is *Chuang Tzu,* translated by Herbert A. Giles (London: Allen and Unwin, 1961); the authorship of this work is a very controversial matter. Most scholars accept the first seven chapters, the so-called inner chapters, as authentic and the remaining 26 chapters as later additions, either partly or in whole dating from the third to the first century BCE. Selections have been translated in Wing-tsit Chan, *A Source Book in Chinese Philosophy* (Princeton, NJ: Princeton University Press, 1963). See also Fung Yu-lan, *A Short History of Chinese Philosophy* (New York: Macmillan, 1948).

Fung, Yu-lan. *A Taoist Classic: Chuang-tzu*; A New Selected Translation with an Exposition of the Philosophy of Kuo Hsiang. Beijing, Foreign Language Press, 1998.

Graham, Angus C. *The Inner Chapters*. Boston: Unwin Paperbacks, 1989.

Kjellberg, Paul and Ivanhoe, Philip J. eds. *Essays on Skepticism, Relativism, and Ethics in the Zhuangzi*. Albany: State University of New York Press, 1996.

Mair, Victor Mair ed. *Experimental Essays on Chuang-tzu*. Honolulu: University of Hawaii Press, 1983.

Mair, Victor. *Wandering on the Way: Early Taoist Tales and Parables of Chuang Tzu*. Honolulu: University of Hawaii Press, 1998.

Watson, Burton. *The Complete Works of Chuang Tzu*. New York: Columbia University Press, 1968.

Wu, Kuang-ming. *Chuang Tzu: World Philosopher at Play*. American Academy of Religion, Studies in Religion, 26. New York: Crossroad, 1982.

Wing-tsit Chan (1967)
Bibliography updated by Huichieh Loy (2005)

ZHU XI (CHU HSI)

(1130–1200)

Zhu Xi was a leading scholar, thinker, and teacher of the revival of philosophical Confucianism known at the time as *Daoxue* (learning of the way), often referred to as neo-Confucianism. The prolific author of texts synthesizing the views of his immediate predecessors and reinterpreting the classical canon, Zhu Xi attained a status in the Chinese tradition comparable to that of Thomas Aquinas in the European world. Zhu's influence has been even more pervasive and long-lived, however; from 1313 until their abolition in 1905, China's civil service examinations took Zhu's commentaries to be the authoritative interpretations of the classics. Hence for nearly a millennium every literate individual in China had at least some familiarity with Zhu's teachings.

Zhu was born into turbulent times. In 1127 Jurchen people conquered northern China. Zhu's father was among many who protested the humiliating peace treaty that China was forced to accept, and he was demoted to a rural position in Anhui, where Zhu was born. Zhu took up his father's politics as he matured, committing himself to the hawkish group that wanted to take back the north. Partly out of disenchantment with the regime's failure to follow such policies, Zhu never played a significant role in the national bureaucracy despite having passed the highest-level civil service exam and having received his *jinshi* degree at the age of nineteen.

At first Zhu was quite eclectic in his intellectual and spiritual interests, but several encounters in his twenties with the staunch Confucian Li Tong (1095–1163) convinced him to commit himself wholeheartedly to the Confucianism associated with two celebrated thinkers from the eleventh century, the brothers Cheng Hao (1032–1085) and Cheng Yi (1033–1107). Over much of the rest of his life, Zhu held sinecure positions as a temple guardian and devoted himself to study, writing, and teaching. He produced a huge corpus of essays and commentaries that, together with the voluminous recorded and published conversations between Zhu and his students, articulated and defended a creative synthesis that has come to define mainstream neo-Confucianism.

Zhu's philosophical system was the product of the range of interlocking areas his writings encompassed: ontology, cosmology, nature (human and otherwise), psychology, epistemology, moral cultivation, ethics, and politics. In addition, despite his distance from national politics, he was deeply concerned with the practical import of his views; among other things, he worked to revitalize independent academies and advocated a form of village self-government known as a "community compact." Like most long-lived and prolific thinkers, Zhu revised his outlook over time, and many expressions of his ideas are highly contextual, depending on the circumstances he was addressing.

The central concepts in Zhu's ontology are *li* (pattern or principle) and *qi* (material force). Zhu saw that the patterns followed by one thing or in one affair interact with those of countless others, as when the unchecked growth of one tree stunts the growth of others nearby, and argued that there is an all-encompassing *li* in accord with which the myriad subsidiary patterns are able to develop in order and harmony. *Li* are the patterns underlying the constant change of the psychological and material world; *qi* is the dynamic stuff of which this world is composed. *Qi*, in turn, can be analyzed as either *yin* or *yang*, depending on whether it is contracting or expanding, soft or hard, dark or light, and so on. Each thing or affair has its own *li*, which in one sense can be understood as the possibilities for that thing: the patterns of change it can instantiate. Zhu held that the patterns followed by one thing interact with those of countless others, as when the unchecked growth of one tree stunts the growth of others nearby; he argued that there is an all-encompassing *li* in accord with which the myriad subsidiary patterns are able to develop in order and harmony.

From the human perspective, this all-encompassing *li* is called "moral pattern (*yi li*)"; applied to the cosmos, it is "nature's pattern (*tian li*)." Zhu believed *li* to have logical priority over *qi* but to have no existence independent of *qi*. He borrowed the term "Great Ultimate (*taiji*)" from Zhou Dunyi (1017–1073) to refer to the source of all creativity, the not-yet-material totality of all patterns in which *qi* has yet to be differentiated into *yin* and *yang*. The ideas of unceasing creativity and its original goodness lie at the heart of Zhu's metaphysics.

The view that nature has at its core goodness, harmony, creativity, and order applies equally to humans and to the cosmos at large. Zhu developed ideas of Cheng Yi and others to explain how we can be said to have good natures yet regularly have problematic thoughts and feelings. He also discussed the things we need to do to realize the pure goodness of our original natures. One core idea is that problems occur when our "unactualized (*weifa*)" minds become "actualized (*yifa*)" via our real and imperfect bodies and their desires. Our moral natures themselves have some reality, as can be seen by the near-ubiquitous spontaneous compassionate response we have to the suffering of innocents, but our *qi*—the psycho-physical reality of our emotions, habits, and so on—is not, except in sages, purely expressive of the equilibrium in our unactualized minds.

What is to be done? Zhu believed that education should begin with a period of "lesser learning" in which one learns good habits without delving into the reasoning that justifies them. In the subsequent "greater learning," one continues to nurture the "reverence (*jing*)" for moral pattern while beginning to investigate the theoretical grounding of those patterns. This "investigation of things (*gewu*)," which relied in part on a controversial redaction of the brief classic text *Greater Learning*, was the subject of much subsequent debate. Zhu seems to have had two kinds of investigation foremost in mind: the patterns observed in peoples' interactions with one another and the patterns instantiated by ancient sages and worthies, as recorded in the classics and histories. Indeed, reading was a central focus of his teaching, just as textual scholarship was a central focus of his scholarship. Zhu believed that without reference to external models of proper patterns, students would be too easily misled by introspection into their own reactions and motivations, which might be clouded by the impurities of one's *qi*. The goal of Zhu's teachings was practical: Given the centrality of "benevolence (*ren*)" in the life of the morally worthy person (and in the acme of human personality, the sage), he sought to motivate people to improve themselves by the most reliable method.

See also Cheng Hao; Cheng Yi; Chinese Philosophy; Confucius; Cosmology; Ontology, History of; Thomas Aquinas, St.; Zhou Dunyi.

Bibliography

ORIGINAL SOURCES

A great deal of work has been done since 1990 in compiling modern editions of Zhu's corpus. Most comprehensive is the *Complete Works of Zhu Xi* (*Zhuzi Quanshu*) from Shanghai Classics Press and Anhui Education Press; the first volumes began to appear in 2002. In 1996 Sichuan Education Press published *Collected Works of Zhu Xi* (*Zhu Xi Ji*) in ten volumes, which contains all of Zhu's formal writings. In addition, there are numerous editions of Zhu's collected sayings (*Zhuzi Yulei*) and other monographs available.

TRANSLATIONS

Daniel Gardner has provided perhaps the best introduction to Zhu's thought by translating selections from Zhu's conversations about learning, in *Learning to Be a Sage* (Berkeley: University of California Press, 1990). Further depth is provided by Allen Wittenborn's excellent, complete translation of Zhu's *Further Reflections on Things at Hand* (Lanham, MD: University Press of America, 1991). Wing-tsit Chan's translations are still quite helpful; see both the section on Zhu in his *Sourcebook in Chinese Philosophy* (Princeton, NJ: Princeton University Press, 1963) and the numerous comments from Zhu included in the important collection of earlier neo-Confucian writings that Zhu coedited, *Reflections on Things at Hand* (New York: Columbia University Press, 1967).

SECONDARY STUDIES

The closest thing to a general, book-length study of Zhu in English is Julia Ching, *The Religious Thought of Chu Hsi* (Oxford: Oxford University Press, 2000). Donald Munro's *Images of Human Nature* (Princeton, NJ: Princeton University Press, 1988) critically engages a range of Zhu's ideas by focusing on the images he uses to structure his thinking. The best work in English on the intellectual context in which Zhu's ideas developed is Hoyt Tillman, *Confucian Discourse and Chu Hsi's Ascendancy* (Honolulu, University of Hawaii Press, 1992). There are also a handful of more specialized monographs and many articles devoted to Zhu; a particularly high-quality collection of the latter is Wing-tsit Chan, ed., *Chu Hsi and Neo-Confucianism* (Honolulu, 1986). Chan's *Chu Hsi: New Studies* (Honolulu, 1989) also contains a wide range of helpful essays. Finally, Chinese-language studies of Zhu are flourishing. Two particularly important works are Chen Lai, *A Study of Zhu Xi's Philosophy* (*Zhu Xi zhexue yanjiu*) (Beijing, 1988) and Yu Yingshi, *Zhu Xi's Historical World* (*Zhu Xi de lishi shijie*) (Taibei, 2003).

Stephen C. Angle (2005)

ZIEHEN, THEODOR

(1862–1950)

Theodor Ziehen, the German psychologist and philosopher, was born in Frankfurt am Main and served as professor of psychiatry at the universities of Jena, Utrecht, Halle, and Berlin. He lived as a private scholar in Wiesbaden from 1912 to 1917, when he returned to teaching as professor of philosophy and psychology at the University of Halle. He retired in 1930.

Ziehen's viewpoint in epistemology is in the broadest sense positivistic. Knowledge must start with that which is experientially given, which Ziehen termed "becomings" (*gignomene*). From this "gignomenal principle" follows the "principle of immanence," according to which there is no such thing as metaphysical knowledge of the transcendental, and therefore it is nonsensical to want to know that which is not given. The first task of philosophy thus consists in seeking the laws of all that is given (the "positivistic" or "nomistic" principle). According to Ziehen, such a "gignomenological" investigation leads to the conclusion that the traditional antithesis between the subjective, mental world of consciousness and the objective, material external world is inadmissible because the given is "psychophysically neutral." We must, however, distinguish two kinds of law-governed relations: The *gignomene* are to be called mental insofar as they are considered with regard to their "parallel components" (the mental, subjective ingredients of experiences, which parallel certain physiological processes); and the *gignomene* are to be understood as physical insofar as attention is fixed on their "reduction ingredients" ("reducts"), which are subject to causal laws.

Thus, Ziehen did not distinguish in the customary manner between material and mental reality; rather, he sought to understand the structure of the given, which he claimed to be the sole reality, in terms of two kinds of regularities—causal laws and parallel laws. Viewed from this "binomistic" standpoint, which assumes a twofold conformance to law in the given, real things appear as possibilities of perception, as potential perceptions, as "virtual reducts" that are both "transgressive" and "intramental." They lie beyond the boundaries of the individual content of consciousness, but they are nevertheless not situated "behind" experience but are immanent in it. Thus, real things represent certain aspects of experience that are determined by the causal type of laws. The processes governed by causal law ("the laws of nature") go along specific paths with a specific velocity; through the parallel laws that direct mental life, the *gignomene* are transformed into individual experiences.

Thus, for Ziehen psychology stood in contrast with the other natural sciences—the causal sciences—as the science of the "parallel component" of the given. Ziehen combated what he considered to be mythologizing faculty psychology, including Wilhelm Wundt's theory of apperception. He advocated a physiologically oriented, analytic, serial, or associationist approach to the subject. To association he added a second factor regulating the course of consciousness—the "constellation." A constellation arises at a given time from the mutual inhibition and stimulation of ideas, and it selects from the many ideas that are associated and, hence, ready for reproduction. In addition to association and constellation, Ziehen assumed three other basic mental functions—synthesis, analysis, and comparison.

Besides the causal laws and the parallel laws, Ziehen assumed a third, more general kind of regularity—conformity to logical laws—common to and set above the two other kinds of laws.

Ziehen also wrote on the philosophy of religion. He identified God with the regularity governing the world. God must be thought of as the essence or embodiment of "regularity in general"; as the totality of logical regularity, of natural laws, and of the laws of mental and spiritual life. It would be an inadmissible anthropomorphism to look beyond the regularities for a personal source of them.

See also Basic Statements; Epistemology; Laws of Nature; Philosophy of Religion; Positivism; Psychology; Wundt, Wilhelm.

Bibliography

WORKS BY ZIEHEN

Leitfaden der physiologischen Psychologie. Jena, Germany: Fischer, 1891; 12th ed., 1924.

Erkenntnistheorie auf psychophysiologischer und physikalischer Grundlage. Jena, Germany: Fischer, 1913; 2nd ed., published as *Erkenntnistheorie,* 2 parts, 1934–1939.

Die Grundlagen der Psychologic, 2 vols. Leipzig and Berlin: Teubner, 1915.

Lehrbuch der Logik. Bonn: A. Marcus and E. Weber, 1920.

Autobiography in *Die Philosophie der Gegenwart in Selbstdarstellungen,* Vol. IV. Edited by Raymond Schmidt. Leipzig: Meiner, 1923.

Vorlesungen über die Ästhetik, 2 parts. Halle: Niemeyer, 1923–1925.

Die Grundlagen der Religionsphilosophie. Leipzig, 1928.

WORKS ON ZIEHEN

Graewe, H. "Theodor Ziehen zum 90. Geburtstag." *Die Pyramide* (Innsbruck) (11) (1952): 201–202.

Peters, Wilhelm. "Theodor Ziehen zum 70. Geburtstag." *Kant-Studien* 37 (1932): 237–240.

Ulrich, Martha. "Der Ziehen'sche Binomismus und sein Verhältnis zur Philosophie der Gegenwart." *Kant-Studien* 25 (1921): 366–395.

Franz Austeda (1967)
Translated by Albert E. Blumberg

ZOROASTRIANISM

"Zoroastrianism," for more than a thousand years the dominant religion of Persia, is founded on the teachings of the prophet Zarathustra. (Zoroaster is an often used version of his name, and from it the name of the religion is derived; this version reflects ancient Greek transliteration.) Four main stages in the religion's history can be distinguished: the early faith as promulgated by Zarathustra himself; the religion of the Persian Empire under Darius I (who ruled 521–486 BCE) and his Achaemenid successors; its renewal under the Arsacid (250 BCE–226 CE) and Sassanian (226–641) dynasties; and the late period, when the religion was swamped by Islam but continued as the faith of a minority, some of whom settled in India and are known as Parsis (literally "Persians").

SCRIPTURES

The scriptures are known as the Avesta (or Zend-Avesta) and consist of various hymns, treatises, and poems. They comprise the Yasna, a collection of liturgical writings that contains the important Gāthās (literally "songs"), possibly written by Zarathustra himself; the Yashts, hymns to various divinities; and the Vendidād, which contains prescriptions for rituals of purification and so on. Many of these writings belong to a period when Zoroastrianism had become overlaid by polytheistic elements; some may date from as late as the fourth century, although the majority were composed much earlier. From the fourth century a further and extensive set of writings, which expressed the reformed theology of the Sassanian period, was compiled in the later language of Pahlavi.

ZARATHUSTRA AND HIS TEACHING

There is considerable dispute and uncertainty about the date and place of the prophet's life. Although Greek sources mention dates of up to several thousand years BCE, the most plausible theories are that he lived in the tenth or ninth century BCE or in the sixth or fifth. Although certain evidence points to his having lived in eastern Iran, the language of the Gāthās has been found to belong to northwest Iran. According to the traditions surrounding Zarathustra's life, he converted King Vishtaspa (Hystaspes in Greek transliteration), which proved decisive for the spread of the new religion. Vishtaspa ruled parts of eastern Iran and was the father of Darius the Great, a strong exponent and protector of the faith. These facts lend some support to the hypothesis that Zarathustra lived at the later date and in eastern Iran.

Although traditional accounts of Zarathustra's life are heavily overlaid by legend, it is probable that he was the son of a pagan priest of a pastoral tribe. At the age of thirty or a little later, he had a powerful religious experience, probably of a prophetic nature, analogous to the inaugural visions of such Old Testament prophets as Isaiah. He is reported to have encountered the angel Vohu Manah ("Good Thought"), who took him to the great spirit Ahura-Mazda ("The Wise Lord"), Zarathustra's name for God. Other revelations combined to induce him to preach a purified religion, combating the existing Persian polytheism, which had similarities to the Vedic religion of India. At first he met with considerable opposition, but the conversion of Vishtaspa paved the way for Zarathustra's wide influence, despite the king's later defeat in war and the occupation of his capital. Zarathustra is said to have been killed at the age of seventy-seven during Vishtaspa's defeat, but according to

later accounts, he died while performing the fire sacrifice, an important element in the new cultus.

Zarathustra's God had the attributes of a sky god, like the Indian god Varuna. Both were ethical and celestial and were worshiped by the Indo-European Mitanni of the mountainous region to the north of the Mesopotamian plain during the latter part of the second millennium BCE. Zarathustra strongly denounced the cult of the gods of popular religion, equating such beings with evil spirits who seduced men from the worship of the one Spirit. The belief in the malicious opposition to the purified religion that he preached and the incompatibility of Ahura-Mazda's goodness with the creation of evil led Zarathustra to conceive of a cosmic opposition to God. He mentions Drūj ("The Lie"), an evil force waging war against Ahura-Mazda. From this early concept developed the later Zoroastrian theology of dualism.

Although Zarathustra attacked the existing religion, he also compromised with it. A slight concession to polytheism was involved in the doctrine of the Amesha-Spentas ("Immortal Holy Ones"), such as Dominion and Immortality, which were personified qualities of Ahura-Mazda. It is probable that Zarathustra was making use of certain aspects of the existing mythology and transforming them into attributes and powers of God. He seems to have used the fire sacrifice, a prominent feature of later and modern Zoroastrianism, transforming what had previously been part of the fabric of the polytheistic cultus. Zarathustra's fire sacrifice was also related in origin to the ritual surrounding the figure of Agni (Fire) in ancient Indian religion.

He preached an ethic based on the social life of the husbandman, the good man being one who tends his cattle and tills the soil in a spirit of peace and neighborliness. The good man must also resist worshipers of the *daevas* (gods), who, together with the evil spirit opposed to Ahura-Mazda, threaten the farmer's livelihood. These ideas probably reflected the social conditions of Zarathustra's time and country, when there was a transition from the nomadic to the pastoral life. The *daeva*-worshipers would then represent bands of nomadic raiders, and the new purified religion would be a means of cementing a settled, pastoral fabric of society. One of the Gāthās is a dialogue in which there figures a mysterious being called the Ox Soul, who complains of the bad treatment meted out to cattle upon the earth. The angel Vohu Manah promises that they will be protected by Zarathustra, who prays earnestly to Ahura-Mazda for assistance. These connections between the new religion and a settled cattle-raising society later became obscured when Zoroastrianism became the religion of the Persian Empire and when they were no longer relevant.

The moral life, however, was not confined to neighborliness and resistance to evil *daeva*-worshipers. It was part of a much wider cosmic struggle, in which the good man participates in the battle of Ahura-Mazda against the evil Angra Mainyu, the chief agent of The Lie (in later language, these were called, respectively, Ormazd and Ahriman). The battle will consummate in a final judgment, involving the resurrection of the dead and the banishment of the wicked to the regions of punishment. This notion of a general judgment was supplemented by a dramatic picture of the individual's judgment. He must cross to Ahura-Mazda's paradise over the narrow bridge called Chinvat. If his bad deeds outweigh his good ones, he will topple into the dreadful, yawning abyss. Some of this Zoroastrian eschatology came to influence Jewish eschatology, partly through the contact with Persia consequent to the Exile and partly because of the succeeding Persian suzerainty over Israel. Zoroastrianism, therefore, indirectly influenced Christianity.

DEVELOPMENT OF RITUAL

When Zoroastrianism came to be the dominant religion of the Persian Empire during the Achaemenid dynasty, there was an increasing trend toward restoring the cult of lesser deities. This was a partial consequence of the adoption of Zoroastrianism as the state cult. Artaxerxes II, for instance, caused images of the goddess Anahita (connected in origin to Ishtar, the Babylonian fertility deity) to be set up in the chief cities of the empire. The cultus came to be administered, in some areas at least, by the priestly class known as the Magi, from which term the word *magic* is derived; the Magi also came to figure in Christian legend about the birth of Christ. This priestly class was probably of Median origin. At first, the Magi had opposed the new faith, but after having adopted it, they began to change its character by importing extensive magical and ritual practices into it. Thus, the later portions of the Avesta contain spells and incantations. Further, the Gāthās were no longer treated simply as expressing Zarathustra's religion and teachings but as having intrinsic magical powers. Their proper repetition could combat the evil powers by which men were beset. However, the full history of the development of Zoroastrianism toward a ritualistic cult has never been fully disentangled, partly because of the intervening changes brought about in the late fourth century BCE by Alexander's conquest of the Persian Empire and its subsequent division among Greek dynasties. This Hellenistic period,

lasting until the Parthian era in the second century BCE (begun by Mithridates I of the Arsacid dynasty), saw further syncretism, an offshoot of which was Mithraism, the cult of Mithra or Mithras, which later became important in the Roman Empire as a mystery religion.

DEVELOPMENT OF COSMOLOGY

While Zarathustra had stressed the ethical dimension of religion and the Mazdaism, as Zoroastrianism was later called, of the Achaemenid period had emphasized its ritual dimension, the reformed Zoroastrianism established in the Sassanian period displayed a strong interest in the doctrinal dimension of the faith. It is chiefly in this phase of Zoroastrianism that we discover a speculative interest in the workings of the universe. A theory of history was worked out that divided historical time into four eras, each lasting 3,000 years. In the first era, God brings into existence the angelic spirits and *fravashis,* which are the eternal prototypes of creatures (and, preeminently, of human beings). Since Ahura-Mazda creates by means of thought and since he foresees Angra Mainyu, the latter comes into existence. During the second period, the primeval man, Gayomard, and the primeval Ox (the prototype of the animal realm) exist undisturbed, but at the beginning of the third epoch the Evil Spirit, Angra Mainyu, succeeds in attacking and destroying them. From the seed of these two primeval beings men and animals arise, and there is a mixture of good and evil in the world. The last era begins with Zarathustra's mission; it will culminate in the final divine victory, which will occur partly through the agency of Soshyans, a semidivine savior. The universe will then be restored to an everlasting purified state in which the saved, now immortal, sing the praises of Ahura-Mazda. In this theory of history, the individual's life is linked to the unfolding cosmic drama.

The theory, while assigning the final victory to God, allows the nature and scale of the Evil One's operations to be alarming. Further, if Angra Mainyu arises through the thought of Ahura-Mazda, then evil comes from the Creator. This put the Zoroastrian theologians in a dilemma, and so attempts were made to work out doctrines that would more consistently explain the existence of evil. For instance, the movement known as Zurvanism held that both Ahura-Mazda and Angra Mainyu issued from a first principle, Zurvān (Infinite Time). Zurvān is beyond good and evil; only with the realm of finite time is the contrast between good and evil meaningful. On the other hand, Zurvān, the Supreme Being, dwells in an eternal state, raised beyond the conflicts and contrasts that exist in the temporal world.

INFLUENCE AND SURVIVAL

Elements of Zoroastrian teaching and mythology entered into Mithraism and Manichaeanism, and its eschatology had a marked influence on the Judeo-Christian tradition. However, the Muslim conquest of Persia in the seventh century largely destroyed the religion in its home country. Its survival in India was due to the Zoroastrians who emigrated in order to escape Muslim persecution. This Parsi community, centered chiefly on the west coast in and around Bombay, has maintained the cultus and interprets the faith in a strictly monotheistic sense. Their emphasis on education has given them an influence out of all proportion to their numbers.

See also Cosmology; Dualism in the Philosophy of Mind; Evil, The Problem of; Freud, Sigmund; Mani and Manichaeism.

Bibliography

Duchesne-Guillemin, J. *Zoroastre.* Paris: Maisonneuve, 1948.

Ghirshman, R. *Iran.* Baltimore, 1954.

Modi, J. J. *Religious Ceremonies and Customs of the Parsis,* 2nd ed. London, 1954.

Zaehner, R. C. *The Dawn and Twilight of Zoroastrianism.* London: Weidenfeld and Nicolson, 1961.

Zaehner, R. C. *Zurvān: A Zoroastrian Dilemma.* Oxford: Clarendon Press, 1955.

Zend-Avesta. Translated by J. Darmesteter in *Sacred Books of the East,* edited by F. Max Müller. Oxford: Clarendon Press, 1883.

RECOMMENDED READING

Boyce, Mary. *A History of Zoroastrianism.* 3 vols. Leiden: Brill, 1975–1982.

Boyce, Mary. *A Persian Stronghold of Zoroastrianism.* Oxford: Clarendon Press, 1977.

Boyce, Mary. *Zoroastrians, Their Beliefs and Practices,* reprint edition. London: Routledge, 2001.

Clark, Peter. *Zoroastrianism: An Introduction to an Ancient Faith.* Brighton, UK, and Portland, OR: Sussex Academic Press, 1998.

Hinnells, John. *Zoroastrian and Parsi Studies: Selected Works of John Hinnells.* Burlington, VT: Ashgate, 2000.

Insler, S. H. *The Gāthās of Zarathustra.* Leiden: Brill, 1975.

Kriwaczek, Paul. *In Search of Zarathustra: The First Prophet and the Ideas that Changed the World.* New York: Knopf, 2003.

Mahalingam, Indira, and Brian Carr, eds. *Companion Encyclopedia of Asian Philosophy.* London: Routledge, 1997.

Ninian Smart (1967)

Bibliography updated by Christian B. Miller (2005)

ZUBIRI, XAVIER

(1898–1983)

Xavier Zubiri, the Spanish Christian ontologist, was born in San Sebastián. He was professor of the history of philosophy in Madrid from 1926 to 1936 and in Barcelona from 1940 to 1942, after an absence abroad during the Spanish Civil War. He then left university teaching to give well-attended "private courses" in Madrid. His influence in Spain has been out of all proportion to the scanty amount of his published work.

Zubiri has been called a Christian existentialist, and indeed that is one aspect of his effort to synthesize neoscholastic theology with certain contemporary philosophies (those of Edmund Husserl, Martin Heidegger, and José Ortega y Gasset) and with modern science. To achieve this harmonizing of separate disciplines, Zubiri undertook studies in theology, philosophy, and natural science that could well have occupied three scholarly lives. He took a doctorate of theology in Rome and of philosophy in Madrid (where he studied under Ortega) before attending Heidegger's lectures in Freiburg and studying physics, biology, and Asian languages in various European centers. He translated into Spanish not only metaphysical works by Heidegger but also texts on quantum theory, atomic science, and mathematical physics generally.

From this extensive study Zubiri concluded that positive science and Catholic philosophy were separate points of view concerning the same reality. The philosopher-theologian cannot dispute, correct, or complete anything in science, but neither does he have to accept the philosophical opinions of scientists. The connection between these two parallel approaches to reality is simply that the sciences always leave us metaphysically hungry and with the feeling that they have not exhausted all the possibilities of knowledge, so they impel us to turn to philosophy. It is only when we come to philosophy in this way that it is really valuable; any philosophy that is undertaken without being forced upon us by scientific study is insipid.

What the sciences must get from philosophy, Zubiri claims, is an idea of nature, a theory of being to delimit their ontological horizons. They cannot themselves build such an idea out of positive facts, although they can criticize and reject unsuitable concepts of nature offered by philosophers. Aristotle provided an idea of nature adequate for the founding of physics, and Scholasticism did the same for modern science: Without John Duns Scotus and William of Ockham, Galileo Galilei's work would have been impossible. Physics is again in crisis, facing problems that cannot be solved by physicists, logicians, or epistemologists but only by ontologists, who can supply a fresh idea of nature within which quantum physics can progress.

In his philosophy of existence, Zubiri accepts the "radical ontological nullity" of man, who is nothing apart from the tasks he has to wrestle with. It is in dealing with his tasks that man comes to be. His nature consists in the mission of being sent out into existence to realize himself as a person. These views Zubiri read into Heidegger and Ortega, but he added a doctrine of "religation." (*Religation* was coined by Zubiri from the Latin *religare*, "to tie," which may also be the root of "religion.") According to this doctrine, we are not simply thrown into existence, as atheistic existentialists say, but are impelled into it by something that we feel all the time as an obligation, a force imposing on us the task of choosing and realizing ourselves. That something is deity, to which we are bound, or tied. Religation, the relation to deity, is the "fundamental root of existence" and the "ontological structure of personality."

See also Aristotle; Duns Scotus, John; Existentialism; Galileo Galilei; Heidegger, Martin; Husserl, Edmund; Ortega y Gasset, José; Philosophy of Science, History of; William of Ockham.

Bibliography

Zubiri's works include *Ensayo de una teoría fenomenológica del juicio*, a doctoral thesis on Husserl (Madrid, 1923); *Naturaleza, historia, Dios* (Madrid, 1944); *Sobre la esencia* (Madrid, 1962); and *Cinco lecciones de filosofía* (Madrid: Sociedad de Estudios y Publicaciones, 1963).

For commentary on Zubiri, see Luis Diez del Corral et al., *Homenaje a Xavier Zubiri* (Madrid, 1963), and Julián Marías, "Xavier Zubiri," in *La escuela de Madrid* (Buenos Aires, 1959).

OTHER WORKS BY ZUBIRI

Inteligencia sentiente. Inteligencia y realidad. Madrid: Alianza Editorial/Sociedad de Estudios y Publicaciones, 1980.

Inteligencia y Logos. Madrid Alianza Editorial/Sociedad de Estudios y Publicaciones, 1982.

Siete ensayos de Antropología filosófica. Bogotá: Ed. Universidad de Santo Tomás, 1982.

Inteligencia y Razón. Madrid: Alianza Editorial/Sociedad de Estudios y Publicacione, 1983.

Sobre el hombre. Madrid: Alianza Editorial/Sociedad de Estudios y Publicacione, 1986.

El problema filosófico de la historia de las religions. Madrid: Alianza Editorial/Sociedad de Estudios y Publicacione, 1993.

Los problemas fundamentales de la metafísica occidental. Madrid: Alianza Editorial/Sociedad de Estudios y Publicacione, 1994.

Espacio. Materia. Tiempo. Madrid: Alianza Editorial/Fundación Xavier Zubiri, 1996.

Sobre el problema de la filosofía. Madrid: Alianza Editorial/Fundación Xavier Zubiri, 1996.

El problema teologal del hombre: Cristianismo. Madrid: Alianza Editorial/Fundación Xavier Zubiri, 1997.

El hombre y la verdad. Madrid: Alianza Editorial/Fundación Xavier Zubiri, 1999.

Primeros escritos (1921-1926). Madrid: Alianza Editorial/Fundación Xavier Zubiri, 1999.

TRANSLATIONS

On Essence. Washington, DC: Xavier Zubiri Foundation of North America, 1980.

Nature, History, God. Washington, DC: Xavier Zubiri Foundation of North America, 1981.

Man and God. Washington, DC: Xavier Zubiri Foundation of North America, 1997.

The Philosophical Problem of the History of Religions. Washington, DC: Xavier Zubiri Foundation of North America, 1999.

Sentient Intelligence. Washington, DC: Xavier Zubiri Foundation of North America, 1999.

The Dynamic Structure of Reality. Champagne: University of Illinois Press, 2003.

OTHER

A complete bibliography of Zubiri's writings, the English translations of those writings, and literature on Zubiri is maintained by the Xavier Zubiri Foundation of North America at www.zubiri.org.

Neil McInnes (1967)
Bibliography updated by Thomas Nenon (2005)

File not found (FNF)